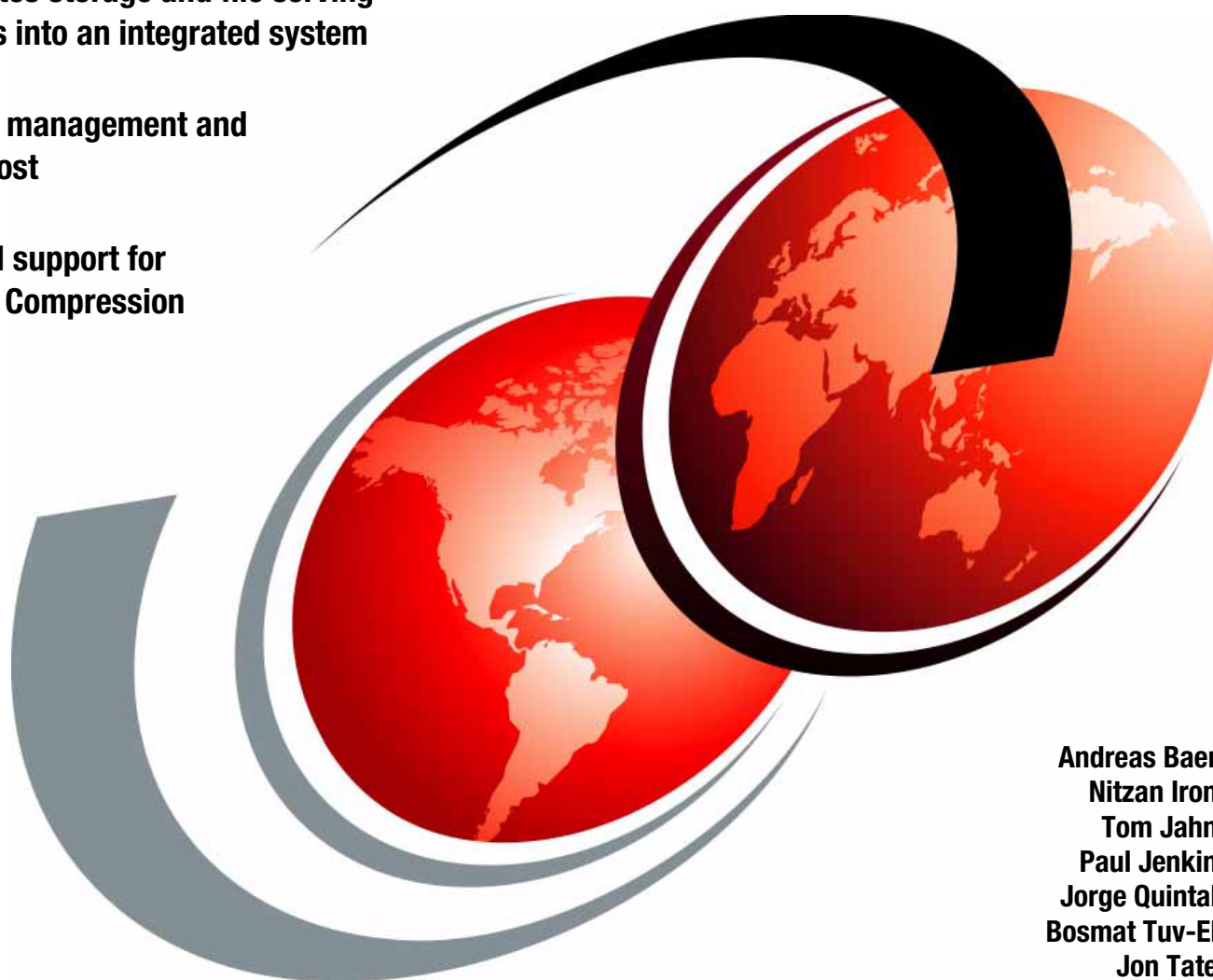


# Implementing the IBM Storwize V7000 Unified

**Consolidates storage and file serving workloads into an integrated system**

**Simplifies management and reduces cost**

**Integrated support for Real-time Compression**



Andreas Baer  
Nitzan Iron  
Tom Jahn  
Paul Jenkin  
Jorge Quintal  
Bosmat Tuv-El  
Jon Tate

# Redbooks





International Technical Support Organization

**Storwize V7000 Unified**

March 2013

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

### **First Edition (March 2013)**

This edition applies to Version 1.4.0.1-5c of the IBM Storwize V7000 Unified and Version 6.4.1.3 (build 75.2.1302012000) of the IBM Storwize V7000.

This document created or updated on March 29, 2013.

© Copyright International Business Machines Corporation 2011. 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> .....	xi
Trademarks .....	xii
<b>Preface</b> .....	xiii
The team who wrote this book .....	xiii
Now you can become a published author, too! .....	xvi
Comments welcome .....	xvi
Stay connected to IBM Redbooks .....	xvii
<b>Chapter 1. Introduction</b> .....	1
1.1 A short history lesson .....	2
1.2 About the rest of this book .....	3
1.3 Latest release highlights .....	5
1.3.1 Real-time Compression for File Systems .....	5
1.3.2 Local Authentication for NAS .....	5
<b>Chapter 2. Terminology and file serving concepts</b> .....	7
2.1 Terminology for storage and file services .....	8
2.1.1 Terminology for random access mass storage .....	8
2.1.2 Terminology for file systems and file sharing, file access and file transfer .....	9
2.2 File serving with file sharing and file transfer protocols .....	13
2.2.1 The Network File System (NFS) protocol .....	13
2.2.2 The Server Message Block (SMB) protocol .....	15
2.2.3 The File Transfer Protocol (FTP) .....	18
2.2.4 The Hypertext Transfer Protocol (HTTP) .....	19
2.2.5 The Secure Copy Protocol (SCP) .....	19
2.2.6 The SSH File Transfer Protocol (SFTP) .....	20
<b>Chapter 3. Architecture and functions</b> .....	21
3.1 High level overview of Storwize V7000 Unified .....	22
3.2 Storwize V7000 Unified system configuration .....	23
3.2.1 Storwize V7000 Unified storage subsystem configuration .....	23
3.2.2 Storwize V7000 Unified file server subsystem configuration .....	24
3.3 Storwize V7000 Unified storage functions .....	25
3.4 Storwize V7000 Unified file serving related functionality .....	26
3.4.1 Storwize V7000 Unified file sharing and file transfer protocols .....	27
3.4.2 Storwize V7000 Unified NFS protocol support .....	28
3.4.3 Storwize V7000 Unified SMB protocol support .....	29
3.4.4 Storwize V7000 Unified SONAS cluster manager .....	32
3.4.5 Storwize V7000 Unified product limits .....	33
<b>Chapter 4. Access control for file serving clients</b> .....	37
4.1 Authentication and authorization in general .....	38
4.1.1 UNIX authentication and authorization .....	38
4.1.2 Windows authentication and authorization .....	38
4.1.3 UNIX and Windows authentication and authorization in the Storwize V7000 Unified	39
4.2 Methods used for access control .....	39
4.2.1 Kerberos .....	39

4.2.2	User names and User IDs. . . . .	39
4.2.3	Group names and GIDs . . . . .	40
4.2.4	Resource names and security identifier (SID). . . . .	40
4.2.5	UID/GID/SID mapping in the Storwize V7000 Unified. . . . .	40
4.2.6	Directory services in general. . . . .	40
4.2.7	Windows NT 4.0 Domain Controller / SAMBA Primary Domain Controller. . . . .	41
4.2.8	Lightweight Directory Access Protocol (LDAP). . . . .	41
4.2.9	Microsoft Active Directory (AD). . . . .	41
4.2.10	Services for UNIX (SFU) and Identity Management for Unix . . . . .	41
4.2.11	Network Information Service (NIS) . . . . .	41
4.2.12	Access control list (ACL) in general . . . . .	42
4.2.13	GPFS NFSv4 ACLs . . . . .	42
4.2.14	POSIX bits . . . . .	42
4.2.15	ACL mapping . . . . .	42
4.3	Access control with Storwize V7000 Unified. . . . .	43
4.3.1	Authentication methods supported . . . . .	43
4.3.2	AD authentication . . . . .	43
4.3.3	AD with SFU authentication or with Identity Management for Unix. . . . .	44
4.3.4	SAMBA PDC authentication . . . . .	44
4.3.5	LDAP authentication . . . . .	44
4.3.6	Network Information Service. . . . .	45
4.4	Access control limitations and considerations. . . . .	45
4.4.1	Authentication limitations . . . . .	45
4.4.2	Authorization limitations . . . . .	46
<b>Chapter 5. Storage Virtualization. . . . .</b>		<b>49</b>
5.1	User requirements driving storage virtualization. . . . .	50
5.2	Storage virtualization terminology. . . . .	50
5.2.1	Realizing the benefits of Storwize V7000 Unified storage virtualization . . . . .	53
5.2.2	Using internal physical disk drives in the Storwize V7000 Unified . . . . .	53
5.2.3	Using external physical disk drives in the Storwize V7000 Unified. . . . .	54
5.3	Summary. . . . .	56
<b>Chapter 6. NAS use cases, differences between SONAS and Storwize V7000 Unified</b>		<b>57</b>
6.1	Use cases for Storwize V7000 Unified . . . . .	58
6.1.1	Unified Storage with both File and Block access . . . . .	58
6.1.2	Multi User File Sharing with centralized Snapshots and Backup . . . . .	59
6.1.3	Availability and Data Protection . . . . .	60
6.1.4	ILM, HSM and Archiving solution . . . . .	60
6.2	Storwize V7000 Unified and SONAS . . . . .	61
6.2.1	SONAS brief overview . . . . .	61
6.2.2	Implementation differences between the Storwize V7000 Unified and SONAS . . . . .	63
<b>Chapter 7. IBM General Parallel File System . . . . .</b>		<b>65</b>
7.1	Overview . . . . .	66
7.2	GPFS technical concepts and architecture. . . . .	66
7.2.1	Split brain situations and GPFS . . . . .	68
7.2.2	GPFS file system pools and Storwize V7000 storage pools. . . . .	68
7.2.3	File system pools in GPFS . . . . .	69
7.2.4	GPFS file sets . . . . .	71
7.2.5	GPFS parallel access and byte-range locking . . . . .	71
7.2.6	GPFS synchronous internal replication. . . . .	72
7.2.7	Active Cloud Engine . . . . .	72
7.2.8	GPFS and Hierarchical Storage Management . . . . .	73

7.2.9 GPFS Snapshots .....	73
7.2.10 GPFS quota management .....	74
<b>Chapter 8. Copy services overview</b> .....	<b>75</b>
8.1 Storage copy services of the Storwize V7000 Unified .....	76
8.1.1 FlashCopy for creating point-in-time copies of volumes .....	76
8.1.2 Metro Mirror and Global Mirror for remote copy of volumes .....	78
8.2 File system level copy services of the Storwize V7000 Unified file modules .....	79
8.2.1 Snapshots of file systems and file sets .....	79
8.2.2 Asynchronous replication .....	80
8.3 Managing Asynchronous Replication .....	82
<b>Chapter 9. GUI and CLI</b> .....	<b>85</b>
9.1 Graphical User Interface setup .....	86
9.1.1 Web server .....	86
9.1.2 Management GUI .....	87
9.1.3 Web browser and settings .....	88
9.1.4 Starting the browser connection .....	88
9.2 Command Line Interface setup .....	90
9.3 Using the GUI .....	91
9.3.1 Menus .....	92
9.4 Using the CLI .....	96
9.4.1 File commands .....	97
9.4.2 Block Commands .....	98
<b>Chapter 10. Planning for implementation</b> .....	<b>99</b>
10.1 Planning steps sequence .....	100
10.1.1 Get preliminary planning data .....	100
10.1.2 Determine the system configuration <b>to order</b> .....	<b>101</b>
10.1.3 Perform the physical hardware planning .....	101
10.1.4 Define the environment and services needed .....	101
10.1.5 Plan for system implementation .....	102
10.1.6 Plan for Data Migration .....	102
10.2 Support, limitations, and tools .....	102
10.3 Storwize V7000 Unified advanced features and functions .....	104
10.3.1 Licensing for advanced functions .....	104
10.3.2 Planning guidelines for using Real-time compression .....	104
10.3.3 Asynchronous Replication .....	106
10.3.4 Snapshots .....	107
10.3.5 Information Lifecycle Management .....	107
10.3.6 Hierarchical Storage Management (HSM) .....	107
10.3.7 Data Backup and Recovery: TSM or NDMP .....	107
10.3.8 Antivirus .....	108
10.3.9 External Virtualization of SAN attached back-end storage: .....	108
10.3.10 Remote Copy Services (for block I/O access only) .....	108
10.3.11 FlashCopy (block volumes only) .....	109
10.3.12 General GPFS recommendation .....	109
10.3.13 GPFS internal synchronous replication, also known as 'NSD failure groups' .....	109
10.3.14 Manage the write caching options in Storwize V7000 Unified and on client side .....	110
10.3.15 Redundancy .....	110
10.4 Checkpoints and considerations for authentication .....	111
10.4.1 Active Directory (includes Kerberos) .....	111
10.4.2 LDAP (Lightweight Directory Access Protocol) .....	112

10.4.3 Samba PDC (NT4 mode) . . . . .	112
10.4.4 Local Authentication . . . . .	112
10.5 SAN considerations . . . . .	112
10.5.1 Zoning considerations . . . . .	113
10.6 LAN considerations . . . . .	113
10.7 Miscellaneous configuration planning . . . . .	114
10.7.1 Set up local users to manage the Storwize V7000 Unified system . . . . .	115
10.7.2 Define call home and/or event notifications . . . . .	115
10.7.3 Storage pool layout . . . . .	115
10.7.4 File access protocols required for client access . . . . .	116
10.7.5 Multiple simultaneous exports of same subset of data via different protocols . . . . .	116
10.8 Physical hardware planning . . . . .	117
10.8.1 Plan for space and layout . . . . .	118
10.8.2 Planning for Storwize V7000 Unified environment . . . . .	119
10.9 System implementation planning . . . . .	120
10.9.1 Configuration details and settings required for setup . . . . .	121
10.9.2 Configuration options for file access only . . . . .	129
10.9.3 Configuration options for Block I/O access only . . . . .	131
10.10 Planning for Data Migration . . . . .	132
<b>Chapter 11. Implementation . . . . .</b>	<b>133</b>
11.1 Process overview . . . . .	134
11.2 Task checklist . . . . .	134
11.3 Hardware unpack, rack and cable . . . . .	136
11.3.1 Preparation . . . . .	136
11.3.2 Review packing slips and check components . . . . .	136
11.3.3 Confirm environmentals and planning . . . . .	138
11.3.4 Rack controller enclosures . . . . .	138
11.3.5 Rack expansion enclosures . . . . .	138
11.3.6 Rack file modules . . . . .	139
11.3.7 Cabling . . . . .	139
11.4 Power on and checkout . . . . .	146
11.4.1 Network . . . . .	146
11.4.2 Power on expansions and controllers . . . . .	146
11.4.3 Power on file modules . . . . .	146
11.5 Install latest software . . . . .	147
11.5.1 Determine current firmware and code levels . . . . .	147
11.5.2 Preparation for reload . . . . .	148
11.5.3 Re-install the software . . . . .	148
11.6 Initialize the system . . . . .	149
11.6.1 Configure USB key . . . . .	149
11.6.2 Initialize the Storwize V7000 controller . . . . .	152
11.6.3 Initialize the file modules . . . . .	153
11.7 Base configuration . . . . .	155
11.7.1 Connect to GUI interface . . . . .	155
11.7.2 Easy Setup wizard . . . . .	156
11.7.3 Setup periodic configuration backup . . . . .	171
11.8 Manual setup and configuration changes . . . . .	171
11.8.1 System names . . . . .	171
11.8.2 System licenses . . . . .	171
11.8.3 Support . . . . .	172
11.9 Network . . . . .	174
11.9.1 Public Networks . . . . .	174



11.9.2 Service ports . . . . .	175
11.9.3 iSCSI . . . . .	176
11.9.4 Fibre Channel ports . . . . .	177
11.9.5 Fibre Channel ports . . . . .	177
11.10 Alerting . . . . .	178
11.10.1 E-mail . . . . .	178
11.10.2 SNMP . . . . .	180
11.10.3 Syslog Server . . . . .	181
11.11 Directory Services and Authentication . . . . .	182
11.11.1 DNS . . . . .	182
11.11.2 Authentication . . . . .	183
11.11.3 NTP server . . . . .	184
11.12 Health check . . . . .	185
11.13 User Security . . . . .	186
11.13.1 Change passwords . . . . .	186
11.13.2 Create cluster users . . . . .	188
11.13.3 Create local users using Local Authentication for NAS access . . . . .	189
11.14 Storage controller configuration . . . . .	191
11.14.1 External SAN requirements . . . . .	191
11.14.2 Configure storage . . . . .	192
11.15 Block configuration . . . . .	192
11.15.1 Copy Services . . . . .	192
11.16 File Services configuration . . . . .	197
11.16.1 File service components . . . . .	197
11.16.2 File Systems examples . . . . .	199
<b>Chapter 12. Antivirus . . . . .</b>	<b>207</b>
12.1 Overview . . . . .	208
12.2 Scanning individual files . . . . .	209
12.3 Batch scan . . . . .	209
12.4 Setup and configure Antivirus . . . . .	210
12.4.1 Antivirus setup steps . . . . .	210
<b>Chapter 13. Performance and Monitoring . . . . .</b>	<b>219</b>
13.1 Monitoring . . . . .	220
13.2 TPC . . . . .	222
13.2.1 Configuration . . . . .	222
<b>Chapter 14. Backup and Recovery . . . . .</b>	<b>223</b>
14.1 Cluster backup . . . . .	224
14.1.1 Philosophy for file and block . . . . .	224
14.1.2 Storage enclosure backup (Storwize V7000) . . . . .	224
14.1.3 File module backup . . . . .	225
14.2 Cluster recovery . . . . .	226
14.2.1 Storage enclosure recovery (V7000) . . . . .	226
14.3 Data backup . . . . .	229
14.3.1 Data backup philosophy . . . . .	229
14.3.2 TSM . . . . .	229
14.3.3 NDMP . . . . .	232
14.4 Data recovery . . . . .	238
14.4.1 TSM . . . . .	238
14.4.2 Asynchronous data recovery . . . . .	239
14.4.3 NDMP . . . . .	239

<b>Chapter 15. Troubleshooting and Maintenance</b>	241
15.1 Maintenance philosophy	242
15.2 Event logs	242
15.2.1 Storwize V7000 Storage Controller Event log (block)	244
15.2.2 V7000 Unified File Module Event log file	247
15.2.3 Block	248
15.2.4 File	250
15.2.5 Working with compressed volumes out of space conditions	252
15.3 Collect Support Package	256
15.4 Information Center	259
15.4.1 Contents	259
15.4.2 Index	259
15.4.3 Search results	259
15.4.4 Offline Information Center	260
15.5 Call home and alerting	261
15.5.1 SNMP	261
15.5.2 Email	261
15.5.3 Call Home	261
15.6 IBM Support Remote Access	262
15.7 Changing parts	262
15.8 Preparing for recovery	263
15.8.1 Superuser password	263
15.8.2 admin password	263
15.8.3 Root password	264
15.8.4 Service ip addresses	264
15.8.5 Test GUI connection to the Storwize V7000 Storage enclosure	264
15.8.6 Assist-on-site	264
15.8.7 Backup config saves	264
15.9 Software	265
15.9.1 Software package	265
15.9.2 Software upgrade	267
 <b>Chapter 16. Real-time Compression in the IBM Storwize V7000 Unified</b>	 275
16.1 General Considerations for compression and block volume compression use cases	276
16.2 Compressed File System Pool Configurations	276
16.2.1 Selectively compressed file system with two pools	276
16.2.2 Configuring a selective compressed file system	278
16.2.3 Fully Compressed File system	281
16.2.4 Compression Rules by File Set	282
16.3 Capacity Planning	286
16.3.1 Planning Capacity with the Comprestimator tool	286
16.3.2 Capacity planning for selectively compressed file systems	286
16.4 Compression Metrics for File Systems	288
16.5 Managing compressed file systems	290
16.5.1 Adding a compressed pool to a file system	290
16.5.2 Making a selectively compressed file system uncompressed	293
16.6 Compression Saving Reporting	295
16.6.1 Reporting basic overview	295
16.6.2 Reporting of compression in the GUI	296
16.6.3 Compression reporting using the CLI	300
 <b>Related publications</b>	 303
IBM Redbooks	303

Other publications .....	303
Online resources .....	305
Help from IBM .....	305
<b>Index</b> .....	<b>307</b>



# Notices

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

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

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

*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

Any references in this information to non-IBM 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 supply in any way it believes appropriate without incurring any obligation to you.

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

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


## COPYRIGHT LICENSE:

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

## Trademarks

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

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

Active Cloud Engine™	OS/390®	System x®
AIX®	Real-time Compression™	Tivoli®
Easy Tier®	Redbooks®	XIV®
FlashCopy®	Redbooks (logo)  ®	z/OS®
GPFS™	Storwize®	
IBM®	System Storage®	

The following terms are trademarks of other companies:

Intel, Intel Xeon, 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, Windows NT, 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.

# Preface

IBM® Storwize® V7000 Unified is a virtualized storage system designed to consolidate block and file workloads into a single storage system for simplicity of management, reduced cost, highly scalable capacity, performance and high availability. IBM Storwize V7000 Unified Storage also offers improved efficiency and flexibility through built-in solid state drive (SSD) optimization, thin provisioning, Real-time Compression™, and non-disruptive migration of data from existing storage. The system can virtualize and reuse existing disk systems offering a greater potential return on investment.

We suggest that readers familiarize themselves with the following books in order to get the best from this book:

*Implementing the IBM Storwize V7000 V6.3*, SG24-7938

*Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933

*Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859

*SONAS Implementation Guide and Best Practices Guide*, SG24-7962

*SONAS Concepts, Architecture, and Planning Guide*, SG24-7963

## The team who wrote this book



**Andreas Baer** is an IBM Certified Consulting IT Specialist with Advanced Technical Support at the European Storage Competence Center (ESCC) in Mainz, Germany. He joined IBM in 1995 and has worked as PFE on several disk subsystem products before joining ATS specializing in High End Disk solutions, Open Systems attachments, and Virtualization solutions with SAN Volume Controller (SVC) and Storwize V7000. He coordinated the product introductions for several ESS, DS6000/DS8000, and SVC/Storwize V7000 generations from an ATS perspective. Andreas is also working with multiple Development, Program Management, and Technical Support Management teams with focus on topics like testing, interoperability, product introductions, and special customer support requests. He holds a Masters degree in Physics from the Technical University of Darmstadt, Germany.



**Nitzan Iron** is a Level 3 Development Support Team Leader Engineer for IBM Real-time Compression in Israel. Nitzan has 12 years of IT, support and pre sales in Network Attached Storage. He joined IBM through the acquisition of Storwize in 2010 and his knowledge extends to various storage systems, operating systems, network switching and scripting. His current responsibilities include worldwide product support of IBM Real-time Compression.



**Tom Jahn** is a Senior Client Technical Specialist for IBM Technical Sales System Storage in Germany. He has 16 years of experience providing technical sales support in IBM. Tom provided technical sales support for networking and server consolidation on OS/390® UNIX for IBM and its clients before he joined the storage brand twelve years ago. He is engaged in providing technical support for Open Systems storage solutions across multiple platforms and a wide client base, currently with a focus on next generation storage and storage virtualization. He holds a Dipl. Ing. degree in Computer Science from the Staatliche Studienakademie Sachsen.



**Paul Jenkin** currently works as a Level 2 Technical Support Specialist in the Australia and New Zealand geography covering SAN and Storage Virtualization. He has worked in maintenance and support for all of his 32 years in IBM, having specialist and technical support roles for most of that time and across most of IBM's product range including PCs, networking, midrange and z-series enterprise systems. Since the introduction of SAN and networked storage systems, Paul has been involved in the solution design, implementation and support of SAN products. Paul holds a Diploma in Computer Engineering.



**Jorge Quintal** is a Storage Managing Consultant currently providing Development Support for IBM Real-time Compression. He joined IBM through the acquisition of Sequent Computer Systems in 1999 and during his time at IBM he has worked for Storage Lab Services as one of the original members working with SAN File System, SAN Volume Controller, NAS, and as lead for N-Series services development and implementations. Jorge also worked for an extended period of time as an XIV® Technical Advisor.



**Bosmat Tuv-El** is a Level 3 Development Support Engineer for IBM Real-time Compression in Israel. Bosmat joined the support team after 5 years in the QA department, where she was a team leader. She joined IBM through the acquisition of Storwize in 2010. Through her eight years of IT experience, Bosmat has gained a vast knowledge of various storage systems, operating systems, network switching and scripting. Her current responsibilities include worldwide product support of IBM Real-time Compression and product documentation. Bosmat is about to graduate in Computer Science and Management from The Open University in Israel.





**Jon Tate** is a Project Manager for IBM System Storage® SAN Solutions at the International Technical Support Organization, San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2 support for IBM storage products. Jon has 27 years of experience in storage software and management, services, and support, and is both an IBM Certified IT Specialist and an IBM SAN Certified Specialist. He is also the UK Chairman of the Storage Networking Industry Association.

This book was produced by a team of specialists from around the world working at Brocade Communications Systems, San Jose, IBM Israel, Tel Aviv, and the International Technical Support Organization, San Jose Center.

We extend our thanks to the following people for their contributions to this project, including the development and PFE teams in Hursley.

Chris Canto  
 Peter Eccles  
 Robin Findlay  
 Carlos Fuente  
 Geoff Lane  
 Richard Macauley  
 Andrew Martin  
 Cameron McAllister  
 Paul Merrison  
 Lucy Harris  
 Sukhi Sohal  
 Matt Smith  
 Barry Whyte  
 Muhammad Zubair  
**IBM Hursley**

Eyal Traitel  
**IBM Israel**

Duane Bolland  
 Jackson Shea  
**IBM Beaverton**

Norm Bogard  
**IBM Orlando**

Chris Saul  
**IBM San Jose**

Tina Sampson  
**IBM Tucson**

Sangam Racherla  
**IBM ITSO**

Achim Christ  
 Nils Haustein  
 Michael Jahn

Alexander Saupp  
Thomas Luther  
**IBM Germany**

Special thanks to the Brocade staff for their unparalleled support of this residency in terms of equipment and support in many areas:

Jim Baldyga  
Mansi Botadra  
Silviano Gaona  
Brian Steffler  
Marcus Thordal  
Steven Tong  
*Brocade Communications Systems*

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

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

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

[ibm.com/redbooks/residencies.html](http://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:

[ibm.com/redbooks](http://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>





# Introduction

In this IBM Redbook we introduce a new product, the IBM Storwize V7000 Unified (V7000U). The Storwize V7000 Unified integrates the serving of storage and file related services, such as file sharing and file transfer capabilities, in one system. The Storwize V7000 Unified is capable of storage system virtualization as well, using the mature virtualization capabilities of the IBM SAN Volume Controller (SVC). It is an integrated storage server, storage virtualization and file server appliance.

## 1.1 A short history lesson

IT infrastructure concepts and products advance and change over time, adjusting to changing requirements and tasks at hand. For instance, in many computing environments, a centralised approach with a strictly specified design and infrastructure components has been superseded by a client-server approach, with a decentralized and less proprietary, more interoperable infrastructure, easier to be adopted by new concepts and tasks.

As always IT architects have to work with the concepts, technology and designs that are available at a given point-in-time. For instance, it used to be the case that the servers in client-server environments could only use internal storage. This changed with the dawn of external RAID and RAID storage systems. If we look at storage as a service, this previously internal only, service was now out-tasked to a specialized device, enabling new infrastructure concepts.

The serving of storage from a central element in the infrastructure to many storage clients (multi-purpose servers that previously accessed their own storage) and the implementation of high availability in regards to storage with that now had the ability to mirror data to two independent storage servers, now became more prevalent. These specialized storage servers are in essence *storage server appliances*. What was simply called a 'server', because it housed all elements in one device and provided a service to clients, now became a client itself - a storage client. They need to use the storage service provided by the storage servers, to be able to build their own value-added services on top.

Servers were becoming more and more used for specialized tasks and subsequently the hardware and software was adapted to support this specialization to a certain degree.

One of these tasks was organizing files in file systems and making the file system space and the files themselves accessible by file sharing clients. These adapted servers are known as 'file servers'. With external storage, these devices are now storage clients on the one hand and file servers on the other. They *use* a lower level service (storage) and *provide* a higher level service (file serving).

To enhance the functionality, availability, disaster tolerance and other aspects of a file serving sub-infrastructure, new types of devices were developed and introduced. These were specialized single purpose file servers (a *file server appliance*) which are solely designed to provide only the functionality that is implemented into the product of the given vendor - without the explicit ability to use them for different tasks (as with multi purpose servers).

This specialization, which can have many other advantages, led to appliances that were called Network Attached Storage (NAS). Although strictly speaking it is not serving storage but files.

This gave us a modular, layered infrastructure in which each logical layer also corresponds with a class of device(s) which only provides services related to that layer. A storage server serves storage to a storage client. A file server serves files to a file client. These devices are connected to each other using different forms of external networks, using specialized protocols to provide their service.

When file server appliances started to be used to generate storage volumes from files in their file system, they made these volumes accessible to storage clients using iSCSI and FC. This meant that these devices provided services belonging to two different functional layers, and they acted as both file servers and storage servers. To make the distinction they were called 'Unified Storage'. The IBM approach to this is to take a storage server that is capable of storage virtualization and *integrates* it with a file server product in one device.

And this product is the IBM Storwize V7000 Unified Storage as shown in Figure 1-1.

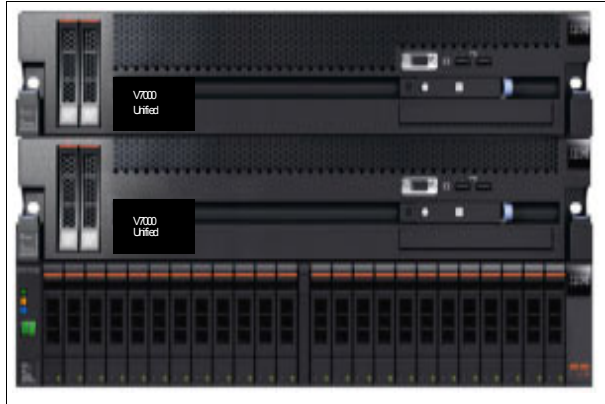


Figure 1-1 IBM Storwize V7000 Unified Storage

Internally this is still built with a layered approach in that a storage server serves storage to external storage clients as well as to the internal file server, and the file server provides file services to external file clients. This is a truly integrated storage server, storage virtualization and file server appliance.

## 1.2 About the rest of this book

This book starts off providing the basics of the terminology and concepts of storage and file services and how they relate to each other. Building on that we introduce file sharing and file transfer methods in general. The architecture of the Storwize V7000 is briefly explained, as well as the specifics of the implementation of the file related functionality and the access control options. Storage related functions like virtualization and copy services are covered more briefly, as these topics are covered in available ITSO publications already as these functions are the same as in the IBM Storwize V7000. We recommend having a copy of this book available to help your understanding of the Storwize V7000 piece of the equation:

*Implementing the IBM Storwize V7000 V6.3, SG24-7938*

Information of the file server related part of the product, such as the IBM General Parallel File System (GPFS™) and differences to SONAS are included as well. For more in-depth coverage of SONAS we recommend:

*SONAS Implementation Guide and Best Practices Guide, SG24-7962*

*SONAS Concepts, Architecture, and Planning Guide, SG24-7963*

For GPFS we recommend:

*Implementing the IBM General Parallel File System (GPFS) in a Cross Platform Environment, SG24-7844*

The theoretical part of the book and is followed by the implementation chapters of the book. The user interfaces, planning for implementation, and the actual implementation of the Storwize V7000 Unified are described. This includes the Antivirus implementation, performance and monitoring overview, backup and recovery, and troubleshooting and maintenance.

This book aims to help to understand file services and how they are implemented in the Storwize V7000 Unified and to help to successfully implement, run and maintain the product.



## 1.3 Latest release highlights

While the V7000 Unified received various small updates in various areas, there are two significant new features.

### 1.3.1 Real-time Compression for File Systems

It is now possible to create file systems that make use of the real time compression feature. This capability and the combination of the RtC feature together with the Policy language is covered in Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.

### 1.3.2 Local Authentication for NAS

The V7000 Unified can now serve as authentication system for NAS users. This feature is targeted for users that do not run an authentication infrastructure or would use the NAS features of the V7000 Unified in an isolated environment with only few users.

NAS Users can be managed using the CLI or GUI. NAS users can self manage their passwords using a web interface. This feature is covered in Authentication services section of this redbook as well as in the infocenter.





# Terminology and file serving concepts

In this chapter we talk about the terminology we use for storage and file services and discuss client server file sharing and file transfer concepts.

## 2.1 Terminology for storage and file services

With the Storwize V7000 Unified we provide two different service layers to clients. Especially with an integrated product like the Storwize V7000 Unified it is very important to use coherent terminology.

The main focus when designing infrastructures is the service the products provide (for example, is it the kind of service the client asks for, are the requirements met?). Even though our products perform input/output operations (I/O), it is services like the mapping of logical volumes and the sharing of files which are important to have in mind.

### 2.1.1 Terminology for random access mass storage

The lowest level of service provided to permanently store and retrieve data for use in upper layers in computer systems is random access mass storage, hereafter referred to as “**storage**”. Storage, for instance, is provided by Hard Disk Drives (HDD), Solid State Drives (SSD) and (using HDD and SSDs) by Redundant Array of Independent Disks (RAID) controllers and the controllers of external **storage systems** / **storage servers**. Oftentimes the term “block storage” is used, although in many cases it is not needed to use the term “block”.

The prefix “block” originates in the disk data organization architecture. For instance the Fixed Block Architecture (FBA or shorter FB) uses records of data written in “blocks” of a specific size. Logical Block Addressing (LBA), an FBA which is common today, uses a fixed sector size of 512 bytes. Referring to storage with the prefix “block” leaves out other architectures such as Count Key Data (CKD), used for instance in IBM z/OS®. In cases where it is needed to differentiate from CKD, it makes sense to refer to the storage as **fixed block storage (FB storage)** without omitting the term “fixed”.

LBA storage, as used by Open Systems, provides the storage to upper layers of the system as a sequence of blocks which is called a **volume**. For the term volume as well, the prefix “block” is *not* needed to differentiate from upper layer facilities and protocols, since these protocols do not provide storage nor volumes in any form - they *use* storage in form of volumes and provide upper layer services.

Storage provided by physical random access mass storage devices such as HDDs or SSDs is referred to as **physical volume (PV)**, for instance in RAID and in Logical Volume Manager (LVM) software components of an operation system.

Using physical volumes to start with, there are volumes which are called **logical volume (LV)** being presented in different places. As such the term logical volume has been overloaded and it depends on either the context it is being used in or upon a further definition to be understood. An LV can be a logical volume provided by a computer systems internal RAID adapter, an external storage system or an LVM. All these LV have in common that they as well present a sequence of blocks which can be used the same way a PV would. For upper layers which use these volumes the source is transparent, that means they do not know if they are using a PV or an LV. For instance an LV, like a PV, may be used by a File System (FS) or a Database Management System (DBMS). The diagram in Figure 2-1 on page 9 shows a storage server providing logical volumes to a storage client.

A computer system accessing a logical volume mapped to it is a **storage client**. A logical volume served to a storage client by a storage server (storage system) is often referred to as a “LUN”. This usage of the term LUN is wrong since LUN is short for Logical Unit Number, an addressing scheme used to *identify* a logical unit (LU). This addressing scheme is used for Small Computer System Interface (SCSI) based systems, using the different flavors of SCSI,

Internet SCSI (iSCSI) and Fibre Channel (FC). A logical volume is only one kind of an SCSI **logical unit** (LU), identified by a LUN - but LUs are *not limited* to be logical volumes. They can be other devices addressed by the SCSI addressing scheme as well, for instance tape devices.

Summarized, with **storage servers** we provide the service **storage** when we map **logical volumes** to **storage clients**.

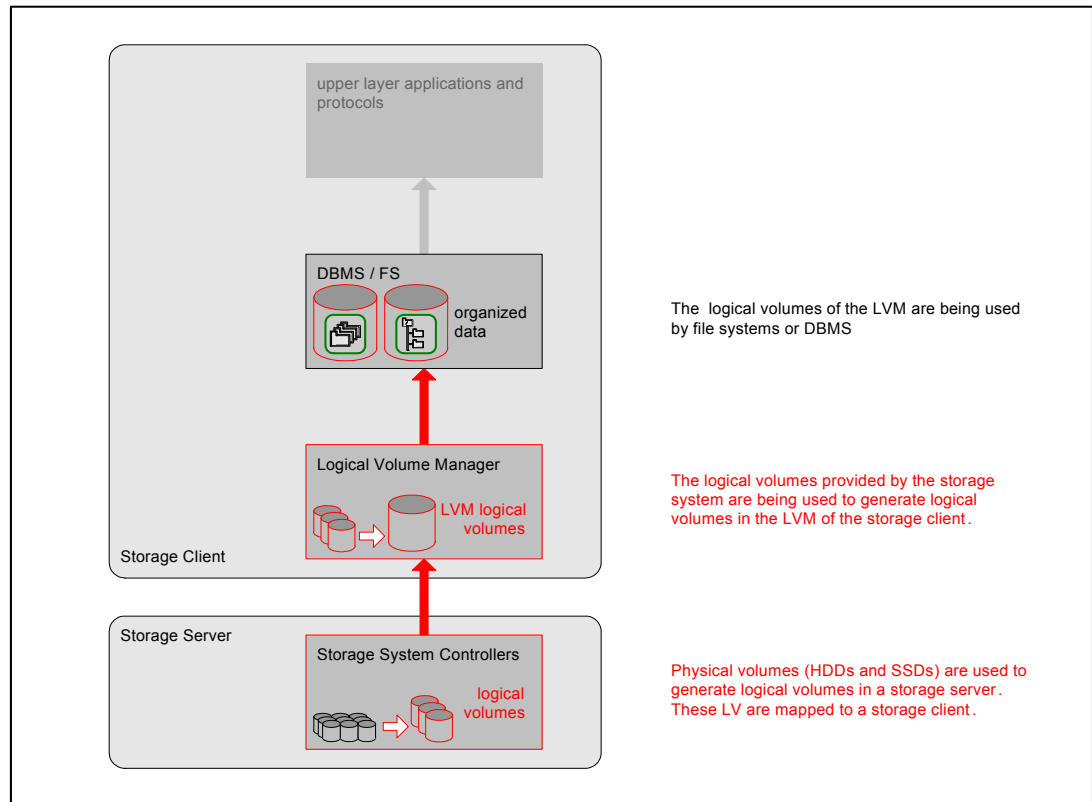


Figure 2-1 Storage server providing logical volumes to a storage client

## 2.1.2 Terminology for file systems and file sharing, file access and file transfer

Operating Systems of storage clients put a structure on and organize the storage space of logical volumes to store and retrieve data. Usually an LVM takes the logical volumes from storage servers to build LVM logical volumes. On top of the LVM logical volumes are the facilities to structure them to enable writing of, accessing and reading of data, such as FS or DBMS.

Sometimes there is a bit of confusion about **file systems** and **file serving and file sharing**, which is partly rooted in the naming of the networking protocols used to share and access file resources.

A **file system** is a means to write (store permanently), organize, find and read data in the form of file system **files** on a computer systems random access mass storage device, which can be a physical volume or a logical volume. The files in file systems are accessed by services, applications and protocols running on the local computer system. Examples for file systems are the Third Extended File System (**ext3**) and the Microsoft New Technology File System (**NTFS**).

**File sharing** is a means of making data *already organized* in a file system **accessible** to users of other (network connected) computer systems. These **file sharing protocols**, also called **file access protocols**, use client server infrastructures. We do not intend to discuss Peer-to-Peer (P2P) file sharing protocols in this publication. The **file server** part of the protocol makes the files accessible for the **file client** part of the protocol. It is common and technically correct to say files are being **shared**, not only because the server shares access to the files with the client, but also because these files may be accessed by multiple clients (shared).

Some common file sharing protocols that allow users to access files on another computer systems file system in a similar way as they access data in their local file system, are the different versions of the **Network File System (NFS)** protocol and the **Server Message Block (SMB)** protocol. The SMB protocol very often is referred to as Common Internet File System (CIFS). CIFS is only a *specific dialect* (one version) of the SMB protocol. See 2.2.2, “The Server Message Block (SMB) protocol” on page 15 for more information about the SMB protocol. The newest versions of the file sharing protocols NFS and SMB are NFSv4.1 and SMB 2.1.

The terms used for NFS and SMB servers and clients are **NFS file server**, **NFS file client** and **SMB file server**, **SMB file client**. Sometimes, the software instances are called **NFS file service** and **SMB file service** to distinguish them from the computer systems hardware they run on. In short, the terms NFS server and NFS client and SMB server and SMB client are used as well.

The diagram in Figure 2-2 on page 10, as an example, shows a file server using logical volumes, the file system ext3 and the file sharing protocol NFS to make files accessible to an NFS client.

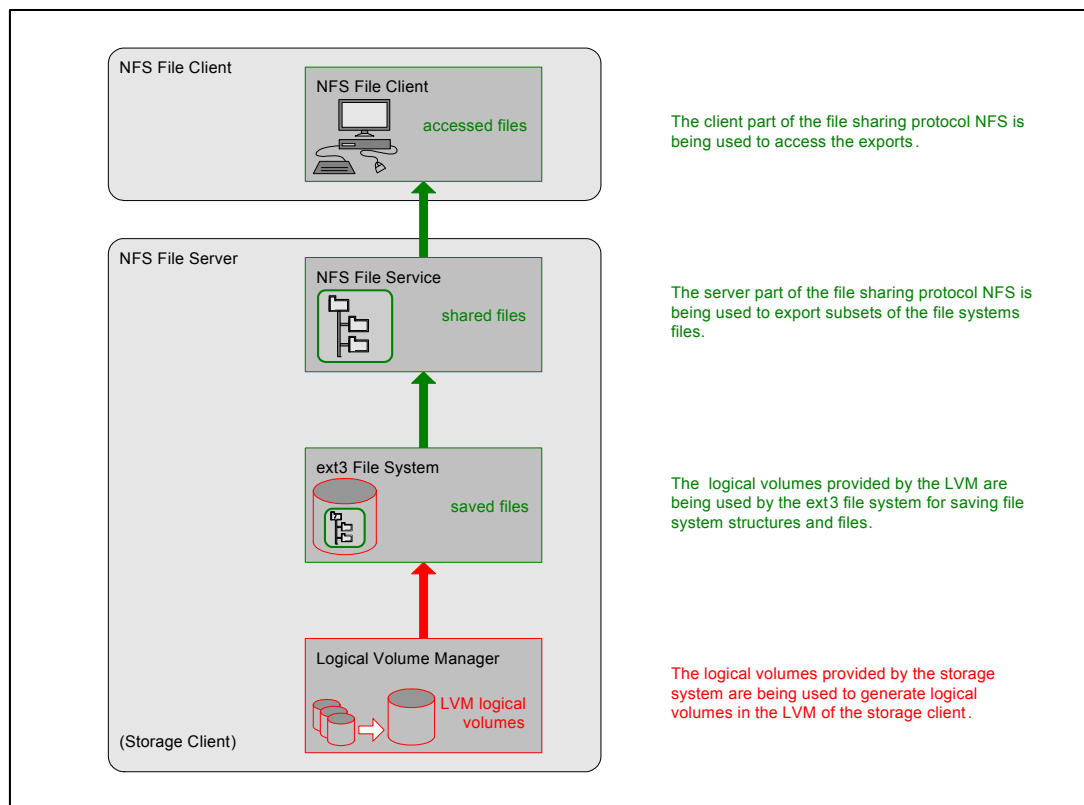


Figure 2-2 File server housing a file system and sharing files, file client accessing files

As we have seen, although the term “File System” is part of both terms, of **NFS** and the SMB protocol dialect **CIFS**, these are *not* file systems. They are networking client server file sharing protocols/file access protocols, which as upper layer protocols rely upon the existence of a file system to begin with.

Other methods of making files accessible and transferring files are the File Transfer Protocol (FTP), the Hypertext Transfer Protocol (HTTP) and its flavors as well as the Secure Shell (SSH) based protocols Secure Copy Protocol (SCP) and Secure FTP (SFTP). The terms used for FTP and HTTP servers and clients are **FTP file server** (short **FTP server**), **FTP file client** (short **FTP client**) and **HTTP server**, **HTTP client**.

The different means of file sharing such as using the NFS protocol and file transfer services such as using the FTP protocol can be called **file serving** in its broadest sense.

File servers are sometimes referred to as “file storage” and logical volumes used to place file systems on are sometimes called “file volumes”. Both terms are wrong and misleading. Since we established that file sharing as an upper layer service is not equal to the provision of storage (the mapping of logical volumes to storage clients is), an easy and correct way of referring to these systems and services would be “file server” and “file serving” or when file sharing protocols are used “file sharing”. This avoids abusing the terms “storage” and “volume” with the prefix “block” (to distinguish it from “file storage” and “file volume”) as well. All volumes for Open Systems are “Fixed Block Architecture” volumes, including all volumes used to place file systems on. There is no such thing as a “file volume”.

Specialized file servers are commonly called Network Attached Storage (NAS). Although NAS products are network attached, they in fact are file servers which use internal or external storage (in which case they themselves are storage clients). Although the term NAS is widely used, it may help to think of them as a File Server Appliance or simply as special purpose file servers.

The diagram in Figure 2-3 on page 12 shows the different layers from storage to the access of files and how they relate to each other. The example shown is using a storage server to provision logical volumes to a storage client (the file server). The file server then uses these logical volumes to organize data in form of files in file systems. The file server then provides access to the content of these file systems to file clients using file access protocols.

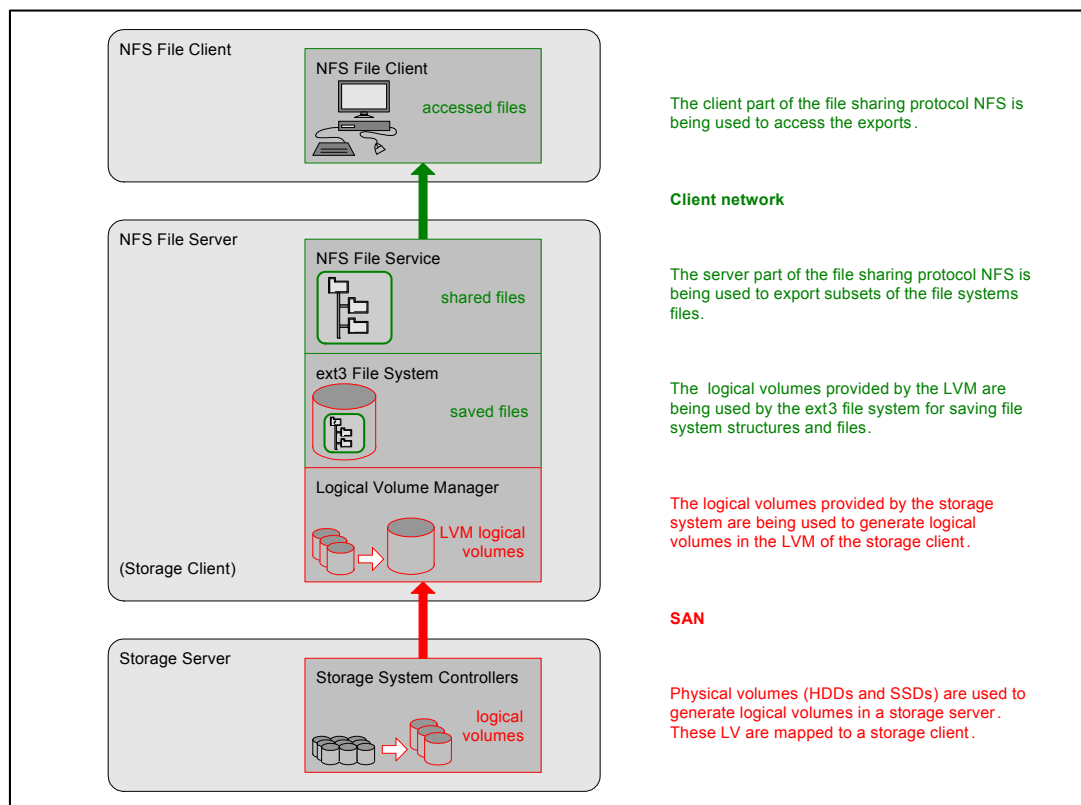


Figure 2-3 Storage server providing LV to file server, file server sharing files, file client accessing files



## 2.2 File serving with file sharing and file transfer protocols

One important aspect of distributed computing is being able to access files located on a remote system. To the user, remotely located files should have a “transparent” image; that is, the user should not be concerned whether the file is on a remote or local system. This also means that only a single set of commands should be defined to control local and remote files. The file sharing protocols NFS and SMB use the concept of integrating access to shared parts of remote file systems into the local file system structures. Some of the design issues for file access protocols are:

- ▶ **Access:** A file should appear to a user as a local file whether it is located on a remote or local machine. The path name to both local and remote files should be identical.
- ▶ **Concurrency:** The access to a file must be synchronized so that a modification by one process does not cause any undesired effect to other processes that depend on the same file.
- ▶ **Failure and recovery:** Both the client and the server must provide a recovery mechanism such that if one of both fails, the other will carry out a set of expected behaviors.
- ▶ **Mobility:** If a file is moved from one system to another, the client must still be able to access that file without any alteration.
- ▶ **Scalability:** A wide range of network sizes and workloads must be supported.

Other ways of serving and transferring files are for instance FTP, HTTP, SCP and SFTP. They are designed with different goals in mind, for instance HTTP is mainly used to serve content to world wide web clients, whereas FTP is being used to transfer files to and from remote locations without the goal of making the remote files available to processes of the local machine transparently.

### 2.2.1 The Network File System (NFS) protocol

The Network File System (NFS) is a file sharing protocol first developed by Sun Microsystems (Sun) in 1984. It is designed to enable computer users to access portions of file systems on remote systems in a similar way as if they were accessing files in a local file system. Initially, NFS was used within Sun only.

#### Overview of the NFS protocol

The portions of a file system tree made accessible are called **exports** on the server side and **mounts** on the client side. Import would have been a fitting term since importing is what happens. The exported parts of a file system get “imported” into the file system structure of the local system. The characteristics, dependencies and limitations of these exports are the same as those of the file system they have been exported from. NFS, due to its roots is found mostly on UNIX and UNIX-like systems and is included in the base OS or distributions. NFS is the standard for these systems to share and access files remotely. There are NFS implementations available for many other OS as well.

NFS initially only used the stateless User Datagram Protocol (UDP) as the transport layer but implementations using the stateful Transmission Control Protocol (TCP) started to appear as well. NFS is based on Remote Procedure Call (RPC), which is an inter-process communication protocol used to invoke routines in another address space, such as in a remote computer system. NFS is described and defined in Request For Comment (RFC) documents and may be implemented free of charge.

**Note:** RFC documents are sometimes referred to as being “internet standards”. This is not the case, although some RFCs become a standard as well. An RFC may, for example, propose and describe internet related protocols and methods. Even though not officially a standard, it allows implementations of protocols to adhere to a specific RFC so that implementations based on a certain RFC may be able to interact and be compatible with each other. The name RFC can be misleading; an RFC, once it is published, is not changed. In case there is a need for a change, a new RFC is being published with a new RFC number.

## NFS Version 2

The first version of NFS available outside of Sun was NFS Version 2 (NFSv2). It was published in 1989 in RFC 1094. NFSv2 needs a port mapper (rpc.portmap, portmap, rpcbind), which assigns the ports on which the NFS services will listen. These ports are temporary and change once the NFS server restarts. Additional configuration is necessary for the ports used to become permanent, which makes NFS usable through firewalls which only allow traffic to specified ports.

NFSv2 has some limitations concerning use cases, scalability and performance. It supports files only up to 2 GB, due to the 32 bit file offset. The size of any single data transfer cannot exceed 8 KB, which hinders performance due to the high amount of NFS requests. Another performance limiting factor is that NFSv2 only works in a synchronous way. Data must be written to the file system by the NFS server before the write is acknowledged to the client (so called stable writes). This limits the scalability of NFSv2.

NFSv2 does not support Kerberos authentication. To grant NFS clients access to exports, the access is granted to the computer system that the NFS client is running on. This means any user on that system is able to access the exports. The limits for users are only file and directory permissions.

## NFS Version 3

NFS Version 3 (NFSv3) was published in 1995 in RFC 1813. It still uses the port mapper. NFSv3 removed some limitations and introduced some changes. The file offset is now 64 bit, allowing the support for large files. The maximum transfer size limit of 8 KB is gone, the client and server can agree upon the transfer size.

NFSv3 introduced an asynchronous method of operation (so called unstable writes). With unstable writes the server does not need to acknowledge the write to the file system to the client immediately and thus can delay the write. The server must acknowledge the write only when it receives a commit request. This speeds up client writes and enables the server to efficiently write the data, mostly independent from the clients write operations. The NFSv3 client is able to detect uncommitted data in an error situation and can recover from that.

NFSv3 still grants access to the computer system and does not authenticate the user, and does not support Kerberos. This limits the use of NFSv3 to (very) trusted networks.

Sun handed over the maintenance of NFS to the Internet Engineering Task Force (IETF) before NFS Version 4 was defined and published.

## NFS Version 4

NFS Version 4 (NFSv4) was published in 2000 in RFC 3010. In 2003 it was revised and published in RFC 3530. The NFSv4 server does not rely on port mapper anymore. NFSv4 requires TCP as the transport layer, it listens on the well known port TCP 2049. It is a stateful protocol, maintaining the state of objects on the server. This way the server knows about the intentions of clients and some problems with stateless operation can be avoided. NFSv4

improves performance and functionality, for instance with file system semantics for the Microsoft Windows operating systems. It supports Windows Access Control Lists (Windows ACL), but not the Portable Operating System Interface based on Unix (POSIX) ACL, in addition to UNIX permissions. The security model in NFSv4 builds upon Kerberos, Low Infrastructure Public Key Mechanism (LIPKEY), and Simple Public Key Mechanism Version 3 (SPKM-3). The method used is agreed upon by the NFS client and NFS server. Also to be negotiated are other security mechanisms such as which encryption algorithm is being used. A recent development is NFS version 4.1 (NFSv4.1), which was published in 2010 in RFC 5661. NFSv4.1 offers new features for instance to leverage clustered installations with parallel processing.

## 2.2.2 The Server Message Block (SMB) protocol

The Server Message Block (SMB) protocol is a client server protocol for sharing file and printing resources in a distributed computing environment. It enables the access to shared remote resources in a similar way as if they were part of the local system.

Initially the SMB protocol was developed by IBM and further developed by Microsoft, Intel, IBM, 3Com and others. An early mention of SMB is in the *IBM Personal Computer Seminar Proceedings* document from October 1984. In 1987, the SMB protocol was officially defined in a Microsoft/Intel document called Microsoft Networks/OpenNET-FILE SHARING PROTOCOL. Thereafter it was developed by Microsoft and others. It has been mainly used with client computers running the IBM OS/2 OS versions and the Microsoft Windows family of OSs, where the SMB protocol with the functionality to act as SMB server and SMB client is build in. SMB server and SMB client implementations for other platforms became available and are in widespread use today.

### Overview of the original SMB protocol

The SMB protocol has been developed further over the years, which results in many variants of the protocol, called dialects. It retained backward compatibility with the ability to negotiate the dialect to be used for a session. The dialects are defined by a standard command set and are identified by a standard string such as PC NETWORK PROGRAM 1.0 (the first dialect of the SMB protocol), MICROSOFT NETWORKS 3.0, DOS LANMAN 2.1, or NT LM 0.12 (the SMB dialect NT LAN Manager, designated as CIFS).

The name of the SMB protocol refers to the packets of data sent between the client and server (the Server Message Blocks). Each SMB contains a request from a SMB client or a response from an SMB server (client server, request response protocol).

The SMB protocol is used to provide access to resources and access resources, called **shares**. Shares can be subsets of file systems and its contents, printers and serial ports and some other resources. The term share is also used because the resources may be accessed by multiple clients (shared between them), with the protocol providing the locking mechanisms. The SMB protocol, being a presentation/application layer protocol, has been implemented on top of various transport layer protocols, with NetBIOS over TCP/IP being the most common. In general it is independent from the transport protocol, but expects it to be connection oriented. With changes, it can be implemented on top of a stateless protocol such as UDP as well.

As stated before, the SMB protocol is backward compatible, for both the server and the client. When an SMB client connects to a SMB server, they identify which dialect of the SMB protocol they both understand and negotiate which one to use. The goal is to agree upon the dialect with the highest level of functionality they both support.

There are two levels of access control. The first level is the share level in which a client needs to provide the password for the share. The second level is the user level, in which the client authenticates to the server with a user name and password. Once authenticated, the server in turn sends a User ID (UID) to the client, which uses this UID in all later SMBs. Now the access to shares (which are not protected at the share level as well) is possible. In case of file serving, the SMB client may now open, read, write and close files by sending the proper SMBs to the SMB server.

SMB provides the locking of files and records. To enhance performance a mechanism called opportunistic locking (oplock) has been introduced. It enables SMB protocol clients to cache data. This mechanism can lead to data loss in case of connection failures or server failures. To prevent data loss in case of failures, an option to disable opportunistic locking may be provided in the SMB protocol implementation. Disabling oplocks will of course remove the performance benefits as well. Figure 2-4 shows how an SMB server manages oplocks.

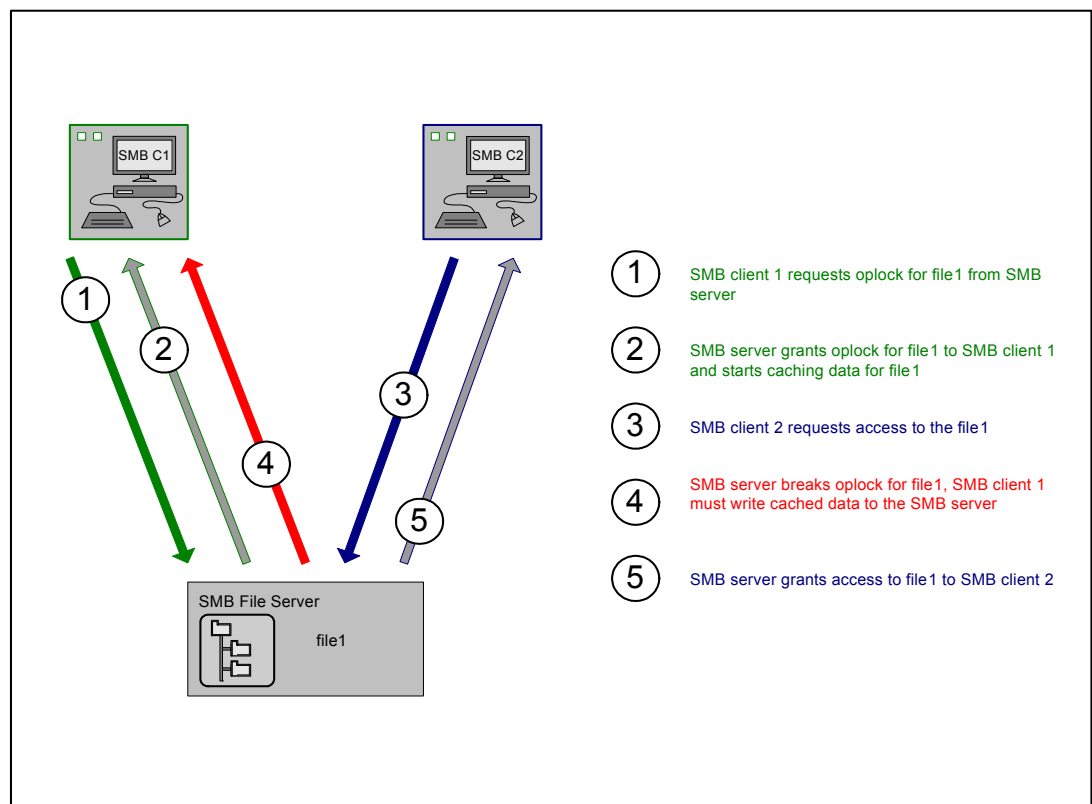


Figure 2-4 SMB server managing oplocks

The SMB protocol enables to set attributes for files, directories and extended attributes as well. It also supports ACLs.

**What is CIFS and what is it not?:** CIFS is *not synonymous* with the SMB protocol, rather it is a *dialect* of the SMB protocol.

Microsoft started to call its versions of the SMB protocol the “Microsoft SMB protocol”. The first dialect of the Microsoft SMB protocol was the “NT LAN Manager” dialect. This dialect of the SMB protocol was based upon the implementation of the SMB protocol in the Microsoft NT4 and Windows 2000 (NT5) OS. Microsoft proposed this specific dialect of SMB to the Internet Engineering Task Force (IETF) to become a standard with the name “CIFS”. During this time, the terms SMB and CIFS started to be used interchangeably, driven by Microsoft and since it was expected by many that CIFS would become a standard and would be the successor (instead of just a dialect) of the SMB protocol. CIFS did not become a standard and has not been published as an RFC document as well. In 2002, the Storage Network Industry Association (SNIA) CIFS Work Group has published a document for CIFS with the title “Common Internet File System (CIFS) Technical Reference Revision: 1.0”, which does neither constitute a standard or specification, nor does it claim to be.

There is no real specification (for instance documented in an RFC) which would enable developers to implement the SMB protocol with (to spec) and be interoperable with various other implementations (which would adhere to this specification as well). For instance, to be interoperable with the SMB protocol implemented in Microsoft products, developers have to adjust to the changes Microsoft may make.

After CIFS (the NT LAN Manager dialect of the SMB protocol), the Microsoft SMB protocol was subsequently developed further and the term “SMB protocol” is being used by Microsoft and others. The extended version of CIFS is the Server Message Block (SMB) Version 1.0 Protocol, (SMB 1). The most recent development, the complete redesign of the SMB protocol, is being officially called the Server Message Block (SMB) Version 2 Protocol.

Unless the *specific dialect* of the SMB protocol which had been proposed to become the IETF standard “CIFS” is being referred to, the usage of the term “CIFS” is not correct.

For SMB 1, the authentication support has been enhanced. It is now compatible with the Generic Security Services Application Program Interface (GSSAPI). It supports Kerberos authentication through GSSAPI.

It is now possible to query for older file versions, in case that is supported by the file system. Other enhancements make the protocol more efficient, such as the implementation of SMB server side only operations (without the need to transfer the file to the SMB client and back to the SMB server. For instance, for an SMB client initiated copy operation from one directory to another it makes no sense to read and write the file (transfer the data back and forth over the network, wasting network resources) as it still is done in CIFS. Concerning the use of TCP as the transport protocol, additionally to NetBIOS over TCP, SMB 1 supports to run directly on TCP (Direct TCP), without the need for NetBIOS.

Quotas may be used by the SMB 1 client to limit file space used if the SMB server supports quotas.

**Kerberos and GSSAPI:** Kerberos is an Internet protocol designed to add security to networked servers. It uses secret-key strong cryptography to deliver user authentication for networked applications. Kerberos has been developed at the Massachusetts Institute of Technology (MIT) and can be freely obtained. It is available in commercial products as well. The Kerberos API is not standardized but Kerberos 5 includes a GSSAPI implementation.

GSSAPI is an Application Programming Interface (API) to enable software vendors to implement security related services with a common API instead of supporting each other directly. GSSAPI is an IETF standard. With GSSAPI being implemented in the SMB 1 protocol, Kerberos can be used for authentication of SMB 1 clients.

### The SMB Version 2 protocol

All versions of the SMB protocol, including CIFS and SMB 1, are evolutionary changes and enhancements of the original SMB protocol. “The Server Message Block (SMB) Version 2 Protocol” (SMB 2), introduced in 2007 by Microsoft, is a newly developed file sharing protocol. It features a different set of commands, but is based on the concepts of the original SMB protocol. The latest version of SMB 2 is SMB 2.1. To be compatible with older SMB versions (including the original SMB protocol versions), the older SMB protocol is used to negotiate the SMB version to be used by the SMB client and the SMB server.

SMB 2 reduces complexity, increases efficiency and thus performance (especially for high latency networks). Also, it is more scalable and introduces other enhancements such as connection error handling, improved message signing and support for symbolic links.

The SMB 2 protocol only supports TCP as the transport protocol, either using Direct TCP or NetBIOS over TCP.

