

# Implementing IBM FlashSystem 900

Karen Orlando

Ingo Dimmer

Detlef Helmbrecht

Jon Herd

Carsten Larsen

Matt Levan



**Storage**





International Technical Support Organization

**Implementing IBM FlashSystem 900**

November 2016

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

**Second Edition (November 2016)**

This edition applies to the IBM FlashSystem 900, Release 1.2.

© Copyright International Business Machines Corporation 2015, 2019. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.



# Contents

<b>Notices</b> .....	ix
Trademarks .....	x
<b>IBM Redbooks promotions</b> .....	xi
<b>Preface</b> .....	xiii
Authors .....	xiv
Now you can become a published author, too! .....	xvi
Comments welcome .....	xvi
Stay connected to IBM Redbooks .....	xvii
<b>Summary of changes</b> .....	xix
April 2019, Second Edition update .....	xix
November 2016, Second Edition .....	xix
May 2015, First Edition .....	xix
<b>Chapter 1. FlashSystem storage introduction</b> .....	1
1.1 FlashSystem storage overview .....	2
1.2 IBM FlashCore technology .....	2
1.2.1 Flash wear guarantee .....	4
1.3 Why flash technology matters .....	4
1.4 IBM FlashSystem family: Product differentiation .....	5
1.5 Technology and architectural design overview .....	6
1.5.1 Hardware-only data path .....	7
1.5.2 The 20nm flash card memory chips .....	8
1.5.3 Flash module capacities .....	8
1.5.4 Gateway interface FPGA .....	8
1.5.5 Flash controller FPGA .....	8
1.5.6 IBM Variable Stripe RAID and 2D Flash RAID overview .....	9
1.6 Variable Stripe RAID .....	10
1.6.1 How Variable Stripe RAID works .....	10
1.7 Two-dimensional (2D) Flash RAID .....	12
<b>Chapter 2. IBM FlashSystem 900 architecture</b> .....	15
2.1 Introduction to the IBM FlashSystem 900 architecture .....	16
2.1.1 Capacity .....	16
2.1.2 Performance and latency .....	17
2.1.3 Power requirements .....	18
2.1.4 Physical specifications .....	18
2.1.5 FlashCore technology .....	18
2.1.6 Scalability .....	20
2.1.7 Protocol support .....	20
2.1.8 Encryption support .....	21
2.1.9 Comparison of IBM FlashSystem models 840 and 900 .....	23
2.1.10 Management .....	23
2.2 Architecture of IBM FlashSystem 900 .....	24
2.2.1 Overview of architecture .....	24
2.2.2 Hardware components .....	26
2.2.3 Canisters .....	26

2.2.4	Interface cards	27
2.2.5	MicroLatency modules	31
2.2.6	Battery modules	34
2.3	Administration and maintenance	36
2.3.1	Serviceability and software enhancements	36
2.3.2	System management	37
2.4	Support matrix	43
2.5	Product integration overview	43
2.5.1	IBM Spectrum Virtualize - SAN Volume Controller	44
2.5.2	IBM Storwize V7000 storage array	45
2.5.3	IBM DB2 database environments	46
2.5.4	IBM Spectrum Scale	46
2.5.5	IBM TS7650G ProtecTIER	47
2.5.6	VMware with IBM Storage Integration Server	48
<b>Chapter 3.</b>	<b>Planning</b>	<b>49</b>
3.1	Prerequisites to installation	50
3.1.1	Contact information and checklist	50
3.1.2	Completing the hardware location chart	51
3.2	Planning cable connections	52
3.2.1	Management port connections	52
3.2.2	Interface card connections	54
3.3	Planning for power	59
3.4	Planning for configuration	59
3.5	The call home option	61
3.6	TCP/IP requirements	62
3.7	Planning for encryption	63
3.8	Checking web browser settings for the management GUI	64
3.9	Licensing	66
3.10	Supported hosts and operating system considerations	66
<b>Chapter 4.</b>	<b>Installation and configuration</b>	<b>67</b>
4.1	First-time installation	68
4.1.1	Installing the hardware	68
4.2	Cabling the system	71
4.2.1	Cabling for Fibre Channel	71
4.2.2	Cabling for FCoE	73
4.2.3	Cabling for iSCSI	73
4.2.4	Cabling for QDR InfiniBand	74
4.2.5	FC cable type	74
4.2.6	Ethernet management cabling	74
4.2.7	Power requirements	75
4.2.8	Cooling requirements	75
4.2.9	Cable connector locations	75
4.3	Initializing the system	76
4.3.1	Prepare for initialization using InitTool	76
4.3.2	Initializing the system through the web management interface	86
4.3.3	Service Assistant Tool	97
4.4	RAID storage modes	98
4.5	Connectivity guidelines for improved performance	98
4.5.1	Interface card configuration guidelines	98
4.5.2	Host adapter guidelines	99
4.5.3	Cabling guidelines	100

4.5.4 Zoning guidelines . . . . .	100
<b>Chapter 5. IBM FlashSystem 900 client host attachment and implementation . . . .</b>	<b>101</b>
5.1 Host implementation and procedures . . . . .	102
5.2 Host connectivity . . . . .	102
5.2.1 Fibre Channel SAN attachment . . . . .	102
5.2.2 Fibre Channel direct attachment . . . . .	103
5.2.3 General Fibre Channel attachment rules . . . . .	104
5.3 Operating system connectivity and preferred practices . . . . .	104
5.3.1 FlashSystem 900 sector size . . . . .	104
5.3.2 File alignment for the best RAID performance . . . . .	105
5.3.3 IBM AIX and FlashSystem 900 . . . . .	105
5.3.4 IBM i and FlashSystem 900 . . . . .	109
5.3.5 FlashSystem 900 and Linux client hosts . . . . .	115
5.3.6 FlashSystem 900 and Microsoft Windows client hosts . . . . .	117
5.3.7 FlashSystem 900 and client VMware ESX hosts . . . . .	122
5.3.8 FlashSystem 900 and IBM SAN Volume Controller or Storwize V7000 . . . . .	123
5.3.9 FlashSystem iSCSI host attachment . . . . .	123
5.3.10 FlashSystem iSCSI configuration . . . . .	123
5.3.11 Windows 2008 R2 and Windows 2012 iSCSI attachment . . . . .	125
5.3.12 Linux iSCSI attachment . . . . .	128
5.4 Miscellaneous host attachment . . . . .	131
5.4.1 FlashSystem 900 and Solaris client hosts . . . . .	131
5.4.2 FlashSystem 900 and HP-UX client hosts . . . . .	139
5.5 FlashSystem 900 preferred read and configuration examples . . . . .	140
5.5.1 FlashSystem 900 deployment scenario with preferred read . . . . .	140
5.5.2 Implementing preferred read . . . . .	142
5.5.3 Linux configuration file multipath.conf example . . . . .	152
5.5.4 Example of a VMware configuration . . . . .	157
5.6 FlashSystem 900 and Easy Tier . . . . .	157
5.7 Troubleshooting . . . . .	158
5.7.1 Troubleshooting Linux InfiniBand configuration issues . . . . .	158
5.7.2 Linux fdisk error message . . . . .	159
5.7.3 Changing FC port properties . . . . .	160
5.7.4 Changing iSCSI port properties . . . . .	160
<b>Chapter 6. Using IBM FlashSystem 900 . . . . .</b>	<b>161</b>
6.1 Overview of IBM FlashSystem 900 management tools . . . . .	162
6.1.1 Access to the graphical user interface . . . . .	162
6.1.2 Graphical user interface (GUI) layout . . . . .	163
6.1.3 Navigation . . . . .	165
6.1.4 Selecting multiple items . . . . .	166
6.1.5 Status indicators . . . . .	167
6.2 Monitoring menu . . . . .	168
6.2.1 Monitoring System menu . . . . .	168
6.2.2 Monitoring events . . . . .	179
6.2.3 Monitoring performance menu . . . . .	190
6.3 Volumes menu . . . . .	199
6.3.1 Navigating to the Volumes menu . . . . .	200
6.3.2 Volumes menu . . . . .	200
6.3.3 Volumes by Host . . . . .	206
6.4 Hosts menu . . . . .	210
6.4.1 Navigating to the Hosts menu . . . . .	211

6.4.2 Volumes by Host . . . . .	218
6.5 Access menu . . . . .	218
6.5.1 Navigating to the Access menu . . . . .	219
6.5.2 Users window . . . . .	219
6.5.3 Accessing CLI by using PuTTY . . . . .	223
6.5.4 User groups . . . . .	226
6.5.5 Audit log menu . . . . .	229
<b>Chapter 7. Configuring settings . . . . .</b>	<b>231</b>
7.1 Settings menu . . . . .	232
7.1.1 Navigating to the Settings menu . . . . .	232
7.1.2 Notifications menu . . . . .	233
7.1.3 Security menu . . . . .	243
7.1.4 Network menu . . . . .	260
7.1.5 Support menu . . . . .	261
7.1.6 System menu . . . . .	265
7.2 Service Assistant Tool . . . . .	276
7.2.1 Accessing Service Assistant Tool . . . . .	276
7.2.2 Log in to Service Assistant Tool . . . . .	277
<b>Chapter 8. Product integration . . . . .</b>	<b>279</b>
8.1 Running FlashSystem 900 with IBM Spectrum Virtualize - SAN Volume Controller . . . . .	280
8.1.1 IBM System Storage SAN Volume Controller introduction . . . . .	280
8.1.2 SAN Volume Controller architecture and components . . . . .	283
8.1.3 SAN Volume Controller hardware options . . . . .	285
8.1.4 IBM Spectrum Virtualize - SAN Volume Controller advanced functionality . . . . .	288
8.2 SAN Volume Controller connectivity to FlashSystem 900 . . . . .	290
8.2.1 SAN Volume Controller FC cabling to SAN . . . . .	291
8.2.2 SAN zoning and port designations . . . . .	292
8.2.3 Port designations . . . . .	293
8.2.4 Verifying FlashSystem 900 connectivity in SAN Volume Controller . . . . .	295
8.2.5 Import and export . . . . .	307
8.3 Integration considerations: FlashSystem 900 and SAN Volume Controller . . . . .	308
8.4 Integration considerations: FlashSystem 900 and IBM Storwize V7000 . . . . .	308
<b>Chapter 9. Use cases and solutions . . . . .</b>	<b>311</b>
9.1 Introduction to the use cases . . . . .	312
9.2 Tiering . . . . .	313
9.2.1 Easy Tier or block-level tiering . . . . .	313
9.2.2 Information Life Management or file-level tiering . . . . .	315
9.3 Preferred read . . . . .	315
9.3.1 Implementing preferred read . . . . .	318
9.4 Flash only . . . . .	321
9.5 Comparison . . . . .	321
<b>Chapter 10. Hints and tips . . . . .</b>	<b>323</b>
10.1 Encryption hints . . . . .	324
10.2 Setting up call home . . . . .	324
10.3 System check . . . . .	325
10.3.1 Checking the Fibre Channel connections . . . . .	325
10.4 Host attachment hints . . . . .	326
10.4.1 Fibre Channel link speed . . . . .	326
10.4.2 Host is in a degraded state . . . . .	327
10.4.3 FlashSystem port status . . . . .	327

10.4.4 AIX multipathing .....	328
10.4.5 Direct attach hints .....	328
10.5 General guidelines for testing a specific configuration .....	328
10.5.1 Save the default configuration .....	329
10.5.2 Test scenarios .....	330
10.5.3 Supported System Configurations .....	330
10.5.4 Secure erase of data .....	330
10.5.5 Performance data gathering basics .....	331
10.6 Troubleshooting .....	334
10.6.1 Troubleshooting prerequisites and information to record .....	334
10.6.2 User interfaces for servicing your system .....	337
10.6.3 Event reporting .....	340
10.6.4 Resolving a problem .....	343
10.7 IBM System Storage Interoperation Center (SSIC) .....	343
<b>Related publications</b> .....	345
IBM Redbooks .....	345
Online resources .....	346
Help from IBM .....	346



# Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US*

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

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

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

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

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

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

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

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


## COPYRIGHT LICENSE:

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

# Trademarks

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

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

AIX®	IBM Spectrum™	PowerVM®
BigInsights®	IBM Spectrum Accelerate™	ProtecTIER®
DB2®	IBM Spectrum Control™	Real-time Compression™
developerWorks®	IBM Spectrum Protect™	Redbooks®
DS8000®	IBM Spectrum Scale™	Redbooks (logo)  ®
Easy Tier®	IBM Spectrum Storage™	Storwize®
FlashCopy®	IBM Spectrum Virtualize™	System Storage®
GPFS™	MicroLatency®	Tivoli®
IBM®	NetView®	Tivoli Enterprise Console®
IBM FlashCore®	Power Systems™	Variable Stripe RAID™
IBM FlashSystem®	POWER7®	XIV®
IBM Flex System®	POWER8®	
IBM SmartCloud®	PowerPC®	

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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.

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.



## Find and read thousands of IBM Redbooks publications

- ▶ Search, bookmark, save and organize favorites
- ▶ Get personalized notifications of new content
- ▶ Link to the latest Redbooks blogs and videos

Get the latest version of the Redbooks Mobile App



iOS

Download  
Now

Android



## Promote your business in an IBM Redbooks publication

Place a Sponsorship Promotion in an IBM® Redbooks® publication, featuring your business or solution with a link to your web site.

Qualified IBM Business Partners may place a full page promotion in the most popular Redbooks publications. Imagine the power of being seen by users who download millions of Redbooks publications each year!



[ibm.com/Redbooks](http://ibm.com/Redbooks)

About Redbooks → Business Partner Programs

THIS PAGE INTENTIONALLY LEFT BLANK

# Preface

Today's global organizations depend on being able to unlock business insights from massive volumes of data. Now, with IBM® FlashSystem 900, powered by IBM FlashCore® technology, they can make faster decisions based on real-time insights and unleash the power of the most demanding applications, including online transaction processing (OLTP) and analytics databases, virtual desktop infrastructures (VDIs), technical computing applications, and cloud environments.

This IBM Redbooks® publication introduces clients to the IBM FlashSystem® 900. It provides in-depth knowledge of the product architecture, software and hardware, implementation, and hints and tips. Also illustrated are use cases that show real-world solutions for tiering, flash-only, and preferred-read, and also examples of the benefits gained by integrating the FlashSystem storage into business environments.

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

This book describes the following offerings of the IBM Spectrum™ Storage family:

- ▶ IBM Spectrum Storage™

The IBM Spectrum Storage family is based on proven technologies and designed specifically to simplify storage management, scale to keep up with data growth, and optimize data economics. It represents a new, more agile way of storing data, and helps organizations prepare themselves for new storage demands and workloads. The software-defined storage solutions included in the IBM Spectrum Storage family can help organizations simplify their storage infrastructures, cut costs, and start gaining more business value from their data.

For details about the entire IBM Spectrum Storage family, see the following web page:

<http://www.ibm.com/systems/storage/spectrum/>

- ▶ IBM Spectrum Control™

IBM Spectrum Control 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 SmartCloud® Virtual Storage Center, Tivoli® Storage Productivity Center, IBM Storage Integration Server, and others.

For more information about IBM Spectrum Control, see the following web page:

[https://www.ibm.com/developerworks/servicemanagement/sm/spectrum\\_control/](https://www.ibm.com/developerworks/servicemanagement/sm/spectrum_control/)

- ▶ IBM Spectrum Virtualize™

IBM Spectrum Virtualize industry-leading storage virtualization enhances existing storage to improve resource utilization and productivity to achieve a simpler, more scalable and cost-efficient IT infrastructure. The functionality of IBM Spectrum Virtualize is provided by IBM SAN Volume Controller.

See the IBM Spectrum Virtualize - SAN Volume Controller web page:

<http://www.ibm.com/systems/storage/software/virtualization/svc/>

► IBM Spectrum Scale™

IBM Spectrum Scale is a proven, scalable, high-performance data and file management solution, based on IBM General Parallel File System or GPFS™. IBM Spectrum Scale technology is a high-performance enterprise file management platform. It can help you move beyond simply adding storage to optimize data management.

See the IBM Spectrum Scale web page:

<http://www.ibm.com/systems/storage/spectrum/scale/>

► IBM Spectrum Accelerate™

IBM Spectrum Accelerate is a software-defined storage solution, born of the proven IBM XIV® integrated storage offering. It is designed to help speed delivery of data across the organization and add extreme flexibility to cloud deployments

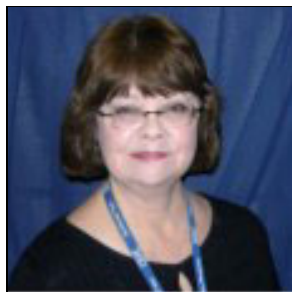
IBM Spectrum Accelerate delivers hotspot-free performance, easy management scaling, and proven enterprise functionality such as advanced mirroring and flash caching to different deployment platforms.

See the IBM Spectrum Accelerate web page:

<http://www.ibm.com/systems/storage/spectrum/accelerate/>

## Authors

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



**Karen Orlando** is a Project Leader at the International Technical Support Organization, Tucson Arizona Center. Karen has over 25 years in the IT industry with extensive experience in open systems management, and information and software development of IBM hardware and software storage. She holds a degree in Business Information Systems from the University of Phoenix and is Project Management Professional (PMP), certified since 2005.



**Ingo Dimmer** is an IBM Senior Certified IT Specialist and a Project Management Institute (PMI) Certified Project Manager Professional in the EMEA IBM System Storage® Advanced Technical Support organization based in Kelsterbach, Germany. Ingo has 17 years of experience in enterprise storage. He provides pre-sales and post-sales technical support focusing on the IBM Power Systems™ platform. His areas of expertise include IBM i external disk and tape storage solutions, I/O performance, virtualization and high availability. Ingo is the author of several IBM technical papers and IBM Redbooks publications.



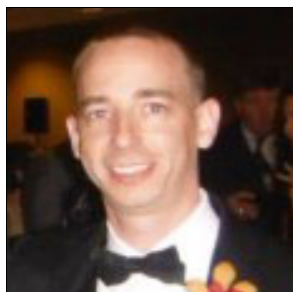
**Detlef Helmbrecht** is an Advanced Technical Skills (ATS) IT Specialist working for the IBM Systems & Technology Group. He is located in the European Storage Competence Center (ESCC), Germany. Detlef has over 25 years of experience in IT, performing various roles, including software design, sales, and solution architect. His areas of expertise include high-performance computing (HPC), disaster recovery, archiving, application tuning, and FlashSystem.



**Jon Herd** is an IBM Storage Technical Advisor working for the European Storage Competence Center (ESCC), Germany. He covers UK and Ireland, advising customers on a portfolio of IBM storage products including FlashSystem products. Jon has been with IBM for more than 40 years and has held various technical roles, including EMEA level support on mainframe servers and technical education development. He holds IBM certifications in Supporting IT Solutions at expert level, and Actualizing IT Solutions experienced level. He is also a certified member of the British Computer Society (MBCS CITP) and the Institute of Engineering and Technology (MIET).



**Carsten Larsen** is an IBM Certified Senior IT Specialist working for the Technical Services Support organization in IBM Denmark, delivering consultancy services to IBM clients within the storage arena. Carsten joined IBM in 2007, leaving behind a job at HP where he worked with storage arrays and UNIX for 10 years, holding, among others, the HP certification: Master Certified Systems Engineer (MASE). While working for IBM, Carsten obtained Brocade BCFP and BCSD certifications and NetApp NCD and NCIE certifications. Carsten is the author of several Redbooks publications related to these product certifications.



**Matt Levan** is an IBM FlashSystem Corporate Solutions Architect. Matt has been involved in IT and Storage Technologies for over 15 years. He has performed innumerable roles, including storage administration, technical support, and solutions architect. Matt spent many years at technology-leading companies, such as Novus Consulting Group, VeriSign, EMC, and Innovative Data Solutions. Matt's current role is to help worldwide field technical sellers for IBM Flash Storage products. His primary responsibilities are to provide technical sales engineering, competitive winning strategies, and cohesive technical sales solutions to a multitude of IBM internal organizations and IBM clients worldwide.

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

- ▶ From IBM i Storage Development:  
Kristopher Whitney  
Joe Writz
- ▶ From IBM Systems, Flash System:  
James Cioffi  
Mark Fleming  
Brian Groff  
Gen Jaquette  
Matt Key  
Kim Miller  
Greg Windsor  
Brent Yardley
- ▶ From IBM Systems, User Experience Design:  
Elizabeth Hatfield

## Now you can become a published author, too!

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

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

<http://www.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 found at this website:  
<http://www.ibm.com/redbooks>
- ▶ Send your comments in an email to:  
[redbooks@us.ibm.com](mailto:redbooks@us.ibm.com)
- ▶ Mail your comments to:  
IBM Corporation, International Technical Support Organization  
Dept. HYTD Mail Station P099  
2455 South Road  
Poughkeepsie, NY 12601-5400

## Stay connected to IBM Redbooks

- ▶ Find us on Facebook:  
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:  
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:  
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:  
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:  
<http://www.redbooks.ibm.com/rss.html>





# Summary of changes

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

Summary of Changes for *Implementing IBM FlashSystem 900*  
for SG24-8271-01, as created or updated on April 12, 2019.

## April 2019, Second Edition update

Added note on page 122 stating the FlashSystem 900 does not support the Atomic Test and Set (ATS) API of VMware.

## November 2016, Second Edition

The second edition includes information about attaching the IBM FlashSystem 900 to IBM i. Chapter 5, “IBM FlashSystem 900 client host attachment and implementation” on page 101 was updated with considerations that are specific to IBM i. The following topics that are relative to IBM i and IBM FlashSystem 900 are new to this chapter:

- ▶ Attachment methods
- ▶ Zoning rules
- ▶ Hardware and software requirements
- ▶ Configuration and performance considerations

## May 2015, First Edition

This first edition reflects the information that is implemented for the initial release of IBM FlashSystem 900, which is based on FlashSystem 840 release 1.3.

The changes described are implemented since IBM FlashSystem 840 release 1.3, which is document in *Implementing IBM FlashSystem 840*, SG24-8189.

This edition might include minor corrections and editorial changes that are not identified.

### New information

This edition includes the following new information.

#### ***IBM FlashCore technology***

At the heart of FlashSystem 900 lies IBM FlashCore technology, which consists of the following three key elements:

- ▶ Hardware-accelerated architecture engineered for flash, with hardware only data path including Hardware RAID, Non-blocking Crossbar Switch, Hardware Only Data Path, Single Box Highly Available Architecture, Concurrent Code Load and Concurrent Maintenance

- ▶ IBM MicroLatency® modules designed for low latency, high density, and reliability with health binning and heat segregation which continually assesses health by monitoring several key parameters and proactively performing in-flight adjustments.
- ▶ IBM Advanced Flash Management improves flash endurance 9x over standard implementations with IBM engineered ECC, advanced wear leveling, and proprietary garbage collection, relocation, and block picking algorithms. See “Terminology” on page 19.

### ***IBM MicroLatency module granularity***

IBM has enhanced the flexibility of the IBM FlashSystem 900 with new, more granular capacity offerings starting as low as a 2.4 TB RAID 5 configuration, providing an attractive entry point so that you can start with smaller, application-specific flash deployments and scale up to Tier 1 disk replacement as business needs grow.

This book includes the following new information based on Release 2.0:

- ▶ IBM MicroLatency modules incorporating IBM enhanced multi-level cell (MLC) chips and IBM FlashCore technology
- ▶ Up to 12 IBM MicroLatency modules in increments of 4, 6, 8, 10, or 12 for the 1.2 TB module and 6, 8, 10, or 12 for the 2.9 TB and 5.7 TB modules
- ▶ Capacity points:
  - RAID 5: 2.2, 4.5, 6.8, 9.1, 11.4, 11.4, 17.1, 22.8, 28.5, 22.8, 34.2, 45.6, and 57 TB

**Note:** IBM FlashSystem 900 only supports a RAID 5 configuration.

### ***FlashSystem 900 GUI management***

Additional FlashSystem 900 management software improvements make system management and performance health monitoring even more effective in the already-intuitive FlashSystem GUI. Additional features include these:

- ▶ Selectable front and rear views of the machine through a rotate icon, with hot spot highlight displays for each of the major components status.
- ▶ A tri-pane indicator at the bottom center on the main GUI, displaying the values of bandwidth, IOPS, and latency.
- ▶ New function (Modify Reserve Capacity) for changing the amount of space reserved for LUNs. Reserving less space for LUNs allows the FlashSystem 900 to allocate more space for housekeeping operations, and as a result improves the overall machine performance.

### **Retained Features**

The following features are included with FlashSystem 900 as of support of FlashSystem 840 release 1.3.

#### ***1300 watt power supply***

The 1300 watt power supply for high-line voltage is now standard in IBM FlashSystem 900. This provides a high-power output supply, to run at maximum performance for longer durations during power supply servicing, resulting in more predictable performance under unexpected failure conditions.

#### ***Encryption***

FlashSystem 900 has retained these new encryption functions, which are performed from the FlashSystem GUI:

- ▶ Hot Encryption Activation: Adding an encryption license to a previously initialized system

- Encryption Rekey: Changing the encryption key on a previously initialized system

### ***Battery reconditioning***

IBM FlashSystem 900 incorporates the battery reconditioning feature to calibrate the gauge that reports the amount of charge on the batteries. This ensures that the batteries are kept at the optimal level to support the IBM FlashSystem 900, and conduct a controlled and orderly shutdown, in the event of an external power outage.

### **Changed information**

This revision also includes additional feedback from the field including Removal of RAID 0 support. RAID 0 configuration is no longer supported on the IBM FlashSystem 900. Only RAID 5 configurations are supported.





# FlashSystem storage introduction

Flash technology in the data center is too relevant to be ignored for a few simple reasons:

- ▶ Since its introduction, flash storage has improved across all metrics: higher performance, density, and reliability, all of which translate to improved business efficiency.
- ▶ Flash cost per capacity and cost per transaction relative to hard disk storage make it extremely attractive to businesses that are attempting to maintain pace in a 24x7 competitive marketplace.
- ▶ Flash is easily integrated into existing data center environments and provides an instant boost to the mission-critical applications.

Although flash in storage is pervasive in the data center, its implementation varies considerably among competitors and technologies. Some use it as a simple cache accelerator; others implement it as yet another permanent data tier. The reality is that flash only matters when two conditions in the data center are met:

- ▶ Flash eliminates I/O bottlenecks while generating higher levels of application efficiency (improved performance).
- ▶ Storage economics are improved by its use. That is, it provides lower total cost of ownership (TCO) and faster return on investment (ROI) to the existing environment (enables new business opportunities).

The IBM FlashSystem storage delivers high performance, efficiency, and reliability for shared enterprise storage environments. It helps clients address performance issues with their most important applications and infrastructure.

This chapter introduces the IBM FlashSystem storage system and its core value, benefits, and technological advantages. It covers the following topics:

- ▶ FlashSystem storage overview
- ▶ IBM FlashCore technology
- ▶ Why flash technology matters
- ▶ IBM FlashSystem family: Product differentiation
- ▶ Technology and architectural design overview
- ▶ Variable Stripe RAID
- ▶ Two-dimensional (2D) Flash RAID

## 1.1 FlashSystem storage overview

Flash technology fundamentally changed the paradigm for IT systems, enabling new use cases and unlocking the scale of enterprise applications. Flash technology enhances the performance, efficiency, reliability, and design of essential enterprise applications and solutions by addressing the bottleneck in the IT process (data storage), enabling truly optimized information infrastructure.

The IBM FlashSystem shared flash storage systems offer affordable, high-density, ultra low-latency, highly reliable and scalable performance in a storage device that is both space efficient and power efficient. IBM Flash products, which can either augment or replace traditional hard disk drive (HDD) storage systems in enterprise environments, empower applications to work faster and scale further.

In addition to optimizing performance, the IBM FlashSystem family helps bring enterprise reliability and macro efficiency to the most demanding data centers so that businesses can see the following benefits:

- ▶ Reduce customer complaints by improving application response time.
- ▶ Service more users with less hardware.
- ▶ Reduce I/O wait and response times of critical applications.
- ▶ Simplify solutions.
- ▶ Reduce power and floor space requirements.
- ▶ Speed up applications, therefore enhancing the pace of business.
- ▶ Improve the utilization of the existing infrastructure.
- ▶ Extend the existing infrastructure.
- ▶ Mitigate risk.

From the client business perspective, an IBM FlashSystem provides benefits and value in four essential areas:

- ▶ **Extreme performance**  
Enables businesses to unleash the power of performance, scale, and insight to drive services and products to market faster
- ▶ **MicroLatency**  
Achieves competitive advantage through applications that enable faster decision making because of microsecond response times
- ▶ **Macro efficiency**  
Decreases costs by getting more from the efficient use of the IT staff, IT applications, and IT equipment because of the efficiencies flash brings to the data center
- ▶ **Enterprise reliability**  
Enhances customer experience through durable and reliable designs that use enterprise class flash and patented data protection technology

## 1.2 IBM FlashCore technology

The IBM FlashCore technology, used in IBM FlashSystem 900, employs several new and patented mechanisms to achieve greater capacity and throughput, at a lower cost than the previous IBM FlashSystem 840. Figure 1-1 on page 3 shows the three major areas within IBM FlashCore technology and the unique IBM attributes of each one.

# IBM FlashCore™ Technology

## *The DNA of the FlashSystem family*

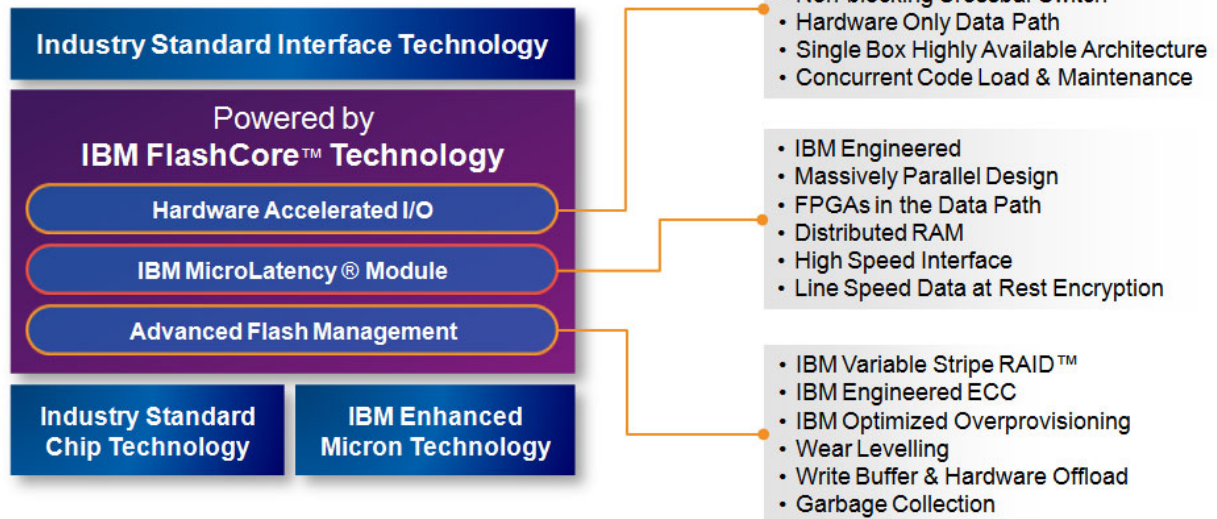


Figure 1-1 IBM FlashCore Technology

### ► Hardware Accelerated I/O

IBM FlashSystem 900 hardware design offers several unique IBM components including Hardware RAID, Non-Blocking Crossbar Switch, Hardware-Only Data Path, Single-Box Highly Available Architecture, Concurrent Code Load, and Concurrent Maintenance.

### ► IBM MicroLatency modules

IBM FlashSystem 900 uses the new 20nm IBM enhanced multi-level cell (MLC) flash card memory chips and either 1.2 TB, 2.9 TB, or 5.7 TB capacity IBM MicroLatency modules. The IBM FlashSystem 900 design also employs the use of IBM Engineered Massively Parallel Design, FPGAs in the Data Path, Distributed RAM, and High-Speed Interfaces plus Hardware-based Data-at-Rest Encryption.

### ► Advanced Flash Management

IBM FlashSystem 900 has unique patented designs to ensure maximum availability and which include IBM Variable Stripe RAID™, IBM Engineered ECC (error correction code), IBM Optimized Over-Provisioning, advanced wear leveling on IBM MicroLatency modules, Write Buffer and Hardware Offload, and IBM Garbage Collection. See “Terminology” on page 19.

All this is made possible because of the following IBM patented and world class innovations:

- ECC algorithms that correct very high bit-error rates
- Variable voltage and read level shifting which help to maximize flash endurance
- Health binning and heat segregation, which continually monitor the health of flash blocks and perform asymmetrical wear leveling and sub-chip tiering

This all results in giving up to 57% improvement in endurance with a potential 45% reduction in write amplification.

### 1.2.1 Flash wear guarantee

Through close collaboration between IBM research, software development, and flash engineering, IBM created an advanced flash characterization platform to test thousands of flash devices over a lifetime of wear.

As a result, IBM has made major strides in advanced flash management. This improves MLC flash endurance 9x over standard implementations, and provides enterprise reliability and performance with IBM MicroLatency.

IBM FlashSystem 900, including its MicroLatency Flash Modules, is covered by up to seven years of total hardware support through the applicable warranty period plus up to six years of optional post-warranty hardware maintenance. Clients may purchase the post-warranty hardware maintenance either at the time of system purchase or up until IBM announces withdrawal from marketing or withdrawal from service.

## 1.3 Why flash technology matters

Flash technology is vibrant and fast growing. 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 the storage economics:

- ▶ Reduces application and server licensing costs, especially those related to databases and virtualization solutions.
- ▶ Improves application efficiency, that is, 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 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: central processing units, network, storage area networks (SANs), and memory. All of them have improved their speeds by a minimum of 10X; some of them by 100X, such as data networks. However, spinning disk has only increased its performance 1.2 times.

IBM FlashSystem 900 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 because of the high density and high performance per density flash solutions that are replacing racks 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 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, and cooling. Flash is one of the best tools for the data center manager for improving data center economics.

## 1.4 IBM FlashSystem family: Product differentiation

Flash is used widely in the data center, either within a server (Peripheral Component Interconnect Express (PCIe) cards) or internal solid-state drives (SSDs), in storage arrays (hybrid or all-flash), appliances, or platform solutions (hardware, software, network). Flash can be used as cache or as a data tier. Because of the vast and wide adoption of flash, several flash architectures exist and, therefore, criteria that can be applied to compare flash options. See Figure 1-2.

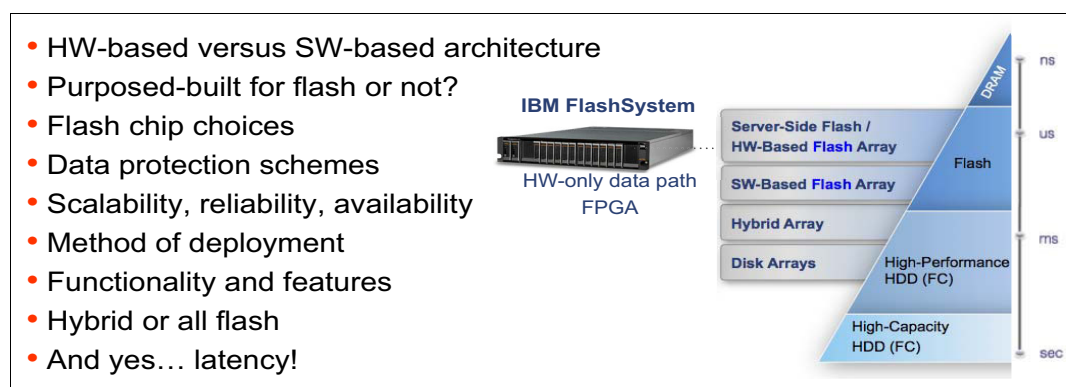


Figure 1-2 The different deployments of Flash

Most storage vendors use and promote flash. The difference is how it is implemented, and the impact 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 it provides extremely low latency and consistent response times. It is designed and built specifically for flash.

Some other vendors create flash appliances based on commodity server platforms and use software-heavy stacks. Some suppliers use hardware technologies that are designed and created for disk, not flash. Some hybrid arrays combine existing storage designs, spinning HDDs, and solid-state disk (SSD).

The IBM storage portfolio includes SSD and flash on various storage platforms; however, these alternative solutions do not have the same low latency (MicroLatency) as the hardware-accelerated 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 needed by applications, and the demand for lower response times in particular tasks. An implementation example is the IBM Easy Tier® technology. For an overview of this technology, see “Easy Tier” on page 289.

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.

For more information about IBM Easy Tier, see the following sources:

- ▶ *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933 (see the topic about advanced features for storage efficiency)
- ▶ *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521 (see the chapter about IBM System Storage Easy Tier function)
- ▶ *IBM System Storage DS8000 Easy Tier*, REDP-4667

IBM FlashSystem storage provides a hardware-only data path that realizes all of the potential of flash memory. These systems differ from traditional storage systems, in both 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 products are fast and efficient. The hardware-only data path has a minimum number of software layers, which are mostly firmware components, and management software that is separated from the data path (out-of-band). 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.

## **1.5 Technology and architectural design overview**

The IBM FlashSystem, with an all-hardware data path using field programmable-gate array (FPGA) modules, is engineered to deliver the lowest possible latency. The modules incorporate proprietary flash controllers and use numerous patented technologies. The FlashSystem controllers have a proprietary logic design, firmware, and system software.

There are no commodity 2.5-inch SSDs, PCIe cards, or any other significant non IBM assemblies within the system. The flash chips, FPGA chips, processors, and other semiconductors in the system are carefully selected to be consistent with the purpose-built design, which is designed for high performance, reliability, and efficiency.

The IBM FlashSystem storage systems offer the following notable architectural concepts:

- ▶ Hardware-only data path.
- ▶ Use of FPGAs extensively.
- ▶ Field-upgradable hardware logic.
- ▶ Less expensive design cycle.
- ▶ Extremely high degree of parallelism.
- ▶ Intelligent flash modules.
- ▶ Distributed computing model.
- ▶ Low-power IBM PowerPC® processors (PPCs).
- ▶ Interface and flash processors run thin real-time operating systems.
- ▶ With minimal management communication, the management processor communicates with the interface and flash processors through an internal network.

### 1.5.1 Hardware-only data path

The hardware-only data path design of the IBM FlashSystem eliminates software layer latency. To achieve extremely low latencies, the IBM FlashSystem advanced software functions are carefully assessed and implemented on a limited basis. For environments that require advanced storage services, implementing the IBM FlashSystem with IBM SAN Volume Controller, which delivers IBM Spectrum Virtualize technology, can offer an unmatched combination of performance, low latency, and rich software functionality.

#### Notes:

- ▶ IBM SAN Volume Controller (SVC) delivers the functions of IBM Spectrum Virtualize, part of the IBM Spectrum Storage family, and has been improving infrastructure flexibility and data economics for more than 10 years. Its innovative data virtualization capabilities provide the foundation for the entire IBM Storwize® family. SAN Volume Controller provides the latest storage technologies for unlocking the business value of stored data, including virtualization and IBM Real-time Compression™.
- ▶ In addition, the latest system includes the new SAN Volume Controller Data Engine to help support the massive volumes of data created by today's demanding enterprise applications. SAN Volume Controller is designed to deliver unprecedented levels of efficiency, ease of use and dependability for organizations of all sizes.

In the IBM FlashSystem, data traverses the array controllers through FPGAs and dedicated, low-power CPUs. There are no wasted cycles on *interface* translation, protocol control, or tiering.

IBM FlashSystem 900, with an all-hardware data path design, has an internal architecture that differs from other hybrid (SSD and HDD) or SSD-only-based disk systems.

## 1.5.2 The 20nm flash card memory chips

The *flash chip* is the basic storage component of the IBM MicroLatency module. A maximum of 64 multi-level cell (MLC) flash chips can exist for each flash module. Combining flash chips of different flash technologies is not supported in the same flash module or storage system to maintain consistent wearing and reliability.

The IBM FlashSystem 900 employs the new 20nm MLC chips, which are of a higher density than the 24nm eMLC chips that are used in the IBM FlashSystem 840. This new design of chips allows the IBM FlashSystem 900 to package greater densities of flash memory per card whilst retaining the same if not better performance and wear.

IBM patented ECC correction and checking algorithms ensure the same or greater performance from the MLC based chips, with a greater capacity for the same footprint and at a lower cost per TB.

## 1.5.3 Flash module capacities

IBM FlashSystem 900 uses either 1.2 TB, 2.9 TB, or 5.7 TB installed IBM MicroLatency modules. This is a 40% increase in capacity per module over the IBM FlashSystem 840. The modules must be of the same capacity throughout the machine and cannot be intermixed with the older 24nm flash modules.

Only RAID 5 is supported on the IBM FlashSystem 900 with configurations of 4, 6, 8, 10, and 12 modules when using the 1.2 TB IBM MicroLatency modules, and 6, 8, 10, and 12 modules when using the 2.9 TB or 5.7 TB IBM MicroLatency modules

**Empty bays:** If fewer than 12 modules are installed, flash module fillers must be installed in the empty bays to maintain cooling airflow in the system enclosure.

## 1.5.4 Gateway interface FPGA

The gateway interface FPGA is responsible for providing I/O to the flash module and direct memory access (DMA) path. It is located on the flash module and has two connections to the backplane.

## 1.5.5 Flash controller FPGA

The flash controller FGPA of the flash module is used to provide access to the flash chips and is responsible for the following functions:

- ▶ Provides data path and hardware I/O logic
- ▶ Uses lookup tables and a write buffer
- ▶ Controls 13 or 16 chips (module size dependent)
- ▶ Operates independently of other controllers
- ▶ Maintains write ordering and layout
- ▶ Provides write setup
- ▶ Maintains garbage collection
- ▶ Provides error handling

Figure 1-3 on page 9 shows the flash controller design details.

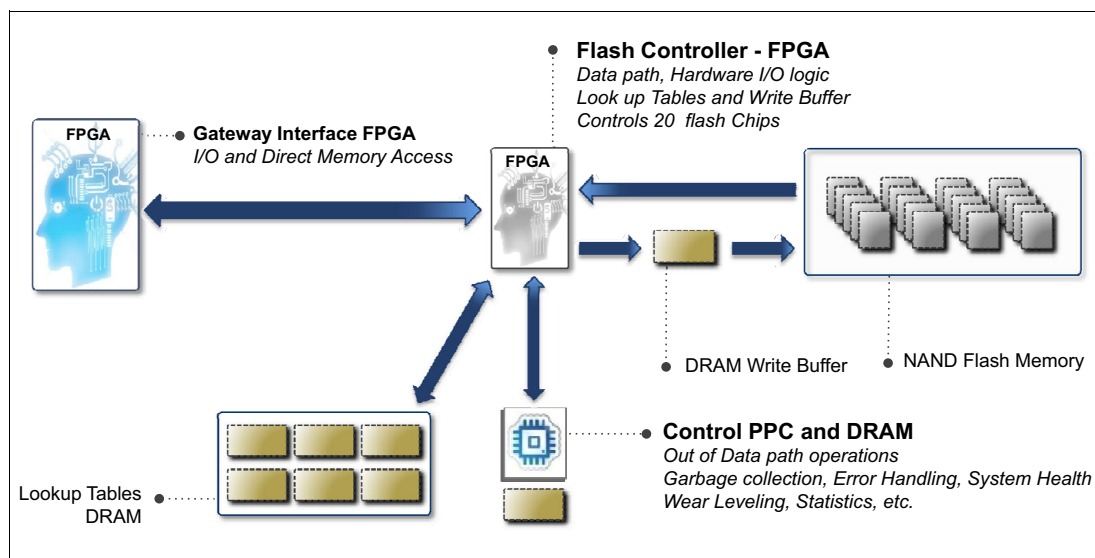


Figure 1-3 IBM FlashSystem controller details

The concurrent operations performed on the flash chips include moving data in and out of the chip through DMA, and by internally moving data and performing erasures. While actively transferring user data in the service of host-initiated I/O, the system can simultaneously run garbage collection activities without affecting the I/O. The ratio of transparent background commands running concurrent to active data transfer commands is 7:1.

A maximum of four flash controllers per IBM MicroLatency module exist: two for each primary board and two for each expansion board.

### 1.5.6 IBM Variable Stripe RAID and 2D Flash RAID overview

Storage systems of any kind are typically designed to perform two main functions: to store and protect data. IBM FlashSystem 900 includes the following options for data protection:

- ▶ RAID data protection:
  - IBM Variable Stripe RAID
  - Two-dimensional (2D) Flash RAID
- ▶ Flash memory protection methods
- ▶ Optimized RAID rebuild times

Table 1-1 lists the various methods of protection.

Table 1-1 Various types of IBM FlashSystem protection

Layer	Managed by	Protection
System-level RAID 5	Centralized RAID controllers	Module failure
Module-level RAID 5	Each module across the chips	Chip failure and page failure
Module-level Variable Stripe RAID	Each module across the chips	Subchip, chip, or multi-chip failure
Chip-level error correction code (ECC)	Each module using the chips	Bit and block error

**Note:** The proprietary 2D Flash RAID data protection scheme of the IBM FlashSystem 900 storage system combines system-level RAID 5 and module-level Variable Stripe RAID (not only module-level RAID).

## 1.6 Variable Stripe RAID

*Variable Stripe RAID* is a unique IBM technology that provides data protection of the memory page, block, or whole chip, which eliminates the necessity to replace a whole flash module in a single memory chip failure or plane failures. This, in turn, expands the life and endurance of flash modules and reduces considerably maintenance events throughout the life of the system.

Variable Stripe RAID provides high redundancy across chips within a flash module. RAID is implemented at multiple addressable segments within chips, in a  $n+1$  RAID 5 fashion, and it is controlled at the flash controller level (up to four in each flash module). Because of the massive parallelism of DMA operations controlled by each FPGA and parallel access to chip sets, dies, planes, blocks, and pages, the implementation of Variable Stripe RAID has minimal impact on performance.

The following information describes some of the most important aspects of Variable Stripe RAID implementation:

- ▶ Variable Stripe RAID is managed and controlled by each of the four flash controllers within a single module.
- ▶ A flash controller is in charge of only 13 or 16 flash chips. (IBM MicroLatency module capacity size dependent).
- ▶ Data is written on flash pages of 8 KB and erased in 1 MB flash blocks.
- ▶ Variable Stripe RAID is implemented and managed at flash chip *plane* levels.
- ▶ There are 16 planes per chip.
- ▶ Before a plane fails, at least 256 flash blocks within a plane must be deemed *failed*.
- ▶ A plane can also fail in its entirety.
- ▶ Up to 64 planes can fail before a whole module is considered failed.
- ▶ Up to four chips can fail before a whole module is considered failed.
- ▶ When a flash module is considered failed, 2D Flash RAID takes control of data protection and recovery.
- ▶ When a plane or a chip fails, Variable Stripe RAID activates to protect data while maintaining system-level performance and capacity.

### 1.6.1 How Variable Stripe RAID works

Variable Stripe RAID is an IBM patented technology. It includes but is more advanced than a simple RAID of flash chips. Variable Stripe RAID introduces two key concepts:

- ▶ The RAID stripe is not solely across chips; it actually spans across flash layers.
- ▶ The RAID stripe can automatically vary based on observed flash plane failures within a flash module. For example, stripes are not fixed at  $n+1$  RAID 5 stripe members, but they can go down to  $n-1$ ,  $n-2$ , or even  $n-3$  based on plane failures.

This ability to protect the data at *variable* stripes effectively maximizes flash capacity even after flash component failures. Figure 1-4 shows an overview of the IBM FlashSystem Variable Stripe RAID.

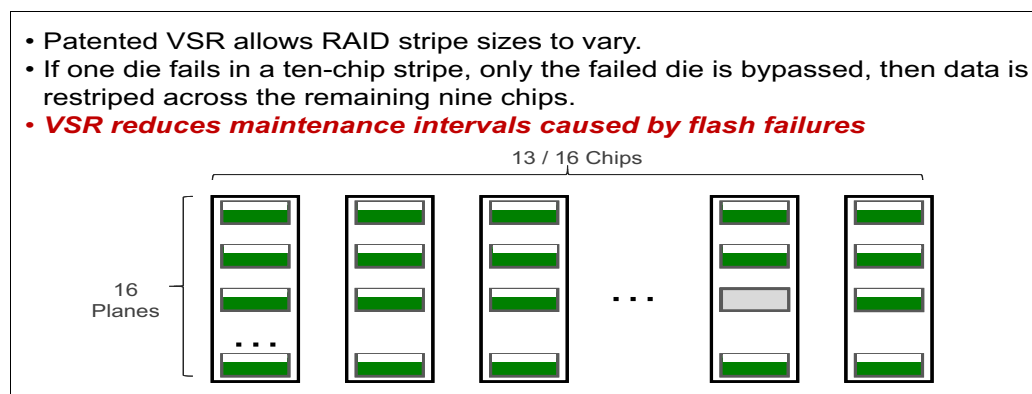


Figure 1-4 IBM FlashSystem Variable Stripe RAID (VSR)

Figure 1-5 shows the benefits of IBM Variable Stripe RAID.

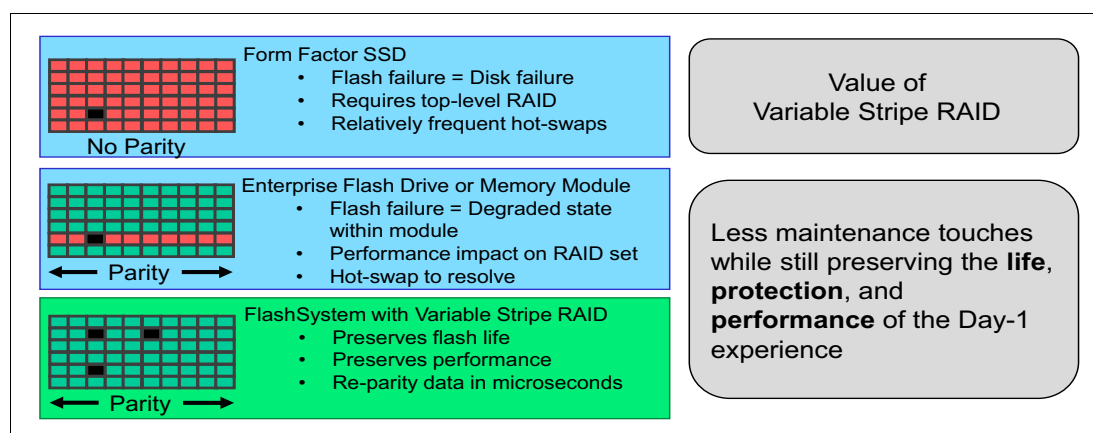


Figure 1-5 The value of the IBM FlashSystem Variable Stripe RAID

An important aspect to emphasize is that Variable Stripe RAID has an effect at only the *plane* level. Therefore, *only* the affected planes within a plane failure are converted to (N-1). Variable Stripe RAID maintains the current stripe member count (n+1) layout through the rest of the areas of all other planes and chips that are not involved in the plane failure.

To illustrate how Variable Stripe RAID functions, assume that a plane fails within a flash chip and is no longer available to store data. This might occur as a result of a physical failure within the chip, or some damage is inflicted on the address or power lines to the chip. The plane failure is detected and the system changes the format of the page stripes that are used. The data that was previously stored in physical locations across chips in all ten lanes using a page stripe format with ten pages is now stored across chips in only nine lanes using a page stripe format with nine pages. Therefore, no data stored in the memory system was lost, and the memory system can self-adapt to the failure and continue to perform and operate by processing read and write requests from host devices.

This ability of the system to automatically self-adapt, when needed, to chip and intra-chip failures makes the FlashSystem flash module extremely rugged and robust, and capable of operating despite the failure of one or more chips or intra-chip regions. This ability also makes the system easier to use because the failure of one, two, or even more individual memory chips or devices does not require the removal and potential disposal of previously used memory storage components.

The reconfiguration or reformatting of the data to change the page stripe formatting to account for chip or intra-chip failures might reduce the amount of physical memory space that is held in reserve by the system and available for the system for background operation. In all but the most extreme circumstances (in which case the system creates alerts), it does not affect usable capacity or performance.

### **Reliability, availability, and serviceability**

The previous explanation points out an increase in reliability, availability, and serviceability (RAS) levels and the IBM FlashSystem RAS levels over other technologies.

In summary, Variable Stripe RAID has these capabilities:

- ▶ Patented Variable Stripe RAID allows RAID stripe sizes to vary.
- ▶ If one plane fails in a chip stripe, only the failed plane is bypassed, and then data is restriped across the remaining chips. No system rebuild is needed.
- ▶ Variable Stripe RAID reduces maintenance intervals caused by flash failures.

## **1.7 Two-dimensional (2D) Flash RAID**

*Two-dimensional (2D) Flash RAID* refers to the combination of Variable Stripe RAID (at the flash module level) and system-level RAID 5.

The second dimension of data protection is implemented across flash modules of RAID 5 protection. This system-level RAID 5 is striped across the appropriate number of flash modules in the system based on the selected configuration. System-level RAID 5 can stripe across any of these flash modules:

- ▶ Four (2D+1P+1S: note IBM MicroLatency 1.2 TB module only)
- ▶ Six (4D+1P+1S)
- ▶ Eight (6D+1P+1S)
- ▶ Ten (8D+1P+1S)
- ▶ Twelve (10D+1P+1S).

The architecture allows you to designate a dynamic flash module hot spare.



Figure 1-6 shows the IBM FlashSystem 2D RAID.

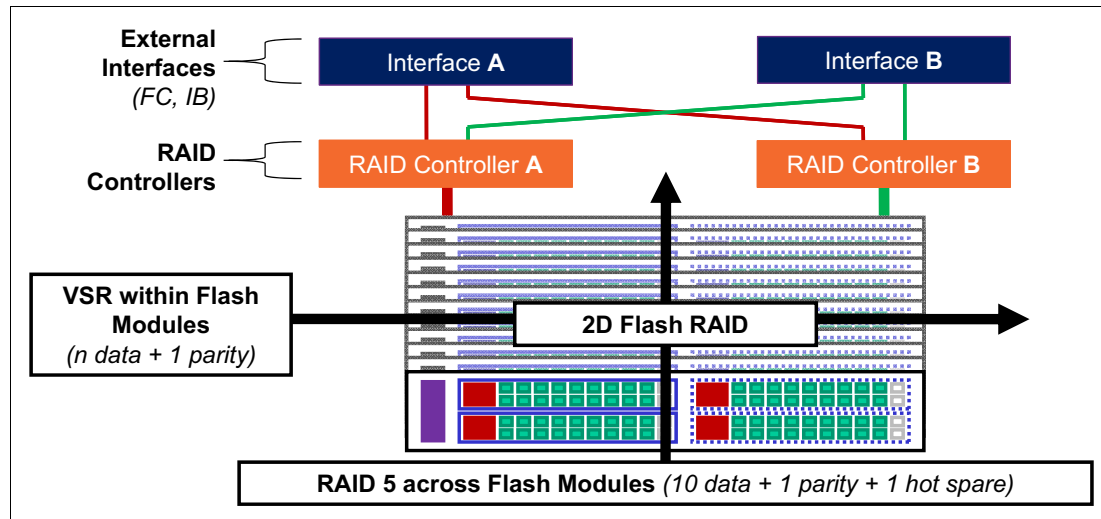


Figure 1-6 IBM FlashSystem 2D RAID

The 2D Flash RAID technology within the IBM FlashSystem provides two independent layers of RAID 5 data protection within each system: the module-level Variable Stripe RAID technology and an additional system-level RAID 5 across flash modules. When operating in system-level RAID 5 mode, redundant centralized RAID controllers create a stripe arrangement across the four, six, eight, ten, or twelve flash modules in the system.

The system-level RAID 5 complements the Variable Stripe RAID technology implemented within each flash module, and it provides protection against data loss and data unavailability resulting from flash module failures. It also allows data to be rebuilt onto a hot spare flash module, so that flash modules can be replaced without data disruption.

In addition to 2D Flash RAID and Variable Stripe RAID data protection, the IBM FlashSystem family storage systems incorporate other reliability features:

- ▶ Error correction codes to provide bit-level reconstruction of data from flash chips.
- ▶ Checksums and data integrity fields designed to protect all internal data transfers within the system.
- ▶ Overprovisioning to enhance write endurance and decrease write amplification.
- ▶ Wear-leveling algorithms balance the number of writes among flash chips throughout the system.
- ▶ Sweeper algorithms help ensure that all data within the system is read periodically to avoid data fade issues.

Understanding 2D Flash RAID allows you to visualize the advantage over other flash storage solutions. Both Variable Stripe RAID and 2D Flash RAID are implemented and controlled at FPGA hardware-based levels. The 2D Flash RAID eliminates single points of failure and provides enhanced system-level reliability.





# IBM FlashSystem 900 architecture

The IBM FlashSystem 900 architecture is described here. It includes an introduction and product features, a comparison to its predecessor, the IBM FlashSystem 840, and an overview of the architecture and hardware. Also included is an overview of the administration and serviceability of the IBM FlashSystem 900, interoperability, and integration with other IBM products.

For more details about IBM FlashSystem architecture, see the IBM FlashSystem 900 web page at the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

This chapter covers the following topics:

- ▶ Introduction to the IBM FlashSystem 900 architecture
- ▶ Architecture of IBM FlashSystem 900
- ▶ Administration and maintenance
- ▶ Support matrix
- ▶ Product integration overview

## 2.1 Introduction to the IBM FlashSystem 900 architecture

The IBM FlashSystem 900 is an all-flash storage array that provides extreme performance and large capacity while also providing enterprise class reliability and “green” data center power and cooling requirements. The IBM FlashSystem 900 holds up to twelve 5.7 TB IBM MicroLatency modules in only 2U of rack space, making it an extremely dense all-flash storage array solution.

The IBM FlashSystem 900 also provides performance up to 1,100,000 I/O per second (IOPS), bandwidth up to 10 GBps, and latency as low as 90 microseconds. This high capacity, extreme performance, and enterprise reliability are powered by the patented IBM FlashCore Technology. The IBM FlashSystem 900 supports several protocols, including FC (16, 8, and 4 Gbps), Fibre Channel over Ethernet (FCoE), iSCSI, and InfiniBand, enabling connections to high performance servers and storage area networks.

The IBM FlashSystem 900 core attributes are described next. Figure 2-1 shows the front view of the IBM FlashSystem 900.



Figure 2-1 IBM FlashSystem 900 front view

### 2.1.1 Capacity

IBM FlashSystem 900 supports a maximum of twelve 5.7 TB IBM MicroLatency modules, which provide a maximum capacity of 57 TB (RAID 5). The IBM FlashSystem 900 can be ordered with four, six, eight, ten, or twelve MicroLatency modules. The available MicroLatency modules are either 1.2 TB, 2.9 TB, or 5.7 TB storage capacity.

**Important:** 1.2 TB, 2.9 TB, and 5.7 TB IBM MicroLatency modules cannot be intermixed in the same IBM FlashSystem 900 chassis.

IBM FlashSystem 900 supports RAID 5 configurations.

IBM FlashSystem 900 supports the creation of up to 2,048 logical unit numbers (LUNs). The size of the LUNs can be 1 MiB - 51.8 TiB in size (not to exceed the total system capacity). The IBM FlashSystem 900 supports up to 2,048 host connections and up to 256 host connections for each interface port. IBM FlashSystem 900 allows the mapping of multiple LUNs to each host for Fibre Channel, Fibre Channel over Ethernet (FCoE), iSCSI, and InfiniBand protocols.

IBM FlashSystem 900 supports up to 256 host connections for the iSCSI protocol. Table 2-1 lists all combinations of storage capacities for various configurations of the IBM FlashSystem 900.

Table 2-1 IBM FlashSystem 900 capacity in TB for RAID 5

IBM FlashSystem 900 configuration	RAID 5 TB
Four 1.2 TB flash modules	2.2
Six 1.2 TB flash modules	4.5
Eight 1.2 TB flash modules	6.8
Ten 1.2 TB flash modules	9.1
Twelve 1.2 TB flash modules	11.4
Six 2.9 TB flash modules	11.4
Eight 2.9 TB flash modules	17.1
Ten 2.9 TB flash modules	22.8
Twelve 2.9 TB flash modules	28.5
Six 5.7 TB flash modules	22.8
Eight 5.7 TB flash modules	34.2
Ten 5.7 TB flash modules	45.6
Twelve 5.7 TB flash modules	57

## Exact byte counts

The following exact byte counts are for the flash modules that are used in the IBM FlashSystem 9000:

- 5.7 TB module:
  - Presented: 5717176090624 bytes (after Variable Stripe RAID and reserve).
  - Raw flash: 8796093022208 bytes (64 chips x 128 GiB).
  - The IBM FlashSystem V9000 GUI reports 5.24 TiB usable.
- 2.9 TB module:
  - Presented: 2858588045312 bytes (after Variable Stripe RAID and reserve).
  - Raw flash: 4398046511104 bytes (32 chips x 128 GiB).
  - The IBM FlashSystem V9000 GUI reports 2.62 TiB usable.
- 1.2 TB module:
  - Presented: 1143435427840 bytes (after Variable Stripe RAID and reserve).
  - Raw flash: 1786706395136 bytes (13 chips x 128 GiB).
  - The IBM FlashSystem V9000 GUI reports 1.06 TiB usable.

## 2.1.2 Performance and latency

The IBM FlashSystem 900 uses all hardware field-programmable gateway array (FPGA) components in the data path, which enables fast I/O rates and low latency. The IBM FlashSystem 900 provides extreme performance of up to 1,100,000 IOPS and bandwidth up to 10 GBps. The IBM FlashSystem 900 provides write latency as low as 90  $\mu$ s and read latency as low as 155  $\mu$ s.

### 2.1.3 Power requirements

The IBM FlashSystem 900 is *green data center friendly*. The IBM FlashSystem 900 consumes only 625 W of power (steady state RAID 5 configuration for a 70/30 read/write workload on an eight-module 2.9 TB flashcard system) and uses two standard single-phase (100V - 240V) electrical outlets.

Consider the following information:

- ▶ Plan to attach each of the two power supplies in the enclosure to separate main power supply lines.
- ▶ The 1300 W power supply for high-line voltage provides IBM FlashSystem 900 with high-power to run at maximum performance for longer durations during power supply servicing, resulting in more predictable performance under unexpected failure conditions. Optimal operation is achieved when operating between 200V - 240V (Nominal). These are the maximum and minimum voltage ranges (Vrms) and associated high line AC ranges:
  - Minimum: 180V.
  - Nominal: 200V - 240V.
  - Maximum: 265V.
  - Using two power sources provides power redundancy. Place the two power supplies on separate circuits.

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

### 2.1.4 Physical specifications

The IBM FlashSystem 900 installs in a standard 19-inch equipment rack and is 2U high and 19 inches wide. A standard data 42U 19-inch data center rack can be fully populated with 21 IBM FlashSystem 900 storage systems.

The IBM FlashSystem 900 has the following physical dimensions:

- ▶ Height: 8.62 mm (3.39 inches)
- ▶ Width: 445 mm (17.5 in; 19 inch Rack Standard)
- ▶ Depth: 761 mm (29.96 in)
- ▶ Weight (maximum configuration - 12 flash modules): 17.69 kg (39 lb)
- ▶ Airflow path: Cool air flows into the front of the unit (intake) to the rear of the unit (exhaust)
- ▶ Heat: 2133 BTU per hour (assuming 625W)

### 2.1.5 FlashCore technology

The IBM FlashSystem 900 provides enterprise class reliability and serviceability that are unique for the all-flash storage arrays. The IBM FlashSystem 900 uses the FlashCore technology to provide data protection and maximum system uptime.

- ▶ IBM Advanced Flash Management improves flash endurance 9x over standard implementations.
  - Proprietary garbage collection, relocation, and block picking algorithms invented by IBM. See “Terminology” on page 19.
  - Flash wear leveling includes the following functions:
    - ECC algorithms that correct very high bit error rates.
    - Variable voltage and read level shifting to maximize flash endurance.

- Health binning and heat segregation continually monitors the health of flash blocks and performs symmetrical wear leveling and sub-chip tiering.
  - Hot Data Placement provides up to 57% improvement in endurance. Heat Level Grouping provides up to 45% reduction in write amplification.
- ▶ Variable Stripe RAID: A patented IBM technology that provides an intramodule RAID stripe on each flash module.
  - ▶ Two-dimensional (2D) Flash RAID: System-wide RAID 5 along with Variable Stripe RAID helps reduce downtime and maintains performance and allows the provisioning of an entire flash module as a spare to be used in another flash module failure.

## Terminology

The following terms are mentioned in this book:

<b>Wear leveling</b>	An algorithm that assures even usage of all blocks.
<b>Garbage collection</b>	Erasing blocks, which are not used anymore, so that they can be rewritten.
<b>Relocation</b>	Moving a block to another location.
<b>Block picking</b>	The first step of the garbage collection process. Using proprietary algorithms, the best block is picked for garbage collection.

## Extra reliability and serviceability features of IBM FlashSystem 900

In addition to the standard features, the IBM FlashSystem 900 includes the following reliability and serviceability features:

- ▶ Hot-swappable IBM MicroLatency modules by way of tool-less front panel access  
If a MicroLatency module failure occurs, critical client applications can remain online while the defective module is replaced.  
  
Because client application downtime does not need to be scheduled, you can typically perform this service immediately instead of waiting days for a service window. The *directed maintenance procedure*, accessible from the GUI, can be used to prepare the IBM FlashSystem 900 for a MicroLatency module replacement. You can remove the MicroLatency modules easily from the front of the IBM FlashSystem 900 unit without needing to remove the top access panels or extend cabling.
- ▶ Concurrent code loads  
The IBM FlashSystem 900 supports concurrent code load, enabling client applications to remain online during firmware upgrades to all components, including the flash modules.
- ▶ Redundant hot-swappable components  
RAID controllers called *canisters*, management modules, and interface cards (all contained in the canister), batteries, fans, and power supplies are all redundant and hot swappable. All components are easily accessible through the front or rear of the unit so the IBM FlashSystem 900 does not need to be moved in the rack, and top access panels or cables do not need to be extended. This makes servicing the unit easy.

**Tip:** Concurrent code loads require that all connected hosts have at least two connections (at least one to each canister) to the FlashSystem 900. For more information, see 10.4, “Host attachment hints” on page 326.

## 2.1.6 Scalability

The IBM FlashSystem 900, which supports the growth of the storage capacity after deployment, supports a maximum configuration of twelve 1.2 TB, 2.9 TB, or 5.7 TB IBM MicroLatency modules. The IBM FlashSystem 900 can be purchased with 2, 4, 6, 8, 10, or 12 1.2 TB, 2.9 TB, or 5.7 TB modules.

The IBM FlashSystem 900 offers these upgrade options:

- ▶ Systems that are purchased with 4 MicroLatency modules can be expanded to 6, 8, 10, or 12 of the same capacity MicroLatency modules.
- ▶ Systems that are purchased with 6 MicroLatency modules can be expanded to 8, 10, or 12 of the same capacity MicroLatency modules.
- ▶ Systems that are purchased with 8 MicroLatency modules can be expanded to 10 or 12 of the same capacity MicroLatency modules.
- ▶ Systems that are purchased with 10 MicroLatency modules can be expanded to 12 of the same capacity MicroLatency modules.

### **Important considerations:**

- ▶ Mixing different capacity MicroLatency modules (1.2 TB, 2.9 TB, or 5.7 TB) in any configuration on the IBM FlashSystem 900 is not supported.
- ▶ If an IBM FlashSystem 900 is purchased with 1.2 TB MicroLatency modules, all system expansions must be with 1.2 TB MicroLatency modules.
- ▶ If an IBM FlashSystem 900 is purchased with 2.9 TB MicroLatency modules, all system expansions must be with 2.9 TB MicroLatency modules.
- ▶ If an IBM FlashSystem 900 is purchased with 5.7 TB MicroLatency modules, all system expansions must be with 5.7 TB MicroLatency modules.
- ▶ Expanding an IBM FlashSystem 900 unit with 2, 4, 6, or 8 additional MicroLatency modules requires that the system is reconfigured. A backup of the system configuration and data migration, if needed, must be planned before the expansion.

## 2.1.7 Protocol support

The IBM FlashSystem 900 supports the following interface protocols and number of connections:

- ▶ Fibre Channel (16 ports of 4 Gbps or 8 Gbps)
- ▶ Fibre Channel (8 ports of 16 Gbps; these ports also support 8 Gbps and 4 Gbps)
- ▶ Fibre Channel over Ethernet (FCoE) (16 ports of 10 Gbps FCoE)
- ▶ InfiniBand (8 ports of quad data rate (QDR) 40 Gbps)
- ▶ iSCSI (16 ports of 10 Gbps Ethernet)

### **Important considerations:**

- ▶ The IBM FlashSystem 900 only supports one interface type per system. For example, having two FC interface cards and two InfiniBand interface cards in the same IBM FlashSystem 900 storage system is not possible.
- ▶ The IBM FlashSystem 900 supports eight active ports across the entire system if 16 Gbps FC is enabled. These eight ports can operate at 16, 8, or 4 Gbps.



## 2.1.8 Encryption support

IBM FlashSystem 900 provides optional encryption of data at rest, which protects against the potential exposure of sensitive user data and user metadata that are stored on discarded or stolen flash modules. Encryption of system data and metadata is not required, so system data and metadata are not encrypted.

**SEDs:** Some IBM products, which implement encryption of data at rest, stored on a fixed block storage device, implement encryption using self-encrypting disk drives (SEDs). The IBM FlashSystem 900 flash module chips do not use SEDs. The IBM FlashSystem 900 data encryption and decryption are performed by the IBM MicroLatency modules, which can be thought of as the functional equivalent of Self-Encrypting Flash Controller (SEFC) cards.

The general encryption concepts and terms for the IBM FlashSystem 900 are as follows:

- ▶ *Encryption-capable* refers to the ability of the IBM FlashSystem 900 to optionally encrypt user data and metadata by using a secret key.
- ▶ *Encryption-disabled* describes a system where no secret key is configured. The secret key is neither required, nor used, to encrypt or decrypt user data. Encryption logic is actually still implemented by the IBM FlashSystem 900 while in the encryption-disabled state, but uses a default, or well-known, key. Therefore, in terms of security, encryption-disabled is effectively the same as not encrypting at all.
- ▶ *Encryption-enabled* describes a system where a secret key is configured and used. This does not necessarily mean that any access control was configured to ensure that the system is operating securely. Encryption-enabled only means that the system is encrypting user data and metadata using the secret key.
- ▶ *Access-control-enabled* describes an encryption-enabled system that is configured so that an access key must be provided to authenticate with an encrypted entity, such as a secret key or flash module, to unlock and operate that entity. The IBM FlashSystem 900 permits access control enablement only when it is encryption-enabled. A system that is encryption-enabled can optionally also be access-control-enabled to provide functional security.
- ▶ *Protection-enabled* describes a system that is both encryption-enabled and access-control-enabled. An access key must be provided to unlock the IBM FlashSystem 900 so that it can transparently perform all required encryption-related functionality, such as encrypt on write and decrypt on read.
- ▶ *Protection Enablement Process (PEP)* transitions the IBM FlashSystem 900 from a state that is not protection-enabled to a state that is protection-enabled. The PEP requires that the client provide a secret key to access the system, and the secret key must be resiliently stored and backed up externally to the system, for example, on a USB flash drive.

PEP is not merely activating a feature through the GUI or CLI. To avoid the loss of data that was written to the system before the PEP occurs, the client must move all of the data to be retained off the system before the PEP is initiated, and then must move the data back onto the system after the PEP completes. The PEP is performed during the system initialization process, if encryption is activated.
- ▶ *Application-transparent encryption* is an attribute of the IBM FlashSystem 900 encryption architecture; it means that applications are not aware that encryption and protection are occurring. This is in contrast to application-managed encryption (AME), where an application must serve keys to a storage device.

- ▶ *Hot Key Activation* is the process of changing an *encryption-disabled* FlashSystem 900 to *encryption-enabled* while the system is running.
- ▶ *Non-Disruptive Rekey* is the process of creating a new encryption key that supersedes the existing key on a running FlashSystem 900.

**License:** The IBM FlashSystem 900 requires a license for encryption. If encryption is required, validate with IBM marketing or your IBM Business Partner that the license is ordered with the equipment.

## Configuring encryption

You can activate encryption with the easy setup wizard during initialization or the Hot Key Activation process after the FlashSystem 900 has already been initialized, when an encryption feature code is purchased. If encryption is activated, an encryption key is generated by the system to be used for access to the system. The process invokes a wizard that guides the user through the process of copying the encryption key to multiple USB keys.

The IBM FlashSystem 900 provides support for Encryption Rekey to create new encryption keys that supersede the existing encryption keys.

**Note:** If you plan to implement either Hot Key Activation or Encryption Rekey, we suggest that you inform IBM Support to monitor the operation.

## Accessing an encrypted system

At system start-up (power on) or to access an encrypted system, the encryption key must be provided by an outside source so that the IBM FlashSystem 900 can be accessed. The encryption key is provided by inserting the USB flash drives that were created during system initialization into a canister.

## Encryption technology

Key encryption is protected by an Advanced Encryption Standard (XTS-AES) algorithm key wrap using the 256-bit symmetric option in XTS mode, as defined in the IEEE1619-2007 standard. An HMAC-SHA256 algorithm is used to create a hash message authentication code (HMAC) for corruption detection, and it is additionally protected by a system-generated cyclic redundancy check (CRC).

## 2.1.9 Comparison of IBM FlashSystem models 840 and 900

Table 2-2 lists differences between the IBM FlashSystem 840 and the IBM FlashSystem 900 modules.

Table 2-2 IBM FlashSystem 840 and 900 comparison

Feature to compare	IBM FlashSystem 840	IBM FlashSystem 900
<b>Storage capacity options (TB)</b>	2, 4, 6, 8, 10, 12, 16, 20, 24, 32, 40, and 48	2.2, 4.5, 6.8, 9.1, 11.4, 11.4, 17.1, 22.8, 28.5, 22.8, 34.2, 45.6, and 57
<b>Form factor</b>	2U	2U
<b>Performance (IOPS)</b>	1,100,000	1,100,000
<b>Bandwidth</b>	8 GBps	10 GBps
<b>Latency (read/write)</b>	135µs/90µs	155µs/90µs
<b>Available interfaces</b>	16, 8, and 4 Gbps FC 40 Gbps QDR InfiniBand 10 Gbps FCoE 10 Gbps iSCSI	16, 8, and 4 Gbps FC 40 Gbps QDR InfiniBand 10 Gbps FCoE 10 Gbps iSCSI
<b>Chip type</b>	eMLC	IBM enhanced MLC
<b>Chip RAID</b>	Yes	Yes
<b>System RAID</b>	Yes	Yes
<b>Power consumption (steady-state RAID 5)</b>	625W <sup>a</sup>	625W <sup>b</sup>
<b>LUN masking</b>	Yes	Yes
<b>Management</b>	IBM SAN Volume Controller CLI HTML 5.0 c2011 IBM Storage GUI SNMP Email alerts Syslog redirect	IBM SAN Volume Controller CLI HTML 5.0 c2011 IBM Storage GUI SNMP Email alerts Syslog redirect

a. 625W is for a 70/30 read/write workload on an eight-module 2 TB flashcard system.

b. 625W is for a 70/30 read/write workload on an eight-module 2.9 TB flashcard system.

## 2.1.10 Management

The IBM FlashSystem 900 includes state-of-the-art IBM storage management interfaces. The IBM FlashSystem 900 graphical user interface (GUI) and command-line interface (CLI) are updated from previous versions of the IBM FlashSystem products to include the IBM SAN Volume Controller GUI and SAN Volume Controller CLI, which deliver the functionality of IBM Spectrum Virtualize.

The IBM FlashSystem 900 also uses a USB key for system initialization, similar to the IBM V7000 disk system. The IBM FlashSystem 900 also supports Simple Network Management Protocol (SNMP), email notification (Simple Mail Transfer Protocol (SMTP)), and syslog redirection.

Figure 2-2 shows a contextual view of the IBM FlashSystem 900 GUI. For more details about the use of the FlashSystem 900 GUI and CLI, see 2.3.2, “System management” on page 37.

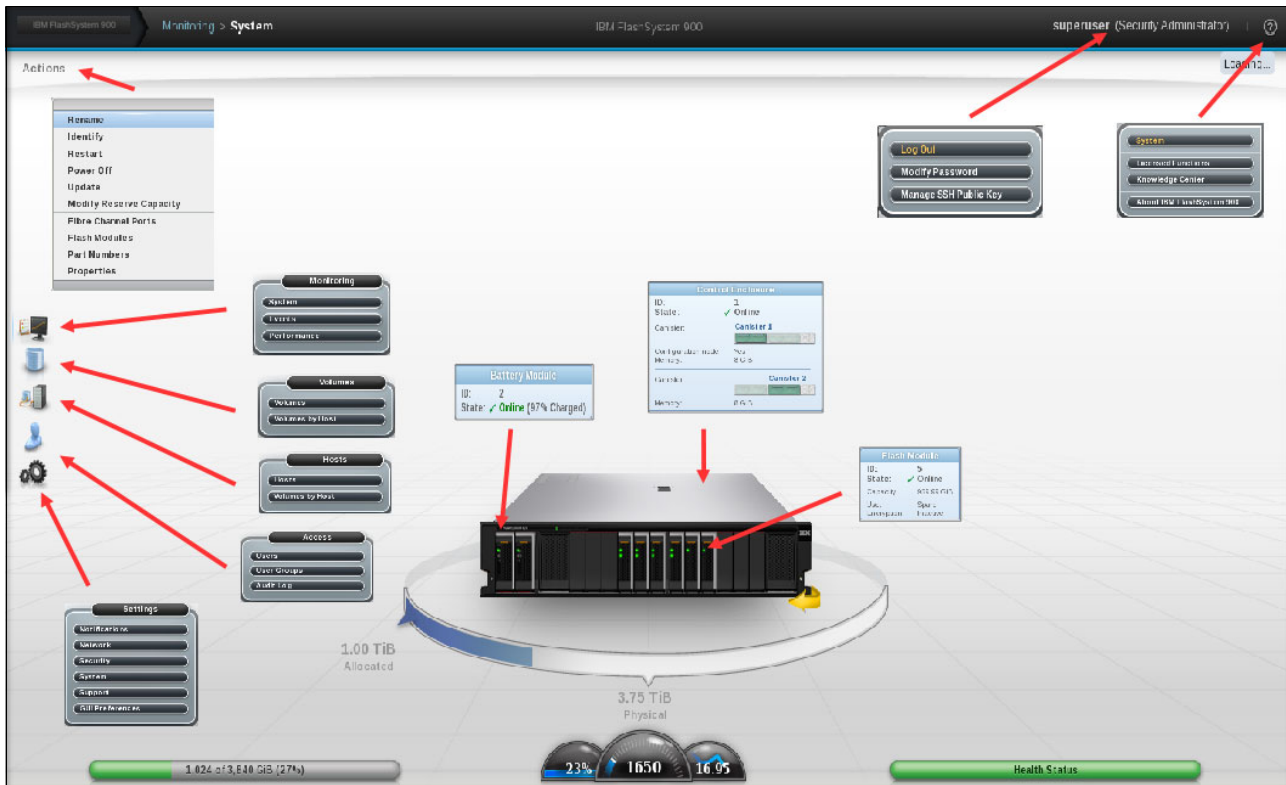


Figure 2-2 IBM FlashSystem 900 GUI

## 2.2 Architecture of IBM FlashSystem 900

The IBM FlashSystem architecture is explained. Key product design characteristics, including performance, reliability, and serviceability, are described. Hardware components are also described.

### 2.2.1 Overview of architecture

The design goals for the IBM FlashSystem 900 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 as little downtime as possible. The IBM FlashSystem 900 uses many FPGA components and as little software as possible, keeping I/O latency to a minimum and I/O performance to a maximum.

Figure 2-3 on page 25 illustrates the IBM FlashSystem 900 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 canisters and all the flash modules connect, and also connections to battery modules, fan modules, and power supply units. The two RAID canisters contain crossbar controllers, management modules, interface controllers and interface adapters, and fan modules. The two RAID canisters 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).

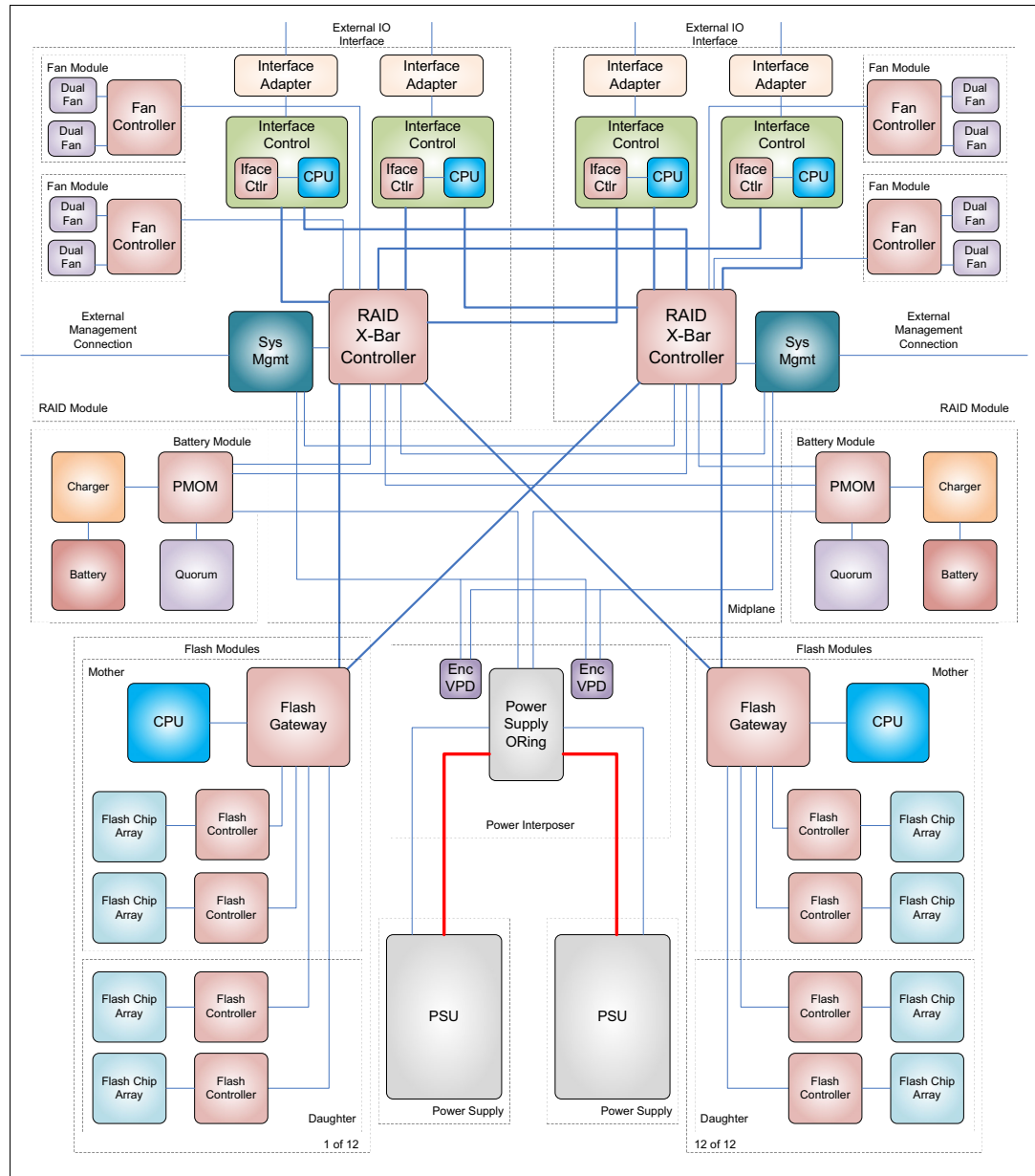


Figure 2-3 IBM FlashSystem 900 architecture

## 2.2.2 Hardware components

The following list shows the core IBM FlashSystem 900 components:

- ▶ Canisters
- ▶ Interface cards
- ▶ IBM MicroLatency modules
- ▶ Battery modules
- ▶ Power supply units
- ▶ Fan modules

Figure 2-4 shows the IBM FlashSystem 900 front view. The two battery modules are to the left and the twelve MicroLatency modules are to the right.



Figure 2-4 IBM FlashSystem 900 front view

Figure 2-5 shows the IBM FlashSystem 900 rear view. The canisters are to the left (large units) and the two power supply units are to the right (small units).



Figure 2-5 IBM FlashSystem 900 rear view

## 2.2.3 Canisters

Each IBM FlashSystem 900 storage system contains two fully redundant canisters. The fan modules are at the bottom and the interface cards are at the top. Each canister contains a RAID controller, two interface cards, and a management controller with an associated 1 Gbps Ethernet port. Each canister also has a USB port and two hot swappable fan modules.

Figure 2-6 on page 27 shows the components of the IBM FlashSystem 900 from the rear. One of the two canisters was removed, and you see two interface cards and two fan modules. The power supply unit to the right of the fans provides redundant power to the system. All components are concurrently maintainable except the midplane and the power interposer, which has no active components. All external connections are from the rear of the system.

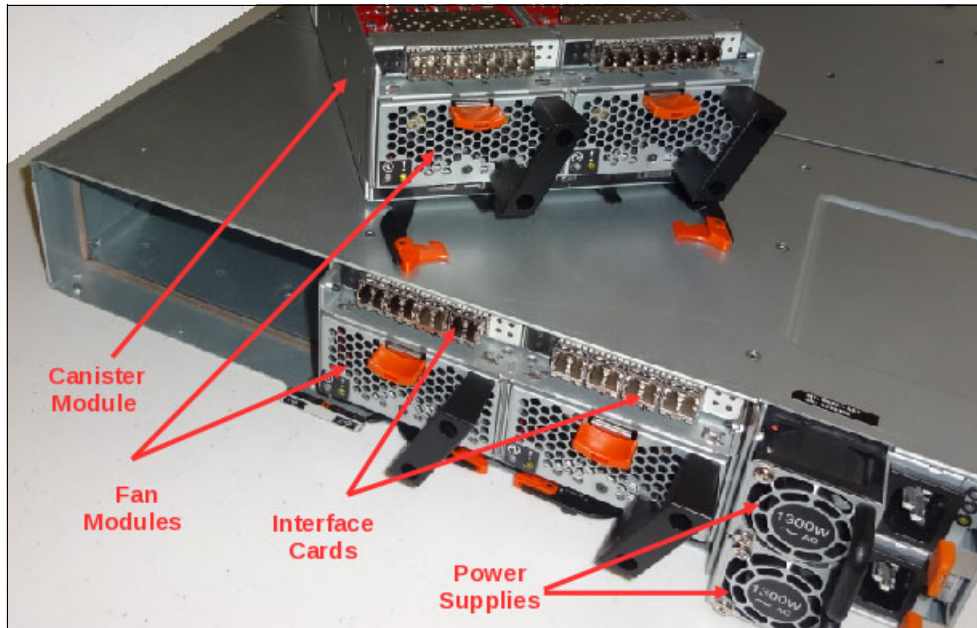


Figure 2-6 Rear view of the FlashSystem 900 with canister removed

To maintain redundancy, the canisters are hot-swappable. If any of the components (except the fans) within a canister fail, the entire canister is replaced as a unit. Both fan modules in each canister are hot-swappable.

**Notes:** If either interface card in a canister fails, the entire canister (minus the fans) must be replaced as an entire unit. When replacing hardware in the IBM FlashSystem 900, follow the *directed maintenance procedure* that is accessible through the GUI. For more details, see “Directed maintenance procedures (DMP)” on page 185.

For more details about the IBM FlashSystem canisters, including canister state LEDs, see the IBM FlashSystem 900 web page in the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## 2.2.4 Interface cards

The IBM FlashSystem 900 supports the following protocol interface cards:

- ▶ Fibre Channel (16 Gbps, 8 Gbps, and 4 Gbps)
- ▶ Fibre Channel over Ethernet (FCoE) (10 Gbps)
- ▶ InfiniBand QDR (40 Gbps)
- ▶ iSCSI (10 Gbps)

**Note:** Although FC and FCoE protocols use the same interface card, only one protocol is supported for each IBM FlashSystem 900 unit.



In the IBM FlashSystem 900, both the FC and FCoE protocols use the same interface card type. The InfiniBand and iSCSI protocols require different module types. Figure 2-7 shows a four-port FC interface card, which is used for 16 Gbps FC (two ports used only), 8 Gbps, 4 Gbps FC (four ports used), and FCoE (four ports used at 10 Gbps each).

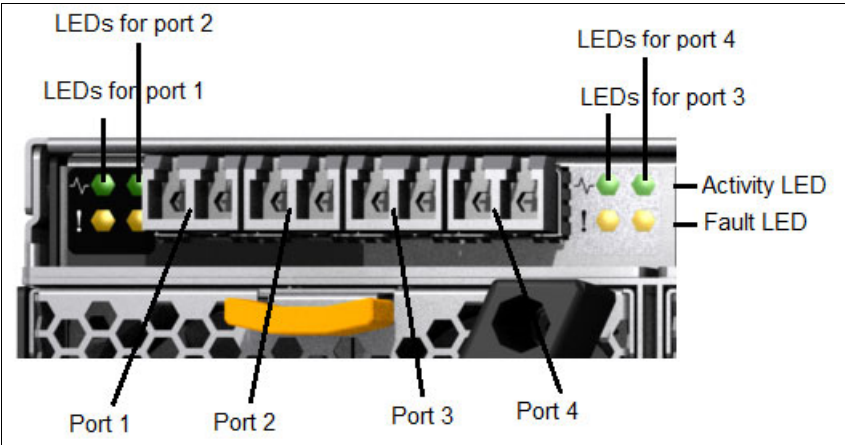


Figure 2-7 IBM FlashSystem 900 FC and FCoE interface card

The two LEDs per port (Figure 2-7) have different status meanings, depending on whether the interface is FC or FCoE:

- ▶ Fibre Channel LED description:
  - A: Link state
  - B: Link speed
- ▶ Fibre Channel over Ethernet LED description:
  - A: Activity
  - B: Link state

### Fibre Channel and FCoE interface card ports and indicators

The FC and FCoE ports on each interface card are numbered 1 - 4, starting from the left.

There are two LED indicators for each FC or FCoE port, or a total of four pairs per interface card.

### Fibre Channel and FCoE port LED descriptions

Each FC or FCoE interface port in the IBM FlashSystem 900 has a set of LEDs to indicate status. Figure 2-7 shows the locations of the port LED.

Table 2-3 lists the LED states and descriptions for FC ports.

Table 2-3 FC LED port descriptions

LED name	Color	States
Link state	Green	<ul style="list-style-type: none"> <li>▶ Off: No small form-factor pluggable (SFP) transceiver installed</li> <li>▶ Slow flash: SFP transceiver installed, no link</li> <li>▶ Solid: Link connected</li> </ul>
Link speed	Amber	<ul style="list-style-type: none"> <li>▶ Off: No link.</li> <li>▶ Two fast flashes: 4 Gb FC connection</li> <li>▶ Three fast flashes: 8 Gb FC connection</li> <li>▶ Four fast flashes: 16 Gb FC connection</li> </ul>



Table 2-4 lists the LED states and descriptions for FCoE ports.

Table 2-4 FCoE LED port descriptions

LED name	Color	States
Activity	Green	<ul style="list-style-type: none"><li>▶ Flashing: Activity on the link</li><li>▶ Off: No activity</li></ul>
Link state	Amber	<ul style="list-style-type: none"><li>▶ Off: No link established</li><li>▶ Solid: Link established</li></ul>

## Support of 16 Gbps Fibre Channel

The IBM FlashSystem 900 supports the new 16 Gbps FC connection speed through the standard FC interface card. The following rules apply to supporting 16 Gbps FC on the IBM FlashSystem 900:

- ▶ If using 16 Gbps FC, only two (of the four) ports on the FC modules can be used. The two leftmost ports (1 and 2) on each interface card are used for 16 Gbps support. The two rightmost ports (3 and 4) are disabled when 16 Gbps is sensed on any port in the IBM FlashSystem 900.
- ▶ If using 16 Gbps FC, the interface is configured as either 16 Gb FC (only two ports active), 8 Gb FC (4 ports active), or 10 Gb FCoE (4 ports). This is configured at the factory and is not changeable by the client. Direct-attach can be supported through point-to-point topology in 16 Gb FC, provided that the host supports direct attach.
- ▶ Four Gbps and 8 Gbps FC connections are supported on the same system connecting to 16 Gbps devices, but there will still only be a total of eight available active ports (ports 1 and 2 on each interface card).

For example, an IBM FlashSystem 900 storage system can have four FC connections at 16 Gbps and four FC connections at 8 Gbps.

- ▶ FC interfaces support Fibre Channel Protocol (FCP) only, with point-to-point (FC-P2P), arbitrated loop (FC-AL), and switched fabric (FC-SW) topologies. FC interfaces can be configured as N\_port or NL\_port types.
- ▶ Sixteen Gbps FC ports do not work in FC-AL mode, and need to be connected to a storage area network (SAN) fabric.
- ▶ Two Gbps FC ports are not supported directly by the IBM FlashSystem 900; a SAN fabric must be used to support these older hosts.

For details about a high-level design for your SAN environment and preferred practices guidance based on IBM 16 Gbps b-type products and features, focusing on FC SAN design details, see *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186.

## InfiniBand interface card

The IBM FlashSystem 900 supports four 2-port InfiniBand 40 Gbps interface cards. A total of eight ports of 40 Gbps InfiniBand connections are supported per IBM FlashSystem 900. Figure 2-8 shows a two-port IBM FlashSystem 900 module.

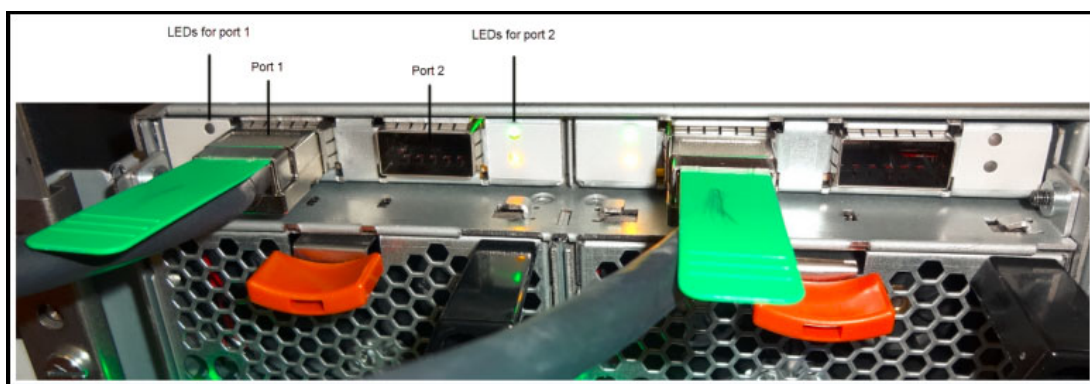


Figure 2-8 Two-port IBM FlashSystem module with InfiniBand

## InfiniBand support

The IBM FlashSystem 900 InfiniBand interface cards have two 4X QDR ports each. The InfiniBand interface card ports are capable of connecting to quad data rate (QDR), Double Data Rate (DDR), or Single Data Rate (SDR) InfiniBand host channel adapters (HCAs) using the SCSI Remote Direct Memory Access (RDMA) Protocol Secure Remote Password (SRP). The IBM FlashSystem 900 InfiniBand interfaces support SCSI RDMA Protocol (SRP) only.

## InfiniBand interface card port LED descriptions

Each InfiniBand interface port in the IBM FlashSystem 900 has a set of LEDs to indicate the status. Table 2-5 lists the InfiniBand LED port descriptions.

Table 2-5 InfiniBand LED port descriptions

LED name	Color	States
Link state	Green	<ul style="list-style-type: none"><li>▶ Off: No link established</li><li>▶ Solid: Link is established</li></ul>
Activity	Amber	<ul style="list-style-type: none"><li>▶ Off: No physical link</li><li>▶ Solid: Link is established, no activity</li><li>▶ Flashing: Activity on the link</li></ul>

## iSCSI interface card

The IBM FlashSystem 900 supports four of the 4-port iSCSI 10 Gbps interface cards. A total of 16 ports of 10 Gbps iSCSI connections are supported per IBM FlashSystem 900.

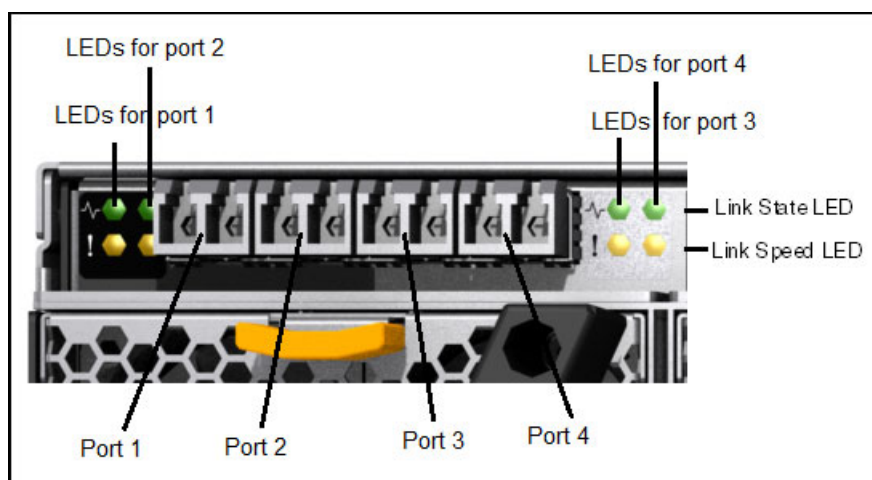


Figure 2-9 IBM FlashSystem 900 iSCSI interface card

## iSCSI support

The IBM FlashSystem 900 iSCSI interface cards have four ports each. The iSCSI interface card ports are capable of connecting to Ethernet switches using 10 Gbps with either optical or copper interconnects. The IBM FlashSystem 900 supports 256 hosts connections using iSCSI. The IBM FlashSystem 900 assigns a T11 Network Address Authority (NAA) name to the system, which, by default, is in this form:

naa.<WWNN>

## iSCSI interface card port LED descriptions

Each iSCSI interface port in the IBM FlashSystem 900 has a set of LEDs to indicate the status. Table 2-6 shows the iSCSI LED port descriptions.

Table 2-6 iSCSI LED port descriptions

LED name	Color	States
Link state	Green	<ul style="list-style-type: none"> <li>▶ Flashing: Activity on the link</li> <li>▶ Off: No activity</li> </ul>
Activity	Amber	<ul style="list-style-type: none"> <li>▶ Off: No link established</li> <li>▶ Solid: Link established</li> </ul>

## 2.2.5 MicroLatency modules

The IBM FlashSystem 900 supports up to 12 IBM MicroLatency modules, accessible from the enclosure front panel. Each IBM MicroLatency module has a usable capacity of either 1.06 TiB (1.2 TB), 2.62 TiB (2.9 TB), or 5.24 TiB (5.7 TB) of flash storage. IBM MicroLatency modules without the daughter board are either half-populated with 1.06 TiB (1.2 TB) or fully populated with 2.62 TiB (2.9 TB). The optional daughter board adds another 2.62 TiB (2.9 TB) for a total of 5.24 TiB (5.7 TB).

Figure 2-10 illustrates an IBM FlashSystem 900 MicroLatency module (base unit and optional daughter board).



Figure 2-10 IBM FlashSystem 900 MicroLatency module

**Note:** All MicroLatency modules in the IBM FlashSystem 900 must be ordered as 1.2 TB, 2.9 TB, or 5.7 TB. IBM MicroLatency modules types cannot be mixed. The daughter board *cannot* be added after deployment.

The maximum storage capacity of the IBM FlashSystem 900 is based on the following factor:

- ▶ In a RAID 5 configuration, one IBM MicroLatency module is reserved as an active spare, and capacity equivalent to one module is used to implement a distributed parity algorithm. Therefore, the maximum usable capacity of a RAID 5 configuration is 57 TB (51.8 TiB; 10 MicroLatency modules x 5.7 TB (5.184 TiB)).

IBM MicroLatency modules are installed in the IBM FlashSystem 900 based on the following configuration guidelines:

- ▶ A minimum of four MicroLatency modules must be installed in the system. RAID 5 is the only supported configuration of the IBM FlashSystem 900.
- ▶ The system supports configurations of 4, 6, 8, 10, and 12 MicroLatency modules in RAID 5.
- ▶ All MicroLatency modules that are installed in the enclosure must be identical in capacity and type.
- ▶ For optimal airflow and cooling, if fewer than 12 MicroLatency modules are installed in the enclosure, populate the module bays beginning in the center of the slots and adding on either side until all 12 slots are populated.

Table 2-7 has suggestions for populating MicroLatency module bays.

Table 2-7 Supported MicroLatency module configurations

No. of installed flash modules <sup>a</sup>	Flash mod. slot 1	Flash mod. slot 2	Flash mod. slot 3	Flash mod. slot 4	Flash mod. slot 5	Flash mod. slot 6	Flash mod. slot 7	Flash mod. slot 8	Flash mod. slot 9	Flash mod. slot 10	Flash mod. slot 11	Flash mod. slot 12
Four					X	X	X	X				
Six				X	X	X	X	X	X			
Eight			X	X	X	X	X	X	X	X		
Ten		X	X	X	X	X	X	X	X	X	X	
Twelve	X	X	X	X	X	X	X	X	X	X	X	X

a. RAID 5 is supported with configurations of 4, 6, 8, 10, and 12 MicroLatency modules.

Consider the following information:

- ▶ If fewer than 12 modules are installed, module blanks must be installed in the empty bays to maintain cooling airflow in the system enclosure.
- ▶ During system setup in the storage enclosure management GUI, the system automatically configures RAID settings based on the number of flash modules in the system. For systems with four or more MicroLatency modules, RAID 5 is used.
- ▶ All MicroLatency modules installed in the enclosure must be identical in capacity and type.

**Important:**

- ▶ MicroLatency modules are hot swappable. However, to replace a module, you must power down the MicroLatency module by using the management GUI *before* you remove and replace the module. This service action does not affect the active LUNs, and I/O to the connected hosts can continue while the MicroLatency module is replaced. Be sure to follow the *directed maintenance procedure* from the IBM FlashSystem 900 GUI before any hardware replacement. See Chapter 6, “Using IBM FlashSystem 900” on page 161.
- ▶ The suggestion is for the storage enclosure to remain powered on, or be powered on periodically, to retain array consistency. The storage enclosure can be safely powered off for up to 90 days, in temperatures up to 40 degrees C. Although the MicroLatency modules will retain data if the enclosure is temporarily disconnected from power, if the system is powered off for a period of time exceeding 90 days, data might be lost.
- ▶ When replacing this part, you must follow the recommended procedures for handling devices that are sensitive to electrostatic discharge (ESD).

## 2.2.6 Battery modules

The IBM FlashSystem 900 contains two hot-swappable battery modules. The function of the battery modules is to ensure that the system is gracefully shut down (write cache fully flushed and synchronized) when AC power is lost to the unit. The IBM FlashSystem 900 battery modules are hot-swappable. Figure 2-11 shows battery module 1, which is in the leftmost front of the IBM FlashSystem 900. An IBM FlashSystem 900 battery module can be hot-swapped without software intervention; however, be sure to follow the *directed maintenance procedure* from the IBM FlashSystem 900 GUI before any hardware replacement.



Figure 2-11 IBM FlashSystem 900 Battery Module 1

### Battery reconditioning

A battery reconditioning feature on the FlashSystem 900 calibrates the gauge to report the amount of charge on the batteries. On systems that have been installed for 10 months or more, or systems that have experienced several power outages, the recommendation to run “battery reconditioning” will appear in the Event Log shortly after upgrading. For more information, see the IBM FlashSystem 900 page in the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## Power supply units

The IBM FlashSystem 900 contains two hot-swappable power supply units. The system can remain fully online if one of the power supply units fails. The IBM FlashSystem 900 power supply units are accessible from the rear of the unit and are fully hot-swappable. Figure 2-12 shows the two IBM FlashSystem 900 hot-swappable power supply units. The IBM FlashSystem 900 GUI and alerting systems (SNMP and so on) will report a power supply fault. The power supply can be hot-swapped without software intervention; however, be sure to follow the *directed maintenance procedure* from the IBM FlashSystem 900 GUI before any hardware replacement.

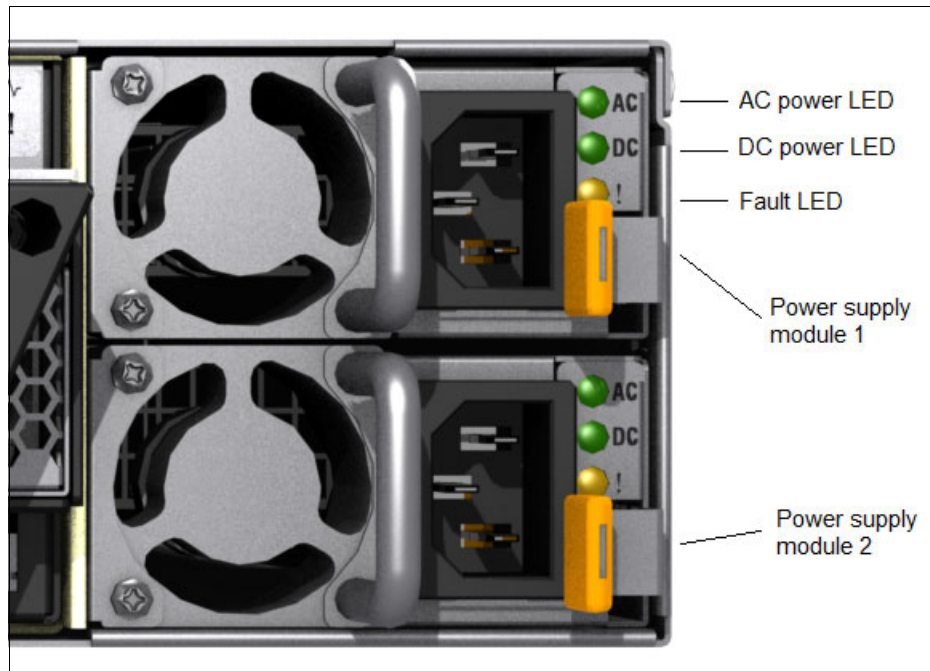


Figure 2-12 IBM FlashSystem 900 hot swappable power supply units

## Fan modules

The IBM FlashSystem 900 contains four hot-swappable fan modules. Each FlashSystem 900 canister holds two hot-swappable fan modules. Each fan module contains two fans. The system can remain fully online if one of the fan modules fails. The IBM FlashSystem 900 fan modules are accessible from the rear of the unit (in each canister) and are fully hot-swappable.



Figure 2-13 shows an IBM FlashSystem 900 hot-swappable fan module. The IBM FlashSystem 900 GUI and alerting systems (SNMP and so on) will report a fan module fault. The fan module can be hot-swapped without software intervention; however, be sure to follow the *directed maintenance procedure* from the IBM FlashSystem 900 GUI before any hardware replacement.

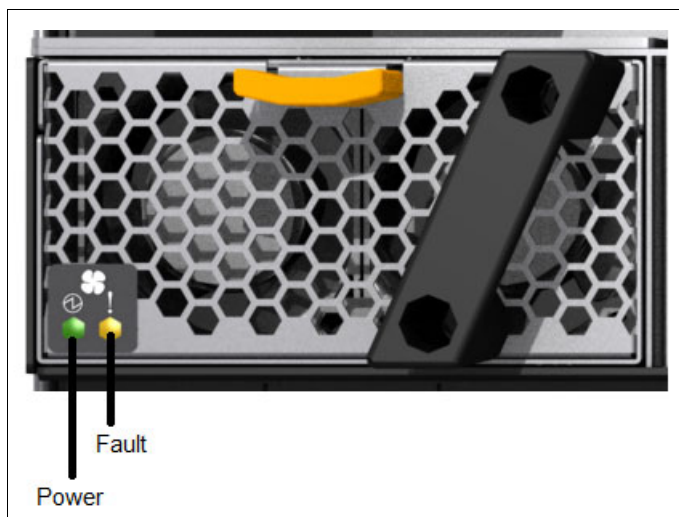


Figure 2-13 IBM FlashSystem 900 fan module

## 2.3 Administration and maintenance

The IBM FlashSystem 900 storage system capabilities for administration, maintenance, and serviceability are described in this section.

### 2.3.1 Serviceability and software enhancements

IBM FlashSystem 900 includes several design enhancements for the administration, management, connectivity, and serviceability of the system:

- ▶ Concurrent code load: The IBM FlashSystem 900 supports the ability to upgrade the system firmware on the canister (RAID controllers, management modules, and interface cards) and flash modules without affecting the connected hosts or their applications.
- ▶ Easily accessible hot-swappable modules with no single point of failure: The IBM FlashSystem 900 design enables the easy replacement of any hardware module through the front or rear of the unit. The IBM FlashSystem 900 does not require the top panel to be removed and does not need to be moved in the rack to replace any component.
- ▶ Standard IBM SAN Volume Controller CLI and GUI: The IBM FlashSystem 900 uses the latest SAN Volume Controller CLI and GUI for simple and familiar management of the unit.
- ▶ Encryption support: The IBM FlashSystem 900 supports hardware encryption of the flash modules to meet the audit requirements of enterprise, financial, and government clients.
- ▶ Sixteen Gbps FC and FCoE support: The IBM FlashSystem 900 supports 16 Gbps FC and 10 Gbps FCoE, enabling clients to take advantage of the latest high-speed networking equipment while increasing performance.



## 2.3.2 System management

IBM FlashSystem 900 includes the use of the common IBM SAN Volume Controller CLI and the popular IBM SAN Volume Controller GUI, which is based on the user-friendly IBM XIV GUI. The IBM FlashSystem 900 supports SNMP, email forwarding (SMTP), and syslog redirection for complete enterprise management access.

### USB key initialization process

IBM FlashSystem 900 uses a new Universal Serial Bus (USB) key initialization process, which is similar to the IBM V7000 disk systems initialization. A USB key has an initialization file on it and is placed in a Microsoft Windows or Linux workstation to program the initial IP address information into a utility. The USB key is then placed into the IBM FlashSystem 900 on the first start-up and the initialization file is read and applied. The IBM FlashSystem 900 can then be managed through the GUI and CLI. For more information about the USB key initialization process, see Chapter 4, “Installation and configuration” on page 67.

### Graphical user interface (GUI)

IBM FlashSystem 900 includes the use of the standard IBM SAN Volume Controller GUI. This GUI is simple to use and based on the popular IBM XIV GUI.

The IBM FlashSystem 900 GUI is launched from a supported Internet browser when you enter the systems management IP address. Then, the login window opens (Figure 2-14) where you enter a valid user name and password.



Figure 2-14 IBM FlashSystem 900 GUI login window

The system overview window opens (Figure 2-15). In the middle of the window is a real-time graphic depicting the IBM FlashSystem 900.

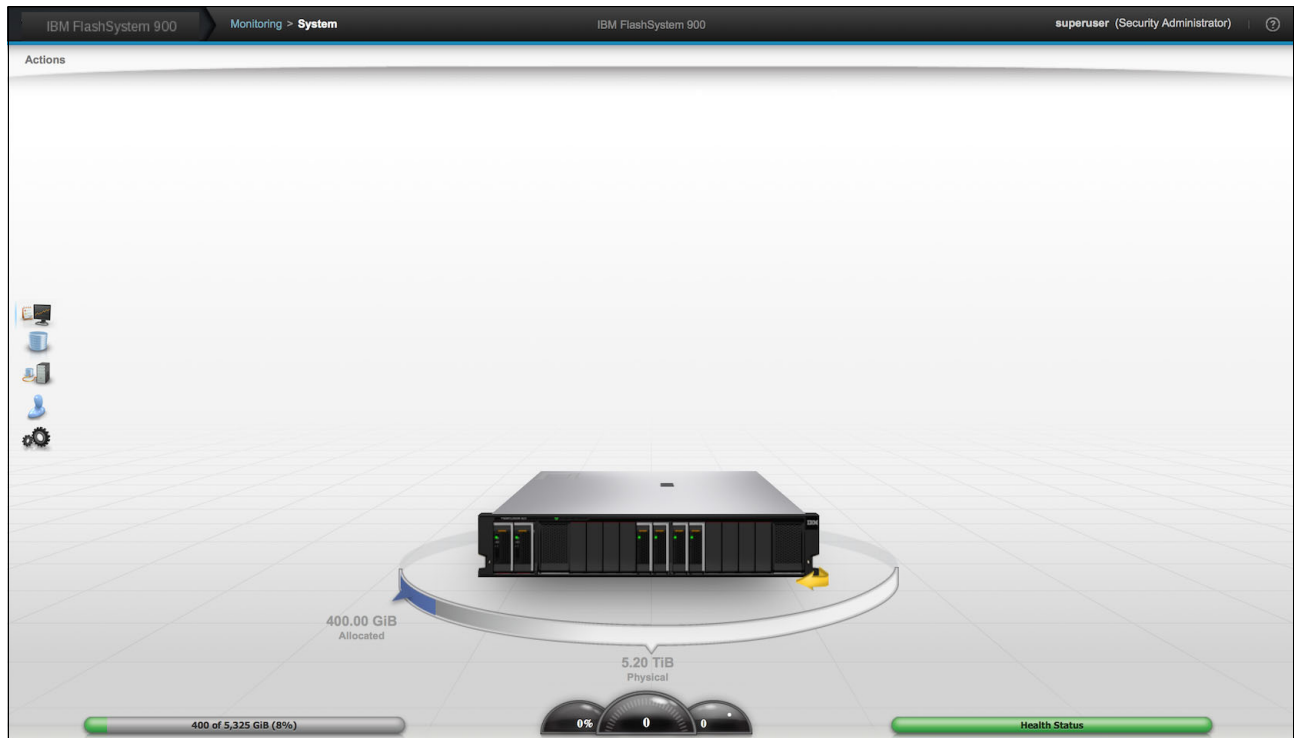


Figure 2-15 System overview window

At the bottom of the window are three dashboard icons:

- ▶ Capacity
- ▶ Performance
- ▶ System status

At the left of the window are five function icons:

- ▶ Monitoring function
- ▶ Volumes function
- ▶ Hosts function
- ▶ Access function
- ▶ Settings function

Those icons are briefly described next. Also see these chapters:

- ▶ Details about the GUI are in Chapter 6, “Using IBM FlashSystem 900” on page 161.
- ▶ Details about the Settings element are in Chapter 7, “Configuring settings” on page 231.

