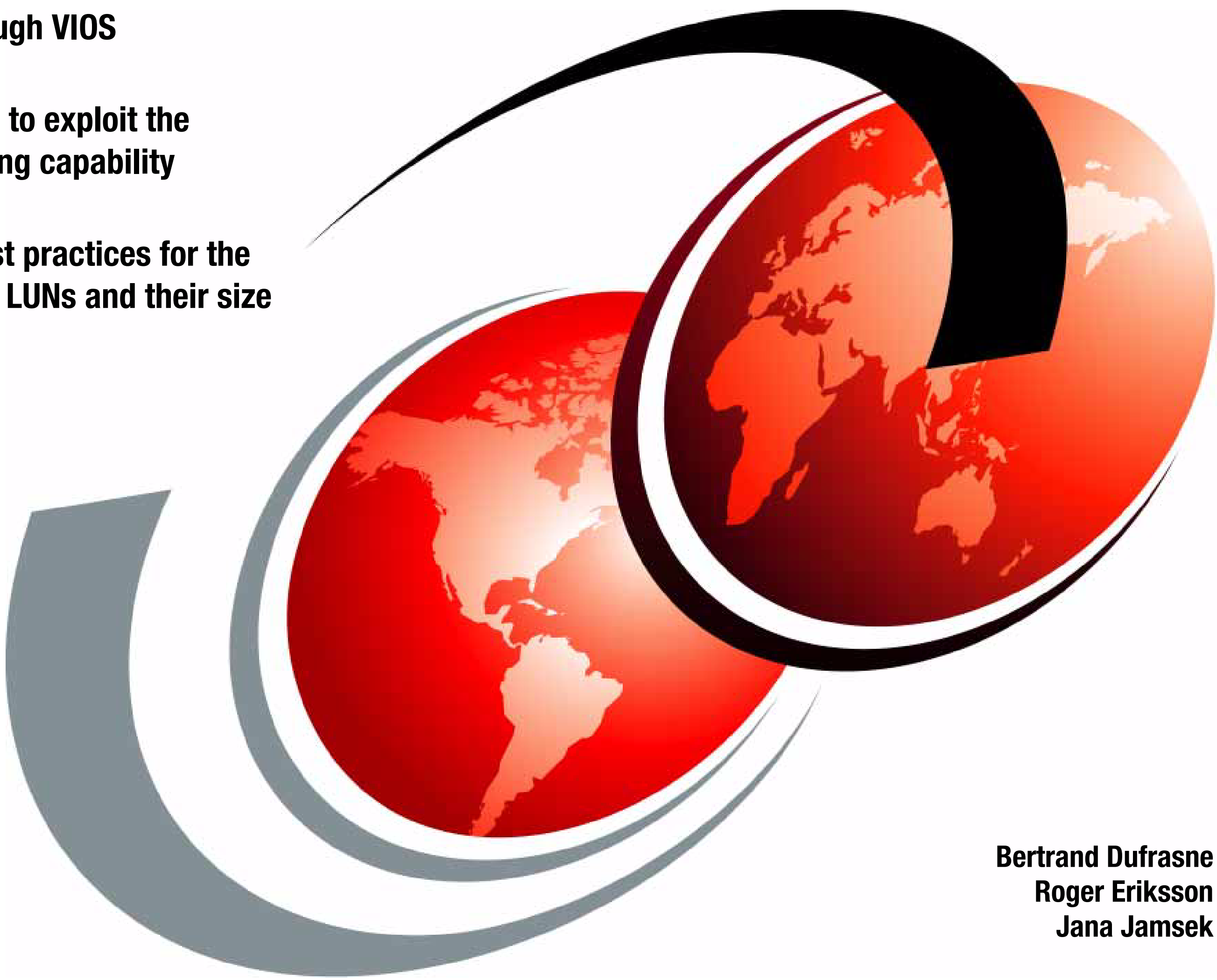
IBM XIV Storage System with the Virtual I/O Server and IBM i

Understand how to attach XIV to IBM i through VIOS

Learn how to exploit the multipathing capability

Follow best practices for the number of LUNs and their size



Bertrand Dufrasne
Roger Eriksson
Jana Jamsek

Redpaper

ibm.com/redbooks

International Technical Support Organization

IBM XIV Storage System with the Virtual I/O Server and IBM i

February 2012

REDP-4598-01

Note: Before using this information and the product it supports, read the information in "Notices" on page v.

Second Edition (February 2012)

This edition applies to version 11 of the IBM XIV System Storage software, when attaching the XIV Storage System server to an IBM i client through a Virtual I/O Server 2.1.2 partition.

This document was created or updated on February 20, 2012.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

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 give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: 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 states 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 Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

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

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.

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 the names and addresses used by an actual business enterprise 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.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
BladeCenter®
DS8000®
i5/OS®
IBM®
Micro-Partitioning®
Power Systems™
POWER6+™
POWER6®
POWER7®
PowerVM®
Power®
POWER®
Redbooks®
Redpaper™
Redbooks (logo) ®
System i®
System p®
System Storage®
System x®
XIV®

The following terms are trademarks of other companies:

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

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

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

Preface

This IBM® Redpaper™ publication explains how you can connect the IBM XIV® Storage System server to the IBM i operating system through the Virtual I/O Server (VIOS). A connection through the VIOS is especially of interest to IT centers that have many small IBM i partitions. When using the VIOS, the Fibre Channel host adapters can be installed in the VIOS and shared by many IBM i clients by using virtual connectivity to the VIOS.

This paper also includes guidance for sizing XIV logical unit numbers (LUNs) when attached to the IBM i client.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO) in San Jose, CA.

Bertrand Dufrasne is an IBM Certified Consulting IT Specialist and Project Leader for IBM System Storage® disk products at the ITSO in San Jose, CA. He has worked for IBM in various IT areas. He has authored many IBM Redbooks® publications and has developed and taught technical workshops. Before joining the ITSO, Bertrand worked for IBM Global Services as an Application Architect. He holds a master degree in electrical engineering.

Roger Eriksson is an STG Lab Services consultant, based in Stockholm, Sweden, and works for the European Storage Competence Center in Mainz, Germany. He is a Senior Accredited IBM Product Service Professional. Roger has over 20 years of experience in working on IBM servers and storage, including Enterprise and Midrange disk, network-attached storage (NAS), storage area network (SAN), IBM System x®, System p®, and BladeCenter®. He has done consulting, proof of concepts, and education mainly with the XIV product line since December 2008, working with clients and various IBM teams. He holds a degree in mechanical engineering.

Jana Jamsek is an IT Specialist for IBM Slovenia. She works in Storage Advanced Technical Support for Europe as a specialist for IBM Storage Systems and the IBM i (i5/OS®) operating system. Jana has eight years of experience working with the IBM System i® platform and its predecessor models, and eight years of experience working with storage. She has a master degree in computer science and a degree in mathematics from the University of Ljubljana in Slovenia.

Thanks to the following people for their contributions to this project:

John Bynum
Ingo Dimmer
Adrian Flores
Robert Gagliardi
Aviad Offer
Alison Pate
Brian Sherman
Wesley Varela
Anthony Vattathil
IBM Corporation

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 in length, 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 at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks
- Send your comments in an e-mail to:

  redbooks@us.ibm.com
- Mail your comments to:

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

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- 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

1

Overview of the Virtual I/O Server with IBM i

This chapter provides an overview of IBM PowerVM® and the Virtual I/O Server (VIOS). It describes Fibre Channel connectivity configurations in the VIOS and provides guidance for setting up the VIOS to connect the IBM XIV Storage System server to an IBM i client.

This chapter includes the following sections:

- Introduction to IBM PowerVM
- PowerVM client connectivity to the IBM XIV Storage System

1.1 Introduction to IBM PowerVM

Virtualization on IBM Power® Systems™ servers can provide a rapid and cost-effective response to many business needs. Virtualization capabilities have become an important element in planning for IT floor space and servers. Growing commercial and environmental concerns create pressure to reduce the power footprint of servers. Also, with the escalating cost of powering and cooling servers, consolidation and efficient utilization of the servers are becoming critical.

Virtualization on Power Systems servers enables an efficient utilization of servers by reducing the following areas:

- Server management and administration costs, because of fewer physical servers
- Power and cooling costs with increased utilization of existing servers
- Time to market, because of immediate deployment of virtual resources

1.1.1 Overview of IBM PowerVM

IBM PowerVM is a software appliance tied to IBM Power Systems, that is, the converged IBM i and IBM p server platforms. It is licensed on a POWER® processor basis. PowerVM is a virtualization technology for IBM AIX®, IBM i, and Linux environments on IBM POWER processor-based systems.

PowerVM offers a secure virtualization environment with the following major features and benefits:

- Consolidates diverse sets of applications that are built for multiple operating systems (AIX, IBM i, and Linux) on a single server
- Virtualizes processor, memory, and I/O resources to increase asset utilization and reduce infrastructure costs
- Dynamically adjusts server capability to meet changing workload demands
- Moves running workloads between servers to maximize availability and avoid planned downtime

Virtualization technology is offered in three editions on Power Systems:

- PowerVM Express Edition
- PowerVM Standard Edition
- PowerVM Enterprise Edition

These editions provide logical partitioning technology by using the Hardware Management Console (HMC) or the Integrated Virtualization Manager (IVM), dynamic logical partition (LPAR) operations, IBM Micro-Partitioning® and VIOS capabilities, and Node Port ID Virtualization (NPIV).

PowerVM Express Edition

PowerVM Express Edition is available only on the IBM Power 520 and Power 550 servers. It is designed for clients who want an introduction to advanced virtualization features at an affordable price.

With PowerVM Express Edition, clients can create up to three partitions on a server (two client partitions and one for the VIOS and IVM). They can use virtualized disk and optical devices, in addition to trying the shared processor pool. All virtualization features, such as

Micro-Partitioning, shared processor pool, VIOS, PowerVM LX86, shared dedicated capacity, NPIV, and virtual tape, can be managed by using the IVM.

PowerVM Standard Edition

For clients who are ready to gain the full value from their server, IBM offers the PowerVM Standard Edition. This edition provides the most complete virtualization functions for UNIX and Linux in the industry and is available for all IBM Power Systems servers.

With PowerVM Standard Edition, clients can create up to 254 partitions on a server. They can use virtualized disk and optical devices and try the shared processor pool. All virtualization features, such as Micro-Partitioning, Shared Processor Pool, VIOS, PowerVM Lx86, Shared Dedicated Capacity, NPIV, and Virtual Tape, can be managed by using HMC or IVM.

PowerVM Enterprise Edition

PowerVM Enterprise Edition is offered exclusively on IBM POWER6® servers. It includes all the features of the PowerVM Standard Edition, plus the *PowerVM Live Partition Mobility* capability.

With PowerVM Live Partition Mobility, you can move a running partition from one POWER6 or IBM POWER7® technology-based server to another with no application downtime. This capability results in better system utilization, improved application availability, and energy savings. With PowerVM Live Partition Mobility, planned application downtime because of regular server maintenance is no longer necessary.

1.1.2 Virtual I/O Server

The VIOS is Virtualization software that runs in a separate partition of the POWER system. VIOS provides virtual storage and networking resources to one or more client partitions.

The VIOS owns the physical I/O resources such as Ethernet and SCSI/FC adapters. It virtualizes those resources for its client LPARs to share them remotely by using the built-in hypervisor services. These client LPARs can be created quickly, typically owning only real memory and shares of processors without any physical disks or physical Ethernet adapters.

With Virtual SCSI support, VIOS client partitions can share disk storage that is physically assigned to the VIOS LPAR. This support of VIOS is used to make storage devices, such as the IBM XIV Storage System server. Such devices do not support the IBM i proprietary 520-byte/sectors format that is available to IBM i clients of VIOS.

VIOS owns the physical adapters, such as the Fibre Channel storage adapters that are connected to the XIV system. The logical unit numbers (LUNs) of the physical storage devices that are detected by VIOS are mapped to VIOS virtual SCSI (VSCSI) server adapters that are created as part of its partition profile.

The client partition with its corresponding VSCSI client adapters defined in its partition profile connects to the VIOS VSCSI server adapters by using the hypervisor. VIOS performs SCSI emulation and acts as the SCSI target for the IBM i operating system.

Figure 1-1 shows an example of the VIOS owning the physical disk devices and its virtual SCSI connections to two client partitions.

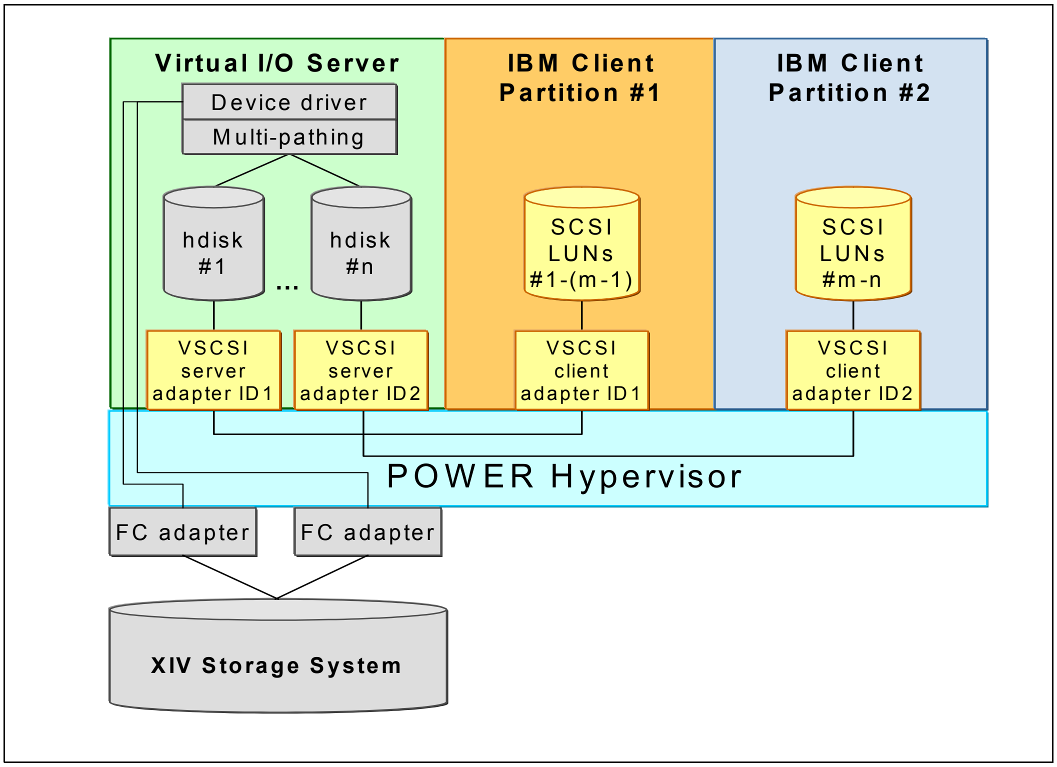


Figure 1-1 VIOS virtual SCSI support

1.1.3 Node Port ID Virtualization

The VIOS technology is enhanced to boost the flexibility of IBM Power Systems servers with support for NPIV. NPIV simplifies the management and improves performance of Fibre Channel SAN environments by standardizing a method for Fibre Channel ports to virtualize a physical node port ID into multiple virtual node port IDs. The VIOS takes advantage of this feature and can export the virtual node port IDs to multiple virtual clients. The virtual clients see this node port ID and can discover devices as though the physical port was attached to the virtual client.

The VIOS does not do any device discovery on ports by using NPIV. Thus no devices are shown in the VIOS connected to NPIV adapters. The discovery is left for the virtual client and all the devices found during discovery are detected only by the virtual client. This way, the virtual client can use FC SAN storage-specific multipathing software on the client to discover and manage devices.

For more information about PowerVM virtualization management, see the IBM Redbooks publication *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590.

VIOS NPIV connection: Connection through VIOS NPIV to an IBM i client is possible only for storage devices that can attach natively to the IBM i operating system, such as the IBM System Storage DS8000® or DS5000. To connect other storage devices, use VIOS with virtual SCSI adapters.

1.2 PowerVM client connectivity to the IBM XIV Storage System

The XIV system can be connected to an IBM i partition through VIOS. To set up the environment on an IBM POWER6 system to connect the XIV system to an IBM i client with multipath through two VIOS partitions, see 3.1, “Connecting a PowerVM IBM i client to XIV” on page 16.

2

Planning for the IBM XIV Storage System server with IBM i

This chapter assists you in properly planning for the use of an IBM XIV Storage System server in an IBM i environment. It includes the following sections:

- Planning preparation
- Best practices

Operating system setup: In this paper, the IBM i operating system is in a logical partition (LPAR) on an IBM Power Systems server or Power Blade platform.

2.1 Planning preparation

You can connect the XIV Storage System to an IBM i partition through a Virtual I/O System (VIOS).

> **Important:** Although PowerVM and VIOS are supported on IBM POWER5, POWER6, and POWER7 systems, IBM i, which is a client of VIOS, is supported only on POWER6 and POWER7 systems.

The information in this chapter is based on code that was available at the time of writing this book. For the latest support information and instructions, see the following resources:

- System Storage Interoperability Center (SSIC)

  http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

- The Host Attachment publications

  http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

2.1.1 Requirements

When attaching an XIV Storage System to an IBM i VIOS client, follow the general requirements in Table 2-1, which were current at the time of writing this paper. These requirements offer guidance about the necessary hardware and software levels for the XIV Storage System with IBM i. For detailed requirements, see the System Storage Interoperation Center at:

http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

Table 2-1 IBM i and VIOS requirements for XIV attachment

XIV attach	Server	VIOS level	IBM i 6.1	IBM i 7.1
VIOS VSCSI	POWER7	2.2 or later	Yes (IBM i 6.1.1)	Yes
VIOS VSCSI	Blade servers based on POWER7 and BladeCenter H Chassis	2.2 or later	Yes (IBM i 6.1.1)	Yes
VIOS VSCSI	IBM POWER6+™	2.1.1 or later	Yes	Yes
VIOS VSCSI	Blade servers based on POWER6 and BladeCenter H Chassis	2.1.1 or later	Yes	Yes
VIOS VSCSI	POWER6	2.1.1 or later	Yes	Yes

For the latest information about the environments that are used when connecting the XIV Storage System to IBM i, see the following websites:

- System i Storage Solutions

  http://www.ibm.com/systems/i/hardware/storage/index.html

- Virtualization with IBM i, PowerVM, and Power Systems

  http://www.ibm.com/systems/i/os/

- IBM Power Systems Hardware Information Center

  http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/iphdx/550_m50_landing.htm

- IBM Power Blade servers

  http://www.ibm.com/systems/power/hardware/blades/index.html
- IBM i and System i Information Center

  http://publib.boulder.ibm.com/iseries/
- IBM Support Portal

  http://www.ibm.com/support/entry/portal/
- System Storage Interoperation Center

  http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

2.1.2 Supported SAN switches

For a list of supported SAN switches when connecting the XIV Storage System to the IBM i operating system, see the System Storage Interoperation Center at:

http://www.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss?start_over=yes

2.1.3 Physical Fibre Channel adapters and virtual SCSI adapters

You can connect up to 4,095 logical unit numbers (LUNs) per target and up to 510 targets per port on a VIOS physical Fibre Channel (FC) adapter. Because you can assign up to 16 LUNs to one virtual SCSI (VSCSI) adapter, you can use the number of LUNs to determine the number of virtual adapters that you need.

> **Important:** When the IBM i operating system and VIOS are on an IBM Power Blade server, you can define only one VSCSI adapter in the VIOS to assign to an IBM i client. Consequently the number of LUNs to connect to the IBM i operating system is limited to 16.

2.1.4 Queue depth in the IBM i operating system and Virtual I/O Server

When connecting the IBM XIV Storage System server to an IBM i client through the VIOS, consider the following types of queue depths:

IBM i queue depth to a virtual LUN
: SCSI command tag queuing in the IBM i operating system that enables up to 32 I/O operations to one LUN at the same time.

Queue depth per physical disk (hdisk) in the VIOS
: Indicates the maximum number of I/O requests that can be outstanding on a physical disk in the VIOS at a particular time.

Queue depth per physical adapter in the VIOS
: Indicates the maximum number of I/O requests that can be outstanding on a physical adapter in the VIOS at a particular time.

The IBM i operating system has a fixed queue depth of 32, which is *not* changeable. However, a user can set up the queue depths in the VIOS. The default setting in the VIOS varies based on the type of connected storage, type of physical adapter, and type of multipath driver or host attachment kit that is used. Typically the XIV Storage System has the following queue depths:

- The queue depth per physical disk is 40.
- The queue depth per 4-Gbps FC adapter is 200.
- The queue depth per 8-Gbps FC adapter is 500.

Check the queue depth on physical disks by entering the following VIOS command:

```
lsdev -dev hdiskxx -attr queue_depth
```

If needed, set the queue depth to 32 by using the following command:

```
chdev -dev hdiskxx -attr queue_depth=32
```

This command ensures that the queue depth in the VIOS matches the IBM i queue depth for an XIV LUN.

2.1.5 Multipath with two Virtual I/O Servers

The IBM XIV Storage System server is connected to an IBM i client partition through the VIOS. For redundancy, you connect the XIV Storage System to an IBM i client with two or more VIOS partitions, with one VSCSI adapter in the IBM i operating system assigned to a VSCSI adapter in each VIOS. The IBM i operating system then establishes a multipath to an XIV LUN, with each path using one VIOS. For XIV attachment to the VIOS, the VIOS integrated native MPIO multipath driver is used. You can use up to eight VIOS partitions in such a multipath connection. However, most installations use multipath with two VIOS partitions.

For more information, see 3.1.3, “IBM i multipath capability with two Virtual I/O Servers” on page 21.

2.2 Best practices

This section presents general guidance for IBM XIV Storage System servers that are connected to a host server. This guidance also applies to the IBM i operating system. With the grid architecture and massive parallelism inherent to XIV system, the general approach is to maximize the utilization of all the XIV resources at all times.

