Redbooks

ibm.com/redbooks

Implementing IBM FlashSystem V9000 AE3

Detlef Helmbrecht

James Cioffi

Jon Herd

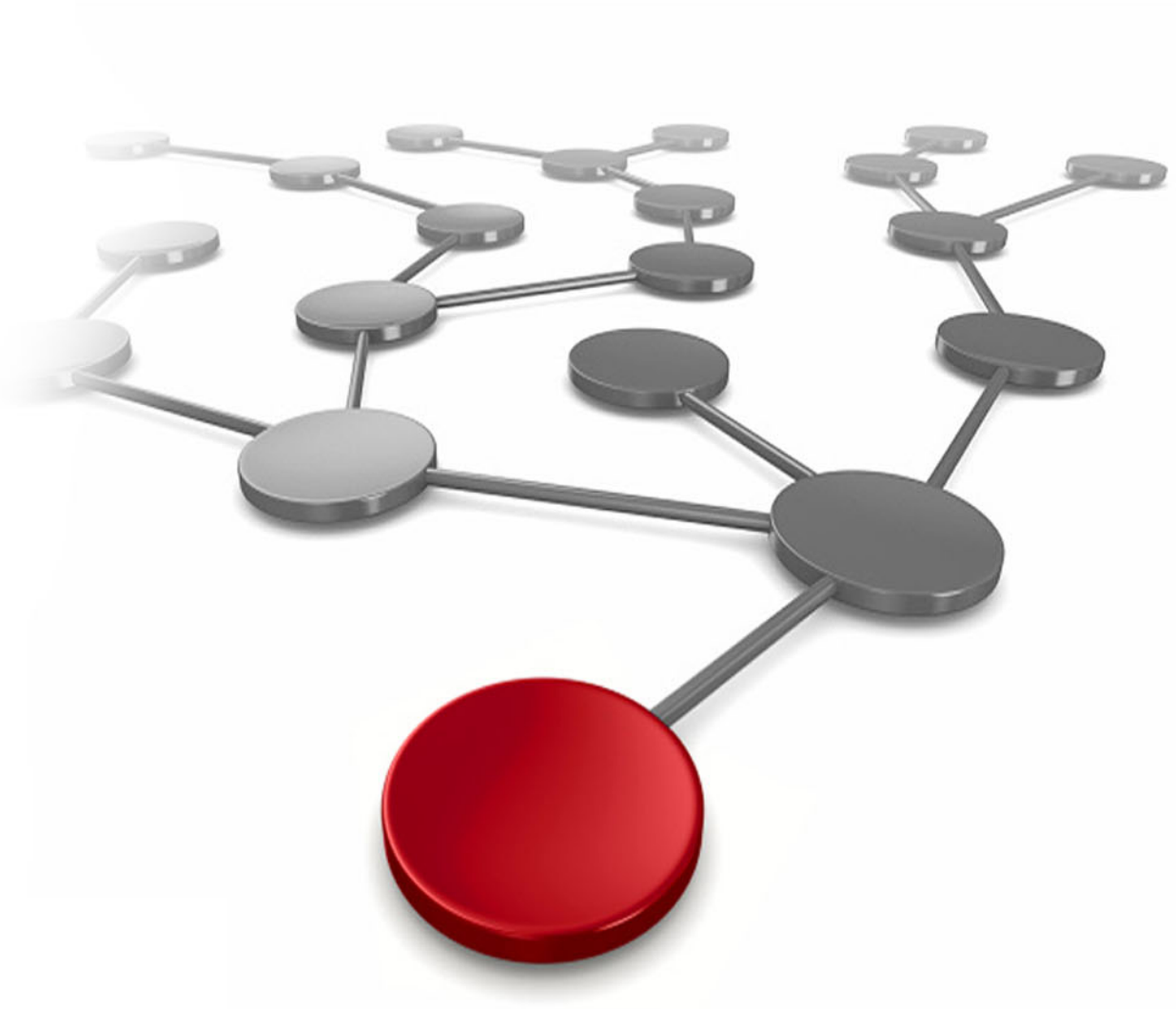
Jeffrey Irving

Christian Karpp

Volker Kiemes

Carsten Larsen

Adrian Orben



IBM

Redbooks

International Technical Support Organization

Implementing IBM FlashSystem V9000 AE3

April 2018

SG24-8413-00

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

First Edition (April 2018)

This edition applies to IBM FlashSystem V9000 AE3.

Contents

Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®
DS8000®
Easy Tier®
FlashCopy®
GDPS®
Geographically Dispersed Parallel Sysplex™
HyperSwap®
IBM®
IBM FlashCore®
IBM FlashSystem®
IBM Spectrum™
IBM Spectrum Control™
IBM Spectrum Storage™
IBM Spectrum Virtualize™
IBM Watson®
MicroLatency®
Parallel Sysplex®
Passport Advantage®
PowerHA®
Real-time Compression™
Redbooks®
Redpaper™
Redbooks (logo) ®
Storwize®
System Storage®
Tivoli®
Watson™
XIV®
z/OS®

The following terms are trademarks of other companies:

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

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

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

Preface

The success or failure of businesses often depends on how well organizations use their data assets for competitive advantage. Deeper insights from data require better information technology.

As organizations modernize their IT infrastructure to boost innovation rather than limit it, they need a data storage system that can keep pace with several areas that affect your business:

- Highly virtualized environments
- Cloud computing
- Mobile and social systems of engagement
- In-depth, real-time analytics

Making the correct decision on storage investment is critical. Organizations must have enough storage performance and agility to innovate when they need to implement cloud-based IT services, deploy virtual desktop infrastructure, enhance fraud detection, and use new analytics capabilities. At the same time, future storage investments must lower IT infrastructure costs while helping organizations to derive the greatest possible value from their data assets.

The IBM® FlashSystem V9000 is the premier, fully integrated, Tier 1, all-flash offering from IBM. It has changed the economics of today's data center by eliminating storage bottlenecks. Its software-defined storage features simplify data management, improve data security, and preserve your investments in storage. The IBM FlashSystem® V9000 SAS expansion enclosures provide new tiering options with read-intensive SSDs or nearline SAS HDDs.

IBM FlashSystem V9000 includes IBM FlashCore® technology and advanced software-defined storage available in one solution in a compact 6U form factor. IBM FlashSystem V9000 improves business application availability. It delivers greater resource utilization so you can get the most from your storage resources, and achieve a simpler, more scalable, and cost-efficient IT Infrastructure.

This IBM Redbooks® publication provides information about IBM FlashSystem V9000 Software V8.1. It describes the core product architecture, software, hardware, and implementation, and provides hints and tips.

The underlying basic hardware and software architecture and features of the IBM FlashSystem V9000 AC3 control enclosure and on IBM Spectrum™ Virtualize 8.1 software are described in these publications:

- *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414
- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933

Using IBM FlashSystem V9000 software functions, management tools, and interoperability combines the performance of IBM FlashSystem architecture with the advanced functions of software-defined storage to deliver performance, efficiency, and functions that meet the needs of enterprise workloads that demand IBM MicroLatency® response time.

This book offers IBM FlashSystem V9000 scalability concepts and guidelines for planning, installing, and configuring, which can help environments scale up and out to add more flash capacity and expand virtualized systems. Port utilization methodologies are provided to help you maximize the full potential of IBM FlashSystem V9000 performance and low latency in your scalable environment.

This book is intended for pre-sales and post-sales technical support professionals, storage administrators, and anyone who wants to understand how to implement this exciting technology.

Authors

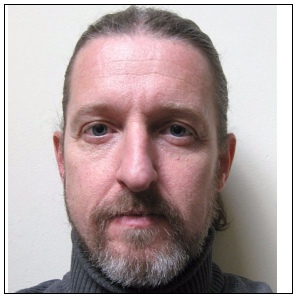
This book was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO), San Jose Center.

Detlef Helmbrecht is an Advanced Technical Skills (ATS) IT Specialist working for the IBM Systems. He is in the EMEA Storage Competence Center (ESCC) in Kelsterbach, Germany. Detlef has over 30 years of experience in IT, performing various roles including software engineer, sales, and solution architect. His areas of expertise include high-performance computing (HPC), disaster recovery, archiving, application tuning, and FlashSystem.

James Cioffi is a Senior Product Engineer with IBM FlashSystems. He has 33 years with IBM, mostly in technical support. At various times, he has been a course developer and trainer, project manager, technical support lead, and programmer. His current role is Product Engineering and the focal point for the FlashSystem 900. His current job involves providing level 3 support for FlashSystems, which includes writing and reviewing technical documentation, IBM Redbooks publications, and FlashSystem training.

Jon Herd is an IBM Storage Technical Advisor working for the ESCC, Germany. He covers the United Kingdom and Ireland, advising customers on a portfolio of IBM storage products, including IBM FlashSystem products. Jon has been with IBM for more than 40 years, and has held various technical roles, including Europe, Middle East, and Africa (EMEA)-level support on mainframe servers and technical education development. He holds IBM certifications in Supporting IT Solutions at an expert level, and Actualizing IT Solutions at an experienced level. He is also a certified Member of the British Computer Society (MBCS) Chartered IT Professional (CITP), and a certified Member of the Institution of Engineering and Technology (MIET).

Jeffrey Irving is an IBM Storage Technical Advisor who supports multiple storage products. He has more than 30 years of experience in IT. Beginning his career with AT&T Bell Labs testing UNIX operating systems, he moved on to supercomputers, leading several software test teams in the process. Jeff has been with IBM 15 years, specializing in storage virtualization and customer support. He also leads the Technical Advisor team for FlashSystem products. Jeff works in Wisconsin, US.



Christian Karpp is a Senior IT Specialist and Security Consultant at IBM in Mannheim, Germany. He has been with IBM for 22 years, with the last 10 years working as a Client Technical Specialist at the Storage Brand's technical sales force. Christian has an extensive background in IT, architecture, and concepts, with IBM Spectrum products, with IBM Spectrum Virtualize™ and IBM Storwize® being his current focus. He studied Computer Science and graduated from the University of Applied Science in Mannheim, Germany and is the author of "qperf", which is a tool to collect and analyze performance data of Storwize and Virtualize family members.

Volker Kiemes is an IBM Certified Senior IT Specialist in the Technical Support Organization in Kelsterbach, Germany. Volker has 20 years experience in disk drive storage technology. Since 2001, Volker works for IBM Enterprise Class Storage in support, development, testing, product documentation, and education. Volker shifted his primary work scope as Product Field Engineer in 2013 from IBM DS8000® to IBM FlashSystem, which are Enterprise Class storage products.

Carsten Larsen is an IBM Certified Senior IT Specialist working for the Technical Services Support organization in IBM Denmark and delivering consultancy services to IBM clients within the storage arena. Carsten joined IBM in 2007 when he left HP, where he worked with storage arrays and UNIX for 10 years. While working for IBM, Carsten obtained several Brocade and NetApp certifications. Carsten is the author of several IBM Redbooks publications.

Adrian Orban is Product Field Engineer for IBM Flash System. He joined IBM in 2003. Beginning in 2006, he is a member of the Storage Support Team and is responsible for IBM high-end disk products, such as IBM DS8000 and IBM XIV®. During this time, Adrian completed several storage-related certifications. Adrian holds a Bachelor's degree in Business Informatics at University of Applied Science Mainz, Germany.

The project that produced this publication was managed by Detlef Helmbrecht.

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

Bertrand Dufrasne
International Technical Support Organization

Erik Franz
IBM Systems, EMEA Storage Competence Center

Gordon Bloxham
Shelly Howrigon
Lee Sanders
Mark Sistrunk
Matt Smith
Michael Warburton
IBM Systems, Storage

Robert Wallis, storage software
IBM Systems, Storage Software

Philip Clark
Dave Gimple
Kelly Groff
Matt Key
Kim Miller
Brent Yardley
AI IBM Watson®
IBM Systems, Flash System

Thanks to the following authors of the previous edition of this book:

- Marcela Adan
- Christophe Fagiano
- Jana Jamsek
- Renato Santos
- James Thompson

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 two - six weeks in length, and you can participate either in person or as a remote resident working from your home base.

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us.

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

- Use the online **Contact us** review 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:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

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

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

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

- Stay current on recent Redbooks publications with RSS Feeds:

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

1

IBM FlashSystem V9000 introduction

This chapter introduces the IBM FlashSystem V9000 storage system and its core values, benefits, and technological advantages.

This chapter includes the following topics:

- IBM FlashSystem V9000 overview
- IBM FlashSystem family
- FlashSystem V9000 architecture
- Advanced software features

The detailed description of the IBM FlashSystem V9000 AC3 control enclosure and on Spectrum Virtualize 8.1 software are described in these publications:

- *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414
- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933

1.1 IBM FlashSystem V9000 overview

With the release of FlashSystem V9000 Software version 8.1, extra functions and features are available, including support for new and more powerful FlashSystem V9000 storage enclosure model AE3. Software features added include GUI enhancements, new dashboard, Support Assistance and inline hardware compression in the model AE3 storage enclosure.

1.1.1 Hardware components

Each IBM FlashSystem V9000 AC3 control enclosures is an individual server in an IBM FlashSystem V9000 clustered system on which the IBM FlashSystem V9000 software runs. These control enclosures are organized into I/O groups; each I/O group is made up of a pair of AC3 control enclosures.

An I/O group takes the storage that is presented to it by the AE3 storage enclosures as MDisks, adds these to pools, and translates the storage into logical disks (volumes) that are used by applications on the hosts. An AC3 control enclosure is in only one I/O group and provides access to the volumes in that I/O group.

These are the core IBM FlashSystem V9000 components:

- FlashSystem V9000 AC3 control enclosure:
  - Power supply units
  - Battery modules
  - Fan modules
  - Interface cards
- FlashSystem V9000 AE3 storage enclosure:
  - Control canisters
  - IBM MicroLatency modules
  - Power supply units
  - Battery modules
  - Fan modules
  - Interface cards
- SAN switches for flexible building blocks
- FlashSystem V9000 expansion enclosures (SAS attached)

The IBM FlashSystem V9000 delivers high capacity for the enterprise data center. IBM FlashSystem V9000 uses a fully featured and scalable all-flash architecture that performs at up to 5.2 million input/output operations per second (IOPS) with IBM MicroLatency modules, is scalable to 80 gigabytes per second (GBps), and delivers internal flash capacity up to 1.7 PB (maximum effective with inline hardware compression) or 7.2 PB (effective, assuming 5:1 data reduction using software compression, the IBM Real-time Compression™ technology (RtC).

Note: Applications may not benefit from the lowest FlashSystem 900 AE3 latency if data is compressed using RtC, they may benefit a higher compression rate. Details can be found in section “Physical and effective capacity based on compression rates” in the IBM Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

Figure 1-1 shows the IBM FlashSystem V9000 front view. The 12 IBM MicroLatency modules are in the middle of the unit.

Figure 1-1 IBM FlashSystem V9000

Beyond its base all-flash architecture, the IBM FlashSystem V9000 also addresses tiered capabilities as described below.

The IBM FlashSystem V9000 Control Enclosure model AC3 delivers advanced software functionality with Spectrum Virtualize 8.1. With the release of IBM FlashSystem V9000 Software version 8.1, extra functions and features are available, including new Graphical User Interface (GUI) and support for new and more powerful IBM FlashSystem V9000 storage enclosure model AE3 with inline hardware compression.

SAS-based small form factor (SFF) and large form factor (LFF) expansion enclosures are providing a mixture of nearline SAS hard disk drives (HDDs) and flash managed disks (MDisks) in a pool, which can be used for IBM Easy Tier®.

Up to 20 serial-attached SCSI (SAS) expansion enclosures are supported per IBM FlashSystem V9000 controller pair, providing up to 240 drives with expansion enclosure model 12F, and up to 480 drives with expansion enclosure model 24F.

The new IBM FlashSystem V9000 LFF expansion enclosure model 92F supports up to 92 drives per enclosure, with a mixture of HDD and SSD drives in various capacities.

Using its flash-optimized design, IBM FlashSystem V9000 can provide response times of 180 microseconds. It delivers better acquisition costs than a high-performance spinning disk for the same effective capacity while achieving five times the performance when using RtC, making it ideal for environments that demand extreme performance.

The IBM FlashSystem V9000 LFF expansion enclosure model 12F offers tiering options with 8 TB or 10 TB nearline SAS hard disk drives (HDDs).

The IBM FlashSystem V9000 SFF expansion enclosure model 24F offers tiering options with low-cost solid-state drives (SSDs).

Figure 1-2 shows IBM FlashSystem V9000 expansion enclosure model 12F.

Figure 1-2 IBM FlashSystem V9000 expansion enclosure model 12F

Figure 1-3 shows IBM FlashSystem V9000 expansion enclosure model 24F.

Figure 1-3 IBM FlashSystem V9000 expansion enclosure model 24F

Figure 1-4 shows IBM FlashSystem V9000 expansion enclosure model 92F.

Figure 1-4 IBM FlashSystem V9000 expansion enclosure model 92F

IBM FlashSystem V9000 with the new model EA3 storage enclosure has built in and always on inline hardware compression.

Furthermore with optional RtC technology, IBM FlashSystem V9000 provides up to 180 terabtyes (TB) usable flash capacity, and up to 900 TB effective flash capacity in only 6U (assumes a 5:1 reduction in data using RtC). This scales to 1.44 PB usable flash capacity and up to 7.2 PB effective flash capacity in only 34U. The FlashSystem V9000 inline hardware compression will not compress already compressed RtC workloads.

IBM FlashSystem V9000 delivers enterprise-class advanced storage capabilities, including these among others:

- IBM Real-time Compression Accelerators
- IBM Easy Tier
- Thin provisioning
- Copy services
- Data virtualization
- IBM HyperSwap® Split-Clusters
- Highly available configurations
- N_Port ID Virtualization (NPIV) support
- Distributed redundant array of independent disks (DRAID) Component in Doubt (CID)
- iSCSI virtualization support
- SKLM Encryption support (at code level 8.1)
- Transparent Cloud Tiering (at code level 8.1)

Advanced data services that are provided include copy services, mirroring, replication, external virtualization, IBM HyperSwap capabilities, Microsoft Offloaded Data Transfer (ODX)-capable features, and VMware vSphere Storage application programming interfaces (APIs) Array Integration (VAAI) support.

Host interface support includes 8 gigabit (Gb) and 16 Gb Fibre Channel (FC), and 10 Gb Fibre Channel over Ethernet (FCoE) or Internet Small Computer System Interface (iSCSI). Advanced Encryption Standard (AES) 256 hardware-based encryption adds to the rich feature set.

Note: The AC3 control enclosure supports only the 16 Gb 4-port Fibre Channel adapter, however it can negotiate down to both 8 Gb and 4 Gb, so this book uses the reference “16/8/4” to indicate that the three speeds are supported by this adapter.

IBM FlashSystem V9000, including its IBM MicroLatency module (flash modules), is covered by up to seven years of total hardware support through the applicable warranty period.

1.1.2 Power requirements

IBM FlashSystem V9000 is *green data center friendly*. The IBM FlashSystem V9000 building block uses only 2700 watts of power under maximum load, and uses six standard single phase (100v - 240v) electrical outlets, two per AC3 storage controller and two for the AE3 storage enclosure. Plan to attach each of the two power supplies in each of the enclosures, to separate main power supply lines.

The IBM FlashSystem V9000 maximum configuration, with four scalable building blocks and four additional AE3 storage enclosures, consumes 16 kilo watts of power under maximum load.

AE3 storage enclosure: The 1300 W power supply provides the AE3 storage enclosure with high power to run at maximum performance for longer durations during power supply servicing, resulting in more predictable performance under unexpected failure conditions. The operating environment electrical power is as follows:

- Voltage range: 100-240 V AC
- Frequency: 50 - 60 Hz

Using two power sources provides power redundancy. The suggestion is to place the two power supplies on different circuits.

Important: The power cord is the main power disconnect. Ensure that the socket outlets are located near the equipment and are easily accessible.

1.1.3 Physical specifications

The IBM FlashSystem V9000 installs in a standard 19-inch equipment rack. The IBM FlashSystem V9000 building block is 6U high and 19 inches wide. A standard 42U 19-inch data center rack can be used to be populated with the maximum IBM FlashSystem V9000 configuration to use up to 36U.

The IBM FlashSystem V9000 has the following physical dimensions:

- IBM FlashSystem V9000 control enclosure (AC3) each:
  - Width: 447.6 mm (17.62 in); 19-inch Rack Standard
  - Depth:801 mm (31.54 in)
  - Height: 87.5 mm (3.44 in.)
  - Weight: 23.8 kg (52.47 lb)
  - Airflow path: Cool air flows into the front of unit (intake) to rear of unit (exhaust)
  - Power consumption: 700 watts maximum, 450 watts typical operation
  - Heat dissipation: 512 BTU per hour
- IBM FlashSystem V9000 storage enclosure (AE3):
  - Width: 445 mm (17.6 in); 19-inch rack standard
  - Depth: 761 mm (29.96 in)
  - Height: 86.2 mm (3.39 in)
  - Weight (maximum configuration is 12 flash modules): 34 kg (75 lb)
  - Airflow path: Cool air flows into the front of unit (intake) to rear of unit (exhaust)
  - Power consumption: 1300 watts maximum, 625 watts typical operation
  - Heat dissipation: 1194 BTU (maximum configuration RAID 5)

For more information about IBM FlashSystem V9000 configuration options, see the IBM FlashSystem V9000 Model AE3 Product Guide.

1.2 IBM FlashSystem family

The success or failure of businesses often depends on how well organizations use their data assets for competitive advantage. Deeper insights from data require better information technology. As organizations modernize their IT infrastructure to boost innovation rather than limit it, they need a data storage system that can keep pace with highly virtualized environments, cloud computing, mobile and social systems of engagement, and in-depth, real-time analytics.

1.2.1 Why flash matters

Flash is a vibrant and fast growing technology. Clients are looking to solve data center problems, optimize applications, reduce costs, and grow their businesses.

Here are several reasons why flash is a *must* in every data center, and why an IBM FlashSystem changes storage economics:

- Reduces application and server licensing costs, especially those related to databases and virtualization solutions.
- Improves application efficiency, an application's ability to process, analyze, and manipulate more information, faster.
- Improves server efficiency. Helps you get more out of your existing processors, use less random access memory (RAM) per server, and consolidate operations by having server resources spend more time processing data as opposed to waiting for data.
- Improves storage operations. Helps eliminate costly application tuning, wasted developer cycles, storage array hot spots, array tuning, and complex troubleshooting. Decreases floor space usage and energy consumption by improving overall storage environment performance.
- Enhances performance for critical applications by providing the lowest latency in the market.

Almost all technological components in the data center are getting faster, including central processing units, network, storage area networks (SANs), and memory. All of them have improved their speeds by a minimum of ten times (10x); some of them by 100x, such as data networks. However, spinning disk has only increased its performance 1.2x.

The IBM FlashSystem V9000 provides benefits that include a better user experience, server and application consolidation, development cycle reduction, application scalability, data center footprint savings, and improved price performance economics.

Flash improves the performance of applications that are critical to the *user experience*, such as market analytics and research applications, trading and data analysis interfaces, simulation, modeling, rendering, and so on. Server and application consolidation is possible because of the increased process utilization resulting from the low latency of flash memory, which enables a server to load more users, more databases, and more applications.

Flash provides or gives back *time* for further processing within the existing resources of such servers. Clients soon realize that there is no need to acquire or expand server resources as often or as soon as was previously expected.

Development cycle reduction is possible because developers spend less time designing an application to work around the inefficiencies of HDDs and less time tuning for performance.

Data center footprint savings are realized due to the high density and high performance of the IBM flash solutions, these systems are replacing racks and cabinet bays of spinning HDDs. Reducing the data center footprint also translates into power and cooling savings, making flash one of the greenest technologies for the data center.

Improved price: Performance economics are because of the low cost for performance value from the IBM FlashSystem. The cost savings result from deploying fewer storage enclosures, fewer disk drives, fewer servers with fewer processors, and less RAM while using less power, space, cooling and fewer processor licenses. Flash is one of the best tools for the data center manager for improving data center economics.

1.2.2 Product differentiation

Flash is used widely in the data center in many ways: In a server (Peripheral Component Interconnect Express (PCIe) cards or internal SSDs), in storage arrays (hybrid or all-flash), appliances, or platform solutions (hardware, software, and network). Flash can be used as cache or as a data tier. Because of the vast and wide adoption of flash, several flash architectures and, therefore, criteria can be applied to compare flash options. See Figure 1-5.

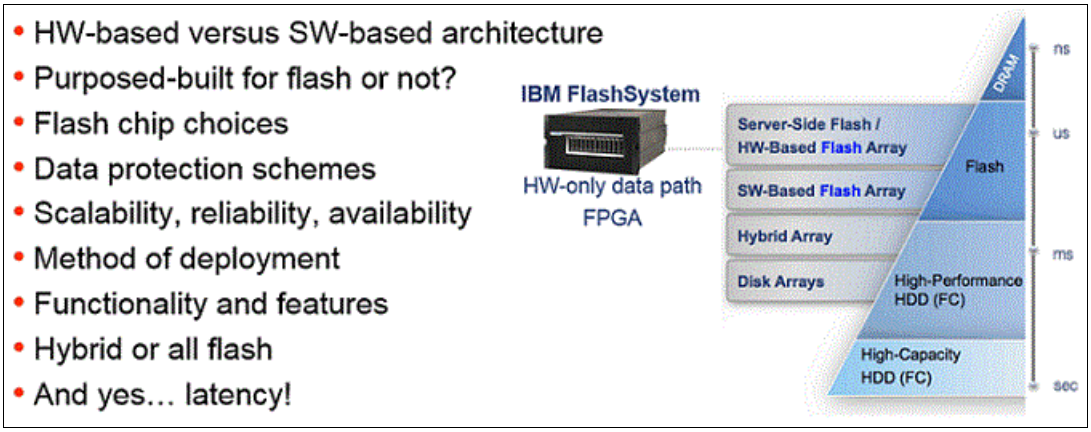


Figure 1-5 The various deployments of flash

Most storage vendors manufacture and market some type of flash memory system. The difference is how it is implemented, and the effect that such implementation has on the economics (cost reduction and revenue generation) for clients.

Flash technology is used to eliminate the storage *performance bottleneck.* The IBM FlashSystem family is a key shared-storage market leader and provides extremely low latency and consistent response times. It is designed and purpose-built specifically to harness what flash technology has to offer.

Some other vendors create flash appliances based on commodity server platforms and use software-heavy stack. Also, they use hardware technologies designed and created for disk, not flash. Others have designed their products using hybrid arrays combining existing storage designs, spinning HDDs, and SSD. The IBM storage portfolio includes SSD and flash on a variety of storage platforms; however, these alternative solutions do not have the same low latency (MicroLatency) as the hardware-accelerated IBM FlashSystem.

IBM FlashSystem family versus SSD-based storage arrays

Flash memory technologies appeared in the traditional storage systems some time ago. These SSD-based storage arrays help to successfully address the challenge of increasing I/Os per second (IOPS) needed by applications, and the demand for lower response times in particular tasks. An implementation example is the IBM Easy Tier technology.

However, these technologies typically rely on flash in the format of Fibre Channel (FC), serial-attached SCSI (SAS), or Serial Advanced Technology Attachment (SATA) disks, placed in the same storage system as traditional spinning disks, and using the same resources and data paths. This approach can limit the advantages of flash technology because of the limitations of traditional disk storage systems.

SAS attached storage (called expansion enclosures) still has advantages under the control of the IBM FlashSystem V9000, and can be considered as a good second tier level of storage in this environment.

IBM FlashSystem storage enclosures provide a hardware-only data path that realizes all of the potential of flash memory. These systems differ from traditional storage systems, both in the technology and usage.

An SSD device with an HDD disk form factor has flash memory that is put into a carrier or tray. This carrier is inserted into an array, such as an HDD. The speed of storage access is limited by the following technology because it adds latency and cannot keep pace with flash technology:

- Array controllers and software layers
- SAS controllers and shared bus
- Tiering and shared data path
- Form factor enclosure

IBM FlashSystem purpose-built MicroLatency modules are fast and efficient, designed using hardware-only data path technology that has a minimum number of software layers. Using this technology, IBM implements a mostly firmware component data path, and management software that is separated from the data path enabling the lowest latency modules on the market.

The only other family of products with hardware-only access to flash technology is the PCI Express (PCIe) flash product family, where products are installed into a dedicated server. With the appearance of the IBM FlashSystem, the benefits of PCIe flash products to a single server can now be shared by many servers.

1.2.3 The IBM Piece of Mind Initiative

IBM Storage has developed three new programs anchored to all-flash IBM storage offerings including:

- Data reduction program is designed to reduce planning risks and help lower storage costs by meeting baseline levels of data compression effectiveness in IBM Spectrum Virtualize based offerings.
- Controller upgrade program enables customers of designated all-flash IBM storage systems to reduce costs while maintaining leading-edge controller technology for essentially the cost of ongoing system maintenance.
- A new high-availability program helps enterprises avoid the costs and risks related to business downtime by ensuring the availability of business-critical data and storage systems.

Separately, the Data Reduction Guarantee, Controller Upgrade Program and High-Availability Guarantee each offer many benefits. But when combined as part of an IBM all-flash storage solution, the power of all three to help customers lower costs, reduce business risk and maintain the most current technologies can be even more significant. For example:

- Flash endurance coverage while hardware maintenance is current, ensuring that Flash wear never becomes a problem.
- Enhanced support is available from IBM via the Enterprise Class Support service that comes with the 9848 machine type.

Confidence. Trust. Peace of mind. IBM understands that real solutions include more than simply great engineering.

For further information on the IBM Piece of Mind Initiative please see the Technical White Paper of the same title.

1.3 FlashSystem V9000 architecture

The IBM FlashSystem V9000 architecture is explained in the following section together with key product design characteristics, performance, and serviceability. Hardware components are also described.

1.3.1 Overview of architecture

The IBM FlashSystem V9000 AC3 control enclosure combines software and hardware into a comprehensive, modular appliance that uses symmetric virtualization. Single virtualization engines, which are known as AC3 control enclosures, are combined to create clusters. In a scalable solution, each cluster can contain between two and eight control enclosures.

Symmetric virtualization is achieved by creating a pool of managed disks (MDisks) from the attached storage systems. Those storage systems are then mapped to a set of volumes for use by attached host systems. System administrators can view and access a common pool of storage on the storage area network (SAN). This functionality helps administrators to use storage resources more efficiently and provides a common base for advanced functions.

The design goals for the IBM FlashSystem V9000 are to provide the client with the fastest and most reliable all-flash storage array on the market, while making it simple to service and support with no downtime. The IBM FlashSystem V9000 uses hardware acceleration techniques incorporating Field Programmable Gate Array (FPGA) components to reduce the software stack which keeps I/O latency to a minimum and I/O performance to a maximum.

IBM Spectrum Virtualize software

IBM FlashSystem V9000 is built with IBM Spectrum Virtualize software, which is part of the IBM Spectrum Storage™ family.

Virtualization is a radical departure from traditional storage management. In traditional storage management, storage is attached directly to a host system that controls the storage management. SAN introduced the principle of networks of storage, but storage is still primarily created and maintained at the RAID system level. Multiple RAID controllers of different types require knowledge of, and software that is specific to, the specific hardware. Virtualization provides a central point of control for disk creation and maintenance.

IBM Spectrum Virtualize is a key member of the IBM Spectrum Storage portfolio. It is a highly flexible storage solution that enables rapid deployment of block storage services for new and traditional workloads, on-premises, off-premises, and in a combination of both, and it is designed to help enable cloud environments.

For more information about the IBM Spectrum Storage portfolio, see the IBM Storage website.

AE3 storage enclosure architecture

Figure 1-6 on page 12 illustrates the IBM FlashSystem V9000 AE3 storage enclosure design. At the core of the system are the two high-speed non-blocking crossbar buses. The crossbar buses provide two high-speed paths, which carry the data traffic, and they can be used by any host entry path into the system. There is also a slower speed bus for management traffic.

Connected to the crossbar buses are high-speed non-blocking RAID modules and IBM MicroLatency modules. There is also a passive main system board (midplane) to which both the RAID controllers (called canisters) and all the flash modules connect, and also connections to battery modules, fan modules, and power supply units.

The two RAID controllers contain crossbar controllers, management modules, interface controllers and interface adapters, and fan modules. The two RAID controllers form a logical cluster, and there is no single point of failure in the design (assuming that all host connections have at least one path to each canister).

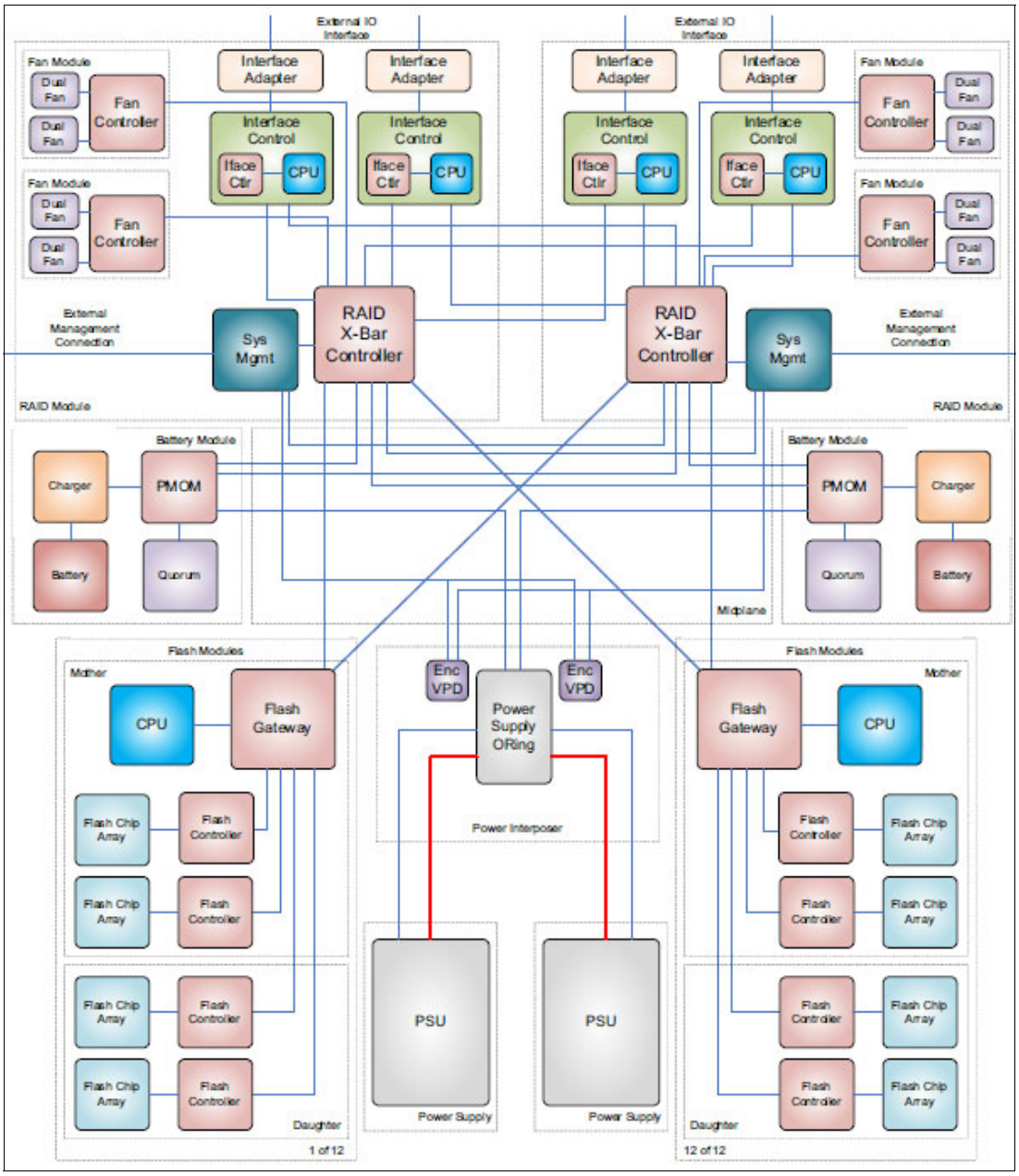


Figure 1-6 AE3 storage enclosure architecture

IBM FlashSystem V9000 software

The IBM FlashSystem V9000 software provides the following functions for the host systems that attach to IBM FlashSystem V9000:

- Creates a pool of storage. Two choices are available when the system consists of more than one AE3 storage enclosure:
  - Create a separate pool for each AE3 storage enclosure.
  - Create one storage pool that spans all AE3 storage enclosures.

Important: Before deciding whether to create a single storage pool or multiple storage pools, carefully evaluate which option best fits your solution needs, considering data availability and recovery management.

- Provides logical unit virtualization.
- Manages logical volumes.

IBM FlashSystem V9000 software also provides these advanced functions:

- Large scalable cache
- Copy services:
  - IBM FlashCopy® (point-in-time copy) function, including thin-provisioned FlashCopy to make multiple targets affordable
  - Metro Mirror (synchronous copy)
  - Global Mirror (asynchronous copy)
- Data migration
- Space management
- IBM Easy Tier function to automatically migrate the most frequently used data to higher-performance storage
- Thin-provisioned logical volumes
- Compressed volumes to consolidate storage
- HyperSwap, which enables each volume to be presented by two I/O groups
- Microsoft Offloaded Data Transfer (ODX)
- VMware and vSphere 6.0 support
- Enhanced FlashCopy bitmap space increased

For more information about the IBM FlashSystem V9000 advanced software features, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

MDisks

A managed disk (MDisk) is a logical unit of physical storage. MDisks are either arrays (RAID) from internal storage or volumes from external storage systems. MDisks are not visible to host systems.

An MDisk might consist of multiple physical disks that are presented as a single logical disk to the storage area network (SAN). An MDisk always provides usable blocks of physical storage to the system even if it does not have a one-to-one correspondence with a physical disk.

Each MDisk is divided into a number of extents, which are sequentially numbered starting at 0 (zero), from the start to the end of the MDisk. The extent size is a property of pools. When an MDisk is added to a pool, the size of the extents that the MDisk is divided into depends on the attribute of the pool to which it was added. The access mode determines how the clustered system uses the MDisk.

Attention: If you observe intermittent breaks in links or if you replaced cables or connections in the SAN fabric or LAN configuration, you might have one or more MDisks in degraded status. If an I/O operation is attempted when a link is broken and the I/O operation fails several times, the system partially excludes the MDisk and changes the status of the MDisk to *excluded*. You must include the MDisk to resolve the problem.

The MDisks are placed into storage pools where they are divided into several extents, which are 16 - 8192 MB, as defined by the IBM FlashSystem V9000 administrator. For more information about the total storage capacity that is manageable per system regarding the selection of extents, see the IBM Storwize V7000 support topic.

A volume is host-accessible storage that was provisioned from one *storage pool*. Alternatively, if it is a mirrored volume, it was provisioned from two storage pools. The maximum size of an MDisk is 1 PB. One IBM FlashSystem V9000 supports up to 4096 MDisks.

MDisks consideration for IBM FlashSystem V9000

Each MDisk from external storage has an online path count, which is the number of nodes that have access to that MDisk. The path count represents a summary of the I/O path status between the system nodes and the storage device. The maximum path count is the maximum number of paths that were detected by the system at any point in the past. If the current path count is not equal to the maximum path count, the MDisk might be degraded. That is, one or more nodes might not see the MDisk on the fabric.

Previously with IBM Spectrum Virtualize (2145 SAN Volume Controller model DH8) and previously with IBM FlashSystem V840, the leading practices stated that the back-end storage (on SAN Volume Controller) or internal storage (in FlashSystem V840) should be divided into 16 MDisks for the best performance. For the IBM FlashSystem V9000 storage enclosure AE2 this changed to be a single MDisk per storage enclosure.

For the AE3 storage enclosure, attached to IBM FlashSystem V9000, eight MDisks per storage enclosure should be created for best performance.

The reason for this change can be explained in the relationship of the I/O throughput on the machine, versus the number of cores and threading on the control enclosure architecture.

The control enclosures assign workloads to different cores, depending on the object that is associated with the workload. The three categories of objects are as follows:

- Interface channel (I/O) ports
- VDisks (volumes)
- MDisks (managed disks)

When an I/O comes in, this input is assigned to the core associated with an interface channel port. It moves to the VDisk thread and then to the MDisk thread and finally back to an interface channel thread, for de-staging back out of the system.

The VDisk has the most amount of work associated with it.

For additional information of Storage pools and Volumes see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

1.3.2 IBM FlashSystem V9000 building blocks

The IBM FlashSystem V9000 building block consists of two AC3 control enclosures, one AE3 storage enclosure, and software and hardware features. A building block can be either fixed or scalable. You can combine scalable building blocks to create larger clustered systems in such a way that operations are not disrupted.

Figure 1-7 shows a single IBM FlashSystem V9000 fixed versus multiple scalable building blocks.

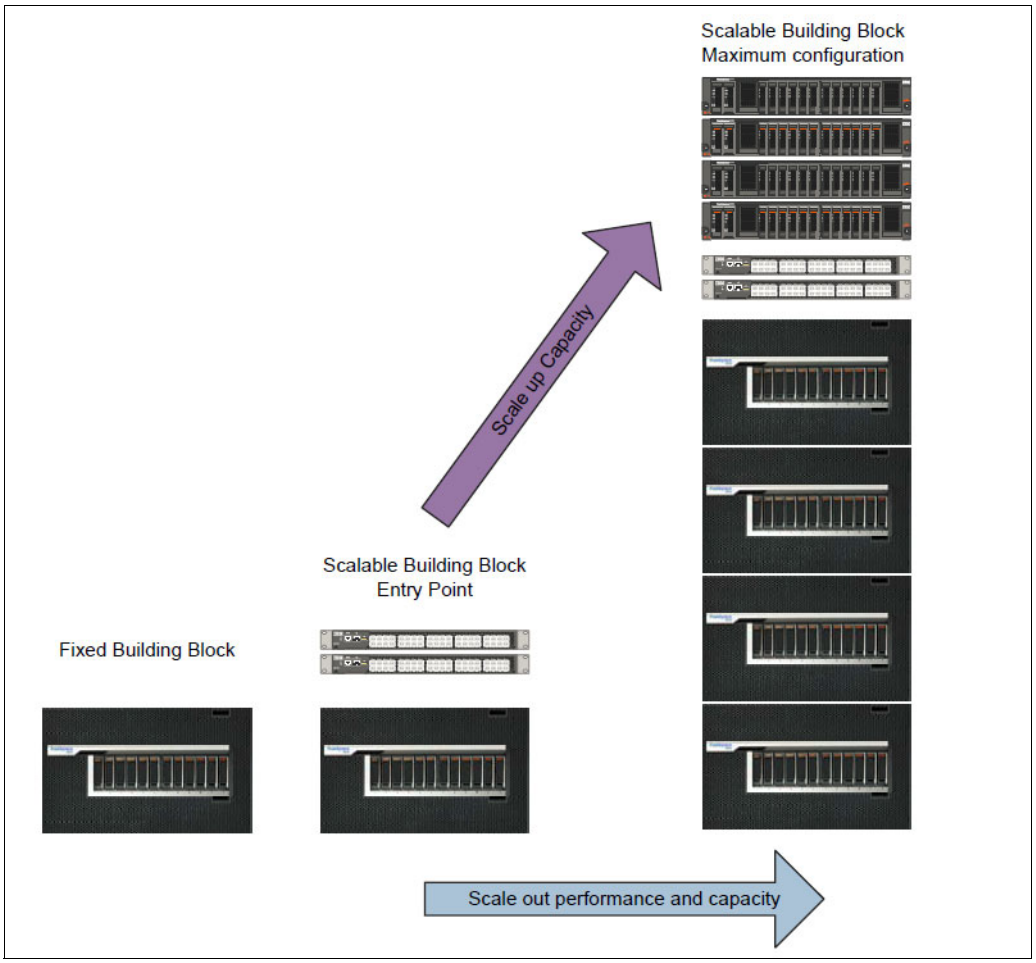


Figure 1-7 IBM FlashSystem V9000 scalability

Note: You can mix the AC2 control enclosure-based building blocks with the newer AC3 control enclosure building blocks, but each building block must have either two AC2s or two AC3s. The control enclosure types *cannot* be inter-mixed within a building block.

Scale up or scale out

A scalable building block can be scaled up by adding IBM FlashSystem V9000 AE3 storage enclosures for increased storage capacity. You can add a maximum of four extra storage enclosures.

A scalable building block can be scaled out by combining up to four building blocks to provide higher IOPS and bandwidth needs for increased performance.

Figure 1-7 illustrates the scalable capacity of IBM FlashSystem V9000. It also shows that extra IBM FlashSystem V9000 storage enclosures (SEs) can be added to a single building block, and also to two, three, or four building blocks.

1.3.3 IBM FlashSystem V9000 expansion enclosures

With the introduction of IBM FlashSystem V9000 expansion enclosures, even greater capacity offerings are now available.

The IBM FlashSystem V9000 large form factor (LFF) expansion enclosure model 12F offers new tiering options with high capacity nearline SAS hard disk drives (HDDs). Each LFF expansion enclosure supports up to twelve 8 TB or 10 TB drives.

The IBM FlashSystem V9000 small form factor (SFF) expansion enclosure model 24F offers new tiering options with low-cost SSDs. Each SFF expansion enclosure supports up to 24 2.5-inch low-cost SSD drives.

Up to 20 LFF or SFF expansion enclosures are supported per IBM FlashSystem V9000 controller pair, providing up to 480 drives with expansion enclosure model 24F (SFF) and up to 240 drives with expansion enclosure model 12F (LFF).

IBM FlashSystem V9000 HD expansion enclosure model 92F delivers increased storage density and capacity in a cost-efficient way. The IBM FlashSystem HD expansion enclosure model 92F offers the following features:

- 5U, 19-inch rack mount enclosure with slide rail and cable management assembly
- Support for up to ninety-two 3.5-inch LFF 12 Gbps SAS top-loading drives
- High-performance disk drives, high-capacity nearline disk drives, and flash drive support:
  - High-capacity, archival-class nearline disk drives in 8 TB and 10 TB 7,200 rpm
  - Flash drives in 1.92 TB, 3.84 TB, 7.68 TB, and 15.36 TB
- Dual redundant 200 - 240V AC power supplies (new PDU power cord required)

Up to eight model 92F high-density (HD) expansion enclosures are supported per IBM FlashSystem V9000 controller pair, providing up to 736 drives with expansion model 92F. With four controller pairs, a maximum of 32 HD expansion enclosures with up to 2,944 drives can be attached.

If a mix of SFF, LFF, and HD enclosures is required, see 2.2.4, “SAS expansion enclosures” on page 39.

Figure 1-8 on page 17 shows the maximum possible configuration with a single building block (controller pair) using a combination of native IBM FlashSystem V9000 storage enclosures and LFF or SFF expansion enclosures.

Note: If you require a mix of storage expansion enclosure types, see the supported configurations in 2.2.4, “SAS expansion enclosures” on page 39.

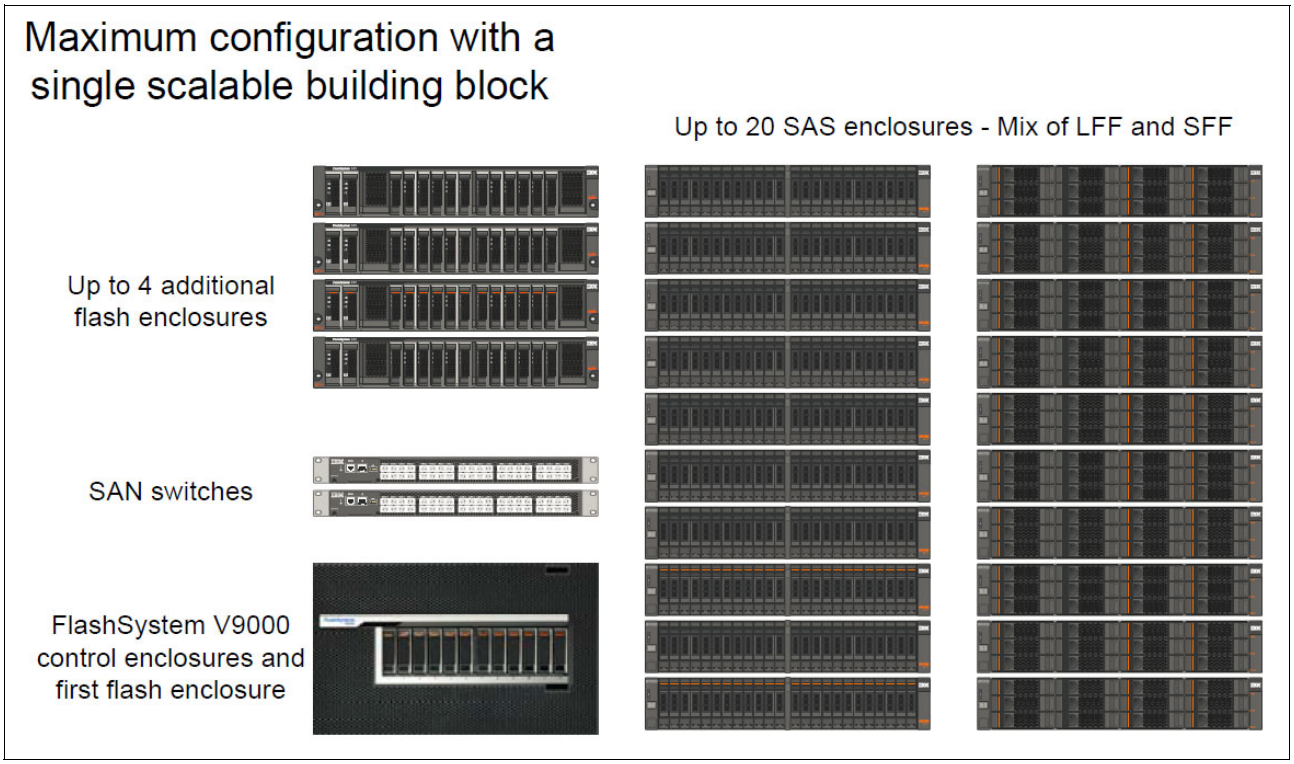


Figure 1-8 Single scalable building block maximum configuration

High-density, low-cost SSDs allow applications to scale and achieve high read performance while maintaining traditional reliability and endurance levels. The 1.92 TB and 3.84 TB SAS 2.5-inch low-cost SSDs options are available for IBM FlashSystem V9000 SFF expansion enclosure.

High-capacity nearline SAS drives enables high value tiered storage with hot data stored in flash and warm data on lower cost nearline SAS HDDs all managed by IBM Easy Tier. The 8 TB and 10 TB SAS 3.5-inch nearline drives are available for IBM FlashSystem V9000 LFF storage expansion enclosure model 12F.

The IBM FlashSystem V9000 with software version 8.1 also supports model 92F 5U-high, 92-drive bay, and supports the following drive types:

- High-capacity, nearline HDDs in 8 TB and 10 TB 7,200 rpm
- SSD Flash drives in 1.92 TB, 3.84 TB, 7.68 TB, and 15.36 TB

For IBM FlashSystem V9000 AE3 storage enclosure, only RAID5 with a standby hot spare is supported. However, the SAS attached expansion enclosures can be configured with various RAID options. The preference for SAS attached storage is distributed RAID DRAID 6, which offers improved redundancy and rebuild times.

Note: To support SAS expansion enclosures, an AH13 - SAS Enclosure Attach adapter card must be installed in slot 2 of each AC3 control enclosure in the building block.

1.3.4 IBM FlashSystem V9000 AE3 storage enclosure

The new IBM FlashSystem V9000 storage enclosure AE3 looks similar to the IBM FlashSystem V9000 storage enclosure AE2. The physical enclosure is the same, power supplies, batteries, fans and interface cards also remain the same. However, new RAID canisters and new 3D TLC Flash memory modules either 3.6 TB, 8.6 TB, or 18 TB provides more IOPS and larger capacity.

Furthermore, the AE3 enclosure has built in and always on inline hardware compression which offers 2:1 reduction in data, using inline hardware compression on the 3.6 TB and 8.5 TB modules. For less compressible workloads, IBM offers the 18 TB Flash memory modules for a maximum of up to 180 TB usable and up to 219 TB maximum effective capacity within a single AE3 enclosure.

Figure 1-9 shows the FlashSystem V9000 AE3 storage enclosure. One visible difference from the EA2 storage enclosure is the blue handlebars on the flash modules.

Figure 1-9 FlashSystem V9000 AE3 storage enclosure

IBM FlashSystem V9000 Storage Enclosure AE3 provides the following advanced features:

- IBM FlashCore technology
- Hardware accelerated I/O
- IBM MicroLatency module
- Inline hardware compression
- Advanced flash management
- Flash wear assurance

For details of the features mentioned above see the *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414 Redbooks publication.

1.3.5 Fixed and scalable configurations

IBM FlashSystem V9000 can be configured as a fixed building block or a scalable building block.

Fixed building block

A fixed building block contains one IBM FlashSystem V9000. The AE3 storage enclosure is cabled directly to each model AC3 control enclosure using 16 Gb links. Each control enclosure is connected to switches or directly attached to a host. The control enclosures internal connections are directly connected, without the use of switches or a SAN fabric, to form the cluster links. A fixed building block can be upgraded to a scalable building block, but the upgrade process is disruptive to operations.

Scalable building block

Scalable building blocks can contain from one and up to four control enclosure pairs (I/O groups) and from one and up to eight AE3 storage enclosures. In a scalable building block, the control enclosures are not cabled directly to each other. Instead Fibre Channel switches are used to create a private storage fabric. This infrastructure means that you can add building blocks or storage enclosures non disruptively.

The Fibre Channel switch fabrics do not have to be dedicated, and can be shared with hosts or server-side storage area networks (SANs). Care must be taken to ensure correct zoning for optimal interaction between hosts or server-side storage to FlashSystem V9000 control enclosures and backend storage.

For more guidelines of port utilization techniques in a scalable environment, see Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245.

The back-end storage switch fabric is isolated, through the zoning, from the host or server-side SAN for these reasons:

- So that any host or server does not have access to the AE3 storage enclosures directly
- So that the I/O from the controller to the storage does not interfere with the I/O from the host to the controllers

1.3.6 Scale-up and scale-out solutions

IBM FlashSystem V9000 offers the flexibility of the purchase of an all flash solution and hybrid enclosures that can be upgraded in the future, by the ability to scale-up for increased capacity, scale-out for increased performance, or both.

Clients can start with a fixed building block, or opt for a *scale-up scale-out* (SUSO) solution, that includes two 16 Gb FC switches, which enables you to add extra storage enclosures and building blocks with minimal effect to the existing systems.

Figure 1-10 shows the IBM FlashSystem V9000 scale-up and scale-out capabilities.

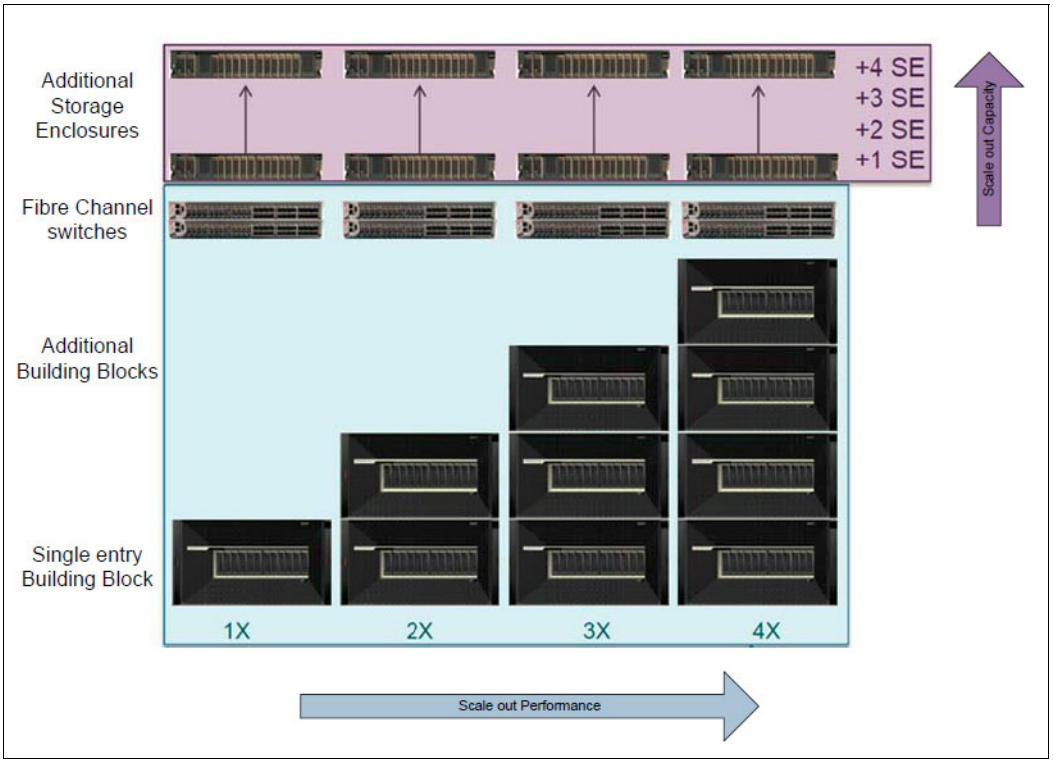


Figure 1-10 IBM FlashSystem V9000 scale-up and scale-out capabilities

Figure 1-11 shows a scale-out solution with four IBM FlashSystem V9000 building blocks, using the 16 Gb FC switches for interconnections.

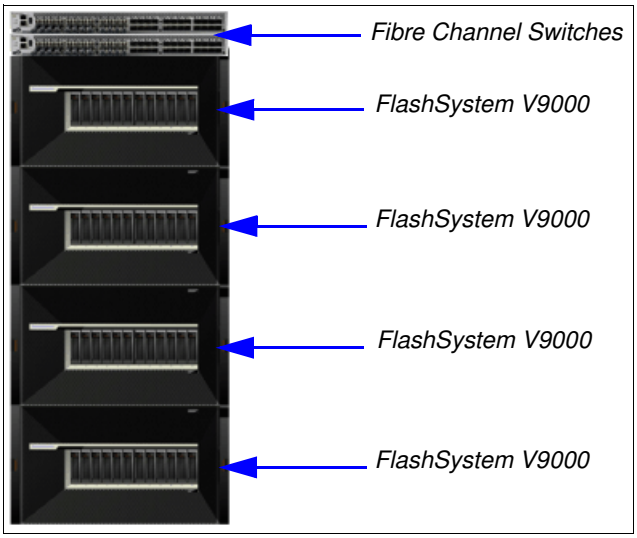


Figure 1-11 Scale out IBM FlashSystem V9000 solution

Figure 1-12 shows a scale-up solution with one IBM FlashSystem V9000 scalable building block and four additional IBM FlashSystem V9000 AE3 storage enclosures.

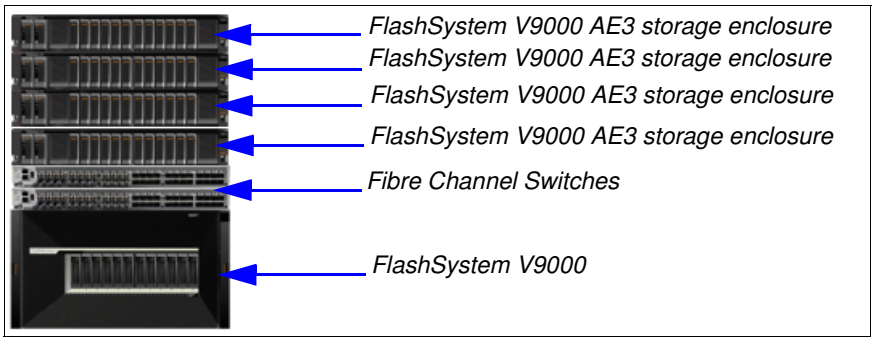


Figure 1-12 Scale up IBM FlashSystem V9000 solution

Figure 1-13 shows a scale-up and scale-out solution with four IBM FlashSystem V9000 building blocks and additional four IBM FlashSystem V9000 AE3 storage enclosures, indicating a maximum supported configuration.

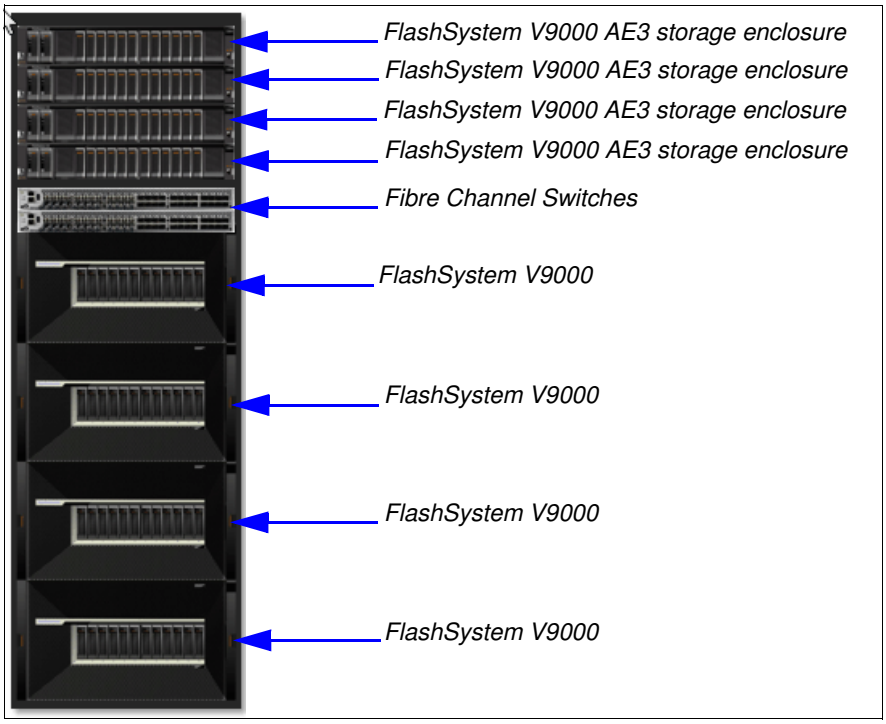


Figure 1-13 Scale-up and scale-out IBM FlashSystem V9000 solution

Note: The FC internal connection switches are ordered together with the first IBM FlashSystem V9000 scalable building block. IBM also supports the use of customer-supplied FC switches and cables, if they are supported by IBM. See the latest information about supported FC switches at the IBM System Storage® Interoperation Center (SSIC).

The IBM FlashSystem V9000 capacity can be expanded further by the inclusion of additional SAS attached expansion enclosures. See more details in 1.3.3, “IBM FlashSystem V9000 expansion enclosures” on page 16.

For more details about IBM FlashSystem V9000 scale-up or scale-out solutions, see Chapter 3, “Scalability” on page 61.

1.4 Advanced software features

The IBM FlashSystem V9000 builds on the IBM SAN Volume Controller nodes in conjunction with the FlashSystem V9000 AE3 storage enclosure. FlashSystem V9000 delivers the same advanced software features as IBM SAN Volume Controller. These include:

- IBM Real-time Compression
- FlashCopy
- Thin provisioning
- Thin-provisioned flash copies
- Built in data migration tools
- Advanced copy services
- Remote mirroring:
  - Metro Mirror
  - Global Mirror
  - Global Mirror with Change Volumes
- External virtualization
- Easy Tier
- IBM HyperSwap
- Transparent cloud tiering (V8.1)
- Remote Support capability
- Upload support package to directly to open PMR

For more information of software features see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

2

Planning

This chapter describes the steps that are required when you plan the installation of the IBM FlashSystem V9000 in your environment. This chapter considers the implications of your storage network from both the host attachment side and the virtualized storage expansion side. This chapter also describes all the environmental requirements that you must consider.

This chapter includes the following topics:

- General planning introduction
- Physical planning
- Logical planning
- License features
- IBM FlashSystem V9000 configuration backup procedure

This planning guide is based on the IBM FlashSystem V9000 AC3 control enclosure and the AE3 storage enclosure. Details about AE3 storage enclosure and about IBM Spectrum Virtualize 8.1 are described in these publications:

- *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933

For any configurations based on the AC2 or AC3 and the AE2 storage enclosure, you should see the previous editions of the Redbooks publications:

- *Introducing and Implementing IBM FlashSystem V9000*, SG24-8273
- *IBM Storwize V7000, Spectrum Virtualize, HyperSwap, and VMware Implementation*, SG24-8317

2.1 General planning introduction

The new IBM FlashSystem V9000 AC3 / AE3 combination is fundamentally different from the previous AC2 / AE2 configuration. The IBM FlashSystem V9000 AE3 storage enclosure is now a virtualized enclosure. On the older IBM FlashSystem V9000 AC2 or AC3 and the AE2 combinations, the AE2 storage enclosure was managed by the AC2 or AC3 control enclosures. Why this change was made is described in *IBM FlashSystem V9000 Model AE3 Product Guide*, REDP-5468.

Figure 2-1 shows the relation of the IBM FlashSystem V9000 AC3 control enclosures, and the new virtualized AE3 storage enclosure.

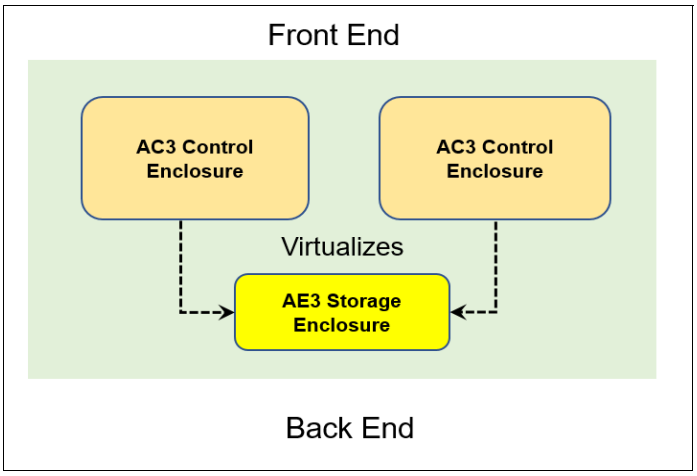


Figure 2-1 IBM FlashSystem V9000 AC3 control enclosures, and virtualized AE3 storage enclosure

Figure 2-2 shows the relation of the V9000 AC3 control nodes, the managed AE2 storage enclosure, and the new virtualized AE3 storage enclosure, in a mixed storage enclosure environment.

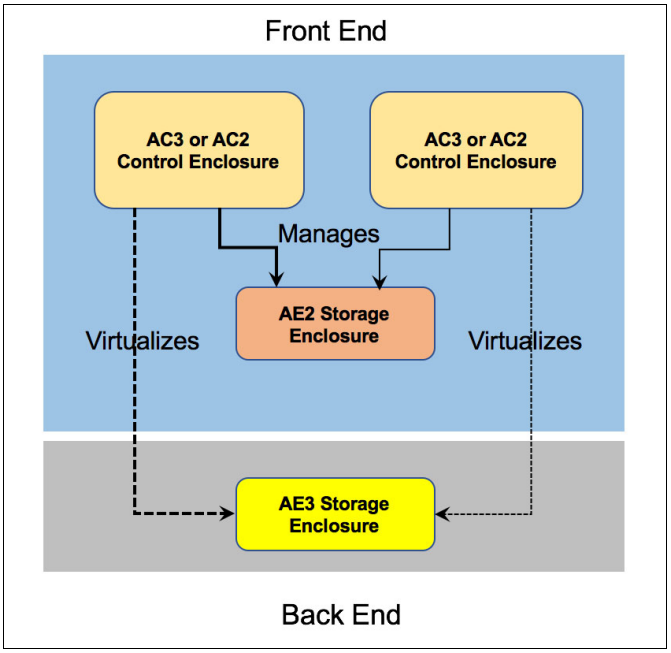


Figure 2-2 IBM FlashSystem V9000 AC3 / AE2 / AE3 enclosure combinations

To achieve the most benefit from the IBM FlashSystem V9000, preinstallation planning must include several important steps. These steps can ensure that the IBM FlashSystem V9000 provides the best possible performance, reliability, and ease of management to meet the needs of your solution. Proper planning and configuration also helps minimize future downtime by avoiding the need for changes to the IBM FlashSystem V9000 and the storage area network (SAN) environment to meet future growth needs.

Important steps include planning the IBM FlashSystem V9000 configuration and completing the planning tasks and worksheets before system installation.

An IBM FlashSystem V9000 solution is sold in what is referred to as a *building block*, as shown in Figure 2-3. A single building block consists of two AC3 control enclosures and one AE3 storage enclosure. Each building block is an *I/O Group* in the solution.

Note: In this chapter, the FlashSystem V9000 AE3 storage enclosure is called *external storage* because it is not managed using the FlashSystem V9000 controller GUI, but using its own FlashSystem V9000 AE3 storage enclosure GUI.

Figure 2-3 IBM FlashSystem V9000 base building block

IBM FlashSystem V9000 can be grown in two directions depending on the needs of the environment. This is known as the *scale-up, scale-out* capability:

- It can have all its capabilities increased by adding up to four total building blocks to the solution. This increases both the capacity and the performance alike.
- If just capacity is needed, it can be increased by adding up to four total AE3 storage enclosures beyond the single AE3 contained within each building block.

  A fully configured IBM FlashSystem V9000 consists of eight AC3 control enclosures and eight AE3 storage enclosures, sometimes referred to as an *eight by eight* configuration.

This chapter covers planning for the installation of a single IBM FlashSystem V9000 solution, consisting of a single building block (two AC3 control enclosures and one AE3 storage enclosure). When you plan for larger IBM FlashSystem V9000 configurations, consider the required SAN and networking connections for the appropriate number of building blocks and scale-up expansion AE3 storage controllers.

Note: If you have an existing V9000 comprised of two AC2 or AC3 control enclosures and one, or more AE2 storage enclosures, and are adding a new additional AE3 enclosure, then please refer to the section in this chapter concerning mixed storage enclosure types.

For details about scalability and multiple building blocks, see Chapter 3, “Scalability” on page 61.

Requirement: A pre-sale Technical Delivery Assessment (TDA) must be conducted to ensure that the configuration is correct and the solution being planned for is valid. A pre-install TDA must be conducted shortly after the order is placed and before the equipment arrives at the customer’s location to ensure that the site is ready for the delivery and that roles and responsibilities are documented regarding all the parties who will be engaged during the installation and implementation.

Before the system is installed and configured, you must complete all the planning worksheets. When the planning worksheets are completed, you submit them to the IBM service support representative (SSR).

Follow these steps when you plan for an IBM FlashSystem V9000 solution:

1. Collect and document the number of hosts (application servers) to attach to the IBM FlashSystem V9000, the traffic profile activity (read or write, sequential, or random), and the performance expectations for each user group (input/output (I/O) operations per second (IOPS) and throughput in megabytes per second (MBps)).
2. Collect and document the storage requirements and capacities:
   - Total external storage that will be attached to the IBM FlashSystem V9000
   - Required storage capacity for local mirror copy (Volume mirroring)
   - Required storage capacity for point-in-time copy (IBM FlashCopy)
   - Required storage capacity for remote copy (Metro Mirror and Global Mirror)
   - Required storage capacity for use of the IBM HyperSwap function
   - Required storage capacity for compressed volumes
   - Per host for storage capacity, the host logical unit number (LUN) quantity, and sizes
   - Required virtual storage capacity that is used as a fully managed volume and used as a thin-provisioned volume
3. Define the local and remote IBM FlashSystem V9000 SAN fabrics to be used for both the internal connections and the host and external storage. Also plan for the remote copy or the secondary disaster recovery site as needed.
4. Define the number of building blocks and additional expansion AE3 storage controllers required for the site solution. Each building block that makes up an I/O Group is the container for the volume. The number of necessary I/O Groups depends on the overall performance requirements.
5. If applicable, also consider any 12F, 24F, or 92F expansion enclosure requirements and the type of drives need in each enclosure. See section 2.2.4, “SAS expansion enclosures” on page 39 for more details.
6. Design the host side of the SAN according to the requirements for high availability and best performance. Consider the total number of ports and the bandwidth that is needed between the host and the IBM FlashSystem V9000, and the IBM FlashSystem V9000 and the external storage subsystems.
7. Design the internal side of the SAN according to the requirements as outlined in the cabling specifications for the building blocks being installed. This SAN network is used for IBM FlashSystem V9000 control nodes, and the expansion storage data transfers. Connecting this network across inter-switch links (ISL) is not supported.

Important: Check and carefully count the required ports for the wanted configuration. Equally important, consider future expansion when planning an initial installation to ensure ease of growth.

8. If your solution uses Internet Small Computer System Interface (iSCSI), design the iSCSI network according to the requirements for high availability (HA) and best performance. Consider the total number of ports and bandwidth that is needed between the host and the IBM FlashSystem V9000.
9. Determine the IBM FlashSystem V9000 cluster management and service Internet Protocol (IP) addresses needed. The V9000 system requires the following addresses:
   a. Cluster IP address for the V9000 System
   b. Service IP addresses for each AC3 control enclosures
   c. Cluster IP address for each virtualized AE3 storage enclosure
   d. Service IP addresses for each of the AE3 storage enclosures nodes
10. Determine the IP addresses for the IBM FlashSystem V9000 system and for the hosts that connect through the iSCSI network.
11. Define a naming convention for the IBM FlashSystem V9000 AC3 control enclosures, host, and any external storage subsystem planned. For example, *ITSO_V9000-1* shows that the IBM FlashSystem V9000 is mainly used by the International Technical Support Organization (ITSO) Redbooks team, and is the first IBM FlashSystem V9000 in the department.
12. Define the managed disks (MDisks) from external storage subsystems.

    Note: IBM FlashSystem V9000 AE3 storage enclosures must have eight volumes assign by the SSR. These then become the MDisks assigned, per enclosure, to the V9000. Assignment of these volumes is an IBM SSR install task.

13. Define storage pools. The use of storage pools depend on the workload, any external storage subsystem connected, more expansions or building blocks being added, and the focus for their use. There might also be a need for defining pools for use by data migration requirements or EasyTier. EasyTier is discussed in detail in 2.3.6, “EasyTier” on page 52.
14. Plan the logical configuration of the volumes within the I/O Groups and the storage pools to optimize the I/O load between the hosts and the IBM FlashSystem V9000.
15. Plan for the physical location of the equipment in the rack. IBM FlashSystem V9000 planning can be categorized into two types:
    – Physical planning
    – Logical planning

    The following sections describe these planning types in more detail.

    Note: IBM FlashSystem V9000 V8.1.0 provides GUI management of the HyperSwap function. HyperSwap enables each volume to be presented by two I/O groups. If you plan to use this function, you must consider the I/O Group assignments in the planning for the IBM FlashSystem V9000.

    For more details about the HyperSwap function, see Chapter 6, “IBM HyperSwap” on page 127, and Chapter 6, “IBM HyperSwap” on page 127.

2.2 Physical planning

Use the information in this section as guidance when you are planning the physical layout and connections to use for installing your IBM FlashSystem V9000 in a rack and connecting to your environment.

Industry standard racks are defined by the Electronic Industries Alliance (EIA) as 19-inch wide by 1.75-inch tall rack spaces or units, each of which is commonly referred to as *1U of the rack*. Each IBM FlashSystem V9000 building block requires 6U of contiguous space in a standard rack. Additionally, each add-on expansion enclosure requires another 2U of space.

When growing the IBM FlashSystem V9000 solution, by adding building blocks and expansions, the best approach is to plan for all of the members to be installed in the same rack for ease of cabling the internal dedicated SAN fabric connections. One 42U rack can house an entire maximum configuration of an IBM FlashSystem V9000 solution, and also its SAN switches and an Ethernet switch for management connections.

Figure 2-4 shows a fully configured solution of four building blocks plus four additional scale out storage enclosures in a 42U rack. This is known as an 8 x 8 configuration because it contains a total of eight control enclosures and eight storage enclosures.

Figure 2-4 Maximum full configuration of IBM FlashSystemV9000 fully scaled-out and scaled-up

The AC3 control enclosures

Each AC3 control enclosure can support up to eight PCIe expansion I/O cards, as identified in Table 2-1, to provide a range of connectivity and capacity expansion options. However at this time of writing, only seven slots are used (slot 1 use is not supported).

Table 2-1 Layout of expansion card options for AC3 control enclosures

PCIe slot	Adapter Type
1	Not supported for use
2	SAS
3	Fibre Channel or Ethernet
4	Fibre Channel or Ethernet
5	SAS or Compression accelerator
6	Fibre Channel or Ethernet
7	Fibre Channel or Ethernet
8	Compression Accelerator

Five I/O adapter options can be ordered:

- Feature code AH10: Four-port 8 gigabits per second (Gbps) FC Card:
  - Includes one four-port 8 Gbps FC Card with four Shortwave Transceivers
  - Maximum feature quantity is three
- Feature code AH11: Two-port 16 Gbps FC Card:
  - Includes one two-port 16 Gbps FC Card with two Shortwave Transceivers
  - Maximum feature quantity is four
- Feature code AH12: 4-port 10 Gbps Ethernet (iSCSI/FCoE):
  - Includes one four-port 10 GbE Card with four small form-factor pluggable plus (SFP+) transceivers
  - Maximum feature quantity is one
- Feature code AH13: 4-port 12 Gbps SAS
- Feature code AF44: 4-port 16 Gbps Fibre Channel

There is also an option for ordering the compression accelerator feature, which is included by default with IBM Real-time Compression software:

- Feature code AH1A: Compression Acceleration Card:
  - Includes one Compression Acceleration Card
  - Maximum feature quantity is two

Note the following information about the AC3 control enclosure PCIe adapters and slots:

- A maximum of four 4-port 16 Gbps Fibre Channel adapters can be installed in each control enclosure
- A maximum of one 4-port 10 Gbs Ethernet (iSCSI/FCoE) adapter can be installed in each control enclosure

- The 4-port SAS adapter can connect to V9000 standard or high-density expansion enclosures only. Only ports 1 and 3 can be used to provide the connections to each of the expansion enclosures.
- The compression accelerator adapter has no external ports. Compression adapters can be installed in PCIe slots 5 and 8 only. Two adapters can be installed offering improved I/O performance when using compressed volumes.

For more IBM FlashSystem product details, see *IBM FlashSystem V9000 Model AE3 Product Guide*, REDP-5468.

Figure 2-5 shows the rear view of an AC3 control enclosure with the eight available (only seven operational) PCIe adapter slots locations identified.

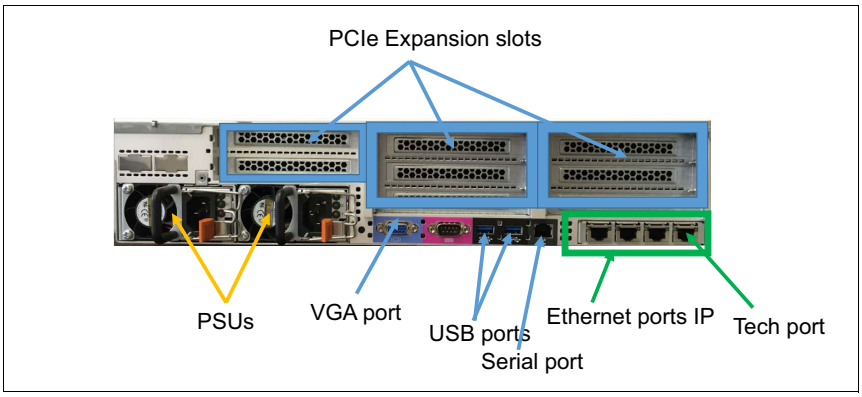


Figure 2-5 AC3 control enclosure rear view

The AE3 storage enclosure is a flash memory enclosure that can house up to 12 modules of 3.6 TB, 8.5 TB, and 18 TB capacities. The enclosure is equipped with either four FC adapters configured with four 8 Gbps ports, or configured with two 16 Gbps ports. There are two adapters per canister for a total of sixteen or eight ports. The AE3 storage enclosure also has two redundant 1300 W power supplies.

Figure 2-6 shows locations of these components.

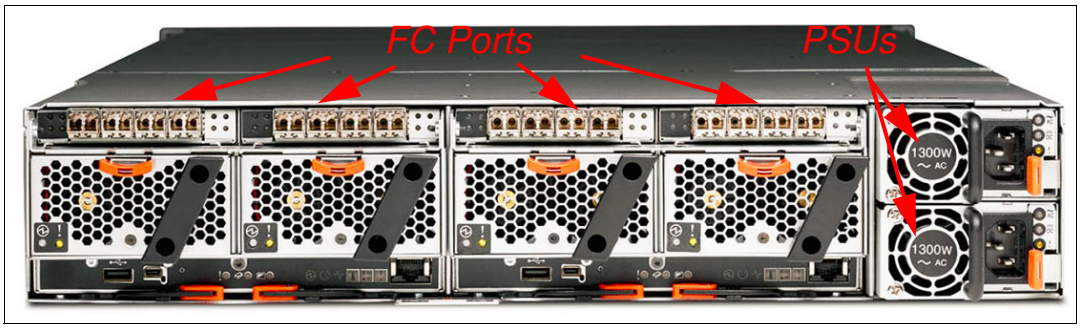


Figure 2-6 AE3 rear view

For more detailed information about the AE3 storage enclosure, please refer to the FlashSystem 900 Implementation Guide: *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

2.2.1 Racking considerations

IBM FlashSystem V9000 must be installed in a minimum of a one building block configuration. Each building block is designed with the two AC3 control enclosures and the AE3 enclosure in the middle. These enclosures must be installed contiguously and in the proper order for the system bezel to be attached to the front of the system. A total of 6U is needed for a single building block. Ensure that the space for the entire system is available.

Location of IBM FlashSystem V9000 in the rack

Because the IBM FlashSystem V9000 AC3 control enclosures and AE3 storage enclosure must be racked together behind their front bezel, all the members of the IBM FlashSystem V9000 must be interconnected together; the location where you rack the AC3 and the AE3 enclosures is important.

Use Table 2-2 to help plan the rack locations that you use for up to a 42U rack. Complete the table for the hardware locations of the IBM FlashSystem V9000 system and other devices.

Table 2-2 Hardware location planning of the IBM FlashSystem V9000 in the rack

Rack unit	Component
EIA 42	
EIA 41	
EIA 40	
EIA 39	
EIA 38	
EIA 37	
EIA 36	
EIA 35	
EIA 34	
EIA 33	
EIA 32	
EIA 31	
EIA 30	
EIA 29	
EIA 28	
EIA 27	
EIA 26	
EIA 25	
EIA 24	
EIA 23	
EIA 22	
EIA 21	

Rack unit	Component
EIA 20	
EIA 19	
EIA 18	
EIA 17	
EIA 16	
EIA 15	
EIA 14	
EIA 13	
EIA 12	
EIA 11	
EIA 10	
EIA 9	
EIA 8	
EIA 7	
EIA 6	
EIA 5	
EIA 4	
EIA 3	
EIA 2	
EIA 1	

Figure 2-7 shows a single base building block IBM FlashSystem V9000 rack installation with an additional AE3 storage enclosure plus space for future growth.