### **Monitoring function**

Figure 2-16 on page 39 shows the Monitoring icon and the associated branch-out menu. Click the **Monitoring** icon if you want to do any of these actions:

- ▶ System: Monitor the system health of the IBM FlashSystem 900 hardware.
- ▶ Events: View the events log of the IBM FlashSystem 900.
- ▶ Performance: Launch the system I/O performance graphs.



Figure 2-16 IBM FlashSystem 900 GUI: Monitoring icon and branch-out menu

### **Volumes function**

Figure 2-17 shows the Volumes icon and the associated branch-out menu. Click the **Volumes** icon if you want to do any of these actions:

- ▶ Volumes: View a list of all system storage volumes (LUNs), create new volumes, edit existing volumes, and delete volumes.
- ▶ Volumes by Host: View a list of volumes that are associated with hosts, create new associations, or delete associations.



Figure 2-17 IBM FlashSystem 900 GUI: Volumes icon and branch-out menu

### **Hosts function**

Figure 2-18 on page 40 shows the Hosts icon and the associated branch-out menu. Click the **Hosts** icon if you want to do any of these actions:

- ▶ Hosts: View a list of all hosts, create new hosts, edit existing hosts, and delete hosts.
- ▶ Volumes by Host: View a list of volumes that are associated with hosts, create new associations, or delete associations.

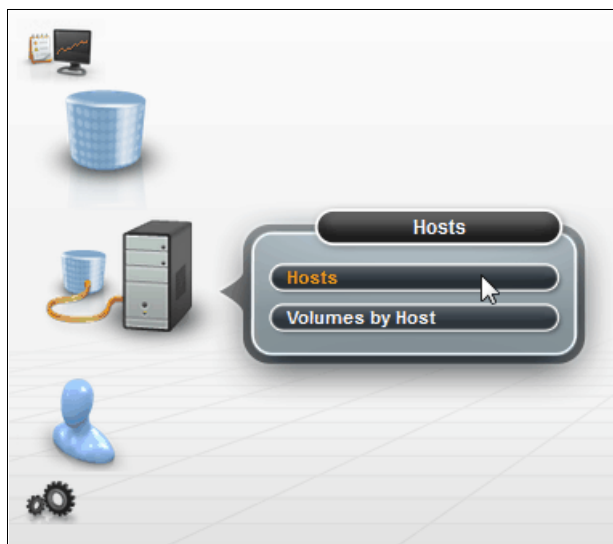


Figure 2-18 IBM FlashSystem 900 GUI: Hosts icon and branch-out menu

### **Access function**

Figure 2-19 shows the Access icon and associated branch-out menu. Click the **Access** icon if you want to do any of these actions:

- ▶ Users: View a list of current users, create new users, edit existing users, and delete users.
- ▶ User Groups: Create user groups (based on access rights) and associate users with groups.
- ▶ Audit Log: View the system access log and view actions by individual users.



Figure 2-19 IBM FlashSystem 900 GUI: Access icon and branch-out menu

## Settings function

Figure 2-20 shows the Settings icon and associated branch-out menu. Click the **Settings** icon if you want to configure system parameters, including alerting, open access, GUI settings, and other system-wide configurations. For more information, see Chapter 7, “Configuring settings” on page 231.



Figure 2-20 IBM FlashSystem 900 GUI: Settings icon and branch-out menu

## Command-line interface (CLI)

IBM FlashSystem 900 uses the standard IBM SAN Volume Controller storage CLI. This CLI is common among several IBM storage products, including the IBM SAN Volume Controller and the IBM Storwize family of products: the V7000, IBM V5000, IBM V3700, and IBM V3500 disk systems. IBM SAN Volume Controller CLI is easy to use with built-in help and hint menus.

To access the IBM FlashSystem 900 SAN Volume Controller CLI, a Secure Shell (SSH) session to the management IP address must be established (Telnet is not enabled on the IBM FlashSystem 900). The client is then prompted for a user name and password.

## Call home email SMTP support

IBM FlashSystem 900 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. These emails can also be sent within the client’s enterprise to other email accounts that are specified. After it is set up, system events that might require service are emailed automatically to an IBM Service account automatically specified in the IBM FlashSystem 900 code. The email alerting can be set up as part of the system initialization process or added or edited at anytime through the IBM FlashSystem 900 GUI. Also, a test email can be generated at anytime to test the connections.

Figure 2-21 shows the IBM FlashSystem 900 email setup window.

Figure 2-21 IBM FlashSystem 900 email alerting setup window

## SNMP support

IBM FlashSystem 900 supports SNMP versions 1 and 2. The GUI is used to set up SNMP support on the IBM FlashSystem 900.

To set up SNMP support on the IBM FlashSystem 900, click the **Settings** icon at the left side of the window, click the **Notifications** tab and click the **SNMP** tab to enter the SNMP trap receiver IP address and community access information. Figure 2-22 shows the IBM FlashSystem 900 SNMP setup window.

Figure 2-22 IBM FlashSystem 900 SNMP setup window

**Note:** The IBM FlashSystem 900 CLI, based on technology from SAN Volume Controller can also be used to program the SNMP settings.

## Redirection of syslog

You can redirect syslog messages to another host for system monitoring. Use the GUI to set up syslog redirection on the IBM FlashSystem 900.

To set up syslog redirection, click the **Settings** icon on the lower left of the window, click the **Event Notifications** tab, and then click the **Syslog** tab to enter the remote host trap IP address and directory information. Figure 2-23 shows the Syslog redirection setup window.



Figure 2-23 IBM FlashSystem 900 Syslog redirection setup window

**CLI:** The IBM FlashSystem 900 CLI can also be used to set up syslog redirection.

## 2.4 Support matrix

IBM FlashSystem 900 supports a wide range of operating systems (Windows Server 2008 and 2012, Linux, IBM AIX® and VMware vSphere/ESXi), hardware platforms (IBM System x, IBM Power Systems, and x86 servers not from IBM), host bus adapters (HBAs), and SAN fabrics.

For specific information, see the IBM System Storage Interoperation Center (SSIC):

<http://ibm.com/systems/support/storage/ssic>

Also, consider using the IBM SAN Volume Controller as a front-end host-facing interface for the IBM FlashSystem 900. Therefore, if the IBM FlashSystem 900 is used with the IBM SAN Volume Controller, the host interoperability matrix for the IBM SAN Volume Controller is relevant. Obtain the IBM SAN Volume Controller interoperability information, also from the SSIC web page:

<http://ibm.com/systems/support/storage/ssic>

Contact your IBM sales representative or IBM Business Partner for assistance or questions about the IBM FlashSystem 900 or IBM SAN Volume Controller interoperability.

## 2.5 Product integration overview

IBM FlashSystem 900 is an all-flash storage system that can enhance the performance of almost any application. A high-level overview of how the IBM FlashSystem 900 works with a list of IBM products and applications is described. In addition to the products described here, the IBM FlashSystem 900 also works with a wide variety of other IBM software applications and hardware products, and also products from third-party vendors. For more detail, consult with your IBM salesperson or IBM Business Partner for advice about incorporating the IBM FlashSystem 900 into any of these or other applications.

For more detail about deploying the IBM FlashSystem 900 with the IBM products described here, see Chapter 8, “Product integration” on page 279.

For more information about using the IBM FlashSystem with other IBM and third-party solutions and applications, see this website:

<http://www.ibm.com/storage/flash>

## 2.5.1 IBM Spectrum Virtualize - SAN Volume Controller

The IBM FlashSystem 900 all-flash storage array provides clients with storage that is fast and has low latency.

IBM SAN Volume Controller delivers the functions of IBM Spectrum Virtualize, part of the IBM Spectrum Storage family, and has been improving infrastructure flexibility and data economics for more than 10 years. Its innovative data virtualization capabilities provide the foundation for the entire IBM Storwize family. SVC provides the latest storage technologies for unlocking the business value of stored data, including virtualization and Real-time Compression. IBM SAN Volume Controller enriches any storage environment by adding storage management functions and a wide variety of features:

- ▶ IBM FlashCopy® point-in-time copies
- ▶ Local and remote mirroring
- ▶ Thin provisioning
- ▶ Real-time Compression
- ▶ EasyTier support (automatically directs “hot I/O” to fastest storage)
- ▶ Support for virtual environment APIs
- ▶ Support for OpenStack Cloud environments
- ▶ Support for IBM Storage Integration Server
- ▶ Storage consolidation

IBM SAN Volume Controller provides these features with minimal delay or latency in the I/O path. The combination of the IBM FlashSystem 900 and IBM SAN Volume Controller enables clients to take advantage of the speed of the IBM FlashSystem 900 and the robust storage management capabilities of the IBM SAN Volume Controller.

**Note:** If you want the advanced software features and low latency of the IBM FlashSystem 900 combined with SAN Volume Controller functions and services, such as mirroring, IBM FlashCopy, thin provisioning, IBM Real-time Compression Copy Services, and broader host support, you may purchase IBM FlashSystem V9000. For product details, see the IBM Redbooks Product Guide, *IBM FlashSystem V840 Enterprise Performance Solution*, TIPS1158:

<http://www.redbooks.ibm.com/abstracts/tips1158.html>

### Additional information

See the following resources for more information:

- ▶ Information about FlashSystem 900 and the IBM FlashSystem resources (including Redbooks publications, white papers, demos and other information about this topic):  
<http://www.ibm.com/storage/flash>
- ▶ Information about FlashSystem running in an IBM Spectrum Virtualize environment is in *Implementing FlashSystem 840 with SAN Volume Controller*, TIPS1137.



- ▶ Information about the IBM SAN Volume Controller:
  - *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
  - *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933
- ▶ An in-depth description of combining the IBM FlashSystem 900 and IBM Spectrum Virtualize - SAN Volume Controller is in 8.1, “Running FlashSystem 900 with IBM Spectrum Virtualize - SAN Volume Controller” on page 280.

## 2.5.2 IBM Storwize V7000 storage array

Similar to the IBM Spectrum Virtualize - SAN Volume Controller product, the IBM Storwize V7000 storage array can provide storage management services (FlashCopy, thin-provisioning, mirroring, replication, Real-time Compression, support for virtual environments, and Easy Tier support) to externally connected storage systems.

The IBM FlashSystem 900 can be connected to the IBM V7000 storage array and provide high performance and low latency to connected hosts, while taking advantage of the IBM V7000 storage management services.

The IBM Storwize V7000 FlashSystem Edition and FlashSystem Enterprise Addition enable you to accelerate your mid-range storage solution by taking advantage of the extreme performance and low latency of the FlashSystem. For more details, see the following website:

<http://www.ibm.com/systems/storage/flash/>

### Additional information

See the following resources for more information:

- ▶ To deploy the IBM V7000 and IBM FlashSystem 900 together, contact your IBM sales representative or IBM Business Partner.
- ▶ Information about the IBM V7000 and IBM FlashSystem:
 

<http://www.ibm.com/storage/flash>
- ▶ Information about the IBM Storwize V7000:
  - *IBM Storwize V7000 and SANSlide Implementation*, REDP-5023
  - *IBM Flex System V7000 Storage Node Introduction and Implementation Guide*, SG24-8068
  - *Implementing the IBM Storwize V7000 Gen2*, SG24-8244
  - *IBM Flex System V7000 Storage Node*, TIPS1039
- ▶ An in-depth description of combining the IBM FlashSystem 900 and the IBM V7000 is in 8.3, “Integration considerations: FlashSystem 900 and SAN Volume Controller” on page 308.

### 2.5.3 IBM DB2 database environments

IBM FlashSystem 900 enables clients to speed up their databases dramatically. IBM DB2® is a high-performance, enterprise-scale database that is used by several of the largest IBM clients worldwide. Moving some or all of an IBM DB2 database onto the IBM FlashSystem 900 accelerates performance and increases CPU utilization at the same time. Also, moving a small portion of an IBM DB2 database onto the IBM FlashSystem 900 can have dramatic results.

#### Additional information

To use the IBM FlashSystem 900 and IBM DB2 products together:

- ▶ Contact your IBM sales representative or IBM Business Partner.
- ▶ Obtain information at this website:  
<http://www.ibm.com/storage/flash>
- ▶ *Faster DB2 Performance with IBM FlashSystem*, TIPS1041

### 2.5.4 IBM Spectrum Scale

IBM Spectrum Scale is a proven, scalable, high-performance data and file management solution, based upon IBM General Parallel File System or GPFS. IBM Spectrum Scale technology is a high-performance enterprise file management platform, and it can help you move beyond simply adding storage to optimize data management. With IBM Spectrum Scale, businesses can achieve higher performance while reducing the footprint by placing *elastic* storage metadata on the FlashSystem 900. Running IBM Spectrum Scale *virtualizes* IBM FlashSystem for file-based access in much the same way that SAN Volume Controller and the V7000 virtualize it for block-based access.

IBM Spectrum Scale is also a key component for many big data products. IBM BigInsights®, DB2 PureScale, and SAP HANA all use IBM Spectrum Scale. In addition, IBM Spectrum Scale is used for many high-performance storage applications, such as media, life sciences, high-performance computing (HPC), and so on. Some applications need fast I/O for data and some applications need fast I/O for metadata. There is a mix of random and streaming I/O for data, depending on the type of workload.

IBM Spectrum Scale integrates with the IBM FlashSystem 900 and offers your business environment the following potential benefits:

- ▶ IBM Spectrum Scale enables the IBM FlashSystem to be used as a storage *tier* under user control, scheduled control, or dynamically, where files are moved to and from flash and disk (and even tape) under policy control or when they are used.
- ▶ You can use the IBM FlashSystem to support either data, metadata, or both, and it can support millions of file creations and deletions per minute.
- ▶ A small part of the IBM FlashSystem can be partitioned for metadata use and the remainder can be used for hot data.
- ▶ This integration provides the capability to replicate data sync or async to another site, and it supports full active-active two-site configurations (sync replication only).
- ▶ IBM Spectrum Scale can perform mirroring of two FlashSystem 900s, either at a single site, or across multiple sites.
- ▶ IBM Spectrum Scale is used at many sites where InfiniBand is used. IBM Spectrum Scale and InfiniBand are a good match for the IBM FlashSystem 900.

In this scenario, IBM Spectrum Scale and DataDirect Networks (DDN) communicate using the same SCSI Remote Direct Memory Access (RDMA) Protocol (SRP) supported by the IBM FlashSystem 900. SRP is similar to iSCSI protocol. SRP provides access to LUNs across networks, and SRP protocol is used across InfiniBand networks. GPFS was historically implemented using SRP protocols to access DDN disk subsystems from GPFS Network Shared Disks (NSD) servers.

## 2.5.5 IBM TS7650G ProtecTIER

IBM TS7650G ProtecTIER® is an enterprise class deduplication system that is used by IBM clients to make their backup environments more efficient. IBM TS7650G ProtecTIER uses IBM System x servers, IBM TS7650G ProtecTIER software, and a Fibre Channel-attached storage array to create a deduplication storage unit. This deduplication storage unit can be used by backup applications, such as IBM Spectrum Protect™ (formerly IBM Tivoli Storage Manager), to back up and restore data. IBM TS7650G ProtecTIER also provides disaster recovery by replicating backup data to another location for safekeeping. The IBM TS7650G ProtecTIER back-end storage array is configured with two types of LUNs:

- ▶ Metadata LUNs are used to record where data is kept.
- ▶ User data LUNs are used to store the actual data.

Metadata LUN performance is critical and 15 K RPM spindles in a RAID 10 4+4 configuration are commonly used.

IBM FlashSystem 840 can be used by IBM TS7650G ProtecTIER as the back-end storage device for metadata and user data LUNs. A common use for the IBM FlashSystem 840 is for IBM TS7650G ProtecTIER metadata LUNs. Compared to the cost of dozens of 15 Kb or 10 Kb HDD spindles, a more cost-effective way might be to use the IBM FlashSystem 840 for IBM TS7650G ProtecTIER metadata LUNs. Also more cost-effective might be to use the IBM FlashSystem 840 for the entire IBM TS7650G ProtecTIER repository if high performance, but small capacity is needed.

### Notes:

- ▶ Consult your IBM sales representative or IBM Business Partner for more information about deploying the IBM FlashSystem 900 with IBM TS7650G ProtecTIER.
- ▶ If the configuration that you want is not available for selection on the IBM System Storage Interoperation Center (SSIC) web page, request approval by submitting a Solution for Compliance in a Regulated Environment (SCORE) request to IBM. A similar process is the request for price quotation (RPQ).

### Additional information

See the following resources for more information:

- ▶ More details about IBM TS7650G ProtecTIER running with IBM FlashSystem storage:
  - *Implementing ProtecTIER 3.3 and IBM FlashSystem*, TIPS1140
  - *IBM ProtecTIER Implementation and Best Practices Guide*, SG24-8025
- ▶ Information about IBM TS7650G ProtecTIER:  
<http://www.ibm.com/systems/storage/tape/ts7650g/>

**Notes:** IBM TS7650G ProtecTIER only supports the IBM FlashSystem in RAID 5 mode. If FlashSystem is used for the IBM TS7650G ProtecTIER metadata LUNs, all future expansions must be with the IBM FlashSystem storage.

## 2.5.6 VMware with IBM Storage Integration Server

IBM FlashSystem 900 supports VMware vSphere Storage APIs - Array Integration (VAAI) and VMware vSphere API for Storage Awareness (VASA) through integration with the IBM Storage Integration Server. IBM FlashSystem 900 support for VAAI includes using the Block Reclaim/UNMAP primitive to inform the FlashSystem 900 to reclaim storage. IBM FlashSystem 900 integrates with the IBM Storage Integration Server, to provide a consolidated method for system monitoring, automation, and provisioning.

For more details about IBM Storage Integration Server running with the FlashSystem, see *Deploying IBM FlashSystem V840 Storage in a VMware and Cloud Environment*, REDP-5148.



# Planning

This chapter provides planning information and general considerations for you to review before you install the IBM FlashSystem 900 (FlashSystem 900). This information includes connectivity, supported host environments, and IP addresses.

This chapter covers the following topics:

- ▶ Prerequisites to installation
- ▶ Planning cable connections
- ▶ Planning for power
- ▶ Planning for configuration
- ▶ The call home option
- ▶ TCP/IP requirements
- ▶ Planning for encryption
- ▶ Checking web browser settings for the management GUI
- ▶ Licensing
- ▶ Supported hosts and operating system considerations

## 3.1 Prerequisites to installation

Understand the contact information and checklists before you start. Plan to provide the required network infrastructure and the storage network infrastructure as described in the following sections.

### 3.1.1 Contact information and checklist

Before installing the system, collect the necessary information that will be used during the installation and setup of the system.

#### Contact information for the call home feature

The FlashSystem 900 is customer installable. Enabling the *call home* feature allows IBM personnel to be notified of any critical hardware problems.

**Important:** A strong suggestion is to enable the call home feature so that IBM personnel can be notified 24x7 about any critical hardware problems. Not enabling the call home feature can result in long delays for any necessary service actions.

When the IBM Support Center receives a call home report, an IBM service representative contacts your company to work on resolving the problem. You need basic information, such as an email address and a Simple Mail Transfer Protocol (SMTP) gateway address to set up call home. You can complete the necessary information in Table 3-13 on page 62. Also see 3.5, “The call home option” on page 61 to set up call home before you install the system.

#### Checklist before you start

Review this checklist to be sure that you have the latest information for planning the installation:

- The initialization tool is available on the USB flash drive that is shipped with the system. The name of the application file is `InitTool.exe`. If you cannot locate the USB flash drive or you want to use the latest code, you can download the application from the Fix Central support website (search for `initialization tool`):

<http://www.ibm.com/support/fixcentral>

The `InitTool` uses features of Microsoft Internet Explorer (IE). Make sure to use the latest IE version if you encounter problems using the `InitTool`.

- If the FlashSystem 900 will be used with IBM Spectrum Virtualize, which is based on technology from SAN Volume Controller, the SAN Volume Controller version must support the FlashSystem 900. Check the SAN Volume Controller supported hardware list for the SAN Volume Controller version that you want to use:

- The IBM SAN Volume Controller web page:

<http://www.ibm.com/systems/storage/software/virtualization/svc/>

- SAN Volume Controller 7.4 supports FlashSystem 900. See *V7.4.x Supported Hardware List, Device Driver, Firmware and Recommended Software Levels for SAN Volume Controller*.

<http://www.ibm.com/support/docview.wss?uid=ssg1S1004946>

- References to all SAN Volume Controller versions:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1003658>

- ▶ The FlashSystem 900 will use 2U of rack space. Check for sufficient free space in the rack and for round or square holes in the rack to attach the enclosure rails.
- ▶ Always check the IBM System Storage Interoperation Center (SSIC) to get the latest information about supported operating systems, hosts, switches, and so on:  
<http://www.ibm.com/systems/support/storage/ssic/interoperability.wss>
- ▶ If the host uses clustering software, for example, Microsoft Cluster Server (MSCS) or Oracle Automatic Storage Management (ASM), make sure that it is supported with the FlashSystem 900.
- ▶ Set up a risk mitigation plan that lists all unresolved action items.
- ▶ Connect the Ethernet ports to a switch with offline alert notification enabled.
- ▶ Collect all necessary information, for example IP addresses, port information, user and email names.

### 3.1.2 Completing the hardware location chart

Planning for the physical location includes documenting the rack locations of the IBM FlashSystem 900 enclosure and other devices, based on the requirements of each device.

The hardware location chart in Table 3-1 represents the rack into which the enclosure is installed. Each row of the chart represents one Electronic Industries Alliance (EIA) 19-inch wide by 1.75-inch tall rack space or unit, each of which is commonly referred to as *1U of the rack*. As you design your rack, use the hardware location chart to record the physical configuration of the 2U enclosure and other devices in your system.

**Note:** Install the enclosure where it can be easily serviced, but ensure that the rack is kept stable, for example, by installing enclosures beginning from the bottom.

Complete Table 3-1 for the hardware locations of the FlashSystem 900 enclosure and other devices.

Table 3-1 Hardware location of the FlashSystem 900 enclosure and other devices

Rack unit	Component
EIA 36	
EIA 35	
EIA 34	
EIA 33	
EIA 32	
EIA 31	
EIA 30	
EIA 29	
EIA 28	
EIA 27	
EIA 26	
EIA 25	

Rack unit	Component
EIA 24	
EIA 23	
EIA 22	
EIA 21	
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	

## 3.2 Planning cable connections

This section provides the steps to plan for setting up cable connections for the enclosure. You can use the tables to record the details of the cable connections for your system.

### 3.2.1 Management port connections

Each of the two canisters that are in the FlashSystem 900 includes an Ethernet port for accessing the management GUI.

Use Table 3-2 on page 53 to record the management port connection information.



Table 3-2 Enclosure management port Ethernet connections

Canister	Management port	
<b>Canister 1 Ethernet Management Port</b>	Switch:	
	Port:	
	Speed:	
<b>Canister 2 Ethernet Management Port</b>	Switch:	
	Port:	
	Speed:	

**Important:** Three IP addresses are required for managing the storage enclosure: a cluster IP address and two service IP addresses. Each of the three IP addresses must be a unique value.

The left canister, when viewed from the back, is the first node. If the left canister is not available at system start-up, the right canister defaults to be the first node.

You use the USB-Key **InitTool** to set up the cluster IP address. You must set the two service IP addresses after the installation of the system. Use Table 3-3 to record the cluster and service IP address settings for the storage enclosure.

Table 3-3 IP addresses for the storage enclosure

<b>Cluster name:</b>	
<b>Cluster IP address:</b>	
IP:	
Subnet mask:	
Gateway:	
<b>Service IP address 1:</b>	
IP:	
Subnet mask:	
Gateway:	
<b>Service IP address 2:</b>	
IP:	
Subnet mask:	
Gateway:	

## 3.2.2 Interface card connections

This section describes the information needed for various interface card connections and for the connection switches.

**Important:** Each host must be connected to both canisters to add redundancy and improve performance. If a host is not connected to both canisters, its state shows as *degraded*.

### Fibre Channel

Each canister supports two optional Fibre Channel (FC) interface cards:

- ▶ 8 Gb FC cards support four ports.
- ▶ 16 Gb FC cards support the two leftmost ports on the card.

**Important:** Small form-factor pluggable (SFP) transceivers must be obtained and deployed in pairs to support multipathing.

Use Table 3-4 to record the FC port connection information.

Table 3-4 Fibre Channel 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)
Canister 1 Fibre Channel card 1 (left)	Switch or host:				
	Port:				
	Speed:				
Canister 1 Fibre Channel card 1 (right)	Switch or host:				
	Port:				
	Speed:				
Canister 2 Fibre Channel card 2 (left)	Switch or host:				
	Port:				
	Speed:				
Canister 2 Fibre Channel card 2 (right)	Switch or host:				
	Port:				
	Speed:				

## Fibre Channel over Ethernet

Each canister supports two optional Fibre Channel over Ethernet (FCoE) interface cards. There are four ports on each card.

### Notes:

- An FCoE switch is required when connecting the FCoE adapter ports. Direct attach to hosts is not supported.
- SFP transceivers must be obtained and deployed in pairs to support multipathing.

Use Table 3-5 to record the FCoE port connection information.

Table 3-5 FCoE port connections

Location	Item	FCoE port 1	FCoE port 2	FCoE port 3	FCoE port 4
Canister 1 FCoE card 1 (left)	Switch or host:				
	Port:				
	Speed:				
Canister 1 FCoE card 1 (right)	Switch or host:				
	Port:				
	Speed:				
Canister 2 FCoE card 2 (left)	Switch or host:				
	Port:				
	Speed:				
Canister 2 FCoE card 2 (right)	Switch or host:				
	Port:				
	Speed:				

## Quad data rate InfiniBand

Each canister supports two optional InfiniBand interface cards. Two ports are on each card.

Use Table 3-6 to complete the InfiniBand port connection information.

Table 3-6 *InfiniBand port connections*

Location	Fibre Channel port 1	InfiniBand port 1	InfiniBand port 2
Canister 1 InfiniBand card 1 (left)	Switch or hosts:		
	Port:		
	Speed:		
Canister 1 InfiniBand card 1 (right)	Switch or hosts:		
	Port:		
	Speed:		
Canister 2 InfiniBand card 2 (left)	Switch or hosts:		
	Port:		
	Speed:		
Canister 2 InfiniBand card 2 (right)	Switch or hosts:		
	Port:		
	Speed:		

## Internet Small Computer System Interface (iSCSI)

iSCSI is an IP-based standard for transferring data that supports host access by carrying SCSI commands over IP networks. The iSCSI standard is defined by Request for Comments memorandum (RFC) 3720. For the FlashSystem 900, connections from iSCSI-attached hosts to nodes are supported. The FlashSystem 900 supports up to sixteen 10 Gb Ethernet ports, using small form-factor pluggable transceiver (SFPs) with optical cabling or direct attach copper (DAC) cabling. iSCSI uses iSCSI qualified name (IQN), extended unique identifier (EUI), or T10 Network Address Authority (NAA) identifiers.

Table 3-7 shows that iSCSI and Fibre Channel terms have analogous components.

*Table 3-7 Comparison of iSCSI and Fibre Channel components*

iSCSI components	Fibre Channel components
iSCSI host bus adapter	Fibre Channel host bus adapter
Network interface controller (NIC) and iSCSI software initiator	Fibre Channel host bus adapter
IP switch	Fibre Channel switch
IP router	N/A
iSCSI name, such as iSCSI qualified name (IQN), extended unique identifier (EUI), or T11 Network Address Authority (NAA) identifiers	Worldwide node name (WWNN)

### ***iSCSI initiators and targets***

In an iSCSI configuration, the iSCSI host or server sends requests to a node. The host contains one or more initiators that attach to an IP network to initiate requests to send and receive responses from an iSCSI target, for example, the FlashSystem 900. Each initiator and target are given a unique iSCSI name, such as an IQN, EUI, or NAA identifier. An IQN is a 223-byte ASCII name. An EUI is a 64-bit identifier. A NAA is a 64 or 128-bit identifier. An iSCSI name represents a worldwide unique naming scheme that is used to identify each initiator or target in the same way that worldwide node names (WWNNs) are used to identify devices in a Fibre Channel fabric.

iSCSI targets are devices that respond to iSCSI commands. An iSCSI device can be an end node, such as a storage device, or it can be an intermediate device, such as a bridge between IP and Fibre Channel devices. Each iSCSI target is identified by a unique iSCSI name.

Use Table 3-8 to complete the iSCSI port connection information.

Table 3-8 iSCSI port connections

Location	Item	iSCSI port 1	iSCSI port 2	iSCSI port 3	iSCSI port 4
Canister 1 iSCSI card 1 (left)	Switch or host: Port:				
	IP address:				
	Subnet:				
	Gateway:				
Canister 1 iSCSI card 1 (right)	Switch or host: Port:				
	IP address:				
	Subnet:				
	Gateway:				
Canister 2 iSCSI card 1 (left)	Switch or host: Port:				
	IP address:				
	Subnet:				
	Gateway:				
Canister 2 iSCSI card 2 (right)	Switch or host: Port:				
	IP address:				
	Subnet:				
	Gateway:				

### Switch information

Each of the two canisters includes an Ethernet port that is attached to an Ethernet switch. These two ports are used for the management of the system.

Use Table 3-9 to complete the switch information.

Table 3-9 Switch information

	IPv4 address	IPv6 address	Media Access Control (MAC) address	Physical location
Switch 1				
Switch 2				
Switch 3				
Switch 4				

### 3.3 Planning for power

This section describes necessary information for planning to attach each of the two power supplies in the enclosure to separate main power supply lines.

Plan to connect the power cords on the right side of the rack (when viewed from the rear) to power sources that provide power in the 100 - 127V / 200 - 240V ac range at 10.0/50A 50/60 Hz.

Each 1300W power supply that is available for use can supply this power:

- ▶ A maximum output power of 900W when using a 100 - 127V ac power source at 10.0A, 50/60 Hz
- ▶ A maximum output power of 1300W when using a 200 - 240V ac power source at 6.9A, 50/60 Hz (recommended power configuration)

Using two power sources provides power redundancy. The power supplies must be matched; you cannot have one 900W and one 1300W power scenario.

**Note:** We suggest that you place the two power supplies on different circuits.

Figure 3-1 shows two power connections.

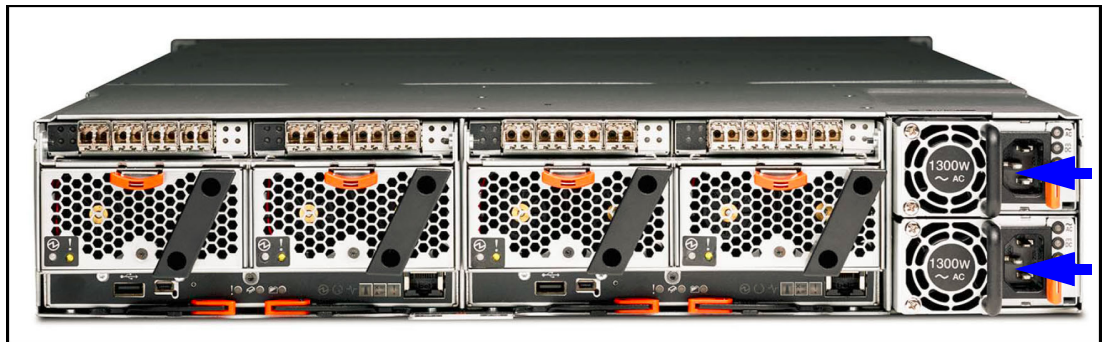


Figure 3-1 FlashSystem 900 rear view and power connections

The power cables are specific to the power requirements of your country or region.

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

### 3.4 Planning for configuration

You must plan for the management address and service address data before the system is installed. A management address must be allocated for the system. The *management address* provides access to system configuration and administration functions, such as the management GUI.

**Important:** The management and service addresses for the enclosure must be allocated within the same network.

The management IP address is required when the system is initialized. The system initialization tool (**InitTool**) that is provided on a USB flash drive provides a convenient wizard for initializing the system and configuring the management (cluster) IP addresses.

Use Table 3-10 to record the management IP address that is allocated for use by the system.

Table 3-10 Management IP address configuration

Configuration item	Value
Management IP address:	
Subnet mask:	
Gateway address:	

Two service addresses must be allocated to the enclosure. The enclosure canisters retain the service IP addresses, allowing convenient access to node configuration and service functions, such as the service assistant GUI and CLI for that node. IPv4 and IPv6 protocols can be used.

Use Table 3-11 to plan the service IP addresses required to perform service actions.

Table 3-11 Service IP address configuration

Configuration item	Value
<b>Service address 1</b>	
Management IP address:	
Subnet mask:	
Gateway address:	
<b>Service address 2</b>	
Management IP address:	
Subnet mask:	
Gateway address:	

Use Table 3-12 to configure the system for event notification.

Table 3-12 Event notification settings

Configuration item	Value
Email server address	
Simple Network Management Protocol (SNMP) server address	
SNMP community strings	
Syslog servers	



## 3.5 The call home option

FlashSystem 900 supports the setting up of 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.

To send email, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the management IP address. You can enable this option during setup. To set up call home, provide the following information:

- ▶ System location:
  - Company name
  - Street address
  - City
  - State or province
  - Postal code
  - Country or region
- ▶ Contact details information that IBM Support center can use to contact the customer to resolve system errors:
  - Contact name
  - Contact Email address
  - Telephone (primary)
  - Telephone (alternate, optional)
  - Machine location
- ▶ Email Server, call home, and event notifications are routed through this email server:
  - Email Serve IP Address
  - Email Serve Server Port

**Note:** For call home, the following two email addresses are automatically created:

- ▶ flash-sc1@vnet.ibm.com
- ▶ flash-sc2@vnet.ibm.com

## 3.6 TCP/IP requirements

To plan your installation, consider the TCP/IP address requirements of the system and the requirements to access other services. You must also plan for the Ethernet address allocation and for the configuration of the Ethernet router, gateway, and firewall.

Table 3-13 lists the TCP/IP ports and services that are used.

Table 3-13 TCP/IP ports and services

Service	Traffic direction	Protocol	Port	Service type
Email (SMTP) notification and inventory reporting	Outbound	TCP <sup>a</sup>	25	Optional
SNMP event notification	Outbound	UDP <sup>b</sup>	162	Optional
Syslog event notification	Outbound	UDP	514	Optional
IPv4 Dynamic Host Configuration Protocol (DHCP) (Node service address)	Outbound	UDP	68	Optional
IPv6 DHCP (Node service address)	Outbound	UDP	547	Optional
Network Time Protocol (NTP) server	Outbound	UDP	123	Optional
Secure Shell (SSH) for CLI access	Inbound	TCP	22	Mandatory
HTTPS for GUI access	Inbound	TCP	443	Mandatory
HTTPS for new firmware check by the GUI	Outbound	TCP	443	Optional
Remote user authentication service - HTTP	Outbound	TCP	16310	Optional
Remote user authentication service - HTTPS	Outbound	TCP	16311	Optional
Remote user authentication service - Lightweight	Outbound	TCP	389	Optional
Wake On LAN	Inbound	N/A	N/A	Mandatory

a. Transmission Control Protocol (TCP)

b. User Datagram Protocol (UDP)

**Note:** Both IPv4 and IPv6 addresses are supported.

For configuration and management, you must allocate an IP address to the Ethernet management port on each canister, which is referred to as the *management IP address*. If both IPv4 and IPv6 operate concurrently, an address is required for each protocol.

You can configure the enclosure for event notification by SNMP, syslog, or email. To configure notification, you must ensure that the SNMP agent, syslog IP addresses, or SMTP email server IP addresses can be accessed from all management addresses.

The system does not use name servers to locate other devices. You must supply the numeric IP address of the device. To locate a device, the device must have a fixed IP address.

For example, when you click **GUI** → **Settings** → **General** → **Upgrade Software**, the software level is checked. This check is done by the GUI by using the specific URL:

<https://public.dhe.ibm.com/storage/flash/9840.js>

To be able to perform this check, the system that runs the GUI must have access to this URL using port 443.

## 3.7 Planning for encryption

Planning for encryption involves purchasing a licensed function and then activating and enabling the function on the system.

To encrypt data that is stored on drives, they must contain an active license and be configured to use encryption. When encryption is activated and enabled on the system, valid encryption keys must be present on the system when the system unlocks the drives or the user generates a new key. The encryption key must be stored on USB flash drives that contain a copy of the key that was generated when encryption was enabled. Without these keys, user data on the drives cannot be accessed. The encryption key is read from the USB flash drives that were created during system initialization.

Before activating and enabling encryption, determine which method to use for accessing key information during times when the system requires the presence of an encryption key. The system requires an encryption key to be present during the following operations:

- ▶ System power-on
- ▶ System reboot
- ▶ User initiated rekey operations
- ▶ Firmware update

Consider the following factors when planning for encryption:

- ▶ Physical security of the system
- ▶ Need and benefit of manually providing encryption keys when the system requires
- ▶ Availability of key data
- ▶ Encryption license is purchased, activated, and enabled on the system

Two options are available for accessing key information on USB flash drives:

- ▶ USB flash drives are inserted in the system at all times

If you want the system to unlock the drives automatically when the system requires an encryption key to be present, a USB flash drive must be left inserted in both of the two canisters. This way both canisters have access to the encryption key. This method requires that the physical environment where the system is located be secure. A secure location prevents an unauthorized person from making copies of the encryption keys, stealing the system, or accessing data that is stored on the system. If a USB flash drive containing valid encryption keys is left inserted in both of the two canisters, the system will always have access to the encryption keys and the user data on the drives will always be accessible.

**Note:** This method requires that the physical environment where the system is located is secure. If the location is secure, it prevents an unauthorized person from making copies of the encryption keys, stealing the system, or accessing data that is stored on the system.

- USB flash drives are never inserted into the system except as required

For the most secure operation, do not keep the USB flash drives inserted in the canisters on the system. However, this method requires that you manually insert the USB flash drives that contain copies of the encryption key in the canisters during operations for which the system requires an encryption key to be present. USB flash drives that contain the keys must be stored securely to prevent theft or loss. During the operations for which the system requires an encryption key to be present, the USB flash drives must be inserted manually into each canister so data can be accessed. After the system completes unlocking the drives, the USB flash drives must be removed and stored securely.

**Note:** The USB flash drives must be stored securely to prevent theft or loss.

You can enable encryption and change the encryption key stored on the USB device at any time by using the GUI or the CLI.

**Note:** You can encrypt an existing FlashSystem 900. For assistance or questions about purchasing this licensed function, see your IBM sales representative or IBM Business Partner.

## 3.8 Checking web browser settings for the management GUI

To access the management GUI, you must ensure that your web browser is supported and that the correct settings are enabled.

At the time of writing this book, the management GUI supports the following web browsers and versions:

- Mozilla Firefox 32
- Mozilla Firefox Extended Support Release (ESR) 31
- Microsoft Internet Explorer (IE) 10 and 11
- Google Chrome 37

For a list of supported web browsers, go to the “Planning” topic of the IBM FlashSystem 900 web page (in the IBM Knowledge Center) and click **Checking your web browser settings for the management GUI**:

<http://ibm.co/1Gduc12>

IBM supports later versions of the browsers if the vendors do not remove or disable functionality that the product relies on. For browser levels later than the versions that are certified with the product, IBM customer support accepts usage-related and defect-related service requests. As with operating system and virtualization environments, if the support center cannot re-create the issue in our lab, we might ask the client to re-create the problem on a certified browser version to determine whether a product defect exists. Defects are not accepted for cosmetic differences between browsers or browser versions that do not affect the functional behavior of the product. If a problem is identified in the product, defects are accepted. If a problem is identified with the browser, IBM might investigate potential solutions or workarounds that the client can implement until a permanent solution becomes available.

## Procedure to configure your web browser

To configure your web browser, follow these steps:

### 1. Enable JavaScript for your web browser:

- For Mozilla Firefox, JavaScript is enabled by default and requires no extra configuration.
- For Microsoft Internet Explorer (IE) running on Microsoft Windows 7:
  - i. In Internet Explorer, click **Tools** → **Internet Options**.
  - ii. Click **Security Settings**.
  - iii. Click **Internet** to choose the Internet zone.
  - iv. Click **Custom Level**.
  - v. Scroll to the Scripting section, and then in Active Scripting, click **Enable**.
  - vi. Click **OK** to close Security Settings.
  - vii. Click **Yes** to confirm the change for the zone.
  - viii. Click **OK** to close Internet Options.
  - ix. Refresh your browser.
- For Microsoft Internet Explorer (IE) running on Microsoft Windows Server 2008:
  - i. In Internet Explorer, click **Tools** → **Internet Options**.
  - ii. Click **Security**.
  - iii. Click **Trusted sites**.
  - iv. On the **Trusted sites** dialog, verify that the web address for the management GUI is correct and click **Add**.
  - v. Verify that the correct web address was added to the Trusted sites window.
  - vi. Click **Close** on the Trusted sites window.
  - vii. Click **OK**.
  - viii. Refresh your browser.
- For Google Chrome:
  - i. On the menu bar in the Google Chrome browser window, click **Settings**.
  - ii. Click **Show advanced settings**.
  - iii. In the **Privacy** section, click **Content settings**.
  - iv. In the **JavaScript** section, select **Allow all sites to run JavaScript**.
  - v. Click **OK**.
  - vi. Refresh your browser.

### 2. Enable cookies in your web browser:

- For Mozilla Firefox:
  - i. On the menu bar in the Firefox browser window, click **Tools** → **Options**.
  - ii. On the Options window, select **Privacy**.
  - iii. Set “Firefox will” to **Use custom settings for history**.
  - iv. Select **Accept cookies from sites** to enable cookies.
  - v. Click **OK**.
  - vi. Refresh the browser.
- For Microsoft Internet Explorer:
  - i. In Internet Explorer, click **Tools** → **Internet Options**.
  - ii. Click **Privacy**. Under Settings, move the slider to the bottom to allow all cookies.
  - iii. Click **OK**.
  - iv. Refresh your browser.

- For Google Chrome:
  - i. On the menu bar in the Google Chrome browser window, click **Settings**.
  - ii. Click **Show advanced settings**.
  - iii. In the **Privacy** section, click **Content settings**.
  - iv. In the **Cookies** section, select **Allow local data to be set**.
  - v. Click **OK**.
  - vi. Refresh your browser.
- 3. For Mozilla Firefox *only*, enable scripts to disable or replace context menus:
  - a. On the menu bar in the Firefox browser window, click **Tools** → **Options**.
  - b. On the Options window, select **Content**.
  - c. Click **Advanced** by the Enable JavaScript setting.
  - d. Select **Disable or replace context menus**.
  - e. Click **OK** to close the Advanced window.
  - f. Click **OK** to close the Options window.
  - g. Refresh your browser.
- 4. Enable TLS 1.1/1.2 (Microsoft Internet Explorer 10 only; for Internet Explorer 11 and later TLS 1.1 or 1.2 enabled by default):
  - a. Open Internet Explorer.
  - b. Select **Tools** → **Internet Options**.
  - c. Select the **Advanced** tab.
  - d. Scroll to the **Security** section.
  - e. Check **Use TLS 1.1** and **Use TLS 1.2**.

## 3.9 Licensing

There is only one license with the FlashSystem 900. The FlashSystem 900 storage system provides support for AES XTS 256 data-at-rest encryption when the Encryption Enablement Pack, Feature Code AF14 is ordered.

For assistance or questions about the Encryption Enablement Pack for the FlashSystem 900, see your IBM sales representative or IBM Business Partner.

## 3.10 Supported hosts and operating system considerations

Always check the IBM System Storage Interoperation Center (SSIC) to get the latest information about supported operating systems, hosts, switches, and so on:

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

If a configuration that you want is not available at the SSIC, a Solution for Compliance in a Regulated Environment (SCORE) or a request for price quotation (RPQ) must be submitted to IBM requesting approval. To submit a SCORE or RPQ, contact your IBM representative.



# Installation and configuration

In this chapter, you learn how to install and configure the IBM FlashSystem 900. The chapter describes system cabling and management, and demonstrates how the initial setup procedure prepares the system for use.

This chapter covers the following topics:

- ▶ First-time installation
- ▶ Cabling the system
- ▶ Initializing the system
- ▶ RAID storage modes
- ▶ Connectivity guidelines for improved performance

## 4.1 First-time installation

The initial installation of the IBM FlashSystem 900 includes unpacking the system and installing it in a rack. When the system is physically installed in rack, it must be connected to power and cabled for management. Also, it must be cabled for communication with hosts.

This chapter describes unpacking the system to getting it operating in a functional state so it is ready for use.

### 4.1.1 Installing the hardware

The IBM FlashSystem 900 has its replaceable components installed from either the front or the back side of the enclosure.

#### **Installation poster**

The installation poster (Figure 4-1 on page 69), which is included with the system when it is delivered from the factory, is an overview of how to prepare the system for first-time use.





# IBM FlashSystem 900 Installation Poster

## ① Prepare for setup

You will need:



Installation Information:



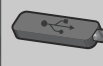
<https://ibm.biz/BdEANE>

Be aware that:



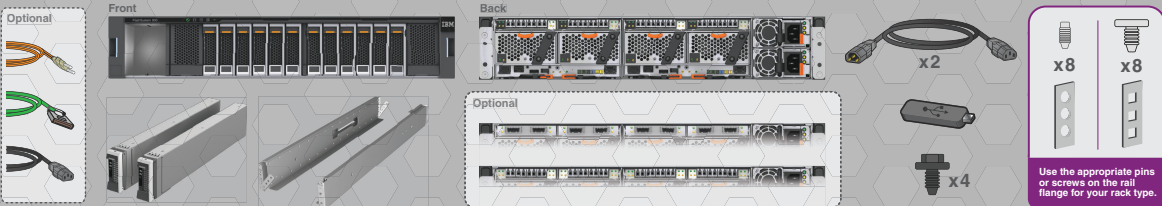
Before installing your FlashSystem 900 components, read the IBM Systems Safety Notices in the publication package.

Important:

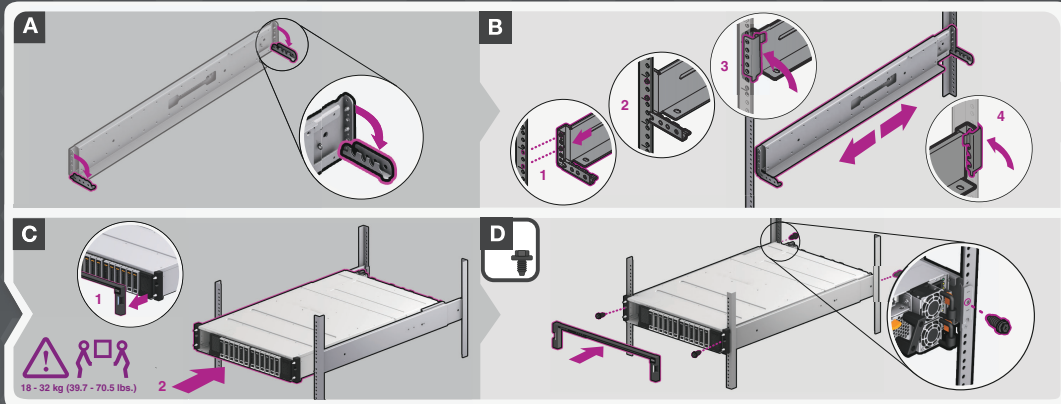


If this system will use encryption, ensure that you always know where the USB flash drives are located. During initialization, encryption keys are copied on the three USB flash drives, and these drives should be stored securely.

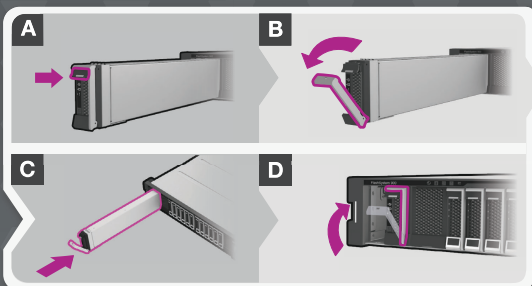
## ② Verify the enclosure package



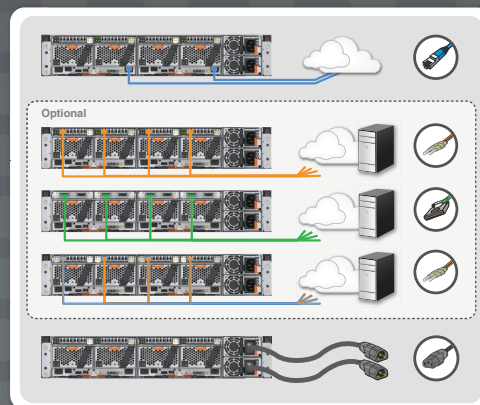
## ③ Install enclosures in a rack



## ④ Install batteries (x2)



## ⑤ Connect cables and power on



## ⑥ Set up the system (You will need: IP address / Subnet mask / Gateway)



© Copyright IBM Corporation 2015. Printed in Canada

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., 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 [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml). Linux is a registered trademark of Linux Torvalds in the United States, other countries, or both. Windows is a trademark of Microsoft Corporation in the United States, other countries, or both.

IBM FlashSystem 900 Installation Poster



Figure 4-1 Installation poster

## Rack installation

To install an enclosure in a rack, complete the following steps:

1. Install the rack mount rails.
2. Align the enclosure with the front of the rack cabinet.
3. Carefully slide the enclosure into the rack along the rails until the enclosure is fully inserted.
4. Secure the enclosure to the rack with a screw in the rack mounting screw hole on each side of the enclosure.
5. Pull off the front bezel to reveal the holes for the screws.
6. Insert the screws.
7. Replace the bezel by snapping it in place.
8. Install the two battery modules.

The rails are not designed to hold an enclosure that is partially inserted. The enclosure must therefore always be in a fully inserted position. To reduce the weight of the enclosure before lifting it, you can temporarily remove the two battery modules and the flash modules from the front of the enclosure.

**Note:** Install the enclosure where it can be easily serviced, but ensure that the rack is kept stable, for example, by installing enclosures beginning from the bottom.

## Installing batteries

The IBM FlashSystem 900 storage system contains two redundant backup batteries. In a loss of power to the enclosure, the batteries will supply power so that any volatile data is written to the flash modules and the system will shut down in an orderly manner.

The IBM FlashSystem 900 batteries install in the left front side of the enclosure and plug into the midplane, as shown in Figure 4-2.

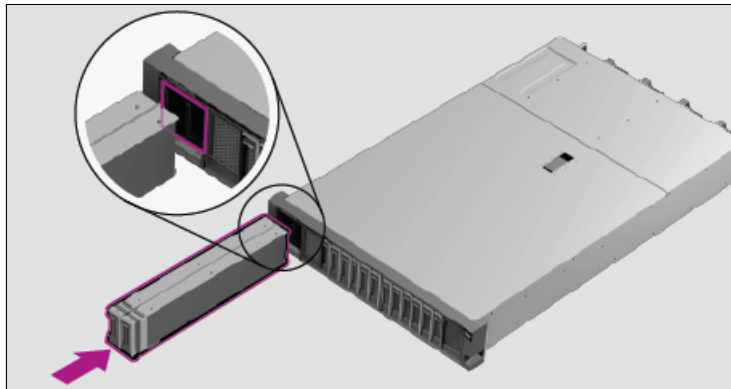


Figure 4-2 Install the batteries

**Note:** The IBM FlashSystem 900 does not have power switches, so whenever power is applied to the power inlets, the system will be powered on. Batteries and flash cards must be installed before power cables are connected to the power inlets.

## 4.2 Cabling the system

Various considerations exist when you plan and physically install the IBM FlashSystem 900. This section lists them and includes several notes and preferred practice considerations for Fibre Channel cabling, network cabling, physical installation, and configuration.

### 4.2.1 Cabling for Fibre Channel

IBM FlashSystem 900 supports either 16 Gbps Fibre Channel (FC), 8 Gbps FC, 10 Gbps Fibre Channel over Ethernet (FCoE), 10 Gbps iSCSI, or Quadruple Data Rate (QDR) InfiniBand interfaces.

In this example environment, the FlashSystem 900 is configured with four dual port 16 Gbps FC cards, for a total of eight 16 Gbps FC ports. *Only two ports for each card are active for 16 Gbps.* The optimal cabling scenario with this hardware configuration is to cable Port 1 (P1) on each interface card to storage area network (SAN) Fabric A and Port 2 (P2) on each card to SAN Fabric B, as shown in Figure 4-3.

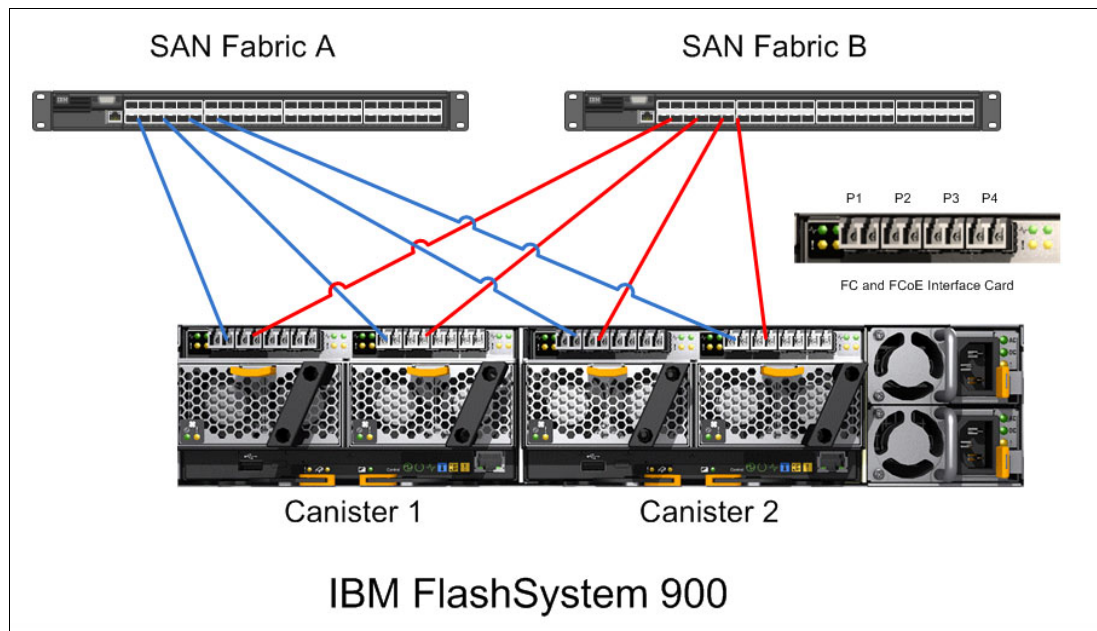


Figure 4-3 Cabling for 16 Gbps FC connectivity

The FC interface cards physically have four ports, but only port P1 and P2 are enabled in the 16 Gbps configuration. The same type of interface card hardware is used for 8 Gbps FC connectivity and for 10 Gbps FCoE connectivity. In the 8 Gbps FC and the 10 Gbps FCoE configuration, all four ports on each interface card are enabled. This configuration achieves SAN fabric switch-level redundancy and IBM FlashSystem 900 interface card-level redundancy, providing protection in an issue or failure.

A further consideration is to distribute the IBM FlashSystem 900 FC ports onto different SAN switch port groups to distribute the workload and provide the best available bandwidth. For more information about SAN switch port groups for IBM/Brocade switches, see the book about the IBM b-type Gen 5, 16 Gbps switches and Network Advisor, *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116.

If the environment contains only a single SAN switch, all FC ports from the FlashSystem 900 are cabled to this switch. Although this configuration is technically possible, it does not provide any redundancy in a SAN switch failure.

**Note:** The 16 Gbps FC, 8 Gbps FC, 10 Gbps FCoE, and iSCSI interface cards look similar and are all equipped with four physical ports. However, these interface cards are not interchangeable and must be ordered with the correct IBM feature code to provide the needed functionality.

## FC port speed settings

When running FC, the IBM FlashSystem 900 can be configured with up to 8 ports at 16 Gbps or 16 ports at 8 Gbps.

Depending on the configuration, the connected SAN switches can be capable of 16 Gbps or 8 Gbps. However, the IBM FlashSystem 900 FC ports can also run at 4 Gbps, if required. The preferred practice is to manually configure and fix the speed of the SAN switch ports to the highest mutually available speed, rather than using auto negotiate. This is done purely for consistency and stability.

Example 4-1 shows a truncated **switchshow** command, output from an IBM SAN switch, showing ports connected to the IBM FlashSystem 900 set to 16 Gbps.

*Example 4-1 The switchshow output from an IBM SAN switch that shows ports connected at 16 Gbps*

---

```
IBM_2498_F48:FID128:admin> switchshow
```

---

Index	Port	Address	Media	Speed	State	Proto	
=====							
0	0	010000	id	N8	Online	FC	F-Port 21:01:00:1b:32:2a:23:b1
1	1	010100	id	<b>N16</b>	<b>Online</b>	<b>FC</b>	<b>F-Port 50:05:07:60:5e:fe:0a:dd</b>
2	2	010200	--	N16	No_Module	FC	
3	3	010300	--	N16	No_Module	FC	
4	4	010400	id	N8	Online	FC	F-Port 21:00:00:24:ff:22:f9:ea
5	5	010500	id	<b>N16</b>	<b>Online</b>	<b>FC</b>	<b>F-Port 50:05:07:60:5e:fe:0a:d9</b>
6	6	010600	--	N16	No_Module	FC	
7	7	010700	--	N16	No_Module	FC	
8	8	010800	id	N16	Online	FC	F-Port 10:00:8c:7c:ff:0b:0f:00
9	9	010900	id	N16	Online	FC	F-Port 10:00:8c:7c:ff:0b:78:81
10	10	010a00	--	N16	No_Module	FC	
11	11	010b00	id	N8	Online	FC	F-Port 10:00:00:00:c9:d4:94:11

---

The N16 identifier under the Speed column in Example 4-1 indicates that the system negotiates to 16 Gbps. To fix the ports at 16 Gbps, use the command shown in Example 4-2 on page 73.

#### Example 4-2 Fixing port speed at 16 Gbps on the IBM/Brocade SAN switch

```
IBM_2498_F48:FID128:admin> portcfgspeed 1 16;  
IBM_2498_F48:FID128:admin> portcfgspeed 5 16;
```

```
IBM_2498_F48:FID128:admin> switchshow
```

Index	Port	Address	Media	Speed	State	Proto		
0	0	010000	id	N8	Online	FC	F-Port	21:01:00:1b:32:2a:23:b1
1	1	010100	id	<b>16G</b>	<b>Online</b>	<b>FC</b>	<b>F-Port</b>	<b>50:05:07:60:5e:fe:0a:dd</b>
2	2	010200	--	N16	No_Module	FC		
3	3	010300	--	N16	No_Module	FC		
4	4	010400	id	N8	Online	FC	F-Port	21:00:00:24:ff:22:f9:ea
5	5	010500	id	<b>16G</b>	<b>Online</b>	<b>FC</b>	<b>F-Port</b>	<b>50:05:07:60:5e:fe:0a:d9</b>
6	6	010600	--	N16	No_Module	FC		
7	7	010700	--	N16	No_Module	FC		
8	8	010800	id	16G	Online	FC	F-Port	10:00:8c:7c:ff:0b:0f:00
9	9	010900	id	16G	Online	FC	F-Port	10:00:8c:7c:ff:0b:78:81
10	10	010a00	--	N16	No_Module	FC		
11	11	010b00	id	N8	Online	FC	F-Port	10:00:00:00:c9:d4:94:11

The ports are no longer negotiating speed but are fixed to 16 Gbps.

For more details of how to configure the IBM and Brocade SAN switches for the correct and optimal interconnection with the IBM FlashSystem 900, see *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186.

## 4.2.2 Cabling for FCoE

The IBM FlashSystem 900 supports Fibre Channel over Ethernet (FCoE) when ordered with the correct interface cards for FCoE. Cables used for FCoE are the Fibre Optical Multi Mode type that is also used for FC and iSCSI connectivity. The IBM FlashSystem 900 FCoE ports connect to switches, such as the Cisco Nexus 5000 Series Switches and the Cisco MDS 9000 Family configured with Fibre Optical ports and connectors. The operating speed for FCoE is 10 Gbps. The IBM FlashSystem 900 supports 16 FCoE ports.

For optimal redundancy and performance, the IBM FlashSystem 900 FCoE ports must connect to different and redundant FCoE switches in the same way that is illustrated in Figure 4-3 on page 71, except that the switches are FCoE-capable switches and not SAN switches.

## 4.2.3 Cabling for iSCSI

The IBM FlashSystem 900 supports iSCSI when ordered with the correct interface cards for iSCSI. Cables used for iSCSI are the same Fibre Optical Multi Mode type that is also used for FC and FCoE connectivity. The IBM FlashSystem 900 iSCSI ports connect to switches, such as the Cisco Nexus 5000 Series Switches equipped with Fibre Optical ports. The operating speed for iSCSI is 10 Gbps. The IBM FlashSystem 900 supports 16 iSCSI ports and IPV4 connectivity.

For optimal redundancy and performance, the FlashSystem 900 iSCSI ports must connect to different and redundant LAN switches. The hosts that connect to the FlashSystem 900 storage must access iSCSI ports on different iSCSI I/O cards on both of the FlashSystem 900 controller canisters.

## 4.2.4 Cabling for QDR InfiniBand

The IBM FlashSystem 900 supports Quadruple Data Rate (QDR) InfiniBand when ordered with the correct interface cards for this protocol. Each interface card has two ports operating at 40 Gbps. The IBM FlashSystem 900 supports four QDR InfiniBand adapters for a total of eight ports, each operating at 40 Gbps.

QDR InfiniBand is used in situations where powerful, high-demand servers need high-bandwidth access to the IBM FlashSystem 900.

The 40 Gbps QDR InfiniBand host bus adapters (HBAs) can be ordered from IBM or from Mellanox Technologies. Mellanox Technologies also offers InfiniBand switches with which you can connect multiple hosts, through a redundant switch configuration, to a single IBM FlashSystem 900.

Cables used for QDR InfiniBand can be ordered from IBM and can be up to 10 meters (32.8 feet) long.

Figure 4-4 shows a FlashSystem 900 canister mounted with a four-port QDR InfiniBand interface card.

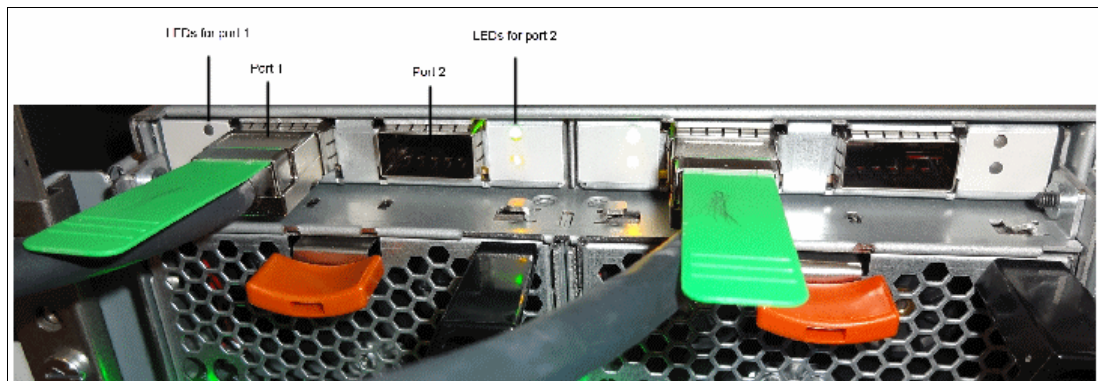


Figure 4-4 QDR InfiniBand interface card

## 4.2.5 FC cable type

OM4 standard cabling must be used where possible to provide the clearest connection. All of the connectors need to be the LC-LC connector standard.

## 4.2.6 Ethernet management cabling

The IBM FlashSystem 900 contains dual management control processors, each with its own Ethernet management port. These two Ethernet management ports operate with a single clustered IP address for system management. If one control canister is taken out for service or is inoperable, the other control canister services the clustered IP address.

Individual canister service IP addresses are configurable through the Ethernet management ports. These IP addresses provide access to each of the canister modules and are, among other features, capable of setting a canister into a service state or rebooting a specified canister.

Figure 4-5 on page 76 identifies the Ethernet management ports on the system.

The default speed setting of the IBM FlashSystem 900 Ethernet management network interface is auto, allowing the port to negotiate speed and duplex settings with the switch. The maximum configurable speed of the interface is 1 Gbps full duplex.

## 4.2.7 Power requirements

The IBM FlashSystem 900 includes dual, redundant AC power modules. Plan to attach each of the two power supplies in the enclosure to separate main power supply lines and to power sources that provide power in the 100 - 240V AC range, depending on the country.

A single IBM FlashSystem 900 has a power consumption rating of 1300 watts maximum with 625 watts typically seen operating in RAID 5 (625 watts for a 70/30 read/write workload on an eight 2 TB flashcard system).

Although the system operates when only one power supply is connected, this configuration is not advised. Using the power cords that are provided, connect each IBM FlashSystem 900 power inlet to an uninterruptible power supply (UPS) battery-backed power source. If possible, always connect each of the power cords to separate circuits.

### **Tips:**

- ▶ The IBM FlashSystem 900 device must be connected to a UPS-protected power source and each power supply must be on a different phase of power to provide power redundancy.
- ▶ So that the 1300 watt power supply can supply the full 1300 watts for which it is rated, it must be attached to a high-line voltage (220 - 240 volts).

## 4.2.8 Cooling requirements

The IBM FlashSystem 900 storage system has a British Thermal Unit (BTU) per-hour rating of approximately 2133 BTU. For maximum configurations, it can be as high as 3753 BTU.

It is suggested that the cooling vents in the room be located at the front of the rack because the air flows from the front of the rack to the back.

## 4.2.9 Cable connector locations

The IBM FlashSystem 900 has cable connections for power and management and optional connections for either FC, FCoE, iSCSI, or QDR InfiniBand interface cards.

The FC, FCoE, and iSCSI interface cards are typically used for connecting to switches where multiple servers will be able to connect to the IBM FlashSystem 900. The QDR InfiniBand interfaces can be connected directly to a host, or they can be connected through QDR InfiniBand switches.

Various options for host connectivity are described in Chapter 2, “IBM FlashSystem 900 architecture” on page 15.

Planning for Ethernet connections and connectivity is described in Chapter 3, “Planning” on page 49.



Figure 4-5 shows the rear side of the IBM FlashSystem 900 and its connectors.

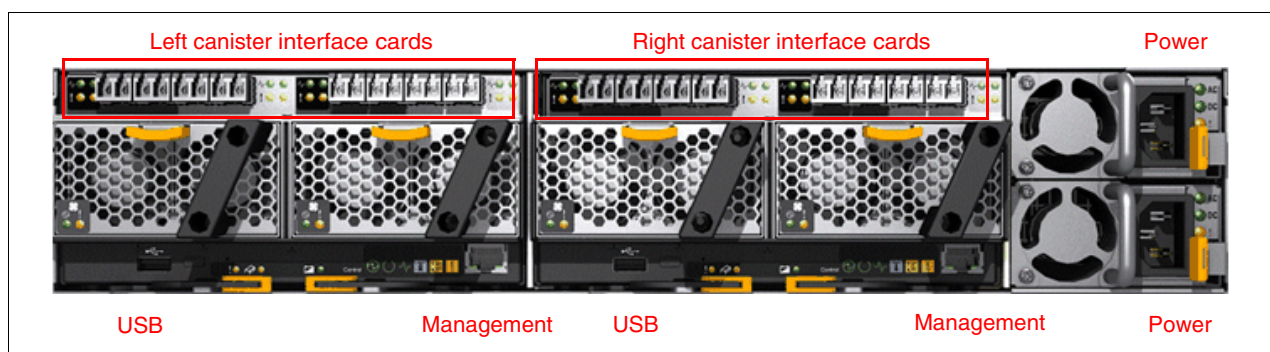


Figure 4-5 Rear view of the IBM FlashSystem 900

The IBM FlashSystem 900 is configured with either 16 Gbps FC, 8 Gbps FC, 10 Gbps FCoE, 10 Gbps iSCSI, or 40 Gbps QDR InfiniBand interface cards. Combinations of interface card types in a single system are not supported. The FlashSystem 900 depicted in Figure 4-5 is configured with FC/FCoE interface card types.

## 4.3 Initializing the system

After installing and powering on the new system, you must initialize it to be able to manage it. You will be using InitTool, and the System Setup wizard in the web management interface, FlashSystem 900 Storage Management, to initialize your system.

1. Prepare a USB flash drive with the correct IP address using InitTool by inserting the provided USB flash drive into a personal computer.
2. Run the InitTool and follow the instructions for using the System Initialization wizard.
3. Initialize the system by inserting the USB flash drive into the left canister USB port.
4. Log on to the system through your web browser to the FlashSystem 900 Storage Management GUI and continue with the System Setup wizard (shown in Figure 4-15 on page 87).
5. Upon completion of the wizard, insert the USB flash drive back into the same computer used in step 1.
6. Use the InitTool final window to review the results of the initialization (a summary window is shown in Figure 4-13 on page 84 and Figure 4-14 on page 85).

### 4.3.1 Prepare for initialization using InitTool

The supplied USB flash drive contains an initialization tool called InitTool, which is used to initialize the system. After initializing, you will be able to access the web management interface to complete the configuration procedures.

**Note:** If you do not have the original USB flash drive that is provided with your system and that contains `InitTool.bat`, download the latest software from the Fix Central website and create a new one:

<http://www.ibm.com/support/fixcentral>



The initialization tool (**InitTool**) can be used with the following systems:

- ▶ Microsoft Windows computer
- ▶ Linux computer
- ▶ Apple Macintosh computer

Before you begin:

- ▶ Ensure that the physical installation of the enclosure is complete.
- ▶ You need a computer to complete the initialization procedure.
- ▶ The computer must have a USB 2.0 port.

To initialize the system, complete the following steps:

1. Gather the information that you will use to configure the system:
  - a. You must have the IP network address that you will use to manage the system:
    - IP address
    - Subnet mask
    - Gateway
  - b. Optional: The following information is useful for enabling additional capabilities:
    - IP address of a Network Time Protocol (NTP) server for automatically setting date and time
    - IP address of a Simple Mail Transfer Protocol (SMTP) server for sending alert notifications
  - c. In the documentation package, find the USB flash drive that is included with your order.
  - d. Insert the USB flash drive into a USB port on the personal computer.
  - e. To launch the tool, open the USB flash drive, and double-click **InitTool.bat**. The initialization tool wizard starts.
  - f. In the wizard, click **Next** and select **Create a new system**.
  - g. In the wizard, follow the instructions, which are to complete these actions:
    - i. Enter the details of the system management address that you want to use.
    - ii. Plug in both power cables into the power supply units.  
Wait for the status LED to flash. This process can take up to 10 minutes.
    - iii. Take the USB flash drive and insert it into the IBM FlashSystem 900 and allow it to initialize.
    - iv. Return the flash drive to the workstation to check that the initialization completed.
2. If the system initialization completed successfully, click **Finish**.
  - If you have a network connection to the system, the system management GUI is displayed.
  - If the workstation does not have a network connection to the system, go to the workstation that you will use to manage the system and start a supported browser. Direct the browser to the management address that you specified for the system.
3. Log in with this user name and password:
  - User name: superuser
  - Password: passw0rd (The “0” character in the password is a zero.)
4. Follow the instructions in the window to begin setting up your system.

## Example: Using Microsoft Windows to initialize

The following example shows the steps involved in initializing the system using a Microsoft Windows computer to prepare the supplied USB flash drive using InitTool.

The first window in the System Initialization wizard shows that the tool can be used for different purposes:

- ▶ Install a new system
- ▶ Reset the superuser password
- ▶ Edit the service IP address

### ***Start the InitTool process***

Figure 4-6 shows the initial step in using InitTool.



Figure 4-6 System Initialization (InitTool) Welcome

## Tasks

In the second step of the System Initialization wizard, select **Yes** to configure a new system (Figure 4-7).

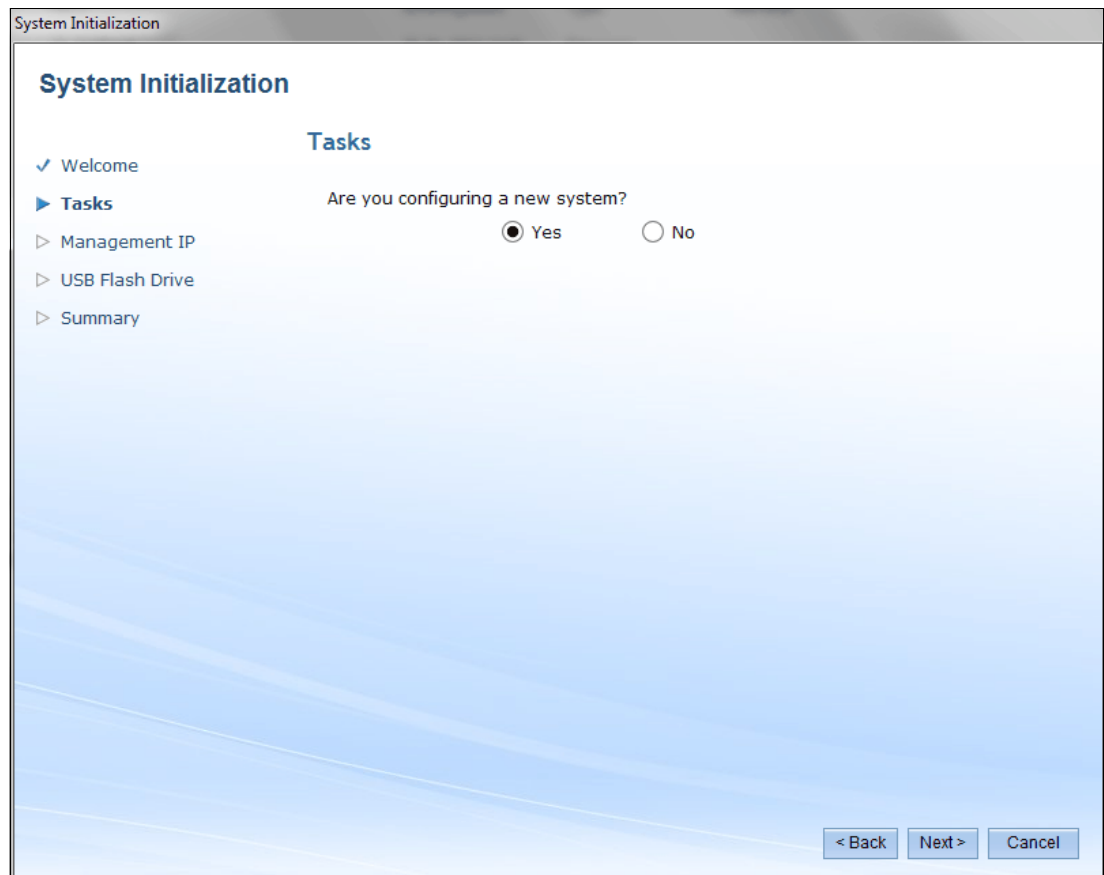


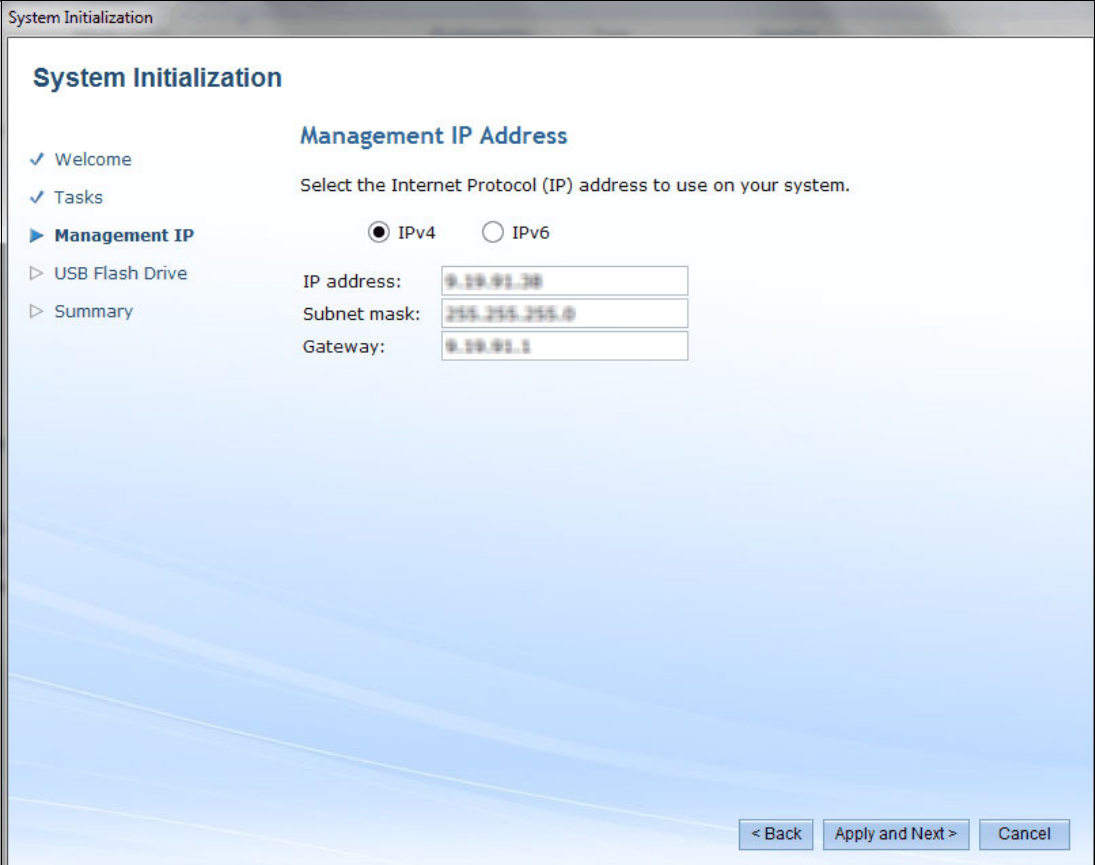
Figure 4-7 *InitTool Tasks*

If you selected **No**, the InitTool assumes that you are unable to access your system and gives you two options:

- ▶ Reset the superuser password
- ▶ Set the service IP address

### **Management IP address**

In the next step of the System Initialization wizard, type the IP address for managing the IBM FlashSystem 900 (Figure 4-8).



The image shows a screenshot of the 'System Initialization' window, specifically the 'Management IP Address' step. The window has a title bar 'System Initialization' and a sidebar on the left with a tree view containing 'Welcome', 'Tasks', 'Management IP' (selected), 'USB Flash Drive', and 'Summary'. The main area is titled 'Management IP Address' and contains the instruction 'Select the Internet Protocol (IP) address to use on your system.' Below this are two radio buttons: 'IPv4' (selected) and 'IPv6'. Under the 'IPv4' section, there are three input fields: 'IP address:' with the value '9.19.91.38', 'Subnet mask:' with the value '255.255.255.0', and 'Gateway:' with the value '9.19.91.1'. At the bottom right of the window are three buttons: '< Back', 'Apply and Next >', and 'Cancel'.

Figure 4-8 InitTool Management IP Address window

## Power on

In the next step, you are instructed to power on the IBM FlashSystem 900 and wait for the status LED to flash (Figure 4-9). Powering on can take up to 10 minutes.



Figure 4-9 InitTool Power on

### ***Insert USB into canister***

In the next step, you are instructed to remove the USB flash drive from the PC and insert it into the USB port in the *left* IBM FlashSystem 900 canister (controller) as shown in Figure 4-10. At the end of the process, the Identify LED turns on and then off.



Figure 4-10 *InitTool Insert USB flash drive into canister*

**Note:** The USB flash drive must be inserted into the left canister to correctly initialize the IBM FlashSystem 900.

When the system initialization process finishes, a results file, `satask_result.html`, is written to the USB flash drive by the IBM FlashSystem 900 canister. The results file indicates success or failure of the process. **InitTool** can be used to verify this result.

### Reinsert USB flash drive into personal computer

You are instructed to reinsert the USB flash drive into the Microsoft Windows computer (Figure 4-11).

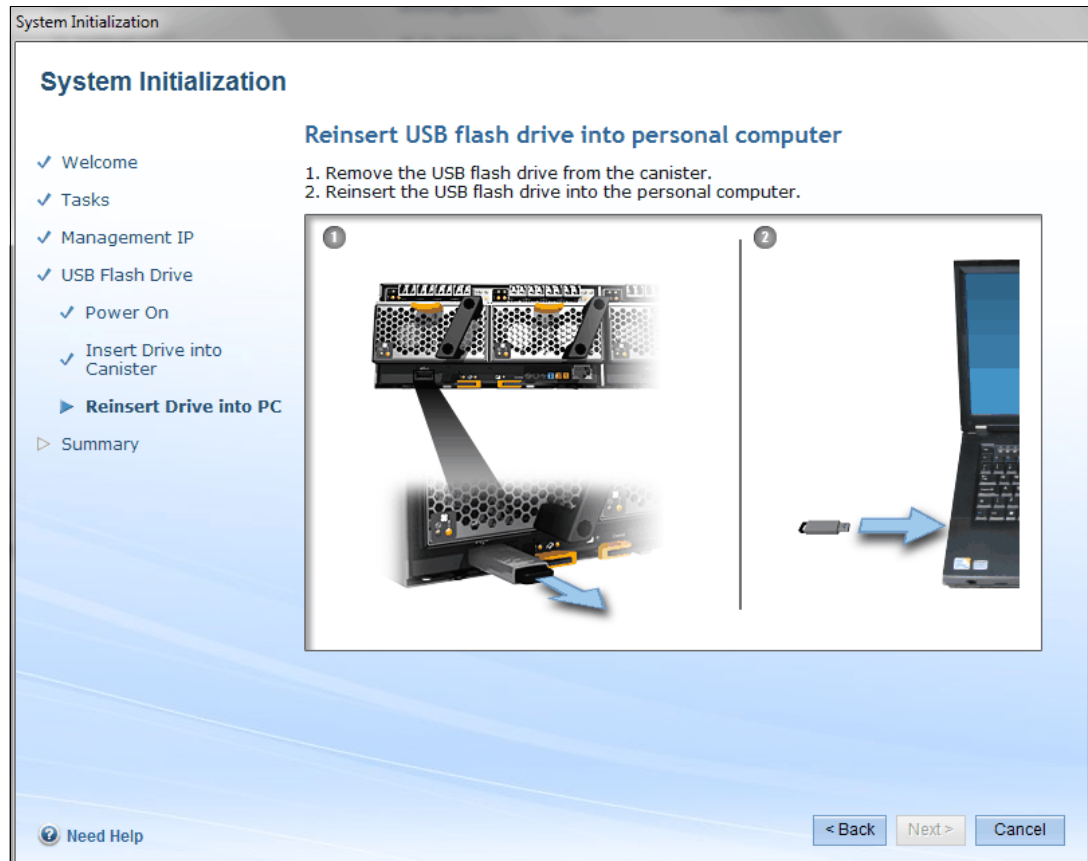


Figure 4-11 Reinsert USB flash drive into personal computer

### System initialization failed

If system initialization fails, a message is displayed (Figure 4-12).

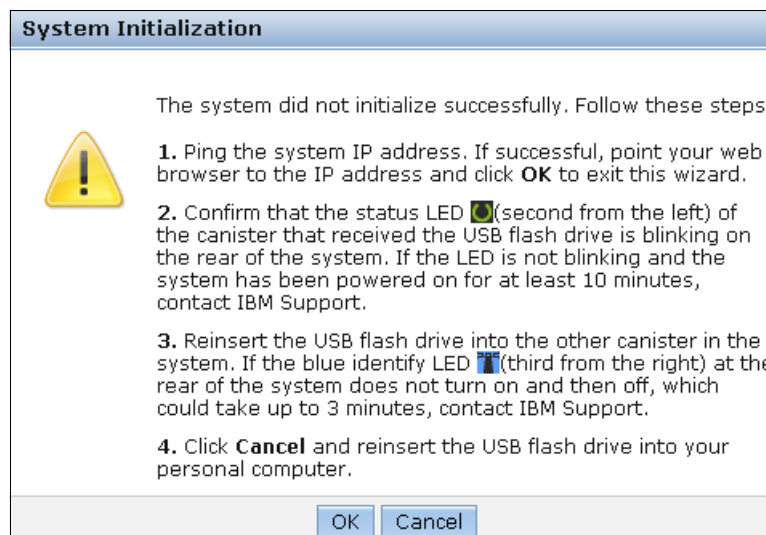


Figure 4-12 System initialization failed

System initialization can fail for multiple reasons. Follow the steps shown to resolve the situation. The initialization process checks the content of the `satask_result.html` file. Only if a cluster was successfully created, the Next option is enabled for the initialization process to continue.

### **Check connectivity**

Initialization now checks connectivity to the IBM FlashSystem 900 (Figure 4-13).

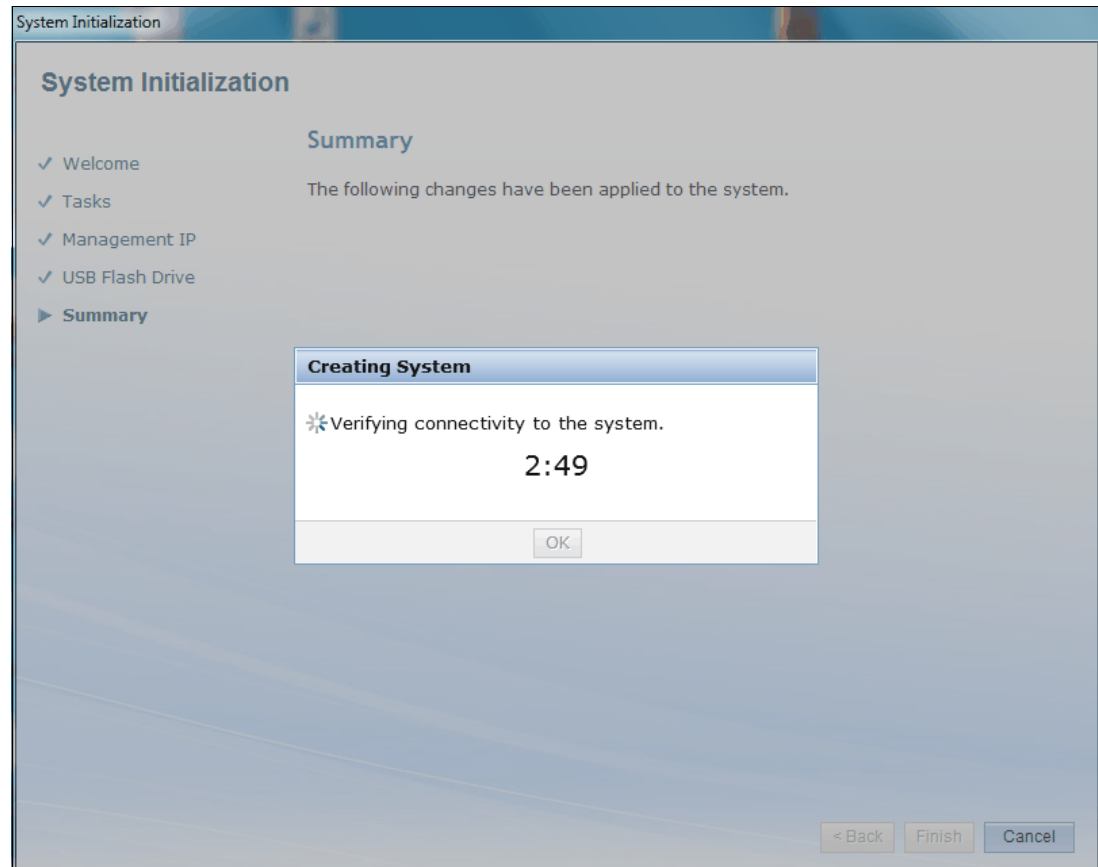


Figure 4-13 *InitTool: Verify connectivity*

### **Summary**

If the Ethernet ports of the newly initialized FlashSystem 900 are attached to the same network as the personal computer where **InitTool** was run, **InitTool** checks connectivity to the system and displays the result of the system initialization process.

When connectivity is successful, the Summary window opens (Figure 4-14 on page 85).



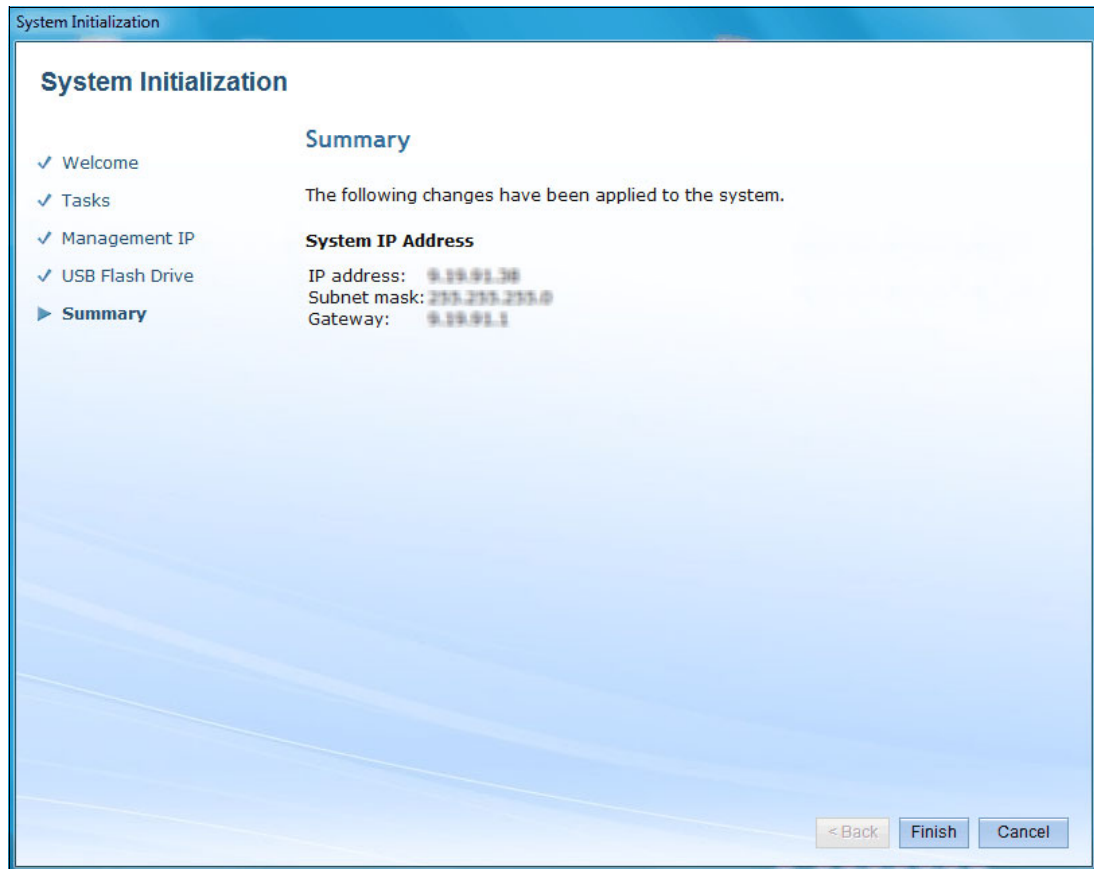


Figure 4-14 InitTool Summary

Click **Finish**. A message indicates that the initialization process completed successfully and the InitTool procedure completes.

The first part of the InitTool process is completed. You will be able to access the FlashSystem Storage Management GUI (Figure 4-15 on page 87) to complete the configuration process.

The next step in the IBM FlashSystem 900 installation process is to open a web browser and log on to the system at the selected IP address, which is demonstrated in 4.3.2, “Initializing the system through the web management interface” on page 86.

### Other purposes of InitTool

The InitTool can be used for other purposes in addition to initializing IBM FlashSystem 900. InitTool can also be used if you are unable to access the system. InitTool provides two options:

- ▶ Reset the superuser password:  
You get to the login prompt, but you do not know the superuser password. InitTool can, in this case, modify the superuser password.
- ▶ Set the service IP address:  
You want to access a specific canister but you do not know the service IP addresses. Specifying `/service` in the URL does not open the IBM FlashSystem 900 Service Assistant Tool. InitTool can, in this case, modify the service IP addresses.

As with the IBM FlashSystem 900 initialization process, **InitTool** creates a file (`satask.txt`) on the USB flash drive. This file contains a command that is read and executed by the IBM FlashSystem 900 canister when the US flash drive is inserted into the USB port of the system.

### 4.3.2 Initializing the system through the web management interface

After the IBM FlashSystem 900 is initialized using the USB flash drive, you use a supported web browser to point it to the selected address.

#### Supported web browsers

The web-based GUI simplifies storage management and provides a fast and efficient management tool. It is loosely based on the IBM System Storage XIV software and has a similar look and behavior.

The management GUI requires a supported web browser. At the time of writing this book, the management GUI supports the following web browsers with equal or later versions:

- ▶ Mozilla Firefox 32
- ▶ Mozilla Firefox Extended Support Release (ESR) 31
- ▶ Microsoft Internet Explorer (IE) 10 and 11
- ▶ Google Chrome 37

IBM supports later versions of the browsers if the vendors do not remove or disable functionality that the product relies upon.

Requirements for web browsers are as follows:

- ▶ JavaScript must be enabled.
- ▶ Cookies must be allowed.
- ▶ Enable the scripts to disable or replace context menus. (Mozilla Firefox only).
- ▶ Enable TLS 1.1/1.2 (Microsoft Internet Explorer 9 and 10 only).

For a list of supported web browsers, and how to configure them, see *Web Browser Requirements* in the IBM FlashSystem 900 Knowledge Center:

<https://ibm.biz/BdsjJ7>

## Log on and change password

In this step of the IBM FlashSystem 900 initialization procedure, log on to the web management GUI by specifying passw0rd (with a zero) as the password (Figure 4-15).



Figure 4-15 Initialization procedure logon as superuser

The IBM FlashSystem 900 at this point allows only the superuser user to access the system and therefore does not prompt for a user name.

The next step is to change the password (Figure 4-16).



Figure 4-16 Initialization procedure to change the password

The IBM FlashSystem 900 administrators are encouraged not to leave the superuser password as the default but instead to create individual users with their own passwords for security reasons. You can also configure authentication and authorization for users of the clustered system as described in “Configure remote authentication” on page 243.

## System Setup wizard

After the password is changed, the System Setup wizard starts. While using the wizard, you, as the administrator, are prompted for the following configuration information:

- ▶ Provide a system name.
- ▶ Configure the date and time in one of these ways:
  - NTP server (preferred)
  - Manual
- ▶ Configure encryption.
- ▶ Configure the access type:
  - Open Access Yes: All hosts have access to all volumes.
  - Open Access No: Volumes must be mapped to hosts.
- ▶ Set up call home.
- ▶ Confirm the number of flash modules.
- ▶ Confirm the summary of changes.

All of these steps can also be configured after the System Setup wizard is finished. The options for changing any of these configuration settings are to use either the GUI or CLI.

Figure 4-17 shows the first step of the System Setup wizard, which is invoked automatically after the change password step.

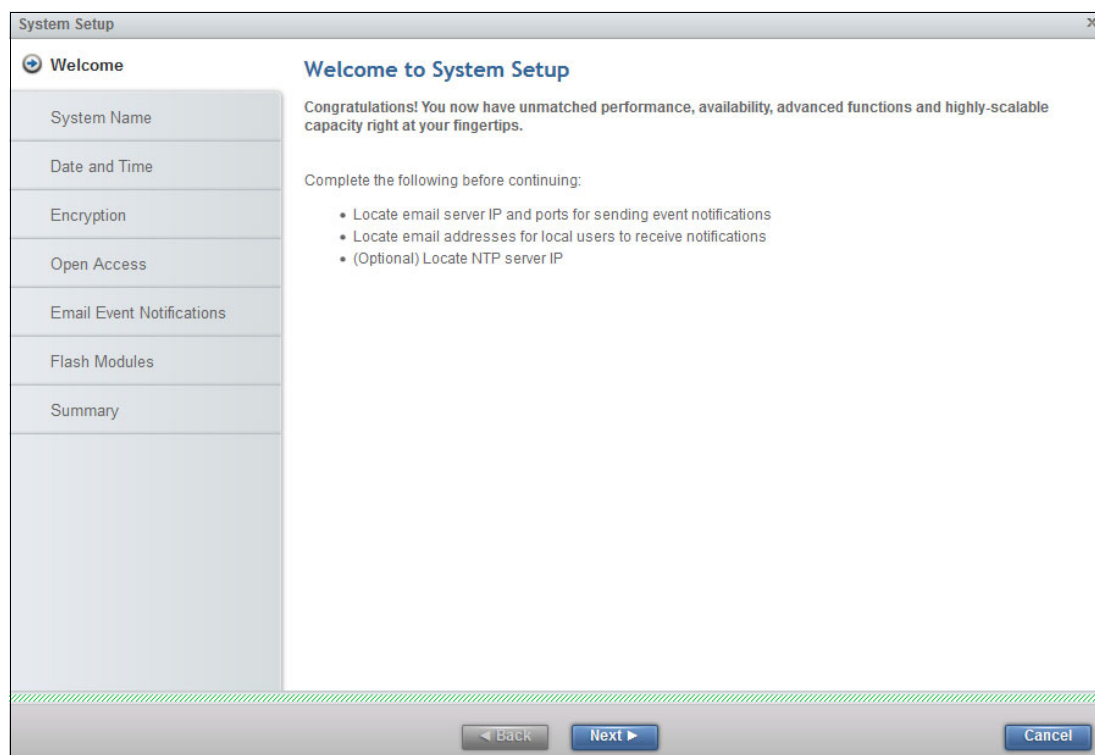


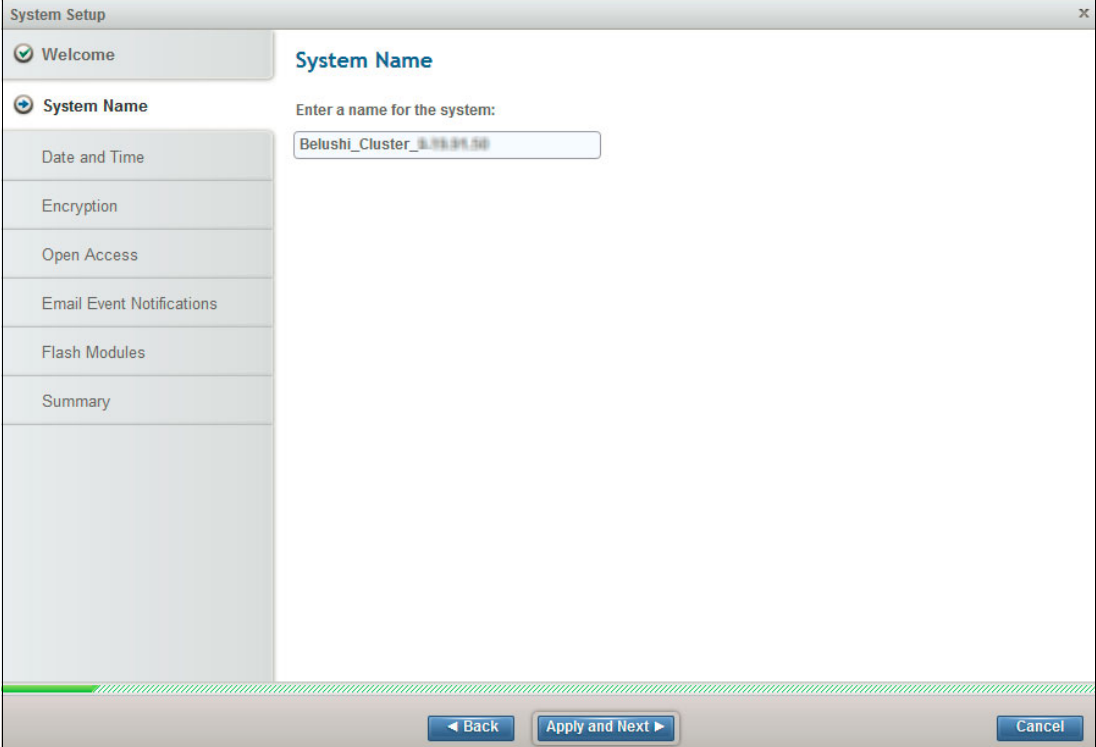
Figure 4-17 Initialization procedure welcome window

The Welcome window of the System Setup wizard tells you that the following information is needed:

- ▶ NTP server IP address for automatic and accurate system time
- ▶ SMTP server IP address for sending emails with warnings and alerts
- ▶ Email addresses for local users who will receive warnings and alerts

### ***Configure a system name***

Next, provide a host name for the system. This example uses the system name, Belushi\_Cluster\_9.xx.xx.xx (Figure 4-18). The host name can be changed later at any time from the main window at the FlashSystem 900 GUI.



The screenshot shows a window titled "System Setup" with a sidebar on the left containing the following options: Welcome (selected with a checkmark), System Name (selected with a blue arrow), Date and Time, Encryption, Open Access, Email Event Notifications, Flash Modules, and Summary. The main area is titled "System Name" and contains the text "Enter a name for the system:" followed by a text input field containing "Belushi\_Cluster\_9.99.99.99". At the bottom of the window are three buttons: "Back", "Apply and Next", and "Cancel".

*Figure 4-18 Initialization procedure to name the system*

## Configure date and time

Next, configure the system date and time. The preferred practice is to configure the system with an NTP server. By using an NTP server, the date and time settings are always correct, which is useful in log analysis and troubleshooting. If an NTP server is not available at the time of installation, it can be added later. You can then set the date and time manually as shown in Figure 4-19.

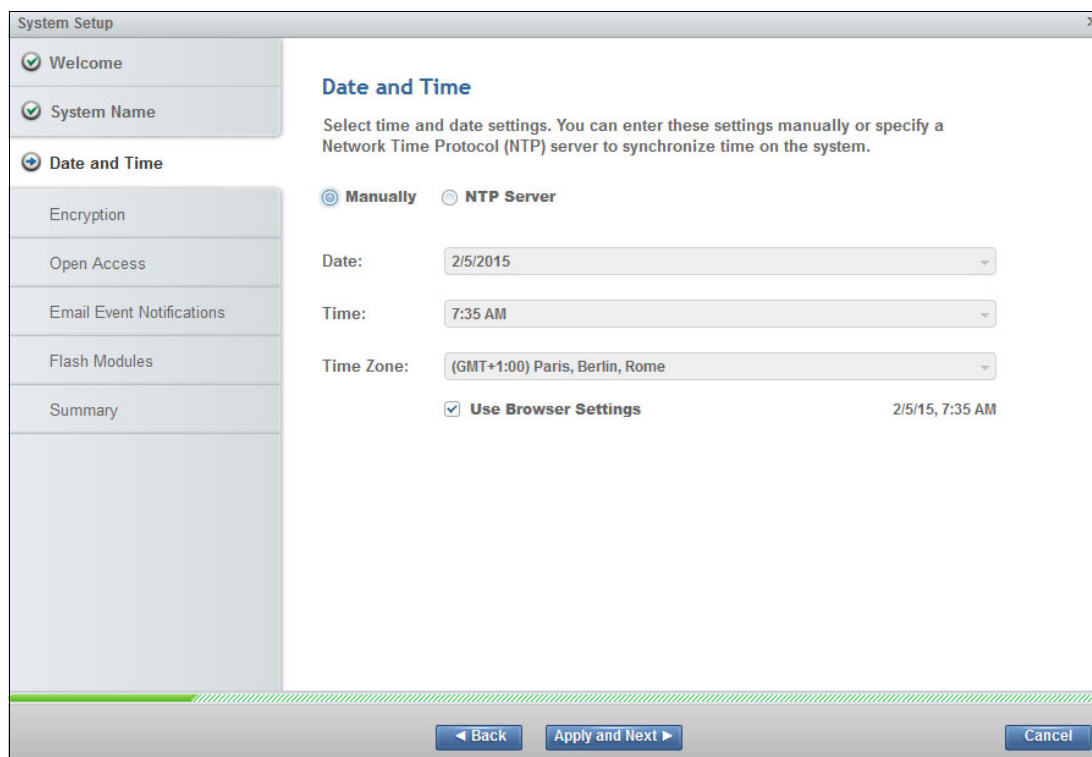


Figure 4-19 Initialization procedure date and time

Each time that a new command executes, the Task window opens. At first, the Task window shows collapsed output. By clicking **View more details**, the expanded output shows the command that is run on the IBM FlashSystem 900 (Figure 4-20).

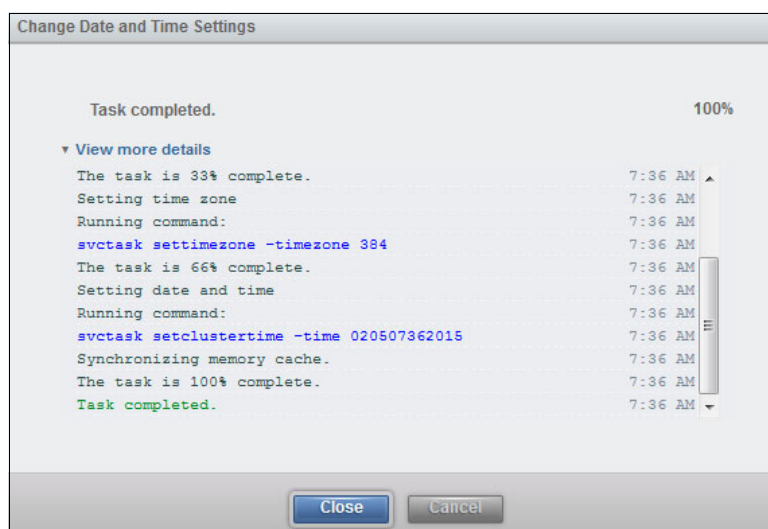


Figure 4-20 Initialization procedure task executing

A shortcut to setting the date and time manually is to select the **Use Browser Settings** check box, which causes the date and time to be inherited from the web browser in use.

### Encryption

When the IBM FlashSystem 900 initializes, it provides the option to encrypt the system and generate encryption keys. Users must purchase the Feature Code AF14 to enable encryption. Two USB flash drives are required and the encryption enablement procedure creates and store these keys on two USB flash drives which the user must install at the rear side into the FlashSystem 900 controllers.

If the system is encrypted the system can function without the USB flash drives containing the encryption key, but it cannot be rebooted, repaired or upgraded and the internal flash disks cannot be formatted or erased without access to the USB flash drives holding the encryption key.

**Note:** With FlashSystem 900, the encryption keys are cached. There is now the capability to run dumps (DMPs) to replace flash modules or canisters if this procedure is done concurrently.

FlashSystem 900 encryption can also be enabled after the system is initialized and while it is running full production. Starting with FlashSystem 840 release 1.3, it is not a disruptive operation to encrypt a FlashSystem 840 or 900.

At this time, select **No** to the system encryption question (Figure 4-21). For more information about how to enable encryption see 7.1.3, “Security menu” on page 243.

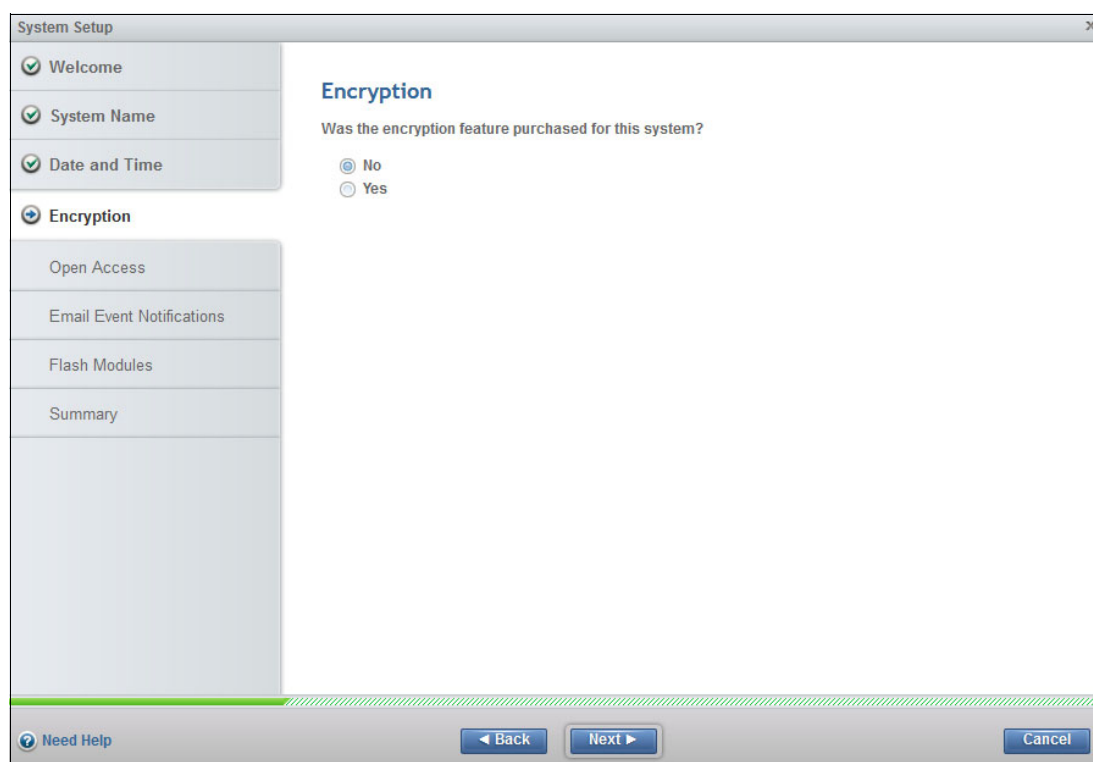


Figure 4-21 Initialization procedure encryption

### Allow or disallow open access

The IBM FlashSystem 900 can be configured to allow (enable) or disallow (do not enable) open access. Open access is feasible when the system is connected to a host or group of hosts that all need the same access to the FlashSystem 900 volumes.

An example is a FlashSystem 900 that functions as a fast data store for a number of VMware servers in a cluster. By zoning the SAN switches to connect the IBM FlashSystem 900 to the VMware servers, they are all able to access all volumes.

Allowing open access is used where the IBM FlashSystem 900 is connected to correctly zoned SAN switches. However, disallowing open access and forcing the system to map its volumes only to selected hosts provides an extra layer of security.

When the system is configured to disallow open access, several extra buttons become available in the GUI of the managed system. These buttons include the Hosts group of buttons and the Volumes by Host from the Volumes group of buttons. These buttons are for selective mapping of volumes to hosts.

Figure 4-22 shows the System Setup wizard where open access is disallowed (**No**), which is the option selected for this example.

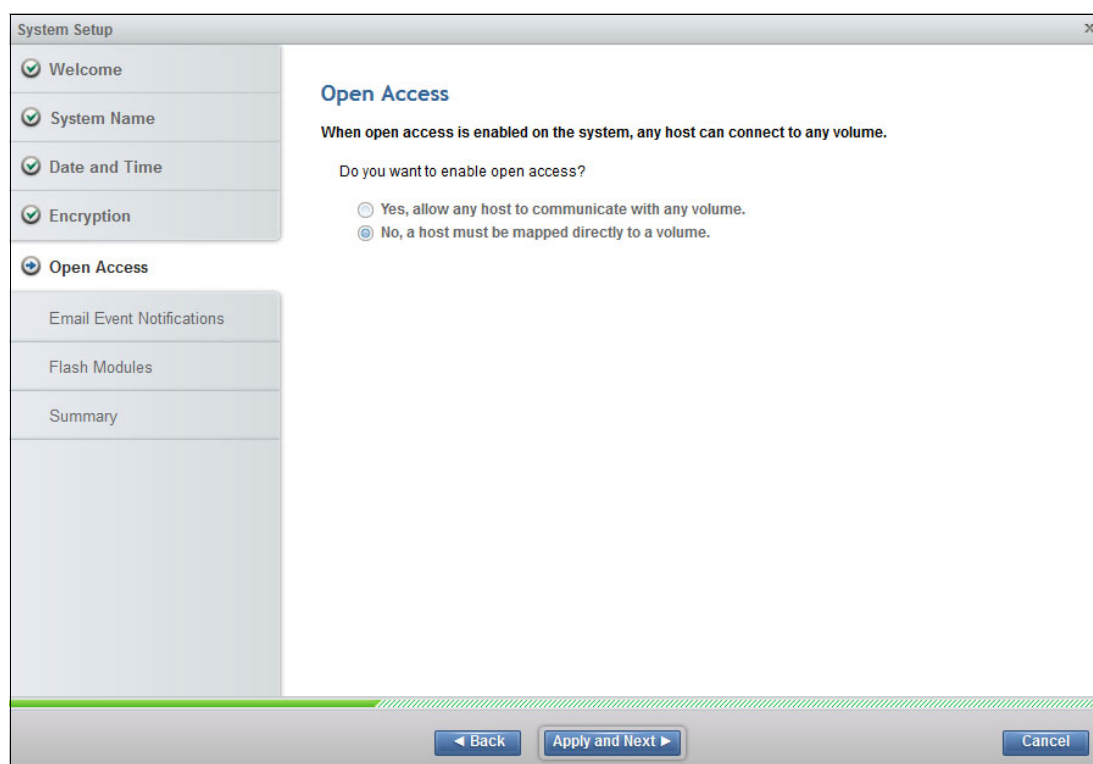


Figure 4-22 Initialization procedure denying open access



### Configure the call home feature

Next, configure the call home feature. This feature from the IBM FlashSystem 900 is invoked by sending an email with messages and alerts to IBM Support.

