IBM GDPS

An Introduction to Concepts and Capabilities

Mairi Jane Lee

Rosazila Musa

Marie-France Narbey

Ayrton Gomes Ribeiro

John Thompson

IBM

Redbooks

International Technical Support Organization

IBM GDPS: An Introduction to Concepts and Capabilities

August 2024

Twentieth Edition (August 2024)

Contents

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®	IBM Z®	VTAM®
CICS®	IBM z Systems®	WebSphere®
Db2®	IBM z16™	z Systems®
DS8000®	InfoSphere®	z/OS®
Easy Tier®	OMEGAMON®	z/VM®
Enterprise Storage Server®	Parallel Sysplex®	z/VSE®
FICON®	RACF®	z13®
FlashCopy®	Redbooks®	z15®
GDPS®	Redbooks (logo) ®	z16™
HyperSwap®	System z®	zEnterprise®
IBM®	Tivoli®	

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Red Hat, OpenShift, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

Other company, product, or service names may be trademarks or service marks of others.

Preface

This IBM Redbooks® publication presents an overview of the IBM Geographically Dispersed Parallel Sysplex® (IBM GDPS®) offerings and the roles they play in delivering a business IT resilience solution.

The book begins with general concepts of business IT resilience and disaster recovery (DR), along with issues that are related to high application availability, data integrity, and performance. These topics are considered within the framework of government regulation, increasing application and infrastructure complexity, and the competitive and rapidly changing modern business environment.

Next, it describes the GDPS family of offerings with specific reference to how they can help you achieve your defined goals for high availability and disaster recovery (HADR). Also covered are the features that simplify and enhance data replication activities, the prerequisites for implementing each offering, and tips for planning for the future and immediate business requirements. Tables provide easy-to-use summaries and comparisons of the offerings. The extra planning and implementation services available from IBM® also are explained.

Then, several practical client scenarios and requirements are described, along with the most suitable GDPS solution for each case.

The introductory chapters of this publication are intended for a broad technical audience, including IT System Architects, Availability Managers, Technical IT Managers, Operations Managers, System Programmers, and Disaster Recovery Planners. The subsequent chapters provide more technical details about the GDPS offerings, and each can be read independently for those readers who are interested in specific topics. Therefore, if you read all of the chapters, be aware that some information is intentionally repeated.

Authors

This book was produced by a team of specialists from around the world working with the IBM Redbooks Poughkeepsie Center.

Mairi Jane Lee is an independent technical writer with over 15 years of experience working as an Information Developer for IBM. Before working as a member of the GDPS Information Development team, she worked as an Information Developer on the IBM CICS® Transaction Server and CICS TX Series group of products.

Rosazila Musa is an MBA-qualified Business Operations Professional with 10 years' experience in IBM Malaysia. During this time, she has managed services for global telecommunication business under IBM Telecom Expense Management (TEM), including end-to-end telecommunication ordering, inventory, and invoicing processes across multiple telecommunication vendors globally. Currently with the GDPS team, she took part in the transformation efforts for business operations and developed an interest in technical content development.

Marie-France Narbey is a certified project manager with over 20 years of experience in different areas of IBM mainframe. Her different positions in EMEA Product Engineering, Poughkeepsie Labs, Customer Satisfaction Project Office, pre-sales, and field enablement activities led her to join the GDPS test team as Information Development team lead.

Ayrton Gomes Ribeiro is a Level 2 Certified IT Specialist and IT Architect, He has 18 years of Mainframe experience as an IBMer and as a customer, mostly working with IBM z/OS® Automation, implementing IBM Z® System Automation, converting other automation products to SA, as well as implementing and supporting GDPS at customers. He has been actively working as a GDPS Specialist and Infrastructure Architect for the last 3 years.

John Thompson is a Senior Technical Staff Member with over 30 years of experience in IBM z/OS software design and development. He is currently involved in IBM Z architecture and strategy, focused on resilience. John is a member of the leadership team for the GDPS Customer Design Council.

Now you can become a published author, too

Here's an opportunity to spotlight your skills, grow your career, and become a published author, all at the same time. Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us.

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

► Use the online **Contact us** review form:

ibm.com/redbooks

► Send your comments by email:

redbook@us.ibm.com

► Mail your comments:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

► Look for us on LinkedIn:

http://www.linkedin.com/groups?home=&gid=2130806

► Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

► Stay current on recent Redbooks publications with RSS Feeds:

http://www.redbooks.ibm.com/rss.html

Summary of changes

This section describes the technical changes that were made in this edition of the book and in previous editions. This edition also includes minor corrections and editorial changes that are not identified.

Summary of Changes
for IBM Redbooks publication SG24-6374-19
for IBM GDPS: An Introduction to Concepts and Capabilities
as created or updated on August 5, 2024.

March 2024, Twentieth Edition

New and changed information

► As part of the Words Matter initiative, which aims to remove vocabulary that contains a negative racial or cultural bias, the term "master" is replaced with the term "primary" wherever possible.

► "System and hardware management" on page 74 was updated.

► New section "Autonomous Path Reconfiguration" on page 104 was added.

► New section "Autonomous Path Reconfiguration" on page 141 was added.

► "GDPS tools for GDPS Metro HyperSwap Manager" on page 143 was updated.

► "Other GM considerations" on page 154 is updated.

► New section "Enhanced resiliency with Region Switch and GM Bidirectional support" on page 165 was added.

► New section "Dynamic PPRC Link Configuration Management" on page 173 was added.

► "GDPS tools for GDPS GM" on page 175 is updated.

► New section "Securing the GDPS environment" on page 204 was added.

► The following enhancements are made to the LCP chapter:

– New management profile functions were added in "LCP terminology" on page 256.

– A significant update of "GDPS Metro Global - GM 4-site" on page 271.

– An update of "Role-based security" on page 287.

– The addition of "Dual control" on page 288, which describes the need for two users to cooperate to perform potentially disruptive actions.

March 2023, Nineteenth Edition

New and changed information

► All references to "XRC," "Metro Global - XRC," and "MzGM" were removed. The following chapters and sections were deleted because the XRC function was withdrawn from service:

  – Section 2.4.2 XRC (z/OS Global Mirror (GM)).

  – Chapter 5, "Introducing GDPS Global - XRC".

  – Section 9.5 GDPS Metro Global - XRC 3-site.

  – Section 9.6 GDPS Metro Global - XRC 4-site solution.

  – Section 10.3.4 GDPS Metro Global - XRC.

► Section 3.5.6, "Reconfiguring the remote copy environment" on page 98 was added to describe the introduction of the IBM Geographically Dispersed Parallel Sysplex (GDPS) Reconfiguration wizard.

► "External LCP in GDPS Metro environments" on page 262 was updated to describe the following new capabilities:

  – External Logical Corruption Protection (LCP) from both volume sets in a single-leg environment.

  – External LCP from the RS3 volume set in a dual-leg environment.

  – Definition of the GDPS Metro volume sets to the GDPS GM LCP Manager system in alternative subchannel sets for scalability.

► "Internal LCP in GDPS Metro Global - GM 4-site environments" on page 271 was added to describe the new supported LCP operational model.

► Section 9.5, "Monitoring" on page 288 was added to describe the monitoring that is performed by the LCP manager.

► "What is a workload" on page 181 was updated to reflect that a workload can now consist of both IBM Db2® and VSAM data.

► The following updates were made to describe the new *Linux in LPAR* function:

  – Section 3.3.2, "Multiplatform Resiliency for IBM Z" on page 70 was updated to include Linux in LPAR as supported.

  – "Multiplatform Resiliency for Linux in LPAR" on page 78 was added to describe the support for Linux in LPAR systems.

► The maximum number of IBM z/VM® systems that can participate in an SSI cluster was updated to eight, as described in "GDPS Metro xDR support for z/VM Single System Image clustering" on page 74.

► All occurrences of Customization Verification Program and CVP were changed to Installation Verification Program (IVP).

► The capability of the IVP to also validate the installation of the GUI setup is described in "Installation Verification Program" on page 75.

► Section 2.8.5, "Flexible Capacity for Cyber Resiliency" on page 37 was added to describe GDPS support for IBM Z Flexible Capacity for Cyber Resilience.

► Section 2.8.6, "GDPS CBU, OOCoD, CPE, SRB, and Flexible Capacity" on page 37 was updated to describe GDPS support for IBM Z Flexible Capacity for Cyber Resilience.

► Section 11.1.1, "Logical view" on page 308 was updated to present the extension of the GDPS Enterprise Portal to GDPS Continuous Availability for the logical view.

► Section 11.1.2, "Physical view" on page 311 was updated to show that the Server Time Protocol (STP) roles are now visible on the physical view of the GDPS Enterprise Portal.

► Several small changes were made for clarity, consistency, or to correct minor errors.

March 2022, Eighteenth Edition

New and changed information
► The section titled "GDPS xDR support for Red Hat OpenShift Container Platform on page 75 was added to describe the GPDS xDR support for Red Hat OpenShift Container Platform.

► The section titled "Remote copy windows" on page 86 was updated to describe the new dynamic PPRC link management capability.

► The section titled "GDPS graphical user interface" on page 154 was updated to indicate that the GUI is now supported in Metro Global Mirror (MGM) environments.

► Section 5.3.5, "Securing the GDPS environment" on page 163 was added to describe the new role-based security capability of GDPS Global - GM.

► The section titled "What is a workload" on page 175 was updated to reflect that a workload can consist of both IMS and IBM Db2.

► Section 6.2.7, "Replication software" on page 182 was updated to describe Global Consistency Groups (GCGs) for IMS, which provides scalability improvements to IIDR for IMS.

► The section titled "Planned Actions" on page 204 was updated to reflect the following changes:

  – VSAM workloads can now be switched with a single ROUTING SWITCH command.

  – During switch operations, data that is associated with all update workload types (Db2, IMS, and VSAM) is fenced in the original active site and unfenced in the new active site.

► Section 9.3.2, "GDPS Global - GM" on page 265 was updated to reflect the new capability to implement External LCP in GDPS GM 2-site environments.

► The section titled "External LCP in GDPS GM environments" on page 266 was added.

► Section 9.4.1, "Scripting" on page 277 was updated to describe the new incremental restore to production capability and to provide additional information about the existing restore functions in LCP environments.

► Section 9.4.3, "Securing the GDPS LCP environment" on page 289 was added to describe the new role-based security capability of the GDPS LCP Manager.

► Chapter 11, "GDPS Enterprise Portal" on page 309 was added to describe the new GDPS Portal.

► Several small changes were made for clarity, consistency, or to correct minor errors.

March 2021, Seventeenth Edition

New and changed information

► The section "Disk and LSS configuration" on page 73 was updated to reflect that disk sharing is now allowed between systems inside and outside of the GDPS sysplex that use hardware reserves.

► 3.4.1, "z/OS Proxy disk and disk subsystem sharing" on page 82The following changes were made to describe the IBM Easy Tier® Heat Map Transfer functions available in GDPS MGM and GDPS MzGM environments, including the enhanced capabilities that were added to GDPS MGM:

  – The information describing the Easy Tier Heat Map Transfer function in GDPS Metro was relocated to new 3.7.5, "Easy Tier Heat Map Transfer" on page 102.

  – Section 8.3.11, "Easy Tier Heat Map Transfer in a GDPS MGM 3-site environment" on page 252 was added.

  – Section 8.4, "GDPS Metro Global - GM 4-site solution" on page 253 was updated.

  – Section 9.5.8, "Easy Tier Heat Map Transfer in a GDPS MzGM 3-site environment" on page 291 was added.

  – Section 9.6, "GDPS Metro Global - XRC 4-site solution" on page 292 was updated.

► Section 3.5.2, "GDPS scripts" on page 90 was updated to describe the new Central Processor Complex (CPC) failure takeover script capability.

► Section 3.6.1, "GDPS Metro health checks" on page 99 and "GDPS Metro HyperSwap Manager health checks" on page 135 were updated to indicate that Status Display Facility (SDF) alerts can now optionally be raised for health check exceptions.

► The following changes were made in support of the new RESTful application programming interface (API) support in GDPS GM:

  – New section 5.3.3, "Application programming interfaces" on page 162 was added to describe the APIs available to manage the GDPS GM environment.

  – The information describing the Query Services API was moved to new 5.3.3, "Application programming interfaces" on page 162.

  – A new section "RESTful APIs" on page 96 was added to describe the new RESTful API feature for GDPS GM.

► The following sections were updated to indicate that the GM IBM FlashCopy® disks are no longer required to be defined to the GDPS GM R-sys:

  – "Addressing z/OS device limits in a GDPS Metro environment" on page 25
  – 5.6, "Flexible testing and Logical Corruption Protection" on page 168
  – 8.3.5, "Scalability in a GDPS MGM 3-site environment" on page 249

► Section 8.4, "GDPS Metro Global - GM 4-site solution" on page 253 was updated to reflect that Managing devices defined in an alternative subchannel set is now supported in GDPS MGM 4-site environments.

► "GDPS graphical user interface" on page 84 is updated to indicate that the GDPS GUI is able to manage GDPS MGM 3-site and 4-site environments.

► Updates were made to the following sections to describe the new support for internal and external LCP in GDPS GM 3-site and 4-site environments:

  – Section 8.4, "GDPS Metro Global - GM 4-site solution" on page 253
  – Section 9.3.3, "GDPS Metro Global - GM" on page 267

- ► Section 9.1, "LCP terminology" on page 258 was updated to describe the new ability to dynamically expand Safeguarded Copy (SGC) backup capacity devices in LCP environments.
- ► The following sections were updated to describe the ability to now have multiple recovery copy sets (CS) in LCP environments:
  - 9.2, "Introduction to LCP and Testcopy Manager" on page 259
  - 9.2.1, "Internal LCP" on page 259
  - 9.2.2, "External LCP" on page 260
- ► Section 8.4, "GDPS Metro Global - GM 4-site solution" on page 253 was updated to reflect that IBM Db2 Analytics Accelerator on IBM Z running in Secure Service Containers (SSCs) is now supported in GDPS MGM 4-site environments.
- ► The following updates were made to describe the integration of GDPS MGM 3-site and GDPS CA to provide a zero data loss (ZDL) capability over regional distances:
  - Section 8.3.13, "GDPS MGM 3-site integration with GDPS Continuous Availability" on page 252 was updated
  - "Zero Data Loss at distance" on page 253 was added
  - Section 8.4, "GDPS Metro Global - GM 4-site solution" on page 253 was updated
- ► The following updates were made to describe GDPS support for Capacity for Planned Events (CPE) and System Recovery Boost (SRB):
  - Section 2.8.3, "Capacity for Planned Events" on page 36 was added to describe CPE support.
  - Section 2.8.4, "System Recovery Boost" on page 36 was added to describe SRB support.
  - Information that is related to CPE and SRB was added to 2.8.6, "GDPS CBU, On/Off CoD, CPE, SRB, and Flexible Capacity" on page 37.
- ► "GDPS scripts" on page 90 was updated to describe the enhancements that were made to planned actions that allow currently running scripts to be stopped and scripts that were previously stopped or that failed to be restarted.
- ► Section 6.2.5, "IBM Multi-site Workload Lifeline for z/OS" on page 181 was updated to describe the availability of user exits provided by the Multi-site Workload Lifeline manager to allow management of workloads that use unsupported connectivity methods.
- ► The following changes were made in support of the new RESTful API support in GDPS XRC:
  - New section 5.3.3, "Application programming interfaces" on page 162 was added to describe the APIs available to manage the GDPS XRC environment.
  - The information describing the Query Services API was moved to new section 5.3.3, "Application programming interfaces" on page 162.
  - A new section "RESTful APIs" on page 162 was added to describe the new RESTful API feature for GDPS XRC.
- ► Several small changes were made for clarity, consistency, or to correct minor errors.

Introducing business resilience and the role of IBM GDPS

In this chapter, we discuss the objective of this book and briefly introduce the contents and layout. We discuss the topic of business IT resilience from a technical perspective (we refer to it as *IT resilience*).

The chapter includes a general description that is not specific to mainframe platforms, although the topics are covered from an enterprise systems and mainframe perspective. Finally, we introduce the members of the IBM Geographically Dispersed Parallel Sysplex (GDPS) family of offerings and provide a brief description of the aspects of an IT resilience solution that each offering addresses.

This chapter includes the following topics:

1.1 Objective

Business IT resilience is a high profile topic across many industries and businesses. Apart from the business drivers requiring near-continuous application availability, government regulations in various industries now take the decision about whether to have an IT resilience capability out of your hands.

This book was developed to provide an introduction to the topic of business resilience from an IT perspective, and to share how GDPS can help you address your IT resilience requirements.

1.2 Layout of this book

This chapter starts by presenting an overview of IT resilience and disaster recovery (DR). These practices have existed for many years. However, recently they became more complex because of a steady increase in the complexity of applications, the increasingly advanced capabilities of available technology, competitive business environments, and government regulations.

In Chapter 2, "Infrastructure planning for availability and IBM GDPS" on page 13, we briefly describe the available technologies that are typically used in a GDPS solution to achieve IT resilience goals. To understand the positioning and capabilities of the various offerings (which encompass hardware, software, and services), it is also useful to have at least a basic understanding of the underlying technology.

Following these two introductory chapters and starting with Chapter 3, "IBM GDPS Metro" on page 45, we describe the capabilities and prerequisites of each offering in the GDPS family of offerings. Because each offering addresses fundamentally different requirements, each member of the GDPS family of offerings is described in a chapter of its own.

Finally, we include a section with examples illustrating how the various GDPS offerings can satisfy your requirements for IT resilience and DR.

1.3 IT resilience

IBM defines *IT resilience* as the ability to rapidly adapt and respond to any internal or external disruption, demand, or threat, and continue business operations without significant impact.

IT resilience is related to, but broader in scope, than *DR*. DR concentrates solely on recovering from an *unplanned* event.

When you investigate IT resilience options, these two terms must be at the forefront of your thinking:

► Recovery time objective (RTO)

This term refers to *how long* your business can afford to wait for IT services to be resumed following a disaster.

If this number is not clearly stated now, think back to the last time that you had a significant service outage. How long was that outage, and how much difficulty did your company suffer as a result? This information can help you get a sense of whether to measure your RTO in days, hours, or minutes.

► Recovery point objective (RPO)

This term refers to *how much data* your company is willing to re-create following a disaster. In other words, what is the acceptable time difference between the data in your production system and the data at the recovery site?

As an example, if your DR solution depends on daily full volume tape dumps, your RPO is 24 - 48 hours depending on when the tapes are taken offsite. If your business requires an RPO of less than 24 hours, you will almost certainly be forced to do some form of offsite real-time data replication instead of relying on these tapes alone.

The terms *RTO* and *RPO* are used repeatedly in this book because they are core concepts in the methodology that you can use to meet your IT resilience needs.

1.3.1 Disaster recovery

As mentioned, the practice of preparing for DR is something that has been a focus of IT planning for many years. In turn, there is a wide range of offerings and approaches available to accomplish DR. Several options rely on offsite or even outsourced locations that are contracted to provide data protection or even servers if there is a true IT disaster. Other options rely on in-house IT infrastructures and technologies that can be managed by your own teams.

There is no one correct answer for which approach is better for every business. However, the first step in deciding what makes the most sense for *you* is to have a good view of your IT resiliency objectives, specifically your RPO and RTO.

Although Table 1-1 does not cover every possible DR offering and approach, it does provide a view of what RPO and RTO might typically be achieved with some common options.

Table 1-1   Typical achievable RPO and RTO for some common DR options

Description	Typically achievable RPO	Typically achievable RTO
No DR plan	Not applicable: all data is lost	Not applicable
Tape vaulting	Measured in days since last stored backup	Days
Electronic vaulting	Hours	Hours (hot remote location) to days
Active replication to remote site (without recovery automation)	Seconds to minutes	Hours to days (dependent on availability of recovery hardware)
Active storage replication to remote "in-house" site	Zero to minutes (dependent on replication technology and automation policy)	One or more hours (dependent on automation)
Active software replication to remote "active" site	Seconds to minutes	Seconds to minutes (dependent on automation)

Generally a form of real-time software or hardware replication is required to achieve an RPO of minutes or less, but the only technologies that can provide an RPO of zero are synchronous replication technologies (see 2.3, "Synchronous versus asynchronous data transfer" on page 20) coupled with automation to ensure that no data is written to one location and not the other.

The recovery time is largely dependent on the availability of hardware to support the recovery and control over that hardware. You might have real-time software or hardware-based replication in place, but without server capacity at the recovery site you have hours to days before you can recover this previously current data.

Furthermore, even with all the spare capacity and current data, you might find that you are relying on people to perform the recovery actions. In this case, you will undoubtedly find that these same people are not necessarily available in a true disaster or even more likely, they find that processes and procedures for the recovery are not practiced or accurate. This is where automation comes in to mitigate the risk that is introduced by the human element and to ensure that you meet the RTO required of the business.

Also, you might decide that one DR option is not appropriate for all aspects of the business. Various applications might tolerate a greater loss of data and might not have an RPO as low as others. At the same time, some applications might not require recovery within hours whereas others most certainly do.

Although there is obvious flexibility in choosing different DR solutions for each application, the added complexity these different DR solutions can bring needs to be balanced carefully against the business benefit. The preferred approach, supported by GDPS, is to provide a single optimized solution for the enterprise. This solution generally leads to a simpler solution and, because less infrastructure and software might need to be duplicated, often a more cost-effective solution, too. Consider a different DR solution only for your most critical applications, where their requirements cannot be catered for with a single solution.

1.3.2 The next level

In addition to the ability to recover from a disaster, many businesses now look for a greater level of availability covering a wider range of events and scenarios. This larger requirement is called *IT resilience*. In this book, we concentrate on two aspects of IT resilience: *DR*, and *continuous availability* (CA), which encompasses recovering from disasters and keeping your applications up and running throughout the far more common planned and unplanned outages that do not constitute an actual disaster.

For some organizations, a proven DR capability that meets their RTO and RPO can be sufficient. Other organizations might need to go a step further and provide near-continuous application availability.

There are several market factors that make IT resilience imperative:

► High and constantly increasing client and market requirements for CA of IT processes

► Financial loss because of lost revenue, punitive penalties or fines, or legal actions that are a direct result of disruption to critical business services and functions

► An increasing number of security-related incidents, causing severe business impacts

► Increasing regulatory requirements

► Major potential business impact in areas such as market reputation and brand image from security or outage incidents

For a business today, few events affect a company as much as having an IT outage, even for a matter of minutes, and then finding a report of the incident splashed across the newspapers and the evening news. Today, your clients, employees, and suppliers expect to be able to do business with you around the clock and from around the globe.

To help keep business operations running 24x7, you need a comprehensive business continuity plan that goes beyond DR. Maintaining high availability (HA) and continuous operations in normal day-to-day operations are also fundamental for success. Businesses need resiliency to help ensure two essentials:

► Key business applications and data are protected and available
► If a disaster occurs, business operations continue with a minimal impact

Regulations

In some countries, government regulations specify how organizations must handle data and business processes. An example is the Health Insurance Portability and Accountability Act (HIPAA) in the United States. This law defines how an entire industry, the US healthcare industry, must handle and account for patient-related data.

Other well-known examples include the US government-released *Interagency Paper on Sound Practices to Strengthen the Resilience of the US Financial System*, which loosely drove changes in the interpretation of IT resilience within the US financial industry, and the Basel II rules for the European banking sector, which stipulate that banks must have a resilient back-office infrastructure.

This area is one that accelerates as financial systems around the world become more interconnected. Although a set of recommendations published in Singapore (such as *S 540-2008 Standard on Business Continuity Management*) might be directly addressing only businesses in a relatively small area, it is common for companies to do business in many countries around the world, where these requirements might be ones for ongoing business operations of any kind.

Business requirements

An important concept to understand is that the cost and complexity of a solution can increase as you get closer to true CA, and that the value of a potential loss must be borne in mind when deciding which solution you *need*, and which one you can *afford*. You do not want to spend more money on a CA solution than the financial loss you can incur as a result of an outage.

A solution must be identified that balances the costs of the solution with the financial impact of an outage. Several studies have been done to identify the cost of an outage; however, most of them are several years old and do not accurately reflect the degree of dependence most modern businesses have on their IT systems.

Therefore, your company must calculate the impact in your specific case. If you have not already conducted such an exercise, you might be surprised at how difficult it is to arrive at an accurate number. For example, if you are a retailer and you suffer an outage in the middle of the night after all the batch work completes, the financial impact is far less than if you had an outage of equal duration in the middle of your busiest shopping day. Nevertheless, to understand the value of the solution, you must go through this exercise, using assumptions that are fair and reasonable.

1.3.3 Other considerations

In addition to the increasingly stringent availability requirements for traditional mainframe applications, there are other considerations, including the ones that are described in this section.

Increasing application complexity

The mixture of disparate platforms, operating systems, and communication protocols that are found within most organizations intensifies the already complex task of preserving and recovering business operations. Reliable processes are required for recovering the mainframe data and also, perhaps, data accessed by multiple types of UNIX, Microsoft Windows, or even a proliferation of virtualized distributed servers.

It is becoming increasingly common to have business transactions that span and update data on multiple platforms and operating systems. If a disaster occurs, your processes must be designed to recover this data in a consistent manner.

Just as you would not consider recovering half an application's IBM Db2 data to 8:00 AM. and the other half to 5:00 PM, the data that is touched by these distributed applications must be managed to ensure that *all* of this data is recovered with consistency to a single point in time. The exponential growth in the amount of data that is generated by today's business processes and IT servers compounds this challenge.

Increasing infrastructure complexity

Have you looked in your computer room recently? If you have, you probably found that your mainframe systems are only a small part of the equipment in that room. How confident are you that all those other platforms can be recovered? And if they can be recovered, will it be to the same point in time as your mainframe systems? And how long will that recovery take?

Figure 1-1 shows a typical IT infrastructure. If you have a disaster and recover the mainframe systems, will you be able to recover your service without all the other components that sit between the user and those systems? It is important to remember why you want your applications to be available, that is, so that users can access them.

Therefore, part of your IT resilience solution must include more than addressing the non-mainframe parts of your infrastructure. It must also ensure that recovery is integrated with the mainframe plan.
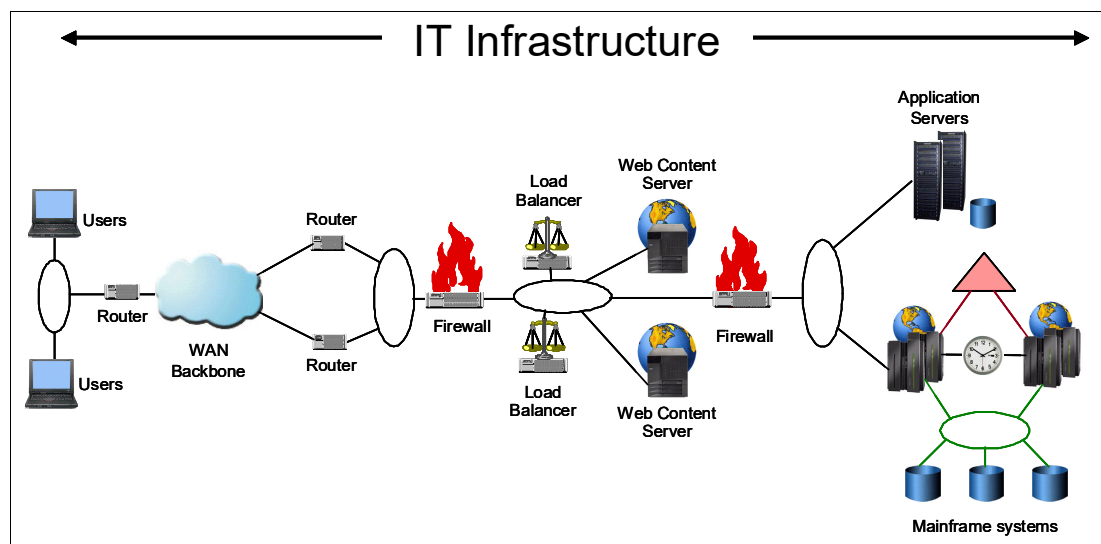
Figure 1-1   Typical IT infrastructure

Outage types

In the early days of computer data processing, planned outages were relatively simple to schedule. Most of the users of your systems were within your company, so the impact to system availability was able to be communicated to all users before the outage. Examples of planned outages are software or hardware upgrades that require the system to be brought down. These outages can take minutes or even hours.

Most outages are planned, and even among unplanned outages, most are not disasters. However, in the current business world of 24x7 internet presence and web-based services that are shared across and also between enterprises, even planned outages can be a serious disruption to your business.

Unplanned outages are unexpected events. Examples of unplanned outages are software or hardware failures. Although various of these outages might be quickly recovered from, others might be considered a disaster.

You will have both planned and unplanned outages while running your organization, and your business resiliency processes must cater to both types. However, you will likely find that coordinated efforts to reduce the numbers of and impacts of unplanned outages often are complementary to doing the same for planned outages.

Later in this book we discuss the technologies available to you to make your organization more resilient to outages, and perhaps avoid them altogether.

1.4 Characteristics of an IT resilience solution

As the previous sections demonstrate, IT resilience encompasses more than the ability to get your applications up and running after a disaster with "some" amount of data loss, and after "some" amount of time.

When investigating an IT resilience solution, keep in mind the following points:

► Support for planned system outages

  Does the proposed solution support stopping a system in an orderly manner? Does it support moving a system from the production site to the backup site in a planned manner? Does it support server clustering, data sharing, and workload balancing, so the planned outage can be masked from users?

► Support for planned site outages

  Does the proposed solution support moving the entire production environment (systems, software subsystems, applications, and data) from the production site to the recovery site? Does it support moving production systems back and forth between production and recovery sites with minimal or no manual intervention?

► Support for data that spans more than one platform

  Does the solution support data from more systems than just z/OS? Does it provide data consistency across all supported platforms, or only within the data from each platform?

► Support for managing the data replication environment

  Does the solution provide an easy-to-use interface for monitoring and managing the data replication environment? Will it automatically react to connectivity or other failures in the overall configuration?

- ► Support for data consistency

  Does the solution provide consistency across all replicated data? Does it support protecting the consistency of the second copy if it is necessary to resynchronize the primary and secondary copy?

- ► Support for continuous application availability

  Does the solution support continuous application availability? From the failure of any component? From the failure of a complete site?

- ► Support for hardware failures

  Does the solution support recovery from a hardware failure? Is the recovery disruptive (restart or IPL again) or transparent (IBM HyperSwap®, for example)?

- ► Support for monitoring the production environment

  Does the solution provide monitoring of the production environment? Is the operator notified in a failure? Can recovery be automated?

- ► Dynamic provisioning of resources

  Can the solution dynamically allocate resources and manage workloads? Will critical workloads continue to meet their service objectives, based on business priorities, if there is a failure?

- ► Support for recovery across database managers

  Does the solution provide recovery with consistency independent of the database manager? Does it provide data consistency across multiple database managers?

- ► End-to-end recovery support

  Does the solution cover all aspects of recovery, from protecting the data through backups or remote copy, through to automatically bringing up the systems following a disaster?

- ► Cloned applications

  Do your critical applications support data sharing and workload balancing, enabling them to run concurrently in more than one site? If so, does the solution support and use this capability?

- ► Support for recovery from regional disasters

  What distances are supported by the solution? What is the impact on response times? Does the distance that is required for protection from regional disasters permit a continuous application availability capability?

You then need to compare your company's requirements in each of these categories against your existing or proposed solution for providing IT resilience.

1.5 GDPS offerings

GDPS is a collection of several offerings, each addressing a different set of IT resiliency goals that can be tailored to meet the RPO and RTO for your business. Each offering uses a combination of server and storage hardware or software-based replication and automation and clustering software technologies, many of which are described in more detail in Chapter 2, "Infrastructure planning for availability and IBM GDPS" on page 13.

In addition to the infrastructure that makes up a GDPS solution, IBM also includes services, particularly for the first installation of GDPS and optionally for subsequent installations to ensure that the solution meets and fulfills your business objectives.

The following list briefly describes each offering, with a view of which IT resiliency objectives it is intended to address. Extra details are included in separate chapters of this book:

► GDPS Metro

A near-CA and DR solution across two sites that are separated by metropolitan distances. The solution is based on the IBM Metro Mirror synchronous disk mirroring technology.

► GDPS Metro HyperSwap Manager

A near-CA solution for a single site or an entry-level DR solution that is across two sites that are separated by metropolitan distances. The solution is based on the same mirroring technology as GDPS Metro, but does not include much of the system automation capability that makes GDPS Metro a complete DR solution.

► IBM GDPS Virtual Appliance

A near-CA and DR solution across two sites that are separated by metropolitan distances. The solution is based on the IBM Metro Mirror synchronous disk mirroring technology. The solution provides Near-CA *and* DR protection for IBM z/VM and Linux on IBM Z in environments that do not have IBM z/OS operating systems.

► GDPS Global - GM (also known as GDPS GM)

A DR solution across two regions that are separated by virtually unlimited distance. The solution is based on the IBM System Storage Global Mirror (GM) technology, which is a disk subsystem-based asynchronous form of remote copy.

► GDPS Metro Global - GM (also known as GDPS MGM)

A 3-site or a symmetrical 4-site configuration is supported:

– GDPS MGM 3-site

A 3-site solution that provides CA across two sites within metropolitan distances in one region and DR to a third site, in a second region, at virtually unlimited distances. It is based on a combination of the Metro Mirror and GM technologies.

– GDPS MGM 4-site

A symmetrical 4-site solution that is similar to the 3-site solution in that it provides CA within region and DR cross region. In addition, in the 4-site solution, the two regions are configured symmetrically so that the same levels of CA and DR protection are provided, no matter which region production runs in.

► GDPS Continuous Availability

A multisite CA/DR solution at virtually unlimited distances. This solution is based on software-based asynchronous mirroring between two active production sysplexes running the same applications with the ability to process workloads in either site.

As mentioned briefly at the beginning of this section, each of these offerings provides the following benefits:

► GDPS automation code

This code has been developed and enhanced over several years to use new hardware and software capabilities to reflect best practices, based on IBM experience with GDPS clients since the inception of GDPS, in 1998, and to address the constantly changing requirements of our clients.

► Can use underlying hardware and software capabilities

IBM software and hardware products have support to surface problems that can affect the availability of those components, and to facilitate repair actions.

► Services

There is only one factor in common across all GDPS implementations: Each has a unique requirement or attribute that makes it different from every other implementation. The services aspect of each offering provides you with invaluable access to experienced GDPS practitioners.

The amount of service that is included depends on the scope of the offering. For example, more function-rich offerings, such as GDPS Metro, include a larger services component than GDPS Metro HyperSwap Manager.

Note: Detailed information about each of the offerings is provided in the following chapters. It is not necessary to read all chapters if you are interested only in a specific offering. If you do read all the chapters, you might notice that some information is repeated in multiple chapters.

1.6 Automation and disk replication compatibility

The GDPS automation code relies on the runtime capabilities of IBM Z NetView and IBM System Automation. Although these products provide tremendous first-level automation capabilities in and of themselves, there are alternative solutions that you might already have from other vendors.

GDPS continues to deliver features and functions that take advantage of properties unique to the IBM Tivoli® products (such as support for alert management through IBM System Automation for Integrated Operations Management), but Z NetView and IBM System Automation also work well alongside other first-level automation solutions. Therefore, although benefits exist to having a comprehensive solution from IBM, you do not have to replace your current automation investments before moving forward with a GDPS solution.

Most of the GDPS solutions rely on the IBM developed disk replication technologies[1] of Metro Mirror, for GDPS Metro, and GM, for GDPS GM. These architectures are implemented on IBM enterprise disk storage products. Also, the external interfaces for all of these disk replication technologies (Metro Mirror, GM, and FlashCopy) were licensed by many major enterprise storage vendors.

This approach gives clients the flexibility to select the disk subsystems that best match their requirements and to mix and match disk subsystems from different storage vendors within the context of a single GDPS solution. Although most GDPS installations do rely on IBM storage products, there are several production installations of GDPS around the world that rely on storage products from other vendors.

IBM has a GDPS Qualification Program for other enterprise storage vendors to validate that their implementation of the advanced copy services architecture meets the GDPS requirements.

The GDPS Qualification Program offers the following arrangement to vendors:

► IBM provides the system environment.
► Vendors install their disks in this environment.
► Testing is conducted jointly.
► A qualification report is produced jointly, describing details of what was tested and the results.

---

[1] Disk replication technology is independent of the GDPS Continuous Availability solution, which uses software replication.

Recognize that this qualification program does not imply that IBM provides defect or troubleshooting support for a qualified vendor's products. However, it does indicate at least a point-in-time (PiT) validation that the products are functionally compatible and demonstrates that they work in a GDPS solution.

Check directly with non-IBM storage vendors if you are considering using their products with a GDPS solution because they can share their own approaches and capabilities to support the specific GDPS offering you are interested in.

1.7 Summary

We discussed why it is important to have an IT resilience solution, and have provided information about key objectives to consider when developing your own solution. We have also introduced the GDPS family of offerings with a brief description of which objectives of IT resiliency each offering is intended to address.

In Chapter 2, "Infrastructure planning for availability and IBM GDPS" on page 13 we introduce key infrastructure technologies related to IT resilience focused on the mainframe platform. After that, we describe how the various GDPS offerings use those technologies. Finally, we position the various GDPS offerings against typical business scenarios and requirements.

We intend to update this book as new GDPS capabilities are delivered.

2

Infrastructure planning for availability and IBM GDPS

In this chapter, we discuss several technologies that are available to help you achieve your goals that are related to IT resilience, recovery time, and recovery point objectives (RPOs). To understand how the IBM Geographically Dispersed Parallel Sysplex (GDPS) offerings that are described in this book can help you, it is important to have at least conceptual knowledge of the functions, capabilities, and limitations of these underlying technologies.

This chapter includes the following topics:

13

2.1 Parallel Sysplex overview

As discussed in Chapter 1, "Introducing business resilience and the role of IBM GDPS" on page 1, *IT resilience* covers more than just recovery from a disaster. It also encompasses ensuring high availability (HA) on a day-to-day basis, protecting your applications from normal planned and unplanned outages. You cannot expect to be able to provide continuous or near-continuous application availability across a disaster if you are unable to provide that in normal operations.

IBM Parallel Sysplex® is the primary mechanism that is used by IBM to provide the highest levels of application availability on the IBM Z[1] platform. The logical first step in a business resiliency project is to do all you can to deliver the highest levels of service from your existing configuration. Implementing Parallel Sysplex with data sharing and dynamic workload routing provides higher levels of availability now. It also provides a foundation to achieve greater resiliency if you implement GDPS.

In the following sections, we briefly discuss Parallel Sysplex, the benefits you can derive by using the technology, and the points to consider if you decide to implement GDPS Metro or GDPS Continuous Availability. Because GDPS GM does not have a continuous availability (CA) aspect, there are no Parallel Sysplex considerations specifically relating to GDPS GM. There are also no Parallel Sysplex considerations for the IBM GDPS Virtual Appliance because the GDPS Virtual Appliance protects only IBM z/VM and Linux on IBM Z platforms.

2.1.1 Maximizing application availability

There is only one way to protect applications from the loss of a single component (such as an IBM CICS region or a z/OS system), and that is to run multiple, failure-isolated copies. This supposition infers an ability to share data at the record level, with integrity, and to *dynamically* route incoming work requests across the available servers. Parallel Sysplex uses hardware and software components to link individual systems together in a cluster. Because all systems in the sysplex are able to share resources and data, they appear as a single image to applications and users, while eliminating single points of failure.

Having more than one instance of an application within the sysplex can shield your users from both planned and unplanned outages. With Parallel Sysplex, parts of the cluster can be brought down for maintenance, upgrades, or any other type of outage, while the applications continue to be available on other members of the sysplex.

GDPS Continuous Availability further extends this concept with the ability to switch the workload between two sysplexes that are separated by virtually unlimited distance for both planned and unplanned outage situations.

Although it is not necessary to have a Parallel Sysplex before implementing most GDPS solutions, it is important to understand the role that Parallel Sysplex plays in supporting the CA aspect of IT resilience. Technical information about implementing and using Parallel Sysplex is available in other IBM Documentation, so it is not covered in this book.

---

[1] In this book, we use the term *IBM Z* to refer to the IBM z Systems®, IBM z Systems®, IBM System z®, and IBM zSeries® ranges of processors. If something applies only to System z or zSeries processors, we point that out at the time.

2.1.2 Multisite sysplex considerations

The considerations for a multisite sysplex depend on whether you plan to run production systems in both sites at the same time or if all the production systems are in a single site at any one time. Configurations where production systems can run in both sites at the same time are referred to as *multisite workload configurations*. Configurations where the production systems run together in one site or the other (but not split across multiple sites) are referred to as *single-site workload configurations* or sometimes as *Active/Standby configurations*. Other variations on this configuration, where production systems are predominantly running at one site but where partially active systems or systems that are enabled only for queries are running at the other site, are still considered multisite workloads.

> **Terminology:** This section is focused on a *multisite sysplex*, which is a single sysplex spread across multiple (typically two) sites, and how the workload is configured to run in those sites to provide near-CA and metro distance disaster recovery (DR).
>
> Do not confuse it with the GDPS Continuous Availability solution that uses some of the same terminology, but is related to multiple sysplexes (limited to two, currently) and how the workload is configured between the two sysplexes, not within any single sysplex.
>
> In a GDPS Continuous Availability environment, it is anticipated that each of the participating sysplexes is in an Active/Active configuration. This configuration provides local and CA with GDPS Metro and GDPS Continuous Availability, which provides a solution for unlimited distance CA/DR. For more information about the GDPS Continuous Availability solution, see Chapter 6, "IBM GDPS Continuous Availability solution" on page 179.

Several phrases are often used to describe variations of multisite workload. Brief definitions are included here for the more commonly implemented variations.

Active/Active   This refers to a multisite workload configuration where z/OS systems are actively running in the same sysplex with active subsystems in more than one site at the same time. Typically this term also implies that applications take advantage of data sharing and dynamic workload routing in such a way that applications can freely move from one site to another. Finally, critical Parallel Sysplex resources are duplexed or replicated in such a way that if one site fails, the remaining site can recover workload within minutes after contending locks and communications timeouts clear. When combined with HyperSwap, an Active/Active configuration has the potential to provide near-CA for applications even in a site outage.

Active/Warm   This refers to a multisite workload configuration that is similar to the Active/Active configuration, with production systems running at more than one site. The difference is that the workload generally runs in one site at a time, with the systems in the other site started without subsystems or other resources active.

This configuration is intended to save IPL time when moving workload between sites. It can be most effective for supporting the planned movement of workload because in many unplanned scenarios, the "warm" systems might also not survive.

Active/Query   This refers to a multisite workload configuration that is quite close to the Active/Active configuration, but where workload at the second site is partitioned or restricted (possibly to queries only) in such a way as to limit impacts because of serialization, thus protecting shared resources when delays because of distance between the sites is a concern. Again, depending on the configuration of the coupling facility (CF) structures (that is, whether they are duplexed across sites or basically in one site at a time), this configuration might provide value only for planned scenarios because in many unplanned scenarios the "query" or "hot standby" subsystems might not survive.

You can devise potentially many more configuration variations, but from a Parallel Sysplex and GDPS[2] perspective, all of them fall into either the single-site or the multisite workload category.

Single-site or multisite workload configuration

When first introduced, Parallel Sysplexes were typically contained within a single site. Extending the distance between the operating system images and the CF has an impact on the response time of requests that use that CF. Also, even if the systems sharing the data are spread across more than one site, all the primary disk subsystems are normally contained in the same site, so a failure affecting the primary disks affects the systems in both sites. As a result, a multisite workload configuration does not, in itself, provide greater availability than a single-site workload configuration during unplanned outages. To achieve the optimal benefit from a multisite workload configuration for planned outages, use HyperSwap to move applications and their data from one site to the other nondisruptively.

More specifically, be careful when planning a multisite workload configuration if the underlying Parallel Sysplex cannot be configured to spread the important CF structures across the sites and still achieve the required performance. As discussed later in this chapter and illustrated in Table 2-1 on page 39, the Coupling Link technology can support links upwards of 100 km with qualified Dense Wavelength Division Multiplexing (DWDM). However, this situation does not mean that your workload tolerates even 1 km of distance between the z/OS images and the CF. Individual coupling operations are delayed by 10 microseconds per kilometer. Although this time can be calculated, there is no safe way to predict the increased queuing effects that are caused by the increased response times and the degree of sharing that is unique to each environment. In other words, you must run your workload with connections at distance to evaluate the tolerance and impact of distance.

The benefits of a multisite workload come with more complexity. This complexity must be accounted for when weighing the benefits of such configurations.

---

[2] Not including the GDPS Continuous Availability solution, which relates to a multiple sysplex configuration that can be either single-site or multisite workloads.

CF structure duplexing

Two mechanisms exist for duplexing CF structures:

► User-Managed Structure Duplexing is supported for use only with Db2 group buffer pool (GBP) structures. Duplexing the GBP structures can significantly reduce the time to recover the structures following a CF or CF connectivity failure. The performance impact of duplexing the GBP structures is small. Therefore, it is best to duplex the GBP structures used by a production Db2 data sharing group.

► System-Managed Coupling Facility Structure Duplexing (referred to as *SM duplexing*) provides a general purpose, hardware-assisted and easy-to-use mechanism for duplexing CF structures. This feature is primarily intended to allow installations to do data sharing without having to have a failure-isolated CF. However, the design of SM duplexing means that having the CFs a significant distance (kilometers) apart can have a dramatic impact on CF response times for the duplexed structures, and thus your applications, and needs careful planning and testing.

In addition to the response time question, there is another consideration relating to the use of cross-site SM Duplexing. Because communication between the CFs is independent of the communication between mirrored disk subsystems, a failure that results in remote copy being suspended would not necessarily result in duplexing being suspended at the same instant. In a potential disaster, you want the data in the "remote" CF to be frozen in time at the same instant the "remote" disks are frozen, so you can restart your applications from the moment of failure.

If you are using duplexed structures, it might seem that you are ensured to be able to use the duplexed instance of your structures if you must recover and restart your workload with the frozen secondary copy of your disks. However, this situation is not always the case. There can be rolling disaster scenarios where before, after, or during the freeze event, an interruption occurs (perhaps failure of CF duplexing links) that forces CFRM to drop out of duplexing. There is no guarantee that the structure instance in the surviving site is the one that is kept. It is possible that CFRM keeps the instance in the site that is about to totally fail. In this case, there will not be an instance of the structure in the site that survives the failure.

Furthermore, during a rolling disaster event, if you freeze secondary disks at a certain point but continue to update the primary disks and the CF structures, then the CF structures, whether duplexed or not, will not be usable if it is necessary to recover on the frozen secondary disks. This situation depends on some of your installation's policies.

To summarize, if there is a surviving, accessible instance of application-related structures, it might or might be consistent with the frozen secondary disks and therefore might or might not be usable. Furthermore, depending on the circumstances of the failure, even with structures duplexed across two sites, you are not 100% guaranteed to have a surviving, accessible instance of the application structures. Therefore, you must have procedures in place to restart your workload without the structure contents.

For more information, see the white paper titled System-Managed CF Structure Duplexing, GM13-0103.

2.2 Data consistency

In an unplanned outage or disaster situation the ability to perform a database restart, rather than a database recovery, is essential to meet the recovery time objective (RTO) of many businesses, which typically are less than an hour. Database restart allows starting a database application (as you would follow a database manager abend or system abend) without having to restore it from backups. Database recovery is normally a process that is measured in many hours (especially if you have hundreds or thousands of databases to recover), and it involves restoring the last set of image copies and applying log changes to bring the databases up to the point of failure.

But, there is more to consider than the data for one data manager. What if you have an application that updates data in IMS, Db2, and VSAM? If you need to do a recover for these products, will your recovery tools recover them to the same point in time and to the level of granularity that ensures that either all or none of the updates that are made by one transaction are recovered? Being able to do a restart rather than a recover avoids these issues.

Data consistency across all copies of replicated data, spread across any number of storage subsystems, and sometimes across multiple sites, is essential to providing data integrity and the ability to perform a normal database restart if there is a disaster.

2.2.1 Dependent write logic

Database applications commonly ensure the consistency of their data by using *dependent write logic* regardless of whether data replication techniques are being used. Dependent write logic states that if I/O B must logically follow I/O A, so B does not start until A completes successfully. This logic would normally be included in all software to manage data consistency. There are numerous instances within the software subsystem, such as databases, catalog/VTOC, and VSAM file updates, where dependent writes are issued.

As an example, in Figure 2-1 on page 19, LOG-P is the disk subsystem containing the database management system (DBMS) logs, and DB-P is the disk subsystem containing the DBMS data segments. When the DBMS updates a database, it also performs the following process:

1. Write an entry to the log about the intent of the update.
2. Update the database.
3. Write another entry to the log indicating that the database was updated.

If you are doing a remote copy of these volumes, be sure that *all* the updates are mirrored to the secondary disks.
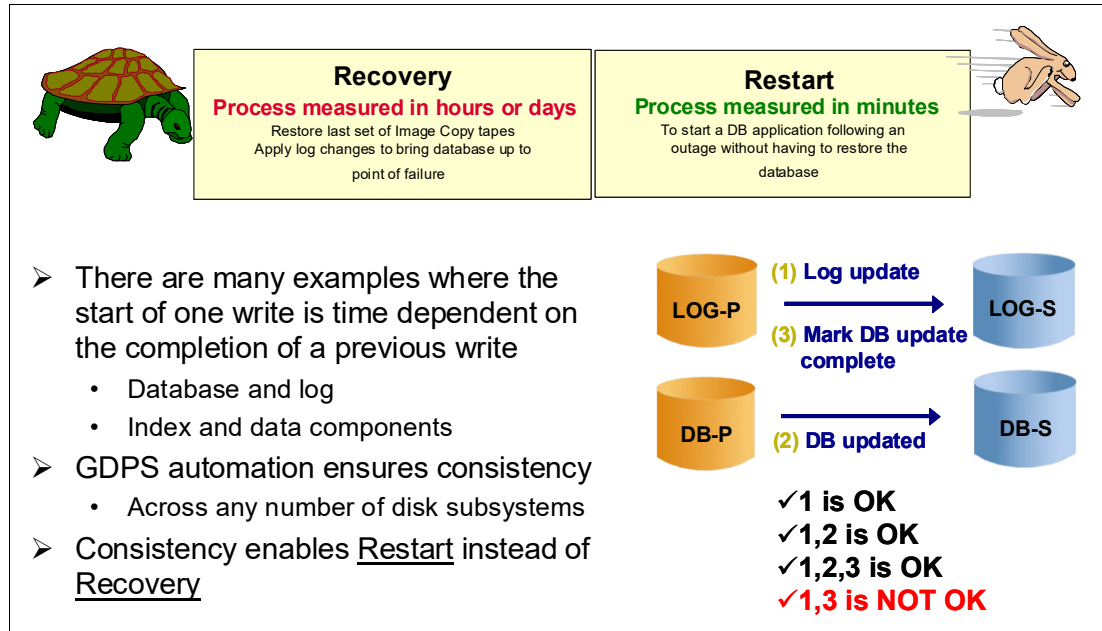
Figure 2-1   Need for data consistency

It is unlikely that all the components in a data center will fail at the same instant, even in the rare case of a full data center outage. The networks might fail first, or possibly one disk subsystem, or any other component in unpredictable combinations. No matter what happens, the remote image of the data must be managed so that cross-volume and subsystem data consistency is preserved during intermittent and staged failures that might occur over many seconds, even minutes. Such a staged failure is referred to as a *rolling disaster*.

Data consistency during a rolling disaster is difficult to achieve for synchronous forms of remote copy because synchronous remote copy is entirely implemented within disk subsystem pairs.

For example, in Figure 2-1 the synchronously mirrored data sets are spread across multiple disk subsystems for optimal performance. The volume containing the DBMS log on the LOG-P disk subsystem in Site1 is mirrored to the secondary volume in the LOG-S disk subsystem in Site2, and the volume containing the data segments in the DB-P disk subsystem in Site1 is mirrored to the secondary volume in the DB-S disk subsystem in Site2.

Assume that a disaster is in progress in Site1, causing the link between DB-P and DB-S to be lost before the link between LOG-P and LOG-S is lost. With the link between DB-P and DB-S lost, a write sequence of (1), (2), and (3) might be completed on the primary devices (depending on how the remote copy pair was defined) and the LOG writes (1) and (3) would be mirrored to the LOG-S device, but the DB write (2) would not have been mirrored to DB-S. A subsequent DBMS restart using the secondary copy of data in Site2 would clean up in-flight transactions and resolve in-doubt transactions, but the missing DB write (2) would not be detected. In this example of the missing DB, write the DBMS integrity was compromised.[3]

---

[3] The way the disk subsystem reacts to a synchronous IBM Metro Mirror remote copy failure depends on the options you specify when setting up the remote copy session. The behavior that is described here is the default if no overrides are specified.

For more information about data consistency for synchronous remote copy, see "Metro Mirror data consistency" on page 24 and for Global Mirror (GM) in 2.4.2, "Global Mirror" on page 26.

For GDPS Continuous Availability, which relies on asynchronous software replication as opposed to the use of Metro Mirror or GM, consistency is managed within the replication software products. For more information, see 2.4.4, "IBM software replication products" on page 29.

2.3 Synchronous versus asynchronous data transfer

Synchronous data transfer and *asynchronous* data transfer are two methods that are used to replicate data. Before selecting a data replication technology, you must understand the differences between the methods that are used and the business impact.

When using synchronous data transfer, as shown in Figure 2-2 by using IBM Metro Mirror, the application writes are first written to the primary disk subsystem (1) and then forwarded on to the secondary disk subsystem (2). When the data is committed to *both* the primary and secondary disks (3), an acknowledgment that the write is complete (4) is sent to the application. Because the application must wait until it receives the acknowledgment before running its next task, there is a slight performance impact. Furthermore, as the distance between the primary and secondary disk subsystems increases, the write I/O response time increases because of signal latency[4].

The goals of synchronous replication are zero or near-zero loss of data, and quick recovery times from failures that occur at the primary site. Synchronous replication can be costly because it requires high-bandwidth connectivity.

One other characteristic of synchronous replication is that it is an enabler for nondisruptive switching between the two copies of the data that are known to be identical.
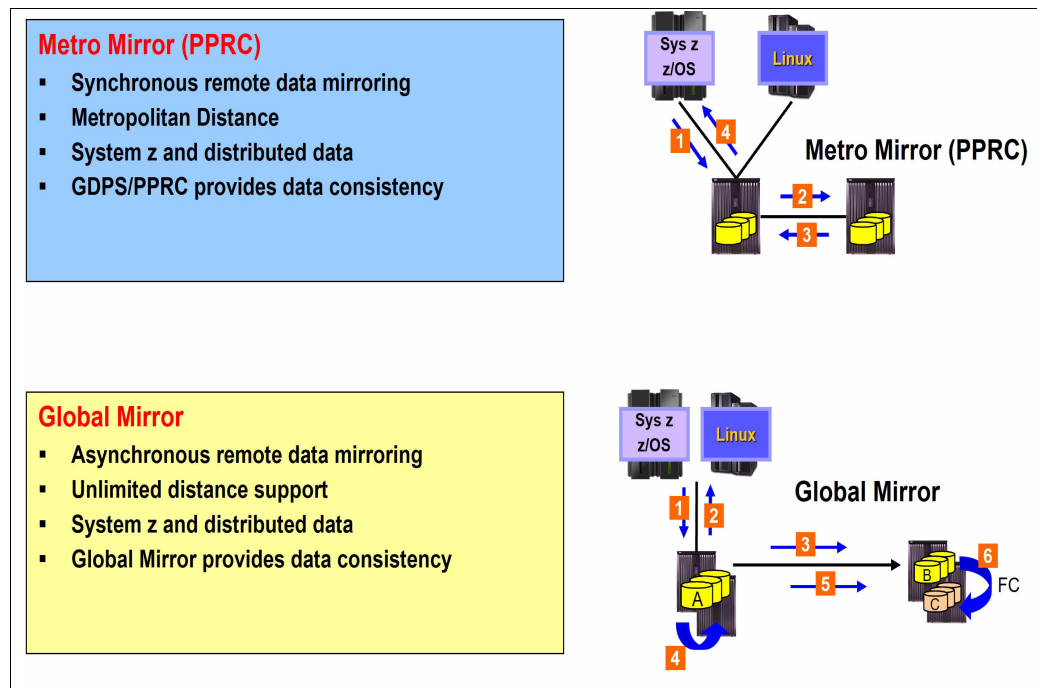
Figure 2-2   Synchronous versus asynchronous storage replication

---

[4] Signal latency is related to the speed of light over fiber and is 10 microseconds per km, round trip.

With asynchronous replication (GM) (Figure 2-2 on page 20), the application writes to the primary disk subsystem (1) and receives an acknowledgment that the I/O is complete when the write is committed on the primary disk (2). The write to the secondary disk subsystem is completed in the background. Because applications do not have to wait for the completion of the I/O to the secondary device, asynchronous solutions can be used at virtually unlimited distances with negligible impact to application performance. In addition, asynchronous solutions do not require as much bandwidth as the synchronous solutions.

With software-based asynchronous replication, as used in a GDPS Continuous Availability environment, data is captured from the database subsystem logs at the source copy when a transaction commits data to the database. That captured data is then sent asynchronously to a second location where it is applied to the target copy of the database in near real time.

When selecting a data replication solution, perform a business impact analysis to determine which solution meets the businesses requirements while ensuring that your service delivery objectives continue to be met; see Figure 2-3. The maximum amount of transaction loss that is acceptable to the business (RPO) is one measurement that is used to determine which remote copy technology should be deployed. If the business is able to tolerate the loss of committed transactions, then an asynchronous solution likely provides the most cost-effective solution. When no loss of committed transactions is the objective, then synchronous remote copy must be deployed. In this case, the distance between the primary and secondary remote copy disk subsystems, and the application's ability to tolerate the increased response times, must be factored into the decision process.
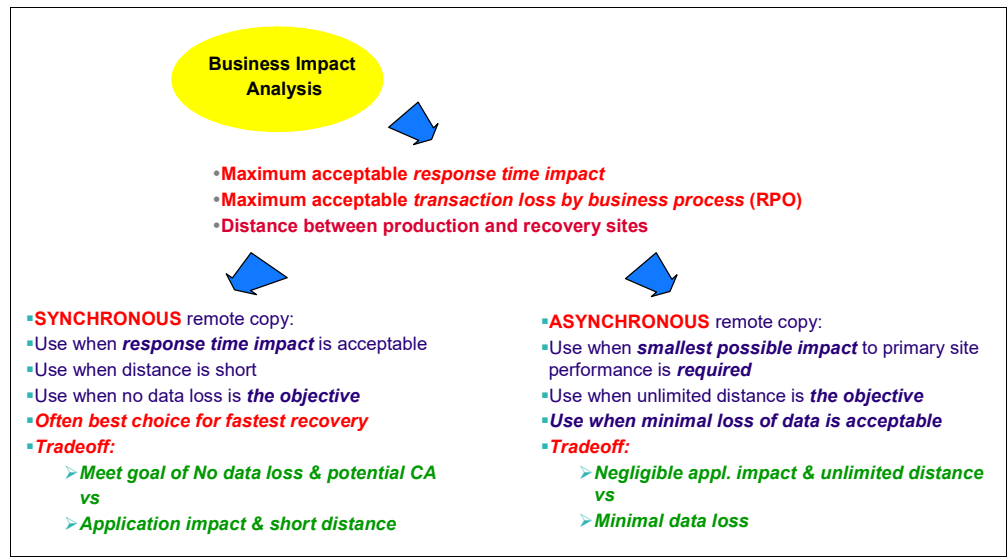
Figure 2-3   Business impact analysis

Many enterprises have both business and regulatory requirements to provide near-continuous data availability, without loss of transactional data, while protecting critical business data if there is a wide-scale disruption. This situation can be achieved by implementing three-copy (sometimes referred to as *3-site*) mirroring solutions that use both synchronous and asynchronous replication technologies. Synchronous solutions are used to protect against the day-to-day disruptions with no loss of transactional data. Asynchronous replication is used to provide out-of-region data protection, with some loss of committed data, for wide-spread disruptions. The key is to ensure cross-disk subsystem data integrity and data consistency are maintained through any type of disruption.

For more information about three-copy replication solutions, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237.

2.4 Data replication technologies

The two primary ways to make your data available following a disaster are as follows:

- ► By using a form of tape-based backup
- ► By using data replication to a recovery site (also known as remote copy)

    This replication can be hardware-based or software-based.

For companies with an RTO of a few hours or less, a tape-based solution is unlikely to be acceptable because it is not possible to restore all your volumes and apply all database logs in the time available. Therefore, we are assuming that if you are reading this book that you already have or are planning to implement some form of data replication technology.

Remotely copying your data eliminates the time that would be required to restore the data from tape and addresses the problem of having to recover data that is generated between the last backup of an application system and the time when the application system fails. Depending on the technology used, remote copy implementations provide a real-time (or near real-time) continuing copy of data between a source and a target.

IBM offers two technologies to provide this type of mirroring for disk storage:

- ► Metro Mirror: Updates to the primary volumes are synchronously mirrored to the remote volumes and all interactions that are related to this activity are done between the disk subsystems. Multi-Target Metro Mirror (MTMM) is based on Metro Mirror and allows multiple secondary copies from the same primary.
- ► GM: This offering mirrors the data asynchronously, and like Metro Mirror, all interactions are done between the disk subsystems; no host interaction is involved.

These technologies are described more fully in the following sections.

IBM also offers several software-based replication products. Unlike the technologies listed for mirroring disk storage (which are application independent), most software replication products are specific to the database source and target in use. The following products are supported in a GDPS Continuous Availability environment:

- ► IBM InfoSphere® Data Replication for IMS for z/OS
- ► IBM InfoSphere Data Replication for VSAM for z/OS
- ► IBM InfoSphere Data Replication for Db2 for z/OS

These products are introduced in the following sections. For more information, see IBM Documentation.

2.4.1 Metro Mirror

Metro Mirror ensures that after the volume pair is established and remains synchronized that the secondary volume always contains the same data as the primary. The IBM implementation of Metro Mirror provides synchronous data mirroring at distances up to 300 km (and potentially even greater distances after technical review and approval).

Important: Always use caution when considering long distances. When we say that something is "supported up to xx km," it means that the technology works at that distance if you have qualified cross-site connectivity technology that supports that protocol. See 2.9, "Cross-site connectivity considerations" on page 37 for more details.

Also consider the impact the increased response time has on your applications. Some applications can tolerate the response time increase that is associated with cross-site distances of 100 km, but the same distance in another installation might make it impossible for the applications to deliver acceptable levels of performance.

So, carefully evaluate the projected response time impact, and apply that increase to your environment to see whether the result is acceptable. Your vendor storage specialist can help you determine the disk response time impact of the proposed configuration.

Recovery point objective with Metro Mirror

If you have an RPO of zero, meaning zero data loss (ZDL), Metro Mirror is the only IBM remote copy option that can achieve that objective.

You might not always have ZDL if you use Metro Mirror. ZDL means that there will never be updates that are made to the primary disks that are not mirrored to the secondaries. The only way to ensure that ZDL is to immediately stop all update activity to the primary disks if the remote copy relationship ceases to exist (if you lose connectivity between the primary and secondary devices), for example).

Thus, choosing to have ZDL really means that you must have automation in place that stops all update activity in the appropriate circumstances. It also means that you accept the possibility that the systems can be stopped for a reason other than a real disaster; for example, if the failure was caused by a broken remote copy link rather than a fire in the computer room. However, completely avoiding single points of failure in your remote copy configuration can reduce the likelihood of such events to an acceptably low level.

Supported platforms with Metro Mirror

Metro Mirror replication is supported for any IBM or non-IBM disk subsystem that supports the Metro Mirror architecture, specifically the Freeze/Run capability. Metro Mirror can mirror Fixed-Block (FB) devices that are used by IBM Z and platforms other than IBM Z and Count-Key-Data (CKD) devices that are used by mainframe operating systems, such as IBM z/OS, IBM z/VM, and IBM z/VSE®.

Not all operating systems necessarily support an interface to control the remote copy function. However, the Metro Mirror function for FB devices can be controlled from a connected z/OS system if the disk storage subsystem supports the IBM z/OS Fixed-Block Architecture (zFBA) feature (as described in "FB disk management prerequisites" on page 70 for GDPS Metro, and in 4.3.1, "FB disk management prerequisites" on page 129 for GDPS Metro HyperSwap Manager).

With current implementations of Metro Mirror, the primary and secondary disk subsystems must be from the same vendor, although vendors (including IBM) often support Metro Mirror between different disk subsystem models of their own product lines. This situation can help with migrations and technology upgrades.

Distance with Metro Mirror

The maximum distance that is supported for IBM Metro Mirror is 300 km (without an RPQ). Typical GDPS Metro and GDPS Metro HyperSwap Manager configurations are limited to distances less than 300 km because of Coupling Link configurations. For more information about the supported distances for these Parallel Sysplex connections, see 2.9.3, "Coupling links" on page 39. You also need to contact other storage vendors to understand the maximum distances supported by their Metro Mirror compatible mirroring implementations.

Performance with Metro Mirror

As the distance between your primary and secondary disk subsystems increases, the time it takes for your data to travel between the subsystems also increases. This situation might have a performance impact on your applications because they cannot proceed until the write to the secondary device completes.

As response times increase, link use also increases. Depending on the type and number of Metro Mirror links you configured, more links and the use of Parallel Access Volumes (PAVs) might help to provide improved response times at longer distances.

Disk Magic, a tool available to your IBM storage specialist, can be used to predict the impact of various distances, link types, and link numbers for IBM disk implementation. We consider access to the information provided by such a tool essential to a GDPS project that uses Metro Mirror.

Metro Mirror connectivity

Connectivity between the primary and secondary disk subsystems can be provided by direct connections between the primary and secondary disk subsystems, by IBM FIbre Connection (IBM FICON®) switches, by DWDMs, and by channel extenders.

The type of intersite connection (dark fiber or telecommunications link) available determines the type of connectivity that you use: telecommunication links can be used by channel extenders, and the other types of connectivity require dark fiber.

For more information about connectivity options and considerations for IBM Z, see the most recent version of *IBM System z Connectivity Handbook,* SG24-5444.

Metro Mirror data consistency

When using Metro Mirror, the following sequence of actions occurs when an update I/O is issued to a primary volume:

1. Write to the primary volume (disk subsystem cache and non-volatile storage (NVS)).

   Your production system writes data to a primary volume and a cache hit occurs.

2. Write to the secondary (disk subsystems cache and NVS).

   The primary disk subsystem's microcode then sends the update to the secondary disk subsystem's cache and NVS.

3. Signal write is complete on the secondary.

   The secondary disk subsystem signals write complete to the primary disk subsystem when the updated data is in its cache and NVS.

4. Post I/O is complete.

   When the primary disk subsystem receives the write complete from the secondary disk subsystem, it returns `Device End (DE)` status to your application program. Now, the application program can continue its processing and move on to any dependent writes that might be waiting for this one to complete.

However, Metro Mirror on its own provides this consistency only for a single write. Ensuring consistency across multiple logical subsystems (LSSs) and even across multiple disk subsystems requires automation on top of the Metro Mirror function. This situation is where GDPS comes in with freeze automation, which is described in the following sections:

► 3.1.1, "Protecting data integrity and data availability with GDPS Metro" on page 46 for GDPS Metro
► 4.1.1, "Protecting data integrity and data availability with GDPS HM" on page 112 for GDPS Metro HyperSwap Manager

Metro Mirror transparent disk swap

Because under normal conditions the primary and secondary disks are known to be identical, with Metro Mirror it is possible to swap to using the secondary copy of the disks in a manner that is transparent to applications that are using those disks. This task is not simple. It requires tight control and coordination across many devices that are shared by multiple systems in a timely manner. GDPS Metro and GDPS Metro HyperSwap Manager automation, with support that is provided in z/OS, z/VM, and specific distributions of Linux on IBM Z, provide such a transparent swap capability that is known as *HyperSwap*.

HyperSwap is a key availability-enabling technology. For more information about GDPS HyperSwap, see the following sections:

► "GDPS HyperSwap function" on page 52 for GDPS Metro
► "GDPS HyperSwap function" on page 117 for GDPS Metro HyperSwap Manager
► "GDPS HyperSwap function" on page 220 for the GDPS Virtual Appliance.

Addressing z/OS device limits in a GDPS Metro environment

As clients implement IT resiliency solutions that rely on multiple copies of data, more are finding that the z/OS limit of 64 K (65,536) devices is limiting their ability to grow or even to take advantage of technologies like HyperSwap. Clients can consolidate data sets to fewer, large volumes, but even with them, sometimes they might not make operational sense for all types of data.

As a result, z/OS introduced the concept of an "alternative subchannel set," which can include the definition for certain types of disk devices. An alternative subchannel set provides another set of 64 K devices for the following device types:

► PAV alias devices
► Metro Mirror secondary devices (defined as 3390D)
► FlashCopy target devices

Including PAV alias devices in an alternative subchannel set is transparent to GDPS and is common practice for current GDPS Metro HyperSwap Manager and GDPS Metro environments.

Support is included in GDPS Metro HyperSwap Manager and GDPS Metro to allow definition of Metro Mirror secondary devices in an alternative subchannel set. With this feature, GDPS can support Metro Mirror configurations with nearly 64 K device pairs. GDPS Metro HyperSwap Manager allows the secondary devices for z/OS systems in the GDPS sysplex, and for managed z/VM systems (and guests) to be defined in an alternative subchannel set. GDPS Metro only supports alternative subchannel sets for z/OS systems in the sysplex.

There are limitations to keep in mind when considering the use of this feature. Specifically, enhanced support is provided in IBM zEnterprise® 196 or 114 servers that allow the Metro Mirror secondary copy of the IPL, IODF, and stand-alone dump devices for z/OS systems in the GDPS sysplex to also be defined in the alternative subsystem set (MSS1).

With this support, a client can define all z/OS Metro Mirrored devices belonging to the GDPS sysplex uniformly with their secondary in the alternative subchannel set. This situation removes the necessity to define IPL, IODF, and stand-alone dump devices differently in MSS0.

The use of alternative subchannel sets for the FlashCopy target devices that are managed by GDPS is not necessary because no requirement exists to define unit control blocks (UCBs) for these devices (they can be in any subchannel set or not defined at all). This issue contributes to the ability of GDPS to support Metro Mirror configurations with nearly 64 K device pairs because no device numbers or UCBs are used by the FlashCopy target devices.

Multi-Target Metro Mirror

> **Note:** There is no requirement to define UCBs for the FlashCopy target devices that are managed by GDPS.

Multi-target PPRC, also known as MT-PPRC, is based on the PPRC (Metro Mirror) technology. The MT-PPRC architecture allows multiple secondary, synchronous, or asynchronous secondary devices for a single primary device.

MTMM is a specific topology that is based on the MT-PPRC technology, which allows maintaining two synchronous Metro Mirror secondary targets (two Metro Mirror legs) from a single primary device. Each leg is tracked and managed independently. Consider the following points:

► Data is transferred to both targets in parallel.
► Pairs operate independent of each other.
► Pairs may be established, suspended, or removed separately.
► A replication problem on one leg does not affect the other leg.
► HyperSwap is possible on either leg.

MTMM provides all the benefits of Metro Mirror plus has the extra protection of a second synchronous leg.

Summary

Metro Mirror synchronous replication gives you the ability to remote copy your data in real time, with the potential for no data loss at the recovery site. Metro Mirror is your only choice if your RPO is zero. Metro Mirror is the underlying remote copy capability that the GDPS Metro, GDPS Metro HyperSwap Manager, and GDPS Virtual Appliance offerings are built on.

2.4.2 Global Mirror

GM is an asynchronous remote copy technology that enables a 2-site DR and backup solution for the IBM Z and distributed systems environments. Using asynchronous technology, GM operates over Fibre Channel Protocol (FCP) communication links and maintains a consistent and restartable copy of data at a remote site that can be at virtually unlimited distances from the local site.

GM works by using three sets of disks, as shown in Figure 2-4 on page 27. Global Copy (PPRC Extended Distance, or PPRC-XD), which is an asynchronous form of PPRC (Metro Mirror), is used to continually transmit data from the primary (A) to secondary (B) volumes, by using the out-of-sync bitmap to determine what needs to be transmitted. Global Copy does not guarantee that the arriving writes at the local site are applied to the remote site in the same sequence. Therefore, Global Copy by itself does not provide data consistency.

If there are multiple physical primary disk subsystems, one of them is designated as the *Master* and is responsible for coordinating the creation of consistency groups. The other disk subsystems are subordinates to this Master.

Each primary device maintains two bitmaps. One bitmap tracks incoming changes. The other bitmap tracks which data tracks must be sent to the secondary before a consistency group can be formed in the secondary.

Periodically, depending on how frequently you want to create consistency groups, the Master disk subsystem signals the subordinates to pause application writes and swap the change recording bitmaps. This situation identifies the bitmap for the next consistency group. While the I/Os are paused in all LSSs in the GM session, any dependent writes are not issued because the CE/DE has not been returned. This situation maintains consistency across disk subsystems. The design point to form consistency groups is 2 - 3 ms.

After the change recording bitmaps are swapped, write I/Os are resumed and the updates that remain on the GM primary for the current consistency group will be drained to the secondaries. After all the primary devices are drained, a FlashCopy command is sent to the GM secondaries (B), which are also the FlashCopy source volumes, to perform a FlashCopy to the associated FlashCopy target volumes (C). The tertiary or C copy is a consistent copy of the data.

The B volumes are secondaries to Global Copy and are not guaranteed to be consistent. The C copy provides a "gold copy" which can be used to make the B volumes consistent in case recovery is required. Immediately after the FlashCopy process is logically complete, the primary disk subsystems are notified to continue with the Global Copy process. For more information about FlashCopy, see 2.6, "FlashCopy" on page 32.

After Global Copy is resumed, the secondary or B volumes are inconsistent. However, if recovery is needed, the FlashCopy target volumes provide the consistent data for recovery.

All this processing is done under the control of microcode in the disk subsystems. You can have up to 16 mirrored pairs in a pool, one of which is the Master primary and secondary pair (see Figure 2-4).
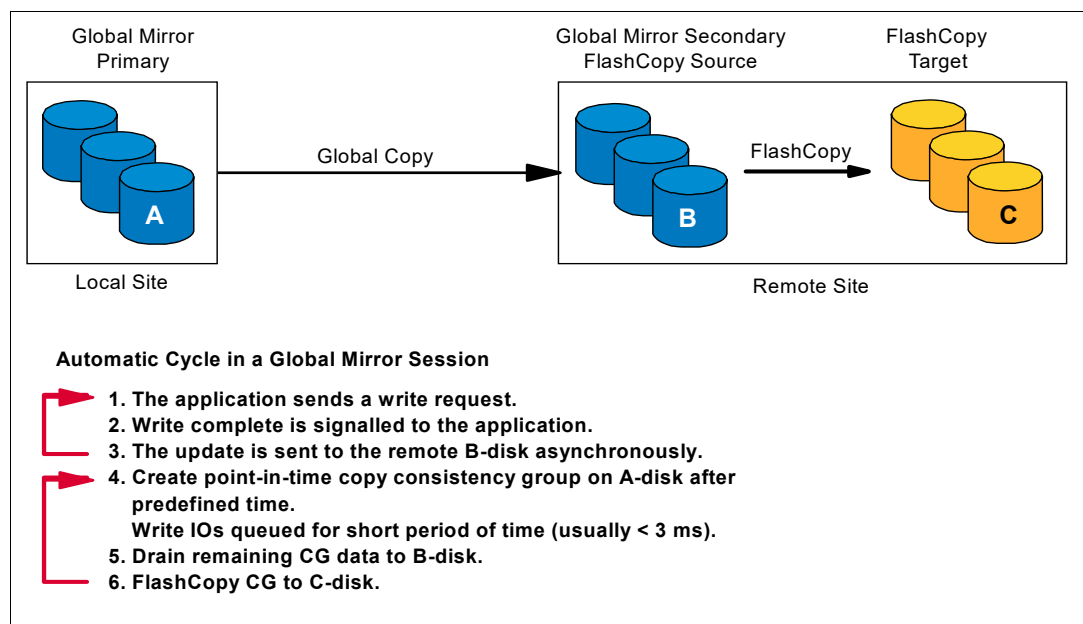
Figure 2-4   Global Mirror: How it works

Recovery point objective

Because GM is an asynchronous remote copy solution, there always is an amount of data that must be re-created following a disaster. As a result, GM can be used only when your RPO requirement is greater than zero.The amount of time that the FlashCopy target volumes lag behind the primary depends mainly on the following items:

► How often consistency groups are built

  This situation is controlled by the installation and can be specified in terms of seconds.

► The amount of bandwidth

  If there is insufficient bandwidth to transmit the updates in a timely manner, contention on the remote copy links can cause the secondary volumes to drift further behind at peak times. The more frequently you create consistency groups, the more bandwidth you require.

Although it is not unusual to have an average RPO of 2 - 3 seconds with GM, it is possible that the RPO increases if production write rates exceed the available resources. However, the mirroring session is not suspended and the production workload is not impacted if the capacity of the replication environment is exceeded because of unexpected peaks in the workload or an underconfigured environment.

To maintain a consistent lag between the primary and secondary disk subsystems, you must have sufficient connectivity. For more information about planning for the performance aspects of your GM configuration, see *IBM DS8870 Copy Services for IBM z Systems*, SG24-6787.

Supported platforms

The IBM Enterprise Storage Server® and IBM DS8000® families of disk subsystems support GM. For other enterprise disk vendors, contact your vendor to determine whether they support GM and if so, on which models.

Distance and connectivity

Because GM is an asynchronous remote copy capability, the amount of time it takes to mirror the update to the remote disks does not affect the response times to the primary volumes. As a result, virtually unlimited distances between the primary and secondary disk subsystems are supported.

GM requires FCP links on the disk subsystems. If the recovery site is within the distance that is supported by FCP direct connect, switches, or DWDM, you can use one of those methods to connect the primary and secondary disk subsystems. Otherwise, you must use network extension technology that supports FCP links.

Addressing z/OS device limits in a GDPS GM environment

As clients implement IT resiliency solutions that rely on multiple copies of data, more are finding that the z/OS limit of 64 K (65,536) devices is limiting their ability to grow. Clients can consolidate data sets to fewer, large volumes, but even with these volumes, sometimes they might not make operational sense for all types of data.

To this end, z/OS introduced the concept of an *alternative subchannel set*, which can include the definition for certain types of disk devices. An alternative subchannel set provides another set of 64 K devices for the following device types:

► PAV alias devices
► Metro Mirror secondary devices (defined as 3390D)
► FlashCopy target devices

Including PAV alias devices in an alternative subchannel set is transparent to GDPS and is common practice for many client configurations.

The application site controlling system performs actions against the GM primary devices and can address up to nearly 64 K devices. The recovery site controlling system performs actions against the GM secondary and the GM FlashCopy devices. GDPS supports defining the GM FlashCopy devices in an alternative subchannel set (MSS1) or not defining them at all (which is known as *no-UCB FlashCopy*). This ability allows up to nearly 64 K devices to be replicated in a GDPS GM environment.

Summary

GM provides an asynchronous remote copy offering that can support an RPO of two to three seconds at virtually unlimited distance. GM is the replication technology that GDPS Global - GM is built on.

2.4.3 Combining disk remote copy technologies for CA and DR

In this section, we briefly describe Metro Global Mirror (MGM). For more information, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237. Combining the technologies of Metro Mirror and HyperSwap with GM enables clients to meet requirements for CA with ZDL locally within metropolitan distances for most failures, along with providing a DR solution for a region-wide disaster. This combination often enables clients to meet increasing regulatory requirements.

Metro Global Mirror

MGM is a data replication solution that combines the capabilities of Metro Mirror and GM.

Synchronous replication between a primary and secondary disk subsystem either within a single data center, or between two data centers within metropolitan distances, is implemented by using Metro Mirror. GM is used to asynchronously replicate data from the secondary disks to a third disk subsystem in a recovery site typically out of the local metropolitan region. As described in 2.4.2, "Global Mirror" on page 26, a fourth set of disks, also in the recovery site, are the FlashCopy targets used to provide the consistent data for DR.

MGM provides a comprehensive three-copy (or four-copy) data replication strategy to protect against day-to-day disruptions, while protecting critical business data and functions if there is a wide-scale disruption.

2.4.4 IBM software replication products

This section does not aim to provide a comprehensive list of all IBM software-based replication products. Instead, it provides an introduction to the following supported products within the GDPS Continuous Availability solution:

► InfoSphere Data Replication for IMS for z/OS
► InfoSphere Data Replication for VSAM for z/OS
► InfoSphere Data Replication for Db2 for z/OS

These products provide the capability to asynchronously copy changes to data held in IMS or Db2 databases or VSAM files from a source to target copy. Use fine-grained controls to precisely define what data is critical to your workload and needs to be copied in real time between the source and target. Unlike disk replication solutions that are application or data-neutral and work at the z/OS volume level, software replication does not provide a mechanism for copying all possible data types in your environment. As such, it is suited to provide only a CA/DR solution for specific workloads that can tolerate only the IMS, Db2 or VSAM database-resident information to be copied between locations. This situation is also discussed in Chapter 6, "IBM GDPS Continuous Availability solution" on page 179.

InfoSphere Data Replication for IMS for z/OS

IMS Replication provides the mechanisms for producing copies of your IMS databases and maintaining the currency of the data in near real time, typically between two systems separated by geographic distances. There is essentially no limit to the distance between source and target systems because the copy technique is asynchronous and uses TCP/IP as the protocol to transport the data over your wide area network (WAN).

IMS replication employs Classic data servers in the source and target systems to provide the replication services.

Classic source server

The Classic source server reads the IMS log data and packages changes to the specified databases into messages that are then sent through TCP/IP to the target location.

Classic target server

The Classic target server, running in the target location, receives messages from the source server and applies the changes to a replica of the source IMS database in near real time. IMS replication provides mechanisms to ensure that updates to a record in the source database are applied in the same sequence in the target replica. Furthermore, IMS replication maintains a *bookmark* to know where it reached in processing the IMS log data so that if any planned or unplanned outage occurs, it can later catch up knowing where it was at the time of the outage.

For more information, IBM Documentation.

InfoSphere Data Replication for VSAM for z/OS

VSAM replication is similar in structure to IMS replication. For CICS/VSAM workloads, the transaction data for selected VSAM data sets is captured by using the CICS log streams as the source. For non-CICS workloads, CICS VSAM Recovery (CICS VR) logs are used as the source for capturing VSAM update information. The updates are transmitted to the target by using TCP/IP, where they are applied to the target data sets upon receipt.

InfoSphere Data Replication for Db2 for z/OS

InfoSphere Replication Server for z/OS, as used in the GDPS Continuous Availability solutions, is also known as *Q replication*. It provides a high capacity and low latency replication solution that uses IBM WebSphere® MQ message queues to transmit data updates between source and target tables of a Db2 database.

Q replication is split into two distinct pieces:

► Q capture program or engine
► Q apply program or engine

Q capture

The Q capture program reads the Db2 logs or changes to the source table or tables that you want to replicate. These changes are then put into WebSphere MQ messages and sent across the WebSphere MQ infrastructure to the system where the target table is. There, they are read and applied to the target table by the Q apply program.

The Q capture program is flexible in terms of what can be included or excluded from the data that is sent to the target and even the rate at which data is sent can be modified if required.

By the nature of the method of Q replication, the replication of data is an asynchronous process. Even so, an RPO of a few seconds is possible even in high update environments.

Q apply

The Q apply program takes WebSphere MQ messages from a receive queue, or queues and then applies the changes that are held within the message to the target tables. The Q apply program is designed in such a way to use parallelism to keep up with updates to multiple targets while maintaining any referential integrity constraints between related target tables.

Both the Q capture and Q apply programs have mechanisms to track what was read from the logs and sent to the target site, and what was read from the receive queues and applied to the target tables, including any dependencies between updates.

This situation in turn provides data consistency and allows for restart of both the capture and apply programs, if required or in failures.

For more information about Q replication, see IBM Documentation.

2.5 Tape-resident data

Operational data, that is, data that is used directly by applications supporting users, is normally found on disk. However, there is another category of data (called *support data*) that supports the operational data; this data often is in tape subsystems. Support data typically covers migrated data, point-in-time (PiT) backups, archive data, and other data. For sustained operation in the failover site, the support data is indispensable. Furthermore, some enterprises have mission-critical data that is only on tape. You need a solution to ensure that tape data is readily accessible at your recovery site.

Just as you mirror your disk-resident data to protect it, you can mirror your tape-resident data. GDPS supports management of the IBM TS7700[5]. See section 3.1.2, "Protecting tape data" on page 59 for details about GDPS TS7700 support. The IBM TS7700 provides comprehensive support for replication of tape data. For more information about the TS7700 technology that complements GDPS for tape data, see *IBM TS7700 Release 5.3 Guide*, SG24-8464.

---

[5] At the time of writing, the TS7700 management support is available only in GDPS Metro.

2.6 FlashCopy

FlashCopy provides a PiT copy of a volume, with almost instant availability for the user of both the source and target volumes. There is also a data set-level FlashCopy supported for z/OS volumes. Only a minimal interruption is required for the FlashCopy relationship to be established. Then, the copy is created by the disk subsystem, with minimal impact on other disk subsystem activities. The volumes created when you use FlashCopy to copy your secondary volumes are called *tertiary volumes*.

FlashCopy and disaster recovery

FlashCopy has specific benefits in relation to DR. For example, consider what happens if you temporarily lose connectivity between primary and secondary Metro Mirror volumes. At the point of failure, the secondary volumes are consistent. However, during the period when you are resynchronizing the primary and secondary volumes, the secondary volumes are inconsistent (because the updates are not applied in the same time sequence that they were written to the primaries). So, what happens if you have a disaster during this period? If it is a real disaster, your primary disk subsystem is a smoldering lump of metal on the computer room floor. And your secondary volumes are inconsistent, so those volumes are of no use to you either.

So, how do you protect yourself from such a scenario? One way (a best practice) is to take a FlashCopy of the secondary volumes just before you start the resynchronization process. This situation at least ensures that you have a consistent set of volumes in the recovery site. The data might be several hours behind the primary volumes, but even data a few hours old that is consistent is better than current, but unusable, data.

An extra benefit of FlashCopy is that it performs DR tests while still retaining DR readiness. The FlashCopy volumes that you created when doing the resynchronization can be used to enable frequent testing (thus ensuring that your recovery procedures continue to be effective) without having to use the secondary volumes for that testing.

FlashCopy can operate in several modes. GDPS uses one of the following modes of FlashCopy, depending on the GDPS offering:

COPY   When the volumes are logically copied, the FlashCopy session continues as a background operation, physically copying all the data from the source volume to the target. When the volumes are physically copied, the FlashCopy session ends. In this mode, the FlashCopy target physical volume is a mirror image of the source volume at the time of the FlashCopy.

NOCOPY   When the volumes are logically copied, a FlashCopy session continues as a background operation, physically copying only those tracks later updated by write operations to the source volume. In this mode, the FlashCopy target physical volume contains only data that was changed on the source volume after the FlashCopy.

NOCOPY2COPY   Change existing FlashCopy relationship from NOCOPY to COPY. This action can be done dynamically. When one or more NOCOPY relationships exist for a source volume, NOCOPY2COPY initiates a background copy for all target relationships with intersecting source extents from the point in time the NOCOPY was issued. Upon completion of the background copy, the converted relationship or relationships are terminated.

INCREMENTAL   Allows repetitive FlashCopies to be taken, but only the tracks that changed since the last FlashCopy are copied to the target volume. This mode *refreshes* a FlashCopy relationship and bring the target up to the source's newly established PiT. Incremental FlashCopy helps reduce the background copy completion time when only a subset of data on either the source or target has changed, thus giving you the option to perform a FlashCopy on a more frequent basis.

CONSISTENT   This option is applicable to GDPS Metro and GDPS Metro HyperSwap Manager environments. It creates a consistent set of tertiary disks without suspending Metro Mirror. It uses the FlashCopy Freeze capability, which puts all source disks in Extended Long Busy (ELB) to ensure that the FlashCopy source disks are consistent before the PiT copy is made. After the source disks are consistent, the FlashCopy is taken (fast) and the Freeze is thawed.

Without this support, you would need to suspend Metro Mirror (planned freeze) and then resynchronize Metro Mirror to produce a consistent PiT copy of the secondary disks. HyperSwap would remain disabled from the time you suspended Metro Mirror until the mirror is full-duplex again; however, this process can take a long time depending on how much data was updated while Metro Mirror remained suspended. In comparison, with Consistent FlashCopy, HyperSwap is only disabled during the FlashCopy Freeze, which should be a few seconds.

GDPS gives you the capability to restrict the FlashCopy Freeze duration and to stop the FlashCopy operation if the FlashCopy Freeze time exceeds your threshold.

To create a consistent PiT copy of the primary disks without Consistent FlashCopy, you would need to somehow make sure that there is no I/O on the primary disks (effectively, you would need to stop the production systems). With Consistent FlashCopy, production systems continue to run and I/O is prevented during the few seconds until the FlashCopy Freeze completes. After the FlashCopy Freeze completes, the primary disks are in a consistent state, the FlashCopy operation itself is fast, and then the freeze is thawed and production systems resume I/O. Consistent FlashCopy can be used with COPY, NOCOPY, or INCREMENTAL FlashCopy.

If you plan to use FlashCopy, the source and target volumes must be within the same physical disk subsystem. This situation is a capacity planning consideration when configuring and planning for the growth of your disk subsystems.

If you performed a site switch to run in the recovery site, at some point you will want to return to the production site. To provide equivalent protection and testing capability no matter which site you are running in, consider providing FlashCopy capacity in both sites.

Furthermore, GDPS does not perform FlashCopy for selected volumes. The GDPS use of FlashCopy is for the purposes of protection during resynchronization and for testing. Both of these tasks require that a PiT copy for the entire configuration is made. GDPS FlashCopy support assumes that you provide FlashCopy target devices for the entire configuration and that every time GDPS performs a FlashCopy, it is for all secondary devices (GDPS Metro also supports FlashCopy for primary devices).

An exception to this rule is that GDPS can perform FlashCopy for a subset of the production volumes when FlashCopy is used for the purposes of Logical Corruption Protection (LCP). For more information about how GDPS uses FlashCopy technology to provide flexible testing and protection against various types of logical data corruption, including cyberattacks and internal threats, see section 9.2, "Introducing LCP and Testcopy Manager" on page 257.

User-initiated FlashCopy

User-initiated FlashCopy supports FlashCopy of all defined FlashCopy volumes by using panel commands, GDPS scripts, or GDPS Z NetView for z/OS commands, depending on which GDPS product is used.

Space-efficient FlashCopy

Space-efficient FlashCopy (FlashCopy SE) is functionally not much different from the standard FlashCopy. The concept of *space-efficient* with FlashCopy SE relates to the attributes or properties of a DS8000 volume. As such, a space-efficient volume can be used like any other DS8000 volume.

When a normal volume is created, it occupies the defined capacity on the physical drives. A space-efficient volume does not occupy physical capacity when it is initially created. Space gets allocated when data is written to the volume. This situation allows the FlashCopy target volume capacity to be thinly provisioned (that is, smaller than the full capacity of the source volume). In essence this means that when planning for FlashCopy, you may provision less disk capacity when using FlashCopy SE than when using standard FlashCopy, which can help lower the amount of physical storage that is needed by many installations.

All GDPS products support FlashCopy SE. Details of how FlashCopy SE is used by each offering is described in the chapter that is related to that offering.

2.7 Automation

If you have challenging recovery time and RPOs, implementing disk remote copy, software-based replication, tape remote copy, FlashCopy, and other processes are prerequisites for you to be able to recover from a disaster and meet your objectives. However, be sure that you realize that they are only enabling technologies. To achieve the stringent objectives placed on many IT departments today, it is necessary to tie those technologies together with automation and sound systems management practices. In this section, we discuss your need for automation to recover from an outage.

2.7.1 Recovery time objective

If you reached this point in the document, we presume that your RTO is a "challenge" to you. If you performed tape-based DR tests, you know that ensuring that all your data is backed up is only the start of your concerns. In fact, even getting all those tapes restored does not result in a mirror image of your production environment. You also need to get all your databases up to date, get all systems up and running, and then start all your applications.

Trying to drive all these tasks manually will, without question, prolong the whole process. Operators must react to events as they happen, while consulting recovery documentation. However, automation responds at machine speeds, meaning that your recovery procedures run without delay, resulting in a shorter recovery time.

2.7.2 Operational consistency

Imagine an average computer room scene immediately following a system failure. All the phones are ringing. Every manager within reach moves in to determine when everything will be recovered. The operators are frantically scrambling for procedures that are more than likely outdated. And the systems programmers are all vying with the operators for control of the consoles; in short, chaos.

Imagine, instead, a scenario where the only manual intervention is to confirm how to proceed. From that point on, the system recovers itself using well-tested procedures. How many people watch it does not matter because it will not make mistakes. And you can yell at it all you like, but it still behaves in exactly the manner it was in which it was programmed to behave. You do not need to worry about outdated procedures being used. The operators can concentrate on handing calls and queries from the assembled managers. And the systems programmers can concentrate on pinpointing the cause of the outage, rather than trying to get everything up and running again.

And this process is just for a system outage. Can you imagine the difference that well-designed, coded, and tested automation can make in recovering from a real disaster? Apart from speed, perhaps the biggest benefit that automation brings is consistency. If your automation is thoroughly tested, you can be assured that it will behave in the same way, time after time. When recovering from as rare an event as a real disaster, this consistency can be a lifesaver.

2.7.3 Skills impact

Recovering a computing center involves many complex activities. Training staff takes time. People come and go. You cannot be assured that the staff that took part in the last DR test are available to drive recovery from this real disaster. In fact, depending on the nature of the disaster, your skilled staff might not even be available to drive the recovery.

The use of automation removes these concerns as potential pitfalls to your successful recovery.

2.7.4 Summary

The technologies that you use to recover your systems all have various control interfaces. Automation is required to tie them all together so they can be controlled from a single point and your recovery processes can run quickly and consistently.

Automation is one of the central tenets of the GDPS offerings. By using the automation provided by GDPS, you save all the effort to design and develop this code yourself, and also benefit from the IBM experience with hundreds of clients across your industry and other industries.

2.8 Flexible server capacity

In this section, we discuss options for increasing your server capacity concurrently, for either planned upgrades or unplanned upgrades, to quickly provide the additional capacity you require on a temporary basis. These capabilities can be used for server or site failures, or they can be used to help meet the temporary peak workload requirements of clients.

The only capabilities that are described in this section are the ones that are used by GDPS. Other capabilities exist to upgrade server capacity, either on a temporary or permanent basis, but they are not covered in this section.

For more information about Capacity BackUp (CBU), see the *System z Capacity on Demand User's Guide,* SC28-6846.

2.8.1 Capacity BackUp upgrade

A CBU upgrade for IBM Z processors provides extra capacity (reserved processors) on a Central Processor Complex (CPC) that can be non-disruptively activated and brought into your configuration for use in an emergency. Use the CBU capability to run with only the capacity that you require for normal operations, and then quickly add capacity to the recovery site if you lose capacity at the application site.

The CBU contract also enables an agreed-on number of DR tests to run over the period of the contract. GDPS supports activating CBU for test purposes.

2.8.2 On/Off Capacity on Demand

On/Off Capacity on Demand (OOCoD) is a function that enables concurrent and temporary capacity growth of the server. The difference between CBU and OOCoD is that OOCoD is for planned capacity increases, and CBU is intended to replace capacity that is lost as a result of an unplanned event elsewhere. OOCoD can be used for client peak workload requirements, for any length of time, and it has a daily hardware and software charge.

OOCoD helps clients, with business conditions that do not justify a permanent upgrade in capacity to contain workload spikes that might exceed permanent capacity so that Service Level Agreements cannot be met. OOCoD can concurrently add processors (CPs, IFLs, ICFs, zAAPs, and zIIPs) up to the limit of the installed books of an existing server. It is restricted to double the currently installed capacity.

2.8.3 Capacity for Planned Events

Capacity for Planned Events (CPE) can be used to replace capacity because of relocation of workloads, such as during system migrations, data center or server relocation, recabling, or general work on the physical infrastructure of the data processing environment.

CPE can concurrently and temporarily (for 72 hours) activate more CPs, ICFs, IFLs, zAAPs, zIIPs, and SAPs to increase the CP capacity level, or a combination of these processors.

2.8.4 System Recovery Boost

System Recovery Boost (SRB) delivers substantially faster system shutdown and restart, short duration recovery process boosts for sysplex events (such as HyperSwap events), and enables faster catch up of the accumulated backlog of work after specific events, such as system restart.

SRB is available starting with the IBM z15® IBM Z processor.

2.8.5 Flexible Capacity for Cyber Resiliency

Use this offering for IBM servers beginning with the IBM z16™ to shift production capacity between participating IBM z16 servers at different sites. Production capacity can be moved to other servers to accommodate planned events such as server maintenance and for unplanned events like data center outages.

With Flexible Capacity for Cyber Resiliency, CBU may not be necessary, but it works with OOCoD to temporarily increase production capacity for managing peak workloads.

2.8.6 GDPS CBU, OOCoD, CPE, SRB, and Flexible Capacity

The GDPS temporary capacity management capabilities are related to the capabilities provided by the particular server system being provisioned. Processors before the IBM Z10 required that the full capacity for a CBU upgrade or OOCoD be activated, even though the full capacity might not be required for the particular situation at hand.

GDPS, with IBM Z10 and later generation systems, supports activating temporary capacity, such as CBU and OOCoD, based on a preinstalled capacity-on-demand record. In addition to the capability to activate the full record, GDPS also defines profiles that determine what will be activated. The profiles are used with a GDPS script statement and provide the flexibility to activate the full record or a partial record.

When temporary capacity upgrades are performed by using GDPS facilities, GDPS tracks activated CBU and OOCoD resources at a CPC level.

GDPS provides keywords in GDPS scripts to support activation and deactivation of the CBU, OOCoD, CPE, SRB, and Flexible Capacity functions.

GDPS allows definition of capacity profiles to add capacity to already running systems. Applicable types of reserved engines (CPs, zIIPs, zAAPs, IFLs, and ICFs) can be configured online to GDPS z/OS systems, to xDR-managed z/VM systems, and to CFs that are managed by GDPS.

When a GDPS z/OS system starts, GDPS automatically configures online any applicable reserved engines (CPs, zIIPs, and zAAPs) based on the LPAR profile. The online configuring of reserved engines is done only if temporary capacity was added to the CPC where the system is started by using GDPS facilities.

2.9 Cross-site connectivity considerations

When setting up a recovery site, there might be a sizeable capital investment to get started, but you might find that one of the largest components of your ongoing costs is related to providing connectivity between the sites. Also, the type of connectivity available to you can affect the recovery capability that you can provide. Conversely, the type of recovery capability you want to provide affects the types of connectivity you can use.

In this section, we list the connections that must be provided, from a simple disk remote copy configuration through to an Active/Active workload configuration. We briefly review the types of cross-site connections that you must provide for the different GDPS solutions and the technology that must be used to provide that connectivity. All of these descriptions relate solely to cross-site connectivity. We assume that you already have whatever intrasite connectivity is required.

2.9.1 Server-to-disk links

If you want to be able to use disks installed remotely from a system in the production site, you must provide channel connections to those disk control units.

Metro Mirror and MTMM-based solutions

For Metro Mirror and MTMM with GDPS, all secondary disks (both sets for MTMM) must be defined to and channel-accessible to the production systems for GDPS to be able to manage those devices.

If you foresee a situation where systems in the production site will be running off the secondary disks (for example, if you use HyperSwap), you need to provide connectivity equivalent to that provided to the corresponding primary volumes in the production site. The HyperSwap function nondisruptively swaps from the primary volume of a mirrored pair to what was the secondary volume.

If you do not have any cross-site disk accessing, minimal channel bandwidth (two FICON channel paths from each system to each disk subsystem) is sufficient.

Depending on your director and switch configuration, you might be able to share the director-to-director links between channel and Metro Mirror connections. For more information, see *IBM System z Connectivity Handbook,* SG24-5444.

HyperSwap across sites with less than full channel bandwidth

You might consider enabling unplanned HyperSwap to the secondary disks in the remote site even if you do not have sufficient cross-site channel bandwidth to sustain your production workload for normal operations. Assuming that a disk failure is likely to cause an outage and you must switch to using a disk in the other site, the unplanned HyperSwap might at least give you the opportunity to perform an orderly shutdown of your systems first. Shutting down your systems cleanly avoids the complications and longer restart time that is associated with crash-restart of application subsystems.

For GDPS Metro environments, the same consideration applies to enabling HyperSwap to the remote secondary copy: Channel bandwidth to the local secondary copy should not be an issue.

Global Mirror based solutions

For GM, the production systems would normally not have channel access to the secondary volumes.

Software replication-based solutions

As with other asynchronous replication technologies, if effectively unlimited distances are supported, there is no requirement for the source systems to have host channel connectivity to the data in the target site.

2.9.2 Data replication links

You need connectivity for your data replication activity for the following circumstances:

► Between storage subsystems (for Metro Mirror or GM)
► Across the WAN for software-based replication

Metro Mirror based and Global Mirror based solutions

The IBM Metro Mirror (including MTMM) and GM implementations use FCP links between the primary and secondary disk subsystems. The FCP connection can be direct, through a switch, or through other supported distance solutions (for example, Dense Wave Division Multiplexer, DWDM, or channel extenders).