2.2.1 Distributing connectivity

The goal for host connectivity is to create a balance of the resources in the IBM XIV Storage System server. Balance is achieved by distributing the physical connections across the interface modules. A host usually manages multiple physical connections to the storage device for redundancy purposes by using a SAN connected switch. The ideal is to distribute these connections across each interface module. This way, the host uses the full resources of each module to which it connects, for maximum performance.

It is *not* necessary for each host instance to connect to each interface module. However, when the host has more than one physical connection, it is beneficial to have the connections (cabling) spread across separate interface modules.

Similarly, if multiple hosts have multiple connections, you must distribute the connections evenly across the interface modules.

Several configurations using Fibre Channel are technically possible, and they vary in terms of their cost and the degree of flexibility, performance, and reliability that they provide.

Important: Production environments must always have a redundant (high availability) configuration. No single points of failure can exist. Hosts must have as many host bus adapters (HBAs) as needed to support the operating system, application, and overall performance requirements.

The following configurations, all of which have no single point of failure, are possible:

- If a module fails, each host remains connected to all other interface modules.
- If an FC switch fails, each host remains connected to at least three interface modules.
- If an HBA on a host fails, each host remains connected to at least three interface modules.
- If a host cable fails, each host remains connected to at least three interface modules.

Configuration 1: Redundant configuration with 12 paths to each volume

Figure 2-1 illustrates the fully redundant configuration.

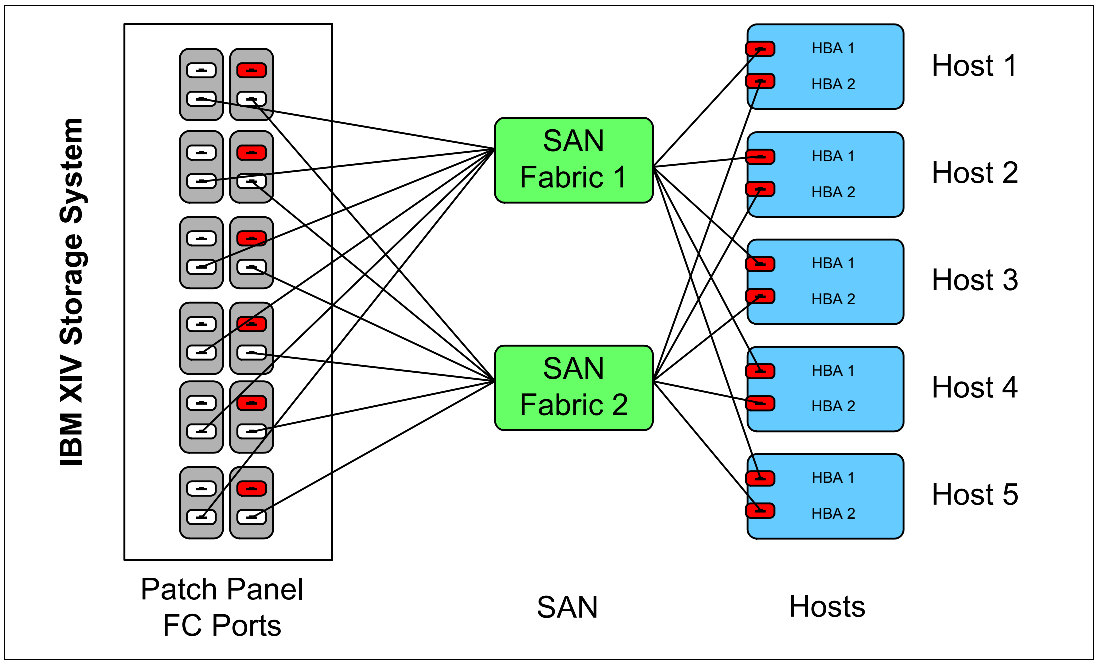


Figure 2-1 FC fully redundant configuration

This configuration has the following setup:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each FC switch has a connection to a separate FC port of each of the six interface modules.
- Each volume can be accessed through 12 paths, which is the maximum number to consider. Using more than 12 paths offers no benefit, because it might cause issues with host processor utilization and server reliability if a path failure occurs (such as the failure of an HBA on a host).

Configuration 2: Redundant configuration with six paths to each volume

Figure 2-2 illustrates a redundant configuration that accesses all interface modules, but uses the ideal number of six paths per LUN on the host.

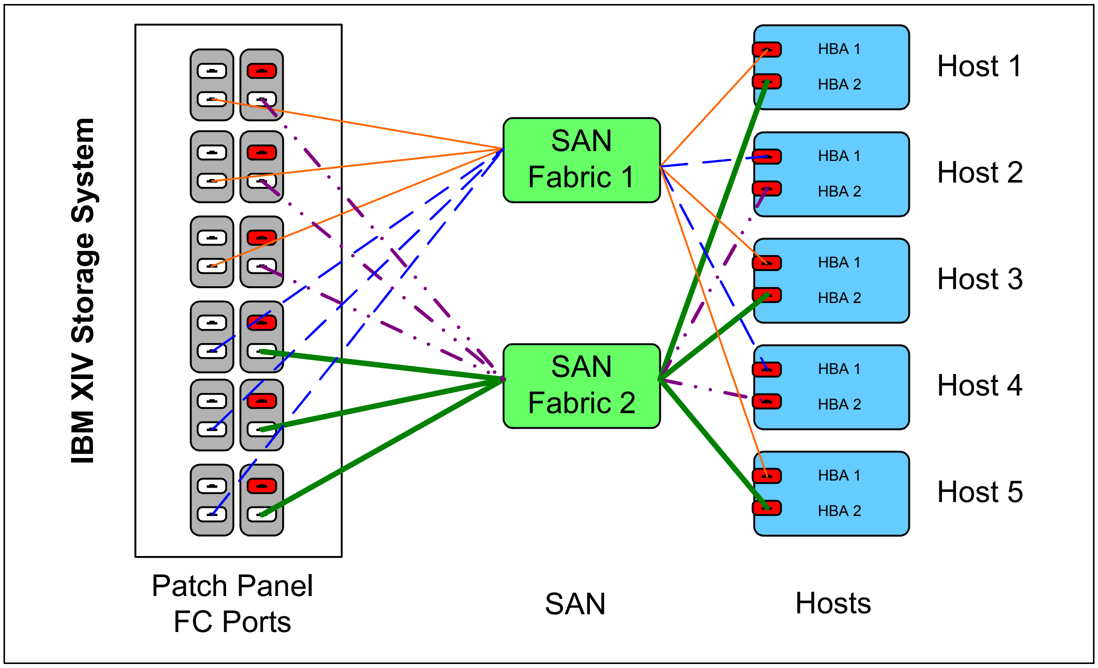


Figure 2-2 FC redundant configuration

This configuration has the following setup:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each FC switch has a connection to a separate FC port of each of the six interface modules.
- One host uses the first three paths per fabric, and the next host uses the three other paths per fabric.
- If a fabric fails, all interface modules are still used.
- Each volume has six paths, making the most ideal configuration.

Configuration 3: Redundant configuration with minimal cabling

An even simpler redundant configuration is illustrated in Figure 2-3.

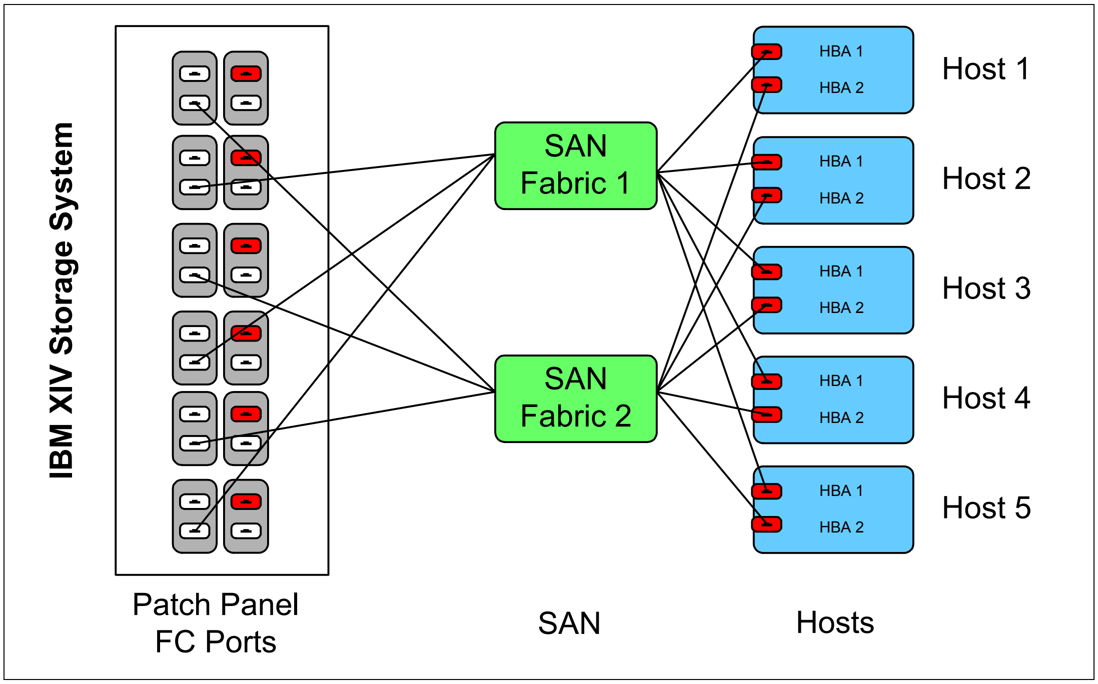


Figure 2-3 FC simple redundant configuration

This configuration has the following setup:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each FC switch has a connection to three separate interface modules.
- Each volume has six paths.

2.2.2 Determining the ideal path count

In each example in this chapter, you can use SAN zoning to control the number of paths configured per volume. Because the XIV can have up to 24 Fibre Channel ports, an administrator might be tempted to configure many paths. Configurations that use up to 48 paths suddenly become possible, which is not a good practice. No performance or reliability benefits are gained by using many paths. In fact, no benefit is gained by going beyond 12 paths per volume and rarely much benefit is gained by going beyond 6 paths per volume. Consequently, aim for 4 or 6 paths per volume as a standard.

2.2.3 Zoning SAN switches

To maximize balancing and distribution of host connections to an IBM XIV Storage System server, create a zone for the SAN switches so that each host adapter connects to each XIV interface module and through each SAN switch.

> **Appropriate zoning**: Use a separate zone for each host adapter (initiator). For each zone that contains the host adapter, add all switch port connections from the XIV Storage System (targets).

2.2.4 Queue depth

SCSI command tag queuing for LUNs on the IBM XIV Storage System server enables multiple I/O operations to one LUN at the same time. The LUN queue depth indicates the number of I/O operations that can be done simultaneously to a LUN.

The XIV architecture eliminates the existing storage concept of a large central cache. Instead, each module in the XIV grid has its own dedicated cache. The XIV algorithms that stage data between disk and cache work most efficiently when multiple I/O requests are coming in parallel. This area is where the host queue depth becomes an important factor in maximizing XIV I/O performance. Therefore, configure the queue depths of the HBA on the host as large as possible.

2.2.5 Number of application threads

The overall design of the IBM XIV Storage System grid architecture excels with applications that employ threads to handle the parallel execution of I/Os. The multithreaded applications will profit the most from XIV performance.

3

Implementing the IBM XIV Storage System server with IBM i

This chapter uses examples to illustrate the tasks for configuring the IBM XIV Storage System server for use with the IBM i operating system. In the examples, all of the IBM i disk space, including the load source disk (boot disk), consists of the XIV logical unit numbers (LUNs).

The LUNs are connected to the IBM i partition in multipath with two Virtual I/O Servers (VIOSs). For this setup, see 3.1, "Connecting a PowerVM IBM i client to XIV" on page 16.

In addition, this chapter describes the following tasks:

- Creating a storage pool on the XIV system
- Defining XIV volumes (LUNs) for the IBM i operating system and connecting them to both VIOS partitions
- Assigning the corresponding VIOS disk devices to virtual SCSI (VSCSI) adapters in the IBM i client partition
- Installing the IBM i operating system on XIV volumes

This chapter includes the following sections:

- Connecting a PowerVM IBM i client to XIV
- Mapping the volumes in the Virtual I/O Server
- Installing the IBM i client

3.1 Connecting a PowerVM IBM i client to XIV

The XIV system can be connected to an IBM i partition through the VIOS. The following sections explain how to set up the environment on a POWER6 system to connect the XIV system to an IBM i client with multipath through two VIOS partitions. Setting up a POWER7 system to an XIV Storage System is similar.

3.1.1 Creating the Virtual I/O Server and IBM i partitions

This section explains how to create a VIOS partition and an IBM i partition through the POWER6 Hardware Management Console (HMC). It also explains how to create VSCSI adapters in the VIOS and the IBM i partition and how to assign them so that the IBM i partition can work as a client of the VIOS.

For more information about creating the VIOS and IBM i client partitions in the POWER6 processor-based server, see 6.2.1, “Creating the VIOS LPAR,” and 6.2.2, “Creating the IBM i LPAR,” in the IBM Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

Creating a Virtual I/O Server partition in a POWER6 server

To create a POWER6 logical partition (LPAR) for the VIOS:

1. Insert the PowerVM activation code in the HMC if you have not already done so. Select **Tasks** → **Capacity on Demand (CoD)** → **Advanced POWER Virtualization** → **Enter Activation Code**.
2. Create the partition. In the left pane, select **Systems Management** → **Servers**. In the right pane, select the server to use for creating the VIOS partition. Then select **Tasks** → **Configuration** → **Create Logical Partition** → **VIO Server** (Figure 3-1).

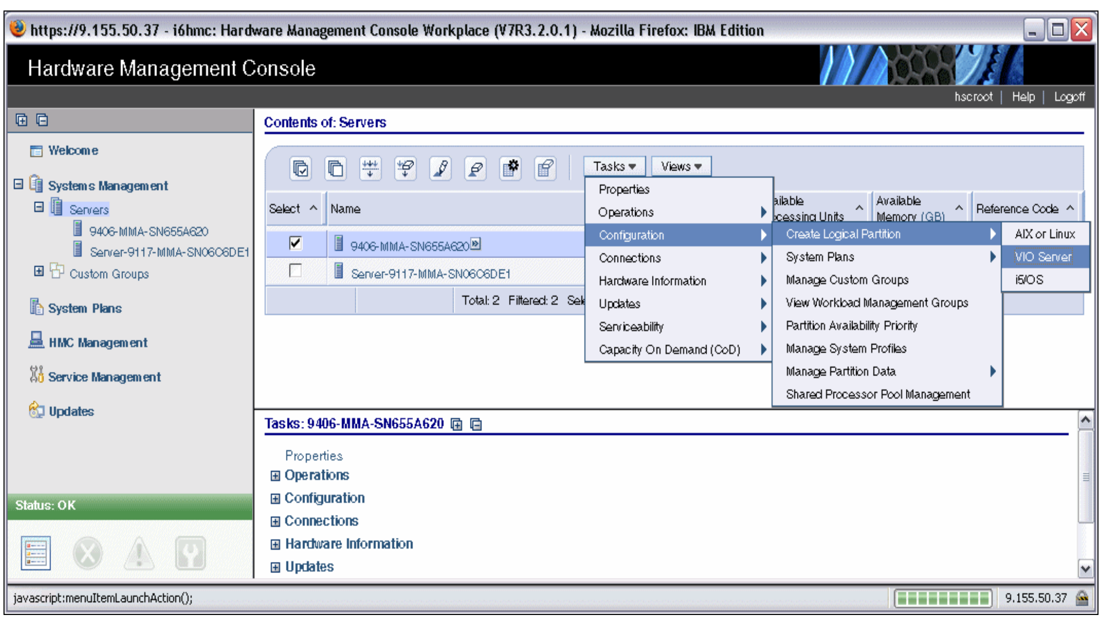


Figure 3-1 Creating the VIOS partition

3. In the Create LPAR wizard:
   a. Enter the partition ID and name.
   b. Enter the partition profile name.
   c. Select whether the processors in the LPAR will be dedicated or shared. Consider selecting **Dedicated**.
   d. Specify the minimum, desired, and maximum number of processors for the partition.
   e. Specify the minimum, desired, and maximum amount of memory in the partition.
4. In the I/O panel (Figure 3-2), select the I/O devices to include in the new LPAR. In our example, we include the RAID controller to attach the internal SAS drive for the VIOS boot disk and DVD_RAM drive. We include the physical Fibre Channel (FC) adapters to connect to the XIV server. As shown in Figure 3-2, we add them as *Required*.

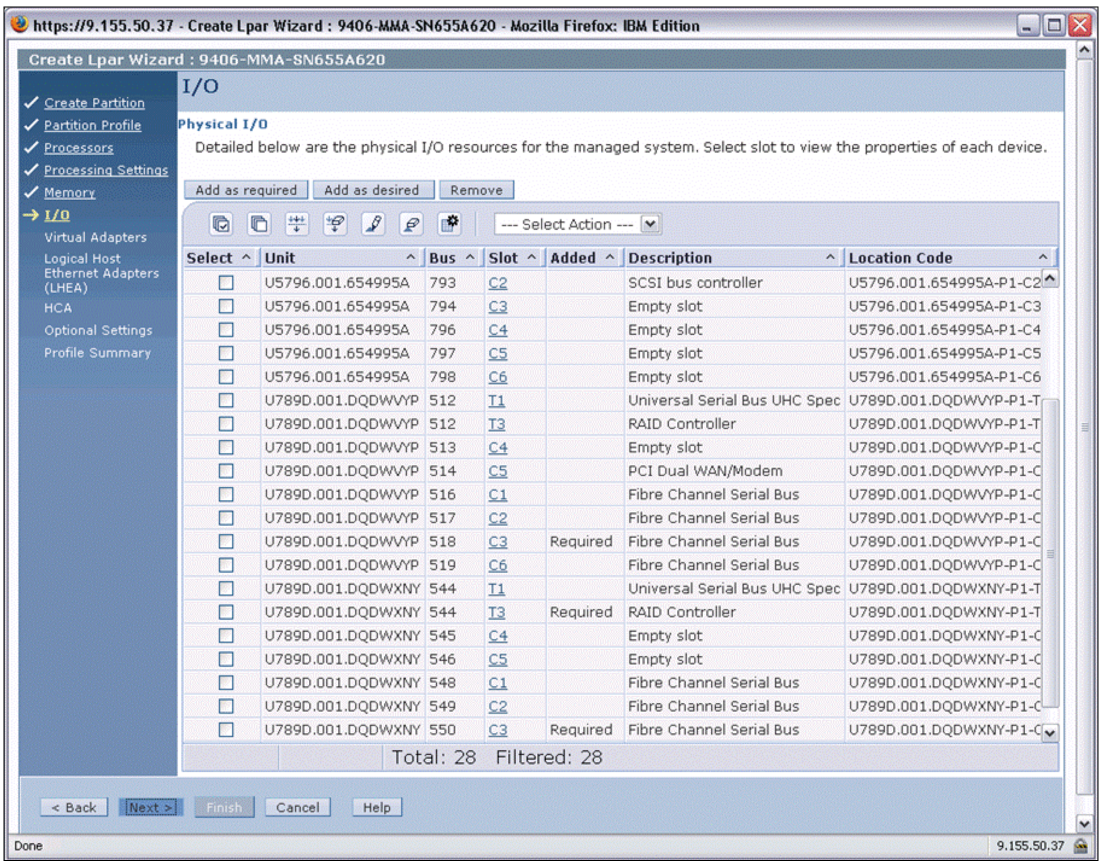


Figure 3-2 Adding the I/O devices to the VIOS partition

5. In the Virtual Adapters panel, create an Ethernet adapter by selecting **Actions** → **Create** → **Ethernet Adapter**. Mark it as **Required**.
6. Create the VSCSI adapters that will be assigned to the virtual adapters in the IBM i client:
   a. Select **Actions** → **Create** → **SCSI Adapter**.
   b. In the next window, leave the **Any Client partition can connect** selected, or limit the adapter to a particular client.

   If DVD-RAM will be virtualized to the IBM i client, you might want to create another VSCSI adapter for DVD-RAM.

7. Configure the logical host Ethernet adapter:
   a. Select the logical host Ethernet adapter from the list.
   b. In the next window, click **Configure**.
   c. Verify that the selected logical host Ethernet adapter is not selected by any other partitions, and select **Allow all VLAN IDs**.
8. In the Profile Summary panel, review the information, and then click **Finish** to create the LPAR.

Creating an IBM i partition in the POWER6 processor-based server

To create an IBM i partition (that will be the VIOS client):

1. From the HMC, select **Systems Management** → **Servers**. In the right panel, select the server in which you want to create the partition. Then select **Tasks** → **Configuration** → **Create Logical Partition** → **i5/OS**.
2. In the Create LPAR wizard:
   a. Enter the Partition ID and name.
   b. Enter the partition Profile name.
   c. Select whether the processors in the LPAR will be dedicated or shared. Consider selecting **Dedicated**.
   d. Specify the minimum, desired, and maximum number of processors for the partition.
   e. Specify the minimum, desired, and maximum amount of memory in the partition.
   f. In the I/O panel, if the IBM i client partition is not supposed to own any physical I/O hardware, click **Next**.
3. In the Virtual Adapters panel, select **Actions** → **Create** → **Ethernet Adapter** to create a virtual Ethernet adapter.
4. In the Create Virtual Ethernet Adapter panel, accept the suggested adapter ID and the VLAN ID. Select **This adapter is required for partition activation**, and then click **OK** to continue.
5. Still in the Virtual Adapters panel, select **Actions** → **Create** → **SCSI Adapter** to create the VSCSI client adapters on the IBM i client partition that is used for connecting to the corresponding VIOS.
6. For the VSCSI client adapter ID, specify the ID of the adapter:
   a. For the type of adapter, select **Client**.
   b. Select **Mark this adapter is required for partition activation**.
   c. Select the VIOS partition for the IBM i client.
   d. Enter the server adapter ID to which you want to connect the client adapter.
   e. Click **OK**.