## 2.2.3 The File Transfer Protocol (FTP)

The File Transfer Protocol (FTP) is an OSI application layer protocol designed to transfer files between computer systems. It was initially written by Abhay Bhushan and was developed further and defined in RFC documents. FTP is a client server protocol with the **FTP server** providing access to files (and file space) and **FTP clients** browsing the content made available, downloading files from and uploading files to the FTP server.

FTP is historically command line based but GUI implementations across many OS became available and are in widespread use. FTP clients may mimic the look and behavior of file managers or file browsers in operating systems GUIs, but FTP is not designed to integrate transparently into the representation of file system trees as file access protocols like the NFS protocol and the SMB protocol do. There is no share to connect to and no export to mount. Instead, a FTP client connects to an FTP server and either has to authenticate or may be able to connect as ‘anonymous’ (no authentication needed). Then the client may browse the FTP repository and transfer files. Once files have been downloaded to the FTP clients, users of the clients system may use the files from within the file system (as opposed to the file access protocols NFS and SMB, which are used to work with files still residing in the file system of the remote system).

FTP uses TCP with the server using two well known ports, the “control” (or command) port TCP 21 for commands and authentication and the “data” port TCP 20 for data transfers.

There are two modes of operation with FTP, **active mode** and **passive mode**. With active mode, the FTP client uses a random unprivileged port (above TCP 1023) to connect to the FTP servers control port (TCP 21, where the FTP server is listening on for incoming FTP client requests). Using this control connection, the FTP client informs the FTP server about its



*own* listening port for the transfer of data. The client side listening port known to the FTP server is not a well known port, instead it is the random port the FTP client used to connect to the FTP server plus 1 (the FTP client only tells the FTP server the port, it does *not initiate* the data transfer connection). Then, the server connects with its data port (TCP 20) to this connection specific port of the client. As we see this data connection is *initiated by the server* and the client side is “listening”, a behavior normally expected from a server. This mode of operation was called active mode in retrospect. It is the FTP server who *actively* tries to establish the data transfer connection to a client.

When servers or daemons “listen” to ports for connection attempts by clients from outside a firewall secured network, they need these ports opened or forwarded by firewalls or they would wait for eternity. Therefore, firewalls are usually configured to allow incoming traffic to servers *on specific well known* ports. The behavior of the FTP protocol in active mode, where the FTP client is listening to a random and temporary non privileged port, leads to implications with client side firewalls. Connection attempts from outside, if not specifically allowed, are usually blocked.

To overcome this issue, the passive mode has been introduced to FTP. In this mode, the FTP client uses two non privileged ports above port TCP 1023, usually consecutive, to initiate two connections to the FTP server. It uses the first port to establish the control connection to the FTP server control port TCP 21. It tells the FTP server to use passive mode (the FTP server will be listening on a port for the data transfer connection). The server side data port is a non privileged port above port TCP 1023 on the server side, randomly chosen by the FTP server. The FTP server sends the information about which port it listens on to the FTP client. With this knowledge the FTP client can now establish the data transfer connection from its second port to the FTP server listening port. Thus, only the client is initiating connections and the client side firewall should let them pass. The server side firewall needs to be configured to allow the incoming connection attempts to the non privileged ports to pass, which is the drawback to that method from the server side point of view. The issue is that the firewall needs to be configured to allow connection attempts to the ports above 1023. One way to reduce that problem is FTP servers supporting the limiting of listening ports to a range of ports. Only these ports would need to be configured by the firewall to allow incoming connection attempts.

## 2.2.4 The Hypertext Transfer Protocol (HTTP)

The Hypertext Transfer Protocol (HTTP) is an OSI application layer client server networking protocol. It is an IETF standard with the version HTTP/1.1 being the most recent. HTTP is used to transfer files between HTTP servers and HTTP clients. The HTTP server software makes the content available to the HTTP clients, using port TCP 80 by default (other TCP ports can be used as well but must be specified by the HTTP client). The HTTP client software may have functionality to interpret these files and display the result as layout. Such HTTP client software is commonly called “web browser”. As such, HTTP is a core technology for the World Wide Web (WWW).

To encrypt the requests by a client and the actual content while in transit, the Secure Sockets Layer (SSL) and its successor Transport Layer Security (TLS) can be used. This is called HTTP Secure (HTTPS) or HTTP over SSL/TLS. HTTPS uses port TCP 443 by default. HTTPS is specified in RFC documents.

## 2.2.5 The Secure Copy Protocol (SCP)

The Secure Copy Protocol (SCP) is a client server protocol used by SCP clients to copy files to and retrieve files from SCP servers. It is also possible to initiate a transfer of files between

two remote systems. It uses the Secure Shell infrastructure for encrypted transfer. Usually the SSH daemon provides the SCP functionality and may act as both, the SCP client and the SCP server. The port used is TCP 22. There is no official standard for SCP.

### **2.2.6 The SSH File Transfer Protocol (SFTP)**

SSH File Transfer Protocol (SFTP), also sometimes referred to as Secure FTP, is a file transfer protocol developed in the context of SSH2 and is the standard file transfer protocol to be used with SSH2. It expects to run over a secured connection that SSH provides. It is not equal to Simple FTP or to FTP over SSH, it is a newly developed protocol. It provides functionality similar to FTP as compared to the SSH based SCP protocol.





# Architecture and functions

In this chapter we provide an overview of the architecture of Storwize V7000 Unified and its functions.

## 3.1 High level overview of Storwize V7000 Unified

To be able to serve logical volumes and files, the hardware and software to provide these services is integrated into one product. Viewed from its clients one part of Storwize V7000 Unified is a storage server and the other part is a file server, therefore it is called 'Unified'.

### **Storwize V7000 Unified storage subsystem - the Storwize V7000**

Storwize V7000 Unified uses internal storage to generate and provide logical volumes to storage clients, thus acting as a storage system. It is capable of the virtualization of external storage systems as well.

The storage subsystem of Storwize V7000 Unified consists of the hardware and software of the IBM Storwize V7000 storage system (Storwize V7000). Initially, it runs the SAN Volume Controller (SVC)/ Storwize V7000 code level 6.4 (at the time of writing).

The storage subsystem is used for:

- ▶ The provision of logical volumes to external storage clients
- ▶ The provision of logical volumes to the internal storage clients, the File Modules

### **Storwize V7000 Unified file server subsystem - the File Modules**

In addition to providing logical volumes, Storwize V7000 Unified is used to provide access to file system space and thus to files residing in these file systems. It utilizes file sharing protocols/file access protocols and file transfer/file copy protocols, thus acting as a file server.

The file server subsystem of Storwize V7000 Unified consists of two IBM Storwize V7000 File Modules (File Modules or FM). These File Modules perform the functions of the IBM Scale Out Network Attached Storage (SONAS) software, initially running code level 1.4 (at the time of writing).

The File Modules of Storwize V7000 Unified are internal storage clients of Storwize V7000 Unified. They use the logical volumes provided by the Storwize V7000 to save files and share files to file clients. The base operating system (OS) of the File Modules is RedHat 6.1. They use a distributed file system, the IBM General Parallel File System (GPFS), to store and retrieve files. To make the content of the GPFS accessible by file clients, the File Modules use the file sharing protocols / file access protocols NFS, SMB, FTP, HTTPS, SCP and SFTP.

A high level system diagram of Storwize V7000 Unified with a virtualized storage system, a storage client and a file client is shown in Figure 3-1.

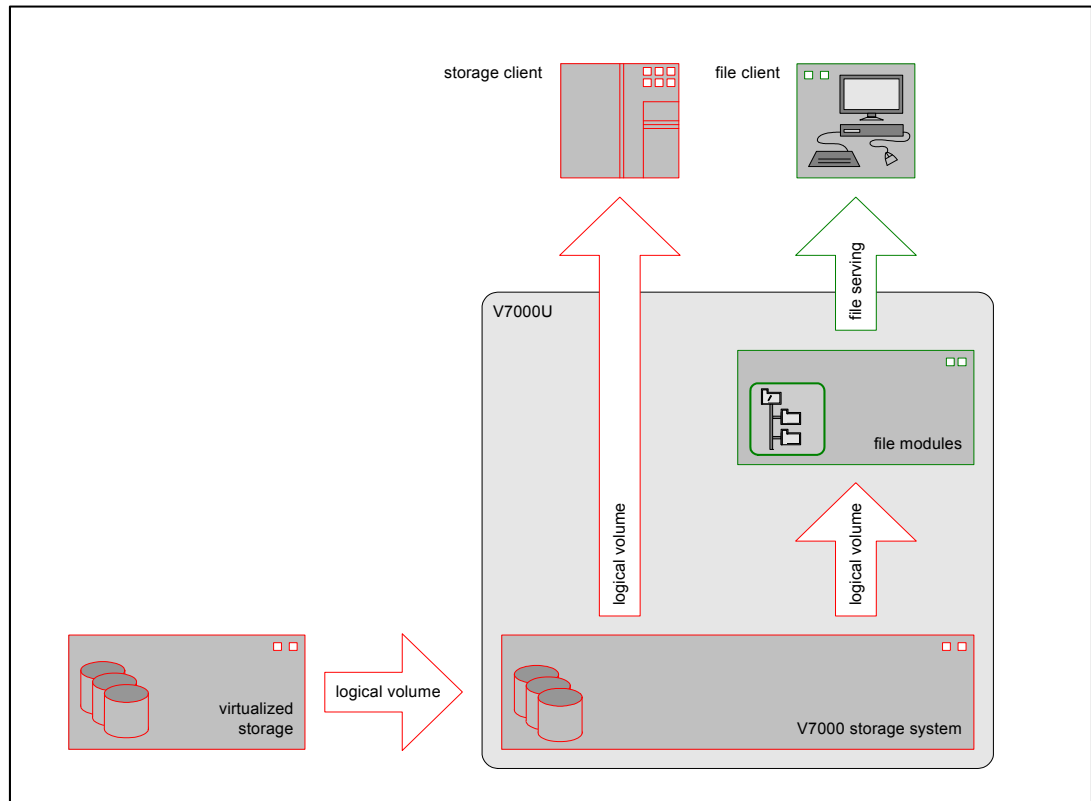


Figure 3-1 Storwize V7000 Unified, high level system diagram

## 3.2 Storwize V7000 Unified system configuration

Storwize V7000 Unified consist of a Storwize V7000 controller enclosure, 0-9 Storwize V7000 expansion enclosures (storage server subsystem), two File Modules (file server subsystem), and the inter-connecting cables.

### 3.2.1 Storwize V7000 Unified storage subsystem configuration

The Storwize V7000 controller enclosure houses two controller canisters, the redundant power supplies (764W with battery backup) and the internal drive bays of which there are two types:

- ▶ Enclosure for 12 3.5" drives
- ▶ Enclosure for 24 2.5" drives

The expansion enclosures contain two Switched Bunch Of Drives (SBOD) canisters of either type (which can be mixed freely) and two 500W PSUs. The Storwize V7000 supports up to 120 3.5" drives, 240 2.5" drives or a mixture of both. Storwize V7000 Unified supports the same dual-port 6Gb Serial Attached SCSI (SAS) HDDs and SSDs as the Storwize V7000:

- ▶ 200, 400 GB 2.5" E-MLC SSD
- ▶ 146, 300 GB 2.5" 15k SAS
- ▶ 300, 600, 900 GB 2.5" 10k SAS
- ▶ 2, 3 TB 3.5" 7200 SAS
- ▶ 1TB 2.5" 7200 SAS

The Storwize V7000 subsystem of Storwize V7000 Unified is used to create virtual volumes and provides them as logical volumes to storage clients. The protocols used are FC and iSCSI.

The interfaces on the two controller canisters are:

- ▶ Eight 2/4/8 Gb FC ports, fitted with short wave transceivers, SFP+.
  - Four ports for external connectivity to FC storage clients.
  - Four ports for internal connectivity to the File Modules.
- ▶ Four 1 GbE ports for external connectivity to iSCSI storage clients and for management (at least one 1 GbE port of each Storwize V7000 controller canister must be connected to the client network).
- ▶ Four 10 GbE ports for connectivity to iSCSI storage clients. This is optional, it requires a Host Interface Module (HIM) in each Storwize V7000 controller canister.
- ▶ Four 4x 6Gb SAS connectors for up to 5 expansion enclosures in the first SAS chain and for up to 4 expansion enclosures in the second SAS chain.

### 3.2.2 Storwize V7000 Unified file server subsystem configuration

The file server subsystem of Storwize V7000 Unified consists of two File Modules. They are System x® servers x3650M3.

Details for one File Module are:

- ▶ Form factor: 2U
- ▶ Processor: Single Four Core Intel Xeon C3539 2.13GHz, 8G L3 cache
- ▶ Cache per controller: 8GB, 16GB or 64GB
- ▶ Storage: Two 600 GB 10K SAS drives, RAID 1
- ▶ Power Supply Units: Two (redundant), 675 W

The interfaces on one File Module are:

- ▶ Four 1 GbE ports.
  - Two ports for external connectivity to file clients and file level remote copy.
  - Two ports for the management network between the FM for SONAS clustering.
- ▶ Two ports 10 GbE for external connectivity to file clients and file level remote copy.
- ▶ Two ports 8Gb FC, one port is internally connected to each Storwize V7000 controller canister.

The internal and external interfaces of the V7000k, including the optional 10 GbE interfaces on the Storwize V7000, are shown in Figure 3-2 on page 25.

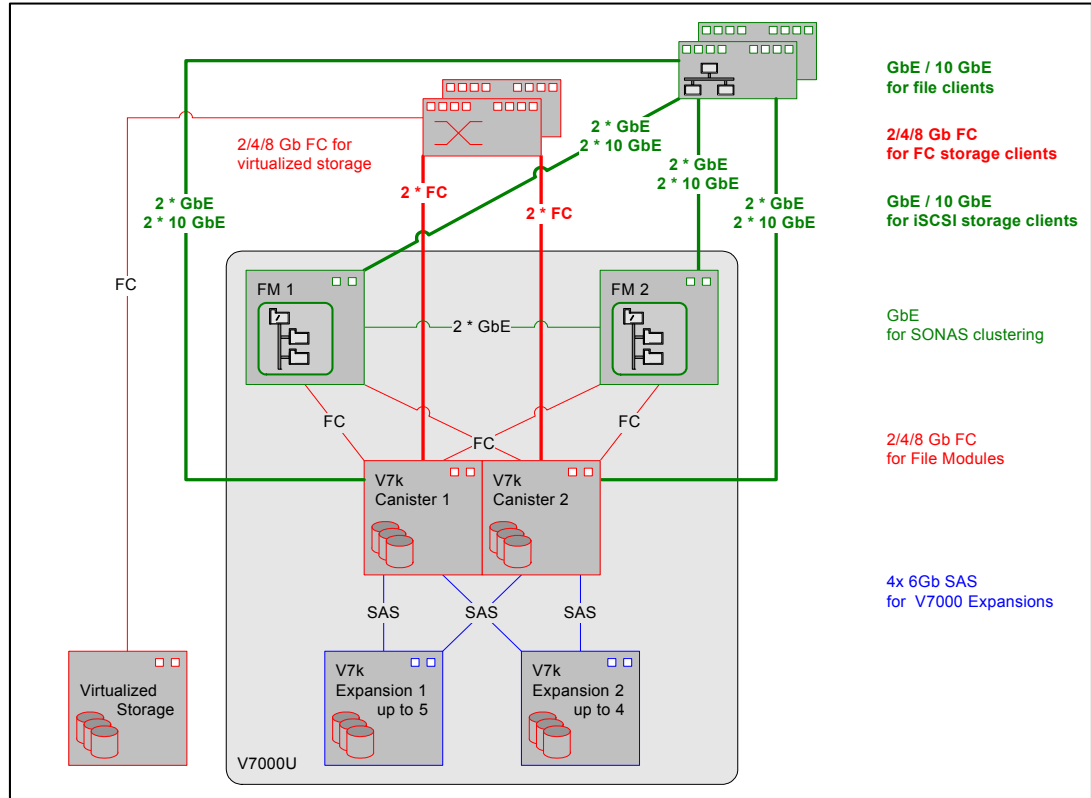


Figure 3-2 Storwize V7000 Unified, internal and external interfaces

### 3.3 Storwize V7000 Unified storage functions

Storwize V7000 Unified provides the storage service by mapping logical volumes to storage clients.

The storage functions implemented in and supported by Storwize V7000 Unified are the same as in the Storwize V7000 release 6.4. The exception is the clustering of two or more Storwize V7000, which is not supported. The Storwize V7000 uses RAID to protect against drive failures for the internal SAS attached storage. The other storage functions of the Storwize V7000 can be used on both internal and external virtualized storage systems.

The protocol used to obtain access to external storage systems is the Fibre Channel Protocol (FCP). The protocols used by storage clients to access the logical volumes mapped to them are FC and iSCSI. The externally connected FC ports on the Storwize V7000 controllers work as initiator and target, the iSCSI ports act as target only. Storwize V7000 Unified supports up to 256 connections to storage clients per node with each, FC and iSCSI.

For storage system virtualization, the Storwize V7000 supports up to 128 storage pools (MDisk Groups). The extent sizes for these pools may be configured to be between 16MB and 8GB. System can manage  $2^{22}$  extents. For example, with 16 MB extent size, the system can manage up to  $16 \text{ MB} \times 4,194,304 = 64 \text{ TB}$ . The maximum number of volumes for use by the file modules and for external storage clients is 2048. Volumes may be in striped, sequential or image mode (for online volume migration). Volume mirroring can be used to protect against failures of MDisk Groups, for instance two external virtualized storage systems. Thin provision can be used with 64KB or 256KB grain size. Easy Tier® hybrid storage pools with two tiers

may be used (SSD and HDD). The automatic extent level migration works based on access history in the previous rolling 24 hour period.

For an in depth discussion of the storage functions of Storwize V7000 Unified refer to *Implementing the IBM Storwize V7000 V6.3*, SG24-7938.

### 3.4 Storwize V7000 Unified file serving related functionality

Additionally to storage, Storwize V7000 Unified provides file sharing and file transfer services, broadly called file serving. The protocols implemented in Storwize V7000 Unified are described in general in Chapter 2, “Terminology and file serving concepts” on page 7. Here we list specifics of the software architecture, operational characteristics, and components of Storwize V7000 Unified SONAS software.

The file related functionality of IBM Storwize V7000 Unified is provided by the SONAS software which runs on the file modules. The file modules integrate the functions of the SONAS software model with the SONAS interface node, SONAS storage node, and SONAS management node functions running on the same hardware.

- ▶ The interface node provides the means for file serving capabilities (NFS, SMB, HTTPS, FTP, SCP and SFTP).
- ▶ The management node provides the interface for configuring, administering and monitoring the SONAS system.
- ▶ The storage node provides the connection to the storage.

Both file modules of Storwize V7000 Unified work together as a cluster, working in parallel to provide the functions of the SONAS software.

The SONAS software provides multiple elements and integrated components that work together in a coordinated manner to provide the file related services to clients. A basic overview of the software components running on Storwize V7000 Unified file modules is shown in Figure 3-3 on page 27.

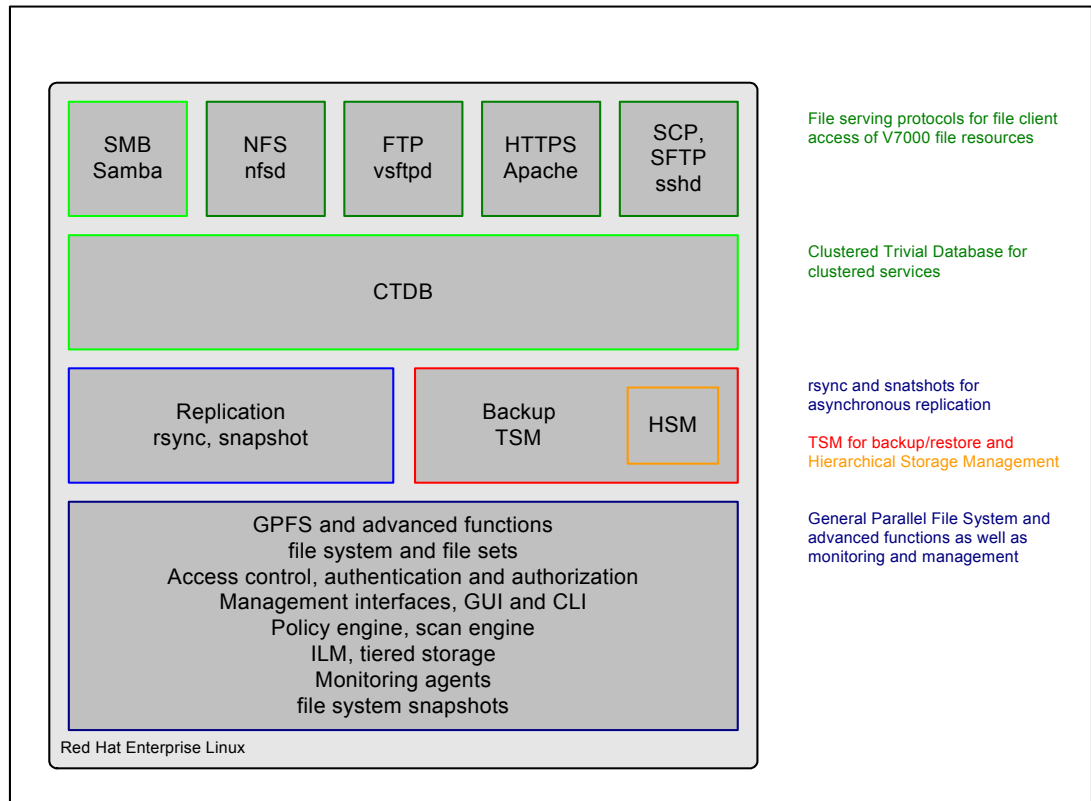


Figure 3-3 Basic overview of Storwize V7000 Unified file modules software

The software running on Storwize V7000 Unified file modules, SONAS, provides integrated support of policy-based automated placement and subsequent tiering and migration of data. We can provision storage pools and store file data according to its importance to the organization or according to performance requirements. For example, we can define multiple storage pools with various drive types and performance profiles. We can create a higher performance storage pool with fast drives and define a less expensive (and lower performance) storage pool with higher capacity Nearline drives. Sophisticated policies are built into SONAS which can transparently migrate data between pools based on many characteristics, such as capacity threshold limits and age of the data for use in a Information Lifecycle Management (ILM) strategy.

The SONAS software supports remote replication, point-in-time copy (file system-level snapshots), and automated storage tiering, all managed as a single instance within a global name space. Asynchronous replication is specifically designed to cope with connections that provide low bandwidth, high latency, and low reliability. The asynchronous scheduled process will pick-up the updates on the source Storwize V7000 Unified system and write them to the target Storwize V7000 Unified system using snapshots and the rsync tool.

### 3.4.1 Storwize V7000 Unified file sharing and file transfer protocols

The network file sharing protocols / file transfer protocols that are supported by Storwize V7000 Unified today are NFS, SMB, FTP, SCP, SFTP and HTTPS. Storwize V7000 Unified uses GPFS to organize the data and save it to the storage of the Storwize V7000 part of the system. The SONAS cluster manager is used to provides cross-node and cross-protocol locking services for the file serving functions in NFS, SMB, FTP, SCP, SFTP and HTTPS. The

SMB file sharing function maps semantics and access control to the POSIX based GPFS with native NFSv4 ACL.

### 3.4.2 Storwize V7000 Unified NFS protocol support

The NFS protocol functionality of Storwize V7000 Unified is provided by nfsd. The following characteristics apply to the NFS implementation of Storwize V7000 Unified:

- ▶ Supports NFSv2 and NFSv3
- ▶ Support of normal NFS data access functions.
- ▶ Supports client system authentication through NFS host lists.
- ▶ Supports authorization through POSIX bits and enforcement of ACLs.
- ▶ Supports reading and writing of POSIX bits.
- ▶ Supports the NFSv3 advisory locking mechanism.
- ▶ Semi-transparent node failover if the application supports network retry.

Storwize V7000 Unified implements NFSv4 ACLs, regardless of the actual file serving protocol used. This provides the strength of the NFSv4 ACLs even to clients that access Storwize V7000 Unified by the NFSv2, NFSv3, SMB, FTP, and HTTPS protocols.

#### NFS protocol limitations and considerations

The flexibility of the NFS protocol export options allows for potentially unsafe configurations. Adhering to good practice guidelines reduces the potential for data corruption.

- ▶ NFSv4 is not supported.
  - SecureNFS is not supported.
- ▶ Do not mount the same NFS export on one client from both Storwize V7000 Unified File Modules as data corruption might occur.
- ▶ Do not mount the same export twice on the same client.
- ▶ Do not export both a directory and any of its subdirectories from a server if both are part of the same file system.
- ▶ Do not export the same file system, or the same file, through multiple exports to the same set of clients.
- ▶ A client should never access the same file through two different server:export paths. The client cannot distinguish that the two objects are the same, so write ordering is not possible and client-side caching will be affected.
- ▶ Do not export the same file system, or the same file, through multiple exports to the same set of clients. A client should never access the same file through two different server:export paths. The client cannot distinguish that the two objects are the same, so write ordering is not possible and client-side caching will be affected. In the Storwize V7000 Unified system, each export is assigned a new file system ID even if the exports are from the same file system. This could lead to data corruption, which is why it is not good practice.

While the use of nested mounts on the same filesystem is strongly discouraged, it is possible to create nested mounts using Storwize V7000 Unified system. If nested mounts are configured on Storwize V7000 Unified system, it is the customer's responsibility to exercise extreme caution to avoid any possibility of corruption.

POSIX ACLs for NFSv2/v3 are not supported on Storwize V7000 Unified system. Clients should only mount Storwize V7000 Unified NFS exports using an IP address. Do not mount a Storwize V7000 Unified NFS export using a DNS Resource Record entry name. If you mount



a Storwize V7000 Unified NFS export using a host name, ensure that the name is unique and remains unique as this restriction prevents data corruption and data unavailability.

When an NFS client detects an NFS server change, such as an NFS server reboot or a new NFS server assuming NFS server responsibility from the previous NFS server, while writing data asynchronously, the NFS client is responsible for detecting whether it is necessary to retransmit data and for retransmitting all uncommitted cached data to the NFS server if retransmission is required.

Storwize V7000 Unified system failover is predicated on this expected client behavior. For example, when an NFS client is writing data asynchronously to one of Storwize V7000 Unified File Modules, if the other File Module assumes the NFS server role, the NFS client must detect the server change and retransmit all uncommitted cached data to the other File Module to ensure that all of the data is safely written to stable storage.

### 3.4.3 Storwize V7000 Unified SMB protocol support

The SMB protocol functionality of Storwize V7000 Unified is provided by an implementation of Samba and it is clustered using the Clustered Trivial Data Base (CTDB).

Samba is a software package made available under the GNU General Public License (GPL) aimed to provide file sharing functions to SMB file clients (for instance such as Windows systems). Samba provides the following functions:

- ▶ Name resolution.
- ▶ Access control through authentication and authorization.
- ▶ Service announcement for browsing of resources.
- ▶ File sharing and print queue sharing.

SMB protocol access in SONAS has been explicitly tested from SMB file clients running Microsoft Windows (2000, XP, Vista 32-bit, Vista 64-bit, 2008 Server), Linux with SMBClient, Mac OS X 10.5, and Windows 7.

GPFS is a POSIX-compliant Unix style file system. For SBM file clients, Storwize V7000 Unified maps Unix ACLs to Windows access control semantics. A multitude of file access concurrency and cross-platform mapping functions are done by the SONAS software, especially in the SONAS cluster manager. The SONAS / Storwize V7000 Unified implementation for SMB file access includes the following characteristics:

- ▶ File access using SMB is only supported for file systems residing on internal storage of the Storwize V7000 Unified storage part and not for external virtualized storage systems.
- ▶ SMB protocol version support:
  - The SMB 1 protocol is supported.
  - SMB 2.0 and later is not fully supported, see the note in “SMB protocol limitations” on page 30.
- ▶ SMB data access and transfer capabilities are supported with normal locking semantics.
- ▶ Supports consistent locking across platforms.
  - By supporting mandatory locking mechanisms and strict locking.
- ▶ User authentication is provided through Microsoft AD or through LDAP.
- ▶ Consistent central ACLs enforcement across all platforms.
- ▶ ACLs are enforced on files and directories.
  - Can be modified using Windows tools.
- ▶ Semi transparent fail-over if the SMB implementation supports network retry.
- ▶ Supports the win32 share modes for opening and creating files.
- ▶ Supports case insensitive file lookup.
- ▶ Support for DOS attributes on files and directories.
- ▶ Archive bit, ReadOnly bit, System bit and other semantics not requiring POSIX attributes.

- ▶ MS-DOS / 16 bit Windows short file names.
- ▶ Supports generation of 8.3 character file names.
- ▶ Notification support of changes to file semantics to all clients in session with the file.
- ▶ Opportunistic locks and leases are supported, for enabling client side caching
- ▶ Offline or de-staged file support (by the SONAS HSM function through TSM):
  - Offline files will be displayed with the hourglass symbol in the Windows Explorer.
  - Recall to disk is transparent to the application, no additional operation is needed.
  - Windows Explorer can display file properties without the need to recall offline files.
- ▶ SONAS Snapshots are integrated into the Volume Shadow Services (VSS) interface.
  - Allows users with proper authority to recall older file versions from SONAS snapshots.
  - Supports file version history for file versions created by SONAS Snapshots.
- ▶ File time stamps:
  - Created Time stamp:
    - The time when the file was created in the current directory.
    - When the file is copied to a new directory, a new value will be set.
  - Modified Time stamp:
    - The time when the file was last modified.
    - When the file is copied, it will keep the value in the new directory.
  - Accessed Time stamp:
    - The time when the file was last accessed.
    - This value is set by the application, but not all applications modify it.

## SMB protocol limitations

Consider the following SMB protocol limitations when configuring and managing Storwize V7000 Unified system:

Limitations for SMB and SMB2:

- ▶ Alternate data streams are not supported. One example is an NTFS alternate data stream from a Mac OS X operating system.
- ▶ Server side file encryption is not supported.
- ▶ Level 2 opportunistic locks (oplocks) are currently not supported. This means that level 2 oplock requests are not granted.
- ▶ Symbolic links cannot be stored or changed and are not reported as symbolic links, but symbolic links created via NFS will be respected as long they point to a target under the same exported directory.
- ▶ SMB signing for attached clients is not supported.
- ▶ SSL secured communication to Active Directory is not supported.
- ▶ Storwize V7000 Unified acting as Distributed File System (DFS) root is not supported.
- ▶ Windows Internet Name Service (WINS) is not supported.
- ▶ Retrieving Quota information using NT\_TRANSACT\_QUERY\_QUOTA is not supported.
- ▶ Setting Quota information using NT\_TRANSACT\_SET\_QUOTA is not supported.
- ▶ Managing Storwize V7000 Unified system using the Microsoft Management Console Computer Management Snap-in is not supported, with the following exceptions:
  - Listing shares and exports
  - Changing share or export permissions
- ▶ Users can traverse through each directory only. Listing of files in the directory must still be permitted by the ACL (READ). It solves the problem where you have the directory structure "/A/B/C" , and you have read permissions on "C" but no permissions on "A" and "B". By introducing BypassTraversalCheck you can access C without having "EXECUTE/SEARCH" permissions on A/B, but you are still not allowed to browse the content of A and B.

If the value is "yes", IBM SONAS will grant "SEARCH/EXECUTE" access for the directory tree. Allowed values are "yes" or "no". The default is "yes". Setting the

--bypassTraversalCheck option allows a user to directly access files and folders that the

user owns, and that are contained under parent folders for which the user does not have Read or Write permissions. Users without Read and Execute access to the share or export in which the user-owned files and folders are located can Read and Modify the files inside the export for which the user has permissions granted by the **--bypassTraversalCheck option**. However, in this case, operations like Rename file and Delete file are not granted by default. This is normal CIFS behavior. Modify ACLs as required to enable these operations.

**SMB 1 specific limitations:**

- ▶ CIFS Extensions for UNIX are not supported.

**SMB 2 specific limitations:**

- ▶ SMB 2.1 is not supported.
  - Storwize V7000 Unified system does not grant durable or persistent file handles.

**Storwize V7000 Unified FTP support**

Storwize V7000 Unified provides FTP access from FTP clients using vsftpd. The following characteristics apply:

- ▶ Supports file transfer to and from any FTP client.
- ▶ Supports user authentication through AD and LDAP.
- ▶ Supports enforcement of ACLs and retrieval of POSIX attributes.
  - ACLs cannot be modified using FTP as there is no support for the chmod command.
- ▶ Supports FTP resume for clients which support network retry on node fail over.
- ▶ Characters for file names and directory names are UTF 8 encoded.

**Storwize V7000 Unified HTTPS support**

Storwize V7000 Unified supports simple read only file transfer of files through the HTTPS protocol from any HTTP client using Apache. All transfers are using HTTPS to provide access control. The following features are supported through HTTPS:

- ▶ Supports read only file transfer of appropriately formatted files.
- ▶ Supports user authentication through AD and LDAP
- ▶ Supports enforcement of ACLs.
  - ACLs cannot be viewed nor modified through HTTPS.
- ▶ On node fail-over during a file transfer, the transfer is cancelled.
  - It must be retried on the other file module.
  - Partial retrieve is supported, minimizing duplicate transfers in a fail-over situation.
- ▶ Characters for file names and directory names are UTF 8 encoded.

The SONAS software uses HTTP aliases as the vehicle to emulate the share concept. For example, share XYZ will be accessible by <https://server.domain/XYZ>.

The Web-based Distributed Authoring and Versioning (WebDAV) and the Representational State Transfer (REST) API are not supported at this time in Storwize V7000 Unified, they are known requirements.

**Storwize V7000 Unified SCP and SFTP support**

The Storwize V7000 supports the transfer of files between an SCP client and Storwize V7000 Unified using sshd. All the default options implemented in this protocol are supported. Also, SFTP is available by sshd to transfer files in a manner similar to using FTP.

**Storwize V7000 Unified locking characteristics**

POSIX byte range locks set by NFS clients are stored in GPFS, and Windows clients accessing Storwize V7000 Unified using the SMB protocol honor these POSIX locks. The

mapping of SMB protocol locks to POSIX locks is updated dynamically on each locking change.

Unless the application specifically knows how to handle byte range locks on a file or are architected for multiple concurrent writes, concurrent writes to a single file are not desirable in any operating system. To maintain data integrity, locks are used to guarantee that only one process can write to a file (or to a byte range in a file) at a time. Although file systems traditionally locked the entire file, newer ones such as GPFS support the ability for a range of bytes within a file to be locked. Byte range locking is supported for both, the SMB protocol and the NFS protocol, but this does require the application to know how to exploit this capability.

If another process attempts to write to a file (or a section of one) that is already locked, it will receive an error and will wait until the lock is released.

Storwize V7000 Unified supports the standard DOS and NT file system (deny-mode) locking requests, which allow only one process to write to an entire file at a give time, as well as byte range locking. In addition, Storwize V7000 Unified supports the Windows locking known as opportunistic locking or oplock.

SMB protocol byte range locks set by Windows SMB file clients are stored both in the SONAS interface node cluster wide database, and by mapping them to POSIX byte range locks in GPFS. This mapping ensures that NFS file clients see relevant SMB protocol locks as POSIX advisory locks, and NFS file clients honor these locks.

### 3.4.4 Storwize V7000 Unified SONAS cluster manager

The SONAS cluster manager is a core SONAS component that coordinates and orchestrates SONAS functions and advanced functions such as the byte range locking available in GPFS. It coordinates the access of the various file sharing protocols to GPFS files. The SMB file sharing function maps to SMP protocol semantics and access control to the POSIX based GPFS and NFSv4 ACLs.

The SONAS cluster manager provides the clustered implementation and management of the file modules (interface node function) including tracking and distributing record updates across both file modules in the cluster. It controls the public IP addresses used to publish the file services, and moves them as necessary between the file modules. By monitoring scripts, the SONAS cluster manager monitors and determines the health state of the file modules. If a file module has a problem, such as a hardware failure, or a software failure the SONAS cluster manager will dynamically migrate the affected public IP addresses and in-flight workloads to the other file module. It uses the “tickle-ACK” method with the affected clients, so that they re-establish the TCP connection to the other file module. With this method acknowledgement packets get exchanged which allows for the remaining file module to send a proper reset packet so that clients know the connection to the original file module is to be reset. Otherwise clients would time out after a possibly very long time.

The SONAS software works in an active-active, high-available and workload-sharing manner with the clustering functionality provided by the SONAS cluster manager. If a file module fails, the SONAS software will automatically fail over the workload to the remaining file module. From a workload allocation standpoint, SONAS uses the Domain Name System (DNS) to perform round-robin access, to spread workload as equally as possible on an IP address basis across the file modules.

The SONAS allocates a single network client to one file module. The SONAS software does not rotate a single client’s workload across file modules. That is not only unsupported by DNS or the SMB protocol, but would also decrease performance, because caching and

read-ahead is done in the file module. It is for this reason that any one individual client is going to be assigned, for the duration of their session, to one file module.

One of the primary functions of the SONAS cluster manager is to support concurrent access from concurrent users, spread across multiple various network protocols and platforms to many files. SONAS software also supports, with proper authority, concurrent read and write access to the same file, including byte-range locking. Byte-range locking means that two users can access the same file concurrently, and each user can lock and update a subset of the file.

We see that all file accesses from the users to GPFS will logically traverse the SONAS cluster manager. It logically implies that the cluster manager will handle metadata and locking, but will not handle data transfer; in other terms, the cluster manager is not in-band in regards to data transfer.

The CTDB functionality provides important capabilities for the SONAS cluster manager to provide a global name space to all users from any file access protocol, in which both file modules appear as a single file server. The CTDB also assures that all the SONAS SMB components on both file modules are able to talk to each other in a high performance manner, and update each other about the locking and other information.

### 3.4.5 Storwize V7000 Unified product limits

Consider the following limits when configuring and managing Storwize V7000 Unified system.

- ▶ Non-Storwize V7000 Unified application installation is not supported on Storwize V7000 Unified File Modules.
- ▶ The number of shares and exports that can be created is limited to 1000 per protocol.
- ▶ If naming a share using the CLI, you can use up to 80 characters, however the GUI limits you to 72 characters.
- ▶ Restricting ports by VLAN, service, or other criteria, is not possible.
- ▶ VLAN 1 is not supported for Storwize V7000 Unified client traffic.
  - This restriction is intended to prevent security exposure and reduce the probability of network configuration errors. VLAN 1 has been used within the industry as the default or native VLAN. Many vendors use VLAN ID value 1 for management traffic by default. Configuring VLAN 1 as available within the network can be a security exposure because VLAN 1 might span large parts of the switched network by default. Common practice in the industry strongly discourages the use of VLAN 1 for user client traffic. Setting VLAN 1 for user client traffic can require very explicit steps that differ by vendor and can be prone to configuration error.
- ▶ The access control lists (ACL) READ\_NAMED, WRITE\_NAMED and SYNCHRONIZE have no effect on Storwize V7000 Unified system.
- ▶ Task scheduling has the following limitations:
  - Only schedules in equal space increments of three hours are supported.
  - Space increments other than three hour increments are not supported.
- ▶ It is strongly suggested that only the UTF-8 character set is selected when connecting to Storwize V7000 Unified CLI via SSH, to Storwize V7000 Unified GUI via browser, or when connecting to Storwize V7000 Unified shares and exports via NFS and FTP:
  - By selecting UTF-8 encoding in the SSH client for the connection to Storwize V7000 Unified CLI.

- By selecting UTF-8 as locale for the connection to Storwize V7000 Unified in the FTP client.
- All Storwize V7000 Unified internal scripts and tools currently use `LANG=en_US.UTF8`, and handle file names and directory names as though they contained only UTF-8 characters. Users can create files and directories using different locales; for example, by using an external linux client that is set to `LANG=ISO-8859-1`, `LANG=is_IS` or `LANG=de_DE`, or DBCS locales like `LANG=euc_JP`. Storwize V7000 Unified kernel NFS daemon simply treats file and directory names as a stream of bytes, so by using NFS mounts, you can theoretically copy those files and directories into Storwize V7000 Unified system. Storwize V7000 Unified kernel NFS daemon is not aware of locales, and therefore could copy files or directories with non-UTF-8 characters into Storwize V7000 Unified system.
- UTF-8 uses the most-significant bit to encode characters that are not in the ASCII character set, which includes only characters with hexadecimal values 0x01-0x7f; decimal values 1-127. The UTF-8 encoding enforces that if one byte in a file or directory name is greater than hexadecimal 0x7f, that a second, and maybe a third, and maybe a forth, byte must follow to complete a valid character. Therefore, files and directories that are created in a non-UTF-8 locale that have such a byte greater than 0x7f in their name would be invalid when interpreted as UTF-8.
- The command-line interface (CLI) command input, the graphical user interface (GUI), and some output such as messages and log entries, currently require UTF-8 format only.
- Multibyte character set (MBCS) support is limited to file and directory names. MBCS characters in object names (for example, user names) are not supported.
- Current limitations:
  - Non-UTF-8 characters in file and directory names are not displayed correctly in the CLI, the GUI, messages or log entries.
  - Non-UTF-8 file and directory names can only be read from clients that have the same language setting. For example, if an NFS client defined as ISO-8859-1 is used to create a file, a CIFS client or a different NFS client using UTF-8 cannot see or access that file.
  - Non-UTF-8 file and directory names cannot be backed up or restored.
  - Non-UTF-8 file and directory names cannot be specified when using Storwize V7000 Unified CLI, which interprets characters only as UTF-8. Attempting to restore a file name that contains a non-UTF-8 character would not restore the file with that file name because the byte representation is different.
  - Non-UTF-8 file and directory names might cause problems in other Storwize V7000 Unified areas, including asynchronous replication, backup, and file access methods such as FTP, HTTPS, SCP, SMB and NFS.
  - Non-UTF-8 file and directory names might get represented differently in different locales. Some locales might not even be able to represent the byte combination at all, might treat the filenames as invalid, and might not process them correctly, if at all.
  - Object names using multi-byte non-UTF-8 characters might be limited to as few as 25% of the maximum number of characters that are supported for the names of the same object that are composed of only 1-byte UTF-8 characters.
  - A directory that contains non-UTF-8 characters in its path cannot be the root of a share or export, or of a file set.

- The following characters cannot be backed up or restored: chr(24), chr(25) and newline.
- Wild cards and double quotes are NOT officially supported for backup. If you require that those characters be backed up, contact your IBM representative.

Note that:

- ▶ Windows Service for Unix (SFU)/Subsystem for UNIX-based Applications (SUA) NFS does not handle non-ASCII UTF-8 file names correctly.
- ▶ Internet Explorer (IE) version 8 does not correctly display non-ASCII characters in FTP file names and directory names that use UTF-8 for file name encoding. Such files or directories cannot be accessed. For example, IE does not correctly parse FTP directory listings that contain space characters within user or group names.









## Access control for file serving clients

In this chapter we discuss access control to resources for file serving clients. Access control is a broad term about controlling who (user) or what (system) is granted access to which resources and can have many criteria. Two key concepts used to control access are authentication and authorization.

Authentication is a means to provide and verify credentials to ensure the identity of a user. Authorization is a means to grant access to a specific service or specific resources to a user, usually after successful authentication.

## 4.1 Authentication and authorization in general

The objective of authentication is to verify the claimed identity of users and components. Authentication methods include, for example, unique user IDs, keys and digital certificates. As the first process, authentication provides a way of identifying a user, typically by having the user provide his credentials before access is granted. This can be done by entering a valid user name and valid password. Typically the process of authentication is based on each user having a unique set of criteria for gaining access. The authentication server compares a users authentication credentials with other user credentials stored in a database. If the credentials match, the user is deemed to have identified successfully. If the credentials do not match, authentication fails.

After a successful authentication the user may get access to the service or resource based upon the authorization associated to him. Authorization may be based upon a users UID and matching UID in ACLs or other means of mapping a specific user to a specific resource or service.

### 4.1.1 UNIX authentication and authorization

Unix authentication is system based. Granting access to resources within the system or for shared resources being accessed from the system, the authorization, is UID/GID based. A user uses their user name and a credential (for instance a password or a private SSH key) to logon on a Unix system. The system looks up the users UID in local files or an external directory service such as LDAP and then verifies the received credential. The information for credential verification might be stored locally (for instance hashes of passwords are stored in `/etc/shadow`, public SSH keys are stored in `.ssh/authorized_keys`) or the external directory service (for instance LDAP).

Once a user has successfully logged on to a Unix system, they are trusted (authenticated) on this system but also by all other systems which trust the particular system the user just logged on. For example, for file sharing via the NFS file sharing protocol, a NFS file server administrator creates an NFS export and grants access to the users *system*. For NFS file access, the Unix NFS client running on the users system sends the users UID with each file access request and the NFS service running on the NFS file server considers this UID as authenticated, assuming that the system correctly authenticated the users UID. The UID is not used for *another specific* authentication by the NFS server, but for the *authorization* of the users access to file resources. For instance, if a users system is authenticated but the UID of the user does not match the UID for resources on the NFS server system, the user still has no access to the resources of the NFS server since he is not *authorized* to access them.

This means, to be able to access remote NFS resources, the user's system must be authenticated to the remote server and the user must be authenticated to the local system and authorized to access its resources (by a successful logon to a system) *and* the UID (the user's systems NFS client provides) must match a valid UID on the NFS servers system for specific resources.

### 4.1.2 Windows authentication and authorization

Windows authentication and authorization is session based. A user logs on using their user name and password on the Windows system. The system looks up the user's SID (Security Identifier) in the local Windows registry or on the Windows domain controller, for example an Active Directory Server, and then verifies the received credential. The information for

credential verification might be stored locally (in the Windows registry) or on the external Windows domain controller (the aforementioned Active Directory Server).

Once a user is successfully logged on to a Windows system they are trusted on this system, but the user still must authenticate and authorize to other services provided by other network connected systems before they can access them. For instance, to access shares via the SMB protocol, an SMB file server administrator creates an SMB share and customizes the ACL (for authorization) of this share to grant the user access to the share.

For access to SMB shares, the SMB client running on the user's system sends an authentication request to an authentication server (for instance Active Directory). The authentication server checks if the requesting user is allowed to use the service and then returns a session credential (which is encrypted with the key of the SMB server) to the SMB client. The SMB client sends the session credential to the SMB server. The SMB server decrypts the session credential and verifies its content. Once the session credential is verified, the SMB server knows that the user was authenticated and authorized by the authentication server.

### **4.1.3 UNIX and Windows authentication and authorization in the Storwize V7000 Unified**

To provide heterogeneous file sharing for Unix and Windows, the Storwize V7000 Unified must support the authentication methods for Unix and Windows as described above. The Storwize V7000 Unified uses Windows authentication for incoming SMB connection requests and Unix authentication for incoming NFS, HTTP, SCP and SFTP requests.

## **4.2 Methods used for access control**

Depending on the dominating operating system environment, the size of the infrastructure and other variables, different methods to enforce access control are employed.

### **4.2.1 Kerberos**

Kerberos is a network authentication protocol for client server applications by using symmetric key cryptography. User password in clear text format is never sent over the network. The Kerberos server grants a ticket to the client for a short span of time. This ticket is used by the client of a service while communicating with the server to get access to the service for instance access to SMB file server shares. Windows Active Directory authentication is based on Kerberos. MIT Kerberos is a free implementation of Kerberos protocol, which is provided by the Massachusetts Institute of Technology.

### **4.2.2 User names and User IDs**

Unix system and Unix based appliances such as the Storwize V7000 Unified use user names and user identifiers (UID) to represent users of and to the system. The user name is typically a human readable sequence of alphanumeric characters and the UID is a positive integer value. When a users logs on to a UNIX system, the operating systems looks up the UID and then uses this UID for further representation of the user.

User names, UIDs, and the mapping of user names to UIDs are stored locally in the `/etc/passwd` file or on an external directory service such as AD, LDAP or NIS.

### 4.2.3 Group names and GIDs

Unix systems use groups to maintain sets of users which have the same permissions to access certain system resources. Similar to user names and UIDs, a Unix system also maintains group names and group identifiers (GID). A Unix user can be member of one or more groups, where one group is the primary or default group. Unix groups are not nested, they contain users only but not other groups.

Group names, GIDs, the mapping of group names to GIDs, and the memberships of users to groups are stored locally in the `/etc/group` file or on an external directory service such as AD, LDAP or NIS. The primary group of an user is stored in `/etc/passwd` or in an external directory service.

### 4.2.4 Resource names and security identifier (SID)

Windows refers to all operating system entities as resource, including users, groups, computers and other resources. Each resource is represented by a so called security identifier (SID). Windows groups may be nested, for instance one group may include one or more users and or more groups. Resource names and SIDs are stored locally in the Windows registry or in an external directory service such as Active Directory or LDAP.

### 4.2.5 UID/GID/SID mapping in the Storwize V7000 Unified

The Storwize V7000 Unified stores all user data in the GPFS file system which uses UIDs and GIDs for authorization. For SMB share access, the Storwize V7000 Unified needs to map SIDs to UIDs and GIDs to enforce access control. NFS clients send the UID and GID of a user which requests access to a file. The Storwize V7000 Unified uses Linux default access control mechanism by comparing the received UID and GID with the UIDs and GIDs stored in GPFS.

The UIDs and GIDs used by the NFS clients must match the UIDs and GIDs stored inside GPFS. There is a requirement to allow the remapping of external UIDs and GIDs used by the NFS client to different UIDs and GIDs stored on GPFS.

For HTTP, SFTP and SCP access, the Storwize V7000 Unified requires users to authenticate via a user name. The Storwize V7000 Unified needs to map the user name to one UID and one or more GIDs for GPFS access control.

When SMB clients using Windows connect to the Storwize V7000 Unified, it first contacts the Active Directory to check for username and password combination. The UID/GID pair created is then stored in the `idmap` database in the Storwize V7000 Unified. At the first time a user logs in, the `id` mapping is created. After that, it is picked up from the database directly.

For NFS access from Unix clients, the UID is provided by the Unix clients itself. In case of mixed access from Windows and Unix, Active Directory with SFU can be used.

### 4.2.6 Directory services in general

Storing user and group information in local files works well for small organizations which operate only a very few servers. Whenever a user is added or deleted, the group membership is changed, or a password is updated, this information must be updated on all servers. Storing this information in local files does not scale for large organizations having many users which need selected access to many servers and services.

Directory services allow you to store and maintain user and group information centrally on an external server. Servers look up this information in the directory server instead of storing this information in local files.

#### **4.2.7 Windows NT 4.0 Domain Controller / SAMBA Primary Domain Controller**

A domain is a concept introduced in Windows NT where a user may be granted access to a number of computer resources with the use of user credentials. A Domain Controller (DC) is a server that responds to authentication requests and controls access to a variety of computer resources. Windows 2000 and later versions introduced Active Directory, which largely eliminated the concept of primary and backup domain controllers. Primary Domain Controller (PDC) are still used by customers: The SAMBA software can be configured as primary domain controller and customer run SAMBA on Linux, acting as PDC. The Samba4 project has the goal to run SAMBA as an AD server.

#### **4.2.8 Lightweight Directory Access Protocol (LDAP)**

The Lightweight Directory Access Protocol (LDAP) is directory service access protocol using TCP/IP. LDAP is a lightweight alternative to the traditional Directory Access Protocol (DAP), therefore it is called LDAP.

An LDAP directory is usually structured hierarchically, as a tree of nodes. Each node represents an "entry" within the LDAP database. A single LDAP entry may consists of multiple attribute value pairs, and is uniquely identified by a distinguished name.

#### **4.2.9 Microsoft Active Directory (AD)**

Active Directory (AD) is a Microsoft created technology introduced from Windows 2000 onwards and provides the following network services:

- ▶ Directory Service, based upon LDAP
- ▶ Authentication service, based upon Kerberos
- ▶ Domain Name System (DNS) function

#### **4.2.10 Services for UNIX (SFU) and Identity Management for Unix**

Services for Unix (SFU) is a Microsoft Windows component for Windows Server 2003 with AD and Identity Management for Unix is used instead in Windows Server 2008 with AD, which provides interoperability between Microsoft Windows and Unix environments. The Storwize V7000 Unified uses it primarily for UID/GID/SID mapping.

#### **4.2.11 Network Information Service (NIS)**

Network Information Service (NIS) is a directory service protocol for centrally storing configuration data of a computer network. NIS protocols and commands were originally defined by Sun Microsystems, the service is now widely implemented. Originally called Yellow Pages or YP, some of the binary names still start with yp. The original NIS design was seen to have inherent limitations, specifically in the areas of scalability and security. Therefore modern and secure directory systems, primarily LDAP are used as alternative. The NIS information are stored in so called NIS maps, typically providing the following information:

- ▶ Password related data similar to data stored in /etc/passwd
- ▶ Group related data similar to data stored in /etc/group

- Network configuration such as netgroups

#### 4.2.12 Access control list (ACL) in general

Generally, an access control list (ACL) is a list of permissions which is attached to a resource. An ACL describes which identities are allowed to access the respective resource (for instance read, write, execute). ACLs are the built in access control mechanism of Unix and Windows systems. SONAS uses a Linux built in ACL mechanism for access control to files which are stored on GPFS.

#### 4.2.13 GPFS NFSv4 ACLs

There are a broad range of ACL formats which differ in syntax and semantics. The ACL format defined by NFSv4 is also called NFSv4 ACL. GPFS ACLs implement the NFSv4 style ACL format which is sometimes referred as GPFS NFSv4 ACL. The Storwize V7000 Unified stores all user files in GPFS. The GPFS NFSv4 ACLs are used for access control of files stored on the Storwize V7000 Unified. The implementation of NFSv4 ACLs in GPFS does not imply that GPFS or the Storwize V7000 Unified support NFSv4. NFSv4 support of the Storwize V7000 Unified is planned for a future release.

#### 4.2.14 POSIX bits

The POSIX bits of a file are a way to specify access permissions to files. Unix file systems allow you to specify the owner and the group of a file. The POSIX bits of a file allow to configure access control for the owner, the group and for all other users to read, write to or execute the file. POSIX bits are less flexible than ACLs.

The change of the POSIX bits of a GPFS file system will trigger a modification of its GPFS NFSv4 ACL. Since the Storwize V7000 Unified uses GPFS NFSv4 ACLs for access control, the Storwize V7000 Unified administrators and IBM service personnel should never change the POSIX bits of files stored in GPFS.

#### 4.2.15 ACL mapping

GPFS NFSv4 ACLs and Windows ACLs are not compatible. For instance, Windows supports unlimited nested groups which is not fully supported by GPFS NFSv4 ACLs. The Storwize V7000 Unified maps Windows ACLs on a best fit basis to GPFS NFSv4 ACLs which results in some limitations. It is a known limitation that in this aspect the current Storwize V7000 Unified is not fully compatible with Windows SMB file sharing.

## 4.3 Access control with Storwize V7000 Unified

The authentication configuration of the Storwize V7000 Unified consists of two elements, the configuration of a directory service and the refinement of the ID mapping. In the Storwize V7000 Unified implementation it is essential to define an initial owner when creating a share. Only this owner has initial access from a file client and can start to define directory structures and associated ACLs for all other designated users of this share. It can not be displayed and listed afterwards, and cannot be changed if there is any data stored on the share.

### 4.3.1 Authentication methods supported

The Storwize V7000 Unified supports the following authentication methods:

- ▶ Microsoft Active Directory (AD)
- ▶ Microsoft Active Directory (AD) and Services for Unix (SFU) and Identity Management for Unix
- ▶ Microsoft Windows NT 4.0 Domain Controller / SAMBA Primary Domain Controller (PDC)
- ▶ Lightweight Directory Access Protocol (LDAP)
- ▶ Lightweight Directory Access Protocol (LDAP) with MIT Kerberos
- ▶ Network Information Service (NIS)

The Storwize V7000 Unified uses the following other authentication elements within the methods supported:

- ▶ Netgroups: It is a groups of systems used to restrict access for mounting NFS exports on a set of systems and deny mounting on rest of the systems. The Storwize V7000 Unified supports netgroup being stored in NIS.
- ▶ Kerberos: The Storwize V7000 Unified supports Kerberos with AD (mandatory) and LDAP (optional).
- ▶ Secure Sockets Layer / Transport Level Security (SSL / TLS): These protocols are primarily used to increase the confidentiality and integrity of data being send over the network. These protocols are based on public-key cryptography and uses Digital Certificate based on X509 for identification.

With the Storwize V7000 Unified, we can only configure one authentication method, for instance AD, at one time. The authentication server is external to the Storwize V7000 Unified. The authentication server needs to be configured separately and the Storwize V7000 Unified GUI or CLI will not provide any means to configure or manage the external authentication server. This is true even for the Kerberos server.

The Storwize V7000 Unified provides server side authentication configuration for various protocols, which include NFS, SMB, FTP, SCP, SFTP and HTTP. For NFSv3, only the protocol configuration is performed, Kerberos has a few special steps to be performed on the V7000 for NFSv3, though. Since authentication happens on the NFSv3 client side, authentication needs to be configured on the client side mainly.

It is required that the V7000 is synchronized in time with the authentication servers. Authentication will not work if time is not synchronized. Authentication configuration does not ensure this and hence this needs to be ensured manually.

### 4.3.2 AD authentication

To use Active Directory, the V7000 must be configured for and joined to the Active Directory domain, this will automatically create the required computer account in the AD. The “public



clustername" specified during installation is used as the computer account name. File sharing protocol access should always be done with this name (in order for Kerberos to work). Authentication is provided for all supported file access protocols except NFS. AD with SFU must be configured for access via NFS for Windows Server 2003 with AD. On Windows Server 2008 with AD, the Identity Management for Unix has to be enabled in AD.

### 4.3.3 AD with SFU authentication or with Identity Management for Unix

AD with SFU or with Identity Management for Unix is the right choice for customers with the following conditions:

- ▶ Customer uses Windows Server 2003 or Windows Server 2008 with AD to store user information and user passwords.
- ▶ Customer plans to use NFS.
- ▶ Customer plans to use asynchronous replication.

The primary Windows group assigned to a AD user must have a GID assigned. Otherwise the user will be denied access to the system. Each user in AD must have a valid UID and GID assigned in order to be able to mount and access exports. SFU should not be added, once data is stored on the Storwize V7000 Unified.

The primary unix group setting in active directory will not be used by the Storwize V7000 Unified. The Storwize V7000 Unified will always use the primary Windows group as the primary group for the user. This results in new files and directories created by a user via the SMB protocol being owned by their primary windows group and not by the primary UNIX group. For this reason, it is recommended that the Unix primary group to be the same as the Windows primary group defined for the user.

Once data is stored on the Storwize V7000 Unified with AD, it is difficult to add Services for Unix (SFU) later on, because the UIDs and GIDs used internally by GPFS must match the UIDs and GIDs stored in SFU. If already conflicting UIDs and GIDs are stored in SFU this is not possible so clients should configure the Storwize V7000 Unified with AD and SFU from the very beginning.

The ACLs copied from the source Storwize V7000 Unified system to the target Storwize V7000 Unified system include UIDs and GIDs of the source Storwize V7000 Unified system which are inconsistent with the UIDs and GIDs of the target Storwize V7000 Unified system.

To enable all NFS users of more than one Storwize V7000 Unified to access all systems with the same identity, the authentication schema must be changed to UID/GID in AD. Existing users will have to be edited (mapping from SID to UID/GID) and new users need to get the UNIX user information added while creating the user in AD.

### 4.3.4 SAMBA PDC authentication

NT4 / SAMBA PDC is the old domain controller concept used by Windows NT and Windows 2000. This is not supported by Microsoft any more. To support this method, the open source / Samba community developed the Samba PDC.

### 4.3.5 LDAP authentication

LDAP can be used in environments where Windows and UNIX clients are being used. The Storwize V7000 Unified supports LDAP with Kerberos for SMB protocol access. LDAP is not supported for secured NFS, FTP, HTTP, SCP protocol access.



### 4.3.6 Network Information Service

NIS is used in UNIX based environments for centralized user and service management. NIS is used for keeping user, domain and netgroup information. Netgroup is used to group client machine IP / hostname which can be specified while creating NFS exports. NIS is also used for user authentication for services like ssh, ftp, http and more. However in the V7000, we do not support NIS as authentication method. In the Storwize V7000 Unified we use NIS for netgroup support and ID mapping. We use the NIS default domain to resolve the netgroup even though we support multiple NIS domains. The NIS client configuration needs server and domain details of the NIS server.

Three different modes of NIS configuration are supported.

- ▶ NIS for netgroup and AD or PDC/NT4 for authentication and AD increment ID mapping.
  - Used for mixed environments where there will be Windows and Unix users. In this mode The Storwize V7000 Unified supports both the SMB and the NFS protocol and netgroups.
- ▶ NIS with ID mapping as extension to AD or Samba PDC/NT4 and netgroup support.
- ▶ Plain NIS without any authentication, just for netgroup support (only NFS).

## 4.4 Access control limitations and considerations

The following limitations for authentication and authorization apply.

### 4.4.1 Authentication limitations

Consider the following authentication limitations when configuring and managing the Storwize V7000 Unified system.

For Active Directory (AD) with the Services for Unix (SFU) UID/GID/SID mappings extension:

- ▶ Enabling SFU for a trusted domain requires a two-way trust between the principal and the trusted domain.
- ▶ To access the Storwize V7000 Unified system, users and groups must have a valid UID/GID assigned to them in Active Directory. Allowed range is between 1 and 4294967295, both inclusive. It is advisable to keep the lower range greater than 1024 to avoid conflict with the CLI users. Invoking the command with lower range less than 1024 will generate a warning message and ask for confirmation. Use the **--force** option to override it.
- ▶ For user access, the primary group on the Storwize V7000 Unified system is the Microsoft Windows Primary group, not the Unix primary group that is listed in the Unix attribute tab in the user's properties. Therefore, the user's primary Microsoft Windows group must be assigned a valid GID.

For Active Directory (AD) with the Network Information Service (NIS) mappings extension:

- ▶ Since Unix style names do not allow spaces in the name, the following conventions for mapping Active Directory users and groups to NIS are implemented:
  - Convert all upper case characters to lower case characters.
  - Replace every space character with the underscore character. For example, an Active Directory user named **CAPITAL** Name has the corresponding name **capital\_name** on NIS.

- If Active Directory is already configured on the Storwize V7000 Unified system, you can only use the **--idMapConfig** option of the **cfgad** Storwize V7000 Unified CLI command to change the high value of the range, and the high value of the range can only be changed to a higher value. You cannot change the high value of the range to a lower value. You cannot change the low value of the range, and you cannot change the range size.

For example, if you used the **cfgad** Storwize V7000 Unified CLI command with the **--idMapConfig** option to configure Active Directory specifying a value for the **--idMapConfig** option as 3000-10000:2000, you can only use the **cfgad** Storwize V7000 Unified CLI command with the **--idMapConfig** option to increase the value 10000 for the high value of the range. You cannot decrease the value of 10000 for the high value of the range. You cannot change the value 3000 for the low value of the range, and you cannot change the value 2000 for the range size.

- To change from NIS ID mappings to Active Directory ID mappings, or to change the ID mapping parameters of an already existing Active Directory configuration by using the **--idMapConfig** option of the **cfgad** Storwize V7000 Unified CLI command, either to change the low value of the range, decrease the high value of the range, or change the range size, you must perform the following steps in the following sequence:

1. Submit the **cleanupauth** Storwize V7000 Unified CLI command and do not specify the **--idmapDelete** option.
2. Submit the **cleanupauth** Storwize V7000 Unified CLI command and do specify the **--idmapDelete** option.
3. Submit the **cfgad** Storwize V7000 Unified CLI command with the options and values that you want for the new Active Directory configuration.

If you do not perform the above steps in sequence, results are unpredictable and can include loss of data access.

- UIDs and GIDs less than 1024 are denied access for the FTP, SCP and HTTPS protocols for all of the supported authentication schemes other than Active Directory with SFU.
- Authentication configuration commands stop and restart the services SMB, NFS, FTP, SCP and HTTPS. This action is disruptive for clients that are connected. Connected clients lose their connection, and file operations are interrupted. File services resume a few seconds after an authentication configuration command completes.

## 4.4.2 Authorization limitations

When managing authorization, the following Storwize V7000 Unified system implementation details apply:

- When a child file or child directory is created, the ACL that a file is initially assigned depends on the ACL type, the file system settings, and the ACL of the parent directory. Depending on these variables, the results in GPFS might be slightly different than in Microsoft Windows. For example, if the parent directory is set to have two ACLs, for example, full access for owner and for everyone, the Windows default is to create two ACLs for the child: allow full access for owner and allow full access for everyone. The GPFS default creates six ACLs: allow and deny ACLs for owner, group and everyone.
- The special permissions Write Data/Create File and Create Folder/Append Data cannot be set separately for files. If either of these permissions is set, both are set. Enabling one always enables the other, and disabling one always disables the other. For directories, they can be set separately as long as these access control entries (ACE) are not inherited by files. You can configure two separate ACEs, where the ACE that is inherited by files has both special permissions enabled or both disabled, and another ACE that is inherited by directories where one of the above special permissions is enabled and the other disabled.