Software-based solutions

Both IMS replication and Db2 replication use your WAN connectivity between the data source and the data target. Typically, for both, either natively or through IBM MQ for Z/OS, TCP/IP is the transport protocol that is used, although other protocols, such as LU6.2, are supported. It is beyond the scope of this book to go into detail about WAN design, but ensure that any such connectivity between the source and target have redundant routes through the network to ensure resilience from failures. There are effectively no distance limitations on the separation between source and target. However, the greater the distance between them affects the latency and the RPO that can be achieved.

2.9.3 Coupling links

Coupling links are required in a Parallel Sysplex configuration to provide connectivity from the z/OS images to the CF. Coupling links are also used to transmit timekeeping messages when Server Time Protocol (STP) is enabled. If you have a multisite Parallel Sysplex, you must provide coupling link connectivity between sites.

For distances greater than 10 km, either ISC3 or Parallel Sysplex InfiniBand Long Reach links must be used to provide this connectivity. The maximum supported distance depends on several things, including the particular DWDMs that are being used and the quality of the links.

Table 2-1 lists the distances that are supported by the various link types.

Table 2-1   Supported CF link distances

Link type	Link data rate	Maximum unrepeated distance	Maximum repeated distance
ISC-3	2 Gbps[a] 1 Gbps[b]	10 km 20 km[c]	200 km
Parallel Sysplex InfiniBand Long Reach 1X	5.0 Gbps 2.5 Gbps[d]	10 km	175 km
Parallel Sysplex InfiniBand 12X, for use within a data center	6 GBps 3 GBps[e]	150 meters	Not applicable

a. Gbps (gigabits per second).
b. RPQ 8P2197 provides an ISC-3 Daughter Card that clocks at 1 Gbps.
c. Requires RPQ 8P2197 and 8P2263 (IBM Z Extended Distance).
d. The Parallel Sysplex InfiniBand Long Reach feature negotiates to 1x InfiniBand single data rate link data rate of 2.5 Gbps if connected to qualified DWDM infrastructure that cannot support the 5 Gbps (1x InfiniBand double data rate) rate.
e. The Parallel Sysplex InfiniBand links negotiate to 12x InfiniBand single data rate link data rate of 3 GBps when connected to IBM Z9 servers.

2.9.4 Server Time Protocol

STP is a server-wide facility that is implemented in the Licensed Internal Code (LIC) of the IBM Z servers. It provides the capability for multiple servers to maintain time synchronization with each other. STP is the successor to the 9037 Sysplex Timer.

STP is designed for servers that are configured to be in a Parallel Sysplex or a basic sysplex (without a CF), and servers that are not in a sysplex, but need to be time-synchronized. STP is a message-based protocol in which timekeeping information is passed over data links between servers. The timekeeping information is transmitted over externally defined coupling links. Coupling links are used to transport STP messages.

If you are configuring a sysplex across two or more sites, you need to synchronize servers in multiple sites. For more information about STP, see *Server Time Protocol Planning Guide*, SG24-7280, and *Server Time Protocol Implementation Guide*, SG24-7281.

2.9.5 XCF signaling

One of the requirements for being a member of a sysplex is the ability to maintain XCF communications with the other members of the sysplex. XCF uses two mechanisms to communicate between systems: XCF signaling structures in a CF and channel-to-channel adapters. Therefore, if you are going to have systems in both sites that are members of the same sysplex, you must provide CF connectivity, CTC connectivity, or preferably both, between the sites.

If you provide both CF structures and CTCs for XCF use, XCF dynamically determines which of the available paths provides the best performance and use that path. For this reason, and for backup in a failure, we suggest providing *both* XCF signaling structures and CTCs for XCF cross-site communication.

2.9.6 HMC and consoles

To be able to control the processors in the remote center, you need to have access to the local area network (LAN) containing the SEs and HMCs for the processors in that location. Such connectivity is typically achieved by using bridges or routers.

If you are running systems at the remote site, you also want to be able to have consoles for those systems. Two options are 2074 control units and OSA-ICC cards. Alternatively, you can use SNA consoles, but they cannot be used until IBM VTAM® is started, so they cannot be used for initial system loading.

2.9.7 Connectivity options

> **Note:** WAN connectivity options are not covered in this book. Table 2-2, except for HMC connectivity, is predominantly related to disk replication solutions.

Now that we explained what you need to connect across the two sites, we briefly review the most common options for providing that connectivity. There are several ways to provide all this connectivity, from direct channel connection through to DWDMs. Table 2-2 on page 41 lists the different options. The distance that is supported varies by device type and connectivity method.

Table 2-2   Cross-site connectivity options

Connection type	Direct (unrepeated)	Switch and director or cascaded directors	DWDM	Channel extender
Server to disk	Yes	Yes	Yes	Yes
Disk Remote copy	Yes	Yes	Yes	Yes
Coupling links	Yes	No	Yes	No
STP (coupling links)	Yes	No	Yes	No
XCF signaling	Yes	Yes (CTC) No (coupling links)	Yes	Yes (CTC only) No (coupling links)
HMC/consoles	Yes	Yes	Yes	Yes

For more information about options and distances that are possible, see *IBM System z Connectivity Handbook,* SG24-5444.

FICON switches/directors

For more information about IBM Z qualified FICON and FCP products and products that support mixing FICON and FCP within the same physical Fibre Channel switch or FICON director, see the I/O Connectivity web page.

The maximum unrepeated distance for FICON is typically 10 km. However, FICON switches can be used to extend the distance from the server to the control unit further with the use of a cascaded configuration. The maximum supported distance for the interswitch links (ISLs) in this configuration is technology- and vendor-specific.

No matter what the case might be, if the property between the two sites is not owned by your organization, you need a vendor to provide dark fiber between the two sites because FICON switches/directors cannot be directly connected to telecommunication lines.

For more information, see *IBM System z Connectivity Handbook,* SG24-5444.

Wavelength Division Multiplexing

A Wavelength Division Multiplexor (WDM) is a high-speed, high-capacity, scalable fiber optic data transport system that uses DWDM or Course Wavelength Division Multiplexing (CWDM) technology to multiplex several independent bit streams over a single fiber link, thus making optimal use of the available bandwidth.

WDM solutions that support the protocols that are described in this book generally support metropolitan distances in the range of tens to a few hundred kilometers. The infrastructure requirements and the supported distances vary by vendor, model, and even by features on a model.

More specifically, several qualified WDM solutions support the following key protocols that are used in a GDPS solution:

► FICON
► InterSystem Channel (ISC-3)
► Parallel Sysplex InfiniBand Long Reach links
► STP over ISC-3 Peer Mode or Parallel Sysplex InfiniBand Long Reach
► Potentially, protocols that are not IBM Z protocols

Given the criticality of these links for transport of data and timing information, it is important to use only qualified WDM vendor solutions when extending Parallel Sysplexes to more than one site (as is often done as part of a GDPS configuration).

The latest list of qualified WDM vendor products, along with links to corresponding IBM Redpaper publications for each product, is available at the IBM Resource Link web page (sign-in required).

Also see "Hardware products for servers" on the Library page.

Channel extenders

Channel extenders are special devices that are connected in the path between a server and a control unit, or between two control units. Channel extenders extend connections over greater distances than that provided by DWDM. Distances that are supported with channel extenders are virtually unlimited.

Unlike DWDMs, channel extenders support connection to telecom lines, removing the need for dark fiber. This situation can make channel extenders more flexible because access to high-speed telecoms is often simpler to obtain than access to dark fiber.

However, channel extenders typically do not support the same range of protocols as DWDMs. In a IBM Z context, channel extenders support IP connections (for example, connections to OSA adapters), FCP and FICON channels, but not coupling links or time synchronization-related links.

For much more detailed information about the options and distances that are possible, see *IBM System z Connectivity Handbook,* SG24-5444.

More information about channel extenders that are qualified to work with IBM storage is available to download from the DS8000 Series Copy Services Fibre Channel Extension Support Matrix web page.

2.9.8 Single points of failure

When planning to connect systems across sites, it is vital to do as much as you possibly can to avoid all single points of failure. Eliminating all single points of failure makes it simpler to distinguish between a connectivity failure and a failure of the remote site. The recovery actions that you take are different, depending on whether the failure you just detected is a connectivity failure or a real site failure.

If you have only a single path, you do not know whether it was the path or the remote site that went down. If you have no single points of failure and everything disappears, there is an good chance that it was the site that went down. Any other mechanism to distinguish between a connectivity failure and a site failure (most likely human intervention) cannot react with the speed that is required to drive effective recovery actions.

2.10 Testing considerations

Testing your DR solution is a required and essential step in maintaining DR readiness. Many enterprises have business or regulatory requirements to conduct periodic tests to ensure that the business can recover from a wide-scale disruption and recovery processes meet RTO and RPO requirements. The only way to determine the effectiveness of the solution and your enterprise's ability to recover from a disaster is through comprehensive testing.

One of the most important test considerations in developing a DR test plan is to make sure that the testing you conduct truly represents the way you would recover your data and enterprise. This way, when you must recover following a disaster, you can recover the way you were testing, thus improving the probability that you will be able to meet the RTO and RPO objectives that are established by your business.

Testing disk mirroring-based solutions

When conducting DR drills to test your recovery procedures, without extra disk capacity to support FlashCopy, the mirroring environment is suspended so the secondary disks can be used to test your recovery and restart processes. When testing is completed, the mirror must be brought back to a duplex state again. During this window, until the mirror is back to a duplex state, the enterprises ability to recover from a disastrous event is compromised.

If this situation is not acceptable or your enterprise must perform periodic DR tests while maintaining a disaster readiness posture, you must provide more disk capacity to support FlashCopy. The additional FlashCopy device can be used for testing your recovery and restart procedures while the replication environment is running. This situation ensures that a current and consistent copy of the data is available, and that disaster readiness is maintained throughout the testing process.

The additional FlashCopy disk can also be used to create a copy of the secondary devices to ensure that a consistent copy of the data is available if a disaster-type event occurs during primary and secondary volume resynchronization.

From a business perspective, installing the additional disk capacity to support FlashCopy means incurring extra expense. However, not having it can result in compromising the enterprise's ability to recover from a disastrous event, or in extended recovery times and exposure to more data loss.

Testing software replication solutions

Similar in some instances to the situation described for testing disk-based mirroring solutions, if you test on your target copy of your database or databases, you must pause the replication process. Potentially, you might have to also re-create the target copy from scratch by using the source copy as input when the test is complete.

It would be normal to test the recovery procedures and operational characteristics of a software replication solution in a pre-production environment that as close as possible reflects the production environment.

However, because of the nature of software replication solutions, there is limited recovery that is required in the target site. Updates will either have been sent (and applied) from the source site, or they will not; the apply process is based on completed units of work, so there should be no issue with incomplete updates arriving from the source site. The testing is more likely to be related to the process for handling the potential data loss and any possible handling of collisions that are caused by the later capture/apply of stranded transactions with other completed units of work that might have occurred following an outage or disaster.

Testing methodology

How you approach your DR testing is also an important consideration. Most enterprises aim to do most disruptive testing in a test or "sandbox" environment. This testing ideally will closely resemble the production environment so that the testing scenarios that are done in the sandbox are representative of what is applicable also in your production environment.

Other enterprises might decide to simulate a disaster in the production environment to really prove that the processes and technology deliver what is required. A disaster can surface to the technology in different ways (for example, different components failing in different sequences), so the scenarios you devise and test should consider these possible variations.

A typical approach to DR testing in production is to perform some form of a planned site switch. In such a test, the production service is closed down in a controlled manner where it normally runs, and then restarted in the DR site. This type of test demonstrates that the infrastructure in the DR site can run the services within the scope of the test, but given the brief duration of such tests (often over a weekend only) not all possible workload scenarios can be tested.

For this reason, consider the ability to move the production services to the DR site for an extended period (weeks or months) to give an even higher degree of confidence. This ability to "toggle" production and DR locations can provide other operational benefits, such as performing a preemptive switch because of an impending event, along with increased confidence in being able to run following a DR invocation.

With this approach, it is important to continue to test the actual DR process in your test environment because a real disaster is unlikely to happen in a way where a controlled shutdown is possible. Those processes must then be carefully mapped across to the production environment to ensure success in a DR invocation.

In some industries, regulation might dictate or at least suggest guidelines about what constitutes a valid DR test, and this situation also needs to be considered.

2.11 Summary

In this chapter, we covered the major building blocks of an IT resilience solution. We discussed providing CA for normal operations, the options for keeping a consistent offsite copy of your disk and tape-based data, the need for automation to manage the recovery process, and the areas you need to consider when connecting across sites.

In the next few chapters, we discuss the functions that are provided by the various offerings in the GDPS family.

3

IBM GDPS Metro

In this chapter, we discuss the capabilities and prerequisites of the GDPS Metro offering. GDPS Metro supports both planned and unplanned situations, helping to maximize application availability and providing business continuity. A GDPS Metro solution delivers the following benefits:

► Near-continuous availability (CA)

► Disaster recovery (DR) across metropolitan distances

► Protection against multiple failures

 GDPS Metro implemented in a dual leg configuration maintains three copies of your data so that even if one copy becomes unavailable, IBM Geographically Dispersed Parallel Sysplex (GDPS) can continue to provide near-CA and DR by using the remaining two copies.

► Recovery time objective (RTO) less than an hour

► Recovery point objective (RPO) of zero

The functions provided by GDPS Metro fall into two categories: Protecting your data and controlling the resources that are managed by GDPS. The following functions are among the ones that are included:

► Protecting your data:

 – Ensuring the consistency of the secondary copies of your data in a disaster or suspected disaster, including the option to also ensure zero data loss (ZDL)

 – Transparent switching to either of the secondary disk sets by using HyperSwap

► Controlling the resources managed by GDPS during normal operations, planned changes, and following a disaster:

 – Monitoring and managing the state of the production z/OS systems and LPARs (shutdown, activating, deactivating, IPL, and automated recovery)

 – Monitoring and managing z/VM guests (shutdown, activating, deactivating, IPL, and automated recovery)

 – Managing the couple data sets (CDS) and coupling facility (CF) recovery

- Support for switching your disk, or systems, or both, to another site
- User-customizable scripts that control how GDPS Metro reacts to specified error situations, which can also be used for planned events

This chapter includes the following topics:

3.1 Introducing GDPS Metro

GDPS Metro is a CA and DR solution that handles many types of planned and unplanned outages. As described in Chapter 1, "Introducing business resilience and the role of IBM GDPS" on page 1, most outages are planned, and even among unplanned outages, most are not disasters. GDPS Metro provides capabilities to address the required levels of availability across these outages and in a disaster scenario. These capabilities are described in this chapter.

3.1.1 Protecting data integrity and data availability with GDPS Metro

In 2.2, "Data consistency" on page 18, we point out that data integrity across primary and secondary volumes of data is essential to perform a database restart and accomplish an RTO of less than an hour. This section includes details about how GDPS Metro automation provides both data consistency if there are mirroring problems and data availability if there are primary disk problems.

The following types of disk problems trigger a GDPS automated reaction:

- Mirroring problems (Freeze triggers). No problem exists writing to the primary disk subsystem, but a problem exists mirroring the data to one or both of the secondary disk subsystems. For more information, see "GDPS Freeze function for mirroring failures" on page 47.
- Primary disk problems (HyperSwap triggers). There is a problem writing to the primary disk: either a hard failure, or the disk subsystem is not accessible.

GDPS Freeze function for mirroring failures

GDPS uses automation, which is keyed off events or messages, to stop all mirroring for a replication leg when a remote copy failure occurs between one or more of the primary/secondary disk subsystem pairs on that replication leg. In particular, the GDPS automation uses the IBM PPRC Freeze and Run architecture, which is implemented as part of Metro Mirror on IBM disk subsystems and also by other enterprise disk vendors. In this way, if the disk hardware supports the Freeze and Run architecture, GDPS can ensure consistency across all data in the sysplex (consistency group), regardless of disk hardware type.

This preferred approach differs from proprietary hardware approaches that work only for one type of disk hardware. For more information about data consistency with synchronous disk mirroring, see "Metro Mirror data consistency" on page 24.

When a mirroring failure occurs, this problem is classified as a Freeze trigger and GDPS stops activity across *all* disk subsystems for the affected replication leg at the time the initial failure is detected, thus ensuring that the dependent write consistency of the secondary disks for that replication leg is maintained. In a dual-leg environment, mirroring activity for the other replication leg is not affected by the freeze.

The following process occurs when a GDPS performs a Freeze:

1. Remote copy is suspended for all device pairs on the affected replication leg.

2. While the suspend command is being processed for each logical subsystem (LSS), each device goes into a long busy state. When the suspend completes for each device, z/OS marks the device unit control block (UCB) in all connected operating systems to indicate an Extended Long Busy (ELB) state.

3. No I/Os can be issued to the affected devices until the ELB is thawed with the PPRC Run (or "thaw") action or until it times out. (The consistency group timer setting commonly defaults to 120 seconds, although for most configurations a longer ELB is preferable.)

4. All paths between the Metro Mirrored disks on the affected replication leg are removed, which prevents further I/O to the associated secondary disks if Metro Mirror is accidentally restarted.

Because no I/Os are processed for a remote-copied volume during the ELB, dependent write logic ensures the consistency of the affected secondary disks. GDPS performs a Freeze for all LSS pairs that contain GDPS managed mirrored devices.

> **Important:** Because of the dependent write logic, it is not necessary for all LSSs to be frozen at the same instant. In a large configuration with many thousands of remote copy pairs, it is not unusual to see short gaps between the times when the Freeze command is issued to each disk subsystem. However, because of the ELB such gaps are not a problem.

After GDPS performs the Freeze and the consistency of the secondary disks on the affected leg is protected, the action GDPS takes next depends on the client's PPRCFAILURE policy (also known as Freeze policy). For more information about the actions GDPS takes based on this policy, see "Freeze policy (PPRCFAILURE policy) options" on page 48.

GDPS Metro uses a combination of storage subsystem and sysplex triggers to automatically secure, at the first indication of a potential disaster, a data-consistent secondary copy of your data using the Freeze function. In this way, the secondary copy of the data is preserved in a consistent state, even before production applications are aware of any issues.

Ensuring the data consistency of the secondary copy ensures that a normal system restart can be performed instead of having to perform database management system (DBMS) forward recovery actions. This approach is an essential design element of GDPS to minimize the time to recover the critical workloads in a disaster in the primary site.

You can appreciate why such a process must be automated. When a device suspends, there is not enough time to start a manual investigation process. The entire mirror for the affected leg must be frozen by stopping further I/O to it, and then the policy indicates whether production continues to run with mirroring temporarily suspended, or whether all systems should be stopped to ensure ZDL.

In summary, a freeze is triggered as a result of a Metro Mirror suspension event for any primary disk in the GDPS configuration; that is, at the first sign that a duplex mirror that is going out of the duplex state. When a device suspends, all attached systems are sent a *State Change Interrupt* (SCI). A message is issued in all of those systems and then each system must issue multiple I/Os to investigate the reason for the suspension event.

When GDPS performs a freeze, all primary devices in the Metro Mirror configuration suspend for the affected replication leg. This suspension can result in significant SCI traffic and many messages in all systems. GDPS, with z/OS and microcode on the DS8000 disk subsystems, supports reporting suspensions in a summary message per LSS instead of at the individual device level. This feature is known as *Summary Event Notification for PPRC Suspends (PPRCSUM)*. When compared to reporting suspensions on a per device basis, PPRCSUM dramatically reduces the message traffic and extraneous processing that is associated with Metro Mirror suspension events and freeze processing.

Freeze policy (PPRCFAILURE policy) options

When a mirroring failure is detected on a replication leg, GDPS automatically and unconditionally performs a Freeze of that leg to secure a consistent set of secondary volumes in case the mirroring failure might be the first indication of a site failure. Because the primary disks are in the ELB state as a result of the freeze and the production systems are locked out, GDPS must act. Here, there is no time to interact with the operator on an event-by-event basis. The action must be taken immediately. The action to be taken is determined by a customer policy setting, that is, the PPRCFAILURE policy option (also known as the Freeze policy option). GDPS will use this same policy setting after every Freeze event to determine what its next action should be. The policy can be specified at a leg level allowing a different policy specification for each of the replication legs. The following options are available:

► `PPRCFAILURE=GO` (Freeze and Go)

   GDPS allows production systems to continue operation after mirroring is suspended.

► `PPRCFAILURE=STOP` (Freeze and Stop)

   GDPS resets production systems while I/O is suspended.

► `PPRCFAILURE=STOPLAST`

   This option is only relevant to dual-leg configurations. When it is specified, GDPS checks the mirroring status of the other replication leg. If the status of the other leg is OK, GDPS performs a Go. If not, and this leg is the last viable leg that GDPS has frozen, GDPS performs a Stop.

► `PPRCFAILURE=COND` (Freeze and Stop conditionally)

GDPS tries to determine whether a secondary disk caused the mirroring failure. If so, GDPS performs a Go. If not, GDPS performs a Stop.

► `PPRCFAILURE=CONDLAST`

This option is only relevant to dual-leg configurations. When it is specified, GDPS checks the mirroring status of the other replication leg. If the status of the other leg is OK, GDPS performs a Go. If not (the freeze was performed on the last viable leg), GDPS tries to determine whether a secondary disk caused the mirroring failure. If so, GDPS performs a Go. If not, GDPS performs a Stop.

Freeze and Go

With this policy, after performing the Freeze, GDPS performs a Run action against all primary LSSs, which is also known as performing a Go. Performing a Go removes the ELB and allows production systems to continue to use these devices. The devices are in remote copy-suspended mode in relation to the secondary devices on the affected leg, so any further writes to these devices are no longer being mirrored to the secondary devices on that leg (writes continue to be mirrored to the secondary devices on the other leg in dual-leg configurations, assuming that mirroring on that leg is in duplex status at the time). However, changes are tracked by the hardware so that later only the changed data is resynchronized to the secondary disks and the affected leg.

With this policy, you avoid an unnecessary outage for a false freeze event, that is, if the trigger is a transient event. However, if the trigger turns out to be the first sign of an actual disaster, you might continue operating for an amount of time before all systems fail. Any updates made to the primary volumes during this time are not replicated to the secondary disk on the affected leg, and therefore are lost if you end up having to recover on those secondary disks. In addition, because the CF structures were updated after the secondary disks were frozen, the CF structure content is not consistent with the secondary disks. Therefore, the CF structures in either site cannot be used to restart workloads and log-based restart must be used when restarting applications.

This recovery is not a full forward recovery. It is forward recovery of any data, such as Db2 group buffer pools (GBPs) that might have existed in a CF but might not have been written to disk yet. This recovery results in prolonged recovery times. The duration depends on how much such data existed in the CFs then. With a Freeze and Go policy, you might consider tuning applications such as Db2, which can harden such data on disk more frequently than otherwise.

Freeze and Go is a high availability (HA) option that avoids production outage for false freeze events. However, it carries a potential for data loss.

Freeze and Stop

With this policy, you can be assured that no updates are made to the primary volumes after the Freeze because all systems that can update the primary volumes are reset. This situation ensures that no more updates can occur to the primary disks because such updates would not be mirrored to the affected secondary disk, meaning that it would not be possible to achieve ZDL if a failure occurs (or if the original trigger was an indication of a catastrophic failure) and recovery on the affected secondary disk is required.

You can choose to restart the systems when you want. For example, if this freeze was a false freeze (that is, a false alarm), then you can quickly resynchronize the mirror and restart the systems only after the mirror is duplexed.

If you are using duplexed CF structures along with a Freeze and Stop policy, it might seem that will use the duplexed instance of your structures if you must recover and restart your workload with the frozen secondary copy of your disks. However, this case is not always true. There can be rolling disaster scenarios where before, after, or during the freeze event, there is an interruption (perhaps failure of CF duplexing links) that forces CFRM to drop out of duplexing.

There is no guarantee that it is the structure instance in the surviving site that is kept. It is possible that CFRM keeps the instance in the site that is about to totally fail. In this case, there is not an instance of the structure in the site that survives the failure.

To summarize, with a Freeze and Stop policy, if there is a surviving, accessible instance of application-related CF structures, this instance is consistent with the frozen secondary disks. However, depending on the circumstances of the failure, even with structures that are duplexed across two sites, you are not 100% guaranteed to have a surviving, accessible instance of the application structures. Therefore, you must have the procedures in place to restart your workloads without the structures.

Although a Stop policy can be used to ensure no data loss, if a failure occurs that is a false freeze event, that is, it is a transient failure that did not necessitate recovery by using the frozen disks, it results in unnecessarily stopping the systems.

Freeze and Stop last

For dual-leg configurations, when this policy option is specified, after the Freeze, GDPS checks the status of mirroring on the other replication leg (the leg other than the one that was frozen) to determine whether the leg that frozen was the last leg actively replicating data. If the other leg is still actively replicating data, GDPS performs a Go. But if the other leg is already frozen or mirroring status is not OK, GDPS performs a Stop.

When only one replication leg is defined in your configuration (you have only one secondary copy of your data), using this policy specification is the same as using a Freeze and Stop policy.

Freeze and Stop conditional

Field experience has shown that most of the Freeze triggers are not necessarily the start of a rolling disaster, but are "False Freeze" events that do not necessitate recovery on the secondary disk. Examples of such events include connectivity problems to the secondary disks and secondary disk subsystem failure conditions.

With a COND policy, the action that GDPS takes after it performs the Freeze is conditional. GDPS tries to determine whether the mirroring problem was as a result of a permanent or temporary secondary disk subsystem problem:

► If GDPS can determine that the freeze was triggered as a result of a secondary disk subsystem problem, GDPS performs a Go. That is, it allows production systems to continue to run by using the primary disks. However, updates are not mirrored until the secondary disk can be fixed and Metro Mirror can be resynchronized.

► If GDPS cannot find out that the cause of the freeze was a secondary disk subsystem, GDPS operates on the assumption that it is still the beginning of a rolling disaster in the primary site and performs a Stop, resetting all the production systems to ensure ZDL. GDPS cannot always detect that a particular freeze trigger was caused by a secondary disk, and that some freeze events that are in fact caused by a secondary disk could still result in a Stop.

For GDPS to determine whether a freeze trigger might have been caused by the secondary disk subsystem, the IBM DS8000 disk subsystems provide a special query capability that is known as the *Query Storage Controller Status* microcode function. If all disk subsystems in the GDPS managed configuration support this feature, GDPS uses this special function to query the secondary disk subsystems in the configuration to understand the state of the secondaries and if one of these secondaries might have caused the freeze. If you use the `COND` policy setting but all disks in your configuration do not support this function, GDPS cannot query the secondary disk subsystems, and the resulting action is a Stop.

This option can provide a good compromise where you can minimize the chance that systems are stopped for a false freeze event and increase the chance of achieving ZDL for a real disaster event.

Freeze and Stop conditional last

For dual-leg configurations, when this policy option is specified, after the Freeze, GDPS checks the status of mirroring on the other replication leg (the leg other than the one that was frozen) to determine whether the leg frozen was the last leg actively replicating data. If the other leg is still actively replicating data, GDPS performs a Go. If the other leg is already frozen or mirroring status is not OK, GDPS performs conditional Stop processing; that is, it queries the secondary disk subsystem and performs a Go if, as a result of the query, it determines that the freeze was caused by the secondary, but performs a Stop if it cannot determine with certainty that the problem was caused by the secondary.

When you have only one replication leg that is defined in your configuration (you have only one secondary copy of your data), using this policy specification is the same as using a Freeze and Stop conditional policy.

PPRCFAILURE policy selection considerations

The PPRCFAILURE policy option specification directly relates to RTOs and RPOs, which are business objectives. Therefore, the policy option selection is really a business decision rather than an IT decision. If data associated with your transactions is high-value, it might be more important to ensure that no data that is associated with your transactions is ever lost, so you might decide on a Freeze and Stop policy.

If you have huge volumes of relatively low-value transactions, you might be willing to risk some lost data in return for avoiding unnecessary outages with a Freeze and Go policy. The Freeze and Stop Conditional policy attempts to minimize the chance of unnecessary outages and the chance of data loss, but there is still a risk of either, however small.

The various PPRCFAILURE policy options, which are combined with the fact that the policy options are specified on a per replication leg basis (different policies can be specified for different legs), gives you the flexibility to refine your policies to meet your unique business goals.

For example, if your RPO is zero, you can use the following PPRCFAILURE policy:

► For RL2, Freeze and Stop (`PPRCFAILURE=STOP`)

Because RS3 is your DR copy and you must ensure that you never lose data should you ever have to recover and run on the RS3 disk, you must always unconditionally stop the systems to ensure that no further updates occur to the primary disks that might be lost in a recovery scenario.

► For RL1, Freeze and Stop on the last leg only (`STOPLAST`)

You do not need to take a production outage when Metro Mirror freezes on the HA leg if RL2 is still functional and continues to provide DR protection. However, if RL2 is not functional when Metro Mirror on RL1 suspends, you might want to at least retain the capability to recover on RS2 disk with ZDL if it becomes necessary.

However, if you want to avoid unnecessary outages at the risk of losing data if there is an actual disaster, you can specify Freeze and Go for both of your replication legs.

GDPS HyperSwap function

If there is a problem writing or accessing the *primary disk* because of a failing, failed, or inaccessible or non-responsive disk, there is a need to swap from the primary disks to one of the sets of secondary disks.

GDPS Metro delivers a powerful function that is known as HyperSwap. HyperSwap swaps from using the primary devices in a mirrored configuration to using what was one of the sets of secondary devices, in a manner that is transparent to the production systems and applications that use these devices. Before the availability of HyperSwap, a transparent disk swap was not possible. All systems that use the primary disk would have been shut down (or might have failed, depending on the nature and scope of the failure) and would have been started by using the secondary disks. Disk failures were often a single point of failure for the entire sysplex.

With HyperSwap, such a switch can be accomplished without IPL and with just a brief hold on application I/O. The HyperSwap function is controlled by automation, thus allowing all aspects of the disk configuration switch to be controlled through GDPS.

HyperSwap can be started in two ways:

► Planned HyperSwap

A planned HyperSwap is started by operator action by using GDPS facilities. One example of a planned HyperSwap is where a HyperSwap is initiated before planned disruptive maintenance to a disk subsystem.

► Unplanned HyperSwap

An unplanned HyperSwap is started automatically by GDPS, triggered by events that indicate the primary disk problem.

Primary disk problems can be detected as a direct result of an I/O operation to a specific device that fails because of a reason that indicates a primary disk problem such as:

– No paths available to the device
– Permanent error
– I/O timeout

In addition to a disk problem being detected as a result of an I/O operation, it is also possible for a primary disk subsystem to proactively report that it is experiencing an acute problem. The IBM DS8000 provides a special microcode function that is known as the *Storage Controller Health Message Alert* capability. Problems of different severity are reported by disk subsystems that support this capability. Those problems classified as *acute* are also treated as HyperSwap triggers. After systems are swapped to use the secondary disks, the disk subsystem and operating system can try to perform recovery actions on the former primary without impacting applications since the applications are no longer using those disks.

Planned and unplanned HyperSwap have requirements in terms of the physical configuration, such as having to be symmetrically configured. While a client's environment meets these requirements, there is no special enablement required to perform planned swaps. Unplanned swaps are not enabled by default and must be enabled explicitly as a policy option. This is described in more detail in "Preferred Swap Leg and HyperSwap (Primary Failure) policy options" on page 54.

When a swap is initiated, GDPS always validates various conditions to ensure that it is safe to swap. For example, if the mirror is not fully duplex on a leg, that is, not all volume pairs are in a duplex state, a swap cannot be performed on that leg. The way that GDPS reacts to such conditions changes depending on the condition that is detected and whether the swap is a planned or unplanned swap.

Assuming that there are no show-stoppers and the swap proceeds, for both planned and unplanned HyperSwap, the systems that are using the primary volumes experience a temporary pause in I/O processing. GDPS blocks I/O both at the channel subsystem level by performing a Freeze, which results in all disks going into ELB, and also in all systems, where I/O is quiesced at the operating system (UCB) level. This action ensures that no systems use the disks until the switch is complete. During the time when I/O is paused, the following process is completed:

1. The Metro Mirror configuration is *physically switched*. This process includes physically changing the secondary disk status to primary. Secondary disks are protected and cannot be used by applications. Changing their status to primary allows them to come online to systems and be used.

2. The disks are logically switched in each of the systems in the GDPS configuration. This involves switching the internal pointers in the operating system control blocks (UCBs). After the switch, the operating system will point to the former secondary devices which will be the new primary devices.

3. Finally, the systems resume operation by using the new, swapped-to primary devices. The applications are not aware of the fact that different devices are now being used.

This brief pause during which systems are locked out of performing I/O is known as the *User Impact Time*. In benchmark measurements at IBM that use currently supported releases of GDPS and IBM DS8000 disk subsystems, the User Impact Time to swap 10,000 pairs across 16 systems during an unplanned HyperSwap was less than 10 seconds. Most implementations are much smaller than this and typical impact times in a well-configured environment that uses the most current storage and server hardware are measured in seconds. Although the results depend on your configuration, these numbers give you a high-level idea of what to expect.

HyperSwap can run on either replication leg in a GDPS Metro dual-leg environment. For a planned swap, you must specify which leg you want to use for the swap. For an unplanned swap, which leg is chosen depends on many factors, including your HyperSwap policy. This is described in more detail in "Preferred Swap Leg and HyperSwap (Primary Failure) policy options" on page 54.

After a replication leg is selected for the HyperSwap, GDPS swaps all devices on the selected replication leg. Just as the Freeze function applies to the entire consistency group, HyperSwap is for the entire consistency group. For example, if a single mirrored volume fails and HyperSwap is started, processing is swapped to one of the sets of secondary devices for *all* primary volumes in the configuration, including those primary volumes in other, unaffected, disk subsystems. This process ensures that all primary volumes remain in the same site. If HyperSwap swaps only the failed LSS, you would then have several primaries in one location, and the remainder in another location. This situation would make for a complex environment to operate and administer I/O configurations.

Incremental Resynchronization

For dual-leg configurations, when a disk switch or recovery on one of the secondaries occurs, Multi-Target Metro Mirror (MTMM) provides for a capability that is known as "Incremental Resynchronization" (IR). Assume that your RS1 disks are the current primaries and the RS2 and RS3 disks are the current secondaries. If you switch from using RS1 to using RS2 as your primary disks, to maintain a multi-target configuration, you must establish replication on RL1, between RS2 and RS1, and on RL3, between RS2 and RS3. Use a feature of the Metro Mirror copy technology that is known as Failover/Failback, together with the MTMM IR capability to establish replication for RL1 and RL3 without having to copy all data from RS2 to RS1 or from RS2 to RS3. Only the changes that occur on B after the switch to B are copied to resynchronize the two legs.

If there is an unplanned HyperSwap from RS1 to RS2, because RS1 failed, replication can be established on RL3 between RS2 and RS3 to restore DR readiness. Again, this is an IR (only changed tracks are copied), so the duration to get to a protected position is faster compared to performing an initial copy for the leg.

HyperSwap with less than full channel bandwidth

You might consider enabling unplanned HyperSwap on the cross-site replication leg (RL2), even if you do not have sufficient cross-site channel bandwidth to sustain the full production workload for normal operations. Assuming that a disk failure is likely to cause an outage and that you must switch to using the RS3 disk in the other site (because the RS2 disks in the same site are down at the time), the unplanned HyperSwap to RS3 might at least present you with the opportunity to perform an orderly shutdown of your systems first. Shutting down your systems cleanly avoids the complications and restart time elongation that is associated with a crash-restart of application subsystems.

Preferred Swap Leg and HyperSwap (Primary Failure) policy options

Clients might prefer not to immediately enable their environment for unplanned HyperSwap when they first implement GDPS. For this reason, unplanned HyperSwap is not enabled by default. However, we strongly suggest that all GDPS Metro clients enable their environment for unplanned HyperSwap, at a minimum, on the local replication leg (RL1) if it is configured. Both copies of disk on the RL1 leg (RS1 and RS2) are local, so distance and connectivity should not be an issue.

You control the actions that GDPS takes for primary disk problems by specifying a Primary Failure policy option. This option is applicable to both replication legs. However, you have the option of overriding this specification at a leg level and request a different action based on which leg is selected by GDPS to act upon. Furthermore, there is the Preferred Swap Leg policy, which is factored in when GDPS decides which leg to act upon as a result of a primary disk problem trigger.

Preferred Swap Leg selection for unplanned HyperSwap

In a dual-leg configuration, a primary disk problem trigger is common to both replication legs because the primary disk is common to both legs. Before acting on the trigger, GDPS first needs to select which leg to act upon. GDPS provides you with the ability to *influence* this decision by specifying a Preferred Swap Leg policy. GDPS attempts to select the leg that you identified as the Preferred Swap Leg first. However, if this leg is not eligible for the action that you specified in your Primary Failure policy, GDPS attempts to select the other active replication leg. These reasons are among the ones that your Preferred Swap Leg might not be eligible for selection:

► It is the MTIR leg.
► All pairs for the leg are not in a duplex state.
► It is not HyperSwap enabled.

HyperSwap retry on non-preferred leg

If the preferred leg is viable and selected for an unplanned swap, there is still a possibility (albeit small) that the swap on this leg fails for some reason. When the swap on the first leg fails, if the other replication leg is enabled for HyperSwap, GDPS retries the swap on the other leg. This action maximizes the chances of a successful swap.

Primary failure policy options

After GDPS selects which leg that it will act on when a primary disk problem trigger occurs, the first thing it will do will be a Freeze on the selected leg (the same as is performed when a mirroring problem trigger is encountered). GDPS then applies the Primary Failure policy option that is specified for that leg. The Primary Failure policy for each leg can specify a different action. You can specify the following Primary Failure policy options:

► `PRIMARYFAILURE=GO`

No swap is performed. The action GDPS takes is the same as for a freeze event with policy option `PPRCFAILURE=GO`. A Run action is performed, which enables systems to continue by using the original primary disks. Metro Mirror is suspended and updates are not being replicated to the secondary. depending on the scope of the primary disk problem, it might be that some or all production workloads cannot run or cannot sustain required service levels. Such a situation might necessitate restarting the systems on the secondary disks. Because of the freeze, the secondary disks are in a consistent state and can be used for restart. However, any transactions that ran after the Go action are lost.

► `PRIMARYFAILURE=STOP`

No swap is performed. The action GDPS takes is the same as for a freeze event with policy option `PPRCFAILURE=STOP`. GDPS system-resets all the production systems. This action ensures that no further I/O occurs. After performing situation analysis, if it is determined that this issue was not transient and that the secondaries should be used to IPL the systems again, no data will be lost.

► `PRIMARYFAILURE=SWAP,`*swap_disabled_action*

The first parameter, SWAP, indicates that after performing the Freeze, GDPS will proceed with performing an unplanned HyperSwap. When the swap is complete, the systems are running on the new, swapped-to primary disks (former secondaries). Mirroring on the selected leg is in a suspended state because the primary disks are known to be in a problematic state, there is no attempt to reverse mirroring. After the problem with the primary disks is fixed, you can instruct GDPS to resynchronize Metro Mirror from the current primaries to the former ones (which are now considered to be secondaries).

The second part of this policy, `swap_disabled_action`, indicates what GDPS should do if HyperSwap was temporarily disabled by an operator action at the time the trigger was encountered. Effectively, an operator action instructed GDPS not to perform a HyperSwap, even if there is a swap trigger. GDPS already performed a freeze. The second part of the policy control what action GDPS takes next.

The following options (which are in effect only if HyperSwap is disabled by the operator) are available for the second parameter (the disk is already frozen):

GO   This is the same action as GDPS would have performed if the policy option had been specified as `PRIMARYFAILURE=GO`.

STOP   This is the same action as GDPS would have performed if the policy option had been specified as `PRIMARYFAILURE=STOP`.

Preferred Swap Leg and Primary Failure policy selection considerations

For the Preferred Swap Leg policy, consider whether you can tolerate running with disk and systems in opposite sites with no or minimal performance impact. If that is acceptable, you can choose either leg, although it might be better to prefer the RL2 (Site1-Site2) leg. If you cannot tolerate running with disks and systems in opposite sites, choose the RL1, local leg.

For the Primary Failure policy, as a best practice, specify SWAP for the first part of the policy option to enable HyperSwap, at least on the local replication leg (RL1). If distance and connectivity between your sites is not an issue, consider specifying SWAP for the first part of the policy on the remote replication leg (RL2) also.

For the Stop or Go choice, either as the second part of the policy option or if you are not using SWAP, similar considerations apply as for the PPRCFAILURE policy options to Stop or Go. Go carries the risk of data loss if it is necessary to abandon the primary disk and restart systems on the secondary. Stop carries the risk of taking an unnecessary outage if the problem was transient. The key difference is that with a mirroring failure, the primary disks are not broken. When you enable the systems to continue to run on the primary disk with the Go option, other than a disaster (which is a low probability), the systems are likely to run with no problems. With a primary disk problem, with the Go option, you are allowing the systems to continue running on what are known to be disks that experienced a problem just seconds ago. If this situation was a serious problem with widespread impact, such as an entire disk subsystem failure, the applications experience severe problems. Some transactions might continue to commit data to those disks that are not broken. Other transactions might be failing or experiencing serious service time issues. Also, if there is a decision to restart systems on the secondary because the primary disks cannot support the workloads, there will be data loss. The probability that a primary disk problem is a real problem that necessitates a restart on the secondary disks is higher when compared to a mirroring problem. A Go specification in the Primary Failure policy increases your risk of data loss.

If the primary failure was of a transient nature, a Stop specification results in an unnecessary outage. However, with primary disk problems, the probability that the problem might necessitate restart on the secondary disks is high, so a Stop specification in the Primary Failure policy avoids data loss and facilitates faster restart.

The considerations relating to CF structures with a PRIMARYFAILURE event are similar to a PPRCFAILURE event. If there is an actual swap, the systems continue to run and continue to use the same structures as they did before the swap. The swap is transparent. With a Go action, because you continue to update the CF structures along with the primary disks after the Go, if you need to abandon the primary disks and restart on the secondary, the structures are inconsistent with the secondary disks and are not usable for restart purposes. This action prolongs the restart and your recovery time. With Stop, if you decide to restart the systems by using the secondary disks, there is no consistency issue with the CF structures because no further updates occurred on either set of disks after the trigger was captured.

GDPS use of DS8000 functions

GDPS strives to use (when it makes sense) enhancements to the IBM DS8000 disk technologies. In this section, we provide information about the key DS8000 technologies that GDPS supports and uses.

PPRC Failover/Failback support

When a primary disk failure occurs and the disks are switched to the secondary devices, PPRC Failover/Failback (FO/FB) support eliminates the need to do a full copy when reestablishing replication in the opposite direction. Because the primary and secondary volumes are often in the same state when the freeze occurred, the only differences between the volumes are the updates that occur to the secondary devices after the switch.

Failover processing sets the secondary devices to primary suspended status and starts change recording for any subsequent changes made. When the mirror is reestablished with failback processing, the original primary devices become secondary devices and a resynchronization of changed tracks takes place.

GDPS Metro requires PPRC FO/FB capability to be available on all disk subsystems in the managed configuration.

PPRC eXtended Distance (PPRC-XD)

PPRC-XD (also known as Global Copy) is an asynchronous form of the PPRC copy technology. GDPS uses PPRC-XD rather than Metro Mirror (which is the synchronous form of PPRC) to reduce the performance impact of certain remote copy operations that potentially involve a large amount of data. For more information, see 3.7.2, "Reduced impact on initial copy and resynchronization" on page 102.

Storage Controller Health Message Alert

This feature facilitates triggering an unplanned HyperSwap proactively when the disk subsystem reports an acute problem that requires extended recovery time.

PPRC Summary Event Messages

GDPS supports the DS8000 PPRC Summary Event Messages (PPRCSUM) function, which is aimed at reducing the message traffic and the processing of these messages for Freeze events. This function is described in "GDPS Freeze function for mirroring failures" on page 47.

Soft Fence

Soft Fence provides the capability to block access to selected devices. As discussed in "Protecting secondary disks from accidental update" on page 58, GDPS uses Soft Fence to avoid write activity on disks that are exposed to accidental update in certain scenarios.

On-demand dump (also known as non-disruptive statesave)

When problems occur with disk subsystems such as those which result in an unplanned HyperSwap, a mirroring suspension or performance issues, a lack of diagnostic data from the time the event occurs can result in difficulties in identifying the root cause of the problem. Taking a full statesave can lead to temporary disruption to host I/O and is often frowned upon by clients for this reason. The on-demand dump (ODD) capability of the disk subsystem facilitates taking a non-disruptive statesave (NDSS) at the time that such an event occurs. The microcode does this task automatically for certain events such as a dump of the primary disk subsystem that triggers a Metro Mirror freeze event and also allows an NDSS to be requested by an exploiter. This capability enables first failure data capture (FFDC) and thus ensures that diagnostic data is available to aid problem determination. Not all information that is contained in a full statesave is contained in an NDSS and therefore there may still be failure situations where a full statesave is requested by the support organization.

GDPS supports taking an NDSS by using the remote copy panels. In addition to this support, GDPS autonomically takes an NDSS if there is an unplanned Freeze or HyperSwap event.

Query Host Access function

When a Metro Mirror disk pair is being established, the device that is the target (secondary) must not be in use by any system. The same is true when establishing a FlashCopy relationship to a target device. If the target is in use, the establishment of the Metro Mirror or FlashCopy relationship fails. When such failures occur, it can be a tedious task to identify which system is holding up the operation.

The Query Host Access disk function provides the means to query and identify what system is using a selected device. GDPS uses this capability and adds usability in several ways:

► Query Host Access identifies the LPAR that is using the selected device through the Central Processor Complex (CPC) serial number and LPAR number. It is still a tedious job for operations staff to convert this information to a system or CPC and LPAR name. GDPS does this conversion and presents the operator with more readily usable information, thus avoiding this additional conversion effort.

► Whenever GDPS is requested to perform a Metro Mirror or FlashCopy establish operation, GDPS first performs Query Host Access to see whether the operation is expected to succeed or fail as a result of one or more target devices being in use. GDPS alerts the operator if the operation is expected to fail, and identifies the target devices in use and the LPARs holding them.

► GDPS continually monitors the target devices that are defined in the GDPS configuration and alerts operations to the fact that target devices are in use when they should not be. This alert enables operations to fix the reported problems in a timely manner.

► GDPS provides the ability for the operator to perform ad hoc Query Host Access to any selected device by using the GDPS panels.

Protecting secondary disks from accidental update

A system cannot be started by using a disk that is physically a Metro Mirror secondary disk because Metro Mirror secondary disks cannot be brought online to any systems. However, a disk can be secondary from a GDPS (and application use) perspective but physically, from a Metro Mirror perspective, have simplex or primary status.

For both planned and unplanned HyperSwap, and a disk recovery, GDPS changes former secondary disks to primary or simplex state. However, these actions do not modify the state of the former primary devices, which remain in the primary state.

Therefore, the former primary devices remain accessible and usable even though they are considered to be the secondary disks from a GDPS perspective. This configuration makes it is possible to accidentally update or IPL from the wrong set of disks. Accidentally using the wrong set of disks can potentially result in a loss of data integrity or data.

GDPS Metro provides protection against using the wrong set of disks in different ways:

► If you attempt to load a system through GDPS (either script or panel) by using the wrong set of disks, GDPS rejects the load operation.

► If you used the HMC rather than GDPS facilities for the load, then early in the IPL process, during initialization of GDPS, if GDPS detects that the system coming up was started by using the wrong set of disks, GDPS quiesces that system, preventing any data integrity problems that might be experienced had the applications been started.

► GDPS uses a DS8000 disk subsystem capability, which is called Soft Fence for configurations where the disks support this function. Soft Fence provides the means to *fence* (that is, block) access to a selected device. GDPS uses Soft Fence when appropriate to fence devices that would otherwise be exposed to accidental update.

3.1.2 Protecting tape data

Although most of your critical data is on disk, it is possible that other data you require following a disaster is on tape. Just as you mirror your disk-resident data to protect it, equally you can mirror your tape-resident data. GDPS Metro supports management of the IBM TS7700. GDPS provides TS7700 configuration management and displays the status of the managed TS7700s on GDPS panels.

TS7700 libraries that are managed by GDPS are monitored and alerts are generated for non-normal conditions. The capability to control TS7700 replication from GDPS scripts and panels by using TAPE ENABLE and TAPE DISABLE by library, grid, or site is provided for managing TS7700 during planned and unplanned outage scenarios.

Another important aspect of replicated tape is the identification of "in-doubt" tapes. Tape replication is not exactly like disk replication in that the replication is not done every time that a record is written to the tape. The replication is typically performed at tape unload rewind time or even later. This means that if there is an unplanned event or interruption to the replication, some volumes might be one level behind in one or more libraries in the grid. If you must perform a recovery operation in one site because the other site failed, it is important to identify if any of the tapes in the library in the site where you are recovering are a level earlier. Depending on the situation with any in-doubt tapes in the library or libraries you use in the recovery site, you might need to perform special recovery actions. For example, you might need to rerun one or more batch jobs before resuming batch operations.

GDPS supports identifying in-doubt tapes in a TS7700 library. The TS7700 provides a capability that is called Bulk Volume Information Retrieval (BVIR). By using this BVIR capability, GDPS automatically collects information about all volumes in all libraries in the grid where the replication problem occurred if there is an unplanned interruption to tape replication. GDPS can then use this information to report on in-doubt volumes in any library in that grid if the user requests a report. In addition to this automatic collection of in-doubt tape information, it is possible to request GDPS to perform BVIR processing for a selected library by using the GDPS panel interface at any time.

The IBM TS7700 provides comprehensive support for replication of tape data. For more information about the TS7700 technology that complements GDPS for tape data, see *IBM TS7700 Release 5.3 Guide*, SG24-8464.

3.1.3 Protecting distributed (Fixed-Block) data

> **Terminology:** The following definitions describe the terminology that we use in this book when referring to the various types of disks:
>
> ► IBM Z or Count-Key-Data (CKD) disks
>
>   GDPS can manage disks that are formatted as CKD disks (the traditional mainframe format) that are used by any of the following IBM Z operating systems: z/VM, VSE, KVM, and Linux on IBM Z.
>
>   We refer to the disks that are used by a system running on the mainframe as *IBM Z disks*, *CKD disks*, or *CKD devices*. These terms are used interchangeably.
>
> ► Fixed-Block (FB) disks
>
>   Disks that are used by systems other than ones that are running on IBM Z are traditionally formatted as FB and are referred to as *FB disks* or *FB devices* in this book.

GDPS Metro can manage the mirroring of FB devices that are used by non-mainframe operating systems. The FB devices can be part of the same consistency group as the mainframe CKD devices, or they can be managed separately in their own consistency group.

For more information about FB disk management, see 3.3.1, "Fixed-Block disk management" on page 69.

3.1.4 Protecting other CKD data

Systems that are fully managed by GDPS are known as *GDPS managed systems* or *GDPS systems*. The following types of GDPS systems are available:

► z/OS systems that are in the GDPS sysplex
► z/VM systems that are managed by GDPS Metro MultiPlatform Resiliency for IBM Z (xDR)
► KVM systems that are managed by GDPS Metro MultiPlatform Resiliency for IBM Z (xDR)
► IBM Db2 Analytics Accelerator on IBM Z running in Secure Service Container (SSC) LPARs that are managed by GDPS Metro MultiPlatform Resiliency for IBM Z (xDR)
► z/OS systems that are outside of the GDPS sysplex that are managed by the GDPS Metro z/OS Proxy (z/OS Proxy)

GDPS Metro can also manage the disk mirroring of CKD disks that are used by systems outside of the sysplex: other z/OS systems, Linux on IBM Z, virtual machine (VM), VSE, and KVM systems that are not running any GDPS Metro or xDR automation. These systems are known as "foreign systems."

Because GDPS manages Metro Mirror for the disks that are used by these systems, their disks are attached to the GDPS controlling systems. With this setup, GDPS is able to capture mirroring problems and perform a freeze. All GDPS managed disks belonging to the GDPS systems and these foreign systems are frozen together, regardless of whether the mirroring problem is encountered on the GDPS systems' disks or the foreign systems' disks.

GDPS Metro is not able to directly communicate with these foreign systems. For this reason, GDPS automation will not be aware of certain other conditions such as a primary disk problem that is detected by these systems. Because GDPS will not be aware of such conditions that would have otherwise driven autonomic actions such as HyperSwap, GDPS does not react to these events.

If an unplanned HyperSwap occurs (because it was triggered on a GDPS managed system), the foreign systems cannot and will not swap to using the secondaries. Mechanisms are provided to prevent these systems from continuing to use the former primary devices after the GDPS systems are swapped. You can then use GDPS automation facilities to reset these systems and restart them using the swapped-to primary disks.

3.2 GDPS Metro configurations

At its most basic, a GDPS Metro configuration consists of at least one production system, at least one controlling system in a sysplex, primary disks, and at least one set of secondary disks. The actual configuration depends on your business and availability requirements.

One aspect of availability requirements has to do with the availability of the servers, systems, and application instances. The following configurations that address this aspect of availability are most common:

► Single-site workload configuration

   In this configuration, all production systems normally run in the same site, referred to as Site1, and the GDPS controlling system runs in Site2. In effect, Site1 is the active site for all production systems. The controlling system in Site2 is running and resources are available to move production to Site2, if necessary, for a planned or unplanned outage of Site1. Although you might also hear this referred to as an *Active/Standby* GDPS Metro configuration, we avoid the Active/Standby term to avoid confusion with the same term used in conjunction with the GDPS Continuous Availability product.

► Multisite workload configuration

   In this configuration, the production systems run in both sites, Site1 and Site2. This configuration typically uses the full benefits of data sharing available with a Parallel Sysplex. Having two GDPS controlling systems, one in each site, is preferable. Although you might also hear this configuration referred to as an *Active/Active* GDPS Metro configuration, we avoid the Active/Active term to avoid confusion with the same term used with the GDPS Continuous Availability product.

► Business Recovery Services (BRS) configuration

   In this configuration, the production systems and the controlling system are all in the same site, referred to as Site1. Site2 can be a client site or can be owned by a third-party recovery services provider (thus the name BRS). You might hear this configuration referred to as an *Active/Cold* configuration.

Another aspect of availability requirements has to do with the availability of data. The most basic configuration of GDPS Metro consists of to copies of data, a set of primary disks and one set of secondary disks. This configuration is known as a *single-leg* configuration.

GDPS Metro also leverages the IBM MTMM disk mirroring technology to maintain two synchronous secondary copies of your data. This configuration, which is known as a *dual-leg* configuration, provides an extra level of availability because data resiliency can be maintained, even when one copy of data is lost.

These configuration options are described later in this section.

3.2.1 Controlling systems

Why does a GDPS Metro configuration need a controlling system? At first, you might think that the controlling system is extra infrastructure overhead. However, when you have an unplanned outage that affects production systems or the disk subsystems, it is crucial to have a system such as the controlling system that can survive failures that might impact other portions of your infrastructure. Use the controlling system to perform situation analysis after the unplanned event to determine the status of the production systems or the disks, and then to drive automated recovery actions. The controlling system plays a vital role in a GDPS Metro configuration.

The controlling system must be in the same sysplex as the production system (or systems) so it can see all the messages from those systems and communicate with those systems. However, it shares an absolute minimum number of resources with the production systems (typically just the CDS). By being configured to be as self-contained as possible, the controlling system is unaffected by errors that can stop the production systems (for example, an ELB event on a primary volume).

The controlling system must have connectivity to all the Site1 and Site2 primary and secondary devices that it manages. If available, it is preferable to isolate the controlling system infrastructure on a disk subsystem that is not housing mirrored disks that are managed by GDPS.

The controlling system is responsible for carrying out all recovery actions following a disaster or potential disaster, for managing the disk mirroring configuration, for initiating a HyperSwap, for initiating a freeze and implementing the freeze/swap policy actions, for reassigning Server Time Protocol (STP) roles, and for restarting failed system.

> **Note:** The availability of the dedicated GDPS controlling system (or systems) in *all* configurations is a fundamental requirement of GDPS. It is not possible to merge the function of the controlling system with any other system that accesses or uses the primary volumes or other production resources.

Configuring GDPS Metro with two controlling systems, one in each site is a best practice because a controlling system is designed to survive a failure in the opposite site of where the primary disks are. Primary disks are normally in Site1 and the controlling system in Site2 is designed to survive if Site1 or the disks in Site1 fail. However, if you reverse the configuration so that the primary disks are now in Site2, the controlling system is in the same site as the primary disks. It cannot survive a failure in Site2 and might not survive a failure of the disks in Site2 depending on the configuration. Configuring a controlling system in both sites ensures as much protection, no matter which site is the primary disk site. When two controlling systems are available, GDPS manages to assign a Master role to the controlling system that is in the same site as the secondary disks and switching the Master role if there is a disk switch.

Improved controlling system availability: Enhanced timer support

Normally, a loss of synchronization with the sysplex timing source generates a disabled console write to operator with reply (WTOR) message that suspends all processing on the LPAR until a response is made to the WTOR. The WTOR message is IEA394A in STP timing mode.

In a GDPS environment, z/OS is aware that a system is a GDPS controlling system and allows a GDPS controlling system to continue processing even when the server it is running on loses its time source and becomes unsynchronized. Therefore, the controlling system is able to complete any freeze or HyperSwap processing it might have started and is available for situation analysis and other recovery actions, instead of being in a disabled WTOR state.

In addition, because the controlling system is operational, it can be used to help in problem determination and situation analysis during the outage, thus further reducing the recovery time that is needed to restart applications.

The controlling system is required to perform GDPS automation in a failure. Actions might include these tasks:

► Reassigning STP roles
► Performing the freeze processing to ensure secondary data consistency
► Coordinating HyperSwap processing
► Running a takeover script
► Aiding with situation analysis

Because the controlling system needs to run with only a degree of time synchronization that allows it to correctly participate in heartbeat processing regarding the other systems in the sysplex, this system should be able to run unsynchronized for 80 minutes by using the local time-of-day (TOD) clock of the server (referred to as *local timing mode*), instead of generating a WTOR.

Automated response to STP sync WTORs

GDPS on the controlling systems, by using the BCP Internal Interface (BCPii), provides automation to reply to WTOR IEA394A when the controlling systems are running in local timing mode. See "Improved controlling system availability: Enhanced timer support" on page 62. A server in an STP network might have recovered from an unsynchronized to a synchronized timing state without client intervention. By automating the response to the WTOR messages, potential time outs of subsystems and applications in the client's enterprise might be averted, thus potentially preventing a production outage.

If WTOR IEA394A is posted for production systems, GDPS uses the BCPii to automatically reply RETRY to the WTOR. If z/OS determines that the CPC is in a synchronized state, either because STP recovered or the Coordinated Timing Network (CTN) was reconfigured, it no longer spins and continues processing. If the CPC is still in an unsynchronized state when GDPS automation responded with RETRY to the WTOR, however, the WTOR is reposted.

The automated reply for any system is retried for 60 minutes. After 60 minutes, you will need to manually respond to the WTOR.

3.2.2 Single-site workload configuration

A GDPS Metro single-site workload environment typically consists of a multisite sysplex, with all production systems running in a single site, normally Site1, and the GDPS controlling system in Site2. The controlling system (or systems because you may have two in some configurations) normally runs in the site containing the secondary disk volumes.

The multisite sysplex can be a base sysplex or a Parallel Sysplex; a CF is not strictly required. The multisite sysplex must be configured with redundant hardware (for example, a CF and a Sysplex Timer in each site), and the cross-site connections must also be redundant. Instead of using Sysplex Timers to synchronize the servers, you can also use STP to synchronize the servers.

Figure 3-1 shows a typical GDPS Metro single-site workload configuration. LPARs P1 and P2 are in the production sysplex, as are the CFs CF1, CF2, and CF3. The primary (RS1) disks are in Site1, with a set of secondaries (RS2) also in Site1 and another set of secondaries (RS3) in Site2. All the production systems are running in Site1, with only the GDPS controlling system (K1) running in Site2. You notice that system K1's disks (those marked K1) are also in Site2 and are not mirrored.

The GDPS Metro code itself runs under NetView and System Automation, and runs in every system in the GDPS sysplex.
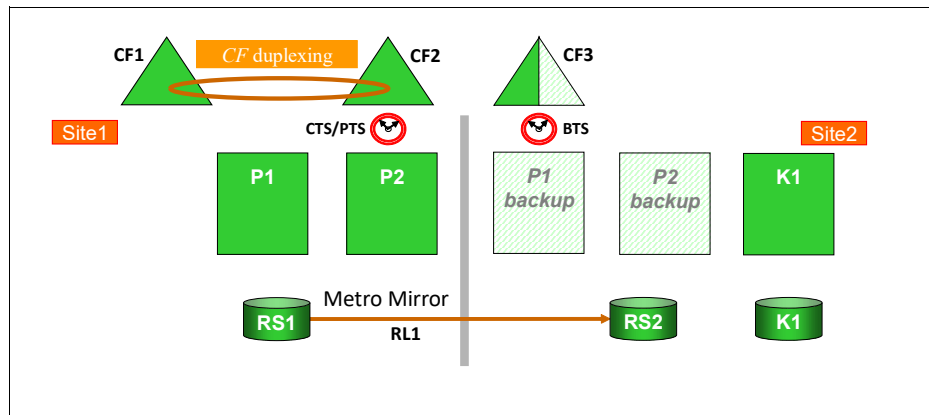
Figure 3-1   GDPS Metro single site workload configuration

3.2.3 Multisite workload configuration

A multisite workload configuration, which is shown in Figure 3-2, differs from a single-site workload in that production systems are running in *both* sites. Although running a multisite workload as a base sysplex is possible, seeing this configuration as a base sysplex (that is, without CFs) is unusual because a multisite workload is usually a result of higher availability requirements, and Parallel Sysplex and data sharing are core components of such an environment.
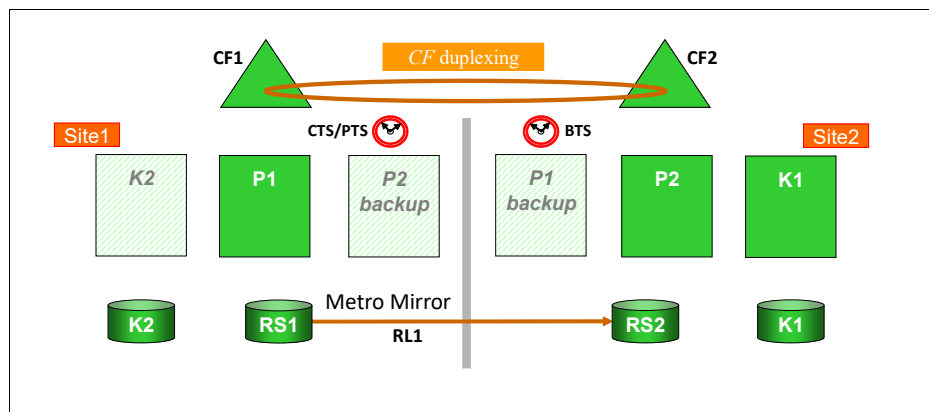
Figure 3-2   GDPS Metro multisite workload configuration

Because in this example we have production systems in both sites, we need to provide the capability to recover from a failure in either site. So, in this case, there is also a GDPS controlling system with its own local (not mirrored) disk running in Site1, namely System K2. Therefore, if there is a disaster that disables Site2, there will still be a GDPS controlling system available to decide how to react to that failure and what recovery actions are to be taken.

3.2.4 Business Recovery Services configuration

A third configuration is known as the *BRS configuration*, and is shown in Figure 3-3 on page 66. In this configuration, all the systems in the GDPS configuration, including the controlling systems, are in a sysplex in the same site, namely Site1. The sysplex does not span the two sites. The second site, Site2, might be a client site or might be owned by a third-party recovery services provider; thus the name BRS.

Site2 contains the secondary disks and the alternative CDS, and might also contain processors that are available in a disaster, but are not part of the configuration. This configuration can also be used when the distance between the two sites exceeds the distance that is supported for a multisite sysplex, but is within the maximum distance that is supported by FICON and Metro Mirror.

Although there is no need for a multisite sysplex with this configuration, you must have channel connectivity from the GDPS systems to the secondary disk subsystems. Also, as explained in the next paragraph, the controlling system in Site1 needs channel connectivity to its disk devices in Site2. Therefore, FICON link connectivity from Site1 to Site2 is required.

For more information about options that are available to extend the distance of FICON links between sites, see 2.9.7, "Connectivity options" on page 40, and *IBM Z Connectivity Handbook*, SG24-5444.

In the BRS configuration one of the two controlling systems must have its disk devices in Site2. This configuration permits that system to be restarted manually in Site2 after a disaster is declared. After it restarts in Site2, the system runs a GDPS script to recover the secondary disk subsystems, reconfigure the recovery site, and restart the production systems from the disk subsystems in Site2.

If you have only a single controlling system and you have a total cross-site fiber connectivity failure, the controlling system running on Site2 disks might not be able to complete the Freeze operation because it loses access to its disk in Site2. Having a second controlling system running on Site1 local disks ensures that the freeze operation completes successfully if the controlling system running on Site2 disks is down or cannot function because of a cross-site fiber loss.

GDPS attempts to maintain the primary system in the controlling system by using the secondary disks (see Figure 3-3).
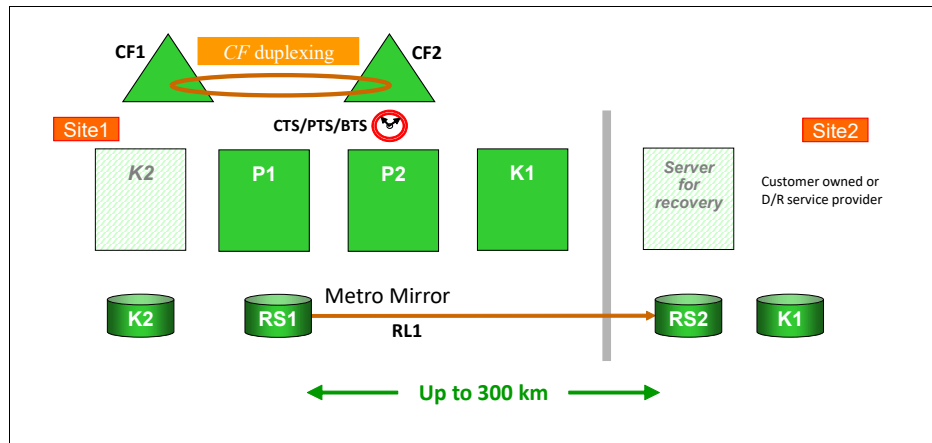
Figure 3-3   GDPS Metro BRS configuration

3.2.5 Single-leg configuration

The previous sections showed GDPS Metro in single-leg configurations. That is, the configurations consisted of only two copies of the production data: a primary copy and a secondary copy. The primary and secondary copies of data are called *disk locations* or *replication sites* (RSs). The copy in Site1 is known as *RS1* and the copy in Site2 is known as *RS2*.

The replication connection between the RSs is called a *replication leg* or a *leg*. A replication leg has a fixed name that is based on the two disk locations that it connects. In a single-leg configuration, there is only one replication leg, called RL1, and it connects the two disk locations RS1 and RS2.

3.2.6 Dual-leg configuration

In addition to providing the single-leg configuration that was described in the previous section, GDPS Metro also uses the IBM MTMM disk mirroring technology to provide a dual-leg configuration that maintains two synchronous secondary copies of your data to provide an extra level of data resiliency.

With a dual-leg configuration, an RS is added and two replication legs are also added. The three disk locations, or copies, are known as RS1, RS2, and RS3. RS1 and RS2 are assumed to be "local" and are fixed in Site1, and RS3 is fixed in Site2.

Although any of the three RSs can assume the primary disk role, in a typical configuration:

► The primary disk is in Site1, that is, either RS1 or RS2.
► The other disk copy in Site1 provides HA protection.
► The copy in Site2 (RS3) provides DR protection.

The replication legs in a dual-leg configuration have fixed names that again, are based on the two disk locations that they connect:

► The RS1-RS2 (or RS2-RS1) leg is RL1.
► The RS1-RS3 (or RS3-RS1) leg is RL2.
► The RS2-RS3 (or RS3-RS2) leg is RL3.

The name of a replication leg never changes, even if the replication direction is reversed for that leg. However, the *role* of a leg can change, depending on the primary disk location. The two legs from the current primary to each of the two secondaries serve as the *active* legs whereas the leg between the two secondary locations serves as the *incremental resync* or *MTIR* leg.

To illustrate this concept, consider the sample dual-leg configuration that is shown in Figure 3-4.
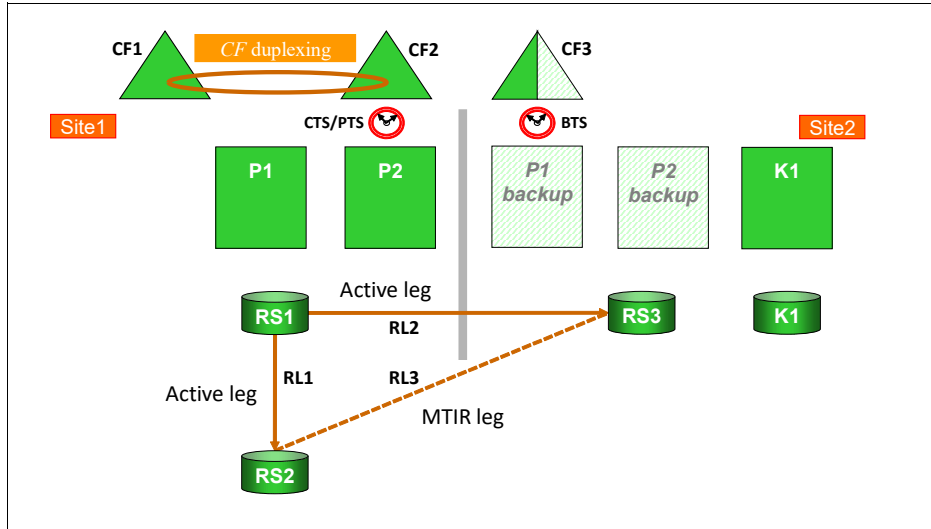
Figure 3-4   Typical GDPS Metro dual-leg configuration

In this sample configuration, RS1 is the primary disk location, RL1 and RL2 are the active replication legs, and RL3 is the MTIR leg.

If a disk switch exists and RS2 becomes the new primary disk, RL1 and RL3 become the active replication legs and RL2 then becomes the MTIR leg, as shown in Figure 3-5.
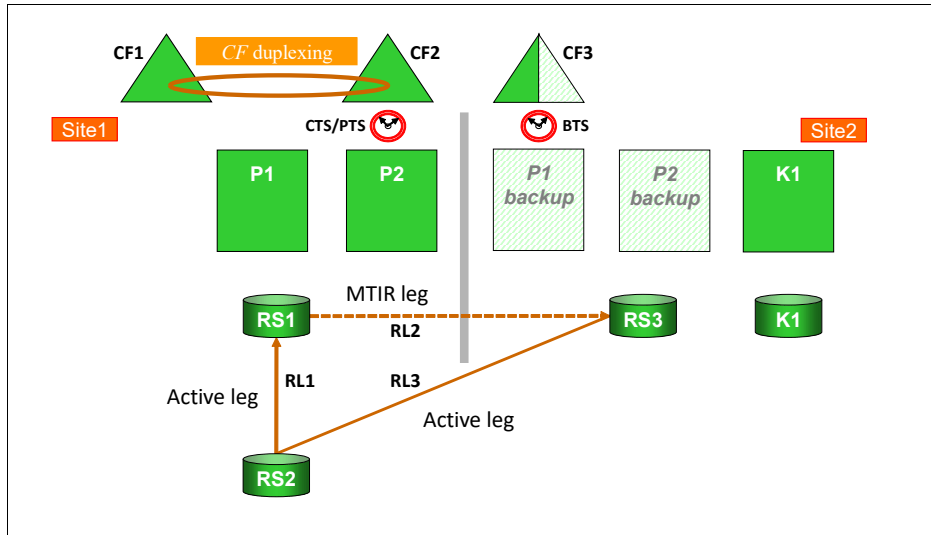
Figure 3-5   GDPS Metro dual-leg configuration after a switch to RS2

3.2.7 Combining GDPS Metro with GDPS GM

When configured in a single-leg topology, GDPS Metro can be combined with GDPS Global - GM (GDPS GM) in 3-site and 4-site configurations. In such configurations, GDPS Metro (when combined with Parallel Sysplex use and HyperSwap) in one region provides CA across a metropolitan area or within the same local site, and GDPS GM provides DR capability using a remote site in a different region.

The 4-site environment is configured in a symmetric manner so that there is a GDPS Metro-managed replication leg available in both regions to provide CA within the region, with GDPS GM providing cross-region DR, no matter in which region production is running at any time.

Combining GDPS Metro and GDPS GM in this fashion is referred to as GDPS Metro Global - GM (GDPS MGM). For more information about GDPS MGM configurations, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237.

3.2.8 GDPS Metro in a single site

The final configuration is where you want to benefit from the capabilities of GDPS Metro to extend the CA attributes of a Parallel Sysplex to planned and unplanned disk reconfigurations, but you do not have the facilities to mirror disk across two sites. In this case, you can implement GDPS Metro HyperSwap Manager (GDPS HM).

GDPS HM is similar to the full-function GDPS Metro offering, except that it does not include the scripts for management of the LPARs and workloads. GDPS HM is upgradeable to a full GDPS Metro implementation. For more information about GDPS HM, see Chapter 4, "IBM GDPS Metro HyperSwap Manager" on page 111.

Because configuring GDPS HM (or GDPS Metro) within a single site does not provide protection against site failure events, such a configuration is likely to be used within the context of a GDPS MGM multi-site solution rather than a stand-alone solution.

Another possibility is that this configuration is for a client environment that has aggressive RTOs for failures other than a disaster event and some mechanism such as tape vaulting is used for disaster protection. This means that long recovery times and a fair amount of data loss can be tolerated during a disaster.

3.2.9 Other considerations

The availability of the dedicated GDPS controlling system (or systems) in *all* scenarios is a fundamental requirement in GDPS. Merging the function of the controlling system with any other system that accesses or uses the primary volumes is not possible.

Equally important is that certain functions (stopping and restarting systems and changing the couple data set configuration) are done through the scripts and panel interface that are provided by GDPS. Because events such as systems going down or changes to the couple data set configuration are indicators of a potential disaster, such changes must be initiated by using GDPS functions so that GDPS understands that they are planned events.

3.3 GDPS Metro management of distributed systems and data

Most enterprises today have a heterogeneous IT environment where the applications and data are on various hardware and software platforms, such as IBM Z, IBM Power, UNIX, Windows, and Linux. Such an environment can benefit greatly if a single point of control manages the data across all the platforms, and for the DR solution to coordinate the recovery across multiple platforms.

In this section, we describe the following functions that are provided by GDPS Metro that are available for clients to manage data and coordinate DR across multiple platforms:

► FB disk management
► Multiplatform Resiliency for IBM Z (also known as xDR)

3.3.1 Fixed-Block disk management

Most enterprises today run applications that update data across multiple platforms. For these enterprises, there is a need to manage and protect not just the data that is on CKD devices, but also the data that is on FB devices for IBM Z and non IBM Z servers. GDPS Metro provides the capability to manage a heterogeneous environment of IBM Z and distributed systems data through a function that is called FB disk management.

The FB Disk Management function allows GDPS to be a single point of control to manage business resiliency across multiple tiers in the infrastructure, which improves cross-platform system management and business processes. GDPS Metro can manage the Metro Mirror remote copy configuration and FlashCopy for distributed systems storage.

Specifically, FB disk support extends the GDPS Metro Freeze capability to FB devices that are in supported disk subsystems to provide data consistency for the IBM Z data and the data on the FB devices.

With FB devices included in your configuration, you can select one of the following options to specify how Freeze processing is to be handled for FB disks and IBM Z (CKD disks), when mirroring or primary disk problems are detected:

► You can select to Freeze all devices managed by GDPS.

  If this option is used, the CKD and FB devices are in a single consistency group. Any Freeze trigger for the IBM Z or FB devices results in the FB and the IBM Z LSSs managed by GDPS being frozen. Use this option to have consistent data across heterogeneous platforms in a disaster to restart systems in the site where secondary disks are located.

  This option is especially suitable when distributed units of work are on IBM Z and distributed servers that update the same data; for example, by using the IBM Distributed Relational Database Architecture (IBM Db2 DRDA).

► You can select to Freeze devices by group.

  If this option is selected, the CKD devices are in a separate consistency group from the FB devices. Also, the FB devices can be separated into Multiple Consistency Groups (MCG); for example, by distributed workloads. The Freeze is performed on only the group for which the Freeze trigger was received. If the Freeze trigger occurs for an IBM Z disk device, only the CKD devices are frozen. If the trigger occurs for an FB disk, only the FB disks within the same group as that disk are frozen.

FB disk management prerequisites

GDPS requires the disk subsystems that contain the FB devices to support the z/OS Fixed-Block Architecture (zFBA) feature. GDPS runs on z/OS and therefore communicates to the disk subsystems directly over a channel connection. The zFBA provides GDPS the ability to send the commands that are necessary to manage Metro Mirror and FlashCopy directly to FB devices over a channel connection. It also enables GDPS to receive notifications for certain error conditions (for example, suspension of an FB device pair). These notifications allow the GDPS controlling system to drive autonomic action such as performing a freeze for a mirroring failure.

> **Note:** HyperSwap for FB disks is not supported for any IBM Z or non IBM Z servers.

3.3.2 Multiplatform Resiliency for IBM Z

GDPS Metro includes a function that is known as Multiplatform Resiliency for IBM Z (also known as xDR). This function extends the near-CA and DR capabilities that are provided by GDPS Metro to other platforms, or operating systems, running on IBM Z servers.

For example, to reduce IT costs and complexity, many enterprises are consolidating open servers into Linux on IBM Z servers. Linux on IBM Z can be implemented as guests that are running under z/VM, as servers that are running natively on IBM Z, or as servers that are running under the KVM Hypervisor on IBM Z. This configuration results in a multitiered architecture in which the application server and the database server are running on different IBM Z platforms. Several examples exist of an application server running on Linux on IBM Z and a database server running on z/OS, including the following examples:

► WebSphere Application Server running on Linux and CICS, Db2 running under z/OS
► SAP application servers running on Linux and database servers running on z/OS

For such multitiered architectures, Multiplatform Resiliency for IBM Z provides a coordinated near-CA and DR solution for the z/OS and the Linux on IBM Z tiers. It can be implemented if the Linux on IBM Z servers run as guests under z/VM, as servers that are running under the KVM Hypervisor on IBM Z, or running natively in an IBM Z LPAR (known as *Linux in LPAR*), and if the disks being used are CKD disks.

> **Note:** For the remainder of this section, Linux on IBM Z is also referred to as *Linux*. The terms are used interchangeably.

Another IBM Z platform that requires coordinated near-CA and DR protection is known as the IBM SSC, or SSC platform. The IBM SSC is a container technology through which you can more quickly and securely deploy firmware and software appliances on IBM Z and IBM LinuxONE servers. An SSC partition (or LPAR) is a specialized container for installing and running specific firmware or software appliances. An appliance is an integration of operating system, middleware, and software components that work autonomously and provide core services and infrastructures that focus on consumability and security.

One such appliance that was deployed by way of an SSC is the IBM Db2 Analytics Accelerator on IBM Z, which is a workload optimized appliance add-on to Db2 for z/OS. It dramatically speeds up queries, and offers a unified homogeneity of service, support, and operations and deeper integration with operational processes. Multiplatform Resiliency for IBM Z provides a coordinated near-CA and DR solution for the IBM Db2 Analytics Accelerator on IBM Z *as a priced feature* of GDPS Metro.

In this section, we describe the following functions that are provided by GDPS Metro that are available for clients to coordinate DR across multiple IBM Z platforms:

► Multiplatform Resiliency for z/VM
► Multiplatform Resiliency for KVM
► Multiplatform Resiliency for Linux in LPAR
► Multiplatform Resiliency for IBM Db2 Analytics Accelerator on IBM Z

Multiplatform Resiliency for z/VM

This section describes the following topics:

► Protecting data integrity across multiple platforms

► GDPS xDR configuration

► Disk and LSS configuration

► System and hardware management

► GDPS controlled shutdown of z/VM

► GDPS Metro xDR support for z/VM Single System Image clustering

► GDPS xDR support for z/VSE guests of z/VM

► GDPS xDR support for Red Hat OpenShift Container Platform

► Installation Verification Program

► xDR Extended Monitoring

Protecting data integrity across multiple platforms

z/VM provides a HyperSwap function. With this capability, the virtual device that is associated with one real disk can be swapped transparently to another disk. GDPS Metro coordinates planned and unplanned HyperSwap for z/OS and z/VM disks, which provide continuous data availability spanning the multitiered application. It does not matter whether the first disk failure is detected for a z/VM disk or a z/OS disk; all are swapped together.

For site failures, GDPS Metro provides a coordinated Freeze for data consistency across z/VM and z/OS. Again, it does not matter whether the first freeze trigger is captured on a z/OS disk or a z/VM disk; all are frozen together.

Note: Most xDR functions, including HyperSwap, benefit non-Linux guests of z/VM also. In fact, having no "production" Linux guests at all is possible. The only requirement for Linux guests is for the xDR proxy nodes, which must be dedicated Linux guests.

However, a z/VM host running z/OS guests is not supported by xDR.

GDPS xDR configuration

In a GDPS xDR-managed z/VM system, you must configure a special Linux guest, which is known as the *proxy* guest. The proxy is a guest that is dedicated to providing communication and coordination with the GDPS Metro controlling system. It must run System Automation for Multiplatforms (SA MP) with the separately licensed xDR feature.

The proxy guest serves as the middleware for GDPS. It communicates commands from GDPS to z/VM, monitors the z/VM environment, and communicates status information and failure information, such as a HyperSwap trigger that is affecting z/VM disk back to the GDPS Metro controlling system. GDPS Metro uses SA MP to pass commands to z/VM and Linux guests.

GDPS xDR supports the definition of two proxy nodes for each z/VM host: one proxy node running on Site1 disk and the other running on Site2 disk. This support extends the two-controlling systems model to the xDR proxy nodes, so it provides a HA proxy design. At any particular time, the proxy node running on disk in the Metro Mirror secondary site is the primary proxy, which is the proxy node that the GDPS primary K-sys coordinates actions with. Similar to the controlling system primary role, the proxy node primary role is switched automatically when Metro Mirror disk is switched (or recovered) or when the primary proxy fails.

It is not mandatory to manage z/OS production systems by using GDPS. The only z/OS systems that are mandatory are the GDPS controlling systems. Originally, xDR supported only one GDPS Controlling system (also referred to as the GDPS primary K-sys).

xDR functions were processed only by the single GDPS primary K-sys. In a planned or unplanned outage of the GDPS primary K-sys, the primary function switched to a production system but xDR processing was interrupted because production systems cannot perform xDR functions.

xDR now supports two GDPS Controlling systems. If your SA MP xDR environment is configured to support two GDPS Controlling systems, xDR processing is protected in a planned or unplanned outage of the Controlling system that is the current primary because the alternative Controlling system takes over the primary responsibility and the alternative Controlling system can perform xDR functions.

Also, if an autonomic primary switch as a result of a disk swap occurs, xDR functions are protected because the alternative primary is a Controlling system and can manage xDR resources.

During cluster initialization, the proxy and non-proxy nodes send their initialization signal to both GDPS Controlling systems. Only the GDPS system that is the current primary responds to the initialization signal, which is how the Linux nodes know which of the Controlling systems is the current primary. Certain events (such as heartbeating and communication of an I/O error) are sent to the current primary, and certain other events (such as initialization) are communicated to both Controlling systems.

In a primary K-sys switch, GDPS informs the Linux nodes of the switch and the Linux nodes then resume relevant communications with the new primary K-sys.

As a best practice, run GDPS with two Controlling systems and enable xDR to support two Controlling systems.

Figure 3-6 on page 73 shows an xDR configuration with two GDPS Controlling systems after a HyperSwap of the primary disks from Site1 to Site2. The primary K-sys was moved to K2-sys in Site1. xDR functions can still be performed by K2-sys, for example, a subsequent disk failure in Site2.
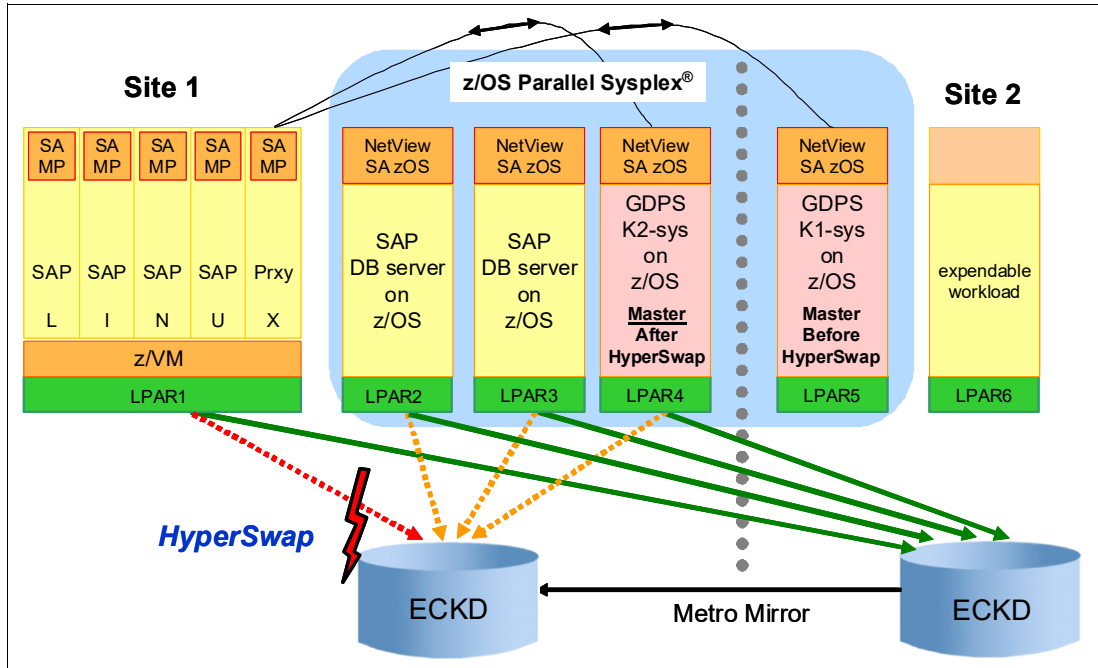
Figure 3-6   xDR configuration with two Controlling systems after a HyperSwap

Also in Figure 3-6, several Linux nodes are running as guests under z/VM. One of the Linux guests is the proxy. The non-proxy SAP Linux guests are shown as also running SA MP, which is not mandatory. If you do run SA MP in the production Linux guest systems, GDPS provides extra capabilities for such guests.

Disk and LSS configuration

The disks being used by z/VM, the guest machines, and the proxy guest in this configuration must be CKD disks. GDPS xDR support allows for the definition of the Metro Mirror secondary devices that are configured to z/VM and the guest machines to be defined in an alternative subchannel set[1]. This approach can simplify definitions and provide high scalability for your disk configuration. For more information, see "Addressing z/OS device limits in a GDPS Metro environment" on page 25.

GDPS xDR also supports the sharing of a logical disk control unit (LSS) by multiple z/VM systems, which facilitates the efficient sharing of resources, provides configuration flexibility, and simplifies the setup that would be required to keep the LSSs separate. It also enables xDR environments to use the z/VM Cross System Extension (CSE) capability.

For example, suppose you have more than one z/VM system and want to perform the following tasks:

► Share the IBM RACF® database across your systems.

► Manage one VM Directory for all the systems.

► Ensure that a minidisk is linked only RW on one guest on one system, and have all the systems enforce that.

► Share the z/VM System Residence volumes.

---

[1] Only alternative subchannel set 1 (MSS1) is supported for defining the Metro Mirror secondary devices.

z/VM CSE can perform all of these tasks for you. Use the z/VM CSE to treat separate VM systems as a Single System Image (SSI), which lowers your system management workload and provides higher availability. For more information about CSE, see *z/VM CP Planning and Administration*, SC24-6083.

If you want to share LSSs and disks, consider the following points:

- ▶ In one LSS, you can place disks into as many xDR-managed z/VM systems as you want.

- ▶ If you want, any z/VM disk that is managed by GDPS can be shared by multiple xDR-managed z/VM systems. This approach requires that you also implement z/VM CSE.

Serialization for disk is supported through the Reserve/Release mechanism for minidisks under z/VM control.

In addition to various z/VMs sharing an LSS, having z/OS and z/VM disks in the same LSS is possible. This configuration allows the LSS capacity to be split between z/OS and z/VM and with the use of hardware reserves, individual disks can be shared by z/VM and z/OS systems.

System and hardware management

System and hardware management capabilities that are similar to the ones that are available for z/OS systems are also available for z/VM systems. GDPS xDR can perform a graceful shutdown of z/VM and its guests and perform hardware actions, such as LOAD and RESET against the z/VM system's partition. GDPS supports taking a stand-alone dump of a z/VM system and in a HyperSwap, it automatically switches the pointers of the z/VM dump volumes to the swapped to site. GDPS can manage Capacity BackUp (CBU) and On/Off Capacity on Demand (OOCoD) for IFLs and CPs on which z/VM systems are running.

GDPS controlled shutdown of z/VM

Graceful shutdown of a z/VM system involves multiple virtual servers. It is a complex process and GDPS has special automation to control this shutdown. The GDPS automated process occurs in multiple phases:

- ▶ During the first phase, all the SA MP clusters with all the nodes for these clusters are stopped. The primary proxy is the only guest running SA MP that is not stopped. When all clusters and nodes running SA MP are successfully stopped, GDPS proceeds to the next phase.

- ▶ During the second phase, all remaining guests that can process the *shutdown signal* are stopped.

- ▶ In phase three, the primary proxy server and z/VM are shut down.

When an xDR-managed z/VM system is shut down by using the GDPS Stop Standard Action (or equivalent script statement), all xDR-managed guests are stopped in parallel. GDPS controls the sequence in which you stop guest systems during a z/VM shutdown.

GDPS Metro xDR support for z/VM Single System Image clustering

z/VM provides a function that is called SSI clustering where up to eight z/VM systems can be clustered to provide more effective resource sharing and other capabilities.

GDPS xDR supports z/VM systems that are members of an SSI cluster. GDPS are aware of the fact that a z/VM system is a member of an SSI. It allows GDPS to perform certain system control actions for these z/VM systems correctly while observing SSI rules.

Linux guests can be transparently moved from one z/VM system in an SSI cluster to another; that is, without requiring the guest to be stopped. This capability, which is called *Live Guest Relocation*, provides CA for Linux guests of z/VM in planned outage situations. If a z/VM system is going to be shut down, for disruptive software maintenance for example, the relocatable Linux guests can first be moved to other z/VM systems in the cluster to avoid an outage to these guests. Similarly, for an entire site shutdown, the guests under all z/VM systems in the site to be shut down can first be moved to z/VM systems in the other site.

GDPS supports performing Live Guest Relocation for xDR-managed z/VM systems. GDPS provides a relocation test capability that tries to assess whether a particular relocation action is likely to be successful. For example, the target z/VM system might not have sufficient resources to host the guest to be moved. Such a test function is useful because it can rectify potential problems before they are encountered. GDPS management for CPs and IFLs that use OOCoD is complementary to this function. You can use GDPS to first increase IFL capacity on the target CPC before performing the actual move.

GDPS xDR support for z/VSE guests of z/VM

GDPS provides specific support for z/VSE guest systems. GDPS monitoring of z/VSE guests requires z/VSE 5.1 with the GDPS Connector (also known as the GDPS Client) enabled for GDPS monitoring. z/VSE guests of xDR-managed z/VM systems can be enabled for special GDPS xDR monitoring and management:

► GDPS can detect the failure of a z/VSE guest and automatically restart it.

► z/VSE guests can be gracefully shut down as part of the graceful shutdown of the hosting z/VM system initiated by GDPS.

GDPS xDR support for Red Hat OpenShift Container Platform

GDPS supports the Red Hat OpenShift Container Platform V4.7.13 application environment like any other z/VM LPAR-based application workload. GDPS xDR supports Red Hat OpenShift Container Platform that is hosted on z/VM guests on a z/VM SSI running on CKD DASD, without persistent storage.

Installation Verification Program

The xDR Installation Verification Program (IVP) verifies that installation and customization activities were done correctly for Linux on IBM Z environments. This helps identify any issues with the customization of the environment where many components exist with specific setup and customization requirements. It also helps identify aspects of the xDR customization that do not adhere to best practices.

IVP is an operator-initiated program that can be used after initial setup, and periodically thereafter, to ensure that changes to the environment have not broken the xDR setup. Two separate programs are provided: One to run on the controlling systems and another to run on the Linux server to ensure that both ends of the implementation are verified.

xDR Extended Monitoring

The GDPS HyperSwap Monitor provides checking for z/OS systems to find out whether the z/OS systems that are managed by GDPS meet the required conditions. Any system that does not meet the required conditions is marked as "not HyperSwap ready." A planned HyperSwap may not run unless all systems are HyperSwap ready. If an unplanned swap is triggered, systems that are not HyperSwap ready are reset and the swap is performed with the participation of only those systems that are HyperSwap ready.

GDPS also performs similar HyperSwap monitoring for xDR systems. Several environmental conditions that are required for HyperSwap for xDR systems are checked. If an xDR system does not meet one or more environmental conditions, GDPS attempts to autonomically fix the detected issue. If it is not possible to autonomically fix the issue, alerts are raised.

Also, any such xDR system that does not meet all environmental conditions that are monitored is marked as "not HyperSwap ready." Raising alerts during monitoring allows an installation to act on the alert and to fix the reported problems in a timely manner to avoid having the system reset if an unplanned swap is triggered.

Multiplatform Resiliency for KVM

In this section, we discuss protecting data integrity across platforms and multiplatform resiliency for KVM configurations.

Protecting data integrity across multiple platforms

KVM does not provide a HyperSwap function. However, GDPS Metro coordinates planned and unplanned HyperSwap for z/OS, Linux under z/VM, Linux under KVM, and Linux in LPAR CKD disks to maintain data integrity and control the shutdown and restart in place of the KVM LPARs. For disk or site failures, GDPS Metro provides a coordinated Freeze for data consistency on CKD disks across KVM, z/VM, Linux in LPAR, and z/OS.

Multiplatform resiliency for KVM configuration

Multiplatform Resiliency for KVM uses the xDR protocol to communicate with an xDR KVM proxy to send Libvirt commands, so xDR must be enabled in GDPS to support KVM. The xDR KVM Proxy is delivered as a Linux RPM for SLES or a DEB package for Ubuntu. The proxy serves as the middleware for GDPS. It communicates commands from GDPS to KVM, monitors the KVM environment, and communicates status information back to the GDPS Metro controlling system.

Figure 3-7 shows a GDPS Metro configuration with a mix of z/OS, Linux on z/VM and Linux on KVM, all managed by a single GDPS.
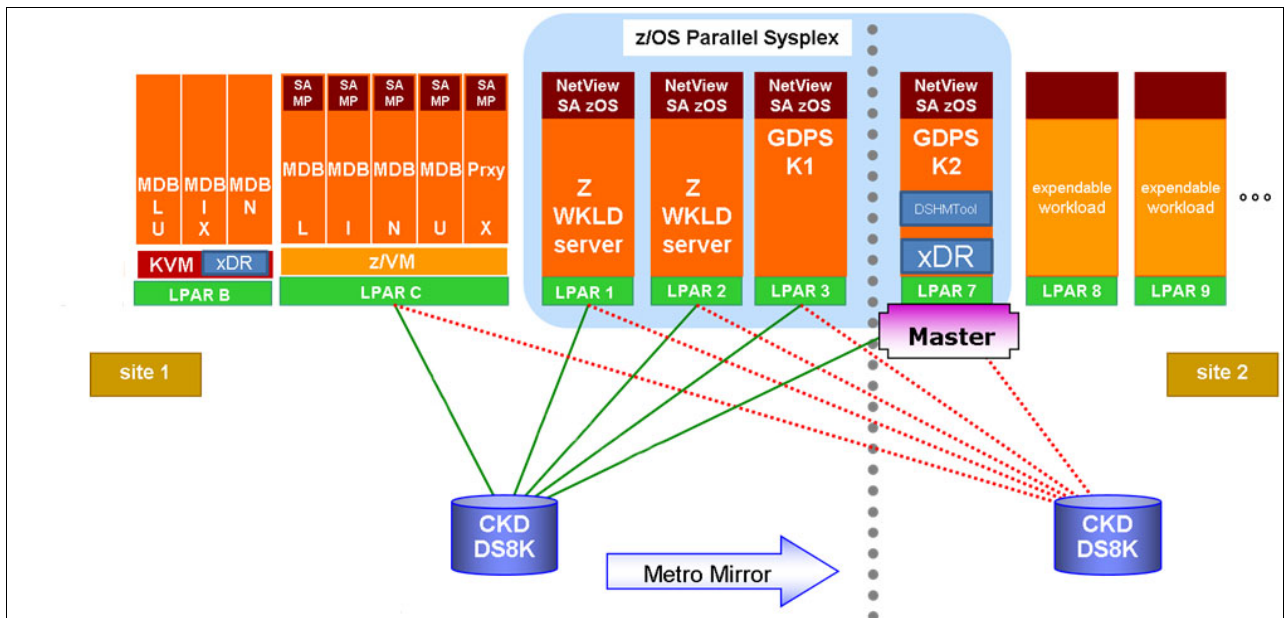
Figure 3-7   Multiplatform Resiliency for KVM

System and hardware management

System and hardware management capabilities similar to the ones that are available for z/OS systems are also available for KVM systems. GDPS Metro Multiplatform Resiliency for KVM provides support for:

- ► IBM z16, IBM z15, IBM z14, IBM z13®, IBM EC12, and IBM BC12 systems.
- ► LinuxONE Emperor and RockHopper Systems (based on IBM PR/SM BCPii support, not PR/SM 2 DPM).
- ► Planned switch (graceful shutdown and startup, or restart in place) for VMs and the KVM hypervisor.
- ► Planned HyperSwap for IBM Z coexistence. The user script that performs the planned HyperSwap must first shut down the KVM LPARs and restart them in place after the HyperSwap.
- ► Unplanned HyperSwap for IBM Z coexistence. The SWAPSITE12, or whichever script has control after an unplanned HyperSwap, must restart the KVM LPARs because the script was reset by the unplanned HyperSwap.
- ► Support for multiple KVMs on IBM Z, multiple z/VM systems, and multiple Linux LPARs, and the coexistence of KVM and z/VM.
- ► Single point of control to manage disk mirroring configuration (GDPS initiates replication direction changes).
- ► IBM ECKD Freeze.
- ► Support for Libvirt management interfaces for Hypervisor and VM management.
- ► ECKD disks on an IBM DS8000 storage subsystem.
- ► GDPS maintains its methods for managing z/OS, z/VM, and their respective disk replication.
- ► The VM workload startup is handled by using a guest Linux OS autostart policy.

Multiplatform Resiliency for KVM uses GDPS PR/SM BCPII commands and GDPS script statements for LPAR management. It does not use DPM on LinuxONE; only the PR/SM BCPII interface is used for LPAR management.

For remote management of KVM and its guests, Multiplatform Resiliency for KVM uses Libvirt commands, which are run by calling `libvirt virsh` commands from an xDR proxy that is running in the KVM Linux instance.

KVM HyperSwap status

The GDPS HyperSwap Monitor provides checking for z/OS systems to find out whether the z/OS systems that are managed by GDPS meet the required conditions for HyperSwap. Any z/OS system that does not meet the required conditions is marked as "not HyperSwap ready."

A planned HyperSwap may not run unless all systems are HyperSwap ready. If an unplanned swap is triggered, systems that are not HyperSwap ready are reset and the swap is performed with the participation of only those systems that are HyperSwap ready. KVM does not support HyperSwap and so always has a HyperSwap status of INHIBITED. Therefore, all KVM systems must be shutdown before a planned HyperSwap.

Summary

The Multiplatform Resiliency for KVM function of GDPS can provide a single point of control to monitor and manage Linux systems that are running under the KVM Hypervisor on IBM Z alongside z/OS systems and other Linux systems that are running under z/VM. It can also provide a coordinated failover for planned and unplanned events that can affect any of the KVM, z/VM, or z/OS resources. In short, you can achieve business resiliency across your entire enterprise.