   If necessary, you can repeat this step to create another VSCSI client adapter to connect to the VIOS VSCSI server adapter that is used to virtualize the DVD-RAM.
7. Configure the logical host Ethernet adapter:
   a. Select the logical host Ethernet adapter from the list, and click **Configure**.
   b. In the next panel, ensure that no other partitions have selected the adapter, and select **Allow all VLAN IDs**.
8. In the OptiConnect Settings panel, if OptiConnect is not used in IBM i, click **Next**.

9. In the Load Source Device panel, if the connected XIV system will be used to boot from a storage area network (SAN), select the virtual adapter that connects to the VIOS.

IBM i Load Source Device: The IBM i Load Source device is on an XIV volume.

10. In the Alternate Restart Device panel, if the virtual DVD-RAM device will be used in the IBM i client, select the corresponding virtual adapter.
11. In the Console Selection panel, select the default of HMC for the console device. Click **OK**.
12. Depending on the planned configuration, click **Next** in the three panels that follow until you reach the Profile Summary panel.
13. In the Profile Summary panel, review the specified configuration, and then click **Finish** to create the IBM i LPAR.

3.1.2 Installing the Virtual I/O Server

This section explains how to install the VIOS. For information about how to install the VIOS in a partition of the POWER6 processor-based server, see the Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

If the disks that you are going to use for the VIOS installation were previously used by an IBM i partition, you must reformat them for 512 bytes per sector.

Installing the Virtual I/O Server software

To install the VIOS software:

1. Insert the VIOS installation disk into the DVD drive of the VIOS partition.
2. Activate the VIOS partition:
   a. In the left pane, select **Systems Management** → **Servers**, and then select the server.
   b. Select the VIOS partition, and then select **Operations** → **Activate**.
3. In the Activate Logical Partition window (Figure 3-3):
   a. Select the partition profile.
   b. Select the **Open a terminal window or console session** check box.
   c. Click **Advanced**.

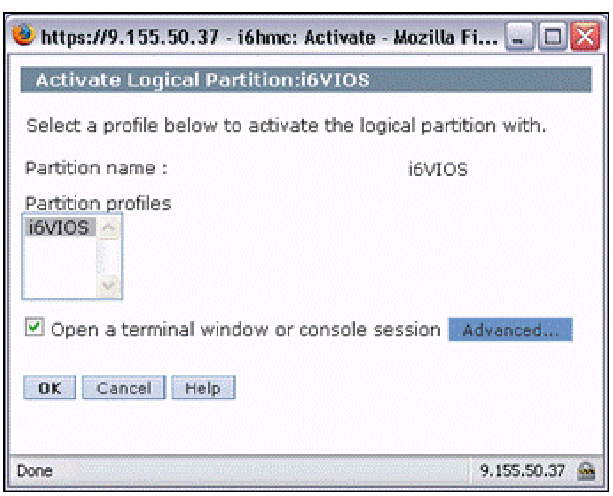


Figure 3-3 Activate LPAR window

4. In the next window, choose boot mode **SMS** and click **OK**. Then, click **OK** to activate the VIOS partition in SMS boot mode when a terminal session window is open.

> **Console window:** Perform the following steps in the console window.

5. In the SMS main menu (Figure 3-4), select option **5. Select Boot Options**.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Main Menu
1.   Select Language
2.   Setup Remote IPL (Initial Program Load)
3.   Change SCSI Settings
4.   Select Console
5.   Select Boot Options



-------------------------------------------------------------------------------
Navigation Keys:

                                       X = eXit System Management Services


-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:5
```

Figure 3-4 SMS main menu

6. From the Multiboot menu, select option **1. Select Install/Boot Device**.
7. From the Select Device Type menu, select option **3. CD/DVD**.
8. From the Select Media Type menu, select option **5. SATA**.
9. From the Select Media Adapter menu, select the media adapter.
10. From the Select Device menu, select the device **SATA CD-ROM**.
11. From the Select Task menu, select option **2. Normal boot mode**.
12. In the next console window, when prompted by the question “Are you sure you want to exit System Management Services?”, select option **1. Yes**.

    The VIOS partition is automatically rebooted in normal boot mode.
13. In the VIOS installation panel that opens, at the prompt for the system console, type `1` to use this terminal as the system console.
14. At the installation language prompt, type `1` to use English as the installation language.
15. In the VIOS welcome panel, select option **1 Start Install Now with Default Settings**.
16. From the VIOS System Backup Installation and Settings menu, select option **0 Install with the settings listed above**.

    Installation of the VIOS starts, and its progress is shown in the Installing Base Operation System panel.
17. After successful installation of the VIOS, log in with the prime administrator user ID `padmin`. Enter a new password and type `a` to accept the software terms and conditions.

18. Before you run any VIOS command other than the **chlang** command to change the language setting, accept the software license terms by entering the following command:

```
license -accept
```

To view the license before you accept it, enter the following command:

```
license -view
```

Using LVM mirroring for the Virtual I/O Server

Set up LVM mirroring to mirror the VIOS root volume group (rootvg). In our example, we mirror it across two RAID 0 arrays of hdisk0 and hdisk1 to help protect the VIOS from potential POWER6 server internal SAS disk drive failures.

Configuring Virtual I/O Server network connectivity

To set up network connectivity in the VIOS:

1. In the HMC terminal window, logged in as padmin, enter the following command:

```
lsdev -type adapter | grep ent
```

Look for the logical host Ethernet adapter resources. In our example, it is ent1, as shown in Figure 3-5.

```
$ lsdev -type adapter | grep ent
ent0            Available   Virtual I/O Ethernet Adapter (l-lan)
ent1            Available   Logical Host Ethernet Port (lp-hea)
```

Figure 3-5 Available logical host Ethernet port

2. Configure TCP/IP for the logical Ethernet adapter entX by using the **mktcpip** command syntax and specifying the corresponding interface resource enX.
3. Verify the created TCP/IP connection by pinging the IP address that you specified in the **mktcpip** command.

Upgrading the Virtual I/O Server to the latest fix pack

As the last step of the installation, upgrade the VIOS to the latest fix pack.

3.1.3 IBM i multipath capability with two Virtual I/O Servers

The IBM i operating system provides multipath capability for allowing access to an XIV volume (LUN) through multiple connections. One path is established through each connection. Up to eight paths to the same LUN or set of LUNs are supported. Multipath provides redundancy in case a connection fails, and it increases performance by using all available paths for I/O operations to the LUNs.

With Virtual I/O Server release 2.1.2 or later, and IBM i release 6.1.1 or later, you can establish multipath to a set of LUNs, with each path using a connection through a different VIOS. This topology provides redundancy in case either a connection or the VIOS fails. Up to eight multipath connections can be implemented to the same set of LUNs, each through a different VIOS. However, we expect that most IT centers will establish no more than two such connections.

3.1.4 Virtual SCSI adapters in multipath with two Virtual I/O Servers

In our setup, we use two VIOS and two VSCSI adapters in the IBM i partition, where each adapter is assigned to a virtual adapter in one VIOS. We connect the same set of XIV LUNs to each VIOS through two physical FC adapters in the VIOS multipath and map them to VSCSI adapters serving IBM i partition. This way, the IBM i partition sees the LUNs through two paths, each path by using one VIOS. Therefore, multipathing is started for the LUNs. Figure 3-6 shows our setup.

For our testing, we did not use separate switches as shown in Figure 3-6, but rather we used separate blades in the same SAN Director. In a real, production environment, use separate switches as shown in Figure 3-6.

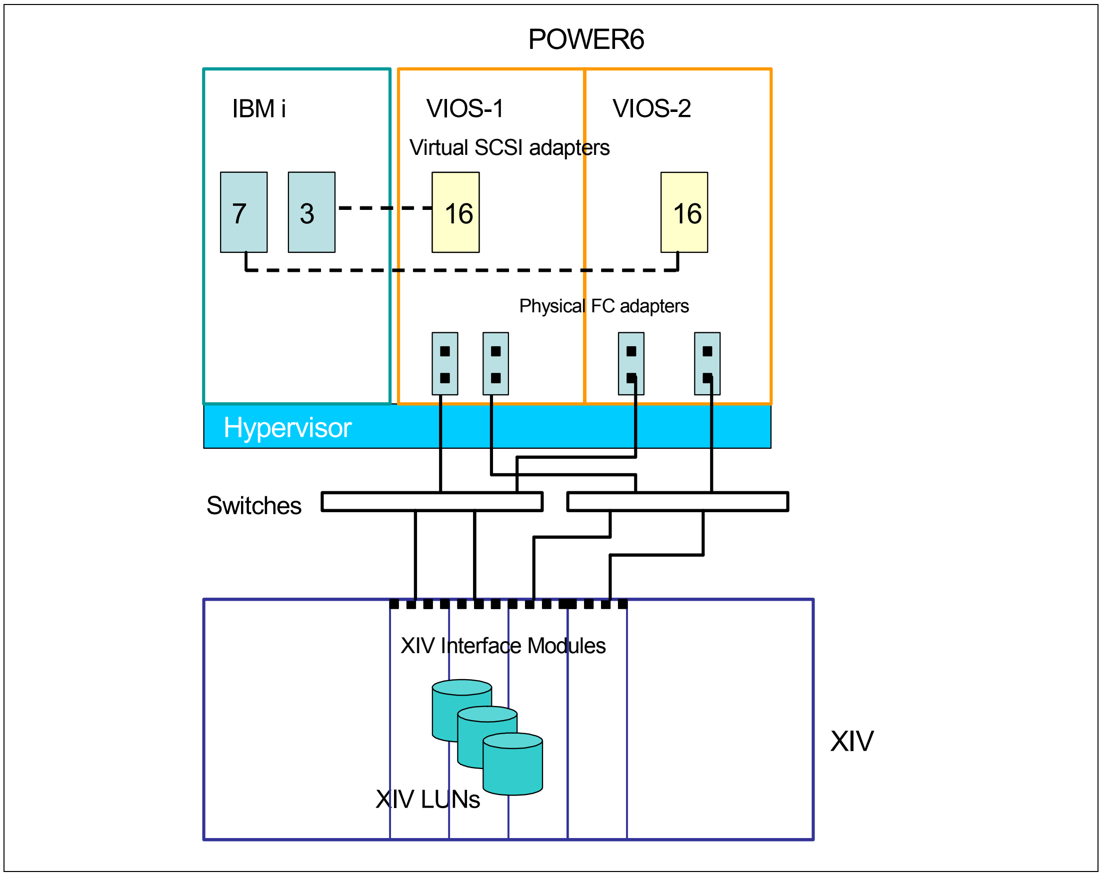


Figure 3-6 Setup for multipath with two VIOS

To connect XIV LUNs to an IBM i client partition in multipath with two VIOS:

Important: Perform steps 1 through 5 in each of the two VIOS partitions.

1. After the LUNs are created in the XIV system, use the XIV Storage Management GUI or Extended Command Line Interface (XCLI) to map the LUNs to the VIOS host. For more information, see 3.2, “Mapping the volumes in the Virtual I/O Server” on page 24.
2. Log in to VIOS as administrator. In our example, we use PuTTY to log in as described in 6.5, “Configuring VIOS virtual devices,” of the Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

   Type the **`cfgdev`** command so that the VIOS can recognize newly attached LUNs.

3. In the VIOS, remove the SCSI reservation attribute from the LUNs (hdisks) that will be connected through two VIOS by entering the following command for each hdisk that will connect to the IBM i operating system in multipath:

```
chdev -dev hdiskX -attr reserve_policy=no_reserve
```

4. To get more bandwidth by using multiple paths, enter the following command for each hdisk (hdiskX):

```
chdev -dev hdiskX -perm -attr algorithm=round_robin
```

5. Set the queue depth in VIOS for IBM i to a value of 32 or higher. The default value for XIV Storage Systems is 40 in VIOS. Higher values use more memory. A value of 40 is the recommended value for AIX, Linux, and IBM i under VIOS.
6. Verify the attributes:

```
lsdev -dev hdiskX -attr
```

Figure 3-7 shows this command and its output.

```
$ lsdev -dev hdisk94 -attr
attribute       value                              description
user_settable

PCM             PCM/friend/fcpother                Path Control Module               False
algorithm       round_robin                        Algorithm                         True
clr_q           no                                 Device CLEARS its Queue on error True
dist_err_pcnt   0                                  Distributed Error Percentage      True
dist_tw_width   50                                 Distributed Error Sample Time     True
hcheck_cmd      inquiry                            Health Check Command              True
hcheck_interval 60                                 Health Check Interval             True
hcheck_mode     nonactive                          Health Check Mode                 True
location                                           Location Label                    True
lun_id          0x1000000000000                    Logical Unit Number ID            False
lun_reset_spt   yes                                LUN Reset Supported               True
max_retry_delay 60                                 Maximum Quiesce Time              True
max_transfer    0x40000                            Maximum TRANSFER Size             True
node_name       0x5001738000cb0000                 FC Node Name                      False
pvid            none                               Physical volume identifier        False
q_err           yes                                Use QERR bit                      True
q_type          simple                             Queuing TYPE                      True
queue_depth     40                                 Queue DEPTH                       True
reassign_to     120                                REASSIGN time out value           True
reserve_policy  no_reserve                         Reserve Policy                    True
rw_timeout      30                                 READ/WRITE time out value         True
scsi_id         0xa1400                            SCSI ID                           False
start_timeout   60                                 START unit time out value         True
unique_id       261120017380000CB1797072810XIV03IBMfcp Unique device identifier      False
ww_name         0x5001738000cb0150                 FC World Wide Name                False
$
```

Figure 3-7 Output for the lsdev --dev hdiskx -attr command

7. Map the disks that correspond to the XIV LUNs to the VSCSI adapters that are assigned to the IBM i client. First, check the IDs of assigned virtual adapters. Then complete the following steps:
   a. In the HMC, open the partition profile of the IBM i LPAR, click the **Virtual Adapters** tab, and observe the corresponding VSCSI adapters in the VIOS.
   b. In the VIOS, look for the device name of the virtual adapter that is connected to the IBM i client. You can use the **`lsmap -all`** command to view the virtual adapters.
   c. Map the disk devices to the SCSI virtual adapter that is assigned to the SCSI virtual adapter in the IBM i partition by entering the following command:

      ```
      mkvdev -vdev hdiskxx -vadapter vhostx
      ```

Upon completing these steps, in each VIOS partition, the XIV LUNs report in the IBM i client partition by using two paths. The resource name of disk unit that represents the XIV LUN starts with `DMPxxx`, which indicates that the LUN is connected in a multipath.

3.2 Mapping the volumes in the Virtual I/O Server

For more information about the XIV Storage management GUI and XCLI, see the Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

After the XIV volumes are defined, add them to both VIOS partitions. To make them available for the IBM i client, perform the following tasks in each VIOS:

1. Connect to the VIOS partition. In our example, we use a PuTTY session to connect.
2. In the VIOS, enter the **`cfgdev`** command to discover the newly added XIV LUNs, making the LUNs available as disk devices (hdisks) in the VIOS. In our example, the LUNs that we added correspond to hdisk132 - hdisk140, as shown in Figure 3-8.

```
9.11.230.185 - PuTTY
hdisk122        Available   IBM 2810XIV Fibre Channel Disk
hdisk123        Available   IBM 2810XIV Fibre Channel Disk
hdisk124        Available   IBM 2810XIV Fibre Channel Disk
hdisk125        Available   IBM 2810XIV Fibre Channel Disk
hdisk126        Available   IBM 2810XIV Fibre Channel Disk
hdisk127        Available   IBM 2810XIV Fibre Channel Disk
hdisk128        Available   IBM 2810XIV Fibre Channel Disk
hdisk129        Available   IBM 2810XIV Fibre Channel Disk
hdisk130        Available   IBM 2810XIV Fibre Channel Disk
hdisk131        Available   IBM 2810XIV Fibre Channel Disk
hdisk132        Available   IBM 2810XIV Fibre Channel Disk
hdisk133        Available   IBM 2810XIV Fibre Channel Disk
hdisk134        Available   IBM 2810XIV Fibre Channel Disk
hdisk135        Available   IBM 2810XIV Fibre Channel Disk
hdisk136        Available   IBM 2810XIV Fibre Channel Disk
hdisk137        Available   IBM 2810XIV Fibre Channel Disk
hdisk138        Available   IBM 2810XIV Fibre Channel Disk
hdisk139        Available   IBM 2810XIV Fibre Channel Disk
hdisk140        Available   IBM 2810XIV Fibre Channel Disk
hdisk196        Available   IBM 2810XIV Fibre Channel Disk
hdisk197        Available   IBM 2810XIV Fibre Channel Disk
hdisk198        Available   IBM 2810XIV Fibre Channel Disk
hdisk199        Available   IBM 2810XIV Fibre Channel Disk
```

Figure 3-8 The hdisks of the added XIV volumes

As previously explained, to realize a multipath setup for IBM i, we connected (mapped) each XIV LUN to both VIOS partitions. Before assigning these LUNs (from any of the VIOS partitions) to the IBM i client, make sure that the volume is not SCSI reserved.

3. Because a SCSI reservation is the default in the VIOS, change the reservation attribute of the LUNs to *non-reserved*. First, check the current reserve policy by entering the following command:

   ```
   lsdev -dev hdiskx -attr reserve_policy
   ```

   Here *hdiskx* represents the XIV LUN.

   If the reserve policy is not `no_reserve`, change it to `no_reserve` by entering the following command:

   ```
   chdev -dev hdiskX -attr reserve_policy=no_reserve
   ```

4. Before mapping hdisks to a VSCSI adapter, check whether the adapter is assigned to the client VSCSI adapter in IBM i and whether any other devices are mapped to it.

   a. Enter the following command to view the virtual slot of the adapter and to see any other devices assigned to it:

   ```
   lsmap -vadapter <name>
   ```

   In our setup, no other devices are assigned to the adapter, and the relevant slot is C16 (Figure 3-9).

   ```
   $ lsmap -vadapter vhost5
   SVSA            Physloc
   --------------- -------------------------------------------- --
   vhost5          U9117.MMA.10AF384-V2-C16

   VTD                   NO VIRTUAL TARGET DEVICE FOUND
   ```

   Figure 3-9 Virtual SCSI adapter in the VIOS

   b. From the HMC, edit the profile of the IBM i partition. Select the partition, and choose **Configuration** → **Manage Profiles**. Then select the profile and click **Actions** → **Edit**.

c. In the partition profile, click the **Virtual Adapters** tab, and make sure that a client VSCSI adapter is assigned to the server adapter with the same ID as the virtual slot number. In our example, client adapter 3 is assigned to server adapter 16 (thus matching the virtual slot C16) as shown in Figure 3-10.

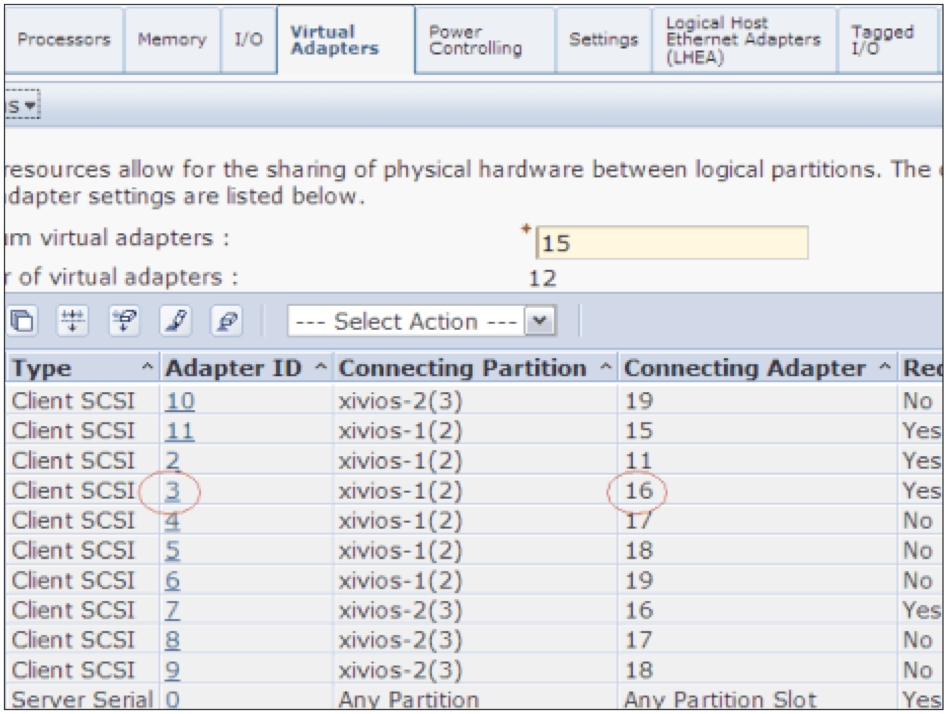


Figure 3-10 Assigned virtual adapters

5. Map the relevant hdisks to the VSCSI adapter by entering the following command:

   `mkvdev -vdev hdiskx -vadapter <name>`

   In our example, we map the XIV LUNs to the adapter vhost5, and we give to each LUN the virtual device name by using the `-dev` parameter as shown in Figure 3-11.