The following information must be provided to set up call home correctly:

- ▶ System Location: Where is the system physically installed?
- ▶ Contact Details: Who is the IBM Support contact for issues that require attention:
  - Name
  - Phone
  - Email
- ▶ Email server or servers (SMTP server):
  - IP address
  - Port (defaults to port 25)
- ▶ Event notification: Who else is to be notified of alerts, warnings, and errors?
- ▶ Summary: Confirm the settings.

Figure 4-23 shows the Email Event Notifications page. (The call home option can be configured after the initial configuration by using the menu **Security** → **Notifications**, which is demonstrated in 7.1.2, “Notifications menu” on page 233.)

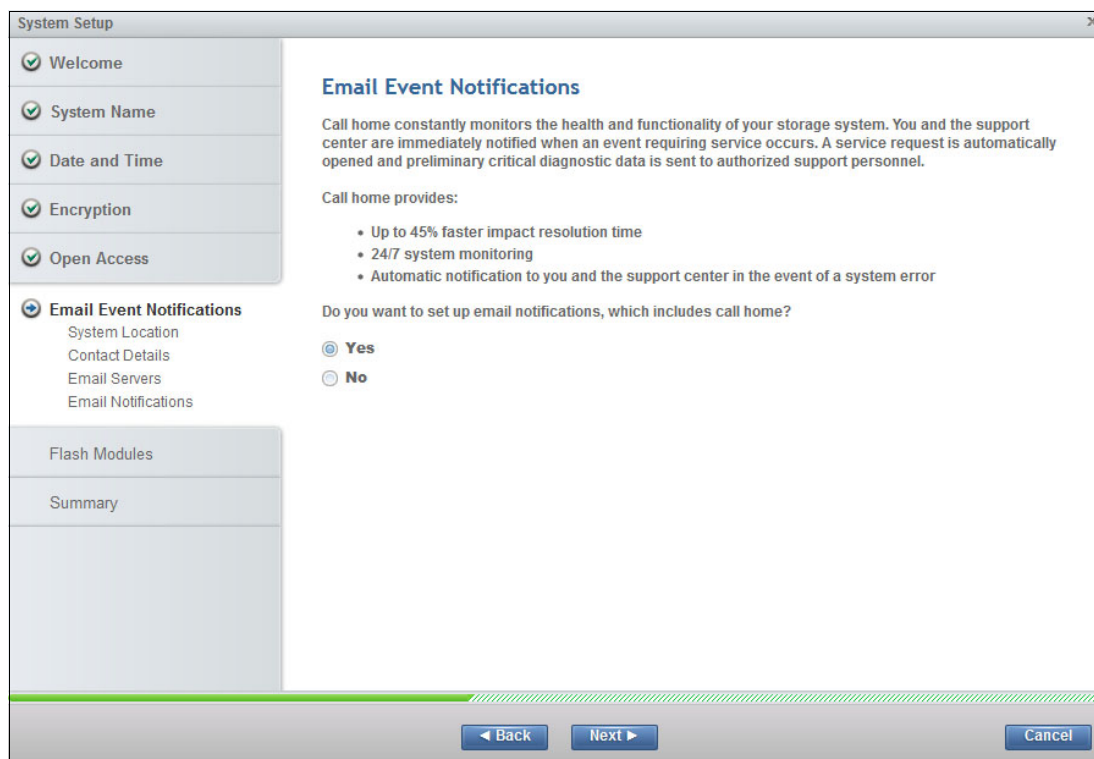


Figure 4-23 Initialization procedure: Call home

If you select **No** (at the question Do you want to set up email notifications, which includes call home?), a warning message is displayed (Figure 4-24 on page 94).

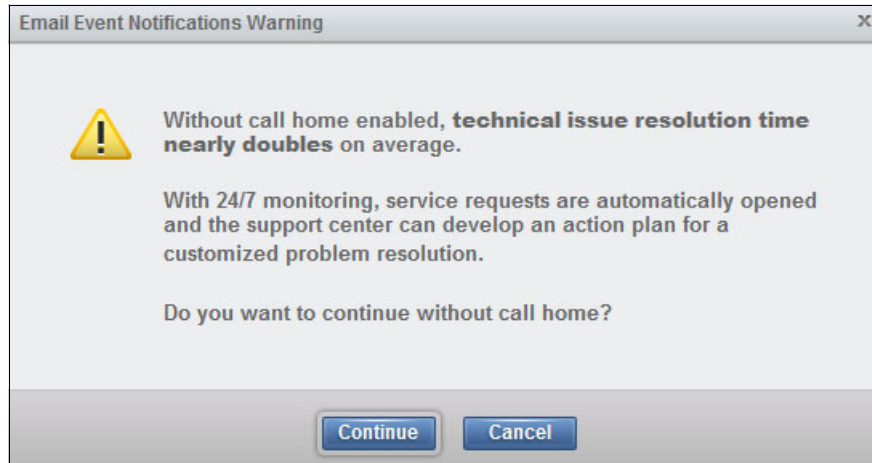


Figure 4-24 Initialization procedure: Call home warning

### Flash modules

During system setup FlashSystem 900 discovers the quantity and size of flash modules installed in the system. At the final step of the system setup procedure a single RAID 5 array is created on these flash modules. One flash module is reserved and acts as an active spare while the remaining modules forms the RAID 5 array.

Figure 4-25 shows that the system discovered four online flash modules.

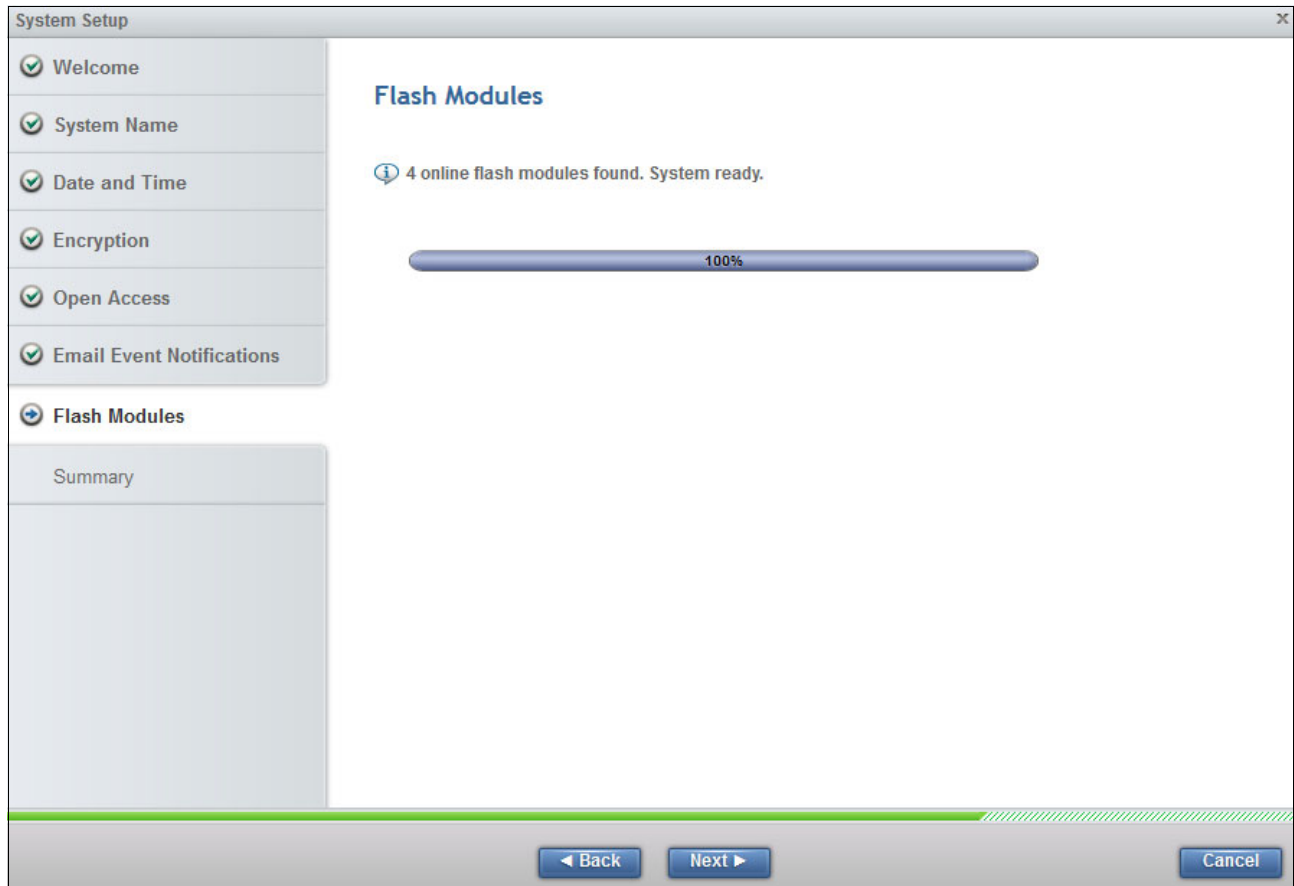


Figure 4-25 Initialization procedure: Flash Modules

With four flash modules, the usable capacity consists of the capacity of two modules because one is used for spare and the other for RAID 5 protection.

Finally, when all system setup and call home details are configured, the System Setup wizard shows a summary of the configuration (Figure 4-26). Click **Finish** to complete the System Setup wizard.

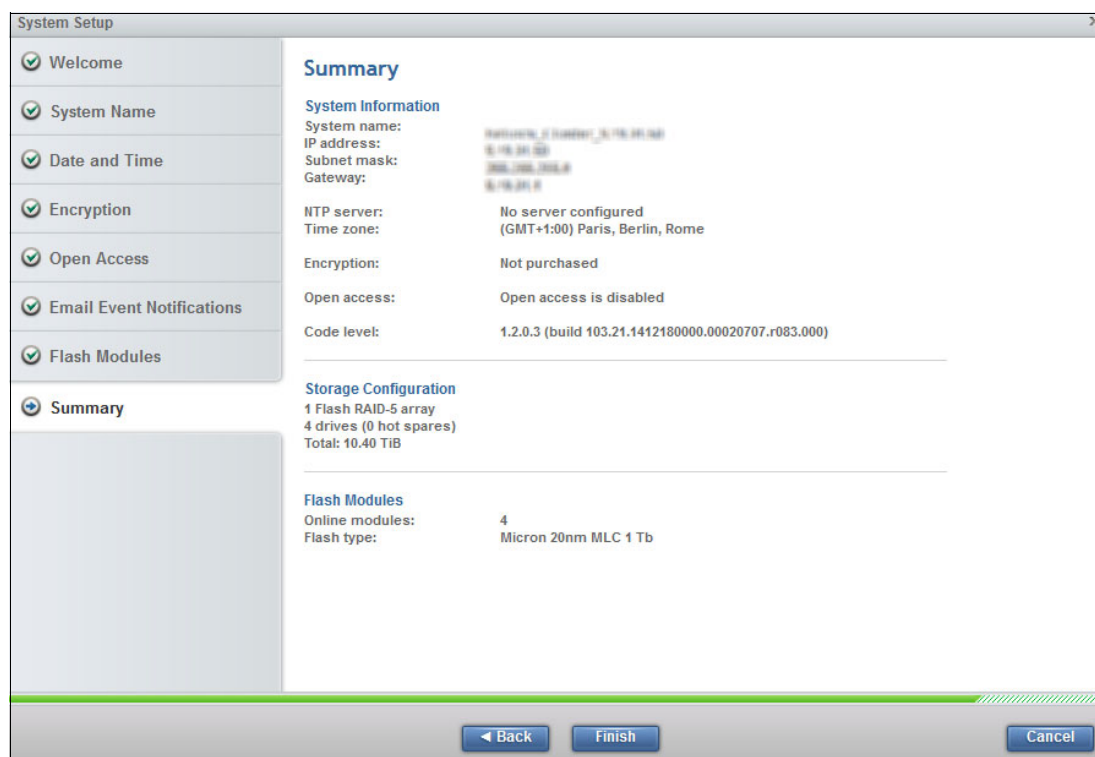


Figure 4-26 Initialization procedure: Summary

### ***Initialization complete***

The IBM FlashSystem 900 is now initialized and ready to use. The web browser automatically opens to the main page of the management GUI (Figure 4-27).

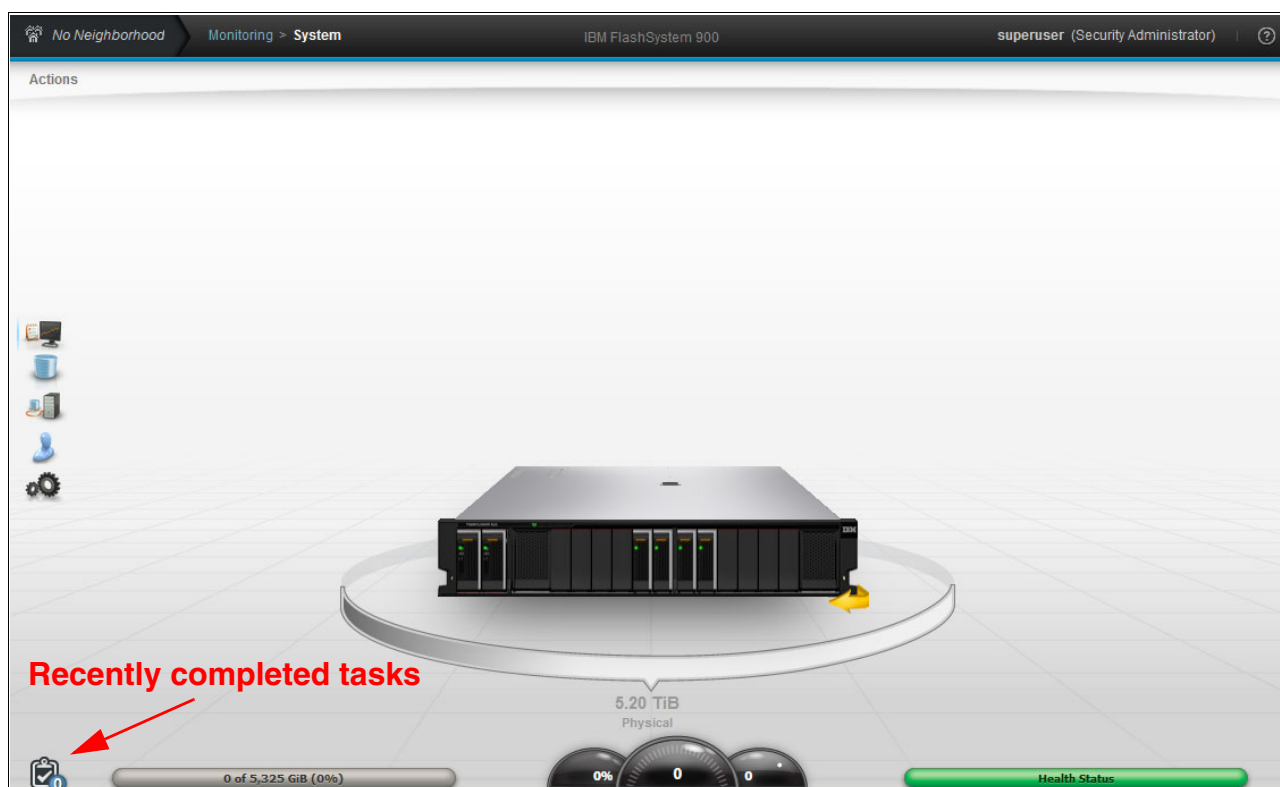


Figure 4-27 IBM FlashSystem 900 GUI main page

For more information about how to use the FlashSystem 900 GUI, see Chapter 6, “Using IBM FlashSystem 900” on page 161.

The icon in the lower left corner indicates recently completed tasks. Click it to see that the FlashSystem 900 initialized array0 (Figure 4-28 on page 97).

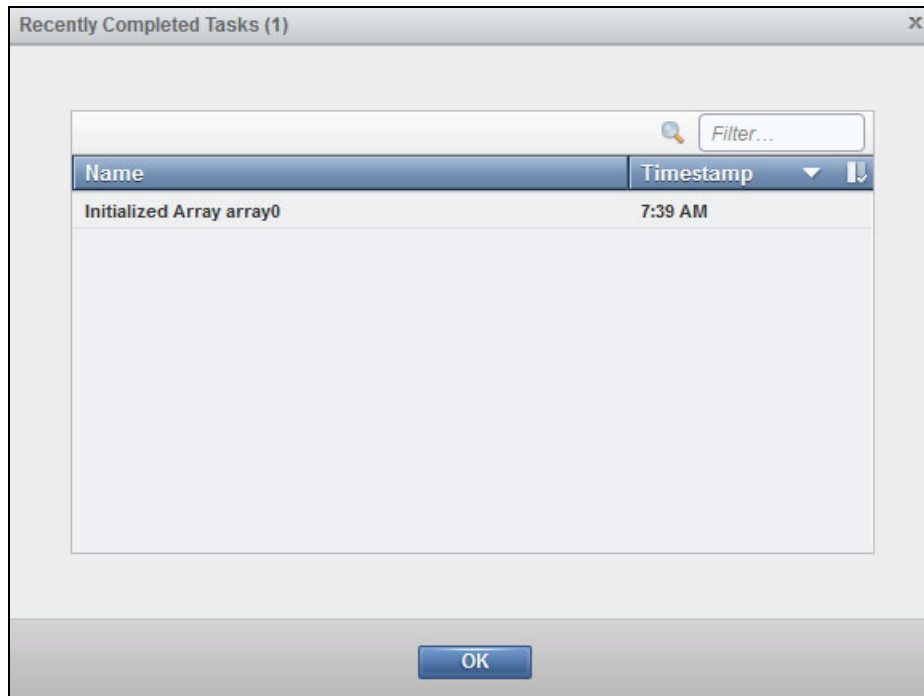


Figure 4-28 Recently completed tasks: array0 created

### 4.3.3 Service Assistant Tool

IBM FlashSystem 900 Service Assistant Tool can be used in various service event cases. Service Assistant Tool is normally only used in cases where the client is instructed to use it by IBM Support because Service Assistant Tool contains destructive and disruptive functions.

**Note:** Only the superuser account is allowed to log on to the Service Assistant Tool.

The following examples show the functions of IBM FlashSystem 900 Service Assistant Tool:

- ▶ Review installed hardware and firmware
- ▶ Review Ethernet ports and IP addresses
- ▶ Review worldwide names (WWNs)
- ▶ Change WWNs
- ▶ Canister enters the Service state
- ▶ Canister reboot
- ▶ Collect logs
- ▶ Reinstall software
- ▶ Configure CLI access
- ▶ Restart web service
- ▶ Recover system

There are two ways to open Service Assistant Tool:

- ▶ Point your web browser directly to the service IP address of each canister.
- ▶ Point your web browser to the management IP address of your IBM FlashSystem 900 and to specify service in the URL:

`https://192.168.10.10/service`

**Opening the tool:** Service Assistant Tool has destructive and disruptive functions. Open the Service Assistant Tool *only when you are instructed to do so by IBM Support*.

## 4.4 RAID storage modes

The process to implement the various RAID protection technologies that are used in the IBM FlashSystem 900 is described in 2.2, “Architecture of IBM FlashSystem 900” on page 24.

### Variable Stripe RAID

In addition to RAID 5, which protects the IBM FlashSystem 900 against an entire flash module failure, the IBM FlashSystem 900 also supports Variable Stripe RAID.

Variable Stripe RAID is a built-in data technology that protects data from subcomponent failure on a flash module.

The subcomponent failure Variable Stripe RAID protects against a *plane*. When a bad plane occurs, Variable Stripe RAID technology allows a plane to be removed from use without affecting the available capacity of other devices within the RAID stripe. Upon detection of a failure, the failing plane is removed from use (no further writes are allowed) and all used pages within the affected stripe are marked as “critical to move.” Information from the affected stripe is then gradually relocated to stripes that are known to be good, which is a process that is performed as a background task to minimize the required processing power.

How Variable Stripe RAID works is described in greater detail in 1.6, “Variable Stripe RAID” on page 10.

## 4.5 Connectivity guidelines for improved performance

You can configure the various network connections to improve the overall performance of the IBM FlashSystem 900. Considerations for planning the installation of the storage system are described next.

### 4.5.1 Interface card configuration guidelines

You can improve reliability and performance by following specific network connection guidelines for the interface cards in the enclosure.

Figure 4-29 on page 99 shows the rear side of the IBM FlashSystem 900 when installed with four FC or FCoE interface cards.

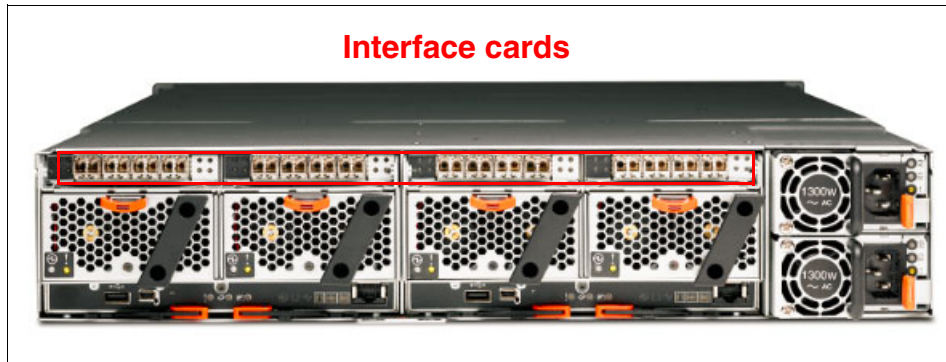


Figure 4-29 Interface cards

The FlashSystem 900 enclosure includes two canisters, each containing two FC, FCoE, iSCSI, or QDR InfiniBand interface cards. To eliminate any single failure point, use multipathing to span across interface cards in both canisters. Most major operating systems incorporate a multipathing option.

The storage system allows the flexible presentation of volumes, which are also called *logical unit numbers* (LUNs), through the interface cards. A volume or LUN can be presented and accessed through every port simultaneously. The loss of one interface card does not affect the I/O performance of the other interface cards. This configuration is commonly referred to as *active/active* or *active/active symmetric multipathing*.

**Note:** The IBM FlashSystem 900 interface cards are equipped with either two 16 Gbps FC, four 8 Gbps FC, four 10 Gbps FCoE, four 10 Gbps iSCSI, or four QDR InfiniBand ports.

Each canister mounts with two interface cards. Although having up to 16 ports on a single IBM FlashSystem 900 is possible, it is not a cabling or zoning requirement that host communication be established through all available ports. Therefore, an attached host can connect to fewer than all four ports on the interface cards.

## 4.5.2 Host adapter guidelines

To improve bandwidth, install dual-port host bus adapters and host channel adapters (HCAs) in the servers that are used with the storage system.

Dual-port HBAs and HCAs provide more ports for aggregating bandwidth. In an ideal case, multiple dual-port HBAs and HCAs are installed in each server for redundancy.

QDR InfiniBand HCAs maximize the bandwidth to the storage system. Most QDR InfiniBand HCAs use PCI Express 2.0, instead of version 1.0, to allow for the higher-rated bandwidth. Servers with PCI Express 2.0, or later, expansion slots provide the highest bandwidth.

The storage system is tested for interoperability against all major HBA and HCA vendors.

To verify valid and supported configurations, see the SSIC web page:

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

### 4.5.3 Cabling guidelines

As described in 4.2, “Cabling the system” on page 71, several considerations exist that are related to the correct cabling.

In addition, design cabling to provide resiliency across storage system management ports and FC, FCoE, iSCSI, or QDR InfiniBand interface card ports. Also, consider the related network switches and server HBA and HCA ports, and create a design that provides redundancy and has enough paths for optimal performance.

All high-availability concepts of a dual-switched fabric setup apply to the storage system. One key element to recognize when cabling the storage system is the use of available paths. The storage system is designed to deliver I/O throughput through all connected ports. Use all ports if the application can benefit from more bandwidth. To take advantage of these ports on the storage system, there must be an equal number of server ports. Otherwise, there are underutilized ports on one side of the fabric.

### 4.5.4 Zoning guidelines

In a switched fabric environment, implementing zoning when you connect the storage system to the network can improve performance and simplify upgrades.

In a switched fabric deployment, a common approach is to isolate one application’s server storage devices from server storage devices of other applications. This practice prevents cross-traffic and helps ease maintenance. Therefore, employ zoning in all multiple server environments.

Zoning a server’s HBA or HCA is best deployed when only a single HBA or HCA is contained in a zone. This approach is also referred to as *Single Initiator Zoning*. Having more than one HBA or HCA in a zone can lead to unexpected errors and can cause unplanned downtime. HBAs or HCAs are called *initiators*, and these initiators can be zoned to a single or multiple *target* ports. A target port is a storage controller port, such as the ports in the IBM FlashSystem 900.

The host initiators must be zoned to sufficient storage target ports to carry the expected workload. Host initiator ports must however not be zoned to more than the necessary number of storage target ports. Zoning host initiator ports to more than the necessary number of storage target ports can result in excessive paths to storage, which can decrease performance.

**Note:** When you connect the storage system to a dual-switched fabric architecture, ensure that each zoned server uses interface cards in both canisters in the storage system. Ensure that each zoned server does not depend on a single interface card or a single canister for availability.

For information about how to zone an IBM FlashSystem to IBM SAN Volume Controller, see 8.2, “SAN Volume Controller connectivity to FlashSystem 900” on page 290.

For information about how to zone storage devices in general to IBM Spectrum Virtualize, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.





# IBM FlashSystem 900 client host attachment and implementation

This chapter has installation, implementation, and other general information and guidelines for connecting client host systems to the IBM FlashSystem 900 (FlashSystem 900).

This chapter covers the following topics:

- ▶ Host implementation and procedures
- ▶ Host connectivity
- ▶ Operating system connectivity and preferred practices
- ▶ Miscellaneous host attachment
- ▶ FlashSystem 900 preferred read and configuration examples
- ▶ FlashSystem 900 and Easy Tier
- ▶ Troubleshooting

The chapter provides this information:

- ▶ FlashSystem 900 sector size and host block size considerations
- ▶ Partition and file alignment for best performance
- ▶ Multipath support implementation for various host operating systems
- ▶ Necessary drivers for several operating systems
- ▶ Host integration for various operating systems

## Notes about IBM SAN Volume Controller and IBM Spectrum Virtualize:

- ▶ Some of the following sections mention IBM SAN Volume Controller, which delivers the functions of IBM Spectrum Virtualize, part of the IBM Spectrum Storage family.
- ▶ IBM Spectrum Virtualize is industry-leading storage virtualization that enhances existing storage to improve resource utilization and productivity so you can achieve a simpler, more scalable and cost-efficient IT infrastructure. The functionality of IBM Spectrum Virtualize is provided by IBM SAN Volume Controller.
- ▶ For more details, see the following web page:

<http://www.ibm.com/systems/storage/software/virtualization/svc/index.html>

## 5.1 Host implementation and procedures

The procedures to connect the IBM FlashSystem 900 to client hosts using various operating systems are described in the following sections.

## 5.2 Host connectivity

The IBM FlashSystem 900 can be attached to a client host by four methods:

- ▶ Fibre Channel (FC)
- ▶ Fibre Channel over Ethernet (FCoE)
- ▶ InfiniBand
- ▶ iSCSI

Always check the IBM System Storage Interoperation Center (SSIC) to get the latest information about supported operating systems, hosts, switches, and more:

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

If a configuration that you want is not available on the SSIC, request approval from IBM by submitting a Solution for Compliance in a Regulated Environment (SCORE) or request for price quotation (RPQ). To submit a SCORE/RPQ, contact your IBM FlashSystem marketing representative or IBM Business Partner.

The IBM FlashSystem 900 can be SAN-attached by using a switch or can be directly attached to a client host. Check the IBM SSIC for specific details. Several operating system and Fibre Channel (FC) driver combinations allow point-to-point direct access with 16 Gbps FC. Check your environment and check the SSIC to use 16 Gbps direct attachment to the host.

**Note:** The FlashSystem 900 16 Gbps FC attachment does not support arbitrated loop topology (direct connection to client hosts). The IBM FlashSystem 900 must be connected to a SAN switch when using 16 Gbps FC if the host operating system does not support point-to-point FC direct connections. At the time that this book was written, IBM AIX did not support point-to-point FC direct connections.

### 5.2.1 Fibre Channel SAN attachment

If you attach a host using a SAN switch to the FlashSystem 900, ensure that each host port is connected and zoned to both canisters of the FlashSystem 900. If only one FlashSystem 900 canister is connected to a host port, the host state will be shown as *degraded*. This will be referred to in the remainder of this chapter as the *switch rule*.

**Note:** When you use a switch, you must zone host ports according to the switch rule. With the exception of IBM i, a host port must be connected to each FlashSystem 900 canister. For IBM i, different switch zoning rules apply as described in 5.3.4, “IBM i and FlashSystem 900” on page 109.

Figure 5-1 on page 103 shows the correct SAN connection of an AIX server with two ports to the FlashSystem 900. In this example, four zones are set up:

- ▶ AIX port 8a FlashSystem 900 port 41
- ▶ AIX port 8a FlashSystem 900 port 61

- ▶ AIX port 27 FlashSystem 900 port 51
- ▶ AIX port 27 FlashSystem 900 port 71

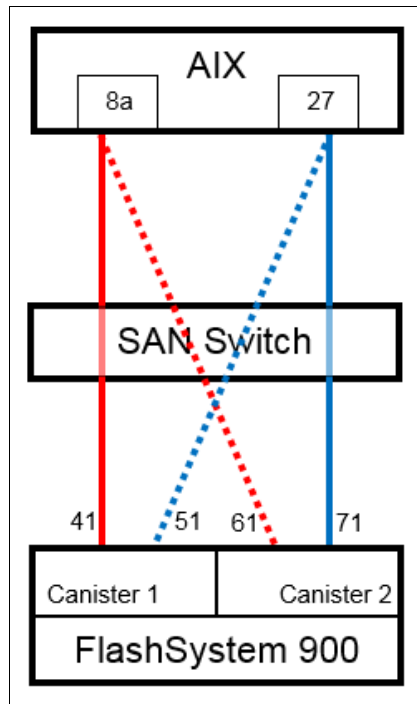


Figure 5-1 SAN attachment

## 5.2.2 Fibre Channel direct attachment

If you attach the FlashSystem 900 directly to a host, the host must be attached to both canisters. If the host is not attached to both canisters, the host will be shown as *degraded*.

Figure 5-2 shows the correct direct attachment of an AIX server with two ports to the FlashSystem 900. This example shows two connections:

- ▶ AIX port 8a directly attached to FlashSystem 900 port 41
- ▶ AIX port 27 directly attached to FlashSystem 900 port 71

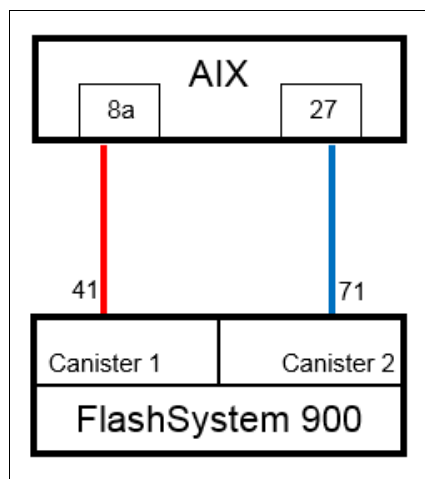


Figure 5-2 Direct attachment

If you use SAN attachment and direct attachment simultaneously on a FlashSystem 900, the direct-attached host state will be *degraded*. Using a switch will enforce the *switch rule* for all attached hosts, which means that a host port has to be connected to both FlashSystem canisters. Because a direct-attached host cannot connect one port to both canisters, it will not meet the *switch rule* and its state will be *degraded*.

**Note:** You can attach a host through a switch and simultaneously attach a host directly to the FlashSystem 900. But then, the direct-attached host will be shown as *degraded*.

### 5.2.3 General Fibre Channel attachment rules

These rules apply to FC connections:

- ▶ If directly attached, a host must have ports connected to both canisters.
- ▶ If connected to a switch, all host ports must have paths to both canisters. This is the switch rule.
- ▶ If any port is connected to a switch, the switch rule applies to all hosts except IBM i, regardless of whether that host is connected through a switch.

## 5.3 Operating system connectivity and preferred practices

Detailed information about the IBM FlashSystem 900 client host connections using various operating systems is described in the following sections.

### 5.3.1 FlashSystem 900 sector size

In a traditional spinning disk, a *sector* refers to a physical part of the disk. The size of the sector is defined by the disk manufacturer and most often set to 512 bytes. The 512-byte sector size is supported by most operating systems.

The FlashSystem 900 does not have fixed physical sectors like spinning disks do. Data is written in the most effective way on flash. To maintain compatibility however, the sector size is the same as the sector size of most traditional spinning disk sectors. Therefore, the default sector size is presented to the host by the FlashSystem 900 is 512 bytes. Starting with firmware version 1.1.3.0, you can create volumes with a sector size of 4096 bytes by using the CLI **mkvdisk** command and the new **-blocksize** parameter. For details about this command, see the FlashSystem 900 web page in the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

The **mkvdisk** command **-blocksize** parameter specifies the SCSI logical unit sector size. The only two possible values are 512 (the default) and 4096:

- ▶ Size 512 is the default. It is supported by most operating systems.
- ▶ Size 4096 provides better performance but it might not be supported by your host operating system or application.

**Note:** Format all client host file systems on the storage system at 4 KB or at a multiple of 4 KB. Do this for a used sector size of 512 and 4096 bytes. For example, file systems that are formatted at an 8 KB allocation size or a 64 KB allocation size are satisfactory because they are a multiple of 4 KB.

### 5.3.2 File alignment for the best RAID performance

File system alignment can improve performance for storage systems by using a RAID storage mode. File system alignment is a technique that matches file system I/O requests with important block boundaries in the physical storage system. Alignment is important in any system that implements a RAID layout. I/O requests that fall within the boundaries of a single stripe have better performance than an I/O request that affects multiple stripes. When an I/O request crosses the endpoint of one stripe and into another stripe, the controller must then modify both stripes to maintain their consistency.

Unaligned accesses include those requests that start at an address that is not divisible by 4 KB, or are not a multiple of 4 KB in size. These unaligned accesses are serviced at much higher response times, and they can also significantly reduce the performance of aligned accesses that were issued in parallel.

The IBM FlashSystem 900 provides 512-byte sector size support that greatly improves response times for I/O requests that cannot be forcibly aligned. However, alignment to 4 KB must be maintained whenever possible.

### 5.3.3 IBM AIX and FlashSystem 900

The IBM FlashSystem 900 can be attached to AIX client hosts by using the following Fibre Channel (FC) method:

- The IBM FlashSystem 900 connects to AIX through Node Port Identifier Virtualization (NPIV) and Virtual I/O Server (VIOS) modes.

#### Directly attached Fibre Channel topology for AIX

Configure the FlashSystem 900 FC controllers to arbitrated loop topology when the controllers are directly attached to the AIX hosts. You must check the SSIC for supported configurations. For more details about SSIC, see 5.2, “Host connectivity” on page 102.

**Note:** The FlashSystem 900 16 Gbps FC ports do not support direct connection to AIX client hosts. A SAN switch must be placed between the IBM FlashSystem 900 and any 16 Gbps-attached client host. If arbitrated loop is required by the client host, connect at 8 Gbps FC to the IBM FlashSystem 900.

#### Optimal logical unit number configurations for AIX

The number of logical unit numbers (LUNs) that you create on the IBM FlashSystem 900 can affect the overall performance of AIX.

Applications perform optimally if at least 32 LUNs are used in a volume group. If fewer volumes are required by an application, use the Logical Volume Manager (LVM) to map fewer logical volumes to 32 logical units. This does not affect performance in any significant manner (LVM overhead is small).

**Note:** Use at least 32 LUNs in a volume group because this number is the best balance between good performance (the more queued I/Os the better FlashSystem 900 performs) and minimizing overhead and complexity.

#### Sector size restrictions for AIX

The AIX operating system supports the 512-byte sector size, which the IBM FlashSystem 900 supports.

## Auto Contingent Allegiance support

Certain host systems require the Auto Contingent Allegiance (ACA) support to run multiple concurrent commands. When using the round-robin multipathing algorithm, IBM AIX sends out extraneous ACA task management commands. ACA support on logical units is always enabled on the IBM FlashSystem 900.

## Volume alignment

The IBM AIX operating system volumes align to 4 KB boundaries.

## Implementing multipathing for IBM AIX hosts

*Multipathing* enables the host to access the FlashSystem 900 LUNs through different paths. This architecture helps to protect against I/O failures, such as port, cable, or other path issues.

**Important:** The latest updates for multipathing support on the IBM AIX operating system, are at IBM Fix Central:

<http://www.ibm.com/support/fixcentral>

## Resetting the host bus adapter and disk configuration

The following sections describe how to reconfigure the host bus adapters (HBAs) to implement multipathing. After you install the latest IBM AIX updates for support of the IBM FlashSystem 900, AIX must rescan the SCSI bus for the LUNs to recognize them as devices that support multipathing. Begin by reconfiguring the HBA and its attached disks.

To reset the HBA or disk configuration, complete the following steps:

**Important:** If other disks are attached to any HBA devices, the following commands remove the configuration for those disks and the HBA. If you are attempting to save the current configuration, skip these steps.

1. Determine the device names of the HBAs to which the storage system is connected, by entering this command:  

```
lsdev -t efscsi
```
2. For each HBA device name, enter the following command to remove the HBA and the disk configuration that is associated with it:  

```
rmdev -l <device name> -R
```
3. Determine whether any disks are already defined that must be removed before rescanning, by entering this command:  

```
lsdev -C -c disk
```
4. If any LUNs are already defined as Other FC SCSI Disk Drive, remove the old definitions. For each disk name, enter the following command:  

```
rmdev -l <disk name> -R
```

## Setting the fast fail recovery flag for the host bus adapter

You can set the fast fail recovery flag for the HBA to improve the failover response.

For the multipath I/O (MPIO) driver to fail over to an available path in a timely manner after a path failure, set the **fast\_fail** recovery flag for the HBA devices to which the storage system is connected.

At a command prompt, enter this command:

```
chdev -a fc_err_recov=fast_fail -l <device name>
```

The <device name> is the device name of the HBA that is connected to the system.

### ***Rescanning for the storage system logical unit numbers***

After the host system is configured to recognize that the storage device supports multipathing, you must rescan for the LUNs.

At a command prompt, enter this command:

```
cfgmgr -vl <device name>
```

The <device name> is the device name of the HBA connected to the system.

### ***Confirming the configuration***

After you change the configuration to support multipathing, confirm that the configuration is working correctly.

To confirm the new configuration, complete the following steps:

1. Ensure that the configuration is successful by entering the following command to list all disks available to the system:

```
lsdev -C -c disk
```

All LUNs must use MPIO. They must show as MPIO IBM FlashSystem Disk.

The following command gives you detailed information about a LUN:

```
lscfg -vl <LUN>
```

Example 5-1 on page 108 shows the output of those two commands. This AIX system has four disks attached:

- FlashSystem 820 LUN using MPIO
- A LUN without multipathing
- Another LUN without multipathing
- FlashSystem 900 LUN using MPIO

Both IBM FlashSystem units are shown as MPIO IBM FlashSystem Disks. But, you see the different models when you look at the Machine Type and Model attribute of the **lscfg** command output, for example, the fourth LUN is a FlashSystem 900.

Some output lines were removed for clarity (Example 5-1 on page 108).

#### *Example 5-1 Check the AIX MPIO configuration*

---

```
# lsdev -C -c disk
hdisk0 Available          Virtual SCSI Disk Drive
hdisk1 Available 00-00-02 MPIO IBM FlashSystem Disk
hdisk2 Available 00-01-02 Other FC SCSI Disk Drive
hdisk3 Available 01-01-02 Other FC SCSI Disk Drive
hdisk4 Available 00-01-02 MPIO IBM FlashSystem Disk

# lscfg -vl hdisk1
hdisk1      U78C0.001.DBJ2497-P2-C1-T1-W20040020C2117377-L0  MPIO IBM FlashSystem Disk

      Manufacturer.....IBM
      Machine Type and Model.....FlashSystem
      ...

# lscfg -vl hdisk4
hdisk4      U78C0.001.DBJ2497-P2-C1-T2-W500507605EFE0AD1-L0  MPIO IBM FlashSystem Disk

      Manufacturer.....IBM
      Machine Type and Model.....FlashSystem-9840
      ...
```

---

2. If disks are missing, or are extra, or the LUNs do not show as an MPIO IBM FlashSystem Disk, check that the connections and the storage system configuration are correct. You must then remove the configuration for the HBAs and complete the rescan again. For more information, see “Resetting the host bus adapter and disk configuration” on page 106.
3. To ensure that all the connected paths are visible, enter the following command:  
`lspath`

Example 5-2 shows the paths for the IBM FlashSystem 900 used in Example 5-1.

#### *Example 5-2 AIX lspath output*

---

```
# lspath -l hdisk4
Enabled hdisk4 fscsi1
Enabled hdisk4 fscsi1
Enabled hdisk4 fscsi3
Enabled hdisk4 fscsi3
```

---

4. If paths are missing, check that the connections and the storage system configuration are correct. You must then remove the configuration for the HBAs and perform the rescan again. For more information, see “Resetting the host bus adapter and disk configuration” on page 106.

### **Configuring path settings**

All paths on the IBM FlashSystem 900 are equal. All ports have access to the LUNs, and there is no prioritized port. Therefore, you use them all at the same time. You have to set the distribution of the I/O load at the operating system level. The round-robin distribution is the ideal way to use all of the ports equally.

Set the algorithm attribute to `round_robin` before you add the hdisk to any volume group. All outgoing traffic is then spread evenly across all of the ports, as shown in the following example:

```
chdev -l <LUN> -a algorithm=round_robin
```



The `shortest_queue` algorithm is available in the latest technology levels of AIX for some devices. The algorithm behaves similarly to `round_robin` when the load is light. When the load increases, this algorithm favors the path that has the fewest active I/O operations. Therefore, if one path is slow because of congestion in the SAN, the other less congested paths are used for more of the I/O operations. The `shortest_queue` (if available) or `round_robin` enables the maximum use of the SAN resources. You can use the `load_balance` algorithm to spread the load equally across the paths.

To list the attributes of the LUN, enter the following command:

```
lsattr -El <LUN>
```

Example 5-3 shows output of the `chdev` and the `lsattr` commands of the FlashSystem 900 that is used in Example 5-1 on page 108. The number of spaces in the output is changed for clarity.

*Example 5-3 AIX `chdev` and `lsattr` commands*

---

```
# chdev -l hdisk4 -a algorithm=round_robin
hdisk4 changed

# lsattr -El hdisk4
PCM                PCM/friend/fcpothe          Path Control Module          False
PR_key_value       none                        Persistent Reserve Key Value True+
algorithm          round_robin                Algorithm                     True+
clr_q              no                          Device CLEARS its Queue on error True
dist_err_pcmt      0                          Distributed Error Percentage  True
dist_tw_width      50                         Distributed Error Sample Time True
hcheck_cmd         test_unit_rdy              Health Check Command          True+
hcheck_interval    60                         Health Check Interval         True+
hcheck_mode        nonactive                  Health Check Mode             True+
location           Location                    Label                          True+
lun_id             0x0                        Logical Unit Number ID        False
lun_reset_spt      yes                        LUN Reset Supported           True
max_coalesce       0x40000                    Maximum Coalesce Size         True
max_retry_delay    60                         Maximum Quiesce Time          True
max_transfer       0x80000                    Maximum TRANSFER Size         True
node_name          0x500507605efe0ad0         FC Node Name                  False
pvid               none                        Physical volume identifier    False
q_err              yes                         Use QERR bit                  True
q_type             simple                      Queuing TYPE                  True
queue_depth        64                         Queue DEPTH                   True
reassign_to        120                        REASSIGN time out value       True
reserve_policy     no_reserve                 Reserve Policy                 True+
rw_timeout         30                         READ/WRITE time out value     True
scsi_id            0x10100                    SCSI ID                       False
start_timeout      60                         START unit time out value     True
timeout_policy     fail_path                  Timeout Policy                 True+
unique_id          54361IBM      FlashSystem-9840041263a20412-0000-0004-00006410FlashSystem-984003IBMfcp
ww_name            0x500507605e800e41         Unique device identifier      False
                   FC World Wide Name          False
```

---

### 5.3.4 IBM i and FlashSystem 900

This section covers the following topics pertaining to IBM i and FlashSystem 900:

- ▶ Attachment methods
- ▶ Zoning rules
- ▶ Hardware and software requirements
- ▶ Configuration and performance considerations

## Attachment methods

The IBM FlashSystem 900 can be attached to IBM i hosts by using one of the following Fibre Channel (FC) methods:

- ▶ *NPIV-attachment* through the IBM Virtual I/O Server (VIOS) using Node Port Identifier Virtualization (NPIV) in conjunction with the required NPIV-capable SAN switch.
- ▶ *Native-attachment*, that is, without the IBM VIOS, either with a SAN switch or without one. The latter (without a SAN switch) is also known as *direct-attachment*.

**Note:** IBM i attachment to the IBM FlashSystem 900 through the IBM VIOS using virtual SCSI is *not* supported.

## Zoning rules

The same rules apply for IBM i direct-attachment to an IBM FlashSystem 900 as for other host operating systems. However, different zoning rules apply for SAN switch-attached IBM i host Fibre Channel initiators. Unlike with other operating systems, these initiators must not be zoned to both FlashSystem canisters.

For IBM FlashSystem 900 native or NPIV SAN switch attachment, a one-to-one zoning should be used such that one IBM i Fibre Channel initiator port is zoned with one FlashSystem target port from a single FlashSystem canister, as shown in Figure 5-3 on the right. Note that this is different also to IBM FlashSystem V9000, IBM Storwize series, or IBM SAN Volume Controller attachment where a single IBM i initiator should be zoned to both storage controller nodes to support the SCSI Asymmetrical Logical Unit Access (ALUA) LUN affinity concept of these controller nodes with preferred (active) and non-preferred (passive) paths. With the IBM FlashSystem 900, all paths are active.

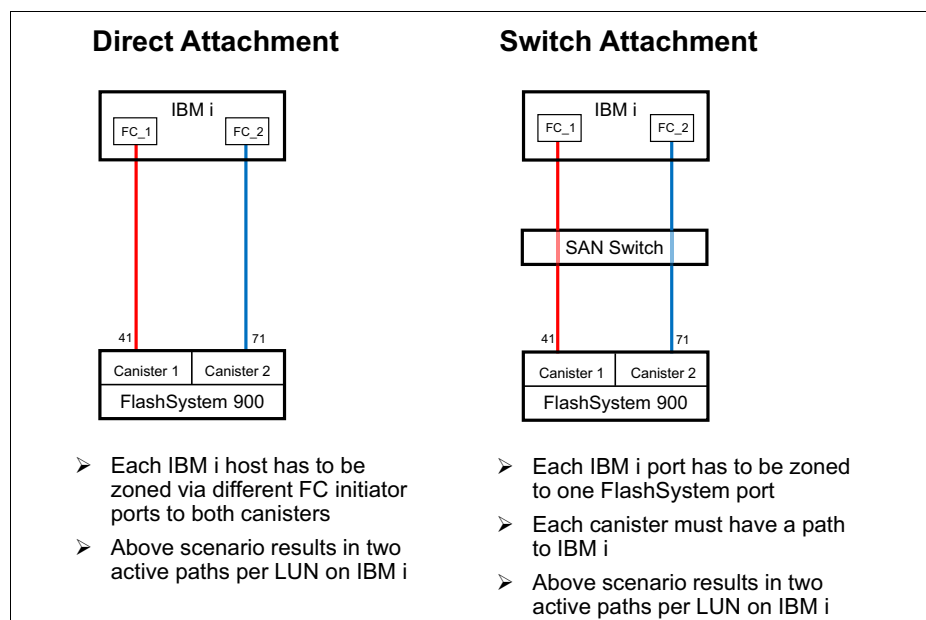


Figure 5-3 IBM i FlashSystem 900 attachment and zoning rules

## Hardware and software requirements

The minimum hardware and software requirements for IBM i attachment to the IBM FlashSystem 900 are summarized in Table 5-1 on page 111. For further interoperability information, see the IBM System Storage Interoperation Center (SSIC):

<https://www.ibm.com/systems/support/storage/ssic/interoperability.wss>

Table 5-1 IBM i and IBM FlashSystem 900 Minimum Requirements

Requirement	Requirement details and versions
IBM i version and release	IBM i 7.2 Technology Refresh 2 or later + latest HIPER PTF group <sup>a</sup>
IBM Power Systems server	IBM POWER7® firmware level FW780_40 or later IBM POWER8® firmware level FW810 or later <sup>b</sup>
Attachment	VIOS NPIV Native using switches Native direct <sup>c</sup>
VIOS level	2.2.3.4 or later
Fibre Channel adapters and Fibre Channel over Ethernet adapters	Native connection or VIOS NPIV connection: <ul style="list-style-type: none"> <li>▶ 8 Gb 2-port FC adapter #5735/#5273</li> <li>▶ 8 Gb 2-port FC adapter #EN0G/#EN0F (VIOS NPIV only)</li> <li>▶ 8 Gb 4-port FC adapter #5729 (VIOS NPIV only)</li> <li>▶ 8 Gb 4-port FC adapter #EN12/#EN0Y (VIOS NPIV only)</li> <li>▶ 16 Gb 2-port FC adapter #EN0A/#EN0B</li> </ul> VIOS NPIV connection via FCoE: <ul style="list-style-type: none"> <li>▶ 2x 10 Gb FCoE 2x 1GbE SFP+ adapter #EN0H/#EN0J</li> <li>▶ 4-port (10 Gb FCoE and 1 GbE) LR and RJ45 adapter #EN0M/#EN0N</li> <li>▶ 4-port (10 Gb FCoE and 1 GbE) copper and RJ45 adapter #EN0K/#EN0L</li> </ul>
SAN switches	Brocade or Cisco
FlashSystem firmware	1.2.0.11 or later

a. See the IBM TechNote *PTF listing for 4096 disk sector support at 7.2 with D/TD840 or D/T6B4E-050 drives* for additional information about recommended PTFs:

<http://www.ibm.com/support/docview.wss?uid=nas8N1020957>

b. Attachment of IBM FlashSystem 900 to IBM i with Boot from SAN is supported only in Power Systems models that support the stated firmware levels

c. Direct attachment can be done with: 8 Gb ports in IBM i and FlashSystem – 8 Gb ports in FlashSystem must be configured as Fibre Channel Arbitrated Loop; 16 Gb ports in IBM i and FlashSystem

## Configuration and performance considerations

Any volume (LUN) configured and attached natively or through a VIOS NPIV connection to an IBM i partition must be created with 4096 byte sectors by using the FlashSystem CLI command **mkvdisk** with the parameter **-blocksize 4096** or by using the GUI version 1.3 or later which supports the 4096 byte block size as shown in Figure 5-4 on page 112. IBM i compresses its tagged pages as a minor subset of its 4160 byte pages to fit into the 4096 byte sector format supported by the IBM FlashSystem 900.

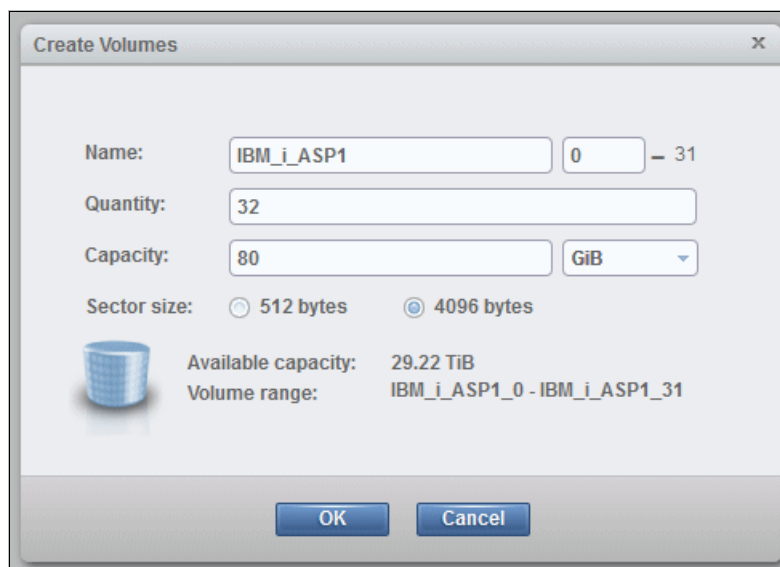


Figure 5-4 FlashSystem 900 GUI creating 4096 bytes sector LUNs

**Note:** The 512 byte sector LUNs on an IBM FlashSystem 900 are *not* supported by IBM i unless this FlashSystem is virtualized by an IBM SAN Volume Controller.

Because IBM FlashSystem 900 generally expects a host initiator to log into both of its canisters, which is not applicable to IBM i, the IBM i host is shown with a state of *degraded* by the FlashSystem as shown in Figure 5-5. This is not a failure indication in this case and should be ignored.

+ Add Host    Actions    Filter    [Icon]			
Name	State	Host Ports	Host Mappings
IBM i	<span style="border: 1px solid red; padding: 2px;">⚠ Degraded</span>	24	Yes

Figure 5-5 FlashSystem 900 GUI reported IBM i Host State

Similar to other SAN storage systems for the IBM FlashSystem 900, a moderate LUN size should be chosen for IBM i in the approximate range of 40 - 300 GB because IBM i uses a fixed queue depth per disk unit and path. I/O concurrency and performance typically benefit from having a reasonable number of LUNs configured in an IBM i auxiliary storage pool (ASP) especially for applications known to perform a lot of file creates, opens, and closes.

Up to 64 LUNs are supported per IBM i physical or virtual Fibre Channel adapter port.

For further details about IBM PowerVM® Virtual I/O Server planning and implementation including NPIV attachment for IBM i see *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

With IBM PowerVM Virtual I/O Server NPIV attachment, usually no storage performance tunable parameters are available because VIOS merely acts as a Fibre Channel I/O pass-through from its owned physical Fibre Channel adapter through the IBM Power Systems hypervisor to a VIOS client partition such as IBM i. VIOS does not even “see” the NPIV client LUNs and thus it does not perform I/O multi-pathing for them, which should be done by the IBM i client, preferably across two VIOS partitions as shown in Figure 5-6 on page 113.

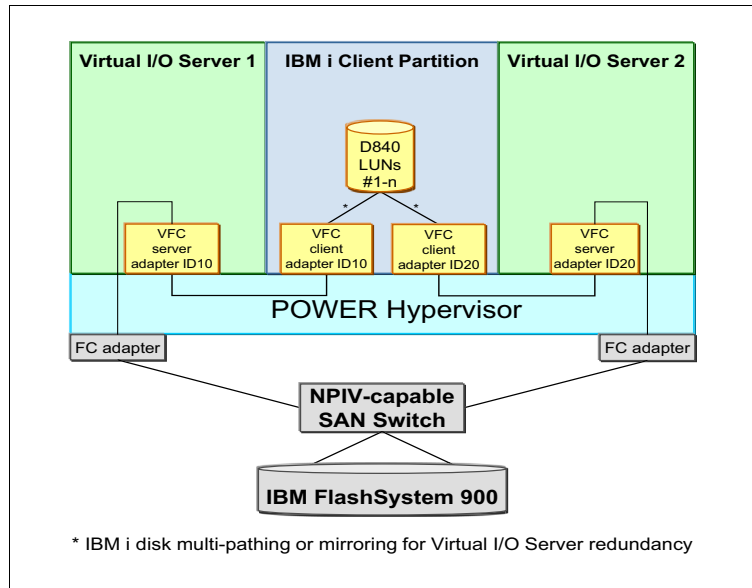


Figure 5-6 IBM i NPIV-attachment with two redundant Virtual I/O Servers

For VIOS attachment (NPIV-attachment), the Fibre Channel adapter queue depth is a storage performance-related tunable parameter. It can be considered to be increased if not already using VIOS version 2.2.4.10 or later and having deployed its default *rule set* for applying IBM recommended device settings.

The Fibre Channel adapter model-dependent current setting for its adapter queue depth, its allowed range, and its increase to its maximum supported value is shown in Example 5-4. With the fcsX FC adapter port resource usually in use, as implied in the example, VIOS would still need to be restarted for the permanent change of the adapter queue depth that is made to its resource database to become effective.

*Example 5-4 Displaying and changing the VIOS FC adapter queue depth*

---

```
$ lsdev -dev fcs0 -attr | grep num_cmd_elems
num_cmd_elems 500          Maximum number of COMMANDS to queue to the adapter True

$ lsdev -dev fcs0 -range num_cmd_elems
20...4096 (+1)

$ chdev -dev fcs0 -perm -attr num_cmd_elems=4096
fcs0 changed
```

---

For both native and VIOS NPIV-attachment, the IBM FlashSystem 900 LUNs report to IBM i as device type D840 disk units, as shown in Figure 5-7 on page 114.

Figure 5-7 on page 114 shows the disk configuration of a newly set up IBM i partition with its load source unit, that is, disk unit 1 being the only disk unit configured in the system ASP (ASP 1). All disk units are accessible from IBM i through two active Fibre Channel paths. The IBM i integrated multi-path driver distributes the I/O across all active paths of a disk unit using a round-robin algorithm with some load-balancing applied. Non-configured and non-initialized disk units, that is, those not assigned to an IBM i ASP yet, are still reported with DPHxxx resource names.

Display Disk Path Status						
ASP	Unit	Serial Number	Type	Model	Resource Name	Path Status
1	1	Y85FB50002E7	D840	040	DMP002	Active
		Y85FB50002E7	D840	040	DMP001	Active
*	*	Y85FB50002E8	D840	040	DPH001	Active
		Y85FB50002E8	D840	040	DPH010	Active
*	*	Y85FB50002E9	D840	040	DPH002	Active
		Y85FB50002E9	D840	040	DPH011	Active
*	*	Y85FB50002EA	D840	040	DPH003	Active
		Y85FB50002EA	D840	040	DPH012	Active
*	*	Y85FB50002EB	D840	040	DPH004	Active
		Y85FB50002EB	D840	040	DPH013	Active
*	*	Y85FB50002EC	D840	040	DPH005	Active
		Y85FB50002EC	D840	040	DPH014	Active
*	*	Y85FB50002ED	D840	040	DPH006	Active
		Y85FB50002ED	D840	040	DPH015	Active

More...

Press Enter to continue.

F3=Exit

F5=Refresh

F9=Display disk unit details

F11=Display encryption status

F12=Cancel

The least significant six digits of an IBM i disk unit serial number for a FlashSystem 900 LUN come from the volume unique identifier (UID) as assigned by the FlashSystem 900 as shown in Figure 5-8. The first five digits of the serial number following the letter “Y” are a unique hash value built by IBM i, which cannot be used to identify a particular storage system.

Figure 5-8 FlashSystem 900 GUI Properties for Volume

Figure 5-9 shows the example of 16x 80 GiB volumes from an IBM FlashSystem 900 with its 84577 MB usable capacity (80 x 1024 x 1024 x 1024 x 4096/4160 bytes) reported on IBM i.

Display Disk Configuration Capacity									
ASP	Unit	Type	Model	Threshold	Overflow	----Protected----		---Unprotected---	
						Size	%Used	Size	%Used
1				90%	No	0	0.00%	1353245	1.19%
	1	D840	040			0	0.00%	84577	8.30%
	2	D840	040			0	0.00%	84577	0.71%
	3	D840	040			0	0.00%	84577	0.71%
	4	D840	040			0	0.00%	84577	0.71%
	5	D840	040			0	0.00%	84577	0.71%
	6	D840	040			0	0.00%	84577	0.72%
	7	D840	040			0	0.00%	84577	0.72%
	8	D840	040			0	0.00%	84577	0.72%
	9	D840	040			0	0.00%	84577	0.71%
	10	D840	040			0	0.00%	84577	0.72%
	11	D840	040			0	0.00%	84577	0.71%
More...									
Press Enter to continue.									
F3=Exit                    F5=Refresh                    F10=Display disk unit details									
F11=Display disk configuration protection    F12=Cancel									

Figure 5-9 IBM i SST Display Disk Configuration Capacity

IBM FlashSystem 900 LUNs, despite their inherent RAID protection by the FlashSystem, are reported as *unprotected* disk units to IBM i as shown in Figure 5-9. Thus, additional storage system level protection can be implemented by using IBM i mirroring to a second IBM FlashSystem.

### 5.3.5 FlashSystem 900 and Linux client hosts

The FlashSystem 900 can be attached to Linux client hosts with the following methods:

- ▶ Fibre Channel (FC)
- ▶ InfiniBand
- ▶ FCoE
- ▶ iSCSI

The FlashSystem 900 benefits the most from operating systems where multipathing and logical volumes are supported. Most Linux distributions have the same optimum configurations. Specific Linux configuration settings are shown.

#### Network topology guidelines

You can use an arbitrated loop or point-to-point topology on FC configurations for Linux hosts.

**Note:** The FlashSystem 900 16 Gbps FC ports do not support direct connection to client hosts. A SAN switch must be placed between the IBM FlashSystem 900 and any 16 Gbps-attached client host.

## Aligning a partition using Linux

Use this procedure to improve performance by aligning a partition in the Linux operating system.

The Linux operating system defaults to a 63-sector offset. To align a partition in Linux using **fdisk**, complete the following steps:

1. At the command prompt (#), enter the **fdisk /dev/mapper/<device>** command.
2. To change the listing of the partition size to sectors, enter **u**.
3. To create a partition, enter **n**.
4. To create a primary partition, enter **p**.
5. To specify the partition number, enter **1**.
6. To set the base sector value, enter **128**.
7. Press **Enter** to use the default last sector value.
8. To write the changes to the partition table, enter **w**.

**Note:** The *<device>* is the FlashSystem 900 volume. Example 5-24 on page 153 shows how to create device names for a FlashSystem 900 volume.

The newly created partition now has an offset of 64 KB and works optimally with an aligned application.

If you are installing the Linux operating system on the storage system, create the partition scheme before the installation process. For most Linux distributions, this process requires starting at the text-based installer and switching consoles (press **Alt+F2**) to get the command prompt before you continue.

## Multipathing information for Linux

You can use MPIO to improve the performance of the Linux operating system. Linux kernels of 2.6, and later, support multipathing through device-mapper-multipath. This package can coexist with other multipathing solutions if the other storage devices are excluded from device-mapper.

For a template for the `multipath.conf` file, see Example 5-23 on page 152.

Because the storage system controllers provide true active/active I/O, the `rr_min_io` field in the `multipath.conf` file is set to 4. This results in best distribution of I/O activity across all available paths. You can set it to 1 for a round-robin distribution or if the I/O activity is more sequential in nature, you can increase the `rr_min_io` field by factors of 2 for a performance gain by using buffered I/O (non-direct).

## Integrating InfiniBand controllers

To integrate with InfiniBand technology, the storage system provides block storage by using the SCSI Remote Direct Memory Access (RDMA) Protocol (SRP).

The Linux operating system requires several software modules to connect to the storage system through InfiniBand technology and SRP. In particular, make sure that you install the *srp* and *srptools* modules, and install drivers for the server's host channel adapter (HCA). Use the OpenFabrics Enterprise Distribution (OFED) package from the following website to install these modules, either individually or by using the Install All option:

<http://www.openfabrics.org>



Figure 5-10 shows the setting of the InfiniBand `/etc/infiniband/openib.conf` configuration file for SRP.

```
# Load SRP module
SRP_LOAD=yes #
Enable SRP High Availability daemon
SRPHA_ENABLE=yes
SRP_DAEMON_ENABLE=yes
```

*Figure 5-10 InfiniBand configuration file*

These settings cause the SRP and the SRP daemons to load automatically when the InfiniBand driver starts. The SRP daemon automatically discovers and connects to InfiniBand SRP disks.

Use the `SRPHA_ENABLE=yes` setting. This setting triggers the multipath daemon to create a multipath disk target when a new disk is detected.

InfiniBand technology also requires a Subnet Manager (SM). An existing InfiniBand network already has an SM. In many cases, an InfiniBand switch acts as the SM. If an SM is needed, install OpenSM, which is included with the OFED package, and start it on a single server in the network by entering the following command:

```
# /etc/init.d/opensmd start
```

This script opens an SM on a single port only. If multiple ports directly connect to the storage system, a custom script is needed to start the SM on all ports.

### 5.3.6 FlashSystem 900 and Microsoft Windows client hosts

The FlashSystem 900 can be attached to Windows client hosts with the following methods:

- ▶ Fibre Channel (FC)
- ▶ FCoE

The IBM FlashSystem 900 sees the most benefit from operating systems where multipathing and logical volumes are supported. However, certain applications depend on operating systems that are designed for workstations, and they can still benefit from the storage system performance.

#### Network topologies for Windows hosts

Arbitrated loop or point-to-point topology can be used on the FC configuration for Windows hosts.

**Note:** The FlashSystem 900 16 Gbps FC ports do not support direct connection to client hosts. A SAN switch must be placed between the IBM FlashSystem 900 and any 16 Gbps-attached client host.

#### Implementing 4 KB alignment for Windows Server 2003

Use this procedure to improve performance by establishing a 4 KB alignment on a Windows operating system.

Before Windows Vista and Windows Server 2008, systems running the Windows operating systems offset the partition by 63 sectors, or 31.5 KB.

To align to the preferred 4 KB sector size, implement the offset by using the `diskpart.exe` utility:

1. Start the Windows `diskpart.exe` utility to open the DISKPART line prompt.
2. To view the list of available LUNs, enter the following command:  
`DISKPART> list disk`
3. To select the LUN that holds the file system, enter the following command:  
`DISKPART> select disk <disk number>`
4. To create a partition on the selected LUN, enter the following command:  
`DISKPART> create partition primary align=64`
5. Use the Microsoft Management Console (MMC) or other method to assign a file system or drive letter (raw access) to the partition.

If you are installing the Windows XP or Windows Server 2003 operating system, create the partition on the LUN before the installation of the operating system. You can create the partition by using a Linux Live CD, or by presenting the LUN to another Windows host and disconnecting the drive after the partitioning is complete.

### **Windows Server 2003 multipathing**

Configuring MPIO on a Windows Server 2003 operating system can improve reliability. Windows Server 2003 has a built-in MPIO driver that is provided by Microsoft. This driver is responsible for aggregating the links of storage systems and reporting the addition or removal of links to the kernel while online. To use this feature, a Device Specific Module (DSM) is provided to identify the storage system to the MPIO driver. All major storage vendors have support for the multipathing function and coexist safely because of the common driver in the Windows operating system.

To obtain a copy of the FlashSystem 900 driver for Windows Server 2003, a SCORE or RPQ must be submitted to IBM to request approval. To submit a SCORE or RPQ, contact your IBM representative or IBM Business Partner.

If you are using EMC PowerPath data path management software, it must be at version 4.6, or later, before the storage system DSM can be used.

### **Windows Server 2008 and Windows Server 2012 multipathing**

Windows Server operating system versions that begin with Windows Server 2008 no longer require a separate DSM. Instead, the MPIO function must be installed on the server. For more information, see the Microsoft TechNet website:

[http://technet.microsoft.com/en-us/library/ee619752\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/ee619752(WS.10).aspx)

You can enable multipathing by selecting the **Server Manager** → **Features** option. See Figure 5-11.

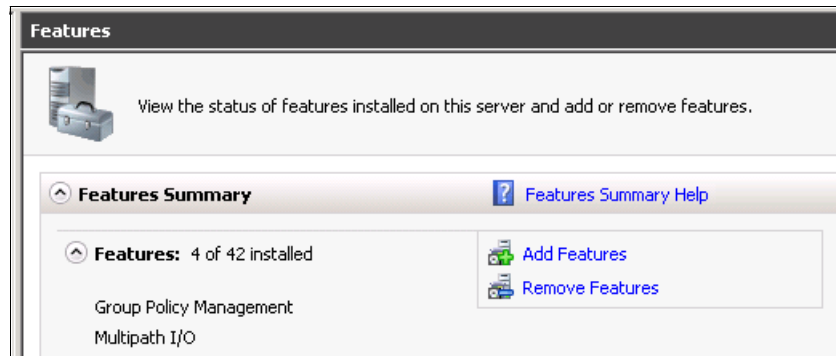


Figure 5-11 Windows 2008 example of activated multipathing

You can set vendor ID (IBM) and product ID (FlashSystem-9840) using **Administrative Tools** → **MPIO**. You enter the eight-character vendor ID and the 16-character product ID by using the MPIO Devices pane (Figure 5-12).

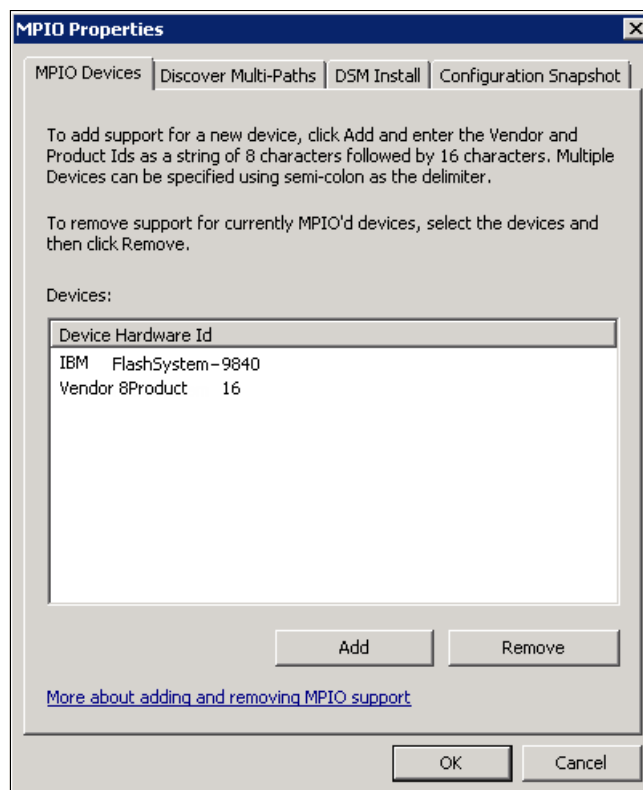


Figure 5-12 Windows MPIO vendor ID and product ID for the FlashSystem 900

**Note:** The vendor ID must be eight characters in length, including spaces. The product ID must be 16 characters in length, including spaces.

Table 5-2 shows the correct vendor ID and product ID for different IBM FlashSystem products.

Table 5-2 IBM FlashSystem SCSI standard inquiry data

IBM FlashSystem	Vendor identification	Product identification
IBM FlashSystem 900	IBM	FlashSystem-9840
IBM FlashSystem 840	IBM	FlashSystem-9840
IBM FlashSystem 820	IBM	FlashSystem
IBM FlashSystem 720	IBM	FlashSystem
IBM FlashSystem 810	IBM	FlashSystem
IBM FlashSystem 710	IBM	FlashSystem

After you install the MPIO function, set the load balance policy on all storage system LUNs to **Least Queue Depth** (Figure 5-13). All available paths to the LUNs are then used to aggregate bandwidth. The load balance policy is set through the Properties pane of each multipath disk device in the Windows Device Manager.

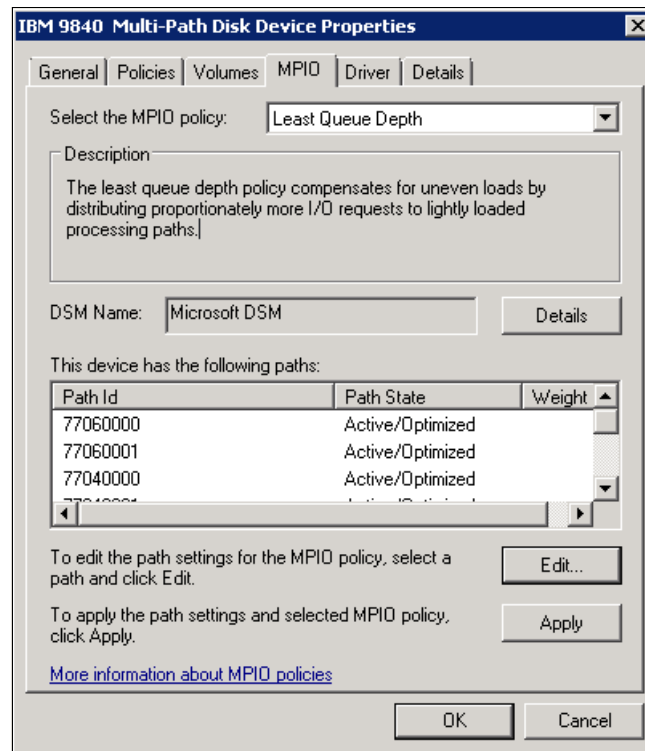


Figure 5-13 Windows MPIO queue configuration

## Power option setting for the highest performance

Select **Control Panel** → **Hardware** → **Power Options**. In the window, set the Windows power plan to **High performance** (Figure 5-14).

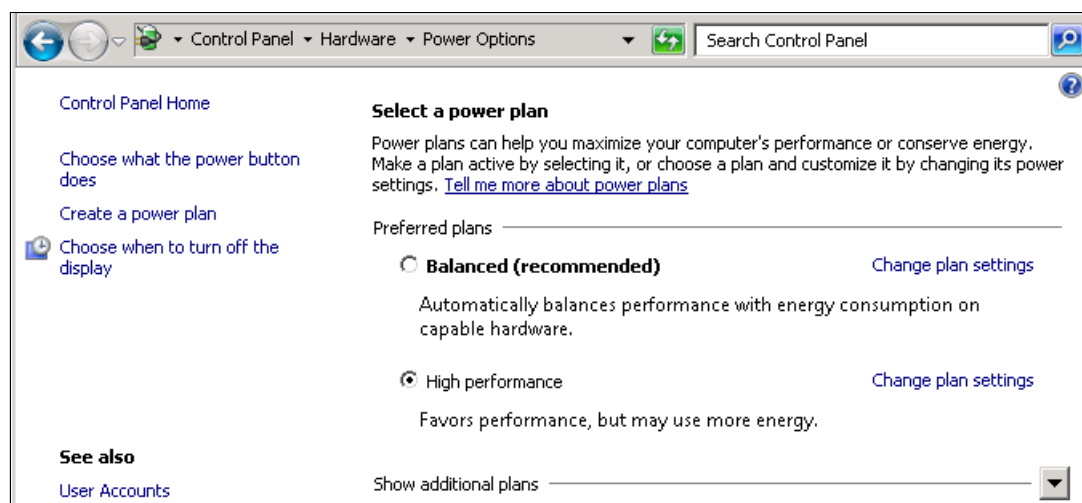


Figure 5-14 Windows Power Options

## Optimum disk command timeout settings

Adjust the disk **TimeOutValue** parameter on the Windows operating system for more reliable multipath access.

Windows operating systems have a default disk command **TimeOutValue** of 60 seconds. If a SCSI command does not complete, the application waits 60 seconds before an I/O request is tried again. This behavior can create issues with most applications, so you must adjust the disk **TimeOutValue** in the registry key to a lower value. For more information about setting this value, see the Microsoft TechNet website.

Adjust this key to your needs:

HKLM\System\CurrentControlSet\Services\Disk\TimeOutValue

- ▶ For a disk in a *nonclustered* configuration, set TimeOutValue to 10.
- ▶ For a disk in a *clustered* configuration, set TimeOutValue to 20.

## Windows 2012 R2 fast volume formatting

Formatting a FlashSystem volume in Windows 2012 r2 can take a long time to complete. For example, a *quick format* of a 1.5 TB volume can take 30 minutes or more to complete.

This behavior is related to the way the Windows **trim** command (known as **TRIM**) works in conjunction with the **TRIM** feature of FlashSystem 900. **TRIM**, also known as **UNMAP** in the SCSI command set, allows an operating system to inform a solid-state drive (SSD) or flash drive which blocks of data are no longer considered in use and can be wiped internally.

On a new uninitialized FlashSystem 900, you can disable **TRIM** and the format will take 1 - 2 seconds to complete. **TRIM** should be re-enabled after the format is completed.

**Important:** The suggestion is to format a volume with **Trim** enabled, except in the case of a new uninitialized FlashSystem 900.

On a new uninitialized FlashSystem 900, *without any previously stored data*, a customer can disable, then later re-enable **TRIM** from the Windows host:

**DISABLE**                      fsutil behavior set DisableDeleteNotify 1

**RE-ENABLE**                fsutil behavior set DisableDeleteNotify 0

Changes will take effect immediately, with the exception of in-flight requests. Implementing this change does not require a reboot and it affects all volumes on the Windows server. After formatting of the volume is complete, be sure to **RE-ENABLE TRIM**.

### 5.3.7 FlashSystem 900 and client VMware ESX hosts

The FlashSystem 900 can be attached to VMware ESX client hosts by using the following methods:

- ▶ Fibre Channel (FC)
- ▶ FCoE

Arbitrated loop topology is required when you attach the IBM FlashSystem 900 directly to the client VMware ESX hosts.

**Note:** The FlashSystem 900 16 Gbps FC ports do not support direct connection to client hosts. A SAN switch must be placed between the IBM FlashSystem 900 and any 16 Gbps-attached client host. If arbitrated loop connection is required by the host, 8 Gbps FC must be used.

To configure round-robin multipathing in a VMware ESX environment, complete these steps:

1. In the vSphere client, select the **Configuration** tab.
2. In the Devices view, select each disk on which you want to change the path selection.
3. In the Manage Paths pane, change the Path Selection setting to **Round Robin** (VMware).

You must set the alignment in the guest operating system. VMware aligns its data stores to 64 KB, but guest virtual machines must still align their own presentation of the storage. Before you continue the installation of a guest Linux operating system or a guest Windows Server 2003 operating system, partition the storage to the aligned accesses.

#### VAAI unmap support

FlashSystem 900 supports the VMware ESXi **VAAI unmap** command. The ESXi host informs the storage system that files or VMs had been deleted or moved from a VMFS data store using **VAAI unmap**. The **VAAI unmap** is then using the **SCSI UNMAP** command. This VAAI command is often used with Thin Provisioned VMFS datastores but this command is not restricted to Thin Provisioned volumes. FlashSystem 900 does not provide thin provisioned volumes.

**Note:** The FlashSystem 900 does not support the Atomic Test and Set (ATS) API of VMware. To avoid unnecessary SCSI reservation conflicts and performance impacts, this feature should not be used and must be disabled in the clients VMware configuration

### 5.3.8 FlashSystem 900 and IBM SAN Volume Controller or Storwize V7000

For details about IBM SAN Volume Controller or Storwize V7000 product integration, considerations, and configuration with the IBM FlashSystem 900, see Chapter 8, “Product integration” on page 279.

### 5.3.9 FlashSystem iSCSI host attachment

Support for iSCSI hosts is planned for these operating systems:

- ▶ Red Hat Enterprise Linux 6.5
- ▶ SUSE Linux Enterprise Server 11 Service Pack 3
- ▶ Microsoft Windows 2012 R2

You can find an excellent description about the host site for iSCSI connection in *Implementing the IBM Storwize V7000 V7.2*, SG24-7938.

**Note:** Always check the IBM System Storage Interoperation Center (SSIC) to get the latest information about supported operating systems, hosts, switches, and so on:

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

### 5.3.10 FlashSystem iSCSI configuration

Figure 5-15 shows the GUI I/O port information. In this picture, ports 3 and 11 are online.

**Note:** The FlashSystem 900 iSCSI port ID numbering starts with ID 3. The two ports with ID 1 and 2 are the management ports of the two canisters as shown in Example 5-5.

Port	State	IP	Subnet	Gateway
	Online	192.168.61.215	255.255.255.0	192.168.61.1
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Offline	192.168.1.10	255.255.255.0	0.0.0.0
	Online	192.168.61.216	255.255.255.0	192.168.61.1
	Offline	192.168.1.10	255.255.255.0	0.0.0.0

Figure 5-15 FlashSystem 900 iSCSI properties

Example 5-5 shows the configuration by using the **lsportip** command.

*Example 5-5 FlashSystem 900 lsportip information*

---

```
>lsportip -delim :
id:node_id:node_name:canister_id:adapter_id:port_id:IP_address:mask:gateway:IP_address_6:prefix_6:gateway_6:MAC:duplex:state:speed:failover:link_state:host:host_6
1:1:node1:0:0:0:::0::40:f2:e9:4a:a3:34::management_only::no:active::
2:2:node2:0:0:0:::0::40:f2:e9:4a:a2:84::management_only::no:active::
3:1:node1:1:1:1:192.168.6x.xxx:255.255.255.0:192.168.61.1::0::40:F2:E9:4A:01:99::online:10Gb:no:active::
4:1:node1:1:1:2:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:9B::offline:NONE:no:inactive::
5:1:node1:1:1:3:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:9D::offline:NONE:no:inactive::
6:1:node1:1:1:4:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:9F::offline:NONE:no:inactive::
7:1:node1:1:2:1:192.168.1.10:255.255.255.0:0.0.0.0::0::40:16:88:06:00:01::offline:NONE:no:inactive::
8:1:node1:1:2:2:192.168.1.10:255.255.255.0:0.0.0.0::0::40:16:88:06:00:03::offline:NONE:no:inactive::
9:1:node1:1:2:3:192.168.1.10:255.255.255.0:0.0.0.0::0::40:16:88:06:00:05::offline:NONE:no:inactive::
10:1:node1:1:2:4:192.168.1.10:255.255.255.0:0.0.0.0::0::40:16:88:06:00:07::offline:NONE:no:inactive::
11:2:node2:2:1:1:192.168.6x.xxx:255.255.255.0:192.168.61.1::0::40:F2:E9:4A:01:91::online:10Gb:no:active::
12:2:node2:2:1:2:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:93::offline:NONE:no:inactive::
13:2:node2:2:1:3:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:95::offline:NONE:no:inactive::
14:2:node2:2:1:4:192.168.1.10:255.255.255.0:0.0.0.0::0::40:F2:E9:4A:01:97::offline:NONE:no:inactive::
15:2:node2:2:2:1:192.168.1.10:255.255.255.0:0.0.0.0::0::CC:CC:CC:CC:CC:CD::offline:NONE:no:inactive::
16:2:node2:2:2:2:192.168.1.10:255.255.255.0:0.0.0.0::0::CC:CC:CC:CC:CC:CF::offline:NONE:no:inactive::
17:2:node2:2:2:3:192.168.1.10:255.255.255.0:0.0.0.0::0::CC:CC:CC:CC:CC:D1::offline:NONE:no:inactive::
18:2:node2:2:2:4:192.168.1.10:255.255.255.0:0.0.0.0::0::CC:CC:CC:CC:CC:D3::offline:NONE:no:inactive::
```

---



### 5.3.11 Windows 2008 R2 and Windows 2012 iSCSI attachment

In Windows 2008 R2 and 2012, the Microsoft iSCSI software initiator is preinstalled. Enter `iscsi` in the search field of the Windows start menu (Figure 5-16) and click **iSCSI Initiator**.

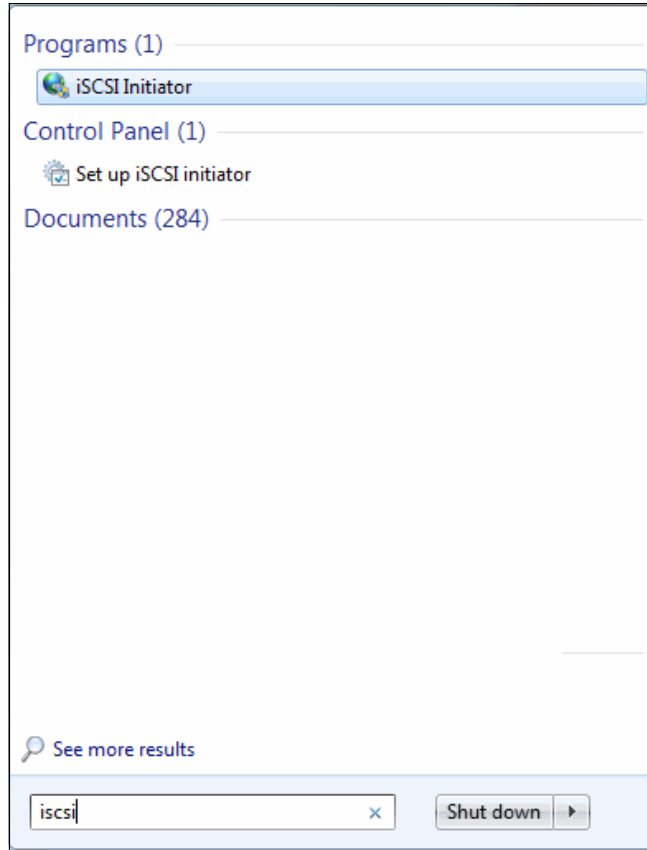


Figure 5-16 Windows iSCSI initiator

Confirm the automatic startup of the iSCSI service (Figure 5-17).

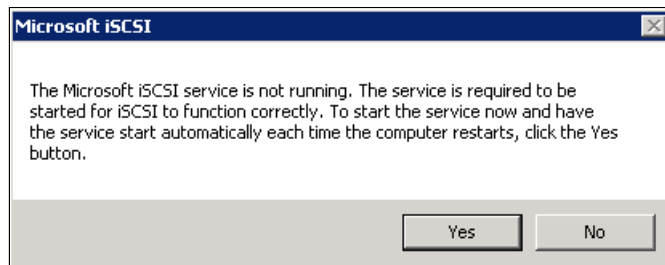


Figure 5-17 Automatic startup of the iSCSI service

The iSCSI Configuration window opens. Select the **Configuration** tab (Figure 5-18). Write down the initiator name of your Windows host, you will use it later in the host configuration part when configuring FlashSystem 900.

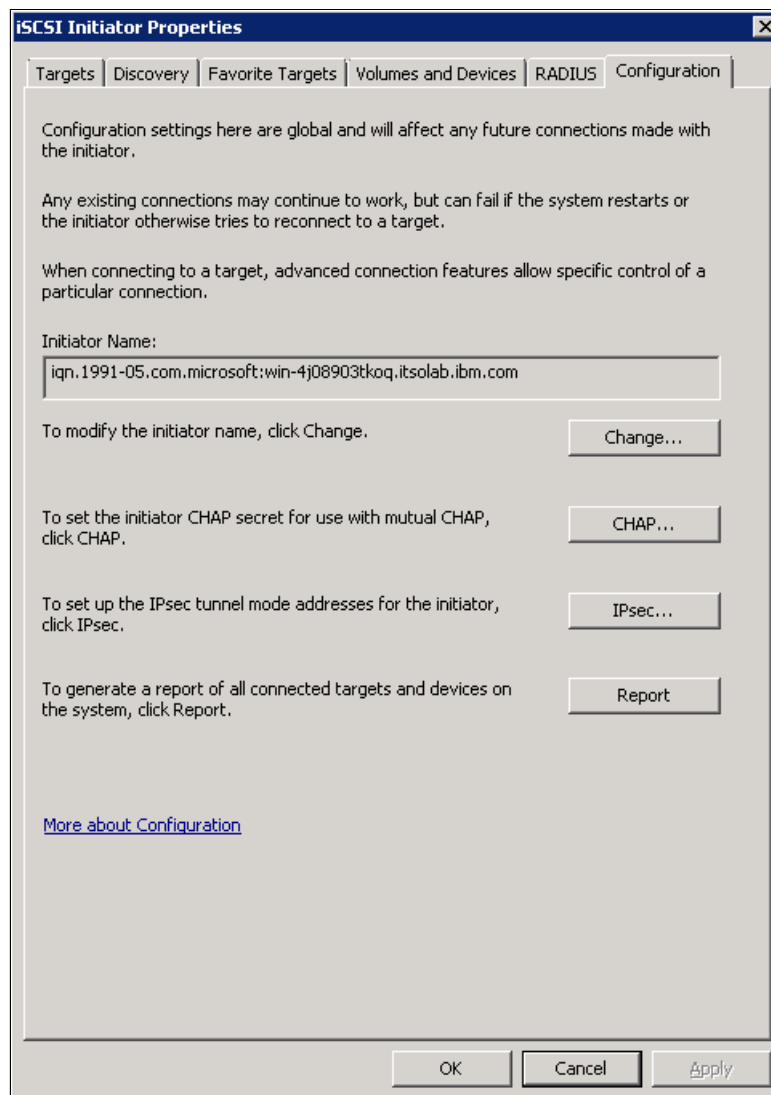


Figure 5-18 iSCSI Initiator Properties window

Open the FlashSystem 900 GUI and open the Add Host dialog by selecting **Hosts** → **Hosts** → **Add Hosts**.

Enter a name for the Windows host and add the initiator name of the Windows host that is shown in Figure 5-19 on page 127, which shows the dialog with values.

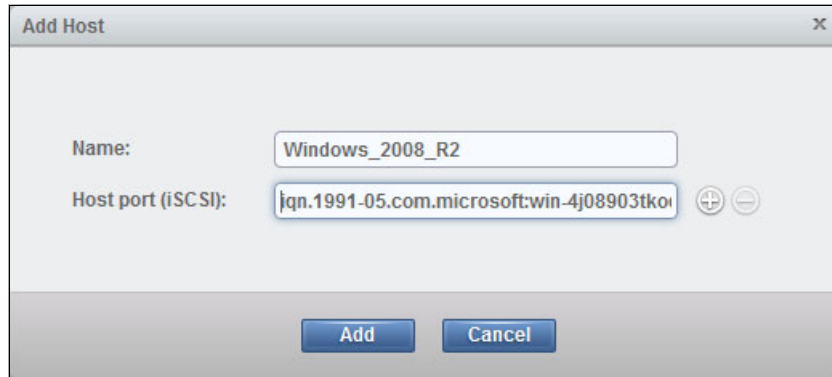


Figure 5-19 iSCSI add host dialog

Create the FlashSystem 900 LUNs for this host and map them to this host as described in 6.3, “Volumes menu” on page 199.

Return to the Windows system’s iSCSI Initiator Properties window and select the **Targets** tab. Click **Refresh**; the FlashSystem 900 is now listed in the Discovered targets section (Figure 5-20). If FlashSystem 900 is not listed, you can add its IP addresses in the Target field and click **Quick Connect**.

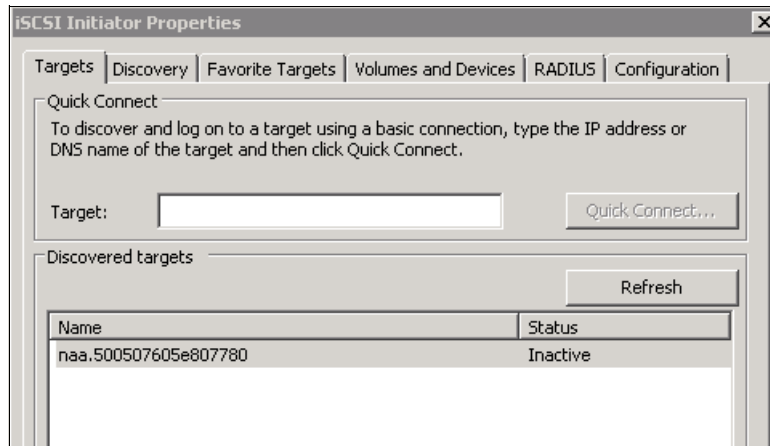


Figure 5-20 iSCSI discovered targets

Click **Connect** and select **Enable multi-path** if the host is connected with multiple Ethernet port to FlashSystem 900 (Figure 5-21).

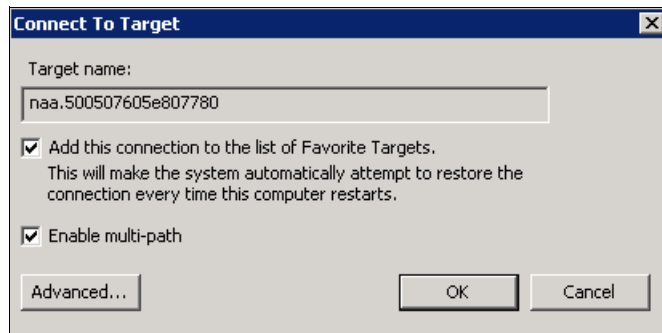


Figure 5-21 Connect to target dialog

### 5.3.12 Linux iSCSI attachment

Install or set up the iSCSI initiator software according to the instructions of your Linux distribution and version.

To configure the FlashSystem 900, you need the iSCSI initiator name. Example 5-6 shows how to find the initiator name on Red Hat Enterprise Linux 6.2.

*Example 5-6 Red Hat 6.2 iSCSI initiator name*

---

```
# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:21774cad37da
```

---

Open the FlashSystem 900 GUI and select **Hosts** → **Hosts** → **Add Hosts**. The Add Host dialog opens (Figure 5-22). Enter a name for the Linux host and add the initiator name of the host as shown in Example 5-6.

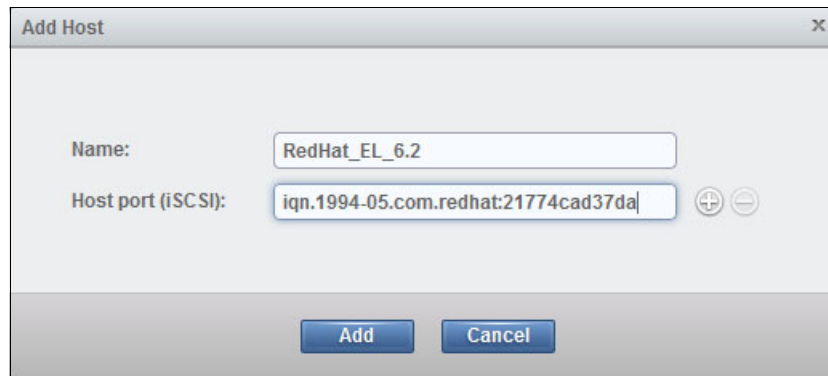


Figure 5-22 iSCSI Add Host dialog

Create the FlashSystem 900 LUNs for this host and map them to this host as described in 6.3, “Volumes menu” on page 199.

Return to the Linux host and start the iSCSI disk discovery. Example 5-7 shows the command that is used to discover the LUNs created on the FlashSystem 900 for this host. In this example, two FlashSystem LUNs will be attached to the Linux host. The Linux host will have one path to each canister. Multipathing will be set up.

*Example 5-7 Discover iSCSI targets and devices*

---

```
#
# # use iscsiadm --mode discoverydb to detect iSCSI targets
# # detect a connection to each FlashSystem 900 to be able to set up multipathing

# # first FlashSystem canister
# iscsiadm --mode discoverydb --type sendtargets --portal 192.168.61.215 --discover
192.168.6x.xxx:3260,-1 naa.500507605e807780

# # second FlashSystem canister
# iscsiadm --mode discoverydb --type sendtargets --portal 192.168.61.216 --discover
192.168.6x.xxx:3260,-1 naa.500507605e807780

# # login to first target to detect iSCSI disks
# iscsiadm --mode node --targetname naa.500507605e807780 --portal 192.168.6x.xxx --login
Logging in to [iface: default, target: naa.500507605e807780, portal: 192.168.6x.xxx,3260] (multiple)
Login to [iface: default, target: naa.500507605e807780, portal: 192.168.6x.xxx,3260] successful.
```

```
# # login to second target to detect iSCSI disks
# iscsiadm --mode node --targetname naa.500507605e807780 --portal 192.168.6x.xxx --login
Logging in to [iface: default, target: naa.500507605e807780, portal: 192.168.6x.xxx,3260] (multiple)
Login to [iface: default, target: naa.500507605e807780, portal: 192.168.6x.xxx,3260] successful.
```

```
# # use iscsiadm to get detailed information about the target
# # both targets will only differ in the node addresses
# iscsiadm -m node --targetname=naa.500507605e807780 --portal=192.168.6x.xxx --op=show
# BEGIN RECORD 6.2.0-873.2.el6
node.name = naa.500507605e807780
... (lines left out)
node.discovery_address = 192.168.6x.xxx
... (lines left out)
node.conn[0].address = 192.168.6x.xxx
... (lines left out)
```

```
# # use iscsiadm -m session to get the attached iSCSI disk
# # they are listed at the end of the command output
# # the two falsystem LUN's are seen by each portal
# iscsiadm -m session -P 3
iSCSI Transport Class version 2.0-870
version 6.2.0-873.2.el6
Target: naa.500507605e807780
    Current Portal: 192.168.6x.xxx:3260,50
    Persistent Portal: 192.168.6x.xxx:3260,50
```

... (lines left out)

```
*****
Attached SCSI devices:
*****
Host Number: 16 State: running
scsi16 Channel 00 Id 0 Lun: 0
scsi16 Channel 00 Id 0 Lun: 2
    Attached scsi disk sds                State: running
scsi16 Channel 00 Id 0 Lun: 3
    Attached scsi disk sdt                State: running
Current Portal: 192.168.61.216:3260,18
Persistent Portal: 192.168.61.216:3260,18
```

... (lines left out)

```
*****
Attached SCSI devices:
*****
Host Number: 17 State: running
scsi17 Channel 00 Id 0 Lun: 0
scsi17 Channel 00 Id 0 Lun: 2
    Attached scsi disk sdu                State: running
scsi17 Channel 00 Id 0 Lun: 3
    Attached scsi disk sdv                State: running
```

---

You can set up Linux multipathing with the configuration file `multipath.conf` and the udev rules (as described in 5.5.3, “Linux configuration file `multipath.conf` example” on page 152 and Example 5-27 on page 156) by adding the FlashSystem 900 LUN Volume Unique Identifier to the `blacklist_exceptions` in the `multipath.conf` file.

Select **Hosts** → **Volumes by Host**, right-click the volume name you want, and then select **Properties**. Figure 5-23 shows this value in the properties window of the FlashSystem 900 LUN using the GUI.



Figure 5-23 FlashSystem 900 LUN Volume Unique Identifier

When using the CLI, you get this information from the `lsvdisk` command. Example 5-8 shows this value in the `multipath.conf` file.

*Example 5-8 Blacklist exception for iSCSI*

---

```
blacklist_exceptions {
    wwid                "36005076*"
}
```

---

This entry will allow all LUNs starting with "36005076" to be used by the Linux multipath daemon. This will also include the FlashSystem 900 FC LUNs. You must restart the Linux multipath daemon for this change to take effect, and then check for new devices as shown in Example 5-9.

*Example 5-9 Restarting the Linux multipath daemon*

---

```
# # add the Volume Unique Identifier to /multipath.conf
# vi /etc/multipath.conf

# # restart multipathd
# service multipathd restart
ok
Stopping multipathd daemon:          [ OK ]
Starting multipathd daemon:          [ OK ]

# # check for new devices
# multipath -v 2
create: mpathd (36005076aa18e082aa000000006000007) undef IBM,FlashSystem-9840
size=130G features='0' hwhandler='0' wp=undef
`-+- policy='queue-length 0' prio=1 status=undef
   | 16:0:0:2 sds 65:32 undef ready running
   `-- 17:0:0:2 sdu 65:64 undef ready running
create: mpathe (36005076aa18e082aa000000007000008) undef IBM,FlashSystem-9840
```

```

size=131G features='0' hwhandler='0' wp=undef
`-+- policy='queue-length 0' prio=1 status=undef
  | 16:0:0:3 sdt 65:48 undef ready running
  ` 17:0:0:3 sdv 65:80 undef ready running

# # list devices
# multipath -ll
mpathe (36005076aa18e082aa000000007000008) dm-5 IBM,FlashSystem-9840
size=131G features='0' hwhandler='0' wp=rw
`-+- policy='queue-length 0' prio=1 status=active
  | 16:0:0:3 sdt 65:48 active ready running
  ` 17:0:0:3 sdv 65:80 active ready running
mpathd (36005076aa18e082aa000000006000007) dm-4 IBM,FlashSystem-9840
size=130G features='0' hwhandler='0' wp=rw
`-+- policy='queue-length 0' prio=1 status=active
  | 16:0:0:2 sds 65:32 active ready running
  ` 17:0:0:2 sdu 65:64 active ready running

#

```

---

## 5.4 Miscellaneous host attachment

This section provides implementation and other general information for connecting client host systems to IBM FlashSystem 900.

**Note:** Always check the IBM SSIC website to get the latest information about supported operating systems, hosts, switches, adapters, and more:

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

If the IBM SSIC does not list the support, submit a SCORE or RPQ to IBM to request approval. To submit a SCORE or RPQ, contact your IBM representative or IBM Business Partner.

### 5.4.1 FlashSystem 900 and Solaris client hosts

The Oracle Solaris operating system has slight differences between x86 and SPARC support when the disks are partitioned. The Solaris Multiplexed I/O (MPxIO) multipathing setups, however, are nearly identical.

#### Network topology guidelines

Configure the settings of the FC ports in the system for the network topology that is used with the storage system.

If they are directly attached to a server, configure the FC ports in the storage system to an arbitrated loop topology. If configured through a switch, set the FC ports to point-to-point to correctly negotiate with the switch.

**Note:** The 16 Gbps FC ports of the FlashSystem 900 do not support direct connection to client hosts. A SAN switch must be placed between the FlashSystem 900 and any 16 GBps-attached client host.

## Sector size information for Solaris

The Solaris operating system supports a 512-byte sector size. The FlashSystem 900 uses a 512-byte sector size.

## Aligning the partition for Solaris

Aligning the partition to a 4 KB boundary improves performance of the storage system.

Solaris SPARC aligns slices on logical unit numbers (LUNs) when the LUN is using an Oracle SUN label for a LUN smaller than 2 TB. If the 2 TB capacity is exceeded, the operating system must use an Extensible Firmware Interface (EFI) label. If you use the EFI label, the partitions that are created do not start on a 4 KB boundary. EFI disks default to 34-sector offsets that have the default partition table. To align the partition to a 4 KB boundary for optimal performance, complete the following steps:

1. After formatting the disk, select the partition option.
2. Choose the **All Free Hog** partitioning base option:

```
partition> 0
Select partitioning base:
0. Current partition table (unnamed)
1. All Free Hog
Choose base (enter number) [0]? 1
```

3. Change the base sector to 40.
4. Press Enter to use the default values for the remaining options:

```
partition> 0
Enter partition id tag[usr]:
Enter partition permission flags[wm]:
Enter new starting Sector[34]: 40
Enter partition size[10066067388b, 10066067427e, 4915071mb, 4799gb, 4tb]:
  Part Tag Flag First Sector Size Last Sector
  0 usr wm 40 4.69TB 10066067427
```

5. After the first sector is changed, save the configuration and then continue by using the newly created partition.

## Multipathing information for Solaris 11 hosts

The method of implementing multipathing depends on the HBA used in the server. MPxIO is the built-in multipathing mechanism for Solaris and no longer requires an HBA that is branded Oracle SUN.

To enable MPxIO support, copy the `/kernel/drv/scsi_vhci.conf` file to the `/etc/driver/drv/scsi_vhci.conf` file and modify the `/etc/driver/drv/scsi_vhci.conf` file.

Example 5-10 shows the FlashSystem entry for Solaris 11.

*Example 5-10 Example of `/etc/driver/drv/scsi_vhci.conf` for Solaris 11 that is reduced for clarity*

---

```
name="scsi_vhci" class="root";
load-balance="round-robin";
auto-failback="enable";
ddi-forceload =
    "misc/scsi_vhci/scsi_vhci_f_asym_sun",
    "misc/scsi_vhci/scsi_vhci_f_asym_lsi",
    "misc/scsi_vhci/scsi_vhci_f_asym_emc",
    "misc/scsi_vhci/scsi_vhci_f_sym_emc",
```



```

        "misc/scsi_vhci/scsi_vhci_f_sym_hds",
        "misc/scsi_vhci/scsi_vhci_f_sym",
        "misc/scsi_vhci/scsi_vhci_f_tpgs";

scsi-vhci-failover-override =
    "IBM FlashSystem-9840", "f_sym";

spread-iport-reservation = "yes";
ipport-rlentime-snapshot-interval = 30;

```

---

**Note:** The "IBM FlashSystem-9840" entry in Example 5-10 contains exactly five spaces.

## Preferred read with Solaris

You can increase the speed of an application by accelerating the read I/Os. Implementing preferred read with the FlashSystem 900 gives you an easy way to deploy the FlashSystem 900 in an existing environment. This section describes how to set up preferred read with Solaris.

Use Solaris Volume Manager (SVM) to create mirrored volumes and to set up preferred read to the first disk in the mirrored volume. Use the following command:

```
metaparam -r first <device>
```

This command modifies the read option for a mirror. The option *first* specifies reading only from the first submirror.

FlashSystem must be the first disk in the mirrored volume. If you create a new volume, you have to select the FlashSystem 900 metadvice as the first metadvice for the volume. If you have an existing mirrored volume, you can use these steps:

1. Add FlashSystem metadvice to the existing volume.
2. Synchronize the data.
3. Destroy the mirrored volume without losing data.
4. Re-create the mirrored volume and use the FlashSystem 900 metadvice as the first entry.

The next example shows the steps in creating a Solaris file system with mirrored devices and preferred read on the FlashSystem 900:

1. Check for the attached FlashSystem 900.
2. Create the partition on the attached FlashSystem 900.
3. Create a mirrored file system.
4. Set preferred read on the FlashSystem 900.

## Check for attached FlashSystem 900

Use the **fcinfo** command to get information about the local FC ports and the devices that are attached to this port. Example 5-11 shows these two use cases.

*Example 5-11 Solaris FC port information*

---

```

#
# # list local FC ports
# fcinfo hba-port
HBA Port WWN: 2100001b320f324e
    Port Mode: Initiator
    Port ID: 10800
    OS Device Name: /dev/cfg/c9

```

```
Manufacturer: QLogic Corp.  
Model: QLE2462  
Firmware Version: 5.6.4  
FCode/BIOS Version: BIOS: 1.29; fcode: 1.27; EFI: 1.09;  
Serial Number: RFC0802K03058  
Driver Name: qlc  
Driver Version: 20120717-4.01  
Type: N-port  
State: online  
Supported Speeds: 1Gb 2Gb 4Gb  
Current Speed: 4Gb  
Node WWN: 2000001b320f324e  
Max NPIV Ports: 127  
NPIV port list:
```

```
# # list systems attached to local FC -ports  
# fcinfo remote-port -s -p 2100001b320f324e  
Remote Port WWN: 500507605efe0ac2  
Active FC4 Types: SCSI  
SCSI Target: yes  
Port Symbolic Name: IBM      FlashSystem-9840 0020  
Node WWN: 500507605efe0ad0  
LUN: 0  
Vendor: IBM  
Product: FlashSystem-9840  
OS Device Name: /dev/rdisk/c10t500507605EFE0AC2d0s2
```

---

## Create a partition on attached FlashSystem 900

The Solaris Server in this example is based on the x86 architecture. Before creating Solaris partitions, the disk must be partitioned for possible operating systems on the x86 server. In Example 5-12, the disk is used only for Solaris.

Some lines were removed for clarity.

### *Example 5-12 Create Solaris partition*

---

```
#  
# # Format disk  
# format  
Searching for disks...done  
  
AVAILABLE DISK SELECTIONS:  
0. c7d0 ...  
   /pci@0,0/pci-ide@1f,2/ide@0/cmdk@0,0  
1. c9t0d1 ...  
   /pci@0,0/pci8086,d13a@5/pci1077,138@0/fp@0,0/disk@w20080020c2117377,1  
2. c9t0d2 ...  
   /pci@0,0/pci8086,d13a@5/pci1077,138@0/fp@0,0/disk@w20080020c2117377,2  
3. c10t500507605EFE0AC2d0 <IBM-9840-0020 cyl 16716 alt 2 hd 224 sec 56>  
   /pci@0,0/pci8086,d13a@5/pci1077,138@0,1/fp@0,0/disk@w500507605efe0ac2,0  
Specify disk (enter its number): 3  
selecting c10t500507605EFE0AC2d0  
[disk formatted]  
No Solaris fdisk partition found.
```

# FORMAT MENU:

- disk - select a disk
- type - select (define) a disk type
- partition - select (define) a partition table
- current - describe the current disk
- format - format and analyze the disk
- fdisk - run the fdisk program
- repair - repair a defective sector
- label - write label to the disk
- analyze - surface analysis
- defect - defect list management
- backup - search for backup labels
- verify - read and display labels
- save - save new disk/partition definitions
- inquiry - show disk ID
- volname - set 8-character volume name
- !<cmd> - execute <cmd>, then return
- quit

format> fdisk

No fdisk table exists. The default partition for the disk is:

a 100% "SOLARIS System" partition

Type "y" to accept the default partition, otherwise type "n" to edit the partition table.

y

format> partition

# PARTITION MENU:

- 0 - change `0' partition
- 1 - change `1' partition
- 2 - change `2' partition
- 3 - change `3' partition
- 4 - change `4' partition
- 5 - change `5' partition
- 6 - change `6' partition
- 7 - change `7' partition
- select - select a predefined table
- modify - modify a predefined partition table
- name - name the current table
- print - display the current table
- label - write partition map and label to the disk
- !<cmd> - execute <cmd>, then return
- quit

partition> print

Current partition table (default):

Total disk cylinders available: 16715 + 2 (reserved cylinders)

Part	Tag	Flag	Cylinders	Size	Blocks
0	unassigned	wm	0	0	(0/0/0) 0
1	unassigned	wm	0	0	(0/0/0) 0
2	backup	wu	0 - 16714	99.98GB	(16715/0/0) 209672960
3	unassigned	wm	0	0	(0/0/0) 0
4	unassigned	wm	0	0	(0/0/0) 0
5	unassigned	wm	0	0	(0/0/0) 0
6	unassigned	wm	0	0	(0/0/0) 0
7	unassigned	wm	0	0	(0/0/0) 0
8	boot	wu	0 - 0	6.12MB	(1/0/0) 12544
9	unassigned	wm	0	0	(0/0/0) 0

```
partition> 0
Part      Tag      Flag      Cylinders      Size      Blocks
  0 unassigned      wm          0          0      (0/0/0)          0
```

```
Enter partition id tag[unassigned]:
Enter partition permission flags[wm]:
Enter new starting cyl[0]:
Enter partition size[0b, 0c, 0e, 0.00mb, 0.00gb]: 16713e
partition> print
Current partition table (unnamed):
Total disk cylinders available: 16715 + 2 (reserved cylinders)
```

Part	Tag	Flag	Cylinders	Size	Blocks
0	unassigned	wm	1 - 16713	99.97GB	(16713/0/0) 209647872
1	unassigned	wm	0	0	(0/0/0) 0
2	backup	wu	0 - 16714	99.98GB	(16715/0/0) 209672960
3	unassigned	wm	0	0	(0/0/0) 0
4	unassigned	wm	0	0	(0/0/0) 0
5	unassigned	wm	0	0	(0/0/0) 0
6	unassigned	wm	0	0	(0/0/0) 0
7	unassigned	wm	0	0	(0/0/0) 0
8	boot	wu	0 - 0	6.12MB	(1/0/0) 12544
9	unassigned	wm	0	0	(0/0/0) 0

... < create a partition on slice 7, size 1 cylinder, to be used for metadb >

```
partition> label
Ready to label disk, continue? yes
```

... < partition written, quit format >

## Create a mirrored file system

You can easily set up a mirrored file system using the Solaris Volume Manager (SVM) Soft Partitioning. Example 5-13 shows the steps to create SVM metadevices and use them to set up a mirrored file system. One mirror is on a spinning disk and the other mirror is on the FlashSystem 900.

Some lines were removed for clarity.

### Example 5-13 Solaris mirrored file system

```
#
# # create an SVM database with the attached devices
# metadb -f -a c9t0d0s7 c10t500507605EFE0AC2d0s7
# metadb -i
      flags      first blk      block count
      a          u          16          8192      /dev/dsk/c9t0d0s7
      a          u          16          8192      /dev/dsk/c10t500507605EFE0AC2d0s7
...
u - replica is up to date
...
a - replica is active, commits are occurring to this replica
...

# # create a one-on-one concatenation on spinning disk
# metainit d_disk 1 1 c9t0d0s0
d_disk: Concat/Stripe is setup
```

```

# # create a one-on-one concatenation mn flashsystem
# metainit d_FlashSystem 1 1 c10t500507605EFE0AC2d0s0

# # mirror setup. first on spinning disk
# metainit d_mirror -m d_disk
d_mirror: Mirror is setup

# # attach second disk
# metattach d_mirror d_FlashSystem
d_mirror: submirror d_FlashSystem is attached

# # check
# metastat -p d_mirror
d_mirror -m /dev/md/rdisk/d_disk /dev/md/rdisk/d_FlashSystem 1
d_disk 1 1 /dev/rdisk/c9t0d0s0
d_FlashSystem 1 1 /dev/rdisk/c10t500507605EFE0AC2d0s0

# # create FS
# newfs /dev/md/rdisk/d_mirror
newfs: construct a new file system /dev/md/rdisk/d_mirror: (y/n)? y
Warning: 512 sector(s) in last cylinder unallocated
/dev/md/rdisk/d_mirror: 115279360 sectors in 18763 cylinders of 48 tracks, 128 sectors
56288.8MB in 1173 cyl groups (16 c/g, 48.00MB/g, 5824 i/g)
super-block backups (for fsck -F ufs -o b=#) at:
32, 98464, 196896, 295328, 393760, 492192, 590624, 689056, 787488, 885920,
Initializing cylinder groups:
.....
super-block backups for last 10 cylinder groups at:
114328992, 114427424, 114525856, 114624288, 114722720, 114821152, 114919584,
115018016, 115116448, 115214880

# # mount
# mount /dev/md/dsk/d_mirror /mnt

# # check
# df -h | grep mirror
/dev/md/dsk/d_mirror 54G 55M 54G 1% /mnt

```

---

The file system is now on two mirrored devices.

## Set preferred read on FlashSystem 900

Solaris Volume Manager uses the round-robin read algorithm as a default. To read from only one device, you can choose the first option. Then, the reads will be performed only from the first device. In this example, the first device that was used was on spinning disk. To use the FlashSystem 900 as the first device in the mirror, you must destroy the mirror and re-create the mirror with the FlashSystem 900 as the first disk.

**Attention:** Be extremely careful while performing the steps described in Example 5-14 because the data can be at risk. Ensure that you back up any critical data before performing any file system management activities.