Multiplatform Resiliency for Linux in LPAR

In this section, we describe xDR for Linux in LPAR, which is a separately priced feature of GDPS Metro. In this configuration, Linux runs natively in its own partition in an IBM Z server.

xDR for Linux in LPAR configuration

Figure 3-9 on page 80 shows a GDPS Metro configuration that includes one instance of Linux in LPAR.
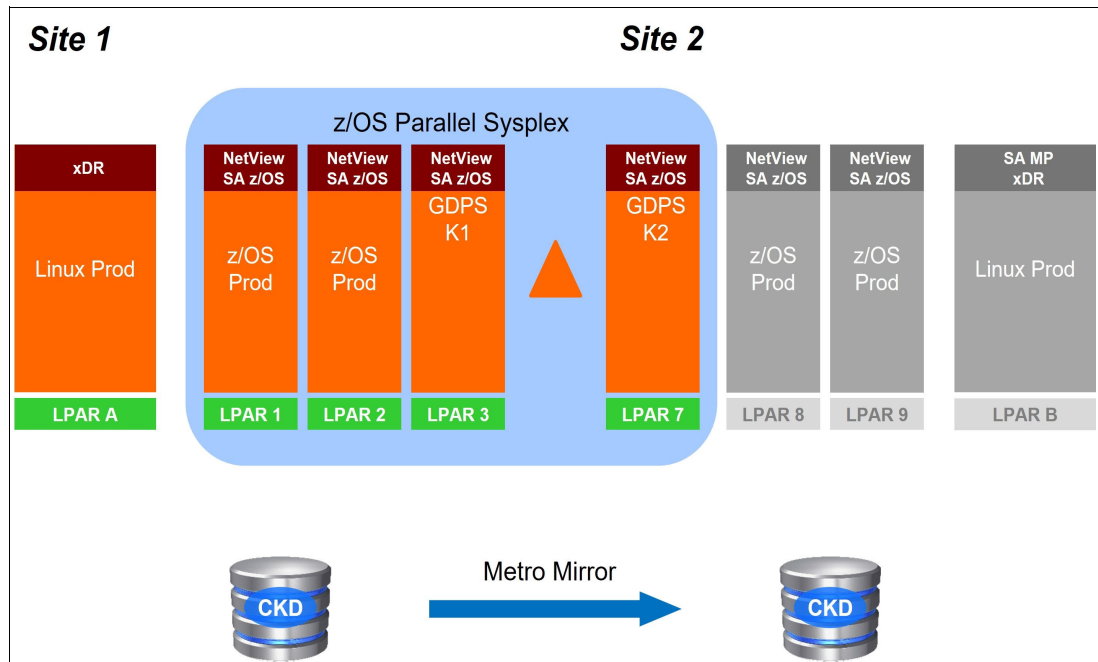
Figure 3-8   Multiplatform Resiliency for Linux in LPAR

In the figure, there is a z/OS Parallel Sysplex that consists of two z/OS production systems running in LPARs 1 and 2 at Site1, a GDPS controlling system K1 running in LPAR 3 at Site 1, and another GDPS Controlling system K2 running in LPAR 7 at Site 2. There is also a Linux in LPAR system hosting production Linux workloads running in LPAR A at Site 1. In Site 2, there are backup LPARs that are available for the two z/OS production systems and the Linux in LPAR system to take over the production workloads if a disaster occurs at the primary site. Finally, the data for all systems is on CKD disk and mirrored from Site 1 to Site 2 by Metro Mirror.

SA MP, along with the xDR agent, must be running in each xDR-managed Linux in LPAR system. SA MP on each system monitors that system and reports status information to the GDPS Metro controlling system, along with alerts for any disk errors that are encountered by that system. Also, the controlling system communicates commands to the Linux in LPAR system through SA MP.

Disk configuration and replication management

The disk devices that are used by the Linux in LPAR system are defined to GDPS, which enables GDPS to manage Metro Mirror for those devices. The disks that are used by the Linux in LPAR system must be CKD disks.

GDPS xDR support enables the definition of the Metro Mirror secondary devices that are configured to the Linux in LPAR system to be defined in an alternative subchannel set. This approach can simplify definitions and provide high scalability for your disk configuration. For more information, see "Addressing z/OS device limits in a GDPS Metro environment" on page 25.

Protecting data integrity across multiple platforms

The HyperSwap function is not available for Linux in LPAR systems. However, GDPS Metro provides coexistence for planned and unplanned HyperSwap events.

For planned HyperSwap, the user-written script that performs the operation must first shut down the Linux in LPAR systems and restart them in place after the HyperSwap is complete. For an unplanned HyperSwap, GDPS Metro provides takeover scripts that gain control after an unplanned HyperSwap occurs. These takeover scripts can be used to automatically restart the Linux in LPAR systems, which were reset by the unplanned HyperSwap operation.

In either case, GDPS Metro coordinates the operation across the CKD disks for all systems in the GDPS configuration, whether they are z/OS systems, Linux systems running under z/VM or KVM, Linux in LPAR systems, or IBM Db2 Analytics Accelerator on IBM Z. This configuration maintains data consistency across the entire environment.

For site failures, GDPS Metro provides a coordinated Freeze across the CKD devices for all systems in the configuration to provide a consistent copy of data on the secondary devices to enable recovery.

System and hardware management

The LPAR that the Linux in LPAR system is running in is also defined to GDPS, which enables GDPS to manage the Linux in LPAR system.

System and hardware management capabilities that are similar to the ones that are available for z/OS systems are also available for Linux in LPAR systems. Capabilities are provided for the following tasks:

► Load the Linux in LPAR system.
► Stop the Linux in LPAR system.
► Reset the Linux in LPAR system.
► Activate the Linux in LPAR system LPAR.
► Deactivate the Linux in LPAR system LPAR.
► Query the status of the Linux in LPAR system.

Use these capabilities to manage the Linux in LPAR system from the GDPS Standard Actions panel and to automate the following workflows:

► Planned site switch to Site 2

► Unplanned failover to Site 2, which is triggered by an Linux in LPAR system failure

► Unplanned failover to Site2, which is triggered by a PPRC primary problem

► Unplanned freeze of PPRC mirroring, which is triggered by a PPRC mirroring problem

► Return home to Site 1 as a planned action

► DR testing

Multiplatform Resiliency for Linux in LPAR uses GDPS PR/SM BCPII commands and GDPS script statements for LPAR management. It does not use DPM on LinuxONE; only the PR/SM BCPII interface is used for LPAR management.

Multiplatform Resiliency for IBM Db2 Analytics Accelerator on IBM Z

In this section, we discuss xDR for IBM Db2 Analytics Accelerator on an IBM Z configuration.

xDR for IBM Db2 Analytics Accelerator on IBM Z configuration

Figure 3-9 shows a GDPS Metro configuration with an IBM Db2 Analytics Accelerator on IBM Z SSC.
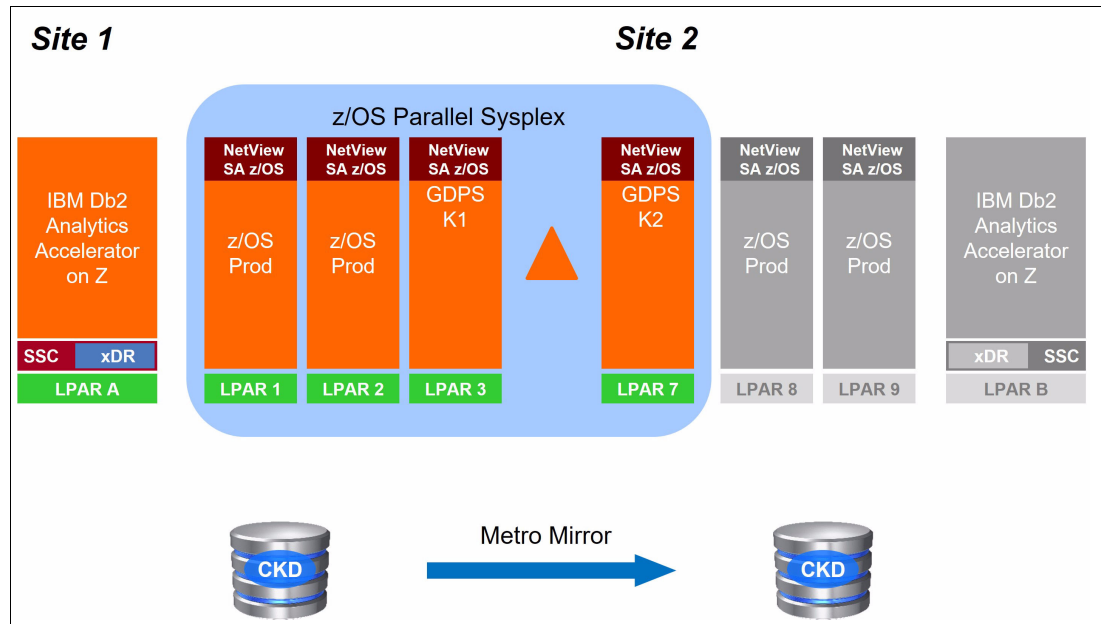
Figure 3-9   Multiplatform Resiliency for IBM Db2 Analytics Accelerator on IBM Z

GDPS Metro supports IBM Db2 Analytics Accelerator on IBM Z in Active/Passive mode. Active/Passive mode means that only one active instance of IBM Db2 Analytics Accelerator on IBM Z exists at any time (LPAR A in Figure 3-9). A backup LPAR is also available at the DR site to take over the IBM Db2 Analytics Accelerator on IBM Z workload if a disaster occurs at the primary site (LPAR B in Figure 3-9).

GDPS Metro uses the xDR protocol to communicate with the IBM Db2 Analytics Accelerator on IBM Z SSC. An xDR agent, running on the SSC, serves as the middleware for GDPS. It communicates commands from GDPS to the SSC, monitors the SSC environment, and communicates status information back to the GDPS Metro controlling system.

Disk configuration and replication management

The disk devices that are used by the IBM Db2 Analytics Accelerator on IBM Z are defined to GDPS, which enables GDPS to manage Metro Mirror for those devices. The disks that are being used by the IBM Db2 Analytics Accelerator on IBM Z must be CKD disks.

GDPS xDR support allows for the definition of the Metro Mirror secondary devices that are configured to the IBM Db2 Analytics Accelerator on IBM Z to be defined in an alternative subchannel set[2]. This approach can simplify definitions and provide high scalability for your disk configuration. For more information, see "Addressing z/OS device limits in a GDPS Metro environment" on page 25.

Protecting data integrity across multiple platforms

The HyperSwap function is not available for IBM Db2 Analytics Accelerator on IBM Z. However, GDPS Metro provides coexistence for planned and unplanned HyperSwap events.

For planned HyperSwap, the user-written script that performs the operation must first shutdown the IBM Db2 Analytics Accelerator on IBM Z LPARs and restart them in place after the HyperSwap is complete. For an unplanned HyperSwap, GDPS Metro provides takeover scripts that get control after an unplanned HyperSwap occurs. These takeover scripts can be used to automatically restart the IBM Db2 Analytics Accelerator on IBM Z LPARs, which were reset by the unplanned HyperSwap operation.

In either case, GDPS Metro coordinates the operation across the CKD disks for all systems in the GDPS configuration, whether they be z/OS systems, Linux systems running under z/VM or KVM, Linux in LPAR systems, or IBM Db2 Analytics Accelerator on IBM Z. This configuration maintains data consistency across the entire environment.

For site failures, GDPS Metro provides a coordinated Freeze across the CKD devices for all systems in the configuration to provide a consistent copy of data on the secondary devices to enable recovery.

System and hardware management

The IBM Db2 Analytics Accelerator on IBM Z SSC LPAR is also defined to GDPS, which enables GDPS to manage the IBM Db2 Analytics Accelerator on IBM Z.

System and hardware management capabilities that are similar to the ones that are available for z/OS systems are also available for IBM Db2 Analytics Accelerator on IBM Z SSC LPARs. Capabilities are provided for the following tasks:

► Stop the IBM Db2 Analytics Accelerator on IBM Z
► Reset the IBM Db2 Analytics Accelerator on IBM Z
► Activate the IBM Db2 Analytics Accelerator on IBM Z SSC LPAR
► Deactivate the IBM Db2 Analytics Accelerator on IBM Z SSC LPAR
► Query the status of the IBM Db2 Analytics Accelerator on IBM Z

Use these capabilities to manage the IBM Db2 Analytics Accelerator on IBM Z from the GDPS Standard Actions panel and to automate the following workflows:

► Planned site switch to Site 2

► Unplanned failover to Site 2, which is triggered by an IBM Db2 Analytics Accelerator on IBM Z failure

► Unplanned failover to Site2, which is triggered by a PPRC primary problem

► Unplanned freeze of PPRC mirroring, which is triggered by a PPRC mirroring problem

► Return home to Site 1 as a planned action

► DR testing

---

[2] Only alternative subchannel set 1 (MSS1) is supported for defining the Metro Mirror secondary devices.

GDPS xDR for IBM Db2 Analytics Accelerator on IBM Z uses PR/SM BCPII commands and IBM Z Hardware Management Console Web Services application programming interface (API) requests for managing the SSC LPAR. Therefore, the GDPS Metro controlling systems must be configured to enable communications for these interfaces.

3.4 Managing z/OS systems outside of the GDPS sysplex

In 3.1.4, "Protecting other CKD data" on page 60, we describe a method that allows GDPS to monitor and manage Metro Mirror on behalf of systems that are not running in the GDPS sysplex. We refer to such non-GDPS systems outside of the sysplex as *foreign systems* and we refer to the disk of these systems as *foreign disk*.

Managing foreign systems and the foreign disk by using the method that is described in 3.1.4, "Protecting other CKD data" on page 60 has a key limitation in that this method does not support HyperSwap for the foreign systems.

Although the foreign disks are included in the swap scope, the foreign systems must stop before a planned swap and are denied access to the swapped-from disks by hanging on the ELB (or by the Soft Fence that is established by GDPS) as a result of an unplanned swap after, which they (the foreign systems) must be reset and reloaded from the swapped-to disk.

However, GDPS Metro provides a feature that is known as the *z/OS Proxy* that extends the near CA protection of HyperSwap to z/OS systems that are running outside of the GDPS sysplex, which includes stand-alone z/OS systems (MONOPLEX or XCFLOCAL) and systems that are running in a multi-system sysplex other than the GDPS sysplex.

In a z/OS Proxy environment, a GDPS Metro agent runs in each of the z/OS Proxy-managed systems that are outside of the GDPS sysplex. This agent, which is known as the *z/OS Proxy*, communicates with the primary GDPS controlling system, which facilitates coordinated planned and unplanned HyperSwap, and coordinated freeze processing across the systems in the GDPS sysplex and all z/OS systems that are managed by the z/OS Proxy.

In addition to Metro Mirror, Freeze and HyperSwap management, much of the hardware management (for example, automated system resets and IPLs) of the z/OS Proxy-managed systems is provided. However, some GDPS Metro functions, such as the CDS and CF management functions for z/OS Proxy-managed systems running in foreign sysplexes, are not available.

Figure 3-10 on page 83 shows a basic configuration to help explain the support that GDPS provides in monitoring and managing the z/OS Proxy-managed systems and the mirrored disks that are used by these systems.
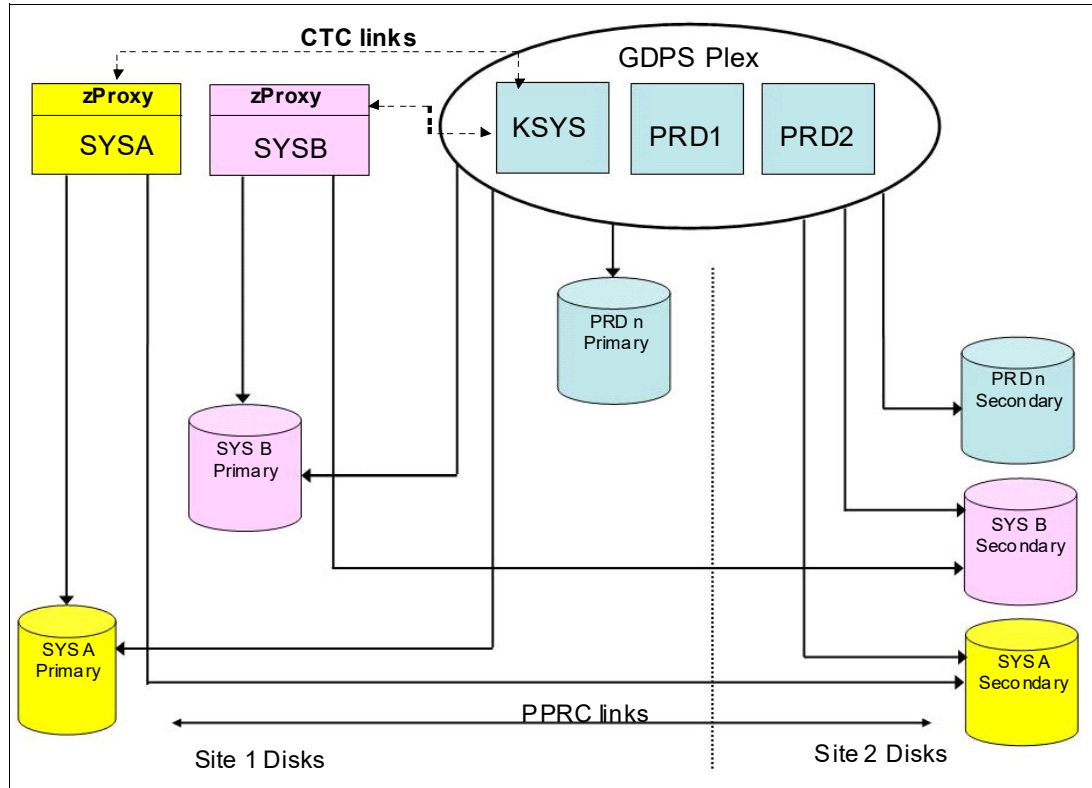
Figure 3-10   Sample z/OS Proxy Environment

As shown in Figure 3-10, the traditional GDPS sysplex environment consists of production systems PRD1 and PRD2 and the controlling system KSYS. The primary disks for these GDPS production systems in Site1 are mirrored to Site2 using Metro Mirror. This environment represents a standard GDPS Metro installation.

The systems SYSA and SYSB are z/OS Proxy-managed systems. They are outside of the GDPS sysplex and do not run GDPS NetView or System Automation code. Instead, they run the z/OS Proxy agent, which communicates and coordinates actions with the Master GDPS controlling system.

The z/OS Proxy-managed systems are connected to the controlling systems by using FICON Channel-to-Channel connections.

The z/OS Proxy-managed systems do not need host attachment to the disks that belong to the systems in the GDPS sysplex and do not need to define those disks. However, the systems in the GDPS sysplex *do* need to have UCBs for and have host channel attachment to all Metro Mirrored disk, their own and all disks that belong to the z/OS Proxy-managed disks.

3.4.1 z/OS Proxy disk and disk subsystem sharing

The mirrored disk attached to the z/OS Proxy-managed systems can be on separate physical disk subsystems or in the same physical disk subsystems as the disk belonging to the systems in the GDPS sysplex. Mirrored disks for the systems in the GDPS sysplex and the z/OS Proxy-managed systems can also be colocated in the same LSS.

Because hardware reserves are now allowed in a GDPS HyperSwap environment, GDPS systems and z/OS Proxy-managed systems can share the GDPS-managed Metro Mirrored disks. Systems in the GDPS sysplex can share disks among themselves more efficiently by converting reserves to global enqueues and this configuration is a best practice for devices that are shared only within the GDPS sysplex. Similarly, systems in any foreign sysplex can share disks with each other if reserves in the foreign sysplex are converted to global enqueues.

3.5 Managing the GDPS environment

We saw how GDPS Metro can protect almost any type of data that can be in a disk subsystem. It can also provide data consistency across multiple platforms. However, as discussed in Chapter 1, "Introducing business resilience and the role of IBM GDPS" on page 1, most IBM Z outages are not disasters. Instead, they are planned outages, with a small percentage of unplanned outages.

In this section, we describe the other aspect of GDPS Metro; that is, its ability to monitor and manage the resources in its environment. GDPS provides several mechanisms to help you manage the GDPS sysplex and resources within that sysplex. These mechanisms include user interfaces, scripts, and APIs. We review these mechanisms and provide more information about the management of the GPS environment in the following sections.

3.5.1 User interfaces

Two primary user interface options are available for GDPS Metro: The NetView 3270 panels and a browser-based GUI (also referred to as *GDPS GUI* in this book).

An example of the main GDPS Metro 3270-based panel is shown in Figure 3-11.

```
VPCPPNLN                 GDPS Metro (DOTANKINIENAPLUS)            GDPS V4.R7.M0
User ID: FREDO
                       ---- GDPS Status Indicators ----


  System          = GBC1    - A6PB1      PPRC and HyperSwap status = OK
  Current Primary = GBC2    - A6PB4      Primary Dasd = RS1
  Debug           = ON


                                         Tape Status  = OK


                     ---- GDPS Options ----

      1       Dasd Remote Copy      7       Sysplex Resource Management
      2       Tape Remote Copy      8       Debug ON/OFF
      3       Standard Actions      9       View Definitions
                                    H       Health Checks and Diagnostics
                                    C       Config Management
      6       Planned Actions       M       Run Monitor1/Monitor3
                                    L       Logical Corruption Protection



Selection ===>  _
              Licensed Materials - Property of IBM
        6942-35B © Copyright IBM Corp. 1998, 2024  All Rights Reserved.
```

Figure 3-11   Main GDPS Metro 3270-based panel

The panel that is shown in Figure 3-11 includes a summary of configuration status at the top, and a menu of selectable choices. As an example to view the disk mirroring (Dasd Remote Copy) panels, enter 1 at the Selection prompt, and then click **Enter**.

GDPS GUI

The GDPS GUI is a browser-based interface that is designed to improve operator productivity. The GDPS GUI provides the same functional capabilities as the 3270-based panel, such as providing management capabilities for Remote Copy Management, Standard Actions, Sysplex Resource Management, Status Display Facility (SDF) Monitoring, and browsing the CANZLOG by using simple point-and-click procedures. Advanced sorting and filtering is available in most of the views that are provided by the GDPS GUI. In addition, users can open multiple windows or tabs to allow for continuous status monitoring, while performing other GDPS Metro management functions.

The GDPS GUI is available in stand-alone GDPS Metro environments and GDPS MGM 3-site and 4-site environments (for more information about GDPS MGM 3-site and 4-site environments, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237).

The GDPS GUI display has four main sections:

► The application header at the top of the page that provides an Actions button for carrying out some GDPS tasks, along with the help function and the ability to logoff or switch between target systems.

► The application menu that is down the left side of the window. This menu gives access to various features and functions that are available through the GDPS GUI.

► The active window that shows context-based content, depending on the selected function. This tabbed area is where the user can switch context by clicking a different tab.

► A status summary area that is shown at the bottom of the display.

Note: For the remainder of this section, only the GDPS GUI is shown to illustrate the various GDPS management functions. The equivalent traditional 3270 panels exist but are not shown here.

The initial status window (known as the dashboard) of the GDPS Metro GUI is shown in Figure 3-12. This window provides an instant view of the status and direction of replication, HyperSwap status, and systems and systems availability. Hovering over the various icons provides more information through windows.
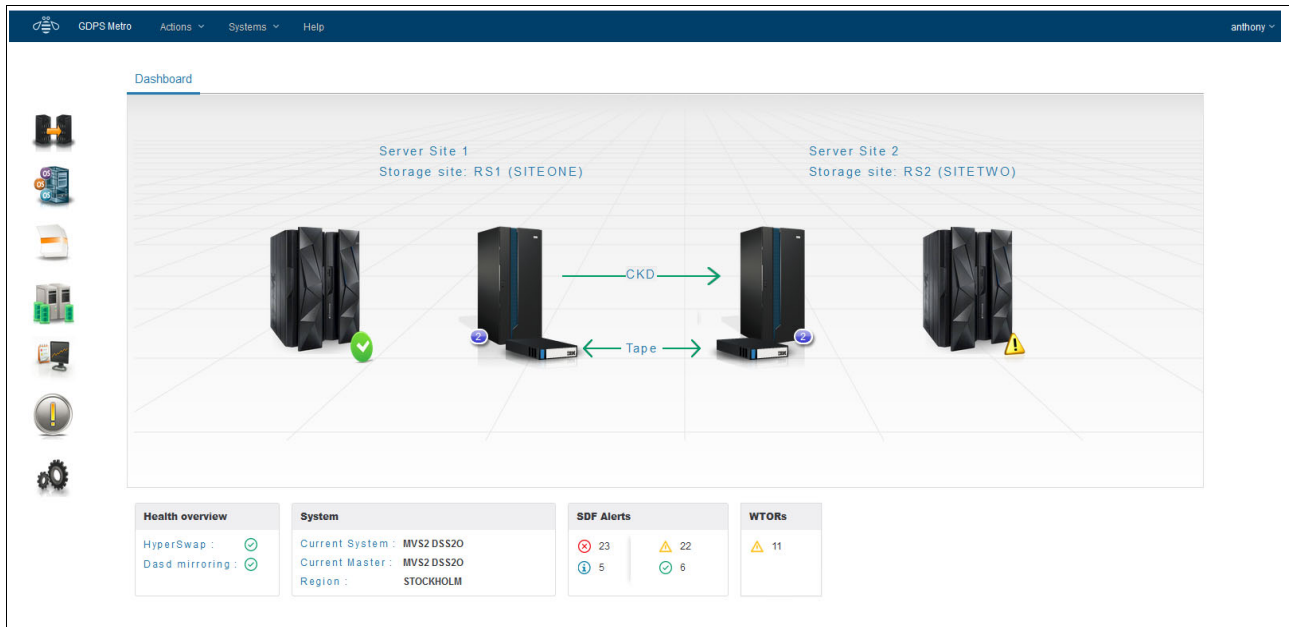
Figure 3-12   GDPS GUI initial window

Monitoring function: Status Display Facility

GDPS also provides many monitors to check the status of disks, sysplex resources, and other items. Anytime there is a configuration change, or something in GDPS that requires manual intervention, GDPS raises an alert. GDPS uses the SDF that is provided by System Automation as the primary status feedback mechanism for GDPS.

GDPS provides a dynamically updated window, as shown in Figure 3-13 on page 87. There is a summary of all current alerts at the bottom of each window. The initial view that is presented is for the SDF trace entries so you can follow; for example, script execution. Click one of the icons representing the other alert categories to view the different alerts that are associated with automation or remote copy in either site, or click **All** to see all alerts. You can sort and filter the alerts based on several fields that are presented, such as severity.
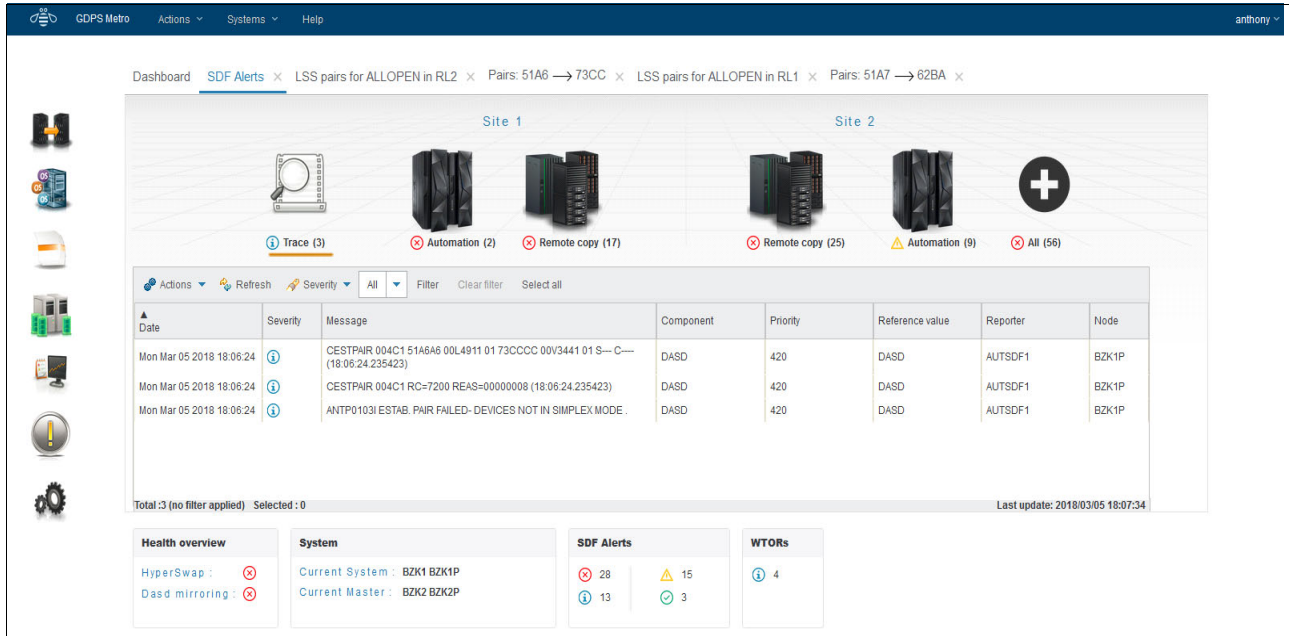
Figure 3-13   GDPS GUI SDF window

The GDPS GUI refreshes the alerts automatically every 10 seconds by default. As with the 3270 panel, if there is a configuration change or a condition that requires special attention, the color of the fields change based on the severity of the alert. By pointing to and clicking any of the highlighted fields, you can obtain detailed information regarding the alert.

The color of the fields change based on the severity of the alert. By pointing to and clicking any of the highlighted fields, you can obtain detailed information regarding the alert.

Remote copy windows

The z/OS Advanced Copy Services capabilities are powerful, but the native command-line interface (CLI), z/OS TSO, and ICKDSF interfaces are not as simple as the DASD remote copy panels are. To more easily check and manage the remote copy environment, use the DASD remote copy windows that are provided by GDPS.

For GDPS to manage the remote copy environment, you must first define the configuration (primary and secondary LSSs, primary and secondary devices, and PPRC links) to GDPS in a file called the GEOPARM file. This GEOPARM file can be edited and introduced to GDPS directly from the GDPS GUI.

After the configuration is known to GDPS, you can use the GUI to check that the current configuration matches the one you want. You can start, stop, suspend, and resynchronize mirroring and you can perform these actions at the device level, the LSS level, or both.

You can also manage the PPRC links dynamically, which means you do not have to update the GEOPARM file and then initiate the process to load a new DASD configuration. This approach is helpful because you might need to temporarily add PPRC links to handle an increase in update activity, or you might need to remove failing PPRC links that can cause significant mirroring delays and with the capability to dynamically manage the links, you can make these changes while avoiding the temporary disruption to storage availability (HyperSwap) that happens when a new DASD configuration is loaded.

Figure 3-14 shows the mirroring window for CKD devices at the LSS level.
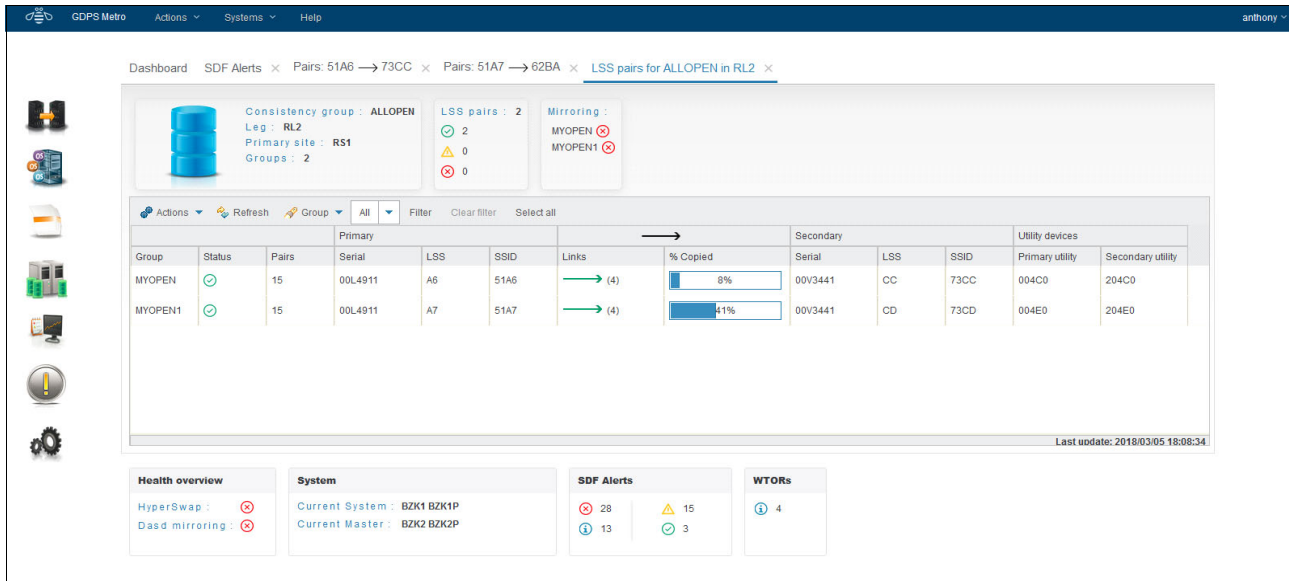
Figure 3-14   GDPS GUI DASD Remote Copy: LSS-level detail window

The top section of the DASD Remote Copy LSS-level window summarizes the number of LSS pairs and their status, including their mirroring status, in the selected consistency group.

The middle section of the window contains a table with one row for each LSS pair in the selected consistency group. In addition to the rows for each LSS, there is a header row containing an Action menu that you use to perform the various DASD management tasks, and a filter menu to filter the information presented.

To perform an action on a single LSS-pair (SSID-pair), double-click a row in the table. The frame that is shown in Figure 3-15 on page 89 is then displayed. The table in this frame shows each of the mirrored device pairs within a single LSS-pair, along with the status of each pair. In this example, one of the pairs is fully synchronized and in duplex status and the rest of the pairs are in pending status, as summarized in the top section of the window. More information can be viewed for each pair by double-clicking the row, or by selecting the row with a single click and then selecting **Query** from the Actions menu.
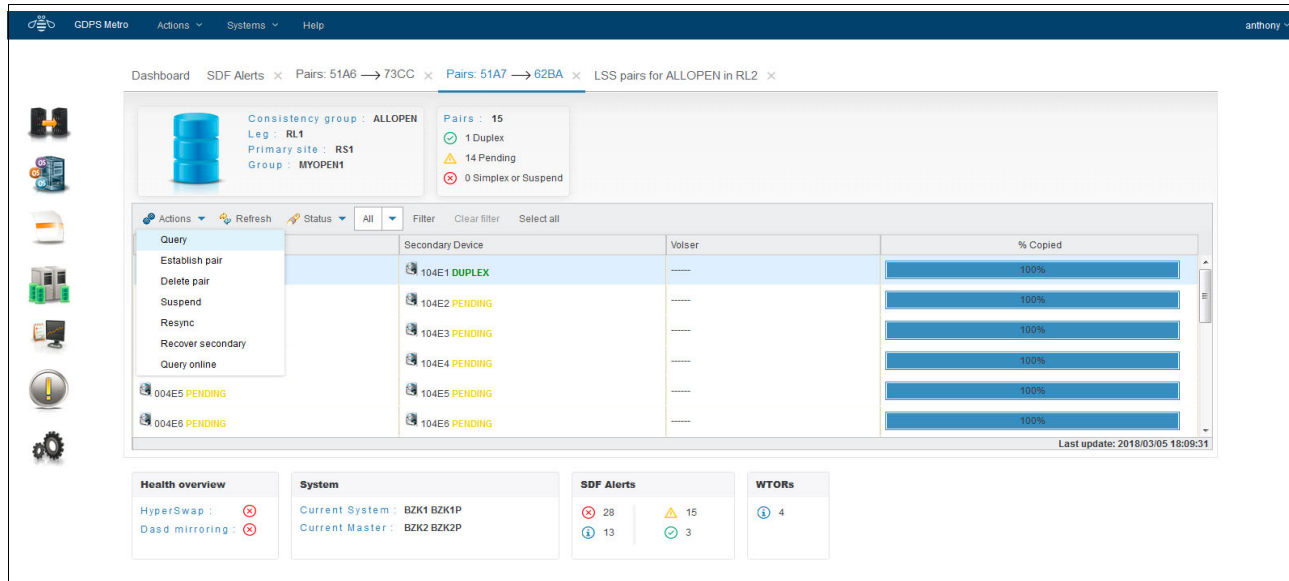
Figure 3-15 GDPS GUI DASD Remote Copy: Device-level detail window

If you are familiar with using the TSO or ICKDSF interfaces, you might appreciate the ease of use of the DASD remote copy panels.

The GUI that is provided by GDPS is *not* intended to be a remote copy monitoring tool. Because of the overhead that is involved in gathering the information for every device to populate the windows, GDPS gathers this data only on a timed basis, or on demand following an operator instruction. The normal interface for finding out about remote copy status or problems is the SDF.

Standard Actions

GDPS provides facilities to help manage many common system-related planned actions. There are two reasons to use the GDPS facilities to perform these actions, which are known as *Standard Actions*:

► They are tested and based on IBM preferred procedures.

► Using the GDPS interface lets GDPS know that the changes that it is seeing (for example, a system being partitioned out of the sysplex) are planned changes, and therefore GDPS does not react to these events.

Standard Actions are single-step actions and are intended to impact only one resource. Examples are starting a system IPL, maintaining the various IPL address and load parameters that can be used to IPL a system, selecting the IPL address and load parameters to be used the next time a system IPL is performed, or activating/deactivating an LPAR.

If you want to stop a system, change its IPL address, then perform an IPL, you start three separate Standard Actions, one after the other. GDPS scripting, as described in 3.5.2, "GDPS scripts" on page 91, is a facility that is suited to multi-step, multi-system actions.

The GDPS Metro Standard Actions GUI window is shown in Figure 3-16. It displays all the systems that are managed by GDPS Metro. It shows the status and various IPL information for each system. To perform actions on a system, select the row with a single click and then, select the wanted action from the Actions menu.
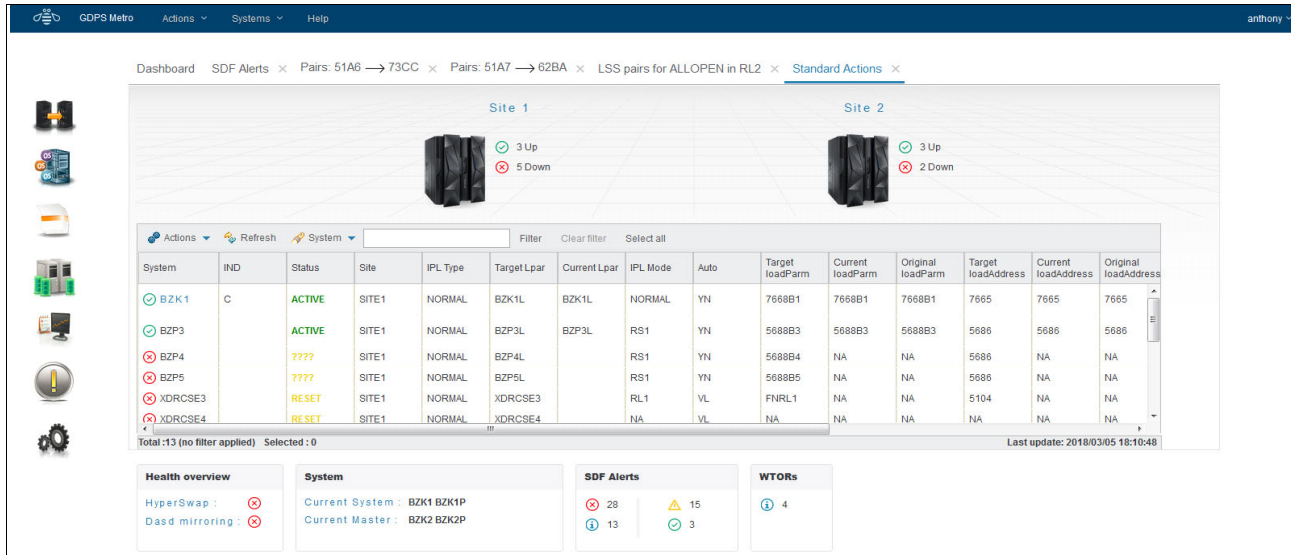
Figure 3-16   GDPS GUI Standard Actions window

GDPS supports taking a stand-alone dump by using the GDPS Standard Actions window. Clients that use GDPS facilities to perform HMC actions no longer need to use the HMC for taking stand-alone dumps.

Sysplex resource management

There are certain resources that are vital to the health and availability of the sysplex. In a multisite sysplex, it can be complex trying to manage these resources to provide the required availability while ensuring that any changes do not introduce a single point of failure.

The GDPS Metro Sysplex Resource Management window, as shown in Figure 3-17 on page 91, provides you with the ability to manage the sysplex resources without having knowledge about where the resources exist. Click the resource type (CDS or CFs) to open a panel to manage each resource type.

For example, normally you have primary CDS in Site1, and your alternates in Site2. However, if you are shutting down Site1, you still want to have a Primary and Secondary set of CDS, but both must be in Site2. The GDPS Sysplex Resource Management panels provide this capability, without you having to know specifically where each CDS is located.

GDPS also provides facilities to manage the CFs in your sysplex. These facilities allow for isolating all of your structures in the CF or CFs in a single site and returning to your normal configuration with structures spread across (and possibly duplexed across) the CFs in the two sites.

Use the maintenance mode switch to start or stop maintenance mode on a single CF (or multiple CFs, if all selected CFs are in the same site). DRAIN, ENABLE, and POPULATE functions are still available for single CFs.

Managing your sysplex resources can also be accomplished through GDPS scripts, which provide an automated means for managing CDSs and CFs for planned and unplanned site or disk subsystem outages. For more information about GDPS scripting capability, see 3.5.2, "GDPS scripts" on page 91.
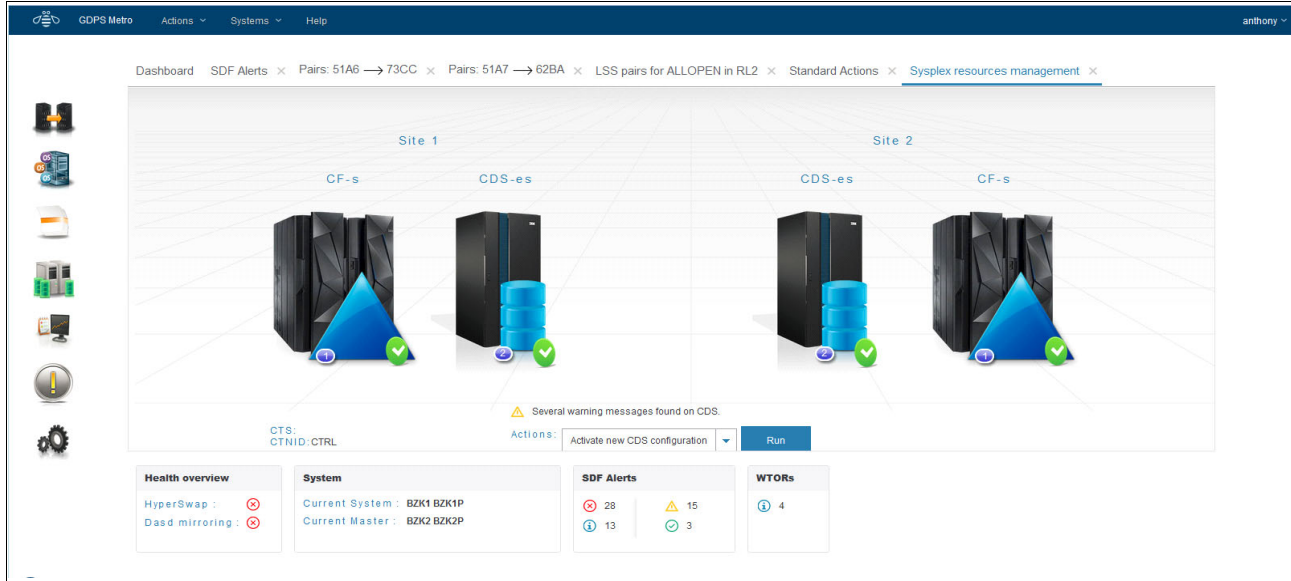
Figure 3-17   GDPS GUI Sysplex Resource Management window

Installation Verification Program

The GUI installation requires many customization steps to configure many separate components. Failure to perform the required steps correctly can result in the GDPS functions failing when you need to use them. This situation can potentially result in outages of systems or delays in recovery and loss of availability. If you cannot identify the source of your problems, the resolution time can be longer because you must report the problem to defect support. More time is spent debugging and fixing the environment.

IVP ("Installation Verification Program" on page 75) validates the installation and customization of the xDR environment and the installation of the GDPS GUI. It also verifies that the major functions and features are working correctly.

IVP is an operator-initiated program that you can use after initial setup and later to ensure that changes to the environment did not break the GUI setup.

3.5.2 GDPS scripts

We showed how the GDPS GUI and the GDPS 3270-based panels provide powerful functions to help you manage GDPS resources. However, using GDPS panels is only one way of accessing this capability. Especially when you need to initiate what might be a complex, compound, multistep procedure involving multiple GDPS resources, it is simpler to use a script that, in effect, is a workflow.

Nearly all main functions that can be initiated through the GDPS panels are also available by using GDPS scripts. Scripts also provide extra capabilities that are not available using the panels.

A "script" is a procedure, which is written by you, that pulls together one or more GDPS functions. Scripts can be initiated manually for a planned activity through the GDPS GUI or panels (by using the Planned Actions interface), automatically by GDPS in response to an event (such as a HyperSwap), or through a batch interface. GDPS performs the first statement in the list, checks the result, and only if it is successful, proceeds to the next statement. If you perform the same steps manually, you would have to check the results, which can be time-consuming, and initiate the next action. With scripts, the process is automated.

Automating complex tasks can sometimes require scripts to contain numerous steps and these scripts can run for significant amounts of time. This issue increases the possibility that a running script might need to be stopped intentionally or fail unexpectedly because of unusual environmental conditions, for example.

To maximize efficiency in these cases, GDPS tracks how far a script progressed. If a script fails or is stopped manually, the script can be restarted at the suitable point, which eliminates unnecessary duplicate processing and saves time.

Scripts can easily be customized to automate the handling of various situations, both to handle planned changes and unplanned situations. A script is an important aspect of GDPS. Scripts are powerful because they can access the full capability of GDPS. The ability to start all the GDPS functions through a script provides the following benefits:

► Speed

The script runs the requested actions and check the results at machine speeds. Unlike a human, it does not need to search for the latest procedures or the commands manual.

► Consistency

If you look into most computer rooms immediately following a system outage, what would you see? Mayhem, with operators frantically scrambling for the latest system programmer instructions. All the phones ringing. Every manager within reach asking when the service will be restored. And every systems programmer with access is vying for control of the keyboards. This chaos results in errors because humans naturally make mistakes when under pressure. But with automation, your well-tested procedures runs in exactly the same way, time after time, regardless of how much you shout at them.

► Thoroughly tested procedures

Because they behave in a consistent manner, you can test your procedures over and over until you are sure that they do everything that you want, in exactly the manner that you want. Also, because you need to code everything and cannot assume a level of knowledge (as you might with instructions that are intended for a human), you are forced to thoroughly think out every aspect of the action the script is intended to undertake. And because of the repeatability and ease of use of the scripts, they lend themselves more easily to frequent testing than manual procedures.

Planned Actions

Planned Actions are GDPS scripts that are started from the panels (option 6 on the main GDPS panel, as shown in Figure 3-11 on page 84) or from the GUI. GDPS scripts are procedures that pull together into a list one or more GDPS functions. Scripted procedures that you use for a planned change are known as *control scripts*.

A control script that is running can be stopped if necessary. Control scripts that were stopped or that failed can be restarted at any step of the script. These capabilities provide a powerful and flexible workflow management framework.

For example, you can have a short script that stops a system and then restarts it in an alternative LPAR location, as shown in Example 3-1. The sample also handles deactivating the original LPAR after the system is stopped and activating the alternative LPAR before the system is started in this location.

Example 3-1   Sample script to restart a system

```
COMM='Example script to restart system SYS1 on alternative ABNORMAL LPAR location'
SYSPLEX='STOP SYS1'
SYSPLEX='DEACTIVATE SYS1'
IPLTYPE='SYS1 ABNORMAL'
SYSPLEX='ACTIVATE SYS1 LPAR'
SYSPLEX='LOAD SYS1'
```

A more complex example of a Planned Action is shown in Figure 3-18.
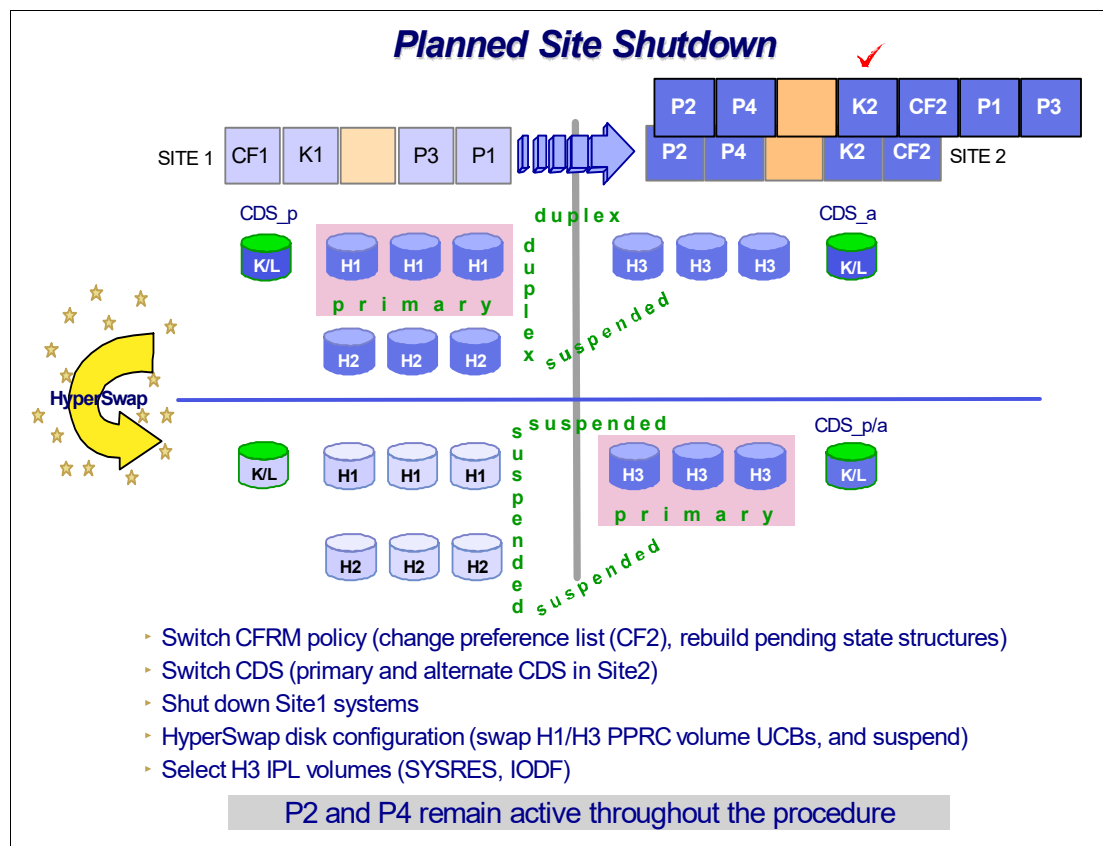
Figure 3-18   GDPS Metro Planned Action

In this example, a single action in GDPS running a planned script of only a few lines results in a complete planned site switch. Specifically, the following actions are done by GDPS:

► The systems in Site1, P1 and P3, are stopped (P2 and P4 remain active in this example).

► The sysplex resources (CDS and CF) are switched to use only the resources in Site2.

► A HyperSwap is ran to use the disk in Site2 (RS3 disk). As a result of the swap GDPS automatically switches the IPL parameters (IPL address and load parameters) to reflect the new configuration.

► The IPL location for the P1 and P3 systems are changed to the backup LPAR location in Site2.

- ► The backup LPAR locations for P1 and P3 systems are activated.
- ► P1 and P3 are started in Site2 using the disk in Site2.

Using GDPS removes the reliance on out-of-date documentation, provides a single repository for information about IPL addresses and load parameters, and ensures that the process is done the same way every time with no vital steps accidentally overlooked.

STP CTN role reassignments: Planned operations

GDPS provides a script statement to reconfigure an STP-only CTN by reassigning the STP-only CTN server roles. In an STP CTN servers (CPCs) are assigned special roles to identify which CPC is preferred to be the clock source (Preferred Time Server (PTS)), which CPC is able to take over as the clock source for planned and unplanned events (Backup Time Server (BTS)), which CPC is the active clock source (Current Time Server (CTS)), and which CPC helps with STP recovery (Arbiter).

As a best practice, reassign the server roles be reassigned before performing planned disruptive actions on any of these special role servers. Examples of planned disruptive actions are power-on reset (POR) and Activate/Deactivate. The script statement can be integrated as part of your existing control scripts to perform these planned disruptive actions.

For example, if you are planning to deactivate the CPC that is the PTS/CTS, you can now run a script to perform the following tasks:

- ► Reassign the PTS/CTS role to a different CPC in the CTN
- ► Optionally, also reassign the BTS and Arbiter roles if required
- ► Run script statements that you might already have in place today to deactivate the PTS/CTS CPC

After the disruptive action is completed you can run a second script to restore the STP roles to their normal operational state, as listed here:

- ► Script statement to activate the CPC
- ► Reassign the STP server roles to their normal operational state
- ► Statements that you might already have in existing scripts to perform starts

Takeover scripts

Takeover scripts define actions that GDPS runs automatically after specific unplanned events occur. A reserved name is defined for each takeover script that correlates it to the specific unplanned event that it addresses. When one of the unplanned events occurs, GDPS Metro automatically runs the suitable takeover script if it was defined.

Two types of takeover scripts are available: post swap and CPC failure. The following sections provide more information about each type of script.

Post-swap scripts

Post-swap scripts define actions that GDPS runs after an unplanned HyperSwap. Several specific unplanned HyperSwap scenarios are available and for each one, a reserved name for the associated takeover script is used. In an unplanned HyperSwap trigger, GDPS Metro immediately and automatically runs an unplanned HyperSwap. Following the HyperSwap operation, GDPS then runs the suitable takeover script if it was defined.

The post-swap takeover scripts include reserved names that help GDPS determine the applicability of the script for the unplanned swap situation. For example, if an unplanned swap from RS1 to RS3 occurs, GDPS automatically schedules a script that is named SWAPSite13 if you defined it. Typical actions that you might want to perform following an unplanned HyperSwap include resynchronizing mirroring for the MTIR replication leg in a dual-leg environment and changing the couple data set configuration.

For HyperSwap operations that swap production from one site to another, you might want to reconfigure STP to keep the CTS role on the CPC that is in the same site as the swapped-to, new primary devices.

CPC failure scripts

GDPS monitors data-related events and also performs system-related monitoring. When GDPS detects that a z/OS system is no longer active, it verifies whether the policy definition indicates that Auto IPL was enabled, that the threshold of the number of IPLs in the predefined time window was not exceeded, and that no planned action is active. If these conditions are met, GDPS can automatically restart the system in place, bring it back into the Parallel Sysplex, and restart the application workload (see Figure 3-19).

IPL and Restart in Place

AUTOMATION

ANALYSIS
- AutoIPL on?
- Threshold not exceeded?
- No planned action active?

SITUATION MANAGEMENT
- IPL system into sysplex
- Initiate application startup

System Failure

Adequate, fast response to exception condition
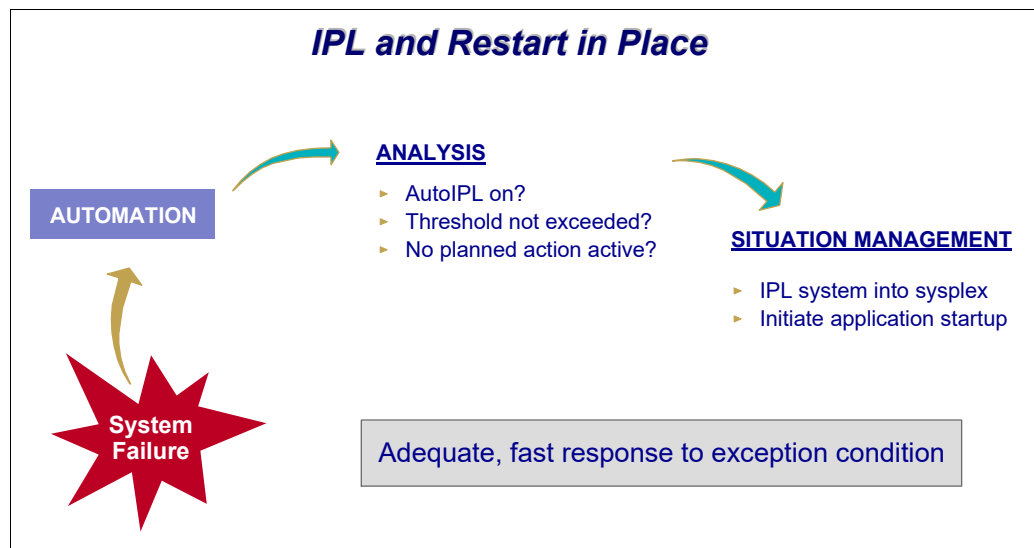
Figure 3-19   Recovering a failed image

Although Auto IPL processing occurs automatically based on policy and does not require a script, you can define CPC failure scripts to run specific actions other than restarting in place when one or more systems fail as part of a complete CPC failure. In such a script, you might want to activate backup partitions on another CPC for all the systems on the failing CPC, activate CBU if suitable, and IPL these systems on the alternative CPC. You can have one such script that is defined in advance for every server in your configuration.

For example, if you have a CPC that is named CECA1, you can define a script that is named CECFAIL_CECA1. If GDPS Metro detects a complete failure of CECA1, GDPS Metro automatically runs the script that is named CECFAIL_CECA1 to handle the unplanned event.

Scripts for other unplanned events

The following sections describe other unplanned events for which you might want to define actions for GDPS Metro to run. In these cases, GDPS Metro does *not* automatically run the script after the event occurs. Rather, these scripts must be run from the panels or the GUI; therefore, they are considered *control scripts*.

STP CTN role reassignments: Unplanned failure

If a failure condition results in the PTS, BTS, or Arbiter no longer being an operational synchronized CPC in the CTN, a suggestion is that after the failure and possible STP recovery action, the STP roles be reassigned to operational CPCs in the CTN. The reassignment reduces the potential for a sysplex outage in a second failure or planned action affects one of the remaining special role CPCs.

The script statement capability that is described in "STP CTN role reassignments: Planned operations" on page 94 can be used to integrate the STP role reassignment as part of an existing script and eliminate the requirement for the operator to perform the STP reconfiguration task manually at the HMC.

STP WTOR IEA394A response: Unplanned failure

As described in "Improved controlling system availability: Enhanced timer support" on page 62, a loss of synchronization with the sysplex timing source generates a disabled console WTOR. This suspends all processing on the LPAR until a response to the WTOR is provided. The WTOR message is IEA394A if the CPC is in STP timing mode (either in an STP Mixed CTN or STP-only CTN).

GDPS, by using scripts, can reply (either ABORT or RETRY) to the IEA394A sync WTOR for STP on systems that are spinning because of a loss of synchronization with their CTS. As described in "Automated response to STP sync WTORs" on page 63, autonomic function exists to reply RETRY automatically for 60 minutes on any GDPS systems that posted this WTOR.

The script statement complements and extends this function, as described:

► It provides the means to reply to the message after the 60-minute automatic reply window expires.

► It can reply to the WTOR on systems that are not GDPS systems (foreign systems) that are defined to GDPS; the autonomic function replies only on GDPS systems.

► It can reply ABORT on any systems you do not want to restart for a failure scenario before reconfiguration and synchronization of STP.

Batch scripts

GDPS also provides a flexible batch interface to start scripts from outside of GDPS. These scripts are known as *batch scripts* and can be started in the following ways:

► As a REXX program from a user terminal
► By using the IBM MVS MODIFY command to the NetView task
► From timers in NetView
► Triggered through the SA automation tables

This capability, along with the Query Services interface that is described in 3.7.4, "Concurrent Copy cleanup" on page 103, provides a rich framework for user-customizable systems management procedures.

3.5.3 Application programming interfaces

GDPS provides two primary programming interfaces to allow other programs that are written by clients, independent software vendors (ISVs), and other IBM product areas to communicate with GDPS. These APIs allow clients, ISVs, and other IBM product areas to complement GDPS automation with their own automation code. The following sections describe the APIs provided by GDPS.

Query Services

GDPS maintains configuration information and status information in NetView variables for the various elements of the configuration that it manages. Query Services is a capability that allows NetView REXX programs to query the values for numerous GDPS internal variables. The variables that can be queried pertain to the Metro Mirror configuration, the system and sysplex resources that are managed by GDPS, and other GDPS facilities, such as HyperSwap and the GDPS Monitors.

In addition to the Query Services function that is part of the base GDPS product, GDPS provides several samples in the GDPS SAMPLIB library to demonstrate how Query Services can be used in client-written code.

GDPS also makes available to clients a sample tool called the Preserve Mirror Tool (PMT), which facilitates adding new disks to the GDPS Metro Mirror configuration and bringing these disks to duplex. The PMT tool, which is provided in source format, makes extensive use of GDPS Query Services and thus provides clients with an excellent example of how to write programs to benefit from Query Services.

RESTful APIs

As described in "Query Services" on page 97, GDPS maintains configuration information and status information about the various elements of the configuration that it manages. Query Services can be used by REXX programs to query this information.

The GDPS RESTful API also provides the ability for programs to query this information. Because it is a RESTful API, it can be used by programs that are written in various programming languages, including REXX that are running on various server platforms.

In addition to querying information about the GDPS environment, the GDPS RESTful API allows programs that are written by clients, ISVs, and other IBM product areas to start actions against various elements of the GDPS environment. Examples of these actions include starting and stopping Metro Mirror, starting and stopping systems, managing sysplex resources, and starting GDPS monitor processing. These capabilities enable clients, ISVs, and other IBM product areas to provide an even richer set of functions to complement the GDPS functions.

GDPS provides samples in the GDPS SAMPLIB library to demonstrate how the GDPS RESTful API can be used in programs.

3.5.4 Additional system management information

Most of the GDPS Standard Actions and several script commands require GDPS to communicate with the HMC. The interface GDPS uses to communicate with the HMC is called the BCPii. This interface allows GDPS to automate many of the HMC actions, such as Load, Stop (graceful shutdown), Reset, Activate LPAR, and Deactivate LPAR. GDPS can also perform ACTIVATE (POR), CBU ACTIVATE/UNDO, OOCoD ACTIVATE/UNDO, and STP role reassignment actions against an HMC object that represents a CPC.

The GDPS LOAD and RESET Standard Actions (available through the GUI and NetView Standard Actions panels, the SYSPLEX script statement, and the RESTful APIs) allow specification of a CLEAR or NOCLEAR operand. This configuration provides operational flexibility to accommodate client procedures, which eliminate the requirement to use the HMC to perform specific LOAD and RESET actions.

Also, when you LOAD a system by using GDPS (by way of the GUI, NetView panels, scripts, or RESTful APIs), GDPS can listen for operator prompts from the system being started and reply to such prompts. GDPS optionally replies to such IPL-time prompts automatically, removing reliance on operator skills and eliminating operator error for selected messages that require replies.

SYSRES Management

Today many clients maintain multiple alternative z/OS SYSRES devices (also known as *IPLSET*s) as part of their maintenance methodology. GDPS provides special support to allow clients to identify IPLSETs. This support removes the requirement for clients to manage and maintain their own procedures when starting a system on a different alternative SYSRES device.

GDPS can automatically update the IPL pointers after any disk switch or disk recovery action that changes the GDPS primary disk location indicator for Metro Mirror disks. This update removes the requirement for clients to perform extra script actions to switch IPL pointers after disk switches, and greatly simplifies operations for managing alternative SYSRES "sets."

3.5.5 Securing the GDPS environment

GDPS uses RACF XFACILIT and GXFACILI resource classes to create a role-based security model for controlling access to the resources in your GDPS environment that is customized to your specific environment. Simple definitions can be used to control access at the panel option level or more granular definitions can be used to control access to specific types of resources, or even all the way down to the specific resource level.

With the role-based security model, you can create your own roles or you can use the common roles that GDPS recommends that include GDPS Administrator, GDPS Operator, GDPS User, and Non-GDPS User. You define the resources that these roles can access and the type of access they have to those resources by granting them access to the resource profiles that represent the various resources in your environment. Finally, you can grant access to various resources to users by adding them to the appropriate roles.

When you use the role-based security model, GDPS ensures that the user has sufficient authority to take a specific action against a specific resource, regardless of whether they are attempting to act by using the panels directly or by running a GDPS script.

Finally, the GDPS Security Definition Utility tool is available to assist you with implementing your role-based security environment. For more information, see the GDPS Security Definition Utility guide.

A detailed GDPS security checklist is also available in the *GDPS Metro 4.7 Installation and Customization Guide* ZG24-6755

3.5.6 Reconfiguring the remote copy environment

The GDPS Reconfiguration wizard can perform complex operations while making modifications to your remote copy environment. Provide some initial information, and then it automates the major required tasks:

► Configuring new storage servers
► Initiating and managing mirroring to new storage servers
► Updating the GDPS configuration, including the DASD configuration (GEOPARM)

This function aids with the replacement of DS8000 auxiliary storage subsystems that are managed by GDPS.

The Reconfiguration Wizard is started from the GDPS GUI by using the cog icon (Figure 3-20).

Figure 3-20   Selecting the GDPS Reconfiguration wizard

Operations

Reconfiguration Wizard operations consist of a validation phase and an execution phase. In the validation phase, the Reconfiguration Wizard determines the changes that are required to perform the operation and validates that there are no conditions that preclude the running of the operation.

Then, the Reconfiguration Wizard provides a change report that details the changes that are made as a part of the operation and waits for permission from the user to proceed with the operation. Now, the change report is used to obtain authorization to proceed with the operation from the appropriate change control personnel.

Throughout the Reconfiguration Wizard, you can access help information by clicking the **Help** tab at the top of the window.

The user can disconnect from the reconfiguration operation now without ending the operation.

The execution phase begins when the user instructs the Reconfiguration Wizard to continue. During the execution phase, the changes to the environment to perform the reconfiguration are made.

3.6 GDPS Metro monitoring and alerting

The GDPS SDF panel, discussed in "Monitoring function: Status Display Facility" on page 86, is where GDPS dynamically displays color-coded alerts.

Alerts can be posted as a result of an unsolicited error situation that GDPS listens for. For example, if one of the multiple PPRC links that provide the path over which Metro Mirror operations take place is broken, there is an unsolicited error message issued. GDPS listens for this condition and raises an alert on the SDF panel, notifying the operator of the fact that a PPRC link is not operational. Clients run with multiple PPRC links and if one is broken, Metro Mirror continues over any remaining links. However, it is important for operations to be aware that a link is broken and fix this situation because a reduced number of links results in reduced Metro Mirror bandwidth and reduced redundancy. If this problem is not fixed in a timely manner and more links fail, it can result in production impact because of insufficient mirroring bandwidth or total loss of Metro Mirror connectivity (which results in a freeze).

Alerts can also be posted as a result of GDPS periodically monitoring key resources and indicators that relate to the GDPS Metro environment. If any of these monitoring items are found to be in a state that is deemed to be *not normal* by GDPS, an alert is posted on SDF.

Various GDPS monitoring functions are ran on the GDPS controlling systems and on the production systems because from a software perspective, it is possible that different production systems have different views of some of the resources in the environment, and although status can be normal in one production system, it can be not normal in another. All GDPS alerts that are generated on one system in the GDPS sysplex are propagated to all other systems in the GDPS. This propagation of alerts provides for a single focal point of control. It is sufficient for the operator to monitor SDF on the master controlling system to be aware of all alerts generated in the entire GDPS complex.

When an alert is posted, the operator must investigate (or escalate) and corrective action must be taken for the reported problem as soon as possible. After the problem is corrected, it is detected during the next monitoring cycle and the alert is cleared by GDPS automatically.

GDPS Metro monitoring and alerting capability is intended to ensure that operations are notified of and can take corrective action for any problems in their environment that can affect the ability of GDPS Metro to do recovery operations. This approach maximizes the chance of achieving your availability and RPO/RTO commitments.

3.6.1 GDPS Metro health checks

In addition to the GDPS Metro monitoring described, GDPS provides health checks. These health checks are provided as a plug-in to the z/OS Health Checker infrastructure to check that certain settings related to GDPS adhere to best practices.

The z/OS Health Checker infrastructure is intended to check various settings to determine whether these settings adhere to z/OS optimum values. For settings found to be not in line with best practices, exceptions are raised in the Spool Display and Search Facility (SDSF) and optionally, SDF alerts are also raised. If these settings do not adhere to recommendations, this issue can hamper the ability of GDPS to perform critical functions in a timely manner.

Often, if there are changes in the client environment, they might necessitate adjustment of various parameter settings that are associated with z/OS, GDPS, and other products. It is possible that you can miss making these adjustments, which can affect GDPS. The GDPS health checks are intended to detect such situations and avoid incidents where GDPS is unable to perform its job because of a setting that is less than ideal.

For example, GDPS Metro provides facilities for management of the CDS for the GDPS sysplex. One of the health checks provided by GDPS Metro checks that the CDS are allocated and defined to GDPS in line with the GDPS best practices.

Similar to z/OS and other products that provide health checks, GDPS health checks are optional. Several optimum values that are checked and the frequency of the checks can be customized to cater to unique client environments and requirements.

There are a few z/OS best practices that conflict with GDPS best practices. The related z/OS and GDPS health checks result in conflicting exceptions being raised. For such health check items, to avoid conflicting exceptions, z/OS defines a *coexistence policy* where you can indicate which practice is to take precedence; GDPS or z/OS. GDPS provides sample coexistence policy definitions for the GDPS checks that are known to be conflicting with z/OS.

GDPS also provides a convenient interface for managing the health checks by using the GDPS panels (a similar interface is available by using the GDPS GUI). You can use it to perform actions such as activate/deactivate or run any selected health check, view the customer overrides in effect for any optimum values, and other actions.

Figure 3-21 shows a sample of the GDPS Health Checks Information Management panel. In this example, you see that all the health checks are enabled. The status of the last run is also shown, which indicates that some were successful and some resulted in raising a medium exception. The exceptions can also be viewed by using other options on the panel.

```
VPC8PHC0              GDPS Health Checks Information        A6P92   G9C2       (1/15)

HealthChecker:  STARTED   Procname: HZSPROC     Policy Name: DEFAULT

Actions:   S elect   R un      A ctivate  D eactivate  P urge  I nstall
           Q uery    QP olicy PR int
Cmd Check name                          State              Status             Policy
                                                                              Stmnt
_   GDPS_CHECK_DEVICE                   ACTIVE(ENABLED)    EXCEPTION-HIGH     No
_   GDPS_CHECK_SPOF                     ACTIVE(ENABLED)    EXCEPTION-MEDIUM   No
_   GDPS_CHECK_GRS                      ACTIVE(ENABLED)    EXCEPTION-MEDIUM   No
_   GDPS_CHECK_XCF_CDS                  ACTIVE(ENABLED)    EXCEPTION-MEDIUM   No
_   GDPS_CHECK_CONSOLE                  ACTIVE(ENABLED)    EXCEPTION-MEDIUM   No
_   GDPS_CHECK_K_SYS_LPAR               ACTIVE(ENABLED)    SUCCESSFUL         No
_   GDPS_CHECK_CONFIG                   ACTIVE(ENABLED)    SUCCESSFUL         No
_   GDPS_CHECK_XCF                      ACTIVE(ENABLED)    SUCCESSFUL         No
_   GDPS_CHECK_MAXSYS                   ACTIVE(ENABLED)    SUCCESSFUL         No
_   GDPS_CHECK_DASDMIH                  ACTIVE(ENABLED)    SUCCESSFUL         No


Commands: 1 Run     2 Activate   3 Deactivate 4 Purge    5 Install   6 Summary
          7 GEOHCP 8 Parameter 9 Exit List 10 Policy 11 Remote
Selection ===>  _
F1=Help  F3=Return  F5=Refresh  F6=Roll        F8=Down  F10=Left  F11=Right
```

Figure 3-21   GDPS Metro Health Checks Information Management panel (VPC8PHC0)

3.7 Other facilities that are related to GDPS

Miscellaneous facilities that GDPS Metro provides can help in various ways, such as reducing the window during which DR capability is not available.

3.7.1 HyperSwap coexistence

In the following sections, we discuss the GDPS enhancements that remove some of the restrictions that existed regarding HyperSwap coexistence with products such as Softek Transparent Data Migration Facility (TDMF) and IMS Extended Recovery Facility (XRF).

HyperSwap and TDMF coexistence

To minimize disruption to production workloads and service levels, many enterprises use TDMF for storage subsystem migrations and other disk relocation activities. The migration process is transparent to the application, and the data is continuously available for read and write activities throughout the migration process.

However, the HyperSwap function is mutually exclusive with software that moves volumes around by switching UCB pointers. The currently supported versions of TDMF and GDPS allow operational coexistence. With this support, TDMF automatically temporarily disables HyperSwap as part of the disk migration process only during the brief time when it switches UCB pointers.

Manual operator interaction is not required. Without this support, through operator intervention, HyperSwap is disabled for the entire disk migration, including the lengthy data copy phase.

HyperSwap and IMS XRF coexistence

HyperSwap also has a technical requirement that RESERVEs cannot be allowed in the hardware because the status cannot be reliably propagated by z/OS during the HyperSwap to the new primary volumes. For HyperSwap, all RESERVEs must be converted to GRS global enqueue through the GRS RNL lists.

IMS/XRF is a facility by which IMS can provide one active subsystem for transaction processing, and a backup subsystem that is ready to take over the workload. IMS/XRF issues hardware RESERVE commands during takeover processing, which cannot be converted to global enqueues through GRS RNL processing. This coexistence problem also is resolved so that GDPS is informed before IMS issuing the hardware RESERVE, allowing it to automatically disable HyperSwap. After IMS finishes processing and releases the hardware RESERVE, GDPS is again informed and re-enables HyperSwap.

3.7.2 Reduced impact on initial copy and resynchronization

Performing Metro Mirror copy of a large amount of data across many devices while the same devices are used in production by application workloads can potentially affect production I/O service times if such copy operations are performed synchronously. Your disk subsystems and PPRC link capacity are typically sized for steady state update activity, but not for bulk, synchronous replication. Initial copying of disks and resynchronization of disks are examples of bulk copy operations that can affect production if performed synchronously.

There is no need to perform initial copy or resynchronizations by using synchronous copy because the secondary disks cannot be made consistent until all disks in the configuration reach the duplex state.

GDPS supports initial copy and resynchronization by using asynchronous PPRC-XD (also known as Global Copy). When GDPS initiates copy operations in asynchronous copy mode, GDPS monitors progress of the copy operation and when the volumes are near full duplex state, GDPS converts the replication from the asynchronous copy mode to synchronous PPRC (Metro Mirror). Initial copy or resynchronization by using PPRC-XD eliminates the performance impact of synchronous mirroring on production workloads.

Without asynchronous copy, it might be necessary to defer these operations or reduce the number of volumes being copied at any time. This approach delays the mirror from reaching a duplex state, thus impacting a client's ability to recovery. Using the XD-mode asynchronous copy allows clients to establish or resynchronize mirroring during periods of high production workload, and can potentially reduce the time during which the configuration is exposed.

This function requires that all disk subsystems in the GDPS configuration support PPRC-XD.

3.7.3 Reserve Storage Pool

Reserve Storage Pool (RSP) is a type of resource that was introduced with the z/OS Management Facility (z/OSMF) that can simplify the management of defined but unused volumes. GDPS supports including RSP volumes in the Metro Mirror configuration that is managed by GDPS. Metro Mirror primary volumes are expected to be online in controlling systems, and GDPS monitoring on the GDPS controlling systems results in an alert being raised for any Metro Mirror primary device that is found to be offline. However, because z/OS does not allow RSP volumes to be brought online to any system, GDPS monitoring recognizes that an offline primary device is an RSP volume and suppresses alerting for these volumes.

3.7.4 Concurrent Copy cleanup

The DFSMS Concurrent Copy (CC) function uses a "sidefile" that is kept in the disk subsystem cache to maintain a copy of changed tracks that are not yet copied. For a Metro Mirrored disk, this sidefile is not mirrored to the secondary subsystem. If a HyperSwap runs while a CC operation is in progress, the application that uses CC will fail after the completion of the HyperSwap. GDPS does not allow a planned swap when a CC session exists against your primary Metro Mirror devices. However, unplanned swaps still are allowed. Therefore, if you plan to use HyperSwap for primary disk subsystem failures (unplanned HyperSwap), try to eliminate any use of CC, because you cannot plan when a failure occurs.

Checking for CC is performed by GDPS immediately before performing a planned HyperSwap. SDF trace entries are generated if one or more CC sessions exist and the swap command ends with no Metro Mirror device pairs being swapped. Identify and terminate any CC against the Metro Mirror primary devices before the swap.

When attempting to resynchronize your disks, checking is performed to ensure that the secondary devices do not retain CC status from the time when they were primary devices. They are not supported as Metro Mirror secondary devices. Therefore, GDPS will not attempt to establish a duplex pair with secondary devices if it detects a CC session.

GDPS provides a function to discover and terminate CC sessions that would otherwise cause errors during a resync operation. The function is controlled by a keyword that provides options to disable, to conditionally enable, or to unconditionally enable the cleanup of CC sessions on the target disks. This capability eliminates the manual task of identifying and cleaning up orphaned CC sessions before resynchronizing a suspended Metro Mirror relationship.

3.7.5 Easy Tier Heat Map Transfer

IBM DS8000 Easy Tier optimizes data of logical volumes across the various physical tiers of storage within a disk subsystem to optimize application performance. The placement decisions are based on learning the data access patterns and can be changed dynamically and transparently to the applications by using this data.

Metro Mirror replicates the data from the primary to the secondary disk subsystem; however, the Easy Tier learning information is not included in Metro Mirror scope. The secondary disk subsystems are optimized according to the workload on these subsystems, which is different than the activity on the primary (there is only a write workload on the secondary whereas there is read/write activity on the primary). As a result of this difference, during a disk switch or disk recovery, the secondary disks that you switch to are likely to display different performance characteristics compared to the former primary.

Easy Tier Heat Map Transfer is the DS8000 capability to transfer the Easy Tier learning from a Metro Mirror primary to the secondary disk subsystems so that the secondary disk subsystems can also be optimized (based on this learning) and has similar performance characteristics if it is promoted to become the primary.

GDPS integrates support for Heat Map Transfer. In a dual-leg configuration, Heat Map Transfer is established for both secondary targets. The Heat Map Transfer actions (such as start/stop of the processing and reversing transfer direction) are incorporated into the GDPS managed processes. For example, if Metro Mirror is temporarily suspended on a leg by GDPS for a planned or unplanned secondary disk outage, Heat Map Transfer is also suspended on that leg, or if Metro Mirror direction is reversed as a result of a HyperSwap, Heat Map Transfer direction is also reversed.

3.7.6 Autonomous Path Reconfiguration

The GDPS Autonomous Path Reconfiguration feature circumvents potential problems during the run of a START SECONDARY command, which is used to establish PPRC paths for disk replication (asynchronous and synchronous).

In a configuration where this feature is not enabled, a START SECONDARY command stops immediately if one of the PPRC links defined to GDPS is not operational.

With this feature enabled, if GDPS detects a failure while attempting to establish PPRC paths, it tries to collect information from the path in error and remove the bad links from the configuration, and then re-attempts to establish the path. If GDPS is not successful in establishing the PPRC paths after the autonomous reconfiguration, it ends the START SECONDARY command with RC=8 and generates a detailed report in the NetView log for investigation.

3.8 Flexible testing, resync protection, and Logical Corruption Protection

Configuring point-in-time (PiT) copy (FlashCopy or Safeguarded Copy (SGC)) capacity in your GDPS Metro environment provides several significant benefits:

► You can conduct regular DR drills or other tests by using a copy of production data while production continues to run.

► You can save a consistent, "golden" copy of the Metro Mirror secondary data, which can be used if the primary disk or site is lost during a Metro Mirror resynchronization operation.

► You can save multiple consistent copies of all or a subset of the devices in your configuration, at different points in time that can be used to recover from logical corruption events including cyberattacks and internal attacks.

FlashCopy and the various options that are related to FlashCopy are discussed in 2.6, "FlashCopy" on page 32. GDPS Metro supports taking a FlashCopy of the current primary or either of the current secondary disks sets. The COPY, NOCOPY, NOCOPY2COPY, and INCREMENTAL options are supported. CONSISTENT FlashCopy is supported in conjunction with COPY, NOCOPY, and INCREMENTAL FlashCopy.

FlashCopy can also be used, for example, to back up data without the need for extended outages to production systems, and to provide data for data mining applications and batch reporting.

GDPS Metro uses the FlashCopy technology and the SGC technology to provide a powerful solution for protecting against various types of logical data corruption, including cyberattacks and internal threats. This capability is referred to as Logical Corruption Protection (LCP). For more information about LCP, see Chapter 9, "IBM GDPS Logical Corruption Protection and Testcopy Manager" on page 255.

3.8.1 Using space-efficient FlashCopy volumes

As discussed in "Space-efficient FlashCopy" on page 34, by using space-efficient volumes, you might be able to lower the amount of physical storage needed and thus reduce the cost associated with providing a tertiary copy of the data. GDPS provides support allowing FlashCopy SE volumes to be used as FlashCopy target disk volumes. Whether a target device is space-efficient or not is transparent to GDPS; if any of the FlashCopy target devices defined to GDPS are space-efficient volumes, GDPS uses them. All GDPS FlashCopy operations, whether through GDPS scripts, the GUI, panels, or FlashCopies automatically taken by GDPS, can use space-efficient targets.

space-efficient volumes are ideally suited for FlashCopy targets when used for resync protection. The FlashCopy is taken before the resync and can be withdrawn when the resync operation is complete. As changed tracks are sent to the secondary for resync, the time-zero (T0) copy of this data is moved from the secondary to the FlashCopy target device. This means that the total space requirement for the targets is equal to the number of tracks that were out of sync, which typically are less than a full set of fully provisioned disks.

Another potential use of space-efficient volumes is if you want to use the data for limited DR testing.

Understanding the characteristics of FlashCopy SE is important to determine whether this method of creating a PiT copy satisfies your business requirements. For example, will it be acceptable to your business if, because of an unexpected workload condition, the extent pool on the disk subsystem for the space-efficient devices becomes full and your FlashCopy is invalidated so that you are unable to use it? If your business requirements dictate that the copy must always be guaranteed to be usable, space-efficient might not be the best option and you can consider using standard FlashCopy instead.

3.9 GDPS tools for GDPS Metro

GDPS Metro includes tools that provide a function that is complementary to the GDPS function. The tools represent a function that all or many clients are likely to develop themselves to complement GDPS. The use of these tools eliminates the need for you to develop similar functions yourself. The tools are provided in source code format, which means that you can modify the code to suit your needs if the tool does not meet your requirements exactly.

The following tools are available with GDPS Metro:

► GDPS Console Interface Tool (also known as and referred to as GCI)

This tool facilitates the use of the MVS system console as an interface for submitting GDPS scripts for execution or running script commands.

► PMT

This tool simplifies and automates the process of adding PPRC disk devices to a running GDPS environment. PMT is designed to minimize the time during which GDPS mirroring status is Not OK (NOK), which maximizes the duration a duplex mirror (along with Freeze and HyperSwap capability) is preserved.

► GDPS XML Conversion (GeoXML) Tool

This tool helps you to convert a GDPS/PPRC GEOPARM configuration definition file for a single replication leg to GDPS Metro XML-format GEOPARM definitions. This process simplifies the task of defining the GDPS Metro configuration for GDPS/PPRC clients who are moving to the use of GDPS Metro.

► GDPS EasyLog Tool

This Microsoft Windows-based tool helps you to extract and easily download the MVS Syslog and NetView logs from a z/OS environment. It also helps in analyzing the Netlog after it is downloaded to a workstation.

► GDPS Security Definition Utility tool

The GDPS Security Definition Utility tool helps you to implement enhancements to GDPS options that use role-based security by helping with the definition of appropriate profiles in the XFACILIT class of RACF, and by assigning relevant access to them.

3.10 GDPS Metro co-operation with GDPS Continuous Availability

GDPS Metro provides facilities for co-operation with GDPS Continuous Availability if GDPS Continuous Availability is used to provide workload level protection for selected workloads that are running on the systems that are in the GDPS Metro sysplex. See 6.5, "GDPS Continuous Availability co-operation with GDPS Metro" on page 205 for details.

3.11 Services component

GDPS provides more than remote copy management. It also includes system, server hardware and sysplex management, automation, testing processes, and DR processes.

Most installations do not have skills in all these areas readily available. It is also rare to find a team that has this range of skills across many implementations. However, the GDPS Metro offering includes exactly that: access to a global team of specialists in all the disciplines you need to ensure a successful GDPS Metro implementation.

Specifically, the Services component includes several or all the following services:

► Planning to determine availability requirements, configuration recommendations, and implementation and testing plans

► Installation and necessary customization of NetView and System Automation

► Remote copy implementation

► GDPS Metro automation code installation and policy customization

► Assistance in defining RPOs and RTOs

► Education and training on GDPS Metro setup and operations

► Onsite implementation assistance

► Project management and support throughout the engagement

The sizing of the Services component of each project is tailored for that project, based on many factors including what automation is already in place, whether remote copy is already in place, and whether the two centers are already in place with a multisite sysplex. This situation means that the skills provided are tailored to the specific needs of each particular implementation.

3.12 GDPS Metro prerequisites

For more information about the latest GDPS Metro prerequisites, see this web page.

3.13 Comparing GDPS Metro versus other GDPS offerings

So many features and functions are available in the various members of the GDPS family that recalling them all and remembering which offerings support them is sometimes difficult. To position the offerings, Table 3-1 lists the key features and functions and indicates those features and functions that are delivered by the various GDPS offerings.

Table 3-1   Supported features matrix

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
CA	Yes	Yes	Yes	No
DR	Yes	Yes	Yes	Yes
CA/DR protection against multiple failures	Yes	No	No	No
CA for foreign z/OS systems	Yes with z/OS proxy	No	No	No
Supported distance	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	Virtually unlimited
Consistent FlashCopy support	Yes, using CONSISTENT	Yes, using CONSISTENT for secondary only	No	Yes, using CGPause
Reduced impact on initial copy/resync	Yes	Yes	Yes	Not applicable
Tape replication support	Yes	No	No	No
Production sysplex automation	Yes	No	Not applicable	No
Span of control	Both sites	Both sites (disk only)	Both sites	Disk at both sites, and the recovery site (CBU or LPARs)
GDPS scripting	Yes	No	Yes	Yes
Monitoring, alerting and health checks	Yes	Yes	Yes (except health checks)	Yes
Query Services	Yes	Yes	No	Yes
MSS support for added scalability	Yes (RS2 in MSS1, RS3 in MSS2)	Yes (secondary in MSS1)	No	Yes (Global Mirror (GM) FlashCopy and Primary for Metro Global Mirror (MGM) in MSS1)
MGM 3-site and 4-site	Yes (all configurations)	Yes (3-site only and non-IR only)	No	Yes (all configurations)
FB disk	Yes	Yes	No	Yes

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
z/OS equivalent function for Linux on IBM Z	Yes (Linux on IBM Z running as a z/VM guest only)	No	Yes (Linux on IBM Z running as a z/VM guest only)	Yes
GDPS GUI	Yes	Yes	Yes	Yes

3.14 Summary

GDPS Metro is a powerful offering that provides DR, CA, and system/sysplex resource management capabilities. HyperSwap, available with GDPS Metro, transparently swaps disks between disk locations. The power of automation tests and perfects the actions to be taken, either for planned or unplanned changes, thus minimizing or eliminating the risk of human error.

This offering is one of the offerings in the GDPS family, along with GDPS Metro HyperSwap Manager, and GDPS Virtual Appliance that offers the potential of ZDL, and that can achieve the shortest RTO, typically less than 1 hour after a complete site failure.

It is also one of the only members of the GDPS family, along with GDPS Virtual Appliance that is based on hardware replication *and* provides the capability to manage the production LPARs. Although GDPS GM offers LPAR management, the scope of system management includes only the systems in the recovery site, and not the production systems running in Site1.

GDPS Metro, in a dual-leg configuration, is the only GDPS offering that can provide zero-data-loss DR protection, even *after* a primary disk failure.

In addition to the DR and planned reconfiguration capabilities, GDPS Metro also provides simple interfaces for monitoring and managing the various elements of the GDPS configuration.

4

IBM GDPS Metro HyperSwap Manager

In this chapter, we discuss the capabilities and prerequisites of the GDPS Metro HyperSwap Manager (GDPS HM) offering.

GDPS HM extends the availability attributes of a Parallel Sysplex to disk subsystems, whether the Parallel Sysplex and disk subsystems are in a single site, or whether the Parallel Sysplex and the primary/secondary disk subsystems span across two sites.

It transparently switches primary disk subsystems with the secondary disk subsystems for either a planned or unplanned disk reconfiguration. It also supports disaster recovery (DR) capability across two sites by enabling the creation of a consistent set of secondary disks in a disaster or potential disaster.

However, unlike the full IBM Geographically Dispersed Parallel Sysplex (GDPS) Metro offering, GDPS HM does not provide any resource management or recovery management capabilities.

The following functions are for protecting data that is provided by GDPS HM:

► Ensuring the consistency of the secondary data in case there is a disaster or suspected disaster, including the option to also ensure zero data loss (ZDL)

► Switching to the secondary disk by using HyperSwap

► Managing the remote copy configuration for IBM Z and other platform data

Because GDPS HM is a subset of the GDPS Metro offering, you might want to review the comparison that is presented in Table 4-1 on page 144 if you read Chapter 3, "IBM GDPS Metro" on page 45.

This chapter includes the following topics:

4.1 Introducing GDPS HM

GDPS HM provides a subset of GDPS Metro capability with the emphasis more on the remote copy and disk management aspects. At its most basic, GDPS HM extends Parallel Sysplex availability to disk subsystems by delivering the HyperSwap capability to mask disk outages that are caused by planned disk maintenance or unplanned disk failures. It also provides monitoring and management of the data replication environment, including the freeze capability.

In the multisite environment, GDPS HM provides an entry-level DR offering. Because GDPS HM does not include the systems management and automation capabilities of GDPS Metro, it cannot provide in and of itself the short RTO that is achievable with GDPS Metro. However, GDPS HM does provide a cost-effective route into full GDPS Metro later if your recovery time objectives (RTOs) change.

4.1.1 Protecting data integrity and data availability with GDPS HM

In 2.2, "Data consistency" on page 18, we point out that data integrity across primary and secondary volumes of data is essential to perform a database restart and accomplish an RTO of less than an hour. This section provides details about how GDPS automation in GDPS HM provides both data consistency if there are mirroring problems and data availability if there are disk problems.

The following types of disk problems trigger a GDPS automated reaction:

► Metro Mirror replication problems (Freeze triggers)

There is no problem with writing to the primary disk subsystem, but there is a problem mirroring the data to the secondary disk subsystem. For more information, see "GDPS Freeze function for mirroring failures" on page 113."

► Primary disk problems (HyperSwap triggers)

There is a problem writing to the primary disk: Either a hard failure, or the disk subsystem is not accessible or not responsive. For more information, see "GDPS HyperSwap function" on page 117.

GDPS Freeze function for mirroring failures

GDPS uses automation, which is keyed off events or messages, to stop all mirroring when a remote copy failure occurs. In particular, the GDPS automation uses the IBM PPRC Freeze and Run architecture, which was implemented as part of Metro Mirror on IBM disk subsystems and also by other enterprise disk vendors. In this way, if the disk hardware supports the Freeze/Run architecture, GDPS can ensure consistency across all data in the sysplex (consistency group) regardless of disk hardware type. This preferred approach differs from proprietary hardware approaches that work only for one type of disk hardware. For more information about data consistency with synchronous disk mirroring, see "Metro Mirror data consistency" on page 24.

When a mirroring failure occurs, this problem is classified as a Freeze trigger and GDPS stops activity across *all* disk subsystems at the time the initial failure is detected, thus ensuring that the dependent write consistency of the remote disks is maintained. This result is what happens when a GDPS performs a Freeze:

► Remote copy is suspended for all device pairs in the configuration.

► While the suspend command is processed for each logical subsystem (LSS), each device goes into a long busy state. When the suspend completes for each device, z/OS marks the device unit control block (UCB) in all connected operating systems to indicate an Extended Long Busy (ELB) state.

► No I/Os can be issued to the affected devices until the ELB is thawed with the PPRC Run (or "thaw") action or until it times out (the consistency group timer setting commonly defaults to 120 seconds, although for most configurations a longer ELB is a best practice).

► All paths between the Metro Mirrored disks are removed, preventing further I/O to the secondary disks if Metro Mirror is accidentally restarted.

Because no I/Os are processed for a remote-copied volume during the ELB, dependent write logic ensures the consistency of the remote disks. GDPS performs Freeze for all LSS pairs that contain GDPS managed mirrored devices.

> **Important:** Because of the dependent write logic, it is not necessary for all LSSs to be frozen at the same instant. In a large configuration with many thousands of remote copy pairs, it is not unusual to see short gaps between the times when the Freeze command is issued to each disk subsystem. However, because of the ELB, such gaps are not a problem.

After GDPS performs the Freeze and the consistency of the remote disks is protected, what GDPS does depends on the client's PPRC Failure policy (also known as Freeze policy). The policy, as described in "Freeze policy (PPRC Failure policy) options" on page 115, tells GDPS to take one of these three possible actions:

► Perform a **Run** action against all LSSs. This action removes the ELB and enables production systems to continue using these devices. The devices are in remote copy-suspended mode, meaning that any further writes to these devices are no longer being mirrored. However, the changes are tracked by the hardware so that, later, only the changed data is resynchronized to the secondary disks. For more information about this policy option, see "Freeze and Go" on page 115.

► System-reset all production systems. This action ensures that no more updates can occur to the primary disks because such updates are not mirrored and it is not possible to achieve recovery point objective (RPO) zero (ZDL) if a failure occurs (or if the original trigger was an indication of a catastrophic failure). For more information about this option, see "Freeze and Stop" on page 115.

► Try to determine whether the cause of the Metro Mirror suspension event was a permanent or temporary problem with any of the secondary disk subsystems in the GDPS configuration. If GDPS can determine that the Metro Mirror failure was caused by the secondary disk subsystem, it is not a potential indicator of a disaster in the primary site. In this case, GDPS performs a Run action and allows production to continue using the suspended primary devices. However, if the cause cannot be determined to be a secondary disk problem, GDPS resets all systems, which guarantee ZDL. For more information, see "Freeze and Stop conditionally" on page 116.

GDPS HM uses a combination of storage subsystem and sysplex triggers to automatically secure, at the first indication of a potential disaster, a data-consistent secondary site copy of your data using the Freeze function. In this way, the secondary copy of the data is preserved in a consistent state, even before production applications are aware of any issues. Ensuring the data consistency of the secondary copy ensures that a normal system restart can be performed instead of having to perform database management system (DBMS) forward recovery actions. This element is an essential design element of GDPS to minimize the time to recover the critical workloads if there is a disaster in the primary site.

You will appreciate why such a process must be automated. When a device suspends, there is not enough time to start a manual investigation process. The entire mirror must be frozen by stopping further I/O to it, and then allowing production to run with mirroring temporarily suspended, or stopping all systems to ensure ZDL based on the policy.

In summary, freeze is triggered as a result of a Metro Mirror suspension event for any primary disk in the GDPS configuration, that is, at the first sign of a duplex mirror that is going out of duplex state. When a device suspends, all attached systems are sent a State Change Interrupt (SCI). A message is issued in all of those systems and then each system must issue multiple I/Os to investigate the reason for the suspension event.

When GDPS performs a freeze, all primary devices in the Metro Mirror configuration suspend. This suspension results in significant SCI traffic and many messages in all systems. GDPS, with z/OS and microcode on the DS8000 disk subsystems, supports reporting suspensions in a summary message per LSS instead of at the individual device level. When compared to reporting suspensions on a per device basis, the Summary Event Notification for PPRC Suspends (PPRCSUM) dramatically reduces the message traffic and extraneous processing that is associated with PPRC suspension events and freeze processing.

Freeze policy (PPRC Failure policy) options

As described, when a mirroring failure is detected, GDPS automatically and unconditionally performs a Freeze to secure a consistent set of secondary volumes in case the mirroring failure might be the first indication of a site failure. Because the primary disks are in an ELB state as a result of the freeze and the production systems are locked out, GDPS must act. There is no time to interact with the operator on an event-by-event basis. The action must be taken immediately and is determined by a customer policy setting, namely the PPRC Failure policy option (also known as the Freeze policy option). GDPS will use this same policy setting after every Freeze event to determine what its next action should be. The options are listed here:

- ► `PPRCFAILURE=STOP` (Freeze and STOP)

  GDPS resets production systems while I/O is suspended.

- ► `PPRCFAILURE=GO` (Freeze and Go)

  GDPS allows production systems to continue operation after mirroring is suspended.

- ► `PPRCFAILURE=COND` (Freeze and Stop, conditionally)

  GDPS tries to determine whether a secondary disk caused the mirroring failure. If so, GDPS performs a Go. If not, GDPS performs a Stop.

Freeze and Stop

If your RPO is zero (that is, you cannot tolerate any data loss), you must select the Freeze and Stop policy to reset all production systems. With this setting, you can be assured that no updates are made to the primary volumes after the Freeze because all systems that can update the primary volumes are reset. You can choose to restart them when you see fit. For example, if this freeze was a false freeze (that is, a false alarm), then you can quickly resynchronize the mirror and restart the systems only after the mirror is duplexed.

If you are using duplexed coupling facility (CF) structures along with a Freeze and Stop policy, it might seem that you are guaranteed to be able to use the duplexed instance of your structures if you must recover and restart your workload with the frozen secondary copy of your disks. However, this result is not always the case. There can be rolling disaster scenarios where before, following, or during the freeze event, there is an interruption (perhaps a failure of CF duplexing links) that forces CFRM to drop out of duplexing.

There is no guarantee that the structure instance in the surviving site is the one that is kept. It is possible that CFRM keeps the instance in the site that is about to totally fail. In this case, there is not an instance of the structure in the site that survives the failure.

To summarize, with a Freeze and Stop policy, if there is a surviving, accessible instance of application-related CF structures, that instance is consistent with the frozen secondary disks. However, depending on the circumstances of the failure, even with structures that are duplexed across two sites you are not 100% guaranteed to have a surviving, accessible instance of the application structures. You must have the procedures in place to restart your workloads without the structures.

A Stop policy ensures no data loss. However, if this event was a false Freeze event, that is, it was a transient failure that did not necessitate recovery by using the frozen disks, it stops the systems unnecessarily.

Freeze and Go

If you can accept an RPO that is not *necessarily* zero, you might decide to let the production systems *continue* operation after the secondary volumes are protected by the Freeze. In this case, you use a Freeze and Go policy. With this policy, you avoid an unnecessary outage for a false freeze event, that is, if the trigger is a transient event.

However, if the trigger turns out to be the first sign of an actual disaster, you might continue operating for an amount of time before all systems fail. Any updates made to the primary volumes during this time are not replicated to the secondary disk, and therefore are lost. In addition, because the CF structures were updated after the secondary disks were frozen, the CF structure content is not consistent with the secondary disks. Therefore, the CF structures in either site cannot be used to restart workloads and log-based restart must be used when restarting applications.

This recovery is not a full forward recovery. It is forward recovery of any data such as Db2 group buffer pools (GBPs) that might have existed in a CF but might not have been written to disk yet. This recovery results in prolonged recovery times. The extent of this elongation depends on how much such data existed in the CFs then. With a Freeze and Go policy, you might consider tuning applications such as Db2 to harden such data on disk more frequently than otherwise.

Freeze and Go is a high availability (HA) option that avoids production outage for false Freeze events. However, it carries a potential for data loss.

Freeze and Stop conditionally

Field experience shows that most occurrences of freeze triggers are not necessarily the start of a rolling disaster, but are "false freeze" events, which do not necessitate recovery on the secondary disk. Examples of such events include connectivity problems to the secondary disks and secondary disk subsystem failure conditions.

With a `COND` (conditional) specification, the action that GDPS takes after it performs the Freeze is conditional. GDPS tries to determine whether the mirroring problem was as a result of a permanent or temporary secondary disk subsystem problem:

► If GDPS can determine that the freeze was triggered as a result of a secondary disk subsystem problem, then GDPS performs a Go. That is, it allows production systems to continue to run by using the primary disks. However, updates are not mirrored until the secondary disk can be fixed and Metro Mirror can be resynchronized.

► If GDPS cannot find out that the cause of the freeze was a secondary disk subsystem, then GDPS deduces that it might be the beginning of a rolling disaster in the primary site. Therefore, it performs a Stop, resetting all production systems to ensure ZDL. GDPS cannot always detect that a particular freeze trigger was caused by a secondary disk, and some freeze events that are truly caused by a secondary disk might still result in a Stop.

For GDPS to determine whether a freeze trigger might have been caused by the secondary disk subsystem, the IBM DS8000 disk subsystems provide a special query capability that is known as the *Query Storage Controller Status* microcode function. If all disk subsystems in the GDPS managed configuration support this feature, GDPS uses this special function to query the secondary disk subsystems in the configuration to understand the state of the secondaries and whether one of those secondaries might have caused the freeze. If you use the COND policy setting but all disks your configuration do not support this function, then GDPS cannot query the secondary disk subsystems, and the resulting action is a Stop.

This option might provide a useful compromise where you can minimize the chance that systems stop for a false freeze event, and increase the chance of achieving ZDL for a real disaster event.

PPRC Failure policy selection considerations

As described, the PPRC Failure policy option specification directly relates to Recovery Time and RPOs, which are business objectives. Therefore, the policy option selection is really a business decision, rather than an IT decision. If data associated with your transactions is of high value, it might be more important to ensure that no data that is associated with your transactions is ever lost, so you might decide on a Freeze and Stop policy. If you have huge volumes of relatively low value transactions, you might be willing to risk some lost data in return for avoiding unneeded outages with a Freeze and Go policy. The Freeze and Stop Conditional policy attempts to minimize the chance of unnecessary outages and the chance of data loss, but there is still a risk of either outcome.

Most installations start with a Freeze and Go policy. Companies that have an RPO of zero typically then move on and implement a Freeze and Stop Conditional or Freeze and Stop policy after the implementation is proven to be stable.

GDPS HyperSwap function

If there is a problem writing or accessing the *primary* disk because of a failing, failed, or non-responsive primary disk, there is a need to swap from the primary disks to the secondary disks.

