Redbooks
ibm.com/redbooks

# IBM DS8000 Easy Tier

## (Updated for DS8000 Release 9.0)

Bert Dufrasne

Peter Kimmel

Matthew Houzenga

Dennis Robertson

Storage

IBM

Redpaper

International Technical Support Organization

# IBM DS8000 Easy Tier (Updated for DS8000 Release 9.0)

November 2019

REDP-4667-08

**Note:** Before using this information and the product it supports, read the information in “Notices” on page vii.

**Ninth Edition (November 2019)**

This edition applies to the IBM DS8000 series with DS8000 LMC 7.9.0.xxx (bundle version 89.0.xxx.x).

This document was created or updated on October 7, 2019.

# Contents

# Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

# Trademarks

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

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

| | | |
|---|---|---|
| Db2® | IBM Spectrum™ | Redbooks (logo)  ® |
| DS8000® | IBM Spectrum Control™ | Resource Measurement Facility™ |
| Easy Tier® | IBM Z® | RMF™ |
| FlashCopy® | Power Systems™ | System Storage® |
| GDPS® | Redbooks® | z Systems® |
| IBM® | Redpaper™ | z/OS® |

The following terms are trademarks of other companies:

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation 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 describes the concepts and functions of IBM System Storage® Easy Tier®, and explains its practical use with the IBM DS8000® series and License Machine Code 7.9.0.xxx (also known as R9.0).

Easy Tier is designed to automate data placement throughout the storage system disks pool. It enables the system to (automatically and without disruption to applications) relocate data (at the extent level) across up to three drive tiers. The process is fully automated. Easy Tier also automatically rebalances extents among ranks within the same tier, removing workload skew between ranks, even within homogeneous and single-tier extent pools.

Easy Tier supports a Manual Mode that enables you to relocate full volumes. Manual Mode also enables you to merge extent pools and offers a rank depopulation function. Easy Tier fully supports thin-provisioned Extent Space Efficient fixed block (FB) and count key data (CKD) volumes in Manual Mode and Automatic Mode.

Easy Tier also supports extent pools with small extents (16 MiB extents for FB pools and 21 cylinders extents for CKD pools). Easy Tier also supports high-performance and high-capacity flash drives in the High-performance flash enclosure, and it enables additional user controls at the pool and volume levels.

This paper is aimed at those professionals who want to understand the Easy Tier concept and its underlying design. It also provides guidance and practical illustrations for users who want to use the Easy Tier Manual Mode capabilities.

Easy Tier includes additional capabilities to further enhance your storage performance automatically: Easy Tier Application, and Easy Tier Heat Map Transfer.

For more information about these features, see the following publications:

- *DS8870 Easy Tier Application*, REDP-5014
- *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015.

## Authors

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

**Bert Dufrasne** is an IBM Certified IT Specialist and Project Leader for IBM System Storage disk products at the ITSO, San Jose Center. He has worked at IBM in various IT areas. He has written many IBM publications, and has developed and taught technical workshops. Before Bert joined the ITSO, he worked for IBM Global Services as an Application Architect. He holds a Master's degree in Electrical Engineering.

**Peter Kimmel** is an IT Specialist and ATS team lead of the Enterprise Storage Solutions team at the EMEA Storage Competence Center (ESCC) in Kelsterbach, Germany. He joined IBM Storage in 1999, and since then has worked with all the various IBM Enterprise Storage Server and System Storage DS8000 generations, with a focus on architecture and performance. He was involved in the Early Shipment Programs (ESPs) of these early installs, and co-authored several DS8000 IBM publications. Peter holds a Diploma (MSc) degree in Physics from the University of Kaiserslautern.

Thanks to the authors of the previous editions of this paper:

Matthew Houzenga, Antonio Rainero, Dennis Robertson.

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

Nick Clayton
Ping He
Yang Liu (Loren)
Paul Spagnolo
**IBM**

## 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 2 - 6 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:

**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:

  **ibm.com**/redbooks

- Send your comments in an email:

  redbooks@us.ibm.com

- Mail your comments:

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

## Stay connected to IBM Redbooks

- Find us on Facebook:

  https://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  https://twitter.com/ibmredbooks

- Look for us on LinkedIn:

https://www.linkedin.com/groups/2130806/

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

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

- Stay current on recent Redbooks publications with RSS Feeds:

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

# Summary of changes

This section describes the technical changes made in this edition of the paper. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for *IBM DS8000 Easy Tier*
*(Updated for DS8000 Release 9.0)*,
as created or updated on February 23, 2022.

## October 2019, Ninth Edition

This revision reflects information regarding the new features that were introduced with the IBM DS8000 with Licensed Machine Code (LMC) 7.9.0 (bundle version 89.0), referred to as Release 9.0, along with the new model IBM DS8900F.
While the DS8900F is an all-flash model, some references to the hybrid DS8880 models are still included to keep IBM Redbooks publication relevant to previous DS8000 models.

### New or changed information

The following new or updated topics are covered:

- New accelerated-mode option for increasing the speed of Easy Tier in certain exceptional situations.
- New DS GUI with Release 9.0.
- Removed most referrals to STAT.

1

# IBM DS8000 Easy Tier

This chapter introduces the basic concepts of dynamic data relocation, and shows how these concepts are implemented by IBM Easy Tier in the IBM DS8000.

This chapter includes the following topics:

- Introduction
- Easy Tier evolution
- Easy Tier license
- Capabilities

## 1.1 Introduction

In modern, complex application environments, the demands for storage capacity and performance lead to issues about planning and optimizing storage resources.

Consider the following typical storage management issues:

- Usually when a storage system is implemented, only a portion of the configurable physical capacity is deployed. When the storage system runs out of the installed capacity and more capacity is requested, a hardware upgrade is implemented to add new physical resources to the storage system.

  This new physical capacity hardly can be configured to keep an even spread of the overall storage resources. Typically, the new capacity is allocated to fulfill only new storage requests. The existing storage allocations do not benefit from the new physical resources. Similarly, the new storage requests do not benefit from the existing resources; only new resources are used.

- In a complex production environment, it is not always possible to optimize storage allocation for performance. The unpredictable rate of storage growth and the fluctuations in throughput requirements, which are I/O per second (IOPS), often lead to inadequate performance. Furthermore, the tendency to use even larger volumes to simplify storage management works against the granularity of storage allocation, and a cost-efficient storage tiering becomes difficult to achieve.

  With the introduction of high performing, but expensive, technologies, such as flash drives, this challenge becomes even more important.

- Any business has applications that are more critical than others, and a need exists for specific application optimization. Therefore, you need to be able to relocate specific application data to faster storage media or exclude slow storage media from a specific volume.

All of these issues deal with data placement and relocation capabilities. Most of these challenges can be managed by having spare resources available and by moving data to optimize storage configurations. However, all of these corrective actions are expensive in terms of hardware resources, labor, and service availability. Relocating data among the physical storage resources dynamically, that is, transparently to the attached host systems, is becoming increasingly important.

Easy Tier can specifically address these issues. In the DS8000, Easy Tier is a built-in dynamic data relocation feature that allows the host-transparent movement of data among the storage system resources as shown in Figure 1-1 on page 3. This feature significantly improves configuration flexibility and performance tuning and planning.

**No charge:** Easy Tier is a no-charge feature of the IBM DS8000.

For more information about the DS8000 series, and the DS8900F in particular, see *IBM DS8900F Architecture and Implementation*, SG24-8456.

**Note:** New systems shipped from manufacturing have Easy Tier set by default to manage all tiered pools.

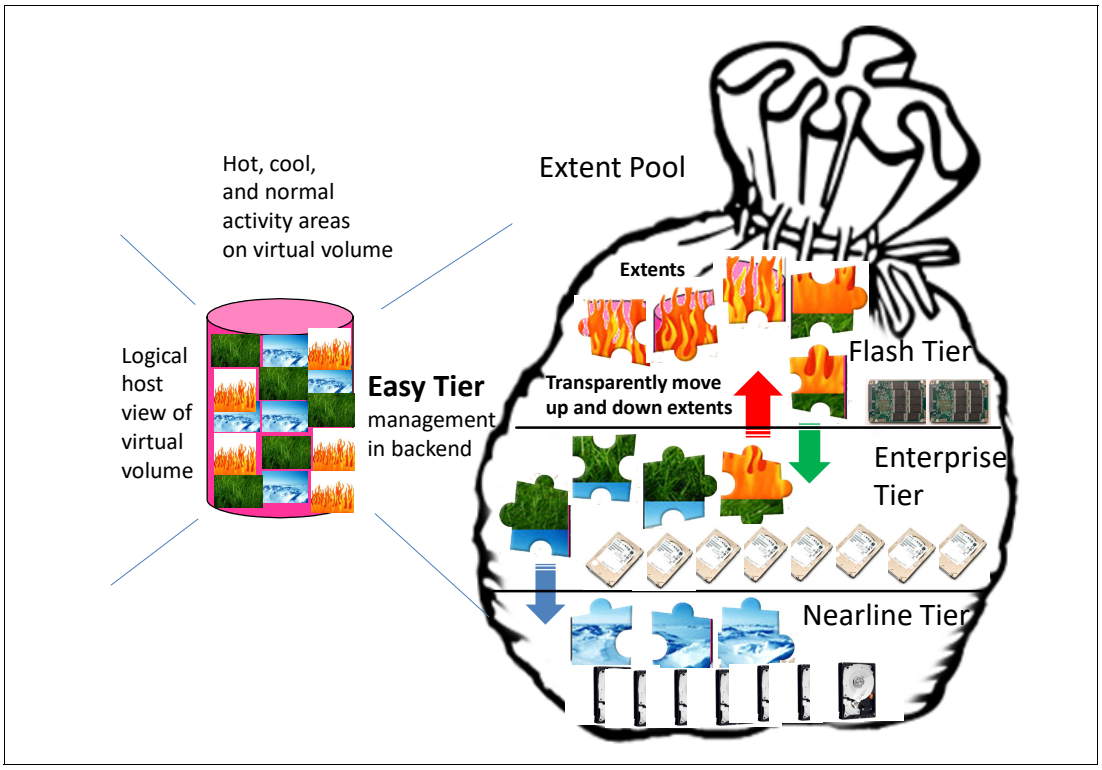



*Figure 1-1 Easy Tier concept*

## 1.2 Easy Tier evolution

The evolution of Easy Tier is summarized in Figure 1-2 on page 5.

DS8000 Release 5.1 introduced automated storage performance management by efficiently boosting enterprise-class performance with flash drives (solid-state drives (SSDs)) and automating storage tiering from enterprise-class drives to flash drives, optimizing flash deployments with minimal costs. It also introduced dynamic volume relocation and dynamic extent pool merge.

Release 6.1 added automated storage economics management by combining enterprise-class drives with nearline drives to maintain enterprise-tier performance while shrinking the footprint and reducing costs with large capacity nearline drives. The second generation also introduced intra-tier performance management (*auto-rebalance*) for hybrid pools and manual volume rebalance and rank depopulation, along with warm and cold demote.

Release 6.2 introduced further enhancements to provide automated storage performance and storage economics management across all three drive tiers (flash, enterprise, and nearline storage tiers). These enhancements allow you to consolidate and efficiently manage more workloads on a single DS8000 system. It also introduced support for auto-rebalance in homogeneous pools and support for thin-provisioned extent space-efficient (ESE) volumes.

Release 6.3 (DS8700 and DS8800) and release 7.0 (DS8870) enhanced the support of Full Disk Encryption (FDE) drives. IBM Easy Tier can perform volume migration, auto performance rebalancing in homogeneous and hybrid pools, hotspot management, rank depopulation, and thin provisioning (ESE volumes only) on encrypted drives and non-encrypted drives.

Release 7.1 introduced the following advanced features:

- Easy Tier Application: Enables administrators to directly affect the placement of logical volumes on specific tiers of a DS8000 based on application and business requirements. Easy Tier Application is described in *DS8870 Easy Tier Application*, REDP-5014.
- Easy Tier Heat Map Transfer: Designed to maintain application-level performance at the secondary site of a DS8000 by transferring the Easy Tier information to the secondary site. See*IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015.

Release 7.3 introduced the following advanced features:

- Support for the flash drives installed in the high-performance flash enclosure (HPFE) only. The use of such flash drives in the DS8900F and DS8880 provides up to four times the performance of other flash drives (SSDs).
- (DS8870 only:) Support for the IBM Flash Adapter 90 attached to an IBM Power Systems™ server, in addition to the EXP30 Ultra SSDs that DS8000 already supported.

Release 7.4 adds the following advanced and innovative features:

- Easy Tier Application control: Easy Tier Application controls provide a more granular and flexible control of workload learning and data migration. They also provide volume-level tier restriction where a volume can be excluded from the nearline tier. A user can manage and control Easy Tier at the pool or volume level.
- Easy Tier Application for IBM Z®: Easy Tier Application for IBM Z provides comprehensive data placement management policy support between an application and storage. With this feature, you need to program the policy only once to enforce it automatically. With hints about data usage and performance expectations, storage is optimized automatically toward higher performance and efficiency.

Release 8.1.1 provides support for small extents. It also allows initial extent allocation on the flash tier, rather than on the enterprised tier. This function is known as the *SSD Home Tier*. It can be enabled or disabled. Finally, a warm promote is now possible to allow promoting extents on a nearline rank when there is sudden significant activity on that nearline rank.

Release 8.3 introduces support for the following features:

- Support for High-Capacity flash (HCF) drives in the DS8880. High-Capacity flash drives are a new large-capacity flash drive option in the DS8880 Gen-2 and DS8900F Flash Enclosures that provide higher capacity than the High-Performance flash drives, but with lower performance.
- Easy Tier reporting in the DS8880 Storage Management GUI. This replaces the function previously provided by the Storage Tier Advisor Tool (STAT). STAT can no longer be used with DS8000 Release 8.3 and later microcodes, but remains available with earlier microcode levels.

Release 8.4 introduces a version of High-Capacity flash drives with 7.68 TB. At the same time, a new tier has been introduced: The 7.68 TB flash, as well as potential further and larger flash drive types, are considered to be the new Flash Tier 2. The 3.84 TB HCF drives make up the new Flash Tier 1.
Additional new Flash Tier 2 drive options like 15.36 TB, as well as (for the small models) 1.92 TB, have been introduced with Release 8.5.

Figure 1-2 shows the evolution of Easy Tier.

| DS8000 Model | Release | Tier support | Auto Mode (Sub Volume) | Manual Mode (Full Volume) |
|---|---|---|---|---|
| DS8700 | R5.1 | Two tier<br>SSD + ENT<br>SSD + NL | • Promote<br>• Swap | • Dynamic extent pool merge<br>• Dynamic volume relocation |
| DS8700<br>DS8800 | R6.1<br>R6.1 | Any two tiers<br>SSD + ENT<br>SSD + NL<br>ENT + NL | • Promote and Swap<br>• Warm and Cold Demote<br>• Auto Rebalance (Hybrid pool only) | • Rank depopulation<br>• Manual volume rebalance |
| DS8700<br>DS8800 | R6.2<br>R6.2 | Any three tiers<br>SSD + ENT + NL | • Three tier support<br>• Auto Rebalance (Homogeneous Pool)<br>• ESE Volume support | |
| DS8700<br>DS8800<br>DS8870 | R6.3<br>R6.3<br>R7.0 | Full support for FDE (encryption) drives | • Automatic data relocation capabilities for all FDE disk environments | • Support for all manual mode command for FDE environments |
| DS8870 | R7.1 | Easy Tier Application<br>Easy Tier Heat Map Transfer<br>Easy Tier Server | • Storage administrators can control data placement with DS CLI<br>• Data placement API to enable software integration<br>• Learning data capture and apply for heat map transfer for remote copy environments<br>• Unified storage caching and tiering capability for AIX servers | |
| DS8870 | R7.3 | Tier 0 support for High-Performance Flash Enclosures | • Intra-tier rebalance for heterogeneous Flash storage pools (HPFE and SSD) | |
| DS8870 | R7.4 | Easy Tier Application API for Z<br>Easy Tier Control | z/OS Db2 hints of data placement at dataset level<br>Allow customer to control learning and migration behavior at pool/volume level<br>Can exclude volumes from Nearline tier | |
| DS8880 | R8.1.1 | Support for small extents | Easy Tier can manage extent pools with small extents | |
| DS8880 | R8.3 | Support for High-Capacity Flash Drives, and tier<br>Easy Tier reporting in DSGUI | | |
| DS8880 | R8.4 | High-Capacity Flash further divided into two tiers: Flash Tier 1 and Flash Tier 2 | | |
| DS8880 | R8.5 | More Flash Tier 2 drive options | | |
| DS8900F | R9.0 | New all-flash model, new GUI | Accelerated Mode option | |

*Figure 1-2 Easy Tier evolution*

Release 9.0 introduces the new model DS8900F as an all-flash model, with no options for spinning drives anymore. The DS8900F has a new DS GUI which it shares with other IBM storage systems, and that also includes the design of the Easy Tier functions. This release also introduces an acceleration mode for Easy Tier, which can be used in certain exceptional situations.

## 1.3 Easy Tier license

For DS8900F storage systems, the Easy Tier function is part of the Basic Function License.

## 1.4 Capabilities

Easy Tier offers the following capabilities to help you optimize the data placement for various scenarios and workloads:

- Automatic Mode
- Manual Mode
- Easy Tier control
- Storage Tier Advisor Tool (for code levels up to release 8.2)
- DS8000 Storage Management GUI

### 1.4.1 Easy Tier Automatic Mode

In Automatic Mode, Easy Tier dynamically manages the capacity in single-tier (homogeneous) and multi-tier (hybrid) extent pools, by auto-rebalancing extents to achieve better performance.

For DS8880, multi-tier extent pools can consist of up to seven drive types which are managed in up to three different tiers. For DS8900F, as there are no spinning disk drives anymore, we have up to three different drive types, in up to three different tiers.

Easy Tier Automatic Mode manages the data relocation across different tiers (inter-tier or cross-tier management) and within the same tier (intra-tier management) without any disruption to host operations.

The cross-tier or inter-tier capabilities aim to relocate the extents of each logical volume to the most appropriate storage tier within the extent pool to improve the overall storage cost-to-performance ratio. In Automatic Mode, Easy Tier also provides an auto-rebalance capability that adjusts the system to continuously provide excellent performance by balancing the load on the ranks within a certain tier in an extent pool.

Easy Tier Automatic Mode manages any combination of up to three tiers using the five drive classes:

- High-Performance Flash, *Flash Tier 0*: Flash drives (in the High-Performance Flash Enclosures), and legacy solid-state drives (SSDs, for older DS8000 models)
- High-Capacity Flash, *Flash Tier 1*: The high-capacity flash drives which are treated as a lower tier than the high-performance flash. Currently, the 3.84 TB are the only flash drive option in this category.
- High-Capacity Flash, *Flash Tier 2*: The high-capacity flash drives of an even lower category, which are treated again as its own and different tier. Currently, the 7.68 TB and

15.36 TB flash drives fall into that category, but also the small 1.92 TB flash. Further higher-capacity flash beyond the 15-TB size will also likely be *Flash Tier 2*.

- Enterprise: 15K and 10K RPM 2.5-inch hard disk drives (HDD - DS8880 and older models only)
- Nearline: 7.2K RPM 3.5-inch HDD (DS8880 and older models only)

Figure 1-3 shows the valid storage pool options. Up to three different tiers are allowed in a single extent pool, with Flash Tier 1 and Flash Tier 2 of the high-capacity flash being two independent tiers. For DS8900F, only the Flash Tiers apply.

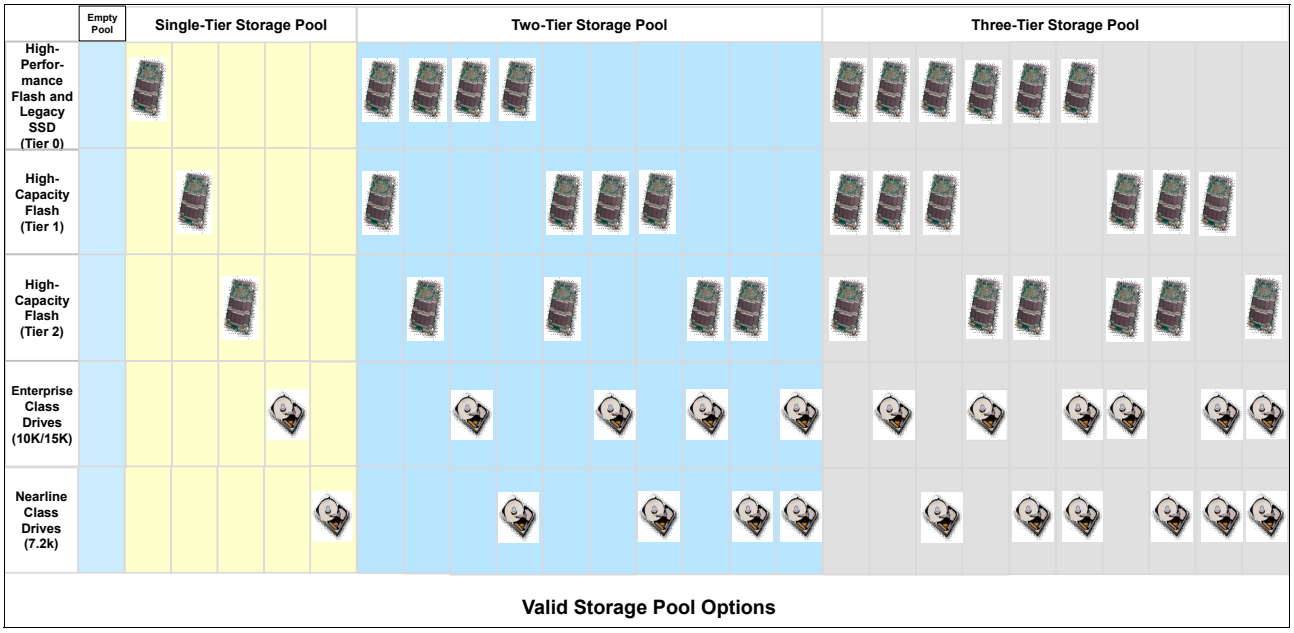



*Figure 1-3 Valid storage pool options*

Figure 1-4 shows a slightly simplified version of the valid combinations of drive classes and their tier assignments in an Easy Tier pool, to maintain readability: High-Capacity Flash consists of potentially two different tiers: Flash Tier 1, and Flash Tier 2.

**Important:** Of the *five* tiers, Flash Tier 0 (800/1600/3200 GB), Flash Tier 1 (3.84 TB), Flash Tier 2 (1.92/3.84/15.36/30.72 TB), Enterprise (10K/15K legacy), and Nearline (7.2K legacy), only *up to three tiers maximum* can coexist in one single extent pool.

| Physical Devices | HP + HC + ENT + NL | HP + HC + ENT | HP + HC + NL | HP + HC | HC + ENT + NL | HC + ENT | HC + NL | HP + ENT + NL | HP + ENT | HP + NL | ENT + NL |
|---|---|---|---|---|---|---|---|---|---|---|---|
| High-Performance Flash (HP) | Not Allowed | 0 | 0 | 0 | | | | 0 | 0 | 0 | |
| High-Capacity Flash (HC) | | 1 | 1 | 1 | 0 | 0 | 0 | | | | |
| ENT 15K/10K HDD (ENT) | | 2 | | | 1 | 1 | | 1 | 1 | | 1 |
| NL HDD (NL) | | | 2 | | 2 | | 1 | 2 | | 1 | 2 |

*Figure 1-4 Easy Tier 3-Tiering Mapping (including legacy hybrid DS8000 models)*

## 1.4.2 Easy Tier Manual Mode

Easy Tier Manual Mode enables a set of manually started actions to manage data among the storage system resources in a dynamic fashion (without any disruption of the host operations).

Manual Mode includes the following capabilities:

- Dynamic volume relocation: Allows a DS8000 to migrate volume to another extent pool.
- Dynamic volume redistribution: Allows a DS8000 to redistribute volume capacity within the pool over the arrays.
- Dynamic extent pool merge: Allows an extent pool to be merged with another extent pool. However, you can only merge extent pools that have the same extent size.
- Rank depopulation: Allows the removal of an allocated rank from an extent pool and relocates the allocated extents to the other ranks in the pool.

Combining these capabilities greatly improves the configuration flexibility of the DS8000 and provides ease of use.

## 1.4.3 Easy Tier control

Easy Tier was designed to make storage management easy and automate everything. However, some control mechanisms were wanted by users. Therefore, Easy Tier has enabled some control at the system, pool, and volume levels. If you know that an unusual workload will run soon, you can disable Easy Tier learning for some time at the pool level and restart learning when the systems returns to normal workload conditions.

You can, for example, suspend Easy Tier to avoid including a temporary batch workload at the end of every month and quarter that is different from the regular workload. After the batch activity completes, the Easy Tier learning process and migration for the impacted pools can be resumed.

You can also move volumes to the highest tier and pin them there for some time to provide the best performance for some special jobs that run, for example, at the end of every month. Then you can unpin the data and let Easy Tier decide where to move the data.

In summary, the following advanced and innovative features are possible with Easy Tier:

- Pause the migration process
- Resume the migration process
- Cancel the migration process
- Restart the migration process
- Assign the tier to the volume
- Unassign the tier from the volume
- Exclude the nearline tier for the volume
- Pause the learning (monitoring) process
- Resume the learning (monitoring) process
- Erase all monitoring data, including migration plans
- Change the tier assignment order

### 1.4.4 Storage Tier Advisor Tool

**Note:** The Storage Tier Advisor Tool has been replaced by integrated reporting in the DS8000 Storage Management GUI (DS GUI) since Release 8.3. For more information, see 1.4.5, “DS GUI: Easy Tier reporting”.

In addition to the host-transparent data relocation capability, Easy Tier provides monitoring capabilities that allow back-end workload data collection. The Easy Tier performance data collection is available on any DS8000.

These Easy Tier performance statistics are used by the Easy Tier Automatic Mode facility to identify data that might benefit from a relocation to a higher or lower storage tier, and to create the appropriate migration plans.

The statistical data can be downloaded and further processed with the IBM DS8000 Storage Tier Advisor Tool (STAT). The STAT tool can still be downloaded from the IBM Fix Central support site for older DS8000 models, for instance DS8870:

https://www.ibm.com/support/fixcentral/swg/selectFixes?parent=Enterprise%20Storage%20Servers&product=ibm/Storage_Disk/DS8870

STAT provides information about the workload heat distribution in the data at the volume level, and it provides configuration recommendations with performance improvement predictions. It allows a closer analysis of the workload characteristics and helps you to evaluate the potential benefits of adding new drive resources, such as flash technology, to the system. Additionally, reports that are generated by STAT can be useful for users to estimate further storage expansion needs.

### 1.4.5 DS GUI: Easy Tier reporting

Easy Tier Reporting is available in the DS GUI. This feature was introduced with Release 8.3 and replaces functions previously provided by the STAT.

In the DS GUI, the Easy Tier reports can be found in the performance reporting section. See Chapter 5, “DS8000 Easy Tier reporting” on page 109 for more details.

2

# DS8000 Easy Tier concepts, design, and implementation

This chapter introduces the basic concepts, design, and implementation of the Easy Tier features that are available on the IBM DS8900F, and some of the earlier DS8000 models.

Depending on the model and release level, some features might not be available.

This chapter includes the following topics:

- Easy Tier features
- General I/O characteristics
- Easy Tier concepts
- Easy Tier support for small extents
- Easy Tier Report integrated in DS GUI
- Easy Tier operating modes
- Easy Tier allocation controls
- Easy Tier design
- Easy Tier control

This chapter and the remainder of this book cover the general functions of Easy Tier Automatic Mode and Manual Mode. For more information about Easy Tier Application, and Easy Tier Heat Map Transfer, see the following publications:

- *DS8870 Easy Tier Application*, REDP-5014
- *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015

## 2.1 Easy Tier features

Easy Tier has two modes:

- Automatic Mode (optional and must be enabled) optimizes storage performance and storage economics management automatically across drive tiers through data placement on a subvolume level in multi-tier or hybrid extent pools. *Multi-tier* or *hybrid extent* pools are storage pools that contain a mix of different drive tiers.

  Easy Tier can nondisruptively relocate data at extent level across different drive tiers or even within the same drive tier to optimize performance and resource usage. The auto-rebalance capability automatically rebalances the workload, even across the physical resources within a drive tier, which reduces the occurrence of hotspots.

- Manual Mode (always active) includes user-requested functions to manually perform tasks on single-tier or multi-tier pools without any service outage. A user can merge extent pools, depopulate ranks, migrate volumes, and control Easy Tier. Also, you can relocate volumes within the same extent pool to evenly restripe the volumes and balance the volumes' capacities across all the ranks in that extent pool.

  Manual volume redistribution applies to only single-tier or homogeneous extent pools that are not already enabled for Easy Tier Automatic Mode management.

**Modes:** Easy Tier Automatic Mode and Manual Mode are not exclusive, which means that you can use Manual Mode capabilities even if Automatic Mode is active.

All of these functions are performed dynamically without interrupting host I/Os within the DS8000.

You can configure Easy Tier to improve performance by managing or monitoring volume capacity placement in pools. You can enable the allocation order that Easy Tier uses to select the drive classes when capacity is allocated in pools.

Additional Easy Tier settings and controls are available at the system or pool level. These settings and controls are explained in Chapter 4, "DS GUI and DS CLI support for IBM DS8000 Easy Tier" on page 63.

## 2.2 General I/O characteristics

The workload and I/O characteristics of any installation are unique, even though general or similar patterns exist.

Many installations generally perform more random I/Os during the day when online transactions are the main activity. After the main online period finishes, the batch period starts, during which time the I/Os become more sequential in nature.

A volume or logical unit number (LUN) that is active during the online period might become almost idle during the batch period. A busy volume during the online period on one day might not have much activity on the following day.

Financial institutions are a good example. The market-open period is a critical time because of the high volume transactions during a short period. Month-end, quarter-end, and year-end periods are generally times when I/O transactions might double, or more, compared to normal days.

With this variability, tracking the changes in the I/O characteristics of each volume is nearly impossible, in addition to managing the location or placement of those volumes based on their level of activity. Even within a volume, hotspots exist, which means that the I/Os are not evenly distributed across all extents within the volume.

Certain extents are hotter (more active) compared to other extents within the same volume. Therefore, any attempt for manual performance micro-management at the extent level with a manual placement of data extents on the appropriate storage resource is an impossible task.

**Tip:** you can control Easy Tier by suspending the learning (monitoring) process to avoid the effects of high transaction loads that are generated by specific, one-time processing. To suspend monitoring, set the Easy Tier mode to Disable (see “Easy Tier Controls” on page 64).

## 2.3 Easy Tier concepts

The availability of storage drives with various performance characteristics allows the user to directly place data on the most appropriate storage according to performance and relevant business requirements. You can complete this task with a manual approach by using existing data migration tools or the Easy Tier Manual Mode facilities, as described in 2.6.2, “Easy Tier Manual Mode” on page 20. Nevertheless, the manual approach includes the following limitations:

- Data granularity: In Manual Mode, the data placement is at the volume level. But with the ever increasing size of the volumes, this level of granularity can lead to inefficient usage of valuable and costly higher-performance storage resources.
- Performance requirement variability: In complex application environments, the workload characteristics are variable. This situation leads to more efforts to move the volumes in a timely manner according to the workload performance changes.

Furthermore, moving the data across the storage system resources (that is, physical drives) is always an expensive task in terms of labor and service availability. Nonetheless, data relocation is often required to adapt to physical configuration changes and new performance requirements, especially when various drive technologies are implemented on the same hardware.

**Important:** The Easy Tier Application feature can also be considered a way to manually control the data placement (at the volume and extent level), and it cooperates with Easy Tier Automatic Mode. For more information about Easy Tier Application, see *DS8870 Easy Tier Application*, REDP-5014.

With the availability of high-performance flash drives, high-capacity flash drives, and solid-state drives (SSD) in the DS8000, the question arises about how to use such faster drives effectively and efficiently. Considering the superior performance of flash compared to spinning drives (enterprise or nearline drives), replacing all spinning drives with flash on the storage system is advantageous.

The cost of flash is becoming more affordable, and IBM now moves to all-flash DS8000 systems. From an economic point of view, a balance must exist between cost and performance. With flash drives existing in different categories and tiers, the right tiering and price mix can be found in the course of the initial sizing, while often having the bulk of the capacity in lower-tier storage. And still, taking advantage of low flash response times.

For hybrid (legacy) DS8000 systems having the capacity bulk in HDDs, already a smaller flash percentage like 10...20% gives a bigger latency improvement and often allows a combination with larger-capacity second-tier HDDs, like 1.8 TB/10K rpm disk drives. With the introduction of the High-Capacity Flash drives and their tiers, you have more options to combine High-Performance flash and High-Capacity flash drives economically.

The 7 drive types (=including those for older hybrid DS8000 types) are grouped into 5 different drive classes:

- High-performance flash drives, *Flash Tier 0* (this is usually smaller-capacity flash)
- High-capacity flash drives, *Flash Tier 1* (currently 3.84 TB only)
- High-capacity flash drives, *Flash Tier 2* (flash drives of usually a very large capacity, RAID-6)
- Enterprise drives: This class includes the SAS (10K or 15K RPM) disk drives (legacy DS8000 models only)
- Nearline drives: This class is for the Nearline (7.2K RPM) disk drives, which provide large data capacity but lower performance (legacy DS8000 models only).

A hybrid extent pool can hold up to three different tiers of these drive classes.

Extent pools that contain ranks of more than one drive class are called *hybrid* or *multi-tier* pools. Extent pools that contain ranks of only a single drive class are referred to as *homogeneous* or *single-tier* pools.

Although Easy Tier can manage four drive classes, you are only allowed to have three drive classes in any hybrid pool.

**Important:** A hybrid pool can hold two or three drive classes, but not more.

If RAID-10 and RAID-6 drives are in the same extent pool, or if both 10K and 15K RPM disk drives are in the same extent pool, the drives are managed as a single tier. In both cases, the rank saturation for different rank types (for example, 10K RAID-5 and 15K RAID-5) can differ. The workload rebalancing within a single tier takes the rank saturation into consideration when attempting to achieve an equal level of saturation across the ranks within a tier.

**Note:**

- High-capacity flash drives are treated as a lower tier than high-performance flash drives, when those are present.
- If both 3.84 TB and 7.68/15.36/1.92 TB drives are present, they will be treated as different tiers.
- The 15K RPM enterprise disks are treated as the same tier as 10K RPM drives. However, the intra-tier auto-rebalance (micro-tiering) function of Easy Tier can use the higher IOPS capability of the 15K enterprise disks.

Other terms that are often used in this book are *hot data*, *warm data*, and *cold data*. These terms match the three priority levels that can be applied to extents that are monitored by Easy Tier Automatic Mode to define a migration plan for those extents.

**Buckets:** The Easy Tier Automatic Mode internal algorithm manages a higher number of priority levels (also called *buckets*) and then a finer granularity for extent temperature monitoring. Extent distribution in those buckets is computed dynamically for each migration plan, depending on the extents' activity. For more information about migration plans, see 2.8, "Easy Tier design" on page 28.

*Hot data* refers to data that has more I/O workload compared to other extents in the same extent pool and in the same tier. The lower the percentage of hot data, the greater the skew in the environment, and the higher the benefit of a few flash drives. Hot data is promoted to a higher tier with a significant performance improvement.

*Cold data* is data that has low workload levels or no workload activity. Cold data does not benefit from a higher tier and is not promoted. It is demoted to the lowest available tier.

*Warm data* is the rest of the workload that is not hot or cold and might be promoted (if this action leads to a performance improvement) or demoted, depending on the precise workload levels and configuration.

When we refer to *data* here, we mean *extent* because when Easy Tier is used in Automatic Mode, it performs its operations at the extent level.

Business users know the applications that are more important and business-critical. Based on that information, the storage administrator can determine the relevant volumes that are used by those applications. After the volumes are identified and their sizes are known, the storage administrator can determine how many higher-performance tier ranks are needed to accommodate all of those volumes. Placing those volumes on higher-performance tier ranks can provide a significant performance improvement to the application.

The drawback is that a significant number of higher performance ranks will likely be required, which makes them cost prohibitive. In addition, not all volumes are hot, so you might be inefficiently placing cold volumes on higher performance ranks. Furthermore, even on a volume that is considered hot, often not all extents on that volume are hot. This situation can mean that you are not using the available storage on higher performance ranks as effectively as you might prefer.

## 2.4 Easy Tier support for small extents

Two different extent sizes are supported for extent pools:

- An extent pool can consist of all large extents, where a large extent is 1 GiB for FB pools and 1113 cyl extents for CKD pools.
- An extent pool can consist of all small extents, where a small extent is 16 MiB for FB pools and 21 cyl for CKD pools.

It is not practical to monitor and manage each small extent individually as Easy Tier did when only large extents were supported. With the arrival of small extents, the concept of *track group* was introduced.

For large extents, a track group is identical to a large extent. For small extents, a track group is the collection of 64 consecutive small extents of a virtual FB volume, or a collection of 53 consecutive small extents of a virtual CKD volume.

So, a track group has the same size as a large extent because 64 x 16 MiB = 1 GiB and 53 x 21 cyl = 1113 cyl, as illustrated in Figure 2-1. This implementation is a compromise because it allows some management at the extent and at the track group level.

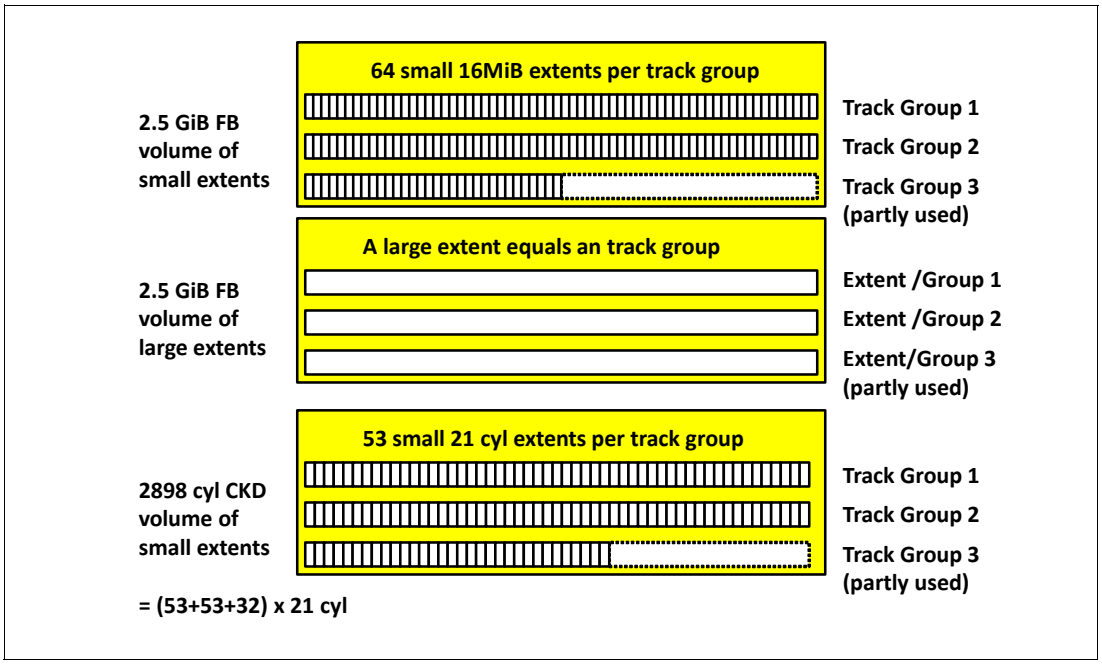



*Figure 2-1 Easy Tier's concept of a track group*

When a volume is initially created in a hybrid pool made of small extents, any track group will reside on a specific tier. However, over time, Easy Tier promotes or demotes extents in that track group based on I/O activity.

This process means that a particular track group will eventually exist on three tiers, each of which is independently monitored. Easy Tier manages statistics for each track group and for every tier that the track group is present on.

Each track group has three heat values, one for small extents in Tier 0 of that track group, one for all small extents in Tier 1 of that track group and one for all small extents in Tier 2 of that track group (see Figure 2-2).

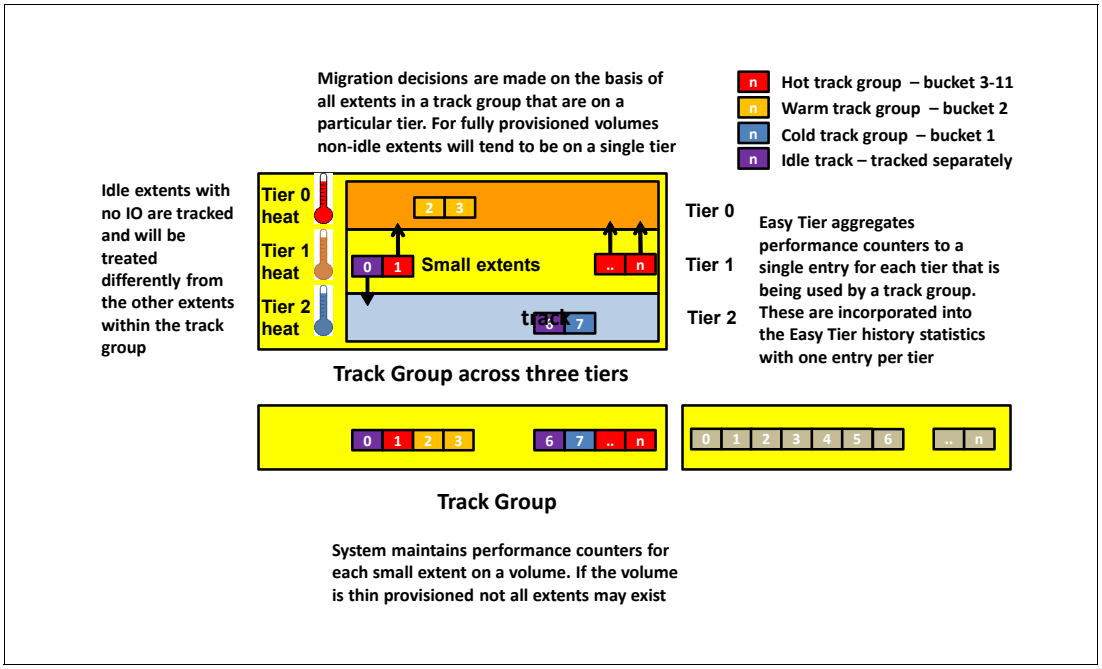



*Figure 2-2 Heat of small extents in track groups*

To further optimize efficiency, Easy Tier also tracks *idle* extents at the small extent level. These are extents that have not had any I/O within a defined time period.

If a track group is promoted from one tier to another, then the idle extents in the track group will *not* be promoted. For demotion of cold extents, the idle extents for all track groups will be demoted first before the entire track groups are considered. Idle extents will also be the first candidates for demotion selection as part of a swap.

**Note:** When in the following sections we talk about extents or track groups being moved around by Easy Tier, we also mean the small extents of a tier of a track group.

### *Capacity limits with small extents*

Many clients now prefer the small extents, given that they offer better space efficiency for instance when working with extent space-efficient volumes (ESE). In addition, for machine capacities of up to a few hundred TB, this is a very good choice. Please note, however, the monitoring capability of Easy Tier with small extents is limited in the sense that for total volume capacities of 4 PB and more, the large extents (1 GiB/1113 cylinders) should be used, when creating pools.

## 2.5 Easy Tier Report integrated in DS GUI

Easy Tier enables dynamic data relocation to achieve the optimum performance on the DS8000. Now it can be of interest to study the behavior and actions of Easy Tier itself, for a certain storage system and its respective pools.

Up until DS8000 Release 8.2 code, the IBM Storage Tier Advisor Tool (STAT) for DS8000 was available to analyze the data that is collected by the Easy Tier monitoring function to produce a Volume Heat Distribution or heat map.

Starting with DS8000 R8.3, you are not able to offload the binary heat data and use the STAT to parse it. STAT remains available for use with the older code levels, for instance for gathering information when replacing these legacy models.