#### Example 5-14 Solaris preferred read setup

---

```
#
# # destroy mirror, recreate mirror
# # first umount
# umount /dev/md/dsk/d_mirror

# # delete the mirror
# metaclear d_mirror
d_mirror: Mirror is cleared

# # recreate mirror, FlashSystem as first entry
# metainit d_mirror -m d_FlashSystem
d_mirror: Mirror is setup

# # attach spinning disk
# metattach d_mirror d_disk
d_mirror: submirror d_disk is attached

# # check
# metastat
d_mirror: Mirror
    Submirror 0: d_FlashSystem
        State: Okay
    Submirror 1: d_disk
        State: Resyncing
    Resync in progress: 20 % done
    Pass: 1
    Read option: roundrobin (default)
    Write option: parallel (default)
    Size: 115279360 blocks (54 GB)

d_FlashSystem: Submirror of d_mirror
    State: Okay
    Size: 115279360 blocks (54 GB)
    Stripe 0:
        Device          Start Block  Dbase    State Reloc Hot Spare
        c10t500507605EFE0AC2d0s0 0          No       Okay   Yes

d_disk: Submirror of d_mirror
    State: Resyncing
    Size: 115279360 blocks (54 GB)
    Stripe 0:
        Device          Start Block  Dbase    State Reloc Hot Spare
        c9t0d0s0         0          No       Okay   Yes

# # first disk preferred read
# # metaparam -r first d_mirror

# # check changed read algorithm
# metaparam d_mirror
d_mirror: Mirror current parameters are:
    Pass: 1
    Read option: first (-r)
    Write option: parallel (default)

# # mount
# mount /dev/md/dsk/d_mirror /mnt

# # check
# df -h | grep mirror
```

```

/dev/md/dsk/d_mirror    54G    10G    44G    19%    /mnt

# # create a read process and check preferred read with iostat:
# # for example: iostat -xMnz 10
# # only md/d_FlashSystem and md/d_mirror will show activity.

```

**Note:** You must destroy and re-create a metadevice mirror to change the device that is used as the first disk for the preferred read.

## 5.4.2 FlashSystem 900 and HP-UX client hosts

This section discusses HP-UX client host with FlashSystem 900 configurations.

**Note:** Always check the IBM SSIC to ensure that the IBM FlashSystem 900 supports your required client host and version required.

### HP-UX operating system configurations

Consider this checklist when you use HP-UX client hosts:

- ▶ If using Veritas File System (VxFS), select a block size of 4 Kb or greater.
  - ▶ If using HP physical volume links (PVLlinks), add each disk and path into the same Volume Group under LVM.
  - ▶ For directly attaching FlashSystem, FC ports on the FlashSystem must be set to arbitrated loop (AL) topology (not for 16 Gb attachment).
  - ▶ For fabric-attached, use point-to-point topology or Auto.
  - ▶ LUNs greater than 2 TB have not been tested.
  - ▶ For LUNs that require sequential detection, you can use the **mkvdiskhostmap** command to explicitly assign LUNs; LUNs on a bus (FC port) start at LUN 0, LUN 1, and so on.
  - ▶ LUNs are historically limited to 0 - 7 on a bus.
- When HP-UX is on traditional SCSI-2 storage architecture, you will not see LUN 8+ unless the SCSI driver in HP-UX is updated to SCSI-3 compliance.
- ▶ No virtual bus or virtual target architecture in the FlashSystem; more than eight LUNs require LUN masking to overlapping I/O paths.

### Detecting LUNs on HP-UX server

Changes in the FlashSystem LUN configuration do not generate a CHECK\_CONDITION to the host. You must rescan the SCSI bus manually in either of the following ways:

- ▶ # ioscan -fnC disk
- ▶ # ioinit -i (HP-UX 11v1, 11v2)

You might need to force CHECK\_CONDITION by a link reset. You can use the **chportfc** command and its **-reset** option to perform a link reset on a dedicated port.

### Alignment

HP-UX volumes will align to 4 KB boundaries. The VxFS file system must be set to a block size of 4096 or higher to keep alignment:

```
# mkfs -F vxfs -o bsize=4096 <disk>
```

## 5.5 FlashSystem 900 preferred read and configuration examples

Examples of implementing preferred read in different environments and of the Linux `multipath.conf` configuration file are shown in the following sections.

### 5.5.1 FlashSystem 900 deployment scenario with preferred read

Implementing preferred read with the IBM FlashSystem 900 gives you an easy way to deploy the IBM FlashSystem 900 in an existing environment. The data is secured by writing it to two separate storage systems. Data is read at the FlashSystem 900 speed, because it is always read from the FlashSystem 900. This implementation does not change the existing infrastructure concepts, for example, data security, replication, backup, disaster recovery, and so on. Preferred read can be implemented with the following techniques:

- ▶ IBM SAN Volume Controller Virtualize/V7000:
  - Virtual disk mirroring (also known as *volume mirroring*)
- ▶ At the volume manager or operating system level:
  - IBM AIX
  - Linux LVM (native least queue read)
- ▶ At the application level:
  - Oracle Automatic Storage Management (ASM)
  - Standby or reporting instance
  - SQL Server: AlwaysOn Availability Groups maximizes the availability of a set of user databases for an enterprise. An availability group supports a failover environment for a discrete set of user databases, known as availability databases, that fail over together.

The following examples are schemas that show the logical setup. You have to plan the FC, FCoE, or InfiniBand cabling and SAN setup, depending on your environment and needs.

An example of implementing preferred read with the operating system volume manager is shown in Figure 5-24 on page 141. It represents a schema of IBM AIX LVM mirroring.



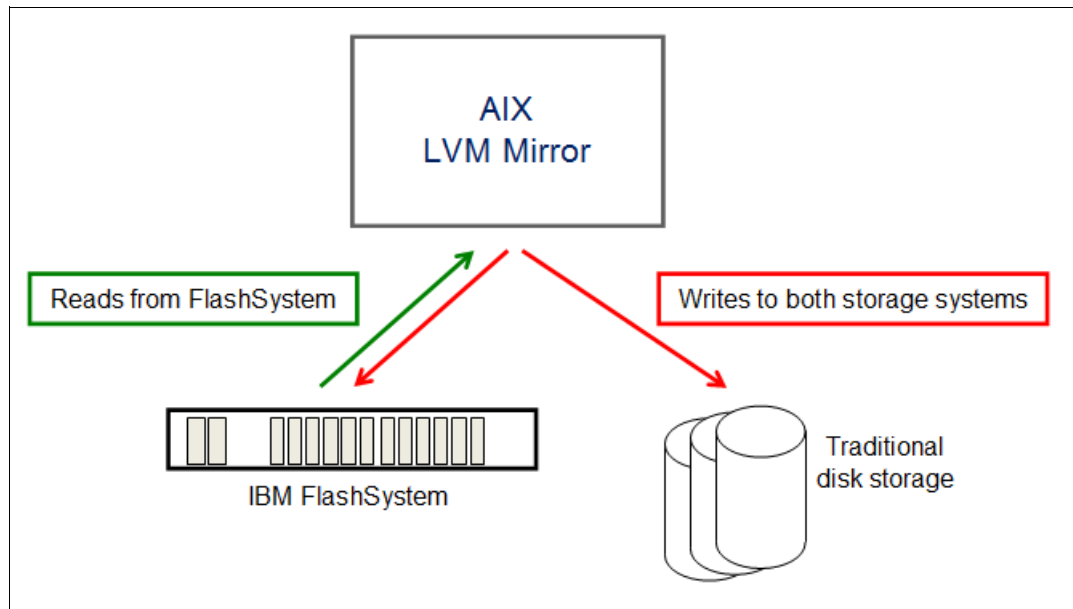


Figure 5-24 Preferred read with AIX

An example of implementing preferred read on the application level is shown in Figure 5-25. It represents a schema of Oracle Automatic Storage Management (ASM) mirroring.

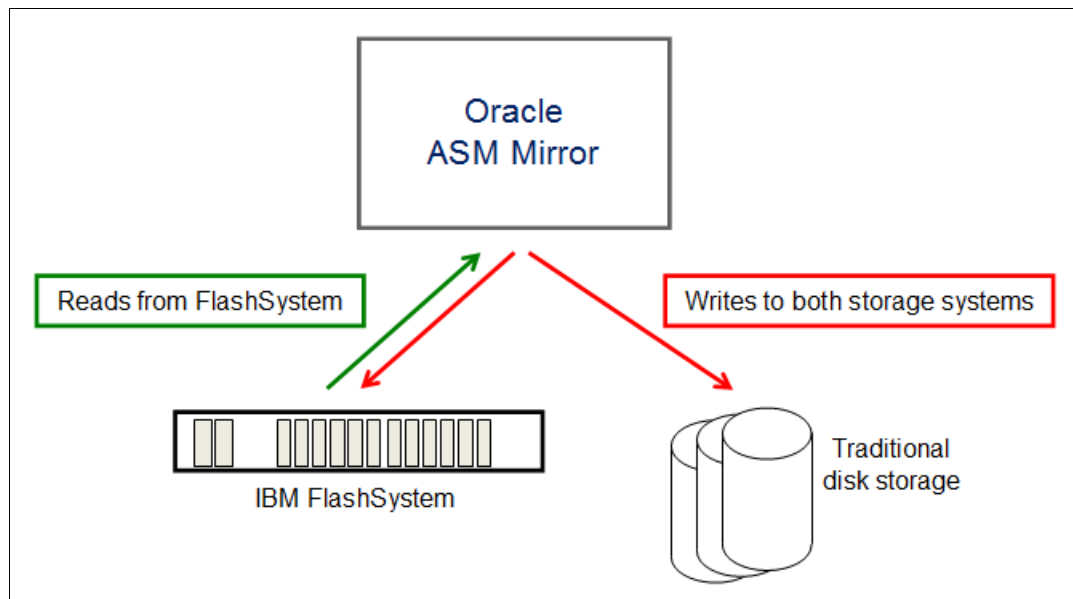


Figure 5-25 Preferred read with Oracle ASM

An example of implementing preferred read on a virtualization layer is shown in Figure 5-26 on page 142. It represents a schema of the IBM SAN Volume Controller.

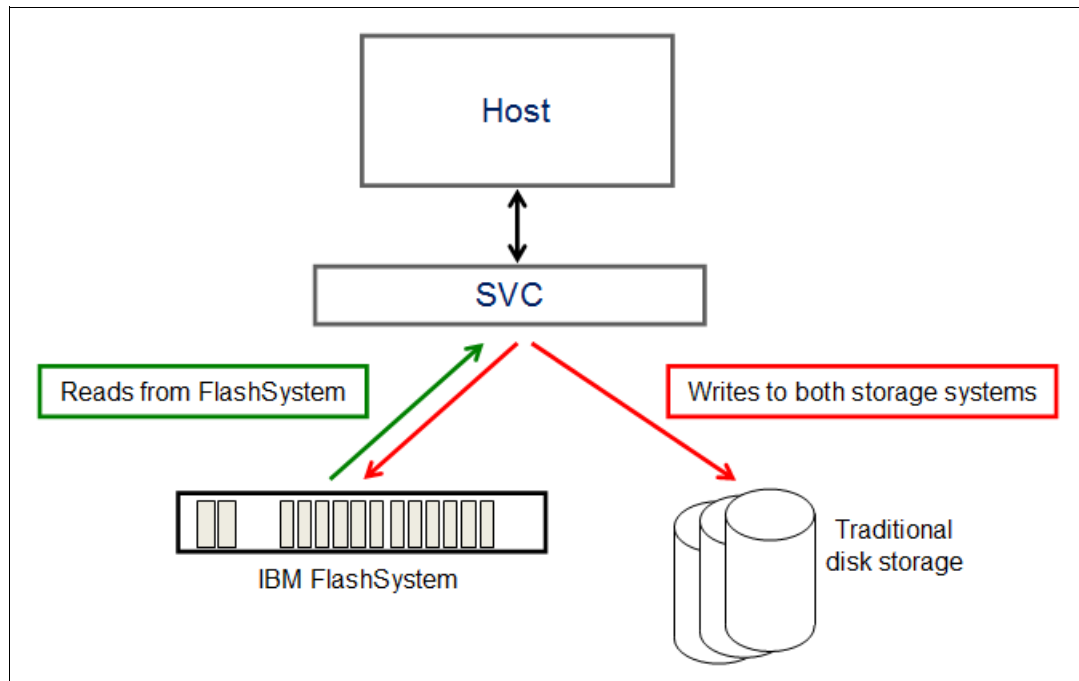


Figure 5-26 Preferred read with the IBM SAN Volume Controller

## 5.5.2 Implementing preferred read

You can increase the speed of an application by accelerating the read I/Os. Implementing preferred read with the IBM FlashSystem 900 gives you an easy way to deploy the IBM FlashSystem 900 in an existing environment. The data is secured by writing it to two separate storage systems. Data is read at the FlashSystem 900 speed, because it is always read from the IBM FlashSystem 900. This implementation does not change the existing infrastructure concepts, for example, data security, replication, backup, disaster recovery, and so on. Preferred read is described in 5.5.1, “FlashSystem 900 deployment scenario with preferred read” on page 140.

### Preferred read with AIX

On AIX, preferred read is implemented by the AIX Logical Volume Manager (LVM).

The following steps are illustrated in Example 5-15 on page 143 through Example 5-21 on page 150. The examples show the process of creating a preferred read configuration with the FlashSystem 900. The steps assume that the AIX server is cabled and zoned correctly.

1. Create a file system on spinning disk.
2. Add the IBM FlashSystem 900 as a mirrored copy to this file system.
3. Set the correct read and write policy.
4. Set preferred read to the IBM FlashSystem 900.

In the following steps (Example 5-15 on page 143 through Example 5-21 on page 150), two systems, which are attached through a SAN to the AIX host, are used. Their AIX hdisk information is listed:

- ▶ hdisk1 - hdisk4: IBM MPIO FC 2145 (V7000)
- ▶ hdisk5 - hdisk8: IBM FlashSystem 900 Storage

The following steps are based on AIX 7.1.

## Create a file system on spinning disk

The steps in Example 5-15 create a file system on AIX. In this example, hdisk1 - hdisk4 are used. All commands are preceded by a comment to the next action. Always check the command parameters against your current AIX version.

### Example 5-15 AIX file system creation

---

```
#
# # Create a file system on normal disks

# # list physical disks
# lsdev -C -c disk
hdisk0 Available Virtual SCSI Disk Drive

# # attach Disksystem to AIX server and check for new disks
# cfgmgr

# # list physical disks
# lsdev -C -c disk
hdisk0 Available Virtual SCSI Disk Drive
hdisk1 Available 00-00-02 MPI0 FC 2145
hdisk2 Available 00-00-02 MPI0 FC 2145
hdisk3 Available 00-00-02 MPI0 FC 2145
hdisk4 Available 00-00-02 MPI0 FC 2145

# # set path policy to your needs: round_robin, load_balance, or shortest_queue
# # check path for all disks, hdisk1 as an example
# lsattr -El hdisk1 | grep algorithm
algorithm load_balance

# # use chdev if needed
# chdev -l hdisk1 -a algorithm=round_robin
# chdev -l hdisk1 -a algorithm=shortest_queue
# chdev -l hdisk1 -a algorithm=load_balance

# # create a volume group with the four IBM 2145 FC Disks
# mkvg -B -y test_vg_1 -t 8 hdisk1 hdisk2 hdisk3 hdisk4
0516-1254 mkvg: Changing the PVID in the ODM.
0516-1254 mkvg: Changing the PVID in the ODM.
0516-1254 mkvg: Changing the PVID in the ODM.
0516-1254 mkvg: Changing the PVID in the ODM.
test_vg_1

# # list the information
# lsvg
rootvg
test_vg_1
# lsvg test_vg_1
VOLUME GROUP: test_vg_1 VG IDENTIFIER:
00f6600100004c00000001464e967f3d
VG STATE: active PP SIZE: 64 megabyte(s)
VG PERMISSION: read/write TOTAL PPs: 3196 (204544 megabytes)
MAX LVs: 512 FREE PPs: 3196 (204544 megabytes)
LVs: 0 USED PPs: 0 (0 megabytes)
OPEN LVs: 0 QUORUM: 3 (Enabled)
TOTAL PVs: 4 VG DESCRIPTORS: 4
STALE PVs: 0 STALE PPs: 0
ACTIVE PVs: 4 AUTO ON: yes
MAX PPs per VG: 130048
MAX PPs per PV: 8128 MAX PVs: 16
```

```

LTG size (Dynamic): 256 kilobyte(s)      AUTO SYNC:      no
HOT SPARE:      no                      BB POLICY:      relocatable
PV RESTRICTION: none                    INFINITE RETRY: no
DISK BLOCK SIZE: 512

# # create a logical volume with file system type jfs2
# # name will be test_lv_1
# mklv -y test_lv_1 -t'jfs2' test_vg_1 3096 hdisk1 hdisk2 hdisk3 hdisk4
test_lv_1

# # create a file system on the logical volume
# # mount point /test/preferred_read will be created at the same time
# crfs -v jfs2 -d test_lv_1 -m /test/preferred_read
File system created successfully.
202893060 kilobytes total disk space.
New File System size is 405798912

# # mount new created file system
# mount /test/preferred_read

# # check
# df -g /test/preferred_read

```

Filesystem	GB blocks	Free	%Used	Iused	%Iused	Mounted on
/dev/test_lv_1	193.50	193.47	1%	4	1%	/test/preferred_read

---

### ***Add the FlashSystem 900 as a mirrored copy to this file system***

The steps in Example 5-16 extend an existing file system and use this extension to create a mirror. In this example, the second disk, hdisk4, is used.

All commands are preceded by a comment to the next action.

#### *Example 5-16 Create a mirrored file system on AIX*

---

```

#
# # Add FlashSystem 900 as a mirrored copy to this file system

# # attach FlashSystem 900 to AIX server and check for new disks
# cfgmgr

# # check for new FlashSystem 900 disk, will be hdisk5 hdisk6 hdisk7 hdisk8
# lsdev -C -c disk
hdisk0 Available          Virtual SCSI Disk Drive
hdisk1 Available 00-00-02 MPI0 FC 2145
hdisk2 Available 00-00-02 MPI0 FC 2145
hdisk3 Available 00-00-02 MPI0 FC 2145
hdisk4 Available 00-00-02 MPI0 FC 2145
hdisk5 Available 00-00-02 MPI0 IBM FlashSystem Disk
hdisk6 Available 00-00-02 MPI0 IBM FlashSystem Disk
hdisk7 Available 00-00-02 MPI0 IBM FlashSystem Disk
hdisk8 Available 00-00-02 MPI0 IBM FlashSystem Disk

# # set path policy to your needs:round_robin or shortest_queue
# # check path for all disks, hdisk5 as an example
# lsattr -El hdisk5 | grep algorithm
algorithm  shortest_queue

# # use chdev if needed
# chdev -l hdisk5 -a algorithm=round_robin
# chdev -l hdisk5 -a algorithm=shortest_queue

```

```

# # list used Physical volume names
# lslv -m test_lv_1 | awk '{print $3, "\t", $5, "\t", $7}' | uniq

PV1      PV2      PV3
hdisk1
hdisk2
hdisk3
hdisk4

# # add FlashSystem 900 disk to volume group
# extendvg test_vg_1 hdisk5 hdisk6 hdisk7 hdisk8
0516-1254 extendvg: Changing the PVID in the ODM.
0516-1254 extendvg: Changing the PVID in the ODM.
0516-1254 extendvg: Changing the PVID in the ODM.
0516-1254 extendvg: Changing the PVID in the ODM.

# # create a mirror
# mklvcopy test_lv_1 2 hdisk5 hdisk6 hdisk7 hdisk8

# # list used Physical volume names and check mirror
# lslv -m test_lv_1 | awk '{print $3, "\t", $5, "\t", $7}' | uniq

PV1      PV2      PV3
hdisk1   hdisk5
hdisk2   hdisk6
hdisk3   hdisk7
hdisk4   hdisk8

# # check mirror state
# lsvg -l test_vg_1
test_vg_1:
LV NAME          TYPE      LPs      PPs      PVs  LV STATE  MOUNT POINT
test_lv_1        jfs2      3096     6192     8    open/stale /test/preferred_read
loglv00          jfs2log   1        1        1    open/syncd  N/A

# # the mirror is stale, synchronize it
# # this command will take some time depending on volume size
# syncvg -P 32 -v test_vg_1

# # check mirror state
# lsvg -l test_vg_1
test_vg_1:
LV NAME          TYPE      LPs      PPs      PVs  LV STATE  MOUNT POINT
test_lv_1        jfs2      3096     6192     8    open/syncd /test/preferred_read
loglv00          jfs2log   1        1        1    open/syncd  N/A

# # turn VG quorum off
# # always check your business needs, if VG quorum should be enabled or disabled
# # do this to ensure the VG will not go offline if a quorum of disks goes missing
# chvg -Q n test_vg_1

# # check VG state
# lsvg test_vg_1
VOLUME GROUP:    test_vg_1          VG IDENTIFIER:
00f6600100004c00000001464e967f3d
VG STATE:        active          PP SIZE:        64 megabyte(s)
.
OPEN LVs:        2              QUORUM:         1 (Disabled)
.

```

---

Now, the file system data is mirrored onto two separate physical locations. The first copy is on spinning disk; the second copy is on the FlashSystem 900.

### ***Set the correct read and write policy***

IBM AIX LVM sets the scheduling policy for reads and writes to the storage systems. If you use mirrored logical volumes, the following scheduling policies for writing to disk can be set for a logical volume with multiple copies:

- ▶ **Sequential scheduling policy**

Performs writes to multiple copies or mirrors in order. The multiple physical partitions representing the mirrored copies of a single logical partition are designated primary, secondary, and tertiary. In sequential scheduling, the physical partitions are written to in sequence. The system waits for the write operation for one physical partition to complete before starting the write operation for the next one. When all write operations are complete for all mirrors, the write operation is complete.

- ▶ **Parallel scheduling policy**

Simultaneously starts the write operation for all the physical partitions in a logical partition. When the write operation to the physical partition that takes the longest to complete finishes, the write operation is complete. Specifying mirrored logical volumes with a parallel scheduling policy might improve I/O read-operation performance because multiple copies allow the system to direct the read operation to the least busy disk for this logical volume.

- ▶ **Parallel write with sequential read scheduling policy**

Simultaneously starts the write operation for all the physical partitions in a logical partition. The primary copy of the read is always read first. If that read operation is unsuccessful, the next copy is read. During the read retry operation on the next copy, the failed primary copy is corrected by the LVM with a hardware relocation. This patches the bad block for future access.

- ▶ **Parallel write with round-robin read scheduling policy**

Simultaneously starts the write operation for all the physical partitions in a logical partition. Reads are switched back and forth between the mirrored copies.

To get the preferred read performance of the IBM FlashSystem 900, set the policy to parallel write with sequential read. With this option, you get these functions:

- ▶ Write operations are done in parallel to all copies of the mirror.
- ▶ Read operations are always done on the primary copy of the devices in the mirror set.

Example 5-17 shows how to change the LVM scheduler to the parallel write with sequential read scheduling policy.

**Important:** Downtime of the file system is required when you change the LVM scheduler policy.

#### *Example 5-17 Changing the scheduler to parallel write with the round-robin read scheduling policy*

```
# # check current state of the LVM scheduler
# lslv test_lv_1
LOGICAL VOLUME:      test_lv_1          VOLUME GROUP:  test_vg_1
LV IDENTIFIER:       00f6600100004c00000001464e967f3d.1  PERMISSION:    read/write
VG STATE:            active/complete    LV STATE:       opened/syncd
TYPE:                jfs2               WRITE VERIFY:   off
MAX LPs:             3096               PP SIZE:       64 megabyte(s)
COPIES:              2                  SCHED POLICY:   parallel
```

```

LPs: 3096 PPs: 6192
STALE PPs: 0 BB POLICY: relocatable
INTER-POLICY: minimum RELOCATABLE: yes
INTRA-POLICY: middle UPPER BOUND: 16
MOUNT POINT: /test/preferred_read LABEL: /test/preferred_read
DEVICE UID: 0 DEVICE GID: 0
DEVICE PERMISSIONS: 432
MIRROR WRITE CONSISTENCY: on/ACTIVE
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?: NO
INFINITE RETRY: no
# lslv test_lv_1 | grep SCHED POLICY
COPIES: 2 SCHED POLICY: parallel

# # Logical volume must be closed.
# # If the logical volume contains a file system,
# # the umount command will close the LV device.
# umount /test/preferred_read

# # set scheduler to parallel write with sequential read-scheduling policy
# # (parallel/sequential)
# # Note: mklv and chlvs: The -d option cannot be used with striped logical volumes.
# chlvs -d ps test_lv_1
# # check changed state of the LVM scheduler
# lslv test_lv_1
LOGICAL VOLUME: test_lv_1 VOLUME GROUP: test_vg_1
LV IDENTIFIER: 00f6600100004c00000001464e967f3d.1 PERMISSION: read/write
VG STATE: active/complete LV STATE: closed/syncd
TYPE: jfs2 WRITE VERIFY: off
MAX LPs: 3096 PP SIZE: 64 megabyte(s)
COPIES: 2 SCHED POLICY: parallel/sequential
LPs: 3096 PPs: 6192
STALE PPs: 0 BB POLICY: relocatable
INTER-POLICY: minimum RELOCATABLE: yes
INTRA-POLICY: middle UPPER BOUND: 16
MOUNT POINT: /test/preferred_read LABEL: /test/preferred_read
DEVICE UID: 0 DEVICE GID: 0
DEVICE PERMISSIONS: 432
MIRROR WRITE CONSISTENCY: on/ACTIVE
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?: NO
INFINITE RETRY: no
# lslv test_lv_1 | grep SCHED POLICY
COPIES: 2 SCHED POLICY: parallel/sequential

# # mount file system
# mount /test/preferred_read
# # write some data to the filesystem /test/preferred_read
# # then read this data and check with iostat

Disks: % tm_act Kbps tps Kb_read Kb_wrtn
hdisk1 0.0 147495.2 1036.4 737476 0
hdisk2 0.0 134964.8 736.6 674824 0
hdisk4 0.0 85035.2 448.6 425176 0
hdisk3 0.0 118422.4 684.2 592112 0

```

The setup of the logical volume is now preferred read from the first copy.

### ***Set preferred read to the FlashSystem 900***

Check which logical volumes are primary, which are secondary, and which are tertiary, if any. This list might be a long list. Example 5-18 lists the first 10 lines of the command output and a command for a quick overview.

*Example 5-18 Get first, second, and third logical volume physical disk*

---

```
#
# # Get first, second, and third logical volume's physical disk
# # Get reduced list of all involved hdisks
# # list used Physical volume names
# lslv -m test_lv_1 | awk '{print $3, "\t", $5, "\t", $7}' | uniq

PV1      PV2      PV3
hdisk1   hdisk5
hdisk2   hdisk6
hdisk3   hdisk7
hdisk4   hdisk8
```

---

Now, the spinning disk devices in the PV1 column are the primary devices. The reads will all be supported by the PV1 devices. During boot, the PV1 devices are the primary copy of the mirror, and they will be used as the sync point.

You must remove the PV1 disks, so that the PV2 disks will be primary. Then, you add the removed disk and set up a mirror again.

Example 5-19 shows how to make the FlashSystem 900 the primary copy.

*Example 5-19 Make the FlashSystem 900 the primary disk*

---

```
# # remove the primary copy which is on spinning disk
# rmlvcopy test_lv_1 1 hdisk1 hdisk2 hdisk3 hdisk4

# # list used Physical volume names ; no mirror
# lslv -m test_lv_1 | awk '{print $3, "\t", $5, "\t", $7}' | uniq

PV1      PV2      PV3
hdisk5
hdisk6
hdisk7
hdisk8

# # add removed spinning disk to mirror data
# mklvcopy test_lv_1 2 hdisk1 hdisk2 hdisk3 hdisk4

# # list used Physical volume names ; first copy now on FlashSystem
# lslv -m test_lv_1 | awk '{print $3, "\t", $5, "\t", $7}' | uniq

PV1      PV2      PV3
hdisk5   hdisk1
hdisk6   hdisk2
hdisk7   hdisk3
hdisk8   hdisk4

# # check mirror state
# lsvg -l test_vg_1
test_vg_1:
LV NAME          TYPE      LPs      PPs      PVs  LV STATE  MOUNT POINT
test_lv_1        jfs2      3096     6192     8    open/stale /test/preferred_read
```



```

loglv00          jfs2log    1      1      1    open/syncd    N/A

# # the mirror is stale, synchronize it
# # this command will take some time depending on volume size
# syncvg -P 32 -v test_vg_1

# # check mirror state
# lsvg -l test_vg_1
test_vg_1:
LV NAME          TYPE      LPs      PPs      PVs  LV STATE    MOUNT POINT
test_lv_1        jfs2      3096     6192     8    open/syncd   /test/preferred_read
loglv00          jfs2log    1        1        1    open/syncd   N/A

# # turn VG quorum off
# # always check your business need if VG quorum should be enabled or disabled
# # do this to ensure the VG will not go offline if a quorum of disks goes missing
# chvg -Q n test_vg_1

# # check VG state
# lsvg test_vg_1
VOLUME GROUP:    test_vg_1          VG IDENTIFIER:  00f6600100004c00000001464e967f3d
VG STATE:        active          PP SIZE:        64 megabyte(s)
.
OPEN LVs:        2              QUORUM:         1 (Disabled)
.

# # again read the data and check with iostat

Disks:          % tm_act      Kbps      tps      Kb_read  Kb_wrtn
hdisk0           0.0          8.8        1.6         44         0
hdisk5           0.0      281169.6    5695.8    1405848         0
hdisk6           0.0      261426.4    5057.2    1307132         0
hdisk7           0.0      228819.2    4434.0    1144096         0
hdisk8           0.0      115506.4    2221.0     577532         0

```

---

You can use the **iostat** command to see the effects of the parallel/sequential scheduler settings and the performance of the spinning disk or the FlashSystem 900 as the primary disk. Execute this command with the normal disk and again later with the FlashSystem 900 as the primary disk:

```
iostat -DR1TV 3
```

Notice that only the first disk is used for reading and that you get a big performance increase with the FlashSystem 900.

Example 5-20 shows the **iostat** output before changing the primary disk. It shows **hdisk1**, which is a spinning disk. Example 5-21 on page 150 shows the result after changing the primary. You see that **hdisk5**, which is the FlashSystem disk, is used only for reading.

Output is shortened for clarity.

*Example 5-20 The iostat command checks for preferred read on the spinning disk*

```

#
# # the volume group has to be in syncd state
# # use dd command to read from the mirrored logical volume
# dd if=/dev/test_lv_1 of=/dev/zero bs=16k count=100000

# # execute iostat in another windows

```

```
# iostat -DRITV 3
Disks:
-----
hdisk1
```

---

*Example 5-21 The iostat command checks for preferred read on the IBM FlashSystem 900*

---

```
#
# # the volume group has to be in syncd state
# # use dd command to read from the mirrored logical volume
# dd if=/dev/test_lv_1 of=/dev/zero bs=16k count=100000

# # execute iostat in another windows
# iostat -DRITV 3

Disks:
-----
hdisk5
```

---

## Setting the FC topology on the FlashSystem 900 and AIX

The IBM FlashSystem 900 can be directly attached to an AIX host without a switch. In this case, the FC ports of the FlashSystem 900 are changed to arbitrated loop (AL) topology. You can use the **chportfc** command to change port settings on the IBM FlashSystem 900. On the AIX system, the ports also needed to be changed to AL.

Example 5-22 shows changing two ports, **fscsi0** and **fscsi2**, on the AIX system. This system has four FC ports. Ports 0 and 2 are directly attached to the FlashSystem 900 using AL and ports 1 and 3 are attached to a switch. The **cfgmgr** command detects the correct topology.

*Example 5-22 Set the AIX port to arbitrated loop*

---

```
# # before using these commands
# # you must first alter the port topology on the FlashSystem 900
# # all traffic has to be stopped before using this command
#
# # remove FC port fscsi0 and then configure it using cfgmgr
# rmdev -Rdl fscsi0
# cfgmgr -vl fcs0

# # remove FC port fscsi2 and then configure it using cfgmgr
# rmdev -Rdl fscsi2
# cfgmgr -vl fcs2

# check all 4 ports
# lsattr -El fscsi0 | grep attach
attach      al

# lsattr -El fscsi1 | grep attach
attach      switch

# lsattr -El fscsi2 | grep attach
attach      al

# lsattr -El fscsi3 | grep attach
attach      switch
```

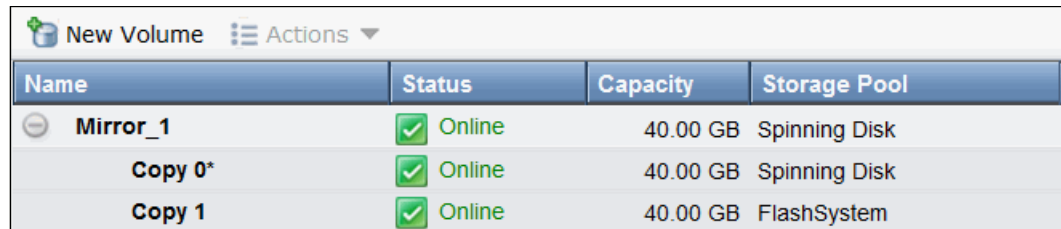
---

**Before you run the commands:** The topology must be set on an attached system, switch, or storage device before you run these commands.

## Preferred read with the IBM SAN Volume Controller

You can set up preferred read on the IBM FlashSystem 900 with the IBM SAN Volume Controller with only one mouse-click. In the IBM SAN Volume Controller GUI, go to the **Volumes** menu and right-click the FlashSystem 900 disk of the mirrored volume. Then, select **Make Primary**. Notice the start asterisk (\*) next to your primary disk, which is now the preferred read disk.

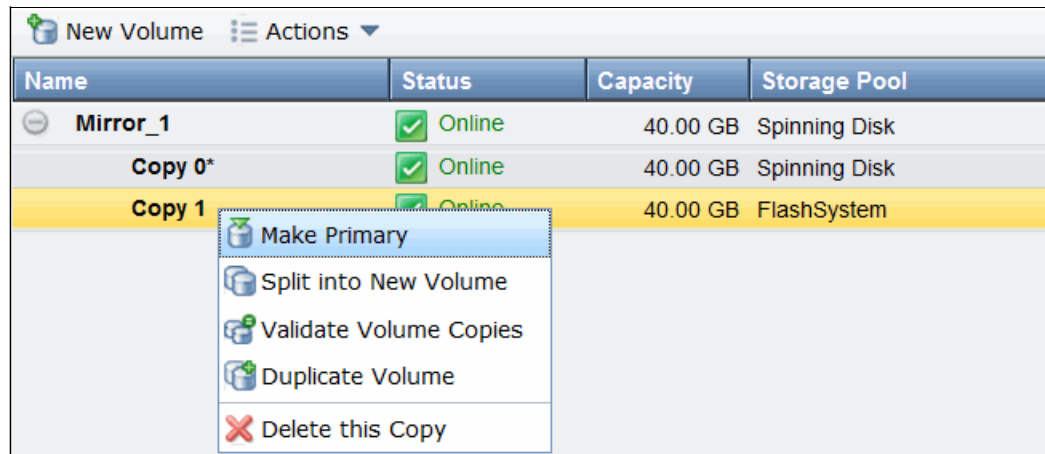
Figure 5-27 shows the mirrored IBM SAN Volume Controller volume with preferred read on a spinning disk.



Name	Status	Capacity	Storage Pool
<b>Mirror_1</b>	Online	40.00 GB	Spinning Disk
Copy 0*	Online	40.00 GB	Spinning Disk
Copy 1	Online	40.00 GB	FlashSystem

Figure 5-27 SAN Volume Controller mirrored VDisk with preferred read on spinning disk

Figure 5-28 shows the option menu to select the primary copy. The primary copy is identical to the preferred read disk.

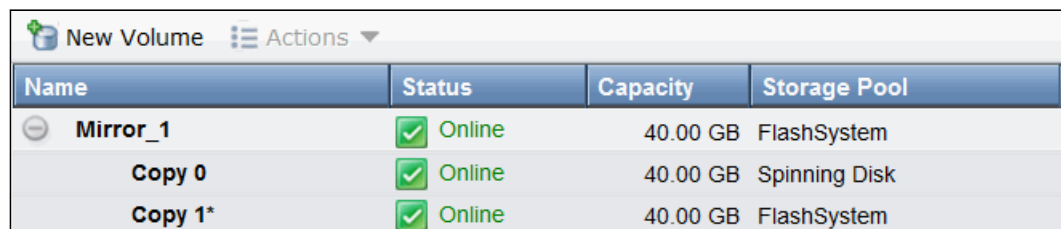


Name	Status	Capacity	Storage Pool
<b>Mirror_1</b>	Online	40.00 GB	Spinning Disk
Copy 0*	Online	40.00 GB	Spinning Disk
Copy 1	Online	40.00 GB	FlashSystem

- Make Primary
- Split into New Volume
- Validate Volume Copies
- Duplicate Volume
- Delete this Copy

Figure 5-28 SAN Volume Controller option Make Primary

Figure 5-29 shows the mirrored IBM SAN Volume Controller volume with preferred read using the FlashSystem 900.



Name	Status	Capacity	Storage Pool
<b>Mirror_1</b>	Online	40.00 GB	FlashSystem
Copy 0	Online	40.00 GB	Spinning Disk
Copy 1*	Online	40.00 GB	FlashSystem

Figure 5-29 SAN Volume Controller mirrored VDisk with preferred read on the FlashSystem 900

## Preferred read with Oracle ASM

Oracle Automatic Storage Management (ASM) in Oracle 11g includes advanced features that can use the performance of the IBM FlashSystem 900. The features of Preferred Mirror Read and Fast Mirror Resync are the two most prominent features that fit in this category.

You can set up preferred read by using the following two features. You can get detailed information about these two features in the Oracle documentation:

- ▶ Preferred Mirror Read
- ▶ Fast Mirror Resync

You can read a detailed guide about Oracle ASM in the “Administering ASM Disk Groups” topic of the *Database Storage Administrator’s Guide* from the Oracle Help Center:

[http://docs.oracle.com/cd/B28359\\_01/server.111/b31107/asmdiskgrps.htm#OSTMG137](http://docs.oracle.com/cd/B28359_01/server.111/b31107/asmdiskgrps.htm#OSTMG137)

## 5.5.3 Linux configuration file multipath.conf example

The name and the value of several `multipath.conf` file attributes are changed from version 5 to version 6, or from version 6 to version 6.2. The example is based on Linux 6.2 and includes commented values for other versions. Check your Linux version to set the correct `wwid` values.

Example 5-23 shows a Linux `multipath.conf` file for the IBM FlashSystem 900.

**Important:** Check for the correct SCSI inquiry string. The actual string is as follows:

"FlashSystem-9840"

*Example 5-23 IBM FlashSystem 840 multipath.conf file for Linux 6.2*

```
# multipath.conf for Linux 5.x and Linux 6.x
#
# Always check for the correct parameter names, other versions may use a different name.
# Check the correct names and values for your Linux environment.
#

defaults {
    udev_dir          /dev
    polling_interval  30
    checker_timeout   10
}

blacklist {
    wwid              "*"
}

blacklist_exceptions {
    wwid              "36005076*"
}

devices {
    device {
        vendor          "IBM"
        product          "FlashSystem-9840"
        path_selector    "round-robin 0"           # Linux 5, Linux 6
        path_selector    "queue-length 0"         # Linux 6.2, if available
        path_grouping_policy multibus
        path_checker      tur
        rr_min_io_rq      4                       # Linux 6.x
    }
}
```

```
#
    rr_min_io          4                # Linux 5.x
    rr_weight          uniform
    no_path_retry      fail
    failback           immediate
    dev_loss_tmo       300
    fast_io_fail_tmo   25
}

}

multipaths {
# Change these example WWID's to match the FlashSystem LUN.s.
    multipath {
        wwid          360050768018e9fc15000000006000000
        alias          FlashSystem_900_6
    }
    multipath {
        wwid          360050768018e9fc15000000007000000
        alias          FlashSystem_900_7
    }
}
}
```

---

## Example of Linux commands to configure the FlashSystem 900

Example 5-24 shows the commands and their results after attaching two 103 GB FlashSystem 900 volumes to the Linux 6.2 host. In this example, a 100 GB FlashSystem 820 LUN was already attached to the system, and the two FlashSystem 900 LUNs are new. The multipath.conf file of Example 5-23 on page 152 must be extended by entries for the IBM FlashSystem 900, which are shown in Example 5-26 on page 155.

*Example 5-24 Commands to create the FlashSystem 840 devices*

---

```
#
# # list current devices
# multipath -l
mpatha (1ATA      SAMSUNG HE161HJ                S209J90S) dm-4 ATA,SAMSUNG
HE161HJ
size=149G features='0' hwhandler='0' wp=rw
`-+- policy='round-robin 0' prio=0 status=active
   `- 0:0:0:0 sda 8:0  active undef running
FlashSystem_820_1 (20020c24001117377) dm-1 IBM,FlashSystem
size=100G features='0' hwhandler='0' wp=rw
`-+- policy='queue-length 0' prio=0 status=active
   |- 5:0:0:1 sdc 8:32  active undef running
   `- 4:0:0:1 sdj 8:144 active undef running

# # check for new devices
# multipath
create: FlashSystem_900_7 (360050768018e9fc15000000007000000) undef IBM,FlashSystem-9840
size=103G features='0' hwhandler='0' wp=undef
`-+- policy='queue-length 0' prio=1 status=undef
   |- 5:0:1:1 sde 8:64  undef ready running
   |- 6:0:0:1 sdg 8:96  undef ready running
   |- 7:0:0:1 sdi 8:128 undef ready running
   `- 4:0:1:1 sdl 8:176 undef ready running
create: FlashSystem_900_6 (360050768018e9fc15000000006000000) undef IBM,FlashSystem-9840
size=103G features='0' hwhandler='0' wp=undef
`-+- policy='queue-length 0' prio=1 status=undef
   |- 6:0:0:0 sdf 8:80  undef ready running
   |- 5:0:1:0 sdd 8:48  undef ready running
   |- 7:0:0:0 sdh 8:112 undef ready running
```

```

~ 4:0:1:0 sdk 8:160 undef ready running

# # list multipath devices
# multipath -l
FlashSystem_900_6 (360050768018e9fc15000000006000000) dm-2 IBM,FlashSystem-9840
size=103G features='0' hwhandler='0' wp=rw
~+- policy='queue-length 0' prio=0 status=active
  |- 6:0:0:0 sdf 8:80 active undef running
  |- 5:0:1:0 sdd 8:48 active undef running
  |- 7:0:0:0 sdh 8:112 active undef running
  ~- 4:0:1:0 sdk 8:160 active undef running
mpatha (1ATA SAMSUNG HE161HJ S209J90S) dm-4 ATA,SAMSUNG
HE161HJ
size=149G features='0' hwhandler='0' wp=rw
~+- policy='round-robin 0' prio=0 status=active
  ~- 0:0:0:0 sda 8:0 active undef running
FlashSystem_820_1 (20020c24001117377) dm-1 IBM,FlashSystem
size=100G features='0' hwhandler='0' wp=rw
~+- policy='queue-length 0' prio=0 status=active
  |- 5:0:0:1 sdc 8:32 active undef running
  ~- 4:0:0:1 sdj 8:144 active undef running
FlashSystem_900_7 (360050768018e9fc15000000007000000) dm-0 IBM,FlashSystem-9840
size=103G features='0' hwhandler='0' wp=rw
~+- policy='queue-length 0' prio=0 status=active
  |- 5:0:1:1 sde 8:64 active undef running
  |- 6:0:0:1 sdg 8:96 active undef running
  |- 7:0:0:1 sdi 8:128 active undef running
  ~- 4:0:1:1 sdl 8:176 active undef running

# # list devices with ls
# ls -l /dev/mapper/
lrwxrwxrwx. 1 root root      7 Oct 16 09:30 FlashSystem_820_1 -> ../dm-1
lrwxrwxrwx. 1 root root      7 Oct 16 09:58 FlashSystem_900_6 -> ../dm-2
lrwxrwxrwx. 1 root root      7 Oct 16 09:58 FlashSystem_900_7 -> ../dm-0

```

---

Example 5-25 shows the creation of an aligned partition in Linux 6.2.

#### *Example 5-25 Creating a Linux partition*

---

```

#
# fdisk /dev/mapper/FlashSystem_900_6

```

The device presents a logical sector size that is smaller than the physical sector size. Aligning to a physical sector (or optimal I/O) size boundary is recommended, or performance may be impacted.

WARNING: DOS-compatible mode is deprecated. It's strongly recommended to switch off the mode (command 'c') and change display units to sectors (command 'u').

```

Command (m for help): u
Changing display/entry units to sectors

Command (m for help): n
Command action
  e   extended
  p   primary partition (1-4)
p
Partition number (1-4): 1
First sector (63-216006655, default 1024): 128

```

Last sector, +sectors or +size{K,M,G} (128-216006655, default 216006655):  
Using default value 216006655

Command (m for help): w  
The partition table has been altered!

Calling ioctl() to reread partition table.  
Syncing disks.

---

## Using FlashSystem 820 and FlashSystem 900 with Linux client hosts

The FlashSystem 820 and IBM FlashSystem 900 have different product names that must be configured in the `multipath.conf` file. FlashSystem 840 and FlashSystem 900 have the same product string ("FlashSystem-9840") and use the same `multipath.conf` configuration. Example 5-26 shows the configuration file for the FlashSystem 820 and FlashSystem 900.

*Example 5-26 Multipath.conf file extension for the FlashSystem 820 and FlashSystem 900*

---

```
# multipath.conf for Linux 5.x and Linux 6.x
#
# Always check for the correct parametername, another version may use a different name.
# Check the correct names and values for your Linux environment
#

defaults {
    udev_dir                /dev
    polling_interval        30
    checker_timeout         10
}

blacklist {
    wwid                    "*"
}

blacklist_exceptions {
    wwid                    "36005076*"          # FlashSystem 900
    wwid                    "20020c24*"          # FlashSystem 710/810/720/820
}

devices {
    # FlashSystem 840 and 900
    device {
        vendor                "IBM"
        product                "FlashSystem-9840"
#        path_selector          "round-robin 0"      # Linux 5, Linux 6
#        path_selector          "queue-length 0"     # Linux 6.2, if available
        path_grouping_policy   multibus
        path_checker           tur
#        rr_min_io_rq           4                    # Linux 6.x
#        rr_min_io              4                    # Linux 5.x
        rr_weight              uniform
        no_path_retry          fail
        failback               immediate
        dev_loss_tmo           300
        fast_io_fail_tmo       25
    }
    # FlashSystem 710/810/720/820
    device {
        vendor                "IBM"
```

```

        product          "FlashSystem"
#       path_selector    "round-robin 0"      # Linux 5, Linux 6
        path_selector    "queue-length 0"     # Linux 6.2, if available
        path_grouping_policy multibus
        path_checker      tur
#       rr_min_io_rq      1                   # 6.x, FlashSystem 710/810
#       rr_min_io         1                   # 5.x, FlashSystem 710/810
        rr_min_io_rq      4                   # 6.x, FlashSystem 720/820
        rr_min_io         4                   # 5.x, FlashSystem 720/820
        rr_weight          uniform
        no_path_retry      fail
        failback           immediate
        dev_loss_tmo       300
        fast_io_fail_tmo   25
    }
}

multipaths {

# Change these example WWID's to match the FlashSystem LUN.s.
    multipath {
        wwid          360050768018e9fc15000000006000000
        alias          FlashSystem_900_6
    }
    multipath {
        wwid          360050768018e9fc15000000007000000
        alias          FlashSystem_900_7
    }
}

```

---

## Linux tuning

The Linux kernel buffer file system writes data before it sends the data to the storage system. With the IBM FlashSystem 900, better performance can be achieved when the data is not buffered but is directly sent to the IBM FlashSystem 900. When setting the scheduling policy to no operation (NOOP), the fewest CPU instructions possible are used for each I/O. Setting the scheduler to NOOP gives the best write performance on Linux systems. You can use the following setting in most Linux distributions as a boot parameter:

```
elevator=noop
```

Current Linux devices are managed by the device manager *Udev*. You can define how Udev will manage devices by adding rules to the `/etc/udev/rules.d` directory.

Example 5-27 shows the rules for the IBM FlashSystem 900 with Linux 6. This udev rules file contains two lines originally. It is divided into multiple lines for better readability. If you use this file, be sure that each line starts with the keyword `ACTION`.

*Example 5-27 Linux device rules Linux 6.x*

```

#
cat 99-IBM-FlashSystem.rules

ACTION=="add|change", SUBSYSTEM=="block",ATTRS{device/model}=="FlashSystem-9840",
ATTR{queue/scheduler}="noop",ATTR{queue/rq_affinity}="1",
ATTR{queue/add_random}="0",ATTR{device/timeout}="5"

ACTION=="add|change", KERNEL=="dm-*",
PROGRAM="/bin/bash -c 'cat /sys/block/$name/slaves/*/device/model | grep FlashSystem-9840'",
ATTR{queue/scheduler}="noop",ATTR{queue/rq_affinity}="1",ATTR{queue/add_random}="0"

```

---



Example 5-28 shows the rule for the IBM FlashSystem 900 with Linux 5. This udev rules file contains one line originally. It is divided into multiple lines for better readability. If you use this file, be sure that the line starts with the keyword ACTION.

*Example 5-28 Linux device rules Linux 5.x*

---

```
#
cat 99-IBM-FlashSystem.rules

ACTION=="add|change", SUBSYSTEM=="block",SYSFS{model}=="FlashSystem-9840",
RUN+="/bin/sh -c 'echo noop > /sys/$DEVPATH/queue/scheduler'"
```

---

You can apply the new rules by using the commands in Example 5-29.

*Example 5-29 Restarting udev rules*

---

```
# linux 6.2
/sbin/udevadm control --reload-rules
/sbin/start_udev
```

---

### 5.5.4 Example of a VMware configuration

You can set the number of I/Os for each path on VMware with this command, which sets 10 I/Os for each path:

```
esxcli nmp roundrobin setconfig --device <device> --iops=10 --type "iops"
```

## 5.6 FlashSystem 900 and Easy Tier

You can implement the IBM FlashSystem 900 with IBM SAN Volume Controller Easy Tier. SAN Volume Controller Easy Tier automatically moves hot, frequently used data to the FlashSystem 900 and cold less frequently or never used data to the traditional disk system. An example of implementing SAN Volume Controller Easy Tier is shown in Figure 5-30 on page 158.

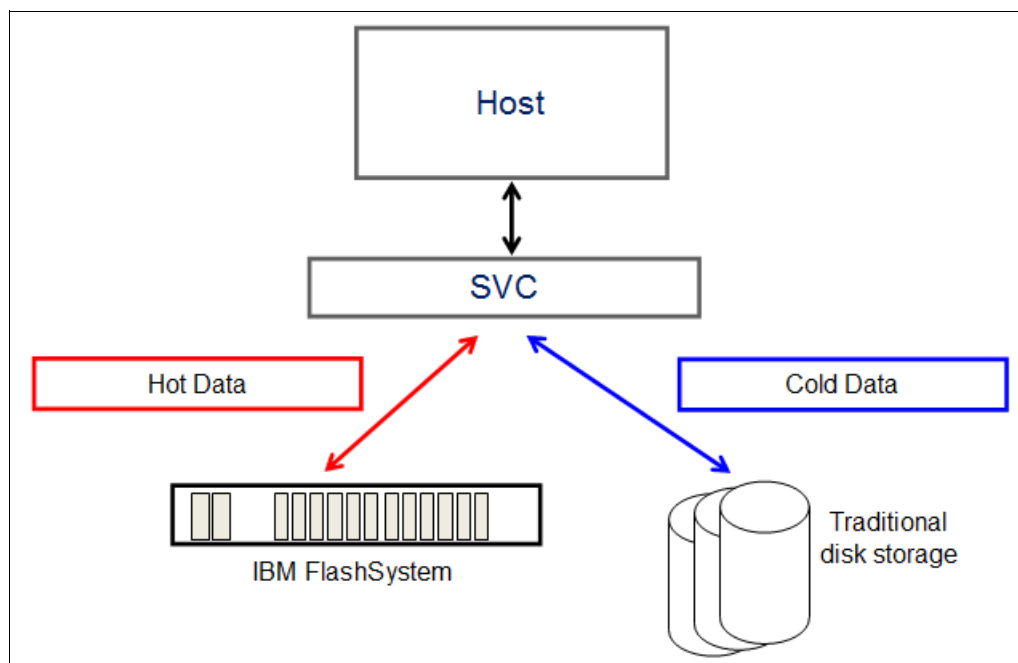


Figure 5-30 Easy Tier with SAN Volume Controller

For more details about the IBM FlashSystem 900 and SAN Volume Controller solution, see these sections:

- ▶ 8.1.4, “IBM Spectrum Virtualize - SAN Volume Controller advanced functionality” on page 288
- ▶ 9.2, “Tiering” on page 313

## 5.7 Troubleshooting

Troubleshooting information for issues that you might encounter when configuring InfiniBand or creating file systems is described in this section.

### 5.7.1 Troubleshooting Linux InfiniBand configuration issues

The following list describes potential Linux configuration issues, troubleshooting guidance, and resolutions:

- ▶ When you install `OFED-X.X.X`, an error occurs, which indicates that you failed to build the `ofa_kernel` RPM Package Manager (originally Red Hat Package Manager).

The kernel that is used by the server might not be supported by OpenFabrics Enterprise Distribution (OFED). If the Install All option was chosen, try the Customize option in the OFED installation menu and select only the components that are needed. If this process does not work, try installing a different version of OFED.

- ▶ Loading the driver module fails.

The HCA used might not be supported by OFED, or the driver was not installed correctly. Obtain the latest drivers for the host channel adapters (HCAs) from the HCA vendor’s website.

- ▶ When you try to install OFED, an error message is displayed:

```
"<module 1> is required to build <module 2>"
```

This error means that certain dependencies that are required by OFED are not installed on the server. You must install all of the required dependencies:

- To search for the necessary RPM (if the **yum** package-management tool is available), enter this command:

```
# yum provides <dependency_name>
```

- To install the RPM, enter this command:

```
# yum install <dependency_rpm>
```

**Note:** If **yum** is not installed on the server, each dependency must be manually downloaded and installed.

- ▶ When you try to run the **srp\_daemon** command, a message indicates that an operation failed. Check these items:
  - Make sure that the storage system is physically connected to the network and that all components are powered on.
  - Make sure that the correct cable is used and that OpenSM is running. To confirm whether OpenSM is running, enter this command:

```
# /etc/init.d/opensmd status
```

- ▶ Loading the **ib\_srp** module fails.

Verify that OFED is installed correctly and that the necessary device drivers are also installed. If a custom OFED installation was done, be sure that **ibutils** and all packages that are related to **srp** were selected.

## 5.7.2 Linux fdisk error message

You might receive the following messages when you create a partition in Linux:

```
"Re-reading the partitioning table failed"
```

Also, the corresponding device is not created in the **/dev/mapper** directory. Solve this problem by issuing the **partprobe** command, as depicted in Example 5-30.

*Example 5-30 Solving the partition table failed with error 22*

---

```
[root@localhost ~]# fdisk /dev/mapper/FlashSystem_900_3
```

```
<partition creation lines left out for clarity>
```

```
...
```

```
Command (m for help): w
```

```
The partition table has been altered!
```

```
Calling ioctl() to re-read partition table.
```

```
WARNING: Re-reading the partition table failed with error 22: Invalid argument.
The kernel still uses the old table. The new table will be used at
the next reboot or after you run partprobe(8) or kpartx(8)
Syncing disks.
```

```
[root@localhost ~]# ls -l /dev/mapper/
lrwxrwxrwx. 1 root root      7 Oct  8 08:10 FlashSystem_900_2 -> ../dm-0
lrwxrwxrwx. 1 root root      7 Oct  8 09:20 FlashSystem_900_3 -> ../dm-2

[root@localhost ~]# partprobe

[root@localhost ~]# ls -l /dev/mapper/
lrwxrwxrwx. 1 root root      7 Oct  8 08:10 FlashSystem_900_2 -> ../dm-0
lrwxrwxrwx. 1 root root      7 Oct  8 09:20 FlashSystem_900_3 -> ../dm-2
brw-rw----. 1 root disk 253,  7 Oct  8 09:21 FlashSystem_900_3p1
```

---

You can search the `/dev/mapper` directory for newly generated partitions. After using the **partprobe** command, the new partition is generated.

### 5.7.3 Changing FC port properties

The FlashSystem 900 automatically detects the SAN topology and speed. If you want to set the speed, for example, 16 Gbps, or the topology, such as arbitrated loop or fabric explicitly, use the **chportfc** command. The **lspportfc** command lists the current settings.

**Note:** The FlashSystem 900 16 Gbps FC attachment uses fabric topology only.

### 5.7.4 Changing iSCSI port properties

The FlashSystem 900 uses 10 Gb Ethernet for iSCSI connections. Use the **chportip** command to set the IP address, netmask, and gateway values of the iSCSI ports. The **lspportip** command lists the current settings.



## Using IBM FlashSystem 900

This chapter shows how to operate IBM FlashSystem 900 in your business environment. Both the graphical user interface (GUI) and the command-line interface (CLI) are used to demonstrate how to monitor the system and work with volumes, hosts, and user security.

This chapter covers the following topics:

- ▶ Overview of IBM FlashSystem 900 management tools
- ▶ Monitoring menu
- ▶ Volumes menu
- ▶ Hosts menu
- ▶ Access menu

## 6.1 Overview of IBM FlashSystem 900 management tools

The FlashSystem 900 can be managed from either the built-in GUI, which is a web-based management tool, or from the CLI.

The web-based GUI is designed to simplify storage management and to provide a fast and more efficient management tool. It is based on the IBM System Storage XIV software and has a similar look and behavior.

To use the GUI to manage the FlashSystem 900, be sure to have a supported web browser. For a list of supported web browsers, see “Supported web browsers” on page 86.

**JavaScript:** You might need to enable JavaScript in your browser. Additionally, if you are using Firefox, under Advanced JavaScript Settings, click **Disable or replace context menus** and allow cookies.

### 6.1.1 Access to the graphical user interface

To log on to the GUI, point your web browser to the management IP address that was set during the initial setup of the FlashSystem 900. The default credentials are as follows:

- ▶ User name: superuser
- ▶ Password: passw0rd (with a zero in place of the letter “O”)

Figure 6-1 shows the login window.



Figure 6-1 Login window

After logging in, the Monitoring > System home window opens (Figure 6-2).



Figure 6-2 System overview window

### 6.1.2 Graphical user interface (GUI) layout

The GUI has three main sections for navigating through the management tool (Figure 6-3 on page 164):

- ▶ Function icons (left side), also referred to as menu selection icons
- ▶ Status bars (bottom)
- ▶ Actions menu (upper left or right-click in the home window)

At the left side of the window are five *function icons*. The five function icons represent these areas:

- ▶ Monitoring menu
- ▶ Volumes menu
- ▶ Hosts menu
- ▶ Access menu
- ▶ Settings menu

Also in the window are these controls:

- ▶ Actions menu (upper left or right-click in the home window): For the list of actions on this menu, see Figure 6-4 on page 165.
- ▶ Current menu indicators: Shows the path where the menu is located (as shown in Figure 6-3 on page 164, IBM FlashSystem 900 > Monitoring > System).
- ▶ User security: This key is for managing the security of the user that is currently logging in.
- ▶ The Help icon (question mark (?)): Provides information about licenses and gives access to the IBM FlashSystem 900 web page at the IBM Knowledge Center.

- ▶ Menu selector icons: These are also referred to as *function icons*.
- ▶ Capacity indicators: Figure 6-3 shows 5.20 TiB of physical Flash module space and 2.98 TiB of allocated space. The bottom left shows capacity with 3,048 used GiB of 5,325 GiB installed capacity, with 57% used.
- ▶ Performance: This is also referred to as Performance pod; it displays bandwidth (MBps), I/O per second (IOPS), and latency (ms).
- ▶ Health Indicator: Shows the health of the system. Figure 6-3 indicates that the system is healthy (green).
- ▶ Component status indicators: These show the status of the components of the FlashSystem 900. See Figure 6-4 on page 165 for more details.
- ▶ Rotate for rear side view: Rotates the FlashSystem 900 to see a rear view of the components.

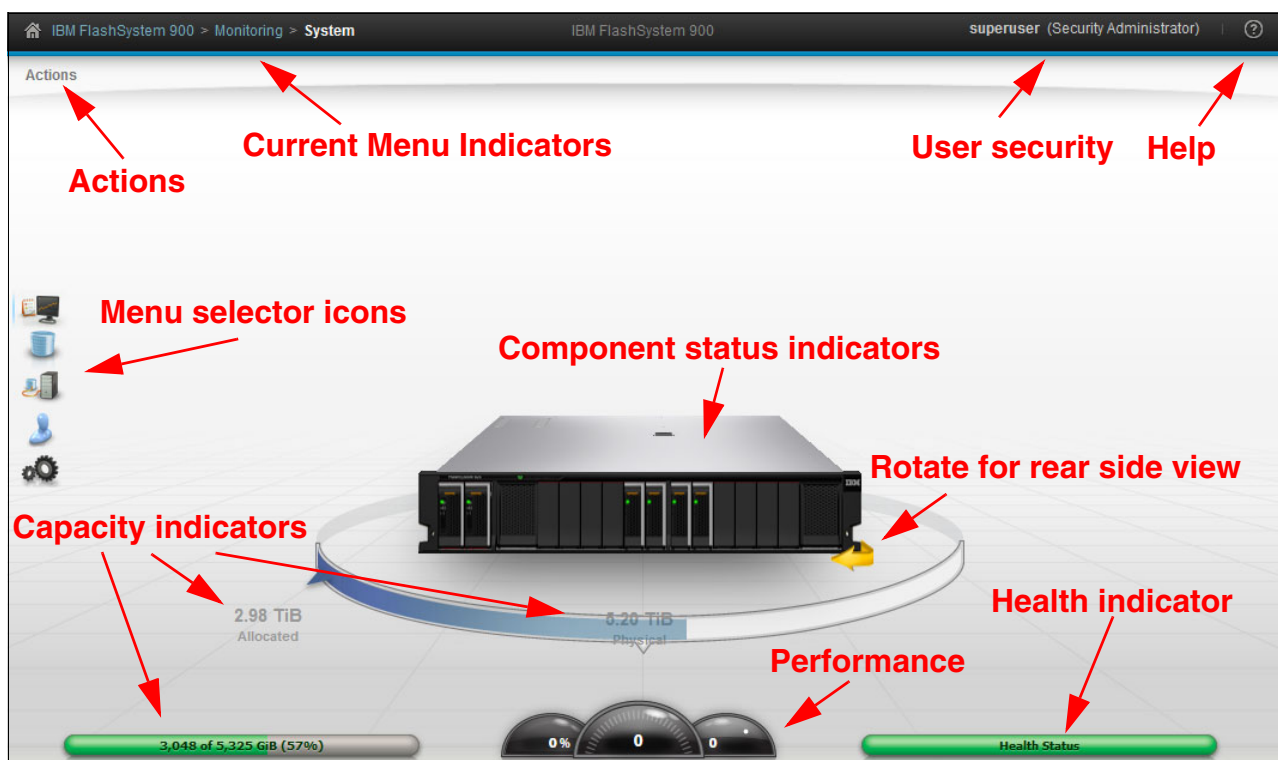


Figure 6-3 Home window: Main areas



## 6.1.3 Navigation

Navigating the management tool is simple. You can hover the cursor over one of the five function icons on the left side of the window, which highlights the function icon and opens a list of options. Figure 6-4 shows a list of the FlashSystem 900 software function icons and the associated menu options.

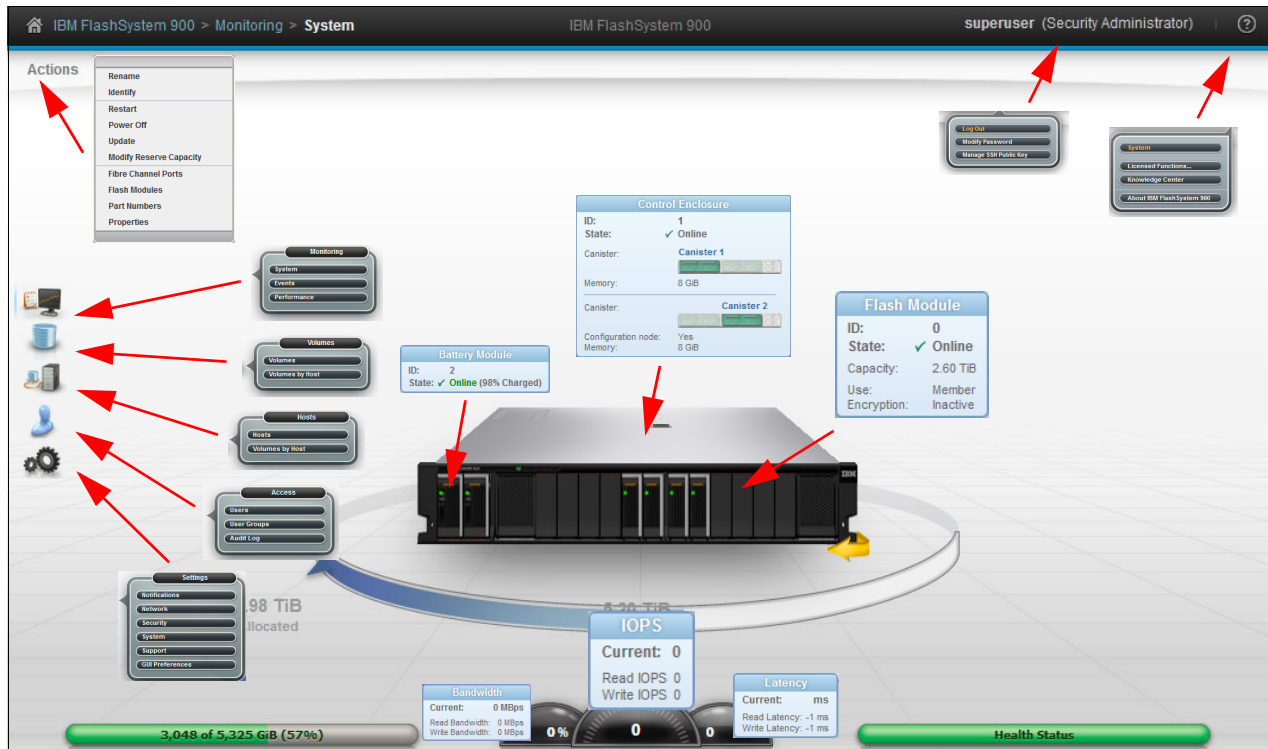


Figure 6-4 FlashSystem 900: Menu options

The following sections describe the five function icons and their associated menu options.

Figure 6-5 shows the home window with a rear view of the FlashSystem 900. To see this view click the **Rotate** icon located at the right side of the image.

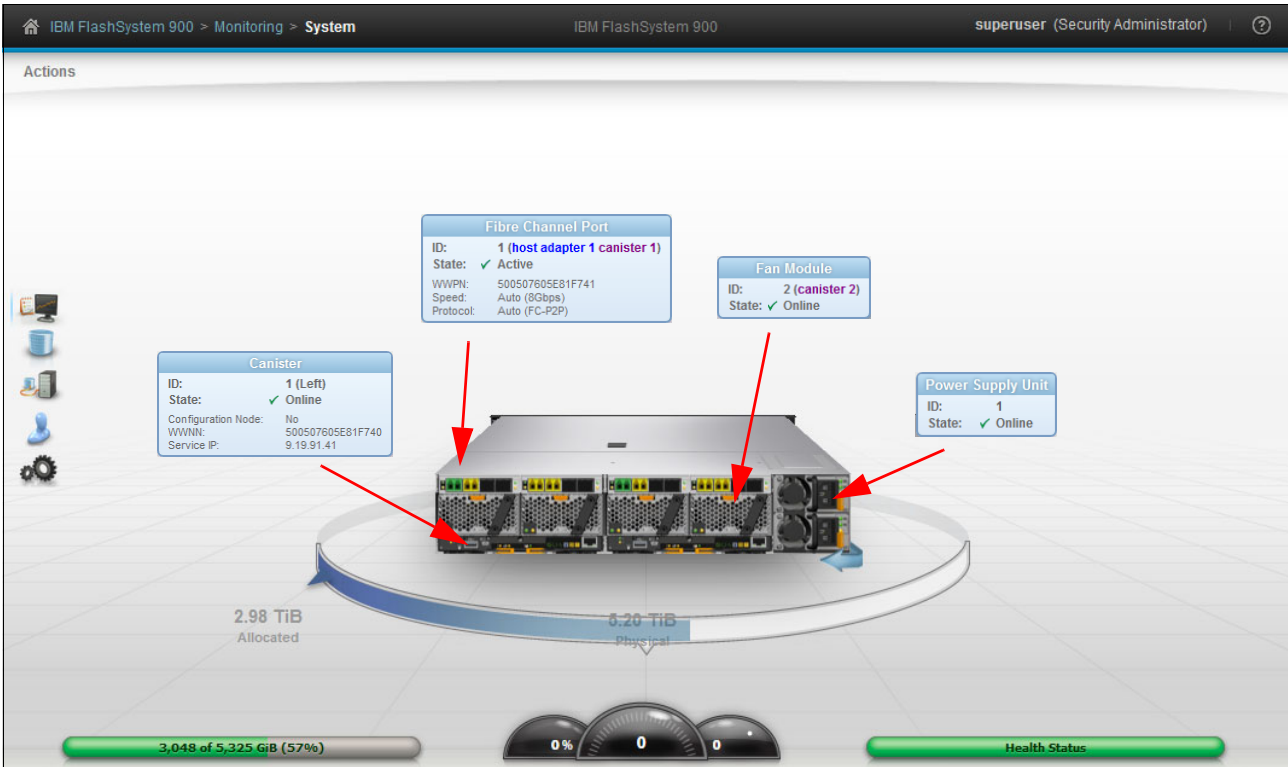


Figure 6-5 FlashSystem 900: Menu options Rear view

### 6.1.4 Selecting multiple items

With the FlashSystem 900 management tool, you can select multiple items by using a combination of the Shift keys or Ctrl keys. To select multiple items in a display, click the first item, press and hold the Shift key, and click the last item in the list that you require. All the items in between those two items are selected. For example, Figure 6-6 illustrates multiple selections from the **Volumes** → **Volumes** menu.

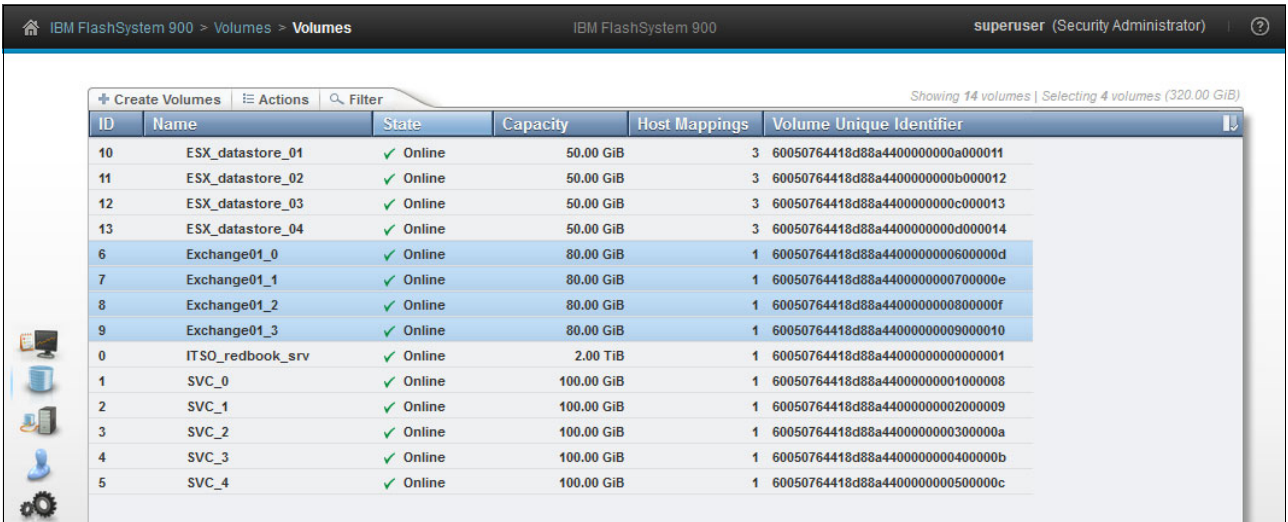


Figure 6-6 Multiple selections by using the Shift key

This function is useful if you want to expand multiple volumes at the same time.

If you want to select multiple items that are not in sequential order, click the first item, press and hold the Ctrl key, and click the other items that you require (Figure 6-7).

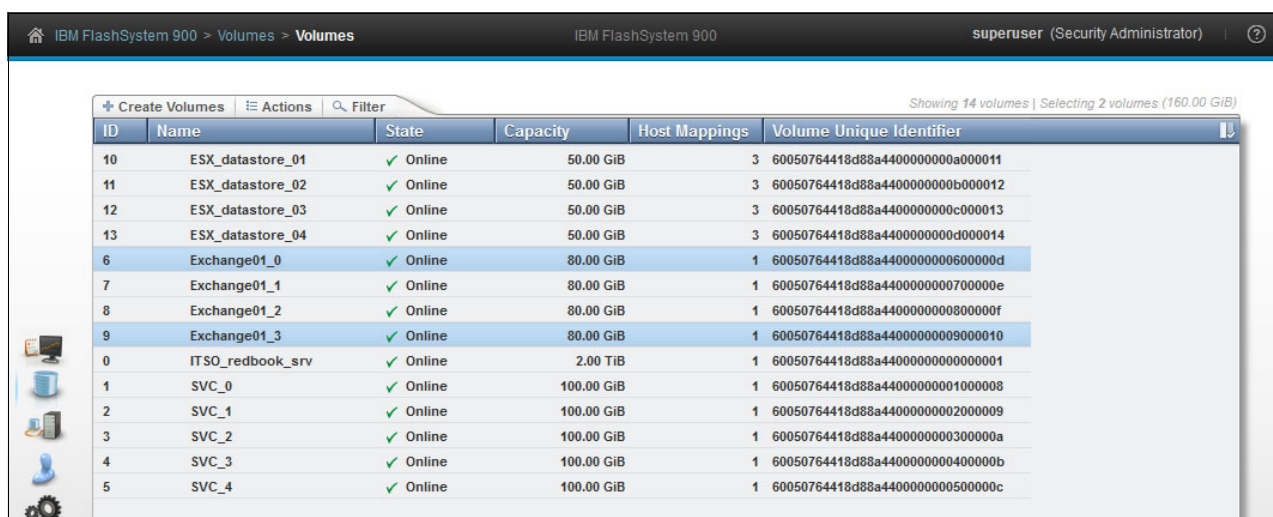


Figure 6-7 Multiple selections by using the Ctrl key

## 6.1.5 Status indicators

Other useful tools are the Status indicators that appear at the bottom of the window (Figure 6-8). These indicators provide information about capacity usage, throughput in megabytes per second and I/O per second (IOPS), and the health status of the system. The status indicators are visible from all windows in the FlashSystem 900 GUI.



Figure 6-8 Status indicators at the bottom of the GUI window

The status indicators show the following information:

- ▶ Capacity:
  - Used gigabytes
  - Installed gigabytes
  - Percentage used
- ▶ Performance:
  - Bandwidth (MBps)
  - I/O per second (IOPS)
  - Latency
- ▶ Health; one of these options:
  - Healthy (green)
  - Warning (yellow) and a link to Monitoring → Events is provided
  - Error (red) and a link to Monitoring → Events is provided
  - Upgrade status percentage

## 6.2 Monitoring menu

The **Monitoring** → **System** menu is the default menu and home page for the FlashSystem 900 GUI. It has three options as shown in Figure 6-4 on page 165:

- ▶ System
- ▶ Events
- ▶ Performance

Part of the default window or home page is also the Actions menu where the system can be managed and information about the system can be obtained.

### 6.2.1 Monitoring System menu

In the home window of the FlashSystem 900 GUI, you can select Actions in the upper-left corner. Actions can also be activated by right-clicking anywhere in the GUI.

#### Actions menu

The following information can be retrieved in the Properties window by selecting **Actions** → **Properties** (Figure 6-9 on page 169):

- ▶ System name
- ▶ System state
- ▶ Hardware type
- ▶ Firmware version
- ▶ Serial number
- ▶ Model and type number
- ▶ Worldwide name
- ▶ Storage capacity
- ▶ Power on days

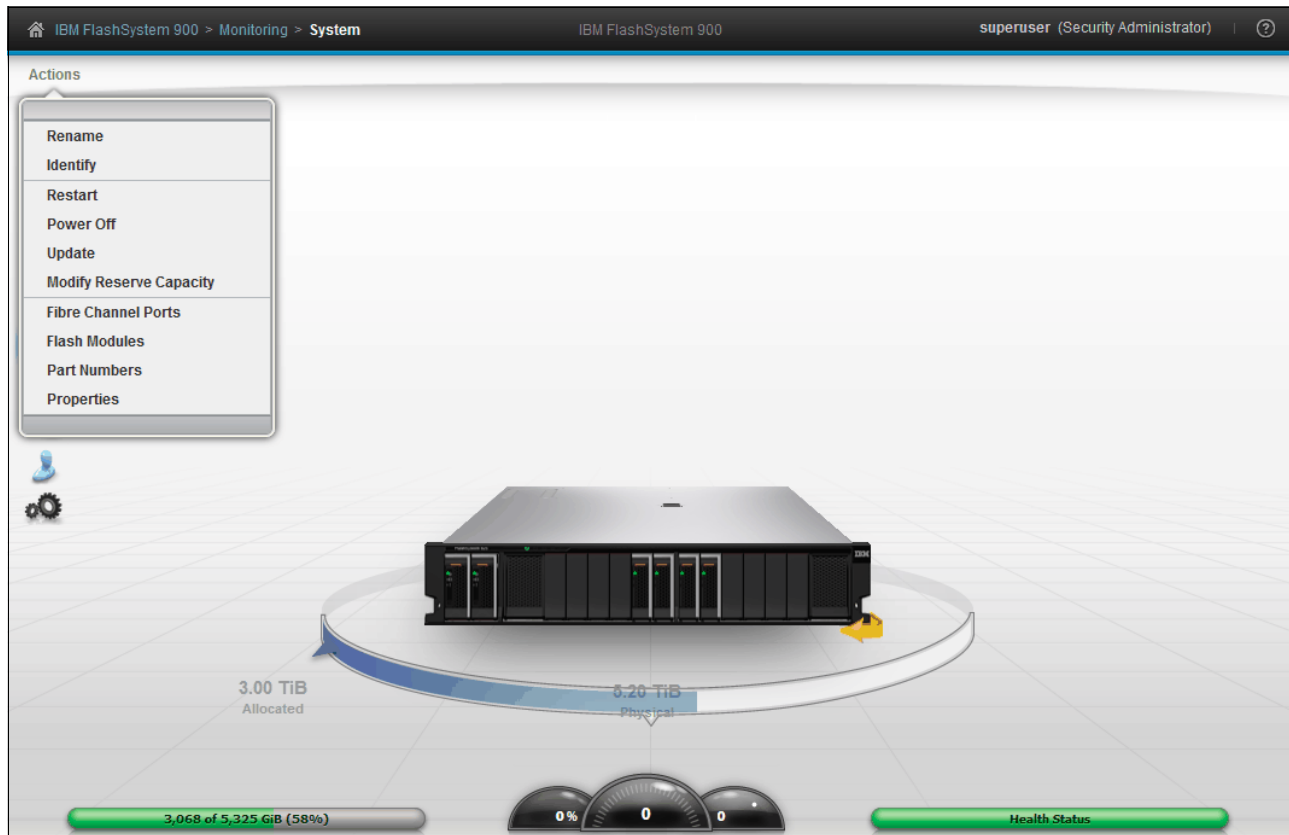


Figure 6-9 Monitoring System menu with actions displayed

The Properties window opens and shows the system's properties (Figure 6-10).

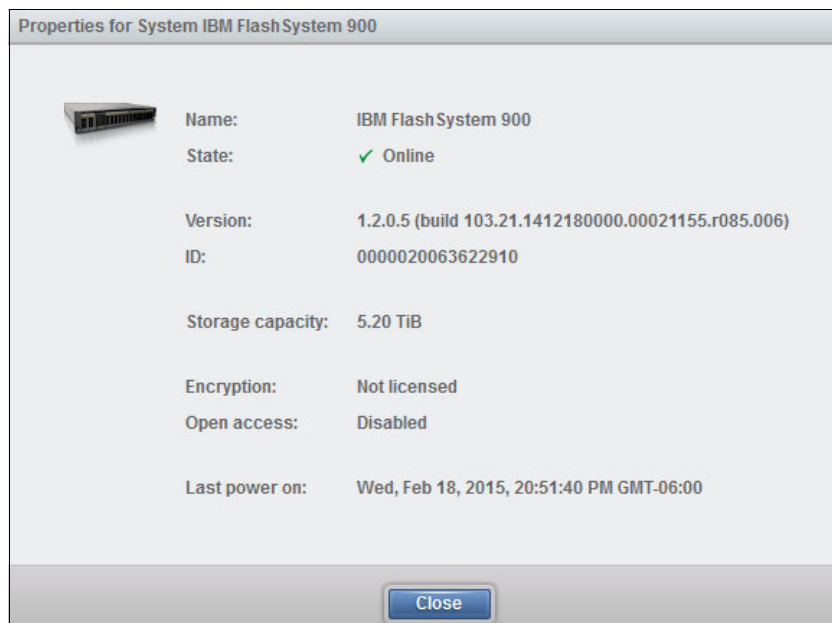


Figure 6-10 Properties for the system cluster

## Rename system

If you want to change the system name, click **Actions** → **Rename** to open the Rename System dialog (Figure 6-11). Enter a new name and click **Rename**.

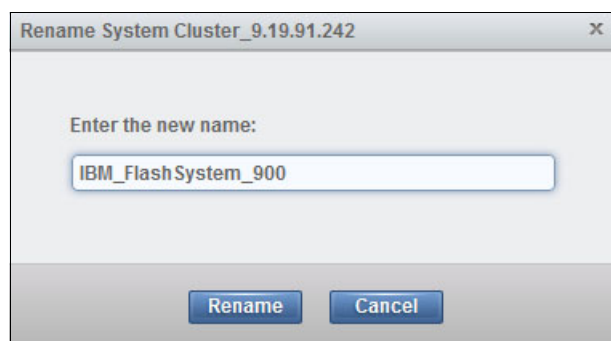


Figure 6-11 Rename the system

When the system is renamed, the task window (Modify System Properties) opens (Figure 6-12). It shows the CLI command that the system uses to make the change.

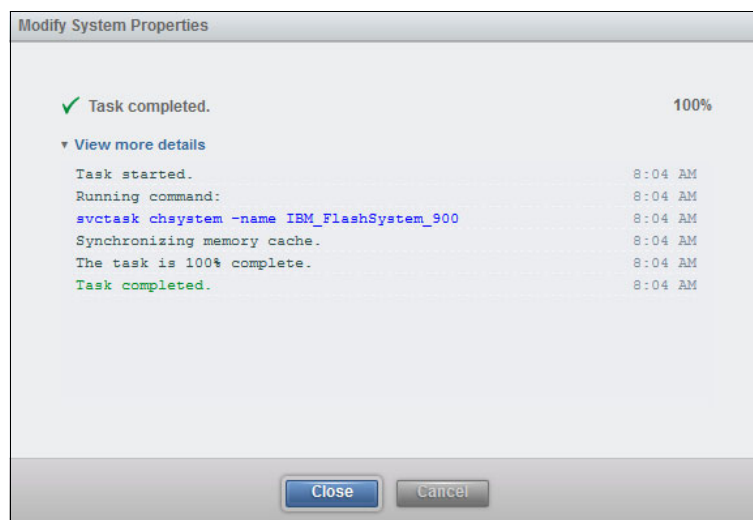


Figure 6-12 CLI command displays

The CLI commands that are displayed can also be executed by the user from within an open CLI window by using PuTTY or a similar terminal emulation tool. For more information about how to use the CLI, see 6.5.3, “Accessing CLI by using PuTTY” on page 223.

## Rename system using CLI

When system properties and settings are changed from the GUI, commands are executed on the system. In the preceding example, you renamed the system host name by using the GUI, and the Modify System Properties window opened (Figure 6-12). In that window, the CLI commands that the system uses to change system properties are displayed.

Example 6-1 shows the use of the CLI command to change system properties. The output is shortened for clarity.

*Example 6-1 Change the system name using the CLI*

---

```
IBM_FlashSystem:Cluster_9.xx.xx.xx:superuser>svctask chsystem -name
IBM_FlashSystem_900
IBM_FlashSystem:Cluster_9.xx.xx.xx:superuser>svcinfo lssystem
id 0000020063622910
name IBM_FlashSystem_900
```

---

In the example, the prompt shows the new system name at the next CLI login.

## Identify LED

Another function of the Actions menu is the Identify function. When the Identify function is enabled, the Identify LED on the front side of the IBM FlashSystem 900 and both controller canisters turn on their blue Identify LED. The canisters are mounted from the rear side of the FlashSystem 900; the canister Identify LEDs are shown from the rear side of the unit.

Figure 6-13 shows the Identify LED when it is on.

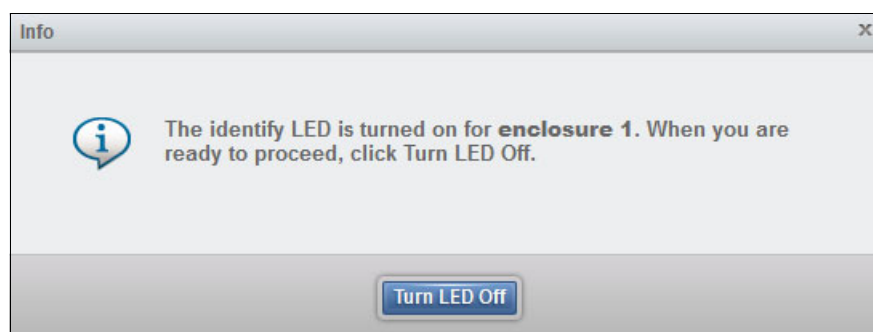


Figure 6-13 Identify LED is turned on

Also, each canister can be identified through the IBM FlashSystem 900 Service Assistant Tool. The Service Assistant Tool is described in 7.2, “Service Assistant Tool” on page 276.

## Power off

The IBM FlashSystem 900 can be turned off through the Actions menu. Many reasons exist to turn off the unit. One reason might be that you need to relocate the system to another site or shut it down for scheduled power maintenance. The power off function ensures that the system is turned off securely so that data is preserved.

When you click **Actions** → **Power Off**, a warning message displays. Figure 6-14 on page 172 shows that the Power Off window requires the administrator to type a confirmation code to prevent an accidental power-off cycle of the device.

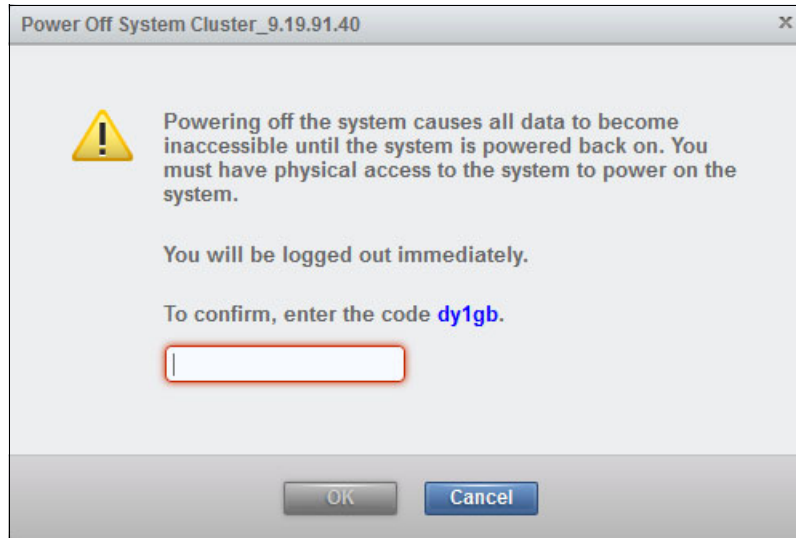


Figure 6-14 Power off entire system or a single canister

Individual controller canisters can be rebooted or placed into service state from the Service Assistant Tool under guidance from IBM Support. One reason for placing a controller into service state might be if the system is to be reinitialized or if a controller is to be replaced.

For more information about the Service Assistant Tool, see 7.2, “Service Assistant Tool” on page 276.

## Modify capacity reservation

You can modify the amount of flash storage capacity that is reserved for system management tasks, which may improve system performance.

Out of the usable flash storage capacity that is available, the system sets aside a certain amount of that capacity for system management tasks, which reduces the overall usable capacity that is available for volumes.

Select **Reserve capacity for performance** and click **Modify** to allocate a portion of the usable capacity to enhance system performance as shown in Figure 6-15.

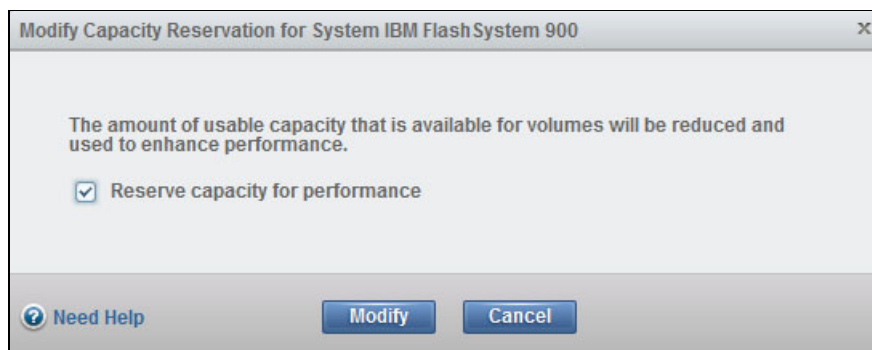


Figure 6-15 Modify capacity reservation

The amount of usable capacity that is available for volumes will be reduced and used to enhance performance.



## Fibre Channel ports

To view the Fibre Channel ports of the IBM FlashSystem 900 interface cards, select **Actions** → **Fibre Channel Ports**. The Fibre Channel Ports window opens.

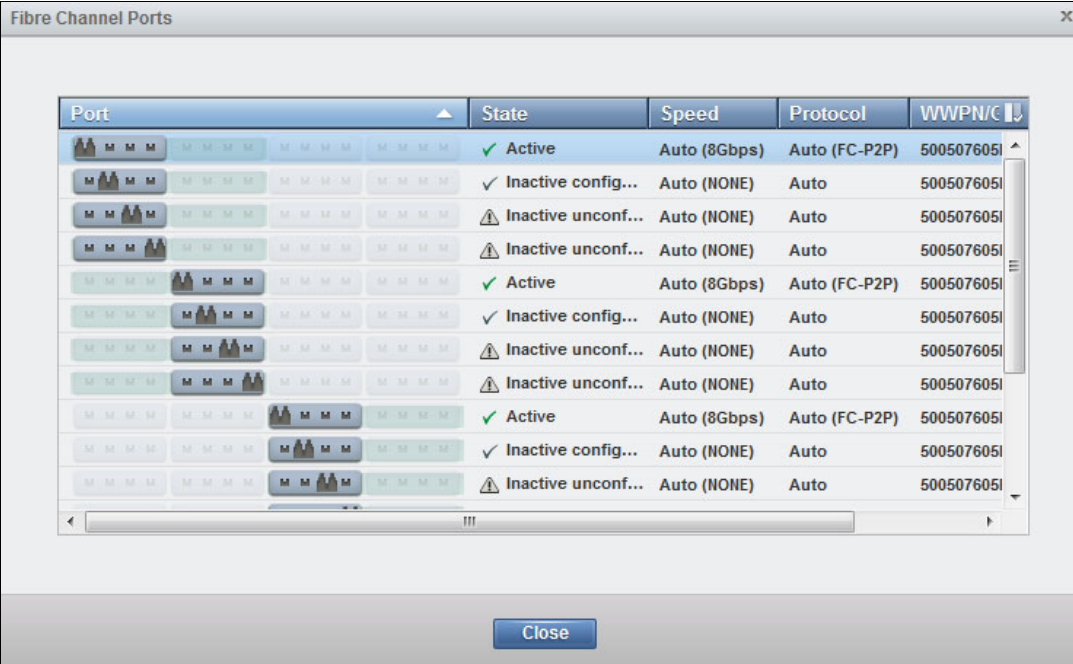
### Fibre Channel and FCoE

In this example, eight 16 Gbps Fibre Channel (FC) ports are displayed, each showing the following information:

- Depiction of port location
- State
- Port: Fibre Channel (FC) or Fibre Channel over Ethernet (FCoE)
- Port speed: Auto, 16 Gbps, 8 Gbps, 4 Gbps, or 2 Gbps
- WWPN/GUID: Worldwide port name (WWPN) and globally unique identifier (GUID)
- Topology: Fibre Channel Arbitrated Loop (FC-AL) or Fibre Channel-Peer to Peer (FC-P2P)

Topology and speed for the ports are automatically detected for FlashSystem 900.

Figure 6-16 shows the Fibre Channel I/O ports in the system.



Port	State	Speed	Protocol	WWPN/GUID
	✓ Active	Auto (8Gbps)	Auto (FC-P2P)	500507605I
	✓ Inactive config...	Auto (NONE)	Auto	500507605I
	⚠ Inactive unconf...	Auto (NONE)	Auto	500507605I
	⚠ Inactive unconf...	Auto (NONE)	Auto	500507605I
	✓ Active	Auto (8Gbps)	Auto (FC-P2P)	500507605I
	✓ Inactive config...	Auto (NONE)	Auto	500507605I
	⚠ Inactive unconf...	Auto (NONE)	Auto	500507605I
	⚠ Inactive unconf...	Auto (NONE)	Auto	500507605I
	✓ Active	Auto (8Gbps)	Auto (FC-P2P)	500507605I
	✓ Inactive config...	Auto (NONE)	Auto	500507605I
	⚠ Inactive unconf...	Auto (NONE)	Auto	500507605I

Figure 6-16 The status of the I/O ports

Any FC or FCoE port that is not connected and online has an Inactive status. The FC-P2P (point-to-point) topology is used in situations where two FC ports connect directly to each other. FC-P2P is the default for a host that is directly connected to the FlashSystem 900, and FC-P2P is also used for a FlashSystem 900 FC port that is connected to a Fibre Channel switch.

The FC-AL (arbitrated loop) topology is also used to attach a host directly to the FlashSystem 900 in cases where the host only supports FC-AL, for example, when connecting a VMware ESX server directly to the FlashSystem 900.

**Note:** FC-AL is not supported for ports that are connected at 16 Gbps.

Figure 6-17 shows the logical numbering of the FC ports.

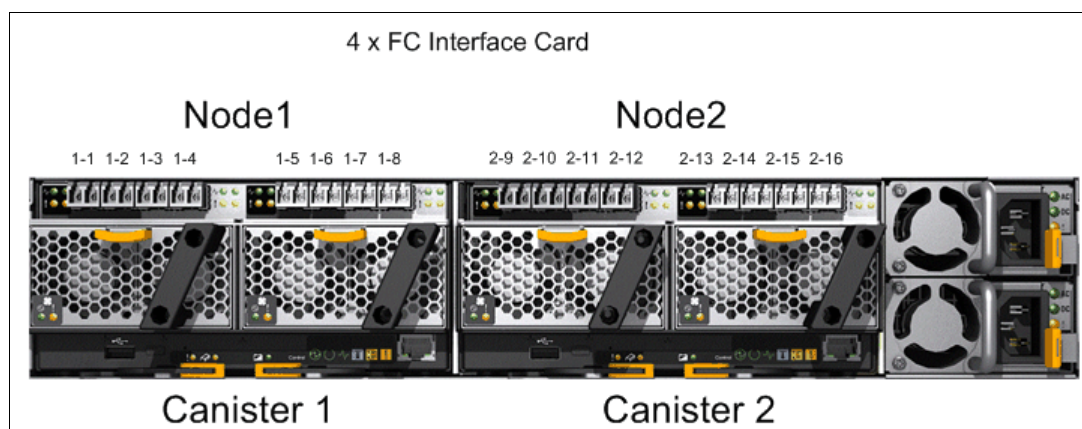


Figure 6-17 I/O port logical numbering

For a description of the physical numbering of the ports of the interface cards, see Chapter 4, “Installation and configuration” on page 67.

**Note:** The physical numbering and the logical numbering of the ports are not the same. The physical numbering of the ports of the interface cards starts from the left with P1 and P4 to the right, depending on the configuration. Logical numbering depends on the node name and these node names can swap, so the port names depicted in Figure 6-17 are the default names, but they might change.

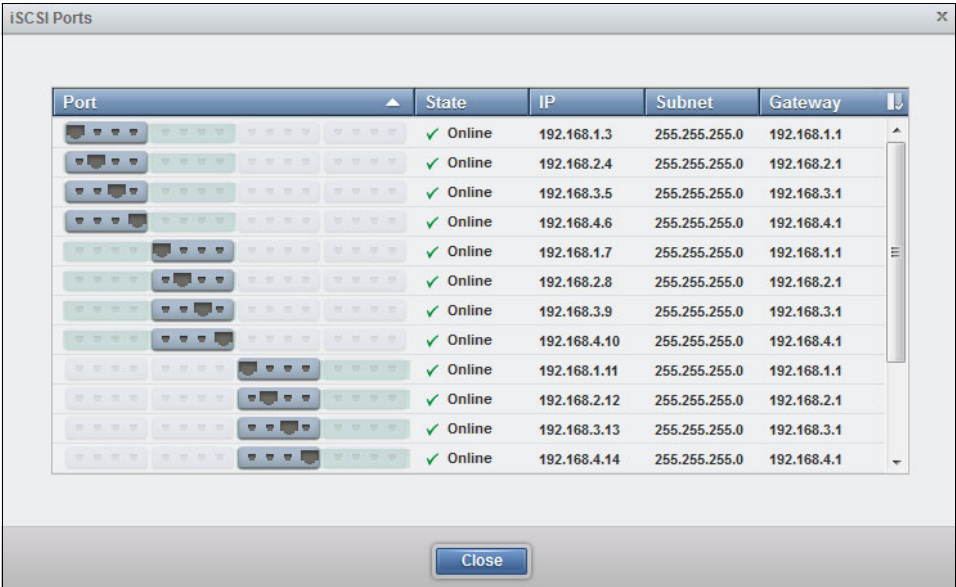
For more details, see 7.2, “Service Assistant Tool” on page 276.

### ***iSCSI***

In the FlashSystem 900 GUI home window, click **Actions** → **iSCSI Ports**. Sixteen iSCSI ports are displayed, each indicating the following information:

- ▶ Depiction of port location
- ▶ State
- ▶ IP address
- ▶ Subnet mask
- ▶ Gateway

The result is shown in Figure 6-18.



Port	State	IP	Subnet	Gateway
	✓ Online	192.168.1.3	255.255.255.0	192.168.1.1
	✓ Online	192.168.2.4	255.255.255.0	192.168.2.1
	✓ Online	192.168.3.5	255.255.255.0	192.168.3.1
	✓ Online	192.168.4.6	255.255.255.0	192.168.4.1
	✓ Online	192.168.1.7	255.255.255.0	192.168.1.1
	✓ Online	192.168.2.8	255.255.255.0	192.168.2.1
	✓ Online	192.168.3.9	255.255.255.0	192.168.3.1
	✓ Online	192.168.4.10	255.255.255.0	192.168.4.1
	✓ Online	192.168.1.11	255.255.255.0	192.168.1.1
	✓ Online	192.168.2.12	255.255.255.0	192.168.2.1
	✓ Online	192.168.3.13	255.255.255.0	192.168.3.1
	✓ Online	192.168.4.14	255.255.255.0	192.168.4.1

Figure 6-18 Display iSCSI I/O ports

To display the IP address of individual iSCSI ports, first rotate the depiction of the FlashSystem 900 by clicking the **Rotate** icon. Then, hover the mouse cursor over the preferred iSCSI port, right-click, and select **Properties** (Figure 6-19).

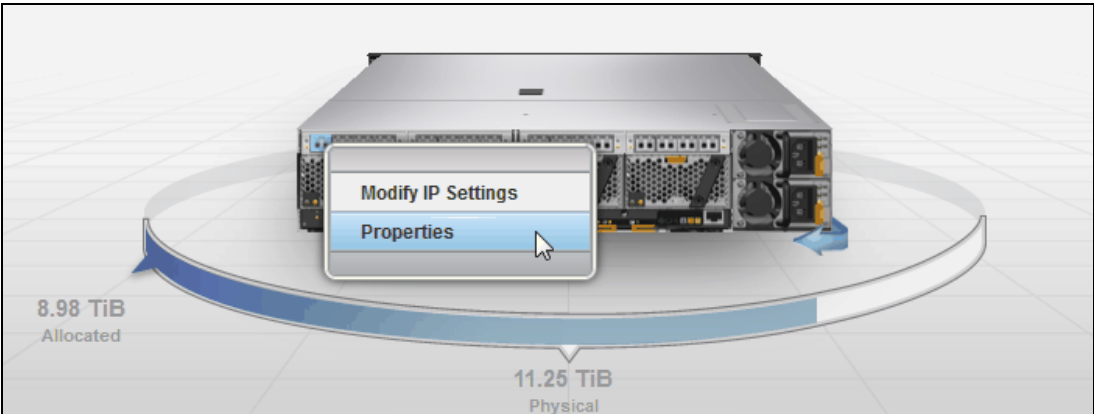


Figure 6-19 Display properties of an iSCSI port

The Properties window of the selected iSCSI port opens and you can review the network address (Figure 6-20).

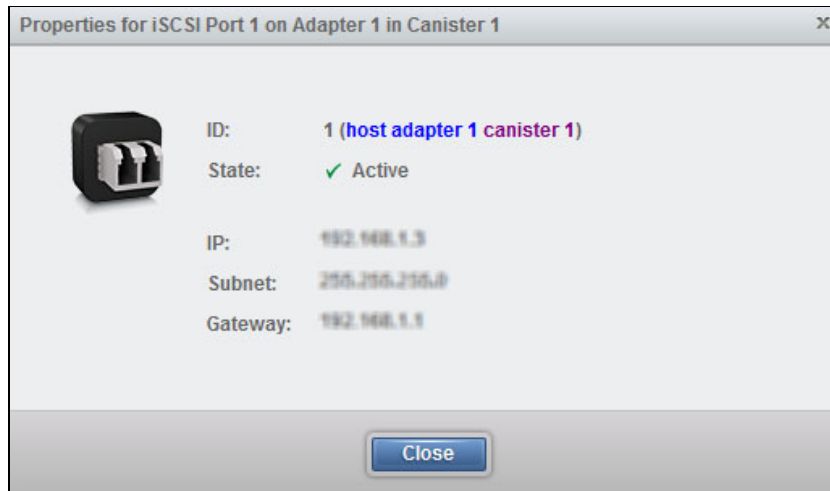


Figure 6-20 Review iSCSI IP address of a single port

To modify an IP address of the selected port instead, right-click and select **Modify IP Settings**. The Modify IP Settings window opens (Figure 6-21). Modify the information you want and click **Modify**.

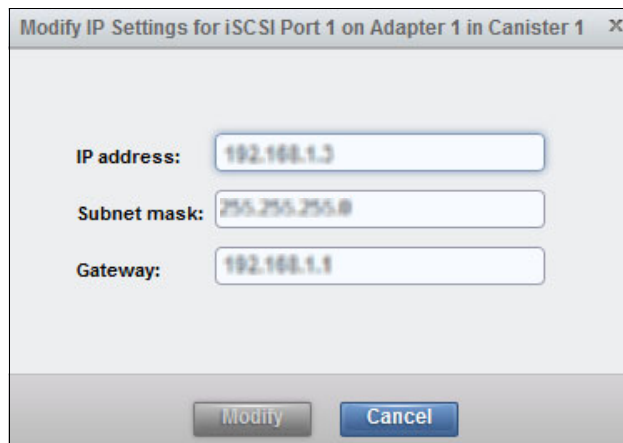


Figure 6-21 Modify iSCSI IP address

The changes are committed. The selected iSCSI port now has the selected IP address.

## Flash module properties

Click **Actions** → **Flash Modules** to open the Flash Module Properties information window (Figure 6-22).

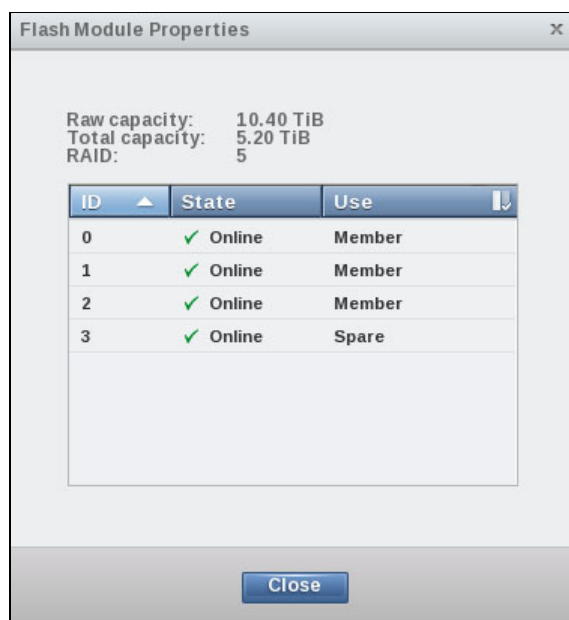


Figure 6-22 Flash Module Properties display

Flash modules within an initialized FlashSystem 900 must always be *online*, except when a flash module is in the *failed* state.

The Use column (parameter) can have these values:

- ▶ Candidate (ready to be a RAID 5 member or spare)
- ▶ Member
- ▶ Spare

RAID 5 provides redundancy for failed flash modules and keeps one flash module as a *spare*. The only situation in which a flash module can be a *candidate* is when there is no RAID configuration on the flash module. Including a candidate flash module into the RAID configuration requires the reinitialization of the array, which is a data destructive action. For instructions of how to reinitialize the RAID configuration, see 4.4, "RAID storage modes" on page 98.

**Note:** The only situation in which a flash module can be a candidate is when there is no RAID configuration on the flash modules.

To view properties of the individual flash modules, from the home window, hover the mouse over the FlashSystem 900 depiction, right-click a flash module, and select properties. The flash module properties are displayed (Figure 6-23).

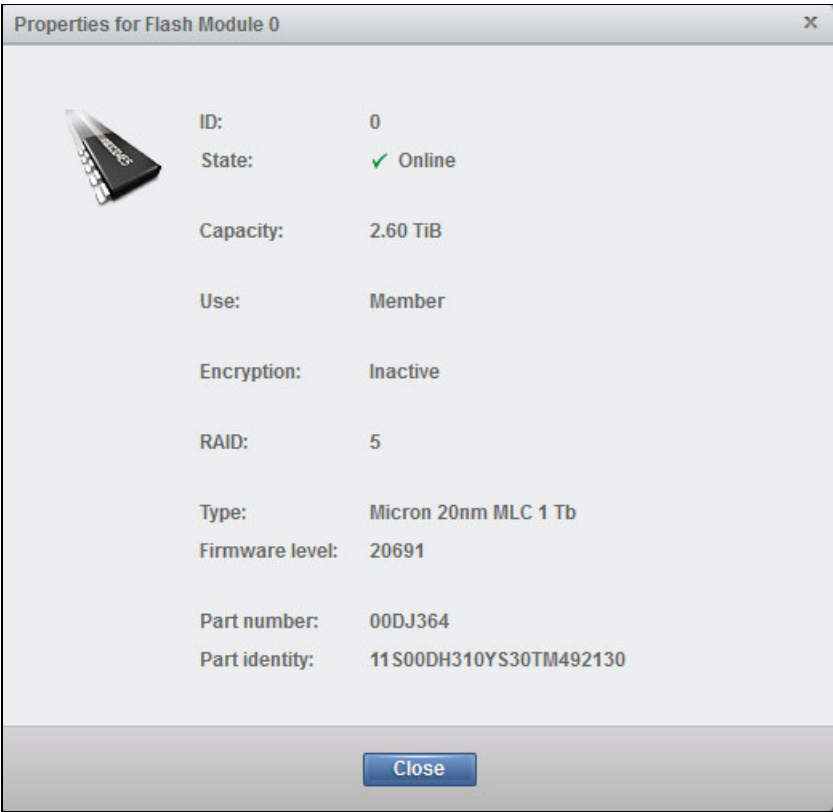
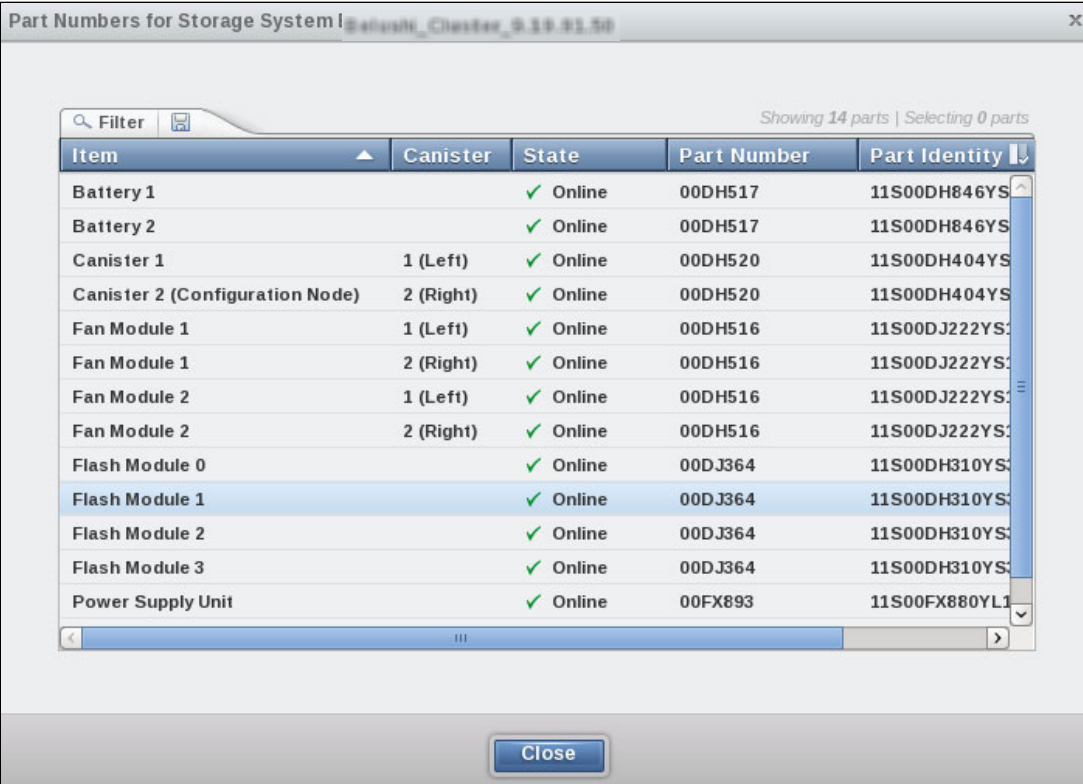


Figure 6-23 Flash Module 0 properties

## Part numbers

To review a list of part numbers for customer-replaceable units (CRUs) and field-replaceable units (FRUs), select **Actions** → **Part Numbers**. CRUs can be replaced by IBM clients; FRUs are replaced by either IBM Support or an IBM Service Partner.

Figure 6-24 shows the part numbers that are available for the IBM FlashSystem 900.



Part Numbers for Storage System | @flash900\_Cluster\_9.19.91.50

Showing 14 parts | Selecting 0 parts

Item	Canister	State	Part Number	Part Identity
Battery 1		✓ Online	00DH517	11S00DH846YS
Battery 2		✓ Online	00DH517	11S00DH846YS
Canister 1	1 (Left)	✓ Online	00DH520	11S00DH404YS
Canister 2 (Configuration Node)	2 (Right)	✓ Online	00DH520	11S00DH404YS
Fan Module 1	1 (Left)	✓ Online	00DH516	11S00DJ222YS
Fan Module 1	2 (Right)	✓ Online	00DH516	11S00DJ222YS
Fan Module 2	1 (Left)	✓ Online	00DH516	11S00DJ222YS
Fan Module 2	2 (Right)	✓ Online	00DH516	11S00DJ222YS
Flash Module 0		✓ Online	00DJ364	11S00DH310YS
Flash Module 1		✓ Online	00DJ364	11S00DH310YS
Flash Module 2		✓ Online	00DJ364	11S00DH310YS
Flash Module 3		✓ Online	00DJ364	11S00DH310YS
Power Supply Unit		✓ Online	00FX893	11S00FX880YL1

Close

Figure 6-24 List of part numbers

## 6.2.2 Monitoring events

The IBM FlashSystem 900 might show the Health Status indicator as these colors:

- ▶ Green (healthy)
- ▶ Yellow (degraded or warning)
- ▶ Red (critical)

Events on the IBM FlashSystem 900 storage system are logged in the event log of the **Monitoring** → **Events** menu.

## Navigating to events

To navigate to the event log, hover the cursor over the **Monitoring** icon and then click **Events** (Figure 6-25).