GDPS HM delivers a powerful function that is known as *HyperSwap*, which swaps from using the primary devices in a mirrored configuration to using what were the secondary devices, apparent to the production systems and applications that are using these devices. Before the availability of HyperSwap, a transparent disk swap was not possible. All systems using the primary disk would have been shut down (or might have failed, depending on the nature and scope of the failure) and would have been restarted by using the secondary disks. Disk failures were often a single point of failure for the entire sysplex.

With HyperSwap, such a switch can be accomplished without IPL and with a brief hold on application I/O. The HyperSwap function is completely controlled by automation, thus allowing all aspects of the disk configuration switch to be controlled through GDPS.

HyperSwap can be started in two ways:

► Planned HyperSwap: A planned HyperSwap is started by operator action by using GDPS facilities. One example of a planned HyperSwap is where a HyperSwap is initiated before planned disruptive maintenance to a disk subsystem.

► Unplanned HyperSwap: An unplanned HyperSwap is started automatically by GDPS, triggered by events that indicate a primary disk problem.

Primary disk problems can be detected as a direct result of an I/O operation to a specific device that fails because of a reason that indicates a primary disk problem, such as:

– No paths available to the device
– Permanent error
– I/O timeout

In addition to a disk problem being detected as a result of an I/O operation, it is also possible for a primary disk subsystem to proactively report that it is experiencing an acute problem. The IBM DS8000 models have a special microcode function that is known as the *Storage Controller Health Message Alert* capability. Problems of different severity are reported by disk subsystems that support this capability. Those problems classified as acute are also treated as HyperSwap triggers. After systems are swapped to use the secondary disks, the disk subsystem and operating system can try to perform recovery actions on the former primary without impacting the applications that use those disks.

Planned and unplanned HyperSwap have requirements in terms of the physical configuration, such as having a symmetrically configured configuration. If a client's environment meets these requirements, there is no special enablement that is required to perform planned swaps. Unplanned swaps are not enabled by default and must be enabled explicitly as a policy option. For more information, see "HyperSwap (Primary Failure) policy options" on page 119.

When a swap is initiated, GDPS always validates various conditions to ensure that it is safe to swap. For example, if the mirror is not fully duplex, that is, not all volume pairs are in a duplex state, then a swap cannot be performed. The way that GDPS reacts to such conditions changes depending on the condition that is detected and whether the swap is a planned or unplanned swap.

Assuming that there are no show-stoppers and the swap proceeds, for both planned and unplanned HyperSwap, the systems that are using the primary volumes experience a temporary pause in I/O processing. GDPS blocks I/O both at the channel subsystem level by performing a Freeze, which results in all disks going into ELB, and also in all systems I/O being quiesced at the operating system (UCB) level. This action ensures that no systems use the disks until the switch is complete. During this time when I/O is paused, the following actions occur:

► The Metro Mirror configuration is *physically switched*. This action involves physically changing the secondary disk status to primary. Secondary disks are protected and cannot be used by applications. Changing their status to primary allows them to come online to systems and be used.

► The disks are logically switched in each of the systems in the GDPS configuration. This involves switching the internal pointers in the operating system control blocks (UCBs). The operating system points to the former secondary devices instead of the current primary devices.

► For planned swaps, the mirroring direction can be reversed (optional).

► Finally, the systems resume operation by using the new, swapped-to primary devices even though applications are not aware of the fact that different devices are now being used.

This brief pause during which systems are locked out of performing I/O is known as the User Impact Time. In benchmark measurements at IBM that use currently supported releases of GDPS and IBM DS8000 disk subsystems, the User Impact Time to swap 10,000 pairs across 16 systems during an unplanned HyperSwap was less than 10 seconds. Most implementations are much smaller than this and typical impact times that use the most current storage and server hardware are measured in seconds. Although the results depend on your configuration, these numbers give you a high-level idea of what to expect.

GDPS HM HyperSwaps all devices in the managed configuration. Just as the Freeze function applies to the entire consistency group, HyperSwap is similar for the entire consistency group. For example, if a single mirrored volume fails and HyperSwap is started, processing is swapped to the secondary copy of *all* mirrored volumes in the configuration, including the ones in other, unaffected, subsystems. To maintain disaster readiness, all primary volumes *must* be in the same site. If HyperSwap were to swap only the failed LSS, you would then have several primaries in one site, and the remainder in the other site. This also makes for a complex environment to operate and administer I/O configurations.

Why is this necessary? Consider the configuration that is shown in Figure 4-1 on page 119. This configuration is what might happen if only the volumes of a single LSS or subsystem were hyperswapped without swapping the whole consistency group. What happens if a remote copy failure occurs at 15:00? The secondary disks in both sites are frozen at 15:00 and the primary disks (in a Freeze and Go policy) continue to receive updates.
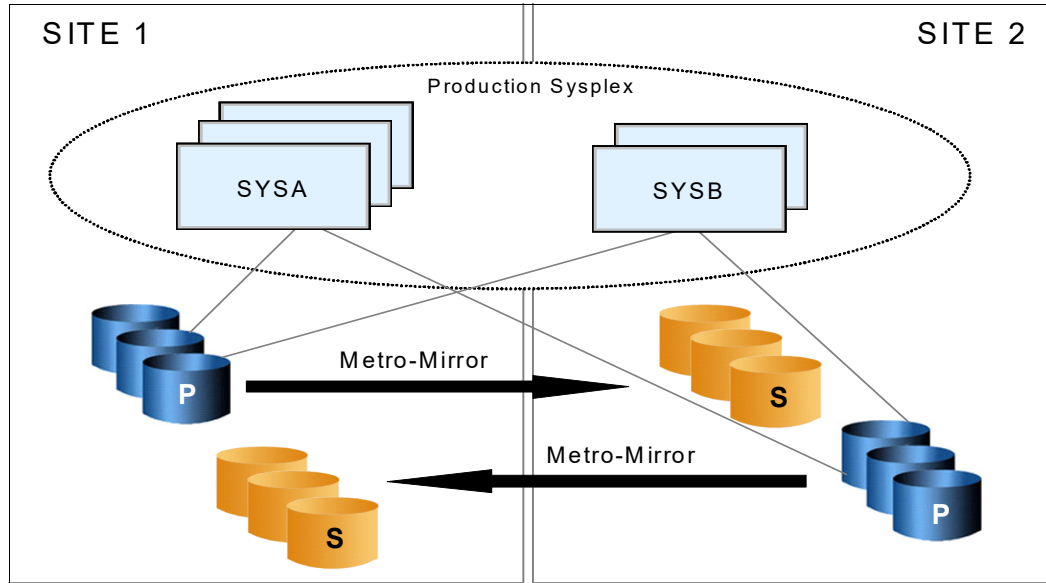
Figure 4-1   Unworkable Metro Mirror disk configuration

Now, assume that either site is hit by another failure at 15:10. What do you have? Half the disks are now at 15:00 and the other half are at 15:10 and *neither* site has consistent data. In other words, the volumes are of virtually no value to you.

If you had *all* secondaries in Site2, all volumes in that site would be consistent. If you had the disaster at 15:10, you would lose 10 minutes of data with the Go policy, but at least all the data in Site2 would be usable. Using a Freeze and Stop policy is no better for this partial swap scenario because, with a mix of primary disks in either site, you must maintain I/O configurations that can match *every* possible combination to IPL any systems.

More likely, you must first restore mirroring across the entire consistency group before recovering systems, which is not practical. Therefore, for DR readiness, it is necessary that *all* the primary volumes are in one site and *all* the secondaries in the other site.

HyperSwap with less than full channel bandwidth

You might consider enabling unplanned HyperSwap even if you do not have sufficient cross-site channel bandwidth to sustain the full production workload for normal operations. Assuming that a disk failure is likely to cause an outage and you must switch to using disk in the other site, the unplanned HyperSwap might at least present you with the opportunity to perform an orderly shutdown of your systems first. Shutting down your systems cleanly avoids the complications and restart time elongation that is associated with a crash-restart of application subsystems.

HyperSwap (Primary Failure) policy options

Clients might prefer not to immediately enable their environment for unplanned HyperSwap when they first implement GDPS. For this reason, HyperSwap is not enabled by default. However, we strongly suggest that all GDPS HM clients enable their environment for unplanned HyperSwap.

An unplanned swap is the action that makes most sense when a primary disk problem is encountered. However, other policy specifications that do not result in a swap are available. When GDPS detects a primary disk problem trigger, the first thing that it does is a Freeze (the same as it performs when a mirroring problem trigger is detected).

GDPS then uses the selected Primary Failure policy option to determine what action it takes next:

► `PRIMARYFAILURE=GO`

No swap is performed. The action GDPS takes is the same as for a freeze event with policy option `PPRCFAILURE=GO`. A Run action is performed, which allows the systems to continue by using the original primary disks. Metro Mirror is suspended, so updates are not replicated to the secondary. However, depending on the scope of the primary disk problem, it might be that all or some production workloads cannot run or cannot sustain required service levels. Such a situation might necessitate restarting the systems on the secondary disks. Because of the freeze, the secondary disks are in a consistent state and can be used for restart. However, any transactions that ran after the Go action will be lost.

► `PRIMARYFAILURE=STOP`

No swap is performed. The action GDPS takes is the same as for a freeze event with policy option `PPRCFAILURE=STOP`. GDPS resets all the production systems. This action ensures that no further I/O occurs. After performing situation analysis, if it is determined that this issue was not transient and that the secondaries should be used to restart the systems, no data will be lost.

► `PRIMARYFAILURE=SWAP,`*swap_disabled_action*

The first parameter, `SWAP`, indicates that after performing the Freeze, GDPS will proceed with an unplanned HyperSwap. When the swap is complete, the systems are running on the new, swapped-to primary disks (the former secondaries). Metro Mirror is in a suspended state. Because the primary disks are known to be in a problematic state, there is no attempt to reverse mirroring. After the problem with the primary disks is fixed, you can instruct GDPS to resynchronize Metro Mirror from the current primaries to the former ones (which are now considered to be secondaries).

The second part of this policy, *swap_disabled_action*, indicates what GDPS should do if HyperSwap was temporarily disabled by operator action at the time the trigger was encountered. Effectively, an operator action instructed GDPS not to perform a HyperSwap, even if there is a swap trigger. GDPS already did a freeze. So, the second part of the policy says what GDPS should do next.

The following options are available for the second parameter, which comes into play only if HyperSwap was disabled by the operator (the disk is already frozen):

GO   This is the same action as GDPS would perform if the policy option had been specified as `PRIMARYFAILURE=GO`.

STOP   This is the same action as GDPS would perform if the policy option had been specified as `PRIMARYFAILURE=STOP`.

Primary Failure policy specification considerations

As indicated previously, the action that best serves RTO/RPO objectives when there is a primary disk problem is to perform an unplanned HyperSwap. Therefore, the SWAP policy option is the best practice policy option.

For the Stop or Go choice, either as the second part of the SWAP specification or if you will not be using SWAP, similar considerations apply as discussed for the PPRC Failure policy options to Stop or Go. Go carries the risk of data loss if it becomes necessary to abandon the primary disk and restart systems on the secondary. Stop carries the risk of taking an unnecessary outage if the problem was transient.

The key difference is that with a mirroring failure, the primary disks are not broken. When you allow the systems to continue to run on the primary disk with the Go option, other than a disaster, which is a low probability, the systems are likely to run with no problems. With a primary disk problem, with the Go option, you are allowing the systems to continue running on what are known to be disks that experienced a problem just seconds ago. If this problem was a serious with widespread impact, such as an entire disk subsystem failure, the applications are going to experience severe problems. Some transactions might continue to commit data to those disks that are not broken. Other transactions might be failing or experiencing serious service time issues.

Finally, if there is a decision to restart systems on the secondaries because the primary disks are not able to support the workloads, there is data loss. The probability that a primary disk problem is a real problem that necessitates a restart on the secondary disks is higher when compared to a mirroring problem. A Go specification in the Primary Failure policy increases your overall risk for data loss.

If the primary failure was of a transient nature, a Stop specification results in an unnecessary outage. However, with primary disk problems it is likely that the problem might necessitate a restart on the secondary disks. Therefore, a Stop specification in the Primary Failure policy avoids data loss and facilitates faster restart.

The considerations relating to CF structures with a PRIMARYFAILURE event are similar to a PPRCFAILURE event. If there is an actual swap, the systems continue to run and continue to use the same structures as they did before the swap; the swap is transparent. With a Go action, you continue to update the CF structures along with the primary disks after the Go. If you need to abandon the primary disks and restart on the secondary, the structures are inconsistent with the secondary disks and are not usable for restart purposes. This action prolongs the restart and your recovery time. With Stop, if you decide to restart the systems by using the secondary disks, there is no consistency issue with the CF structures because no further updates occurred on either set of disks after the trigger was detected.

GDPS usage of DS8000 functions

GDPS strives to use (when it makes sense) enhancements to the IBM DS8000 disk technologies. In this section, we provide information about the key DS8000 technologies that GDPS supports and uses.

Failover/Failback support

When a primary disk failure occurs and the disks are switched to the secondary devices, PPRC Failover/Failback (FO/FB) support eliminates the need to do a full copy when reestablishing replication in the opposite direction. Because the primary and secondary volumes are often in the same state when the freeze occurred, the only differences between the volumes are the updates that occur to the secondary devices after the switch. Failover processing sets the secondary devices to primary suspended status and starts change recording for any subsequent changes made. When the mirror is reestablished with failback processing, the original primary devices become secondary devices and a resynchronization of changed tracks takes place.

GDPS HM requires Metro Mirror FO/FB capability to be available on all disk subsystems in the managed configuration.

PPRC Extended Distance

PPRC Extended Distance (PPRC-XD) (also known as Global Copy) is an asynchronous form of the PPRC copy technology. GDPS uses PPRC-XD rather than synchronous PPRC (Metro Mirror) to reduce the performance impact of certain remote copy operations that potentially involve a large amount of data. For more information, see section 4.6.2, "GDPS HM reduced impact on initial copy and resynchronization" on page 140.

Storage Controller Health Message Alert

This feature facilitates triggering an unplanned HyperSwap proactively when the disk subsystem reports an acute problem that requires extended recovery time. For more information about unplanned HyperSwap triggers, see "GDPS HyperSwap function" on page 117.

PPRCS Summary Event Messages

GDPS supports the DS8000 PPRC Summary Event Messages (PPRCSUM) function, which is aimed at reducing the message traffic and the processing of these messages for Freeze events. For more information, see "GDPS Freeze function for mirroring failures" on page 113.

Soft Fence

Soft Fence provides the capability to block access to selected devices. As discussed in "Protecting secondary disks from accidental update" on page 123, GDPS uses Soft Fence to avoid write activity on disks that are exposed to accidental update in certain scenarios.

On-demand dump (also known as non-disruptive statesave)

When problems occur with disk subsystems, such as those that result in an unplanned HyperSwap, mirroring suspension or performance issues can happen. The lack of diagnostic information can be associated with any of the things that can happen. This function is designed to reduce the likelihood of missing diagnostic information.

Taking a full statesave can lead to temporary disruption to the host I/O and is often disliked by clients for this reason. The on-demand dump (ODD) capability of the disk subsystem facilitates taking a non-disruptive statesave (NDSS) at the time that such an event occurs. The microcode does this task automatically for certain events, such as a dump of the primary disk subsystem that triggers a Metro Mirror freeze event, and also allows an NDSS to be requested. This process enables first failure data capture (FFDC) and thus ensures that diagnostic data is available to aid problem determination. Not all information that is contained in a full statesave is contained in an NDSS. Therefore, there might still be failure situations where a full statesave is requested by the support organization.

GDPS supports taking an NDSS by using the remote copy panels (or GDPS GUI). In addition to this support, GDPS autonomically takes an NDSS if there is an unplanned freeze or HyperSwap event.

Query Host Access

When a Metro Mirror disk pair is being established, the device that is the target (secondary) must not be used by any system. The same is true when establishing a FlashCopy relationship to a target device. If the target is in use, the establishment of the Metro Mirror or FlashCopy relationship fails. When such failures occur, it can be a tedious task to identify which system is holding up the operation.

The Query Host Access disk function provides the means to query and identify what system is using a selected device. GDPS uses this capability and adds usability in several ways:

► Query Host Access identifies the LPAR that is using the selected device through the Central Processor Complex (CPC) serial number and LPAR number. It is still a tedious job for operations staff to convert this information to a system or CPC and LPAR name. GDPS does this conversion and presents the operator with more readily usable information, avoiding this additional conversion effort.

► Whenever GDPS is requested to perform a Metro Mirror or FlashCopy establish operation, GDPS first performs Query Host Access to see whether the operation is expected to succeed or fail because of one or more target devices are in use. It alerts operations if the operation is expected to fail, and identifies the target devices in use and the LPARs holding them.

► GDPS continually monitors the target devices that are defined in the GDPS configuration and alerts operations that target devices are in use when they should not be. This process enables operations to fix the reported problems in a timely manner.

► GDPS provides the ability for the operator to perform ad hoc Query Host Access to any selected device by using the GDPS panels (or GDPS GUI).

Easy Tier Heat Map Transfer

IBM DS8000 Easy Tier optimizes data of logical volumes across the various physical tiers of storage within a disk subsystem to optimize application performance. The placement decisions are based on learning the data access patterns and can be changed dynamically and transparently to the applications by using this data.

Metro Mirror mirrors the data from the primary to the secondary disk subsystem; however, the Easy Tier learning information is not included in Metro Mirror scope. The secondary disk subsystems are optimized according to the workload on these subsystems, which differs from the activity on the primary (there is only a write workload on the secondary; there is read/write activity on the primary). As a result of this difference, during a disk switch or disk recovery, the secondary disks that you switch to are likely to display different performance characteristics compared to the former primary.

Easy Tier Heat Map Transfer is the DS8000 capability to transfer the Easy Tier learning from a Metro Mirror primary to the secondary disk subsystem so that the secondary disk subsystem can also be optimized based on this learning and has similar performance characteristics if it is promoted to become the primary.

GDPS integrates support for Heat Map Transfer. The Heat Map Transfer actions (such as start or stop of the processing and reversing transfer direction) are incorporated into the GDPS managed processes. For example, if Metro Mirror is temporarily suspended by GDPS for a planned or unplanned secondary disk outage, Heat Map Transfer is also suspended; or if Metro Mirror direction is reversed as a result of a HyperSwap, Heat Map Transfer direction is also reversed.

Protecting secondary disks from accidental update

A system cannot be started by using a disk that is physically a Metro Mirror secondary disk because Metro Mirror secondary disks cannot be brought online to any systems. However, a disk can be secondary from a GDPS (and application use) perspective but physically have a simplex or primary status from a Metro Mirror perspective.

For planned and unplanned HyperSwap, and a disk recovery, GDPS changes former secondary disks to primary or simplex state. However, these actions do not modify the state of the former primary devices, which remain in the primary state. Therefore, the former primary devices remain accessible and usable even though they are considered to be the secondary disks from a GDPS perspective. This configuration makes it is possible to accidentally update or IPL from the wrong set of disks. Accidentally using the wrong set of disks can result in a potential data integrity or data loss problem.

GDPS HM provides IPL protection early in the IPL process. During initialization of GDPS, if GDPS detects that the system coming up was started by using the wrong set of disks, GDPS quiesces that system, preventing any data integrity problems that might be experienced if the applications were started.

GDPS also uses an IBM DS8000 disk subsystem capability, which is called Soft Fence, for configurations where the disks support this function. Soft Fence provides the means to fence, that is, to block access to a selected device. GDPS uses Soft Fence when appropriate to fence devices that might otherwise be exposed to accidental update.

4.1.2 Protecting distributed (Fixed-Block) data

> **Terminology:** The following definitions describe the terminology that we use in this book when referring to the various types of disks:
>
> ► IBM Z or Count-Key-Data (CKD) disks
>
>   GDPS can manage disks that are formatted as CKD disks (the traditional mainframe format) that are used by any of the following IBM Z operating systems: z/VM, VSE, KVM, and Linux on IBM Z.
>
>   We refer to the disks that are used by a system that is running on the mainframe as *IBM Z disks*, *CKD disks*, or *CKD devices*. These terms are used interchangeably.
>
> ► Fixed-Block (FB) disks
>
>   Disks that are used by systems other than those systems that are running on IBM Z are traditionally formatted as FB and are referred to as *FB disks* or *FB devices* in this book.

GDPS HM can manage the mirroring of FB devices that are used by non-mainframe operating systems. The FB devices can be part of the same consistency group as the mainframe CKD devices, or they can be managed separately in their own consistency group.

For more information about FB disk management, see 4.3, "Fixed-Block disk management" on page 129.

4.1.3 Protecting other CKD data

Systems that are fully managed by GDPS are known as *GDPS managed systems* or *GDPS systems*. These systems are the z/OS systems in the GDPS sysplex.

GDPS HM can also manage the disk mirroring of CKD disks that are used by systems outside the sysplex: Other z/OS systems, Linux on IBM Z, virtual machine (VM), and VSE systems that are not running any GDPS HM or xDR automation. These systems are known as "foreign systems."

Because GDPS manages Metro Mirror for the disks that are used by these systems, these disks must be attached to the GDPS controlling systems. With this setup, GDPS is able to capture mirroring problems and performs a freeze. All GDPS managed disks belonging to the GDPS systems and these foreign systems are frozen together, regardless of whether the mirroring problem is encountered on the GDPS systems' disks or the foreign systems' disks.

GDPS HM cannot directly communicate with these foreign systems. For this reason, GDPS automation are not aware of certain other conditions, such as a primary disk problem that is detected by these systems. Because GDPS is not aware of such conditions that would have otherwise driven autonomic actions such as HyperSwap, GDPS cannot react to these events.

If an unplanned HyperSwap occurs (because it triggered on a GDPS managed system), the foreign systems cannot and do not swap to using the secondaries. Mechanisms are provided to prevent these systems from continuing to use the former primary devices after the GDPS systems are swapped. You can then use GDPS automation facilities to reset these systems and restart them by using the swapped-to primary disks.

4.2 GDPS Metro HyperSwap Manager configurations

A basic GDPS Metro HyperSwap Manager configuration consists of at least one production system, at least one controlling system, primary disks, and secondary disks. The entire configuration can be in either a single site to provide protection from disk outages with HyperSwap, or it can be spread across two data centers within metropolitan distances as the foundation for a DR solution. The actual configuration depends on your business and availability requirements.

4.2.1 Controlling system

Why does a GDPS Metro HyperSwap Manager configuration need a controlling system? At first, you might think that a controlling system is an extra infrastructure overhead. However, when you have an unplanned outage that affects production systems or the disk subsystems, it is crucial to have a system such as the controlling system that can survive failures that might impact other portions of your infrastructure. The controlling system perform situation analysis after the unplanned event to determine the status of the production systems or the disks. The controlling system plays a vital role in a GDPS Metro HyperSwap Manager configuration.

The controlling system must be in the same sysplex as the production system (or systems) so it can see all the messages from those systems and communicate with those systems. However, it shares an absolute minimum number of resources with the production systems (typically only the sysplex couple data sets (CDS)). By being configured to be as self-contained as possible, the controlling system is unaffected by errors that might stop the production systems (for example, an ELB event on a primary volume).

The controlling system must have connectivity to all the Site1 and Site2 primary and secondary devices that it manages. If available, it is preferable to isolate the controlling system infrastructure on a disk subsystem that is not housing mirrored disks that are managed by GDPS.

The controlling system is responsible for carrying out all Metro Mirror and Server Time Protocol (STP)-related recovery actions following a disaster or potential disaster, for managing the disk mirroring configuration, for initiating a HyperSwap, for initiating a freeze and implementing the freeze policy actions following a freeze event, for reassigning STP roles, and other actions.

The availability of the dedicated GDPS controlling system (or systems) in *all* configurations is a fundamental requirement of GDPS. It is not possible to merge the function of the controlling system with any other system that accesses or uses the primary volumes or other production resources.

Especially in 2-site configurations, configuring GDPS HM with two controlling systems, one in each site is a best practice because a controlling system is designed to survive a failure in the opposite site of where the primary disks are. Primary disks are normally in Site1 and the controlling system in Site2 is designed to survive if Site1 or the disks in Site1 fail. However, if you reverse the configuration so that the primary disks are in Site2, the controlling system is in the same site as the primary disks. It cannot survive a failure in Site2 and might or might not survive a failure of the disks in Site2, depending on the configuration. Configuring a controlling system in both sites ensures as much protection no matter which site is the primary disk site. When two controlling systems are available, GDPS manages by assigning a primary role to the controlling system that is in the same site as the secondary disks and switching the primary role if there is a disk switch.

Improved controlling system availability: Enhanced timer support

Normally, a loss of synchronization with the sysplex timing source generates a disabled console write to operator with reply (WTOR) that suspends all processing on the LPAR until a response is made to the WTOR. The WTOR message is IEA394A in STP timing mode.

In a GDPS environment, z/OS is aware that a system is a GDPS controlling system and allows a GDPS controlling system to continue processing even when the server it is running on loses its time source and becomes unsynchronized. Therefore, the controlling system is able to complete any freeze or HyperSwap processing it might start, and it is available for situation analysis and other recovery actions instead of being in a disabled WTOR state.

In addition, because the controlling system is operational, it can be used to help in problem determination and situation analysis during the outage, thus reducing further the recovery time that is needed to restart applications.

The controlling system is required to perform GDPS automation if there is a failure, which might include these actions:

► Performing the freeze processing to ensure secondary data consistency
► Coordinating HyperSwap processing
► Aiding with situation analysis

Because the controlling system needs to run only with a degree of time synchronization that allows it to correctly participate in heartbeat processing regarding the other systems in the sysplex, this system should be able to run unsynchronized for 80 minutes by using the local time-of-day (TOD) clock of the server (referred to as *local timing mode*), rather than generating a WTOR.

Automated response to STP sync WTORs

GDPS on the controlling systems, by using the BCP Internal Interface (BCPii), provides automation to reply to WTOR IEA394A when the controlling systems are running in local timing mode (for more information, see "Improved controlling system availability: Enhanced timer support" on page 126). A server in an STP network might recover from an unsynchronized to a synchronized timing state without client intervention. By automating the response to the WTORs, potential timeouts of subsystems and applications in the client's enterprise might be averted, thus potentially preventing a production outage.

If either WTOR IEA394A is posted for production systems, GDPS uses the BCPii to automatically reply RETRY to the WTOR. If z/OS determines that the CPC is in a synchronized state, either because STP recovered or the Coordinated Timing Network (CTN) was reconfigured, it no longer spins and continues processing. If the CPC is still in an unsynchronized state when GDPS automation responded with RETRY to the WTOR, however, the WTOR is reposted.

The automated reply for any system is retried for 60 minutes. After 60 minutes, you will need to manually respond to the WTOR.

4.2.2 GDPS Metro HyperSwap Manager in a single site

In the single-site configuration, the controlling systems, primary disks, and secondary disks are all in the same site, as shown in Figure 4-2. This configuration benefits from the capabilities of GDPS Metro HyperSwap Manager to manage the mirroring environment, and HyperSwap across planned and unplanned disk reconfigurations. A single site configuration does *not* provide DR capabilities because all the resources are in the same site, and if that site suffers a disaster, then the systems and disk are all gone.

> **Note:** We continue to refer to Site1 and Site2 in this section, although this terminology here refers to the two copies of the production data in the same site.

Although having a single controlling system might be acceptable, we suggest having two controlling systems to provide the best availability and protection. The K1 controlling system can use Site2 disks, and K2 can use the Site1 disks. In this manner, a single failure does not affect availability of at least one of the controlling systems, and it is available to perform GDPS processing.
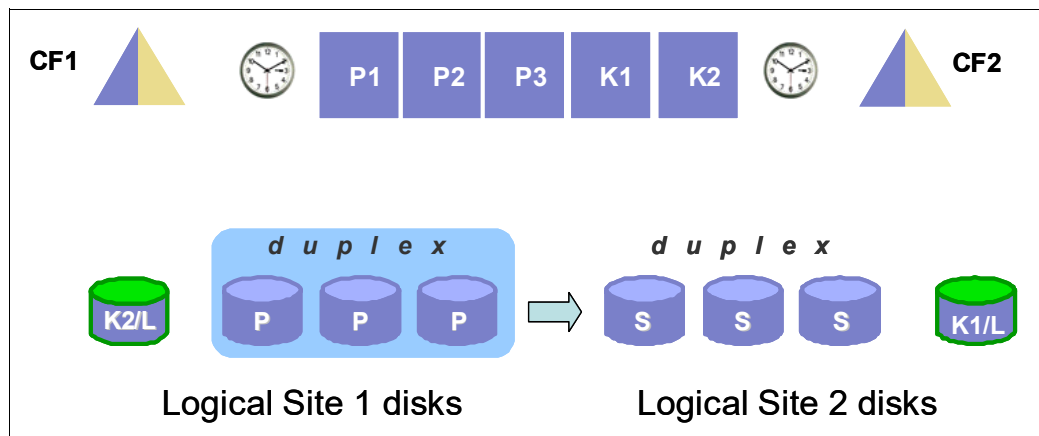
Figure 4-2   GDPS Metro HyperSwap Manager single-site configuration

4.2.3 GDPS Metro HyperSwap Manager in a 2-site configuration

Another option is to use GDPS Metro HyperSwap Manager with the primary disk in one site, and the secondaries in a second site, as shown in Figure 4-3. This configuration *does* provide the foundation for DR because the secondary copy of disk is in a separate site that is protected from a disaster in Site1. GDPS Metro HyperSwap Manager also delivers the freeze capability, which ensures a consistent set of secondary disk in a disaster.
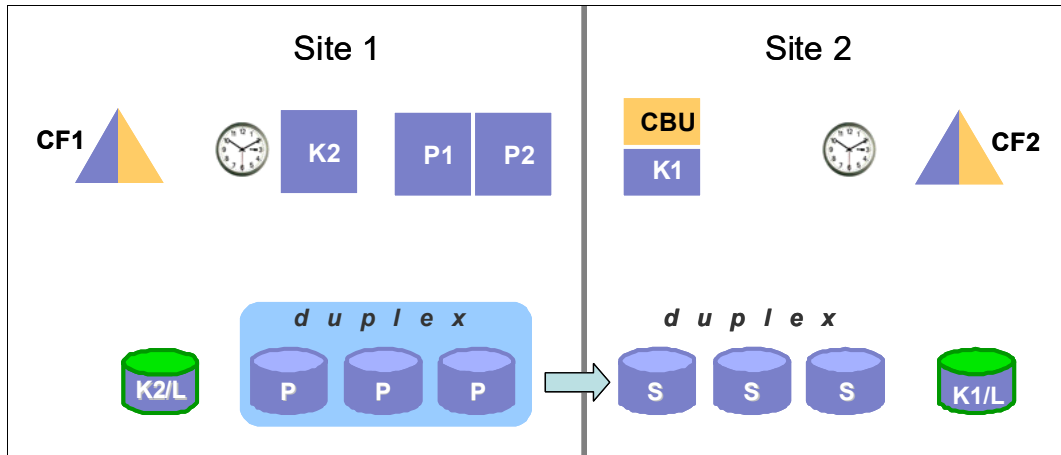
Figure 4-3   GDPS Metro HyperSwap Manager 2-site configuration

If you have a 2-site configuration, and chose to implement only one controlling system, it is a best practice that you place the controlling system in the recovery site. The advantage of this is that the controlling system continues to be available even if a disaster takes down the whole production site. Placing the controlling system in the second site creates a multisite sysplex, meaning that you must have the appropriate connectivity between the sites. To avoid cross-site sysplex connections, you might also consider the Business Recovery Services (BRS) configuration that is described in 3.2.4, "Business Recovery Services configuration" on page 65.

To get the full benefit of HyperSwap and the second site, ensure that there is sufficient bandwidth for the cross-site connectivity from the primary site servers to the secondary site disk. Otherwise, although you might be able to successfully perform the HyperSwap to the second site, the I/O performance following the swap might not be acceptable.

4.2.4 GDPS Metro HyperSwap Manager in a 3-site configuration

GDPS Metro HyperSwap Manager can be combined with GDPS GM in a 3-site configuration. In this configuration, GDPS Metro HyperSwap Manager provides protection from disk outages across a metropolitan area or within the same local site, and GDPS GM provides DR capability in a remote site.

We call this combination GDPS Metro Global - GM (GDPS MGM). For more information about the capabilities and limitations of using GDPS Metro HyperSwap Manager in a GDPS MGM solution, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237.

4.2.5 Other important considerations

The availability of the dedicated GDPS controlling system (or systems) in *all* scenarios is a fundamental requirement in GDPS. It is not possible to merge the function of the controlling system with any other system that accesses or uses the primary volumes.

4.3 Fixed-Block disk management

As discussed in 3.3.1, "Fixed-Block disk management" on page 69, most enterprises today run applications that update data across multiple platforms. For these enterprises, there is a need to manage and protect not just the IBM Z data, but also the data that is on FB devices for IBM Z and non IBM Z servers. GDPS HM provides the capability to manage a heterogeneous environment of IBM Z and distributed systems data through a function that is called that is FB disk management.

The FB Disk Management function allows GDPS to be a single point of control to manage business resiliency across multiple tiers in the infrastructure, which improves cross-platform system management and business processes. GDPS HM can manage the Metro Mirror remote copy configuration and FlashCopy for distributed systems storage.

Specifically, FB disk support extends the GDPS HM Freeze capability to FB devices that are in supported disk subsystems to provide data consistency for the IBM Z data and the data on the FB devices.

With FB devices included in your configuration, you can select one of the following options to specify how Freeze processing is to be handled for FB disks and IBM Z (CKD disks), when mirroring or primary disk problems are detected:

► Freeze all devices that are managed by GDPS.

   If this option is used, the CKD and FB devices are in a single consistency group. Any Freeze trigger (for the IBM Z or FB devices) results in the FB and the IBM Z LSSs managed by GDPS being frozen. This option provides consistent data across heterogeneous platforms if a disaster occurs so that you can restart systems in the site where secondary disks are located. This option is especially suitable when there are distributed units of work on IBM Z and distributed servers that update the same data; for example, by using IBM Db2 DRDA, which is the IBM Distributed Relational Database Architecture.

► Freeze devices by group.

   If this option is selected, the CKD devices are in a separate consistency group from the FB devices. Also, the FB devices can be separated into Multiple Consistency Groups (MCG) by distributed workloads, for example. The Freeze is performed on only the group for which the Freeze trigger was received. If the Freeze trigger occurs for an IBM Z disk device, only the CKD devices are frozen. If the trigger occurs for an FB disk, only the FB disks within the same group as that disk are frozen.

4.3.1 FB disk management prerequisites

GDPS requires the disk subsystems that contain the FB devices to support the z/OS Fixed-Block Architecture (zFBA) feature. GDPS runs on z/OS and therefore communicates to the disk subsystems directly over a channel connection. The zFBA provides GDPS the ability to send the commands that are necessary to manage Metro Mirror and FlashCopy directly to FB devices over a channel connection.

It also enables GDPS to receive notifications for specific error conditions (for example, suspension of an FB device pair). These notifications allow the GDPS controlling system to drive autonomic action, such as performing a freeze for a mirroring failure.

Note: HyperSwap for FB disks is not supported for any IBM Z or non IBM Z servers.

4.4 Managing the GDPS Metro HyperSwap Manager environment

The bulk of the functions that are delivered with GDPS Metro HyperSwap Manager relate to maintaining the integrity of the secondary disks and being able to nondisruptively switch to the secondary volume of the Metro Mirror pair.

However, there is an extra aspect of remote copy management that is available with GDPS Metro HyperSwap Manager, namely the ability to query and manage the remote copy environment using the GDPS panels.

In this section, we describe this other aspect of GDPS Metro HyperSwap Manager. Specifically, GDPS Metro HyperSwap Manager provides multiple mechanisms including user interfaces, a command-line interface (CLI), and multiple application programming interfaces (APIs) to let you:

► Be alerted to any changes in the remote copy environment.
► Display the remote copy configuration.
► Start, stop, and change the direction of remote copy.
► Perform HyperSwap operations.
► Start and stop FlashCopy.

Note: GDPS Metro HyperSwap Manager does not provide script support. For scripting support with added capabilities, the full-function GDPS Metro product is required.

4.4.1 User interfaces

Two primary user interface options are available for GDPS HM: the NetView 3270 panels and a browser-based GUI (also referred to as the GDPS GUI in this book).

An example of the main GDPS HM 3270-based panel is shown in Figure 4-4.

```
VPCPPNLN               GDPS Metro HyperSwap Manager            GDPS V4.R4.M
User ID: GAEL
                        ---- GDPS Status Indicators ----

  System       =  G1C1    - A6P11   PPRC and HyperSwap status = OK
  Current Master =  G1C2    - A6P13   Primary Dasd = RS1
  Debug        =  ON



                        ---- GDPS Options ----

     1     Dasd Remote Copy      7      Sysplex Resource Management
                                 8      Debug ON/OFF
     3     Standard Actions      9      View Definitions
                                 H      Health Checks and Diagnostics
                                 C      Config Management
     6     Planned Actions       M      Run Monitor1/Monitor3



 Selection ===>   _
                 Licensed Materials - Property of IBM
          6942-35B © Copyright IBM Corp. 1998, 2021  All Rights Reserved.
```

Figure 4-4   GDPS Metro HyperSwap Manager main GDPS panel

Notice that several option choices are in blue instead of black. These blue options are supported by the GDPS Metro offering, but are not part of GDPS Metro HyperSwap Manager.

This panel includes a summarized configuration status at the top and a menu of choices. For example, to view the disk mirroring (Dasd Remote Copy) panels, enter `1` at the `Selection` prompt, and then press Enter.

GDPS GUI

The GDPS GUI is a browser-based interface that is designed to improve operator productivity. The GDPS GUI provides the same functional capabilities as the 3270-based panels, such as providing management capabilities for Remote Copy Management, Configuration Management, Status Display Facility (SDF) Monitoring, and browsing the CANZLOG by using simple point-and-click procedures. Advanced sorting and filtering is available in most of the views that are provided by the GDPS GUI. In addition, users can open multiple windows or tabs to allow for continuous status monitoring, while performing other GDPS Metro HyperSwap Manager management functions.

The GDPS GUI display has four main sections:

► The application header at the top of the page that provides an Actions button for carrying out various GDPS tasks, along with the help function and the ability to log off or switch between target systems.

► The application menu is down the left side of the window. This menu gives access to various features and functions available through the GDPS GUI.

► The active window that shows context-based content depending on the selected function. This tabbed area is where the user can switch context by clicking a different tab.

► A status summary area is shown at the bottom of the display.

The initial status panel of the GDPS Metro HyperSwap Manager GDPS GUI is shown in Figure 4-5. This panel provides an instant view of the status and direction of replication, and HyperSwap status. Hovering over the various icons provides more information by using windows.

> **Note:** For the remainder of this section, only the GDPS GUI is shown to illustrate the various GDPS management functions. The equivalent traditional 3270 panels are not shown here.

Figure 4-5   Full view of GDPS GUI main panel

Monitoring function: Status Display Facility

GDPS also provides many monitors to check the status of disks, sysplex resources, and other components. Anytime there is a configuration change, or something in GDPS that requires manual intervention, GDPS raises an alert. GDPS uses the SDF provided by System Automation as the primary status feedback mechanism for GDPS.

GDPS provides a dynamically updated window, as shown in Figure 4-6. A summary of all current alerts is provided at the bottom of each window. The initial view that is presented is for the SDF trace entries so you can follow, for example, script execution. Click one the other alert categories to view the different alerts that are associated with automation or remote copy in either site, or select **All** to see all alerts. You can sort and filter the alerts based on various fields that are presented, such as severity.

Figure 4-6   GDPS GUI SDF panel

By default, the GDPS GUI refreshes the alerts automatically every 10 seconds. As with the 3270 window, if there is a configuration change or a condition that requires special attention, the color of the icons changes based on the severity of the alert. By pointing to and clicking any of the highlighted fields, you can obtain detailed information regarding the alert.

Remote copy panels

The z/OS Advanced Copy Services capabilities are powerful, but the native CLI, z/OS TSO, and ICKDSF interfaces are not as simple as the DASD remote copy panels are. To more easily check and manage the remote copy environment, use the DASD remote copy panels that are provided by GDPS.

For GDPS to manage the remote copy environment, you must first define the configuration (primary and secondary LSSs, primary and secondary devices, and PPRC links) to GDPS in a file called the GEOPARM file. This GEOPARM file can be edited and introduced to GDPS directly from the GDPS GUI.

After the configuration is known to GDPS, you can use the panels to check that the current configuration matches the one you want. You can start, stop, suspend, and resynchronize mirroring at the volume or LSS level. These actions can be done at the device or LSS level, or both. Figure 4-7 shows the mirroring panel for CKD devices at the LSS level.

Figure 4-7   GDPS GUI Dasd Remote Copy SSID panel

The Dasd Remote Copy panel is organized into three sections:

► Upper left provides a summary of the device pairs in the configuration and their status.

► Upper right starts GDPS-managed FlashCopy operations.

► A table with one row for each LSS pair in your GEOPARM. In addition to the rows for each LSS, there is a header row with an **Action** menu to perform the various DASD management tasks, and the ability to filter the information presented.

To perform an action on a single SSID-pair, double click a row in the table. A panel is then displayed, where you can perform the same actions as those available as line commands on the top section of the 3270 panel.

After an individual SSID-pair is selected, the frame that is shown in Figure 4-8 is displayed. The table in this frame shows each of the mirrored device pairs within a single SSID-pair, along with the status of each pair. In this example, all the pairs are fully synchronized and in duplex status, as summarized in the upper left area. More details can be viewed for each pair by double-clicking the row, or selecting the row with a single click and then selecting **Query** from the Actions menu.
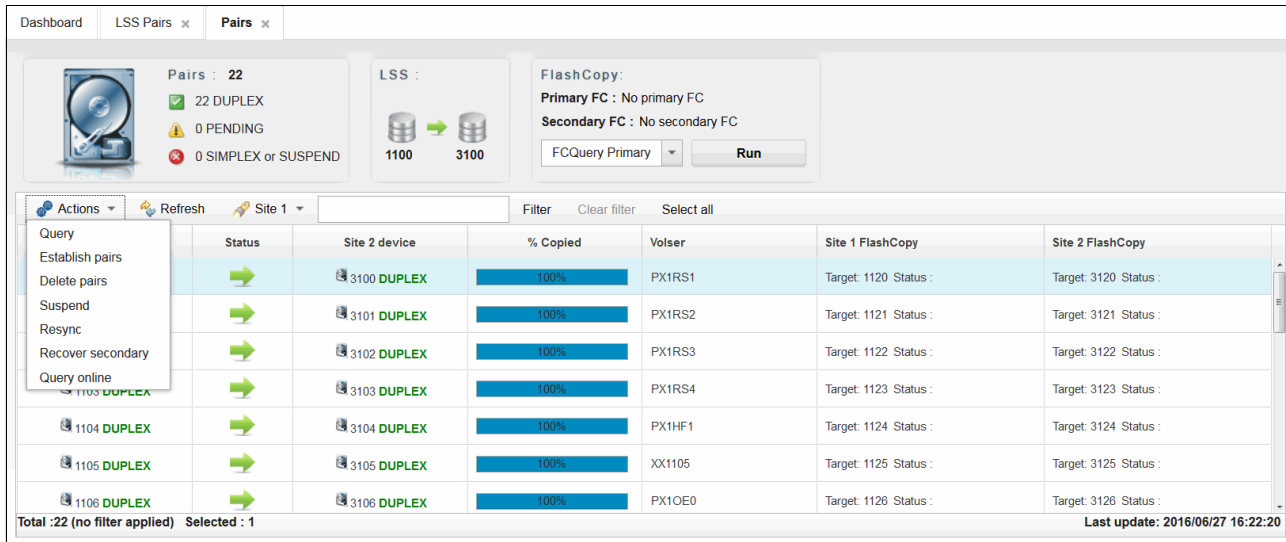
Figure 4-8   GDPS GUI Dasd Remote Copy: View Devices detail panel

If you are familiar with using the TSO or ICKDSF interfaces, you might appreciate the ease of use of the DASD remote copy panels.

These panels provided by GDPS are *not* intended to be a remote copy monitoring tool. Because of the overhead that is involved in gathering the information for every device to populate the panels, GDPS gathers this data only on a timed basis, or on demand following an operator instruction. The normal interface for finding out about remote copy status or problems is the SDF.

Similar panels are provided for controlling the Open LUN devices.

4.4.2 NetView commands

Even though GDPS Metro HyperSwap Manager does not support using scripts as GDPS Metro does, certain GDPS operations are started by using NetView commands that perform similar actions to the equivalent script command. These commands are entered at a NetView command prompt.

There are commands to perform the following types of actions (not an all-inclusive list). One command would accomplish all the following actions:

► Temporarily disable HyperSwap and re-enable HyperSwap.

► List systems in the GDPS and identify which are controlling systems.

► Perform a planned HyperSwap disk switch.

► Perform a planned freeze of the disk mirror.

► Make the secondary disks usable through a PPRC failover or recover action.

► Restore Metro Mirror mirroring to a duplex state.

► Take a point-in-time (PiT) copy of the current set of secondary CKD devices. FlashCopy uses COPY or NOCOPY options and the NOCOPY2COPY option to convert an existing FlashCopy taken with NOCOPY to COPY are supported. The CONSISTENT option (FlashCopy Freeze) is supported with the COPY. and NOCOPY options.

► Reconfigure an STP-only CTN by reassigning the Preferred Time Server (PTS) and Current Time Server (CTS) roles and the Backup Time Server (BTS) and Arbiter roles to one or more CPCs.

► Unfence disks that were blocked by Soft Fence.

4.4.3 Application programming interfaces

GDPS provides two primary programming interfaces to allow other programs that are written by clients: independent software vendors (ISVs), and other IBM product areas to communicate with GDPS. These APIs allow clients, ISVs, and other IBM product areas to complement GDPS automation with their own automation code. The following sections describe the APIs that are provided by GDPS.

GDPS Metro HyperSwap Manager Query Services

GDPS maintains configuration information and status information in NetView variables for the various elements of the configuration that it manages. Query Services is a capability that allows client-written NetView REXX programs to query the value of numerous GDPS internal variables. The variables that can be queried pertain to the Metro Mirror configuration, the system and sysplex resources that are managed by GDPS, and other GDPS facilities, such as HyperSwap and GDPS Monitors.

In addition to the Query Services function that is part of the base GDPS product, GDPS provides several samples in the GDPS SAMPLIB library to demonstrate how Query Services can be used in client-written code.

GDPS also makes available to clients a tool that is called the *Preserve Mirror Tool* (PMT), which facilitates adding disks to the GDPS Metro HyperSwap Manager configuration and bringing these disks to duplex.

RESTful APIs

As described in "GDPS Metro HyperSwap Manager Query Services", GDPS maintains configuration information and status information about the various elements of the configuration that it manages. Query Services can be used by REXX programs to query this information.

The GDPS RESTful API also provides the ability for programs to query this information. Because it is a RESTful API, it can be used by programs that are written in various programming languages, including REXX that are running on various server platforms.

In addition to querying information about the GDPS environment, the GDPS RESTful API allows programs that are written by clients, ISVs, and other IBM product areas to start actions against various elements of the GDPS environment. Examples of these actions include starting and stopping Metro Mirror, running HyperSwap operations, and starting GDPS monitor processing. These capabilities enable clients, ISVs, and other IBM product areas to provide an even richer set of functions to complement the GDPS functions.

GDPS provides samples in the GDPS SAMPLIB library to demonstrate how the GDPS RESTful API can be used in programs.

4.5 GDPS Metro HyperSwap Manager monitoring and alerting

The GDPS SDF panel, as discussed in "Monitoring function: Status Display Facility" on page 132, is where GDPS dynamically displays color-coded alerts.

Alerts can be posted as a result of an unsolicited error situation for which GDPS listens. For example, if one of the multiple PPRC links that provide the path over which Metro Mirror operations take place is broken, there is an unsolicited error message issued. GDPS listens for this condition and raises an alert on the SDF panel notifying the operator of the fact that a PPRC link is not operational.

Clients run with multiple PPRC links and if one is broken, Metro Mirror still continues over any remaining links. However, it is important for the operations staff to be aware of the fact that a link is broken and fix this situation because a reduced number of links results in reduced Metro Mirror bandwidth and reduced redundancy. If this problem is not fixed in a timely manner, and more links have a failure, it can result in production impact because of insufficient mirroring bandwidth or total loss of Metro Mirror connectivity (which results in a freeze).

Alerts can also be posted as a result of GDPS periodically monitoring key resources and indicators that relate to the GDPS Metro HyperSwap Manager environment. If any of these monitoring items are found to be in a state that is deemed to be not normal by GDPS, an alert is posted on SDF.

Various GDPS monitoring functions run on the GDPS controlling systems and on the production systems because, from a software perspective, it is possible that different production systems have a different view of some of the resources in the environment and although status can be normal in one production system, it might be not normal in another. All GDPS alerts that are generated on one system in the GDPS sysplex are propagated to all other systems in the GDPS. This propagation of alerts provides a single focal point of control. Operators monitor SDF on the primary K-sys for all alerts that are generated in the entire GDPS complex.

When an alert is posted, the operator must investigate (or escalate) and corrective action must be taken for the reported problem as soon as possible. After the problem is corrected, this correction is detected during the next monitoring cycle and the alert is cleared by GDPS automatically.

GDPS Metro HyperSwap Manager monitoring and alerting capability is intended to ensure that operations are notified of and can take corrective action for any problems in their environment that can affect the ability of GDPS Metro HyperSwap Manager to do recovery operations. This process maximizes the chances of achieving your IT resilience commitments.

4.5.1 GDPS Metro HyperSwap Manager health checks

In addition to the GDPS Metro HyperSwap Manager monitoring described, GDPS provides health checks. These health checks are provided as a plug-in to the z/OS Health Checker infrastructure to check that certain settings related to GDPS adhere to best practices.

The z/OS Health Checker infrastructure is intended to check various settings to see whether these settings adhere to z/OS optimum values. For settings that are not inline with best practices, exceptions are raised in the Spool Display and Search Facility (SDSF) and optionally, SDF alerts are also raised. If these settings do not adhere to recommendations, this issue can hamper the ability of GDPS to perform critical functions in a timely manner.

Often, if there are changes in the client environment, they might necessitate adjustment of various parameter settings that are associated with z/OS, GDPS, and other products. It is possible that you can miss making these adjustments, which might affect GDPS. The GDPS health checks are intended to detect such situations and avoid incidents where GDPS is unable to perform its job because of a setting that is less than ideal.

For example, GDPS Metro HyperSwap Manager requires that the controlling systems' data sets are allocated on non-mirrored disks in the same site where the controlling system runs. The Site1 controlling systems' data sets must be on a non-mirrored disk in Site1 and the Site2 controlling systems' data sets must be on a non-mirrored disk in Site2. One of the health checks provided by GDPS Metro HyperSwap Manager checks that each controlling system's data sets are allocated in line with the GDPS best practices.

Similar to z/OS and other products that provide health checks, GDPS health checks are optional. Several optimum values that are checked and the frequency of the checks can be customized to cater to unique client environments and requirements.

Several z/OS best practices conflict with GDPS best practices. The z/OS and GDPS health checks for these practices result in conflicting exceptions being raised. For such health check items, to avoid conflicting exceptions, z/OS provides the capability to define a coexistence policy where you can indicate which best practice is to take precedence; GDPS or z/OS. GDPS includes sample coexistence policy definitions for the GDPS checks that are known to be conflicting with those definitions for z/OS.

GDPS also provides a useful interface for managing the health checks by using the GDPS panels. You can perform actions such as activate/deactivate or run any selected health check, view the customer overrides in effect for any best practices values, and other actions.

Figure 4-9 on page 139 shows a sample of the GDPS Health Check management panel. In this example, you see that all the health checks are enabled. The status of the last run is also shown indicating that some were successful and some resulted in a medium exception. The exceptions can also be viewed by using other options on the panel.

```
VPC8PHC0         GDPS Health Checks Information              G7C2          (1/14)


HealthChecker:  STARTED    Procname: HZSPROC    Policy Name: GDPS

Actions:      S elect   R un   A ctivate   D eactivate   I nfo

Cmd Check name                       State              Status

  _    GDPS_CHECK_CONFIG            ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_SPOF              ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_GRS               ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_XCF_CDS           ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_CONSOLE           ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_DASDMIH           ACTIVE(ENABLED)    EXCEPTION-MEDIUM
  _    GDPS_CHECK_DEVICE            ACTIVE(ENABLED)    SUCCESSFUL
  _    GDPS_CHECK_XCF               ACTIVE(ENABLED)    SUCCESSFUL
  _    GDPS_CHECK_MAXSYS            ACTIVE(ENABLED)    SUCCESSFUL
  _    GDPS_CHECK_K_SYS_LPAR        ACTIVE(ENABLED)    SUCCESSFUL


 Command Issued
 Commands:  R Run  A Activate  D Deactivate  S Summary  G GEOHCP00  P Parameter
 Selection ===>
 F1=Help  F3=Return  F5=Refresh  F6=Roll        F8=Down  F10=Left  F11=Right
```

Figure 4-9   GDPS Metro HyperSwap Manager Health Check management panel

4.6 Other facilities that are related to GDPS

In this section, we describe miscellaneous facilities that are provided by GDPS HM that can help in various ways, such as reducing the window during which DR capability is not available.

4.6.1 HyperSwap coexistence

In the following sections, we discuss the GDPS enhancements that remove various restrictions that existed regarding HyperSwap coexistence with products such as Softek Transparent Data Migration Facility (TDMF) and IMS Extended Recovery Facility (XRF).

HyperSwap and TDMF coexistence

To minimize disruption to production workloads and service levels, many enterprises use TDMF for storage subsystem migrations and other disk relocation activities. The migration process is transparent to the application, and the data is continuously available for read and write activities throughout the migration process.

However, the HyperSwap function is mutually exclusive with software that moves volumes around by switching UCB pointers. The good news is that currently supported versions of TDMF and GDPS allow operational coexistence. With this support, TDMF automatically temporarily disables HyperSwap as part of the disk migration process only during the short time where it switches UCB pointers. Manual operator interaction is not required. Without this support, through operator intervention, HyperSwap is disabled for the entire disk migration, including the lengthy data copy phase.

HyperSwap and IMS XRF coexistence

HyperSwap also has a technical requirement that RESERVEs cannot be allowed in the hardware because the status cannot be reliably propagated by z/OS during the HyperSwap to the new primary volumes. For HyperSwap, all RESERVEs must be converted to GRS global enqueue through the GRS RNL lists.

IMS/XRF is a facility by which IMS can provide one active subsystem for transaction processing, and a backup subsystem that is ready to take over the workload. IMS/XRF issues hardware RESERVE commands during takeover processing, which cannot be converted to global enqueues through GRS RNL processing. This coexistence problem also was resolved so that GDPS is informed before IMS issuing the hardware RESERVE, allowing it to automatically disable HyperSwap. After IMS finishes processing and releases the hardware RESERVE, GDPS is again informed and reenables HyperSwap.

4.6.2 GDPS HM reduced impact on initial copy and resynchronization

Performing Metro Mirror copy of a large amount of data across many devices while the same devices are used in production by application workloads can potentially affect production I/O service times when those copy operations are performed synchronously. Your disk subsystems and PPRC link capacity are typically sized for steady state update activity but not for bulk, synchronous replication. Initial copy of disks and resynchronization of disks are examples of bulk copy operations that can affect production if performed synchronously.

There is no need to perform initial copy or resynchronizations by using synchronous copy because the secondary disks cannot be made consistent until all disks in the configuration reach the duplex state.

GDPS supports initial copy and resynchronization by using asynchronous PPRC-XD (also known as Global Copy). When GDPS initiates copy operations in asynchronous copy mode, GDPS monitors progress of the copy operation. When the volumes are near full duplex state, GDPS converts the replication from the asynchronous copy mode to synchronous PPRC. Performing the initial copy or resynchronization by using PPRC-XD eliminates the performance impact of synchronous mirroring on production workloads.

Without asynchronous copy, it might be necessary to defer these operations or reduce the number of volumes that are copied at any time, which might delay the mirror from reaching a duplex state, impacting a client's ability to recovery. Using the XD-mode asynchronous copy allows clients to establish or resynchronize mirroring during periods of high production workload, and can potentially reduce the time during which the configuration is exposed.

This function requires that all disk subsystems in the GDPS configuration support PPRC-XD.

4.6.3 Reserve Storage Pool

Reserve Storage Pool (RSP) is a type of resource that was introduced with the z/OS Management Facility (z/OSMF) that can simplify the management of defined but unused volumes. GDPS support includes RSP volumes in the Metro Mirror configuration that are managed by GDPS. Metro Mirror primary volumes are expected to be online in controlling systems. GDPS monitoring on the GDPS controlling systems results in an alert being raised for any Metro Mirror primary device that is found to be offline. However, because z/OS does not allow RSP volumes to be brought online to any system, GDPS monitoring recognizes that an offline primary device is an RSP volume and suppresses alerting for these volumes.

4.6.4 Concurrent Copy cleanup

The DFSMS Concurrent Copy (CC) function uses a "sidefile" that is kept in the disk subsystem cache to maintain a copy of changed tracks that are not yet copied. For a Metro Mirrored disk, this sidefile is not mirrored to the secondary subsystem. If you perform a HyperSwap while a CC operation is in progress, the application that uses CC fails after the completion of the HyperSwap. GDPS does not allow a planned swap when a CC session exists against your primary Metro Mirror devices. However, unplanned swaps are still allowed. Therefore, if you plan to use HyperSwap for primary disk subsystem failures (unplanned HyperSwap), try to eliminate any use of CC because you cannot plan when failures occur.

Checking for CC is performed by GDPS immediately before performing a planned HyperSwap. SDF trace entries are generated if one or more CC sessions exist, and the swap command ends with no Metro Mirror device pairs swapped. Identify and end any CC sessions against the Metro Mirror primary devices before the swap.

When attempting to resynchronize your disks, checking is performed to ensure that the secondary devices do not retain CC status from the time when they were primary devices. These devices are not supported as Metro Mirror secondary devices. Therefore, GDPS does not attempt to establish a duplex pair with secondary devices if it detects a CC session.

GDPS can discover and terminate CC sessions that might otherwise cause errors. The function is controlled by a keyword that provides options to disable, to conditionally enable, or to unconditionally enable the cleanup of CC sessions on the target disks. This capability eliminates the manual task of identifying and cleaning up orphaned CC sessions.

4.7 Autonomous Path Reconfiguration

The GDPS Autonomous Path Reconfiguration feature circumvents potential problems during the run of a START SECONDARY command, which is used to establish PPRC paths for disk replication (asynchronous and synchronous).

In a configuration where this feature is not enabled, a START SECONDARY command stops immediately if one of the PPRC links defined to GDPS is not operational.

With this feature enabled, if GDPS detects a failure while attempting to establish PPRC paths, it tries to collect information from the path in error and remove the bad links from the configuration, and then re-attempts to establish the path. If GDPS is not successful in establishing the PPRC paths after the autonomous reconfiguration, it ends the START SECONDARY command with RC=8 and generates a detailed report in the NetView log for investigation.

4.8 GDPS Metro HyperSwap Manager flexible testing and resync protection

Configuring PiT copy (FlashCopy) capacity in your Metro Mirror environment provides the following main benefits:

► You can conduct regular DR drills or other tests by using a copy of production data while production continues to run.

► You can save a consistent, "golden" copy of the Metro Mirror secondary data, which can be used if the primary disk or site is lost during a Metro Mirror resynchronization operation.

FlashCopy and the various options that are related to FlashCopy are discussed in 2.6, "FlashCopy" on page 32. GDPS Metro HyperSwap Manager supports taking a FlashCopy of the current secondary CKD disks. The COPY, NOCOPY, and NOCOPY2COPY options are supported. CONSISTENT FlashCopy is supported with COPY and NOCOPY FlashCopy.

In addition, FlashCopy can be used to provide a consistent PiT copy of production data to be used for nondisruptive testing of your system and application recovery procedures. FlashCopy can also be used, for example, to back up data without the need for extended outages to production systems; to provide data for data mining applications; and for batch reporting and other uses.

4.8.1 Using space-efficient FlashCopy volumes

As discussed in "Space-efficient FlashCopy" on page 34, by using space-efficient volumes, you might be able to lower the amount of physical storage that is needed and thus reduce the cost that is associated with providing a tertiary copy of the data. GDPS supports FlashCopy SE volumes as FlashCopy target disk volumes. Whether a target device is space-efficient or not is transparent to GDPS; if any of the FlashCopy target devices defined to GDPS are space-efficient volumes, GDPS uses them. All GDPS FlashCopy operations with the NOCOPY option, through panels or by using the FLSHCOPY command or FlashCopies automatically taken by GDPS can use space-efficient targets.

space-efficient volumes are ideally suited for FlashCopy targets when used for resync protection. The FlashCopy is taken before the resync and can be withdrawn when the resync operation is complete. As changed tracks are sent to the secondary for resync, the time-zero (T0) copy of this data is moved from the secondary to the FlashCopy target device. This means that the total space requirement for the targets is equal to the number of tracks that were out of sync, which is typically going to be less than a full set of fully provisioned disks.

Another potential use of space-efficient volumes is if you want to use the data for limited DR testing.

Understand the characteristics of FlashCopy SE to determine whether this method of creating a PiT copy satisfies your business requirements. For example, is it acceptable to your business if, because of some unexpected workload condition, the repository on the disk subsystem for the space-efficient devices gets full and your FlashCopy is invalidated so that you are unable to use it? If your business requirements dictate that the copy must always be guaranteed to be usable, space-efficient might not be the best option and you can consider using standard FlashCopy instead.

4.9 GDPS tools for GDPS Metro HyperSwap Manager

GDPS HM also includes tools that provide functions that are complementary to GDPS functions. The tools represent the functions that all or many clients are likely to develop themselves to complement GDPS. Using tools eliminates the necessity for you to develop similar functions yourself. The tools are provided in source code format, which means that if the tool does not meet your requirements completely, you can modify the code to tailor it to your needs.

The following tools are available with GDPS Metro HyperSwap Manager:

▶ GDPS XML Conversion (GeoXML) Tool is a tool that helps you to convert a GDPS/PPRC (or GDPS/HM) GEOPARM configuration definition file for a single replication leg to XML format GEOPARM definitions. This tool simplifies the task of defining the Metro Mirror configuration for GDPS/PPRC (or GDPS/HM) clients that are moving to the use of GDPS Metro environment.

▶ GDPS EasyLog Tool is a Windows-based tool that is intended to help you extract and easily download the Syslog and Netlog from a z/OS environment. It also provides help in analyzing the Netlog when downloaded onto a workstation.

PMT is intended to simplify and automate to a great extent the process of bringing new devices to Metro Mirror duplex state. It also adds these devices to your running GDPS environment, while reducing the time during which the GDPS managed Metro Mirror is not full-duplex (and therefore not protected by Freeze and HyperSwap) to a minimum. PMT also provides facilities to aid with migration procedures when Global Copy (PPRC-XD) and Metro Mirror are used to migrate data to new disk subsystems.

▶ The GDPS Security Definition Utility tool helps you to implement enhancements to GDPS options that use role-based security by helping with the definition of appropriate profiles in the XFACILIT class of RACF, and by assigning relevant access to them.

4.10 Services component

GDPS touches on more than remote copy. It also includes automation, testing processes, DR processes, and other areas.

Most installations do not have skills in all these areas readily available. And it is rare to find a team that has this range of skills across many implementations. However, the GDPS Metro HyperSwap Manager offering includes exactly that: access to a global team of specialists in all the disciplines you need to ensure a successful GDPS Metro HyperSwap Manager implementation.

Specifically, the Services component includes some or all of the following services:

► Planning to determine availability requirements, configuration recommendations, implementation and testing plans

► Assistance in defining RPOs

► Installation and necessary customization of the special GDPS Metro HyperSwap Manager versions of NetView and System Automation

► Remote copy implementation

► GDPS Metro HyperSwap Manager automation code installation and policy customization

► Education and training on GDPS Metro HyperSwap Manager setup and operations

► Onsite implementation assistance

► Project management and support throughout the engagement

GDPS Metro HyperSwap Manager projects are typically much smaller than projects for the other GDPS offerings. Nevertheless, the sizing of the services component of each project can be tailored for that project based on many factors including what automation is already in place, whether remote copy is already in place, or whether the two centers are already in place with a multisite sysplex if required. The skills that are provided are tailored to the specific needs of each implementation.

4.11 GDPS Metro HyperSwap Manager prerequisites

For more information about the latest GDPS Metro HyperSwap Manager prerequisites, see this GDPS web page.

4.12 Comparing GDPS Metro HyperSwap Manager to other GDPS offerings

So many features and functions are available in the various members of the GDPS family that recalling them all and remembering which offerings support them is sometimes difficult. To position the offerings.

Table 4-1 lists the key features and functions and indicates which ones are delivered by the various GDPS offerings.

Table 4-1   Supported features matrix

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
Continuous availability (CA)	Yes	Yes	Yes	No
DR	Yes	Yes	Yes	Yes
CA/DR protection against multiple failures	Yes	No	No	No
CA for foreign z/OS systems	Yes with z/OS proxy	No	No	No

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
Supported distance	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	Virtually unlimited
Consistent FlashCopy support	Yes, using CONSISTENT	Yes, using CONSISTENT for secondary only	No	Yes, using CGPause
Reduced impact on initial copy/resync	Yes	Yes	Yes	Not applicable
Tape replication support	Yes	No	No	No
Production sysplex automation	Yes	No	Not applicable	No
Span of control	Both sites	Both sites (disk only)	Both sites	Disk at both sites; recovery site (Capacity BackUp (CBU) or LPARs)
GDPS scripting	Yes	No	Yes	Yes
Monitoring, alerting and health checks	Yes	Yes	Yes (except health checks)	Yes
Query Services	Yes	Yes	No	Yes
MSS support for added scalability	Yes (RS2 in MSS1, RS3 in MSS2)	Yes (secondary in MSS1)	No	Yes (Global Mirror (GM) FlashCopy and Primary for Metro Global Mirror (MGM) in MSS1)
MGM 3-site and 4-site	Yes (all configurations)	Yes (3-site only and non-Incremental Resynchronizati on (IR) only)	No	Yes (all configurations)
FB disk	Yes	Yes	No	Yes
z/OS equivalent function for Linux on IBM Z	Yes (Linux on IBM Z running as a z/VM guest only)	No	Yes (Linux on IBM Z running as a z/VM guest only)	Yes
GDPS GUI	Yes	Yes	Yes	Yes

4.13 Summary

GDPS Metro HyperSwap Manager is a powerful offering that can extend Parallel Sysplex availability to disk subsystems by delivering the HyperSwap capability to mask planned and unplanned disk outages. It also provides monitoring and management of the data replication environment, including the freeze capability. It can provide these capabilities either in a single site, or when the systems and disks are spread across two data centers within metropolitan distances.

In a multisite configuration, GDPS Metro HyperSwap Manager can also be an entry-level offering, which can provide ZDL. The RTO is typically longer than what can be obtained with a full GDPS Metro offering. As time goes by, if your business needs to migrate from GDPS Metro HyperSwap Manager to the full GDPS Metro offering, this task can be achieved also.

In addition to DR and CA capabilities, GDPS Metro HyperSwap Manager provides a simple interface for monitoring and managing the remote copy configuration.

IBM GDPS Global - GM

In this chapter, we discuss the capabilities and prerequisites of the GDPS Global - GM (GM) offering.

The GDPS GM offering provides a disaster recovery (DR) capability for businesses that have an RTO of as little as one hour, and a recovery point objective (RPO) as low as five seconds. It is often deployed in configurations where the application and recovery sites are more than 200 km (124 miles) apart and want to have integrated remote copy processing for mainframe and non-mainframe data.

The functions that are provided by GDPS GM fall into the following categories:

► Protecting your data:

– Protecting the integrity of the data on the secondary data in a disaster or suspected disaster.

– Managing the remote copy environment through IBM Geographically Dispersed Parallel Sysplex (GDPS) scripts and NetView panels or the web interface.

– Optionally supporting remote copy management and consistency of the secondary volumes for Fixed-Block (FB) data. Depending on your application requirements, the consistency of the FB data can be coordinated with the Count-Key-Data (CKD) data.

► Controlling the disk resources that are managed by GDPS during normal operations, planned changes, and following a disaster:

– Support for recovering the production environment after a disaster.

– Support for switching your data and systems to the recovery site.

– Support for testing recovery and restart by using a practice FlashCopy point-in-time (PiT) copy of the secondary data while live production continues to run in the application site and remains protected with the secondary copy.

This chapter includes the following topics:

5.1 Introducing GDPS Global - GM

GDPS GM is a DR solution that supports virtually unlimited distances between the application and recovery sites. The underlying IBM Global Mirror (GM) remote copy technology is a disk subsystem-based remote copy technology that requires that the primary and secondary disk subsystems are from the same vendor. GM supports both IBM Z CKD data and distributed data, and GDPS GM also includes support for both.

Unlike GDPS GM, GDPS GM does not provide any automation or management of the production systems. Instead, its focus is on managing the Global Mirror remote copy environment and automating and managing recovery of data and systems in the case of a disaster.

GDPS GM supports remote copy data from multiple systems and sysplexes.

The capabilities and features of GDPS GM are described in this chapter.

5.1.1 Protecting data integrity

Because the role of GDPS GM is to provide DR support, its highest priority is protecting the integrity of the data, CKD and FB, in the recovery site. This section discusses the support that is provided by GDPS for these various data types.

Traditional IBM Z (CKD) data

As described in 2.4.2, "Global Mirror" on page 26, GM protects the integrity of the remote-copied data by creating consistency groups, continuously or at intervals that are specified by the installation. The process is managed by the Master disk subsystem, which is based on the GDPS GM configuration.

There are no restrictions relating to which operating systems' data can be supported; any system that writes to CKD devices (z/OS, z/VM, z/VSE, and Linux for System z) is supported. Regardless of which systems are writing to the devices, all management control is from the z/OS system that is running the GDPS GM local controlling system, also known as the *K-sys*.

How frequently a consistency group can be created depends on the bandwidth that is provided between the application and recovery site disks. IBM can perform a bandwidth analysis for you to help you identify the required capacity.

GDPS Global - GM uses devices in the primary and secondary disk subsystems to run the commands to manage the environment. Some of these commands directly address a primary device, whereas others are directed to the logical subsystem (LSS). To run these LSS-level commands, you must designate at least one volume in each primary LSS as a GDPS *utility device*, which is the device that serves as the "go-between" between GDPS and the LSS. These utility devices do not need to be dedicated devices; that is, they can be one of the devices that are being mirrored as part of your GM session. In fact, the utility devices also need to be mirrored.

Distributed (FB) data

GDPS GM provides the capability to manage a heterogeneous environment of IBM Z and distributed systems data. Through a function called FB Disk Management, GDPS GM can manage the GM remote copy configuration and FlashCopy for distributed systems data. The GM remote copy technology (see 2.4.2, "Global Mirror" on page 26) inherently provides data consistency for IBM Z and distributed systems data.

The FB devices can be in the same GM session as the CKD devices or in a separate session. If the FB devices and CKD devices are in the same session, they have the same consistency point and they must be recovered together. If they are in a different session, they have a different consistency point (the data for each session is consistent within itself, but the data for the two sessions is inconsistent with each other) and can be recovered separately.

FB Disk Management prerequisites

GDPS requires that the disk subsystems that contain the FB devices to support specific architectural features. The following architectural features are supported by all IBM disk subsystems:

▶ The ability to manage FB devices through a CKD utility device

GDPS runs on z/OS and can communicate GM commands to the disk subsystem directly over a channel connection to CKD devices only. To communicate commands to the FB LSS and devices, the architecture allows the use of a CKD utility device in the same disk subsystem as a go-between to send commands and to monitor and control the mirroring of the FB devices. GDPS needs at least one CKD utility device in each hardware cluster of the storage subsystem where FB devices are located.

▶ The ability to send SNMP traps to report certain errors

The FB LSS and devices must communicate certain error conditions back to GDPS (for example, an abnormal state of a GM session in GDPS GM). This status is communicated to the z/OS host that is running the GDPS that is controlling the system through an IP connection by using SNMP traps. GDPS captures these traps and drives autonomic action, such as performing a freeze for a mirroring failure.

A sample FB device GDPS GM configuration is shown in Figure 5-1. Not shown are the IP connections from the attached disks to the z/OS host where GDPS is running.
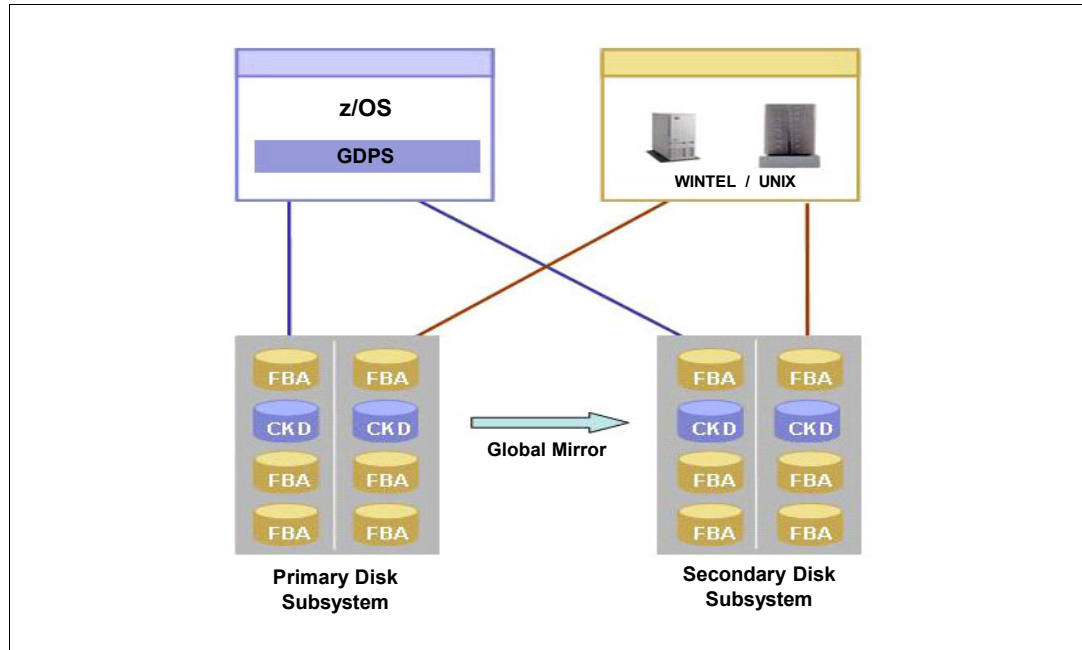
Figure 5-1   Fixed-Block disk support

5.2 GDPS Global - GM configuration

At its most basic, a GDPS GM configuration consists of one or more production systems, an application site controlling system (K-sys), a recovery site controlling system (R-sys), primary disks, and two sets of disks in the recovery site.

The GM copy technology uses three sets of disks. For more information about how GM works and how the disks are used to provide data integrity, see 2.4.2, "Global Mirror" on page 26.

The K-sys is responsible for controlling all remote copy operations and for sending configuration information to the R-sys. In normal operations, most operator and system programmer interaction with GDPS GM is through the K-sys.

The K-sys role is related to remote copy; it does not provide any monitoring, automation, or management of systems in the application site, or any FlashCopy support for application site disks. There is no requirement for the K-sys to be in the same sysplex as the system or systems for which it is managing data. In fact, the K-sys is placed in a monoplex on its own.

You can also include the K-sys disks in the GDPS managed GM configuration and replicate them. The K-sys does not have the isolation requirements of the controlling system in a GDPS Metro configuration.

The R-sys is primarily responsible for validating the configuration, monitoring the GDPS managed resources (such as the disks in the recovery site), and carrying out all recovery actions for test purposes or if a real disaster occurs. For more information about testing by using FlashCopy, see 5.9, "Flexible testing and Logical Corruption Protection" on page 173.

The K-sys and R-sys communicate information to each other by using a NetView-to-NetView network communication mechanism over the wide area network (WAN). K-sys and R-sys are dedicated to their roles as GDPS controlling systems.

GDPS GM can control multiple GM sessions. Each session can consist of a maximum of 17 disk subsystems (combination of primary and secondary). All the members of the same session have the same consistency point.

Typically, the data for all systems that must be recovered together is managed through one session. For example, a z/OS sysplex is an entity where the data for all systems in the sysplex must be in the same consistency group.

If you have two production sysplexes under GDPS GM control, the data for each can be managed through a separate GM session, in which case they can be recovered individually. You can also manage the entire data for both sysplexes in a single GM session, in which case if one sysplex fails and you must start recovery, you also must recover the other sysplex.

Information about which disks are to be mirrored as part of each session and the intervals at which a consistency point is to be created for each session is defined in the GDPS remote copy configuration definition file (GEOMPARM). GDPS GM uses this information to control the remote copy configuration. As with the other GDPS offerings, the NetView panel interface (or the web interface) is used as the operator interface to GDPS.

Although the panel interface or web interface support management of GM, they are primarily intended for viewing the configuration and performing some operations against single disks. GDPS scripts are used for actions against the entire configuration because this is simpler (with multiple panel actions that are combined into a single script command) and less error-prone.

The actual configuration depends on your business and availability requirements, the amount of data you are remote copying, the types of data you are remote copying (only CKD or both CKD and FB), and your RPO.

Figure 5-2 shows a typical GDPS GM configuration.
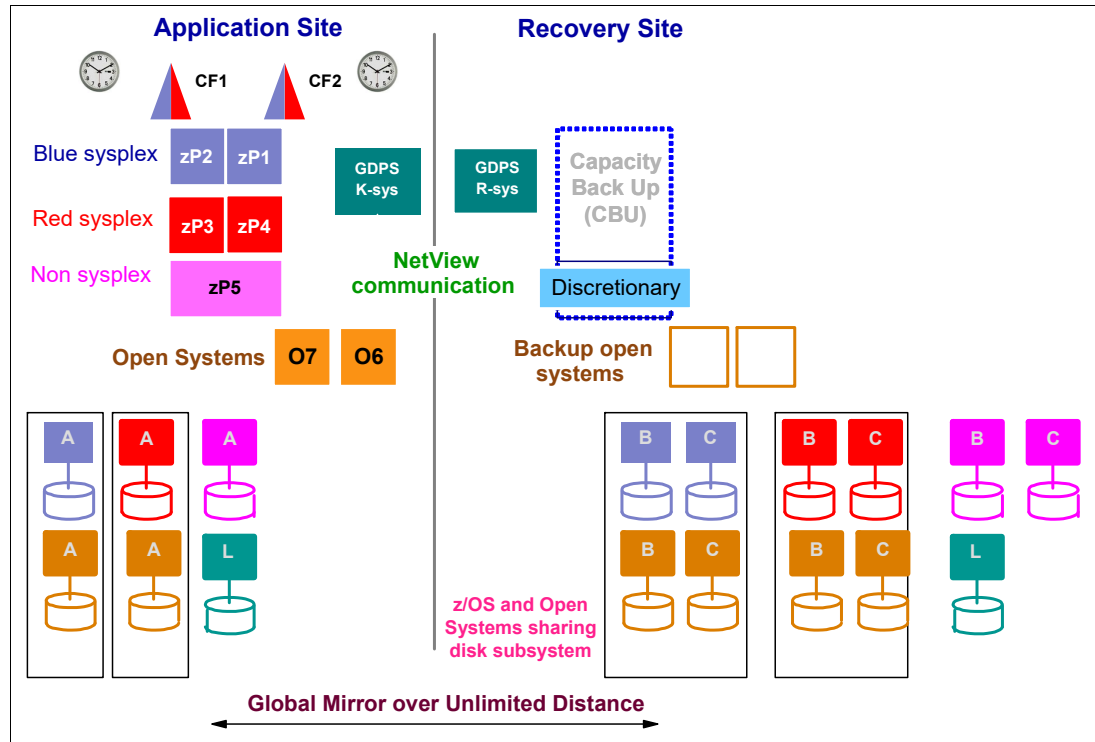
Figure 5-2   GDPS GM configuration

The application site, as shown in Figure 5-2, features the following items:

- ► z/OS systems spread across several sysplexes
- ► A non-sysplexed z/OS system
- ► Two distributed systems
- ► The K-sys
- ► The primary disks (identified by A)
- ► The K-sys' own disks (marked by L)

The recovery site includes the following items:

- ► The R-sys

- ► A Central Processor Complex (CPC) with the Capacity BackUp (CBU) feature that also contains expendable workloads that can be displaced

- ► Two backup distributed servers

- ► The GM secondary disks (marked by B)

- ► The GM FlashCopy targets (marked by C)

- ► The R-sys disks (marked by L)

Although there is great flexibility in terms of the number and types of systems in the application site, several items are fixed. Consider the following points:

- ► All the GM primary disks and the K-sys *must* be in the application site[1].

- ► All the GM secondary disks, the FlashCopy targets used by GM, and the GDPS R-sys *must* be in the recovery site[2].

---

[1] The application site is where production applications whose data is to be mirrored normally run, and it is the site where the GM primary disks are. You might also see this site referred to as the *local site* or the *A-site*.

With GDPS GM, although the K-sys should be dedicated to its role as a controlling system, it is not necessary to provide as much isolation for the K-sys as required in a GDPS Metro or GDPS HM configuration.

If there is insufficient bandwidth for GM, the consistency points fall behind, which means that the RPO might not be achieved, but there is no performance degradation that is associated with the primary devices (application performance is not affected). If you want to protect RPO in a GM environment, you must provide sufficient bandwidth to handle the peak write load.

The GDPS GM code runs under NetView and System Automation, and is run in the K-sys and R-sys *only*.

GDPS GM multiple R-sys colocation

GDPS GM can support multiple sessions. Therefore, the same instance of GDPS GM can be used to manage GM replication and recovery for several diverse sysplexes and systems. However, there are certain cases where different instances of GDPS GM are required to manage different sessions. One example is the GDPS GM leg of a GDPS MGM configuration: in such a configuration, GDPS GM is restricted to managing only one session. Clients might have other requirements that are based on workloads or organizational structure for isolating sessions to be managed by different instances of GDPS GM.

When you have multiple instances of GDPS GM, each instance needs its own K-sys. However, the R-sys "functions" of each instance can be combined to run in the same z/OS image. Each R-sys function runs in a dedicated NetView address space in the same z/OS. Actions, such as running scripts, can be done simultaneously in these NetView instances. This ability reduces the overall cost of managing the remote recovery operations for customers that require multiple GDPS GM instances.

5.2.1 GDPS GM in a 3-site or 4-site configuration

GDPS GM can be combined with GDPS Metro (or GDPS HM) in a 3-site or 4-site configuration, where GDPS Metro (or GDPS HM) is used across two sites within metropolitan distances (or even within a single site) to provide continuous availability (CA) through Parallel Sysplex use and GDPS HyperSwap. GDPS GM provides DR in a remote region.

We call this combination the GDPS Metro Global - GM (GDPS MGM) configuration. In such a configuration, GDPS Metro and GDPS GM provide more automation capabilities.

After you understand the base capabilities that are described in 2.4.3, "Combining disk remote copy technologies for CA and DR" on page 29, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237 for more information about GDPS MGM.

---

[2] The recovery site is where the mirrored copies of the production disks are located, and it is the site to which production systems are failed over in a disaster. You might also see this site referred to as the *remote site* or the *R-site*.

5.2.2 Other GM considerations

The availability of the GDPS K-sys in *all* scenarios is a fundamental requirement in GDPS. The K-sys monitors the remote copy process, implements changes to the remote copy configuration, and sends GDPS configuration changes to the R-sys.

Although the main role of the R-sys is to manage recovery after a disaster or to enable DR testing, it is important that the R-sys also is available always because the K-sys sends changes to GDPS scripts and changes to the remote copy or remote site configuration to the R-sys at the time the change is introduced on the K-sys. If the R-sys is not available when such configuration changes are made, it is possible that it might not have the latest configuration information in a subsequent disaster, resulting in an impact to the recovery operation.

Also, the R-sys plays a role in validating configuration changes. Therefore, it is possible that a change that contains errors that are rejected by the R-sys (if it was running) are not detected. This issue affects the remote copy or recovery operation.

Although GDPS GM is in essence a DR offering rather than a CA offering, it is enhanced to provide the Region Switch capability so that customers may run production in the recovery site can do so without introducing gaps in the resiliency capabilities while running in this alternative configuration. They can switch back by leveraging GDPS functions.

If you want to perform a Region Switch from the primary site to a recovery site and then return to the primary site, set up GDPS GM in the opposite direction (which means that you also need two sets of disks in the application site), and set up alternative GDPS configurations that activate on switching from the primary site to a recovery site.

Starting with GDPS Global 4.7, there is a separately licensed feature that is called GDPS GM Bidirectional, which enhances the ability to switch back and forth with new GDPS functions. For more information, see 5.4.2, "GDPS GM Bidirectional support" on page 168

5.3 Managing the GDPS environment

GDPS Global - GM automation code runs in one system in the application site only (the K-sys) and it does not provide for any monitoring or management of the production systems in this site. The K-sys has the following responsibilities:

► It is the primary point of GDPS GM control for operators and system programmers in normal operations.

► It manages the remote copy environment. Changes to the remote copy configuration (adding new devices into a running GM session or removing devices from a running session) are driven from the K-sys.

► Changes to the configuration definitions or scripts (including configuration definitions for recovery site resources and scripts that will run on the R-sys) are defined in the K-sys and automatically propagated to the R-sys.

In the recovery site, GDPS GM runs only in one system: the R-sys. However, the role and capabilities of the R-sys are different from the ones of the K-sys. Although both are GDPS controlling systems, there are fundamental differences between them. The R-sys has the following responsibilities:

► Validate the remote copy configuration in the remote site. This responsibility is a key role. GM is a hardware replication technology. Just because the GM primary disks can communicate to the GM secondary disks over remote copy links does not mean that in a recovery situation, systems can use these disks. The disks must be defined in that site's I/O configuration. If you are missing some disks, this issue can cause recovery to fail because you cannot properly restart systems that need those disks.

► Monitor the GDPS-managed resources in the recovery site and raise alerts for not-normal conditions. For example, GDPS uses the BCP Internal Interface (BCPii) for performing hardware actions, such as adding temporary CBU capacity to CPCs, deactivating LPARs for discretionary workloads, and activating LPARs for recovery systems. The R-sys monitors that it has BCPii connectivity to all CPCs that it must perform actions against.

► Communicate status and alerts to the K-sys that is the focal management point during normal operations.

► Automate reconfiguration of the recovery site (recovering the GM, taking a FlashCopy, activating CBU, activating backup partitions, and other actions) for recovery testing or in a true disaster.

The R-sys has no relation to any application site resources. The only connection that it has to the application site is the network connection to the K-sys for exchanging configuration and status information.

5.3.1 User interfaces

The operator interface for GDPS GM is provided through NetView 3270 panels or a .browser-based GUI, which is also referred to as the GDPS GUI (see "GDPS GUI" on page 156). In normal operations, the operators interact mainly with the K-sys, but there is also a similar set of interfaces for the R-sys.

> **Note:** The GDPS GUI that is described in this chapter is new and replaces the former GDPS Web GUI that was described in the previous releases of this book. The GDPS Web GUI (which was based on the NetView Web Application) was removed from the GDPS solution V4R1.

The NetView interface for GDPS consists of two parts. The first and potentially the most important part is the Status Display Facility (SDF). GDPS sends an alert to SDF whenever there is a change of status to something that GDPS does not consider "normal" and that can affect the ability to recover so it is something that requires investigation and manual intervention.

SDF provides a dynamically updated color-coded panel that provides the status of the systems and highlights any problems in the remote copy configuration. If something changes in the environment that requires attention, the color of the associated field on the panel changes. K-sys sends alerts to the R-sys and R-sys sends alerts to K-sys so that both controlling systems are aware of any problems always.

During normal operations, the operators should always have a K-sys SDF panel within view so that they are immediately aware of anything that requires intervention or action. When R-sys is being used for managing testing or recovery operations, operators should also have access to the R-sys SDF panel.

The other part of the NetView interface consists of the panels that are provided by GDPS to help you manage and inspect the environment. The main GDPS panel is shown in Figure 5-3.

```
VPCPPNLM  GDPS Global - GM  Kg-sys (STOCKHOLM)                  GDPS V4.R4.M0
User ID:  SIIK
                       ---- GDPS Status Indicators ----


  Kg-sys          =  MVS4     DSS4C      Mirroring      =  OK
  Debug           =  ON                  Dasd Config    =  2021-03-25  15:03:41
  Topology        =  GM2SITE




                       ---- GDPS Options ----

        1        Dasd Remote Copy
                                      8        Debug ON/OFF
        3        Standard Actions     9        View Definitions
                                      H        Health Checks and Diagnostics
                                      C        Config Management
        6        Planned Actions      M        Run Monitor1/Monitor3


Selection ===>  _
   F1=Help            F3=Return              F6=Roll
```

Figure 5-3   GDPS main panel (K-sys)

From this panel, you can perform the following actions:

► Query and control the disk remote copy configuration.

► Start GDPS standard actions (the ability to control and initiate actions against LPARs).

    On the K-sys, the only standard action that is supported is the ability to update IPL information for the recovery site LPARs.

    On the R-sys, all standard actions are available.

► Start GDPS scripts (Planned Actions).

► Manage GDPS Health Checks.

► View and refresh the definitions of the remote copy configuration.

► Run GDPS monitors.

GDPS GUI

The GDPS GUI is a browser-based interface that improves operator productivity. The GDPS GUI provides the same functional capabilities as the 3270-based panel, such as providing management capabilities for Remote Copy Management, Standard Actions, Sysplex Resource Management, SDF Monitoring, and browsing the CANZLOG, by using simple point-and-click procedures.

Advanced sorting and filtering is available in most of the views that are provided by the GDPS GUI. In addition, users can open multiple windows or tabs to allow for continuous status monitoring while performing other GDPS GM management functions.

The GDPS GUI is available in stand-alone GDPS GM environments and GDPS MGM 3-site and 4-site environments (for more information about GDPS MGM 3-site and 4-site environments, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237).

The GDPS GUI display features the following main sections, as shown in Figure 5-4:

► The application header at the top of the page includes an Actions button with which various GDPS tasks can be performed, along with the Help function and the ability to log off or switch between target systems.

► The application menu is on the left side of the window. This menu gives access to various features and functions that are available through the GDPS GUI.

► The active window shows context-based content, depending on the selected function. This tabbed area is where you can switch context by clicking a different tab.

► A status summary area is shown at the bottom of the display.

Note: For the remainder of this section, only the GDPS GUI is shown to highlight the various GDPS management functions. The equivalent traditional 3270 panels are not shown here.

The initial status window (known as the *dashboard*) of the GDPS Global - GM GUI is shown in Figure 5-4. This window provides an instant view of the status and direction of replication, and disks and systems availability. Hovering over the various icons provides more information through windows.

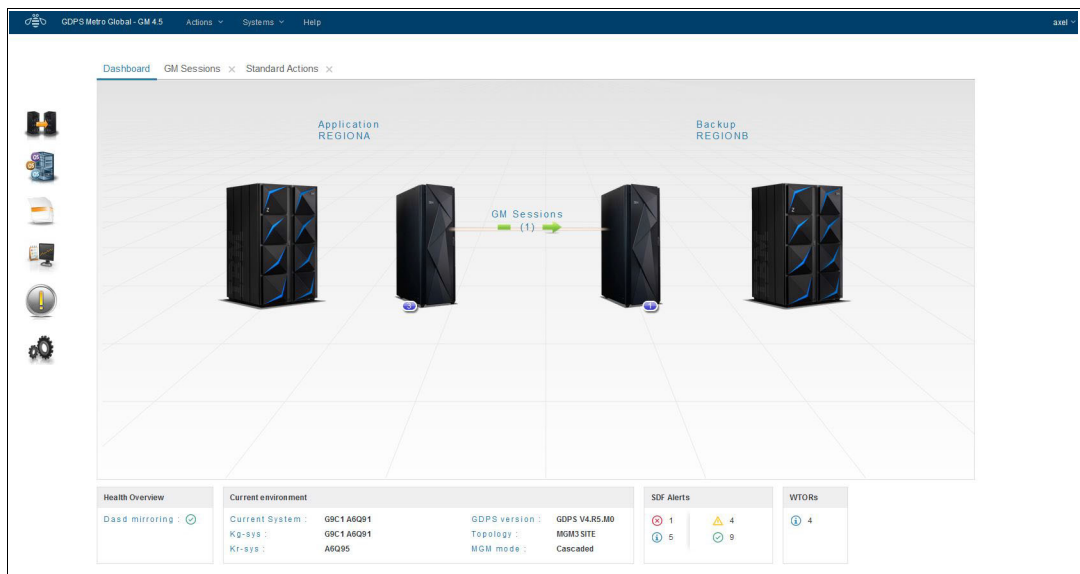*Figure 5-4   GDPS GUI Dashboard (initial window)*

Monitoring function: Status Display Facility

GDPS also provides many monitors to check the status of disks, sysplex resources, and so on. GDPS raises an alert whenever a configuration change occurs, or something in GDPS that requires manual intervention. GDPS uses the SDF that is provided by System Automation as the primary status feedback mechanism for GDPS.

GDPS provides a dynamically updated window, as shown in Figure 5-5. A summary of all current alerts is shown at the bottom of each window.

The initial view that is presented is for the SDF trace entries, meaning that you can follow, for example, script execution. Click one of the icons that represents the other alert categories to view the different alerts that are associated with automation or remote copy in either site, or click **All** to see all alerts. You can sort and filter the alerts based on several fields that are presented, such as severity.
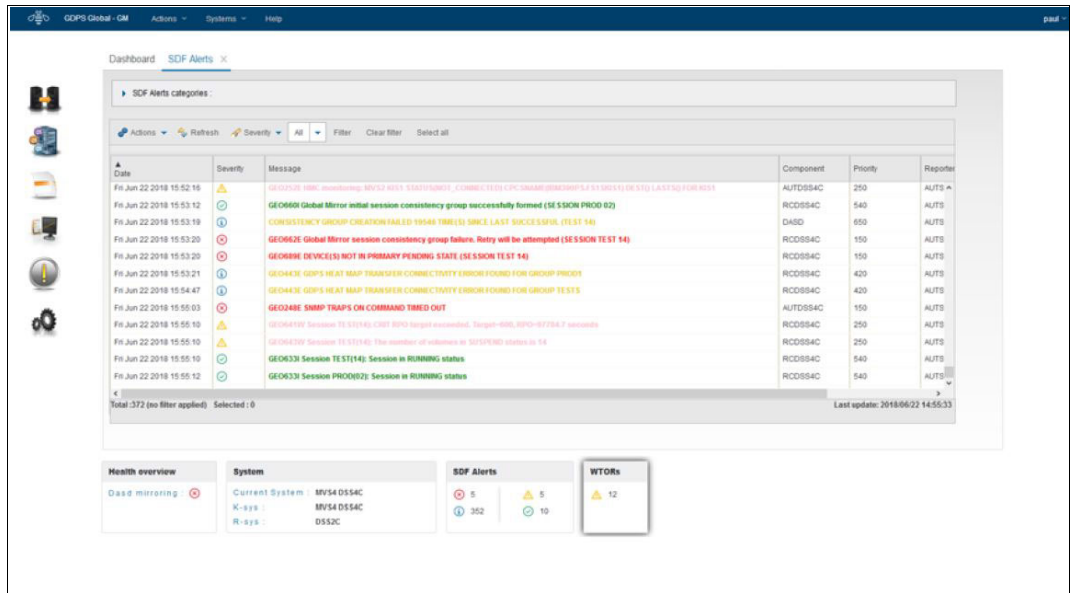
Figure 5-5   GDPS GUI SDF window (Trace entries))

Remote copy operations

Although GM is a powerful copy technology, the z/OS operator interface to it is not intuitive. Use the Disk Remote Copy panels that are provided by GDPS to make it simpler for operators to check and manage the remote copy environment.

For GDPS to manage the remote copy environment, you first define the configuration to GDPS in the GEOMPARM file on the K-sys. The R-sys always receives the configuration information from the K-sys and validates the remote site disk configuration.

After the configuration is known to GDPS, you can use the GUI to check that the current configuration matches the one you want. You can start, stop, pause, and resync mirroring. These actions can be done at the device, LSS, or session level. However, we suggest that GDPS control scripts are used for actions at the session level.

Figure 5-6 on page 159 shows the GM sessions status panel for GDPS GM as viewed on the K-sys. By using the panel, you can review the status of the GM sessions and obtain more information about individual LSS or device pairs if required.
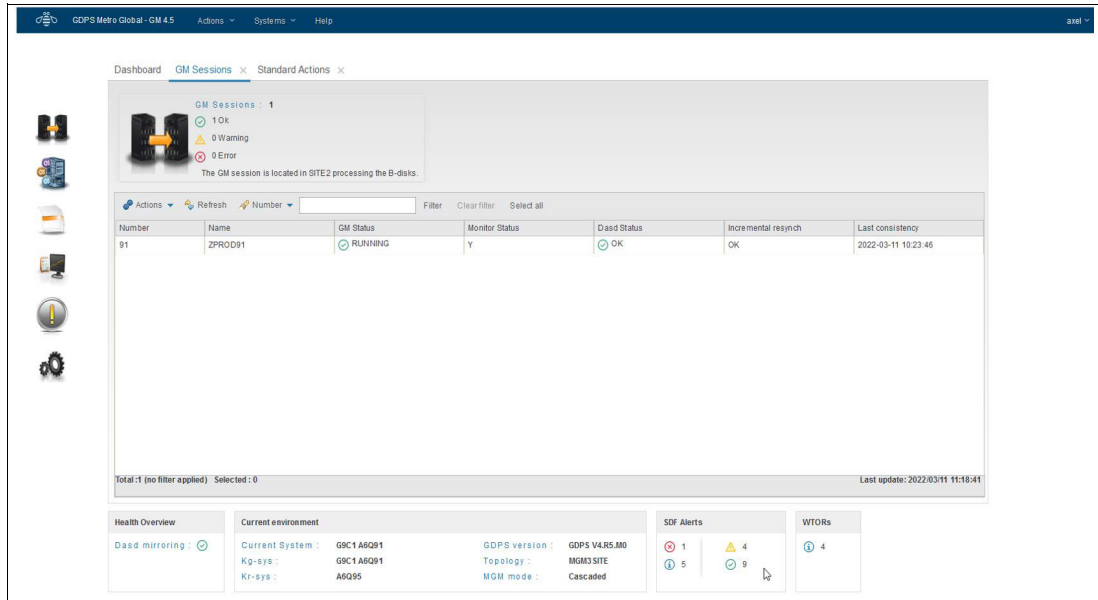
Figure 5-6   GM sessions status panel as viewed on the K-sys

The panel for the R-sys is similar, except that the R-sys can perform only a limited number of actions (typically only those actions that are necessary to take corrective action) against the devices in the recovery site. Control of the GM session can be done from the K-sys only; the R-sys can control only the devices in the recovery site.

Figure 5-7 shows an example of our panel that displays the LSS pairs.
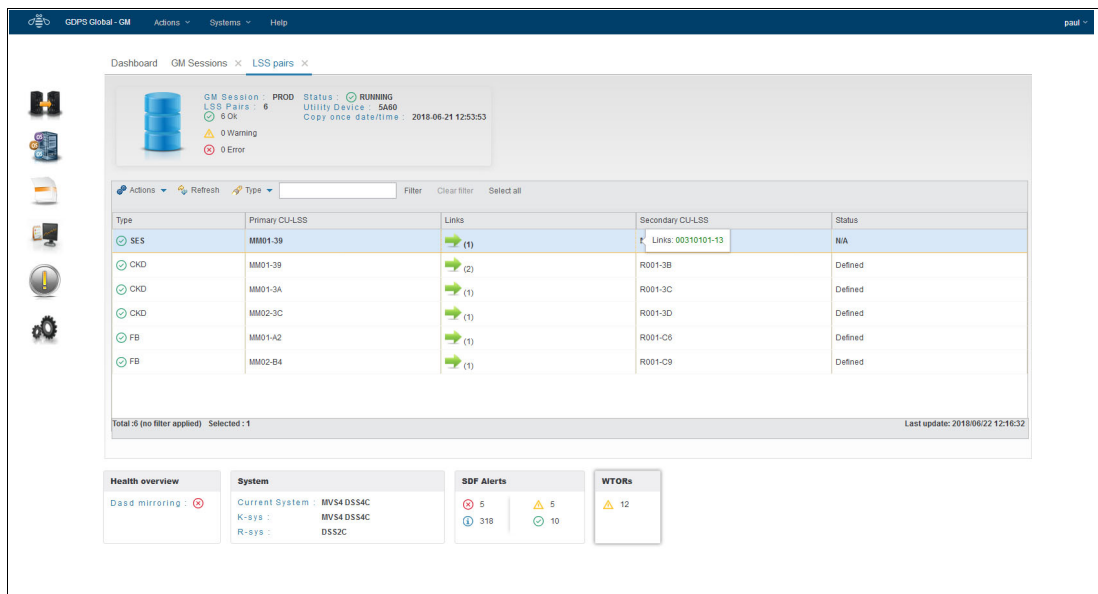
Figure 5-7   Sample panel for the LSS pairs

Figure 5-8 shows the device pairs.
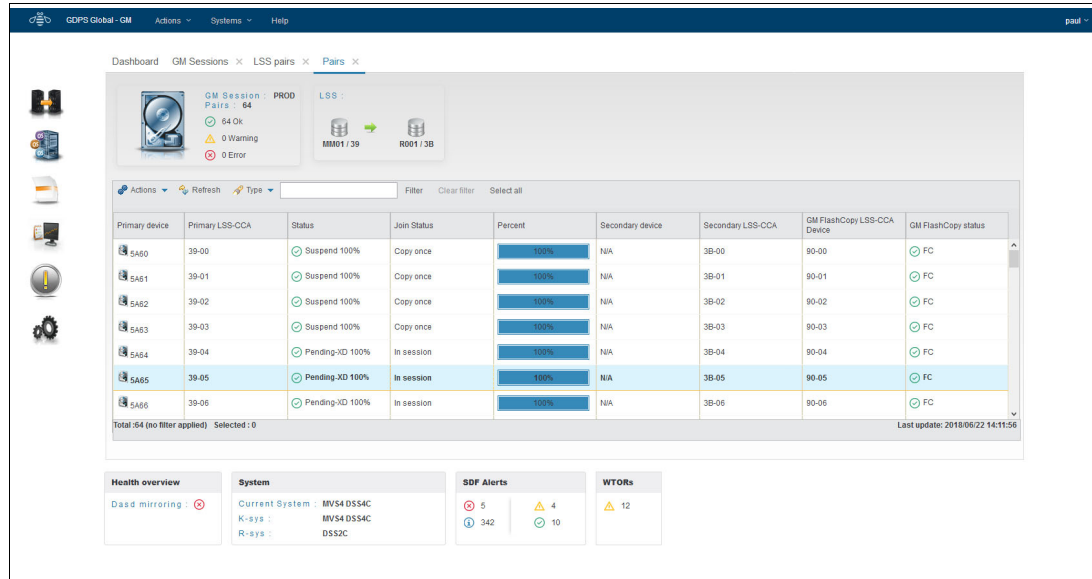
Figure 5-8   Sample panel for the device pairs

The GUI that is provided by GDPS is *not* intended to be a remote copy monitoring tool. Because of the overhead that is involved in gathering information about every device in the configuration to populate the windows, GDPS gathers this data on a timed basis only, or on demand following an operator instruction.