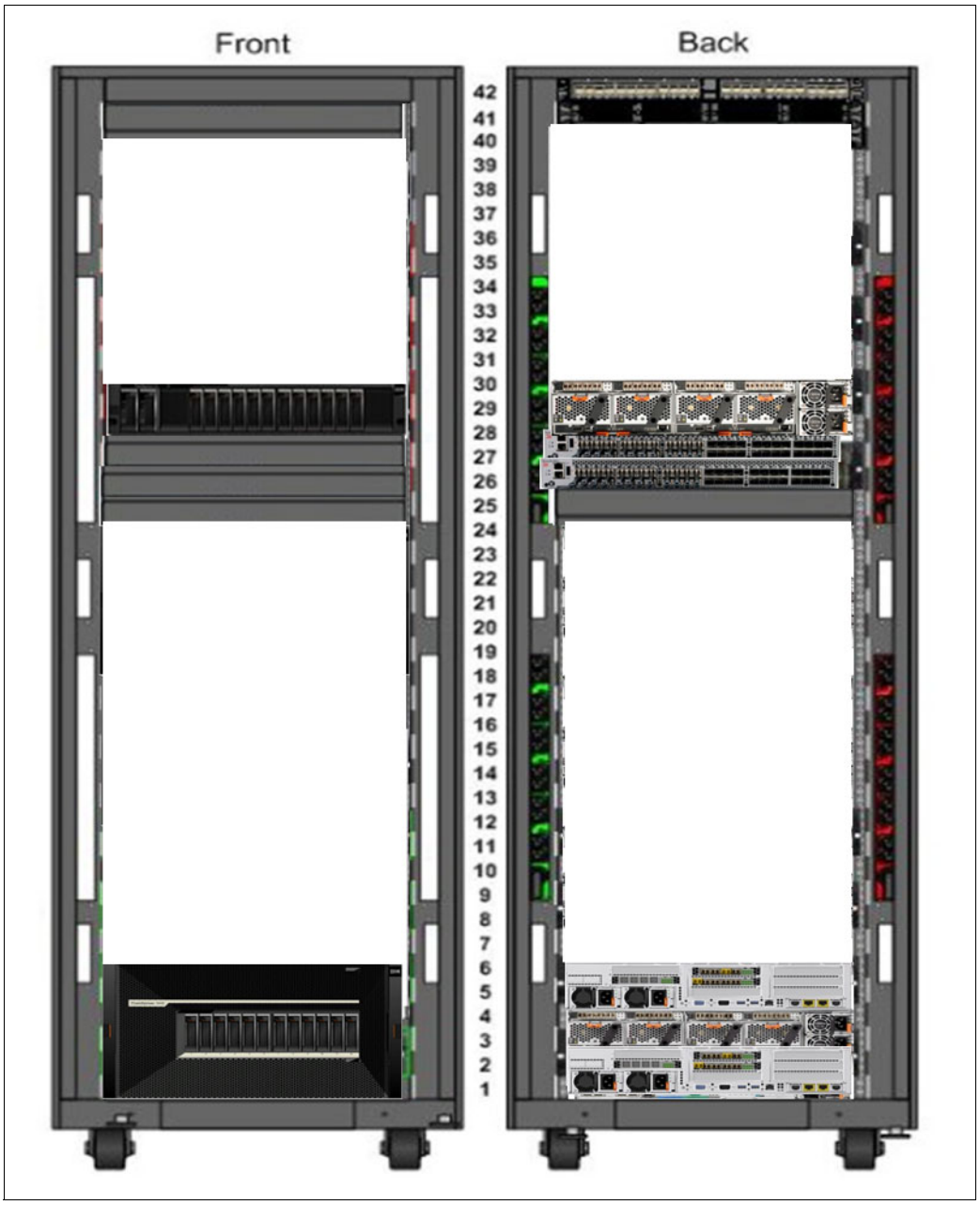


Figure 2-7 Sample racking of an IBM FlashSystemV9000 single building block with an add-on expansion for capacity

2.2.2 Power requirements

Each AC3 and AE3 enclosures requires two IEC-C13 power cable connections to connect to their 750 W and 1300 W power supplies. Country specifics power cables are available for ordering to ensure that proper cabling is provided for the specific region. A total of six power cords are required to connect each IBM FlashSystem V9000 building block to power.

Figure 2-8 shows an example of a base building block with the two AC2s, with two 750-W power supplies in each, and the AE3 with two 1300-W power supplies. There are six connections that require power for the IBM FlashSystem V9000 system.

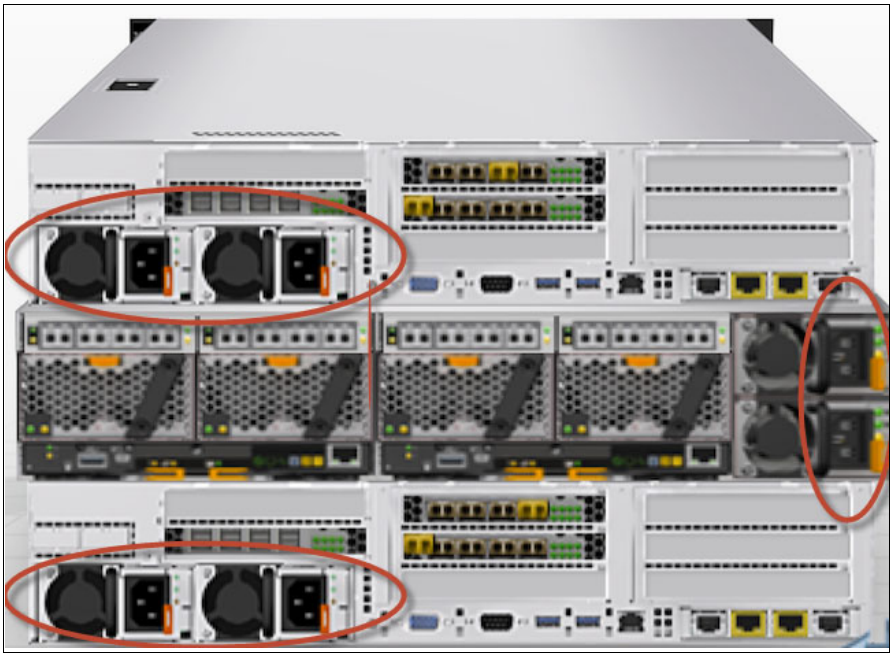


Figure 2-8 IBM FlashSystemV9000 fixed building block power cable connections

Upstream redundancy of the power to your cabinet (power circuit panels and on-floor Power Distribution Units (PDUs)), within cabinet power redundancy (dual power strips or in-cabinet PDUs), and upstream high availability structures (uninterruptible power supply (UPS), generators, and so on) influence your power cabling decisions.

If you are designing an initial layout that will have future growth plans to follow, you should plan to allow for the additional building blocks to be co-located in the same rack with your initial system for ease of planning for the additional interconnects required. A maximum configuration of the IBM FlashSystem V9000 with dedicated internal switches for SAN and local area network (LAN) can almost fill a 42U 19-inch rack.

Figure 2-7 on page 33 shows a single 42U rack cabinet implementation of a base building block IBM FlashSystem V9000 and also one optional IBM FlashSystem V9000 AE3 expansion add-on, all racked with SAN and LAN switches capable of handling additional future scaled out, scaled up additions with the 16 Gb switches for the SAN.

Tip: When cabling the power, connect one power cable from each AC3 control enclosures and AE3 storage enclosure to the left side internal PDU and the other power supply power cable to the right side internal PDU. This enables the cabinet to be split between two independent power sources for greater availability. When adding more IBM FlashSystem V9000 building blocks to the solution, continue the same power cabling scheme for each additional enclosure.

You must consider the maximum power rating of the rack: *do not exceed it*. For more power requirement information, see IBM FlashSystem V9000 at IBM Knowledge Center.

2.2.3 Network cable connections

As shown in Figure 2-9 the FC ports for this example (an 8 Gbps *fixed building block*) are identified for all the connections of the internal (back-end) fiber connections.

Figure 2-9 IBM FlashSystemV9000 fixed building block 8 Gbps FC cable connections

Create a cable connection table or similar documentation to track all of the connections that are required for the setup of these items:

- AC3 controller enclosures
- AE3 storage enclosures
- Ethernet
- FC ports: Host and internal
- iSCSI and Fibre Channel over Ethernet (FCoE) connections: Host Only

Figure 2-10 shows the back of the AC3 control enclosure with PCIe slots information.

Figure 2-10 AC3 control enclosure rear view with PCIe slots information

Slot numbers and adapter types are listed in Table 2-3.

Table 2-3 AC3 control enclosure PCIe slot numbers and adapter type

PCIe slot	Adapter types
1	Not supported for use

PCIe slot	Adapter types
2	SAS
3	Fibre Channel or Ethernet
4	Fibre Channel or Ethernet
5	SAS or Compression accelerator
6	Fibre Channel or Ethernet
7	Fibre Channel or Ethernet
8	Compression accelerator

You can download a sample cable connection table from the IBM FlashSystem V9000 page of IBM Knowledge Center by using the following steps:

1. Go to the IBM FlashSystem V9000 page in IBM Knowledge Center.
2. Click **Planning** on the right side panel.
3. In the list of results, select **Planning for the hardware installation (customer task)**.
4. Here you can select either option for download:
   - Planning worksheets for fixed building blocks
   - Planning worksheets for scalable building blocks

Use Table 2-4 to document the management and service IP address settings for the V9000 building block (both AC3 and AE3) in your environment.

Table 2-4 Management IP addresses for the IBM FlashSystem V9000 building block.

Cluster name:	
IBM FlashSystem V9000 AC3 Control Enclosures:	
Cluster IP address:	
IP:	
Subnet mask:	
Gateway:	
Node #1 Service IP address:	
IP:	
Subnet mask:	
Gateway:	
Node #2 Service IP address:	
IP:	
Subnet mask:	
Gateway:	
IBM FlashSystem V9000 AE3 Storage Enclosure	
Management IP address:	
IP:	
Subnet mask:	
Gateway:	
Canister 1 Service IP address:	
IP:	
Subnet mask:	
Gateway:	
Canister 2 Service IP address:	
IP:	
Subnet mask:	
Gateway:	

Note: If you have more than one AE3 storage enclosure to configure, then you will need a extra management IP and two service IP addresses, per additional storage enclosure.

Use Table 2-5 to document FC port connections for a single building block in your environment.

Table 2-5 Fibre Channel (FC) port connections

Location	Item	Fibre Channel port 1	Fibre Channel port 2	Fibre Channel port 3 (8 Gb FC only)	Fibre Channel port 4 (8 Gb FC only)
AC3 - Node1 Fibre Channel card 1	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 1 Fibre Channel card 2	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 1 Fibre Channel card 3	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 1 Fibre Channel card 4 (16 Gbps only)	AE3, Switch host:				
	Port:				
	Speed:				
AE3 - Canister 1 Fibre Channel card 1 (left)	AC3, Switch host:				
	Port:				
	Speed:				
AE3 - Canister 1 Fibre Channel card 2 (right)	AC3, Switch host:				
	Port:				
	Speed:				

Location	Item	Fibre Channel port 1	Fibre Channel port 2	Fibre Channel port 3 (8 Gb FC only)	Fibre Channel port 4 (8 Gb FC only)
AE3 - Canister 2 Fibre Channel card 1 (left)	AC3, Switch host:				
	Port:				
	Speed:				
AE3 - Canister 2 Fibre Channel card 2 (right)	AC3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 2 Fibre Channel card 1	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 2 Fibre Channel card 2	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 2 Fibre Channel card 3	AE3, Switch host:				
	Port:				
	Speed:				
AC3 - Node 2 Fibre Channel card 4 (16 Gbps only)	AE3, Switch host:				
	Port:				
	Speed:				

A complete suggested cabling guide is in the installation section of the IBM FlashSystem V9000 in IBM Knowledge Center.

2.2.4 SAS expansion enclosures

Three models of SAS expansion enclosures are offered:

- 9846/9848-12F
- 9846/9848-24F
- 9846/9848-92F

Expansion enclosure models 12F and 24F

To support a flash-optimized tiered storage configuration for mixed workloads, up to 20 9846/9848-12F or 9846/9848-24F SAS expansion enclosures can be connected to each building block in the system.

Maximum expansion enclosure capacity:

- A 9846/9848-12F SAS expansion enclosure contains up to 12 3.5 inch nearline SAS drives, and up to 9.6 PB raw capacity using 3.5 inch nearline SAS drives
- A 9846/9848-24F SAS expansion enclosure contains up to 24 2.5 inch high capacity SSDs, and up to 29.4 PB raw capacity
- Each building block supports up to 480 drives with expansion enclosure Model 24F (SFF) and up to 240 drives with expansion enclosure Model 12F (LFF)

Expansion enclosure model 92F

IBM FlashSystem V9000 High-Density (HD) Expansion Enclosure Model 92F delivers increased storage density and capacity for IBM FlashSystem V9000 with cost-efficiency while maintaining its highly flexible and intuitive characteristics:

- A 9846/9848-92F IBM FlashSystem HD expansion
- Expansion enclosure Model 92F offers the following features:
  - 5U, 19-inch rack mount enclosure with slide rail and cable management assembly
  - Support for up to ninety-two 3.5-inch large-form factor (LFF) 12 Gbps SAS top-loading drives
  - High-performance disk drives, high-capacity nearline disk drives, and flash drive support
  - High-capacity, archival-class nearline disk drives in 8 TB and 10 TB 7,200 rpm
  - High capacity SSDs in 1.92 TB, 3.84 TB, 7.68 TB, and 15.36 TB
  - Redundant 200 - 240VA power supplies (new PDU power cord required)
  - Up to 8 HD expansion enclosures are supported per IBM FlashSystem V9000 building block, providing up to 368 drives with expansion Model 92F for up to 7.36 PB of raw SAS HDD or 11.3 PB SSD capacity in each building block (up to a maximum of 32 PB total)
  - With four building blocks, a maximum of 32 HD expansion enclosures can be attached giving a maximum 29.4 PB of raw SAS capacity and 32PB of raw SSD capacity is supported

All drives within an enclosure must be the same model, but, a variety of drive models are supported for use in the IBM FlashSystem expansion enclosures, including SAS flash drives or SAS hard disk drives. These drives are hot swappable and have a modular design for easy replacement.

Note: To support SAS expansion enclosures, an AH13 - SAS Enclosure Attach adapter card must be installed in expansion slot 2 of each AC3 control enclosure in the building block.This is only for version 8.1 or higher.

Expansion enclosure worksheet

If the system includes optional SAS expansion enclosures, you must record the configuration values that will be used by the IBM SSR during the installation process.

Complete Table 2-6 based on your particular system and provide this worksheet to the IBM SSR prior to system installation.

Table 2-6 Configuration values: SAS enclosure x, building block x, and SAS enclosure n, building block n

Configuration setting	Value	Usage in CLI
MDisk group name	xxxx	`mkmdiskgrp -name mdisk_group_name -ext extent_size`
MDisk extent size in MB	xxxx	
RAID level (RAID5 or RAID6)	xxxx	`mkdistributedarray -level raid_level -driveclass driveclass_id -drivecount x -stripewidth x -rebuildareas x mdiskgrp_id | mdiskgrp_name`
driveclass_id: The class that is being used to create the array, which must be a numeric value.	xxxx	
drivecount: The number of drives to use for the array. The minimum drive count for RAID5 is 4; the minimum drive count for RAID6 is 6.	xxxx	
stripewidth: The width of a single unit of redundancy within a distributed set of drives. For RAID5, it is 3 - 16; for RAID6, it is 5 - 16.	xxxx	
rebuildareas: The reserved capacity that is distributed across all drives available to an array. Valid values for RAID5 and RAID6 are 1, 2, 3, and 4.	xxxx	

SAS expansion enclosures intermix enclosures is required, see 3.2.4, “SAS expansion enclosures intermix” on page 69.

2.3 Logical planning

Each IBM FlashSystem V9000 building block creates an I/O Group for the IBM FlashSystem V9000 system. IBM FlashSystem V9000 can contain up to four I/O Groups, with a total of eight AC3 control enclosures in four building blocks.

This section includes the following topics:

- Management IP addressing plan
- SAN zoning and SAN connections
- Call home option
- Remote Support Assistance
- IBM FlashSystem V9000 system configuration
- EasyTier
- Volume configuration
- SAN boot support

2.3.1 Management IP addressing plan

To manage the IBM FlashSystem V9000 system, you access the management GUI of the system by directing a web browser to the cluster’s management IP address.

In addition to this, the IBM FlashSystem V9000 Model AE3 storage enclosure attached to the V9000 AC3 control enclosures, also uses its own management GUI, accessed via its own management IP address.

The IBM FlashSystem V9000 AC3 uses a *technician port* feature. This is defined on Ethernet port 4 of any AC3 control enclosures and is allocated as the technician service port (and marked with the letter “T”). All initial configuration for the IBM FlashSystem V9000 AC3’s is performed through a technician port. The port broadcasts a Dynamic Host Configuration Protocol (DHCP) service so that any notebook or computer with DHCP enabled can be automatically assigned an IP address on connection to the port.

IBM FlashSystem V9000 AE3 storage enclosure uses a USB key initialization tool process. All initial configuration for the IBM FlashSystem V9000 AE3 is performed through the use of this USB key process.

Note: The hardware installation process for the V9000 is completed by the IBM SSR. If the V9000 is a scalable solution, then the SSR will work in conjunction with the IBM Lab Services team, to complete the installation.

See FlashSystem 900 Implementation Guide *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414, Installation Chapter, for further details on how to use this configuration process.

After the initial cluster configuration has been completed, the technician port automatically routes the connected user directly to the service GUI for the specific AC3 control enclosure if attached.

Note: The default IP address for the technician port on a V9000 AC3 node is 192.168.0.1. If the technician port is connected to a switch, it is disabled and an error is logged.

Each IBM FlashSystem V9000 AC3 control enclosure requires one Ethernet cable connection to an Ethernet switch or hub. The cable must be connected to port 1. For each cable, a 10/100/1000 Mb Ethernet connection is required. Both Internet Protocol Version 4 (IPv4) and Internet Protocol Version 6 (IPv6) are supported.

Note: For increased redundancy, an optional second Ethernet connection is supported for each AC3 control enclosure. This cable can be connected to Ethernet port 2.

To ensure system failover operations, Ethernet port 1 on all AC3 control enclosures, and on the AE3 storage enclosures, must be connected to the common set of subnets. If used for increased redundancy, Ethernet port 2 on all AC3 enclosures must also be connected to a common set of subnets. However, the subnet for Ethernet port 1 does not have to be the same as the subnet for Ethernet port 2.

Each IBM FlashSystem V9000 cluster must have a cluster management IP address and also a service IP address for each of the AC3 control enclosures in the cluster. Similarly each AE3 storage enclosure must have a management IP address and two service IP addresses assigned and cables connected, per enclosure.

Example 2-1 shows details. of the AC3 addresses

Example 2-1 AC3 Management IP address example (two building block configuration)

```
management IP add. 10.11.12.120
node 1 service IP add. 10.11.12.121
node 2 service IP add. 10.11.12.122
node 3 service IP add. 10.11.12.123
node 4 service IP add. 10.11.12.124
```

Requirement: Each AC3 control enclosure in an IBM FlashSystem V9000 clustered system must have at least one Ethernet connection.

Support for iSCSI on the IBM FlashSystem V9000 is available from only the optional 10 GbE adapters and would require extra IPv4 or extra IPv6 addresses for each of those 10 GbE ports used on each of the AC3 nodes. These IP addresses are independent of the IBM FlashSystem V9000 clustered system configuration IP addresses on the 10 GbE port 1 and port 2 for AC3 control enclosures.

When accessing the IBM FlashSystem V9000 through the GUI or Secure Shell (SSH), choose one of the available management or service IP addresses to connect to. In this case, no automatic failover capability is available. If one network is down, use an IP address on the alternative network.

Note: The Service Assistant tool described in the (reference to SVC section 13.10 and 900 Chapter 10 (verify) Redbooks is a web-based GUI that is used to service individual nodes and/or canisters, primarily when a node has a fault and is in a service state. This GUI is usually only used with guidance from IBM remote support. On the V9000 with AE3 storage enclosures, the service ports in the canisters of the AE3 enclosures should be assigned IP addresses and connected to the network. In prior versions of the V9000, with an AE2 storage enclosure, it was possible to manage the canisters of the AE2 by logging in to one of the nodes of control enclosures. With the AE3, a dedicated Ethernet connection to each canister of each AE3 storage enclosure is required to use the service assistant interface.

2.3.2 SAN zoning and SAN connections

IBM FlashSystem V9000 can connect to 8 Gbps or 16 Gbps Fibre Channel (FC) switches for SAN attachments. From a performance perspective, connecting the IBM FlashSystem V9000 to 16 GBps switches is better. For the internal SAN attachments, 16 Gbps switches are both better-performing and more cost-effective, because the 8 Gbps solution requires four switch fabrics, compared to the 16 Gbps needing only two.

Note: In the internal (back-end) fabric, ISLs are not allowed in the data path.

Both 8 Gbps and 16 Gbps SAN connections require correct zoning or VSAN configurations on the SAN switch or directors to bring security and performance together. Implement a dual-host bus adapter (HBA) approach at the host to access the IBM FlashSystem V9000. This example shows the 16 Gbps connections; details about the 8 Gbps connections are at IBM Knowledge Center.

Note: The IBM FlashSystem V9000 V8.1 supports 16 Gbps direct host connections without a switch.

Port configuration

With the IBM FlashSystem V9000 there are up to sixteen 16 Gbps Fibre Channel (FC) ports per building block used for the AE3 (eight ports) and internal AC3 communications (four per AC3, back-end) traffic. There are also two adapters, which if FC type, can be divided between the Advanced Mirroring features, host, and external virtualized storage (front-end) traffic.

If you want to achieve the lowest latency storage environment, the “scaled building block” solution provides the most ports per node to intercluster and inter-I/O group traffic with all the back-end ports zoned together. When creating a scaled out solution, the same port usage model is repeated with all building blocks. When creating a scaled up solution, you will add the new AE3 ports to the zone configurations equally so that the traffic load and redundancy are kept equally balanced.

For cabling and port utilization tables and suggestions, see Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245:

- A.3, “Guidelines: The performance method” on page 247
- A.4, “Guidelines: The infrastructure savings method” on page 249

Note: Connecting the AC3 control enclosures FC ports and the AE3 FC ports in an IBM FlashSystem V9000 scalable environment is an IBM lab-based services task. For details, see the IBM FlashSystem V9000 web page at IBM Knowledge Center.

Customer provided switches and zoning

This topic applies to anyone using customer-provided switches or directors.

External virtualized storage systems are attached along with the host on the front-end FC ports for access by the AC3 control enclosures of the IBM FlashSystem V9000. Carefully create zoning plans for each additional storage system so that these systems will be properly configured for use and best performance between storage systems and the IBM FlashSystem V9000. Configure all external storage systems with all IBM FlashSystem V9000 AC3 control enclosures; arrange them for a balanced spread across the system.

All IBM FlashSystem V9000 AC3 control enclosures in the IBM FlashSystem V9000 system must be connected to the same SANs, so that they all can present volumes to the hosts. These volumes are created from storage pools that are composed of the virtualized AE3 storage enclosure MDisks, and if licensed, the external storage systems MDisks that are managed by the IBM FlashSystem V9000.

For suggested fabric zoning see Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245 (and specifically A.5, “Guidelines: Zoning and pathing” on page 250).

2.3.3 Call home option

IBM FlashSystem V9000 supports setting up a Simple Mail Transfer Protocol (SMTP) mail server for alerting the IBM Support Center of system incidents that might require a service event. This is the *call home* option. You can enable this option during the setup.

Tip: Setting up call home involves providing a contact that is available 24 x 7 if a serious call home issue occurs. IBM support strives to report any issues to clients in a timely manner; having a valid contact is important to achieving service level agreements (SLAs). For more detail about properly configuring call home, see section Notifications menu in the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Table 2-7 lists the necessary items for the AC3.

Table 2-7 Call home options for AC3 control enclosures

Configuration item	Value
Primary Domain Name System (DNS) server	
SMTP gateway address	
SMTP gateway name	
SMTP “From” address	Example: `V9000_name@customer_domain.com`
Optional: Customer email alert group name	Example: `group_name@customer_domain.com`
Network Time Protocol (NTP) manager	
Time zone	

In addition to the IBM FlashSystem V9000 AC3 call home set-up, you also have to perform a similar set-up on the IBM FlashSystem V9000 AE3 storage enclosure, as its performs its own call home functions.

Table 2-8 lists the necessary items for the AC3.

Table 2-8 Call home options for AE3 storage enclosure

Configuration item	Value
Primary Domain Name System (DNS) server	
SMTP gateway address	
SMTP gateway name	
SMTP “From” address	Example: `FS900_name@customer_domain.com`
Optional: Customer email alert group name	Example: `group_name@customer_domain.com`
Network Time Protocol (NTP) manager	
Time zone	

See the Installation chapter in the FlashSystem 900 AE3 implementation Guide, *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414, for information in setting up the IBM FlashSystem V9000 AE3 storage enclosure call home function.

2.3.4 Remote Support Assistance

The IBM FlashSystem V9000 AC3 control enclosure and the IBM FlashSystem V9000 AE3 storage enclosure both support the new remote support assistance (RSA) feature. This function is enabled, and available, at code level 8.1.x and higher for the AC3 and code level 1.5.x and higher or the AE3.

By using the Remote Support Assistance (RSA), the customer is able to initiate a secure connection from FlashSystem V9000 to IBM when problems arise. An IBM remote support specialist can then connect to the system to collect system logs, analyze a problem, if possible run repair actions remotely, or assist the client or an IBM SSR who is on site.

Important: IBM encourages all customers to use the high-speed remote support solution that is enabled by RSA. Problem analysis and repair actions without a remote connection can get more complicated and time-consuming.

The RSA uses a high-speed internet connection, but it gives the customer the ability to initiate an outbound Secure Shell (SSH) call to a secure IBM server. Fire wall rules might need to be configured at the customer's fire wall to allow the FlashSystem V9000 Cluster and Service IPs to establish a connection to the IBM Remote Support Center via SSH. This applies to both, to the V9000 AC3 control enclosure and also the IBM FlashSystem V9000 AE3 storage enclosure.

Note: The type of access that is required for a remote support connection is outbound port TCP/22 (SSH) from the FlashSystem V9000 Cluster and Service IPs (both AC3 and AE3).

The RSA consists of FlashSystem V9000 internal functions with a set of globally deployed supporting servers. Together, they provide secure remote access to the FlashSystem V9000 when necessary and when authorized by the customer's personnel.

Figure 2-11 shows the overview of the V9000 RSA set-up, which has three major components.

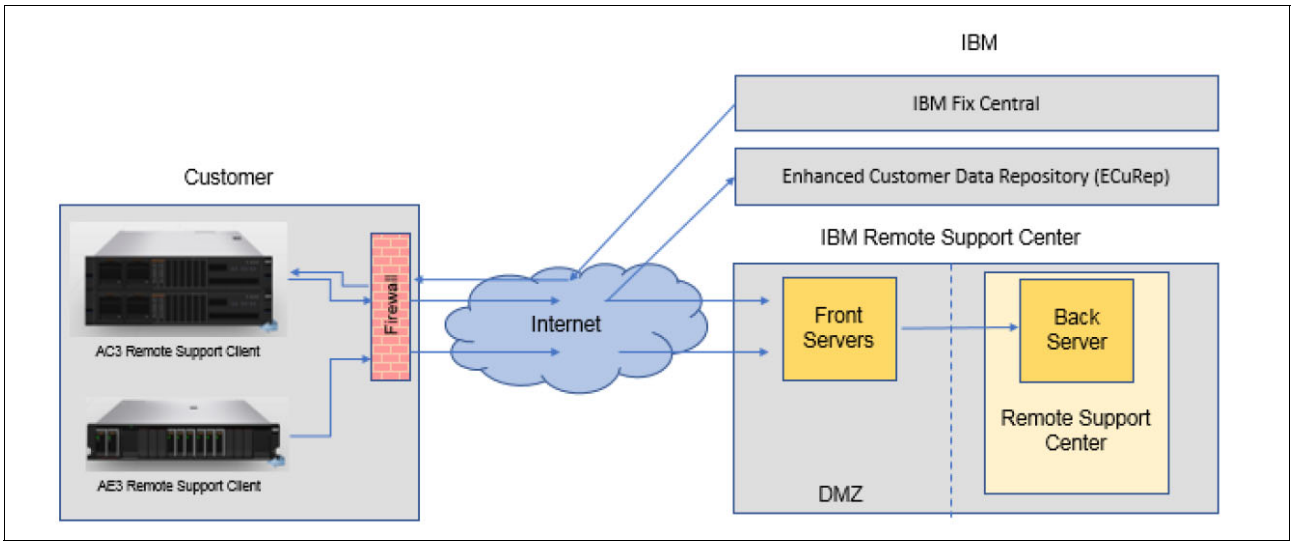


Figure 2-11 Overview of the V9000 RSA set-up without proxy

Remote Support Client (machine internal)

The Remote Support Client is a software component inside FlashSystem V9000 that handles remote support connectivity. It resides on all nodes of the V9000 AC3 control enclosure and also the on the nodes of the IBM FlashSystem V9000 AE3 storage. The software component relies only on a single outgoing Transmission Control Protocol (TCP) connection, and it cannot receive inbound connections of any kind. The Remote Support Client is controlled either by using CLI or the GUI.

Remote Support Center Front Server (Internet)

Front Servers are on an IBM Demilitarized Zone (DMZ) of the internet and receive connections from the Remote Support Client and the IBM Remote Support Center Back Server. Front Servers are security-hardened machines that provide a minimal set of services, such as maintaining connectivity to connected Clients and to the Back Server.

They are strictly inbound, and never initiate anything on their own accord. No sensitive information is ever stored on the Front Server, and all data that passes through the Front Server from the client to the Back Server is encrypted so that the Front Server cannot access this data.

Note: When activating Remote Support Assistant, the following four Front Servers will be used via port TCP/22 (SSH):

- 204.146.30.139
- 129.33.206.139
- 204.146.30.157 (by default for V9000 AE3 storage enclosures only)
- 129.33.207.37 (by default for V9000 AE3 storage enclosures only)

Remote Support Center Back Server (IBM Intranet)

The Back Server manages most of the logic of the Remote Support Assistance system. It is located within the IBM Intranet. The Back Server maintains connection to all FrontServers and is access-controlled. Only IBM employees who are authorized to perform remote support of FlashSystem V9000 are allowed to use it. The Back Server is in charge of authenticating a support person.

It provides the support person with an user interface (UI) through which to choose a system to support based on the support person's permissions. It also provides the list of systems that are currently connected to the Front Servers, and it manages the remote support session as it progresses (logging it, allowing additional support persons to join the session, and so on).

In addition the V9000 remote support solution can take advantage of the following two IBM internet support environments.

IBM Enhanced Customer Data Repository (ECuRep)

Further, if a remote connection exists, the IBM remote support specialists can off load the required support logs by themselves. For additional information about ECuRep, see the support web page Overview.

IBM Fix Central

FixCentral provides fixes and updates for IBM system's software, hardware, and operating system. The V9000 AC3 control enclosure provides the possibility to allow an IBM remote support specialist to perform software updates remotely. During this process the V9000 control enclosure automatically downloads the required software packages from the IBM.

Note: To download software update packages, the following six IP addresses are used via outbound port TCP/22 (SSH) from the V9000 AC3 control enclosure to FixCentral:

- 170.225.15.105
- 170.225.15.104
- 170.225.15.107
- 129.35.224.105
- 129.35.224.104
- 129.35.224.107

Firewall rules might need to be configured. Further it is required to configure a DNS server to allow the download function to work.

Remote Support Proxy.

Optionally, an application called Remote Support Proxy can be used when one or more FlashSystem V9000 systems do not have direct access to the Internet (for example, because of firewall restrictions). The Remote Support Client within the FlashSystem will then connect through this optional proxy server to the Remote Support Center Front Servers. The Remote Support Proxy runs as a service on a Linux system that has Internet connectivity to the Remote Support Center and local network connectivity to the FlashSystem V9000.

Figure 2-12 illustrates the connection through the Remote Support Proxy.

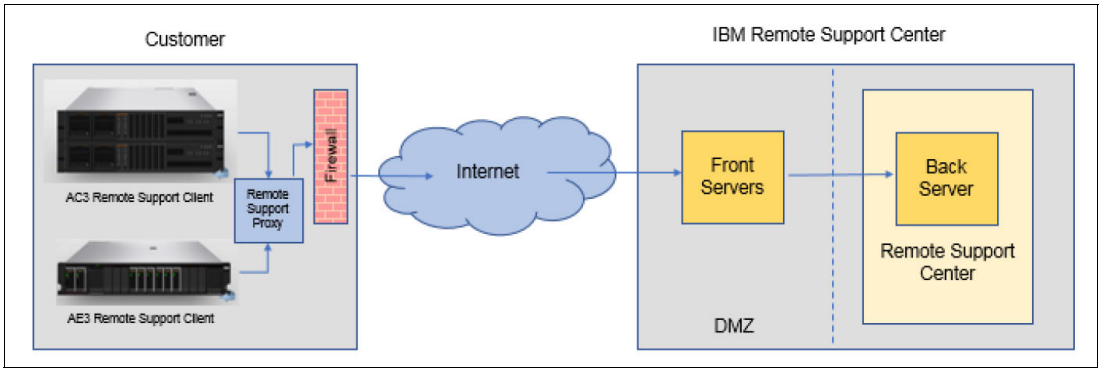


Figure 2-12 Remote Support Proxy set-up

The communication between the Remote Support Proxy and the Remote Support Center is encrypted with an additional layer of Secure Sockets Layer (SSL).

Note: The host that is running the Remote Support Proxy must have TCP/443 (SSL) outbound access to Remote Support Front Servers.

Remote Support Proxy software.

The Remote Support Proxy is a small program which is supported on some Linux versions. The software is also used for other IBM Storage Systems like IBM XIV or FlashSystem A9000. The installation files and documentations are available at the storage portal website.

“Setting Up Remote Support” in the Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414 shows how to setup the Remote Support Proxy.

Note: At the time of writing the Redbook, The Remote Support Proxy does not support a connection to IBM Enhanced Customer Data Repository (ECuRep) for automatically uploading logs. Also the software download from FixCentral is not supported through the optional Remote Support proxy.

For the AC3 set-up and RSA configuration, please refer to the “Remote Support Assistance” in *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

For the AE3 set-up and RSA configuration, please refer to the FlashSystem 900 Implementation Guide *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

2.3.5 IBM FlashSystem V9000 system configuration

To ensure proper performance and high availability in the IBM FlashSystem V9000 installations, consider the following guidelines when you design a SAN to support the IBM FlashSystem V9000:

- All nodes in a clustered system must be on the same LAN segment, because any node in the clustered system must be able to assume the clustered system management IP address. Make sure that the network configuration allows any of the nodes to use these IP addresses. If you plan to use the second Ethernet port on each node, it is possible to have two LAN segments. However, port 1 of every node must be in one LAN segment, and port 2 of every node must be in the other LAN segment.
- To maintain application uptime in the unlikely event of an individual AC3 control enclosure failing, IBM FlashSystem V9000 control enclosures are always deployed in pairs (I/O Groups). If a control enclosure fails or is removed from the configuration, the remaining control enclosures operates in a degraded mode, but the configuration is still valid for the I/O Group.

  Important: The IBM FlashSystem V9000 V8.1 release includes the HyperSwap function, which allows each volume to be presented by two I/O groups. If you plan to use this function, you must consider the I/O Group assignments in the planning for the IBM FlashSystem V9000.

  For more details about the HyperSwap function, 1.3, “FlashSystem V9000 architecture” on page 10 and see Chapter 6, “IBM HyperSwap” on page 127.
- The FC SAN connections between the AC3 control enclosures and the switches are optical fiber. These connections can run at either 8 or 16 Gbps depending on your switch hardware.
- The AC3 control enclosures ports can be configured to connect either by 8 Gbps direct connect, known as the *fixed building block* configuration, or by 16 Gbps to an FC switch fabric, known as a *scalable building block*.
- Direct connections between the AC3 control enclosures and hosts are supported with some exceptions.
- Direct connection of AC3 control enclosures and external storage subsystems are not supported.

  Exception: The IBM FlashSystem V9000 AE3 storage enclosure can be direct attached to the AC3 control enclosure in a fixed block configuration.

- Two IBM FlashSystem V9000 clustered systems cannot have access to the same external virtualized storage LUNs within a disk subsystem.

  Attention: Configuring zoning so that two IBM FlashSystem V9000 clustered systems have access to the same external LUNs (MDisks) can result in data corruption.

- The IBM FlashSystem V9000 enclosures within a building block must be co-located (within the same set of racks) and in a contiguous 6U section.
- The IBM FlashSystem V9000 uses three MDisks as quorum disks for the clustered system. A preferred practice for redundancy is to have each quorum disk in a separate storage subsystem, where possible. The current locations of the quorum disks can be displayed using the **`lsquorum`** command and relocated using the **`chquorum`** command.

The storage pool and MDisk

The storage pool is at the center of the relationship between the MDisks and the volumes (VDisk). It acts as a container from which MDisks contribute chunks of physical capacity known as *extents*, and from which VDisks are created.

The internal MDisks in the IBM FlashSystem V9000 are created on a basis of *eight* MDisk's for managed expansion AE3 enclosures attached to the IBM FlashSystem V9000 clustered system. These AE3 storage enclosures can be part of a building block, or an add-on expansion in a *scale-up* configuration.

Additionally, MDisks are also created for each external storage attached LUN assigned to the IBM FlashSystem V9000 as a managed or as unmanaged MDisk for migrating data. A managed MDisk is an MDisk that is assigned as a member of a storage pool:

- A *storage pool* is a collection of MDisks. An MDisk can only be contained within a single storage pool
- IBM FlashSystem V9000 can support up to 1,024 storage pools
- The number of volumes that can be allocated per system limit is 10,000
- Volumes are associated with a single storage pool, except in cases where a volume is being migrated or mirrored between storage pools

> **Information:** For more information about the MDisk assignments and explanation of why eight MDisk's per AE3 storage enclosure is used, see "MDisks" on page 13.
>
> For the most up-to-date SAN Volume Controller configuration limits, search for the "Configuration Limits and Restrictions" topic for the latest SAN Volume Controller version.

Extent size

Each MDisk is divided into chunks of equal size called *extents*. Extents are a unit of mapping that provides the logical connection between MDisks and volume copies.

The extent size is a property of the storage pool and is set when the storage pool is created. All MDisks in the storage pool have the same extent size, and all volumes that are allocated from the storage pool have the same extent size. The extent size of a storage pool cannot be changed. If you want another extent size, the storage pool must be deleted and a new storage pool configured.

The IBM FlashSystem V9000 supports extent sizes of 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, and 8192 MB. By default, the MDisk created for the internal expansions of flash memory in the IBM FlashSystem V9000 building block are created with an extent size of 1024 MB. To use a value that differs from the default requires the use of CLI commands to delete and re-create with different value settings. For information about the use of the CLI commands, search for CLI commands in IBM Knowledge Center.

Table 2-9 lists all of the extent sizes that are available in an IBM FlashSystem V9000.

Table 2-9 Extent size and maximum clustered system capacities

Extent size	Maximum clustered system capacity
16 MB	64 TB
32 MB	128 TB
64 MB	256 TB
128 MB	512 TB
256 MB	1 petabyte (PB)
512 MB	2 PB
1,024 MB	4 PB
2,048 MB	8 PB
4,096 MB	16 PB
8,192 MB	32 PB

Consider the following information about storage pools:

- Maximum clustered system capacity is related to the extent size:
  - 16 MB extent = 64 TB and doubles for each increment in extent size; for example, 32 MB = 128 TB. For the internal expansion enclosure MDisk, the default extent size is 1024 MB.
  - You cannot migrate volumes between storage pools with separate extent sizes. However, you can use volume mirroring to create copies between storage pools with separate extent sizes.
- Storage pools for performance and capacity:
  - Before deciding whether to create a single or multiple storage pools, carefully evaluate which option best fits the solution needs, considering data availability and recovery management. Storage pool design affects the extents that make up a volume. The extents are the mapping to the disk storage that affects performance of the volume.
- Reliability, availability, and serviceability (RAS):
  - With external storage license, it might make sense to create multiple storage pools in circumstances where a host only gets its volumes built from one of the storage pools. If the storage pool goes offline, it affects only a subset of all the hosts using the IBM FlashSystem V9000.
  - If you do not isolate hosts to storage pools, create one large storage pool. Creating one large storage pool assumes that the MDisk members are all of the same type, size, speed, and RAID level.
  - The storage pool goes offline if any of its MDisks are not available, even if the MDisk has no data on it. Therefore, do *not* put MDisks into a storage pool until they are needed.
  - If needed, create at least one separate storage pool for all the image mode volumes.
  - Make sure that the LUNs that are given to the IBM FlashSystem V9000 have all host-persistent reserves removed.

2.3.6 EasyTier

IBM EasyTier is a function that automatically and non disruptively moves frequently accessed data from HDD MDisks to flash drive MDisks, thus placing such data in a faster tier of storage. With version 7.8 and higher, EasyTier supports 4 tiers of storage.

The IBM FlashSystem V9000 supports these tiers:

- Tier 0 flash: Specifies a tier0_flash IBM FlashSystem MicroLatency module or an external MDisk for the newly discovered or external volume.
- Tier 1 flash: Specifies a tier1_flash (or flash SSD drive) for the newly discovered or external volume.
- Enterprise tier: Enterprise tier exists when the pool contains enterprise-class MDisks, which are disk drives that are optimized for performance.
- Nearline tier: Nearline tier exists when the pool contains nearline-class MDisks, which are disk drives that are optimized for capacity.

All MDisks belong to one of the tiers, which includes MDisks that are not yet part of a pool.

If the AE3 storage enclosure is used in an EasyTier pool and encryption is enabled in V9000 AC3 on the pool, then the AC3 nodes will send encrypted, incompressible data to the AE3. The Spectrum Virtualize software detects, if an mdisk is encrypted by the FlashSystem. Therefore, if an AE3 flash enclosure will be part of an encrypted easy-tier pool, encryption must be enabled on the AE3 BEFORE it is enabled in the EasyTier pool.

IBM Spectrum Virtualize does not attempt to encrypt data in an array that is already encrypted, This will allows the hardware compression of the AE3 to be effective. However, there are cases in which using SVC's software compression is preferred, such as if there is highly compressible data, (for example 3:1 or higher). In these cases, both encryption and compression can be done by the AC3 control nodes.

For more information about EasyTier, see the IBM redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Storage pools have an EasyTier setting that controls how EasyTier operates. The setting can be viewed through the management GUI but can only be changed by the CLI.

By default the storage pool setting for EasyTier is set to `Auto` (`Active`). In this state, storage pools with all managed disks of a single tier have EasyTier status of `Balanced`.

If a storage pool has managed disks of multiple tiers, the EasyTier status is changed to `Active`. The **`chmdiskgrp -easytier off 1`** command sets the EasyTier status for storage pool 1 to `Inactive`. The **`chmdiskgrp -easytier measure 2`** command sets the EasyTier status for storage pool 2 to `Measured`.

Figure 2-13 shows four possible EasyTier states.

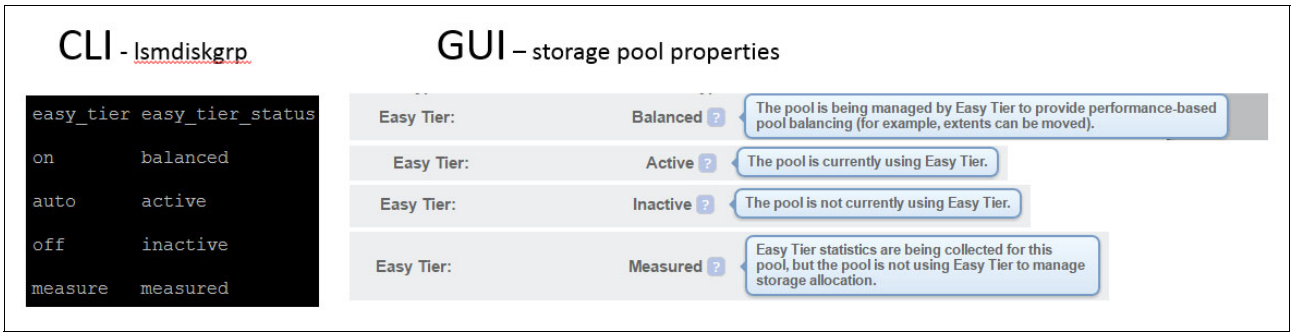


Figure 2-13 EasyTier status for CLI and GUI

EasyTier evaluation mode

EasyTier evaluation mode is enabled for a storage pool with a single tier of storage when the status is changed with the command line to `Measured`. In this state, EasyTier collects usage statistics for all the volumes in the pool. These statistics are collected over a 24-hour operational cycle, so you will have to wait several days to have multiple files to analyze. The statistics are copied from the control enclosures and viewed with the IBM Storage Tier Advisor Tool.

Instructions for downloading and using the tool are available in the “Extracting and viewing performance data with the IBM Storage Tier Advisor Tool” topic at IBM Knowledge Center.

This tool is intended to supplement and support, but *not* replace, detailed preinstallation sizing and planning analysis.

EasyTier considerations

When a volume is created in a pool that has EasyTier active, the volume extents are initially be allocated only from the Enterprise tier. If that tier is not present or all the extents have been used, the volume will be assigned extents from other tiers.

To ensure optimal performance, all MDisks in a storage pool tier must have the same technology and performance characteristics.

EasyTier functions best for workloads that have hot spots or data. Synthetic random workloads across an entire tier are not a good fit for this function. Also, you should not allocate all the space in the storage pool to volumes. You should leave some capacity free on the fastest tier for EasyTier to use for migration.

EasyTier volume planning

FlashSystem V9000 Flash Enclosure Model AE3 has the capability to alter the default quantity and size of managed disks. By default, the AE3 storage enclosure is configured with eight equal sized volumes referred to as managed disks (MDisks). For example, a 200TiB maximum effective capacity AE3 would then have eight 25TiB volumes.

This is different from the prior generation AE2 storage enclosure, which presented one large volume which was the entire capacity of the AE2 flash enclosure (for example 57TB). The capability to use individual MDisks (AE3 storage enclosure volumes) is important as it allows V9000 configurations that are intended to be used for EasyTier.

The following example assumes that only two storage tiers are used and the cold to hot capacity ratio is 5:1.

Setup 1: Two pools

AE3 storage enclosure with 200 TiB effective capacity, consisting of eight 25 TiB volumes, comprising the pool `Pool0`.

- Pool0 (200 TiB) has the following MDisks:

  mdisk0 (25 TiB), mdisk1(25 TiB), mdisk2 (25 TiB), mdisk3 (25 TiB), mdisk4 (25 TB), mdisk5 (25 TiB), mdisk6 (25 TiB), mdisk7(25 TiB)

  A frequent practice with EasyTier is to have a particular ratio of slower lower tier to that of high performance flash tier. If the goal is to use a 5 to 1 (5:1) ratio and you have 100TiB of spinning nearline HDD, then a 5:1 ratio would suggest that adding about 20TiB of AE3 flash then satisfies that ratio.

  To illustrate this configuration, consider the externally virtualized HDD capacity is a 100TiB known as mdisk10. The pool `Pool1` is formed with this mdisk.

- Pool1 (100 TiB) has the following MDisk:

  mdisk10 (100 TiB)

Setup 2: Two pools, one pool with an EasyTier configuration

AE3 with 200TB effective capacity, and mdisk10 as a virtualized HDD.

AE3 storage enclosure with 200TiB effective capacity, consisting of eight 25 TiB volumes, comprising the pool `Pool0`. Externally virtualized HDD capacity os one MDisk of 100TiB, comprising the pool `Pool1`. The corresponding MDisks are listed in Setup 1.

With the default AE3 being configured as eight equally sized mdisks, simply remove one of the mdisks from Pool0, then add it to Pool1. The resulting configuration then looks like:

- Pool0 (175 TiB) has following seven (7) MDisks:

  mdisk0 (25 TiB), mdisk1(25 TiB), mdisk2 (25 TiB), mdisk3 (25 TiB), mdisk4 (25 TB), mdisk5 (25 TiB), mdisk6 (25 TiB)

- Pool1 (125 TiB) has following MDisks:

  mdisk10 (100 TiB), mdisk7(25 TiB)

  The changed pool `Pool1` has a 4:1 ratio.

Setup 3: Two pools, one pool with an optimized EasyTier configuration

To get a more precise setting, the AE3 storage enclosure GUI will need to be used to alter MDisk sizes. The goal in this example is to get to a 5:1 ratio exactly, which means that mdisk7 as 25 TiB is too big. To solve this, we can go to the AE3 storage enclosure GUI and delete mdisk7, after unconfiguring it from any V9000 control enclosure pool.

Then, create a new mdisk7 as 20 TiB, and to not be wasteful of capacity, create a 5 TiB mdisk8. From the V9000 cluster GUI, detect these changes in the mdisks, then add mdisk7 (20 TiB) to Pool1 and mdisk8 (5iTB) to Pool0. The resulting configuration then looks like:

- Pool0 (180 TiB) has following seven (7) MDisks:

  mdisk0 (25 TiB), mdisk1(25 TiB), mdisk2 (25 TiB), mdisk3 (25 TiB), mdisk4 (25 TB), mdisk5 (25 TiB), mdisk6 (25 TiB), mdisk8 (5 TiB)

- Pool1 (125 TiB) has following MDisks:

  mdisk10 (100 TiB), mdisk7(20 TiB)

  Here, Pool1 is now exactly the desired 5:1 ratio, and all remaining other AE3 capacity of 180TB is in Pool0.

2.3.7 Volume configuration

An individual volume is a member of one storage pool and one I/O Group:

- The storage pool defines which MDisks provided by the disk subsystem make up the volume.
- The I/O Group (two nodes make an I/O Group) defines which IBM FlashSystem V9000 nodes provide I/O access to the volume.

Important: No fixed relationship exists between I/O Groups and storage pools.

Perform volume allocation based on the following considerations:

- Optimize performance between the hosts and the IBM FlashSystem V9000 by attempting to distribute volumes evenly across available I/O Groups and nodes in the clustered system.
- Reach the level of performance, reliability, and capacity that you require by using the storage pool that corresponds to your needs (you can access any storage pool from any node). Choose the storage pool that fulfills the demands for your volumes regarding performance, reliability, and capacity.
- I/O Group considerations:
  - With the IBM FlashSystem V9000, each building block that is connected into the cluster is an additional I/O Group for that clustered V9000 system.
  - When you create a volume, it is associated with one node of an I/O Group. By default, every time that you create a new volume, it is associated with the next node using a round-robin algorithm. You can specify a *preferred access node*, which is the node through which you send I/O to the volume rather than using the round-robin algorithm. A volume is defined for an I/O Group.
  - Even if you have eight paths for each volume, all I/O traffic flows toward only one node (the preferred node). Therefore, only four paths are used by the IBM Subsystem Device Driver (SDD). The other four paths are used only in the case of a failure of the preferred node or when concurrent code upgrade is running.
- Thin-provisioned volume considerations:
  - When creating the thin-provisioned volume, be sure to understand the utilization patterns by the applications or group users accessing this volume. You must consider items such as the actual size of the data, the rate of creation of new data, and modifying or deleting existing data.
  - Two operating modes for thin-provisioned volumes are available:
    - *Autoexpand volumes* allocate storage from a storage pool on demand with minimal required user intervention. However, a misbehaving application can cause a volume to expand until it has consumed all of the storage in a storage pool.
    - *Non-autoexpand volumes* have a fixed amount of assigned storage. In this case, the user must monitor the volume and assign additional capacity when required. A misbehaving application can only cause the volume that it uses to fill up.
  - Depending on the initial size for the real capacity, the grain size and a warning level can be set. If a volume goes offline, either through a lack of available physical storage for autoexpand, or because a volume that is marked as non-expand had not been expanded in time, a danger exists of data being left in the cache until storage is made available. This situation is not a data integrity or data loss issue, but you must not rely on the IBM FlashSystem V9000 cache as a backup storage mechanism.

Important:

- Keep a warning level on the used capacity so that it provides adequate time to respond and provision more physical capacity.
- Warnings must not be ignored by an administrator.
- Use the autoexpand feature of the thin-provisioned volumes.

  - When you create a thin-provisioned volume, you can choose the grain size for allocating space in 32 kilobytes (KB), 64 KB, 128 KB, or 256 KB chunks. The grain size that you select affects the maximum virtual capacity for the thin-provisioned volume. The default grain size is 256 KB, and is the preferred option. If you select 32 KB for the grain size, the volume size cannot exceed 260,000 GB. The grain size cannot be changed after the thin-provisioned volume is created.

    Generally, smaller grain sizes save space but require more metadata access, which could adversely affect performance. If you *will not be* using the thin-provisioned volume as a FlashCopy source or target volume, use 256 KB to maximize performance. If you *will be* using the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.
  - Thin-provisioned volumes require more I/Os because of directory accesses. For truly random workloads with 70% read and 30% write, a thin-provisioned volume requires approximately one directory I/O for every user I/O.
  - The directory is two-way write-back-cached (just like the IBM FlashSystem V9000 fast write cache), so certain applications perform better.
  - Thin-provisioned volumes require more processor processing, so the performance per I/O Group can also be reduced.
  - A thin-provisioned volume feature called *zero detect* provides clients with the ability to reclaim unused allocated disk space (zeros) when converting a fully allocated volume to a thin-provisioned volume using volume mirroring.

- Volume mirroring guidelines:
  - With the IBM FlashSystem V9000 system in a high performance environment, this capability is only possible with a *scale up* or *scale out* solution as the single expansion of the first building block only provides one MDisk in one storage pool. If you are considering volume mirroring for data redundancy, a second expansion with its own storage pool would be needed for the mirror to be on.
  - Create or identify two separate storage pools to allocate space for your mirrored volume.
  - If performance is of concern, use a storage pool with MDisks that share the same characteristics. Otherwise, the mirrored pair can be on external virtualized storage with lesser-performing MDisks.

2.3.8 SAN boot support

The IBM FlashSystem V9000 supports SAN boot or startup for IBM AIX®, Microsoft Windows Server, and other operating systems. SAN boot support can change, so check the following SSIC web page regularly.

2.4 License features

All FlashSystem V9000 model AE3 systems have the FlashSystem V9000 software pre-installed. One 5639-RB8 license is required for each model AE3, 12F, and 24F storage enclosure, with four 5639-RB8 licenses being required for each model 92F storage enclosure. These models are the FlashSystem virtualized storage and expansion enclosures.

The system requires Storage Capacity Units (SCU) licenses for any external systems that are being virtualized. More more information on IBM Spectrum Virtualize licensing, see the sales manual.

With the AE3 storage enclosure, there is also a licensed feature code for hardware assisted encryption:

- Feature code AF14 - Encryption Enablement Pack.

More more information on licensing, see the IBM FlashSystem V9000 Model AE3 Product Guide.

2.4.1 Encryption feature

The IBM FlashSystem V9000 Encryption feature is offered with the IBM FlashSystem V9000 under the following feature:

- Feature code AF14 - Encryption Enablement Pack:
  - Includes three USB keys on which to store the encryption key
  - Maximum feature quantity is eight (for a full *scale up and scale out* solution)
  - Enables data encryption at rest on the AE3 storage enclosure assigned MDisks

There are two ways to install encryption feature on the IBM FlashSystem V9000 as follows:

- USB Keys on each of the AE3 storage enclosures
- IBM Security Key Lifecycle Manager (SKLM)

You can use one or both ways to install encryption. Using both USB and SKLM methods together gives the most flexible availability of the encryption enablement.

Note: To invoke either method requires the purchase of the Feature code AF14 - Encryption Enablement Pack.

USB Keys

This feature supplies three USB keys to store the encryption key when the feature is enabled and installed. If necessary, there is a rekey feature that can also be performed. When the UBS keys encryption feature is being installed, the IBM FlashSystem V9000 AE3 GUI is used for each AE3 that will have the encryption feature installed. The USB keys must be installed in the USB ports in the rear of the AE3 storage enclosure.

Figure 2-14 (rear view) shows the location of USB ports on the AE3 storage enclosure.

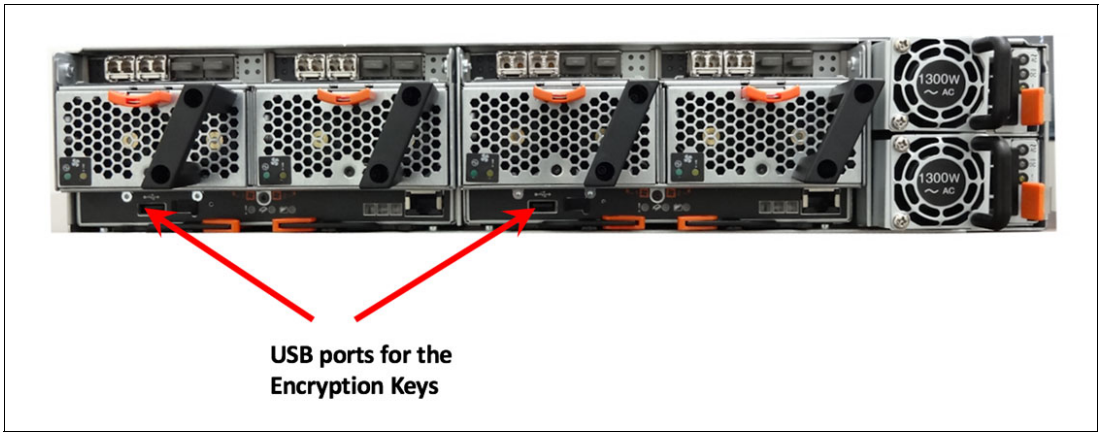


Figure 2-14 Location of USB ports on the AE3 storage enclosure

IBM Security Key Lifecycle Manager (SKLM)

IBM FlashSystem V9000 Software V8.1 adds improved security with support for encryption key management software that complies with the Key Management Interoperability Protocol (KMIP) standards, such as IBM Security Key Lifecycle Manager (SKLM) to help centralize, simplify, and automate the encryption key management process.

Before IBM FlashSystem V9000 Software V8.1, you could enable encryption by using USB flash drives to copy the encryption key to the system.

Note: If you are creating a new cluster with V 8.1, you have the option to either use USB encryption or key server encryption or both. The USB flash drive method and key server method can be used in parallel on the same system. Existing clients that are currently using USB encryption will be able to move to key server encryption. The migration of a local (USB) key to a centrally managed key (SKLM key server) is also available.

Encryption summary

Encryption can take place at the hardware or software level.

Encryption at the AE3 level (hardware)

A IBM FlashSystem V9000 with an AE3 enclosure supports hot encryption activation when enabling encryption in the storage enclosure. With hot encryption activation, you can enable encryption on an existing flash array without having to remove the data. Enabling encryption this way is a non-destructive process.

Hardware encryption is the preferred method for AE3 storage enclosures because this method works with the hardware compression that is built in to the Flash Modules of the AE3 flash storage enclosure.

Note: When an AE3 storage enclosure is configured with 18 TB modules, then compression using RtC may be more effective for data with compression rations greater than 1.2:1. Using RtC may not deliver the lowest FlashSystem 900 AE3 latency.

Encryption at the AC3 level (software)

With highly compressible data, (greater than a 2.5:1 ratio), compression using the RtC engine in the IBM FlashSystem V9000 control nodes makes more effective use of space on the flash enclosure. If using this compression option, then either software or hardware encryption is acceptable.

Software encryption should be used with other storage that does not support its own hardware encryption. For more information about encryption technologies supported by other IBM storage devices, see the *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

2.4.2 Compression

There are two ways to compress data on the V9000 depending on the type of storage attached to the system as follows:

- Real-time Compression (RtC)
- AE3 In-line Hardware Compression

FlashSystem V9000 AE3 storage enclosure in-line hardware compression is always on. The best usable to maximum effective capacity ratio is depending on the MicroLatency module capacity. Some workload not demanding lowest latency and having a good possible compression rate could be a candidate for using RtC. Also see “Physical and effective capacity based on compression rates” in the IBM Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

Real-time Compression

The IBM FlashSystem V9000 Real-time Compression (RtC) feature uses additional hardware that is dedicated to the improvement of the Real-time Compression functionality. When ordered, the feature includes two Compression Acceleration Cards per control enclosure for the I/O Group to support compressed volumes.

The compression accelerator feature is ordered, by default, with RtC software: Feature code AH1A - Compression Acceleration Card: The quantity of the Compression Acceleration Cards per controller is either zero (0) or two (2).

With two Compression Acceleration cards in each node (a total of four cards per I/O group), the total number of managed compressed volumes is up to 512 per I/O group. RtC type compression would be used on the previous generation IBM FlashSystem V9000 AC2 or AC3 and AE2 combinations or for externally virtualized storage systems that do not support their own compression function.

AE3 Inline Hardware Compression

The IBM FlashSystem V9000 AE3 storage enclosure has in-line hardware compression as part of its architecture. This type of compression is “always on” and cannot be switched off. For further details of the AE3 compression, its architecture and operation, please see the Architecture topic in *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

2.5 IBM FlashSystem V9000 configuration backup procedure

Configuration backup is the process of extracting configuration settings from a clustered system and writing it to disk. The configuration restore process uses backup configuration data files for the system to restore a specific system configuration. Restoring the system configuration is an important part of a complete backup and disaster recovery solution.

Only the data that describes the system configuration is backed up. You must back up your application data by using the appropriate backup methods.

To enable routine maintenance, the configuration settings for each system are stored on each node. If power fails on a system or if a node in a system is replaced, the system configuration settings are automatically restored when the repaired node is added to the system. To restore the system configuration in a disaster (if all nodes in a system are lost simultaneously), plan to back up the system configuration settings to tertiary storage. You can use the configuration backup functions to back up the system configuration. The preferred practice is to implement an automatic configuration backup by applying the configuration backup command.

The virtualization map is stored on the quorum disks of external MDisks, and is accessible to every IBM FlashSystem V9000 control enclosure.

For complete disaster recovery, regularly back up the business data that is stored on volumes at the application server level or the host level.

Before making major changes to the IBM FlashSystem V9000 configuration be sure to save the configuration of the system. By saving the current configuration, you create a backup of the licenses that are installed on the system. This can assist you in restoring the system configuration. You can save the configuration by using the “svcconfig backup” CLI command.

The next two steps show how to create a backup of the V9000 AC3 configuration file and to copy the file to another system:

1. Log in to the cluster IP using an SSH client and back up the IBM FlashSystem V9000 configuration. Example 2-2 shows the output of the “svcconfig backup” CLI command.

Example 2-2 Output of the “svcconfig backup” CLI command

```
superuser> svcconfig backup
.............................................................
CMMVC6155I SVCCONFIG processing completed successfully
```

2. Copy the configuration backup file from the system. Using secure copy, copy the following file from the system and store it:

   `/tmp/svc.config.backup.xml`

For example, use **`pscp.exe`**, which is part of the PuTTY commands family. Example 2-3 shows the output of the **`pscp.exe`** CLI command.

Example 2-3 Using pscp.exe

```
pscp.exe superuser@<cluster_ip >:/tmp/svc.config.backup.xml .
superuser@ycluster_ip> password:
svc.config.backup.xml | 163 kB | 163.1 kB/s | ETA: 00:00:00 | 100%
```

This process also needs to be completed on each AE3 storage enclosure in the IBM FlashSystem V9000 cluster. You will need to log in to each of the AE3 cluster IP addresses, using an SSH client and run the “svcconfig backup” command on each of the FlashSystem AE3 attached storage enclosures.

Note: This process saves only the configuration of the V9000 system. User data must be backed up by using normal system backup processes

3

Scalability

This chapter describes the scaling capabilities of IBM FlashSystem V9000:

- Scale out for capacity
- Scale up for performance

A single IBM FlashSystem V9000 storage building block consists of two IBM FlashSystem V9000 control enclosures (AC3) and one IBM FlashSystem V9000 storage enclosure (AE3). Additionally, the AC3 control enclosures can be configured with SAS-enclosures for capacity expansion.

The examples of scaling in this chapter show how to add control enclosures, a storage enclosure, and an expansion enclosure, and how to configure scaled systems.

This chapter demonstrates scaling out with additional building blocks and adding one additional storage enclosure. This setup consists of two IBM FlashSystem V9000 building blocks configured as one IBM FlashSystem V9000 cluster.

This chapter includes the following topics:

- Overview
- Building block for scaling
- Adding an IBM FlashSystem V9000 AE3 storage enclosure
- Adding a second IBM FlashSystem V9000 building block
- Adding an IBM FlashSystem V9000 expansion enclosure (12F, 24F, and 92F)
- Planning
- Installing
- Operations
- Concurrent code load in a scaled-out system

3.1 Overview

IBM FlashSystem V9000 has a scalable architecture that enables flash capacity to be added (scaled up) to support multiple applications. The virtualized system can also be expanded (scaled out) to support higher IOPS and bandwidth, or the solution can be simultaneously scaled up and out to improve capacity, IOPS, and bandwidth while maintaining MicroLatency. As a result, your organization can gain a competitive advantage through MicroLatency response times and a more efficient storage environment. IBM FlashSystem V9000 has the following scalability features per building block:

- Slots for up to 12 hot-swappable flash memory modules (3.6 TB, 8.5 TB, or 18 TB modules)
- Configurable 14.4 - 180 TB of usable capacity for increased flexibility per storage enclosure
- Up to 20 standard expansion enclosures per controller pair (up to 80 total) with up to 9.6 PB raw capacity using NL-SAS HDDs or 29.4 PB raw capacity using SSDs
- Up to 8 high-density (HD) expansion enclosures per controller pair (up to 32 total) with up to 29.4 PB raw capacity using NL-SAS HDDs or 32 PB raw capacity using SSDs
- IBM FlashSystem V9000 has the following flexible scalability configuration options:
  - Scale up: Add more flash capacity
  - Scale up: Add more SAS capacity
  - Scale out: Expand virtualized system
  - Scale up and out: Add more flash and SAS capacity and expand virtualized system

Four types of storage enclosures are discussed in this chapter:

- IBM FlashSystem V9000 storage enclosure (AE3)
  - Native IBM FlashSystem V9000 storage
  - Fibre channel attached
  - Based on MicroLatency Modules (flash modules)
- IBM FlashSystem V9000 expansion enclosure (12F, 24F, or 92F)
  - SAS drive based either SSD or nearline drives
  - SAS attached
  - Used for capacity expansion

Figure 3-1 on page 63 shows all the types of building blocks and expansion enclosures available on the IBM FlashSystem V9000.

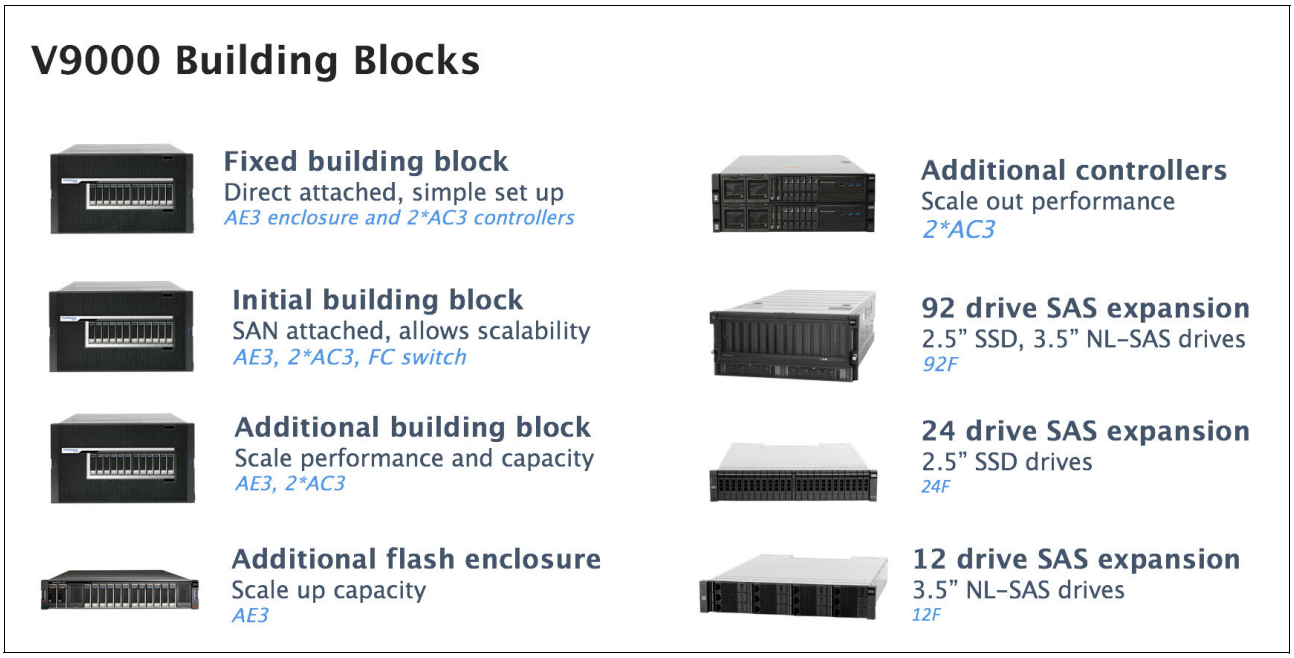


Figure 3-1 Types of building blocks and expansion enclosures available on the IBM FlashSystem V9000

3.2 Building block for scaling

A single IBM FlashSystem V9000 storage platform consists of two IBM FlashSystem V9000 control enclosures (AC3) directly cabled to one IBM FlashSystem V9000 storage enclosure (AE3), representing a *fixed* building block.

For balanced increase of performance and scale, up to four IBM FlashSystem V9000 building blocks can be clustered into a single storage system, multiplying performance and capacity with each addition. The *scalable* building blocks require connectivity through Fibre Channel switches. The scalable building block configurations also support the addition of up to four individual IBM FlashSystem V9000 storage enclosures to be added to the storage system.

If 720 TB of usable capacity from four building blocks is not enough capacity, up to four extra AE3 storage enclosures can then be added. In total, an IBM FlashSystem V9000 storage system can contain a maximum of eight IBM FlashSystem V9000 storage enclosures, offering a potential usable storage capacity of 1.44 PB, and up to 1.76 PB maximum effective capacity is available with in-line hardware compression on the AE3 storage enclosures.

Real-time Compression is available as a software feature, assisted by hardware accelerator cards in the IBM FlashSystem V9000 control enclosures. Real-time Compression enables users to deploy Real-time Compression where it is applicable.

Figure 3-2 on page 64 shows a summary of potential capacities for one IBM FlashSystem V9000 building block.

The scalable building blocks require connectivity through Fibre Channel switches.

A fixed building block uses direct internal connections without any switches. Contact your IBM representative if you want to scale up or scale out from a fixed building block.

V9000 AE3 Storage Capacity

	3.6 TB				8.5 TB			18 TB		
Quantity of Modules	6	8	10	12	8	10	12	8	10	12
Usable Capacity	14.44	21.66	28.88	36.1	51.3	68.4	85.5	108	144	180
Effective Capacity (AE3 and RtC mix)*	28	43	57	72	102	136	171	216	288	360
Effective Maximum (AE3 only)**	43	65	87	109	131	175	219	131	175	219
Effective Maximum (RtC only)***	72	108	144	180	256	342	427	540	720	900

* **Effective (AE3 and RtC mix)** capacity assumes a 2:1 reduction in data, using inline hardware compression on the 3.6 and 8.5TB modules and RtC on the 18TB modules.
** **Effective Maximum (AE3 only)** capacity is the maximum amount of data that can be stored, regardless of the compressibility of the workload.
*** **Effective Maximum (RtC only)** assumes a 5:1 reduction in data using Real-time Compression within the IBM Spectrum Virtualize software

Figure 3-2 Potential capacities for one IBM FlashSystem V9000 building block

Figure 3-3 illustrates the scalable capacity of IBM FlashSystem V9000. It also shows that extra AE3 storage enclosures can be added to a single building block, and also to two, three, or four building blocks.

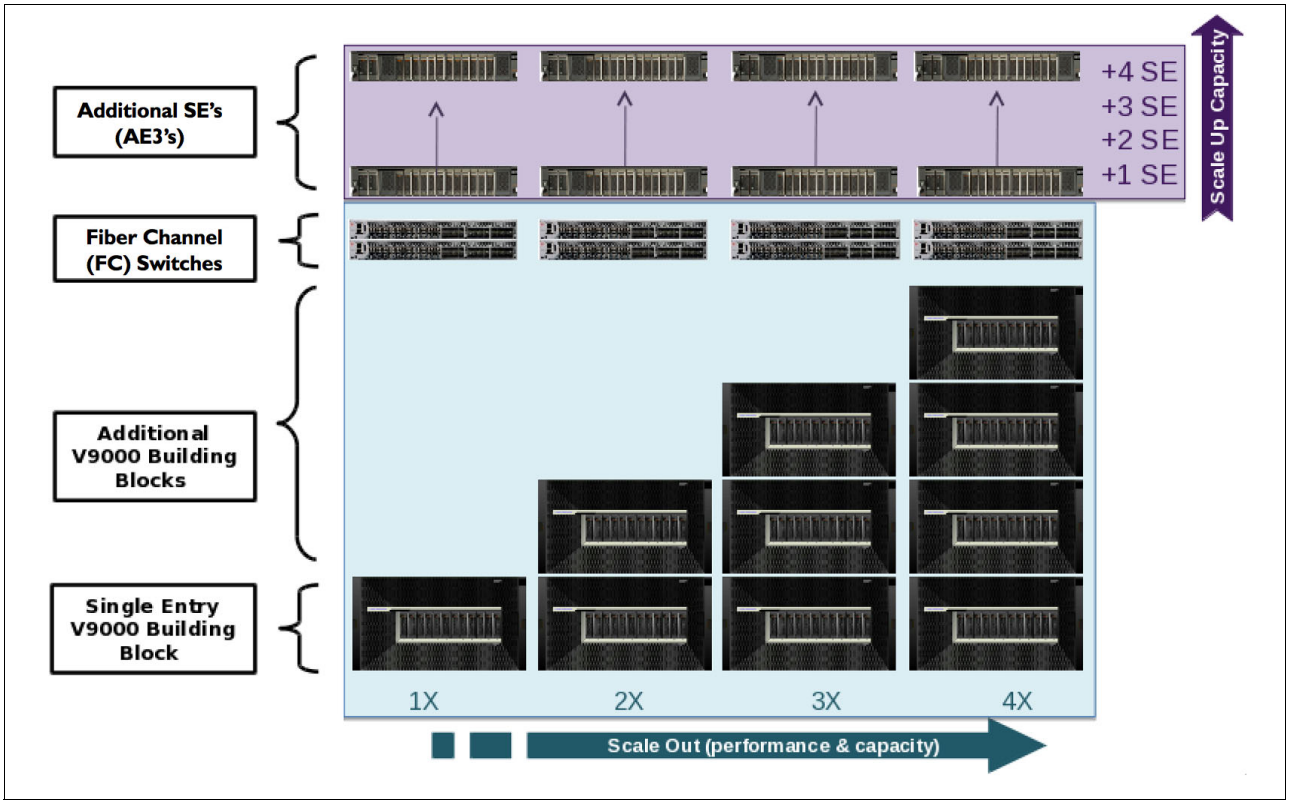


Figure 3-3 Scalable capacity of IBM FlashSystem V9000

3.2.1 Scaling concepts

IBM FlashSystem V9000 provides three scaling concepts:

- Scale up. Add more flash capacity:
  - Add up to four extra IBM FlashSystem V9000 storage enclosures.
- Scale up. Add more SAS capacity:
  - Add up to 80 IBM FlashSystem V9000 model 12F or 24F expansion enclosures.
  - Add up to 32 IBM FlashSystem V9000 model 92F expansion enclosures.
- Scale out. Expand virtualized system:
  - Add up to three IBM FlashSystem V9000 building blocks for extra performance and capacity.

The first scalable IBM FlashSystem V9000 building block consists of two IBM FlashSystem V9000 control enclosures (AC3), one IBM FlashSystem V9000 storage enclosure (AE3), representing a building block and two Fibre Channel switches for the internal 16 Gbps FC cabling. This building block with switches is called scalable building block.

Note: Internal FC speed for the AE3 storage enclosures can be either 16 Gbps or 8 Gbps. For the AC3 control enclosures only 16 Gbps is supported.

IBM FlashSystem V9000 can have up to four extra storage enclosures and scale out to four building blocks as shown in Figure 3-3. The maximum configuration has eight IBM FlashSystem V9000 control enclosures, and eight IBM FlashSystem V9000 storage enclosures.

3.2.2 Scale up for capacity

Scale up for capacity is adding an IBM FlashSystem V9000 AE3 storage enclosure to an existing building block. This storage enclosure will then be managed by its own GUI or CLI as the existing AE3 storage enclosures. This IBM FlashSystem V9000 might be a scalable building block or already be a scaled IBM FlashSystem V9000. Adding other storage to an IBM FlashSystem V9000, such as IBM Storwize V7000 or IBM FlashSystem 900, is not considered as IBM FlashSystem V9000 scale up, because it is not managed by the IBM FlashSystem V9000 and it is attached using the external fabric and not the internal switches.

To add an extra IBM FlashSystem V9000 AE3 storage enclosure, see section 3.3, “Adding an IBM FlashSystem V9000 AE3 storage enclosure” on page 71 To add an extra IBM FlashSystem V9000 expansion enclosure, see 3.5, “Adding an IBM FlashSystem V9000 expansion enclosure” on page 78.

3.2.3 Scale out for performance

Scaling out for performance is equivalent to adding a second, third, or fourth building block to a scalable building block. This additional building block is managed by the same GUI or CLI as the existing IBM FlashSystem V9000. This existing IBM FlashSystem V9000 might be a single scalable building block, so that the switches are already in place, or already be a scaled IBM FlashSystem V9000 of up to three building blocks.

Scale out always adds two controller nodes and one storage enclosure per building block to an existing IBM FlashSystem V9000.

To add another IBM FlashSystem V9000 building block, see 3.4, “Adding a second IBM FlashSystem V9000 building block” on page 78.

Table 3-1 summarizes the minimum and maximum capacity for scalable building blocks, including the addition of AE3 storage enclosures.

Table 3-1 IBM FlashSystem V9000, scalable building blocks including additional storage enclosures

Scalable building blocks (BB)	Minimum usable capacity (TB)	Maximum usable capacity (TB)	Maximum effective capacity (TB) with Inline HW Compression
1 BB	14.4	180	219
1 BB + 1 AE3	28.8	360	438
1 BB + 2 AE3	43.2	540	657
1 BB + 3 AE3	57.6	720	876
1 BB + 4 AE3	72.0	900	1095
2 BB	28.8	360	438
2 BB + 1 AE3	43.2	540	657
2 BB + 2 AE3	57.6	720	876
2 BB + 3 AE3	72.0	900	1095
2 BB + 4 AE3	86.4	1080	1314
3 BB	43.2	540	657
3 BB + 1 AE3	57.6	720	876
3 BB + 2 AE3	72.0	900	1095
3 BB + 3 AE3	86.4	1080	1314
3 BB + 4 AE3	100.8	1260	1533
4 BB	57.6	720	876
4 BB + 1 AE3	72.0	900	1095
4 BB + 2 AE3	86.4	1080	1314
4 BB + 3 AE3	100.8	1260	1533
4 BB + 4 AE3	116.0	1440	1752

PCIe expansion ports

Seven Peripheral Component Interconnect Express (PCIe) slots are available for port expansions in the IBM FlashSystem V9000 AC3 control enclosures.

Table 3-2 shows the host port count per building block configuration (1, 2, 3, or up to 4 building blocks).

Table 3-2 Host port count per building blocks

Building blocks	16 Gbps FC (host and storage)	10 Gbps iSCSI (host and storage)	10 Gbps FCoE (host)
1	32	8	8
2	64	16	16
3	96	24	24
4	128	32	32

For more detailed information about various interface setups, see Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245.

Expansion enclosures

IBM FlashSystem V9000 Software V8.1 introduces support for the addition of expansion enclosures also called tiered solution Models 9846/8-12F, 9846/8-24F and 9846/8-24F, which are available for the AC3 control enclosures.

IBM FlashSystem V9000 Small Form Factor (SFF) expansion enclosure model 24F offers new tiering options with low cost solid-state drives (SSDs). Each SFF expansion enclosure supports up to 24 2.5-inch low cost SSD drives.

Up to 20 expansion enclosures model 12F or 24F are supported per IBM FlashSystem V9000 building block, providing up to 480 drives with expansion enclosure model 24F (SFF) and up to 240 drives with expansion model 12F (LFF) for up to 2.4 PB of raw NL-SAS capacity in each building block. With four building blocks 9.6 PB of raw NL-SAS capacity is supported.

The IBM FlashSystem V9000 High-density (HD) Large Form Factor (LFF) Expansion Enclosure Model 92F supports up to 92 drives per enclosure, with a mixture of rotating disks and SSD drives in various capacities.

The HD Expansion Enclosure Model 92F provides additional configuration options. Up to eight HD expansion enclosures model 92F are supported per IBM FlashSystem V9000 building block, providing up to 736 drives for up to 7.3 PB of raw NL-SAS capacity or 11.3 PB SSD capacity in each building block. With four building blocks a maximum of 32 HD expansion enclosures model 92F can be attached giving a maximum 29.4 PB of raw NL-SAS capacity and 32 PB of raw SSD capacity. For information about the allowed intermix of expansion enclosures, see 2.2.4, “SAS expansion enclosures” on page 39.

Figure 3-4 shows the maximum possible configuration with a single building block using a combination of native IBM FlashSystem V9000 storage enclosures and expansion enclosures.

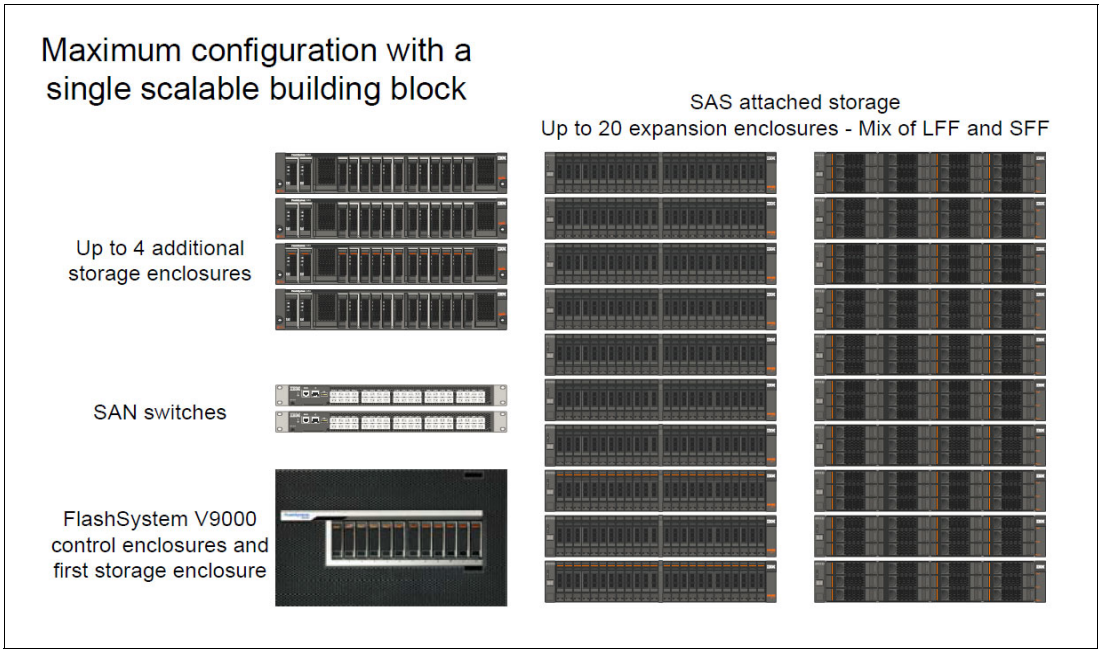


Figure 3-4 Maximum configuration with a single scalable building block using model 12F and 24F expansion enclosures

Table 3-3 shows the maximum capacities for such a configuration.

Table 3-3 IBM FlashSystem V9000 maximum capacities

	Model 12F 10TB NL-SAS	**Model 24F 15.36TB SSD**	**Model 92F 10TB NL-SAS**	**Model 92F 15.36TB SSD**
1 building block (8 x 92F)			7.3 PB	11.3 PB
4 building blocks (32 x 92F)			29.4 PB	32 PB[a]
1 building block (20 x 12F or 24F)	2.4 PB	7.3 PB		
4 building blocks (80 x 12F or 24F)	9.6 PB	29.4 PB		

a. IBM FlashSystem V9000 Version 7.7.1 has a maximum manageable capacity of 32 PB.

High-density (HD) solid-state drives (SSDs)

High-density SSDs allow applications to scale and achieve high performance while maintaining traditional reliability and endurance levels. 1.92 TB, 3.84 TB, 7.68 TB, and 15.36 TB SAS 2.5-inch SSD options are available for IBM FlashSystem V9000 SFF expansion enclosure model 24F for up to 7.3 PB raw SSD capacity in each building block for a maximum 29.4 PB with four building blocks.

With expansion enclosure model 92F 7.68 TB and 15.36 TB SSD drives are available for up to 11.3 PB raw SSD capacity in each building block for a maximum 32 PB with four building blocks.

High capacity nearline drives

High capacity nearline drives enables high value tiered storage with hot data stored in flash and warm data on lower cost NL-SAS HDDs all managed by IBM Easy Tier. The 10 TB SAS 3.5-inch nearline drives are available for IBM FlashSystem V9000 LFF expansion enclosure Model 12F and for Model 92F. Maximum capacities with four building blocks using expansion enclosure Model 12F is 9.6 PB raw nearline capacity and 29.4 PB using model 92F.

RAID types

RAID5 with standby hot spare is the only available RAID option for IBM FlashSystem V9000 native flash storage expansion. However, the additional SAS attached expansion enclosures can be configured with various RAID options. Distributed RAID (DRAID 5 and DRAID 6), which offers improved RAID rebuild times, is preferred for expansion enclosures.