You can now directly offload the three CSV files (data movement, skew curve, and workload) and the Excel tool (STAT Charting Utility) from both the DSGUI and DSCLI. This technique enables you to perform these tasks:

- Get the skew curve CSV file for Disk Magic modeling
- View the detailed data for Easy Tier planning, monitoring, and debugging

The raw data to generate the Easy Tier workload categorization report and migration report can be obtained from either the DSGUI or the DSCLI, as shown in Figure 2-3.

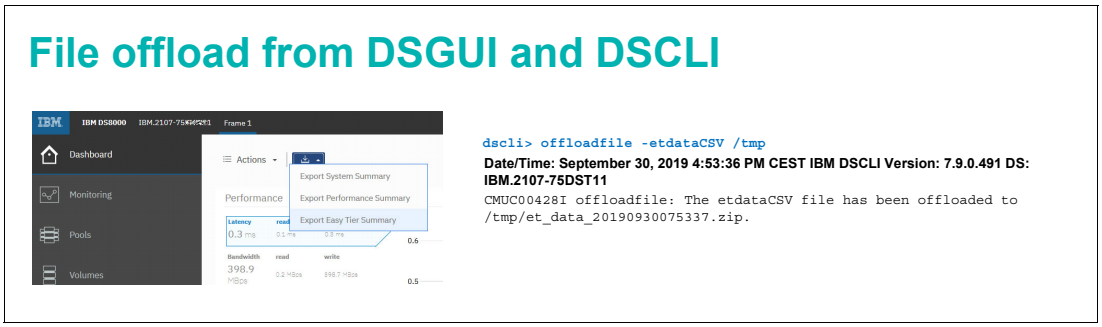



*Figure 2-3 DS GUI and DS CLI File offload*

Additionally, there are performance views in the DSGUI to observe the Easy Tier behavior on pool and volume level directly. See Chapter 5, “DS8000 Easy Tier reporting” on page 109 for more on these views, and how to work with the offloaded data.

## 2.6 Easy Tier operating modes

More details about the Easy Tier Automatic Mode and Manual Mode of operation are described.

### 2.6.1 Easy Tier Automatic Mode

Easy Tier Automatic Mode can optimize the data placement on DS8000 automatically and economically. Easy Tier dynamically manages the capacity in single-tier (homogeneous) extent pools (auto-rebalance) and multi-tier (hybrid) extent pools that contain up to three different drive tiers.

Easy Tier Automatic Mode can be enabled for all extent pools (including single-tier pools) or for multi-tier pools only. It can also be disabled, which means that no extent pools are monitored. Extent pools that are handled by Easy Tier are referred to as *managed* pools, and those that are not handled as *non-managed* pools. Easy Tier Automatic Mode supports regular and thin-provisioned extent space efficient (ESE) volumes with small or large extents. Easy Tier Automatic Mode manages the data relocation both across different tiers (inter-tier or cross-tier management) and within the same tier (intra-tier management).

The cross-tier or inter-tier capabilities deal with the Automatic Data Relocation (ADR) feature that aims to relocate the extents of each logical volume to the most appropriate storage tier within the extent pool to improve the overall storage cost-to-performance ratio. This task is done without any user intervention, and is not apparent to the application host.

Logical volume extents with high latency in the rank are migrated to storage media with higher-performance characteristics. Extents with low latency in the rank are kept in storage media with lower performance characteristics.

After a migration of extents is finished, the degree of hotness for extents does not stay the same over time. Eventually, certain extents on a higher performance tier become cold, and other extents on a lower-cost tier become hotter compared to cold extents on the higher-performance tier. When this event happens, cold extents on a higher-performance tier are eventually demoted or swapped to a lower-cost tier and replaced by new hot extents from the lower tier.

Easy Tier always first evaluates whether the “cost” of moving an extent to a higher-performance tier is worth the expected performance gain. This migration scenario is shown in Figure 2-4.

**Note:** Easy Tier tries to keep extents in a higher-performance tier, if possible. Extent relocation to a lower tier can occur proactively (cold demote), reactively (warm demote), or it can occur if a higher tier is running out of space.

Easy Tier moves data gradually to avoid contention with I/O activity that is associated with production workloads. It does not move extents unless a measurable latency benefit can be realized by the move. The impact that is associated with Easy Tier management is so small that the effect on the overall system performance is nearly undetectable.

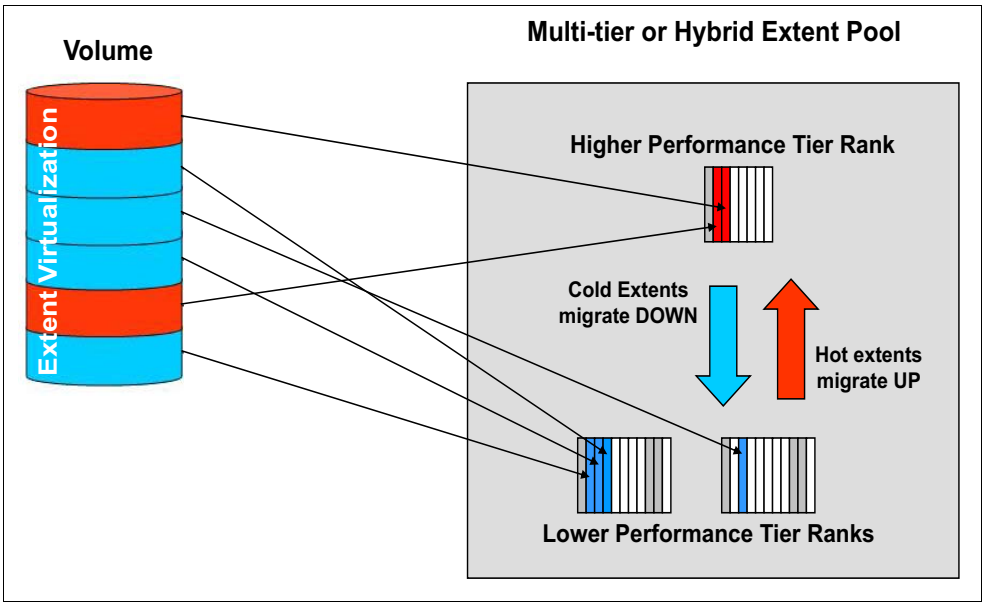



*Figure 2-4 Easy Tier Automatic Mode*

Easy Tier Automatic Mode provides intra-tier capabilities (auto-rebalance) that perform an automatic extent rebalancing within a tier of a managed extent pool. This action allows the systems to avoid hotspots or rank saturation that might result from an unbalanced workload distribution across the ranks of the same drive tier. Auto-rebalance also takes advantage of new ranks when they are added to the storage pool by automatically relocating some data to these ranks and redistributing the workload to achieve a balanced rank usage.

In addition to the inter-tier management, Easy Tier Automatic Mode can handle storage device variations within a tier by using the intra-tier (also called *micro-tiering*) capability. An example of storage device variations within a tier is when an intermix of ranks with different RAID levels or different drive RPM (for legacy DS8000s) exists within the same storage tier of an extent pool. A mix of some RAID-6 tiers of a certain drive kind with some RAID-10 of the same drives would be a typical micro-tiering scenario.

In these configurations, the Easy Tier Automatic Mode micro-tiering capability considers the different performance profiles of each micro-tier. It then performs inter-tier and intra-tier (auto-rebalance) optimizations.

Easy Tier does not handle a micro-tier like an additional tier. It is still part of a specific tier. For this reason, the extent “hotness” does not trigger any promotion or demotion across micro-tiers of a same tier, and the extent relocation across micro-tiers can occur only as part of the auto-rebalance feature.

**Note:** In Automatic Mode, Easy Tier data in memory persists in local storage, ensuring that the Easy Tier configurations are available at failover, cold start, or Easy Tier restart.

### 2.6.2 Easy Tier Manual Mode

Easy Tier Manual Mode provides the following extended capabilities for logical configuration management:

- Dynamic extent pool merge
- Dynamic volume relocation
- Manual volume redistribution
- Rank depopulation capability
- Pool-level Easy Tier control:
  - Suspend or resume Easy Tier learning
  - Reset Easy Tier learning
  - Suspend or resume extent relocation
  - Query pool-level Easy Tier control
- Volume-level Easy Tier control:
  - Suspend or resume Easy Tier learning
  - Reset Easy Tier learning
  - Query volume-level Easy Tier control state
  - Exclude volume from nearline tier control

Easy Tier Manual Mode capabilities are user-initiated. The standard DS8000 series management interfaces (DS command-line interface (DS CLI) and graphical user interface (DS GUI)) can be used to start the Easy Tier Manual Mode functions.

**Note:** The DS8000 GUI does not include all Easy Tier Manual Mode features. Because Easy Tier control is geared toward advanced users, control-related activities are available from the DS CLI.

## Dynamic extent pool merge

The dynamic extent pool merge capability enables a user to initiate a merging process of at least two extent pools. During this process, all volumes inside all extent pools remain accessible to the hosts. The dynamic merge process can leave volumes on current ranks, or redistribute volumes inside a merged pool.

**Limitations:** The dynamic extent pool merge is allowed only among extent pools with the same server affinity or rank group. Dynamic extent pool merge is not allowed in the following circumstances:

- If pools have different storage types (fixed block (FB) and count key data (CKD))
- If the extent pools have different extent sizes
- If you selected an extent pool that contains volumes that are currently migrated
- If the combined extent pools have 4 PB or more of ESE logical capacity

A dynamic extent pool merge consists of the following operations (the following numbers correspond to the numbers in Figure 2-5):

1. Changing the assignment of ranks in the source extent pool to the target extent pool
2. Changing the assignment of volumes and space-efficient repositories in the source extent pool to the target extent pool
3. Merging the repositories of source and target pools
4. Deleting the source extent pool that no longer has ranks or volumes assigned

**Note:** A new, merged extent pool can be assigned a new name.

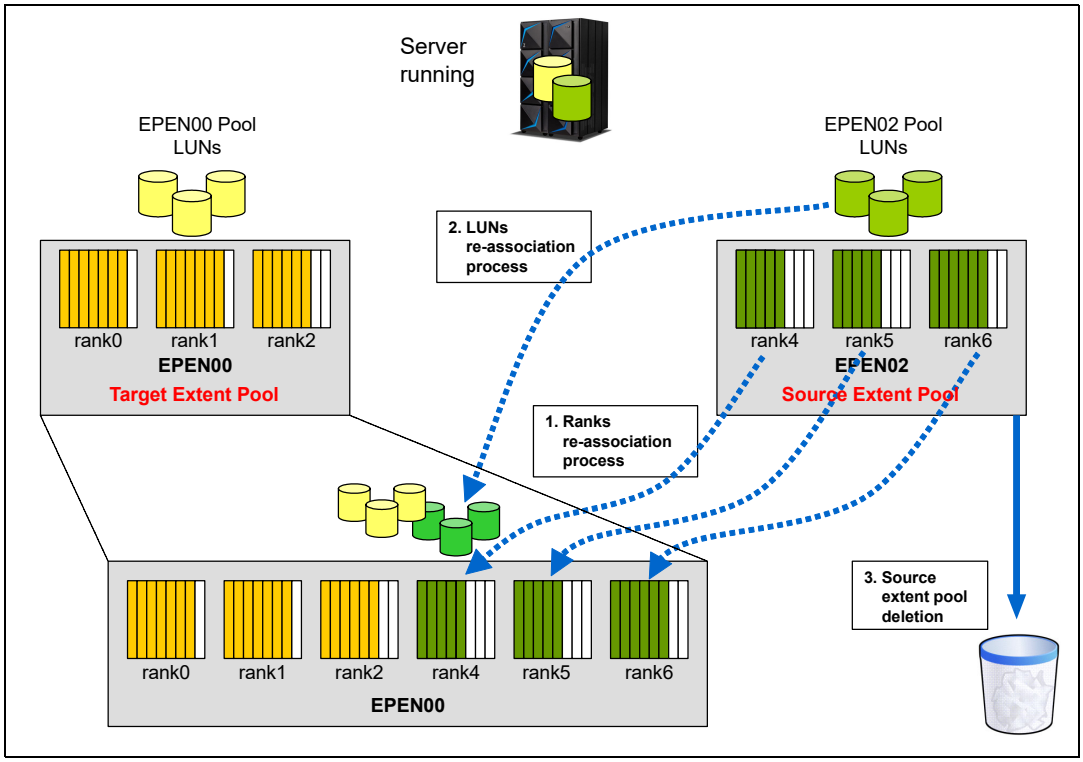



*Figure 2-5 Dynamic extent pool merge process without volume redistribution*

An extent pool merging process can be started to perform the following operations:

- Homogeneous extent pool consolidation

  Creating larger homogeneous pools can enable new volumes to be spread over a greater quantity of physical resources, which improves the overall performance.

- Hybrid extent pool creation

  Consolidate two extent pools to create a merged multitier extent pool with mixed storage technologies that are managed by the Easy Tier Automatic Mode facility.

## Dynamic volume relocation

Dynamic volume relocation, which is also known as *migration*, enables a user to start a volume move from its current extent pool (*source* extent pool) to another extent pool (*target* extent pool). During the volume migration process, the volume remains accessible to hosts.

**Limitations:** The dynamic volume relocation is allowed only among extent pools with the same server affinity or rank group. Additionally, the dynamic volume migration is not allowed in the following circumstances:

- If source and target pools have different storage types (FB and CKD)
- If source and target extent pools have different extent sizes

Figure 2-6 shows a typical scenario where you want to change the underlying storage characteristics for a specific volume. The scenario includes the following process:

1. A host server accesses volume A (the red volume in the picture).
2. You decide to move volume A from its current storage allocation in enterprise ranks to nearline ranks.
3. You start a migration by using dynamic volume relocation of volume A from extent pool EPEN00 to target extent pool EPNL02.
4. The relocation process begins pre-allocating the necessary extents to relocate volume A in the target extent pool. The extent distribution in the target extent pool is determined by the specified extent allocation method (EAM), rotate extents (storage pool striping), or rotate volumes. If the EAM is not changed with the migration, the volume's current EAM is maintained. In this example, the EAM is rotate extents.
5. The actual volume A relocation starts, moving extents from EPEN00 to EPNL02.
6. When the relocation of any single extent is complete, the migrated extent is freed.
7. The process completes when all volume A extents are migrated to EPNL02.

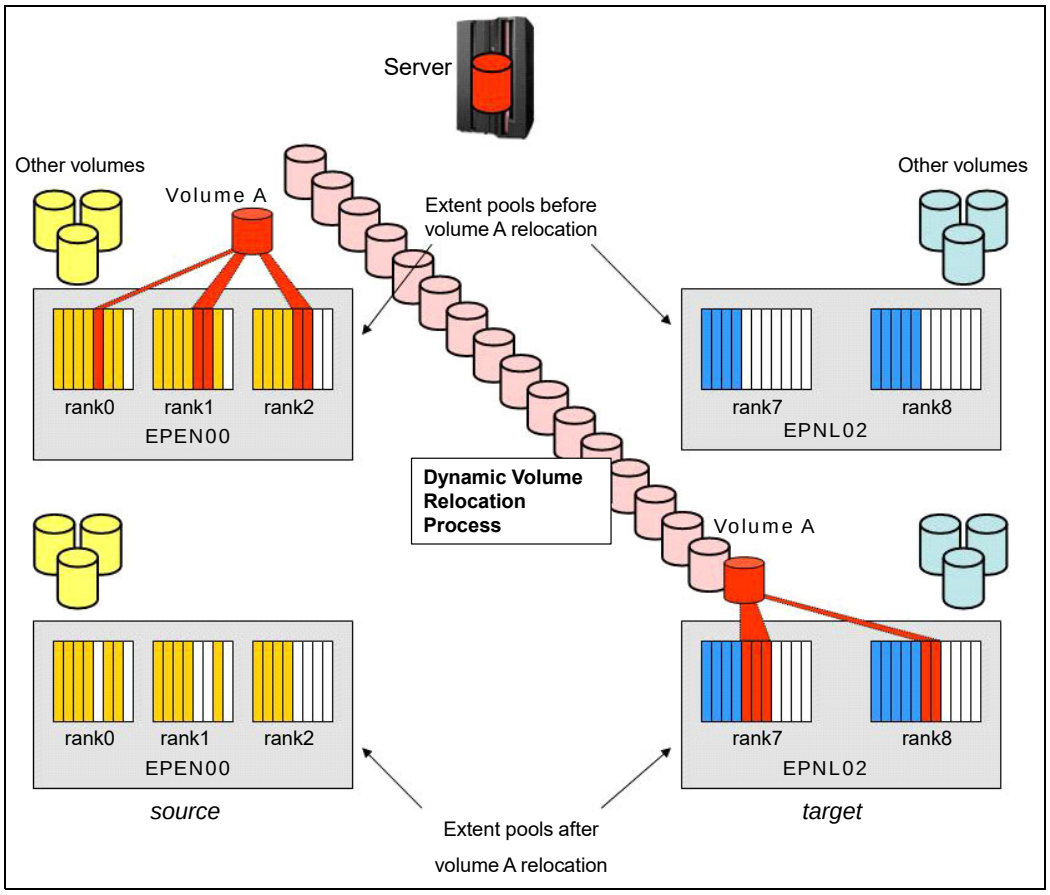



*Figure 2-6 Dynamic volume relocation between two pools*

## Manual volume rebalance

Manual volume rebalance is designed to redistribute the extents of volumes within a non-managed, single-tier (homogeneous) pool so that workload skew and hotspots are less likely to occur on the ranks. Manual volume rebalance is an enhancement of the dynamic volume relocation feature.

A volume is migrated back into the same extent pool with the source extent pool as the target extent pool and the EAM as rotate extents (the use of storage pool striping). In this case, the algorithm tries to evenly spread the volume's extents across all ranks in the extent pool.

This feature is especially useful for redistributing extents after new ranks are added to a non-managed extent pool or after homogeneous extent pools merge (without volume redistribution during the merge) to balance the capacity and the workload of the volumes across all available ranks in a pool.

Manual volume rebalance provides manual performance optimization by redistributing the capacity of a volume within a non-managed homogeneous extent pool. It can also be referred to as *capacity rebalance* because it balances only the extents of the volume across the ranks within an extent pool without actually taking any workload statistics or device usage into account.

A balanced distribution of the extents of a volume across all available ranks in a pool is supposed to provide a balanced workload distribution, and to minimize skew and hotspots. It is also referred to as *volume restriping*, and is supported for standard and ESE thin-provisioned volumes.

During extent redistribution by using manual volume rebalance, only one extent at a time is allocated rather than pre-allocating the full volume. Only a minimum amount of free capacity is required in the extent pool, and only the required number of extents is relocated.

If a manual volume rebalance operation is run in an existing pool with only partial capacity available on some ranks and no capacity available on other ranks, it might not always achieve a balanced distribution of the volume capacity across all ranks. For example, one rank is allocated by another volume by using an EAM of rotate volumes. In this case, you might need to perform the relocation in a certain sequence or in multiple steps to ensure that you have at least some free extents available on each rank in the pool.

See "Manual volume rebalance" on page 86 for an example of the use of the DS CLI.

**Manual volume redistribution:** Manual volume redistribution is not supported in the following situations:

- In managed extent pools, because these pools are already under control of the Easy Tier Automatic Mode algorithm.
- Generally in hybrid or multi-tier pools (managed or non-managed), which are assumed to be prepared for Easy Tier Automatic Mode management.

Figure 2-7 on page 25 shows a typical scenario in which you perform a manual volume redistribution to better distribute the allocated capacity for a specific volume across the physical resources within a non-managed, single-tier, or homogeneous extent pool. The scenario includes the following tasks:

1. A server accesses volume B. Volume B is in a non-managed, single-tier extent pool.
2. You decide to spread volume B from its current storage allocation on one rank (rank 0) to all ranks in the extent pool (rank 0, rank 1, and rank 2).
3. You start a volume migration by using dynamic volume relocation for volume B, specifying the target extent pool to be the same as the source extent pool where volume B currently exists, that is, EPEN00.

To ensure that the extents of the volume are spread evenly across all available ranks in the extent pool, the rotate extents EAM attribute must be specified for storage pool striping. For more information, see 4.1.8, “Dynamic volume relocation” on page 70, or 4.3.2, “Dynamic volume relocation” on page 82.

4. The relocation process begins pre-allocating the extents to relocate volume B on the target ranks. The pre-allocation process mitigates the extent availability requirements, and allocates only one extent at a time rather than pre-allocating the full volume. Only the minimum number of required extents to achieve a balanced extent distribution of the volume is relocated.
5. The actual relocation of volume B starts moving one extent at a time from rank 0 to the other ranks.
6. When the relocation of an extent is completed, the migrated extent is freed.
7. The process ends when all required extents of the volume B extents are migrated.

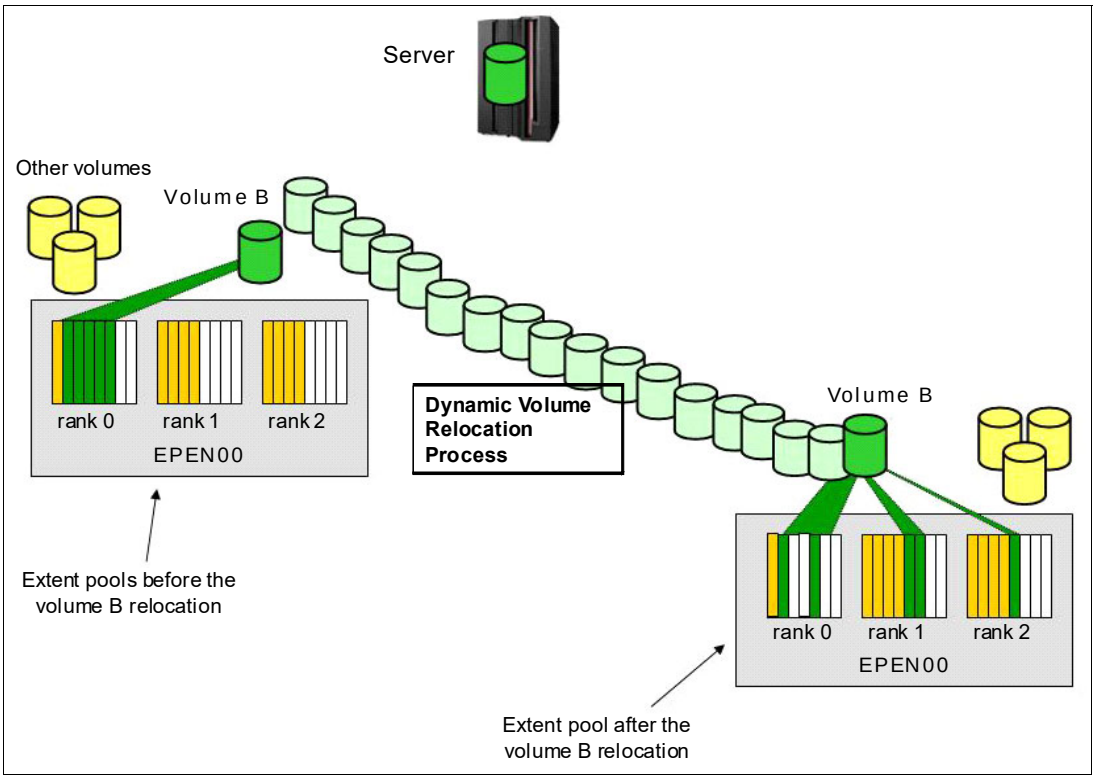



*Figure 2-7 Manual volume redistribution in a non-managed pool*

## Rank depopulation

Rank depopulation is an Easy Tier Manual Mode capability that enables a user to unassign a rank from an extent pool, even if the rank has extents that are allocated by volumes in the pool. For the rank to be unassigned, Easy Tier automatically attempts to migrate all allocated extents to other ranks within the same extent pool.

**Tip:** The rank depopulation feature can also be used to free up ranks and reassign them to another extent pool. It offers you the possibility to migrate rank from extent pools created with large extents to new extent pools using small extents.

During this process, the affected volumes remain accessible to hosts.

**Notes:** If no free space exists inside the extent pool to accommodate the capacity of all of the volumes after rank depopulation, the process cannot be initiated from the GUI. From the DS CLI, the following message is issued:

```
CMUN80204E chrank: Rxx: The unassign rank task could not be initiated because
insufficient free extents are available in the other ranks of the pool.
```

For thin provisioning, the depopulation of a rank that contains ESE volumes is supported.

Figure 2-8 shows the operations that occur when a user wants to unassign a rank in an extent pool that has extents that are already allocated to some volumes.

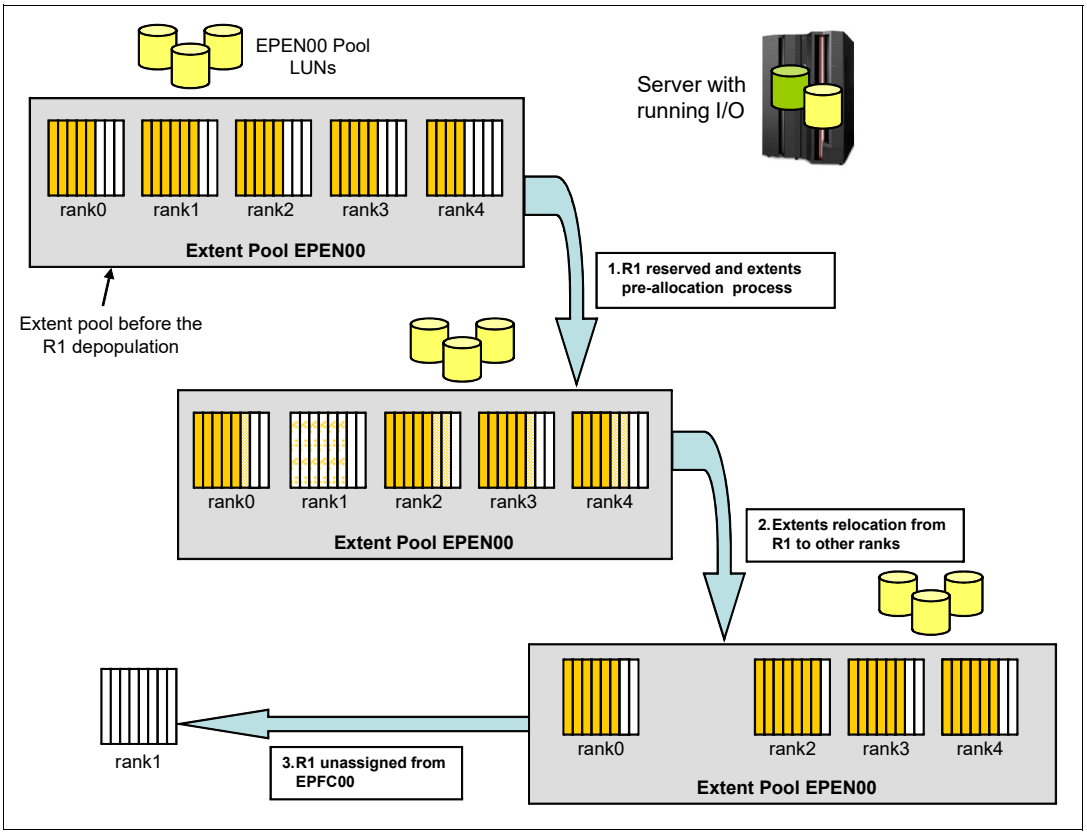



*Figure 2-8 Rank depopulation*

The process includes the following steps:

1. You decide to unassign rank R1 from extent pool EPEN00, for example, to assign it to another extent pool.
2. You start rank unassignment from the extent pool by using the DS GUI or DS CLI.
3. Easy Tier automatically reserves the rank to prevent new extent allocations. The state of the rank then changes to Depopulating.
4. The depopulation process pre-allocates extents in the other ranks of the pool, which are needed to relocate the extents that are allocated in the R1 rank.
5. The actual extent relocation starts moving, in an asynchronous way, a few extents at a time, from R1 to other ranks in the pool.
6. When the relocation of any single extent is completed, the migrated extent is freed.

The process ends when all extents that were on R1 are relocated to the other ranks in the pool. Easy Tier unassigns R1 from the pool, and the state of the R1 rank then changes to `Unassigned`.

### Tools to identify hot volumes

Easy Tier Manual Mode can actively assist with performance tuning by manually moving volumes to the appropriate extent pool. This process is particularly useful for mixed-technology configurations with multiple homogeneous extent pools of different drive tiers.

Dynamic volume relocation can be used to manually manage the volume placement onto the most appropriate storage media and extent pool. To use this mode, you first must identify the hot volumes.

The following tools are available to identify hot volumes:

- You can monitor Easy Tier directly from the DSGUI by using these reports:
  - Workload categorization report
  - Migration report

  For details, see Chapter 5, “DS8000 Easy Tier reporting” on page 109.

  Prior to DS8000 R8.3, you could use the tool STAT for this purpose.
- FLASHDA tool for IBM z/OS®

  This flash device adapter tool uses SAS code (`http://www.sas.com`) to analyze System Management Facilities (SMF) record type 42 subtype 6 and SMF record type 74 subtype 5. The SAS code helps to identify volumes and data sets that are good candidates to be on flash. For more information, see the z/OS Downloads website:

  `https://www.ibm.com/servers/resourcelink/svc00100.nsf/pages/zosDownloads`
- IBM Spectrum™ Control and IBM Storage Insights also allow you to identify hot volumes because they show I/O statistics at the volume level. For more information about IBM Spectrum Control™ or Storage Insights, see the product websites:

  `https://www.ibm.com/marketplace/spectrum-control`
  `https://www.ibm.com/marketplace/analytics-driven-data-management`

Based on the output reports of these tools, you can use the FLASHDA tool for z/OS. The tool output also identifies the “hot” data at the data set level. Based on this data, the flash ranks can be migrated by data set by using the appropriate z/OS tools.

## 2.7 Easy Tier allocation controls

For Easy Tier to function effectively, it needs some free extents in each extent pool to be able to move extents around. A storage administrator can keep an eye on this and make sure that not all capacity is allocated. However, a control lets the system reserve some space for Easy Tier extent movements. By default, the control is set to reserve space. If necessary, the **`chsi`** command with the **`-etsrmode disable`** option can be used to disable automatic space reservation for Easy Tier.

Another control is the allocation policy for new volumes. In implementations before microcode 8.3, new storage was allocated from Tier 1 if available and then from Tier 2, but Tier 0 was not preferred. You can change the allocation policy according to your needs with the **`chsi`** command and the **`-ettierorder highutil`** or **`highperf`** option.

With later code implementations, the allocation policy for new volumes changed according to the allocation order shown in Table 2-1, for hybrid configurations, and Table 2-2, for all-flash configurations.

**Note:** The default allocation order is `High Utilization` for hybrid pools, and `High Performance` for all-flash pools.

*Table 2-1 Hybrid volume allocation order*

| **DS8880 Hybrid** | |
|---|---|
| **High Utilization** | **High Performance** |
| Enterprise tier | Flash tier 0 (high performance) |
| Flash tier 1 (high capacity) | Flash tier 1 (high capacity) |
| Flash tier 2 (high capacity) | Flash tier 2 (high capacity) |
| Nearline tier | Enterprise tier |
| Flash tier 0 (high performance) | Nearline tier |

*Table 2-2 All-Flash volume allocation order*

| **DS8900F, and DS8880 All-Flash** | |
|---|---|
| **High Utilization** | **High Performance** |
| Flash tier 1 (high capacity) | Flash tier 0 (high performance) |
| Flash tier 2 (high capacity) | Flash tier 1 (high capacity) |
| Flash tier 0 (high performance) | Flash tier 2 (high capacity) |

## 2.8 Easy Tier design

Easy Tier enables dynamic data relocation, that is, the capability to move volumes or extents across storage system resources transparently to the host. Dynamic data relocation is the core capability that is introduced by Easy Tier. In Automatic Mode, Easy Tier allows automatic data relocation within managed extent pools to optimize system performance at the extent level. Easy Tier Manual Mode provides dynamic volume relocation to relocate data at the volume level and manually manage volume placement in the appropriate extent pool or storage tier.

The DS8000 can monitor the back-end I/O usage intensity of all allocated extents or track groups. Easy Tier data policies and migration capabilities can be controlled at the extent pool or volume level. The Workload categorization and migration reports are part of the DS8000 DSGUI.

Easy Tier I/O workload monitoring only collects I/O statistics from the physical back end, which means the back-end I/O activity on the rank level. The system monitors the stage and destage activity of each extent group that is allocated to a logical volume in the extent pool and calculates a temperature metric for each extent group (also referred to as a *heat map*). A read cache hit that is satisfied by data that is already in cache is not monitored or recorded by the system for the Easy Tier metrics, and therefore is not reflected in the Easy Tier heat map.

### 2.8.1 Flash, enterprise, and nearline storage tiers

Tiered storage is an approach that uses different types of storage throughout the storage infrastructure or even on a selected storage system, such as the DS8000. It is a mix of higher performing or higher-cost storage with lower performing or lower-cost storage, which places data based on performance needs.

Correctly balancing these tiers leads to the minimal cost, best performance solution. Easy Tier Automatic Mode provides monitoring and dynamic data relocation capabilities to automatically optimize I/O performance across different tiers of storage by placing subvolume data onto the correct tier.

Easy Tier can manage seven drive types:

- High-performance flash drives, *Flash Tier 0*
- Solid-state flash drives, *Flash Tier 0* (legacy models)
- High-capacity flash drives, *Flash Tier 1*
- High-capacity flash drives, *Flash Tier 2*
- Enterprise drives 15K (legacy models)
- Enterprise drives 10K (legacy models)
- Nearline drives 7.2K (legacy models)

The performance characteristics of flash, enterprise, and nearline storage tiers differ considerably. Flash performance, including high-performance flash drives, high-capacity flash drives, and SSDs, is superior compared to hard disk drive (HDD) performance because HDDs are limited by mechanical latencies (rotational latency, seek time, and transfer time) with access times around 4 - 8 ms. Flash provides access times far below 1 ms. Furthermore, enterprise drives perform better than nearline drives, especially for small-block random I/O operations.

Flash gives the most performance improvement over HDDs when they are handling small-block random read I/O requests. Similarly, small block random write I/O requests can also be handled much better by flash compared to HDDs. Flash performance is best for small block random read I/O requests, and typically decreases with increasing I/O request block sizes (32 KB or larger) or increasing write ratios.

Although sequential large-block I/O still performs better on flash, Easy Tier does not consider moving large block I/O requests to flash. In terms of price-performance, it is more beneficial to move small block random I/O requests to flash rather than large-block sequential I/O requests.

Comparing enterprise and nearline drives, conclusions are almost equivalent, even if the performance gain between those two tiers is much smaller. Random I/Os perform about twice as fast on enterprise drives, and sequential I/Os perform about the same on enterprise as on nearline drives.

The introduction of high-capacity flash drive types enables the option of locating large-block sequential I/O on this tier, or to replace HDDs completely with a cost-effective all-flash configuration.

For more information about flash on DS8000, see 3.2, “Physical configuration planning and flash considerations” on page 46.

## 2.8.2 Easy Tier performance metrics

To determine the best storage tier for an extent in an extent pool, Easy Tier uses performance metrics that are collected by the Easy Tier monitoring process. For every extent, the following I/O performance data is recorded by the Easy Tier storage system performance monitor:

- Whether the I/O is a read or a write.
- Whether the I/O is a small block (random) I/O or large block (sequential) I/O.
- Cumulative latency of the I/O, which is the time spent to complete the I/O operations in the rank.

  In a Metro Mirror environment, another time delay exists because of the required data transfer of a write I/O to the secondary site. This other latency or service time is not included in the foregoing performance data, because this I/O activity is not an activity that is occurring at the back end on the rank level.
- I/O rate.
- Amount of data that is transferred.

Easy Tier determines whether an extent or extent group is hot, warm, or cold based on these performance metrics about its automated cross-tier performance management capability in Automatic Mode. Easy Tier calculates the temperature or *heat* for each extent group, which is a metric for the extent's hotness regarding its back-end I/O activity. The heat of an extent group is crucial when determining the next extent migration plan and relocating extents to the appropriate storage tier in managed extent pools.

## 2.8.3 Easy Tier cross-tier and intra-tier performance management

Easy Tier Automatic Mode provides automatic inter-tier or cross-tier performance management and automatic intra-tier performance management in multitier (hybrid) or single-tier (homogeneous) extent pools.

### Easy Tier cross-tier or inter-tier performance management

In configurations with managed multi-tier extent pools that contain ranks from more than one storage tier, Easy Tier Automatic Mode creates a migration plan to migrate the hot extents from a lower performance to a higher performance tier. It gives higher priority to the extents with the highest temperatures.

The system also monitors the workload of each rank in an extent pool to determine whether a rank's workload is causing excessive I/O latency, and to avoid rank overload or resource saturation situations. All performance information is used to generate an extent relocation or migration plan at least once every 24 hours. The goal of the migration plan is to place each extent in the pool onto the most appropriate storage tier available.

The migration plan considers the workload statistics that accumulated for the extent over time. This cross-tier or inter-tier extent migration is triggered by the extent hotness in multitier extent pools.

If an extent group is promoted from one tier to another, then the idle extents in the extent group will not be promoted. For demotion of cold extents, the idle extents for all extent groups are demoted first before entire extent groups are considered. Idle extents are the first candidates for demotion selection as part of a swap.

**Three-tier migration:** In a three-tier extent pool configuration, the cross-tier extent migration only occurs between adjacent tiers. There are no direct migrations from the lowest tier to the highest tier that bypass the middle tier.

However, clients using Nearline drives have occasionally seen problems where a significant amount of data on Nearline suddenly becomes active. In early implementations, Easy Tier did not promote this data until the next analysis interval. Easy Tier now supports a *warm promote*. Warm promote functions in a similar way to warm demote. If the 5-minute average performance shows that a Nearline rank is overloaded, Easy Tier immediately starts to promote extents until the condition is relieved.

### Easy Tier intra-tier performance management

Easy Tier in Automatic Mode performs intra-tier extent migrations by monitoring the overall I/O workload for each rank and creating a migration plan to rebalance the I/O workload evenly across all ranks within a tier.

This feature, which is called *auto-rebalance,* constantly rebalances the I/O workload across all ranks within a storage tier in managed extent pools. This feature is not restricted to multitier or hybrid pools. This feature is also available for single-tier or homogeneous extent pools when they are managed by Easy Tier in Automatic Mode.

The migration is staged so that the impact on host I/Os is minimal (comparable to IBM FlashCopy® background activity).

### Easy Tier migration plan

After an extent migration plan is generated, the storage system migrates a limited number of extents every 5 minutes until all extents in the plan are migrated, or until the next plan is generated. This process continues repeatedly on an ongoing basis. Given a consistent workload, all extents are eventually migrated to their optimum storage type. For a dynamic workload, where the temperature metrics of the extents fluctuate over time, each new extent migration plan can adapt dynamically to workload variations, and can relocate extents appropriately.

The extent migration plan algorithms consider the cost of moving the extent when deciding whether an extent can be relocated. If the benefit of moving the extent does not exceed the cost of the relocation, the extent is not included in the migration plan.

When no more free extents are available on a higher tier but hot extents are still on a lower tier, Easy Tier evaluates whether the hottest extent on the lower tier benefits more when placed on the higher tier compared to the extent on the higher tier with the coldest temperature. If that is the case, the locations of those two extents are swapped.

## 2.8.4 Migration plan creation

A migration plan is only applicable when the Easy Tier monitoring task can sample enough performance data for the volumes. Migration plans are created internally on the DS8000 based on the Easy Tier monitoring data.

Easy Tier Automatic Mode creates intra-tier migration plans (auto-rebalance plans) and cross-tier migration plans. The two kinds of migration plans have two different decision windows during which Easy Tier monitors and collects I/O statistics to identify the extents to be moved. The decision window for the auto-rebalance migration plan is 6 hours, and the decision window for the cross-tier migration plan is approximately 24 hours.

At least once every 24 hours, a cross-tier migration plan is created. However, the start time varies, so you are not always starting at the same time each day for a new data collection. The new plan is immediately carried out. The migration plan creation is independent of when you offload the monitor data from the DS8000.

Easy Tier creates one migration queue per extent pool, allowing each pool to be optimized independently. This situation leads to even better migration plan management, because each extent temperature is computed relatively to the other extents within the same pool. During migration plan execution, extent relocation occurs in a round-robin fashion at the pool and tier level. This action ensures that progress is made on each pool and tier.

Figure 2-9 on page 32 shows the basic timeline for creating and running an inter-tier migration plan. It indicates that monitoring data is constantly and continuously collected. This volume is not a growing data volume; it is an update of the collected performance counters. The size of the data that is created by the Easy Tier monitoring task is related to the number of volumes that are monitored.

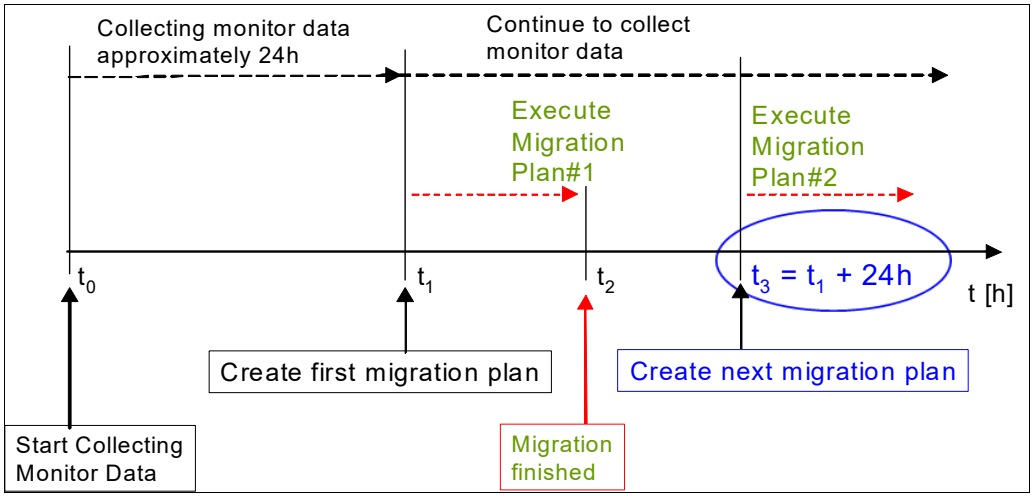



*Figure 2-9 Create inter-tier migration plan cycle*

**Migration plan cycle:** $t_3$ is not exactly $t_1$ + 24h because a new migration plan is created at least every 24 hours, but not necessarily exactly at 24-hour intervals.

The first inter-tier migration plan for Easy Tier cross-tier performance management is created after monitoring data is collected for approximately 24 hours. The first intra-tier migration plan for the Easy Tier auto-rebalance feature is created after monitor data is collected for approximately 6 hours.

After a 24-hour period with monitoring enabled, a good picture is available of what the I/O workload on the physical storage back end looks like, and the first cross-tier migration plan starts to run.

Remember that a cross-tier migration plan starts moving extents across ranks from different tiers within managed extent pools when space is available on a higher performance tier and extents on the lower performance tier are considered hot. Potentially, the extents on the lower performance tier will benefit from being relocated onto a higher tier.

After another 24-hour decision window, a new migration plan is created and starts to run. This new plan is based on performance data that was collected for the past two days.

The relevance of older monitor data is exponentially decreasing over time. Historical data is considered when available, but with diminishing relevance (the system uses a seven-day, long-term decision window to factor in older I/O statistics). The most relevant data is the data that is collected in the last decision window relative to the time when the new migration plan is created.

**Decision window:** The Easy Tier Automatic Mode algorithm uses a short-term and a long-term sliding decision window to determine the intra-tier extent migration plan. Statistics in the short-term window have a higher weight relative to those statistics in the long-term window.

Statistics older than the long-term limit are automatically removed from the DS8000. The short-term window is set to 24 hours, and the long-term window is set to seven days. Those parameters cannot be set by using the DS CLI or DS GUI commands.

When the time to run a migration plan is longer than the decision window, the current migration plan run is stopped and a newly created migration plan is started.