The normal interface for finding out about remote copy problems is the SDF, which is dynamically updated if or when a problem is detected.

Standard Actions

The K-sys does not provide any management functions for any systems in the application site or in the recovery site. The R-sys manages recovery in the recovery site. As a result, the available Standard Actions vary, depending on which type of controlling system you use.

On the K-sys, the only Standard Action that is available is to define the possible IPL address and Loadparms that can be used for recovery systems (production systems when they are recovered in the recovery site) and to select the one to use in a recovery action. Changes that are made on this panel are automatically propagated to the R-sys.

The K-sys Standard Actions panel is shown in Figure 5-9.
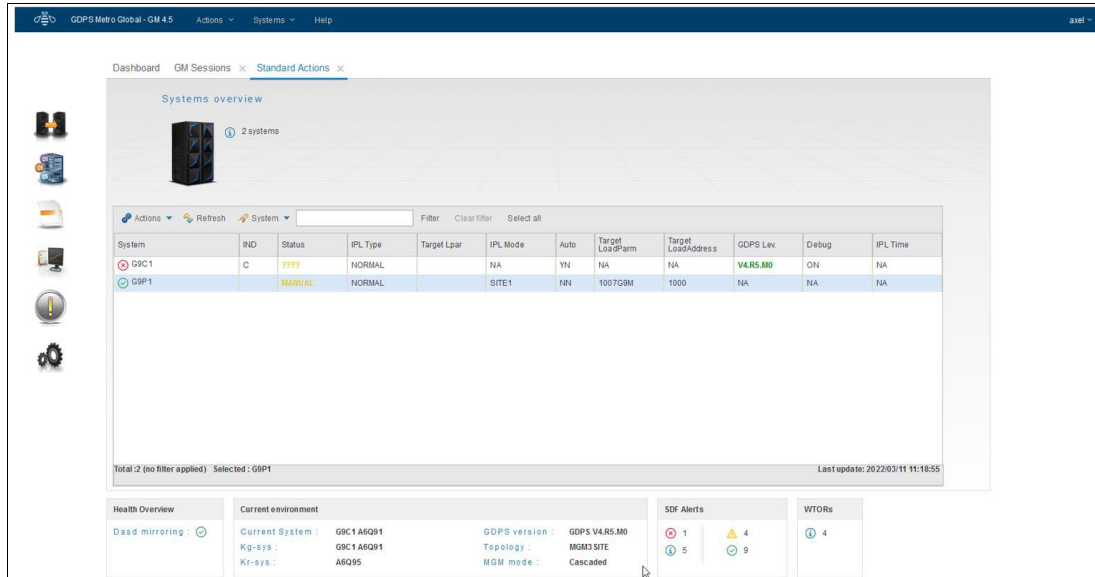
Figure 5-9   GDPS GM K-sys Standard Actions panel

Because the R-sys manages the recovery if a disaster occurs (or IPL for testing purposes) of the production systems in the recovery site, it has a wider range of functions available (see Figure 5-10). Functions are provided to activate and deactivate LPARs, IPL and reset systems, and update the IPL information for each system.

Figure 5-10   Example GDPS GM R-sys Standard Actions panel for a selected system

Standard Actions are single-step actions and are intended to affect only one resource. For example, if you want to reset an expendable test system that is running in the recovery site, deactivate the LPAR of the expendable system, activate the recovery LPAR for a production system, and then, IPL the recovery system into the LPA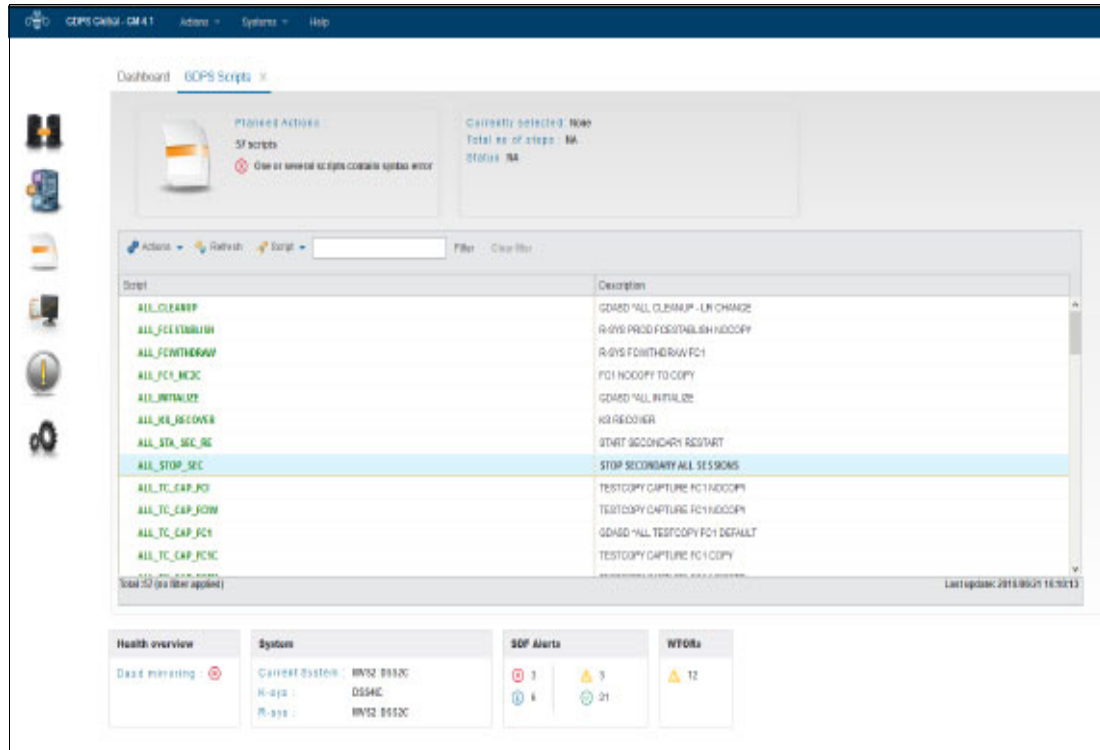R you activated, you start four separate Standard Actions, one after the other. GDPS scripting, as described next, is a facility that is suited to multi-step, multi-system actions.

5.3.2 GDPS scripts

Nearly all the functions that can be started through the panels (and more) are also available from GDPS scripts. A script is a program that consists of one or more GDPS functions to provide a workflow.

In addition to the low-level functions that are available through the panels, scripts can start functions with a single command that might require multiple separate steps if performed through the panels. For example, if you have a new disk subsystem and are adding several LSSs that are populated with many devices to your GM configuration, this process can require a large number of panel actions.

In comparison, this process can be accomplished by using a single script command. It is faster and more efficient to perform compound or complex operations by using scripts.

Scripts can be started manually through the GDPS panels or through a batch job. In GDPS GM, the only way to start the recovery of the secondary disks is through a GDPS script on the R-sys; starting a recovery directly from the mirroring panels is not supported.

Scripts are written by you to automate the handling of certain situations, both planned changes and also error situations. This function is an important aspect of GDPS.

Scripts are powerful because they can access the full capability of GDPS. The ability to start all the GDPS functions through a script provides the following benefits:

► Speed

The script runs the requested actions as quickly as possible. Unlike a human, it does not need to search for the latest procedures or the commands manual.

► Consistency

If you look into most computer rooms immediately following a system outage, what would you see? Likely, mayhem. Operators are frantically scrambling for the latest system programmer instructions. All the phones are ringing. Every manager within reach is asking when the service will be restored. And every system programmer with access is vying for control of the keyboards.

All of this chaos results in errors because humans often make mistakes when under pressure. But with automation, your well-tested procedures run in the same way, time after time, regardless of how much you shout at them.

► Automatic checking of results from commands

Because the results of many GDPS commands can be complex, manual checking of results can be time-consuming and presents the risk of missing something. In contrast, scripts automatically check that the preceding command (remember, that one command can have been six GM commands, each of the six run against thousands of devices) completed successfully before proceeding with the next command in the script.

► Thoroughly tested procedures

Because scripts behave in a consistent manner, you can test your procedures over and over until you are sure that they do everything that you want, in exactly the manner that you want. Also, because you must code everything and cannot assume a level of knowledge (as you might with instructions that are intended for a human), you are forced to thoroughly think out every aspect of the action the script is intended to undertake. Finally, because of the repeatability and ease of use of the scripts, they lend themselves more easily to frequent testing than manual procedures.

Planned Actions

In a GDPS GM environment, all actions affecting the recovery site are considered *planned actions*. You can think of this as pre-planned unplanned actions. GDPS scripts can be started from the Panels (option 6 on the main GDPS panel, as shown in Figure 5-3 on page 156) and from the GUI.

A control script that is running can be stopped if necessary. Control scripts that were stopped or failed can be restarted at any step of the script. These capabilities provide a powerful and flexible workflow management framework.

An example of a planned action in GDPS GM is a script that prepares the secondary disks and LPARs for a DR test.

Such a script performs the following actions:

► Recovers the disks in the disaster site, which makes the B disks consistent with the C disks. The B disks are used for the test and the C disks contain a consistent copy that ages during the test.

► Activates CBU capacity in the recovery site CPCs.

► Activates backup partitions that are predefined for the recovery systems (that is, the production systems that are running in the recovery site).

► Activates any backup coupling facility (CF) partitions in the recovery site.

► Loads the systems into the partitions in the recovery site by using the B disks.

When the test is complete, you run another script in the R-sys to perform the following tasks:

► Reset the recovery systems that were used for the test

► Deactivate the LPARs that were activated for the test.

► Undo CBU on the recovery site CPCs.

► Issue a message to the operators to manually shut down any open systems servers in the recovery site that were used for the test.

► Bring the B disks back into sync with the C disks (which are consistent with the primary disks at the time of the start of the test).

► Finally, you run a script on the K-sys to resynchronize the recovery site disks with the production disks.

Batch scripts

In addition to the ability to start GDPS scripts from the GDPS panel interfaces, a script can be start from outside of GDPS by using a batch interface. These scripts are known as *batch scripts* and they cannot be started from the GDPS panels or GUI. This ability is especially suited to processes that are run regularly, and feature some interaction with the GDPS environment.

5.3.3 Application programming interfaces

GDPS provides two primary programming interfaces to allow other programs that are written by clients, independent software vendors (ISVs), and other IBM product areas to communicate with GDPS. These application programming interfaces (APIs) allow clients, ISVs, and other IBM product areas to complement GDPS automation with their own automation code. The following sections describe the APIs that are provided by GDPS.

Query Services

GDPS maintains configuration information and status information in NetView variables for the various elements of the configuration that it manages. GDPS Query Services is a facility that allows user-written REXX programs that are running under NetView to query and obtain the value of various GDPS variables. This configuration augments GDPS automation with your own automation REXX code for various purposes, such as monitoring or problem determination.

Query Services allows clients to complement GDPS automation with their own automation code. In addition to the Query Services function (which is part of the base GDPS product), GDPS provides several samples in the GDPS SAMPLIB library to demonstrate how Query Services can be used in client-written code.

RESTful APIs

As described in "Query Services" on page 164, GDPS maintains configuration information and status information about the various elements of the configuration that it manages. Query Services can be used by REXX programs to query this information.

The GDPS RESTful API also provides the ability for programs to query this information. Because it is a RESTful API, it can be used by programs that are written in various programming languages, including REXX, that are running on various server platforms.

In addition to querying information about the GDPS environment, the GDPS RESTful API allows programs that are written by clients, ISVs, and other IBM product areas to run actions against various elements of the GDPS environment. Examples of these actions include starting and stopping GM, updating the GM session parameters, and starting GDPS monitor processing. These capabilities enable clients, ISVs, and other IBM product areas to provide an even richer set of functions to complement the GDPS functions.

GDPS provides samples in the GDPS SAMPLIB library to demonstrate how the GDPS RESTful API can be used in programs.

5.3.4 Additional system management information

In a GDPS GM environment, the remote controlling system can use the hardware and system management actions to reconfigure the recovery site by adding temporary capacity, activating backup partitions, and starting production systems. This controlling system can be either for test purposes or for a real recovery. GDPS does not manage the systems or the hardware in the application site.

Most of the GDPS Standard Actions and several script commands require GDPS to communicate with the HMC. The interface GDPS uses to communicate with the HMC is called the BCPii. This interface allows GDPS to automate many of the HMC actions, such as LOAD, RESET, Activate or Deactivate an LPAR, and Activate or Undo CBU or On/Off Capacity on Demand (OOCoD).

The GDPS LOAD and RESET Standard Actions (available through the Standard Actions panel or the SYSPLEX script statement) allow specification of a CLEAR or NOCLEAR operand. These Standard Actions provide the operational flexibility to accommodate client procedures.

Extensive facilities for adding temporary processing capacity to the CPCs in the recovery site are provided by the GDPS scripting capability.

5.3.5 Securing the GDPS environment

GDPS uses RACF XFACILIT and GXFACILI resource classes to create a role-based security model for controlling access to the resources in your GDPS environment that is customized to your specific environment. Simple definitions can be used to control access at the panel option level or more granular definitions can be used to control access to specific types of resources, or even all the way down to the specific resource level.

With the role-based security model, you can create your own roles or use the common roles that GDPS recommends that include GDPS Administrator, GDPS Operator, GDPS User, and Non-GDPS User. You define the resources that these roles can access and the type of access they have to those resources by granting them access to the resource profiles that represent the various resources in your environment. Finally, you can grant access to various resources to users by adding them to the suitable roles.

When you use the role-based security model, GDPS ensures that the user has sufficient authority to take a specific action against a specific resource, regardless of whether they are attempting to act by using the panels directly or by running a GDPS script.

The GDPS Security Definition Utility tool is available to assist you with implementing your role-based security environment.

5.4 Enhanced resiliency with Region Switch and GM Bidirectional support

Companies relying on technology are constantly looking for improved resiliency tools and processes, for which GDPS provides great value. Regulation requirements have also increased over the years to enforce companies to adhere to resiliency standards and protect critical business data and business continuity. As part of these requirements, GDPS Global Mirror has been supporting the capability to perform Region Switch and Return Home for a few years.

The Region Switch process consists of implementing an equivalent infrastructure on both primary and recovery sites to enable customers to switch their production systems to the recovery site, which ensures that the DR capabilities remain available and that the systems can return home when necessary. Companies perform a Region Switch exercise for planned data center maintenance, DR readiness validation, or to prove to regulators that the production workload can successfully run from the recovery site.

5.4.1 GDPS GM Region Switch

The GDPS GM Region Switch and Return Home process consists of a normal GDPS GM configuration, that is, it has a GDPS K-sys running in Region A and a GDPS R-sys in Region B, with its normal GEOGROUP, GEOPARM, and scripts in place. From a disk replication perspective, the primary disk "A" in Region A replicates to the secondary disk "B" in Region B, which can take FlashCopy to the disk "C", which is also in Region B. The IBM Z processors in Region B must have LPAR definitions and infrastructure capabilities that are equivalent to Region A to start and run the systems that are moving from Region A to Region B. Then, you shut down the production systems in Region A, recover the disks at Region B, and start the production systems in Region B. To maintain a DR capability while running in Region B, use a FlashCopy disk "C" in Region A.

To orchestrate the GDPS Global Mirror in the reverse direction, preconfigure the GDPS R-sys in Region B to become the new K-sys, and the GDPS K-sys in Region A to become the new R-sys. GDPS requires a parallel inactive configuration with its own GEOGROUP, GEOPARM, and scripts to support the reconfiguration before GDPS GM can establish the replication in the reverse direction. Whenever GEOPARM is updated to include or remove replication pairs, PPRC links, disk subsystems, or other actions, the equivalent updates must be applied to the inactive GEOPARM configuration. Changes to GEOGROUP and scripts must be updated in both active and inactive configuration. This parallel configuration is activated after the Region Switch process completes, and the original GDPS configuration remains deactivated until the Return Home process takes place.

Figure 5-11 shows a sample environment at the starting point of this procedure.
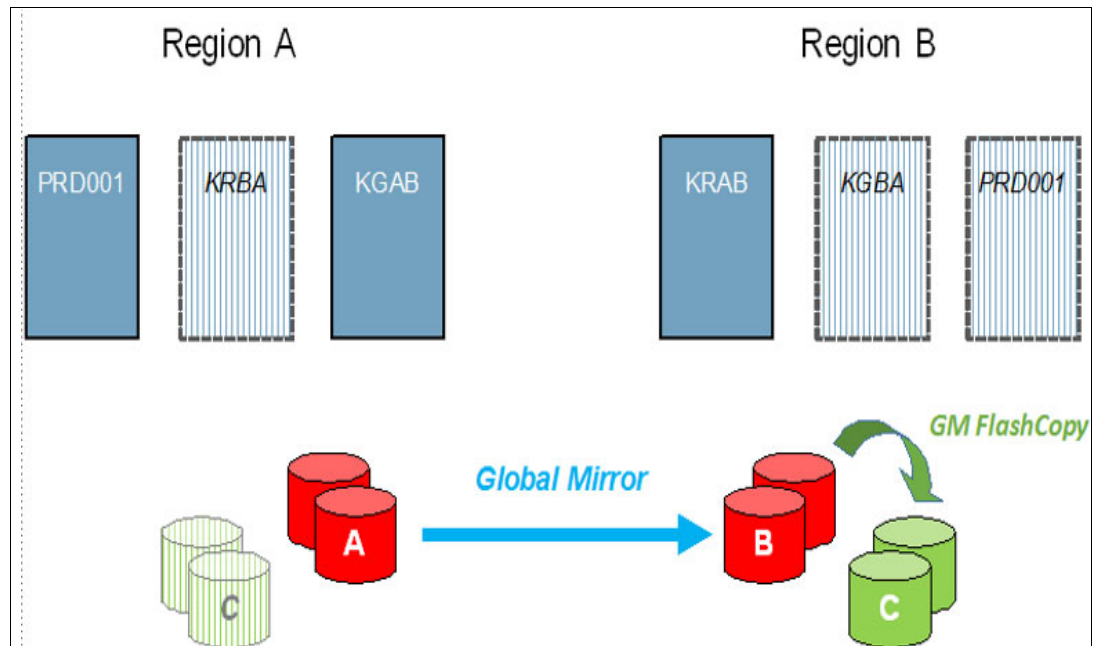
Figure 5-11   Sample environment at the start of the region switch procedure

Figure 5-12 shows the stage where the production systems switched to Region B but the Global Mirror replication is still not reversed.
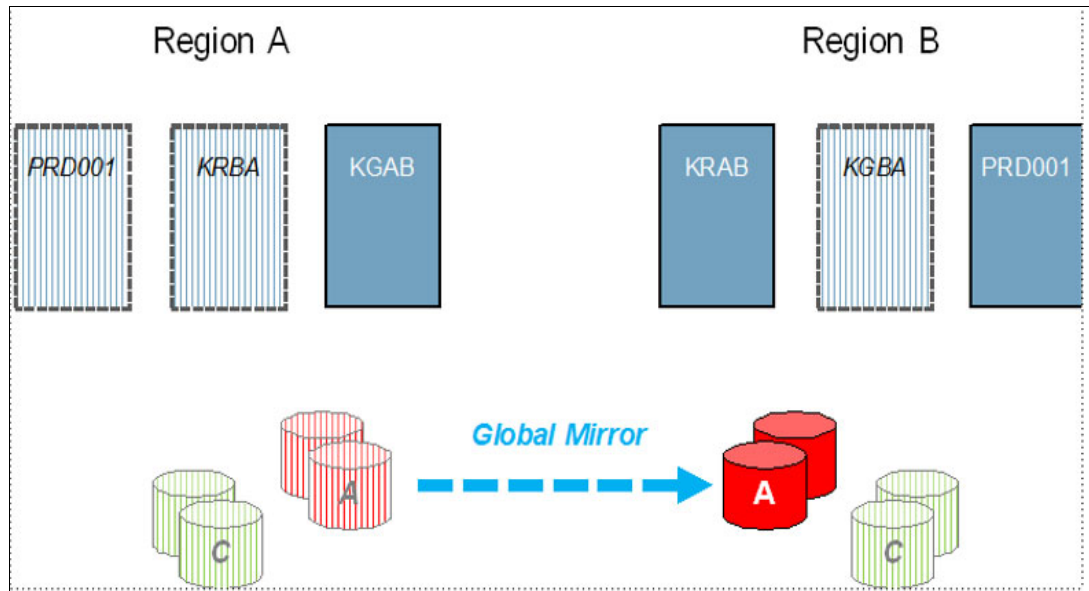

Figure 5-12   Sample environment during the region switch procedure

Figure 5-13 shows the final stage after the Region Switch completes and the Global Mirror replication is running from Region B as primary to Region A as secondary.
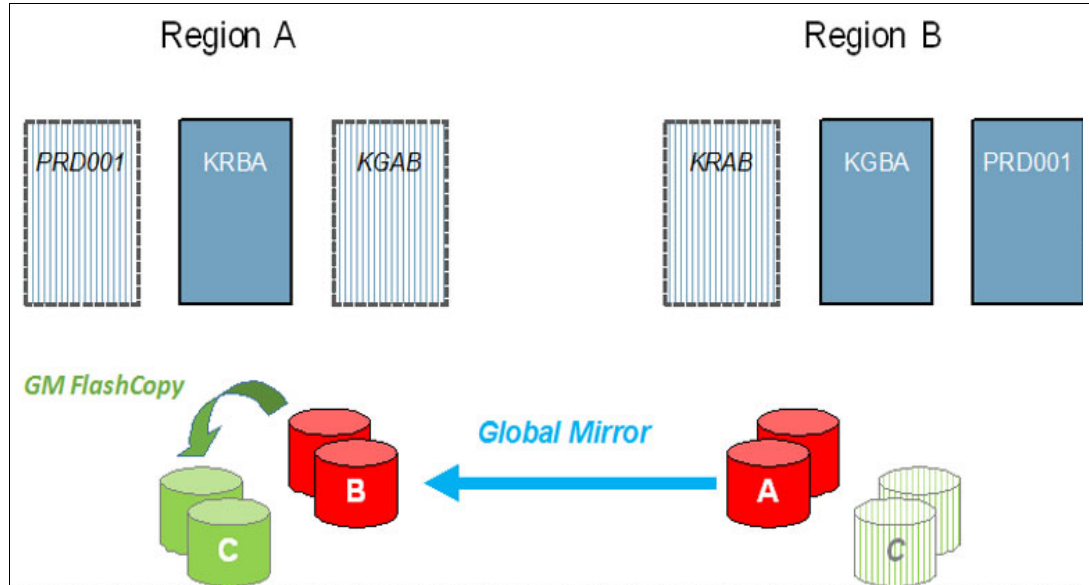

Figure 5-13   Sample environment at the end of the region switch procedure

5.4.2 GDPS GM Bidirectional support

The GDPS GM Bidirectional feature is an enhancement to the previously available Region Switch and Return Home capability. The purpose of this separately priced feature is to improve the efficiency of the Region Switch process by implementing the following features:

► A common GEOGROUP that includes both K-sys and R-sys pairs.
► A common GEOPARM that contains the DASD configuration for both mirroring directions.
► Multiple GM sessions that can run and switch direction independently.
► Region Switch that uses a GDPS procedure.
► Region switch that uses GDPS script and Remote Script Execution.
► All K-sys and R-sys controlling system instances are active always to manage the replication in both directions.

The GDPS GM Bidirectional feature reduces the complexity of maintaining multiple GEOPARMs and GEOGROUPs by using common members that are individually rolled out to both active GDPS GM environments independently, which avoids the situation of not deploying equivalent changes to both active and inactive sites. It also eliminates the risks of non-validated changes that are made to the inactive parameters that could be detected only during the Region Switch process, when such a configuration was made active.

Also, the Region Switch process can leverage the procedure through the Procedure handler to achieve a successful Region Switch, where the applicable regions switch roles, that is, the application region becomes the replication region and vice versa.

Figure 5-14 shows a GDPS GM Bidirectional environment. The environment consists of a server site, SS1, in each region that contains the K-sys and R-sys controlling systems for both directions. Each K-sys and R-sys pair handles the replication for one direction. In this example, KGAB and KRAB handle GM sessions from Region A to Region B, and KGBA and KRBA handle GM sessions from Region B to Region A.
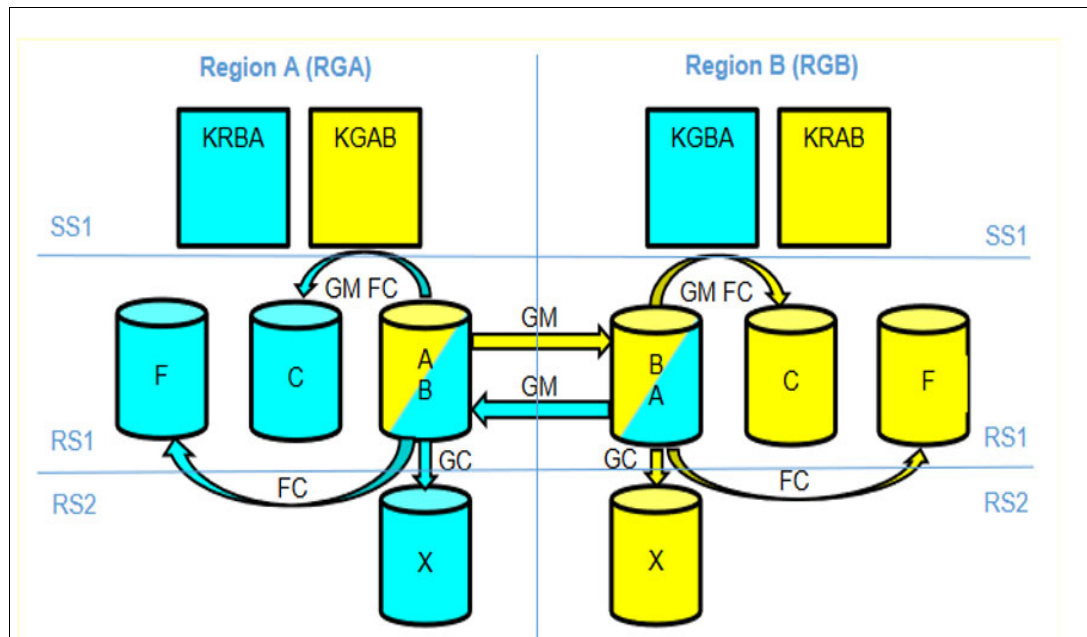

Figure 5-14   Example of a bidirectional environment

A region switch can be done individually per GM session. Any dependency that requires multiple GM sessions to have the same direction must be coordinated by using GDPS scripts or administrative procedures outside of GDPS.

The GDPS Tools utility GDPS XML Conversion Tool for GM2Site can help you generate an XML GEOPARM for a bidirectional environment when converting an existing GM 2-site or GM 2-site region switch-capable environment into a GM 2-site Bidirectional environment.

5.5 GDPS GM monitoring and alerting

The GDPS SDF panel is described in 5.3.1, "User interfaces" on page 155. It is this panel on which GDPS dynamically displays alerts (which are color-coded based on severity) when a non-normal status or situation is detected.

Alerts can be posted as a result of an unsolicited error situation for which GDPS listens. For example, if a problem occurs with the GM session and the session suspends outside of GDPS control, GDPS is aware of this issue because the disk subsystem that is the Master for the GM session posts an SNMP alert. GDPS listens for these SNMP alerts, and in turn posts an alert on the SDF panel that notifies the operator of the suspension event.

Alerts can also be posted as a result of GDPS periodically monitoring key resources and indicators that relate to the GDPS GM environment. If any of these monitoring items are found to be in a state that is deemed to be not normal by GDPS, an alert is posted on SDF.

Because the K-sys and R-sys have different roles and affect different resources, they each monitor a different set of indicators and resources.

For example, the K-sys has TCP/IP connectivity to the A disk through which the GM Master disk subsystem posts SNMP alerts about GM problems. For this reason, it is important that the TCP/IP connectivity between the K-sys and the production disk is functioning properly. The K-sys, among other things, monitors this connection to ensure that it is functional so that if there is a GM problem, the SNMP alert reaches the K-sys.

Likewise, it is the R-sys that uses the BCPii to perform hardware actions to reconfigure the recovery site, for disaster testing or in a real recovery scenario. One of the resources that is monitored by the R-sys is the BCPii connection to all CPCs in the recovery site on which the R-sys can perform hardware operations, such as CBU or LPAR activation.

In addition to posting alerts on their own SDF panel, the K-sys and R-sys forward any alerts to the other system for posting. Because the operator is notified of R-sys alerts on the K-sys SDF panel, it is sufficient for the operator to monitor the K-sys SDF panel during normal operations if the K-sys is up and running.

If an alert is posted, the operator must investigate (or escalate) and corrective action that must be taken for the reported problem as soon as possible. After the problem is corrected, it is detected during the next monitoring cycle and the alert is cleared by GDPS automatically.

GDPS GM monitoring and alerting capability is intended to ensure that operations are notified and can take corrective action for any problems in their environment that can affect the ability of GDPS GM to do recovery operations. This correction maximizes the installation's chance of achieving RPO and RTO commitments.

5.5.1 GDPS GM health checks

In addition to GDPS GM monitoring, GDPS provides health checks. These health checks are provided as a plug-in to the z/OS Health Checker infrastructure to check that certain settings related to GDPS adhere to GDPS best practices.

The z/OS Health Checker infrastructure is intended to check various settings to determine whether these settings adhere to z/OS best practices values. For settings that are found to be not in line with best practices, exceptions are raised in the Spool Display and Search Facility (SDSF).

Many products, including GDPS, provide health checks as a plug-in to the z/OS Health Checker. There are various parameter settings that are related to GDPS, such as z/OS parmlib settings or NetView settings, and the recommendations and best practices for these settings are documented in GDPS publications. If these settings do not adhere to recommendations, this issue can hamper the ability of GDPS to perform critical functions in a timely manner.

Although GDPS monitoring detects that GDPS cannot perform a particular task and raises an alert, the monitor alert might be too late, at least for that instance of an incident. Often, if there are changes in the client environment, this issue might necessitate adjustment of some parameter settings that are associated with z/OS, GDPS, and other products. It is possible that you can miss making these adjustments, which might result in affecting GDPS.

The GDPS health checks are intended to detect such situations and avoid such incidents where GDPS is unable to perform its job because of a setting that is less than ideal.

For example, there are several address spaces that are associated with GDPS GM and best practices are documented for these address spaces. The GDPS code runs in the NetView address space and there are DFSMS address spaces that GDPS interfaces with to perform GM copy services operations.

GDPS recommends that these address spaces are assigned to specific Workload Manager (WLM) service classes to ensure that they are dispatched in a timely manner and do not lock each other out. For example, one of the GDPS GM health checks determines whether these address spaces are set up and running with the characteristics that are recommended by GDPS.

Similar to z/OS and other products that provide health checks, GDPS health checks are optional. Several best practices values that are checked and the frequency of the checks can be customized to cater to unique client environments and requirements.

GDPS also provides a useful interface for managing the health checks by using the GDPS panels. You can perform actions, such as activate or deactivate or run any selected health check, and view the customer overrides in effect for any best practices values.

Figure 5-15 shows a sample of the GDPS Health Check management panel. In this example, you see that all the health checks are enabled. The status of the last run is also shown, which indicates whether the last run was successful or resulted in an exception. Any exceptions can also be viewed by using other options on the panel.

```
VPC8PHC0         GDPS Health Checks Information              G9C1           (1/4)


HealthChecker:  STARTED   Procname: HZSPROC    Policy Name: DEFAULT

Actions:    S elect  R un  A ctivate  D eactivate  I nfo

Cmd Check name                         State              Status

 _    GDPS_CHECK_DASDMIH               ACTIVE(ENABLED)    EXCEPTION-MEDIUM
 _    GDPS_CHECK_NUMUCBS               ACTIVE(ENABLED)    SUCCESSFUL
 _    GDPS_CHECK_JOBS                  ACTIVE(ENABLED)    SUCCESSFUL
 _    GDPS_CHECK_K_SYS_LPAR            ACTIVE(DISABLED)   ENV N/A









Commands:  R Run  A Activate  D Deactivate  S Summary  G GEOHCP00  P Parameter
Selection ===>
F1=Help  F3=Return  F5=Refresh  F6=Roll                 F10=Left  F11=Right
```

Figure 5-15   GDPS GM Health Check management panel

5.6 Other facilities that are related to GDPS

In this section, we describe other facilities that are provided by GDPS Global - GM that can help various ways.

5.6.1 GDPS GM Copy Once facility

GDPS provides a *Copy Once* facility to copy volumes that have data sets on them that are required for recovery but the content is not critical, so they do not need to be copied all the time. Page data sets and work volumes that contain only truly temporary data such as sort work volumes are primary examples. The Copy Once facility can be started whenever required to refresh the information about these volumes.

To restart your workload in the recovery site, you need to have these devices or data sets available (the content is not required to be up to date). If you do not remote copy all of your production volumes, you must either manually ensure that the required volumes and data sets are preallocated and kept up to date at the recovery site or use the GDPS Copy Once function to manage these devices.

For example, if you are not replicating your paging volumes, then you must create the volumes with the proper volume serial with required data sets in the recovery site. Then, each time you change your paging configuration in the application site, you must reflect the changes in your recovery site.

The GDPS Copy Once function provides a method of creating an initial copy of such volumes plus the ability to re-create the copy if the need arises as the result of any changes in the application site.

If you plan to use the Copy Once facility, you need to ensure that no data that needs to be continuously replicated is placed on the volumes you define to GDPS as Copy Once because these volumes will not be continuously replicated. The purpose of Copy Once is to ensure that you have a volume with the correct VOLSER, and with the data sets that are required for recovery are allocated, available in the recovery site. The data in the data sets is not time-consistent with the data on the volumes that are continuously mirrored.

5.6.2 Global Mirror Monitor integration

GDPS provides a Global Mirror Monitor (also referred to as *GM Monitor*) that is fully integrated into GDPS. This function provides a monitoring and historical reporting capability for GM performance and behavior, and some autonomic capability based on performance. The GM Monitor provides the following capabilities:

► Ability to view recent performance data for a GM session, for example to understand if an ongoing incident might be related to GM.

► Generation of alerts and messages for GM behavior based on exceeding thresholds in a defined policy.

► Ability to perform automatic actions such as pausing a GM session or resuming a previously paused session based on a defined policy.

► Creation of SMF records with detailed historical GM performance and behavioral data for problem diagnosis, performance reporting, and capacity planning.

The GM Monitor function runs in the K-sys and supports both CKD and FB environments. An independent monitor can be started for each GM session in your GDPS configuration. GDPS stores the performance data that is collected by each active monitor. Recent data is viewable by using the GDPS 3270 panels.

5.6.3 Easy Tier Heat Map Transfer

IBM DS8000 Easy Tier optimizes data placement of logical volumes across the various physical tiers of storage within a disk subsystem to optimize application performance. The placement decisions are based on learning the data access patterns, and can be changed dynamically and transparently by using this data.

GM copies the data from the primary to the secondary disk subsystem. However, the Easy Tier learning information is not included in the GM scope. The secondary disk subsystems are optimized according to the workload on these subsystems, which is different than the activity on the primary (there is only a write workload on the secondary whereas there is read/write activity on the primary).

Also, there is little activity on the tertiary disk (FlashCopy target disk, or FC1 disk), so it is optimized differently than the primary disk or the secondary disk. As a result of these differences, during a recovery, the disks that you recover on (secondary or tertiary) are likely to display different performance characteristics compared to the former primary.

Easy Tier Heat Map Transfer is the DS8000 capability to transfer the Easy Tier learning from a GM primary disk to a target set of disks. With GDPS GM, the Easy Tier learning can be transferred to the secondary disk and the tertiary disk (FC1 disk) so that whatever disk you recover on can also be optimized based on this learning, and has similar performance characteristics as the former primary.

GDPS integrates support for Heat Map Transfer. The Heat Map Transfer actions (such as start/stop of the processing and reversing transfer direction) are incorporated into the GDPS managed processes. For example, if GM is temporarily suspended for a planned or unplanned secondary disk outage, Heat Map Transfer is also suspended.

5.7 Dynamic Site Table refresh

GDPS Global Mirror manages different systems that are part of the GDPS configuration through the GEOGROUP, which builds the Site Table. With GDPS GM 4.7, it is possible to dynamically add, remove, or change system definitions in the GEOGROUP by using the standard Config Management option in GDPS main panel. This approach reduces the complexity of managing the Site Table and keeps the configuration process standard (that is, under the same menu that is used to refresh Options, Scripts, GEOGROUP, Disk, and others).

5.8 Dynamic PPRC Link Configuration Management

GDPS Global Mirror 4.7 introduces the ability to dynamically manage PPRC links through GDPS panels. Using this new feature, you can add, remove, change, and query PPRC links that are associated to a particular consistency group or replication leg.

Use the K-sys to manage Global Mirror replication legs, and the R-sys to manage X-Disk legs, if they are available.

5.9 Flexible testing and Logical Corruption Protection

If you want to conduct a DR test, you can use GDPS GM to prepare the B disks to be used for the test. However, during the test, remote copying must be suspended because the B disks are being used for the test, and the C disks contain a consistent copy of the production disks at the start of the test. If you have a real disaster during the test, the C disks are used to give you a consistent restart point. However, all updates that are made to the production disks after the start of the test must be re-created. At the completion of the test, GDPS GM uses the Failover/Failback capability to resynchronize the A and B disks without having to do a complete copy.

GDPS GM supports an extra FlashCopy disk device, referred to as *F disks* or *FC1 disks*. F disks are extra "practice" FlashCopy target devices that might optionally be created in the recovery site. These devices might be used to facilitate stand-alone testing of your DR procedures. Disaster testing can be conducted by starting recovery systems on the F disk while live production continues to run in the application site and remains protected by the B and C disks. In addition, the F disk can be used to create a "gold" or insurance copy of the data in a disaster situation. If you have this additional practice FlashCopy, you can schedule disaster tests on demand much more frequently because such tests will have little or no impact on your RPO and DR capability.

For added scalability, GDPS allows the GM FlashCopy disks (C) to be defined in alternative subchannel set MSS1 or to not be defined to the R-sys at all. (For more information, see "Addressing z/OS device limits in a GDPS Metro environment" on page 25.) GDPS GM also supports the use of FC1 disk without having the FC1 disk defined to the R-sys.

By combining GM with FlashCopy, you can create a usable copy of your production data to provide for on-demand testing capabilities and other nondisruptive activities. If there is a requirement to perform DR testing while maintaining the currency of the production mirror or for taking regular extra copies, once or twice a day, for other purposes, then consider installing the additional disk capacity to support F disks in your GM environment.

5.9.1 Using space-efficient FlashCopy

As discussed in "Space-efficient FlashCopy" on page 34, by using space-efficient FlashCopy volumes, you might be able to lower the amount of physical storage needed, and thus reduce the cost that is associated with providing a tertiary copy of the data. GDPS has support to allow FlashCopy SE volumes to be used as FlashCopy target disk volumes.

This support is transparent to GDPS; if the FlashCopy target devices defined to GDPS are space-efficient volumes, GDPS uses them. All GDPS FlashCopy operations with the NOCOPY option, whether through GDPS scripts or panels, can use space-efficient targets.

Because the IBM FlashCopy SE repository is of fixed size, it is possible for this space to be exhausted, thus preventing further FlashCopy activity. Therefore, as a best practice, use space-efficient volumes for temporary purposes, so that space can be reclaimed regularly.

GDPS GM might use space-efficient volumes as FlashCopy targets for either the C-disk or the F-disk. In the GM context, where the C-disk is allocated to space-efficient volumes, each new Consistency Group reclaims used repository space since the previous Consistency Group, as the new flash is established with the C-disk. Therefore, a short Consistency Group Interval in effect ensures the temporary purpose recommendation for FlashCopy data. However, if the Consistency Group Interval grows long because of constrained bandwidth or write bursts, it is possible to exhaust available repository space. This action causes a suspension of GM because any subsequent FlashCopy will not be possible.

Using space-efficient volumes for F disks depends on how you intend to use the F disks. These disks can be used for short-term, less-expensive testing, but are suitable for actual recovery because of their non-temporary nature.

5.9.2 Creating a test copy by using GM CGPause and testing on isolated disks

The most basic GM configuration requires the GM secondary disk and the GM FlashCopy on the secondary disk subsystems. If you use an extra set of practice FlashCopy disks on the same disk subsystems, while you are performing recovery testing, you have the I/O activity for GM mirroring and also the I/O activity generated by recovery testing on the same set of secondary disk subsystems. This I/O activity from the testing can potentially affect the GM mirroring.

GDPS GM supports creating a *test copy* on disk subsystems that are isolated from the secondary disk subsystems. We call these subsystems the X-disks. The GM secondary disks are connected to the X-disks by using the Global Copy (PPRC-XD) asynchronous copy technology. The GM secondary disks are the primary disks for the relationship to the X-disks.

To create a consistent test copy on the X-disks, GDPS GM uses the Consistency Group Pause (CGPause) capability of the DS8000 disk subsystem to make the GM secondary disks consistent. After the GM secondary disks are consistent, GDPS waits until all data on these disks is replicated to the X-disks and isolates the X-disks. GDPS then resumes the GM session.

The entire process of isolating the test copy on X-disks takes place in a short amount of time, which means minimal impact to GM operations occurred during the creation of the test copy. Now, with the test copy isolated on disk subsystems other than the secondary disk subsystems, any testing that is performed does not interfere with or affect GM replication, which continues while you test on the X-disk copy.

GDPS also supports the same technique by using CGPause to create practice FlashCopy. For environments that do not support CGPause, the GM secondary disks must first be recovered to make them consistent to take the practice FlashCopy. This GM session disruption is longer compared to creating the FlashCopy test copy by using CGPause.

In summary, CGPause minimizes the interruption to the GM session when creating a test copy. Isolating the test copy on a separate set of disk subsystems (X-disk) eliminates any impact that the testing operation might have on the resumed GM session.

5.9.3 Logical Corruption Protection

In addition to the use of FlashCopy technology to provide flexible testing capabilities, GDPS GM uses another technology that is called Safeguarded Copy (SGC) to provide a powerful solution for protecting against various types of logical data corruption, including cyberattacks and internal threats. This capability is referred to as *Logical Corruption Protection* (LCP). For more information about LCP, see Chapter 9, "IBM GDPS Logical Corruption Protection and Testcopy Manager" on page 255.

5.10 GDPS tools for GDPS GM

GDPS includes tools that provide functions that are complementary to GDPS function. The tools represent the functions that many clients are likely to develop themselves to complement GDPS. Using the GDPS tools eliminates the necessity for you to develop a similar function yourself. The tools are provided in source code format, which means that if the tool does not completely meet your requirements, you can modify the code to tailor it to your needs.

The GDPS Distributed Systems Hardware Management Toolkit is available for GDPS GM. It provides an interface for GDPS to monitor and control distributed systems' hardware and virtual machines (VMs) by using script procedures that can be integrated into GDPS scripts. This tool provides REXX script templates that show examples of how to monitor/control: IBM AIX® HMC, VMware ESX server, IBM BladeCenter, and stand-alone x86 servers with Remote Supervisor Adapter II (RSA) cards.

The GDPS XML Conversion (GeoXML) tool and GDPS XML Conversion Tool for GM2Site help you to convert GEOPARM from older versions to an XML-based version (starting at GDPS 4.4) and converting a GM 2-site or GM 2-site region switch-capable environment into a GM 2-site Bidirectional environment.

The GDPS Security Definition Utility tool helps you to implement enhancements to GDPS options that use role-based security by helping you with the definition of appropriate profiles in the XFACILIT class of RACF and by assigning relevant access to them.

5.11 Services component

As demonstrated, GDPS touches on more than remote copy. It also includes automation, disk and system recovery, testing processes, and DR processes.

Most installations do not have all these skills readily available. Also, it is rare to find a team that possesses this range of skills across many implementations. However, the GDPS GM offering provides access to a global team of specialists in all the disciplines you need to ensure a successful GDPS GM implementation.

Specifically, the Services component includes some or all of the following services:

► Planning to determine availability requirements, configuration recommendations, implementation and testing plans. Planning session topics include hardware and software requirements and prerequisites, configuration and implementation considerations, cross-site connectivity planning and potentially bandwidth sizing, and operation and control.

► Assistance in defining RPOs and recovery time objectives (RTOs).

► Installation and necessary customization of NetView and System Automation.

► Remote copy implementation.

► GDPS GM automation code installation and policy customization.

► Education and training on GDPS GM setup and operations.

► Onsite implementation assistance.

► Project management and support throughout the engagement.

The sizing of the Services component of each project is tailored for that project based on many factors, including what automation is already in place, whether remote copy is already in place, and other factors. The skills that are provided are tailored to the specific needs of each implementation.

5.12 GDPS GM prerequisites

For more information about GDPS GM prerequisites, see this GDPS web page.

5.13 Comparing GDPS GM versus other GDPS offerings

So many features and functions are available in the various members of the GDPS family that recalling them all and remembering which offerings support them is sometimes difficult. Table 5-1 lists the key features and functions and indicates which are delivered by the various GDPS offerings.

Table 5-1   Supported features matrix

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
CA	Yes	Yes	Yes	No
DR	Yes	Yes	Yes	Yes
CA/DR protection against multiple failures	Yes	No	No	No

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
CA for foreign z/OS systems	Yes with z/OS proxy	No	No	No
Supported distance	200 km 300 km (Business Recovery Services (BRS) configuration)	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	Virtually unlimited
Consistent FlashCopy support	Yes, using CONSISTENT	Yes, using CONSISTENT for secondary only	No	Yes, using CGPause
Reduced impact on initial copy/resync	Yes	Yes	Yes	Not applicable
Tape replication support	Yes	No	No	No
Production sysplex automation	Yes	No	Not applicable	No
Span of control	Both sites	Both sites (disk only)	Both sites	Disk at both sites; recovery site (CBU or LPARs)
GDPS scripting	Yes	No	Yes	Yes
Monitoring, alerting and health checks	Yes	Yes	Yes (except health checks)	Yes
Query Services	Yes	Yes	No	Yes
MSS support for added scalability	Yes (RS2 in MSS1, RS3 in MSS2)	Yes (secondary in MSS1)	No	Yes (GM FlashCopy and Primary for Metro Global Mirror (MGM) in MSS1)
MGM 3-site and 4-site	Yes (all configurations)	Yes (3-site only and non-Incremental Resynchronizati on (IR) only)	No	Yes (all configurations)
FB disk	Yes	Yes	No	Yes
z/OS equivalent function for Linux on IBM Z	Yes (Linux on IBM Z running as a z/VM guest only)	No	Yes (Linux on IBM Z running as a z/VM guest only)	Yes
GDPS GUI	Yes	Yes	Yes	Yes

5.14 Summary

GDPS GM provides automated DR capability over virtually unlimited distances for both CKD and FB devices.

The following controlling systems in a GDPS GM configuration provide different functions:

► The K-sys, in the application site, is used to set up and control all remote copy operations.
► The R-sys, in the recovery site, is used primarily to drive recovery in a disaster.

You define a set of scripts that can reconfigure the servers in the recovery site, recover the disks, and start the production systems. The powerful scripting capability can perfect the actions to be taken, either for planned or unplanned changes, thus eliminating the risk of human error. Both the K-sys and R-sys monitor key indicators and resources in their span of control and alert the operator of any non-normal status so that corrective action can be taken in a timely manner to eliminate or minimize RPO and RTO impact.

The B disks in the recovery site can be used for DR testing. The C disks contain a consistent (although aging) copy of the production volumes. Optionally, a practice FlashCopy (F disks) can be integrated to eliminate the risk of RPO impact associated with testing on the B disks.

In addition to its DR capabilities, GDPS GM also provides a simple interface for monitoring and managing the remote copy configuration.

IBM GDPS Continuous Availability solution

In this chapter, we introduce the GDPS Continuous Availability solution. This solution significantly reduces both the time that is required to recover systems in a disaster recovery (DR) situation, and the time that a workload is unavailable due to a planned outage.

The chapter includes the following topics:

6.1 Overview of GDPS Continuous Availability

In this section, we provide a high-level description of the GDPS Continuous Availability solution and explain where it fits in with the other IBM Geographically Dispersed Parallel Sysplex (GDPS) products.

6.1.1 Positioning GDPS Continuous Availability

Business continuity features the following key metrics:

► Recovery time objective (RTO): How long can you afford to be without your systems?
► Recovery point objective (RPO): How much data can you afford to lose or re-create?
► Network recovery objective (NRO): How long does it take to switch over the network?

Multiple offerings are available in the GDPS family, all of which are covered in this book. The GDPS products other than GDPS Continuous Availability are continuous availability (CA) and DR solutions that are based on synchronous or asynchronous disk hardware replication.

To achieve the highest levels of availability and minimize the recovery for planned and unplanned outages, various clients deployed GDPS Metro multi-site workload configurations, which have the following requirements:

► All critical data must be Metro Mirrored and HyperSwap enabled.
► All critical coupling facility (CF) structures must be duplexed.
► Applications must be Parallel Sysplex enabled.

However, the signal latency between sites can affect online workload throughput and batch duration. This issue results in sites typically being separated by no more than approximately 20 km (12.4 miles) fiber distance[1].

Therefore, the GDPS Metro multi-site workload configuration, which can provide an RPO of zero and an RTO as low as a few minutes, does not provide a solution if an enterprise requires that the distance between the active sites is greater than 20 - 30 km (12.4 - 18.6 miles).

GDPS GM, which is based on asynchronous hardware replication, provides for virtually unlimited site separation. However, it requires that the workload from the failed site is restarted in the recovery site and this process typically takes 30 - 60 minutes in a well-tuned environment. Therefore, GDPS GM cannot achieve the RTO of seconds that is required by various enterprises for their most critical workloads.

When the GDPS products based on hardware replication are used, it is not possible to achieve aggressive RPO and RTO goals while providing the sufficient site separation that is required by some enterprises.

For these reasons, the GDPS Continuous Availability sites concept was conceived.

---

[1] The distance between sites in a GDPS Metro multi-site workload configuration that any client can tolerate depends on the client's application workloads and service level requirements. Each client must test with its own applications and workloads to determine the distance it can achieve. Nearly all clients running GDPS Metro multi-site workload configurations are running their two sites at a 20 km (12.4 miles) distance or less. However, this restriction does not necessarily mean that larger distances are impossible.

6.1.2 GDPS Continuous Availability sites concept

The GDPS Continuous Availability sites concept consists of having two sites that are separated by virtually unlimited distances, running the same applications, and having the same data to provide cross-site workload balancing and CA and DR. This change is a fundamental paradigm shift from a *failover* model to a *CA* model.

GDPS Continuous Availability (GDPS CA) does not use any of the infrastructure-based data replication techniques that other GDPS products rely on, such as Metro Mirror (PPRC) or Global Mirror (GM)[2]. Instead, GDPS Continuous Availability relies on both of the following methods:

► Software-based asynchronous replication techniques for copying the data between sites.

► Automation, primarily operating at a workload level, to manage the availability of selected workloads and the routing of transactions for these workloads.

The GDPS Continuous Availability product, which is a component of the GDPS Continuous Availability solution, acts primarily as the coordination point or controller for these activities. It is a focal point for operating and monitoring the solution and readiness for recovery.

> **Note:** For simplicity, in this chapter we refer to both the solution and the product as GDPS Continuous Availability. We might also refer to the environment managed by the solution, and the solution itself, as GDPS CA.

What is a workload

A *workload* is defined as the aggregation of the following components:

► Software

  User-written applications, such as COBOL programs, and the middleware runtime environment (for example, CICS regions, InfoSphere Replication Server instances and IBM Db2 subsystems).

► Data

  A related set of objects that must preserve transactional consistency and optionally referential integrity constraints (for example, Db2 tables or IMS databases). Network connectivity

  One or more TCP/IP addresses and ports (for example, 10.10.10.1:80), IBM MQ Queue managers, or SNA applids.

> **Note:** Workloads that include resources from multiple DBMSs (Db2 and IMS, or Db2 and VSAM) are supported. However, transactional consistency across multiple DBMSs is *not* provided.

---

[2] Where the GDPS CA Zero Data Loss (ZDL) feature is used, GDPS CA relies on GDPS Metro or GM.

The following workload types are supported and managed in a GDPS Continuous Availability environment:

► Update or read/write workloads

These workloads run in what is known as the *Active/Standby configuration*. In this case, a workload that is managed by GDPS Continuous Availability is active in one sysplex and receives transactions that are routed to it by the workload distribution mechanism that is managed by the IBM Multi-site Workload Lifeline.

The workload also uses software replication to copy changed data to another instance of the workload that is running in a second sysplex where all the infrastructure components (LPARs, systems, middleware, and others) and even the application are ready to receive work in what is termed a standby mode.

The updated data from the active instance of the workload is applied in real time to the database subsystem instance that is running in standby mode.

► Query or read-only workloads

These workloads are associated with update workloads, but they can be actively running in both sites concurrently. Workload distribution between the sites is based on policy options, and accounts for environmental factors, such as the latency for replication that determines the age (or currency) of the data in the standby site.

No data replication is associated with the query workload because no updates are made to the data. You can associate up to two query workloads with a single update workload.

► Crypto workload

These workloads provide the optional capability to keep cryptographic key material synchronized between sites in a GDPS Continuous Availability environment. Only one crypto workload is ever available, which is responsible for monitoring the replication of changes to the VSAM data sets that are used by ICSF in the two sites. It also alerts the operator to any problems with this issue.

Figure 6-1 on page 183 shows these concepts for an update workload at a high level. Transactions arrive at the workload distributor, which is also known as the *load balancer*. Depending on the current situation, the transactions are routed to what is termed the currently active sysplex in the configuration for that particular workload.
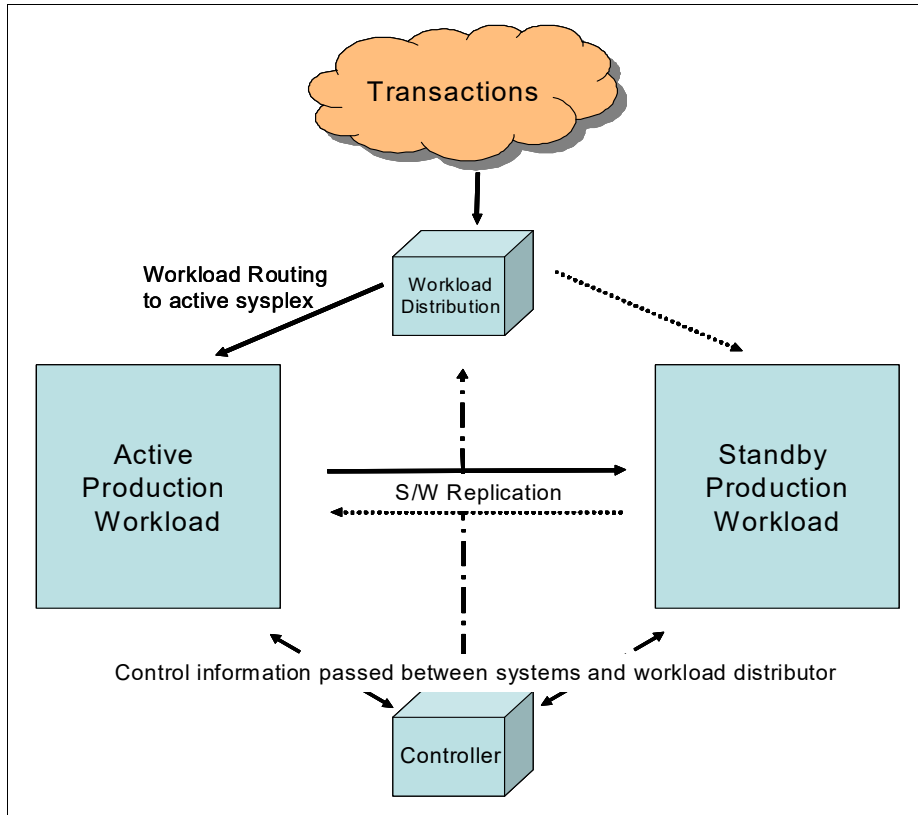
Figure 6-1   GDPS Continuous Availability concept

The environment is constantly being monitored to ensure that workload is being processed in the active sysplex. If GDPS Continuous Availability detects that a workload is not processing normally, a policy-based decision is made to automatically start routing work to the standby sysplex (rather than the currently active sysplex), or to prompt the operator to act. In a similar way, for query workloads, a policy that uses the latency of replication as thresholds trigger GDPS Continuous Availability or other products in the solution to take some action.

Information is constantly being exchanged by the systems in the active and standby sysplexes, the GDPS controllers (one in each location), and the workload distribution mechanism to ensure that an accurate picture of the health of the environment is maintained to enable appropriate decisions from the automation.

In a planned manner, it is also possible to switch each workload from the currently active to the standby sysplex if the need arises, such as for routine maintenance.

In your environment, you are likely to have some applications and data that you do not want to manage with, or that cannot be managed by, GDPS Continuous Availability. For example, you might have an application that uses a data type for which software data replication is not available or is not supported by GDPS Continuous Availability.

You still need to provide high availability (HA) and DR (HADR) for such applications and data. For this task, GDPS Continuous Availability provides for integration and co-operation with other GDPS products that rely on hardware replication and are independent of application and data type.

Specifically, special coordination is provided with GDPS Metro, which is described in 6.5, "GDPS Continuous Availability co-operation with GDPS Metro" on page 205.

6.2 GDPS Continuous Availability solution products

The GDPS Continuous Availability architecture, which is shown at a conceptual level in Figure 6-2, consists of several products coordinating the monitoring and managing of the various aspects of the environment.
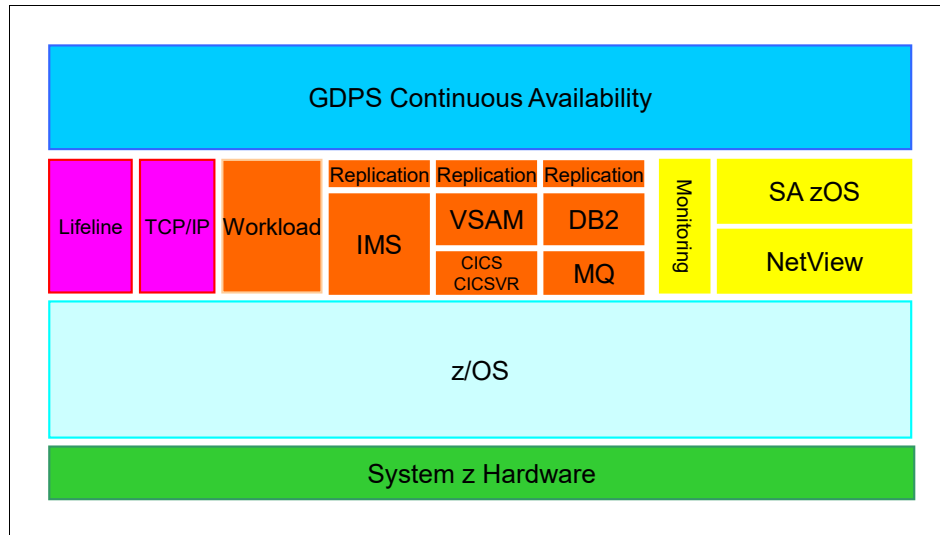
Figure 6-2   GDPS Continuous Availability architecture

This section describes the various products that are required for GDPS Continuous Availability and their role or function within the overall framework. The following products are briefly discussed:

► GDPS Continuous Availability
► BM Z NetView
► IBM Z NetView for z/OS Enterprise Management Agent (NetView agent)
► BM Z Service Management Suite
► IBM Z NetView for Continuous Availability
► IBM Z System Automation for z/OS
► IBM Multi-site Workload Lifeline for z/OS
► Middleware such as CICS, IMS, Db2, MQ to run the workloads
► Replication Software:
  – IBM InfoSphere Data Replication for Db2 for z/OS

    BM MQ for z/OS v9.1 is required for Db2 data replication

  – IBM InfoSphere Data Replication for VSAM for z/OS

    CICS Transaction Server for z/OS, CICS VSAM Recovery for z/OS, or both, are required for VSAM replication

  – InfoSphere IMS Replication for z/OS
► Other optional components

  IBM OMEGAMON® monitoring products for monitoring the various parts of the solution

For more information about a solution view that shows how the products are used in the various systems in which they run, see 6.3, "GDPS Continuous Availability environment" on page 189.

6.2.1 GDPS Continuous Availability product

The GDPS Continuous Availability product provides automation code that is an extension of many of the techniques that were tested in other GDPS products. They also were tested with many client environments around the world for management of their mainframe CA and DR requirements.

The following key functions are provided by the GDPS Continuous Availability code:

► Workload management, such as starting or stopping all components of a workload in a specific sysplex.

► Replication management, such as starting or stopping replication for a specific workload from one sysplex to the other.

► Routing management, such as stopping or starting routing of transactions to one sysplex or the other for a specific workload.

► System and Server management, such as STOP (graceful shutdown) of a system; LOAD, RESET, ACTIVATE, DEACTIVATE the LPAR for a system; and capacity on-demand actions, such as Capacity BackUp (CBU) or On/Off Capacity on Demand (OOCoD) activation.

► Monitoring the environment and alerting for unexpected situations.

► Planned/Unplanned situation management and control, such as planned or unplanned site or workload switches.

► Autonomic actions, such as automatic workload switch (policy-dependent).

► Powerful scripting capability for complex/compound scenario automation.

► Co-operation with GDPS Metro to provide continuous data availability in the GDPS Continuous Availability sysplexes.

► Monitoring of replication across GDPS CA-managed sysplexes of cryptographic data, which is stored by ICSF in VSAM data sets. If a switch occurs from one sysplex to another, all workloads that use ICSF cryptographic services can continue to run.

► Graphical user interface (GUI).

6.2.2 IBM Z NetView

The IBM Z NetView product is a prerequisite for GDPS Continuous Availability automation and management code. In addition to being the operating environment for GDPS, the NetView product provides more monitoring and automation functions that are associated with the GDPS Continuous Availability solution.

Monitoring capability by using the NetView agent is provided for the following items:

► IBM Multi-site Workload Lifeline for z/OS
► IBM InfoSphere Data Replication for Db2 for z/OS
► IBM InfoSphere Data Replication for VSAM for z/OS
► IBM InfoSphere Data Replication for IMS for z/OS

NetView Agent

The IBM Z NetView Enterprise Management Agent (also known as TEMA) is used in the solution to pass information from the z/OS NetView environment to the Tivoli Enterprise Portal, which is used to provide a view of your enterprise. From this portal, you can drill down to more closely examine components of each system that is monitored. The NetView agent requires IBM Z NetView Monitoring for Continuous Availability.

6.2.3 IBM Z NetView Monitoring for Continuous Availability

IBM Z NetView Monitoring for Continuous Availability is a suite of monitoring components to monitor and report on various aspects of a client's IT environment. Several of the IBM Z NetView Monitoring components are used in the overall monitoring of aspects (such as monitoring the workload) within the GDPS Continuous Availability environment.

The specific components that are required for GDPS Continuous Availability are listed here.

Tivoli Enterprise Portal

Tivoli Enterprise Portal (portal client or portal) is a Java-based interface for viewing and monitoring your enterprise. Tivoli Enterprise Portal offers two modes of operation: desktop and browser.

Tivoli Enterprise Portal Server

Tivoli Enterprise Portal Server (portal server) provides the core presentation layer for retrieval, manipulation, analysis, and preformatting of data. The portal server retrieves data from the hub monitoring server in response to user actions at the portal client, and sends the data back to the portal client for presentation. The portal server also provides presentation information to the portal client so that it can render the user interface views suitably.

Tivoli Enterprise Monitoring Server

The Tivoli Enterprise Monitoring Server (monitoring server) is the collection and control point for performance and availability data and alerts that are received from monitoring agents (for example, the NetView agent). It is also responsible for tracking the online or offline status of monitoring agents.

The portal server communicates with the monitoring server, which in turn controls the remote servers and any monitoring agents that might be connected to it directly.

6.2.4 IBM Z System Automation for z/OS

IBM Z System Automation for z/OS is a cornerstone of all members of the GDPS family of products. In GDPS Continuous Availability, it provides the critical policy repository function, in addition to managing the automation of the workload and systems elements. System Automation for z/OS also provides the capability for GDPS to manage and monitor systems in multiple sysplexes.

System Automation for z/OS is required on the Controllers and all production systems that are running GDPS Continuous Availability workloads. If you use an automation product other than System Automation for z/OS to manage your applications, you do not need to replace your entire automation with System Automation. Your automation can coexist with System Automation and an interface is provided to ensure that proper coordination occurs.

6.2.5 IBM Multi-site Workload Lifeline for z/OS

This product provides intelligent routing recommendations to external load balancers or IBM MQ Queue Managers for server instances that can span two sysplexes or sites. Support is also provided for workloads that target SNA applications that are running in an Active/Standby configuration across two sites. Finally, user exits are provided with which you can manage the routing of workloads that use unsupported connectivity methods.

The IBM Multi-site Workload Lifeline for z/OS product consists of Advisors and Agents. One Lifeline Advisor is available that is active in the same z/OS image as the GDPS Primary Controller and assumes the role of primary Advisor. At most, one other Lifeline Advisor is active on the Backup Controller and assumes the role of secondary Advisor.

The two Advisors exchange state information so that the secondary Advisor can take over the primary Advisor role if the current primary Advisor is ended or a failure occurs on the system where the primary Advisor was active.

In addition, a Lifeline Agent is active on all z/OS images where workloads can run. All Lifeline Agents monitor the health of the images that they are running on and the health of the workload. These Agents communicate this information back to the primary Lifeline Advisor, which then calculates routing recommendations.

For TCP/IP-based routing, external load balancers establish a connection with the primary Lifeline Advisor and receive routing recommendations through the open-standard Server/Application State Protocol (SASP) application programming interface (API), which is documented in RFC 4678.

For IBM MQ based routing, the Lifeline Advisor and agents communicate with IBM MQ Queue managers to manage IBM MQ message traffic. Lifeline communicates with VTAM in the z/OS LPARs in which the workloads run to monitor and manage SNA-based traffic to the workloads.

Finally, user exits are started when routing decisions must be made for a workload that relies on connectivity methods that are not directly supported by IBM Multi-site Workload Lifeline.

The Lifeline Advisor also establishes a Network Management Interface (NMI) to allow network management applications (such as NetView) to retrieve internal data that the Advisor uses to calculate routing recommendations.

The Lifeline Advisors and Agents use configuration information that is stored in text files to determine what workloads must be monitored and how to connect to each other and external load balancers and IBM MQ Queue managers.

6.2.6 Middleware

Middleware components, such as CICS regions or Db2 subsystems, form a fundamental part of the GDPS Continuous Availability environment because they provide the application services that are required to process the workload.

To maximize the availability characteristics of the GDPS Continuous Availability environment, applications and middleware must be replicated across multiple images in the active and standby Parallel Sysplexes to cater for local HA if components fail. Automation must be in place to ensure clean start, shutdown, and local recovery of these critical components. CICS and Db2 workloads that are managed by CPSM derive more benefits in a GDPS Continuous Availability environment.

6.2.7 Replication software

Unlike in other GDPS solutions where the replication is based on mirroring the disk-based data at the block level (such as GDPS Metro or GDPS GM), replication in GDPS Continuous Availability is managed by software only. The following products are supported in GDPS Continuous Availability:

► IBM InfoSphere Data Replication for Db2 for z/OS (IIDR for Db2)

This product, also widely known as Q-rep, uses underlying IBM MQ for Z/OS as the transport infrastructure for moving the Db2 data from the source to the target copy of the database. Transaction data is *captured* at the source site and placed in IBM MQ queues for transmission to a destination queue at the target location, where the updates are then *applied* in real time to a running copy of the database.

For large-scale and update-intensive Db2 replication environments, a single pair of capture/apply engines might not be able to keep up with the replication. Q-rep provides a facility that is known as Multiple Consistency Groups (MCG) where the replication work is spread across multiple capture/apply engines, yet the time order (consistency) for the workload across all capture/apply engines is preserved in the target database. GDPS supports and provides specific facilities for workloads that use MCG with Db2 replication.

► IBM InfoSphere IMS Replication for z/OS (IIDR for IMS)

IBM InfoSphere IMS Replication for z/OS is the product that provides IMS data replication and uses a capture and apply technique that is similar to the technique that is outlined for Db2 data. However, IMS Replication does not use IBM MQ as the transport infrastructure to connect the source and target copies. Instead, TCP/IP is used in place of IBM MQ through the specification of hostname and port number to identify the target to the source and similarly to define the source to the target.

For large-scale and update-intensive IMS replication environments, a single pair of capture/apply engines might not be able to keep up with the replication. IIDR for IMS provides a facility that is known as *Global Consistency Groups* (GCGs) where the replication work is spread across multiple capture/apply engines.

> **Note:** For IMS GCG support, transaction consistency is *not* supported across multiple subscriptions.

► IBM InfoSphere VSAM Replication for z/OS (IIDR for VSAM)

IBM InfoSphere VSAM Replication for z/OS is similar in structure to the IMS replication product, except that it is for replicating VSAM data. For CICS VSAM data, the sources for capture are CICS log streams. For non-CICS VSAM data, CICS VSAM Recovery (CICS VR) is required for logging, and is the source for replicating such data. Similar to IMS replication, TCP/IP is used as the transport for VSAM replication.

GDPS Continuous Availability provides high-level control capabilities to start and stop replication between identified source and target instances through scripts and window actions in the GDPS GUI.

GDPS also monitors replication latency and uses this information when deciding whether Query workloads can be routed to the standby site.

6.2.8 Other optional components

Other components can optionally be used to provide specific monitoring, as described in this section.

Tivoli OMEGAMON XE family

Other products, such as IBM Z OMEGAMON Monitor for z/OS, IBM OMEGAMON XE for Db2 Performance Expert, and IBM OMEGAMON for IMS on z/OS, can be deployed to provide specific monitoring of products that are part of the GDPS Continuous Availability sites solution.

6.3 GDPS Continuous Availability environment

In this section, we provide a conceptual view of a GDPS Continuous Availability environment, plugging in the products that run on the various systems in the environment. We then take a closer look at how GDPS Continuous Availability works. Finally, we briefly discuss environments where GDPS Continuous Availability and other workloads coexist on the same sysplex.

Figure 6-3 shows the key components of a GDPS Continuous Availability environment.
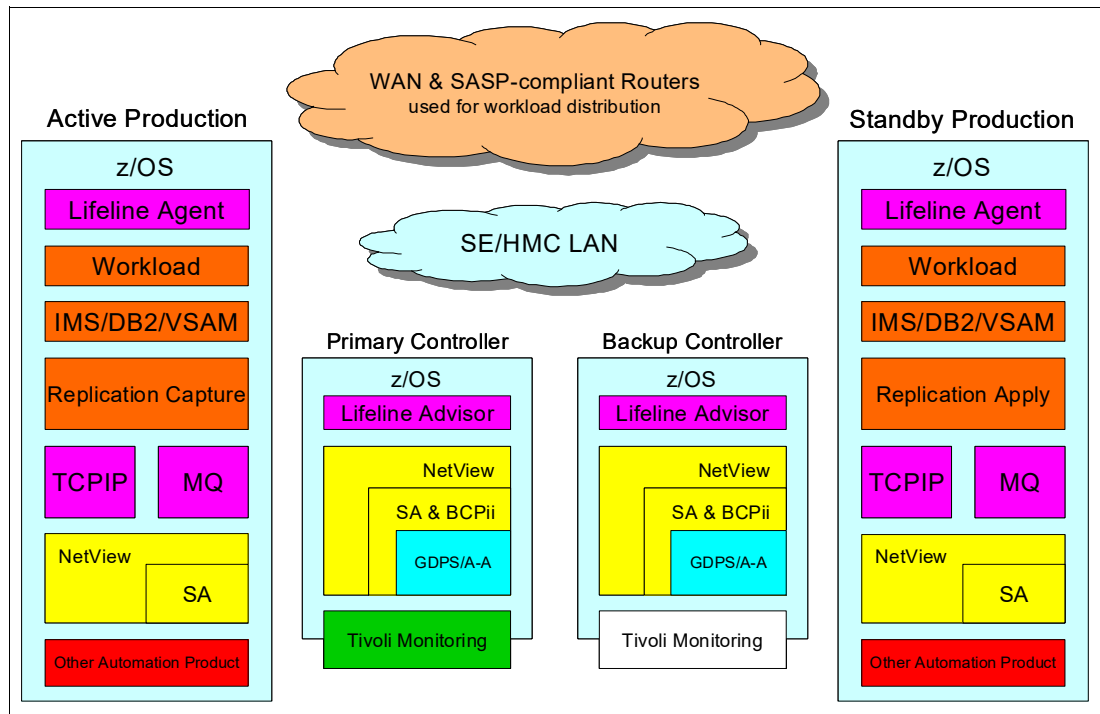

Figure 6-3   GDPS Continuous Availability environment functional overview

The GDPS Continuous Availability environment consists of two production sysplexes (also referred to as *sites*) in different locations. For each update workload that is to be managed by GDPS Continuous Availability, at any time, one of the sysplexes is the *active* sysplex and the other acts as *standby*.

As shown in Figure 6-3 on page 189, one workload and only one active production system is running this workload in one sysplex, and one production system is standby for this workload. However, multiple cloned instances of the active and the standby production systems can exist in the two sysplexes.

When multiple workloads are managed by GDPS, a specific sysplex can be the active sysplex for one update workload, while it is standby for another. It is the routing for each update workload that determines which sysplex is active and which sysplex is standby for a workload. As such, in environments where multiple workloads exist, no concept as an active sysplex is used. A sysplex that is the currently active one *for an update workload* is used.

The production systems (the active and the standby instances) are actively running the workload that is managed by GDPS. What makes a sysplex (and the systems in that sysplex) active or standby is whether update transactions are being routed to that sysplex.

The SASP routers in the network, which are shown in Figure 6-3 on page 189 as the cloud under GDPS and LifeLine Advisor, control routing of transactions for a workload to one sysplex or the other. Although a single router is the minimum requirement, we expect that you configure multiple routers for resiliency.

The workload is actively running on the z/OS system in both sysplexes. The workload on the system that is active for that workload is processing update transactions because update transactions are being routed to this sysplex.

The workload on the standby sysplex is actively running, but is not processing any update transactions because update transactions are not being routed to it. It is waiting for work, and can process work at any time if a planned or unplanned workload switch occurs that results in transactions being routed to this sysplex. If a workload switch occurs, the standby sysplex becomes the active sysplex for the workload.

The workload on the standby sysplex can be actively processing query transactions for the query workload that is associated with an update workload. Replication latency at any time, with thresholds that you specify in the GDPS policy, determines whether query transactions are routed to the standby sysplex.

Software replication policy indicates when the latency or the replication lag is considered to be too high (that is, the data in the standby sysplex is considered to be too far behind) when query transactions are no longer being routed there, but are routed to the active sysplex instead. When query transactions are no longer being routed to the standby sysplex because the latency threshold was exceeded, another threshold specified in the replication policy information indicates to Lifeline when it is OK to route query transactions to the standby sysplex again. Lifeline policy information indicates what percentage of the incoming query transactions are routed to the standby site or whether you want the conditions, such as latency and workload health, to dictate a dynamic decision on which of the two sysplexes query transactions are routed to at any time.

For example, your policy might indicate that query transactions for a workload are not routed to the standby sysplex if latency exceeds 7 seconds and that it is permitted to route to the standby sysplex after latency falls below 4 seconds. Latency is continually monitored to understand whether query transactions can be routed to the standby sysplex.

In addition to the latency control, you can specify a policy to indicate what percentage of the incoming query transactions are routed to the standby site or whether you want the conditions, such as latency and workload health, to dictate a dynamic decision on which of the two sysplexes query transactions are routed to at any time.

A GDPS CA workload constitutes all the subsystems that receive and process updates or query transactions accessing the replicated databases using the defined routing mechanism.

On the active system, you see a replication *capture engine*. One or more such engines can exist, depending on the data being replicated. This software replication component captures all updates to the databases that are used by the workload that is managed by GDPS and forwards them to the standby sysplex.

On the standby sysplex, the counterpart of the capture engine is the *apply engine*. The apply engine receives the updates that are sent by the capture engine and immediately applies them to the database for the standby sysplex.

The data replication in a GDPS environment is asynchronous (not all GDPS types are asynchronous). Therefore, the workload can perform a database update, and this write operation can complete, independent of the replication process.

Replication requires sufficient bandwidth for transmission of the data being replicated. IBM has services that can help you determine the bandwidth requirements based on your workload.

If replication is disrupted for any reason, the replication engines, when restored, include logic to know where they left off and can transmit only those changes that are made after the disruption.

Because the replication is asynchronous, no performance effect is associated with replication. For a planned workload switch, the switch can occur after all updates are drained from the sending side and applied on the receiving side.

For Db2 replication, GDPS provides extra automation to determine whether all updates drained. This feature allows planned switch of workloads by using Db2 replication to be automated.

For an unplanned switch, some data often is captured but not yet transmitted and applied on the target sysplex because replication is asynchronous. The amount of this data effectively converts to RPO.

With a correctly sized, robust transmission network, the RPO during normal operations is expected to be as low as a few seconds. You might also hear the term *latency* used with replication. Latency is another term that is used for the replication lag or RPO.

Although we talk about RPO, data is lost only if the original active site or the disks in this site where some updates were stranded are physically damaged so that they cannot be restored with the data intact. Following an unplanned switch to the standby site, if the former active site is restored with its data intact, any stranded updates can be replicated to the new active site then and no data is lost.

Also, specialized implementations of GDPS Continuous Availability, known as a *ZDL configuration*, can be used in some environments to provide an RPO of zero. This feature is available even when the disks in the site that failed were physically damaged. For more information about the ZDL configuration, see 6.6, "Zero Data Loss configuration" on page 207.

IBM MQ is shown on production systems and is required for Db2 replication. Either CICS or CICS VR are required on the production systems for VSAM replication.

On the production systems on both the active and standby sysplexes, you also see the monitoring and management products. NetView, System Automation, and the LifeLine Agent run on all production systems, monitoring the system, the workload on the system, and replication latency, and provide information to the GDPS Continuous Availability Controllers.

TCP/IP on the production systems is required in support of several functions that are related to GDPS Continuous Availability.

On the production systems, we show that you might have a product other than System Automation to manage your applications. In such an environment, System Automation is still required for GDPS Continuous Availability workload management. However, it is not necessary to replace your automation to use System Automation. A simple process for enabling the coexistence of System Automation and other automation products is available.

Not shown in Figure 6-3 on page 189 is the possibility of running other workloads that are not managed by GDPS Continuous Availability on the same production systems that run GDPS Continuous Availability workloads. For more information about other non-GDPS Continuous Availability workloads, see 6.3.2, "Considerations for other non-CA workloads" on page 196.

Figure 6-3 on page 189 shows two GDPS Controller systems. At any time, one is the Primary Controller, and the other is the Backup. These systems often are in each of the production sysplex locations, but they are not required to be collocated in this way.

GDPS Continuous Availability introduces the term *Controller*, as opposed to the *Controlling System* term that is used within other GDPS solutions. The function of the Primary Controller is to provide a point of control for the systems and workloads that are participating in the GDPS Continuous Availability environment for planned actions (such as IPL and directing, which is the active sysplex for a workload) and for recovery from unplanned outages. The Primary Controller is also where the data that is collected by the monitoring aspects of the solution can be accessed.

Both controllers run NetView, System Automation and GDPS Continuous Availability control code, and the LifeLine Advisor. The Tivoli Monitoring components Tivoli Enterprise Monitoring Server and IBM Z NetView Enterprise Management Agent run on the Controllers. Figure 6-3 on page 189 shows that a portion of Tivoli Monitoring is not running on z/OS. The Tivoli Enterprise Portal Server component can run on Linux on IBM Z or on a distributed server.

Together with System Automation on the Controllers, you see the BCP Internal Interface (BCPii). On the Controller, GDPS uses this interface to perform hardware actions against the LPAR of production systems or the LPAR of the other Controller system, such as LOAD and RESET, and for performing hardware actions for capacity on demand, such as CBU or OOCoD activation.

Figure 6-3 on page 189 also shows the Support Element/Hardware Management Console (SE/HMC) local area network (LAN). This element is key for the GDPS Continuous Availability solution.

The SE/HMC LAN spans the IBM Z servers for both sysplexes in the two sites. This configuration allows for a Controller in one site to act on hardware resources in the other site. To provide a LAN over large distances, the SE/HMC LANs in each site are bridged over the wide area network (WAN).

It is desirable to isolate the SE/HMC LAN on a network other than the client's WAN, which is the network that is used for the GDPS Continuous Availability application environment and connecting systems to each other. When isolated on a separate network, Lifeline Advisor (which is responsible for detecting failures and determining whether a sysplex failed) can try to access the site that appears to fail over the WAN and SE/HMC LAN.

If the site is accessible through the SE/HMC LAN but not the WAN, Lifeline can conclude that only the WAN failed, and not the target sysplex. Therefore, isolating the SE/HMC LAN from the WAN provides another check when deciding whether the entire sysplex failed and whether a workload switch is to be performed.

6.3.1 GDPS Continuous Availability: A closer look

In this section, we examine more closely how GDPS Continuous Availability works by using an example of a GDPS Continuous Availability environment with multiple workloads (see Figure 6-4). In this example, we consider update workloads only. Extending this example with query workloads corresponding to one or more of the update workloads can be a simple matter.
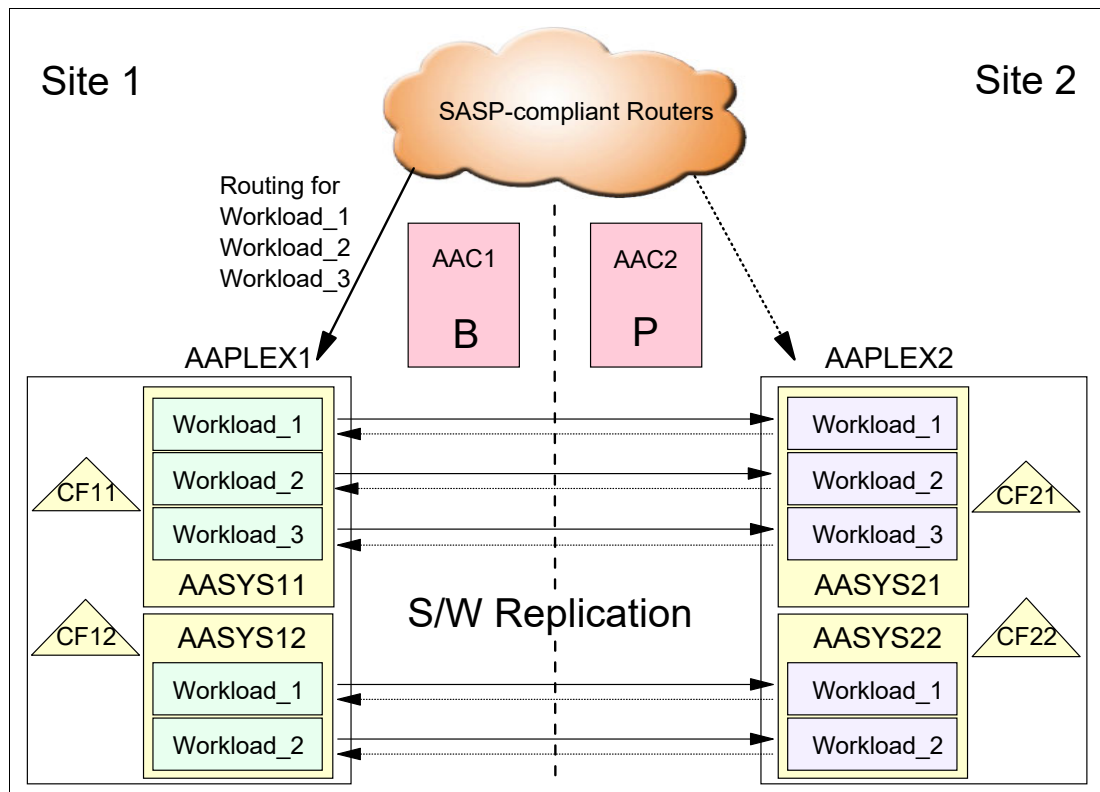
Figure 6-4   GDPS Continuous Availability environment with multiple workloads: All active in one site

Figure 6-4 shows two sites (Site1 and Site2) and a Parallel Sysplex in each site: AAPLEX1 runs in Site1 and AAPLEX2 runs in Site2. CFs CF11 and CF12 serve AAPLEX1 structures. CF21 and CF22 serve AAPLEX2 structures.

Each sysplex consists of two z/OS images. The z/OS images in AAPLEX1 are named AASYS11 and AASYS12. The images in AAPLEX2 are named AASYS21 and AASYS22. Two GDPS Controller systems are also shown: AAC1 in Site1, and AAC2 in Site2.

Three workloads are managed by GDPS in this environment: Workload_1, Workload_2, and Workload_3. Workload_1 and Workload_2 are cloned, which are Parallel Sysplex enabled applications that run on both z/OS images of the sysplexes. Workload_3 runs only in a single image in the two sysplexes.

Now, the transactions for all three workloads are routed to AAPLEX1. The workloads are running in AAPLEX2, but they are not processing transactions because no transactions are routed to AAPLEX2.

AAPLEX1 is the source for data replication for all three workloads, and AAPLEX2 is the target. Also shown in Figure 6-4 on page 193 are reverse replication links from AAPLEX2 toward AAPLEX1. This configuration indicates that if the workload is switched, the direction of replication can be and is switched.

If AASYS12 incurs an unplanned z/OS outage, all three workloads continue to run in AASYS11. Depending on the sizing of the systems, it is possible that AASYS11 does not feature sufficient capacity to run the entire workload.

Also, AASYS11 is now a single point of failure for all three workloads. In such a case where no workload failed but a possible degradation of performance and availability levels exists, you must decide whether you want to continue running all three workloads in AASYS11 until AASYS12 can be restarted or whether you switch one or more (or possibly all three) workloads to run in AAPLEX2 systems. Prepare for these decisions; that is, a so-called *pre-planned unplanned scenario*.

If you decide to switch one or more workloads to run actively in AAPLEX2, you often use a pre-coded planned action GDPS script to perform the switch of the workloads you want. Switching a workload in this case requires the following actions, by issuing a single GDPS command issued from a script, or by selecting the appropriate action from the GDPS GUI:

1. Stop the routing of transactions for the selected workloads to AAPLEX1.
2. Wait until all updates for the selected workloads on AAPLEX1 are replicated to AAPLEX2.
3. Stop replication for the selected workloads from AAPLEX1 to AAPLEX2.
4. Start the routing of transactions for the selected workloads to AAPLEX2.

After such a planned action script is started, it can complete the requested switching of the workloads in a matter of seconds.

As you can see, we do not stop the selected workloads in AAPLEX1. The workload does not need to be stopped for this specific scenario where we toggled the subject workloads to the other site to temporarily provide more capacity, remove a temporary single point of failure, or both.

We assumed in this case that AAPLEX2 had sufficient capacity available to run the workloads that are switched. If AAPLEX2 did not have sufficient capacity, GDPS can also activate OOCoD on one or more servers in Site2 that is running the AAPLEX2 systems before routing transactions there.

Now, assume that you decide to switch Workload_2 to Site2, but you keep Site1/AAPLEX1 as the primary for the other two workloads. When the switch is complete, the resulting position is shown in Figure 6-5 on page 195. In the figure, we assume that you also restarted in place the failed image, AASYS12.
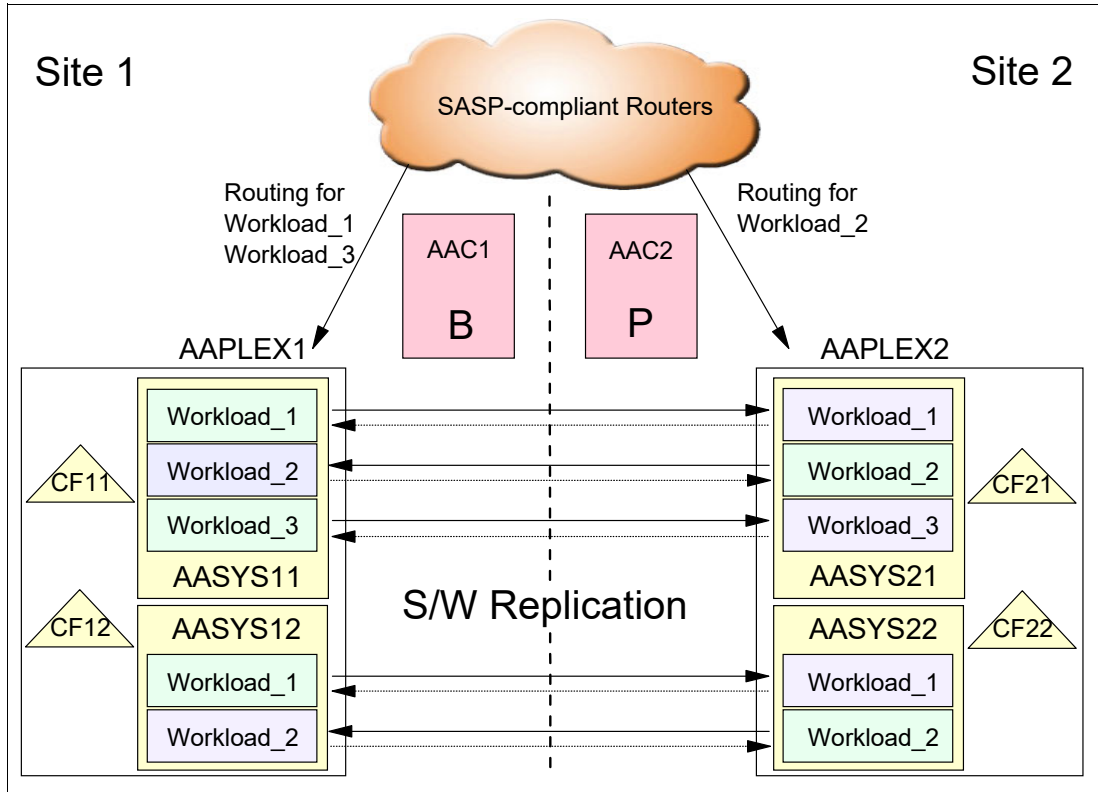
Figure 6-5 GDPS Continuous Availability environment with different workloads active in different sites

The router cloud shows to which site the transactions for each of the workloads are routed. Based on routing, AAPLEX2 is now the active sysplex for Workload_2. AAPLEX1 remains the active sysplex for Workload_1 and Workload_3. Replication for the data for Workload_2 is from AAPLEX2 to AAPLEX1. Replication for the other two workloads is still from AAPLEX1 to AAPLEX2.

The example that we discussed was an outage of AASYS12 that runs only cloned instances of the applications for Workload_1 and Workload_2. In contrast, Workload_3 does not include any cloned instances and runs only on AASYS11.

An unplanned outage of AASYS11 result in a failure of Workload_3 in its current sysplex. This failure is detected and based on your workload failure policy, can trigger an automatic switch of the failed workload to the sysplex that is standby for that workload.

However, if you do not want GDPS to perform automatic workload switch for failed workloads, you can select the option of an operator prompt. The operator is prompted whether GDPS is to switch the failed workload or not. If the operator accepts the switching of the workload, GDPS performs the necessary actions to switch the workload. No pre-coded scripts are necessary for this switch that results from a workload failure (automatic or operator confirmed). GDPS understands the environment and performs all the required actions to switch the workload.

Continuing with the same example where AASYS11 failed, which results in failure of Workload_3 in AAPLEX1, when GDPS performs the workload switch, AAPLEX2 becomes the active sysplex and AAPLEX1 is the standby. However, AAPLEX1 can serve only as standby when AASYS11 is restarted and Workload_3 is started on it.

Meanwhile, transactions are running in AAPLEX2 and updating the data for Workload_3. Until replication components of Workload_3 are restarted in AAPLEX1, the updates are not replicated from AAPLEX2 to AAPLEX1.

When replication components are restored on AAPLEX1, replication must be started for Workload_3 from AAPLEX2 to AAPLEX1. The replication components for Workload_3 on AAPLEX1 now resynchronize, and the delta updates that occurred while replication was down are sent across. When this process is complete, AAPLEX1 can be considered to be ready as the standby sysplex for Workload_3.

For an entire site/sysplex failure, GDPS provides similar capabilities as those capabilities for individual workload failure. In this case, multiple workloads might be affected.

Similar to workload failure, a policy determines whether GDPS is to automatically switch workloads that fail as a result of a site failure or perform a prompted switch. The only difference here is that the policy is for workloads that fail as a result of an entire site failure whereas in the previous example, we discussed the policy for individual workload failure.

You can specify for each workload individually whether GDPS is to perform an automatic switch or prompt the operator. Also, you can select a different option for each workload (automatic or prompt) for individual workload failure versus site failure.

For entire site or sysplex failures where multiple workloads are affected and switched, GDPS provides parallelization. The RTO for switching multiple workloads is much the same as switching a single workload.

Unplanned workload switches are expected to take slightly longer than planned switches because GDPS must wait to ensure that the unresponsive condition of the systems or workloads is not because of a temporary stall that can soon clear itself (that is, a false alarm). However, after the failure detection interval expires and the systems or workloads continue to be unresponsive, the workload switches are fast and performed in parallel for all workloads that are switched.

In summary, GDPS Continuous Availability manages individual workloads. Different workloads can be active in different sites. What is not allowed is for a particular workload to be actively receiving and running transactions in more than one site at any time.

6.3.2 Considerations for other non-CA workloads

In the same sysplex where CA workloads are running, you might have other workloads that are not managed by GDPS Continuous Availability.

In such an environment where CA and non-CA workloads coexist, it is important to provide the necessary level of isolation for the CA workloads and data. The data that belongs to the Active/Active workloads is replicated under GDPS Continuous Availability control and must not be used by non-managed applications.

Assume that a workload is active in Site1 and standby in Site2. Also, assume that a non-managed application is in Site1 that uses the same data that is used by your managed workload.

If you now switch your managed workload to Site2, the non-managed workload that is not included in the GDPS Continuous Availability solution scope continues to update the data in Site1 while the managed workload started to update the database instance in Site2. Such use of data that belongs to Active/Active workloads by non-managed applications can result in data loss, potential data corruption, and serious operational issues.

For this reason, the data that belongs to CA workloads must not be modified in any way by other applications. The simplest way to provide this isolation is to run CA workloads and other workloads in different sysplexes.

It might not be possible to provide sysplex-level isolation. If you are able to isolate your CA workloads and data, you likely still have other non-managed workloads and the data for such workloads coexisting in the same sysplex with your CA workloads. In this case, another solution, beyond GDPS Continuous Availability, such as GDPS Metro or GDPS GM, must be employed to protect the data and manage the recovery process for the non-CA workloads.

GDPS Continuous Availability includes specific functions to cooperate and coordinate actions with GDPS Metro that is running on the same sysplex. GDPS Metro can protect the entire sysplex, not just the systems that are running the CA workloads. For more information about this capability, see 6.5, "GDPS Continuous Availability co-operation with GDPS Metro" on page 205.

Because client environments and requirements vary, no "one size fits all" type of recommendation can be made. It is possible to combine GDPS Continuous Availability with various other hardware-replication-based GDPS products to provide a total recovery solution for a sysplex that houses Active/Active and other workloads.

If you cannot isolate your CA workloads into a separate sysplex, discuss this issue with your IBM GDPS specialist, who can provide you with guidance that is based on your specific environment and requirements.

6.4 GDPS Continuous Availability functions and features

In this section, we provide a brief overview of the following functions and capabilities that are provided by the GDPS Continuous Availability product:

► GDPS GUI
► Standard Actions for system/hardware automation
► Monitoring and Alerting
► GDPS scripts
► GDPS APIs

6.4.1 GDPS Continuous Availability Graphical User Interface (GUI)

GDPS Continuous Availability is operated on the Controller systems by using an operator interface that is provided through a GUI. Unlike other GDPS products, no 3270-based user interface is available with GDPS Continuous Availability.
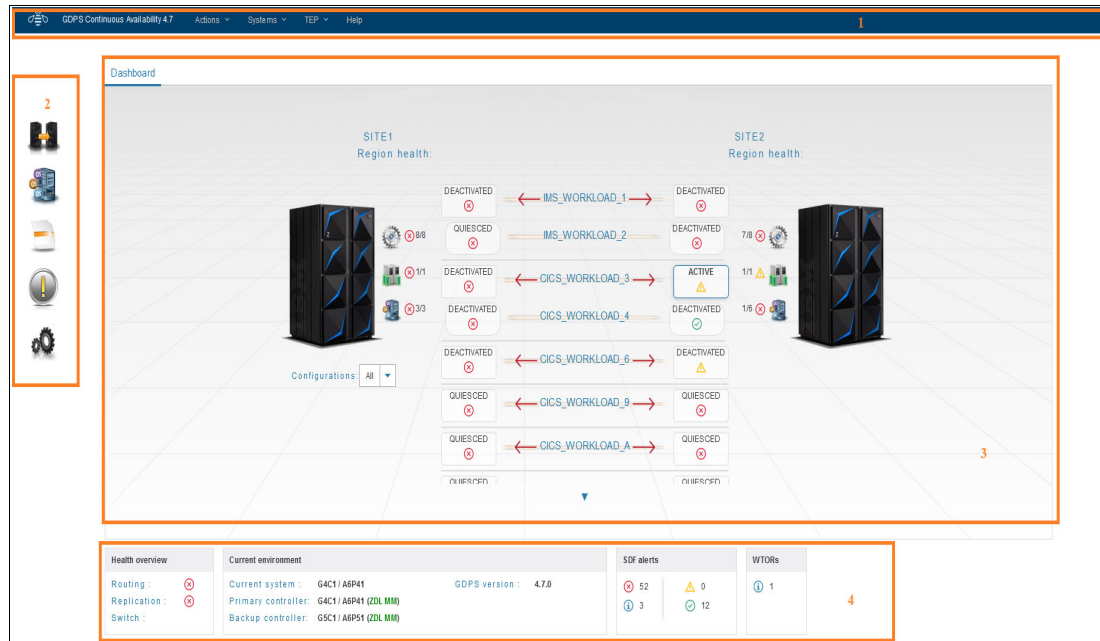
The GUI initial window is shown in Figure 6-6.

Figure 6-6   Initial GUI window

The GUI window features the following sections (the numbers here correspond to the numbers that are shown in Figure 6-6):

1. A header bar that includes the following components:

    –   The name of the GDPS solution.

    –   An Actions button with which you can run commands that are relevant to the current window.

    –   A Systems button that is used to change to a different NetView instance.

    –   A TEP button that is used to access the Tivoli Enterprise Portal. In addition to providing a monitoring interface to the overall solution, the Tivoli Enterprise Portal sets up specific situations for alerting of conditions, such as the replication latency exceeding a certain threshold.

        The workload-related workspaces can also quickly show such things as the number of servers active in both sites and to where the routing is active. This information can be useful to correlate against that information that is shown in the GDPS web interface to confirm the status of any particular resources.

    –   A Help button.

    –   A button that indicates the logged-on user ID, from which you can change the refresh rate and the order of display of workloads, and log out.

2. An application menu, which provides access to the following:

    –   Standard Actions
    –   CPC Operations
    –   Workload Management
    –   Script management
    –   Status Display Facility (SDF) Alerts
    –   CANZLOG, NetView, write to operator with reply (WTOR) messages
    –   Debug settings
    –   Dump settings.

3. The main area of the window, which is known as the Dashboard. It displays information about the state of replication of workloads between sysplexes.

4. The status summary area. It provides general information about the status of routing, replication, and switching, and the system and controllers, plus SDF Alerts and WTOR messages.

Most frames include a Help button to provide extensive help text that is associated with the information that is displayed and the selections that are available on that specific frame.

6.4.2 GDPS Continuous Availability scripts

We reviewed the GDPS web interface, which provides powerful functions to help you manage your workloads and systems in the sites where they are running. However, the GDPS web interface is not the only means for performing these functions.

Nearly all functions that can be manually started by the operator through the web interface are also available through GDPS scripts. Other actions are not available through the web interface, such as activating capacity on demand (CBU or OOCoD) that are only possible by using GDPS scripts. In addition to the set of script commands that are supplied by GDPS, you can integrate your own REXX procedures and run them as part of a GDPS script.