In this case, the Apply onto field of the Permission Entry panel can contain the following values:

- This folder only.
- This folder and subfolders.
- Subfolders only.

If you attempt to specify the values This folder, subfolders and files, This folder and files or Files only, the following security pop-up message is displayed: Unable to save permission changes on folder. The parameter is incorrect.

- ▶ The BypassTraversalCheck privilege that can be used on Windows servers is not supported on the Storwize V7000 Unified system. To read the content of a subdirectory, a user must not only have READ permission in the ACL of this subdirectory, the user must also have traversal permission (SEARCH in Windows, execute in POSIX) for all of the parent directories. You can set the traverse permission in the everyone group ACE at the share root, and inherit this privilege to all subdirectories.
- ▶ ACL management can be done through NAS protocols by an authorized end user.
- ▶ The default ACL on a file system root directory (700 root root) prevents end users from accessing the file system.

ACL inheritance stops at file set junction points; new file sets will always have the default ACL (700 root root).

For security reasons, creating an export does not allow the setting of an owner for existing directories. The owner can only be changed if the directory is empty. Once a directory contains files or directories or linked file sets, you cannot change the owner when creating an export.

If the owner option is omitted when creating an export for a nonexistent directory, the directory is created and inherits the ACL if inheritance is configured. If ACL inheritance is not configured, a new export for a nonexistent directory is assigned the default ACL (700 root root).

Note that if the directory is empty, the owner can be changed by deleting and recreating the export with the owner option.

Using POSIX commands such as **chmod** overwrite any previous ACLs and create a new ACL with entries only for owner, group and everyone.

When you create an export you can create multiple levels of directories if they do not yet exist, but if multiple directory levels are created the owner is set only for the leaf directory. All other directories are created as owner root and can only be accessed if an appropriate ACL inheritance has been previously configured.

SONAS automatically creates owner, group and everyone special entries to the ACL to support interoperability with NFS, unlike Microsoft Windows which does not use these special ACL entries. An inherited ACL might look different than the parent ACL because the owner and group entries have changed. Other ACL entries are not affected by this special behavior.

Only the root user can change ownership to a different user.





# Storage Virtualization

This chapter contains a general description of virtualization concepts and provides an explanation about storage virtualization in Storwize V7000 Unified system.

## 5.1 User requirements driving storage virtualization

In today's environment there is an emphasis on a smarter planet and dynamic infrastructure. Thus, there is a need for a storage environment that is as flexible as the application and server mobility.

In a non-virtualized storage environment, every system is an "island" that needs to be managed separately. The storage virtualization helps to overcome this and improves the management process of different storage systems.

You can see the importance of addressing the complexity of managing storage networks by applying the total cost of ownership (TCO) metric to storage networks. Industry analyses show that storage acquisition costs are only about 20% of the TCO. Most of the remaining costs are related to managing the storage system.

When discussing virtualization it is important to recognize that there are many ways to achieve it, and therefore the best possible option might vary depending on requirements, and because of that the path taken is usually different for each specific environment.

However, there are some overall and general benefits that can address the client concerns and they are described here:

- ▶ Easier administration:
  - Simplified management: the integrated approach means less hardware and software layers to manage separately, and therefore less resource is needed
  - Less management overhead equates to less costs and therefore cost savings
  - Ideally a single level of control for use of advanced functions such as copy services
  - Decoupling usage of advanced functions from, for example, the hardware and technology used as the storage back-end
  - Decoupling usage of multiple, virtualized operating system environments from the actual server and hardware used
- ▶ Improved flexibility:
  - Shared buffer resources: flexible assignment of resources to multiple levels is possible as required, and as needs change over time, there is the potential for automation of this flexible resource assignment
- ▶ Improved resource utilization:
  - Shared buffers plus their flexible assignment can lead to less resources and less resource buffers required
  - Delayed acquisitions with this consolidated approach
  - Cost savings due to consolidated block and file virtualization approach

## 5.2 Storage virtualization terminology

Although *storage virtualization* is a term that is used extensively throughout the storage industry, it can be applied to a wide range of technologies and underlying capabilities. In reality, most storage devices can technically claim to be virtualized in one form or another. Therefore, we must start by defining the concept of storage virtualization as used in this book.

This is how IBM defines storage virtualization:

- ▶ Storage virtualization is a technology that makes one set of resources look and feel like another set of resources, preferably with more desirable characteristics.
- ▶ It is a logical representation of resources that is not constrained by physical limitations:
  - It hides part of the complexity.
  - It adds or integrates new function with existing services.
  - It can be nested or applied to multiple layers of a system.

When discussing storage virtualization, it is important to understand that virtualization can be implemented at various layers within the I/O stack. We have to clearly distinguish between virtualization at the disk layer and virtualization at the file system layer. The focus of this book is virtualization at the disk layer, which is more specifically referred to as block-level virtualization, or block aggregation layer

Virtualization devices can reside inside or outside the data path, and when this is the case they are respectively called in-band/symmetrical or out-of-band/asymmetrical virtualization.

- ▶ Symmetrical: in-band appliance

The device is a SAN appliance that sits in the data path, and all I/O flows through the device. This kind of implementation is also referred to as *symmetric virtualization* or *in-band*.

The device is both target and initiator. It is the target of I/O requests from the host perspective, and the initiator of I/O requests from the storage perspective. The redirection is performed by issuing new I/O requests to the storage. The SVC/Storwize V7000 uses symmetrical virtualization.

- ▶ Asymmetrical: out-of-band or controller-based

The device is usually a storage controller that provides an internal switch for external storage attachment. In this approach, the storage controller intercepts and redirects I/O requests to the external storage as it does for internal storage. The actual I/O requests are themselves redirected. This kind of implementation is also referred to as *asymmetric virtualization* or *out-of-band*.

Figure 5-1 shows variations of the two virtualization approaches.

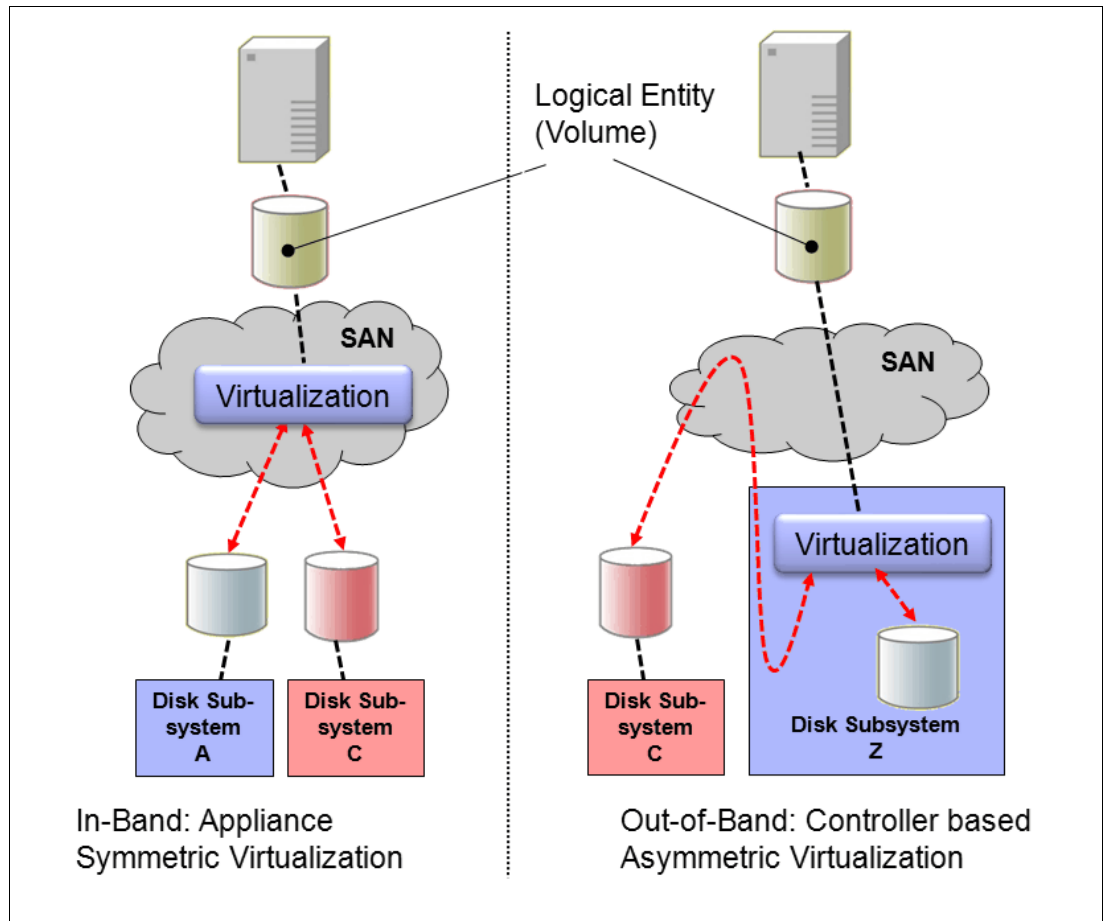


Figure 5-1 Overview of block-level virtualization architectures

In terms of the storage virtualization concept, the focus in this chapter is on block-level storage virtualization in a symmetrical/in-band solution.



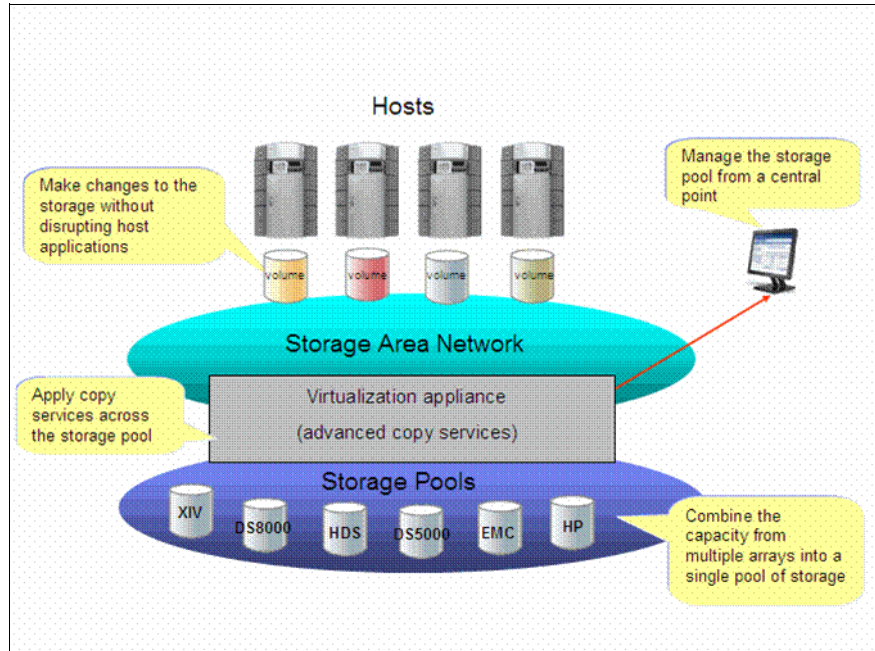


Figure 5-2 In-band Storage Virtualization on the Storage Network layer

The IBM Storwize V7000 Unified inherits its fixed block-level storage virtualization features entirely from the SVC and Storwize V7000 product family. These features can be used independently of the filesystem or enhancements built into the Storwize V7000 Unified based on the software stack that runs on the two file modules.

### 5.2.1 Realizing the benefits of Storwize V7000 Unified storage virtualization

Storwize V7000 Unified system has the ability to manage external storage arrays as well as its own storage.

Managing external storage in the V7000 Unified system reduces the number of separate environments that need to be managed down to a single environment and provides a single interface for storage management.

Moreover, advanced functions such as mirroring and FlashCopy® are provided in this system so there is no need to purchase them again for each new disk subsystem.

Migrating data from external storage to the Storwize V7000 Unified system can be easily done due to the virtualization engine this system offers, by connecting the external storage array to existing LUN on the V7000 Unified system and copy the data with data migration procedure.

In addition, free space does not need to be maintained and managed within each storage subsystem, which further increases capacity utilization

### 5.2.2 Using internal physical disk drives in the Storwize V7000 Unified

The Storwize V7000 Unified recognizes its internal physical disk drives as *drives* and supports RAID arrays.

The *RAID array* could be comprised of different numbers of physical disk drives depending on the RAID level chosen, although in general it can be up to 16 drives in one RAID array. This

RAID array/MDisk is then added to a *Pool* layer which manages performance and capacity, whereby multiple MDisk could belong to one pool. In these pools the logical storage entity, the *volumes*, are created (which reside in the pool and by default are striped across all the MDisk in a pool whereby the stripesize is known as an *extent*). These volumes are then mapped to external hosts to provide storage capacity to them and are seen by the hosts as a SCSI disk (which is a volume with underlying special properties and capabilities, such as two independent copies with volume mirroring, or a much smaller real capacity in the case of a thin-provisioned volume).

These are the logical steps required on V7000 to set up the volumes and make them accessible to a host:

First you need to select the physical drives, create the RAID array (once done this array is now an MDisk), then add this MDisk to a Pool, then create the Volumes in this Pool, and then map the volumes to the external host.

**Note:** In the Storwize V7000 Unified these are the steps required for external hosts attached via iSCSI or via an FC SAN to the Storwize V7000 component. It works the same for the volumes used by the File Modules to host the data for file systems and enable host access via file protocols, but these volumes are not displayed in the standard volume panels/GUI screens just as the File Modules are not displayed as hosts in the standard host panels in the GUI. The volumes used for file systems are only visible in the file system context and the file system related GUI screens.

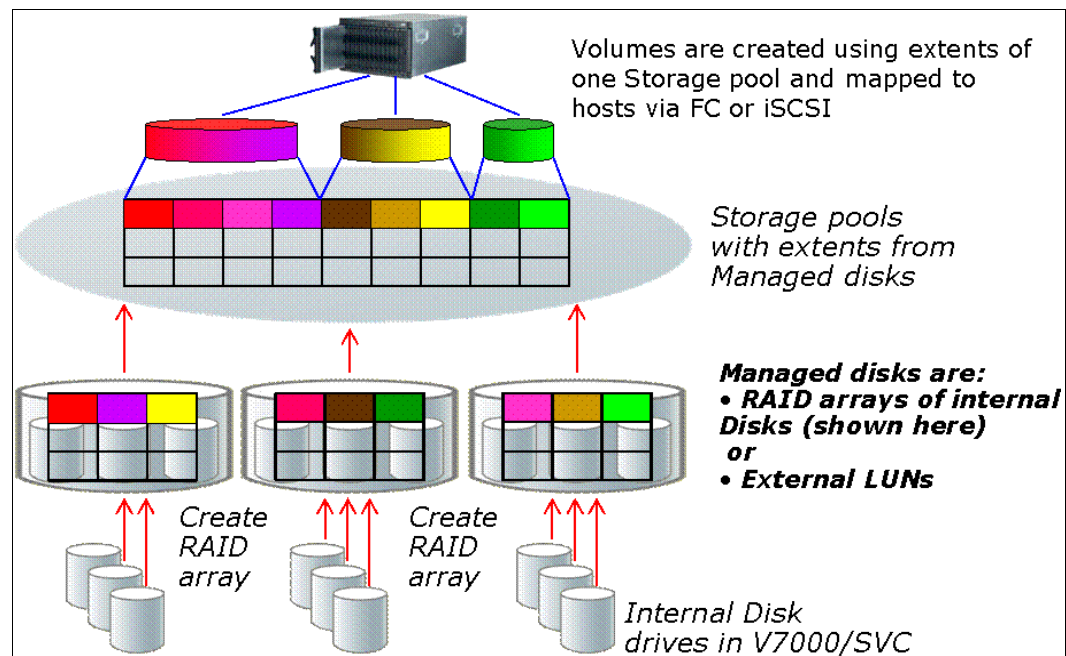


Figure 5-3 Virtualization layers using internal disks

### 5.2.3 Using external physical disk drives in the Storwize V7000 Unified

Generally there is a dependency on the specifics of the external storage subsystem being used and the features that are built in to it.

That being the case these steps provide an overview of how to use external SAN attached storage. Our only assumption is that the storage subsystem provides RAID functionality for data protection against physical disk drive failures.

In the external SAN attached storage system (storage controller), there will be internal logical devices created and presented/mapped to the Storwize V7000 as logical volumes, LUNs or RAID arrays and this depends on the capabilities of the specific storage system. These logical entities will be recognized by the Storwize V7000 as MDisks directly, which will then be added to storage pools in the same fashion as for MDisks based on Storwize V7000 internal disks as described previously. Next, volumes are created in these storage pools and mapped to external hosts, and again this is done in exactly the same way as before.

This is shown in Figure 5-4.

These are the steps required on the external storage system/controller are as follows:

Select the physical disks, create the RAID array, then map the entire array to the Storwize V7000, or create logical volumes within the array and map these to the Storwize V7000.

They will be recognized as MDisks in the Storwize V7000 and treated accordingly in the next logical configuration steps.

To set up the volumes on the Storwize V7000 and make them accessible to a host, you will need to detect the MDisks from the external controller, then add the MDisks that are found to the storage pools, create volumes in the pools and then map the volumes to the external hosts connected via iSCSI or FC.

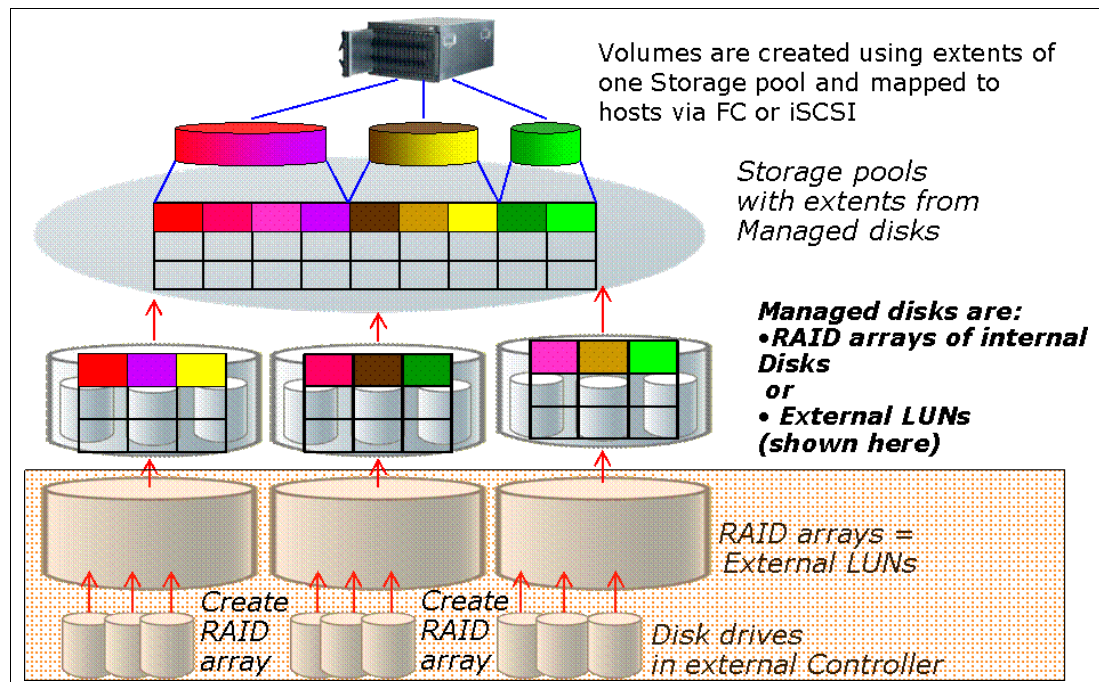


Figure 5-4 Virtualization layers using external disks

In the previous sections we have described how the different layers interact with each other to provide the volume entity that is to be presented to the connected hosts.

In a Storwize V7000 Unified configuration the file system storage uses its own set of volumes created in Storwize V7000 storage pools.

The Storwize V7000 storage pools could be shared flexibly between volumes providing capacity for fixed block hosts attached via iSCSI or FC, and volumes providing capacity for file systems, which are then made available to network attached hosts (file clients) by means of exports/file shares created.

And, vice-versa, the file system storage uses its own set of volumes created in Storwize V7000 pools, where the pools could be shared between volumes for fixed block I/O and file access as well.

**File system:** Creating a file system for CIFS protocol used is not supported. Only NFS and block level devices are supported.

## 5.3 Summary

Storage virtualization is no longer merely a concept or an unproven technology. All major storage vendors offer storage virtualization products. Making use of storage virtualization as the foundation for a flexible and reliable storage solution helps enterprises to better align business and IT by optimizing the storage infrastructure and storage management to meet business demands.

The IBM System Storage SAN Volume Controller, the Storwize V7000, and the Storwize V7000 Unified are built upon a mature, sixth-generation virtualization solution that uses open standards and is consistent with the Storage Networking Industry Association (SNIA) storage model. The appliance-based in-band block virtualization process, in which intelligence, including advanced storage functions, is migrated from individual storage devices to the storage network, can reduce your total cost of ownership and improve your return on investment.

At the same time it can improve the utilization of your storage resources, simplify your storage management, and improve the availability of your applications.

Chapter 6.

## **NAS use cases, differences between SONAS and Storwize V7000 Unified**

In this chapter we will build on the NAS methods discussed earlier in Chapter 4, “Access control for file serving clients” on page 37 and describe some typical use cases for Network Attached Storage using the features and functions built into the Storwize V7000 Unified.

As an add on to this chapter, we list the major differences to the IBM Scale Out NAS solution (SONAS) which is built using the exact same software stack which has been adopted for the Storwize V7000 Unified file access component. Therefore the majority of the actual file access methods and functionality built-in is very similar, however since the SONAS hardware is very different there are also some major differences between the two products. One of the major areas is scalability.

## 6.1 Use cases for Storwize V7000 Unified

Here are some examples of use cases which utilize and benefit from the powerful software features built into the Storwize V7000 Unified. There are many other possibilities and most of the options can be combined to build the tailored solution which fits to the individual needs of a customer.

### 6.1.1 Unified Storage with both File and Block access

Sometimes there are requirements for a storage system which can handle both file access and block access at the same time. An additional benefit then is flexibility regarding the storage assignment - being able to move storage capacity between these two access methods as required and as needs might change over time. At the same time with access to the data, simultaneously using both file and block access methods should not interfere with each other in terms of creating performance dependencies.

The Storwize V7000 Unified provides storage to both worlds in one unified system with the flexibility mentioned and is a good solution for these requirements. It provides dedicated interfaces for both methods of data access and allows to shift storage capacity between them. An overview of the different interfaces is shown in Figure 6-1.

- ▶ Storwize V7000 Unified provides file access via the IP network, for example, for SMB/CIFS and NFS exports and file access is handled by the two File Modules
- ▶ Storwize V7000 Unified provides block storage access via IP for iSCSI and block access is handled by the V7000
- ▶ Storwize V7000 Unified provides block storage access via the SAN using Fibre Channel Protocol and block access is handled by the V7000
- ▶ Storwize V7000 Unified offers the flexibility to use separate or shared storage pools between file and block access — in both cases internally using separate volumes for file access versus block access
- ▶ Flexible usage and moving of storage capacity according to changing needs

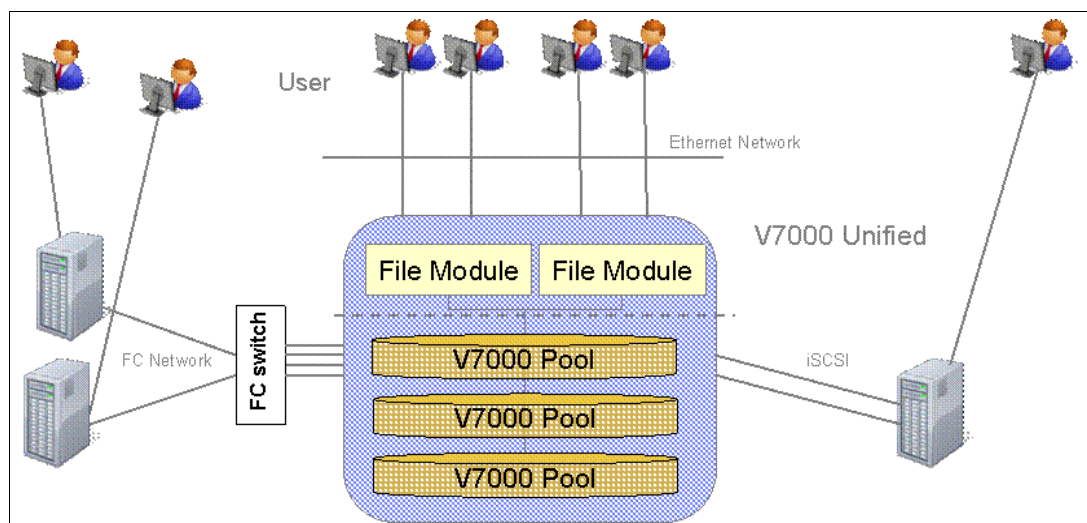


Figure 6-1 Storwize V7000 Unified - unified storage for both file and block access



## 6.1.2 Multi User File Sharing with centralized Snapshots and Backup

In many cases there is a benefit of having a centralized storage solution for multiple users working together, for example, in a project work group. While every user has their own home directory and data to work with, there is also a need to share files between the members of the group and this can be set up with the appropriate share and access control structures. In addition it is more efficient to handle requirements such as data protection and space management at a work group level rather than at an individual user level.

The Storwize V7000 Unified has the appropriate functions to enable centralized management and protection as well as individual data access. It also provides enhanced scalability in a single namespace compared to traditional single NAS file solutions. This is shown in Figure 6-2 on page 59.

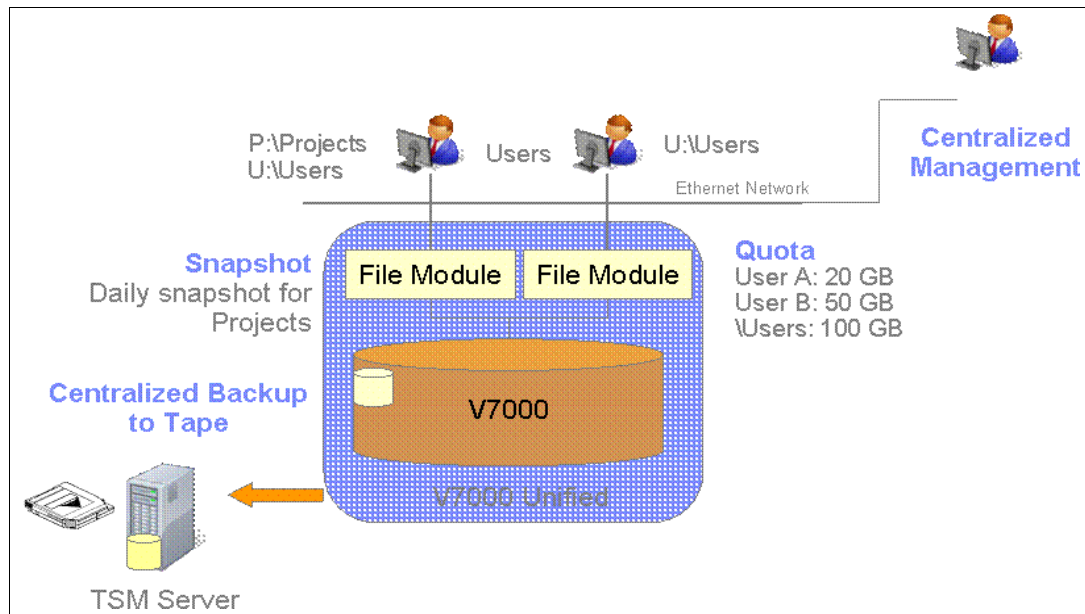


Figure 6-2 Storwize V7000 Unified centralized management for multiple users sharing files

- ▶ Multiple users and/or groups store and share files on a Storwize V7000 Unified
  - Multi user access to a single file is possible using sophisticated GPFS locking mechanisms
  - File sharing between Windows and UNIX client environments is possible
- ▶ Storwize V7000 Unified allows you to set quotas for individual users, groups or at a share level:
  - Providing granular space management as needed
  - Including warning levels - *soft quota* - and hard limits - *hard quota*
- ▶ Storwize V7000 Unified allows for centralized Snapshots, for example, for important data of the entire work group:
  - The administrator uses general Snapshot rules such as scope, frequency and levels of retention, tailored to the needs of, for example, the work group
  - It removes the need for every user to take care of its own data
  - It provides easy recoverability of multiple file versions
- ▶ A centralized backup provides data protection for the entire system, or at a share level meaning:
  - Individual users do not have to define and take care of this themselves
  - More efficient resource usage and scheduling
- ▶ File replication to a second Storwize V7000 Unified system for disaster protection:

- Works at the file system level and protects the entire file system, including all file shares

### 6.1.3 Availability and Data Protection

The Storwize V7000 Unified has multiple built-in options for availability and data protection of important data. Some of these options are illustrated in Figure 6-3 on page 60.

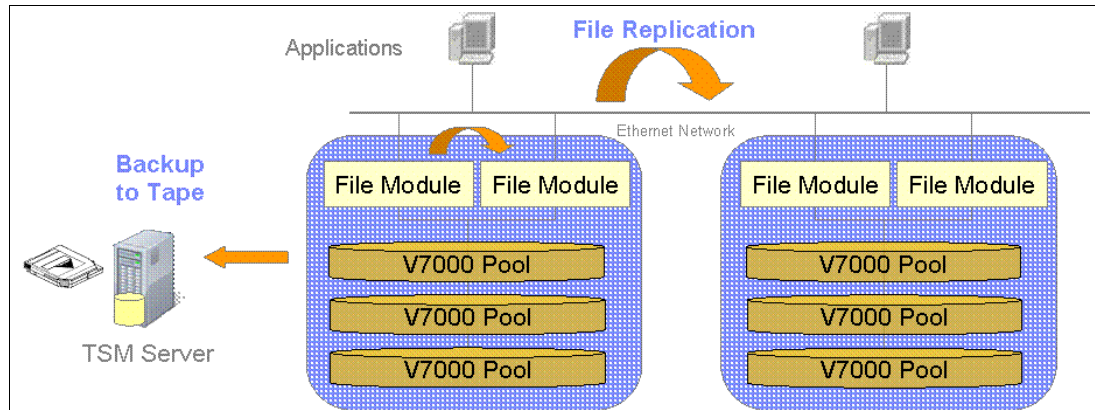


Figure 6-3 Storwize V7000 Unified availability and data protection

- ▶ Clustered File Modules provide high availability and redundancy against failures by using a GPFS based clustered file system to store your data
- ▶ Clustered File Modules are used for load balancing for file access from multiple clients simultaneously
- ▶ The Storwize V7000 Unified provides a highly available storage back-end
- ▶ Data can and should be protected by using backup to tape
- ▶ Disaster protection can be added by using asynchronous replication of files to a distant site

### 6.1.4 ILM, HSM and Archiving solution

It is often a requirement to manage the data placement over the lifetime of a file, and to place/move it to the appropriate storage tier for a cost efficient solution, all preferably in an automated fashion with minimal administrative management efforts (and related costs). At the same time there might be legal requirements to keep certain type of files for compliance, for example for an extended period of time.

The Storwize V7000 Unified is well suited to these requirements and has powerful features built-in. In this regard the policy based ILM and HSM functionality provide a solution to these requirements as shown in Figure 6-4.



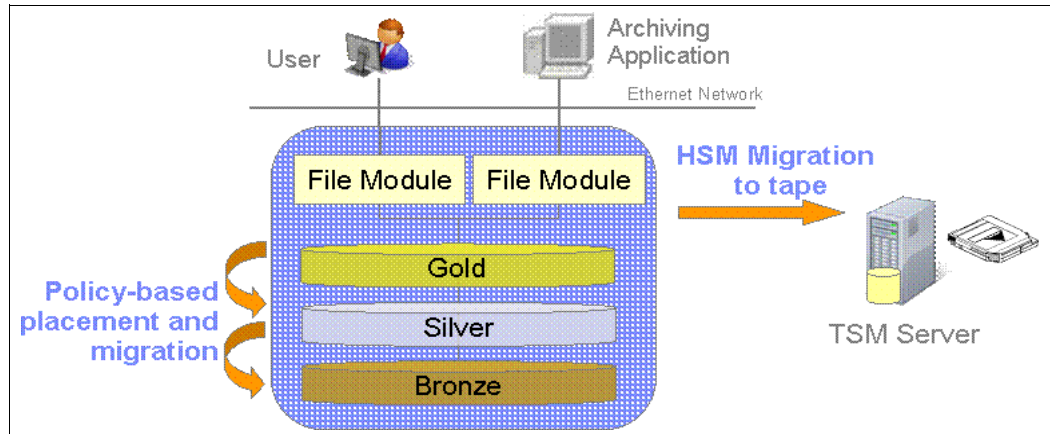


Figure 6-4 Storwize V7000 Unified as ILM, HSM and archiving solution

- ▶ Archive applications and users store data in file shares on the V7000 Unified
- ▶ Policies can be defined for automated data placement, migration and deletion by using:
  - A very powerful SQL-like policy language that is built-in to define these individual, tailored policies as required.
  - Placement policies define the initial placement of a file when it is first created
  - Migration policies are used to move data to its appropriate storage tier over its entire life time
- ▶ Static data could be kept for an extended period of time, for example as defined by legal requirements and deleted automatically
- ▶ HSM could be used to move 'unused data' (according to defined criteria) to tape:
  - This is handled by the TSM HSM server
  - Both migration to external tape and recall of the data on request are fully transparent
  - A *stub file* is left in the file system where the original data resided. If there is a request to access that data again a transparent file recall moves the data back into the file system

## 6.2 Storwize V7000 Unified and SONAS

In order to discuss the differences to Storwize V7000 Unified a better understanding of the SONAS implementation is needed, therefore this chapter starts with a short introduction into the SONAS solution.

### 6.2.1 SONAS brief overview

SONAS is a scale-out NAS implementation, built with a focus on scalability. There is enormous room for growth until the architectural limitations are actually hit, however it still provides a single namespace across the entire configuration. It is a GPFS two-tier implementation, see Chapter 7, "IBM General Parallel File System" on page 65, with separate nodes handling client I/O, known as *Interface Nodes*, while separate nodes provide the Network Shared Disks (NSDs) and handle the back-end storage tasks, and are therefore called *Storage Nodes*. This is illustrated in Figure 6-5.

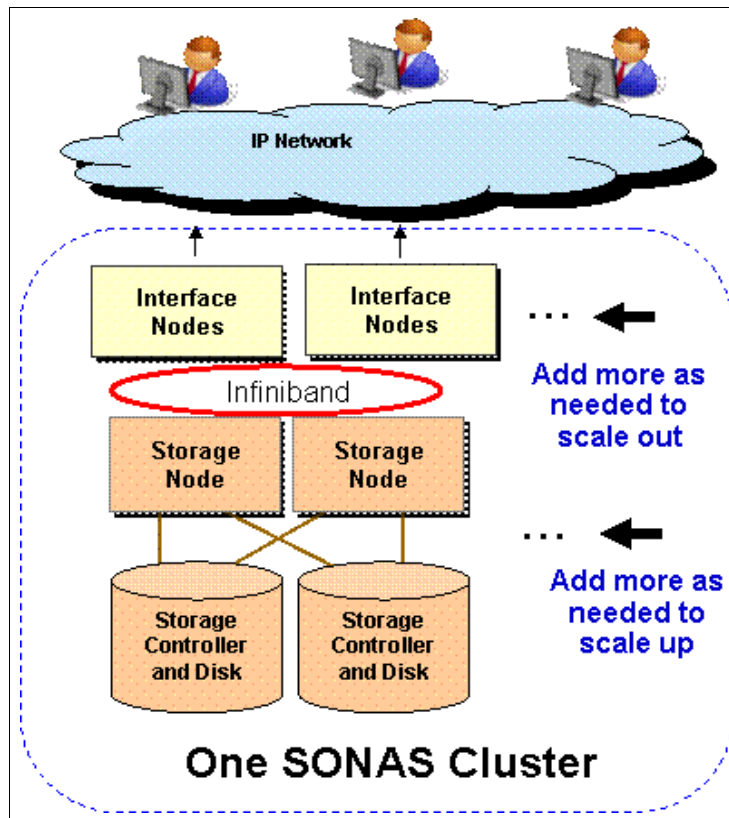


Figure 6-5 SONAS Overview: Two-tier architecture and independent scalability

The scalable high speed connectivity between Interface Nodes and Storage Nodes is implemented via an internal Infiniband network.

The latest SONAS version is 1.4. For SONAS information and administration go to:

[http://pic.dhe.ibm.com/infocenter/sonasic/sonaslic/index.jsp?topic=%2Fcom.ibm.sonas.doc%2Fadm\\_wcache\\_limitations.html](http://pic.dhe.ibm.com/infocenter/sonasic/sonaslic/index.jsp?topic=%2Fcom.ibm.sonas.doc%2Fadm_wcache_limitations.html)

SONAS release 1.3 supports up to:

- ▶ 32 Interface Nodes
  - Providing an aggregated bandwidth of up to 105 GBps
- ▶ 30 Storage Pods
  - Providing up to 7200 disk drives in the storage back-end
  - With 3 TB HDDs this results in a 21.6 PB storage capacity

SONAS uses a centralized management concept with multiple distributed Management Node roles, providing single access using either a Graphical User Interface (GUI) or Command Line Interface (CLI) to the SONAS cluster.

In general, SONAS and Storwize V7000 Unified support the same software features, but there are differences such as the ones listed in 6.2.2, “Implementation differences between the Storwize V7000 Unified and SONAS” on page 63. Therefore make sure to check both products’ support pages as the official references.

- ▶ Support portal for SONAS:

[http://www.ibm.com/support/entry/portal/Overview/Hardware/System\\_Storage/Network\\_Attached\\_Storage\\_%28NAS%29/SONAS/Scale\\_Out\\_Network\\_Attached\\_Storage](http://www.ibm.com/support/entry/portal/Overview/Hardware/System_Storage/Network_Attached_Storage_%28NAS%29/SONAS/Scale_Out_Network_Attached_Storage)

- ▶ Support portal for Storwize V7000 Unified:

<http://www.ibm.com/storage/support/storwize/v7000/unified>

## 6.2.2 Implementation differences between the Storwize V7000 Unified and SONAS

Although both products use the same NAS software stack, there are differences due to the different hardware implementations, scalability and supported software features as well.

The following list reflects differences at a software release for both products and might be subject to change in future releases.

- ▶ GPFS one-tier architecture in Storwize V7000 Unified, two-tier architecture in SONAS
  - SONAS has dedicated, different servers, as interface nodes and storage nodes
- ▶ Scalability: hardware scalability is limited in Storwize V7000 Unified, for example, no independent scalability of resources for interface nodes and storage nodes in Storwize V7000 Unified
  - There is a fixed number of two File Modules in the Storwize V7000 Unified
  - Storage capacity limits in Storwize V7000 Unified: internally one V7000 Control Enclosure and up to nine V7000 Expansion Enclosures (up to 240 disk drives), plus support for external virtualized SAN attached storage
- ▶ Local authentication is only available on V7000 Unified. For more information see Chapter 11, “Implementation” on page 133.
- ▶ Real-time Compression is only available on Storwize V7000 Unified. For more information see Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.





# IBM General Parallel File System

In this chapter we discuss the clustered file system which is built into the Storwize V7000 Unified as one of its foundations: the IBM General Parallel File System (GPFS). The content is focused on GPFS itself and its features, and in terms of the implementation inside the Storwize V7000 Unified there might be some differences to the capabilities which GPFS natively has - check the Storwize V7000 Unified Configuration Limits and Restrictions page accordingly, which can be found at:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003906>

Note that Asynchronous Replication for files, as implemented in Storwize V7000 Unified, is not a generic GPFS function so this feature is not described here. Refer to Chapter 8, “Copy services overview” on page 75.

## 7.1 Overview

GPFS has been available from IBM for a long time dating back to the first releases in the mid 1990's. It has its roots in parallel computing requirements where a scalable, highly available file system was required. GPFS has the parallelism of serving data to (many) clients as well as availability and scalability built-in from design. Therefore it has been used for many years in parallel computing, high performance computing (HPC), Digital Media solutions, and Smart Analytics to name a few.

It also inherited functionality of other projects over time, such as the IBM SAN File System. GPFS is part of many other IBM solution offerings as well, such as IBM Information Archive and Smart Analytics solutions.

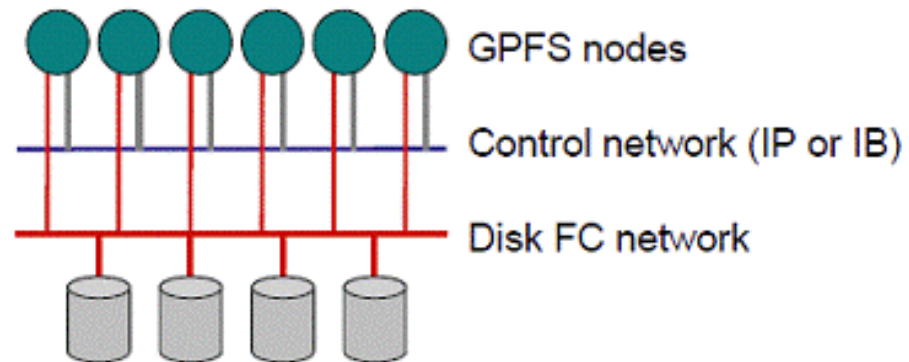
## 7.2 GPFS technical concepts and architecture

The concept of GPFS is a clustered file system built on a grid parallel architecture - parallelism for both host access and data transfers to storage enables its scalability and performance.

The storage entities that GPFS knows are called Network Shared Disk (NSD), as we can see in Figure 7-1. GPFS works with the concept of separate NSD servers and NSD clients in a two-tier architecture, however if only one tier is used both NSD server and client roles are in the same machine. This is the case for the Storwize V7000 Unified implementation and a difference to the SONAS implementation as also shown in Figure 7-1.

## GPFS One Tier Architecture (V7000Unified)

- All Nodes = Client and NSD Server



## GPFS Two Tier Architecture (SONAS)

- Separated Client Nodes and NSD Server Nodes, connected through High Speed Network

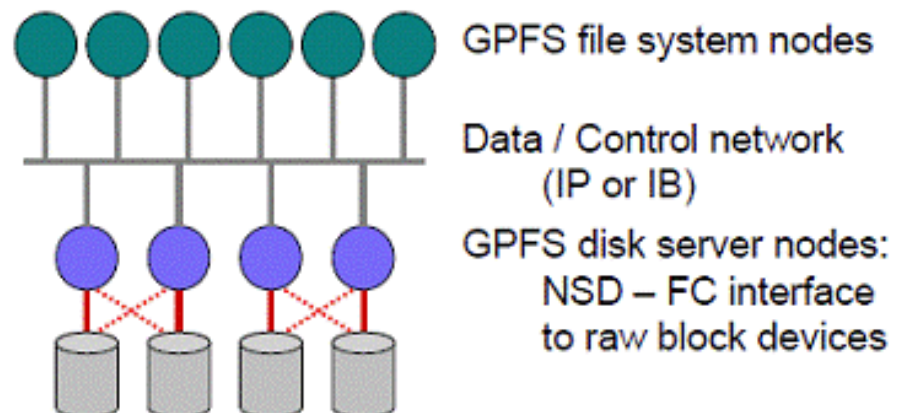


Figure 7-1 GPFS: Examples for One Tier (Storwize V7000 Unified) and Two Tier Architecture (SONAS)

GPFS stripes all the data written across all available NSDs using the defined file system block size as stripe size. This way GPFS ensures that the maximum number of NSDs is contributing to a given I/O, avoiding performance hot spots.

- ▶ Supported file system block sizes in Storwize V7000 Unified: 256 KB, 1 MB, 4 MB
- ▶ The minimum I/O size GPFS is using is called a sub block or fragment (which is 1/32 of the file system block size because GPFS is working with 32 sub blocks per file system block internally)
- ▶ Sub blocks are introduced to combine small files or parts of files into a single block to avoid wasted capacity

For virtualized NSDs presented to the GPFS layer (like NSDs provided by the Storwize V7000 Unified) it is therefore beneficial to optimize the NSD I/O characteristics according to the GPFS I/O pattern.

From a storage subsystem perspective it is desired to get 'full stride writes' to its underlying RAID arrays, which means parity (for RAID 5 and RAID 6) can be calculated on the fly without the need to read from disks before, therefore avoiding the so-called 'RAID penalty'. In conjunction with GPFS this lead to a change of the RAID Presets built into the Storwize V7000 Unified compared to V7000 standalone: The Presets for the RAID arrays in Storwize V7000 Unified will aim to configure 8 data disks plus the required parity disks into one RAID array. For RAID 5 it is an 8+P array, for RAID 6 this is an 8+P+Q array accordingly. These new Presets are reflected in the Sizing tools like Capacity Magic and Disk Magic as well.

### 7.2.1 Split brain situations and GPFS

A GPFS cluster normally requires at least three cluster nodes. An uneven number is selected on purpose by most clustering solutions to still have a quorum of cluster nodes (or other voting members) available if one of them fails — in order to avoid a split brain situation of the cluster. A typical split brain scenario means that there are two parts of the cluster still alive, each one with half of the remaining cluster nodes (so that none of the two parts has a quorum), but they cannot communicate any longer. Due to the loss of communication between the two parts, the parts themselves cannot distinguish which one has the most current information and should continue to operate. One solution for this is to have a 'tie breaker' in the configuration, and the SVC/V7000 clustering implementation uses a quorum disk for that purpose.

In the GPFS cluster implementation in Storwize V7000 Unified it is not possible to have three cluster nodes since there are only two File Modules in the configuration. To help with a split brain situation when the File Modules have lost communication the Storwize V7000 storage system in the back-end acts as the tie breaker. If they lose communication both File Modules, as cluster members, will then communicate with the Storwize V7000 and provide their current status.

The Storwize V7000 then determines which File Module should continue to operate ('survive') as the GPFS cluster and sends an **expelmember** command to the other File Module which then has to leave the cluster. In addition, the V7000 removes the volume mappings of the expelled File Module to guarantee data integrity for the remaining GPFS cluster.

### 7.2.2 GPFS file system pools and Storwize V7000 storage pools

GPFS has the internal concept of pools (in GPFS internal terminology they are called 'storage pools' too) and in order to distinguish them from the pools used in the Storwize V7000 storage layer, we will refer to them as *file system pools* in this book. As described in Chapter 3, "Architecture and functions" on page 21 these GPFS file system pools are



mapped to V7000 storage pools within a Storwize V7000 Unified system. An overview of the different internal structures involved is shown in Figure 7-2 on page 69.

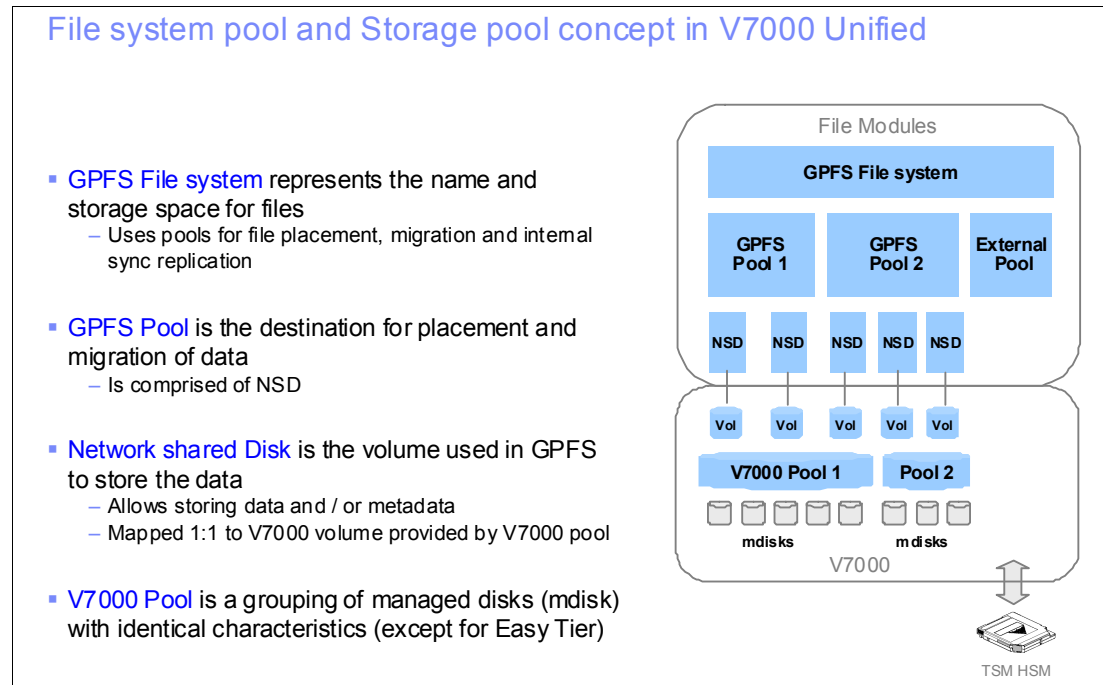


Figure 7-2 File system pool and Storage pool concept in Storwize V7000 Unified

With respect to the mapping between file system pools and storage pools there is one exception: GPFS synchronous internal replication uses a **one to two** mapping (one file system pool to two storage pools) as described in 7.2.6, “GPFS synchronous internal replication” on page 72. Normally there is a **one to one** mapping between file system pools and storage pools, and in either case this mapping is typically established at file system creation.

For a standard file system, not providing ILM functionality, there is one file system pool which is mapped to one storage pool.

For a file system providing ILM functionality, there are multiple file system pools, for example, *system gold*, *silver*, *bronze*, where **each one is mapped to one storage pool**, and where the storage pools have descending tiers, i.e. storage classes descending from fast/expensive to slower/cheaper storage, from the tier mapped to the *system gold* file system pool to the one mapped to the *bronze* file system pool.

**Note:** So as not to confuse these two logical entities within the Storwize V7000 Unified, we will refer to the GPFS pool entity as the *file system pool* and the Storwize V7000 pool is called *storage pool* in this book.

## 7.2.3 File system pools in GPFS

For file system pools in GPFS:

- ▶ There is a maximum of eight internal pools per file system,
- ▶ One pool (default) is always required as the system pool
- ▶ Seven optional pools are available as user pools

- ▶ For configuring ILM, the file system is created with multiple file system pools. That means, beside the one default file system pool called *system* (which exists for every GPFS file system) there are as many *additional file system pools* of descending tiers defined (and mapped to storage pools with corresponding descending tiers) as additional storage tiers are planned to be used through the lifetime of the files in that file system.
  - An example of a typical hierarchy of pools of descending storage tiers/classes is Gold, Silver, Bronze
- ▶ In addition, an external file system pool is possible — which will be used for offloading data as part of an HSM solution (Hierarchical Storage Management) managed by TSM for Space Management as the storage manager application for HSM.

GPFS itself provides a clustered file system layer which is able to support up to 256 file systems in a native GPFS or SONAS implementation, whereas the supported limit in a Storwize V7000 Unified environment is 64 file systems currently.

**Note:** The maximum number of file systems supported in the Storwize V7000 Unified implementation is different - currently up to 64 file systems. For the most current information for not only the file system limit, but for all maxima, refer to:

<http://www-01.ibm.com/support/docview.wss?uid=ssglS1003906>

In addition to the file system pools (with NSDs) and storage pool (with volumes) concept shown in Figure 7-2 on page 69, there are additional configuration layers involved to establish and manage access to the data stored inside the GPFS file system. The essential step is the definition of shares (exports) to be able to access the data from file clients, whereas file sets and directories provide more granularity for the management of that data access (file sets also enable quota management and independent file sets enable snapshots in addition) as shown in Figure 7-3 on page 70.

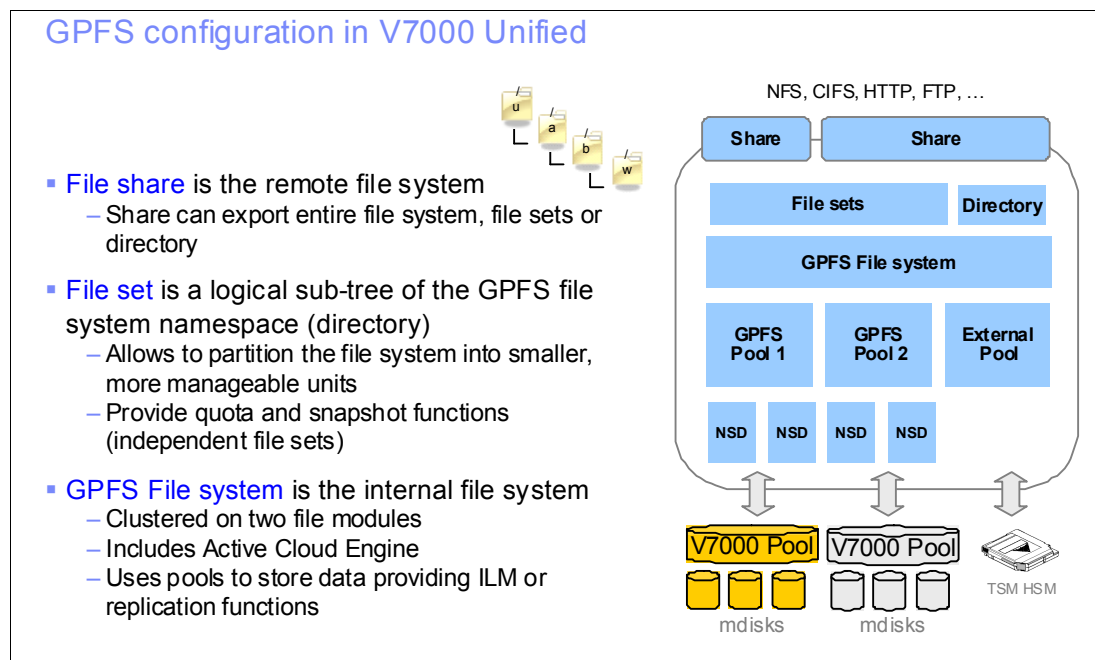


Figure 7-3 Layers involved in managing data access in Storwize V7000 Unified

## 7.2.4 GPFS file sets

A file set is a subtree of a file system namespace that in many respects behaves like a separate file system. File sets provide the ability to partition a file system to allow administrative operations at a finer granularity than the entire file system.

File sets in many aspects behaves like a separate file system *and* available in two types: *dependent file sets* and *independent file sets*.

An independent file set has a separate inode space but shares physical storage with the remainder of the file system.

A dependent file set shares the inode space and snapshot capability of the containing independent file set.

When the file system is first created, only one file set exist which is called the root file set. The root file set contains the root directory and system files such as quota files.

File set details and differences:

- ▶ The default is one 'root' file set per file system
- ▶ Quotas and policies are supported on both *dependent file sets* and *independent file sets*
- ▶ Snapshots are only supported for *independent file sets* (and at the level of the entire *file system* itself) because:
  - Only *independent file sets* provide their own inode space — *dependent file sets* use the inodes of the file system
  - GPFS limit for snapshots is 256 per file system, but 32 are reserved for internal use, for example, for Backup and Async Replication — therefore a maximum of 224 snapshots per file system are available to the user
  - A maximum of 256 snapshots available per independent file set
- ▶ A maximum of 1000 *independent file sets* and 3000 *dependent file sets* is supported per file system

For more information about the differences between dependent and independent file sets refer to:

[http://pic.dhe.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmng\\_filesets\\_topic\\_welcome.html](http://pic.dhe.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmng_filesets_topic_welcome.html)

## 7.2.5 GPFS parallel access and byte-range locking

GPFS uses a distributed cluster manager and various roles distributed across the cluster nodes, both highly scalable and highly available, with transparent adaptive and 'self-healing' capabilities. For that purpose all nodes, or a redundant subset of nodes, have equal roles and run the same daemons, if one node fails other available nodes can then take over its role(s).

GPFS allows parallel access from different client systems to a single file managed by sophisticated locking mechanisms operating GPFS cluster wide on the level of byte-range within a file. There is also support for opportunistic locking (oplocks) which allows client-side caching of data and this can provide a performance benefit.

**Important:** For applications with critical data, all non-mirrored caching options in GPFS, which are:

- ▶ Caching on the client side: controlled via opportunistic locking (*oplocks*) option
- ▶ Caching in file module cache or local interface node cache: controlled via *syncio* option

should be disabled.

This can be done using the CLI command **chexport** for the relevant shares with the parameters '*oplocks=no*' and '*syncio=yes*'.

## 7.2.6 GPFS synchronous internal replication

GPFS provides optional, additional redundancy by means of a synchronous, file system internal replication, based on grouping of NSDs into independent so-called '*failure groups*':

- ▶ A *failure group* is a logical group of NSDs with the same dependencies, for example, failure boundaries
  - Typically two independent storage systems, providing NSDs, to protect the GPFS data against the failure of an entire storage system
  - Within the implementation in Storwize V7000 Unified there is one Storwize V7000 in the back-end which in itself provides redundancy by design and protects against any single failure. Also virtualized external SAN attached storage systems are managed/presented by the Storwize V7000 and can also rely on the failure boundary of an entire Storwize V7000 storage system as well.
- ▶ Setting up the GPFS synchronous internal replication requires the definition of two independent storage pools (as failure groups) per file system pool:
  - Resulting in a two to one mapping between two storage pools and the file system pool
  - Metadata is always stored in the file system pool *system*
- ▶ The replication process is fully synchronous mirroring over two sets of NSDs from the two independent storage pools
- ▶ The configurable options are to replicate *metadata*, *data* or both between two independent storage pools (failure groups)

**Note:** GPFS synchronous internal replication of metadata, data or both is fully supported and configurable on Storwize V7000 Unified. The goal is to perform synchronous mirroring between independent failure groups on two separate storage systems which do not have to provide redundancy. In the Storwize V7000 Unified implementation there is one highly available central storage subsystem with redundancy by design, which is the Storwize V7000 itself, managing all the disk storage in the back-end (Storwize V7000 internal disks and external SAN attached storage systems virtualized by the Storwize V7000).

## 7.2.7 Active Cloud Engine

GPFS provides a very fast and scalable scan engine that is able to scan through all files or subdirectories quickly, and which is used for multiple purposes, for example, identifying files for Antivirus scanning, for ILM and for changed files for incremental backups. Since GPFS is designed for scalability and can grow to a very large file system with a single name space, its scan engine has been designed to be as scalable as well. This is a real competitive advantage of GPFS compared to other large clustered file systems.

In the GPFS implementations in both SONAS and Storwize V7000 Unified this engine is also called the Active Cloud Engine™ (ACE).

The GPFS scan engine is also used to apply user-defined policies to the files stored in GPFS, building the foundation for the Information Lifecycle Management (ILM) of files within GPFS:

- ▶ Enabling ILM, automated migrations of files based on user defined criteria
- ▶ Uses a subset of SQL as policy language
  - User can specify rules grouped in different policies using this policy language
  - Can be scripted
- ▶ Policies and rules for file placement (creation), migration and deletion
- ▶ Rules within the policy are evaluated first to last, and the first one to match will be executed and determines the handling of the relevant file(s)
  - Recommendation is to add a default rule in every case which gets applied when no other rule matches the criteria
  - If no default rule exists and no other rule matches the criteria defined, no action will be taken unless defined otherwise

For more information refer to the *GPFS 3.5: Advanced Administration Guide*, SC23-5182-05:

<http://www-01.ibm.com/support/docview.wss?uid=pub1sc23518205>

## 7.2.8 GPFS and Hierarchical Storage Management

GPFS has support for TSM HSM built-in, which provides a way to offload data from the file system to external storage in conjunction with TSM for Space Management. Internally this is handled by a special file system pool that has been defined called the external pool. Based on the defined criteria, the GPFS policy engine identifies the files to be moved and GPFS will then move these files into this external storage pool from which the TSM HSM server fetches the data and stores it on a TSM HSM supported storage device, usually a tape device.

While the data itself is being offloaded, and is saving space inside the file system, for every file a so-called *stub file* is left inside the file system. It contains all the metadata belonging to that file and which is needed to be read by the scan/policy engine. That means for policy scans the data itself can remain on the external storage outside of the GPFS file system, because all metadata information for this file is still available via the *stub file*. But if a user wants to access the file, or TSM wants to back up the file, or the Antivirus scanner wants to scan the file, it will be recalled and will be loaded back into the file system by HSM. This is all done transparently to the user.

## 7.2.9 GPFS Snapshots

GPFS also offers a space efficient Snapshot technology which is described in Chapter 8, “Copy services overview” on page 75. This is a summary of the main features of GPFS snapshots as implemented in Storwize V7000 Unified:

- ▶ Snapshots use pointers to data blocks based on redirect-on-write, that means a new write coming in and even an update to an existing data block is written to a new data block since the old data block is still contained in the Snapshot.
  - Space efficient as it does not consume space/capacity when invoked
  - Space is only consumed when data changes and new blocks are written based on redirect-on-write
- ▶ Snapshots are available for file systems and independent file sets
- ▶ Allowing a maximum of 256 snapshots of the entire file system plus 256 for independent file sets underneath
- ▶ Reserving 32 of these for internal use, hence 224 are available for client use
- ▶ Snapshot rules allow the scheduling of snapshots if required, and retention rules for Snapshots can be defined

- ▶ The Snapshot manager routine runs once per minute executing Snapshot rules in sequential order of definition

## 7.2.10 GPFS quota management

As mentioned previously there are quotas for file system space management built-in:

- ▶ Quotas can be set at a file set, user, or group level
- ▶ Soft quotas, a grace period, and hard quotas are supported
  - When the soft quota limit is reached, a warning is sent but write access is still possible, until either the grace period expires or the hard quota limit is reached, whichever comes first. Writing of data is then inhibited until space is freed up, for example, by deleting files.
  - Default grace period is 7 days
  - When the hard quota limit is reached while updating a file, writing/closing the currently open file is still possible to protect the data.

More detailed information about GPFS, including different purpose-built configurations, can be found in the following books:

- ▶ *GPFS: A Parallel File System*, SG24-5165
- ▶ *Implementing the IBM General Parallel File System (GPFS) in a Cross Platform Environment*, SG24-7844



# 8

## Copy services overview

In this chapter we provide an overview of the Storwize V7000 Unified storage copy functions provided by the Storwize V7000 storage subsystem and the file level copy functions provided by the file modules. For an in depth discussion about storage copy functions refer to *Implementing the IBM Storwize V7000 V6.3*, SG24-7938.

## 8.1 Storage copy services of the Storwize V7000 Unified

Storwize V7000 Unified provides storage in the form of logical volumes to the internal file modules as well as to external storage clients. In addition it provides the same logical volume based copy services that the standalone Storwize V7000 provides. These are FlashCopy, Metro Mirror, Global Mirror and Global Mirror with Change Volumes.

### 8.1.1 FlashCopy for creating point-in-time copies of volumes

FlashCopy is the point-in-time copy capability of the Storwize V7000. It is used to create instant, complete, and consistent copy from a source volume to a target volume. Often this functionality is called Time-Zero copy, point-In-time copy or Snapshot copy.

#### Creating a copy without Snapshot functionality

Without a function such as FlashCopy, to achieve a consistent copy of data for a specific point-in-time, the I/O of the application that manipulates the data has to be quiesced for the entire time the physical copy process takes place. In such a case, the time that the copy process requires is defined by the amount of data to be copied and the capabilities of the infrastructure to copy the data. Only after the copy process is finished can the application that manipulates the data start to access the volume that was involved. Only then is it ensured that the data on the copy target is self-consistent and identical to the data on the source for a given point in time.

#### Creating copies with FlashCopy

With FlashCopy, this process is different. FlashCopy enables the creation of a copy of a source volume to a target volume in a very short time. Thus, the application has to be prevented from changing the data only for a short period of time. For the FlashCopy function to be executed a FlashCopy mapping needs to be created; two ordinary volumes get “mapped” together for the purpose of creating a point-in-time copy.

After the FlashCopy process has been started on the mapping, the target volume represents the contents of the source volume for the point in time when the FlashCopy was invoked. The target volume does not yet contain all the data of the source volume *physically*. It can be seen as a “virtual” copy, created using bitmaps.