3.2.4 SAS expansion enclosures intermix

IBM FlashSystem V9000 control enclosures with the SAS enclosure attach adapters support up to two SAS chains of expansion enclosures. The SAS chains have limits depending on the number of standard and dense expansion enclosures. Table 3-4 shows the allowed intermix of expansion enclosures per SAS chain.

Table 3-4 Number of expansion enclosures allowed per SAS chain; two chains per building block

Number of expansion enclosures	Config 1	Config 2	Config 3	Config 4	Config 5
1	Standard	Dense	Dense	Dense	Dense
2	Standard	Standard	Dense	Dense	Dense
3	Standard	Standard	Standard	Dense	Dense
4	Standard	Standard	Standard	Standard	Dense
5	Standard	Standard	Standard	Standard	
6	Standard	Standard	Standard		
7	Standard	Standard	Standard		
8	Standard	Standard			
9	Standard				
10	Standard				

Note: *Standard* refers to either Model 12F LFF or Model 24F SFF expansion enclosures. *Dense* refers to the Model 92F HD expansion enclosure.

Table 3-5 shows the allowed intermix of expansion enclosures per building block (control enclosure pair). Any cell in the table is valid but you must balance the enclosures across both chains up to the maximum limits specified per chain.

Table 3-5 Number of expansion enclosure types that can be intermixed on one building block; two SAS chains

Number of expansion enclosures	Config 1	Config 2	Config 3	Config 4	Config 5
1	Standard	Dense	Dense	Dense	Dense
2	Standard	Standard	Dense	Dense	Dense
3	Standard	Standard	Standard	Dense	Dense
4	Standard	Standard	Standard	Standard	Dense
5	Standard	Standard	Standard	Standard	Dense
6	Standard	Standard	Standard	Dense	Dense
7	Standard	Standard	Standard	Dense	Dense
8	Standard	Standard	Dense	Dense	Dense
9	Standard	Dense	Dense	Standard	
10	Standard	Standard	Standard	Standard	
11	Standard	Standard	Standard		
12	Standard	Standard	Standard		
13	Standard	Standard	Standard		
14	Standard	Standard	Standard		
15	Standard	Standard			
16	Standard	Standard			
17	Standard			**Legend:**	
18	Standard			Chain 1	
19	Standard			Chain 2	
20	Standard				

Note: *Standard* refers to either Model 12F LFF or Model 24F SFF expansion enclosures. *Dense* refers to the Model 92F HD expansion enclosure.

Note: To support SAS attached expansion enclosures, an AH13 - SAS Enclosure Attach adapter card must be installed in expansion slot 2 of each AC3 control enclosure in the building block.

3.3 Adding an IBM FlashSystem V9000 AE3 storage enclosure

This section gives an example of adding an extra IBM FlashSystem V9000 AE3 storage enclosure to a single scalable building block. Before scaling a building block, be sure that the FC cabling is set up and zoning on the switches has been implemented.

Note: The Fibre Channel internal connection switches are ordered together with the first IBM FlashSystem V9000 scalable building block. You can also supply your own Fibre Channel switches and cables, if they are supported by IBM. See the list of supported Fibre Channel switches at the SSIC web page.

Figure 3-5 shows a scalable building block before adding an extra IBM FlashSystem V9000 storage enclosure. Because the AE3 storage enclosure is now a virtualized enclosure, it does not show up on the AC3 GUI, and has its own GUI, as shown in Figure 3-6 on page 72.

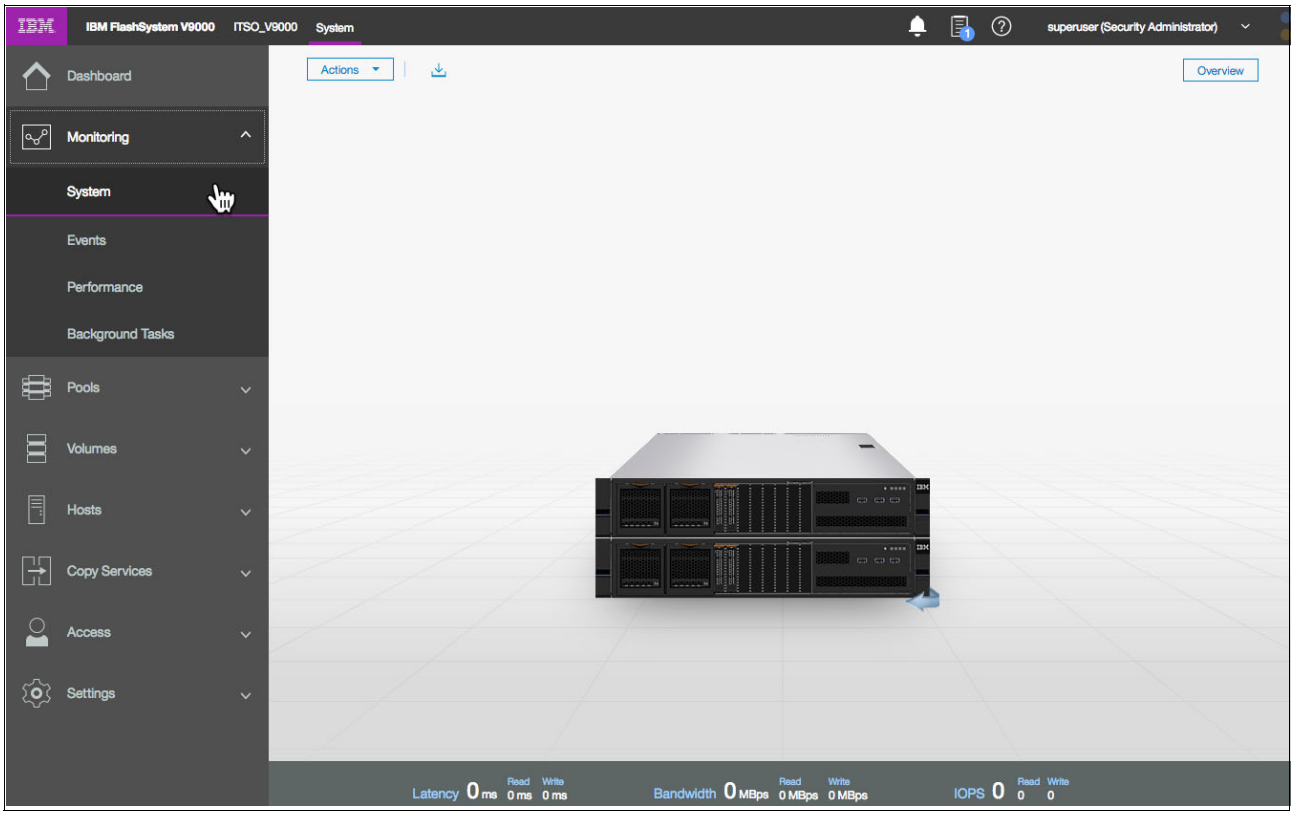


Figure 3-5 Single scalable building block showing the AC3 control enclosures

Figure 3-6 shows a single scalable building block for the AE3 storage enclosure.

Figure 3-6 Single scalable building block showing the AE3 storage enclosure

Note: The V9000 AC3 and AE3 enclosures now have a separate GUI for each element. The figures above show the pair of AC3's in one GUI and the AE3 in the other. The new additional AE3 being added will also have its own GUI as well.

To add an additional IBM FlashSystem V9000 AE3 storage enclosure to the existing AC3 cluster, the SSR will complete the following steps:

- Physically position and install the additional AE3 storage enclosure into the IBM FlashSystem V9000 rack.
- Initialize the V9000 AE3 using the USB initialization process and start the AE3 in the IBM FlashSystemV9000 GUI as shown in Figure 3-6. For further information see chapter Chapter 4, "Installation and configuration" on page 83.
- Create eight volumes on the AE3, using up all the available capacity of the storage enclosure. These volumes will be used by the V9000 AC3 control enclosures when discovering the volumes later in the initialization process.

  Using eight volumes enables you to use a part of the AE3 storage enclosure in an EasyTier pool. See the Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 for details about discovering these volumes and how to use them.

  The Lab Services or customer team can then plug the FC cables from the AE3 to the switches and verify that the V9000 AC3 cluster is able to see the new AE3 enclosure.

The following screens show how to add this new AE3 storage enclosure to the existing AC3 cluster.

Figure 3-7 shows the main dashboard GUI Page for the V9000 AC3.

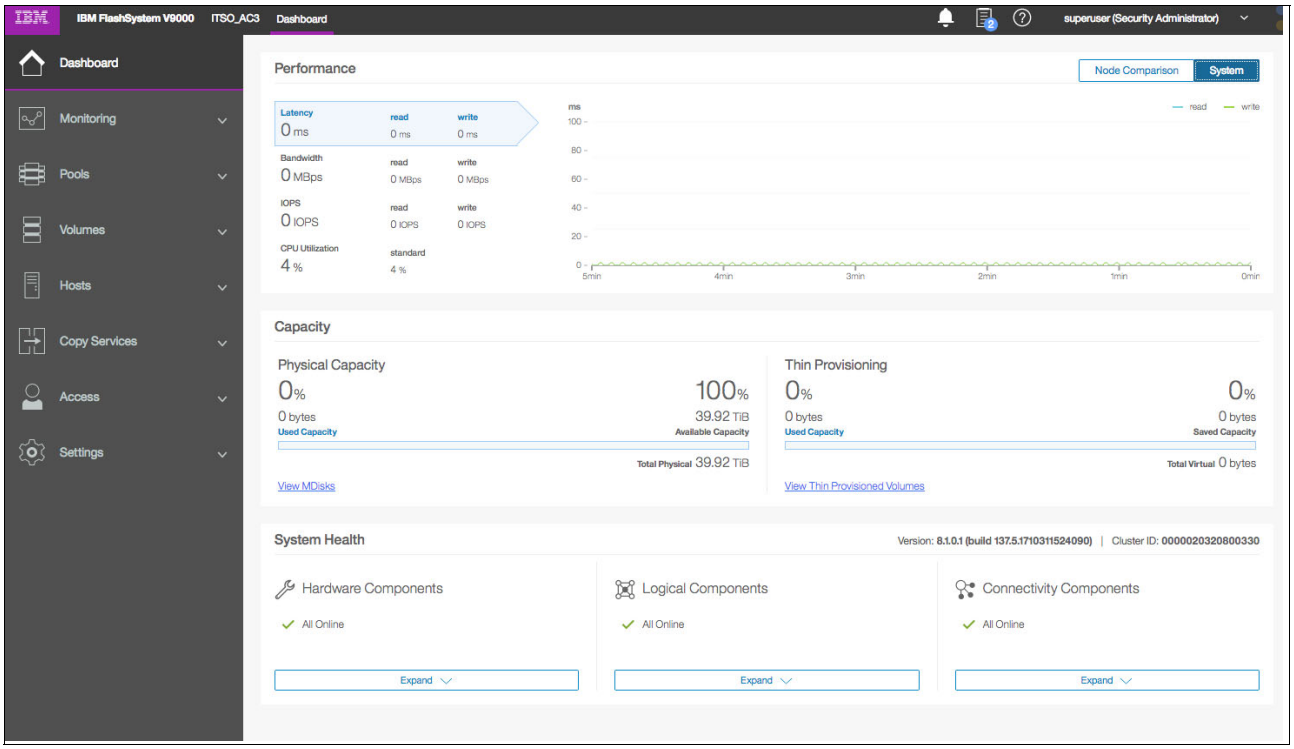


Figure 3-7 Dashboard page of the V9000 AC3 GUI

To add the AE3, complete the following steps:

1. From this screen shown in Figure 3-7, select the **Volumes** → **Volumes by Pool** option from the left hand menu. The new screen shown in Figure 3-8 displays, and the AE3 should show tn the available storage list. If it does not, then select the “Discover Storage” drop down to start this process.

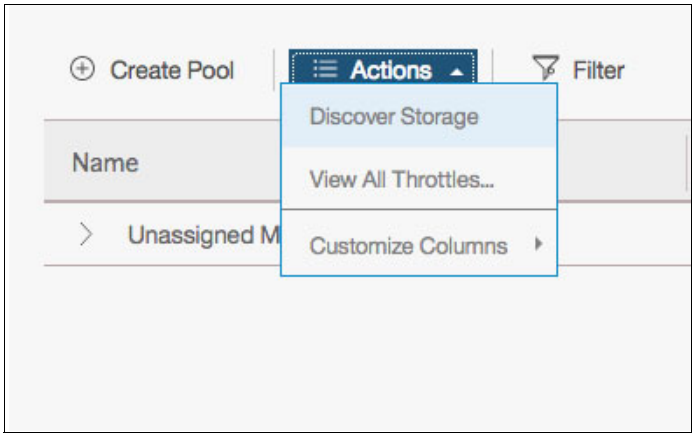


Figure 3-8 Discover Storage menu item

2. When the AE3 MDisks are visible on the **MDisks by Pools** screen, as shown in Figure 3-9, select the **Create Pool** option.

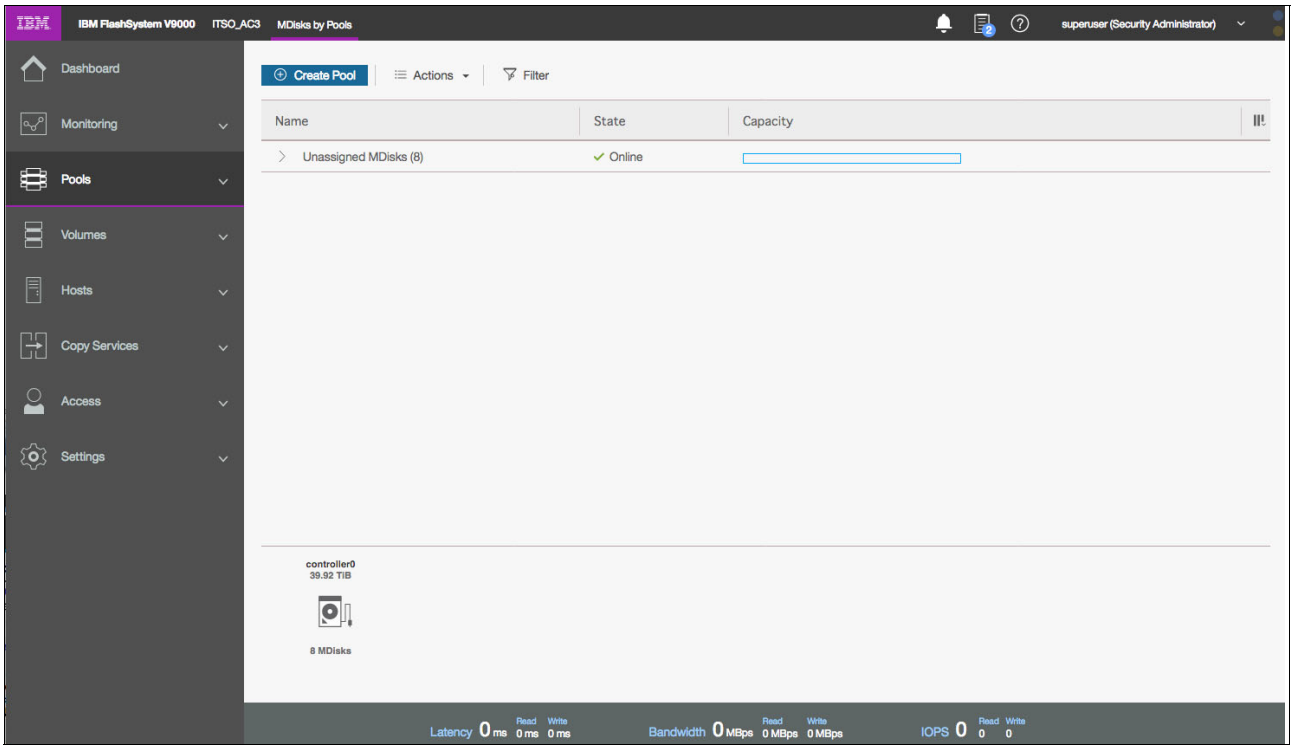


Figure 3-9 MDisk's by Pools screen

3. Click the **Create Pool** button using a name of your choice, as shown in Figure 3-9. When the Create Pool operation is running you see the display shown in Figure 3-10.

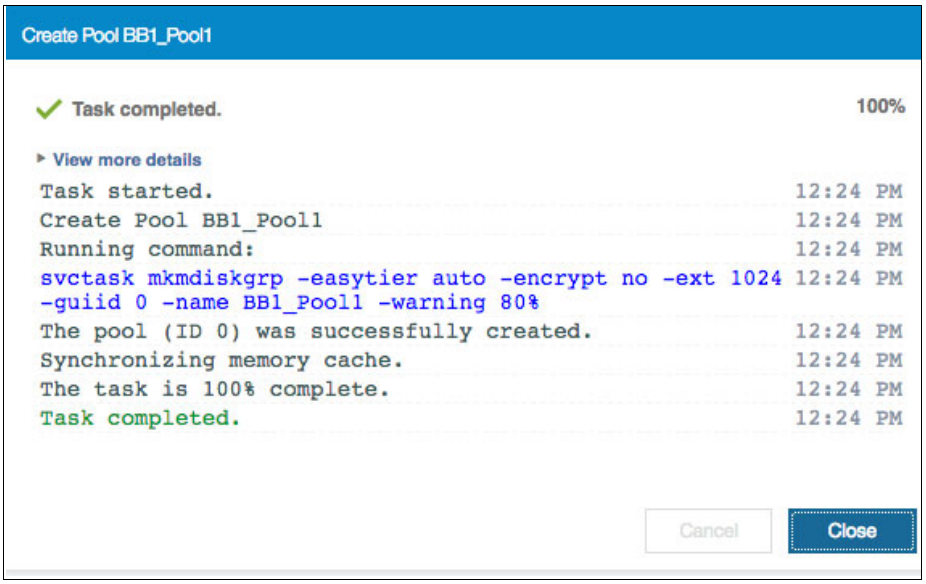


Figure 3-10 Create Pool operation

4. When the pool has been successfully created, you see the message shown in Figure 3-11. This informs you that you are required to **Add Storage** to the pool before you can create volumes.

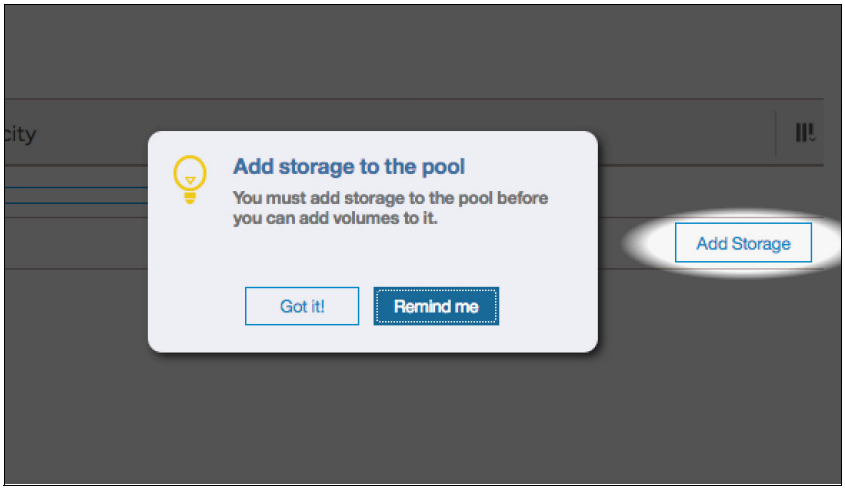


Figure 3-11 Add Storage message

5. Add storage to the pool you just created by clicking the **Add Storage** button on the right, as shown in Figure 3-11. You can also see a small icon at the bottom of the page in Figure 3-12 that shows the MDisks that are available to select in the next few screens, when creating this storage.

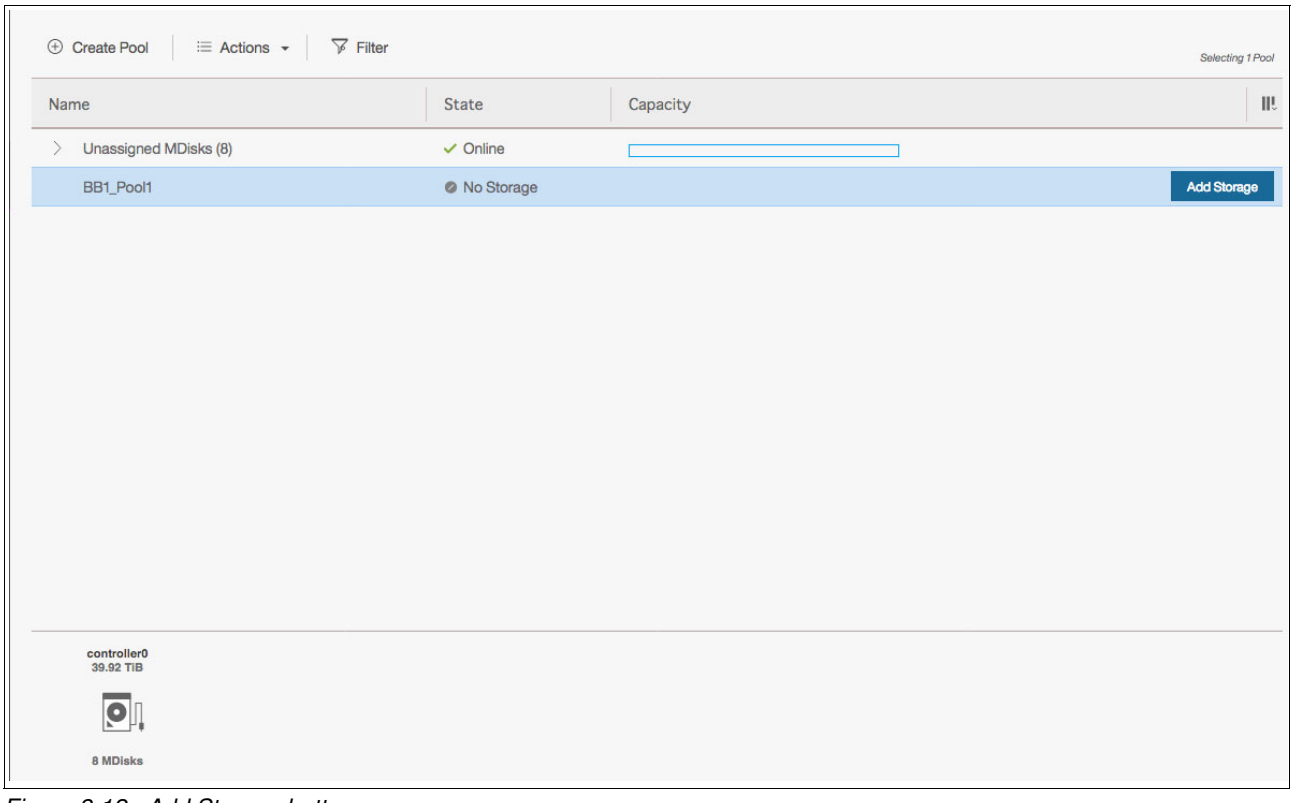


Figure 3-12 Add Storage button

6. Select the **External** option for the storage assignment and the **controller 0** in the menu shown in Figure 3-13.
7. The next menu is to select which MDisks you want the storage to reside on. Select the all checkbox to include all of the MDisks from the menu, as shown in Figure 3-13. This selects all of the 8 available MDisks.

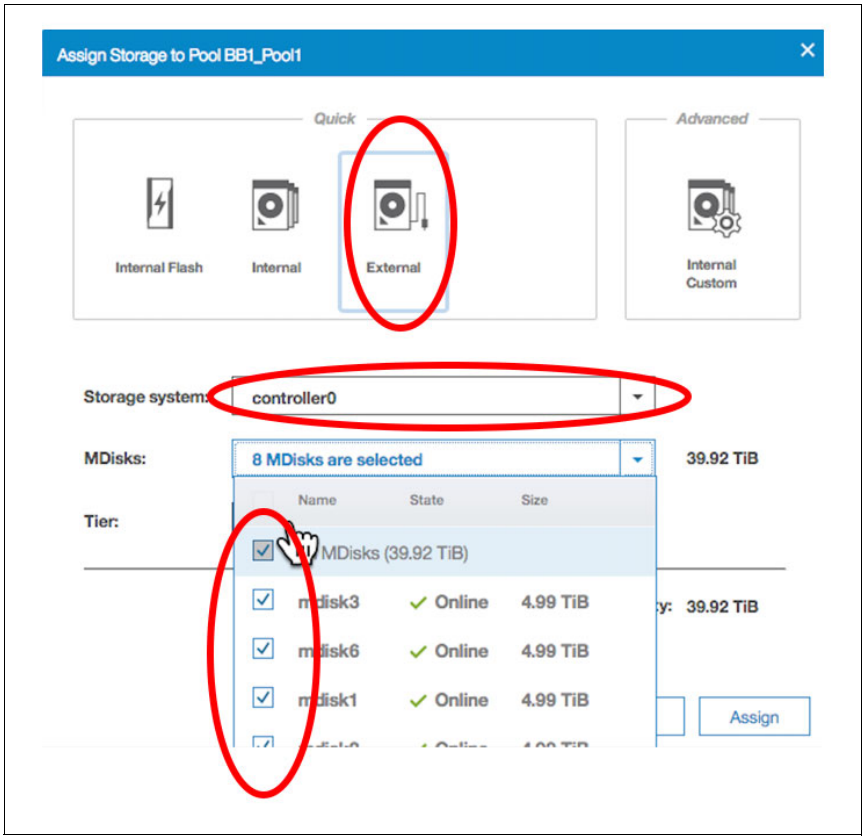


Figure 3-13 MDisk selection

8. Move to the next drop down to select the EastTier tier level. For V9000 AE3 enclosures, this is Tier 0 Flash, as shown in Figure 3-14 on page 77.

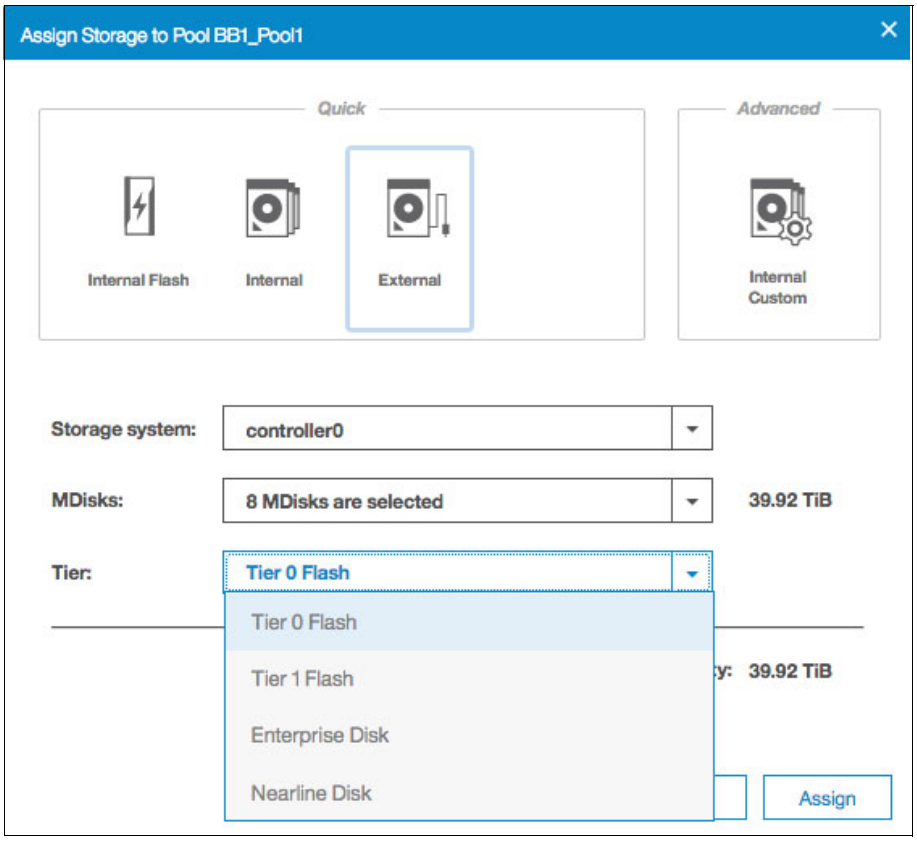


Figure 3-14 EasyTier level selection

9. After selecting the Storage system, MDisks, and Tier attributes for the storage, click the **Assign** button on the right hand side to start the process running. You will see the progress as shown in Figure 3-15.

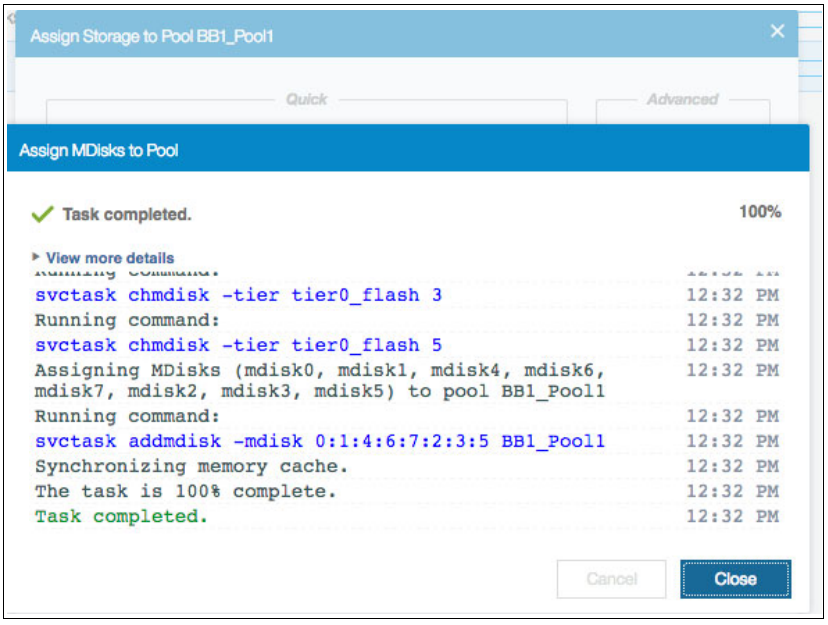


Figure 3-15 Assign MDisk to pool progress

10. Click the **Close** button to close this window once you see the “Task Completed” message.
11. The screen shown in Figure 3-16 shows the pool “BB_Pool” now has storage assigned from the 8 MDisks that were previously selected, and is now ready to add volumes as required.

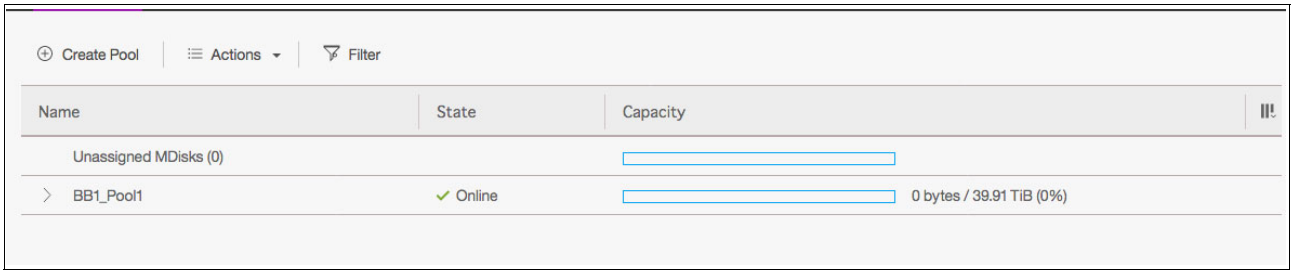


Figure 3-16 Pool storage assigned

Important: Before deciding whether to create a single or multiple storage pools, carefully evaluate which option best fits your solution needs, considering data availability and recovery management.

3.4 Adding a second IBM FlashSystem V9000 building block

This section discusses adding an extra IBM FlashSystem V9000 AC3-based building block to a single IBM FlashSystem V9000 AC3-based scalable building block.

For instructions about how to add a second IBM FlashSystem V9000 AC3-based building block to the first building block, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

3.5 Adding an IBM FlashSystem V9000 expansion enclosure

This section discusses adding an IBM FlashSystem V9000 expansion enclosure (12F, 24F or 92F) to a scalable or fixed building block. The expansion enclosure is added to a building block which has AC3 controller nodes.

For instructions about how to add an expansion enclosure to the IBM FlashSystem V9000 AC3-based building block, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

3.6 Planning

For more information about planning, see the following areas of this book:

- Chapter 2, “Planning” on page 23 describes details for planning the set up of a scaled IBM FlashSystem V9000.
- Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245 provides examples and guidelines for configuring port utilization and zoning to optimize performance and properly isolate the types of Fibre Channel traffic.

Guidelines are provided for two suggested methods of port utilization in an IBM FlashSystem V9000 scalable environment, dependent on customer requirements:

– IBM FlashSystem V9000 port utilization for *infrastructure savings*

  This method reduces the number of required Fibre Channel ports attached to the customer's fabrics. This method provides high performance and low latency, but performance might be port-limited for certain configurations. Intra-cluster communication and AE3 storage traffic occur over the internal switches.

– IBM FlashSystem V9000 port utilization for *performance*

  This method uses more customer switch ports to improve performance for certain configurations. Only ports that are designated for intra-cluster communication are attached to private internal switches. The private internal switches are optional and all ports can be attached to customer switches.

3.7 Installing

Chapter 4, "Installation and configuration" on page 83 includes details of how to install and configure IBM FlashSystem V9000. It describes the tasks that are completed by the IBM Service Support Representative or IBM lab-based services to set up the system and the follow-on task done by the customer.

3.8 Operations

The IBM FlashSystem V9000 GUI is the focal point for operating the system. However as the IBM FlashSystem V9000 AC3 and AE3 have separate GUI's for control of each of the elements, the majority of the operations for volumes and hosts manipulations are very similar to the IBM SVC Volume Controller.

For information about host and volume creation, see chapters 7 and 8 in the *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

3.9 Concurrent code load in a scaled-out system

This section demonstrates the IBM FlashSystem V9000 software update. Before you start a system update, be sure that the system has no problems that might interfere with a successful update. When the system uses HyperSwap volumes, make sure that all HyperSwap relationships have a status of `Online` by running the **`lsrcrelationship`** command or by using the GUI. Hosts must be configured with multipathing between the nodes of the accessing I/O group or groups when using HyperSwap.

Note: The software release notes contain the current information about the update.

This section includes a brief description of the update on an IBM FlashSystem V9000 scaled out system.

Figure 3-17 shows the System view of an IBM FlashSystem V9000 full scaled-out cluster using four building blocks and four additional IBM FlashSystem V9000 storage enclosures (in total, eight controller nodes and eight storage enclosures).

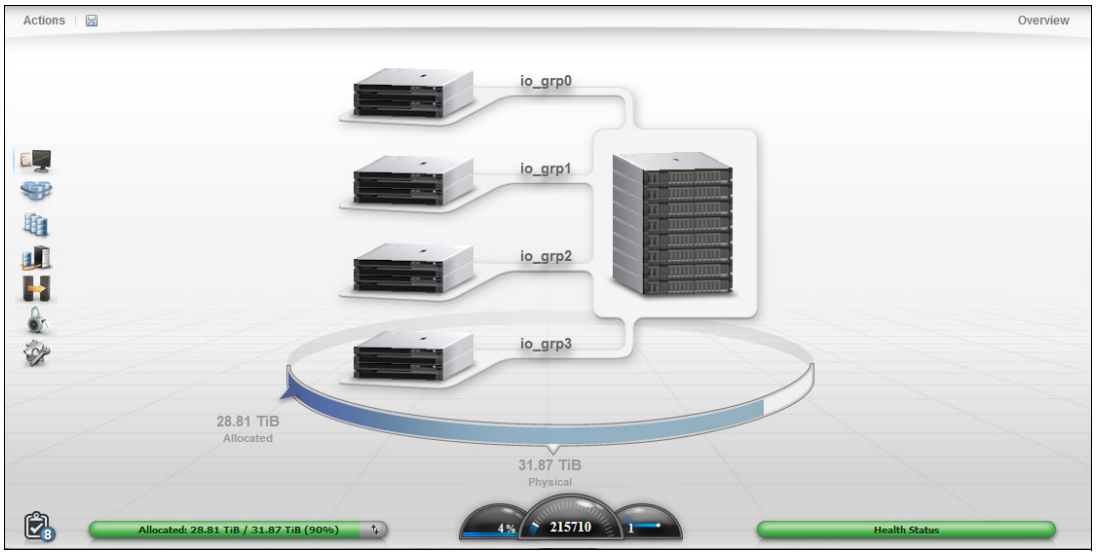


Figure 3-17 Full scaled-out and scaled-up IBM FlashSystem V9000

IBM FlashSystem V9000 AC3 update consists of three automated steps:

a. Update of one controller node per I/O group, one controller at a time.

b. Pause for approximately 30 minutes for host path discovery; hosts have to reconnect to the updated controller nodes.

c. Update of the other controller nodes of an I/O group, one controller at a time.

The update takes about 2.5 hours for a cluster with one building block. You can add 10 - 15 minutes per additional node.

Note: The IBM FlashSystem V9000 AE3 storage enclosures do not get upgraded by this process. They have to be upgraded as a stand-alone unit. See the upgrade section in *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

To upgrade the IBM FlashSystem V9000 AC3 control enclosures complete the following initial steps:

1. To start the concurrent code load (CCL), From the main dashboard screen click **Settings** → **System** → **Update System**. You are presented the screen shown in Figure 3-18.

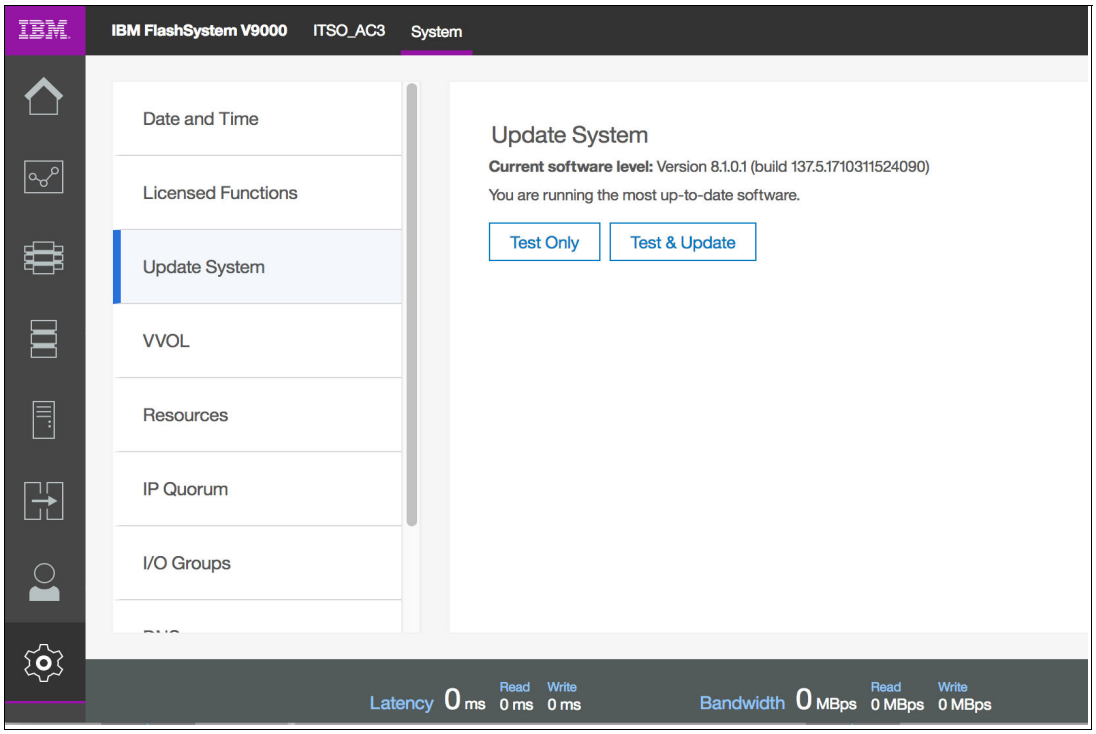


Figure 3-18 Update system

2. We now have two options as follows:
   - **Test Only**: this runs only the code upgrade test utility, but does not upgrade the V9000 code. This is useful if you want to check out the machine to see if it is capable of a code upgrade. It is totally concurrent and only does some checking and does not proceed with the code upgrade.
   - **Test & Update**: this runs the test code upgrade test utility, and if it passes, then proceeds to upgrade the code.
3. Click the **Test & Update** button.

4. The Update System wizard opens (Figure 3-19). Provide a name for the test utility and the update package. Use the folder buttons to select the correct file names and click **Update**.

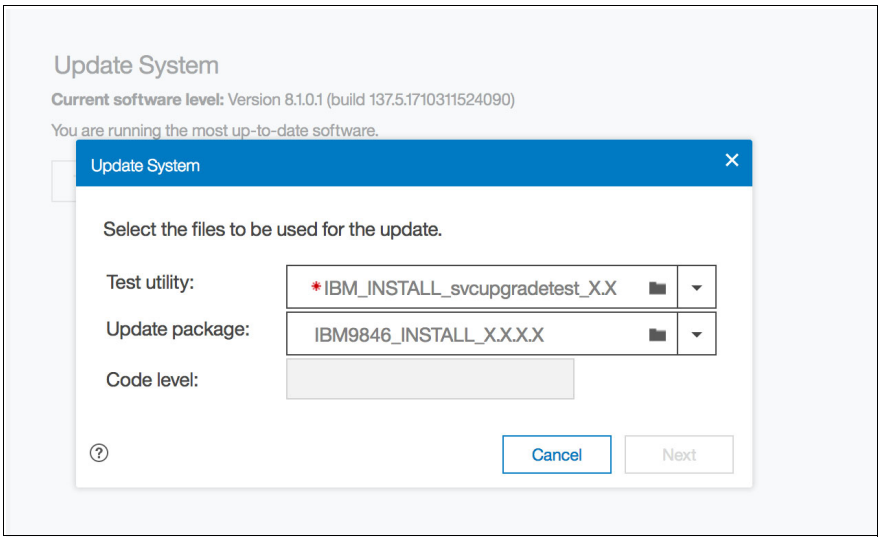


Figure 3-19 Update system file selection

As the upgrade progresses, you are presented with the screen show in Figure 3-20 where the overall progress bar is updated periodically, giving the most current status of the upgrade.

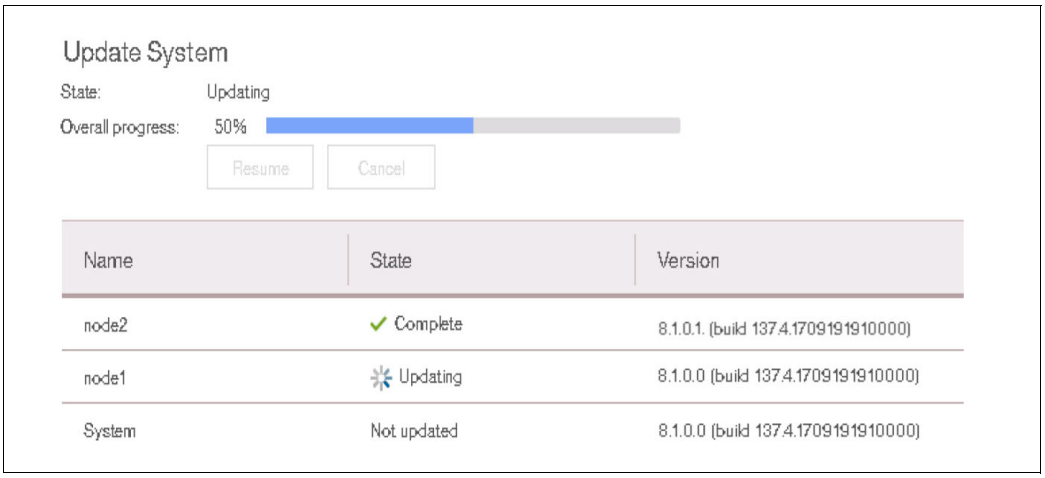


Figure 3-20 Overall progress and status bar

The full process for the code upgrade can be found in the Upgrading the System section in the Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

When completed, all of the IBM FlashSystem V9000 controller nodes are updated to the current software level.

4

Installation and configuration

This chapter shows how to install and configure IBM FlashSystem V9000. The system environmental requirements, cabling, and management are described. Installation from the initial setup procedure through configuring the system for use is demonstrated.

This chapter includes the following topics:

- Installation overview
- IBM FlashSystem V9000 physical specifications
- Installing the hardware
- Connecting the components
- Initial customer setup

4.1 Installation overview

Installation and initial configuration of IBM FlashSystem V9000 requires the completion of various tasks. An IBM Service Support Representative (SSR) installs a single fixed building block without switches. For scaled building blocks all configuration is done by IBM lab services. An SSR is responsible for the physical installation only. After this is done, the customer can then set up the system.

Important: The customer worksheets in 2.2.1, “Racking considerations” on page 31 must be completed before the installation because they determine the location of the components in the rack.

4.1.1 Tasks for the IBM SSR or IBM lab-based services

During installation, the IBM SSR or the IBM lab-based services perform the tasks described in this section. To learn about the steps that will be performed, limitations, and requirements, see the “Installing” topic at IBM Knowledge Center.

Hardware installation

To install the IBM FlashSystem V9000 hardware, an IBM SSR or IBM lab-based services must complete the following tasks:

1. Install the AC3 control enclosures and AE3 storage enclosure in the rack.
2. Connect the components as planned using Fibre Channel (FC) either in a direct attach configuration, or with FC switches for scaled configurations, and with switched fabrics for the host, the Ethernet management switch, and to the power distribution units.
3. Connect the components, using FC with or without switches, the Ethernet management switch, and connecting the components to the power distribution units:
   - For a fixed building block installation, the SSR completes the cabling.
   - For a scalable building block installation, IBM lab-based services completes the cabling.

Note: For details about the procedure to power up IBM FlashSystem V9000, see the topic about powering on and powering off an IBM FlashSystem V9000 building block.

Initial setup tasks

After the hardware is installed in the rack and cabled to meet the configuration you want, an IBM SSR or IBM lab-based services conducts the initial installation and configuration by performing the following tasks:

1. Initialize the AE3 storage enclosure using a USB key and performing initial setup.
2. Connect a workstation to an AC3 control enclosure technician port.
3. Configure the IBM FlashSystem V9000 clustered system with a name and management IP address.
4. Use the web browser to go to the management IP address of the cluster, and follow the steps of the setup wizard in the management GUI to set up call home and remote support by using information from the customer-supplied worksheets, shown in Chapter 2, “Planning” on page 23.

4.1.2 First customer involvement

After the IBM SSR or IBM lab-based services completes the service setup process, you log in to the AE3 storage and AC3 control enclosures and complete the following tasks by using the customer setup wizard as described in 4.5, “Initial customer setup” on page 91. The following steps are an overview of the process:

1. Confirm Settings on AE3 storage enclosure
2. Connect to the AC3 Control Enclosure to run the installation wizard
3. Change the system password
4. Change the system name
5. Configure licensed functions
6. Set the date and time
7. Confirm the call home settings that the IBM SSR or IBM lab-based services entered
8. Customer Creates Storage Pool
9. Add the AE3 storage enclosure managed disks

4.2 IBM FlashSystem V9000 physical specifications

The IBM FlashSystem V9000 is installed in a standard 19-inch equipment rack. The IBM FlashSystem V9000 building block is 6U high and 19 inches wide. A standard data 42U 19-inch data center rack can be used to populate with the maximum IBM FlashSystem V9000 configuration to use up to 36U. For a description of physical dimensions, see 1.1.3, “Physical specifications” on page 6.

4.3 Installing the hardware

The IBM SSR is responsible for physically installing IBM FlashSystem V9000 components in the rack.

This process involves installing the two control enclosures and the storage enclosure. These three enclosures form a 6U building block. For a scalable building block, additional components, such as Fibre Channel switches and an Ethernet switch are also installed.

When installed, the components of the IBM FlashSystem V9000 appear as shown in Figure 4-1. Only an 8 or 16 Gbps Fibre Channel attachment can be used for direct attach.

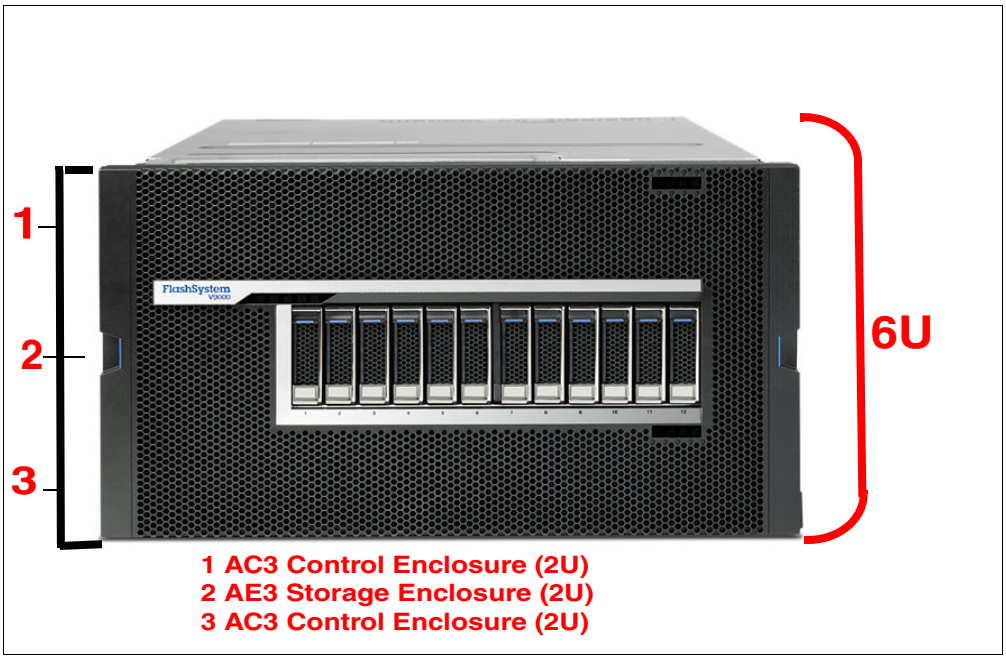


Figure 4-1 View of the IBM FlashSystem V9000 front panel

4.4 Connecting the components

The AE3 storage enclosure and AC3 control enclosures can be connected to create a fixed building block or a scalable building block:

- The IBM SSR is responsible for connecting (cabling) the three enclosures in a fixed building block.
- Lab-based services is responsible for cabling the AC3 control enclosures, the AE3 storage enclosure, and the Fibre Channel switches and Ethernet switch in a scalable building block.

For interface protocol and connection tables, see 2.2.3, “Network cable connections” on page 35. A complete suggested cabling guide is in the installation section of the IBM FlashSystem V9000 in IBM Knowledge Center.

4.4.1 Connecting the components in a fixed building block

The IBM SSR is responsible for connecting the components of a fixed building block.

To create a fixed building block:

1. Create direct links between the AE3 storage enclosure and the two AC3 control enclosures.
2. Create a direct link between the control enclosures.
3. Connect the hosts or external storage to the AC3 control enclosures.

To improve performance and provide redundancy, both AC3 control enclosures are connected to both canisters in the AE3 storage enclosure.

Figure 4-2 illustrates the interface protocols and maximum number of connections supported for an IBM FlashSystem V9000 fixed building block for a five-host, 16 Gbps Fibre Channel configuration. There are also redundant cluster links between the two AC3 control enclosures.

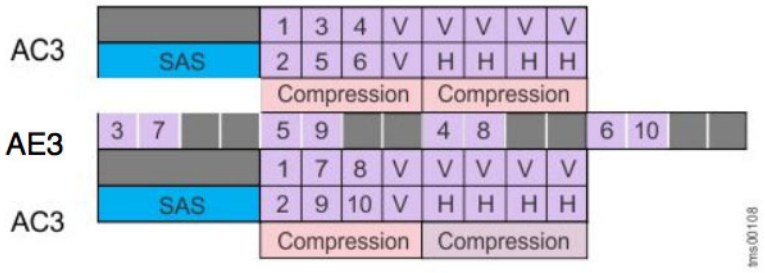


Table 1. Fixed building block with FC host connections legend

1 - 10	Connections between corresponding ports within fixed building block
H	Host port
V	Variable use - host connection, cluster, external virtualization, remote mirroring
SAS	SAS adapter in AC3 slot 2
Compression	Compression adapter in AC3 slots 5 and 8

Figure 4-2 Fixed building block protocols and connections for 16 FC Gbps configuration using AC3 control enclosures

Information about supported protocols and connections for IBM FlashSystem V9000 fixed and scalable building blocks is in For interface protocol and connection tables, see 2.2.3, “Network cable connections” on page 35. A complete suggested cabling guide is in the installation section of the IBM FlashSystem V9000 in IBM Knowledge Center.

Note: For the fixed building block 16 Gbps configuration, you need two FC ports (AC3 to AC3) and six FC ports (AC3 to AE3). The remaining six FC ports can be used for host or external storage attachment.

4.4.2 Connecting the components in a scalable building block

IBM lab-based services is responsible for cabling the building block components.

Cabling involves installing the network and storage data cables that are internal to the IBM FlashSystem V9000 scaled building blocks.

Figure 4-3 shows a conceptual view of a typical cabling example of one building block with an added AE3 storage enclosure.

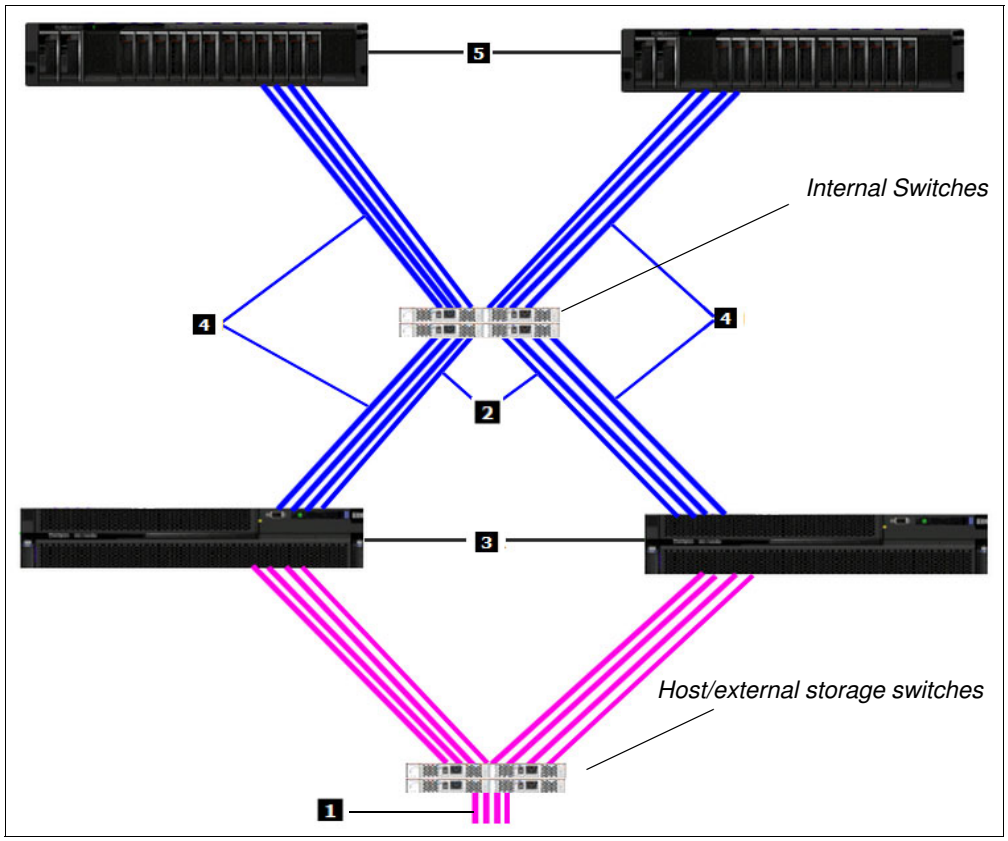


Figure 4-3 One building block and one additional AE3 storage enclosure connection

The numbers in the figure have the following meanings:

1. Host and external storage connections
2. Redundant cluster links between the AC2 or AC3 control enclosures
3. AC2 or AC3 control enclosures
4. Fibre Channel connections to the AE3 storage enclosures
5. AE3 storage enclosures

Remember that one building block contains two AC3 control enclosures and one AE3 storage enclosure. Your final configuration can be up to four building blocks and four additional AE3 storage enclosures.

For internal Fibre Channel switches, a minimum of two switches are needed for redundancy purposes on a 16 Gbps configuration and four switches for an 8 Gbps configuration (for AC3 control enclosure only). Switches that IBM suggests can be used, or any other switches can be used with the relevant attention on the open access method that is used by the AC3 control enclosure to reach the AE3 storage enclosure.

Note: When using customer SAN Switches, plan carefully to ensure that optimal performance is available to the V9000 internal communications.

Host connections to the AC3 control enclosures can be these connections types:

- Direct/switched Fibre Channel connections (16 Gbps or 8 Gbps)
- Fibre Channel Over Ethernet or ISCSI (10 Gbps Ethernet links)

The rear view of the AC3 control enclosure is shown in Figure 4-4.

Figure 4-4 AC3 controller enclosure rear view

For more information about connecting components, see the following IBM Knowledge Center topics:

- Connecting the components in a fixed building block (IBM SSR task)
- Connecting the components in a scalable building block (IBM lab-based services task)

Comprehensive examples and configuration guidelines of the following two preferred methods for port utilization in an IBM FlashSystem V9000 scalable environment are in Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245:

- IBM FlashSystem V9000 port utilization for *infrastructure savings*

  This method reduces the number of required customer Fibre Channel ports attached to the customer fabrics. This method provides high performance and low latency but performance might be port limited for certain configurations. Intra-cluster communication and AE3 storage traffic occur over the internal switches.

- IBM FlashSystem V9000 port utilization for *performance*

  This method uses more customer switch ports to improve performance for certain configurations. Only ports designated for intra-cluster communication are attached to private internal switches. The private internal switches are optional and all ports can be attached to customer switches. By following the cabling guidelines in the appendix you can see performance improvements up to 40% for sequential reads and up to 80% for sequential writes.

4.4.3 Ethernet cabling

IBM lab-based services is responsible for cabling components to the Ethernet switch.

The internal connectivity Ethernet switch provides management connections for the IBM FlashSystem V9000.

Figure 4-5 shows a typical wiring diagram for cabling the components to the Ethernet switch.

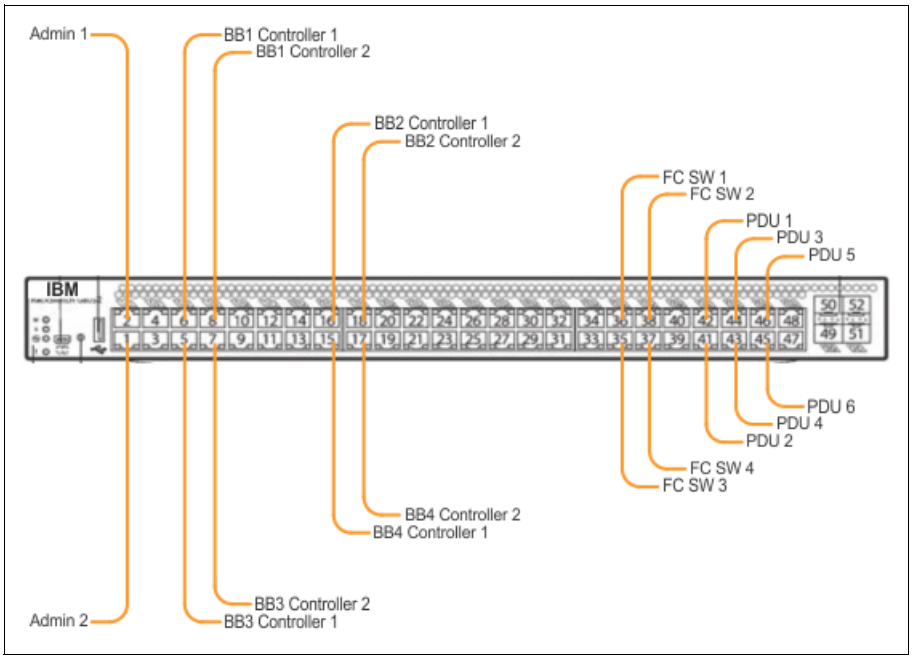


Figure 4-5 Ethernet switch wiring diagram

Switch port usage depends on the speed:

- If two 16 Gbps FC switches are used rather than four 8 Gbps FC switches, only two of the four Ethernet connections for FC switches are needed.
- Redundant Ethernet management links for AC3 control enclosures are not necessary to protect them from management link and port failures. The AC3 control enclosures function as a pair in an I/O group, and, if a node failure occurs, the partner AC3 control enclosure in the I/O group (or another AC3 control enclosure in the cluster) takes over management of the I/O group or cluster.

4.4.4 Scaling from one to two, three, or four building blocks

IBM lab-based services is responsible for the configuration tasks that are involved in scaling from one scalable building block to two, three, or four scalable building blocks.

Before you begin

Be sure these actions are complete before you scale the environment:

- The building block components must be installed in the rack in the specified locations.
- The building block components must be cabled.
- The first scalable building block must be operational.

Procedure

Follow this procedure to add a new building block:

1. Power on the components for the new building blocks.
2. Log on to the management GUI of the existing system. The system recognizes that there are new candidate enclosures and displays a graphic that shows the new enclosures.
3. Place the cursor over the graphic of a new enclosure, and then click to run the wizard that adds the new enclosure to the system.
4. From the networking section of the GUI, add the management IP addresses for the new building blocks. For details, see 2.3.1, “Management IP addressing plan” on page 41.

More information about various interface setups is in Appendix A, “Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment” on page 245.

4.5 Initial customer setup

After the service setup of the new system is complete, you can use the management GUI to input the initial setup information. FlashSystem V9000 with AE3 Storage enclosures have 2 setups:

- AE3 storage enclosure configuration setup is performed primarily by the IBM SSR using information provided by the customer during the planning process. The customer must confirm this information has been entered correctly.
- AC3 control enclosure setup involves the IBM SSR initializing the system and the customer completing the installation wizard which.

Before you begin, be sure that you have the following information:

- The cluster management IP address of both system entered by the IBM SSR
- The IP address of a Network Time Protocol (NTP) server, for automated setting of date and time
- Licensed function information
- The Simple Mail Transfer Protocol (SMTP) server IP address
- Any email contacts you want to be added in the notification mails

4.5.1 FlashSystem V9000 AE3 Setup Confirmation

Complete the following steps to confirm that the setup of the system is correct for Notifications, Network, Security, System, Support, and GUI Preferences.

In the following procedure we’ll touch briefly on each of the items for the FlashSystem V9000 AE3 Storage Enclosure, detailed explanations of items can be found in the Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

Start by connecting to the Management Interface:

1. Open a supported browser and go to the following address:

   `http://<AE3_management_ip_address>`

2. Login using User ID `superuser` and default password of `passw0rd`.

The AE3 dashboard displays, as shown in Figure 4-6. Notice on the left panel we've expanded the 1 **Settings** menu. We'll step through each of the items.

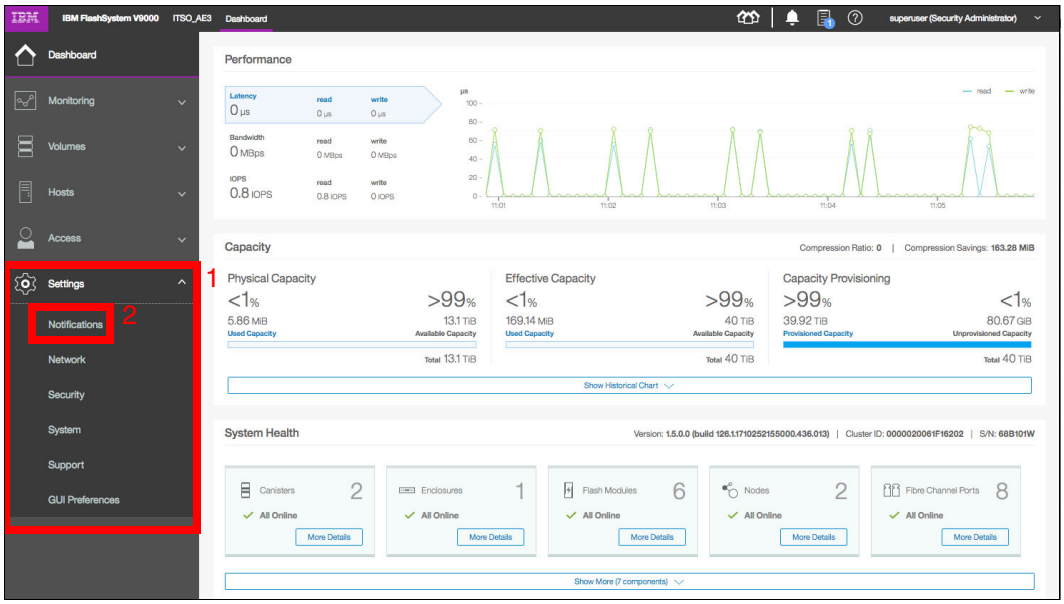


Figure 4-6 FlashSystem V9000 AE3 Management Dashboard

Notifications

In this section, Call Home is reviewed using the Notification selection 2 in the setting section of the main menu shown in Figure 4-6. IBM cannot overstate the value of call home for storage devices in your environment. Simply put, call home ensures IBM can automatically open service tickets for your system. IBM uses advanced analytical based on this call home information to open proactive tickets to replace components before they fail. It is the best way to get optimal support for your device. Ensuring your call home contact information is accurate and testing the process is highly recommended.

Email

In the notifications panel shown in Figure 4-7 on page 93, review the call home settings.

1. Notice that 1 Email Users **Email Address** is blank. This means that although call home message errors are sent to IBM 2, the customer would not receive notifications of errors. Create an email user for notifications by putting the form into 3 **Edit** mode and adding an email address for the notifications.

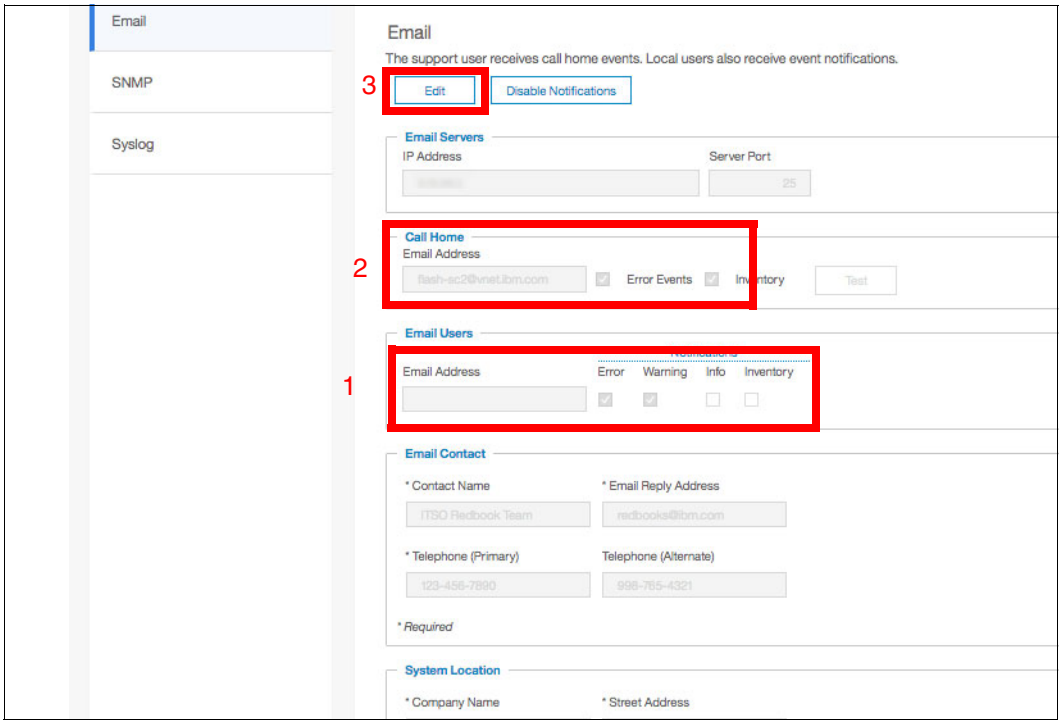


Figure 4-7 Discovered no email notification to customer

2. Most customers want an email when errors are sent to IBM. This is easily corrected by adding the team alias as an email user, as shown in Figure 4-8.

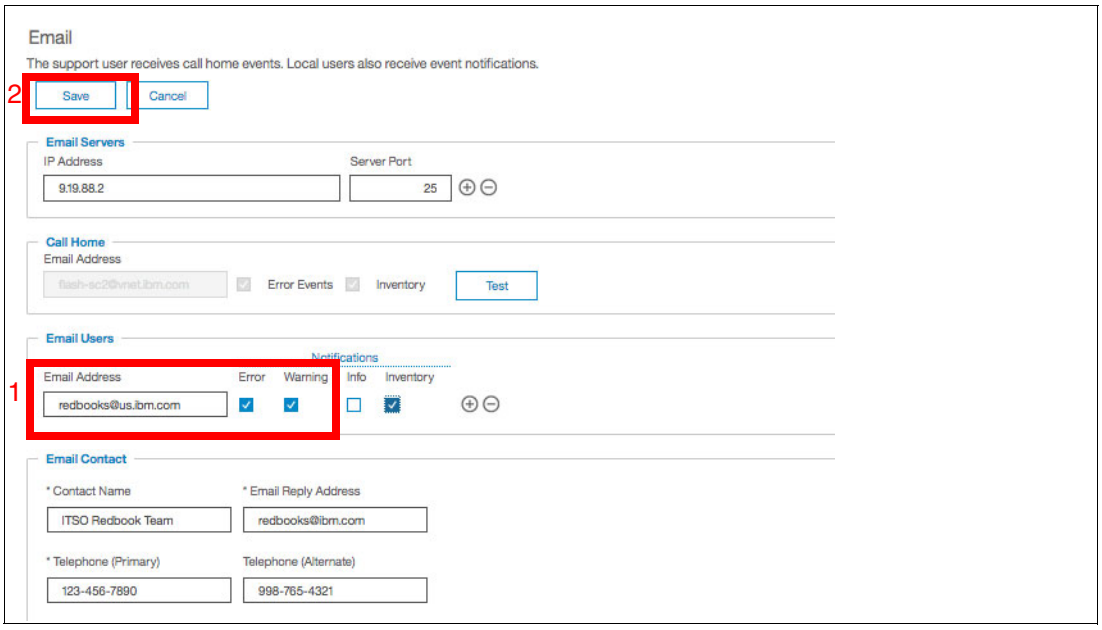


Figure 4-8 Adding Customer email Notification and Saving

3. With the form in Edit mode (Figure 4-8), add the user to be notified as shown in 1. Leaving the default warning level selected generates additional email traffic, so most customers do not leave this selected. It is necessary to 2 **Save** the form and re-edit the form before you can test the connections.

4. With the form again in edit mode in Figure 4-9 notice how both Call Home and Email Users each have a **Test** button enabled. Make sure to test both of them (1 & 2). In the case of call home, a Support ticket is automatically opened, and the contact in this form receives a communication from IBM. Use **Cancel** 3 to close the form.

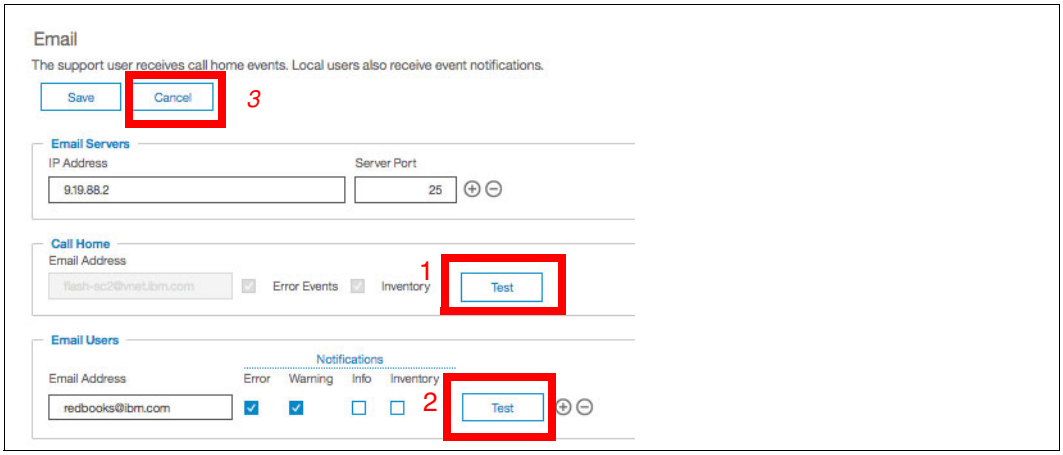


Figure 4-9 Testing both call home and email notification

Note: Testing Call Home exercises the support process by ensuring that your system entitles with IBM support. Contact information is also validated, because support reaches out to the primary contact.

SNMP

The SNMP panel is used to activate and download the MIB file to support SNMP, as shown in Figure 4-10. This is a post-install configuration usually set by the customer.

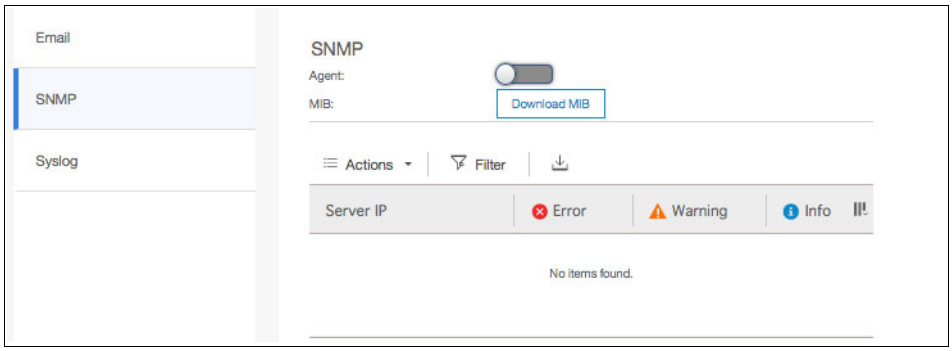


Figure 4-10 SNMP Setup Panel

Syslog

Syslog is a server used to receive log messages from systems. This is a post-install configuration performed by the customer, as shown in Figure 4-11.

Figure 4-11 Syslog Setup Panel

Network

Confirming the network configuration, shown in Figure 4-12, is important to ensure access to the system in the future.

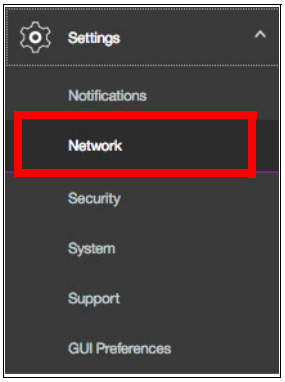


Figure 4-12

This section describes the Management and Service IP addresses, it is important to confirm these are the correctly configured.

Management IP

Your FlashSystem V9000 AE3 storage enclosures is managed using the management IP address. This panel, shown in Figure 4-13, can be used to change it in the future.

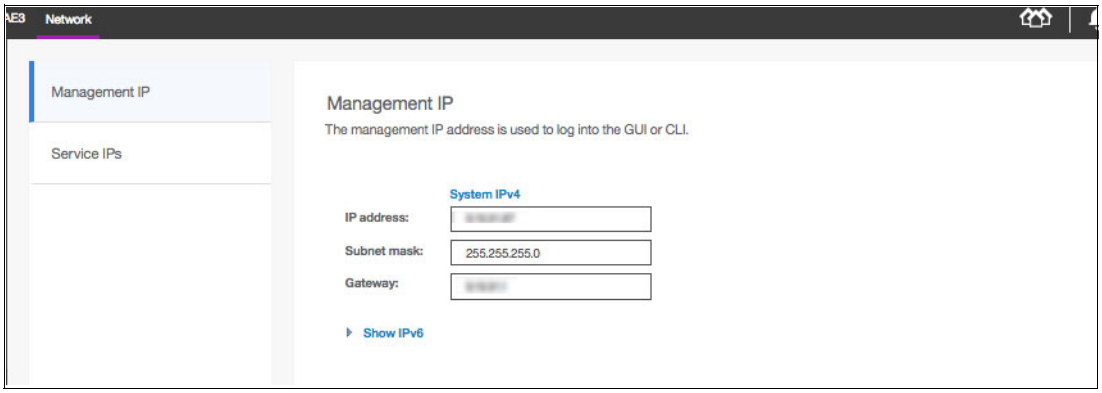


Figure 4-13 Management IP address panel

Note: It is important that you can reach the system on the Service IP addresses before the management IP address is changed.

Service IP

Setting the Service IP addresses is important to the serviceability of your system. IBM advises you to set these to accessible addresses on the network with your Management IP, as shown in Figure 4-14. Confirm that these are correct for your environment.

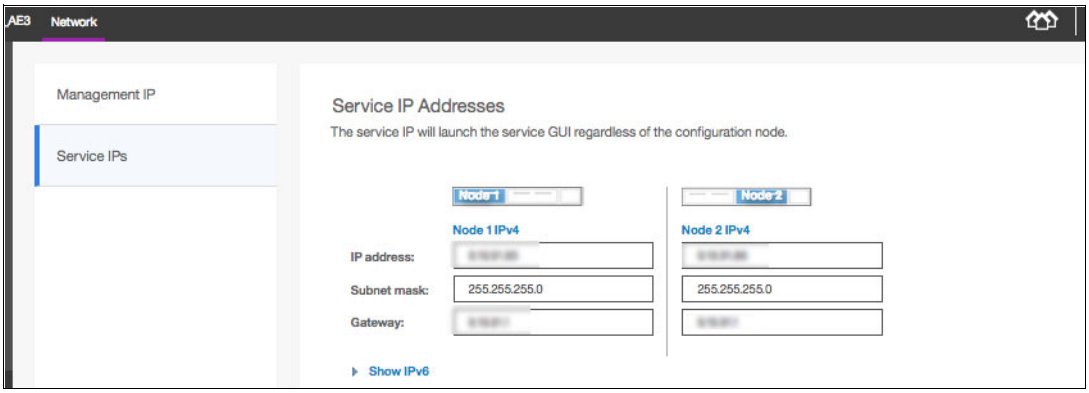


Figure 4-14 Confirm Service IP Addresses are configured

Note: The management IP address is the same as the configuration node. Either canister can be the configuration node.

Security

The security selections in Figure 4-15 are all post-install customer configurations. You can skip this for your initial installation, however the tabs are listed here with brief explanations.

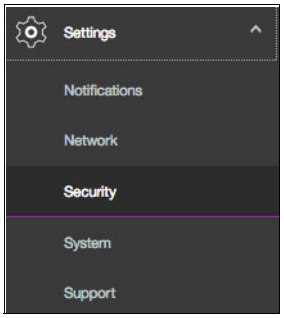


Figure 4-15 Security Pane Selection

Remote Authentication

Remote authentication is selecting the type of LDAP server to be used. This is normally configured by the customer after the install. Figure 4-16 shows the dialog that appears after the configure remote authentication option is selected.

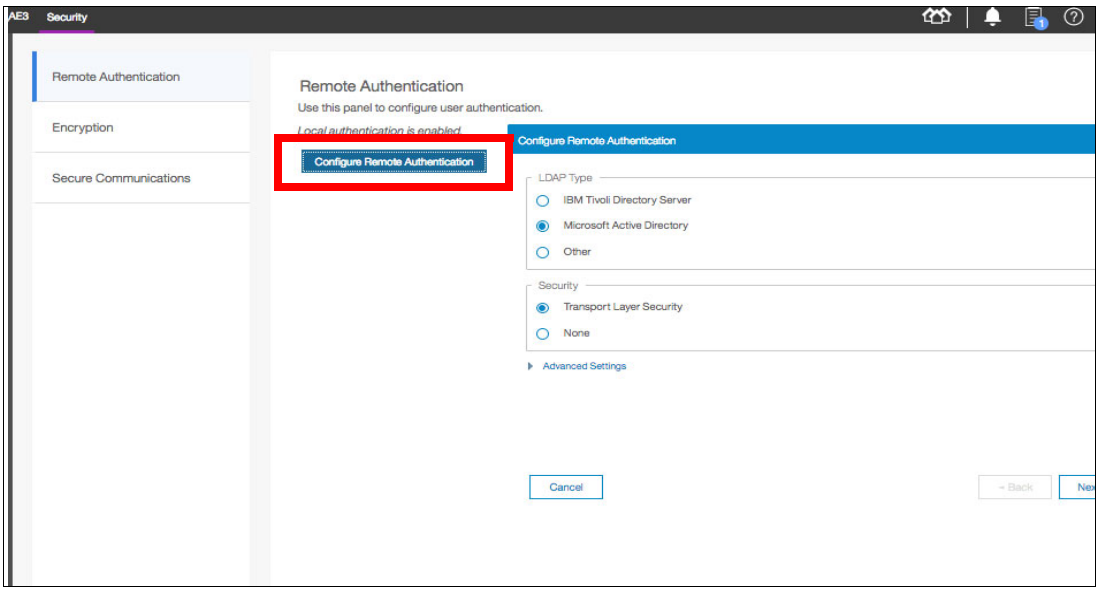


Figure 4-16 Configure Remote Authentication

Encryption

Encryption is available for customers that purchased a license, this is shown in Figure 4-17. Both USB-based and SKLM-based encryption are available with the FlashSystem V9000 AE3 storage enclosure.

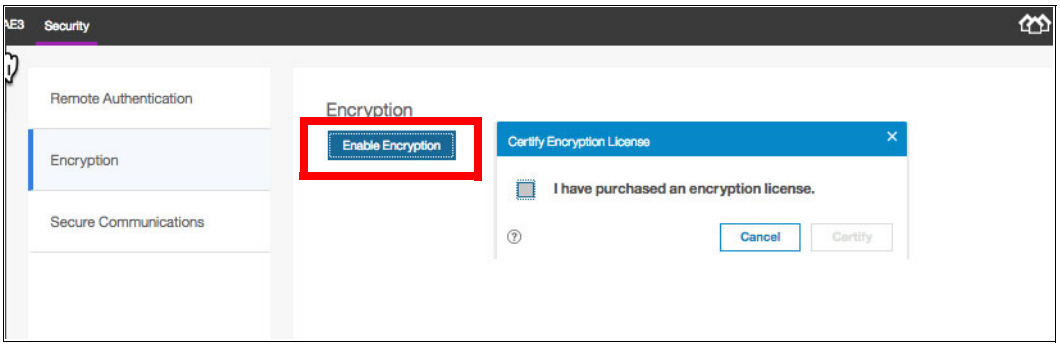


Figure 4-17 Enable Encryption Panel

Secure Communications

To ensure that the certificate and public ID of the system are valid and secure, each system has a digital Secure Sockets Layer (SSL) certificate. This is accomplished using setting in below in Figure 4-18. This is a customer configuration task usually done after the system is installed.

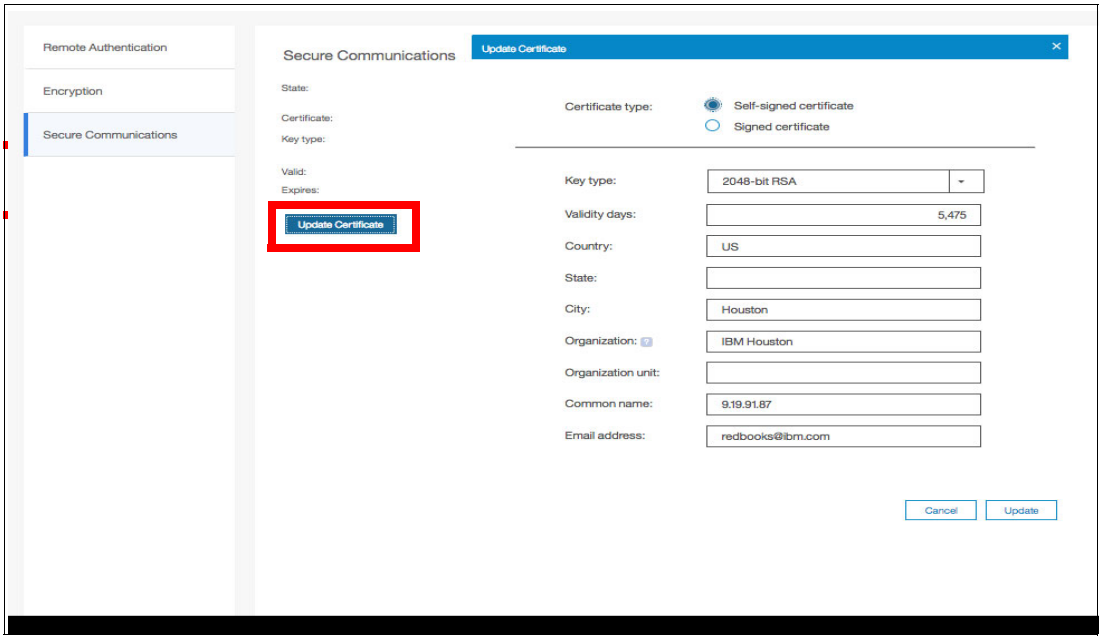


Figure 4-18 Updating SSL Certificates

System

The **System** menu in Figure 4-19 allows you to confirm the time zone and network Network Time Protocol (NTP) Settings.