A *script* is a procedure that is recognized by GDPS that pulls together into a workflow (or a list) one or more GDPS functions to be run one after the other. GDPS checks the result of each command and proceeds with the next command only if the previous command ran successfully. Scripts can be started manually through the GDPS windows (by using the Planned Actions interface), automatically by GDPS in response to an event (Unplanned Actions), or through a batch interface.

Scripts are simple to code. The use of scripts forces you to plan for the actions you must take for various planned and unplanned outage scenarios, and how to bring the environment back to normal. In this sense, when you use scripts, you plan even for an unplanned event and are not caught unprepared. This aspect is an important aspect of GDPS. Scripts are powerful because they can use the full capability of GDPS.

The ability to plan and script your scenarios and start all GDPS functions provides the following benefits:

► Speed

A script runs the requested actions as quickly as possible. Unlike a human, it does not need to search for the latest procedures or the commands manual. It can check results fast and continue with the next statement immediately when one statement is complete.

► Consistency

If you look into most computer rooms immediately following a system outage, what would you see? Mayhem. Operators frantically scrambling for the latest system programmer instructions. All the phones ringing. Every manager within reach asking when the service will be restored. And every systems programmer with access vying for control of the keyboards. All this chaos results in errors because humans naturally make mistakes when under pressure. But with automation, your well-tested procedures run in the same way, time after time, regardless of how much you shout at them.

► Thoroughly thought-out and tested procedures

Because they behave in a consistent manner, you can test your procedures over and over until you are sure that they do everything that you want, in exactly the manner that you want. Also, because you must code everything and cannot assume a level of knowledge (as you might with instructions that are intended for a human), you are forced to thoroughly think out every aspect of the action the script is intended to undertake. Because of the repeatability and ease of use of scripts, they lend themselves more easily to frequent testing than manual procedures.

► Reduction of requirement for onsite skills

How many times have you seen DR tests with large numbers of people onsite for the test and many more standing by for a call? How realistic is this scenario? Can all these people be onsite on short notice if a catastrophic failure occurred?

The use of GDPS automation and scripts removes the need for the numbers and the range of skills that enterprises traditionally needed to do complex or compound reconfiguration and recovery actions.

Planned Actions

Planned Actions are GDPS scripts that are started from the GUI by using the Planned Actions frame, as described in Figure 6-6 on page 198. GDPS scripts are procedures that pull together into a list one or more GDPS functions to be run sequentially. Scripted procedures that you use for a planned change to the environment are known as *control scripts*.

A control script that is running can be stopped if necessary. Control scripts that were stopped or that failed can be restarted at any step of the script. These capabilities provide a powerful and flexible workflow management framework.

As a simple example, you can have a script that recycles a z/OS system. This action is performed if you apply maintenance to the software that required a restart of the system. The script runs the STOP standard action, which performs an orderly shutdown of the target system followed by a LOAD of the same system.

However, it is possible that in your environment you use alternative system volumes. While your system runs on one set of system volumes, you perform maintenance on the other set. So, assuming that you are running on alternative SYSRES1 and you apply this maintenance to SYSRES2, your script also must point to SYSRES2 before it performs the LOAD operation.

As part of the customization that you perform when you install GDPS, you can define entries with names of your choice for the load address and load parameters that are associated with the alternative SYSRES volumes for each system. When you want to LOAD a system, you use a script statement to point to one of these pre-customized entries by using the entry name that you used when defining them to GDPS.

Example 6-1 shows a sample script to perform this action. In this example, MODE=ALTRES2 points to the load address and load parameters that are associated with alternative SYSRES2 where you applied your maintenance.

Example 6-1   Sample script to restart a system on an alternative SYSRES

```
COMM='Re-IPL system AASYS11 on alternate SYSRES2'
SYSPLEX='STOP AASYS11'
IPLTYPE='AASYS11 MODE=ALTRES2'
SYSPLEX='LOAD AASYS11'
```

Example 6-2 shows a sample script to switch a workload from its current active site to its standby site.

Example 6-2   Sample script to switch a workload between sites

```
COMM='Switch WORKLOAD_1'
ROUTING 'SWITCH WORKLOAD=WORKLOAD_1'
```

No target site is specified in the `ROUTING SWITCH` statement. This specification is not made because GDPS is aware of where WORKLOAD_1 is active and GDPS switches it to the other site. The single `ROUTING SWITCH` statement performs the following actions:

► Stops routing of update transactions to the original active site.

► Fences the DBMS tables, data sets, and objects in the original active site to prevent updates other than those updates that are made by replication.

► Waits for replication of the final updates in the current active site to drain.

► Unfences the DBMS tables, data sets, and objects in the new active site to updates by the applications.

► Starts routing update transactions to the former standby site, which now becomes the new active site for this workload.

► If a query workload is associated with this update workload, and if, for example, 70% of queries were being routed to the original standby site, the routing for the query workload is changed to send 70% of queries to the new standby site after the switch.

All of these actions are done as a result of running a single script with a single command. This feature demonstrates the simplicity and power of GDPS scripts.

Our final example for the use of a script can be for shutting down an entire site, perhaps in preparation for disruptive power maintenance at that site. For this example, we use the configuration with three workloads, all active in Site1, as shown in Figure 6-7.
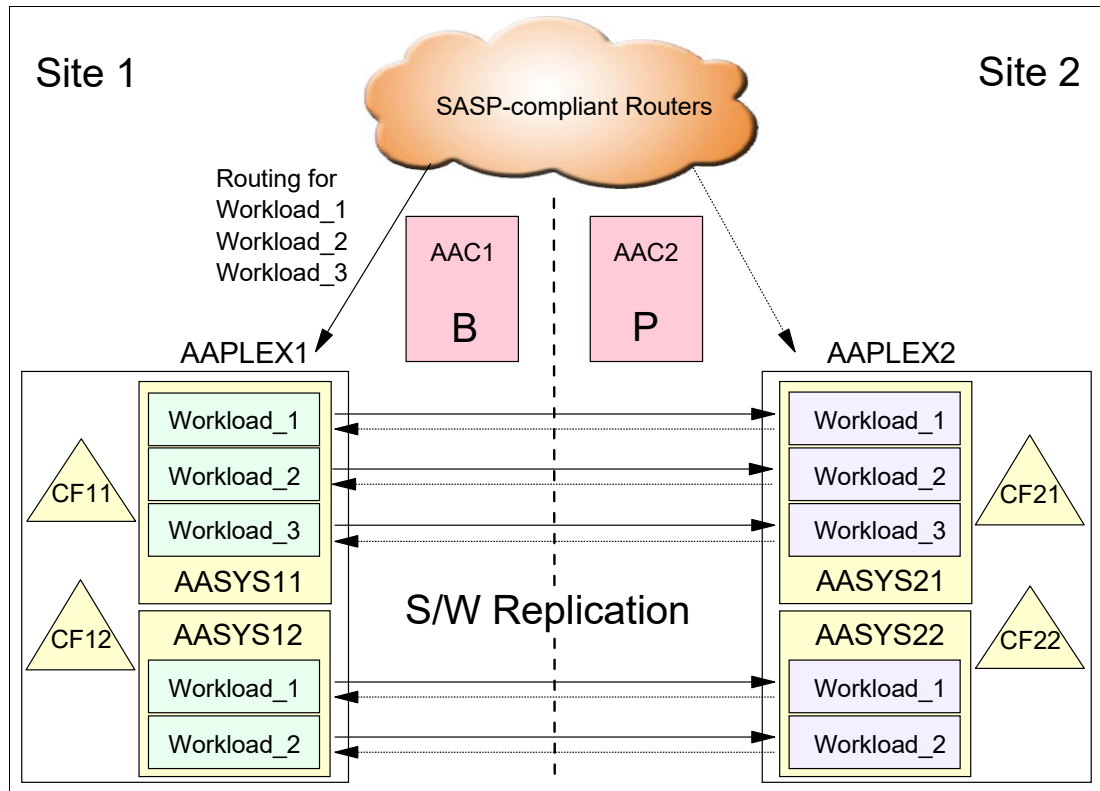

Figure 6-7   GDPS Continuous Availability environment sample for Site1 shutdown script

The following sequence is used to completely shut down Site1:

1. Activate OOCoD on the CPCs that are running the AAPLEX2 systems and CFs (although not shown in Figure 6-7, we assume that the CPCs are named `CPC21` and `CPC22` for this example).

2. Switch transaction routing for all workloads to `AAPLEX1`.

3. Stop replication from `AAPLEX1` to `AAPLEX2`.

4. Stop the `AASYS11` and `AASYS12` systems.

5. Deactivate the system and CF LPARs in Site1.

The planned action script to accomplish the Site1 shutdown for this environment is shown in Example 6-3.

Example 6-3   Sample Site1 shutdown script

```
COMM='Switch all workloads to Site2 and Stop Site1'
OOCOD='ACTIVATE CPC=CPC21 ORDER=order#'
OOCOD='ACTIVATE CPC=CPC22 ORDER=order#'
ROUTING='SWITCH WORKLOAD=ALL'
REPLICATION='STOP WORKLOAD=ALL FROM=AAPLEX1 TO=AAPLEX2'
SYSPLEX='STOP SYSTEM=(AASYS11,AASYS12)'
SYSPLEX='DEACTIVATE AASYS11'
SYSPLEX='DEACTIVATE AASYS12'
SYSPLEX='DEACTIVATE CF11'
SYSPLEX='DEACTIVATE CF12'
```

These sample scripts demonstrate the power of the GDPS scripting facility. Simple, self-documenting script statements drive compound and complex actions. A single script statement can operate against multiple workloads or multiple systems. A complex procedure can be described in a script by coding only a handful of statements.

Another benefit of such a facility is the reduction in skill requirements to perform the necessary actions to accomplish the task at hand. For example, in the workload switch and the site shutdown scenarios (depending on your organizational structure within the IT department), you might require database, application/automation, system, and network skills to be available to perform all required steps in a coordinated fashion.

Batch scripts

GDPS also provides a flexible batch interface to initiate scripts to make planned changes to your environment. These scripts, which are known as *batch scripts,* cannot be started from the GDPS GUI. Instead, they are started from some other planned event that is external to GDPS. For example, the starting event can be a job or messages that are triggered by a job scheduling application.

This capability, along with the Query Services that are described in "GDPS Continuous Availability Query Services" on page 204, provides a rich framework for user-customizable automation and systems management procedures.

Switch scripts

As described in 6.3.1, "GDPS Continuous Availability: A closer look" on page 193, if a workload or entire site fails, GDPS performs the necessary steps to switch one or more workloads to the standby site. This switching, which is based on the selected policy, can be automatic with no operator intervention or can occur after operator confirmation. However, in either case, the steps that are required to switch any workload are performed by GDPS and no scripts are required for this process.

Although GDPS performs the basic steps to accomplish switching affected workloads, you might want GDPS to perform more actions that are specific to your environment along with the workload switch steps. One such example can be activating CBU for more capacity in the standby site.

Switch scripts are unplanned actions that run as a result of a workload failure or site failure that is detected by GDPS. These scripts cannot be activated manually. They are started automatically if you coded them as a result of an automatic or prompted workload or site switch action that is started by GDPS. The intent of Switch scripts is to complement the standard workload or site switch processing that is performed by GDPS.

6.4.3 Application programming interfaces

GDPS provides two primary programming interfaces to allow other programs that are written by clients: independent software vendors (ISVs), and other IBM product areas to communicate with GDPS. These APIs allow clients, ISVs, and other IBM product areas to complement GDPS automation with their own automation code. The following sections describe the APIs provided by GDPS.

GDPS Continuous Availability Query Services

GDPS maintains configuration information and status information in NetView variables for the various elements of the configuration that it manages. GDPS Query Services is a capability that allows client-written NetView REXX programs to query the value for numerous GDPS internal variables. The variables that can be queried pertain to the GDPS environment (such as the version and release level of the GDPS control code), sites, sysplexes, and workloads that are managed by GDPS Continuous Availability.

In addition to the Query Services function, which is part of the base GDPS product, GDPS provides several samples in the GDPS SAMPLIB library to demonstrate how Query Services can be used in client-written code.

RESTful APIs

As described in "GDPS Continuous Availability Query Services", GDPS maintains configuration information and status information about the various elements of the configuration that it manages. Query Services can be used by REXX programs to query this information.

The GDPS RESTful API also provides the ability for programs to query this information. Because it is a RESTful API, it can be used by programs that are written in various programming languages, including REXX, that are running on various server platforms.

In addition to querying information about the GDPS environment, the GDPS RESTful API allows programs that are written by clients, ISVs, and other IBM product areas to start actions against various elements of the GDPS environment. These actions include the following examples:

- ▶ Starting and stopping workloads
- ▶ Starting, stopping, and switching routing for one or more workloads
- ▶ Starting and stopping software replication
- ▶ Starting and stopping systems
- ▶ Running scripts
- ▶ Starting GDPS monitor processing.

These capabilities enable clients, ISVs, and other IBM product areas to provide an even richer set of functions to complement the GDPS functions.

GDPS provides samples in the GDPS SAMPLIB library to demonstrate how the GDPS RESTful API can be used in programs.

6.4.4 Securing the GDPS environment

GDPS uses RACF XFACILIT and GXFACILI resource classes to create a role-based security model for controlling access to the resources in your GDPS environment that is customized to your specific environment. Simple definitions can be used to control access at the panel option level or more granular definitions can be used to control access to specific types of resources, or even all the way down to the specific resource level.

With the role-based security model, you can create your own roles or use the common roles that GDPS recommends, which include GDPS Administrator, GDPS Operator, GDPS User, and Non-GDPS User. You define the resources that these roles can access and the type of access they have to those resources by granting them access to the resource profiles that represent the various resources in your environment. Finally, you can grant access to various resources to users by adding them to the appropriate roles.

When you use the role-based security model, GDPS ensures that the user has sufficient authority to take a specific action against a specific resource, regardless of whether they are attempting to take the action by using the GUI directly or by running a GDPS script.

Finally, the GDPS Security Definition Utility tool is available to help you with implementing your role-based security environment. For more information, see the *GDPS Security Definition Utility* guide, which can be found by licensed users on the GDPS FTP server.

6.5 GDPS Continuous Availability co-operation with GDPS Metro

In a GDPS Continuous Availability environment, each of the sysplexes that are running the CA workloads must be is as highly available as possible. As such, we suggest that the CA workloads are Parallel Sysplex enabled data sharing applications. Although this configuration eliminates a planned or unplanned system outage from being a single point of failure, disk data within each local sysplex is not protected by Parallel Sysplex alone.

To protect the data for each of the two sysplexes that comprise the GDPS Continuous Availability environment, these sysplexes can be running GDPS Metro with Metro Mirror replication and HyperSwap, which complement and enhance local high, CA for the sysplex or sysplexes. For more information about the various capabilities that are available with GDPS Metro, see Chapter 3, "IBM GDPS Metro" on page 45.

With GDPS Continuous Availability and GDPS Metro monitoring and managing the same production systems for a particular sysplex, certain actions must be coordinated. This requirement is necessary so that the GDPS that is controlling systems for the two environments do not interfere with each other or that one environment does not misinterpret actions that are taken by the other environment.

For example, it is possible that one of the systems in the sysplex needs to be restarted for a software maintenance action. The restart of the system can be performed from a GDPS Continuous Availability Controller or by using GDPS Metro that is running on all systems in the same sysplex.

Assume that you start the restart from the GDPS Continuous Availability Controller. GDPS Metro detects that this system is no longer active. It interprets what was a planned restart of a system as a system failure and issues a takeover prompt.

The GDPS Continuous Availability co-operation with GDPS Metro provides coordination and serialization of actions across the two environments to avoid issues that can stem from certain common resources being managed from multiple control points. In our example, when you start the restart from the CA Controller, it communicates this action to the GDPS Metro controlling system.

The GDPS Metro controlling system then locks this system as a resource so that no actions can be performed against it until the CA Controller signals completion of the action. This same type of coordination occurs regardless of whether the action is started by GDPS Continuous Availability or GDPS Metro.

GDPS Continuous Availability can support coordination with GDPS Metro that is running in or both of the CA sites.

In Figure 6-8, we show a GDPS Continuous Availability environment across two regions, Region A and Region B. SYSPLEXA in Region A and SYSPLEXB in Region B comprise the two sysplexes that are managed by GDPS Continuous Availability. Systems AAC1 and AAC2 are the GDPS Continuous Availability Controller systems. Also, each of these sysplexes is managed by an instance of GDPS Metro, with systems KP1A/KP2A being the GDPS Metro controlling systems for SYSPLEXA and KP1B/KP2B being the GDPS Metro controlling systems in SYSPLEXB.

The GDPS Continuous Availability Controllers communicate to each of the GDPS Metro controlling systems in both regions. It is this communication that makes the cooperation possible.
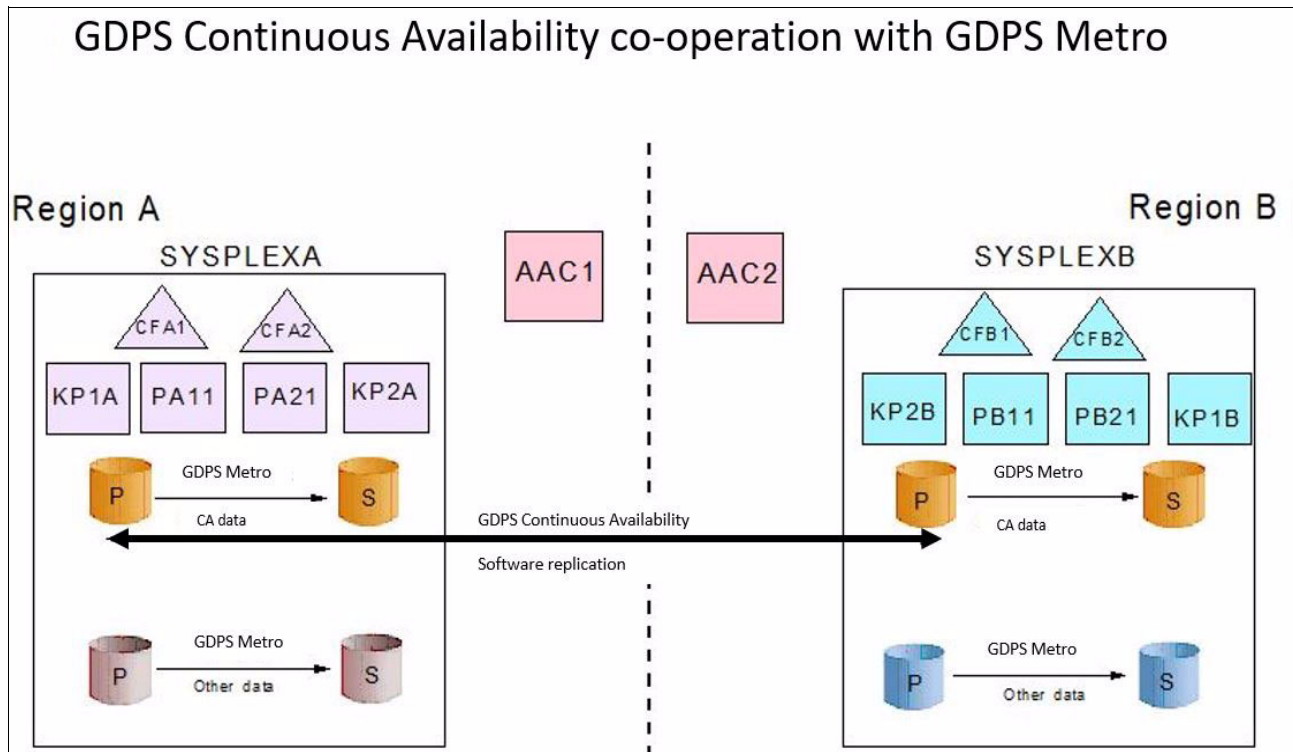

Figure 6-8   GDPS Continuous Availability co-operation with GDPS Metro

SYSPLEXA contains the data for the CAworkloads, and other data for applications that are running in the same sysplex but not managed by CA. Also included are the various system infrastructure data that is also not managed by CA.

All of this data that belongs to SYSPLEXA is replicated within Region A by using Metro Mirror and is HyperSwap protected and managed by GDPS Metro. The CA data is replicated through software to SYSPLEXB.

Similarly, another instance of GDPS Metro is available that is managing SYSPLEXB with the Active/Active data and also any non-CA data belonging to SYSPLEXB being replicated through Metro Mirror and HyperSwap protected within Region B.

Each SYSPLEXA and SYSPLEXB can be running in a single physical site or across two physical sites within their respective regions.

All data within both sysplexes is HyperSwap protected, meaning that a disk within a region is not a single point of failure and the sysplex can continue to function during planned or unplanned disk outages. HyperSwap is transparent to all applications that are running in the sysplex (assuming that the data for all applications is replicated with Metro Mirror). Therefore, it is also transparent to all subsystems in charge of running the Active/Active workloads, replicating the Active/Active data, and monitoring the Active/Active environment.

HyperSwap of disks within a region is transparent to the cross-region software replication process. Software replication knows about and captures data from the logs on the current primary Metro Mirror volumes only. If a HyperSwap exists, software replication continues capturing data from the logs that are now on the new primary volumes.

In addition to the HyperSwap protection, GDPS Metro provides several other facilities and benefits, which are described in Chapter 3, "IBM GDPS Metro" on page 45.

Because of the capabilities of GDPS Metro, we expect that clients perform most of the day-to-day system, sysplex, and Metro Mirror management activities for each of the two sysplexes that use GDPS Metro facilities. However, GDPS Continuous Availability must be used for management and switching of the Active/Active workloads and replication.

Finally, management (actions, such as STOP and IPL) of the Active/Active controllers can be performed only by using GDPS Continuous Availability because these systems are outside of the respective sysplexes and GDPS Metro scope of control is limited to the systems in the sysplex.

In summary, GDPS Continuous Availability and GDPS Metro can be deployed in a complementary fashion. These products provide the necessary controls to facilitate any coordination that is required when operating on common resources.

6.6 Zero Data Loss configuration

The ZDL configurations are specialized implementations of GDPS Continuous Availability that can support ZDL for an unplanned outage in the active site. With a ZDL implementation, if the active site suffers an outage, the latest updates are available on disk in the standby site, and therefore are not lost as they are in a non-ZDL configuration.

At a high level, ZDL is achieved by using disk mirroring to maintain a copy of the primary disk in the standby site and performing the software replication capture and apply process from that secondary disk in the standby site. By using this configuration, if the active site suffers an outage, the latest updates are available on disk in the standby site and are not lost because any such updates are in the "normal" or non-ZDL model[3].

---

[3] To achieve ZDL in site outage scenarios, a STOP policy must be in effect for Metro Mirror replication and primary DASD failure events. No update can be made to the primary copy of the data if it cannot be replicated to the secondary copy that is used for the replication capture process in the Standby site.

There are two possible implementations of ZDL:

Metro ZDL (ZDL MM)   This implementation allows for ZDL when the two GDPS CA sites are within supported Metro Mirror distance[4].

Achieved by placing a secondary copy of the primary disk for the active workloads in the standby site and performing both the software replication capture and apply process in the standby site (see Figure 6-10 on page 210).

Global ZDL (ZDL GM)   This implementation allows for ZDL when the distance between the two GDPS CA sites exceeds that supported for Metro Mirror.

Achieved by placing a secondary copy of the primary disk for the active workloads in an intermediate site, called SiteB for the purposes of this discussion. SiteB is within Metro Mirror distance of the Active site and updates to the Active site are copied to SiteB by using Metro Mirror, under the management of GDPS Metro. GDPS Global - GM (GM) then replicates data that was copied to SiteB to the Standby site, across unlimited distance, by using GM (see Figure 6-11 on page 211).

SiteB is logically isolated and also physically secure enough to survive a disaster in Sysplex A. For ZDL GM, the assumption is that SiteB survives long enough to mirror all updates to the GM secondaries in the Standby site.

The ZDL configurations are defined at a workload level, allowing ZDL and non-ZDL workloads to be operated in the same GDPS CA environment.

ZDL is a separately priced feature of GDPS Continuous Availability.

---

[4] The maximum supported Metro Mirror replication distance without RPQ is 300KM. However, depending on your configuration, the response time might not be acceptable at such a distance.

6.6.1 Db2 replication in a non-ZDL environment

Figure 6-9 shows an overview of the normal software replication model in GDPS AA where the capture process runs in the same site as the active workload.
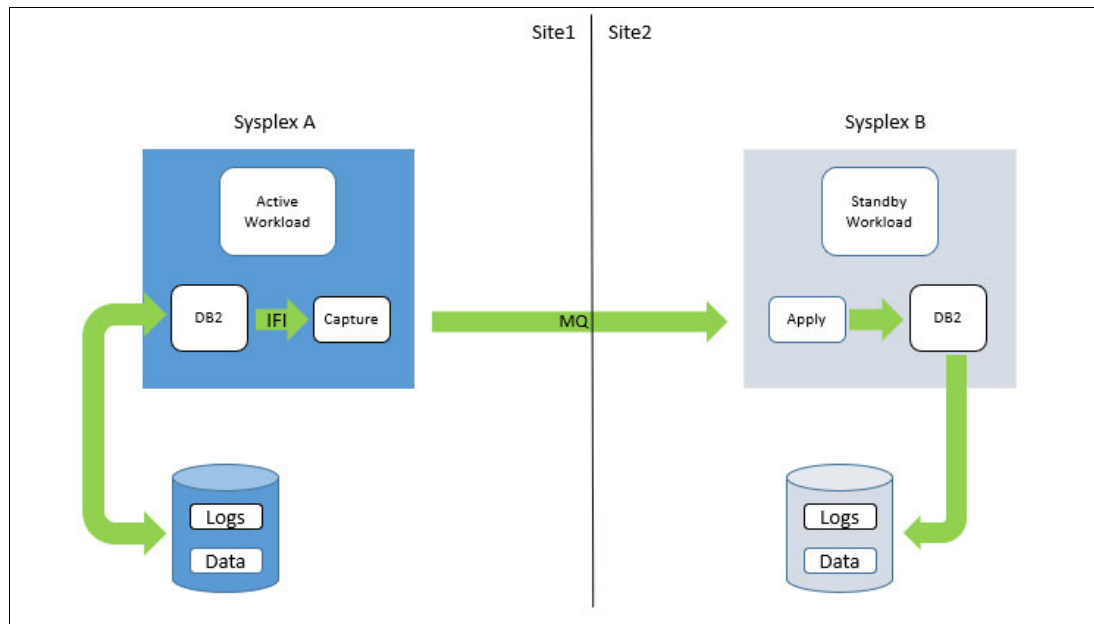

Figure 6-9   Overview of Db2 replication in a non-ZDL environment

Because software replication is asynchronous in nature, an unplanned loss of the active site (Site1) results in loss of the workload and the capture process and is also highly likely to leave so-called stranded transactions that were not yet sent to the standby site (Site2) to be applied. These stranded transactions can become "lost" transactions if the unplanned outage is of a catastrophic nature and the data cannot be retrieved later.

6.6.2 Db2 ZDL replication within Metro Mirror distances

For Db2 workloads where such potential for data loss is unacceptable, if the two GDPS AA sites are within acceptable Metro Mirror distances for your configuration, Metro ZDL can be used to deliver ZDL. Figure 6-10 shows an overview of the Metro ZDL configuration.
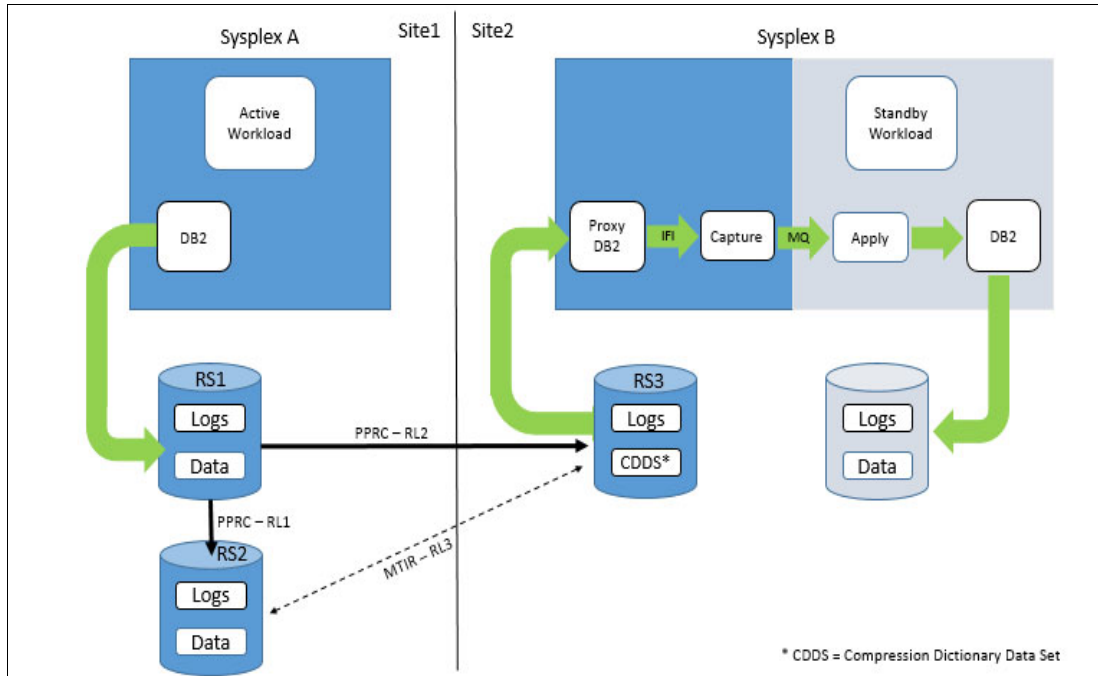

Figure 6-10   Overview of ZDL configuration within Metro Mirror distances

In Figure 6-10, RS3, which is in the standby site (Site2), is a PPRC secondary copy of the primary data (RS1) from the active site (Site1). There is also a PPRC secondary copy (RS2) of the active site data in Site1 for HA purposes.

RS3 is *not* a copy of all the data on RS1; only specific data that is required for the software replication capture process needs to be replicated by using Metro Mirror to Site2. This subset of data that is required is the Db2 logs and the Db2 compression dictionary. RS2 is a full copy of RS1 for HA (HyperSwap) purposes. All three of these copies of data are managed by using the GDPS Metro solution.

Running in Site2 are several Proxy systems (at least 2 for HA) that are driving the software replication capture process. These systems are part of the same sysplex (shown as Sysplex B in Figure 6-10) as the systems running the apply process and the standby workloads.

The systems that are running the capture process use specialized capabilities to read from the Metro Mirror secondary volumes to access the Db2 log information that is required for the capture process. This information determines the changes that need to be sent (still over IBM MQ as in a normal Db2 replication implementation) to the apply process for writing to the standby copy of the data.

Symmetrical Metro ZDL configuration

Figure 6-10 on page 210 shows an *asymmetrical* ZDL configuration. That is, it supports the ZDL function for workloads that are active in Site1 and standby in Site2 only. Because no Metro Mirror relationship exists to mirror the Sysplex B Database updates back to Site1, traditional software replication (as shown in Figure 6-9 on page 209) must be used to mirror the updates, which are made by workloads that are active in Site2, back to the Sysplex A Database in Site1.

A *symmetrical* ZDL configuration is supported for Metro ZDL. For a symmetrical configuration, the disks that contain the Sysplex B Database in Site2 also are in a Multi-Target Metro Mirror (MTMM) relationship such that the logs and CDDS for the Sysplex B database are mirrored synchronously to Site1.

Several systems that are running in Sysplex A drive the software replication capture process by using the specialized capabilities to read from the Metro Mirror secondary volumes. As such, the Db2 log information can be accessed, which is required for the capture process to determine the changes that must be mirrored to the Sysplex A Database in Site1.

6.6.3 Db2 replication over Global Mirror distances

For workloads where the potential for data loss is unacceptable but the distance between the two GDPS AA sites exceeds Metro Mirror distances, Global ZDL can be used to deliver ZDL. Figure 6-11 provides an overview of the Metro ZDL configuration.
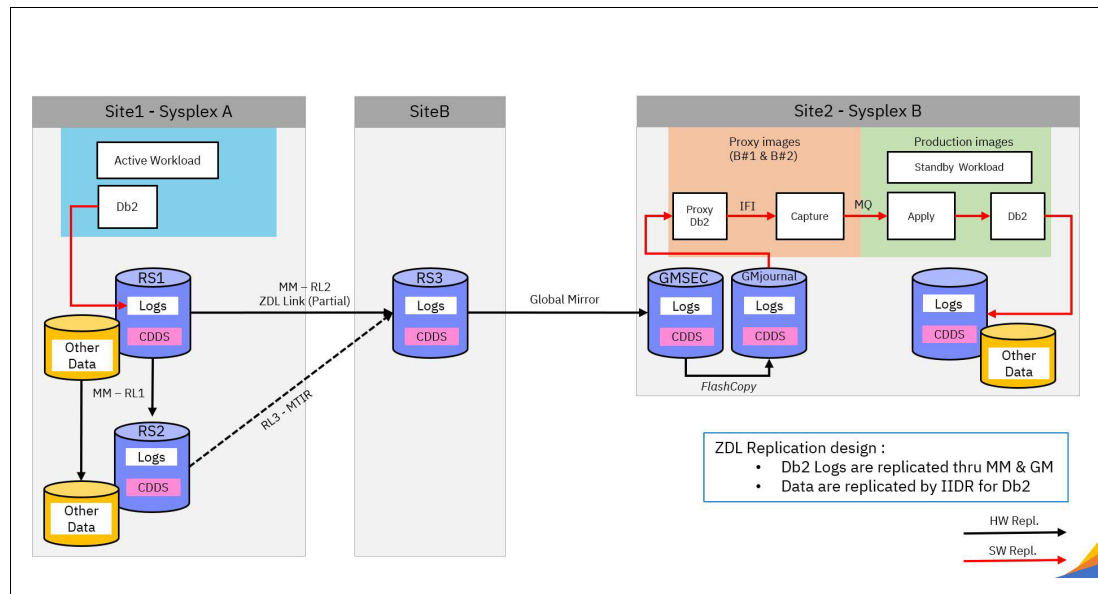

Figure 6-11   ZDL configuration over Global Mirror distances

In Figure 6-11, RS3, which is in an intermediate site (called SiteB for the purposes of this discussion), is a PPRC secondary copy of the primary data (RS1) from the active site (Site1). There is also a PPRC secondary copy (RS2) of the active site data in Site1 for HA purposes.

RS3 is *not* a copy of all the data on RS1 because only specific data that is required for the software replication capture process must be replicated to SiteB by using Metro Mirror. The subset of data that is required is the Db2 logs and the Db2 compression dictionary. RS2 is a full copy of RS1 for HA (HyperSwap) purposes. All three copies of data are managed by using the GDPS Metro solution.

The GDPS GM Kg system that is managing the GM is running in Site1. The GM DS8K Master is in SiteB. After the GM session is set up correctly, it does not need the Kg system in an unplanned switch. GM replicates the data from RS3 to the GM secondary volumes that are in Site2. GM flash copies the copied data to the GM journal volumes. There is not GDPS Metro code, GDPS Global - GM code, or user code running in SiteB.

In Site2, several systems are running (there must be at least two Proxy systems for HA) that drive the software replication capture process. These systems are part of the same sysplex (shown as Sysplex B in Figure 6-11 on page 211) as the systems that are running the apply process and the standby workloads. The systems running the capture process use specialized capabilities to read from the GM journal volumes to access the Db2 log information that is required for the capture process to determine the changes that must be sent (over IBM MQ as in a normal Db2 replication implementation) to the apply process for writing to the standby copy of the data.

> **Note:** ZDL GM is an asymmetric configuration that allows only ZDL replication from Site1 to Site2. GDPS CA does not support ZDL GM from Site2 to Site1.

6.7 Flexible testing with GDPS Continuous Availability

The best testing possible to understand if a workload can run in the recovery location is to run it there. GDPS Continuous Availability is well positioned for this task because the application is running in both locations and providing some level of confidence that the infrastructure in either site can sustain the workload. For complete confidence, you must also switch the workload so that the previously standby instance becomes active and processes transactions for some time.

Toggling a workload between sites in a GDPS Continuous Availability setup can be simple. The workload can be periodically switched to run in one site or other site (in a planned manner) in a matter of seconds, with no data loss. Running the workload live in the other site with transactions being routed to it gives you the best opportunity to assess whether adjustments are needed to your infrastructure or operating procedures and ensures that both of your sites are ready to assume live workloads.

Creating workload failures to test unplanned workload switching can also be simple. However, we know that not all clients are willing to do such tests live in their production environments, no matter how small the expected service disruption might be.

A best practice is to maintain a sandbox testing environment that closely represents the production environment. A sandbox testing environment for Parallel Sysplex can be extended so that you have a test GDPS Continuous Availability environment by adding another sandbox sysplex in the other site and a couple of Controllers for the sandbox GDPS.

If you do not have a sandbox sysplex but have, for example, a development sysplex, this sandbox can be extended to serve as a testing environment. With such a test environment, you can test new levels of software components or maintenance to these components before you introduce such changes into production. Also, use test GDPS Continuous Availability to test planned and at least some portion of unplanned outage and switch scenarios before they are tested in production. Various scenarios might never get tested in production, in which case the testing in the test environment can still provide an indication about whether the solution is set up properly and can be expected to work.

6.8 GDPS Continuous Availability services

GDPS Continuous Availability includes more than data replication. It also includes many other aspects of your environment, such as sysplex, automation, network, workload routing, workload management, testing processes, and planned and unplanned outage scenario testing.

Most installations do not have all of these skills readily available. It is rare to find a team with this range of skills across many implementations. However, the GDPS Continuous Availability offering includes access to a global team of specialists in all the disciplines you need to ensure a successful GDPS implementation.

The most successful GDPS projects are those projects in which IBM and client skills form a unified team to perform the implementation.

Specifically, the Services component of GDPS Continuous Availability includes some or all of the following tasks:

► Planning to determine availability requirements, configuration recommendations, implementation, and testing plans

► Installation and necessary customization of:
  – NetView
  – System Automation (customization for coexistence with other automation products)
  – Multi-site Workload Lifeline Advisor
  – Tivoli Monitoring

► Data replication implementation:
  – Bandwidth analysis
  – Installation and necessary customization of InfoSphere Data Replication Server for z/OS (Db2, IMS, or VSAM)

► Setup of SASP-compliant routers and switches

► Assistance with cross-site connectivity for WAN and SE/HMC LAN

► GDPS Continuous Availability automation code installation and customization:
  – Training on GDPS Continuous Availability setup and operations
  – Assistance with planning, coding, and testing GDPS scripts and scenarios
  – Assistance with planning and implementing GDPS Continuous Availability cooperation, integration with GDPS Metro, GDPS GM, or all of them

► Project management and support throughout the engagement

The services that IBM can provide with a HADR project are not restricted to those services that we described in this chapter. Here, we provided a list of services that specifically relate to GDPS Continuous Availability implementation.

The sizing of the services component of each project is tailored for that project based on many factors, including what automation or replication is in place and which of the prerequisite products are installed. Therefore, the services and skills that are provided as part of those services are tailored to the specific needs of each particular client and implementation.

6.9 GDPS Continuous Availability prerequisites

For more information about the current list of prerequisites for GDPS Continuous Availability, see this web page.

6.10 Comparing GDPS Continuous Availability to other GDPS offerings

In each of the chapters that describe the other GDPS products that are based on hardware replication, we provide a table to compare the characteristics of these solutions against each other at a high level. We do not include GDPS Continuous Availability in these comparisons because it is a somewhat "apples-to-oranges" comparison.

You saw that GDPS Continuous Availability is fundamentally different than the other GDPS products. It is based on software replication rather than hardware. It also is workload-level management and switch rather than system-level management and restart.

Also, we discussed how GDPS Continuous Availability is not necessarily mutually exclusive with other GDPS products and how GDPS Metro or GDPS MGM can be combined with GDPS Continuous Availability to provide a comprehensive, robust near-CA and DR solution for your enterprise.

For more information about basic positioning and comparison for GDPS Continuous Availability against the other GDPS products, see 6.1.1, "Positioning GDPS Continuous Availability" on page 180.

6.11 Summary

GDPS Continuous Availability is a powerful offering that facilitates near-instantaneous switching of workloads between two sites that can be separated by virtually unlimited distances. Based on asynchronous software replication, planned switches can be accomplished with no data loss (RPO 0). When sufficient replication bandwidth is provided, the RPO can be as low as a few seconds for an unplanned workload switch.

GDPS Continuous Availability provides a range of capabilities, through an intuitive web interface or by using simple and powerful scripting, for workload management, workload routing, data replication management, management of system and hardware resources for planned and unplanned events. Through extensive monitoring and failure detection mechanisms, unplanned workload switches can be automated, which removes human intervention and optimizes RPO.

For enterprises that require high levels of protection with near zero RPO and RTO at distances beyond the practical reach of a GDPS Metro multi-site workload configuration, GDPS Continuous Availability is uniquely positioned to meet these requirements for critical workloads.

7

IBM GDPS Virtual Appliance

In this chapter, we provide an overview of the IBM Geographically Dispersed Parallel Sysplex (GDPS) Virtual Appliance offering. The GDPS Virtual Appliance supports both planned and unplanned situations, which helps to maximize application availability and provide business continuity. In particular, a GDPS Virtual Appliance solution can deliver the following capabilities:

- ► Near-continuous availability (CA) solution
- ► Disaster recovery (DR) solution across metropolitan distances
- ► Recovery time objective (RTO) less than an hour
- ► Recovery point objective (RPO) of zero

The main objective of the GDPS Virtual Appliance is to provide these capabilities to clients that use z/VM and Linux on IBM Z and do not have z/OS in their environments[1]. The virtual appliance models that are used by this offering results in a solution that is easily managed and operated without requiring z/OS skills.

The functions provided by the GDPS Virtual Appliance fall into two categories: protecting your data and controlling the resources that are managed by GDPS. These functions include the following items:

- ► Protecting your data:

  – Ensures the consistency of the secondary data if there is a disaster or suspected disaster, including the option to also ensure zero data loss (ZDL)

  – Transparent switching to the secondary disk by using HyperSwap

  – Management of the remote copy configuration

---

[1] For clients who run z/OS and have z/OS skills, equivalent capabilities exist by using the GDPS Metro Multiplatform Resiliency for IBM Z as described in "Multiplatform Resiliency for z/VM" on page 71.

215

- Controlling the resources managed by GDPS during normal operations, planned changes, and following a disaster:
  - Monitoring and managing the state of the production Linux for IBM Z guest images and LPARs (shutdown, activating, deactivating, IPL, and automated recovery)
  - Support for switching your disk, or systems, or both, to another site
  - User-customizable scripts that control the GDPS Virtual Appliance action workflow for planned and unplanned outage scenarios

This chapter includes the following topics:

7.1 Introducing the GDPS Virtual Appliance

The GDPS Virtual Appliance is a CA and DR solution that handles many types of planned and unplanned outages. As mentioned in Chapter 1, "Introducing business resilience and the role of IBM GDPS" on page 1, most outages are planned, and even among unplanned outages, most are not disasters. The GDPS Virtual Appliance provides capabilities to help provide the required levels of availability across these outages and in a disaster scenario. This chapter describes the data integrity and availability protection and the systems management capabilities provided by the GDPS Virtual Appliance.

The term *production system* is used throughout this chapter to refer to any z/VM images together with the Linux on IBM Z guests that are being managed by this instance of the GDPS Virtual Appliance.

7.2 GDPS Virtual Appliance configuration components

This section contains a high-level description of the components in a GDPS Virtual Appliance configuration. The components consist of both hardware and software. The hardware includes the disk subsystems that contain the production data and the remote copy services that perform the data replication. The software components include GDPS and other automated management code that runs on the GDPS Virtual Appliance, and GDPS Multiplatform Resiliency for IBM Z (also known as xDR), which runs on the z/VM systems that are managed by the GDPS Virtual Appliance. Figure 7-1 on page 217 shows an example of a GDPS Virtual Appliance environment.
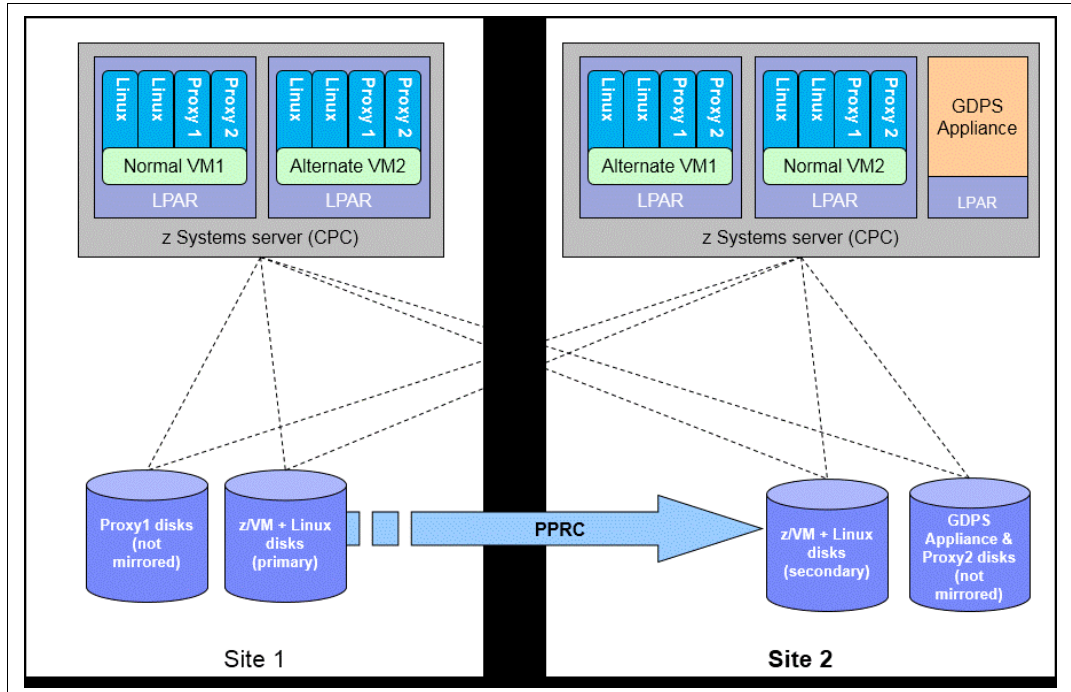
Figure 7-1   GDPS Virtual Appliance Environment

7.2.1 GDPS Virtual Appliance

The GDPS Virtual Appliance is a self-contained system that includes the GDPS Metro software that provides monitoring and management of Metro Mirror replication of production disk and monitoring and management of the z/VM systems that are using the production disk. Use the GDPS Appliance to initiate planned events and to perform situation analysis after an unplanned event to determine the status of the production systems or the disks, and then to drive automated recovery actions. The GDPS Virtual Appliance is responsible for carrying out all actions during a planned event or following a disaster or potential disaster, such as managing the disk mirroring configuration, initiating a HyperSwap, initiating a freeze and implementing the freeze/swap policy actions, restarting failed systems, and others.

A GDPS Virtual Appliance environment is typically spread across two data centers (Site1 and Site2) where the primary copy of the production disk is normally in Site1. The GDPS Appliance must have connectivity to all the Site1 and Site2 primary and secondary devices that it manages. For availability reasons, the GDPS Virtual Appliance runs in Site2 on a local disk that is not mirrored with Metro Mirror, which provides failure isolation for the appliance system to ensure that it is not impacted by failures that affect the production systems and remains available to automate any recovery action.

7.2.2 Multiplatform Resiliency for IBM Z

The GDPS Virtual Appliance provides automated management of z/VM systems with a function called "Multiplatform Resiliency for IBM Z (also known as xDR)". To provide these capabilities, the GDPS Virtual Appliance communicates and coordinates with System Automation for Multiplatforms (SA MP) running on Linux on IBM Z.

In each GDPS xDR-managed z/VM system, you must configure two special Linux guests, which are known as the *proxy* guests, as shown in Figure 7-1 on page 217. One proxy node is configured on Site1 disk and the other is configured on Site2 disk. The proxies are guests that are dedicated to providing communication and coordination with the GDPS Virtual Appliance They must run SA MP with the separately licensed xDR feature.

The proxy guests serve as the middleman for GDPS. They communicate commands from GDPS to z/VM, monitor the z/VM environment, and communicate status information and failure information (such as a HyperSwap triggers affecting the z/VM disk) back to the GDPS Virtual Appliance. At any time, the proxy node that is running on disk in the Metro Mirror secondary site is the *Master proxy,* which is the proxy node with which the GDPS Virtual Appliance coordinates actions. The proxy node Master role is switched automatically when Metro Mirror disk is switched (or recovered) or when the Master proxy fails.

The disks that is being used by z/VM, the guest machines, and the proxy guest in this configuration must be Count-Key-Data (CKD) disks.

z/VM provides a HyperSwap function. With this capability, the virtual device that is associated with one real disk can be swapped transparently to another disk. GDPS coordinates planned and unplanned HyperSwap for z/VM disks, providing continuous data availability For site failures, GDPS provides a coordinated Freeze for data consistency across all z/VM systems.

GDPS can perform a graceful shutdown of z/VM and its guests and perform hardware actions such as LOAD and RESET against the z/VM system's partition. GDPS supports taking a PSW restart dump of a z/VM system. Also, GDPS can manage Capacity BackUp (CBU) and On/Off Capacity on Demand (OOCoD) for IFLs and CPs on which z/VM systems are running.

7.3 Protecting data integrity and data availability with the GDPS Virtual Appliance

In 2.2, "Data consistency" on page 18, we point out that data integrity across primary and secondary volumes of data is essential to perform a database restart and accomplish an RTO of less than an hour. This section provides details about how GDPS automation in the GDPS Virtual Appliance provides both data consistency if there are mirroring problems and data availability if there are disk problems.

The following types of disk problems trigger a GDPS automated reaction:

► Mirroring problems (Freeze triggers)

No problem exists writing to the primary disk subsystem, but a problem exists mirroring the data to the secondary disk subsystem. For more information, see 7.3.1, "GDPS Freeze function for mirroring failures" on page 219.

► Primary disk problems (HyperSwap triggers)

There is a problem writing to the primary disk: Either a hard failure, or the disk subsystem is not accessible or not responsive. For more information, see 7.3.2, "GDPS HyperSwap function" on page 220.

7.3.1 GDPS Freeze function for mirroring failures

GDPS uses automation to stop all mirroring when a remote copy failure occurs. In particular, the GDPS automation uses the IBM PPRC Freeze/Run architecture, which is implemented as part of Metro Mirror on IBM disk subsystems and also by other enterprise disk vendors. In this way, if the disk hardware supports the Freeze/Run architecture, GDPS can ensure consistency across all data for the managed systems (consistency group) regardless of disk hardware type. This preferred approach differs from proprietary hardware approaches that work only for one type of disk hardware. For more information about data consistency with synchronous disk mirroring, see "Metro Mirror data consistency" on page 24.

When a mirroring failure occurs, this problem is classified as a Freeze trigger and GDPS stops activity across *all* disk subsystems at the time the initial failure is detected, thus ensuring that the dependent write consistency of the remote disks is maintained.

The following actions occur when a GDPS performs a Freeze:

► Remote copy is suspended for all device pairs in the configuration.

► While the suspend command is being processed, each device goes into a long busy state.

► No I/Os can be issued to the affected devices until the long busy state is thawed with the PPRC Run (or "thaw") action or until it times out. The consistency group timer setting commonly defaults to 120 seconds, although for most configurations a longer or Extended Long Busy (ELB) setting is preferred.

► All paths between the Metro Mirrored disks are removed, preventing further I/O to the secondary disks if Metro Mirror is accidentally restarted.

Because no I/Os are processed for a remote-copied volume during the ELB, dependent write logic ensures the consistency of the remote disks. GDPS performs a Freeze for all Metro Mirrored devices in the GDPS managed configuration.

> **Important:** Because of the dependent write logic, it is not necessary for all devices to be frozen at the same instant. In a large configuration with many thousands of remote copy pairs, it is not unusual to see short gaps between the times when the Freeze command is issued to each disk subsystem. However, because of the ELB such gaps are not a problem.

After GDPS automation performs the Freeze and the consistency of the remote disks is protected, the GDPS Virtual Appliance will perform a Run action against all logical subsystems (LSSs), which removes the ELB and allows production systems to continue using these devices. The devices are in remote copy-suspended mode, meaning that any further writes to these devices are no longer being mirrored. However, changes are tracked by the hardware so that only the changed data is resynchronized to the secondary disks later.

If the Freeze trigger turns out to be the first sign of an actual disaster, your z/VM systems might continue operating for an amount of time before those systems fail. Any updates made to the primary volumes during this time will not be replicated to the secondary disk, and are therefore lost.

The GDPS Virtual Appliance uses a combination of storage subsystem and production system triggers to automatically secure, at the first indication of a potential disaster, a data-consistent secondary site copy of your data by using the Freeze function. In this way, the secondary copy of the data is preserved in a consistent state, even before production applications are aware of any issues. Ensuring the data consistency of the secondary copy ensures that a normal system restart can be performed instead of having to perform database management system (DBMS) forward recovery actions, which is an essential design element of GDPS to minimize the time to recover the critical workloads if there is a disaster at the primary site.

You can appreciate why such a process must be automated. When a device suspends, there is not enough time to start a manual investigation process.

In summary, freeze is triggered as a result of a Metro Mirror suspension event for any primary disk in the GDPS Virtual Appliance configuration; that is, at the first sign that a duplex mirror that is going out of the duplex state. When a device suspends, all attached systems are sent a State Change Interrupt (SCI). A message is issued in all of those systems and then each virtual machine (VM) system must issue multiple I/Os to investigate the reason for the suspension event.

When GDPS performs a freeze, all primary devices in the Metro Mirror configuration suspend, which can result in significant SCI traffic and many messages in all systems. With z/VM and microcode on the DS8000 disk subsystems, GDPS supports reporting suspensions in a summary message per LSS instead of at the individual device level. When compared to reporting suspensions on a per device basis, the Summary Event Notification for PPRC Suspends (PPRCSUM) dramatically reduces the message traffic and extraneous processing that are associated with Metro Mirror suspension events and freeze processing.

7.3.2 GDPS HyperSwap function

If there is a problem writing or accessing the *primary* disk because of a failing, failed, or non-responsive primary disk, then there is a need to swap from the primary disks to the secondary disks.

The GDPS Virtual Appliance delivers a powerful function that is known as HyperSwap. HyperSwap swaps from using the primary devices in a mirrored configuration to using what were the secondary devices, transparent to the production systems and applications using these devices.

Without HyperSwap, a transparent disk swap is not possible. All systems that use the primary disk must be shut down (or might fail, depending on the nature and scope of the failure) and must restart by using the secondary disks. Disk failures are often a single point of failure for the entire production environment.

With HyperSwap, such a switch can be accomplished without IPL and with just a brief hold on application I/O. The HyperSwap function is completely controlled by automation, thus allowing all aspects of the disk configuration switch to be controlled through GDPS.

HyperSwap can be started in two ways:

► Planned HyperSwap

A planned HyperSwap is started by operator action by using GDPS facilities. One example of a planned HyperSwap is where a HyperSwap is initiated before planned disruptive maintenance to a disk subsystem.

► Unplanned HyperSwap

An unplanned HyperSwap is started automatically by GDPS, triggered by events that indicate the primary disk problem.

Primary disk problems can be detected as a direct result of an I/O operation to a specific device that fails because of a reason that indicates a primary disk problem such as:

– No paths available to the device
– Permanent error
– I/O timeout

In addition to a disk problem being detected as a result of an I/O operation, it is also possible for a primary disk subsystem to proactively report that it is experiencing an acute problem. The IBM DS8000 family has a special microcode function that is known as the *Storage Controller Health Message Alert* capability. Problems of different severity are reported by disk subsystems that support this capability. Those problems classified as acute are also treated as HyperSwap triggers. After systems are swapped to use the secondary disks, the disk subsystem and operating system can try to perform recovery actions on the former primary without impacting the applications that use those disks.

Planned and unplanned HyperSwap have requirements in terms of the physical configuration, such as having it symmetrically configured.

When a swap is initiated, GDPS always validates various conditions to ensure that it is safe to swap. For example, if the mirror is not fully duplex, that is, not all volume pairs are in a duplex state, a swap cannot be performed. The way that GDPS reacts to such conditions changes depending on the condition that is detected and whether the swap is a planned or unplanned swap.

Assuming that there are no show-stoppers and the swap proceeds, for both planned and unplanned HyperSwap, the systems that are using the primary volumes experience a temporary pause in I/O processing. GDPS blocks I/O both at the channel subsystem level by performing a Freeze that results in all disks going into ELB, and also in all systems, where I/O is quiesced at the operating system (UCB) level. This action ensures that no systems use the disks until the switch is complete.

During the time when I/O is paused, the following process is completed:

1. The Metro Mirror configuration is *physically switched*, which includes physically changing the secondary disk status to primary. Secondary disks are protected and cannot be used by applications. Changing their status to primary allows them to come online to systems and be used.

2. The disks are *logically switched* in each of the systems in the GDPS configuration. This involves switching the internal pointers in the operating system control blocks. After the switch, the operating system will point to the former secondary devices instead of the current primary devices.

3. For planned swaps, optionally, the mirroring direction can be reversed.

4. Finally, the systems resume operation by using the new, swapped-to primary devices. The applications are not aware of the fact that different devices are now being used.

This brief pause during which systems are locked out of performing I/O is known as the *User Impact Time*.

The GDPS Virtual Appliance HyperSwaps all devices in the managed configuration. Just as the Freeze function applies to the entire consistency group, HyperSwap is for the entire consistency group. For example, if a single mirrored volume fails and HyperSwap is started, processing is swapped to the secondary copy of *all* mirrored volumes for all managed systems in the configuration, including volumes in unaffected subsystems because to maintain disaster readiness all primary volumes *must* be in the same site. If HyperSwap were to swap the only failed devices, you would then have several primaries in one site, and the remainder in the other site. This action would also make for a complex environment to operate and administer I/O configurations.

HyperSwap with less than full channel bandwidth

You might consider enabling unplanned HyperSwap even if you do not have sufficient cross-site channel bandwidth to sustain the full production workload for normal operations. Assuming that a disk failure is likely to cause an outage and you need to switch to using disk in the other site, the unplanned HyperSwap might at least present you with the opportunity to perform an orderly shutdown of your systems first. Shutting down your systems cleanly avoids the complications and restart time elongation that is associated with a crash-restart of application subsystems.

7.3.3 GDPS usage of DS8000 functions

GDPS strives to use (when it makes sense) enhancements to the IBM DS8000 disk technologies. In this section, we provide information about the key DS8000 technologies that the GDPS Virtual Appliance supports and uses.

Metro Mirror Failover/Failback support

When a primary disk failure occurs and the disks are switched to the secondary devices, Metro Mirror Failover/Failback (FO/FB) support eliminates the need to do a full copy when reestablishing replication in the opposite direction. Because the primary and secondary volumes are often in the same state when the freeze occurred, the only differences between the volumes are the updates that occur to the secondary devices after the switch. Failover processing sets the secondary devices to primary suspended status and starts change recording for any subsequent changes made. When the mirror is reestablished with failback processing, the original primary devices become secondary devices and a resynchronization of changed tracks takes place.

The GDPS Virtual Appliance requires Metro Mirror FO/FB capability to be available on all disk subsystems in the managed configuration.

PPRC eXtended Distance (PPRC-XD)

PPRC-XD (also known as Global Copy) is an asynchronous form of the PPRC copy technology. GDPS uses PPRC-XD rather than synchronous PPRC (Metro Mirror) to reduce the performance impact of certain remote copy operations that potentially involve a large amount of data. For more information, see 4.6.2, "GDPS HM reduced impact on initial copy and resynchronization" on page 140.

Storage Controller Health Message Alert

This feature facilitates triggering an unplanned HyperSwap proactively when the disk subsystem reports an acute problem that requires extended recovery time. For more information about unplanned HyperSwap triggers, see 7.3.2, "GDPS HyperSwap function" on page 220.

PPRCS Summary Event Messages

GDPS supports the DS8000 PPRC Summary Event Messages (PPRCSUM) function, which is aimed at reducing the message traffic and the processing of these messages for Freeze events. For more information, see 7.3.1, "GDPS Freeze function for mirroring failures" on page 219.

Soft Fence

Soft Fence provides the capability to block access to selected devices. As discussed in 7.3.4, "Protecting secondary disks from accidental update" on page 223, GDPS uses Soft Fence to avoid write activity on disks that are exposed to accidental update in certain scenarios.

On-demand dump (also known as non-disruptive statesave)

When problems occur with disk subsystems such as those which result in an unplanned HyperSwap, a mirroring suspension or performance issues, a lack of diagnostic data from the time the event occurs can result in difficulties in identifying the root cause of the problem. Taking a full statesave can lead to temporary disruption to host I/O and is often frowned upon by clients for this reason. The on-demand dump (ODD) capability of the disk subsystem facilitates taking a non-disruptive statesave (NDSS) when such an event occurs. The microcode does this task automatically for certain events, such as a dump of the primary disk subsystem that triggers a Metro Mirror freeze event. It also allows an NDSS to be requested. This enables first failure data capture (FFDC) and thus ensures that diagnostic data is available to aid problem determination. Not all information that is contained in a full statesave is contained in an NDSS and therefore there might still be failure situations where a full statesave is requested by the support organization.

GDPS supports taking an NDSS by using the GDPS GUI. In addition to this support, GDPS autonomically takes an NDSS if there is an unplanned Freeze or HyperSwap event.

7.3.4 Protecting secondary disks from accidental update

A system cannot be restarted by using a disk that is physically a Metro Mirror secondary disk because Metro Mirror secondary disks cannot be brought online to any systems. However, a disk can be secondary from a GDPS (and application use) perspective but physically have a simplex or primary status from a Metro Mirror perspective.

For both planned and unplanned HyperSwap, and a disk recovery, GDPS changes former secondary disks to primary or simplex state. However, these actions do not modify the state of the former primary devices, which remain in the primary state. Therefore, the former primary devices remain accessible and usable even though they are considered to be the secondary disks from a GDPS perspective, which makes it is possible to accidentally update or IPL from the wrong set of disks. Accidentally using the wrong set of disks can result in a potential data integrity or data loss problem.

The GDPS Virtual Appliance provides protection against using the wrong set of disks in the following ways:

► If you attempt to load a system through GDPS (either script or panel or GUI) by using the wrong set of disks, GDPS rejects the load operation.

► GDPS uses a DS8000 disk subsystem capability, which is called Soft Fence for configurations where the disks support this function. Soft Fence provides the means to *fence*, which means blocking access to a selected device. GDPS uses Soft Fence to fence devices that would otherwise be exposed to accidental update.

7.4 Managing the GDPS environment

You saw how the GDPS Virtual Appliance can protect your data during unplanned outages. However, as discussed in Chapter 1, "Introducing business resilience and the role of IBM GDPS" on page 1, most IBM Z outages are not disasters. Most are planned outages, with a small percentage of unplanned outages.

In this section, we describe other aspects of the GDPS Virtual Appliance, that is, its ability to monitor and manage the resources in its environment.

7.4.1 GDPS graphic user interface

The user interface that is used for managing the GDPS Virtual Appliance environment is known as the *GDPS graphic user interface* or *GDPS GUI*. Figure 7-2 shows the GDPS GUI home page.

Figure 7-2   GDPS GUI home page

As you can see, there are four distinct areas of the page:

1. Page Header

   Use this areas to start the following GDPS actions on demand:

   – Running a GDPS monitoring process. For more information about GDPS monitors, see 7.5, "GDPS monitoring and alerting" on page 232.

   – Temporarily disabling or reenabling HyperSwap.

2. Navigation menu

   This area contains icon links to the other panels that are available. Clicking an icon link opens a new tab in the Main Window and displays the corresponding panel in the new tab, including the following examples:

   – Dashboard: This panel is described in "Dashboard panel" on page 225.
   – Standard Actions: This panel is described in "Standard Actions panel" on page 225.
   – Planned Actions: This panel is described in "Planned Actions panel" on page 227.
   – Status Display Facility (SDF) Alerts: This panel is described in"SDF panel" on page 228.

3. Main Window

   This area is a tabbed workspace where the various GDPS panels are displayed. The Dashboard panel is displayed by default and is described in "Dashboard panel" on page 225. Other tabs are added to this area as extra panels are displayed. Inactive/hidden tabs can be brought to the foreground by clicking the associated tab.

4. Status Summary

   This area contains a graphical summary of the status and health of the GDPS managed environment, including the HyperSwap status, the disk mirroring status, the number of alerts of each severity that are displayed on the appliance, and the number of outstanding operator replies currently displayed on the appliance.

Dashboard panel

The Dashboard panel is the anchor content for the main window. This panel tab is always available to be made active. It shows at a glance the status of the components in your GDPS environment. Figure 7-2 on page 224 shows an example of the Dashboard panel. It includes icons that can be selected for the processors and disk in both Site1 and Site2. It also graphically shows the current direction and the status of Metro Mirror, plus the percentage of volume pairs that are in duplex state.

Clicking the arrow indicating the status and direction of the mirror opens the LSS Pairs panel. This panel is described in "LSS Pairs panel" on page 229. Clicking the Site1 or Site2 processor icon opens the Standard Actions panel. This panel is described next.

Standard Actions panel

GDPS provides facilities to help manage many common system-related actions. There are two reasons to use the GDPS facilities to perform these *Standard Actions*:

► They are tested and based on IBM best practices.

► Using the GDPS interface informs GDPS that the changes that it is seeing are planned changes, and therefore GDPS is not to react to these events.

Standard actions are performed by using the Standard Actions panel, which is shown in Figure 7-3. This panel is displayed by clicking one of the processor icons that are displayed on the Dashboard panel (as described in "Dashboard panel" on page 225), or by clicking the Standard Actions icon on the navigation menu (as described in 7.4.1, "GDPS graphic user interface" on page 224).
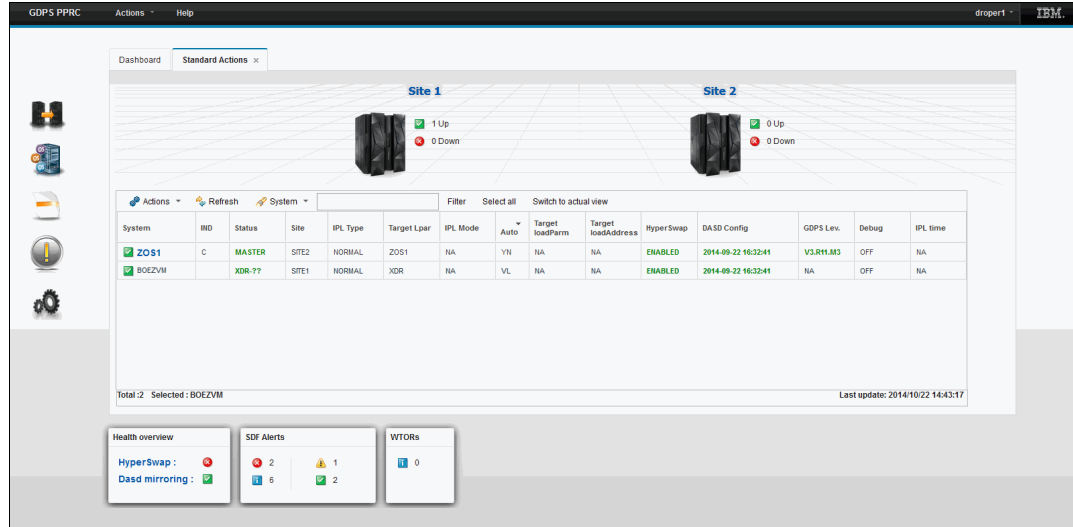

Figure 7-3   GDPS GUI Standard Actions panel

The panel displays a list of all systems that are defined to GDPS. The upper portion of the panel contains site icons with a summary count of the number of systems up and down in each Site.

Above the system list header is a toolbar that you can use to perform actions such as stopping, loading, and resetting systems and activating and deactivating LPARs.

If you double-click a z/VM system in the list that is presented, another panel opens in which you can operate at the *cluster* or Linux on IBM Z *guest* level within that z/VM image.

Planned Actions panel

The Planned Actions panel (see Figure 7-4) is displayed by clicking the Planned Actions icon on the Navigation menu, as described in 7.4.1, "GDPS graphic user interface" on page 224.
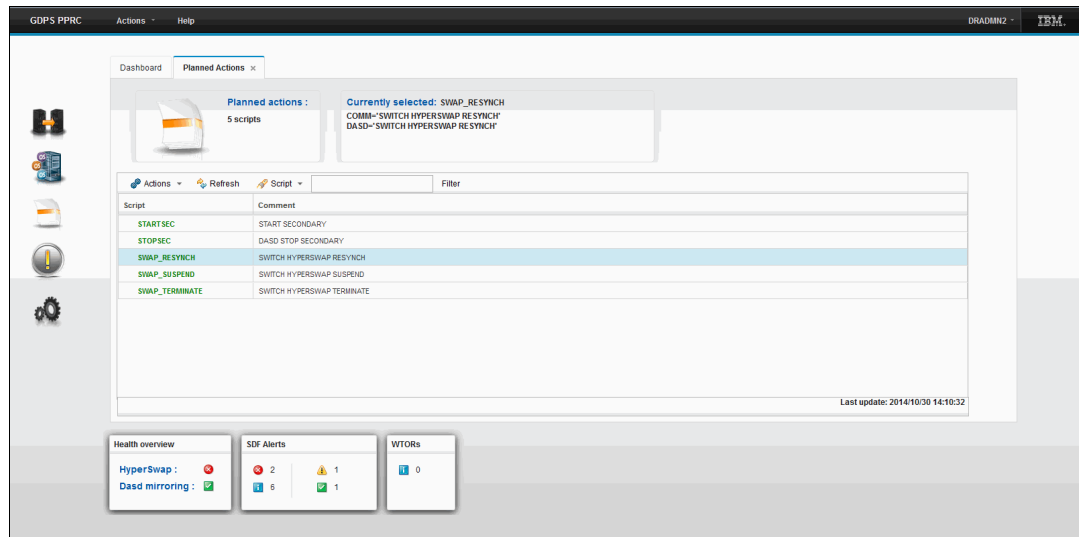

Figure 7-4   GDPS GUI Planned Actions panel

The panel displays a list of all *Control scripts* that were defined to GDPS. A Control script is a procedure that is recognized by GDPS that pulls together one or more GDPS functions. Control scripts perform complex, multi-step operations without having to run each step individually by using various panel options. For more information about Control scripts, see 7.4.2, "GDPS scripts" on page 230.

The upper portion of the Planned Actions panel contains a display box that contains the statements that are defined for any script that is selected. To run a script, double-click the script.

SDF panel

The SDF panel is the main panel for monitoring the status of GDPS managed resources. You can navigate to this panel by clicking the SDF alert icon that is displayed on the Navigation menu, as described in 7.4.1, "GDPS graphic user interface" on page 224. An example of the SDF panel is shown in Figure 7-5.
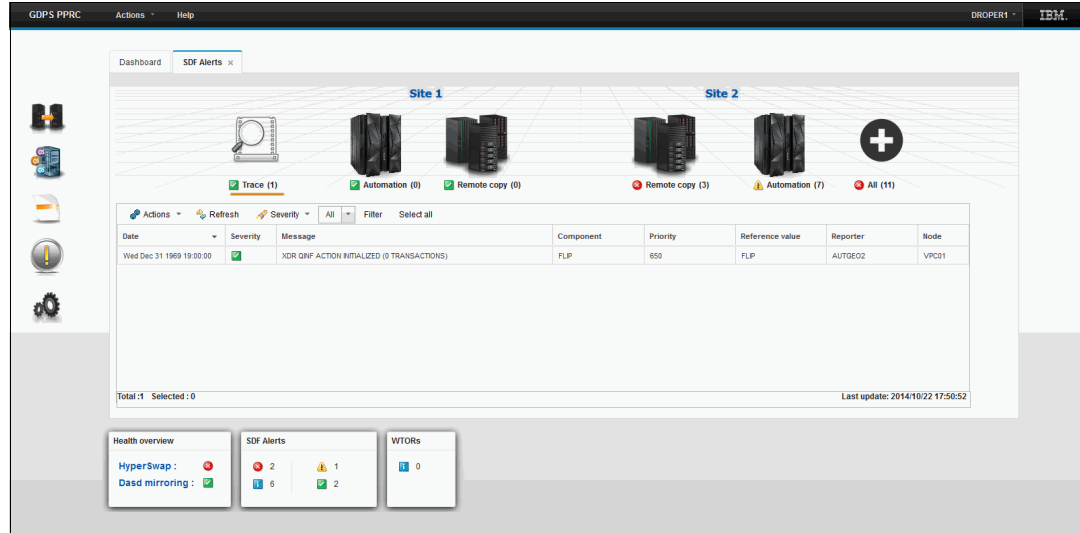

Figure 7-5   GDPS GUI SDF panel

The panel is divided horizontally into two sections. The upper section contains icons that can be clicked for filtering the SDF entry list that is displayed in the lower section based on the type of alert. The filtering icon labels indicate how many alerts of that type and location in parentheses.

Any SDF alerts that pass the applied filtering are displayed in the SDF entry list at the bottom of the panel.

Above the entry list header is a toolbar that you can use to delete alerts, display help associated with alerts, and others.

Remote Copy management panels

To manage the remote copy environment by using the GDPS Virtual Appliance, you first define your entire remote copy configuration, including your primary and secondary LSSs, your primary and secondary devices, and your PPRC links to GDPS in a file called the GEOPARM file. This enables GDPS to provide you with the capability to perform actions against all devices/pairs in your environment with a single action, rather than having to run an action against each device/pair.

This section describes the panel options that are provided by GDPS to manage your Remote Copy environment.

LSS Pairs panel

The initial panel for Remote Copy management is the LSS pairs panel. You navigate to this panel by clicking the mirroring status and direction arrow that is displayed on the Dashboard panel, as described in "Dashboard panel" on page 225. An example of the LSS Pairs panel is shown in Figure 7-6.
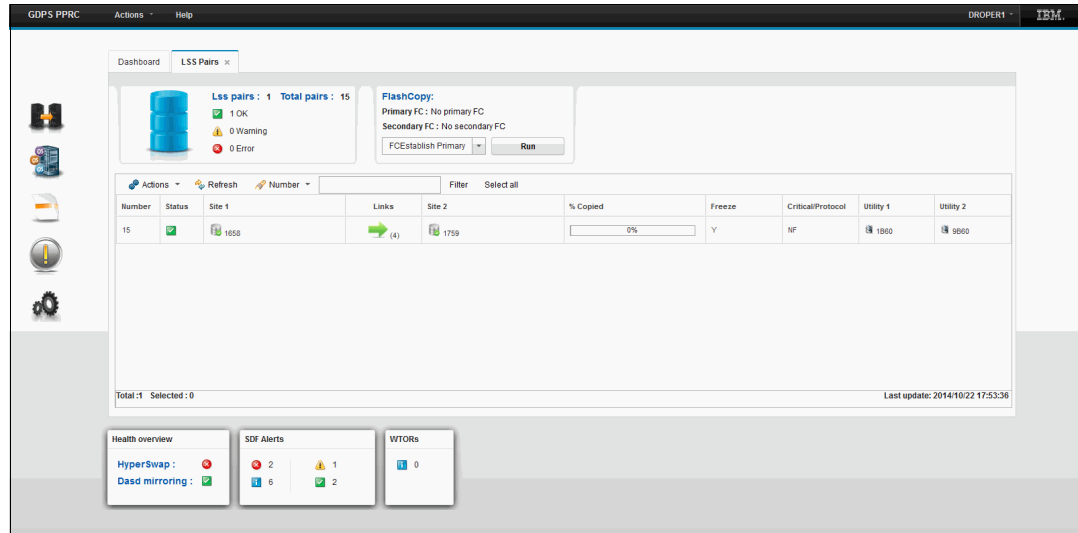

Figure 7-6   GDPS GUI LSS Pairs panel

The panel displays a list of all LSS pairs that are defined in the GDPS mirror. The upper left contains a summary count of the total number of LSS pairs and the number of LSS pairs by status severity. Double-clicking an LSS pair opens the Pairs panel for the LSS pair, as described in "Pairs panel" on page 230.

Above the LSS pair list header is a toolbar that you can use to perform various functions against all volume pairs in the selected LSS pairs. Examples of the functions you can perform by using the toolbar include querying the status of the pairs, suspending mirroring for the pairs, restarting mirroring for the pairs, and recovering the secondary devices for the pairs.

Pairs panel

Use the Pair panel to perform Remote Copy management at the volume pair level, rather than at the LSS pair level. An example of the Pairs panel is shown in Figure 7-7.
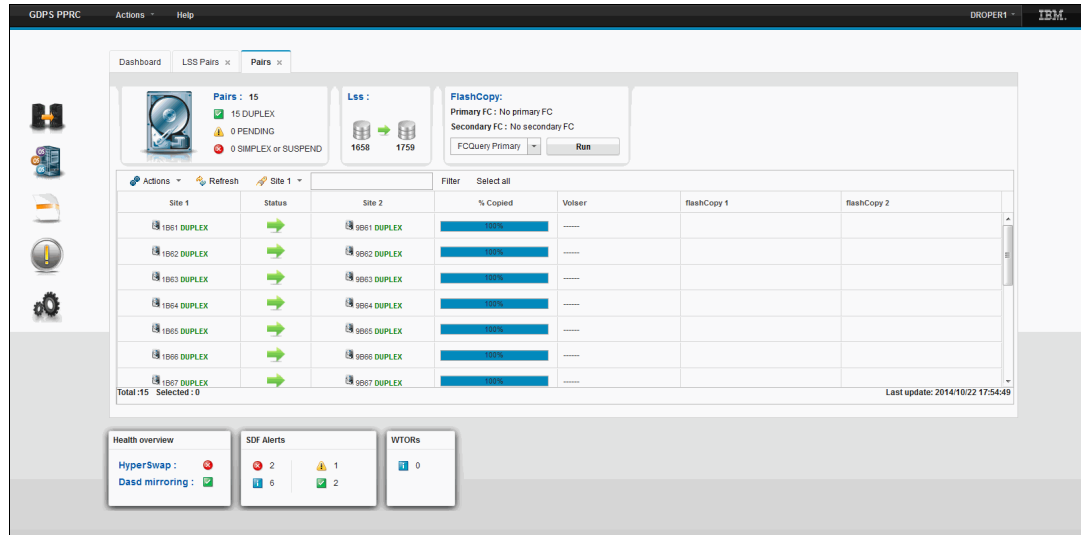

Figure 7-7   GDPS GUI Pairs panel

The panel displays a list of all volume pairs that are defined in the selected LSS. The upper left contains a summary count of the total number of volume pairs and the number of volume pairs by status severity. Double-clicking a volume pair issues a query for the pair and display the resulting output in a dialog box.

Above the volume pair list header is a toolbar that you can use to perform various functions against all selected volume pairs. Examples include querying the status of the pairs, suspending mirroring for the pairs, restarting mirroring for the pairs, and recovering the secondary devices for the pairs.

7.4.2 GDPS scripts

GDPS can automate complex, multi-step planned operations against your Remote Copy environment and against the production systems in your environment by using Control scripts.

Again, a script is a procedure that is recognized by GDPS that pulls together one or more GDPS functions. When running a script, GDPS performs the first statement in the list, checks the result, and only if it is successful, proceeds to the next statement. If you perform the same steps manually, you must check the results, which can be time-consuming, and then initiate the next action. With scripts, the process is automated.

Scripts are powerful because they can access the full capability of GDPS. The ability to start GDPS functions through a script provides the following benefits:

► Speed

The script runs the requested actions and check the results at machine speeds. Unlike a human, it does not need to search for the latest procedures or the commands manual.

► Consistency

With automation, your procedures run in the same way, time after time.

► Thoroughly tested procedures

Because they behave in a consistent manner, you can test your procedures over and over until you are sure that they do everything that you want, in exactly the manner that you want. Also, because you need to code everything and cannot assume a level of knowledge (as you might with instructions that are intended for a human), you are forced to thoroughly think out every aspect of the action the script is intended to undertake. And because of the repeatability and ease of use of the scripts, they lend themselves more easily to frequent testing than manual procedures.

7.4.3 System Management actions

Most of the GDPS Standard Actions require actions to be done on the HMC. The interface between the GDPS Virtual Appliance and the HMC is through a facility that is called the BCP Internal Interface (BCPii). GDPS uses BCPii to communicate directly with the hardware for automation of HMC actions such as LOAD, RESET, Activate LPAR, and Deactivate LPAR. GDPS can also perform ACTIVATE (power-on reset (POR)), CBU ACTIVATE/UNDO, and OOCoD ACTIVATE/UNDO.

Furthermore, when you LOAD a z/VM system by using GDPS (panels or scripts), GDPS listens for certain[2] operator prompts from the system being started and reply to the prompts. This support for replying to these IPL-time prompts automatically, helps to remove reliance on operator skills and eliminating operator error for any messages that require replies.

SYSRES Management

Today many clients maintain multiple alternative z/VM SYSRES devices (also known as *IPLSET*s) as part of their maintenance methodology. GDPS provides special support to allow clients to identify IPLSETs, which removes the requirement for clients to manage and maintain their own procedures when starting a system on a different alternative SYSRES device.

GDPS can automatically update the IPL pointers after any disk switch or disk recovery action that changes the GDPS primary site indicator for Metro Mirror disks, which removes the requirement for clients to perform extra script actions to switch IPL pointers after disk switches, and greatly simplifies operations for managing alternative SYSRES "sets."

---

[2] Only operator prompts that can be safely replied to in a consistent manner are candidates for automatic replies.

7.5 GDPS monitoring and alerting

As discussed in "SDF panel" on page 228, the GDPS SDF panel is where GDPS dynamically displays color-coded alerts.

Alerts can be posted as a result of an unsolicited error situation for which GDPS listens. For example, if one of the multiple PPRC links that provide the path over which Metro Mirror operations take place is broken, an unsolicited error message is issued.

GDPS listens for this condition and raises an alert on the SDF panel, which notifies the operator of the fact that a PPRC link is not operational. Clients run with multiple PPRC links and if one is broken, Metro Mirror continues over any remaining links.

However, it is important for operations to know that a link is broken and fix this situation because a reduced number of links results in reduced Metro Mirror bandwidth and reduced redundancy. If this problem is not fixed in a timely manner and more links fail, it can result in production impact because of insufficient mirroring bandwidth or total loss of Metro Mirror connectivity (which results in a freeze).

Alerts can also be posted as a result of GDPS periodically monitoring key resources and indicators that relate to the GDPS Virtual Appliance environment. If any of these monitoring items are found to be in a state that is deemed to be not normal by GDPS, an alert is posted that can be viewed by using the GDPS GUI on the appliance system.

When an alert is posted, the operator must investigate (or escalate) and corrective action must be taken for the reported problem as soon as possible. After the problem is corrected, this correction is detected during the next monitoring cycle and the alert is cleared by GDPS automatically.

The GDPS Virtual Appliance monitoring and alerting capability is intended to ensure that operations are notified of and can take corrective action for any problems in their environment that can affect the ability of the appliance to do recovery operations. This capability maximizes the chance of achieving your availability and RPO and RTO commitments.

7.6 Services component

GDPS touches on more than remote copy. It also includes automation, database management and recovery, testing processes, DR processes, and other areas.

Most installations do not have skills in all these areas readily available. It is rare to find a team that has this range of skills across many implementations. However, the GDPS Virtual Appliance offering includes exactly that: Access to a global team of specialists in all the disciplines that you need to ensure a successful GDPS Virtual Appliance implementation.

Specifically, the Services component includes several or all the following services:

► Planning to determine availability requirements, configuration recommendations, and implementation and testing plans

► Remote copy implementation

► GDPS Virtual Appliance installation and policy customization

► Assistance in defining RPOs and RTOs

► Education and training on the GDPS Virtual Appliance setup and operations

- ► Onsite implementation assistance
- ► Project management and support throughout the engagement

The sizing of the Services component of each project is tailored for that project, based on many factors, which include what automation is already in place, whether remote copy is already in place, the cross-site connectivity in place, and others. The skills provided are tailored to the specific needs of each particular implementation.

7.7 GDPS Virtual Appliance prerequisites

For more information about GDPS Virtual Appliance prerequisites, see this website.

7.8 Comparing GDPS Virtual Appliance to other GDPS offerings

So many features and functions are available in the various members of the GDPS family that recalling them all and remembering which offerings support them is sometimes difficult. To position the offerings, Table 7-1 lists the key features and functions and indicates which ones are delivered by the various GDPS offerings.

Table 7-1   Supported features matrix

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
CA	Yes	Yes	Yes	No
DR	Yes	Yes	Yes	Yes
CA/DR protection against multiple failures	Yes	No	No	No
CA for foreign z/OS systems	Yes with z/OS proxy	No	No	No
Supported distance	200 km 300 km (Business Recovery Services (BRS) configuration)	200 km 300 km (BRS configuration)	200 km 300 km (BRS configuration)	Virtually unlimited
Consistent FlashCopy support	Yes, using CONSISTENT	Yes, using CONSISTENT for secondary only	No	Yes, using CGPause
Reduced impact on initial copy/resync	Yes	Yes	Yes	Not applicable
Tape replication support	Yes	No	No	No
Production sysplex automation	Yes	No	Not applicable	No

Feature	GDPS Metro	GDPS HM	GDPS Virtual Appliance	GDPS GM
Span of control	Both sites	Both sites (disk only)	Both sites	Disk at both sites; recovery site (CBU or LPARs)
GDPS scripting	Yes	No	Yes	Yes
Monitoring, alerting and health checks	Yes	Yes	Yes (except health checks)	Yes
Query Services	Yes	Yes	No	Yes
MSS support for added scalability	Yes (RS2 in MSS1, RS3 in MSS2)	Yes (secondary in MSS1)	No	Yes (Global Mirror (GM) FlashCopy and Primary for Metro Global Mirror (MGM) in MSS1)
MGM 3-site and 4-site	Yes (all configurations)	Yes (3-site only and non-Incremental Resynchronization (IR) only)	No	Yes (all configurations)
Fixed-Block (FB) disk	Yes	Yes	No	Yes
z/OS equivalent function for Linux on IBM Z	Yes (Linux on IBM Z running as a z/VM guest only)	No	Yes (Linux on IBM Z running as a z/VM guest only)	Yes
GDPS GUI	Yes	Yes	Yes	Yes

7.9 Summary

The GDPS Virtual Appliance is a powerful offering that provides DR, CA, and system resource management capabilities for z/VM and Linux on IBM Z. GDPS Appliance is the only GDPS offering that is packaged in a virtual appliance, eliminating the necessity for z/OS and sysplex skills to manage and operate the solution.

HyperSwap, available with the GDPS Virtual Appliance, transparently can swap disks between two sites. The power of automation tests and perfects the actions to be taken, either for planned or unplanned changes, thus minimizing or eliminating the risk of human error.

The GDPS Virtual Appliance is one of the offerings in the GDPS family, along with GDPS HMHM and GDPS Metro, that offers the potential of ZDL that offers the potential of ZDL, and that can achieve the shortest RTO, typically less than 1 hour after a complete site failure.

It is also one of the only members of the GDPS family, again along with GDPS Metro, that is based on hardware replication and that provides the capability to manage the production LPARs. Although GDPS GM offers LPAR management, the scope of system management is limited, and includes only the systems in the recovery site, and not the production systems that are running in Site1.

In addition to the DR and planned reconfiguration capabilities, the GDPS Virtual Appliance also provides an interface for monitoring and managing the various elements of the GDPS configuration.

Combining local and metro continuous availability with out-of-region disaster recovery

In this chapter, we discuss the capabilities and considerations for implementing GDPS Metro Global - GM (GDPS MGM). It is of interest to clients who have requirements for both continuous availability (CA) local and regional disaster recovery (DR) protection.

GDPS MGM combines the CA attributes of GDPS Metro, or GDPS Metro HyperSwap Manager with the out-of-region DR capabilities of GDPS Global - GM to protect critical business data during a wide-scale disruption. They also provide for fast automated recovery under various, smaller-scale failure conditions.

> **Note:** GDPS Metro and GDPS Metro HyperSwap Manager can be combined with GDPS GM as described in this chapter. To aid in readability, only GDPS Metro is used in the text for most of the descriptions. If a particular function is not supported by GDPS Metro HyperSwap Manager, it is mentioned.

The following functions are provided by GDPS MGM:

► Three-copy disk mirroring using GDPS Metro to support zero data loss (ZDL) for day-to-day disruptions at metropolitan distances, and GDPS GM for long distance, out-of-region data protection, with limited data loss during a wide-scale disruption.

► Four-copy[1] disk mirroring combining GDPS Metro in the production region to support ZDL for day-to-day disruptions at metropolitan distances, GDPS GM between the two regions, and another instance of GDPS Metro in the recovery region to manage Global Copy (PPRC-XD) that can be switched to synchronous-mode while moving production to the recovery region in a planned or unplanned manner.

► Multisite management of the remote copy environment to maintain data integrity and data consistency across all disk copies.

---

[1] Incremental Resynchronization (IR) of GDPS GM and management of four copy configurations are not supported with GDPS Metro HyperSwap Manager.

237

- ► Support for transparent switching to secondary disks if there is a primary disk storage subsystem failure by using GDPS Metro with HyperSwap.

  This support offers the ability to incrementally resynchronize the GDPS GM mirror after a Metro Mirror HyperSwap.

- ► Fast, automated recovery for recovery time objective (RTO) of less than an hour for site and regional disasters.

- ► ZDL protection for both open systems and IBM Z by using GDPS Metro and GDPS GM, assuming that only one site is lost during a disaster.

- ► Use of FlashCopy to facilitate nondisruptive functions (such as backups, data mining, application testing, DR testing), and to provide a consistent copy of the data during remote copy synchronization to ensure that disaster readiness is maintained always.

- ► Planned switch to run production in the recovery region and then return home.

This chapter includes the following topics:

- ► 8.1, "Introduction" on page 238
- ► 8.2, "Design considerations" on page 239
- ► 8.3, "GDPS Metro Global - GM 3-site solution" on page 243
- ► 8.4, "GDPS Metro Global - GM 4-site solution" on page 252

8.1 Introduction

Enterprises running highly critical applications have an increasing need to improve the overall resilience of their business services and functions. Enterprises already doing synchronous replication are accustomed to the availability benefits of relatively short distance synchronous replication, which is especially true in mainframe environments where the capabilities of HyperSwap handle disk subsystem failures without an outage and to use server capacity in both sites.

Regulatory bodies (both governmental and industry-based) in various countries are requiring enterprises to maintain a significant distance between their primary and disaster locations to protect against wide-scale disruptions. For some organizations, these regulations can result in a requirement to establish backup facilities well outside the range of synchronous replication capabilities, thus driving the need to implement asynchronous disk mirroring solutions.

From a business perspective, this setup might mean compromising CA to comply with regulatory requirements. With a three-copy disk mirroring solution, the availability benefits of synchronous replication can be combined with the distance that is allowed by asynchronous replication to meet both the availability expectations of the business and the requirements of the regulator. Further extension to four-copy configurations allows for equivalent high availability (HA) characteristics when running in either region.

8.2 Design considerations

In the following sections, we describe design considerations that include three-copy solutions versus 3-site solutions, multi-target and cascading topologies, four-copy solutions, and cost considerations.

8.2.1 Three-copy solutions versus 3-site solutions

It is not always the case that clients implementing a three-copy mirroring solution have three independent data centers (as shown in Figure 8-1), each with the capability to run production workloads.
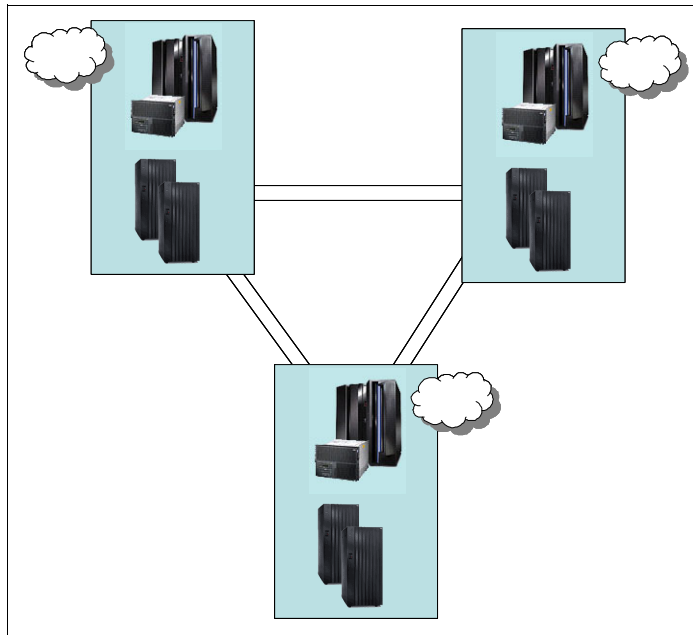

Figure 8-1   Three-site solution

Having three distinct locations with both the connectivity that is required for the replication and connectivity for user access is expensive and might not provide sufficient cost justification. As the distance between the locations connected with synchronous mirroring increases, the ability to provide CA features such as cross-site disk access, HyperSwap, or coupling facility (CF) duplexing diminishes.

Having a production location with two copies of data within a single data center (shown in Figure 8-2), along with a third copy of the data at a remote recovery location, provides you with many of the benefits of a full 3-site solution while allowing for a reduced overall cost. Disk subsystem failures are handled as local failures and if the single site has some degree of internal resilience, then even minor "disaster-type" events can perhaps be handled within the single location.
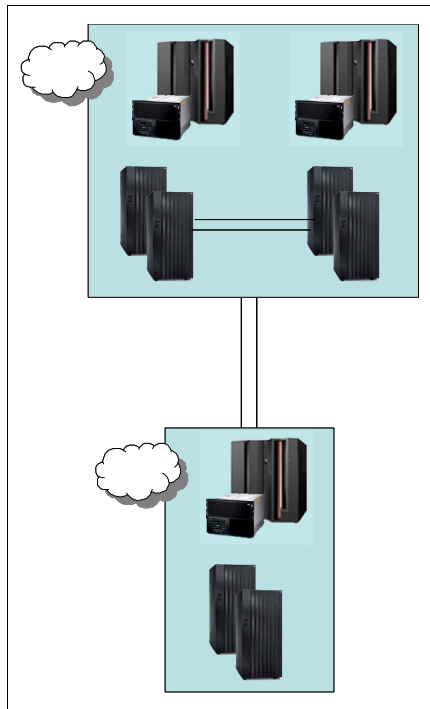

Figure 8-2   A 2-site solution

Another benefit of the two-data center solution, especially in an IBM Z environment, is that you can realize the full benefit of features such as HyperSwap and CF duplexing to provide CA features without provisioning significant additional and expensive cross-site connectivity, or having concerns regarding the impact of extended distance on production workloads.

Figure 8-3 on page 241 shows another variation of this scenario, in which the primary data center is a campus location with separate machine rooms or buildings, each with the ability to run production workloads.
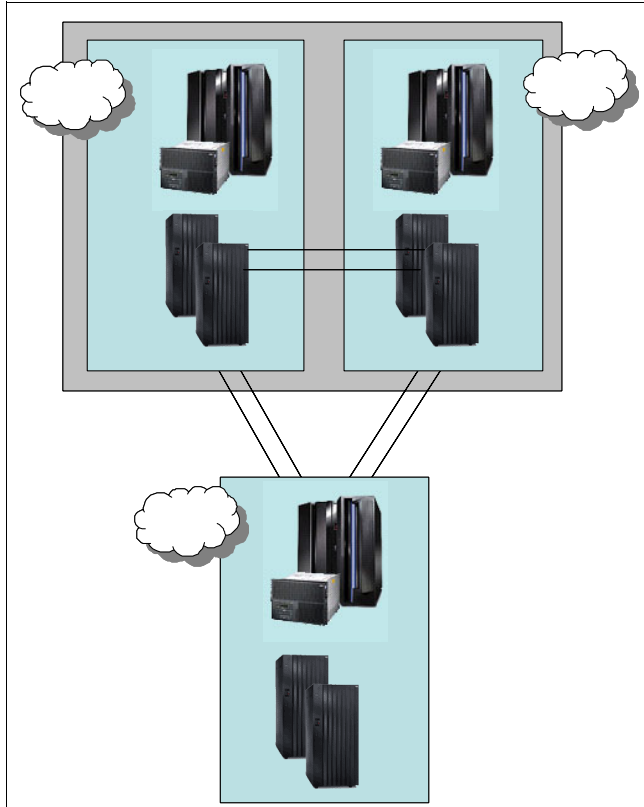
Figure 8-3   A 2-site solution: Campus and Recovery site

In the past, clients often used the bunker topology (as shown in Figure 8-4) to create a solution that might provide mirroring at extended distances, but still handle a primary site failure without data loss.
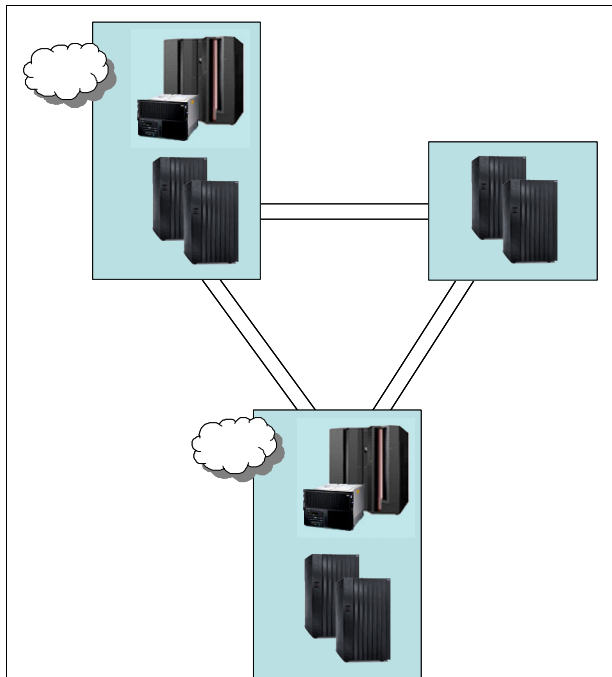

Figure 8-4   Two sites and an intermediate bunker

There are several arguments against this approach:

► For ZDL, you need a policy in which, if the mirroring stops, the production applications are also stopped. There are clients who implement such a policy, but it is not a common policy. If production is allowed to continue after a local mirroring failure, then ZDL cannot be guaranteed in all situations.

► If the disaster event also affects the bunker site or affects the bunker site first, then ZDL is again not guaranteed. If the reason for the extended distance to the recovery site was to handle regional events, then this possibility cannot be excluded.

► The networking and hardware costs of the bunker site are probably still considerable despite there being no servers present. Further investment in the availability characteristics of the primary location or in a campus-type solution in which the synchronous secondary disk subsystems can be used for production services might provide a greater return on investment for the business.

8.2.2 Multi-target and cascading topologies

Multi-target and cascading topologies are similar in terms of capabilities in that both provide a synchronous and an asynchronous copy of the production data. Certain failure scenarios are handled by multi-target solutions and other scenarios by cascading solutions.

The key requirements for either topology are as follows:

► A viable recovery copy (RC) and recovery capability is available always in a location other than where production is running. It is possible that regulatory requirements demand this capability. This requirement implies that no situations exist in which both offsite copies are compromised.

► Any single site failure results in only at most a short outage of the replication capability between the surviving sites to ensure minimal exposure where there might be increased data loss for a second failure.

With this requirement, being able to do IR between any two copies of the data is desirable. The absence of this requirement can result in an extended period of exposure to more data loss in a second failure.

8.2.3 Four-copy solutions

Many parallels can be drawn between 4-site solutions and the 3-site solutions that are presented in the previous section. Clients are unlikely to have four discrete physical data centers because of the cost implications. In all probability, the four-copy solution is most likely to be implemented in two physical locations where each location has two "hardened" data center facilities in the one location. For example, in two adjacent buildings on a campus, or even two separate data center halls within a single building where the halls are separated by fire-resistant barriers and are independently provided with power.

8.2.4 Cost considerations

The third and potentially fourth locations are insurance that meets regulatory compliance. This setup might imply that costs for this location are kept to an absolute minimum.

Reducing the network bandwidth to remote locations can provide significant cost savings for the overall cost of the solution. Given that a synchronous copy is already available 'locally', trading off the recovery point objective (RPO) versus the cost of the network might be a useful compromise especially if the times of increased RPO are during periods of batch processing or database maintenance where the transactional data loss would be smaller.

Using a DR service provider such as IBM BCRS is one method of reducing the costs of the third location, fourth location, or both. Shared hardware assets and the removal of the requirement to invest in extra physical locations can provide significant cost benefits, and with most events that are expected to be handled in the two main locations, the disadvantages of a shared facility are reduced.

8.2.5 Operational considerations

When running in multiple locations and combining different techniques together to provide an overall solution, there can be the requirement to do synchronized actions in both regions. To facilitate this requirement from an operational standpoint, IBM Geographically Dispersed Parallel Sysplex (GDPS) provides a Remote Script Execution function so that from a single point of control you are able to initiate actions in any of the individual GDPS environments that make up the overall solution.

8.3 GDPS Metro Global - GM 3-site solution

This section describes the capabilities and requirements of the GDPS Metro Global - GM 3-site (GDPS MGM 3-site) solution. This solution combines the capabilities of GDPS Metro and GDPS GM. Synchronous replication between a primary and a secondary disk subsystem that are within a single data center, or between two data centers within metropolitan distances, is implemented with GDPS Metro in a single-leg configuration. GDPS GM is used to asynchronously replicate data to a third disk subsystem in a recovery site that is typically out of the local metropolitan region.