After FlashCopy has started, but before it has finished physically copying the data to the target, the copy can be accessed in read/write mode. From that point on, data that has to be changed on the source volume (by the applications that manipulates the source volume) is written to the target volume beforehand, thus ensuring that the representation of the data on the target volume for that point in time is valid.

It is also possible to copy all the data of the source volume to the target volume through a background copy process. The target volume, although not fully copied yet, represents a clone of the source volume as long as the relationship between the source and the target exists. Once all data has been copied to the target volume, the relationship between the volumes can be removed, and both volumes become normal. The former target volume is now a physical clone of the source volume for the point in time that the FlashCopy was invoked.

To create consistent copies of data that spans multiple volumes, consistency groups can be used. Consistency groups are sets of FlashCopy mappings, which get copied at the same point in time, thus creating a consistent snapshot of the data across all volumes.



FlashCopy is very flexible. It is possible for a volume to be a source volume in one FlashCopy mapping and for the volume to be the target volume in another FlashCopy mapping. This is called Cascaded FlashCopy. Also, one volume can have the role as source volume in multiple FlashCopy mappings with *different* target volumes. This is called Multiple Target FlashCopy. Another feature of FlashCopy is the capability of incrementally “updating” a fully copied target volume with only the changes that have been made to the source volume of the same mapping. This is called Incremental FlashCopy. Another feature of FlashCopy is the possibility to reverse the direction of the mapping, thus making it possible to restore a source volume from a target volume while retaining the original target volume. This is called Reverse FlashCopy. The flexibility of FlashCopy can be enhanced further with the use of thin provisioned volumes, this is called Space Efficient FlashCopy.

## FlashCopy usage cases

FlashCopy has many uses. One obvious use is for backing up a consistent set of data without requiring a long backup window. The application manipulating the data must make sure the data is consistent, and the application must be suspended for a short period of time. Once the copy is started, the backup application can access the target while the applications can resume manipulating the live data. No full volume copy is needed.

One very useful case for FlashCopy is to create a full consistent copy of production data for a given point in time at a remote location. In this case, we combine Metro Mirror/Global Mirror and FlashCopy, and we take a FlashCopy from the Metro Mirror/Global Mirror secondary volumes. We can take a consistent backup of our production data on the second location, or create a clone of the data so it is available if anything should happen to our production data.

FlashCopy can also be used as a “safety net” for operations that make copies of data inconsistent for longer-than-normal periods of time. For example, if Global Mirror was to get out of synchronization, the auxiliary volume is still consistent in itself, but the process of resynchronization renders the auxiliary volume inconsistent as long as it is not finished. To obtain a consistent copy of the data of the auxiliary volume while it is being synchronized, a FlashCopy of this volume can be created.

Another use for FlashCopy is to create clones of data for application development testing or for application integration testing. FlashCopy is also useful when a set of data has to be used for different purposes - for example, a FlashCopy database can be used for data mining.

## FlashCopy presets

The IBM Storwize V7000 storage subsystem provides three FlashCopy presets, named Snapshot, Clone and Backup, to simplify the more common FlashCopy operations, as shown in Table 8-1.

Table 8-1 FlashCopy presets

Preset	Purpose
Snapshot	Creates a point-in-time view of the production data. The snapshot is not intended to be an independent copy, but is used to maintain a view of the production data at the time the snapshot is created. This preset automatically creates a thin-provisioned target volume with 0% of the capacity allocated at the time of creation. The preset uses a FlashCopy mapping with 0% background copy so that only data written to the source or target is copied to the target volume.

Preset	Purpose
Clone	Creates an exact replica of the volume, which can be changed without affecting the original volume. After the copy operation completes, the mapping that was created by the preset is automatically deleted. This preset automatically creates a volume with the same properties as the source volume and creates a FlashCopy mapping with a background copy rate of 50. The FlashCopy mapping is configured to automatically delete itself when the FlashCopy mapping reaches 100% completion
Backup	Creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume. This preset automatically creates a volume with the same properties as the source volume. The preset creates an incremental FlashCopy mapping with a background copy rate of 50.

## 8.1.2 Metro Mirror and Global Mirror for remote copy of volumes

Metro Mirror and Global Mirror are IBM branded terms for the functions Synchronous Remote Copy and Asynchronous Remote Copy, respectively. We use the term “Remote Copy” to refer to both functions where the text applies to each equally. These functions are used to maintain a copy of logical volumes held by one Storwize V7000, Storwize V7000 Unified or SVC in another Storwize V7000, Storwize V7000 Unified or SVC at a remote location. This copy can be either synchronous or asynchronous. A new enhancement as of SVC version 6.3.0 is Global Mirror with Change Volumes to support the use low bandwidth links.

### Metro Mirror

Metro Mirror works by establishing a *synchronous* copy relationship between two volumes of equal size. This can be an intracluster relationship established between 2 nodes within the same I/O group of one cluster, or an intercluster relationship, which means a relationship between two clusters that are separated by distance. Those relationships can be standalone or in a consistency group.

Metro Mirror functionality ensures that updates are committed to both the primary and secondary volumes before sending confirmation of the completion to the server. This ensures that the secondary volume is synchronized with the primary volume in case of a failure. The secondary volume is in a read-only state, and manual intervention is required to change that access to read/write state. The server administrator also has to mount the secondary disk so the application can start to use that volume.

### Global Mirror

Global Mirror copy relationships work in a similar way as Metro Mirror does but by establishing an *asynchronous* copy relationship between two volumes of equal size. This is mostly intended for intercluster relationships over long distances.

With Global Mirror, a confirmation is sent to the server before it has received good completion at the secondary volume. When a write is sent to a primary volume, it is assigned a sequence number. Mirror writes sent to the secondary volume are committed in sequential number order. If a write is issued while another write is outstanding, it may be given the same sequence number.

This functionality operates so as to maintain a consistent image at the secondary at all times. It identifies sets of I/Os that are active concurrently at the primary, assigning an order to those

sets, and applying these sets of I/Os in the assigned order at the secondary. If a further write is received from a host while the secondary write is still active for the same block, even though the primary write might have completed, the new host write on the secondary will be delayed until the previous write has been completed.

### **Global Mirror with Change Volumes**

Global Mirror with Change Volumes is an added piece of functionality for Global Mirror designed to assist in attainment of consistency on lower-quality network links.

Change Volumes leverage the FlashCopy functionality, but cannot be manipulated as FlashCopy volumes, as they are special purpose only. Change Volumes provide the ability to replicate point-in-time images on a cycling period (default 300 seconds.) This means that the change rate will only need to include the condition of the data at the point-in-time the image was taken, instead of all the updates during the period.

With Change Volumes, a FlashCopy mapping exists between the primary volume and the primary Change Volume. The mapping is updated on the cycling period (60 seconds to 1 Day.) The primary Change Volume is then replicated to the secondary Global Mirror volume at the target site, which is then captured in another change volume on the target site. This provides an always consistent image at the target site and protects the data from being inconsistent during resynchronization.

### **Copy Services interoperability between SVC, Storwize V7000 and Storwize V7000 Unified**

With 6.3.0 a new concept is introduced to the Storwize V7000 and Storwize V7000 Unified called *layers*. Layers determine how the Storwize V7000 and Storwize V7000 Unified interacts with the IBM SAN Volume Controller (SVC). Currently there are two layers, **replication** and **storage**. All devices must be at least at the 6.3.0 code level and the Storwize V7000 and Storwize V7000 Unified must be set to be to the replication layer when in a copy relationship with the SVC.

The replication layer is for when we want to use the Storwize V7000 or the Storwize V7000 Unified with one or more SVCs as a remote copy partner. The storage layer is the default mode of operation for the Storwize V7000 and is for when we want to use the Storwize V7000 to present storage to an SVC.

## **8.2 File system level copy services of the Storwize V7000 Unified file modules**

In this section we overview how the SONAS software of the Storwize V7000 Unified file modules implements copy services. It implements two main features, snapshots and asynchronous replication.

### **8.2.1 Snapshots of file systems and file sets**

The Storwize V7000 Unified implements space-efficient Snapshots. Snapshots enable online backups to be maintained, providing near instantaneous access to previous versions of data without requiring complete, separate copies or resorting to offline backups.

In the current version the Storwize V7000 Unified has the capability to offer 256 Snapshots per file system and 256 per fileset. The Snapshots can be scheduled or performed by

authorized users or by the Storwize V7000 Unified administrator. SONAS Snapshot technology makes efficient use of storage by storing only block-level changes between each successive Snapshot. Only the changes made to the original file system consume additional physical storage, thus reduce physical space requirements and maximizing recoverability.

Snapshots also support the integration with Microsoft Volume Shadow copy Services (VSS). The VSS allows you to display an older file or a folder version, through Microsoft Windows Explorer. Snapshots are exported to Windows SMB clients by the Volume Shadow Service (VSS) API. This means that Snapshot data can be accessed and copied back, through the previous versions dialog in the Microsoft Windows Explorer.

## 8.2.2 Asynchronous replication

Another important feature of the Storwize V7000 Unified file module software is asynchronous replication. In this section we overview how asynchronous replication is designed to provide a bandwidth friendly mechanism.

Asynchronous replication is available for replicating incremental changes at the file system level to one other site. Asynchronous replication is done using an IBM enhanced and IBM supported version of open source tool 'rsync'. The enhancements include the ability to have more than one file module in parallel able to work on the rsync transfer of the files.

The asynchronous replication is unidirectional, changes on the target site are not replicated back. Async replication can be scheduled. Depending on the number of files included in the replication, the minimal interval will vary depending on the amount of data and files to be sent.

The asynchronous replication is space efficient, that means that it transfers only the changed blocks of a file, not the entire file again. Resource efficiency and high performance is achieved by using both interface nodes in parallel, to transfer the data.

Asynchronous replication is useful for disaster tolerance and disaster recovery capabilities, in other words, using incremental change replication to a disaster recovery remote site. This is particularly important when the raw amount of data for backup/restore for large amounts of data, is so large that a tape restore at a disaster recovery site might be unfeasible from a time-to-restore standpoint.

The first step in performing an async replication is to execute a central policy engine scan for async replication. The high performance scan engine is used for this scan. As part of the asynchronous replication, an internal snapshot will be made of both the source file system and the target file system. The next step is to make a mathematical hash of the source and target snapshots, and compare them. The final step is to exploit the parallel data transfer capabilities by having both file modules participate in the transfer of the async replication changed blocks to the target remote file systems. The internal snapshot at the source side assures that data being transmitted is in data integrity and consistency, and is at a single point in time. The internal snapshot at the target is there to provide a fallback point in time capability, if for any reason the drain of the changes from source to target fails before it is complete.

The basic steps in of SONAS asynchronous replication are as follows:

- ▶ Take a snapshot of both the local and remote file system(s). This ensures first that we are replicating a frozen and consistent state of the source file system.
- ▶ Collect a file path list with corresponding stat information, by comparing the two with a mathematical hash, in order to identify changed blocks.
- ▶ Distribute the changed file list to a specified list of source interface node(s).

- ▶ Run a scheduled process that performs rsync operations on both file modules, for a given file list, to the destination Storwize V7000 Unified. Rsync is a well-understood open source utility, that will pick-up the changed blocks on the source Storwize V7000 Unified file system, and stream those changes in parallel to the remote, and write them to the target Storwize V7000 Unified file system.
- ▶ The snapshot at the remote Storwize V7000 Unified system insures that a safety fallback point is available if there is a failure in the drain of the new updates.
- ▶ When the drain is complete, then the remote file system is ready for use.
- ▶ Both snapshots are automatically deleted after a successful replication run.

A simple diagram of asynchronous replication is shown in Figure 8-1.

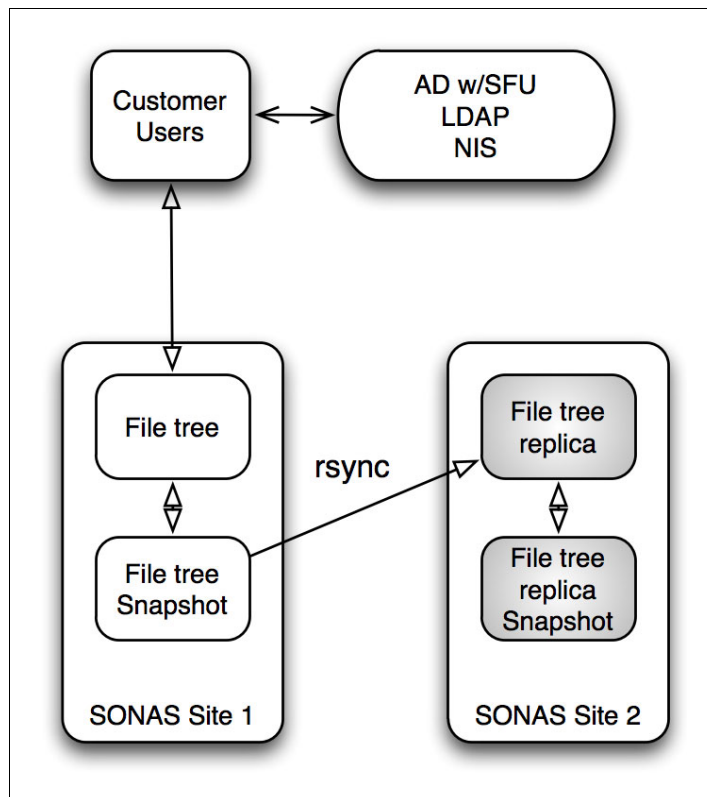


Figure 8-1 Asynchronous replication

### Asynchronous replication limitations

There are limitations that should be kept in mind when using the asynchronous replication function.

- ▶ The asynchronous replication relationship is configured as a one-to-one relationship between the source and target.
- ▶ The entire file system is replicated in asynchronous replication. While you can specify paths on the target system, you cannot specify paths on the source system.
- ▶ The source and target cannot be in the same system.
- ▶ Asynchronous replication processing on a file system can be impacted by the number of migrated files within the file system. Asynchronous replication on a source file system causes migrated files to be recalled and brought back into the source file system during the asynchronous replication processing.

- ▶ File set information on the source system is not copied to the target system. The file tree on the source is replicated to the target, but the fact that it is a file set is not carried forward to the target system's file tree. File sets must be created and linked on the target system before initial replication, because a file set cannot be linked to an existing folder.
- ▶ Quota information is also not carried forward to the target system's file tree. Quotas can be set after initial replication as required, using quota settings from the source system.
- ▶ Active Directory (AD) Only, and AD with NIS using Storwize V7000 Unified internal UID/GID mapping, are not supported by asynchronous replication because the mapping tables in the Storwize V7000 Unified system clustered trivial database (CTDB) are not transferred by asynchronous replication. If asynchronous replication is used, the user ID mapping must be external to the Storwize V7000 Unified system.

### ***Considerations***

For the first occurrence of running asynchronous replication, you might want to consider transporting the data to the remote site physically at first and have replication take care of changes to the data. Asynchronous replication is no faster than a simple copy operation. Ensure that adequate bandwidth is available to finish replications on time.

There is no mechanism for throttling on asynchronous replication. GPFS balances the load between asynchronous replication and other processes.

Source and target root paths passed as parameters must not contain a space, comma, parenthesis, single or double quotation mark characters, "\", "\n", "\r", "\t", "?", "!", "%", or any whitespace characters.

## **8.3 Managing Asynchronous Replication**

Managing asynchronous replication is described in the IBM Storwize V7000 Unified Information Center at:

<http://ibm.biz/BdxFDP>

These topics are discussed:

Configuring asynchronous replication:

<http://ibm.biz/BdxFDy>

Starting and stopping asynchronous replication:

<http://ibm.biz/BdxFDM>

Listing asynchronous replications:

<http://ibm.biz/BdxFDS>

Removing and changing the asynchronous replication configuration:

<http://ibm.biz/BdxFDv>

Asynchronous replication disaster recovery:

<http://ibm.biz/BdxFDM>

Cleaning up asynchronous replication results

<http://ibm.biz/BdxFDK>

Scheduling an established asynchronous replication task

<http://ibm.biz/BdxFDa>







## GUI and CLI

The primary interface for the Storwize V7000 Unified is the Graphical User Interface (GUI) where all configuration and administration functions can be performed. All functions can also be performed using the terminal based Command Line Interface (CLI). A few specialized commands are only available in the CLI which may also be required during recovery should the GUI be unavailable. Both methods are required for management of the cluster.

In this chapter we demonstrate how to set up both methods of access, show how to use them and when each is appropriate.

## 9.1 Graphical User Interface setup

Almost all of the IP addresses in the cluster have a web interface running behind them, but each has a specific purpose.

### 9.1.1 Web server

Each node in the cluster has a web server running. What is presented by each of these web servers depends on the functional status and configuration of the particular node at any given time. Note that all web connections use the HTTPS protocol. If a connection is attempted using HTTP, then it will usually be redirected.

#### Storage node Canisters

Both the storage nodes in the Controller Enclosure can be connected to on their service IP address. This will display the Service Assistant (SA) panel for that node. This is a direct connection to the node software and does not require that the cluster is active or operational, only that the node has booted its OS.

One of the nodes will assume the role of config node when the cluster is active. That node will present the storage management IP address and will present the storage system management GUI. Only one of the nodes will present this and there is only one address.

#### File nodes

For management functions, one of the file nodes is the active management node. This node will present the management IP address and the management GUI for the entire cluster.

Both the file nodes will be connectable with HTTPS over the other IP addresses assigned to their interfaces. What is presented depends on the user configuration of the cluster.

The IP Report shown in Figure 9-1 can be accessed using **Settings** → **Network** → **IP Report** and it will show all the IPs being used to manage the File Module nodes and Control Enclosure nodes.

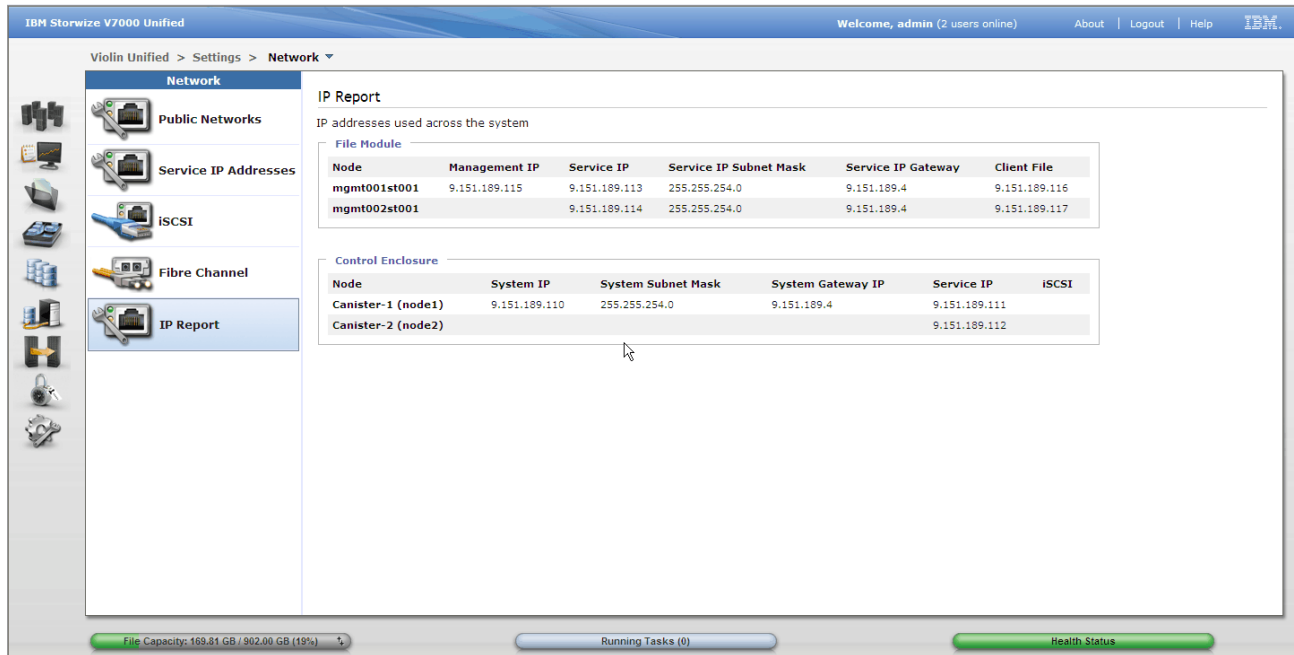


Figure 9-1 IP Report

### 9.1.2 Management GUI

The primary management interface for the Storwize V7000 Unified cluster is the management IP address that is assigned to the file modules. This GUI combines all management functions and can be used for both file and block storage management.

The storage system or control enclosure also has a management interface which is the same as the management GUI found on the standalone Storwize V7000. This can also be connected to at any time, but provides management of the storage function only. Access to resources directly used by the file modules is prohibited, but normal block configuration can be done. It is suggested that you use only the full cluster GUI presented from the file module to avoid confusion, although there may be times during complex recovery when IBM Support will ask you to connect to this interface. There is a warning given when an attempt is made to connect to the control enclosure GUI as seen in Figure 9-2 on page 88 which states to “Use the Storwize V7000 Unified management GUI for normal system management”.

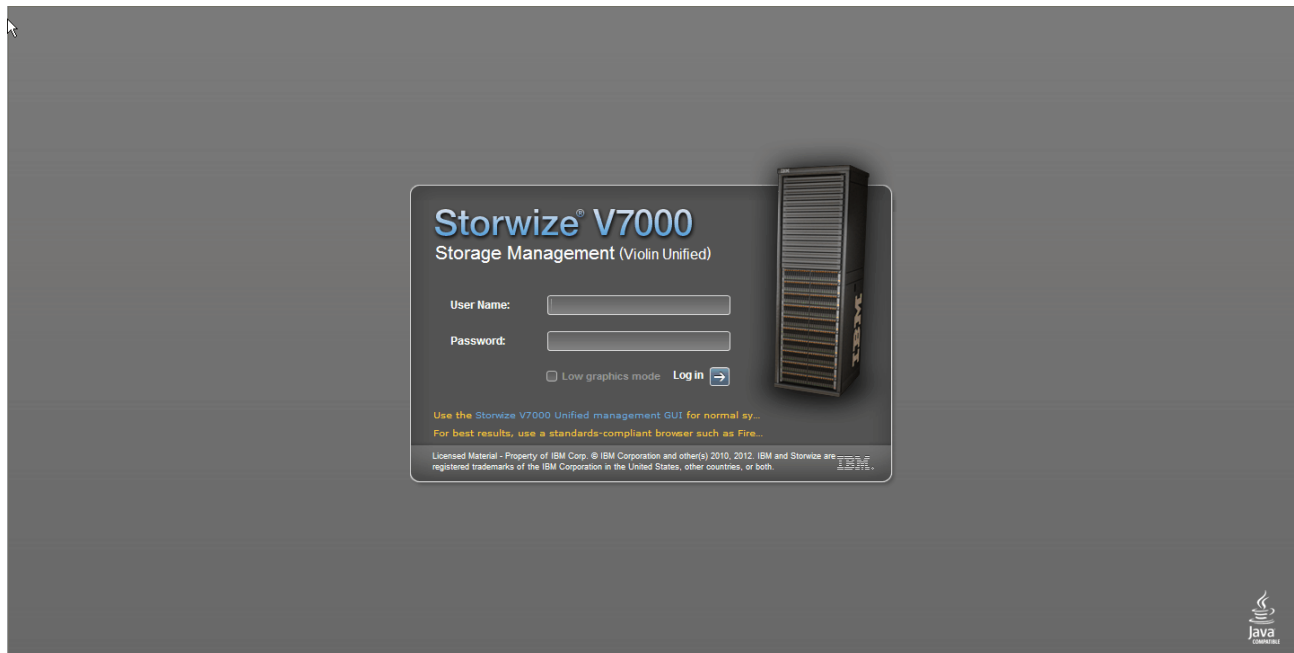


Figure 9-2 Enclosure GUI warning

You will need to access the storage GUI during implementation to set passwords and test its functionality.

### 9.1.3 Web browser and settings

To connect to the GUI, you need a workstation running an approved web browser. Generally any current browser is supported but to see the current list of supported browsers go to the Storwize V7000 Unified support website:

<http://ibm.biz/BdxFXf>

At the time of writing Firefox 3.5 or higher and (IE) 8.x or higher are listed as supported.

To access the management GUI, you must ensure that your web browser is supported and has the appropriate settings enabled. For browser settings refer to:

<http://ibm.biz/BdxFX2>

### 9.1.4 Starting the browser connection

Start the browser application and enter the management IP address assigned to the file modules. If you used `http://<ip_address>` then you will be redirected to `https://<ip_address>`. You will now be warned that there is a security exception and you need to approve the exception to continue. This is normal for this type of https connection.

This will now present you with the logon page, as shown in Figure 9-3.



Figure 9-3 GUI logon

Enter your User Name and Password. If this is a new install, then the default is admin/admin, otherwise you need to use the userid and password assigned to you by your storage administrator.

Note the box on this screen labeled “*Low Graphics Mode*”. This option will disable the animated graphics on the management pages and provide a simplified graphics presentation. This is very useful if connecting remotely as it reduces the traffic and increases response time. Some users simply prefer to disable the animation using this option.

With animation on, simply hover over the icons on the left side and the submenu choices are presented as seen in Figure 9-4 on page 90. Using the mouse select the submenu to launch the desired page.



Figure 9-4 GUI animation

Alternatively, if you have disabled animation, first click on the icon to show that section, then using the pull downs at the top of the page, select the subheading as seen in Figure 9-5 on page 90.

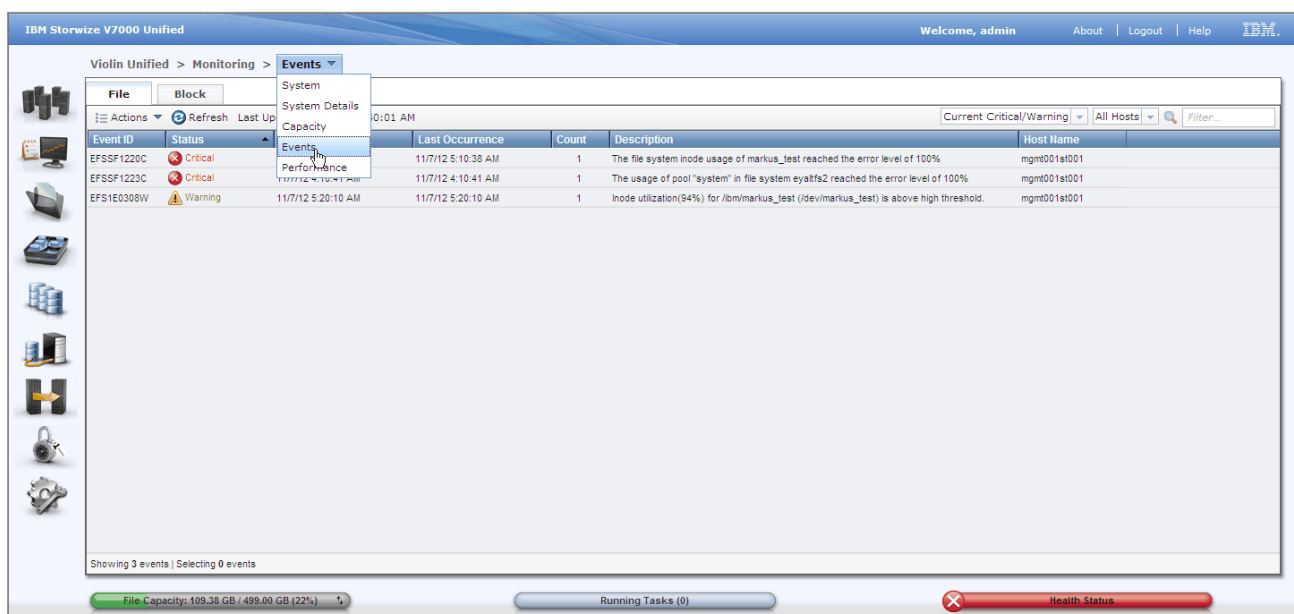


Figure 9-5 GUI no animation

## 9.2 Command Line Interface setup

Using a suitable terminal client such as PuTTY, connect to the management IP address using SSH (port 22). you will then get a login prompt as shown in Figure 9-6.



Figure 9-6 CLI - login prompt

If this is the first time a connection has been made from this workstation, then you may get asked to accept a security key as shown in Figure 9-7 on page 91. Click Yes to tell PuTTY to save the rsa key for future connections.

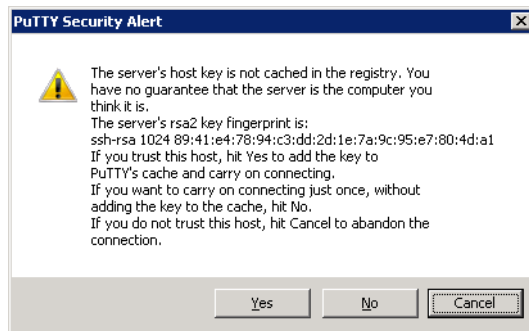


Figure 9-7 PuTTY rsa key

Save the connection definition in PuTTY so it can easily be launched in the future.

You should also connect, test and save a session to the storage management IP address. This will only be used in a recovery situation, but it is a good practice to have it tested and easy to start beforehand. If you are accustomed to earlier code levels of SVC, you will note that the requirement to create and store a key file has been dropped. Authentication is now by userid and password.

## 9.3 Using the GUI

All management and configuration functions for both file and block are available through the Storwize V7000 Unified management GUI interface. For this reason, there is no need to connect to the GUI interface of the storage module for normal operations.

Once logged on to the Management GUI the main window is displayed. This has 5 main areas.

Top action bar	This area is the blue bar across the top. It has a welcome message and includes links to help and information. It also has a logout link to close all functions and log off the GUI.
Main Section Icons	Down the left hand side of the window are a number of icons. Each represents a main section of management. Hovering the mouse over each icon will cause it to become larger and a submenu will appear to the right. If low graphics mode was selected, then you click on the icon to display the topic and chose the submenu using the navigation.
Navigation menu	Along the top of the main window there is a menu showing the currently displayed panel and which section and menu item it belongs to. If the submenu has multiple choices, then it will be shown as a pull-down which allows you to select the submenu.
Main window	The current window is displayed in the right (largest) panel, based on the selections made. The content of this window vary depending on the action being performed.
Bottom status bar	At the bottom are 3 bars. <ul style="list-style-type: none"> <li>- The left bar indicates the current file capacity of the cluster and how much has been used.</li> <li>- The middle bar indicates the number of background tasks running.</li> <li>- On the right side, the bar give information about the health of the cluster. Normally this is colored green, but will change to yellow or red if there are exceptions. Hovering over the "X" at the left hand end will pop up a list of the major components that have unhealthy status and indicate the highest priority status on that component.</li> </ul>

### 9.3.1 Menus

We describe the menus available in the sections that follow.

#### Home

The home menu has only one submenu, the Overview. The Overview window displays a graphical view of the entire cluster from a data point of view. It shows the major areas where data is managed and each icon also gives the number of resources defined in that area. The icons are arranged to show the relationship between each and the data flow.

The suggested tasks button gives a list of shortcuts to common tasks.

Clicking on an icon will display a brief description of that resource at the bottom of the window.

#### Monitoring

The monitoring menu displays the following submenus:

**System:** This gives a graphical view of the cluster showing each major component. The graphic for each component indicates its current status with colored indicators. Hovering over the links or clicking on the item will display popup windows giving the status and configuration details. The identify button will turn on the attention light on each component to assist in locating the physical device.

**System details:** This option gives a tree view of each component. Clicking on the entry in the tree view will display details of that component in the right panel. There is an action pull down to select from the available actions for that component. Each component view is unique and gives detailed information about that component, including its status. Where applicable, the event logs relating to that component will be listed.



**Events:** This menu option is discussed in detail in 15.2, “Event logs” on page 240. There are two tabs which display the two independent event logs, File and Block. The view of each log can be customized and filtered by selecting the desired options in the filter controls at the top of the panel. Both logs provide an action pull down which acts on the highlighted line entry in the view below. It is also possible to right click on a log entry directly to show this action list. The choices in the action list vary between the two logs.

**Capacity:** This menu options gives a window with five tabs. The “*File Systems*” tab shows a list of the file systems, their total capacity and usage details. Clicking on each file system causes it to be included in the graph displayed at the bottom of the window which tracks historic usage over time, or if desired, the percentage. The “*File System Pools*” tab shows a list of the filesystems, their total capacity and specific usage details. It also shows details related to Compression and Thin Provisioning. The “*File Sets*” tab lists the file sets defined and gives usage metrics on each one. The “*Users*” tab give file usage by each user defined to the cluster and “*User Groups*” gives a higher level view based on the groups the users belong to.

**Performance:** The performance option has three tabs and gives a simple view of some key performance indicators. The graphs shown are not meant to provide detailed tuning information, but to show at a quick glance areas of immediate concern that may need further investigation or to quickly identify a problem area during a performance impact on the cluster. The “*File*” tab shows four graphs and the scale can be altered using the pull down menu on the right side. The “*Block*” tab shows four graphs with the scale fixed to 5 minutes. The scope can be changed to show the whole cluster or one node. The “*File Modules*” tab shows graphs and the scale can also be altered using the pull down menu on the right side.

## Files

All configuration actions for the File Services are performed in this menu. These functions are covered in detail in the Chapter 11, “Implementation” on page 133. The files menu presents the following submenus:

**File Systems:** Use this option to view the status and manage the file systems configured on the cluster. Use the “New File System” button to create a file system, or the “Actions” pull down to perform management functions on an existing one. You can determine whether the file system is compressed or not and also see capacity information. You can also filter whether to show NSD or storage pool details for each file system.

**Shares:** This option lists all shares or exports defined on the cluster, the path for their root and the protocol they are able to be accessed with. Use the “New Share” button to create a new share, this will launch a window to enter the details. Or, use the “Actions” pull down to manage an existing share.

**File Sets:** This menu option shows the defined file sets in a list detailing their type, what the path is and in which file system and statistical details. Use the “New File Set” button to define a new file set and the “Actions” pull down to manage an existing one.

**Snapshots:** In this option you can create a snapshot, or mange an existing one from the list displayed using the “New Snapshot” and “Actions” pull down.

**Quotas:** In this option you can create a quota using the “New Quota” pull down, or mange an existing one from the list displayed using the “Actions” pull down.

**Services:** The services tab is used to configure and manage the additional tools provided for the file service.

- ▶ Backup selection gives a choice of which backup technology will be used to backup the file service. At the time of writing, two options are available, Tivoli® Storage Manager and Network Data Management Protocol (NDMP)”
- ▶ The backup option display is technology specific and is used to configure the backup process.
- ▶ The Antivirus selection is used to configure the external antivirus server if Antivirus scanning is being used.

## Pools

The storage pools are a pool of storage from which volumes are provisioned and used as block storage by servers directly and also used by the file server to form file systems. Note that resources owned by the file server will not show in all the GUI views, but the capacity used will be seen. The display for each pool also displays details related to compression. This menu gives several view as follows:

**Volumes by Pool:** Clicking on the pool (or mdiskgroup) in the left panel will display the volumes in that group and their details in the right panel. You can use the “New Volume” tab to create new volumes for blocked storage to servers and also the “Actions” tab to manage these same volumes. You can only monitor NSDs which are the volumes assigned to filesystems such as examining properties.

**Internal Storage:** The Storwize V7000 has internal disk drives in the enclosures. This option displays and manages these drives. Click on the drive class in the left panel to display the drives in that class.

**External Storage:** Storwize V7000 can also manage external storage subsystems if desired using the SAN connection. If any are attached, they are managed in this option. Click on the storage system controller in the left panel to display the volumes presented in the right panel.

**MDisks by Pools:** This option gives a different view of the pools. Here we can see which MDisks are in each pool.

**System Migration:** This wizard is to assist with migrating an external storage system to be managed by the Storwize V7000.

## Volumes

The volumes are built from extents in the storage pools and presented to hosts as external disks. There are several types of block volumes such as thin-provisioned, compressed, uncompressed or generic, and mirrored volumes. In this view we can create, list and manage these volumes.

**Volumes:** This is a listing of all volumes.

**Volumes by Pool:** By selecting the pool in the left panel, we can display the volumes that are built from that pool.

**Volumes by Host:** The hosts that are defined to the cluster are listed in the left panel. By clicking on a host, we can see which volumes are mapped to that host. Note that the File Modules are hosts also and use block volumes (NSDs) but do not appear as hosts.

## Hosts

Each host that will be accessing block volumes on the cluster needs to be defined. In each definition, there also needs to be defined the WWN or iSCSI details of the ports of that host. Once the host is defined, then volumes can be mapped to it, which are then visible to the ports with the WWNs listed.

**Hosts:** This is a list of all defined hosts. Here we can add and manage these definitions.

**Ports by Host:** This view allows us to see the ports defined on each host. The hosts are listed in the left panel, clicking on the host will display the ports in the right panel.

**Host Mappings:** Each mapping showing the host and the volume mapped is listed, one per line.

**Volumes by Host:** In this view we can select the host from the left panel and see volumes that are mapped to it in the right panel.

## Copy Services

Storwize V7000 Unified provide a number of different methods of coping and replicating data. FlashCopy is provided for instant copy of block volumes within the cluster. Remote copy is used to copy block volumes to another location on another cluster and this can be done synchronously (Metro Mirror) or asynchronously (Global Mirror). File systems can be replicated to another file system using the “File Copy Services” submenu described bellow.

**FlashCopy:** In this option, all the volumes in the cluster are listed. Here, we can create and manage copies and view the status of each volume.

**Consistency Groups:** These are used to group multiple copy operations together that have a need to be controlled at the same time. In this way the group can be controlled by starting, stopping, and so on, with a single operation. Additionally, the function will ensure that when stopped for any reason, the IOs to all group members have all stopped at the same “point-in-time” in terms of the host writes to the primary volumes, ensuring time consistency across volumes.

**FlashCopy Mappings:** This option allows us to create and view the relationship (mapping) between the FlashCopy source and target volumes.

**Remote Copy:** In this option we can create remote copies and consistency groups. We can then view and manage these.

**Partnerships:** For a remote copy to be used, there must be a partnership set up between two or more clusters. This option is used to create and manage these partnerships.

**File Copy Services:** Use this panel to select different methods to replicate data to between different file systems.

## Access

There are a number of levels of user access to the cluster, which are managed in this option. The access levels are divided into groups each having a different level of access and authority. If desired, multiple users can be defined and their access assigned to suit the tasks they perform.

**Users:** This option lists the user groups in the left panel, and the users in that group in the right panel. New users can be added to a group and managed.

**Audit Log:** All commands issued on the cluster are logged in this log. Note that even if initiated from the GUI, most actions cause a CLI command to be run, so this will also be logged.

**Local Authentication:** The system supports user authentication and ID mapping using local authentication server for NAS data access. Using local authentication eliminates the need for

a remote authentication service, such as Active Directory or Samba Primary Domain Controller (PDC), thus simplifying authentication configuration and management.

## Settings

Use the Settings panel to configure system options for event notifications, directory services, IP addresses, and preferences related to display options in the management GUI

**Event Notifications:** This option is used to configure the alerting and logging. Here we define the e-mail and SNMP servers and the levels of alerting as desired. This is covered in detail in 15.5, “Call home and alerting” on page 259.

**Directory Services:** Directory services defines the fundamental settings for the file server. We need to define the DNS domain and servers. We define the authentication method that will be used by the file server and the authentication server.

**Network Protocol:** If HTTP is configured as an access protocol on any shares, we need to define the HTTPS security. Note web access is by HTTPS only, pure HTTP protocol is not allowed. Here we set up the authentication method and keys if desired.

**Network:** The network setup for all the interfaces in the cluster is configured here. Use the buttons in the left panel to select the interface and view or modify the values in the right panel.

- ▶ Public Networks defines the file access IP addresses that are presented on the client facing interfaces. These addresses will float across the file module ports as needed.
- ▶ Service IP Addresses are for the storage enclosure only. Define a unique address for port 1 on each node canister. This address is used only for support and recovery.
- ▶ iSCSI defines settings for the cluster to attach iSCSI-attached hosts.
- ▶ Use the Fibre Channel panel to display the Fibre Channel connectivity between nodes, storage systems, and hosts.

**IP Report:** The IP Report panel displays all the IP addresses that are currently configured on the system.

**Support:** This option allows us to define connectivity for sending alerts to IBM Support and allowing IBM Support to connect to the cluster. We can also create, off-load and manage the data collections needed by support.

**General:** In this option we can set the time and date for the cluster, enter licensing details if needed and perform software upgrades for the cluster. The software process is covered in detail in 15.9, “Software” on page 263.

## 9.4 Using the CLI

**Note:** We suggest using the GUI instead of the CLI. The GUI builds the CLI commands needed and automatically includes the correct parameters that are needed. We also suggest using the GUI to determine the best way to use CLI commands. One example is when creating a compressed volume which requires some specific parameters such as **rsize** and **autoexpand** in order to avoid having the volume going offline prematurely because these parameters are missing or mis-configured.

Like the GUI, there is a CLI connection to the Storwize V7000 Unified management address and also to the Storwize V7000 storage enclosure management address. All functions can be performed on the Storwize V7000 Unified, so the only access required for normal operation is

this single CLI session. The CLI session to the storage is only needed in recovery situations, but it is a good practice to have set it up and tested it.

The commands are unique, so storage commands can be issued on the unified CLI using the same syntax. Note that most block commands can be prefixed with `svcinfo` or `svctask` as has been the case on SVC and Storwize V7000 previously. Where there is ambiguity, this prefix needs to be added. This ensures the command is unique and gets the desired result.

For example, **lsnode** displays information about the file modules as shown in Example 9-1

*Example 9-1 lsnode file module information*

```
[7802378.ibm]$ lsnode
Hostname IP Description Role Product version Connection status GPFS status CTDB status
Last updated
mgmt001st001 172.31.8.2 active management node management,interface,storage 1.4.0.0-37b OK active active
11/8/12 2:11 PM
mgmt002st001 172.31.8.3 passive management node management,interface,storage 1.4.0.0-37b OK active active
11/8/12 2:11 PM
EFSSG1000I The command completed successfully.
[7802378.ibm]$
```

and **svcinfo lsnode** displays information about the Storwize V7000 nodes as shown in Example 9-2.

*Example 9-2 lsnode Storwize V7000 node information*

```
[7802378.ibm]$ svcinfo lsnode
id name UPS_serial_number WNNN status IO_group_id IO_group_name config_node UPS_unique_id hardware iscsi_name
iscsi_alias panel_name enclosure_id canister_id enclosure_serial_number
1 node1 50050768020023DC online 0 io_grp0 no 50050768020023DC 100
iqn.1986-03.com.ibm:2145.violinunified.node1 01-1 1 1 78M01FK
3 node2 50050768020023DD online 0 io_grp0 yes 50050768020023DD 100
iqn.1986-03.com.ibm:2145.violinunified.node2 01-2 1 2 78M01FK
[7802378.ibm]$
```

The Information Center has detailed information about the use and syntax of all commands. Most commands are available to all users, but some commands are dependent on the authority level of the userid that is logged on.

Scripting of CLI commands is supported provided the scripting tool supports SSH calls. Refer to the Information Center for details on generating a key and using scripting.

Listed below are some commands that you might find useful during recovery. Always refer to the information center for syntax and expected results.

## 9.4.1 File commands

<b>lscluster</b>	lists the clusters managed
<b>lsnode</b>	List the nodes in the cluster
<b>lsnwmgt</b>	Shows the configuration of the management ports
<b>lsnwinterface</b>	Lists the physical client facing interfaces
<b>lsnw</b>	List the networks (or subnets) defined
<b>chnwmgt</b>	Set or change the addressing of the file module management ports
<b>chrootpwd</b>	Change the root password across all nodes. Need root logon.
<b>initnode</b>	Stop or restart a file node.
<b>resumenode</b>	Resume a node that has been suspended or banned.
<b>stopcluster</b>	Shuts down a cluster or node.
<b>suspendnode</b>	Suspends a node.

<b>lsfs</b>	List the file systems
<b>lsmount</b>	List the mount status of all file systems
<b>mountfs</b>	Used to mount a file system, only used during recovery.
<b>unmountfs</b>	Unmounts a file system.

## 9.4.2 Block Commands

<b>svc_snap</b>	Gathers a data collect from the block storage Storwize V7000.
<b>lssystem</b>	Lists the Storwize V7000 storage system.
<b>svcinfo lsnode</b>	Lists the nodes in the storage system.
<b>lsdumps</b>	Lists dump files saved on the storage system.
<b>lsfabric</b>	Produces a list (often very long) of all the fibre channel paths known to the storage system.
<b>lsmdisk</b>	Lists all the mdisks visible to the storage system. Useful if you need to also see mdisks owned by the file storage, which are hidden in the GUI.
<b>detectmdisk</b>	Re-scans and rebalances fibre channel paths. Use with care as this will reconfigure the pathing to the current visible paths and drop failed paths.
<b>chsystemip</b>	Change or set the IP addresses of the storage system.
<b>stopsystem</b>	Allow you to stop a node or the entire storage system.



# Planning for implementation

In this chapter we describe the planning steps required to select a suitable Storwize V7000 Unified configuration and advise on the information that you need to consider to implement it successfully in your environment.

We will provide checklists for the detailed information needed to set up the Storwize V7000 Unified system, as well as decision steps along the way.

## 10.1 Planning steps sequence

Our intent is to provide a checklist of the different steps that need to be considered in what we have identified as a logical order.

**Note:** There is planning information in the Storwize V7000 Unified Information Center available at:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp)

For the Storwize V7000 Unified library and related publications go to:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmlt\\_relatedinfo\\_224agr.html](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmlt_relatedinfo_224agr.html)

### 10.1.1 Get preliminary planning data

Before implementing a new Storwize V7000 Unified system, it is important to gather all the information about the current environment, and the expectations from the system according to client needs. This is essential in order to plan the system configuration and to get the maximum benefits from it.

It is important to understand the workload of the current system or the new Storwize V7000 Unified system in order to plan the configuration of the storage system:

Collect the following data:

- ▶ Storage capacity required for file and block
- ▶ Planned capacity growth/buffer/estimate over the lifetime of the system
- ▶ Known performance requirements for Block I/O: IOPS - response time, MB/s - Throughput, I/O size, read/write ratio, sequential or random
- ▶ Hosts to be connected for Block I/O: number, platforms, LUNs per host, number and type of ports per host, multipathing requirements
- ▶ Remote Copy requirements for Block I/O: Metro/Global Mirror, capacities required for each, peak write requirements, distances, distance technologies used
- ▶ FlashCopy requirements for Block I/O: Capacity, Flashcopy features used
- ▶ Clients to be connected for File I/O: number, interfaces, protocols used (CIFS, NFS...), I/O patterns
- ▶ Number of users for simultaneous access
- ▶ Known performance requirements for File I/O per protocol (read/write ratio, sequential or random, I/O size)
- ▶ Asynchronous Replication requirements for file storage
- ▶ Snapshot requirements (frequency, retention period, estimated change rate of data)
- ▶ External SAN virtualized storage to be connected
- ▶ Backup requirements
- ▶ Define the local and remote (if needed) SAN fabrics required
- ▶ Define the networks required



If it is an existing environment or there is test environment available, the workloads experienced can be measured and projected. There are free tools available to analyze workloads and gather the necessary information about the system components, I/O patterns and network traffic. For Windows environments *perfmon* could be used, for example AIX® and Linux environments *nmon* is one of the options to use. There are others like *traceroute*, *netstat*, *tcptrace*, *tcpdump*, *iozone*, *iorate*, *netperf*, *nfsstat*, *iostat*, and others which could be used as well.

This might be an iterative step to be revisited again if required, for example as requirements change later on, or as a result of discussions within the other steps in this sequence.

### 10.1.2 Determine the system configuration to order

Storwize V7000 unified storage system is a flexible system that supports different configurations and sizing.

Use the following to decide on the system configuration that matches your needs. If you need help, then ask your IBM representative for assistance.

- ▶ Use Capacity Magic to determine the configuration for the required capacity based on input data, adding buffer for metadata, snapshots etc. - as a start configuration for the Disk Magic modeling
- ▶ Use Disk Magic to verify the performance requirements can be met with the system configuration determined with Capacity Magic - adjust drive types, RAID levels and number of drives required accordingly
- ▶ Plan for the software licences needed, e.g. base V7000 SW license, file module licenses, external virtualization, remote copy, TSM backup, TSM for space management and real-time compression.

These may have to be revisited several times as requirements change.

### 10.1.3 Perform the physical hardware planning

It is important to take into account the physical components that will be needed to ensure that everything that is required is ordered ahead of time. This includes:

- ▶ Storwize V7000 Unified system, cables, connectors
- ▶ SAN and Network switches required, cables, connectors,
- ▶ File access clients required, I/O adapter cards
- ▶ FC/iSCSI hosts required, I/O adapter cards
- ▶ Power and cooling requirements for all hardware involved
- ▶ Plan for the lab floor space and rack layout for all the hardware identified

### 10.1.4 Define the environment and services needed

Plan for the environment and the services that will be required:

- ▶ IP addresses needed for management and service of Storwize V7000 and both File Modules, public IP addresses to serve file I/O, client IP addresses
- ▶ Authentication service: server(s) needed according to the selected method and netgroup/ID mapping support
- ▶ Time synchronization: Network Time Protocol (NTP) server(s)

- ▶ Domain Name System: DNS server(s)
- ▶ Copy Services and Async Replication, including required connectivity and remote target systems
- ▶ Backup server(s) according to the method chosen and storage
- ▶ TSM HSM server(s) and storage, if required
- ▶ Antivirus scan engine(s)

### 10.1.5 Plan for system implementation

The closer you get to the actual implementation the more important it is that you have considered:

- ▶ Define the local and remote (if needed) SAN zoning requirements
- ▶ Define the network requirements for management and data access
- ▶ Define the network interfaces of V7000 and File modules, including subnets and VLANs
- ▶ Define the logical configuration of the system (both File and Block access)
- ▶ Define the pools and LUN layout for Block access
- ▶ Define the pools, exports/shares, file systems, file sets and directory structures for File access
- ▶ Define users required - for management/monitoring roles of the Storwize V7000 Unified itself and for file based access requiring authentication and configure them within the authentication service/directory server
- ▶ Plan for the user ID mapping method (external/mixed)
- ▶ Define authorizations required for every file access user within the file system/file set/directory structures

### 10.1.6 Plan for Data Migration

Based on the features of the Storwize V7000 Unified there are two different options for data migration:

- ▶ Migrating data from existing SAN attached storage to the Storwize V7000 using the built-in migration wizard and image mode volumes
- ▶ Migrating data from existing NAS systems to the Storwize V7000 Unified using file based migration options

## 10.2 Support, limitations, and tools

Always verify your environment against the latest support information for Storwize V7000 Unified and be sure to use the latest versions of modelling tools for capacity and performance.

Determine lists of hosts and platforms to be attached and verify interoperability support and any restrictions for:

- ▶ FC attachments
- ▶ Network attachments/file access
- ▶ iSCSI attachments

Determine your requirements and verify they are within the capabilities and limitations of the system. The Technical Delivery Assessment (TDA) checklist provides additional useful considerations. Use the modelling tools available (with help from your IBM or Business Partner support if needed) to determine the system configuration which is able to fulfill your capacity and performance requirements.

Here are some useful links for these purposes:

- Support portal for Storwize V7000 Unified:

<http://www.ibm.com/storage/support/storwize/v7000/unified>

- Interoperability support pages for Storwize V7000 Unified:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1003911>

- Configuration Limits and Restrictions:

<http://www.ibm.com/support/docview.wss?uid=ssg1S1003906>

- The general 'Limitations' section in the Infocenter is very useful as preparation for the planning and implementation decisions:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fadm\\_limitations.html](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fadm_limitations.html)

- Verify the planned setup and environment using the Pre-Sales Technical Delivery Assessment (TDA) checklist. The checklists for TDA for both Pre-Sales and Pre-Install can be found here (IBM and Business Partner (BP) internal link only, contact your IBM or BP support for help if you do not have access):

<http://w3.ibm.com/support/assure/assur30i.nsf/WebIndex/SA986>

- Determine your capacity requirements including Asynchronous Replication for files, Snapshots, FlashCopy, Remote Copy for block I/O, and GPFS internal replication requirements and verify the system configuration using **Capacity Magic**.
- Determine all the workload parameters required such as the number of I/Os per second (IOPS), throughput in MB/s, I/O transfer sizes for both file I/O and block I/O workloads, number of clients (for file access), number of hosts (for block I/O access, both iSCSI and FC), copy services requirements for both block I/O and file I/O, and verify the system configuration using **Disk Magic** modelling.

Note that the accuracy of the input data determines the quality of the output regarding the system configuration required.

It is also important to note that there are influencing factors outside of the Storwize V7000 Unified system modeled which can lead to a different performance experienced after implementation, like network setup and I/O capabilities of the clients used.

The IBM modelling tools for capacity (Capacity Magic) and performance (Disk Magic) can be found here (IBM internal link only - contact your IBM or Business Partner support for help with the modelling if you do not have access, Business Partner have access to these tools via Partnerworld):

<http://w3.ibm.com/sales/support/ShowDoc.wss?docid=SSPQ048068H83479I86>

## 10.3 Storwize V7000 Unified advanced features and functions

In this section we summarize the main considerations involved in the setup and usage of every individual feature and function listed. As usual, the individual ranking and importance of these features and their corresponding considerations might vary, so this list is neither necessarily complete for all possible environments, nor is the order of the individual considerations by any means representative.

### 10.3.1 Licensing for advanced functions

- ▶ Almost all advanced functions are included in the two base licences required for Storwize V7000 Unified, which are: 5639-VM1 (V7000 Base, one license per V7000 enclosure required) and 5639-VF1 (File Module Base, two licences required)
- ▶ Exception: External Virtualization requires a 5639-EV1 license by enclosure
- ▶ Exception: Remote Copy Services for Block I/O access requires a 5639-RM1 license by enclosure
- ▶ Exception: Real-time compression requires license by enclosure.

### 10.3.2 Planning guidelines for using Real-time compression

It is important to understand the requirements for using compression before enabling it.

- ▶ **Hardware requirements:** Compression required dedicated hardware resources within the node which is assigned/de-assigned when compression is enabled/disabled. when you create the first compressed volume in an I/O group, hardware resources are assigned as there are less cores available for the Fast Path I/O code. Therefore you should not create compressed volume/file system if the CPU utilization is consistently sustained above 25%.
- ▶ **Data type** - best candidate for data compression are data types that are not compressed by nature. do not use compression in volumes/file system that contains data that compressed by nature. Selecting such data to be compressed will provide very little savings or no savings at all while consuming CPU resources by generating additional I/Os. Avoid compressing data with less than 25% compression ratio. Data with at least 45% compression ratio is the best candidate for compression.
- ▶ **Compression ratio estimation** - In order to estimate the compression ratio of a volume, use the Comprestimator tool. this is a command line host-based utility that scan the volume and return the compression ratio that can be achieved while using compression. for more information refer to:

<http://www.redbooks.ibm.com/redpieces/pdfs/redp4859.pdf>

**Comprestimator:** Comprestimator can be used only on devices mapped to hosts as block devices. Therefore it cannot be used in file servers and file systems in the V7000 unified. For more information about estimating compression ratio of files refer to Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.

- ▶ **Mixture of comprmissible and uncomprmissible data in a file system - Placement policy:**

When a file system contains a mixture of compressible and non-compressible data, it is possible to create a file system with two file system pools. One for the compressible files - configured with compression enabled, and the other for the non-compressible files. The **placement policy** option is configures to place the compressible files in the compressed

filesystem pool. The policy is based on a list of file extensions of compressed file type that are defined as an exclude list. This extensions list is edited manually when configuring the policy. Using the placement policy will avoid spending system resources on non-compressible data. The policy is only affecting files that created after the management policy has changed.

For more information about the placement policy and file types, refer to Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.

- ▶ **License** - compression has limited access in the current version. Therefore a code should be entered when configuring a new compressed file system. In order to get the code to enable compression contact IBM at NEWDISK@us.ibm.com and a specialist will contact you to provide the code.
- ▶ **Amount of compressed volumes:** The number of compressed volumes is limited to 200 per I/O group. This number includes compressed file systems. When a file system is created, 3 compressed volumes are created for it, and they are counted in the 200 compressed volumes limitation. For example, if you created 195 compressed volumes and then created one compressed file system, you will actually have 198 compressed volume in use and you will not be able to create another compressed file system (only 2 compressed volumes will be left and it takes 3). Therefore it is important to understand this limitation and plan the number of compressed volumes and file systems in the entire system before using it.
- ▶ **Plan the amount of pools you need in order to use compression:** When using compression the consideration about the MDisk that should be created are different. There are several items that should be considered first:
  - **Compressed and non-ncompressed volumes should not be stored in the same MDisk group.** In mixed volume types, the compressed and non-compressed volumes will share the same cache and it may increase the response time. Therefore, it is not recommended to create compressed volumes in an MDisk group that contains non-compressed volumes.
  - **Create different pools for data and metadata.** In order to use compression, the data and metadata should be separated because the metadata should not be compressed. Therefore, at least two storage pools should be created. for more information about configuration refer to Chapter 11, “Implementation” on page 133.
- ▶ **Balanced system:** When creating the first compressed volume, CPU and memory resources are allocated for compression. When the system contains more than one IO group It is recommended to create a balanced system. If you create a low amount of compressed volumes it is recommended to create them all in one IO group. For larger numbers of compressed volumes, the general recommendation, in systems with more than one I/O group, would be to distribute compressed volumes across I/O groups. For example, a clustered pair of Storwize V7000 control enclosures requires 100 compressed volumes, it is better to configure 50 volumes per I/O group, instead of 100 compressed volumes in one I/O group. You should also ensure the preferred nodes are evenly distributed.
- ▶ **Easy tier:** Real-time compression does not support Easy tier. Easy Tier is a performance function that will automatically migrate or move 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 multi-tiered storage pool over a 24-hour period. Compressed volumes have a unique write pattern to MDisk which would have triggered unnecessary data migrations. For this reason, Easy Tier is **disabled** for compressed volumes and you cannot enable it. You can however create a compressed volume in an Easy Tier storage pool but automatic data placement is not active.

For more information about Real-time compression see Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.

For more information about compression technology refer to *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859

<http://www.redbooks.ibm.com/redpieces/abstracts/redp4859.html>

### 10.3.3 Asynchronous Replication

- ▶ Requires a second Storwize V7000 Unified system
  - Not supported within a single Storwize V7000 Unified system
- ▶ Requires full external User ID mapping, for example, Active Directory with Services for Unix (SFU)
- ▶ Operates at the file system level, with a 1:1 relation between the local and remote file system
- ▶ Sequence of operation:
  - First snapshot is taken on source file system
  - Changes to previous replication are identified and replicated to target file system,
  - Once the replication is complete then a snapshot is taken on the target file system and source snapshot is deleted
- ▶ Timing:
  - Frequency is determined by the interval defined in a scheduled task for Asynchronous Replication. The minimum interval which can be defined is 1 minute. Duration of one run is determined by the time to take a snapshot, scan for changed files, the time it takes to transfer the changes to the remote site given the network bandwidth available, then to take a snapshot at the remote site, and finally the time to delete the source snapshot. In case the first Asynchronous Replication cycle has not been completed before the next scheduled Asynchronous Replication is triggered, the subsequent replication will not start to enable the first one to complete successfully (an error will be logged). After its completion a new Asynchronous Replication will start at the next scheduled replication cycle.

**Note:** Fileset and share definitions, quota and snapshot rules defined are not contained within the replicated data and this information is kept only at the source and is not transferred to the replication target.

These definitions have to be applied to the target file system as needed for a failover scenario (which might be different from the scenario at the source).

For testing Disaster Recovery the target file system could be mounted as read-only to clients on the target side. If writes are allowed and happen to the target file system then there is a potential data integrity issue, since the source file system does not reflect these updates as changes are only tracked at the source.

If write access should be allowed to the target side, for example, as part of a Disaster Recovery test, it is required to create file clones for the affected/targeted data files within the target file systems. This cannot be done using the snapshot on the target side (which can only be accessed as read-only) because it is not possible to create file clones from a snapshot.

### 10.3.4 Snapshots

Snapshots are by design space efficient, working with pointers and redirect on write for updates to existing data blocks which are part of a snapshot. Therefore the rate of changes to the data being part of a snapshot as well as the frequency of creating snapshots and the retention period for the existing snapshots determine the capacity required for snapshots. If a snapshot is just used for backup or asynchronous replication and deleted afterwards, then there is usually no need to include significant extra capacity for snapshots into the planning. But if a number of previous versions of files are kept in snapshots to protect against operational failures (enabling easy file restores for users) this needs to be taken into account along with the expected change rate for this data.

### 10.3.5 Information Lifecycle Management

File systems for ILM require multiple internal file system pools to be defined in different tiers, which are then mapped to different storage pools which should be based on storage tiers for example drive classes, drive technology.

You will need to create a plan for the lifecycle of a file, for example based on file type, time since last modification or time since last access. Based on the file capacities needed for the different tiers the corresponding capacity in storage tiers needs to be provided. This determines the type and number of disk drives to order for the back-end storage.

Determine the desired policy definitions for:

- ▶ Data placement at file creation
- ▶ Data migration between the tiered pools during the lifetime of the file
- ▶ Data deletion after the file's specified expiration criteria are met

### 10.3.6 Hierarchical Storage Management (HSM)

- ▶ HSM works in conjunction with TSM as the backup method
- ▶ HSM is not supported in conjunction with NDMP backup
- ▶ Requires software and license for Tivoli for Space Management
- ▶ Requires an external file system pool to be configured
- ▶ Requires appropriate external storage which is supported by TSM HSM

### 10.3.7 Data Backup and Recovery: TSM or NDMP

- ▶ Only one method is supported, hence a selection of TSM or NDMP is required so you must choose one or the other
- ▶ If NDMP is selected, then HSM is not supported
- ▶ NDMP backup is supported with Netbackup, Commvault Sipana, EMC Networker, and TSM as Data Management Application (DMA)
- ▶ The NDMP data service runs on the Storwize V7000 Unified File Modules
- ▶ Different NDMP topologies are supported: 2-way or 3-way (local is not supported)
  - 2-way: DMA and Tape Service running on same system
  - 3-way: DMA and Tape Service running on different systems, whereby the metadata information is sent to the DMA and the data containers sent to the Tape Service



- ▶ If TSM is selected as the backup method, the preinstalled TSM client on the file modules is used
- ▶ Selection of TSM enables the option to use HSM as well.

### 10.3.8 Antivirus

- ▶ Supported AV product families/access schemes are: Symantec and McAfee (via ICAP on port 1344)
- ▶ Requires external scan engines
- ▶ Configurable options: scan on file open, scan on file close after write, scheduled batch scans (aka bulk scan)

**Note:** Bulk scans do not re-scan HSM-migrated files, no file recall is therefore required.

### 10.3.9 External Virtualization of SAN attached back-end storage:

- ▶ Provides scalability beyond the Storwize V7000 limit for internal storage, which is 360 TB currently
- ▶ Maximum capacity which can be addressed is determined by Storwize V7000 extent size(s) defined at the storage pool layer, with a maximum of 2<sup>22</sup> extents managed
- ▶ External storage is licensed by storage enclosure
- ▶ Same support matrix as Storwize V7000 and the SAN Volume Controller

### 10.3.10 Remote Copy Services (for block I/O access only)

Remote Copy Services (not including Asynchronous file based replication) are the same as available with a Storwize V7000 with the same minimum code release of V6.4. They are not applicable for Storwize V7000 volumes used for file systems.

An important consideration in the Storwize V7000 Unified is the reduced FC fabric connections due to the required direct connections between the File Modules and the canisters.

See more details on this in *Implementing the IBM Storwize V7000 V6.3*, SG24-7938

- ▶ Remote copy partnerships are supported with other SVC, Storwize V7000 or Storwize V7000 Unified systems (with the StorwizeV7000 in the back-end of a Storwize V7000 Unified system)
- ▶ Fibre Channel Protocol support only
- ▶ Licensed by enclosure
- ▶ SAN and SVC/Storwize V7000 Copy Services distance rules apply (maximum 80 ms per round trip)
- ▶ Needs partnerships defined to remote system (SVC/Storwize V7000/Storwize V7000 Unified)
- ▶ A maximum of three partnerships at a time are supported, and that means a maximum of 4 systems can be in one copy configuration. Be aware that not all topologies possible are supported, for example all 4 systems configured in a string A-B-C-D.



- ▶ Within the partnerships defined between systems, the copy services relationships are established at a volume level as a 1:1 relationship between volumes. Each volume can only be in one copy services relationship at a time.
- ▶ Consistency Groups are supported

### 10.3.11 FlashCopy (block volumes only)

The Flashcopy implementation (not including Snapshots as used for file sets and file systems) is the same as available with a standalone Storwize V7000 with the same minimum code release of V6.3. FlashCopy operations are not applicable for Storwize V7000 volumes used for file systems.