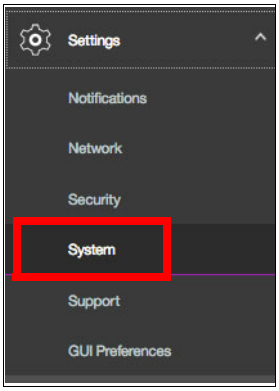


Figure 4-19 Select the System Menu

Date and Time

Figure 4-20 shows the window that enables you to verify the 1 **Time Zone** and 2 **Set Date and Time** or **Set NTP IP Server IP Address** options are set correctly.

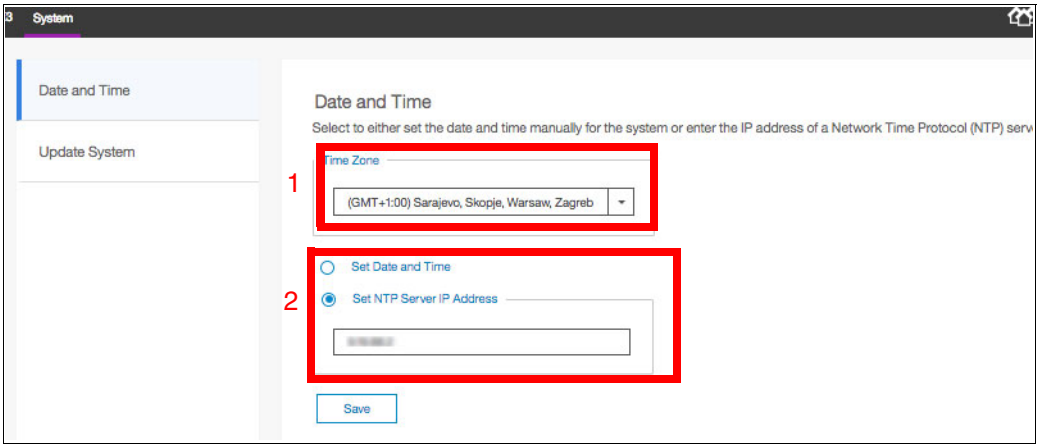


Figure 4-20 Ensure that the date and time are correctly set

Update System

Code currency is managed using the update system tab shown in Figure 4-21. If there is a later version indicated on this panel you should review the release notes on IBM Fix Central to plan an upgrade if appropriate.

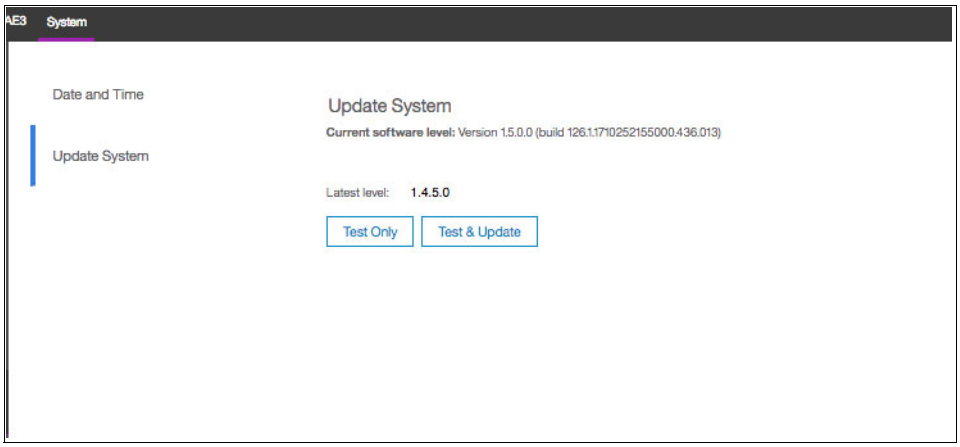


Figure 4-21 Confirm that you are at the current code release

Support

The Support Panel should be reviewed to ensure that **Support** assistance has been configured correctly for your companies policies, as shown in Figure 4-22.

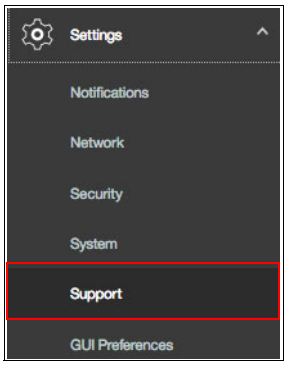


Figure 4-22 Support Menu Selection

Support Assistance window

To set up support, complete the following steps:

1. In Figure 4-23, you select 1 Setup Remote Assistance and the Setup Window 2 with the default settings displays.

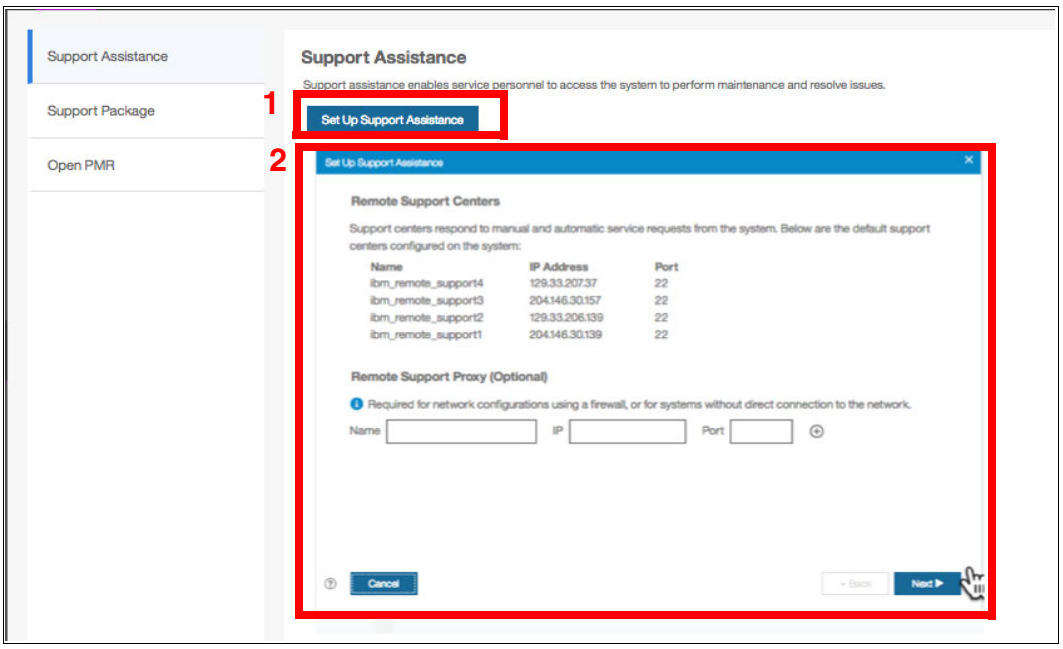


Figure 4-23 Set Up Support Assistance

> **Note:** Customer Firewalls must allow connection for SSH to IBM IP Addresses, as shown in Figure 4-23.

2. The Wizard continues in Figure 4-24 on page 101 allowing selection of how IBM will be allowed to work with this system. The most common selection is At Any Time. On System Error is another popular setting. These connections allow IBM timely access to collect logs. On Permission only requires the customer to open the connect when required.

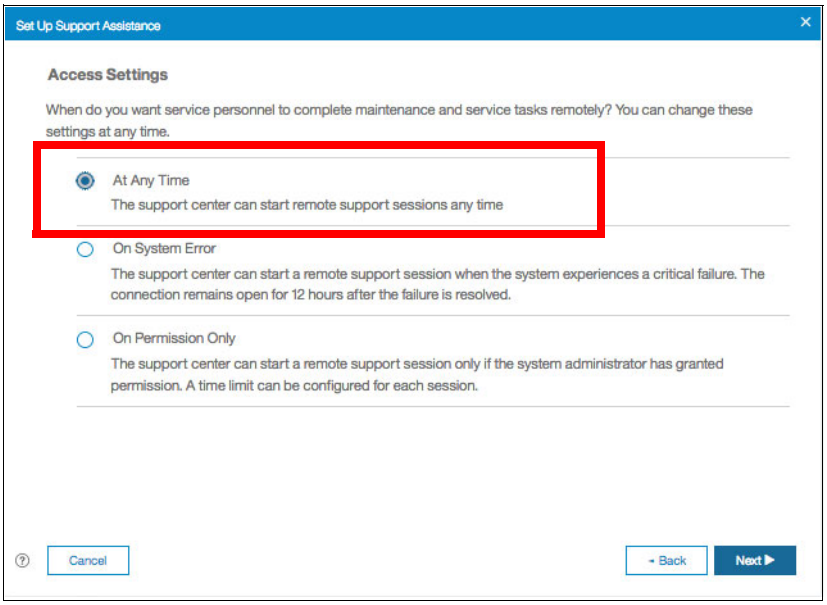


Figure 4-24 Access Settings

Note: IBM always reaches out to the primary contact for permission to collect logs; this is standard practice.

3. The wizard completes with the ability to set an access code, as shown in Figure 4-25. Selecting **Finish** completes the wizard without the access code.

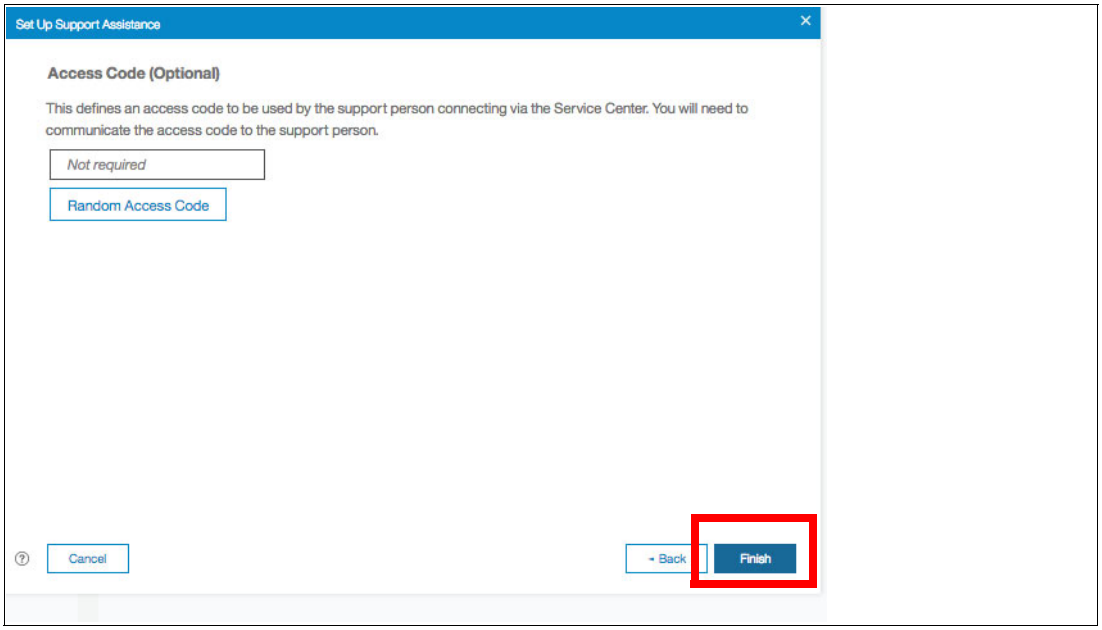


Figure 4-25 Optional Access Code can be Set

Note: Setting an access code requires that the primary contact for the system be available to provide the access code when required.

Support Package

Support package, shown in Figure 4-26, is used when the IBM Support Team needs the customer to collect logs. This is not part of the initial customer setup, and is normally done in response to a support ticket.

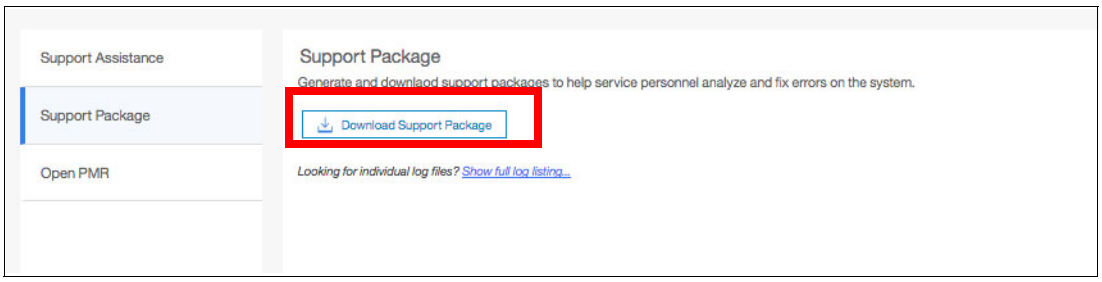


Figure 4-26 Generate a Support Package

Open PMR

The Open PMR capability is used to open a Problem Management Record (PMR) for the current system. Shown in Figure 4-27, a sample PMR is created to illustrate the capability.

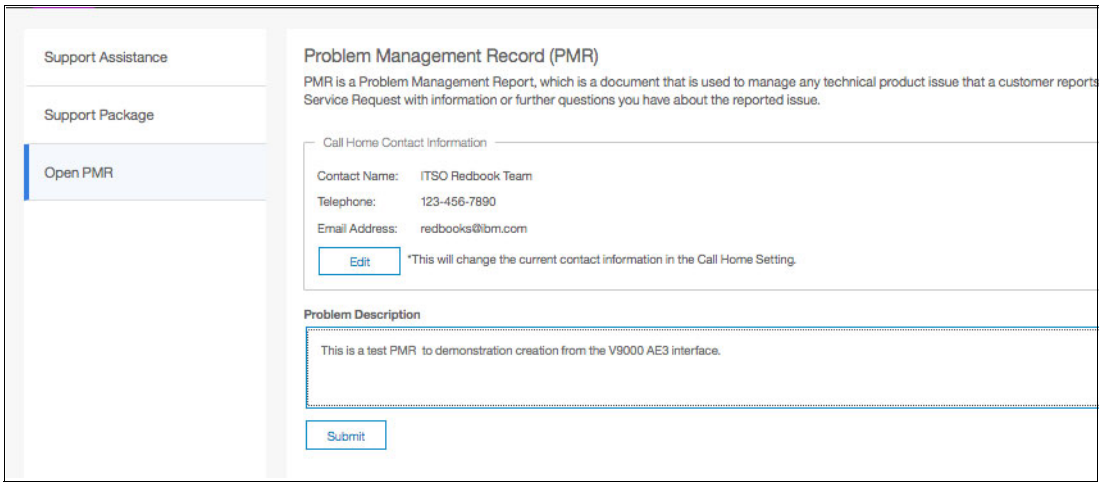


Figure 4-27 Opening a PMR from the AE3 Storage Enclosure GUI

GUI Preferences

GUI Preferences allows customizing of the GUI welcome screen and general characteristics of the GUI. The menu in Figure 4-28 brings up the panel.

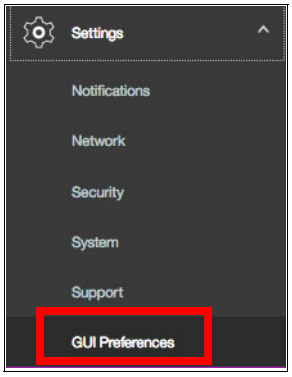


Figure 4-28 Select GUI Preferences

Login Message

Use the Login Message capability to customize the system welcome screen for the GUI and CLI, as shown in Figure 4-29.

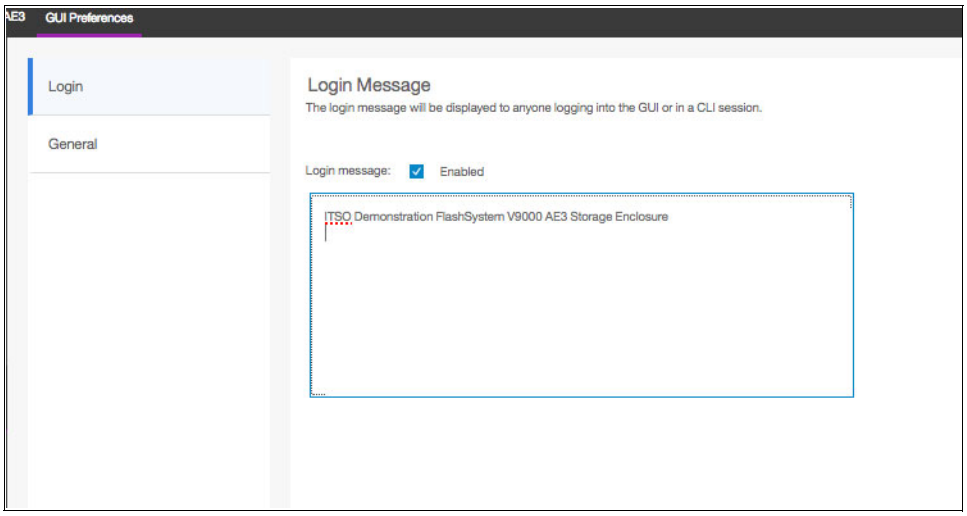


Figure 4-29 Customizing the Login Message

General

The General GUI setting scan be found in Figure 4-30. These settings enable you to set a time-out value for the GUI to lock out. Another customizing option is the access to IBM Knowledge Center documentation.

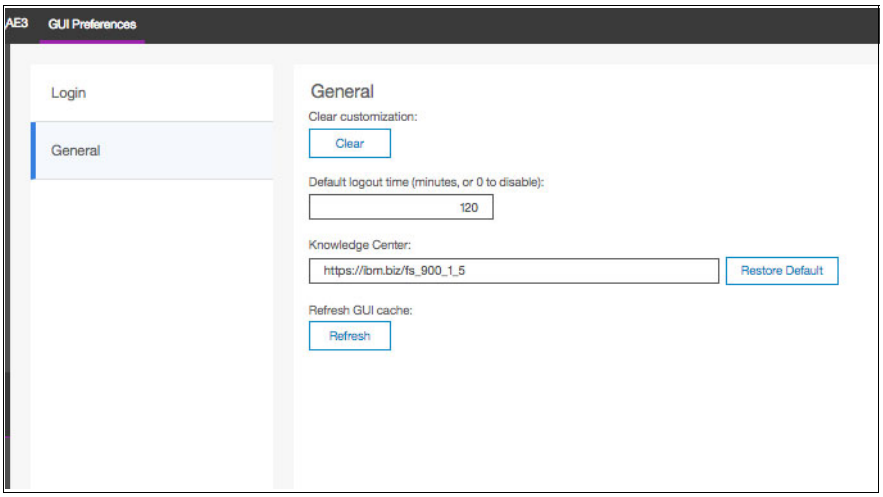


Figure 4-30 Customize GUI properties

This concludes verifying the FlashSystem V9000 AE3 storage enclosure settings. In the following section the AC3 control enclosure setup wizard is run.

4.5.2 FlashSystem V9000 AC3 Setup Wizard

License agreement and password change

This section explains how to start the customer setup wizard.

Tips: Next, you need the cluster management IP address (<*management_ip_address*>) that the IBM SSR entered for the AC3 Control Enclosure and AE3 Storage Enclosure

Complete the following steps:

1. Open a supported browser and go to the following address:

   `http://<AC3_management_ip_address>`

   IBM SSRs will provide login information when their portion of the installation is completed, as shown in Figure 4-31.

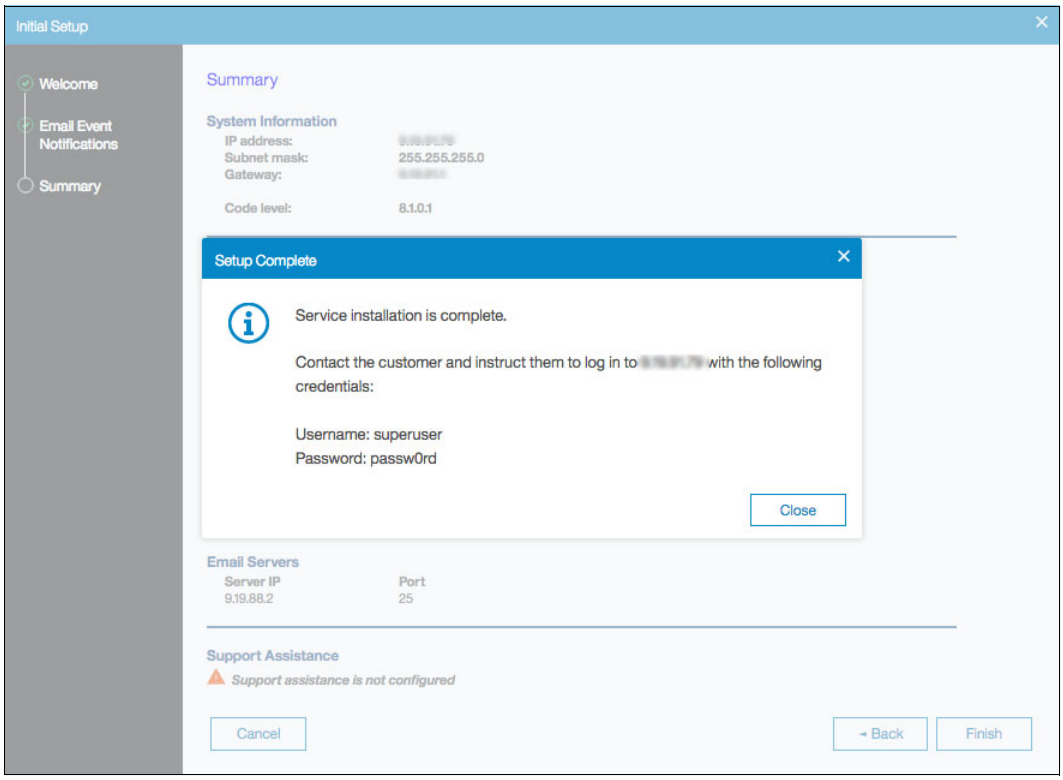


Figure 4-31 SSR customer information screen

Note: IBM FlashSystem V9000 uses a self-signed certificate, which in some browsers causes a warning message, such as `The certificate is only valid for xxxx`. Adding an exception will avoid this warning in the future.

2. To continue, select **Next**. On the V9000 Welcome screen:

   a. Read the IBM FlashSystem V9000 product license.
   b. Select **I agree with the terms in the license agreement**.
   c. Click **Next** (Figure 4-32).

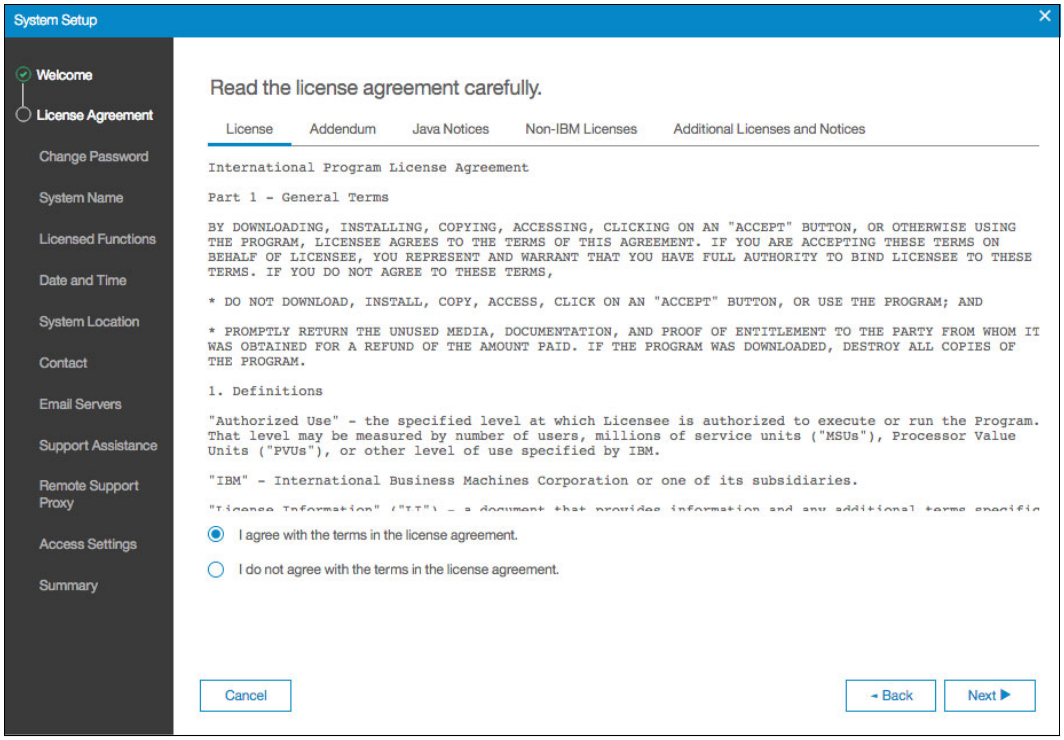


Figure 4-32 License agreement page

3. You can change the default password (`passw0rd` with a zero) in the next window (Figure 4-33).

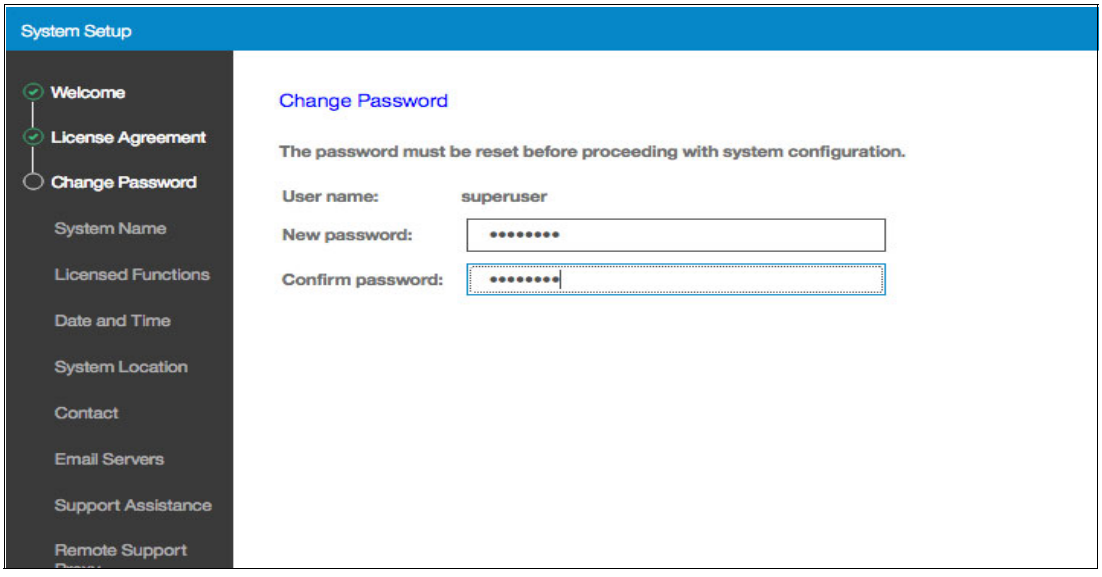


Figure 4-33 Superuser password change panel

Set System Name

Confirm that the system name set by the SSR is correct, as shown in Figure 4-34.

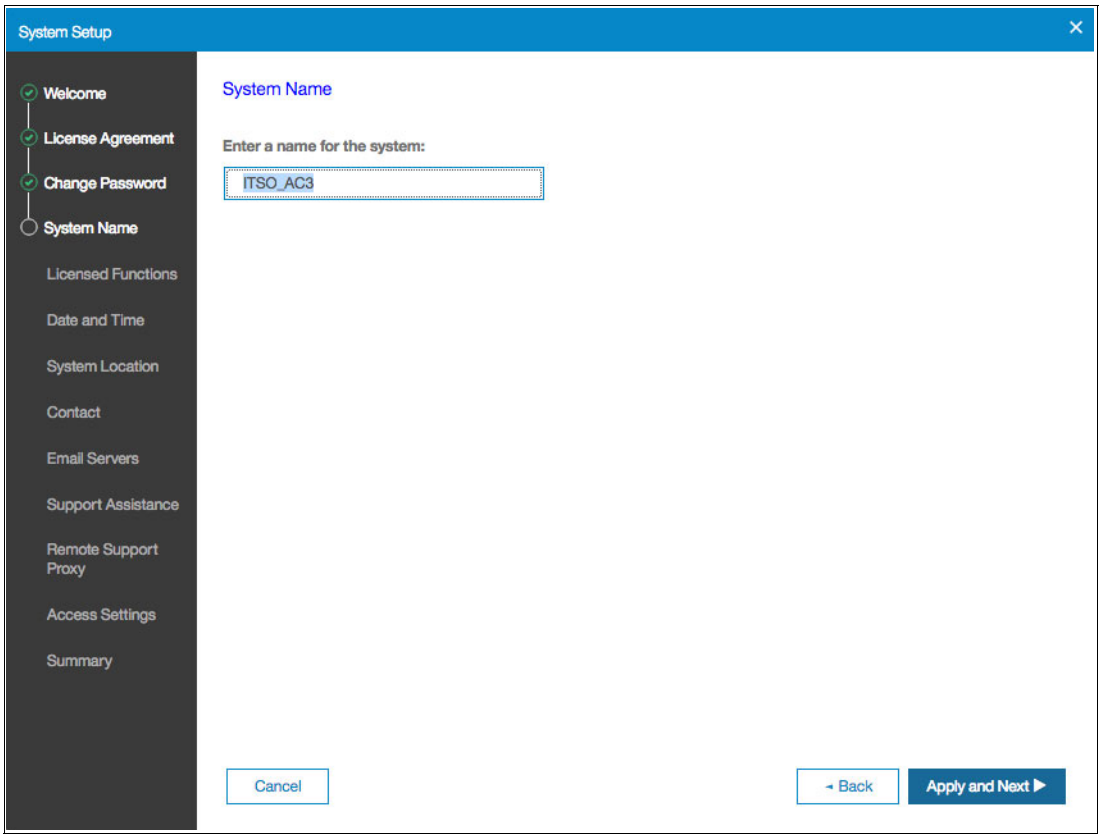


Figure 4-34 Set System Name

Setup Licensed Functions

This window enables you to set the values for licensed functions, as shown in Figure 4-35.

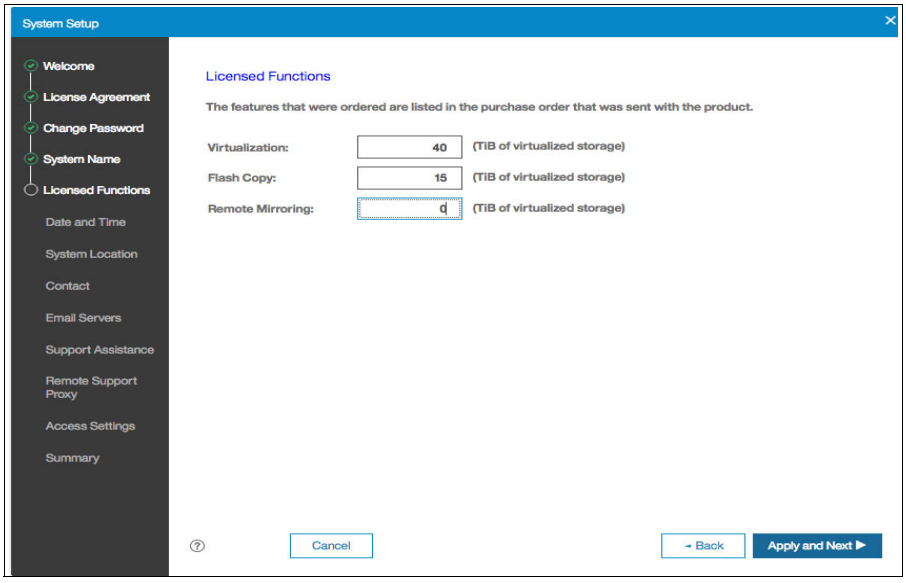


Figure 4-35 Set up Licensed Functions

Note: If you do not have the compression acceleration feature installed, the license value for encryption will be zero.

In Figure 4-36, the license limits entered previously are configured on the system.

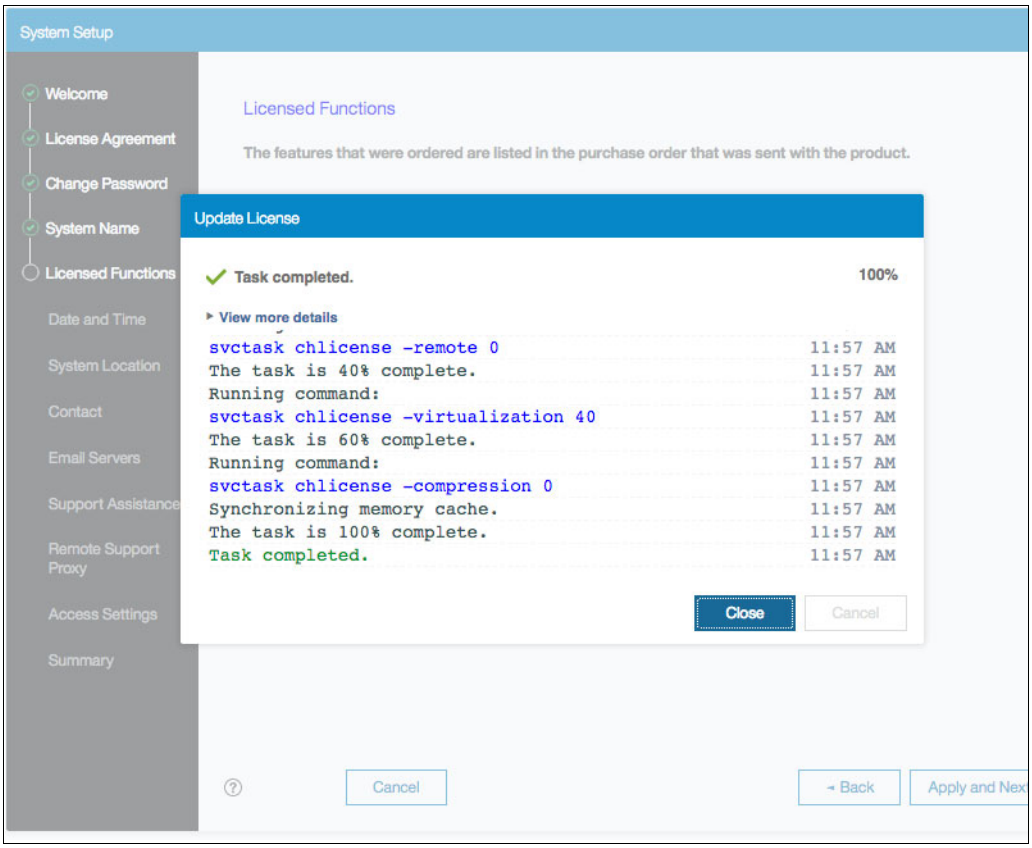


Figure 4-36 Licensed commands being executed

Date and Time

You might need the IP address of an NTP server (if the IBM technician has not yet entered it). In Figure 4-37, the screen for entering the NTP server is shown.

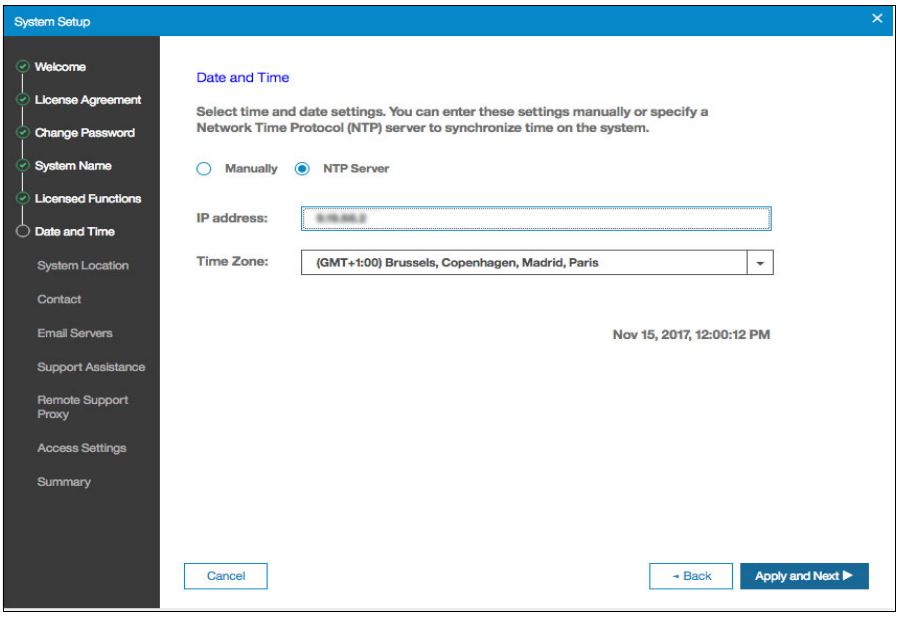


Figure 4-37 Setup the NTP Server and time zone

System Location

System location is used in conjunction with call home information to determine the physical location of the system when IBM Technicians are dispatched. In Figure 4-38, confirm that the System Location information is correct.

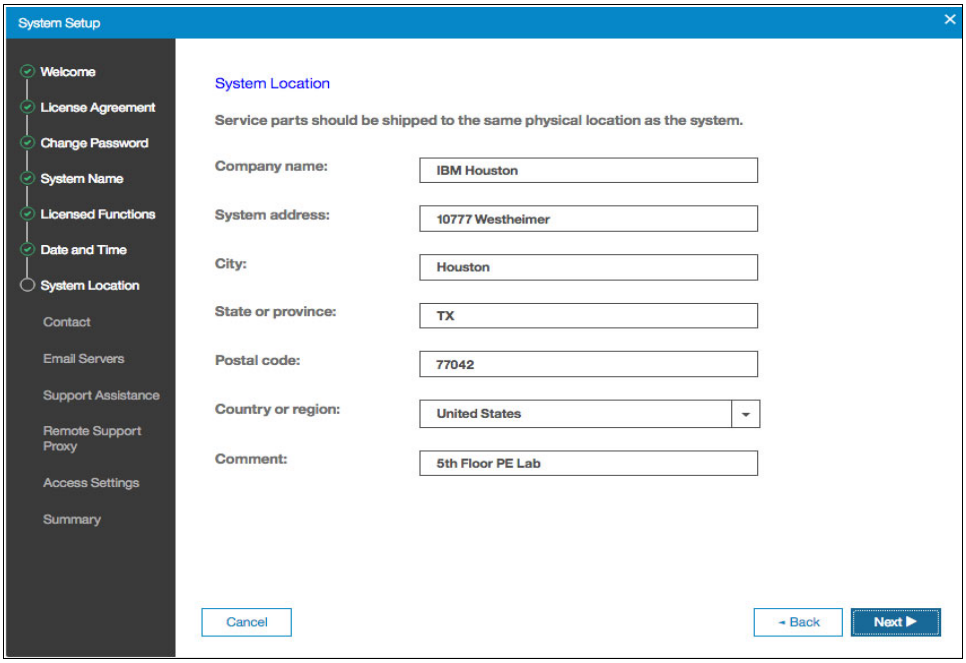


Figure 4-38 Confirm the system location information

Contact

Confirm the contact information is correct for your environment in Figure 4-39. When IBM receives call home messages, this information is part of the support ticket that is automatically generated. No other information is used by the support center to contact your support team, so be sure that the information on this form is accurate.

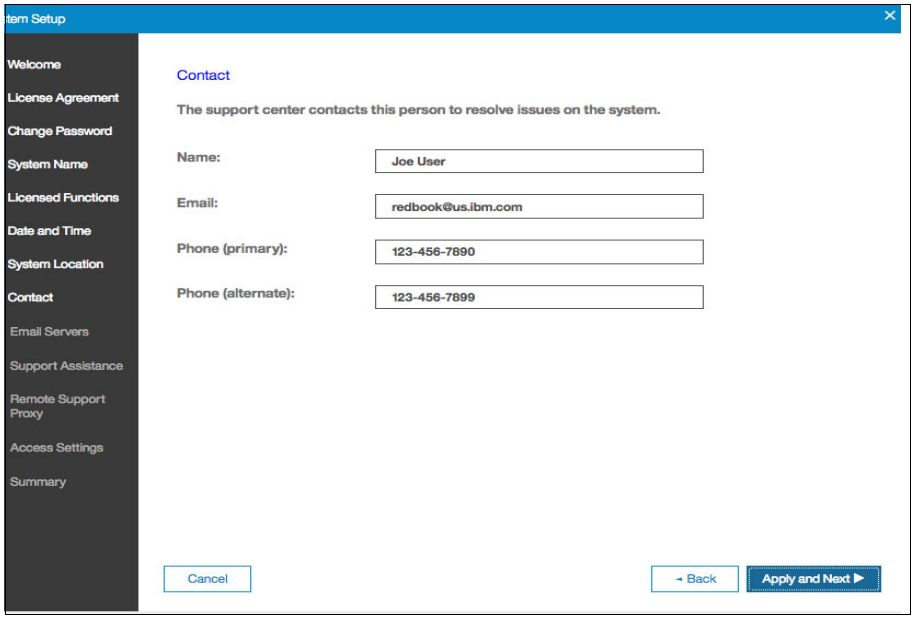


Figure 4-39 Enter accurate contact information

Email Servers

To setup the Email servers you will need IP address of the Simple Mail Transfer Protocol (SMTP) server as shown in Figure 4-40. Select the **Ping** button 1 to test the communication path to the SMTP Gateway.

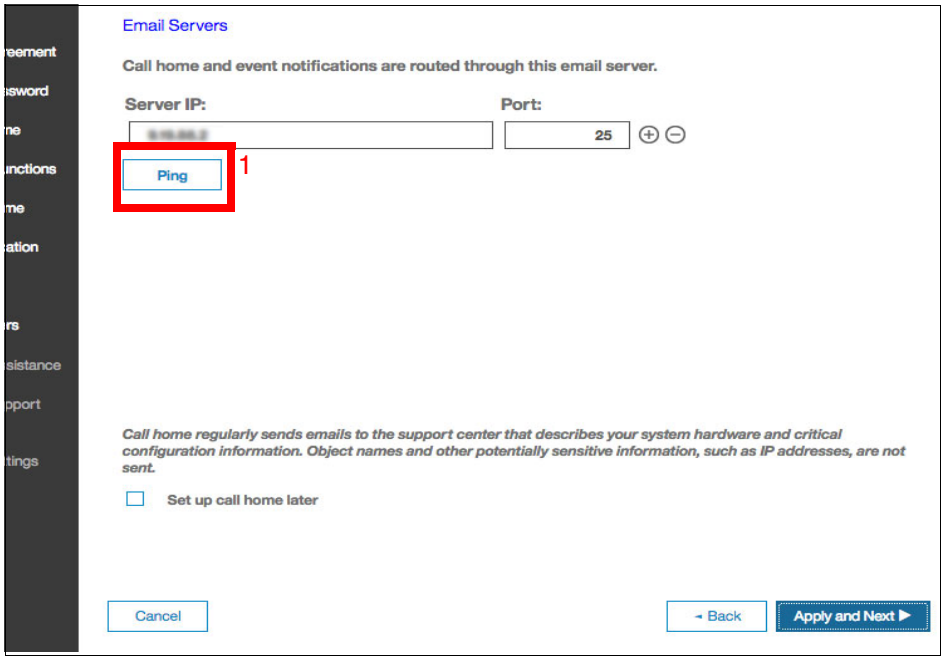


Figure 4-40 Enter your SMTP Gateway

Note: If the IBM SSR entered this information in the initial GUI, these values are displayed and you can verify them.

Support Assistance

Support Assistance enables support personnel to access the system to complete troubleshooting and maintenance tasks. You can configure either local support assistance, where support personnel visit your site to fix problems with the system, or remote support assistance. Both local and remote support assistance use secure connections to protect data exchange between the IBM support center and system. Shown in Figure 4-41 are the Support Assistance options.

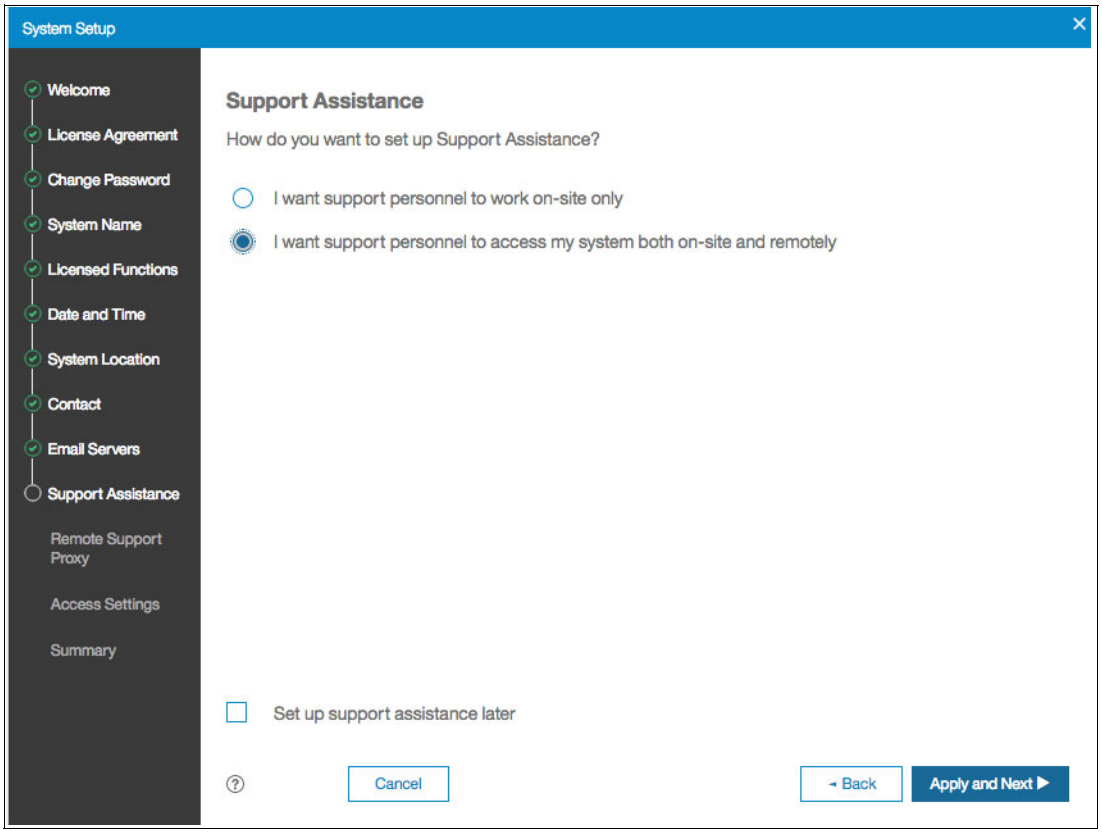


Figure 4-41 Support assistance setup panel

Remote Support Proxy

The Support Assistance panel is presented. Choose which option suits your environment best, as shown in Figure 4-42. The Remote Support Proxy works best for customers with multiple systems, this feature is also covered in the planning chapter 2.3.4, “Remote Support Assistance” on page 45.

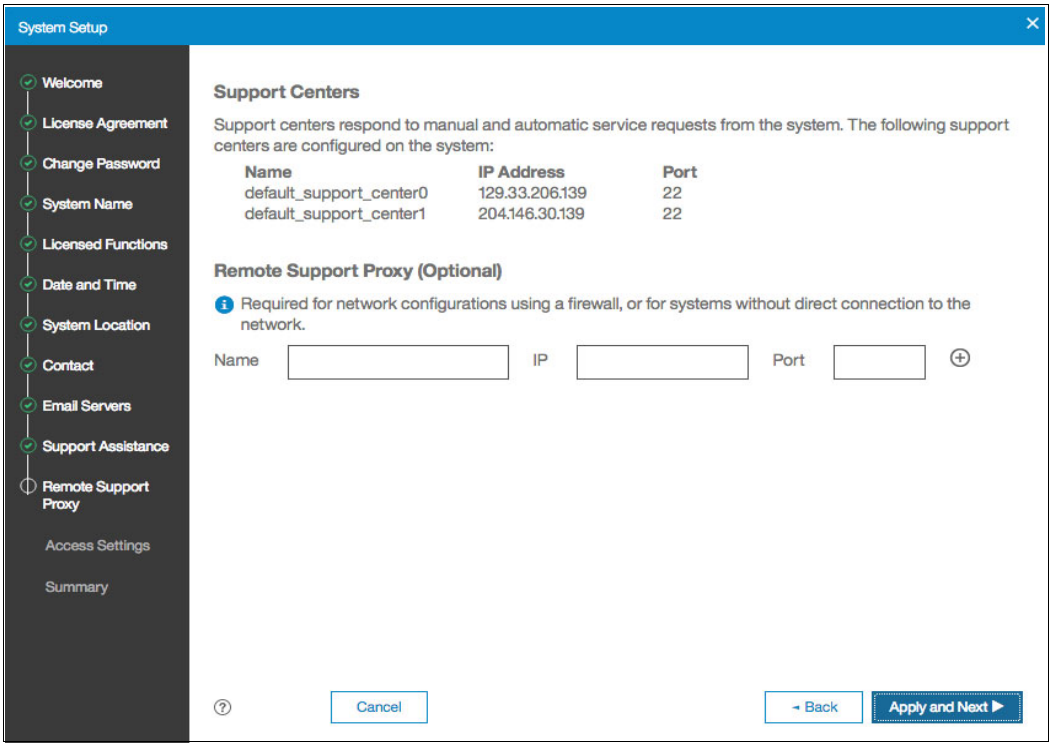


Figure 4-42 Confirm the Support assistant options

Access Settings

Next define how IBM support can access the system. **At Any Time** means the connection is available for IBM to collect logs as required. **On permission Only** means the customer will need to establish the connection before service takes place, as shown in Figure 4-43.

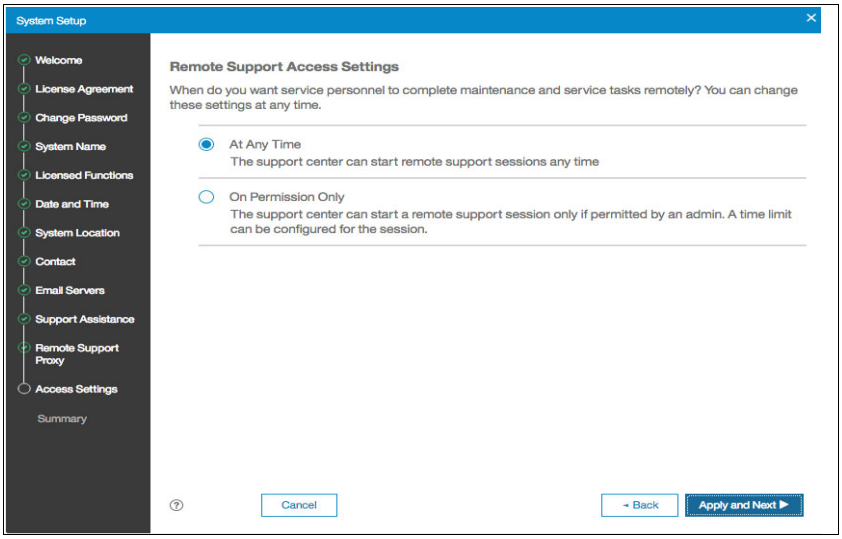


Figure 4-43 Confirm remote support access settings

In Figure 4-44, the results of the remote support configuration are displayed.

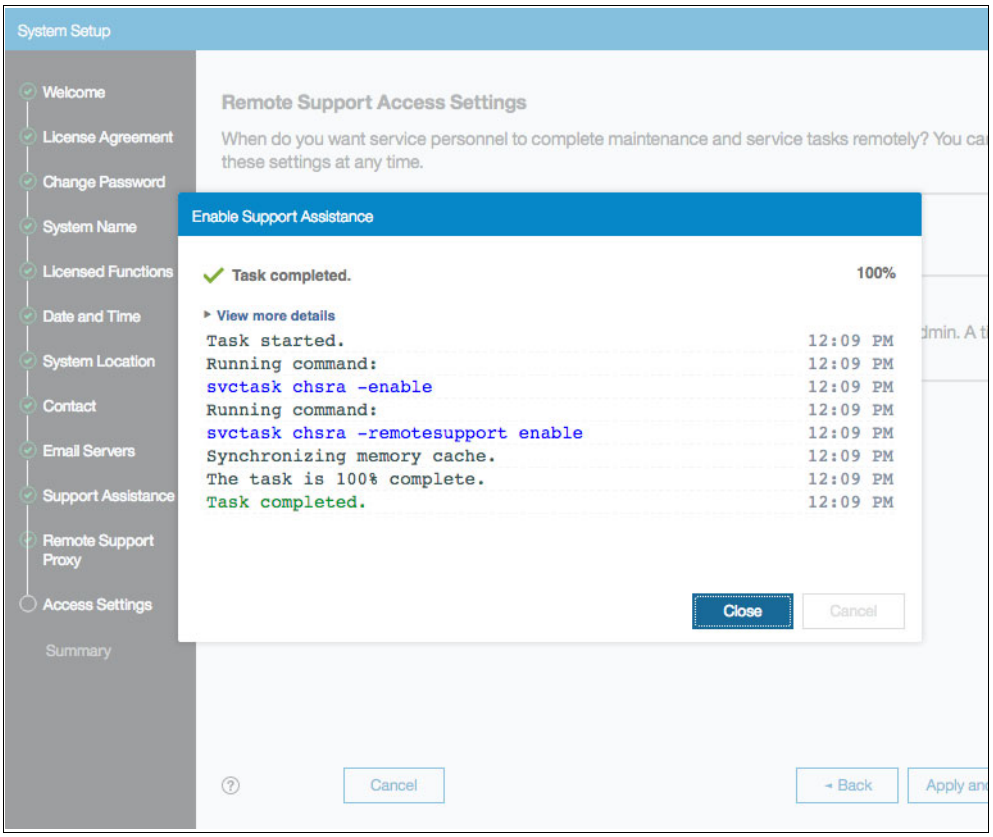


Figure 4-44 Details of the remote support assistance configuration are shown.

Summary

The summary display in Figure 4-45 shows the current configuration of the system.

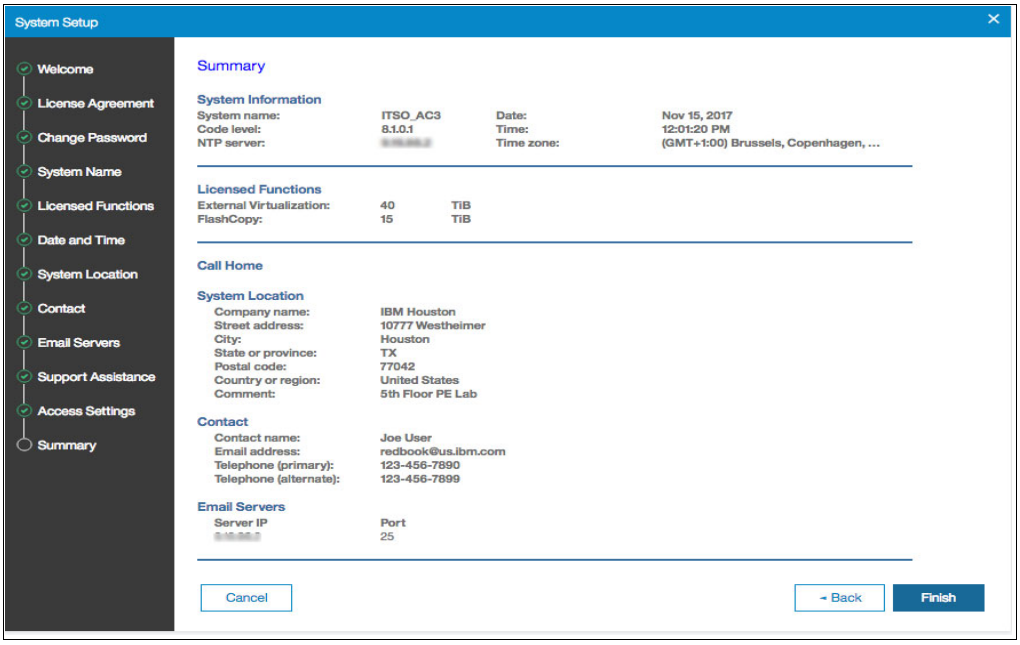


Figure 4-45 Summary display

The last display is confirming that the customer has completed the wizard as shown in Figure 4-46.

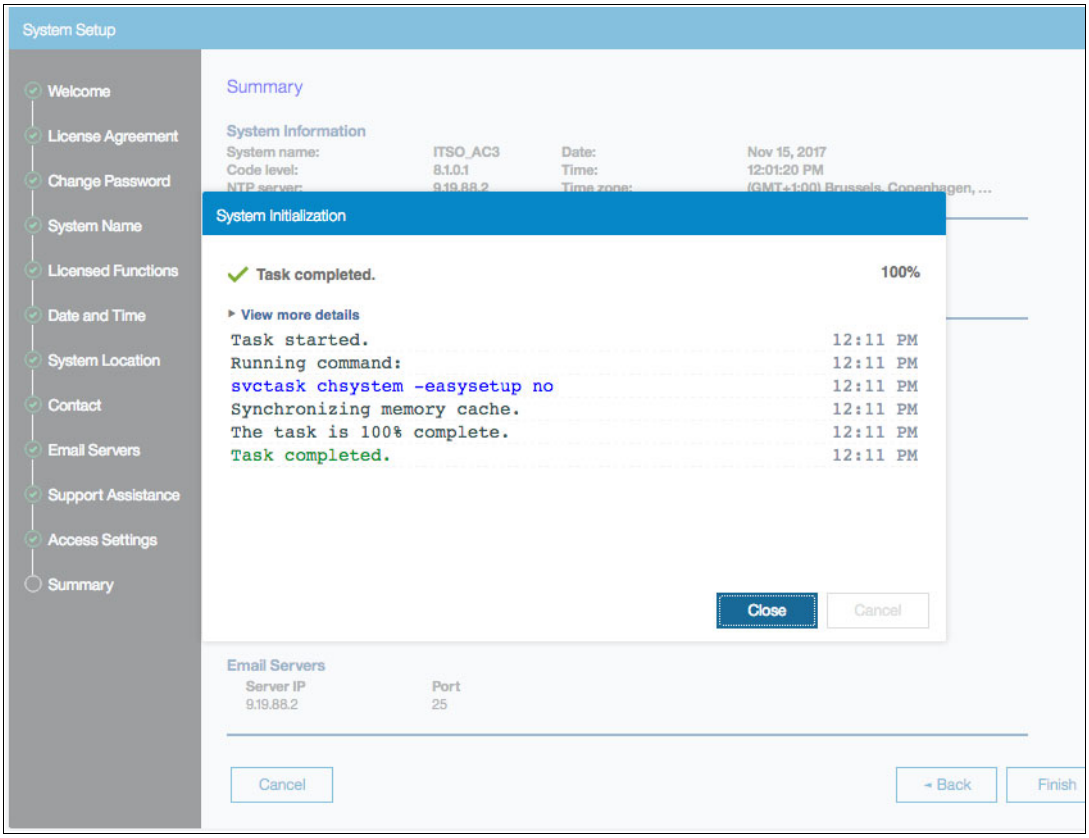


Figure 4-46 Completion of the easy setup wizard

4.5.3 Configure FlashSystem V9000 AC3 Controllers with AE3 Storage

The configuration of the AC3 controllers is completed by recognizing the storage virtualized by the AE3 storage enclosure. AE3 is recognized as external storage, previously the AE2 enclosure was treated as internal (managed) storage so the configuration process has changed. In the following sections we illustrate creating a storage pool, discovering the AE3 storage, and assigning the storage to the pool.

Creating the Storage Pool and adding AE3 storage

All storage resources that are under the system control are managed using storage pools. Storage pools aggregate internal and external capacity and provide the containers in which volumes can be created.

1. Begin by selecting the panel that you want to use to create the pool, as shown in Figure 4-47.

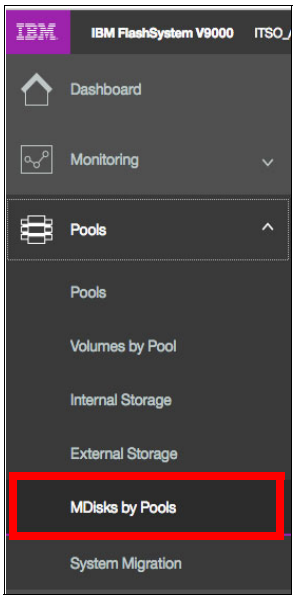


Figure 4-47 Select the MDisks by Pools from the Navigation Menu

2. The main display changes to the MDisk by Pool display similar to Figure 4-48. In this example the system already recognized the storage mapped as indicated by 1 and 2. Create the pool by selecting **Create Pool** 3.

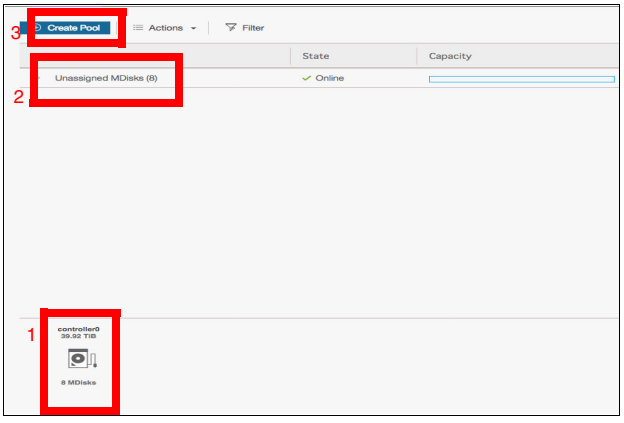


Figure 4-48 Create the Pool

3. Provide a meaningful name for the pool. In this case, we used `BB1_Pool 1` to represent building block 1 pool 1. The extent size defaults to 1 GB and leave this default. For more information about extents see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933. In Figure 4-49 the pool is created.

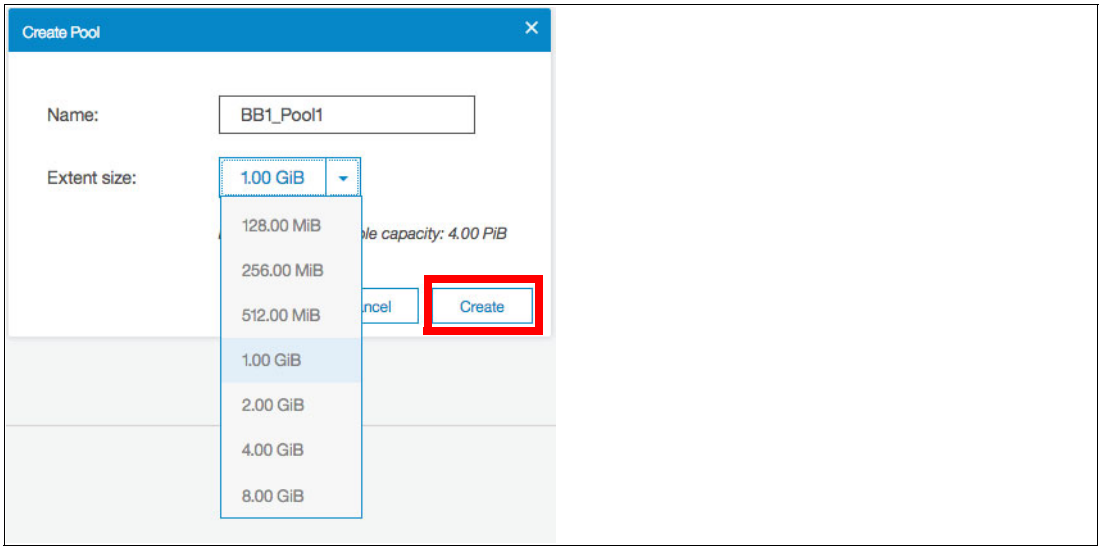


Figure 4-49 Create a Pool to manage the AE3 storage

Figure 4-50 shows the details of the command run to create the pool.

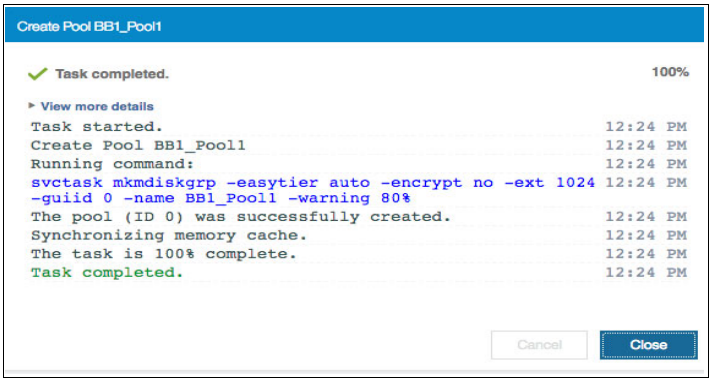


Figure 4-50 Command output from pool creation

4. On return to the MDisk by Pool display you may see this reminder in Figure 4-51 stating that the pool is empty, and needs to have storage added.

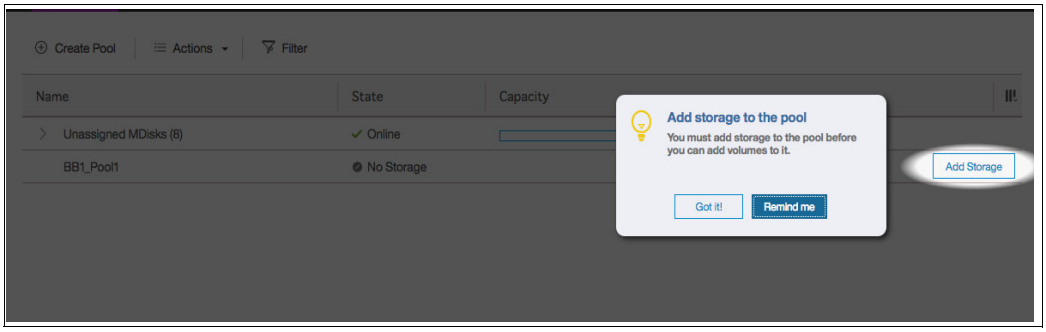


Figure 4-51 Reminder to add storage to the pool

5. In our example the AC3 control enclosure recognized the storage from the AE3, so we can simply add the storage to the pools, as shown in Figure 4-52.

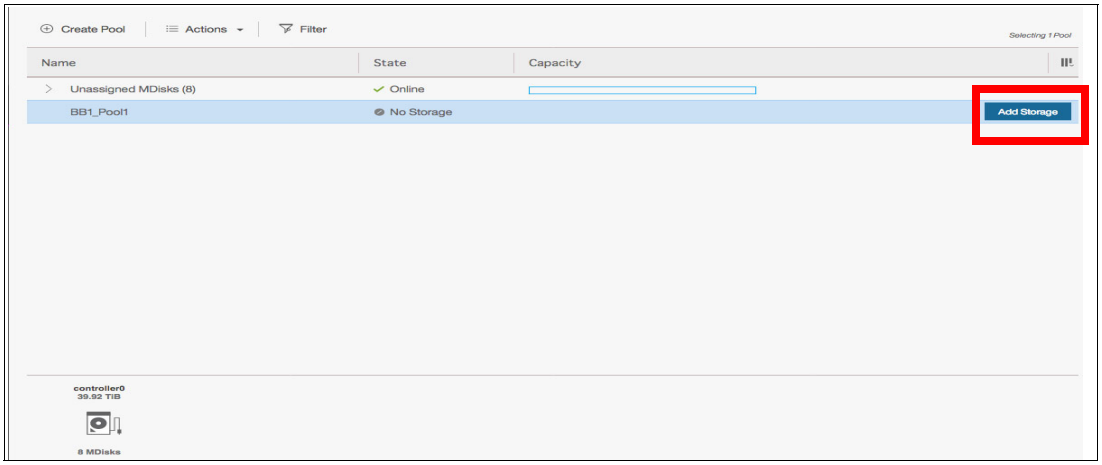


Figure 4-52 Select add storage

6. This will bring up the unused storage, as shown in Figure 4-53. Select **All MDisks** from the pull-down menu, which makes all of the MDisks part of the storage pool.

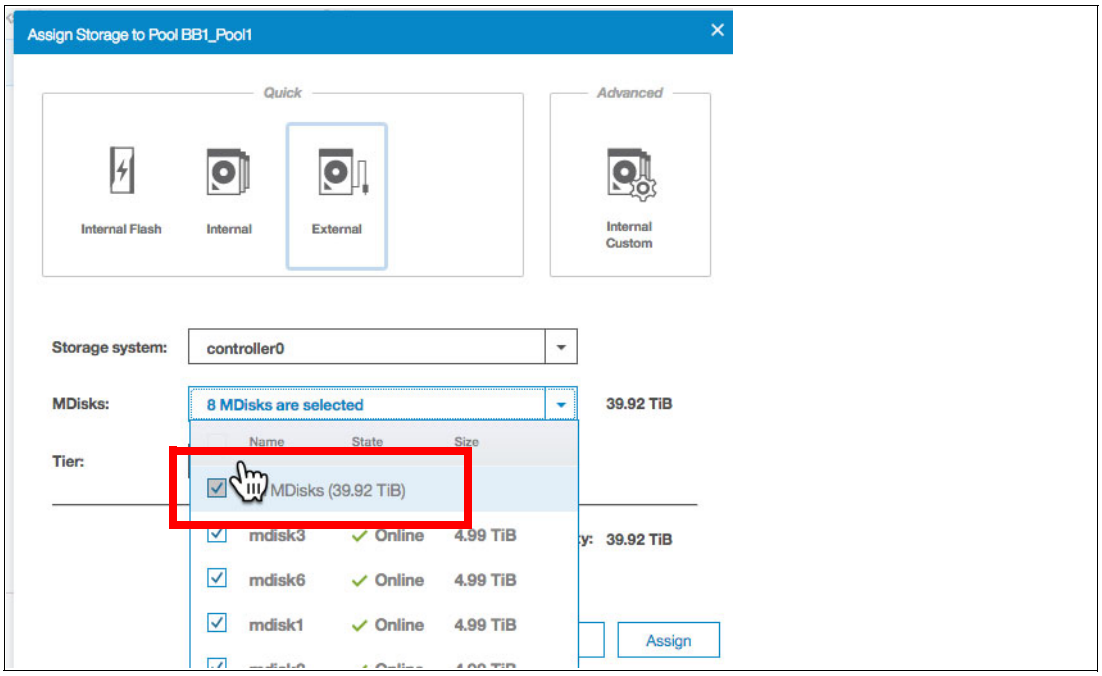


Figure 4-53 Select all MDisks

7. There is also the ability to classify the tier of storage being added to the pool. The AE3 is considered Tier 0 Flash storage. Ensure that **Tier 0 Flash** is selected, as shown in Figure 4-54. Select **Assign** to complete the task.

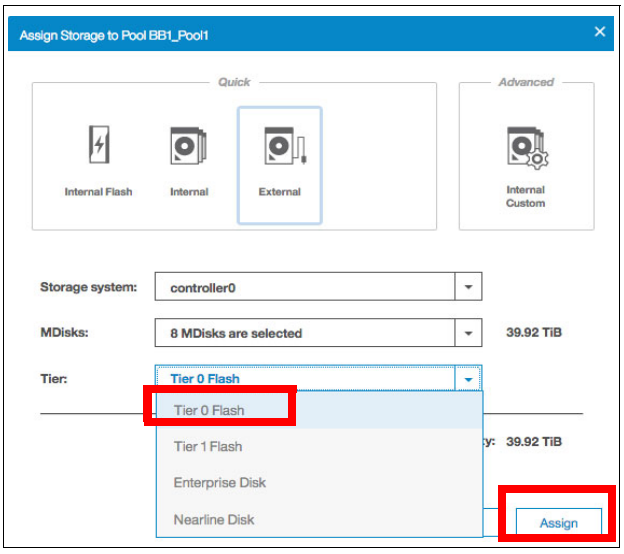


Figure 4-54 Select Tier 0 Flash and Assign storage to the pool

8. In Figure 4-55 the commands run are shown, including changing the storage tier and adding the MDisks to the pool. Click **Close**.

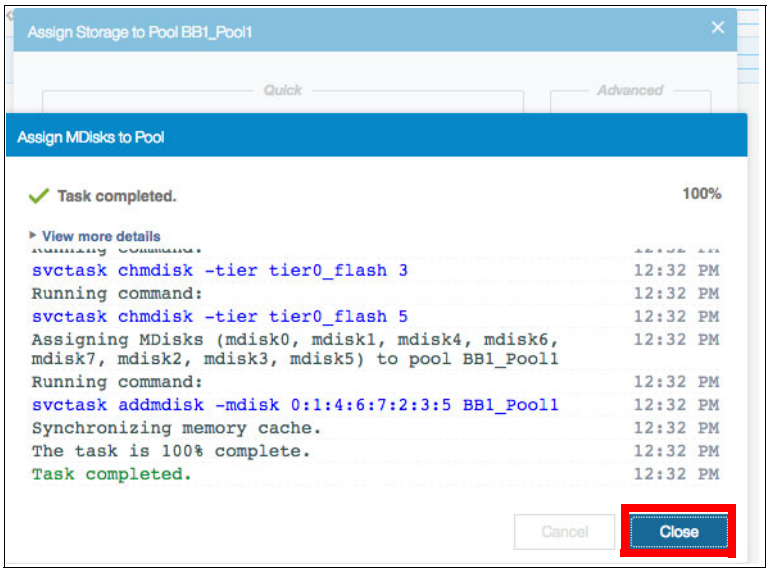


Figure 4-55 Command output from adding the AE3 storage to the pool

Returning to the MDisks by Pools display, shown in Figure 4-56, we can see the storage is available for allocation.

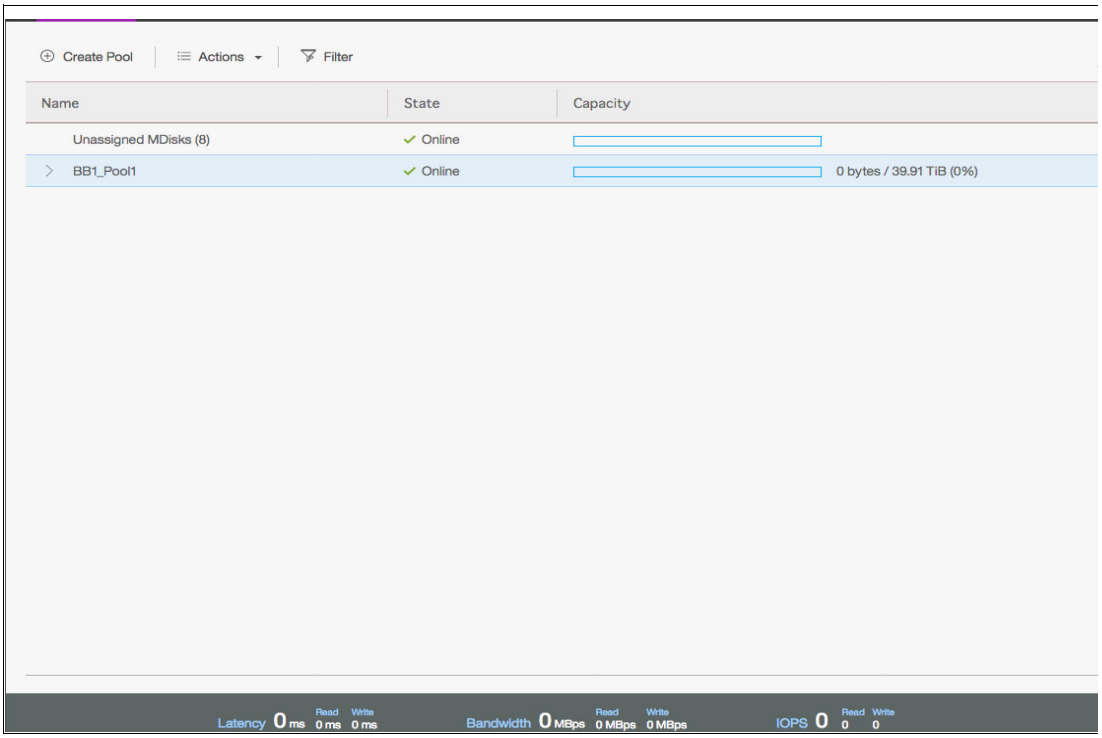


Figure 4-56 Pool is now available for allocation

The initial configuration of the system is complete, next steps are to allocate volumes from the extents in this storage pool. The IBM Redbooks publication, *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933, provides complete details about how this works.

Discovering the AE3 Storage

When you create or remove Logical Units (LU) on an external storage system, the change is not always automatically detected. If that is the case select Discover Storage for the system to rescan the Fibre Channel. The rescan process discovers any new MDisks that were added to the system and rebalances MDisk access across the available ports. It also detects any loss of availability of the controller ports.

Figure 4-57 shows the menu selection from the MDisks by Pools example in the previous section. Select **Actions** → **Discover Storage** to initiate a discovery for new LUs available from attached storage.

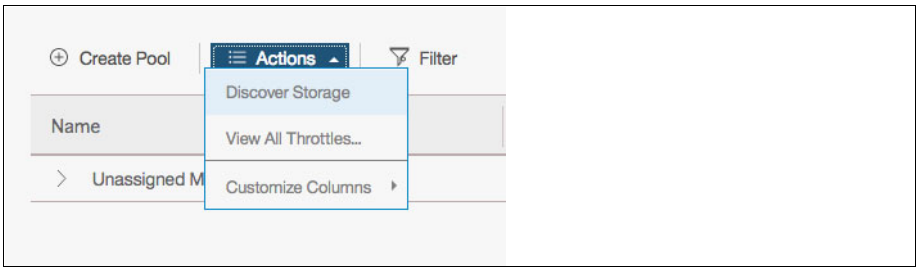


Figure 4-57 Discover storage option

Any unassigned storage is reflected on the panel after about a minute. Figure 4-58 shows the panel with unassigned storage present.

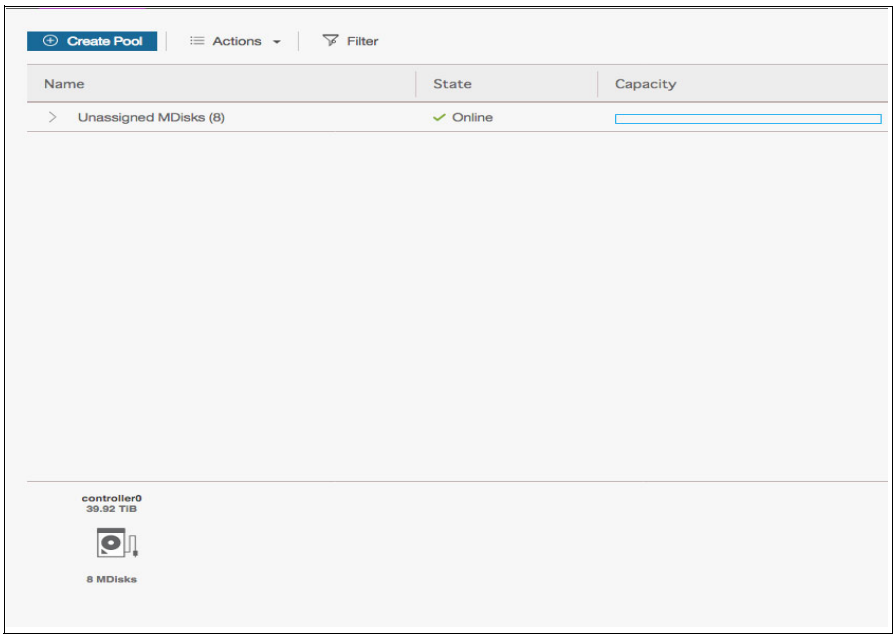


Figure 4-58 Unassigned Mdisks Displayed

When storage is visible, it can be assigned to the pool as previously described in 4.5.3, “Configure FlashSystem V9000 AC3 Controllers with AE3 Storage” on page 113.

5

Configuring settings

The Settings function covers various options for monitoring, configuring interfaces, and extracting support logs. It also covers remote authentication and the firmware update process. The Settings section of the IBM FlashSystem V9000 graphical user interface (GUI) of the FlashSystem V9000 AE3 control enclosure and FlashSystem V9000 AE3 storage enclosure are described in this chapter.

This chapter includes the following topics:

- Settings menu
- Network menu
- Security menu
- System menu
- Support menu
- GUI preference

5.1 Settings menu

The scope of this section is to provide an add-on description for the **Settings** menu in the FlashSystem V9000 GUI, in addition to the IBM Redbooks publications *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 and *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

Note: The control enclosure GUI and CLI are available via system management IP, which is hosted in the config node of the V9000 solution. Each V9000 FlashSystem AE3 storage enclosure will present its own GUI and CLI, that are reachable via their cluster IP. FlashSystem V9000 AE3 storage enclosures are shown in the V9000 GUI under **Pools** → **External Storage**.

5.1.1 FlashSystem AE3 storage enclosure settings

FlashSystem AE3 storage enclosure settings are documented in the Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414 in the “Settings menu” section. For other types of external storage, connected to the V9000, see the appropriate product documentation.

5.1.2 V9000 control enclosure and internal storage settings

V9000 control enclosure and internal storage settings are documented in the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933, “Settings” section.

A typical example for internal storage are FlashSystem AE2 storage enclosure and SAS expansion enclosures. For more details about SAS expansion enclosure, see 2.2.4, “SAS expansion enclosures”.

Note: In previous V9000 solutions, the FlashSystem AE2 storage enclosure are managed as internal storage. The AE2 remains configured as internal storage, even when the V9000 solution is upgraded to 8.1 firmware or further. FlashSystem AE3 components are added as external storage although they are part of the same product. The management coupling has been reduced in order to increase the flexibility of the overall solution.

However, there are exceptions to the previous rule.

5.1.3 Notifications menu

The system can use Simple Network Management Protocol (SNMP) traps, syslog messages, and call home emails to notify you and the support center when significant events are detected. Any combination of these notification methods can be used simultaneously. Notifications are normally sent immediately after an event is raised. However, there are some events that might occur because of active service actions. If a suggested service action is active, these events are notified only if they are still unfixed when the service action completes.

See IBM Redbooks *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 for examples.

5.2 Network menu

Use the Network panel to manage management IP addresses, service IP addresses, and supported network interfaces.

5.3 Security menu

Use the Security panel to configure and manage remote authentication and encryption settings on the system. With remote authentication services, an external authentication server can be used to authenticate users to system data and resources. User credentials are managed externally through various supported authentication services, such as LDAP.

5.3.1 Remote authentication with LDAP

For remote authentication using the Lightweight Directory Access Protocol (LDAP), see the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 chapters:

- “Configuring user authentication”
- “Configuring secure communications”

5.3.2 Encryption provided by the V9000 control enclosure

See the following IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 chapter:

- “Encryption”

5.3.3 Encryption provided by the FlashSystem AE3 storage enclosure

See the following IBM Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414 chapter:

- “Settings menu → Security menu → About encryption”

Note: Data encryption provided from the FlashSystem AE3 storage enclosure is the preferred method in a typical configuration. FlashSystem AE3 provide the benefit of hardware compression and encryption inside the flash module. If encryption is configured on the V9000 control enclosure, make sure to enable encryption on FlashSystem AE3 first. FlashSystem V9000 will not encrypt data in the control enclosure, if it is encrypted on FlashSystem 900 AE3. Otherwise by the nature of encryption, it will end up in incompressible data stored at the FlashSystem AE3 storage enclosure.

5.4 System menu

Use System pages to manage overall system configuration options, such as licenses, updates, and date and time settings.

5.4.1 V9000 control enclosure and internal storage firmware update

See the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 chapter:

- "RAS, monitoring, and troubleshooting → Software update"

Note: The firmware update example in the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 shows an update from 7.6. to 8.1. The FlashSystem V9000 firmware update process from 8.1 firmware to higher levels is similar.

5.4.2 FlashSystem 900 AE3 storage firmware update

To update the firmware on the FlashSystem 900 AE3, connect to the individual storage enclosure GUI. See the IBM Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414, chapter "Settings menu → System menu → Update software".

5.4.3 Some rules of thumb apply for V9000 firmware updates

First update the firmware from the management IP of the V9000 control enclosure. This will update all control enclosure through a single GUI interface, including the internal storage. Complete this update before any other action.