```
$ mkvdev -vdev hdisk132 -vadapter vhost5 -dev vadamaboot132
vadamaboot132 Available
```

Figure 3-11 Mapping the LUNs in the VIOS

After completing these steps for each VIOS, the LUNs are available to the IBM i client in a multipath (one path through each VIOS).

3.2.1 Using the HMC to map volumes to an IBM i client

As an alternative to mapping volumes by using the VIOS commands, you can assign the volumes to the IBM i partition from the POWER6 HMC. This function is available on HMC level EM340-095 from May 2009 or later.

To map the volumes to an IBM i client:

Important: To have a multipath to the IBM i client LUNS through two VIOS, you must perform steps 1 through 4 for each VIOS.

1. Connect to VIOS, discover volumes, and change the reserve_policy attribute of the LUNs. You *cannot* use HMC to do this.
2. In the HMC window, expand **Systems Management** and click **Servers**. In the right pane, select the server on which the VIOS and IBM i client partitions are located, expand the menu at the server, and select **Configuration** → **Virtual Resources** → **Virtual Storage Management** (Figure 3-12).

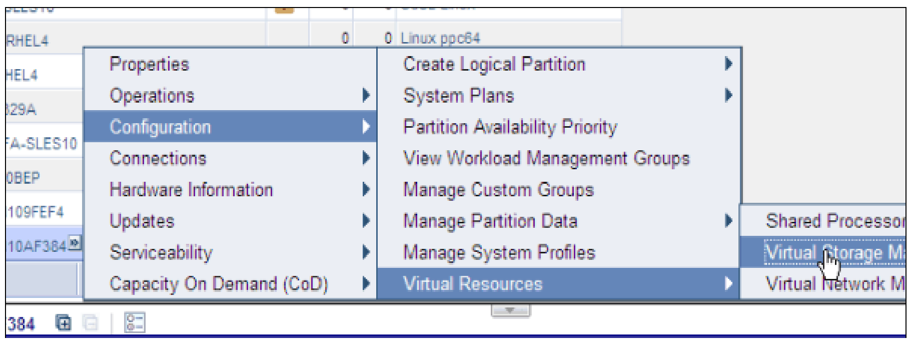


Figure 3-12 HMC: Selecting Virtual Storage Management for the selected server

3. In the Virtual Storage Management window (Figure 3-13), select the VIOS to which to assign volumes. Click **Query VIOS**.

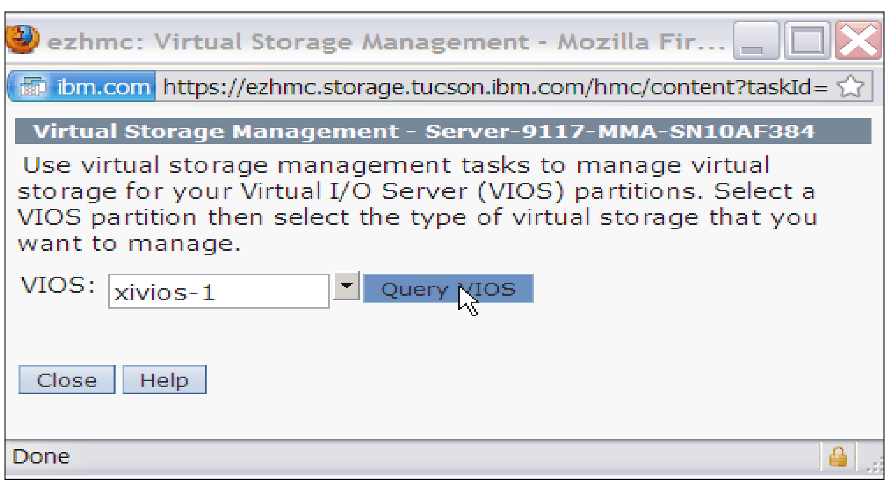


Figure 3-13 HMC: Clicking Query VIOS

4. After the Query VIOS results are shown, click the **Physical Volumes** tab to view the disk devices in the VIOS (Figure 3-14).
   a. Select the hdisk that you want to assign to the IBM i client, and click **Modify assignment**.

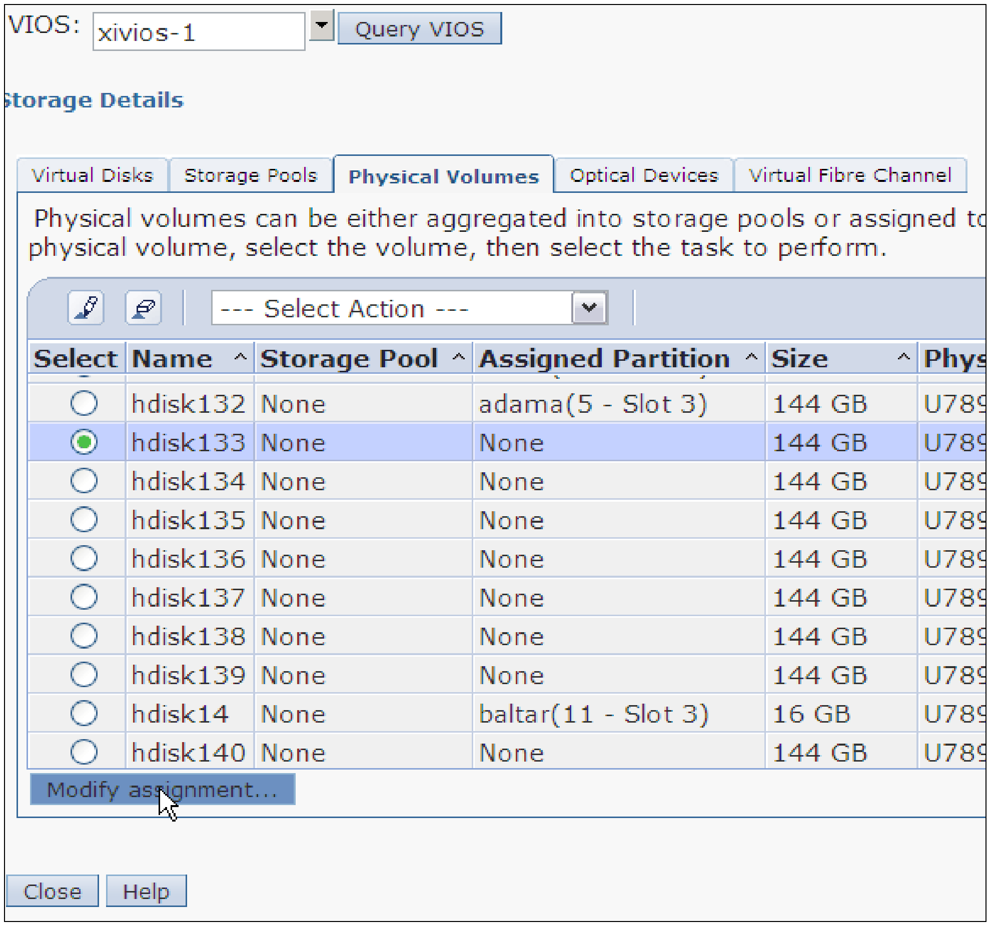


Figure 3-14 HMC: Modifying the assignment of physical volumes in the VIOS

b. In the Modify Physical Disk Partition Assignments window (Figure 3-15), from the list, choose the IBM i client to which you want to assign the volumes.

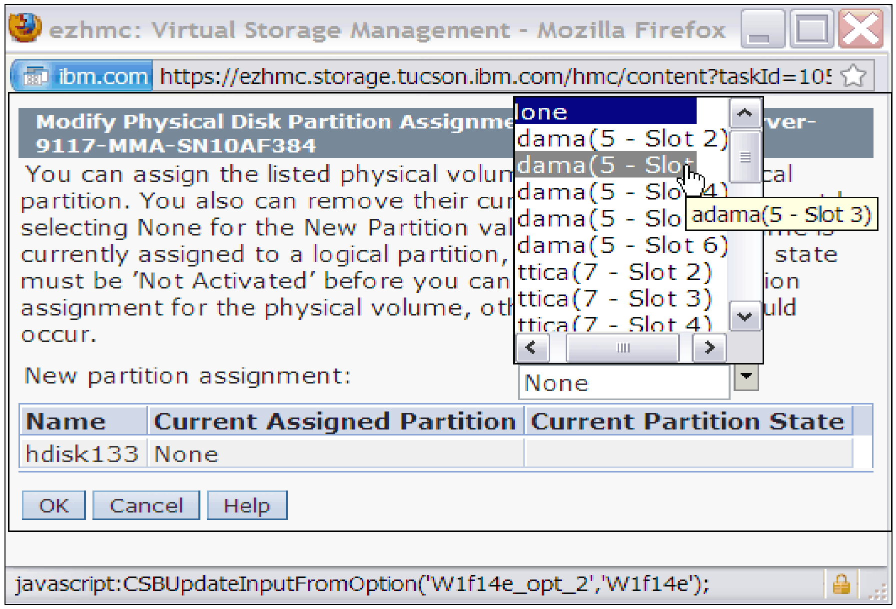


Figure 3-15 HMC: Selecting the partition to which to assign LUNs

c. Click **OK** to confirm the selection for the selected partition (Figure 3-16).

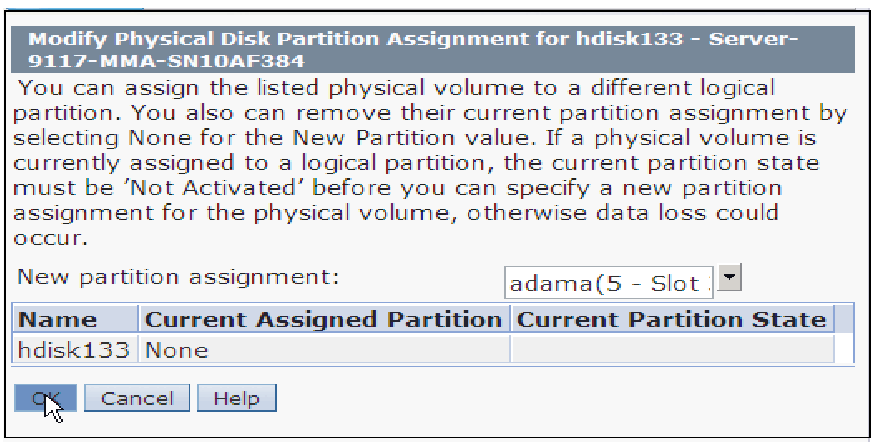


Figure 3-16 HMC: Confirming to assign to the selected partition

You now see the information message indicating that the hdisk is being reassigned (Figure 3-17). The Query VIOS function starts automatically again.

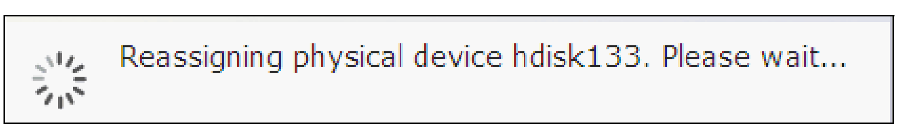


Figure 3-17 HMC: Reassigning the physical device

The newly assigned volumes are displayed in the query results.

3.3 Installing the IBM i client

To install the IBM i client:

1. Insert the installation media *Base_01 6.1.1* in the DVD drive of the client IBM i partition.

   > **Installing IBM i:** To install IBM i, you can use the DVD drive that is dedicated to the IBM i client partition or the virtual DVD drive that is assigned in the VIOS. If you are using the virtual DVD drive, insert the installation media in the corresponding physical drive in the VIOS.

2. In the IBM i partition, make sure that the tagged adapter for load source disk points to the adapter in the VIOS to which the XIV LUNs are assigned. Also, ensure that the tagged adapter for Alternate restart device points to the server VSCSI adapter with the assigned optical drive or to the physical adapter with DVD-RAM.

   To check the tagged adapter:

   a. Select the partition in HMC, and select **Configuration** → **Manage Profiles**.

   b. Select the profile and select **Actions** → **Edit**.

   c. In the partition profile, click the **Tagged I/O** tab.

   As shown in Figure 3-18, we use client adapter 3 for the load source and client adapter 11 for the alternate installation device.

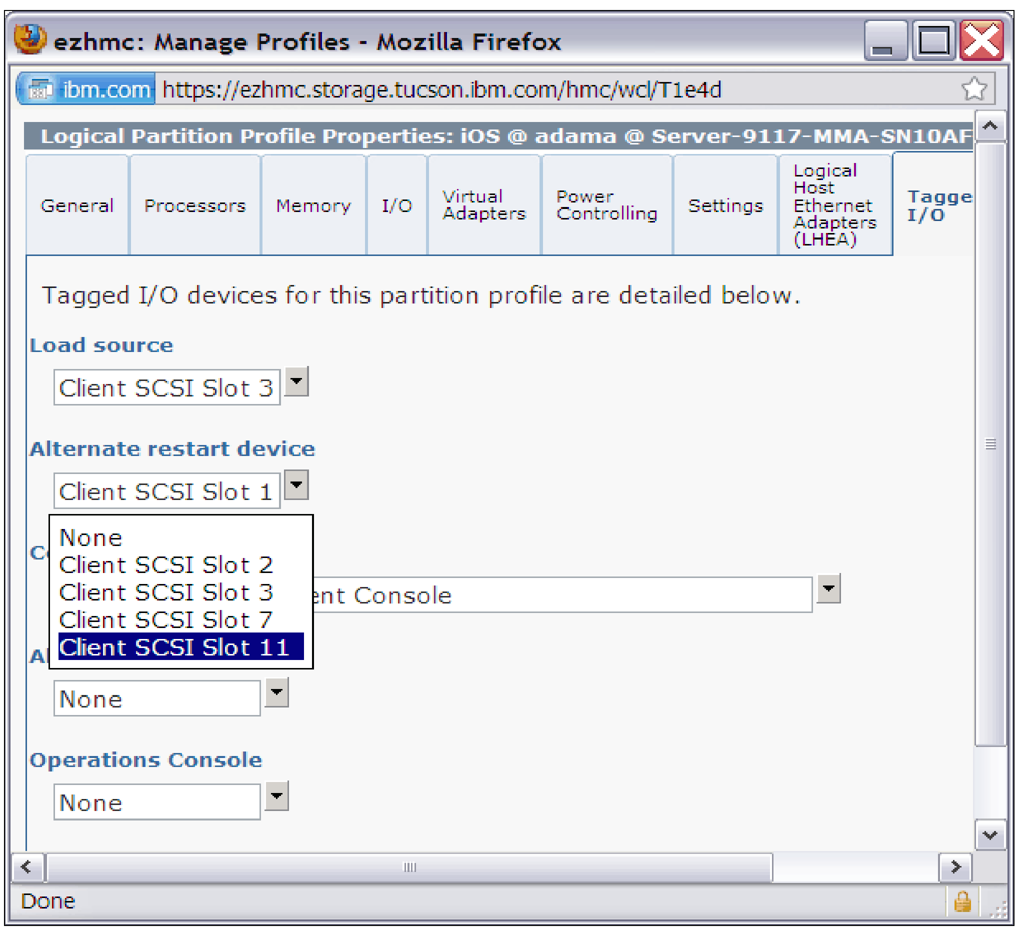


Figure 3-18 Tagged adapters in the IBM i client

d. Still in the partition profile, click the **Virtual Adapters** tab, and verify that the corresponding server virtual adapters are the ones with the assigned volumes and DVD drive. See Figure 3-10 on page 26.

3. In the IBM i client partition, make sure that IPL source is set to **D** and that Keylock position is set to **Manual**.

   To verify these settings, select the IBM i client partition in HMC, choose **Properties**, and click the **Settings** tab to see the currently selected values. Figure 3-19 shows the settings in the client partition used for this example.

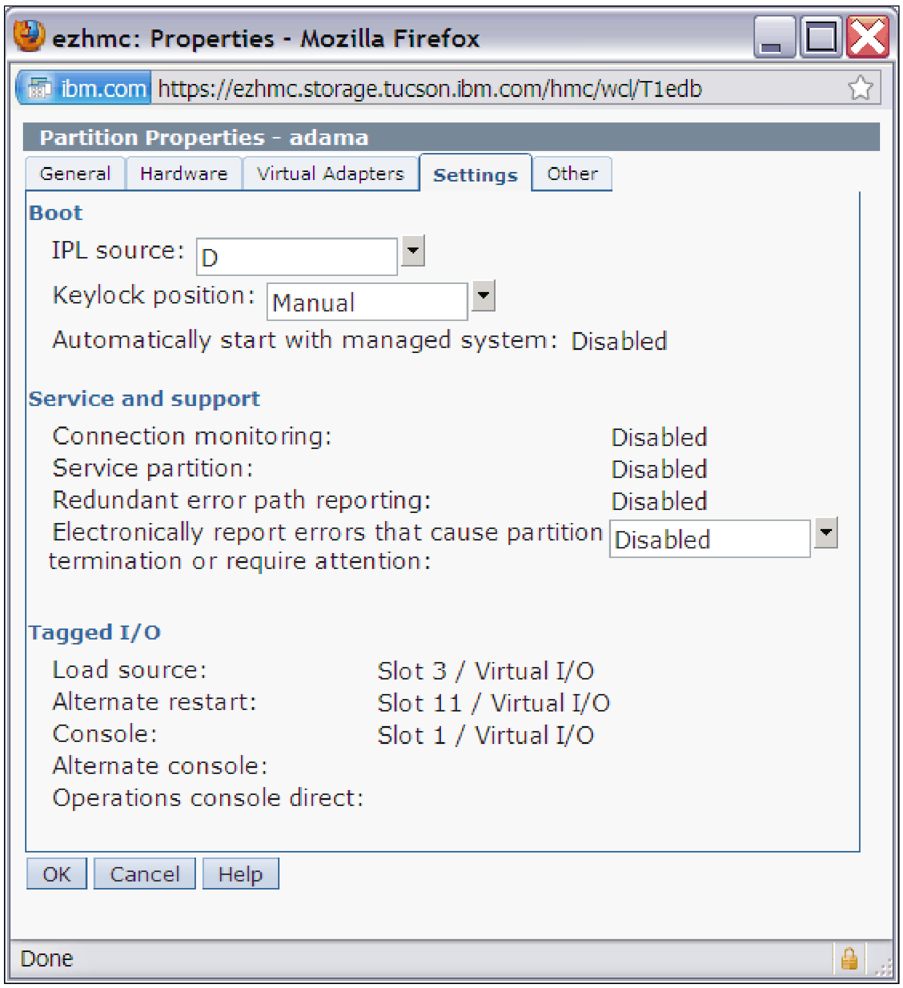


Figure 3-19 Settings in the client partition

4. To activate the IBM i client partition, select this partition in the HMC, and then select **Activate**. You can open the console window in the HMC.

   Alternatively with the IBM Personal Communications tool, you can use Telnet to connect to HMC port 2300 (Figure 3-20).

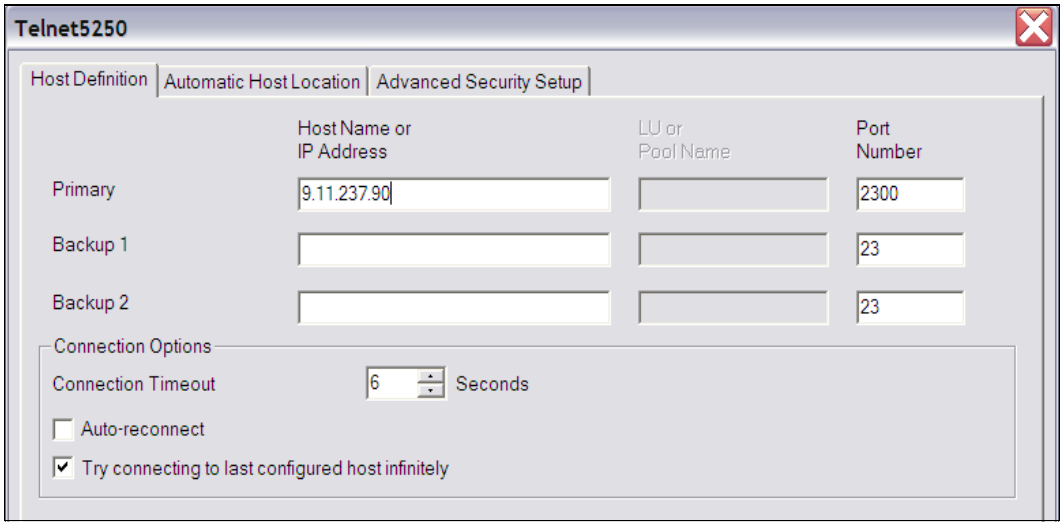


Figure 3-20 Using Telnet to connect to the HMC

5. Select the appropriate language, server, and partition (Figure 3-21).

```
                 HMC Remote 5250 Console System Selection
                                HMC: ezhmc

 Select one of the following and press Enter:


 Option  System Name       Type      Model     Serial      State
  1      Fibonacci-SN1003A 9123      710       1003A0A     Operating
  2      Cauchy-SN10020EA- 9123      710       10020EA     Operating
  3      Newton-SN10D95AC- 9123      710       10D95AC     Operating
  4      Dirac-710-SN10D95 9123      710       10D957C     Operating
  5      Bessel-SN100212A- 9123      710       100212A     Operating
  6      Fourier-SN06EAB2B 9123      710       06EAB2B     Operating
  7      Server-9117-MMA-S 9117      MMA       109FEF4     Operating
  8      Cramer-SN10D958C- 9123      710       10D958C     Operating
  9      Server-9117-MMA-S 9117      MMA       10AF384     Operating
 10      Server-9119-FHA-S 9119      FHA       029F634     Operating
 11      Euler-710-SN06EAB 9123      710       06EAB1B     Operating
 12      Descartes-SN10D92 9123      710       10D921C     Operating
                                                                  more...
 System: 9_
```

Figure 3-21 Selecting the server

6. Open the console window (Figure 3-22).

```
Type option, press Enter.
1=Connect dedicated
2=Connect shared
3=Show Details

                                              Reference Use
Option  Partition        Partition State  Code     Count Console Status
 2      5: adama         Running          00000000 2     Connected
 _      6: boxey         Running          00000000 1     Connected
 _      7: attica        Running          00000000 1     Connected
```

Figure 3-22 Opening the console window of a partition

7. In the first console panel when installing the IBM i, select a language for the IBM i operating system (Figure 3-23).

```
                                                          System:   E10AF384
Language feature  . . . . . . . . . . . . . . :   2924

Press Enter to confirm your choice for language feature.
  Installing the system will continue.
```

Figure 3-23 IBM i installation: Language selection

8. In the next console panel (Figure 3-24), select **1. Install Licensed Internal Code**.