See more details on this in *Implementing the IBM Storwize V7000 V6.3*, SG24-7938

- ▶ Need to take volumes and capacity needed for FlashCopies into account
- ▶ All SVC/V7000 Flashcopy options are fully supported on standard Storwize V7000 volumes not used in file systems
- ▶ Consistency Groups are supported

### 10.3.12 General GPFS recommendation

Every file operation requires access to the metadata associated, therefore it is a general GPFS recommendation to place the metadata on the fastest drive type available. This can be achieved by creating NSDs (Storwize V7000 volumes associated with a file system) based on the fastest drive type and add them to the *system* filesystem pool with the specific usage type of *metadataonly*. In addition the usage type of the other, slower NSDs need to be set to *dataonly*. This ensures that only the fastest disks will host the metadata of the filesystem.

### 10.3.13 GPFS internal synchronous replication, also known as ‘NSD failure groups’

As described in Chapter 7, “IBM General Parallel File System” on page 65 this provides an additional copy of the selected data type (data, metadata or both) in a different storage pool. Therefore the pool configuration and additional capacity required needs to be taken into account.

- ▶ Synchronous replication operates within a file system and provides duplication of the selected data type
- ▶ Requires multiple storage pools to be defined per file system pool (for example only *system* pool of a file system which always contains the metadata or other data pools as well as desired)
- ▶ Ideally file system pools using this functionality are replicated between storage pools in independent failure boundaries
  - This independence defines the level of protection, for example, against storage subsystem failure
  - This independence is compromised here since there is only one Storwize V7000 storage system managing the back-end storage
- ▶ If *metadata*, *data* or *both* to be replicated between the file system pools
  - Defines level of protection
- ▶ Capacity used must be included in planning for total file system capacity

- Approximately 5-10% of filesystem capacity is used for metadata so you will need to adjust overall capacity accordingly

### 10.3.14 Manage the write caching options in Storwize V7000 Unified and on client side

There are different options to enable/disable caching within the layers inside the Storwize V7000 Unified and also outside, for example, on the client side.

On **NFS clients** the options specified with the *mount* command determine if client side caching is allowed, hence this can be changed at the level of each individual export and the Storwize V7000 Unified has no control which option each NFS client is using: Mounting an export with the *'sync'* parameter disables the client side caching and assures the data is sent to the Storwize V7000 Unified after each update immediately.

For **CIFS clients** the Storwize V7000 Unified supports opportunistic locking (oplocks) which enables client side caching for the CIFS clients. That means by default the client side caching is granted by the Storwize V7000 Unified to every client which requests opportunistic locking. This can be changed for every individual export/share via the *chexport* command (see information box below).

**Inside the Storwize V7000 Unified** there could be write caching on the file modules managed by the NFS/CIFS server layer. This will happen by default for all open files for CIFS access while for NFS access the default is already set to *syncio=yes*. As soon as there is a sync command or a file gets closed the updates are written to the NSDs (volumes in the V7000 pools) immediately and will be stored in the mirrored write cache of the V7000 before destaged to disk, this is safe as there is a second copy of the data. The caching in the file modules can be controlled for every export/share via the *chexport* command options (see information box below).

**Important:** For applications with critical data, all non-mirrored caching options in GPFS, which are:

- ▶ Caching on the client side: controlled via opportunistic locking (*oplocks*) option
- ▶ Caching in file module cache: controlled via *syncio* option

should be disabled.

This can be done using the CLI command **chexport** for the relevant shares with the parameters *'oplocks=no'* and *'syncio=yes'*.

### 10.3.15 Redundancy

The Storwize V7000 Unified has been designed to be highly available providing redundancy by design. In order to achieve High Availability for the data access and operations it is required that other parts of the environment provide redundancy as well. This is essential for services like Authentication, NTP and DNS. There is a similar requirement for the networks to be redundant and of course for the power sources of all these components as well.

If there is no redundancy at just one of these levels then there is an exposure to not being able to continue operations when there is just a single failure in the environment. Having redundancy at all these levels ensures that at least a double failure is necessary to create an outage to the operations.

## 10.4 Checkpoints and considerations for authentication

Decide which implementation of authentication service and if only external (recommended) or a mixed external/internal user ID mapping will be used.

If there is an existing authentication infrastructure already, many times this will only include one version (for example Active Directory or LDAP), which will basically determine the decision for the implementation of Storwize V7000 Unified as well.

**Important:** The Storwize V7000 Unified supports only *one* authentication method at a time - and changing it later is *not recommended*, and therefore it is important to carefully decide and select the method at the start.

Make sure long term goals are taken into account. Also the potential usage of certain functions, like Asynchronous Replication as well as the planned future enhancements for WAN caching (see Statement of Direction published at Announcement time of Storwize V7000 Unified), requires an *external only* ID mapping. So if there are chances this might be needed at some point in the future make sure an external only user ID mapping is used from the start.

The details of each of the following authentication options are described in Chapter 4, “Access control for file serving clients” on page 37.

Below is a summary of the available options out of which *one* method needs to be selected:

### 10.4.1 Active Directory (includes Kerberos)

These are the options and considerations:

- ▶ Standard (provides ID mapping for Windows only)
  - in this case the Storwize V7000 Unified will use an internal ID mapping as well, both for UNIX type users (using a User ID/Group ID scheme) as well as mapping Windows SIDs to local UID/GIDs
- ▶ With Services for Unix (RFC2307 schema)
  - Available on domain controllers running Windows 2003 SP2R2 and higher
  - This option will provide a full external User ID mapping for both Windows and UNIX users inside the Active Directory server
- ▶ With Services for Unix (sfu schema)
  - Available on domain controllers running Windows 2000 and 2003
  - Similar option for older domain controllers to provide a full external User ID mapping for both Windows and UNIX users inside the Active Directory server
- ▶ With NIS (netgroup support only)
  - Adding netgroup support via NIS
    - Netgroup is an option to group hosts and manage them as one group
- ▶ with NIS (netgroup support and User ID mapping)
  - Using NIS for the UNIX user ID mapping

### Recommendation for an Active Directory environment

Providing a full external user ID mapping in the Active Directory server is the recommended option in an Active Directory environment.

## 10.4.2 LDAP (Lightweight Directory Access Protocol)

These are the LDAP options:

- ▶ LDAP
- ▶ Secure LDAP (with Kerberos)
- ▶ Secure LDAP (with Kerberos) and SSL/TLS encrypted communication (available via CLI only)

### Recommendation for an LDAP environment

Secure LDAP (with Kerberos) and SSL/TLS encrypted communication provides the most security and is therefore the recommended option in an LDAP environment

## 10.4.3 Samba PDC (NT4 mode)

These are the Samba PDC options:

- ▶ Standalone
- ▶ With NIS (netgroup support only)
- ▶ With NIS (netgroup support and User ID mapping)

This is a legacy implementation which not many environments will require, however it is still supported in conjunction with the Storwize V7000 Unified.

- ▶ **NIS** (NFS only)
  - NIS with netgroup support only

This option does not provide a *user* based authentication but rather a *client* (host or IP address) based authentication as all users connecting from the same NFS client machine will get access.

## 10.4.4 Local Authentication

In this release, version 4.1 local authentication has been added. This gives the ability to create an open LDAP server on the node and replicate the configuration between the nodes using the LDAP MirrorMode.

For more information please see Chapter 11, “Implementation” on page 133.

## 10.5 SAN considerations

The SAN considerations of the Storwize V7000 Unified are very similar to the ones of a standalone Storwize V7000, since all the FC access related functions are exactly the same. The only difference is that the Storwize V7000 Unified only has 4 FC ports available on its V7000 node canisters for SAN connectivity since the other 4 ports are dedicated for, and directly connected to, the two File Modules.

Recommendation is to have a redundant SAN configuration with two independent fabrics, providing redundancy for the V7000 connections, FC host port connections and connections for external SAN virtualized storage systems. All connections should be evenly distributed between both fabrics to provide redundancy in case a fabric, host, or storage adapter goes offline.

### 10.5.1 Zoning considerations

For the Fibre Channel connectivity of the Storwize V7000 Unified the same zoning considerations apply as for a standalone V7000, with the one difference that there are only 4 FC ports available (2 ports per node canister: port 3 and port 4).

Recommendation is to create a node zone in every fabric with two of the four V7000 ports (one per node canister, ports 3 in one fabric, ports 4 in the other fabric) as a means of redundant communication path between the two node canisters, should there be a problem with the communication via the midplane inside the V7000.

If there are hosts to be attached via Fibre Channel to the Storwize V7000, create a host zone per host in each fabric and assign the FC connections of this host in a redundant fashion and zone the V7000 node canisters to ensure redundancy.

If there is external SAN attached storage to be virtualized, create a storage zone in each fabric with half of the ports of the external storage system and one port per node canister in the same fashion.

## 10.6 LAN considerations

The Storwize V7000 Unified uses the Storwize V7000 as the back-end storage system for internal storage, therefore all the LAN considerations for the Storwize V7000 product apply to the back-end part as well:

- ▶ In contrast to the ports on the File Modules, the network ports on the Storwize V7000 are *not* bonded by default
- ▶ 1 Gb Ethernet port 1 on both node canisters is used for management access via the Storwize V7000 cluster IP address by default
  - Optional: use port 2 to define a second management IP address for redundancy
  - As with the standalone Storwize V7000, the management IP address is active on port 1 of the current configuration node canister. The role as configuration node canister can be taken by either one of the two node canisters in a V7000 and is changed/switched between the two for example based on problems, changes or during V7000 code updates.

**Note:** The management communication between the Storwize V7000 and the two File Modules runs via these 1GbE ports on the Storwize V7000 and they must be configured to be in the same subnet as the management ports on the File Modules. The File Modules could optionally use 10 GbE ports for management, but the 1 GbE ports are the default and must be used for initial configuration.

Note also that the Storwize V7000 Unified only uses IPv4 at this time.

- ▶ Both 1 GbE ports can simultaneously be used and configured for iSCSI access from iSCSI hosts

- ▶ The 10 GbE ports in V7000 models 312/324 can be configured for iSCSI access only
- ▶ Every node canister in the V7000 should be configured with a Service IP address which is usable in the given network environment, in case the Storwize V7000 cluster management IP address is not reachable or maybe the cluster itself is no longer working. Then access via the Service Assistant interface to the individual node canisters might be required to debug and solve the situation. There are default addresses of 192.168.70.121 and 192.168.70.122 set on the node canisters, with a subnet mask of 255.255.255.0 and a default gateway of 192.168.70.1. These can be used if appropriate or changed to other valid addresses on the customer network.

**Note:** In contrast to the standalone Storwize V7000, the service IP addresses of the Storwize V7000 node canisters can no longer be changed as part of the USB key initialization process of a Storwize V7000 Unified system. The init tool screens for Storwize V7000 Unified only allow us to set the service IP addresses of the two File Modules during initial install. Therefore it is recommended as the first step, when GUI access is available, to change the service IP addresses of the V7000 node canisters to the desired ones.

In addition, there are two File Modules providing their own, different interfaces and therefore generating additional considerations:

- ▶ All interfaces are bonded, i.e. using a virtual interface bonded on two physical ports.
  - Therefore the two ports belonging to one bonded interface can not be attached to separate networks, this is described in the installation documentation as well.
- ▶ The File Modules use two 1 GbE ports (bonded) for a direct cluster connection between them.
- ▶ Similar to the Storwize V7000, the two remaining 1 GbE ports can be used for both management access and data traffic (difference: the V7000 supports iSCSI traffic only) via TCP/IP.
- ▶ The default bonds configured are: *ethX0* for data traffic on the 1 GbE ports, *mgmt0* for management traffic on the same 1 GbE ports and *ethX1* for the 10GbE ports of the File Modules
- ▶ Default management access is via 1 GbE ports and this is required for initial install.
  - Management via 10 GbE ports is optional and could be configured later.
  - Ensure that the communication with the Storwize V7000 management IP via 1 GbE continues to work since management of the Storwize V7000 storage system is always via 1 GbE.
- ▶ VLANs are supported and can be configured for both 1 GbE and 10 GbE after initial install and Easy Setup steps are complete, as there is no VLAN support during initial install and Easy Setup.

**Note:** Currently there is an open problem when having both 1 GbE data network and 10 GbE data network on the same subnet. The Storwize V7000 Unified then responds only on the 1 GbE interfaces. This is to be fixed in one of the next maintenance releases.

## 10.7 Miscellaneous configuration planning

In the sections that follow we detail some of the other items that need to be configured:

### 10.7.1 Set up local users to manage the Storwize V7000 Unified system

A number of predefined roles are available to define users with different accesses and to tailor access levels to your requirements:

- ▶ Security Administrator rights plus User management
- ▶ Administrator - full administration of the system except User management
- ▶ Export Administrator - export/share related administration only
- ▶ System Administrator - system related administration only
- ▶ Storage Administrator - storage related administration only
- ▶ Snapshot Administrator - snapshot related administration only
- ▶ Backup Administrator - backup and replication related administration only
- ▶ Operator - only has read access to the system
  - The default user of a Storwize V7000 Unified system is *admin*, which has the Security Admin role and can manage other users
    - It is recommended to create other Security Admin users as required, optionally you can increase security by changing the default access for example by changing the password for the user *admin*.

### 10.7.2 Define call home and/or event notifications

Call home requires an SMTP or email server address on the customer LAN that can forward emails to the default IBM service address. Details about the system, customer/administrator contact and phone numbers are needed to establish contact from IBM support personnel in case of problems.

Event notification is supported via

- ▶ Email - requires SMTP or email server address to be specified, multiple levels of notifications can be specified (for example problems, informational events)
- ▶ SNMP - define IP address of server and which kinds of events, for example status changes, utilization should trigger a notification
- ▶ Syslog Server - define IP address of server to receive information, currently only information about the V7000 will be sent.

### 10.7.3 Storage pool layout

In general there are useful default settings, referred to as 'Presets', built into the system which will be used for the automated configuration steps as offered by Easy Setup. If there are standard performance requirements for either file I/O or block I/O workload then these can conveniently be used, creating shared pool(s) containing volumes for both workloads. If there is a significant workload on either file I/O or block I/O side then it is recommended to separate these by using separate pools. The separate pools enable fault isolation, performance predictability by using different physical disk drives in the back-end and easier performance analysis.

An additional criteria is the file protocol used to access the data: Data accessed by CIFS clients **must not** reside on SAN attached, external virtualized storage. Beside the reason of having separate failure boundaries on storage pool level this is another reason to manage separate storage pools for external storage and assign them only to file systems which do not have CIFS shares defined.



Here is a checklist for the storage pool layout:

- ▶ Block, File or mixed storage/workload required
  - Block workload only: no dependencies to file workloads, use GUI or CLI to configure
    - No special performance requirements then use the Presets/Best Practices built into Storwize V7000 by checking 'Auto-configure storage' in Easy Setup
    - Special consideration regarding performance and/or placement optimization: the CLI allows for specially tailored configurations
- ▶ File workload only: operating outside of the general positioning of Storwize V7000 Unified, but there might be good reasons for that.
  - No special performance requirements: use 'Auto-configure storage' in Easy Setup and GUI to configure, this includes the PreSets/Best Practices built-in
  - Special consideration regarding performance and/or placement optimization: the CLI allows for specially tailored configurations
- ▶ If mixed Block/File workload: plan storage layout between the two, including a manual configuration of Mdisks/Pools/Volumes as needed, configure storage layout, first for the file systems (generating file volumes) first using GUI (or CLI), then block volumes. General recommendation: although supported, do not use mixed pools - use separate pools for file access and block I/O for better performance control and bottleneck analysis if required.

#### 10.7.4 File access protocols required for client access

All data subsets only need to be accessed via one file protocol:

- ▶ NFS exports: the default owner of an export is the *root* user. Need *root* user access to the NFS client for initial access to the export and create the directory structures and access permissions for other users as desired.
- ▶ CIFS shares: it is mandatory to **specify an owner** when creating a new share to be able to access it for the first time from the client side. Otherwise the default owner will be the *root* user as with NFS but this user typically does not exist on CIFS client side. This initial owner specified is the one used for initial access from CIFS client side to create the directory structures and ACLs for all users as desired, important is for example the 'TraverseFolder' right to be able to access directories below the original home directory of the share. Once the directory structure for other users and appropriate ACLs are defined, necessary shares can be defined in the Storwize V7000 Unified afterwards. If their access works as needed, the initial owner can be deleted if desired or his access right could be minimized as needed.

#### 10.7.5 Multiple simultaneous exports of same subset of data via different protocols

This is fully supported by the Storwize V7000 Unified. Most likely this multiprotocol export will be using NFS and CIFS: the difficulty is to ensure access rights and ACLs set are compatible from both client sides.



**Note:**

**First time creation of a CIFS share:** It is mandatory to specify an owner for first time access when creating a CIFS share, otherwise only the *root* user has access to that share if no other permissions are set from the client side already (which is not the case if it is really the first time access). An owner of a share can only be set or changed on the Storwize V7000 Unified when there is no data stored in the share yet.

**For managing CIFS ACLs:** Authorization setting to 'Bypass traversal check' is not supported by Storwize V7000 Unified, therefore 'Traverse folder' rights need to be explicitly granted to all users (or for example to 'Everyone') which should have access to directory structures below the level of the current directory. Please note that these users do not see any contents when traversing directories just having 'traverse folder' rights.

**Simultaneous export of the same data via both CIFS and NFS:** Changing ACLs from NFS side will most likely destroy the CIFS ACLs since NFS uses the much simpler POSIX bits for user/group/other to manage access rights. Since CIFS provides much more sophisticated ACL management options, it is recommended to manage ACLs for the common share on the CIFS client side.

**Note:** Volumes for file access (equivalent to NSDs as seen by GPFS) are not explicitly visible in the GUI and can not be modified in the standard GUI screens for example volumes or pools. Only volumes for block I/O can be created and modified in these GUI screens. This is made on purpose since the 'File volumes' get created on file system creation and are always associated with a file system. Therefore they should not be manipulated separately and are hidden from the standard GUI panels involving volumes, they are managed on the file system layer instead (typical CLI commands: *chfs* and *mkdisk*).

The 'File volumes' can be displayed explicitly by listing the details on the respective file system in the GUI or via CLI.

## 10.8 Physical hardware planning

Based on the results of the configuration sizing steps, determine the list of hardware items to order. Table 10-1 tries to pre-empt the main questions to ensure all areas have been thought of. However, due to the complexity and multitude of options this might not be complete in every case and specific items might need to be added as required.

Table 10-1 Checklist for required hardware

Hardware area	Components/Details	Your items/numbers
Storwize V7000 Unified configuration	Base configuration V7000 expansions Connectivity for the different interfaces/protocols for all locations/sites involved	
Network components and connectivity	Ethernet switches 1GbE/10GbE and connectivity, for all locations/sites involved	

Hardware area	Components/Details	Your items/numbers
SAN connectivity	FC switches/directors, ports and connectivity for all locations/sites involved	
Clients for File access	Server HW, 1 GbE NIC or 10 GbE CNA, Connectivity for all locations/sites involved	
Hosts for FC or iSCSI access	Server HW, FC HBAs, 1Gb/10Gb Network Cards, Connectivity for all locations/sites involved	
Services	Server(s) for NTP, DNS, Authentication, Backup, HSM, Antivirus - all including connectivity	
Miscellaneous	SAN attached storage & connectivity Remote V7000/Storwize V7000 Unified systems for Remote Copy or Async Replication	

Make sure that the required power and cooling is verified and provided as well.

### 10.8.1 Plan for space and layout

- ☐ An appropriate 19-inch rack with 6U - 24U of space is required - depending on the number of expansion enclosures to be installed. Each V7000 enclosure and each File Module measures 2U in height. The minimum configuration is one V7000 control enclosure and two File Modules with a total height of 6 U.
- ☐ Redundant power outlets in the rack are required to connect the two power cords per V7000 enclosure and per File Module to independent power sources. The number of power outlets required ranges from 6 to 24 per Storwize V7000 Unified system depending on the number of V7000 expansion enclosures.
- ☐ Regarding the physical hardware placement, layout and connectivity there is a lot of detailed information in Chapter 11, "Implementation" on page 133.
- ☐ Two SAS cables of the appropriate length are required per V7000 expansion enclosure. The individual lengths required are determined by the rack layout and placement chosen for the V7000 control and expansion enclosures.

**Note:** There are two independent SAS chains to connect the V7000 control enclosure to the expansion enclosures. A symmetrical, balanced way to distribute the expansion enclosures on both SAS chains is recommended for performance and availability. Please note that the internal disk drives of the control enclosure belong to SAS chain 2 and therefore a maximum of 4 expansion enclosures can be connected to this chain. On SAS chain 1 a maximum of 5 expansion enclosures can be connected. To ensure a symmetrical, balanced distribution the first expansion enclosure would be connected to SAS chain 1, the second one to SAS chain 2, the third one to SAS chain 1 and so on.

The Storwize V7000 Unified Infocenter provides an overview about aspects of the physical implementation planning here:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fsvc\\_installplan\\_22qgvs.html](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fsvc_installplan_22qgvs.html)

The Storwize V7000 Unified hardware installation is described in the *Storwize V7000 Unified Quick Installation Guide*, GA32-1056, available from:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmlt\\_relatedinfo\\_224agr.html](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fmlt_relatedinfo_224agr.html)

## 10.8.2 Planning for Storwize V7000 Unified environment

Here is a list of the minimum prerequisites to set up and use a Storwize V7000 Unified system. Several services are essential for operating and accessing the system and should therefore be provided in a highly available fashion, like NTP, Authentication and DNS.

**Note:** All these services are required but do not necessarily require a dedicated server, for example in case of Active Directory for authentication the server could provide the NTP and DNS service as well.

- ▶ *Time server(s)* for synchronization according to the Network Time Protocol (NTP): to guarantee common clock/time across the environment, especially between Authentication server and the Storwize V7000 Unified system and TSM backups. Provide two servers for redundancy.
- ▶ *Domain Name System server(s)*
  - Required for DNS round robin for file access, provide two servers for redundancy
  - Required for Active Directory authentication (if this is used)
- ▶ *Authentication server(s)* - depending on choice/decision made in Step 10.4, “Checkpoints and considerations for authentication” on page 111. Provide two servers for redundancy.
  - Select one of these: Active Directory, LDAP, Samba PDC, local authentication, or NIS
  - Optional:
    - NIS server for Active Directory or Samba PDC
    - Kerberos/KDC server for LDAP
- ▶ *Connectivity*
  - 4 x 1 GbE ports for File Modules, min. 2 x 1 GbE ports for V7000
  - optional: 10 GbE ports for File service (4x), or iSCSI attachment (minimum 2x)
- ▶ *IP addresses for management and service access:*
- ▶ *Minimum of six IP addresses required for system management and service:*
  - 1 x Storwize V7000 Unified cluster management IP,
  - 1 x V7000 cluster management IP,
  - 2 x Service IP addresses for the two File Modules (one each),
  - 2 x Service IP addresses for the two V7000 node canisters (one each).
- ▶ *Optional: 10 GbE for management of the Storwize V7000 Unified Cluster and File Modules.*
- ▶ *Storwize V7000 requires 1 GbE for management.*

- Note: Initial setup of Storwize V7000 Unified always requires 1GbE for management and a dedicated port - VLANs are not currently supported for initial setup using Easy Setup
- ▶ VLANs are not supported for the initial setup, but can be configured later
  - VLAN ID =1 must not be used

**Important:** The management and service IP addresses of a Storwize V7000 Unified must all reside on the same subnet.

All management and service IP addresses must be active and network configured correctly at initial install. In case of connectivity problems between the File Modules and the Storwize V7000 the initial install will fail.

- ▶ Optional: IP addresses for iSCSI if required:
  - 1-4 IP addresses for 1 Gb iSCSI and/or 1-4 IP addresses for 10 Gb iSCSI
  - recommended: use minimum of 2 addresses per required interface 1Gb or 10 Gb
- ▶ IP addresses for serving File I/O to clients:
  - *Minimum of two public IP addresses* to be used to have both File Modules active in serving I/O
    - *For each Interface used* (1 GbE, 10 GbE) for file I/O.
    - 1 GbE uses *ethX0* bond, 10 GbE uses *ethX1* bond
    - All network ports on the File Modules are bonded by default -> both ports need to be connected to the same subnet(s) as documented

Note the difference that network ports on Storwize V7000 node canisters are not bonded
  - Note: Minimum is one public IP address per interface used, but then only one File module would serve I/O, second File module would remain passive
- ▶ Optional prerequisites - only needed if these features and functions are going to be used:
  - Backup server(s) - TSM or NDMP, supported storage, licensed by TSM Server Value Units
  - TSM HSM server(s), supported storage, licensed by TSM for Space Management
  - Antivirus scan engine(s)

## 10.9 System implementation planning

If the Storwize V7000 Unified is to be added to an existing environment with all the prerequisites listed in section 10.8.2, “Planning for Storwize V7000 Unified environment” on page 119 then there is only the planning for the physical location, power and connectivity required, since all external services such as time synchronization via NTP servers, Domain Name System (DNS), and Authentication using the existing method set up in the environment are already available. Maybe an add-on like Services for Unix to an existing Active Directory infrastructure is required.

In the same sense it is necessary to start building the infrastructure (physical location, power, cooling, connectivity) and the required external services (like NTP, DNS, Authentication) first before implementing the Storwize V7000 Unified. As shown in Chapter 11, “Implementation” on page 133 and Table 10-2 on page 121 the relevant, correct, and often detailed information

has to be entered during the Easy Setup Wizard. Steps such as specifying the NTP server(s) are mandatory and the Storwize V7000 Unified checks if there is an existing connection during Easy Setup. If there is no response from, for example, NTP or DNS servers then the Easy Setup will fail.

If no authentication method is defined during Easy Setup there will be no data access from the file client side, as the required protocol daemons/services only start within the Storwize V7000 Unified after authentication is configured.

### 10.9.1 Configuration details and settings required for setup

There are a lot of different options involved in implementing the Storwize V7000 Unified and therefore it is difficult to provide a complete list of all the details required for all options and combinations. There are also two classes of settings, *optional* ones and *mandatory* ones. This section tries to summarize all the mandatory information needed and most of the optional ones to have it available as a comprehensive overview when implementing the system. However there might be some optional areas not covered here in detail.

You can find a good description of the steps covered during the Easy Setup Wizard and its related configuration information fields in Chapter 11, "Implementation" on page 133.

The following tables detailing the system setup information required start chronologically with the information required for the init tool to prepare the USB key followed by the information requested during the Easy Setup Wizard as shown in Table 10-2.

Table 10-2 Information for Storwize V7000 Unified setup

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Init Tool:</b> V7000 IP, Gateway, Subnet mask	<ul style="list-style-type: none"> <li>- V7000 management IP address</li> <li>- Subnet mask</li> <li>- Gateway IP address</li> </ul>	<p><b>Mandatory</b></p> <ul style="list-style-type: none"> <li>- IP for the V7000 storage cluster (not accessed directly in normal operations)</li> <li>- Gateway and Subnet mask for entire Unified system</li> </ul> <p><b>Please note:</b> Service IP addresses for the V7000 node canisters can not be set here, need access to GUI or CLI. Recommendation: set them first after completing Easy Setup and get GUI access the first time!</p>	<ul style="list-style-type: none"> <li>-</li> <li>-</li> <li>-</li> </ul> <p>V7000 node service IPs:</p> <ul style="list-style-type: none"> <li>-</li> <li>-</li> </ul>

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Init Tool:</b> Storwize V7000 Unified IP and File Module details	<ul style="list-style-type: none"> <li>- Storwize V7000 Unified cluster management IP address</li> <li>- File Module 1 Service IP address</li> <li>- File Module 2 Service IP address</li> <li>- Internal network IP address range</li> </ul>	<b>Mandatory</b> <ul style="list-style-type: none"> <li>- IP for the Unified cluster -&gt; needed and used for all management operations of the Storwize V7000 Unified</li> <li>- individual Service IP for direct access to a File module for troubleshooting</li> <li>- internal network for direct communication and troubleshooting</li> </ul>	<ul style="list-style-type: none"> <li>-</li> <li>-</li> <li>-</li> <li>-</li> </ul>
<b>Easy Setup Step2:</b> System attributes	<ul style="list-style-type: none"> <li>- System name</li> <li>- NetBIOS name</li> <li>- Time zone</li> <li>- NTP server IP address</li> <li>- Alternate NTP server IP address</li> </ul>	<b>Mandatory</b> <ul style="list-style-type: none"> <li>- Name of the Storwize V7000 Unified cluster</li> <li>- Name by which this cluster will be seen on the network (SMB protocol) and known to an AD domain</li> <li>- Continent and City: different scheme, not sorted by GMT+/-Xh, see link  <a href="http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fappx_timezones.html">http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fappx_timezones.html</a> </li> <li>- NTP server is required!</li> </ul> <b>Optional</b> <ul style="list-style-type: none"> <li>- it is recommended to have a backup/alternate NTP server</li> </ul>	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Easy Setup Step3:</b> Licenses	<ul style="list-style-type: none"> <li>- External Virtualization (by enclosures)</li> <li>- Remote Copy (by enclosures)</li> <li>-Real-time compression by enclosures</li> </ul>	<b>Mandatory</b> <ul style="list-style-type: none"> <li>- number of storage enclosures of virtualized SAN storage behind V7000</li> <li>- number of storage enclosures used for Block I/O based Remote Copy functions of V7000 (via Fibre Channel SAN)</li> <li>-Number of storage enclosures using Real-time compression.</li> </ul>	
<b>Easy Setup Step4:</b> Support Notifications	<p>Step1</p> <ul style="list-style-type: none"> <li>- Email server IP address</li> <li>- Company name</li> <li>- Customer email</li> <li>- Customer telephone number</li> <li>- Off-shift telephone number</li> <li>- IBM support email address</li> </ul> <p>Step2</p> <ul style="list-style-type: none"> <li>- enable a proxy server to access the Internet</li> </ul>	<p><b>Optional</b>, but strongly recommended</p> <ul style="list-style-type: none"> <li>- Storwize V7000 Unified cluster will use this Email server to send email</li> <li>- Company name to appear in the email sent</li> <li>- Customer contact to receive email from Storwize V7000 Unified</li> <li>- prime shift telephone number which will be called by IBM support</li> <li>- off-shift telephone number if prime shift number is not answered 24h</li> <li>- leave at default - it is the default address in IBM for 'call home' alerts</li> </ul> <p>Step2</p> <ul style="list-style-type: none"> <li>- if access via proxy server is required, click 'enable' and provide the proxy detail information</li> </ul>	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Easy Setup Step5:</b> Domain Name System	<ul style="list-style-type: none"> <li>- DNS domain name</li> <li>- DNS server(s)</li> <li>- DNS search domains</li> </ul>	<p><b>Mandatory</b></p> <ul style="list-style-type: none"> <li>- name of public network domain associated with Storwize V7000 Unified operations</li> <li>- IP address(es) of your DNS server(s). One is required, more are recommended for availability/redundancy</li> </ul> <p><b>Optional</b></p> <ul style="list-style-type: none"> <li>- additional domain names to be searched in</li> </ul>	
<b>Easy Setup Step6:</b> Authentication	<ul style="list-style-type: none"> <li>- Active Directory (AD)</li> <li>- LDAP</li> <li>- Samba PDC</li> <li>- NIS (NFS only)</li> <li>- Extended NIS</li> <li>- Local authentication</li> </ul>	<p><b>Mandatory</b> (if not specified here, use GUI/CLI to configure authentication later)</p> <ul style="list-style-type: none"> <li>- Radio buttons, choice between AD, LDAP, Samba PDC and NIS</li> </ul> <p><b>Optional</b></p> <ul style="list-style-type: none"> <li>- Extended NIS can be chosen in conjunction with AD or Samba PDC</li> </ul>	



Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Authentication</b> - Details for <b>Active Directory</b>	<b>(only required if choice is Active Directory)</b> - Server - User ID - Password - Enable Services for UNIX (SFU) - Domain Name, Ranges, Schema Mode  if Extended NIS in addition: - primary NIS domain - Server map - Enable UserID mapping - Domain map - User map - User ID range - Group ID range	<b>(only required if choice is Active Directory)</b> - IP address of Active Directory Server - Administrative User ID - password for Administrative User ID - Check-box, select if support for UNIX is required - (only if SFU selected): name of domain SFU belongs to, Lower to Upper limit of the range for User and Group IDs, SFU schema mode used (SFU or RFC2307)  if Extended NIS in addition: - name of the primary NIS domain - NIS server to NIS domain map - check 'Enable' if NIS UserID mapping to be used -> this enables the next four topics: - mapping of the AD domain to the NIS domain(s) - define how to deal with user exceptions (DENY, AUTO, or DEFAULT) - specify User ID range to be used with AUTO option - specified Group ID range to be used with AUTO option	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Authentication</b> - Details for <b>LDAP</b>	<b>(only required if choice is LDAP)</b> - Specify one or more LDAP servers - Search base for users and groups - Bind distinguished name (DN) - Bind password - User suffix - Group suffix - Workgroup - Security Method - Enable Kerberos - Server name - Realm	<b>(only required if choice is LDAP)</b> - IP address(es) of LDAP server(s) - Search base as defined in LDAP server - DN as defined in the LDAP server(s) - password for this DN - User suffix as defined by the LDAP server - Group suffix as defined by the LDAP server - Domain name - If SSL or TLS is used, a screen to specify certificate will appear. If the setting is 'off', the option for Kerberos appears (GUI). Usage of SSL/TLS <b>and</b> Kerberos can be configured using the CLI - check box, check to enable Kerberos - (only if Kerberos enabled): name of Kerberos server - (only if Kerberos enabled): Kerberos realm	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Authentication</b> - Details for <b>Samba PDC</b>	<b>(only required if choice is Samba PDC)</b> - Server host - Administrative User ID - Administrative password - Domain name - NetBIOS name  if Extended NIS in addition: - Primary NIS domain - Server map - Enable UserID mapping - Domain map - User map - User ID range - Group ID range	<b>(only required if choice is Samba PDC)</b> - IP address of the NT4 PDC server - User ID with admin authority to access the NT4 PDC server - password for this User ID - NT4 domain name - NT4 NetBIOS name  if Extended NIS in addition: - name of the primary NIS domain - NIS server to NIS domain map - check 'Enable' if NIS UserID mapping to be used -> this enables the next 4 entry topics - mapping of the NT4 domain to the NIS domain(s) - define how to deal with user exceptions (DENY, AUTO, or DEFAULT) - specify User ID range to be used with AUTO option - specified Group ID range to be used with AUTO option	
<b>Authentication</b> - Details for <b>NIS (NFS only)</b>	<b>(only required if choice is NIS, (NFS only), a.k.a. Basic NIS)</b> - Primary NIS domain - Server Map	<b>(only required if choice is NIS, (NFS only), a.k.a. Basic NIS)</b> - name of primary NIS domain - NIS server to NIS domain map	
<b>Authentication</b> - Details for Local authentication	- User/group name - Password	<b>Optional:</b> Group Id. Otherwise it will be set automatically	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Easy Setup Step8:</b> Configure Storage	- automatically configure internal storage now	<b>Optional here in Easy Setup, but mandatory</b> to be configured to have V7000 provide storage capacity for the Storwize V7000 Unified system - click yes, if internal storage should be configured as specified in Configuration Summary - if not, use GUI/CLI later to configure the internal storage provided by the V7000. If external, SAN attached storage is used, use its appropriate GUI/CLI interfaces to configure	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
<b>Easy Setup Step9:</b> Public Networks	<ul style="list-style-type: none"> <li>- New network</li> <li>- Subnet</li> <li>- VLAN ID</li> <li>- Default gateway</li> <li>- IP address pool</li> <li>- additional gateways</li> <li>- Interface</li> </ul>	<p><b>Optional here in Easy Setup, but mandatory</b> to be configured to enable access to data for file clients (if not configure here, use GUI/CLI to configure later to enable File I/O)</p> <ul style="list-style-type: none"> <li>- Select new network to get to next screens</li> <li>- Subnet with network mask in CIDR syntax (i.e. number of bits reserved for network mask), see Table 3 in <a href="http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fsvc_hardware_planning.html">http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Fsvc_hardware_planning.html</a></li> <li>- VLANs can not be configured in Easy Setup - later step: enter VLAN number if VLANs are to be used (Note: VLAN 1 is not supported))</li> <li>- IP address of default gateway for this subnet</li> <li>- pool of public IP addresses used to serve File I/O using DNS round-robin, minimum is one, but need at least two to have both File Modules serving I/O</li> <li>- optional, add if there are additional gateways</li> <li>- select ethX1 for 10GbE, ethX0 for 1 GbE interface bond</li> </ul>	

### 10.9.2 Configuration options for file access only

In Table 10-3 we show the file access specific configuration options.

Table 10-3 File access specific configuration options

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
Client systems	<ul style="list-style-type: none"> <li>- 10 GbE attached</li> <li>- 1 GbE attached</li> </ul>	List systems, IP addresses, users	IP addresses, users

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
Users	<ul style="list-style-type: none"> <li>- Local users in Storwize V7000 Unified for management &amp; monitoring</li> <li>- Users by client systems for data access</li> </ul>	<ul style="list-style-type: none"> <li>- create local users and select their role(s)</li> <li>- create users on client systems, specify access rights, create same users within authentication server(s) if applicable</li> </ul>	<p>Local Users Storwize V7000 Unified:</p> <p>Users for data access by client:</p>
File System(s)	<ul style="list-style-type: none"> <li>- Size(s)/Capacity needed</li> </ul>	<ul style="list-style-type: none"> <li>- include ILM and Policies if required</li> </ul>	
ILM	<ul style="list-style-type: none"> <li>- define tiered storage pools</li> <li>- define tiered file system pools</li> <li>- define ILM policies</li> </ul>		
File Sets	<ul style="list-style-type: none"> <li>- Independent</li> <li>- Dependent</li> </ul>	<p>Add more granularity</p> <ul style="list-style-type: none"> <li>- to define Snapshot rules and Quota</li> <li>- to define Quota</li> </ul>	
Exports/Shares	<p>by Protocol:</p> <ul style="list-style-type: none"> <li>- CIFS</li> <li>- NFS</li> <li>- https</li> <li>- ftp</li> <li>- scp</li> </ul> <p>Mixed exports if required, for example</p> <ul style="list-style-type: none"> <li>- CIFS &amp; NFS,...</li> </ul>	<ul style="list-style-type: none"> <li>- define owner at initial CIFS share creation</li> <li>- define extended ACLs for CIFS if required</li> <li>- define authorization/access rights from the client side</li> </ul>	
Snapshots	<ul style="list-style-type: none"> <li>- creation rules</li> <li>- retention rules</li> </ul>	<ul style="list-style-type: none"> <li>- by independent file set</li> <li>- by file system</li> </ul>	
Quota	<ul style="list-style-type: none"> <li>- by user</li> <li>- by group</li> <li>- by file set(s)</li> </ul>	<p>For each entity required</p> <ul style="list-style-type: none"> <li>- define soft limit and hard limit</li> </ul>	
Backup	<ul style="list-style-type: none"> <li>- method</li> <li>- server IP addresses</li> </ul>	<ul style="list-style-type: none"> <li>- choose TSM or NDMP</li> <li>- specify IP addresses of backup server(s)</li> </ul>	
HSM	<ul style="list-style-type: none"> <li>- define external file system pool</li> <li>- define TSM HSM settings</li> </ul>		

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
Async Replication	<ul style="list-style-type: none"> <li>- prepare remote partner system(s)</li> <li>- define file system to be replicated, create target file system on remote system, define replication settings</li> <li>- define schedule/task</li> </ul>		
GPFS internal replication	<ul style="list-style-type: none"> <li>- define multiple storage pools per file system pool(s), min. for <i>system</i> pool, other data pools if required</li> <li>- define if Metadata, Data or both to be replicated</li> </ul>		
Real-time compression	Define separate pools for the data and metadata.		

### 10.9.3 Configuration options for Block I/O access only

In Table 10-4 we show the block I/O specific configuration options.

*Table 10-4 Block I/O specific configuration options*

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
Hosts	<ul style="list-style-type: none"> <li>- host name(s) for FC or iSCSI</li> <li>- WWPNs</li> </ul>	<ul style="list-style-type: none"> <li>- create host objects with associated FC WWPNs or iSCSI IQN</li> <li>- create SAN zoning</li> </ul>	
Storage configuration by host	<ul style="list-style-type: none"> <li>- Capacity</li> <li>- Storage pool layout</li> <li>- Volumes</li> <li>- Thin provisioning</li> <li>- Easy Tier</li> </ul>	<ul style="list-style-type: none"> <li>- Pool and volume layout based on overall planning and modelling results</li> <li>- define Thin Provisioning parameters</li> <li>- define Easy Tier start configurations (hybrid storage pools)</li> </ul>	
Copy Services partnerships	<ul style="list-style-type: none"> <li>- Metro Mirror</li> <li>- Global Mirror</li> <li>- Volumes and Volume pairs</li> <li>- Consistency Groups</li> </ul>	<ul style="list-style-type: none"> <li>- to V7000, SVC or Storwize V7000 Unified systems (but V7000 to V7000 fibre channel attached thereof)</li> </ul>	

Step/Purpose	Entry field/information	Comment/explanation	Your Data/Selection
Flashcopy requirements	<ul style="list-style-type: none"><li>- volumes</li><li>- type and options used</li><li>- Consistency Groups</li></ul>		

## 10.10 Planning for Data Migration

If data migration is needed, there are different ways available to achieve it. In general, if you need IBM assistance with the migration of data into the Storwize V7000 Unified contact IBM regarding the Data Migration Services offerings available which matches your requirements:

<http://www.ibm.com/services/us/en/it-services/data-migration-services.html>

For SAN attached block storage, there is a Migration Wizard built into the GUI which helps in migrating existing data into the Storwize V7000 managed storage. This wizard is described in more detail in *Implementing the IBM Storwize V7000 V6.3*, SG24-7938.

For data migration from existing file storage/network attached storage to the Storwize V7000 Unified the migration has to happen on a file level in order to keep all files aware and the software up to date about the changes, and to maintain the ACLs of files and directories. Therefore the block level migration options built into the V7000 cannot be used for that purpose. We recommend contacting your IBM representative to decide on the best migration policy in this case.





# Implementation

This chapter describes the steps to implement the Storwize V7000 Unified from hardware setup to providing host storage. It is not expected that one resource will perform all the steps as several different skill sets are likely to be required during the process.

## 11.1 Process overview

The installation, implementation and configuration tasks are grouped into major steps in the following processes. Each step is performed sequentially and in most cases must be completed before the next step can begin.

A task checklist is included to assist in the implementation. It serves as a checklist to ensure all steps are completed, as a quick reference for experienced implementers, and can also be useful in planning a timeline and to identify the resources and skills needed.

**Important:** While the intent is to provide a complete end to end checklist and procedures for implementation, the latest product manuals should always be referenced. Where possible, manual references have been included.

## 11.2 Task checklist

The major steps for installing and configuring the Storwize V7000 Unified are listed below to give quick overview and aid in planning. These steps are covered in detail in the following sections. Table 11-1 shows an implementation checklist.

Table 11-1 Implementation checklist

Task	Steps	Complete
<b>Hardware Rack and stack</b>		
Preparation	Complete the planning checklist including IP addresses, protocols and server names and addresses.	
Packing slips	Check all items received against the packing lists.	
Environmentals	Confirm cooling and power, room access and safety.	
Rack Storage Control Enclosure	Rack mount the Storage Enclosure	
Rack Expansion Enclosures	If any expansion enclosures are shipped, rack mount these now using the recommended layout.	
Rack File Modules	Rack the two file modules	
Cabling	Power Control enclosures Expansion enclosures - SAS cables File modules Ethernet	
<b>Power On</b>	Power on and check in this order:	
Network	Switches, routers and devices	
Power on Storage Enclosures	Storage expansions Storage control enclosure	

Task	Steps	Complete
Power on File Modules	Both	
<b>Software</b>	If software is pre-loaded - skip	
Prepare for reload if required		
Re-install software if required		
<b>Initialize</b>		
Configure USB key	Set storage service IPs Run init tool and enter settings	
Initialize the Storage	Insert key into storage enclosure Confirm success	
Initialize the File Modules	Insert key into one file module Confirm success	
<b>Base Configuration</b>		
Configure and Connect to GUI	Setup browser access and PuTTY	
EZ-Setup	Log on to run EZ-Setup Complete as much of the configuration as possible with information provided.	
Backups	Set up scheduled backup cron	
<b>Health Check</b>		
Run Health checks	Confirm system is healthy	
<b>Security</b>		
Change Passwords	Change admin password on file modules and superuser password on block storage	
Create Users	Create additional user logons as desired	
<b>Storage controller</b>	Additional configuration is using block storage	
SAN requirements	Connect to SAN and zone	
Configure Storage	Configure and discover any external storage being used. Discover MDisks and build pools	
<b>Block Storage</b>		
Volumes	Configure volumes as required	
Hosts	Define hosts and host ports to cluster	
Mapping	Map volumes to hosts	
Copy Services	Configure copy services - Flash copy - inter cluster relationships - Remote copy global/metro	

Task	Steps	Complete
<b>File Storage</b>	Configure the file storage	
File Systems	Create file systems from the pools	
Files Sets	Define File sets if desired	
Shares	Create shares Add authorised user to shares	

## 11.3 Hardware unpack, rack and cable

For the following sections refer to the Information Center for details and latest updates at this link:

[http://pic.dhe.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd\\_installhdw\\_2343rc.html](http://pic.dhe.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd_installhdw_2343rc.html)

### 11.3.1 Preparation

Ensure you have reviewed and completed the topics discussed in the planning chapter. This includes the physical environment, allocation of IP addresses and names, identification and preparation of network resources (e.g. DNS, NFS), access to online Information Center.

Also ensure that all personnel are familiar with safety information.

### 11.3.2 Review packing slips and check components

Locate the packing slip in the shipping boxes and review it. Confirm that all the ordered components and features have been received.

Minimum items are:

#### **Control enclosure**

- ▶ Control enclosure (models 2076-112, 2076-124, 2076-312, or 2076-324). The last two digits of the model number identify the number of drive slots, either 12 or 24.
- ▶ Expansion enclosure (models 2076-212 or 2076-224) if ordered.
- ▶ Rack-mounting hardware kit, including per enclosure:
  - Two rails (right and left assembly)
  - Two M5 x 15 Hex Phillips screws per rail (two rails)
  - Two M5 x 15 Hex Phillips screws per chassis
  - Note: Two parts of the rail kit are attached to each side of the enclosure.
- ▶ Two power cords.
- ▶ Drive assemblies or blank carriers (installed in the enclosure).
  - Verify the number of drives and the size of the drives.

#### ***Other items shipped with control enclosure:***

- ▶ Documents

- Read first flyer
- Quality hotline flyer
- Environmental flyers
- Safety notices
- Limited Warranty information
- License information
- License Function authorization document
- IBM Storwize V7000 Quick Installation Guide
- IBM Storwize V7000 Troubleshooting, Recovery, and Maintenance Guide
- ▶ Environmental notices CD
- ▶ Software CD that contains the publication PDFs, and the information center content.
- ▶ One USB key, also known as a flash drive, is located with the publications.

***Additional components for control enclosures:***

- ▶ Fibre Channel cables, if ordered
- ▶ Small form-factor pluggable (SFP) transceivers that are pre installed in the enclosure
- ▶ Longwave SFP transceivers, if ordered

***Additional components for expansion enclosures:***

- ▶ Two SAS cables for each expansion enclosure

**Two File modules**

Each file module box contains:

- ▶ File module (server)
- ▶ Rack-mounting hardware kit, including:
  - Two sets of two rails (right and left assembly)
  - Large cable tie
  - Cable ties
  - Two sets of four M6 screws per rail (two rails)
  - Two sets of two 10-32 screws per chassis
  - Cable management support arm
  - Cable management arm mounting bracket
  - Cable management arm stop bracket
  - Cable management arm assembly
  - Note: The rail kits for the servers differ from the control enclosure.
- ▶ Two power cords

***Additional components for file modules:***

- ▶ Documents
  - Read first flyer
  - Quality hotline flyer
  - Environmental flyers
  - Safety notices
  - Limited warranty information
  - IBM Storwize V7000 Quick Installation Guide
  - IBM Storwize V7000 Troubleshooting, Recovery, and Maintenance Guide
  - License information
  - License Function authorization document
- ▶ Environmental notices CD
- ▶ Software CD that contains the publication PDFs, and the information center content

- ▶ Small form-factor plugable (SFP) transceivers that are pre installed in the enclosure
- ▶ Two USB keys, one for each file module.

### 11.3.3 Confirm environmentals and planning

If not already done, you should review the planning chapter in this book. Ensure you understand locations for each component, that the environment has sufficient capacity in terms of power and cooling, and that rack space is available. Also confirm that the planning worksheet has been completed.

Two people are needed to perform the racking of the modules. It is also recommended that two people are used to perform the cabling as this makes the task much easier.

When racking, do not block any air vents; usually 15 cm (6 inches) of space provides proper airflow.

Do not leave open spaces above or below an installed module in the rack cabinet. To help prevent damage to module components, always install a blank filler panel to cover the open space and to help ensure proper air circulation.

For the following hardware installation tasks, refer to the online information centre for latest details or further clarification.

[http://pic.dhe.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd\\_qiinstref\\_b46700.html](http://pic.dhe.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd_qiinstref_b46700.html)

### 11.3.4 Rack controller enclosures

1. Install the rails in the rack for the controller enclosure. Allow 2U for this module.
2. At this time you may find it easier to install the rails for the expansion enclosures and also the file modules if they are in the same rack, as there is more room if this is done before any modules are installed.
3. Remove the enclosure end caps by squeezing the middle of the cap and pulling.
4. From the front of the rack, with 2 people, lift the enclosure into position and align the rails. Carefully slide it into the rack until it is fully seated.
5. Insert the last 2 screws (one each side) to secure the enclosure to the rack, then replace the end caps.

### 11.3.5 Rack expansion enclosures

Repeat these steps for each expansion enclosure.

1. Install the rails in the rack for the expansion enclosure if not already done. Allow 2U.
2. Remove the enclosure end caps by squeezing the middle of the cap and pulling.
3. From the front of the rack, with 2 people, lift the enclosure into position and align the rails. Carefully slide it into the rack until it is fully seated.
4. Insert the last 2 screws (one each side) to secure the enclosure to the rack, then replace the end caps.

### 11.3.6 Rack file modules

1. Install the rails in the rack for the expansion enclosure if not already done. Allow 2U.
2. Extend the rails fully out the front of the rack.
3. From the front of the rack, with 2 people, lift the enclosure into position with the front slightly higher and align the pins at the rear. Then lower the front and align the front pins. Carefully slide it into the rack until it is fully seated and the latches are secure.
4. If required, insert the 2 screws (one each side) to secure the enclosure to the rack.
5. Install the cable management arm at the rear using the instructions included with the arm. This can be fitted on either side.
6. Repeat the above steps for the second file module.

### 11.3.7 Cabling

There are a number of cables that are needed to interconnect the modules of the Storwize V7000 Unified and to provide connectivity to the network and servers. Most are required before proceeding with the install. Connect the cables as follows.

**Tip:** Install the heaviest cables first (power) and the lightest cables last (fiber) to minimize the risk of damage.

#### **Power**

Each module has two power cords. These should be connected to diverse power supplies and electrically separated as much as possible. Preferably to different power strips in the rack which in turn are powered from different distribution boards.

#### ***Control Enclosures***

1. Ensure the power switches on both power supplies are turned off.
2. On one power supply, prepare the cable retention bracket, release and extend the clip and hold to one side.
3. Attach the power cable and be sure that it is pushed all the way in.
4. Push the retention clip onto the cable, then slide the clip down to fit snugly behind the plug.
5. Tighten the fastener around the plug.
6. Route the cable neatly to the power source and connect. Dress the cable away from the rear of the enclosure and secure any excess so it does not interfere with data cabling.
7. Repeat the above steps for the other power supply and install the second power cable.

#### ***Expansion Enclosures (if present)***

The power supplies are the same as the Control Enclosures. Connect the power cables, two per enclosure using the same procedure, for each Expansion Enclosure.

#### ***File Modules***

1. Attach the first power cable to the File Module and be sure that it is pushed all the way in.
2. Route the cable through the cable management arm, allowing plenty of slack so that cable does not become tight when the arm and module are extended.
3. Connect to the power source and connect. Dress the cable and secure any excess so it does not interfere with data cabling.
4. Repeat for the second power cable and for the two power cables in the second module.

## Ethernet

There are 6 ethernet ports on each File Module and 4 on each control enclosure with an optional 4 more if iSCSI is specified. The 1Gb ports require a copper cable being a minimum of CAT5 UTP. The 10Gb ports are connected using fibre cables and need Multimode (MM) fibre cables with LC connectors. These are connected as follows:

### File Modules

Use Figure 11-1 on page 140 as a reference for plug locations. Route each cable through the cable management arm. Connect the following cables for EACH file module.

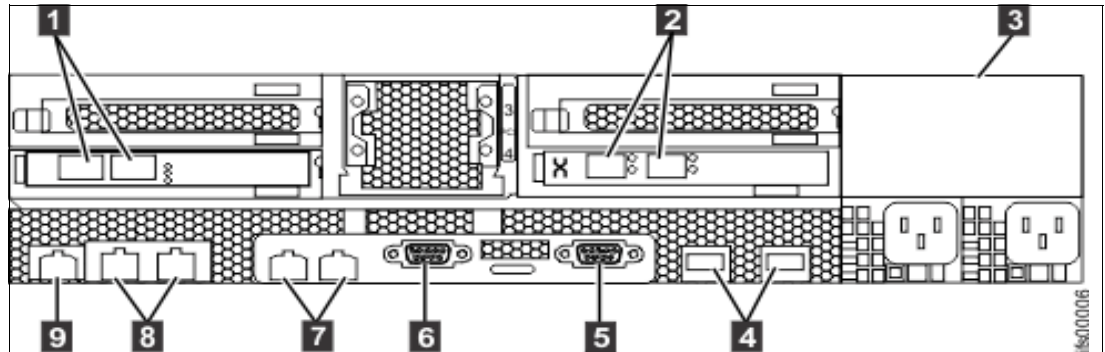


Figure 11-1 File module rear

- Port 1 - 1Gb 7 left <required> Internal connection between file modules. Connect a short cable from this port to port 1 on the other file module.
- Port 2 - 1Gb 7 right <required> Internal connection between file modules. Connect a short cable from this port to port 2 on the other file module.
- Port 3 - 1Gb 8 left <required> Provides management connection and optional data.
- Port 4 - 1Gb 8 right <optional> Alternate management connection and optional data.
- Slot4-0 - 10Gb 2 right <optional> Data connection only.
- Slot4-1 - 10Gb 2 left <optional> Data connection only.

### Control Enclosure

Use Figure 11-2 as a reference for plug locations. Connect the cable to the ethernet port and route the neatly to the rack cable management system.

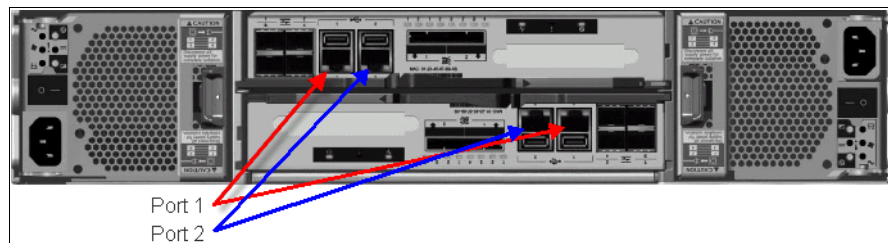


Figure 11-2 Control Module Rear

- Port 1 - 1Gb <required> Management and service connection.
- Port 2 - 1Gb <optional> Alternate management connection.



## SAS cables

If Expansion Enclosures have been installed, then they are connected to the control enclosure using the SAS cables shipped. If no expansion enclosures are installed, then skip this step.

The control enclosure has 2 SAS ports on each node canister. Port 1 from each canister form a pair of chains and these normally connect to the expansion enclosures racked below the control enclosure. Port 2 from each node canister form the second chain which normally connect to the upper expansions. The top canister always connects to the top canister of the next enclosure and the bottom to the bottom.

Connect the Expansion Enclosures using the following chart. Refer to Figure 11-3 on page 141 for port location.

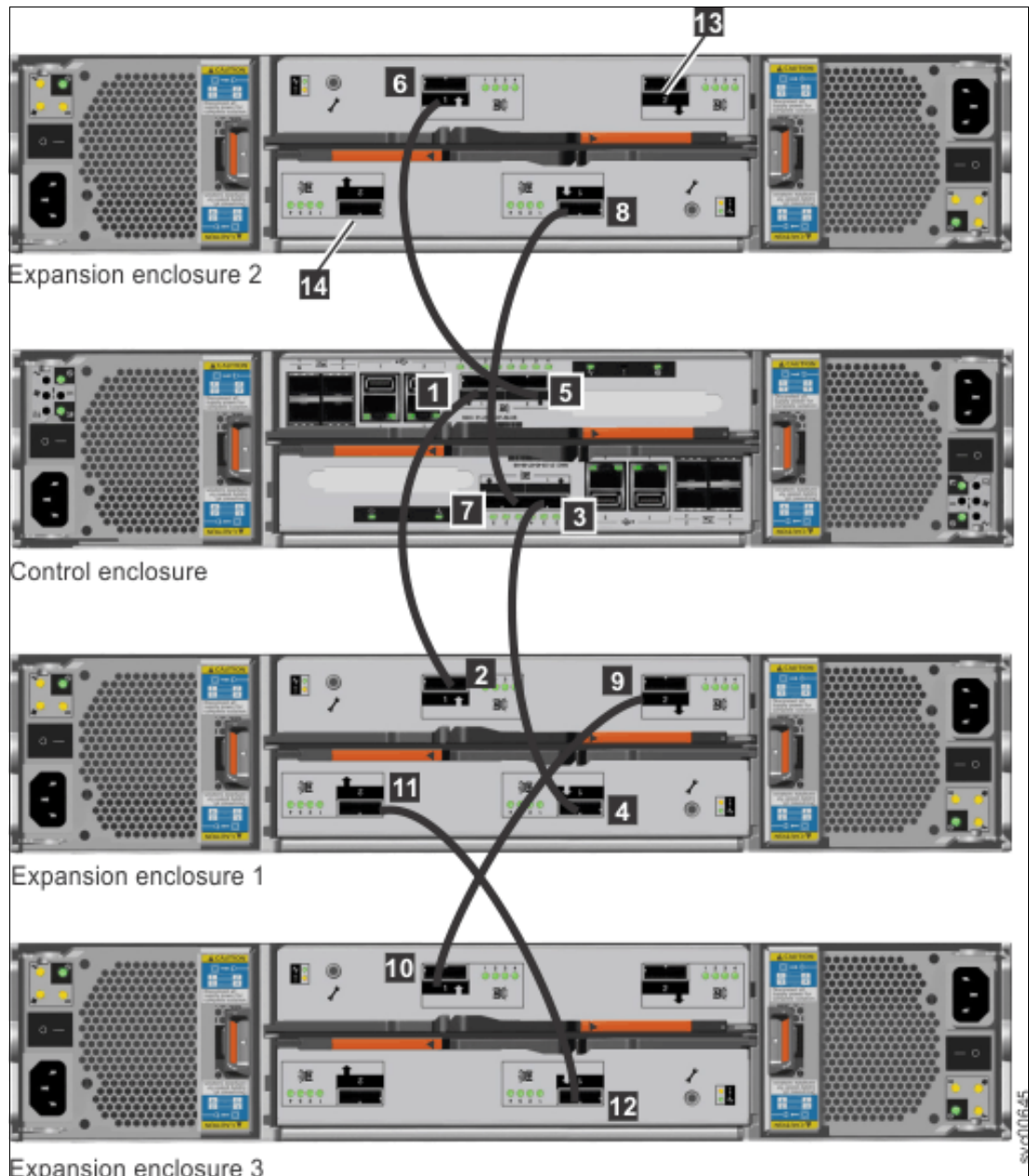


Figure 11-3 SAS cabling for 3 expansions

In Table 11-2 we show the SAS connections

*Table 11-2 SAS connections*

SAS Connections: How each unit connects to the next unit in the chain		
First Unit	Second Unit	Number of Expansions
<b>Controller</b>	<b>Expansion 1</b>	<b>1 Expansions</b>
Upper canister port 1	Upper canister port 1	
Lower canister port 1	Lower canister port 1	
<b>Controller</b>	<b>Expansion 2</b>	<b>2 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 1</b>	<b>Expansion 3</b>	<b>3 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 2</b>	<b>Expansion 4</b>	<b>4 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 3</b>	<b>Expansion 5</b>	<b>5 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 4</b>	<b>Expansion 6</b>	<b>6 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 5</b>	<b>Expansion 7</b>	<b>7 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 6</b>	<b>Expansion 8</b>	<b>8 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	
<b>Expansion 7</b>	<b>Expansion 9</b>	<b>9 Expansions</b>
Upper canister port 2	Upper canister port 1	
Lower canister port 2	Lower canister port 1	

### Fiber optic cables

The default configuration is short wave SFPs which require MM cables with LC connectors. If long distance is required, the SFPs can be specified as long wave, in which case it is important to use the appropriate single mode (SM) cable.

Dress cable gently and carefully being sure to avoid kinks and tight bends. Do not use cable ties or any other “hard” material to hold cables as these cause kinks and signal loss. Velcro cable wraps provide the most economic and safe method of tying cables.

### **File Module**

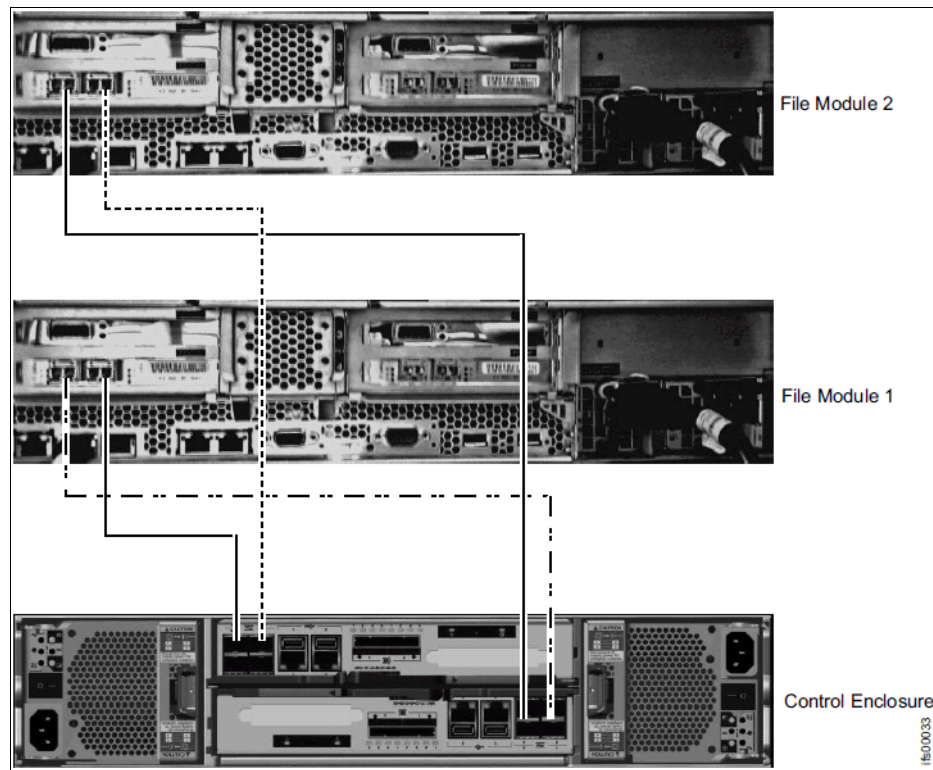
10Gb ethernet cables have already been covered in the ethernet section previously. The other fiber optic cables connected to the file module are Fibre Channel (FC) cables that connect to the control enclosure.

There are 4 FC cables required. Two from each file module to each node canister in the control enclosure. These are connected in a mesh pattern to provide full redundancy of pathing.