GDPS MGM is configured to use the Multi-Target Metro Mirror (MTMM) technology to dynamically switch between a cascaded topology and a multi-target topology as necessary to optimize recovery scenarios such as HyperSwap. This configuration is referred to as a *multi-target GDPS MGM 3-Site* configuration.

8.3.1 GDPS MGM 3-site overview

The GDPS MGM 3-site configuration that is shown in Figure 8-5 is a 3-site CA and DR solution. In this example, Site1 and Site2 are running a multi-site workload configuration (for more information, see 3.2.3, "Multisite workload configuration" on page 64) and are within metropolitan distances to ensure optimal application performance. All data that is required to recover critical workloads is on disk and is mirrored. Each site is configured with sufficient spare capacity to handle failed-over workloads during a site outage.

The third site, or recovery site, can be at virtually unlimited distance from Site1 and Site2 to protect against regional disasters. Asynchronous replication is running between Site2 and the recovery site. Redundant network connectivity is installed between Site1 and the recovery site to provide continued DR protection during a Site2 disaster or a failure of the disk subsystems in Site2. For more information, see "Incremental resynchronization for a GDPS MGM 3-site configuration" on page 245.

There is sufficient CPU capacity that is installed to support the R-sys. Capacity BackUp (CBU) is installed and GDPS starts CBU on IBM Z to provide the extra capacity that is needed to support production workloads if DR is started.
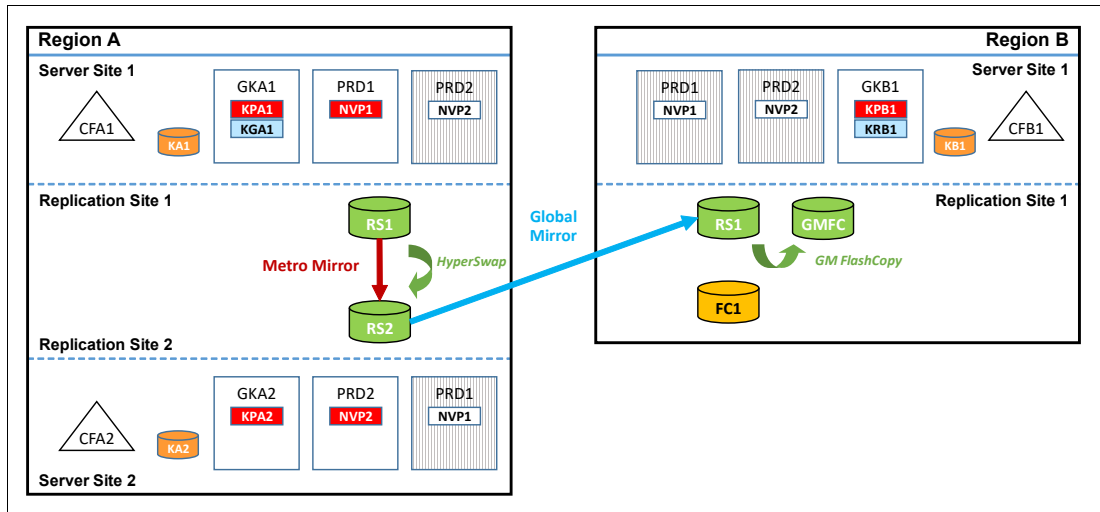

Figure 8-5   GDPS MGM 3-Site cascaded configuration

The RS1 disks in Region A, Site1are synchronously mirrored to the RS2 disks in Region A, Site2 by using Metro Mirror. Then, the RS2 disks in Region A, Site2 are asynchronously mirrored to a third set of disks, RS1, in the recovery region (Region B) by using GM.

A fourth set of disks (GMFC), also in the recovery region, are the FlashCopy targets that are used to provide the consistent data ("journal") for DR. A fifth (FC1) and optional set of disks are used for stand-alone DR testing or, if a real disaster occurs, to create a "golden" or insurance copy of the data. For more information about GM, see Chapter 5, "IBM GDPS Global - GM" on page 147.

Because some distance is likely to exist between the local sites, Site1 and Site2, running the Metro Mirror leg of Metro Global Mirror (MGM), and the remote recovery site that is the GM recovery site, we also distinguish between the local sites and the remote site by using *region* terminology. Site1 and Site2 are in one region, Region A, and the remote recovery site is in another region, Region B.

Incremental resynchronization for a GDPS MGM 3-site configuration

The IR function of MGM enables IR between the surviving Region A replication site (RS) and the recovery site when the disk in the RS in Region A that is hosting the GM session becomes unavailable.

Without this capability, if the disk in the RS in Region A that is hosting the GM session becomes unavailable, the data at the recovery site begins to age because data can no longer be replicated between Region A and Region B. Instead of requiring a new GM session from the production site to the recovery site (and a full copy), the IR capability of GDPS MGM 3-site allows the RC of data in Region B to be resynchronized from the surviving copy in Region A by copying only the changes that occurred since the error event took place.

Figure 8-6 shows how GDPS MGM 3-site can establish a GM session between the production site (RS1 in Region A), and the recovery site (RS1 in Region B) when it detects that the intermediate site (RS2 in Region A), which is hosting the GM session, becomes unavailable.
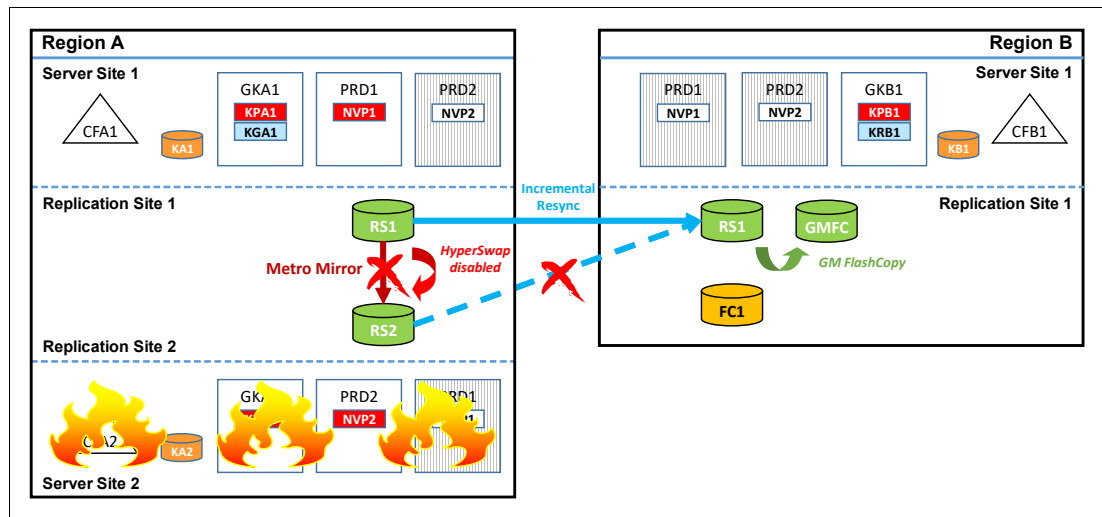

Figure 8-6   GDPS MGM 3-Site cascaded configuration after a Site2 outage

After the GM session is established, only an IR of the changed data needs to be performed, which allows the DR capability to be restored in minutes, instead of hours.

Figure 8-7 shows a GDPS Metro Global - GM 3-site configuration when it is in a multi-target topology. GDPS Metro Global - GM 3-site configurations can dynamically switch between a cascaded topology and a multi-target topology to optimize processing of various recovery scenarios.
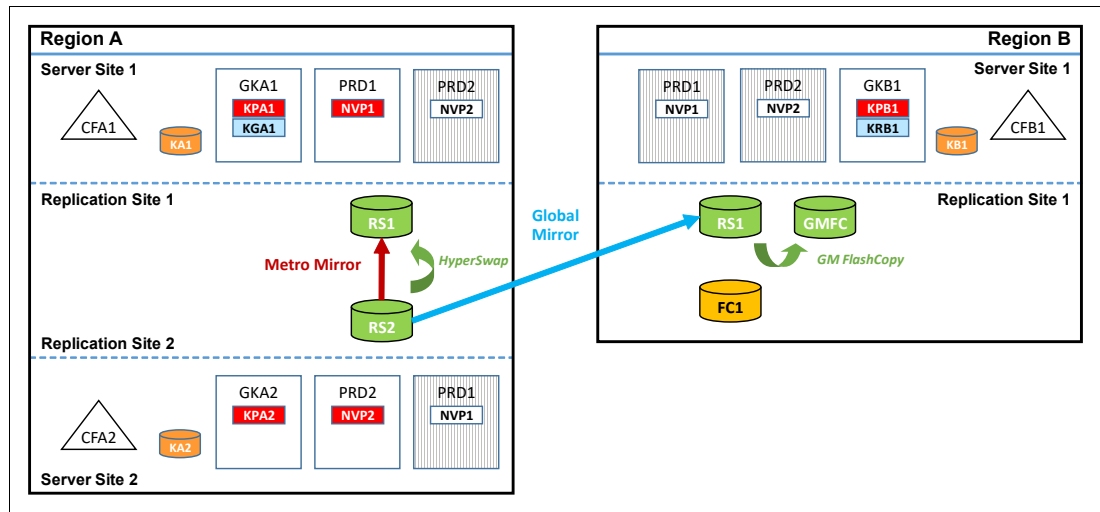

Figure 8-7   GDPS MGM 3-site multi-target configuration

Assume that your GDPS Metro Global - GM 3-Site configuration started out in a cascaded topology, as shown in Figure 8-5 on page 244. If you run a planned HyperSwap to the RS2 disk, followed by a reverse resynchronization of Metro Mirror from the RS2 disk back to the RS1 disk in Region A, the multi-target topology that is shown in Figure 8-7 results.

As shown in Figure 8-7, the RS2 disk is now the primary copy of data that application systems are accessing and the RS1 disk in Region A is the Metro Mirror secondary disk to the RS1 disk. HyperSwap was reenabled to provide HA for the Region A data. This synchronous relationship is managed by GDPS Metro in Region A.

The RS2 disk also remains the GM primary disk that is copied to the RS1 disk in Region B (the GM secondary disk). This asynchronous relationship is managed by using GDPS GM. IR is still enabled from the RS1 disk in Region A to the RS1 disk in Region B to protect from a failure of the RS2 disk and allow the GM copy to be re-established without the need for a full copy if the RS2 disk becomes unavailable.

The advantage of the multi-target capability in this scenario is that, following the HyperSwap, GM from the RS2 disk to the RS1 disk in Region B can remain active and maintain your DR position, while Metro Mirror in Region A is being resynchronized from the RS2 disk back to the RS1 disk.

GDPS MGM Procedure Handler

The GDPS MGM Procedure Handler is a fully integrated component of GDPS for use in 3-site IR or 4-site configurations. The Procedure Handler, along with the provided procedures, can be used to drive several complex scenarios with a single script invocation, as shown in the following examples:

► To incrementally reintroduce the RS2 intermediate disk if GM was incrementally resynchronized from Site1 in Region A to the recovery site.

   A supplied procedure returns to a cascaded RS1-RS2-RS1 configuration when GM processing is moved to RS1 in Region A because of a loss of RS2. Without the procedure, returning to the cascaded configuration requires a full initial copy for both Metro Mirror (RS1 disk to RS2 disk in Region A) and for GM (RS2 disk to RS1 disk in Region B). Therefore, the procedure provides significant availability and DR benefit for IR environments.

   The procedure can be used for this purpose only if the RS2 disk is returned "intact," meaning that metadata on the disk subsystem pertaining to its status as a Metro Mirror secondary and GM primary disk is still available. If you need to introduce a new disk subsystem into the configuration, this process requires a full initial copy of all the data.

► To perform a planned toggle between the RS1 disk and the RS2 disk in Region A.

   This procedure performs periodic "flip/flops" of Site1 and Site2 (or RS1disk and RS2 disk) in Region A. This procedure HyperSwaps the host systems, resynchronizes Metro Mirror, and moves GM to the new Metro Mirror secondary for cascaded-only configurations.

► To incrementally "return home" after recovering production on the RS1 disk in Region B (or after you perform a planned region switch to Region B) by reintroducing both the RS1 disk and the RS2 disk in Region A.

   This capability is possible only if both sets of disks in Region A (RS1 and RS2) are returned intact. Although the MGM mirror can be incrementally reinstated, a production outage is necessary to move production from running in Region B back to running in Region A on either the RS1 or the RS2 disk.

The Procedure Handler supports only CKD disks. IR is not supported by GDPS Metro HyperSwap Manager.

8.3.2 GDPS MGM Site1 failures

The primary role of GDPS Metro is to protect the integrity of the RS2 copy of the data. At the first indication of a failure in Site1, GDPS Metro freezes all RS2 disks to prevent logical contamination of data that is on the RS2 devices. For more information about GDPS Metro processing, see Chapter 3, "IBM GDPS Metro" on page 45.

Now, the GDPS GM session between Site2 and the recovery site is still running, and both locations most likely will have the same set of data after a brief amount of time. The business focus is now on restarting the production systems in either Site2 or the recovery site, depending on the failure scenario. If the systems are started in Site2, the GDPS GM solution is already in place.

8.3.3 GDPS MGM Site2 failures

In this situation, the production systems are still running, so the business requirement is to ensure that DR capabilities are restored as fast as possible. The GDPS GM session should be restarted as soon as possible between Site1 and the recovery site by using IR. For more information, see "Incremental resynchronization for a GDPS MGM 3-site configuration" on page 245. If IR is not configured, a full copy is required.

This scenario has possibly less impact to the business than a failure of the production site, but it depends on the specific environment.

8.3.4 GDPS MGM region switch and return home

It is possible to switch production from running in Region A (in either Site1 or Site2) to Region B. Many GDPS MGM 3-site customers run Site1 and Site2 in the same physical site or on a campus where these two sites are separated by little distance. In such configurations, there might be planned outage events, such as complete power maintenance, that are likely to affect both sites.

Similarly, an unplanned event that impacts both sites forces recovery in Region B.

While production runs in Region B, the disk subsystems in this region track the updates that are made. When Region A is available again, assuming that all disks configured in the region come back intact, it is possible to return production back to Region A using the appropriate supplied procedure without requiring fully copying the data back. Because the updates were tracked, only the data that changed while Region A was down is sent back to the Region A disks to bring them up to date. Then production is shut down in Region B. The final updates are drained to Region A and production can then be restarted in Region A.

Because Region A and Region B are not symmetrically configured, the capabilities and levels of protection that are offered when production runs in Region B is different. Because there is only one copy of the production data in Region B, there is no HyperSwap protection to provide continuous data access. For the same reason, the various operational procedures for GDPS are different when running in Region B. However, even if no outage is planned for Region A, switching production to Region B periodically (for example, once or twice a year) and running live production there for a brief period is the best form of disaster testing because it provides the best indication of whether Region B is properly configured to sustain real, live production workloads.

8.3.5 Scalability in a GDPS MGM 3-site environment

As described in "Addressing z/OS device limits in a GDPS Metro environment" on page 25, GDPS Metro allows defining the Metro Mirror secondary devices in alternative subchannel set 1 (MSS1), which allows up to nearly 64 K devices to be mirrored in a GDPS Metro configuration. The definitions of these devices are in the application site I/O definitions.

Similarly, "Addressing z/OS device limits in a GDPS GM environment" on page 28 describes how GM allows defining the GM FlashCopy target devices in alternative MSS1 (or not defining the GM FlashCopy target devices at all) in the recovery site I/O definitions. It also describes not defining the practice FlashCopy target devices at all to the GDPS GM R-sys, again, allowing up to nearly 64 K devices to be mirrored in a GDPS GM configuration.

In a GDPS MGM 3-site environment where the Metro Mirror secondary devices that are defined in MSS1 are the GM primary devices, there is more support in GDPS GM that allows the GM primary devices to be defined in MSS1. With the combined alternative subchannel set support in GDPS Metro and GDPS GM, up to nearly 64 K devices can be replicated using the MGM technology.

8.3.6 Other considerations in a GDPS MGM 3-site environment

With GM, deliberately underconfiguring the bandwidth that is provided to reduce the total cost of the solution is possible. If significant peaks exist, then this cost savings might be considerable because the network costs are often a significant portion of ongoing costs. The drawback with under-configuring bandwidth is that it might affect the recovery point that can be achieved. If a disaster affects the entire production region, both Site1 and Site2, during any peak when the GM mirror is running behind, there is likely to be more data loss.

8.3.7 Managing the GDPS MGM 3-site environment

GDPS provides a range of solutions for DR and CA in an IBM Z centric environment. GDPS MGM 3-site provides support for MGM within a GDPS environment. GDPS builds on facilities that are provided by System Automation and NetView and uses inband connectivity to manage the MGM relationships.

GDPS MGM 3-site runs two services to manage MGM, both of which run on z/OS systems. The GDPS Metro services run on every z/OS image in the production sysplex and the controlling systems, K1 and K2, in Site1 and Site2. Each controlling system is allocated on its own non-mirrored disk and has access to the primary and secondary disk subsystems.

During normal operations, the master function runs in the controlling system that is located where the secondary disks are. The controlling system is where the day-to-day management and recovery of the Metro Mirror environment is performed. If Site1 or Site2 fails, the Master system manages the recovery of the Metro Mirror disks and production systems.

The second controlling system is an alternative that takes over the master function if the Master controlling system becomes unavailable or a Master switch occurs as a result of, for example, a HyperSwap.

The GDPS GM services run in the Kg and R-sys controlling systems. Kg runs in the production sysplex and is responsible for controlling the GM environment and sending information to the R-sys running in the recovery site. The R-sys is responsible for carrying out all recovery actions during a wide-scale disruption that affect both Site1 and Site2.

In addition to managing the operational aspects of GM, GDPS GM provides facilities to restart IBM Z production systems in the recovery site. By providing scripting facilities, it provides a complete solution for the restart of an IBM Z environment in a disaster situation without requiring expert manual intervention to manage the recovery process.

8.3.8 GDPS MGM 3-site support for multiple IBM Z platforms

As described in 3.3.2, "Multiplatform Resiliency for IBM Z" on page 70, GDPS Metro provides near CA for multiple IBM Z platforms (or operating systems), beyond z/OS, and their disk. Of these extra platforms, z/VM and its guests, including Linux on IBM Z, and IBM Db2 Analytics Accelerator on IBM Z running in Secure Service Containers (SSCs), are also supported by GDPS MGM in 3-site environments. That is, the disk for these systems can be mirrored to the recovery region by using GM and these systems can be brought up in the recovery region for planned and unplanned region switch scenarios and for DR testing.

8.3.9 Flexible testing in a GDPS MGM 3-site environment

To facilitate testing of site failover and failback processing, consider installing extra disk capacity to support FlashCopy in Site1 and Site2. The FlashCopy can be used at both Site1 and Site2 to maintain DR checkpoints during remote copy resynchronization, which ensures that there is a consistent copy of the data available if a disaster-type event occurs while testing your site failover and failback procedures. In addition, the FlashCopy can be used to provide a copy to be used for testing or backing up data without the need for extended outages to production systems.

GDPS MGM 3-site supports another FlashCopy disk device, which is referred to as *F disks* or *FC1 disks*. These disks are other FlashCopy target devices that might be created in the recovery site. The F disks might be used to facilitate stand-alone testing of your DR procedures while the GM environment is running, which ensures that a consistent and current copy of the data is available always. In addition, the F disk can be used to create a "gold" or insurance copy of the data if a disaster situation occurs.

Currently, GDPS MGM 3-site supports the definition and management of a single F device for each MGM *triplet* (RS1, RS2, RS1 disk combinations) in the configuration. To reduce management and operational complexity, support exists in GDPS GM to support the F disk without adding a requirement for these disks to be defined to the I/O configurations of the GDPS systems managing them. Known as "No unit control block (UCB)" FlashCopy, this support allows for the definition of F disks without the need to define extra UCBs to the GDPS management systems.

In addition to the ability to test on the F disks, GDPS MGM 3-site configurations support testing by using X-disk support in GDPS GM, as described in 5.9.2, "Creating a test copy by using GM CGPause and testing on isolated disks" on page 174.

8.3.10 GDPS Query Services in a GDPS MGM 3-site environment

Use GDPS Metro provides Query Services to query various aspects of the Metro Mirror leg of a GDPS MGM 3-site environment. Similarly, use GDPS GM provides Query Services to query various aspects of the GM leg of a GDPS MGM 3-site environment.

The GDPS GM query services also have awareness of the fact that a particular environment is a GDPS MGM 3-site environment that is enabled for IR and returns additional information pertaining to the IR aspects of the environment. In a GM environment, the GM session can be running from Site2 to the recovery site (RS2 to RS1) or from Site1 to the recovery site (RS1 to RS1) at any time.

If GM is running RS2 to RS1, it is the Active GM relationship and the RS1 to RS1 relationship is the Standby GM relationship. The GM query services in an MGM 3-site IR environment return information about both the active and the standby relationships for the physical and logical control units in the configuration and the devices in the configuration.

8.3.11 Easy Tier Heat Map Transfer in a GDPS MGM 3-site environment

As described in 3.7.5, "Easy Tier Heat Map Transfer" on page 104, GDPS Metro manages the transfer of Easy Tier learning within the Metro Mirror environment through a capability that is called Heat Map Transfer. Similarly, as described in 5.6.3, "Easy Tier Heat Map Transfer" on page 172, GDPS GM manages the transfer of Easy Tier learning within the GM environment.

In a GDPS MGM 3-site environment, the Heat Map Transfer functions of GDPS Metro and GDPS GM work together to manage the transfer of Easy Tier learning across the entire environment. When significant events occur, such as planned or unplanned HyperSwaps, planned or unplanned site switch or region switch events, or planned or unplanned removal or loss of the current GM primary disk set, the flow of data across the replication environment is affected. In these cases, GDPS automatically reconfigures the flow of Easy Tier learning information to match the new flow of data across the environment. This ability allows the ongoing Easy Tier learning process taking place in each disk subsystem in the environment to remain current and relevant so that performance effects that are related to subsequent events are minimized.

8.3.12 Prerequisites for a GDPS MGM 3-site configuration

GDPS MGM 3-site has the following prerequisites:

► GDPS Metro or GDPS HM. If GDPS HM is used, the IR function is not available.

► GDPS GM is required and the GDPS GM prerequisites must be met.

► Consult with your storage vendor to ensure that the required features and functions are supported on your disk subsystems.

> **Important:** For the latest GDPS prerequisite information, see the GDPS product website.

8.3.13 GDPS MGM 3-site integration with GDPS Continuous Availability

For GDPS MGM 3-site integration with GDPS Continuous Availability, there are two options to consider.

Disk Replication Integration

It is possible to use cascaded-only GDPS MGM 3-Site to provide local HA and remote disaster recover for the sysplexes that run GDPS Continuous Availability workloads. For such an environment, GDPS Continuous Availability provides special facilities to manage the disk replication aspects of specific planned and unplanned scenarios using the GDPS Continuous Availability Controller system as the single point of control. For more information, see 6.6, "Zero Data Loss configuration" on page 207 for details.

This support is not provided with GDPS MGM Multi-Target 3-Site.

Zero Data Loss at distance

GDPS MGM 3-Site can be integrated with GDPS Continuous Availability to provide a ZDL capability for your most critical workloads, even when your data centers are separated by distances that exceed the capability of synchronous mirroring. For more information, see 6.6, "Zero Data Loss configuration" on page 207.

8.4 GDPS Metro Global - GM 4-site solution

The GDPS MGM 4-site solution is an extension of the GDPS MGM 3-Site solution, which is described in 8.3, "GDPS Metro Global - GM 3-site solution" on page 243. As such, the features and functions that are available with the 3-site solution also are available with the 4-site solution, with a few exceptions that are described later in this section.

The critical difference between the 3-site solution and the 4-site solution is that, with the 4-site solution, a second copy of data is available in the recovery region that can provide a HA copy if you perform a planned or unplanned switch of production to the recovery region. the 4-site configuration is a *symmetrical configuration* because from a data HA perspective, the same capabilities are available whether you are running your production services in Region A or Region B.

This fourth copy of data is created by using asynchronous Global Copy (also known as PPRC-XD) that can be switched to synchronous-mode (that is, Metro Mirror) during a planned or unplanned region switch, which provides the HA copy in that region.

Figure 8-8 shows an MGM 4-site configuration in a cascaded topology that consists of the four copies of data, labeled RS1 and RS2 in Region A and RS1 and RS2 in Region B. The GM FlashCopy target device (or "journal device") is shown in Figure 8-8 as GMFC.
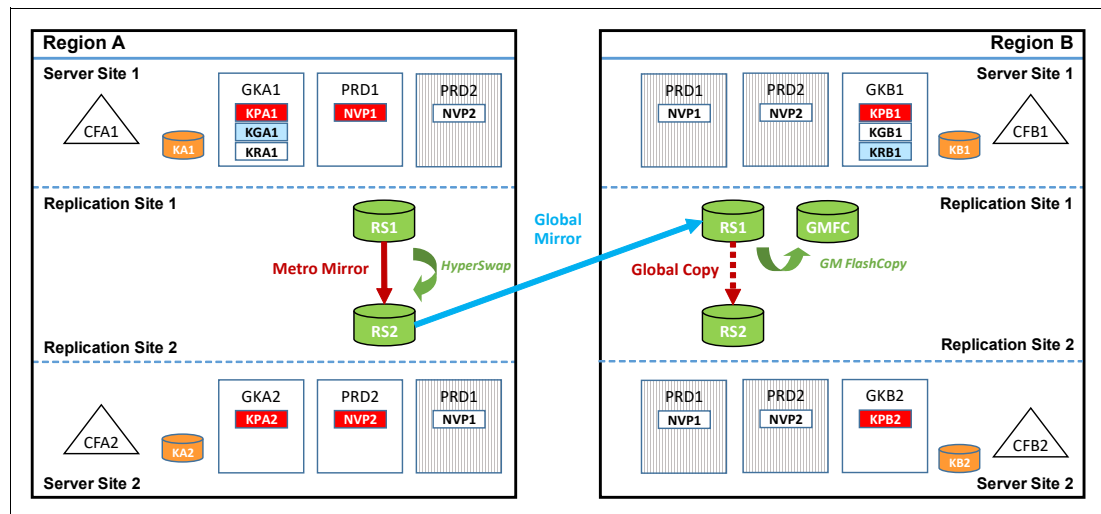

Figure 8-8   GDPS MGM 4-site cascaded configuration

In Figure 8-8, which shows a *steady state* when running in Region A, the RS1 disk in Region A is the primary copy of data that application systems is accessing. The RS2 disk in Region A is the Metro Mirror secondary disk to the RS1 disk, and HyperSwap is enabled to provide HA for the Region A data. This relationship is managed by GDPS Metro running in Region A.

The RS2 disk in Region A is also the GM primary disk, being copied to the RS1 disk in Region B, which is the GM secondary disk. This configuration is managed by using GDPS GM. IR is also enabled from the RS1 disk in Region A to the RS1 disk in Region B to protect from a failure of the RS2 disk in Region A and allow the GM copy to be re-established without the need for a full copy. This configuration is the same as a 3-site configuration. Where it differs is that the RS2 disk in Region B is present and is a Global Copy secondary to the RS1disk in Region B. This relationship, which is managed by GDPS Metro running in Region B, can be converted to fully synchronous Metro Mirror when you perform a switch of production to Region B for whatever reason.

If you switch production to Region B, you use the RS1 disk in Region B as the primary copy, with the RS2 disk in Region B being the Metro Mirror secondary, and the RS1 disk in Region A becomes the GM secondary. Then, the RS2 disk in Region A is the Global Copy secondary disk to the RS1 disk.

Figure 8-9 shows a GDPS MGM 4-site configuration when it is in a multi-target topology. GDPS MGM 4-site configurations can dynamically switch between a cascaded topology and a multi-target topology to optimize processing of various recovery scenarios.
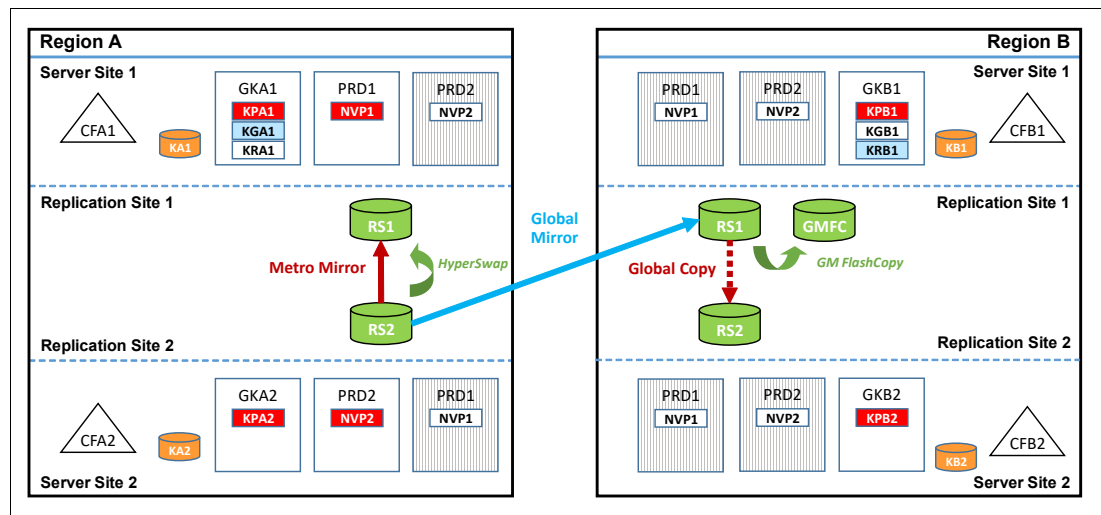

Figure 8-9   GDPS MGM 4-site multi-target configuration

Assume that your GDPS MGM Multi-Site 4-Site configuration started out in a cascaded topology, as shown in Figure 8-8 on page 252. If you run a planned HyperSwap to the RS2 disk in Region A, followed by a resynchronization of Metro Mirror from the RS2 disk back to the RS1 disk, you find yourself in the multi-target topology that is shown in Figure 8-9. In the figure, the RS2 disk in Region A is now the primary copy of data that application systems are accessing and the RS1 disk is the Metro Mirror secondary disk.

HyperSwap was reenabled to provide HA for the Region A data. This relationship is managed by GDPS Metro in Region A. The RS2 disk in Region A is also the GM primary disk, being copied to the RS1 disk in Region B, which is the GM secondary disk. This configuration is managed by using GDPS GM. IR is still enabled from the RS1 disk in Region A to the RS2 disk in Region B to protect from a failure of the RS2 disk in Region A and allow the GM copy to be re-established without the need for a full copy.

Finally, in Region B, the RS2 disk is still a Global Copy secondary to the RS1 disk. Again, this relationship, which is managed by GDPS Metro running in Region B, can be converted to fully synchronous Metro Mirror when you perform a switch of production to Region B for whatever reason.

The advantage of the multi-target capability in this scenario is that, following the HyperSwap, GM from the RS2 disk in Region A to the RS1 disk in Region B can remain active, which maintains your DR position, while Metro Mirror in Region A is being resynchronized from the RS2 disk back to the RS1 disk. In the same situation with cascaded-only MGM 4-Site, GM from the RS2 disk in Region A to the RS1 disk in Region B must be suspended while Metro Mirror in Region A is being resynchronized, which results in your DR position aging until the resync is complete.

The MGM 4-site configurations, as mentioned, remove the single point of failure of disk when you switch to the recovery region. As with GDPS MGM 3-site, precoded procedures are provided by GDPS to manage the following scenarios in the 4-site environments:

► Moving the GM session if there is a GM primary disk subsystem failure.
► Reintroduction of the intermediate disk subsystem.
► Planned Region switch to move production to the opposite region.

However, the following other considerations exist for an MGM 4-site configuration over those considerations for MGM 3-site configurations:

► More disk capacity does not need to be installed for X-disk or FC1 disk in each of the regions to facilitate testing of DR procedures while the GM environment is running. DR testing can be done on the RS2 disk in the current recovery region without affecting the DR position. For this reason, X-disk is not supported in GDPS MGM 3-site configurations.

  However, a test copy can be created on an external disk subsystem other than RS2 in the recovery region, if required for reasons other than DR testing, such as seeding an external test environment with production data. This process can be done by using the Testcopy Manager (TCM) feature. For more information, see 9.2, "Introducing LCP and Testcopy Manager" on page 257.

► The use of asymmetric devices in the remote copy configuration is not supported.

► Use of GDPS HM is not supported in a 4-site configuration because the IR function is required.

► The integration points between GDPS MGM 3-site and GDPS Continuous Availability that is discussed in 8.3.13, "GDPS MGM 3-site integration with GDPS Continuous Availability" on page 251 are *not* available with GDPS MGM 4-site.

8.4.1 Benefits of a GDPS MGM 4-site configuration

You can probably see that in effect, a 4-site configuration is managed as two separate 3-site MGM configurations, where the fourth copy is most relevant when you perform a region switch, or when you want to perform a DR test. A 4-site MGM configuration includes the following key advantages:

► HA capability when running production in either region.

► Retention of DR capability after a region switch. In a 3-site MGM configuration, your DR position ages while running on the region B RS1 disk.

► Nearly identical operational procedures when running in either region.

One final advantage of the GDPS MGM 4-Site configuration is that the TCM feature is available. For more information about the benefits that are provided by this feature, see Chapter 9, "IBM GDPS Logical Corruption Protection and Testcopy Manager" on page 255.

9

IBM GDPS Logical Corruption Protection and Testcopy Manager

In this chapter, we provide an overview of the GDPS Logical Corruption Protection (LCP) and Testcopy Manager (TCM) features, which are separately priced features of IBM Geographically Dispersed Parallel Sysplex (GDPS).

LCP is a set of GDPS capabilities that are provided in response to the growing number of requests for a GDPS managed "Continuous Data Protection" capability and is aimed at helping clients to recover from cyberattacks, internal threats, and other forms of logical data corruption.

At a high level, LCP captures multiple, secure point-in-time (PiT) copies of critical production data (referred to as *protection copies*) to restore the data back into production, if necessary. LCP can also recover a specific PiT copy to another set of devices that can be used to start one or more isolated recovery systems to analyze the scope of a particular logical corruption event.

More security and protection are provided for the LCP protection copies than for copies that are taken with more traditional methods by minimizing host access to these volumes and by providing specific roles and rules for their management.

TCM is a GDPS feature that allows clients to manage, capture, and refresh a test copy for use within an isolated test environment. This isolated copy is created and maintained by using a Global Copy relationship cascaded from an existing GDPS environment.

255

This chapter includes the following topics:

9.1 LCP terminology

The following terms are used when describing LCP:

► Copy sets (CS)

   CS are a grouping of volumes that together provide a consistent PiT copy of data. Taking a copy to a Copy Set is called *capturing* a protection copy (also referred to as a backup copy). The protection copy can be captured by using the Safeguarded Copy (SGC) technology or the FlashCopy technology.

► SGC CS

   The SGC CS are CS that are created with the SGC technology. Protection copies (or backups) taken with SGC require a predefined, virtual capacity, which is known as SGC Backup Capacity, to store the consistent backups of the associated source volumes. The SGC Backup Capacity does not use a device number in the storage system; therefore, it cannot be directly addressed by any host system, isolating them and protecting them from conventional host I/O access. Up to 500 SGC backups per production volume are supported by the SGC technology.

   GDPS supports up to 10 SGC CS. Up to 500 protection copies can be taken for each SGC copy set to separately manage the LCP operations for multiple groups of volumes. For example, you might manage LCP operations separately for production volumes and test volumes.

   Although up to 500 protection copies can be taken for each SGC copy set, the number of protection copies that can be maintained for a specific production device in an environment is dependent upon the amount of physical space that is assigned to the SGC Backup Capacity that is associated with the production device.

   The more updates that occur to the production volume, the more space is required in the SGC Backup Capacity to maintain the copies. For this reason, the SGC Backup Capacity can be dynamically expanded to accommodate changes in update patterns for devices that are protected with SGC.

► FlashCopy CS

   The FlashCopy CS are CS created with the FlashCopy technology. Protection copies (or backups) taken with FlashCopy are isolated and protected from conventional host I/O access by not requiring any unit control blocks (UCBs) to be defined in any host for them and by not allowing writes to them, which in turn, prevent them from being started from.

   GDPS supports up to 10 FlashCopy CS. However, the total number of FlashCopy CS and Recovery CS combined cannot exceed 11.

- ► Recovery copy (RC) CS

  An RC copy set is used for corruption analysis or recovery testing. It is created by FlashCopying from a copy set and is not Target Write Inhibited, which allows it to be used to IPL a recovery or test system.

  GDPS supports up to 10 RC CS. However, the total number of RC CS and FlashCopy CS combined cannot exceed 11.

- ► Replication site (RS) volume sets

  The RS volume sets are the conventional RS volume sets that form the basis of the GDPS replication model and can be directly started from when they are in a primary status.

- ► LCP management profile

  An LCP management profile describes the management characteristics of the protection copies, such as the following items:

  - The RS where the captures are to be taken.
  - The consistency group to be captured.
  - CS that is assigned to this management profile.
  - How long a capture must be retained before it expires and becomes eligible for release.
  - The minimum number of captures to maintain regardless of the retention period.
  - Whether expired captures should be automatically released.
  - How much time must elapse before a new capture can be taken.

  The management profile name is unique within the LCP environment and you can define multiple LCP management profiles.

- ► LCP Restore

  The LCP Restore process is the restoration of data in which the target is a production RS volume set.

- ► LCP Recover

  The recover process is the recovery of data in which the target is the RC copy set. The source of the recovery is a specific FlashCopy or SGC copy set.

9.2 Introducing LCP and Testcopy Manager

Two implementations of LCP are provided: *internal LCP* and *external LCP*. The following sections introduce to these implementations of LCP and to the TCM.

9.2.1 Internal LCP

With internal LCP, the protection copies are within the same storage system as one of the existing production copies in your GDPS environment. This configuration is also referred to as a *virtual airgap* model or a *virtual isolation* model.

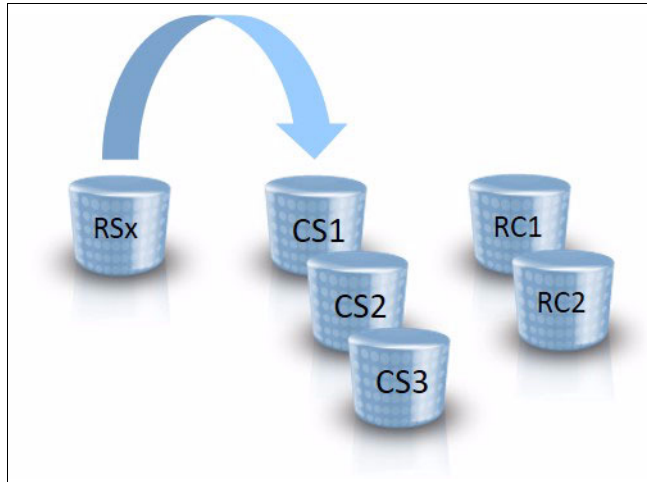Figure 9-1 shows an internal LCP topology.

Figure 9-1   Internal LCP (virtual airgap)

In Figure 9-1, RSx is one of the RS volume sets in your GDPS production environment. CS1, CS2, and CS3 are the CS that contain the protection copies. RC1 and RC2 are RC CS. All of these devices are in the same storage server.

With internal LCP, one of the controlling systems in your GDPS environment manages the LCP environment and is referred to as an *internal LCP Manager*. Therefore, no extra controlling systems are required.

9.2.2 External LCP

External LCP environments differ from internal LCP environments in that the protection copies are isolated from your production environment. An extra copy of data is cascaded from one of the existing production copies in your GDPS environment to a storage system that is external to your production environment.

The protection copies are then captured and maintained within the external storage system. This configuration is also referred to as a *physical airgap* model or a *physical isolation* model. Figure 9-2 shows an external LCP topology.
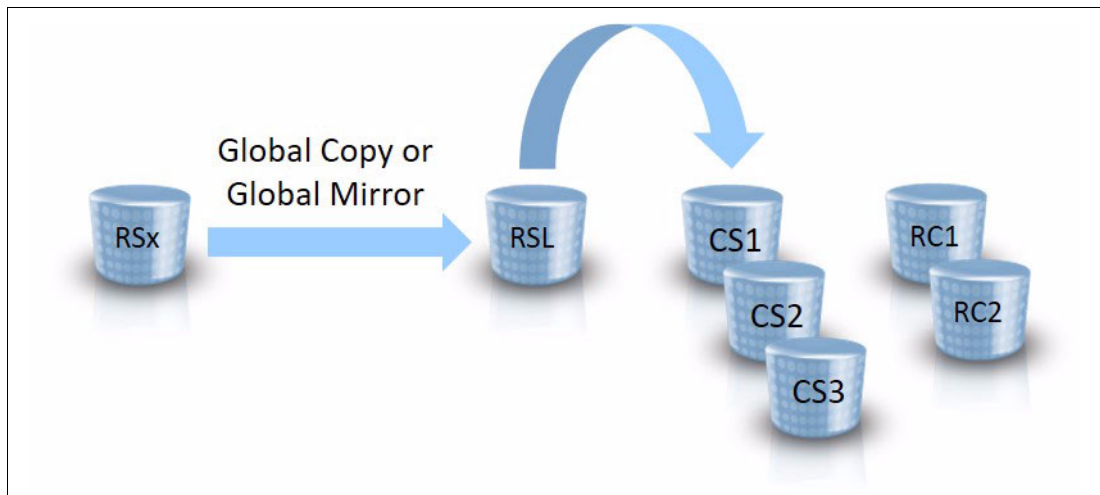

Figure 9-2   External LCP (physical airgap)

In Figure 9-2 on page 258, RSx is again one of the RS volume sets in your GDPS production environment. Global Copy or Global Mirror (GM) (depending on the type of GDPS environment that external LCP is implemented in) is used to replicate the data from the RSx volume set to the volume set labeled RSL in Figure 9-2 on page 258, which is in a different storage system than RSx.

The CS labeled CS1, CS2, and CS3 are the protection copies, which are captured from RSL. The RSL volume set, the CS1, CS2, and CS3 CS, and the recovery volume sets (labeled RC1 and RC2 in Figure 9-2 on page 258) are all in the same storage server.

With external LCP, another external controlling system is required to manage the LCP environment and is referred to as an *external LCP Manager*.

The RC copy set, in both internal and external LCP environments, enables the IPL of systems for forensic analysis or other purposes.

In most client configurations, the FlashCopy CS, the SGC Backup Capacity (when SGC is used), and the RC copy set are all thinly provisioned to minimize space requirements.

9.2.3 Testcopy Manager

The TCM function can be viewed as a subset of the external LCP Manager function. As with the external LCP Manager function, another copy of data is cascaded from one of the existing production copies in your GDPS environment to a storage system external to your production environment.

The main difference between the TCM function and the LCP Manager function is that, with the TCM function, only one PiT copy can be taken and maintained in the external storage system. This copy can then be used in an isolated test environment and can be refreshed with a new copy of production data as required.

Figure 9-3 shows a TCM environment.

Figure 9-3   Testcopy Manager

In Figure 9-3, RSx is again one of the RS volume sets in your GDPS production environment. Global Copy is used to replicate the data from the RSx volume set to the volume set that is labeled RST in Figure 9-3, which is in a different storage system than RSx. The test copy (labeled TC1 in Figure 9-3) is then captured from the RST volume set. The RST volume set and the test copy volume set are in the same storage server.

The TCM feature is included with the External LCP feature. In this case, an LCP environment and a TCM environment can be implemented and maintained from the same production copy of data. The TCM can also be licensed separately from the external LCP Manager. In that case, only a TCM environment can be implemented.

9.3 LCP operational models

The following sections describe the support that is provided for internal and external LCP, and TCM in the various GDPS solution offerings.

9.3.1 GDPS Metro

For more information about the GDPS Metro offering, see Chapter 3, "IBM GDPS Metro" on page 45.

Internal and external LCP is supported in GDPS Metro environments, as described next.

Internal LCP in GDPS Metro environments

Internal LCP can be implemented from any or all RS volume sets in a GDPS Metro environment. Figure 9-4 shows an example of internal LCP in a GDPS Metro single-leg environment.
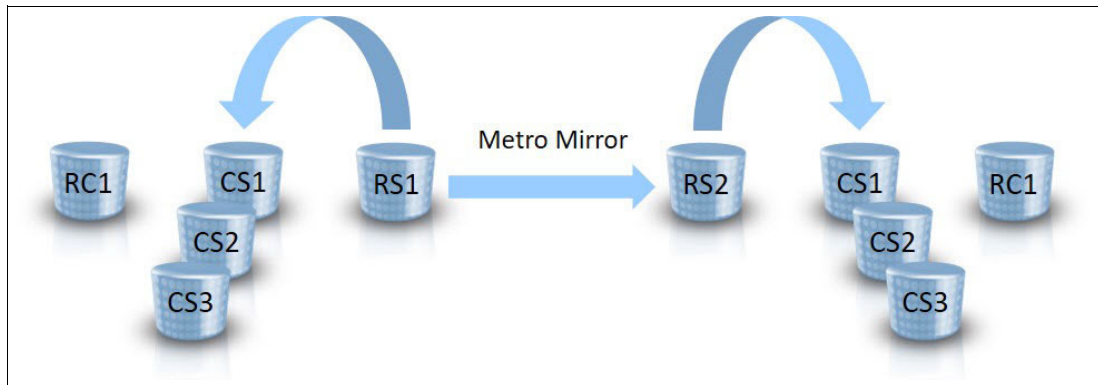

Figure 9-4   Internal LCP in a GDPS Metro single-leg environment

As shown in Figure 9-4, internal LCP is implemented on the RS1 and the RS2 copies of data. Each of these LCP environments consists of three protection copies or CS that are labeled CS1, CS2, and CS3 and one RC copy set that is labeled RC1.

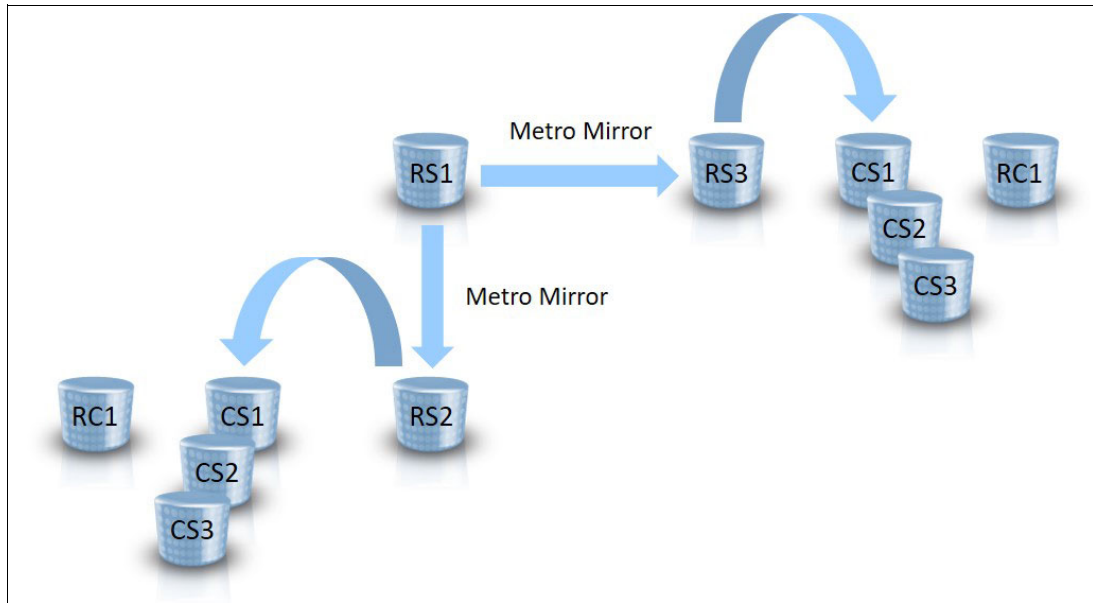Figure 9-5 on page 261 shows an example of internal LCP in a GDPS Metro dual-leg environment.

Figure 9-5   Internal LCP in a GDPS Metro dual-leg environment

As shown in Figure 9-5, internal LCP is implemented on two of the three production copies of data (RS2 and RS3). Again, each of these LCP environments consists of three protection copies or CS that are labeled CS1, CS2, and CS3 and one RC copy set that is labeled RC1.

Consider the following points when internal LCP is implemented in GDPS Metro single-leg or dual-leg environments:

► All internal LCP environments that are implemented in a GDPS Metro configuration are managed by the standard GDPS Metro controlling systems that manage the entire GDPS Metro environment.

► The CS in GDPS Metro internal LCP environments (labeled CS1, CS2, and CS3 in Figure 9-4 on page 260 and in Figure 9-5) can be FlashCopy CS or SGC CS. A mixture of FlashCopy CS and SGC CS is also supported. For example, the CS in the LCP environment that is implemented on the RS1 devices can be FlashCopy CS while the CS that are implemented on the RS2 devices can be SGC CS.

► When an internal LCP protection copy is captured in a GDPS Metro environment, updates to the LCP source devices are temporarily held up so that a consistent copy can be captured. LCP source devices also serve as one of the production RS volume sets in your GDPS Metro environment. Therefore, this temporary freezing of the updates might affect your production applications, which is true whether the LCP source volumes are serving as your Metro Mirror primary devices or your Metro Mirror secondary devices.

The amount of time that updates are frozen depends on many factors, such as the number of logical subsystems (LSSs) and devices in your GDPS Metro configuration. Evaluate whether your applications can tolerate the effect of capturing an LCP copy in your environment.

External LCP (which is described next) avoids this impact to production applications.

External LCP in GDPS Metro environments

External LCP is supported for both single-leg and dual-leg GDPS Metro environments. For single-leg environments, external LCP can be implemented from one or both of the RS volume sets. Figure 9-6 shows an example of external LCP in a GDPS Metro single-leg environment.
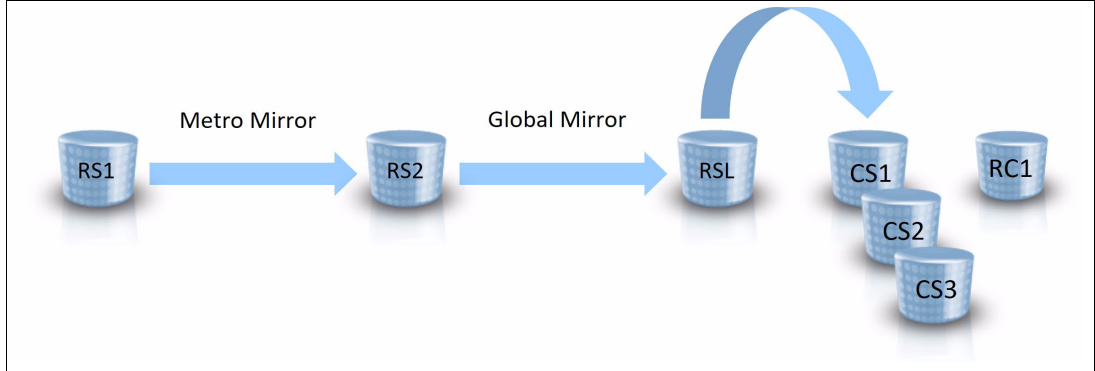

Figure 9-6   External LCP in a GDPS Metro single-leg environment

As shown in Figure 9-6, external LCP is implemented on the RS2 volume set, which is the current set of Metro Mirror secondary devices. GM is used to copy the data from RS2 to the external copy of data (labeled RSL in Figure 9-6). The RSL volume set serves as the source for the LCP protection copy captures. The LCP environment then consists of the RSL devices, along with three protection CS (labeled CS1, CS2, and CS3) and one RC copy set (labeled RC1).

For dual-leg environments, external LCP can be implemented only from the Site2 volumes, which typically are defined as the RS3 volume set. Figure 9-7 shows an example of external LCP in a GDPS Metro dual-leg environment.
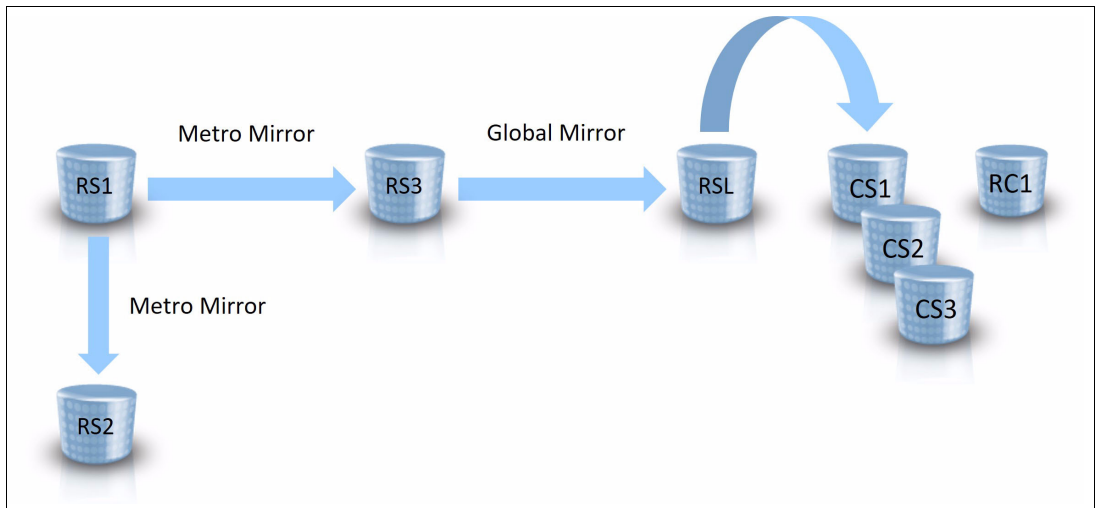

Figure 9-7   External LCP in a GDPS Metro dual-leg environment

As shown in Figure 9-7, external LCP is implemented on the RS3 volume set, which is the set of Metro Mirror secondary devices in Site2. GM is used to copy the data from RS3 to the external copy of data (labeled RSL in Figure 9-7). The RSL volume set serves as the source for the LCP protection copy captures. Then, the LCP environment consists of the RSL devices along with three protection CS (labeled CS1, CS2, and CS3) and one RC copy set (labeled RC1).

The LCP environment in a GDPS Metro external LCP solution is managed by an instance of GDPS Global - GM (GDPS GM) that runs on a separate controlling system from the GDPS Metro controlling system functions (For more information about the GDPS Global - GM offering, see Chapter 5, "IBM GDPS Global - GM" on page 147). This LCP Manager system is used to maintain GM to the RSL volume set, in addition to capturing and recovering the LCP protection copies.

For scalability, the RS2 volume set in a GDPS Metro single-leg environment can be defined to the GDPS GM LCP Manager system in MSS1. Likewise, the RS3 volume set in a GDPS Metro dual-leg environment can be defined to the GDPS GM LCP Manager system in MSS2.

The CS in GDPS Metro external LCP environments (labeled CS1, CS2, and CS3 in Figure 9-6 on page 262 and Figure 9-7 on page 262) can be SGC CS only.

In most client configurations, the SGC Backup Capacity and the RC copy set are thin provisioned to minimize space requirements.

The following steps are taken when a request is made to capture an external LCP protection copy in a GDPS Metro environment:

1. GM is paused on a consistent boundary. This process results in the RSL volume set containing a consistency copy of data and the suspension of the GM session.

2. A protection copy is captured from the RSL volume set to one of the SGC CS.

3. GM is resumed.

This method avoids the potential impact to production applications that is associated with taking an internal LCP copy (for more information, see "Internal LCP in GDPS Metro environments" on page 260).

9.3.2 GDPS Global - GM

For more information about the GDPS Global - GM (GDPS GM) offering, see Chapter 5, "IBM GDPS Global - GM" on page 147.

Internal and external LCP is supported in GDPS GM environments, as described next.

Internal LCP in GDPS GM environments

Internal LCP can be implemented from the GM secondary device set (B.RS1) in GDPS GM environments. Figure 9-8 shows internal LCP in a GDPS GM environment.
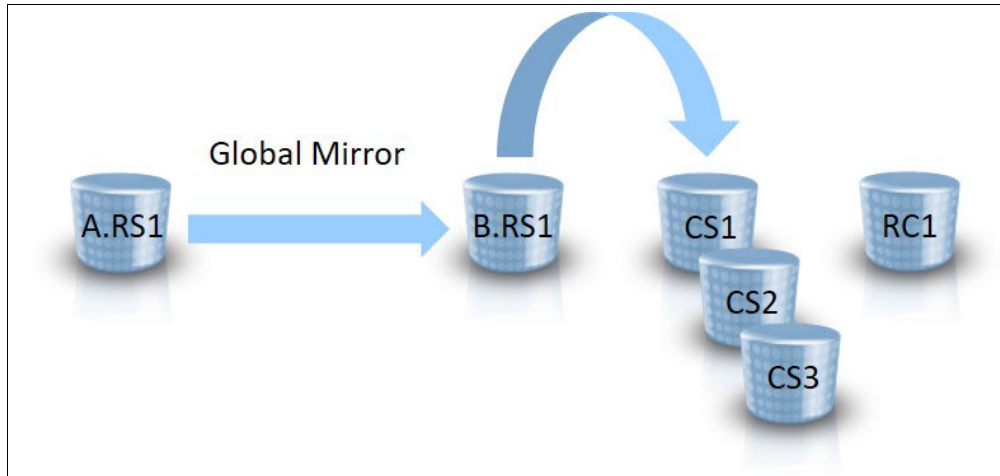

Figure 9-8   Internal LCP in a GDPS Global - GM environment

As shown in Figure 9-8, the RS1 devices in the production region (referred to as Region A) labeled A.RS1, are the GM primary devices that the production applications use.

Internal LCP is implemented from the RS1 devices in the recovery region (referred to as Region B), labeled B.RS1 in the figure, which are also the GM secondary devices. The LCP environment consists of three protection copies or CS that are labeled CS1, CS2, and CS3 and one RC copy set labeled RC1.

The LCP environment in a GDPS GM solution is managed by the GDPS GM controlling system that runs in the recovery region, referred to as the Kr-system, and the CS in a GDPS GM LCP environment must be SGC CS.

The following steps are taken when a request is made to capture a protection copy in a GDPS GM environment:

1. GM is paused on a consistent boundary, which results in the GM secondary devices (B.RS1) containing a consistent copy of data and the GM session being suspended.

2. A protection copy is captured from the B.RS1 volume set to one of the SGC CS.

3. GM is resumed.

This method avoids any impact to production applications during the LCP copy capture process at the expense of a typically minor disruption to your recovery point objective (RPO).

External LCP in GDPS GM environments

External LCP can be implemented from the GM secondary device set (B.RS1) in GDPS GM environments. Figure 9-9 shows external LCP in a GDPS GM environment.
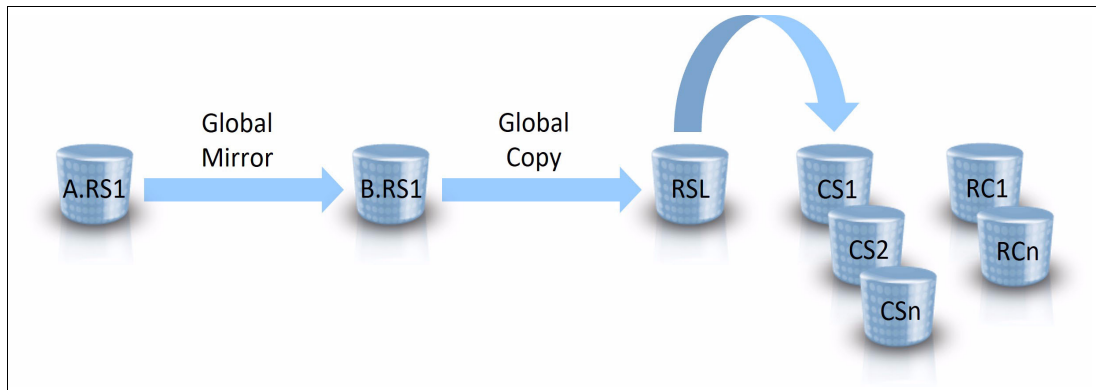

Figure 9-9   External LCP in a GDPS Global - GM environment

As shown in Figure 9-9, a standard GDPS GM 2-site environment exists with GM running between the A.RS1 devices in the production region (referred to as $Region\ A$) and the B.RS1 devices in the recovery region (referred to as $Region\ B$). The A.RS1 devices are the devices that the production applications use.

To seed the external LCP environment, a Global Copy relationship is added to the environment to mirror the data from the B.RS2 devices to the external copy of data (labeled RSL in Figure 9-9). The RSL volume set serves as the source for the LCP protection copy captures. The LCP environment then consists of the RSL devices, along with several protection CS that are labeled CS1, CS2, and CSn and multiple RC CS that are labeled RC1 and RCn.

In this model, the external LCP environment is managed by an instance of GDPS Metro that runs on a separate controlling system from the GDPS GM controlling system functions and that is configured in a single-leg topology.

For more information about the GDPS Metro offering, see Chapter 3, "IBM GDPS Metro" on page 45. This LCP Manager system is used to maintain Global Copy mirroring to the RSL volume set, in addition to capturing and recovering the LCP protection copies.

Consider the following points when external LCP is implemented in GDPS GM 2-site environments:

► The LCP environment can be created and maintained from the B.RS1 devices only.

► The protection copies can be FlashCopy CS or SGC CS.

► Capturing an external LCP protection copy requires task coordination between the LCP Manager controlling system and the controlling systems that make up the GDPS GM 2-site solution. The key coordination task here is creating a consistent data point on the RSL volume set so that it can then be captured to a corresponding copy set. Global Copy is used to maintain the RSL copy and as discussed in 2.4.2, "Global Mirror" on page 26, Global Copy does not provide data consistency.

► The following steps are taken (when a request is made to capture a protection copy) to coordinate consistency across the environment to the RSL volume set and then to capture the consistency group on the suitable copy set devices:

a. GM is paused on a consistent boundary. This process results in the suspension of the GM secondary devices (B.RS1) that contain a consistency copy of data and the GM session.

b. The consistent copy of data is allowed to drain to the RSL volume set.

c. After the consistent copy of data arrives at RSL, the Global Copy relationship between B.RS1 and RSL is suspended to prevent any more updates to the RSL volume set until after the consistent copy of data is captured to a copy set in a subsequent step.

d. GM is resumed, which allows updates to again flow from Region A to the B.RS1 volume set.

e. A protection copy is captured from the RSL volume set to one of the associated CS.

f. The Global Copy relationship between B.RS1 and RSL is resumed to allow updates to again flow to the RSL volume set until the next protection copy is taken.

This approach minimizes the time that it takes to run the capture process each time by keeping the RSL volume set as current as possible. It also avoids any effect on production applications during the LCP copy capture process at the expense of a typically minor disruption to your RPO.

9.3.3 GDPS Metro Global - GM

For more information about the GDPS Metro Global - GM (GDPS MGM) offering, see Chapter 8, "Combining local and metro continuous availability with out-of-region disaster recovery" on page 237.

The following sections describe the LCP and TCM features of the GDPS MGM solutions.

GDPS Metro Global - GM 3-site

Internal and external LCP is supported in GDPS MGM 3-site environments, as described next.

Internal LCP in GDPS Metro Global - GM 3-site environments

Figure 9-10 on page 267 shows an example of internal LCP in a GDPS MGM 3-site environment.
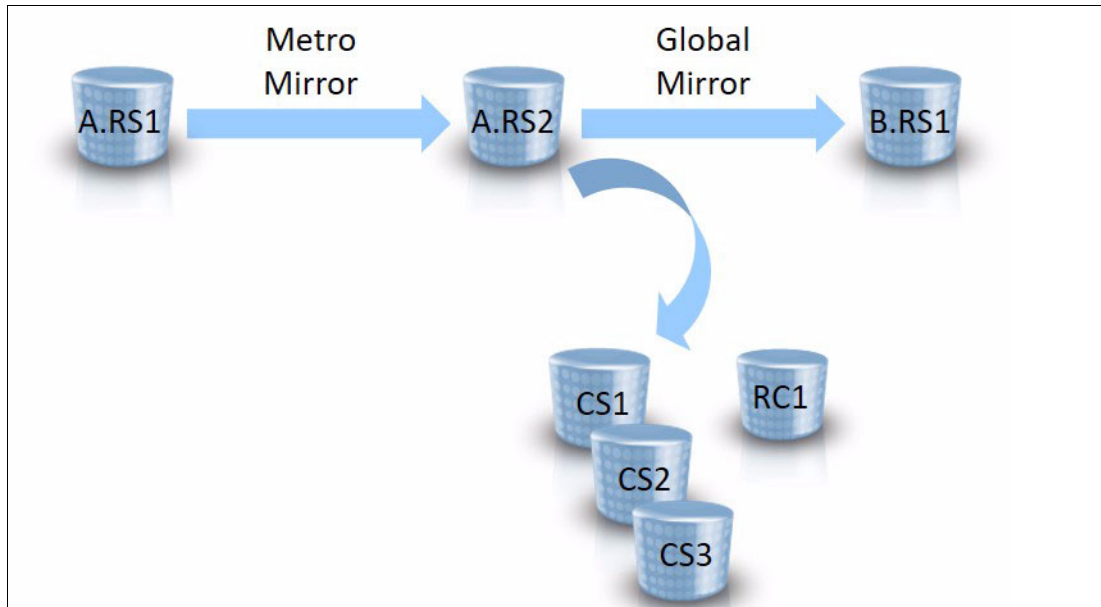
Figure 9-10   Internal LCP on the Metro leg in a GDPS MGM 3-site environment

As shown in Figure 9-10, a standard GDPS MGM 3-site environment exists with Metro Mirror running between the A.RS1 devices and the A.RS2 devices in the production region (referred to as *Region A*) and GM running between the A.RS2 devices in Region A and the B.RS1 devices in the recovery region (referred to as *Region B*).

The A.RS1 devices are the devices that the production applications are using. Internal LCP is implemented on the A.RS2 volume set, which is the current set of Metro Mirror secondary devices. The LCP environment consists of three protection copies or CS that are labeled CS1, CS2, and CS3 and one RC copy set that is labeled RC1.

In addition to implementing internal LCP on the Metro Mirror secondary volume set as shown in this example, you can also implement internal LCP on the Metro Mirror primary volume set (A.RS1 in Figure 9-10) or on both the primary and secondary volume sets in Region A at the same time.

Consider the following points when internal LCP is implemented on the Metro leg (on the Metro Mirror primary or secondary volume sets) in a GDPS MGM 3-site environment:

► The LCP environment is managed by the standard GDPS Metro controlling systems that manage the Metro leg in the active production region.

► The CS can be FlashCopy CS or SGC CS. A mixture of FlashCopy CS and SGC CS is also supported. For example, the CS in the LCP environment that is implemented on the A.RS1 devices can be FlashCopy CS; the CS that are implemented on the A.RS2 devices can be SGC CS.

► When a protection copy is captured, updates to the LCP source devices are temporarily held up so that a consistent copy can be captured. LCP source devices also serve as one of the production RS volume sets in the GDPS Metro leg of your environment. Therefore, this temporary freezing of the updates might affect your production applications. This result is true whether the LCP source volumes are serving as your Metro Mirror primary devices or your Metro Mirror secondary devices.

The amount of time that updates are held up depends on several factors, such as the number of LSSs and devices in your GDPS Metro configuration. Evaluate whether your applications can tolerate the effect of capturing an LCP copy in your environment.

Implementing internal LCP on the GM secondary devices in the recovery region, which is described next, avoids this effect on production applications. Implementing external LCP on the GM secondary devices in the recovery region, which is described in "External LCP in GDPS Metro Global - GM 3-site environments" on page 269, also avoids this effect on production applications.

Figure 9-11 shows an example of internal LCP implemented on the GM secondary devices in the recovery region in a GDPS MGM 3-site environment.
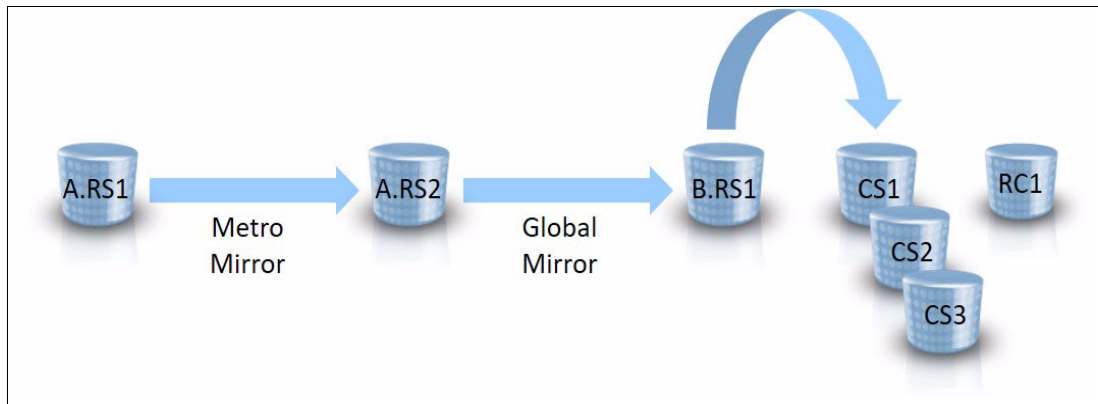

Figure 9-11 Internal LCP on the GM leg in a GDPS MGM 3-site environment

In Figure 9-11, we again have Metro Mirror running between the A.RS1 devices and the A.RS2 devices in Region A and GM running between the A.RS2 devices in Region A and B.RS1 devices in Region B. Internal LCP is implemented on the B.RS1 volume set, which contains the GM secondary devices. The LCP environment consists of three protection copies or CS that are labeled CS1, CS2, and CS3 and one RC copy set that is labeled RC1.

The LCP environment in this configuration is managed by the GDPS GM controlling system that runs in the recovery region, which is referred to as the *Kr-system*, and the CS must be SGC CS.

The following steps are taken when a request is made to capture a protection copy in this environment:

1. Global Mirror is paused on a consistent boundary. This process results in the Global Mirror secondary devices (B.RS1) containing a consistent copy of data and the Global Mirror session being suspended.
2. A protection copy is captured from the B.RS1 volume set to one of the SGC CS.
3. Global Mirror is resumed.

This method avoids any effect on production applications during the LCP copy capture process at the expense of a typically minor disruption to your RPO.

Finally, you might want to maintain LCP protection copies in both regions of your GDPS MGM 3-site environment. This configuration is possible by implementing internal LCP on one of the volume sets in Region A and also implementing internal LCP on the B.RS1 devices in Region B. Figure 9-12 on page 269 shows an example of such an environment.
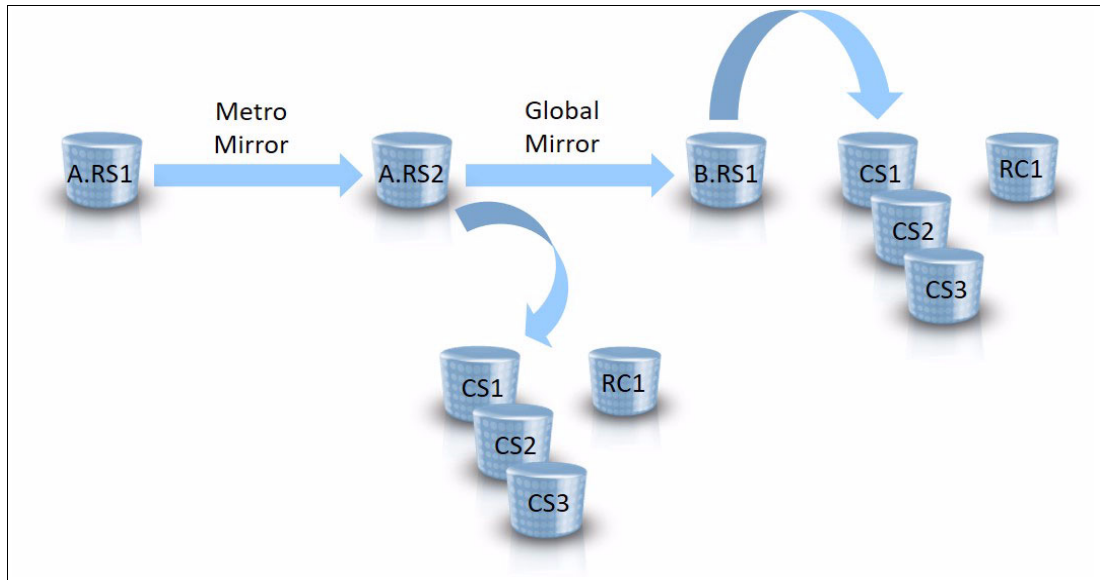
Figure 9-12   Internal LCP in both regions in a GDPS MGM 3-site environment

The environment that is shown in Figure 9-12 looks similar to the environment that is shown in Figure 9-11 on page 268. The key difference is that in addition to having an internal LCP environment implemented on the B.RS1 volume set in Region B, we also have an internal LCP environment that is implemented on the A.RS2 volume set in Region A. Each LCP environment contains a distinct set of protection CS and a distinct recovery volume set.

The LCP environment in Region A is managed by the standard GDPS Metro controlling systems that manage the Metro leg of the environment and the LCP environment in Region B is managed by the GDPS GM Kr-system. These LCP environments are managed separately and protection copies are captured in each one independently.

Having LCP protection copies in both regions ensure that you can recover from logical corruption events, even if one Region failed or connectivity between regions was interrupted.

External LCP in GDPS Metro Global - GM 3-site environments

Figure 9-13 shows an external LCP in a GDPS MGM 3-site environment.
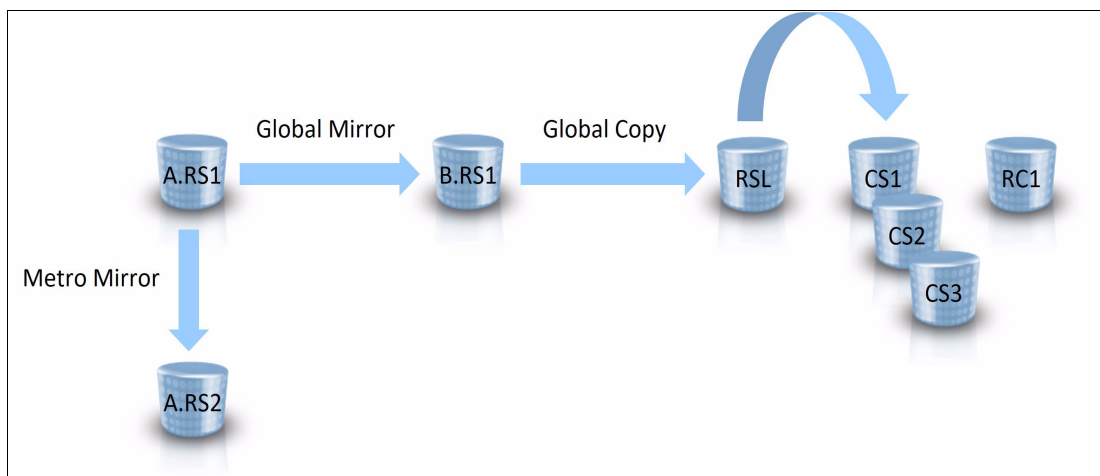

Figure 9-13   External LCP in a GDPS MGM 3-site environment

In Figure 9-13 on page 269, a standard GDPS MGM 3-site environment exists with Metro Mirror running between the A.RS1 devices and the A.RS2 devices in the production region (referred to as *Region A*) and GM running between A.RS1 in Region A and B.RS1 in the recovery region (referred to as *Region B*). The A.RS1 devices are the devices that the production applications are using.

To seed the external LCP environment, a Global Copy relationship is added to the environment to mirror the data from the B.RS2 devices to the external copy of data (labeled RSL in Figure 9-13 on page 269). The RSL volume set serves as the source for the LCP protection copy captures. The LCP environment then consists of the RSL devices, along with three protection CS that are labeled CS1, CS2, and CS3 and one RC copy set labeled RC1.

The environment that is shown in Figure 9-13 on page 269 is in a multi-target configuration, as opposed to the cascaded configurations that the previous GDPS MGM 3-site environments that were shown in thus far. This difference illustrates that the environment can be in either configuration (multi-target or cascaded) and can be switched between the configurations while LCP is active.

In this model, the external LCP environment is managed by an instance of GDPS Metro that runs on a separate controlling system from the GDPS MGM controlling system functions and that is configured in a single-leg topology. For more information about the GDPS Metro offering, see Chapter 3, "IBM GDPS Metro" on page 45. This LCP Manager system is used to maintain Global Copy mirroring to the RSL volume set, in addition to capturing and recovering the LCP protection copies.

Consider the following points when external LCP is implemented in GDPS MGM 3-site environments:

► The LCP environment can be created and maintained from the B.RS1 devices only.

► The protection copies can be FlashCopy CS or SGC CS.

► Capturing an external LCP protection copy requires task coordination between the LCP Manager controlling system and the controlling systems that make up the GDPS MGM 3-site solution. The key coordination task here is creating a consistent data point on the RSL volume set so that it can then be captured to a corresponding copy set. Global Copy is used to maintain the RSL copy and as discussed in 2.4.2, "Global Mirror" on page 26, Global Copy does not provide data consistency.

The following steps are taken (when a request is made to capture a protection copy) to coordinate consistency across the environment to the RSL volume set and then to capture the consistency group on the appropriate copy set devices:

a. GM is paused on a consistent boundary. This process results in the GM secondary devices (B.RS1) containing a consistency copy of data and the GM session being suspended.

b. The consistent copy of data is then allowed to drain to the RSL volume set.

c. After the consistent copy of data arrives at RSL, the Global Copy relationship between B.RS1 and RSL is suspended to prevent any more updates to the RSL volume set until after the consistent copy of data is captured to a copy set in a subsequent step.

d. GM is resumed, which allows updates to again flow from Region A to the B.RS1 volume set.

e. A protection copy is captured from the RSL volume set to one of the associated CS.

f. The Global Copy relationship between B.RS1 and RSL is resumed to allow updates to again flow to the RSL volume set until the next protection copy is taken.

This approach minimizes the time that it takes to run the capture process each time by keeping the RSL volume set as current as possible. It also avoids any effect on production applications during the LCP copy capture process at the expense of a typically minor disruption to your RPO.

GDPS Metro Global - GM 4-site

Internal LCP, External LCP, and TCM are supported in GDPS MGM 4-site environments.

Internal LCP in GDPS Metro Global - GM 4-site environments

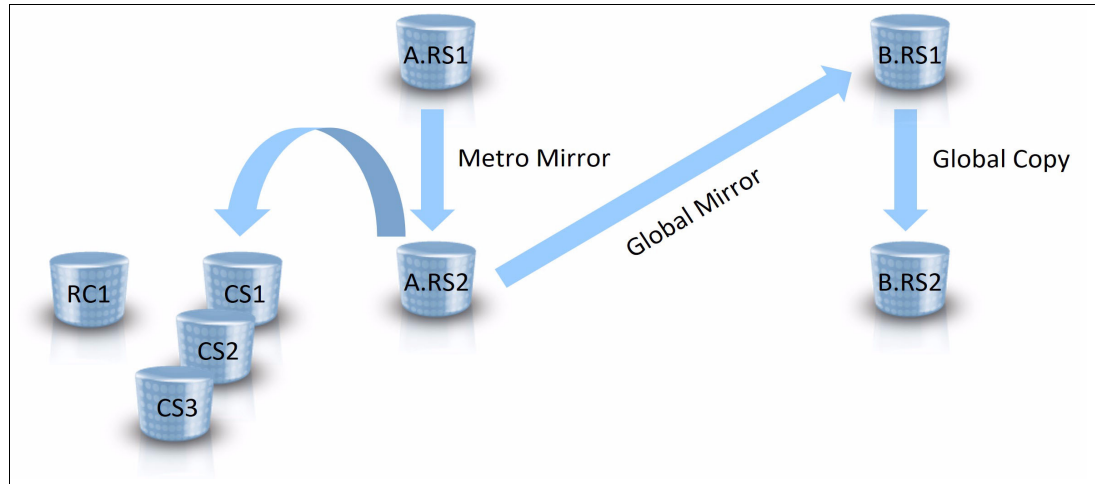Figure 9-14 shows an example of internal LCP in a GDPS MGM 4-site environment.

Figure 9-14   Internal LCP in the current production region in a GDPS MGM 4-site environment

In Figure 9-14, there is a standard GDPS MGM 4-site environment with Metro Mirror running between the A.RS1 devices and the A.RS2 devices in the production region (referred to as Region A); GM running between A.RS1 in Region A and B.RS1 in the current recovery region (referred to as Region B); and Global Copy running between the B.RS1 devices to a second set of devices in Region B, which is labeled B.RS2.

The A.RS1 devices are the devices that the production applications use. Internal LCP is implemented on the A.RS2 volume set, which is the set of Metro Mirror secondary devices in the production region (Region A). The LCP environment consists of three protection copies or CS that are labeled CS1, CS2, and CS3, and one RC copy set that is labeled RC1.

In addition to implementing internal LCP on the Metro Mirror secondary volume set in Region A as shown in this example, you can also implement internal LCP on the Metro Mirror primary volume set in region A (A.RS1 in Figure 9-14) or on both the primary and secondary volume sets in Region A concurrently.

Internal LCP can also be implemented on one or both volume sets in the recovery region, as shown in Figure 9-15.
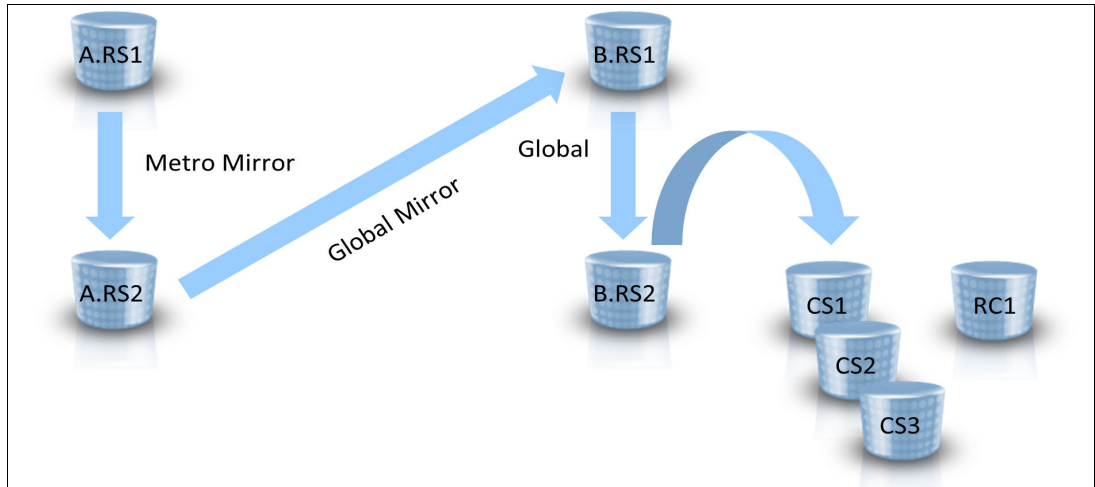

Figure 9-15   Internal LCP in the current recovery region in a GDPS MGM 4-site environment

In Figure 9-15, Metro Mirror is running between the A.RS1 devices and the A.RS2 devices in the production region (Region A); Global Mirror is running between A.RS1 in Region A and B.RS1 in the recovery region (Region B); and Global Copy is running between the B.RS1 devices and the B.RS2 devices in Region B. Internal LCP is implemented on the B.RS2 volume set, which is the set of Global Copy secondary devices in the current recovery region (Region B). The LCP environment consists of three protection copies or copy sets that are labeled CS1, CS2, and CS3, and one RC copy set that is labeled RC1.

You might want to maintain LCP protection copies in both regions of your GDPS MGM 4-site environment. This configuration is possible by implementing internal LCP on one or both volume sets in Region A and on one or both volume sets in Region B. Figure 9-16 shows an example of such an environment.
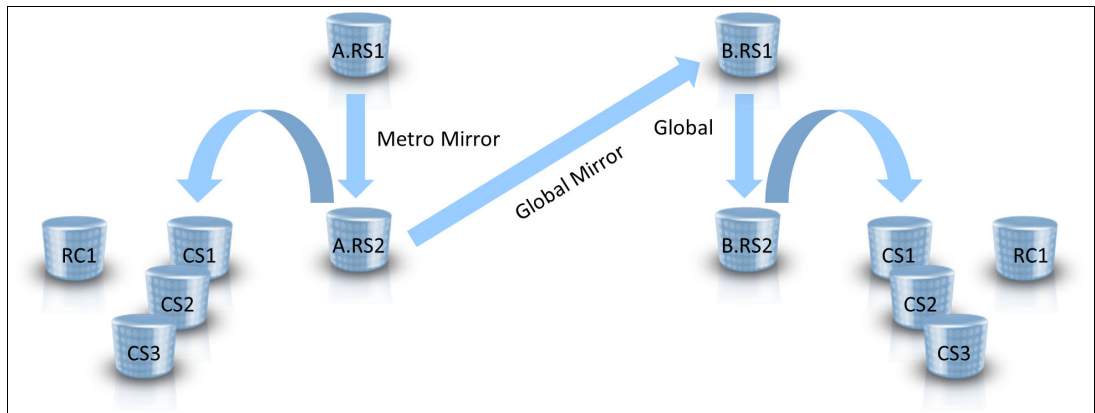
Figure 9-16   Internal LCP in both regions in a GDPS MGM 4-site environment

The environment that is shown in Figure 9-16 looks similar to the environments that are shown in Figure 9-14 on page 271 and Figure 9-15. The key difference is that an internal LCP environment is implemented in both regions: one on the A.RS1 volume set in Region A, and one on the B.RS1 volume set in Region B. Each LCP environment contains a distinct set of protection copy sets and a distinct recovery volume set.

Having LCP protection copies in both regions ensure that you can recover from logical corruption events, even if one Region failed or connectivity between regions was interrupted.

Consider the following points when internal LCP is implemented in a GDPS MGM 4-site environment:

► The LCP environment in Region A is managed by the standard GDPS Metro controlling systems in Region A, and the LCP environment in Region B is managed by the standard GDPS Metro controlling systems in Region B. These LCP environments are managed separately, and protection copies are captured in each one independently.

► The CS can be FlashCopy CS or SGC CS. A mixture of FlashCopy CS and SGC CS is also supported. For example, the CS in the LCP environment that is implemented on the A.RS2 devices can be FlashCopy CS, and the CSs that are implemented on the B.RS2 devices are SGC CS.

► When a protection copy is captured in the production region, updates to the LCP source devices are temporarily frozen so that a consistent copy can be captured. LCP source devices also serve as one of the production RS volume sets in the GDPS Metro leg of your environment. Therefore, this temporary freezing of the updates might affect your production applications. This result is true whether the LCP source volumes are serving as your Metro Mirror primary devices or your Metro Mirror secondary devices.

The amount of time that updates are held up depends on several factors, such as the number of LSSs and devices in your GDPS Metro configuration. Evaluate whether your applications can tolerate the effect of capturing an LCP copy in your environment.

► The following steps occur when a request is made to capture a protection copy in the current recovery region:

a. Global Mirror pauses on a consistent boundary. This process results in the Global Mirror secondary devices (the RS1 devices in the recovery region) that contain a consistent copy of data and the Global Mirror session being suspended.

b. If the capture is being taken on the RS1 volume set, the capture is taken now, Global Mirror is resumed, and the process is completed.

c. If the capture is being taken on the RS2 volume set, then after Global Mirror is suspended, the consistent copy of data may drain to the RS2 volume set.

d. After the consistent copy of data arrives at RS2, the Global Copy relationship between RS1 and RS2 is suspended to prevent any more updates to the RS2 volume set until after the consistent copy of data is captured to a copy set in a subsequent step.

e. Global Mirror is resumed, which enables updates to again flow from Region A to the RS1 volume set.

f. A protection copy is captured from the RS2 volume set to one of the associated copy sets.

g. The Global Copy relationship between RS1 and RS2 is resumed to enable updates to again flow to the RS2 volume set until the next protection copy is taken.

When captures are taken in the recovery region, there is no impact on production applications, but there is an impact to your Recovery Point Objective (RPO), the extent of which again depends on several factors, such as how many LSSs you have in your environment, how many devices you have in your environment, and others.

External LCP in GDPS Metro Global - GM 4-site environments

External LCP is supported only from the RS2 volume set in the recovery region in GDPS MGM 4-site environments. Figure 9-17 shows a GDPS MGM 4-site environment with external LCP implemented.
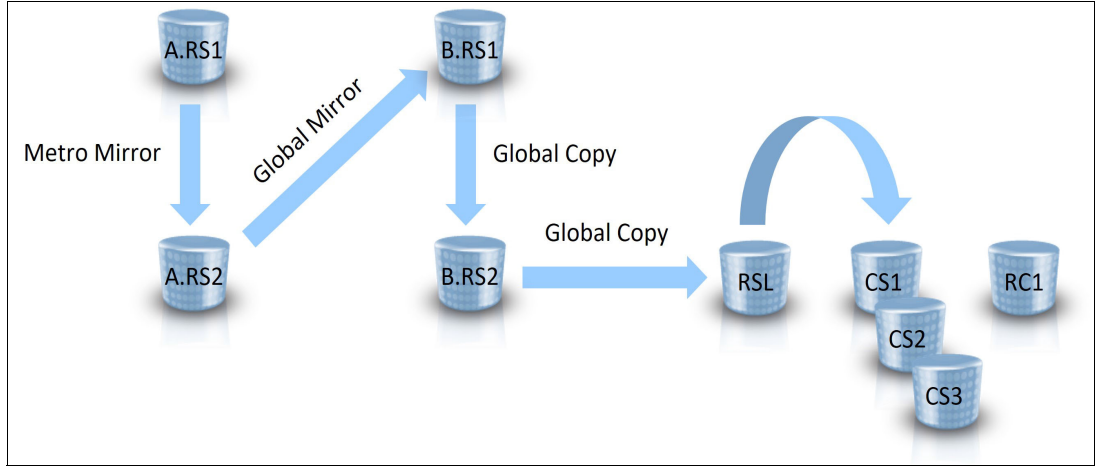

Figure 9-17   External LCP in a GDPS MGM 4-site environment

In Figure 9-17, a standard GDPS MGM 4-site environment exists with Metro Mirror running between the A.RS1 devices and the A.RS2 devices in the production region (referred to as Region A), GM running between A.RS1 in Region A and B.RS1 in the recovery region (referred to as Region B), and Global Copy running between the B.RS1 devices to a second set of devices in Region B, labeled B.RS2. The A.RS1 devices are the devices that the production applications use.

Considering Figure 9-17, to seed the external LCP environment, another Global Copy relationship is used to mirror the data from the B.RS2 devices to the external copy of data, labeled RSL in the figure. The RSL volume set serves as the source for the LCP protection copy captures. The LCP environment then consists of the RSL devices, along with three protection CS labeled CS1, CS2, and CS3 and one RC copy set labeled RC1.

The LCP environment in a GDPS MGM 4-site solution is managed by an instance of GDPS Metro that runs on a separate controlling system from the GDPS MGM controlling system functions and that is configured in a single-leg topology.

For more information about the GDPS Metro offering, see Chapter 3, "IBM GDPS Metro" on page 45. This LCP Manager system is used to maintain Global Copy mirroring to the RSL volume set, in addition to capturing and recovering the LCP protection copies.

Consider the following points when external LCP is implemented in GDPS MGM 4-site environments:

► The LCP environment can be created and maintained from the B.RS2 devices only.

► The protection copies can be FlashCopy CS or SGC CS.

► Capturing an external LCP protection copy requires task coordination between the LCP Manager controlling system and the controlling systems that make up the GDPS MGM 4-site solution. The key coordination task here is creating a consistent data point on the B.RS2 volume set so that it can be replicated to the RSL volume set and then captured to a corresponding copy set. Global Copy is used to maintain the B.RS2 copy and as discussed in 2.4.2, "Global Mirror" on page 26, Global Copy does not provide data consistency.

The following steps are taken when a request is made to capture a protection copy in a GDPS MGM 4-site environment to coordinate consistency across the environment to the RSL volume set and then to capture the consistency group on the appropriate copy set devices:

a. GM is paused on a consistent boundary. This pause results in the suspension of the GM secondary devices (B.RS1) that contain a consistency copy of data and the GM session.

b. The consistent copy of data is then allowed to drain to the B.RS2 volume set and on to the RSL volume set.

c. After the consistent copy of data arrives at RSL, the Global Copy relationship between B.RS2 and RSL is suspended to prevent any more updates to the RSL volume set until after the consistent copy of data is captured to a copy set in a subsequent step.

d. GM is resumed, which allows updates to again flow from Region A to the B.RS1 and B.RS2 volume sets.

e. A protection copy is captured from the RSL volume set to one of the associated CS.

f. The Global Copy relationship between B.RS2 and RSL is resumed to allow updates to again flow to the RSL volume set until the next protection copy is taken. This resumption minimizes the time that it takes to run the capture process each time by keeping the RSL volume set as current as possible.

This method avoids any impact to production applications during the LCP copy capture process at the expense of a typically minor disruption to your RPO.

TCM in GDPS Metro Global - GM 4-site environments

As described in 9.2.3, "Testcopy Manager" on page 259, the TCM feature is similar to the External LCP feature. The key difference being that only one copy set exists in a TCM environment.

Figure 9-18 shows an example of TCM in a GDPS MGM 4-site environment.

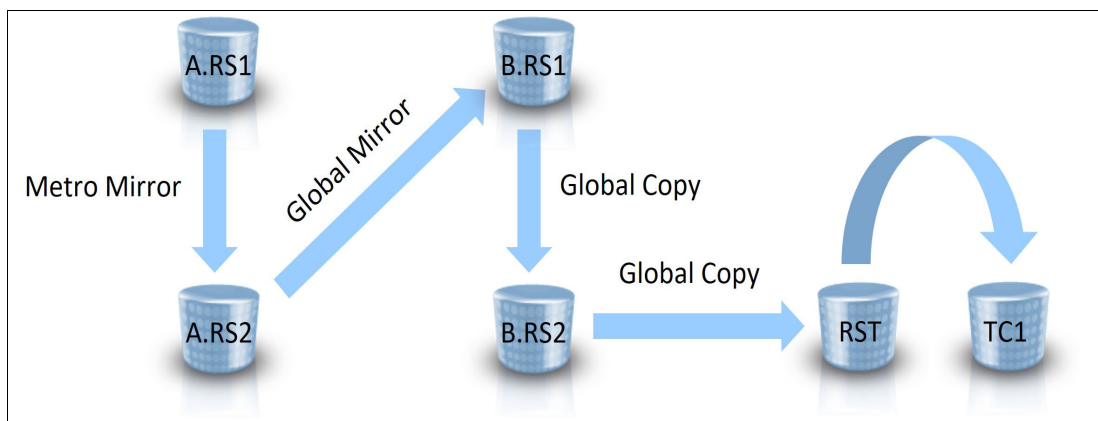*Figure 9-18   Testcopy Manager in a GDPS MGM 4-site environment*

Figure 9-18 looks similar to Figure 9-17 on page 274. A standard GDPS MGM 4-site environment is shown with Global Copy being used to seed the external copy. The external volume set in this case is labeled RST to reflect that it is part of a TCM environment and there is only one copy set, which is labeled TC1. There is no other CS and there is no RC.

As with External LCP, the TCM environment is managed by an instance of GDPS Metro that runs on a separate controlling system from the GDPS MGM controlling system functions and that is configured in a single-leg topology. This TCM system is used to maintain Global Copy mirroring to the RST volume set and to capture the TC1 copy upon request.

In a TCM environment, the TC1 copy set can be a FlashCopy copy set only. SGC is not supported in a TCM environment.

As with external LCP, the TCM environment can be created and maintained from the B.RS2 devices only. Capturing an external TCM test copy requires task coordination between the TCM controlling system and the controlling systems that make up the GDPS MGM 4-site solution. Similar steps are run to capture a test copy as are run to capture a protection copy in an external LCP environment.

Combining external LCP and TCM

As described in 9.2.3, "Testcopy Manager" on page 259, external LCP, and TCM can be implemented in the same GDPS MGM 4-site environment (see Figure 9-19).
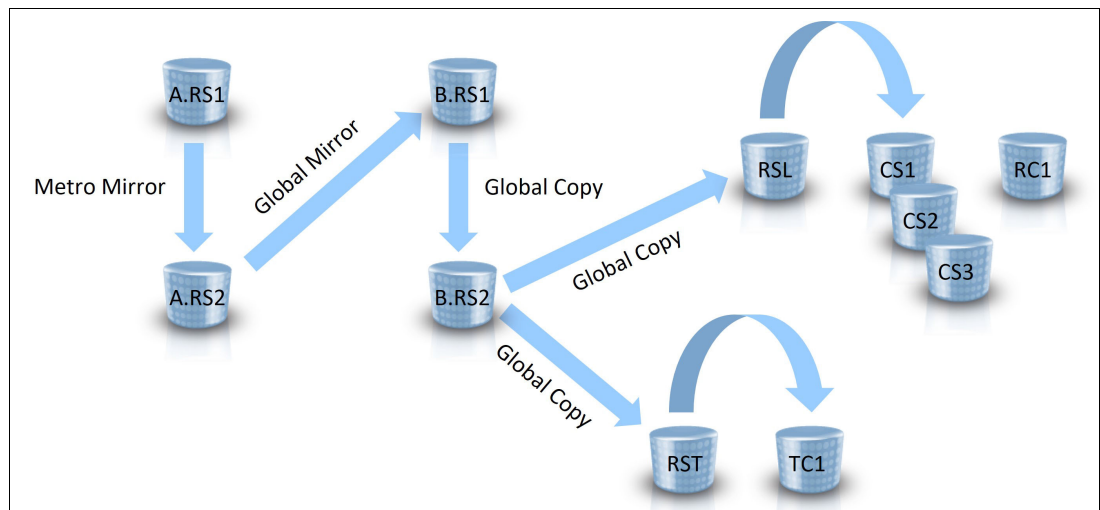
Figure 9-19   External LCP and TCM in a GDPS MGM 4-site environment

Again, Figure 9-19 looks similar to Figure 9-14 on page 271 and Figure 9-18 on page 275 in that a standard GDPS Metro Global Mirror (MGM) 4-site environment is shown. However, in Figure 9-19, *two* Global Copy relationships are used to seed *two* external copies from the B.RS2 volume set: one for the LCP environment and one for the TCM environment.

With this configuration, the external LCP environment and the TCM environment are both managed by a single instance of GDPS Metro that runs on a separate controlling system from the GDPS MGM controlling system functions. However, now this instance of GDPS Metro is configured in a *dual-leg* topology, which allows it to function as the LCP Manager and the TCM Manager.

A capture process in this configuration applies only to one of the external environments. A capture request can capture an LCP protection copy or it can refresh the test copy, but not both at the same time.

9.4 Managing the LCP and TCM environments

This section describes the capabilities that are provided for managing the LCP and TCM environments.

9.4.1 Scripting

Earlier in this book, we described the powerful scripting capability that is provided with most GDPS offerings to enable the automation of complex, multi-step procedures that involve multiple GDPS resources. This scripting capability is extended to enable the automation of the LCP and TCM environments as well.

For LCP environments, script statements are provided to perform the following operations:

- ► Coordinating and securing a consistent copy of data on the RSL volume set (applies to external LCP environments only).

- ► Capturing protection copies from the RSL volume set to the FlashCopy or SGC CS (applies to external LCP environments only).

- ► Capturing protection copies from an RS(n) volume set to the FlashCopy or SGC CS (applies to internal LCP environments only).

- ► Releasing expired FlashCopy and SGC protection copies.

- ► Recovering protection copies from the FlashCopy or SGC CS to an RC copy set.

- ► Capturing a copy from an RS(n) volume set directly to an RC copy set.

  This operation can be useful for taking backups to tape from the RC copy set and for performing validation for early detection of any logical corruption that might occur in the production environment.

- ► Restoring a copy back to production (see ““Restoring data back to production” on page 277”).

- ► Ending (removing) an RC.

Scripting is provided to automate the following operations for TCM environments:

- ► Coordinating and securing a consistent copy of data on the RST volume set
- ► Capturing a test copy from the RST volume set to the test copy volume set
- ► Withdrawing a test copy relationship

Restoring data back to production

GDPS provides scripting capabilities for automating the task of restoring data from the LCP environment back to production. These capabilities apply to the catastrophic use case. In the catastrophic use case, large amounts of data are corrupted, which makes surgical recovery of only the corrupted data infeasible. Therefore, production is assumed to be down while large amounts of data (perhaps the entire environment) are restored.

There are two variations of restoring data back to production, as described next.

Restoring a FlashCopy capture

Scripting can be used to restore a capture that is taken to a FlashCopy copy set. The target of the LCP restore operation is the RS(n) volume set from which the capture was taken. The operation is shown in Figure 9-20.
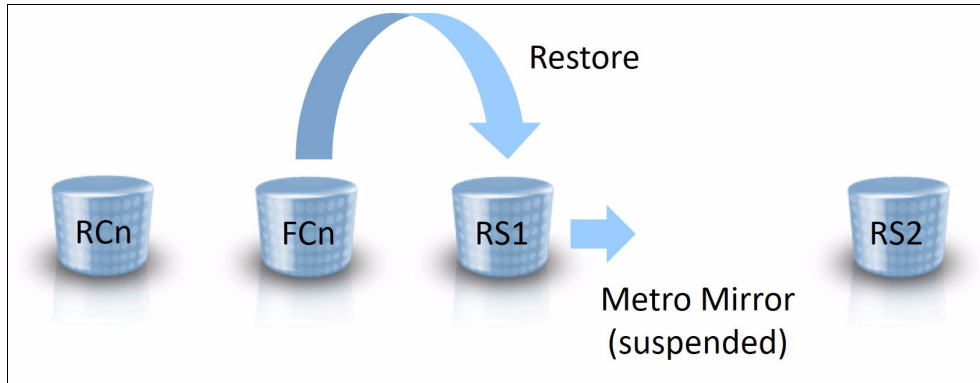

Figure 9-20   FlashCopy restore to production

As shown in Figure 9-20, the RS1 devices are the Metro Mirror primary devices that were used by production before the data corruption event occurred. They are also the devices from which the FlashCopy captures were taken.