Next, log in to the FlashSystem 900 AE3 storage enclosure GUI and start firmware update for each FlashSystem 900 AE3, one at a time. It is a good practice to complete one storage enclosure before starting next enclosure. This will extend firmware update time, but also add extra protection in case of an unexpected error during firmware upgrade.

Always read the appropriate firmware release letter for instructions about how to update V9000 components, and in which sequence the updates need to be done.

Note: For firmware compatibility planning, see the IBM System Storage Interoperability Center (SSIC).

5.5 Support menu

Use the V9000 Support panel GUI to configure and manage connections and upload support packages to the support center. From the FlashSystem 900 AE3 storage enclosure GUI, a manual PMR creation is supported.

5.5.1 Remote Support Assistance

See the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933, chapter "RAS, monitoring, and troubleshooting → Monitoring".

Note: Remote Support Assistance is documented in this IBM Redpaper™ publication in 2.3.4, "Remote Support Assistance" on page 45.

5.5.2 Support package, V9000 control enclosure support logs

See the IBM Redbooks publication *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933, chapter “Collecting support information”.

5.5.3 Support package, FlashSystem V9000 AE3 storage enclosure support logs

The support functions of the FlashSystem V9000 storage enclosure are described in the “Settings → Support Menu” chapter, in the IBM Redbooks publication, *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414.

Note: For a performance-related problem, see the *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 “Performance data and statistics gathering” appendix. For a V9000 volume that is created from an external storage MDisk, also gather support data for that external storage.

5.5.4 PMR creation, FlashSystem V9000 AE3 storage enclosure GUI

A PMR (problem management record) is a document that is used to manage a technical product issue that is reported to IBM. Manual PMR creation is available only from the FlashSystem V9000 AE3 storage enclosure GUI. This feature is useful when an automated PMR does not already exist for a V9000 component.

See the IBM Redbooks publication *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414, “Support Menu → Open PMR” chapter.

5.6 GUI preference

Use the GUI preference page to change settings that are related to how information is displayed on the GUI.

6

IBM HyperSwap

The IBM HyperSwap function is a high availability feature that provides dual-site, active-active access to a volume. This chapter walks you through the process of setting up, configuring and using HyperSwap.

HyperSwap function was introduced with IBM FlashSystem V9000 software V7.5 and it was only available through command-line interface (CLI) commands. IBM FlashSystem V9000 software V7.6 introduced graphical user interface (GUI) support to configure the HyperSwap function and additional CLI commands to configure HyperSwap volumes.

This chapter describes the GUI and CLI commands for FlashSystem V9000 software V8.1. The FlashSystem V9000 software V7.5 commands are still valid and are described in 7.3, “Command-line hints” on page 205.

The HyperSwap function enables each volume to be presented by two I/O groups at two different sites. At the time of publishing, the two I/O groups in a HyperSwap configuration must exist within a single FlashSystem V9000 cluster (in so called SUSO (scale-up, scale-out) configuration). The configuration tolerates combinations of control enclosure, storage enclosures, and site failures, using a flexible choice of host multipathing driver interoperability. This chapter includes the following topics:

- Overview
- HyperSwap design
- Comparison with Enhanced Stretched Cluster
- Planning
- Configuration
- Operations
- HyperSwap with SAS attached expansion enclosures
- Disaster recovery with HyperSwap
- Disaster recovery with consistency groups
- The overridequorum command
- HyperSwap Failure scenarios
- Unconfiguring HyperSwap
- Summary of interesting object states for HyperSwap
- Naming conventions

In this chapter, the term *VDisk* is used for an individual object created with the **mkvdisk** command. The term *HyperSwap volume* is used for the volume with copies on two sites in a HyperSwap relation. The term *basic volume* is used for a volume which is only on one site. For details about these terms, see 6.14, “Naming conventions” on page 184.

Note: At the time of publishing, the two I/O groups in a HyperSwap configuration must exist within a single FlashSystem V9000 cluster. The cluster consists of at least two scalable V9000 building blocks.

For better readability of command-line interface (CLI) examples, the CLI output is shortened in lines, columns, or both in the examples in this chapter.

6.1 Overview

HyperSwap is the high availability (HA) solution for IBM FlashSystem V9000. HyperSwap provides business continuity if hardware failure, power failure, connectivity failure, or disasters occur. HyperSwap is also available on other IBM Spectrum Virtualize products, such as IBM SAN Volume Controller, or IBM Storwize V5000/V7000.

The following list includes general HA requirements:

- Two independent main sites
- Independent infrastructure for power, fire protection, and so on
- Independent servers on each site
- Two independent data copies, one in each site
- Latency optimized intersite traffic to keep both sites’ data copies in sync
- Local high availability in each site
- Application site transparency

Figure 6-1 shows a two-site HA environment.

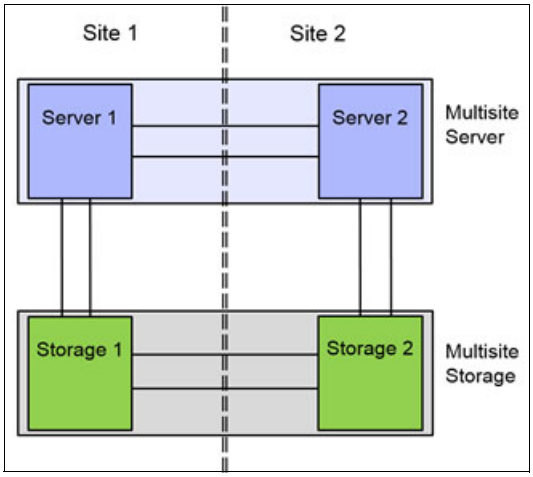


Figure 6-1 High availability environment

The HyperSwap function provides highly available volumes accessible through two sites at a distance up to 300 km apart. A fully-independent copy of the data is maintained at each site. When data is written by hosts at either site, both copies are synchronously updated before the write operation is completed. The HyperSwap function automatically optimizes itself to minimize data transmitted between sites and to minimize host read and write latency.

If the control enclosures, storage enclosures, or any other attached storage at either site go offline, leaving an online and accessible up-to-date copy, the HyperSwap function automatically fails over access to the online copy. The HyperSwap function also automatically resynchronizes the two copies when possible.

6.1.1 HyperSwap implementations

The decision for a HyperSwap failover can be managed by the host or by the storage system. IBM currently has two main solutions:

- Host-based HyperSwap. The host handles storage failures.
- Storage-based HyperSwap. The storage system handles storage failures.

The next sections describe these two solutions.

Host-based HyperSwap

A HyperSwap function is available when using the IBM DS8000 family of products together with IBM PowerHA® System Mirror for AIX or IBM Geographically Dispersed Parallel Sysplex™ (IBM GDPS®) for IBM z/OS®. The HyperSwap functions on those environments use specific software on that host system. All decisions in split scenarios are made by the host.

Figure 6-2 shows a host-based HyperSwap example of an IBM AIX PowerHA and IBM System Storage DS8000 HyperSwap setup.

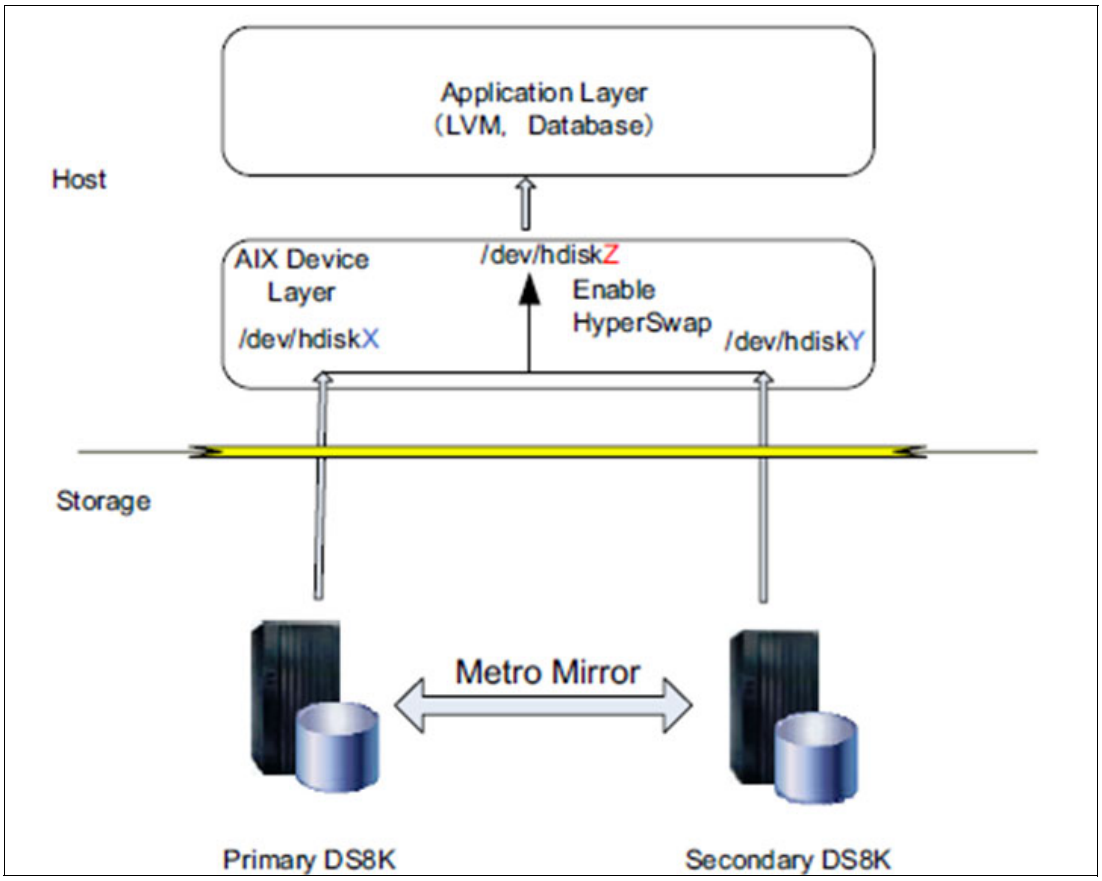


Figure 6-2 Host-based HyperSwap example

Storage-based HyperSwap

IBM Spectrum Virtualize provides the HyperSwap feature in the virtualization layer. It uses technologies from:

- Metro Mirror
- Global Mirror with Change Volumes
- FlashCopy
- Non-disruptive Volume Move

One volume is presented to the host from two different sites. Two I/O groups are presenting the same volume to the host. All decisions in split scenarios are made by IBM Spectrum Virtualize software running on IBM FlashSystem V9000.

The host must detect, accept, and handle HyperSwap changes, and manage the path failover. All FlashSystem V9000 failover decisions are valid for all hosts, or host clusters attached to the FlashSystem V9000 cluster.

Figure 6-3 shows a IBM Spectrum Virtualize-based HyperSwap example. It shows that four VDisks are needed to present one HyperSwap volume to the host.

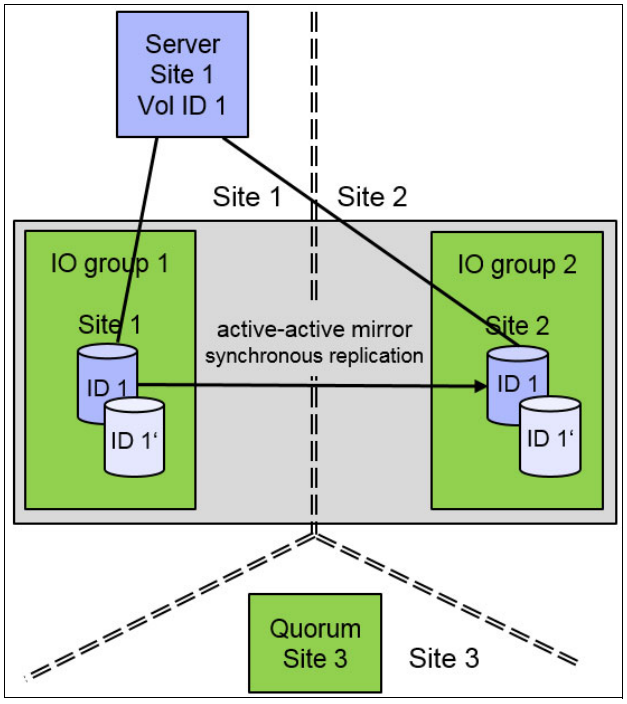


Figure 6-3 IBM Spectrum Virtualize-based HyperSwap example

The HyperSwap function in the FlashSystem V9000 software works with the standard multipathing drivers that are available on a wide variety of host types, with no additional host support required to access the highly available volume. Where multipathing drivers support Asymmetric Logical Unit Assignment (ALUA), the storage system tells the multipathing driver which control enclosures are closest to it, and should be used to minimize I/O latency.

You must assign a site value to the host, to the FlashSystem control enclosures, and storage enclosures. The ALUA supporting multipathing driver configures the host pathing optimally. Details about the configuration are described in 6.5, “Configuration” on page 138.

Tip: When using the HyperSwap function, configure your host multipath driver to use an ALUA-based path policy.

6.2 HyperSwap design

This section provides high-level information about HyperSwap. Details are described throughout the whole chapter.

The IBM FlashSystem HyperSwap function is an active-active mirror based on Metro Mirror technology. It is an *unstoppable* configuration. The relationship of a HyperSwap volume is never in “stopped mode,” except during disaster recovery scenarios.

The LUN ID of the master VDisk is presented to the host as the LUN ID of the HyperSwap volume. The auxiliary VDisk is always seen as offline in the GUI and CLI. The auxiliary VDisk is presented to the host with the same LUN ID as the master VDisk. HyperSwap simulates the master LUN ID for the auxiliary VDisk. The LUN ID of the auxiliary VDisk is not visible to the host.

Figure 6-4 shows the host can access the HyperSwap volume using the master and the auxiliary (aux) VDisk. In the CLI and GUI, the aux VDisk is shown `offline`, but the host can use it with the LUN ID of the master VDisk.

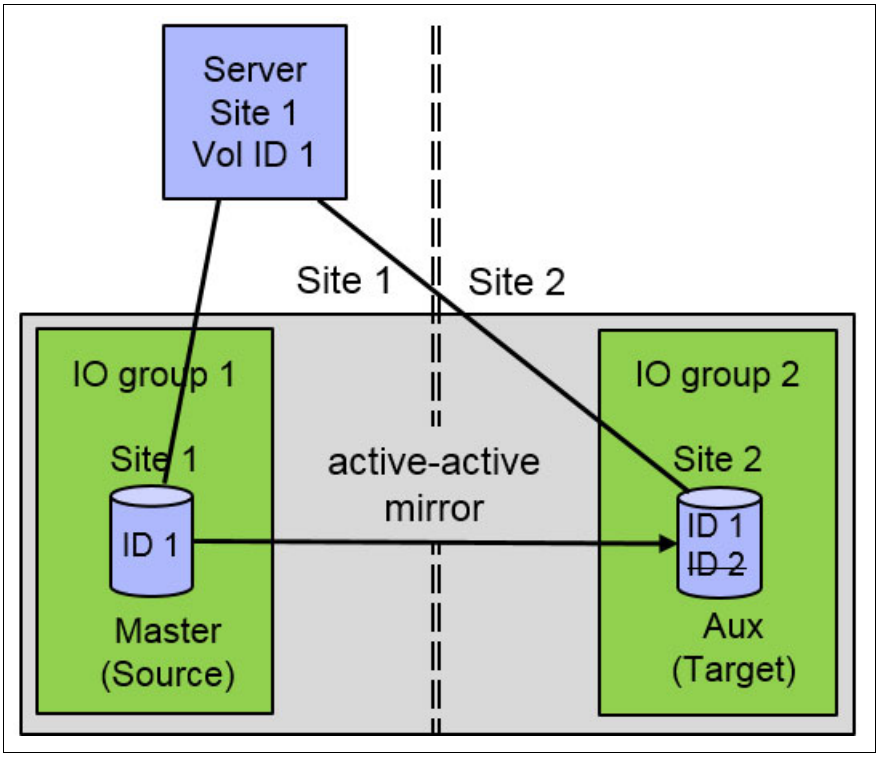


Figure 6-4 HyperSwap LUN ID Simulation

The host can access the HyperSwap volume using the I/O group on site 1, or the I/O group on site 2, or both. The multipath driver of the host is responsible for selecting the optimal paths.

The example in Figure 6-4 shows a host on site 1 accessing the HyperSwap volume using I/O group 1 and the *master* VDisk on site 1. Data is replicated to the *auxiliary* VDisk on site 2.

When the connection from the host to the master VDisk is broken, for example the Fibre Channel (FC) connection between host and I/O group 1 is broken, then the host accesses the data using I/O group 2.

The *master* VDisk is still the *primary* VDisk, so *reads* are serviced by the *master* VDisk and *writes* are forwarded to the *master* VDisk and then replicated to the *auxiliary* or the *secondary* VDisk. If this scenario is running for an adequate amount of time, usually more than 20 minutes, the *auxiliary* VDisk becomes the *primary* VDisk, servicing reads and writes. The *master* VDisk becomes the *secondary* VDisk. I/O arrives in I/O group 2 and is handled by the *auxiliary* (now primary) VDisk and replicated to the master (now secondary) VDisk.

A HyperSwap volume can be accessed concurrently for read and write I/O from any host in any site. All I/O is forwarded to one I/O group in the site with the *primary* VDisk. Using the site with the *non-primary* VDisk increases the long-distance traffic significantly.

HyperSwap Cluster monitors the workload and switches the copy direction if the most workload is arriving on the other site, optimizing performance.

Applications with equal workload pattern to the same HyperSwap volume using both I/O groups, for example Oracle RAC, are currently not optimal for HyperSwap.

Tip: If you are running VMware environment on top of HyperSwap, it is good practice to maintain VMs on the hosts in one site per HyperSwap volume. For example, with VMware Distributed Resource Scheduler (DRS), *should run* VM-host affinity rules.

A host accessing a HyperSwap volume uses two I/O groups. Therefore, the host multipathing must handle two times more paths compared to a normal volume. When changing from standard to HyperSwap topology, the host zoning has to be reviewed. Some hosts have limits for the optimal number of paths to a LUN. If only HyperSwap volumes are configured on FlashSystem V9000, meaning each host accesses a volume using two I/O groups, the maximum number of host objects and the maximum number of volume mappings per host object is cut in half.

HyperSwap volumes use FlashCopy technology to provide consistency protection when synchronizing the *master* and *auxiliary* VDisk after a loss of sync, for example when the link between these VDisks was broken. One change volume per HyperSwap volume must be prepared on each site.

Therefore, a HyperSwap volume requires the configuration of four internal VDisks. Two FlashCopy mappings to each change volume are required (one in each direction), so four FlashCopy maps are required per HyperSwap volume. All of the needed VDisks and remote copy relationships are automatically generated when using the GUI or the CLI when creating a HyperSwap volume.

The *auxiliary* VDisk is offline because its ID is not shown to the host. The **lsvdisk** command shows all VDisks of a HyperSwap volume (Example 6-1).

Example 6-1 The lsvdisk command to display all VDisks of a HyperSwap volume

```
>lsvdisk
id name    IO_group_name status  mdisk_grp_name capacity volume_name function
0 HS_Vol_1 io_grp0       online  Pool_AE3_1     1.00GB   HS_Vol_1    master
1 vdisk1   io_grp1       offline Pool_AE3_2     1.00GB   HS_Vol_1    aux
2 vdisk2   io_grp0       online  Pool_AE3_1     1.00GB   HS_Vol_1    master_change
3 vdisk3   io_grp1       online  Pool_AE3_2     1.00GB   HS_Vol_1    aux_change
```

Figure 6-5 shows the same HyperSwap volume by using the GUI.

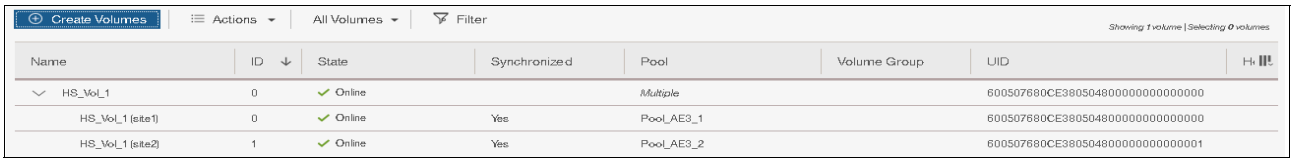


Figure 6-5 HyperSwap volume

The HyperSwap volume master VDisk *status* attribute of the **lsvdisk** command shows whether hosts are able to access data, for example whether the HyperSwap volume has access to up-to-date data or not. It does not show whether the master VDisk itself is actually online. The value *status* for *auxiliary* VDisk is always `offline`, as shown in Example 6-1 on page 132.

The GUI information shows only the HyperSwap volume but not the four VDisks of a HyperSwap volume. The site attributes and the pool of the *master* VDisk and the *auxiliary* VDisk are shown. When creating a HyperSwap volume by using the GUI or the **mkvolume** command, then the *master* VDisk disk will be on site1 and the *auxiliary* VDisk on site 2. Their associated change volumes are created in the same pool as the *master* or the *auxiliary* VDisk.

Use the *status* attribute of the **lsrcrelationship** command to determine if a VDisk is online or offline. Possible values are `online`, `primary_offline`, `secondary_offline`, `io_channel_offline`, `primary_change_offline`, `secondary_change_offline`, and `change_volumes_needed`, as described in 6.13.3, “The lsrcrelationship or lsrcconsistgrp commands” on page 183. Use the **lsvdisk** command to get the name of the remote copy relationship of a HyperSwap volume (Example 6-2).

Example 6-2 HyperSwap volume lsvdisk status information

```
>lsvdisk 0|grep RC_name
...
RC_name rcrel0
...
```

The **lsrcrelationship** command shows the four VDisks of a HyperSwap volume, some lines are omitted for better readability (Example 6-3).

Example 6-3 The lsrcrelationship command

```
>lsrcrelationship rcrel0
id 0
name rcrel0
master_cluster_id 0000020321601412
master_cluster_name HURSVT7
master_vdisk_id 0
master_vdisk_name HS_Vol_1
aux_cluster_id 0000020321601412
aux_cluster_name HURSVT7
aux_vdisk_id 1
aux_vdisk_name vdisk1
primary master
consistency_group_id 0
consistency_group_name HS_Group_AIX
state consistent_synchronized
```

```
bg_copy_priority 50
progress
freeze_time
status online
sync
copy_type activeactive
cycling_mode
cycle_period_seconds 300
master_change_vdisk_id 2
master_change_vdisk_name vdisk2
aux_change_vdisk_id 3
aux_change_vdisk_name vdisk3
```

Figure 6-6 shows the active-active relationship using the GUI.

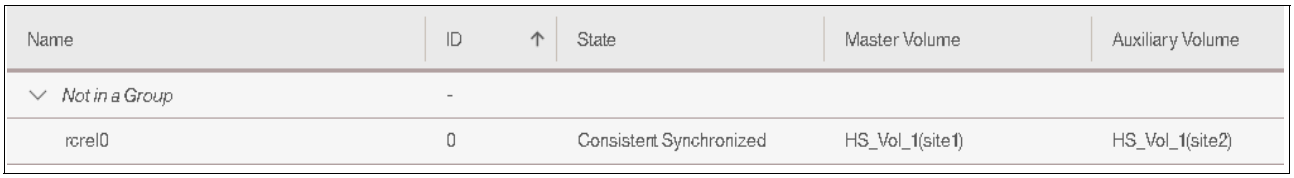


Figure 6-6 HyperSwap volume active-active relationship

The **lsfcmap** command shows the four FlashCopy mappings of a HyperSwap volume, as shown in Example 6-4.

Example 6-4 The lsfcmap command

```
>lsfcmap
id name  source_vdisk_id source_vdisk_name target_vdisk_id target_vdisk_name
0 fcmap0 0               HS_Vol_1          2               vdisk2
1 fcmap1 2               vdisk2            0               HS_Vol_1
2 fcmap2 1               vdisk1            3               vdisk3
3 fcmap3 3               vdisk3            1               vdisk1
```

Figure 6-7 shows the FlashCopy mappings using the GUI.

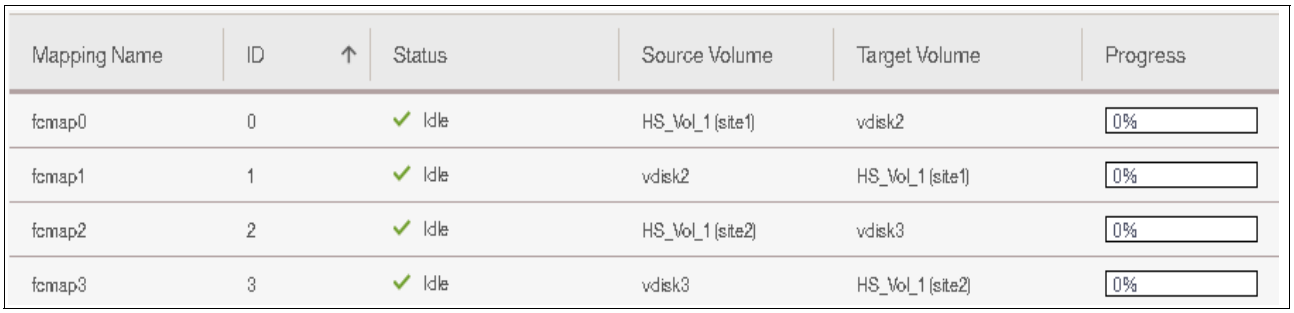


Figure 6-7 HyperSwap volume and its FlashCopy mappings

6.3 Comparison with Enhanced Stretched Cluster

Many of the aspects described so far are the same as those of the existing IBM Spectrum Virtualize Enhanced Stretched Cluster function, introduced in version 7.2 of the software. Table 6-1 provides a list of key differences between the Enhanced Stretched Cluster and HyperSwap topologies.

Note: Enhanced Stretched Cluster is not supported with IBM FlashSystem V9000.

Table 6-1 Enhanced Stretched Cluster and HyperSwap comparison

Description	IBM Spectrum Virtualize Enhanced Stretched Cluster	IBM FlashSystem V9000 HyperSwap
Product availability	SAN Volume Controller only	FlashSystem V9000 with 2 or more I/O groups
Configuration	CLI or GUI	CLI or GUI
Sites	Two for data, third for quorum device and/or DR	Two for data, third for quorum device
Distance between sites	Up to 300 km	Up to 300 km
Independent copies of data maintained	Two to three	Two (Four if additionally Volume Mirroring to two pools in each site)
Host requirements	Standard host multipathing driver	Standard host multipathing driver
Cache retained if only one site online?	No	Yes
Synchronization and resynchronization of copies	Automatic	Automatic
Stale consistent data retained during resynchronization for disaster recovery?	No	Yes
Scope of failure and resynchronization	Single volume	One or more volumes, user configurable
Ability to use FlashCopy together with high availability solution	Yes (although no awareness of site locality of data)	Limited
Ability to use Metro Mirror, Global Mirror, or Global Mirror with change volumes together with high availability solution	One remote copy	No, can use VDisk mirror for additional copies
Maximum highly available volume count	5000	1250
Licensing	Included in the base product	Requires Remote Mirroring license, check licensing Information

An overview of licensing can be found in *FlashSystem V9000 Product Guide*, REDP-5468.

The Enhanced Stretched Cluster function and the HyperSwap function spread the control enclosures of the system across two sites, with additional storage at a third site acting as a tie breaking quorum device.

The topologies differ in how the control enclosures are distributed across the sites:

- Enhanced Stretched Cluster

  For each I/O group in the system, the Enhanced Stretched Cluster topology has one control enclosure (in this topology often called “node”) on one site, and one node on the other site. The topology works with any number (1 - 4) of I/O groups, but because the I/O group is split into two locations, this is only available with SAN Volume Controller, not FlashSystem V9000.

- HyperSwap

  The HyperSwap topology locates both control enclosures of an I/O group in the same site, making this possible to use with either FlashSystem V9000 or SAN Volume Controller products. Therefore, to get a volume resiliently stored on both sites, at least two I/O groups are required.

The Enhanced Stretched Cluster topology uses fewer system resources, enabling a greater number of highly available volumes to be configured. However, during a disaster that makes one site unavailable, the SAN Volume Controller system cache on the nodes of the surviving site is disabled.

Requirement: Using HyperSwap requires the Remote Mirroring license.

6.3.1 Disaster recovery

The HyperSwap function automatically controls synchronization and resynchronization of the *master* VDisk and the *auxiliary* VDisk. If the *master* VDisk and the *auxiliary* VDisk are out of sync, for example the data link to one site had been broken and is fixed again, FlashSystem V9000 will automatically resynchronize the data. Just before resynchronizing data to a VDisk copy, that copy usually contains crash-consistent but *stale* (out-of-date) data. The storage system automatically retains that consistent data during the resynchronization process using change volume technology (which in turn is based on FlashCopy technology).

What this means is that if a problem occurs at the site with the online copy before resynchronization completes, taking that copy offline, you have the opportunity to manually enable read and write access to the consistent, older copy of data, allowing the use of this data for disaster recovery. This option would typically be taken if you know that the offline copy will remain offline for an extended period, and the consistent but older data is useful enough to keep your business running.

As usual with disaster recovery solutions that support business continuity with older data, after the problem is resolved restoring access to the offline copy, you can choose to either revert to that now-online copy, which before the disaster held the latest copy of the data, or continue to work on the stale data used during the disaster. With either choice, the other copy is resynchronized to match the chosen copy.

6.3.2 Consistency Groups

One major advantage of the HyperSwap function, compared to Enhanced Stretched Cluster, is that it is possible to group multiple HyperSwap volumes together for high availability. Using consistency groups to control the synchronization and failover across many HyperSwap volumes in an application ensures that all VDisk copies on a site have data from the same point in time, enabling disaster recovery using that site's VDisk copies.

It also ensures that at least one site has an up-to-date copy of every HyperSwap volume in the consistency group. It further ensures that the other site, if it does not have an up-to-date copy of every VDisk, has a consistent copy of every VDisk for some out-of-date point-in-time.

The following scenario is an example where the data is not consistent across all those volumes and would affect availability:

1. Site 2 goes offline.
2. Application continues to write to its volumes, changes only applied to site 1.
3. Site 2 comes online again.
4. Volumes are resynchronized back to site 2.
5. Site 1 goes offline during the resynchronization, leaving some volumes already resynchronized and some volumes unresynchronized.

Site 1 is the only site that has usable data. Site 2 might have usable data on some VDisks but not others. If this process is continued, it is possible that neither site will have a complete copy of data, making a failure on either site affect production I/O.

Without consistency groups, site 2's data would have been made inconsistent on several of the VDisks, by the attempt to resynchronize, which did not complete. The unresynchronized VDisks contain consistent but old data, as described in 6.3.1, "Disaster recovery" on page 136. Site 2 now has some VDisks with old data and VDisks with resynchronized data. If site 1 data cannot be recovered, another solution is needed to recover business operations.

6.3.3 HyperSwap restrictions for software version 8.1

The following restrictions apply to the HyperSwap function for software version 8.1:

- Cluster internal Metro Mirror is used for replication, so the size of a HyperSwap volume cannot be changed using **`expandvdisksize`** and **`shrinkvdisksize`** commands.
- A cascaded Remote Copy is currently not available. HyperSwap volumes cannot be replicated to a second, independent storage system using Remote Copy functionality.
- Four FlashCopy mappings are required for each HyperSwap volume, which limits the number of HyperSwap volumes to 1250. Additional FlashCopy requirements will reduce the number of possible HyperSwap volumes.

  Check the Configuration Limits and Restrictions.
- FlashCopy usage can be complicated because the Metro Mirror source and target volume can switch during daily operation. For this reason, identification of the copy direction is required for a successful FlashCopy.
- The Remote Copy relationship must be removed first for a reverse FlashCopy operation. After a reverse FlashCopy, all HyperSwap related functions must be manually implemented again (Remote Mirror and FlashCopy relationships).
- IBM FlashCopy Manager is currently not supported with HyperSwap volumes.

6.4 Planning

Two steps are required to configure IBM FlashSystem V9000 for HyperSwap. The first step is to configure the components of the system correctly for the HyperSwap topology. The second step is to create HyperSwap volumes that use that topology.

The first step includes these high-level tasks:

1. Planning the SAN Configuration
2. Defining the sites
3. Configuring the control enclosures
4. Configuring the FlashSystem V9000 internal storage and FlashSystem V9000 AE3 control enclosures
5. Configuring the external storage controllers
6. Defining the quorum device
7. Configuring the hosts
8. Configuring the HyperSwap topology
9. Configuring synchronization rates

You should plan to complete all steps to configure sites for control enclosures, storage enclosures, controllers, hosts, and the system topology in one session. Do not leave a system in production if only some of these steps have been performed.

The IBM FlashSystem V9000 optional SAS attached expansion enclosures are the IBM FlashSystem V9000 internal storage. The AE3 flash storage enclosures and the external storage controllers are the additional storage systems, such as an IBM Storwize V7000, which are attached to IBM FlashSystem V9000.

A 3-site setup is required (Figure 6-3 on page 130). Two sites are used as the main data center to provide two independent data copies. A quorum disk or an IP-based quorum can be used as quorum device. However, the quorum device must be placed in a third, independent failure domain (preferably a third site).

The quorum disk must be supported as an *extended quorum device*. More information about HyperSwap configuration details on extended quorum is in IBM Knowledge Center.

The IP quorum substitutes the active tie-breaker role of the quorum disk. Redundancy can be implemented by using multiple (up to 5) quorum applications, similar to multiple quorum disks. However, only one application is active at a time, the other applications are available if the active quorum device application fails. IBM Knowledge Center has more information about quorum disks.

6.5 Configuration

Several system objects must be configured before selecting the HyperSwap system topology, including sites, AC3 control enclosures, AE3 (or AE2) storage enclosures, storage controllers, and hosts.

6.5.1 SAN Configuration

Two IBM FlashSystem V9000 scalable building blocks in one cluster are required: each scalable building block is placed in one main site. Appropriate Fibre Channel (FC) connections are required between both sites. Check the references on configuration limits and restrictions in 6.3.3, “HyperSwap restrictions for software version 8.1” on page 137.

The two main SAN configurations options are:

- Configuration with inter-switch links (ISLs) between both sites
- Configuration without ISLs between both sites

In a configuration with ISLs between both sites, a division of the SAN infrastructure in a public and a private SAN is required. More details about requirements and use cases for HyperSwap system configuration are available in IBM Knowledge Center.

A quorum device in a third, independent failure domain is required. See 6.4, “Planning” on page 138.

Connection between both main sites

Metro Mirror is used for HyperSwap, so the Metro Mirror bandwidth requirements are the minimum requirements extended by HyperSwap specific requirements. The total required bandwidth between both sites depends on the SAN configuration (with or without ISL), the host peak workload, and the expected growth rate.

FlashSystem V9000 uses IBM FlashSystem MicroLatency modules and is built for low latency requirements. Therefore the SAN must provide lowest latency.

Important: The SAN must sustain IBM FlashSystem V9000 lowest latency.

Configuration with ISL

A bandwidth equal to the peak write bandwidth (as sum from all hosts) is required for intersite communication between I/O groups. This bandwidth must be available in the private SAN. Additionally, you need intersite bandwidth in the public SAN for host-to-control enclosure read and write communication if a host accesses control enclosures in the other sites. For example, after a failure of the local I/O group of the host, or to access volumes that do not use the HyperSwap function.

The guideline for a bandwidth equal to the peak write bandwidth for private SANs gives the minimal bandwidth supported for HyperSwap operations. In some non-optimal configurations, additional bandwidth is required to avoid potential performance issues. For example, if hosts at different sites share a volume, then the private SAN needs bandwidth equal to two times the peak write bandwidth plus the peak read bandwidth. Additional bandwidth is required for initial synchronization or resynchronization.

Therefore, the total required bandwidth can be calculated in the following way:

- Public SAN:
  - Peak host read and write throughput
  - Expected growth
- Private SAN:
  - Peak host write throughput
  - Surcharge for volumes used from hosts in both sites
  - Bandwidth for initial synchronization and resynchronization
  - Expected growth rate

However, in any case at least 4 Gbps are required in the private SAN, if the calculated bandwidth is below 4 Gbps. Consider the following example:

- Assume that the total host peak throughput is 30 GBps, 10 GBps are approximately write throughput.
- Assume that 20% of the peak host workload is created on volumes accessed from hosts at both sites, such as in an Oracle Real Application Clusters (RAC) environment.
- Add 20% for initial synchronization and resynchronization.
- The expected growth is 50% in 3 years.

The simplified bandwidth calculation can be done in the according to the following tables.

Table 6-2 shows the calculation for public SAN.

Table 6-2 Public SAN calculation

Public SAN	**Input**	**Formula**	**Result**
Total host peak read and write throughput	30 GBps		30 GBps
Expected growth	50%	50% of 30 GBps	15 GBps
Total Public SAN bandwidth		30 GBps + 15 GBps	**45 GBps**

Table 6-3 shows the calculation for private SAN.

Table 6-3 Private SAN calculation

Private SAN	**Input**	**Formula**	**Result**
Peak host write throughput	10 GBps		10 GBps
Volumes used from hosts in both sites	20%, half of them use the wrong site, so effective 10%	10% of 10 GBps	1 GBps
Synchronization	20%	20% of 10 GBps	2 GBps
Subtotal		10 GBps + 1 GBps + 2 GBps	**13 GBps**
Expected growth	50%	50% of 13 GBps	6.5 GBps
Total Private SAN bandwidth		13 GBps + 6.5 GBps	**19.5 GBps**

So at least 45 GBps throughput is required for public SAN, and 19.5 GBps is required for the private SAN.

Configuration without ISL

The required bandwidth in a configuration without ISL is significantly lower compared to a configuration with ISL because IBM FlashSystem V9000 does not have to hold bandwidth in different SANs.

The following example uses the example discussed in “Configuration with ISL” on page 139:

- Assume that the total host peak throughput is 30 GBps, of which 10 GBps are write throughput.
- Assume that 20% of the peak host workload is created on volumes accessed from hosts at both sites, such as in an Oracle RAC environment.
- Add 20% for initial synchronization and resynchronization.
- The expected growth is 50% in 3 years.

The simplified bandwidth calculation can be done as shown in Table 6-4.

Table 6-4 Simplified bandwidth calculation

SAN	Input	Formula	Result
Peak host read and write throughput	30 GBps		30 GBps
Peak host write throughput	10 GBps		
Volumes used from hosts in both sites	20%, half of them use the wrong site, so effective 10%	10% of 10 GBps	1 GBps
Synchronization	20%	20% of 10 GBps	2 GBps
Subtotal		30 GBps + 1 GBps + 2 GBps	**33 GBps**
Expected growth	50%	50% of 33 GBps	16.5 GBps
Total Private SAN bandwidth		33 GBps + 16.5 GBps	**49.5 GBps**

At least 49.5 GBps throughput is required in a configuration without ISL.

SAN design

A redundant design is suggested with two independent physical links between both sites, each link should be able to handle the calculated workload separately.

Management of the storage enclosure is using the FC connections between the control enclosure having the configuration node role and the storage enclosures. Every control enclosure can have the configuration node role and therefore every control enclosure must be zoned to every control enclosure.

Figure 6-8 on page 142 shows a schematic zoning design to illustrate the need of FC path from every controller node to every storage enclosure. This is shown as *Enclosure Zone* in the figure.

The *Internal Cluster I/O* zone shown in Figure 6-8 is used for the control enclosure to control enclosure traffic. This includes the active-active mirroring traffic and must be sized accordingly.

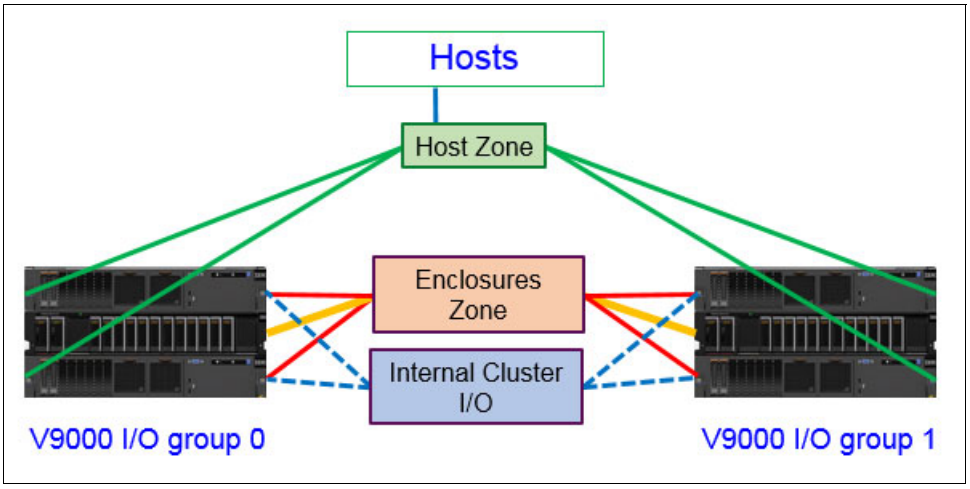


Figure 6-8 Schematic zoning design

6.5.2 Defining the sites

The *site* corresponds to a physical location that houses the physical objects of the system. In a client installation, it can correspond to a separate office, an isolated fire zone, for example a separate office with a unique firewall address, a different data center building, or simply different rooms or racked areas of a single data center that has been planned to have internal redundancy (i.e an independent failure domain).

Parameters that specify that a site exists are used in many of the commands described later. Table 6-5 on page 143 shows the four sites statically defined in the system.

Table 6-5 Site information

Site ID	Default site name	Objects that can be in site	Purpose
None	Has no name, cannot be renamed	Hosts, control enclosures, controllers, storage, and expansion enclosures	The default site for objects when they are not assigned to a specific site. The HyperSwap topology requires objects to be in a site.
1	site1	Hosts, control enclosures, controllers, storage, and expansion enclosures	The first of two sites to perform high availability between. Has no implied preferences compared to site 2.
2	site2	Hosts, control enclosures, controllers, storage, and expansion enclosures	The second of two sites to perform high availability between. Has no implied preferences compared to site 1.
3	site3	Controllers, Fibre Channel external Quorum	A third site providing quorum abilities to act as a tie-break between sites 1 and 2 when connectivity is lost.

Sites 1, 2, and 3 can be renamed from the default name by using the **chsite** command and can be listed by using the **lssite** command (Example 6-5).

Example 6-5 The lssite command to rename default sites

```
chsite -name datacenter_west 1
chsite -name datacenter_east 2
chsite -name quorum_site 3
lssite
ID site_name
1  datacenter_west
2  datacenter_east
3  quorum_site
```

In Example 6-5, the **chsite** command is issued to rename site 1 as `datacenter_west`, site2 as `datacenter_east`, and site 3 as `quorum_site`. This can help you to understand and describe the location of objects in a more meaningful way. This document uses the default names *site1*, *site2*, and *site3* for sites.

6.5.3 Control enclosures

With a HyperSwap system topology, all control enclosures in an I/O group must belong to the same site. You should assign the control enclosures of at least one I/O group to each of sites 1 and 2.

To configure HyperSwap volumes on IBM FlashSystem V9000, you need at least four control enclosures, two scalable IBM FlashSystem V9000 building blocks.

Before the HyperSwap system topology can be selected, the site of every control enclosure must be set by using either of the **chnode** command lines shown in Example 6-6.

Example 6-6 The chnode command to assign a site attribute to a control enclosure

```
chnode -site 1 node1
chnode -site site1 node1
```

This modifies the existing control enclosure `node1` from its current site to site 1. This command must be used for all control enclosure.

Note: Every control enclosure of an I/O group must be in the same site. A control enclosure can never be assigned to site 3.

An I/O group containing copies of HyperSwap volumes can not be moved to the other site. The only way to move an I/O group from one site to the other is to remove all HyperSwap volume copies using that I/O group, delete the control enclosures from that I/O group, then re-add them to the I/O group but with the new site attribute.

Typically, a HyperSwap configuration might contain two or four building blocks, either with one I/O group on site 1 and one I/O group on site 2, or with two I/O groups on each of sites 1 and 2. A possibility is to configure the system with more I/O groups on one site than the other, although the site with fewer control enclosures might become a bottleneck.

Each I/O group should have sufficient bitmap capacity defined using the **`chiogrp`** command for the HyperSwap volumes in addition to the bitmap capacity requirements of other FlashCopy and Global Mirror or Metro Mirror objects needed.

6.5.4 Configuring the AE3 enclosures and external storage controllers

For all AE3 enclosures and virtualized external storage, you must assign the site attribute to all controllers using the following command:

`chcontroller –site <site id> <controller id>`

Controllers should be assigned to site 1, site 2, or site 3 when used for quorum, if they have any managed MDisks and the system is set to use the HyperSwap topology. MDisks can only be assigned to storage pools if they are allocated from a storage controller with a well-defined site that matches that of the storage pool.

You cannot create a HyperSwap volume using a storage controller without a site attribute. The site attribute of a storage controller can be set before or after changing to the HyperSwap topology.

Note: Always assign a site attribute to IBM FlashSystem V9000 control, and storage enclosures *before* changing to HyperSwap topology.

6.5.5 Configuring the IBM FlashSystem V9000 storage enclosures

You must assign the `site` attribute to all IBM FlashSystem V9000 AE2 storage enclosures using the **`chenclosure`** command:

`chenclosure -site <site id> <enclosure id>`

A site attribute for an optional SAS attached expansion enclosure cannot be set because it is attached to control enclosures having a site attribute. Pools of the expansion enclosures will have the site information of the corresponding control enclosure.

You cannot create a HyperSwap volume using a storage enclosure without a site attribute. The site attribute of a storage controller can be set before or after changing to the HyperSwap topology.

6.5.6 Define quorum device

The quorum device can be implemented by using an external storage system, which must be assigned to the third site. Alternatively, an IP quorum application can be used to provide quorum. IBM Knowledge Center has more information about quorum disk.

To assign an external storage system to the third site, use commands in Example 6-7.

Example 6-7 Set site attribute for external storage controller

```
lscontroller
id controller_name   ctrl_s/n     vendor_id product_id_low
0  controller0_V7000 2076         IBM       2145
1  controller1_V7000 2076         IBM       2145

chcontroller -site 3 controller0_V7000
chcontroller -site 3 controller0_V7000
```

The HyperSwap environment requires an active quorum disk in site 3 and one quorum disk candidate in each of the other two main sites. The storage system V7000 is located in the third site and one managed disk is used as the active quorum disk. IBM FlashSystem V9000 storage enclosure spare disks drives are used in site 1 and site 2 as quorum disk candidates.

The IBM FlashSystem V9000 cluster should not be able to change the quorum disk settings, so the **override yes** flag must be used with the **chquorum** command. Also, use the **-active** parameter to be sure that the quorum disk on the tie-breaker site is active. Use the **chquorum** command to set the tie-breaker quorum disk:

```
quorum -active -mdisk <mdisk of site 3> -override yes <quorom id>
```

The controller providing the quorum storage has to specify “extended quorum” support on the SAN Volume Controller supported controller list for the installed software release.

Using the IP quorum application as a tie-breaker

You can us the IP quorum application if no storage is available at the third site.
IBM Knowledge Center has more details about IP quorum configuration.

When the IP quorum application is running, it is automatically detected as third quorum device (Example 6-8).

Example 6-8 IP quorum disk

```
lsquorum
quorum_index status id active object_type override site_id site_name
0            online 11 no     drive       no       1       site1
2            online 23 no     drive       no       2       site2
3            online    yes    device      no               x3690-x5/<ip-address>
```

6.5.7 Configuring the hosts

Host objects have a site parameter. This parameter can be configured on existing host objects as follows:

```
chhost –site 1 Host_AIX
```

This command defines the ports of host `Host_AIX` as being on site 1. If the System is in HyperSwap topology you can use the GUI to assign a site to the host.

Important: The system dynamically configures host multipathing so that hosts in site 1 preferentially send I/O to control enclosures in site 1, and similarly for site 2. So for optimum performance, all of the WWPNs associated with this host object should be on that site. For clustered host systems attached to both sites, you should define a host object per site to optimize the I/O for each physical server in the clustered host system.

New hosts can be added with a defined site by using the following command:

```
mkhost -fcwwpn <WWPN:WWPN> –site <site id>
```

When HyperSwap volumes are mapped to a host by using the **`mkvdiskhostmap`** command, the host must be assigned to either site 1 or site 2.

By default, host objects are associated with all I/O groups. If you use the **`-iogrp`** parameter for the **`mkhost`** command to override this, be sure that hosts accessing HyperSwap volumes are associated with at least the I/O groups that the master and auxiliary VDisk of the HyperSwap volumes are cached in. Missing an association between the host and such an I/O group prevents the host from being able to access HyperSwap volumes through both sites.

Notes:

- A HyperSwap volume has two accessible I/O groups. A HyperSwap volume can be mapped only to a host having access to all accessible I/O groups of the HyperSwap volume.
- A Fibre Channel attached host will show the `degraded` status, if not zoned to all control enclosures belonging to the caching and access I/O groups of a mapped HyperSwap volume.

6.5.8 Configuring the HyperSwap topology

All control enclosures, storage enclosures, and storage controllers can be set to any of sites 1 or 2 (or 3 for controllers) when the system has been set to the standard system topology. Use the command **`lssystem`** to check the current topology, as shown in Example 6-9.

Example 6-9 The lssystem command to check current topology

```
lssystem

...
topology standard
...
```

Figure 6-9 shows the FlashSystem V9000 GUI in standard topology.

Figure 6-9 FlashSystem V9000 GUI with standard topology

Before the system can be set to the HyperSwap system topology, every control enclosure must have a site configured correctly, and it is advisable to set the site of every storage enclosure, expansion enclosure, storage controller, and host too for existing systems. For a new system, you can choose the HyperSwap topology early in your initial configuration, which helps ensure that objects have their sites set correctly.

When all of the sites have been set, the system can be set to use the HyperSwap topology using the **chsystem** command:

```
chsystem –topology hyperswap
```

Figure 6-10 shows the IBM FlashSystem V9000 GUI in HyperSwap topology.

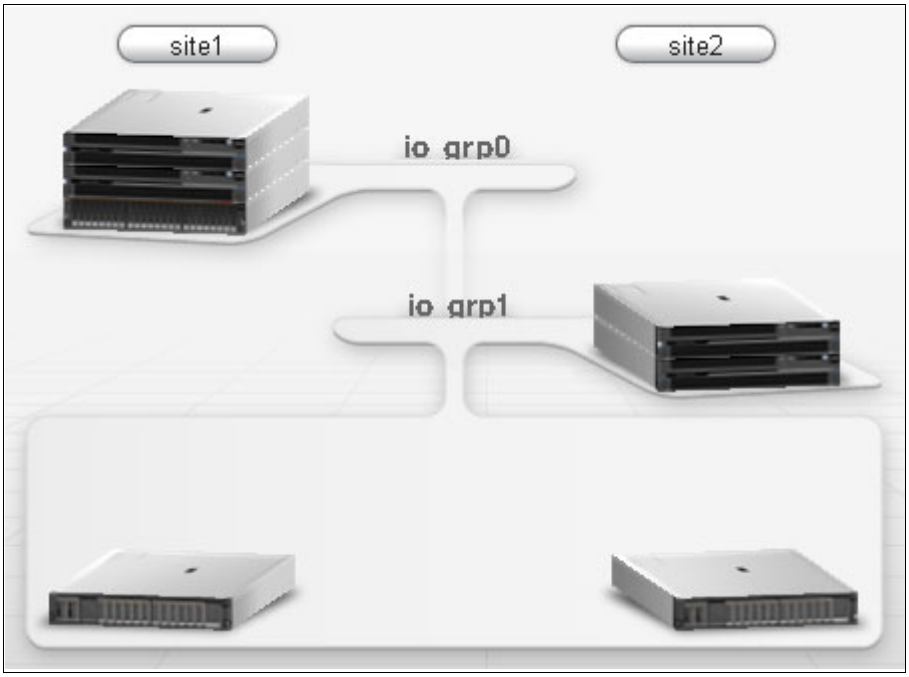


Figure 6-10 FlashSystem V9000 GUI with HyperSwap topology

The site attributes of the control enclosures, and storage enclosures had been set before enabling HyperSwap.

Note: You will not be able to change the topology back to the standard topology if any HyperSwap volumes are defined.

6.5.9 Configuring synchronization rates

Two primary factors affect the synchronization rate and are similar to those for the existing Metro Mirror and Global Mirror replication technologies:

- Partnership bandwidth

  The total bandwidth between site 1 and 2. This is foreground traffic, such as transferring new host writes to the second site, and background traffic, such as synchronization of new HyperSwap volumes or resynchronization.

  You can limit the background traffic of HyperSwap volumes separately from foreground traffic. Limiting the amount of background traffic assures a minimum value for foreground traffic.

- Relationship bandwidth

  This is the background traffic limitation per VDisk.

Partnership bandwidth

The primary attribute to configure is the *partnership bandwidth.* Before the introduction of HyperSwap volumes, this could not be configured for intra-cluster relationships (for example, with both copies in the same system), such as the active-active relationships used for HyperSwap replication. With HyperSwap-capable systems, the *local partnership* bandwidth can be configured, and represents the amount of physical bandwidth between sites used for synchronization.

For compatibility with earlier versions, this defaults to 25 MBps (200 Megabits per second) dedicated to synchronization, which can be appropriate for a small environment. For larger systems, or systems with more bandwidth available between sites, you might want to increase this by using the following command:

```
chpartnership -linkbandwidthmbits 4000 -backgroundcopyrate 20 <localCluster>
```

In this command, you can specify the bandwidth between sites, and how much can be used for synchronization. **<*localCluster*>** should be replaced by the name of the local system.

The **-linkbandwidthmbits** parameter specifies the aggregate bandwidth of the link between two sites in megabits per second (Mbps). It is a numeric value 15 - 100,000. The default is 200, specified in megabits per second (Mbps). This parameter can be specified without stopping the partnership.

The **-backgroundcopyrate** parameter specifies the maximum percentage of aggregate link bandwidth that can be used for background copy operations. It is a numeric value 0 - 100, and the default value is 100, which means that a maximum of 100% of the aggregate link bandwidth can be used for background copy operations.

As with other types of partnership configuration, the system does not yet use the total amount of bandwidth available in any performance tuning, and only uses the resulting background copy bandwidth to determine HyperSwap synchronization rate. So the previous command could also be expressed as the following command:

```
chpartnership -linkbandwidthmbits 800 –backgroundcopyrate 100 <localCluster>
```

This command has the same effect concerning the background traffic, but the earlier command reserves 3,200 MBps for foreground VDisk traffic.

The system will attempt to synchronize at the specified rate for background traffic where possible if there are any active-active relationships that require synchronization (including resynchronization after a copy has been offline for some time). This is true no matter how much new host write data is being submitted requiring replication between sites, so be careful not to configure the synchronization rate so high that this synchronization bandwidth consumption eats up all link bandwidth thus affecting host writes.

Relationship bandwidth

The other control of how fast a relationship can synchronize is the system setting `relationship_bandwidth_limit`. This setting configures the maximum rate at which synchronization I/O is generated for a HyperSwap volume. It is shown with the **`lssystem`** command (Example 6-10).

Example 6-10 Using lssystem, maximum rate synchronization I/O is generated for a volume

```
lssystem
...
relationship_bandwidth_limit 25
...
```

By default this is 25 MBps, this is in megabytes, not in megabits as the partnership configuration. This means that no matter how few relationships are synchronizing, the most synchronization I/O that is generated per HyperSwap volume is 25 MBps (this is 25 MBps of reads on the up-to-date copy, and 25 MBps of writes on the other copy).

If your system has storage that cannot handle the additional 25 MBps of I/O, you can configure this to a lower value using the **`chsystem`** command:

```
chsystem –relationshipbandwidthlimit 10
```

If you want to accelerate synchronization when there aren't many HyperSwap volumes synchronizing, you might want to increase it to a higher value:

```
chsystem –relationshipbandwidthlimit 200
```

The **`-relationshipbandwidthlimit`** parameter specifies the new background copy bandwidth in megabytes per second (MBps), 1 - 1000. The default is 25 MBps. This parameter operates system-wide and defines the maximum background copy bandwidth that any relationship can adopt. The existing background copy bandwidth settings that are defined on a partnership continue to operate, with the lower of the partnership and volume rates attempted.

Note: Do not set this value higher than the default without ensuring that the higher bandwidth can be sustained.

6.5.10 HyperSwap configuration using the GUI wizard

HyperSwap configuration is possible to do by using the GUI. Complete these steps:

1. Start the HyperSwap GUI wizard, select **Actions** → **Modify System Topology** (Figure 6-11).

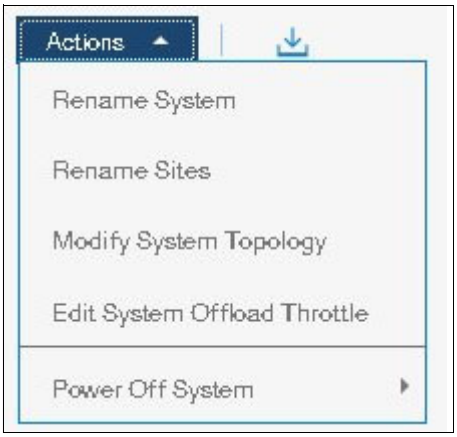


Figure 6-11 Modify System Topology

The Modify System Topology welcome page opens (Figure 6-12).

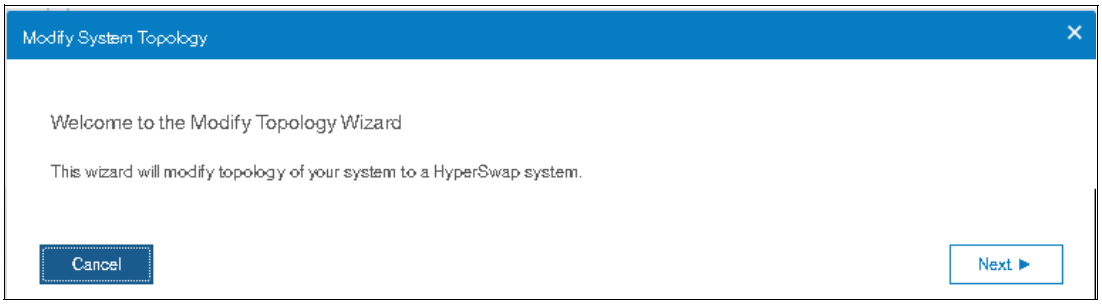


Figure 6-12 Modify System Topology wizard welcome page

2. Assign the site names (Figure 6-13).

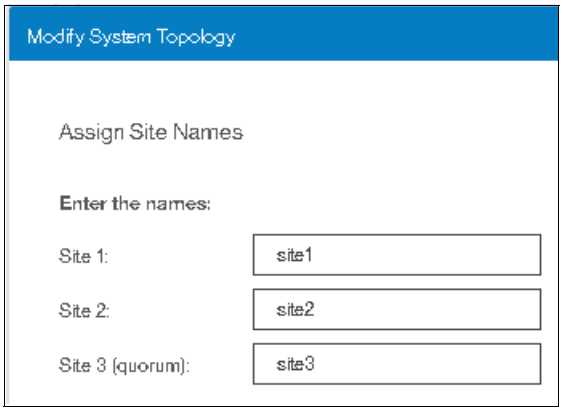


Figure 6-13 Assign site names

3. In this example, the default site names are used. Assign the control enclosures to sites by clicking the dual-arrow radio buttons (Figure 6-14).

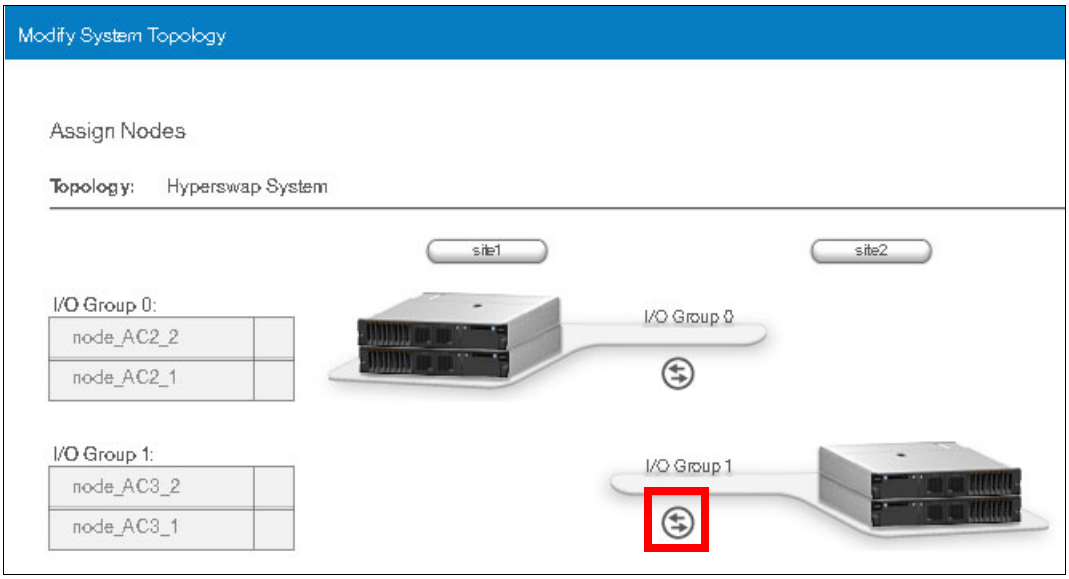


Figure 6-14 Assign control enclosures (nodes) to sites and I/O groups

4. Assign hosts to a site (Figure 6-15).

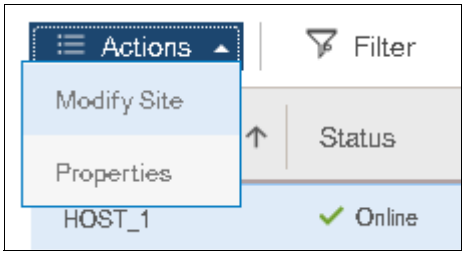


Figure 6-15 Optional host site attributes

5. The site assignment for hosts is mandatory for HyperSwap (optional for Enhanced Stretched Cluster). Assign the site attributes to the storage enclosure (Figure 6-16).

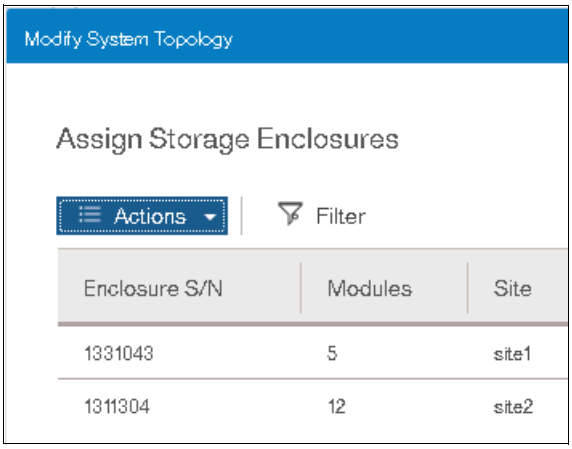


Figure 6-16 Storage enclosure site attributes

6. Assign the site attributes to the external storage controllers (Figure 6-17).

Note: At least one storage system must be assigned to the quorum site. IP quorum can not be defined using the modify topology wizard.

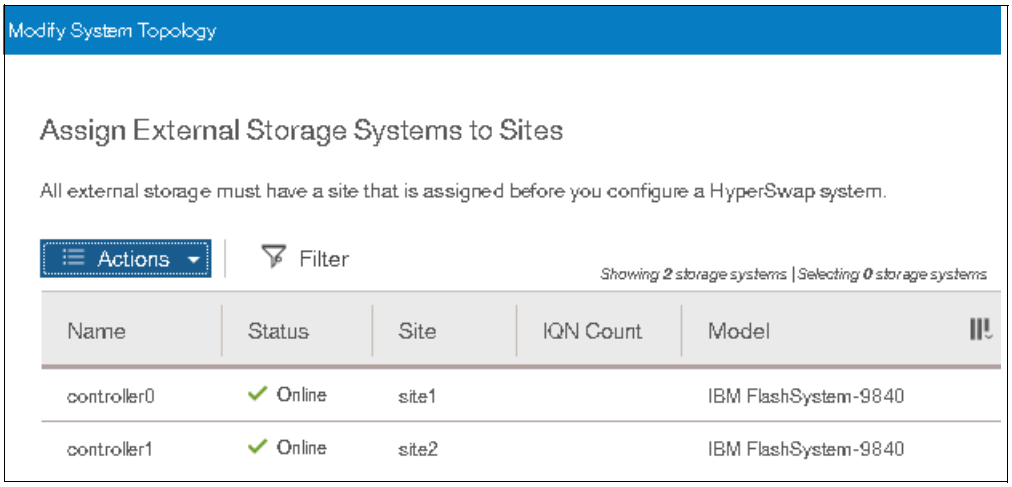


Figure 6-17 Storage controller site attributes

7. Set the bandwidth limits between sites (Figure 6-18).

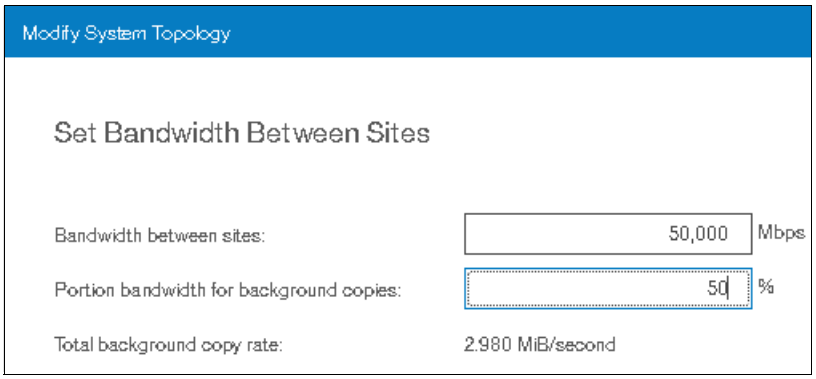


Figure 6-18 Bandwidth settings

Before the changes can be applied, a summary is displayed (Figure 6-19).

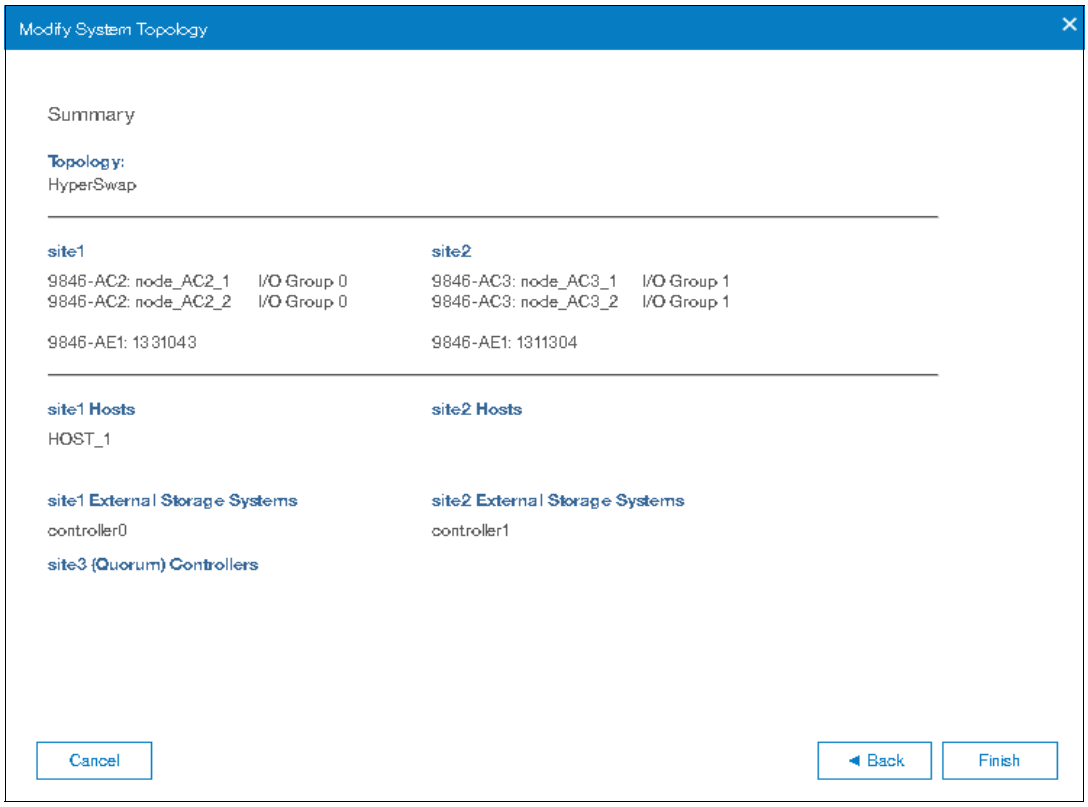


Figure 6-19 Summary of the GUI wizard

8. Click **Finish**. The HyperSwap configuration starts. After the tasks are completed, the system is in HyperSwap topology.
9. Manually set the quorum disk for the third site as described in 6.5.6, “Define quorum device” on page 145.

Note: The quorum device for the third site must be defined manually.

6.5.11 SAN environment for low latency

IBM FlashSystem V9000 provides lowest latency. A HyperSwap volume resides on two sites. Therefore between both sites, a minimal SAN latency is required. The IBM FlashSystem V9000 latency as seen by the host is depending on the SAN latency between the I/O groups on both sites. The write acknowledgement to the host will be done after both sites receive the data. Therefore, the latency of the SAN between both sites is added to the latency as seen by the host.

Note: Lowest SAN latency is needed to preserve IBM FlashSystem V9000 lowest latency.

6.5.12 Creating HyperSwap volumes

HyperSwap capability enables each HyperSwap volume to be presented by two I/O groups. One VDisk on an I/O group of each site stores the data. Each of these two VDisks uses yet another VDisk on the same site as change volume. When the relationship between these four VDisks is defined, one VDisk is the master VDisk, the other VDisk is the auxiliary VDisk, and these two VDisks have an associated Change Volume.

The two VDisks are kept synchronized by the IBM Spectrum Virtualize HyperSwap functions. The host only sees one HyperSwap volume. This HyperSwap volume has the LUN ID from the master VDisk.

Figure 6-20 shows the four VDisks and the HyperSwap volume presented to the host. The host always sees a HyperSwap volume with ID 1. The VDisk with ID 2 is synchronized with the VDisk with ID 1. If the host detects a HyperSwap volume on both I/O groups, both VDisks show ID 1 to the host. If the hosts multipathing driver is Asymmetric Logical Unit Access (ALUA) aware then the hosts multipathing driver detects and uses the preferred control enclosure for I/O.

In case of a failover, for example I/O group 1 is offline, the host accesses site 2 and uses VDisk 2, which presents ID 1 to the host. Even if internally there are different IDs, the host always sees the master ID 1. Therefore, the multipathing driver of the host can switch seamlessly to site 2.

Figure 6-20 shows four VDisk and the HyperSwap volume.

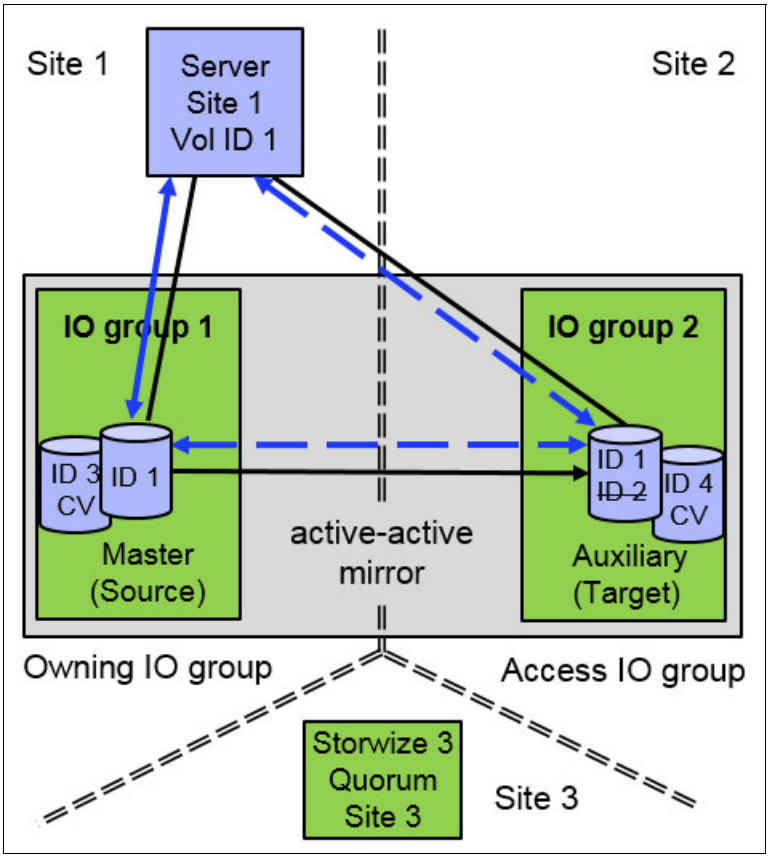


Figure 6-20 The HyperSwap volume build out of four VDisks

The GUI offers an easy-to-use management interface during the HyperSwap volume creation process. Click **Volumes** → **Volumes** → **Create Volumes** and then select **HyperSwap**. The HyperSwap wizard guides you through the volume creation process (Figure 6-21). The volume capacity, the volume name, and the pools in each site must be set. Use of additional functions such as thin provisioning and compression are optional settings. Settings related to I/O groups can be modified.

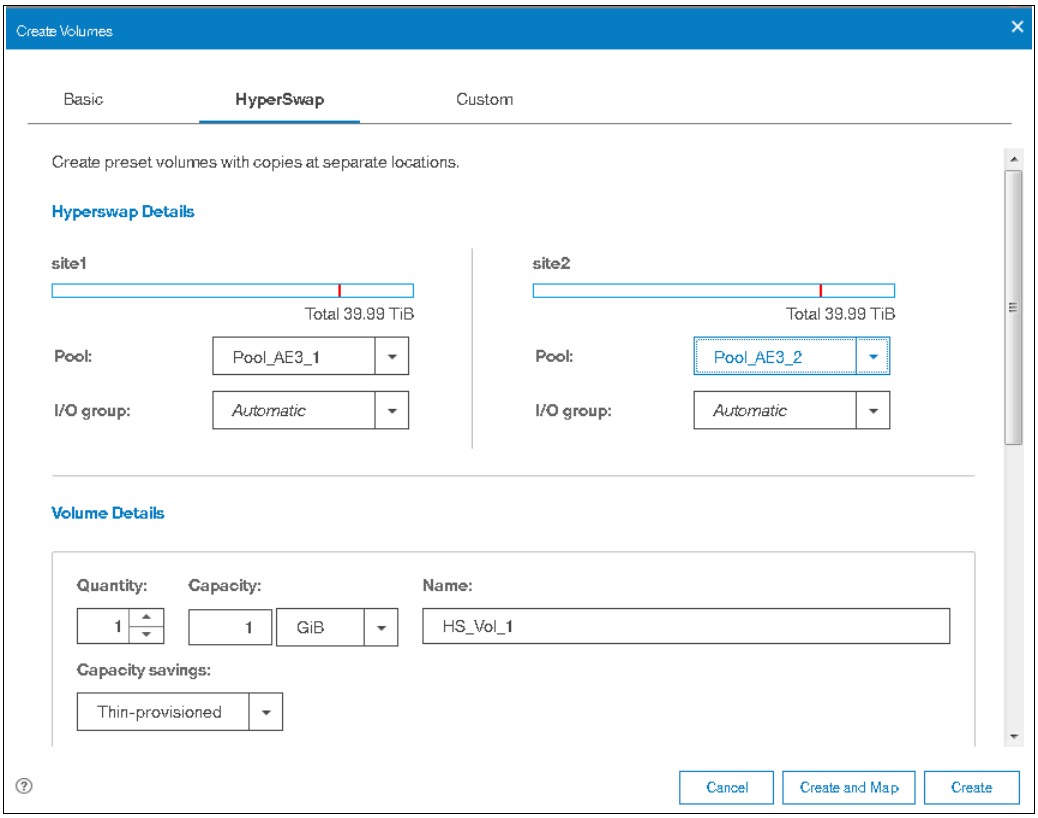


Figure 6-21 HyperSwap select pools

The setup in Figure 6-21 shows select pool *AE3_1* on site 1 and pool *AE3_2* on site 2.

For site 2, two possible pools are shown. Make sure to select pools with the same latency for HyperSwap volumes. This example shows the creation of two HyperSwap volumes (Figure 6-22).

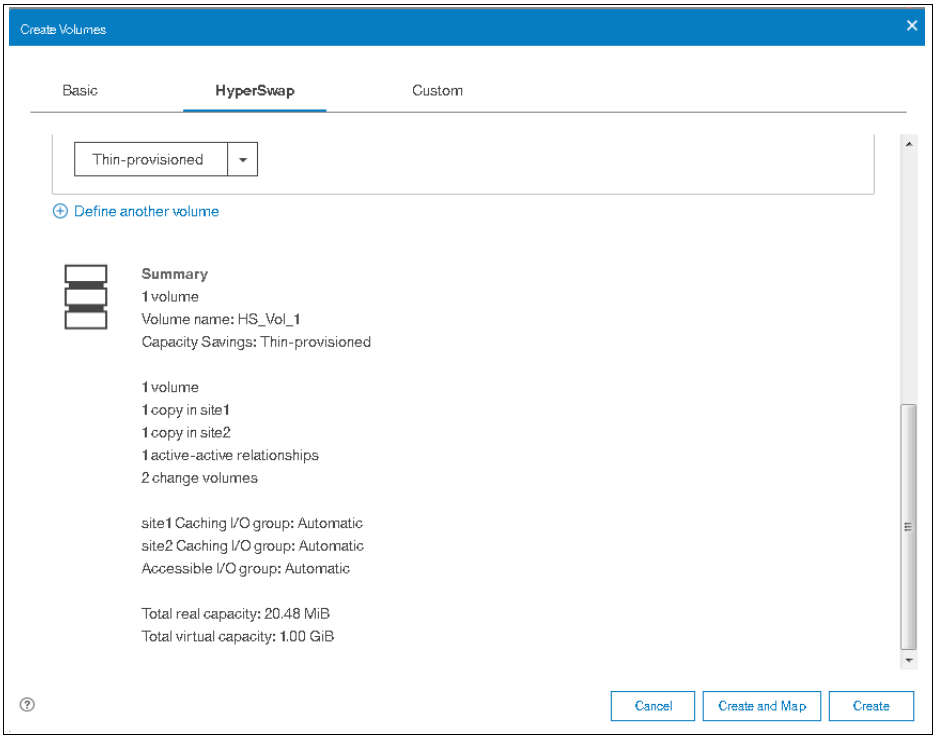


Figure 6-22 HyperSwap volume summary

The new HyperSwap volume is visible in the GUI (Figure 6-23) by selecting **Volumes** → **Volumes**. Only HyperSwap Master and Auxiliary volumes are shown in the GUI; the required FlashCopy volumes are hidden to reduce the level of complexity.

Name	ID	State	Synchronized	Pool
HS_Vol_1	0	Online		*Multiple*
HS_Vol_1 (site1)	0	Online	Yes	Pool_AE3_1
HS_Vol_1 (site2)	1	Online	Yes	Pool_AE3_2

Figure 6-23 HyperSwap volumes

If you create HyperSwap volumes without any space savings, they will always be formatted. Hovering over the state of the HyperSwap volume provides the estimated completion time (Figure 6-24).

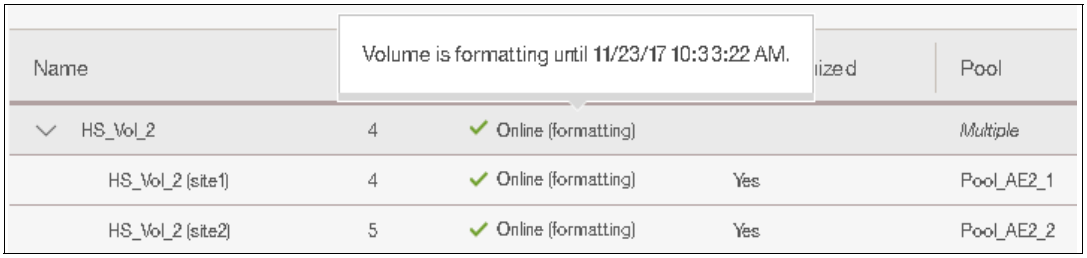


Figure 6-24 Formatting time

The **lsvdiskprogress** command displays the estimated completion time (Example 6-11).

Example 6-11 Estimating formatting completion time

```
lsvdiskprogress
id progress estimated_completion_time
4  12       171123093328
5  12       171123093328
```

Example 6-12 shows the HyperSwap master, auxiliary VDisks, and the associated change volumes. The change volumes will be automatically created using the same pool as the master or auxiliary VDisks respectively.

Example 6-12 lsvdisk information of HyperSwap volumes

```
lsvdisk
id name     IO_group_name status  disk_grp_name  RC_name volume_name    function
0  HS_Vol_1 io_grp0       online  Pool_AE3_1     rcrel0  HS_Vol_1       master
1  vdisk1   io_grp1       offline Pool_AE3_2     rcrel0  HS_Vol_1       aux
2  vdisk2   io_grp0       online  Pool_AE3_1     rcrel0  HS_Vol_1       master_change
3  vdisk3   io_grp1       online  Pool_AE3_2     rcrel0  HS_Vol_1       aux_change
4  HS_Vol_2 io_grp0       online  Pool_AE2_1     rcrel1  HS_Vol_2       master
5  vdisk4   io_grp1       offline Pool_AE2_2     rcrel1  HS_Vol_2       aux
6  vdisk5   io_grp0       online  Pool_AE2_1     rcrel1  HS_Vol_2       master_change
7  vdisk6   io_grp1       online  Pool_AE2_2     rcrel1  HS_Vol_2       aux_change
```

6.5.13 Creating a HyperSwap volume from a basic volume

Existing VDisks can be easily converted to a HyperSwap volume by using the **addvolumecopy** command or the GUI:

1. Use the **Add Volume Copy** option from the context menu of a basic volume (Figure 6-25).

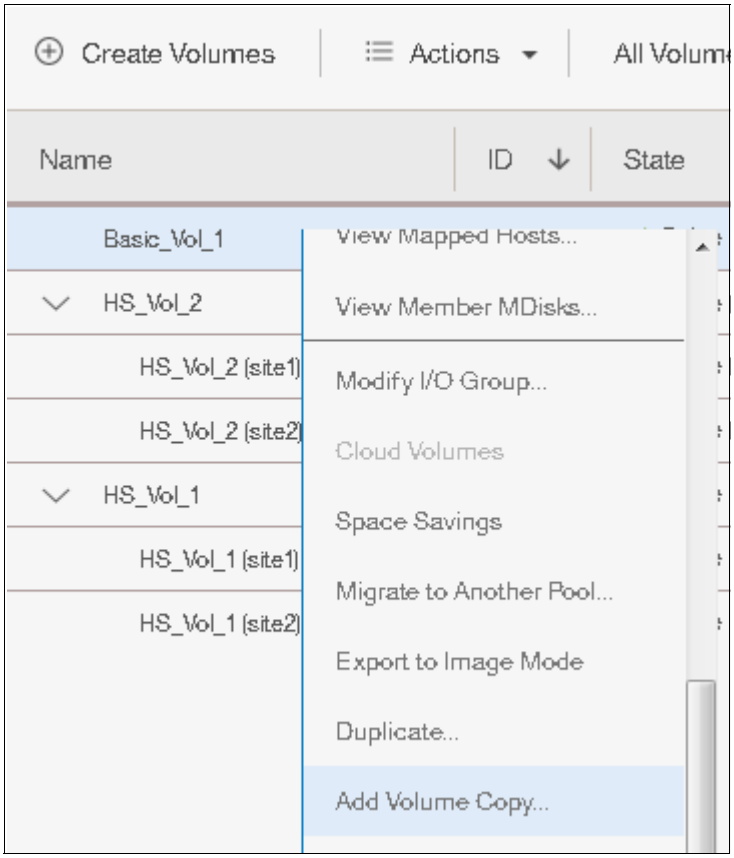


Figure 6-25 Add volume copy

2. Enter the appropriate values for the other site in the Add Volume Copy wizard. Figure 6-26 shows a basic volume on site 2, which will get a volume copy on site 1.