```
                        Install Licensed Internal Code
                                                          System:   E10AF384
Select one of the following:

     1. Install Licensed Internal Code
     2. Work with Dedicated Service Tools (DST)
     3. Define alternate installation device






Selection
     1
```

Figure 3-24 IBM i installation: Selecting Install Licensed Internal Code

9. Select the device for the load source unit. In our installation, we initially assigned only one XIV LUN to the tagged VSCSI adapter in IBM i, and we assigned the other LUNs later in System i DST. Therefore, we only have one unit to select as shown in Figure 3-25.

```
                         Select Load Source Device

Type 1 to select, press Enter.

                                      Sys   Sys    I/O        I/O
Opt  Serial Number     Type  Model    Bus   Card   Adapter    Bus    Ctl    Dev
 1   YH8EJU4FEBQY      6B22   050     255      3         0      0      1      0
```

Figure 3-25   IBM i installation: Selecting the load source unit

10. In the Install Licensed Internal Code (LIC) panel, select option **2. Install Licensed Internal Code and Initialize System** (Figure 3-26).

```
                        Install Licensed Internal Code (LIC)

Disk selected to write the Licensed Internal Code to:
         Serial Number    Type   Model   I/O Bus   Controller   Device
        YH8EJU4FEBQY      6B22    050          0            1        0

Select one of the following:

     1. Restore Licensed Internal Code
     2. Install Licensed Internal Code and Initialize system
     3. Install Licensed Internal Code and Recover Configuration
     4. Install Licensed Internal Code and Restore Disk Unit Data
     5. Install Licensed Internal Code and Upgrade Load Source







Selection
     2
```

Figure 3-26   Installing the Licensed Internal Code and initializing the system

11. When the warning is displayed, press F10 to accept it for the installation to continue. The installation starts, and you can observe the progress as shown in Figure 3-27.

```
                      Install Licensed Internal Code - Status

Install of the Licensed Internal Code in progress.

            +--------------------------------------------------+
Percent     ! ██████████              20%                       !
complete    +--------------------------------------------------+


Elapsed time in minutes  . . . . . . . . :       1.0
```

Figure 3-27   Progress of installing the Licensed Internal Code

12. After the installation of the Licensed Internal Code is completed, access the DST and add more disk units (LUNs) to the System i auxiliary storage pools (ASPs). Select the following options in DST as prompted by the panels:

   a. Select option **4. Work with disk units**.
   b. Select option **1. Work with disk configuration**.
   c. Select option **3. Work with ASP configuration**.
   d. Select option **3. Add units to ASPs**.
   e. Choose an ASP to add disks to or use.
   f. Select option **3. Add units to existing ASPs**.
   g. In the Specify ASPs to Add Units to window, select the disk units to add to the ASP by specifying the ASP number for each disk unit, and then press Enter.

   In our example, we connected eight additional 154-GB LUNs from the XIV system and added them to ASP1. Figure 3-28 shows the load source disk and all the disk units added in ASP1. The disk unit names start with DMP, which indicates that they are connected in a multipath.

```
          Serial                      Resource                          Hot Spare
ASP Unit  Number          Type Model  Name        Status                Protection
  1                                               Unprotected
       1  YH8EJU4FEBQY    6B22  050   DMP001      Configured                N
       2  YV9N8BP2R3Q3    6B22  050   DMP017      Configured                N
       3  YVQ9LKTH24PL    6B22  050   DMP007      Configured                N
       4  YBDGZKX8ES6S    6B22  050   DMP014      Configured                N
       5  YPCZSKB9P95K    6B22  050   DMP003      Configured                N
       6  Y97Y883RM53N    6B22  050   DMP015      Configured                N
       7  Y642ZRE5YHGM    6B22  050   DMP009      Configured                N
       8  YGRU6SVMMLMU    6B22  050   DMP005      Configured                N
       9  YAP3Z6SY6ENR    6B22  050   DMP011      Configured                N




Press Enter to continue.
```

Figure 3-28 LUNs in the System i ASP1

13. Exit the DST, and then choose option **2. Install the operating system** (Figure 3-29).

```
                          IPL or Install the System
                                                            System:   E10AF384
Select one of the following:

     1. Perform an IPL
     2. Install the operating system
     3. Use Dedicated Service Tools (DST)
     4. Perform automatic installation of the operating system
     5. Save Licensed Internal Code






Selection
     2

Licensed Internal Code - Property of IBM 5761-999 Licensed
Internal Code (c) Copyright IBM Corp.  1980, 2009.  All
rights reserved.  US Government Users Restricted Rights -
Use duplication or disclosure restricted by GSA ADP schedule
Contract with IBM Corp.
```

Figure 3-29 Installing the IBM i operating system

14. In the next console window, select the installation device type, and then confirm the language selection. The system starts an IPL from the newly installed load source disk (Figure 3-30).

```
                Licensed Internal Code IPL in Progress
                                                              11/05/09  16:22
IPL:
  Type  . . . . . . . . . . . . . . . . :   Attended
  Start date and time . . . . . . . . . :   11/05/09  16:21:51
  Previous system end . . . . . . . . . :   Abnormal
  Current step / total  . . . . . . . . :     10      16
  Reference code detail . . . . . . . . :   C6004057

IPL step                                    Time Elapsed   Time Remaining
 Journal Recovery                              00:00:01        00:00:00
 IFS Initialization                            00:00:01        00:00:00
>Data Base Recovery
 Journal Synchronization
 Commit Recovery

Item:
  Current / Total . . . . . . . :

Sub Item:
  Identifier  . . . . . . . . . :
  Current / Total . . . . . . . :
```

Figure 3-30 Licensed Internal Code IPL when installing the operating system

15. Upon completion of the IPL, when the system prompts you to load the next installation media (Figure 3-31), insert the media B2924_01 into the DVD drive. For the message at the console, type G.

```
                                  Message

Message ID . . . . . . :   CPA2055      Type . . . . . . . :   Inquiry
From program  . . . . . :   ITLOADFO     Severity . . . . . :   99

Message . . . . . :   The next volume in the sequence is needed to
  continue the installation process.

Cause . . . . . . :   The end of the volume has been reached.

Recovery  . . . . :   Load the next volume on the installation device.






Type choice, press Enter.

  Reply . . . . .   1         1=Continue, 2=End installing
```

Figure 3-31 Loading the next installation media volume

After you insert the date and time, the system shows a progress bar for the IBM i installation, as shown in Figure 3-32.

```
Message ID  . . :   CPI2070                   i5/OS Installation Status

            +-----------------------------------------------------+
   Stage 1  !                         58%                          !
            +-----------------------------------------------------+

Installation                                                 Objects
Stage                                             Completed   Restored
>> 1 Creating needed profiles and libraries . . . . :

   2 Restoring programs to library QSYS . . . . . . :

   3 Restoring language objects to library QSYS . . :

   4 Updating program table . . . . . . . . . . . . :

   5 Installing database files  . . . . . . . . . . :

   6 Installing base directory objects. . . . . . . :

i5/OS is a trademark of IBM in the United States and other countries.
```

Figure 3-32 Progress of the IBM i installation

4

Performance considerations for IBM i with the IBM XIV Storage System

For this paper, we tested IBM i client logical unit numbers (LUNs) defined on the IBM XIV Storage System. The purpose of the testing was to see the performance difference between using a small number of XIV volumes and a large number of XIV volumes when those volumes are on IBM i. In this testing, we used the same capacity. For the capacity, we used large volumes when we defined and connected only a small number of them. We used smaller volumes when we connected a larger number of them. Specifically, we used a 6-TB capacity. The capacity was divided into six 1 TB volumes or into forty-two 154 GB volumes.

We also wanted to see the performance improvement for an IBM i workload when running on an IBM XIV Storage System Gen3 compared to an XIV Gen2 system. For this reason, we ran the test with a large number of LUNs and with a small number of LUNs on both XIV Gen3 and XIV Gen2 systems, both equipped with 15 modules.

Important: These tests show the difference in IBM i performance between using a small number of large LUNs and a large number of smaller LUNs and to compare IBM i performance between XIV Gen3 and XIV Gen2 systems. The goal was *not* to make an overall configuration and setup recommendation for XIV to handle a specific IBM i workload.

This chapter examines the performance impact of a different number of LUNs and their size. It includes the following sections:

- Testing environment
- Testing workload
- Test with 154-GB volumes on XIV Gen2
- Test with 1-TB volumes on XIV Gen2
- Test with 154-GB volumes on XIV Gen3
- Test with 1-TB volumes on XIV Gen3
- Test with a double workload on XIV Gen3
- Conclusions

4.1 Testing environment

The testing environment for this paper uses the following configuration to connect to XIV Storage System Gen2 and Gen3:

- IBM POWER7 system, model 770
- Two IBM i logical partitions (LPARs), named LPAR2 and LPAR3, each of them running with six processing units and 80 GB of memory
- IBM i software level V7.R1, with cumulative PTF package C1116710 and Hyper group PTF SF99709 level 40 installed in each LPAR
- Two Virtual I/O Servers (VIOSs) in the POWER7 system
- VIOS software level 2.2.0.11, with fix pack 24 and service pack 1 in installed in each VIOS system
- XIV Storage System Gen2 with 15 modules with 1-TB disk drives, code level 10.2.4
- XIV Storage System Gen3 with 15 modules with 2-TB disk drives, code level 11.0.0
- In each VIOS, two ports in an 8-Gb Fibre Channel adapter to connect to the XIV, with each port connected to three interface modules in the XIV
- On the XIV Storage System, six 1 TB volumes and forty-two 154 GB volumes, with both volume groups assigned to the VIOS
- In each VIOS, the XIV volumes mapped as follows:
  - Six 1 TB volumes mapped to LPAR2 using two virtual SCSI adapters, with three volumes to each virtual SCSI adapter
  - Forty-two 154 GB volumes to LPAR3 using three virtual SCSI adapters, with 16 volumes to each virtual SCSI adapter

  In each configuration, the number of LUNs is a multiple of six. For a fully configured XIV Storage System with six interface modules, having a number of LUNs that is a multiple of six guarantees an equal distribution of the workload (I/O traffic) across the interface modules.

LPARs: In all tests that were conducted, except the test with a combined double workload, the XIV was used exclusively by one IBM i LPAR. No other applications or server I/Os were running on the XIV.

In the test with a combined double workload, the XIV was used only by two IBM i LPARs. Again, no other workloads were using the XIV.

Figure 4-1 illustrates the testing scenario.

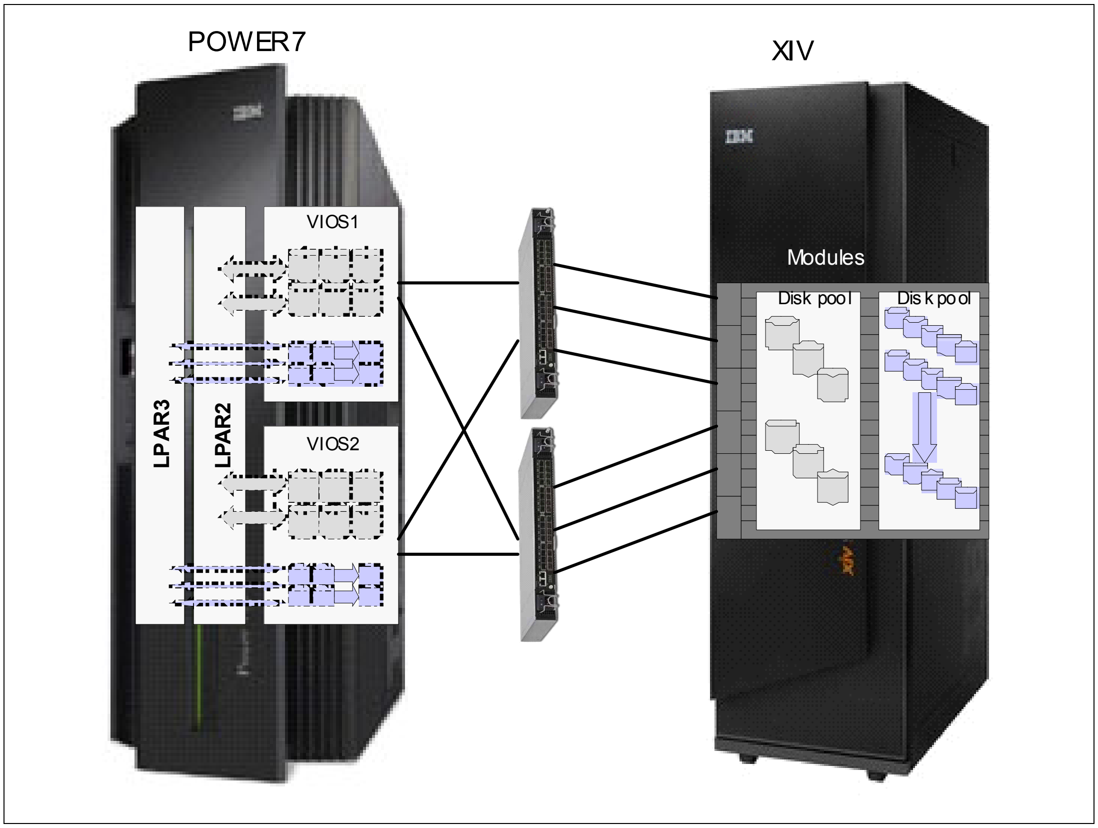


Figure 4-1 Testing environment

Figure 4-2 shows the IBM i LUNs that were defined on the XIV Storage System for each LPAR.

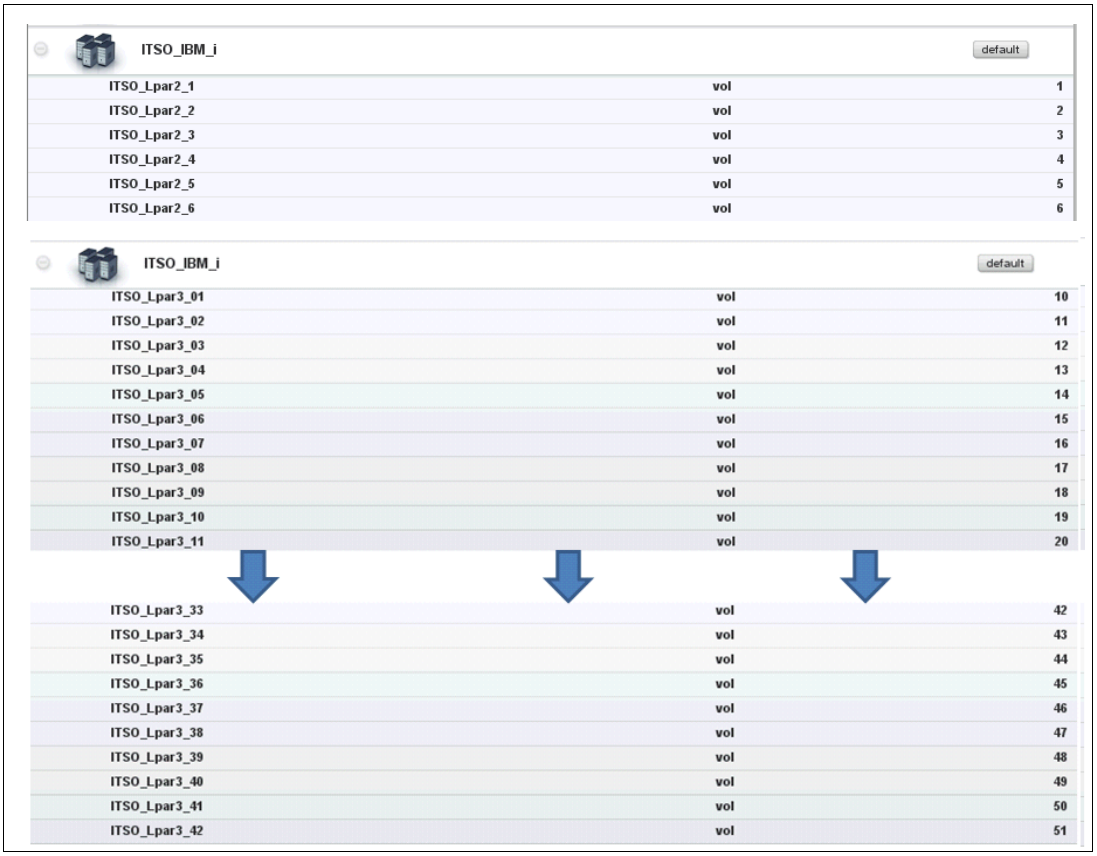


Figure 4-2 The LUNs for IBM i LPARs

Figure 4-3 and Figure 4-4 on page 43 show the XIV volumes reporting in IBM i System Service Tools (SST). Figure 4-3 shows the six 1 TB volumes.

```
                    Display Disk Configuration Status

           Serial                      Resource                             Hot Spare
ASP Unit   Number           Type Model Name          Status                 Protection
  1                                                  Unprotected
       1   Y7WKQ2FQGW5N     6B22   050 DMP001        Configured                  N
       2   Y7Y24LBTSUJJ     6B22   050 DMP026        Configured                  N
       3   Y22QKZEEUB7B     6B22   050 DMP013        Configured                  N
       4   YFVJ4STNADU5     6B22   050 DMP023        Configured                  N
       5   YTXTL2478XA3     6B22   050 DMP027        Configured                  N
       6   YZLEQY7AB82C     6B22   050 DMP024        Configured                  N
```

Figure 4-3 Six 1 TB LUNs reporting in IBM i

Figure 4-4 shows an abbreviated view of the forty-two 154 GB LUNs.

```
                      Display Disk Configuration Status

          Serial                          Resource                             Hot Spare
ASP Unit  Number          Type Model Name          Status                 Protection
  1                                                Unprotected
       1  Y9DY6HCARYRB    6B22  050  DMP001        Configured                  N
       2  YR657TNBKKY4    6B22  050  DMP003        Configured                  N
       3  YB9HSWBCJZRD    6B22  050  DMP006        Configured                  N
       4  Y3U8YL3WVABW    6B22  050  DMP008        Configured                  N
       5  Y58LXN6E3T8L    6B22  050  DMP010        Configured                  N
       6  YUYBRDN3597T    6B22  050  DMP011        Configured                  N

       ..................................................................