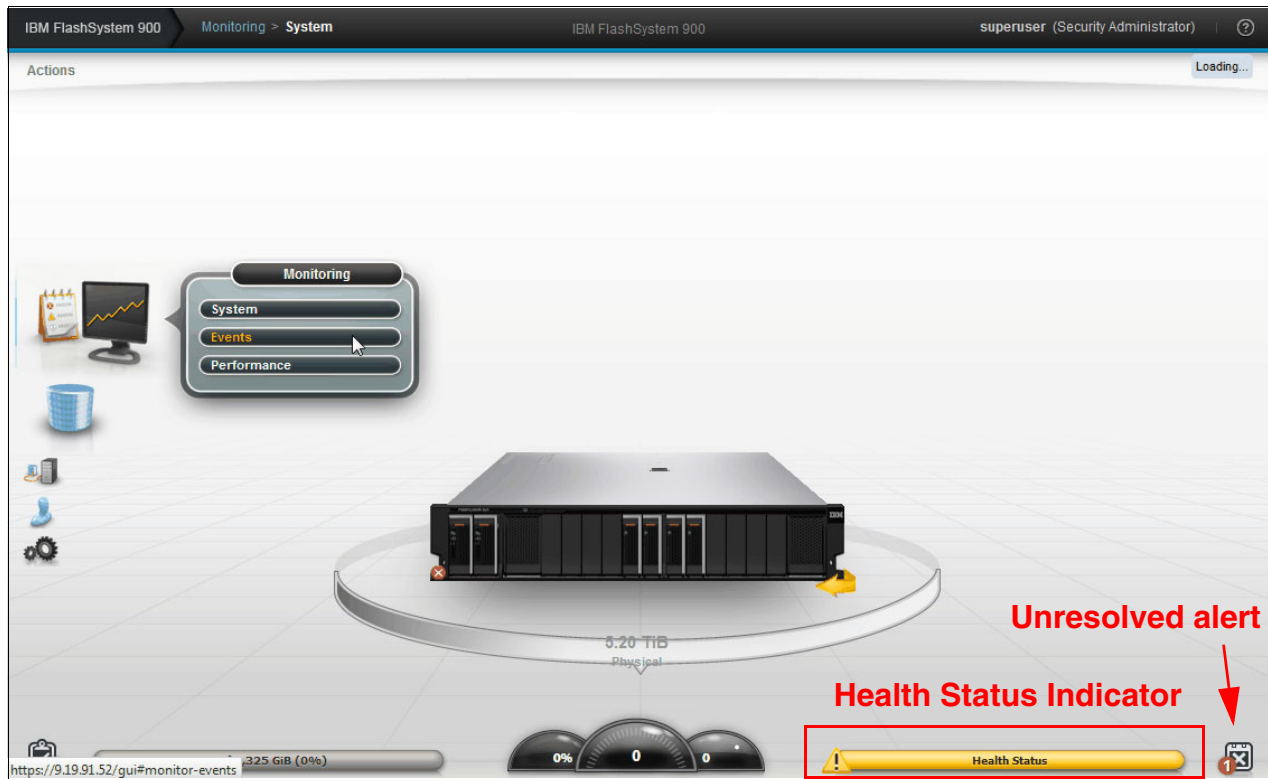


Figure 6-25 Monitoring events

You can also click the attention (!) triangle icon on the left side of the Health Status indicator to get to the Monitoring > Events window as shown in Figure 6-25. A yellow Health Status indicator indicates a warning state.



Figure 6-26 shows the Monitoring > Events window with the Show All mode selected. In the Show All mode, all events, including messages, warnings, and errors are displayed. To get to the Show All mode, first click a *recommended action* and then select **Show All**.

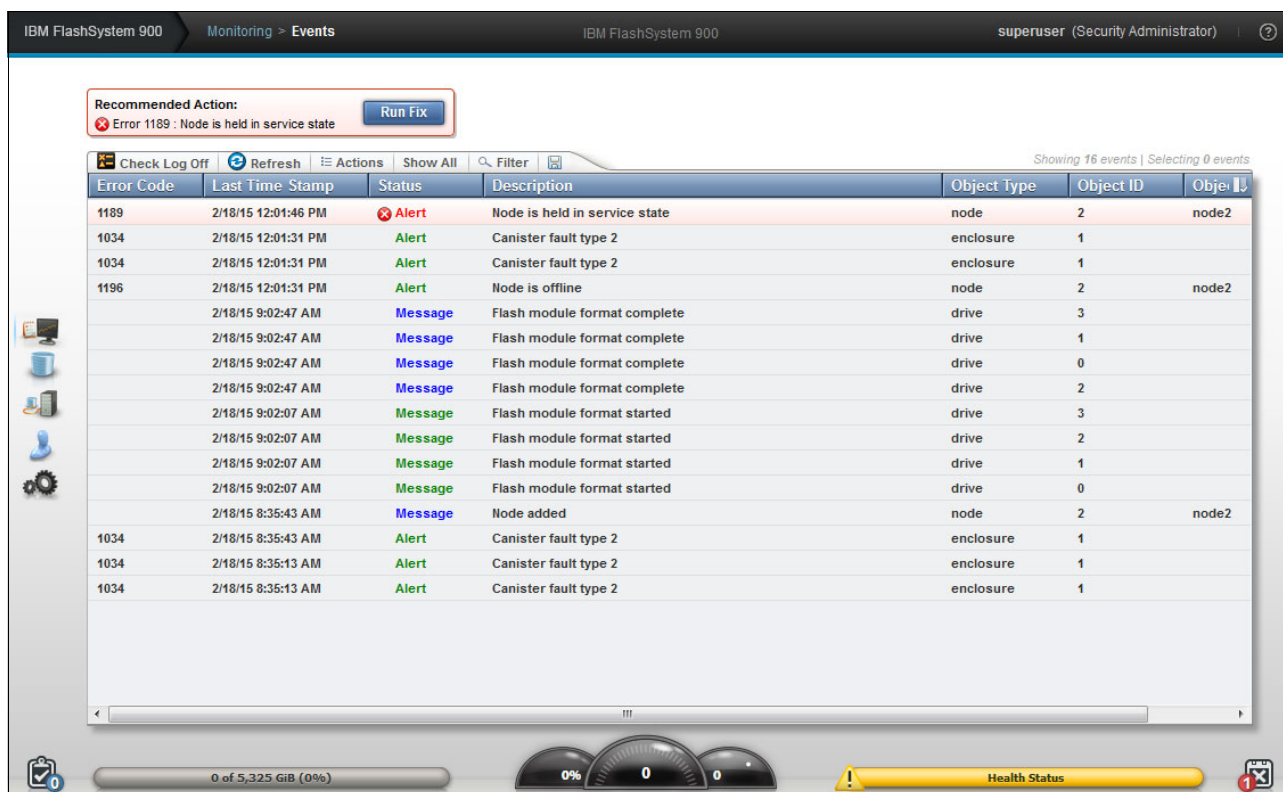


Figure 6-26 The Events window

In the example, a controller node is held in service state, which causes the health status to enter a warning state. Resolving issues is demonstrated in “Directed maintenance procedures (DMP)” on page 185.

The **Monitoring** → **Events** menu can be manipulated in several ways by using the function tabs that are displayed over the list of events.

### Check Log LED off

One function of the FlashSystem 900 is its Check Log LED. This LED illuminates amber for a problem that is not isolated. An error condition results in a *call home* of the problem. There is also a service action; a warning condition results in a service action that the user is expected to fix. Note that there is no correlation between the notification type of “error/warning” and the Check Log LED.

The leftmost function key of the Events menu is the Check Log Off. With Check Log Off, you turn off the Check Log LED on the front of the IBM FlashSystem 900, and only new events turn it on again.

From the Monitoring > Events window, click **Check Log Off**. The window shown in Figure 6-27 opens. To turn off the Check Log LED, click **Yes**.

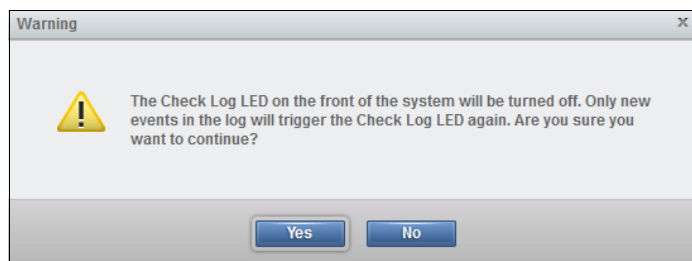


Figure 6-27 Events error LED off

## Change the Events view

You might want more or less information from the Monitoring > Events window. You can change the default view by right-clicking the menu bar or by clicking the check mark icon in the upper-right corner of the Monitoring > Events window, as shown in Figure 6-28.

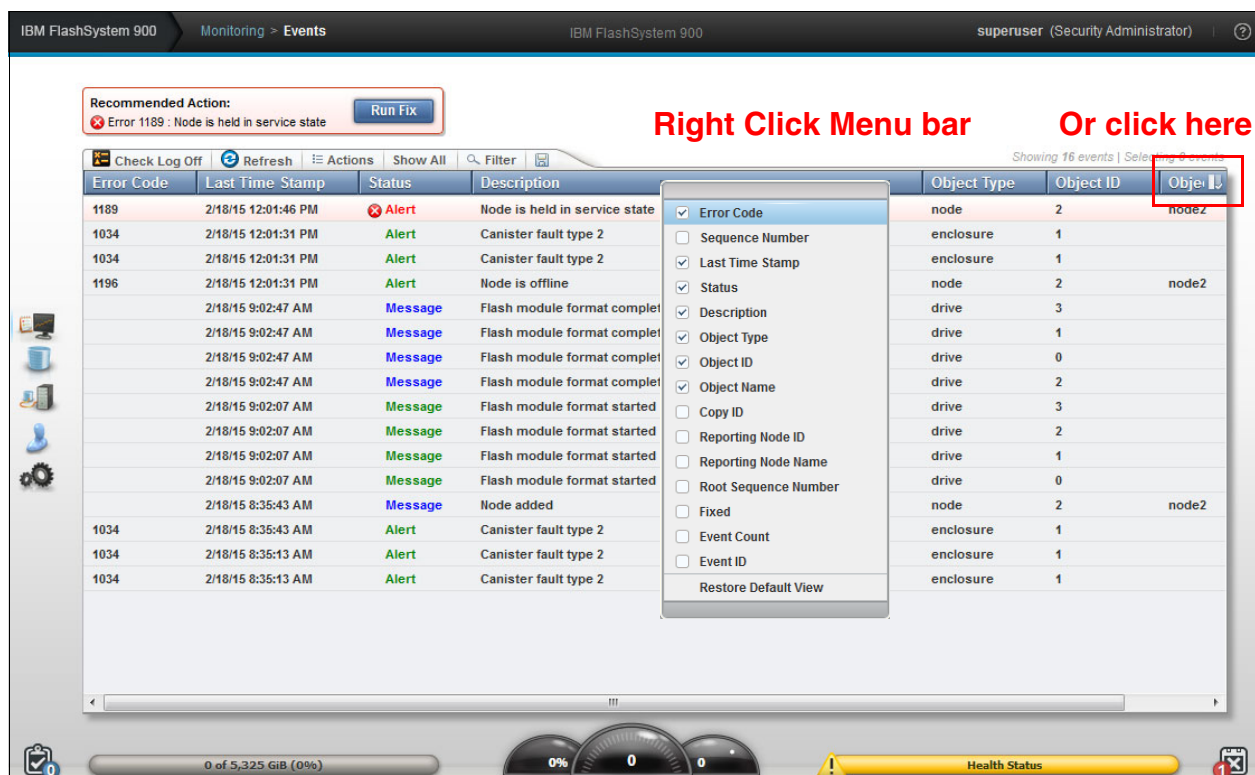


Figure 6-28 Events: Customize columns

From the Actions menu, you can do the following actions:

- ▶ Run fix procedure on error events
- ▶ Mark informational events as fixed
- ▶ Clear the event log
- ▶ Filter displayed events on date
- ▶ Show only events from the last minutes, hours, or days
- ▶ Show the properties of an event

Figure 6-29 shows that only events that are newer than five hours are selected to view.

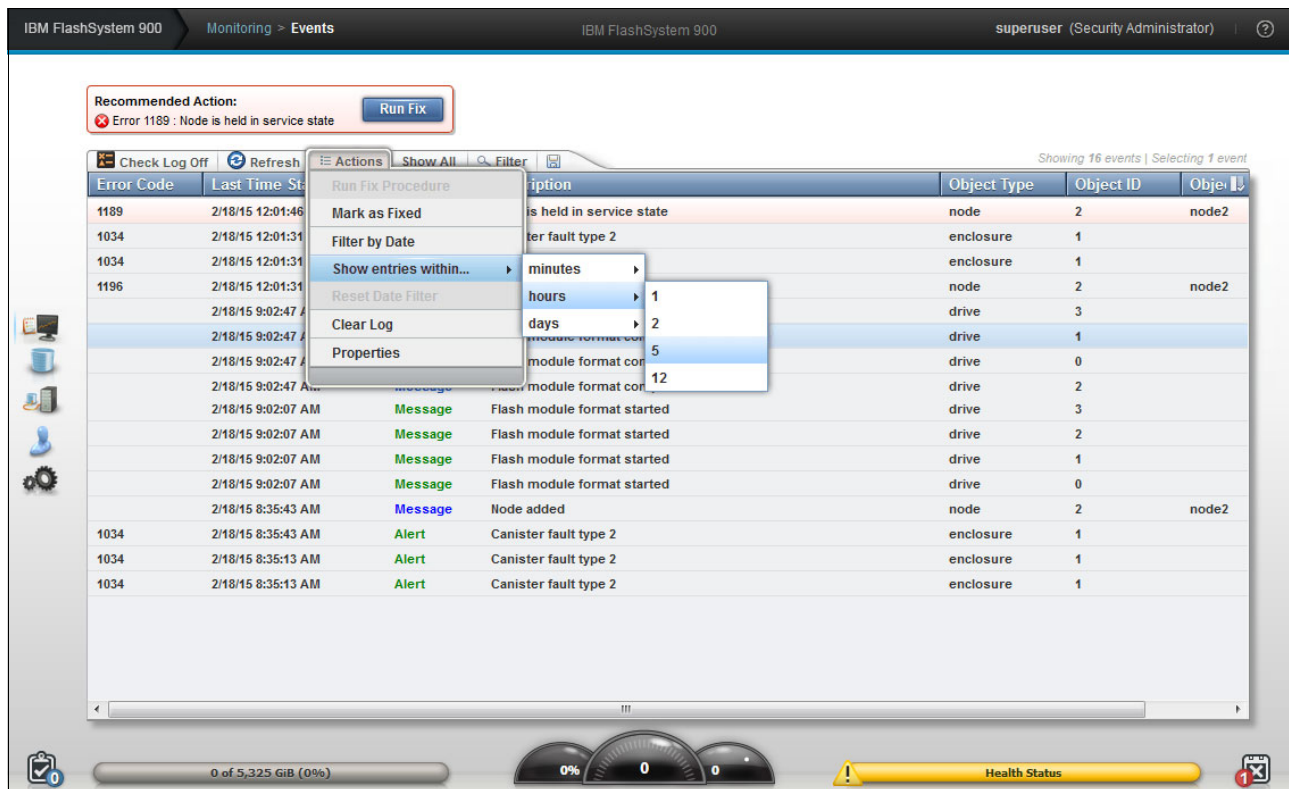


Figure 6-29 Events: Show entries within five hours

## Recommended actions

In the Monitoring > Events window, messages and alerts are displayed. If any unresolved issues exist, the Recommended Actions section opens. You can click **Run Fix** to initiate the fix procedure; the IBM FlashSystem 900 checks whether the problem still exists and fixes the issue, if possible. The fix procedure might bring the system out of a Degraded state and into a Healthy state.

In a normal situation during the daily administration of the FlashSystem 900, you are unlikely to see error events. There might however be a continuing flow of informational messages. The typical Events display is therefore to show only recommended actions.

To show only recommended actions, click **Show All** and select **Recommended Actions** as shown in Figure 6-30.

IBM FlashSystem 900

Monitoring > Events

IBM FlashSystem 900

superuser (Security Administrator)

Recommended Action:  
Error 1189 : Node is held in service state

Run Fix

Check Log Off Refresh Actions Show All Filter

Showing 16 events | Selecting 0 events

Error Code	Last Time Stamp	Status	Recommended Actions	Object Type	Object ID	Object Name
1189	2/18/15 12:01:46 PM	Error	Unfixed Messages and Alerts	node	2	node2
1034	2/18/15 12:01:31 PM	Alert	Canister fault type 2	enclosure	1	
1034	2/18/15 12:01:31 PM	Alert	Canister fault type 2	enclosure	1	
1196	2/18/15 12:01:31 PM	Alert	Node is offline	node	2	node2
	2/18/15 9:02:47 AM	Message	Flash module format complete	drive	3	
	2/18/15 9:02:47 AM	Message	Flash module format complete	drive	1	
	2/18/15 9:02:47 AM	Message	Flash module format complete	drive	0	
	2/18/15 9:02:47 AM	Message	Flash module format complete	drive	2	
	2/18/15 9:02:07 AM	Message	Flash module format started	drive	3	
	2/18/15 9:02:07 AM	Message	Flash module format started	drive	2	
	2/18/15 9:02:07 AM	Message	Flash module format started	drive	1	
	2/18/15 9:02:07 AM	Message	Flash module format started	drive	0	
	2/18/15 8:35:43 AM	Message	Node added	node	2	node2
1034	2/18/15 8:35:43 AM	Alert	Canister fault type 2	enclosure	1	
1034	2/18/15 8:35:13 AM	Alert	Canister fault type 2	enclosure	1	
1034	2/18/15 8:35:13 AM	Alert	Canister fault type 2	enclosure	1	

0 of 5,325 GiB (0%)

0% 0 0

Health Status

Figure 6-30 Events: Show all or only recommended actions

Figure 6-31 shows the resulting error codes. It indicates that a problem in the system needs attention and fixing.

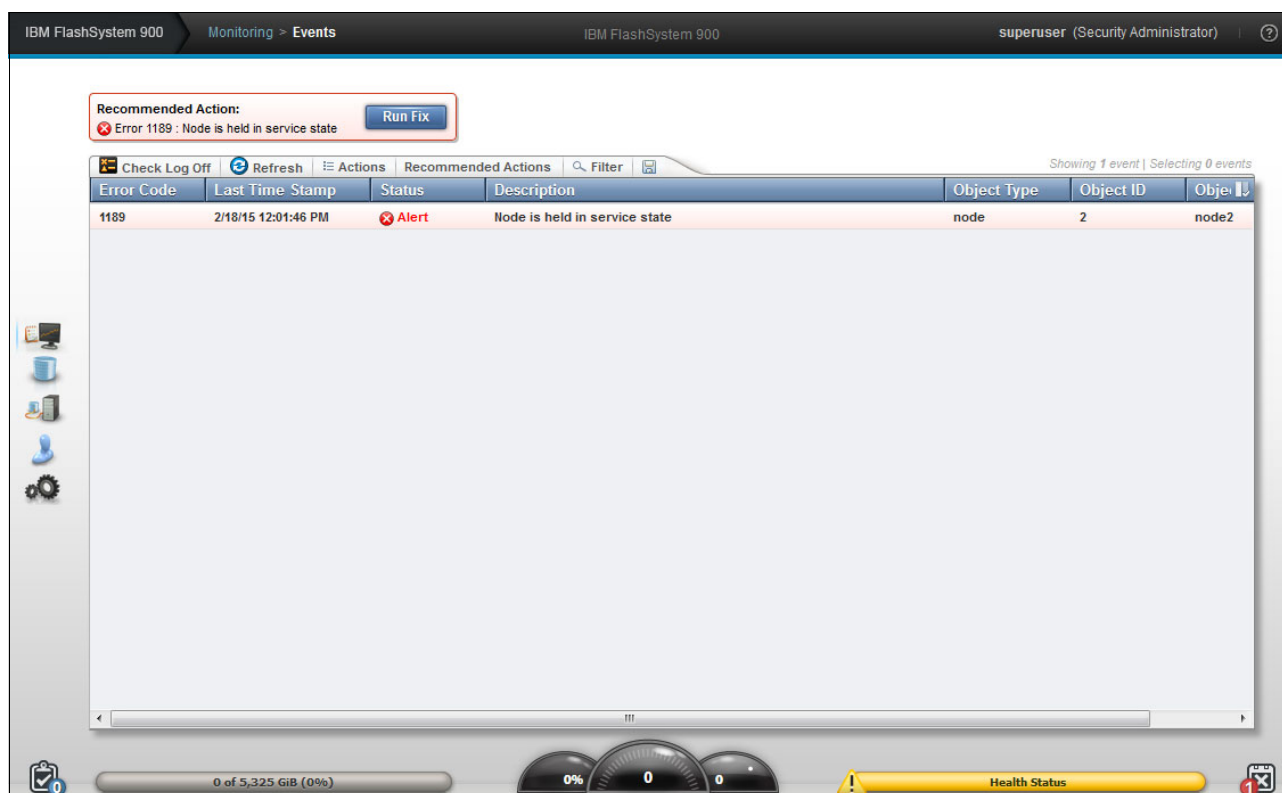


Figure 6-31 Events: Only showing recommended actions

An example of the window that displays the details that are associated with the specific event IDs is in Figure 6-35 on page 187.

## Directed maintenance procedures (DMP)

There are different ways to discover that your system needs attention in a warning or error situation. If the call home feature is configured on your system, which is advised, IBM Support is notified directly from the system and the system administrators are contacted by IBM for corrective actions.

The system administrator might also be in the list of email recipients and therefore is notified directly and immediately from the system as soon as an alert is sent.

For more information about how to configure call home, see Chapter 4, “Installation and configuration” on page 67 and “Email” on page 233.

Another way of getting alert notifications is through Simple Network Management Protocol (SNMP) alerts.

For more information about how to configure SNMP alerts, see “SNMP” on page 241.

When the system administrator logs on to the GUI of the FlashSystem 900, the Status Alerts message is displayed in the lower-right corner. Hovering over the Status Alerts “X” icon shows the unresolved alerts (Figure 6-32).

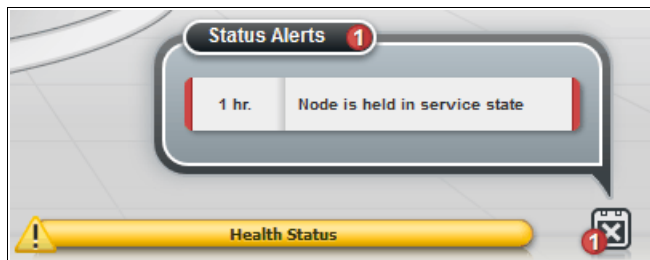


Figure 6-32 Status alerts are displayed

Status Alerts messages are also visible on the picture of the FlashSystem 900 storage in the home window (the Status Alerts is indicated by an X at the right side of the system).

Figure 6-33 shows an X in a red circle where enclosure 1 has an error condition.



Figure 6-33 Status Alerts view displays on the FlashSystem 900 graphic

The Status Alerts in the lower-right corner of the home window is similar to the Status Alerts X icon in the lower-right corner of the IBM FlashSystem 900 depiction. Hovering over the alerts X icon on the graphic of the system shows any unresolved alerts (Figure 6-33).

The situation that was demonstrated can be corrected from only the Service Assistant Tool, which is indicated if you run the Run Fix procedure from the Recommended Actions indicator.

An event that can be fixed from the Recommended Actions indicator is an error situation where the ambient temperature was detected as too high. This is demonstrated next.

### ***Fixing the event***

Clicking any of the Status Alerts displayed takes you to the Monitoring > Events menu where details about the events can be reviewed. The Recommended Actions box displays and indicates that unresolved errors are in the log and need your attention. This method to fix errors is also referred to as the *directed maintenance procedures (DMP)*.

To view the details of a specific event, highlight the event and click **Properties** (Figure 6-34).

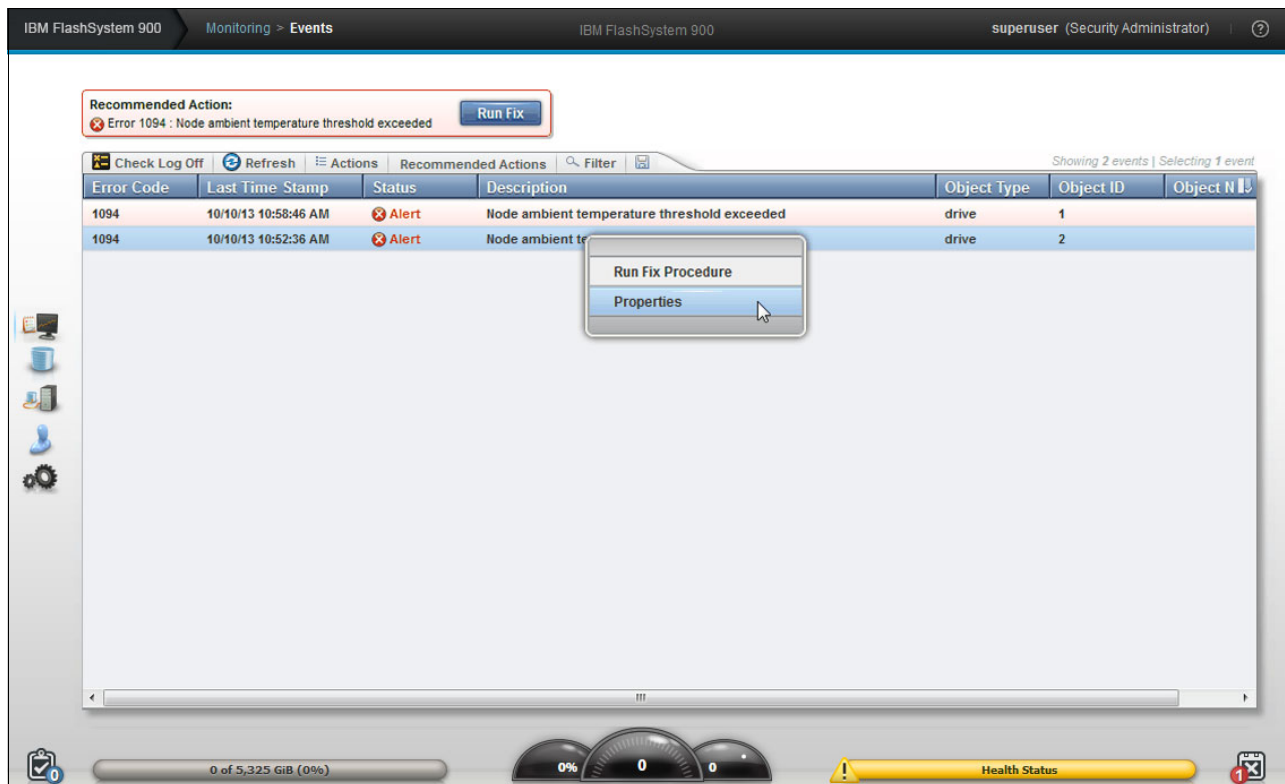


Figure 6-34 Event Properties

The Properties window opens (Figure 6-35) and you can review the details of the error.



Figure 6-35 Properties for an event



Figure 6-36 shows an error. Before starting the fix procedure, filter events to the Recommended Actions so that only errors that require attention are displayed here. Click **Run Fix** to initiate the DMP procedure.

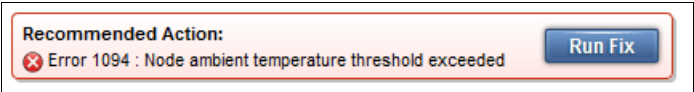


Figure 6-36 Events: Run Fix

The Run Fix procedure guides you through resolving the error event. The error message in this example is caused by a room temperature that is too high, which might cause the system to overheat and eventually shut down if the error situation is not corrected.

Figure 6-37 shows the first step of the DMP procedure. The system reports that drive 2 (flash module 2) in slot 5 is measuring a temperature that is too high. The display also indicates that all four fans in both canisters are operational and online.

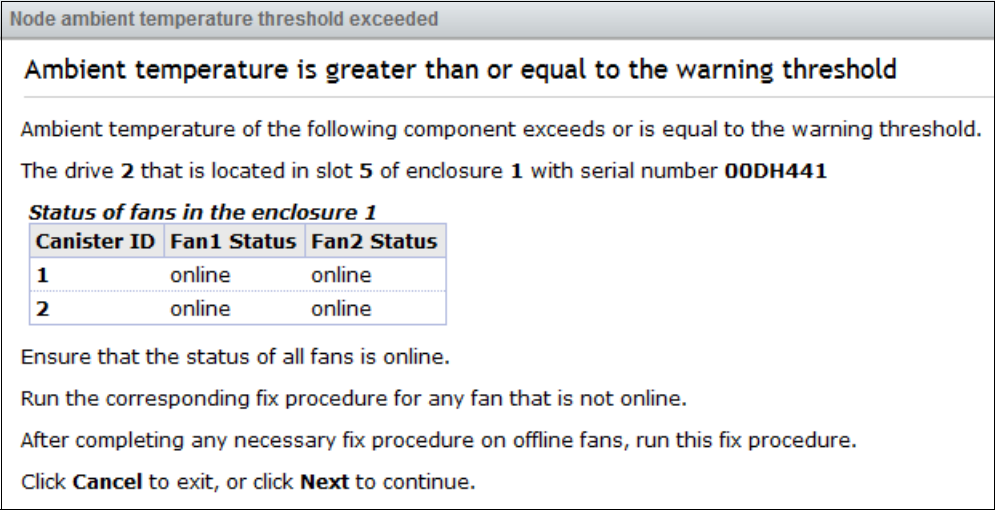


Figure 6-37 DMP procedure step 1

The next step in the DMP procedure is for the administrator to measure the room temperature and to make sure that the ambient temperature is within the specifications for the system. The instructions for this step are shown in Figure 6-38.

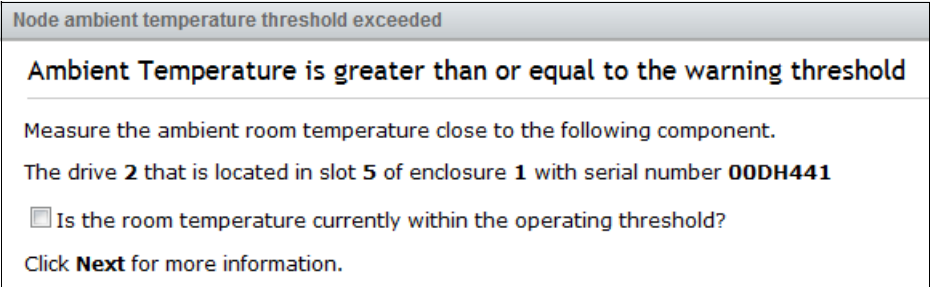


Figure 6-38 DMP procedure step 2



In the third step of the DMP procedure, suggestions about potential causes of overheating are provided. Overheating might be caused by blocked air vents, incorrectly mounted blank carriers in a flash module slot, or a room temperature that is too high. Instructions are displayed as shown in Figure 6-39.

Node ambient temperature threshold exceeded

Check for air flow blockages

Ambient temperature of the following component exceeds or equal to the warning threshold.  
The drive **2** that is located in slot **5** of enclosure **1** with serial number **00DH441**  
Take steps to ensure the environmental requirements of this enclosure are being met. Ensure the following:

- Environmental temperature controls are set to provide the recommended ambient operating temperature.
- Enclosure vents are kept free of dust and debris.
- Air flow in the vicinity of the enclosure is not impeded, for example by cables, equipment, doors, walls or other obstructions in the room.
- A blank carrier is installed in each drive slot that does not contain a drive.

If you make any corrections, allow some time for the node to cool down before continuing this fix procedure.  
Click **Next** for more information.

Figure 6-39 DMP procedure step 3

In this step, the DMP procedure checks whether the error condition is resolved, and all events of the same type are marked as fixed, if possible. The final step is shown in Figure 6-40.

Node ambient temperature threshold exceeded

Event has been marked as fixed

The ambient temperature events have been marked as fixed.  
Click **Close** to exit.

Figure 6-40 DMP procedure step 4

The events indicating an error condition relating to temperature are now gone and the system is back in a healthy state as shown in Figure 6-41.

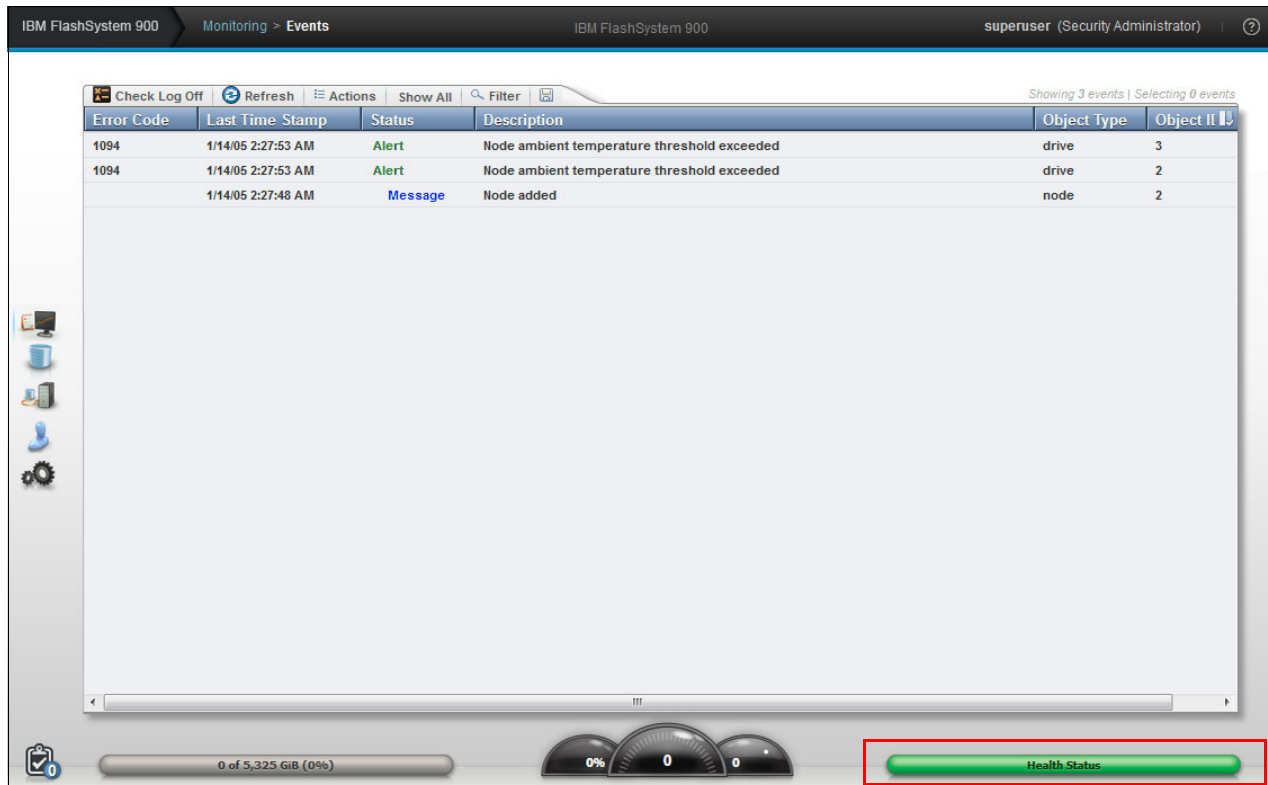


Figure 6-41 Error condition is now resolved

For more information about operational specifications for the IBM FlashSystem 900, see the following web page:

<http://www.ibm.com/systems/storage/flash/>

### 6.2.3 Monitoring performance menu

The IBM FlashSystem 900 Performance menu gives you a good overview of how the system is performing. In the latest firmware release, the performance monitor changed for enhanced functionality. With previous firmware releases, the performance graphs represented five minutes of data. With THE current firmware release, the default performance monitor represents 10 minutes of captured data and the view can expand for showing up to 300 days.

## Performance menu overview

To open the FlashSystem 900 performance monitor, select **Monitoring** → **Performance**. The first time the browser window opens, system latency is displayed (Figure 6-42).

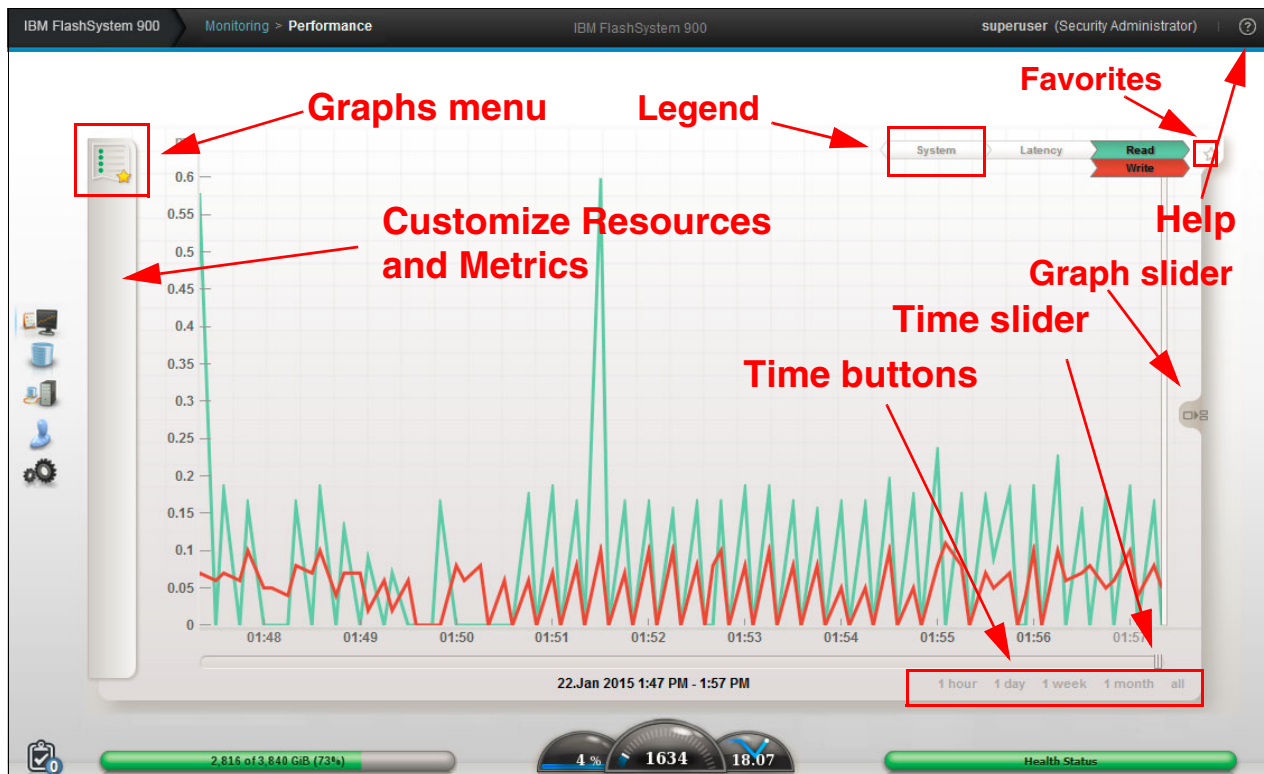


Figure 6-42 Performance monitoring default window

The horizontal part of the graph displays the time line. You can slide the time line to view the past. You can also adjust the granularity of the time line by selecting one hour, one day, one week, one month or all. All displays the year to date.

Five performance charts can be reviewed from the graphs menu. You can save graphs and they are displayed as follows:

- **System I/O**

The System I/O graph displays the average number of read, write, and total I/O requests per second (IOPS) over the sample period. Each request type (read, write, and total) is represented by a different color line.

- **System Latency**

The System Latency graph displays the average amount of time in milliseconds (ms) each read and write I/O request takes over the sampling period. Each request type (read and write) is represented by a different color line.

- **System Bandwidth**

The System Bandwidth graph displays the average number of megabytes per second (MBps) of read, write, total, and rebuild bandwidth over the sample period. Each bandwidth type (read, write, total, and rebuild) is represented by a different color line. There is one line graph for each system that is selected.

- Interface Port Total IOPS

The Total Port IOPS graph displays average number of read, write, and total I/O requests per second (IOPS) over the sample period. There is one line on the graph for each port in each host adapter in each canister. Each adapter has a different color, and all four ports on an adapter have the same color.

- Fibre Channel Port Total Queue Depth

The Total Port Queue Depth graph displays the average number of operations of that type over the sample period. There is one line on the graph for each port in each host adapter in each canister. Each adapter has a different color, and all four ports on an adapter have the same color.

## Graphs menu

The Graphs menu has five default graphs defined. By clicking the icon of the graphs menu, you can select the graphs, as shown in Figure 6-43.



Figure 6-43 Graphs menu default window

By default the performance monitor shows System Latency. Graphs can be customized and added to the menu as described in “Customize graphs menu” on page 195.

The graphs menu lists different resources for the ports, depending on the FlashSystem 900 model (InfiniBand, iSCSI, or Fibre Channel) as shown in Figure 6-44 on page 193.

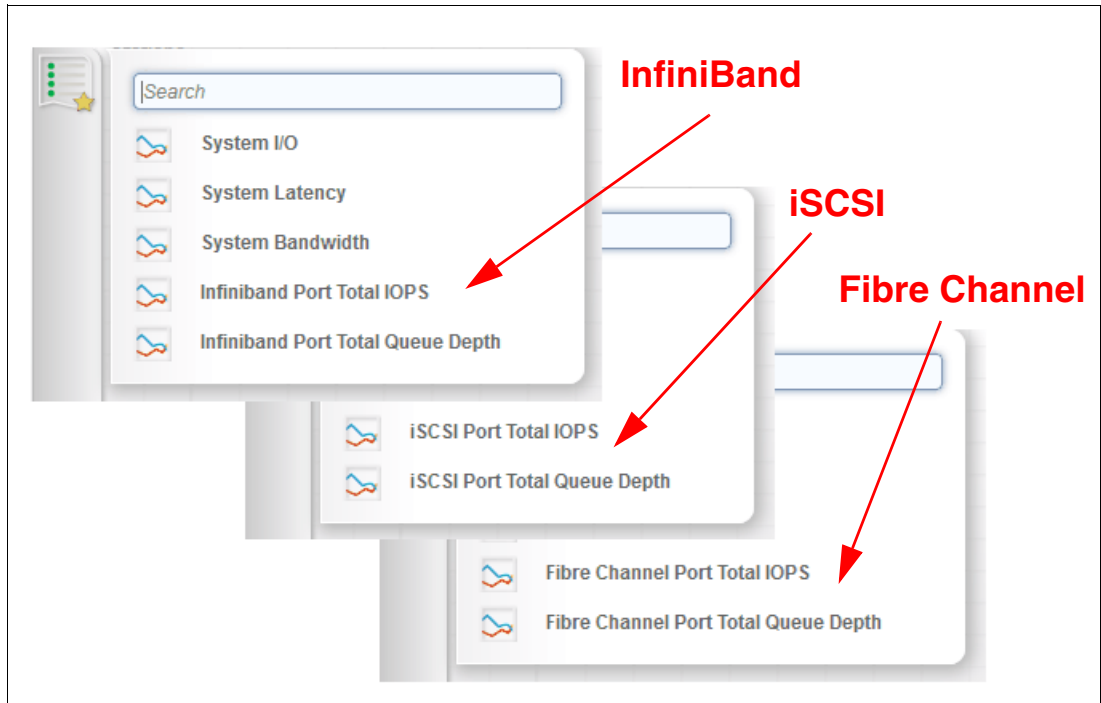


Figure 6-44 Graphs menu of three different models of FlashSystem 900

In the following example, select **System I/O** from the graphs menu. The performance monitor now displays 10 minutes of IOPS (Figure 6-45). Any time of interest can be selected by moving the time slider.

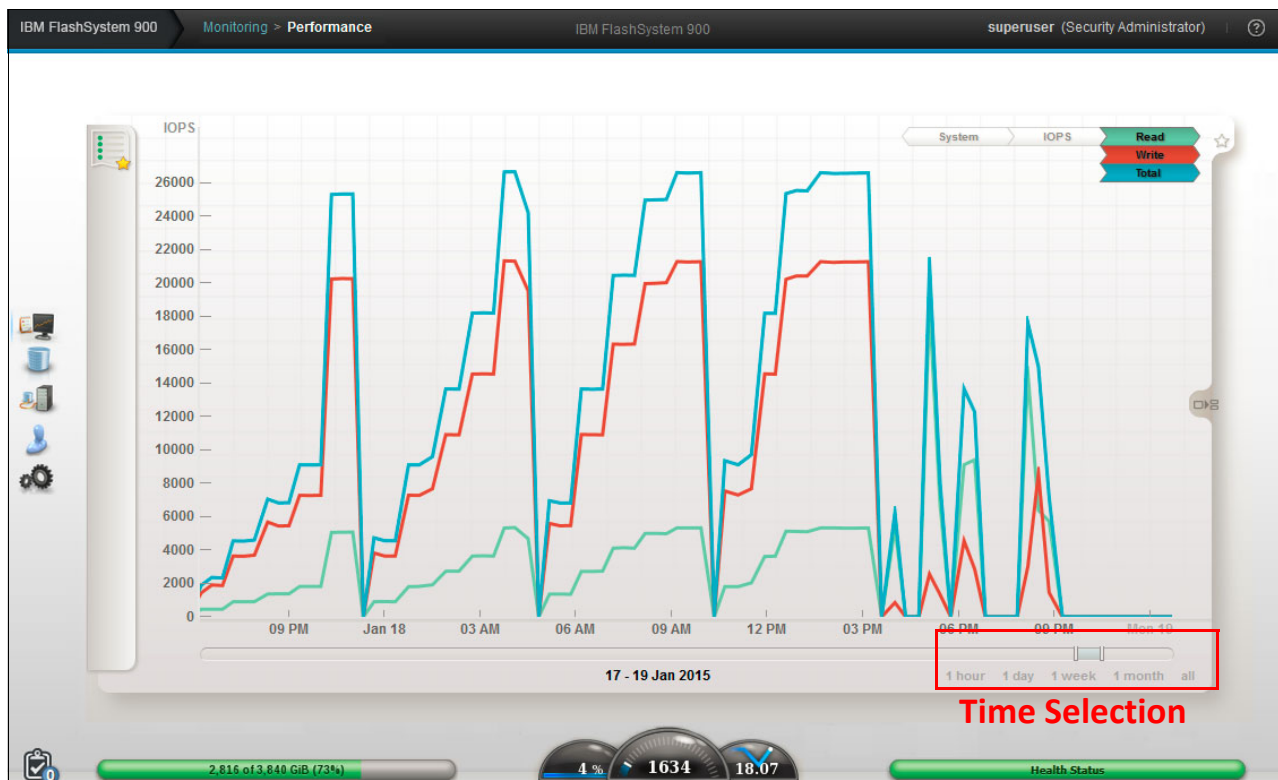


Figure 6-45 Performance graphs showing IOPS at a specific time View two charts

As a storage administrator, you might want to know if there were long response times from storage (latency) at a given time. To compare two metrics, for example IOPS and latency, use the **Click to view two charts** icon at the right side of the graphs window (Figure 6-46).

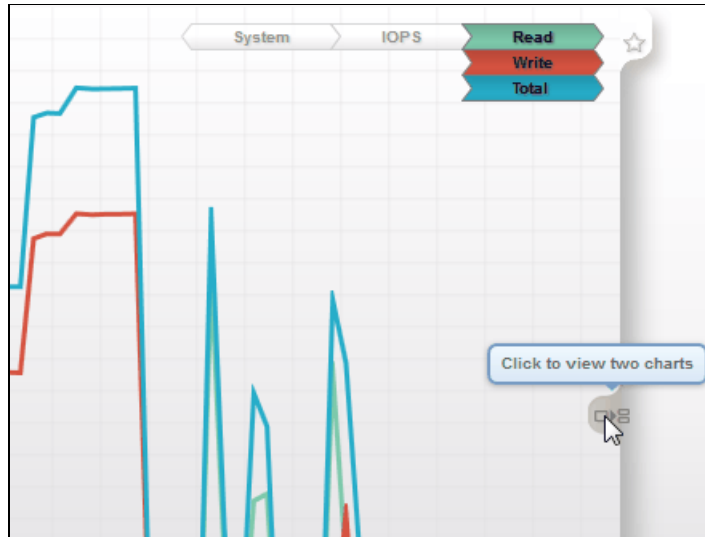


Figure 6-46 Click to view two charts

Then, select the second graph from the graphs menu. You can point to either graph (the upper one or the lower one, whichever you prefer).

Figure 6-47 shows two graphs, IOPS and latency, in the Performance window.

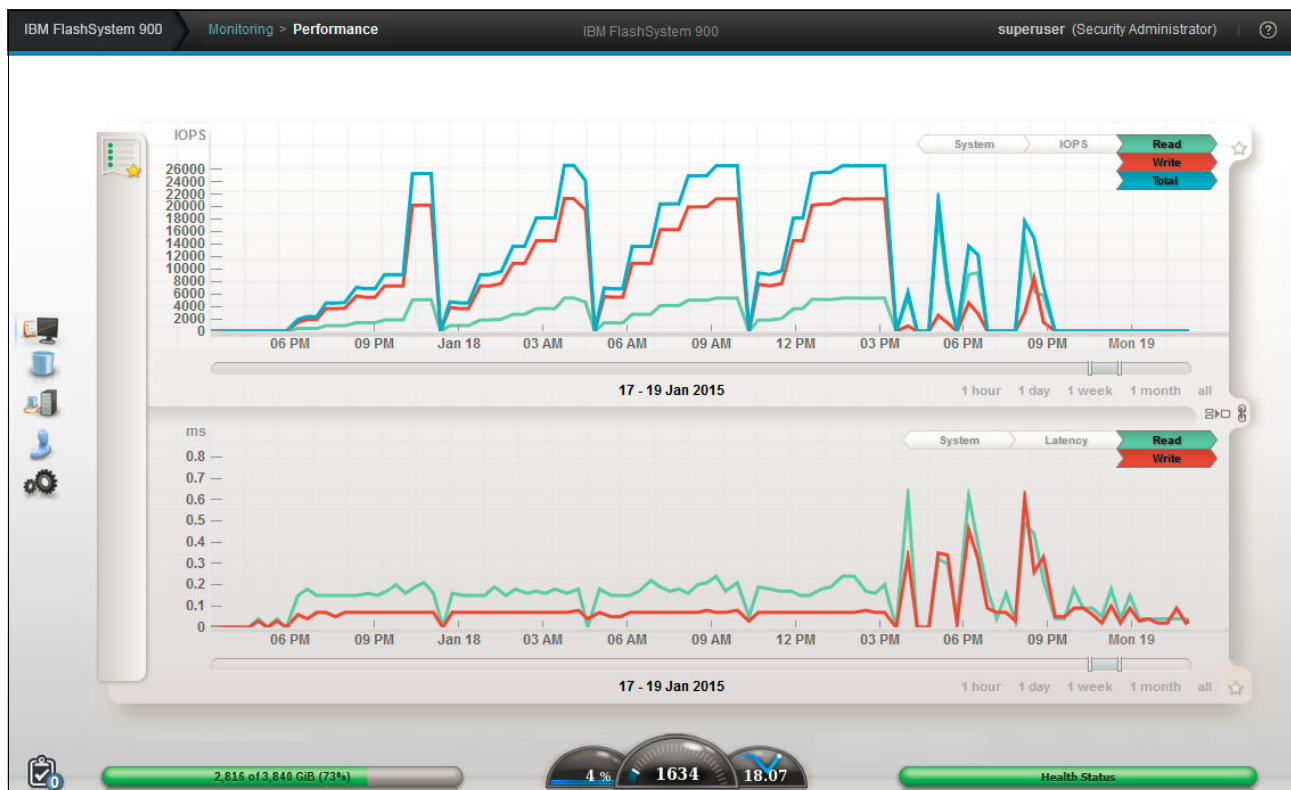


Figure 6-47 Two graphs are displayed

## Customize graphs menu

To customize a graph, click the **Customize Resources and Metrics** bar at the left side of the window. Then, select resources from the upper part of the window and metrics from the lower part of the window (Figure 6-48).

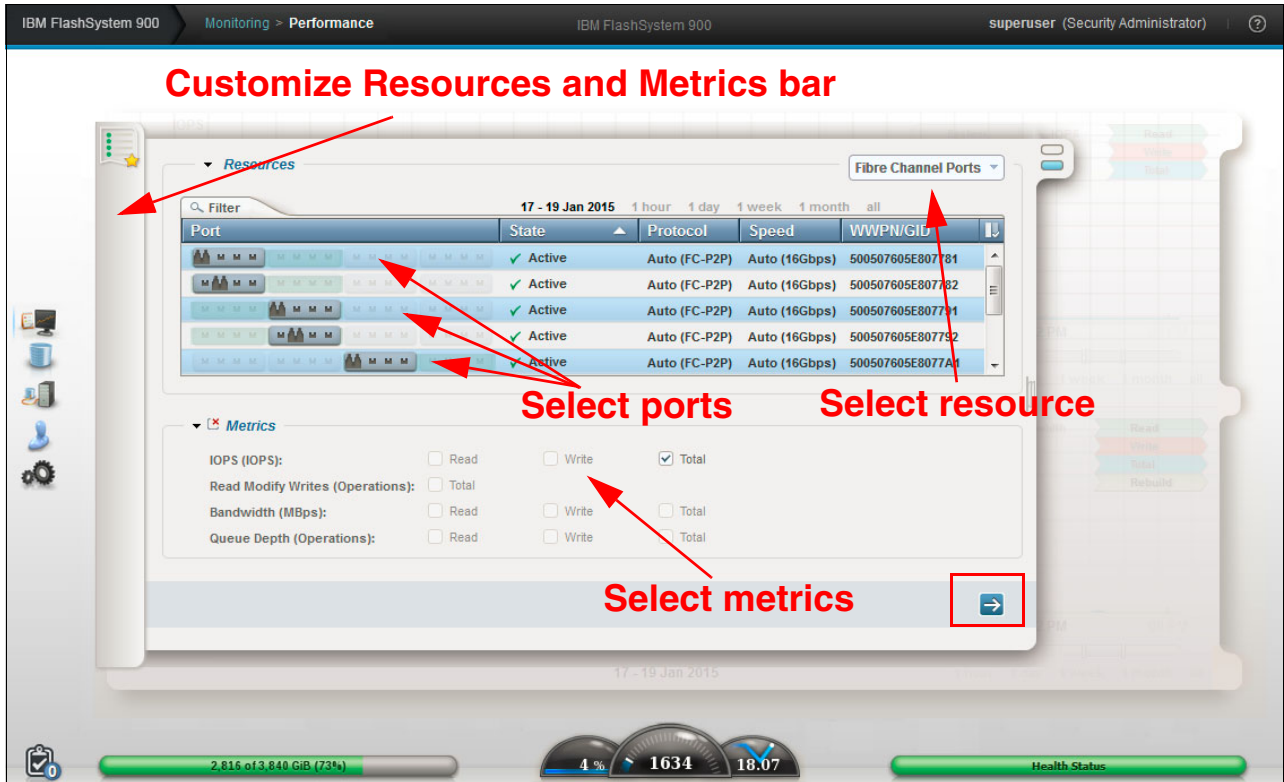


Figure 6-48 Customize graphs



When all resources and metrics are selected, click the blue arrow at the lower right corner; the resulting graphs are displayed (Figure 6-49).



Figure 6-49 Customized graphs are displayed



By clicking the **Add to favorites** icon (Figure 6-50), the two graphs are now listed as favorites on the graphs menu.

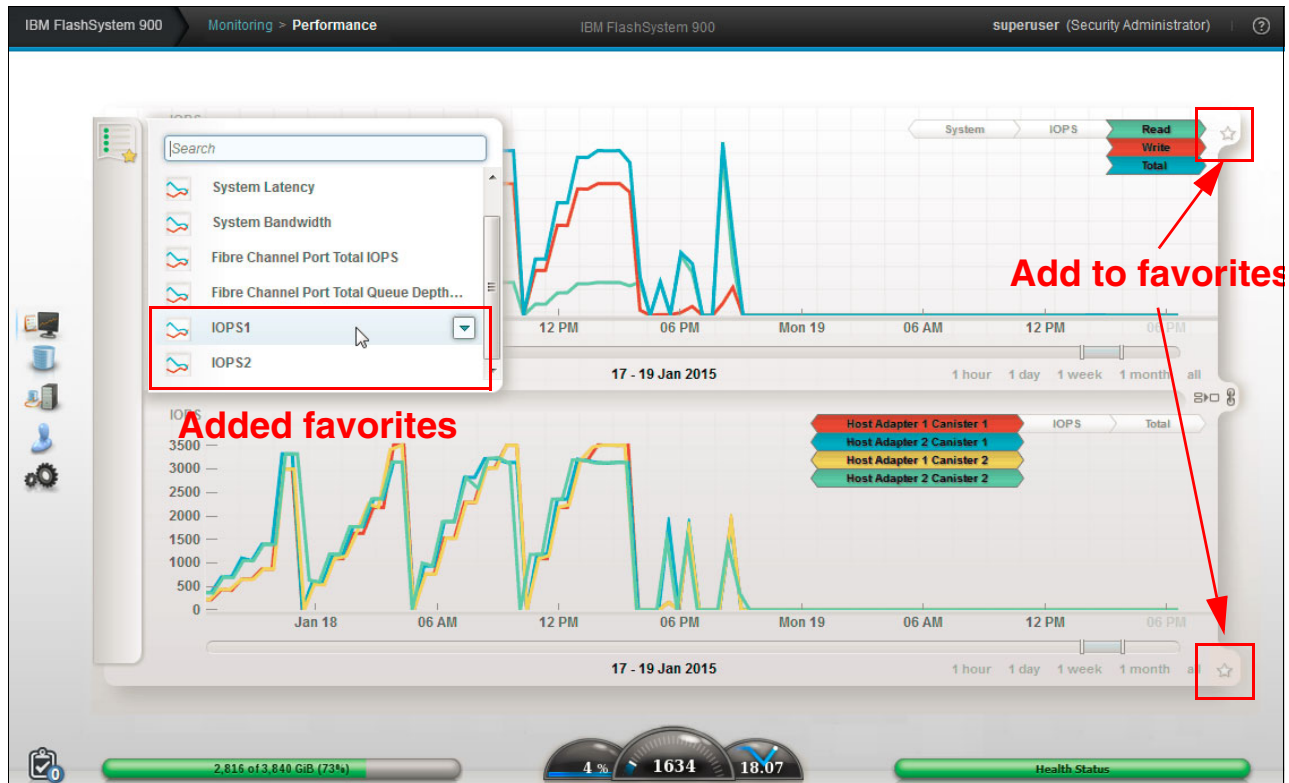


Figure 6-50 Graphs added to favorites

The graphs from the graphs menu can also be pinned to the toolbar where they appear as icons. The customized graphs can also be selected as the default graphs (**Set as Default**) for the FlashSystem 900 performance monitor (Figure 6-51).

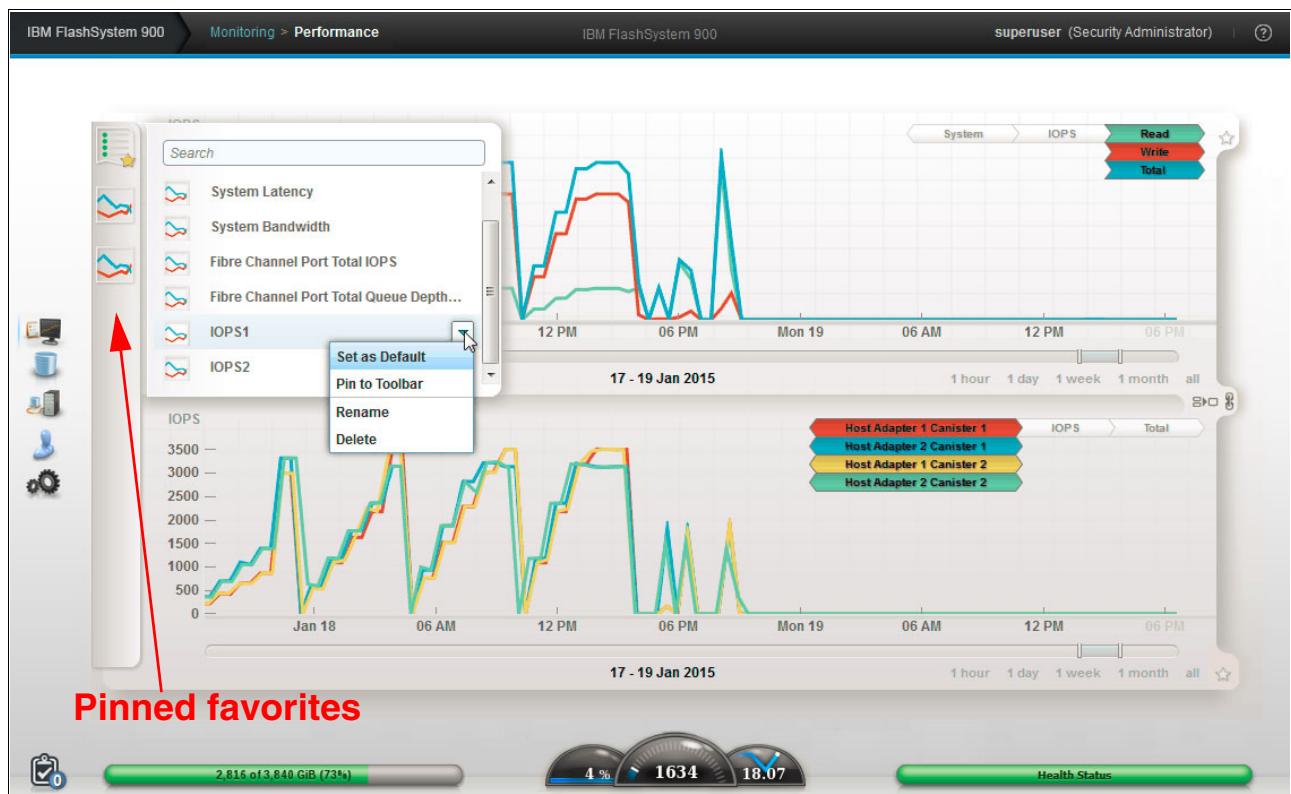


Figure 6-51 Graphs pinned to toolbar

## Host port adapter numbering

Figure 6-52 shows the numbering and naming of the IBM FlashSystem 900 I/O interface ports as they correspond to the performance monitor charts for Interface Queue Depth.

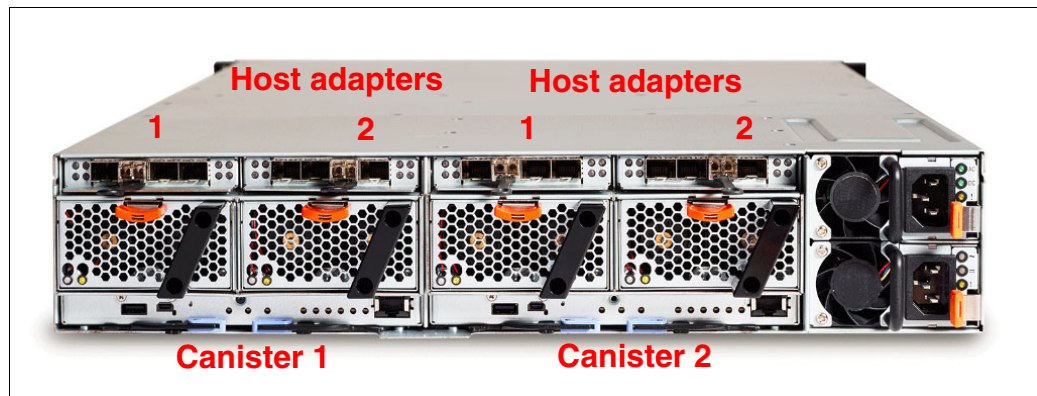


Figure 6-52 Rear-side numbering of interface cards

## IBM data management and storage management

If you need additional performance monitoring, the optimal tool is IBM data management and storage management, which is provided by IBM Spectrum Control. You can manage performance and connectivity from the host file system to the physical disk, including in-depth performance monitoring and analysis of the storage area network (SAN) fabric.

**Name change:** Prior to version 5.2.8, IBM Spectrum Control was known as IBM Tivoli Storage Productivity Center.

IBM data management and storage management solutions deliver the functions of IBM Spectrum Control, a member of the IBM Spectrum Storage family.

For more information about IBM Spectrum Control, see the product details at this web page:

<https://www.ibm.com/support/knowledgecenter/SS5R93>

### Notes:

- ▶ For initial and current releases of the IBM FlashSystem 900, IBM Spectrum Control does not support the product directly. The exception is if the FlashSystem 900 functions as an MDisk for the IBM SAN Volume Controller, in which case, IBM Spectrum Control supports the product through the SAN Volume Controller.
- ▶ SAN Volume Controller delivers the functions of IBM Spectrum Virtualize, part of the IBM Spectrum Storage family.

## 6.3 Volumes menu

This topic provides information about managing volumes.

You can use the FlashSystem 900 GUI or CLI **svctask mkvdisk** command to create a volume. After volumes are created, they can be mapped to a host by using the **mkvdiskhostmap** command.

The volumes are built from extents in the RAID 5 or RAID 0 flash module arrays, and the volumes are presented to hosts as logical units that the host sees as external disks.

The Volumes menu has two options:

- ▶ Volumes
- ▶ Volumes by Host

### 6.3.1 Navigating to the Volumes menu

When you hover the cursor over the **Volumes** function icon, the Volumes menu opens (Figure 6-53).



Figure 6-53 Navigate to the Volumes menu

### 6.3.2 Volumes menu

In the Volumes menu, click **Volumes**; the Volumes window opens (Figure 6-54). You can perform tasks on the volumes, such as create, expand, rename, and delete, or you can review the properties of the volume.

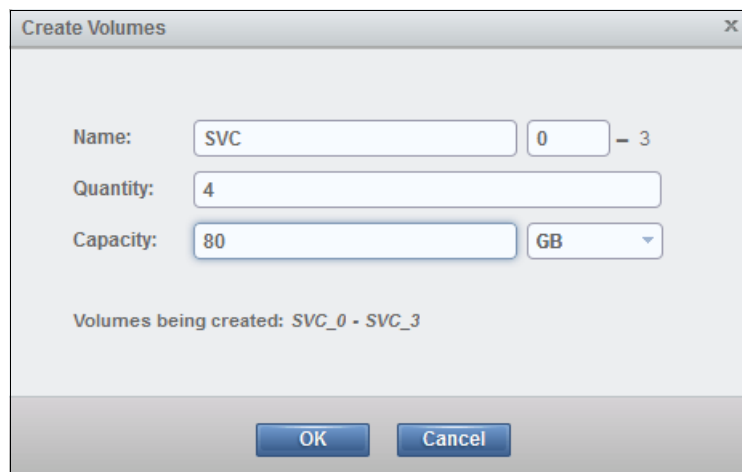
The screenshot shows the 'Volumes' window in the IBM FlashSystem 900 monitoring interface. The breadcrumb navigation reads 'IBM FlashSystem 900 > Volumes > Volumes'. The user is logged in as 'superuser (Security Administrator)'. The window has a toolbar with 'Create Volumes', 'Actions', and 'Filter' buttons. Below the toolbar is a table showing a list of volumes. The table has columns: ID, Name, State, Capacity, Host Mappings, and Volume Unique Identifier. The table contains 10 rows of data. The status bar at the bottom right indicates 'Showing 9 volumes | Selecting 0 volumes'.

Figure 6-54 Volumes window that shows all volumes

## Creating a volume by using the GUI

To create a volume using the GUI from the Volumes menu, click **Create Volumes** (shown in Figure 6-54 on page 200).

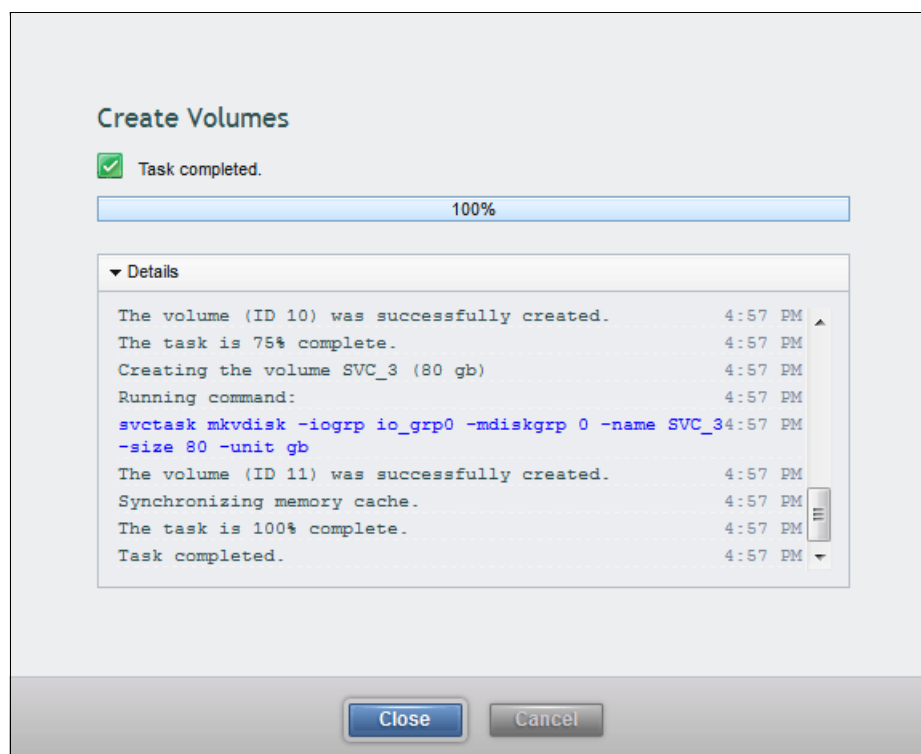
The Create Volumes window opens (Figure 6-55).



The 'Create Volumes' dialog box is shown. It has a title bar with a close button. The main area contains three input fields: 'Name:' with 'SVC' and a range '0 - 3', 'Quantity:' with '4', and 'Capacity:' with '80' and a unit dropdown set to 'GB'. Below these fields, it says 'Volumes being created: SVC\_0 - SVC\_3'. At the bottom are 'OK' and 'Cancel' buttons.

Figure 6-55 Create Volumes window

In this example, use volume name of SVC, a quantity of 4 volumes, and the requested capacity of 80 GB. Then, click **OK**; the task window opens (Figure 6-56).



The 'Create Volumes' task window is shown. It has a title bar. Below the title bar, there is a green checkmark icon and the text 'Task completed.' A progress bar shows '100%'. Below the progress bar is a 'Details' section with a scrollable list of messages and timestamps. The messages are: 'The volume (ID 10) was successfully created. 4:57 PM', 'The task is 75% complete. 4:57 PM', 'Creating the volume SVC\_3 (80 gb) 4:57 PM', 'Running command: 4:57 PM', 'svctask mkvdisk -iogrp io\_grp0 -mdiskgrp 0 -name SVC\_34:57 PM', '-size 80 -unit gb', 'The volume (ID 11) was successfully created. 4:57 PM', 'Synchronizing memory cache. 4:57 PM', 'The task is 100% complete. 4:57 PM', and 'Task completed. 4:57 PM'. At the bottom are 'Close' and 'Cancel' buttons.

Figure 6-56 Create Volumes task window

The Create Volume wizard now creates four volumes of 80 GB each. You can review the resulting volumes in the Volumes window (Figure 6-57).

ID	Name	State	Capacity	Host Mappings	Volume Unique Identifier
10	ESX_datastore_01	Online	50.00 GiB	3	60050764418d88a4400000000a000011
11	ESX_datastore_02	Online	50.00 GiB	3	60050764418d88a4400000000b000012
12	ESX_datastore_03	Online	50.00 GiB	3	60050764418d88a4400000000c000013
13	ESX_datastore_04	Online	50.00 GiB	3	60050764418d88a4400000000d000014
6	Exchange01_0	Online	80.00 GiB	1	60050764418d88a4400000000e00000d
7	Exchange01_1	Online	80.00 GiB	1	60050764418d88a4400000000f00000e
8	Exchange01_2	Online	80.00 GiB	1	60050764418d88a4400000000000000f
9	Exchange01_3	Online	80.00 GiB	1	60050764418d88a44000000001000010
0	ITSO_redbook_srv	Online	2.00 TiB	1	60050764418d88a44000000002000001
1	SVC_0	Online	80.00 GiB	No Host Mappings	60050764418d88a44000000003000015
2	SVC_1	Online	80.00 GiB	No Host Mappings	60050764418d88a44000000004000016
3	SVC_2	Online	80.00 GiB	No Host Mappings	60050764418d88a44000000005000017
4	SVC_3	Online	80.00 GiB	No Host Mappings	60050764418d88a44000000006000018

Figure 6-57 Four SAN Volume Controller volumes created

The newly created volumes have no host mappings at the time of their creation. Host mapping can be performed from the Volumes > Volumes by Host window. For the instructions to map volumes to a host, see “Mapping volumes” on page 207.

## Creating a volume by using the CLI

Example 6-2 shows the CLI process for creating a volume. More or fewer parameters can be applied to the `mkvdisk` command. Example 6-2 specifies the minimum required.

### Example 6-2 Create a volume by using the CLI

```
IBM_Flashsystem:FlashSystem_900:superuser>mkvdisk -size 15 -unit gb -name SVC_4
Virtual Disk, id [7], successfully created
```

```
IBM_Flashsystem:FlashSystem_900:superuser>lsvdisk
id name      IO_group_name status capacity vdisk_UID
0 WIN2008_1 io_grp0      online 40.00GB 0020c24000000000
1 WIN2008_2 io_grp0      online 50.00GB 0020c24001000000
2 WIN2008_3 io_grp0      online 50.00GB 0020c24002000000
3 WIN2008_4 io_grp0      online 39.99GB 0020c24003000000
4 SVC_1      io_grp0      online 15.00GB 0020c24004000000
5 SVC_2      io_grp0      online 15.00GB 0020c24005000000
6 SVC_3      io_grp0      online 15.00GB 0020c24006000000
7 SVC_4      io_grp0      online 15.00GB 0020c24007000000
```

```
IBM_Flashsystem:FlashSystem_900:superuser>
```

## Performing actions on volumes

From the Volumes window, you can perform various actions on the volumes. Click **Actions** to access these operations (Figure 6-58), or you can right-click the volume name, which opens a list of operations you can do to the volume.

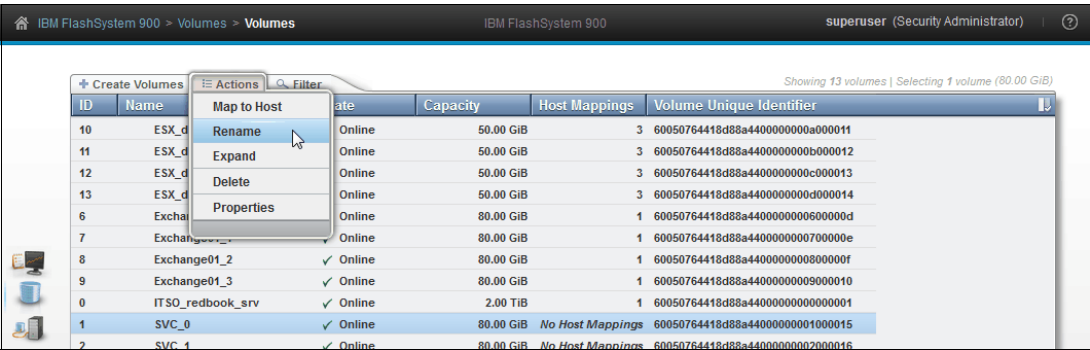


Figure 6-58 Actions of a single volume

Figure 6-59 shows the properties of a volume that indicate the volume name, its capacity, and its sector size. Each volume has a unique ID (UID), which can be discovered from the host side as a property of the logical unit. The volume is currently not mapped to a host.

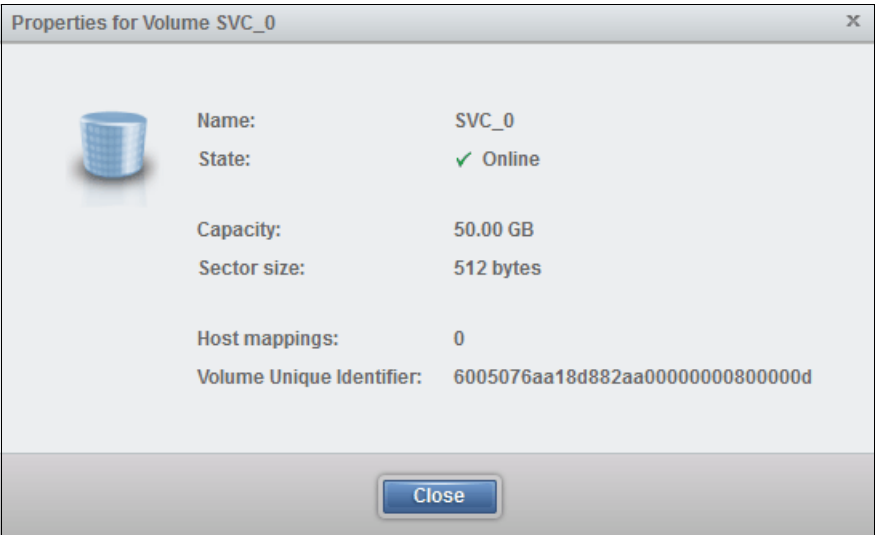


Figure 6-59 Properties of a volume

A volume can be expanded while it is online, therefore maintaining full functionality to the connected hosts. Not all operating systems, however, allow concurrent expansion of their disks so precaution must be taken that the operating system supports it. An alternative to expanding the disk is to create and map a new disk for the host.



## Expanding a volume that is mapped to an AIX host

When more than one volume is selected, the number of enabled actions for the volumes is reduced to only Expand and Delete (The Delete option is shown as disabled in Figure 6-60).

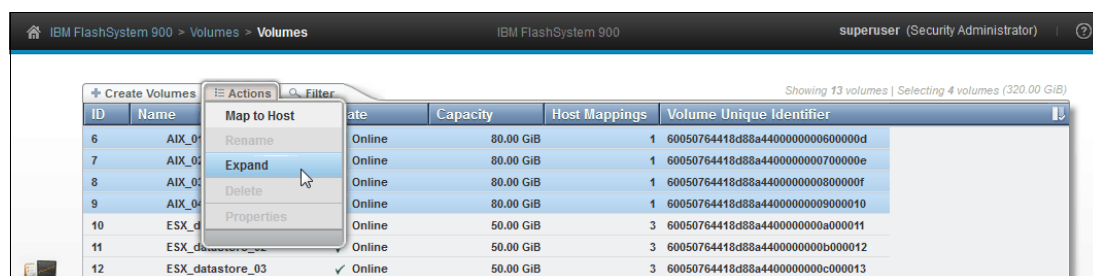


Figure 6-60 Expand four volumes

Figure 6-61 shows that each of the volumes should be expanded to a size of 120 GB.

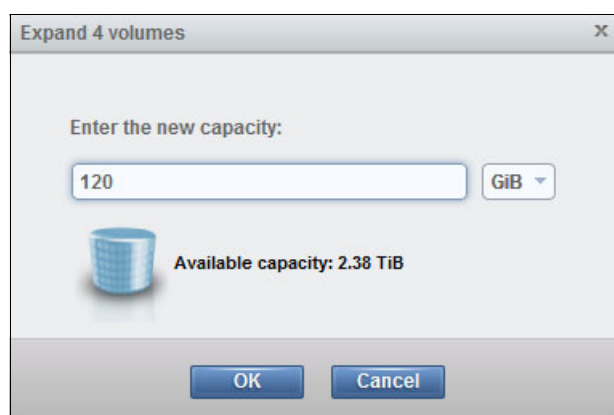


Figure 6-61 Expand to 120 GB

The resulting Volumes window displays the new capacity (Figure 6-62).



Figure 6-62 Four volumes expanded

The IBM FlashSystem 900 supports the ability to dynamically expand the size of a volume if the AIX host is using AIX version 5.2 or later.

Use the AIX **chvg** command options to expand the size of a physical volume that the Logical Volume Manager (LVM) uses without interruptions to the use or availability of the system. For more information, see *AIX 7.1 Operating System and Device Management* at the following web page; at this web page, search for operating system and device management:

<http://ibm.co/1cv2Ac1>



## Expanding a volume that is mapped to a Microsoft Windows host

You can use the GUI and the CLI to dynamically expand the size of a volume that is mapped to a Microsoft Windows host.

After expanding the volume, using the same procedure as shown in the previous examples (Figure 6-60 on page 204 and Figure 6-61 on page 204) for Windows, start the Computer Management application and open the Disk Management window under the Storage branch.

If the Computer Management application was open before you expanded the volume, use the Computer Management application to issue a **rescan** command. You will see the volume that you expanded now has unallocated space at the right side of the disk.

If the disk is a Windows basic disk, you can create a new primary or extended partition from the unallocated space.

If the disk is a Windows dynamic disk, you can use the unallocated space to create a new volume (simple, striped, or mirrored) or add it to an existing volume.

## Shrinking a volume

The shrink volume option is only provided through the CLI and cannot be performed by using the GUI. Volumes can be reduced in size, if necessary. However, if the volume contains data, do not shrink the size of the disk, because shrinking a volume will destroy the data.

When shrinking a volume, be aware of the following considerations:

- ▶ Shrinking a volume removes capacity from the end of the volume's address space. If the volume was used by an operating system or file system, predicting what space was used might be difficult. The file system or OS might be dependent on the space that is removed, even if it is reporting a large amount of free capacity.
- ▶ If the volume contains data that is used, do not attempt under any circumstances to shrink a volume without first backing up your data.

You can use the **shrinkvdiskspace** CLI command to shrink the physical capacity that is allocated to the particular volume by the specified amount.

The **shrinkvdiskspace** command uses this syntax:

```
shrinkvdiskspace -size capacitytoshrinkby -unit unitsforreduction vdiskname/ID
```

**Attention:** Shrinking a volume is a data-destructive action. Only shrink volumes that are not in use and that do not contain data.

Shrinking a volume is shown in Example 6-3. A volume is called a *vdisk* in the CLI.

*Example 6-3 Shrink a volume (vdisk)*

```
IBM_Flashsystem:Cluster_9.xx.xx.242:superuser>lsvdisk
id name      IO_group_name status capacity vdisk_UID
0 WIN2008_1 io_grp0      online 50.00GB 0020c24000000000
1 WIN2008_2 io_grp0      online 50.00GB 0020c24001000000
2 WIN2008_3 io_grp0      online 50.00GB 0020c24002000000
3 WIN2008_4 io_grp0      online 50.00GB 0020c24003000000

IBM_Flashsystem:Cluster_9.xx.xx.242:superuser>shrinkvdisksize -size 10 -unit gb
WIN2008_1

IBM_Flashsystem:Cluster_9.xx.xx.242:superuser>lsvdisk
id name      IO_group_name status capacity vdisk_UID
0 WIN2008_1 io_grp0      online 40.00GB 0020c24000000000
1 WIN2008_2 io_grp0      online 50.00GB 0020c24001000000
2 WIN2008_3 io_grp0      online 50.00GB 0020c24002000000
3 WIN2008_4 io_grp0      online 50.00GB 0020c24003000000

IBM_Flashsystem:Cluster_9.xx.xx.242:superuser>
```

### 6.3.3 Volumes by Host

Clicking the **Volumes** → **Volumes by Host** option opens the window where unmapped and mapped volumes are listed. This window shows which hosts are created on the system and which volumes exist. If the volumes are currently unmapped, they appear as Unmapped Volumes and a plus sign (+) is visible to the left (Figure 6-63).

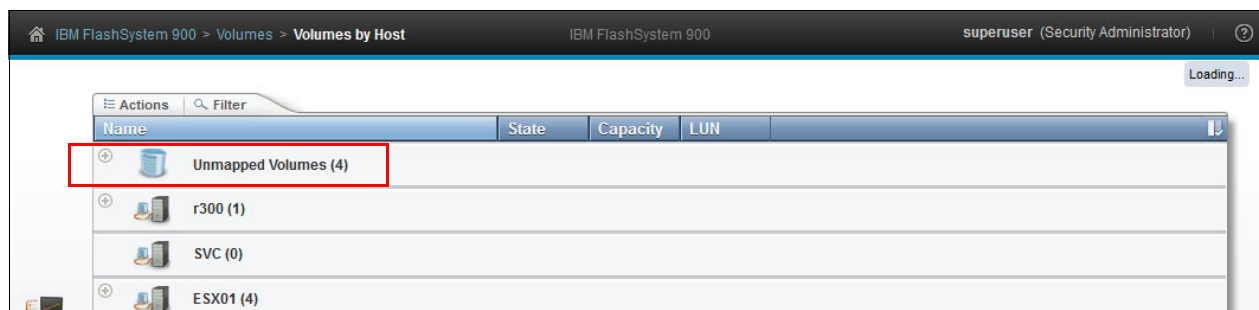


Figure 6-63 Volumes by Host window

## Mapping volumes

When you click the plus sign (+) to expand the Unmapped Volumes window, a list of unmapped volumes is provided. This example has four volumes named SVC\_1, SVC\_2, SVC\_3, and SVC\_4. To map them to the host SAN Volume Controller (SVC), highlight all four Volumes and click **Actions** → **Map to Host** (or right-click) as shown in Figure 6-64.

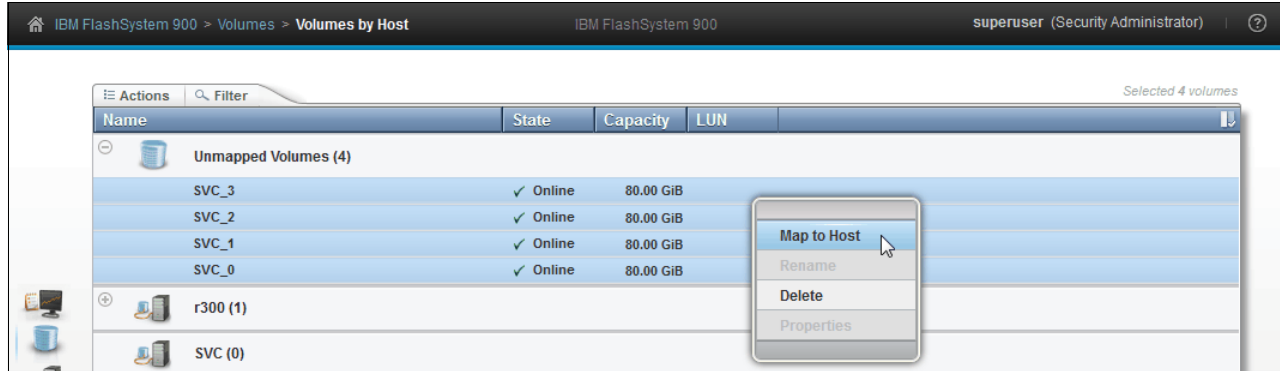


Figure 6-64 Map Volumes to Host

The Map 4 Volumes to Host window opens (Figure 6-65). Select host **SVC** and click **Map**.

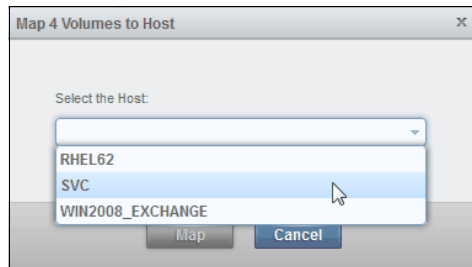


Figure 6-65 Map Volumes: Select the host SVC

Figure 6-66 shows the Modify Mappings window where the CLI commands for mapping the volumes are run.

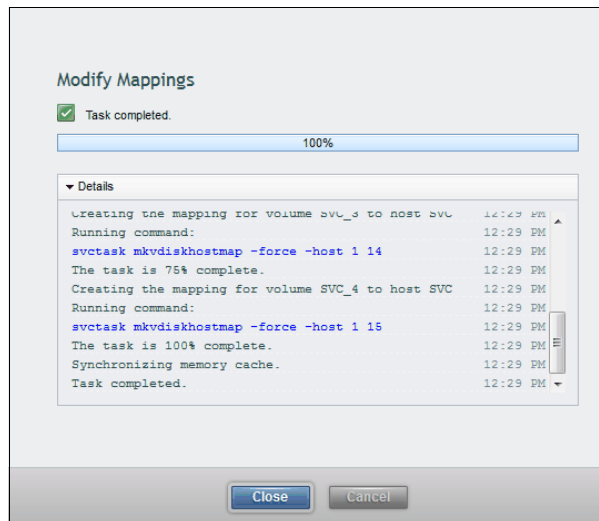


Figure 6-66 Map Volumes: CLI commands display

The Volumes by Host window now shows that the four SVC volumes are mapped and online (Figure 6-67).

Name	ID	State	Capacity	LUN	Volume Unique Identifier
Unmapped Volumes (0)					
r300 (1)					
SVC (4)					
SVC_0	1	✓ Online	80.00 GiB	3	60050764418d88a44000000001000015
SVC_2	3	✓ Online	80.00 GiB	0	60050764418d88a44000000003000017
SVC_1	2	✓ Online	80.00 GiB	1	60050764418d88a44000000002000016
SVC_3	4	✓ Online	80.00 GiB	2	60050764418d88a44000000004000018
ESX01 (4)					

Figure 6-67 Volumes mapped to host SVC

Figure 6-67 also shows that a modified menu bar to include the *LUN ID* and *Volume Unique Identifier* as described in “Change the Events view” on page 182.

## Mapping a volume by using the CLI

A useful way for administrators to perform this task might be to use the CLI. Volumes can be mapped by using the `svctask mkvdiskhostmap` command.

Example 6-4 shows how a volume is mapped to a host by using the CLI.

### Example 6-4 Map volume by using the CLI

```
IBM_Flashsystem:FlashSystem_900:superuser>svctask mkvdiskhostmap -force -host 0 4
Virtual Disk to Host map, id [4], successfully created
```

```
IBM_Flashsystem:FlashSystem_900:superuser>lsvdiskhostmap 4
id name SCSI_id host_id host_name vdisk_UID IO_group_id IO_group_name
4 SVC_5 4 0 SVC 0020c24004000000 0 io_grp0
```

```
IBM_Flashsystem:FlashSystem_900:superuser>
```

In the CLI process for mapping volumes, this example uses the logical number for the host and the logical number for the volume. These logical numbers can be discovered by using the following commands:

- ▶ **lshost**: Shows defined hosts and their status
- ▶ **lsvdisk**: Shows defined volumes and their preferences

## Unmapping volumes

When you delete a volume mapping, you are not deleting the volume, but you are deleting only the connection from the host to the volume. If you mapped a volume to a host by mistake or if you simply want to reassign the volume to another host, click **Volumes** → **Volumes by Host**. Highlight the volume or volumes that you want to unmap, right-click, and click **Unmap from Host** (Figure 6-68).

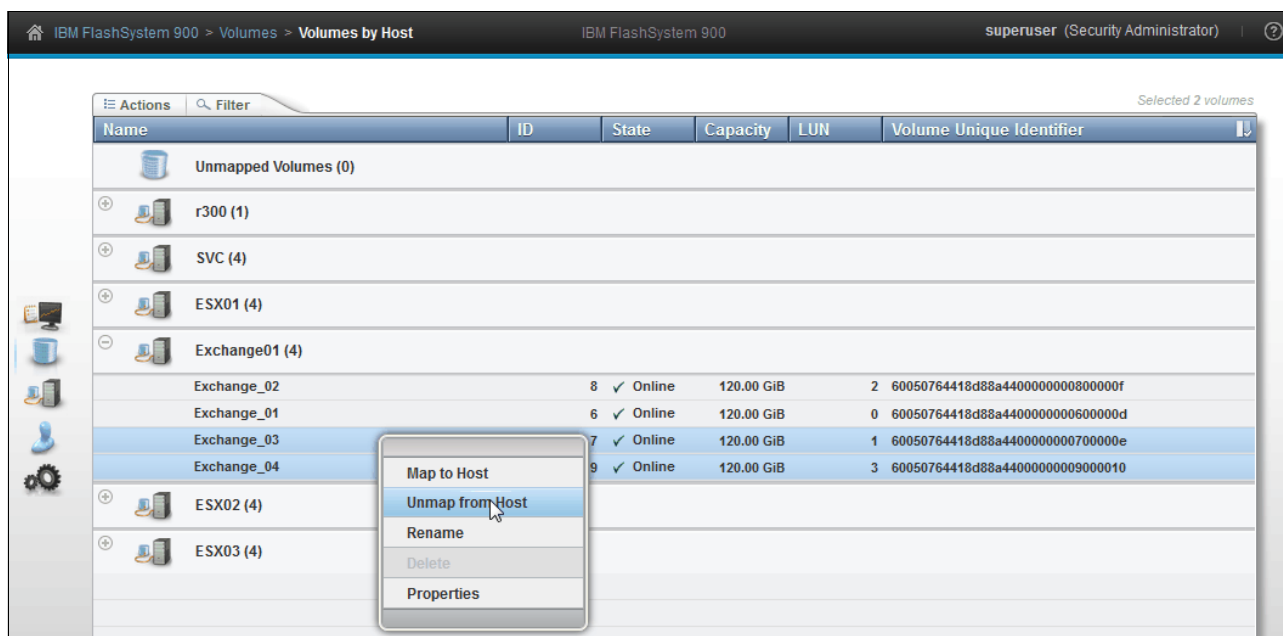


Figure 6-68 Unmap volumes from host

The Unmap 2 Volumes from Host window opens, indicating that the selected volumes will be unmapped (Figure 6-69).

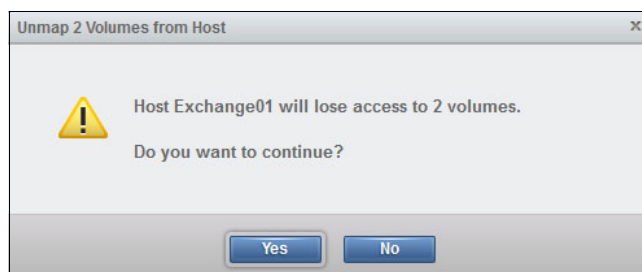


Figure 6-69 Unmapping volumes final step

By unmapping the volumes as shown in these windows, the volumes are made unavailable to the host. If data on the volumes is to be preserved, the host must unmount the disk before the volume is unmapped so that the connection to the disk is closed correctly by the host.

**Note:** Before unmapping a volume from a host, the host must unmount the connection to the disk or I/O errors appear.

After a volume is unmapped, it is listed in under Unmapped Volumes (Figure 6-70) of the FlashSystem 900 GUI. The volumes can now be deleted or can be mapped to another host.

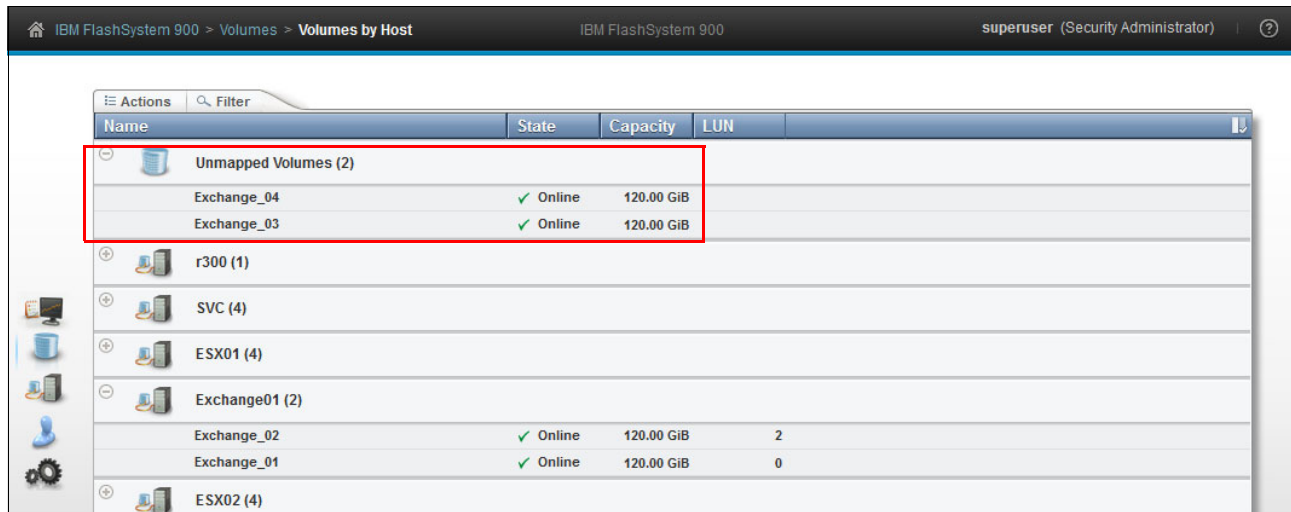


Figure 6-70 Unmapped volumes section

## 6.4 Hosts menu

Use the Hosts menu to manage hosts.

You can use the FlashSystem 900 GUI or the CLI **mkhost** command to create a logical host object. Creating a host object associates one or more worldwide port names (WWPNs) or InfiniBand IDs, of host bus adapters, with a logical host object.

You can then use the created host to map volumes (also called *virtual disks* or *VDisks*) to hosts by using the GUI or CLI **mkvdiskhostmap** command.

The Hosts menu has two options:

- ▶ Hosts
- ▶ Volumes by Host

## 6.4.1 Navigating to the Hosts menu

The GUI Hosts menu and its options are described. When you hover the cursor over the Hosts function icon, the Hosts menu opens (Figure 6-71).



Figure 6-71 Navigate to the Hosts menu

### Adding a host

The process of creating a host object includes specifying the host name and selecting ports for the host.

The FlashSystem 900 models are either InfiniBand, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), or iSCSI capable. However, interface cards cannot be mixed, and a single system must contain only a single type of interface card.

The FlashSystem 900 detects which type of interface cards are installed, and the Add Host wizard automatically adjusts to request the host port type for the actual model. For example, this can be the FC worldwide port name (WWPN), or the iSCSI initiator name or iSCSI qualified name (IQN).

Figure 6-72 shows the Hosts window where the defined hosts display.

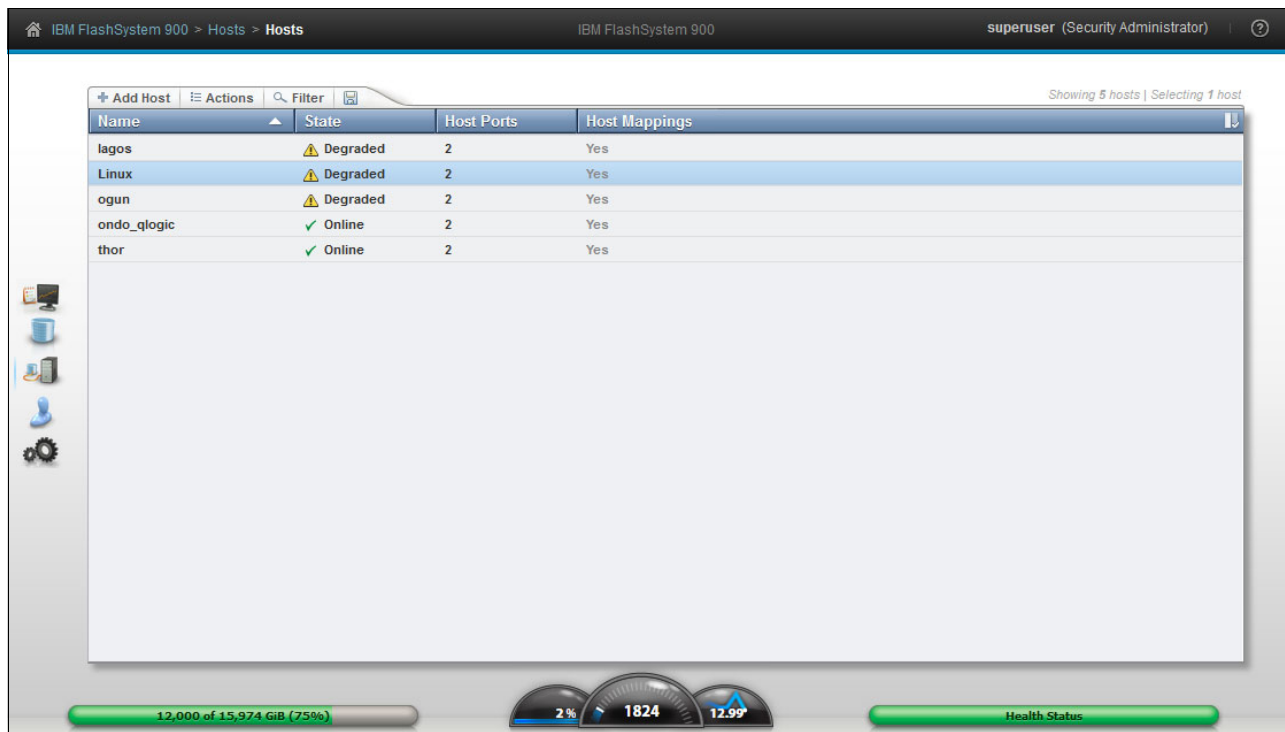


Figure 6-72 Hosts window showing already configured hosts

The Figure 6-72 example shows the Degraded state for the Linux host and other hosts. Each connected host initiator port *must* be zoned and connected to both canisters in the IBM FlashSystem 900. If not, the host reports a Degraded state. The risk of having a host with degraded paths is that the host might lose access to storage in case of canister failure or reboot, which then might cause unplanned downtime.

The FlashSystem 900 uses redundant controller canisters. These might reboot because of system failure. Canister reboot also happens multiple times during firmware update.



### Hosts in a FlashSystem 900 configured with FC interface cards

To create a host object, click **Add Host** in the upper-left corner of the Hosts window. The Add Host window opens. Type the name Exchange01 for the new host and click the arrow at the right side of Host port (WWPN) entry field to select the WWPNs for your Exchange server. Any WWPNs that are zoned to the system, but not already in use by a configured host, appear as shown in Figure 6-73.

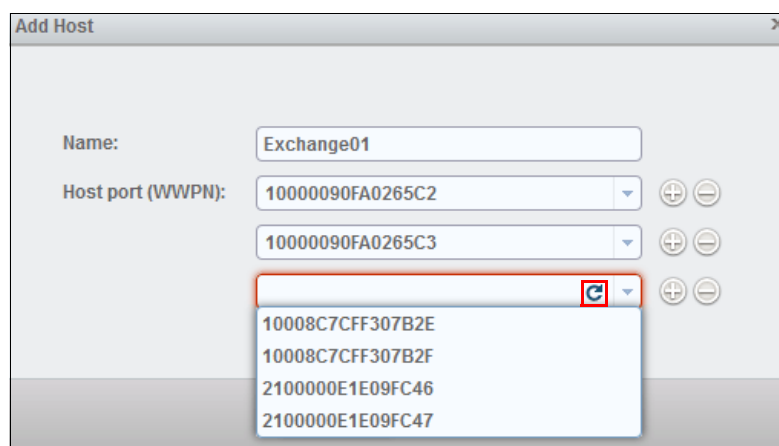


Figure 6-73 Add ports to the new host (the refresh icon is highlighted)

If no WWPNs are listed, the following message is displayed:

No candidate HBA ports were found

In that case, either the WWPN is already used for another host object or the host was not zoned correctly to the FlashSystem 900.

**Tip:** To refresh the list, click the refresh icon, which is highlighted in Figure 6-73.

Up to four WWPNs for each host can be selected. This example uses only two configured WWPNs for the Exchange server; click **Add** (Figure 6-74).

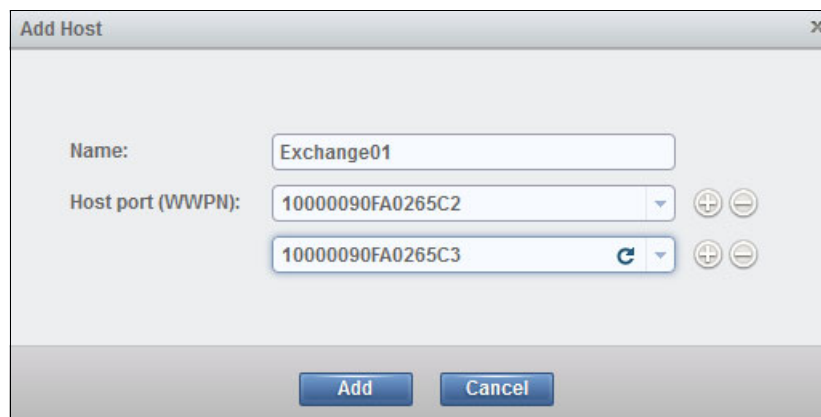


Figure 6-74 Add new host

The newly created Exchange01 host is now online and has two ports (Figure 6-75).

Name	State	Host Ports	Host Mappings
Exchange01	Online	2	No
lagos	Degraded	2	Yes
Linux	Degraded	2	Yes
ogun	Degraded	2	Yes
ondo_qlogic	Online	2	Yes
thor	Online	2	Yes

Figure 6-75 Host Exchange01 is now created and online

All WWPNs in a host object are mapped to the virtual disks.

Click Actions at the Hosts window to see the options for managing and examining the newly created host (Figure 6-76):

- ▶ Rename: rename the host
- ▶ Manage Host Ports: add or remove WWPNs
- ▶ Remove host: this is possibly only when no host mappings exist
- ▶ Host Ports: view status of host ports
- ▶ Properties: view properties of the host

Name	Host Ports	Host Mappings
Exchange01	2	No
lagos	2	Yes
Linux	2	Yes
ogun	2	Yes
ondo_qlogic	2	Yes
thor	2	Yes

Figure 6-76 View status of host ports

To view the properties and status of the host ports, click **Host Ports**. This example has two ports for the host *Exchange01*. The WWPNs are displayed and the status for each port is displayed as shown in Figure 6-77.

WWPN	State
10000090FA0265C2	Active
10000090FA0265C3	Active

Figure 6-77 Host ports for the Exchange01 host

A correctly zoned and configured host port displays as active. Right-click in the blue row at the top of the WWPN column so you can customize columns, to select what you want to see or clear what you do not want to see.

To add or remove host ports, click **Actions** → **Manage Host Ports** (Figure 6-78).

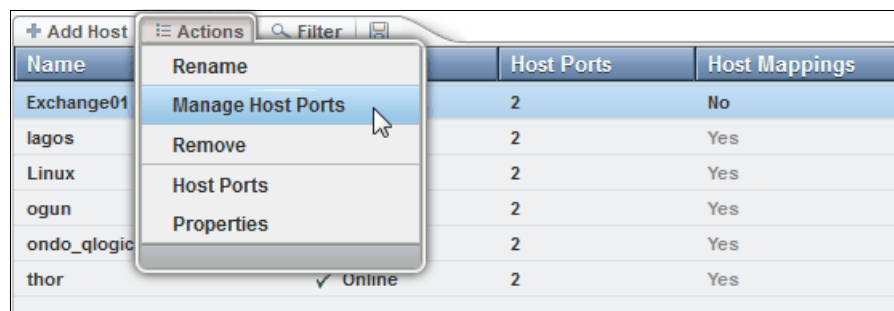


Figure 6-78 Manage Host Ports

In this example, you want to remove host ports, so select one of the Exchange01 host ports and click **Actions** → **Unassign Host Port** (Figure 6-79).

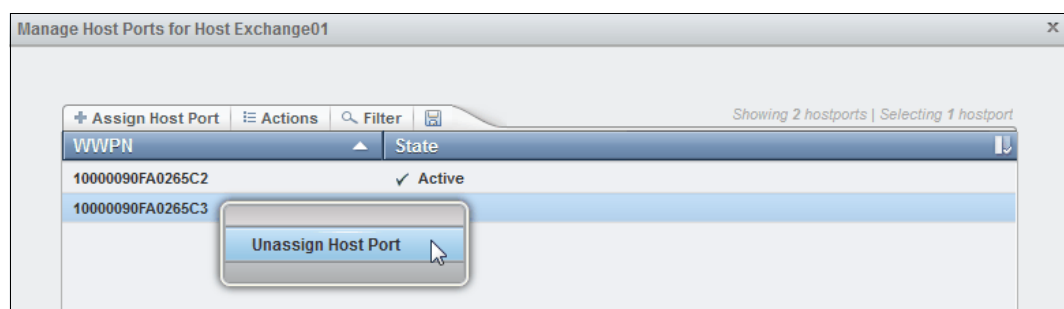


Figure 6-79 Manage host ports for host Exchange01 to remove ports

After unassigning the host port, click **Close** to return to the Hosts view (Figure 6-80) which shows that the Exchange01 host is Online and only a single port is active.

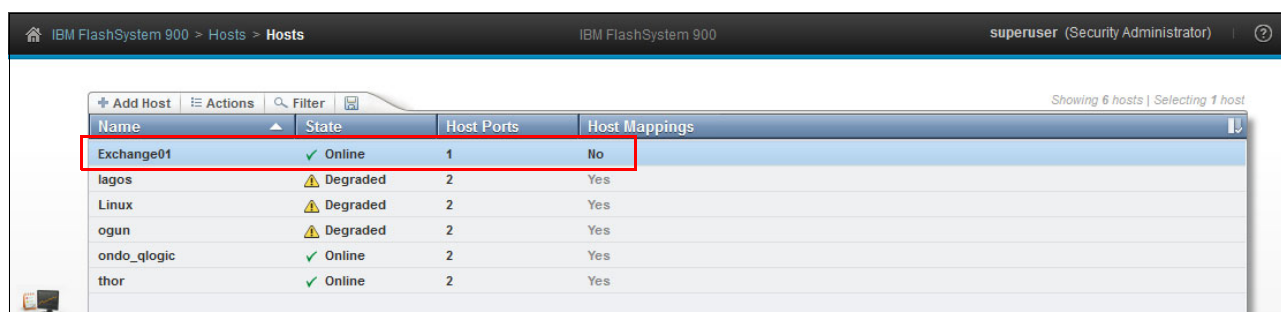


Figure 6-80 Host now only has a single WWPN port configured

Under normal conditions, any configured host should have at least two configured host ports for redundancy. However, you might have reasons for unassigning a host port as demonstrated in this example. One reason might be that if a host bus adapter (HBA) for a host is replaced, SAN zoning and storage systems must be reconfigured for the new host WWPN.

To create a host through the CLI, use the **svctask mkhost** and **svctask addhostport** commands, as demonstrated in Example 6-5.

```
IBM_Flashsystem:Cluster_9.xx.xx.xxx:superuser>svctask mkhost -fcwwpn
5005076801300004 -force -name SVC
Host, id [1], successfully created

IBM_Flashsystem:Cluster_9.xx.xx.xxx:superuser>svctask addhostport -fcwwpn
50050768013A0004 -force SVC
```

[illegible]

The **svctask addhostport** command must be run one time for every host port that you want to define for your host object.

### Hosts in a FlashSystem 900 configured with iSCSI interface cards

To check which iSCSI hosts exist, navigate to the Hosts > Hosts window (Figure 6-81). To create a new host click **Add Host** in the upper left corner of the Hosts window.

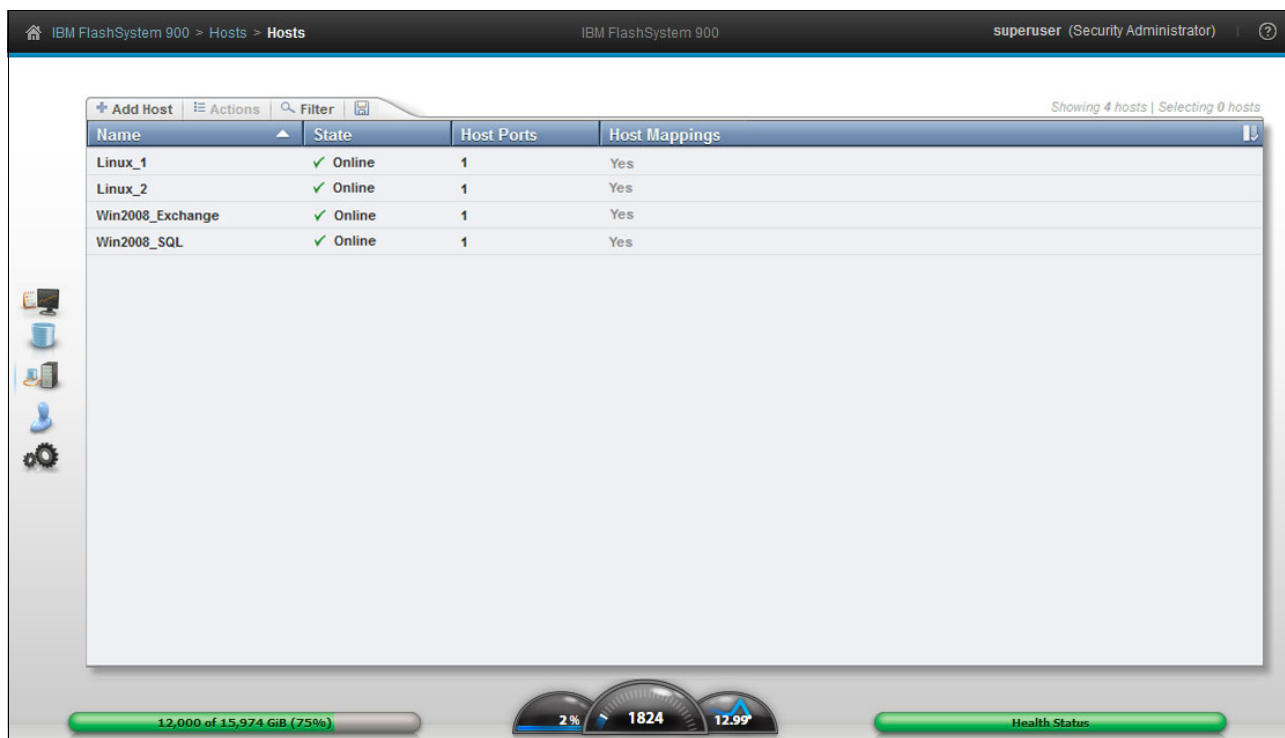


Figure 6-81 Review iSCSI hosts from the Hosts menu

The Add Host window opens (Figure 6-82). Type the host name and the iSCSI IQN name obtained from the iSCSI hosts iSCSI initiator software at the host to be connected. Click **Add** to create the host.

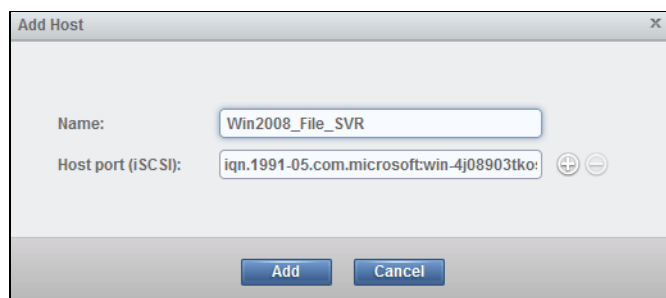


Figure 6-82 Create iSCSI host

Figure 6-83 shows the hosts, including the newly created Win2008\_File\_SRV host that you just created.