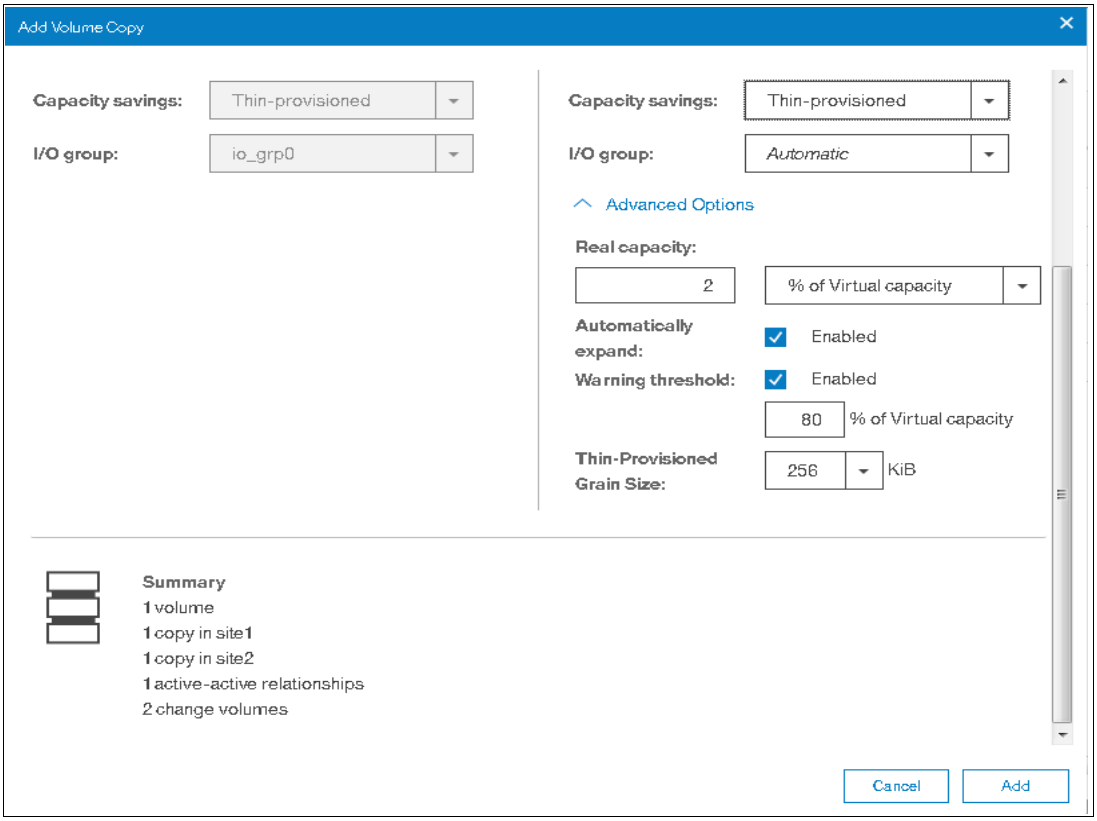


Figure 6-26 Add volume wizard

The running task information displays the progress of the HyperSwap volume synchronization (Figure 6-27).

Name	ID	State	Synchronized	Pool
Basic_Vol_1	8	Synchronizing		*Multiple*
Basic_Vol_1 (site1)	8	Online	Yes	Pool_AE3_1
Basic_Vol_1 (site2)	9	Synchronizing	Yes	Pool_AE3_2

Figure 6-27 Running tasks information

6.5.14 Mapping HyperSwap volumes to a host

Mapping a HyperSwap volumes to a host is the same process as mapping a basic volume to the host by using the CLI or the GUI.

6.6 Operations

The active-active relationship has a **primary** attribute like regular Metro Mirror and Global Mirror relationships. This is set to either `master` or `aux`. With an active-active relationship, the VDisk in one I/O group acts as the primary, supplying data for reads, and serializing writes. All reads and writes must be initially processed by that I/O group. This is the method where writes are consistently applied to the HyperSwap volumes.

The HyperSwap function optimizes the I/O traffic. HyperSwap monitors which I/O group gets most of the host I/O. The VDisk of the HyperSwap volume used by the host with the same site ID as the I/O group with most of the hosts I/O should act as primary.

From an initially created HyperSwap volume, the master VDisk acts as the primary. If the first I/O to this HyperSwap volume is submitted only to the auxiliary VDisk's site for approximately more than 10 minutes, then the system switches the direction of the relationship. The hosts have improved read and write performance. The *secondary* VDisk of this HyperSwap volume is now the *primary* VDisk and the *master* VDisk is now the *auxiliary* VDisk. The active-active relationship is now reversed. This will be indicated by the icon shown in Figure 6-28.

Figure 6-28 Icon to indicate reversed active-active relationship

HyperSwap volumes in consistency groups all switch direction together, so the direction that a set of active-active relationships in a consistency group replicates will depend on which of the two sites has most of the host I/O across all HyperSwap volumes.

To create a consistency group, complete the following steps:

1. Select **Copy Services** → **Remote Copy** → **Create Consistency Group**. The Create Consistency Group wizard opens (Figure 6-29).

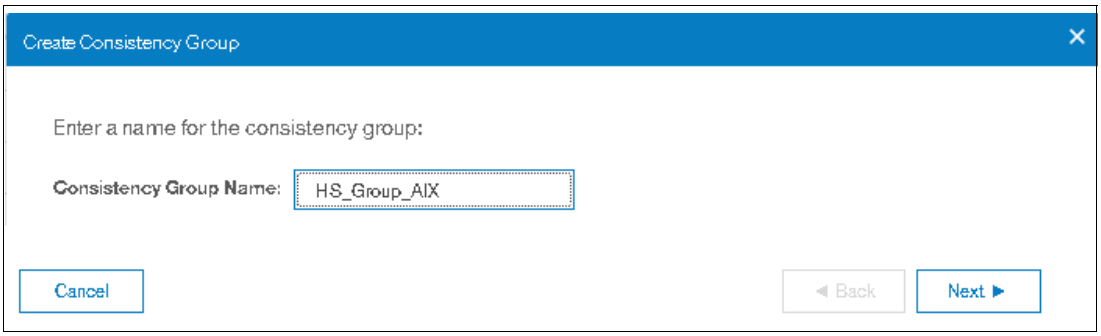


Figure 6-29 HyperSwap consistency group creation

The wizard guides you through the steps.

2. The auxiliary volumes are located on the same system, so select the appropriate options (Figure 6-30) and create an empty consistency group.

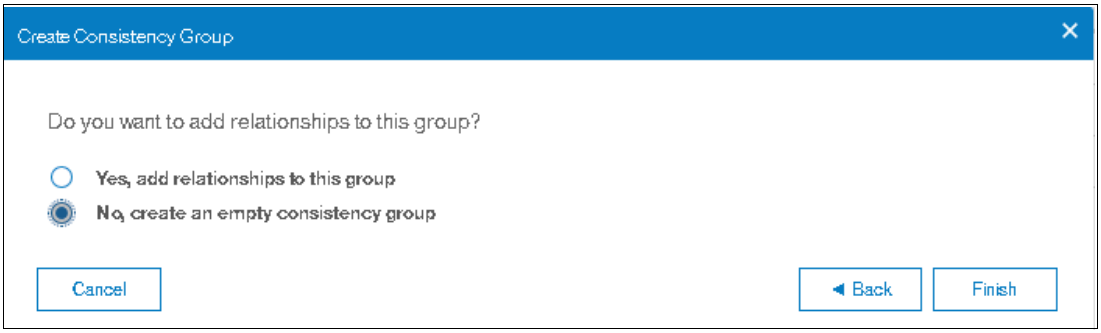


Figure 6-30 Consistency group system selection

3. Select the remote copy relationships you want to move to the consistency group and add them to the consistency group by using the **Actions** menu of the remote copy relationships (Figure 6-31).

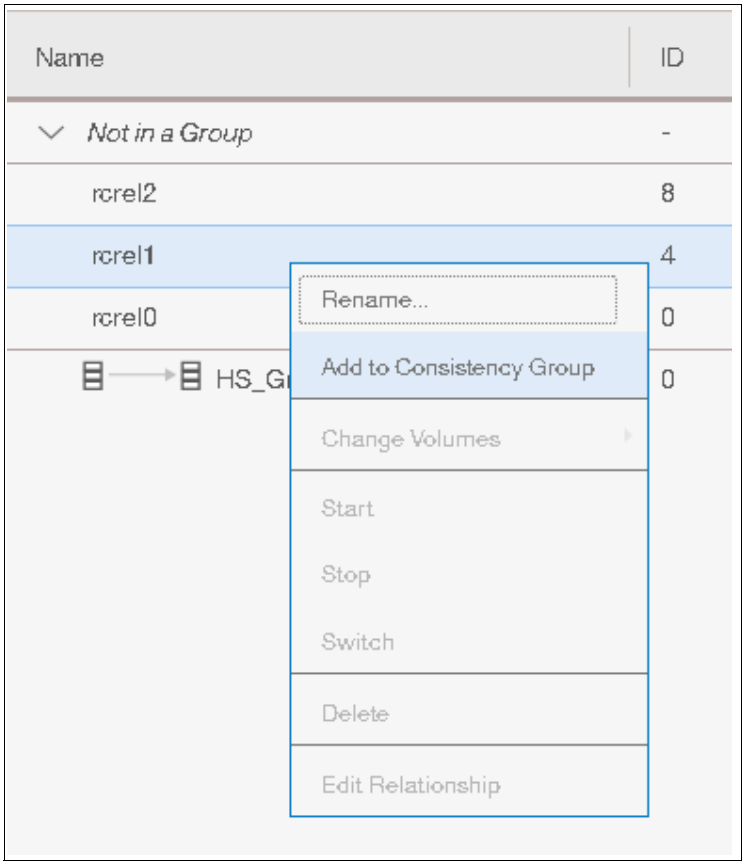


Figure 6-31 Adding remote copy relationships to a consistency group.

Figure 6-32 shows selection of the consistency group for the selected remote copy relationships.

Figure 6-32 Selection the consistency group

Figure 6-33 shows that the master VDisk is the primary VDisk. The consistency group contains the relationships of the two volumes *HS_Vol_1* and *HS_Vol_2*.

HS_Group_AIX	0	Consistent Synchronized	*HURSVT7*	→ *HURSVT7*
rcrel1	4	Consistent Synchronized	HS_Vol_2(site1)	HS_Vol_2(site2)
rcrel0	0	Consistent Synchronized	HS_Vol_1(site1)	HS_Vol_1(site2)

Figure 6-33 HyperSwap normal copy direction

Figure 6-34 shows that the auxiliary VDisk is now the primary VDisk.

HS_Group_AIX	0	Consistent Synchronized	*HURSVT7*	← *HURSVT7*
rcrel1	4	Consistent Synchronized	HS_Vol_2(site1)	HS_Vol_2(site2)
rcrel0	0	Consistent Synchronized	HS_Vol_1(site1)	HS_Vol_1(site2)

Figure 6-34 HyperSwap reversed copy direction

You see the changed direction of the arrow, indicating the reversed copy direction. Next to the state, a sign is added to show that master and auxiliary VDisks are now switched. Hovering over this sign displays a message.

Note: Most of the I/Os are currently a comparison of number of sectors written to rather than a count of I/Os. A 75% majority is required to switch to prevent frequent alternating of direction.

VMware systems can share data stores between multiple virtualized hosts using a single HyperSwap volume. To minimize cross-site I/O traffic, make sure that a data store is only used for virtual machines primarily running on a single site, as this enables HyperSwap to orient the replication optimally.

6.6.1 Site failure

Normally, the storage and control enclosures on both sites are online, and both copies of every HyperSwap volume, the master and auxiliary VDisk of every active-active relationship, contain up-to-date data. If a site fails so that IBM FlashSystem V9000 control enclosures, storage, Fibre Channel connectivity, or a combination is unavailable through hardware failure, power failure, or site inaccessibility, HyperSwap preserves access to VDisks through the remaining site.

A fully synchronized HyperSwap volume has the active-active relationship with the state `consistent_synchronized`. If the storage or control enclosures for the VDisk on one site of a fully synchronized HyperSwap volume goes offline, the following changes occur:

- The state of the active-active relationship becomes `consistent_copying`.
- Host I/O pauses for less than a second in a normal case (this can extend to multiple seconds in some cases, particularly with larger consistency groups).
- If the offline VDisk was the primary copy, the direction of the relationship switches to make the online copy the primary.
- The progress value of the active-active relationship counts down from 100% as the copies become more different (for example, if 10% of the HyperSwap volume was modified while one copy was offline, the progress value shows 90%).
- The master VDisk remains online, and the auxiliary VDisk remains offline when viewed through the **`lsvdisk`** command, regardless of which copy is no longer accessible.

Figure 6-35 show the remote copy relationship in a configuration when the *master* VDisk is the *primary* VDisk.

Name	ID	State	Master Volume	Auxiliary Volume
Not in a Group	-			
rcrel2	9	Consistent Synchronized	HS_Vol_3(site1)	HS_Vol_3(site2)
rcrel1	4	Consistent Synchronized	HS_Vol_2(site1)	HS_Vol_2(site2)
rcrel0	0	Consistent Synchronized	HS_Vol_1(site1)	HS_Vol_1(site2)

Figure 6-35 Remote copy relationships

Example 6-13 shows the same information using the CLI.

Example 6-13 Example of relationship changes after a fail over to the remote site

```
# relationship before fail over to the remote site
>lsrcrelationship rcrel0
name rcrel0
primary master
...
state consistent_synchronized
...
progress
freeze_time
status online

...
```

Example 6-14 on page 164 shows the remote copy relationship after a failover to the remote site. Here the *auxiliary* VDisk is the *primary* VDisk.

The differences between a running relationship shown in Example 6-13 on page 163 and, after a failover, to the remote site are highlighted in Example 6-14.

Example 6-14 Remote copy relationship after fail over to the remote site

```
# relationship after fail over to the remote site
>lsrcrelationship rcrel0
name rcrel0
primary aux
...
state consistent_copying
...
progress 38
freeze_time 2017/11/23/14/48/34
status secondary_change_offline
...
```

Figure 6-36 shows the reversed remote copy relationship after the site failure.

Name	ID	State	Master Volume	Auxiliary Volume
Not in a Group	-			
rcrel0	0	Consistent Copying	HS_Vol_1(site1)	HS_Vol_1(site2)
rcrel1	4	Consistent Copying	HS_Vol_2(site1)	HS_Vol_2(site2)
rcrel2	9	Consistent Copying	HS_Vol_3(site1)	HS_Vol_3(site2)

Figure 6-36 Remote copy reversed relationship

The failover in this example is due to an offline state of the control enclosures on site 1, so the VDisk on site 1 was offline. Figure 6-37 shows the offline information in the volume view.

HS_Vol_1	0	Copy Offline		Multiple
HS_Vol_1 (site1)	0	Offline	Yes	Pool_AE3_1
HS_Vol_1 (site2)	1	Online	Yes	Pool_AE3_2
HS_Vol_2	4	Copy Offline		Multiple
HS_Vol_2 (site1)	4	Offline	Yes	Pool_AE2_1
HS_Vol_2 (site2)	5	Online	Yes	Pool_AE2_2

Figure 6-37 Site failure offline information

Table 6-6 shows the differences between the two **lsrcrelationship** commands. The *master* VDisk is offline and you can see by it looking at the *primary*, and *status* information. The *status* is `secondary_offline` and, because the *auxiliary* VDisk is *primary*, the *secondary* offline VDisk is the *master* VDisk.

Table 6-6 The lsrcrelationship changes

State	Before	After
primary	master	aux
state	consistent_synchronized	consistent_copying
progress	<null>	38

State	Before	After
freeze_time	<null>	2016/10/12/21/23/08
status	online	secondary_change_offline

When that offline copy is restored, the progress value counts back up to 100 as the HyperSwap volume is resynchronized. When it has been resynchronized, the state of the active-active relationship becomes `consistent_synchronized` again. No manual actions are required to make this process occur.

Example 6-15 shows the relationship status after the *master* VDisk is online again. After the resynchronization completes (depending on the amount of data to be replicated, and new data coming in), the status is identical to the status before the *master* VDisk went offline. Only the lines different from the offline master are shown.

Example 6-15 Relationship status after volume is online and synchronized again

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship rcrel0
primary master
state consistent_synchronized
progress
freeze_time
status online
```

During resynchronization, you see the FlashCopy mapping progress filled (Figure 6-38).

Figure 6-38 Flash Copy for volume HS_Vol_1 during resynchronisation

For HyperSwap volumes in consistency groups, a single HyperSwap volume with an inaccessible copy will cause this error recovery procedure to take place on every HyperSwap volume in the consistency group. Each active-active relationship becomes `consistent_copying`. This ensures the consistency of the VDisks on the auxiliary site. If one HyperSwap volume in a group already has one copy offline, and then a different HyperSwap volume VDisks in the same group goes offline, you see two different scenarios.

The `rcrel0` and `rcrel1` are in the same consistency group. Here is the current state of the `HS_Vol_1_rel` HyperSwap volume: The *master* VDisk went offline, then the *auxiliary* VDisk became primary for both HyperSwap volumes (Example 6-16).

Example 6-16 The lsrcrelationship command to see current state of volumes

```
>lsrcrelationship rcrel0
primary aux
state consistent_copying
status secondary_offline
>lsrcrelationship rcrel1
primary aux
state consistent_copying
status online
```

Now, one VDisk of the `rcrel1` HyperSwap volume goes offline:

1. The offline VDisk is on the auxiliary site, the same site with the already-offline VDisk. In Example 6-17, the master VDisk, which is on the secondary site went offline. The HyperSwap volume is still online and accessible from the host.

Example 6-17 Master VDisk offline

```
lsrcrelationship rcrel0
primary aux
state consistent_copying
status secondary_offline
lsrcrelationship rcrel1
primary aux
state consistent_copying
status secondary_offline
```

2. The offline VDisk is on the primary site, the site currently used for I/O. In Example 6-18, the auxiliary VDisk, which is on the current primary site went offline. The HyperSwap volume is now offline and not accessible from the host.

Example 6-18 Auxiliary VDisk offline

```
lsrcrelationship rcrel0
primary aux
state consistent_copying
status secondary_offline
lsrcrelationship rcrel1
primary aux
state consistent_copying
status primary_offline
```

HyperSwap was not able to hide the offline VDisk on the primary site of the `HS_Vol_2_rel` relationship. That HyperSwap volume went offline.

6.6.2 Converting a HyperSwap volume to a basic volume

To remove a copy from a HyperSwap volume, use the **rmvolumecopy** command. Example 6-19 shows two HyperSwap volumes and their VDisks and the result of the **rmvolumecopy** command. The copy on site 2, which is I/O group 1, will be removed from the second HyperSwap volume in this example.

Example 6-19 Removing a HyperSwap volume copy

```
lsvdisk
id name        IO_group_id IO_group_name volume_name function
0  HS_Vol_1    0           io_grp0       HS_Vol_1    master
1  vdisk1      1           io_grp1       HS_Vol_1    aux
2  vdisk2      0           io_grp0       HS_Vol_1    master_change
3  vdisk3      1           io_grp1       HS_Vol_1    aux_change
4  HS_Vol_2 0              io_grp0       HS_Vol_2    master
5  vdisk4      1           io_grp1       HS_Vol_2    aux
6  vdisk5      0           io_grp0       HS_Vol_2    master_change
7  vdisk6      1           io_grp1       HS_Vol_2    aux_change

rmvolumecopy -site 2 HS_Vol_2
```

```
lsvdisk
id name        IO_group_id IO_group_name volume_name function
0  HS_Vol_1    0           io_grp0       HS_Vol_1    master
1  vdisk1      1           io_grp1       HS_Vol_1    aux
2  vdisk2      0           io_grp0       HS_Vol_1    master_change
3  vdisk3      1           io_grp1       HS_Vol_1    aux_change
4  HS_Vol_2    0           io_grp0       HS_Vol_2
```

The **rmvolumecopy** command has removed the auxiliary VDisks, the two change volumes, the four flash copy mappings and the HyperSwap remote copy relationship of the volume HS_Vol_2, thus rendering it to a basic volume.

6.6.3 Deleting HyperSwap volumes

To delete a HyperSwap volume that contains data that is no longer required, use the **rmvolume** command or the GUI as though you were deleting a basic volume (Figure 6-39).

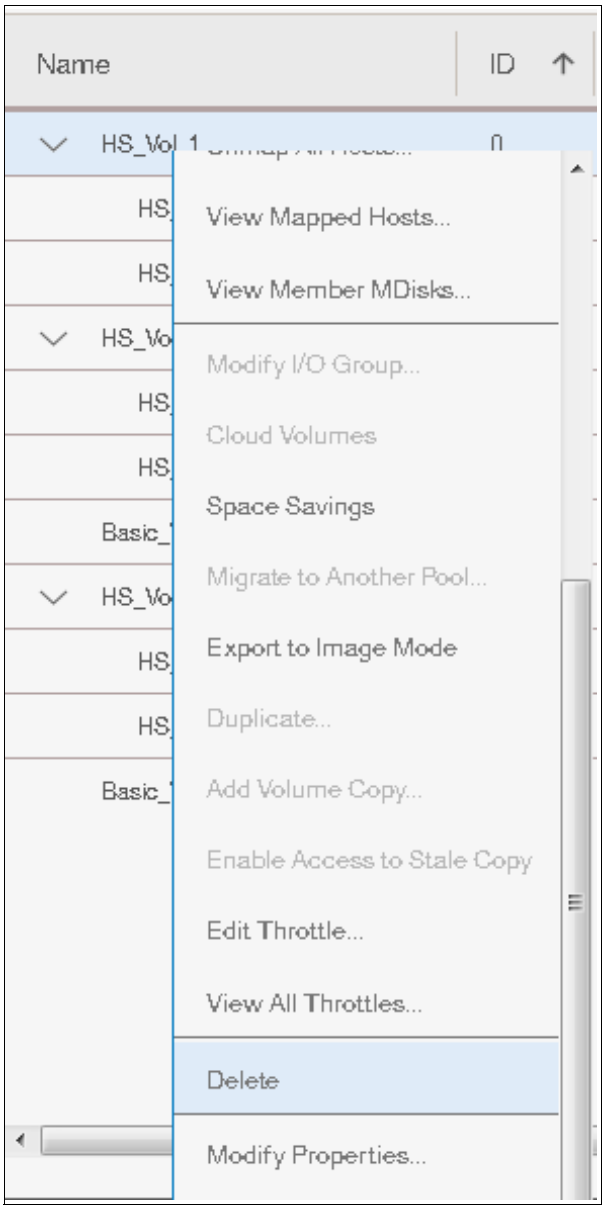


Figure 6-39 Deleting a HyperSwap volume

6.6.4 FlashCopy with HyperSwap volumes

FlashCopy can be used to take point-in-time copies of HyperSwap volumes.

A FlashCopy map with a HyperSwap volume as its source cannot cross sites. Therefore, a FlashCopy mapping where the target VDisk is on site 1 must use the VDisk of the HyperSwap volume on site 1 as its source, and likewise for site 2. It is not possible for a FlashCopy map with a HyperSwap volume as its source to copy data between sites.

For example, if a HyperSwap volume has VDisk 10 providing data on site 1, and VDisk 11 on site 2, FlashCopy maps can be created as follows using the **mkfcmap** command (Example 6-20).

Example 6-20 The mkfcmap command to create FlashCopy maps

```
mkfcmap -source 10 -target 12 ...
mkfcmap -source 11 -target 13 ...
```

In Example 6-20, VDisk 12 is a basic volume already created on site 1, and VDisk 13 on site 2. These two FlashCopy maps can both be used independently to take point-in-time copies of the HyperSwap volume on the two sites. The system provides no coordination of these maps.

When triggering the FlashCopy map, the copy of the HyperSwap volume on the same site as the FlashCopy target VDisk must be either of the following options:

- A *primary* copy of an active-active relationship in any state
- A *secondary* copy of an active-active relationship in a `consistent_synchronized` state

If access has been enabled to an old but consistent copy of the HyperSwap volume, a FlashCopy map can only be triggered on the site that contains that copy. A FlashCopy map cannot be created with a HyperSwap volume as its target. If necessary (for example in a flash-back situation), delete the active-active relationship to convert the HyperSwap volume to a basic volume before creating and triggering the FlashCopy map.

Note: A FlashCopy can only be taken from the VDisks of a HyperSwap volume, not from the HyperSwap volume itself. A FlashCopy cannot be restored on a VDisk of a HyperSwap volume. FlashCopy Manager is currently not supported with HyperSwap Volumes.

6.7 HyperSwap with SAS attached expansion enclosures

IBM FlashSystem V9000 manages internal storage. The internal storage are the optional SAS attached 12F/24F/92F expansion enclosures. Figure 6-40 shows an IBM FlashSystem V9000 setup with one optional storage enclosure on site 1.

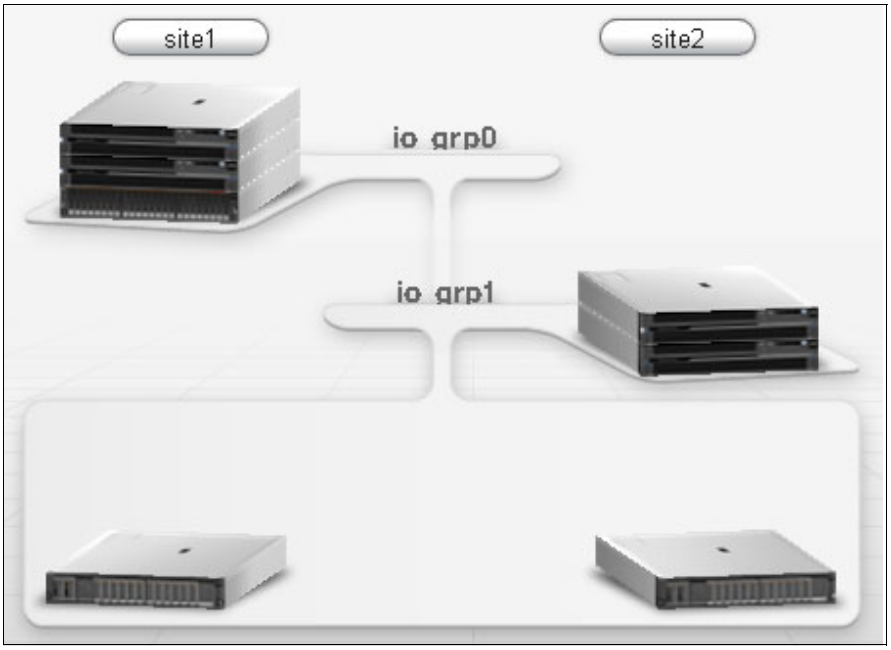


Figure 6-40 FlashSystem V9000 with an optional expansion enclosure connected to io_group0

This configuration uses one storage enclosure and one expansion enclosure on site 1. Site 2 uses one storage enclosure. Figure 6-41 shows the corresponding pools.

Name	State	Capacity
Unassigned MDisks (0)		
Pool_AE2_1	Online	2.00 GiB / 5.62 TiB (0%)
Pool_AE2_2	Online	2.00 GiB / 9.37 TiB (0%)
Pool_AE3_1	Online	3.00 GiB / 39.99 TiB (0%)
Pool_AE3_2	Online	2.00 GiB / 39.99 TiB (0%)
Pool_Local	Online	0 bytes / 1.08 TiB (0%)
mdisk0	Online	1.09 TiB

Figure 6-41 Pools corresponding to the two control enclosures, and one expansion enclosure

You can create a HyperSwap volume by using the HyperSwap volume wizard as described in 6.5.12, “Creating HyperSwap volumes”.

To display the volumes dependent on expansion enclosures, from the GUI select **Monitor** → **System**, right-click the expansion enclosure, and then select **Dependent Volumes** (Figure 6-42).

Figure 6-42 Expansion enclosure information

Figure 6-43 shows the expansion enclosures-dependent volumes.

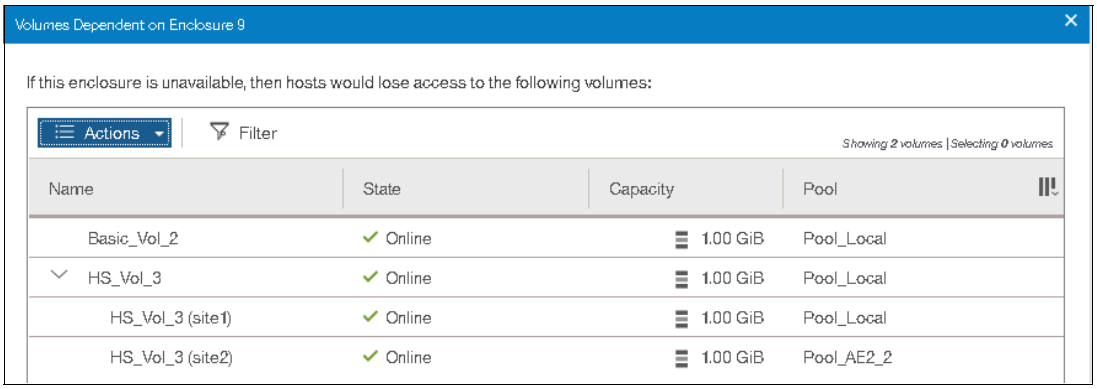


Figure 6-43 Expansion enclosure depending volumes

6.8 Disaster recovery with HyperSwap

The HyperSwap function automatically uses both copies to provide continuous host access to data, providing that both copies are up-to-date. If one copy is up-to-date, and the other is stale, and the up-to-date copy goes offline, the system cannot automatically use the remaining copy to provide high availability to the HyperSwap volume.

However, the user can choose to enable access to that stale copy. This is telling the system to rewind the state of that HyperSwap volume to the point in time of that stale copy.

This rewind to the point in time of that stale copy consists of manual steps, which must be done carefully. Before starting this process, you must make sure that the hosts have not cached data or status of the HyperSwap volumes. Ideally, shut down host systems using the HyperSwap volume before taking these steps. Running these commands without these precautions might crash your applications and corrupt the stale copy.

To demonstrate a stale copy with an up-to-date copy going offline, check active-active relationship of a HyperSwap volume using the **lsrcrelationship** command while the HyperSwap volume is still resynchronizing, after the master had become offline and online again (Example 6-21).

Note: The examples given in this chapter stem from a different system configuration than the examples in previous chapters.

Example 6-21 HyperSwap volume resynchronizing

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship HS_Vol_2_rel
id 6
name HS_Vol_2_rel
master_cluster_id 000002032060460E
master_cluster_name TestCluster
master_vdisk_id 6
master_vdisk_name HS_Vol_2_Mas
aux_cluster_id 000002032060460E
aux_cluster_name TestCluster
aux_vdisk_id 8
aux_vdisk_name HS_Vol_2_Aux
primary aux
consistency_group_id
consistency_group_name
state consistent_copying
bg_copy_priority 50
progress 85
freeze_time 2016/09/29/12/08/31
status online
sync
copy_type activeactive
cycling_mode
cycle_period_seconds 300
master_change_vdisk_id 7
master_change_vdisk_name HS_Vol_2_Mas_CV
aux_change_vdisk_id 9
aux_change_vdisk_name HS_Vol_2_Aux_CV
```

Here, the site of the *master* copy had previously been offline, had returned online, and the HyperSwap volume is resynchronizing. The `consistent_copying` state of the HyperSwap volume shows a resynchronization where the *master* copy contains a stale image, and the value contained in the `freeze_time` field shows when that image dates from. The progress value is increasing toward 100 as the resynchronization process continues.

Now, the site of the *auxiliary* copy goes offline.

Check the active-active relationship using the **lsrcrelationship** command. Only the changes to the previous output are shown in Example 6-22 on page 173. The HyperSwap volume is offline because the primary VDisk went offline during resynchronization.

Example 6-22 The lsrcrelationship command after the site of the auxiliary copy has gone offline

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship HS_Vol_2_rel
...
state consistent_copying
bg_copy_priority 50
progress 87
freeze_time 2016/09/29/12/08/31
status primary_offline
...
```

With the only up-to-date copy of the HyperSwap volume offline, the active-active relationship cannot switch direction to keep the HyperSwap volume online, so the master VDisk is now offline. You see the offline master and auxiliary disk using the **lsvdisk** command (Example 6-23).

Example 6-23 The lsvdisk command shows the master VDisk offline

```
IBM_FlashSystem:TestCluster:superuser>lsvdisk
ID name            IO_group_name status  mdisk_grp_id mdisk_grp_name capacity
6  HS_Vol_2_Mas    io_grp0       offline 0            mdiskgrp_west  35.00GB
7  HS_Vol_2_Mas_CV io_grp0       online  0            mdiskgrp_west  35.00GB
8  HS_Vol_2_Aux    io_grp1       offline 1            mdiskgrp_east  35.00GB
9  HS_Vol_2_Aux_CV io_grp1       online  1            mdiskgrp_east  35.00GB
```

At this point, you look at the `freeze_time` value. If data from that date is not useful, for example it is from too long ago, or before a recent vital update, it might be best to wait until the offline up-to-date copy of the HyperSwap volume can be brought back online.

However, if the stale data is useful, and it is likely that the up-to-date copy of the HyperSwap volume will remain offline for an extended period of time (or will never come online again, for example after a fatal site failure), you can choose to enable access to the stale copy of the HyperSwap volume. Before running this command, make sure that no data or state from this HyperSwap volume is cached on host systems. Stop the active-active relationship using the **stoprcrelationship** command:

```
stoprcrelationship -access <relationship>
```

Check the active-active relationship using the **lsrcrelationship** command. Only the changes to the previous output are shown in Example 6-24. Stopping the relationship will take the HyperSwap volume online using the stale copy. The state of the relationship is `idling`.

Example 6-24 The lsrcrelationship command results after stopping the active-active relationship

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship HS_Vol_2_rel
...
primary master
state idling
bg_copy_priority 50
progress
freeze_time
status
...
```

At this point, the data presented to hosts from this HyperSwap volume immediately changes to that stored on the stale copy. One way to think of this is that the HyperSwap volume has been consistently rolled back to the point in time denoted by the `freeze_time` value.

The HyperSwap volume continues to be readable and writable at this point. You can start your business applications again, and continue from this stale image.

Replication is paused, even if the up-to-date copy becomes online again. This is because the previously stale image, which is now being accessed by hosts, and the previously up-to-date copy, which contains some changes not present on the previously stale image, are now divergent copies. The two copies were the same at the `freeze_time` point in time, but then each had different writes applied. Either copy might be the one that the user wants to keep in the long term.

So the system allows the user to choose which copy is more useful to them. This choice is made based on how much data was missing on the stale copy compared to the up-to-date copy, and how much progress has been made on the stale copy since access was enabled to it.

The first step is determining which copy has the stale copy, which is currently accessible to hosts. This is either the master or auxiliary copy, and is visible under the *primary* attribute of the active-active relationship. You can choose between two scenarios:

- Keep using the copy that hosts are currently accessing, and discard the old up-to-date copy.

  The other, previously up-to-date copy is online again. You decide not to use it, but to keep using the previously stale copy the host is currently accessing. This scenario is described in 6.8.1, “Using the VDisk that the hosts are currently accessing” on page 174.

- Go back to the up-to-date copy and discard the stale copy used for disaster recovery.

  The other, old up-to-date copy is online again, and you decide to discard the changes on this copy and go back to the up-to-date copy. This scenario is described in 6.8.2, “Going back to the up-to-date copy” on page 176.

6.8.1 Using the VDisk that the hosts are currently accessing

This section describes using the stale copy and discarding the old up-to-date copy to start the active-active relationship of the HyperSwap volume.

The disaster recovery using the stale copy was successful and the host is now using that copy. You have decides that this copy, the stale copy, should be used and discarding the up-to-date copy. Use the **`lsrcrelationship`** command to detect the current primary VDisk (Example 6-25).

Example 6-25 The lsrcrelationship command to detect current primary

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship <relationship>
primary master
or
primary aux
```

Use the **`startrcrelationship`** command to start the relationship:

```
startrcrelationship -primary <current_primary> -force <relationship>
```

In this example, <*current_primary*> is the current primary value of the active-active relationship, and is `master` or `aux`. The **`-force`** flag is there because after you make your decision, that loses the ability to use the copy that is not the primary, so it is telling the system that you are aware that this cannot be reverted. In this example, the command is as follows:

```
startrcrelationship -primary master -force HS_Vol_2_rel
```

The host is not affected using this command. There is no need to quiesce host I/O or take any further action. This command resumes HyperSwap replication, and copy across any regions that are different between the two copies to resynchronize as fast as possible. Both copies keep a bitmap of VDisk regions at a 256 KB granularity, used to record writes to that copy that have not yet been replicated to the other copy.

On this resynchronization, both sets of information are used to undo writes applied only to the old up-to-date copy, and also to copy across additional writes made to the stale copy during the disaster recovery. Because the disaster recovery only happened because the copies were resynchronizing before the up-to-date copy went offline, all differences from that interrupted resynchronization process are reverted on the old up-to-date copy now as well.

The active-active relationship goes into an `inconsistent_copying` state, and as copying continues, the progress increases toward 100. At that point, the relationship goes into a `consistent_synchronized` state, showing that both copies are up-to-date, and high-availability is restored.

Use the **`lsrcrelationship`** command to check the status (Example 6-26).

Example 6-26 HyperSwap volume relationship using stale copy, discarding the up-to-date copy

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship HS_Vol_2_rel
ID 6
name HS_Vol_2_rel
master_cluster_id 000002032060460E
master_cluster_name TestCluster
master_vdisk_id 6
master_vdisk_name HS_Vol_2_Mas
aux_cluster_id 000002032060460E
aux_cluster_name TestCluster
aux_vdisk_id 8
aux_vdisk_name HS_Vol_2_Aux
primary master
consistency_group_id
consistency_group_name
state consistent_synchronized
bg_copy_priority 50
progress
freeze_time
status online
sync
copy_type activeactive
cycling_mode
cycle_period_seconds 300
master_change_vdisk_id 7
master_change_vdisk_name HS_Vol_2_Mas_CV
aux_change_vdisk_id 9
aux_change_vdisk_name HS_Vol_2_Aux_CV
```

6.8.2 Going back to the up-to-date copy

This section describes going back to the up-to-date copy and discarding the stale copy to start the active-active relationship of the HyperSwap volume.

The previous section described the steps keeping the stale copy used for disaster recovery used in the long term. It showed how to synchronize to the other copy.

This section describes the scenario when you decide to discard the writings made to the stale copy and go back to up-to-date copy, the one that held the latest data before the disaster recovery.

This scenario is different from the last one, because the image visible by hosts is going to change again. Just as enabling access to the stale copy required hosts to have no cached data from the HyperSwap volume, and ideally they should be fully shut down, the same is true of reverting to the up-to-date copy.

Before going further, make sure that no running hosts are going to be affected by the data changing, and have no stale data that they might corrupt the up-to-date copy with. When applying the **startrcrelationship** command the data visible to hosts instantly reverts to the up-to-date copy. Use the **startrcrelationship** command to start the relationship:

```
startrcrelationship -primary <current_secondary> -force <relationship>
```

In this example, *<current_secondary>* is the copy other than the current primary value of the active-active relationship, and is `master` or `aux`. In other words, if the primary field says `master`, use `aux` here, and vice versa. You cannot get back to the other set of data after you have run this command, and the **-force** flag is there to acknowledge this. In this example, the command is as follows:

```
startrcrelationship -primary aux -force HS_Vol_2_rel
```

The image visible to hosts instantly reverts to the up-to-date copy, so it reverts as soon as you have run this command. You can bring the hosts back online and start using this HyperSwap volume again.

As with the other scenario, the active-active relationship is in an `inconsistent_copying` state while resynchronizing, and again this resynchronization uses the bitmaps of writes to each copy to accelerate this resynchronization process. When the copies are fully synchronized, the relationship goes back to a `consistent_synchronized` state as high availability is restored for the HyperSwap volume.

Use the **lsrcrelationship** command to check the status. The primary VDisk in this example is the *auxiliary* VDisk. While resynchronizing, the state is `inconsistent_copying` until the HyperSwap volume is synchronized (Example 6-27).

Example 6-27 HyperSwap volume relationship using old up-to-date copy and discarding the used stale copy

```
IBM_FlashSystem:TestCluster:superuser>lsrcrelationship HS_Vol_2_rel
id 6
name HS_Vol_2_rel
master_cluster_id 000002032060460E
master_cluster_name TestCluster
master_vdisk_id 6
master_vdisk_name HS_Vol_2_Mas
aux_cluster_id 000002032060460E
aux_cluster_name TestCluster
```

```
aux_vdisk_id 8
aux_vdisk_name HS_Vol_2_Aux
primary aux
consistency_group_id
consistency_group_name
state inconsistent_copying
bg_copy_priority 50
progress 56
freeze_time
status online
sync
copy_type activeactive
cycling_mode
cycle_period_seconds 300
master_change_vdisk_id 7
master_change_vdisk_name HS_Vol_2_Mas_CV
aux_change_vdisk_id 9
aux_change_vdisk_name HS_Vol_2_Aux_CV
```

6.9 Disaster recovery with consistency groups

All the descriptions in 6.7, “HyperSwap with SAS attached expansion enclosures” on page 170 (and after) about enabling access to a stale copy of a HyperSwap volume also apply to HyperSwap consistency groups, for example multiple HyperSwap volumes where the active-active relationships are contained in a single consistency group.

During resynchronization, if any of the up-to-date copies of HyperSwap volumes in a consistency group is offline or unavailable, typically all would be offline in a disaster. You can choose to enable access to the stale copy of every HyperSwap volume in the consistency group. Because the HyperSwap function links replication and failover across HyperSwap volumes in a consistency group, it is assured that during resynchronization, all copies on one site have a stale consistent copy of data, captured at an identical point in time, ideal for disaster recovery.

The **startrcconsistgrp** and **stoprcconsistgrp** commands are the consistency group versions of the **startrcrelationship** and **stoprcrelationship** commands used in section 6.7, “HyperSwap with SAS attached expansion enclosures” on page 170.

The **stoprcconsistgrp** command is used to gain access to the stale copies:

```
stoprcconsistgrp -access <consistency_group>
```

When restarting the consistency group you can either retain the access to the stale copies or revert to the previous up-to-date copies. Use the **lsrcrelationship** command to detect the current primary VDisk (Example 6-28).

Example 6-28 The lsrcrelationship command to detect current primary

```
lsrcrelationship <consistency_group>
primary master
or
primary aux
```

To retain the stale disaster recovery copies currently visible to hosts, and resume HyperSwap replication while discarding the data of the previous up-to-date copies, use the following command:

```
startrcconsistgrp -primary <current_primary> -force <consistency_group>
```

To revert to the previous up-to-date copy and discard the changed data on the stale copies, the following command should be used while the host has no access to the HyperSwap volume, as described in 6.8.2, “Going back to the up-to-date copy” on page 176:

```
startrcconsistgrp -primary <current_secondary> -force <consistency_group>
```

6.10 The overridequorum command

IBM FlashSystem V9000 provides the **overridequorum** command that can be used to override the tie-breaking performed by the system quorum if it left the system in an unusable state.

This command is valid on control enclosures that are in a starting state with either of the following control enclosure errors:

- 551
- 921

You can check for these errors by using the **sainfo lsservicenodes** command. One scenario where using the **overridequorum** command might be useful is if a rolling disaster first breaks the link between the two sites, resulting in the quorum deciding which site’s control enclosures should be allowed to continue. Next, the rolling disaster affects the chosen site’s control enclosures, taking them offline. The entire system is unusable at this point, because of how the tie-break was resolved.

Use the **overridequorum** command on a control enclosure displaying an error code of 551 or 921 on the site you want to start manually:

```
satask overridequorum -force
```

When the **overridequorum** command is issued on a control enclosure displaying a 551 or 921 error, that site’s control enclosures use their cluster state to form a new cluster, with a new cluster ID, based on the system state at the point that the tiebreak stopped that site’s control enclosures from taking part in the cluster.

Other than the new cluster ID, this gives the appearance of reverting the system state to the point in time of that tiebreak, and because the restored control enclosures have system cache and local VDisk copies matching that point in time, the VDisk state is reverted to that point in time as well.

There is no specific interaction between HyperSwap volumes and the **overridequorum** command, if a HyperSwap volume copy local to the site brought online by the **overridequorum** command was up-to-date at the time of the lost tiebreak, it will immediately come online after the **overridequorum** command is run. Alternatively, if the copy was stale at the time of the lost tiebreak, access needs to be enabled to it with the **stoprcrelationship** command.

This command also removes the control enclosures in the other site. This means that HyperSwap volume copies on that site need to be deleted and re-created. More details about unconfiguring HyperSwap are in 6.12.1, “Removing HyperSwap volumes completely” on page 180.

When the HyperSwap volumes are converted to basic volumes on the online site, and the control enclosures in the other site have been re-added to the system, you can then convert the basic volumes back to HyperSwap volumes.

6.11 HyperSwap Failure scenarios

Table 6-7 shows failure scenarios and their effect on HyperSwap, hosts, and applications.

Table 6-7 HyperSwap failure scenarios

Failure scenario	HyperSwap system behavior	Server and application effect
Single switch failure.	System continues to operate by using an alternative path in the same failure.	None.
Slow read or write performance to a copy (giving greater than 30 seconds response time)	System temporarily stops replicating to slow copy, and resynchronizes after 5 minutes.	None.
Single data storage failure.	System continues to operate by using the other data copy.	None.
Single quorum storage failure on site 3.	System continues to operate using alternative storage at site 3.	None.
Failure of either site 1 or 2	System continues to operate on the remaining site.	Servers without high availability (HA) functions in the failed site stop. Servers in the other site continue to operate. Servers with HA software functions are restarted from the HA software. The same disks are seen with the same UIDs in the surviving site, and continue to offer similar read and write performance as before the disaster.
Failure of site 3, containing the active quorum disk	System continues to operate on both sites 1 and 2, selecting a quorum disk from sites 1 and 2 to enable I/O processing to continue.	None.
Access loss between sites 1 and 2	System continues to operate the site that wins the quorum race. The cluster continues with operation, while the control enclosures in the other site stop, waiting for connectivity between sites 1 and 2 to be restored.	Servers without HA functions in the failed site stop. Servers in the other site continue to operate. Servers with HA software functions are restarted from the HA software. The same disks are seen with the same UIDs in the surviving site, and continue to offer similar read and write performance as before the disaster.
Access loss between sites 1 and 2 because of a rolling disaster. One site is down, and the other is still working. Later, the working site also goes down because of the rolling disaster	System continues to operate the site that wins the quorum race. The system continues with operation until the other site goes down. Even if the first site to go down comes back up, the whole system is considered offline until the site that won the quorum race comes back up.	The system can restart using just the site that initially lost the quorum race, by using the **`overridequorum`** command. The HyperSwap volumes revert to the state they were at then that site lost the quorum race. Servers must be stopped before issuing this command, and restarted with the reverted state. Full read and write performance is given.

6.12 Unconfiguring HyperSwap

This section describes the unconfiguring of HyperSwap, including removing the data or keeping the data from the primary or auxiliary site.

6.12.1 Removing HyperSwap volumes completely

If you do not need any data on a HyperSwap volume, and want to delete all objects related to it, use the **rmvolume** command (Example 6-29).

Example 6-29 The rmvdisk command to delete all four disks associated with a HyperSwap volume

```
rmvolume HyperSwap_0
rmvolume HyperSwap_1
```

To use the GUI to delete HyperSwap volumes, right-click the HyperSwap volume, and select **Delete**.

6.12.2 Converting to basic volumes, while retaining access through the master VDisk

If you want to go back to using basic volumes, decide which copy should be retained. Use the **rmvolumecopy** command to delete the VDisk of the site not needed anymore, the two change volumes not needed anymore, and the relationship not needed anymore. Assuming the *master* VDisk is on site 1 and the *auxiliary* disk is on site 2, you can use the **-site** parameter of the **rmvolumecopy** command (Example 6-30).

Example 6-30 The rmvolumecopy command to delete the auxiliary VDisk on site 2, the two change volumes, and the relationship

```
# check the current HyperSwap vdisks using lsvdisk
lsvdisk
id name   IO_group_id IO_group_name status   function
0  AIX_0  0           io_grp0       online   master
1  vdisk1 1           io_grp1       offline  aux
2  vdisk2 0           io_grp0       online   master_change
3  vdisk3 1           io_grp1       online   aux_change
# remove the auxiliary copy
rmvolumecopy -site site2 -removefcmaps AIX_0
# check the changed HyperSwap vdisks using lsvdisk
lsvdisk
id name   IO_group_id IO_group_name status   function
0  AIX_0  1           io_grp0       online
```

The I/O group `io_grp0` is on site 1. Therefore, the remaining volume is on site 1.

6.12.3 Converting to basic volumes, while retaining access through the auxiliary VDisk

Use the **rmvolumecopy** command to delete the VDisk of the site not needed anymore. The previous section (6.12.2, “Converting to basic volumes, while retaining access through the master VDisk” on page 180) describes deleting the auxiliary VDisk on site 2. This section shows the deletion of the *master* disk in site 1.

Assuming the *master* VDisk is on site 1 and the *auxiliary* disk is on site 2, you can use the **-site** parameter of the **rmvolumecopy** command (Example 6-31).

Example 6-31 The rmvolumecopy command to delete the master VDisk on site 1, the two change volumes, and the relationship

```
# check the current HyperSwap vdisks using lsvdisk
lsvdisk
id name   IO_group_id IO_group_name status   function
0  AIX_0  0           io_grp0       online   master
1  vdisk1 1           io_grp1       offline  aux
2  vdisk2 0           io_grp0       online   master_change
3  vdisk3 1           io_grp1       online   aux_change
# remove the auxiliary copy
rmvolumecopy -site site1 -removefcmaps AIX_0
# check the changed HyperSwap vdisks using lsvdisk
lsvdisk
id name   IO_group_id IO_group_name status   function
0  AIX_0  1           io_grp1       online
```

The I/O group `io_grp1` is on site 2. Therefore, the remaining volume is on site 2.

6.12.4 Converting to system topology standard

After all active-active relationships are deleted, IBM FlashSystem V9000 topology can be changed to `standard`. Aspects of the system locked down in the HyperSwap system topology, for example control enclosure and controller sites, can then be changed. The topology of the system can be reverted to the `standard` topology using the **chsystem** command:

```
chsystem –topology standard
```

You can use the GUI to modify IBM FlashSystem V9000 topology (6.5.10, “HyperSwap configuration using the GUI wizard” on page 150). The wizard now lists two prerequisites (Figure 6-44).

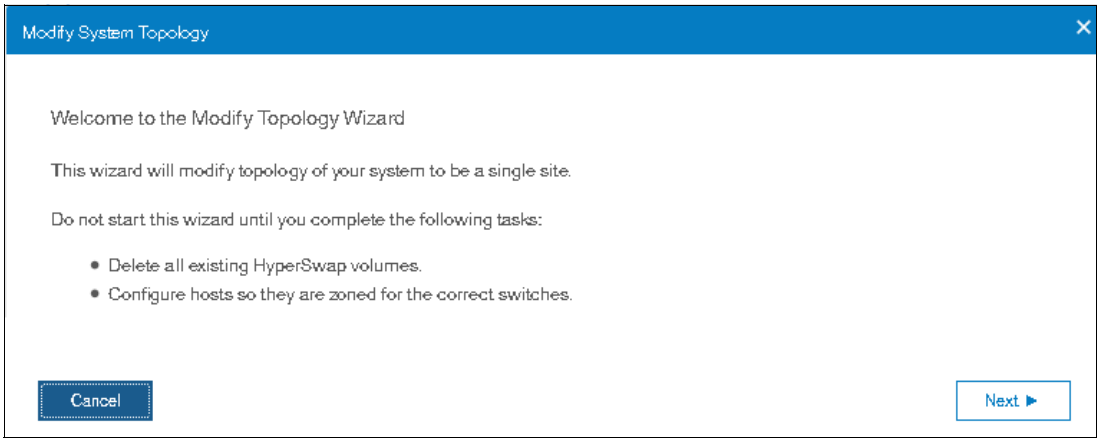


Figure 6-44 Topology wizard when switching to standard topology

Review the Summary page (Figure 6-45 on page 182) and click **Finish**.

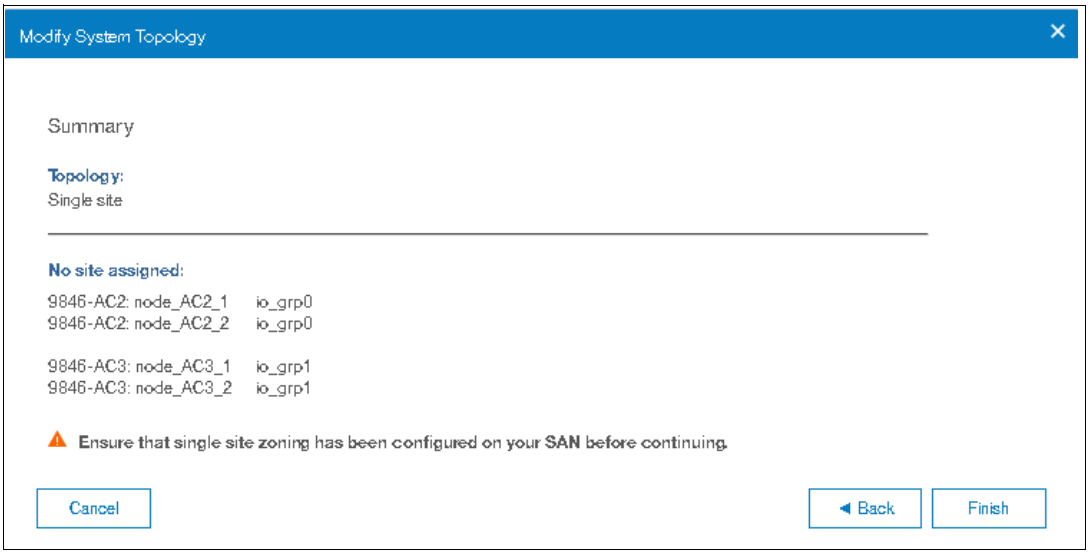


Figure 6-45 HyperSwap wizard summary when switching to standard topology

The wizard will change all controller control enclosure site attributes to `nosite`. You can check the status with the **`lssystem`** command (Example 6-32).

Example 6-32 The lssystem command to check system topology

```
>lssystem | grep topo
topology standard
topology_status
```

6.13 Summary of interesting object states for HyperSwap

This section describes the state values of various commands.

6.13.1 The lsvdisk command

The **`status`** attribute shown for HyperSwap volume master VDisk in **`lsvdisk`** shows whether hosts are able to access data, for example whether the HyperSwap volume has access to up-to-date data or not, not whether the master VDisk itself is actually online. The value `status` for auxiliary VDisk is always offline.

Running **`lsvdisk`** on a specific VDisk to get detailed information also shows the VDisk copy status value. The **`RC_id`** and **`RC_name`** attributes for both master and auxiliary VDisks show the active-active relationship supporting the HyperSwap volume (Example 6-33).

Example 6-33 The lsvdisk command on a specific disk to see VDisk copy status

```
lsvdisk HS_Vol_1_Mas
...
RC_id 0
RC_name HS_Vol_1_rel
RC_change no

...
```

6.13.2 The lsvdiskcopy command

The **status** attribute for VDisk copies supporting HyperSwap volumes shows whether the underlying storage is online (Example 6-34).

Example 6-34 The lsvdiskcopy command shows if underlying storage is online

```
lsvdiskcopy HS_Vol_1_Mas
vdisk_id vdisk_name   copy_id status sync primary mdisk_grp_id mdisk_grp_name
0        HS_Vol_1_Mas 0       online yes  yes     0            mdiskgrp_west
```

6.13.3 The lsrcrelationship or lsrcconsistgrp commands

These values are for stand-alone relationships seen with **lsrcrelationship**. Relationships in a consistency group must all share the same `state`, `primary`, and `freeze_time` field values, so they change value based on the condition of all the relationships in that consistency group. The consistency group itself shows the same values when queried using **lsrcconsistgrp**. The **status** attribute is the key attribute that tells you the HyperSwap volume copying status:

- `inconsistent_stopped`

  This HyperSwap volume only has useful data on the master VDisk, and the relationship's change volumes are not both configured yet.

- `consistent_stopped`

  This HyperSwap volume only has useful data on the master VDisk, the relationship's change volumes are not both configured yet, but the relationship was created with `-sync`, limiting needed copying to only the data written to on the master VDisk since the active-active relationship was created.

- `inconsistent_copying`

  This HyperSwap volume only has useful data on the master VDisk, but it is correctly configured and is performing initial synchronization.

- `consistent_synchronized`

  This HyperSwap volume is correctly configured, has up-to-date data on both VDisks, and is highly available to hosts, if **addvdiskaccess** has been run correctly.

- `consistent_copying`

  This HyperSwap volume is correctly configured, has had or is currently having a period of inaccessibility of one VDisk, leaving that VDisk consistent but stale (the `freeze_time` attribute will show when that stale data dates from). Access to that data can be provided in a disaster with the **stoprcrelationship -access** command, and resynchronization automatically will take place when possible.

- `idling`

  This HyperSwap volume is correctly configured, and has had access enabled to a stale but consistent copy by running **stoprcrelationship -access** when the active-active relationship was in a state of `consistent_copying`. Synchronization is paused, and can be resumed by running **startrcrelationship -primary (master | aux)** according to the direction that the relationship should resynchronize in.

The **primary** attribute will be `master` or `auxiliary`, and tells you which copy is acting as the primary at the moment, and therefore which I/O group is primarily processing I/Os for this HyperSwap volume.

The **status** attribute shows `online` if all needed VDisks are online and are able to synchronize. Otherwise, it shows the reason why synchronization is not possible:

- `primary_offline`
- `secondary_offline`
- `primary_change_offline`
- `secondary_change_offline`

It shows one of the previous statuses if one of the VDisks of the HyperSwap volume is offline, or `change_volumes_needed` if the HyperSwap volume does not have both change volumes configured.

The **progress** attribute shows how similar the two copies are as a percentage, rounded down to the nearest percent. During resynchronization, this counts up to 100 as the HyperSwap volume nears being synchronized.

The **freeze_time** attribute shows at what point the data is frozen on a stale but consistent copy when the relationship has a state of `consistent_copying`. This enables the user to decide if there is value in using the data (with the **stoprcrelationship -access** command) if the up-to-date copy goes offline.

6.13.4 The lsfcmap command

The **status** attribute for FlashCopy mappings used by a HyperSwap volume shows if a FlashCopy mapping is currently used, for example during a resynchronization of VDisks of a HyperSwap volume after a VDisk failure.

Example 6-35 shows a FlashCopy mapping currently used during resynchronization.

Example 6-35 FlashCopy mapping currently used during resynchronization

```
lsfcmap fcmap1
...
name fcmap1
source_vdisk_name HS_Vol_1_Mas
target_vdisk_name HS_Vol_1_Mas_CV
status copying
progress 42
start_time 150804045444
...
```

6.14 Naming conventions

The reference *volume* can refer to a VDisk or to a HyperSwap volume depending the context.

In standard topology the reference *volume* is equivalent to the reference *VDisk*. A *VDisk* can be created by the *mkvdisk* command or the *mkvolume* command in standard topology.

In HyperSwap topology the reference *volume* is ambiguous. It should be specified as *basic volume* or *HyperSwap volume*:

- *Basic volume*: A volume at one site. This can be a VDisk created by the **mkvdisk** command or a VDisk with an mirrored copy. A basic volume is always only on one site.
- *HyperSwap volume*: A volume with two copies on different sites in an active-active relationship. The HyperSwap volume has one basic volume on each site.

7

Hints and tips

FlashSystem V9000 has exceptional capabilities for customers to address data security, redundancy, and application integration. It uses industry-leading IBM support infrastructure, including the IBM Compresstimator utility. This chapter provides helpful hints and tips to use these capabilities in productive ways.

This chapter includes the following topics:

- Performance data and statistics gathering
- Estimating compression savings
- Command-line hints
- Call Home process
- Service support

7.1 Performance data and statistics gathering

This section provides an overview of the performance analysis capabilities of the IBM FlashSystem V9000, and a method for collecting and processing performance statistics. For a more in-depth understanding of performance statistics and interpretation, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

Basically, IBM FlashSystem V9000 does not differ from IBM SAN Volume Controller with IBM FlashSystem 900. IBM FlashSystem V9000 high IOPS and low latency often require host tuning to realize its performance capabilities. For more details regarding IBM FlashSystem 900 running with IBM SAN Volume Controller, see the chapter about product integration in *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414 and *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

7.1.1 IBM FlashSystem V9000 controller performance overview

The caching capability of the IBM FlashSystem V9000 controller and its ability to effectively manage multiple FlashSystem enclosures along with standard disk arrays can provide a significant performance improvement over what can otherwise be achieved when disk subsystems alone are used. To ensure that the wanted performance levels of your system are maintained, monitor performance periodically to provide visibility to potential problems that exist or are developing so that they can be addressed in a timely manner.

Performance considerations

When you are designing an IBM FlashSystem V9000 storage infrastructure or maintaining an existing infrastructure, you must consider many factors in terms of their potential effect on performance. These factors include, but are not limited to, mixed workloads competing for the same resources, overloaded resources, insufficient resources available, poorly performing resources, and similar performance constraints, especially when using external storage.

Remember these rules as you design your SAN and IBM FlashSystem V9000 layout:

- Host to IBM FlashSystem V9000 controller interlink (ISL) oversubscription

  This area is the most significant I/O load across ISLs. The suggestion is to maintain a maximum ratio of 7:1 oversubscription. A higher ratio is possible, but it tends to lead to I/O bottlenecks. This suggestion also assumes a core-edge design, where the hosts are on the edge and the IBM FlashSystem V9000 controller is on the core.

- Storage to IBM FlashSystem V9000 controller ISL oversubscription

  IBM FlashSystem V9000 scale-up scale-out configurations suggest dedicated FC switches (for example SAN48B-5) to support the IBM FlashSystem V9000 controllers and enclosures. Although any supported switch can be used, be careful to avoid bottlenecks.

- Node-to-node ISL oversubscription

  IBM FlashSystem V9000 guidelines do not allow for node-to-node ISL oversubscription.

  This area is the least significant load of the three possible *oversubscription* bottlenecks. In standard setups, this load can be ignored. Although it is not entirely negligible, it does not contribute significantly to ISL load. However, it is mentioned here regarding the split-cluster capability that was made available with SAN Volume Controller technology.

Note: IBM FlashSystem V9000 does not support stretched cluster topology.

- ISL trunking or port channeling

  For the best performance and availability, it is suggested that you use ISL trunking or port channeling. Independent ISL links can easily become overloaded and turn into performance bottlenecks. Bonded or trunked ISLs automatically share load and provide better redundancy in the case of a failure.

- Number of paths per host multipath device

  The maximum supported number of paths per multipath device that is visible on the host is eight. Although the Subsystem Device Driver Path Control Module (SDDPCM), related products, and most vendor multipathing software can support more paths, the V9000 controller expects a maximum of eight paths. In general, you see only a negative effect on performance from more paths than eight. Although the SAN Volume Controller can work with more than eight paths, this design is technically unsupported.

- Do not intermix dissimilar array types or sizes

  Although the IBM FlashSystem V9000 controller supports an intermix of differing storage within storage pools, the best approach is to always use the same array model, RAID mode, RAID size (RAID 5 6+P+S does not mix well with RAID 6 10+P+Q+S), and drive speeds. Given current drive technology, our recommendation for any drive with a capacity larger than 1 TB is to use distributed RAID 6 (DRAID 6) anyway. Mixing standard storage with FlashSystem volumes is not advised unless the intent is to use Easy Tier.

Rules and guidelines are no substitution for monitoring performance. Monitoring performance can provide a validation that design expectations are met and identify opportunities for improvement.

IBM FlashSystem V9000 performance perspectives

IBM FlashSystem V9000 controller consists of software and hardware. The software was developed by the IBM Research Group for IBM Spectrum Virtualize, which delivers the function of SAN Volume Controller and was designed to run on commodity hardware (mass-produced Intel-based CPUs with mass-produced expansion cards), while providing distributed cache and a scalable cluster architecture.

One of the main advantages of this design is the capability to easily refresh hardware. Currently, the IBM FlashSystem V9000 controller is scalable up to four building blocks (eight controllers), and these controllers can be swapped for newer hardware while online. This capability provides a great investment value because the controllers are relatively inexpensive. This capability also provides an instant performance boost with no license changes. Newer controllers can dramatically increase cache per controller, providing an extra benefit on top of the typical refresh cycle.

For more information about the controller replacement and swap, and instructions about adding nodes, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Alternatively you might want to consider the feature called *hot spare node*, which provides a fully automated way to have an extra node in your cluster take over in case an active node should fail or goes offline for maintenance.

Note: To add a hot spare node, the existing enclosures must be configured as one or more scalable building blocks, with the enclosures connected to each other using Fibre Channel (FC) switches.

Important: You cannot add a spare node to a fixed building block with direct connections between the enclosures.

The performance is near linear when controllers are added into the cluster until performance eventually becomes limited by the attached components. This scalability is significantly enhanced using FlashCore technology included with the storage enclosures in each building block.

FlashCore technology is built on three core principles:

- Hardware accelerated I/O
- IBM MicroLatency module
- Advanced flash management.

Partnership with Micron and FlashSystem development teams help to ensure system reliability and optimization for flash. The design goals for IBM FlashSystem V9000 are to provide the customer with the fastest and most reliable all-flash memory arrays on the market, while making it simple to service and support.

Virtualization with the IBM FlashSystem V9000 controller building block design provides specific guidance in terms of the components that are used, so that it can deliver optimal performance. The key item for planning is your SAN layout.

Switch vendors have slightly different planning requirements, but the goal is that you always want to maximize the bandwidth that is available to the IBM FlashSystem V9000 controller ports. The IBM FlashSystem V9000 controller is one of the few devices that can drive ports to their limits on average, so be sure that you put significant thought into planning the SAN layout.

Figure 7-1 on page 189 shows the overall environment with two SAN fabrics:

- Dedicated SAN Switch Fabric for building block communications
- SAN Switch fabric with host zone and an optional storage zone for external storage

Tip: In a production environment, a preferred practice is to use redundant SAN fabrics, which is not shown here.

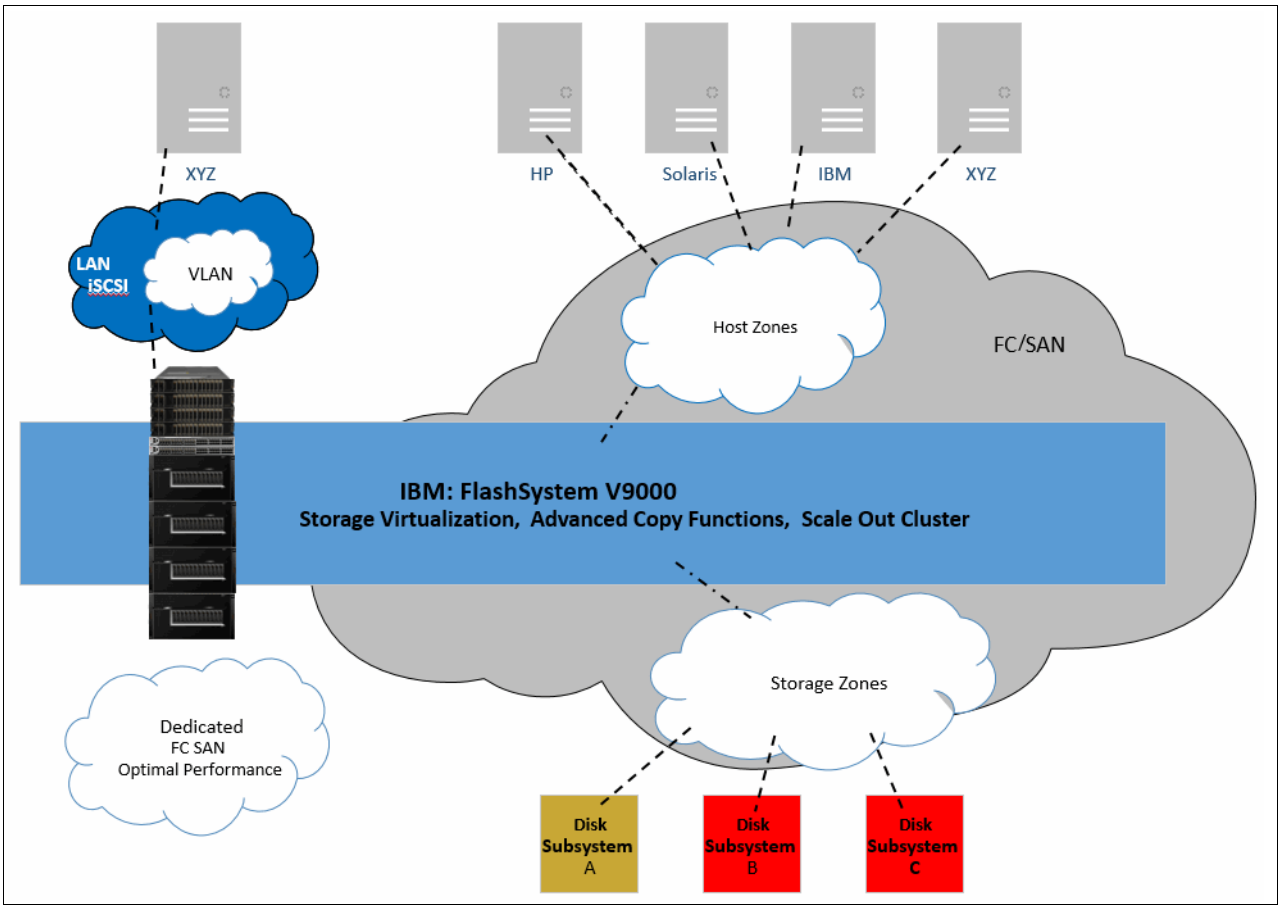


Figure 7-1 IBM FlashSystem V9000 scale up and scale out

A dedicated SAN is suggested but not required; the objective is to *not* introduce external latency between IBM FlashSystem V9000 controllers and their IBM FlashSystem storage enclosures. This can be accomplished through other SAN administration techniques.

Essentially, IBM FlashSystem V9000 controller performance improvements are gained by optimizing delivery of FlashCore technology storage enclosure resources and with advanced functionality that is provided by the IBM FlashSystem V9000 controller cluster. However, the performance of individual resources to hosts on the SAN eventually becomes the limiting factor.

IBM FlashSystem V9000 deployment options

IBM FlashSystem V9000 brings high capacity and integrated management to the enterprise data center. IBM FlashSystem V9000 delivers up to 180 TB usable (219 TB maximum effective) per building block, scales to four building blocks, and offers up to four additional IBM FlashSystem V9000 storage enclosure expansion units for large-scale enterprise storage system capability. External storage and internal SAS expansion enclosures can also be added to increase overall capacity.

IBM FlashSystem V9000 has the following flexible scalability configuration options:

- Base configuration
- Scale up: Add capacity
- Scale out: Add controllers and capacity

For more details, see Chapter 3, “Scalability” on page 61.

The following topics illustrate the performance benefits of two-dimensional scaling in various environments. By offering two-dimensional scaling, the IBM FlashSystem V9000 provides scalable performance that is difficult to surpass. Figure 7-2 shows examples of the maximum performance that can be achieved.

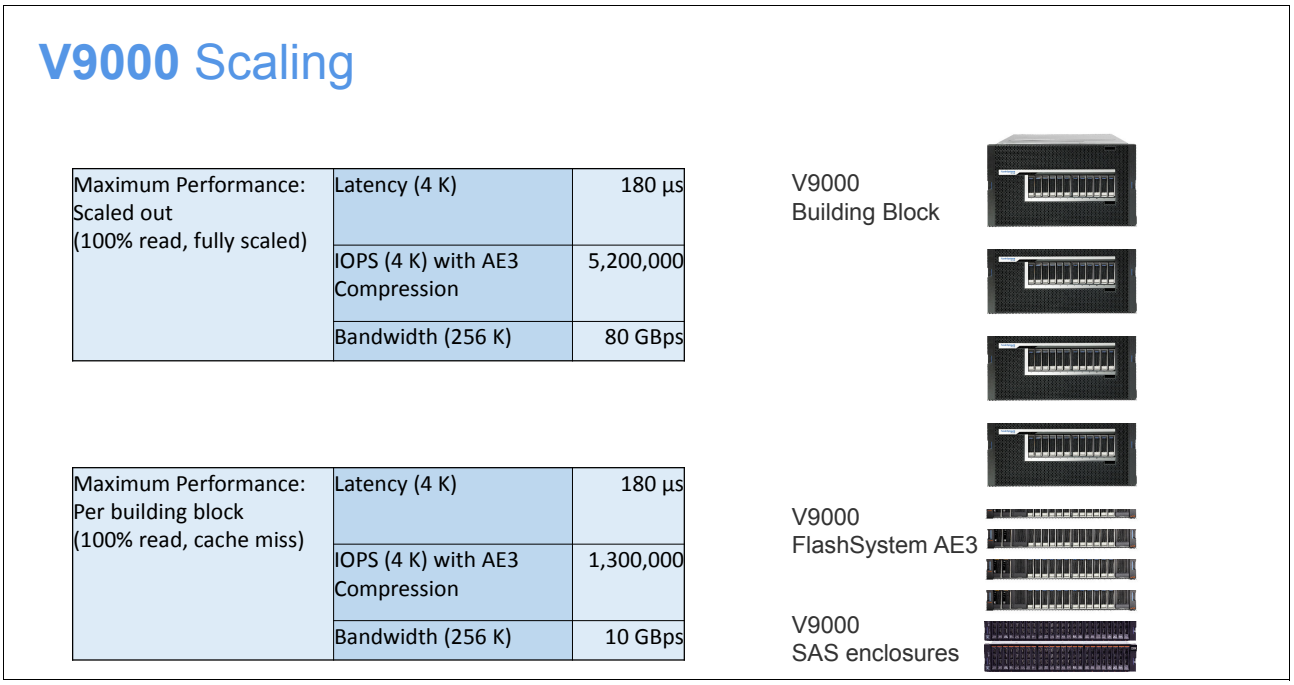


Figure 7-2 Two-dimensional scaling

For more details, see *FlashSystem V9000 Product Guide*, REDP-5468.

7.1.2 Performance monitoring

This section highlights several performance monitoring techniques.

Note: IBM FlashSystem V9000 AE3 *storage enclosure* performance statistics are not included in this section. In most situations, performance measurement data collected by the V9000 *controller enclosure* is the most relevant because it includes data for the entire data path. However, in situations in which you want to isolate the performance of the FlashSystem storage enclosure mdisk, the AE3 performance GUI is more appropriate.

For details about the AE3 storage enclosure performance GUI, See *Monitoring performance menu* in Implementing IBM FlashSystem 900 Model AE3, SG24-8414.

For details about controller enclosure performance for IBM SAN Volume Controller, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Collecting performance statistics

The IBM FlashSystem V9000 components are constantly collecting performance statistics. The default frequency by which files are created is at 5-minute intervals with a supported range of 1 - 60 minutes.

Tip: The collection interval can be changed by using the **`svctask startstats`** command.

The statistics files (VDisk, MDisk, and Node) are saved at the end of the sampling interval and a maximum of 16 files (each) are stored before they are overlaid in a rotating log fashion. This design provides statistics for the most recent 80-minute period if the default 5-minute sampling interval is used.

For more information, see the **startstats** command in IBM Knowledge Center.

Collection intervals: Although more frequent collection intervals provide a more detailed view of what happens within the IBM FlashSystem V9000, they shorten the amount of time that the historical data is available on the IBM FlashSystem V9000 Controller. For example, rather than an 80-minute period of data with the default 5-minute interval, if you adjust to 2-minute intervals, you have a 32-minute period instead.

The IBM FlashSystem V9000 does not collect cluster-level statistics. Instead, you use the per-node statistics that are collected. The sampling of the internal performance counters is coordinated across both nodes of the IBM FlashSystem V9000 controller cluster so that when a sample is taken, all nodes sample their internal counters at the same time.

An important step is to collect all files from all nodes for a complete analysis. Tools, such as IBM Spectrum Control™, can perform this intensive data collection for you. IBM Spectrum Control also makes sure not to miss files, so even with short intervals, its recurring collection will give you historical data at fine granularity. (You need to provide enough space to IBM Spectrum Control to store all this data.)

Note: Starting with IBM Tivoli® Storage Productivity Center (IBM Spectrum Control's former name) version 5.2.6, IBM FlashSystem V9000 is supported. See the current support matrix for more detail.

Statistics file naming

The files that are generated are written to the `/dumps/iostats/` directory. The file name is in the following formats:

- For MDisk statistics:

  `Nm_stats_<node_serial_number>_<date>_<time>`

- For VDisk / volume statistics:

  `Nv_stats_<node_serial_number>_<date>_<time>`

- For node statistics:

  `Nn_stats_<node_serial_number>_<date>_<time>`

- For disk drive statistics (when available through expansion enclosures):

  `Nd_stats_<node_serial_number>_<date>_<time>`

The `node_node_serial_number` is of the node on which the statistics were collected. The date is in the form `<yymmdd>` and the time is in the form `<hhmmss>`. The following example shows an MDisk statistics file name:

`Nm_stats_75AM710_150323_075241`

Example 7-1 shows the typical MDisk, volume, node, and disk drive statistics file names.

Example 7-1 File names of per node statistics

```
IBM_FlashSystem:ITSO_V9000:superuser>svcinfo lsiostatsdumps
id iostat_filename
0  Nv_stats_75AM710_150323_164316
1  Nm_stats_75AM710_150323_164316
2  Nd_stats_75AM710_150323_164316
3  Nn_stats_75AM710_150323_164316
4  Nm_stats_75AM730_150323_164316
5  Nv_stats_75AM730_150323_164316
6  Nd_stats_75AM730_150323_164316
7  Nn_stats_75AM730_150323_164316
```

Tip: The performance statistics files can be copied from the IBM FlashSystem V9000 Controllers to a local drive on your workstation by using the **pscp.exe** (included with PuTTY) from an MS-DOS (or UNIX/Linux) command prompt, as shown in this example:

```
C:\>pscp -unsafe -load ITSOadmin ITSOadmin@ITSO_V9000:/dumps/iostats/*
c:\statsfiles
```

- Specify the **-unsafe** parameter when you use wildcards.
- Use the **-load** parameter to specify the session that is defined in PuTTY.

The qperf utility

The qperf utility is an unofficial (no initial cost and unsupported) collection of *awk* scripts that was made available for download from IBM Techdocs. It provides a *quick performance* overview using the CLI and a UNIX Korn shell (it can also be used with Cygwin on Windows platforms).

You can download qperf from the Quick Performance download site.

The performance statistics files are in `.xml` format. They can be manipulated by using various tools and techniques. Figure 7-3 on page 193 shows the type of chart that you can produce by using the IBM FlashSystem V9000 controller performance statistics.

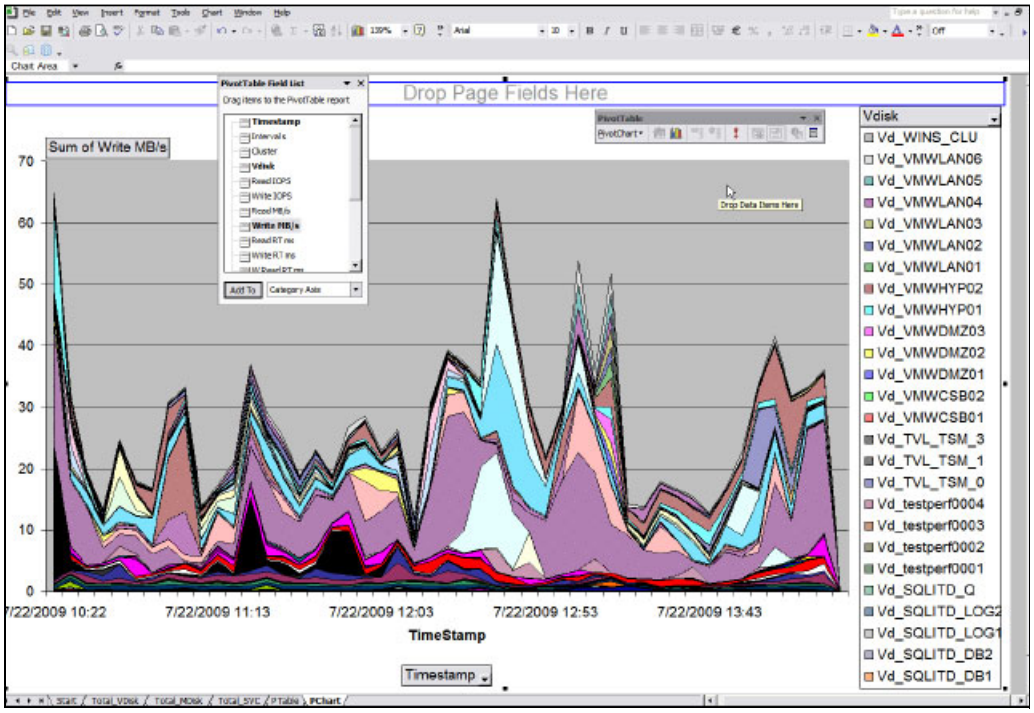


Figure 7-3 Spreadsheet example

Real-time performance monitoring

IBM FlashSystem V9000 controller supports real-time performance monitoring. Real-time performance statistics provide short-term status information for the IBM FlashSystem V9000 controller. The statistics are shown as graphs in the management GUI or can be viewed from the CLI. With system-level statistics, you can quickly view the CPU usage and the bandwidth of volumes, interfaces, and MDisks. Each graph displays the current bandwidth in megabytes per second (MBps) or I/O operations per second (IOPS), and a view of bandwidth over time.

Each control enclosure collects various performance statistics, mostly at 5-second intervals, and the statistics that are available from the config node in a clustered environment. This information can help you determine the performance effect of a specific node. As with system statistics, node statistics help you to evaluate whether the node is operating within normal performance metrics.

Real-time performance monitoring gathers the following system-level performance statistics:

- CPU utilization
- Port utilization and I/O rates
- Volume and MDisk I/O rates
- Bandwidth
- Latency

Note: Real-time statistics are not a configurable option and cannot be disabled. They will provide aggregated metrics, either for the whole system or a single node.

Real-time performance monitoring with the CLI

The **lsnodestats** and **lssystemstats** commands are available for monitoring the statistics through the CLI. Next, examples of how to use them are described.

The **lsnodestats** command provides performance statistics for the nodes that are part of a clustered system, as shown in Example 7-2 (the output is truncated and shows only part of the available statistics). You can also specify a node name in the command to limit the output for a specific node. Statistics field name descriptions are in Table 7-1 on page 195.

Example 7-2 The lsnodestats command output

```
$ ssh superuser@ITSO_V9000 lsnodestats
node_id node_name        stat_name          stat_current stat_peak stat_peak_time
1       BB1ACN1sn75AM710 compression_cpu_pc 10           10        150326170835
1       BB1ACN1sn75AM710 cpu_pc             28           28        150326170835
1       BB1ACN1sn75AM710 fc_mb              351          351       150326170835
1       BB1ACN1sn75AM710 fc_io              109447       111531    150326170805
1       BB1ACN1sn75AM710 drive_io           0            5         150326170820
1       BB1ACN1sn75AM710 drive_ms           0            0         150326170835
2       BB1ACN2sn75AM730 write_cache_pc     34           35        150326170738
2       BB1ACN2sn75AM730 total_cache_pc     80           80        150326170838
2       BB1ACN2sn75AM730 vdisk_mb           212          213       150326170833
2       BB1ACN2sn75AM730 vdisk_io           16272        16389     150326170358
2       BB1ACN2sn75AM730 vdisk_ms           0            0         150326170838
2       BB1ACN2sn75AM730 mdisk_mb           25           27        150326170733
2       BB1ACN2sn75AM730 mdisk_io           1717         2101      150326170423
```

The example shows statistics for the two node members of cluster ITSO_V9000. For each node, the following columns are displayed:

- `stat_name`. The name of the statistic field.
- `stat_current`. The current value of the statistic field.
- `stat_peak`. The peak value of the statistic field in the last 5 minutes.
- `stat_peak_time`. The time that the peak occurred (in YYMMDDhhmmss format).

However, the **lssystemstats** command lists the same set of statistics that is listed with the **lsnodestats** command, but representing all nodes in the cluster. The values for these statistics are calculated from the node statistics values in the following way:

- Bandwidth. Sum of bandwidth of all nodes.
- Latency. Average latency for the cluster, which is calculated by using data from the whole cluster, not an average of the single node values.
- IOPS. Total IOPS of all nodes.
- CPU percentage. Average CPU percentage of all nodes.

Example 7-3 shows the resulting output of the **lssystemstats** command.