      35  YEES6NPSR6MJ    6B22  050  DMP050        Configured                  N
      36  YP5QPYTA89DP    6B22  050  DMP051        Configured                  N
      37  YNTD9ER85M4F    6B22  050  DMP076        Configured                  N
      38  YGLUSQJXUMGP    6B22  050  DMP079        Configured                  N
      39  Y6G7F38HSGQQ    6B22  050  DMP069        Configured                  N
      40  YKGF2RZWDJXA    6B22  050  DMP078        Configured                  N
      41  YG7PPW6KG58B    6B22  050  DMP074        Configured                  N
      42  YP9P768LTLLM    6B22  050  DMP083        Configured                  N
```

Figure 4-4 Forty-two 154 GB LUNs reporting in IBM i

4.2 Testing workload

The commercial processing workload (CPW) was used to conduct the tests. CPW evaluates a computer system and associated software in the commercial environment. It is maintained internally within the IBM i Systems Performance group.

The CPW application simulates the database server of an online transaction processing (OLTP) environment. These transactions are all run by batch server jobs, although they can easily represent the type of transactions that might be done interactively in a client environment. Each transaction interacts with 3 - 8 of the 9 database files that are defined for the workload. The database functions and file sizes vary.

The functions exercised are single and multiple row retrieval, single and multiple row insert, single row update, single row delete, journal, and commitment control. These operations are run against files that vary from hundreds of rows to hundreds of millions of rows. Some files have multiple indexes, and some have only one index. Some accesses are to the actual data, and some take advantage of advanced functions, such as index-only access.

CPW is considered a reasonable approximation of a steady-state, database-oriented commercial application.

After the workload is started, it generates the jobs in the CPW subsystems, with each job generating transactions. The CPW transactions are grouped by regions, warehouses, and users. Each region represents 1000 users or 100 warehouses, and each warehouse supports 10 users. CPW generates commercial types of transactions, such as orders, payments, delivery, and end-stock level.

For these tests, we ran the CPW with 96,000 users, or 9,600 warehouses. After starting the transaction workload, we waited 50 minutes and then started a performance collection that lasted 1 hour. After the performance collection is finished, several other IBM i analyzing tools such as PEX, were run. At the end, the CPW database was restored. The entire CPW run lasted about 5 hours.

4.3 Test with 154-GB volumes on XIV Gen2

The first test was with the 154-GB volumes that were defined on an XIV Gen2 system. Table 4-1 shows the number of different transaction types, the percentage of each type of transaction, the average response time, and the maximum response time. The average response time for most of the transactions was between 0.03 and 0.04 seconds. The maximum response time was 11.4 seconds.

Table 4-1 CPW transaction response times of 154-GB volumes on XIV Gen2

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	7013230	44.33	0.038	2.210
Ordersts	648538	4.10	0.042	2.550
Payment	6846381	43.27	0.033	11.350
Delivery	658587	4.16	0.000	0.250
Stocklvl	655281	4.14	0.083	2.340

Figure 4-5 shows the I/O rate and the disk service time during the collection period.

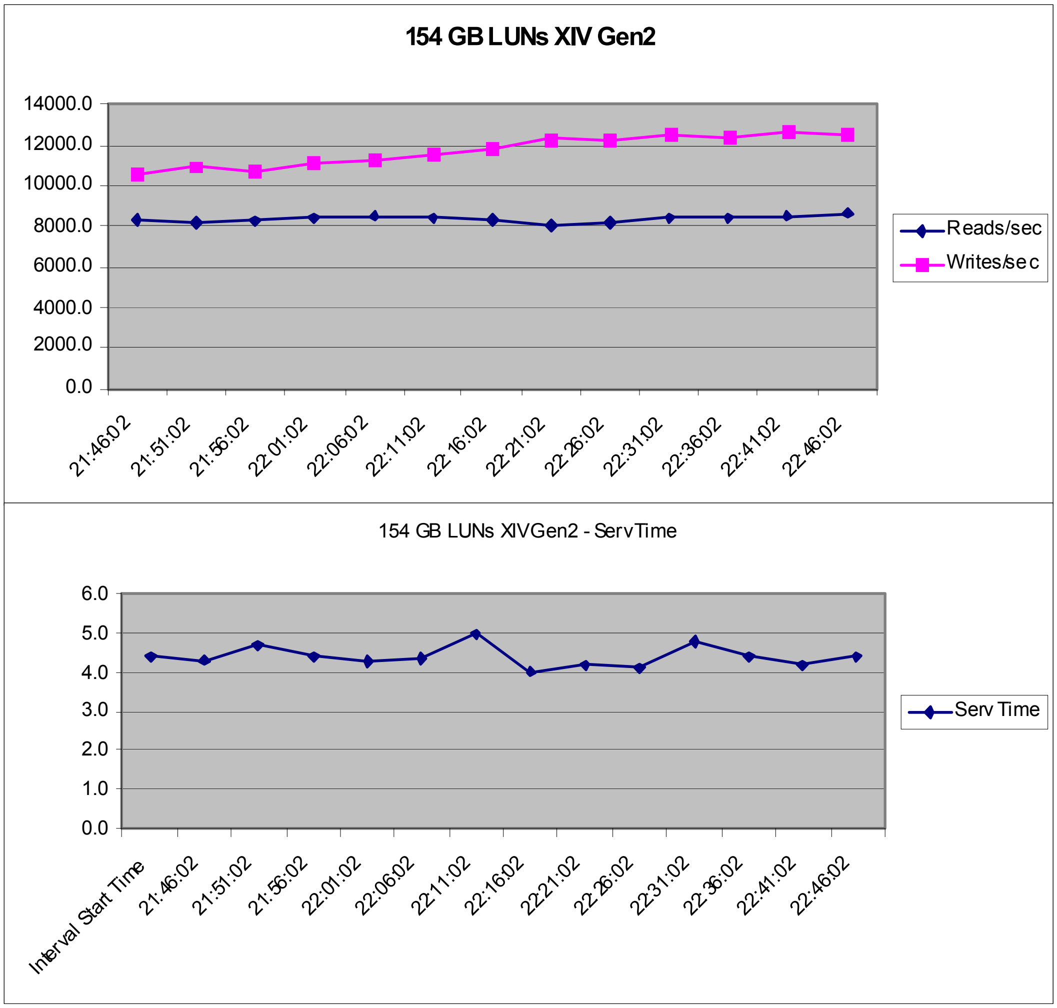


Figure 4-5 I/O rate and disk service time

During this collection period, CPW experienced an average of 8,355 reads/second and 11,745 writes/second. The average service time was 4.4 ms.

Disk wait time: Because the reported disk wait time in IBM i collection services reports was 0 in all the tests, it is not shown in the graphs.

The CPW database took 23 minutes to restore.

Figure 4-6 on page 46 through Figure 4-8 on page 48 show the I/O rate, latency, bandwidth in MBps, read/write ratio, and cache hits reported in XIV during the whole CPW run, during the IBM i collection period, and during the CPW database restore.

Figure 4-6 shows the values during the entire CPW run.

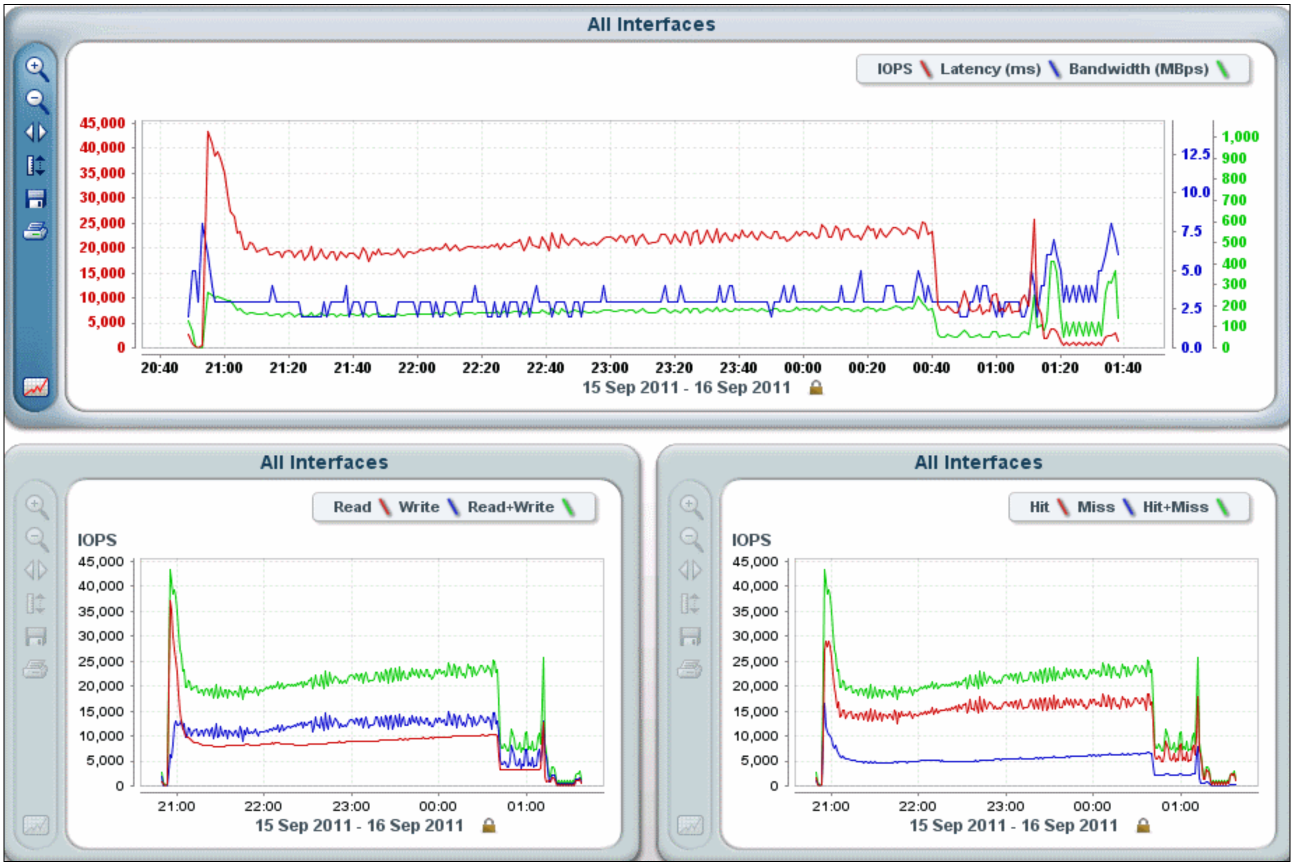


Figure 4-6 XIV values during the entire CPW run

Figure 4-7 shows that the approximate average latency during the collection period was 3 ms, and that the approximate percentage of cache hits was 75%.

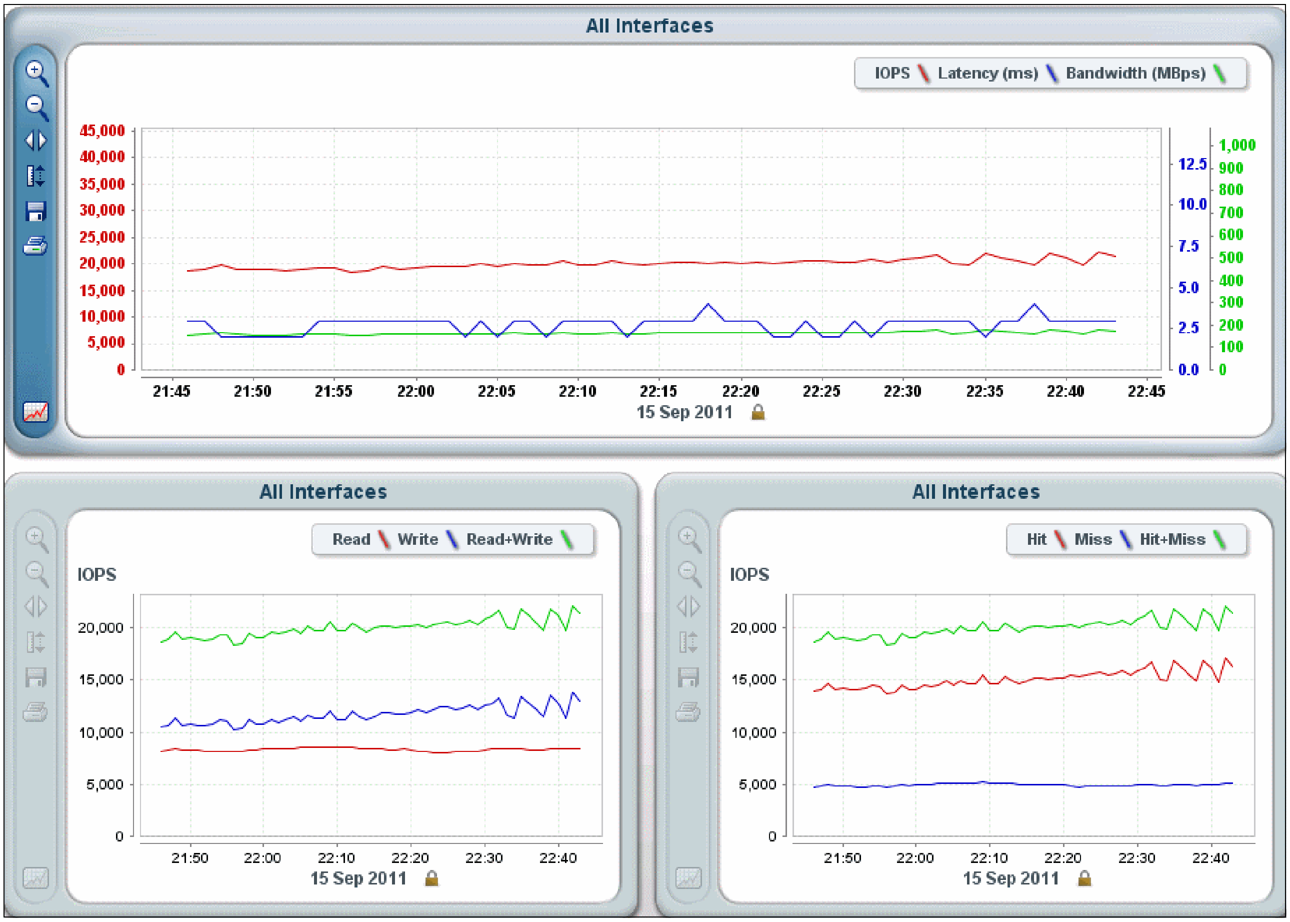


Figure 4-7 XIV values during the collection period

Figure 4-8 shows that, during the database restore, the approximate latency was 6 ms, and the approximate percentage of cache hits was 90 - 100%.

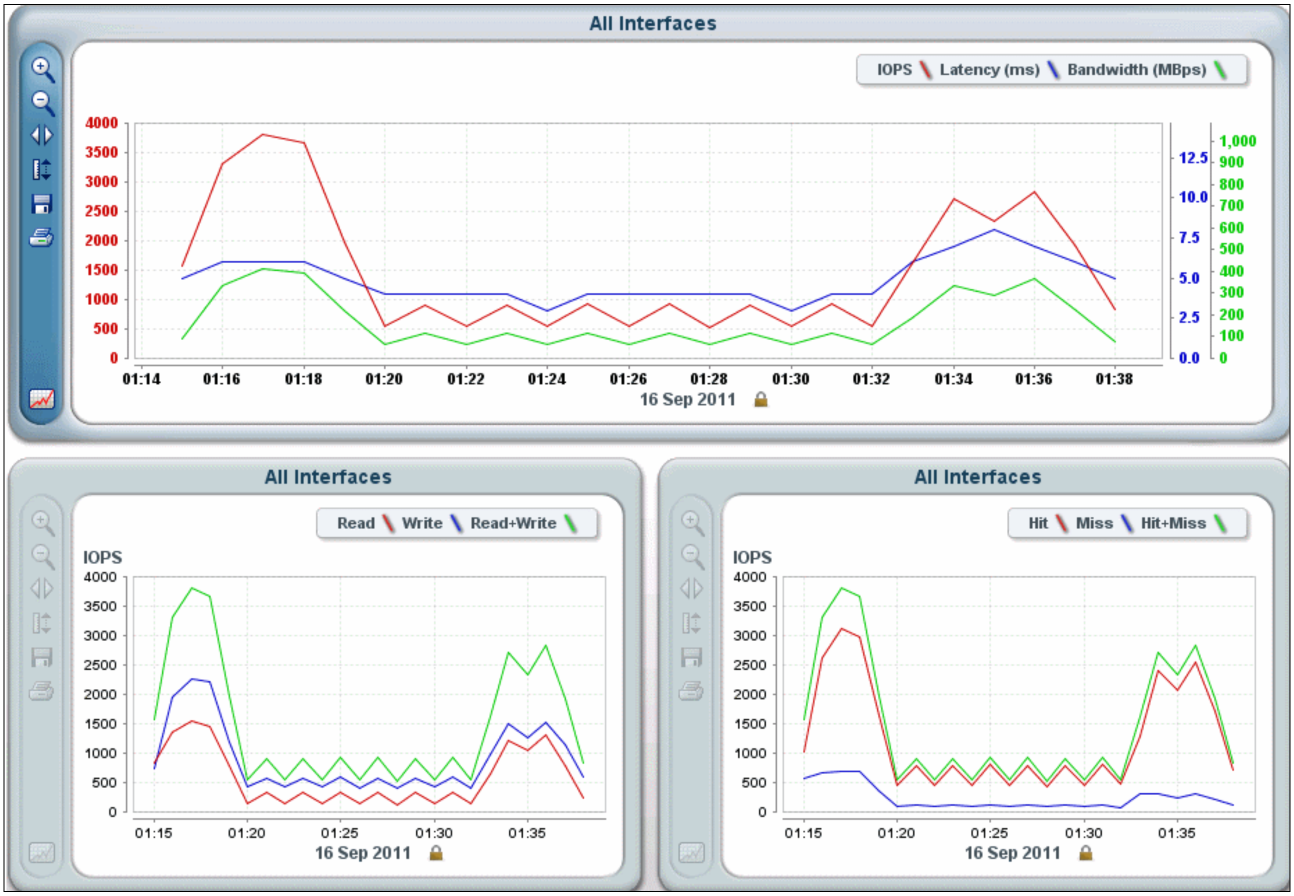


Figure 4-8 XIV values during the database restore

4.4 Test with 1-TB volumes on XIV Gen2

The second test was with 1-TB volumes on the XIV Gen2 system. Table 4-2 shows the number of transaction types, the percentage of each type of transaction, the average response time, and the maximal response time. The average response time for most of the transactions is between 0.3 and 10 seconds. The maximum response time is 984.2 seconds.

Table 4-2 CPW transaction response times of 1-TB volumes on XIV Gen2

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	3197534	46.21	10.219	984.170
Ordersts	271553	3.92	0.422	21.170
Payment	2900103	41.92	0.324	796.140
Delivery	275252	3.98	0.000	0.940
Stocklvl	274522	3.97	1.351	418.640

Figure 4-9 shows the I/O rate and the disk service time during the collection period. During this period, CPW experienced an average of 3,949 reads/second and 3,907 writes/second. The average service time was 12.6 ms.

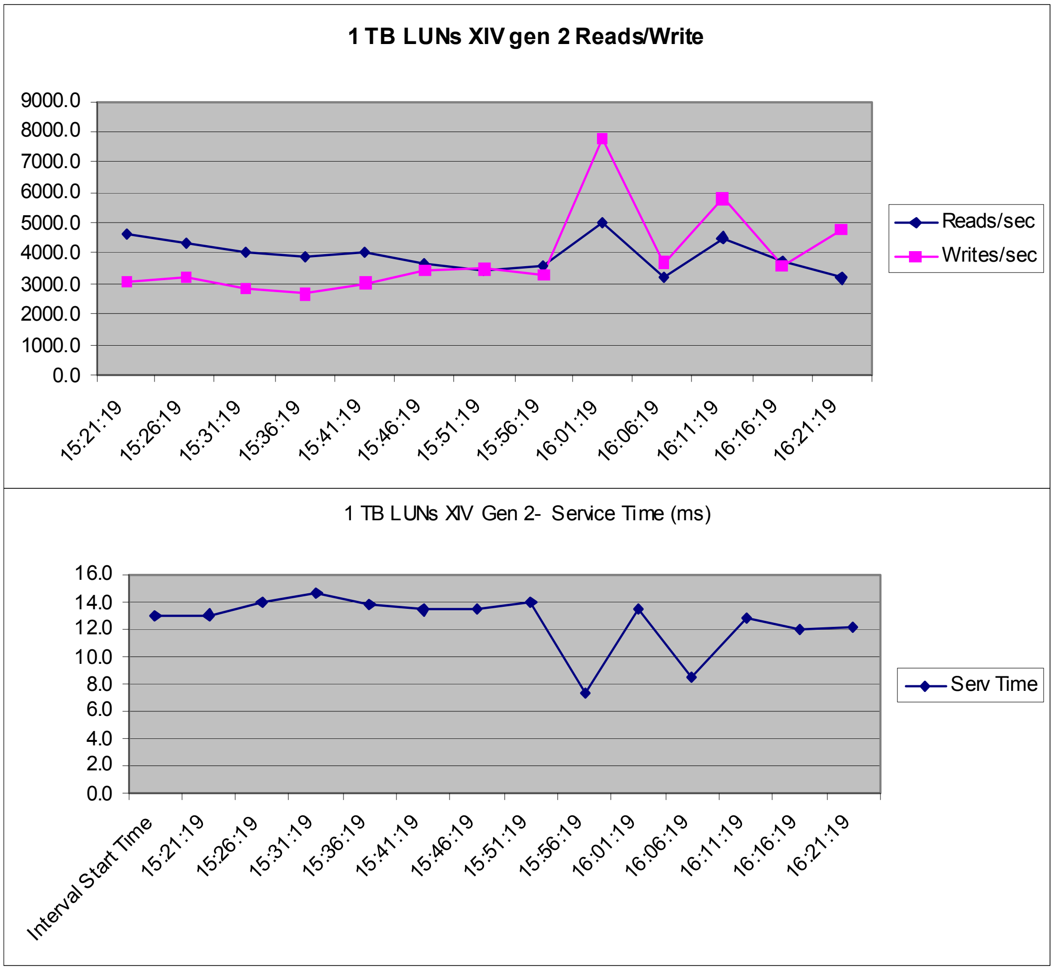


Figure 4-9 I/O rate and disk service time

The CPW database restore that was performed at the end of the run took 24 minutes.

Figure 4-10 on page 50 through Figure 4-12 on page 52 show the I/O rate, latency, bandwidth in MBps, read/write ratio, and cache hits reported in XIV during the whole CPW run, during the IBM i collection period, and for restoring the CPW database.

Figure 4-10 shows the values during the entire CPW run.

Figure 4-10 XIV values during the entire CPW run

As shown in Figure 4-11, the approximate average latency during the collection period was 20 ms, and the approximate percentage of cache hits was 50%.

Figure 4-11 XIV values during the collection period

Figure 4-12 shows that the approximate average of latency during the restore was 2.5 ms, and the approximate cache hit percentage was 90 - 100%.