Because the cross-tier and intra-tier or auto-rebalance plans are independent of each other, the same extent might be eligible for migration in both plans. In this case, the cross-tier migration plan always has the higher priority and prevails.

### 2.8.5 Easy Tier migration types

When Easy Tier determines the correct storage media for an extent based on the extent heat and resource usage, it uses the following extent migration types between storage tiers or within a storage tier, as shown in Figure 2-10:

- Promote and Swap: For moving hot data to higher performing tiers
- Warm Demote: Prevents performance overload of a tier by demoting warm extents to the lower tier and being triggered when bandwidth or IOPS thresholds are exceeded
- Cold Demote on lower tiers: Where coldest data is identified and moved to the lower tier
- Expanded Cold Demote for HDD tiers: To demote part of the sequential workload to use the bandwidth in the nearline tier better
- Auto-Rebalance: Redistributes the extents within a tier to balance usage across all ranks of the same tier for maximum performance
- Warm promote: If drives become suddenly hot, start moving extents immediately to a higher tier

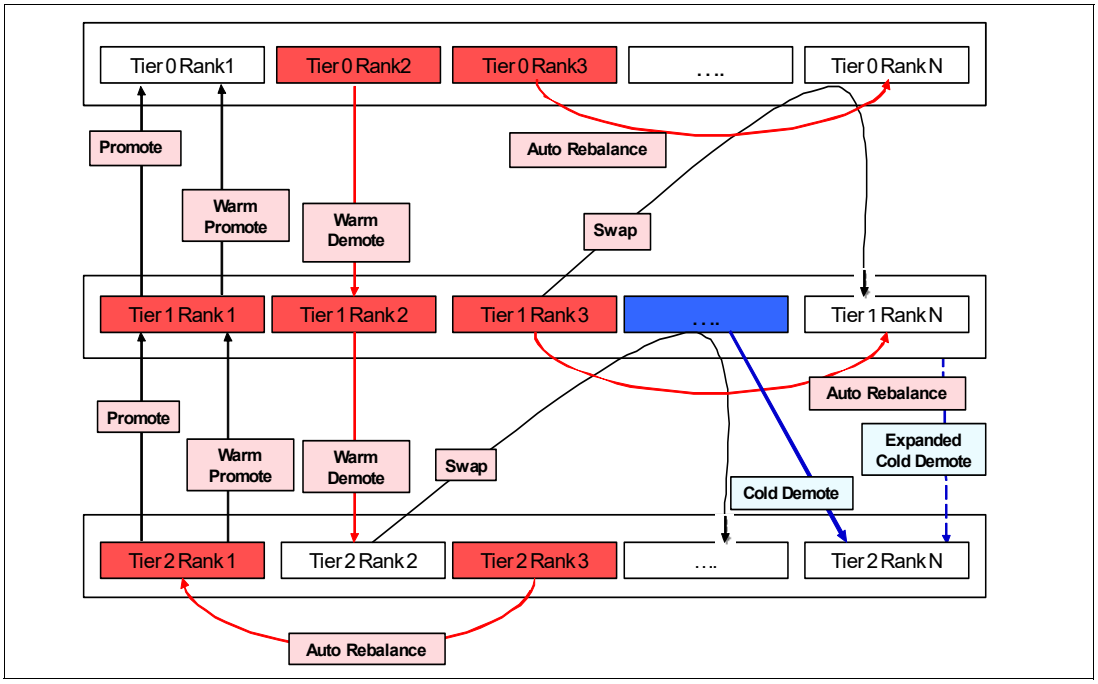



*Figure 2-10 Easy Tier migration types*

**Important:** Migrations can happen only between adjacent tiers. For example, no promotion or demotion occurs from tier 0 to tier 2 or tier 2 to tier 0 in a three-tier extent pool.

For details about tiers and migrations between them, see Figure 2-13 on page 38.

## 2.8.6 Promoting and swapping

The following migration types move hot data to higher performing tiers:

- Promote: Moves the hotter extents to the ranks of the target tier with available capacity.
- Swap: Easy Tier can swap extents between tiers to move hotter data to a higher performing tier when the other it is full.

## 2.8.7 Auto-rebalance

*Auto-rebalance* or *automatic performance rebalance*, which is also known as *intra-tier rebalancing*, is a capability of Easy Tier that automatically rebalances the workload across all ranks of a storage tier within a managed extent pool. Auto-rebalance migrates extents across ranks within a storage tier to achieve a balanced workload distribution across the ranks and to avoid hotspots. By doing so, auto-rebalance reduces the performance skew within a storage tier and provides the best available I/O performance from each tier.

Furthermore, auto-rebalance also automatically populates new ranks that were added to the pool when rebalancing the workload within a tier. Auto-rebalance can be enabled for hybrid and homogeneous extent pools.

Figure 2-11 shows two examples of scenarios where auto-rebalance can be involved.

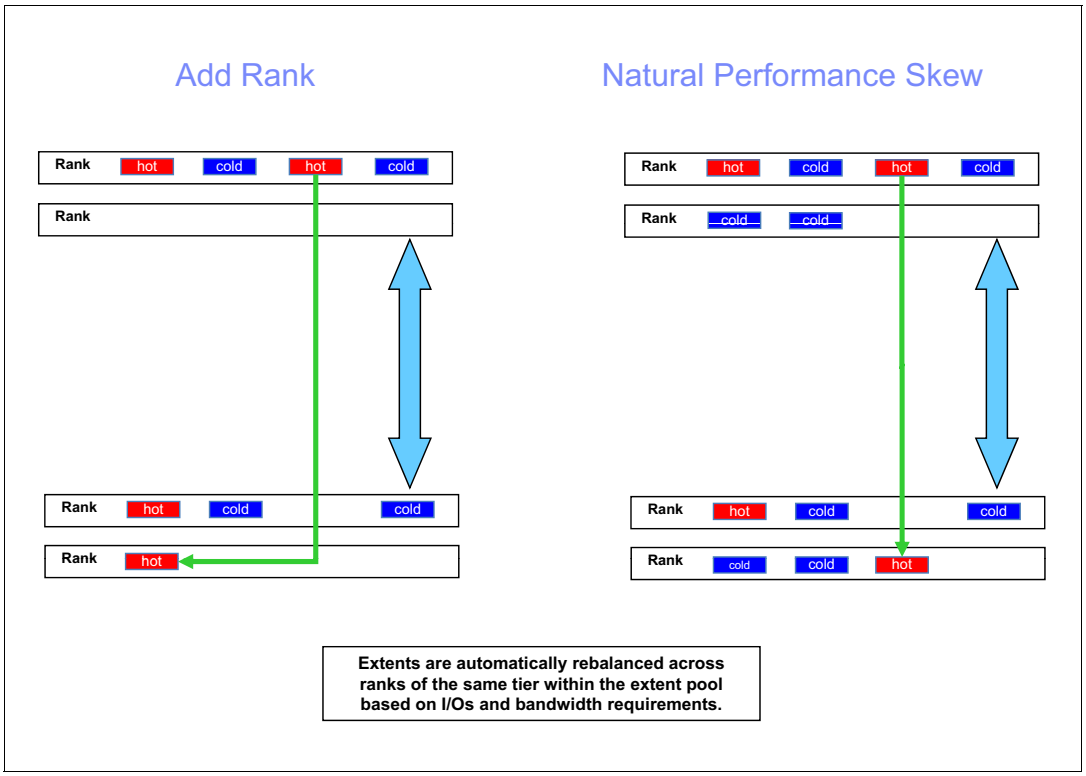



*Figure 2-11 Auto-rebalance scenarios*

The auto-rebalance algorithm maximizes storage performance on both the IOPS and bandwidth levels within the same storage tier. The algorithm uses a threshold at the rank level to determine whether an extent might benefit from a move to another less-used rank within the same tier in the extent pool. Auto-rebalance is triggered every 6 hours, if needed. The rebalance also starts immediately if you add ranks to an existing pool.

Auto-rebalance can occur within a managed extent pool on every tier that has at least two ranks. Device-specific performance variations within a storage tier (micro-tiering) are considered for determining the rank usage based on device-specific performance profiles, considering drive speed, RAID level, and drive size, with workload characteristics.

No rebalancing is done if the skew is insignificant or if migration will cause *toggling*, which means that the migration triggers further rebalancing efforts on the target array.

**Important:**

- Auto-rebalance is based on IOPS rank usage to reduce skew and hotspots within a tier by redistributing extents across ranks of the same tier. Auto-rebalance is not capacity-based, in the sense that it does not try to balance the extents and, therefore, the capacity of the volumes, evenly across the ranks within a tier. This balancing is the way that manual volume rebalance and storage pool striping attempt. Extents are only relocated onto another rank within the same tier for performance reasons.
- Auto-rebalance is not active in non-managed extent pools. Extents within non-managed single-tier extent pools can be rebalanced based on a balanced capacity distribution by using the manual volume rebalance feature, as described in “Dynamic volume relocation” on page 22.

## 2.8.8 Warm and cold demote

In addition to the promote and swap extent migration types that move hot extents to the higher performing tier, the demotion of extents is an essential part of the Easy Tier design.

### Warm demote

Higher-performance tiers can become overloaded in hybrid extent pools, and potentially degrade overall pool performance. To avoid this situation, Easy Tier Automatic Mode monitors the performance of the ranks and triggers the movement of selected extents from the higher-performance tier to the lower-performance tier. This action is based on predefined bandwidth or IOPS overload thresholds.

This automatic operation is rank-based, and the target rank is randomly selected from the lower tier. Warm demote has the highest priority to relieve overloaded ranks quickly. Easy Tier continuously ensures that the higher-performance tier does not suffer from saturation or overload conditions that might affect the overall performance in the extent pool.

A potential issue when Easy Tier attempts a warm demote is that the lower tier might be full (no space available to move an extent, as shown in Figure 2-12).

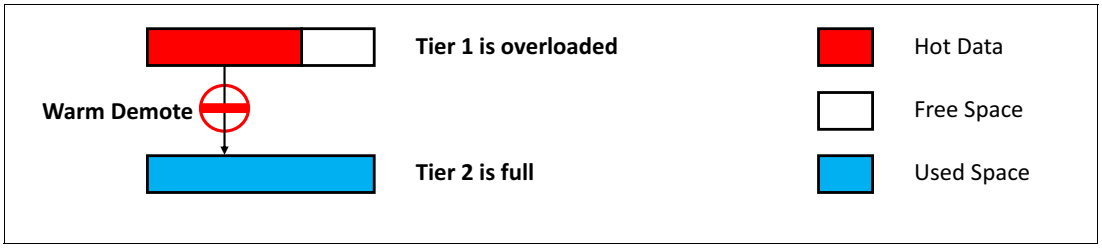



*Figure 2-12 Warm demote space reservation*

To alleviate this situation, Easy Tier uses an algorithm, which is known as the *Space Reservation Manager* (SRM), that reserves space on each tier for Easy Tier activities. The SRM is able to move data to an upper tier or a lower tier based on the space requirement and available space on each pool tier.

In addition, rank performance protection enhancement was introduced on top of the existing protection (rank performance threshold based on Easy Tier performance profile). The enhancement consists of a dynamic measure of rank utilization to detect in real time whether the rank is performance overdriven, and to prevent any new workload from being moved to it. A warm demote is not allowed if the target tier ranks are overdriven.

**Note:** With the warm demote improvement, Easy Tier stops demoting workload to a lower tier if it detects that the lower tier is already performance overloaded. Also, Easy Tier issues a warning when it detects that all tiers are performance overloaded or some ranks are continuously performance overloaded, because that means that Easy Tier is unable to move data around.

### Cold demote

Easy Tier Automatic Mode automatically locates and demotes inactive (or cold) extents that are on a higher performance tier to its adjacent lower-cost tier. In that way, Easy Tier Automatic Mode automatically frees extents on the higher storage tier before the extents on the lower tier become hot, and then helps the system to be more responsive to new hot data.

Cold demote recognizes and demotes cold extents to an appropriate lower-cost tier. Cold extents are demoted in a storage pool to a lower tier if that storage pool is not idle to continuously optimize storage economics.

**Cold demote occurs between drives:** Cold demote occurs only between two adjacent tiers, that is, between flash (Tier 0) and enterprise (Tier 1) drives, and enterprise (Tier 1) and nearline (Tier 2) drives. See Figure 2-10 on page 34.

If all extents in a hybrid storage pool become simultaneously inactive for 15.5 - 24 hours because of a planned or unplanned outage, Easy Tier automatically disables cold demote for this pool. This action helps you schedule extended outages (or experience an outage) without performing cold extent relocation.

**Three-tier extent pool configuration demotion.** In a three-tier extent pool configuration, the extent demotion occurs only between two adjacent tiers.

## Expanded cold demote

Expanded cold demote is an Easy Tier feature that demotes part of the sequential workload to the nearline tier (Tier 2) to use the bandwidth in the nearline tier better. It selects appropriate extents with moderate bandwidth requirements from the enterprise tier for demotion to the nearline tier to use its bandwidth capabilities, and to spread the bandwidth demand across both tiers instead of using only the enterprise tier. Depending on the drive types installed, nearline might not be Tier 2.

### Warm promote

Easy Tier now supports a *warm promote*. Drives in a given tier can suddenly become very active. Instead of waiting for the next migration plan, Easy Tier can react immediately. Warm promote acts in a similar way to warm demote. If the five minute average performance shows that a rank is overloaded, Easy Tier immediately starts to promote extents to the next higher tier until the condition is relieved.

In summary and also considering the legacy hybrid models like DS8880, Easy Tier supports seven drive types, using up to three drive classes, high-performance flash, and solid-state drives (SSD), (both Flash Tier 0), high-capacity flash drives (Flash Tiers 1 and 2), Enterprise drives 15K or 10K, and Nearline drives. With the introduction of the high-capacity flash drives, the drive class in a specific tier can vary depending on the types of drives that are managed by Easy Tier.
To understand the tiers and the migration from one tier to another, see Figure 2-13.

| Source Tier \ Target Tier | *High-Performance Flash* | *High-Capacity Flash* | *ENT HDD* | *NL HDD* |
|---|---|---|---|---|
| *High-Performance Flash* | RB_MOVE<br>RB_SWAP | Warm Demote | Warm Demote<br>Cold Demote*<br>Expanded Cold Demote* | Warm Demote<br>Cold Demote*<br>Expanded Cold Demote* |
| *High-Capacity Flash* | Promote<br>Sequential Promote<br>Swap<br>Warm Promote | RB_MOVE<br>RB_SWAP | Warm Demote<br>Cold Demote*<br>Expanded Cold Demote* | Warm Demote<br>Cold Demote*<br>Expanded Cold Demote* |
| *ENT HDD* | Promote<br>Swap<br>Warm Promote | Promote<br>Swap<br>Warm Promote | RB_MOVE<br>RB_SWAP | Warm Demote<br>Cold Demote<br>Expanded Cold Demote |
| *NL HDD* | Promote<br>Sequential Promote<br>Swap<br>Warm Promote | Promote<br>Sequential Promote<br>Swap<br>Warm Promote | Promote<br>Sequential Promote<br>Swap<br>Warm Promote | RB_MOVE<br>RB_SWAP |

*Among Same Tier (Rank Rebalance)* — *From Higher to Lower T: * Enabled when SSD Home Tier* — *From Lower Tier to Higher*

*Figure 2-13 Easy Tier Migration among tiers*

## 2.8.9 Extent allocation in hybrid and managed extent pools

As described in 2.6.1, “Easy Tier Automatic Mode” on page 18, the Easy Tier Automatic Mode scope can be modified to manage either all extent pools (including single-tier, homogeneous extent pools), or only tiered, multitier extent pools.

When you create a volume in a managed extent pool, that is, an extent pool that is managed by Easy Tier Automatic Mode, the EAM of the volume always becomes managed, no matter which EAM is specified at volume creation. The volume is under the control of Easy Tier, which moves extents to the most appropriate storage tier and rank in the pool based on performance aspects.

Any specified EAM, such as rotate extents or rotate volumes, is ignored. In managed extent pools, an initial EAM that is similar to rotate extents for new volumes is used. The same situation applies whether an existing volume is manually moved to a managed extent pool by using volume migration or dynamic volume relocation.

In hybrid or multi-tier extent pools (whether currently managed or non-managed by Easy Tier), initial volume creation always starts on the ranks of the enterprise tier first, as shown in Example 2-1 on page 39.

In Example 2-1, all extents of the newly created volume are in enterprise tier rank R26. (In a pool, which contains more enterprise arrays, newly created volumes are evenly distributed across enterprise tier ranks only.)

*Example 2-1 Volume creation and initial extent allocation in hybrid extent pools*

```
dscli> lsextpool -l P4
Name         ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols numranks encryptgrp numtiers etmanaged
=============================================================================================================================================
ITSO_EasyTier P4 fb           0  below            10833        50     10833       0     149       4         -       3 yes

dscli> lsrank -l -extpool P4
R27     0 Normal Normal    A25          5 P4        ITSO_EasyTier fb      2122       613          - MA26

dscli> lsarray -l a1 a4 a22 a25
Array State    Data   RAIDtype    arsite Rank DA Pair DDMcap (10^9B) diskclass encrypt
=========================================================================================
A1    Assigned Normal 5 (6+P+S)   S17    R17  2                400.0 SSD       supported
A4    Assigned Normal 6 (5+P+Q+S) S15    R4   2               3000.0 NL        supported
A22   Assigned Normal 5 (6+P+S)   S6     R26  0                900.0 ENT       supported
A25   Assigned Normal 5 (6+P)     S26    R27  10               400.0 Flash     supported

dscli> mkfbvol -cap 1000 -extpool p4 -name ITSO_fbvol a00a
CMUC00025I mkfbvol: FB volume a00a successfully created.

dscli> showfbvol -rank a00a
Name             ITSO_fbvol
ID               A00A
accstate         Online
datastate        Normal
configstate      Normal
deviceMTM        2107-900
datatype         FB 512
addrgrp          A
extpool          P4
exts             1000
captype          DS
cap (2^30B)      1000.0
cap (10^9B)      -
cap (blocks)     2097152000
volgrp           -
ranks            1
dbexts           0
sam              Standard
repcapalloc      -
eam              managed
reqcap (blocks)  2097152000
realextents      1000
virtualextents   0
migrating        0
perfgrp          PG0
migratingfrom    -
resgrp           RG0
tierassignstatus -
tierassignerror  -
tierassignorder  -
tierassigntarget -
```

```
%tierassigned    0
etmonpauseremain -
etmonitorreset   unknown
==============Rank extents==============
rank extents
============
R26     1000

dscli> showfbvol -tier a00a
[...]
===========Tier Distribution============
Tier %allocated
===============
ENT         100
```

Newer code releases have two allocation policy options: High Utilization and High Performance. New volumes are allocated to the *home tier*. When the allocation policy High Utilization is in effect, the home tier is some middle tier, and when High Performance is in effect, the home tier is the highest available flash tier. The extents of a new volume are distributed in a *rotate extents* or *storage pool striping* fashion across all available ranks in this home tier in the extent pool if sufficient capacity is available. Only when all capacity on the home tier in an extent pool is used does volume creation continue on the ranks of another tier.

For High Utilization, the allocation order is for all-flash pools: Flash Tier 1 → Flash Tier 2 → Flash Tier 0, and for hybrid pools: Enterprise → Flash Tier 1 → Flash Tier 2 → Nearline → Flash Tier 0.

For High Performance, the allocation order is for all-flash pools: Flash Tier 0 → Flash Tier 1 → Flash Tier 2, and for hybrid pools: Flash Tier 0 → Flash Tier 1 → Flash Tier 2 → Enterprise → Nearline.

After all capacity on the home tier is exhausted, volume creation continues allocating extents on the next tier, according to the allocation order.

The initial extent allocation in non-managed hybrid pools differs from the extent allocation in single-tier extent pools with rotate extents (the extents of a volume are not evenly distributed across all ranks in the pool because of the different treatment of the different storage tiers). However, the attribute for the EAM of the volume is shown as `rotate extents` if the pool is not under Easy Tier Automatic Mode control. After the pool is managed by Easy Tier Automatic Mode, the EAM becomes `managed`.

**Tip:** A control is available to change the allocation policy. See 2.7, “Easy Tier allocation controls” on page 27 for information about how to change the policy and prefer Tier 0 (high-performance flash tier), for example.

In managed homogeneous extent pools with only a single storage tier, the initial extent allocation for a new volume is the same as with rotate extents or storage pool striping.

For a volume, the appropriate DS CLI commands, **`showfbvol`** or **`showckdvol`**, when used with the **`-rank`** option, allow the user to list the number of allocated extents of a volume on each associated rank in the extent pool, as shown in Example 2-1 on page 39. The commands also show the EAM attribute of the requested volume.

When the Easy Tier Automatic Mode for all extent pools (hybrid and homogeneous extent pools) is enabled, all volumes immediately become managed by Easy Tier. The EAM attribute of all volumes on a DS8000 is changed to `managed`. After it is set to `managed`, the EAM attribute

setting for the volume is permanent and all previous volume EAM attribute information, such as rotate extents or rotate volumes, is lost.

### 2.8.10 Easy Tier considerations

Easy Tier is easy to deploy. With the Easy Tier mode set to `Enable`, no other action is required to activate the Easy Tier Automatic Mode in all hybrid, multitier extent pools to benefit from automated cross-tier and intra-tier storage performance and storage economics management in these pools.

With this setting, Easy Tier takes control of all extent pools and all regular and thin-provisioned volumes (ESE), and automatically optimizes storage performance and storage economics in each pool. It also minimizes skew and avoids hotspots within each tier (auto-rebalance), including single-tier extent pools.

A sufficient monitoring time must pass for the decision window before Easy Tier actually creates migration plans and starts cross-tier and intra-tier extent migrations.

Typically, extent pools are configured in pairs on the DS8000, one for each storage server (server 0 and server 1 or rank group 0 and rank group 1), to balance the workload evenly between the two storage servers and all flash RAID adapters, or device adapters (DA). Flash ranks can also be installed in pairs and assigned to pairs of hybrid pools that are managed by both storage servers.

One consideration with Easy Tier Automatic Mode is that the DS8000 optimizes the data placement based on the current and most recent workload characteristics. In cases where the usage or I/O workload of a logical volume changes considerably from one day to another to a new workload pattern, the achieved performance benefits that result from the previous optimized data placement might be lost for the new workload.

This situation occurs because the I/O access patterns and characteristics changed. It takes time until Easy Tier adapts to the new workload pattern, and the data placement is optimized again for new workload.

In such cases, Easy Tier Application can be an alternative solution to prepare the data placement for the upcoming workload, if the pattern is predictable. For more information about Easy Tier Application, see *DS8870 Easy Tier Application*, REDP-5014.

## 2.9 Easy Tier control

Easy Tier learning and extent migration behaviors can be controlled at the extent pool and volume levels. Easy Tier learning can be suspended, resumed, and reset at the pool or volume level. Easy Tier extent relocation can be suspended or resumed at the pool level. Additionally, a volume can be excluded from being migrated to the nearline tier.

For examples and illustrations, see Chapter 4, “DS GUI and DS CLI support for IBM DS8000 Easy Tier” on page 63.

### 2.9.1 System-level Easy Tier control

Table 2-3 shows the available Easy Tier modes in the DS GUI. For DS CLI settings, see 4.2, “DS CLI support for Easy Tier” on page 79.

*Table 2-3 Easy Tier mode settings in DS GUI*

| Easy Tier Mode | Monitoring |
|---|---|
| Disable | Do not monitor I/O activity or capacity placement. |
| Enable | Monitor I/O activity for capacity and manage capacity placement in all pools. |
| Tiered pools only | Monitor I/O activity for capacity and manage capacity placement within pools that have arrays with different drive classes. |
| Monitor only | Monitor I/O activity for capacity but do not manage placement. |

**Important:** The system-level control switch always has higher priority than the pool-level or volume-level control switch. If any system-level control switch (Easy Tier monitor or Easy Tier management) is changed, the pool or volume-level control switch will be reset.

### 2.9.2 Pool-level Easy Tier control

The following Easy Tier controls can be set at the pool level:

- Suspend or resume Easy Tier learning for a specified pool:
  - Suspend the updating of the long-term average and maximum value of extent statistics for all volumes in the specified pool. A lease time can be set so that after the lease time expires, the learning resumes automatically.
  - Resume updating the long-term average and maximum value of extent statistics with or without history for all volumes in the specified pool.
- Suspend or resume Easy Tier extent relocation for a specified pool:
  - Suspend the extent migration for all volumes in a specified pool for all types of migration, or choose to allow only the warm demote to happen. The goal is to protect system performance. A lease time can be set so that when the lease time expires, the extent migration resumes automatically.
  - Resume the extent migration for the volumes in the specified pool for all types of migrations.
- Reset Easy Tier learning for a specified pool.

  Reset the learning history for all volumes in the pool.
- Query pool state of Easy Tier control switch:
  - Learning state (suspended or normal).
  - Migration state (suspended or normal).
  - Learning reset (the time stamp when the learning of a specified pool is reset).

## 2.9.3 Volume-level Easy Tier control

The following Easy Tier controls can be set at the volume level:

- Suspend and Resume Easy Tier learning for a specified volume:
  - Suspend updating the long-term average and maximum value of extent statistics for the specified volume. A lease time can be set so that after the lease time expires, the learning resumes automatically.
  - Resume updating the long-term average and maximum value of extent statistics with or without history for a specified volume.
- Reset Easy Tier learning for a specified volume.
- Query volume state of Easy Tier control switch:
  - Learning state (suspending or normal).
  - Learning reset (show the time stamp when the learning of a specified volume is reset).
- Exclude a specified volume from the nearline tier. (Only the standard volume and the ESE volume are supported):
  - Exclude a volume from the nearline tier so that the volume will be distributed only on non-nearline tiers, such as the High Performance flash tier and ENT tier, or high capacity flash tier and enterprise tier.
  - Query the assign state for a specified volume.

Figure 2-14 (diagrams on the left) shows the influence of unexpected behavior for unusual temporary workload processing. The diagrams on the right show the ability to pause the Easy Tier learning during an unusual workload and therefore avoid its impact on performance statistics.

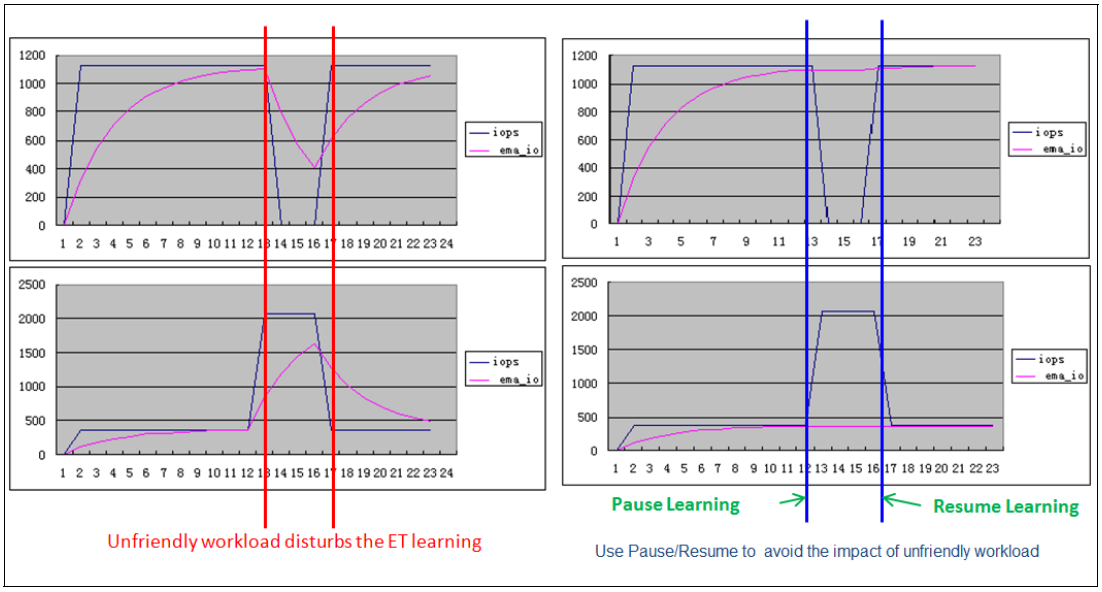



*Figure 2-14 Easy Tier suspend and resume monitoring*

**Note:** Duration is an integer with hour as the unit, and the value definition is [0, 168]. If the value is 0, it means infinite lease. The default is 168 hours (1 week).

- Reset monitoring process

  The ability to reset learning can force Easy Tier to directly and quickly adapt to a new workload.

  The following examples are typical use cases for reset learning:

  – A new or different usage of all volumes within the storage pool, and therefore the original learning data of the volumes is no longer relevant.
  – A change in workload or the deployment of a new application.
  – After a database reorganization.

- Exclude volume from nearline drives

  In certain cases, you might want to exclude volumes to be assigned to a nearline tier to preserve the performance of an important application. The situation is illustrated in Figure 2-15.

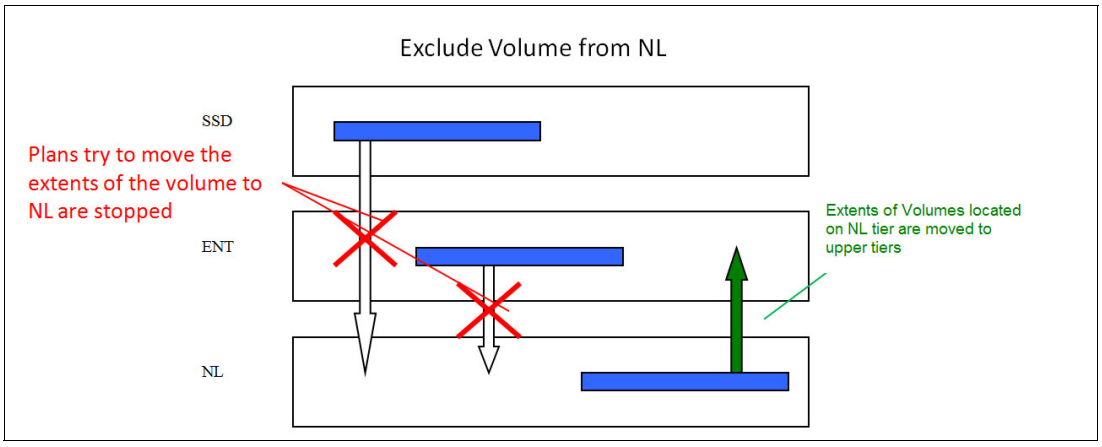



*Figure 2-15 Exclude volume from nearline tier*

This capability is also useful for application initial deployment. Before the actual workload actually starts (pre-production), the volumes that are allocated for the application are idle. However, you want to prevent Easy Tier from demoting the volumes to nearline drives to avoid a performance issue when the application really starts.

With the possibility that volumes are pinned to a specific tier (see 3.4.3, “Working with pinned volumes” on page 58), the new capability of excluding volumes from the nearline tier now allows you, within a single pool that includes three tiers of storage, to choose the following volume behavior by policy:

– On all three tiers (flash, enterprise, and nearline), depending on workload.
– On tier 0/tier 1 only to avoid the potential for lower nearline performance.
– Assigned (pinned) to tier 0 to guarantee flash performance.

3

# Planning for DS8000 Easy Tier

This chapter provides planning information for IBM System Storage Easy Tier. It summarizes the flash technology configuration rules for IBM System Storage DS8000 and offers guidance for implementing Easy Tier Automatic Mode. It also briefly describes available planning tools.

This chapter includes the following topics:

- Prerequisites
- Physical configuration planning and flash considerations
- Logical configuration planning
- Implementation considerations for multi-tier extent pools
- Copy Services considerations
- Workload planning

## 3.1 Prerequisites

For the DS8900F, the Easy Tier function is part of the Basic Function license.

For automatic mode to be active, the following conditions must be met:

- Easy Tier automatic mode monitoring is set to either `All` or `Auto mode`.
- To manage pools, the Auto Mode Volumes must be set to either `Tiered Pools` or `All Pools`.

For more information about how to activate a licensed feature, see *IBM DS8900F Architecture and Implementation*, SG24-8456.

## 3.2 Physical configuration planning and flash considerations

Drives are managed in up to three tiers - usually Flash Tier 0, 1 and 2.

Consider this information about flash drives:

- The *high-performance flash drives* offer the highest storage performance option. The DS8000 high-performance flash enclosure where these drives are installed provides significant performance improvements over prior-generation flash drives (SSDs), and are the best choice for I/O-intensive workloads.
- The *high-capacity flash drives* are a higher-capacity, lower-performing flash drive option in the high-performance flash enclosure. The high-capacity flash drives are treated as a lower performance tier than high-performance flash drives or SSDs, but as a higher tier than HDDs. Because two types of high-capacity flash exist, additional tiering can place between these when mixed. Given the large capacity of some of these in combination with their cost-effective pricing, they can eliminate a need for spinning drives completely, for instance when combined with some portion of high-performance flash.
  The high-capacity flash drives are capacity-optimized, with the 3.84-TB type usually offering an excellent balance between performance and cost. The Flash Tier 2 drives, for instance the 7.68 TB type and larger, provide the highest flash capacity at the lowest cost.

For more information about the drive technologies, see *IBM DS8900F Architecture and Implementation*, SG24-8456.
For more information about the High-Performance Flash Enclosure, see *IBM DS8000 High-Performance Flash Enclosure Gen2*, REDP-5422.

### 3.2.1 Flash technology features and benefits

Flash technology offers the following features and benefits:

- Hard disk drive technology has improved over the years, providing higher bandwidths and higher storage densities at a lower price. However, because of the hard disk drive's moving and spinning mechanical parts, inherent physical limitations exist to the response times (rotational latency, seek time, and transfer time) that a disk drive can achieve.

  Flash drives, compared to enterprise and nearline drives, have low response times with low latencies. With the correct planning, you can benefit from flash technology features to achieve the best performance among different systems and workloads.
- Flash drives give the most performance improvement over HDDs when handling small block random read I/O requests. Small block random write I/O requests also can be

handled much better by flash compared to HDDs. Flash performance is best for small block random read I/O requests and typically decreases with increasing I/O request block sizes (32 KB or larger) or increasing write ratios.

- Flash provides faster data access, higher throughput, and better durability and reliability. Flash drives better tolerate extreme shocks, higher altitude, vibration, and temperature extremes.
- Flash also requires less power consumption. Because no power for the motor is required to spin up the magnetic platters and to move the heads, flash technology uses less energy than an HDD.

## 3.2.2 Flash performance guidelines

The integration of the high-performance flash enclosure provides a high standard of flash performance with the DS8000. The flash enclosure in the DS8000 provides up to three and a half times the performance of traditional SSDs.

High-Performance Flash arrays (or ranks) can coexist within the same extent pool as SSDs. Both are treated by Easy Tier as the highest tier 0. However, Easy Tier recognizes that flash card arrays have higher IOPS and throughput capabilities. The Easy Tier intra-tiering (micro-tiering) functionality accounts for that, and “hotter” I/O-intensive parts of the volumes are moved to the flash cards within that highest tier.

High-Capacity Flash arrays can also coexist with both high-performance flash arrays and SSDs. However, the high-capacity flash arrays are treated as a lower-performance tier than both high-performance flash arrays or SSD arrays. If no high-performance flash arrays or SSD arrays are present, the high-capacity flash arrays are treated as the highest performance tier.

From a performance perspective, flash drives are preferred for your DS8000 drives. Flash drives or solid-state drives (SSDs) have no moving parts (no spinning platters and no actuator arm). They also have lower energy consumption. The performance advantages are the fast seek time and average access time.

SSDs are targeted at applications with heavy IOPS, bad cache hit rates, and random access workload, which necessitates fast response times. Database applications, with their random and intensive I/O workloads, are prime candidates for deployment on flash.

### Flash for metadata

The DS8000 needs to maintain global metadata and metadata for each volume. Fast access to metadata is essential for best performance. This is particularly important with thin-provisioned volumes, because metadata is accessed every time new extent allocations take place. The DS8000 places metadata in the highest tier. Therefore, have some flash storage in each extent pool.

The DS8000 is only allowed to use a fraction of flash capacity for metadata. However, if you have about 10% of your extent pool capacity on flash drives, you can be sure that all metadata is in flash storage. Three copies of the Global Metadata is stored on three separate arrays, for resiliency. If you have at least three flash arrays in your DS8000, all copies of the Global Metadata will be on flash.

### 3.2.3 Data placement guidelines, limitations, and considerations

After you determine the disk system throughput, disk space, and number of disks that are required by your hosts and applications, you must determine the data placement.

As is common for data placement, and to optimize DS8000 resource usage, use the following guidelines:

- Equally spread the workload across the DS8000 internal storage servers. Spreading the volumes equally on rank groups 0 and 1 balances the load across the DS8000. From the DS GUI, the system, by default, uses volume pairs creation.
- Balance the ranks and extent pools between the two DS8000 storage servers to support the corresponding workloads on them. From the DS GUI, the system, by default, uses pool pairs creation.
- Use as many drives as possible. Avoid idle drives, even if all storage capacity is not initially to be used.
- Distribute capacity and workload across device adapter (DA) pairs.
- Use multi-rank extent pools.
- Use some flash capacity in each extent pool.
- Stripe your logical volume across several ranks (the default for multi-rank extent pools).
- Consider placing specific database objects (such as logs) on separate ranks.
- For an application, use volumes from even and odd-numbered extent pools (even-numbered pools are managed by server 0, and odd numbers are managed by server 1).
- For large, performance-sensitive applications, consider the use of two dedicated extent pools (one managed by server 0, the other managed by server 1).
- Consider mixed extent pools with multiple tiers with flash technology as the highest tier, managed by Easy Tier.

For configuration planning purposes, be aware of the following flash configuration limitations on DS8000:

- Since DS8000 Release 8.3, RAID 6 is now the preferred RAID protection and default setting for all drive types.
- The 1.8-inch (flash drives), 2.5-inch (SFF), and 3.5-inch (LFF) drives on DS8000 systems cannot be intermixed in the same storage enclosure pair. They use different storage enclosure technology.

Consider the following information for the placement of flash drives in DS8000 systems:

- SSD-flash drives are installed in default locations by manufacturing, starting with the first storage enclosure pair on each DA pair. This setup spreads the SSD drives over as many DA pairs as possible to achieve an optimal price-to-performance ratio.
- Any one high-performance flash enclosure is considered one DA pair.

For more information about configurations and drive placement, see *IBM DS8900F Architecture and Implementation*, SG24-8456.

## 3.3 Logical configuration planning

Easy Tier is beneficial in multi-tier pools (hybrid pools) and single-tier pools (homogeneous pools). It provides automated cross-tier and intra-tier storage performance management. You can benefit from Easy Tier even if you do not have several storage tiers in one system. However, to get the most benefit from Easy Tier in Automatic Mode, hybrid pools with different storage tiers must be created.

Various flash and HDD combinations are available on DS8000 for use with Easy Tier. You can, for example, use any of the following combinations:

- High-Performance Flash: High-performance flash drives, and flash drives (SSDs)
- High-Capacity Flash: High-capacity flash drives, which are available in two drive classes
- Enterprise: 15K or 10K RPM 2.5-in. HDD
- Nearline: 7.2K RPM 3.5-in. HDD

Figure 1-3 on page 7 and Figure 1-4 on page 8 summarize the possible multi-tier configurations.

You create multi-tier extent pools by combining ranks of various drive technologies, or by merging extent pools with the wanted mix of drive technologies through the dynamic extent pool merge function.

The Easy Tier Automatic Mode fully supports up to three tiers in an extent pool, providing a three-tier storage hierarchy. By using Automatic Mode, you can distribute your entire workload across three-tier extent pools with fully automated storage performance and storage economics management, and efficient distribution of bandwidth and IOPS, across all storage tiers.

A logical configuration of the storage can be based on the capacity requirements and the percentage of flash, enterprise, and nearline drives that are required to satisfy the performance needs. Thin-provisioned extent space efficient (ESE) volumes with large or small extents are supported now. They benefit from Easy Tier automated storage performance and economics management.

**Important:** Consider a little “breathing” space for the Easy Tier to start working; even a few extents are fine.

By default, the DS8000 automatically reserves some storage for Easy Tier extent movements. You can check that this default behavior is enabled through the **`showsi`** command as illustrated in Example 4-1 on page 79. To disable or re-enable the DS8000 to reserve storage for Easy Tier extent movements, use the **`chsi`** command with the **`-etsrmode disable or enable`** option, as shown in Example 4-3 on page 82.

Also, as an extent pool is configured, if Easy Tier is enabled, Easy Tier will eventually start migrating extents to lower tier ranks even if no I/O activity exists. Therefore, you must not wait too many days after completing the system logical configuration before placing it into production. If you must wait for several days, consider disabling Easy Tier on those configured extent pools, or pinning certain volumes to a flash tier.

## 3.3.1 General considerations for DS8000 logical configurations

This section describes general considerations for DS8000 logical configurations.

### Extent pool configuration

Unassigned ranks do not have a fixed or predetermined relationship to any DS8000 storage server. Make sure that the ranks from each DA pair are distributed evenly between both servers, or you might seriously limit the available back-end bandwidth and the system's overall throughput.

For best performance and a balanced configuration, a preferred practice is to create a minimum of two or more extent pools. In these pools, at least one pool for rank group 0 is primarily managed by DS8000 storage server 0, and one pool for rank group 1 is managed by DS8000 storage server 1. Note that the DS GUI does such a balance by default.

This practice now becomes important with flash and small block random I/O workloads. To use the full back-end random I/O performance of two (or more) flash ranks within a certain DA pair, the flash I/O workload must be balanced across DS8000 storage servers.

During DS8000 configuration, you assign half of the ranks of each DA pair to even-numbered extent pools (P0, P2, P4, and so on) that are managed by storage server 0, and half to odd-numbered extent pools (P1, P3, P5, and so on) that are managed by storage server 1. This evenly balances the workload and the available back-end bandwidth between all DAs.

**Important:** Ranks from rank group 0 (even-numbered extent pools, such as P0, P2, and P4) are primarily managed by storage server 0. Ranks from rank group 1 (odd-numbered extent pools, such as P1, P3, and P5) are primarily managed by storage server 1.

Consider the following guidelines for best performance:

- Flash ranks must be properly assigned to even and odd rank groups (storage server 0 or 1).
- Be aware of the balanced distribution of the overall workload across both DS8000 storage servers. Typically, half of the ranks of each DA pair can be assigned to rank group 0 and the other half of the ranks can be assigned to rank group 1. This configuration is also used for the extent pools in most cases.
- In general, always use storage pool striping with the default rotate extents allocation method wherever possible when volumes are created in homogeneous multi-rank extent pools to balance the workload across all ranks in an extent pool.

Following these guidelines provides a balanced usage of the full back-end bandwidth and DS8000 storage servers.

To check for a correct logical configuration of ranks across DA pairs, complete these steps:

1. Identify the DA pair association for each rank by running the **lsarray -l** command, as shown in Example 3-1.

*Example 3-1 Identify flash rank to DA pair association*

```
dscli> lsarray -l
Array State      Data   RAIDtype    arsite Rank DA Pair DDMcap (10^9B) diskclass  encrypt
=========================================================================================
A0    Assigned   Normal 5 (6+P+S)   S11    R2   2               600.0 ENT        supported
A1    Assigned   Normal 5 (6+P+S)   S14    R3   10              400.0 FlashTier0 supported
A2    Assigned   Normal 5 (7+P)     S12    R0   2               600.0 ENT        supported
A3    Assigned   Normal 5 (7+P)     S10    R1   2               600.0 ENT        supported
```

```
A4    Assigned   Normal 6 (5+P+Q+S) S1     R8   0              600.0 ENT        supported
A5    Assigned   Normal 5 (6+P+S)   S9     R4   2              600.0 ENT        supported
A6    Assigned   Normal 5 (6+P+S)   S13    R5   10             400.0 FlashTier0 supported
A7    Assigned   Normal 6 (5+P+Q+S) S2     R9   0              600.0 ENT        supported
A8    Unassigned Normal 6 (5+P+Q+S) S3     -    0              600.0 ENT        supported
A9    Unassigned Normal 6 (5+P+Q+S) S4     -    0              600.0 ENT        supported
A10   Assigned   Normal 5 (7+P)     S5     R6   0              600.0 ENT        supported
A11   Assigned   Normal 6 (6+P+Q)   S6     R7   0              600.0 ENT        supported
```

2. Ensure that half the ranks from each DA pair are assigned or configured to even extent pools (P0, P2, and P4 and so on as rank group 0 that is managed by storage server 0) and half the ranks from the same DA pair are assigned or configured to odd extent pools (P1, P3, and P5 and so on as rank group 1 that is managed by storage server 1), as shown in Example 3-2.