Name	State	Host Ports	Host Mappings
Linux_1	✓ Online	1	Yes
Linux_2	✓ Online	1	Yes
Win2008_Exchange	✓ Online	1	Yes
Win2008_File_SRV	✓ Online	1	No
Win2008_SQL	✓ Online	1	Yes

Figure 6-83 New iSCSI host created

The new host does not yet have any host mappings. Host mappings for iSCSI hosts can be created in the same way as host mappings for FC and FCoE systems, which was demonstrated in “Mapping volumes” on page 207.

As discussed in “Creating a host by using the CLI” on page 216, any single host *initiator* port must be able to communicate with both FlashSystem 900 controllers called the *target* ports. If not, hosts are not able to perform path failover in case of FlashSystem 900 canister reboot, which can cause unplanned downtime, because the iSCSI host loses access to storage.

That means that an iSCSI-attached host must have its iSCSI initiator software configured so that each of its iSCSI initiators connects to at least one iSCSI target port on each FlashSystem 900 canister.

A FlashSystem 900 configured for iSCSI has four interface cards, each with four iSCSI ports. To provide enough bandwidth for failover purposes and redundancy, a more suitable configuration is that each iSCSI initiator is configured to connect to one target port on each FlashSystem 900 interface card for a total of four connections.

For more information about how to configure host connectivity, see Chapter 5, “IBM FlashSystem 900 client host attachment and implementation” on page 101.

## 6.4.2 Volumes by Host

The Volumes by Host option on the *Hosts menu* is functionally identical to the Volumes by Host option on the *Volumes menu* (see 6.3.3, “Volumes by Host” on page 206).

## 6.5 Access menu

Various levels of user access to the FlashSystem 900 system are managed through the Access menu. The access levels are divided into groups. Each group has a different level of access and authority. If you want, multiple users can be defined and their access assigned to suit the tasks that they perform.

The Access menu has the following options:

- ▶ Users
- ▶ User Groups
- ▶ Audit Log

## 6.5.1 Navigating to the Access menu

Hover the cursor over the **Access** function icon, and a menu opens (Figure 6-84).



Figure 6-84 Navigate to the Access menu

The Access menu allows user management and audit log review.

User management includes the creation of new users and the maintenance of roles and passwords for existing users. Also, part of user management is the configuration of Secure Shell (SSH) keys to provide secure access to the CLI for users.

The audit log provides a list of all commands executed on the system and also contains information about which user ran the command.

## 6.5.2 Users window

Figure 6-85 shows the Users window. From this window you can create and delete users, change and remove passwords, and add and remove SSH keys.

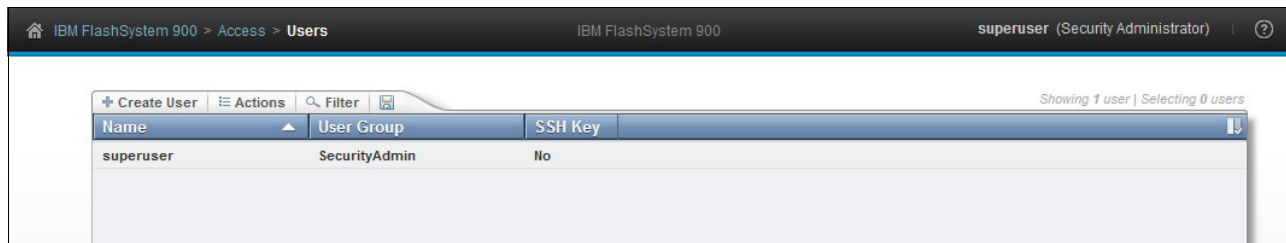


Figure 6-85 Users window

Click **Create User** to open the Create User window (Figure 6-86). You can enter the name of the user and the password, and load the SSH key (if the SSH key was generated). You can choose either to use SSH or a password for CLI authentication.

**Note:** You must have *superuser* authority to perform this action.

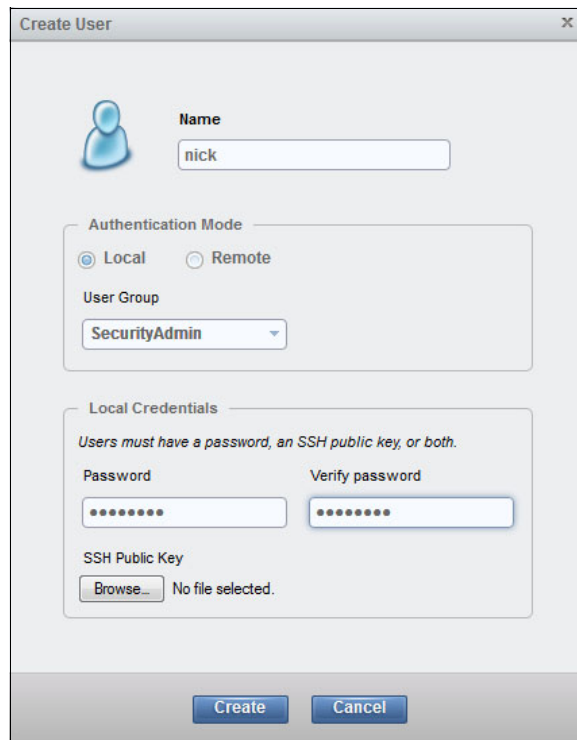
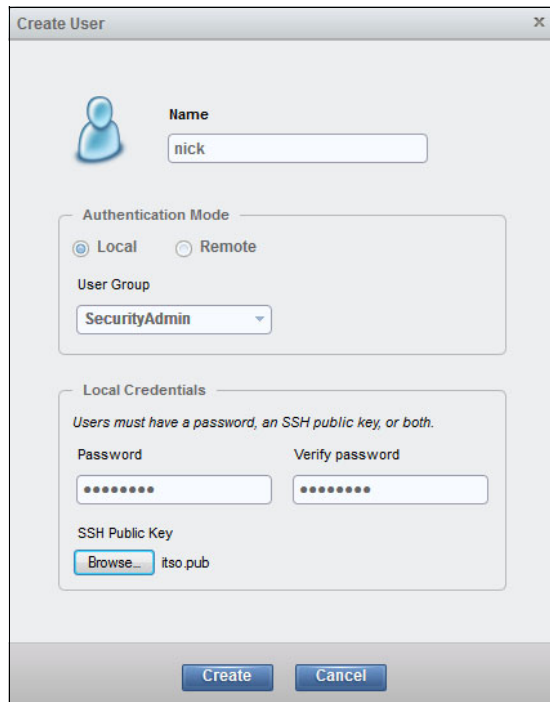


Figure 6-86 Create user

Figure 6-86 shows that a local user `nick` is created and a password is created for that user to authenticate to the system. When user `nick` opens his SSH client and points it to the IP address of the system to which he is granted access, he is prompted for the user name and password.

If the user is required to authenticate by using an SSH key pair, you instead enter the path for the public key in the SSH Public Key field in the Create User window, as shown in Figure 6-87 on page 221.





The 'Create User' dialog box is shown. It has a title bar with 'Create User' and a close button. Inside, there's a user icon and a 'Name' field with 'nick' entered. Below that is the 'Authentication Mode' section with 'Local' selected (radio button) and 'Remote' unselected. Under 'Local', there's a 'User Group' dropdown menu set to 'SecurityAdmin'. The 'Local Credentials' section has a note: 'Users must have a password, an SSH public key, or both.' It contains 'Password' and 'Verify password' fields, both filled with dots. Below these is an 'SSH Public Key' section with a 'Browse...' button and the text 'itso.pub'. At the bottom are 'Create' and 'Cancel' buttons.

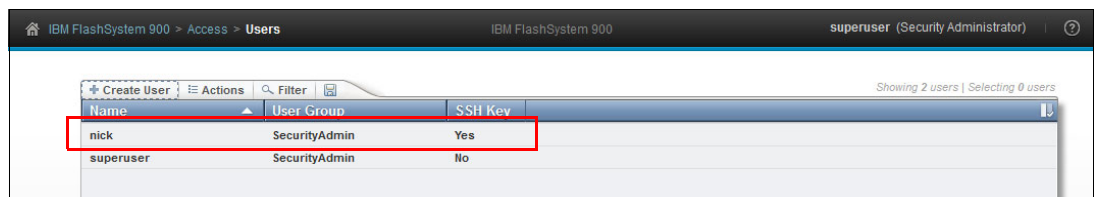
Figure 6-87 Create user as SSH key enabled

The Password and Verify password fields are used for GUI access. If a password is not configured, the user is not able to log in to the GUI.

When the SSH key is generated by using PuTTYgen, you have the choice of configuring a passphrase or not. If the SSH key pair was generated without a passphrase, the user `nick` is not prompted for a password when he opens his SSH client. He is then authenticated with the private key that matches the uploaded public key. If a passphrase was configured when the SSH key pair was created, the user `nick` also needs to type the passphrase password when opening the CLI to access the system.

For information about how to create SSH keys by using PuTTYgen, see “Generating an SSH key pair by using PuTTY” on page 223.

Now, the new user is listed (Figure 6-88) and an SSH key is enabled for that user.



The screenshot shows a web interface for 'IBM FlashSystem 900' with a breadcrumb 'Access > Users'. The user 'superuser (Security Administrator)' is logged in. A table lists users. The first row, for user 'nick', is highlighted with a red rectangle. The table has columns for Name, User Group, and SSH Key.

Name	User Group	SSH Key
nick	SecurityAdmin	Yes
superuser	SecurityAdmin	No

Figure 6-88 User is now created and SSH key enabled

Configuring CLI access, including how to configure SSH keys for secure access only, is described in 6.5.3, “Accessing CLI by using PuTTY” on page 223.

Click **Actions** to see list of options for managing users (Figure 6-89).

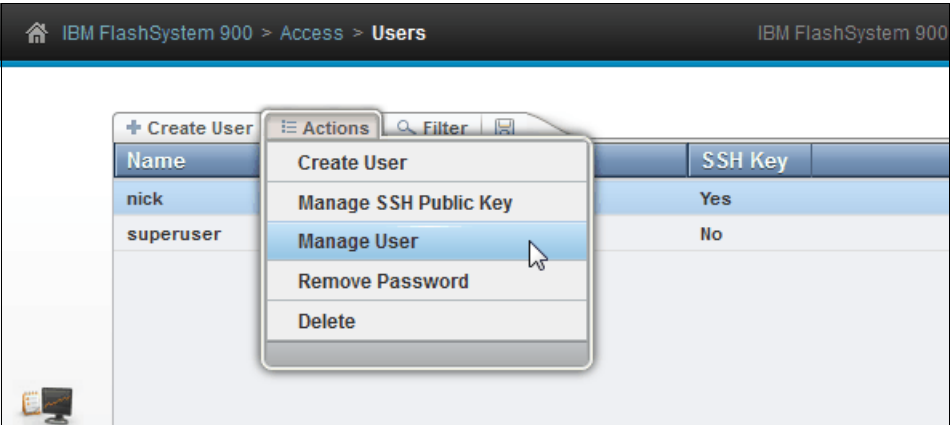


Figure 6-89 Manage user actions

To test whether the new user can log in to the GUI, log out as user superuser and log in as user nick. The login window is shown in Figure 6-90.



Figure 6-90 Log in as user nick

The user area of the GUI in the upper-right corner now shows that user nick is the current user, as shown in Figure 6-91.



Figure 6-91 User nick logged in

### 6.5.3 Accessing CLI by using PuTTY

PuTTY is a no-charge implementation of Telnet and SSH for Windows and UNIX platforms. PuTTY can be downloaded from the following website:

<http://www.putty.org/>

The CLI commands for the IBM FlashSystem 900 use the SSH connection between the SSH client software on the host system and the SSH server on the system.

You must create a system *before* you can use the CLI.

To use the CLI from a client system, follow these steps:

1. Install and set up the SSH client software on each system that you plan to use to access the CLI.
2. Authenticate to the system using a password.
3. If you require command-line access without entering a password, use an SSH public key. Then, store the SSH public key for each SSH client on the system.

#### Generating an SSH key pair by using PuTTY

To use the CLI with SSH keys enabled, you must generate an SSH key pair. SSH keys can be generated from a Windows host using the PuTTY key generator, PuTTYgen, by completing the following steps:

1. Run **puttygen.exe**.
2. Click **SSH-2 RSA** as the type of key to generate.  
Keep the number of bits in a generated key value at 1024.
3. Click **Generate** and then move the cursor around the blank area of the Key section to generate the random characters that create a unique key. When the key is completely generated, the information about the new key is displayed in the Key section.
4. Optional: Enter a passphrase in the Key passphrase and Confirm passphrase fields. The passphrase encrypts the key on the disk; therefore, it is not possible to use the key without first entering the passphrase.
5. Save the public key by following these steps:
  - a. Click **Save public key**. You are prompted for the name and location of the public key.
  - b. Type `icat.pub` as the name of the public key and specify the location where you want to save the public key. On your computer for example, you can create a directory called `keys` to store both the public and private keys.
  - c. Click **Save**.
6. Save the private key by following these steps:
  - a. Click **Save private key**. The PuTTYgen Warning panel is displayed.
  - b. Click **Yes** to save the private key without a passphrase.
  - c. Type `icat` as the name of the private key, and specify the location where you want to save the private key. For example, you can create a directory on your computer called `keys` to store both the public and private keys. We suggest that you save your public and private keys in the same location.
  - d. Click **Save**.
7. Close the PuTTY Key Generator window.

Figure 6-92 shows the folder where you saved the public and private keys that were generated using PuTTYgen.

Name	Date modified	Type	Size
dennis.ppk	01-10-2013 13:31	PPK File	2 KB
dennis.pub	01-10-2013 13:31	PUB File	1 KB
puttygen.exe	01-10-2013 13:30	Application	180 KB

Figure 6-92 SSH keys created using PuTTYgen

## Accessing the CLI with SSH keys enabled

To access the FlashSystem 900 CLI using PuTTY as the SSH client, a several configurations must be performed with PuTTY:

1. Type the IP address or name of the FlashSystem 900. The name can be used if the name resolution, Domain Name System (DNS), is configured. Ensure that **SSH** is selected and that port 22 is used for the connection, as shown in Figure 6-93.

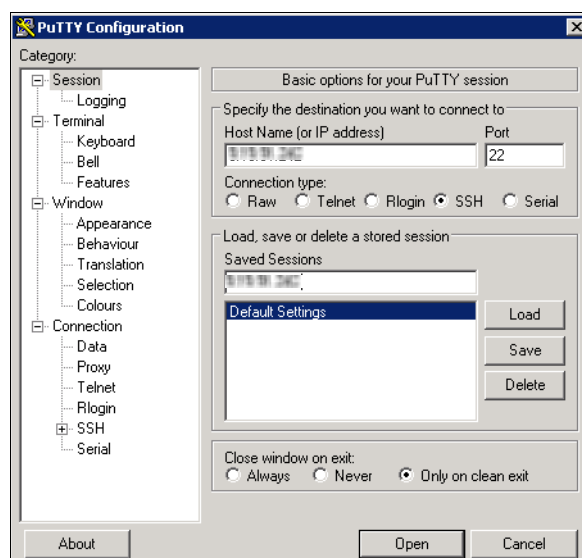


Figure 6-93 Configure PuTTY (step 1 of 3)

2. For configuring the SSH client session with the correct private key for user dennis, complete the following steps.
  - a. Expand **Connection** → **SSH** and click **Auth**.
  - b. In the Private key file for authentication field, type the path for the SSH private key that matches the public key loaded on the FlashSystem 900 for the user dennis.

Figure 6-94 shows to load the private key for the user dennis.

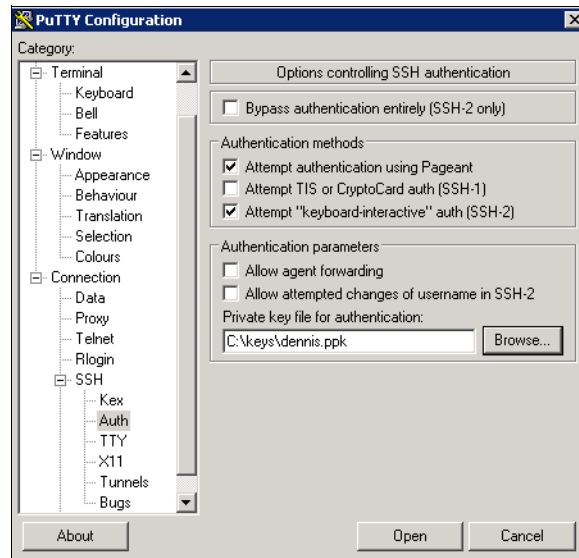


Figure 6-94 Configure PuTTY step 2 of 3

3. To save the connection, type a name and click **Save**. Now, the connection settings are saved for the next time that you use them. Click **Open** to start the CLI session.

Figure 6-95 shows that the configured session is saved.

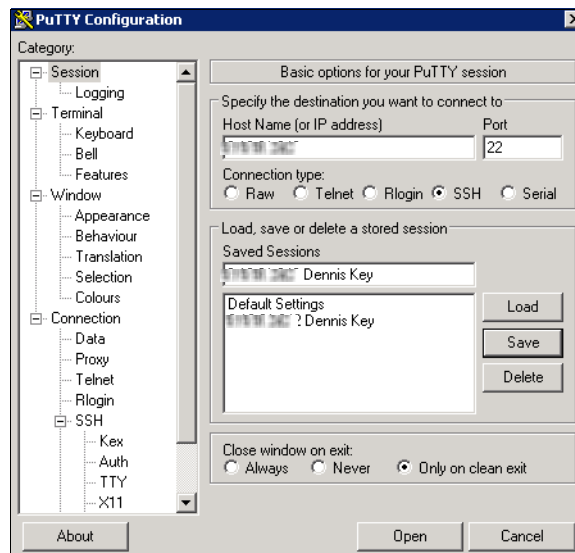


Figure 6-95 Configure PuTTY step 3 of 3

Example 6-6 shows that user dennis logs in to the CLI. He is prompted for a passphrase (password), which was configured when the SSH key pair was generated.

*Example 6-6 User dennis logs in to the CLI by using PuTTY*

```
login as: dennis
Authenticating with public key "rsa-key-20131022"
Passphrase for key "rsa-key-20131022":*****
Last login: Tue Oct 22 08:46:27 2013 from 9.xxx.xxx.xx
```

```
IBM_Flashsystem:FlashSystem_900:dennis>
```

Details about setting up CLI access and generating SSH keys are at the FlashSystem 900 web page in the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## 6.5.4 User groups

Administrators can create role-based user groups where any users who are added to the group adopt the role that is assigned to that group. Roles apply to both local and remote users on the system and are based on the user group to which the user belongs. A local user can belong to only a single group; therefore, the role of a local user is defined by the single group to which that user belongs. Users with the Security Administrator role can organize users of the system by role through user groups.

You can assign the user roles to users of the system, as shown in Table 6-1.

*Table 6-1 User groups available on the IBM FlashSystem 900*

Name	Role
SecurityAdmin	Users with this role can access all functions on the system, including managing users, user groups, and user authentication.
Administrator	Users with this role can access all functions on the system, except those functions that deal with managing users, user groups, and authentication.
Monitor	Users with this role can view objects and system configuration but they cannot configure, modify, or manage the system or its resources.
Service	Users with this role have monitor-role privileges and can view the system information, begin the disk-discovery process, and include disks that are excluded. This role is used by service personnel.

All users must be a member of a predefined user group.

### Creating a user group

To create a new user group, navigate to **Access** → **User Groups** and click **Create New user Group**.

Figure 6-96 on page 227 shows the User Groups window, which lists the default local user groups and one remote user group.

Name	Role	Users	Authentication
Administrator	Administrator	0	Local
Monitor	Monitor	0	Local
SecurityAdmin	SecurityAdmin	2	Local
Service	Service	0	Local

Figure 6-96 User Groups window

For more information about configuring remote authentication, see 7.1.3, “Security menu” on page 243.

The Create User Group window opens. Provide a Group Name and select a Role for the new user. If the group that is created is for remote authentication, select **Enable LDAP for this group** (available only for LDAP authentication; not required for the Operations group).

In this example, you create a new group named Operations and provide the role Monitor for this new group (Figure 6-97). The new group is a local group, and does not require Lightweight Directory Access Protocol (LDAP) authentication.

Group Name: Operations

Role:

- ☒ Monitor
- ☐ Service
- ☐ Administrator
- ☐ Security Administrator

Create Cancel

Figure 6-97 Create a user group called Operations

Click **Create** in the Create User Group window; the new group is created. To view the new group, click **Access** → **User Groups**.

The User Groups window opens (Figure 6-98).

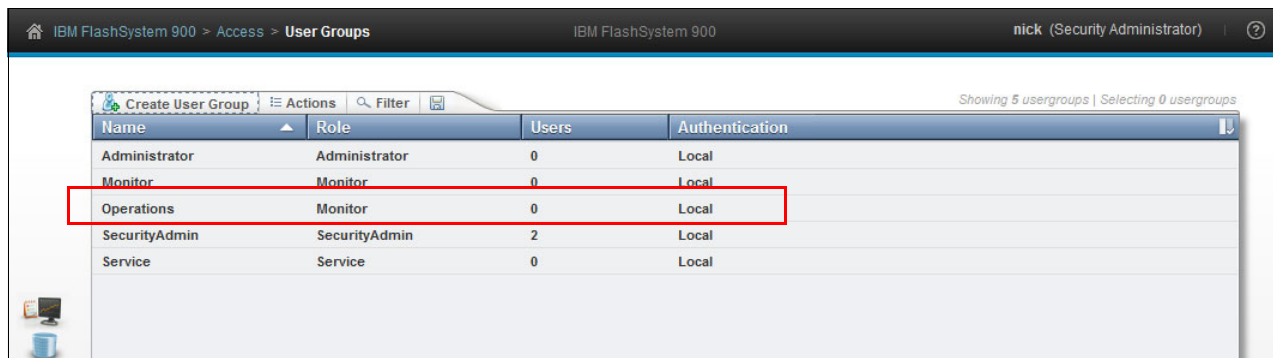


Figure 6-98 User group named Operations was created

After creating a new user group, add users to that group. Navigate to **Access** → **Users** and click **Create User**. The Create User window opens. Type the name of the new user and select the User Group **Operations** (Figure 6-99). A password must be typed and confirmed or the user is not able to log in to the GUI.

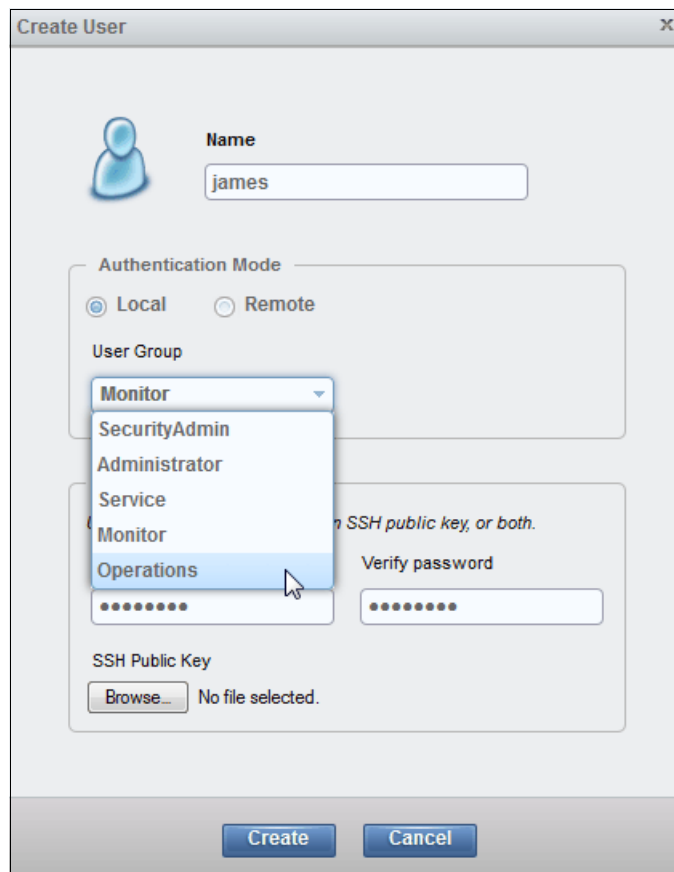


Figure 6-99 Create the user james and add it to a user group

After clicking **Create**, user james is now created and added to the Operations user group. User james can now log in to the GUI and CLI by using the provided password. This user will have the privileges from the user group Operations, which has the role Monitor.



**Note:** If remote authentication is not enabled in the Settings → Security menu, the Create User Group window does not give you the option to create the group for remote authentication and LDAP (Figure 6-97 on page 227).

### 6.5.5 Audit log menu

An *audit log* tracks action commands that are issued through an SSH session or through the management GUI.

The audit log entries provide the following information:

- ▶ Identity of the user who issued the action command
- ▶ Name of the actionable command
- ▶ Time stamp of when the actionable command was issued on the configuration node
- ▶ Parameters that were issued with the actionable command

The following commands are not documented in the audit log:

- ▶ `dumpconfig`
- ▶ `cpdumps`
- ▶ `cleardumps`
- ▶ `finderr`
- ▶ `dumperrlog`
- ▶ `dumpinternallog`
- ▶ `svcservicetask dumperrlog`
- ▶ `svcservicetask finderr`

The following items are also not documented in the audit log:

- ▶ Commands that fail are not logged.
- ▶ A result code of 0 (success) or 1 (success in progress) is not logged.
- ▶ The result object ID of the node type (for the **addnode** command) is not logged.
- ▶ Views are not logged.

#### Reviewing the audit log

To review the audit log, select **Access** → **Audit Log**. The Audit Log window opens (Figure 6-100 on page 230).

Date and Time	User Name	Command	Object ID
2/19/15 12:02:06 PM	nick	svctask mkuser -name james -usergrp 4 -password '#####'	2
2/19/15 11:58:38 AM	nick	svctask mkusergrp -name Operations -role Monitor	5
2/19/15 11:45:08 AM	superuser	svctask mkuser -name nick -usergrp 0 -password '#####'...	1
2/19/15 9:33:40 AM	superuser	svctask rmvdiskhostmap -host 3 7	
2/19/15 9:33:40 AM	superuser	svctask rmvdiskhostmap -host 3 9	
2/19/15 9:32:12 AM	superuser	svctask mkvdiskhostmap -host 3 -force 8	
2/19/15 9:32:12 AM	superuser	svctask mkvdiskhostmap -host 3 -force 9	
2/19/15 9:31:55 AM	superuser	svctask chvdisk -name Exchange_04 9	
2/19/15 9:31:43 AM	superuser	svctask chvdisk -name Exchange_03 7	
2/19/15 9:31:32 AM	superuser	svctask chvdisk -name Exchange_02 8	
2/19/15 9:31:13 AM	superuser	svctask chvdisk -name Exchange_01 6	
2/19/15 9:29:33 AM	superuser	svctask rmvdiskhostmap -host 3 9	
2/19/15 9:29:33 AM	superuser	svctask rmvdiskhostmap -host 3 8	
2/19/15 9:04:58 AM	superuser	svctask mkvdiskhostmap -host 1 -force 3	
2/19/15 9:04:58 AM	superuser	svctask mkvdiskhostmap -host 1 -force 2	
2/19/15 9:04:58 AM	superuser	svctask mkvdiskhostmap -host 1 -force 4	
2/19/15 9:04:58 AM	superuser	svctask mkvdiskhostmap -host 1 -force 1	
2/19/15 9:04:41 AM	superuser	svctask rmvdiskhostmap -host 1 1	
2/19/15 9:04:41 AM	superuser	svctask rmvdiskhostmap -host 1 4	
2/19/15 9:04:40 AM	superuser	svctask rmvdiskhostmap -host 1 3	
2/19/15 9:04:39 AM	superuser	svctask rmvdiskhostmap -host 1 2	

Figure 6-100 Audit Log window

The audit log can be filtered to display only selected users or within a specific time frame. The audit log can also be filtered to show from which IP address a specific command was run.

Figure 6-101 shows filtering options where you add the IP address of the logged in user.

Date and Time	User Name	Command	Object ID
2/19/15 9:33:40 AM		svctask mkuser -name james -usergrp 4 -password '#####'	2
2/19/15 9:33:40 AM		svctask mkusergrp -name Operations -role Monitor	5
2/19/15 9:32:12 AM		svctask mkuser -name nick -usergrp 0 -password '#####'...	1
2/19/15 9:32:12 AM		svctask rmvdiskhostmap -host 3 7	
2/19/15 9:31:55 AM		svctask rmvdiskhostmap -host 3 9	
2/19/15 9:31:43 AM		svctask mkvdiskhostmap -host 3 -force 8	
2/19/15 9:31:32 AM		svctask mkvdiskhostmap -host 3 -force 9	
2/19/15 9:31:13 AM		svctask chvdisk -name Exchange_04 9	
2/19/15 9:29:33 AM		svctask chvdisk -name Exchange_03 7	
2/19/15 9:29:33 AM		svctask chvdisk -name Exchange_02 8	
2/19/15 9:29:33 AM		svctask chvdisk -name Exchange_01 6	
2/19/15 9:04:58 AM	superuser	svctask rmvdiskhostmap -host 3 9	
2/19/15 9:04:58 AM	superuser	svctask rmvdiskhostmap -host 3 8	
2/19/15 9:04:58 AM	superuser	svctask mkvdiskhostmap -host 1 -force 3	

Figure 6-101 Viewing the Audit Log menu by customizing columns

Also, you can export the audit log to a comma-separated file.



## Configuring settings

This chapter describes the Settings function (Figure 7-1 on page 232) of the IBM FlashSystem 900 graphical user interface (GUI). The Settings function offers various options for monitoring, configuring network interfaces, and extracting support logs. It also offers remote authentication and the firmware update process, and explains how to access IBM FlashSystem 900 Service Assistant Tool.

This chapter covers the following topics:

- ▶ Settings menu
- ▶ Service Assistant Tool

## 7.1 Settings menu

You can use the Settings function to configure system options for event notifications, security, IP addresses, FC connectivity, and preferences related to display options in the management GUI.

The Settings menu includes these options:

- ▶ Notifications: Alerting
- ▶ Network: Management and service
- ▶ Security: Remote authentication with Lightweight Directory Access Protocol (LDAP)
- ▶ System: Time settings, firmware update, and so on
- ▶ Support: Extract support logs
- ▶ GUI Preferences

### 7.1.1 Navigating to the Settings menu

If you hover the cursor over the Settings function icon, the Settings menu opens (Figure 7-1).

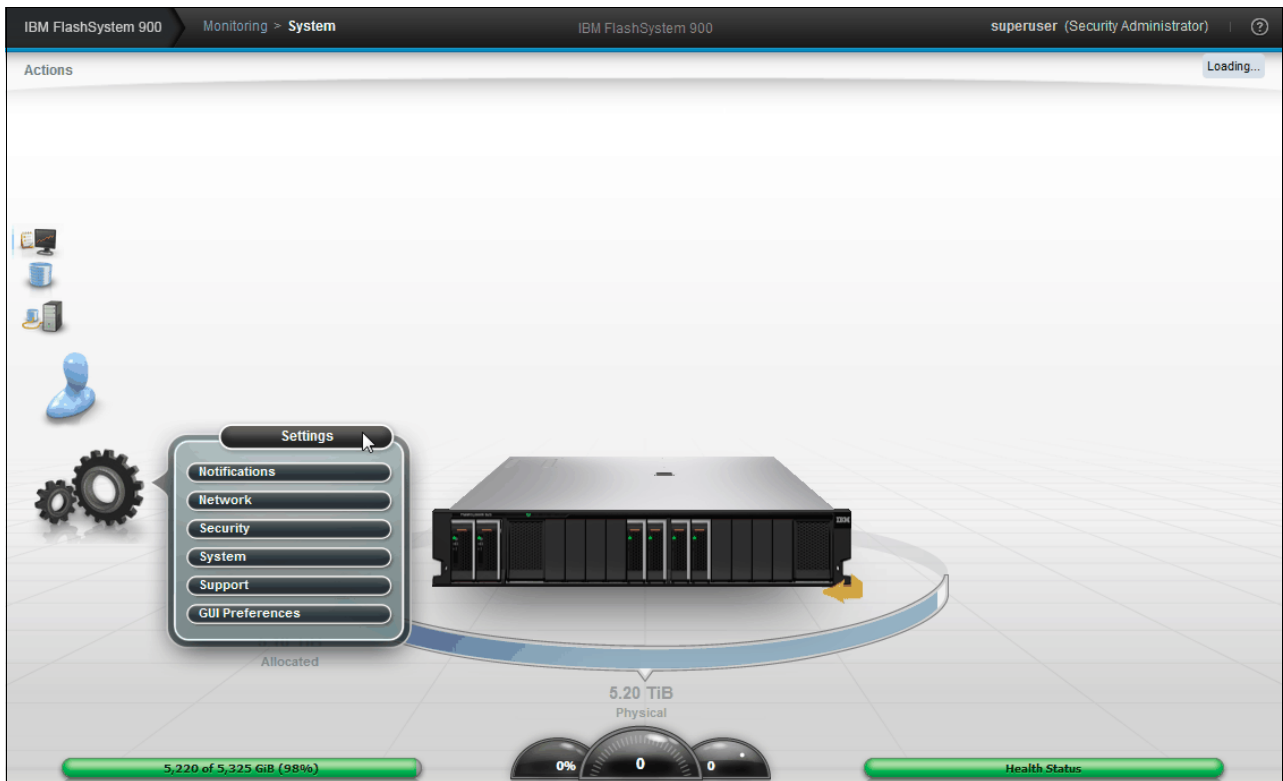


Figure 7-1 Navigate to the Settings menu

## 7.1.2 Notifications menu

FlashSystem 900 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and call home email to notify you and IBM Support 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 exceptions. For example, event 085031 “Array is missing a spare flash module”, (SEC 1690) does not report until the rebuild is complete. Also, certain events might occur because of service actions that are performed. If a recommended service action is active, these events are sent only if they are still not fixed when the service action completes.

### Email

The call home feature transmits operational and event-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues.

To configure email alerts, navigate to **Settings** → **Notifications**. The Notifications window opens (Figure 7-2). From this window, you can configure the email alerts that are included in the call home function. During the configuration of call home, you configure contact information and email receivers for your own notification.

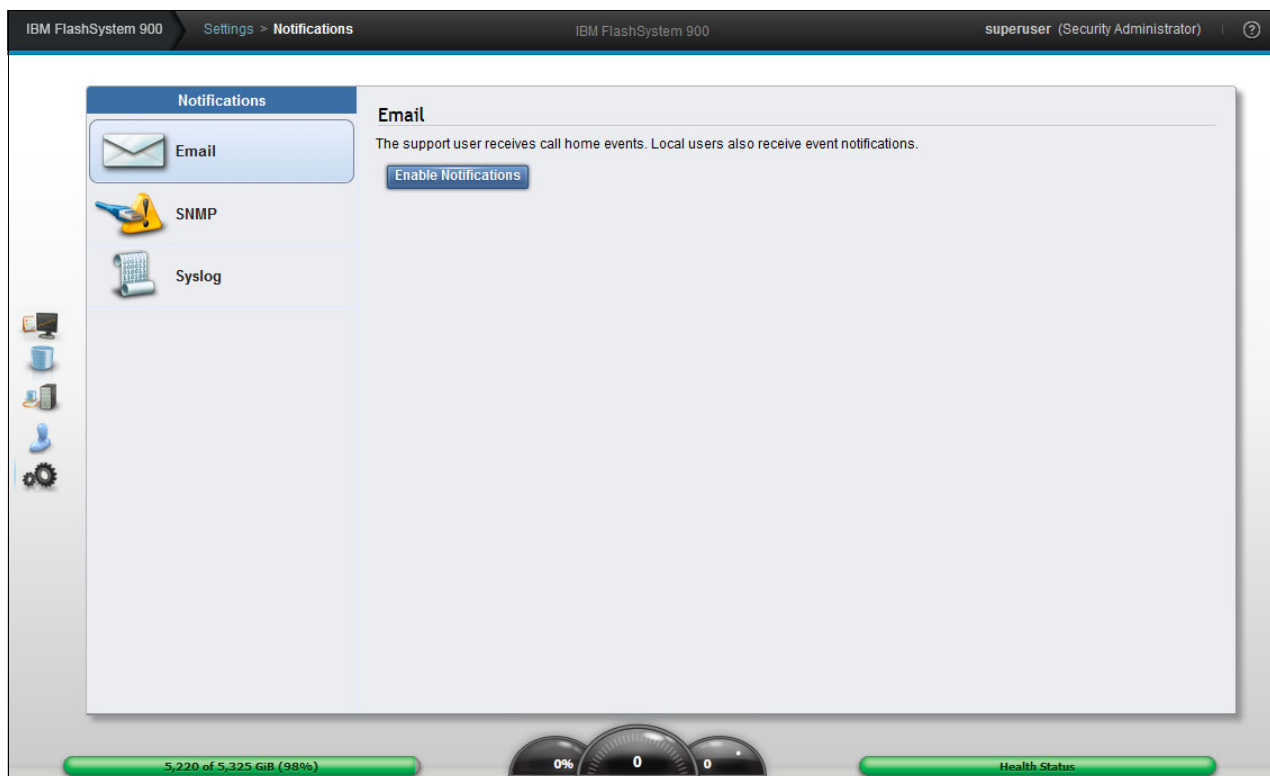


Figure 7-2 Event Notifications Email window

To initiate the configuration wizard, click **Enable Notifications**; the setup call home window opens. The procedure for configuring call home is similar to the initialization of the IBM FlashSystem 900, which also offers configuration of Email Event Notifications. The initial System setup is described in “System Setup wizard” on page 88.

The steps to enable call home involve configuring these items:

- ▶ System Location: Where is the system located?
- ▶ Contact Details: Who should IBM contact in case of a call-home situation?
- ▶ Email Servers: What is the IP address of the SMTP email server?
- ▶ Email Notifications: Who else requires notification through email?

Figure 7-3 shows the Welcome window in the Email Events Notification wizard. Click **Next** to continue.

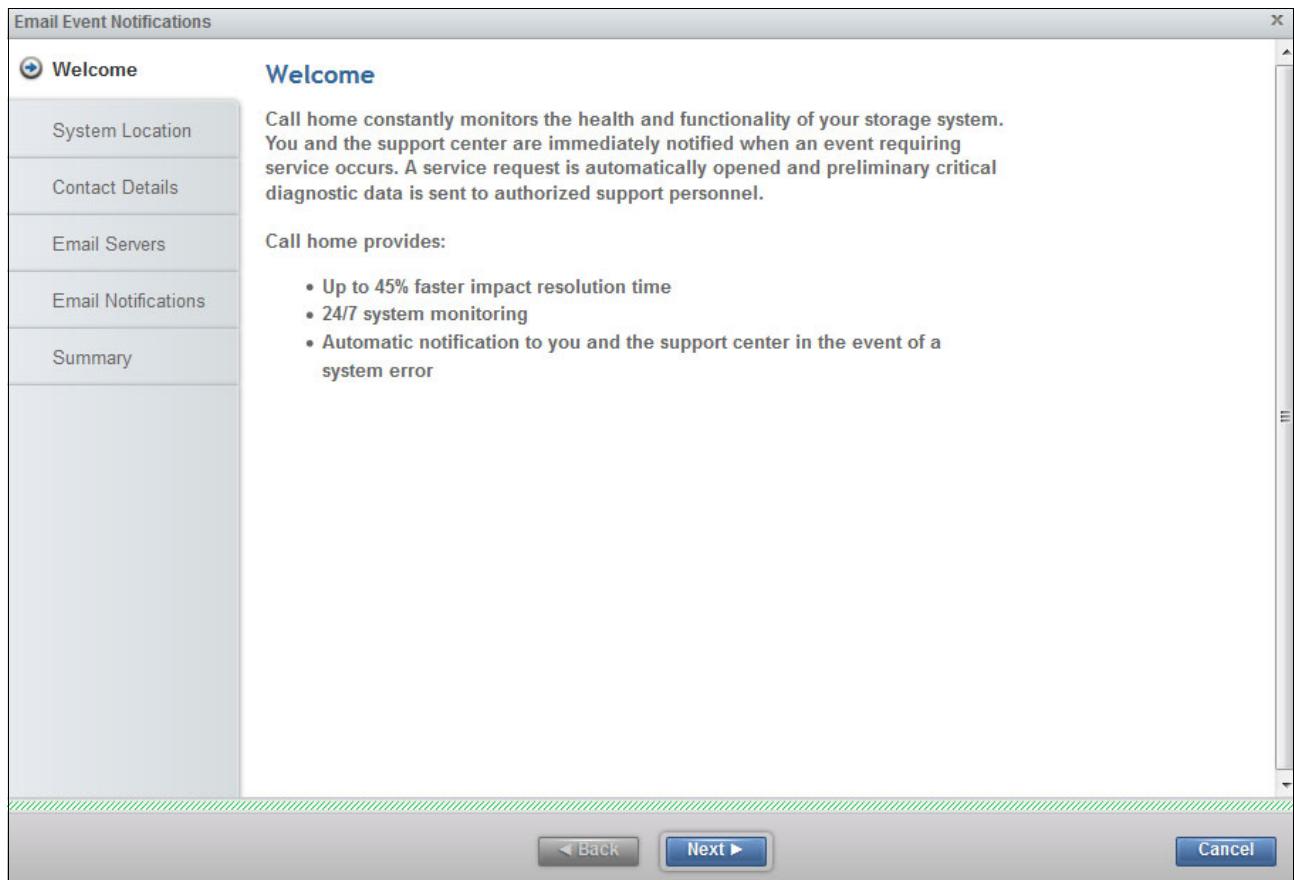


Figure 7-3 Notifications Welcome

Figure 7-4 shows how the System Location details are configured. This is the address and location of the system, and where IBM Support assumes the system is located when a call home is received. Click **Next** to continue.

The screenshot shows a window titled "Email Event Notifications" with a sidebar on the left containing the following links: "Welcome" (with a checkmark icon), "System Location" (with a plus icon and highlighted), "Contact Details", "Email Servers", "Email Notifications", and "Summary". The main content area is titled "System Location" and includes the text "Parts are shipped to this address." Below this, there are several form fields: "Company name:" with the value "IBM ITSO", "Street address:" with the value "10777 Westheimer rd.", "City:" with the value "Houston", "State or province:" with the value "TX", "Postal code:" with the value "77077", and "Country or region:" with a dropdown menu showing "US". At the bottom of the window, there are three buttons: "Back", "Next", and "Cancel".

Figure 7-4 Notifications System Location

**Note:** The *State or province* field and *Country or region* field must be configured correctly, otherwise IBM Support is not able to react on call home calls from the system.

You can test the validity of this field by submitting an Email Notification test. Wait until the Email Notification wizard finishes to submit your test. IBM Support will contact your organization if settings are configured correctly.

Figure 7-5 shows how the client contact is configured. To continue, click **Apply and Next**.

The screenshot shows a web-based configuration window titled "Email Event Notifications". On the left is a vertical sidebar with a list of steps: "Welcome" (checked), "System Location" (checked), "Contact Details" (active, indicated by a blue arrow), "Email Servers", "Email Notifications", and "Summary". The main content area is titled "Contact Details" in blue. Below the title is a descriptive sentence: "The support center contacts this person to resolve issues on the system." There are five input fields with labels to their left: "Contact name:" with the value "John Doe", "Email address:" with the value "jdoe@ibm.com", "Telephone (primary):" with the value "1234567890", "Telephone (alternate):" which is empty, and "Machine location:" with the value "third floor, room 304". At the bottom of the window are three buttons: "Back" (with a left arrow), "Apply and Next" (with a right arrow), and "Cancel".

Figure 7-5 Notifications Contact Details

To send email, the IBM FlashSystem 900 needs to know which mail server can transport the email. It also needs to know the TCP port through which the emails are sent.



Figure 7-6 demonstrates how to configure an SMTP email server. Multiple SMTP email servers can be configured by clicking the plus sign (+) to the right of the Server Port box. Click **Apply and Next** to continue.

The screenshot shows a web-based configuration interface for 'Email Event Notifications'. On the left is a sidebar with a list of steps: 'Welcome', 'System Location', 'Contact Details', and 'Email Servers' (which is highlighted with a blue arrow icon). Below 'Email Servers' are two sub-items: 'Email Notifications' and 'Summary'. The main content area is titled 'Email Servers' in blue. Below the title is a descriptive text: 'Call home and event notifications are routed through this email server.' There are two input fields: 'IP Address' containing '10.10.10.10' and 'Server Port' containing '25'. To the right of the 'Server Port' field are two small circular buttons with '+' and '-' signs. Below these fields is a 'Ping' button. At the bottom of the window, there is a horizontal bar with three buttons: 'Back', 'Apply and Next', and 'Cancel'.

Figure 7-6 Notifications Email Servers

The call home feature from the IBM FlashSystem 900 is enabled by sending an email to IBM at a fixed email destination address that cannot be altered.

Optionally, you can configure event notifications to additional email receivers. Email notification to IBM Support is automatically configured, but typically the client prefers to be notified if any issues occur that need attention. Client event notification is valuable if email transport to IBM fails. An email transport error can occur in an SMTP server outage, or if the SMTP server IP address is changed without the call home function of the IBM FlashSystem 900 being updated correctly.

Additional mail receivers can be configured, as shown in Figure 7-7. Click **Apply and Next** to continue.

Email Event Notifications

- ✓ Welcome
- ✓ System Location
- ✓ Contact Details
- ✓ Email Servers
- ➔ Email Notifications
- Summary

### Email Notifications

The call home email to the support center is automatically configured. You can, optionally, add users to receive event notifications.

Email	Event Type
flash-sc2@vnet.ibm.com	Errors, Inventory
monitor@ibm.com	Alerts

◀ Back    Apply and Next ▶    Cancel

Figure 7-7 Notifications email receivers

A summary is displayed when all configurations are ready to be applied (Figure 7-8). Click **Finish** to apply the Email Notification settings.

**Email Event Notifications**

Summary

**Contact Details**

Contact name:	John Doe
Email address:	jdoe@ibm.com
Telephone (primary):	1234567890
Telephone (alternate):	
Machine location:	third floor, room 304

**System Location**

Company name:	IBM IT SO
Street address:	10777 Westheimer rd.
City:	Houston
State or province:	TX
Postal code:	77077
Country or region:	US

**Email Servers**

IP address:	10.10.10.10
Port:	25

**Call Home**

Support center:	flash-sc2@vnet.ibm.com
Alerts:	Errors, Inventory

**Email Notifications**

Email address:	monitor@ibm.com, Alerts
----------------	-------------------------

◀ Back   Finish   Cancel

Figure 7-8 Notifications summary

Event notifications are now active and the system is reporting to IBM in case of hardware failure. The Email panel of the Settings > Notifications window displays information (Figure 7-9).

**Notifications**

- Email
- SNMP
- Syslog

**Email**

The support user receives call home events. Local users also receive event notifications.

[Edit](#) [Disable Notifications](#)

**Email Servers**

IP Address	Server Port
10.10.10.10	25

**Call Home**

Email Address	Event Type	Inventory
flash-sc2@vnet.ibm.com	Alerts	<input checked="" type="checkbox"/>

[Test](#)

**Email Notifications**

Email Address	Event Type	Inventory
monitor@ibm.com	Alerts	<input type="checkbox"/>

[Test](#)

**Email Contact**

* Contact Name	* Email Reply Address	
John Doe	jdoe@ibm.com	
* Machine Location	* Telephone (Primary)	Telephone (Alternate)
third floor, room 304	1234567890	

0 of 5,325 GiB (0%)    0%    0    0    Health Status

Figure 7-9 Notifications are now enabled and active

Mail receivers can be tested by clicking **Edit** and then **Test** for each of the mail receivers. Testing the *call home* receiver generates a support call at IBM and you are contacted to confirm successful configuration of call home.

## SNMP

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that the system sends.

**Note:** You can download the FlashSystem 900 management information base (MIB) file from the FlashSystem 900 GUI at the **Settings** → **System** → **Notifications** → **SNMP** page.

Figure 7-10 shows the SNMP configuration menu.

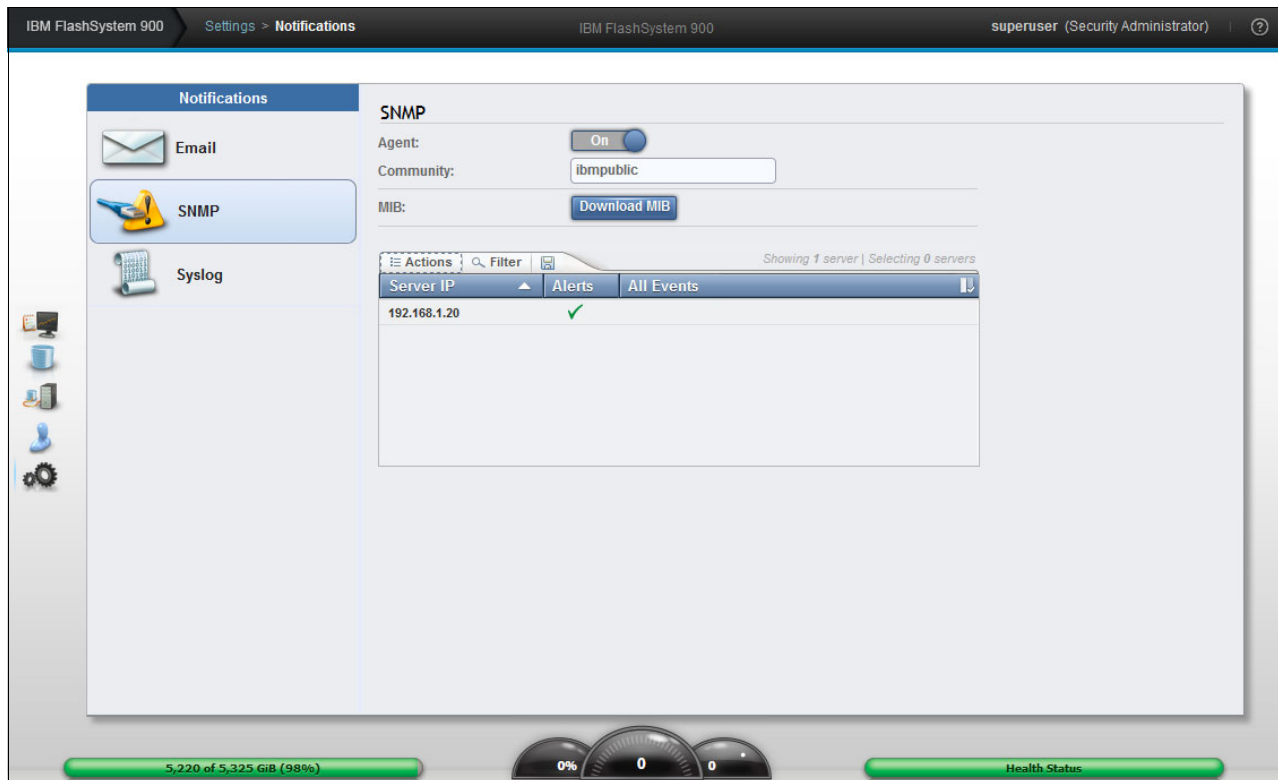


Figure 7-10 Event notifications SNMP window

In the SNMP configuration menu, you can configure one or more SNMP servers. For each of these SNMP servers, you configure the following information:

- ▶ IP address
- ▶ SNMP server port: The default is port 162.
- ▶ SNMP community: The default is public.
- ▶ Event type: The default is Alerts but it can be changed to All events.

Various SNMP trap receiver products are on the market. These are known as *SNMP managers*. IBM Tivoli NetView® or IBM Tivoli Enterprise Console® can be used as IBM SNMP managers.

## Syslog

The *syslog protocol* is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be either IPv4 or IPv6. The system can send syslog messages that notify personnel about an event.

The IBM FlashSystem 900 can transmit syslog messages in either expanded or concise format. You can use a syslog manager to view the syslog messages that the system sends. The system uses the User Datagram Protocol (UDP) to transmit the syslog message. You can specify up to a maximum of six syslog servers (Figure 7-11).

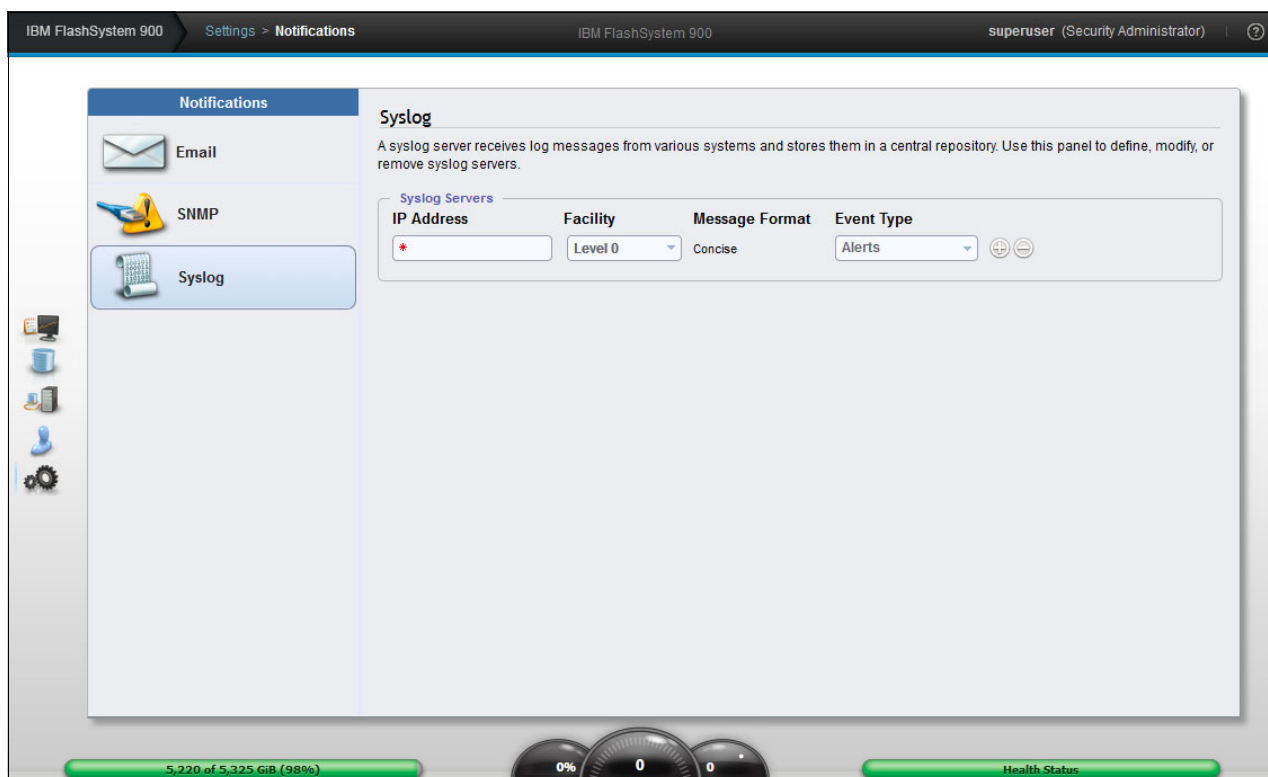


Figure 7-11 Event notifications Syslog window

In the Syslog configuration menu, you can configure one or more syslog servers. For each of these servers, you configure the following information:

- ▶ IP address
- ▶ Facility  
The *facility* determines the format for the syslog messages and can be used to determine the source of the message.
- ▶ Event type (The default is Alerts but it can be changed to All events.)

Various syslog server products are on the market. Many of these products are no-charge products that can be downloaded from the Internet.

### 7.1.3 Security menu

The security menu provides configuration of the following options:

- Remote authentication

When remote authentication is configured, users authenticate with their domain user and password rather than a locally created user ID and password. Remote authentication gives you central access control. If someone leaves the company, you only need to remove access at the domain controller, which means that no orphan user IDs remain on the storage system.

- Encryption

The IBM FlashSystem 900 provides encryption, which protects against the potential exposure of sensitive user data and user metadata that are stored on discarded, lost, or stolen flash modules. Protection of encrypted data is enforced by encryption keys stored on external USB flash drives. A USB flash drive with a valid encryption key must be placed in one of the USB connectors of the FlashSystem 900 during power up, reboot and many repair actions.

#### Configure remote authentication

When a FlashSystem 900 clustered system is created, the authentication settings default to `local`, which means that the IBM FlashSystem 900 contains a local database of users and their privileges. Local users can be created by the superuser account.

You can create two types of users who can access the system. These types are based on how the users authenticate to the system. Local users are authenticated through the authentication methods that are located on the IBM FlashSystem 900.

If the local user needs access to the management GUI, a password is needed for the user. If the user requires access to the command-line interface (CLI) through Secure Shell (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 privileges on the system.

For users of the FlashSystem 900 clustered system, you can configure authentication and authorization using the CLI and the GUI as configured in the Users and User Groups menu.

A *remote user* is authenticated on a remote service with Lightweight Directory Access Protocol (LDAP) as configured in the Settings → Security section of the FlashSystem 900 GUI (Figure 7-1 on page 232). Remote users have their roles defined by the remote authentication service.

Remote authentication is disabled by default and can be enabled to authenticate users against LDAP servers.

A user who needs access to the CLI must be configured as a local user on the IBM FlashSystem 900.

Remote users do not need to be configured locally; they need to be defined only on the LDAP server.

For more information about how to configure remote authentication and authorization for users of the IBM FlashSystem 900, search for the “User authentication configuration” section of the IBM FlashSystem 900 web page of the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## Reasons for using remote authentication

Use remote authentication for the following reasons:

- ▶ Remote authentication saves you from having to configure a local user on every IBM storage system that exists in your storage infrastructure.
- ▶ If you have multiple LDAP-enabled storage systems, remote authentication makes it more efficient to set up authentication.
- ▶ The audit log shows the domain user name of the issuer when commands are executed. The domain user name is more informative than a local user name or just *superuser*.
- ▶ Remote authentication gives you central access control. If someone leaves the company, you only need to remove access at the domain controller, which means that no orphan user IDs remain on the storage system.

## Prepare the LDAP server

The first step in configuring LDAP is to prepare the LDAP server. This example use a Microsoft Windows 2008 R2 Enterprise server, which is promoted to be a Domain Controller by using the **dcpromo** command. Next, the computer role *Active Directory Lightweight Directory Services* is added.

The privileges that the LDAP user gets on the IBM FlashSystem 900 are controlled by user groups on the storage system. There must be matching user groups on the Active Directory (AD) server and on the IBM FlashSystem 900, and the LDAP users must be added to the AD server group.

In this example, you create a group called *FlashAdmin*, which you use to manage your FlashSystem 900 storage device.

1. To create this group, log on to the AD Domain Controller and configure Active Directory. An easy way to configure Active Directory from the AD controller is to select **Start** → **Run**, type **dsa.msc**, and click **OK**. The Active Directory Users and Computers management console opens (Figure 7-12).

Click the **Create a new group in the current container** icon.

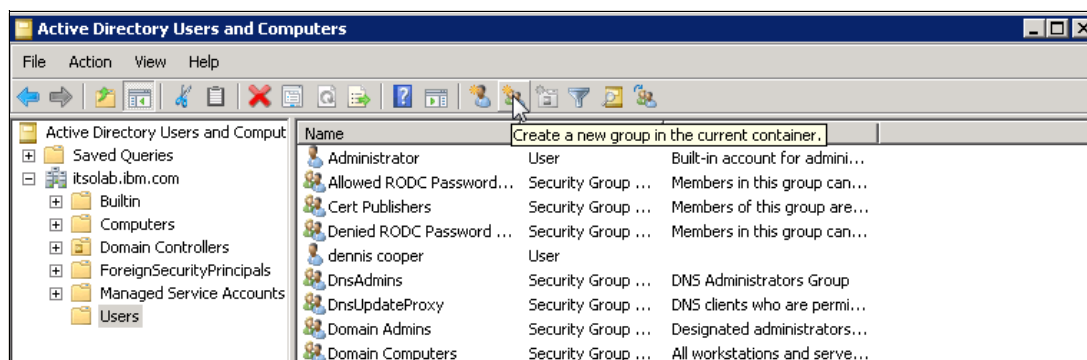


Figure 7-12 Active Directory Users and Computers window to create a new group

2. In the New Object - Group window that opens (Figure 7-13 on page 245), type **FlashAdmin** for the new group name, keep the remaining default values, and click **OK**.



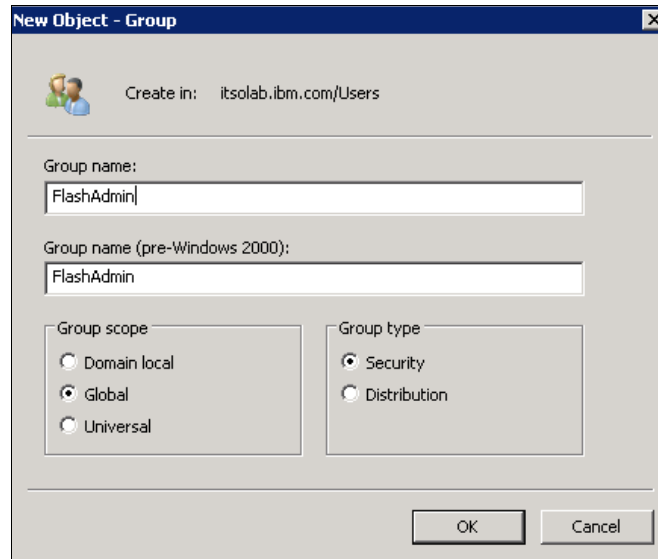


Figure 7-13 Active Directory to create a FlashAdmin group

3. Highlight the users that you want to add to the FlashSystem 900 storage administrator group and click the **Adds the selected objects to a group you specify** icon (Figure 7-14).

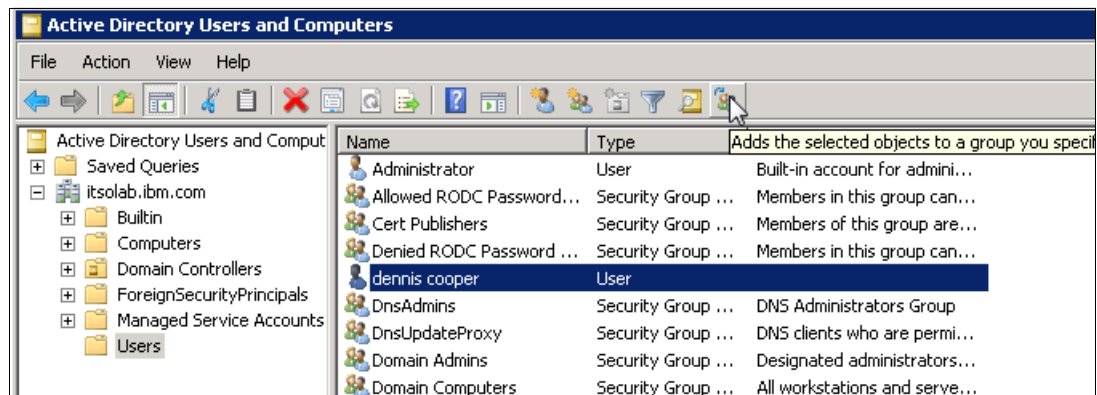


Figure 7-14 Adds the selected objects to a group you specify

4. In the Select Groups window, type FlashAdmin and click **Check Names** (Figure 7-15).

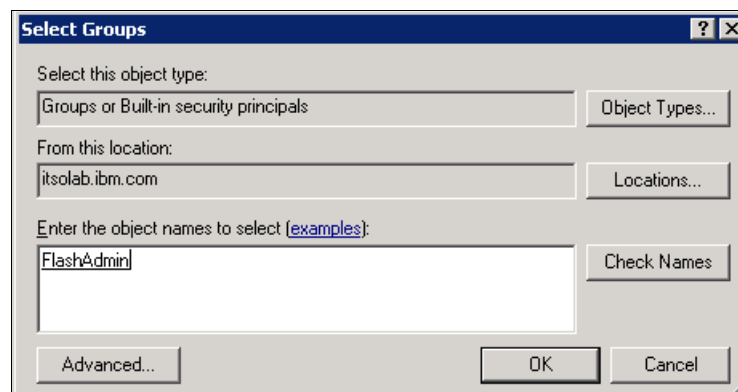


Figure 7-15 Active Directory Select Groups window to add users to the FlashAdmin group

Any other users that might be added to the FlashAdmin group get the same privileges on your FlashSystem 900.

If other users with other privileges are required, another group on the IBM FlashSystem 900 with different privileges is required. A group on the AD server with a matching name is also required.

Your LDAP server is now prepared for remote authentication.

### Enable remote authentication on FlashSystem 900

The next step in configuring remote authentication for the IBM FlashSystem 900 is to specify the authentication server, test connectivity, and test whether users can authenticate to the LDAP server. As shown in Figure 7-1 on page 232, the Settings > Security page has these two options:

- ▶ Remote authentication
- ▶ Encryptions

If you select, **Remote Authentication**, the Remote Authentication panel opens. You can configure user authentication.

The following items are highlighted in Figure 7-16:

1. The default authentication method is local authentication
2. The current authentication method is specified; authentication is enabled.

Click **Configure Remote Authentication**.

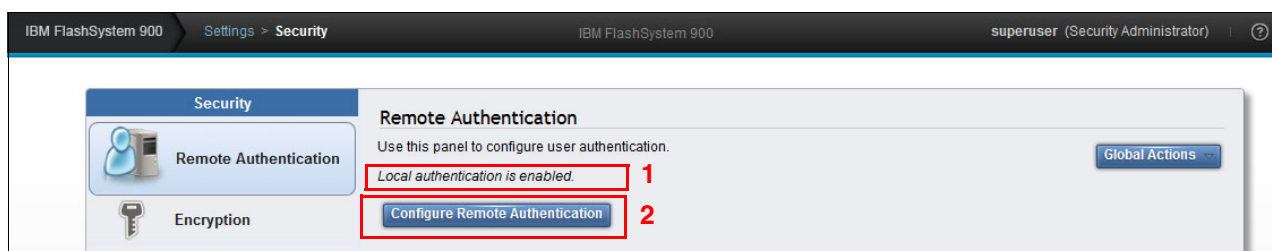


Figure 7-16 Enable remote authentication

The Configure Remote Authentication window opens. Select **LDAP** (Figure 7-17) and click **Next**.

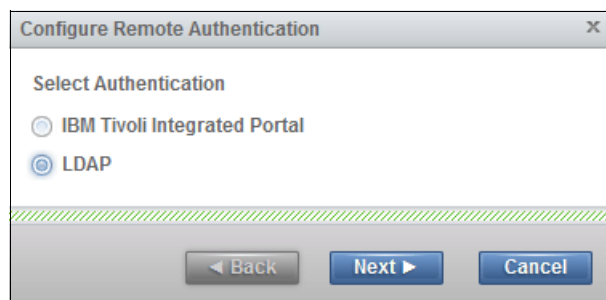


Figure 7-17 Remote Authentication wizard (step 1 of 4)

In the Configure Remote Authentication window (Figure 7-18), select **Microsoft Active Directory**, for Security, select **None**, and click **Advanced Settings** to expand it.

The screenshot shows the 'Configure Remote Authentication' window at step 2 of 4. It has two main sections: 'LDAP Type' and 'Security'. In 'LDAP Type', 'Microsoft Active Directory' is selected with a radio button. In 'Security', 'None' is selected with a radio button. Below these is a button labeled 'Advanced Settings' with a right-pointing arrow. At the bottom are three buttons: 'Back' (with a left arrow), 'Next' (with a right arrow), and 'Cancel'.

Figure 7-18 Remote Authentication wizard (step 2 of 4)

Any user with authority to query the LDAP directory can be used to authenticate. The Active Directory domain is `itsolab.ibm.com`, so use the Administrator login name on the `itsolab.ibm.com` domain to authenticate (Figure 7-19). Then, click **Next**.

The screenshot shows the 'Configure Remote Authentication' window at step 3 of 4. The 'Advanced Settings' button from the previous step is now expanded, revealing two sub-sections. The first, 'Service Credentials (Optional)', contains a 'User Name' field with the text 'itsolabadministrator' and a 'Password' field with masked characters. The second, 'Advanced Settings', contains three fields: 'User Attribute' with 'sAMAccountName', 'Group Attribute' with 'memberOf', and 'Audit Log Attribute' with 'userPrincipalName'. At the bottom are the 'Back', 'Next', and 'Cancel' buttons.

Figure 7-19 Remote Authentication wizard (step 3 of 4)

Type the IP address of the LDAP server, which in this case is `9.xx.xx.xxx`, and the LDAP Group Base Domain Name (DN) for Microsoft Active Directory.

You can obtain the LDAP User and Group Base DN for Microsoft Active Directory by using the following commands:

```
dsquery user -name <username>
dsquery group -name <group name>
```

To look up the Base DN, log on to the LDAP server and run the commands shown in Example 7-1.

*Example 7-1 Checking the LDAP server for the Base DN*

```
C:\Users\Administrator>dsquery group -name FlashAdmin
"CN=FlashAdmin,CN=Users,DC=itsolab,DC=ibm,DC=com"
```

```
C:\Users\Administrator>
```

The Base DN to which you need to enable LDAP authentication requires only the domain part of the output in Example 7-1. In the Base DN (Optional) field of the Configure Remote Authentication window (Figure 7-20), enter the following text:

DC=itsolab,DC=ibm,DC=com

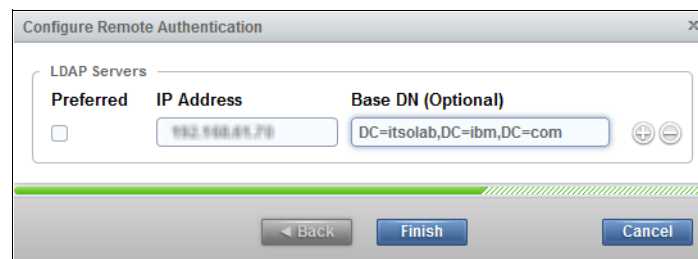


Figure 7-20 Remote Authentication wizard (step 4 of 4)

Click **Finish** to return to the Settings > Security window. Figure 7-21 shows that LDAP is enabled and the window shows the preferences of the configured LDAP server.

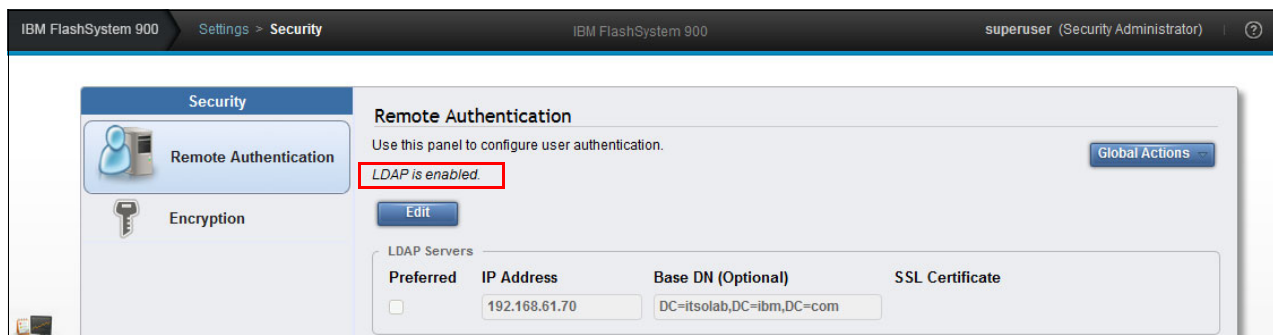


Figure 7-21 Remote Authentication enabled

### Create the FlashSystem 900 LDAP-enabled user group

The first part of the LDAP configuration is complete. You need, however, to create a new user group on your FlashSystem 900 with a name that matches the name that you configured on the LDAP server. For this, configure the name FlashAdmin on the LDAP server.

1. Click **Access** → **User Groups** (Figure 7-22).

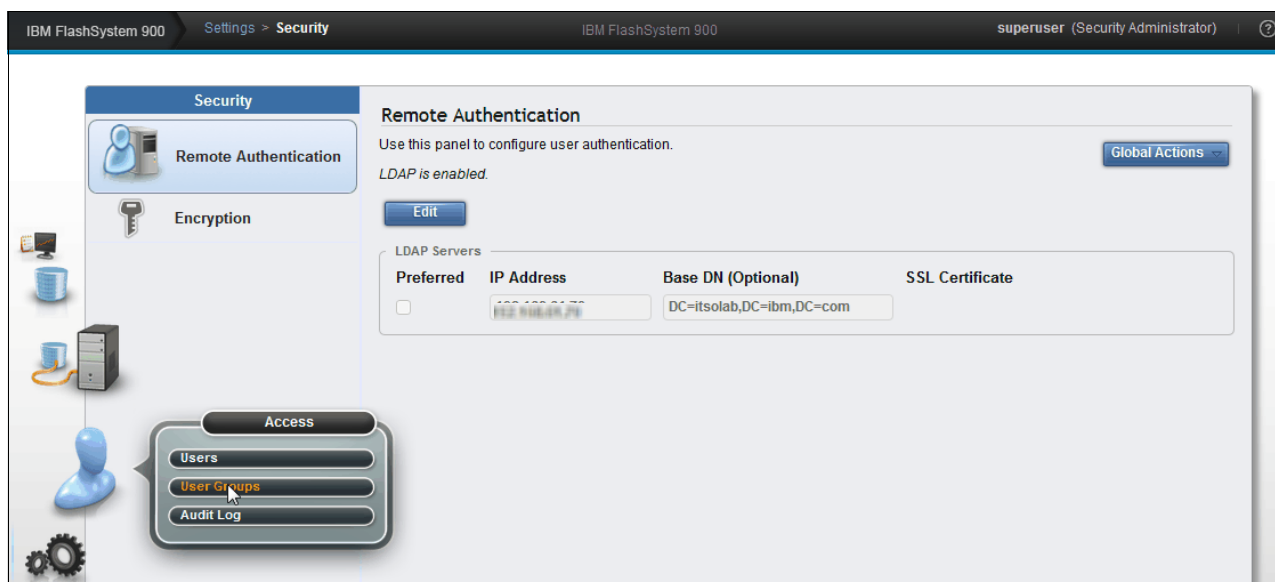


Figure 7-22 Navigate to User Groups

2. Figure 7-23 shows the current configured user groups. Click **Create User Group**.

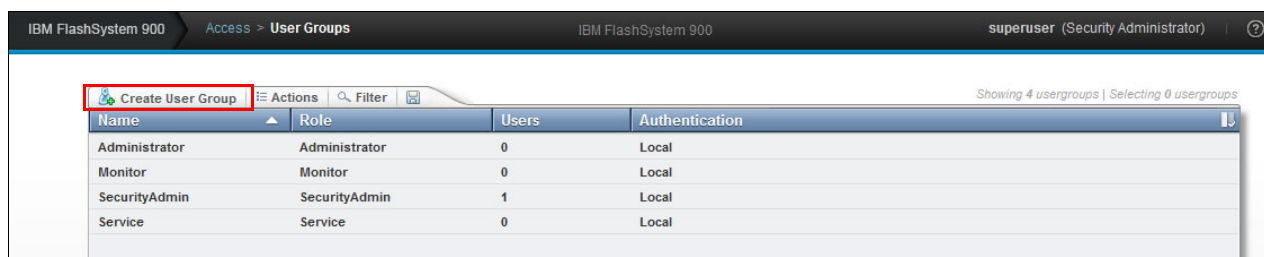
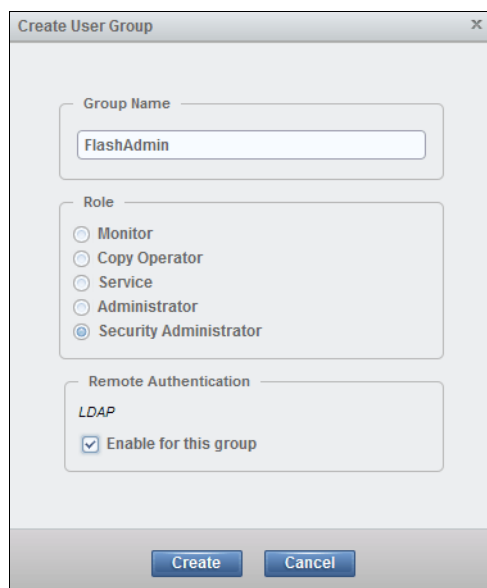


Figure 7-23 Create a new user group

3. Select **Security Administrator** and select **Enable for this group** for the LDAP (Figure 7-24). Type the group name FlashAdmin for the new user group.

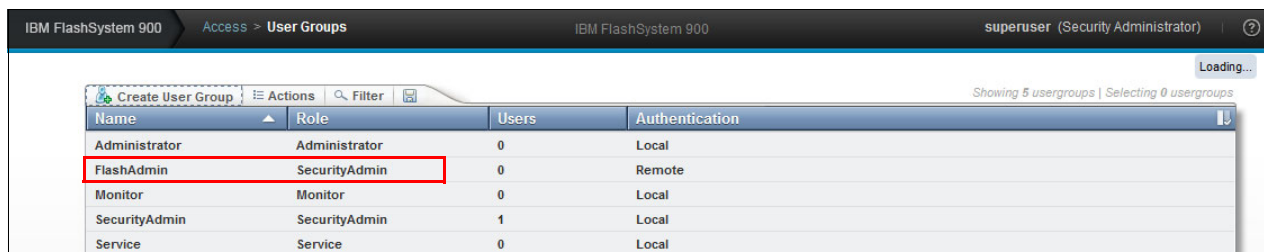


The 'Create User Group' dialog box is shown. It has a title bar with a close button. Inside, there are three sections: 'Group Name' with a text field containing 'FlashAdmin'; 'Role' with five radio buttons: 'Monitor', 'Copy Operator', 'Service', 'Administrator', and 'Security Administrator' (which is selected); and 'Remote Authentication' with a sub-section 'LDAP' containing a checked checkbox 'Enable for this group'. At the bottom are 'Create' and 'Cancel' buttons.

Figure 7-24 Select Security Administrator

**Note:** If the field Remote Authentication is not visible in the Create User Groups window, remote authentication is disabled in **Settings** → **Security**.

The new user group is created and enabled for remote authentication (Figure 7-25).



The screenshot shows the 'User Groups' section of the IBM FlashSystem 900 GUI. A table lists the user groups. The 'FlashAdmin' group is highlighted with a red box. The table has columns for Name, Role, Users, and Authentication.

Name	Role	Users	Authentication
Administrator	Administrator	0	Local
FlashAdmin	SecurityAdmin	0	Remote
Monitor	Monitor	0	Local
SecurityAdmin	SecurityAdmin	1	Local
Service	Service	0	Local

Figure 7-25 Group FlashAdmin created

### Testing LDAP authentication

At this point, you can log out the user superuser and try to log in with the LDAP user. However, before you do that, the Remote Authentication window provides a capability to test LDAP.

1. From the FlashSystem 900 GUI, select **Settings** → **Security** → **Remote Authentication** → **Configure Remote Authentication** → **LDAP**.
2. Click **Global Actions** and select **Test LDAP Connections** (Figure 7-26 on page 251).

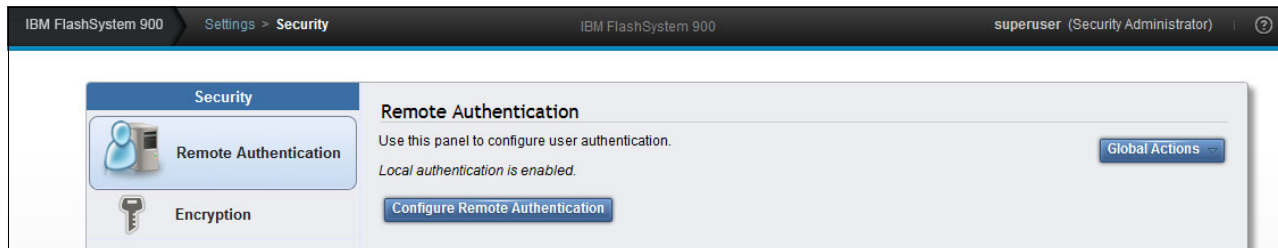


Figure 7-26 Remote Authentication: Test LDAP Connections option

The Test LDAP Connections task window opens and displays the CLI command used to test the connection. In a successful connection to the LDAP server, the output is shown (Figure 7-27).

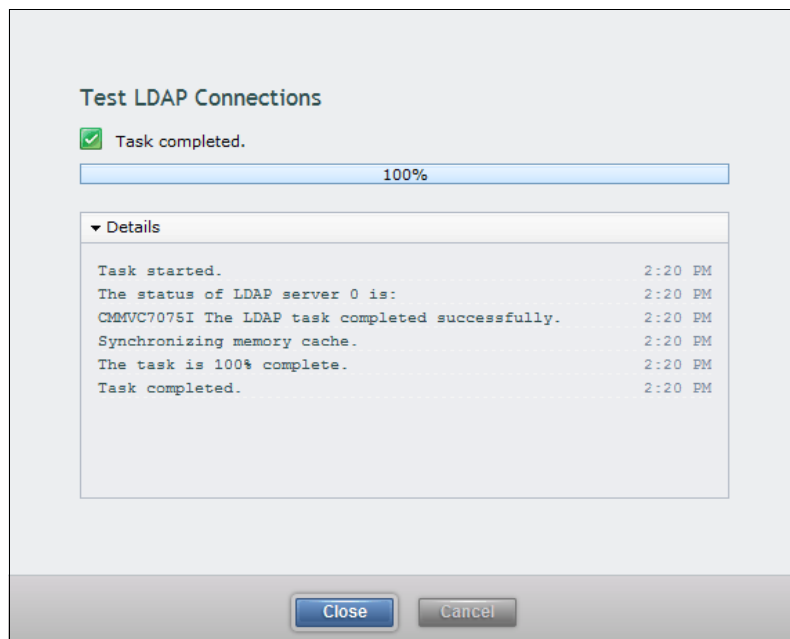


Figure 7-27 Remote Authentication: Test LDAP connections CLI result

From the Global Actions menu, you also can test whether the authentication for a specific user is functional. Click **Test LDAP Authentication**. The window shown in Figure 7-28 opens. Type the credentials of the LDAP user for whom you want to test authentication and click **Test**.

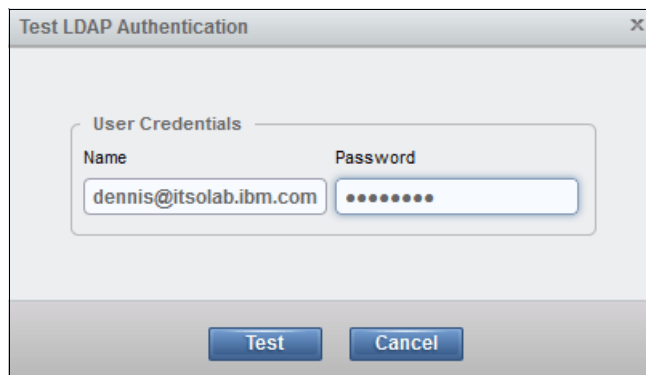


Figure 7-28 Remote Authentication: Test LDAP Authentication

The CLI command window opens. If the authentication is successful, you see the same output as shown in Figure 7-27 on page 251.

If the test is unsuccessful, you see the message shown in Figure 7-29.

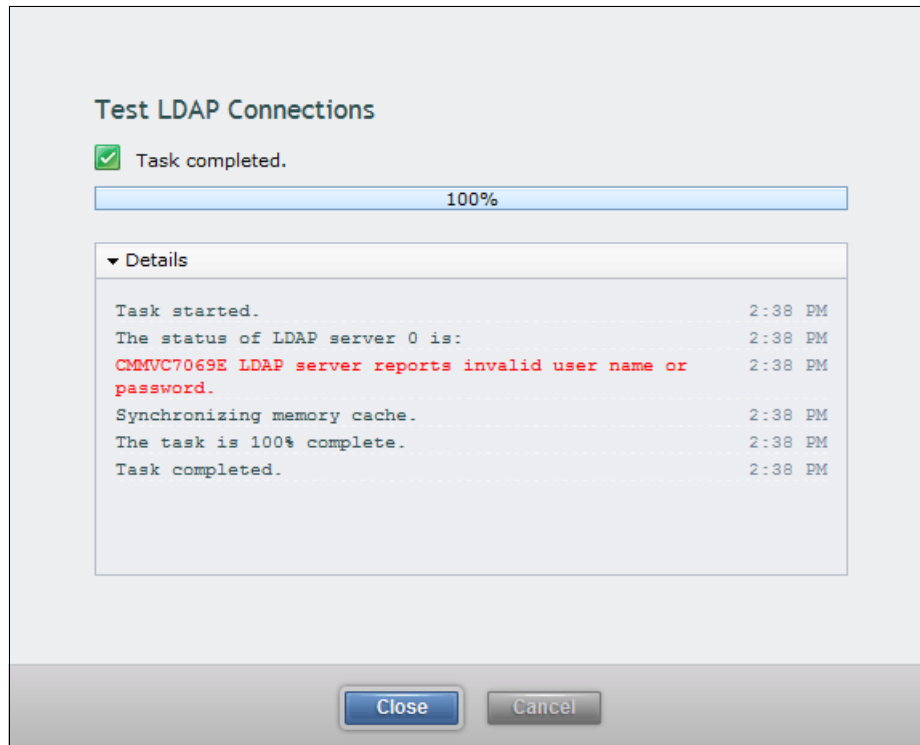


Figure 7-29 Remote Authentication: Test unsuccessful

### Log in as an LDAP user

Assuming that remote authentication is successful, the superuser user can now log out and the LDAP user can log in as shown in Figure 7-30.

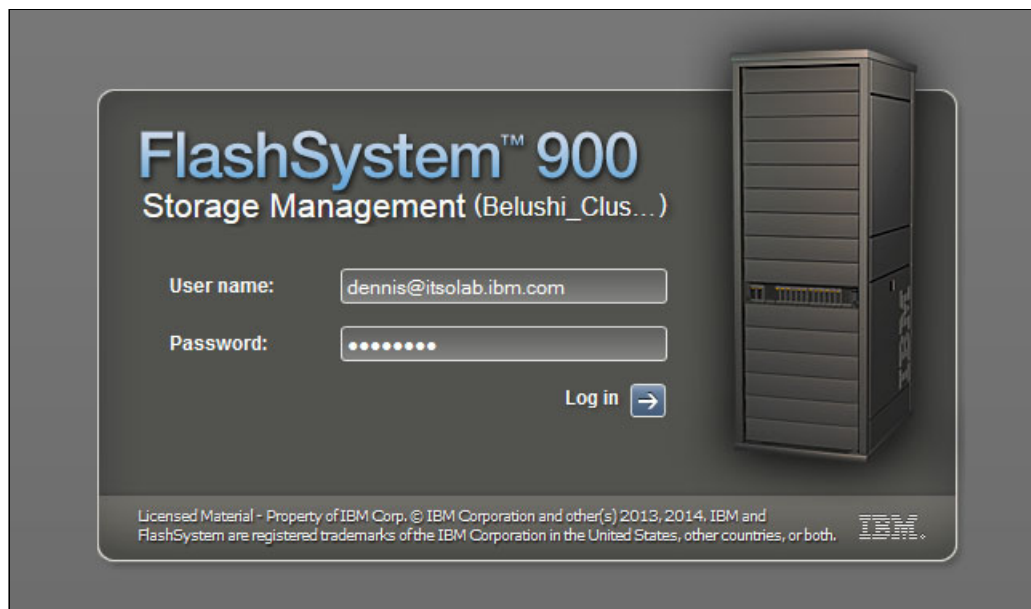


Figure 7-30 Login window for the LDAP user



Figure 7-31 shows the FlashSystem 900 home window. The upper-right corner indicates which user is currently logged in.

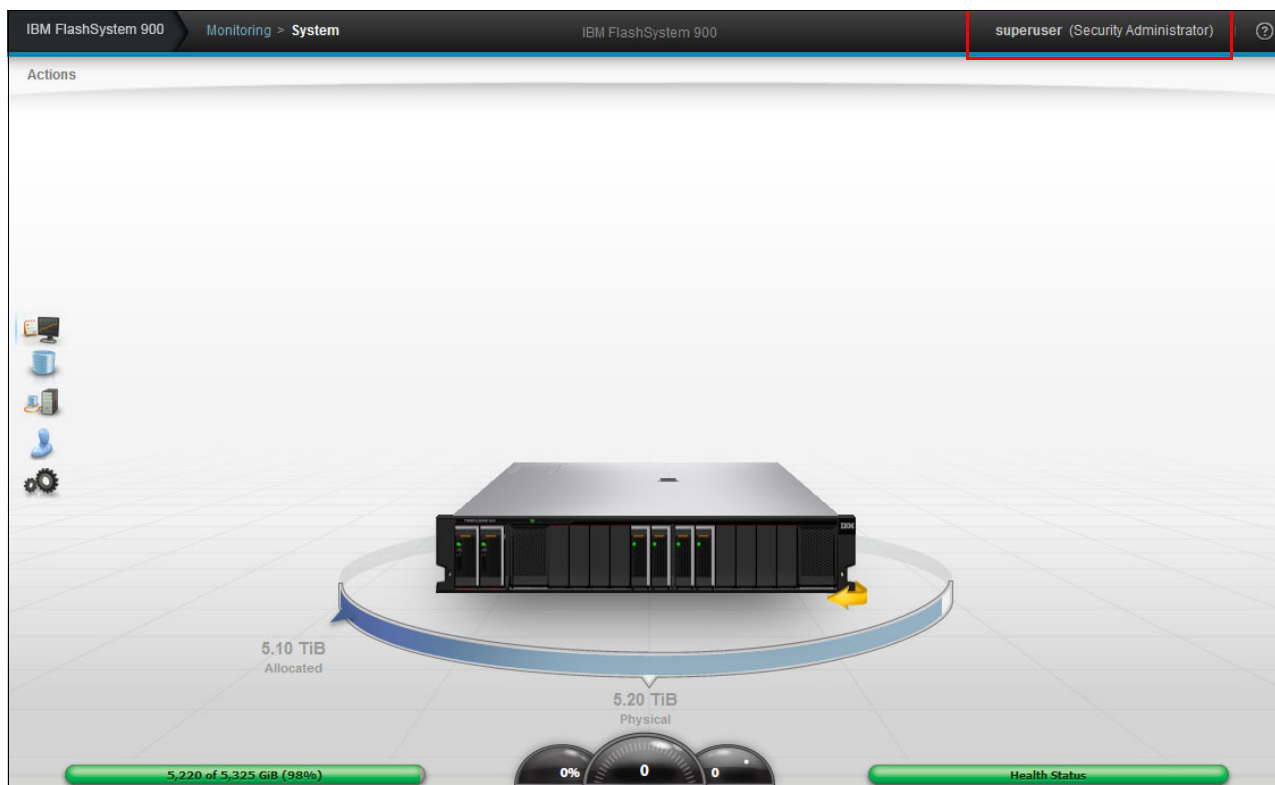


Figure 7-31 Main window LDAP user logged in

Configuring remote authentication is complete.

## About encryption

The IBM FlashSystem 900 provides optional encryption of data at rest, which protects against the potential exposure of sensitive user data and user metadata that are stored on discarded, lost, or stolen flash modules. Encryption of system data and system metadata is not required, so system data and metadata are not encrypted.

### ***AES-XTS 256 bit data-at-rest encryption with local key management***

Two functions are added to the encryption feature:

- ▶ Hot Encryption Activation: Adding an encryption license to a previously initialized system
- ▶ Encryption Rekey: Changing the encryption key on a previously initialized system

If you want to use encryption, ensure that you purchased feature code (FC) AF14: Encryption Enablement Pack (Plant).

### ***Data Encryption Methodology***

The IBM FlashSystem 900 data encryption uses the Advanced Encryption Standard (AES) algorithm, with a 256-bit symmetric encryption key in XTS mode. This encryption mode is known as XTS–AES–256, which is described in the IEEE 1619–2007 data encryption standard. The data encryption key itself is protected by a 256-bit AES key wrap when it is stored in non-volatile form. Two layers of encryption are used with stored data:

- ▶ The data being protected
- ▶ The data encryption key itself

## ***Protection Enablement Process***

The Protection Enablement Process (PEP) transforms a system from a state that is not protection-enabled to a state that is protection-enabled.

The PEP establishes a secret *encryption access key* to access the system, which must be stored and made available for use at a later time, whenever the system needs to be unlocked. The secret *encryption access key* must be stored outside the system on a USB drive, which the system reads to obtain the key. The encryption access key should also be backed up to other forms of storage.

Encryption capability of FlashSystem 900 consists of two functions:

- ▶ Hot Encryption Activation  
Allows an unencrypted FlashSystem to be encryption-enabled while the system is running, without impacting customer data in any way.
- ▶ Non-Disruptive Rekey  
Permits creating a new encryption access key that supersedes the existing key on a running FlashSystem without impacting customer data.

Handling encryption and encryption keys has several aspects:

- ▶ Activating encryption using GUI
- ▶ Activating encryption using CLI
- ▶ Creating new encryption keys (Rekey)
- ▶ Keeping encryption keys from more systems on the same USB flash drives (stacking)
- ▶ Making copies of encryption keys
- ▶ Storing copies of USB flash drives holding encryption keys
- ▶ Leaving encryption keys in or out of the system during normal operation

## ***About USB flash drives and encryption keys***

The FlashSystem 900 predecessor FlashSystem 840 offered the possibility to initialize a new system and enable encryption during the initialization process using InitTool. With FlashSystem 900 and InitTool, enabling encryption during the initialization process is no longer an option. Encryption is enabled after initialization using either the GUI or CLI.

However the same rules for handling encryption keys on USB flash drives are the same.

When the encryption Feature Code AF14 is purchased, IBM sends a total of three USB flash drives: one USB flash drive for the system and two additional USB flash drives for the encryption feature code.

When encryption is activated, an encryption key is generated by the system to be used for access to encrypted data that is stored on the system. The GUI launches a wizard that guides you through the process of copying the encryption key to multiple USB flash drives. The following actions are considered preferred practices for copying and storing encryption keys:

1. Make copies of the encryption key on at least three USB flash drives to access the system.
2. In addition, copy the encryption keys to other forms of storage to provide resiliency and to mitigate risk, if, for example, the three USB flash drives are from a faulty batch of drives.
3. Test each copy of the encryption key before writing any user data to the initialized system.
4. Securely store all copies of the encryption key. As an example, any USB flash drives that do not remain inserted into the system can be locked in a safe. Use comparable precautions to securely protect any other copies of the encryption key stored to other forms of storage.

### Enable the encryption license using GUI

To install the encryption license on a previously initialized system, complete these steps:

1. In the GUI, click **Settings** → **System** → **Licensed Functions** (Figure 7-32).

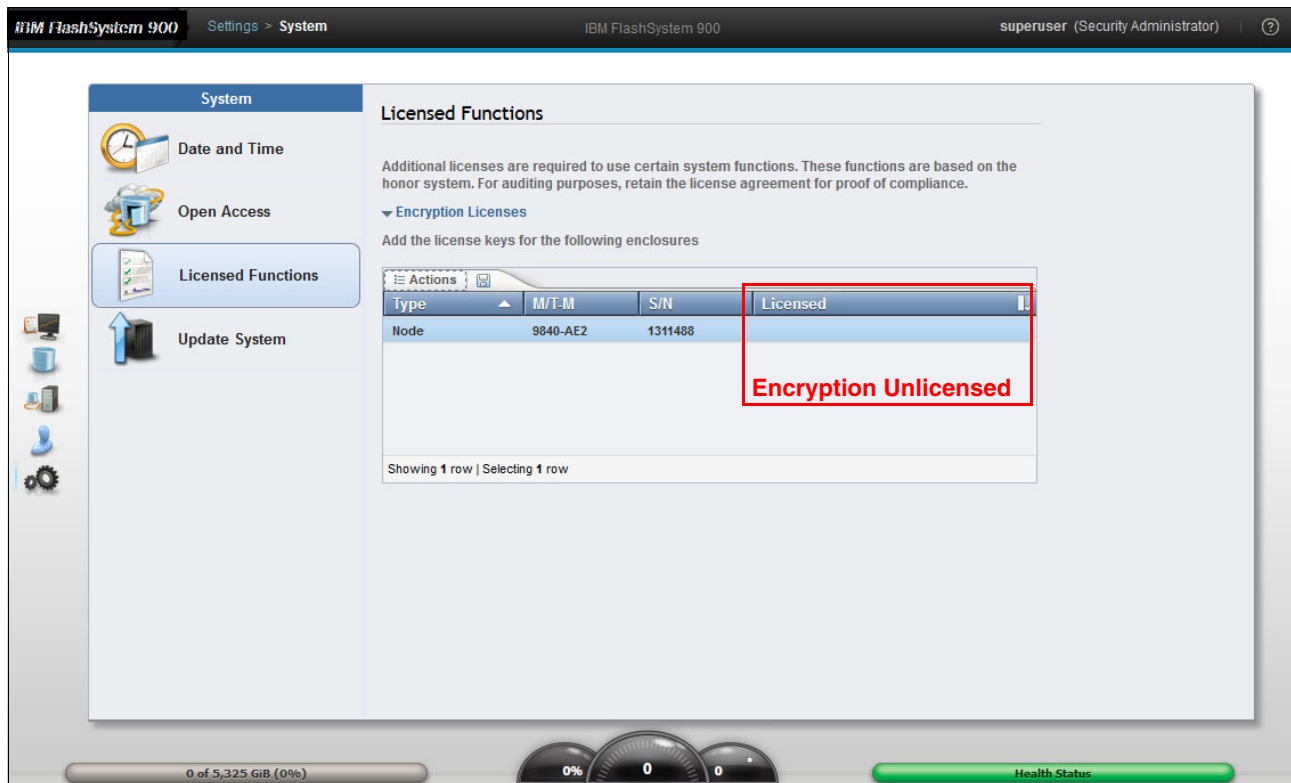


Figure 7-32 Encryption license not enabled

The encryption license is currently not enabled which shows a blank because **Licensed** is not selected. The Licensed column is not a modifiable column. To enable the encryption license, follow the next steps (shown in Figure 7-33, Figure 7-34 on page 256, and Figure 7-35 on page 256). After those steps are complete, a green check mark is indicated under the Licenses column.

2. Click **Actions** → **Modify License**. In the Modify Licenses window (Figure 7-33), select **I have purchased an encryption license** and click **Modify**.

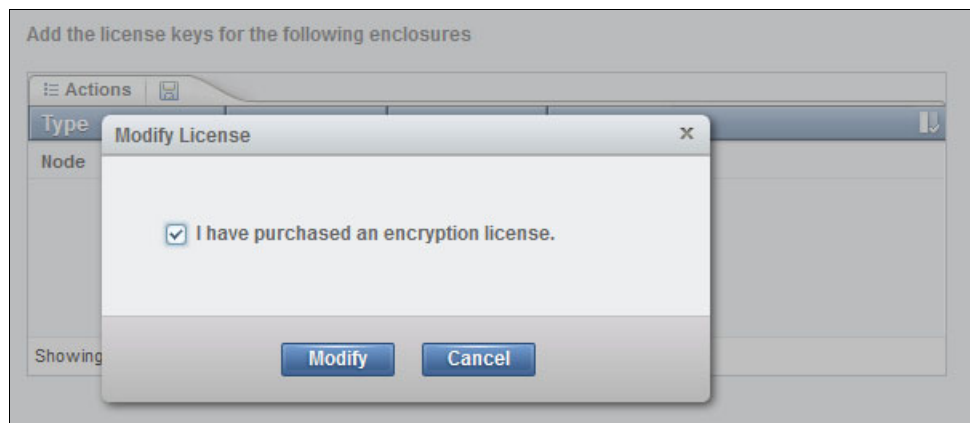


Figure 7-33 Modify license

When you return to the Licensed Functions window, the encryption license is now installed and ready for use (Figure 7-34).

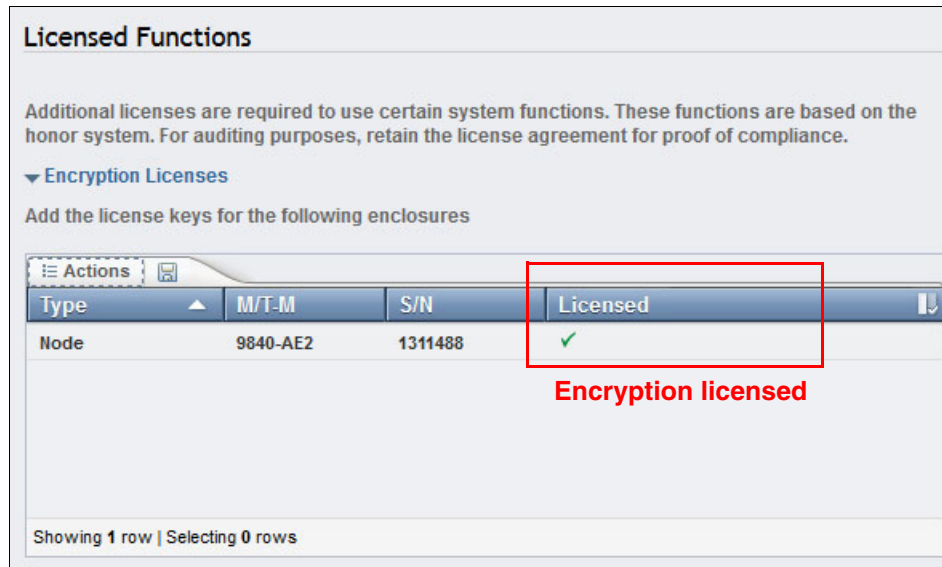


Figure 7-34 Encryption license is now enabled.

### Enable encryption using GUI

Follow these steps to enable encryption using the GUI. For handling encryption using CLI, work with IBM Support for accurate handling procedures.

1. To enable encryption, click **Settings** → **Security** → **Encryption** and click **Enable Encryption** (Figure 7-35).

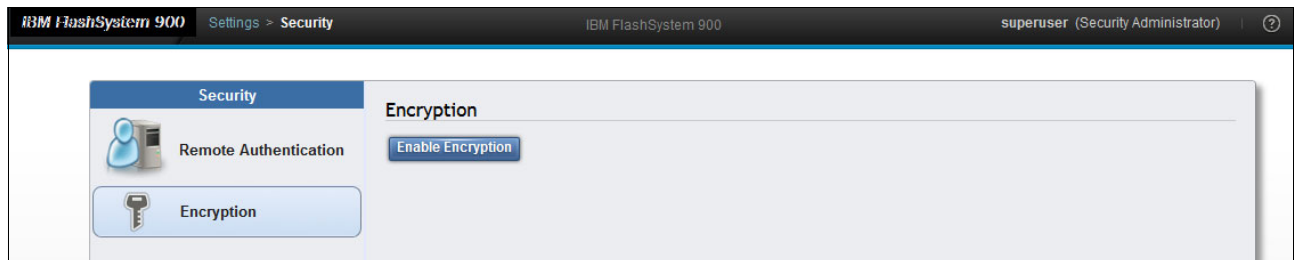


Figure 7-35 Enable encryption

The Enable Encryption wizard opens to the Welcome page (Figure 7-36). The next step requires two USB flash drives.

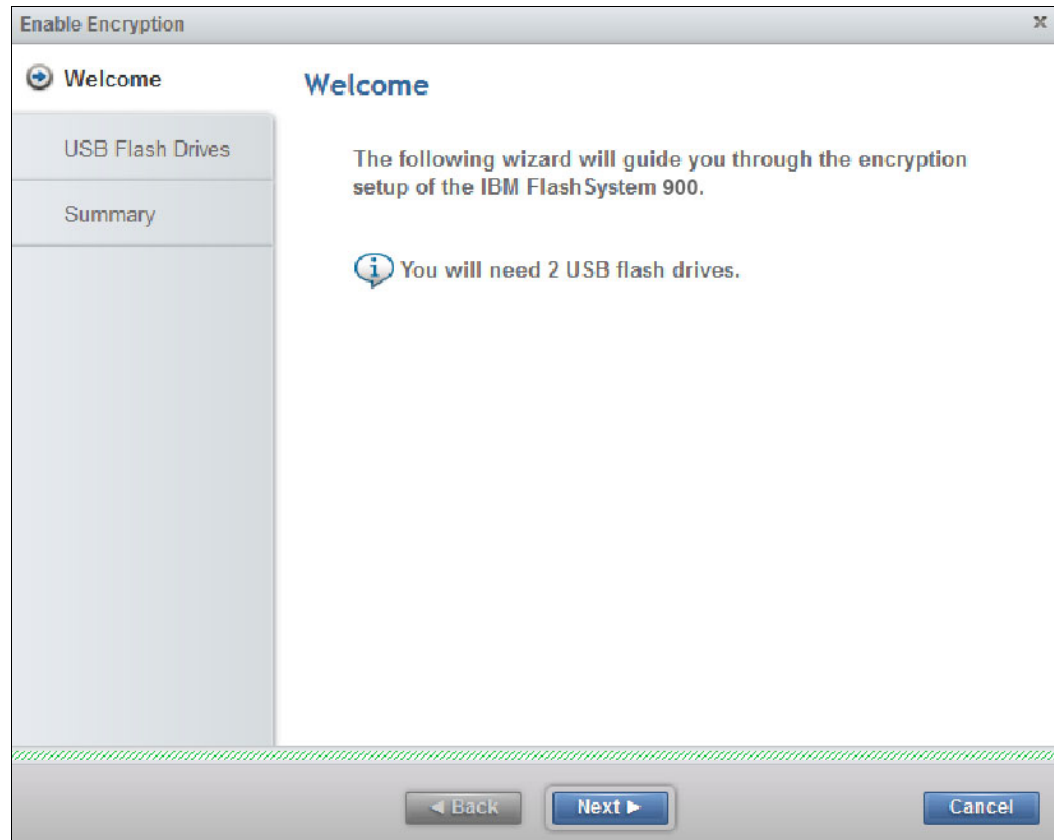


Figure 7-36 Enable Encryption wizard Welcome page

2. Insert two USB flash drives into the FlashSystem 900 controller USB ports and click **Next**. The encryption keys are now generated and stored on both USB keys, and detected (Figure 7-37).

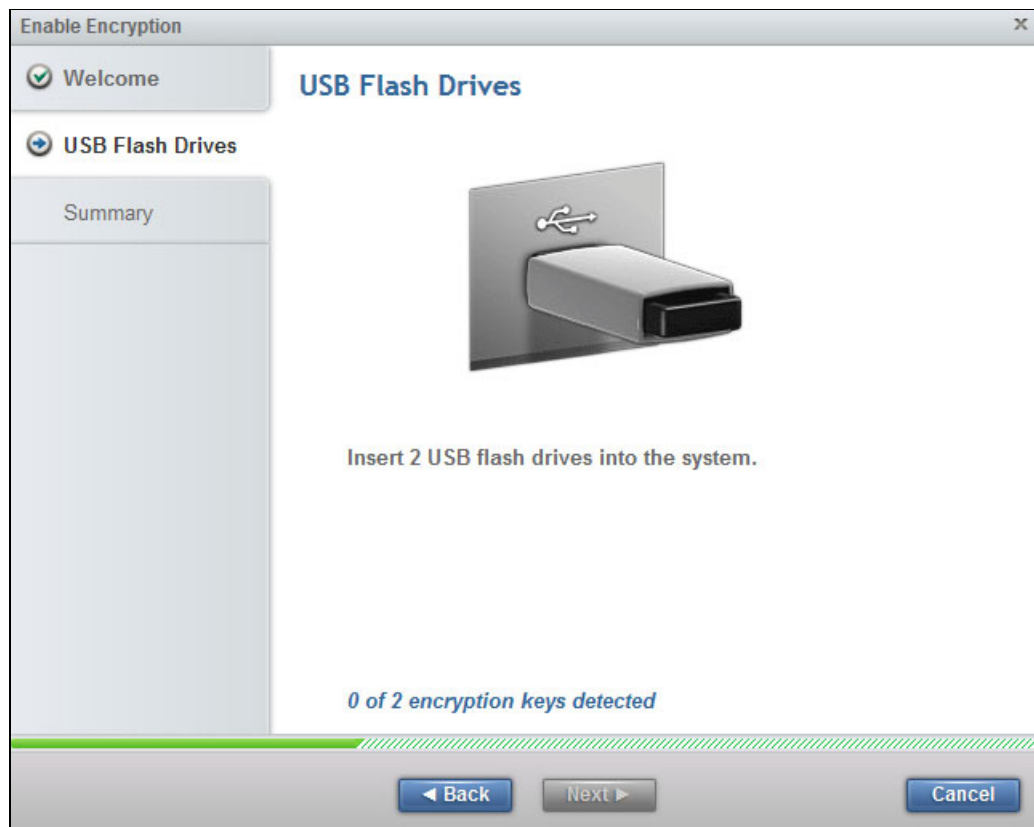


Figure 7-37 Encryption keys writes to USB flash drives

If the system does not recognize USB keys, an error message is issued and the Enable Encryption process cannot continue until USB drives are inserted into the USB ports.

When the process finishes, a message indicates that encryption is enabled (Figure 7-38).



Figure 7-38 Encryption is now enabled

Figure 7-39 shows **Settings** → **Security** → **Encryption** after encryption is enabled. The system reports that USB flash drives are inserted into the USB ports of both nodes and that encryption keys are validated on these drives.

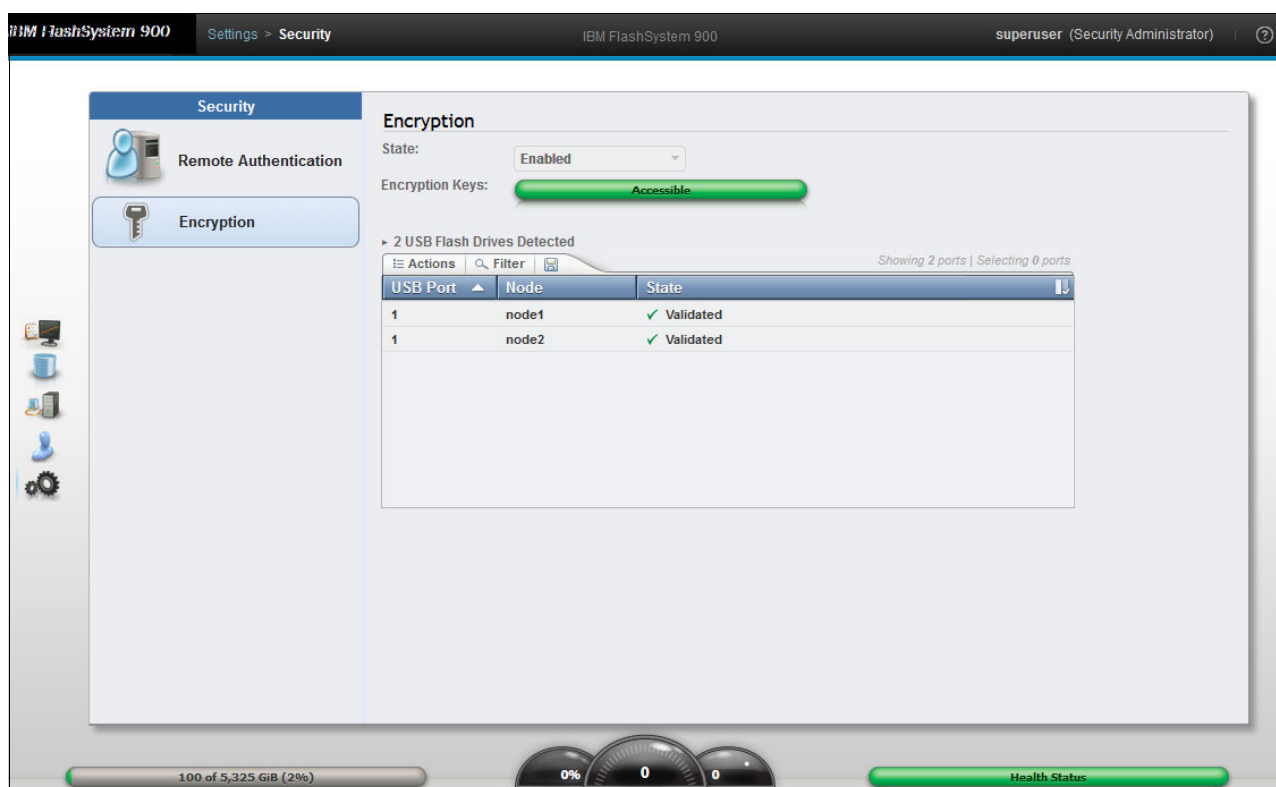


Figure 7-39 Encryption enabled and keys validated

### Working with an encrypted FlashSystem 900

At system start-up (power on) or to access data on an encrypted system, the encryption key must be provided by an outside source so that the system. The encryption key is read from the USB flash drives that store copies of the keys that were created during system initialization. If you want the system to reboot automatically, a USB flash drive with the encryption keys must remain inserted in each canister, so that both canisters have access to the encryption key when they power on.

This method requires that the physical environment where the system is located is secure, so that no unauthorized person can make copies of encryption keys on the USB flash drives and gain access to data stored on the system. For the most secure operation, do not keep the USB flash drives inserted into the canisters on the system. However, this method requires that you manually insert the USB flash drives that contain copies of the encryption key in both canisters before rebooting the system.

The encryption key is required to access encrypted data, and it resides only on the USB flash drive copies and on any additional copies made on other forms of storage. The encryption key cannot be recovered or regenerated by IBM if all user-maintained copies are lost or unrecoverable.

**Attention:** Encryption keys or data from FlashSystem 900 cannot be recovered or regenerated by IBM on an encryption-enabled system if encryption keys are lost.

### Handling encryption using CLI

In addition to using FlashSystem GUI to enable encryption and handle encryption keys, the FlashSystem CLI can be used to enable encryption while the FlashSystem 900 is running. This is a non-destructive procedure when handled correctly.

**Note:** Handling encryption using CLI involves numerous risks. To protect and preserve data, contact IBM Support for assistance with this procedure.