Figure 4-12 XIV values during restore of the database

4.5 Test with 154-GB volumes on XIV Gen3

Table 4-3 shows the number of different transaction types, the percentage of each type of transaction, the average response time, and the maximum response time. The average response time for most of the transactions varies from 0.003 to 0.006 seconds. The maximum response time is 2.5 seconds.

Table 4-3 CPW transaction response times of 154-GB volumes on Gen3

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	7031508	44.32	0.006	0.540
Ordersts	650366	4.10	0.004	0.390
Payment	6864817	43.27	0.003	2.460
Delivery	660231	4.16	0.000	0.010
Stocklvl	656972	4.14	0.031	0.710

Figure 4-13 shows the disk service time response time during the collection period. The average service time of a 1-hour collection period was 0.5 ms.

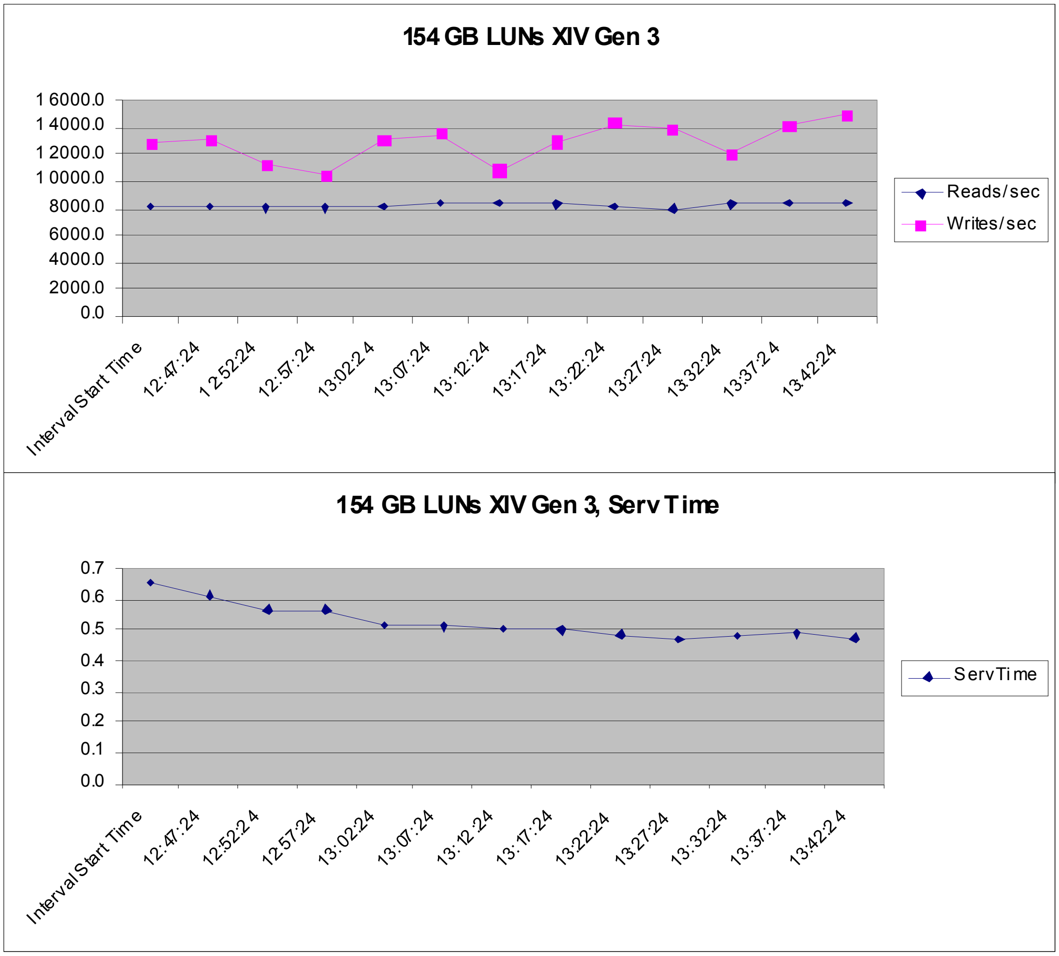


Figure 4-13 I/O rate and disk service times

Figure 4-14 on page 54 through Figure 4-16 on page 56 show the I/O rate, latency, bandwidth in MBps, read/write ratio, and cache hits reported in XIV during the whole CPW run, during the time of the IBM i collection period, and during the CPW database restore.

Figure 4-14 shows the values during the entire CPW run.

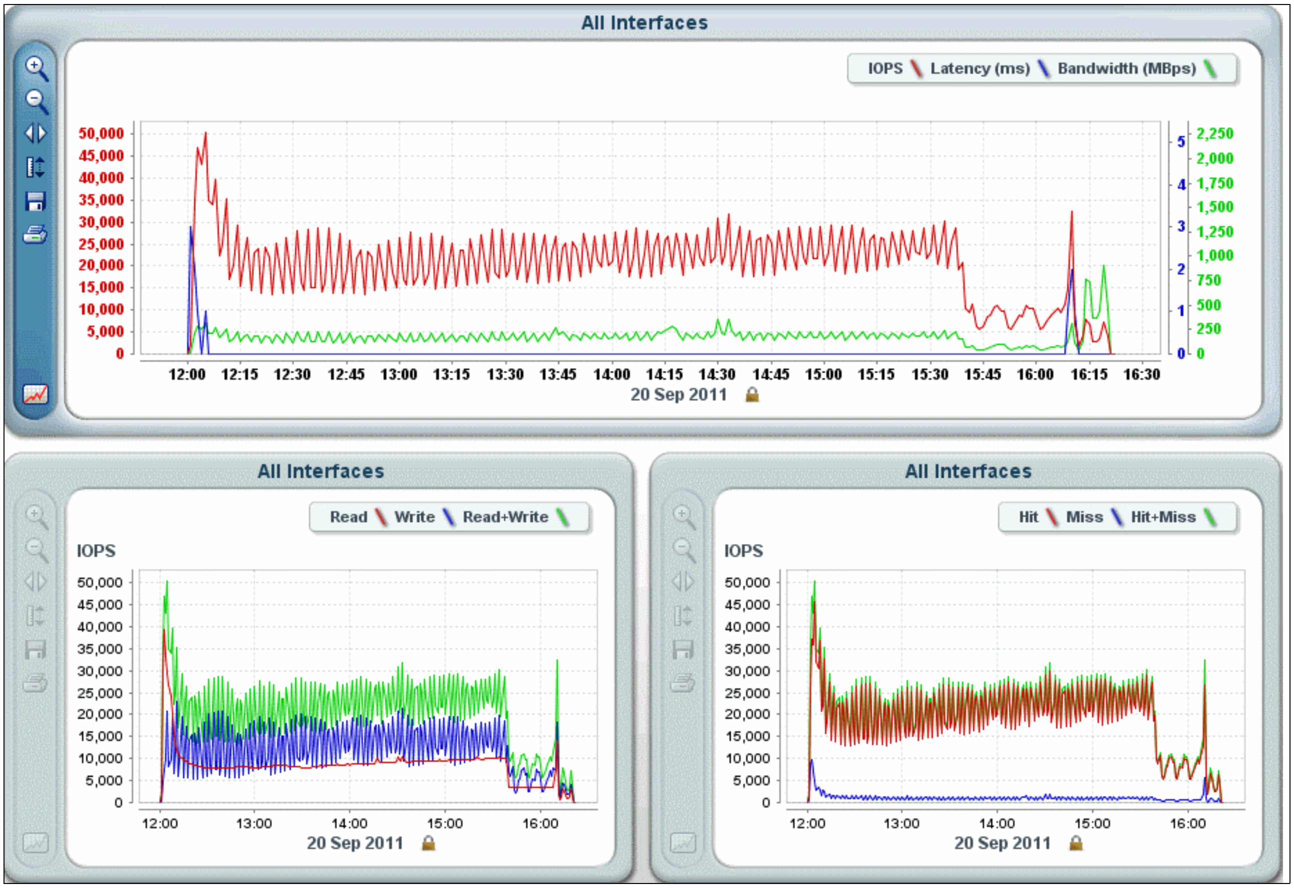


Figure 4-14 XIV values during the entire CPW run

Figure 4-15 shows that the average latency during the collection period is close to 0, and the average percentage of cache hits is close to 100%.

Figure 4-15 XIV values during the collection period

Figure 4-16 shows that, during the database restore, the latency was close to 0, and the percentage of cache hits was close to 100%.

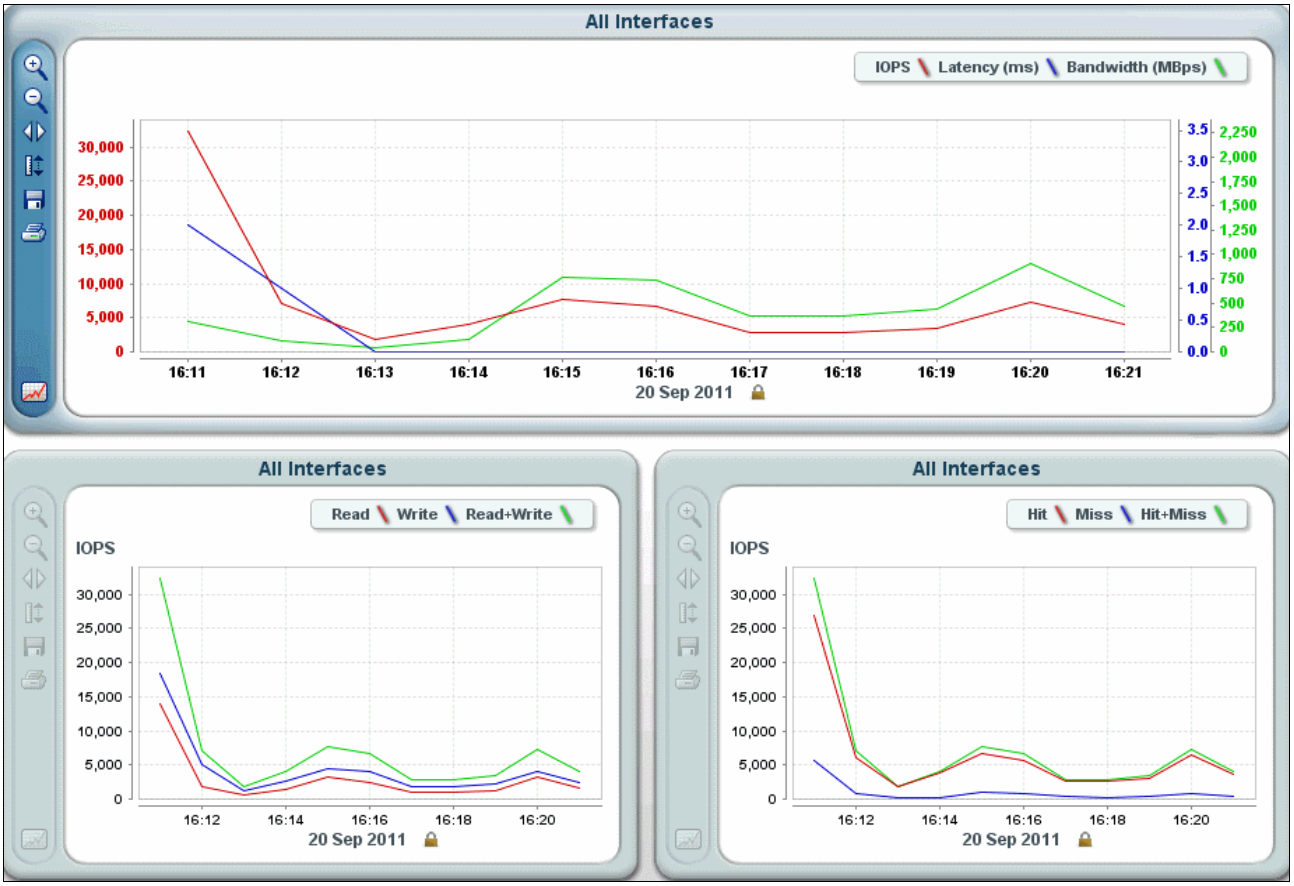


Figure 4-16 XIV values during the database restore

4.6 Test with 1-TB volumes on XIV Gen3

Table 4-4 shows the number of different transaction types, the percentage of each transaction type, the average response time, and the maximum response time. The average response time of most of the transactions varies from 0.003 seconds to 0.006 seconds. The maximum response time is 2.6 seconds.

Table 4-4 CPW transaction response times of 1-TB volumes on XIV Gen3

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	7032182	44.33	0.006	0.390
Ordersts	650306	4.10	0.005	0.310
Payment	6864866	43.27	0.003	2.620
Delivery	660280	4.16	0.000	0.040
Stocklvl	657016	4.14	0.025	0.400

Figure 4-17 shows the disk service time during the collection period. The average service time for a 1-hour collection period was 0.5 ms.

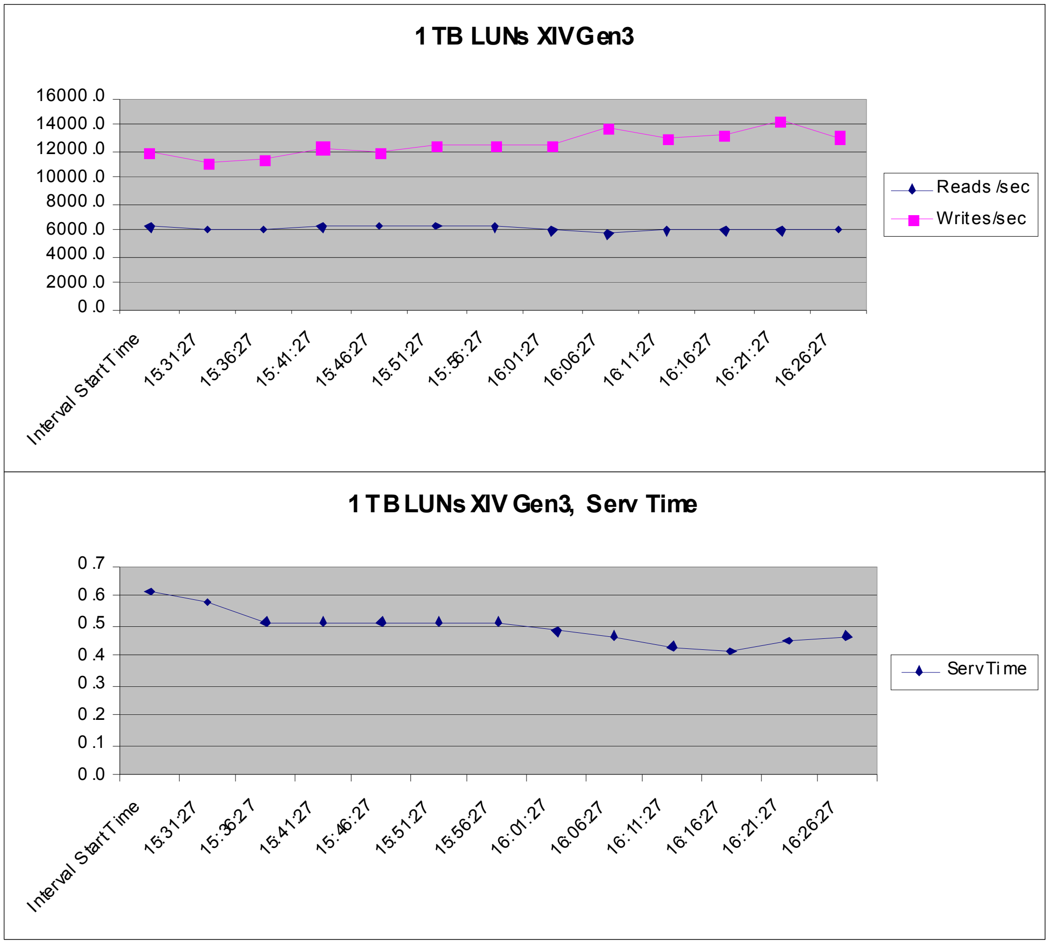


Figure 4-17 I/O rate and disk service time

Figure 4-18 on page 58 through Figure 4-20 on page 60 show the I/O rate, latency, bandwidth in MBps, read/write ratio, and cache hits reported in XIV during the whole CPW run, during the IBM i collection period, and during the CPW database restore.

Figure 4-18 shows the values during the entire CPW run.

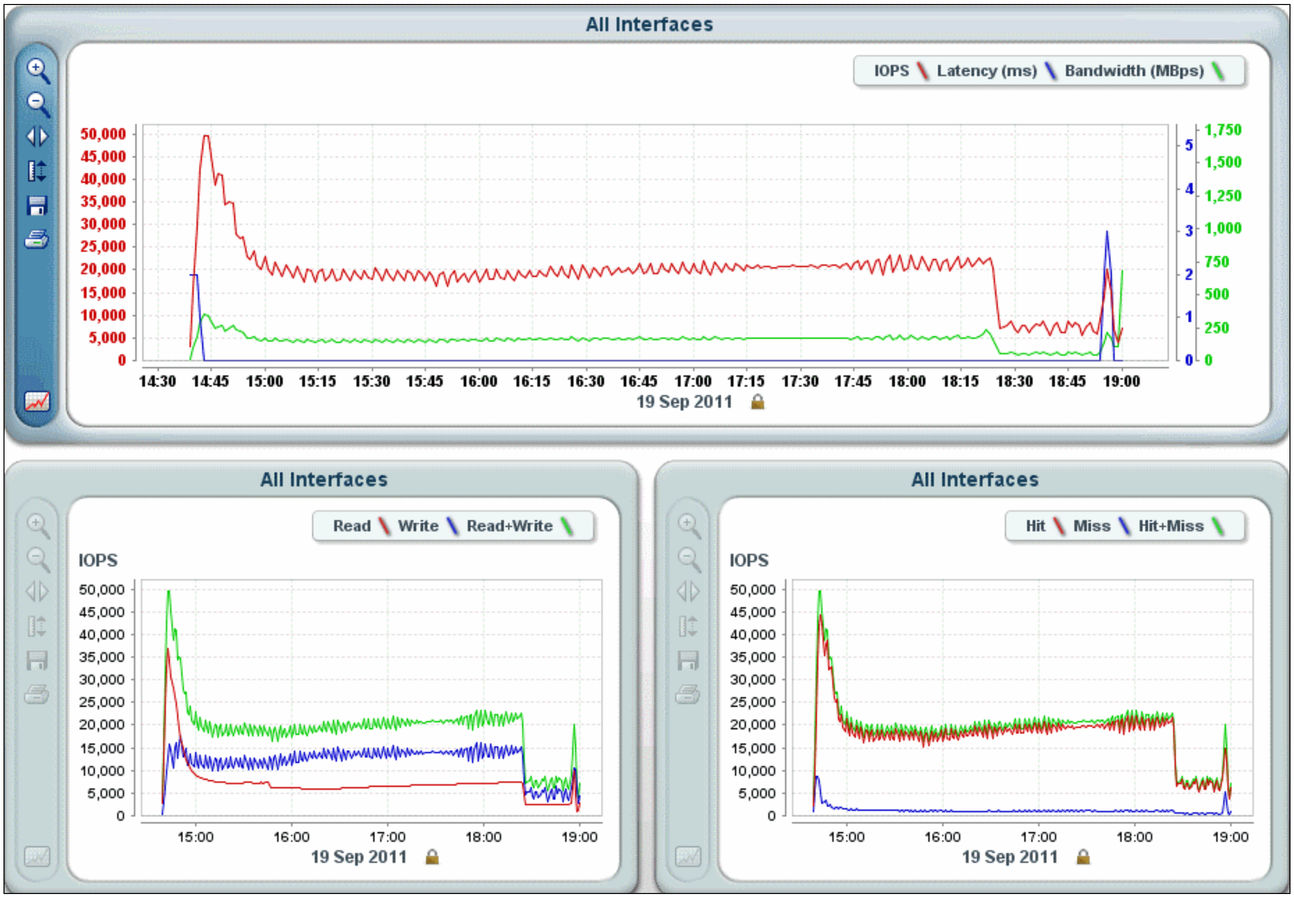


Figure 4-18 XIV values during the entire CPW run

Figure 4-19 shows that, during the collection period, an approximate average latency was 0.2 ms, and the approximate percentage of cache hits was almost 100%.

Figure 4-19 XIV values during the collection period

Figure 4-20 shows that, during the database restore, the latency was close to 0, and the percentage of cache hits was close to 100%.

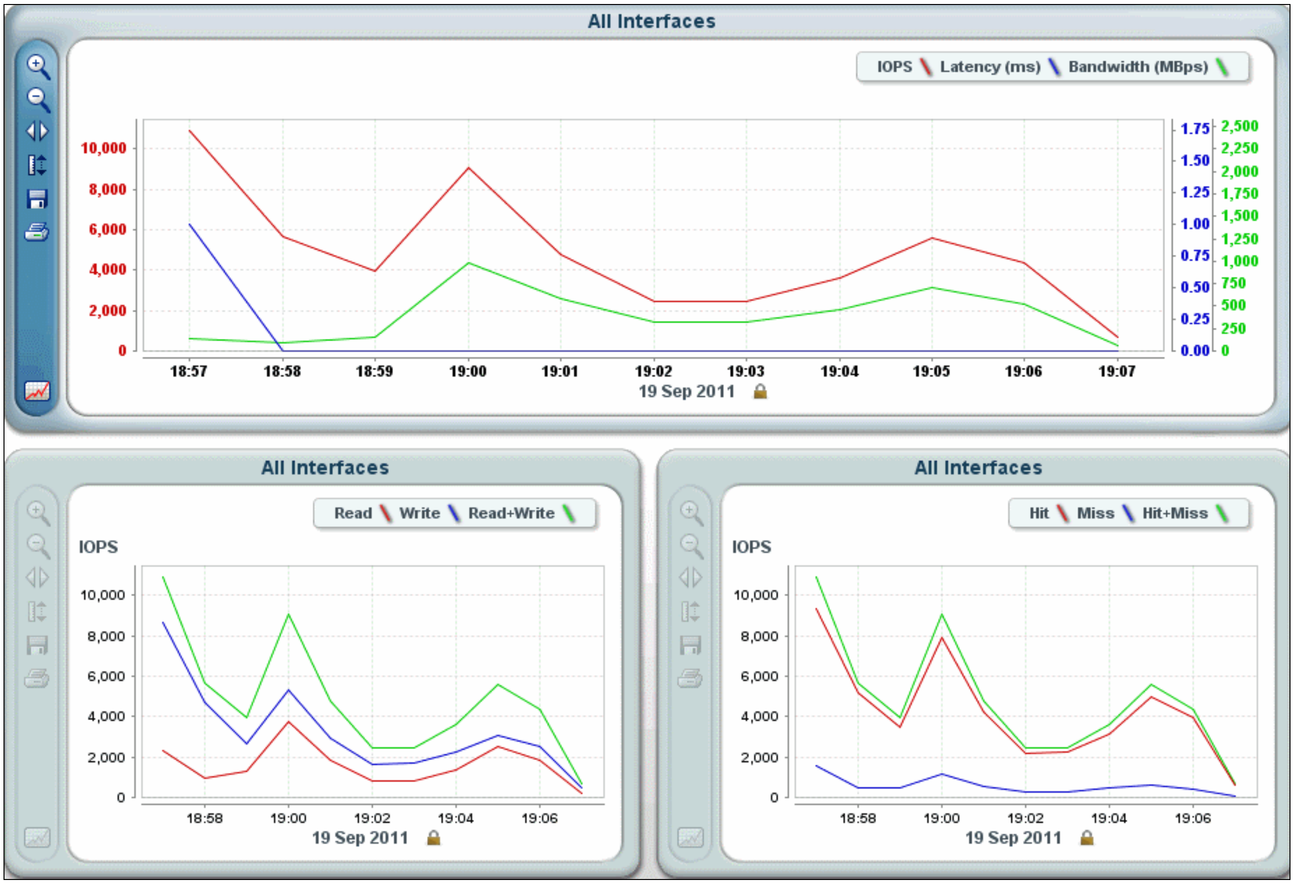


Figure 4-20 XIV values during the database restore

4.7 Test with a double workload on XIV Gen3

In the tests performed on the XIV Storage System Gen3, the workload experienced excellent cache hits close to 100%. Consequently the disk response time was good, and the response time was the same between the environment with the forty-two 154-GB LUNs and the six 1-TB LUNs.

To see the performance difference better between the two LUN sizes and the number of LUNs on XIV Gen3, we increased the I/O rate on the XIV. We ran the CPW with 192,000 users (double compared to the previous runs) in each IBM i LPAR, and we ran the workload in both LPARs at the same time.

Table 4-5 shows the number of different transaction types, the percentage of each type of transaction, the average response time, and the maximum response time. The average response time for most of the transactions was in the range 0.6 - 42 seconds. The maximum response time is 321 seconds.

Table 4-5 CPW transaction response times for 1-TB LUNs with a double workload on XIV Gen3

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	1423884	44.33	42.181	321.330
Ordersts	131548	4.10	0.733	30.150
Payment	1389705	43.27	0.558	38.550
Delivery	133612	4.16	0.000	0.150
Stocklvl	133113	4.14	9.560	44.920

Figure 4-21 shows the response time for the disk service time during the collection period. The average disk service time of a 1-hour collection period was 8.2 ms. The average LUN utilization was 91%.

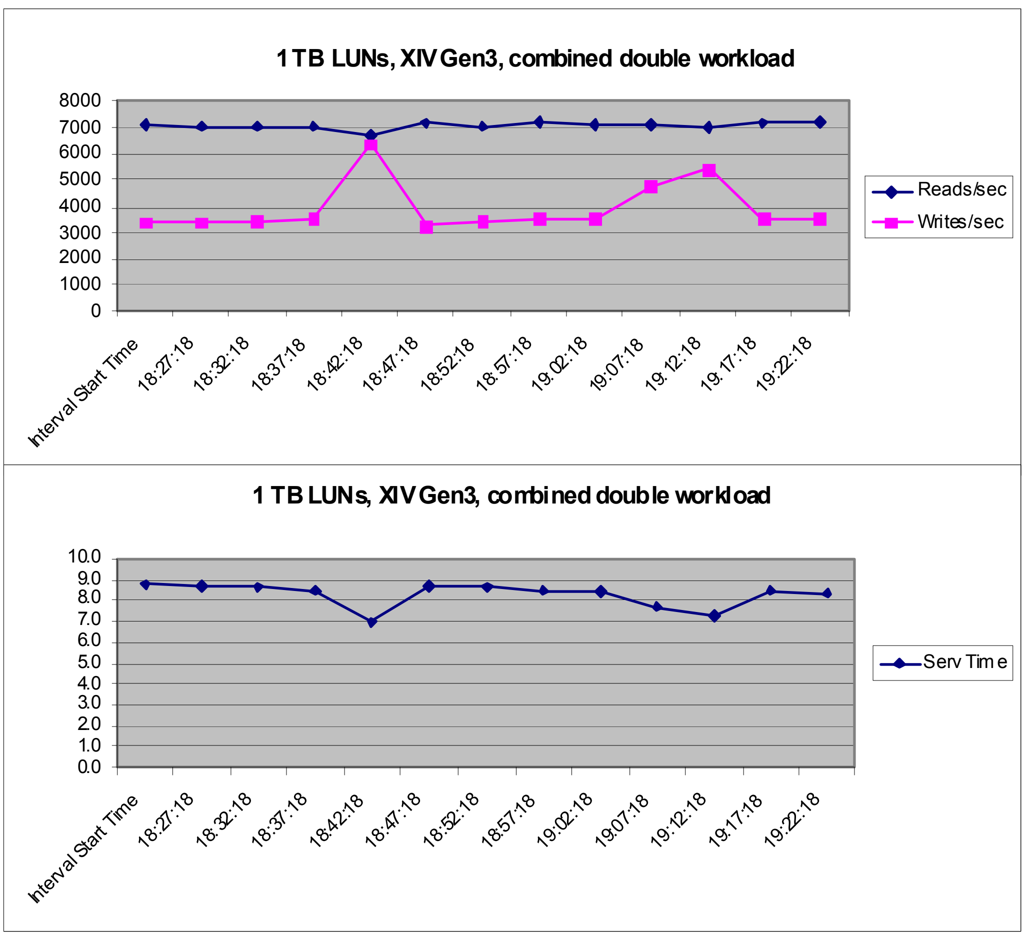


Figure 4-21 1-TB LUNs, combined double workload, I/O rate, and service time

The CPW database restore, which is performed at the end of the run, took 16 minutes.

Table 4-6 shows the transaction response time for the CPW run on the 154-GB LUNs. The average response time for most of the transactions was in the range 1.6 - 12 seconds.

Table 4-6 CPW transaction response times for 154-GB LUNs with a double workload on XIV Gen3

Transaction ID	Count	Percentage	Average response time (seconds)	Maximum response time (seconds)
Neworder	6885794	47.39	12.404	626.260
Ordersts	553639	3.81	0.511	16.850
Payment	5968925	41.08	1.545	178.690
Delivery	560421	3.86	0.042	6.210
Stocklvl	560005	3.85	2.574	21.810

Figure 4-22 shows the disk service time response time during the collection period. The average disk service time of a 1-hour collection period was 4.6 ms. The average LUN utilization was 78.3%.

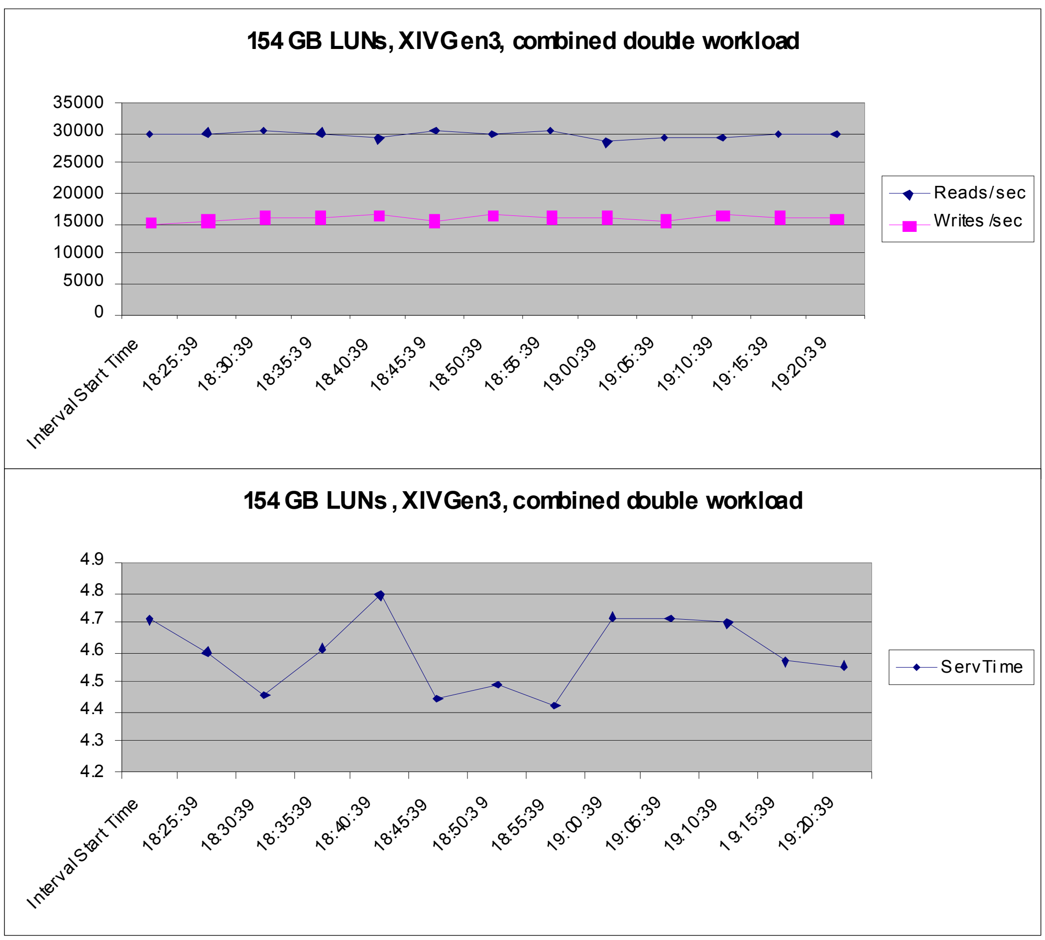


Figure 4-22 154-GB LUNs, combined double workload (with two LPARS), I/O rate, and service time

The CPW database restore took 13 minutes.

In this test, we also ran the workloads in both IBM i LPARs at the same time. Figure 4-23 on page 63 through Figure 4-25 on page 65 show the I/O rate, latency, bandwidth in MBps, and cache hits in XIV during the whole CPW run, during the IBM i collection period, and during the CPW database restore.

The figures show the XIV values of one LUN. Therefore, you must multiply the I/O rate and MBps by the number of LUNs to obtain the overall rates. The latency and cache hits that are shown for one LUN are about the same as the average across all LUNs in the LPAR.