*Example 3-2 Ranks assigned to odd and even extent pools*

```
dscli> lsrank -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts keygrp marray extsize
(cap)
==========================================================================================================
R0     0 Normal Normal    A2           5 P0        PFE_FB     fb      3666       73      1 MA12   1GiB
R1     1 Normal Normal    A3           5 P1        PFE_FB     fb      3666       90      1 MA10   1GiB
R2     0 Normal Normal    A0           5 P2        CKD_Hybrid ckd     3518      824      1 MA11   1113cyl
R3     0 Normal Normal    A1           5 P2        CKD_Hybrid ckd     2377       46      1 MA14   1113cyl
R4     1 Normal Normal    A5           5 P3        CKD_Hybrid ckd     3518      728      1 MA9    1113cyl
R5     1 Normal Normal    A6           5 P3        CKD_Hybrid ckd     2377       40      1 MA13   1113cyl
R6     0 Normal Normal    A10          5 P0        PFE_FB     fb      3666       64      1 MA5    1GiB
R7     0 Normal Normal    A11          6 P0        PFE_FB     fb      3075       54      1 MA6    1GiB
R8     1 Normal Normal    A4           6 P1        PFE_FB     fb      2550       50      1 MA1    1GiB
R9     1 Normal Normal    A7           6 P1        PFE_FB     fb      2550       51      1 MA2    1GiB
```

Remember, the same rule applies in general for all types of disk drives.

## Extent allocation method

Logical volumes can be allocated by using one of two extent allocation methods (EAMs): rotate extents or rotate volumes.

The rotate extents EAM (also known as *storage pool striping*) distributes the capacity that is allocated by one logical volume across the available ranks so that workloads with a random access I/O pattern can benefit from the performance of multiple ranks. The rotate extents EAM is generally preferable because it tends to avoid skew and hotspots on any one rank, because the workload of all volumes tends to be distributed over the available ranks.

This method is the default when new volumes are created and no EAM is specified in non-managed single-tier extent pools. The initial extent allocation that is used in managed extent pools is similar to rotate extents. For more information, see 2.8.9, “Extent allocation in hybrid and managed extent pools” on page 38.

The rotate volumes EAM tries to globalize the capacity that is allocated to one volume on a single rank, but it distributes volumes across the available ranks. This method might be desirable for limiting the number of ranks included in the logical volume regarding its failure boundary.

**Important:** Because Easy Tier applies different migration types to distribute the workload across tiers in hybrid pools, new volumes are always created on the enterprise class drive tier, or SSD first. However, extents can be demoted or promoted after the pool becomes managed by Easy Tier Automatic Mode. This configuration is reflected by the EAM attribute of managed for all volumes that are or were once under the control of Easy Tier Automatic Mode.

The ability to specify the EAM for a volume migration allows the EAM to be changed any time after the volume is created. Volume migration into the same pool specifying the rotate extents EAM can also be used to redistribute extents of a volume in a balanced manner in an extent pool after more ranks are added to the pool. This configuration is also known as *manual volume rebalance*.

In multi-tier or hybrid extent pools, extents are initially allocated by using rotate extents across all ranks of the same storage tier, starting with an extent allocation on the enterprise class, or the so-called *SSD home tier* first. Any specified EAM for the volume creation or migration command is ignored. Furthermore, the EAM of each volume in a managed pool is immediately set to managed after Easy Tier Automatic Mode management for this pool is enabled.

### 3.3.2 Guidelines for creating multi-tier extent pools

In managed *multi-tier* or *hybrid extent pools*, Easy Tier Automatic Mode automatically promotes hot extents from the lower tier to the upper tier (enterprise to flash, and nearline to enterprise), or demotes colder extents from the higher tier to the lower tier (swap extents from flash with hotter extents on enterprise, or demotes cold extents from enterprise to nearline).

Automatic Mode also optimizes the nearline tier by demoting part of the sequential workload to the nearline tier to better balance sequential workloads. Auto-rebalance rebalances extents across the ranks of the same tier based on rank usage to minimize skew and avoid hotspots. Auto-rebalance even takes different device characteristics into account when different devices or even RAID levels are mixed within the same storage tier (*micro-tiering* or *intra-tiering*).

The following guidelines are used for creating multitier or hybrid extent pools that are managed by Easy Tier Automatic Mode based on early experience with Easy Tier, on legacy hybrid DS8000 models:

- Flash drives to HDD drives ratio

  The ratio of flash capacity to HDD capacity in a hybrid pool depends on the workload characteristics and skew. Ideally, there must be enough flash capacity to hold the active (hot) extents in the pool, but not more, to not waste flash capacity. For new DS8880 orders, 10% of flash capacity might be a reasonable percentage to plan for with hybrid pools in typical environments.

  This amount also allows all metadata to be stored in flash for best performance. This configuration can result in the movement of 50% of the small and random I/O workload from enterprise drives to flash. This configuration gives a reasonable initial estimate if measurement data is not available to support configuration planning.

- Enterprise-HDD to nearline-HDD ratio

  For the ratio of enterprise-HDD capacity to nearline-HDD capacity, the previous considerations also apply, with the enterprise tier holding the hot data, and the nearline tier holding the cold data. Use the Easy Tier monitoring data with STAT or the GUI to determine the actual ratio.

- Drive types

  Plan the drives with the intended workload in mind. Already considering DS8880, flash drives are always a good choice, but there are usually storage areas with nearly no I/O activity. Therefore, you may plan for some Nearline storage for each extent pool. But you can also have extent pools with only one tier but with some different drive types. Mixed drive types can be managed by the enhanced micro-tiering capability of Easy Tier.

  For DS8900F and an all-flash approach, a similar reasoning as above will apply when considering the right mix between high-performance flash and the ratio and type of capacity-optimized flash to be chosen.

  Enhanced micro-tiering accounts for device-specific performance characteristics when determining the rank usage for intra-tier performance management (auto-rebalance), such as device variations (speed and RAID level) and workload characteristics.

  However, micro-tiering does not deliberately apply cross-tier extent migration types between the different device types within a tier. It applies auto-rebalance migrations, balancing the workload within the drive tier across ranks based on rank characteristics and usage. This configuration allows your system to avoid skew and hotspots within a storage tier. A mix of drive types and RAID levels in an extent pool impedes performance analysis and makes it more difficult for you to understand what is happening.

- RAID types

  Flash and HDD arrays in a hybrid pool do not need to have the same RAID level. It is likely that all arrays will use RAID 6 in a pool, but there might be circumstances where RAID 10 is also used.

  Whether RAID 10 can be used for nearline drives in a hybrid pool depends on the write content of the workload. RAID 10 typically provides better performance for write-intensive workloads. However, in a DS8880, RAID 10 is not needed as much in a hybrid pool where Easy Tier Automatic Mode migrates active extents to flash ranks. Also the low-latency flash drives in DS8900F make the need for RAID 10 often less pending, since already RAID 6 is giving an excellent performance out of each array. Still, performance benefits might be gained by using RAID 10.

  Also, with Easy Tier, small-block random reads are favored for the flash tier more than random writes. If your workload is characterized by many random writes, you might want to use RAID 10. For enterprise drives, RAID 6 is now the preferable implementation, unless write ratios are higher, for example, exceeding 30% and no flash tier is present. RAID 5 is only available through RPQ, and only for drives below 1 TB of capacity. For large-capacity flash, RAID 6 is mandatory to prevent the danger from double RAID errors that a RAID 5 still potentially has.

- Size and number of extent pools

  With Easy Tier, you can manually redistribute the extents of existing volumes over all ranks in a non-managed, single-tier extent pool by using Easy Tier Manual Mode and the dynamic volume relocation capability (manual volume redistribution). These functions help improve performance by avoiding skew and hotspots, because volume extents can be spread evenly over more ranks and DA pairs in the pool.

  However, hybrid pools can be managed by Easy Tier Automatic Mode. In managed extent pools, no manual restriping is required (or supported) to redistribute the extents after new capacity is added to the pool. This situation occurs because Easy Tier automated cross-tier and intra-tier performance management (auto-rebalance) automatically uses the new capacity.

  Auto-rebalance automatically relocates data based on extent workload and rank usage, providing a balanced resource usage within a drive tier to avoid skew and rank hotspots.

The best approach is to create large extent pools, therefore improving performance and simplifying storage management. Ultimately, you might have only two extent pools for each rank type (fixed block (FB) and count key data (CKD)), one on storage server 0 and one on storage server 1.

However, multiple extent pools are created for the following reasons:

– Drives of the same storage tier with various characteristics (capacity, speed, and RAID type) typically cannot be mixed in the same extent pool, but exceptions might apply under certain circumstances.

– Separate extent pools can be based on failure boundaries. For example, backup volumes, such as permanent IBM FlashCopy volumes that are created with the **copy** option, can be placed in a separate extent pool from the original volumes.

  With this situation, a single rank failure does not affect the original and the backup. However, this operation is unnecessary if the FlashCopy target is protected against rank failures (for example, in a Metro Mirror environment).

– Separate extent pools with different tiers, depending on user needs, such as one extent pool with three tiers (flash + enterprise + nearline), and another extent pool with two tiers (flash + enterprise). With this configuration, you can plan accordingly for specific host needs and different workload characteristics or situations, for example, when you must plan for specific minimum or maximum requirements.

### 3.3.3 Guidelines for creating single-tier (homogeneous) extent pools

Single-tier or homogeneous extent pools contain drives from only one storage tier, as shown in the following examples:

- High-performance flash drives and traditional flash drives (SSDs) (3.2 TB and less)
- High-capacity flash drives, from a Flash Tier 1 type only (3.84 TB)
- High-capacity flash drives, from a Flash Tier 2 type only (7.68 TB+)
- Enterprise drives (15K or 10K RPM - legacy models)
- Nearline drives (7.2K RPM - legacy models)

A DS8000 configuration without any hybrid extent pools can also benefit from Easy Tier automatic performance management in homogeneous single-tier extent pools by using the auto-rebalance feature. Easy Tier monitors the rank usage and rebalances the extents across all ranks of the same storage tier within the extent pool to reduce skew and avoid hotspots, therefore providing the best achievable performance in a homogeneous extent pool.

Note the following information:

- Consider configuring multi-rank extent pools to allow auto-rebalance to operate more effectively.
- Even single-tier pools with different device characteristics benefit from Easy Tier auto-rebalancing, because the algorithms consider the different device characteristics (micro-tiering).
- Consider leaving free capacity in the extent pool to allow auto-rebalance to efficiently relocate extents across ranks or let the system reserve storage for this purpose.
- High-performance flash drives and traditional flash drives (SSD) are considered the same storage tier.
- Enterprise disks with different drive speeds (enterprise 10K RPM and 15K RPM) are considered the same storage tier.

Easy Tier Auto Mode can be set to `all`, and Easy Tier monitoring can be set to `all` or `automode` to enable volume auto-rebalance in all single-tier pools.

In tiered mode, auto-rebalance is enabled in hybrid or multitier pools only. If Easy Tier mode is set to `none`, no volumes are managed by Easy Tier. However, you can manually redistribute a logical volume within a non-managed, single-tier extent pool by using manual volume rebalance. Alternatively, you can migrate a volume from one extent pool to another by using the dynamic volume relocation (DVR) feature of Easy Tier Manual Mode.

The enhanced functions of Easy Tier Manual Mode offer more management capabilities and ease of use. You can use Easy Tier Manual Mode features to easily relocate the extents of a volume within an extent pool, or to relocate an entire volume from one pool to another pool. When a volume migration is targeted to the same pool and the target EAM is rotate extents, the volume migration acts internally as manual volume rebalance.

Use manual volume rebalance to relocate the smallest number of extents of a volume and restripe the extents of that volume on all available ranks in the pool. The behavior of the regular volume migration, which differs from manual volume rebalance, continues to operate as it did in the previous version of Easy Tier.

Manual volume rebalance evenly rebalances the extents of a volume across all ranks within a non-managed, single-tier extent pool, which is helpful, for example, if new ranks were added to the pool or if pools were merged. A hybrid or multitier extent pool is always assumed to be prepared for Easy Tier Automatic Mode management and does not allow the manual redistribution of extents within the pool by using the dynamic volume relocation feature.

Volume restriping within an extent pool can be achieved by initiating manual volume rebalance. Use manual volume rebalance to minimize skew and hotspots in a non-managed, single-tier extent pool when the capacity distribution of the volumes is highly imbalanced across ranks.

However, in managed single-tier extent pools, no manual restriping is required to redistribute the extents after new capacity is added to the pool, because Easy Tier automated intra-tier performance management (auto-rebalance) automatically uses the new capacity. Auto-rebalance automatically relocates data based on extent workload and rank usage, which provides a balanced resource usage within a drive tier to avoid skew and rank hotspots.

**Important:** When you manually relocate the extents of a volume, only the extents of the volume (the volume's capacity) is spread evenly across the available ranks of the target extent pool (volume striping). No workload characteristics are considered for any manual extent relocation.

You can still create standard single-tier extent pools for dedicated workloads that are not managed by Easy Tier, for example, specific optimized configurations where you do not want an automated extent relocation. These configurations might be highly optimized environments with dedicated resources that do not allow resource sharing between applications or different clients, and where host-based data striping and allocation methods are efficiently applied already.

Another example is if you applied dedicated extent pools for dedicated storage tiers and selected workloads, and you manually manage the placement of volumes on the appropriate storage tier. However, you might want to consider enabling and benefiting from the auto-rebalance feature of Easy Tier in single-tier extent pools, to reduce skew and avoid single ranks becoming hotspots.

You can gradually and without interrupting your host workloads move to full Easy Tier Automatic Mode for automated intra-tier and cross-tier performance management whenever you want by using the Easy Tier Manual Mode features, such as dynamic extent pool merge and dynamic volume relocation.

**Use of Easy Tier Automatic Mode is preferred:** This mode provides flexible storage management with ease of use, combined with subvolume storage performance and economics management at an excellent price-to-performance ratio.

Large hybrid pools with Easy Tier Automatic Mode offer a safe and easy configuration if you do not have special requirements to design dedicated highly optimized extent pools. Easy Tier *automatically* handles data placement across tiers and within tiers for the best overall performance.

### 3.3.4 Other considerations for thin-provisioned volumes

Virtual storage and space-efficient repositories do not allow you to indicate a specific EAM, but implicitly use rotate extents as the allocation method. Easy Tier fully supports thin-provisioned ESE volumes with Easy Tier Automatic Mode and Manual Mode.

In hybrid extent pools, thin-provisioned volumes are not limited to non-flash ranks. Easy Tier fully supports ESE volumes in Automatic and Manual Mode, and rank depopulation for ranks that contain ESE volumes.

**Thin provisioning:** Easy Tier supports thin-provisioned ESE volumes in Automatic Mode and Manual Mode (dynamic volume relocation, extent pool merge, and rank depopulation).

In environments that use only FlashCopy copy services that already support the use of thin-provisioned volumes (ESE), you might start by using thin-provisioned volumes in managed hybrid extent pools. This configuration avoids the initial creation of fully provisioned volumes on the nearline drive tier when the storage capacity on the home tier (enterprise or SSD) is exhausted because of fully provisioned volumes with unused allocated capacities being created.

In this case, only used capacity is allocated in the pool, and Easy Tier does not need to move unused extents around, or move hot extents on a large scale from the nearline tier to the enterprise tier and then to the flash tier.

For more information about initial volume allocation in hybrid extent pools, see 2.8.9, “Extent allocation in hybrid and managed extent pools” on page 38.

## 3.4 Implementation considerations for multi-tier extent pools

For multi-tier extent pools that are managed by Easy Tier Automatic Mode, other considerations exist about how to allocate the initial volume capacity and how to gradually introduce new workloads to managed pools.

### 3.4.1 Using thin provisioning with Easy Tier in multi-tier pools

In environments that are using only FlashCopy copy services that already support the usage of thin-provisioned volumes (ESE), you might start by using thin-provisioned volumes in

managed hybrid extent pools (especially with three tiers). You use this configuration to start the initial volume allocation for as many volumes as possible on the enterprise (home) tier.

You also use it to avoid the initial creation of fully provisioned volumes on the nearline drive tier when the storage capacity on the enterprise tier is already exhausted. This exhaustion is because of the creation of fully provisioned volumes with unused allocated capacities.

You can also change the allocation policy according to your needs with the **chsi** command and the **-ettierorder highutil** or **highperf** parameters. The preferred tier, as selected by the allocation policy, is known as the home tier.

With thin-provisioning, only used capacity is allocated in the pool and Easy Tier does not move unused extents around, or move hot extents on a large scale up from the nearline tier to the enterprise tier and to the flash tier. For more information about the initial volume allocation in hybrid extent pools, see 2.8.9, “Extent allocation in hybrid and managed extent pools” on page 38.

### 3.4.2 Staged implementation approach for multi-tier pools

Assume that you put a new DS8000 system into production that replaces an older system that does not have Easy Tier implemented. Put thought into the timeline of the implementation stages, that is, migrating all servers from the older system to the new system.

In a new three-tier environment of a hybrid storage system, volumes are first created on the enterprise tier, and Easy Tier cannot “learn” and optimize before production starts. If you are migrating all the servers at once, some servers’ volumes, for reasons of space, might be placed completely on nearline drives first, although these servers also might require higher performance.

Consider the following staged approach when you migrate servers to a new multi-tier DS8000 system:

- Decide what allocation policy you want to use, `highutil` or `highperf`.
- Decide whether you want to reserve some space in each extent pool manually for Easy Tier extent movements, or, better, let the system decide how much to reserve (**-etsrmode** setting).
- Assign the resources for the high-performing and response time-sensitive workloads first, then add the lower-performance workloads. The use of the opposite configuration might lead to situations where all initial resources, such as the enterprise tier in hybrid extent pools, are allocated already by the secondary workloads.

  This situation does not leave enough space on the enterprise tier for the primary workloads, which then must be on the nearline tier initially.
- Split your servers into several subgroups, where you migrate each subgroup one by one, but not all of them at once. Then, allow Easy Tier several days to learn and optimize. Some extents are moved to the flash tier, and some extents are moved to the nearline tier, and you regain space on the enterprise tier.

  After a server subgroup finishes learning and nearly reaches a steady state, the next server subgroup can be migrated. You gradually allocate the capacity in the hybrid extent pool by optimizing the extent distribution of each application one by one, while space is regained in the enterprise tier (home tier) for the next applications.
- Include the use of advanced Easy Tier Application, and Easy Tier Heat Map Transfer features in this implementation plan to help you optimize the performance.

### 3.4.3 Working with pinned volumes

Consider the situation in which your system has workloads that generate a higher load on your DS8000 occasionally (such as quarterly batch runs or an end-of-year run).

Put your important volumes on the top tier in a multi-tier pool before this type of a run. They can be prestaged to the flash tier, for example, two days in advance, and be pinned there until after the time of the rare batch run, when you release them again.

This idea is outlined in *DS8870 Easy Tier Application*, REDP-5014. This possibility of pinning volumes temporarily now allows you to use large multi-tier extent pools even in situations where previously you laid out such special-load applications onto separate pools.

In addition, you can prevent a volume from being assigned to a nearline tier.

## 3.5 Copy Services considerations

Copy Services are supported by Easy Tier. Copy Services, such as Metro Mirror, Global Mirror, and FlashCopy, do not interfere with the operations of Easy Tier (in Automatic Mode or Manual Mode), and the Copy Services are not aware that Easy Tier optimization is active.

All back-end I/Os, except the extent migration encountered with Easy Tier Manual Mode (dynamic volume relocation) and Automatic Mode (Automatic Data Relocation), are counted in the Easy Tier I/O statistics, including Copy Services back-end I/Os.

However, most of the Copy Services background I/O activity has sequential access patterns and does not contribute to the cross-tier heat calculation of Easy Tier on a bigger scale. However, most of the Copy Services background I/O activity is considered for the bandwidth and rank usage calculations.

For FlashCopy Copy-on-Write and FlashCopy Background Copy, Easy Tier specifically considers these processes and correctly assigns these loads to the respective source and target ranks of a DS8000 server complex, so that a proper heat map is established in these cases.

Like host operations, Copy Services (Metro Mirror, Global Mirror, and z/OS Global Mirror) are unaware of the extent or volume-level relocations that are performed.

In a Metro Mirror environment, another time delay occurs because of the required data transfer of a write I/O to the secondary site. This latency or service time is not included in the performance data that is considered by Easy Tier because this I/O activity is not an activity that occurs in the drive back end on the rank level.

The following considerations apply when Copy Services are used, assuming a symmetrical configuration:

- When volume migration is used, you must have procedures in place to move volumes at the primary and secondary systems to ensure that the same performance levels can be maintained in the event of a failover to the secondary system.
- When Easy Tier is in Automatic Mode, the workloads that are monitored on the primary and the secondary system can differ. Easy Tier at the primary system sees a normal workload, and at the secondary system, it sees only the write workloads.

  This situation means that the optimized extent distribution on the primary system can differ considerably from the one on the secondary system, if no heat map transfer function is

used. The optimized extent reallocation that is based on the workload learning on the primary system is not sent to the secondary system at this time to allow the same extent optimization on both systems based on the primary workload pattern.

In a disaster recovery situation with a failover from the primary site to a secondary site, the extent distribution of the volumes on the secondary system is not optimized to match the primary workload in general.

Easy Tier relearns the production I/O profile and builds a new extent migration plan on the secondary system to adapt to the new production workload before eventually achieving the same optimization and level of performance as on the primary system. This task takes time, so the production workload on the secondary system might not run at its optimum performance.

- FlashCopy copy-on-write activity results in full track writes, which are considered large rather than small. The FlashCopy target is not hot in terms of small I/Os unless it is written to.

Copy Services and Easy Tier can use the Easy Tier Heat Map Transfer Utility (HMTU) to optimize the secondary DS8000. This tool transfers the heat map from the primary site to the secondary site. In addition, IBM GDPS® (for IBM Z clients), or the Copy Service Manager (CSM) can take over the task of transferring the heatmap, thereby eliminating the need for installing HMTU.

Easy Tier at the secondary DS8000 can base its extent movement plan on this heat map, and optimize data placement according to the heat and the tiers available at the secondary site to optimize the performance in a failover. The HMTU can avoid the considerations that were described earlier in this topic.

For more information about identifying the cases where this feature can be used and how to implement it, see *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015.

## Safeguarded Copy

When using the Safeguarded Copy function, the additional workload to be expected is comparable to the equivalent of about one additional FlashCopy running, even with many concurrent target copies specified. So generally, lower performing capacity is fine for such targets, for instance planning with Flash Tier 2 capacity-optimized flash in a DS8900F.

**Important:** Easy Tier does *not* monitor the Safeguarded Copy Backup Capacity.

For more information on Safeguarded Copy, refer to *IBM DS8880 Safeguarded Copy,* REDP-5506.

## 3.6 Workload planning

Easy Tier environments provide greater choice for the number of drive combinations than a single-tier of storage. If you plan to use these environments, consider the following key questions:

- What does an initial single-tier configuration for the planned workload look like?
- How much of my data performs little random I/O or only some sequential I/O and might be considered cold?
- What is the skew of the data and how much flash capacity is needed to handle the major part of the random workload?
- Using flash or nearline drives, what size enterprise drives can I use to handle the workload that remains on the enterprise tier?
- Do I need to provide better back-end performance in the target environment to accelerate application workloads?
- Will flash storage provide significant application acceleration for my type of workload (cache hit percentage sequential workload)?

When you plan drive configurations for Easy Tier, always start planning for a single target drive tier first, which is the tier where the workload is placed initially in a regular single-tier configuration. Typically, this tier is the enterprise tier. Consider the skew of the workload and adapt the initial configuration by adding higher drive tiers for hot data and lower drive tiers for cold data. Adapt the number of drives and drive size of the initial target tier accordingly.

For example, if three tiers are available (flash, enterprise, and nearline), the following planning considerations might be included:

- Databases might be targeted to the enterprise tier first to promote hot extents to the flash tier and demote cold extents to the nearline tier. This configuration uses all three tiers while capacity on the enterprise tier is released.
- File sharing services might be targeted at the nearline tier instead of promoting some of the extents to the enterprise tier and not promoting them to a flash tier to prevent usage of the bandwidth of the flash drives. This configuration uses only two tiers.

Plan the capacity of the target tier and then plan the capacity of the associated tiers. For planning guidelines, see 3.3.2, “Guidelines for creating multi-tier extent pools” on page 52. This approach allows the DS8000 to use its internal Automatic Mode algorithms to balance the performance. It also constantly adapts to the changing application workloads, and leaves the manual control to the administrators only where necessary. You can always add a tier (or ranks) to the existing extent pools or redistribute the workload across the ranks that are already in use.

With three-tier support, most of the existing workloads benefit from Easy Tier Automatic Mode. Consider the following points:

- Online transaction processing (OLTP) and randomized-nature workloads have better response times.
- Large block sequential and batch workloads benefit by moving workloads partly to the nearline tier, which frees space on enterprise drives and releases bandwidth on the enterprise tier.
- Highly mixed and unpredictable workloads are managed and monitored by Easy Tier, which provides heat distribution and cold extent allocation with the recommendations about the performance improvement.

The following workloads might not benefit from the Easy Tier Automatic Mode, but Manual Mode features might be used to support the management of these workloads:

- Hotspots are small and uniformly distributed across extents so that all extents exhibit equal temperatures. This case might not benefit from Automatic Mode, but Manual Mode might support the management of these workloads by using successive dynamic volume relocations.
- Hotspots vary over time so that they are uniformly distributed over a long enough monitoring period. This case also might not benefit from Automatic Mode, but Manual Mode might be used to move these workloads to one or several extent pools that are created for them specifically.

The Easy Tier Application can help in these kinds of workloads and volume placements, moving all extents that are associated with volumes to a higher tier level. For more information about Easy Tier Application, see *DS8870 Easy Tier Application*, REDP-5014.

4

# DS GUI and DS CLI support for IBM DS8000 Easy Tier

This chapter describes the Disk Storage graphical user interface (DS GUI) and command-line interface (DS CLI) support and controls for IBM System Storage Easy Tier.

This chapter includes the following topics:

- DS GUI support for Easy Tier
- DS CLI support for Easy Tier
- Easy Tier functions by using the DS CLI
- Easy Tier control by using the DS CLI

# 4.1 DS GUI support for Easy Tier

You can manage Easy Tier by using the DS Storage Management GUI (for certain functions) or DS CLI. The Easy Tier support in the GUI relates to dynamic volume relocation (or volume migration), dynamic extent pool merge, redistribute volumes, and rank depopulation capabilities.

**Note:** New systems shipped from manufacturing have Easy Tier set by default to manage all tiered pools.

## 4.1.1 Easy Tier controls at the system level

To set Easy Tier controls at the system level, select **Settings** → **System** → **Easy Tier** (Figure 4-1).

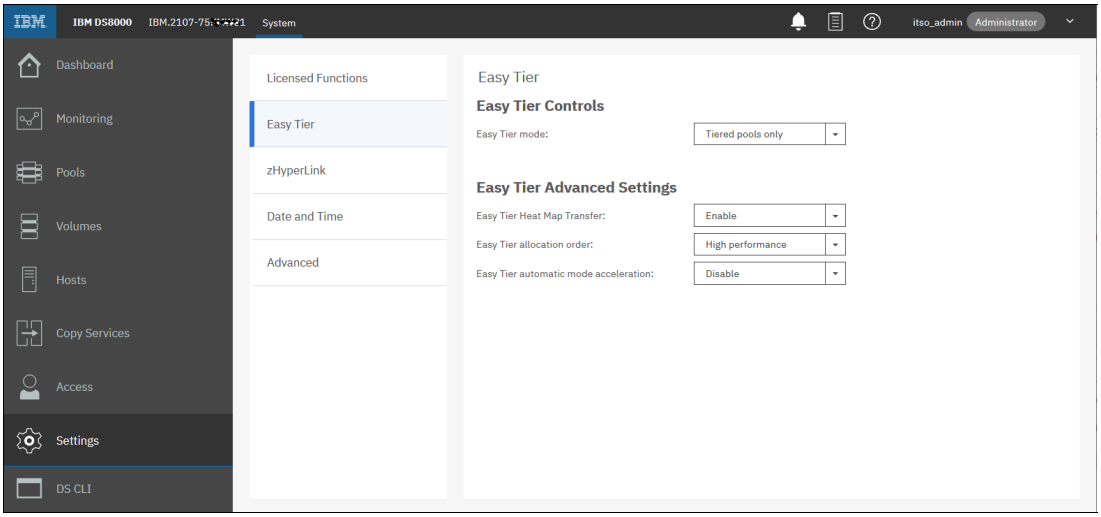



*Figure 4-1 System-level Easy Tier Settings*

Under Easy Tier Controls, specify the **Easy Tier Mode**, and under Easy Tier Advanced Settings**,** specify the **Heat Map Transfer** and **Easy Tier allocation order** settings, as well as, for DS8900, the option for **Easy Tier automatic mode accleration.**

### Easy Tier Controls

Easy Tier Mode determines how Easy Tier monitors I/O activity (Figure 4-2).

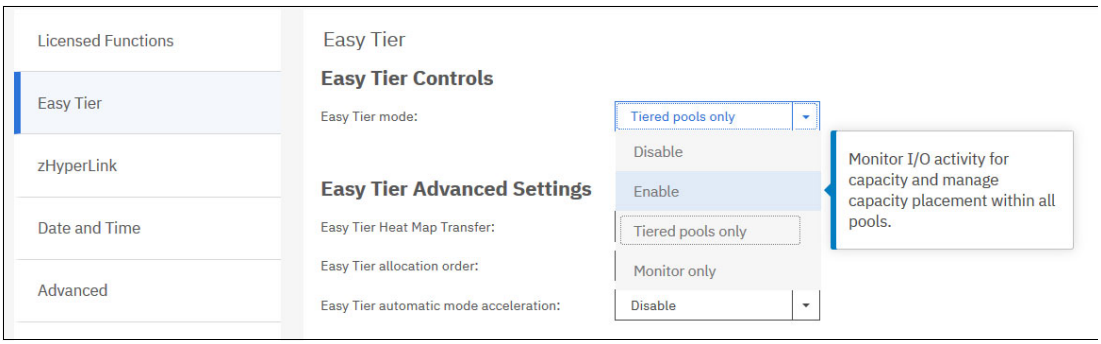



*Figure 4-2 Easy Tier Controls: Easy Tier mode*

One of the following settings can be selected:

- **Enable**: Monitor I/O activity for capacity and manage capacity placement within the pools.
- **Disable**: Do not monitor I/O activity for capacity or manage capacity placement.
- **Tiered pools only**: Monitor I/O activity for capacity and manage capacity placement within pools that have arrays with different drive classes.
- **Monitor only**: Monitor I/O activity for capacity, but do not manage capacity placement.

**Important:** Changing the Easy Tier mode affects the statistics collection. Changing the Easy Tier mode to `Disable` can lead to a reset (reinitialization) of the gathered monitoring data. This situation means that collecting new performance statistics after Easy Tier monitoring is enabled again, when migration plans are created, can take 24 hours.

## Easy Tier Advanced Settings

Under Easy Tier Advanced Settings, you can control the **Easy Tier Heat Map Transfer** and **Easy Tier allocation order** settings.

**Easy Tier Heat Map Transfer** has these options (Figure 4-3):

- **Disable**: Disables the Easy Tier Heat Map utility.
- **Enable**: Enables the Easy Tier Heat Map utility to maintain knowledge of I/O activity between a primary and auxiliary storage system.

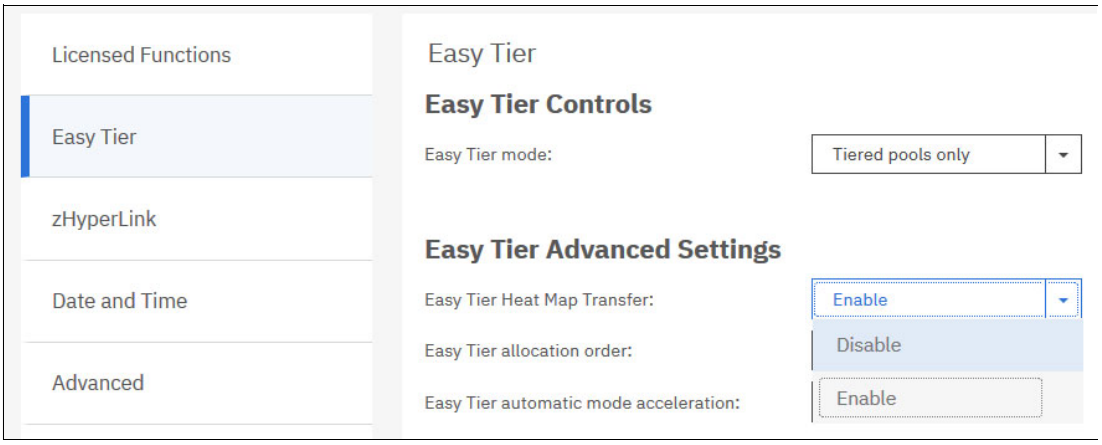



*Figure 4-3 Easy Tier Advanced Settings: Easy Tier Heat Map Transfer*

**Easy Tier Allocation order** has these options:

- **High utilization**: Easy Tier allocates data in drive classes in this order:
  - All-Flash configuration: Flash Tier 1 → Flash Tier 2 → Flash Tier 0
  - Hybrid configuration (DS8880 and older models): Enterprise tier → Flash Tier 1 → Flash Tier 2 → Nearline tier → Flash Tier 0
- **High performance**: Easy Tier allocates data in drive classes in this order:
  - All-Flash configuration: Flash Tier 0 → Flash Tier 1 → Flash Tier 2
  - Hybrid configuration (DS8880 and older models): Flash Tier 0 → Flash Tier 1 → Flash Tier 2 → Enterprise tier → Nearline tier

Figure 4-4 shows the details of the Easy Tier allocation order for **High utilization**.

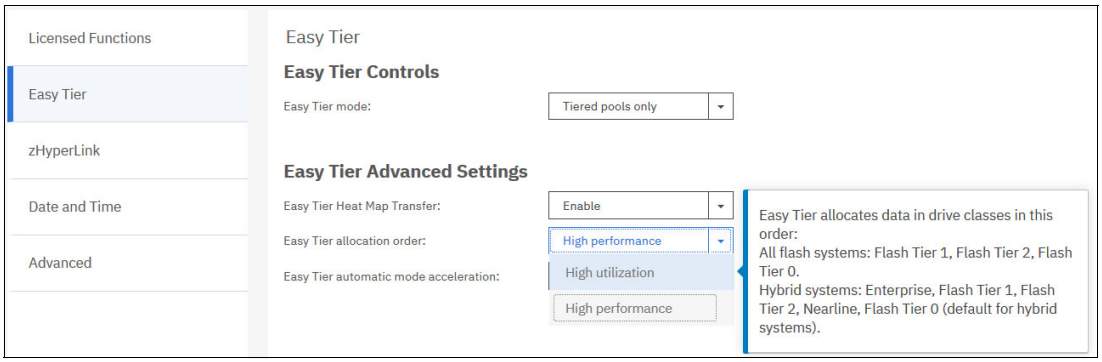



*Figure 4-4 Easy Tier Allocation order: High utilization*

Figure 4-5 shows the details of the Easy Tier allocation order for **High performance** (=default with DS8900F).

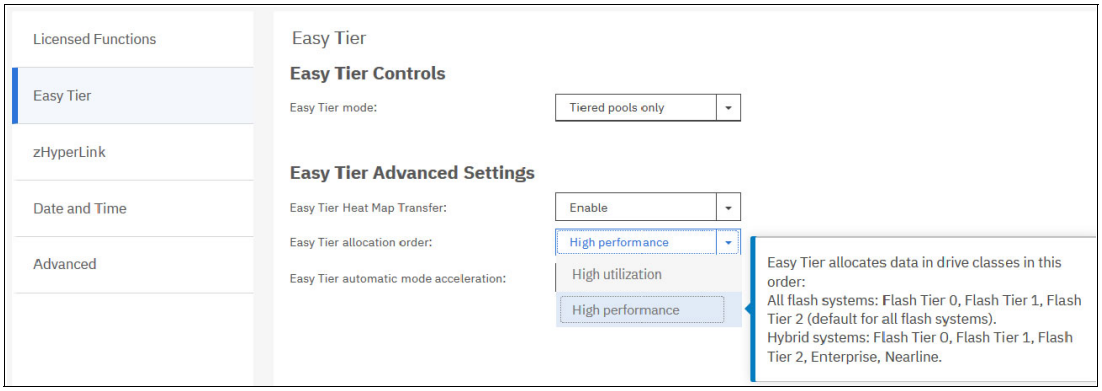



*Figure 4-5 Easy Tier Allocation order: High performance*

**Easy Tier automatic mode acceleration** (DS8900F model only) has the option to disable it (default), or to enable. When having this option enabled, the speed of migrations will be around three times higher. This option is only meant for exceptional one-off situations, not for steady use.

## 4.1.2 Volume Properties window

The Volume Properties window shows details and the status of the selected volumes. Easy Tier processing includes the following configuration states:

- Managed: The volume is managed by Easy Tier.
- Disabled: The volume is not managed by Easy Tier.
- Monitor Only: Easy Tier is monitoring the volume but will not migrate data.
- Assigned to Flash: The volume is assigned to the flash tier.
- Assigned to Enterprise: The volume is assigned to the enterprise tier.
- Assigned to Nearline: The volume is assigned to the nearline tier.
- Excluded from Nearline: The volume is excluded from the nearline tier.
- Learning Paused: The process of learning about the volume is paused.
- Migration Paused: Easy Tier migrations of the volume data are paused.

All of these states relate to the dynamic volume migration facility. For more information, see 4.1.8, “Dynamic volume relocation” on page 70. Figure 4-6 shows the volume, state of the volume, and migration capacity (the size of the remaining data to be migrated), from the volume properties detail window.

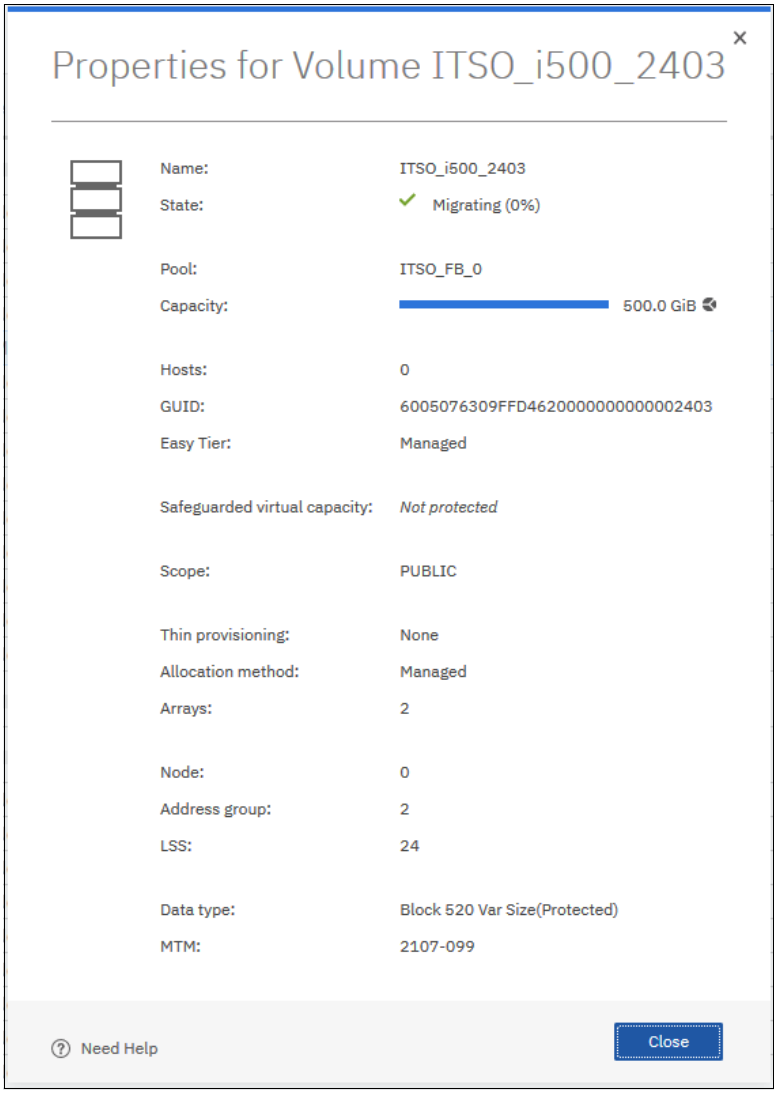



*Figure 4-6 Volume Properties window: Showing the migration status*

The count key data (CKD) Single Volume Property window shows the same fields in relationship to Easy Tier.

## 4.1.3 Volume migration window

The volume migration-related windows support the dynamic volume relocation that is available with Easy Tier mode set to **Enabled**. For a description of these windows, see 4.1.8, “Dynamic volume relocation” on page 70.

## 4.1.4 Disk configuration summary

In the DS8000 GUI, a user can get information about disk configuration by clicking **Pools** → **Arrays by Pool** in the main window. In the Drive Class column on the Arrays by Pool state page (Figure 4-7), the drive class options are Flash Tier 0, Flash Tier 1, Flash Tier 2, and including legacy hybrid machines: Enterprise 15K or 10K, and Nearline 7.2K.

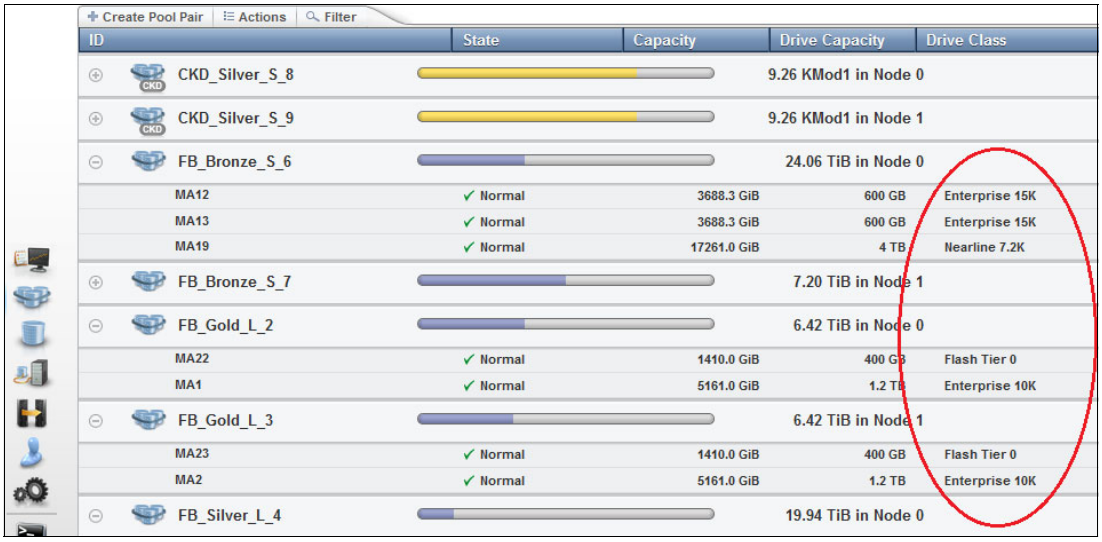



*Figure 4-7 GUI Drive Class representation, and Arrays by pool*

### Extent pools

The extent pools window shows the relevant information about each extent pool from an array perspective. Figure 4-7 also shows the storage capacity as a bar for each extent pool.