## 7.1.4 Network menu

The Network menu is used for the configuration of the network setup for all the interfaces in the cluster.

Click **Settings** → **Network** to open the Network menu. You can update the network configuration, configure Service IP addresses, and view information about the Fibre Channel (FC) connections.

### Management IP addresses

The *Management IP address* is the IP address of the FlashSystem 900 management interface. This interface includes the GUI and the CLI. The GUI is accessed through a web browser and the CLI is accessed through SSH using PuTTY or a similar tool. The Management IP address is a clustered IP address, which means that if any of the canisters are offline for maintenance or for any other reason, the Management IP address is still available on the surviving node.

The configured Management IP address can be reviewed or changed by selecting **Settings** → **Network** → **Management IP Address**. The Management IP Address page opens (Figure 7-40).

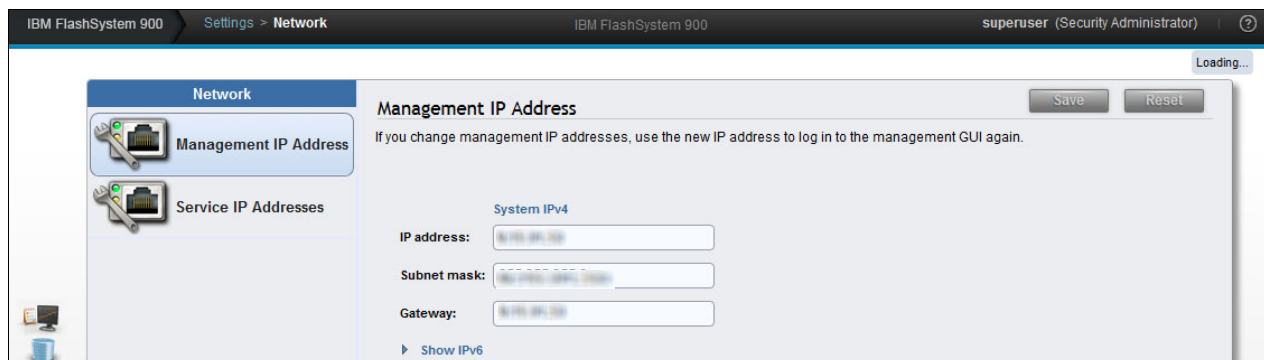


Figure 7-40 Network menu: Set management IP address

### Service IP addresses

The Service IP addresses are the IP addresses of each of the two FlashSystem 900 controllers, which are called *canisters*. These canisters have their own IP addresses where several support actions can be performed, as in these examples:

- ▶ Review installed hardware.
- ▶ Place canister in the Service state.
- ▶ Power cycle canister.
- ▶ Identify canister.
- ▶ Clear all system configuration data.
- ▶ Create new cluster.



- ▶ Recover a failed system (this action is performed *only by IBM Support*).
- ▶ Update firmware manually with the controllers offline.
- ▶ Extract system event logs.

**Note:** The Service IP addresses are normally not used by the IBM FlashSystem 900 administrator. They are used *only* in troubleshooting and scheduled maintenance or when IBM Support performs certain service actions.

To configure the FlashSystem 900 Service IP addresses, click **Settings** → **Network** → **Service IP Address** (Figure 7-41).



Figure 7-41 Configuring Service IP addresses

The default for setting Service IP addresses is to configure an IPv4 address for both FlashSystem 900 nodes. However, IPv6 addresses can also be set by clicking **Show IPv6** and then entering the IPv6 addresses followed by clicking **Save** (This action is not shown in Figure 7-41).

For more information about how to access and use Service Assistant Tool, see 7.2, “Service Assistant Tool” on page 276.

### 7.1.5 Support menu

Click **Settings** → **Support** when log files are requested by IBM Support. IBM Support often requests log files when a support case is opened by the FlashSystem 900 administrators or by the call home function.

The system administrator downloads the requested support package from the system and then uploads it to IBM Support. IBM Support then analyzes the data.

#### Download support package

To download a support package, click **Settings** → **Support** and click **Download Support Package** (Figure 7-42 on page 262).

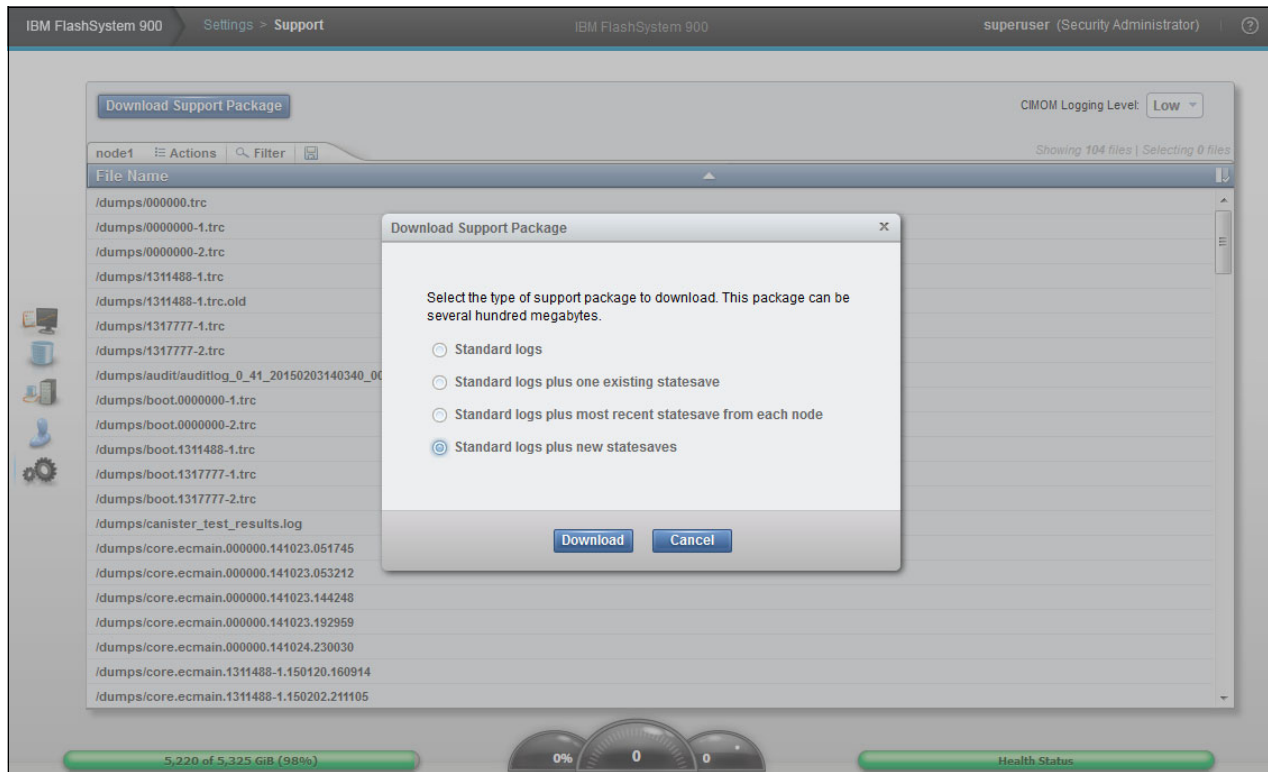


Figure 7-42 Download Support Package

IBM Support usually requests that you select **Standard logs plus new statesaves**. These logs can be downloaded within minutes to hours from the IBM FlashSystem 900, depending on the situation and the size of the support package that is downloaded.

The destination of the support package file is the system where the web browser was launched. Figure 7-43 shows the next step, where you save the support package file on a Windows system.

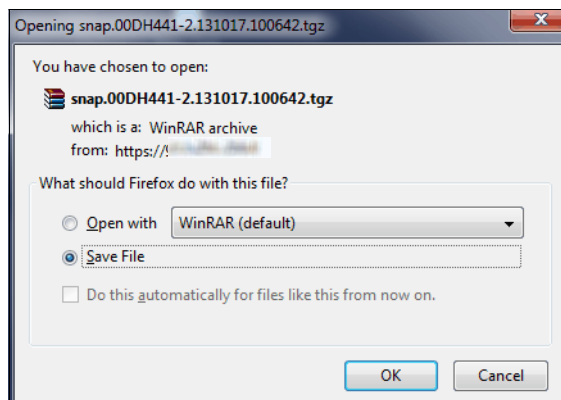


Figure 7-43 Download support package save file

IBM Support usually requests log files to be uploaded to a specific problem management record (PMR) number using Enhanced Customer Data Repository (ECuRep) as the upload media to IBM. See the ECuRep Standard Upload web page:

<http://www.ecurep.ibm.com/app/upload>

## Download individual log files

After analyzing the uploaded support package, IBM Support might request additional files. To locate these files, click **Settings** → **Support** and select **Show full log listing**. This option allows the download of specific and individual log files. An example is shown in Figure 7-44.

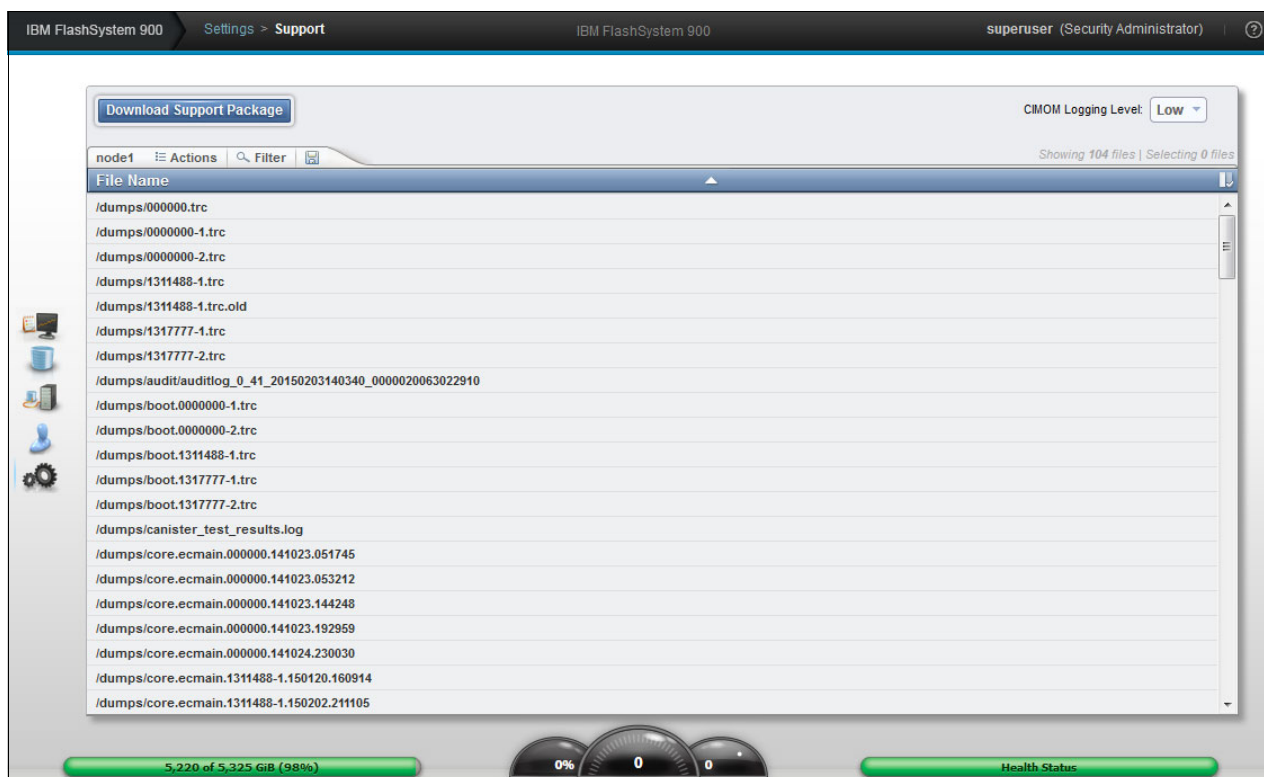


Figure 7-44 Download support: Individual files

You can download any of the various log files by selecting a single item and clicking **Actions** → **Download** (Figure 7-45).

You can also delete a single item in by selecting a single item and clicking **Actions** → **Delete**.

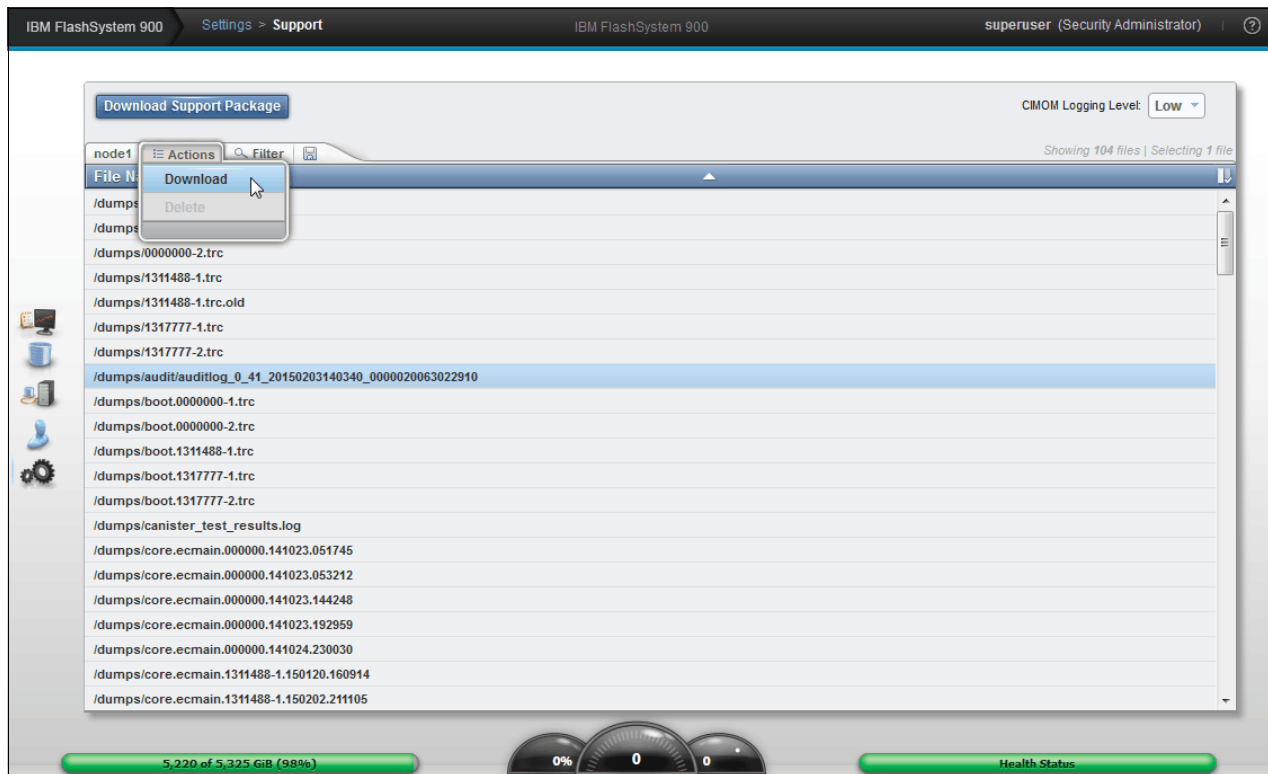


Figure 7-45 Download and Delete options of the Actions menu

**Delete option:** When the Delete option is not available, the file cannot be deleted because it is used by the system.

Log files are saved from each of the installed canisters, which are called *logical nodes*. In the upper left of the window, click the *node* tab to show the node1 or node2 log files (Figure 7-46).

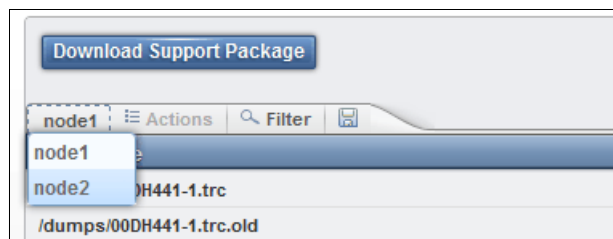


Figure 7-46 Change the log listing of the nodes canister

## 7.1.6 System menu

In the System menu, you can set the time and date for the cluster, enable Open Access, perform software updates for the cluster, and change the preferences for the GUI.

### Date and Time option

Click the **Settings** → **System** option to open the Date and Time page (Figure 7-47), which has options to set the date and time.

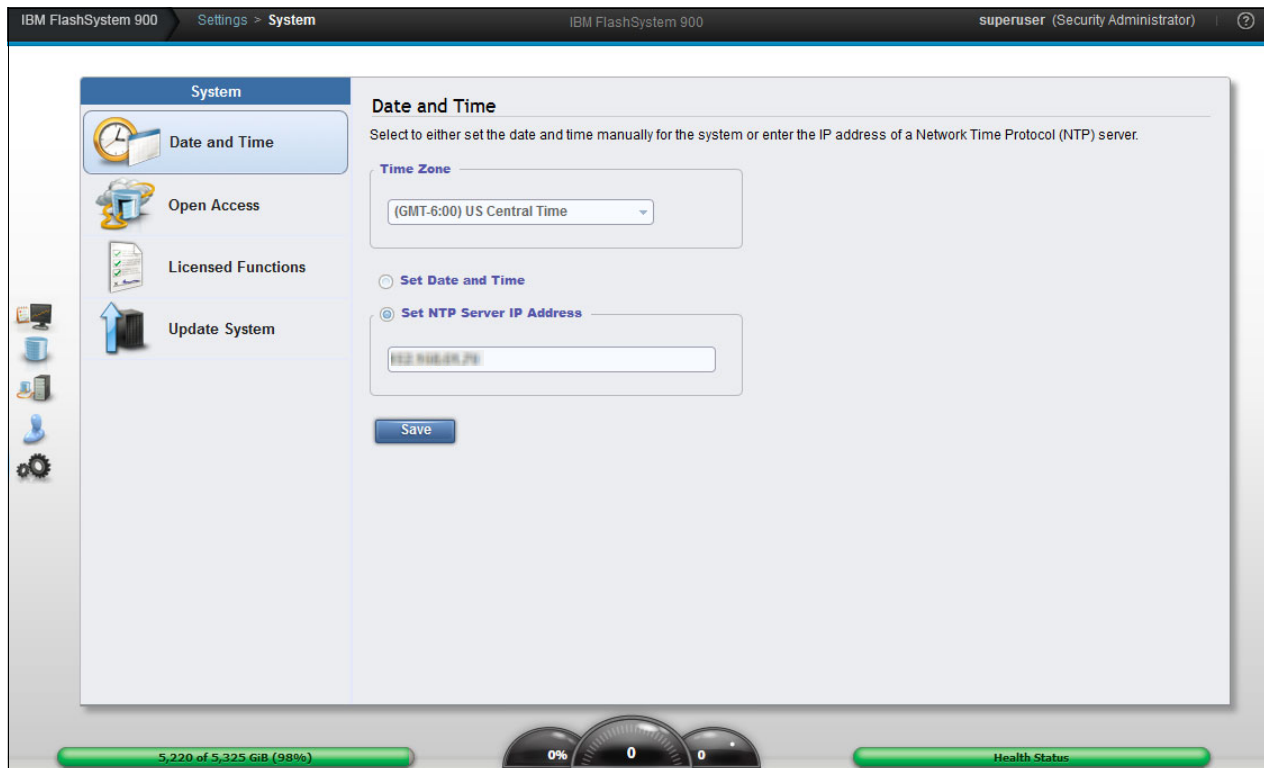


Figure 7-47 Date and time preferences

The preferred method for setting the date and time is to configure a Network Time Protocol (NTP) server. By using an NTP server, all log entries are stamped with an accurate date and time, which is important in troubleshooting. An example might be a temporarily broken Fibre Channel link that caused a path failover at a connected host. To investigate the root cause of this event, logs from the host, logs from the storage area network (SAN) switches, and logs from the IBM FlashSystem 900 must be compared. If the date and time are not accurate, events cannot be compared and matched, which makes a root cause analysis much more difficult.

## Open Access

The IBM FlashSystem 900 can be configured to allow Open Access or to disallow Open Access. Open Access is feasible when the system is directly connected to a host, because then, no other hosts can connect to the system and accidentally read from or write to volumes that belong to other hosts.

Allowing Open Access can also be used in cases where the FlashSystem 900 is connected to correctly zoned SAN switches. However, disallowing Open Access and forcing the system to map its volumes only to selected hosts provides an extra layer of security.

shows the result of Select **Settings** → **System** → **Open Access**. The Open Access page is displayed (Figure 7-48).

**Note:** Open Access can be enabled only when no host mappings are present.

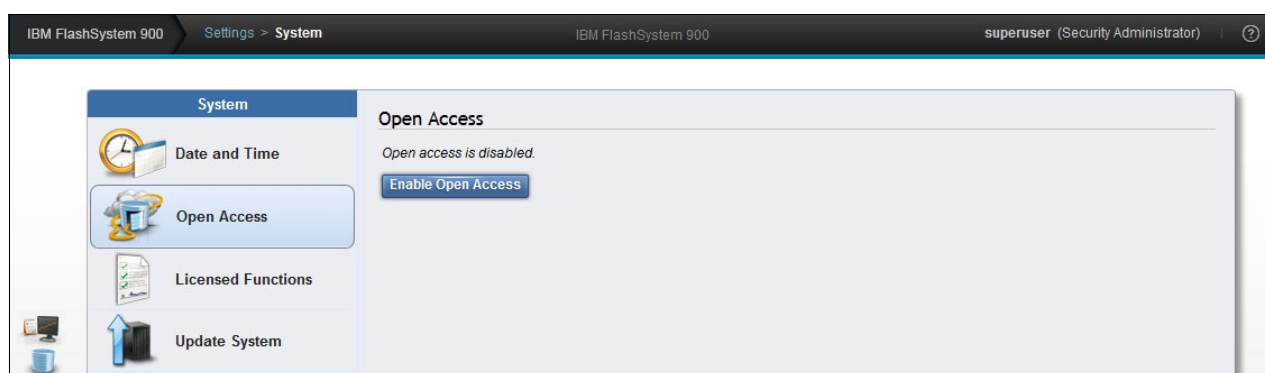


Figure 7-48 Open Access

Your system does not have any hosts defined and Open Access can be enabled. The Enable Open Access button is disabled when hosts are defined and then Open Access is not configurable.

## Update software

**Note:** The upgrade performed on the IBM FlashSystem V840 product is shown and is procedurally the same as the upgrade to the FlashSystem 900.

This section demonstrates how to update firmware through the GUI of the FlashSystem 900. The firmware update that is demonstrated is from version 1.1.2.7 to version 1.1.3.2 and is therefore an update to the current and latest available firmware level. Firmware update is initiated from the Settings > System page.

In the Monitoring > System window, which is also called the Home window of the FlashSystem 900 GUI, right-click and then select **Properties** to get information about the current firmware level (Figure 7-49).

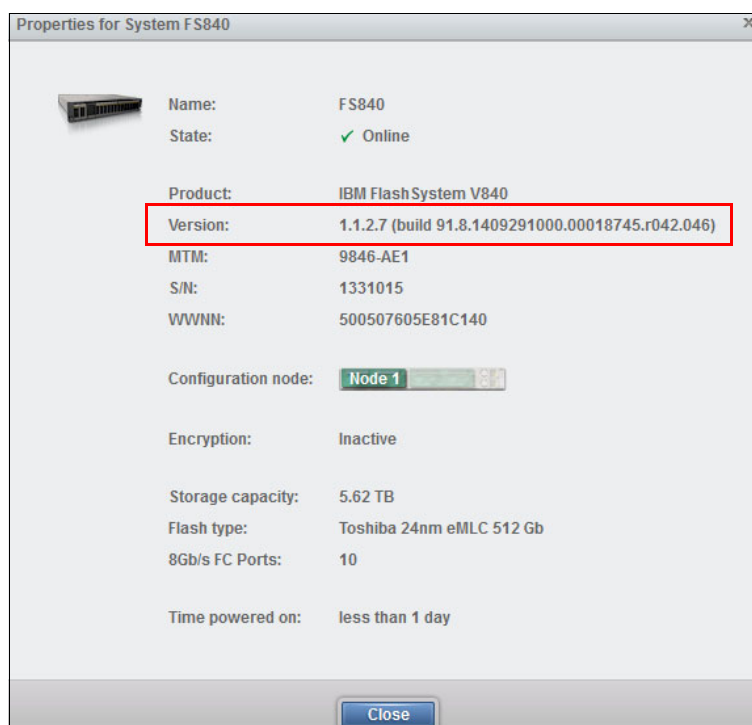


Figure 7-49 The system properties indicate the current firmware version

**Note:** Available for firmware update in this section was a FlashSystem V9000. Regarding the firmware update, the only difference is the name and product number. All steps during the firmware process are exactly the same as with the FlashSystem 900.

Before starting the firmware update, the administrator must download the new firmware image file and the update test utility.

The current firmware for the system can be downloaded from the Internet (if the system has access), or the administrator can download from the following web page:

<http://www.ibm.com/storage/support>

A firmware download requires an appropriate maintenance agreement or indication that the system is covered under warranty. When downloading firmware from IBM, the client must validate coverage by entering the system model number and serial number. The system model number and serial number are on the printed serial number label on the system. You can find them in the GUI, below the firmware version, which as shown in Figure 7-49, they are MTM and S/N.

**Note:** A firmware download from IBM is possible only if the system has an appropriate maintenance agreement or if the machine is under warranty.

The Settings menu allows you to update the FlashSystem 900 firmware. This update is referred to as *Concurrent Code Load* (CCL). Each node in the clustered system automatically updates in sequence while maintaining full accessibility for connected hosts.

To initiate CCL, select **Settings** → **System** → **Update System** → **Update** (Figure 7-50).

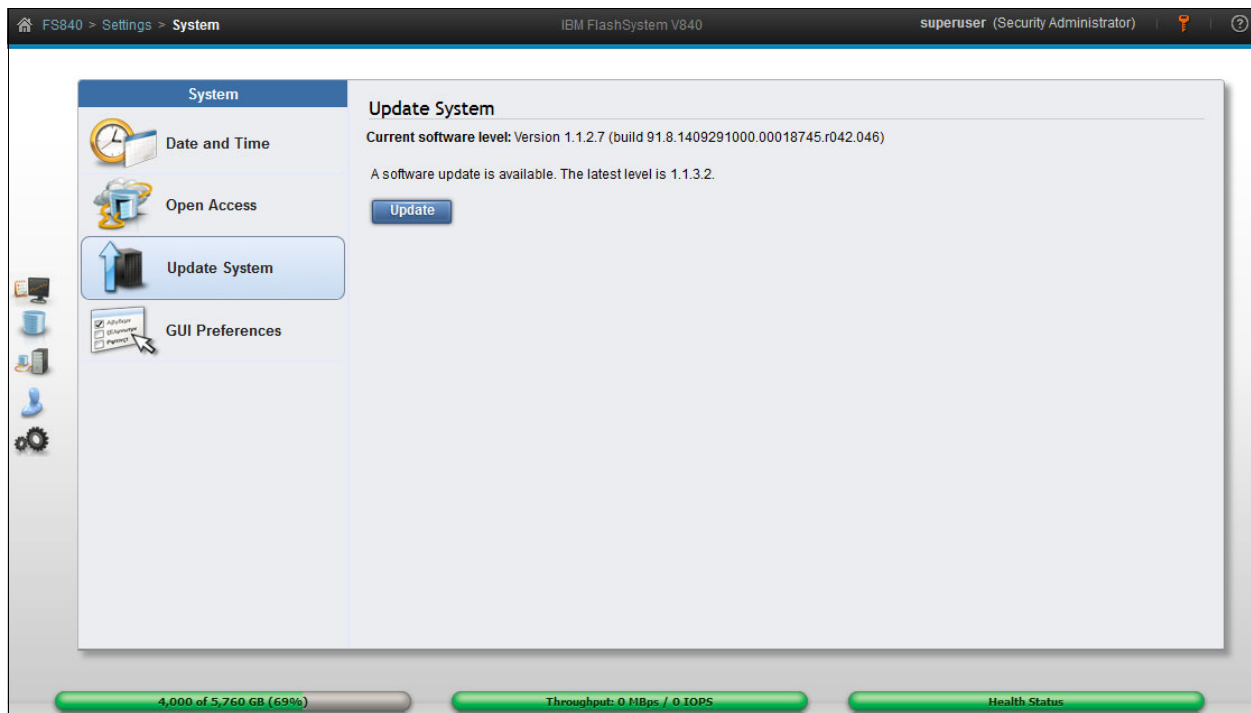


Figure 7-50 Update System

The Update System wizard begins by requesting you to select the test utility and update package. The test utility checks for errors before the firmware update file can be uploaded and the update can be started. Figure 7-51 shows the Update System wizard before files are selected.

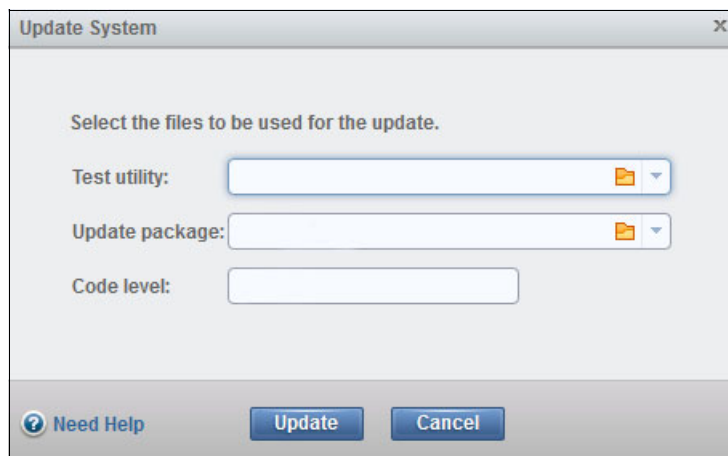


Figure 7-51 Test utility and firmware selection



Click the folder icon to locate the correct test utility and update package. The test utility file is selected (Figure 7-52).

<div>★ Favorites</div> <div>Desktop</div> <div>Downloads</div> <div>Recent Places</div>	Name	Date modified	Type	Size
	IBM9840_INSTALL_1.1.3.2-43.28.tgz.gpg	1/20/2015 11:26 AM	GPG Encrypted File	355,187 KB
	ReleaseNotes_1.1.3.2-43.28-V840.pdf	1/20/2015 11:26 AM	Adobe Acrobat D...	255 KB
	TMS9840_INSTALL_svcupgradetest_1.4	1/20/2015 11:19 AM	4 File	19 KB
	Upgrade_Test_Utility_V1.4_Release_Notes...	1/20/2015 11:19 AM	Adobe Acrobat D...	27 KB

Figure 7-52 Update test utility file selection

The purpose of running the test utility is to verify that no errors exist and that the system is ready to update. If any issues are discovered by the test utility, the firmware update aborts with a message to the administrator of which problems to fix before the update system procedure can be repeated.

The procedure for selecting the update package is shown in Figure 7-53.

<div>★ Favorites</div> <div>Desktop</div> <div>Downloads</div> <div>Recent Places</div>	Name	Date modified	Type	Size
	IBM9840_INSTALL_1.1.3.2-43.28.tgz.gpg	1/20/2015 11:26 AM	GPG Encrypted File	355,187 KB
	ReleaseNotes_1.1.3.2-43.28-V840.pdf	1/20/2015 11:26 AM	Adobe Acrobat D...	255 KB
	TMS9840_INSTALL_svcupgradetest_1.4	1/20/2015 11:19 AM	4 File	19 KB
	Upgrade_Test_Utility_V1.4_Release_Notes...	1/20/2015 11:19 AM	Adobe Acrobat D...	27 KB

Figure 7-53 Update package file selection

Figure 7-54 shows selection of the appropriate files for test utility, update package, and target code level.

Update System

Select the files to be used for the update.

Test utility: TMS9840\_INSTALL\_svcupgradetest\_1.

Update package: IBM9840\_INSTALL\_1.1.3.2-43.28.tgz.gp

Code level: 1.1.3.2

Need Help

Update

Cancel

Figure 7-54 Test utility and firmware selection

The system inserts the latest code level automatically, or the administrator can specify a different firmware level. This example updates to 1.1.3.2.

Click **Update** to proceed. The update test utility and update package files are uploaded to the FlashSystem 900 controller nodes where firmware update begins.

The initial part of the Update System procedure is shown in Figure 7-55.

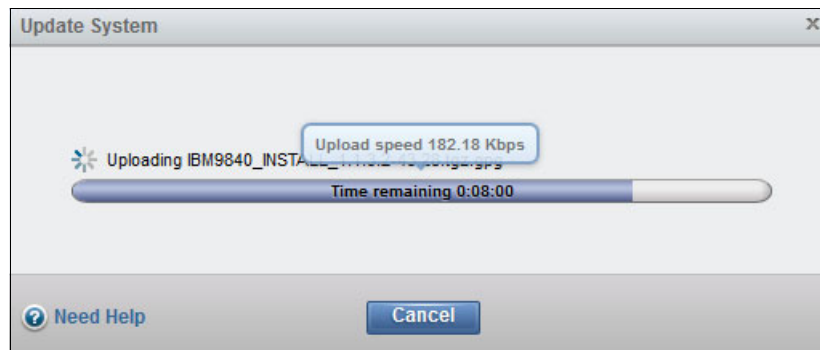


Figure 7-55 Test utility and firmware is uploading

If any errors are identified by the test utility, you must resolve the errors before the firmware update can proceed. Any hardware error prevents System Update from proceeding.

If an error is identified, then take the correct actions to resolve the error identified by the update test utility. You can start troubleshooting in the menu **Monitoring** → **Events**. Use this menu to review and resolve any unfixed errors.

The Concurrent Code Load (CCL) firmware update is now executing in the background. While the system updates, the progress indicators are displayed (Figure 7-56).

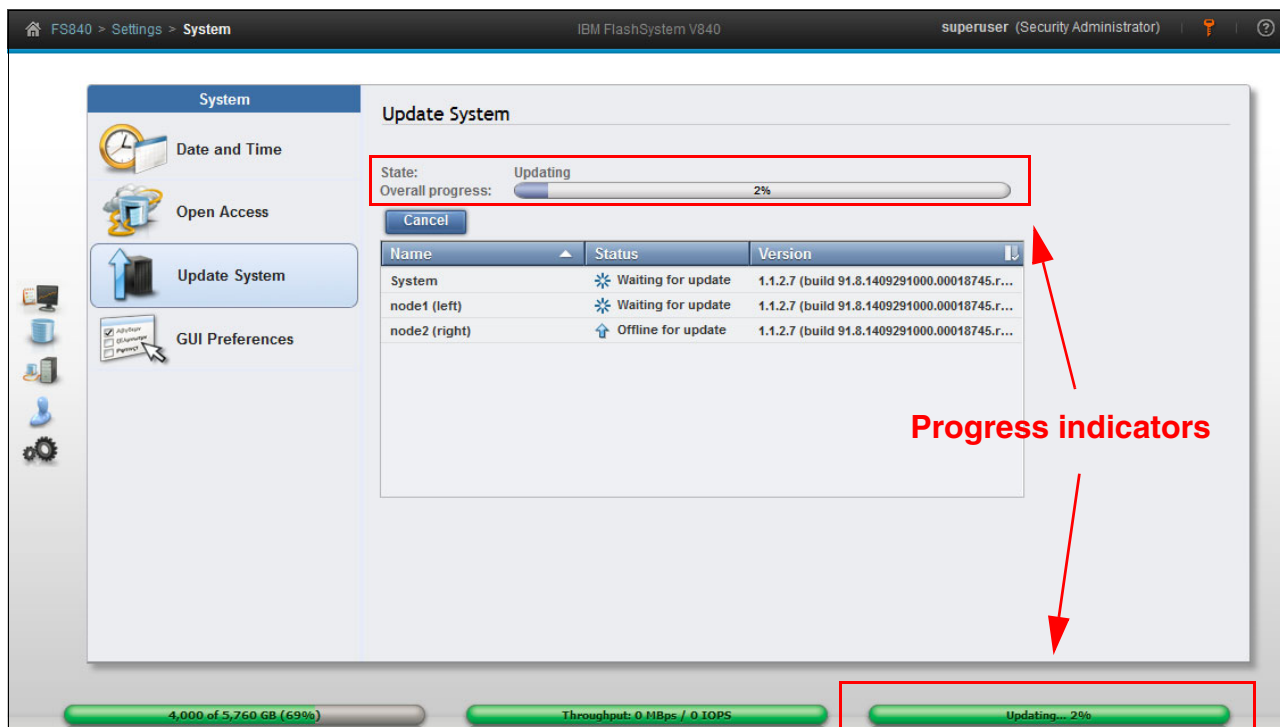


Figure 7-56 Firmware update - node offline for update

The system can be operated normally while it is upgrading; however, no changes can be made until the firmware update completes. If you try to fix any error condition, the Fixes cannot be applied while the system is being upgraded message is displayed (Figure 7-57).

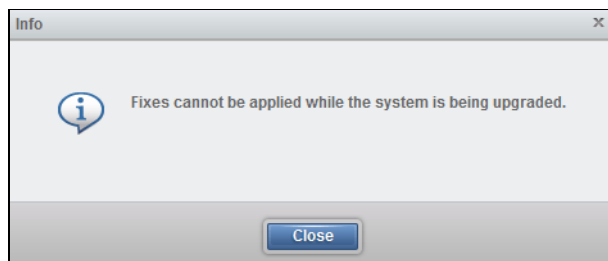


Figure 7-57 Fixes cannot be applied while upgrading

During the update, various messages are displayed in the Update System window. In Figure 7-58, the first controller, which is node2, completes its update and the status of the second controller is Offline for update.

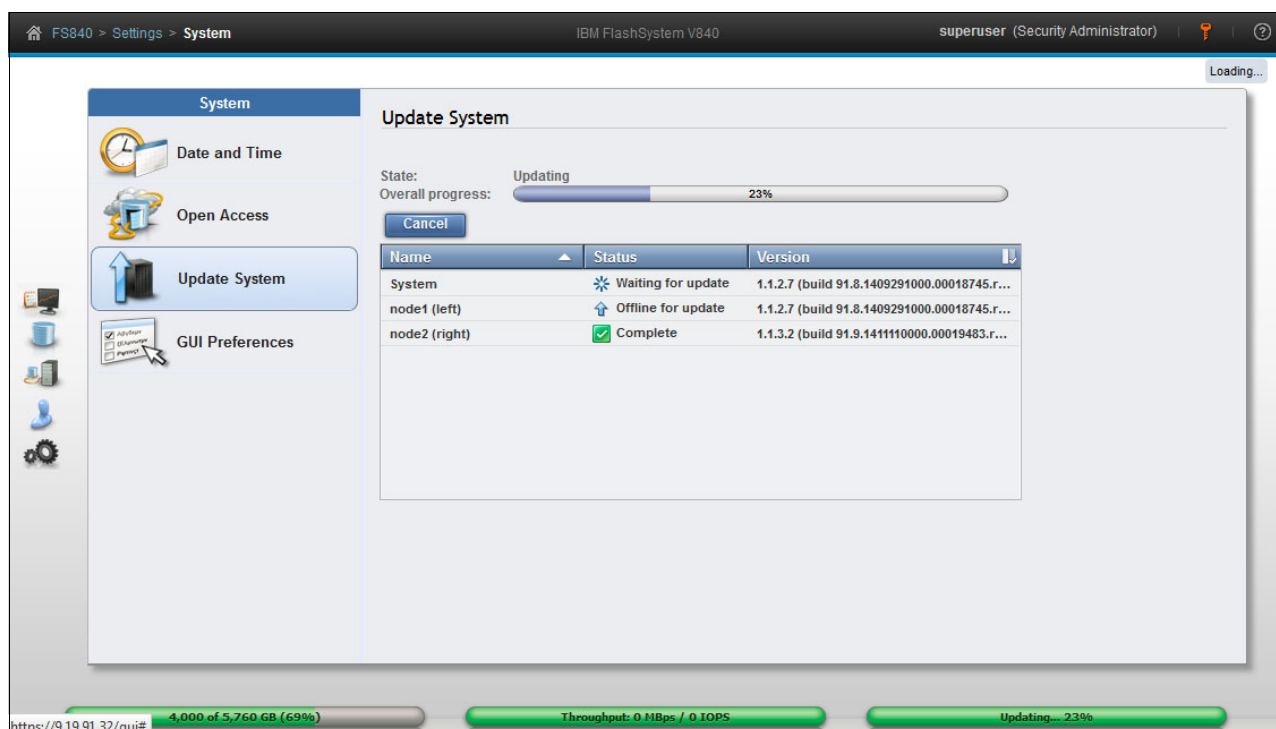


Figure 7-58 Firmware update of the first controller completes

During update, when the node that has the role of configuration node reboots, a message indicates that node failover is detected (Figure 7-59 on page 272). The role of configuration node is now been moved to the other controller and the browser window must be refreshed.

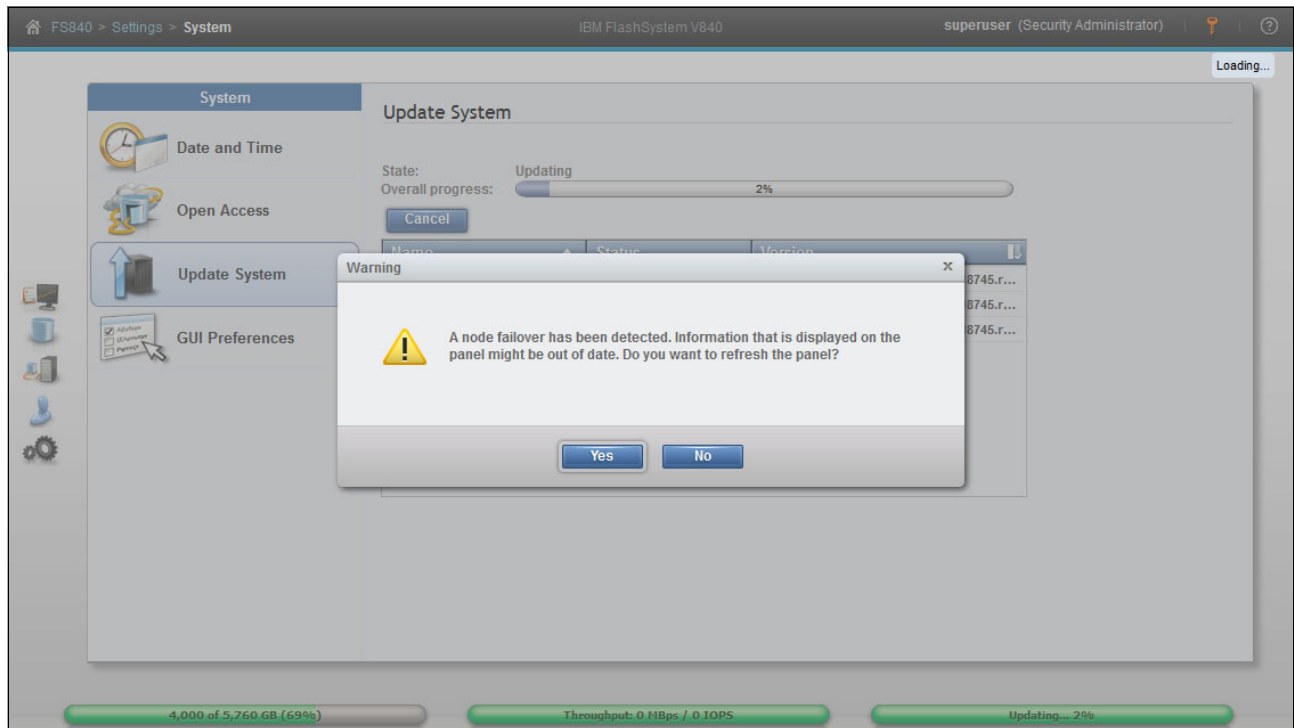


Figure 7-59 Node failover happens during update

When both controllers are updated, the FlashSystem 900 firmware update *commits* the new firmware (Figure 7-60).

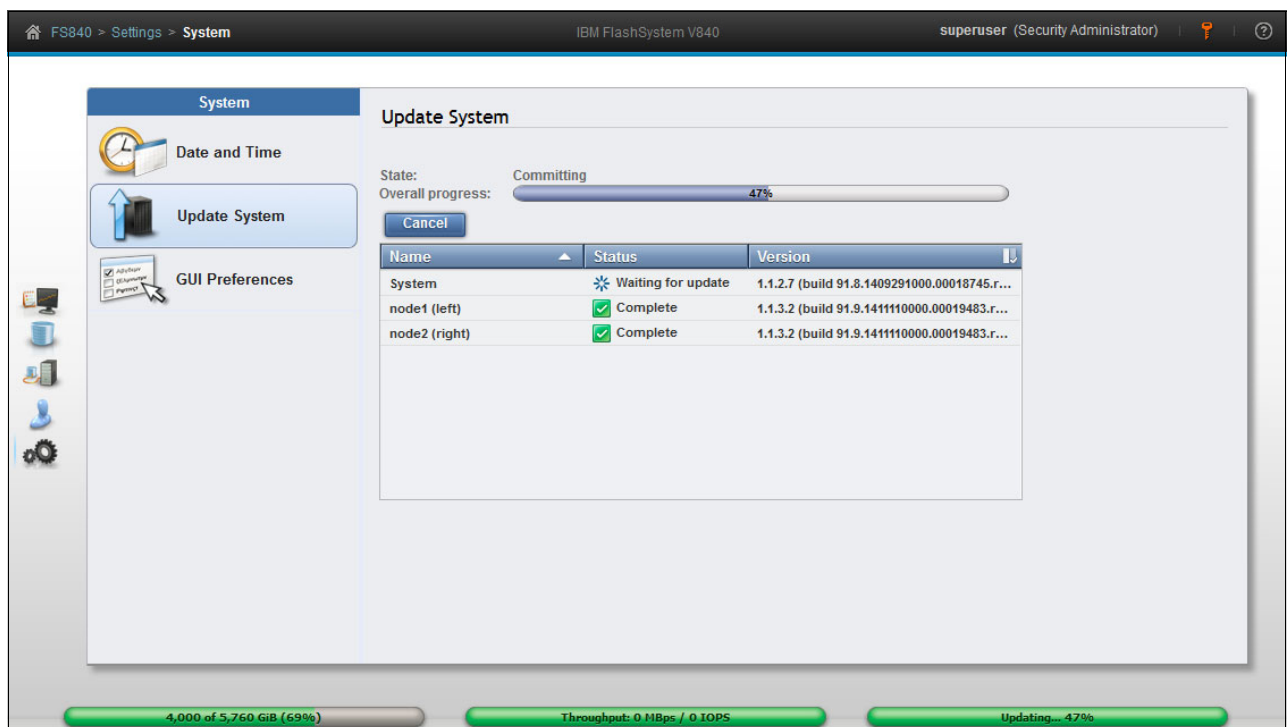


Figure 7-60 Firmware update of both controllers complete

After committing the firmware update, the system starts the *updating hardware* process (Figure 7-61).

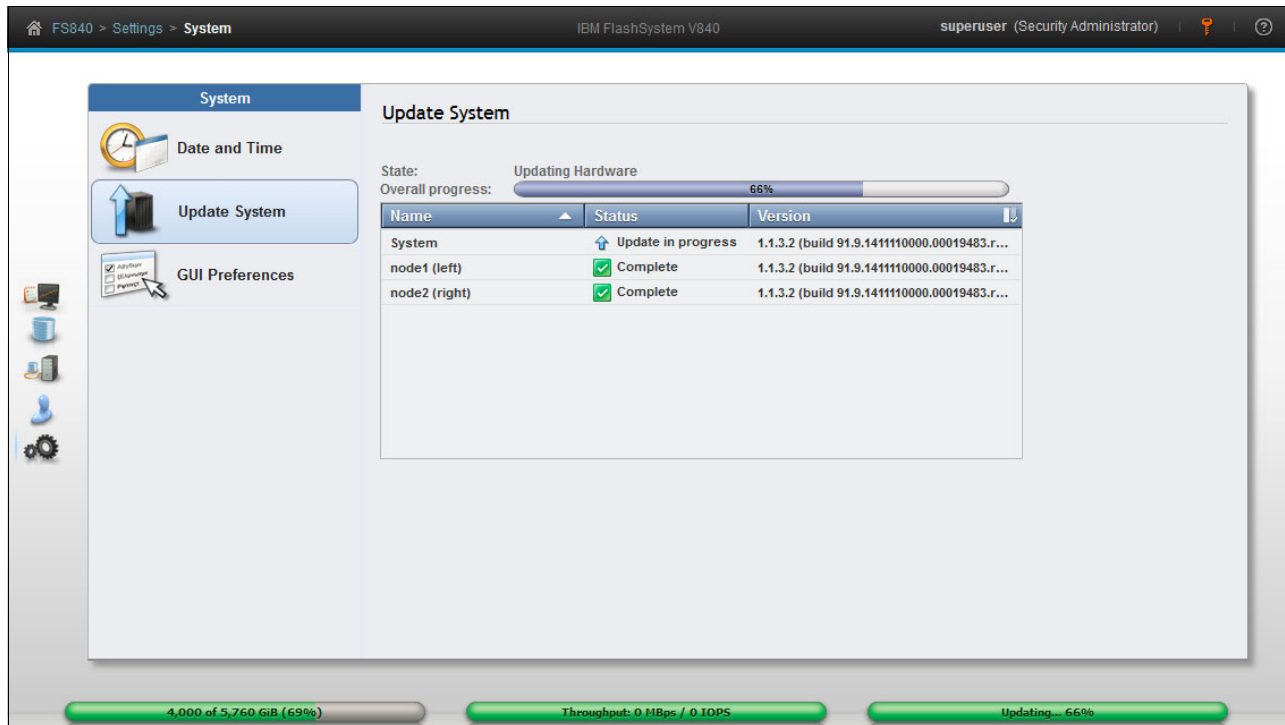


Figure 7-61 Firmware updating the hardware

During the hardware update, all individual components in the system are being firmware-updated. For example, the I/O ports are updated, during which time they are being taken offline, one by one, for update.

When the Update System wizard completes, the system returns to a healthy status. The system now has the latest firmware as shown in Figure 7-62.

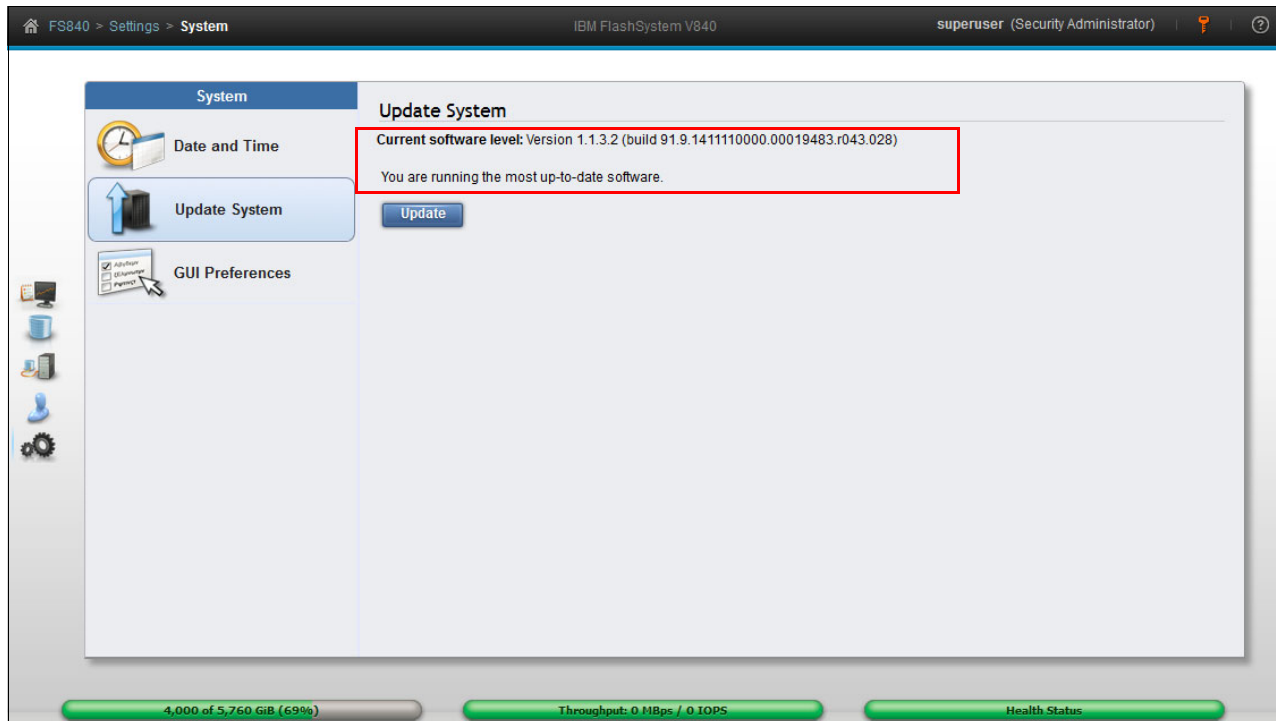


Figure 7-62 Firmware update finished

The new firmware version can be confirmed in the same way, as shown in Figure 7-49 on page 267.

The system properties now display the new firmware version (Figure 7-63).

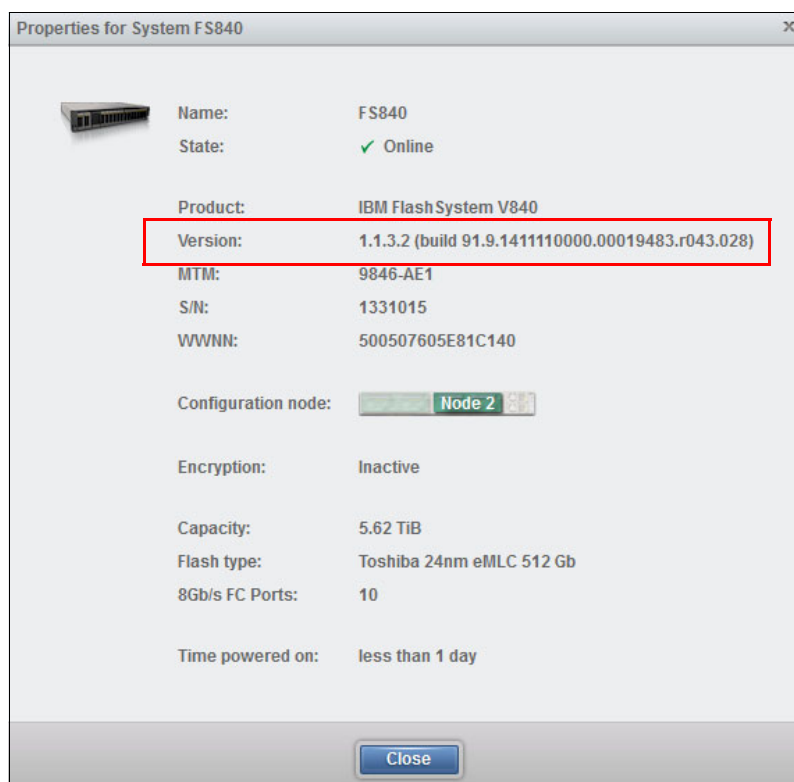


Figure 7-63 System properties with new firmware

The example updated from firmware version 1.1.2.7 to firmware version 1.1.3.2 (Figure 7-63). The entire update process completed in approximately two hours.

As an alternative to upgrading firmware through the GUI, you can use the FlashSystem 900 CLI instead. The process is described at the IBM Knowledge Center:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## GUI preferences

Click **Settings** → **System** → **GUI Preferences** to change the web address of the IBM Knowledge Center, which is the help page for the IBM FlashSystem 900. This help page can be reached from any window in the management GUI by clicking the question mark (?) in the upper-right corner of the GUI (Figure 7-64).

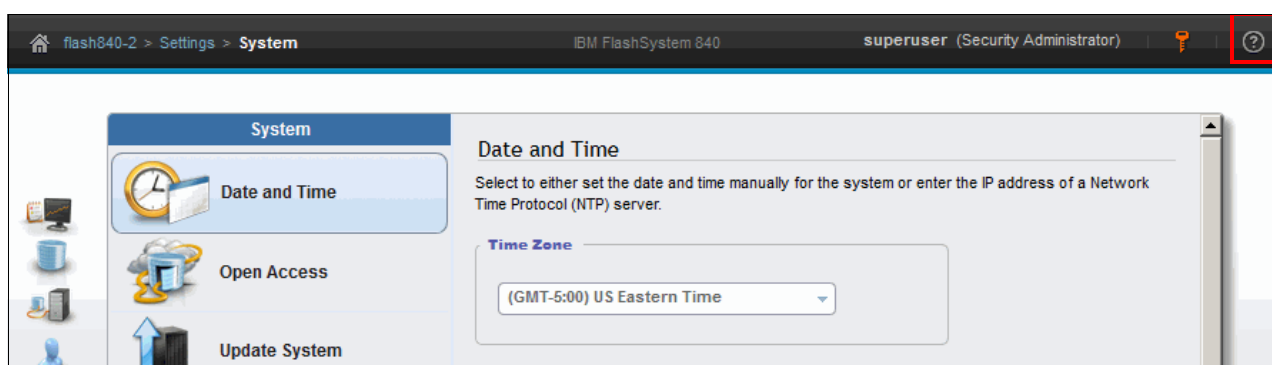


Figure 7-64 IBM Knowledge Center access

The default web address for the IBM FlashSystem 900 web page at the IBM Knowledge Center is the Internet accessible version.

Any address can be configured in the Web Address field. If the system does not have access to the Internet, the web address can be set to access a local version of the IBM FlashSystem 900 web page in the IBM Knowledge Center as shown in Figure 7-65.

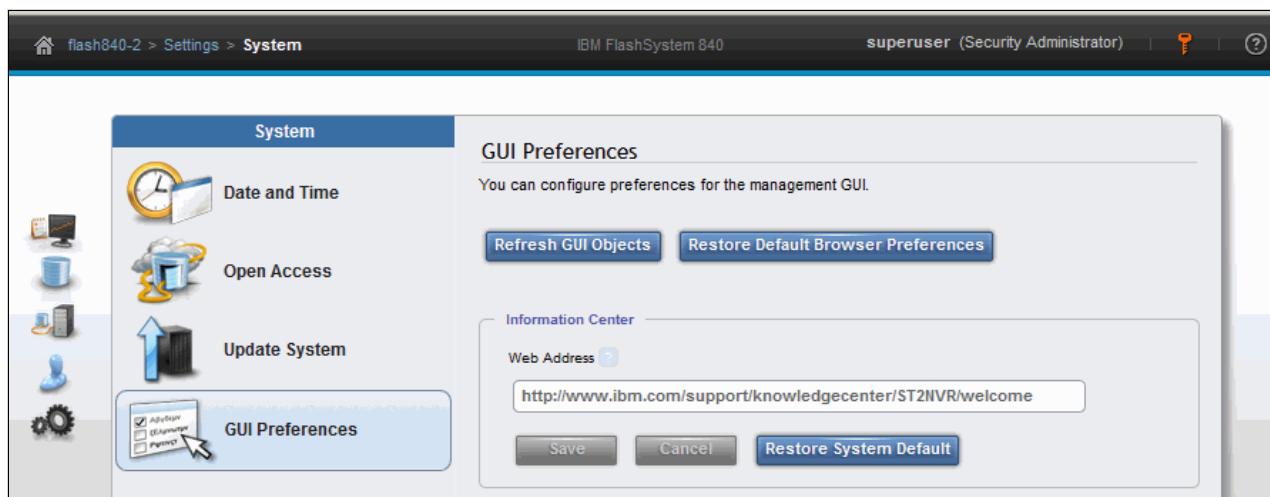


Figure 7-65 System menu: GUI Preferences

The IBM FlashSystem 900 web page in the IBM Knowledge Center is at this web page:

[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)

## 7.2 Service Assistant Tool

*Service Assistant Tool* is used for troubleshooting or when an IBM Support engineer directs you to use it.

### 7.2.1 Accessing Service Assistant Tool

Service Assistant Tool is accessed with a web browser. An example of getting access is to point your browser to the cluster management IP address followed by `/service`. The following URL is an example:

`https://192.168.10.10/service`

Each of the canister's service IP addresses can also be reached. There are different options for getting access to Service Assistant Tool.

The following example lists which IP addresses are configured and how they are accessed:

- ▶ 192.168.10.10 (Service IP address for Canister 1):
  - `https://192.168.10.10/service` opens Service Assistant Tool for Canister 1.
  - `https://192.168.10.10` opens Service Assistant Tool for Canister 1.
- ▶ 192.168.10.11 (Service IP address for Canister 2 (configuration node)):
  - `https://192.168.10.11/service/` opens Service Assistant Tool for Canister 2.
  - `https://192.168.10.11` opens the cluster management GUI.



- ▶ 192.168.10.12 (Cluster IP address):
  - <https://192.168.10.12/service> opens Service Assistant Tool for configuration node.
  - <https://192.168.10.12> opens the cluster management GUI.

**Note:** Canisters are named Canister 1 (view from rear left side) and Canister 2 (view from rear right side). The logical names for canisters in the Service Assistant Tool are node 1 and node 2. Which is node 1 and which is node 2 depends of the actual configuration and in which order the controllers were added to the cluster. If Canister 2 was added first, it gets the logical name node 1. There is, therefore, no direct relation between the canister number and node number.

## 7.2.2 Log in to Service Assistant Tool

The login window of FlashSystem 900 Service Assistant Tool allows only the superuser to log in; therefore, the user name cannot be changed. The Service Assistant Tool login window is shown in Figure 7-66.



Figure 7-66 Service Assistant Tool login

**Attention:** Be careful when you open Service Assistant Tool. Incorrect usage might cause unintended downtime or even data loss. Only use Service Assistant Tool when IBM Support asks you to use it.

After you type the password for the superuser, the Service Assistant Tool displays the information shown in Figure 7-67.

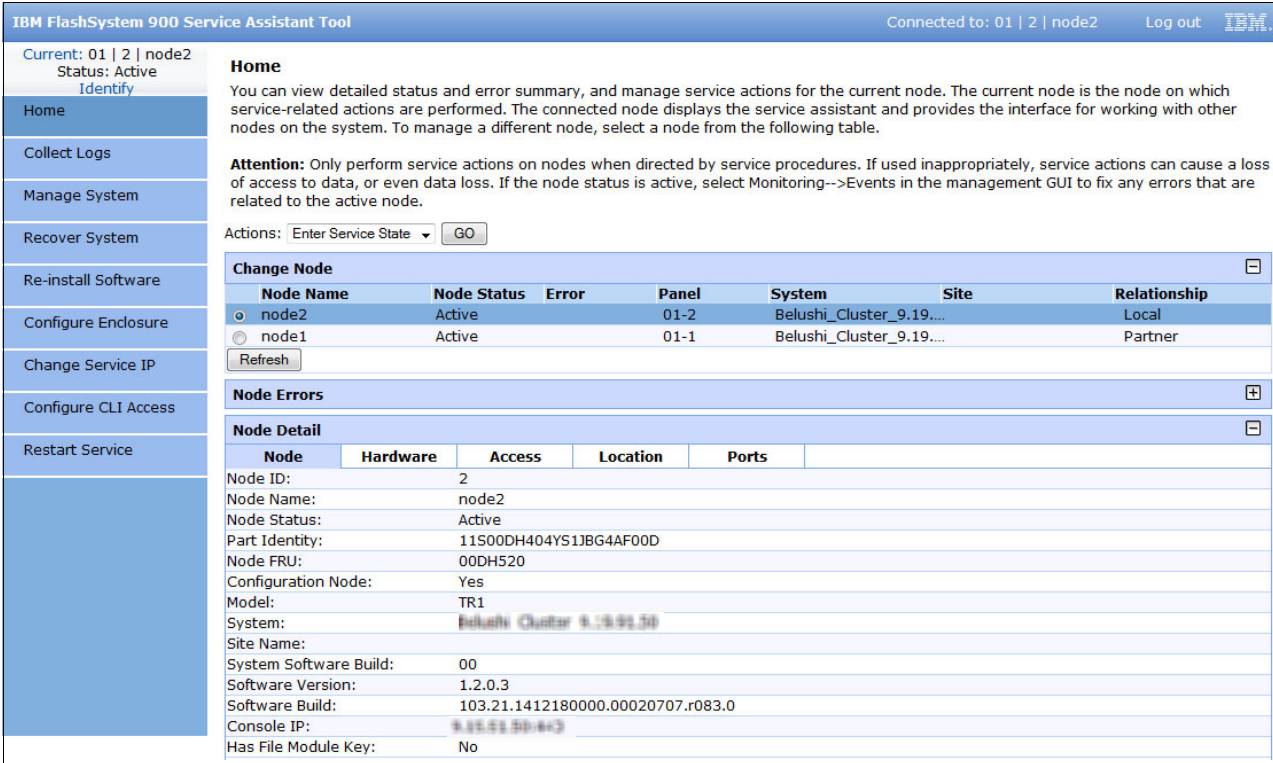


Figure 7-67 Service Assistant Tool main page

The Home window of Service Assistant Tool shows various options for examining installed hardware and revision levels and for identifying canisters or placing these canisters into the service state.



## Product integration

This chapter covers the integration of the IBM FlashSystem 900 storage with the IBM SAN Volume Controller, which delivers the functionality of IBM Spectrum Virtualize, and IBM Storwize V7000. It provides an overview of the main concepts of the products involved, detailed usage considerations, port assignment, port masking, and host connectivity. Additionally, common usage scenarios are described. Throughout, suggestions and preferred practices are identified, where applicable.

This chapter covers the following topics:

- ▶ Running FlashSystem 900 with IBM Spectrum Virtualize - SAN Volume Controller
- ▶ SAN Volume Controller connectivity to FlashSystem 900
- ▶ Integration considerations: FlashSystem 900 and SAN Volume Controller
- ▶ Integration considerations: FlashSystem 900 and IBM Storwize V7000

### Notes:

- ▶ Some of the following sections mention IBM SAN Volume Controller, which delivers the functions of IBM Spectrum Virtualize, part of the IBM Spectrum Storage family.
- ▶ IBM Spectrum Virtualize is industry-leading storage virtualization that enhances existing storage to improve resource utilization and productivity so you can achieve a simpler, more scalable and cost-efficient IT infrastructure.
- ▶ The functionality of IBM Spectrum Virtualize is provided by IBM SAN Volume Controller.
- ▶ For more details, see the following web page:

<http://www.ibm.com/systems/storage/software/virtualization/svc/index.html>

## 8.1 Running FlashSystem 900 with IBM Spectrum Virtualize - SAN Volume Controller

IBM FlashSystem 900 is all about being fast and resilient to minimize latency by using IBM FlashCore hardware-accelerated architecture, IBM MicroLatency modules, and many other advanced flash management features and capabilities.

For clients who want advanced software features, the IBM FlashSystem 900 and IBM System Storage SAN Volume Controller software IBM SAN Volume Controller, which delivers the functions of IBM Spectrum Virtualize, provide an enterprise-class solution by integrating SAN Volume Controller functions and services, such as mirroring, FlashCopy, thin provisioning, Real-time Compression, and broader host support. The best way to achieve all of that function is by deploying the IBM FlashSystem 900 behind a SAN Volume Controller. For clients who need efficiency, this combination also can be used with IBM Easy Tier, with SAN Volume Controller automatically promoting hot blocks to the FlashSystem. The tiered storage solution not only efficiently uses the FlashSystem to increase performance in critical applications, but it also can reduce costs by migrating less critical data to less expensive media.

You can also order IBM FlashSystem V9000, which is a comprehensive all-flash enterprise storage solution. FlashSystem V9000 delivers the full capabilities of IBM FlashCore technology plus a rich set of storage virtualization features.

The FlashSystem V9000 improves business application availability and delivers greater resource utilization so that you can get the most from your storage resources, and achieve a simpler, more scalable, and more cost-efficient IT Infrastructure. Using IBM Storwize family functions, management tools, and interoperability, this product combines the performance of the FlashSystem architecture with the advanced functions of software-defined storage to deliver performance, efficiency, and functions that meet the needs of enterprise workloads demanding IBM MicroLatency response time. For product details of the FlashSystem V840, see the IBM Redbooks Product Guide, *IBM FlashSystem V840 Enterprise Performance Solution*, TIPS1158.

The next topics show how to configure IBM FlashSystem 900 to provide storage to SAN Volume Controller and show how they are designed to operate seamlessly together, reducing management effort. At the time of the writing of this chapter, SAN Volume Controller software version 7.4 was available.

### 8.1.1 IBM System Storage SAN Volume Controller introduction

IBM System Storage SAN Volume Controller is a storage virtualization solution that helps to increase the utilization of existing storage capacity and to centralize the management of multiple controllers in an open system storage area network (SAN) environment.

SAN Volume Controller (machine type 2145 and accompanying software) supports attachment to both IBM and non IBM storage systems. For the most current SAN Volume Controller supported hardware list, see the “Supported Hardware List, Device Driver, and Firmware Levels & Supported Software for SAN Volume Controller” topic at this web page:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1003658>

SAN Volume Controller enables storage administrators to reallocate and scale storage capacity and make changes to underlying storage systems without disruption to applications. SAN Volume Controller also provides the ability to simplify storage infrastructure, use storage

resources more efficiently, improve personnel productivity, and increase application availability.

SAN Volume Controller pools storage volumes from IBM and non IBM disk arrays into a single reservoir of capacity, which can be managed from a central point. SAN Volume Controller also allows data to be migrated between heterogeneous disk arrays without disruption to applications. By moving functionality of copy services into the network, SAN Volume Controller allows you to use a standardized suite of copy services tools that can be applied across the entire storage infrastructure, irrespective of storage vendor restrictions that normally apply for the individual disk controllers in use.

Also, SAN Volume Controller adds functions to the infrastructure that might not be present in each virtualized subsystem. Examples include thin provisioning, automated tiering, volume mirroring, and data compression.

**Note:** Some of the SAN Volume Controller functions that were mentioned are included in the base virtualization license, although for other functions, you might need to purchase an additional license. Contact your IBM representative for assistance about extra licenses.

## **SAN Volume Controller design overview**

The IBM System Storage SAN Volume Controller is designed to handle the following tasks:

- ▶ Combine storage capacity from multiple vendors into a single repository of capacity with a central management point
- ▶ Help increase storage utilization by providing host applications with more flexible access to capacity
- ▶ Help improve the productivity of storage administrators by enabling the management of combined storage volumes from a single, easy to use interface
- ▶ Support improved application availability by insulating host applications from changes to the physical storage infrastructure
- ▶ Enable a tiered storage environment, in which the cost of storage can be better matched to the value of data
- ▶ Support advanced copy services, from higher-cost devices to lower-cost devices and across subsystems from multiple vendors

SAN Volume Controller combines hardware and software into a comprehensive, modular appliance. By using Intel processor-based servers in highly reliable clustered pairs, SAN Volume Controller has no single point of failure. SAN Volume Controller software forms a highly available cluster that is optimized for performance and ease of use.

### ***Storage utilization***

SAN Volume Controller is designed to help increase the amount of storage capacity that is available to host applications. By pooling the capacity from multiple disk arrays within the SAN, it enables host applications to access capacity beyond their island of SAN storage. The Storage Networking Industry Association (SNIA) estimates that open systems disk utilization in a non-virtualized environment is currently only 30 - 50%. With storage virtualization, this utilization can grow up to 80% on average. For more details, see *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933-03.

## **Scalability**

A SAN Volume Controller configuration can start with a single I/O group. An *I/O group* is a pair of high performance, redundant Intel processor-based servers, referred to as *nodes* or *storage engines*. Highly available I/O groups are the basic configuration of a cluster. Adding extra I/O groups (a nondisruptive process) can help increase cluster performance and bandwidth.

SAN Volume Controller can scale out to support up to four I/O groups. SAN Volume Controller supports up to 2,048 host servers. For every cluster, SAN Volume Controller supports up to 8,192 volumes, each one up to 256 TB in size, and a total virtualized capacity of up to 32 petabytes (PB).

Enhanced stretched cluster configurations provide highly available, concurrent access to a single copy of data from data centers up to 300 km apart and enable nondisruptive storage and virtual machine mobility between data centers.

**Note:** 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:  
<http://www.ibm.com/support/docview.wss?uid=ssg1S1003658>

This configuration flexibility means that SAN Volume Controller configurations can start small with an attractive price to suit smaller environments or pilot projects, and then can grow with your business to manage large storage environments.

## **Management**

SAN Volume Controller is managed at the cluster level, providing a single point of control over all the managed storage. SAN Volume Controller provides a comprehensive, easy-to-use graphical user interface (GUI) for central management. This simple interface incorporates the Storage Management Initiative Specification (SMI-S) application programming interface (API), and further demonstrates the commitment of IBM to open standards. SAN Volume Controller cluster can also be managed and monitored through a comprehensive command-line interface (CLI) through Secure Shell (SSH), enabling the use of scripts for automated repeatable operations.

The SAN Volume Controller GUI is designed for ease of use and includes many built-in IBM guidelines that simplify storage provisioning and enable new clients to get started quickly with a rapid learning curve.

Clients using IBM Spectrum Control (formerly IBM Tivoli Storage Productivity Center), IBM Systems Director, and IBM Spectrum Protect Snapshot (formerly IBM Tivoli Storage FlashCopy Manager) can take further advantage of integration points with SAN Volume Controller. Managing SAN Volume Controller under IBM Spectrum Control, enables to management of the most common day-to-day activities for SAN Volume Controller without ever needing to leave the IBM Spectrum Control user interface.

For historic performance and capacity management from the perspectives of both the host and the virtualized storage devices, IBM Spectrum Control helps clients with an end-to-end view and control of the virtualized storage infrastructure. Regarding data protection, IBM Spectrum Protect Snapshot helps integrate SAN Volume Controller FlashCopy function with major applications for consistent backups and restores.

### ***Linking infrastructure performance to business goals***

By pooling storage into a single pool, SAN Volume Controller helps insulate host applications from physical changes to the storage pool, removing disruption. SAN Volume Controller simplifies storage infrastructure by including a dynamic data-migration function, allowing for online volume migration from one device to another. By using this function, administrators can reallocate, scale storage capacity, and apply maintenance to storage subsystems without disrupting applications, increasing application availability.

With SAN Volume Controller, your business can build an infrastructure from existing assets that is simpler to manage, easier to provision, and can be changed without impact to application availability. Businesses can use their assets more efficiently and actually measure the improvements. They can allocate and provision storage to applications from a single view and know the effect on their overall capacity instantaneously. They can also quantify improvements in their application availability to enable better quality of service goals. These benefits help businesses manage their costs and capabilities more closely, linking the performance of their infrastructure to their individual business goals.

### ***Tiered storage***

In most IT environments, inactive data makes up the bulk of stored data. SAN Volume Controller helps administrators control storage growth more effectively by moving low-activity or inactive data into a hierarchy of lower-cost storage. Administrators can free disk space on higher-value storage for more important, active data.

Tiered storage is achieved by easily creating various groups of storage, or *storage pools*, corresponding to underlying storage with various characteristics. Examples are speed and reliability. With SAN Volume Controller, you can better match the cost of the storage used to the value of data placed on it.

### ***Technology for an on-demand environment***

Businesses are facing growth in critical application data that is supported by complex heterogeneous storage environments, while their staffs are overburdened. SAN Volume Controller is one of many offerings in the IBM System Storage portfolio that are essential for an on-demand storage environment. These offerings can help you to simplify your IT infrastructure, manage information throughout its lifecycle, and maintain business continuity.

## **8.1.2 SAN Volume Controller architecture and components**

SAN-based storage is managed by SAN Volume Controller in one or more *I/O groups (pairs)* of SAN Volume Controller *nodes*, referred to as a *clustered system*. These nodes are attached to the SAN fabric, with storage controllers and host systems. The SAN fabric is zoned to allow SAN Volume Controller to “see” the storage controllers, and for the hosts to “see” SAN Volume Controller.

The hosts are not allowed to “see” or operate on the same physical storage (logical unit number (LUN)) from the storage controllers that are assigned to SAN Volume Controller, and all data transfer happens through SAN Volume Controller nodes. This design is commonly described as *symmetric virtualization*.

Storage controllers can be shared between SAN Volume Controller and direct host access as long as the same LUNs are not shared, and both types of access use compatible multipathing drives in the same host or operating system instance. The zoning capabilities of the SAN switch must be used to create distinct zones to ensure that this rule is enforced.

Figure 8-1 shows a conceptual diagram of a storage environment using SAN Volume Controller. It shows several hosts that are connected to a SAN fabric or LAN with SAN Volume Controller nodes and the storage subsystems that provide capacity to be virtualized. In practical implementations that have high-availability requirements (most of the target clients for SAN Volume Controller), the SAN fabric “cloud” represents a redundant SAN. A redundant SAN consists of a fault-tolerant arrangement of two or more counterpart SANs, providing alternate paths for each SAN-attached device.

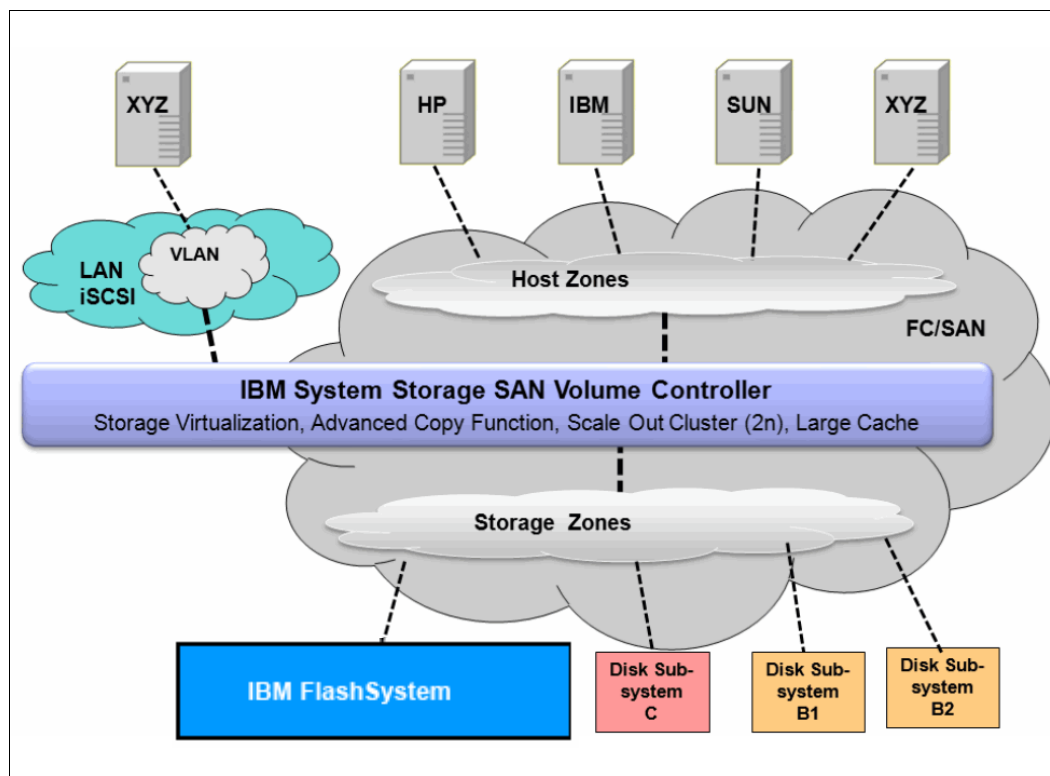


Figure 8-1 Conceptual diagram of SAN Volume Controller and the SAN infrastructure

Both scenarios (using a single network and using two physically separate networks) are supported for Internet Small Computer System Interface (iSCSI)-based and LAN-based access networks to SAN Volume Controller. Redundant paths to volumes can be provided in both scenarios. For iSCSI-based access, using two networks and separating iSCSI traffic within the networks by using a dedicated virtual local area network (VLAN) for storage traffic prevent any IP interface, switch, or target port failure from compromising the host servers' access to the volumes.

A *clustered system* of SAN Volume Controller nodes that are connected to the same fabric presents logical disks or *volumes* to the hosts. These volumes are created from managed LUNs or *managed disks* (MDisks) that are presented to SAN Volume Controller by the storage subsystems and grouped in *storage pools*. Two distinct zone sets are shown in the fabric:

- ▶ Host zones, in which the hosts can see and address SAN Volume Controller nodes and access volumes
- ▶ Storage zones, in which SAN Volume Controller nodes can see and address the MDisks/LUNs that are presented by the storage subsystems



Figure 8-2 shows the logical architecture of SAN Volume Controller, illustrating how different storage pools are built grouping MDisks, and how the volumes are created from those storage pools and presented to the hosts through I/O groups (pairs of SAN Volume Controller nodes). In this diagram, Vol2, Vol7, and Vol8 are mirrored volumes, or volumes with two copies, with each copy residing in a different storage pool. For more information about volume mirroring, see “SAN Volume Controller volume mirroring” on page 304.

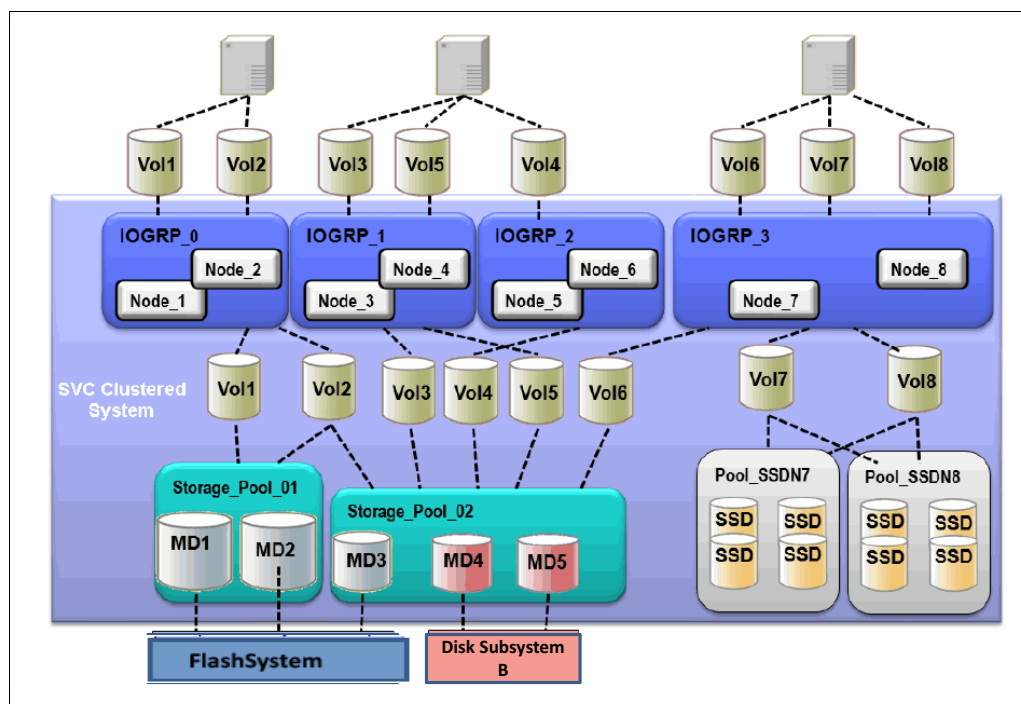


Figure 8-2 Overview of a SAN Volume Controller clustered system, hosts, and storage subsystems

Each MDisk in the storage pool is divided into a number of *extents*. The size of the extent is selected by the administrator at the creation time of the storage pool and cannot be changed later. The size of the extent ranges from 16 MB to 8192 MB. The volume that will reside on the storage pool might be formatted in two types: *sequential* or *striped* (default). For more details, see *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933.

### 8.1.3 SAN Volume Controller hardware options

Throughout its lifecycle, SAN Volume Controller has used IBM System x server technology to offer a modular, flexible platform for storage virtualization that can be rapidly adapted in response to changing market demands and evolving client requirements. This “flexible hardware” design allows you to quickly incorporate differentiating features. Significant hardware options are available.

#### Hardware options for SAN Volume Controller

SAN Volume Controller 2145-DH8 node introduces numerous hardware enhancements. Several of these enhancements relate directly to the Real-time Compression feature and offer significant performance and scalability improvements over previous hardware versions.

## Additional and enhanced CPU options

The 2145-DH8 node offers an updated primary CPU that contains 8 cores as compared to the 4-core and 6-core CPUs available in previous hardware versions. Also, the 2145-DH8 node offers the option of a secondary 8-core CPU for use with Real-time Compression. This additional, compression-dedicated CPU allows for improved overall system performance when using compression over previous hardware models.

**Note:** To use the Real-time Compression feature on 2145-DH8 nodes, the secondary CPU is required.

## Increased memory options

The 2145-DH8 node offers the option to increase the node memory from the base 32 GB to 64 GB, for use with Real-time Compression. This additional, compression-dedicated memory allows for improved overall system performance when utilizing compression over previous hardware models.

**Note:** To use the Real-time Compression feature on 2145-DH8 nodes, the additional 32 GB memory option is required.

## Quick Assist compression acceleration cards

The 2145-DH8 node offers the option to include one or two Intel Quick Assist compression acceleration cards based on the Coletto Creek chipset. The introduction of these Intel based compression acceleration cards in SAN Volume Controller 2145-DH8 node is an industry first, providing dedicated processing power and greater throughput over previous models.

**Note:** To use the Real-time Compression feature on 2145-DH8 nodes, at least one Quick Assist compression acceleration card is required. With a single card, the maximum number of compressed volumes per I/O group is 200. With the addition of a second Quick Assist card, the maximum number of compressed volumes per I/O group is 512.

For more details about the compression accelerator cards, see *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933.

## Additional HBAs

The 2145-DH8 node offers the option to add one or two additional 4-port 8 Gbps or 2-port 16 Gbps FC HBA to improve connectivity options on SAN Volume Controller engine.

The following example scenarios describe where these additional ports can provide benefits:

- ▶ Isolation of node-to-node communication, potentially boosting write performance
- ▶ Isolation of node to IBM FlashSystem communication, allowing for maximum performance
- ▶ Isolation of remote copy traffic, avoiding performance problems

The additional HBA support on 2145-DH8 nodes requires SAN Volume Controller Storage Software Version 7.3 or higher, for 8 Gbps for 16 Gbps FC HBAs.

Support for a second SAN Volume Controller HBA was introduced with Software Version 7.1 and the 2145-CG8 nodes.

**Note:** For information about SAN Volume Controller V7.4 Configuration Limits and Restrictions for IBM System Storage SAN Volume Controller, see this web page:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1004924>

## Port masking

The addition of more Fibre Channel HBA ports allows you to optimize your SAN Volume Controller configuration by using dedicated ports for certain system functions. However, the addition of these ports necessitates the ability to ensure traffic isolation.

These two examples show traffic types that you might want to isolate by using port masking:

- ▶ Local node-to-node communication
- ▶ Remote copy traffic

Port masking was introduced with SAN Volume Controller Storage Software Version 7.1. This feature enables better control of SAN Volume Controller node ports. Host port masking is supported in earlier SAN Volume Controller software versions. In those versions, host port masking provides the ability to define which SAN Volume Controller node ports were used to communicate with hosts.

The enhanced port masking in SAN Volume Controller Storage Software Version 7.1 and later provides the ability to restrict intracluster communication and replication communication to specific ports, ensuring that these traffic types only occur on the ports that you want. This capability eliminates the possibility of host or back-end port congestion because of intracluster communication or replication communication.

### Notes:

- ▶ A SAN Volume Controller node attempts to communicate with other SAN Volume Controller nodes over any available path. Port masking, when enabled, ensures that this will not occur on that port.
- ▶ To use the port masking feature, use the `chsystem -localfcportmask` or `-partnerfcportmask` command.

The features in storage area network (SAN) zoning and the physical port assignment provide greater control and enable less congestion and better usage of SAN Volume Controller ports.

**Order of configuration:** When using port masking with SAN Volume Controller, follow this configuration order:

1. Configure intracluster port masking.
2. Configure replication port masking (if using replication).
3. Configure SAN zones for intracluster communication.
4. Configure SAN zones for replication communication (if using replication).
5. Configure SAN zones for all back-end storage communication.
6. Configure SAN zones for host communication.

For more information about how to configure SAN Volume Controller port masking, see this web page:

<http://www.ibm.com/support/docview.wss?uid=s5g1S1004659>

## SAN Volume Controller Stretched Cluster

The SAN Volume Controller Stretched Cluster configurations are supported by the IBM FlashSystem storage systems. When using the FlashSystem storage systems in the SAN Volume Controller Stretched Cluster environments, follow the guidance in the following Redbooks publications:

- ▶ *IBM SAN and SVC Stretched Cluster and VMware Solution Implementation*, SG24-8072
- ▶ *IBM SAN Volume Controller Stretched Cluster with PowerVM and PowerHA*, SG24-8142

## 8.1.4 IBM Spectrum Virtualize - SAN Volume Controller advanced functionality

The combination of the IBM FlashSystem 900 and IBM Spectrum Virtualize, which delivers the functionality of IBM SAN Volume Controller, help you to take advantage of the speed of the IBM FlashSystem 900 and the robust storage management capabilities of SAN Volume Controller. IBM SAN Volume Controller offers features that enrich any storage environment with introducing minimal delay or latency in the I/O path. This section describes SAN Volume Controller features and benefits.

### Thin provisioning

The *thin provisioning* function helps automate provisioning and improve productivity by enabling administrators to focus on overall storage deployment and utilization, and longer-term strategic requirements, without being distracted by routine everyday storage provisioning.

When creating a thin-provisioned volume, the user specifies two capacities: the real physical capacity allocated to the volume from the storage pool, and its virtual capacity available to the host. Therefore, the real capacity determines the quantity of MDisk extents that is initially allocated to the volume. The *virtual capacity* is the capacity of the volume reported to all other SAN Volume Controller components (for example, FlashCopy, cache, and Remote copy) and to the host servers.

The *real capacity* is used to store both the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value or a percentage of the virtual capacity. Thin-provisioned volumes can be used as volumes assigned to the host, by FlashCopy, and Remote copy, to implement thin-provisioned targets, and also with the mirrored volumes feature.

### FlashCopy

The *FlashCopy* function is designed to create an almost instant copy (or “snapshot”) of active data that can be used for backup purposes or for parallel processing activities. Up to 256 copies of data can be created.

FlashCopy works by creating one or two (for incremental operations) bitmaps to track changes to the data on the source volume. This bitmap is also used to present an image of the source data at the point in time that the copy was taken to target hosts while the actual data is being copied. This capability ensures that copies appear to be instantaneous.

FlashCopy permits the management operations to be coordinated, through a grouping of FlashCopy pairs, so that a common single point in time is chosen for copying target volumes from their respective source volumes. This capability is called *consistency groups* and allows a consistent copy of data for an application that spans multiple volumes.

IBM offers IBM Spectrum Protect Snapshot (formerly Tivoli Storage FlashCopy Manager) that is designed to perform near-instant application-aware snapshot backups using SAN Volume Controller FlashCopy, but with minimal impact to IBM DB2, Oracle, SAP, Microsoft SQL Server, or Microsoft Exchange. IBM Spectrum Protect Snapshot also helps reduce backup and recovery times from hours to a few minutes.

For more information, see the IBM Spectrum Protect Snapshot web page:

<http://www.ibm.com/software/products/en/spectrum-protect-snapshot>

## Easy Tier

SAN Volume Controller Easy Tier function helps improve performance at a lower cost through more efficient use of storage. *Easy Tier* is a performance function that automatically migrates or moves extents off a volume to, or from, one MDisk storage tier to another MDisk storage tier. Easy Tier monitors the host I/O activity and latency on the extents of all volumes with the Easy Tier function turned on in a multitier storage pool over a 24-hour period.

Next, Easy Tier creates an extent migration plan based on this activity and then dynamically moves high activity or hot extents to a higher disk tier within the storage pool. It also moves extents whose activity has dropped off or cooled from the higher-tier MDisk back to a lower-tiered MDisk.

SAN Volume Controller Easy Tier can deliver up to a three-time performance improvement with only 5% flash storage capacity. Easy Tier can use flash storage, whether deployed in SAN Volume Controller nodes or in virtualized disk systems, to benefit all virtualized storage. This approach delivers greater benefits from flash storage than tiering systems that are limited to only a single disk system.

Because the Easy Tier function is so tightly integrated, functions, such as data movement, replication, and management, all can be used with flash in the same way as for other storage. SAN Volume Controller helps move critical data to and from flash storage as needed without application disruptions. Combining SAN Volume Controller with the FlashSystem storage devices delivers the best of both technologies: extraordinary performance for critical applications with IBM MicroLatency, coupled with sophisticated functionality.

## Mirroring and copy services

With many conventional SAN disk arrays, replication operations are limited to in-box or like-box-to-like-box circumstances. Functions from different vendors can operate in different ways, which make operations in mixed environments more complex and increase the cost of changing storage types. But SAN Volume Controller is designed to enable administrators to apply a single set of advanced network-based replication services that operate in a consistent manner, regardless of the type of storage being used.

SAN Volume Controller supports remote mirroring to enable organizations to create copies of data at remote locations for disaster recovery. Metro Mirror supports synchronous replication at distances up to 300 km (186.4 miles). Global Mirror supports asynchronous replication up to 8,000 km (4970.9 miles). Replication can occur between any Storwize family systems, and can include any supported virtualized storage. Remote mirroring works with Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), and IP (Ethernet) networking between sites.

With IP networking, the IBM Storwize family systems support both 1 GbE and 10 GbE connections and use innovative Bridgeworks SANSlide technology to optimize the use of network bandwidth. As a result, the networking infrastructure can require lower speeds (and therefore, lower costs), or users might be able to improve the accuracy of remote data through shorter replication cycles. The remote mirroring functions also support VMware vCenter Site Recovery Manager to help speed up disaster recovery.

*Volume mirroring* is a simple RAID 1 type function that is designed to allow a volume to remain online even when the storage pool backing it becomes inaccessible. Volume mirroring is designed to protect the volume from storage infrastructure failures by providing the ability to seamlessly mirror between storage pools.

Volume mirroring is provided by a specific volume mirroring function in the I/O stack, and it cannot be manipulated like a FlashCopy or other types of copy volumes. This feature does however provide migration functionality, which can be obtained by splitting the mirrored copy from the source or by using the “migrate to” function. Volume mirroring does not have the ability to control back-end storage mirroring or replication.

## Real-time Compression

*Real-time Compression* is designed to enable storing up to five times<sup>1</sup> as much data in the same physical disk space by compressing data as much as 80%. Unlike other approaches to compression, Real-time Compression can be used with active primary data, such as production databases and email systems, which dramatically expands the range of candidate data that can benefit from compression. Real-time Compression operates immediately while data is written to disk, which means that no space is wasted storing decompressed data waiting for post-processing.

The benefits of using Real-time Compression together with other efficiency technologies are significant and include reduced acquisition cost (because less hardware is required), reduced rack space, and lower power and cooling costs throughout the lifetime of the system. Real-time Compression can significantly enhance the usable capacity of your existing storage systems, extending their useful life even more.

By significantly reducing storage requirements with Real-time Compression, you can keep more information online, use the improved efficiency to reduce storage costs, or achieve a combination of greater capacity and reduced cost. Because Real-time Compression can be applied to a much wider range of data, including primary online data, the benefits of compression with SAN Volume Controller can be much greater than with alternative solutions, resulting in much greater savings. Enhancements to SAN Volume Controller nodes support up to three times the performance with Real-time Compression, enabling even larger configurations to experience compression benefits.

Starting on SAN Volume Controller version 7.1, the concurrent use of Easy Tier and Real-time Compression is supported on the same volume.

**Note:** To use the Real-time Compression feature on 2145-DH8 nodes, additional CPU, memory and compression acceleration cards are required as mentioned in 8.1.3, “SAN Volume Controller hardware options” on page 285.

## 8.2 SAN Volume Controller connectivity to FlashSystem 900

The IBM FlashSystem 900 is an all-flash storage array that provides extreme performance and is capable of sustaining highly demanding throughput and low latency across its FC interfaces. It includes up to 16 ports of 8 Gbps or eight ports of 16 Gbps FC. It also provides enterprise-class reliability, large capacity, and “green” data center power and cooling requirements.

To maximize the performance that you can achieve when deploying the FlashSystem 900 with SAN Volume Controller, carefully consider the assignment and usage of the FC HBA ports on SAN Volume Controller. Specifically, SAN switch zoning, coupled with port masking (a feature introduced in SAN Volume Controller Storage Software Version 7.1), can be used for traffic isolation for various SAN Volume Controller functions, reducing congestion and improving latency.

---

<sup>1</sup> Compression data based on IBM measurements. Compression rates vary by data type and content.

After racking, cabling, and powering on the IBM FlashSystem, you must perform several steps to configure the FlashSystem 900 optimally for use with SAN Volume Controller. The first configuration steps are described in Chapter 3, “Planning” on page 49. Follow the procedures in that chapter to set up your IBM FlashSystem 900. You do not need to create any volumes or hosts at this time because the preferred practices of creating volumes and hosts are described in the following topics.

### 8.2.1 SAN Volume Controller FC cabling to SAN

When deploying a new SAN Volume Controller cluster, important steps are to connect the Fibre Channel (FC) ports correctly and to match the port masking and zoning configuration. Figure 8-3 shows the suggested SAN Volume Controller cabling diagram to SAN for dual redundant fabrics.

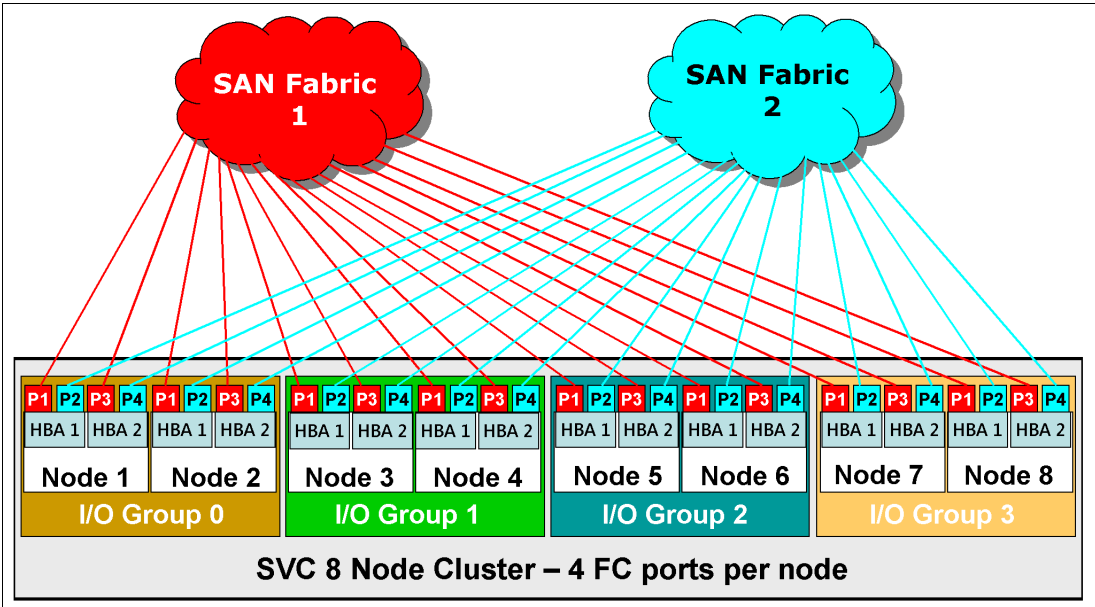


Figure 8-3 Cabling schema for a SAN Volume Controller 8-node cluster with one HBA card

For FlashSystem 900 connectivity, consider using additional HBAs in SAN Volume Controller. Figure 8-4 shows a configuration with one additional SAN Volume Controller HBA in each node, and the cabling schema for cabling to SAN fabric switches.

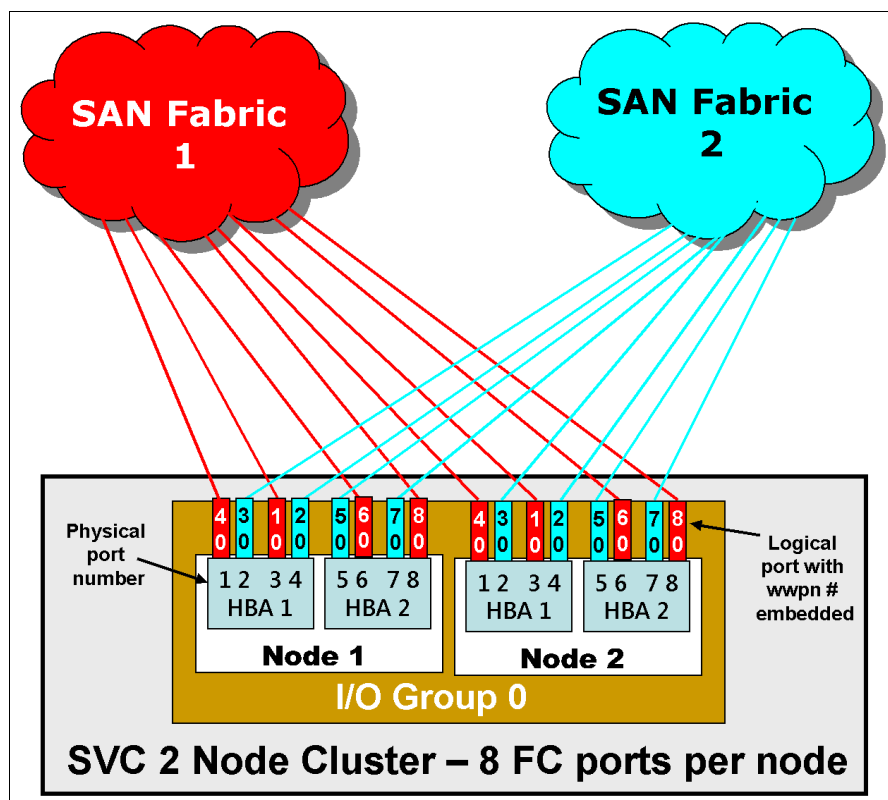


Figure 8-4 Cabling schema for SAN Volume Controller 2-node cluster with two HBA cards per node

## 8.2.2 SAN zoning and port designations

SAN Volume Controller can be configured with two, or up to eight SAN Volume Controller nodes, arranged in a SAN Volume Controller clustered system. These SAN Volume Controller nodes are attached to the SAN fabric, along with disk subsystems and host systems. The SAN fabric is zoned to allow the SAN Volume Controllers to “see” each other’s nodes and the disk subsystems, and for the hosts to “see” the SAN Volume Controllers. The hosts are not able to directly see or operate LUNs on the disk subsystems that are assigned to the SAN Volume Controller system. The SAN Volume Controller nodes within an SAN Volume Controller system must be able to see each other and all of the storage that is assigned to the SAN Volume Controller system.

In an environment where you have a fabric with multiple-speed switches, the preferred practice is to connect the SAN Volume Controller and the disk subsystem to the switch operating at the highest speed. SAN Volume Controller 7.4 with 16 GBps hardware supports 16 GBps, and must be connected to a supported 16 GBps capable switch.

All SAN Volume Controller nodes in the SAN Volume Controller clustered system are connected to the same SANs, and they present volumes to the hosts. These volumes are created from storage pools that are composed of MDisk presented by the disk subsystems.



The zoning capabilities of the SAN switches are used to create three distinct zones:

- ▶ SAN Volume Controller cluster system zones: Create up to two zones per fabric, and include a single port per node, which is designated for intracluster traffic. No more than four ports per node should be allocated to intracluster traffic.
- ▶ Host zones: Create an SAN Volume Controller host zone for each server accessing storage from the SAN Volume Controller system.
- ▶ Storage zones: Create one SAN Volume Controller storage zone for each storage subsystem that is virtualized by the SAN Volume Controller.

Certain limits and restrictions apply for SAN zoning and switch connectivity in SAN Volume Controller 7.4 environments. For more information, see the following web page:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1004924>

For information about supported SAN switches see the IBM storage interoperability matrix:

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

**Note:** The 4 Gbps and 8 Gbps fabrics are not supported when directly connected to the 16 Gbps ports on the node hardware. The 16 Gbps node ports must be connected to a supported 16 Gbps switch.

### 8.2.3 Port designations

The port-to-local node communication is used for mirroring write cache and also metadata exchange between nodes and is critical to the stable operation of the cluster. The DH8 nodes with their 6-port, 8-port, and 12-port configurations provide an opportunity to isolate the port to local node traffic from other cluster traffic on dedicated ports thereby providing a level of protection against misbehaving devices and workloads that can compromise the performance of the shared ports.

Another benefit of isolating remote application traffic on dedicated ports is to ensure that problems impacting the cluster-to-cluster interconnect do not adversely impact ports on the primary cluster and thereby impact the performance of workloads running on the primary cluster.

To isolate both port-to-local and port-to-remote node traffic, be sure to follow the port designations that are listed in Table 8-1 on page 294. Additional port configurations might apply for iSCSI and FCoE connectivity.

Table 8-1 SAN Volume Controller designations

Port	SAN	4-port nodes	8-port nodes	12-port nodes	12-port nodes, write Data Rate > 3 GBps per IO Group
C1P1	A	Host/Storage/ Inter-node	Host/Storage	Host/Storage	Inter-node
C1P2	B	Host/Storage/ Inter-node	Host/Storage	Host/Storage	Inter-node
C1P3	A	Host/Storage/ Inter-node	Host/Storage	Host/Storage	Host/Storage
C1P4	B	Host/Storage/ Inter-node	Host/Storage	Host/Storage	Host/Storage
C2P1	A		Inter-node	Inter-node	Inter-node
C2P2	B		Inter-node	Inter-node	Inter-node
C2P3	A		Replication or Host/Storage	Host/Storage	Host/Storage
C2P4	B		Replication or Host/Storage	Host/Storage	Host/Storage
C5P1	A			Host/Storage	Host/Storage
C5P2	B			Host/Storage	Host/Storage
C5P3	A			Replication or Host/Storage	Replication or Host/Storage
C5P4	B			Replication or Host/Storage	Replication or Host/Storage
localfcportmask		1111	110000	110000	110011

**Notes about Table 8-1:**

- ▶ SAN column assumes an odd/even SAN port configuration. Modifications must be made if other SAN connection schemes are used.
- ▶ Be careful when zoning so that inter-node ports are not used for Host/Storage in the 8-port and 12-port configurations.
- ▶ These options represent optimal configurations based on port assignment to function. Using the same port assignment but different physical locations will not have any significant performance impact in most client environments.

This suggestion provides the traffic isolation that you want while also simplifying migration from existing configurations with only 4 ports, or even later migrating from 8-port or 12-port configurations to configurations with additional ports. More complicated port mapping configurations that spread the port traffic across the adapters are supported and can be considered, but these approaches do not appreciably increase availability of the solution because the mean time between failures (MTBF) of the adapter is not significantly less than that of the non-redundant node components.

Although alternate port mappings that spread traffic across HBAs might allow adapters to come back online following a failure, they will not prevent a node from going offline temporarily to reboot and attempt to isolate the failed adapter and then rejoin the cluster. Our suggestion takes all these considerations into account with a view that the greater complexity might lead to migration challenges in the future and the simpler approach is best.

#### Notes:

- For recent information about the DH8 node as of SAN Volume Controller 7.4, SAN zoning and SAN connections, and port designation recommendations for isolating traffic, see *IBM SAN Volume Controller 2145-DH8 Introduction and Implementation*, SG24-8229 (the chapter about planning and configuration).
- When you attach your IBM FlashSystem to a SAN Volume Controller node that contains a single HBA quad port, follow the zoning and port guidelines that are suggested for any other storage back-end device. See *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

## 8.2.4 Verifying FlashSystem 900 connectivity in SAN Volume Controller

After you activate the zoning, you are to be able to identify the IBM FlashSystem 900 as a controller to SAN Volume Controller. You can use the SAN Volume Controller command **lscontroller** or the SAN Volume Controller GUI to navigate to **Pools** → **External Storage** to verify.

Change the controller name on the SAN Volume Controller system by one of two methods:

- Issue the SAN Volume Controller command **chcontroller** with the **-name** parameter
- Use the SAN Volume Controller GUI to navigate to **Pools** → **External Storage** and then right-click the controller that you want and select **Rename**.

After you change the name, you can easily identify the IBM FlashSystem 900 as a controller to SAN Volume Controller as shown in Figure 8-5.

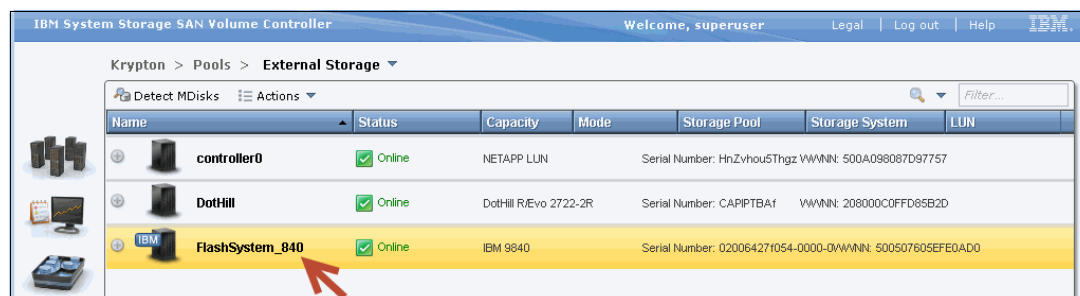


Figure 8-5 IBM FlashSystem 900 as an external storage to SAN Volume Controller

You also need to create zones between SAN Volume Controller and the host. For guidance to configure zoning for host access to SAN Volume Controller, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

### IBM FlashSystem 900 host and volume creation

To provide usable storage (managed disks) to SAN Volume Controller, the IBM FlashSystem 900 must have *volumes* defined. These volumes must be mapped to the SAN Volume Controller *host* that also must be created. For guidance to perform these steps, see Chapter 6, “Using IBM FlashSystem 900” on page 161.

Create one host and name it, for example, SVC. Then, add the SAN Volume Controller worldwide port names (WWPNs) that will communicate with the FlashSystem (in this case, SAN Volume Controller nodes port 1 and port 5) to this newly created host. Then, create the volumes following the guidelines described in the next topic and map them to this newly created host called SVC.

## **SAN Volume Controller managed disk configuration**

The preferred practices and considerations to design and plan the SAN Volume Controller storage pool (MDisk group) setup using the IBM FlashSystem 900 are described. There are several considerations when planning and configuring the MDisks and storage pools for the IBM FlashSystem 900 behind the SAN Volume Controller.

When using the IBM FlashSystem 900 behind the SAN Volume Controller, it is important to note the following points when planning to design and create MDisks for use in storage pools. In this case, the queue assignment and cache assignment are not as relevant as they are with traditional spindle-based disk systems because of the rapid speed with which the IBM FlashSystem 900 is able to process I/O requests.

It is advised to use MDisks in multiples of 4, which is an optimal number for CPU processing, because the MDisks and extents will be equally distributed to the SAN Volume Controller CPU cores.

The second CPU with 32 GB memory feature on SAN Volume Controller Storage Engine Model DH8 provides performance benefit only when Real-time Compression is used. IBM intends to enhance IBM Storwize Family Software for SAN Volume Controller to extend support of this feature to also benefit uncompressed workloads.

### ***Storage pool extent size***

When you work with an all IBM FlashSystem 900 behind a SAN Volume Controller configuration, the extent size can be kept at the default<sup>2</sup> value of 1024 MB (1 GB). The performance of the IBM FlashSystem 900 with random I/O workload does not require the extent size to be smaller. The maximum extent size is 8,192 MB which provides for a maximum Mdisk of 1,024 TB or one petabyte.

If there are existing systems and storage pools, you must have the same extent size for all pools to use transparent volume migration between pools. If you have different extent sizes, you can use volume mirroring to create a copy of the disk and then promote it to the master volume when the copy is completed. However, this manual process takes slightly longer to complete.

### ***All FlashSystem versus mixed storage pools***

If you use the FlashSystem 900 as the primary data storage, add all of the MDisks from the controller to a single *managed disk group* (also known as a *storage pool* in the SAN Volume Controller GUI). However, if more than one FlashSystem 900 is presented to a SAN Volume Controller cluster, a preferred practice is to create a single storage pool per controller.

If you use the FlashSystem 900 with the SAN Volume Controller Easy Tier function, you likely want to create multiple volumes for each *hybrid storage pool*. Create four or more volumes for each hybrid pool, with the combined capacity of these volumes matching the capacity that you want for the *SSD tier* in that pool. More information about SAN Volume Controller Easy Tier with the FlashSystem is in the next topic.

---

<sup>2</sup> Starting on SAN Volume Controller version 7.x

## MDisk mapping, storage pool, and volume creation

This section describes mapping MDisks from an IBM FlashSystem 900 to SAN Volume Controller is defined. Defining a storage pool to accommodate those volumes and the process to provision a volume from the FlashSystem 900 storage pool to the hosts are also explained.

After volumes are created on the IBM FlashSystem 900 and presented to the SAN Volume Controller, they need to be recognized in the SAN Volume Controller. The first step is to click the **Detect MDisks** option from the SAN Volume Controller GUI to detect the newly presented MDisks, as shown in Figure 8-6. Use the following sequence to perform this operation (follow the arrows in Figure 8-6):

1. Select the **Pools** option.
2. Choose **External Storage**.
3. Click the **Detect MDisks** option.
4. When the task completes, click **Close** to see the newly available MDisks.