Select short cables as the file module and control enclosure should normally be close to each other in the same rack. Connect the cables as shown in Table 11-3 and refer to Figure 11-4 for connector locations

*Table 11-3 Cable connections.*

<b>File module</b>	<b>Controller</b>
File module 1 - Fibre Channel slot 2, port 1	Upper canister Fibre Channel port 1
File module 1 - Fibre Channel slot 2, port 2	Lower canister Fibre Channel port 1
File module 2 - Fibre Channel slot 2, port 1	Upper canister Fibre Channel port 2
File module 2 - Fibre Channel slot 2, port 2	Lower canister Fibre Channel port 2



*Figure 11-4 File to Control modules, Fibre Channel cables*

**Control enclosure**

If block storage is being configured, then FC connections are required to your SAN fabrics. Two FC ports are available on each node canister for connection to the SAN. These should be connected in a mesh arrangement to provide fabric path redundancy.

## 11.4 Power on and checkout

We describe how to power on and checkout in the topics that follow.

### 11.4.1 Network

Ensure all network devices are powered on and ready including ethernet and SAN switches. Access should also be available to network resources from the Storwize V7000 Unified, including DNS, NTP and e-mail server.

### 11.4.2 Power on expansions and controllers

Power up the disk storage first. Start with the expansion enclosures, then the control enclosure followed by the file modules. Ensure each group of enclosures is successfully powered on before beginning the next.

**Expansion enclosures**

Power on both power supplies using the switch on each one.

Successful power on is indicated by:

- ▶ Power supply leds, power LED on, 3 fault leds off, on both power supplies
- ▶ Front left end cap, power LED on, all others off
- ▶ Canister leds, status LED on, fault LED off on both canisters
- ▶ Drive leds will come on as each drive becomes ready

**Control enclosure**

Power on both power supplies using the switch on each one.

Successful power on is indicated by:

- ▶ Power supply leds, power LED on, battery good LED on or flashing (if charging) 4 fault leds off, on both power supplies
- ▶ Front left end cap, power LED on, others ignore
- ▶ Canister leds on both canisters, power status LED on, fault LED and system status will depend on current state and can be ignored at this time
- ▶ Drive leds will come on as each drive becomes ready

### 11.4.3 Power on file modules

The power management and support functions are running as long as there is AC power on one or both of the power cables, but the main processor is not power on. When power is applied to the cable, the management functions will start up, this takes approximately 3 minutes. Once complete, as indicated by a flashing power LED on the front of the module, the unit is ready to be powered on.

Connect a keyboard and display to the module and a mouse if a new install. Monitor the boot process on the display.

Power on by pressing the power button on the front of the module. The power LED will light on solid and the boot process will start. If the file module has code loaded, the module will boot and completion will be indicated by a flashing blue attention LED. If no code is present, a message will be displayed on the display to indicate that the boot has failed.

## 11.5 Install latest software

Ensure that as part of the implementation, the latest software levels are installed. In most cases a more recent level will be available than that shipped on the module, so an upgrade will be required. Or if the File Modules have no software installed, then the full package can be installed using the following procedure. The control enclosure (Storwize V7000) will be automatically upgraded, if required, as part of the file module upgrade.

In most cases the Storwize V7000 Unified will be preloaded and a software upgrade can be performed once the cluster initialisation and configuration is complete. In this case you can skip the rest of this section and go to 11.6, “Initialize the system” on page 149.

If the previous state of the cluster is not known or the integrity of the software is suspect, then perform the full restore procedure to begin with a clean software install.

For details on the software structure and concurrent upgrade processes refer to 11.5, “Install latest software” on page 147. As the full DVD restore process is rarely required and if done is part of an initialization, it is included here.

### 11.5.1 Determine current firmware and code levels

This can be difficult to determine, as it generally requires the system to be booted and running and logon user and password details known. If the intent is to restore the code regardless, then the previous level is not important and the restore can proceed.

To determine the file module software level, connect a keyboard and display to the module. Boot and wait for the login prompt. The initial prompt will be preceded by the host name. If console messages have been displayed, press enter to restore the prompt message.

login:

The default user/password is admin/admin. If the cluster has been in use previously the user settings may have been changed, so you will need to know a valid user login.

Once logged in, issue the command **lsnode**. This will display the file nodes. The software level is listed under the heading “Product version”. If the modules are not currently a cluster you may need to repeat the process on the other module. If the modules are at different software levels, the node performing the initialize (the one you put the key in) will reload the other node to the same level as itself.

If it is desired to know the level of the Storwize V7000 control enclosure, connect to the service assistant directly using the service IP address of one of the nodes, or to the management address followed by “/service” and logon using the superuser password. If the password or IP are not known and not default, then wait until the initialization step at 11.6.2, “Initialize the Storwize V7000 controller” below is complete, then you will be able to connect with the management IP address you defined and the default superuser password of *passwd0rd*.

## 11.5.2 Preparation for reload

We discuss preparing to reload the code in the topics that follow.

### Download and prepare software

To do a full DVD restore, you will need to obtain the ISO image from IBM support.

Using a suitable application, burn the DVD from the xxx.iso file. This file will be very large, over 4GB and will create a bootable DVD containing all the required software.

The file modules are restored individually. If time is critical, you may choose to burn 2 DVDs and perform the installs at the same time, although the process ejects the DVD after about 25 minutes, so the second module can be started then, which works quite well.

### Prepare the hardware

**Caution:** This process re-installs the software from scratch. Perform this action only if you have determined it is required.

The file modules need to be installed in the rack and power connected. The power indicator on the front of the module will be flashing to indicate the power control module is booted up and the server is powered off and ready to start.

Connect a standard PC monitor and keyboard (or equivalent KVM tool) to the module.

The following steps are not required with a new install and are only needed if the file module has been redeployed or the install aborted and cleanup is required. These steps may also be required as part of a recovery. The following steps should only be done if required and are included here for completeness.

1. If this module has previously been used as a file module, then to achieve a clean install, the previous configuration should be removed. To achieve this, the software must be booted and logged into from the local keyboard and monitor using access. The entire directory `/persist/` should then be removed. Contact IBM support for assistance if this step is required.
2. If any hardware has been replaced in the server or if the BIOS settings are in doubt, then perform the BIOS reset procedure found in the Information Center. This details the following process:
  - a. Power on the server using the power button on the front
  - b. When the BIOS splash screen is displayed, hit F1 to invoke the BIOS setup.
  - c. Select *Load Default Settings*
  - d. Then select *Boot Manager*
  - e. Select *Add device*
  - f. Select *Legacy only*
  - g. Exit the panels using escape key and save the configuration on exit
3. If necessary, configure the RAID controller on the server for the mirrored disk drive used for the system using the procedure in the Information Center.

## 11.5.3 Re-install the software

Power on the server. If you have just connected the ac power, you will need to wait a few minutes for the power controls to initialize before being able to power on. If already on, reboot the server using Ctl-Alt-Del or if in BIOS, by exiting the BIOS setup utility.



Insert the DVD in the drive. Power needs to be on to open the DVD drive, so as soon as the server is powered up, load the DVD.

Watch the video monitor and wait for the DVD to boot. The software install utility will immediately prompt for confirmation with this message:

- To install software on a node, press the <ENTER> key.

\*NOTE\* - this will destroy all data on the node

- Use the Function keys listed below for more information

[F1 - Main] [F5 - Rescue]

boot:

Press the enter key to begin the install process.

The utility will now build the disk partitions if needed and begin loading the Linux operating system and Storwize V7000 Unified software. There will be at least 2 reboots to complete the install. The first phase copies the software to the disk. Progress can be monitored on the video. After about 25 minutes, the tool will eject the DVD and reboot the server.

**Note:** After 25 minutes the DVD is ejected and can now be used to begin the install process on the other server if required.

Then the server is booted from its disk and Linux installs its components and builds the Linux operating system and installs the Storwize V7000 Unified software. This phase takes about 30 minutes.

When this process is complete the server is booted up on the operational code and is ready for use. This boot takes less than 5 minutes.

Successful install and preparation of the file module is indicated by the blue attention light on the server flashing.

Repeat the process for the other file module.

## 11.6 Initialize the system

The next step is to initialize the system and build a cluster incorporating the Storwize V7000 storage (control enclosure and any expansion enclosures) and the 2 file modules.

### 11.6.1 Configure USB key

A USB key was shipped with the Storwize V7000 Unified. If this has been misplaced then any USB mass storage key can be used, although some models of key are not recognized, so it may require trying a few to succeed. Take care not to use a large capacity key as this can cause the key to not be recognized.

The key must be formatted with a FAT32, EXT2, or EXT3 file system on its first partition.

The shipped key is formatted and preloaded with the initialization tool. If this is missing or you are using a replacement key, then this tool can be downloaded from IBM's support site

<http://www.ibm.com/storage/support/storwize/v7000/unified>

**Note:** There should be a key shipped with the Storwize V7000 controller unit as well as one with each file module. Use the key shipped with a file module as it should have the latest version of the inittool. If the latest tool is downloaded, then any of the keys will work.

Keep the keys stored and secure for later use in case of recovery, rebuild or redeployment.

1. Insert the key into any Windows XP or higher workstation. If the tool does not auto launch, then to the USB key and run *InitTool.exe*. This will launch a window as shown in Figure 11-5 on page 150.

**USB key files:** the USB key will contain the following files

```
autorun.inf ..... windows auto run file
inittool.exe ..... Storwize initialization tool
```

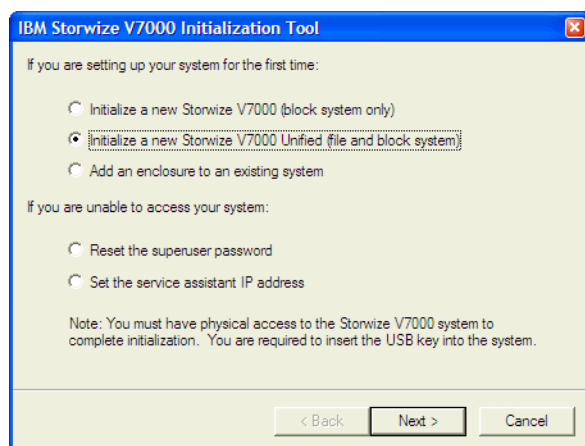


Figure 11-5 Init tool first screen

2. Set the storage Service IP addresses. The service IP addresses assigned to the Storwize V7000 storage enclosure nodes are not considered part of the cluster configuration. These addresses are the base addresses for the nodes and are active independent of the cluster software. They are not set or changed by the cluster configuration. It is important to set these addresses in case access is needed during recovery.

Use the set service assistant IP address option on the init tool to prepare the usb key. Insert the key in the enclosure nodes to set their address. Note, the tool creates one address, so the procedure needs to be done twice, once for each node.

3. Select option "Unified (File and Block)". This universal tool is used for all Storwize V7000 installs. The first option, "block system only" will set up the initialisation process for installing just a Storwize V7000 storage controller. The second option is required to initialize both the storage controller and the file modules for a unified configuration.
4. Click next to get the first setup options as shown in Figure 11-6.

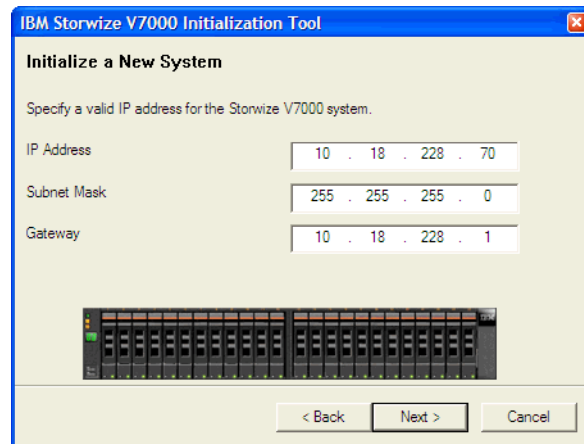


Figure 11-6 Init tool system settings

This IP address is the management address for the “System” component, which is the Storwize V7000 storage controller. As discussed in the planning section, all the management IP addresses must be in the same subnet.

5. Enter the IP address, mask and optional gateway, then click next to get the screen shown in Figure 11-7 on page 151.

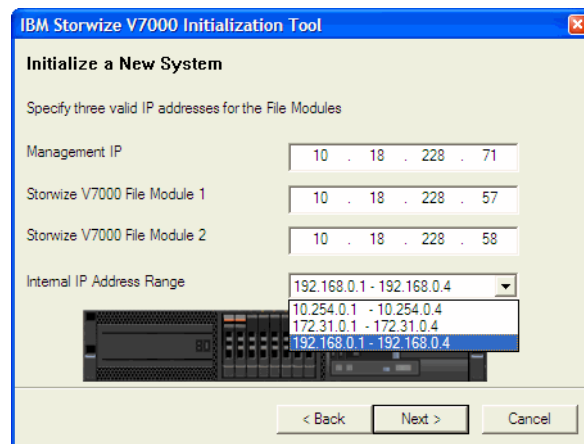


Figure 11-7 Init tool file settings

6. On this panel, give the IP details for the Unified management interface, and also the IP addresses for the File modules 1 and 2. Select from the pull down a range of addresses that are not being used in your network. These addresses are used internally between the modules and cannot be accessed from your network.
7. Click Next to see the instructions to use the key as shown in Figure 11-8.

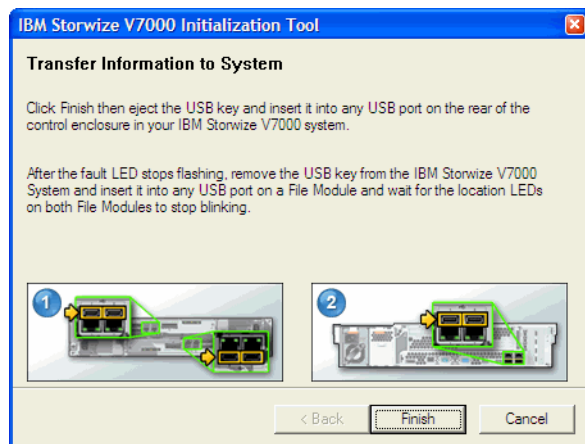


Figure 11-8 Init tool final screen

8. Click Finish. Confirm that the key now has 2 new files, *satask.txt* and *cfgtask.txt*

**Hint:** In case the following steps need to be retried, it is wise to make a copy of the files on the key on to your workstation now so it is easier to resume from this point.

9. Eject and remove the USB key from your workstation.

**USB key files:** the USB key will contain the following files

```
autorun.inf ..... unchanged
inittool.exe ..... unchanged
satask.txt ..... Storage command file. Contains the initialize command
cfgtask.txt ..... File module command file. Contains initialize command
```

## 11.6.2 Initialize the Storwize V7000 controller

1. Ensure that both the node canisters are in “candidate” status before continuing. If the service IP is known on either node then browse to this address and logon using the superuser password (default:passw0rd). The status of the nodes is displayed on the home screen. Or, confirm the status of the three leds on each canister to be on-off-flashing. Read from the right on the top canister and from the left on the bottom one.
2. If the status is incorrect, you will need to connect to the service assistant interface to resolve this. The IP address can be set using the USB key if not known. If the status shows as service, then perform the action to remove from service. If the status is active, then an operational cluster is present and this needs to be removed. Follow the Storwize V7000 procedures to remove the cluster or contact IBM for assistance.
3. Once you have both nodes in candidate status, insert the key in any USB port on the rear of your Storwize V7000 controller enclosure. It is better to use a USB port on the top canister as the canister executing the initialize command will become the first node, node1. This helps prevent confusion later with node1 now being in slot 1. The fault LED will begin flashing. When the fault LED stops flashing the task is complete and the key can be removed.

**Hint:** After the command has been executed, a result file is written back onto the key as well as a SSH key file. You can review the result file *satask\_result.html* to confirm the creation of the nascluster was successful. Also, when successfully executed the command file (satask.txt) is deleted to prevent it accidentally being run again.

4. Confirm also that the three status leds are on-off-on both canisters in the enclosure, indicating that the node canisters are active in a cluster.

**Hint:** In case the following steps need to be retried, make a copy of the files on the key onto your workstation now so it is easier to resume from this point.

**USB key files:** the USB key will contain the following files

```

autorun.inf ..... unchanged
inittool.exe ..... unchanged
..... If successful, satask.txt has been deleted
cfgtask.txt ..... File module command file. Contains initialize command
NAS.ppk ..... SSH key file, needed by the file module
satask_result.html . Result output from the initialize command

```

### 11.6.3 Initialize the file modules

1. Confirm that both the file modules are booted up and ready to be initialized. This is indicated by the flashing blue attention indicator on each file module. Both must have this LED flashing before continuing.
2. Now insert the key into any USB port on one file module.
3. The blue LED will come on solid and remain on until the process has completed. If the blue attention LED on the module the key is inserted in begins flashing again, this indicates a failure. If this happens, remove the key and interrogate the results file. Refer to the Information Centre for analysis of the error and recovery actions. Successful initialisation is indicated by both blue LEDs going off.
4. Wait for the initialization process to complete, (both attention lights going off). Normal initialisation should take approximately 15 minutes, but can be extended if the other file module needs to be upgraded to the same code level (plus 1 hour) or if the control enclosure (Storwize V7000) requires code upgrade (plus 2 hours).
5. Remove the USB key and insert the key into your workstation and either:
  - a. Review the results file in an editor, or
  - b. Start the InitTool program from the key (if it didn't auto run). This will inspect the results file and give a message to say the initialize was successful.

**USB key files:** the USB key will contain the following files

```

autorun.inf ..... unchanged
inittool.exe ..... unchanged
..... If successful, cfgtask.txt has been deleted
..... NAS.ppk will be deleted when copied to File module
satask_result.html . unchanged
SONAS_result.txt ... Result output from the initialize command

```

6. Confirm that you can access the cluster. Using the browser on a workstation that has IP connectivity to the management ports of the Storwize V7000 Unified, point the browser to [https://<management\\_port\\_ip>](https://<management_port_ip>). This is the management IP you assigned to the Storwize V7000 Unified cluster when initializing the usb key.

## 11.7 Base configuration

The Storwize V7000 Unified cluster is now created and consists of two main components; the Storwize V7000 storage consisting of a control enclosure and optional expansion enclosures, and the file server consisting of 2 file modules. These can all be managed from the Storwize V7000 Unified GUI interface, which is presented from the primary File Modules management IP address.

To use the cluster, it must first be configured. A setup tool, called “EZ-Setup”, is provided that will run once only the first time the cluster is logged into after an initialization. The steps to configure the cluster using EZ-Setup are described next.

**Note:** EZ-Setup will only run once and cannot be manually started. Any configuration options that are skipped will need to be manually setup or changed from the appropriate configuration panels later.

### 11.7.1 Connect to GUI interface

Using the browser on a workstation that has IP connectivity to the management ports of the Storwize V7000 Unified, point the browser to `https://<management_port_ip>`. You will be presented with the logon screen as shown in Figure 11-9. Be sure you are connected to the Storwize V7000 Unified and not the Storwize V7000 storage control enclosure directly.

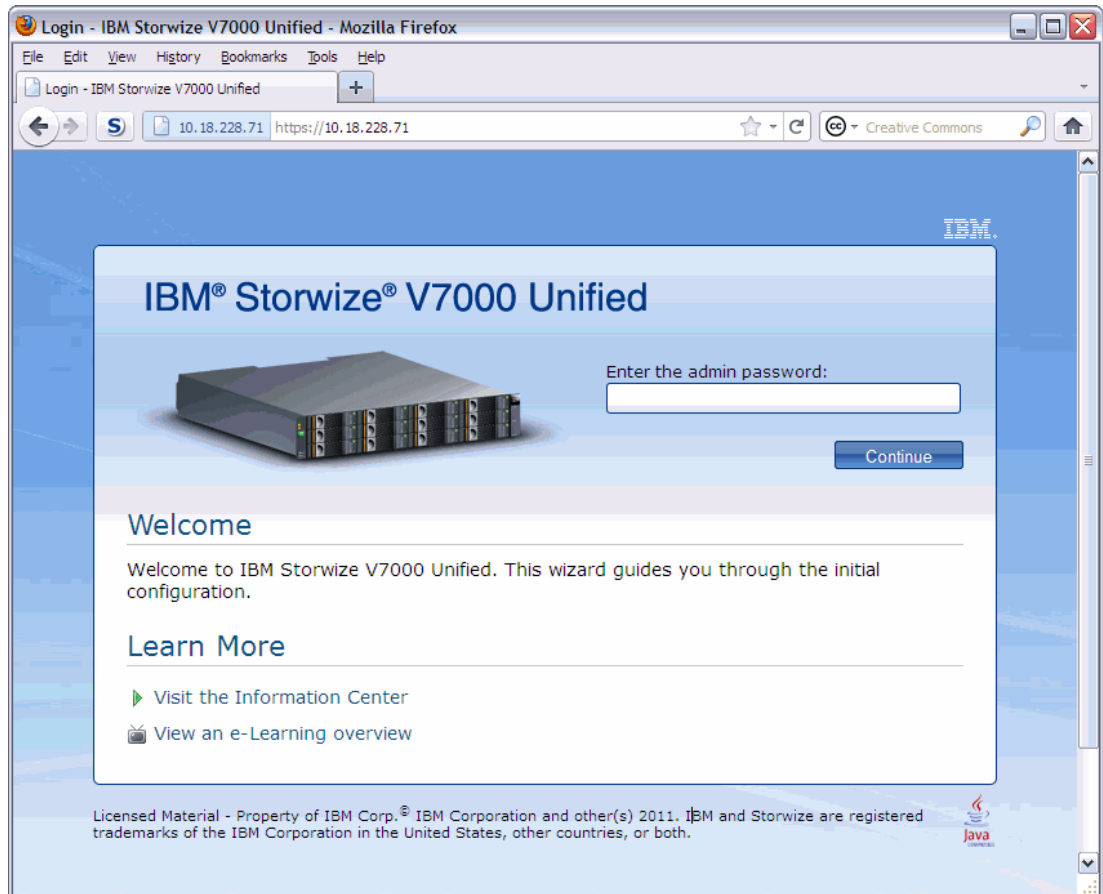


Figure 11-9 Easy Setup login

This login screen leads into the Easy Setup wizard, which will guide you through the base configuration of the machine. It is important to have all the information about how you want to configure the cluster ready before continuing.

Default user is **admin** and the default password is **admin**. For this initial setup login, only admin is available. Enter the password and click continue.

## 11.7.2 Easy Setup wizard

**Tip:** It is important to have all your parameters ready and preferably written out on your planning sheet and also be sure that the various servers are operational and ready. Many of the configuration processes described in the following procedures will test and confirm that the resources you address (e.g. DNS servers, authentication servers, gateways, etc) are actually reachable and operating and the setup steps may fail if they cannot be contacted and connected to.

### License Agreement

The first screens that Easy Setup presents are for the license agreement as shown in Figure 11-10 on page 156. You should read the license agreements on each tab and then click the agree button. Then click Next to continue.

**License Agreement (Step 1 of 2)**

Read the license agreement carefully.

License | IBM Notices | Java Notices | Non-IBM Licenses | IBM ASU License and Notice | Additional Licenses and Notices

**IBM Storwize V7000**

**IBM Storwize V7000 File Module**

International Program License Agreement

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM, LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,

- DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN "ACCEPT" BUTTON, OR USE THE PROGRAM; AND
- PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND PROOF OF ENTITLEMENT TO THE PARTY FROM WHOM IT WAS OBTAINED FOR A REFUND OF THE AMOUNT PAID. IF THE PROGRAM WAS DOWNLOADED, DESTROY ALL COPIES OF THE PROGRAM.

1. Definitions

"Authorized Use" - the specified level at which Licensee is authorized to execute or run the Program. That level may be measured by number of users, millions of service units ("MSUs"), Processor Value Units ("PVUs"), or other level of use specified by IBM.

"IBM" - International Business Machines Corporation or one of its subsidiaries.

"License Information" ("LI") - a document that provides information and any additional terms specific to a Program

☐ I agree with the terms in the license agreement.

☒ I do not agree with the terms in the license agreement.

Figure 11-10 License agreement

### System Attributes

Enter the required fields in the system attributes screen as shown in Figure 11-11.



**System Attributes (Step 2 of 9)**

System name:  
SanJose1

NetBIOS name:  
SanJosev7ku

Time Zone  
America/Los\_Angeles

Network Time Protocol (NTP) server address:  
10.31.2.80

Alternate Network Time Protocol (NTP) server address:  
10.31.2.81

< Back   Next >

Figure 11-11 System attributes

System name	The name of this Storwize V7000 Unified. This name appears on the screens and is the name by which you will know this cluster.
NetBIOS name	The NETBIOS name by which this cluster will be seen on the network. This name is used in the SMB protocol.
Time zone	Choose the time zone from the selection that best represents where this machine is.
NTP server	Enter the IP address of your NTP server. This is a required field. NTP is required to ensure accurate synchronisation between the file modules and to resolve locks.
Alternate NTP	Optional field. If you have a backup NTP server, enter it's address here.

Click next. A progress window will display while configuration changes are made. Wait for "Task Completed", then close.

## System License

If you intend to attach any storage externally to this cluster (internal disks are already licensed), then you require a virtualisation license. This license is enterprise wide, so can be split across several clusters in your business. Enter the value as the number of enclosures being externally managed by this cluster. This value will depend on the type of devices being managed. If unsure, contact your IBM Representative for assistance to determine what value you have purchased.


Enter the value as shown in Figure 11-12. You must have sufficient virtualisation license(s) for all MDiskS being attached to this cluster. If using remote copy features, then also put in this value.

**System License (Step 3 of 9)**

The enclosure license already includes virtualization of internal Serial Attached SCSI (SAS) drives on your IBM Storwize V7000 Unified system. You can use this panel to set any additional options. If you are sharing the total authorized capacities across multiple clusters, enter only the capacities you wish to use on this cluster. The sum of the capacities across all clusters must not exceed your authorized capacities.

**Set License Options**

External Virtualization Limit  
 enclosures

Remote-Copy Limit   
 enclosures

< Back   Next >

Figure 11-12 System license

If not using external storage, remote copy, or compression then leave these values at “0”.

The optional Real-time Compression feature is available by license which is based on a per-enclosure basis.

Click Next.

## Support Notifications

The next screen as shown in Figure 11-13 on page 158 gives the option to configure the support notifications now. This is not mandatory at this time and can be deferred if desired.

**Support Notifications (Step 4 of 9)**

To ensure that your system continues to run smoothly, IBM Support can be automatically notified when serviceable events occur. To enable and configure support notifications, click the button below.

[Configure Support Notifications Now](#)

< Back   Next >

Figure 11-13 Support notifications

To set values, click “Configure Support Notification Now” button.

The next screen asks for your e-mail information as shown in Figure 11-14 on page 159.

**Configure Support Notifications** *Step 1 of 2*

**Contact Information**

Email server IP address: 10.18.228.250 [Test Connection ...](#)

Company name: IBM Redbooks

Customer email: red@books.com

Customer telephone number: +1-555-1234-5678

Off-shift telephone number:

IBM Support email address: callhome1@de.ibm.com

[Next >](#) [Cancel](#)

Figure 11-14 Support notifications

The first field is the IP address of your e-mail server which needs to be accessible to the cluster and will allow the cluster to send e-mail.

Also enter your Company name, contact e-mail address and prime shift telephone number. The off-shift number is optional and only required if the prime number is not 24 hour. These numbers need to be phones that will be answered at anytime so that IBM Support personnel can contact you in the event of the cluster calling call home.

The last field should be left as default. This is the address in IBM for “call home” alerts.

Click next to display the second screen as shown in Figure 11-15 on page 159.

**Configure Support Notifications** *Step 2 of 2*

**Outbound Connectivity**

☐ Enable a proxy server to access the Internet

[< Back](#) [Finish](#) [Cancel](#)

Figure 11-15 Support notifications - proxy

If your network requires proxy authentication for external access, click the enable button and complete the proxy details. When complete, click Finish.

Confirm that the details are correct on the support notifications summary screen as shown in Figure 11-16 on page 160 and then click Next.

**Support Notifications (Step 4 of 9)**

To ensure that your system continues to run smoothly, IBM Support can be automatically notified when serviceable events occur. To enable and configure support notifications, click the button below.

[Configure Support Notifications Now](#)

**Email server IP address:** 10.18.228.250  
**Company name:** IBM Redbooks  
**Customer email:** red@books.com  
**Customer telephone number:** +1-555-1234-5678  
**IBM Support email address:** callhome1@de.ibm.com

[< Back](#)
[Next >](#)

Figure 11-16 Support notifications complete

## Domain Name Service

Next enter your domain name information on the screen as shown in Figure 11-17. Note: the process will test the servers listed are present and fail if they cannot be contacted.

**Domain Name Service (DNS) (Step 5 of 9)**

\* DNS domain name:

redbook.ibm.com

\* DNS servers:

DNS Server		
10.18.228.250	+	-
10.18.228.251	+	-
10.18.228.252	+	-

DNS search domains:

DNS Search Domain		
ibm.com	+	-

[< Back](#)
[Next >](#)

Figure 11-17 Domain name service

- Domain name** This is the public network domain which is appended to your cluster name. This will typically be common to your whole enterprise. E.g. customer.com
- DNS servers** IP address of your DNS server. To add a server IP to the list, click the “+” and use the “X” to delete an entry. Add as many DNS server entries as desired. At least one is required.
- DNS search domains** Optional. Additional domain names that should be searched.
- When complete, click Next.
- A progress panel will display, wait for “Task Completed”, then close.

## Authentication

On this next screen (shown in Figure 11-18) you have the option to set the method that will be used for file access authentication and control. This is optional at this time, so this step can be

deferred and skipped. However, no file access is possible until this section has been configured.

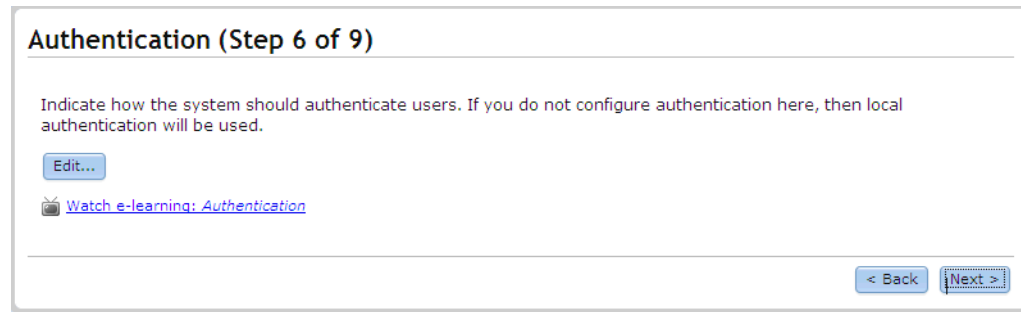


Figure 11-18 Authentication

To continue, click “Edit”, to bypass click “Next”.

After clicking edit, a confirmation screen will appear, click ok.

Choose from the available authentication methods as shown on Figure 11-19 by clicking the appropriate button, then Next.

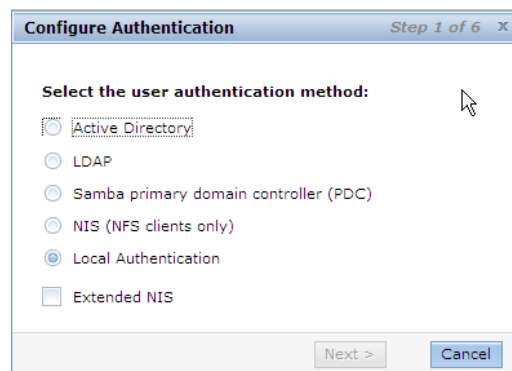


Figure 11-19 Select authentication

Only one form of authentication can be active. Active Directory and PDC both have the option to add Extended NIS if Netgroup support under these authentications is needed. Depending on your choice, you will be taken through a series of setup screens as follows:

### **Active Directory**

If you chose Active Directory then your next screen will be as shown in Figure 11-20.

Active Directory Step 2 of 2

Specify the Active Directory settings:

\* Server:

Administrator's credentials to access the domain:

\* User ID:  \* Password:

☐ Enable Services for UNIX (SFU)

Domain name	Lower Range	Upper Range	Schema Mode	
<input type="text"/>	<input type="text"/>	<input type="text"/>	sfu	<input type="button" value="+"/> <input type="button" value="X"/>

Figure 11-20 Active Directory settings

Enter the required information as follows:

AD Server	This is the IP address of the Active Directory server.
User ID / password	Userid and password that has sufficient authority to connect to the AD server and access authentication and mapping information. Typically this will be the Administrator userid, or an ID with equivalent authority.
Enable SFU	If support for UNIX is required then enable the Service for Unix (SFU) feature and complete the configuration box below
Domain name	This is the domain that SFU belongs to.
UID/GID range	Lower to upper limits of the range of User and Group IDs that will be used by the AD server.
SFU schema	The SFU schema mode being used.

Using the "+" add as many line items as required.

When done click Finish. A progress window will display while the configuration process runs. Wait for Task Completed then close the window.

## LDAP

For LDAP you will see Figure 11-21 on page 163.

**LDAP** Step 2 of 2

\* Specify one or more LDAP servers:

LDAP Server	Port
	389

+ ×

\* Search base for users and groups:

\* Bind distinguished name (DN):

\* Bind password:

\* Confirm bind password:

User suffix:

Group suffix:

Security method: off

☐ Enable Kerberos

\* Server Name:

\* Realm:

\* Key Tab File: Browse...

< Back Finish Cancel

Figure 11-21 LDAP settings

Enter the required information as follows:

LDAP Server	This is the IP address of the LDAP server. Click the “+” to add additional servers if desired.
Search base	The base domain suffix
DN	The root distinguished name.
Bind password	Userid and password required to access the LDAP server.
User/Group suffix	User and Group suffix as defined by the LDAP server.
Workgroup	The domain name for this cluster and the Active Directory.
Security method	Select the SSL mode that will be used. If SSL security is used, then a certificate file is needed. When this option is selected then a new box will appear for the Certificate. Click browse to locate the certificate file on your workstation.
Enable kerberos	If SSL is not used, then the option for kerberos appears. Tick the box to enable.
Kerberos name	Enter the name of the server.
Kerberos realm	Enter the kerberos realm
Key Tab File	Browse to the location of the Kerberos keytab file.

### **Samba PDC**

If Samba PDC was selected, then the PDC configuration screen will be presented as shown in Figure 11-22.

**Samba PDC - NT4** Step 2 of 2

\* Server host:

\* Administrative user ID:       \* Administrative password:

\* Domain name:       \* NetBIOS name:

< Back   Finish   Cancel

Figure 11-22 PDC settings

Enter the required information as follows:

PDC Server	This is the IP address of the NT4 PDC server.
User ID/password	The user id and password used to access the NT4 server that has administrative authority
Domain name	The NT4 Domain name
NetBios name	NT4 NetBIOS name

### **NIS Basic**

If using basic NIS authentication, you will get the setup screen as shown in Figure on page 164.

**NIS Basic** Step 2 of 2

\* Primary NIS domain:

\* Server map:

NIS Server	NIS Domain
<input type="text"/>	<input type="text"/>

+ X

< Back   Finish   Cancel

Figure 11-23 NIS basic settings

Primary domain	The name of the primary NIS domain
NIS Server/Domain	The address of each server. For each server, map the supported domains. Domains are entered as a comma separated list for each server. Use the "+" button to add additional servers.

### **Extended NIS**

For Active Directory and PDC authentication, there is an option to include extended support for NIS. If this was selected then you will also be presented with a configuration settings screen as shown in Figure 11-24



**Extended NIS** Step 3 of 3

\* Primary NIS domain:

\* Server map: ?

NIS Server	NIS Domain	
<input type="text"/>	<input type="text"/>	+ ×

☐ Enable user ID mappings ?

Domain map: ?

Active Directory Domain	NIS Domains	
<input type="text"/>	<input type="text"/>	+ ×

User map:

Active Directory Domain	Action	Mapped User	
<input type="text"/>	AUTO	<input type="text"/>	+ ×

User ID range: From:  To:

Group ID range: From:  To:

Figure 11-24 Extended NIS settings

Primary domain      The name of the primary NIS domain

NIS Server/Domain      The address of each server. For each server, map the supported domains. Domains are entered as a comma separated list for each server. Use the “+” button to add additional servers.

If NIS will be used for user ID mapping then tick the enable box and complete the remaining fields. If not, then this panel is complete.

Domain map      Add entries here to map the AD domain to the NIS domains. Use the “+” button to add lines.

User map      Add entries for user mapping exceptions as required.

### Local Authentication

Select the Local Authentication option to configure the system with an internal authentication mechanism in which users and groups are defined locally on this system. If this was selected then you will also be presented with a the screen as shown in Figure 11-25 on page 166.

**Note:** With Local Authentication you enter user and group information locally which is covered in Section 11.13.3, “Create local users using Local Authentication for NAS access” on page 189.

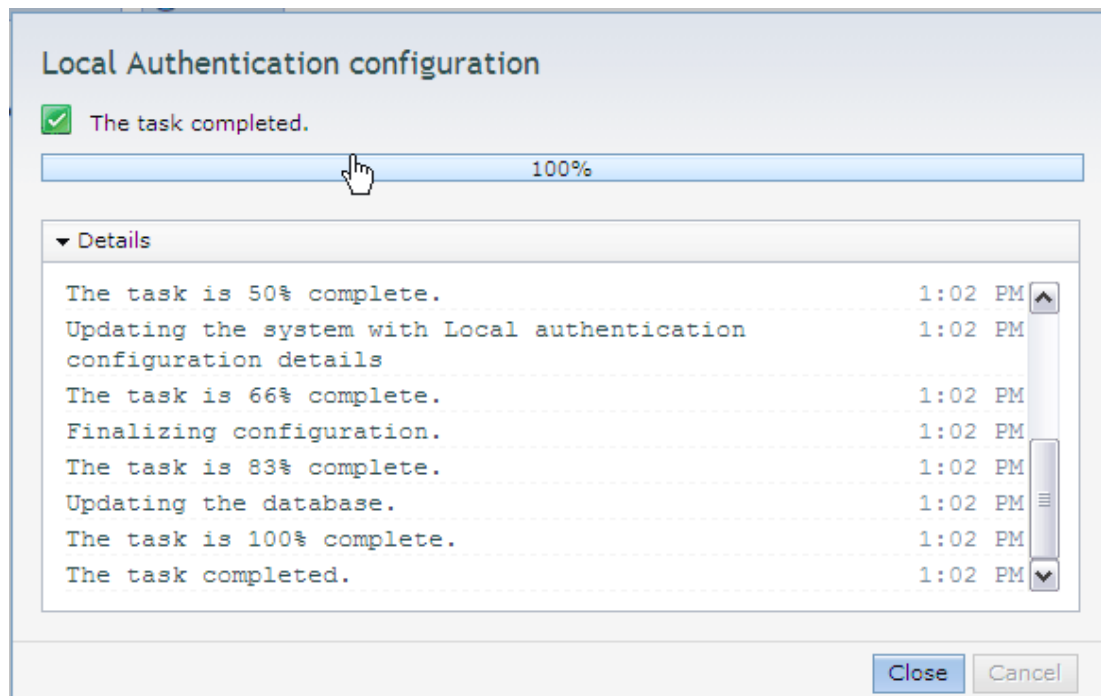


Figure 11-25

### Completion

When the chosen method has been configured and the processing is complete as indicated by “Task Complete”, close the status window. This will return you to the authentication easy setup screen. There will now be a summary of the method that has been configured (see example in Figure 11-26). Check to confirm, then click Next.

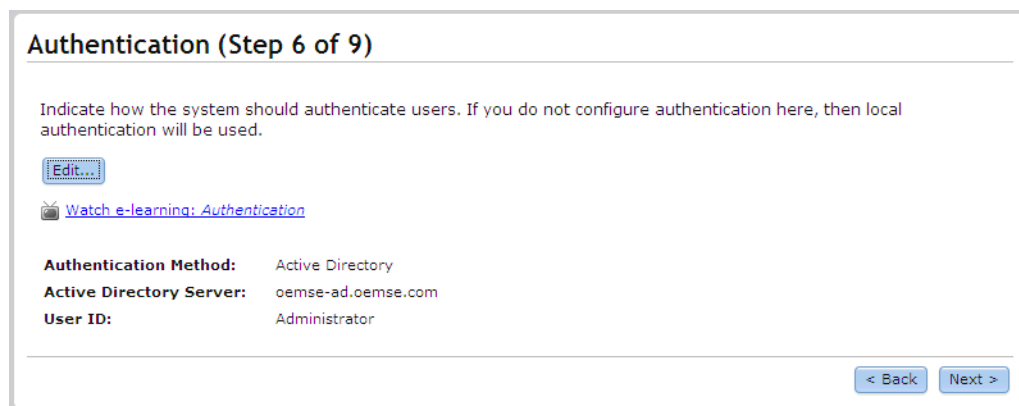


Figure 11-26 Authentication confirmation

### Hardware

You are now presented with a graphical representation of the hardware. Check that all modules and enclosures are correctly shown and that the cabling is correct. See our example in Figure 11-27.

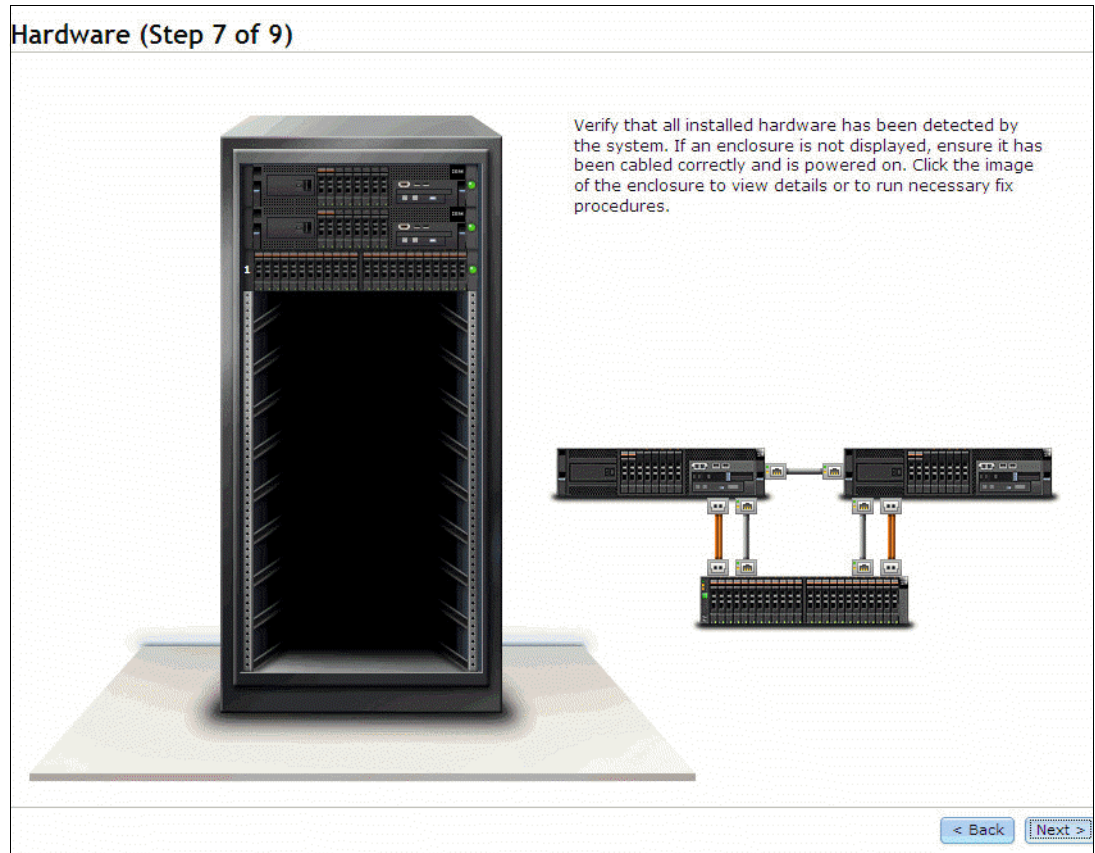


Figure 11-27 Hardware summary

Use the mouse to hover over each component to display more detail as shown in Figure 11-28 on page 168.

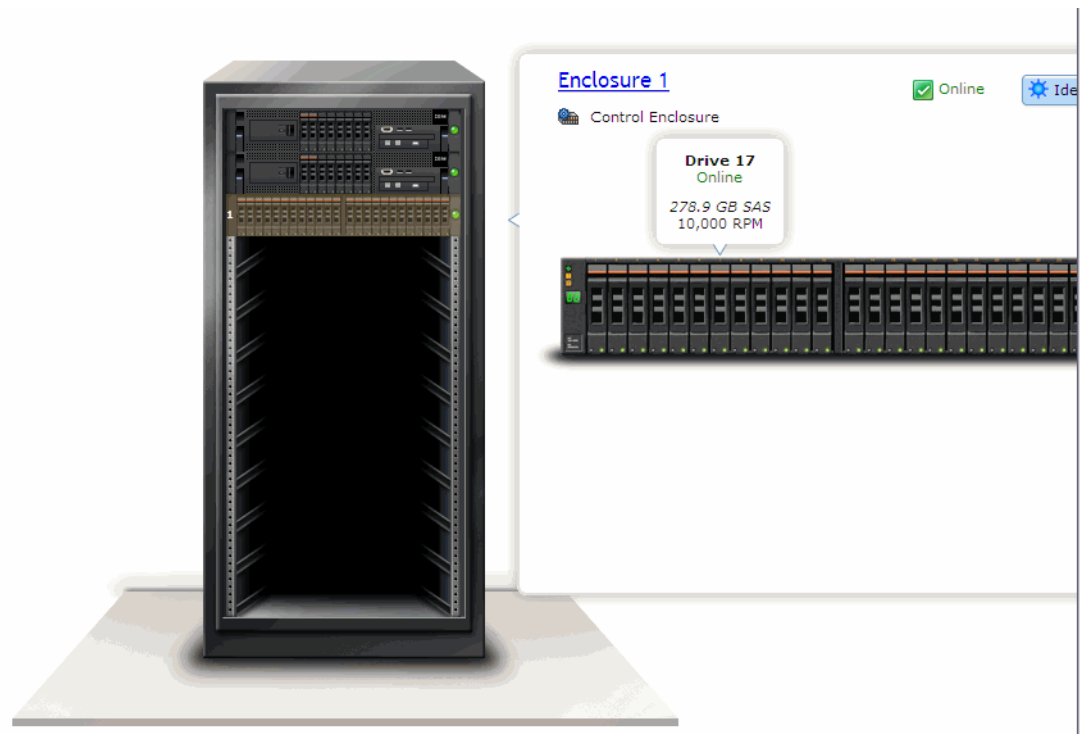


Figure 11-28 Hardware detail

If problems are found, attempt to resolve them now using the options available on this screen. Make sure all expansion enclosures are included in the graphic. If components are missing, ensure they are powered up and cabled correctly.

Click Next to continue

A task status window is displayed, wait for “Task Complete” then close.

## Configure storage

After the storage has been added you will be asked to configure it. The system will suggest a best practice configuration which is detailed on the next screen. If the default configuration is acceptable, tick the yes box to automatically configure the storage as shown in the example in Figure 11-29 on page 169. Otherwise, untick the box to skip auto configuration. You will then need to manually configure the storage later.

**Note:** Storwize V7000 Unified file module and SONAS give best performance using arrays with 8 data spindles due to the GPFS block size. The automatic configuration may therefore suggest 8+P arrays. Where possible, use 8+P or 4+P arrays. For more detail on this topic, refer to SONAS support documentation.

In a 24 drive enclosure, this conveniently gives a 8+P, 8+P, 4+P, S configuration.

### Configure Storage (Step 8 of 9)

Would you like to automatically configure internal storage now?

☒ Yes, automatically configure internal storage now.

**Storage Found:**  
(24 drives) 278.9 GB, SAS, 10000 rpm, io\_grp0

**Configuration Summary:**

**3 x Basic RAID-5** (278.9 GB, SAS, 10000 rpm, io\_grp0):

- 8, 8, 7 drives
- 1 Hot Spares
- 0 Unconfigured Drives

Figure 11-29 Configure Storage - example

## Public networks

You are now presented with a window to define the public networks as shown in Figure 11-30. It is necessary to configure several networks, as each network is tied to an ethernet logical interface.

### Public Networks (Step 9 of 9)

Configure the public network external clients can use to reach the system. This includes all network interfaces used to provide NAS services to external clients in the network.

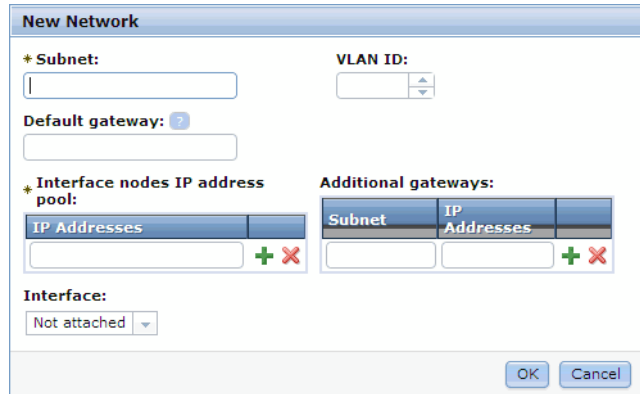
+ New Network
⌵ Actions
🔍 Filter...

Subnet	gateway	IP Addresses
<div style="color: red; font-weight: bold; font-size: 1.2em;">!</div> No items found.		

⏪
Showing 0 items | Selected 0 items
⏩

Figure 11-30 Public Networks

A minimum of 1 address per node is required. Each physical interface of the same speed on the node is logically bonded and the IP addresses are presented across that logical interface. Each network definition is assigned to a logical interface. To create a new network, click on the “New Network” button, which launches a window as shown in Figure 11-31 on page 170.



The 'New Network' window contains the following fields and controls:

- \* Subnet:** A text input field for the subnet definition.
- VLAN ID:** A dropdown menu for selecting a VLAN.
- Default gateway:** A text input field with a help icon (?) next to it.
- \* Interface nodes IP address pool:** A section with a table for IP addresses and a '+' button to add more.
- Additional gateways:** A section with a table for gateways and a '+' button to add more.
- Interface:** A dropdown menu currently set to 'Not attached'.
- Buttons:** 'OK' and 'Cancel' buttons at the bottom right.

Figure 11-31

Subnet	The IP subnet this definition is for. Note that the format of this value is using the CIDR syntax. xxx.xxx.xxx.xxx/yy where yy is the decimal mask given as the number of left justified bits in the mask.
VLAN ID	If VLANs are being used, then enter the vlan number here, otherwise leave the field blank. Valid values are 2-4095. Note VLAN 1 is not supported for security reasons.
Default Gateway	The default gateway (or router) within this subnet for routing. Not a required field if all devices needing to be connected to this interface are in the same subnet.
Interface Pool	Using the “+” button, add IP addresses to the pool for use on this logical interface. These must be in the same subnet as entered above. A minimum of 1 address is required.
Additional Gateways	If more than one gateway (router) exists in the subnet, add the IP addresses here.
Interface	Select the logical interface that this definition is assigned to.

A progress window will display, wait for the completion message. Click OK when complete.

Repeat the process for each network using the “New Network” button until all required networks have been defined. When done, click finish.

A reboot progress window will now display indicating that the file modules are being restarted. When complete close the window.

This is followed by the applying settings status window as shown in Figure 11-32.

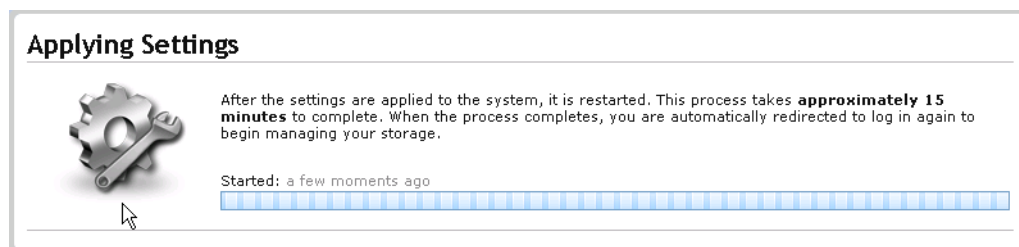


Figure 11-32 Applying settings

Wait for the process to complete. Easy Setup is now complete, the home page for the Storwize V7000 Unified is automatically displayed.

### 11.7.3 Setup periodic configuration backup

Management and configuration information is stored on the File Modules in a Trivial DataBase (TDB). It is recommended that you set up a periodic backup of the TDB at a suitable time, and we recommend daily. This backup may be required by service personnel if the TDB becomes lost or corrupted or in the event of a recovery.

1. Start an ssh session with the cluster management address that you set when initializing the cluster.
2. Logon with user **admin** and password **admin**.
3. Issue the command to perform the periodic configuration backup.

---

#### *Example 11-1*

```
[7802378.ibm]$  
[7802378.ibm]$ mktask BackupTDB --minute 0 --hour 2 --dayOfWeek "*"
EFSSG0019I The task BackupTDB has been successfully created.
EFSSG1000I The command completed successfully.
[7802378.ibm]$
```

---

If you receive an error that the management service is stopped. wait a few minutes for it to complete it's startup from the recent reboot.

4. This command will schedule a backup to run at 02:00 am every day. You can change the time to suit your own environment.
5. Exit the session.

## 11.8 Manual setup and configuration changes

If you are redeploying an existing configuration, or if you skipped any steps in Easy Setup, or you need to alter any values, then use the following procedures to manually set or alter these configuration settings.

If Easy Setup was completed and all values are entered correctly, then this section is for your reference and some of the topics can be skipped. We recommend you set up the optional support details to enable remote support functionality of the cluster, as detailed below in this section.

### 11.8.1 System names

If you need to change the cluster name, use the following CLI command.

```
chsystem -name <new_name>
```

### 11.8.2 System licenses

If the system licences need to be changed, then navigate to the screen as in Figure 11-33 on page 172. Select the settings icon, general. then select update license. Overtyping the values with the new ones to match the enclosure licenses being applied.

Note that while warning messages may be posted if the entered licenses are exceeded, IBM works on an honesty system for licensing. Only enter the value of the license you have purchased.

**Note:** To use the compression function, you must obtain the optional IBM Real-time Compression license.

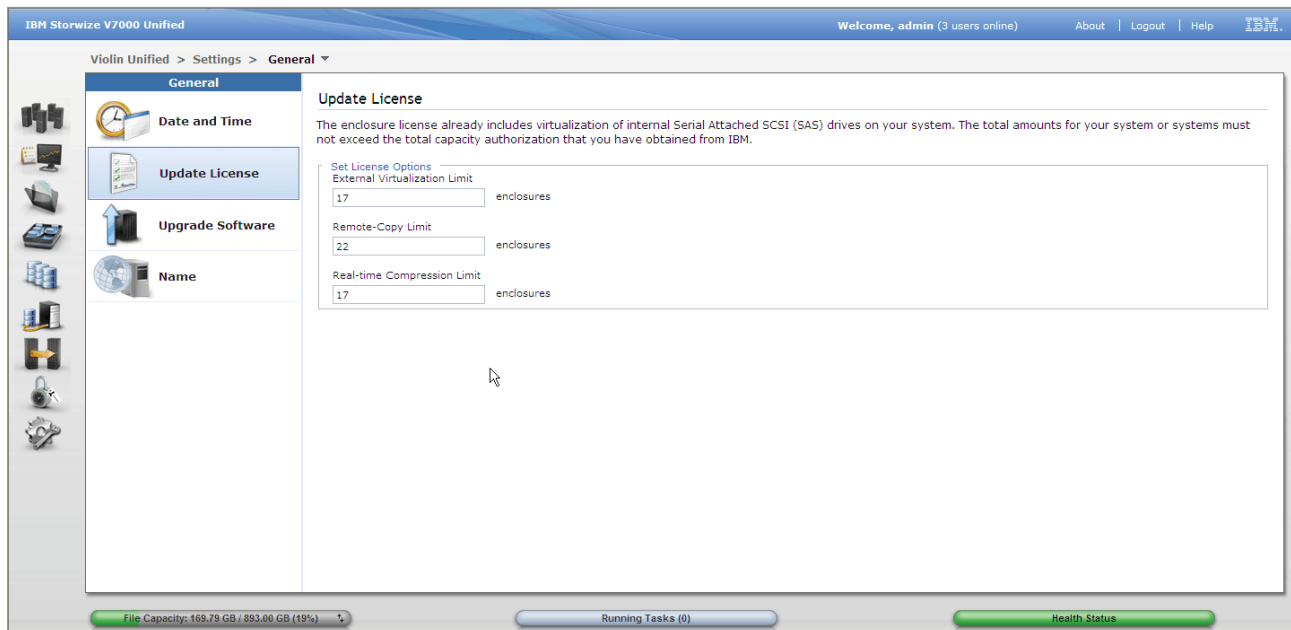


Figure 11-33 Update license

### 11.8.3 Support

Use the following panels to define or update details about remote support functions and also to gather support information. To define call home details, navigate to the screen shown in Figure 11-34 on page 172 by selecting Settings, Support, then click Call Home.

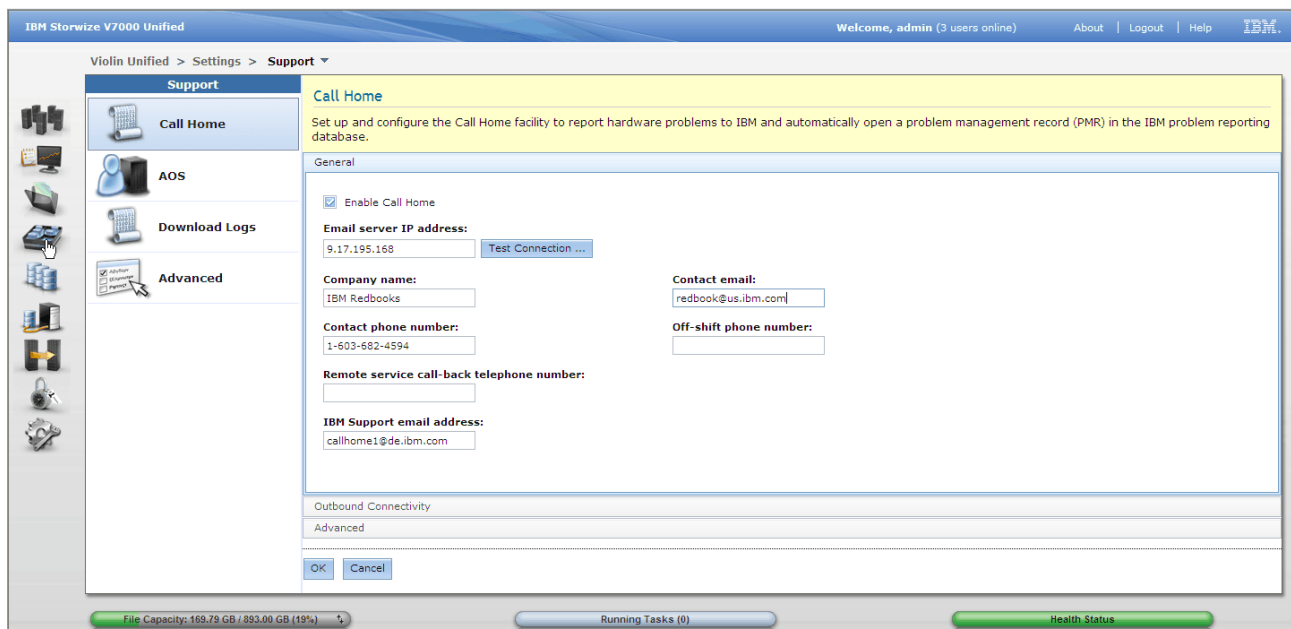


Figure 11-34 Call home definitions



AOS (Assistant On Site) is an IBM Tivoli based product used by IBM support to securely connect to the Storwize V7000 Unified to perform problem determination and engineering actions if required. This can be configured in two ways: lights on and lights out. Lights-on requires manual authorisation from the GUI onsite to allow a connection to complete. Lights-out will connect immediately. This feature requires access to the internet over HTTP and HTTPS protocols.

To configure, navigate to the setup screen as shown in Figure 11-35 on page 173, select the setting icon then support. Click AOS. To enable tick the box and then select lights on or lights out. If your site utilizes a proxy for internet access, enter the IP address and port details. Also a user password if needed. To disable proxy, blank the Proxy Server field.

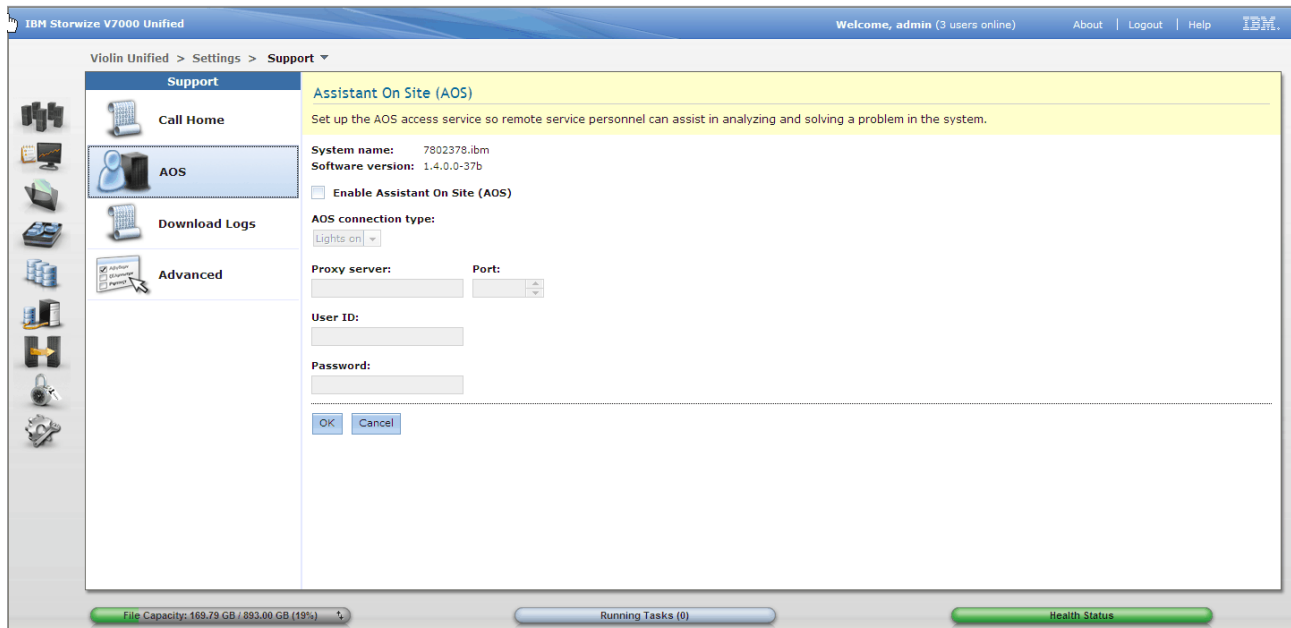


Figure 11-35 AOS configuration

For most problems, IBM support will request logs from the cluster. This is easily done by navigating to the Download Logs screen as seen in Figure 11-36 on page 174. Select Settings, then Support. Then click Download Logs. From this panel you can view the support files by clicking on show full log listing. This will display the contents of the logs directory and will include previous data collects, dumps and other support files. These can be downloaded or deleted from this window.

To create and download a current support package, click the download support package button. This will display a popup window asking for the type of package as shown in Figure 11-37. Select the log type as requested by IBM support. If any doubt take the full logs but also consult the Chapter 15, "Troubleshooting and Maintenance" on page 239 and IBM Support. Then click download. A progress window will appear while the data is created and gathered. When complete the window will close and a download option window will appear. Use this to download the file to the workstation you are browsing from. Note that the file can be retrieved anytime later or by any user with access from show full log listing window.