## 4.1.5 Pool Properties window

Use the Pool Properties window (Figure 4-8 on page 69) to view the extent pool detail. This window shows the following information:

- Name and Pool ID.
- Server affinity node (0 or 1).
- Capacity with or without extent space efficient (ESE) repository. The repository can only be defined from the DSCLI.
- Encryption state of all data in the pool.
- Easy Tier information (Managed, Disabled, Monitor only, Learning paused, Migration paused).
- Number of arrays inside the pool.
- Number of defined volumes inside the pool.
- Number of extents.

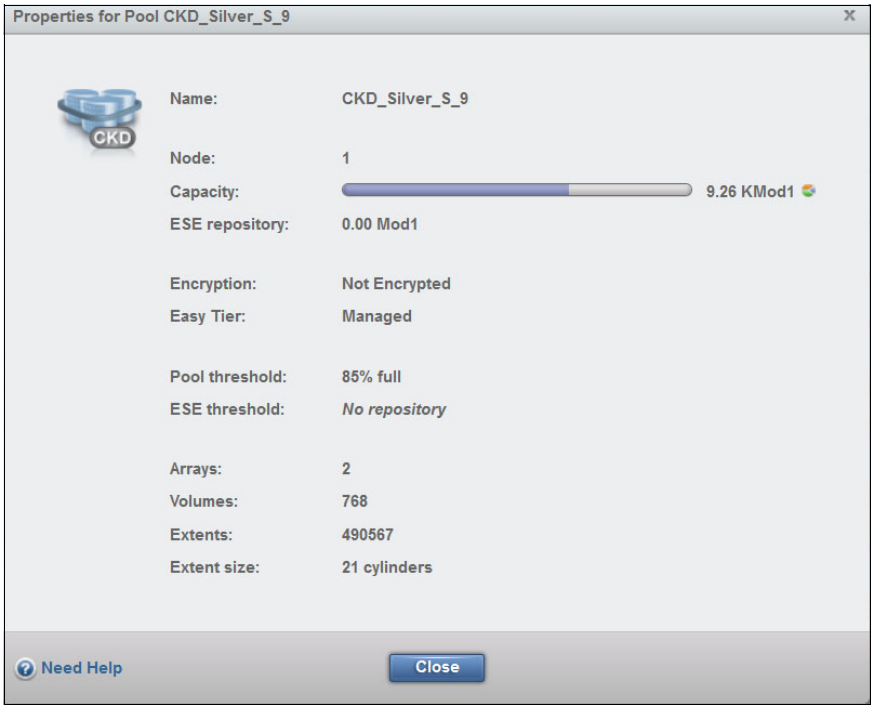



*Figure 4-8 Pool Properties window*

## 4.1.6 Manage Volumes window

To open the Manage Volumes window, click **Volumes** → **Volumes by Pool** (Figure 4-9).


*Figure 4-9 Manage Volumes window*

Figure 4-10 shows the available functions for the selected volume. In relation to Easy Tier, you can choose Migrate or Redistribute tasks. Depending on the state of the volume, the DS Storage Management GUI provides other functions (which are described later).

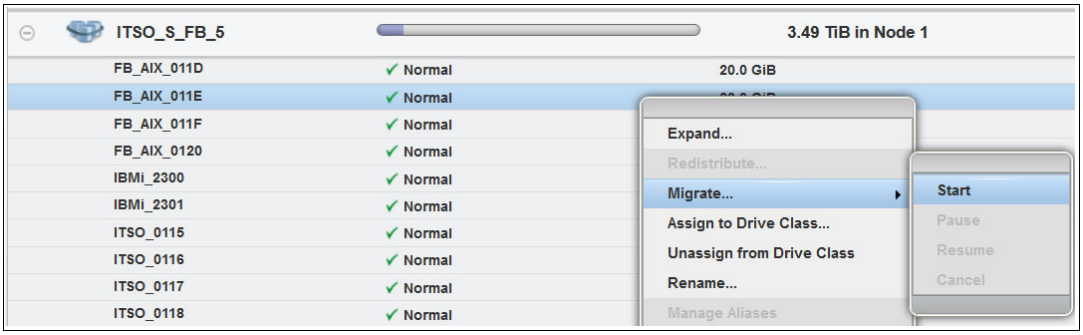



*Figure 4-10 Migration option*

## 4.1.7 Merge Extent Pools window

A user can initialize the extent pool merge by using either path: **Pools** → **Volumes by Pool** or **Pools** → **Arrays by Pools**. Figure 4-11 illustrates the merge process. The source and target pools are selected. Then, the Merge action is selected from the pop-up menu. When this operation finishes, the source extent pool is deleted, leaving only the target extent pool.

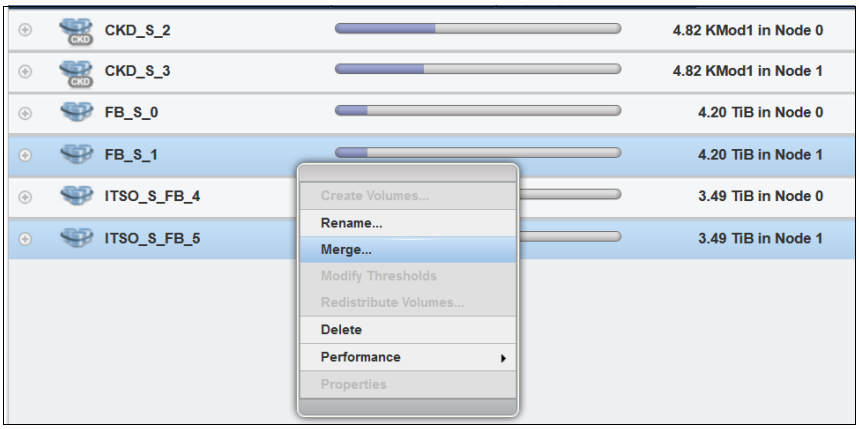



*Figure 4-11 Merge Extent Pools window*

**Restriction:** You can only merge extent pools that contain extents of the same size (small or large) and same type (FB or CKD).

## 4.1.8 Dynamic volume relocation

Dynamic volume relocation (DVR) allows volumes to be migrated (moved) from one extent pool to another (or even back into the same extent pool) transparently to host operations. DVR is an Easy Tier capability.

With the GUI, you can perform these tasks in the Migrate Volume window:

- **Start**: Starts a volume migration
- **Pause**: Temporarily suspends a volume migration
- **Resume**: Resumes a temporarily suspended volume migration
- **Cancel**: Cancels a volume migration

**Consideration:** DVR is allowed only among extent pools of the same volume storage type (count key data (CKD) or fixed block (FB)), of the same server affinity, and of the same extent size. You can migrate volumes between extent pools that are managed by the same storage server (0 or 1).

For more information, see “Dynamic volume relocation” on page 22 and “Manual volume rebalance” on page 23.

## Starting relocation

The volume relocation (migration) can be started from the Volumes by Pool window. Select **Volumes** → **Volumes by Pool**, as shown in Figure 4-12.

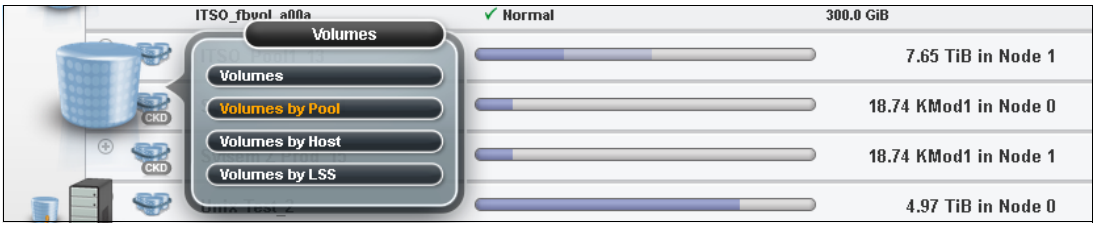



*Figure 4-12 Volumes by Pool window*

To start the volume relocation, select the volume that you want to move, right-click, and click **Migrate** → **Start**, as shown in Figure 4-13.

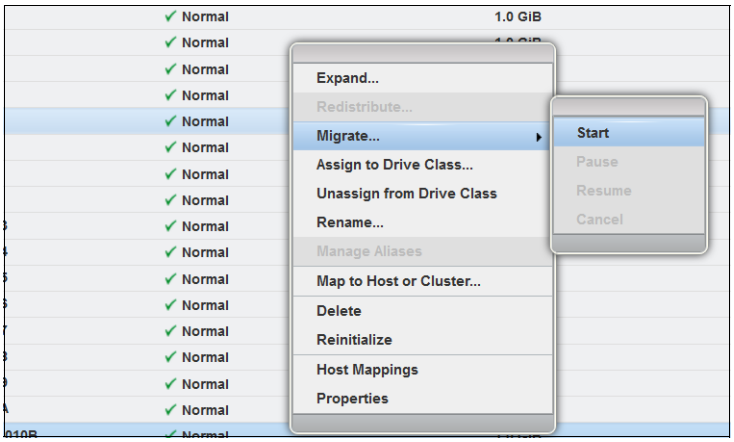



*Figure 4-13 Migration of a selected volume*

In the next dialog window, which is shown in Figure 4-14, choose a target extent pool and click **Migrate**.


*Figure 4-14 Select a target pool for the volume*

The migration process starts, and the volume is moved dynamically to the target pool. The process completes (Figure 4-15).

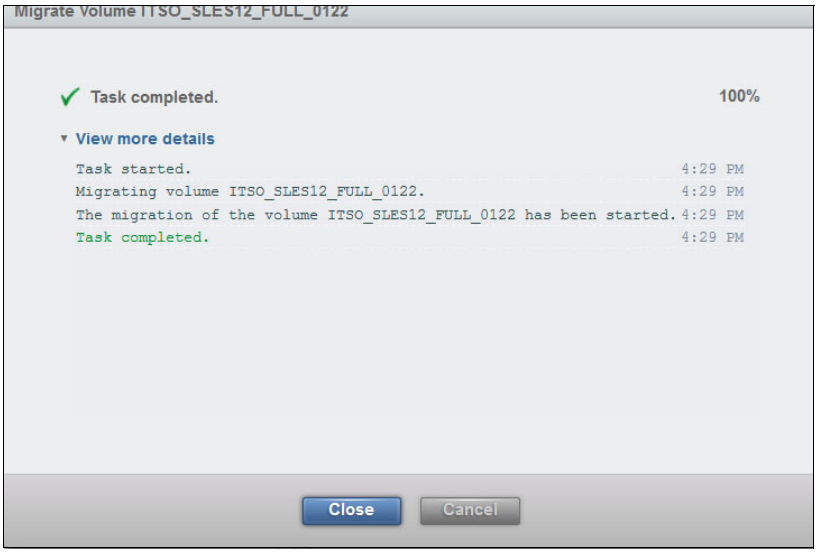



*Figure 4-15 Migration started and completed*

The volume is now visible in the target extent pool. The volume's status changed from Normal to Migrating, as shown in the Properties for Volumes window (Figure 4-16).

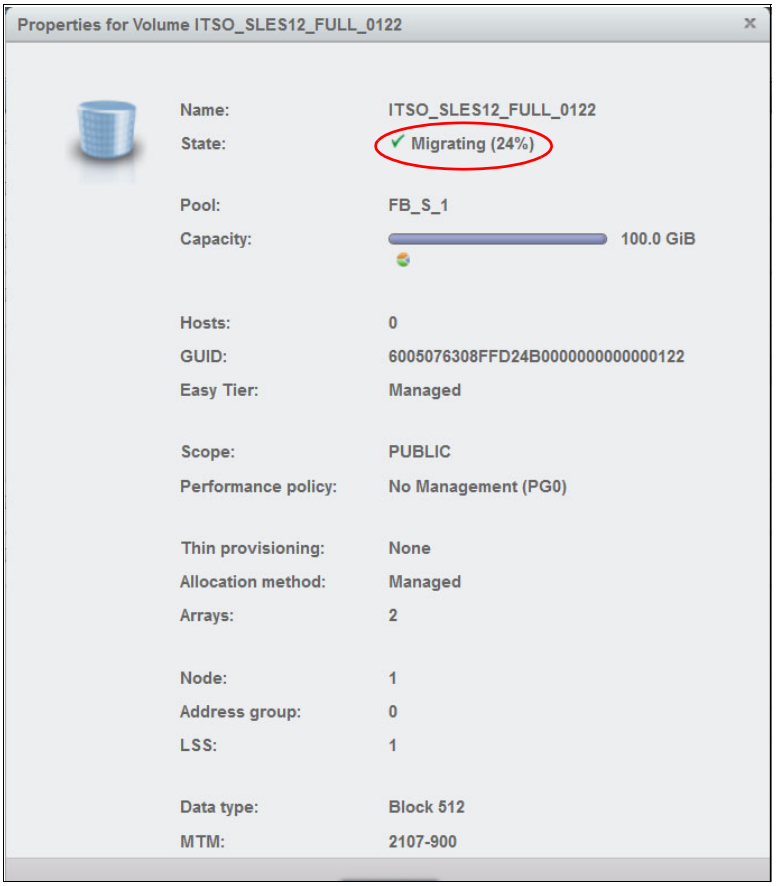



*Figure 4-16 Volume Properties window for the migrated volume*

The state (Migrating) is also reported in the **Volumes by Pool** window, as shown in Figure 4-17.

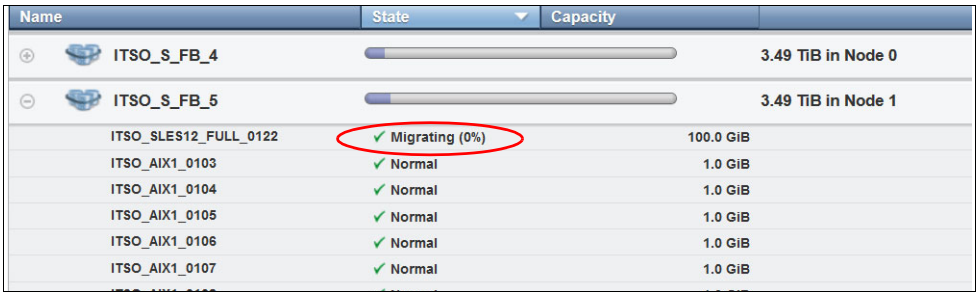



*Figure 4-17 Volume migration state inside the pool*

## Pausing and resuming volume migration

You can pause a migration from the DS GUI, or you can pause a migration only from the DS CLI. You can also resume a paused migration from the DS GUI or the DS CLI.

To pause a migration from the DS GUI, right-click the volume, and select **Migrate** → **Pause**. See Figure 4-18.

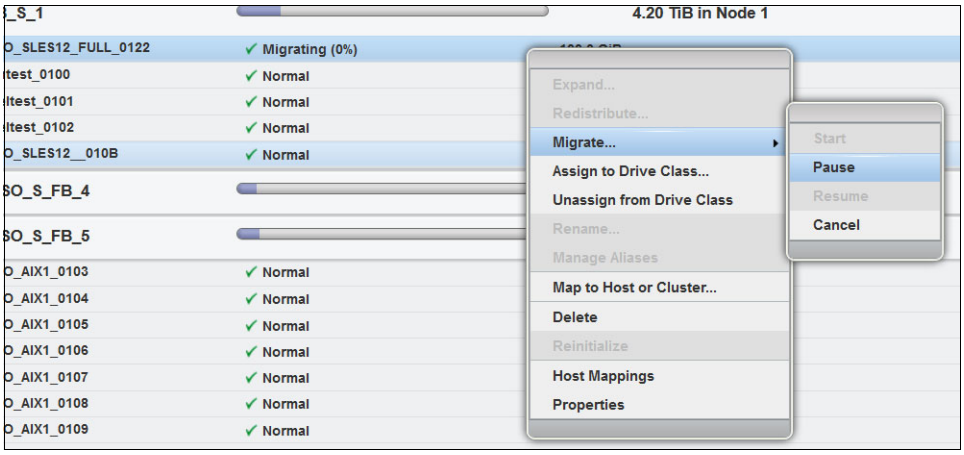



*Figure 4-18 Volume migration pause from DS GUI*

To pause a volume migration process, start a `dscli` session and run the **`managefbvol -action migpause <Volume_ID>`** command, as illustrated in Figure 4-19.

```
dscli> lsfbvol 0122
Date/Time: September 7, 2017 5:20:27 PM CEST IBM DSCLI Version: 7.8.30.470 DS: IBM.2107-75HGX91
Name             ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B) cap (10^9B) cap (blocks)
==========================================================================================================================
ITSO_SLES12_FULL 0122 Online   Normal    Migrating   2107-900  FB 512   P5           100.0           -    209715200
dscli> managefbvol -action migpause 0122
Date/Time: September 7, 2017 5:20:39 PM CEST IBM DSCLI Version: 7.8.30.470 DS: IBM.2107-75HGX91
CMUC00430I managefbvol: The migpause action for FB volume 0122 has completed.
dscli> lsfbvol 0122
Date/Time: September 7, 2017 5:20:51 PM CEST IBM DSCLI Version: 7.8.30.470 DS: IBM.2107-75HGX91
Name             ID   accstate datastate configstate      deviceMTM datatype extpool cap (2^30B) cap (10^9B) cap (blocks)
==========================================================================================================================
ITSO_SLES12_FULL 0122 Online   Normal    Migration Paused 2107-900  FB 512   P5           100.0           -    209715200
```

*Figure 4-19 Migration pause from the DS CLI*

The Migration paused state is also reported in the GUI, for instance, in the **Volumes by Pool** window, as shown in Figure 4-20.

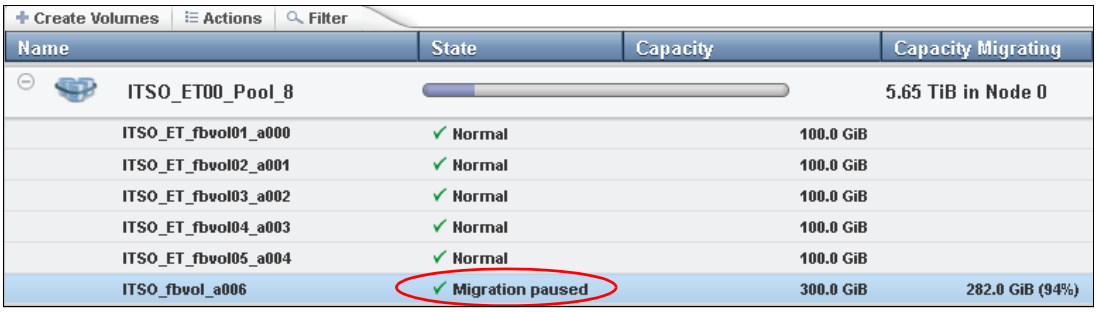



*Figure 4-20 Volume migration paused state inside the pool*

By using the GUI, you can resume the migration process by selecting the volume, right-clicking to open the pop-up menu, and clicking **Resume Migration**, as shown in Figure 4-21.

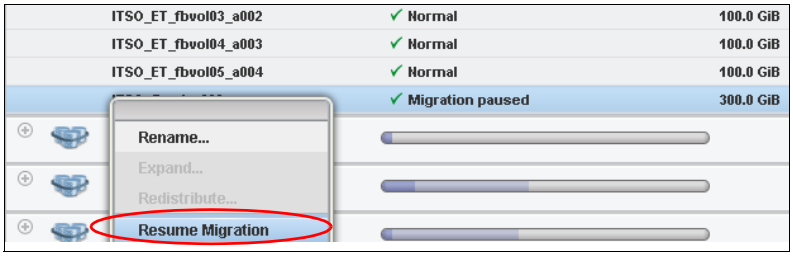



*Figure 4-21 Resume the volume migration*

### Canceling and resuming the volume migration

You cannot cancel a migration from the GUI. You can cancel a migration only from the DS CLI. However, you can resume a canceled migration from the GUI.

To cancel the volume migration process, start a `dscli` session and execute the **`managefbvol -action migcancel <Volume_ID>`** command.

To restart a canceled migration from the GUI, use the same procedure that is shown in “Pausing and resuming volume migration” on page 73.

## 4.1.9 Dynamic extent pool merge

The dynamic extent pool merge allows one extent pool to be merged with another extent pool transparently to the attached host systems. This capability is an Easy Tier Manual Mode capability. From the DS GUI, navigate to the **Array by Pools** or **Volumes by Pool** window to proceed with the merge process.

To start the dynamic extent pool merge, select the extent pools that you want to merge (at least two), right-click to display the pop-up menu, and select **Merge**, as shown in Figure 4-22 on page 75.

**Extent pool merge:** The extent pool merge can be performed only among extent pools with the same server affinity, and between extent pools with the same extent size.

If pools cannot merge due to an unsupported configuration, the DS GUI does not allow the merge process. The option to merge pools is not active.

If you are using the DS CLI, the command finishes with an error, for example:

```
CMUN02952E chextpool: P13: The Merge extent pools task cannot be initiated
because the source extent pool is not serviced by the same logical partition
(LPAR) as the target extent pool.
```

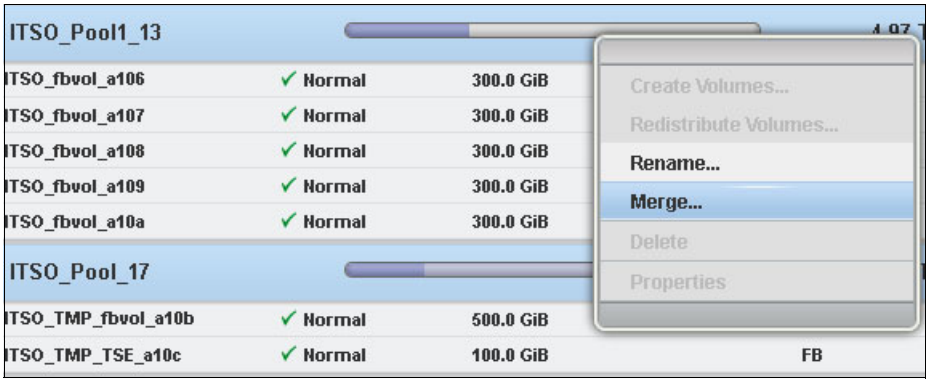



*Figure 4-22 Merge extent pools*

The Merge 2 Pools window that is shown in Figure 4-23 opens. You must assign a name for the merged extent pool.


*Figure 4-23 Merge extent pools characteristics*

When you merge single-tier extent pools, you can select **Redistribute all volumes** to start a volume redistribution as part of the merge operation. This option is not selectable when you merge multi-tier extent pools.

Depending on the choice, when the merge process completes, you get the messages that are shown in either Figure 4-24 or Figure 4-25.

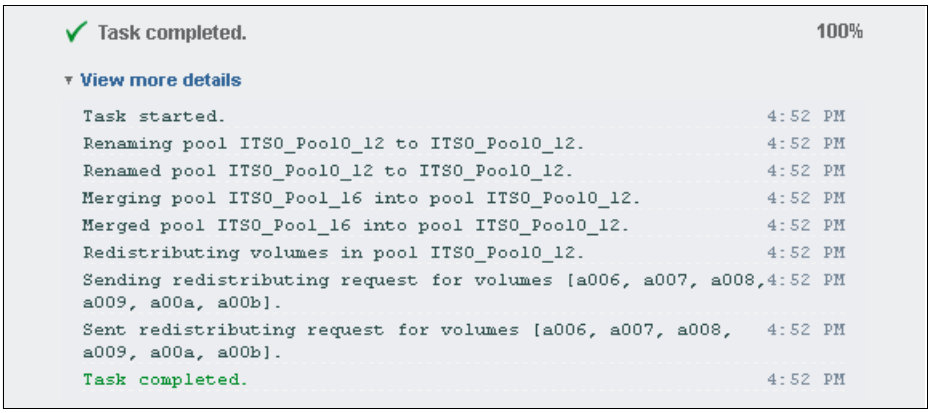



*Figure 4-24 Merge Extent Pools with redistribution*

Figure 4-25 shows the message without redistribution.

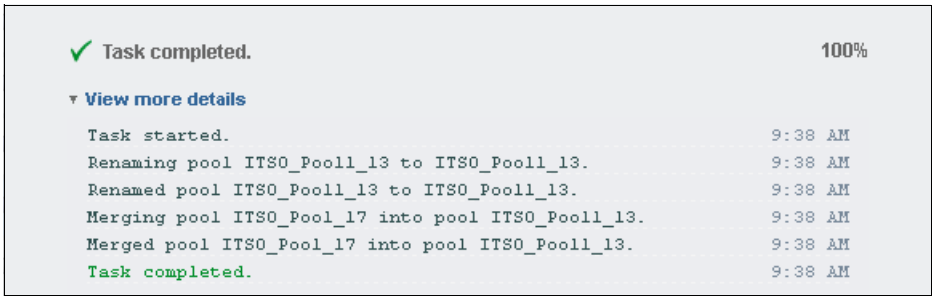



*Figure 4-25 Merge Extent Pools without redistribution*

**Important:** If one of the logical volumes is in the Migrating state, the merge process finishes with error `CMUN80221E` (Figure 4-26).

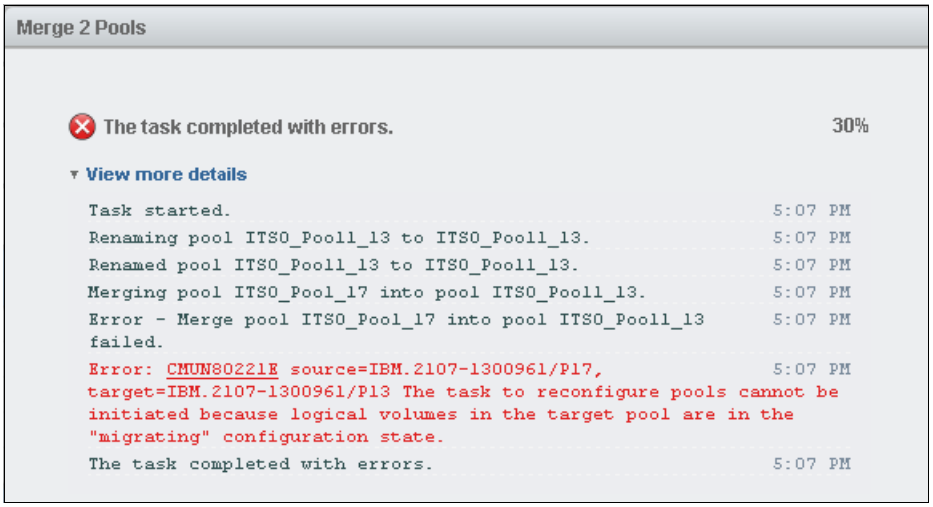



*Figure 4-26 Merge error for pool with migrating volumes*

### 4.1.10 Rank depopulation

Rank depopulation is an Easy Tier Manual Mode feature that allows a user to unassign one rank from an extent pool, even if extents are allocated on the rank. In this case, Easy Tier automatically reassigns all allocated extents to the other ranks in the extent pool. This operation is not apparent to host operations.

**Important:** A rank depopulation operation task fails if insufficient extents are available on the remaining ranks in the extent pool. For details, see “Rank depopulation” on page 25.

To perform the rank depopulation task from the DS GUI, click **Pools** → **Arrays by Pool**. To start rank depopulation, select the rank from the extent pools that you want to depopulate, right-click, and select **Unassign** from the pop-up menu, as shown in Figure 4-27.

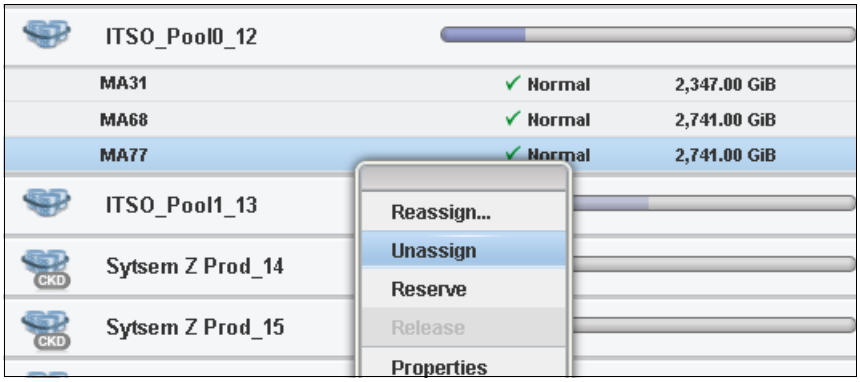



*Figure 4-27 Remove capacity from the pool*

Before depopulation starts, the GUI prompts you to confirm, as shown in Figure 4-28.

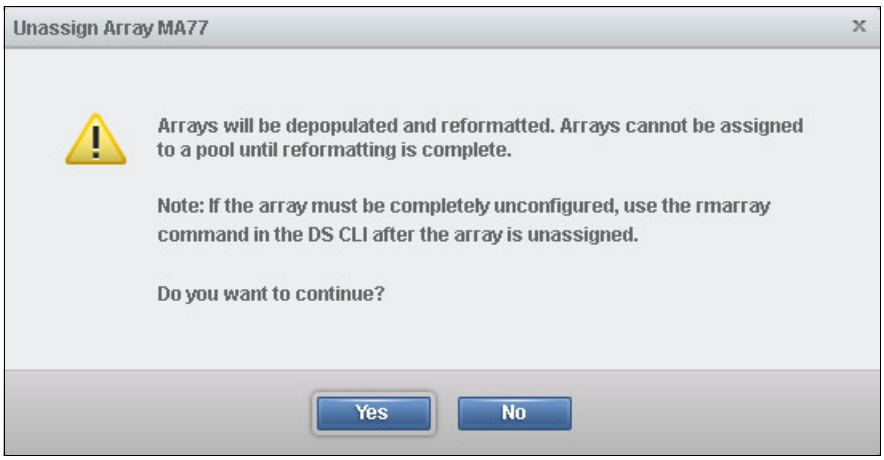



*Figure 4-28 Remove capacity from pool confirmation window*

Click **Yes** to start the depopulation (unassign) process. The action is confirmed, as shown in Figure 4-29 on page 78.

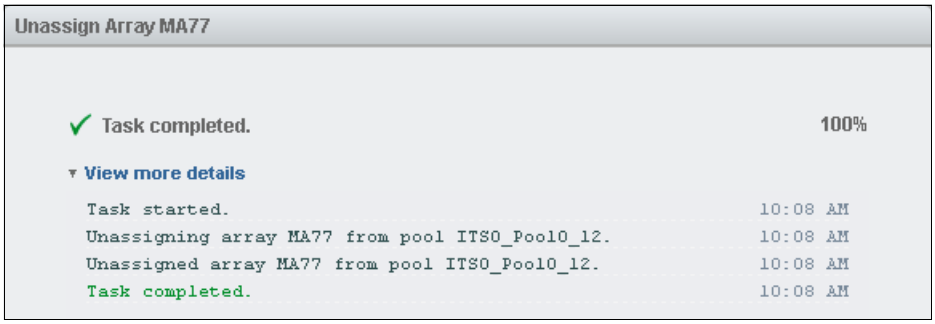



*Figure 4-29 Remove capacity from pool started and completed*

Easy Tier migrates allocated extents from the chosen rank to the remaining ranks in the pool and the rank configuration state changes to Depopulating, as illustrated in Figure 4-30 and Figure 4-31.

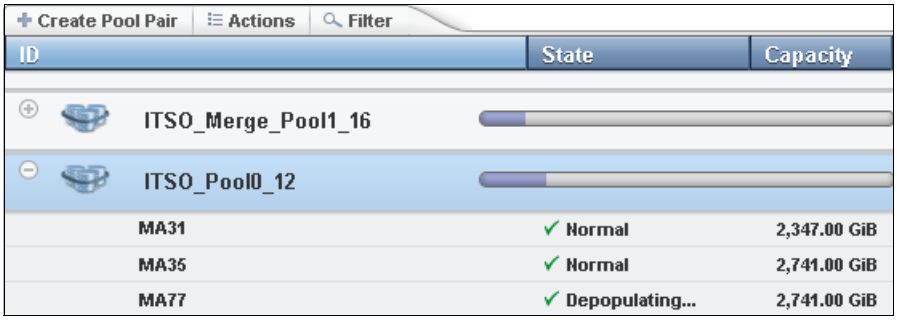



*Figure 4-30 Depopulating rank status*

Figure 4-31 shows the depopulated array.

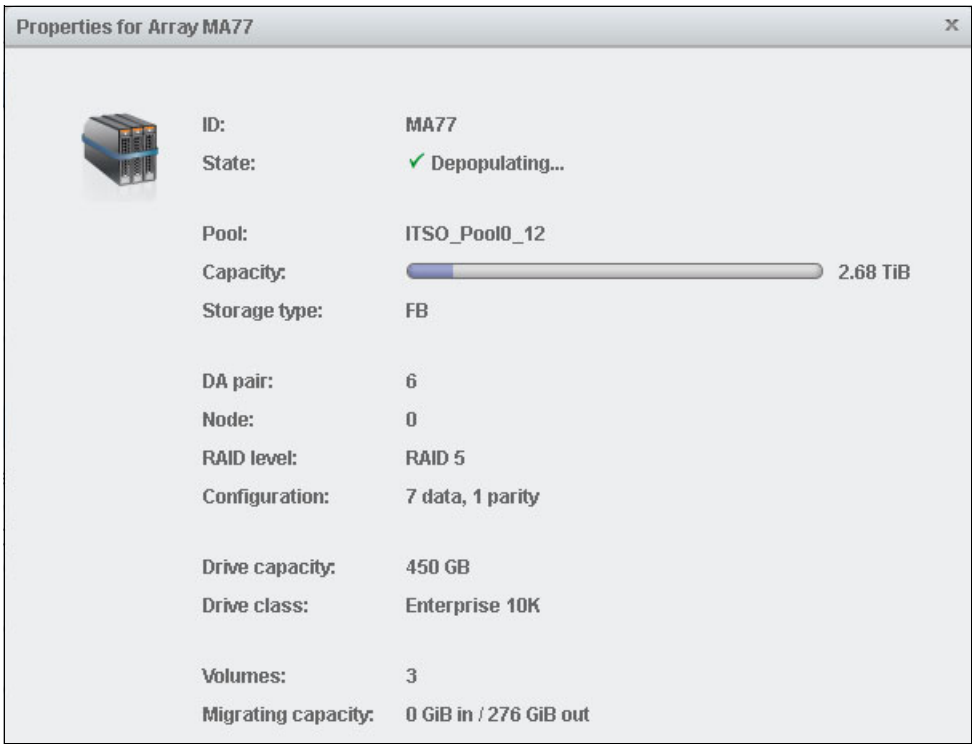



*Figure 4-31 Properties of depopulated array*

In the example, Easy Tier pre-allocated the extents in both remaining ranks and that 10 extents were migrated off the MA77 rank.

This operation can take time, depending on the number of extents to migrate and the overall activity in the extent pool. When all extents are migrated, the rank is unassigned from the extent pool.

## 4.2 DS CLI support for Easy Tier

The following DS CLI commands support Easy Tier features:

| | |
|---|---|
| **chsi** | Controls Easy Tier Automatic Mode and Easy Tier monitoring. The **chsi** command can also reserve some space for Easy Tier extent movements, and can be used to change the allocation order for new volumes. |
| **offloadfile** | Downloads Easy Tier monitoring data. |
| **manageckdvol** | Performs CKD volume migration actions (dynamic volume relocation), and monitoring-related activities. |
| **managefbvol** | Performs FB volume migration action (dynamic volume relocation), and monitoring-related activities. |
| **manageextpool** | Performs dynamic extent pool migration and monitoring-related activities. |
| **chextpool** | Performs dynamic extent pool merge. |
| **chrank** | Performs rank depopulation. |

The following commands show Easy Tier status information:

- **lsckdvol -l**
- **lsfbvol -l**
- **lsextpool -l**
- **lsrank**
- **showckdvol**
- **showfbvol**
- **showextpool**
- **showrank**
- **showsi**

**Command output:** Several of the command output lines in the following examples were truncated to make the examples more readable.

Use the **showsi** command shown in Example 4-1 to check the Easy Tier settings.

*Example 4-1 Checking for Easy Tier control settings (showsi)*

```
dscli> showsi
Date/Time: October 1, 2019 3:06:41 PM CEST IBM DSCLI Version: 7.9.0.491 DS: -
Name             -
desc             -
ID               IBM.2107-75ABCD1
Storage Unit     IBM.2107-75ABCD0
Model            994
WWNN             5005076301234567
Signature        ef12-3456-0123-4567
```

```
State             Online
ESSNet            Enabled
Volume Group      V0
os400Serial       373
NVS Memory        32.0 GB
Cache Memory      440.5 GB
Processor Memory  505.3 GB
MTS               IBM.5334-75ABCDO
numegsupported    16
ETAutoMode        tiered
ETMonitor         automode
IOPMmode          Disabled
ETCCMode          -
ETHMTMode         Enabled
ETSRMode          Enabled
ETTierOrder       High performance
ETAutoModeAccel   Disabled
```

The **chsi** or **showsi** commands change or show the control settings of Easy Tier Automatic Mode (`ETAutoMode=all, tiered`, or `none`) and Easy Tier monitoring (`ETMonitor=automode`, `all`, or `none`).

The ETAutomode specifies the automatic mode of the IBM Easy Tier feature. When the Easy Tier LIC feature is active, the manual mode is always active on all volumes, and the specified value controls only the automatic part of the feature. The following controls are available:

- **all**: The Easy Tier automatic mode is active on all pools.
- **tiered**: The Easy Tier automatic mode is active only on pools with multiple tiers.
- **none**: The Easy Tier automatic mode is not active on.

The ETMonitor specifies whether volumes are monitored by the IBM Easy Tier LIC feature. When the Easy Tier LIC feature is active, the volumes are monitored for the Easy Tier automatic mode. The following values are available:

- **all**: Easy Tier monitors all of the volumes on the DS8000 system, regardless of the state of the Easy Tier LIC feature. This capability demonstrates the potential benefits of Easy Tier. When this value is selected, all values of the **-etautomode** parameter are accepted, if the IBM Easy Tier LIC feature is active.
- **automode**: Easy Tier monitors only those volumes that are managed by the Easy Tier automatic mode as specified by the **-etautomode** parameter.
- **none**: Easy Tier monitors none of the volumes on the DS8000 system. When this value is selected, all values of the **-etautomode** parameter are ignored, and the Easy Tier automatic mode is not active on any pool.

Easy Tier monitoring can be activated on the storage facility image, independently of whether the LIC feature is installed or enabled, and turn on the usage for new Easy Tier features `ETHMTMode=Enable` or `Disable` for Easy Tier Heat Map Transfer Utility.

The **lsextpool -l** and **showextpool** commands show information about the number of tiers in an extent pool and whether the pool is managed by Easy Tier Automatic Mode. For more information, see 4.3.4, “Easy Tier Automatic Mode” on page 98.

Example 4-2 shows the **lsarray -l** command output.

*Example 4-2 Display the diskclass by using lsarray -l*

```
dscli> lsarray -l
Array State    Data   RAIDtype    arsite Rank DA Pair DDMcap (10^9B) diskclass  encrypt
=========================================================================================
A0    Assigned Normal 10 (4x2)    S22    R0   18              800.0 FlashTier0 supported
A1    Assigned Normal 6 (6+P+Q)   S5     R1   2              1200.0 ENT        supported
A2    Assigned Normal 10 (4x2)    S23    R2   18              800.0 FlashTier0 supported
A3    Assigned Normal 6 (6+P+Q)   S6     R3   2              1200.0 ENT        supported
A4    Assigned Normal 10 (3x2+S)  S24    R4   18              800.0 FlashTier0 supported
A5    Assigned Normal 6 (5+P+Q+S) S3     R5   2              1200.0 ENT        supported
A6    Assigned Normal 10 (3x2+S)  S25    R6   18              800.0 FlashTier0 supported
A7    Assigned Normal 6 (5+P+Q+S) S4     R7   2              1200.0 ENT        supported
A8    Assigned Normal 6 (5+P+Q+S) S21    R8   2              6000.0 NL         supported
A9    Assigned Normal 6 (6+P+Q)   S17    R9   2               600.0 ENT        supported
A10   Assigned Normal 5 (6+P+S)   S8     R10  2               600.0 ENT        supported
A11   Assigned Normal 6 (5+P+Q+S) S20    R11  2              6000.0 NL         supported
```

The *volume configuration state* of the volume commands includes appropriate states that reflect volume migration and extent pool merge operations:

- Merging: The volume's extent pool is merged into another pool.
- Migrating: The volume is in the process of migrating.
- Migration Paused: The volume migration is paused.
- Migration Canceled: The volume migration was canceled.
- Migration Error: An error occurred during volume migration.
- Transposition Error: An error occurred while an extent pool was merged.

The rank configuration state includes these appropriate states related to rank depopulation:

- Depopulating: The rank is in the process of depopulating so that it can be unassigned from its extent pool.
- Depopulation Error: An error occurred during rank depopulation.
- Unassigned Reserved: A rank is not assigned to any extent pools and it is reserved.

By using the Easy Tier Application, the volume configuration state of the **volume** commands has the following states that reflect volume reassignment:

- **tierassignstatus**: Indicates the Assign/Unassign status of the volume.
- **tierassignerror**: Helps to identify the possible reason if **tierassignstatus** shows `Error`.
- **tierassignorder**: Shows the data migration type that is specified in the **Assign** command that is issued: `ETdata` or `Access`.
- **tierassigntarget**: Shows the target tier that is specified in the **Assign** command that is issued: `Flash`, `ENT`, or `NL`.
- **%tierassigned**: Shows the percentage of the volume that is assigned.

If you specify the **-tier** parameter, a tier distribution table is appended to the end of the display. It shows the tier assigned percentage allocation among all tiers (flash, enterprise, and nearline) of the volume.

## 4.3 Easy Tier functions by using the DS CLI

This section provides examples of the Easy Tier capabilities by using the DS CLI. For more information, see *IBM System Storage DS Command-Line Interface User's Guide for DS8000 Series,* SC27-8526, or run the **help** command.

### 4.3.1 Change storage system properties for Easy Tier

To function properly, Easy Tier needs some free extents in each extent pool to be able to shift around extents. By default, the system reserves some capacity automatically.

To disable or re-enable the DS8000 to reserve storage for Easy Tier extent movements, use the **chsi** command with the **-etsrmode disable or enable** option, as shown in Example 4-3.

*Example 4-3 Let the system not reserve storage for Easy Tier*

```
dscli> chsi -etsrmode disable
Date/Time: 2019-10-01T14:21:53+0200 IBM DSCLI Version: 7.9.0.491 DS: -
CMUC00042I chsi: Storage image IBM.2107-75DST11 successfully modified.
```

You can change the allocation order of new volumes with the **-ETTierOrder** option of the **chsi** command. You can change the order from the default **highperf** to **highutil** to favor allocation of cheaper lower-tier flash, as shown in Example 4-4.

*Example 4-4 Set the new volume allocation policy*

```
dscli> chsi -ettierorder highutil
Date/Time: 2019-10-01T14:36:58+0200 IBM DSCLI Version: 7.9.0.491 DS: -
CMUC00042I chsi: Storage image IBM.2107-75DST11 successfully modified.
```

To accelerate the Easy Tier automatic migration speed, you can enable the **-ETAutoModeAccel** option, as in Example 4-5, for the DS8900F model. This should be done in one-off situations only, not for usual steady-state operation.

*Example 4-5 Setting Easy Tier migrations to high speed*

```
dscli> chsi -etautomodeaccel enable
Date/Time: 2019-10-01T15:14:42+0200 IBM DSCLI Version: 7.9.0.491 DS: -
CMUC00042I chsi: Storage image IBM.2107-75DST11 successfully modified.
```

### 4.3.2 Dynamic volume relocation

Dynamic volume relocation allows a logical volume to be migrated from one extent pool to another pool (or even the same extent pool by using, for example, manual volume rebalance) transparently to host operations. This capability is an Easy Tier Manual Mode capability. You can perform the following actions:

- Initiate a migration: Start a volume relocation in the same or another extent pool. The volume's extent allocation method (EAM) can be changed as part of the migration.
- Pause or resume a migration: Pause and resume an ongoing volume relocation. You might want to suspend a volume relocation temporarily to avoid more workload on the back-end storage resources.
- Cancel a migration: Cancel an ongoing volume relocation. You might want to cancel a volume relocation because the volume migration is no longer needed. After a volume relocation cancellation, the volume might end up in a transient state (migration canceled) at the time that the volume extents are spread between two extent pools. This situation must be corrected by starting (and completing) a new volume relocation.