The LCP restore operation uses FlashCopy to copy the data back from the FCn copy set to the RS1 device set. The restore can consist of a full push of all data back to the production environment or the restore might be incremental such that only the data that changed since the backup was taken is copied back to production.

Whether the restore is a full push of all data or is incremental depends on the options that are used when the backup was taken.

After the restore is complete, the production systems can be started from the RS1 devices.

The LCP restore operation is rejected if mirroring status is OK. Therefore, it might be necessary to stop Metro Mirror before attempting the LCP restore operation.

When the FlashCopy captures are taken from the Metro Mirror secondary devices (the RS2 device set in Figure 9-20) and are therefore housed in the same disk storage system as the secondary devices, GDPS first converts the secondary devices to suspended primary devices before attempting the FlashCopy restore. After the restore process is complete, the production systems can be started from the RS2 devices.

When restoring a FlashCopy capture, the source of the restore can also be an RC copy set, which is possible only if the FlashCopy capture was taken directly from the RS(n) volume set. That is, if an FlashCopy or SGC capture is recovered to an RC copy set, that capture cannot then be restored back to the RS(n) volume set from the RC copy set.

The ability to capture a copy of production directly to an RC copy set and to restore the copy from the RC copy set back to the RS(n) volume set provides a minimal LCP environment where only one backup copy exists.

Restoring a Safeguarded capture

Scripting can also be used to restore a capture that was taken to an SGC copy set. The target of the LCP restore operation in this case is the device set that is the Metro Mirror peer to the device set from which the SGC captures are taken (see Figure 9-21).
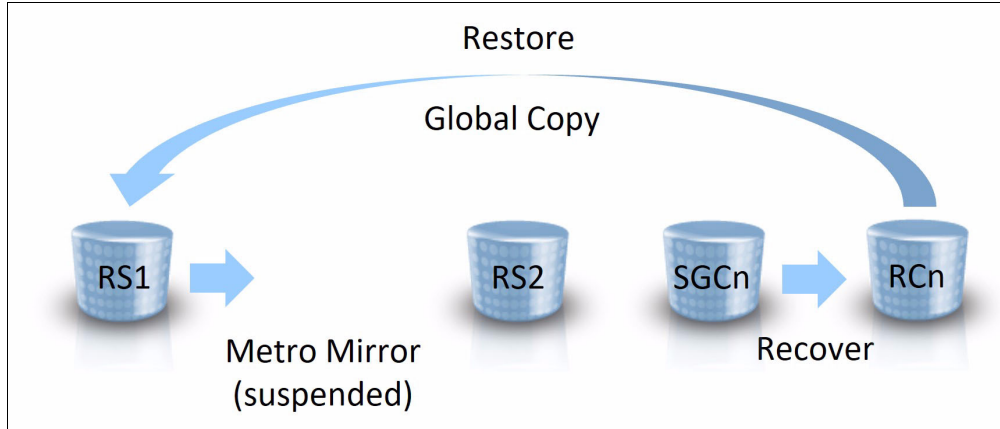

Figure 9-21   Safeguarded Copy restore to production

As shown in Figure 9-21, the RS1 devices are the Metro Mirror primary devices that were used by production before the data corruption event occurred and the RS2 devices are the devices from which the SGC captures were taken. The LCP restore operation consists of first recovering the wanted SGC capture to the RCn device set, and then establishing a Global Copy relationship to incrementally copy the data back to the RS1 device set. After the restore is complete, the production systems can be started from the RS1 devices.

The LCP restore operation is rejected if mirroring status is OK. Therefore, it might be necessary to stop Metro Mirror before attempting the LCP restore operation.

When the SGC captures are taken from the Metro Mirror primary devices (the RS1 device set that is shown in Figure 9-21), the target of the restore is the RS2 device set. Again, the target of the restore of an SGC capture is always the device set that is the Metro Mirror peer to the device set from which the SGC captures are taken. In this case, following the restore, GDPS updates its site indicator to point to the RS2 devices, which results in the RS2 devices being used for the subsequent IPL of the production systems.

Although the restore-to-production scenario that is described in this section is specific to the GDPS Metro single-leg environment, similar functions exist for the other GDPS LCP topologies that are described in 9.3, "LCP operational models" on page 260.

9.4.2 Panels

In the previous sections, we discussed the scripting capability that is provided to manage the various copies in LCP and TCM environments. Because of the simplicity of the TCM environment, little else is required beyond the scripting capability to manage the TCM environment.

However, LCP environments are more complex because they include greater numbers and types of copies and they include management profiles to govern the protection copies. As a result, more tasks are required to manage an LCP environment.

These extra tasks are performed by using the 3270 window interface or the GDPS GUI. This section describes the 3270 window interface. All functions and capabilities that are discussed are also available through the GDPS GUI.

Managing LCP management profiles

When the LCP feature is licensed and configured, option L is provided on the GDPS main window. Selecting option **L** from the GDPS main window displays the management profiles window, as shown in Figure 9-22.

```
VPCPMP00        Logical Corruption Protection Management Profiles

   Actions:   I nfo   C aptures                 S afeGuard   M odify   D el

   MANAGEMENT              CAPTURE      VOLUME   COPY   CAPTURE   EXPIRED
   PROFILE                 TYPE         COUNT    SETS   COUNT     COUNT
   ------------------      ---------    ------   ----   -------   -------
_  PRODUCTI.RS1
_     GOLD_SGC_RS1         SAFEGUARD       12      1        2         2
_     RECOVERY                             12      1        1         0
_     UNASSIGNED           FLASHCOPY       12      3        0         0
_  TEST.RS1
_     RECOVERY                              2      1        0         0
_     SILVER_SGC_RS1       SAFEGUARD        4      1        3         3
_     UNASSIGNED           FLASHCOPY        2      2        0         0




Command ===>  _
F1=Help  F3=Return  F5=Refresh  F6=Roll  F7=Up  F8=Down  F10=Left  F
```

Figure 9-22   VPCPMP00 LCP management profiles window

This window displays all the management profiles that are defined within each consistency group, within each RS. In Figure 9-22, we see the following components:

► Two consistency groups exist in the same RS (RS1). One consistency group is named PRODUCTI and one is named TEST.

► PRODUCTI has one user-defined management profile that is named GOLD_SGC_RS1 and it uses the SGC technology to capture the protection copies. PRODUCTI contains 12 volumes from the RS1 volume set, one SGC copy set, and two captures, both of which are expired.

► TEST has one user-defined management profile that is named SILVER_SGC_RS1 and it also uses the SGC technology to capture the protection copies. TEST contains four volumes from the RS1 volume set, one SGC copy set, and three captures, all of which are expired.

► Each consistency group has three unassigned FlashCopy CS with no active captures and each consistency group also has one RC set (LCP automatically builds internal management profiles for recovery CS and unassigned FlashCopy CS).

Scrolling to the right on this window by using F11 displays the rest of the window, as shown in Figure 9-23.

```
VPCPMP01        Logical Corruption Protection Management Profiles        BZK1

   Actions:   I nfo   C aptures   F lashCopy   S afeGuard   M odify   D elete

   MANAGEMENT              LAST CAPTURE          COPY   RETENTION      FLASHCOPY
   PROFILE                 UTC DATE & TIME       SET    PERIOD         MODE
   ------------------      -------------------   ----   ------------   -----------
_  PRODUCTI.RS1
_     GOLD_SGC_RS1         2020/02/18.11:01:29   001    MINUTE(1)      N/A
_     RECOVERY             2020/03/03.13:13:17   001    N/A            N/A
_     UNASSIGNED                                        N/A            N/A
_  TEST.RS1
_     RECOVERY                                          N/A            N/A
_     SILVER_SGC_RS1       2019/12/17.08:46:56   001    MINUTE(1)      N/A
_     UNASSIGNED                                        N/A            N/A




Command ===>  _                                                    Row 1 of 8
F1=Help  F3=Return  F5=Refresh  F6=Roll  F7=Up  F8=Down  F10=Left  F11=Right
```

Figure 9-23   VPCPMP01 LCP management profiles window continued

This window shows the date and time that the last protection copy was captured for each consistency group and what the retention period is for the protection copies in each of the consistency groups.

Entering **S** against the consistency group name on the LCP management profiles window requests that a new Safeguarded management profile is created and results in the window that is shown in Figure 9-24 being displayed.

```
VPCPMPAS               Add an LCP Management Profile              G9C4

      Capture Type: SAFEGUARD            SafeGuarded capture profile
  Consistency Group: ZPROD91             Consistency Group name
   Replication Site: 2                   Replication site number
 Management Profile: SGCTEST             Profile name
           Copy Set: 1                   Copy set assigned to this profile

   Retention Period: HOUR(5)             Retention period for all captures
  Retention Minimum: 0                   Minimum number of captures to retain
  Automatic Release: NO                  Automatically release captures

   Reservation Time: 0600                Maximum Reservation Scan elapsed time
      Check In Time: 240                 Maximum Check In elapsed time
       CG Pause Time: N/A                Maximum CG Pause elapsed time (GM)

   Minimum Interval: HOUR(4)             Minimum interval between captures
   Monitor Interval: NO                  Enable SGC monitoring

   Enter NO to cancel or YES to proceed with the profile addition

   Selection ===>  YES_
   F1=Help  F3=Return
```

Figure 9-24   VPCPMPAS adding a Safeguard management profile

In this window, the user enters the following information for creating the Safeguard management profile:

Consistency Group The name of the consistency group with which the profile is associated.

Replication Site The RS with which the profile is associated.

Management Profile A unique name identifying the profile.

Copy Set The SGC copy set assigned to the management profile.

Retention Period The amount of time to retain all captures. The retention period can be specified in minutes, hours, or days.

Retention Minimum The minimum number of captures to retain.

Automatic Release Ability to automatically release captures after the retention period expires.

Minimum Interval The minimum capture interval is the amount of time that must elapse before a new capture is permitted. The specified value must be equal to or less than the retention period. The minimum capture interval can be specified in minutes, hours, or days.

The remaining three fields on this window (*Reservation Time, Check In Time,* and *CG Pause Time)* are related to the capturing of an SGC protection copy within this management profile. The process for capturing an SGC protection copy is made up of multiple steps, and these fields specify the amount of time that LCP is allowed to run each step. You have control over how much of an impact the capture process has on your production applications or your disaster recovery (DR) RPO, depending on the specific type of LCP configuration that was implemented and the GDPS solution environment in which it was implemented.

In addition to creating a Safeguard management profile, the following options are available from the LCP management profiles window to manage your management profiles:

Info Displays information about a specific management profile.
Captures Lists all captures that are taken for the management profile.
Modify Modifies a management profile.
Delete Deletes a management profile.

The example that is shown in Figure 9-22 on page 280 and Figure 9-23 on page 281 was taken from a GDPS GM LCP environment, which does not support FlashCopy CS. For environments that support FlashCopy CS, the LCP management profiles window also provides an option to create a FlashCopy management profile.

Displaying information about captures

Entering `C` against the profile name on the LCP management profiles window (as shown in Figure 9-22 on page 280) lists all captures that are assigned to the SafeGuard profile. The SafeGuard Captures panel is shown in Figure 9-25.

```
VPCPLCS0          Logical Corruption Protection SafeGuard Captures          AZK2

Management Profile: GOLD_SGC_RS1                 Capture Type: SAFEGUARD
 Consistency Group: PRODUCTI                    Latest Capture: 2019/11/13.17:19:26
  Replication Site: 1                          Captured/Expired: 2/2

  Actions:  V olumes  T ag Capture  U ntag Capture

  COPY   SEQUENCE  UTC                       VOLUME  CAPTURE   PCD    S2D    CID
  SET    NUMBER    CAPTURE DATE & TIME       COUNT   FLAGS     COUNT  COUNT  COUNT
  ----   --------  -------------------       ------  EITR----  -----  -----  -----
_  001   5DCC3B1E  2019/11/13.17:19:26          18  YNNN....      0      0      0
_  001   5DCC3B01  2019/11/13.17:18:57          18  YNNN....      0      0      0




Command ===>  _                                                     Row 1 of 2
F1=Help  F3=Return              F5=Refresh  F6=Roll  F7=Up  F8=Down
```

Figure 9-25   VPCPLCS0 SafeGuard captures window

The window is divided into two sections. The upper half of the window is static and provides summary statistics for the management profile. This information includes the date and time of the latest capture with the number of captures that are taken and the number of that expired.

The lower, scrollable section of the window includes the following columns:

► Copy Set: The copy set number.

► Sequence Number: The sequence number that is used by the capture.

► Capture Date & Time: Displays the date and time of the capture in Coordinated Universal Time format.

► Volume Count: Indicates the number of volumes that were included in the capture.

► Capture Flags: The following capture flags are available:

  – `Expiration`: Capture the expired flag.

  – `Invalid`: SafeGuard capture was invalidated.

  – `Tagged`: Capture is tagged (for more information about tagging captures, see "Tagging FlashCopy and SGC captures for recovery or restore processing" on page 287).

  – `Recovery`: Capture is in use as the source of an LCP recovery.

► PCD Count: The number of volumes that are flagged `Pending Config Deletion` in this copy set.

► S2D Count: The number of volumes that are flagged `Safe to Delete` in this copy set.

► CID Count: The number of volumes that are flagged `Capture in Doubt` in this copy set.

Regarding the last three fields, `Pending Config Deletion`, `Safe to Delete`, and `Capture in Doubt` are states that copy set volumes can be in during the process of removing the RS volume that is associated with the copy set volumes.

A similar window is available for listing the captures for a user-defined FlashCopy management profile.

LCP internally builds and maintains a management profile that is called UNASSIGNED for each RS. The profile contains all FlashCopy CS that are not yet assigned to a user-defined FlashCopy management profile. Specifying C against the UNASSIGNED profile lists all the unassigned FlashCopy CS in the RS.

The Unassigned Captures window is shown in Figure 9-26.

```
VPCPLCU1         Logical Corruption Protection Unassigned Captures        AZK2

Management Profile: UNASSIGNED                    Capture Type: FLASHCOPY
 Consistency Group: PRODUCTI                     Latest Capture: 2019/11/14.14:14:30
  Replication Site: 1                                  Captured: 1

  Actions:  V olumes

  COPY   SEQUENCE  UTC                       VOLUME  FLASHCOPY
  SET    NUMBER    CAPTURE DATE & TIME       COUNT   MODE
  ----   --------  ------------------        ------  ----------
_  003   5DCD6146  2019/11/14.14:14:30          16   INCREMENTAL




 



 Command ===>  _                                                   Row 1 of 1
 F1=Help  F3=Return              F5=Refresh  F6=Roll  F7=Up  F8=Down
```

Figure 9-26   VPCPLCU1 listing unassigned FlashCopy CS

Because these CS are not under LCP control, the capture flags and the Flag-Set counts are not reported.

LCP internally builds and maintains a management profile that is called RECOVERY to represent all RC CS that were defined for each RS. Specifying C against the RECOVERY profile name lists all recovery CS in the RS.

An example of the Recovery Captures window is shown in Figure 9-27.

```
VPCPLCR1         Logical Corruption Protection Recovery Captures       AZK2

Management Profile: RECOVERY                    Capture Type: RECOVERY
 Consistency Group: PRODUCTI                   Latest Capture: 2019/11/13.17:18:57
   Replication Site: 1                               Captured: 1

   Actions:  V olumes

   COPY  SEQUENCE  UTC                    VOLUME  RECOVERY   RECOVERY      RECOVERY
   SET   NUMBER    CAPTURE DATE & TIME    COUNT   TYPE       MODE          SOURCE
   ----  --------  -------------------    ------  ---------  -----------   --------
_  001   5DCC3B01  2019/11/13.17:18:57       16  SAFEGUARD  NOCOPY        SGC(1)




Command ===>  _                                                    Row 1 of 1
F1=Help  F3=Return              F5=Refresh  F6=Roll  F7=Up  F8=Down
```

Figure 9-27   VPCPLCR1 SafeGuard recovery through the SGC(1) copy set

The window that is shown in Figure 9-27 lists an RC copy set with a Recovery Type of Safeguard, which indicates that it contains an active RC that was captured from a Safeguard copy set.

The recovery type can be one of the following states:

INACTIVE   The RC set is not in use.

DIRECT   The RC set was captured directly from an RS(n) volume.

SAFEGUARD   The RC set was captured by using a Safeguard capture operation.

FlashCopy   The RC set was captured by using a FlashCopy capture operation.

Because these CS are not under LCP control, the capture flags and the Flag-Set counts are not reported.

Displaying information about capture copy set volumes

Entering V against any type of capture lists all volumes that are configured to the associated copy set and their capture status. The Copy Set Volumes window for an SGC capture is shown in Figure 9-28.

```
VPCPLCSS            Logical Corruption Protection Copy Set Volumes         AZK2

Management Profile: GOLD_SGC_RS1                    Copy Set: 1
 Consistency Group: PRODUCTI                    Capture Time: 2019/11/13.17:18:57
  Replication Site: 1                                Expired: YES

  Actions:  Q uery  S et  R eset


                     SAFEGUARD       SEQUENCE  SAFEGUARD  BACKUP    LCP
    UCB    VOLSER   SOURCE           NUMBER    STATE      TRACKS    FLAGS
    -----  ------   -------------    --------  SRCCIR---  --------  PSC---
 _  05071  MM5071   00BAZ11 03.01    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05072  MM5072   00BAZ11 03.02    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05073  MM5073   00BAZ11 03.03    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05074  MM5074   00BAZ11 03.04    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05075  MM5075   00BAZ11 03.05    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05076  MM5076   00BAZ11 03.06    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05077  MM5077   00BAZ11 03.07    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05078  MM5078   00BAZ11 03.08    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  05079  MM5079   00BAZ11 03.09    5DCC3B01  Y..Y.N...  00000000  NNN...
 _  0507A  MM507A   00BAZ11 03.0A    5DCC3B01  Y..Y.N...  00000000  NNN...


Command ===>  _                                              Row 1 of 18
F1=Help  F3=Return              F5=Refresh  F6=Roll  F7=Up  F8=Down
```

Figure 9-28   VPCPLCSS listing Safeguard capture volumes

The LCP Copy Set Volumes window is divided into two sections. The upper half of the window is static and provides summary statistics for the copy set, which includes the date and time of the capture and whether the capture is expired.

The scrollable section of the window contains the following columns:

► UCB: Indicates the UCB device number.

► Volser: The volume serial number.

► Safeguard Source: The serial number, logical subsystem (LSS), and channel connection address for the SafeGuard protected source.

► Sequence Number: The 8-byte sequence number that is used for the LCP capture.

► Safeguard State: The possible values for this field are:

   – SafeGuarded: Y/N - the source volume is in a SafeGuarded relationship
   – Reservation: In Progress or Complete
   – Check In: Is Active - Long Busy
   – Captured: Y - SafeGuard protected volume captured
   – Invalidated: Y - SafeGuard capture has been invalidated
   – Recovery: Y/N - volume is participating in a SafeGuarded recovery

► Backup tracks: The number of tracks that are stored on the SafeGuarded Virtual Backup Capacity.

► LCP Flags: The following flags are available:

   – P: Y/N - Pending Config Deletion
   – S: Y/N - Safe to Delete
   – C: Y/N - Capture in Doubt

Tagging FlashCopy and SGC captures for recovery or restore processing

Before running a script to restore or recover a copy set, the copy set that is to be restored or recovered must be tagged. A capture can be tagged for restore or recover by entering a **T** against it on the Safeguard (or FlashCopy) Captures window. Figure 9-29 shows the same Safeguard Captures window as the window that is shown in Figure 9-25 on page 283; however, this time, one of the SGC captures is tagged for recover as indicated by the T capture flag with a value of Y.

```
 Management Profile: GOLD_SGC_RS1              Capture Type: SAFEGUARD
  Consistency Group: PRODUCTI                 Latest Capture: 2019/11/13.17:19:26
  Replication Site: 1                        Captured/Expired: 2/2

   Actions:  V olumes  T ag Capture  U ntag Capture

   COPY  SEQUENCE  UTC                      VOLUME  CAPTURE    PCD    S2D    CID
   SET   NUMBER    CAPTURE DATE & TIME      COUNT   FLAGS     COUNT  COUNT  COUNT
   ----  --------  -------------------      ------  EITR----  -----  -----  -----
 _  001  5DCC3B1E  2019/11/13.17:19:26        18   YNNN....     0      0      0
 _  001  5DCC3B01  2019/11/13.17:18:57        18   YNYN....     0      0      0




 Sequence number 5DCC3B01 in copy set SGC(1) is now tagged
 Command ===>  _                                              Row 1 of 2
 F1=Help  F3=Return            F5=Refresh  F6=Roll  F7=Up  F8=Down
```

Figure 9-29   VPCPLCS0 Safeguard capture is tagged

A capture can be untagged by entering a **U** against it on the Safeguard (or FlashCopy) Captures window.

A capture must be tagged, which identifies the source for the recovery operation. A recovery operation for a Safeguarded Management Profile without a tagged capture automatically tags the last closed valid Safeguarded capture.

> **Note:** The latest capture is always the open capture and is not eligible for automatic tagging. The use of automatic tagging for Safeguarded captures requires that at least two captures exist. A capture that is flagged as invalid is not eligible for tagging. In this case, automatic tagging evaluates (and where possible) tags the next capture.

9.4.3 Securing the GDPS LCP environment

GDPS secures your LCP environment by using multiple security models: role elevation, which is implemented through role-based security, and dual control (also referred to as maker-checker).

Role-based security

GDPS uses RACF XFACILIT resource classes to create a role-based security model for controlling access to the resources in your GDPS LCP environment. The security model is customized to your specific environment.

Simple definitions can be used to control access to LCP panels, options, and actions that are available on LCP panels, and even specific resources within the LCP environment (such as management profiles). More granular definitions can be used to control access to specific fields that are associated with a management profile, types of resources, or even all the way down to the specific resource level.

With the role-based security model, you can create your own roles or use the common roles that GDPS recommends, which include GDPS Administrator, GDPS Operator, GDPS User, and Non-GDPS User. You define the resources that these roles can access and the type of access that they have to those resources by granting them access to the resource profiles that represent the various resources in your environment. You can grant access to various resources to users by adding them to the suitable roles.

When you use the role-based security model, GDPS ensures that the user has sufficient authority to take a specific action against a specific resource, regardless of whether they are attempting to act by using the panels directly or by running a GDPS script.

Dual control

Although you can control which users may perform various actions in your LCP environment, some actions can be pervasive or destructive enough that you might want to require multiple users to cooperate to perform them. This approach is known as *dual control*. GDPS can define and enforce a dual control policy for the most pervasive actions that can be taken by users in an LCP environment. Specifically, you can prevent a single user from taking specific actions against your management without additional approval from a second user.

As with role-based security, RACF XFACILIT resource classes are used to define the dual control policy. When a user attempts to take an action that is protected by your dual control policy, a ticket is automatically opened for the action. Another user that was granted approver authority can run immediately the action that is requested by the ticket; approve the ticket so that the user that opened the ticket may proceed with the action; or reject the action that is requested by the ticket.

9.5 Monitoring

LCP provides both active and passive monitoring of the storage system resources in the LCP environment. The LCP manager uses an active monitoring capability to query the DS8000 storage system regularly to determine the state of the Safeguarded backup capacity virtual space and the number of captures that were internally rolled off due to out-of-space conditions. Panels display the monitoring activity and control the monitoring intervals.

The LCP manager uses passive monitoring by notifying the DS8000 of the LCP resources (storage pools, SGC virtual capacity, and others) within the DS8000 that it is responsible for managing. Then, the DS8000 notifies the LCP manager when out-of-space conditions occur that are related to the resources it is managing so that the user can be notified and act to rectify the situation.

9.6 Summary

LCP provides a GDPS managed Continuous Data Protection capability and is aimed at helping clients to recover from logical corruption events, whether caused by internal or cyberattacks, or by application or user error. It enables GDPS to regularly create one or more LCP copies of data by using the FlashCopy technology or the SGC technology. These copies can then be used for testing purposes, analysis, and recovery.

Internal LCP is self-contained within the GDPS Controlling systems; all actions are performed from the GDPS Controlling systems. Internal LCP copies are in the same disk storage systems as the production devices with which they are associated.

External LCP extends the internal LCP functions by allowing clients to isolate these LCP volumes from the production environment by the inclusion of a Global Copy or GM relationship that is cascaded from the GDPS environment (the protection copies for LCP are then taken from these Global Copy target devices). External LCP requires another GDPS controlling system, called the LCP Manager Controlling system, to coordinate its activities.

TCM is a function of external LCP that allows clients to manage, capture, and refresh a test copy for use within an isolated test environment. This new isolated copy is a FlashCopy that is taken from a Global Copy secondary volume set that is cascaded from a GDPS MGM 4-site environment.

With external LCP (including TCM), the LCP copies are in different disk storage systems than the production devices with which they are associated.

10

Sample continuous availability and disaster recovery scenarios

In this chapter, we describe several common customer scenarios and requirements, and what we believe to be the most suitable solution for each case.

The following scenarios are described:

► A customer with a single data center that implemented IBM Parallel Sysplex with data sharing and workload balancing wants to move to the next level of availability.

► A customer with two centers needs a disaster recovery (DR) capability that permits application restart in the remote site following a disaster.

► A customer with two sites (but all production systems running in the primary site) needs a proven DR capability *and* a near-continuous availability (CA) solution.

► A customer with two sites at continental distance must provide a DR capability.

► A customer with two sites at relatively long metropolitan distance must provide local CA and remote DR with zero data loss (ZDL).

► A customer who runs only IBM z/VM with Linux on IBM Z guests (no z/OS in their environment) with two sites at metropolitan distance require an automated DR and near CA solution.

The scenarios that are described in this chapter pertain to the use of the IBM Geographically Dispersed Parallel Sysplex (GDPS) products that are based on hardware disk replication.

The scenarios for GDPS CA that use data replication software are described in Chapter 6, "IBM GDPS Continuous Availability solution" on page 179.

This chapter includes the following topics:

10.1 Introduction

In the following sections, we describe how the various GDPS service offerings can address different CA and DR requirements. Because every business is unique, the following sections do not completely list all the ways that the offerings can address the specific needs of your business. However, they do serve to illustrate key capabilities.

In the figures that are included in this chapter, we show minimal configurations for clarity. Many customer configurations are more complex than the examples that are presented here, but both configurations are supported.

10.2 Continuous availability in a single data center

In the first scenario, the customer has only one data center, but wants to have higher availability. The customer-implemented data sharing for their critical applications, and uses dynamic workload balancing to mask the effect of outages.

They mirror all their disks within the same site but must take planned outages when they want to switch from the primary to secondary volumes in preparation for a disk subsystem upgrade or application of a disruptive microcode fix. They are concerned that their disk is their only remaining resource whose failure can take down all their applications.
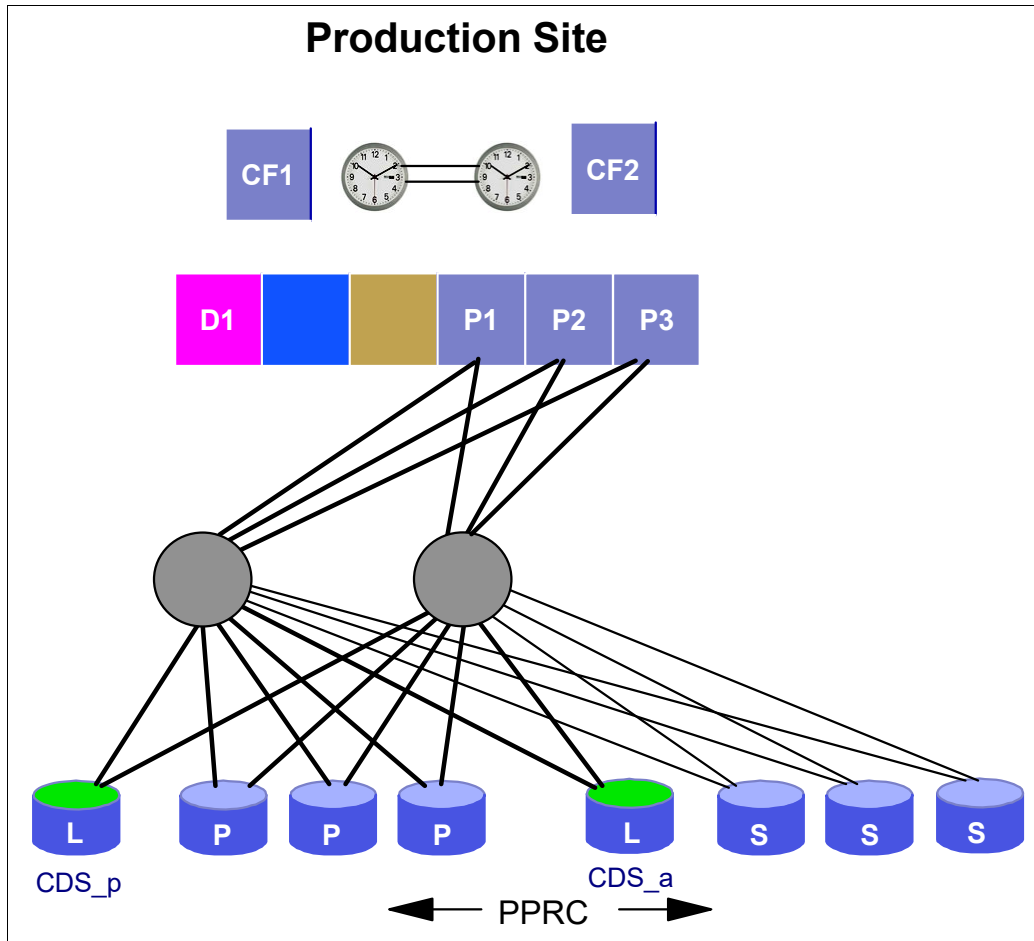
Figure 10-1   Data sharing, workload balancing, and mirroring: Single site

From a DR perspective, the customer relies on full volume dumps. Finding a window of time that is long enough to create a consistent set of backups is becoming a challenge. In the future, they plan to use a second data center to protect them if disaster occurs.

In the interim, they want to investigate the use of FlashCopy to create a consistent set of volumes that they can then dump in parallel with their batch work. But, their current focus is on improving resiliency within their single center.

The customer's situation and requirements are listed in Table 10-1. Also listed are which of those requirements can be addressed by the most suitable GDPS offering for this customer's requirements, namely GDPS Metro HyperSwap Manager.

Table 10-1   Mapping customer requirements to GDPS Metro HyperSwap Manager attributes

Attribute	Supported by GDPS HM
Single site	Y
Synchronous remote copy support	Y (Metro Mirror)
Transparent swap to secondary disks	Y (HyperSwap)
Ability to create a set of consistent tape backups	Y[a]
Ability to easily move to GDPS Metro in the future	Y

a. To create a consistent source of volumes for the FlashCopy in GDPS Metro HyperSwap Manager, you must create a freeze-inducing event and be running with a Freeze and Go policy.

This customer has a primary short-term objective to provide near-CA, but wants to ensure that they address that objective in a strategic way.

In the near term, they need the ability to transparently swap to their secondary devices if a planned or unplanned disk outage occurs. Because they have only a single site, do not have a TS7700, and do not have the time to fully implement the GDPS system and resource management, the full GDPS Metro offering is more than they need.

By implementing GDPS Metro HyperSwap Manager, they can achieve their near-term objectives in a manner that positions them for a move to full GDPS Metro in the future.

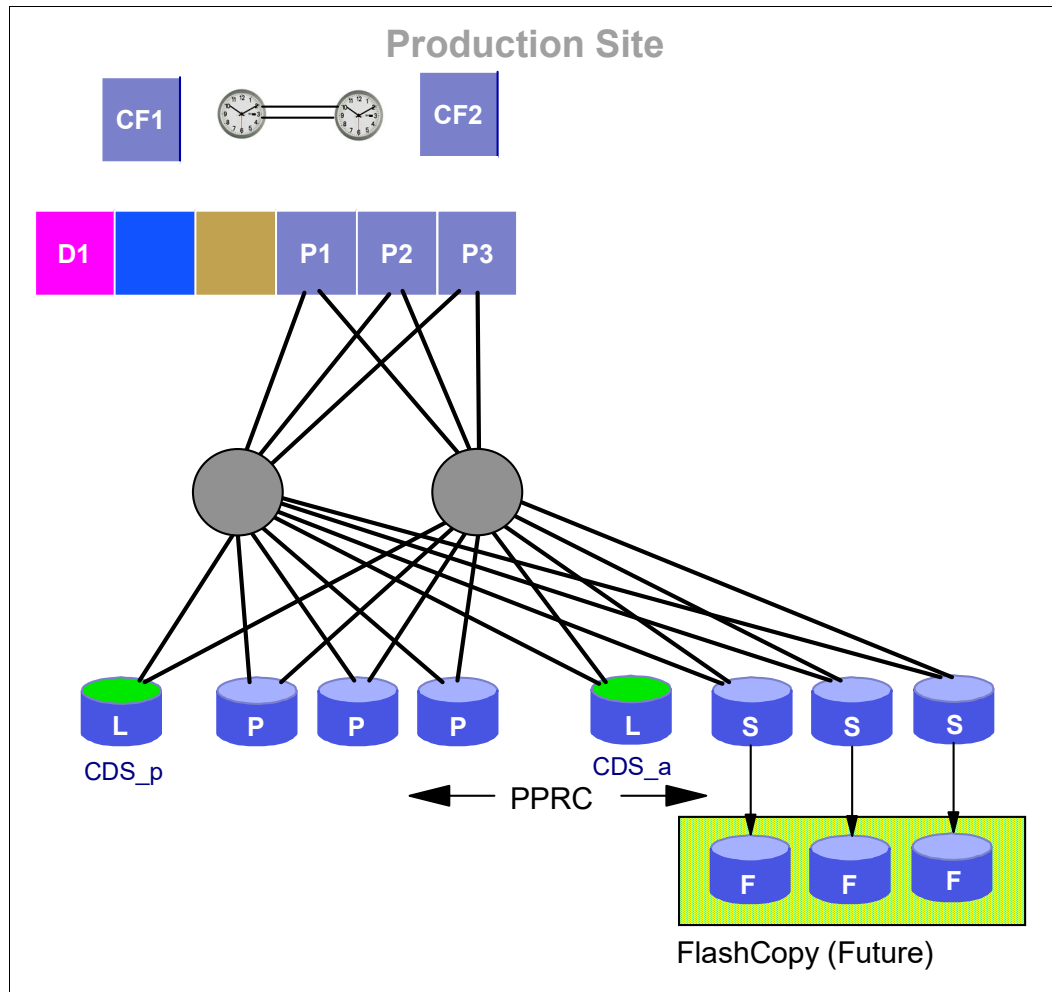Figure 10-2 shows the customer configuration after GDPS Metro HyperSwap Manager is implemented.

Figure 10-2   CA within a single data center

Now, if a failure occurs on the primary disk subsystem, the controlling system starts a HyperSwap in which it transparently switches all systems in the GDPS sysplex over to what were the secondary volumes. The darker lines that are shown in Figure 10-2 that connect the secondary volumes indicate that the processor-to-control unit channel capacity is now similar to what is used for the primary volumes.

After the customer implements GDPS and enabled the HyperSwap function, their next move is to install the extra disk capacity so it can use FlashCopy. The customer then can use the Freeze function to create a consistent view that can be flash-copied to create a set of volumes that can then be full-volume dumped for DR. This configuration creates a more consistent set of backup tapes than the customer has today (because today it is backing up a running system). Also, the backup window now is only a few seconds rather than the hours it that it currently takes. This change enables the customer to make more frequent backups.

10.3 DR across two data centers at metro distance

The next scenario relates to a customer that is under pressure to provide a DR capability in a short time, perhaps for regulatory reasons. The customer has a second data center within metropolitan distances and suitable for synchronous mirroring, but has not yet implemented mirroring between the sites.

Before moving to a full GDPS Metro environment, the customer was going to complete their project to implement data sharing and workload balancing. However, events overtook them and they now need to provide the DR capability sooner than they expected.

The customer can select between the full GDPS Metro offering, as they planned to do in the long term, or to install GDPS Metro HyperSwap Manager now. Because they will not use the extra capabilities that are delivered by GDPS Metro in the immediate future, the customer decides to implement the lower-cost GDPS Metro HyperSwap Manager option.

Table 10-2 summarizes the customer's situation and requirements. It also shows how those requirements can be addressed by GDPS Metro HyperSwap Manager.

Table 10-2   Mapping customer requirements to GDPS Metro attributes

Attribute	Supported by GDPS HM
Two sites, 12 km apart	Y
Synchronous remote copy support	Y (Metro Mirror)
Maintain consistency of secondary volumes	Y (Freeze)
Maintain consistency of secondary volumes during Metro Mirror resync	Y[a] (FlashCopy)
Ability to move to GDPS Metro in the future	Y

  a. FlashCopy is used to create a consistent set of secondary volumes before a resynchronization, following a suspension of remote copy sessions.

This customer must quickly provide a DR capability. Therefore, the primary focus in the near term is to restart its systems at the remote site as though it was restarting off the primary disks after a power failure.

However, the recovery time objective (RTO), which is the time to get the systems up and running again in the remote site, is reduced in the longer term to the point that it cannot be achieved without automation (this issue is addressed by a move to GDPS Metro). The customer also requires a consistent restart point *always* (even during DR testing).

This customer implements GDPS Metro HyperSwap Manager, with the controlling system in the primary site and the secondary disks in the remote site. The auxiliary storage subsystems are configured with sufficient capacity to use FlashCopy for the secondary devices. This configuration allows the customer to run DR tests without affecting its mirroring configuration.

GDPS Metro HyperSwap Manager is installed and the Freeze capability enabled. After the Freeze capability is enabled and tested, the customer installs the extra intersite channel bandwidth that is required to HyperSwap between the sites, as shown in Figure 10-3. Later, in preparation for a move to full GDPS Metro, the customer moves the controlling system (and its disks) to the remote site.
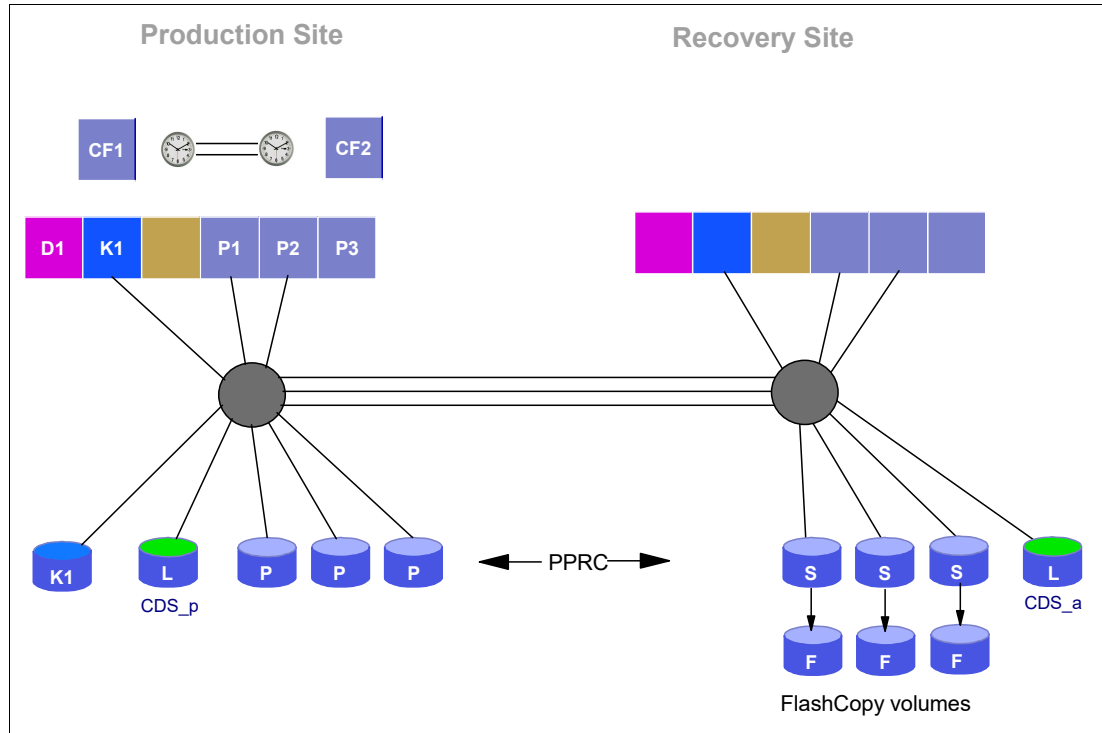

Figure 10-3   GDPS Metro HyperSwap Manager 2-site configuration

10.4 DR and continuous availability across two data centers at metro distance

The customer in this scenario has two centers within metro distance of each other. The customer uses Metro Mirror to remote copy the primary disks to the second site. They also have the infrastructure in place for a cross-site sysplex; however, all production work still runs in the systems in the primary site.

The customer is implementing data sharing, along with dynamic workload balancing, across their production applications. In parallel with the completion of this project, they want to start looking at how the two sites and their current infrastructure can be maximized to provide DR and continuous or near-CA in planned and unplanned outage situations, including the ability to dynamically switch the primary disks back and forth between the two sites.

Because the customer uses remote mirroring, their first priority is to ensure that the secondary disks provide the consistency to allow restart (rather than recovery) if a disaster occurs.

Because of pressure from their business, the customer wants to move to a ZDL configuration as quickly as possible. They also want to investigate other ways to reduce the time that is required to recover from a disaster.

After the DR capability is tested and tuned, the customer's next area of focus is CA across planned and unplanned outages of applications, systems, and complete sites.

This customer also is investigating the use of z/VM and Linux on IBM Z to consolidate several of their thousands of PC servers onto the mainframe. However, this priority is lower than their other tasks.

Because of the DR and CA requirements of this customer, together with the work they did and the infrastructure in place, the GDPS offering for them is GDPS Metro in a single leg configuration. Table 10-3 shows how this offering addresses this customer's needs.

Table 10-3   Mapping customer requirements to GDPS Metro attributes

Attribute	Supported by GDPS Metro
Two sites, 9 km apart	Y
ZDL	Y (Metro Mirror with Freeze policy of `SWAP,STOP`)
Maintain consistency of secondary volumes	Y (Freeze)
Maintain consistency of secondary volumes during Metro Mirror resync	Y[a] (FlashCopy)
Ability to conduct DR tests without impacting DR readiness	Y (FlashCopy)
Automated recovery of disks and systems after a disaster	Y(GDPS script support)
Ability to transparently swap z/OS disks between sites transparently	Y (HyperSwap)
DR and CA support for Linux guests under z/VM	Y

a. FlashCopy is used to create a consistent set of secondary volumes before a resynchronization, following a suspension of remote copy sessions.

Although this customer performed a significant amount of useful work already, fully benefiting from the capabilities of GDPS Metro takes a significant amount of time. Therefore, the project is divided into the following steps:

1. Install GDPS Metro in a single leg configuration, define the remote copy configuration to GDPS, and start the use of GDPS to manage and monitor the configuration.

   This step makes it simpler to implement changes to the remote copy configuration. Rather than issuing many Metro Mirror commands, the GDPS configuration definition must be updated and activated, and the GDPS panels are then used to start the new remote copy sessions.

   Similarly, any errors in the remote copy configuration are brought to the operator's attention by using the NetView Status Display Facility (SDF) facility. Changes to the configuration to stop or restart sessions, or to start a FlashCopy, are simpler by using the NetView interface.

2. After the staff becomes familiar with the remote copy management facilities of GDPS Metro, enable the Freeze capability, initially as `PPRCFAILURE=GO` and then, moving to `PPRCFAILURE=COND` or `STOP` when the customer is confident with the stability of the remote copy infrastructure.

   Because HyperSwap is not implemented immediately, they specify a `PRIMARYFAILURE=STOP` policy to avoid data loss if recovery on the secondary disks becomes necessary after a primary disk problem.

Although the customer has Metro Mirror today, they do not have the consistency on the remote disks that is required to perform a restart rather than a recovery after a disaster. The GDPS Freeze capability adds this consistency and enhances it with the ability to ensure ZDL following a disaster when a `PPRCFAILURE=STOP` policy is implemented.

3. Implement GDPS Sysplex Resource Management to manage the sysplex resources within the GDPS, and start the use of the GDPS Standard actions windows.

   GDPS system and sysplex management capabilities are an important aspect of GDPS. They ensure that all changes to the configuration conform to previously prepared and tested rules, and that everyone can check at any time to see the current configuration; that is, which sysplex data sets and IPL volumes are in use. These capabilities provide the logical equivalent of the whiteboard that is used in many computer rooms to track this type of information.

4. Implement the GDPS Planned and Unplanned scripts to drive down the RTO following a disaster.

   The GDPS scripting capability is key to recovering the systems in the shortest possible time after a disaster. Scripts run at machine speeds, rather than at human speeds. They can be tested over and over until they do precisely what you require. They also always behave in the same way, which provides a level of consistency that is not possible when relying on humans.

   However, the scripts are not limited to DR. This customer sometimes has outages as a result of planned maintenance to its primary site. By using the scripts, they can use HyperSwap to keep its applications available as it moves its systems one by one to the recovery site in preparation for site maintenance, and then back to the normal locations after maintenance is complete.

   Because all production applications still run in the production site now, the processor in the second site is smaller. However, to enable more capacity to quickly be made available if a disaster occurs, the processor has the Capacity BackUp (CBU) feature installed. The GDPS scripts can be used to automatically enable the extra CBU engines as part of the process of moving the production systems to the recovery processor.

   After the DR aspect is addressed, HyperSwap is implemented to provide a near-CA capability for the z/OS systems. A controlling system must be set up in each site when HyperSwap is used to ensure that a system is always available to start a HyperSwap, regardless of where the primary disks might be then.

   The customer uses planned HyperSwap (to move their primary disks before planned maintenance on the primary subsystems) and unplanned HyperSwap (which allows the customer to continue processing across a primary subsystem failure). They test planned HyperSwap while their Primary Failure policy option is still set to `STOP`. However, when they are comfortable and ready, they change to running with a `PRIMARYFAILURE=SWAP,STOP` policy to enable unplanned HyperSwap.

5. Assuming that the consolidation onto Linux on IBM Z proceeded, the heterogeneous DR capability is implemented to manage z/VM systems and its guests and to add planned and unplanned HyperSwap support for z/VM and the Linux guests.

   z/VM systems that are hosting Linux guests who sue CKD disks are placed under GDPS xDR control. This control provides them with near-equivalent management to what is provided for z/OS systems in the sysplex, including planned and unplanned HyperSwap.

   Because it is all managed by the same GDPS, the swap can be started as a result of a problem on a z/OS disk; that is, you do not have to wait for the problem to spread to the Linux disks before the swap is started. Equally, a problem on a CKD Linux disk can result in a HyperSwap of the Linux disks and the z/OS disks.

The projected final configuration is shown in Figure 10-4 (for the sake of clarity, the Linux components are not included).
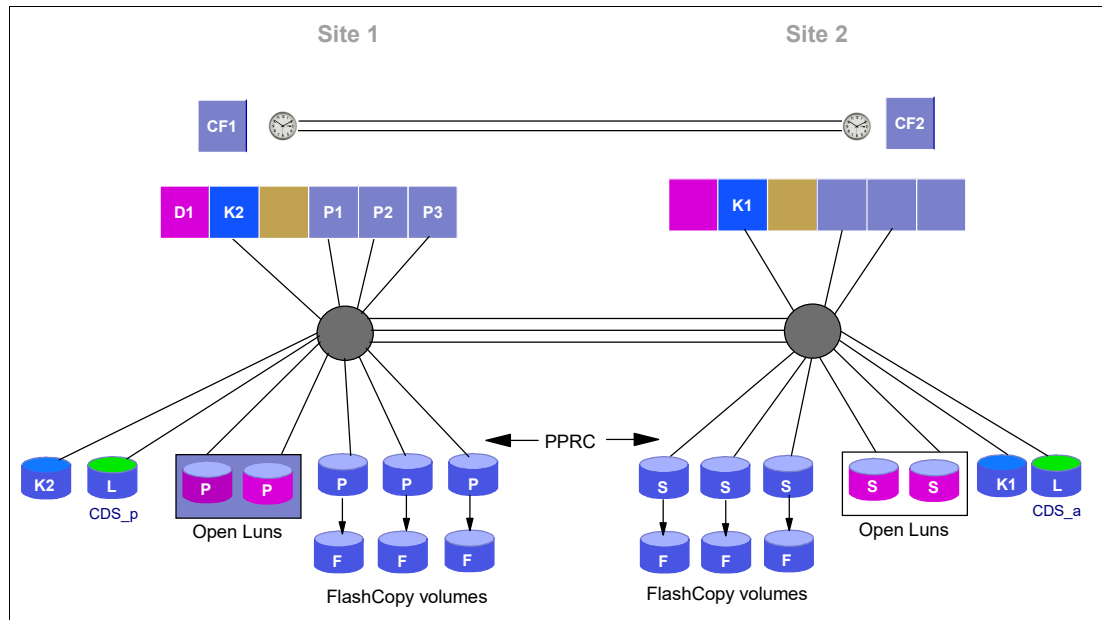

Figure 10-4   Single-site workload GDPS Metro single leg configuration

10.4.1 Multi-site workload

This customer is in the process of enabling all its applications for data sharing and dynamic workload balancing. This project proceeds in parallel with the GDPS project. When the critical applications are enabled for data sharing, the customer plans to move to a multi-site workload configuration, with several production systems in the primary site and others in the recovery site.

To derive the maximum benefit from this configuration, it most likely is possible to transparently swap from the primary to secondary disks. Therefore, it is expected that the move to a multi-site workload does not occur until after HyperSwap is enabled.

The combination of multi-site data sharing and HyperSwap means that the customer's applications remain available across outages that affect a software subsystem (Db2, for example), an operating system, processor, coupling facility (CF), or disk subsystem (primary or secondary). The only event that can potentially result in a temporary application outage is an instantaneous outage of all resources in the primary site, which can result in the database managers in the recovery site having to be restarted.

The move to a multi-site workload might require creating minor changes to the GDPS definitions, several new GDPS scripts, and modifications to existing ones, depending on whether new systems are added or some of the existing systems are moved to the other site. However, apart from that issue, no fundamental change is needed in the way GDPS is set up or operated.

10.5 DR and continuous availability across two data centers at metro distance for z/VM and Linux on IBM Z

The customer in this scenario runs their main production work on Linux on IBM Z, which run as z/VM guests. The production data is on CKD disks. The critical workloads are running on four z/VM systems. Two of the z/VM systems run in one site and the other two in the other site.

They also have a couple of other, less important production z/VM systems that are running Linux guests. The IBM Z server in each site is configured with IFL engines only (no general-purpose CPs), and the customer has no z/OS systems or skills. They have two centers within metro distance of each other. The customer uses Metro Mirror to remote copy the primary disks to the second site. They also have the infrastructure and connectivity in place for the Single System Image (SSI) cluster.

The disk environment is well-structured. Although the various z/VM systems share a physical disk subsystem, the disks for each of the z/VM systems are isolated at a logical subsystem (LSS) level.

Because the customer uses remote mirroring, their first priority is to ensure that the secondary disks provide the consistency to allow restart if a disaster occurs, rather than recovery. Because of pressure from their business, the customer wants to investigate ways to reduce the time that is required to recover from a disaster.

Regulatory pressures also exist that force the customer to periodically demonstrate that they can run their production workload in either site for an extended period. Therefore, they must also use processes to perform planned workload moves between sites as automatically and as fast as possible with minimum operator intervention.

Because of the DR and CA requirements of this customer, together with the work that they did and the infrastructure that is in place, the GDPS offering for them is the GDPS Virtual Appliance. Table 10-4 shows how this offering addresses this customer's needs.

Table 10-4   Mapping customer requirements to GDPS Metro attributes

Attribute	Supported by GDPS Virtual Appliance
Two sites, 9 km apart	Y
Maintain consistency of secondary volumes	Y (Freeze)
Automated recovery of disks and systems after a disaster	Y(GDPS script support)
Ability to transparently swap z/VM (and guest) disks between sites transparently	Y (HyperSwap)
DR and CA support for Linux guests under z/VM	Y
Ability to automate planned move of systems between sites	Y (Script support)
z/OS skills not required	Y

Although this customer performed a significant amount of useful work already and are fully benefiting from the capabilities of the GDPS Virtual Appliance, they are concerned about enabling appliance management for their entire production environment all at once. Because they have their disks isolated in separate LSSs for the SSI and the stand-alone z/VM systems, the following phasing-in of the function is possible:

1. Install a general-purpose CP engine on the Site2 IBM Z server to run the GDPS Virtual Appliance[1].

2. Install the GDPS Virtual Appliance to initially manage one of the stand-alone z/VM systems and the data for this system. Start with the least critical system.

   Define the remote copy configuration to GDPS, and start by using GDPS to manage and monitor the configuration for the first z/VM system.

   In this limited implementation, the customer can test all aspects of the GDPS Virtual Appliance, which is isolated from their more important systems. They can code and test scripts, exercise Freeze, planned and unplanned HyperSwap, refine their operational procedures, and prepare for cutover of their more important z/VM systems.

3. After the staff becomes familiar with the appliance, the customer can then put the second z/VM system and the disks of this system under appliance management. They can perform more tests in this environment to understand how the appliance works when multiple systems are under its control and make final preparations for moving the SSI environment to be under appliance control.

4. The customer adds the 4-way SSI into the appliance-managed environment, perform some more tests, and finalize their implementation.

5. After all systems are under GDPS control, the customer can schedule a test to move their workload to all run in Site2 by using the Site2 disks. Primary disk role is swapped to Site2 by using planned HyperSwap, which makes the move transparent to the systems that were running in Site2.

   Metro Mirror is reversed to run from Site2 disks toward the Site1 disks to retain unplanned HyperSwap capability while the workload is running in Site2. The systems that are running in Site1 are stopped and restarted in Site2 after the disks are swapped. A single planned action script is used to perform this move, which minimizes operator intervention and the time that is required to run the entire process.

   Similarly, a planned action script is used to move the systems back to their "normal" locations.

   The first time that the customer conducts this exercise, they isolate to run production in Site2 over a weekend period, returning to normal before Monday morning. However, by using the same process and scripts, they eventually schedule moves in which they remain in Site2 for a longer period.

---

[1] The option to purchase a IBM Z general-purpose CP engine for customers that require one is included in the GDPS Virtual Appliance deal.

10.6 Local continuous availability and remote disaster recovery across two data centers at a long metropolitan distance

In this scenario, the customer has two data centers (Site1 and Site2) at 100 km (62.1 miles) distance. They run all their systems in Site1 and Site2 is their DR location.

They use Metro Mirror to replicate their data to Site2 and implement the established GDPS/PPRC solution to manage the environment. They use GDPS/PPRC with a Freeze and Stop policy because they have a requirement for ZDL (RPO=0). However, they enabled this environment for unplanned swaps because of the long distance between the sites.

Also, because they do not have sufficient cross-site channel bandwidth between the sites, they cannot run production with their systems in Site1 by using the disks in Site2. The reason that they have HyperSwap enabled is so they can gracefully shut down their systems. After the systems are shut down, they move production to Site2.

The customer has many mirrored devices and defines their Metro Mirror secondary devices in an alternative subchannel that is set to mitigate their unit control block (UCB) constraint. They have FlashCopy devices in Site2, which they use for periodic DR validation testing.

The fact that they cannot fully benefit from HyperSwap means that disk failure is a single point of failure for their sysplex; therefore, they must start DR for a disk failure (a single component failure). They must eliminate this single point of failure by providing a local synchronously mirrored copy of the data, which gives them the full benefit of HyperSwap.

They are due for a disk technology refresh and want to take advantage of this activity to add a local copy of the disk for CA.

Whatever solution they chose, the customer must not be exposed (from a DR risk perspective) as they implement the solution.

Because of their requirement for local synchronous mirroring and HyperSwap, they also must decide how to protect their data for DR purposes. Although the use of Global Mirror (GM) with the Metro Mirror locally in a Metro Global Mirror (MGM) 3-site configuration might be an option, they cannot achieve ZDL for DR events with GM, which is an absolute requirement for their business.

GDPS Metro in a dual leg configuration can provide them with synchronous mirroring, both locally and to the remote data center, and meet their ZDL requirement.

Another key consideration that the customer has is the skills that they built in by using GDPS/PPRC as their DR solution. Although they understand that a new topology with an extra copy of data necessitates changes, they want to avoid reinventing the wheel and not use a radically different solution that voids their investment in the GDPS technology. They want the solution to be phased in.

GDPS Metro (dual-leg) is the ideal solution for this customer. The Multi-Target Metro Mirror (MTMM) copy technology, which is used by GDPS Metro in a dual leg configuration, meets their requirements for local CA and remote DR with minor other skill requirements and their existing Metro Mirror replication can remain functional during the upgrade from GDPS/PPRC to GDPS Metro dual-leg configuration.

In Table 10-5, we show how GDPS Metro in a dual-leg configuration can meet the customer's requirements.

Table 10-5   Mapping customer requirements to GDPS Metro (dual leg) attributes

Attribute	Supported by GDPS Metro
Two sites, 100 km apart	Y
ZDL	Y (Freeze policy with STOP)
Maintain consistency of secondary volumes	Y (Freeze)
Local CA and remote DR	Y (MTMM technology)
Ability to conduct DR tests without impacting DR readiness	Y (FlashCopy)
Automated recovery of disks and systems following a disaster	Y(GDPS script support)
Ability to transparently swap z/OS disks between the local copies of data transparently	Y (HyperSwap, preferred leg)
Ability to transparently swap z/OS disks between the one of the Site1 copies and the Site2 copy transparently to facilitate orderly shutdown	Y (HyperSwap, nonpreferred leg)
Support for a single Metro Mirror leg (Site1-Site2) to facilitate a phased migration to the new topology	Y
Protect investment in GDPS Metro skills	Y
Maintain the existing Site1-Site2 mirror while adding a local mirror	Y

The customer can plan for the following high-level steps when moving their GDPS Metro single leg environment to a GDPS Metro dual leg environment:

1. Refresh the Site1 and Site2 disks with new technology disks that support the MTMM technology. This process is fairly familiar to customers. Often, it can be achieved nondisruptively by using HyperSwap or Transparent Data Migration Facility (TDMF) technologies. Now, the customer also acquires the third set of disks that are to be installed locally.

2. Upgrade GDPS/PPRC to a GDPS Metro single-leg configuration. GDPS Metro in a single-leg configuration functions similar to GDPS/PPRC, with some minor differences.

   The customer has the same protection and capabilities that they had with GDPS/PPRC. The procedural changes that are required to accomplish this implementation step are minor because the overall topology of their mirror did not change. The customer must adjust some of their GDPS scripts and operational procedures, but this adjustment is not a major change.

3. Finalize the implementation by adding the second, local replication leg to the GDPS Metro configuration. This step requires some modifications to the customer's GDPS automation scripts and the addition of some new scripts because the new topology with two replication legs can now cater to more planned and unplanned outage scenarios. The operational procedures must also be changed in parallel.

   Because the customer has familiarized themselves with the high-level differences between GDPS/PPRC and GDPS Metro while running in the single-leg configuration, this second step is not a radical change from a skills perspective. With the completion of this step, the customer meets all of their requirements.

10.7 DR in two data centers at global distance

The customer in this scenario has a data center in Asia and another in Europe. Following the tsunami disaster in 2004, the customer decides to remote copy their production sysplex data to their data center in Europe. The customer is willing to accept the small data loss that results from the use of asynchronous remote copy.

However, a requirement exists that the data in the remote site is consistent to allow application restart. In addition, to minimize the restart time, the solution must automatically recover the secondary disks and restart all the systems.

The customer has approximately 10000 primary volumes that they want to mirror.

GDPS GM is the right GDPS offering for this customer. Because of the long distance between the two sites, which approaches 15000 km (9320.5 miles), a synchronous remote copy method cannot be used.

Table 10-6 shows how the customer's configuration and requirements map to the capabilities of GDPS GM.

Table 10-6   Mapping customer requirements to GDPS GM attributes

Attribute	Supported by GDPS GM
Two sites that are separated by thousands of km/miles	Y
Willing to accept small data loss	Y (The actual amount of data loss depends on several factors, most notably the available bandwidth)
Maintain consistency of secondary volumes	Y
Maintain consistency of secondary volumes during resync	Y[a] (FlashCopy)
Over 10000 volumes	Y
Both z/OS and non z/OS disks need to be mirrored	Y
Automated recovery of disks and systems following a disaster	Y(GDPS script support)

> a. FlashCopy is used to create a consistent set of secondary volumes before a resynchronization, following a suspension of remote copy sessions.

The first step for the customer is to size the required bandwidth for the GM links. This information is used in the tenders for the remote connectivity. Assuming that the cost of the remote links is acceptable, the customer starts installing GDPS GM concurrently with setting up the remote connectivity.

Pending the availability of the remote connectivity, two LPARs are set up for GM testing. The first LPAR hosts the GDPS K-sys controlling system function in Asia. It can be hosted on one of the production system LPARs in the sysplex or in a separate LPAR. The second LPAR hosts the GDPS R-sys controlling system function in the recovery region (Europe). The R-sys runs in a stand-alone LPAR or monoplex. This setup allows the systems programmers and operators to become familiar with GDPS operations and control.

10.8 Other configurations

Many other combinations of configurations are available.

However, we believe that the examples that are provided in this chapter cover the options of one or two sites, short and long distance, and CA and DR requirements.

If you feel that your configuration does not fit into one of the scenarios that are described here, contact your IBM representative for more information about how GDPS can address your needs.

11

IBM GDPS Enterprise Portal

In most cases, GDPS clients deploy more than one IBM Geographically Dispersed Parallel Sysplex (GDPS) environment.

Multiple GDPS environments are deployed for various reasons, including the following examples:

► In most cases, clients have at least one production environment, one development or test environment, and one sandbox (or sandpit) environment.

When a client has multiple production environments, they might have corresponding development or test and sand box environments for each development environment.

► Clients sometimes have a separate production environment for each line of business.

► Global clients often have separate production environments for each geographical location in which they do business.

This need for many environments can add up to a significant number of GDPS environments and as the number of GDPS environments increases, the difficulty of managing all of the environments also increases. Even something as simple as determining whether any environments are experiencing any issues requires a user (or multiple users) to log on to at least one GDPS controlling system per environment, either by using NetView or the GDPS GUI.

Fortunately for clients with multiple GDPS environments, GDPS provides a function that is known as the *GDPS Enterprise Portal* (also referred to as *the GDPS Portal* or *the Portal*). The GDPS Portal provides a single point of control for monitoring and managing all of your GDPS environments.

In this chapter, we provide an overview of the GDPS Portal, which is a feature that is included with a GDPS license, along with the GDPS GUI.

This chapter includes the following topics:

► 11.1, "Viewing your GDPS environments" on page 308

► 11.2, "Managing your GDPS environments" on page 313

11.1 Viewing your GDPS environments

The GDPS Portal provides a logical view and a physical view of your GDPS environments. These views are described next.

11.1.1 Logical view

The logical view provides a representation of the logical entities in each of your GDPS environments. Logical entities include systems (production and GDPS controlling systems), coupling facilities (CFs), replication, sites, and mirroring relationships.

Figure 11-1 shows an example of the logical view.
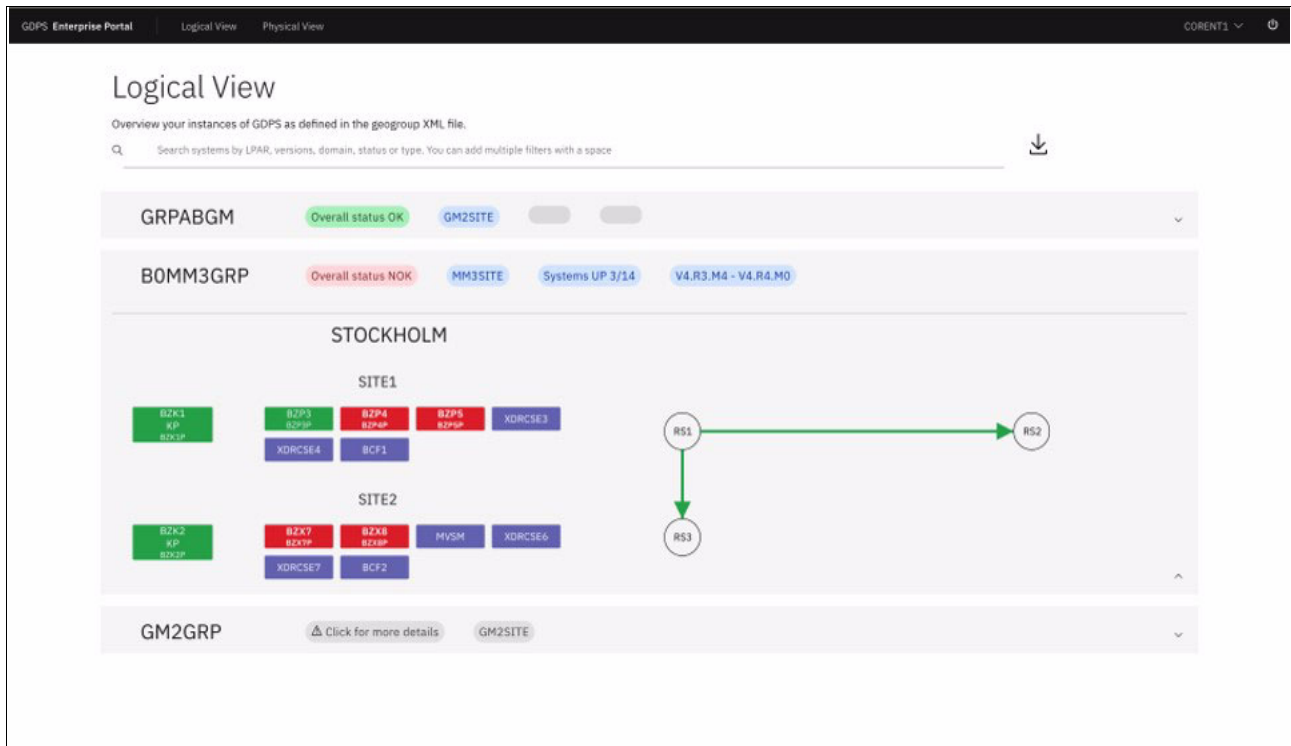
Figure 11-1   Logical view of GDPS Portal

As shown in Figure 11-1, three different GDPS environments are presented:

► The first section represents the environment GRPABGM in our example. The section for GRPABGM is not expanded, but from the summary information that is provided, we can see that GRPABGM is a 2-Site GDPS Global - GM environment with an overall status of OK. That is, all of the systems are active and mirroring (Global Mirror (GM)) is running normally.

► The second section represents the environment B0MM3GRP in our example. From the summary information that is provided for this environment, we can see that it is a 3-site GDPS Metro environment (also known as a *GDPS Metro Dual Leg environment*).

We can also see from the summary information that the overall status for the environment is Not OK (NOK), meaning that at least one unexpected condition exists in the environment.

The section representing B0MM3GRP is expanded to provide the following information about the systems in the environment:

– The region in which the environment exists is called STOCKHOLM.

– Two systems in Site1 and one system in Site2 are green, which indicates that they are active.

– Two systems in Site1 and two systems in Site2 are red, which indicates that they are down.

– Three systems in Site1 and four systems in Site2 are purple, which indicates that they are manual systems; therefore, the status of these systems is unknown.

Clicking a system in the expanded section provides more information about that system. In Figure 11-2, the user clicked the z/OS production system that is named BZP3.
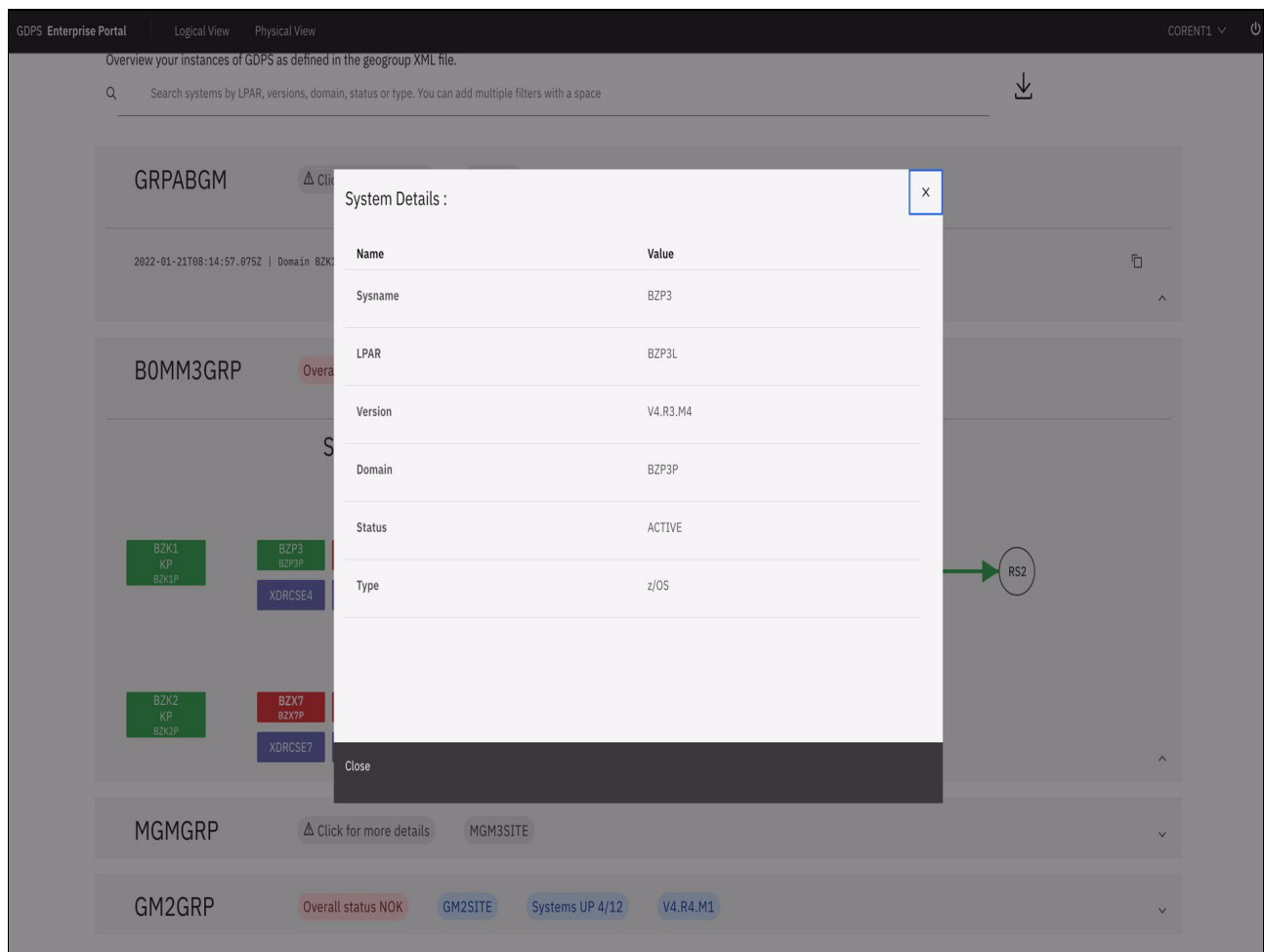
Figure 11-2   More system information

As shown in Figure 11-2, the information box that is presented for system BZ3P provides the following information:

– LPAR in which the system is running.
– Level of GDPS that is running on the system.
– Name of the NetView domain in which the system runs.
– System's status.
– System type.

Also in the expanded section for B0MM3GRP (see Figure 11-1 on page 308), three replication sites (RSs) (RS1, RS2, and RS3) and the arrows between RS1 and RS2 and between RS1 and RS3 represent mirroring (Metro Mirror) for the two replication legs. The fact that the arrows are green indicates that mirroring is running normally.

Hovering over one of the arrows that represents the replication legs provides more information about the replication leg. As shown in Figure 11-3, the user hovered over the replication leg between RS1 and RS2.
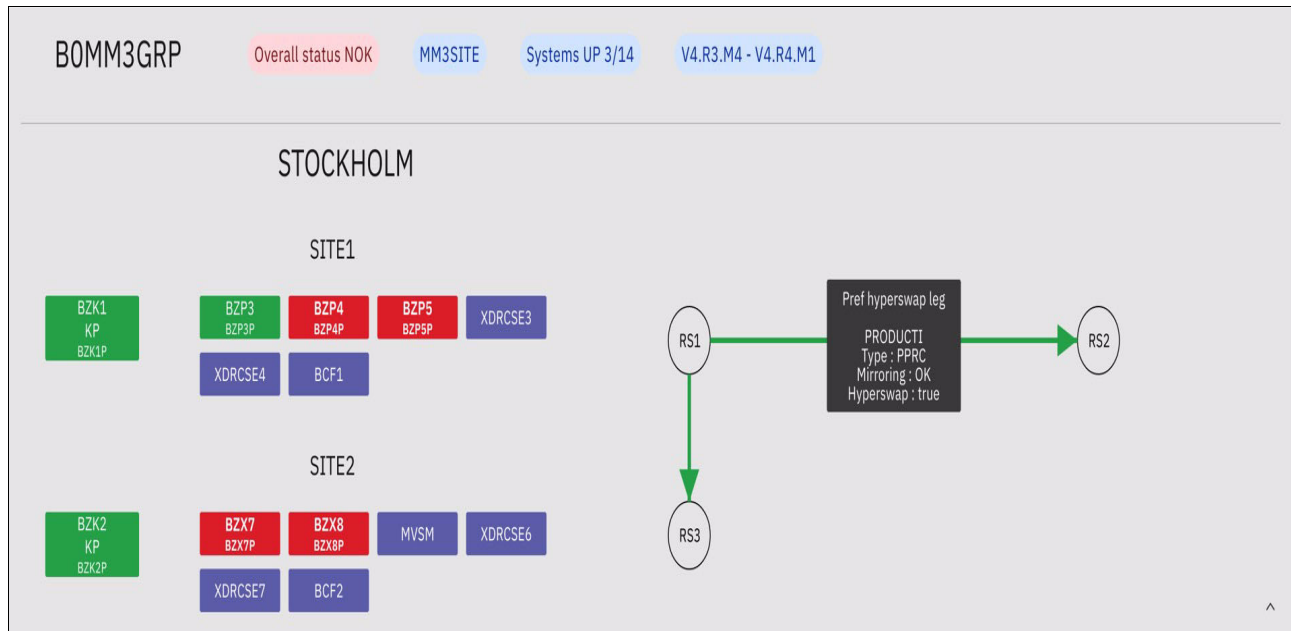
Figure 11-3   More replication leg information

The informational box that is shown provides the following information:

– Leg is the preferred HyperSwap leg.
– Consistency group that is associated with the leg is called PRODUCTI.
– Mirroring status is `OK`.
– HyperSwap is enabled on the leg.

► The third section that is shown in Figure 11-1 on page 308 represents the environment that is called GM2GRP in our example.

In the summary information for GM2GRP, we can see that it is another 2-site GDPS Global - GM; however, the status button is gray and includes the text: `Click for more details.` Clicking this button displays more information about the status. In this case, the GDPS Portal cannot connect to the GDPS controlling systems in the GM2GRP environment.

The GDPS Enterprise Portal support is also available for the GDPS Continuous Availability solution as shown in Figure 11-4 on page 311 for the Logical View.
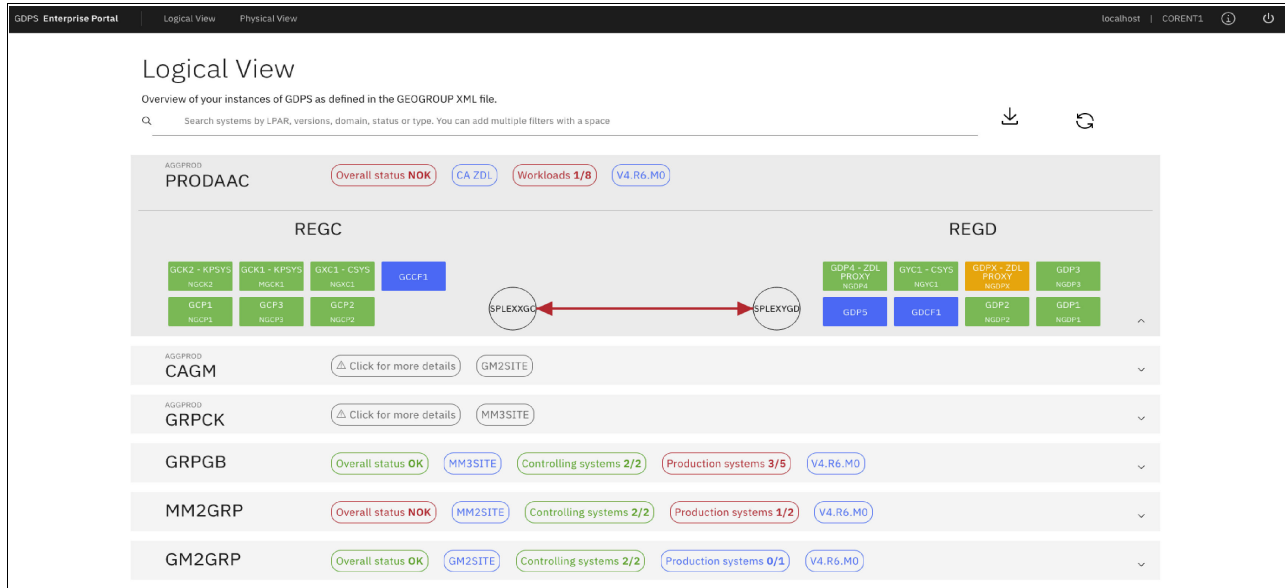
Figure 11-4   GDPS Portal - GDPS Continuous Availability logical view

11.1.2 Physical view

The physical view provides a representation of the physical entities in each of your GDPS environments. Physical entities include host servers (Central Processor Complexes (CPCs)), disk storage systems, and mirroring links.

Figure 11-5 shows an example of the physical view.
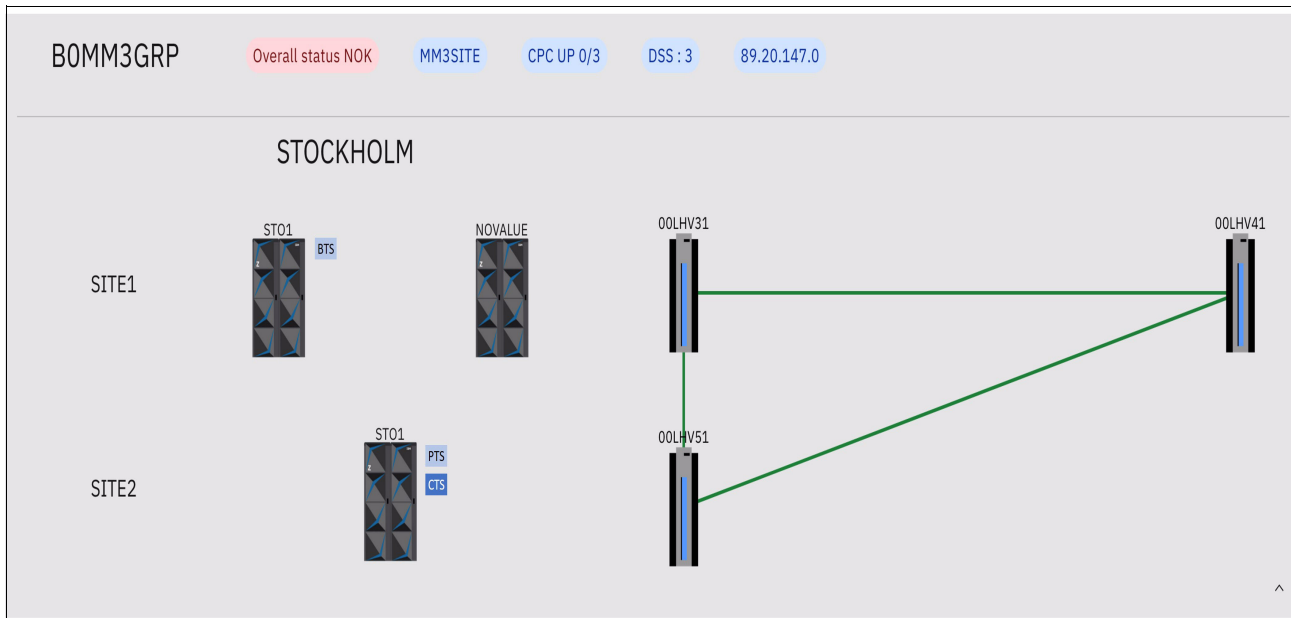
Figure 11-5   GDPS Portal - physical view

As shown in Figure 11-5, we again see a representation of the environment that is named B0MM3GRP, which is the same environment that was represented in the expanded section of the window that is shown in Figure 11-1 on page 308.

We see the same summary information for B0MM3GRP that we saw before, however; this time, we see CPCs instead of systems. Also, we see disk storage systems instead of RSs.

We also see that the systems at Site1 are running on two CPCs and the systems at Site2 are running on a single CPC. The CPC at Site2 serves as both the Server Time Protocol (STP) Preferred Time Server (PTS) and the STP Current Time Server (CTS). One of the CPCs in Site1 serves as the STP Backup Time Server (BTS).

Figure 11-5 on page 311 also shows that two disk storage systems exist in Site1 and one disk storage system in Site2. The two disk storage systems in Site1 correspond to the two logical RSs in Site1, as shown in Figure 11-1 on page 308.

Clicking one of the CPC's displays more information, as shown in Figure 11-6.
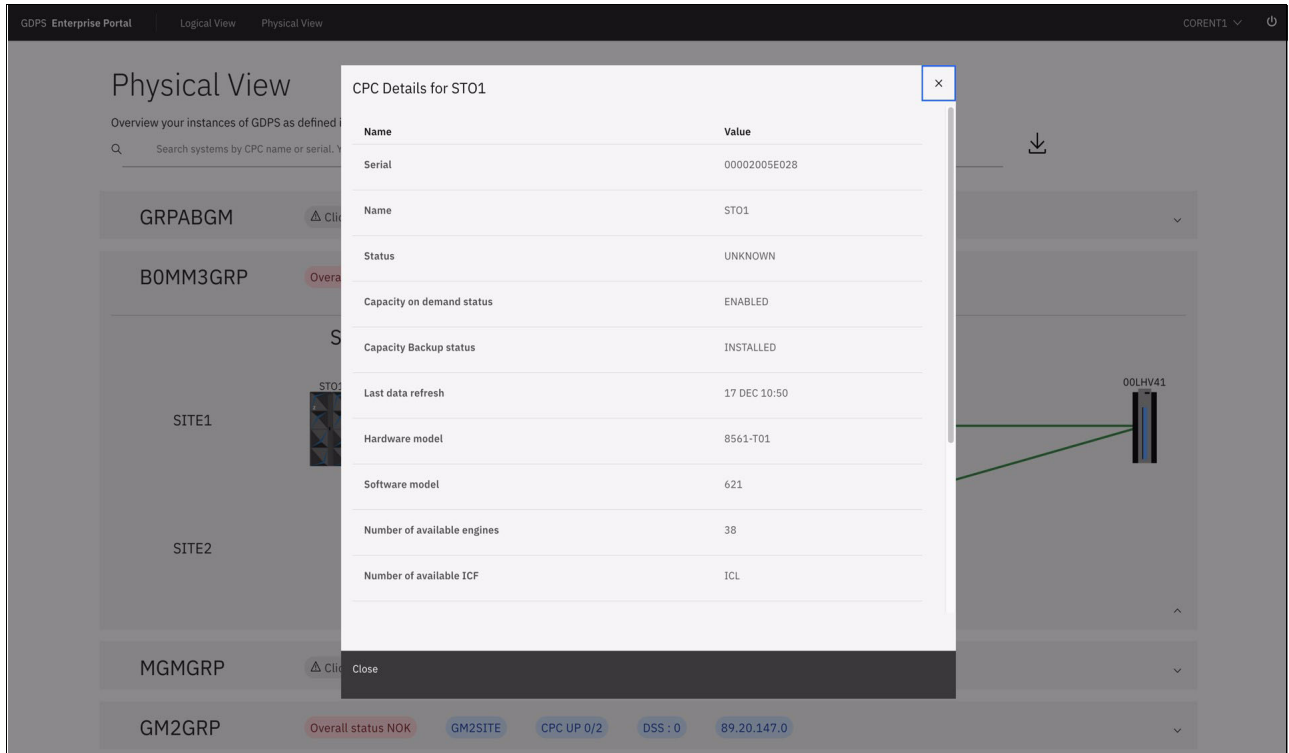
Figure 11-6   More CPC information

In this case, the user clicked the CPC in Site1 named STO1. The extra information that is shown in the box that is shown includes the CPC serial number and status, status of capacity on demand (On/Off Capacity on Demand (OOCoD)) and Capacity BackUp (CBU), the hardware and software model numbers, the number of available engines, and more.

The green lines between the disk storage systems represent the PPRC links over which Metro Mirror is flowing. In this case, the fact that the lines are green indicates that the PPRC links are all active. Hovering over the lines provides more information, such as how many links exist between the primary and secondary disk storage systems and the status of each link.

11.2 Managing your GDPS environments

The GDPS Portal provides a single point of control for monitoring and managing your GDPS environments. When you need to view a specific environment at a more detailed level or you need to change a specific environment, you can "drill down" to the GDPS GUI for that environment where you can then perform these tasks.

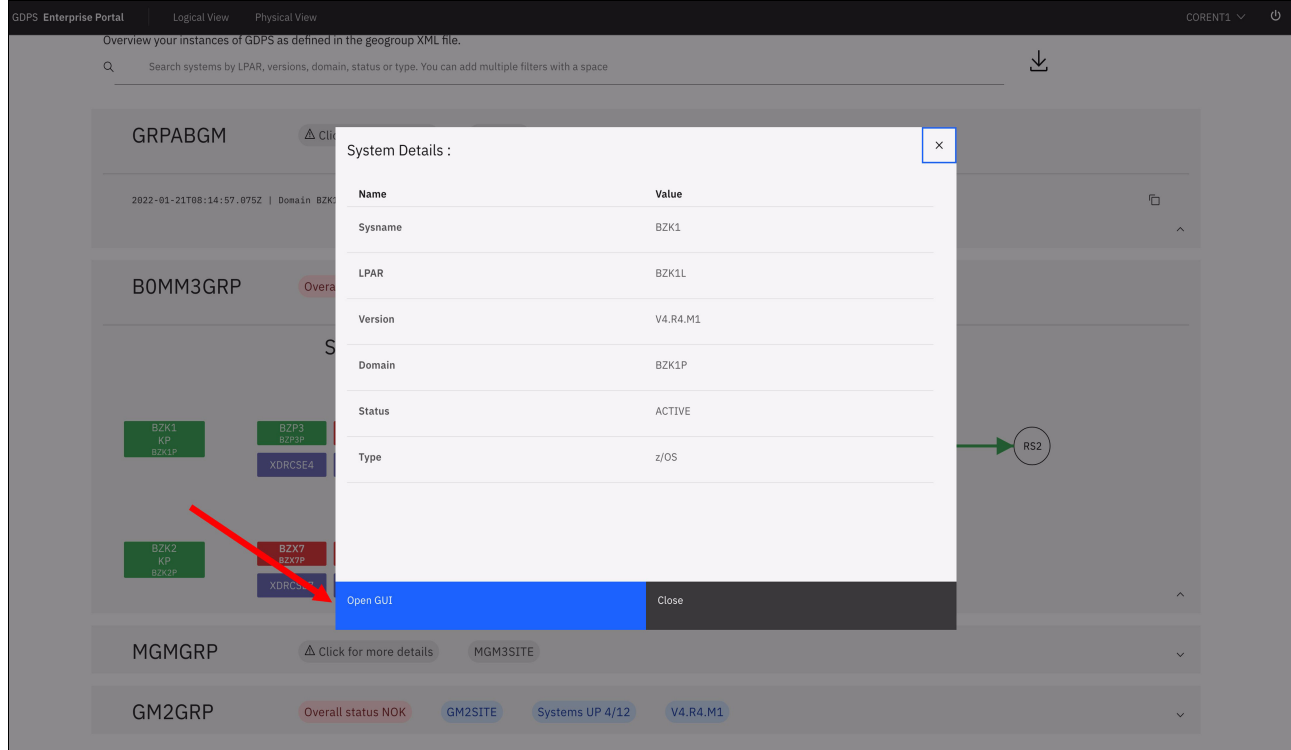To use the GUI, click one of the GDPS controlling systems from the logical view, as shown in Figure 11-7.

Figure 11-7   Drilling down to a specific GDPS environment

As shown in Figure 11-7, the user clicked the system in Site1 that is named BZK1, which is one of the GDPS Metro controlling systems in the B0MM3GRP environment. The information box that is shown includes a button that is labeled `Open GUI`," which is highlighted by a red arrow in Figure 11-7.

Clicking this button takes you to the dashboard of the GDPS Metro GUI for the B0MM3GRP environment.

Abbreviations and acronyms

API	application programming interface		**IBM**	International Business Machines Corporation
BCPii	BCP Internal Interface		**IR**	Incremental Resynchronization
BRS	Business Recovery Services		**ISL**	interswitch links
BTS	Backup Time Server		**ISV**	independent software vendor
BVIR	Bulk Volume Information Retrieval		**IVP**	Installation Verification Program
CA	continuous availability		**LAN**	local area network
CBU	Capacity BackUp		**LCP**	Logical Corruption Protection
CC	Concurrent Copy		**LIC**	Licensed Internal Code
CDS	couple data sets		**LSS**	logical subsystem
CF	coupling facility		**MCG**	Multiple Consistency Groups
CKD	Count-Key-Data		**MGM**	Metro Global Mirror
CLI	command-line interface		**MTMM**	Multi-Target Metro Mirror
CPC	Central Processor Complex		**NDSS**	non-disruptive statesave
CPE	Capacity for Planned Events		**NMI**	Network Management Interface
CS	copy sets		**NOK**	Not OK
CSE	Cross System Extension		**NRO**	network recovery objective
CTN	Coordinated Timing Network		**NVS**	non-volatile store
CTS	Current Time Server		**ODD**	on-demand dump
CWDM	Course Wavelength Division Multiplexing		**OOCoD**	On/Off Capacity on Demand
DBMS	database management system		**PAV**	Parallel Access Volume
DCM	Distributed Cluster Management		**PiT**	point-in-time
DE	Device End		**PMT**	Preserve Mirror Tool
DR	disaster recovery		**POR**	power-on reset
DWDM	Dense Wavelength Division Multiplexing		**PTS**	Preferred Time Server
ELB	Extended Long Busy		**RC**	recovery copy
FB	Fixed-Block		**RPO**	recovery point objective
FCP	Fibre Channel Protocol		**RS**	replication site
FFDC	first failure data capture		**RSA**	Remote Supervisor Adapter
FICON	Fiber Connection		**RSP**	Reserve Storage Pool
GBP	group buffer pool		**RTO**	recovery time objective
GCG	Global Consistency Group		**SA MP**	System Automation for Multiplatforms
GDPS	Geographically Dispersed Parallel Sysplex		**SASP**	Server/Application State Protocol
GeoXML	GDPS XML Conversion		**SCI**	State Change Interrupt
GM	Global Mirror		**SDF**	Status Display Facility
HA	high availability		**SDSF**	Spool Display and Search Facility
HADR	high availability and disaster recovery		**SE**	Support Element
HIPAA	Health Insurance Portability and Accountability Act		**SGC**	Safeguarded Copy
			SRB	System Recovery Boost
			SSC	Secure Service Container

315

SSI	Single System Image
STP	Server Time Protocol
T0	time zero
TCM	Testcopy Manager
TDMF	Transparent Data Migration Facility
TEM	Telecom Expense Management
TOD	time-of-day
UCB	unit control block
VM	virtual machine
WAN	wide area network
WDM	Wavelength Division Multiplexor
WLM	Workload Manager
WTOR	write to operator with reply
zBX	IBM zEnterprise BladeCenter Extension
ZDL	zero data loss
zFBA	IBM z/OS Fixed-Block Architecture

Glossary

A

AOM. asynchronous operations manager.

application system. A system that consists of one or more host systems that perform the main set of functions for an establishment. This system updates the primary disk volumes that are copied by a copy services function.

asynchronous operation. A type of operation in which the remote copy Global Mirror (GM) function copies updates to the secondary volume of a GM pair some time after the primary volume is updated. Contrast with synchronous operation.

B

backup. The process of creating a copy of data to ensure against accidental loss.

C

cache. A random access electronic storage in selected storage controls that is used to retain frequently used data for faster access by the channel.

central processor complex (CPC). The unit within a cluster that provides the management function for the storage server. It consists of cluster processors, cluster memory, and related logic.

channel connection address (CCA). The input/output (I/O) address that uniquely identifies an I/O device to the channel during an I/O operation.

channel interface. The circuitry in a storage control that attaches storage paths to a host channel.

consistent copy. A copy of a data entity (for example, a logical volume) that contains the contents of the entire data entity from a single instant in time.

control unit address. The high-order bits of the storage control address that is used to identify the storage control to the host system.

D

dark fiber. A dedicated fiber link between two sites that is dedicated for use by one client.

DASD. direct access storage device.

data in transit. The update data on application system DASD volumes that is being sent to the recovery system for writing to DASD volumes on the recovery system.

device address. The z/OS term for the field of a Count-Key-Data (CKD) device-level frame that selects a specific device on a control unit image. The one or two leftmost digits are the address of the channel to which the device is attached. The two rightmost digits represent the unit address.

device number. The ESA/390 term for a four-hexadecimal-character identifier (for example, 13A0) that you associate with a device to facilitate communication between the program and the host operator. The device number that you associate with a subchannel.

Device Support Facilities program (ICKDSF). A program that is used to initialize DASD at installation and perform media maintenance.

DFDSS. Data Facility Data Set Services. An IBM licensed program to copy, move, dump, and restore data sets and volumes.

DFSMSdss. A functional component of DFSMS that is used to copy, dump, move, and restore data sets and volumes.

disaster recovery (DR). Recovery after a disaster, such as a fire, that destroys or otherwise disables a system. DR techniques typically involve restoring data to a second (recovery) system and then using the recovery system in place of the destroyed or disabled application system. Also see *recovery*, *backup*, and *recovery system*.

dual copy. A high availability (HA) function that is made possible by the nonvolatile storage in cached IBM storage controls. Dual copy maintains two functionally identical copies of designated DASD volumes in the logical storage subsystem, and automatically updates both copies every time a write operation is issued to the dual copy logical volume.

duplex pair. A volume that consists of two physical devices within the same or different storage subsystems that are defined as a pair by a Metro Mirror and are not in a suspended or pending state. The operation records the same data onto each volume.

DWDM. Dense Wavelength Division Multiplexor. A technique that is used to transmit several independent bit streams over a single fiber link.

F

FlashCopy. A point-in-time (PiT) copy services function that can quickly copy data from a source location to a target location.

K

km. kilometer.

L

Licensed Internal Code (LIC). Microcode that IBM does not sell as part of a machine, but licenses to the customer. LIC is implemented in a part of storage that is not addressable by user programs. Some IBM products use it to implement functions as an alternative to hard-wired circuitry.

link address. On a Fibre Connection (FICON) interface, the portion of a source or destination address in a frame that FICON uses to route a frame through an FICON director. FICON associates the link address with a specific switch port that is on the FICON director. Equivalently, it associates the link address with the channel subsystem or controller link-level functions that are attached to the switch port.

logical partition (LPAR). The ESA/390 term for a set of functions that create the programming environment that is defined by the ESA/390 architecture. ESA/390 architecture uses this term when more than one LPAR is established on a processor. An LPAR is conceptually similar to a virtual machine (VM) environment, except that the LPAR is a function of the processor. Also, the LPAR does not depend on an operating system to create the VM environment.

logical subsystem (LSS). The logical functions of a storage controller that allow one or more host I/O interfaces to access a set of devices. The controller aggregates the devices according to the addressing mechanisms of the associated I/O interfaces. One or more LSSs exist on a storage controller. In general, the controller associates a specific set of devices with only one LSS.

O

orphan data. Data that occurs between the last, safe backup for a recovery system and the time when the application system experiences a disaster. This data is lost when the application system becomes available for use or when the recovery system is used in place of the application system.

P

peer-to-peer remote copy (PPRC). A hardware-based remote copy option that provides a synchronous volume copy across storage subsystems for DR, device migration, and workload migration. This term is synonymous with "Metro Mirror"

pending. The initial state of a defined volume pair before it becomes a duplex pair. During this state, the contents of the primary volume are copied to the secondary volume.

PPRC. See peer-to-peer remote copy.

PPRC dynamic address switching (P/DAS). A software function that dynamically redirects all application I/O from one PPRC volume to another PPRC volume.

primary device. One device of a dual copy or remote copy volume pair. All channel commands to the copy logical volume are directed to the primary device. The data on the primary device is duplicated on the secondary device. See also secondary device.

PTF. program temporary fix.

R

RACF. Resource Access Control Facility.

recovery system. A system that is used in place of a primary application system that is no longer available for use. Data from the application system must be available for use on the recovery system. This task is accomplished through backup and recovery techniques, or through various DASD copying techniques, such as remote copy.

remote copy. A storage-based DR and workload migration function that can copy data in real time to a remote location.

resynchronization. A track image copy from the primary volume to the secondary volume of only the tracks that changed since the volume was last in duplex mode.

S

secondary device. One of the devices in a dual copy or remote copy logical volume pair that contains a duplicate of the data on the primary device. Unlike the primary device, the secondary device may accept only a limited subset of channel commands.

sidefile. A storage area used to maintain copies of tracks within a concurrent copy (CC) domain. A CC operation maintains a sidefile in the storage control cache and another in processor storage.

simplex state. A volume is in the simplex state if it is not part of a dual copy or a remote copy volume pair. Ending a volume pair returns the two devices to the simplex state. In this case, there is no longer any capability for either automatic updates of the secondary device or for logging changes, as is the case in a suspended state.

site table. Entity within GDPS that is created from information in the GEOPLEX DOMAINS. It contains a list of all systems in the GDPS environment.

suspended state. When only one of the devices in a dual copy or remote copy volume pair is being updated because of a permanent error condition or an authorized user command. All writes to the remaining functional device are logged. This function allows for automatic resynchronization of both volumes when the volume pair is reset to the active duplex state.

synchronization. An initial volume copy that is a track image copy of each primary track on the volume to the secondary volume.

synchronous operation. A type of operation in which the remote copy Metro Mirror function copies updates to the secondary volume of a Metro Mirror pair at the same time that the primary volume is updated. Contrast with "asynchronous operation".

T

timeout. The time in seconds that the storage control remains in a "long busy" condition before physical sessions are ended.

Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this book.

IBM Redbooks publications

The following IBM Redbooks publications provide more information about the topics in this book. Some publications that are referenced in this list might be available in softcopy only:

► *IBM DS8870 Copy Services for IBM z Systems*, SG24-6787

► *IBM System Storage Business Continuity: Part 1 Planning Guide*, SG24-6547

► *IBM TotalStorage Enterprise Storage Server Implementing ESS Copy Services with IBM eServer zSeries*, SG24-5680

► *IBM Virtualization Engine TS7700 with R 2.0*, SG24-7975

► *IBM Z Connectivity Handbook*, SG24-5444

► *Server Time Protocol Implementation Guide*, SG24-7281

► *Server Time Protocol Planning Guide*, SG24-7280

The following IBM Redpaper publications contain information about the Dense Wavelength Division Multiplexing (DWDM) products that are qualified for use with GDPS:

► *IBM System z Qualified WDM: Adva FSP 2000 at Release Level 6.2*, REDP-3903

► *IBM System z Qualified WDM: Nortel Optical Metro 5200 at Release Level 10.0*, REDP-3904

► *zSeries Qualified WDM Vendor: Cisco Systems*, REDP-3905

► *zSeries Qualified WDM Vendor: Lucent Techologies*, REDP-3906

You can search for, view, download, or order these documents and other Redbooks publications, Redpaper publications, web docs, drafts, and additional materials, from here:

ibm.com/redbooks

Other publications

The following publications are also relevant as further information sources:

► *Advanced Copy Services,* SC35-0428

► *IBM Tivoli NetView for z/OS Installation: Getting Started V6R2,* GI11-9443

► *System-Managed CF Structure Duplexing Implementation Summary,* GM13-0540

► *System z Capacity on Demand User's Guide,* SC28-6846

► *z/VM CP Planning and Administration*, SC24-6083

Online resources

The following web pages are also relevant as further information sources:

► Interagency Paper on Sound Practices to Strengthen the Resilience of the US Financial System:

http://www.sec.gov/news/studies/34-47638.htm

► Qualified DWDM vendors:

http://ibm.co/1Jia5AJ

► Speech by SEC Staff: Disaster Recovery and Business Continuity Planning:

http://www.sec.gov/news/speech/spch050103mag.htm

► US-EU and US-Swiss Safe Harbor Frameworks, export.gov website:

http://www.export.gov/safeharbor/

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

Redbooks

IBM GDPS: An Introduction to Concepts and Capabilities

(0.5" spine)
0.475"<->0.873"
250 <-> 459 pages

IBM

IBM®

Get connected

in

Redbooks®

ibm.com/redbooks