Example 7-3 The lssystemstats command output

```
$ ssh superuser@ITSO_V9000 lssystemstats
stat_name          stat_current stat_peak stat_peak_time
compression_cpu_pc 9            10        150326172634
cpu_pc             28           28        150326172649
fc_mb              757          780       150326172629
fc_io              217243       219767    150326172454
sas_mb             0            0         150326172649
sas_io             0            0         150326172649
iscsi_mb           0            0         150326172649
iscsi_io           0            0         150326172649
write_cache_pc     34           35        150326172639
total_cache_pc     80           80        150326172649
vdisk_mb           392          414       150326172154
vdisk_io           31891        32894     150326172154
vdisk_ms           0            0         150326172649
mdisk_mb           99           116       150326172439
...
```

Table 7-1 briefly describes each of the statistics that are presented by the **lssystemstats** and **lsnodestats** commands.

Table 7-1 The lssystemstats and lsnodestats statistics field name descriptions

Field name	Unit	Description
cpu_pc	Percentage	Utilization of node CPUs
fc_mb	MBps	Fibre Channel bandwidth
fc_io	IOPS	Fibre Channel throughput
sas_mb	MBps	SAS bandwidth
sas_io	IOPS	SAS throughput
iscsi_mb	MBps	IP-based Small Computer System Interface (iSCSI) bandwidth
iscsi_io	IOPS	iSCSI throughput
write_cache_pc	Percentage	Write cache fullness. Updated every 10 seconds.
total_cache_pc	Percentage	Total cache fullness. Updated every 10 seconds.
vdisk_mb	MBps	Total VDisk bandwidth
vdisk_io	IOPS	Total VDisk throughput
vdisk_ms	Milliseconds	Average VDisk latency
mdisk_mb	MBps	MDisk (SAN and RAID) bandwidth
mdisk_io	IOPS	MDisk (SAN and RAID) throughput
mdisk_ms	Milliseconds	Average MDisk latency
drive_mb	MBps	Drive bandwidth
drive_io	IOPS	Drive throughput
drive_ms	Milliseconds	Average drive latency

Field name	Unit	Description
vdisk_w_mb	MBps	VDisk write bandwidth
vdisk_w_io	IOPS	VDisk write throughput
vdisk_w_ms	Milliseconds	Average VDisk write latency
mdisk_w_mb	MBps	MDisk (SAN and RAID) write bandwidth
mdisk_w_io	IOPS	MDisk (SAN and RAID) write throughput
mdisk_w_ms	Milliseconds	Average MDisk write latency
drive_w_mb	MBps	Drive write bandwidth
drive_w_io	IOPS	Drive write throughput
drive_w_ms	Milliseconds	Average drive write latency
vdisk_r_mb	MBps	VDisk read bandwidth
vdisk_r_io	IOPS	VDisk read throughput
vdisk_r_ms	Milliseconds	Average VDisk read latency
mdisk_r_mb	MBps	MDisk (SAN and RAID) read bandwidth
mdisk_r_io	IOPS	MDisk (SAN and RAID) read throughput
mdisk_r_ms	Milliseconds	Average MDisk read latency
drive_r_mb	MBps	Drive read bandwidth
drive_r_io	IOPS	Drive read throughput
drive_r_ms	Milliseconds	Average drive read latency

Real-time performance monitoring with the GUI

Real-time statistics are also available from the IBM FlashSystem V9000 controller GUI. Select **Monitoring** → **Performance** (Figure 7-4) to open the performance monitoring window.

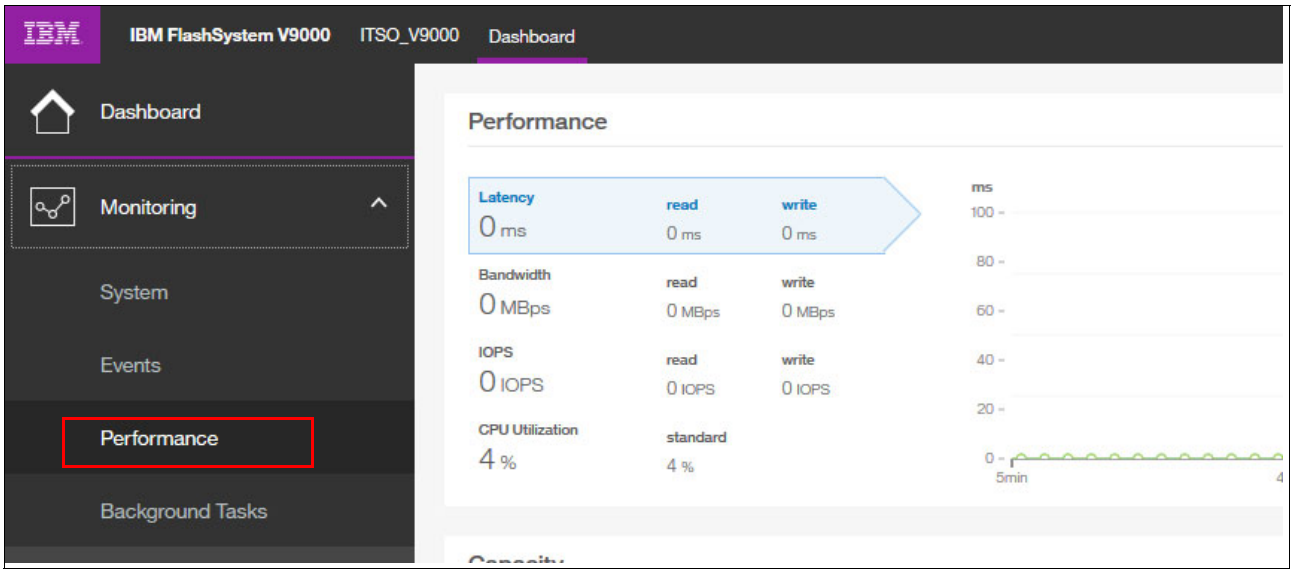


Figure 7-4 IBM FlashSystem V9000 Monitoring menu

As shown in Figure 7-5 on page 198, the Performance monitoring window is divided into the following sections that provide utilization views for the following resources:

- CPU Utilization. Shows the overall CPU usage percentage.
- Volumes. Shows the overall volume utilization with the following fields:
  - Read
  - Write
  - Read latency
  - Write latency
- Interfaces. Shows the overall statistics for each of the available interfaces:
  - Fibre Channel
  - iSCSI
  - SAS
  - IP Remote Copy

  Note: Do not over-engineer the metrics displayed in the interfaces window. Because all traffic of either direction (inbound and outbound) as well of any type (payload I/Os and SCSI commands) is aggregated, along with SAS activities (if used), especially the IOPS view can show much higher numbers than the ones seen in the Volumes windows.
- MDisks. Shows the following overall statistics for the MDisks:
  - Read
  - Write
  - Read latency
  - Write latency

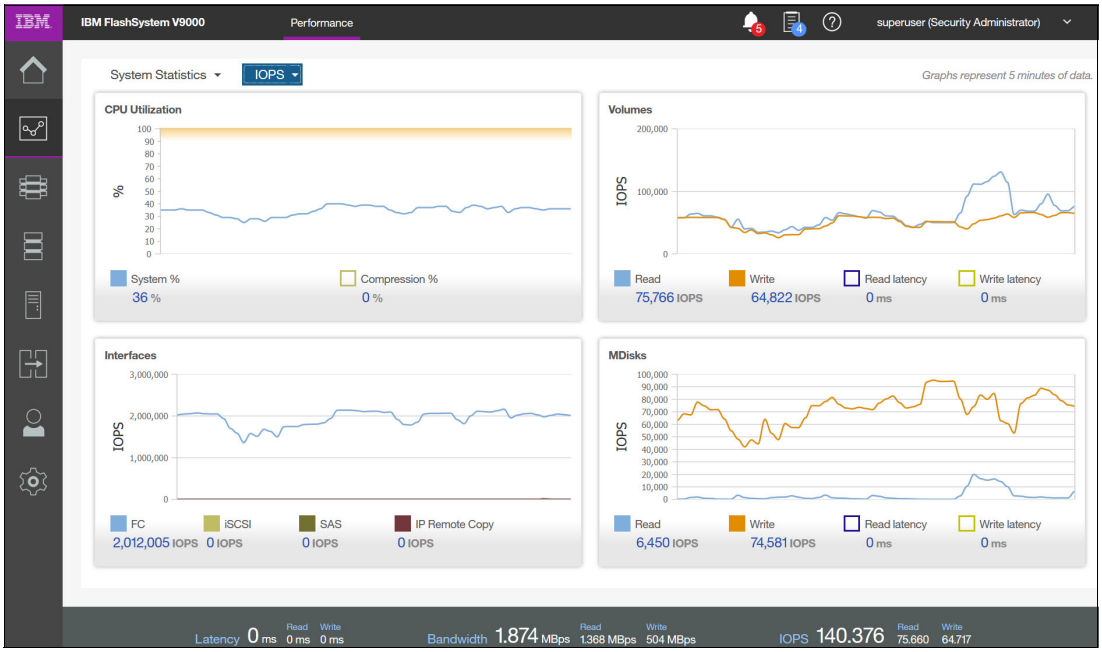


Figure 7-5 Performance monitoring window

You can also select to view performance statistics for each of the available nodes of the system (Figure 7-6).

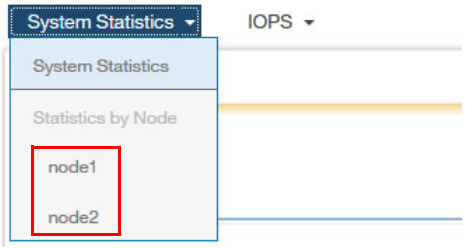


Figure 7-6 Select controller node

Also possible is to change the metric between MBps or IOPS (Figure 7-7).

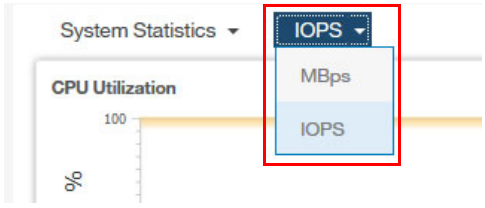


Figure 7-7 Changing to MBps or IOPS

On any of these views, you can select any point with your cursor to know the exact value and when it occurred. When you place your cursor over the time line, it becomes a dotted line with the various values gathered, as shown in Figure 7-8.

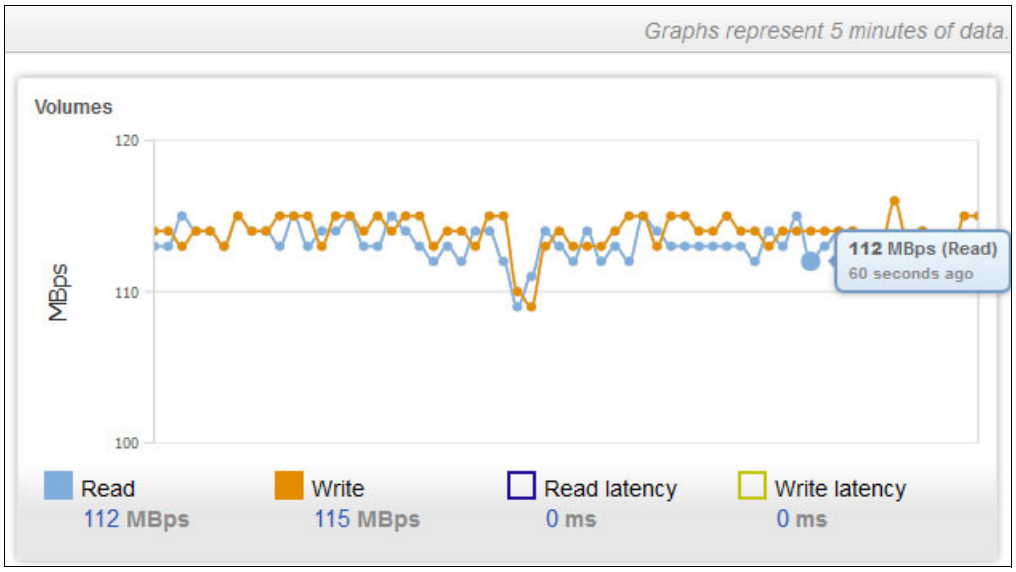


Figure 7-8 Detailed resource use

For each of the resources, you can view various values by selecting the value. For example, as shown in Figure 7-9, the four available fields are selected for the MDisks view: Read, Write, Read latency, and Write latency. In this example, latencies are not selected.

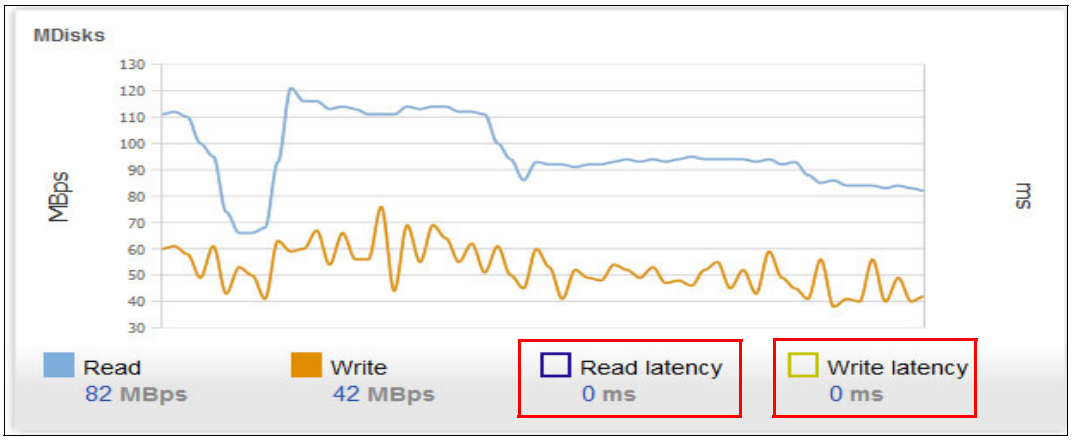


Figure 7-9 Detailed resource use

Performance data collection and IBM Spectrum Control

IBM Spectrum Control (previously known as IBM Tivoli Storage Productivity Center) provides efficient infrastructure management for virtualized, cloud, and software-defined storage to simplify and automate storage provisioning, capacity management, availability monitoring, and reporting.

The functionality of IBM Spectrum Control is provided by IBM Data and Storage Management Solutions and includes functionality delivered by IBM Virtual Storage Center, IBM Storage Integration Server, and others.

Although you can obtain performance statistics in `standard .xml` files, the use of `.xml` files is a less efficient method to analyze the IBM FlashSystem V9000 controller performance statistics. IBM Spectrum Control is the supported IBM tool to collect and analyze IBM FlashSystem V9000 controller performance statistics.

See the following resources:

- IBM Spectrum Control
- Current support matrix
- Using IBM Spectrum Control to manage your storage subsystem:

  IBM Spectrum Family: IBM Spectrum Control Standard Edition, SG24-8321
- Performance data information in IBM Knowledge Center

7.2 Estimating compression savings

Some common data types are good candidates for compression, and others are not. The best candidates for data compression are data types that are not compressed by nature. Viable candidates include data types that are involved in many workloads and applications, such as databases, character/ASCII based data, email systems, server virtualization, CAD/CAM, software development systems, and vector data.

Table 7-2 Typical compression ratios

Type of data	Expected ratio
Databases	50-80%
Server Virtualization	45-70%
Seismic Data	40-70%
Engineering Data	50-80%
E-mail	30-60%

The IBM *Comprestimator* utility can help determine whether your data is a candidate for compression, and to which degree data can be compressed. A host-installed version can be used on any type of attached block devices, and a built-in tool is available for analyzing existing IBM FlashSystem V9000 volumes.

IBM Compretimator can quickly scan existing volumes and provide an accurate estimation of expected compression ratio.

7.2.1 IBM Compretimator: Built-in GUI version

The management GUI has a built-in Compretimator function that uses mathematical and statistical algorithms to estimate the potential savings for the system when using compression at V9000 controller level (not in the AE3 back-end). If you are wondering whether to purchase the compression license for the system, this tool helps determine the capacity that you might save by using compression.

Estimating each volume does not take long to complete; however, analysis is completed on one volume at a time per node. In large configurations, estimation results can take some time.

Note: The built-in Comprestimator can also be accessed from the V9000's CLI. For further details, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Figure 7-10 shows how to start the analysis of the system:

1. In the Volumes > Volumes view, select **Actions** → **Space Savings** → **Estimate Compression Savings**. The system should have volumes with data that has already been written for the analysis to be accurate.

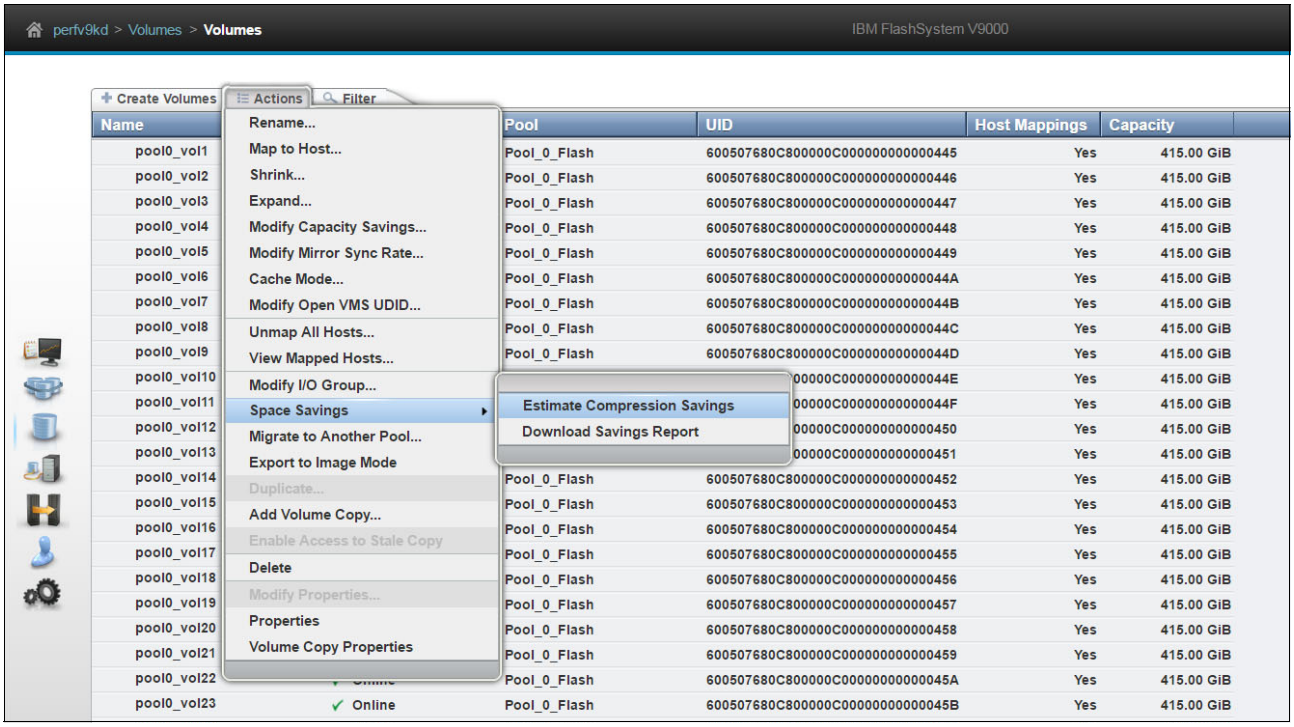


Figure 7-10 Starting the Compression Savings Estimate analysis of existing volumes

2. The same selection to start the analysis is used to determine whether it completed. In the Volumes > Volumes view, select **Actions** → **Space Savings** → **Estimate Compression Savings**. Figure 7-11 shows what occurs when the system is still analyzing the compression savings.

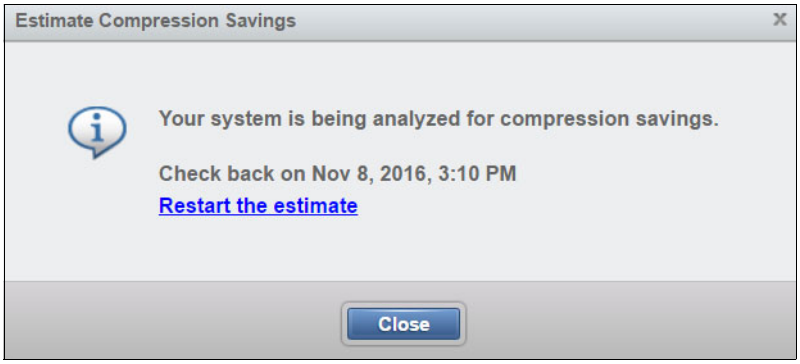


Figure 7-11 Compression Savings Estimate analysis is in progress

Figure 7-12 shows the results of the analysis. This is viewed using the same selection for starting the analysis.

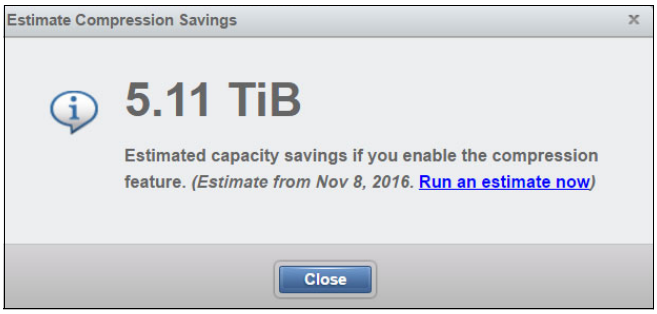


Figure 7-12 Compression Savings Estimate analysis complete

After the Estimate Compression Savings process shows a value for the savings, you can also download a more detailed report that lists the savings per volume.

3. Figure 7-13 shows how to download the Space Savings Report. In the Volumes > Volumes view, select **Actions** → **Space Savings** → **Download Savings Report**.

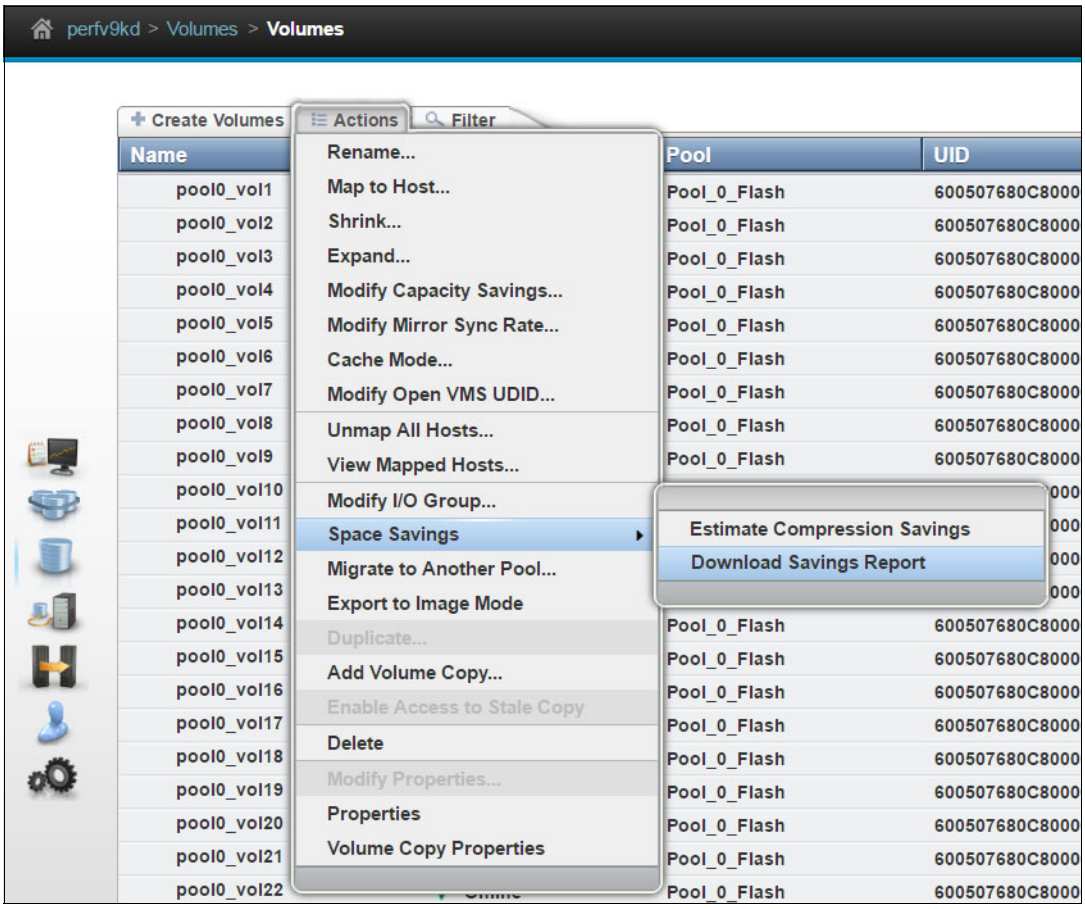


Figure 7-13 Downloading the Space Savings Report

Compression affects the entire system performance, because system resources become dedicated to compression when the first compressed volume is created. These resources are released when the last compression volume is removed. Care should be taken before creating any compressed volumes if the system is already heavily loaded.

7.2.2 IBM Comprestimator utility: Host-installed version

IBM Comprestimator is a command-line, host-based utility that can be used to estimate the compression rate for block-devices. The IBM Comprestimator utility uses advanced mathematical and statistical formulas to perform the sampling and analysis process in a short and efficient way. The utility also displays its accuracy level by showing the maximum error range of the results achieved based on the formulas that it uses. The utility runs on a host that has access to the devices to be analyzed, and only runs read operations, so it has no effect on the data stored on the device.

The following section provides useful information about installing IBM Comprestimator on a host and using it to analyze devices on that host. Depending on the environment configuration, in many cases, IBM Comprestimator is used on more than one host to analyze additional data types.

IBM Comprestimator is supported and, as of the time that this publication was written, can be used on the following client operating system versions:

- Windows 2003 Server, Windows 2008 R2 Server, Windows 2012, Windows 7, Windows 8
- Red Hat Enterprise Linux Version 5.x, 6.x, 7.1 (x86 64bit)
- ESXi 4.1, 5.5, 6.0
- Sun Solaris 10, 11
- AIX 6.1, 7.1
- HPUX 11.31
- SUSE Linux Enterprise Server 11 (x86 64bit)
- Ubuntu 12 (x86 64bit)

Installing IBM Comprestimator

The Comprestimator utility and installation instructions are available from the IBM Comprestimator home page.

IBM Comprestimator can be installed only on supported operating systems (see previous list). After installation completes, the binary files for other supported operating systems are available in the Windows installation folder.

By default, the files are copied to the following locations:

- In Windows 64-bit: `C:\Program Files (x86)\IBM\Comprestimator`
- In Windows 32-bit: `C:\Program Files\IBM\Comprestimator`

After transferring the operating system-dependent IBM Comprestimator tools to your system, follow the installation instructions that are provided on the Comprestimator download page. The program invocation is different on different operating systems, but the output is the same.

Using IBM Comprestimator

This topic describes the syntax, the output, and an explanation of the output of the IBM Comprestimator utility.

IBM Comprestimator syntax

Example 7-4 on page 204 shows a sample syntax for the IBM Comprestimator tool.

Example 7-4 IBM Comprestimator syntax

```
Comprestimator version : 1.5.3.1 (Build w0117)
Usage :
comprestimator <-s storage_type> [ -h | -d device] [-c filename]  [-v] [-p
number_of_threads] [-P] [-I] [--storageVer=version] [--config=task_file]
-d device name       Path of device to analyze (e.g.: /dev/hdisk0)
-p number            Number of threads (default 10)
-c                   Export the results to a CSV file
-v                   Verbose output
-h                   Print this help message
-P                   Display results using a paragraph format
-s,--storageSys      Storage system type. Supported values are: SVC, XIV and FLASHSYSTEM
-I                   Allow larger scale of storage io-error threshold rate (up to 5%)
--flash-modules      Configuration of flash modules (FLASHSYSTEM) (6, 8, 10, 12)
--flash-module-type  Type of flash modules (FLASHSYSTEM).  Supported values are: SMALL,
MEDIUM, LARGE (default is MEDIUM)
--config=file        Configuration file that contains list of devices to analyze
--storageVer=version Target storage system version. Supported Storwize/SVC/Flex options:
6.4, 7.1, 7.2, 7.3; default: 7.3, XIV options: 11.6
```

IBM Comprestimator output

To list storage devices, use the **comprestimator -l** command (Example 7-5).

Example 7-5 Comprestimator: List devices (output shortened for clarity)

```
C:\Program Files (x86)\ibm\Comprestimator\Windows>comprestimator -l
Drive number [0]        \\?\scsi#disk&ven_lsilogic&prod_logical_volume#5&138f362
Drive number [1]        \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&363
Drive number [2]        \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&363
Drive number [3]        \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&363
Drive number [4]        \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&363
Drive number [5]        \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&363
```

This sample analyzes drive number 2 (Example 7-6).

Example 7-6 Analyze Drive number 2 (output shortened for clarity)

```
C:\Program Files (x86)\ibm\Comprestimator\Windows>comprestimator -n 2 -s SAN Volume Controller -v
Version: 1.5.2.2 (Build w0098)
Start time: 15/07/2015 13:48:18.103676
Device name: \\?\mpio#disk&ven_ibm&prod_2145&rev_0000#1&7f6ac24&0&3630303530373
Device size: 100.0 GB
Number of processes: 10

Sample#| Device | Size(GB) | Compressed |   Total    |   Total    | Thin Provisioning | Compression |   Compression
       |  Name  |          |  Size(GB)  | Savings(GB)| Savings(%) |    Savings(%)     | Savings(%)  | Accuracy Range(%)
-------+--------+----------+------------+------------+------------+-------------------+-------------+------------------
32     |********|    100.0 |        6.2 |       93.8 |      93.8% |             75.6% |       74.5% |            51.3%
69     |********|    100.0 |        8.1 |       91.9 |      91.9% |             72.8% |       70.0% |            34.9%
103    |********|    100.0 |        9.1 |       90.9 |      90.9% |             71.3% |       68.2% |            28.6%
```

According to Example 7-6, a savings from compression is in the range 68.2 - 74.5% by enabling compression on the system containing the IBM FlashSystem V9000 volume.

Tip: For IBM FlashSystem V840, select **SAN Volume Controller** as the Storage system type. For IBM FlashSystem V9000, select **FLASHSYSTEM** as the Storage system type.

Explanation of compression output

Table 7-3 explains the output from the IBM Comprestimator.

Table 7-3 IBM Comprestimator output explanations

Header	Explanation
Sample#	The number of the current sample reported.
Device	The device name used in the scan.
Size (GB)	The total size of the device as reported by the operating system, in gigabytes.
Compressed Size (GB)	The estimated size of the device if it is compressed using IBM FlashSystem V9000 Real-time Compression, in gigabytes.
Total Savings (GB)	The total estimated savings from thin-provisioning and compression, in gigabytes.
Total Savings (%)	The estimated savings from thin-provisioning and compression, in percentage of the size of the device. This value is calculated in the following method: Total Savings (%) = 1 - (Compressed Size (GB) / Size (GB)).
Thin Provision Savings (%)	The estimated savings from thin provisioning (areas with zeros are stored using minimal capacity).
Compression Savings (%)	The estimated savings from compression.
Compression Accuracy Range (%)	The accuracy of the estimate provided by Comprestimator. The results provided are estimated based on samples from the device, and therefore might be lower or higher than the actual compression that would be achieved. The approximate accuracy of the results is represented as a percentage of the total size of the device. For example, if the estimated Compression Savings (%) is 67%, and the Compression Accuracy Range is 5%, the actual compression savings (in percentage) if this device is compressed on IBM FlashSystem V9000 is 62% - 72%.

7.3 Command-line hints

IBM FlashSystem V9000 contains a robust command-line interface based on the IBM SAN Volume Controller and Storwize family of products. These command-line scripting techniques can be used to automate the following tasks:

- Running commands on the cluster
- Creating connections
- IBM command-line scripting
- Example commands
- Backing up the Configuration
- Running the Software Upgrade Test Utility
- Secure Erase of Data

7.3.1 Running commands on the IBM FlashSystem V9000

To automate copy services processes, you must connect to the cluster. In normal operations, you connect to the cluster by using the GUI or the command line interface (CLI). The GUI is not an appropriate interface for automating processes, so that alternative is not described here. All automation techniques are achieved through the IBM FlashSystem V9000 CLI or the Common Information Model Object Manager (CIMOM), which currently acts as a proxy to the command line.

This section uses the term *user agent.* The user agent can be the CIMOM, which connects to the cluster by using Secure Shell (SSH). Or the user agent can be a user connecting directly with an SSH client, either in an interactive mode or by using a script.

Running commands to the cluster follows this sequence of steps:

1. Connection
2. Authentication
3. Submission
4. Authorization
5. Running a command (Execution)

Connection

Commands are submitted to the cluster during a connection session to the cluster. User agents make connections through the SSH protocol. FlashSystem has several security features that affect how often you can attempt connections. These security features are in place to prevent attacks (malicious or accidental) that can bring down an IBM FlashSystem V9000 controller node. These features might initially seem restrictive, but they are relatively simple to work with to maintain a valid connection.

When creating automation by using the CLI, an important consideration is to be sure that scripts behave responsibly and do not attempt to breach the connection rules. At a minimum, an automation system must ensure that it can gracefully handle rejected connection attempts.

Figure 7-14 shows how IBM FlashSystem V9000 connection restrictions work.

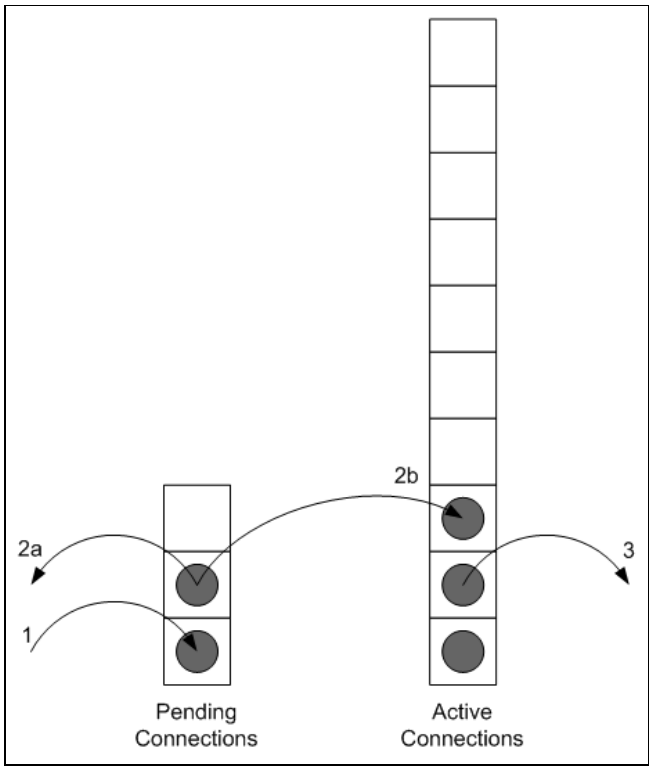


Figure 7-14 IBM FlashSystem V9000 SSH restrictions

Figure 7-14 shows that two queues are in action: *Pending Connections* and *Active Connections*. The connection process follows this sequence:

1. A connection request comes into the IBM FlashSystem V9000. If the Pending Connections queue has a free position, the request is added to it; otherwise, the connection is explicitly rejected.
2. Pending Connections are handled in one of two ways:
   a. If any of the following conditions are true, the connection request is rejected:
      - No key is provided, or the provided key is incorrect.
      - The provided user name is not admin or service.
      - The Active Connections queue is full. In this case, a warning is returned to the SSH client as shown in Example 7-7 on page 208.

   b. If none of the conditions listed in the previous step are true, the connection request is accepted and moved from the Pending Connections queue to the Active Connections queue.
3. Active Connections end after any of the following events:
   – The user logs off manually.
   – The SAN Volume Controller SSH daemon recognizes that the connection has grown idle.
   – The network connectivity fails.
   – The configuration node fails over.

   In this case, both queues are cleared because the SHH daemon stops and restarts on a different node.

Example 7-7 shows an IBM FlashSystem V9000 command-line warning about too many logins. Only 10 concurrent active SSH sessions are allowed.

Example 7-7 IBM FlashSystem V9000 command-line warning about too many logins

```
$ ssh ITSOadmin@ITSO_V9000
CMMVC7017E Login has failed because the maximum number of concurrent CLI sessions
has been reached.

Connection to ITSO_V9000 closed.
```

If the limit of 10 concurrent active SSH sessions is reached, an entry is generated on the error log, as shown in Figure 7-15.

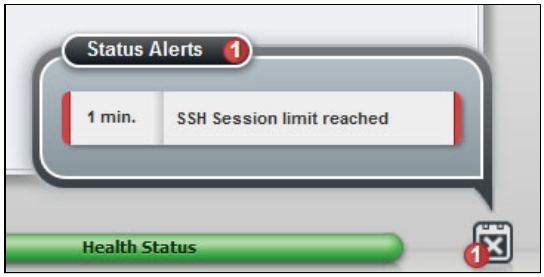


Figure 7-15 GUI console warning the limit was reached

In this case, complete the following steps:

1. Double-click the status alert to see the event panel (Figure 7-16).

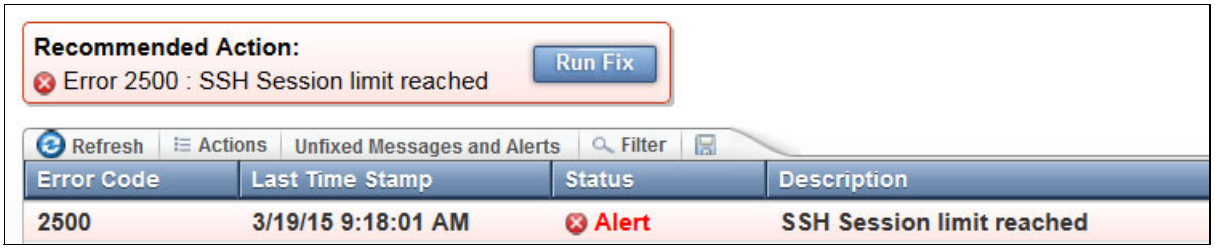


Figure 7-16 Error 2500: SSH Session limit reached

2. To view the details, right-click the error event and select **Properties**. The event details are displayed (Figure 7-17).

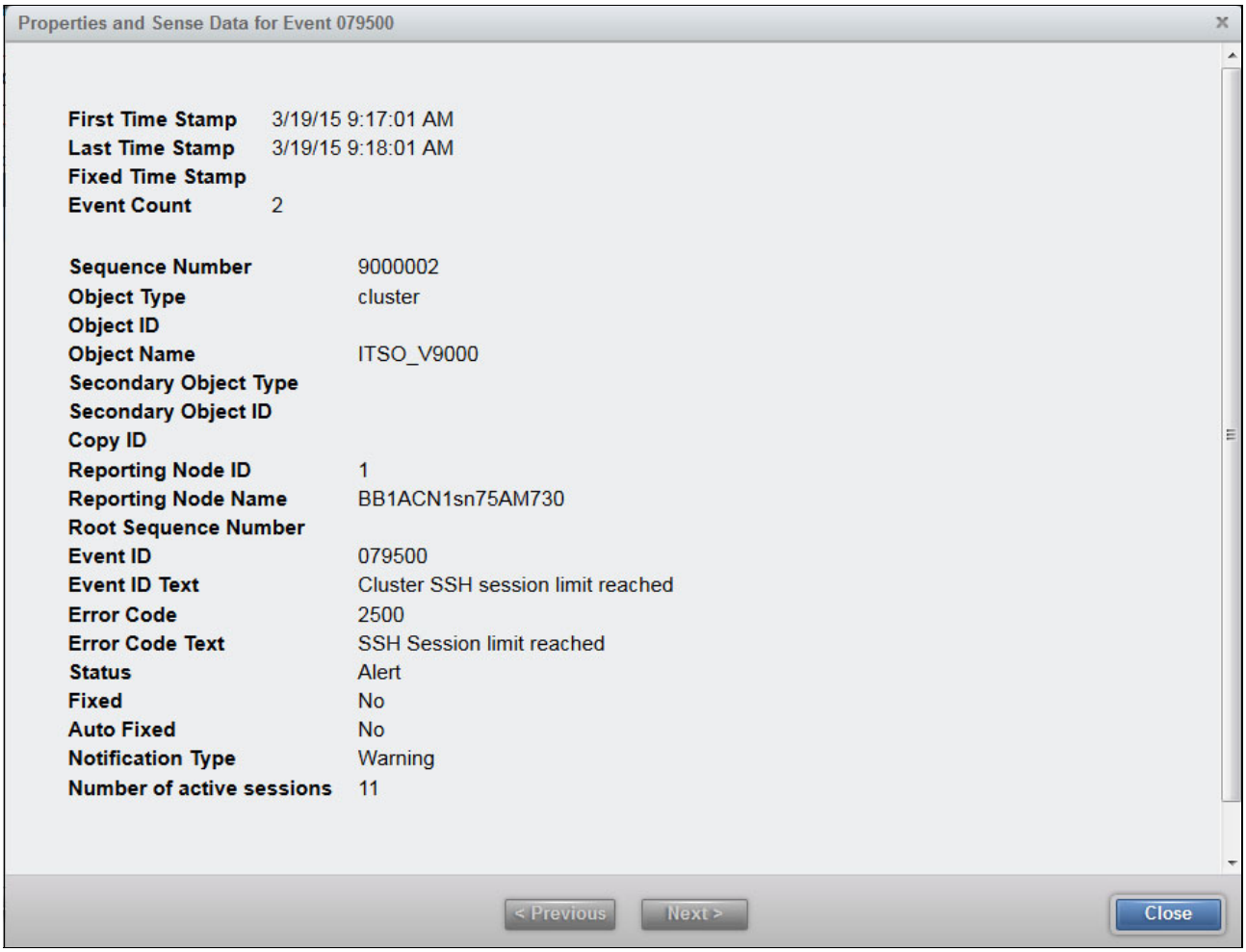


Figure 7-17 Event details

3. To fix this error, choose the **Run Fix Procedure** (described in *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933). A list with active SSH sessions is displayed (Figure 7-18 on page 210). The quickest way to resolve the error is as follows:

   a. Close all connections.

      Selecting **Close all SSH sessions through this fix procedure** closes the listed sessions, and the error is fixed. If you close the active sessions manually on the host side without choosing to close all of the sessions through the Run Maintenance Procedures, you must select **The number of SSH sessions has been reduced, mark this event as fixed**.

   b. Click **Next** to continue.

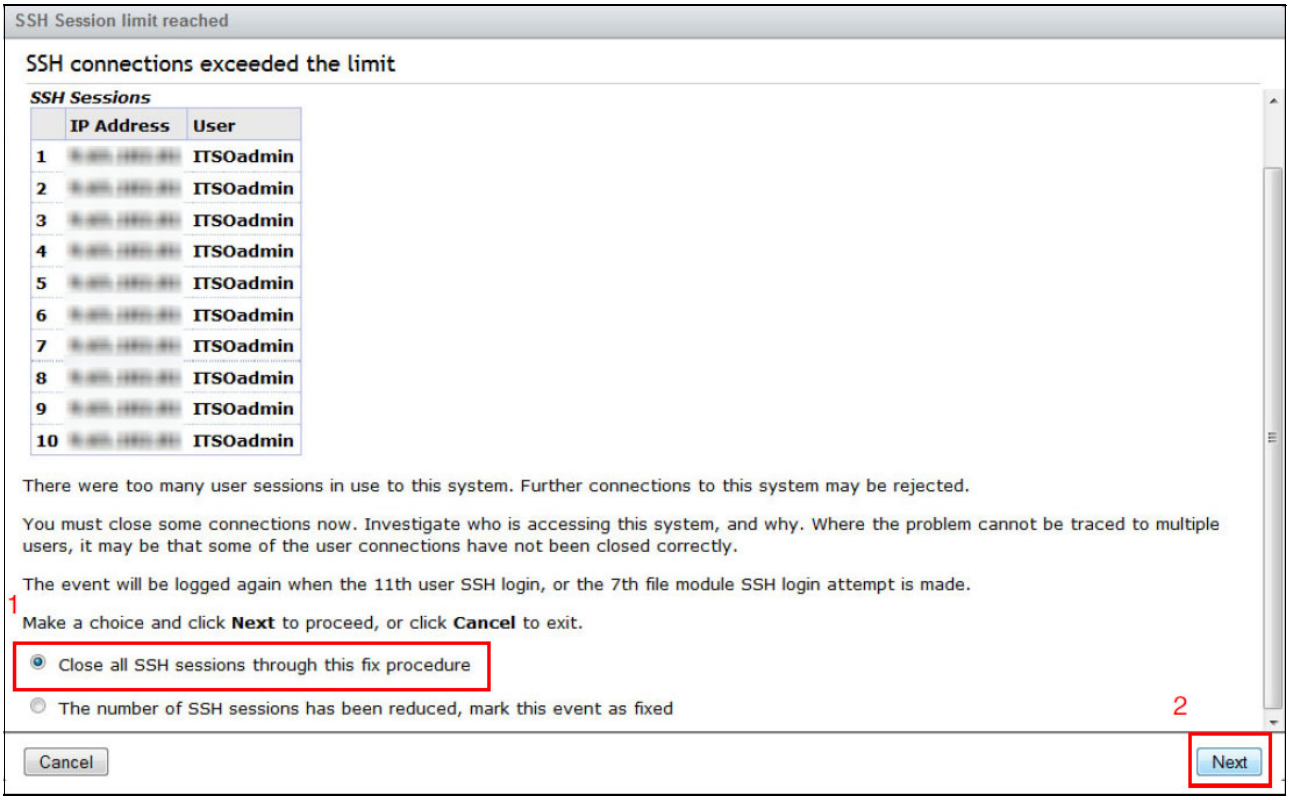


Figure 7-18 SSH Session limit reached

c. A warning that all CLI connections will be closed is displayed (Figure 7-19). Click **Next** to determine whether the process is completed.

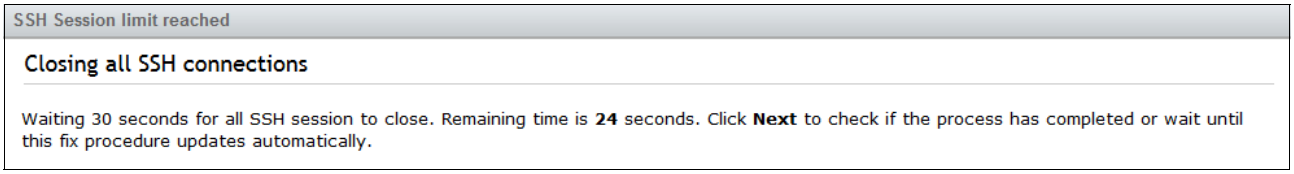


Figure 7-19 Warning about closing all SSH connections

Authentication

IBM FlashSystem V9000 enables you to log in with basically a user name and password. The two types of users who can access the system are local users and remote users. These types are based on how the users are authenticated to the system:

- *Local users* must provide a password, an SSH key, or both. Local users are authenticated through the authentication methods that are in the SAN Volume Controller system. If the local user needs access to the management GUI, a password is needed for the user. If the user requires access to the CLI through SSH, either a password or a valid SSH key file is necessary.

  Local users must be part of a user group that is defined on the system. User groups define roles that authorize the users within that group to a specific set of operations on the system.

- *Remote users* are authenticated on a remote service with either Tivoli Integrated Portal or Lightweight Directory Access Protocol (LDAP v3) support, such as IBM Tivoli Storage Productivity Center, which delivers the functionality of IBM Spectrum Control, or IBM Security Directory Server. A remote user does not need local authentication methods.

  With Tivoli Integrated Portal, both a password and SSH key are required to use the CLI. With LDAP, having a password and SSH key is not necessary, although SSH keys optionally can be configured. Remote users who need to access the system when the remote service is down also need to configure local credentials. Remote users have their groups defined by the remote authentication service.

For more information, see the following resources:

- For details about using the management GUI to manage users and user groups on the system, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.
- To configure remote authentication against servers that implement the Lightweight Directory Access Protocol (LDAP), including IBM Security Services, see 5.3.1, “Remote authentication with LDAP” on page 123.
- For information about the auditing of commands on the IBM FlashSystem V9000 cluster, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Submission

When connected to a cluster, the user agent can start submitting commands. First, the syntax is checked. If the syntax checking fails, an appropriate error message is returned. Any automation implementation must ensure that all submitted commands have the correct syntax. If they do not, they must be designed to handle syntax errors. Designing a solution that does not generate invalid syntax is easier than designing a solution to handle all potential syntax errors.

Authorization

Next, commands with valid syntax are checked to determine whether the user agent has the authority to submit the command. A role is associated with the key that was used to authenticate the connection. IBM FlashSystem V9000 checks the submitted command against the authorization role. If the user agent is *authorized*, the command is sent to be run.

If the user agent is *not authorized* to run this command, the following error is returned:

```
CMMVC6253E The task has failed because the user's role is not authorized to submit
the command.
```

See the following resources:

- For information about authorization and roles, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.
- For more details, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933.

Running a command

When a command is run, it can fail (one possible scenario) or succeed (four possible scenarios):

- The command fails. An error message is written to STDERR.
- The command succeeds. A warning is written to STDERR.
- The command succeeds. A warning is written to STDERR; information is sent to STDOUT.

- The command succeeds. Information is written to STDOUT.
- The command succeeds. Nothing is written to STDOUT.

Note: Data that is written to STDOUT and STDERR by the IBM FlashSystem V9000 is written to STDOUT and STDERR by your SSH client. However, you must manually verify that the data was written to STDOUT and STDERR by your SSH client.

7.3.2 Creating connections

Connecting to the IBM FlashSystem V9000 cluster is the first step in running commands. Any automation solution requires a connection component. This component must be as robust as possible, because it forms the foundation of your solution.

There are two forms of connection solutions:

- Transient: One command is submitted per connection, and the connection is closed after the command is completed.
- Persistent: The connection is made and stays open. Multiple commands are submitted through this single connection, including interactive sessions and the CIMOM.

Transient connections

Transient connections are simple to create. The most common SSH clients enable the user to submit a command as part of the user's invocation. Example 7-8 shows a user submitting two commands as part of the user's invocation using **ssh** on an AIX server. Using the operating system command, the IBM FlashSystem V9000 output can be processed.

Example 7-8 Transient connection to IBM FlashSystem V9000 from AIX Server

```
# ssh -i publickey -l ITSOadmin ITSO_V9000 lsenclosure -delim :
id:status:type:managed:IO_group_id:IO_group_name:product_MTM:serial_number:total_c
anisters:online_canisters:total_PSUs:online_PSUs:drive_slots:total_fan_modules:onl
ine_fan_modules
1:online:expansion:yes:0::9846-AE2:1371006:2:2:2:2:12:0:0

# ssh -i publickey -l ITSOadmin ITSO_V9000 lsenclosure -delim : | cut -f1,2,7,8 -d
:
id:status:product_MTM:serial_number
1:online:9846-AE2:1371006
#
```

Example 7-9 shows a user submitting a command as part of the user's invocation using the **plink** command on a Windows server.

Example 7-9 Transient connection to IBM FlashSystem V9000 from Windows server

```
C:\Program Files\Putty>plink -i private.ppk -l superuser ITSO_V9000 lsenclosure -delim :
id:status:type:managed:IO_group_id:IO_group_name:product_MTM:serial_number:total_canisters:
online_canisters:total_PSUs:online_PSUs:drive_slots:total_fan_modules:online_fan_modules
1:online:expansion:yes:0::9846-AE2:1371006:2:2:2:2:12:0:0

C:\Program Files\Putty>
```

These transient connections go through all five stages of running a command and return to the command line. You can redirect the two output streams (STDOUT and STDERR) using the operating system's standard redirection operators to capture the responses.

These lengthy invocations can be shortened in client-specific ways. User configuration files can be used with the AIX SSH client. The configuration file in Example 7-10 enables you to create a transient connection.

Example 7-10 Sample SSH configuration file saved as sampleCfg

```
# cat sampleCfg
Host ITSO
HostName ITSO_V9000
IdentityFile ./privateKey
User ITSOadmin

Host ITSOsu
HostName ITSO_V9000
IdentityFile .ssh/id_rsa
User superuser
```

The Transient connection is shown in Example 7-11.

Example 7-11 Transient connection to IBM FlashSystem V9000 using SSH and configuration file

```
# ssh -F sampleCFG ITSOsu sainfo lsservicenodes
panel_name cluster_id       cluster_name node_id node_name        relation  node_status
error_data
75AM710    00000203202035F4 ITSO_V9000   1       BB1ACN1sn75AM710 local     Active
75AM730    00000203202035F4 ITSO_V9000   2       BB1ACN2sn75AM730 cluster   Active
01-2       00000203202035F4 ITSO_V9000                            expansion Service     690
01-1       00000203202035F4 ITSO_V9000                            expansion Managed
```

Shortening the **plink** invocation requires the creation of a PuTTY session:

1. First, open the PuTTY application and enter the following line in the Host Name (or IP address) field, as shown in Figure 7-20:

   `superuser@<Host Name or cluster IP address>`

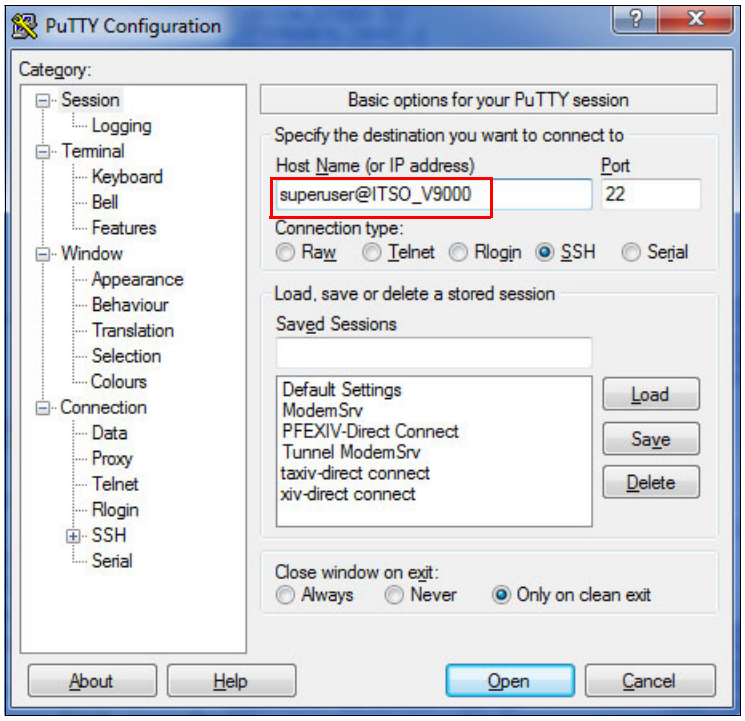


Figure 7-20 Add user name and system name to PuTTY session

2. Configure the private key for this session by making the selections, as shown in steps 1, 2, and 3 of Figure 7-21. click **Browse** (step 4) to locate the private key file.

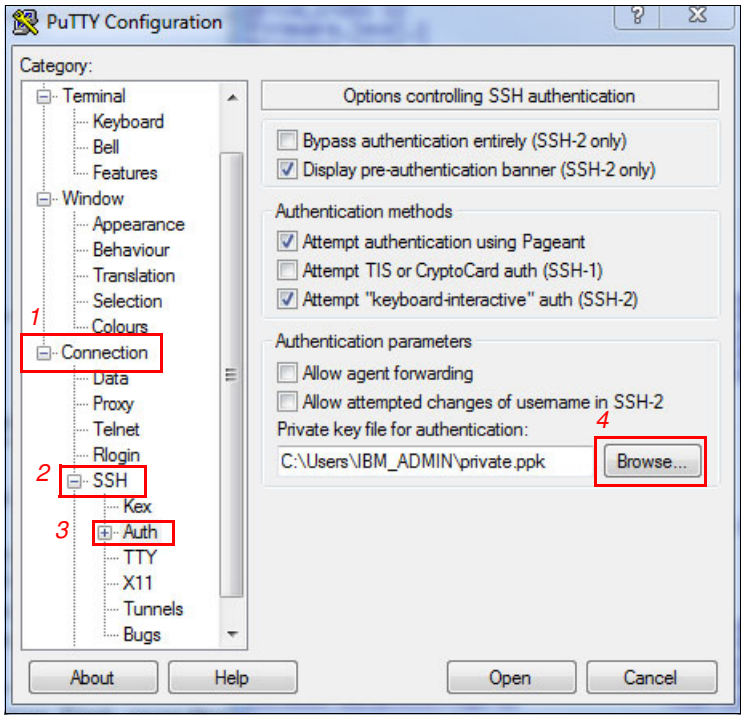


Figure 7-21 Set private key for PuTTY SSH session

3. Complete saving the session (Figure 7-22) by returning to the Session Panel (1), providing a session name (2), and clicking **Save** (3).

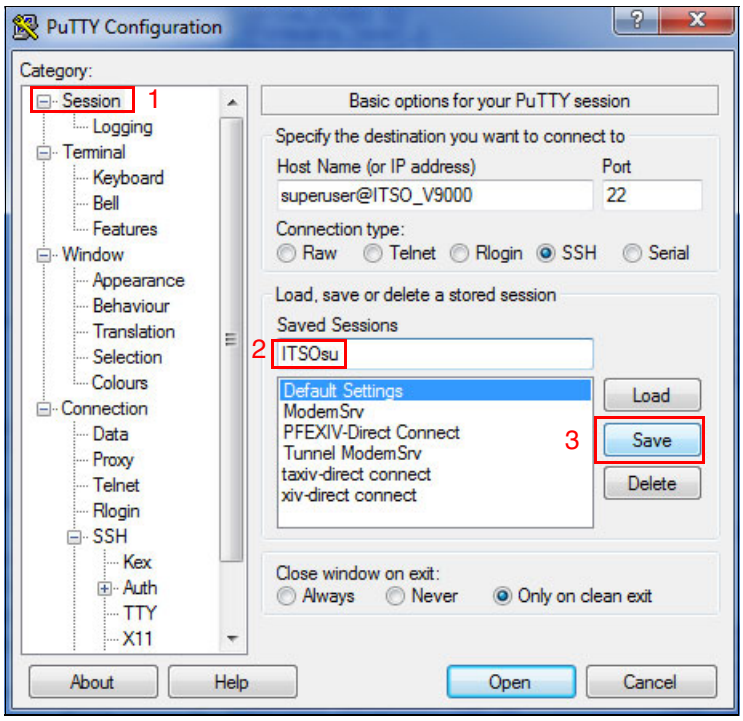


Figure 7-22 Save PuTTY session for use with plink

After a session is saved, you can use it to make transient connections from the command line (Example 7-12).

Example 7-12 Transient connection to IBM FlashSystem V9000 using plink with PuTTY session

```
C:\Users\IBM_ADMIN>plink -load ITSOsu lsenclosurebattery
enclosure_id battery_id status charging_status recondition_needed percent_charged
end_of_life_warning
1            1          online idle            no                 97 no
1            2          online idle            no                 91 no

C:\Users\IBM_ADMIN>
```

Persistent connections

A persistent connection is a connection that exists beyond the submission and execution of a single command. As outlined previously, the CIMOM provides a persistent connection, but it does not provide direct access to the command line. To provide a persistent connection to the command line, you must use multiple processes.

There are as many ways to provide a persistent connection to the command line as there are programming languages. Most methods involve creating a process that connects to the cluster, writing to its STDIN stream, and reading from its STDOUT and STDERR streams.

You can use persistent connections in several ways:

- On a per-script basis

  A script opens a connection that exists for the life of the script, enabling multiple commands to be submitted. The connection ends when the script ends.

- As a stand-alone script

  A connection is opened and other scripts communicate with this script to submit commands to the cluster. This approach enables the connection to be shared by multiple scripts. This in turn enables a greater number of independent scripts to access the cluster without using up all of the connection slots.

For more information about transient and persistent connections, see *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

7.3.3 IBM FlashSystem V9000 command-line scripting

When connected to the cluster command line, you can use small amounts of automation for various purposes, including for the following tasks:

- Repeatedly submitting a single command to a set of IBM FlashSystem V9000 objects
- Searching the configuration for objects conforming to certain criteria

The IBM FlashSystem V9000 command line is a highly restricted Bash shell. You cannot access UNIX commands, such as **cd** or **ls**. The only commands that are available are built-in commands, such as **echo** or **read**. In addition, redirecting inputs and outputs is *not* supported, but you can pipe commands together.

Note: IBM FlashSystem V9000 uses IBM Spectrum Virtualize technology, built on the foundation of the SAN Volume Controller. The command lines function in the same secure way, which enables you to use existing scripting for automation and especially replication.

Example 7-13 shows a script that lists all volumes that are not online. This script complements the **filtervalue** parameter of the **lsvdisk** command. The **filtervalue** parameter provides matches only when a property matches a value. The command-line script in Example 7-13 provides matches according to other criteria.

Example 7-13 IBM FlashSystem V9000 command-line script listing volumes that are not online

```
001. lsvdisk -nohdr | while read id name IOGid IOGname status rest
002. do
003. if [ "$status" != "online" ]
004. then
005.  echo "Volume '$name' \($id\) is $status"
006. fi
007. done
```

Note: The message `vdisks offline` is an error condition. In normal operations, you do not find any that are not online.

Line 001 submits the **lsvdisk** command and pipes the output to the **read** command, which is combined with a **while** command. This combination creates a loop that runs once per line of output from the **lsvdisk** command.

The **read** command is followed by a list of variables. A line is read from the **lsvdisk** command. The first word in that line is assigned to the first variable. The second word is assigned to the second variable, and so on, with any remaining words assigned to the final variable (with intervening spaces included).

In this case, the **-nohdr** parameter is used to suppress display of the headings.

Lines 003 - 006 check the status variable. If it is not equal to `online`, the information is printed to STDOUT.

Submitting command-line scripts

You can submit command-line scripts from an interactive prompt, if required. However, you can also submit the scripts as batch files. Example 7-14 shows how to submit scripts as batch files with **ssh**.

Example 7-14 Submission of batch file to IBM FlashSystem V9000 using SSH

```
ssh superuser@ITSO_V9000 -T < batchfile.sh
Host and WWPN info:

Host 0 (TA_Win2012) : WWPN is =10000000C9B83684
Host 0 (TA_Win2012) : WWPN is =10000000C9B83685
```

Example 7-15 shows how to submit scripts as batch files with **plink**.

Example 7-15 Submission of batch file to IBM FlashSystem V9000 using plink

```
C:\>plink -load ITSOadmin -m batchfile.sh

Host and WWPN info:

Host 0 (RedHat) : WWPN is =2100000E1E302C73
Host 0 (RedHat) : WWPN is =2100000E1E302C72
Host 0 (RedHat) : WWPN is =2100000E1E302C51
Host 0 (RedHat) : WWPN is =2100000E1E302C50
Host 1 (AIX) : WWPN is =10000090FA13B915
Host 1 (AIX) : WWPN is =10000090FA13B914
Host 1 (AIX) : WWPN is =10000090FA0E5B95
Host 1 (AIX) : WWPN is =10000090FA0E5B94
Host 1 (AIX) : WWPN is =10000090FA02F630
Host 1 (AIX) : WWPN is =10000090FA02F62F
Host 1 (AIX) : WWPN is =10000090FA02F621
Host 1 (AIX) : WWPN is =10000090FA02F620
Host 2(TA_Win2012) : WWPN is =10000000C9B83684
Host 2(TA_Win2012) : WWPN is =10000000C9B83685
```

Both commands submit a simple batch file, as shown in Example 7-16. This command lists the WWPN for each host defined in the IBM FlashSystem V9000 (Example 7-15).

Example 7-16 Command-line batch file (batchfile.sh) used in the previous examples

```
echo "Host and WWPN info:"
echo " "
lshost -nohdr | while read name product_name WWPN
do
    lshost $name| while read key value
      do
      if [ "$key" == "WWPN" ]
      then
         echo "Host $name ($product_name) : WWPN is =$value"
      fi
    done
done
```

Server-side scripting

Server-side scripting involves scripting where the majority of the programming logic is run on a server.

Part of server-side scripting is the generation and management of connections to the IBM FlashSystem V9000 system. For an introduction showing how to create and manage a persistent connection to a system and how to manage requests coming from multiple scripts, see “Persistent connections” on page 216.

The Perl module handles the connection aspect of any script. Because connection management is often the most complex part of any script, an advisable task is to investigate this module. Currently, this module uses transient connections to submit commands to a cluster, and it might not be the best approach if you plan to use multiple scripts submitting commands independently.

7.3.4 Sample commands of mirrored VDisks

This section contains sample commands that use the techniques demonstrated in 7.3.3, “IBM FlashSystem V9000 command-line scripting” on page 216. These examples are based on a single building block configuration with sample data designed to support this publication.

Tip: Start with small examples to understand the behavior of the commands.

VDisk mirroring to a second enclosure

This example shows how to mirror all VDisks for redundancy or how to vacate a storage enclosure.

The sync rate

Example 7-17 shows mirroring the VDisks to a new managed disk group. In this example, `sync rate` is low so that it does not adversely affect the load on the system. You can check the progress of synchronization with the **`lsvdisksyncprogress`** command.

Example 7-17 Mirror all VDisks

```
lsvdisk -filtervalue  copy_count=1 -nohdr |
while read id vdiskname rest
do
   addvdiskcopy -mdiskgrp newgroupname -syncrate 30 $id
done
Vdisk [0] copy [1] successfully created
Vdisk [1] copy [1] successfully created
Vdisk [2] copy [1] successfully created
Vdisk [3] copy [1] successfully created
Vdisk [4] copy [1] successfully created
Vdisk [5] copy [1] successfully created
```

Raise the sync rate

Raise the `syncrate` to 80 for all of the VDisks currently not synchronized (Example 7-18).

Example 7-18 Raise syncrate to 80

```
lsvdiskcopy -filtervalue sync=no -nohdr |
while read id vdisk copyid rest
do
   echo “Processing $vdisk”
   chvdisk -syncrate 80 $vdisk
done
```

Tip: Remember, raising the sync rate causes more I/O to be transferred, which can be an issue for a standard disk array.

Change primary in use to the new MDisk group

In Example 7-19, the primary is changed to the copy that was created in Example 7-17 on page 219.

Tip: Remember, all of these volumes must be in a sync state, as shown by the **`lsvdisk`** command output.

Example 7-19 Change VDisk mirror primary to copy in newgroupname

```
lsvdiskcopy -filtervalue mdisk_grp_name=newgroupname -nohdr |
while read id vdisk copyid rest
do
   echo Processing $vdisk
   chvdisk -primary $copyid $vdisk
done
```

Remove all the copies not primary

Example 7-20 removes all VDisk copies in the previous MDisk group.

Example 7-20 Remove VDisk copies

```
lsvdiskcopy -filtervalue mdisk_grp_name=prevmdiskgroup -nohdr |
while read id vdisk copyid rest
do
   echo "Processing rmvdiskcopy -copy $copyid $vdisk"
   rmvdiskcopy -copy $copyid $vdisk
done
```

Tip: Use extreme care when removing a storage enclosure from service. For example, in the case of IBM FlashSystem V9000, the AE2 enclosure should be unmanaged. The V840 equivalent state is to remove all the `mdisk` instances from **`prevmdiskgroup`**; these MDisks become unmanaged.

Create compressed mirrored copies of VDisks not currently mirrored

Example 7-21 looks for all VDisks that have a single copy and creates a mirrored compressed copy.

Example 7-21 Create compressed VDisk mirrors

```
lsvdisk -filtervalue  copy_count=1 -nohdr |
while read id vdiskname rest
do
   addvdiskcopy -mdiskgrp BB1mdiskgrp0 -autoexpand -rsize 50% -syncrate 30
   -compressed $id
done
Vdisk [0] copy [1] successfully created
Vdisk [1] copy [1] successfully created
Vdisk [2] copy [1] successfully created
Vdisk [3] copy [1] successfully created
Vdisk [4] copy [1] successfully created
Vdisk [5] copy [1] successfully created
```

Tip: From the CLI, issue the **`help addvdiskcopy`** command or look in IBM Knowledge Center for details of parameters for this command. All options that are available in the GUI can be issued from the CLI, which helps you more easily work with large numbers of volumes.

Turn on autoexpand for all offline volumes

During testing, an out-of-space condition was encountered with multiple mirrored copies. The condition is the result of the `autoexpand` option not being used when the volumes were mirrored. See Example 7-22.

Example 7-22 Activate autoexpand for all offline VDisk copies

```
lsvdiskcopy -filtervalue status=offline -nohdr | while read vid name copyid rest
do
   chvdisk -autoexpand on -copy $copyid $vid
done
```

Summary

This section presented simple scripts that can be used to automate tasks using the IBM FlashSystem V9000 command-line interface. These concepts can be applied to other commands including backup process, such as creating flash copies of volumes.

7.3.5 Recover lost superuser password

Use the following steps to reset the IBM FlashSystem V9000 superuser password to the factory default value:

1. Locate a blank USB stick and write a file named `satask.txt` into the root directory of the first partition of the USB stick. The file should contain the single **`satask resetpassword`** command.
2. Plug the USB stick into a free USB port on an active AC2 or AC3 control enclosure.
3. Wait for the identification blue led to turn on then off.
4. Unplug the USB stick. The command output is written to the USB key in a file named `satask_result.html`. This is successful if no errors are returned (Figure 7-23).

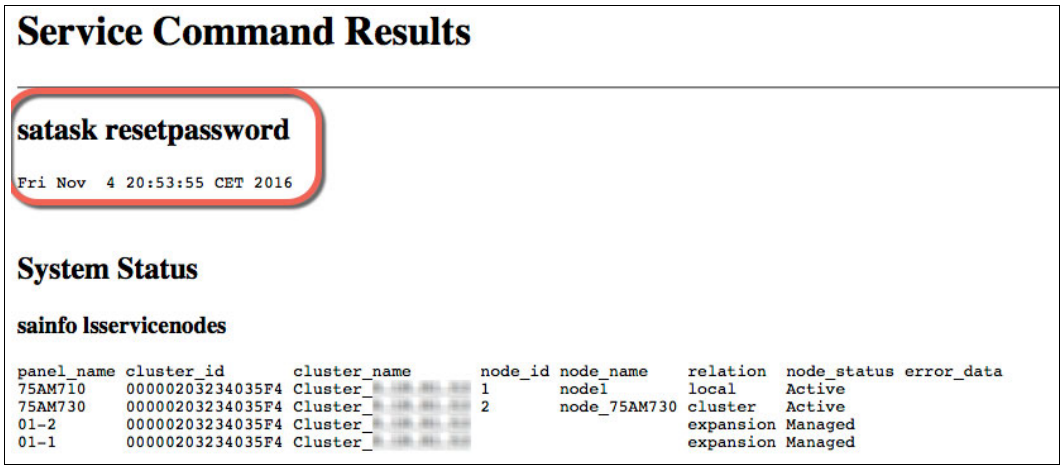


Figure 7-23 Command output in satask_result.html on USB key

Tip: The `satask_result.html` file also contains a report of the system status with several lines of output. The same system status can be obtained at any time by inserting a blank USB key into a AC2 or AC3 control enclosure.

5. Log in to the GUI by using superuser and `passw0rd`, the default password. A prompt guides you to change the default password.

7.3.6 Internal Fibre Channel switch maintenance

Most scalable building block configurations are purchased with IBM System Storage SAN48B-5 (2498-F48) switches that are used for storage controller cluster and AE2 expansion storage connections. These switches are private or internal to the product and are not attached to the customer SAN. These switches use open zoning and are initialized during installation as an IBM lab based services task.

After installation, maintenance of the internal switches is a customer responsibility.

Microcode updates

Fix central for the 2498-F48 switch directs you to the following Brocade Fabric OS website.

From there you can download a current level of microcode and the documentation to help you perform the update.

Example 7-23 shows how to check the firmware level from the switch CLI.

Example 7-23 Checking the firmware level with firmwareshow

```
IBM_2498_F48:FID128:admin> firmwareshow
Appl     Primary/Secondary Versions
----------------------------------------
FOS      v7.4.1c
         v7.4.1c
IBM_2498_F48:FID128:admin>
```

Troubleshooting and monitoring

The CLI and GUI have troubleshooting and monitoring capabilities.

Health monitor

The Monitoring and Alerting Policy Suite (MAPS) is a health monitor supported with Fabric OS 7.2.0 or later. The **`mapsdb --show`** command shows you the health of the switch, as shown in Example 7-24.

Example 7-24 CLI command to view the health of the switch

```
IBM_2498_F48:FID128:admin> mapsdb --show

1 Dashboard Information:
========================

DB start time:                  Thu Jun 23 22:34:59 2016
Active policy:                  dflt_conservative_policy
Configured Notifications:       None
Fenced Ports :                  None
Decommissioned Ports :          None
```

```
Quarantined Ports :             None

2 Switch Health Report:
=======================

Current Switch Policy Status: HEALTHY


3.1 Summary Report:
===================

Category                |Today                    |Last 7 days              |
-------------------------------------------------------------------------------
Port Health             |No Errors                |Out of operating range   |
BE Port Health          |No Errors                |No Errors                |
Fru Health              |In operating range       |In operating range       |
Security Violations     |No Errors                |No Errors                |
Fabric State Changes    |No Errors                |No Errors                |
Switch Resource         |In operating range       |In operating range       |
Traffic Performance     |In operating range       |In operating range       |
FCIP Health             |Not applicable           |Not applicable           |
Fabric Performance Impact|In operating range      |In operating range       |
```

Troubleshooting port errors

If you suspect or want to check for port errors, you can use the **porterrshow** command. It can be helpful to clear the counters during a workload and to check the command after a few minutes to determine if any ports are seeing increasing errors. To clear the counters for all ports on the switch use the **portstatsclear -i 0-47** command.

Example 7-25 shows the **porterrshow 8-16** command to view error counters for ports 8-16.

Example 7-25 CLI command to view port error counters

```
IBM_2498_F48:FID128:admin> porterrshow 8-16
          frames      enc    crc    crc    too    too    bad    enc    disc   link   loss   loss   frjt   fbsy  c3timeout     pcs
        tx     rx      in    err    g_eof  shrt   long   eof    out    c3     fail   sync   sig                  tx     rx      err
  8:   39.4m  39.1m   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
  9:   13.0k  12.4k   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 10:   39.5m  39.2m   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 11:   13.0k  12.4k   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 12:   39.3m  40.1m   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 13:    1.3k 781      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 14:   39.6m  39.8m   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 15:   13.0k  12.4k   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
 16:   39.4m  39.6m   0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
```

Viewing port performance

Use the **portperfshow** command to view the throughput per port on the switch. This can be useful to ensure that the load is balanced across the ports.

Example 7-26 shows restricting of the ports monitor to 8 - 16 and setting the refresh duration to 5 seconds with the **portperfshow 8-16 -t 5** command.

Example 7-26 CLI command to view port performance

```
IBM_2498_F48:FID128:admin> portperfshow 8-16 -t 5
  8       9       10      11      12      13      14      15      16       Total
============================================================================
 217.6k   0     281.5k    0     259.5k    0     248.0k    0     344.9k    1.3m
  8       9       10      11      12      13      14      15      16       Total
============================================================================
 180.0k   0     229.0k    0     211.4k    0     244.0k    0     209.4k    1.0m
  8       9       10      11      12      13      14      15      16       Total
```

Figure 7-24 shows how to view the port performance using the GUI capabilities of the internal switch.

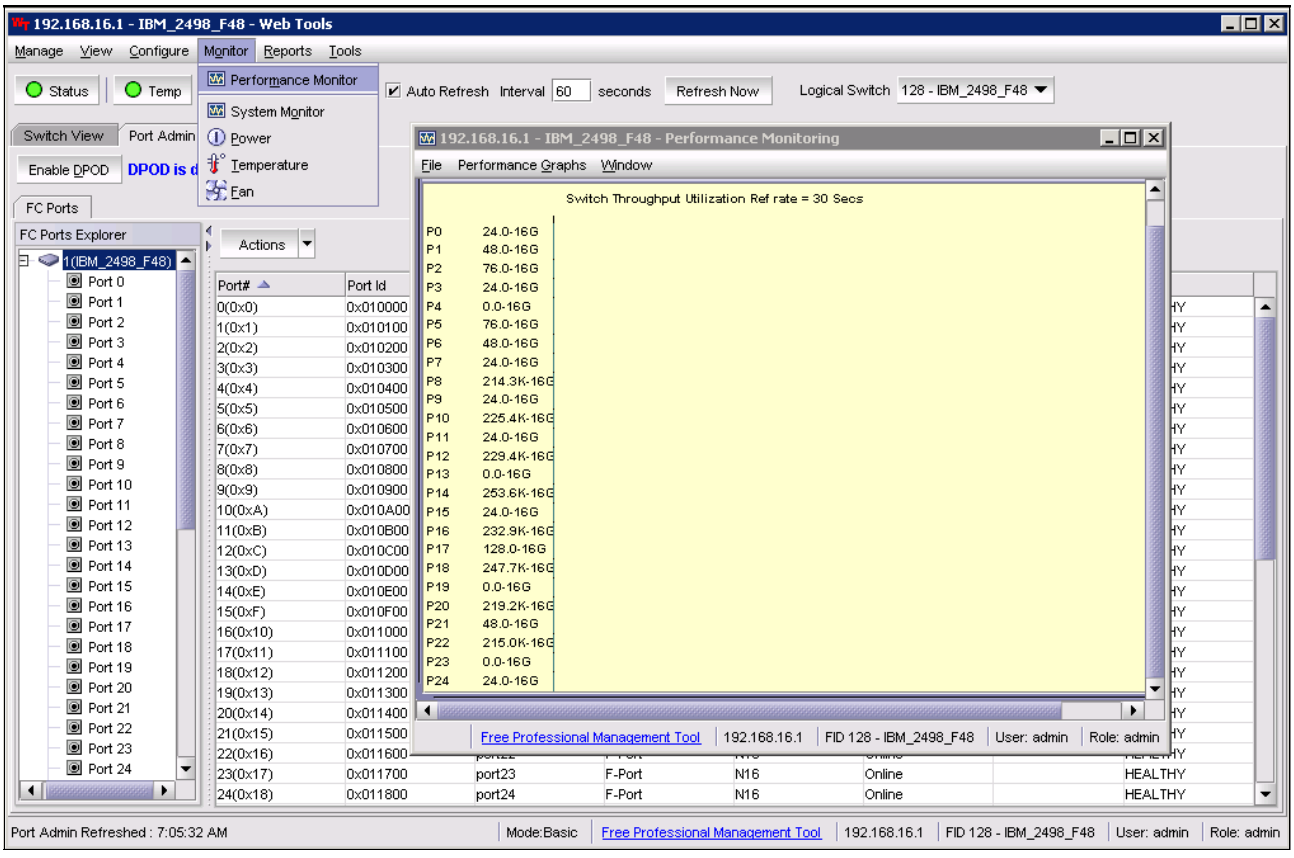


Figure 7-24 Internal switch Performance Monitor

7.3.7 Back up IBM FlashSystem V9000 configuration

Before making major changes to the IBM FlashSystem V9000 configuration, be sure to save the configuration of the system. By saving the current configuration, you create a backup of the licenses that are installed on the system. This can assist you in restoring the system configuration. You can save the configuration by using the **svcconfig backup** command.

The next two steps show how to create a backup of the configuration file and to copy the file to another system:

1. Log in to the cluster IP using an SSH client and back up the FlashSystem configuration:

   ```
   superuser> svcconfig backup
   ..........................................................
   CMMVC6155I SVCCONFIG processing completed successfully
   ```

2. Copy the configuration backup file from the system. Using secure copy, copy the following file from the system and store it:

   ```
   /tmp/svc.config.backup.xml
   ```

   For example, use **pscp.exe**, which is part of the PuTTY commands family:

   ```
   pscp.exe superuser@<cluster_ip>:/tmp/svc.config.backup.xml .
   superuser@ycluster_ip> password:
   svc.config.backup.xml     | 163 kB | 163.1 kB/s | ETA: 00:00:00 | 100%
   ```

   The use of the CLI is described in 7.3, “Command-line hints” on page 205.

Tip: This process saves only the configuration of the system. User data must be backed up by using normal system backup processes.

7.3.8 Using the IBM FlashSystem V9000 Software Upgrade Test Utility

In preparation for upgrading firmware on an IBM FlashSystem V9000, be sure to run the Software Upgrade Test Utility before any upgrade. This step ensures that your system configuration is supported and identifies any potential issue during your upgrade change window.

Overview of the Software Upgrade Test Utility

You can download this small utility from IBM Fix Central (see 7.5.5, “Downloading from IBM Fix Central” on page 240).

The utility is run on the IBM FlashSystem V9000 before a firmware upgrade. The purpose of the utility is to check for known issues and system configurations that might present a problem for the upgrade, and to warn you of conditions that might need to be corrected before running the upgrade. It can be run as many times as needed. The utility is run automatically as part of the GUI upgrade process. Alternatively, it can be run stand-alone, to assess the readiness of a system for upgrade as part of the upgrade planning process.

In releases before 7.6, the Software Upgrade Test Utility is run automatically before upgrading. Later releases provide the capability to run the utility from the GUI stand-alone. See 5.4.1, “V9000 control enclosure and internal storage firmware update” on page 124 for more details. This section is provided to document past releases where this capability was not possible. In earlier IBM Spectrum Virtualize firmware releases, it was optional, but strongly suggested.

When an upgrade is initiated using the IBM FlashSystem web management GUI, you are prompted to download the utility and firmware from Fix Central. Most users prefer to do this before the upgrade and select the files, as shown in Figure 7-25, during the upgrade process. If you click **Update**, the utility runs. If successful, the firmware upgrade process runs.

Tip: During the IBM FlashSystem V9000 GUI upgrade process, the upgrade proceeds immediately after the Software Upgrade Test Utility completes successfully.

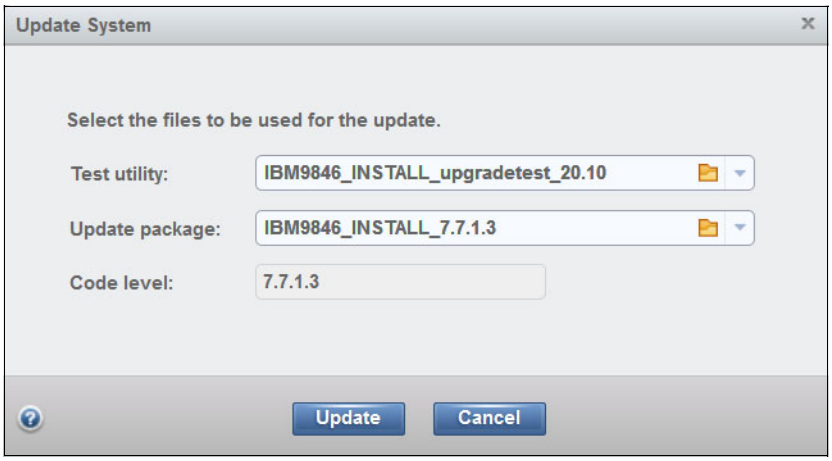


Figure 7-25 Update System

Tip: Always ensure that the upgrade package for the IBM FlashSystem V9000 is the correct package, as indicated by the IBM9846 prefix on the file name.

A preferred practice is for firmware upgrades to be initiated using the GUI so that the Software Upgrade Test Utility is uploaded and run during the upgrade process. The upgrade process is stopped if any issue is identified. If that happens, examine the output of the utility before proceeding. If the utility provides any warnings, correct them before continuing the upgrade.

Upgrades can also be run by using the **`applysoftware`** command using the CLI.

Using the IBM FlashSystem V9000 Software Upgrade Test Utility from the command line

The installation and use of this utility is nondisruptive and does not require any node to be restarted, so there is no interruption to host I/O. The utility will be installed only on the current configuration node. To install and run the utility, complete the following steps:

1. Copy the utility to the `/upgrade` directory on the IBM FlashSystem V9000 using a secure copy utility, such as Secure Copy Protocol (SCP) or **`pscp.exe`**:

   ```
   pcsp <test_utility_filename> superuser@<cluster_ip_address>:/upgrade
   ```

2. Install the utility:

   ```
   applysoftware -file <test_utility_filename>
   ```

3. Run the test utility:

   ```
   svcupgradetest -v 7.7.1.3
   ```

The output is displayed (Example 7-27).

Example 7-27 Output from test utility

```
svcupgradetest version 20.10

Please wait, the test may take several minutes to complete.


****************** Warning found ******************

The audit log of the system was cleared to prepare for the upgrade.
The old audit log can be found in /dumps/audit/ on this node.

Results of running svcupgradetest:
==================================

The tool has found 0 errors and 1 warnings.
```

Tip: There is a message "`****************** Warning found ******************`" in the output for each warning detected.