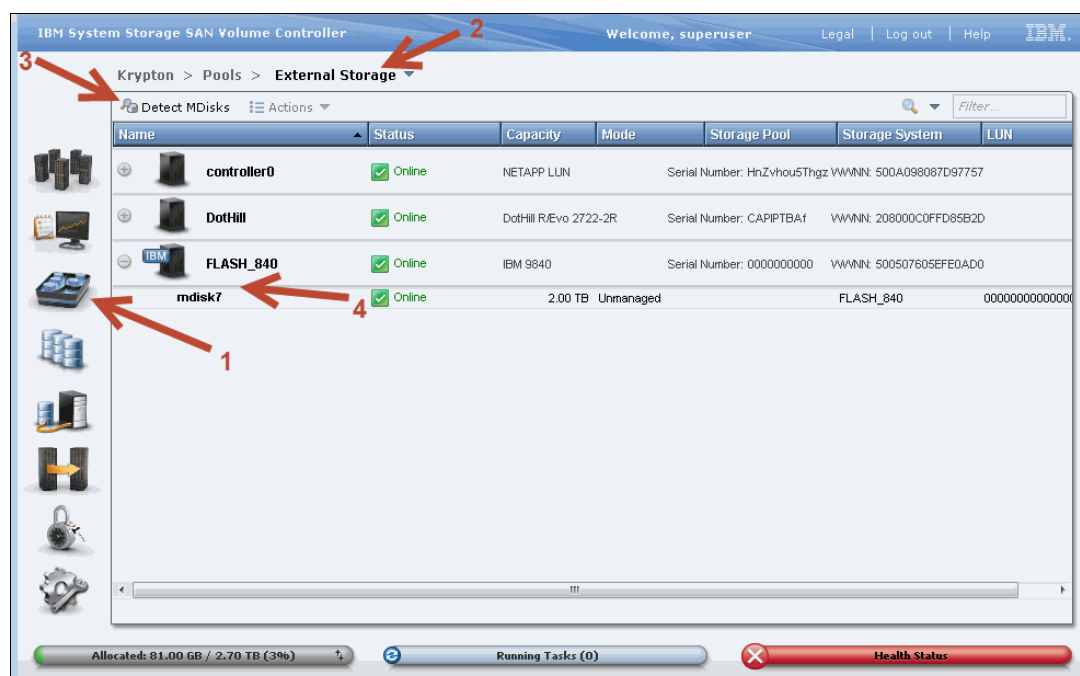


Figure 8-6 Detect MDisks option in the GUI

When this detection completes, a list of MDisks appears, as shown in Figure 8-6 (see *mdisk7*). It is important to rename the MDisks when they are added, for identification purposes. When naming an MDisk, the suggested naming convention is as follows:

%controller name\_lun id on disk system%

This defines the name of the controller from which the LUN is presented and the local identifier that is referenced on the source disk system. This information is helpful when troubleshooting.

To rename an MDisk, select the MDisk that you want, right-click, and select **Rename**. A Rename MDisk window opens. Type the new names of the MDisks and then click **Rename**.

The next step is to create a new storage pool, if you want, as shown in Figure 8-7. A pop-up menu is displayed that you can use to configure the storage pool name and extent size. Ensure that **Advanced Settings** is clicked to show the extent size menu. The default is 1 GB (1024 MB), which can be left as the default, as explained in “SAN Volume Controller managed disk configuration” on page 296. You enter the name of the storage pool, and when complete, click **Next**.

Use the following sequence to perform this operation (follow the arrows in Figure 8-7):

1. Select the **Pools** icon.
2. Choose **MDisks by Pools**.
3. Click **New Pool**.
4. Enter the pool name and extent size that you want.
5. Click **Next** to advance.

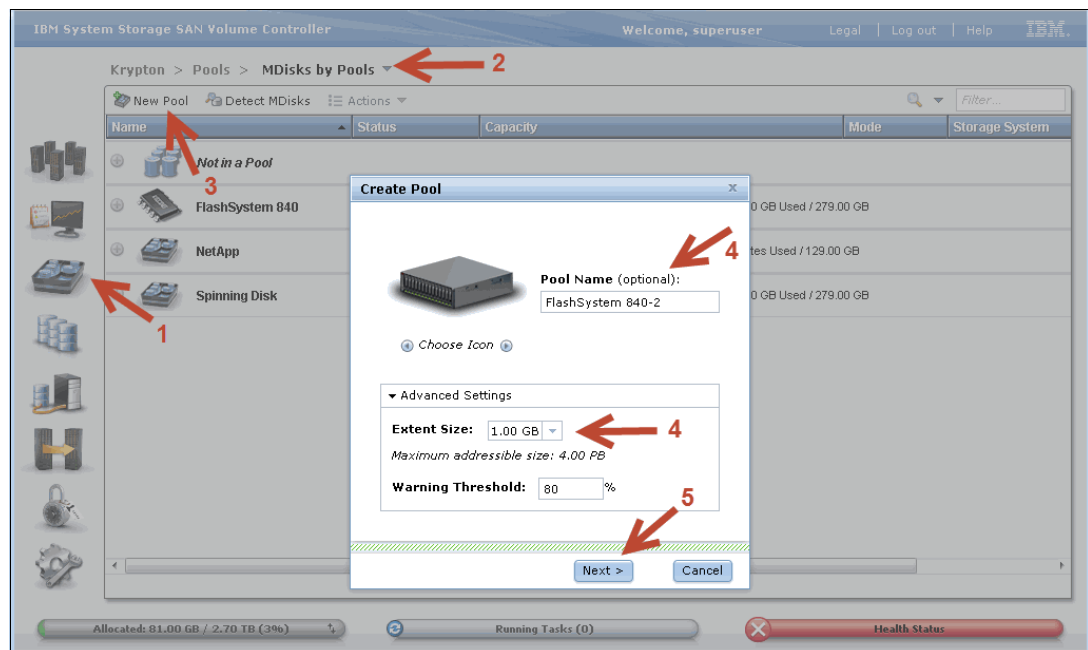


Figure 8-7 The new pool option

The MDisk that will be in this storage pool needs to be added. Select the MDisk to be a member of this group, as shown in Figure 8-8, and then click **Create**.

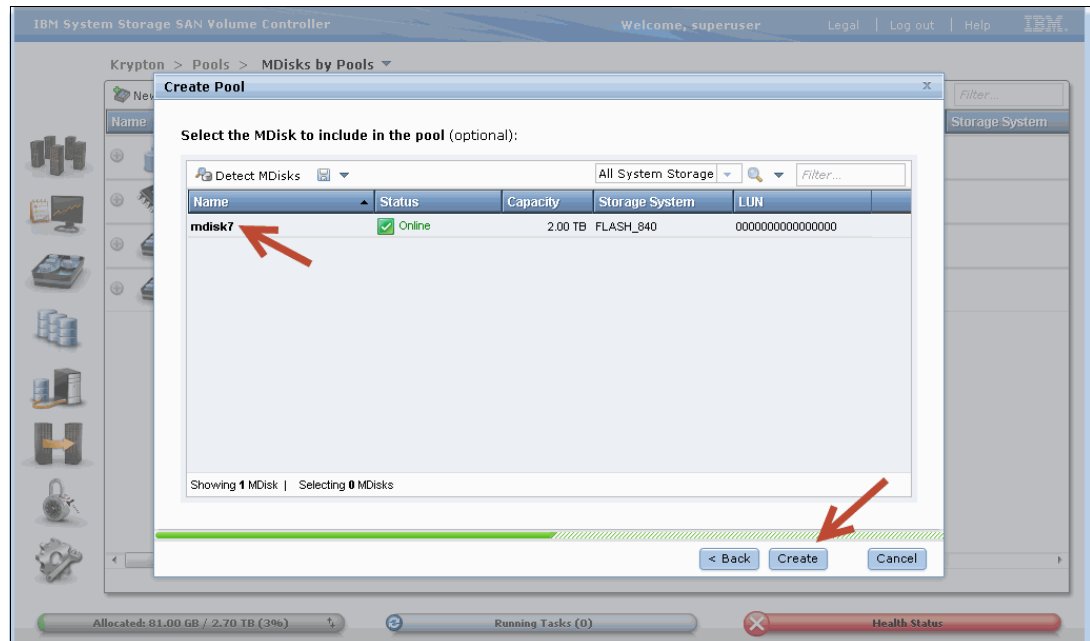


Figure 8-8 Showing the MDisk candidate to include in the storage pool being created

The process begins, as shown in Figure 8-9. Click **Close** when completed to return to the main page.

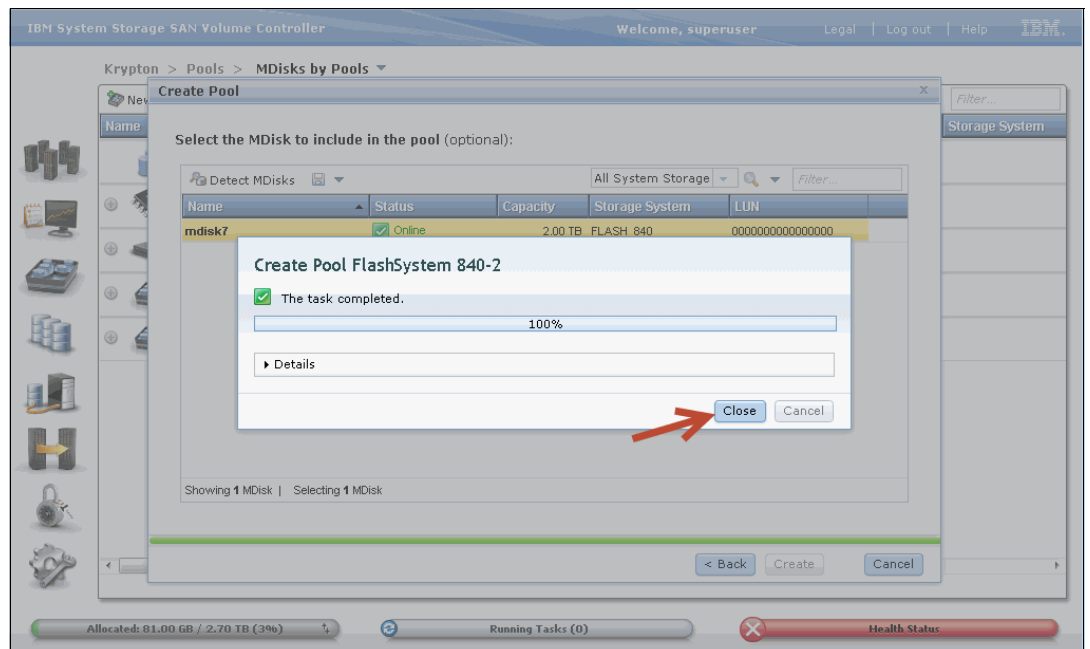


Figure 8-9 Command to create a storage pool being processed

After the command completes, the storage pool is created and the MDisk is a member of the storage pool. To confirm, review the details of the new storage pool just created, as shown in Figure 8-10.

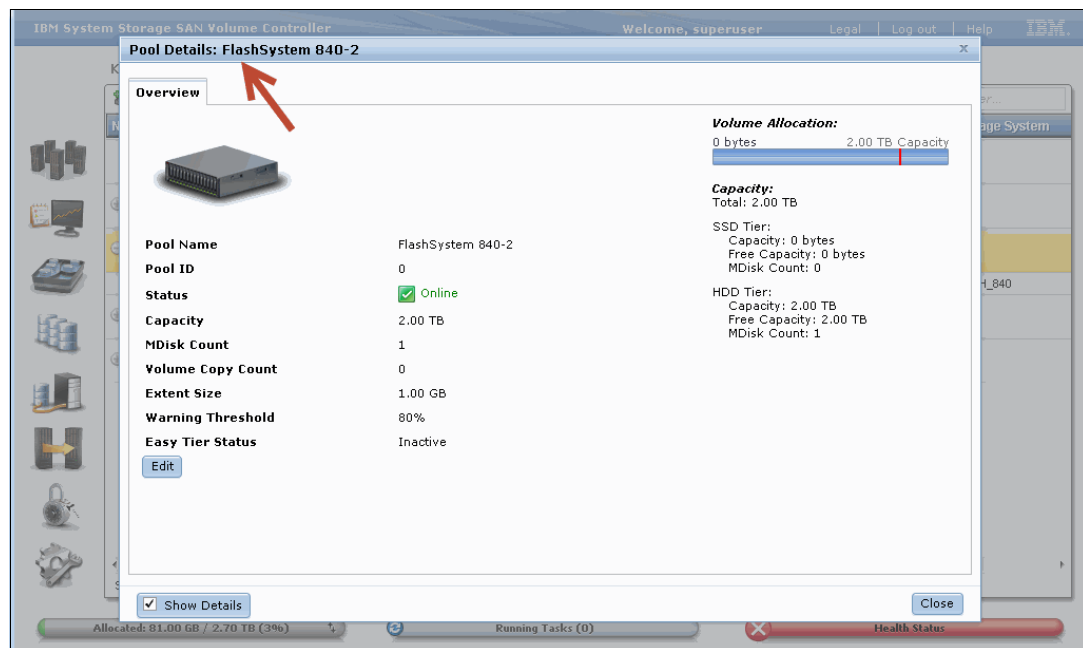


Figure 8-10 Checking storage pool details for the newly created pool

The definition of the MDisk and one storage pool is complete. If a second storage pool is necessary for volume mirroring or for more volume allocation, create the second storage pool by repeating the steps in these figures: Figure 8-7 on page 298, Figure 8-8 on page 299, Figure 8-9 on page 299, and Figure 8-10.

**Suggestion:** Add all MDisk to the defined storage pool at creation time, or before starting volume allocation on that pool. You can select more than one MDisk candidate during the storage pool creation process, or you can add it manually by using the `addmdisk` command (or by using the GUI, also).



Now that you have an available storage pool, you can create a host. Assuming that zoning is already configured between the host and SAN Volume Controller, Figure 8-11 shows the main menu to create a host on SAN Volume Controller. Use the following sequence to perform this operation (follow the arrows in Figure 8-11):

1. Select the **Hosts** icon.
2. Click **Hosts**.
3. Click **New Host**.
4. Select the Fibre Channel port to be added to that host.
5. Click **Add Port to List**.
6. Check the Port Definitions to ensure that you added the port.
7. Click **Create Host** to complete the process.

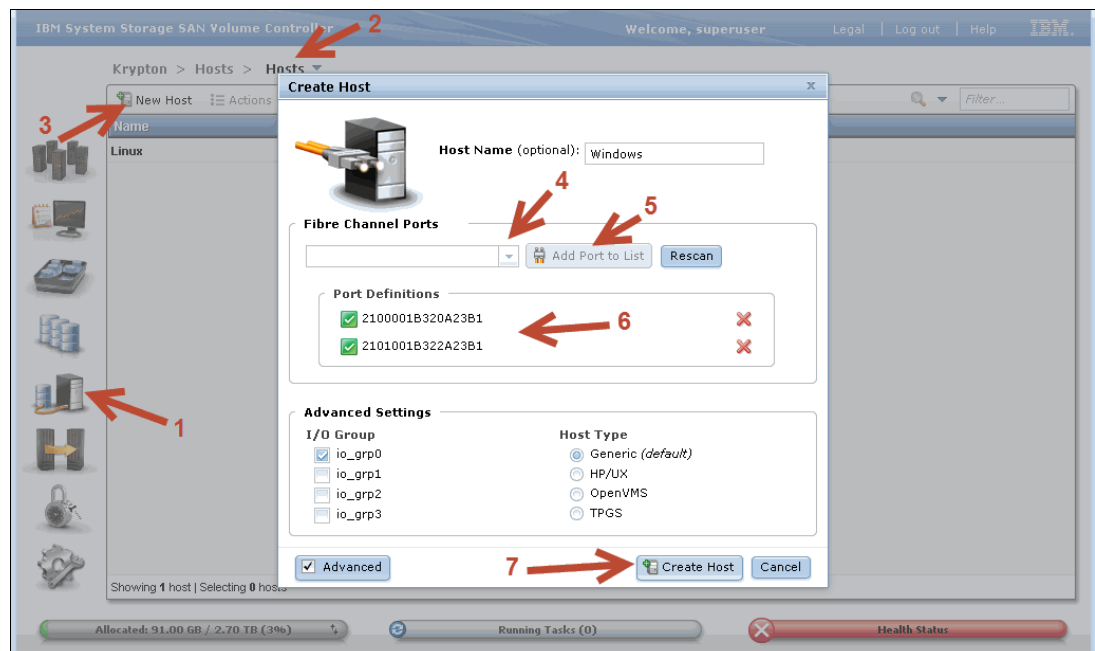


Figure 8-11 The New Host option

After you create the host, you can now create a volume (shown in Figure 8-12). Note that the selected pool is the FlashSystem 900 pool that you created previously. Use the following sequence to perform this operation (follow the arrows in Figure 8-12):

1. Select the **Volumes** icon.
2. Choose **Volumes**.
3. Click the **New Volume** option.
4. Choose a volume preset. In this case, use the Generic icon.
5. Select a pool (in this case, use the FlashSystem pool), and enter the quantity, capacity, and name in the Volume Details section.
6. Click **Create** to complete the process.

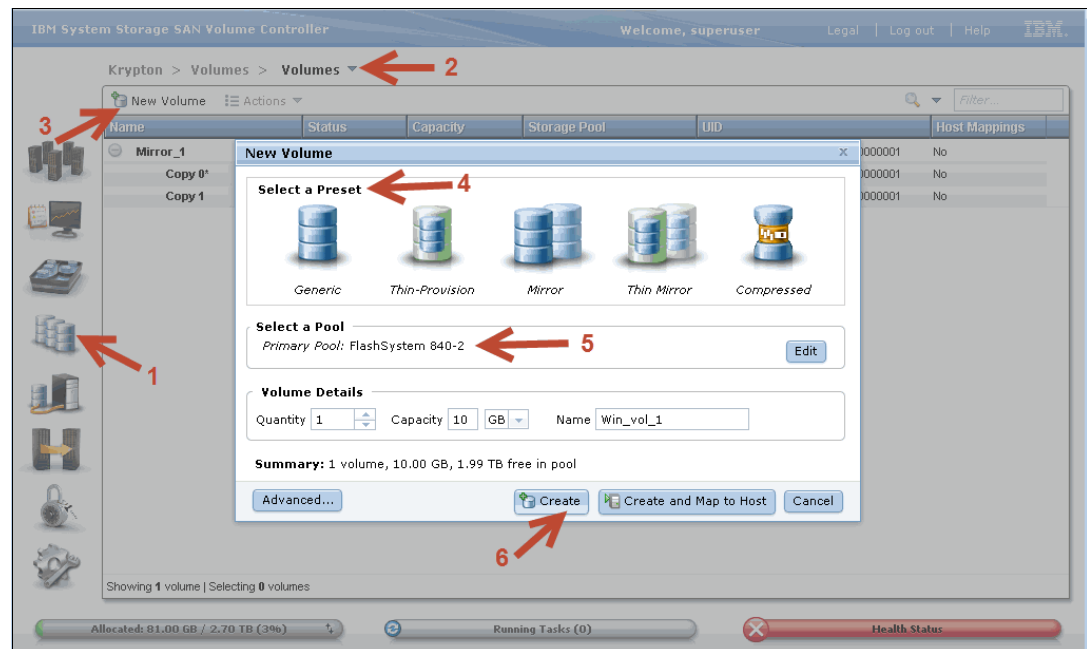


Figure 8-12 The New Volume option

The final step is to assign the recently created volume to the host that is already created. Use the following sequence to perform this operation (follow the arrows in Figure 8-13):

1. Select the **Volumes** icon.
2. Choose **Volumes**.
3. Select the volume that you want to map.
4. Click **Actions**.
5. Select the option **Map to Host**.

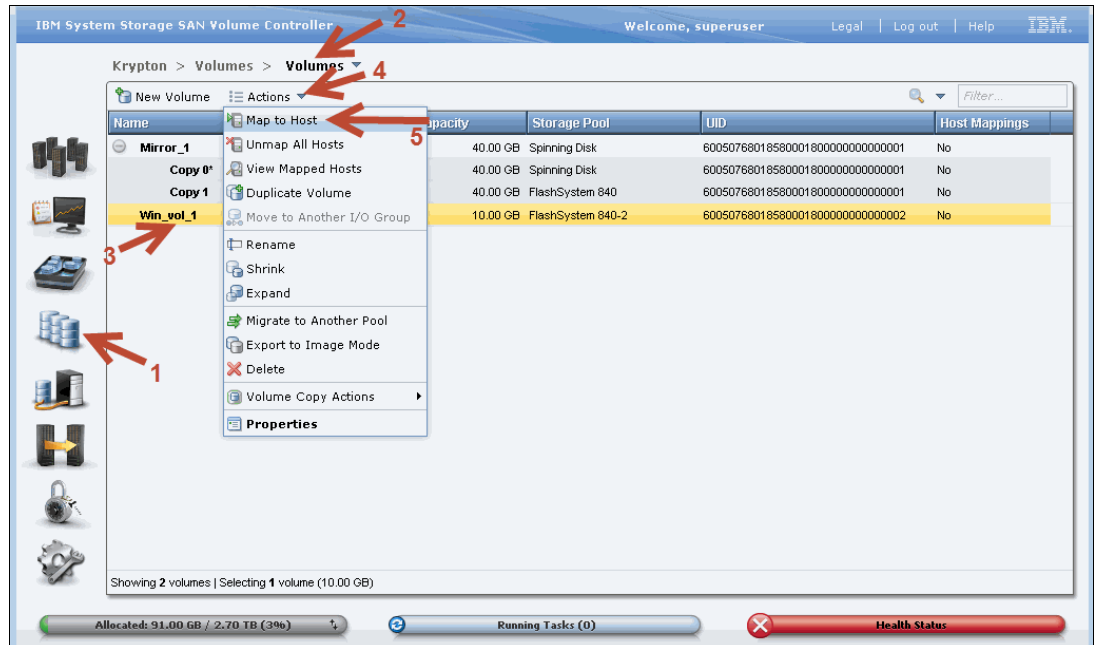


Figure 8-13 The Map to Host option

6. Another pop-up menu is displayed. Select the volumes that you want to map to the host (Figure 8-14). Click **Apply** to complete the assignment.

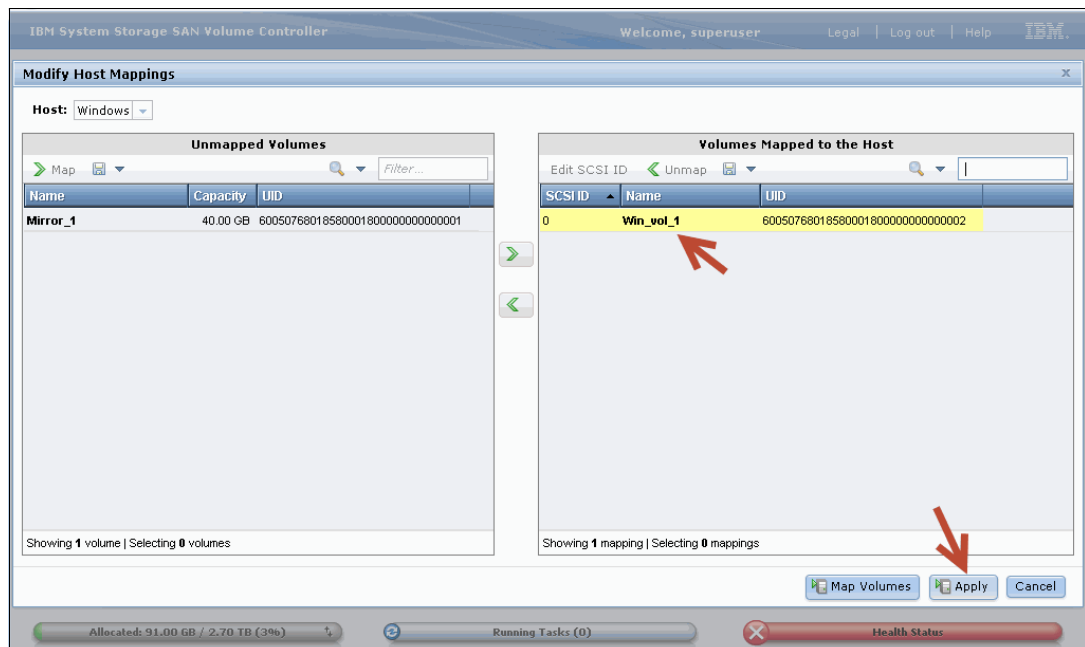


Figure 8-14 The Map to Host steps

The tasks on SAN Volume Controller side are complete. At this point, the server is able to recognize the volume, assuming that the server is already prepared in terms of multipathing software and HBA firmware.

## **SAN Volume Controller volume mirroring**

The SAN Volume Controller *volume mirroring* feature allows the creation of one volume with two copies of its extents. The two data copies, if placed in separate storage pools, allow volume mirroring to eliminate the impact to volume availability if one or more MDisks, or the entire storage pool, fail.

When you design for a highly available SAN Volume Controller and FlashSystem deployment, a good approach is to have each storage pool that is built with MDisks come from separate back-end storage subsystems. This allows the mirrored volumes to have each copy in a different storage pool. In this manner, volume mirroring can provide protection against planned or unplanned storage controller outages because the volume continues to be available for read/write operations from hosts with the surviving copy.

SAN Volume Controller *volume mirroring* is a simple RAID 1-type function that allows a volume to remain online even when the storage pool backing it becomes inaccessible. Volume mirroring is designed to protect the volume from storage infrastructure failures by seamless mirroring between storage pools.

Volume mirroring is provided by a specific volume mirroring function in the I/O stack, and it cannot be manipulated like a FlashCopy or other types of copy volumes. However, this feature provides migration functionality, which can be obtained by splitting the mirrored copy from the source or by using the “migrate to” function. Volume mirroring cannot control back-end storage mirroring or replication.

With volume mirroring, host I/O completes when both copies are written. Before version 6.3.0, this feature took a copy offline when it had an I/O time-out, and then resynchronized with the online copy after it recovered. With 6.3.0, this feature is enhanced with a tunable latency tolerance. This tolerance provides an option to give preference to losing the redundancy between the two copies. This tunable time-out value is Latency or Redundancy.

- ▶ The Latency tuning option, which you set with the **chvdisk -mirrowritepriority latency** command, is the default. This behavior was available in releases before 6.3.0. It prioritizes host I/O latency, which yields a preference to host I/O over availability.
- ▶ However, you might have a need in your environment to give preference to Redundancy when availability is more important than I/O response time. Use the **chvdisk -mirror writepriority redundancy** command.

Regardless of which option you choose, volume mirroring can provide extra protection for your environment.

Migration offers the following options:

- ▶ Volume migration by using *volume mirroring* and then by using *Split into New Volume*  
By using this option, you can use the available RAID 1 functionality. You create two copies of data that initially has a set relationship (one volume with two copies: one primary and one secondary) but then break the relationship (two volumes: both primary and no relationship between them) to make them independent copies of data. You can use this to migrate data between storage pools and devices. You might use this option if you want to move volumes to multiple storage pools. Volume can have two copies at a time, which means you can add only one copy to the original volume and then you have to split those copies to create another copy of the volume.

► Volume migration by using *Move to Another Pool*

By using this option, you can move any volume between storage pools without any interruption to the host access. This option is a quicker version of the previous option (volume mirroring and Split into New Volume). You might use this option if you want to move volumes in a single step or you do not have a volume mirror copy already.

**Note:** Volume mirroring does not create a second volume before you split copies. Volume mirroring adds a second copy of the data under the same volume; as a result, you have one volume presented to the host with two copies of data connected to this volume. Only splitting copies creates another volume and then both volumes have only one copy of the data.

With volume mirroring, you can move data to various MDisk within the same storage pool or move data between different storage pools. There is a benefit of using volume mirroring compared to volume migration: With volume mirroring, storage pools do not need the same extent size as is the case with volume migration.

Starting with firmware 7.3 and the introduction of the new cache architecture, mirrored volume performance is significantly improved. Now, lower cache is beneath the volume mirroring layer, which means each copy has its own cache. This helps in cases of having copies of different types, for example generic and compressed, because now each copy uses its independent cache and each copy can do its own read prefetch now. Destaging of the cache can now be done independently for each copy, so one copy does not affect performance of a second copy.

Also, because the Storwize destage algorithm is MDisk-aware, it can tune or adapt the destaging process, depending on MDisk type and utilization; this can be done for each copy independently.

### **Use cases: SAN Volume Controller volume mirroring with FlashSystem**

Depending on the storage system chosen for each of the two volume mirror copies, certain details must be considered to meet the high performance and low latency expectations of the solution. This section describes usage scenarios for the SAN Volume Controller volume mirroring when virtualizing one or more FlashSystem storage systems, and the potential benefits and preferred practices for each scenario.

#### ***Use case 1: Volume mirroring between two FlashSystem storage systems***

This case is the basic scenario for SAN Volume Controller volume mirroring because both copies reside on the same type of storage systems. In this scenario, there is no expected impact to performance in a failure if the two subsystems are close to the SAN Volume Controller nodes. This scenario indicates a non-stretched cluster setup. In this case, FC link latency is not likely to present an issue, and the default configuration for preferred node, primary copy, and mirroring priority is simplified.

Be careful to ensure a balanced workload between all the available resources (nodes, paths, fabrics, and storage subsystems).

However, if a SAN Volume Controller *Stretched Cluster* architecture is deployed in this scenario, plan carefully to ensure that the link latency between the SAN Volume Controller nodes, the hosts, and the two FlashSystem subsystems does not negatively affect the I/O operations from the applications.

Follow these suggestions, when possible:

- ▶ The preferred node for each mirrored volume needs to be kept at the same site as the subsystem that contains the primary copy of that mirrored volume.
- ▶ The host that is performing the I/O operations to the mirrored volume needs to reside in the same site as the preferred node for that volume.
- ▶ The **-mirrorwritepriority** flag for the mirrored volume must be set to 1 latency if the access to the volume copy across the stretched cluster link represents a significant percentage of the total I/O latency. This can compromise the cache usage. Otherwise, the suggested value of redundancy still applies for every stretched cluster architecture to offer the highest level of data concurrency at both sites for protection against failure in the environment.

**Note:** If the primary purpose of the vdisk copy is data migration, then the mirror write priority should be set to latency. If, however, the primary purpose of the mirror is to provide protection against storage failure, the option should be set to redundancy.

### ***Use case 2: Volume mirroring between a FlashSystem and a non-flash storage system***

In this scenario, usually adopted for cost-reduction reasons, plan to avoid the penalty that is represented by the slowest subsystem to the overall I/O latency. Follow these suggestions:

- ▶ The primary copy of each mirrored volume must be set for the copy that resides in the FlashSystem subsystem so that all the reads are directed to it by SAN Volume Controller. This configuration is commonly referred to as a *preferred read* configuration.
- ▶ If both subsystems are close to the SAN Volume Controller nodes, a *non-Stretched Cluster* setup, the **-mirrorwritepriority** flag for the mirrored volume needs to be set to 1 latency. Therefore, destaging to the volume copy in the slowest subsystem does not introduce a negative impact in the overall write latency and, consequently, to cache usage.

### **Using the FlashSystem with SAN Volume Controller Easy Tier**

In this scenario, the IBM FlashSystem 900 is used with SAN Volume Controller Easy Tier to improve performance on a storage pool. Because of the complexity of this scenario, there are important considerations when you design and plan to add a FlashSystem to an existing environment.

A common question is when to use the FlashSystem storage over internal solid-state drives (SSDs) in SAN Volume Controller. The suggestion is to use the FlashSystem storage when your capacity requirements for Easy Tier exceed five SSDs. At this point, consider the FlashSystem storage systems for cost efficiency and performance. For more information, see *Flash or SSD: Why and When to Use IBM FlashSystem*, REDP-5020.

When planning to use the FlashSystem 900 with SAN Volume Controller Easy Tier, first use the IBM Storage Tier Advisor Tool (STAT) to obtain a comprehensive analysis of hot extents. This will allow you to estimate the amount of required FlashSystem capacity. For more information about using this tool, see *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933.

Ensure that you add these MDisks as *SSD MDisks*; otherwise, SAN Volume Controller Easy Tier cannot distinguish between the spindle-based generic hard disk drives (HDDs) and the FlashSystem SSD MDisks. This task can be done by first creating the new MDisks and then change tier level using the **chmdisk -tier ssd <mdisk>** command after adding the MDisks to a storage pool.

## FlashSystem with SAN Volume Controller replication

Consider the following information when you use the IBM FlashSystem 900 with the SAN Volume Controller and replication:

- ▶ Additional latency overhead with synchronous Metro Mirror
- ▶ Distance of cross-site links and additional latency overhead
- ▶ Adequate bandwidth for cross-site links
- ▶ Amount of data to replicate and its I/O rate
- ▶ Dedicated replication ports

The IBM FlashSystem storage systems provide extremely low latency. The latency of Metro Mirror links might affect the FlashSystem storage systems to a greater degree than other traditional disk systems used with SAN Volume Controller. Metro Mirror replication distances in excess of 10 km (6.2 miles) must be carefully analyzed to ensure that they do not introduce bottlenecks or increase response times when used with the FlashSystem storage systems.

**Tip:** Dedicating ports for replication is strongly advised with this solution. Isolating replication ports can disperse congestion and reduce latency for replication, while protecting other ports from the impacts of increased amounts of replication traffic.

FC is the preferred connectivity method using a Dense Wavelength Division Multiplexer (DWDM), or equivalent device, between the source and target sites. Furthermore, the use of inter-switch links (ISLs) in a trunk or port channel to increase the aggregate bandwidth between the two sites is advised. Also, size the connectivity between sites according to the amount of bandwidth that you are allocating on SAN Volume Controller for replication, with additional bandwidth for future growth and peak loads. To summarize, consider the following factors when you design the replication connectivity:

- ▶ Current average and peak write activity to the volumes that you plan to replicate
- ▶ Desired recovery point objective (RPO) and recovery time objective (RTO) values
- ▶ Existing preferred practices for SAN Volume Controller replication (in *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574)

**Note:** When using Fibre Channel over IP (FCIP) for replication in this scenario, consider using at least 10 Gb links between sites.

## 8.2.5 Import and export

Import and export options are helpful if you use SAN Volume Controller as a migration device in the following ways:

- ▶ Import to Image mode

By using this option, you can import an existing storage MDisk or logical unit number (LUN) with its existing data from an external storage system without putting metadata on it so that the existing data remains intact. After you import it, all copy services functions can be used to migrate the storage to other locations while the data remains accessible to your hosts.

- ▶ Export to Image mode

By using this option, you can move storage from managed mode to image mode, which is useful if you are using the SAN Volume Controller as a migration device. For example, vendor A's product cannot communicate with vendor B's product, but you must migrate existing data from vendor A to vendor B. By using Export to image mode, you can migrate data by using Copy Services functions and then return control to the native array while maintaining access to the hosts.

## 8.3 Integration considerations: FlashSystem 900 and SAN Volume Controller

Figure 8-15 shows an example of a usage scenario that combines SAN Volume Controller and the IBM FlashSystem with a tiered approach to the storage management of a storage FlashSystem. In this solution, write I/O is performed to both the FlashSystem and disk storage for both VDisk Mirror copy 1 and copy 2. Read I/O is performed from the FlashSystem to boost performance with microsecond latency.

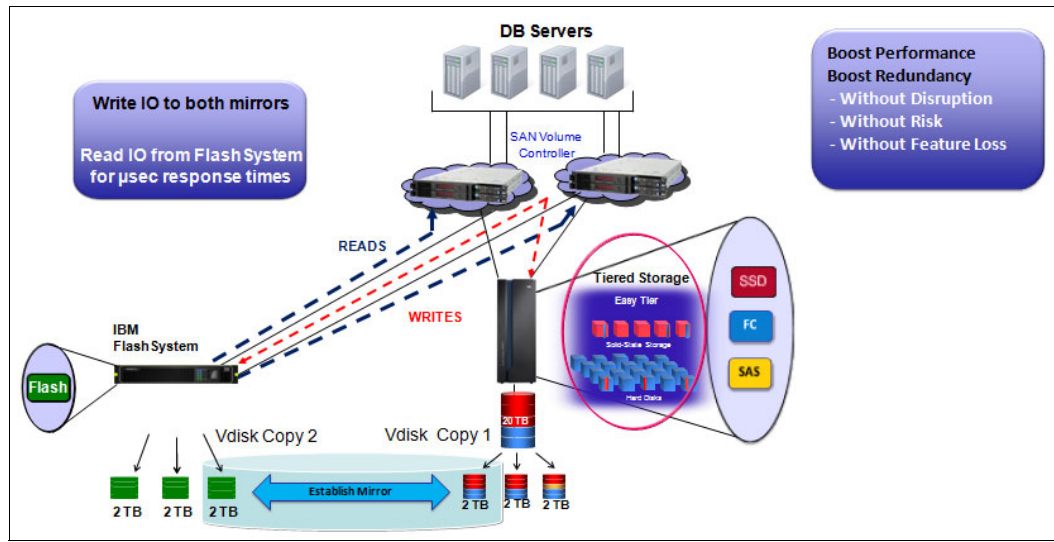


Figure 8-15 IBM FlashSystem and SAN Volume Controller tiered scenario with mirroring

## 8.4 Integration considerations: FlashSystem 900 and IBM Storwize V7000

IBM Storwize V7000 is a virtualized software-defined storage system that can consolidate workloads for simplicity of management, reduced cost, highly scalable capacity, performance, and high availability. This system offers improved efficiency and flexibility with built-in flash storage optimization, thin provisioning, and nondisruptive migration from existing storage.

A mid-range storage system, either a racked IBM Storwize V7000 or a chassis model (IBM Flex System® V7000 Storage Node), provides rich functionality and capacity. When you examine either type of storage node, which includes the usage of Easy Tier, you want to match a certain amount of fast storage with spinning storage. This is typically a 10:1 ratio. For example, if you wanted to achieve 100 TB of capacity with Easy Tier, you can get 90 TB of spinning capacity and 10 TB of flash capacity.

In the Storwize V7000, you can achieve 10 TB of capacity with SSDs in the controller or expansion unit, but if you price the solution, you might find that 10 TB of the FlashSystem 900 are more economical. To gain the maximum benefits in this solution, you can use the Storwize V7000 disk slots for the spinning disk capacity, and use the FlashSystem 900 as the fast tier.

Similar to SAN Volume Controller, the IBM Storwize V7000 offers the same functionalities, such as mirroring, FlashCopy, thin provisioning, Real-time Compression, and remote copy. All suggestions described in SAN Volume Controller section can also be applied to the V7000 integration.



Another similarity with SAN Volume Controller is that you can integrate the FlashSystem 900 with an existing Storwize V7000 environment and a bundled solution: the FlashSystem 900 plus the Storwize V7000. For more information about ordering these bundled solutions, contact IBM services, your IBM representative, or your IBM Business Partner for assistance.

For information about how to deploy the IBM Storwize V7000, see *Implementing the IBM Storwize V7000 V7.2*, SG24-7938 and *Implementing the IBM Storwize V7000 Gen2*, SG24-8244.





## Use cases and solutions

The IBM FlashSystem 900 is the perfect solution for a wide variety of use cases and business needs. The FlashSystem 900 can address I/O issues of applications that require high performance and low latency in relation to storage access.

Knowing how to apply the advantages that this latest technology can offer to IT market scenarios and solutions is important.

This chapter gives an overview of how to take advantage of the benefits of the IBM MicroLatency, macro efficiency, enterprise reliability, and extreme performance. This chapter explains where and how to use the FlashSystem as part of your business solution.

The use cases provide examples of real-world implementations, design solutions, and scenarios. The use cases highlight the benefits that each scenario can offer to clients.

The basics of three solution scenarios are described:

- ▶ Tiering
- ▶ Preferred read
- ▶ Flash only

This chapter covers the following topics:

- ▶ Introduction to the use cases
- ▶ Tiering
- ▶ Preferred read
- ▶ Flash only
- ▶ Comparison

## 9.1 Introduction to the use cases

Certain applications have a natural affinity with the FlashSystem 900. These include applications that have a low tolerance to latency, are I/O per second (IOPS) intense, and need to scale in size and performance. These types of applications are key candidates for the FlashSystem because of their sometimes unique requirements for storage needs.

A set of applications exists that especially benefit from the FlashSystem solutions. The following applications are ideal candidates for the FlashSystem storage:

- ▶ Online transaction processing (OLTP)
- ▶ Online analytical processing (OLAP)
- ▶ Virtual desktop infrastructure (VDI)
- ▶ Cloud-scale infrastructure
- ▶ Computational applications

An entire book might be written about each one of these applications. The objective of this chapter is to briefly describe how the IBM FlashSystem 900 can provide value for the special challenges of these applications.

### Online transaction processing (OLTP)

OLTP applications usually rely on a database that needs to be able to serve the application as fast as possible. In OLTP, parallelism allows the application to run as many transactions as possible at the same time. From a storage perspective, too many parallel I/O requests can be challenging because the parallelism limitation of traditional storage directly related to its number of disks.

Because the FlashSystem 900 is not limited by mechanical disk, there is no longer a limitation to allow the application to reach a new level of parallelism, and to be able to run up to 10x or 20x more transactions in up to one-third of the time, on average.

The FlashSystem 900 allows databases to run more operations, on average up to more than 10x because of the extreme performance. More transactions can occur in less time because of the MicroLatency benefits.

Data can be collected and analyzed to see the FlashSystem benefits. For DB2, performance can be measured by extracting a DB2 MONREPORT. Or, in an Oracle database, you can use an Automatic Workload Repository (AWR) report.

For more information about how to benefit from an IBM FlashSystem in a specific OLTP environment, contact your IBM representative. For more details about running the IBM FlashSystem in OLTP environments, see the IBM Redbooks Solution Guide, *IBM FlashSystem in OLTP Database Environments*, TIPS0973.

### Online analytical processing (OLAP)

OLAP applications are important for the business environment because OLAP is a key player in a wide range of business intelligence (BI) tools. In many industries, OLAP is increasingly taking a more important role.

Because the business environment is more complex, decisions increasingly rely on results from BI tools. BI provides the decision maker the opportunity to have a broader vision of the market. Response time makes a big difference. In OLAP, it is even more evident. A decision made faster can help position you better in the market, but a delayed response might cause a significant negative impact to your business.

OLAP applications use a mathematics model to predict and to calculate behavior, but for that, accessing a large amount of data (*big data*) is necessary; this is where FlashSystem is a “game changer” for organizations. The FlashSystem allows clients to run their analyses and predictions at a new level of speed, helping companies to go to the market faster. It is not only a powerful processor, but the speed of access to big data plays a significant role in OLAP applications.

The FlashSystem 900 with its MicroLatency and extreme performance provides a new competitive advantage to clients.

For more details about running the IBM FlashSystem in OLAP environments, see the IBM Redbooks Solution Guide, *IBM FlashSystem in OLAP Database Environments*, TIPS0974.

## 9.2 Tiering

A *tiering approach*, as used in *IBM Easy Tier*, is a solution that combines functionalities that can add value to other storage, in combination with the highly advanced key points that the FlashSystem 900 can offer, such as MicroLatency and maximum performance.

The great advantage of the tiering approach is the capability to automatically move the most frequently accessed data to the highest performing storage system. In this case, the FlashSystem 900 is the highest performing storage, and the less frequently accessed data can be moved to slower performing storage, which can be solid-state drive (SSD) based storage or disk-based storage. The next topic discusses the data movement that can be done at block level or at file level.

**Important:** Consider that IBM Easy Tier latency with the FlashSystem is much lower than the majority of the high-end storage devices on the market.

### 9.2.1 Easy Tier or block-level tiering

Easy Tier is the IBM implementation of *block tiering*. Other vendors have other names for the block movement between different tiers of storage. This topic does not discuss the details of how tiering works. The main objective is to explain how to take the advantage of this implementation by using the FlashSystem 900 and to describe which use case best fits this solution.

Some tiering solutions do not offer the possibility of *external tiering*. No external tiering means that the block movement can be done only inside the storage device. Because this is a limited approach and it does not work with the FlashSystem 900, only external tiering is discussed.

The IBM solution for external tier, known as IBM Easy Tier, is in IBM SAN Volume Controller and the IBM Storwize V7000 (V7000). SAN Volume Controller offers excellent synergy with the FlashSystem 900.

SAN Volume Controller and the FlashSystem address the combination of the lowest latency with the highest functionality. The V7000 and the FlashSystem address the best cost-benefit solution. Both provide the lowest latency for clients that use traditional disk array storage and need to increase the performance of their critical applications.

Use of the IBM Easy Tier solution is indicated when a need exists to accelerate general workloads. SAN Volume Controller will have a map of the *hot data* (more frequently accessed) and the *cold data* (less frequently accessed), and it will move the hot data to the FlashSystem and the cold data to the conventional storage. When data that was previously hot becomes cold, it moves from the FlashSystem to other storage. The inverse process occurs when cold data becomes hot (or more frequently accessed) and will be moved from a traditional storage to the FlashSystem (Figure 9-1).

This solution will focus on accelerating and consolidating the storage infrastructure. It might not reach the same lowest latency that a flash-only solution offers, but it can be used to improve the overall performance of the infrastructure, avoiding the acquisition of SSDs. The main requirement in this solution is to add more performance to an application or to multiple applications, especially existing applications whose performance is costly to improve by application tuning. The overall latency that SAN Volume Controller adds to the data path does not compromise the expected latency.

SAN Volume Controller adds other important features to the solution, such as replication at the storage layer, snapshots, and Real-time Compression.

Figure 9-1 illustrates how Easy Tier works with the FlashSystem 900. The more frequently accessed hot data is moved to the FlashSystem 900 and the less frequently accessed cold data is moved back to the slower disk array storage.

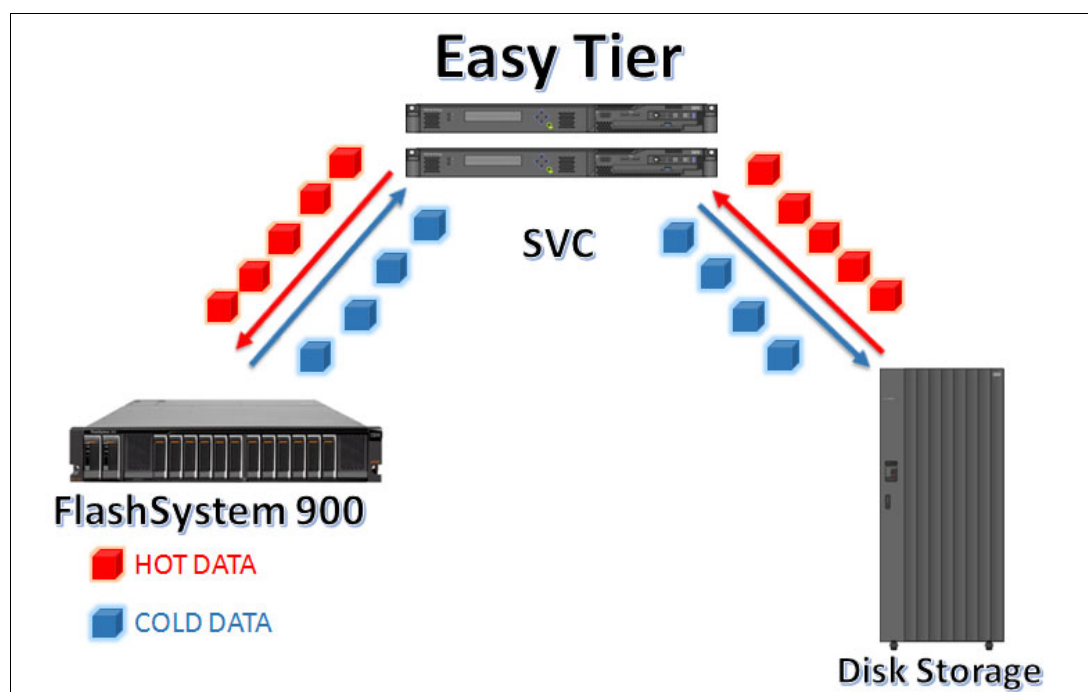


Figure 9-1 Easy Tier moving data between the tiers

Easy Tier with the FlashSystem 900 offers the following benefits to the applications:

- ▶ Capacity efficiency
- ▶ Acceleration to all applications in the same pool
- ▶ Transparent for the application
- ▶ Smart storage determining which is the hot data

## 9.2.2 Information Life Management or file-level tiering

Another tiering approach treats the information at the file level. This type is referred to as *Information Life Management (ILM)*. When using the FlashSystem 900 in an environment where the file is the crucial part of the solution, a preferred practice is to keep the most frequently accessed data in the FlashSystem storage. For more details about ILM, see the ILM fact sheet:

<http://www.ibm.com/services/ch/gts/pdf/br-storage-ilm-factsheet-en.pdf>

Some file systems can move the files between storage pools that contain different kinds of storage, for example, an SSD storage pool and a hard disk drive (HDD) storage pool. In IBM Spectrum Scale (formerly IBM General Parallel File System (GPFS)), you can create policies and move the files between the pools according to a designed policy. Ideally, keep the core files on a faster disk solution, such as the FlashSystem, which provides a significant performance improvement for the application. Also, file systems typically use the concept of *metadata*, which is where the file's descriptive information is located. Considering that the most accessed information in a file system is the metadata, keeping the metadata in a FlashSystem 900 storage pool can significantly improve the overall performance of a file system. IBM Spectrum Scale (formerly GPFS) is considered highly advantageous for the FlashSystem because it can operate in a parallel fashion. For more information, see the IBM Spectrum Scale web page:

<http://www.ibm.com/systems/storage/spectrum/scale/>

## 9.3 Preferred read

A well-known concept among data administrators is that most applications have an I/O profile that has more read events than write events. Traditionally, the majority of databases have an I/O profile of 70/30, which means that the database spends 70% of the time reading and only 30% of the time writing.

This section describes how a combination of the FlashSystem and another disk storage can add redundancy and allow the usage of software functionalities in other storage to provide outstanding performance for the applications with standard I/O behavior.

The following examples illustrate the possible combinations in a *preferred read* implementation and also can be expanded to other applications.

The preferred read solution approach allows the application to use the lowest latency for reads, or read at the speed of the FlashSystem, because the main goal of this solution is to send all read requests to the FlashSystem 900.

Preferred read presumes a *write mirroring* relationship, where the I/O is written to one or more types of storage, and all the read I/O requests are served from the FlashSystem 900.

Figure 9-2 gives a high-level example of how preferred read works.

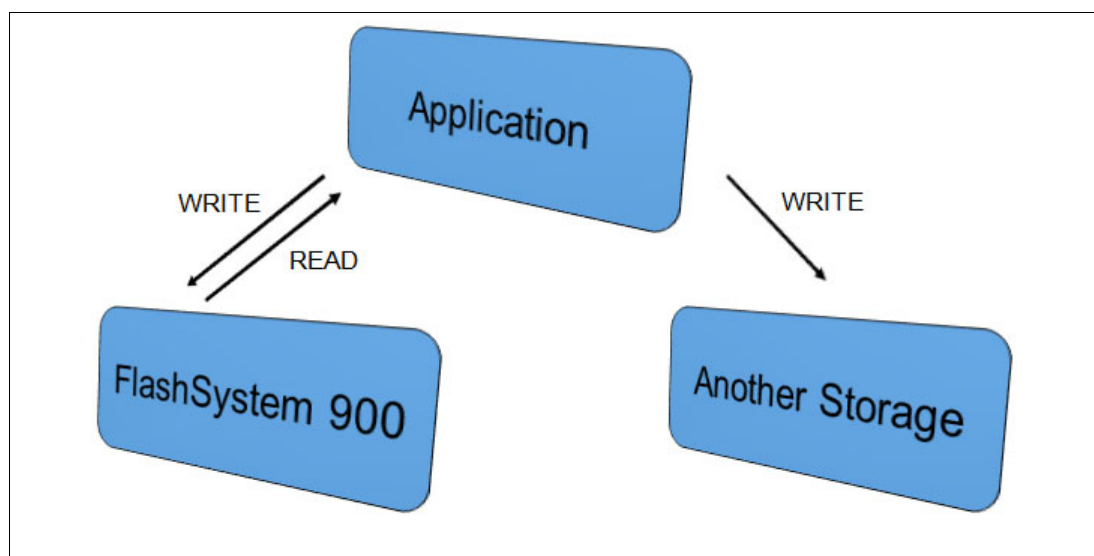


Figure 9-2 Preferred read overview

The preferred read solution adds another level of protection, as Figure 9-3 shows. In a disaster situation, for example, if a power failure occurs on the FlashSystem 900 and it becomes unavailable, the application continues to access the data without interruption and without data loss because, in preferred read, all the data will be mirrored (in both storage products).

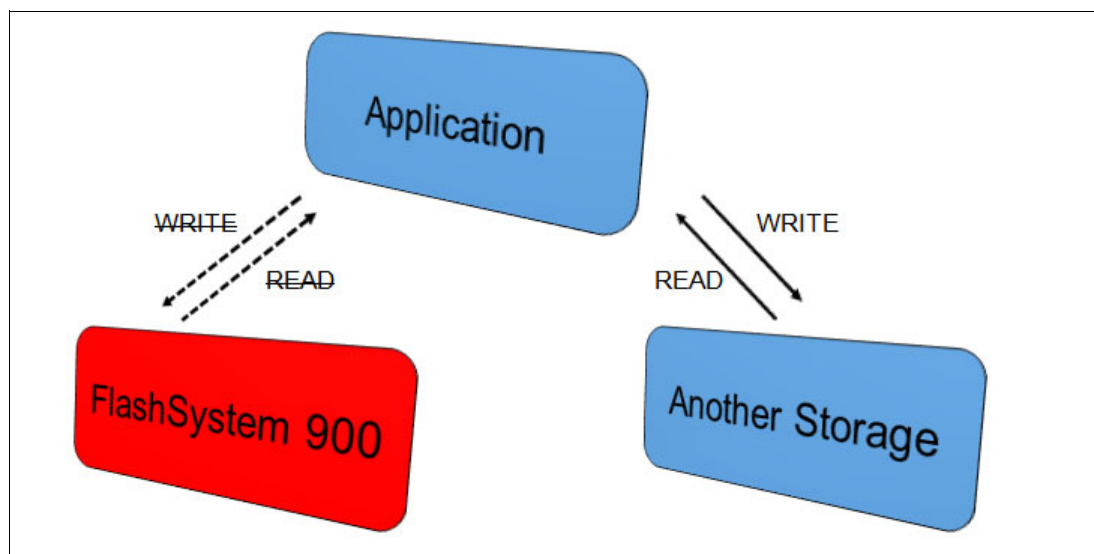


Figure 9-3 Preferred in a disaster situation on the FlashSystem 900

**Important:** For the situation in Figure 9-3, the FlashSystem is the failed component. An important consideration in a disaster recovery plan is that the other storage can also fail. In this case, the replication is stopped. Consider your plan of action in this situation because preferred read continuously works with the FlashSystem. The suggestion is to set an alert in case one of the components in a preferred read relationship stops responding. Include this scenario in the disaster recovery plan.



There is no mirroring between the FlashSystem and other storage. The writes are mirrored, or *split*, which means that every write request goes to both storage systems (the FlashSystem 900 and other storage) that are part of a preferred read relationship. The latency for write requests is always at the latency of the slower storage. You might think that this solution can seriously compromise the overall performance, but it is important to remember that, typically, in a traditional storage system, the writes are done in *cache memory*. So, the majority of the time, write latency is in microseconds because the cache memory has a low response time.

Therefore, the reads performed at the FlashSystem 900 speed (average 70% of the time) and the writes performed with little latency make this solution a good combination of performance and flexibility.

Preferred read can often be implemented on live systems easily and transparently and with no downtime. For more information about how to configure preferred read, see 5.5, “FlashSystem 900 preferred read and configuration examples” on page 140.

The preferred read solution allows the use of features of other storage systems. For example, if a mirroring relationship exists, such as Metro Mirror or Global Mirror, the preferred read does not affect this functionality because the replication will continue sending the write request to the secondary site. This scenario now has three copies of the data: one in the FlashSystem and two copies in the other storage systems.

Figure 9-4 shows how replication can work in a preferred read relationship.

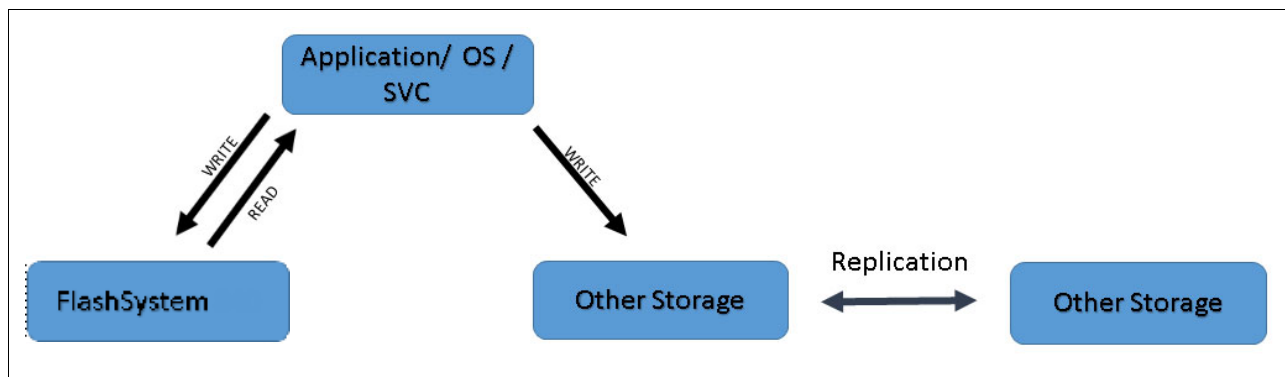


Figure 9-4 Preferred read with replication

**Important:** In the scenario with replication, consider that the write latency will be the latency to write the data on the secondary site and receive the acknowledgment.

Figure 9-5 shows a solution of using Logical Volume Manager (LVM) on site A in a preferred read design. A Metro Mirror relationship exists between the two other disk storage systems, going from site A to site B.

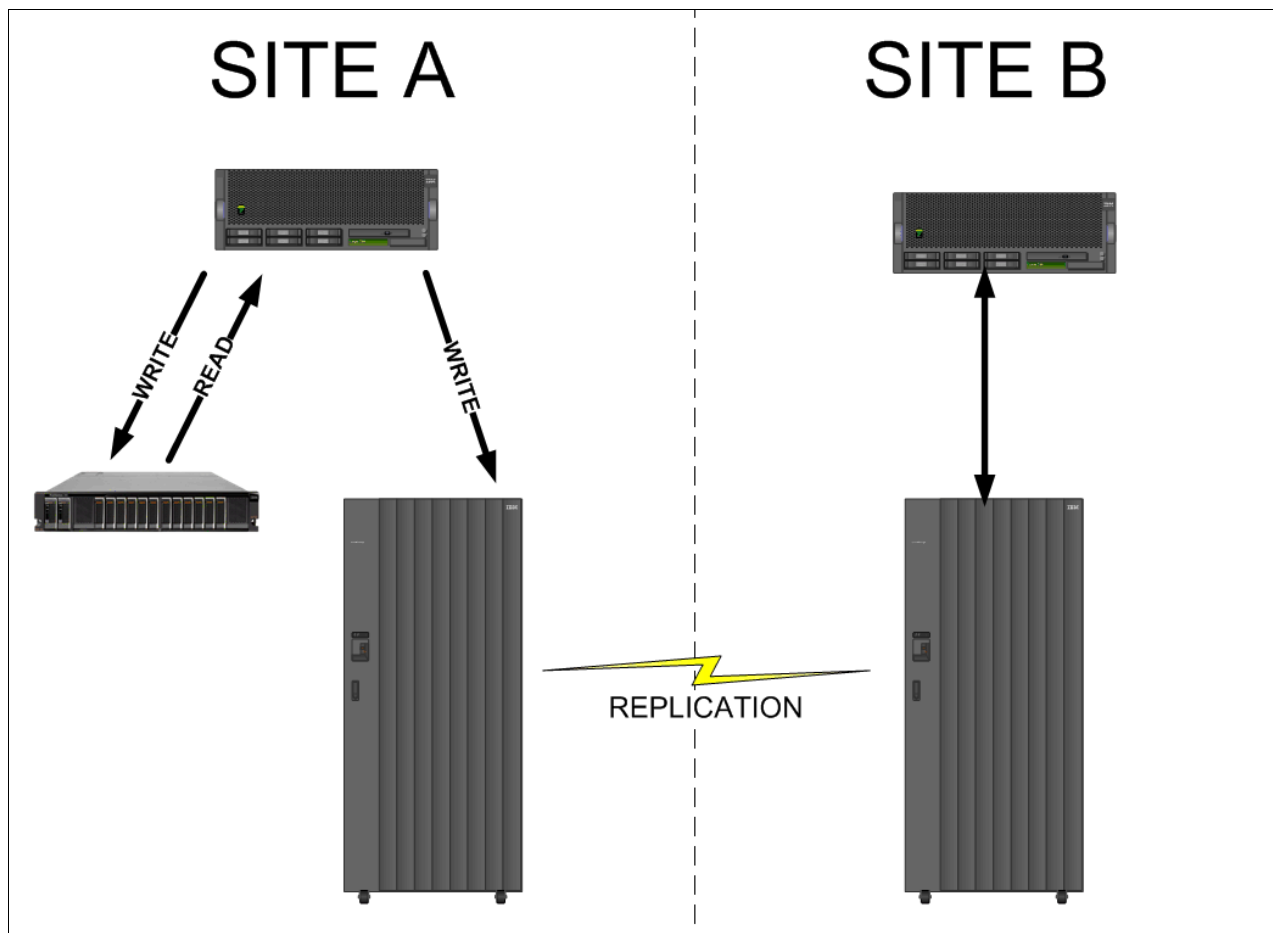


Figure 9-5 Preferred read with Metro Mirror in an IBM DS8000®

### 9.3.1 Implementing preferred read

Further examples of preferred read implementation in some of the more frequently used scenarios are described in this section.

The three ways to implement a preferred read are as follows:

- ▶ SAN Volume Controller and Storwize V7000
- ▶ Application
- ▶ Operating system

For more information about LVM, see 5.5, “FlashSystem 900 preferred read and configuration examples” on page 140.

## Preferred read using SAN Volume Controller and Storwize V7000

IBM SAN Volume Controller and the IBM Storwize V7000 in combination with the FlashSystem 900 add functionality to the solution that is not included with only the FlashSystem 900.

**Note:** In this section, concepts that apply to SAN Volume Controller also apply to the IBM Storwize V7000, because the Storwize V7000 has SAN Volume Controller functionality included as part of its code.

SAN Volume Controller and the IBM Storwize V7000 can be used to implement a preferred read solution with the FlashSystem 900. Preferred read is implemented in SAN Volume Controller by using Volume Mirroring or virtual disk (VDisk) mirroring.

Figure 9-6 shows this scenario. For more information about implementing this feature, see *IBM SAN Volume Controller and IBM FlashSystem 820: Best Practices and Performance Capabilities*, REDP-5027.

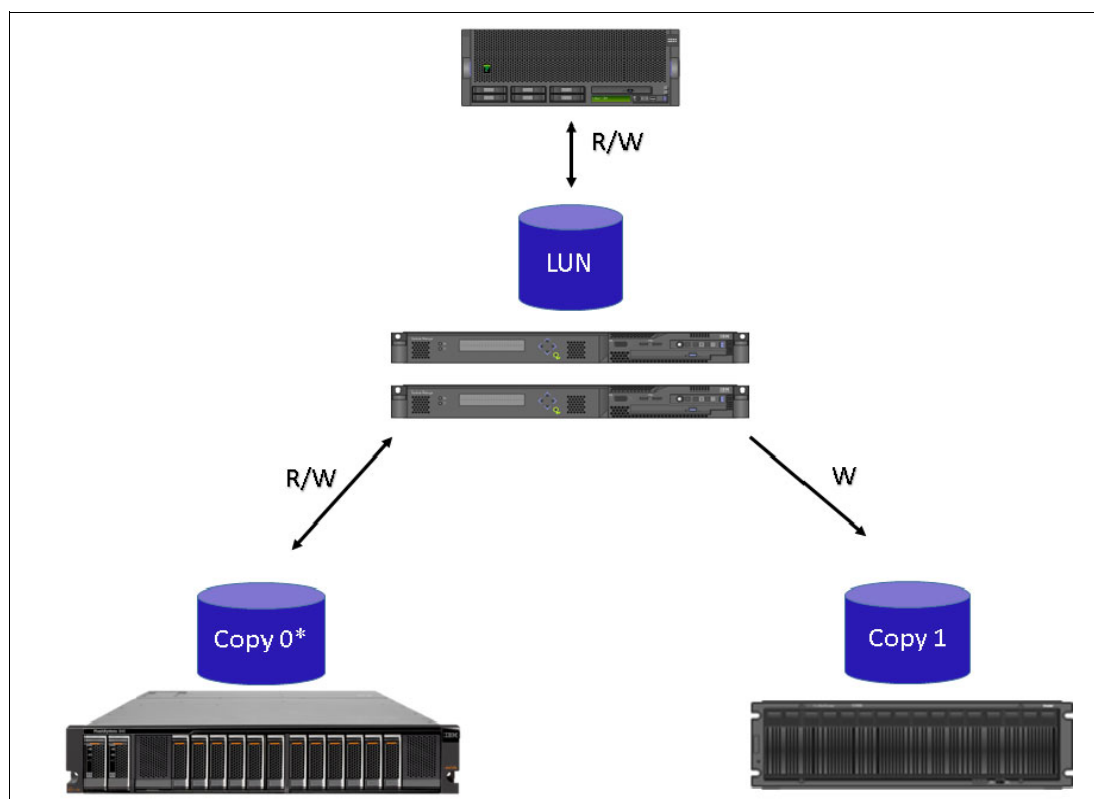


Figure 9-6 Preferred read using SAN Volume Controller

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

- ▶ For SAN Volume Controller functionality and the preferred read:
  - 5.5, “FlashSystem 900 preferred read and configuration examples” on page 140
  - Chapter 8, “Product integration” on page 279
- ▶ For SAN Volume Controller:
  - 8.1, “Running FlashSystem 900 with IBM Spectrum Virtualize - SAN Volume Controller” on page 280

## Preferred read using applications

A wide range of applications can implement mirroring of write I/Os and use a preferred read design to read at the FlashSystem speed. Because the applications can be designed and implemented according to the application developer, see the documentation for your application and search by preferred read implementation, split I/O, or I/O mirroring.

For instructions about how to implement Automatic Storage Management (ASM) preferred read on Oracle environments, see 5.5, “FlashSystem 900 preferred read and configuration examples” on page 140.

**Note:** Always see the most recent documentation for your application for the latest supported configurations for your environment.

## Preferred read using the operating system

Many operating systems, such as Linux, IBM AIX, and Microsoft Windows, have built-in capability to split the I/O between more than one volume by performing a volume mirror at the operating system (OS) level. This kind of functionality usually allows the administrator to decide where the OS will read the data. Other operating systems perform the decision based on the speed of the volumes, using the faster path to storage, in this case, the FlashSystem.

Figure 9-7 shows LVM from AIX controlling the disk and performing a volume mirror at the OS level.

Linux LVM can also be used in this relationship. In the Solaris OS, Solaris Volume Manager (SVM) can be used. In Windows environments, you must use Veritas File System to handle the mirroring of the I/O.

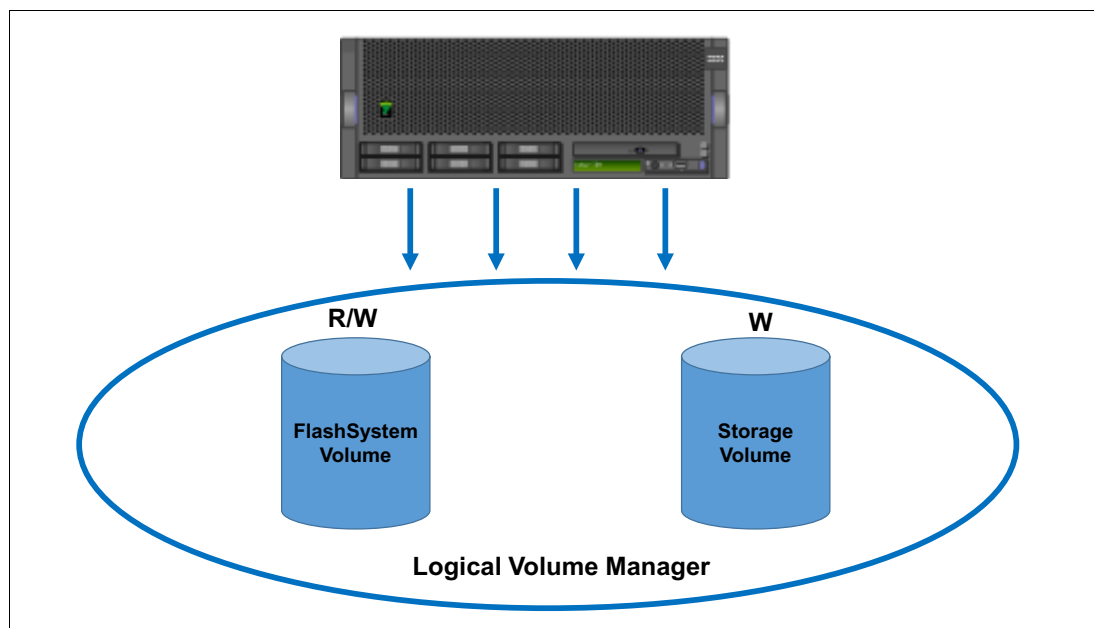


Figure 9-7 Preferred read performed by LVM in an AIX environment

## 9.4 Flash only

This solution design is also known as *manual placement*. Work with the FlashSystem as a dedicated approach. Either place the entire application or the most important part of the application inside the FlashSystem 900. This way, the applications benefit the most from the extreme performance and MicroLatency.

Certain databases have components that when the speed of components increases, they accelerate the application without needing to accelerate the data.

Because the FlashSystem 900 has the potential to be configured with up to 48 TB (RAID 0) of capacity, usually a more efficient way is to place the entire application in the FlashSystem 900 and let the application accelerate at the maximum possible speed.

Figure 9-8 is an example of where the entire database is placed inside the FlashSystem 900.

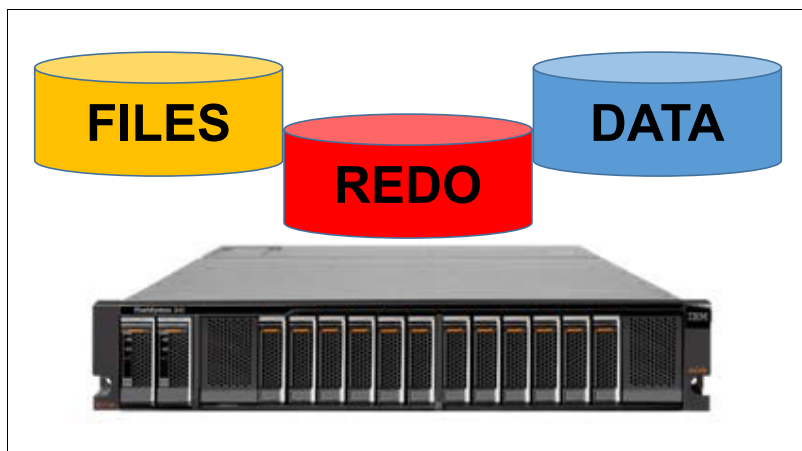


Figure 9-8 Manual data placement example

## 9.5 Comparison

Various solution scenarios were described. The advantages of each were highlighted to help you decide on the best solution approach to your specific application and business environment.

All applications have unique functionality, which you must consider when you design a solution to use the FlashSystem 900. When you compare solutions, consider these items:

- ▶ Number of I/O per second (IOPS)
- ▶ Flexibility
- ▶ Latency

The IBM FlashSystem 900 adds the following benefits to the solution:

- ▶ Extreme performance
- ▶ MicroLatency
- ▶ Macro efficiency
- ▶ Enterprise reliability

In some situations, flexibility might be a priority over latency. In other situations, many IOPS and latency are critical and there is less need for flexibility.

Table 9-1 compares the benefits of each type of solution.

Table 9-1 Solution benefits comparison

Solution scenario	Number of IOPS	Flexibility	Latency
Tiering	High	Very high	Low
Preferred read	Very high	High	Very low
Flash-only or manual placement	Extremely high	Medium	Lowest possible

Figure 9-9 shows the advantages of each solution described in Table 9-1.

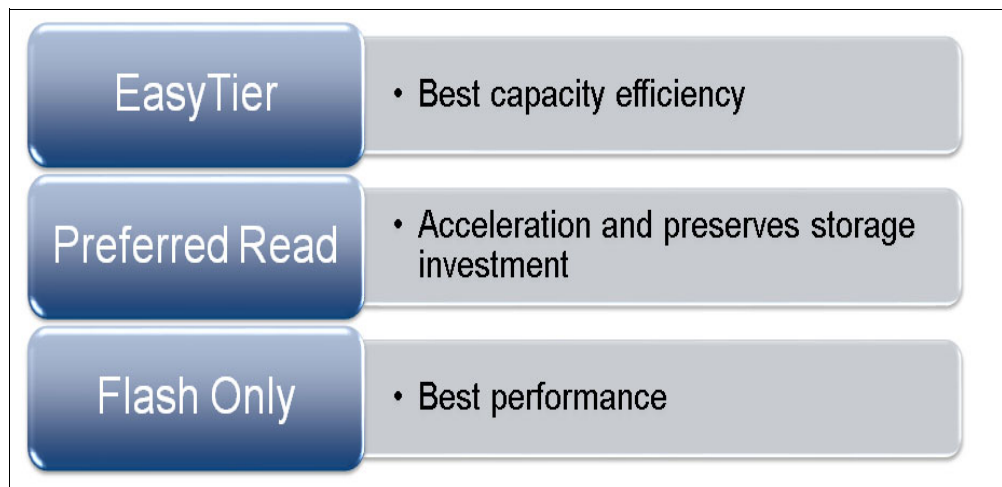


Figure 9-9 Highlights of the benefits that each solution can provide

When performance is important, but you need flexibility, use an IBM Easy Tier solution, because all of SAN Volume Controller features can be used and the administrator does not need to worry about controlling which data is hot.

When every microsecond is important, a flash only solution is the best approach. Systems that need all the capacity of the extreme performance and the MicroLatency typically handles the application flexibility on another level, not on the storage level.

When a balance between the maximum possible performance and some level of flexibility is required, a preferred read solution is the best option. It helps you preserve your investment and offers good flexibility.

To further help you understand the suggestions and options when you are designing a solution, you can also engage the IBM client representative to help you to find the best solution for your application.



## Hints and tips

This chapter provides information about enabling or changing encryption, the importance of setting up the call home facility, Fibre Channel (FC) attachment, hints, and tips for setting up a test environment for the IBM FlashSystem 900 (FlashSystem 900). The chapter offers guidance to help you understand basic performance and troubleshooting information.

This chapter covers the following topics:

- ▶ Encryption hints
- ▶ Setting up call home
- ▶ System check
- ▶ Host attachment hints
- ▶ General guidelines for testing a specific configuration
- ▶ Troubleshooting
- ▶ IBM System Storage Interoperation Center (SSIC)

## 10.1 Encryption hints

You can enable and disable encryption. On a system with enabled encryption, you can change the encryption access keys (rekey).

- ▶ Enabling or disabling encryption

You can enable or disable encryption after the FlashSystem 900 is initialized. The encryption license is required before you enable encryption. The encryption can be enabled on an encryption-disabled FlashSystem 900 Storage Enclosure while it is running. This is a non-destructive procedure that does not impact customer data.

- ▶ Encryption rekey

The encryption access keys can be changed so that the old encryption access key no longer functions, and a new key is required to unlock the system. Rekeying can be performed on an encryption-enabled FlashSystem 900 while it is running. This is a non-destructive procedure that does not impact customer data.

Both operations can be accomplished either on the FlashSystem 900 GUI display (through the **Settings** → **Security** → **Encryption** option) or through the command-line interface (CLI), with the **lsencryption** command to display the settings and **chencryption** to alter the settings.

**Attention:** If you use the CLI commands for these changes, be sure to contact IBM Support for assistance with this procedure. Mistakes made with these commands can render the data permanently inaccessible.

## 10.2 Setting up call home

FlashSystem 900 has a call home facility that alerts IBM Support if a failure occurs in the machine. The call home data is sent to IBM, in the form of an email with the error information and failure codes included.

The installation instructions guide you through setting up the call home facility. After you set up the facility and your system is connected to the IP network, be sure to test the call home facility.

To verify or alter any of the call home settings after the installation is complete, use these steps:

1. From the FlashSystem 900 main GUI window select **Settings** → **Notifications**.
2. In the next panel that opens, select **Email** (Figure 10-1 on page 325). The Email panel lists the settings for email servers, call home notification address, and SMTP server. These settings can all be verified and altered here as required.



Notifications

Email

SNMP

Syslog

Email

The support user receives call home events. Local users also receive event notifications.

Edit

Disable Notifications

Email Servers

IP Address

10.10.10.10

Server Port

25

Call Home

Email Address

flash-sc2@vnet.ibm.com

Event Type

Alerts

Inventory

☒

Test

Email Notifications

Email Address

monitor@ibm.com

Event Type

Alerts

Inventory

☐

Test

Email Contact

\* Contact Name

John Doe

\* Email Reply Address

jdoe@ibm.com

\* Machine Location

third floor, room 304

\* Telephone (Primary)

1234567890

Telephone (Alternate)

\* Required

Figure 10-1 Notifications: Email call home settings

To further enhance the call home importance, the installation process now has a reminder message to complete the set-up. This message is displayed during installation, if the call home function is not enabled. If you decide to disable the facility at any time in the future, a warning message is displayed to stress the importance of call home.

## 10.3 System check

Before investigating performance problems, be sure that the health status of the FlashSystem 900 environment is good. You must look at these components:

- ▶ Fibre Channel connections
- ▶ Switch level connection
- ▶ Host level connections

You can use the FlashSystem 900 GUI to access the event messages by selecting **Monitoring** → **Events** and checking for errors. You can also use the `lseventlog` CLI command. Make sure that no errors exist.

### 10.3.1 Checking the Fibre Channel connections

FlashSystem 900 may be attached with up to sixteen 8 Gbps or eight 16 Gbps adapters. Example 10-1 on page 326 shows the FlashSystem 900 `lsfabric840` command that lists the local and remote WWPNs. In this example, the system is attached to an IBM San Volume Controller.

#### Example 10-1 FlashSystem 900 lsfabric840 information

```
IBM_Flashsystem:FlashSystem-900-03:superuser>lsfabric840
remote_uid      local_uid      canister_id adapter_id port_id node_id node_name type
500507680C120951 500507605E800E41 1          1          1          1      node1     fc
500507680C120951 500507605E800E51 1          2          1          1      node1     fc
500507680C120951 500507605E800E61 2          1          1          2      node2     fc
500507680C120951 500507605E800E71 2          2          1          2      node2     fc
500507680C110951 500507605E800E41 1          1          1          1      node1     fc
500507680C110951 500507605E800E51 1          2          1          1      node1     fc
500507680C110951 500507605E800E61 2          1          1          2      node2     fc
500507680C110951 500507605E800E71 2          2          1          2      node2     fc
500507680140F941 500507605E800E41 1          1          1          1      node1     fc
500507680140F941 500507605E800E51 1          2          1          1      node1     fc
500507680140F941 500507605E800E61 2          1          1          2      node2     fc
500507680140F941 500507605E800E71 2          2          1          2      node2     fc
500507680C220951 500507605E800E42 1          1          2          1      node1     fc
500507680C220951 500507605E800E52 1          2          1          1      node1     fc
500507680C220951 500507605E800E62 2          1          2          2      node2     fc
500507680C220951 500507605E800E72 2          2          2          2      node2     fc
500507680C210951 500507605E800E42 1          1          2          1      node1     fc
500507680C210951 500507605E800E52 1          2          2          1      node1     fc
500507680C210951 500507605E800E62 2          1          2          2      node2     fc
500507680C210951 500507605E800E72 2          2          2          2      node2     fc
500507680110F941 500507605E800E42 1          1          2          1      node1     fc
500507680110F941 500507605E800E52 1          2          1          1      node1     fc
500507680110F941 500507605E800E62 2          1          2          2      node2     fc
500507680110F941 500507605E800E72 2          2          2          2      node2     fc
```

Also check all connections at the switch level and host level.

## 10.4 Host attachment hints

This section lists common host attachment issues and how to resolve them.

### 10.4.1 Fibre Channel link speed

A preferred practice is to set the FlashSystem 900 port link speed and topology to fixed values.

Example 10-2 shows fixed values for a SAN-attached FlashSystem 900. The first port of every Fibre Channel (FC) interface card is connected to a host, and each has a speed that is set to 8 Gbps. The second port of every FC interface card is connected to a switch, and has a speed that is set to 16 Gbps.

#### Example 10-2 SAN-attached FlashSystem 900 with fixed values

```
#
>lsportfc
id canister_id adapter_id port_id type port_speed node_id node_name WWPN nportid status attachment topology
0 1 1 1 fc 8Gb 1 node1 5005076000000041 000001 active host pp
1 1 1 2 fc 16Gb 1 node1 5005076000000042 530500 active switch pp
4 1 2 1 fc 8Gb 1 node1 5005076000000051 000001 active host pp
5 1 2 2 fc 16Gb 1 node1 5005076000000052 570800 active switch pp
8 2 1 1 fc 8Gb 2 node2 5005076000000061 000001 active host pp
9 2 1 2 fc 16Gb 2 node2 5005076000000062 570A00 active switch pp
12 2 2 1 fc 8Gb 2 node2 5005076000000071 000001 active host pp
13 2 2 2 fc 16Gb 2 node2 5005076000000072 530400 active switch pp
```

You can set the speed and topology of the ports by using the **chportfc** command. Example 10-3 shows the setting of a port to point-to-point topology and an 8 Gbps speed.

*Example 10-3 Using the chportfc command to set a port topology and speed*

---

```
>chportfc -topology pp 12
>chportfc -speed 8 12
>chportfc -reset 12
>lsportfc 12
id 12
canister_id 2
adapter_id 2
port_id 1
type fc
port_speed 8Gb
node_id 2
node_name node2
WWPN 5005076000000071
nportid 000001
status active
switch_WWPN 0000000000000000
fpma
vlanid
fcf_MAC
attachment host
topology pp
topology_auto no
speed_auto no
```

---

The use of the command-line interface (CLI) is described in 6.5.3, “Accessing CLI by using PuTTY” on page 223.

## 10.4.2 Host is in a degraded state

A host is shown in a *degraded* state if the host or the port is not connected to both FlashSystem 900 canisters. For details and examples, see 5.2.1, “Fibre Channel SAN attachment” on page 102.

## 10.4.3 FlashSystem port status

You can use the **lsportfc** command to check the status of a FlashSystem 900 FC port. Three other values are possible for the state:

- ▶ **active**  
The port is online and the link to the host/switch is established.
- ▶ **inactive\_configured**  
This port is online and the link to the host/switch is not working. You can check different points to get more information:
  - Cabling
  - Small form-factor pluggable (SFP)
  - Host’s ports
  - Switch for disabled or incorrectly configured ports
  - Link speed and topology of all involved ports

► inactive\_unconfigured

The port is disabled. You can verify that an SFP is installed. Ports might be disabled if an SFP was not present on the controller start-up. Check the logs for error port-related messages.

**Tip:** Another reason for an inactive\_unconfigured port or ports is that you have not created an array yet. Always be sure an array exists before debugging port issues.

Use the `lsportib` or `lsportip` command to check the status of InfiniBand (IB) or iSCSI ports. The output is similar to the `lsportfc` command output. You can also check for a correct IB or iSCSI host setup.

### 10.4.4 AIX multipathing

You must update AIX Object Data Manager (ODM) to recognize the FlashSystem 900 device as a multipath I/O (MPIO) device if FlashSystem 900 is recognized as another disk instead of as an MPIO disk. You can start by using the information at the following website:

<http://www.ibm.com/support/docview.wss?uid=isg1IV50154>

Then, to get the correct ODM driver for your AIX operating system level, follow the references in the line that starts with the following text:

APAR is sysrouted TO

### 10.4.5 Direct attach hints

Always use the latest host driver for direct attachment, especially when connecting with 16 Gb FC. Check the operating system. Certain operating systems currently do not support 16 Gb direct attachment.

**Note:** Fibre Channel over Ethernet (FCoE) does not support direct attachment.

## 10.5 General guidelines for testing a specific configuration

You can evaluate the performance gain when you use FlashSystem 900 by setting up a test environment. This environment can be a dedicated testing environment, or the existing environment. If you use your existing environment, you can use *preferred read* as an unobtrusive way to demonstrate the performance of FlashSystem 900. For more information, see 10.5.2, “Test scenarios” on page 330. Document the test by gathering information before the test, during the test, and at the end of the test.

Follow these steps to document the necessary information:

1. Specify expectations.

You can have a successful test when you set the correct expectations about the performance gain of the test. You can set the correct expectations by analyzing the current environment on the operating system level, or by analyzing the application. For example, the Oracle Automatic Workload Repository (AWR) contains statistics about the Oracle database I/O. The commands for gathering statistical data at the operating system level are in 10.5.5, “Performance data gathering basics” on page 331.

2. Create a baseline.

A *baseline* contains all the information that you will compare after the test to check the success of the test, and to verify whether you fulfilled the expectations. You create the baseline with the current storage environment. The baseline contains values for the latency, I/O per second (IOPS), bandwidth, and run time of batch jobs. This is the first baseline created in the current environment without a FlashSystem.

3. Diagram the current configuration and the testing configuration.

A diagram helps you to visualize the current configuration and the setup of the FlashSystem 900.

4. Set up a test plan.

The test plan contains milestones for every test.

5. Have the correct data.

Ensure that you have the correct data and enough data to run the tests.

6. Create a second baseline in the testing configuration.

The first baseline was created in the configuration without FlashSystem. This second baseline is created in the configuration with FlashSystem. You compare these two baselines to check the success of the test.

7. Document your results.

The result section of your documentation is the most important part.

## 10.5.1 Save the default configuration

Before testing the FlashSystem 900 performance, 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 will assist you in restoring the system to the default settings at the end of the testing. You can save the configuration by using the **svconfig backup** command.

The next 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 a Secure Shell (SSH) client and back up the FlashSystem configuration:

```
superuser>svconfig backup
```

```
.....  
CMMVC6155I SVCCONFIG processing completed successfully
```

2. Be sure to save your default configuration before you create a logical unit number (LUN) or a host.

3. Copy the configuration backup file from the system. Use secure copy to 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 command 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%
```

To use the CLI, see 6.5.3, “Accessing CLI by using PuTTY” on page 223.

## 10.5.2 Test scenarios

Test scenarios are provided that illustrate some of the simplest implementations of the FlashSystem 900, an optimal test scenario, and implementation for preferred read.

The simplest implementations of the FlashSystem 900 include the following functions:

- ▶ Easy Tier
- ▶ Preferred read (FlashSystem 900 is the faster half of mirroring with another type of storage.)
- ▶ Manual tiering (FlashSystem 900 is used as exclusive storage for critical data.)

In test situations, the most important tests are the preferred read and manual tiering variants. An optimal scenario might include these tests:

- ▶ Test the baseline using a standard disk array.
- ▶ Test using a preferred read deployment.
- ▶ Test using a manual tiering deployment.

Implementing preferred read with FlashSystem 900 gives you an easy way to deploy FlashSystem 900 in an existing environment. The data is secured by writing it to two separate storage systems. Data is read at FlashSystem 900 speed because it is always read from FlashSystem 900. This implementation does not change the existing infrastructure concepts, for example, data security, replication, backup, disaster recovery, and so on.

Examples of preferred read configurations are given in 5.5.1, “FlashSystem 900 deployment scenario with preferred read” on page 140.

## 10.5.3 Supported System Configurations

Always check the IBM System Storage Interoperation Center (SSIC) to get the latest information about supported operating systems, hosts, switches, and so on:

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

For information about the prerequisites for the host operating system, see Chapter 5, “IBM FlashSystem 900 client host attachment and implementation” on page 101.

## 10.5.4 Secure erase of data

Some clients, especially in the financial sector, are extremely concerned about data confidentiality. FlashSystem 900 uses encryption to secure data. If you have a license for FlashSystem 900 encryption, you can prevent unauthorized access to FlashSystem data.

**Important:** Deleting FlashSystem 900 encryption key will prevent any access to the data on FlashSystem 900.

The flash cards can be decommissioned by using the **chdrive -task erase** CLI command. The erasure task has a quick mode (approximately 4 minutes) that uses a crypto-erase algorithm and a normal task that erases and overwrites flashcards with patterned data.

## 10.5.5 Performance data gathering basics

All technical resources (application, database, server, and storage) are described. The type of information needed to understand the performance characteristics of the environment is explained.

The focus is performance data gathering from the OS and server perspective, the database perspective, and the storage I/O profile perspective.

The following metrics are used to determine the performance of a storage subsystem:

- ▶ Average read and write IOPS
- ▶ Average read and write bandwidth (MBps)
- ▶ Average latency, which includes wait time and response time
- ▶ Average read/write ratio
- ▶ Average I/O queue depth
- ▶ CPU utilization

The following parameters are optional:

- ▶ Service time (response time minus queue time)
- ▶ Percent (%) busy
- ▶ Average request size
- ▶ Block sizes and block size ratios
- ▶ Average read and write IOPS per block size

You can check for CPU bottlenecks by gathering CPU information with the I/O information.

To evaluate the effects of FlashSystem 900, perform a comparison before and after implementing FlashSystem 900. For an initial performance evaluation or to estimate the performance improvement regarding storage only, three parameters are needed:

- ▶ Average read/write latency
- ▶ Average read/write IOPS
- ▶ CPU usage

More parameters help you to understand the I/O profile of an application.

### Simulating the performance gain with FlashSystem 900

You can simulate the workload by using tools, such as *IOmeter* or *VDBench*. You need a high performance server and FlashSystem 900 to simulate the workload. You take the performance value's average read and write IOPS per block size and the block size ratio to enter these values as a workload description in *IOmeter* or *vdbench*.

When you run this workload pattern in the original environment (the environment in which you gathered the performance data), you see the same IOPS numbers as before. If you now run this workload pattern on a system with FlashSystem, you see a better IOPS number. The ratio of the new IOPS to the original IOPS is the performance gain.

### Server-side information

This section lists operating systems and their various methods to gather performance data. Operating systems have different commands and programs to collect performance data.

To estimate the performance improvement, you gather these parameters during the peak usage of an application. The time frame is 1 - 2 hours.

### **AIX iostat command**

Use the following syntax to create **iostat** information about AIX:

- ▶ `iostat -DRt1 <INTERVAL> <COUNT>`
- ▶ `iostat -T <INTERVAL> <COUNT>`

Example 10-4 shows the commands.

#### *Example 10-4 AIX and iostat*

---

```
# iostat -DRt1 3 2400 > iostat_disk.txt
# iostat -T 3 2400 > iostat_cpu.txt
```

---

Set the values *INTERVAL* and *COUNT* accordingly; an interval of 3 - 5 seconds is enough. Ensure that you start both commands at the same point in time. You can add the **-V** option to the first command to get only the active disks or only the disks that are used. Example 10-5 shows using the **iostat** command with the **-V** option.

#### *Example 10-5 AIX and iostat with the -V option*

---

```
# iostat -DRt1V 3 2400 > iostat_disk.txt
# iostat -T 3 2400 -V > iostat_cpu.txt
```

---

The main information from the output is the average response time and the CPU usage. Look for these values:

- ▶ CPU = % user + % system
- ▶ Await = Average Wait Time per I/O (IOPS/AvgRespTm); Average Response Time
- ▶ rps = The number of read transfers per second (read IOPS)
- ▶ avgserv (read) = The average service time per read transfer  
Different suffixes are used to represent the unit of time. The default is in milliseconds (read latency).
- ▶ \$wps = The number of writes/transfers per second (write IOPS)
- ▶ avgserv (write) = The average service time per read transfer  
Different suffixes are used to represent the unit of time. The default is in milliseconds (write latency).

### **AIX nmon**

The **nmon** command displays local system statistics in interactive mode and records system statistics in recording mode. To collect performance data, use the **nmon** command in recording mode for 1 - 2 hours. You can download the **nmon** tool from the IBM developerWorks® site:

<http://www.ibm.com/developerworks/>

Use the following syntax to create **nmon** information about AIX:

```
# nmon -F <FILENAME> -T -d -A -^ -s <INTERVAL> -c <COUNT>
```

Example 10-6 shows the use of the **nmon** command.

#### *Example 10-6 AIX and nmon*

---

```
# nmon -F /tmp/host1_01122012.nmon -T -d -A -^ -s 3 -c 1200
```

---

Set the values *FILENAME*, *INTERVAL*, and *COUNT* accordingly; an interval of 3 - 5 seconds is enough.



### Windows **perfmon**

The Windows **perfmon** program monitors system activities and resources, such as the CPU, memory, network, and disk. You can use it by starting **perfmon.msc** or **perfmon.exe**.

Table 10-1 lists items that you select in the **perform** program to gather performance values.

Table 10-1 Windows **perfmon** program

Group	Item	Description
Processor	Processor: % Processor Time	CPU time spent
Queue depth	Physical Disk: Avg. Disk Queue Length	Queue length
	Physical Disk: Avg. Disk Write Queue Length	Queue length for writes
	Physical Disk: Avg. Disk Read Queue Length	Queue length for reads
	Physical Disk: Current Disk Queue Length	Current queue length
Block sizes	Physical Disk: Avg. Disk Bytes/Read	Read block size in bytes
	Physical Disk: Avg. Disk Bytes/Write	Write block size in bytes
	Physical Disk: Avg. Disk Bytes/Transfer	R/W block size in bytes
Latency (seconds)	Physical Disk: Avg. Disk Sec/Read	Read latency
	Physical Disk: Avg. Disk Sec/Write	Write latency
	Physical Disk: Avg. Disk Sec/Transfer	Read/write latency
Bandwidth (bytes)	Physical Disk: Disk Read Bytes/sec	Bandwidth (reads)
	Physical Disk: Disk Write Bytes/sec	Bandwidth (writes)
	Physical Disk: Disk Bytes/sec	Bandwidth (total)
IOPS	Physical Disk: Disk Reads/sec	IOPS (reads)
	Physical Disk: Disk Writes/sec	IOPS (writes)
	Physical Disk: Disk Transfers/sec	IOPS (total)

### Citrix **iostat**

Example 10-7 shows how to use the **iostat** command on Citrix.

Example 10-7 Citrix **iostat**

```
# iostat -x 3 1200 > results.txt
```

### Solaris **iostat**

Example 10-8 shows how to use the **iostat** command on Solaris.

Example 10-8 Solaris **iostat**

```
# iostat-xMnz 3 1200 > results.txt
```

## Linux iostat

Example 10-9 shows how to use the **iostat** command for disk and CPU information on Linux.

Example 10-9 Linux iostat

```
# iostat -xkNt 3 1200 > results.txt
# iostat -c 3 1200 > results-cpu.txt
```

Table 10-2 describes the **iostat** return values.

Table 10-2 Description of iostat return values

Item	Description
rrqm/s	The number of read requests merged per second queued to the device.
wrqm/s	The number of write requests merged per second queued to the device.
r/s	The number of read requests issued to the device per second.
w/s	The number of write requests issued to the device per second.
kB/s	The number of kilobytes written to the device per second.
rMB/s	The number of megabytes read from the device per second.
wMB/s	The number of megabytes written to the device per second.
avgrq-sz	Average size (in sectors) of the requests issued to the device.
avgqu-sz	Average queue length of the requests issued to the device.
await	Average time (ms) of I/O requests issued to the device to be served. This includes time spent in the queue and time spent servicing them.
svctm	Average service time (ms) for I/O requests issued to the device. <b>Note:</b> This field will be removed in a future sysstat version of FlashSystem advanced software function.
%util	Percentage of CPU time during which I/O requests were issued to the device (bandwidth utilization for the device).  Device saturation occurs when this value is close to 100%.

## 10.6 Troubleshooting

The information in this section can help you determine, report, and resolve problems.

### 10.6.1 Troubleshooting prerequisites and information to record

Taking advantage of certain configuration options and ensuring that vital system access information is recorded helps ease the process of troubleshooting.

Anyone who has responsibility for managing the system must know how to connect to and log on to the system. This is critical during times when system administrators are not available, for example because of vacation.

**Record the information:** Record the information in the following sections and ensure that authorized people know how to access the information.

## Management IP address

The management IP address connects to the system using the management GUI or starts a session that runs the CLI commands. Record this address and any limitations regarding where it can be accessed from within your Ethernet network.

## Follow power management procedures

Access to your volume data can be lost if you incorrectly power off all or part of a system.

Use the management GUI or the CLI commands to power off a system. Using either of these methods ensures that any volatile data is written to the flash modules and the system is shut down in an orderly manner.

## Set up event notifications

Configure your system to send notifications when a new event is reported.

Correct any issues reported by your system as soon as possible. To avoid monitoring for new events by constantly monitoring the management GUI, configure your system to send notifications when a new event is reported. Select the type of event that you want to be notified about. For example, restrict notifications to only events that require immediate action. Several event notification mechanisms exist:

- Email

An event notification can be sent to one or more email addresses. This mechanism notifies individuals of problems. Individuals can receive notifications wherever they have email access, including mobile devices.

- Simple Network Management Protocol (SNMP)

An SNMP trap report can be sent to a data center management system that consolidates SNMP reports from multiple systems. By using this mechanism, you can monitor your data center from a single workstation.

You can download the FlashSystem 900 management information base (MIB) file from the **Notifications** → **SNMP** GUI page.

- SNMP Agent

An SNMP Server can be enabled on FlashSystem 900. Using this mechanism, SNMP clients can request SNMP data from FlashSystem 900.

- Syslog

A syslog report can be sent to a data center management system that consolidates syslog reports from multiple systems. Using this mechanism, you can monitor your data center from a single workstation.

If your system is within warranty or you have a hardware maintenance agreement, configure your system to send email events to IBM Support if an issue that requires hardware replacement is detected. This mechanism is referred to as *call home*. When this event is received, IBM Support automatically opens a problem report, and if appropriate, contacts you to verify whether replacement parts are required.

If you set up call home to IBM Support, ensure that the contact details that you configure are correct and kept up-to-date as personnel changes.

## **Set up inventory reporting**

Inventory reporting is an extension of the call home email.

Rather than reporting a problem, an email is sent to IBM Support that describes your system hardware and critical configuration information. Object names and other information, such as IP addresses, are not sent. The inventory email is sent on a regular basis. Based on the information that is received, IBM Support can inform you whether the hardware or software that you are using requires an upgrade because of a known issue.

## **Back up your configuration**

The storage system needs to back up your control enclosure configuration data to a file every day. This data is replicated on each control node canister in the system. Download this file regularly to your management workstation to protect the data. This file must be used if there is a serious failure that requires you to restore your system configuration. It is important to back up this file after modifying your system configuration.

## **Resolve alerts in a timely manner**

Your system reports an alert when there is an issue or a potential issue that requires user attention. Perform the recommended actions as quickly as possible after the problem is reported. Your system is designed to be resilient to most single hardware failures. However, if you operate for any period of time with a hardware failure, the possibility increases that a second hardware failure can result in some unavailable volume data.

If there are a number of unfixed alerts, fixing any one alert might become more difficult because of the effects of the other alerts.

## **Keep your software up-to-date**

Check for new code releases and update your code on a regular basis. Use the management GUI or check the IBM Support website to see if new code releases are available.

The release notes provide information about new function in a release, plus any issues that are resolved. Update your code regularly if the release notes indicate an issue to which you might be exposed.

## **Subscribe to support notifications**

Subscribe to support notifications so that you are aware of best practices and issues that might affect your system. Subscribe to support notifications by visiting the IBM Support page on the IBM website. By subscribing, you are informed of new and updated support site information, such as publications, hints and tips, technical notes, product flashes (alerts), and downloads.

## **Know your IBM warranty and maintenance agreement details**

If you have a warranty or maintenance agreement with IBM, know the details that must be supplied when you call for support. Support personnel also ask for your customer number, machine location, contact details, and the details of the problem.

## 10.6.2 User interfaces for servicing your system

Your system provides the following user interfaces to troubleshoot, recover, or maintain your system. The interfaces provide various sets of facilities to help resolve situations that you might encounter.

- ▶ Management graphical user interface (GUI)
- ▶ Command-line interface (CLI)

### Management graphical user interface (GUI)

The management GUI is a browser-based GUI for configuring and managing all aspects of your system. It provides extensive facilities to help troubleshoot and correct problems.

You use the management GUI to manage and service your system. Click **Monitoring** → **Events** for access to problems that must be fixed and maintenance procedures that step you through the process of correcting the problem.

The information on the Events panel can be filtered three ways:

- ▶ Recommended action (default)

This method shows only the alerts that require attention. Alerts are listed in priority order and need to be fixed sequentially by using the available fix procedures. For each problem that is selected, you can perform these tasks:

- Run a fix procedure.
- View the properties.

- ▶ Unfixed messages and alerts

This method displays only the alerts and messages that are not fixed. For each entry that is selected, you can perform these tasks:

- Run a fix procedure.
- Mark an event as fixed.
- Filter the entries to show them by specific minutes, hours, or dates.
- Reset the date filter.
- View the properties.

- ▶ Show all

This method displays all event types, whether they are fixed or unfixed. For each entry that is selected, you can perform these tasks:

- Run a fix procedure.
- Mark an event as fixed.
- Filter the entries to show them by specific minutes, hours, or dates.
- Reset the date filter.
- View the properties.

Some events require a certain number of occurrences in the 25 hours before they are displayed as unfixed. If they do not reach this threshold in 25 hours, they are flagged as expired.

You can also sort events by time or error code. When you sort by error code, the most serious events, those with the lowest numbers, are displayed first. You can select any event that is listed and select **Actions** → **Properties** to view details about the event. You can view the following information:

- ▶ Recommended actions filter. For each problem that is selected, you can do these tasks:
  - Run a fix procedure.
  - View the properties.
- ▶ Event log. For each entry that is selected, you can do these tasks:
  - Run a fix procedure.
  - Mark an event as fixed.
  - Filter the entries to show them by specific minutes, hours, or dates.
  - Reset the date filter.
  - View the properties.

### ***When to use the management GUI***

The management GUI is the primary tool that is used to service your system.

Regularly monitor the status of the system using the management GUI. If you suspect a problem, use the management GUI first to diagnose and resolve the problem.

Use the views that are available in the management GUI to verify the status of the system, the hardware devices, the physical storage, and the available volumes. Click **Monitoring** → **Events** for access to all problems that exist on the system. Use the Recommended Actions filter to display the most important events that need to be resolved.

If there is a service error code for the alert, you can run a fix procedure that assists you in resolving the problem. These fix procedures analyze the system and provide more information about the problem. They suggest actions to take and step you through the actions that automatically manage the system where necessary. Finally, the fix procedure checks that the problem is resolved. If there is an error that is reported, always use the fix procedures within the management GUI to resolve the problem. Always use the fix procedures for both system configuration problems and hardware failures. The fix procedures analyze the system to ensure that the required changes do not cause volumes to be inaccessible to the hosts. The fix procedures automatically perform configuration changes that are required to return the system to its optimum state.

### ***Accessing the management GUI***

This procedure describes how to access the management GUI:

1. Use a supported web browser to point to the management IP address of your system.
2. When the connection is successful, you see a login panel.
3. Log on with your user name and password.
4. Select **Monitoring** → **Events**.
5. Ensure that the events log is filtered by selecting **Recommended Actions**.
6. Select the recommended action and run the fix procedure.
7. Continue to work through the alerts in the order suggested, if possible.

After all the alerts are fixed, check the status of your system to ensure that it is operating as expected.

## Using fix procedures

You can use fix procedures to diagnose and resolve problems with the system.

For example, to repair the system, you might do the following tasks:

- ▶ Analyze the event log.
- ▶ Replace failed components.
- ▶ Verify the status of a repaired device.
- ▶ Restore a device to an operational state in the system.
- ▶ Mark the error as fixed in the event log.

Fix procedures help simplify these tasks by automating as many of the tasks as possible.

The example uses the management GUI to repair a system. Perform the following steps to start the fix procedure:

1. Click **Monitoring** → **Events** and ensure that you are filtering the event log to display the recommended actions.

The list might contain any number of errors that must be repaired. If multiple errors are on the list, the error at the top of the list is the highest priority and must always be fixed first. If you do not fix the higher priority errors first, you might not be able to fix the lower priority errors.

2. Select the error at the top of the list or the subsequent errors to repair.
3. Click **Run Fix Procedure**.

The panel displays the error code and provides a description of the condition.

4. Click **Next** to go forward or **Cancel** to return to the previous panel.
5. One or more panels might be displayed with instructions for you to replace parts or perform other repair activity. If you are not able to complete the actions at this time, click **Cancel** until you return to the previous panel. Click **Cancel** until you are returned to the Next Recommended Actions panel. When you return to the fix procedures, the repair can be restarted from step 1. When the actions that you are instructed to perform are complete, click **OK**. When the last repair action is completed, the procedures might attempt to restore failed devices to the system.

6. After you complete the fix, you see the following statement:

Click OK to mark the error as fixed.

Click **OK**. This action marks the error as fixed in the event log and prevents this instance of the error from being listed again.

7. When you see the following statement, click **Exit**:

The repair has been completed.

If other errors must be fixed, those errors are displayed and the fix procedures continue.

8. If no errors remain, the following statement is displayed:

There are no unfixed errors in the event log.

## Command-line interface (CLI)

Use the CLI to manage a system using the task commands and information commands.

### 10.6.3 Event reporting

Events that are detected are saved in an event log. As soon as an entry is made in this event log, the condition is analyzed. If any service activity is required, a notification is sent.

#### Event reporting process

The following methods are used to notify you and the IBM Support Center of a new event:

- ▶ If you enabled SNMP, an SNMP trap is sent to an SNMP manager that is configured by the client.
- ▶ If enabled, log messages can be forwarded on an IP network by using the syslog protocol.
- ▶ If enabled, event notifications can be forwarded by email by using Simple Mail Transfer Protocol (SMTP).
- ▶ The *call home* feature can be enabled so that critical faults generate a problem management record (PMR) that is then sent directly to the correct IBM Support Center by using email.

#### Understanding events

When a significant change in status is detected, an event is logged in the event log.

Events are classified as either alerts or messages:

- ▶ An *alert* is logged when the event requires an action. Certain alerts have an associated error code that defines the service action that is required. The service actions are automated through the fix procedures. If the alert does not have an error code, the alert represents an unexpected change in the state. This situation must be investigated to see if it is expected or represents a failure. Investigate an alert and resolve it as soon as it is reported.
- ▶ A *message* is logged when a change that is expected is reported, for instance, an array build completes.

#### Viewing the event log

You can view the event log by using the management GUI or the CLI.

You can view the event log by clicking **Monitoring** → **Events** in the management GUI. The event log contains many entries. You can, however, select only the type of information that you need.

You can also view the event log by using the `lseventlog` command.

#### Managing the event log

The event log has a limited size. After it is full, newer entries replace entries that are no longer required. To avoid having a repeated event that fills the event log, certain records in the event log refer to multiple occurrences of the same event. When event log entries are coalesced in this way, the time stamp of the first occurrence and the time stamp of the last occurrence of the problem are saved in the log entry. A count of the number of times that the error condition occurred is also saved in the log entry. Other data refers to the last occurrence of the event.



### ***Describing the fields in the event log***

The event log includes fields with information that you can use to diagnose problems.

Table 10-3 describes several fields that are available to assist you in diagnosing problems.

*Table 10-3 Description of data fields for the event log*

<b>Data field</b>	<b>Description</b>
Event ID	A number that precisely identifies why the event was logged.
Error code	A number that describes the service action that needs to be followed to resolve an error condition. Not all events have error codes that are associated with them. Many event IDs can have the same error code because the service action is the same for all the events.
Sequence number	A number that identifies the event.
Event count	The number of events coalesced into this event log record.
Object type	The object type to which the event log relates.
Object ID	The object ID to which the event log relates.
Fixed	When an alert is shown for an error condition, it indicates whether the reason for the event was resolved. In many cases, the system automatically marks the events fixed when appropriate. There are certain events that must be manually marked as fixed. If the event is a message, this field indicates that you have read and performed the action. The message must be marked as read.
First time	The time when this error event was reported. If events of a similar type are coalesced together so that one event log record represents more than one event, this field is the time that the first error event was logged.
Last time	The time when the last instance of this error event was recorded in the log.
Root sequence number	If set, this number is the sequence number of an event that represents an error that probably caused this event to be reported. Resolve the root event first.
Sense data	Additional data that gives the details of the condition that caused the event to be logged.

### **Event notifications**

Your system can use SNMP traps, syslog messages, emails, and call home notifications to notify you and IBM Remote Technical Support 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 certain events that might occur because of service actions that are being performed. If a recommended service action is active, these events are notified only if they are still unfixed when the service action completes.

Only events recorded in the event log can be notified. Most CLI messages in response to certain CLI commands are not recorded in the event log so they do not cause an event notification.

Table 10-4 describes the levels of event notifications.

Table 10-4 Notification levels

Notification level	Description
Error	<p>Error notification is sent to indicate a problem that must be corrected as soon as possible.</p> <p>This notification indicates a serious problem with the system. For example, the event that is being reported might indicate a loss of redundancy in the system, and it is possible that another failure might result in loss of access to data. The typical reason for sending this type of notification is a hardware failure, but certain configuration errors or fabric errors also are included in this notification level. Error notifications can be configured to be sent as a call home to IBM Remote Technical Support.</p>
Warning	<p>A warning notification is sent to indicate a problem or unexpected condition with the system. Always immediately investigate this type of notification to determine the effect that it might have on your operation, and make any necessary corrections. A warning notification does not require any replacement parts and therefore does not require IBM Support Center involvement. The allocation of notification type Warning does not imply that the event is less serious than an event that has notification level Error.</p>
Information	<p>An informational notification is sent to indicate that an expected event has occurred.</p> <p>No remedial action is required when these notifications are sent.</p> <p>Informational events provide information about the status of an operation. Information events are recorded in the error event log and, depending on the configuration, can be notified through email, SNMP, and syslog.</p>

## Power-on self-test (POST)

A series of tests is performed to check the operation of the components and several of the options that are installed when the system is first turned on. This series of tests is called the *power-on self-test* (POST).

When the code is loaded, additional testing occurs, which ensures that all of the required hardware and code components are installed and functioning correctly.

## Understanding the error codes

Error codes are generated by the event-log analysis and system configuration code.

Error codes help you to identify the cause of a problem, a failing component, and the service actions that might be needed to solve the problem.

## Viewing logs and traces

The system maintains log files and trace files that can be used to manage your system and diagnose problems.

## 10.6.4 Resolving a problem

The management GUI provides extensive facilities to help you troubleshoot and correct problems on your system.

To run the management GUI, start a supported web browser and point it to the management IP address of your system.

After the connection is successful, you see a login panel. Log on using your user name and password. After you log on, select **Monitoring** → **Events** to view the system event log, and then, select the **Recommended Actions** filter.

An event in the log can be an informational message, or it can alert you to an error that requires fixing. Errors are prioritized by their error code, and each has a fix procedure that can be run.

A *fix procedure* is a wizard that helps you troubleshoot and correct the cause of an error. Certain fix procedures will reconfigure the system, based on your responses; ensure that actions are carried out in the correct sequence, and, prevent or mitigate the loss of data. For this reason, you must always run the fix procedure to fix an error, even if the fix might seem obvious.

To run the fix procedure for the error with the highest priority, click **Recommended Action** at the top of the Event page and click **Run This Fix Procedure**. When you fix higher priority events first, the system can often automatically mark lower priority events as fixed.

While the Recommended Actions filter is active, the event list shows only alerts for errors that have not been fixed, sorted in order of priority. The first event in this list is the same as the event displayed in the Recommended Action panel at the top of the Event page of the management GUI.

If correcting errors in a different order is necessary, select an error alert in the event log and then click **Action** → **Run Fix Procedure**.

## 10.7 IBM System Storage Interoperation Center (SSIC)

IBM continuously tests and approves the interoperability of IBM products in different environments. You can search the interoperability results at the IBM System Storage Interoperation Center (SSIC) website.

Check the IBM SSIC website to get the latest information about supported operating systems, hosts, switches, and more:

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

If a configuration that you want is not available at the SSIC, contact your IBM representative about submitting one of these requests to IBM:

- ▶ Solution for Compliance in a Regulated Environment (SCORE)
- ▶ Request for price quotation (RPQ)



# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion 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.

- ▶ *Fabric Resiliency Best Practices*, REDP-4722
- ▶ *Faster DB2 Performance with IBM FlashSystem*, TIPS1041
- ▶ *Flash or SSD: Why and When to Use IBM FlashSystem*, REDP-5020
- ▶ *FlashSystem 900 Product Guide*, TIPS1261
- ▶ *IBM FlashSystem in OLAP Database Environments*, TIPS0974
- ▶ *IBM FlashSystem in OLTP Database Environments*, TIPS0973
- ▶ *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186
- ▶ *IBM FlashSystem V840 Enterprise Performance Solution*, TIPS1158
- ▶ *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940
- ▶ *IBM SAN Volume Controller 2145-DH8 Introduction and Implementation*, SG24-8229
- ▶ *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- ▶ *Implementing FlashSystem 840 with SAN Volume Controller*, TIPS1137
- ▶ *Implementing ProtecTIER 3.3 and IBM FlashSystem*, TIPS1140
- ▶ *Implementing the IBM Storwize V7000 V7.2*, SG24-7938
- ▶ *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933

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](http://ibm.com/redbooks)

## Online resources

These websites are also relevant as further information sources:

- ▶ IBM FlashSystem family products  
<http://www.ibm.com/storage/flash>
- ▶ IBM FlashSystem 900 in the IBM Knowledge Center  
[https://ibm.biz/fs\\_900\\_kc](https://ibm.biz/fs_900_kc)
- ▶ IBM FlashSystem 900 support portal and product documentation (requires an IBMid)  
<http://ibm.co/1bPQyZI>
- ▶ IBM Redbooks Solution and Product Guides for the IBM FlashSystem family  
<http://www.redbooks.ibm.com/redbooks.nsf/searchsite?SearchView&query=flashss>
- ▶ 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](http://ibm.com/support)

IBM Global Services

[ibm.com/services](http://ibm.com/services)

**Redbooks**

## **Implementing IBM FlashSystem 900**

SG24-8271-01

ISBN 0738442259



(0.5" spine)

0.475" <-> 0.873"

250 <-> 459 pages









SG24-8271-01

ISBN 0738442259

Printed in U.S.A.

Get connected