Similarly to the volumes themselves, also the Safeguarded Copy Backup Capacity of certain specified volumes can be migrated between pools, when using the Safeguarded Copy function. Respective options for the **manageckdvol** and **managefbvol** commands are available. For information on Safeguarded Copy, refer to *IBM DS8880 Safeguarded Copy,* REDP-5506.

**Dynamic volume relocation considerations:** Dynamic volume relocation is allowed only among extent pools of the same volume storage type (CKD or FB) and of the same server affinity, and of the same extent size. For example, you can migrate volumes between extent pools P0 and P2 (rank group 0 managed by storage server 0) or P1 and P3 (rank group 1 managed by storage server 1) but not between P0 and P1.

Migration to the same extent pool or manual volume rebalance is not supported for any managed or hybrid (multi-tier) extent pool, even if the hybrid extent pool is not managed by Easy Tier Automatic Mode. For more information, see “Dynamic volume relocation” on page 22 and “Manual volume rebalance” on page 23.

Dynamic volume relocation is performed by using the **managefbvol** command for FB volumes and the **manageckdvol** command for CKD volumes. The command parameters are the same for FB and CKD environments. The following syntax is used:

```
managefbvol -action [ migstart | migcancel | migresume | migpause ]
-eam [  rotatevols | rotateexts ] -extpool TgtPoolID VolID
```

When a volume reassignment to other tiers is performed manually by using Easy Tier Application, the following syntax is used:

```
managefbvol -action tierassign [-assignorder etdata|access] [-tier
flashtier0|flashtier1|flashtier2|ent|nl|nlexclude|hddexclude] <volume ID>
```

In a volume reassignment to use only flash or Enterprise tiers but not a Nearline tier, the following syntax is used:

```
managefbvol -action tierassign [-assignorder etdata | access] -tier nlexclude
<volume ID>
```

One of the following actions can be used:

| | |
|---|---|
| **migstart** | Starts a volume migration. |
| **migpause** | Pauses a volume migration. |
| **migresume** | Resumes a paused volume migration. |
| **migcancel** | Cancels a volume migration. |
| **tierassign** | Reassigns a volume for a different tier. |
| **tierunassign** | Releases a previous volume assignment. |

For more information about how to use the **managefbvol** and **manageckdvol** commands in an Easy Tier Application environment, see *DS8870 Easy Tier Application*, REDP-5014.

The **eam** parameter is optional. It can be used to change the volume’s EAM. It defaults to the volume’s current EAM. The **extpool** parameter is optional and defaults to the volume’s current extent pool.

The target extent pool must belong to the same rank group (0/1) as the volume’s current extent pool, which means that the source and target extent pools must be managed by the same DS8000. For example, a volume of LSS 10 that is in Extent Pool P0 (rank group 0), which is managed by storage server 0, cannot be migrated to Extent Pool P1 (rank group 1), which is managed by storage server 1.

## Volume migration

Example 4-6 shows volume migration from one extent pool to another by using dynamic volume relocation. It also demonstrates the cancel operation. The source extent pool P0 contains three ranks and only the one volume to be migrated (volume ID 1000). The target extent pool P2 contains two ranks, and it is empty. This setup makes it easy to see where the volume extents are allocated.

*Example 4-6 Manual volume migration to another extent pool*

```
dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==========================================================================================
ITSO_FB_A P0 fb            0  below             4734          0      4734        0       1
ITSO_FB_B P2 fb            0  below             1558          0      1558        0       0
ITSO_FB_C P4 fb            0  below              388          0       388        0       0

dscli> lsrank -extpool p0 -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R2     0 Normal Normal    A2           5 P0        ITSO_FB_A   fb      1582        4
R3     0 Normal Normal    A3           5 P0        ITSO_FB_A   fb      1582        4
R4     0 Normal Normal    A4           5 P0        ITSO_FB_A   fb      1582        4

dscli> lsrank -extpool p2 -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R0     0 Normal Normal    A0           5 P2        ITSO_FB_B  fb      779         0
R1     0 Normal Normal    A1           5 P2        ITSO_FB_B  fb      779         0

dscli> managefbvol -action migstart -extpool p2 1000
CMUC00430I managefbvol: The migstart action for FB volume 1000 has completed.

dscli> lsfbvol -config not_normal
Name   ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B)
==============================================================================
ITSO_J 1000 Online   Normal    Migrating   2107-900  FB 512   P0             12.0

dscli> managefbvol -action migcancel 1000
CMUC00430I managefbvol: The migcancel action for FB volume 1000 has completed.

dscli> lsfbvol -config not_normal
Name   ID   accstate datastate configstate         deviceMTM datatype extpool cap (2^30B)
==========================================================================================
ITSO_J 1000 Online   Normal    Migration Cancelled 2107-900  FB 512   P2             12.0

dscli> lsrank -extpool p0 -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R2     0 Normal Normal    A2           5 P0        ITSO_FB_A  fb      1582        2
R3     0 Normal Normal    A3           5 P0        ITSO_FB_A  fb      1582        2
R4     0 Normal Normal    A4           5 P0        ITSO_FB_A  fb      1582        1

dscli> lsrank -extpool p2 -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R0     0 Normal Normal    A0           5 P2        ITSO_FB_B  fb      779         6
R1     0 Normal Normal    A1           5 P2        ITSO_FB_B  fb      779         1

dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
```

```
==========================================================================================
ITSO_FB_A P0 fb            0  below             4741          0      4741        0       0
ITSO_FB_B P2 fb            0  below             1551          0      1551        0       1
ITSO_FB_C P4 fb            0  below              388          0       388        0       0

dscli> managefbvol -action migstart -extpool p4 1000
CMUN01131E managefbvol: 1000: The Migrate volume task cannot be initiated because the
volume that you have selected has a volume configuration state of Migration_cancelled, and
therefore the volume can only be migrated to the source or target extent pool of the
cancelled migration task.

dscli> managefbvol -action migstart -extpool p2 1000
CMUC00430I managefbvol: The migstart action for FB volume 1000 has completed.

dscli> lsfbvol -config not_normal
CMUC00234I lsfbvol: No FB Volume found.

dscli> lsfbvol 1000
Name   ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B)
==========================================================================================
ITSO_J 1000 Online   Normal    Normal      2107-900  FB 512   P2             12.0

dscli> lsrank -extpool p2 -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R0      0 Normal Normal    A0           5 P2        ITSO_FB_B  fb      779         6
R1      0 Normal Normal    A1           5 P2        ITSO_FB_B  fb      779         6
```

Complete the following steps to migrate the volume:

1. Check the initial extent pool status by running the **lsextpool** and **lsrank** commands.
2. Start the migration by running **managefbvol -action migstart**. The **lsfbvol -config not_normal** command can be used to quickly list all volumes in a configuration state other than Normal, such as volumes in a Migrating state.
3. Suspend the migration by running **managefbvol -action migpause**. The **lsfbvol** command shows that the volume is in a Migration Paused state. The **lsrank** commands show that the volume's extents are now in both extent pools. You also can verify this situation by running **showfbvol -rank A00A** instead. The **lsextpool** command shows that the volume is considered in the target extent pool even though it has extents in both pools.
4. Restart migration to the original target pool P2 by running **managefbvol -action migstart**. Eventually, the migration completes and the volume returns to a Normal state.
5. Cancel the migration by running **managefbvol -action migcancel**. The **lsfbvol** command shows that the volume is in a Migration Canceled state. The **lsrank** commands show that the volume's extents are now in both extent pools. You also can verify this situation by running **showfbvol -rank A00A** instead. The **lsextpool** command shows that the volume is considered in the target extent pool even though it has extents in both pools.
6. Restart the migration to the original target pool by running **managefbvol -action migstart**. Eventually, the migration completes and the volume returns to a Normal state. The last **lsrank** command shows that all extents are now on the two ranks of P2.

**Volume migration:** If the same extent pool is used as the source and target pool of a volume migration (for example, with manual volume rebalance), the migration cancel action changes the state from Migrating to Normal instead of Migration Canceled. All volume extents still belong to only one extent pool.

## Manual volume rebalance

Manual volume rebalance is designed to redistribute the extents of volumes within a non-managed, single-tier (homogeneous) pool so that workload skew and hotspots are less likely to occur on the ranks. Manual volume rebalance is an enhancement of the dynamic volume relocation feature.

This feature is useful when a volume is migrated back into the same extent pool with the source extent pool the same as the target extent pool and the EAM is rotate extents (which means the use of storage pool striping). In this case, the algorithm tries to spread the volume's extents evenly across all ranks in the extent pool.

Use manual volume rebalance when a rank is added to a pool, homogeneous extent pools were merged, or when large volumes with an EAM of rotate volumes are deleted and a rebalance can optimize a non-balanced extent distribution within a homogeneous pool. For more information, see "Manual volume rebalance" on page 23.

With manual volume rebalance, only one extent at a time is allocated (rather than preallocating the full volume), so only the smallest number of required extents is relocated. Only a small amount of free capacity needs to be available on the ranks to achieve a balanced capacity distribution across all ranks.

If a manual volume rebalance operation is run in an existing pool with only partial capacity available on certain ranks and no capacity available on other ranks, the distribution of the volume capacity across all ranks might not always be balanced, for example, if one rank is fully allocated by another volume by using an EAM of rotate volumes.

In this case, you might need to perform the relocation in a certain sequence or in multiple steps to ensure that at least some free extents are available on each rank in the pool.

Example 4-7 shows the migration of nine FB volumes back into the same extent pool (P6). This migration shows the benefits of manual volume rebalance. For newly created volumes, the default EAM is rotate extents. The example also demonstrates the use of migration pause and resume.

*Example 4-7 Extent pool capacity distribution before adding new capacity and starting the manual volume rebalance*

```
dscli> lsextpool -l p6
ID stgtype rankgrp status availstor (2^30B) %allocated available numvols numranks numtiers etmanaged
=============================================================================================================
P6 fb            0   full                 0        100         0       9        2        1 no

dscli> lsfbvol -l -extpool p6
ID   accstate datastate configstate deviceMTM datatype extpool sam      captype cap (2^30B) eam
=============================================================================================================
4000 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4001 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4002 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4003 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4004 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4005 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4006 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4007 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4008 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
```

Complete the following steps to manually rebalance the volume:

1. Check the initial extent pool and volume status by running the **lsextpool -l** and **lsfbvol -l** commands. There are nine volumes (4000 - 4008), and each volume uses the rotate extents EAM. The extent pool (P6) has two ranks and no free extents available.
2. Add an empty rank to the extent pool (P6). The **lsrank -l** command shows that a new empty rank was added to the extent pool, as shown in Example 4-8.

*Example 4-8 Add a rank to the extent pool and check the extent allocation across ranks before manual volume rebalance*

```
dscli> chrank -extpool p6 r13
CMUC00008I chrank: Rank R13 successfully modified.

dscli> lsrank -l -extpool p6
ID  Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts encryptgrp
==========================================================================================
R8      0 Normal Normal    A10          5 P6        tier1      fb      900       900          -
R10     0 Normal Normal    A18          5 P6        tier1      fb      900       900          -
R13     0 Normal Normal    A21          5 P6        tier1      fb      900         0          -
```

3. Check that the extent pool (P6) is not managed by Easy Tier Automatic Mode by running the **lsextpool -l** command, and start manual volume rebalance by running the **managefbvol -action migstart** command. Because a target extent pool is not specified on the command, the target is the same as the current extent pool. The **lsfbvol -l** command shows that the volumes' configstate is Migrating.

   The **lsrank -l** command shows that more extents were allocated on the ranks to facilitate the volume rebalance, as shown in Example 4-9.

*Example 4-9 Manual volume rebalance action is started*

```
dscli> lsextpool -l p6
ID stgtype rankgrp status availstor (2^30B) %allocated available  numvols numranks numtiers etmanaged
=========================================================================================================
P6 fb            0  below               900         67       900        9        3        1 no

dscli> managefbvol -action migstart 4000-4008
CMUC00430I managefbvol: The migstart action for FB volume 4000 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4001 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4002 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4003 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4004 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4005 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4006 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4007 has completed.
CMUC00430I managefbvol: The migstart action for FB volume 4008 has completed.

dscli> lsfbvol -l -extpool p6 -config not_normal
ID   accstate datastate configstate deviceMTM datatype extpool sam      captype cap (2^30B) eam
=============================================================================================================
4000 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4001 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4002 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4003 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4004 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4005 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4006 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4007 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4008 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts

dscli> lsrank -l -extpool p6
ID  Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts encryptgrp
```

```
=============================================================================================
R8      0 Normal Normal    A10          5 P6        tier1 fb      900       894          -
R10     0 Normal Normal    A18          5 P6        tier1 fb      900       893          -
R13     0 Normal Normal    A21          5 P6        tier1 fb      900        14          -
```

4. Pause the volume rebalance by running the **managefbvol -action migpause** command. The **lsfbvol** command shows that the volumes are in a Migration Paused state, as shown in Example 4-10.

*Example 4-10 Manual volume rebalance action is paused*

```
dscli> managefbvol -action migpause 4000-4008
CMUC00430I managefbvol: The migpause action for FB volume 4000 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4001 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4002 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4003 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4004 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4005 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4006 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4007 has completed.
CMUC00430I managefbvol: The migpause action for FB volume 4008 has completed.

dscli> lsfbvol -l -extpool p6
ID   accstate datastate configstate      deviceMTM datatype extpool sam      captype cap (2^30B) eam
=============================================================================================================
4000 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4001 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4002 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4003 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4004 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4005 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4006 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4007 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4008 Online   Normal    Migration Paused 2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
```

5. Resume the volume rebalance by running the **managefbvol -action migresume** command. The **lsfbvol -l** command shows that the volumes' configstate is Migrating, as shown in Example 4-11.

*Example 4-11 Manual volume rebalance action is resumed*

```
dscli> managefbvol -action migresume 4000-4008
CMUC00430I managefbvol: The migresume action for FB volume 4000 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4001 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4002 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4003 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4004 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4005 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4006 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4007 has completed.
CMUC00430I managefbvol: The migresume action for FB volume 4008 has completed.

dscli> lsfbvol -l -extpool p6
ID   accstate datastate configstate deviceMTM datatype extpool sam      captype cap (2^30B) eam
=============================================================================================================
4000 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4001 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4002 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4003 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4004 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4005 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
```

```
4006 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4007 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4008 Online   Normal    Migrating   2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
```

6. The **lsfbvol -l** command shows that the configstate is Normal and no longer Migrating. The **lsrank** command shows that the extents now are allocated evenly on all the ranks, as shown in Example 4-12.

*Example 4-12 Extent pool capacity distribution after adding new capacity and initiating manual volume rebalance*

```
dscli> lsfbvol -l -extpool p6
ID   accstate datastate configstate deviceMTM datatype extpool sam      captype cap (2^30B) eam
==========================================================================================================
4000 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4001 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4002 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4003 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4004 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4005 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4006 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4007 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts
4008 Online   Normal    Normal      2107-900  FB 512   P6      Standard      DS       200.0 rotateexts

dscli> lsrank -l -extpool p6
ID  Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts encryptgrp
==========================================================================================
R8      0 Normal Normal    A10          5 P6        tier1 fb      900       600          -
R10     0 Normal Normal    A18          5 P6        tier1 fb      900       600          -
R13     0 Normal Normal    A21          5 P6        tier1 fb      900       600          -
```

The **showfbvol**, **showckdvol**, **showextpool**, and **showrank** commands show volume migration and rank allocation-related information, as shown in Example 4-13.

*Example 4-13 Volume status information that relates to volume migration*

```
dscli> showfbvol -rank 5200
Name                 Broker_4
ID                   5200
accstate             Online
datastate            Normal
configstate          Migrating
deviceMTM            2107-900
datatype             FB 512
addrgrp              5
extpool              P10
exts                 325
cap (MiB)            332800
captype              DS
cap (2^30B)          325.0
cap (10^9B)          -
cap (blocks)         681574400
volgrp               V3
ranks                4
dbexts               0
sam                  Standard
repcapalloc          -
eam                  managed
reqcap (blocks)      681574400
realextents          325
virtualextents       34
```

```
realcap (MiB)              332800
migrating                  310
migratingcap (MiB)         317440
perfgrp                    PG0
migratingfrom              P2
resgrp                     RG0
tierassignstatus           Unknown
tierassignerror            -
tierassignorder            Unknown
tierassigntarget           Unknown
%tierassigned              0
etmonpauseremain           -
etmonitorreset             unknown
[...]
==============Rank extents==============
rank extents capacity (MiB/cyl) metadata
========================================
R4       103 105472             no
R6       103 105472             no
R7       103 105472             no
R8        16 16384              no

dscli> showextpool p10
Name                   HPFE_0
ID                     P10
stgtype                fb
totlstor (2^30B)       2132
availstor (2^30B)      577
resvdstor (2^30B)      1
rankgrp                0
numranks               1
numvols                19
status                 below
%allocated             72
%available             27
configured             2132
allowed                1995
available              577
allocated              1418
reserved               1
configuredCap(MiB/cyl) 2183168
allowedCap(MiB/cyl)    2042880
availableCap(MiB/cyl)  590848
allocatedCap(MiB/cyl)  1452032
reservedCap(MiB/cyl)   1024
%limit                 100
%threshold             15
virextstatus           below
%virallocated          37
%viravailable          62
virconfigured          4300
virallowed             3960
viravailable           2490
virallocated           1470
virreserved            0
%virextlimit           -
%virextthreshold       -
keygrp                 -
opratio                0.73
opratiolimit           -
```

```
%allocated(ese)        0
%allocated(rep)        0
%allocated(std)        66
%allocated(over)       0
%virallocated(ese)     0
%virallocated(tse)     0
%virallocated(init)    0
%migrating(in)         7
%migrating(out)        0
numtiers               1
etmanaged              yes
etmigpauseremain       -
etmonpauseremain       -
etmonitorreset         unknown
extsize                1GiB
```

```
dscli> showrank r8
ID             R8
SN             KAFAW8563A1C92K
Group          0
State          Normal
datastate      Normal
Array          A8
RAIDtype       5
extpoolID      P10
extpoolnam     HPFE_0
volumes        000A,000B,0012,0013,0014,0015,0018,0019,001A,001B,001D,001E,0036,0037,501B,5050,5200,6403,6405
stgtype        fb
exts           2132
usedexts       1525
widearrays     0
nararrays      1
trksize        128
strpsize       512
strpesize      0
extsize        16384
keygrp         -
migrating(in)  0
migrating(out) 296
marray         MA14
extsize (cap)  1GiB
```

## 4.3.3 Dynamic extent pool merge

Dynamic extent pool merge allows one extent pool (the source extent pool) to be merged into another extent pool (the target extent pool) transparently to the attached host systems. This capability is an Easy Tier Manual Mode capability.

The operation is performed by running the **chextpool** command, as shown in the following syntax example:

```
chextpool -merge <sourcepool_ID> -quiet <targetpool_ID>
```

Both source and target extent pools can contain regular or thin-provisioned (ESE) logical volumes. An extent pool merge is not supported if they have different extent sizes.

**Situations in which you cannot merge extent pools:** You cannot merge two extent pools if any of the following conditions exist:

- The extent pools are of different type (FB and CKD).
- Both extent pools have different server affinity.
- Both extent pools have different extent sizes.
- You selected an extent pool that contains volumes that are being migrated.

When the merge is complete, the ranks and volumes that are assigned to the source pool are reassigned to the target pool, and the source extent pool is removed. The optional **-quiet** parameter disables the confirmation prompt for the **-merge** parameter.

Example 4-14 shows a dynamic extent pool merge. The source extent pool P2 contains two ranks and one logical volume (1000). The target extent pool P0 contains three ranks and two logical volumes (1001-1002).

*Example 4-14 Dynamic extent pool merging*

```
dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==========================================================================================
ITSO_FB_A P0 fb            0  below              4716          0      4716        0       2
ITSO_FB_B P2 fb            0  below              1546          0      1546        0       1

dscli> lsrank -l -extpool p0
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R2     0 Normal Normal    A2           5 P0        ITSO_FB_A  fb      1582       10
R3     0 Normal Normal    A3           5 P0        ITSO_FB_A  fb      1582       10
R4     0 Normal Normal    A4           5 P0        ITSO_FB_A  fb      1582       10

dscli> lsrank -l -extpool p2
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R0     0 Normal Normal    A0           5 P2        ITSO_FB_B  fb      779         6
R1     0 Normal Normal    A1           5 P2        ITSO_FB_B  fb      779         6

dscli> lsfbvol 1000-1002
Name   ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B)
==========================================================================================
ITSO_J 1000 Online   Normal    Normal      2107-900  FB 512   P2             12.0
ITSO_J 1001 Online   Normal    Normal      2107-900  FB 512   P0             15.0
ITSO_J 1002 Online   Normal    Normal      2107-900  FB 512   P0             15.0

dscli> chextpool -merge p2 p0
CMUC00422W chextpool: Are you sure that you want to merge source extent pool p2 into target
extent pool p0? [y/n]: y
CMUC00001I chextpool: Extent pool P0 successfully modified.

dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==========================================================================================
ITSO_FB_A P0 fb            0  below              6262          0      6262        0       3

dscli> lsrank -l -extpool p0
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R0     0 Normal Normal    A0           5 P0        ITSO_FB_A  fb      779         6
R1     0 Normal Normal    A1           5 P0        ITSO_FB_A  fb      779         6
```

```
R2      0 Normal Normal    A2           5 P0        ITSO_FB_A  fb      1582       10
R3      0 Normal Normal    A3           5 P0        ITSO_FB_A  fb      1582       10
R4      0 Normal Normal    A4           5 P0        ITSO_FB_A  fb      1582       10

dscli> lsfbvol 1000-1002
Name   ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B)
==========================================================================================
ITSO_J 1000 Online   Normal    Normal      2107-900  FB 512   P0             12.0
ITSO_J 1001 Online   Normal    Normal      2107-900  FB 512   P0             15.0
ITSO_J 1002 Online   Normal    Normal      2107-900  FB 512   P0             15.0
```

Complete the following steps to merge the dynamic extent pool:

1. Check the initial status by running the **lsextpool**, **lsrank**, and **lsfbvol** commands.
2. Merge the extent pool by running the **chextpool -merge p2 p0** command.
3. After the merge completes, check the final status. The source extent pool P2 is removed and all five ranks are now assigned to extent pool P0. Volume extents are on the same ranks as before.

**Merging extent pools:** After you merge extent pools, you might want to redistribute volume extents on the merged extent pool by using the manual volume rebalance function, as described in “Manual volume rebalance” on page 86, especially if the extents are unevenly distributed on the ranks. Enough free space must exist on the ranks.

You can use rank depopulation to unassign a rank from an extent pool, even if extents are allocated on that rank. In that case, Easy Tier automatically reassigns all allocated extents to the other ranks within the same extent pool. This operation is not apparent to the host operations. This operation is available for all types of ranks.

To perform the operation, run the **chrank** command, as shown in the following example:

```
chrank -unassigned -quiet <rank_ID>
```

**Extent allocation:** If no extents are allocated on the specified rank in the extent pool, the rank is directly unassigned from the pool.

Ranks to be depopulated can contain extents from regular and thin-provisioned logical TSE or ESE volumes. When the depopulation is complete, these extents are migrated to the other ranks in the extent pool and the source extents are removed. The optional **-quiet** parameter disables the confirmation prompt for the **-unassign** parameter.

**Important:** A rank depopulation operation is not allowed if insufficient extents are available on the other remaining ranks of the extent pool. However, a rank depopulation operation is supported by regular volumes, and ESE volumes in the pool. For more information, see “Rank depopulation” on page 25.

Example 4-15 shows a rank depopulation operation. The hybrid extent pool P4 contains 12 ranks (including R18, which is a flash rank) and several logical volumes, including 4008, which is spread across the non-flash ranks in the pool.

*Example 4-15 Rank and volume status before rank depopulation*

```
dscli> lsfbvol -l -extpool p4
ID   accstate datastate configstate deviceMTM datatype extpool sam      captype cap (2^30B) eam
=============================================================================================
1000 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        25.0 managed
4000 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        10.0 managed
4001 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        10.0 managed
4002 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS       100.0 managed
4003 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        50.0 managed
4004 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        80.0 managed
4005 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS       150.0 managed
4006 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        20.0 managed
4007 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS        25.0 managed
4008 Online   Normal    Normal      2107-900  FB 512   P4      Standard      DS      2400.0 managed

dscli> showfbvol -rank 4008
Name             -
ID               4008
accstate         Online
datastate        Normal
configstate      Normal
deviceMTM        2107-900
datatype         FB 512
addrgrp          4
extpool          P4
exts             2400
captype          DS
cap (2^30B)      2400.0
cap (10^9B)      -
cap (blocks)     5033164800
volgrp           V1
ranks            12
dbexts           0
sam              Standard
repcapalloc      -
eam              managed
reqcap (blocks)  5033164800
realextents      2400
virtualextents   0
migrating        0
perfgrp          PG0
migratingfrom    -
resgrp           RG0
==============Rank extents==============
rank extents
============
R7       200
R9       200
R10      200
R19      200
R20      200
R21      200
R22      200
R23      200
R24      200
R25      200
```

```
R26      200
R27      200

dscli> lsrank -l -extpool p4
ID  Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R7      0 Normal Normal    A13          5 P4         ITSO_FB_P4 fb      384       207
R9      0 Normal Normal    A15          5 P4         ITSO_FB_P4 fb      448       207
R10     0 Normal Normal    A16          6 P4         ITSO_FB_P4 fb      376       206
R18     0 Normal Normal    A3           5 P4         ITSO_FB_P4 fb      1574      192
R19     0 Normal Normal    A17          5 P4         ITSO_FB_P4 fb      448       361
R20     0 Normal Normal    A18          5 P4         ITSO_FB_P4 fb      448       349
R21     0 Normal Normal    A19          5 P4         ITSO_FB_P4 fb      448       205
R22     0 Normal Normal    A10          6 P4         ITSO_FB_P4 fb      312       206
R23     0 Normal Normal    A11          6 P4         ITSO_FB_P4 fb      312       209
R24     0 Normal Normal    A12          6 P4         ITSO_FB_P4 fb      312       208
R25     0 Normal Normal    A21          5 P4         ITSO_FB_P4 fb      448       207
R26     0 Normal Normal    A26         10 P4         ITSO_FB_P4 fb      255       206
R27     0 Normal Normal    A27         10 P4         ITSO_FB_P4 fb      255       207
```

Complete the following steps to depopulate the rank:

1. Check the initial status by running the **lsrank**, **lsfbvol**, and **showfbvol** commands. Volume 4008 has 200 extents that are allocated on rank R27.

2. Perform the rank depopulation on rank R27 by running the **chrank -quiet -unassign r27** command, as shown in Example 4-16.

*Example 4-16 Use the chrank command to initiate rank depopulation*

```
dscli> chrank -quiet -unassign  r27
CMUC00008I chrank: Rank R27 successfully modified.
```

3. Check the rank status after the depopulation starts (Example 4-17). The used extents on all remaining ranks increase immediately after the **chrank** command is run. This increase is caused by the Easy Tier pre-allocation process. Only the number of used extents on rank R27 decreased.

*Example 4-17 Checking the rank status after the rank depopulation starts*

```
dscli> lsrank -l -extpool p4
ID  Group State       datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R7      0 Normal      Normal    A13          5 P4         ITSO_FB_P4 fb      384       226
R9      0 Normal      Normal    A15          5 P4         ITSO_FB_P4 fb      448       226
R10     0 Normal      Normal    A16          6 P4         ITSO_FB_P4 fb      376       225
R18     0 Normal      Normal    A3           5 P4         ITSO_FB_P4 fb      1574      192
R19     0 Normal      Normal    A17          5 P4         ITSO_FB_P4 fb      448       379
R20     0 Normal      Normal    A18          5 P4         ITSO_FB_P4 fb      448       367
R21     0 Normal      Normal    A19          5 P4         ITSO_FB_P4 fb      448       224
R22     0 Normal      Normal    A10          6 P4         ITSO_FB_P4 fb      312       225
R23     0 Normal      Normal    A11          6 P4         ITSO_FB_P4 fb      312       228
R24     0 Normal      Normal    A12          6 P4         ITSO_FB_P4 fb      312       227
R25     0 Normal      Normal    A21          5 P4         ITSO_FB_P4 fb      448       226
R26     0 Normal      Normal    A26         10 P4         ITSO_FB_P4 fb      255       225
R27     0 Depopulating Normal   A27         10 P4         ITSO_FB_P4 fb      255       205
```

**Important:** The number of used extents on the flash rank R18 did not change after **chrank -unassign** was run. In a hybrid pool, extent allocation always allocates extents on the ranks of the enterprise class tier first until no space is left. If no space is left on the enterprise tier, extent allocation continues on the ranks of the nearline tier and finally on the flash tier. However, you can change the new volume allocation policy. For more information, see 3.3.2, “Guidelines for creating multi-tier extent pools” on page 52.

Extent allocation automatically is balanced, in a rotate-extent fashion, across the ranks within a storage tier. For more information, see 2.8.9, “Extent allocation in hybrid and managed extent pools” on page 38.

4. Check the rank status by running the **lsrank** command (Example 4-18). The used extents on all remaining ranks did not change and the number of used extents on rank R27 is still decreasing.

*Example 4-18 Checking the rank status while rank depopulation is running*

```
dscli> lsrank -l -extpool p4
ID  Group State        datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
==========================================================================================
R7      0 Normal       Normal    A13          5 P4         ITSO_FB_P4 fb      384       230
R9      0 Normal       Normal    A15          5 P4         ITSO_FB_P4 fb      448       231
R10     0 Normal       Normal    A16          6 P4         ITSO_FB_P4 fb      376       230
R18     0 Normal       Normal    A3           5 P4         ITSO_FB_P4 fb      1574      192
R19     0 Normal       Normal    A17          5 P4         ITSO_FB_P4 fb      448       384
R20     0 Normal       Normal    A18          5 P4         ITSO_FB_P4 fb      448       373
R21     0 Normal       Normal    A19          5 P4         ITSO_FB_P4 fb      448       230
R22     0 Normal       Normal    A10          6 P4         ITSO_FB_P4 fb      312       230
R23     0 Normal       Normal    A11          6 P4         ITSO_FB_P4 fb      312       232
R24     0 Normal       Normal    A12          6 P4         ITSO_FB_P4 fb      312       231
R25     0 Normal       Normal    A21          5 P4         ITSO_FB_P4 fb      448       229
R26     0 Normal       Normal    A26         10 P4         ITSO_FB_P4 fb      255       228
R27     0 Depopulating Normal    A27         10 P4         ITSO_FB_P4 fb      255       185
```

5. Check the status of volume 4008 by running **showfbvol -rank** (Example 4-19). The rank depopulation already started to migrate some of the volume extents on R27 to the other ranks.

*Example 4-19 Checking the status of volume 4008 while rank depopulation is running*

```
dscli> showfbvol -rank 4008
Name            -
ID              4008
accstate        Online
datastate       Normal
configstate     Normal
deviceMTM       2107-900
datatype        FB 512
addrgrp         4
extpool         P4
exts            2400
captype         DS
cap (2^30B)     2400.0
cap (10^9B)     -
cap (blocks)    5033164800
volgrp          V1
ranks           12
dbexts          0
sam             Standard
```

```
repcapalloc     -
eam             managed
reqcap (blocks) 5033164800
realextents     2400
virtualextents  0
migrating       0
perfgrp         PG0
migratingfrom   -
resgrp          RG0
==============Rank extents==============
rank extents
============
R7       202
R9       202
R10      202
R19      202
R20      202
R21      202
R22      201
R23      201
R24      202
R25      202
R26      202
R27      180
```

6. After a short pause, check the extent pool P4 rank status by running the **lsrank -l** command (Example 4-20). Rank R27 is no longer part of the extent pool. Running the **lsrank -l r27** command shows that the rank is now unassigned, which means that the depopulation process ended.

*Example 4-20 After rank depopulation rank is removed from the extent pool*

```
dscli> lsrank -l -extpool p4
ID  Group State       datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
=========================================================================================
R7      0 Normal      Normal    A13          5 P4         ITSO_FB_P4 fb      384       230
R9      0 Normal      Normal    A15          5 P4         ITSO_FB_P4 fb      448       231
R10     0 Normal      Normal    A16          6 P4         ITSO_FB_P4 fb      376       230
R18     0 Normal      Normal    A3           5 P4         ITSO_FB_P4 fb      1574      192
R19     0 Normal      Normal    A17          5 P4         ITSO_FB_P4 fb      448       384
R26     0 Normal      Normal    A26         10 P4         ITSO_FB_P4 fb      255       228
R22     0 Normal      Normal    A10          6 P4         ITSO_FB_P4 fb      312       230
R23     0 Normal      Normal    A11          6 P4         ITSO_FB_P4 fb      312       232
R24     0 Normal      Normal    A12          6 P4         ITSO_FB_P4 fb      312       231
R25     0 Normal      Normal    A21          5 P4         ITSO_FB_P4 fb      448       229
R26     0 Normal      Normal    A26         10 P4         ITSO_FB_P4 fb      255       228
```

7. Check the status of volume 4008 by running the **showfbvol -rank** command (Example 4-21). The extents that were depopulated from rank R27 were migrated in a balanced manner across the other ranks of the enterprise tier in the extent pool.

*Example 4-21 Checking the volume status after rank depopulation*

```
dscli> showfbvol -rank 4008
Name            -
ID              4008
accstate        Online
datastate       Normal
configstate     Normal
deviceMTM       2107-900
datatype        FB 512
```

```
addrgrp         4
extpool         P4
exts            2400
captype         DS
cap (2^30B)     2400.0
cap (10^9B)     -
cap (blocks)    5033164800
volgrp          V1
ranks           11
dbexts          0
sam             Standard
repcapalloc     -
eam             managed
reqcap (blocks) 5033164800
realextents     2400
virtualextents  0
migrating       0
perfgrp         PG0
migratingfrom   -
resgrp          RG0
==============Rank extents==============
rank extents
============
R7       218
R9       218
R10      218
R19      218
R20      218
R21      218
R22      218
R23      218
R24      218
R25      219
R26      219
```

**Low priority:** The rank depopulation is a low-priority background process, which is optimized to avoid affecting performance on the extent pool. For that reason, the amount of time that is necessary for the rank depopulation is not predictable.

### 4.3.4 Easy Tier Automatic Mode

Easy Tier Automatic Mode manages the volume capacity within managed extent pools to place the extents on the most appropriate storage tiers and ranks. When Easy Tier Automatic Mode is enabled, Easy Tier automatically handles data placement to optimize performance and storage economics. For more information about Easy Tier concepts and Easy Tier Automatic Mode, see Chapter 2, “DS8000 Easy Tier concepts, design, and implementation” on page 11.

Run the **chsi** command to turn on or off Easy Tier Automatic Mode and to start or stop Easy Tier monitoring, as shown in the following example:

```
chsi -ETAutoMode tiered|all|none -ETMonitor automode|all|none storage_image_ID
```

The **chsi** command parameters are described:

- The **-ETAutoMode** parameter specifies whether the IBM Easy Tier feature is active and controlling the automated portion of this feature. If the value is set to `all`, the Easy Tier Automatic Mode is active on all pools, and the Easy Tier Manual Mode is also active. If the value is set to `tiered`, the Easy Tier Automatic Mode is active on pools with multiple tiers, and the Easy Tier Manual Mode is also active.

  If the value is set to `none`, the Easy Tier Manual Mode is active, but the Easy Tier Automatic Mode is inactive. You can still relocate data manually at a volume level by using Easy Tier Manual Mode features if the Easy Tier LIC feature is installed. When the Easy Tier Automatic Mode is turned off, all Easy Tier migration controls at the pool level are reset.

- The **-ETMonitor** parameter specifies whether volumes are monitored for potential benefits by Easy Tier. If the value is set to `all`, all of the volumes on the DS8000 are monitored. If the value is set to `automode` and Easy Tier automatic migration mode is active, only those volumes that are managed by the Automatic Mode of Easy Tier are monitored. If the value is set to `automode` and Easy Tier is inactive, no volumes are monitored. If the value is set to `none`, no volumes are monitored.

  When the Easy Tier monitor is turned off, all the Easy Tier controls (learning/migration) at the pool level and volume level are reset. For more information, see Table 2-3 on page 42.

**Important:** Do not turn off Easy Tier monitoring if Easy Tier Automatic Mode (**ETAutoMode**) is enabled. When Easy Tier monitoring is turned off, no new performance statistics are collected and Easy Tier Automatic Mode cannot create migration plans. As a result, Easy Tier Automatic Mode stops managing volumes in all managed extent pools.

Changing the Easy Tier monitoring mode affects the statistics collection and can lead to a reset (reinitialization) of the gathered monitoring data. This situation means that it might take up to 24 hours of collecting new performance statistics after Easy Tier monitoring is enabled again until migration plans are created. Easy Tier Automatic Mode can continue to automatically optimize storage performance and economics by relocating extents across and within storage tiers.

Example 4-22 shows the enablement of Easy Tier Automatic Mode for hybrid or multi-tier extent pools and sets monitoring to `on` for all volumes. Run the **showsi** command to view Easy Tier Automatic Mode information.

*Example 4-22 Using the DS CLI to enable Easy Tier Automatic Mode*

```
dscli> chsi -ETAutoMode tiered -ETMonitor all
CMUC00042I chsi: Storage image IBM.2107-75DST11 successfully modified.

dscli> showsi
Name             -
desc             Sand Shark
ID               IBM.2107-75DST11
Storage Unit     IBM.2107-75DST10
Model            996
WWNN             5005076301234567
Signature        89ab-cdef-0123-4567
State            Online
ESSNet           Enabled
Volume Group     V0
os400Serial      462
NVS Memory       32.0 GB
Cache Memory     430.3 GB
```

```
Processor Memory 505.8 GB
MTS              IBM.5331-75DST10
numegsupported   16
ETAutoMode       tiered
ETMonitor        all
IOPMmode         Disabled
ETCCMode         -
ETHMTMode        Enabled
ETSRMode         Enabled
ETTierOrder      High performance
ETAutoModeAccel  Disabled
```

The **lsextpool -l** and **showextpool** commands show information about the number of tiers in an extent pool and whether the pool is managed by Easy Tier Automatic Mode. Example 4-23 on page 100 shows a DS8000 configuration with the **ETAutoMode** control switch set to `tiered`. Only hybrid, multi-tier extent pools are actively managed by Easy Tier Automatic Mode.

All homogeneous single-tier extent pools (`numtiers=1` or `numtiers=0`) are not managed by Easy Tier, which means that no automatic intra-tier performance management (auto-rebalance) is active in these single-tier extent pools.

*Example 4-23 Easy Tier status and number of tiers information*

```
dscli> lsextpool -l
Name       ID stgtype rankgrp  [...]  numvols numranks encryptgrp numtiers etmanaged
==========================================================================================
ckd_hybrid P0 ckd           0            2688        2          -        2 yes
ckd_600_1  P1 ckd           1            2690        1          -        1 no
tst_62esp  P2 fb            0               6        3          -        2 yes
pool_1tier P3 fb            1               0        0          -        0 no
pool_3tier P4 fb            0              44        4          -        3 yes

dscli> showextpool p4
Name                pool_3tier
ID                  P4
stgtype             fb
[...]
numtiers            3
etmanaged           yes

dscli> showsi IBM.2107-75TV181
Name             ATS_04
[...]
ETAutoMode       tiered
ETMonitor        all
```

# 4.4 Easy Tier control by using the DS CLI

This section provides examples of Easy Tier control capabilities by using the DS CLI. For more information, see *IBM System Storage DS Command-Line Interface user's Guide for DS8000 series,* SC27-8526, or run the **help** command.

**Command output:** Some of the command output lines in the following examples were truncated to make the examples more readable.

### 4.4.1 Suspend or resume the Easy Tier migration process for an extent pool

You can pause the extent pool-level migration process from the DS CLI. The default duration is 168 hours, but the user can specify their own settings by using the **duration** option in **manageextpool**. Example 4-24 shows how to pause and resume monitoring at the extent pool level. The user can monitor the **etmigpauseremain** parameter from the **shwextpool** command.

*Example 4-24 Suspending and resuming migration at the pool level*

```
dscli> manageextpool -action etmigpause -duration 96H P9
IBM.2107-xxxxxxx CMUC00477I manageextpool: The etmigpause action for the pool P9 has
completed.

dscli> showextpool P9
[...]
etmanaged           yes
etmigpauseremain    96HOM
etmonpauseremain    -
etmonitorreset      unknown

dscli> manageextpool -action etmigresume  P9
CMUC00477I manageextpool: The etmigresume action for the pool P9 has completed.
```

### 4.4.2 Suspend and resume the Easy Tier monitoring process for extent pool

You can suspend the monitoring process for a specific extent pool. The default duration is 168 hours, but you can also use your own settings, through the **duration** option in the **manageextpool** command. Example 4-25 shows how to pause and resume monitoring for an extent pool. You can watch the **etmonpauseremain** output field from the **showextpool** command to see how long monitoring remains suspended. You can resume monitoring any time by using the **etmonresume** parameter with the **manageextpool** command.

*Example 4-25 Suspending and resuming monitoring at the pool level*

```
dscli> showextpool P9
Name                ITSO_ET1_Pool
ID                  P9
[...]
etmanaged           yes
etmigpauseremain    -
etmonpauseremain    -
etmonitorreset      unknown

dscli> manageextpool -action etmonpause -duration 1h P9
CMUC00477I manageextpool: The etmonpause action for the pool P9 has completed.

dscli> showextpool P9
Name                ITSO_ET1_Pool
ID                  P9
etmanaged           yes
etmigpauseremain    -
etmonpauseremain    1HOM
etmonitorreset      unknown

dscli> manageextpool -action etmonresume P9
CMUC00477I manageextpool: The etmonresume action for the pool P9 has completed.
```

```
dscli> showextpool P9
Name                 ITSO_ET1_Pool
ID                   P9
[...]
etmanaged            yes
etmigpauseremain     -
etmonpauseremain     -
etmonitorreset       unknown
```

## 4.4.3 Reset the Easy Tier monitoring process for the extent pool

You can also reset the monitoring process (that is, reset the statistics and restart the learning from scratch). Use the **etmonreset** parameter with the **manageextpool** command, as shown in Example 4-26. You can check the **etmonitorreset** `Suspend` and `resume` output value from the **showextpool** command to see when the monitoring was last reset.

*Example 4-26 Reset the Easy Tier monitoring data at the pool level*

```
dscli> showextpool P9
Name                 ITSO_ET1_Pool
ID                   P9
[...]
etmanaged            yes
etmigpauseremain     -
etmonpauseremain     -
etmonitorreset       unknown

dscli> manageextpool -action etmonreset P9
CMUC00477I manageextpool: The etmonreset action for the pool P9 has completed.

dscli> showextpool P9
Name                 ITSO_ET1_Pool
ID                   P9
[...]
etmanaged            yes
etmigpauseremain     -
etmonpauseremain     -
etmonitorreset       2018-08-15T11:23:39+0200
```

## 4.4.4 Suspend and resume the Easy Tier monitoring process for a volume

Suspend and resume monitoring actions are also possible at the volume level by using the **manageckdvol** (for CKD volumes) and **managefbvol** (for fixed block volumes) commands (Example 4-27).