4. You can rerun the utility (step 3 on page 226) after it is installed. Installing it again is unnecessary. Installing a new version overwrites the old version.

7.3.9 Secure erase of data

Some clients, especially in the healthcare sector, are concerned about data confidentiality. IBM FlashSystem V9000 uses encryption to secure data. If you have a license for IBM FlashSystem V9000 encryption, you can prevent unauthorized access to IBM FlashSystem data.

Secure erase of IBM FlashSystem V9000 MicroLatency modules

Important: Deleting the IBM FlashSystem V9000 encryption key prevents any access to the data on IBM FlashSystem V9000 when the encryption feature is enabled.

Flash modules can be securely decommissioned by using the **`chdrive -task erase`** command. IBM has certified this erasure process; contact your IBM representative or IBM Business Partner for documentation regarding this process.

Example erasure procedure

These steps can be used to decommission and entire IBM FlashSystem V9000 enclosure.

Attention: This procedure is designed to securely destroy data. There is no recovery, so be careful to identify the correct enclosure.

Complete these steps:

1. Start by confirming the enclosure ID to be erased using the IBM FlashSystem V9000 GUI. Select **Pools** → **MDisks by Pool** (Figure 7-26).

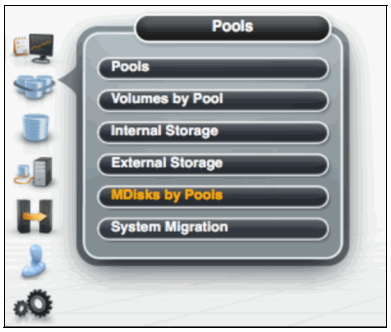


Figure 7-26 Select MDisks by Pools

2. Right-click an MDisk and select **Drives** (Figure 7-27).

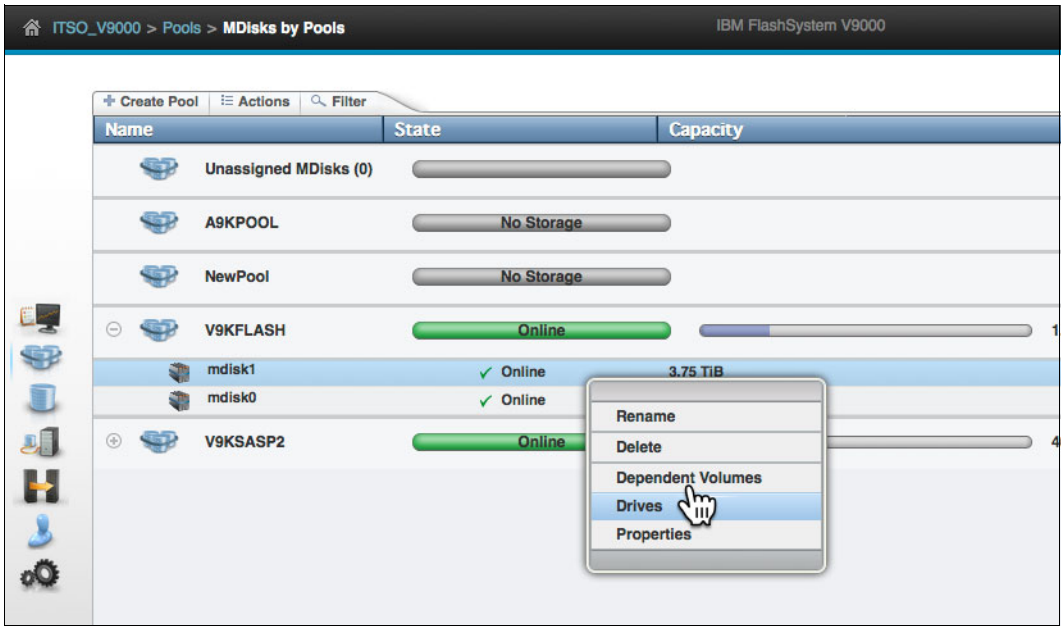


Figure 7-27 Select Drives

3. In the Member Drives panel, note that the Enclosure ID to be erased is enclosure 1 (Figure 7-28).

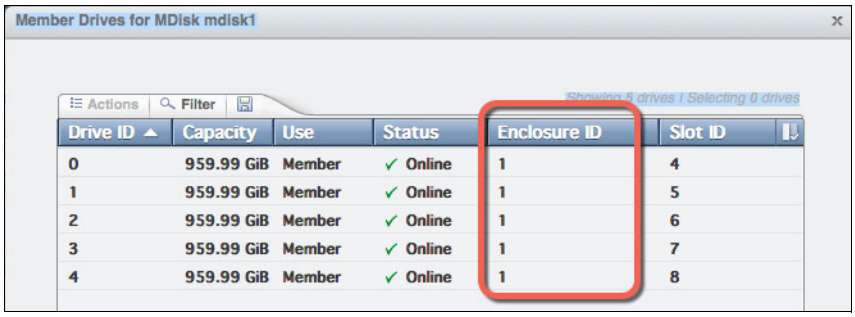


Figure 7-28 Identify the enclosure

4. Right-click the managed disk (mdisk1) and select **Delete** (Figure 7-29).

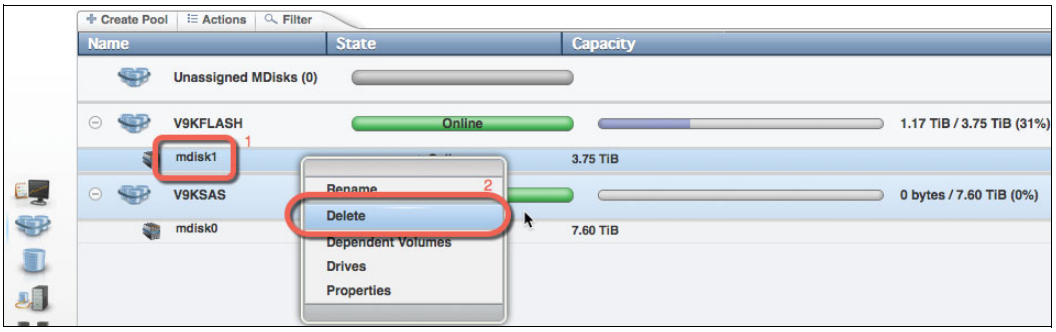


Figure 7-29 Delete the MDisk from the pool

)

Tip: The pool in this example does not have any VDisks allocated. If VDisks are present, there would have to be other managed disks in the pool with enough free space to hold the VDisks. The operation is not allowed without adequate space.

5. The drives that made up the MDisk are now in candidate state (1) for Enclosure ID 1 (2) shown in Figure 7-30.

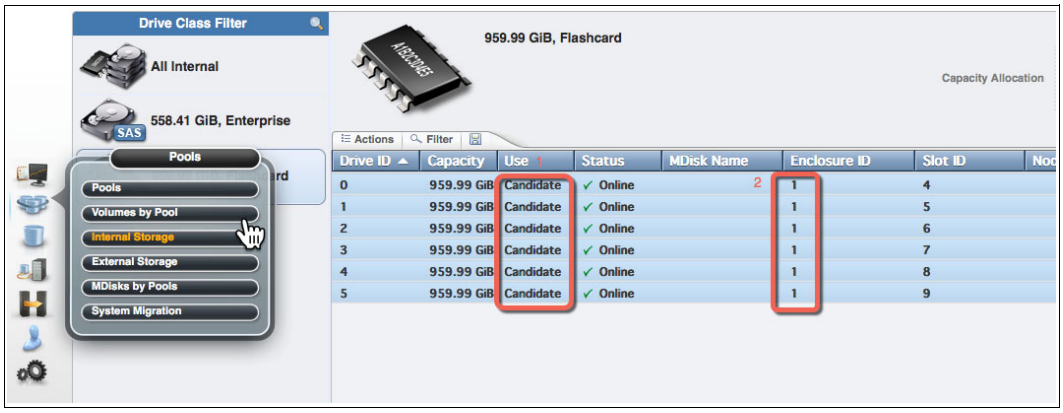


Figure 7-30 Drives in Candidate state for Enclosure ID 1

6. This series of commands (Figure 7-31) securely erases the drives for Enclosure ID 1. The loop (highlighted by number 1) filters just enclosure iD 1. The **`lsdrive`** command (2) shows the drive status as `offline` while the erase process occurs.

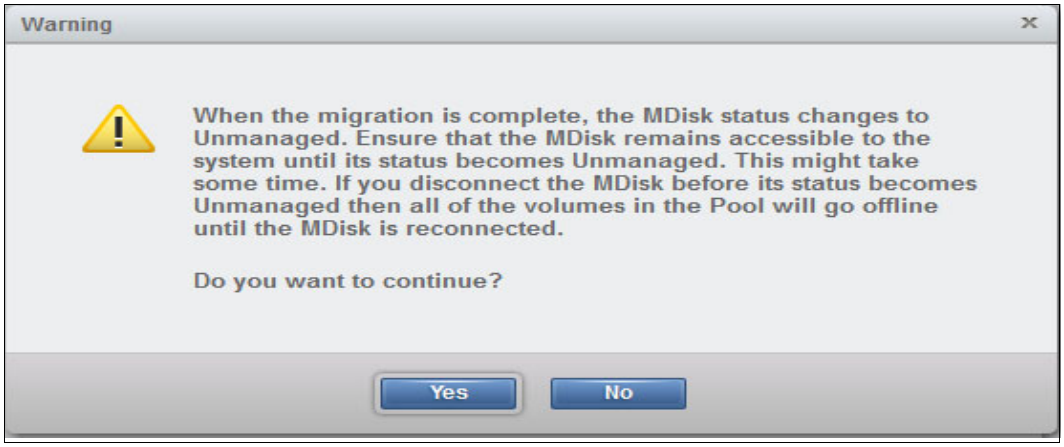


Figure 7-31 Secure erase of the drives in Enclosure ID 1

7. Confirm that the erase operation is complete by verifying that the drives are `online` (Figure 7-32).

```
IBM_FlashSystem:ITSO_V9000:superuser>lsdrive -filtervalue enclosure_id=1
id status error_sequence_number use       tech_type capacity mdisk_id mdisk_name member_id enclos
ure_id slot_id node_id node_name health_state encrypted rekey auto_manage drive_class_id
0  online                       candidate flashcard 960.0GB                                  1
         4                       good         no        no    inactive
1  online                       candidate flashcard 960.0GB                                  1
         5                       good         no        no    inactive
2  online                       candidate flashcard 960.0GB                                  1
         6                       good         no        no    inactive
3  online                       candidate flashcard 960.0GB                                  1
         7                       good         no        no    inactive
4  online                       candidate flashcard 960.0GB                                  1
         8                       good         no        no    inactive
5  online                       candidate flashcard 960.0GB                                  1
         9                       good         no        no    inactive
```

Figure 7-32 Drives have returned to online status

8. Return to the GUI and use the system panel to complete removing the storage enclosure. Confirm that you are working with the correct enclosure by hovering over the device, as shown in Figure 7-33.

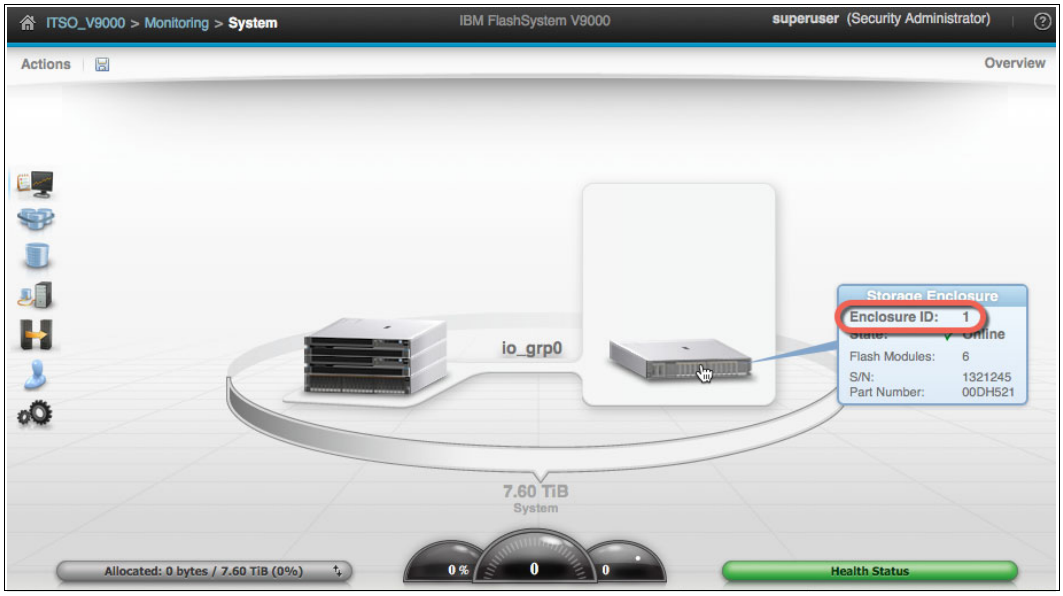


Figure 7-33 Confirm the correct enclosure

9. Complete the removal of the storage enclosure as shown in Figure 7-34:

   a. Select the enclosure (1).
   b. Right-click over the exploded view (2).
   c. Select **Remove Expansion Enclosure** (3).

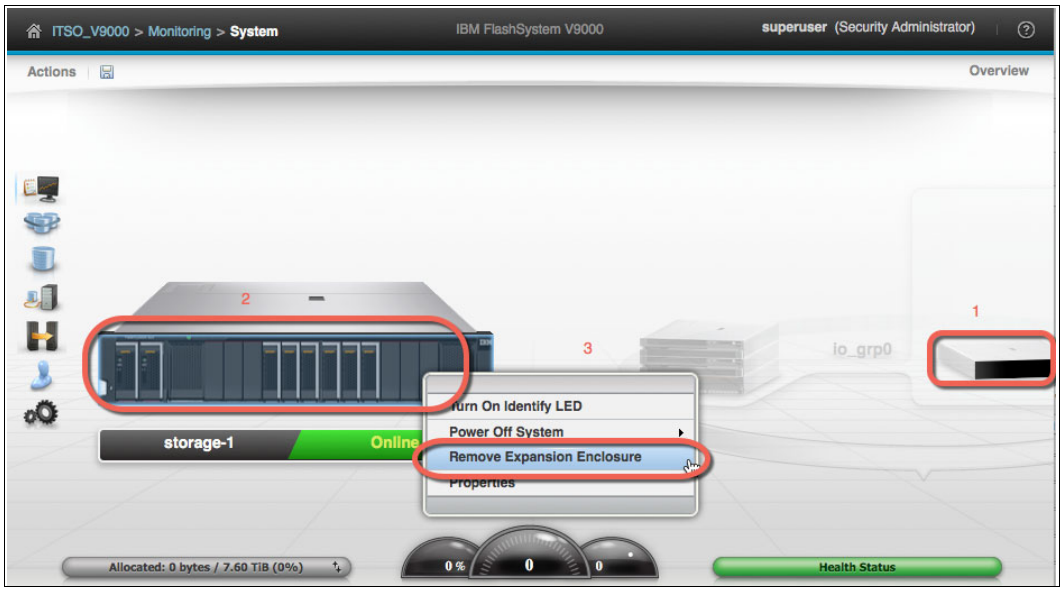


Figure 7-34 Complete removal of storage enclosure

TIP: If the GUI does not show the menu to remove the expansion enclosure, refresh your browser by re-logging into the system.

10. The storage enclosure is now removed from the system. The wizard (Figure 7-35) will lead you through deinstalling the storage enclosure.

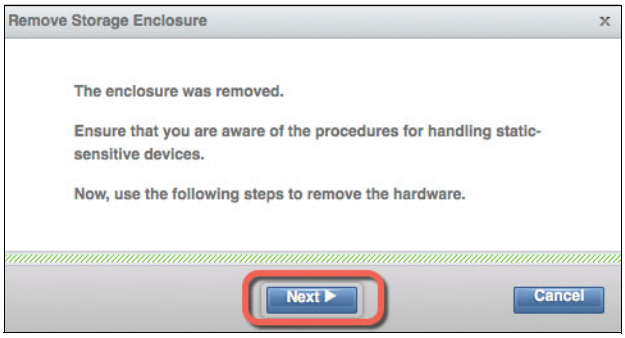


Figure 7-35 Remove storage enclosure

Tip: The Secure Erase process is currently only available for the IBM FlashSystem V9000 AE2 storage enclosure. In version 7.7.x.x, adding and removing the storage enclosure is automated in the GUI.

Erasing flash drives is not a normal operation. By using the encryption features of the IBM FlashSystem V9000, the erasure can be avoided because the data cannot be read, even on a failed flash module. This hint was provided for customers that require an extra layer of security and certification.

7.4 Call Home process

IBM encourages all clients to take advantage of the following settings to enable you and IBM to partner for your success. With the Call Home feature enabled, your system is effectively monitored 24 x 7 x 365. As an IBM client, you can enjoy faster response times, faster problem determination and effectively reduced risk over an unmonitored system. In the future, IBM plans to use inventory report data to directly notify clients who are affected by known configuration or code issues.

While enabling Call Home reporting, IBM encourages clients to also enable inventory reporting in order to take advantage of this future offering. For a more detailed explanation, followed by steps to configure, see 5.1.3, “Notifications menu” on page 122. The configuration setup is a simple process and takes several minutes to complete.

7.4.1 Call Home details

The Call Home function opens a service alert if a serious error occurs on the system, automatically sending details of the error and contact information to IBM Service personnel. If the system is entitled for support, a problem management record (PMR) is automatically created and assigned to the appropriate IBM Service personnel.

The information provided to IBM in this case might be an excerpt from the Event Log containing the details of the error and client contact information from the system. This enables IBM Service personnel to contact you and arrange service on the system, which can greatly improve the speed of resolution by removing the need for you to detect the error and raise a support call.

7.4.2 Email alert

Automatic email alerts can be generated and sent to an appropriate client system administrator or distribution list. This is effectively the same as Call Home, but you can be additionally notified about error, warning, and information messages when they occur. Also, you can receive inventory emails (see 7.4.3, “Inventory” on page 233).

You can view IBM Knowledge Center documentation for your specific IBM FlashSystem V9000 product to determine whether a particular event is classified as error, warning, or informational. Look for the Notification type for each error to determine which you want to be notified for. Because you can customize this, based on the individual, maximum flexibility exists.

7.4.3 Inventory

Rather than reporting a problem, an email is sent to IBM that describes your system hardware and critical configuration information. Object names and other potentially sensitive information, such as IP addresses, are not sent.

IBM suggests that the system inventory be sent on a one-day or seven-day interval for maximum benefit.

7.5 Service support

Understanding how support issues are logged is important information. This section describes support for the IBM FlashSystem V9000, including the IBM Technical Advisor role, Enterprise Class Support, support entitlement, registering components in the Service Request Tool, and calling IBM for support.

7.5.1 IBM Storage Technical Advisor

The IBM Storage Technical Advisor (TA) enhances end-to-end support for complex IT solutions. Each IBM FlashSystem V9000 includes an IBM TA for the initial hardware warranty period. This section describes the IBM TA program in general with specifics on how customers can work with their TA.

The TA service is built around three value propositions:

- Proactive approach to ensure high availability for vital IT services
- Client Advocate that manages problem resolution through the entire support process
- A trusted consultant for both storage hardware and software

Technical Advisors benefit customers by providing a consultant for questions on the IBM FlashSystem V9000. Most customers meet their TA during a Technical Delivery Assessment (Solution Assurance Meeting) before the initial installation.

After this initial meeting, the TA is the focal point for support related activities as follows:

- Maintains a support plan that is specific to each client. This support plan contains an inventory of equipment including customer numbers and serial numbers.
- Coordinates service activities, working with your support team in the background. Monitors progress of open service requests, escalation, expert consultation on problem avoidance.

- Communicates issues with customers, IBM Business Partners, and IBM Sales teams.
- Periodically reviews and provides reports about hardware inventories and service requests. This includes using call home information to provide customer reports on the state of the customer systems.
- Provides oversight of IBM support activities to help companies anticipate and respond to new problems and challenges faster.
- Provides proactive planning, advice, and guidance to improve availability and reliability.

The IBM Storage Technical Advisor is an effective way to improve total cost of ownership (TCO) and free up customer resources. Customers have options to extend the Technical Advisor service beyond the initial hardware warranty using IBM Technical Support Services (TSS) offerings.

Contact your IBM Sales Team or IBM Business Partner for details.

7.5.2 Enterprise Class Support

IBM Enterprise Class Support (ECS) delivers improved response times, hardware and software installation assistance, onsite code upgrades and service coordination across IBM. This enhanced support is available to IBM FlashSystem V9000 customers with the 3-year warranty machine type 9848. This service extends the Technical Advisor service described in 7.5.1, “IBM Storage Technical Advisor” on page 233.

During the Enterprise Class Support warranty period you receive these items:

- Priority Access

  IBM Direct Access Code (DAC) allows you to call IBM and access a specialized team that serves as the initial point of contact for any issues that arise.

- Priority Response

  IBM helps you get to the root of issues faster. A support team, available 24 x 7, returns your call within 30 minutes of receiving the issue notification.

- Software updates

  IBM will coordinate and complete up to six software updates over the warranty period (requires applicable Software Maintenance Agreement, SWMA). This provides you the option of having an IBM System Support Representative (SSR) perform upgrades. Customers were responsible for this task in the past.

- Services coordination

  Each customer is assigned an Account Advocate in addition to the Technical Advisor. These technical experts coordinate service activities working closely with the customer and IBM SSR throughout the 3 year warranty period.

IBM Enterprise Support is an evolving service designed to assist you with the support of your storage products. IBM FlashSystem V9000 with release 7.8 includes the optional capability for remote support. As with call home, by choosing to enable this capability, you can benefit from Enterprise Class Support as features are developed to enhance the support received.

7.5.3 How an IBM FlashSystem V9000 is entitled for support

IBM FlashSystem V9000 systems consist of various hardware and software components, each carrying their own unique requirements for proper entitlement. IBM FlashSystem V9000 systems consist of at least three unique hardware machine types (models) for each building block. Each building block carries its own serial number. Warranty and maintenance entitled software, requires customer ID, product description, and storage enclosure serial number to properly entitle it.

Tip: Customer ID and customer number are the same. The customer number is included in the customer support plan you receive from your Technical Advisor.

Calling IBM Support

Consider this information when you call for support:

- For problems known to be hardware-related, place calls against the affected 9846, 9848-AC2, or AE2 machine type and serial number. Using the correct machine type, model, and serial number avoids service delays.

  Note: Most hardware support tickets are opened by Call Home events. However, if there are effects because of component failure, you can open an appropriate severity support ticket independently.

- For software problems, navigate through the Automated Voice Response for *Software* and provide customer ID, product description, and 9846 or 9848-AE2 storage enclosure, plus serial number.
- If you are unsure whether the issue is hardware or software, call in for *Storage Support* (option 3 is US only). Provide customer ID, product description, and 9846 or 9848-AE2 storage enclosure.

Scenario 1: Host connectivity issue

The customer is experiencing difficulty with attaching a host to the IBM FlashSystem V9000 AC2 control enclosures. The customer opens a *Software* issue (bullet item 2 in “Calling IBM Support” on page 235) against the IBM FlashSystem V9000 AE2 storage enclosure. The customer describes that the issue is the host attachment to the controller enclosure.

Scenario 2: Performance

The customer reports that performance of the IBM FlashSystem V9000 solution is not meeting expectations. The customer opens a *Storage Support* issue (bullet item 3 in “Calling IBM Support” on page 235) against the IBM FlashSystem V9000 AE2 storage enclosure. The customer can save time by uploading snaps from both controller AC2 and storage enclosure AE2 after the PMR number is obtained.

Scenario 3: Hardware issue

The customer reports that email alerts indicate a failed hard disk in the AC2 control enclosure. The customer opens a *hardware* issue (bullet item 1 in “Calling IBM Support” on page 235) against the IBM FlashSystem V9000 AC2 controller serial number reporting the error. This is processed as a standard hardware failure.

7.5.4 Providing logs to IBM ECuRep

IBM Enhanced Customer Data Repository (ECuRep) is a secure and fully supported data repository with problem determination tools and functions. It updates problem management records (PMR) and maintains full data lifecycle management.

This server-based solution is used to exchange data between IBM customers and IBM Technical Support. Do not place files on or download files from this server without prior authorization from an IBM representative. The representative is able to provide further instructions as needed.

To use ECuRep, you need a documented problem management record (PMR) number either provided by the IBM support team with a *call home*, or issued by using the IBM Service Request tool on the IBM support portal.

IBM provides the service request (SR) problem submission tool (the link is highlighted in Figure 7-36) to electronically submit and manage service requests on the web. This tool replaces the Electronic Service Request (ESR) tool.

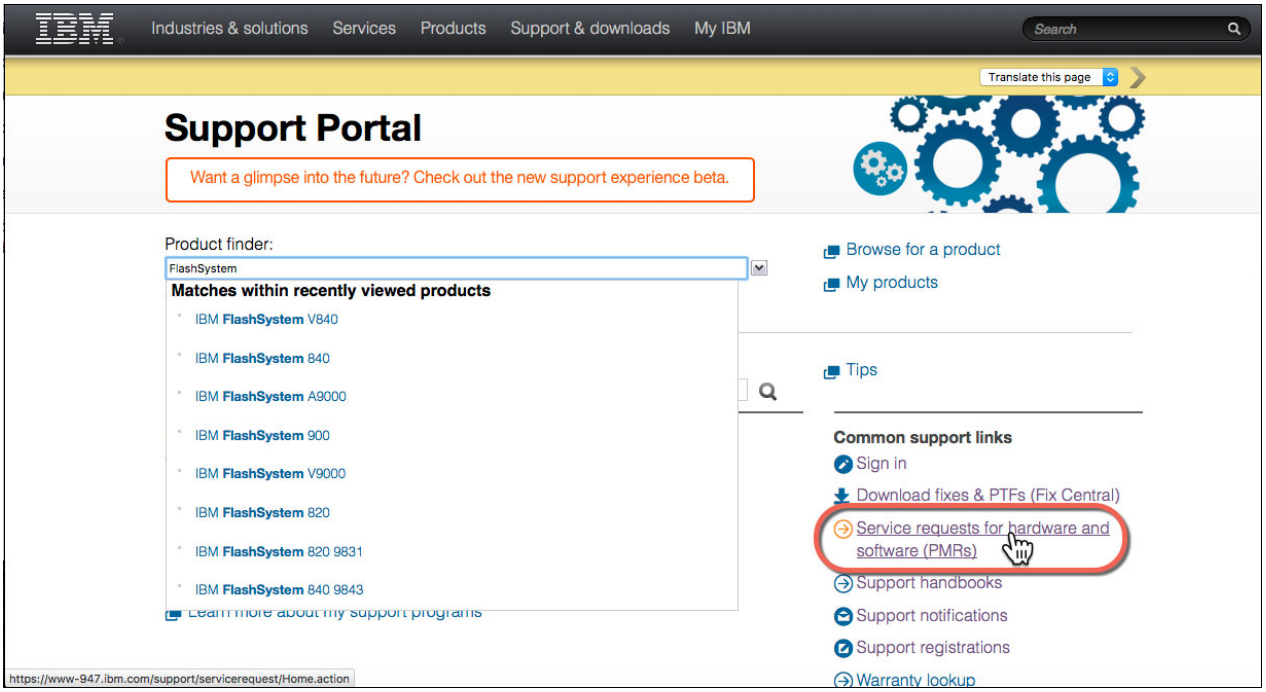


Figure 7-36 Service request link on Support Portal

To provide logs to IBM ECuRep, complete the following steps:

1. Go to the Enhanced Customer Data Repository (ECuRep) web page (Figure 7-37).

   This web page provides information about the repository, instructions for preparing files for upload, and multiple alternatives for sending data. For details, you can click **Help**.

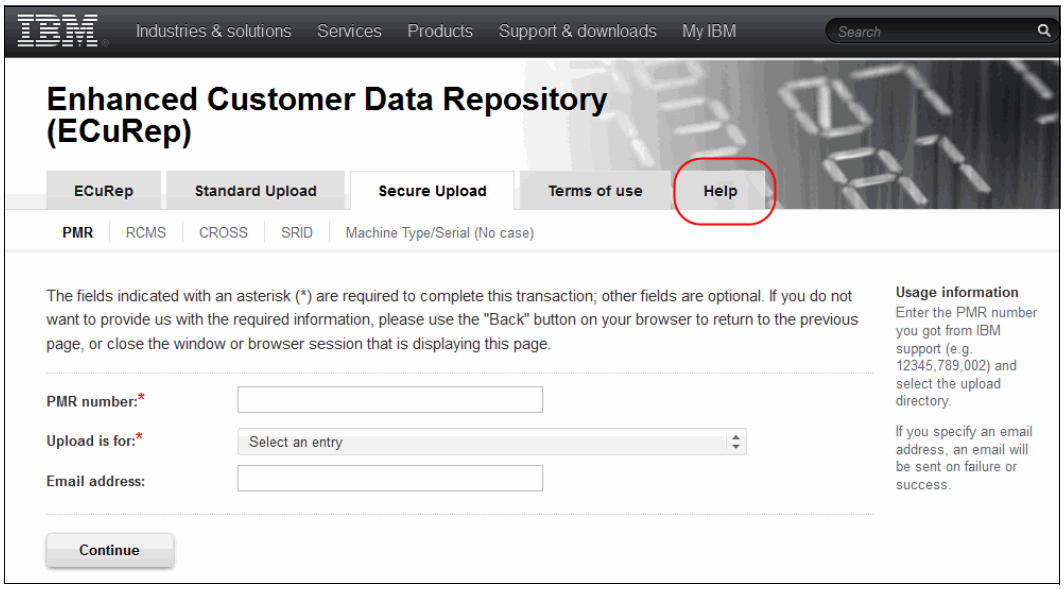


Figure 7-37 Select Help for detailed upload options

Tip: This system is connected to the IBM Problem Management Record. Support tickets are automatically updated, with the files uploaded and queued for an IBM support representative response.

2. IBM provides multiple options for uploading data. Review the options for sending data before you complete the PMR number. The following options are shown in Figure 7-38:
   - Notice in this description that the Send Data tab is selected.
   - As a way to upload a file you can select either FTP (1) or the Java utility (2). The Java utility is the most efficient method to upload file.
   - Select **Prepare data** tab (3) to see the details about file name conventions.
   - The HTTPS (4) option eliminates the file naming requirement.

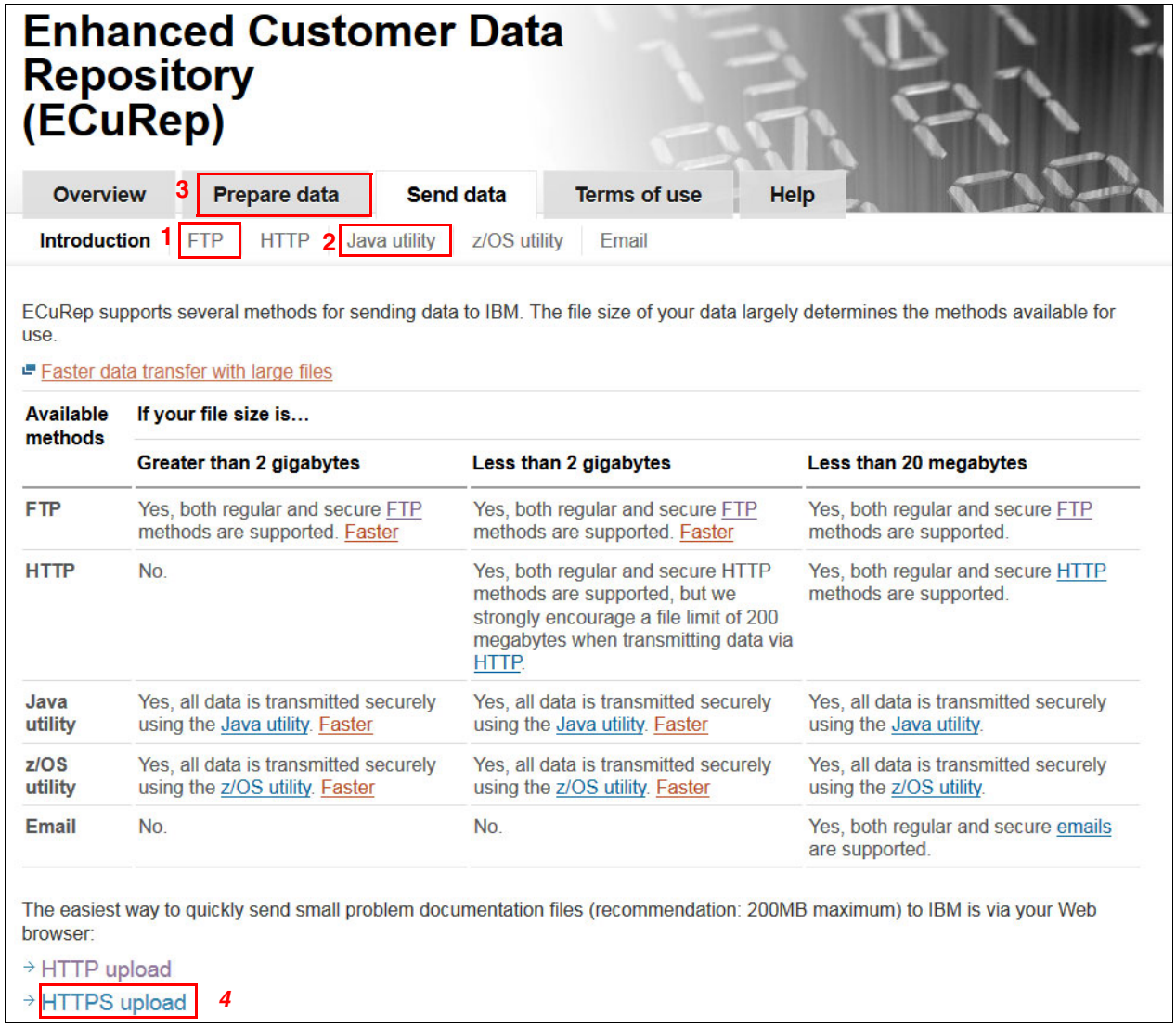


Figure 7-38 Options for sending data

3. Standard upload (Figure 7-39) is the default upload selection. Complete these fields and then click **Continue**:
   - PMR: Using the PMR number on this form accurately logs the files uploaded to the correct PMR.
   - Upload is for: Select **Hardware** for the IBM FlashSystem V9000.
   - Email address: (Optional) Provide your email address for a confirmation.

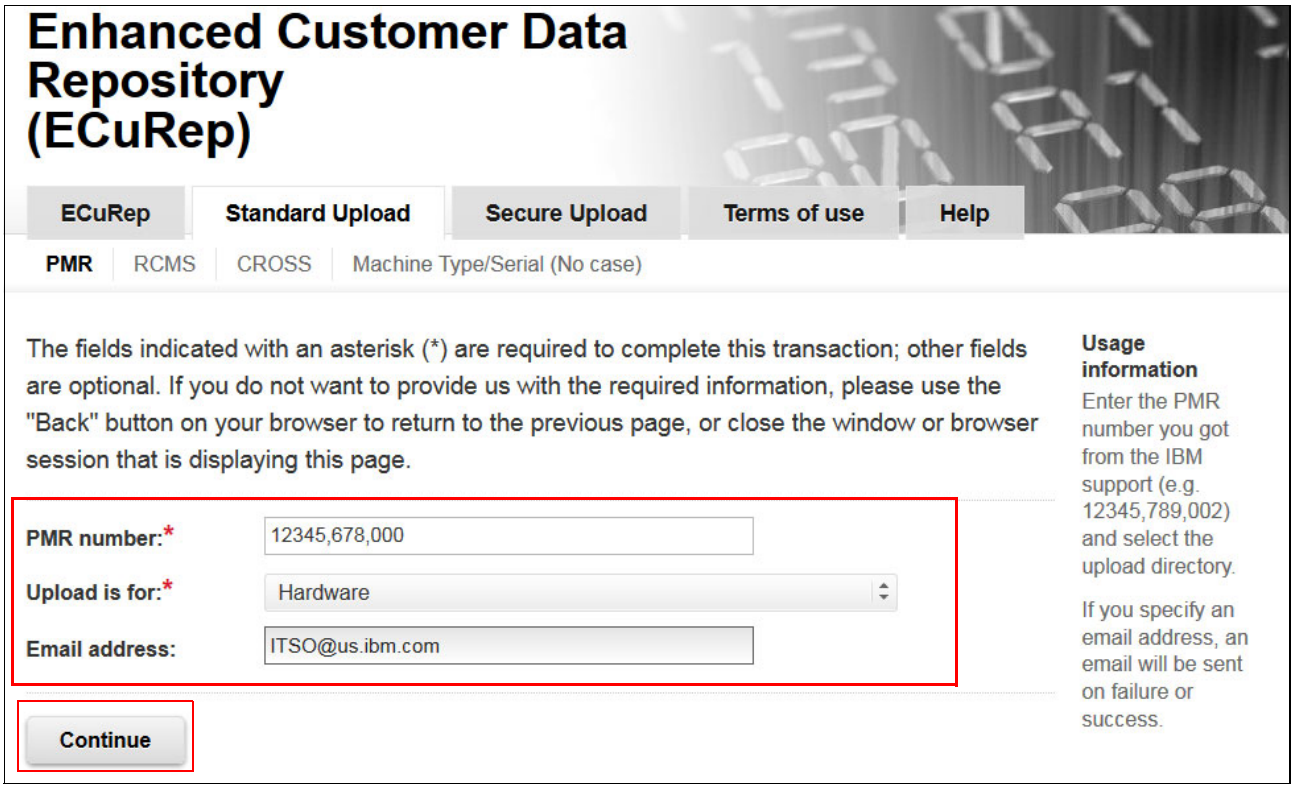


Figure 7-39 Using the HTTP option

4. The file selection panel opens (Figure 7-40). Select files and click **Upload**.

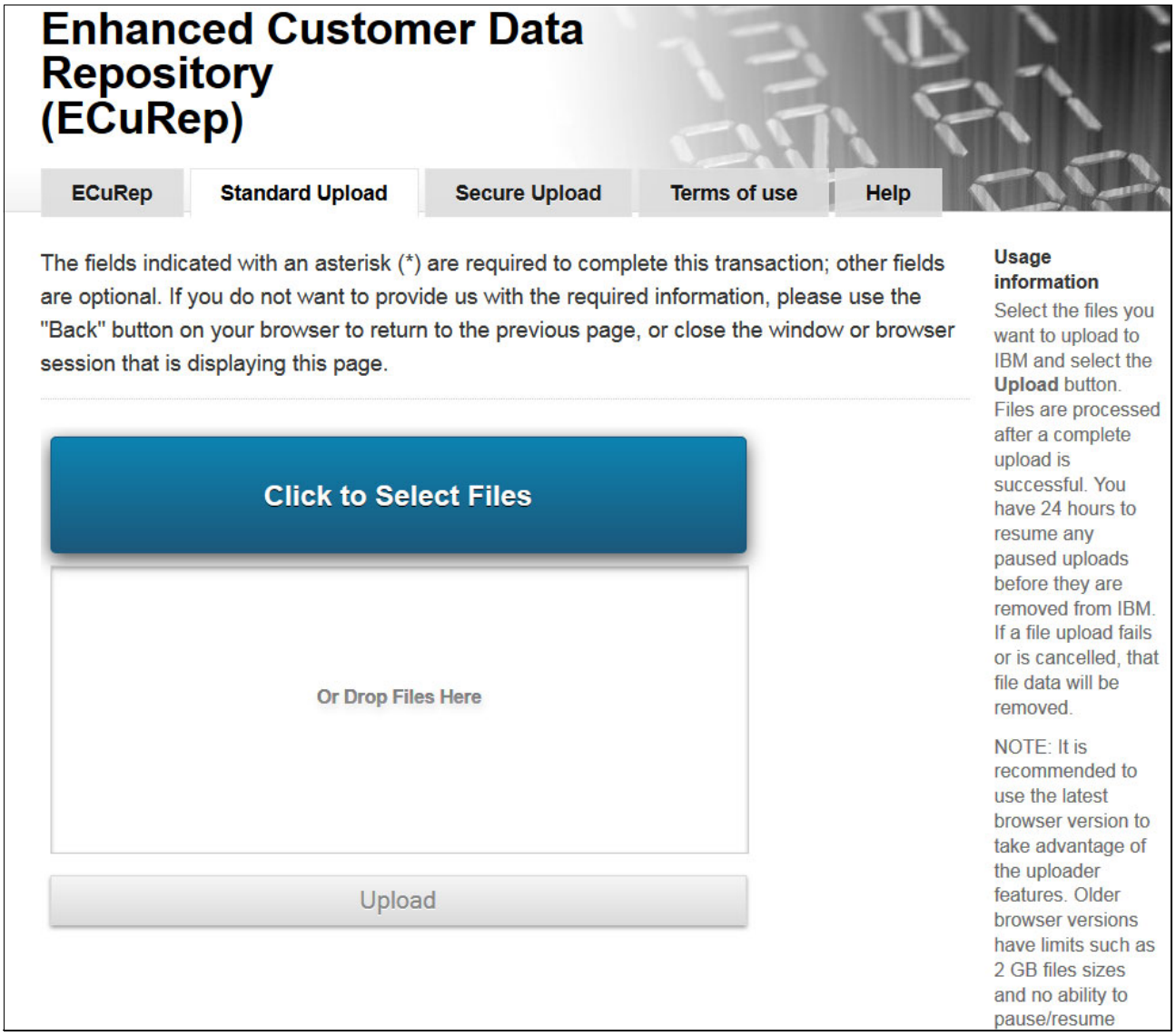


Figure 7-40 File Selection dialog

Tip: Most clients find this way the most effective method to upload logs. IBM suggests understanding the best method for your organization *in advance* and *documenting the process* to save precious time during a crisis.

7.5.5 Downloading from IBM Fix Central

Fix Central provides fixes and updates for your system's software, hardware, and operating system. Go to the IBM Fix Central web page.

If you are not looking for fixes or updates, go to IBM Passport Advantage® to download most purchased software products, or My Entitled Systems Support to download system software.

The following sections describe downloading code from IBM.

Using an IBMid

To use the IBM Fix Central website, you must have an IBMid. Your IBMid provides access to IBM applications, services, communities, support, online purchasing, and more. Additionally, your information is centralized so you can maintain it in a convenient and secure location. The benefits of having an IBMid will increase over time as more applications migrate to IBMid.

The login window is shown in Figure 7-41.

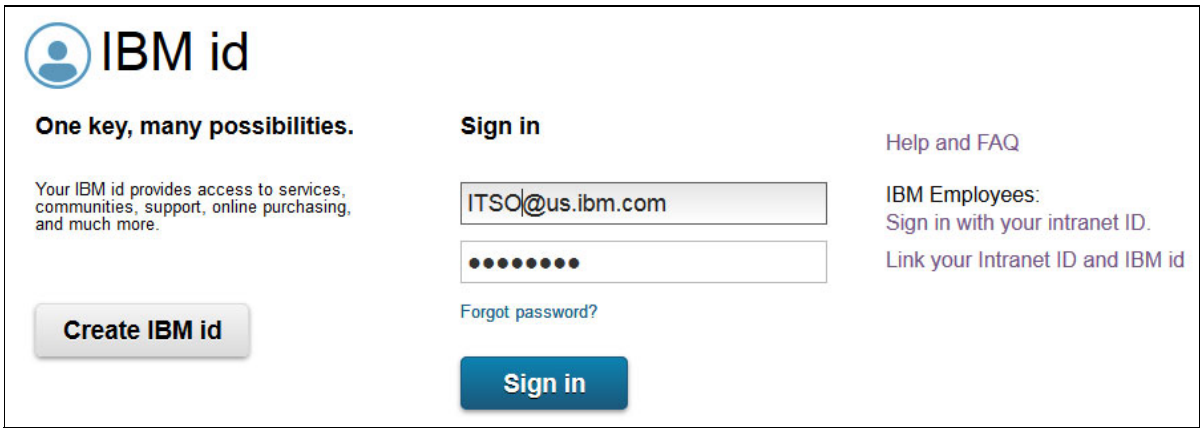


Figure 7-41 IBM id login window

Fix Central

The following steps document the current process for obtaining updates for your IBM FlashSystem V9000. This site is frequently updated based on customer feedback and as IBM documents field experience. It is a good practice to review the support information on this site on a regular basis.

1. After signing in with your IBMid, a page opens to the Support Portal (Figure 7-42). In the Product finder field, start typing `FlashSystem` (1). Select **IBM FlashSystem V9000** from the list (2). The IBM FlashSystem V9000 specific information is displayed.

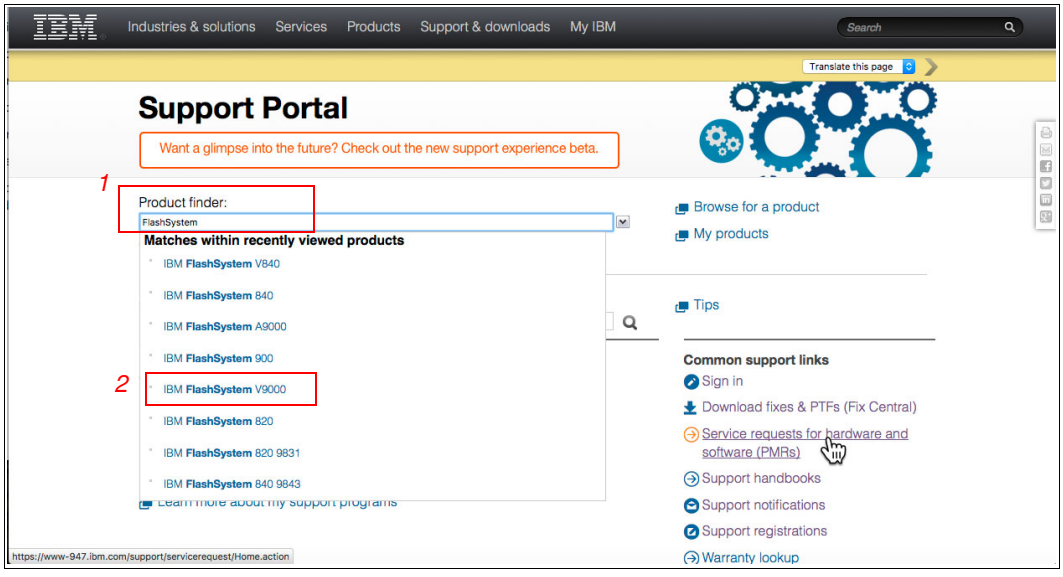


Figure 7-42 Support Portal

2. The product page provides capabilities to download product updates and review other product documentation. In Figure 7-43 new product firmware is displayed by selecting **Downloads (drivers, firmware, PTFs)**.

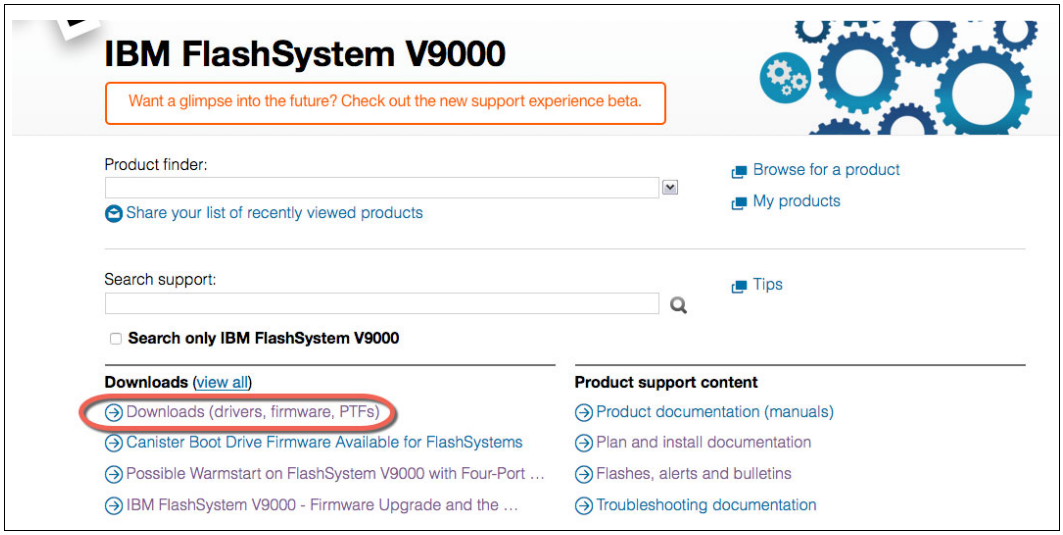


Figure 7-43 Download updates for IBM FlashSystem V9000

3. The Select fixes panel (Figure 7-44) provides a selection of release families. IBM FlashSystem V9000 Release 7.7 and later is required for IBM FlashSystem V9000 System using 9846/9848 model AC3 control enclosure and 12F, 24F expansion enclosures. IBM model 92F requires version 7.8.

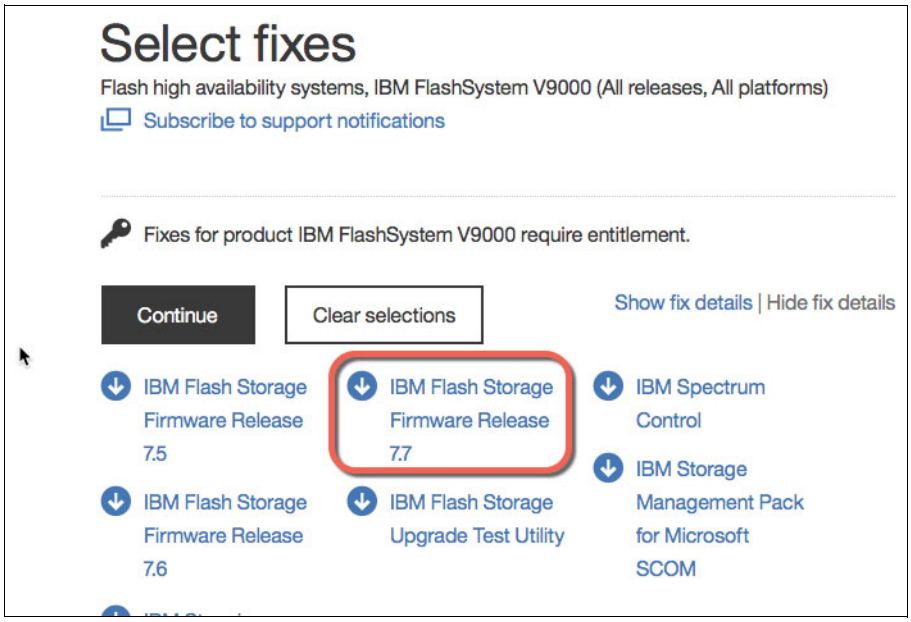


Figure 7-44 Select the Firmware Release family

4. The license agreement is presented. At this time entitlement is confirmed through the acceptance of this agreement shown in Figure 7-45 on page 243.

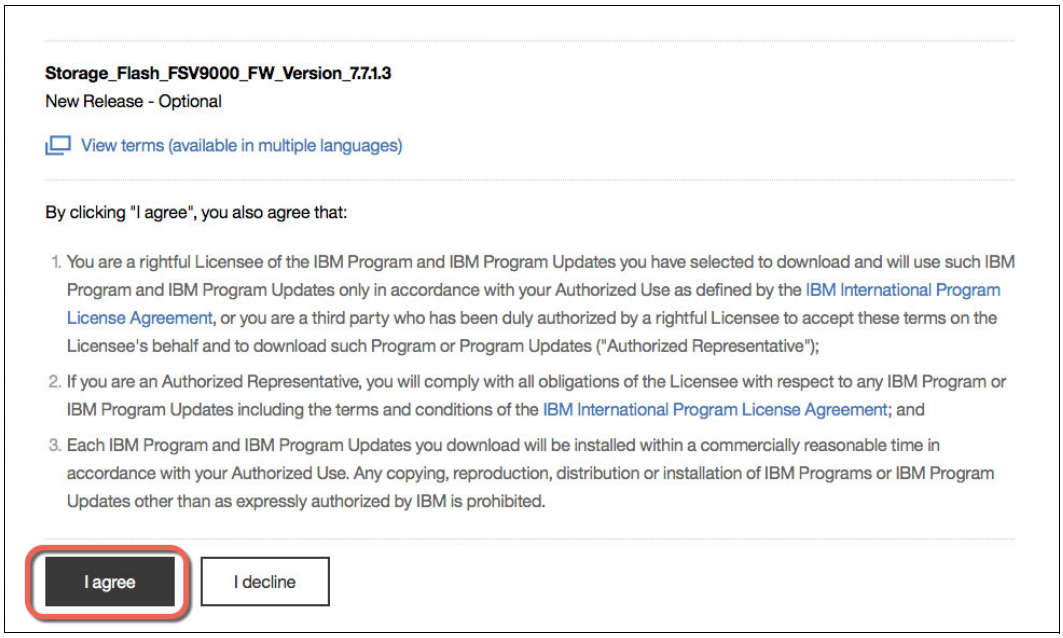


Figure 7-45 Support Portal agreement example

TIP: Previously the support portal prompted for the country, machine type and serial number to entitle the download. It is important to enter accurate information so that the entitlement lookup is successful.

5. Read the release notes in Figure 7-46 to determine the best fix pack for your environment, Select the fix pack (1) and click **Download now** (2) to initiate the file transfer. Download director is the preferred method for download because it is multi-threaded.

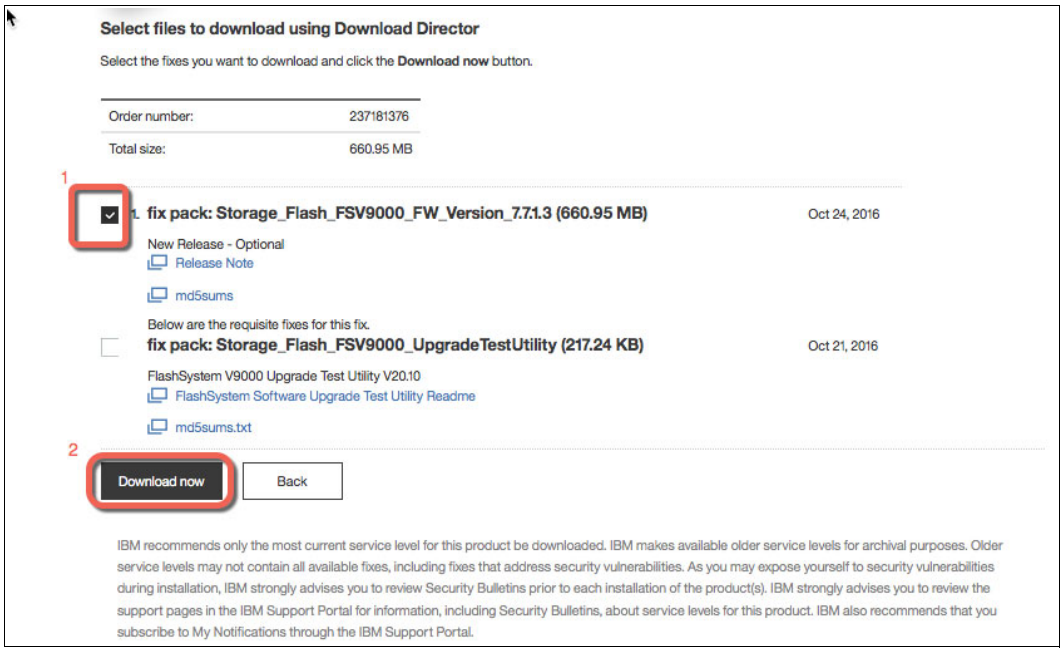


Figure 7-46 Download the best fix pack for your environment

Note: Always read the release notes. They often contain special instructions related to the upgrade that should be part of your planning.

6. To change download options, use the panel shown in Figure 7-47. This panel is available on the Download page.

Figure 7-47 Select Change download options to your preferred method

Summary

This section showed an example of downloading system firmware for the IBM FlashSystem V9000. The test and upgrade package is all that is needed to update all components of the system no matter how many building blocks are present.

A

Guidelines: Port utilization in an IBM FlashSystem V9000 scalable environment

To maximize the full potential of performance and low latency of IBM FlashSystem V9000 with IBM FlashCore technology in a scalable environment, several important items must be considered in configuring your environment. For example, host, FlashSystem storage, additional storage, intra-cluster, and optional remote copy ports must be properly designated and zoned in order to optimize performance and properly isolate the types of Fibre Channel traffic.

This appendix covers the following topics:

- Overview
- Comparison of port utilization methods
- Guidelines: The performance method
- Guidelines: The infrastructure savings method
- Guidelines: Zoning and pathing
- Summary
- Supported environments

A.1 Overview

Note: When an FlashSystem V9000 with AE2 storage enclosure is upgraded to 8.1.x or higher, AE2 will remain configured as internal storage, while AE3 storage enclosure are mapped to the V9000 control enclosure as external storage. This section will use the term *FlashSystem storage enclosure* when there is no difference between AE2 and AE3 storage enclosure from a configuration point of view.

IBM FlashSystem V9000 provides a flexible architecture for assigning port resources. Two primary methods of port utilization that can be implemented in a Fibre Channel environment are suggested depending on your needs. This appendix compares and provides guidelines for these two methods:

- IBM FlashSystem V9000 port utilization for *infrastructure savings*

  This method reduces the number of required customer Fibre Channel ports that are attached to the customer fabrics. This method provides high performance and low latency but performance might be port-limited for certain configurations. Intra-cluster communication and FlashSystem storage enclosure traffic occur over the internal switches.

- IBM FlashSystem V9000 port utilization for *performance*

  This method uses more customer switch ports to improve performance for certain configurations. Only ports that are designated for intra-cluster communication are attached to private internal switches. The private internal switches are optional and all ports can be attached to customer switches.

Random workloads can also experience performance benefits with the performance method because more host ports are available when compared with the infrastructure savings method. This benefit is more pronounced for a IBM FlashSystem V9000 solution that includes dedicated ports for remote copy.

Note: AC3 controller can have up to four 4-port 16 Gb cards. This appendix focuses on the three card configuration.

A.2 Comparison of port utilization methods

The *infrastructure savings* method has dedicated internal switches for the IBM FlashSystem V9000 storage enclosures connections and also intra-cluster communication with a reduced number of customer host facing ports. Eight ports per I/O group are available for storage traffic to the FlashSystem storage enclosures.

The *performance* method uses the customer fabric for all connections (with the option to use dedicated internal switches for intra-cluster communication). The ports have designated purposes based on fabric attachment, zone assignments, and port masking. This method provides shared-use ports that use the full bidirectional capabilities of Fibre Channel along with the V9000 excellent load-balancing capabilities across all available ports. Up to sixteen ports per I/O group are available for storage traffic to FlashSystem storage enclosures.

Both methods can designate host ports for remote copy and mirroring. The performance method is more efficient when ports are designated to remote copy.

Both methods support attachment to additional external storage. In both cases, zones in the customer fabric are required for attaching external storage.

The *infrastructure savings* method requires up to four types of zones:

- Open zoning
- Host zones
- Optional remote copy zones
- External storage zones

The *performance* method requires up to four types of zones; host zones, storage zones (internal and external), intra-cluster zones, with optional remote copy zones.

The *infrastructure savings* method has specific port cabling suggestions for the FlashSystem storage enclosures and AC3 control enclosures that are calculated to support connections and non disruptive scalability for up to eight AC3 controller enclosure and eight FlashSystem storage enclosures (four building blocks plus four storage enclosures). For details, see A.4, “Guidelines: The infrastructure savings method” on page 249.

The *performance* method requires planning to ensure non disruptive scalability. For details, see A.3, “Guidelines: The performance method” on page 247.

A.3 Guidelines: The performance method

The *performance* method for port utilization provides shared use of ports that use the full bidirectional capabilities of Fibre Channel.

Table A-1 on page 248 lists the connections to the two AC3 control enclosures, and host and storage connections to switches that are external to the building block. The building blocks connect to two customer fabrics external to the building block.

If you are doing remote copy (Metro Mirror or Global Mirror), then some ports designated for host and storage can be used for remote copy traffic. Another possibility is to get a fourth 16 Gb card and use some of those ports for remote copy. Depending on the remote copy bandwidth, you might use only one port per node.

Table A-1 provides an example of port utilization that designates ports for intra-cluster traffic, ports for host and storage traffic, and optional ports for remote copy.

Note: The performance method requires customer planning to ensure nondisruptive scalability.

Table A-1 Port connections in scalable environment with port utilization for performance

	12 x 16 Gb per AC3			
Source port on AC3 (slot:port)	**#**	**Type**	**SAN**	**Usage**
3:1	1	16 Gb	C (S1)	Intra-cluster
3:2	2	16 Gb	D (S2)	Intra-cluster
3:3	3	16 Gb	C	Host/Storage (or remote copy)
3:4	4	16 Gb	D	Host/Storage
4:1	5	16 Gb	C (S1)	Intra-cluster
4:2	6	16 Gb	D (S2)	Intra-cluster
4:3	7	16 Gb	C	Host/Storage
4:4	8	16 Gb	D	Host/Storage (or remote copy)
7:1	9	16 Gb	C	Host/Storage
7:2	10	16 Gb	D	Host/Storage
7:3	11	16 Gb	C	Host/Storage
7:4	12	16 Gb	D	Host/Storage
localfcportmask			000000110011	
partnerfcportmask			000010000100	
SAN - C / D customer fabric SAN - Optional S1 / S2 internal switches			32 ports per building block connect to customer fabric: 12 ports per AC3 and 8 ports per FlashSystem storage enclosures.	

Note: When comparing the port utilization for *infrastructure savings* in Table A-2 on page 250 with the port utilization for *performance* listing in Table A-1 on page 248, you see that the customer fabric port connection mapping and usage is much higher in the port utilization for performance method. This method provides shared-use ports that take advantage of the full bidirectional capabilities in Fibre Channel, resulting in higher performance.

Port mask

With Fibre Channel port masking, you can control whether the ports are used to communicate to other local nodes and if they are used to communicate to nodes in remote, partnered systems. Fibre Channel port masking does not affect host or storage traffic.

The system has two Fibre Channel port masks. The local port mask controls connectivity to other nodes in the same system, and the partner port mask control connectivity to nodes in partnered systems. By default, all ports are enabled for both local and partner connectivity. The port masks apply to all nodes in the same local system.

For example, if the local port mask is set to `101101` on a node with eight Fibre Channel ports, ports 1, 3, 4 and 6 are able to connect with other local nodes in the same system. Ports 2, 5, 7 and 8 are not used for connection to other local nodes.

Table A-1 on page 248 also lists the port masks (`localfcportmask` and `partnerfcportmask`) to use with these port designations**.**

Figure A-1 shows an example from the IBM FlashSystem V9000 GUI that displays Fibre Channel port masks (`localfcportmask` and `partnerfcportmask`). This information is displayed when you select **Settings** → **Network** → **Fibre Channel Ports** → **Actions** → **Modify Connection** → **View details**.

Deviations from the port assignments shown in Table A-1 on page 248 require corresponding changes to the port masks. Proper zoning is still required.

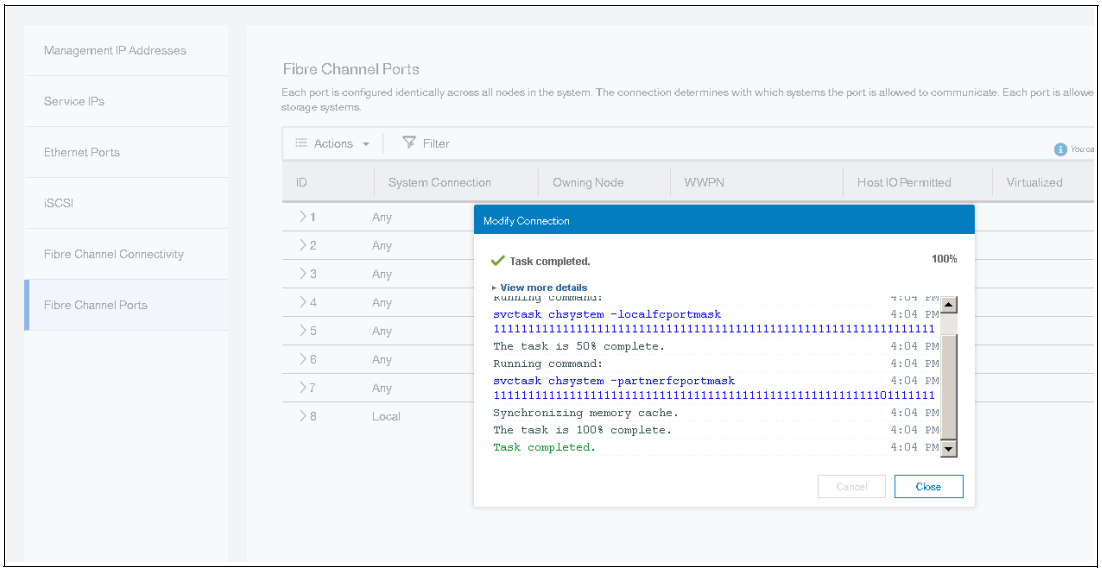


Figure A-1 Modify Fibre Channel port mask example

A.4 Guidelines: The infrastructure savings method

The *infrastructure savings* method for port utilization provides a dedicated port setup. The ports are used either for host, storage, or intra-cluster communication. This method has dedicated internal switches for the IBM FlashSystem storage enclosures connections and also intra-cluster communication with a reduced number of customer host facing ports.

Table A-2 shows the connections to the two AC3 control enclosures, including the cluster connections to the switch that is internal to the building block. Host, remote copy, and external storage connections to switches that are external to the building block are shown.

The building blocks connect to two SAN fabrics within the building block, and two customer fabrics external to the building block.

Table A-2 Port connections in scalable environment for minimum infrastructure effect

	12 x 16 Gb per AC3			
Source port on AC3 (slot:port)	**#**	**Type**	**SAN**	**Usage**
3:1	1	16 Gb	S1	Intra-cluster/FlashSystem storage
3:2	2	16 Gb	S2	Intra-cluster/FlashSystem storage
3:3	3	16 Gb	C	Remote copy (variable)
3:4	4	16 Gb	D	Host/External storage (Variable)
4:1	5	16 Gb	S1	Intra-cluster/FlashSystem storage
4:2	6	16 Gb	S2	Intra-cluster/FlashSystem storage
4:3	7	16 Gb	C	Host/External storage (variable)
4:4	8	16 Gb	D	Remote copy (variable)
7:1	9	16 Gb	C	Host
7:2	10	16 Gb	D	Host
7:3	11	16 Gb	C	Host
7:4	12	16 Gb	D	Host
localfcportmask			000000110011	
partnerfcportmask			000010000100	
SAN - C / D Customer Fabric SAN - S1 / S2 Internal Switches			16 ports per building block connect to customer fabric: 8 per AC3	

The infrastructure savings method has specific port cabling guidelines for the FlashSystem storage enclosures and AC3 controller enclosures. The guidelines are calculated to support connections and non disruptive scalability for up to eight AC3 controller enclosures and eight FlashSystem storage enclosures (four building blocks plus four FlashSystem storage enclosures).

A.5 Guidelines: Zoning and pathing

This section covers general guidelines concerning N_Port ID Virtualization (NPIV) and pathing and also zoning guidelines for the *performance* and *infrastructure savings* methods.

A.5.1 NPIV

N-Port Virtualization ID (NPIV) is a method for virtualizing a physical Fibre Channel port that is used for host I/O. When NPIV is enabled, you must use the virtualized WWPN in your host zones. The non-virtualized WWPN continue to be used for the storage, cluster, and remote copy zones.

From the CLI, the **`lsiogrp`** command returns a detailed view of I/O groups that are visible to the system. If the resulting output is `fctargetportnode:enabled` then NPIV is enabled.

Figure A-2 shows the Fibre Channel Ports view that can be displayed from the GUI by selecting **Settings** → **Network** → **Fibre Channel Ports** and then opening the selection for ports with `ID 1`. The figure shows the four worldwide port names (WWPNs) that correspond to ID 1 for node 1. The WWPN that must be used for the host zones is indicated by `Yes` in the *Host IO Permitted* and *Virtualized* columns.

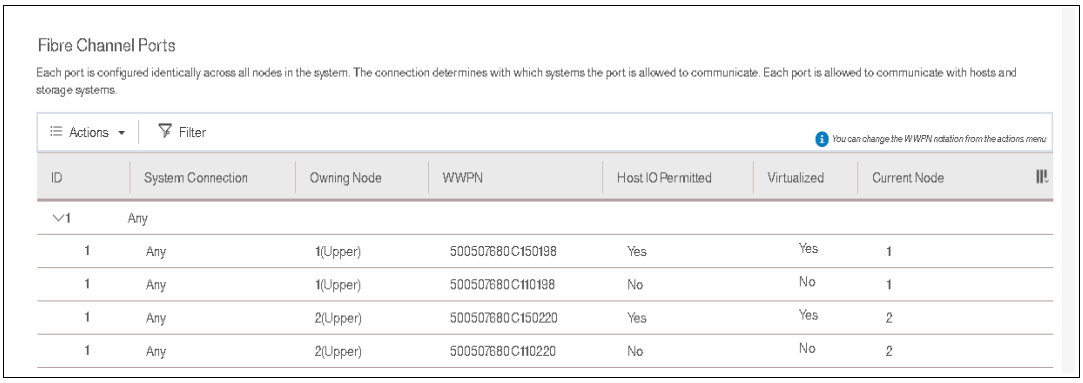


Figure A-2 Fibre Channel Ports view

Note: For more details about NPIV, see the *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 chapter about N-Port Virtualization ID (NPIV) support.

A.5.2 Pathing

The number of paths to a storage volume, commonly referred to as a *logical unit number (LUN)* is a composition of a number of physical and logical elements:

- Physical number of these:
  - IBM FlashSystem V9000 host accessible ports (target ports)
  - Ports per host (initiator ports)
- Logical number of these:
  - IBM FlashSystem V9000 host objects
  - Initiator ports (WWPNs) per IBM FlashSystem V9000 host object
  - IBM FlashSystem V9000 mappings per volume
  - Target ports per initiator ports (zoning)

Calculating the number of paths to a storage volume

The goal is to have no more than a maximum of eight redundant paths between the hosts and the volumes. This provides good performance and protection from single point of failures. Increasing the number of zones or increasing the number of host objects per host decreases the number of paths to a volume.

The first row in Table A-3 represents a way to keep the number of paths to four. This scales with hosts with more than two ports. In this case, there would be a host object for every pair of ports. These port pairs would each be connected to a separate fabric. An equal portion of volumes would be mapped to each host object.

Although eight redundant paths are supported, limiting the number of paths from each host to four reduces the performance impact of error handling by the host, and demands on resources within the host. Limiting the number of paths to four also helps avoid issues with high port fan-outs and fabric state changes.

The last row is not a recommended configuration and shows a high number of paths that can result from lumping all the host ports and all the target ports into one zone per fabric.

Table A-3 Calculating the number of paths to a volume

Host with 16 initiator ports AC3s with 16 target ports 2 redundant fabrics	Number of initiator ports (WWPNs) per host object	Number of host object mappings per volume	Number of target ports per initiator port	Number of paths
4 host zones, 8 host objects	02	1	2	04
2 host zones, 1 host object	16	1	8	128

The calculation is simple arithmetic as follows:

```
# of initiator ports per host object * number of host object mappings per volume *
number of target ports per initiator port = number of paths
```

Note: For a host with more than two ports, create multiple host definitions, each with no more than two ports.

For management simplification, multiple host objects may grouped into a single host cluster, as documented in *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933 chapter about Hosts.

A.5.3 Port utilization method for performance

The method for *performance* requires up to four types of zones: host zones, storage zones (internal and external), intra-cluster zones, with optional remote copy zones. They can be defined as follows:

- One intra-cluster zone per fabric
- Two host zones per fabric
- One storage zone per fabric
- One remote copy zone per fabric (if needed)

Figure A-3 shows an overview of the zoning that is configured for the performance method.

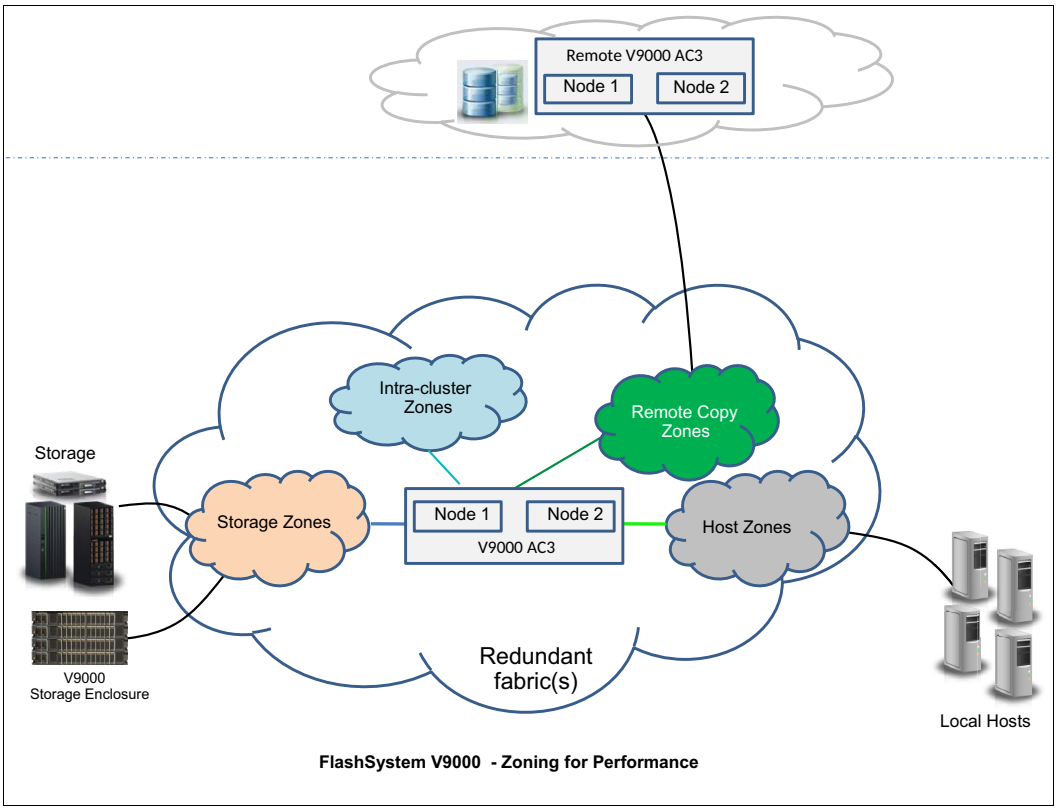


Figure A-3 High level zone requirements

Figure A-4 shows an example for zoning storage and host, based on Table A-1 on page 248 for 12 x 16 Gbps building blocks.

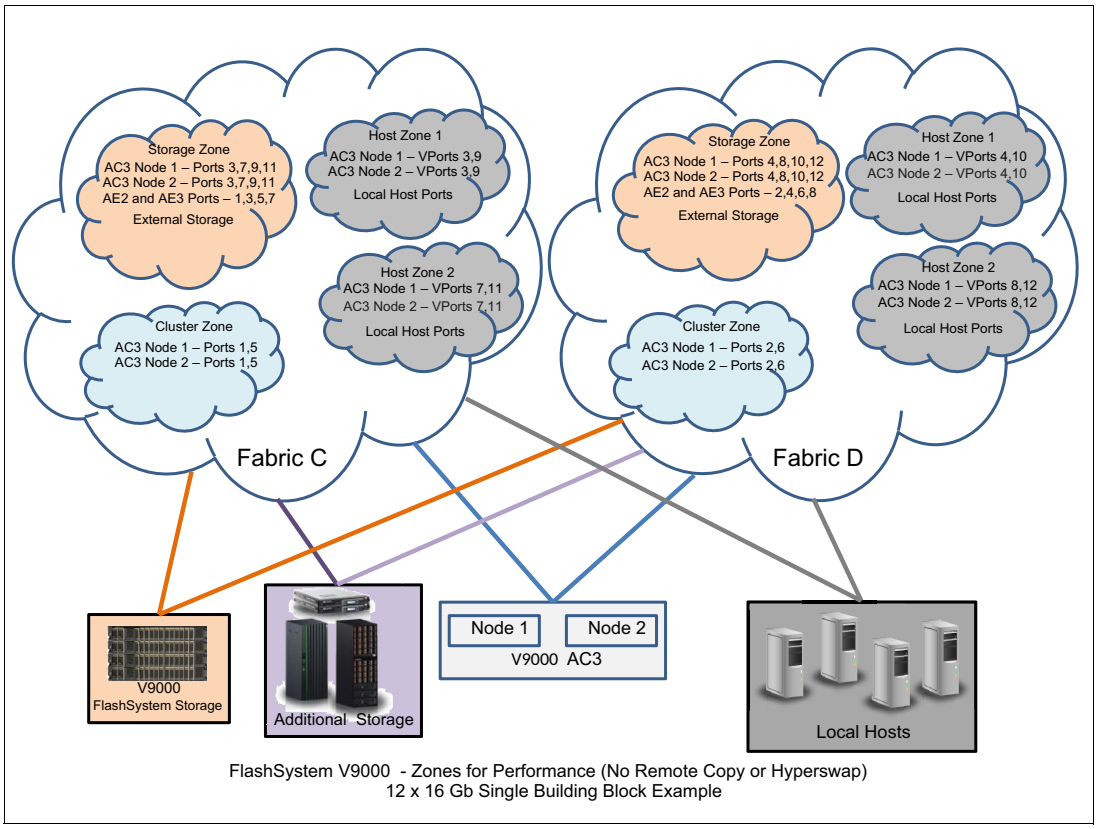


Figure A-4 Zoning for performance detailed example

The same physical AC3 ports are used for both storage and host connections but the WWPNs are different due to NPIV.

Tip: You must use the virtualized WWPN for the host zones with NPIV enabled.

A.5.4 Port utilization method for infrastructure savings

The infrastructure savings method requires up to four types of zones:

- Open zoning and FlashSystem storage
- Host zones
- Remote copy zones (optional)
- Additional external storage zones

These zones are described as follows:

- Open zoning on internal switches

  AC3 and Flash storage enclosures components are connected as documented in the “Installing” topic of IBM Knowledge Center.

- Two host zones per fabric
- One external storage zone per fabric
- One remote copy zone per fabric (if needed)

Figure A-5 shows an overview of the zoning configured for the infrastructure savings method.

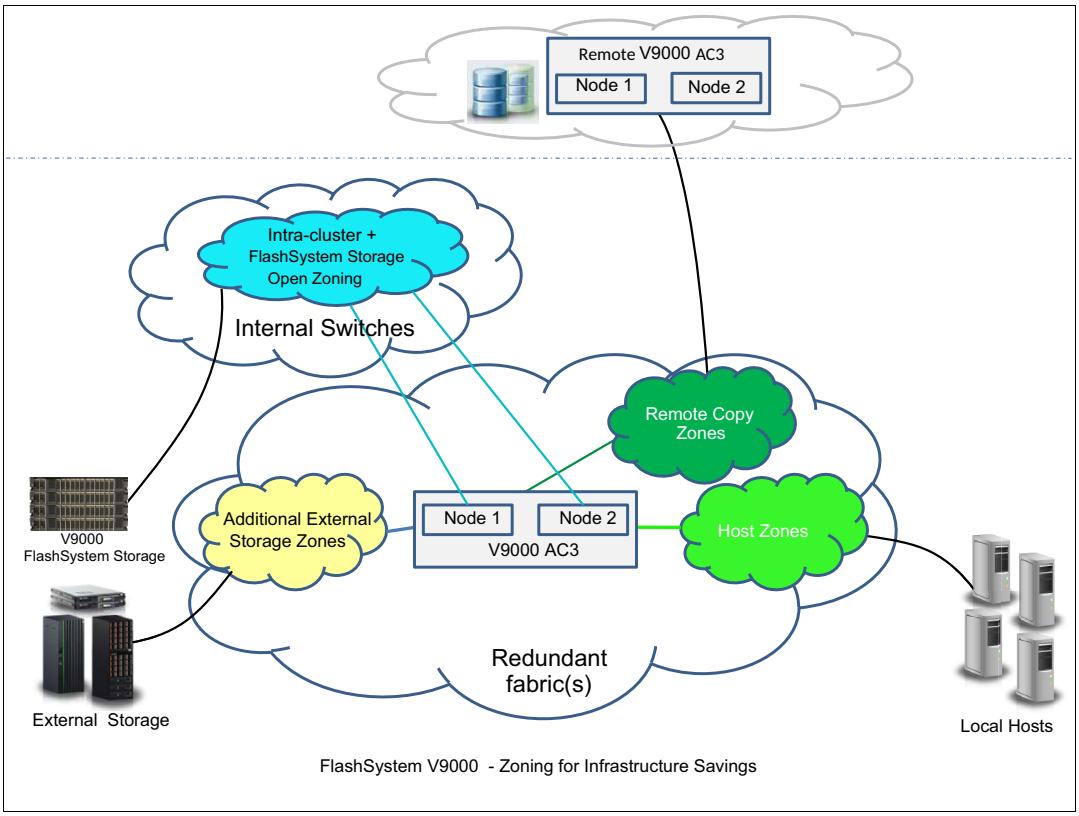


Figure A-5 IBM FlashSystem V9000 scalable solution: Zoning for infrastructure savings

A.6 Summary

This appendix provides guidelines for two methods or port utilization in a scalable IBM FlashSystem V9000 environment. Depending on customer requirements, the following methods of port utilization can be put into practice:

- IBM FlashSystem V9000 port utilization for *performance*

  This method uses more customer switch ports to improve performance for certain configurations. Only ports designated for intra-cluster communication are attached to private internal switches. The private internal switches are optional and all ports can be attached to customer switches.

- IBM FlashSystem V9000 port utilization for *infrastructure savings*

  This method reduces the number of required customer Fibre Channel ports that are attached to the customer fabrics. This method provides high performance and low latency but performance might be port-limited for certain configurations. Intra-cluster communication and FlashSystem storage traffic occur over the internal switches.

By following the port utilization cabling and zoning guidelines in this appendix, you can maximize the full potential of performance and low latency of IBM FlashSystem V9000 in enterprise-scalable environments.

Note: It is beyond the intended scope of this IBM Redpaper publication to provide an in-depth understanding of port configuration and SAN zoning setup. For more information, see the *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933, chapter about SAN configuration planning.

A.7 Supported environments

IBM FlashSystem V9000 can be used in different port utilizations. To check for supported switches, see the IBM System Storage Interoperation Center (SSIC).

For your search, use the values listed in Table A-4.

Table A-4 Values for the search fields

Field	Value to select
Storage Family	IBM System Storage Enterprise Flash
Storage Version	FlashSystem V9000 8.1.x
Connection Protocol	Fibre Channel

Validate that the intended FC switches are listed in the SAN or Networking field.

If there are additional host-side details, such as platform, OS, or multipathing, that restricts your switch selection, then choose the selections in the appropriate lists.

Related publications

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

IBM Redbooks

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:

- *Accelerate with IBM FlashSystem V840 Compression*, REDP-5147
- *Deploying IBM FlashSystem V840 Storage in a VMware and Cloud Environment*, REDP-5148
- *IBM FlashSystem 900 Model AE3 Product Guide*, REDP-5467
- *Implementing IBM FlashSystem 900 Model AE3*, SG24-8414
- *FlashSystem V9000 Product Guide*, REDP-5468
- *IBM FlashSystem V9000 and VMware Best Practices Guide*, REDP-5247
- *IBM FlashSystem V9000 in a VersaStack Environment*, REDP-5264
- *IBM FlashSystem V9000 Version 7.7 Product Guide*, REDP-5409
- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933
- *Introducing and Implementing IBM FlashSystem V9000*, SG24-8273

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

ibm.com/redbooks

Other publications

The *IBM FlashSystem V9000 Quick Start Guide,* GI13-2894 publication is also relevant as a further information source.

Online resources

These websites are also relevant as further information sources:

- IBM FlashSystem V9000 in IBM Knowledge Center:

  https://ibm.biz/fs_V900O_kc

- IBM FlashSystem family:

  https://ibm.biz/BdsaFF

  https://ibm.biz/BdsaFH
- IBM FlashSystem Ecosystem solutions:

  http://www.ibm.com/systems/storage/flash/ecosystem/isv.html
- IBM System Storage Interoperation Center (SSIC):

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

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

Implementing IBM FlashSystem V9000 AE3

SG24-8413-00

ISBN 0738442992

IBM

(0.5” spine)
0.475”<->0.873”
250 <-> 459 pages

SG24-8413-00

ISBN 0738442992

Printed in U.S.A.