Figure 4-23 shows the XIV values for the entire run of both CPW workloads.

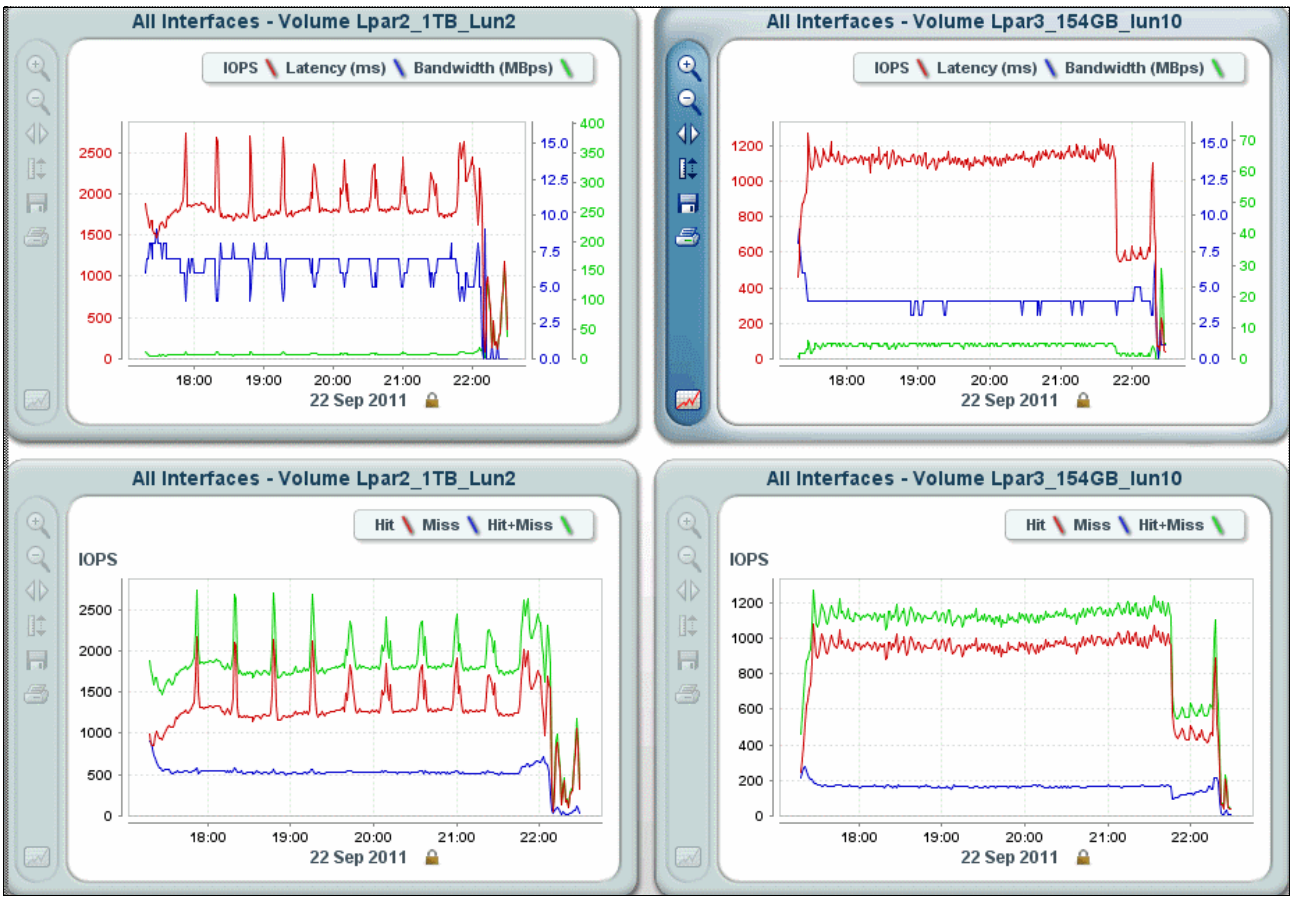


Figure 4-23 XIV values for the entire run of both CPW workloads

Figure 4-24 shows that the average latency of 1-TB LUNs was about 7 ms, and the latency of the 154-GB LUNs was close to 4 ms. On the 1-TB LUNs, the workload experienced about a 60% cache hit. On the 154-GB LUNs, the cache hits were about 80%.

Figure 4-24 XIV values for data collection of both CPW workloads

Figure 4-25 shows that, during the database restore, almost no latency occurred on the 1-TB LUNs, but the latency on the 154-GB LUNs was about 1 ms. Figure 4-25 also shows that the cache hits on both the 1 TB and 154-GB LUNs are close to 100%.

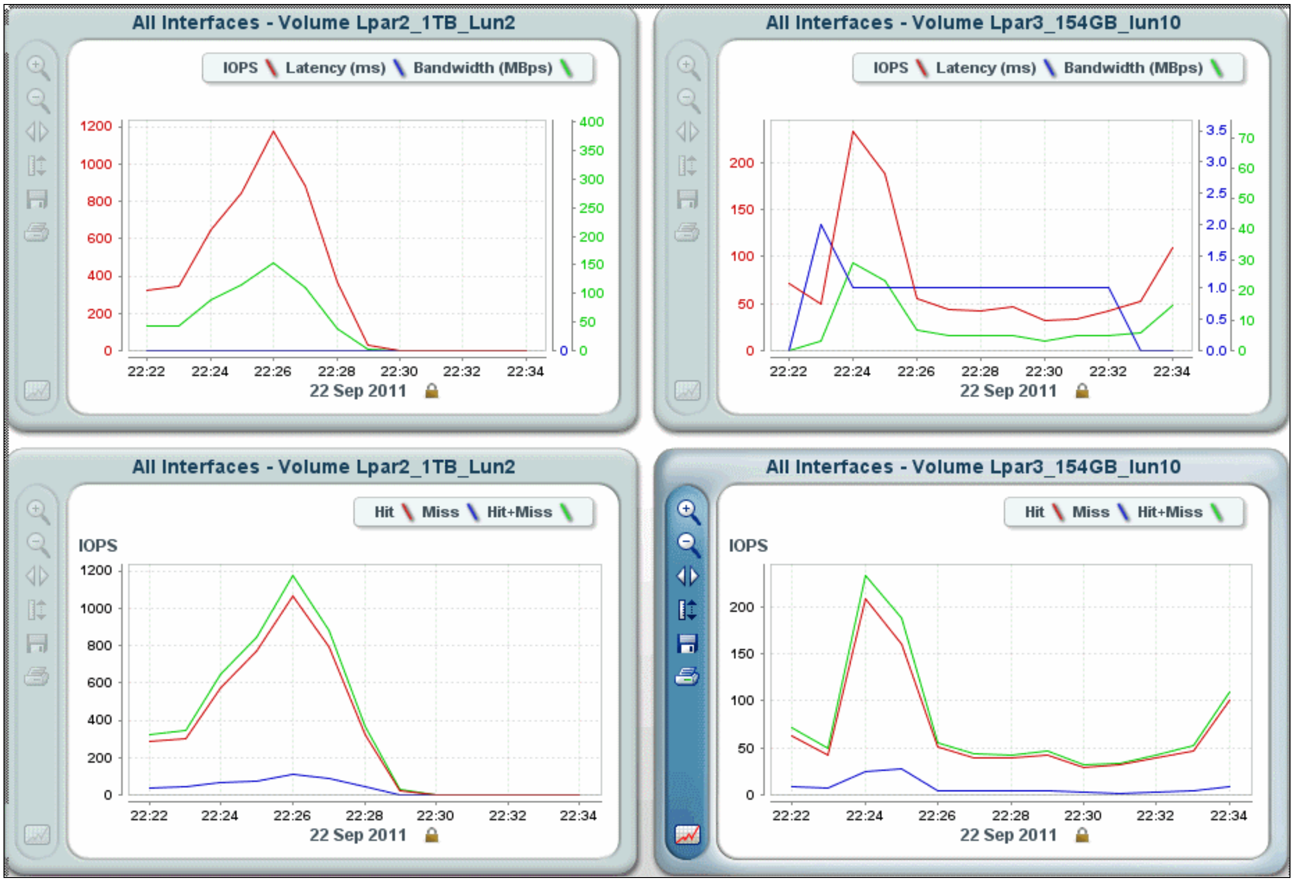


Figure 4-25 XIV values for database restore of both CPW workloads

4.8 Conclusions

Table 4-7 shows the results of the tests that we ran. It includes the average transaction response time and the disk service time reported by IBM i Performance Tools during the CPW data collection period. It also shows the latency and cache hits reported by XIV statistics during the same period.

Table 4-7 Tests results

	Average transaction response time (seconds)	Average disk service time (milliseconds)	Application average latency on XIV (milliseconds)	Application percentage of average cache hits on XIV
XIV Gen2				
Forty-two 154 GB LUNs	0.036	4.4	3	75
Six 1 TB LUNs	4.9	12.6	20	50
XIV Gen3				
Forty-two 154 GB LUNs	0.005	0.5	Near 0	Near 100
Six 1 TB LUNs	0.005	0.5	Near 0	Near 100
XIV Gen3 concurrent double workload				
Forty-two 154 GB LUNs	6.6	4.6	4	80
Six 1 TB LUNs	19.3	8.2	7	60

4.8.1 Comparison of many small LUNs to a few large LUNs

Table 4-7 shows that, on an XIV Gen2, the workload experienced much better response times when using many smaller LUNs compared to using a few large LUNs.

When running the same workload with many small LUNs or a few large LUNs on an XIV Gen3 system, the performance is good in both cases. Also, the difference in response time is minimal between the two environments.

However, when we stressed the XIV Gen3 more, by simultaneously running a double workload in both LPARs, we saw a large difference in response times between the two environments. The response times were faster for the configuration using many small LUNs.

The usage of many small LUNs offers better performance for the following reasons:

- The maximum queue depth (the number of I/O operations that can be done concurrently to a volume) for an IBM i volume in VIOS is 32, which is rather modest compared to the maximum queue depths for other open servers. Therefore, for IBM i, consider defining a larger amount of small LUNs than for the other open system servers, to achieve a comparable number of concurrent I/O operations for the disk space available.
- The more LUNs that are available to an IBM i system, the more server tasks (similar to threads in open systems) IBM i storage management uses to manage the I/O operations to the disk space. Consequently better I/O performance is achieved.

4.8.2 Comparison of XIV Storage System Gen2 and Gen3

When running with the same LUN configuration (size and number of LUNs) on XIV Gen2 and XIV Gen3, we saw much better response times on XIV Gen3. This result is explained by the larger cache and enhanced back-end architecture of the XIV Gen3.

The CPW workload with 96000 users running on XIV Gen3 experienced almost 100% cache hits and good disk response time, which was below 1 ms.

Running two such workloads at the same time did not stress the XIV much more. The cache hits were in the range 90 - 100%, and the service times were in the range 1.2 - 1.6 ms.

When we increased the workload to 192,000 users and ran it concurrently in the two LPARs, we saw lower cache hits in the range 60 - 80% and disk service times in range of about 4 - 8 ms.

4.8.3 Conclusion about the XIV LUN size for IBM i

In our tests, when comparing the configurations of six 1 TB LUNs with the configuration of forty-two 154 GB LUNs (of equal disk capacity in each environment), we observed significantly better performance with the forty-two 154 GB LUN configuration.

Considering the performance and keeping the number of LUNs at a reasonable amount for ease of management in XIV, VIOS, and IBM i, we conclude that, in most cases, a LUN size in the 100 - 150 GB range is appropriate.

In each configuration, the number of LUNS is a multiple of 6. For a fully configured XIV Storage System with six interface modules, having a number of LUNs that is a multiple of six guarantees an equal distribution of the workload (I/O traffic) across the interface modules.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120
- *IBM i and Midrange External Storage*, SG24-7668
- *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590
- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- *IBM XIV Storage System Host System*, GC27-2215
- *IBM XIV Storage System Installation and Service Manual*, GA32-0590
- *IBM XIV Storage System Introduction and Theory of Operations*, GC27-2214
- *IBM XIV Storage System Model 2810 Installation Planning Guide,* GC52-1327-01
- *IBM XIV Storage System Pre-Installation Network Planning Guide for Customer Configuration,* GC52-1328-01
- *IBM XIV Storage System XCLI Manual*, GC27-2213
- *XCLI Reference Guide*, GC27-2213

Online resources

These Web sites are also relevant as further information sources:

- IBM XIV Storage System overview

  http://www.ibm.com/systems/storage/disk/xiv/index.html
- IBM XIV Storage System Information Center

  http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

- System Storage Interoperability Center (SSIC)
  http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

IBM XIV Storage System with the Virtual I/O Server and IBM i

Redpaper™

Understand how to attach XIV to IBM i through VIOS

Learn how to exploit the multipathing capability

Follow best practices for the number of LUNs and their size

This IBM Redpaper publication explains how you can connect the IBM XIV Storage System server to the IBM i operating system through the Virtual I/O Server (VIOS). A connection through the VIOS is especially of interest to IT centers that have many small IBM i partitions. When using the VIOS, the Fibre Channel host adapters can be installed in the VIOS and shared by many IBM i clients by using virtual connectivity to the VIOS.

This paper also includes guidance for sizing XIV logical unit numbers (LUNs) when attached to the IBM i client.

REDP-4598-01