*Example 4-27 Suspending and resuming monitoring at the volume level*

```
dscli> showfbvol 3200
Name                       e980_p
ID                         3200
tierassignstatus           -
tierassignerror            -
tierassignorder            -
tierassigntarget           -
%tierassigned              0
etmonpauseremain           -
```

```
etmonitorreset            unknown

dscli> managefbvol -action etmonpause 3200
CMUC00430I managefbvol: The etmonpause action for FB volume 3200 has completed.

dscli> showfbvol 3200
Name                      e980_p
ID                        3200
tierassignstatus          -
tierassignerror           -
tierassignorder           -
tierassigntarget          -
%tierassigned             0
etmonpauseremain          168H0M
etmonitorreset            unknown

dscli> showfbvol 3200
Name                      e980_p
ID                        3200
tierassignstatus          -
tierassignerror           -
tierassignorder           -
tierassigntarget          -
%tierassigned             0
etmonpauseremain          167H58M
etmonitorreset            unknown

dscli> managefbvol -action etmonresume 3200
CMUC00430I managefbvol: The etmonresume action for FB volume 3200 has completed.

dscli> showfbvol 3200
Name                      e980_p
ID                        3200
tierassignstatus          -
tierassignerror           -
tierassignorder           -
tierassigntarget          -
%tierassigned             0
etmonpauseremain          -
etmonitorreset            unknown
```

## 4.4.5 Reset the Easy Tier monitoring process for a volume

You can also reset monitoring for individual volumes. Use the **manageckdvol** (for CKD volumes) and **managefbvol** (for fixed block volumes) commands, as shown in Example 4-28.

*Example 4-28 Reset the Easy Tier monitoring data at the volume level*

```
dscli> showfbvol 3200
Name                      e980_p
ID                        3200
tierassignstatus          -
tierassignerror           -
tierassignorder           -
tierassigntarget          -
%tierassigned             0
etmonpauseremain          -
etmonitorreset            unknown
```

```
dscli> managefbvol -action etmonreset 3200
CMUC00430I managefbvol: The etmonreset action for FB volume 3200 has completed.

dscli> showfbvol 3200
Name                     e980_p
ID                       3200
tierassignstatus         -
tierassignerror          -
tierassignorder          -
tierassigntarget         -
%tierassigned            0
etmonpauseremain         -
etmonitorreset           2018-08-15T12:54:11+0200

dscli> showfbvol 3201
Name                     e970_p
ID                       3201
tierassignstatus         -
tierassignerror          -
tierassignorder          -
tierassigntarget         -
%tierassigned            0
etmonpauseremain         -
etmonitorreset           2018-08-15T11:23:39+0200
```

## 4.4.6 Assigning a volume to a tier (pinning a volume to a tier)

You can pin a group of volumes manually to a certain tier in a multi-tier pool and later release them again back into the automated Easy Tier management. This kind of pinning might be useful in special workload tasks that seldom run (for example, once a quarter). In that case, pinning volumes to the highest tier, such as two days before an end-of-quarter run starts, can help to pre-stage the volumes to the highest performance level.

Also, with this function, you can put more applications and volumes into large multi-tier extent pools. Before this function was available, you might have been required to set aside separate higher-tier pools for certain applications.

To pin a group of volumes to a tier, use the DS CLI **managefbvol/manageckdvol -action tierassign -tier *xxx*** command. To release the group of volumes again, use the **-action tierunassign** command. The commands can be applied to a whole range of volumes at once. Example 4-29 shows assigning a group of FB volumes to the `FlashTier0` tier.

*Example 4-29 Assigning a group of FB volumes to the flash tier in a large multi-tier extent pool*

```
dscli> managefbvol -action tierassign -tier flashtier0 47a2-47a5
CMUC00430I managefbvol: The tierassign action for FB volume 47A2 has completed.
CMUC00430I managefbvol: The tierassign action for FB volume 47A3 has completed.
CMUC00430I managefbvol: The tierassign action for FB volume 47A4 has completed.
CMUC00430I managefbvol: The tierassign action for FB volume 47A5 has completed.
```

Also, an assignment to the nearline (`nl`) tier or to the enterprise (`ent`) tier is possible, and these assignments can also be permanent.

With the **showfbvol/showckdvol -tier** command, you can control this assignment for a volume. The **showextpool -tier** (or **-rank**) command can control this assignment for the entire extent pool.

**Note:** Assigning a volume to a tier starts the process to migrate the data in the volume to the specified tier.

When assigning a new tier for the volume, Easy Tier automatically starts migrating data to the new tier. Check the following output parameters from the **showfbvol/showckdvol** command to get the detailed tier assignment information (Example 4-30).

*Example 4-30 Output parameters to get the detailed tier assignment information*

```
tierassignstatus Assign Pending/Assigning/Assigned
tierassignerror -
tierassignorder ETdata
tierassigntarget FlashTier0/FlashTier1/FlashTier2/SSD/ENT/NL/NLExcluded
%tierassigned
```

## 4.4.7 Exclude a volume from the nearline tier

Also, you can exclude a volume from being assigned to a nearline tier.

**Note:** If a volume was initially on a nearline rank, restricting a volume from using the nearline tier starts a process to migrate the volume extents away from the nearline ranks.

To exclude a volume from the nearline tier, use the DS CLI **managefbvol/manageckdvol -action tierassign -tier nlexclude *xxx*** command.

Example 4-31 shows a scenario where volume A000, which initially was on a nearline rank, is excluded from the nearline rank. Over time, you can check the **tierassignstatus**, **tierassigntarget**, and **%tierassign** output parameters from the **showfbvol/showckdvol** commands to see the volume that is being reassigned until it moves completely to a non-nearline tier (ENT in this example).

*Example 4-31 Exclude nearline disks from the volume*

```
dscli> showfbvol -tier A000
[...]
==========Tier Distribution============
Tier %allocated
===============
NL          100

dscli> managefbvol -action tierassign -tier nlexclude A000
CMUC00430I managefbvol: The tierassign action for FB volume A000 has completed.

dscli> showfbvol -rank a000
Name             ITSO_ET_fbvol
ID               A000
.
.
.
tierassignstatus Assign Pending
tierassignerror  -
tierassignorder  ETdata
```

```
tierassigntarget NLExcluded
%tierassigned    0
dscli> showfbvol -tier A000
[...]
===========Tier Distribution============
Tier %allocated
===============
NL          100


dscli> showfbvol -rank a000
Name             ITSO_ET_fbvol
ID               A000
[...]
ranks              2
[...]
tierassignstatus Assigning
tierassignerror  -
tierassignorder  ETdata
tierassigntarget NLExcluded
%tierassigned    24
etmonpauseremain -
etmonitorreset   unknown
==============Rank extents==============
R4        66
R26       34
dscli> showfbvol -tier A000
[...]
===========Tier Distribution============
Tier %allocated
===============
NL           66
ENT          34

dscli> lsrank R26
ID  Group State  datastate Array RAIDtype extpoolID stgtype
============================================================
R26     0 Normal Normal    A22          5 P4        fb

dscli> lsarray -l A22
Array State    Data   RAIDtype  arsite Rank DA Pair DDMcap (10^9B) diskclass encrypt
=========================================================================================
A22   Assigned Normal 5 (6+P+S) S6     R26  0                600.0 ENT       supported

dscli>  showfbvol -rank a000
Name             ITSO_ET_fbvol
ID               A000
tierassignstatus Assigned
tierassignerror  -
tierassignorder  ETdata
tierassigntarget NLExcluded
%tierassigned    100
etmonpauseremain -
etmonitorreset   unknown
==============Rank extents==============
rank extents
============
R26      100
dscli> showfbvol -tier A000
[...]
```

```
==========Tier Distribution============
Tier %allocated
===============
ENT         100
```

## 4.4.8 Unassigning a certain tier from a volume

Also, you can unassign a certain tier from a volume.

**Note:** When a user unassigns a tier from a volume, Easy Tier does not initiate an immediate migration task.

To unassign a volume from a certain tier, use the DS CLI **managefbvol/manageckdvol -action tierunassign *xxx*** command.

Example 4-32 shows volume A001, which initially was assigned to the flash tier, which is released again.

*Example 4-32 Unassign tier from volume that was assigned to a flash tier*

```
dscli>  showfbvol a001
Name             ITSO_ET_fbvol02
ID               A001
[...]
tierassignstatus Assigned
tierassignerror  -
tierassignorder  ETdata
tierassigntarget FlashTier0
%tierassigned    100
etmonpauseremain -
etmonitorreset   unknown

dscli> showfbvol -tier A001
[...]
==========Tier Distribution============
Tier       %allocated
=====================
FlashTier0        100

dscli> managefbvol -action tierunassign A001
CMUC00430I managefbvol: The tierunassign action for FB volume A001 has completed.

dscli>  showfbvol a001
Name             ITSO_ET_fbvol02
ID               A001
[...]
tierassignstatus -
tierassignerror  -
tierassignorder  -
tierassigntarget -
%tierassigned    0
etmonpauseremain -
etmonitorreset   unknown
```

5

# DS8000 Easy Tier reporting

This chapter describes how to use IBM System Storage Easy Tier reporting in both the DS8000 Storage Management GUI and the STAT Charting Utility. These reports show the value of Easy Tier from both an operational and usability viewpoint.

Easy Tier can help the storage administrator configure and allocate the storage capacity for the optimal cost-to-performance ratio based on the performance data that is collected by the Easy Tier monitoring task.

This chapter includes the following topics:

- Easy Tier usage
- Easy Tier Reporting in the DS
- Easy Tier Reporting in the DS8000 GUI
- Using the Easy Tier data files
- STAT Charting Utility
- Disk Magic and IBM Storage Modeller

**Note:** The STAT reporting tool itself can no longer be used with DS8000 code R8.3 or later, but remains available for DS8000 R8.2 code and prior code levels.
This can be useful for instance when replacing legacy DS8870 systems.

## 5.1 Easy Tier usage

Easy Tier provides an easy solution for managing DS8000 storage capacity for the optimal usage of available storage tiers and resources within a tier.

In Easy Tier Automatic Mode, Easy Tier automatically and transparently relocates hot extents in hybrid extent pools to the flash, enterprise, or nearline storage tier. Within one tier, Easy Tier can rebalance individual overloaded ranks. You do not need to make complex decisions or logical unit number (LUN) configurations. You can leave all the complex work to the system, which can perform self-tuning within a certain hardware configuration. When configuration or workload changes occur, the system automatically adjusts over time.

In Easy Tier Manual Mode, Easy Tier also offers significant advantages. You can relocate volumes, merge extent pools, rebalance ranks on request, and depopulate ranks in a way that is transparent to the host applications. Additionally, you can also control Easy Tier on the pool level or the volume level.

All Easy Tier functions also work for flash cards.

Easy Tier provides other capabilities by integrating more closely with external components. For more information, see the following publications:

- *DS8870 Easy Tier Application*, REDP-5014
- *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015

By using the statistical usage data that is collected, you can understand the application data that can benefit the most from being relocated to flash (flash drives and solid-state drives), enterprise (serial-attached SCSI (SAS)) drives, or nearline drives.

## 5.2 Easy Tier Reporting in the DS8000 GUI

Since DS8000 Release 8.3 Easy Tier reporting is integrated in the DS GUI. With earlier microcode releases, users can still use the STAT tool to produce reports.

The Easy Tier reports can be found in the performance monitoring section of the DS GUI. To navigate to this section, select **Monitoring** and then **Performance**, as shown in Figure 5-1.

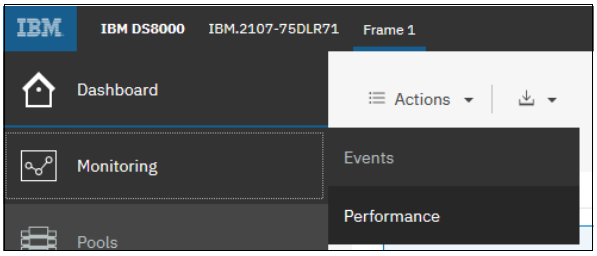



*Figure 5-1 DS8000 GUI: Performance*

In the Performance window, click the toolbar and select either **Pool** or **Volume** to make the Easy Tier metrics available, as shown in Figure 5-2.

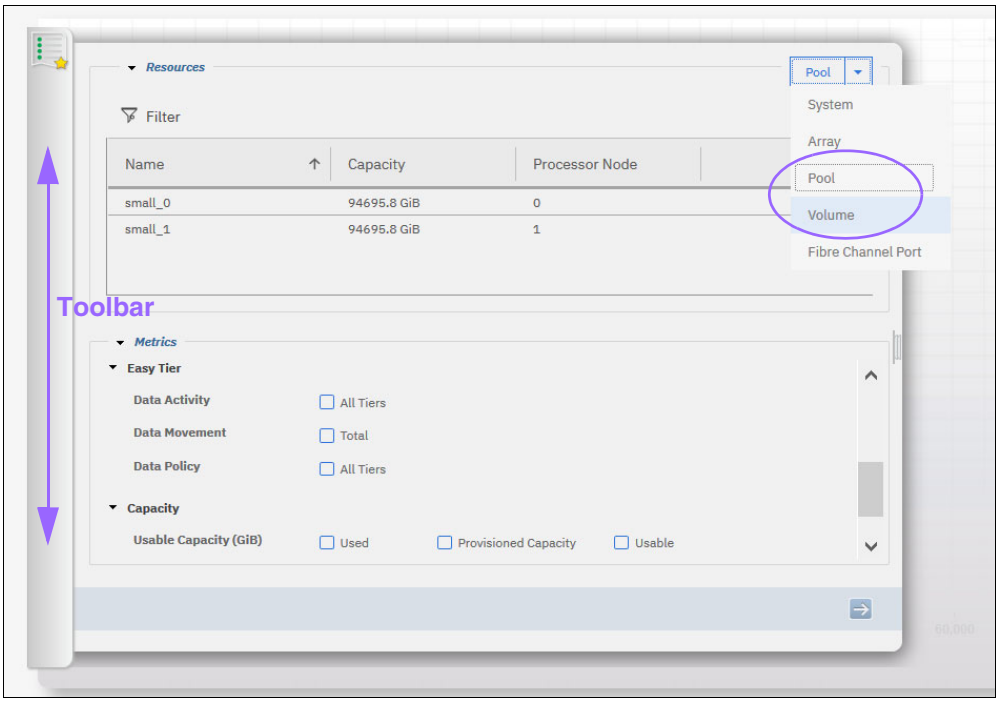



*Figure 5-2 DS8000 GUI: Easy Tier reporting*

## 5.2.1 Easy Tier pool level reporting

These Easy Tier reports are available at the pool level:

- Data Activity for All Tiers
- Data Movement Total
- Data Policy for All Tiers

### Easy Tier Data Activity report

The Data Activity report shows the activity within each tier of a pool as measured by Easy Tier. The most recent day is shown by default, but the report can be changed to any of the previous six days by clicking the date at the bottom of the window.

The data is classified as follows, using different colors as shown in Figure 5-3 on page 112:

- **Available**: Capacity *not* allocated
- **Inactive**: Data with zero activity
- **Low activity**: Data with activity below the cost-to-benefit ratio threshold for the total I/Os
- **Sequential Activity**: All data that is not classified by the other categories
- **High Activity**: Data with more than 0.1 IOPS per GB access density for small I/Os

Select a pool to monitor and then in the Easy Tier section shown in Figure 5-2, select **All Tiers** for **Data Activity**.

Click the arrow at the bottom of the window to get the chart shown in Figure 5-3.

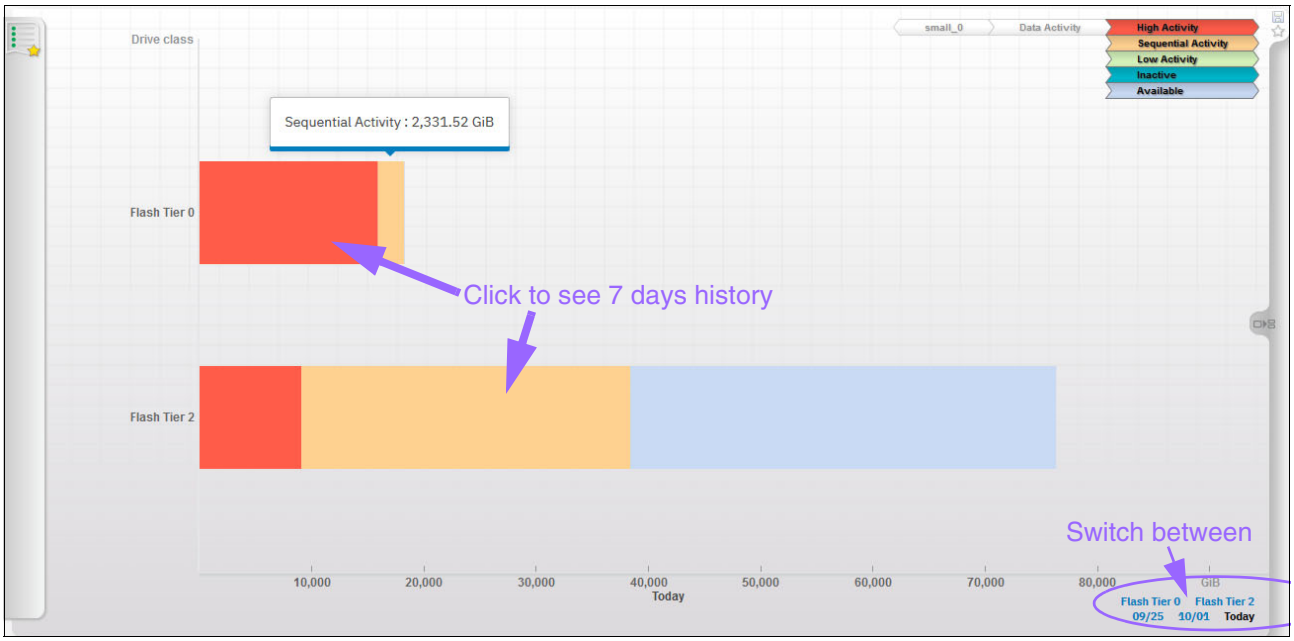



*Figure 5-3 DS8000 GUI: Easy Tier Pool Data Activity*

Clicking a tier gives a detailed 7-day view for that tier.

## Easy Tier Data Movement Report

The data movement report shows the amount of data that has been moved by Easy Tier within a pool over the past hour, day, or week, as shown in Figure 5-4.

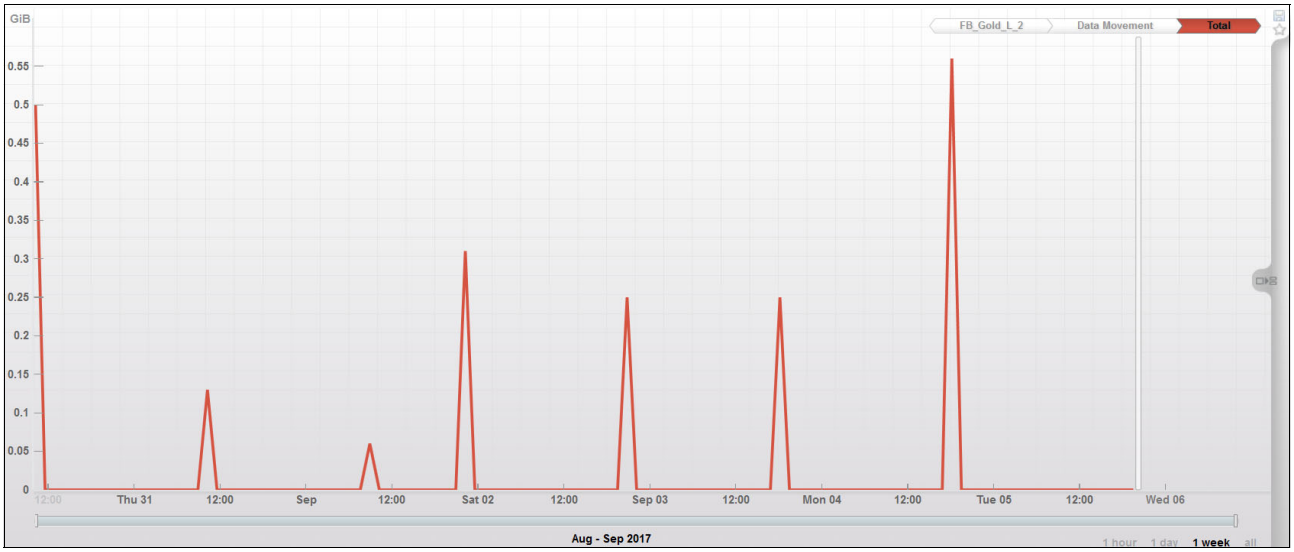



*Figure 5-4 DS8000 GUI: Easy Tier Pool Data Movement*

### Easy Tier Data Policy Report

The data policy report displays how the capacity breaks down by policies for each tier within a pool, as shown in Figure 5-5.

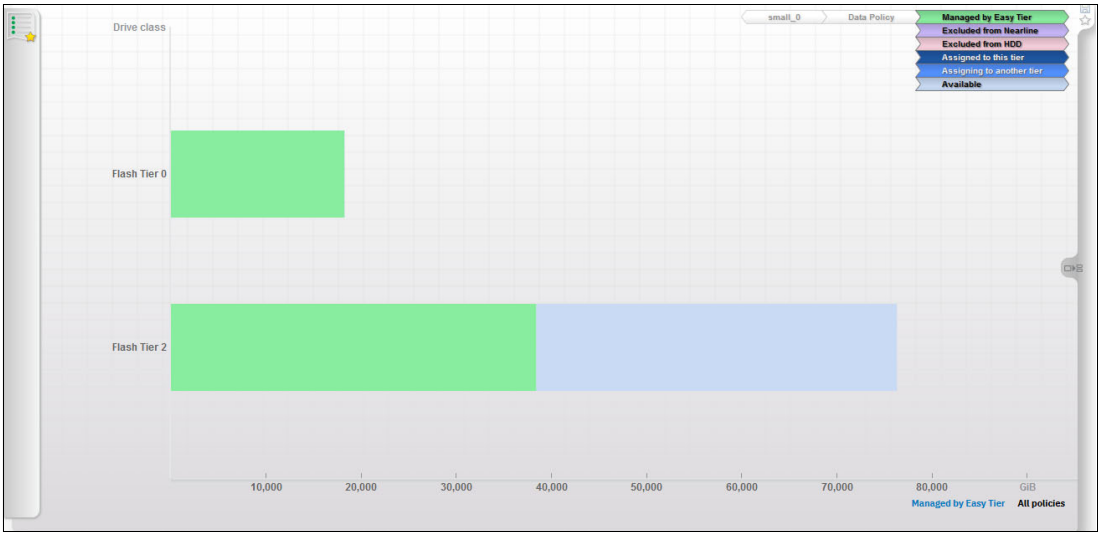



*Figure 5-5 DS8000 GUI: Easy Tier Pool Data Policy*

Clicking the managed capacity within a tier brings you to a chart similar to the Data Activity chart in Figure 5-3 on page 112.

## 5.2.2 Easy Tier Volume-Level Reporting

These two Easy Tier reports are available at the volume level:

- Data Activity
- Data Policy

### Easy Tier Data Activity report

The Data Activity report shows the activity within each volume as measure by Easy Tier (Figure 5-6). If a volume is spread across multiple tiers, all tiers are shown. The most recent day is shown by default, but the report can be changed to any of the previous six days by clicking the date at the bottom of the window.

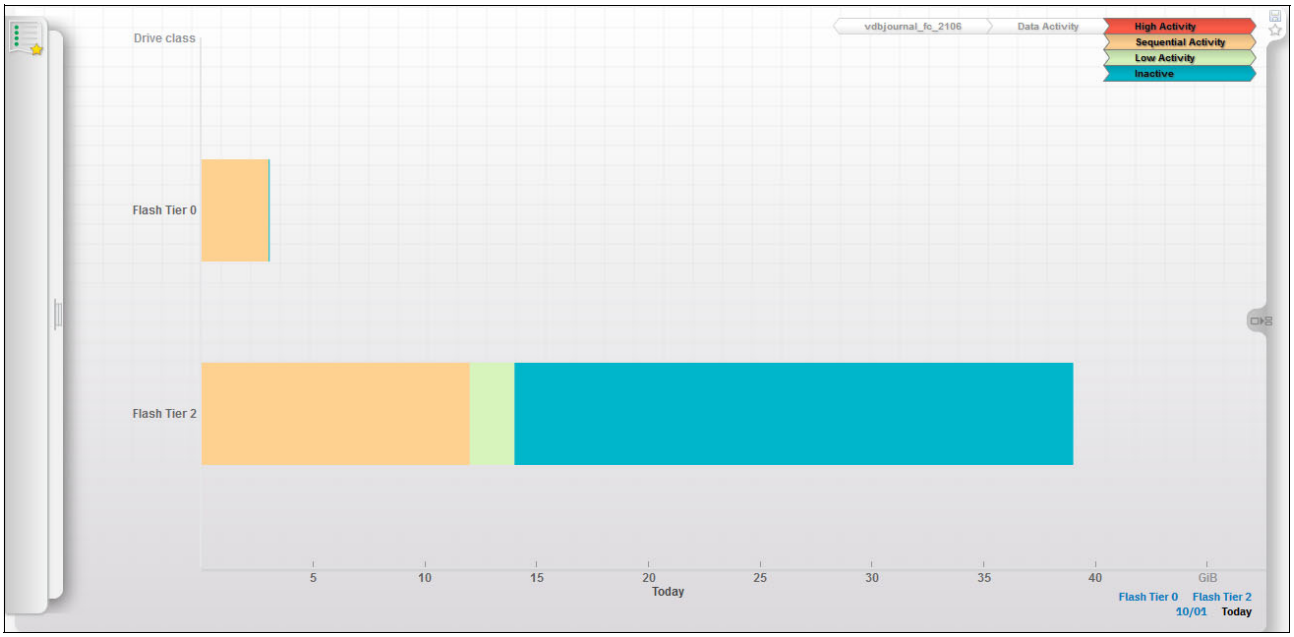



*Figure 5-6 DS8000 GUI: Easy Tier Data Volume Data Activity*

Clicking a tier shows 7-day workload activity history of the selected volumes within that tier, as shown in Figure 5-7.

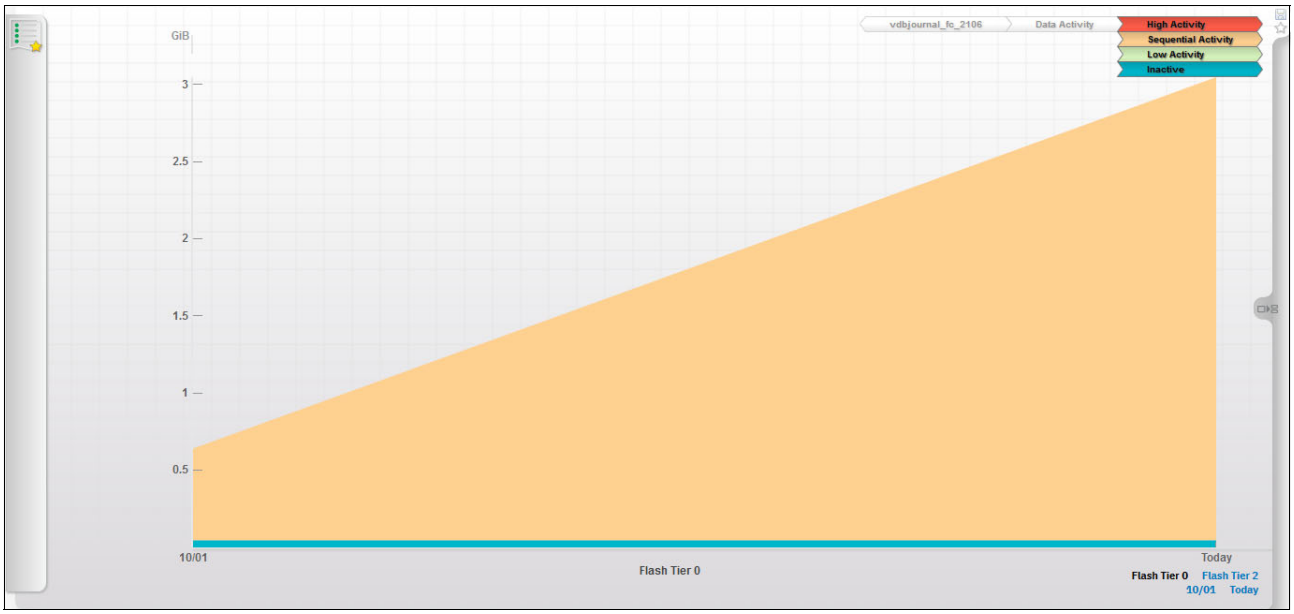



*Figure 5-7 DS8000 GUI: Easy Tier Data Volume Data Activity: 7-day tier history*

## Easy Tier Data Policy report

The Data Policy report for a volume shows the data policy and distribution between tiers for a volume or multiple volumes, as seen in Figure 5-8.

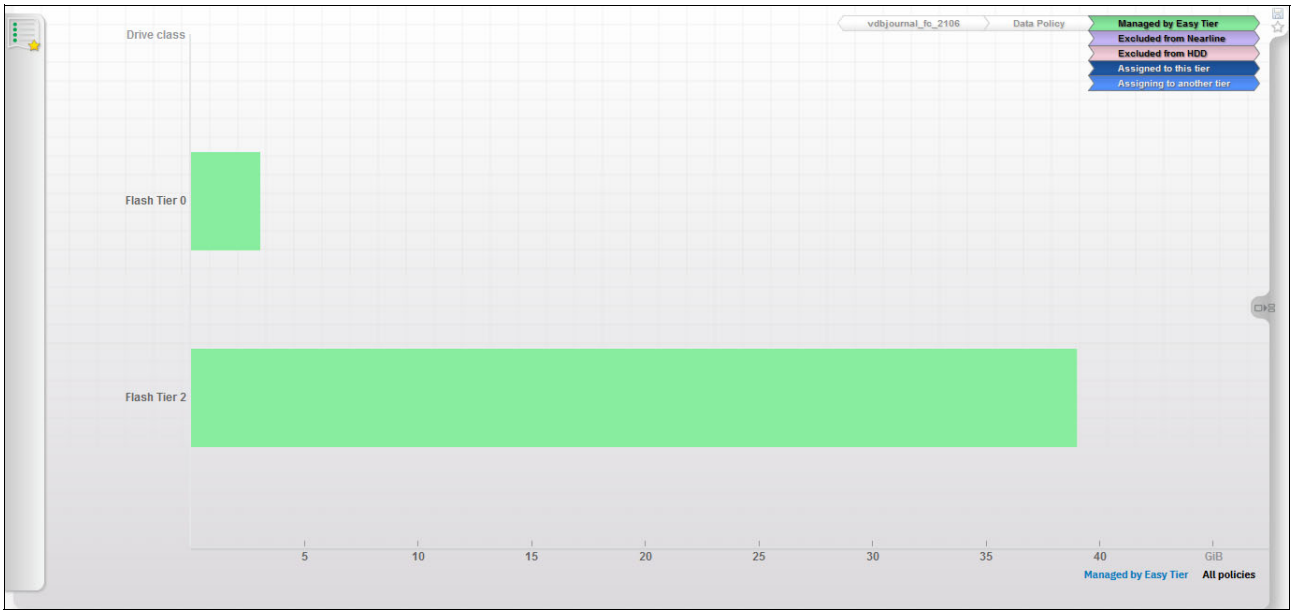



*Figure 5-8 DS8000 GUI: Easy Tier Data Volume Data Policy*

# 5.3 Using the Easy Tier data files

Easy Tier creates data files that can be used for external reporting and planning.

## 5.3.1 Easy Tier data files

The DS GUI or DS CLI can be used to offload the Easy Tier data files that were previously produced by the STAT tool. The data files include data for these items:

- skew_curve
- data_movement
- workload

To offload the Easy Tier data files from the DS GUI, click the download symbol (or for older DS8000 models: the disk icon) and select **Export Easy Tier Summary** as shown in Figure 5-9.

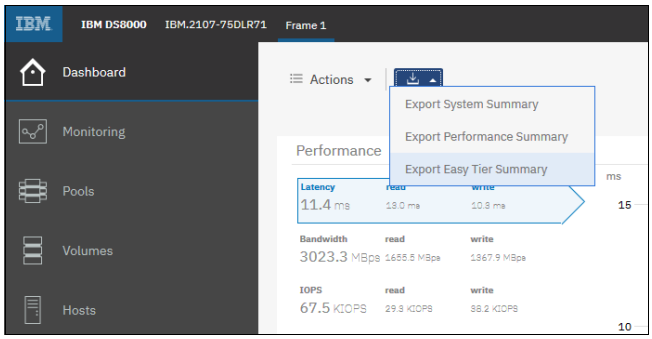



*Figure 5-9 Export Easy Tier files*

A compressed file is downloaded that contains the three Easy Tier data files along with the STAT Charting Utility.

To offload the Easy Tier data files using the DSCLI, issue the **offloadfile** command. As with the GUI, a compressed file is downloaded that contains the three Easy Tier data files along with the STAT Charting Utility. See Figure 5-10.

```
dscli> offloadfile -etdataCSV \tmp
Date/Time: 2019-09-13T18:38:50+0200 IBM DSCLI Version: 7.9.0.491 DS:
IBM.2107-75ZA571
CMUC00428I offloadfile: The etdataCSV file has been offloaded to
\tmp\et_data_20190913181514.zip.
```

*Figure 5-10 DSCLI offloadfile command*

The contents of these files are covered in 5.4, “STAT Charting Utility”.

## 5.4 STAT Charting Utility

The STAT Charting Utility is a Microsoft Excel worksheet that was developed by the IBM Advanced Technical Skills team to provide reporting based on the three Easy Tier CSV files.

You can obtain the utility in these ways:

- It can be downloaded along with the Easy Tier data using the DS GUI.
- It can be downloaded by using the DSCLI **offloadfile** command.
- It can be downloaded at this URL:
  http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5251

When you first open the utility, you will see two tabs. The Reference tab provides help and definitions. The Import tab can be used to import the three Easy Tier CSV files into the utility. As each of the files is imported, new tabs are created (Figure 5-11).

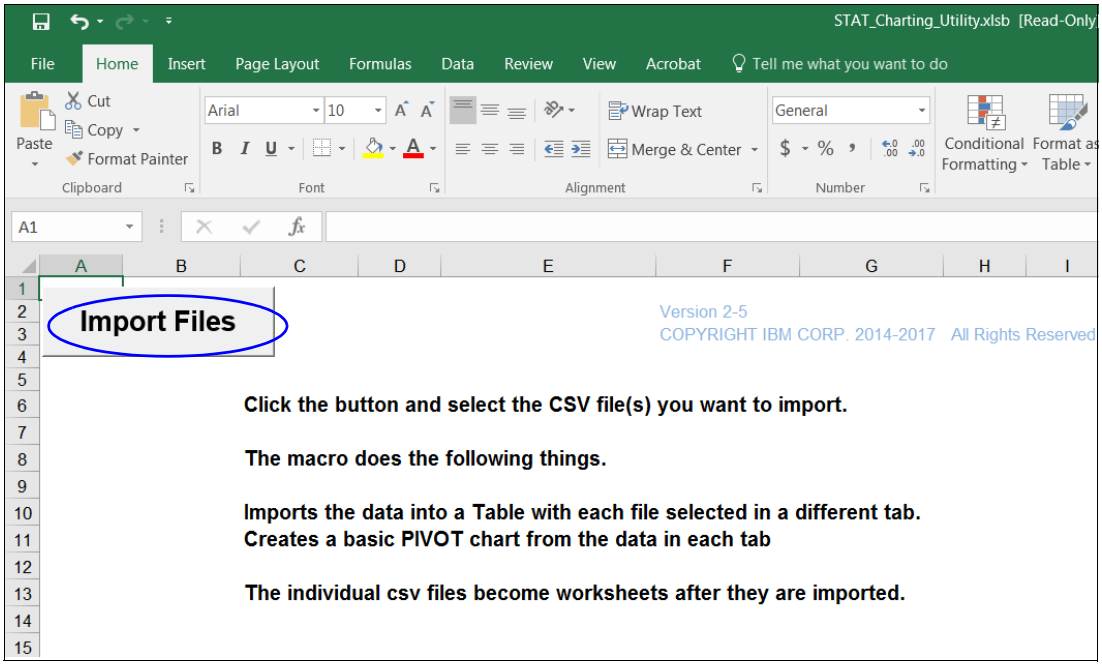



*Figure 5-11 STAT_Charting_Utility: Importing files*

## 5.4.1 Easy Tier skew curve

Skew curves can provide better guidelines for the tier configuration. Looking deeper into the file structure of the `skew_curve.csv` file, you see that it lists all extent pools and tiers and provides information in columns, as listed in Table 5-1.

*Table 5-1 Columns in the skew_curve.csv file*

| Column name | Meaning |
|---|---|
| smallwps | Small-block write IOPS, typically random |
| smallrps | Small-block read IOPS, typically random |
| smallmbs | Small-block MBps |
| largewps | Large-block write IOPS, typically sequential |
| largerps | Large-block read IOPS, typically sequential |
| largembs | Large-block MBps |
| smallms | Small-block back-end response time |
| largems | Large-block back-end response time |

Figure 5-12 shows one of the available spreadsheets.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | index | dss | datetime | extpool | tier | Percent | Write I/O | Total I/O | MB/sec | Small I/O | extent_cap(MB) | smallwps | smallrps | smallmbs |
| 2 | 1 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 0 | 373.81 | 458.59 | 12.87586 | 343.91 | 102400 | 334.72 | 9.19 | 3.78404 |
| 3 | 2 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 0 | 320.92 | 402.19 | 11.82935 | 293.55 | 102400 | 285.86 | 7.69 | 3.27477 |
| 4 | 3 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 0 | 295.28 | 376.31 | 11.57817 | 267.7 | 102400 | 260.84 | 6.86 | 3.00548 |
| 5 | 4 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 276.88 | 356.82 | 11.30579 | 249.69 | 102400 | 243.34 | 6.35 | 2.83088 |
| 6 | 5 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 263.26 | 339.22 | 10.4964 | 239.33 | 102400 | 233.3 | 6.03 | 2.68857 |
| 7 | 6 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 250.45 | 328.81 | 10.76596 | 224.85 | 102400 | 219.22 | 5.63 | 2.57199 |
| 8 | 7 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 232.08 | 311.36 | 10.66797 | 206.89 | 102400 | 201.72 | 5.17 | 2.40013 |
| 9 | 8 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 214.01 | 289.33 | 10.01104 | 190.19 | 102400 | 185.56 | 4.63 | 2.23058 |
| 10 | 9 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 1 | 208.12 | 287.11 | 10.30831 | 183.79 | 102400 | 179.4 | 4.39 | 2.15997 |
| 11 | 10 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 192.81 | 271.93 | 10.21385 | 167.95 | 102400 | 164.07 | 3.88 | 2.00723 |
| 12 | 11 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 182.98 | 260.72 | 9.95118 | 158.47 | 102400 | 154.86 | 3.61 | 1.92318 |
| 13 | 12 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 169.11 | 245.49 | 9.62163 | 145.63 | 102400 | 142.38 | 3.25 | 1.77511 |
| 14 | 13 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 158 | 237.04 | 9.67029 | 135.1 | 102400 | 132.12 | 2.98 | 1.65502 |
| 15 | 14 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 149.87 | 224.72 | 9.15088 | 128.18 | 102400 | 125.33 | 2.85 | 1.58031 |
| 16 | 15 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 2 | 142.52 | 216.5 | 9.03836 | 120.54 | 102400 | 117.9 | 2.64 | 1.52439 |
| 17 | 16 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 3 | 123.46 | 199.94 | 9.16604 | 100.7 | 102400 | 98.73 | 1.97 | 1.32437 |
| 18 | 17 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 3 | 84.15 | 346.3 | 19.42765 | 86.29 | 102400 | 63.44 | 22.85 | 1.59011 |
| 19 | 18 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 3 | 63.95 | 253.12 | 14.19543 | 66.56 | 102400 | 46.31 | 20.25 | 1.29676 |
| 20 | 19 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 3 | 23.64 | 153.06 | 8.98111 | 32.44 | 102400 | 12.22 | 20.22 | 0.86478 |
| 21 | 20 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 3 | 27.78 | 52.73 | 4.28041 | 17.96 | 102400 | 10.32 | 7.64 | 0.49356 |
| 22 | 21 | IBM.2107-SF75DLR70 | 2019-09-25T01:49:49-07:00 | P0 | 0 | 4 | 28.42 | 52.07 | 4.23733 | 17.71 | 102400 | 10.52 | 7.19 | 0.48443 |

Import | Movement | Movement_Chart1 | Skew | Skew_Chart1 | Skew_Chart2 | Workload | Workload_By_Policy | Workload_By_Activity

*Figure 5-12 File structure of a skew_curve.csv file*

Importing the `skew_curve.csv` file creates two new tabs.

The Skew_Chart1 tab includes a workload distribution graph similar to Figure 5-13 on page 118. This graph illustrates by extent pool the percentage of workload against a percentage of capacity.

A steep curve on this chart indicates that a small percentage of the capacity within a pool is making up most of the workload. A more gradual climb indicates that the workload is more evenly distributed.

The spreadsheet includes support for several skew curves, for instance, to compare IOPS to MBps, or to compare the skew curves of small and large blocks.

Typically, these graphs show that the skew for MBps is not as steep as the skew for the small-block IOPS. Also, you might be able to verify the Easy Tier behavior by putting small blocks and the most IOPS to the flash tier. The enterprise and nearline tiers are more suitable for larger blocks.

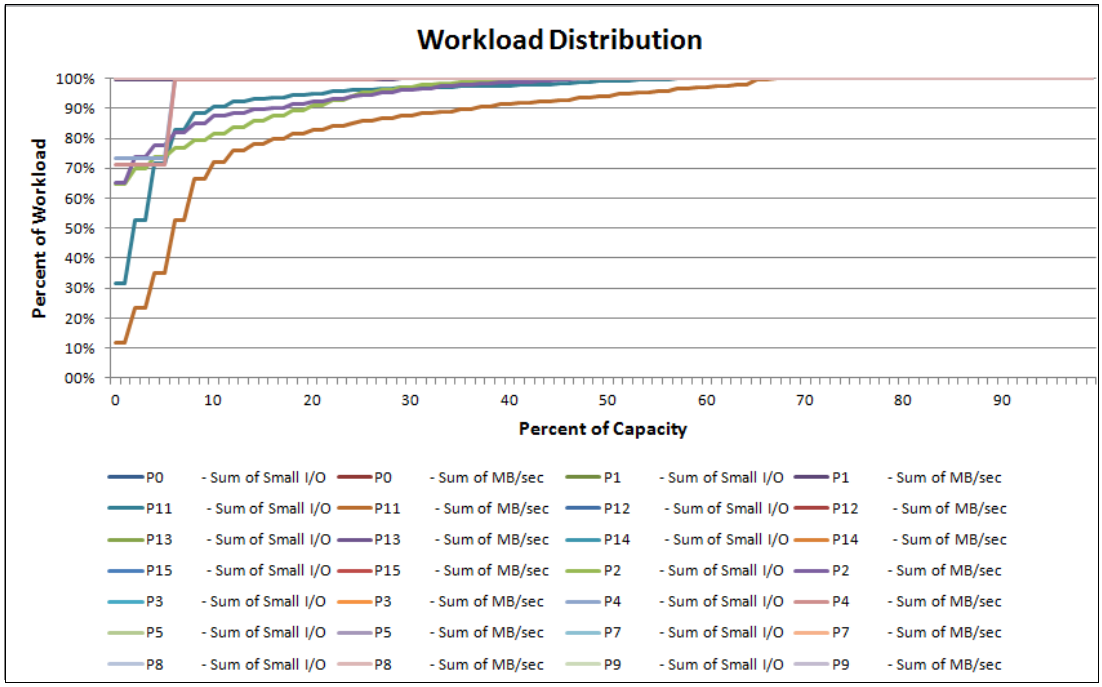



*Figure 5-13 Detailed skew curves for small IOPS and for MBps*

The Skew_Chart2 tab includes a chart that shows for each tier within each pool how the overall IOPS breaks down into small writes (smallwps), small reads (smallrps), large writes (largewps), and large reads (largerps).

An example of this chart is shown in Figure 5-14.