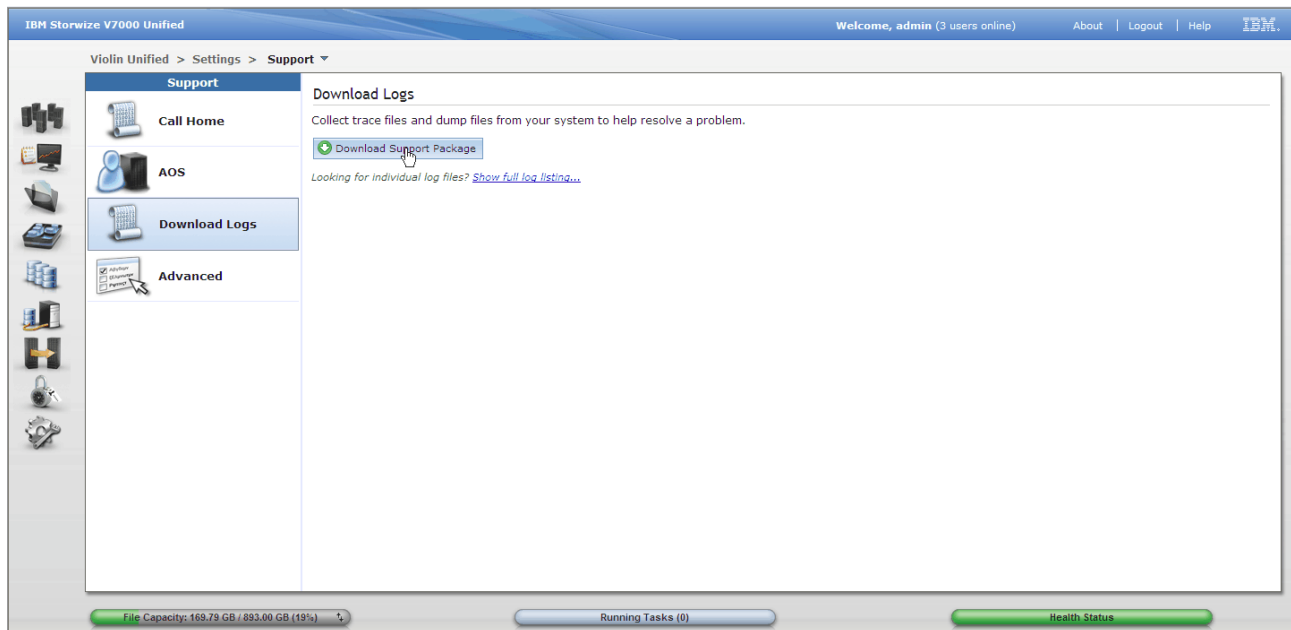


Figure 11-36 Support logs

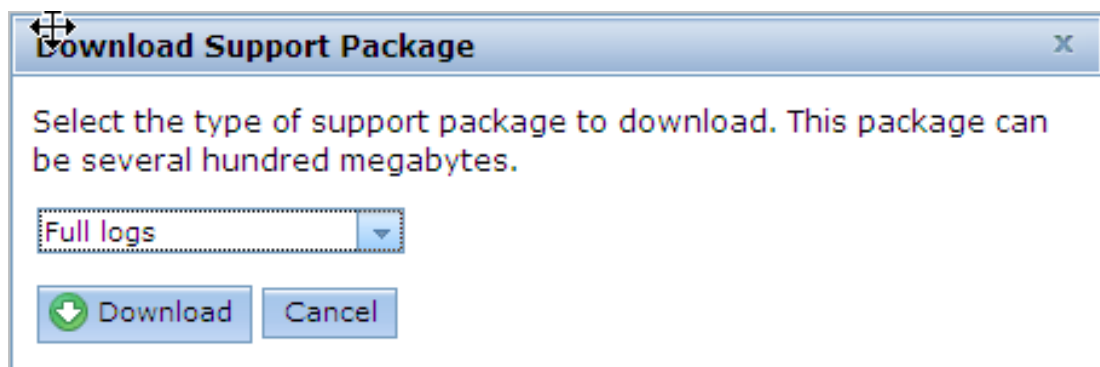


Figure 11-37 Download support package

## 11.9 Network

There are a number of network adapters in the Storwize V7000 Unified. Each is configured with IP addresses and these can be changed if required. Here we go through the various adapters and where to go to reconfigure them.

### 11.9.1 Public Networks

These addresses are on the 10Gb ethernet ports and also the 2 client side 1Gb ports. These are the addresses that the hosts and clients use to access the Storwize V7000 Unified and open file shares. Navigate to the Settings icon and then take the networks option. This will give you the screen as shown in Figure 11-38 on page 175. Click Public Networks.

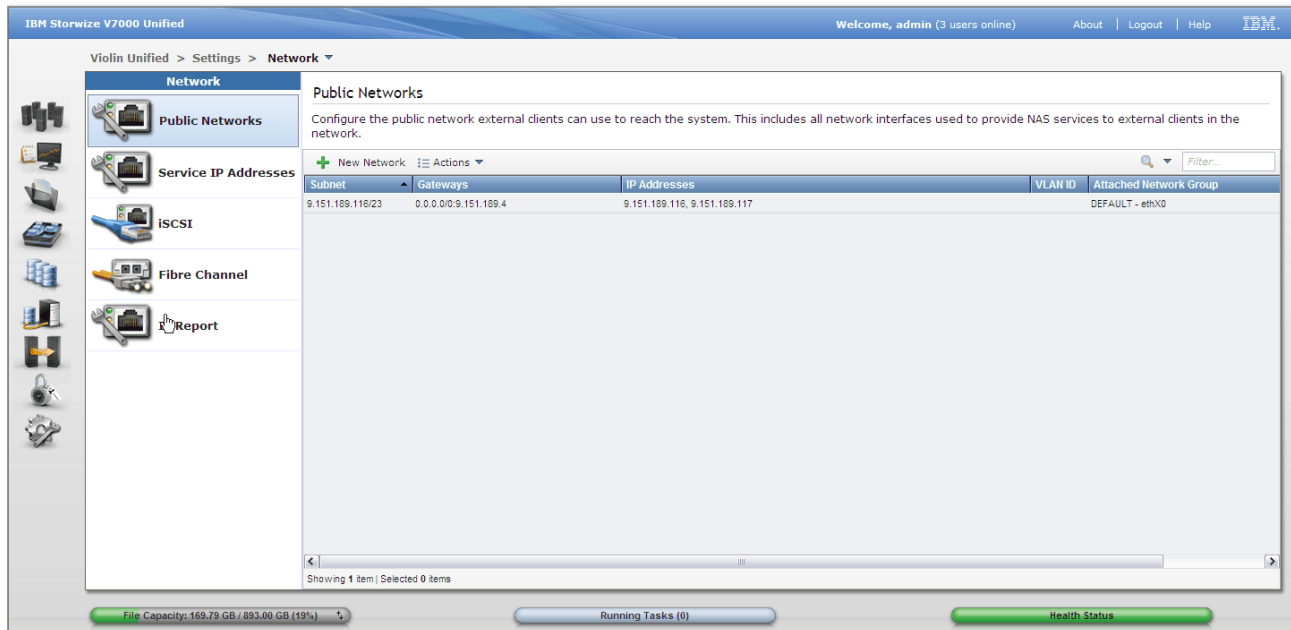


Figure 11-38 Public networks configuration

From here you can add new network definitions and delete existing ones. Each definition is defined to one of the virtual adapters. Refer to “Public networks” on page 169 for details on configuring public networks.

## 11.9.2 Service ports

It is important that the service ports on the control enclosure are set to an IP address. These are seldom used in normal operation but are important in times of a problem and it is better to have them set up beforehand. One IP address is needed for each node canister on it's port 1. This IP will share the port and coexist with the management IP that also is presented on the same port but only from the config node.

The set/alter these IP addresses, select the Settings icon, then Network. Click on the Service IP Addresses to display the configuration window as shown in Figure 11-39 on page 176. Hover the mouse over port 1 to display the settings window. If you need to change it, click in the fields and type the new values. Click OK to save. You must set both canisters, so use the pull down to display the other canisters ports and configuration.

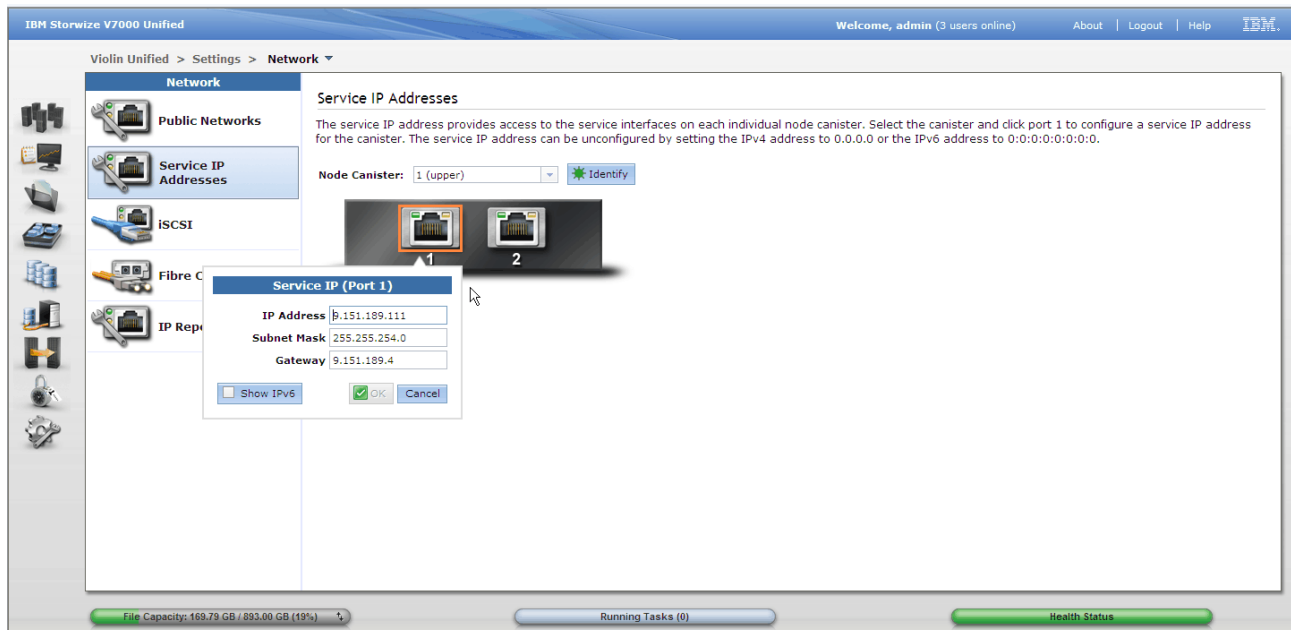


Figure 11-39 Service IP

### 11.9.3 iSCSI

If you have the feature to add the additional ethernet ports to the control enclosure for iSCSI service, then you need to set the addresses for these ports. Note this is an optional feature and only used for iSCSI connection to block storage. Navigate to the settings icon, then network. Select iSCSI from the list. You will get the windows in Figure 11-40 on page 176.

Here we see a graphical representation of the ports. Note that there are 2 diagrams, one for each node canister. Hover over each port to display the configuration panel and enter the IP address details. OK to save.

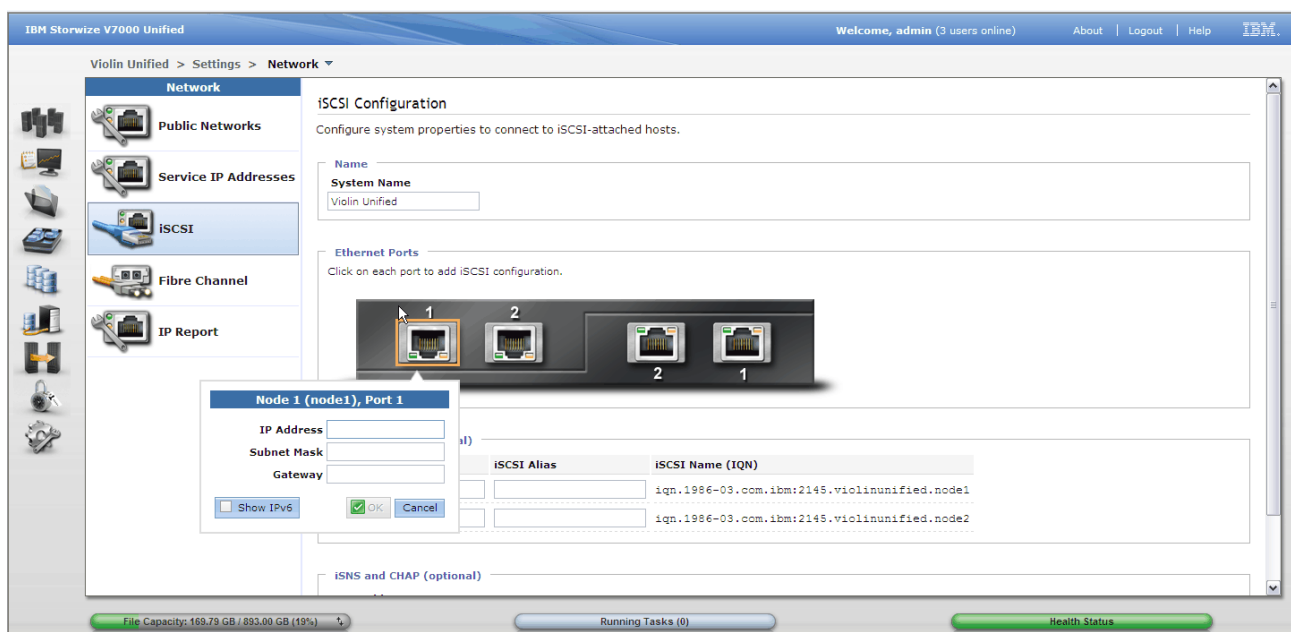


Figure 11-40 iSCSI IP addresses

## 11.9.4 Fibre Channel ports

Use the Fibre Channel panel to displays the Fibre Channel connectivity between nodes, storage systems, and hosts. The results for the ports for all nodes, storage systems, and hosts can be displayed. Navigate to the settings icon, then network. Select Fibre Channel from the list. The display in Figure 11-41 shows the results for all hosts attached to block volumes.

The screenshot shows the IBM Storwize V7000 Unified web interface. The left sidebar contains icons for various settings: Network, Public Networks, Service IP Addresses, iSCSI, Fibre Channel (selected), and IP Report. The main content area is titled 'Fibre Channel' and includes a description: 'Display the connectivity between nodes and other storage systems and hosts that are attached through the Fibre Channel network.' Below this, there are dropdown menus for 'View connectivity for:' (set to 'Hosts') and 'Charlie\_Harper', and a 'Show Results' button. A table displays the connectivity results for four hosts.

Name	System Name	Remote WWPN	Remote...	Local WWPN	Local Port	Local NP...	State	Node Na...	Type
Charlie_Harper		21000024FF35D977	BE1000	500507680240230D	4	BE1C00	Active	node2	Host
Charlie_Harper		21000024FF35DA8E	BE1100	500507680230230D	3	BE1B00	Inactive	node2	Host
Charlie_Harper		21000024FF35D978	BE1300	500507680230230D	3	BE1B00	Active	node2	Host
Charlie_Harper		21000024FF35DA8F	BE1200	500507680240230D	4	BE1C00	Inactive	node2	Host

Showing 4 items

At the bottom of the interface, there are three status bars: 'File Capacity: 169.79 GB / 893.00 GB (19%)', 'Running Tasks (0)', and 'Health Status'.

Figure 11-41 Shows fibre results for hosts

## 11.9.5 Fibre Channel ports

The IP Report panel displays all the IP addresses that are currently configured on the system. The File module table provides details on the all IP addresses that are currently configured to manage NAS services and public networks that clients use to connect to the system. File modules provide the services to access the file data from outside the system and provide the back-end storage and file system that store the file system data. The Control enclosure area show details on all the IP addresses related to managing block storage. The control enclosure provides the services to access the block storage for block clients. A sample display is shown in Figure 11-42 on page 178

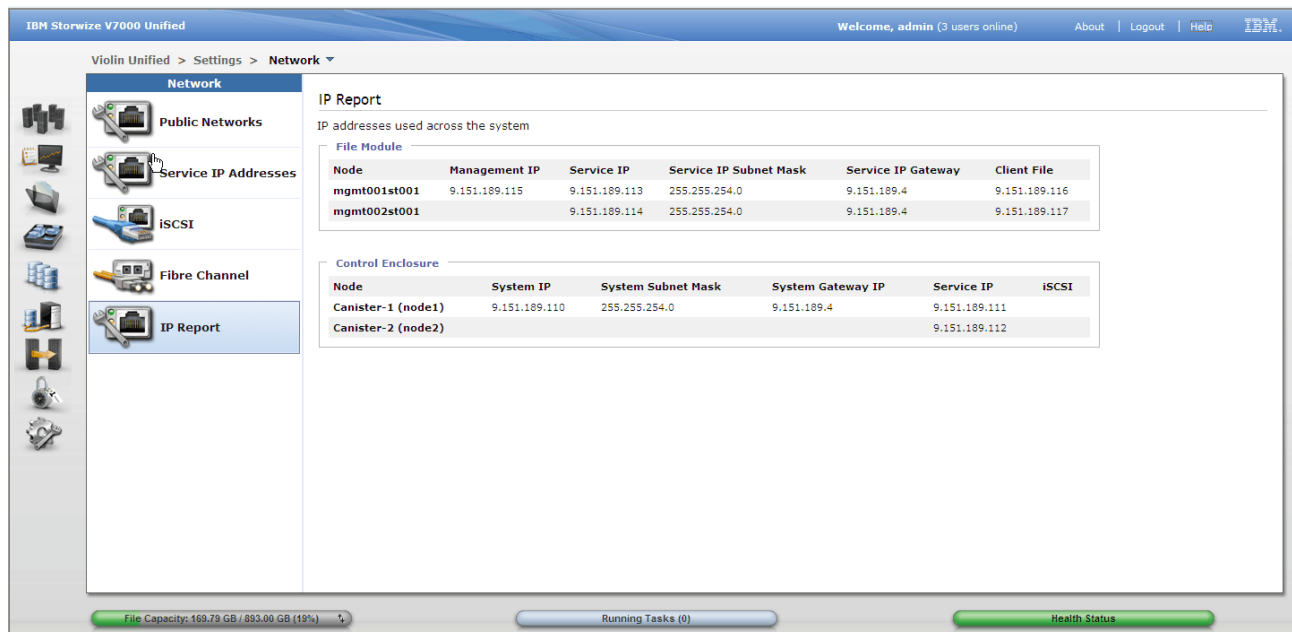


Figure 11-42 IP Report

## 11.10 Alerting

Storwize V7000 Unified supports several methods of alerting events and problems. The call home for IBM support has been discussed in 11.8.3, “Support” on page 172. Additionally alerts can be sent to an e-mail address and/or to an SNMP server. To configure alerting navigate to Settings, then Event notifications.

### 11.10.1 E-mail

For e-mail, first set up the details of your SMTP mail server on the screen as shown in Figure 11-43 on page 179.

Click enable e-mail notifications, then insert the IP address of your SMTP server. Also you must enter the reply address for the e-mail and a senders name. These will show on the e-mail header and identify to the recipient where the e-mail is from, so use a easily recognized name to make alerts clear.

Using the “...” button define as many subject options as needed, then select the desired subject content from the pull down. Fill in optional header and footer details to be included around the event text if desired.

A test e-mail can be sent at any time to confirm the settings and prove the alerting path. Just enter a valid e-mail target and click test e-mail. We recommend you send a test e-mail when complete.

IBM Storwize V7000 Unified

Welcome, admin | About | Logout | Help

SanJose1 > Settings > Event Notifications

**Event Notifications**

- Email Server
- Email Recipients
- SNMP Server
- Syslog Server

**Email Server**

Configure an email server that is used by your site. Ensure that the email server is valid.

☒ Enable email notifications

\* Email server (IP address):  \* Sender's email address:

\* Sender's name:  Subject:

Header:

Footer:

Test email address:

☐ Maximum emails sent per hour:

File Capacity: 1.7 GB / 1.5 TB (0%) | Running Tasks (0) | Health Status

Figure 11-43 E-mail server configuration

Next click on e-mail recipients to enter where the e-mails will be sent. A list of current recipients is displayed. You can select any line to edit or delete from the action drop down. To add a new entry click "new recipient". This will display the entry dialogue as seen in Figure 11-44.

**Event Recipient**

☒ **Enable recipient's email notification**

**\* Name:** Jorge Quintal

**\* Email address:** jquintal@us.ibm.com

**Status change:**

☒ Events Critical  
☐ Report Info, Warning, Critical  
Warning, Critical

**Utilization thresholds:**

☐ Events Info, Warning, Critical  
☐ Report

**GUI:**

☐ Events Info, Warning, Critical  
☐ Report

**Backup:**

☐ Events Info  
☐ Report

**Storage:**

☐ Events Info, Warning, Critical  
☐ Inventory

**Quotas:**

☐ Report 100

OK Cancel

Figure 11-44 Event recipient

Enter a name for the recipient and the e-mail address. Tick the events box for each type of event type you wish this recipient to receive an alert for. Several events have a criticality level, select the desired threshold. Note, critical only is best for most users to reduce the number of e-mails received.

Click the reports box if reports are also required. For quota, choose the percentage of the hard limit for the threshold.

## 11.10.2 SNMP

If you use an SNMP server to monitor your environment and wish to receive alerts from the Storwize V7000 Unified, then navigate to settings, then event notifications and click on SNMP. Complete the form as shown in Figure 11-45.



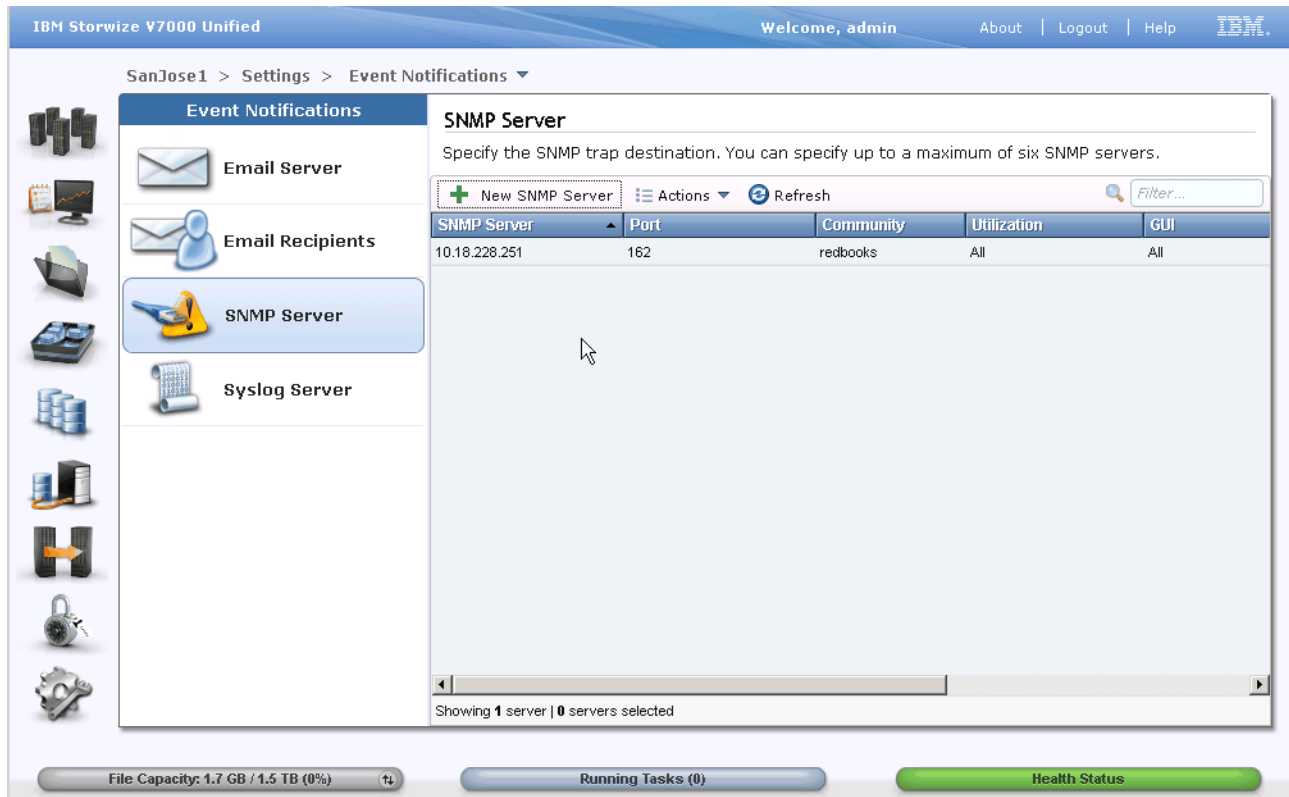


Figure 11-45 SNMP server settings

Use the New SNMP Server button to add a server or highlight the line and use actions to edit or delete. The add/edit popup window is shown in Figure 11-46.

Figure 11-46 Edit SNMP server

Enter the server IP address and port. Fill in the SNMP community name. Tick the event type that are to be sent and for each select the severity.

### 11.10.3 Syslog Server

If you wish to off-load cluster messages to a logging server, then enter the details in the panel as shown in Figure 11-47. Navigate to the Settings icon then Event Notifications. Select Syslog Server.

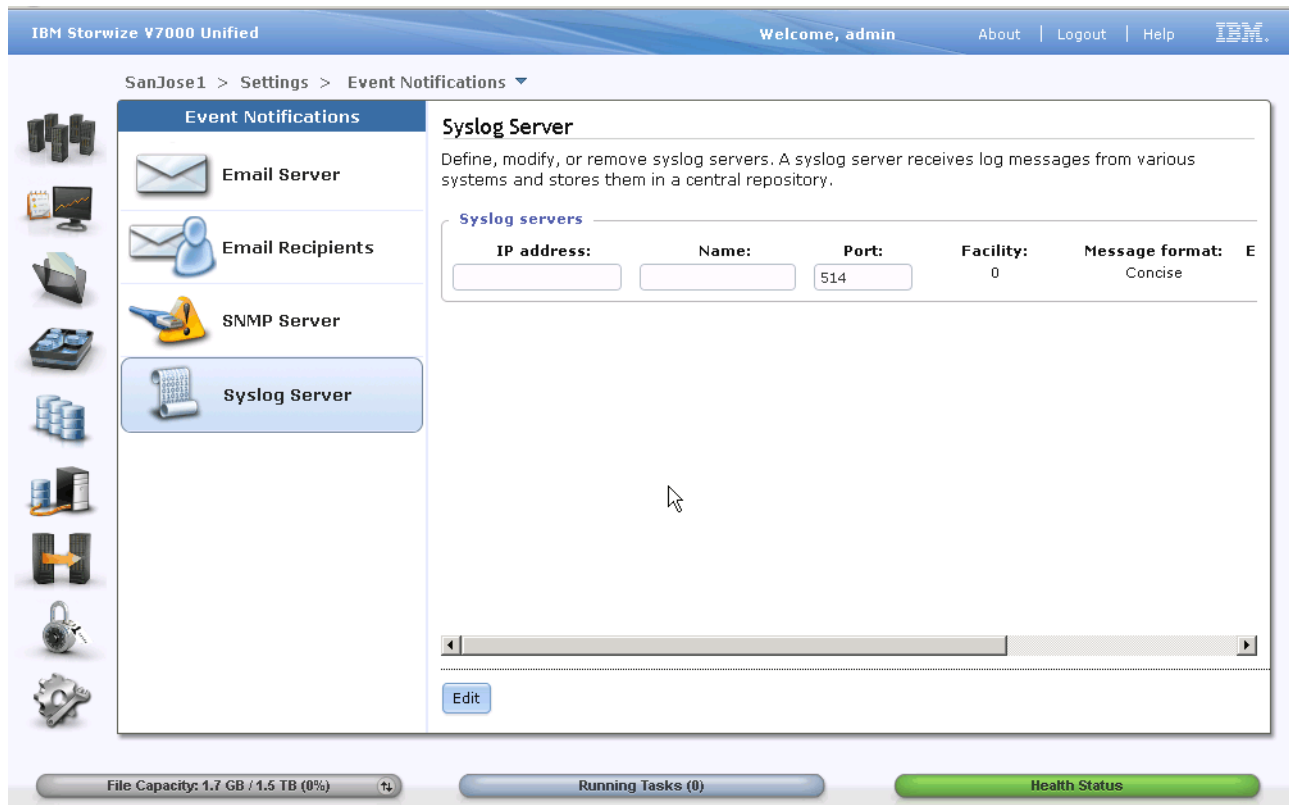


Figure 11-47 Syslog settings

## 11.11 Directory Services and Authentication

Navigate to Settings → Directory Services. First set up the DNS server settings by clicking on the DNS icon in the left panel.

### 11.11.1 DNS

In the DNS settings panel as shown in Figure 11-48 on page 183, click on the Edit button to make changes. First make sure the DNS domain name is correct. This is very important with Active Directory and is a key component with authentication. Define your DNS server and any backup DNS servers using the “+” button to create more entries. Next add any search domains that are outside the primary domain name, that will be involved in accessing this cluster.

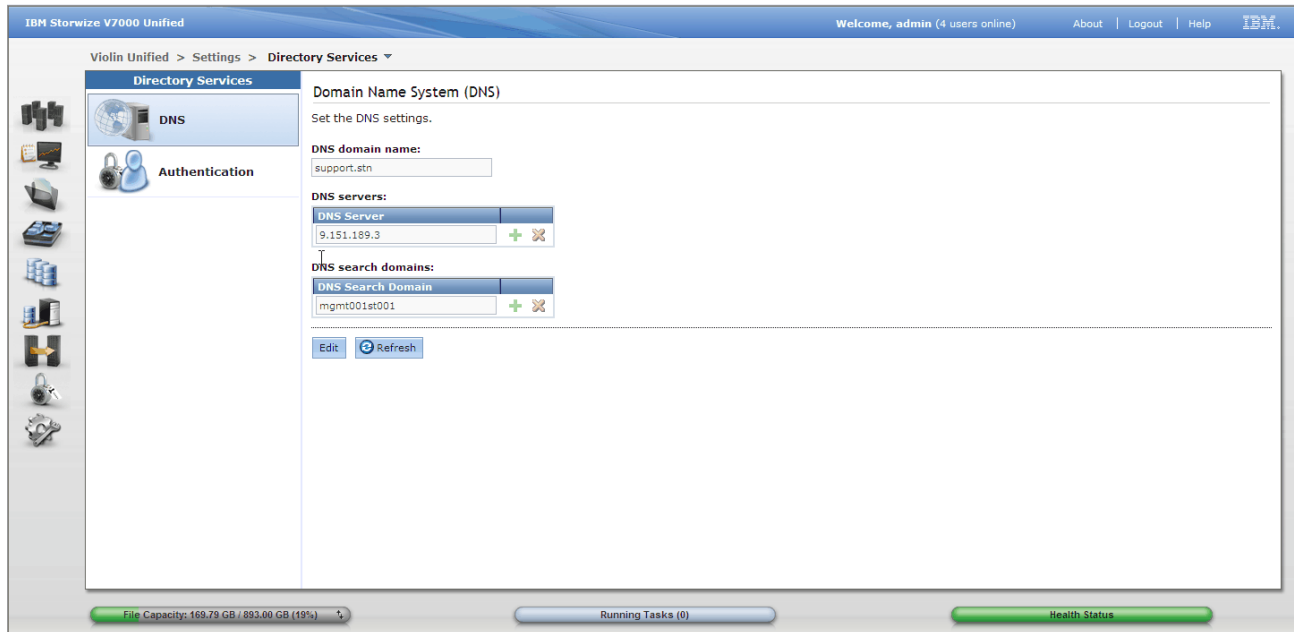


Figure 11-48 DNS Settings

### 11.11.2 Authentication

Now click on the Authentication icon as seen in Figure 11-49 on page 183. The system requires that you determine a method of authentication for users of the system. The system supports either a remote authentication service or a local authentication. Remote authentication is provided by an external server that is dedicated to authenticate users on the system. Before configuring remote authentication on your system, ensure that the remote authentication service is set up correctly. Local authentication is provided by an internal mechanism in the system which is dedicated to authentication.

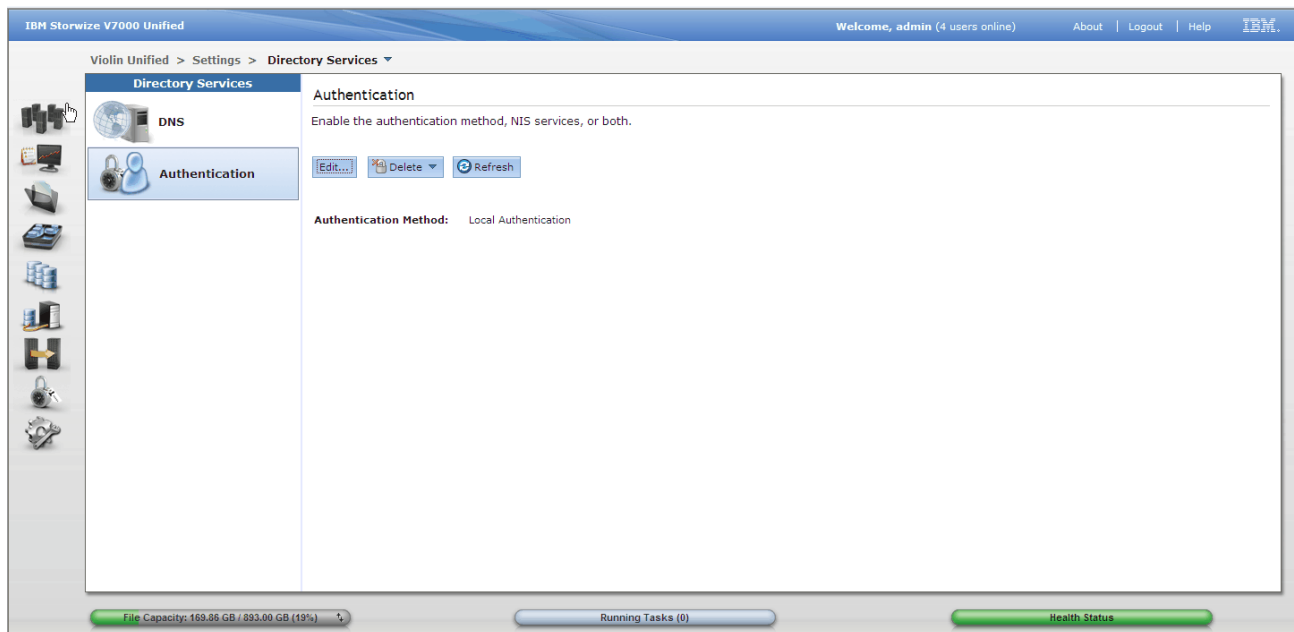


Figure 11-49 Authentication Settings

**Warning:** Use extreme care editing the authentication settings on a running cluster as this may cause loss of access to shares.

If you need to make any changes, then click edit which will launch the authentication screens as seen in Figure 11-50. You will be asked to acknowledge your desire to edit these settings.

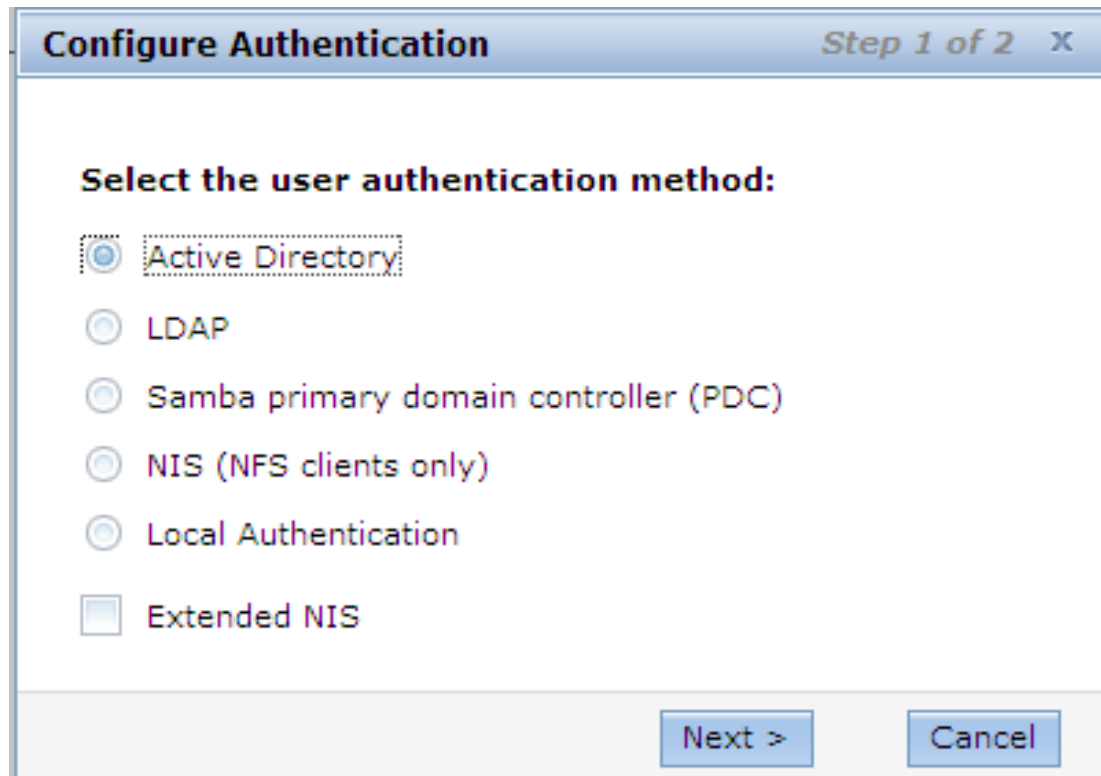


Figure 11-50 Presented Edit screen

The cluster GUI will now take you through the setup panels for the authentication setup. These are the same panels as the EZ-Setup wizard guided you through and are described in detail in “Authentication” on page 160 earlier in this chapter. If any configuration has already been set then the panel entry fields will auto fill with the existing values.

**Note:** With Local Authentication you enter user and group information locally which is covered in Section 11.13.3, “Create local users using Local Authentication for NAS access” on page 189

### 11.11.3 NTP server

It is essential that an NTP is defined in the cluster to ensure consistent clocks in the file modules. This is needed for recovery and resolving deadlocks.

Navigate to General Settings → Date and Time, and click on the NTP icon as shown in Figure 11-51 on page 185 and enter the server details.

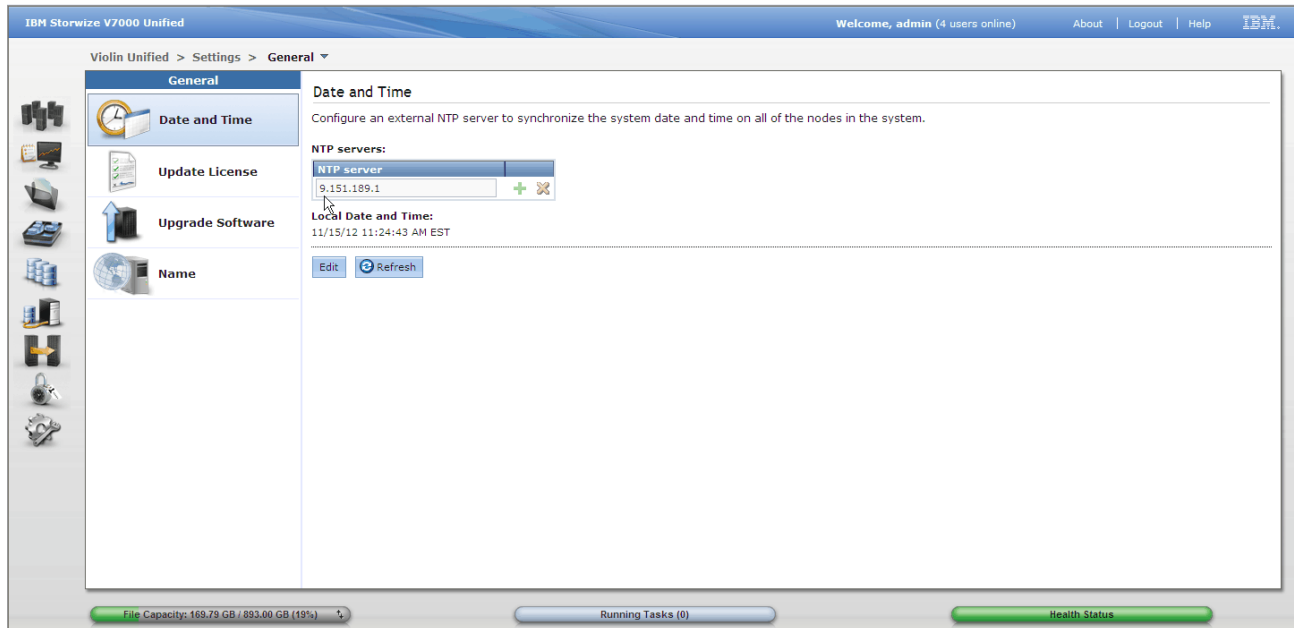


Figure 11-51 NTP Server

Or use the CLI

```
setnwnntp ip,xxx.xxx.xxx.xxx,xxx.xxx.xxx.xxx <= separate addresses with comma.
lsnwnntp
rmnwnntp
```

## 11.12 Health check

Check the health of the system and correct any problems before continuing by checking the bottom of the screen to see the status.

First confirm that the color of the status indicator at the bottom of the GUI screen, this is a good indicator of any serious issues. If not green, click on the X to get a summary of the areas unhealthy.

To get more detail, use the CLI command `lshealth` will give a listing of each function in the cluster and its current status. Use the `-r` parameter to force the task to refresh the information. You can also drill down into a particular function using the `-i` parameter.

Next review the event logs for the file and block storage, and review the logs for each individual file module in system details. Ensure there are no unfixed events.

**Important:** Make sure that storage pools are maintained in a green state specially if compression is used. If the storage pool is allowed to run out of physical space then it will cause compressed volumes to go offline.

## 11.13 User Security

It is important to change the default passwords and define profiles for administrators of the cluster. Change the passwords for the following userids:

admin                      This is the default userid for the Storwize V7000 Unified cluster

superuser                This is the default userid for the Storwize V7000 storage enclosure.

You can also define additional users as required.

At the time of writing the root password for the file modules is widely known to IBM support and implementation teams to assist in setup and recovery. In the future, when all needed functions have been made available through the GUI and CLI, this password will be changed. If this has occurred, then this comment can be ignored. As the password is still the widely known default, ask your IBM team to assist you to change it. The CLI command **chrootpwd** will change it across both nodes.

**Warning:** It is very important to change the default passwords on both the cluster and the Storwize V7000 storage enclosure.

### 11.13.1 Change passwords

In the following sections we show how to change the passwords.

#### Change storage enclosure password

You will need to logon to the Storwize V7000 storage GUI directly. Navigate to Access → Users. The All Users view will be displayed by default. Highlight the “superuser” user definition and use the actions pull down or right click to select properties. This will launch the user properties screen as shown in Figure 11-52 on page 187. Click the change button in the User’s password section and enter a new password. Click OK to action and complete.

**Edit User**



**\* User's name:**

	User Group	Role
<input type="checkbox"/>	Administrator	Administrator
<input type="checkbox"/>	CopyOperator	Backup Administrator
<input type="checkbox"/>	ExportAdmin	Export Administrator
<input type="checkbox"/>	Monitor	Operator
<input checked="" type="checkbox"/>	SecurityAdmin	Security Administrator
<input type="checkbox"/>	SnapAdmin	Snapshot Administrator

**\* User's password:**

Configured

Change

OK

Cancel

Figure 11-52 Edit user - block

--or--

```
svctask chuser -password <xxxxxxx> superuser
```

### Change cluster admin password

On the Storwize V7000 Unified GUI, navigate to Access → Users. The All Users view will be displayed by default. Highlight the “admin” user definition and use the actions pull down or right click to select edit. This will launch the edit screen as shown in Figure 11-53. Click the change button in the User’s password section and enter a new password. Click OK to action and complete.



**Edit User**

 \* User's name:  
admin

	User Group	Role
<input type="checkbox"/>	Administrator	Administrator
<input type="checkbox"/>	CopyOperator	Backup Administrator
<input type="checkbox"/>	ExportAdmin	Export Administrator
<input type="checkbox"/>	Monitor	Operator
<input checked="" type="checkbox"/>	SecurityAdmin	Security Administrator
<input type="checkbox"/>	SnapAdmin	Snapshot Administrator

\* User's password:  
Configured

Figure 11-53 Edit user - Unified

Alternatively you can use the following command:

```
chuser admin -p <xxxxxxxx>
```

### 11.13.2 Create cluster users

Navigate to Access → Users and click on the New User button as seen in Figure 11-54. This will launch the New User window as shown in Figure 11-55. Type in the users name and select the level of authority from the list. Set a password and type it again to confirm. Click OK to create.

With Local Authentication you enter user and group information using the method described in Section 11.13.3, “Create local users using Local Authentication for NAS access” on page 189

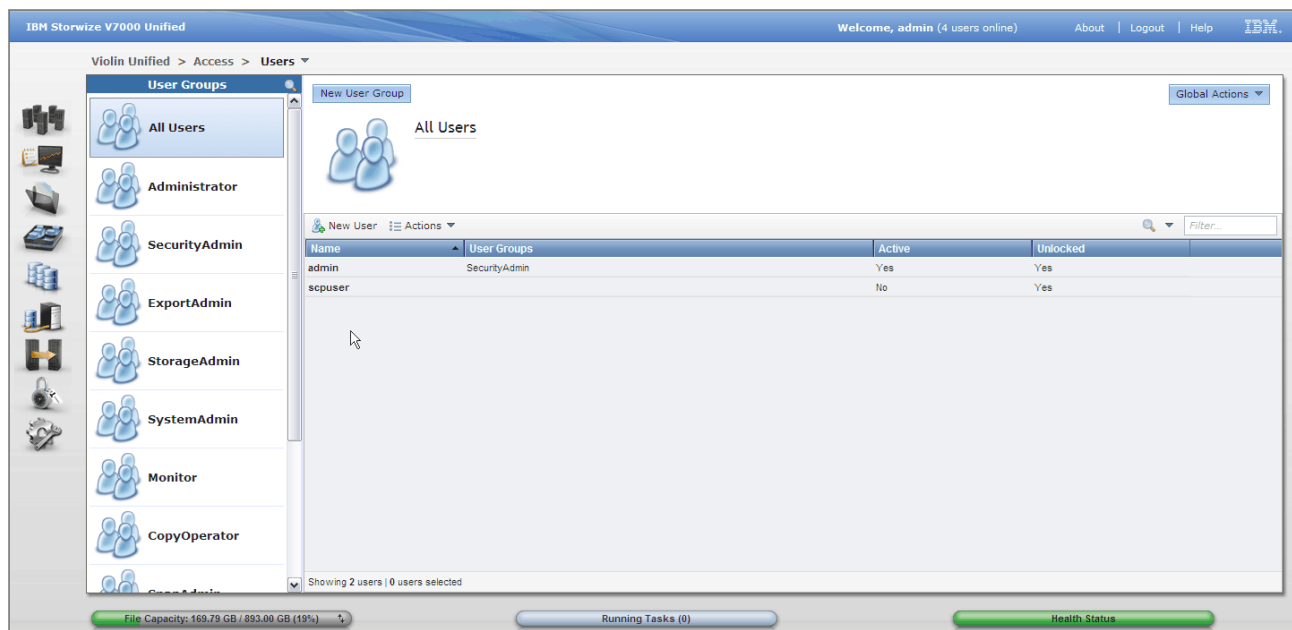


Figure 11-54 Cluster users



**New User**

\* User's name: Trevor

User Group	Role
<input checked="" type="checkbox"/> Administrator	Administrator
<input type="checkbox"/> CopyOperator	Backup Administrator
<input type="checkbox"/> ExportAdmin	Export Administrator
<input type="checkbox"/> Monitor	Operator
<input type="checkbox"/> SecurityAdmin	Security Administrator
<input type="checkbox"/> SnapAdmin	Snapshot Administrator

\* User's password: ..... \* Confirm user's password: .....

OK Cancel

Figure 11-55 New User

Alternatively you can use the following command:

```
mkuser Trevor -p passw0rd -g Administrator
```

### 11.13.3 Create local users using Local Authentication for NAS access

Besides external authentication such as Active Directory or LDAP, the system supports user authentication and ID mapping using local authentication server for NAS data access. After you have configured local authentication on your system, you need to define groups and users who are registered in the local authentication service to access data using the NAS protocols that are supported by this system.

Using local authentication eliminates the need for a remote authentication service, such as Active Directory or Samba Primary Domain Controller (PDC), thus simplifying authentication configuration and management. Local authentication is best used for environments where no external authentication service is present or if the number of users are relatively small. Local authentication supports up to 1000 users and 100 user groups. For configurations where all users are in a single group, the system supports 1000 users. A single user can belong to 16 groups. For larger numbers of users and group, remote authentication should be used to minimize performance impacts to the system.

When creating users and groups for local authentication, ensure user, groups names, and IDs are consistent across multiple systems in your environment. If NAS users accesses data from two or more systems, ensure that those users have the same user name, user ID, and primary group on each system. Consistent user and group attributes are required for using advanced functions such as IBM Advanced Cloud Engine and asynchronous replication. In addition it also provides flexibility in moving data between systems in your environment and simplifies migration to an external Lightweight Directory Access Protocol (LDAP) server.

When managing multiple systems, administrator should designate a primary system that contains all users and groups. Any new user and group should be created first on the primary system and then created on any other systems, using the same user ID or group ID that was defined on the primary system. This practice helps ensure that a user ID or group ID is not overloaded. When specifying user IDs and group IDs, you can have the system automatically generate these values; however, to ensure control over these values and to minimize any authorization problems that might be introduced overtime, assign these values manually.

When creating user and groups for local authentication, the user and group names are case insensitive. For example, if a user exists on the system, named "John", a new user named "john" cannot be created. This is a limitation of some of the supported protocols. In addition, NAS user and group names cannot be the same as CLI users and system users.

Navigate to Access → Local Authentication as shown in Figure 11-56. New local users and new local groups can be create using this process. You can modify a current user using the Action pull-down button or by right-clicking on a user or group.

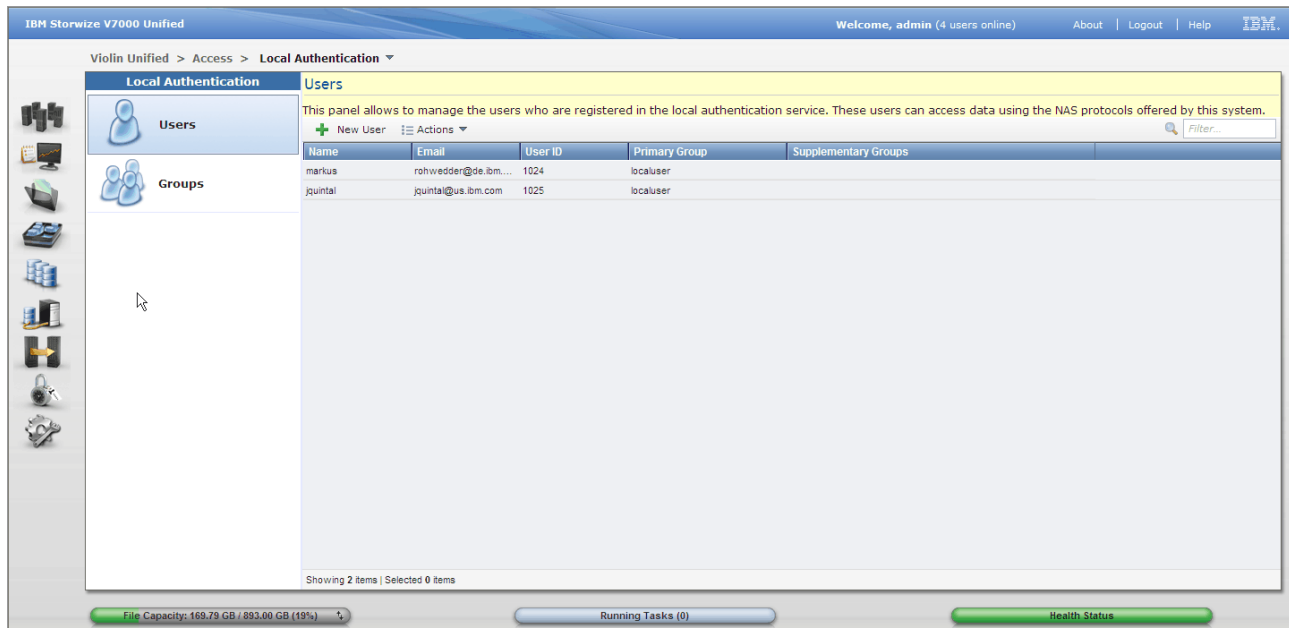


Figure 11-56 User using Local Authentication

The popup in Figure 11-57 on page 190 shows how to add a New User for NAS access.

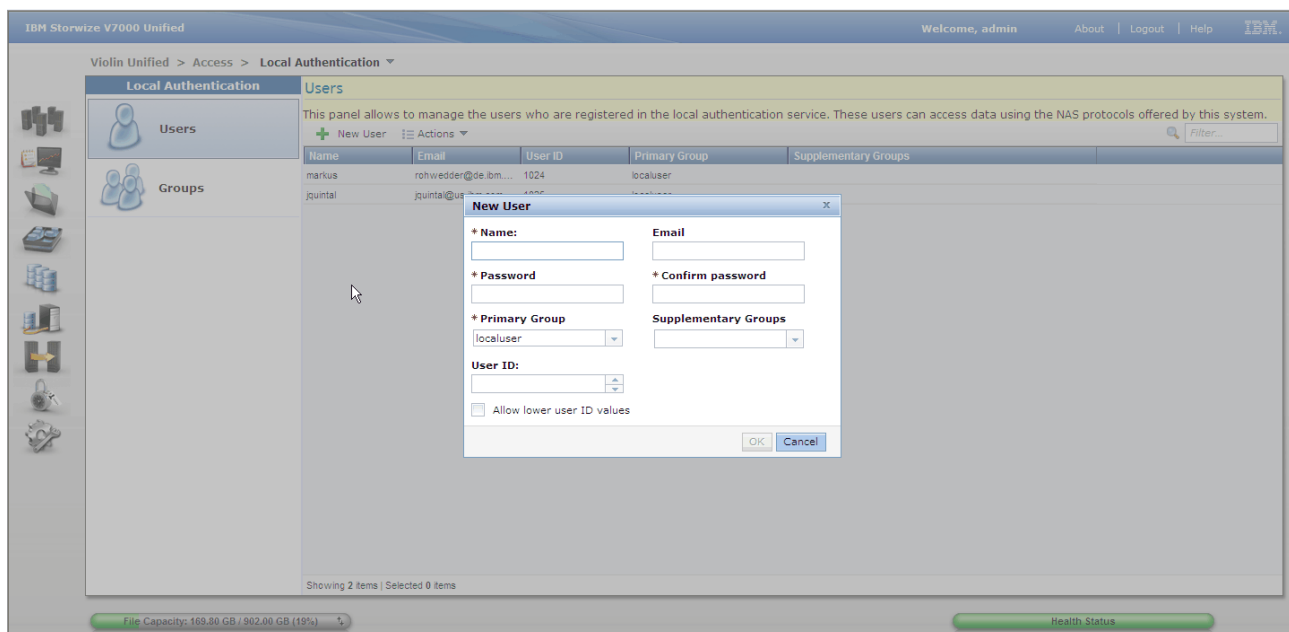


Figure 11-57 Add New User

The popup in Figure 11-58 shows how to add a New Group for NAS access.

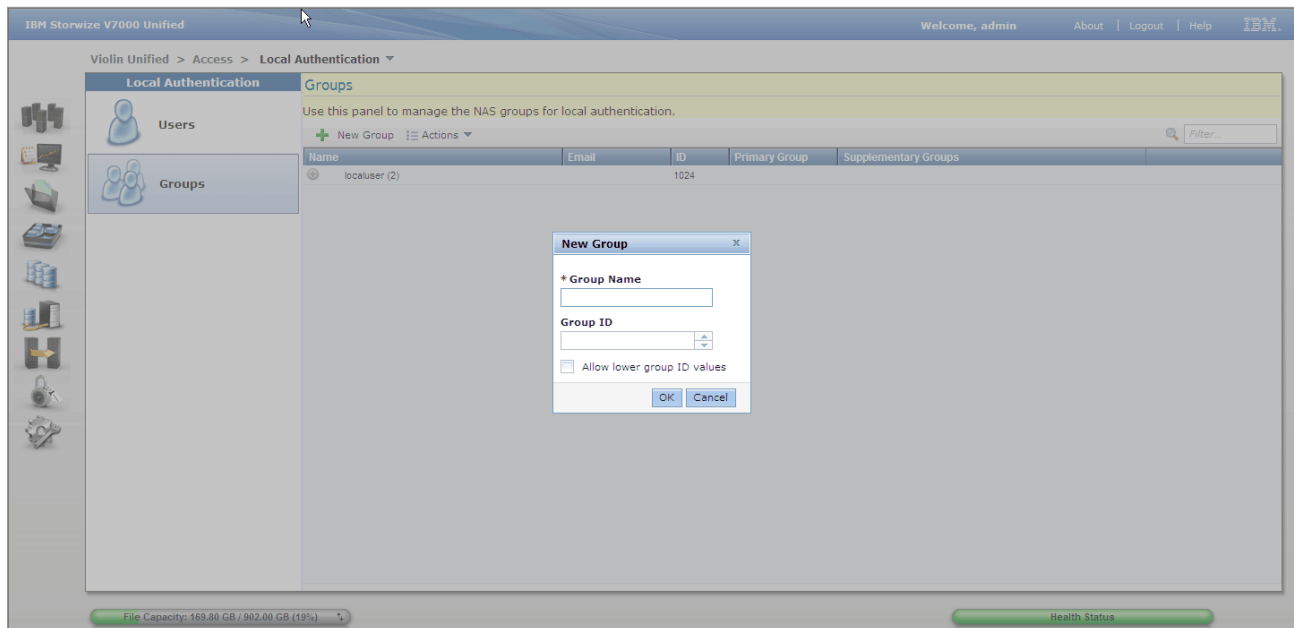


Figure 11-58 Add New Group

## 11.14 Storage controller configuration

The Storwize V7000 storage component provides the storage arrays for the file systems. It is necessary to configure the controllers even if not using the Storwize V7000 for block devices. The basic configuration and initialization should have been completed using the USB key in the implementation steps above. If the storage was auto configured during the Easy Setup process and no block volumes are required, then the configuration is complete and you can skip this step, otherwise continue.

Reference Information Center:

[http://pic.dhe.ibm.com/infocenter/storwize/unified\\_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd\\_installhdw\\_2343rc.html](http://pic.dhe.ibm.com/infocenter/storwize/unified_ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.unified.doc%2Ftbrd_installhdw_2343rc.html)

### 11.14.1 External SAN requirements

Block volumes are accessed over the Fibre Channel Storage Area Network (SAN). The Storwize V7000 is fitted with 4 Fibre Channel ports on each node. When part of the unified cluster, two from each node are dedicated to the file module connections. The remaining two are for connection to the SAN. As is normal practice, two SAN fabrics are needed in parallel to form redundancy and manageability. Connect one port from each node to each fabric.

Zone these ports to each other in each fabric to create additional inter-node links.

Host HBAs on the SAN can now be zoned to the Storwize V7000 ports as required for access to the storage.

## 11.14.2 Configure storage

Presenting block volumes to hosts is the same for the Storwize V7000 Unified as it is for the standard Storwize V7000. The only difference is that the Storwize V7000 Unified GUI is used for both file and block, but the interface and functions are still the same.

For this reason, we recommend using the *Implementing the IBM Storwize V7000 V6.3*, SG24-7938 for guidance on configuration and management of the block storage. We will give a summary of the steps involved below to assist experienced users.

## 11.15 Block configuration

Summarized below are the steps to configure block storage volumes for access by hosts over the SAN.

External storage	If external SAN attached storage is being managed by the Storwize V7000, this must be added first. The storage once installed and connected to the SAN, needs to be configured to present volumes (LUNs) to the cluster. The storage should map the volume as if the Storwize V7000 cluster were a host. The cluster will then discover these as MDisks.
Arrays	The physical drives fitted to the Storwize V7000 must be built into raid arrays. These arrays become MDisks and will be visible to the cluster once built.
MDisks	The array MDisks and the external MDisks can now be defined. They can be given descriptive names and are ready for use.
Pools	Extent pools, which are used to build the volumes, must now be defined. At least one pool is required. There are a variety of reason why multiple pools might be used, the main one is to separate disks into groups of the same performance. Previously known as MDiskgroups. Another reason could be to separate uncompressed from compressed volumes.
Volumes	From the pools, the volumes that the hosts will be using can be defined. A volume must get all its extents from one pool. The volume can be generic, thin-provisioned, mirror, or compressed.
Hosts	Each host that will be accessing the cluster needs to be defined. This requires describing the host and giving it a name, defining the access protocols it will use based on it's operating system type and defining its FC ports. The port definition includes the ports WWN.
Mapping	The last step is to create mappings for the volumes to the hosts. A volume can be mapped to multiple hosts, provided the hosts support disk sharing and are aware the disks is shared, and hosts can see multiple volumes.

### 11.15.1 Copy Services

Several methods of copying volumes are available to suit your business requirements. These are covered in detail in Chapter 8, "Copy services overview" on page 75.

As copy services are covered in depth in several manuals and redbooks, we will summarize the setup steps here and recommend you consult those books for concise detail. For the

Storwize V7000 we recommend you refer to *Implementing the IBM Storwize V7000 V6.3*, SG24-7938.

## Partnership

To work with block remote copy services to another cluster, you must have defined a relationship with at least one other cluster. Partnerships can be used to create a disaster recovery environment or to migrate data between clusters that are in different locations. Partnerships define an association between a local clustered system and a remote system.

**SAN connection** The cluster must be visible to each other over the SAN network. Create zoning so that at least one port from each node can see at least one port on all the remote nodes. Ensure if dual fabric, that this is true on both fabrics for redundancy. It is not necessary, nor recommended to zone all ports.

**Partnership (local)** After the zoning is in place, navigate to Copy Services → Partnerships to display the partnership list. Click *New Partnership*. This will begin a discovery process and provided there is at least one working path to a remote cluster, it will be considered a candidate. Select the partner cluster from the selection list.

You must also set the bandwidth setting. This is tunable at any time, but it is very important to set it correctly. The bandwidth defines to the cluster the maximum aggregate speed that is available to the other cluster. This is the speed data will be sent at. Do not over rate this setting or the link will be flooded. Do not under rate it or throughput will be affected as this caps the data rate. Note that this setting is for data from this cluster only, not received, which must be set at the remote cluster.

The cluster will show as partially configured. It must now also be defined from the remote site to be complete.

**Partnership (remote)** Go to the remote cluster and perform the same operation as above. The partnership must be set up from both clusters to be operational.

Note that the bandwidth setting on the remote cluster is an independent setting and controls the data flow rate from the remote back to the local. Normally these will be the same, but they can be set differently if desired.

When established, the partnership is now in a state of fully configured.

## Remote Copy

Navigate to Copy Services → Remote Copy to display the list of remote copy relationships. To create a new relationship, click on the New Relationship button. The Metro Mirror and Global Mirror Copy Services features enable you to set up a relationship between two volumes, so that updates that are made by an application to one volume are mirrored on the other volume. The volumes can be in the same system or on two different systems.

Now select the type of mirror relationship from Metro, Global and Global with Change Volumes.

On the next panel you have the choice of where the auxiliary volume is located. It is possible to set it on the same cluster, which you might find useful for some applications.

The next panel lets you choose the volumes and will list all eligible volumes on the specified cluster. Use the pull downs to select the volume you want in this relationship.

You are then asked if the relationship is already synchronized. Normally this is no, but there are situations during recovery and build when a relationship is being re-established between volumes and the data has not changed on either.

The last question is to begin starting the copy. This is the initial synchronisation process that will run in the background until complete.

## FlashCopy

As FlashCopy is only available on the same cluster, there is no requirement for a relationship. To create a copy, navigate to Copy Services → FlashCopy Mappings. Click New FlashCopy Mapping. This will launch the window to create a new mapping.

Eligible volumes are listed, select the source and target volumes using the pull down.

On the next screen you need to select the preset type for the copy.

Snapshot	This gives a point in time copy of the volume, thin provisioned. No background copy. Intended for temporary use to freeze the volume.
Clone	A one time use, full copy of the volume.
Backup	A full, point in time copy of the volume that can be repeatedly refreshed.

Using the advanced settings tab, you can adjust the background copy rate, set the mapping to being incremental (only new changes are written), delete after complete and the cleaning rate.

Then you are asked if you want to add the mapping to a consistency group.

## Consistency Groups

There are two main reasons for using consistency groups, by grouping a number of copies together, management actions such as start and stop can be applied to the group, reducing workload. But more importantly, the cluster will ensure that all members perform the action at the same point in time. This means that for a stop, the IOs completed to the remote volumes or copy targets are stopped in sync across all members of the group, which is important for system and application recovery.

To create a FlashCopy group, navigate to Copy Services → Consistency Groups. Select the New Consistency Group button to launch the create window. Give the group a name and that's it. To add a mapping to this group, either chose the group while creating the mapping or highlight the mapping in the listing and use the actions to move it to a consistency group.

To create a Remote Copy group, navigate to Copy Services → Remote Copy. Click on the New Consistency Group button. Give the group a name and define if the relationships are local to local or remote. You are able to create new relationships to be members now, or add then later. To add a relationship, simply highlight the entry and use the actions to add to a consistency group.