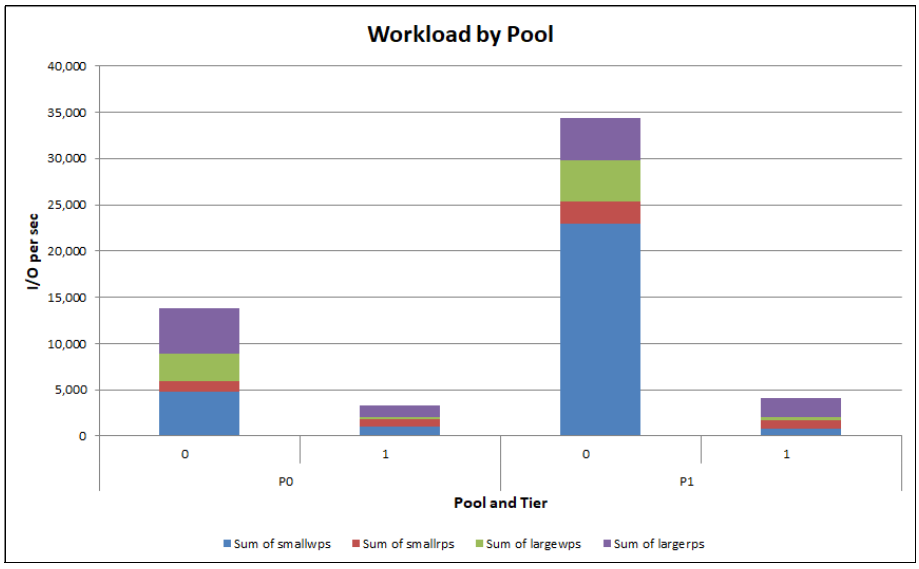



*Figure 5-14 Skew chart2: Workload by pool*

## 5.4.2 Easy Tier workload categorization

A second CSV file that is generated through Easy Tier processing with the latest microcode is named <*SFName*>_workload_ctg.csv. This file helps to categorize the heat data and compares it across tiers and pools.

This new categorization of an Easy Tier workload is defined in the following manner:

- Unallocated: Capacity *not* allocated
- Inactive: Data with zero activity
- Low activity: Data with activity below the cost-to-benefit ratio threshold for the total I/Os
- High Activity: Data with more than 0.1 IOPS per GB access density for small I/Os
- Sequential Activity: All data that is not included in the other categories

The following Easy Tier policy definitions are used in the reports:

- Managed by Easy Tier: Data that is being managed by Easy Tier
- Excluded from Nearline: Data that has been excluded from the nearline tier
- Assigned to this tier: Data that has been assigned to a tier
- Assigning to another tier: Data that is in the process of assigning to another tier
- Excluded from HDD: Data that has been excluded from the HDD tiers

Table 5-2 lists the columns in a workload_ctg.csv file.

*Table 5-2 Columns in the workload_ctg.csv file*

| Column name | Meaning |
|---|---|
| pool_id | DS8000 extent pool ID. |
| lun_id | DS8000 Volume ID. |
| tier | Tier for this row of the volume. There will be three rows per volume. |
| active_cap | Active capacity for this volume in GB. |
| inactive_cap | Inactive capacity for this volume. |
| low_activity_cap | Low activity capacity. |
| active_large_cap | Active large capacity. |
| active_read | IOPS for active read workload on this tier of the volume. |
| inactive_read | IOPS for inactive read workload on this tier of the volume. |
| low_activity_read | IOPS for low activity read workload on this tier of the volume. |
| active_large_read | IOPS for active large read workload on this tier of the volume. |
| active_write | IOPS for active write workload on this tier of the volume. |
| inactive_write | IOPS for inactive write workload on this tier of the volume. |
| low_activity_write | IOPS for low activity write workload on this tier of the volume. |
| active_large_write | IOPS for active large write workload on this tier of the volume. |
| active_kib | Active extents' bandwidth. |
| inactive_kib | Inactive extents' bandwidth. |
| low_activity_kib | Low activity extents' write IOPS. |
| active_large_kib | Active large extents' write IOPS. |

| Column name | Meaning |
|---|---|
| active_heat | Active extents' average response time. |
| inactive_heat | Inactive extents' average response time in ms. |
| low_activity_heat | Low activity extents' average response time in ms. |
| active_large_heat | Active large extents' average response time in ms. |
| unallocated_cap | The unallocated capacity of this pool on this tier. |
| volume_status | The volume assign status. |

Figure 5-15 shows the data structure of a `workload_ctg.csv` file. Information is available on all volumes and all tiers.

| index | pool_id | lun_id | tier | active_cap | inactive_cap | low_activity_cap | active_large_cap | active_read | inactive_read | low_activity_read |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | P4 | 0x1000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | P4 | 0x1000 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | P4 | 0x1000 | 2 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| 4 | P4 | 0x1001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | P4 | 0x1001 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | P4 | 0x1001 | 2 | 0 | 92 | 0 | 0 | 0 | 0 | 0 |
| 7 | P4 | 0x1004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | P4 | 0x1004 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | P4 | 0x1004 | 2 | 0 | 400 | 0 | 0 | 0 | 0 | 0 |
| 10 | P4 | 0x1005 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | P4 | 0x1005 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | P4 | 0x1005 | 2 | 0 | 400 | 0 | 0 | 0 | 0 | 0 |
| 13 | P0 | 0x2800 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | P0 | 0x2800 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 15 | P0 | 0x2800 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | P0 | 0x2801 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | P0 | 0x2801 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 18 | P0 | 0x2801 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | P0 | 0x2802 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | P0 | 0x2802 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 21 | P0 | 0x2802 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | P0 | 0x2803 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | P0 | 0x2803 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

*Figure 5-15 Example for a workload_ctg.csv file*

After the `workload_ctg.csv` file has been imported into the spreadsheet, several new tabs are created:

- Workload_by_Policy: For each tier within each extent pool, the capacity is broken down by policy.
- Workload_by_Activity: For each tier within each extent pool, the capacity is broken down into workload categories.
- Workload_Managed_by_Easy_Tier: The capacity that is managed by Easy Tier for each tier within each pool is broken down into workload categories.
- Workload_Excluded_from_Nearline: For each tier within each extent pool, the capacity that has been excluded from the nearline tier is broken down into workload categories.
- Workload_Excluded_from_HDD: For each tier within each extent pool, the capacity that is excluded from the HDD tier is broken down into workload categories.
- Workload_Assigned: For each tier within each extent pool, the capacity that has been assigned to a specific tier by the user is broken down into workload categories.
- Workload_Assigning: For each tier within each extent pool, the capacity that is in the process of assigning to that tier is broken down into workload categories.
- Workload_High_Activity: For each tier within each extent pool, the capacity that has been identified as having high activity is broken down by workload policies.

- Workload_Sequential_Activity: For each tier within each extent pool, the capacity that has been identified as having sequential activity is broken down by workload policies.
- Workload_Low_Activity: For each tier within each extent pool, the capacity that has been identified as having low activity is broken down by workload policies.
- Workload_Inactive: For each tier within each extent pool, the capacity that has been identified as inactive is broken down by workload policies.
- Workload_WriteIO: For each tier within each extent pool, the write activity is broken down.

An example chart for the Workload_by_Activity can be found in Figure 5-16.

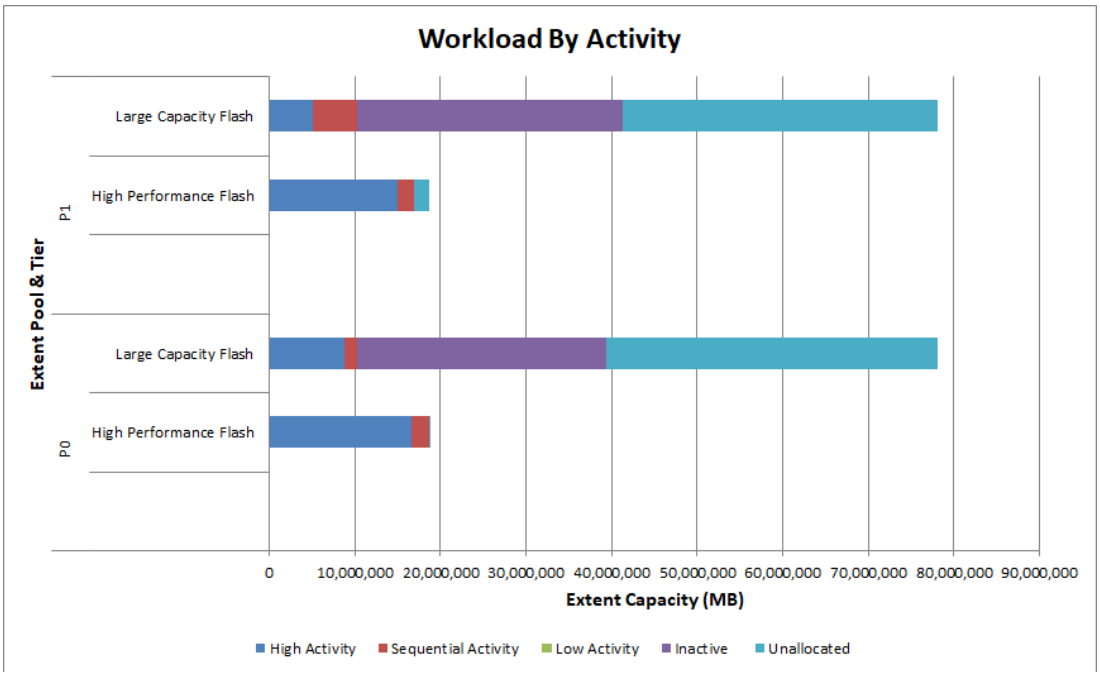



*Figure 5-16 Workload by Activity*

## 5.4.3 Easy Tier data movement report

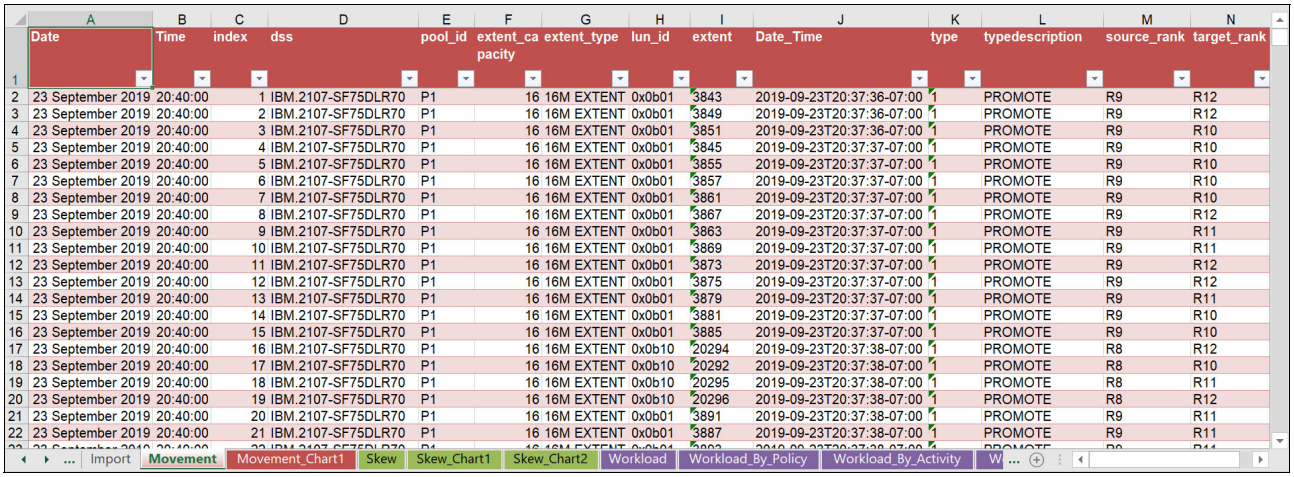
The third CSV file, which ends with `data_movement.csv`, provides an overview of all migration activities that occurred since the last heat map migration plan was generated. The plan is generated in the last 24 hours so that you can gain more understanding about the recent migrations, what triggered these migrations, and from where and to where each extent was migrated. This information is listed in detail with a time stamp for each extent. See Table 5-3.

*Table 5-3 Columns in the data_movement.csv file*

| Column name | Meaning |
|---|---|
| pool_id | ID of the pool where the movement happened |
| extent_capacity | The capacity of this extent in MiBs for FB or cylinders for CKD |
| extent_type | The type of extent |
| lun_id | DS8000 Volume ID |
| extent | The ID of the extent |

| Column name | Meaning |
|---|---|
| time | The time stamp |
| typedescription | Description of data movement type |
| source_rank | The rank ID where the data was moved from |
| target_rank | The rank ID where the data was moved to |

Figure 5-17 shows the data structure of this file. The time is the exact second. The volume and volume extent are listed with from where and to where (rank) it was migrated. The type of migration algorithm that triggered the migration at that particular time is shown. Migration types can be Promote, Demote, Swap, Auto-Rebalance, Warm-Demote, or types that are enforced by Easy Tier Application.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Date | Time | index | dss | pool_id | extent_capacity | extent_type | lun_id | extent | Date_Time | type | typedescription | source_rank | target_rank |
| 2 | 23 September 2019 | 20:40:00 | 1 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3843 | 2019-09-23T20:37:36-07:00 | 1 | PROMOTE | R9 | R12 |
| 3 | 23 September 2019 | 20:40:00 | 2 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3849 | 2019-09-23T20:37:36-07:00 | 1 | PROMOTE | R9 | R12 |
| 4 | 23 September 2019 | 20:40:00 | 3 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3851 | 2019-09-23T20:37:36-07:00 | 1 | PROMOTE | R9 | R10 |
| 5 | 23 September 2019 | 20:40:00 | 4 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3845 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 6 | 23 September 2019 | 20:40:00 | 5 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3855 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 7 | 23 September 2019 | 20:40:00 | 6 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3857 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 8 | 23 September 2019 | 20:40:00 | 7 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3861 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 9 | 23 September 2019 | 20:40:00 | 8 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3867 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R12 |
| 10 | 23 September 2019 | 20:40:00 | 9 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3863 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R11 |
| 11 | 23 September 2019 | 20:40:00 | 10 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3869 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R11 |
| 12 | 23 September 2019 | 20:40:00 | 11 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3873 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R12 |
| 13 | 23 September 2019 | 20:40:00 | 12 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3875 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R12 |
| 14 | 23 September 2019 | 20:40:00 | 13 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3879 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R11 |
| 15 | 23 September 2019 | 20:40:00 | 14 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3881 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 16 | 23 September 2019 | 20:40:00 | 15 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3885 | 2019-09-23T20:37:37-07:00 | 1 | PROMOTE | R9 | R10 |
| 17 | 23 September 2019 | 20:40:00 | 16 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b10 | 20294 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R8 | R12 |
| 18 | 23 September 2019 | 20:40:00 | 17 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b10 | 20292 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R8 | R10 |
| 19 | 23 September 2019 | 20:40:00 | 18 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b10 | 20295 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R8 | R11 |
| 20 | 23 September 2019 | 20:40:00 | 19 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b10 | 20296 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R8 | R12 |
| 21 | 23 September 2019 | 20:40:00 | 20 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3891 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R9 | R11 |
| 22 | 23 September 2019 | 20:40:00 | 21 | IBM.2107-SF75DLR70 | P1 | 16 | 16M EXTENT | 0x0b01 | 3887 | 2019-09-23T20:37:38-07:00 | 1 | PROMOTE | R9 | R11 |



*Figure 5-17 Data structure of a data_movement.csv file*

Using the worksheet that is provided, this table can also be graphical, for instance, listing the activities that occurred at certain times during the day, as shown in Figure 5-18.

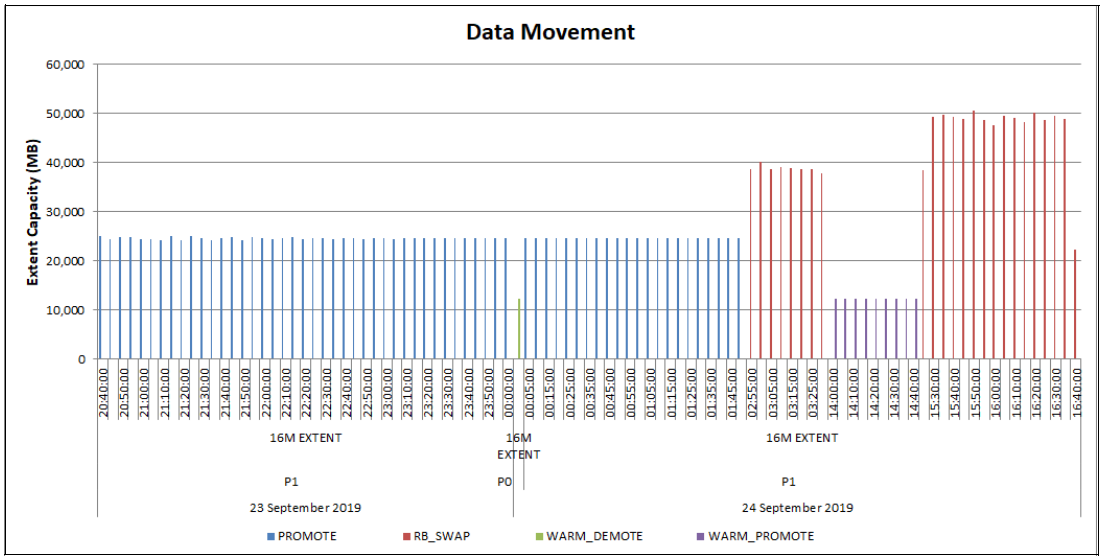



*Figure 5-18 Graphical interpretation of a data_movement.csv file*

## 5.5 Disk Magic and IBM Storage Modeller

Disk Magic as well as its successor IBM Storage Modeller are erformance modeling tools that are used by IBM and IBM Business Partners to model flash and disk storage system performance. They support are larger set of IBM storage systems. Contact your IBM representative or IBM Business Partner to evaluate a respective study.

Disk Magic was enhanced to support Easy Tier. Disk Magic can be used to model the benefit of moving a workload from a system that does not use Easy Tier to an Easy Tier system. Support is available for open systems, IBM Z, and IBM i models.

The workload can be modeled by manually entering the workload characteristics, **`iostat`** outputs, IBM Resource Measurement Facility™ (IBM RMF™) data files, IBM Spectrum Control or Storage Insight reports, and so on. An example is shown in Figure 5-19.

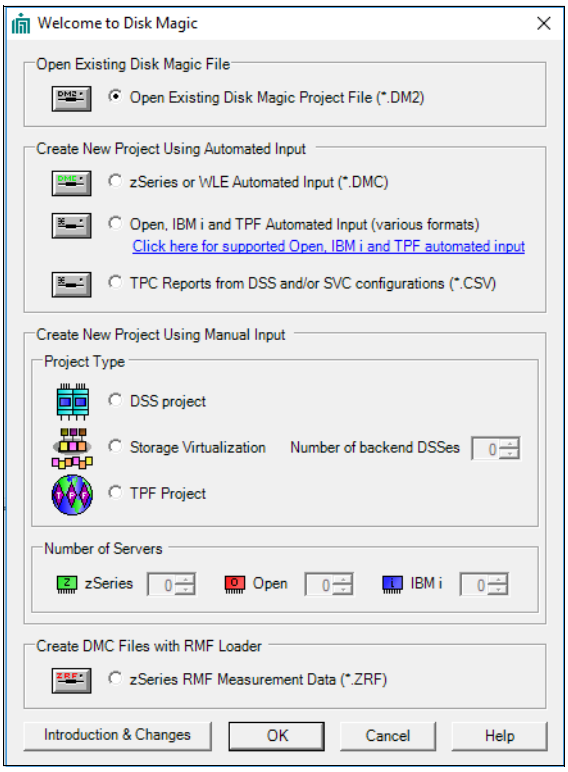



*Figure 5-19 Disk Magic welcome window*

If you manually enter the workload definition into the model, you see various options to define the workload: The number of ranks, the capacity and type of the drives, and the definition of the workload, as shown in Figure 5-20.

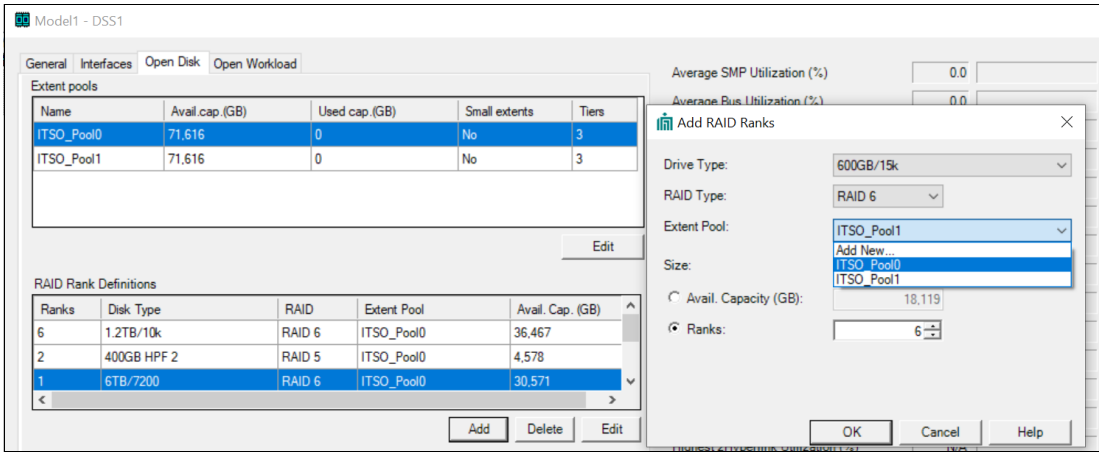



*Figure 5-20 Defining the disk system configuration window*

Easy Tier modeling for the DS8000 is on a skew-factor basis. The workload skew is the most important parameter to the model input because depending on the workload skew, various results can be achieved with the Easy Tier modeling.

The three predefined types of skew level for the workload are heavy, medium, and light skew, as shown in Figure 5-21 for random I/O.

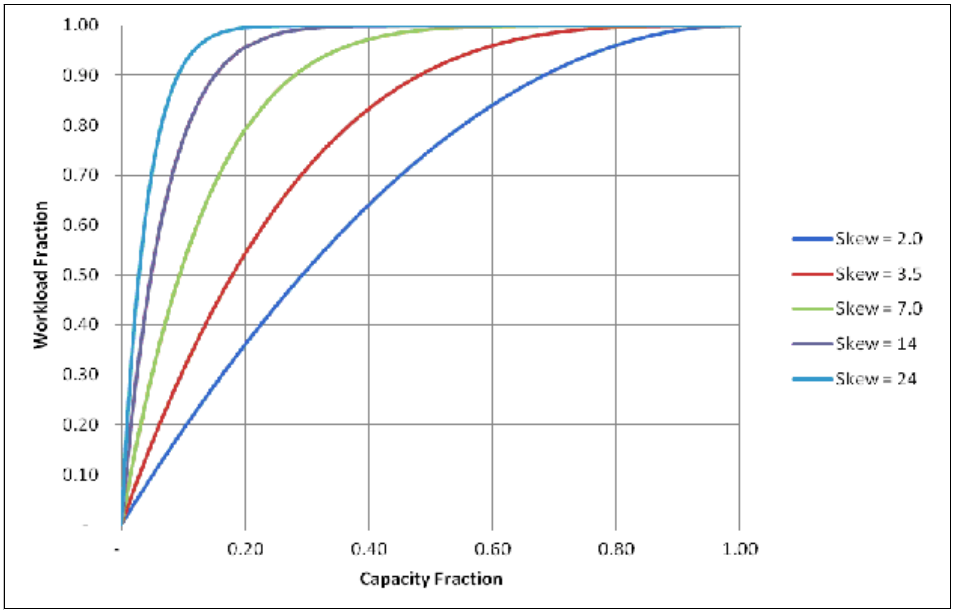



*Figure 5-21 Example of Disk Magic skew factors*

A heavily skewed workload has many I/Os bound to only a small portion of the data, which means that moving only a few extents to flash tier (the hottest extents) greatly improves performance. For a lightly skewed workload, a much smaller difference exists in the workload intensity on the busiest extents versus the most inactive extents. This situation means that the expected performance improvement is smaller.

When Easy Tier is selected, ensure that you choose the best match for your workload. Choosing the heavily skewed profile results in a more aggressive sizing, so you must be sure that the workload is indeed heavily skewed. To get a more conservative prediction, use the lightly skewed option.

The workload can benefit from the use of Easy Tier with flash drives when the skew factor is heavy. Before creating any proposals, the system's current workload and data placement must be analyzed in relationship to the skew factor. The higher the skew factor, the more the system benefits from Easy Tier and flash drive technology.

### 5.5.1 Reading the skew curve into Disk Magic

When you do not specify otherwise, the simulation tool Disk Magic works with predefined skew levels when it models multi-tier pools, which have the following skew values (for random I/O):

- For IBM Z or Open Servers:
  - Very high (24.0)
  - High (14.0)
  - Intermediate (7.0)
  - Low (3.5)
  - Very low (2.0)
- For IBM i Servers, the value is predefined:
  - Very low (2.0)

**Note:** The default value (random I/O) for Open Systems and IBM Z is High (14.0).

These values are often good first guesses. However, the values differ in an actual client environment. Use the `..._skew_curve_...csv` file to read the actual skew curve into Disk Magic instead because it is measured and established from the real workload that runs on the DS8000.

Use the following process:

1. Create a model with the exact configuration of your DS8000, regarding its hardware (amount and type of ranks, cache, host bus adapters (HBAs), and so on). From the IBM Spectrum Control measurements of a peak workload, enter the workload profile (reads, writes, and block sizes) and measured response time into Disk Magic. Then, click **Base** to establish a model base with a predefined skew level.
2. Now, overwrite this predefined skew level. Open the Easy Tier Settings window (Figure 5-22) in the Disk Magic model. Click **Read Heatmap**.

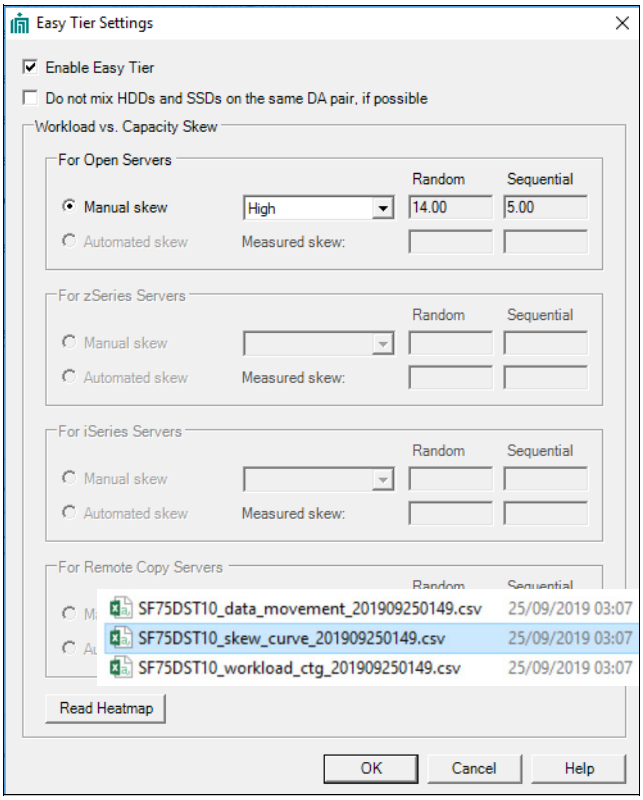



*Figure 5-22 Reading in a skew_curve.csv file*

3. From the menu, select the `.._skew_curve_...csv` file in the `Data_files` directory. You are prompted to confirm replacing the current disk storage system (DSS) name and model with the new IBM.2107-*xxxxxxx* model and heat map.

As a result, we override the former assumed skew level value of 14.00 (random) and continue from now on with the actual skew level of 2.13 (random) (Figure 5-23).

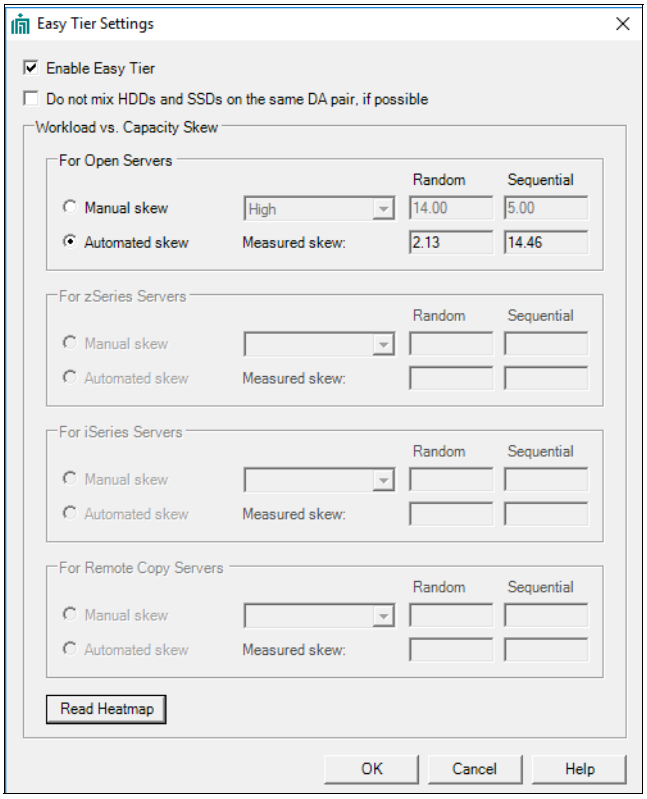



*Figure 5-23 The actual skew curve is loaded*

Because this exercise can lead to a changed response time, you might want to establish another base again with the real response time, but based on this real skew level curve.

4. From here, you can create more models, such as adding more drive ranks to a certain tier and varying the RPM or RAID types. Solve these models to see the projected effects of the potential upgrades and changes.

For help with this process, use the help system in the Disk Magic tool or ask your IBM technical sales support.

A

# IBM DS8000 Easy Tier and IBM z/OS system-managed storage

This appendix compares system-managed storage (SMS) in z/OS configurations with IBM System Storage Easy Tier Automatic Data Relocation (ADR). This brief review of system-managed storage shows how Easy Tier, with its ability to dynamically relocate extents, can complement, or even be a substitute, for SMS.

This appendix includes the following topics:

- z/OS system-managed storage
- Easy Tier storage management
- Hybrid pools in z/OS system-managed storage environments
- z/OS and thin-provisioning
- z/OS Db2 synergy

# z/OS system-managed storage

The concept of storage tiering and managing various storage technologies in an automated fashion was implemented in IBM MVS/Data Facility Product (DFP) V3.0. It was introduced to the marketplace in 1989 as system-managed storage (SMS) and Data Facility Storage Management Subsystem (DFSMS), which consist of the following components:

- MVS/DFP: Software kernel of SMS with volume selection in addition to “I/O drivers” that provide data and device management
- DFSMShsm: Hierarchical storage manager
- DFSMSdss: Highly tuned data set services to copy and move data from any device to any other device
- DFSMSrmm: Removable media manager that keeps track where all the files are on tape

DFSMS is a part of z/OS.

SMS provides support for any level of tiered storage, including disk, tape, and optical storage systems.

By using Automatic Class Selection (ACS) routines, SMS automatically assigns service level constructs to newly allocated data sets. These constructs imply certain service level requirements. Storage Class constructs are used to indicate service levels in terms of required response times or bandwidth requirements. Based on these Storage Class attributes, SMS evaluates by using the Storage Group ACS routine, where storage tier best matches the service level requirements and selects the appropriate volumes.

Figure A-1 shows the concept of SMS with its ACS routines to assign service level-based constructs to a created file, data set, or an object.

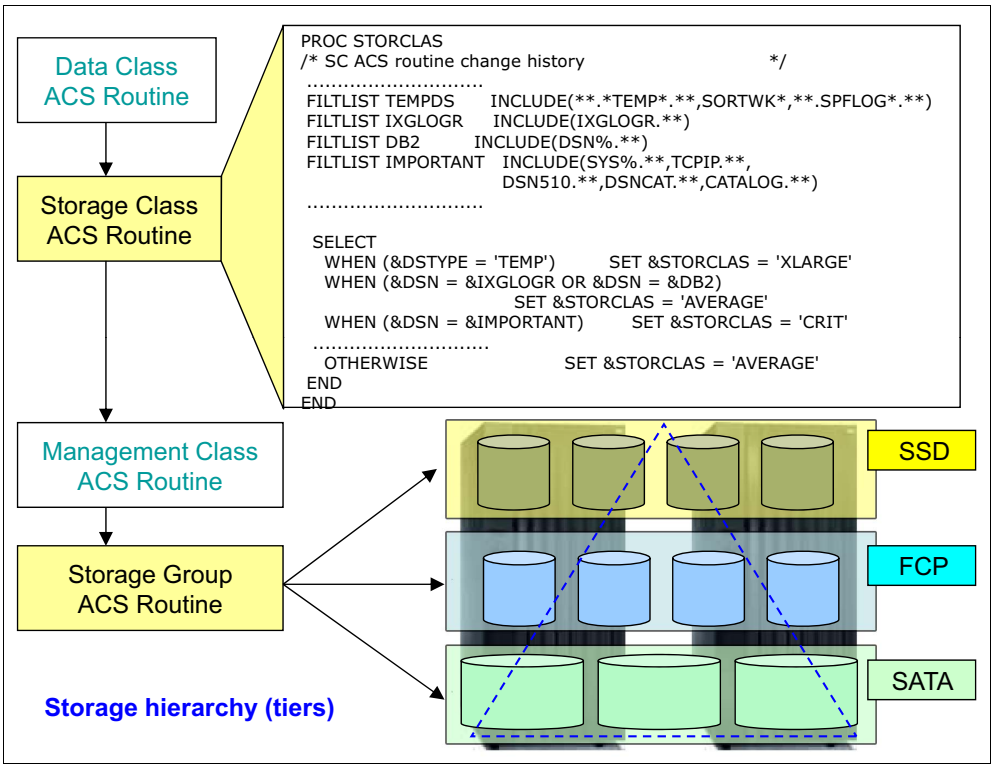



*Figure A-1 System-managed storage in z/OS: Assign a performance service level to a new file*

Particular data sets, such as system data sets or IBM Db2® data sets that are critical to performance, are automatically considered for placement on the best performing devices that are available in the configuration.

Figure A-2 shows a fragment of a Storage Group ACS routine that maps the service level requirements (expressed through Storage Classes) to the existing storage subsystems and technology within the configuration.

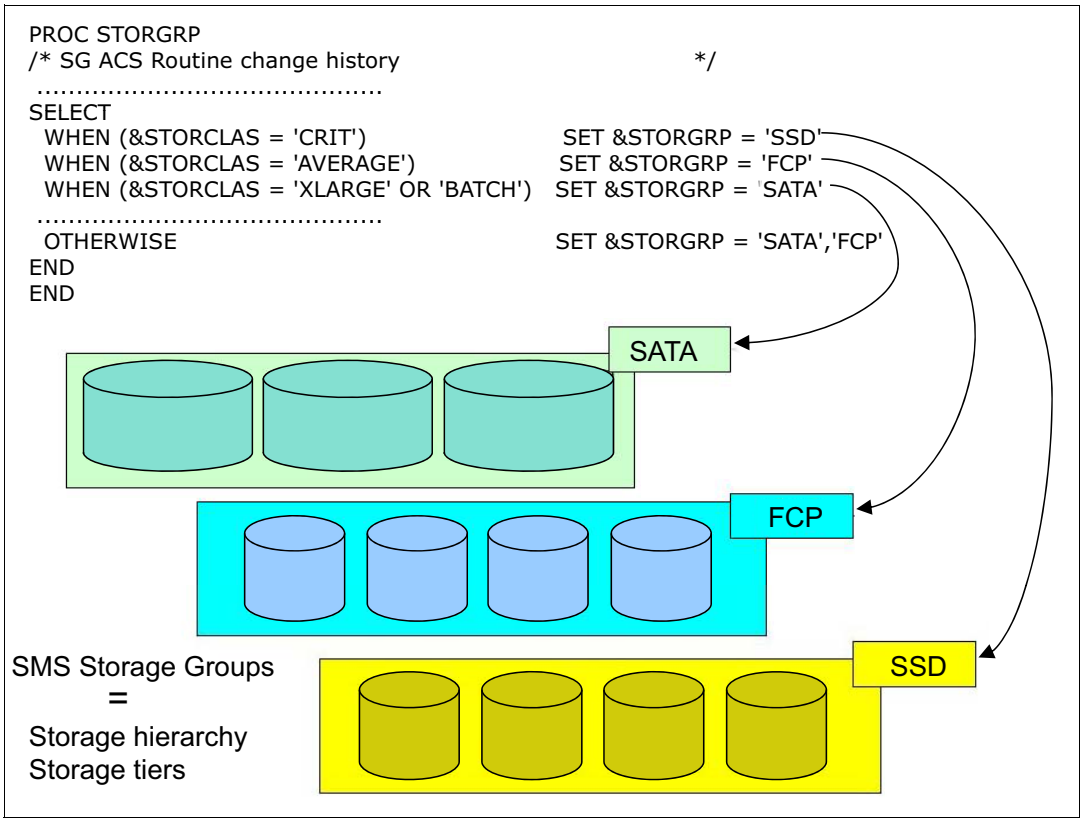



*Figure A-2 Map performance service level requirements to I/O environment*

The actual volume selection is a sophisticated process in DFSMSdfp and takes storage technologies into consideration to map the file to the correct storage tier.

For the purpose of this publication, granularity is at the level of a z/OS data set or file.

A data set can be as small as a track (56 KiB) or in the range of terabytes. Volumes and technology are selected for the entire data set. Placing a large data set on expensive high-performance devices can affect many smaller data sets because the high-performance space is taken. Additionally, DFSMS does not generally automatically move data sets through the storage hierarchy when the need for high performance changes over time.

## Easy Tier storage management

Compared to SMS, the advantage of Easy Tier is the finer granularity (one extent) that it uses to manage the storage tiers within the DS8000. An extent can be 1113 cylinders large, if large extents are used in the extent pool or 21 cyl small, if small extents are used. Also, Easy Tier can dynamically manage the I/O activity and adapt data placement over time.

You might still need to ensure that certain data must be on the most responsive device, no matter how active (or inactive) the data set is. You might want to dedicate a flash rank in an isolated extent pool and deliberately create only volumes that exclusively hold these data sets and no other data. Or you can pin a volume to the flash tier. However, this method is manual and implies intensive and time-consuming system management tasks, which can be done only by highly skilled experts. From this perspective, Easy Tier is a better approach.

## Hybrid pools in z/OS system-managed storage environments

In a z/OS configuration, you can combine the SMS approach with its storage groups and the Easy Tier capabilities, as shown in Figure A-3.

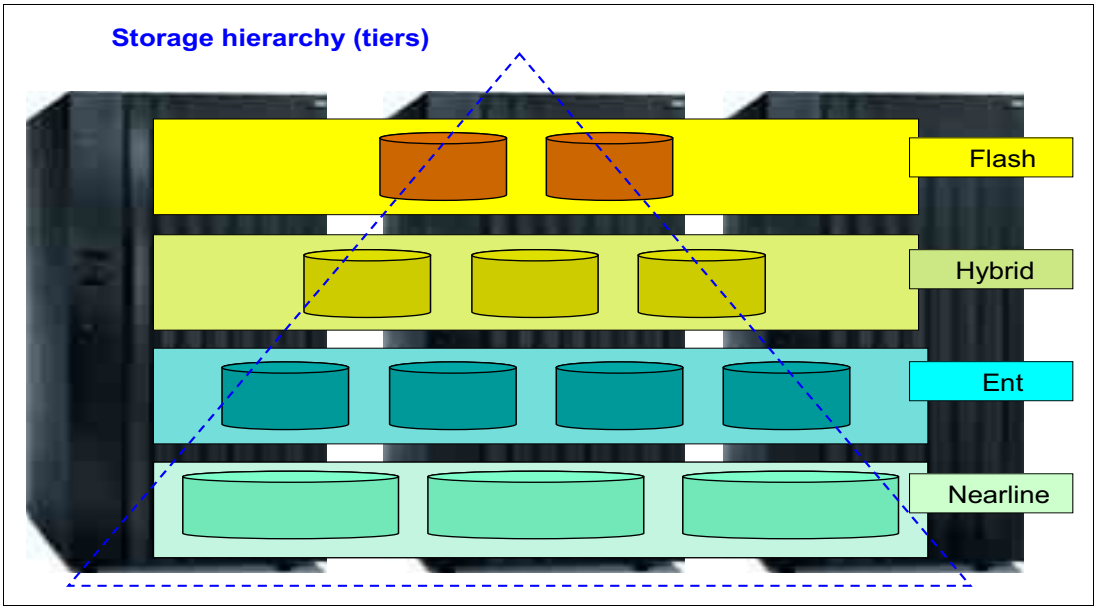



*Figure A-3 Include the new storage group in system-managed configurations*

Isolate *flash* logical volumes for particularly demanding service levels for specific data and applications in the configuration. SMS strictly controls the applications that can have data placed in this expensive storage group. The next storage group, *HYBRID*, combines enterprise and flash ranks. Within that storage group, Easy Tier can be used to relocate data at an extent level to flash or enterprise drives. In the SMS *HYBRID* storage group, you can combine all hybrid extent pools (even from multiple DS8000 systems).

The next level in this storage hierarchy might be a storage group named *Ent*, which contains all the logical volumes that are in extent pools with enterprise drives for all of your DS8000 systems.

Finally, the *nearline* storage group is composed of all the extent pools with nearline drives.

## z/OS and thin-provisioning

With DFSMShsm, you can do effective capacity management in z/OS. With the support of thin-provisioned extent space efficient volumes for production use, the DS8000 provides another mean to more efficient capacity management. Thin-provisioned volumes consume only that space on disk that has data on it. Thin-provisioning allows over-provisioning. You can define large volumes, perhaps all of the same size, where the sum of capacity of these virtual volumes is larger than the physically available storage.

The use of thin-provisioned volumes can save even more storage than DFSMShsm can. Thin-provisioned volumes can be created in extent pools with large 1113 cyl extents or in extent pools with small 21 cyl extents. The use of extent pools with small extents makes thin-provisioning even more efficient.

Although thin-provisioning benefits most from volumes that are not managed with HSM, you can still use it for HSM managed volumes. However, in this case you might need to adjust your migration thresholds when you define large virtual volumes.

## z/OS Db2 synergy

On a DS8000, you can use the Easy Tier Application API with Db2 for IBM Z. Applications are able to proactively advise (by communicating *hints*) Easy Tier on the intended use of data sets.

For more information, see *DS8870 Easy Tier Application*, REDP-5014.

# Related publications

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

## IBM Redbooks publications

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

- *IBM DS8900 Architecture and Implementation*, SG24-8456
- *IBM DS8880 Architecture and Implementation (Release 8.3)*, SG24-8323
- *DS8870 Easy Tier Application*, REDP-5014
- *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015
- *IBM DS8000 High-Performance Flash Enclosure Gen2*, REDP-5422

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

`http://www.ibm.com/redbooks`

## Other publications

These publications are also relevant as further information sources:

- *DS8000 Command-Line Interface User's Guide*, SC27-8526
- *DS8900F Introduction and Planning Guide,* SC27-9560
- *DS8880 Introduction and Planning Guide,* GC27-8525

## Online resources

These websites are also relevant as further information sources:

- IBM DS8000 Knowledge Center:

  `https://www.ibm.com/support/knowledgecenter/ST5GLJ_8.5.3`
- IBM data storage feature activation (DSFA):

  `http://www.ibm.com/storage/dsfa`
- IBM System Storage Interoperation Center (SSIC):

  `http://www.ibm.com/systems/support/storage/ssic/interoperability.wss`
- IBM Techdocs Library - The IBM Technical Sales Library:

  `http://www.ibm.com/support/techdocs/atsmastr.nsf/Web/Techdocs`

# Help from IBM

IBM Support and downloads:

http://www.ibm.com/support

IBM Global Services:

http://www.ibm.com/services

REDP-4667-08

ISBN 0738457159

Printed in U.S.A.