## FlashCopy Mappings

A FlashCopy mapping defines the relationship between a source volume and a target volume.

The FlashCopy feature makes an instant copy of a volume at the time that it is started. To create an instant copy of a volume, you must first create a mapping between the source

volume (the disk that is copied) and the target volume (the disk that receives the copy). The source and target volumes must be of equal size.

A mapping can be created between any two volumes in a system. The volumes do not have to be in the same I/O group or storage pool. When a FlashCopy operation starts, a checkpoint is made of the source volume. No data is actually copied at the time a start operation occurs. Instead, the checkpoint creates a bitmap that indicates that no part of the source volume has been copied. Each bit in the bitmap represents one region of the source volume. Each region is called a grain.

After a FlashCopy operation starts, read operations to the source volume continue to occur. If new data is written to the source or target volume, the existing data on the source is copied to the target volume before the new data is written to the source or target volume. The bitmap is updated to mark that the grain of the source volume has been copied so that later write operations to the same grain do not recopy the data.

## File Copy Services

Use this function to select different methods to replicate data to and from different file systems. The system supports two types of file system replication: Replicate file system and remote caching.

File system replication provides asynchronous replication of all file system data on one system to another file system located remotely over an IP network. The two systems should be separated geographically to provide data recovery and high availability. Asynchronous replication allows one or more file systems in a Storwize V7000 Unified file name space to be defined for replication to another Storwize V7000 Unified system over the customer network infrastructure. Files that have been created, modified or deleted at the primary location are carried forward to the remote system at each invocation of the asynchronous replication.

Asynchronous replication is configured in a single direction one-to-one relationship, such that one site is considered the source of the data, and the other is the target. The replica of the file system at the target remote location is intended to be used in read-only mode until a system or network failure or other source file system downtime occurs. During a file system failure recovery operation, failback is accomplished by defining the replication relationship from the original target back to the original source.

Remote caching provides transparent data distribution among data centers and multiple remote locations over a wide area network (WAN). Remote caching provides local access to centrally stored files and allows users in remote locations to work with files without creating inconsistencies. Data created, maintained, updated, and changed on the home system can be viewed and used on a cache system located anywhere in the WAN.

Replication	Replicating a file system creates copies of a file system between two systems, separated by some geographical distance to provide disaster recovery and business continuity. This is called asynchronous replication. Asynchronous replication is normally used between source and target systems where distance might affect response time because of bandwidth shortages.
Remote caching	Remote caching fetches files on demand from the home system file set to the remote cluster cache in real time during normal operations. Users can see only the files they have permission to see and access the files as if they are present locally. Data is revalidated when accessed to ensure the most recent file versions are available. Data can be prefetched by setting policies or files can be pulled into cache on demand.



## 11.16 File Services configuration

To create a share (or export), we need to have a number of building blocks in place. We will describe these here and give some examples of how to create each step.

### 11.16.1 File service components

The file services function is built up in the following layers. We have already completed the storage hardware and depending on the choices during Easy Setup, the storage pool(s) may have also been built.

#### MDisks

At the bottom is the physical storage, based on the Storwize V7000 storage controller. This may include in its configuration additional expansion enclosure or external storage systems. The physical storage is made up of raid arrays and from these arrays, MDisks are created. In the case of external storage, LUNs are created from the arrays and these are presented to the Storwize V7000 and managed as MDisks.

If the automatically configure storage box was ticked during the Easy Setup procedure, then the storage will have been configured with a best practice approach using all the available disks. They will have been grouped into arrays as a single MDisk and added to the same pool.

#### Pools

These MDisks are grouped together into pools and the logical blocks of storage are merged together creating a single sequential pool of blocks (or extents). The system will stripe the sequence across the MDisks to increase performance. From these pools, volumes can be created and presented to hosts as fibre channel attached SCSI LUNs or to the file modules to build a file system on. This is the block storage layer.

#### Volumes

Block volumes that are to be presented as FC attached are discussed in 11.15, “Block configuration” on page 192. Volumes are defined for file systems by the file system build process, so there is no need to create them. The file system build process creates a number of volumes based depending on its size and internal requirements. Also as they are only used by the file system, the volumes are not visible in the GUI. Volumes were previously known as vdisks. There are two types of volumes that the build process can use, uncompressed and compressed volumes.

#### File systems

A minimum of one file system is required. This provides the base file system and global namespace from which shares and exports can be created. Each file system therefore comprises a structure of directories or paths based on a single root. Any of these directories can be used to create a share or export and the directory name that is being shared is known as a junction point. This file system is the IBM GPFS. A minimum of one storage pool is required to build a file system, although a file system can use multiple pools and multiple file systems can be created from a pool. A special type of filesystem that can also be created which uses two storage pools with a default system pool for uncompressed data such as metadata and uncompressable data and a compressed pool for compressable data.

#### File sets

To improve the manageability of the file system, GPFS has a methodology that creates subsets of the file system to allow for control of file system management at a more granular



level. These are called file sets, which behave similarly to a file system. The file sets root can be anywhere in the directory structure of the parent file system and includes all files and directories above that junction point.

When creating a file set, the base directory path must exist, but the directory (or junction point) being defined must not exist as it will be created as part of the file set creation process.

You must also define the file set as dependent or independent. A dependent file set will share the same file system and inode definitions as the parent “independent” file set that contains it. If set to independent, then the file set has its own inode space allowing for independent management, such as quotas, etc.

When a file system is created an initial file set is also created automatically at the root directory.

## Shares/Exports

These are basically the same thing, the terms simply being used differently for CIFS Windows (share) and UNIX (export). Each export picks up a junction point in the file system. It then presents these files and subdirectories that are in the directory as allowed by the authentication process. These directory structures and files are “shared” using the selected network sharing protocol(s) that the hosts will be using to access the export.

For each export we need the following...

File system	The base file system as configured on the file module.
Junction point	This is the directory in the file system that will contain the data and subdirectories being shared. If files already exist in the directory or subdirectory, then for security reasons, access controls must be obeyed to create the share. If the junction does not exist, it will be created.
Network protocol	Chose from the available supported protocols. E.g. CIFS, NFS, FTP

Depending on the protocol some of the following parameters will also be needed.

User	This is the owner of this file set. Once there are files in the file set, this cannot be changed from the Storwize V7000 Unified. This user must be known to the authentication server.
Access control	Provided this is a new and empty file set, you can define the ACL controls at export creation time.

## Access control

When complete, this now presents this share to the network over the defined protocol(s). Before this share can be accessed and files read or written, the access control server must be updated. All users needing access to the export should exist. Access control of the files and directories is defined at the authentication server. While we allow access to be defined while no files exist during export creation, after files are written, the Storwize V7000 Unified administrator cannot change access control, nor can they define a new access to avoid this restriction.

Operation and administration of the access control server is beyond the intended scope of this book.

## 11.16.2 File Systems examples

The following examples are given to demonstrate the processes, show the steps involved and show the GUI windows that should be expected. These are limited examples and do not cover all possible configurations. To create a filesystem that uses compression, see Chapter 16, “Real-time Compression in the IBM Storwize V7000 Unified” on page 273.

### Create a standard uncompressed file system

Navigate to Files → File Systems. This will display the File Systems panel. Click on New File System to create a file system as shown in Figure 11-59.

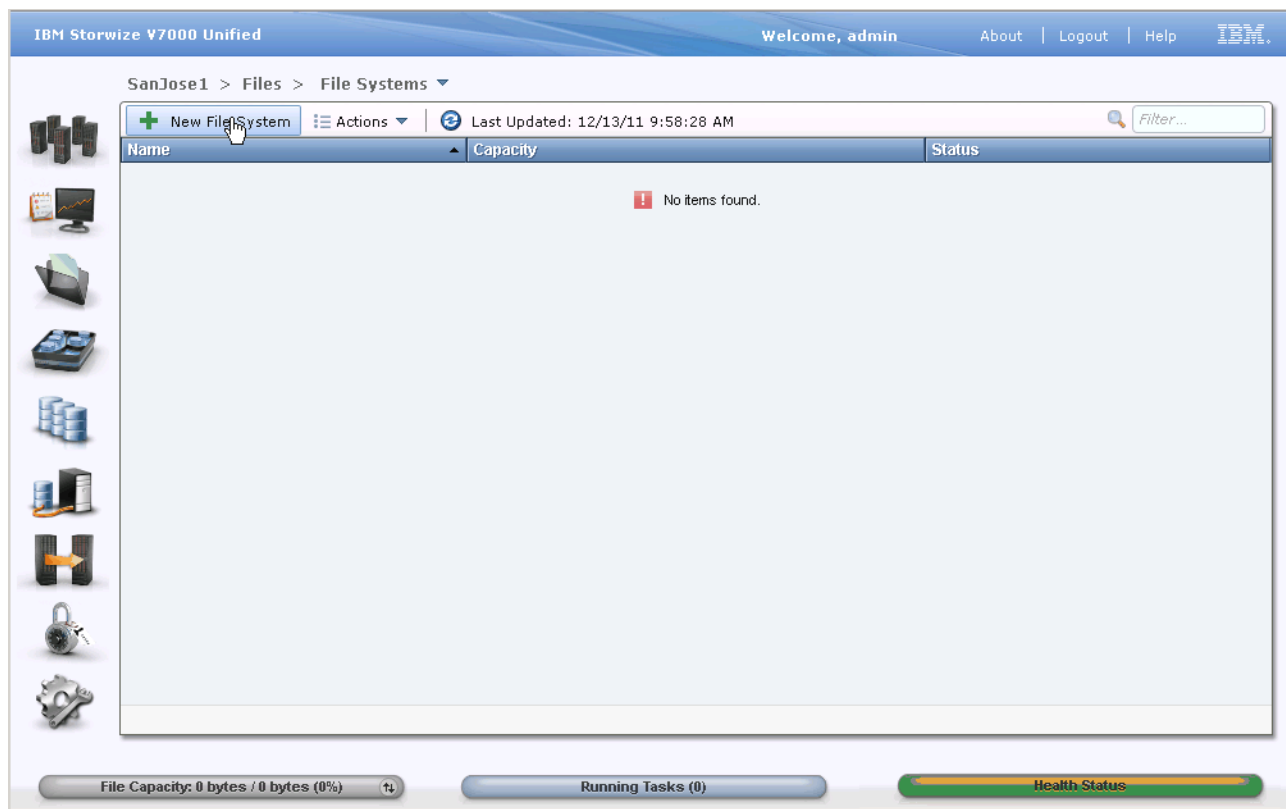



Figure 11-59 File systems

This will launch the New File System window as shown in Figure 11-60 on page 199. There are three options to build a file system, which are selected by clicking on the appropriate icon at the top of the page.


The Single Pool is the most commonly used. Here you select from the list of defined storage pools, a single pool to provide the extents for this file system. You are therefore limited to the free space of that pool. After entering a name for the file system and selecting the desired pool, you then have the option to choose how much of the available free space will be used for this file system. Do this by adjusting the slider at the bottom of the window.

When done, click OK to build the file system. A popup window will show the progress of the build. The process will first build the NSDs (Network Storage Devices) as the file module knows them, or volumes as they are to the block storage. Typically five NSD are created and their size is set to make up the specified File System size.


**New File System**



*Single Pool*



*Migration-ILM*



*Custom*

\* **File system name:**

**Pool name:**  
 system

**Select a storage pool:**

Name	Status	Free Capacity	Capacity
mdiskgrp1	✓ Online	5.4 TB	5.4 TB

**Size:**  
 2.5 GB 5.4 TB

OK Cancel

Figure 11-60 New File System

Another option is to create a File System as a “Migration-ILM”. The SONAS software has a feature that will allow the file system to use multiple pools of different performance. This is used to build a file system that will automatically migrate data to the slower storage based on predefined thresholds.

Using this feature is beyond the intended scope of this book. If it is desired to use this, contact your IBM representative for assistance and refer to the SONAS documentation:

*SONAS Implementation Guide and Best Practices Guide, SG24-7962*

*SONAS Concepts, Architecture, and Planning Guide, SG24-7963*

The third option is Custom. This gives greater flexibility in creating the file system. The storage can be obtained from multiple pools and you have the ability to manually define the file system policies. This feature requires in depth knowledge of the GPFS file system which provides the ability to set policies to automate management of data in a file system. Policies can be set and run manually from command line. If you wish to use policies seek assistance from IBM.

### Create a file set

There is no need to create file sets, but they are very useful in helping with defining and managing the data shares. To work with file sets, navigate to Files → Files Sets. The cluster has automatically defined a file set on the base of each file system, called root as shown in Figure 11-61 on page 200.

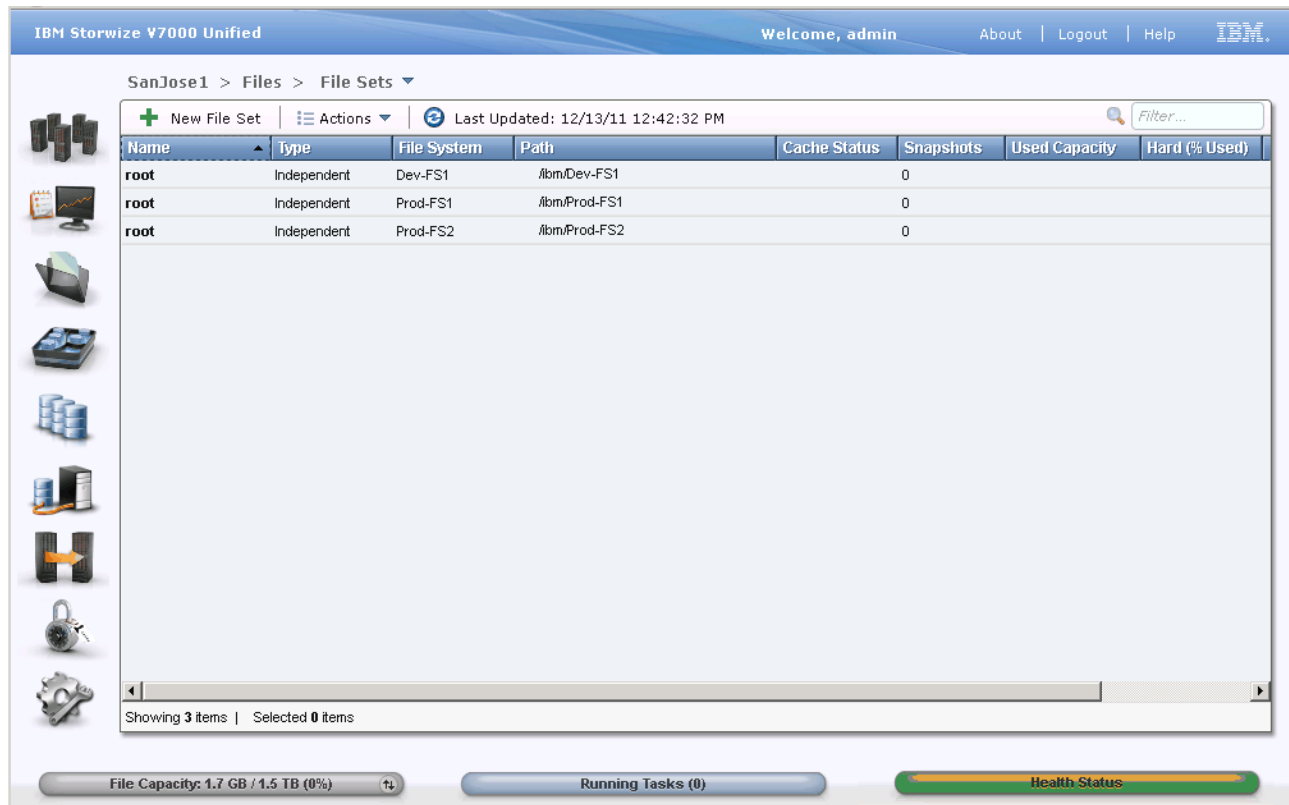


Figure 11-61 File Sets

Additional file sets can be created at any junction point in a file system. To create a new file set, click the “New File Set” button at the top of the window. This will launch the New File Set window. You can use the basic option to simply define the file set and its path, or use the custom option to fill in the quota and snapshot details now. These can be added later if desired.

**Tip:** When browsing for the path, you can right click on a directory to add a new subdirectory under it.

## Shares

To manage shares, navigate to Files → Shares. This will display the main shares window and lists all shares on the cluster. To add a new share, click the New Share button at the top of the window which will launch the New Share window as shown in Figure 11-62 on page 201.

A share can be accessed using any of the supported access methods or any combination of these. Use buttons at the top of the screen to simplify the configuration entry. Use the custom function to use multiple access methods.

All methods have the same common information...

Share name	The name by which the share is known and accessed on the network.
Path	The directory path to the shared directory, which includes the file system.
Owner	The userid of the owner of this share.

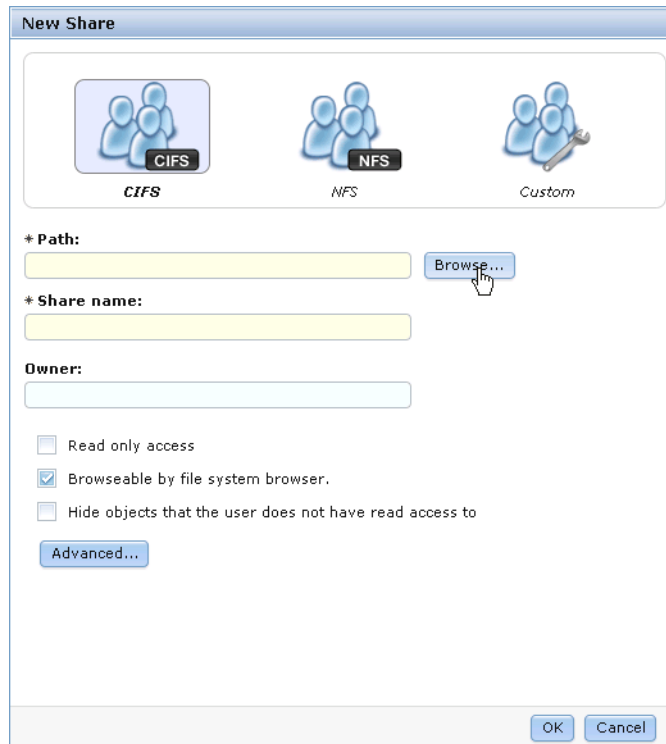


Figure 11-62 New Share

**Tip:** When browsing for the path, you can right click on a directory to add a new subdirectory under it.

### Create a CIFS share

When creating a CIFS share, first complete the common information as noted above. If using the custom option, then click the CIFS tab and click enable CIFS. If creating CIFS only, this information will be on the main panel. Set the read only, browse-able and hide objects to the desired choice and then click advanced. This will launch the window shown in Figure 11-63 on page 202.

There is a field here to enter a verbose comment about the share. This panel also allows you to add the CIFS users or groups that will have access to this share and what that access is. You can add as many entries as desired. Note that security policy allows this access to be defined on a new share, but once data has been written to the directory (or subdirectories) then access control can only be altered using the Active Directory and its associated security processes.



Figure 11-63 CIFS Advanced

## Create an NFS export

To configure a share for NFS access, click the NFS icon at the top of the New Share window, or if using custom for multi-protocol, click the NFS tab as shown in Figure 11-64 on page 203. Use the add NFS clients section to add an entry for each client ID that will have NFS access to the share. Use the “+” button to add each new entry. Set the access to read only or not read only as desired. Unless you have a good reason not to, leave the *Root squash* option ticked. If not ticked, this allows the root user on any system to receive root privileged access to the share, so full authority. Using root squash removes all access from a remote connection from root, unless specifically allowed.

**New Share**

CIFS NFS Custom

Name  
CIFS

NFS

☒ **Enable NFS**

Enable NFS

Add NFS clients

Client name or ID:	Read only access	Root squash	Secure
andreas	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> + -

Summary

OK Cancel

Figure 11-64 NFS share

**Caution:** Always enable root squash to prevent unauthorized access.

Tick the secure box if you will be using NFS secure connection to access this share.

### Create an HTTP, FTP, SCP export

If this share is already defined, then use the edit function to add any of these protocols. If a new share, create a new share and enter the common information as described above. Click the custom icon to show the options as shown in Figure 11-65 on page 204.

Now tick the HTTP, FTP, and/or SCP options as desired. User access and authentication will be the same as for the other protocols using the configured authentication method.

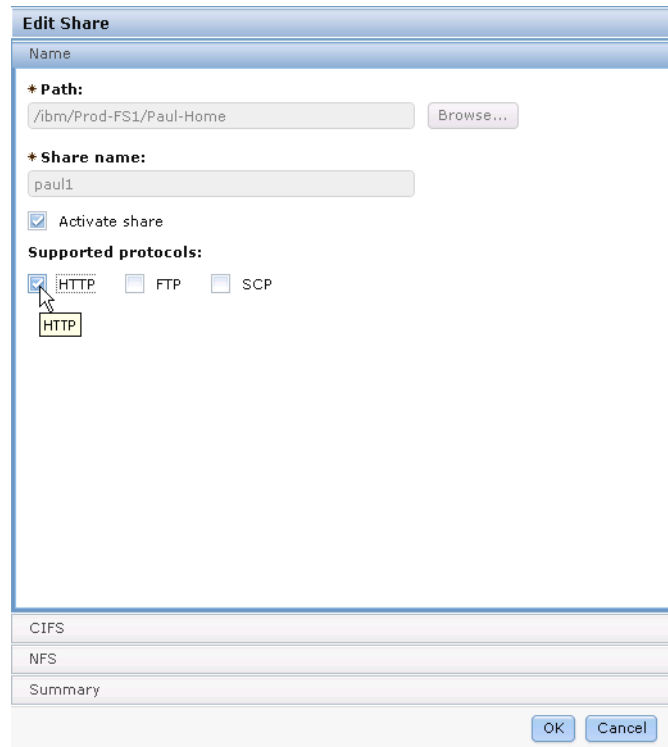


Figure 11-65 HTTP share





# Antivirus

In this chapter we describe the Antivirus feature built into the IBM Storwize V7000 Unified, how the Storwize V7000 Unified interacts with the optional external scan engines for Antivirus protection, what configuration options exist and how to set it up.

## 12.1 Overview

In general the Antivirus (AV) protection is intended for Windows/CIFS users who require an additional level of data protection, for example, against malicious, virus type software. Since UNIX environments are much less exposed, typically there is no Antivirus protection required for these. Consequently this Antivirus feature of the Storwize V7000 Unified is not supported for NFS.

The IBM Storwize V7000 Unified is provided with an AV-Connector interface to communicate with the external scan engines. Here is a summary of the supported options:

- ▶ Antivirus protection for data in CIFS file shares only
- ▶ McAfee and Symantec Antivirus products are supported
- ▶ Scalability and high availability of the Antivirus scanning can be achieved through definition of multiple scan nodes
- ▶ Scan on individual files: on file open command
- ▶ Scan on individual files: on file close command (after creation or modification)
- ▶ Batch scans on all files: configurable options
  - manually (e.g. after update of Antivirus software itself or the known virus signatures)
  - on defined schedule

The cooperation between the V7000 Unified and the external AV scan engines is shown in Figure 12-1 on page 206.

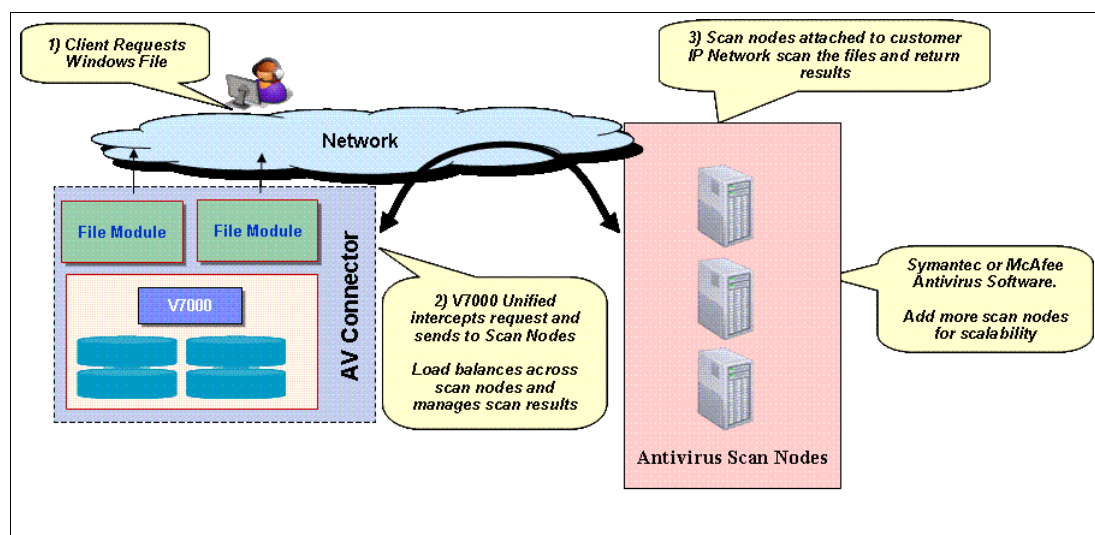


Figure 12-1 V7000 Unified intercepts I/O to enable Antivirus scan first

If Antivirus is configured, the AV connector in Storwize V7000 Unified intercepts the defined I/O operations (on open for read, if configured: on close after write) on individual files and sends a scan request to the defined external scan nodes. These use a stored file signature to verify if the file needs to be scanned. This is required if the signature of a file has changed, either because the file itself has changed (which could potentially be a sign of an infection) or an Antivirus software update has invalidated the associated signatures of all files. In this case, the scan engines will perform a scan of the file to make sure there is no infection and will update the file's signature after completing the scan successfully.

## 12.2 Scanning individual files

For individual files there is a mandatory scan on file open and a configurable option to scan the file on close after write. The AV settings use an inheritance model, and this means the settings will be applied to all subdirectories of the specified path as well. For individual file scans the scope of files to be scanned can be configured using an inclusion list or exclusion list. It is also possible to specify if access to a file is to be denied if it can not be scanned, e.g. if no scan nodes are currently available or if there is no further bandwidth to be able to scan this file in parallel.

The AV connector intercepts the file open (for read) and close (after write, if configured) in Samba:

- ▶ Scan on file open for read:
  - Only if file has changed or virus signature has changed since last scan
  - Result of last scan (signature) stored in extended attributes of a file
    - Includes time of last scan, virus signature definition used
- ▶ Optional: scan on file close after write
  - This is a proactive scan operation, and may improve performance of next file open for read if no other changes occur in the meantime
- ▶ Optional: deny access if file scanning is not available (increases security)

The performance requirements for the file scans determine the required scalability:

- ▶ The ability to scan the files needed fast enough on open for read determine the bandwidth and number of scan nodes required

**Note:** After scanning a file when it is opened the next scan can only occur (if configured accordingly) when the file gets closed. If there are simultaneous read and write accesses to the same file with byte-range locking while it continues to stay open the safety of the individual write updates cannot be guaranteed by Antivirus, a write process could write suspicious data into the file which a subsequent read could pick up unscanned / unprotected while the file is still open. This is, for example, the case for a VMware VMDK file while the virtual system is running.

## 12.3 Batch scan

In addition to the scan operations on individual files a batch scan of all files in a given path can be configured. This could be an entire file export or a subdirectory tree therein. The scan then includes all files and subdirectories of the path specified.

All these files and subdirectories in the path will then be re-scanned after the AV vendor updates its software or the AV virus signatures.

This proactive batch scan eliminates the need that all files have to be scanned on first access, when users have requested them and are waiting for it. It:

- ▶ Helps to mitigate impact on performance on file access after an AV virus signature update, for example when everyone logs in the next morning
- ▶ Updates interval of virus signatures and the AV software determines the required scan interval which might be required every night

- ▶ Stores the result of batch scan in the extended attributes (EA) of the files which means there is no need to scan the file on first access anymore if its signature has not changed in the meantime
- ▶ Verifies the files signatures/EA on every file open to provide a guarantee that there has been no change to the file since the last scan

**Important:** when using HSM and Antivirus bulk scans a bulk scan does not rescan files which have been migrated off the file system using HSM. This means that no file recall is required, preserving the results of HSM policies defined. Scanning a file will update its 'last access time' property.

## 12.4 Setup and configure Antivirus

In this section we describe the steps and options available to set up the Antivirus configuration according to specific needs. The options are described using the Graphical User Interface (GUI), but similarly the setup could also be done using the Command Line Interface (CLI).

The infrastructure for the actual Antivirus scanning process is provided outside of the Storwize V7000 Unified system.

Prerequisites of that infrastructure are:

- ▶ Client supplies and maintains scan nodes
- ▶ Client installs supported Antivirus vendor product on scan nodes:
  - Symantec
  - McAfee
- ▶ The scan nodes are attached to customer's IP network and can communicate to the V7000 Unified system(s).

There are other important considerations related to this topic:

- ▶ Availability and scalability is achieved by provisioning multiple scan nodes:
  - During AV setup a pool of scan nodes is configured to be used in the V7000 Unified
  - Each AV connector instance randomly chooses a scan node from this pool
  - If a scan node fails, it is temporarily removed from the pool:
    - The AV connector checks regularly to see if the node is back online
    - If it is back online again, it will be re-added to the pool of available scan nodes automatically

### 12.4.1 Antivirus setup steps

The Antivirus configuration panels can be found in the **Files** → **Services** → **Antivirus** panels as shown in Figure 12-2 on page 209:

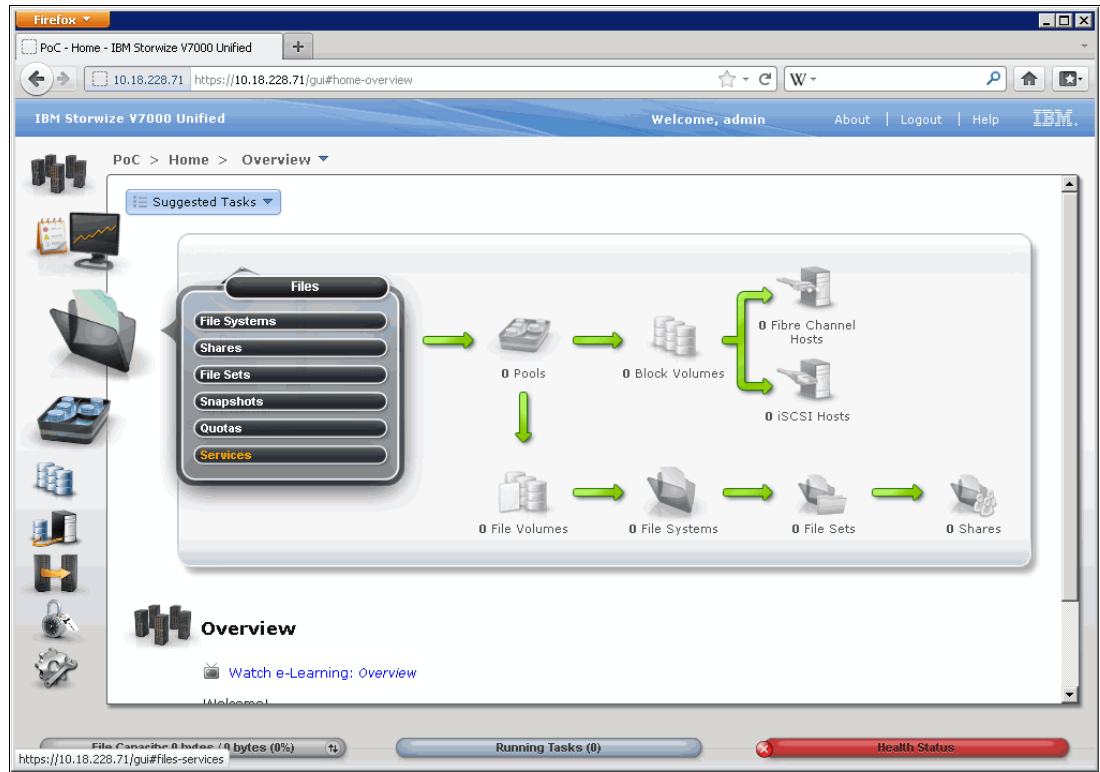


Figure 12-2 Menu path for Antivirus: Files -> Services

Figure 12-3 on page 210 shows the available services, Antivirus is already selected but in this example, it has not been configured yet

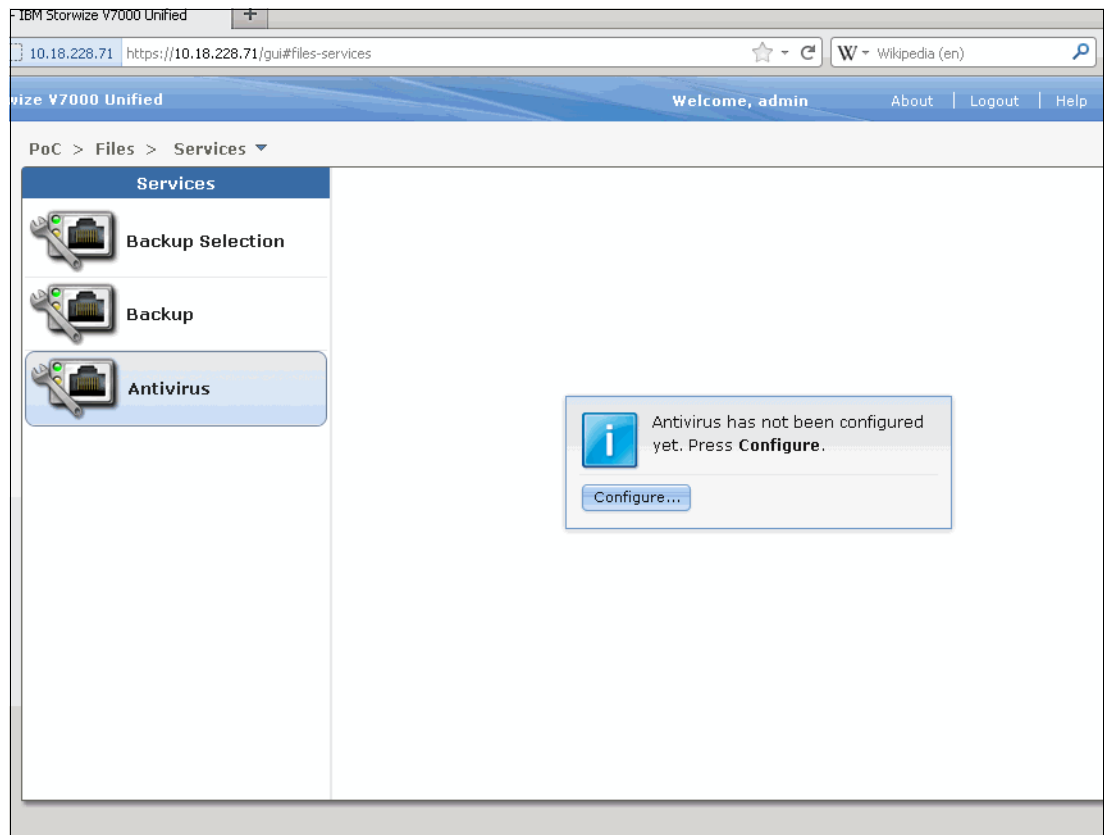


Figure 12-3 Available Services and Antivirus selection

After selecting 'Configure' the next step is to specify three settings according to your needs:

- ▶ List of available scan nodes
- ▶ Scan protocol they are using
- ▶ Global time-out for every scan

The corresponding GUI panels are shown in Figure 12-4 on page 211 and Figure 12-5 on page 211.

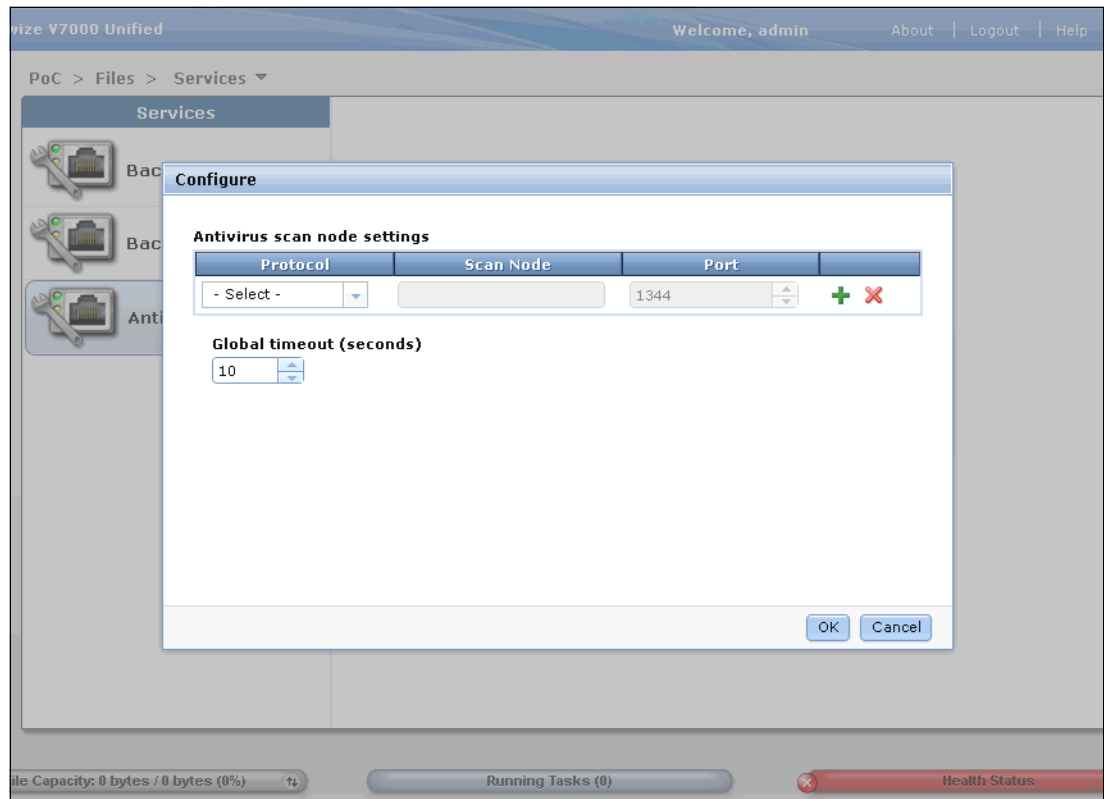


Figure 12-4 Definition of Scan nodes, protocols and time-out value

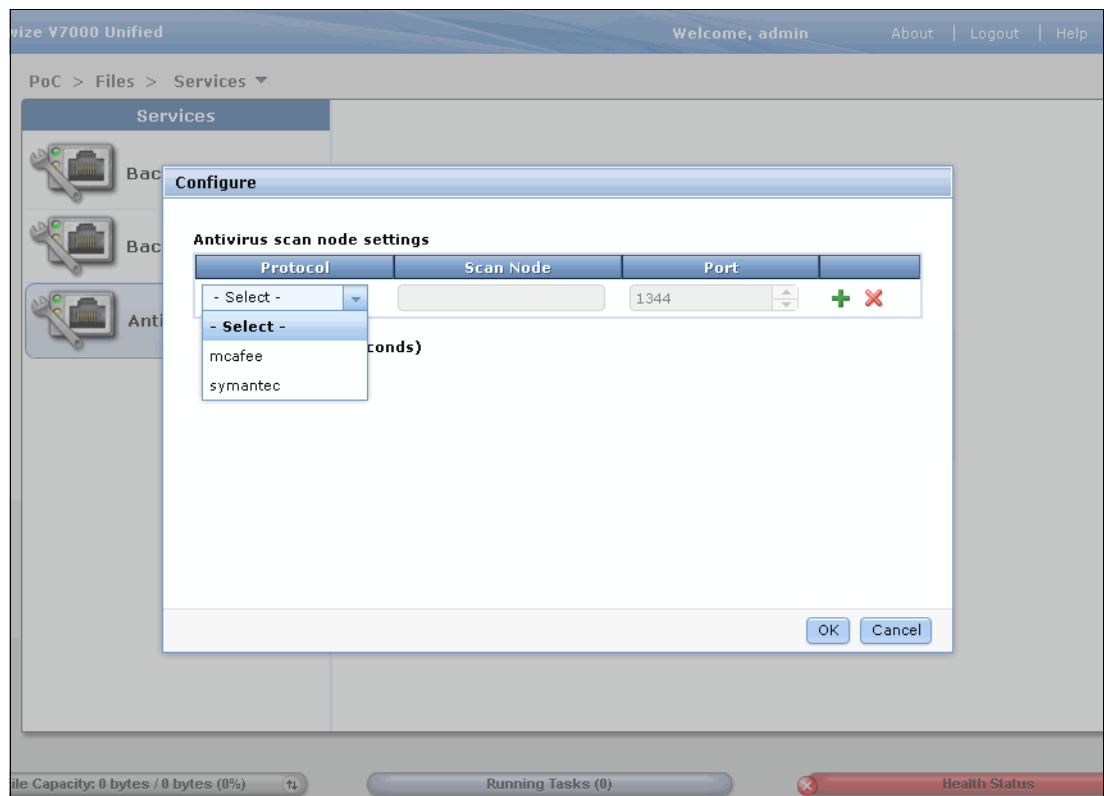


Figure 12-5 Selection of supported scan protocols

After saving the Scan node settings, the main screen for Antivirus is launched, showing two tabs: *Definitions* and *Batch Scans*.

The settings for the scan nodes are not displayed on the main screen, they will be shown or can be changed using the 'Actions' drop-down menu.

Figure 12-6 on page 212 shows the main screen.

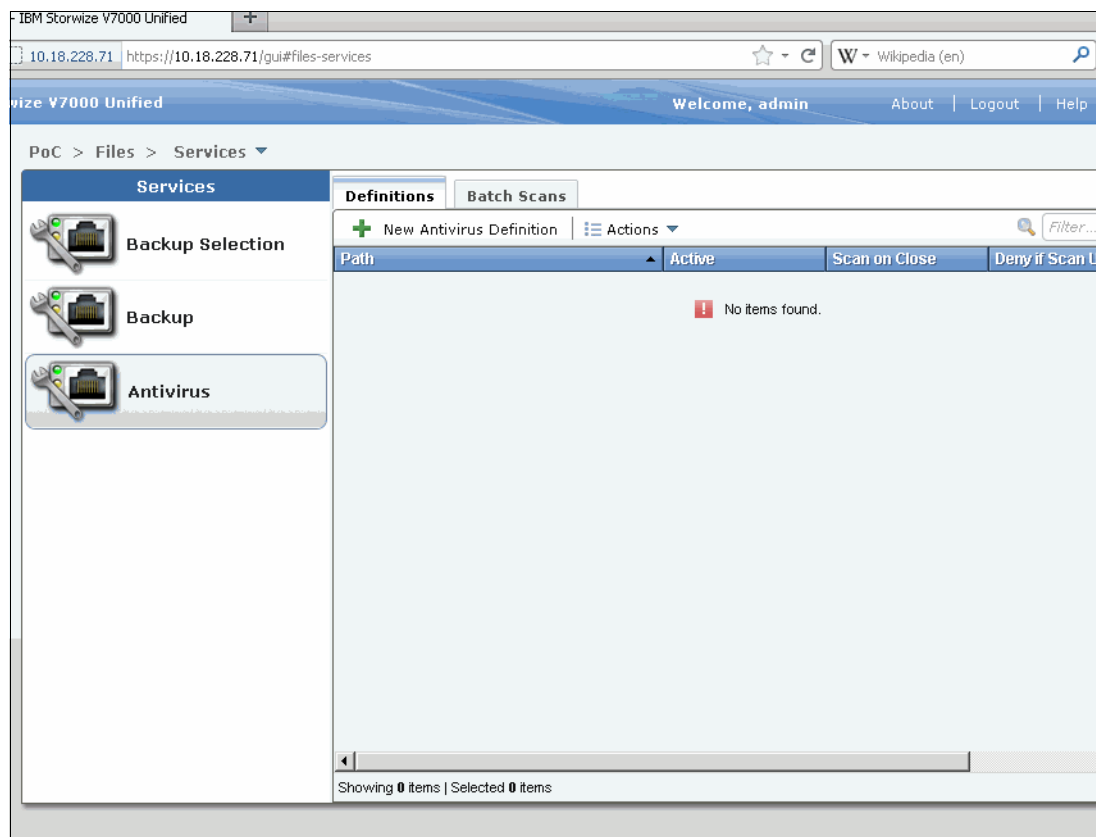


Figure 12-6 Antivirus overview screen showing the two main tabs 'Definitions' and 'Batch Scans'

Selection of Adding a 'New Antivirus Definition' starts the next screen to specify multiple options, as shown in Figure 12-7 on page 213:

- ▶ Specify the path to be scanned:
  - Allows you to browse for available paths
- ▶ Each Antivirus definition can be enabled or disabled
- ▶ Enable scan of files on close after write if required
  - Provides additional security
- ▶ Deny client access if the file could not be validated
  - Provides additional security
  - Will prevent access even if just scan nodes not available or not reachable on the network
- ▶ Specify which default action to take for infected files
  - Options are *No action*, *Delete* or *Quarantine*
- ▶ Definition of scope for the scan:



- Options are *All files*, *Include files with extensions specified* (which means: scan just these), *Exclude files with extensions specified* (which means: scan all others)

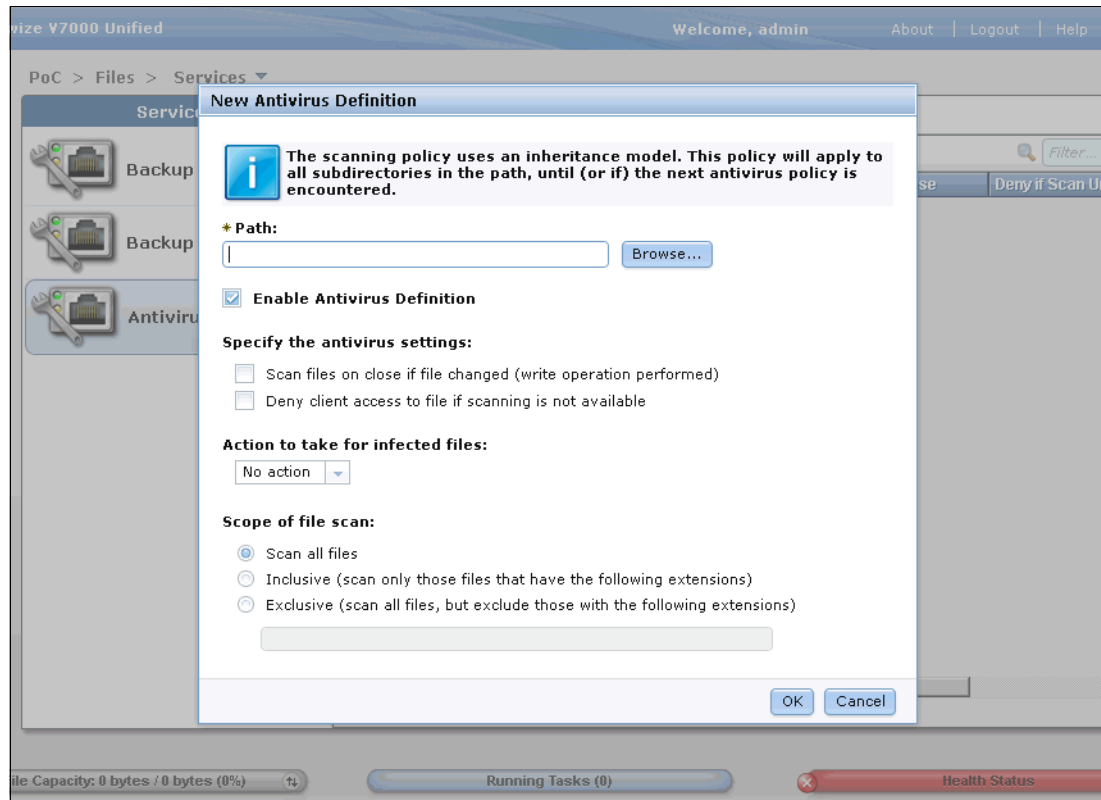


Figure 12-7 Create new Antivirus definition: configurable options

Once a new definition is created, it is shown on the main screen for Antivirus as shown in Figure 12-8.

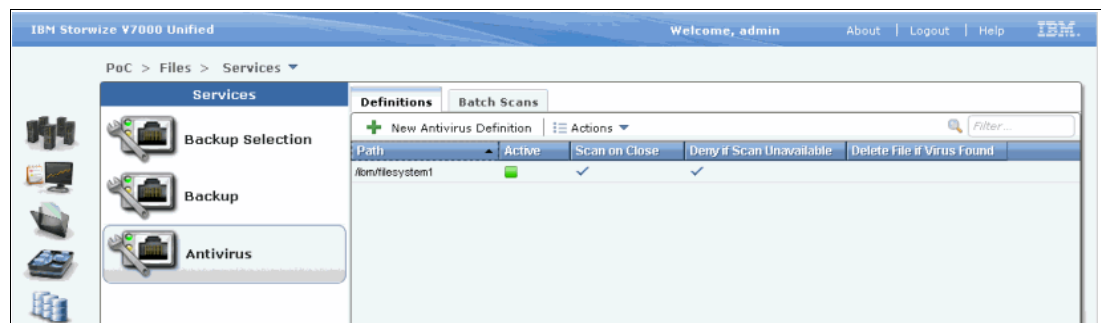


Figure 12-8 List of stored Antivirus definitions

Beside the Antivirus definition there is an optional feature to define a *Batch Scan*.

This can be setup in the corresponding tab as shown in Figure 12-9 on page 214.

The available options are:

- Frequency: *Once a Day, Multiple Days a Week, Multiple Days a Month, Once a Week, Once a Month*

**Note:** Due to the usual change rate for virus definitions of at least once a day for all major vendors of Antivirus software, the most common setting for this is expected to be *Once a Day*

- Time of day: Presets built-in with steps of 15 minutes, but any value could be entered
- Paths to scan: specify for which path(s) this batch scan definition should be used

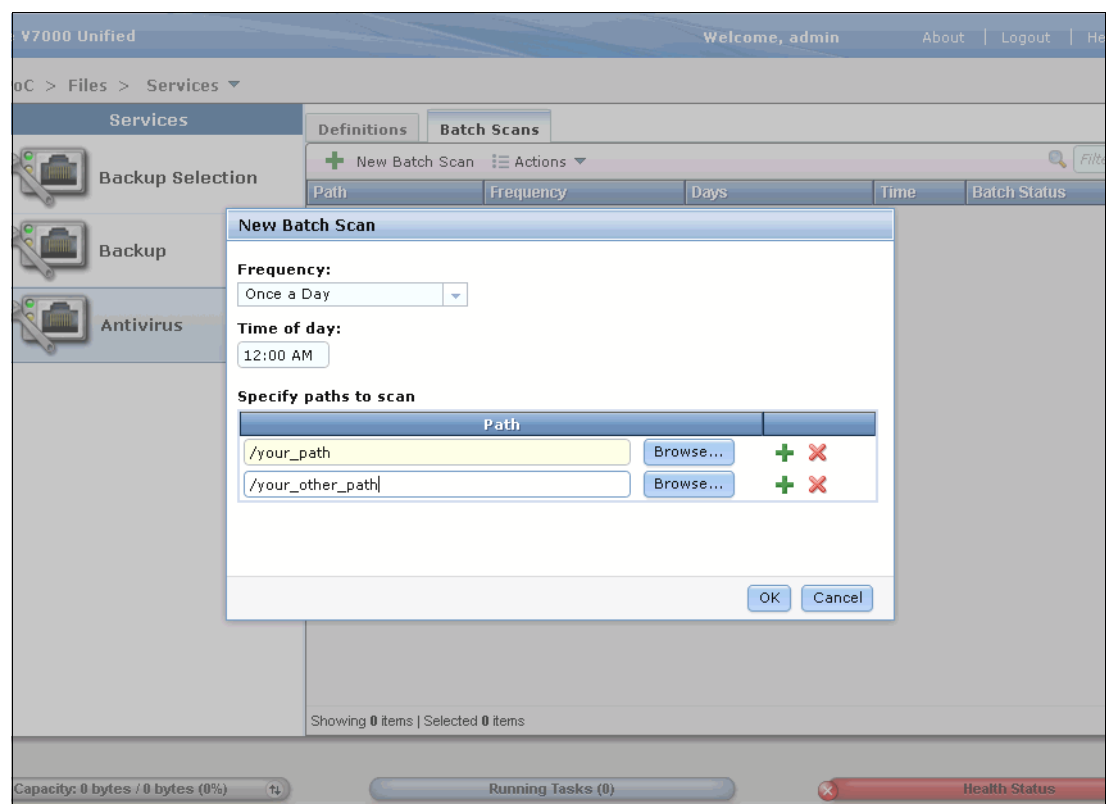


Figure 12-9 Definition of new Batch Scan

The Batch Scans defined are listed on the main page for Batch Scans as shown in Figure 12-10 on page 215.



Figure 12-10 List of stored Batch Scan definitions

Antivirus setup is now complete.

We suggest that you test the antivirus scan engine to ensure correct and expected operation.





## Performance and Monitoring

While the Storwize V7000 Unified incorporates a lot of automatic load balancing technology, there are some areas that can be manually tuned to obtain the best performance, especially under heavy loads. The file and block components need to be treated independently with regard to monitoring and tuning performance. IBM Tivoli Productivity Center (TPC) can monitor the entire Storwize V7000 Unified cluster and detail performance in individual areas of block and file storage.

**RtC:** If compression is being used or considered, we strongly suggest that you consult *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859, especially the chapter on Performance guidelines, available at:

<http://www.redbooks.ibm.com/redpieces/abstracts/redp4859.html>

An in-depth discussion on performance is beyond the intended scope of this book and we only briefly describe the manner of monitoring and the data that should be looked at.

## 13.1 Monitoring

Monitoring needs to be done in several places as it often requires analysis of several areas to locate a performance issue.

The health status indicator at the bottom of the GUI screen, alerts, event logs and status details for the file modules provide the first point of investigation for a problem.

**Important:** It is very important to address all issues and maintain the health status of system as green in order to prevent issues such as volumes going offline due to lack of available physical storage.

The Storwize V7000 Unified provides a quick-look set of graphs that show functions for file and block performance. Note that these graphs have differing collection time frames with block only showing 5 minutes which is not recorded. They are intended to highlight problem areas in real time and are very useful in narrowing down the search during sudden and critical performance hits.

To display the graphs, navigate to Monitoring → Performance. The three tabs available for performance monitoring are for File, Block, and File Modules.

The **File** performance graphs show metrics about client throughput, file throughput, latency, and operations. The time frame for these graphs can be selected from minutes to a year. See Figure 13-1 for details.

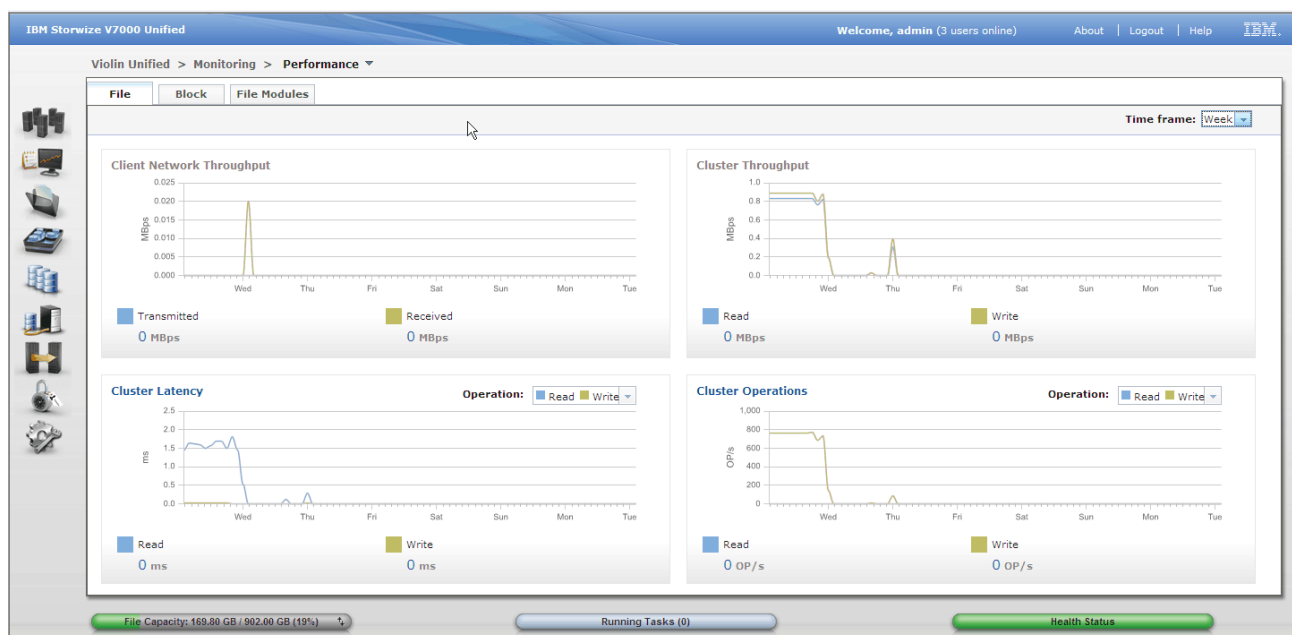


Figure 13-1 File performance

The **Block** performance graphs show real-time statistics that monitor CPU utilization, volume, interface, and MDisk bandwidth of your system and nodes. The CPU utilization graphs also show separate system and compression statistics. Each graph represents five minutes of collected statistics and provides a means of assessing the overall performance of your system. See Figure 13-2 on page 219 for details.

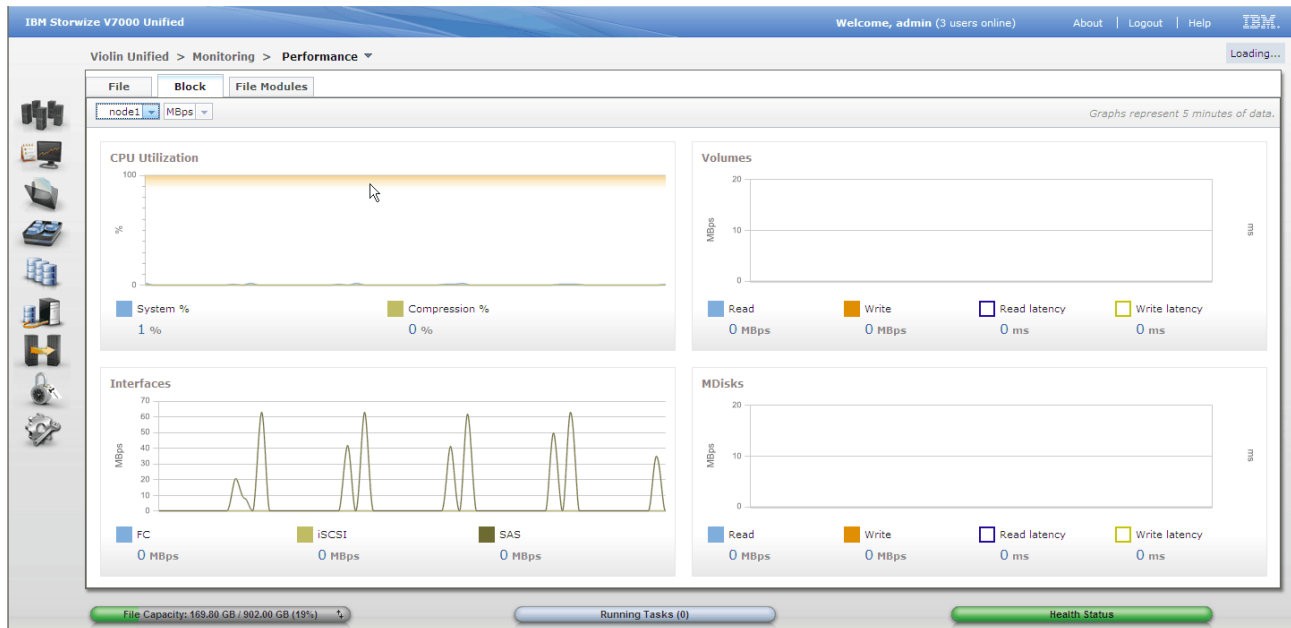


Figure 13-2 Block performance

The **File Modules** performance graphs show the performance metrics for each file module for CPU, memory, and public network and you can also select numerous data types for each item selected such as collisions, drops, errors, and packets for public network statistics and other related data types for CPU and Memory statistics. The time frame for these graphs can be selected from minutes to a year. See Figure 13-3 for details.

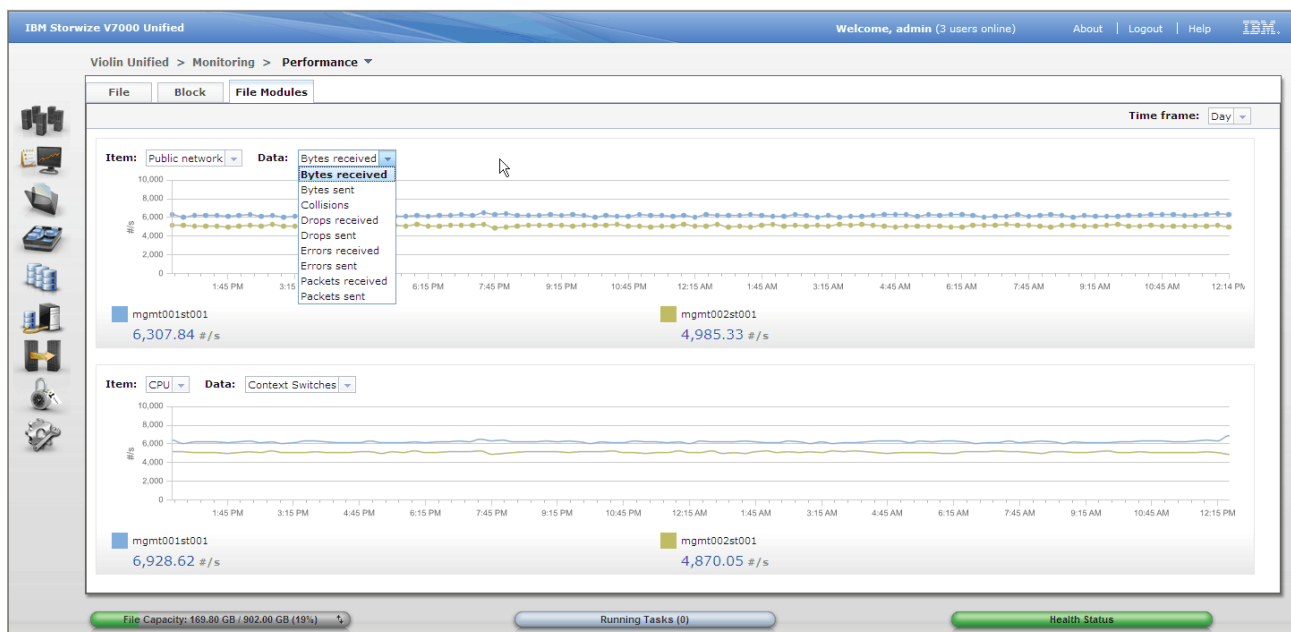


Figure 13-3 File Module performance

The block storage system is continuously logging raw statistical data to files. Depending on the setting for the statistics interval (default = 5 minutes) all the statistical counters are collected and saved regularly. Only the last 15 files are kept and the oldest ones are purged on a first in last out (FILO) basis. TPC and other performance tools collect these raw files.

These stats files are also collected by the support data collection, which will collect all current files, that is, the last 15 saves. IBM technical support will use this data if required, to analyze problems. In smaller installations where TPC is not used, IBM support may be able to assist with identifying a problem performance area.

Most performance issues are found outside the cluster. Analysis of SAN switch logs, external storage controller logs and host logs is usually required to fully identify a performance problem.

File performance issues will typically be network related and may also be due to file module resources. Use the cluster performance graphs and your network monitoring tools as well as the host logs to monitor metrics.

There are many areas that need to be considered when analyzing a performance problem. Again, the health status indicator at the bottom of the GUI screen, alerts and the event logs provide the first point of investigation for a problem, which is discussed in the Troubleshooting chapter. Using the `!health` command to expand in detail a unhealthy system can be very useful. By using the `-i` parameter, more detail can be seen on individual components.

Next is to ensure the traffic to the two file modules is balanced. Because the IP addresses are shared across the file module interfaces, are dynamically assigned by DNS and can dynamically move, it is possible for the traffic to become unbalanced. Also, using the graphs, you may be able to identify if an internal resource is overloaded or faulty, but this will not locate an overloaded export or file system. There are commands that can be run to interrogate software components and their health, but these will require a level of experience to use quickly and successfully. IBM technical support will also have additional tools such as tracing facilities that can help identify problem.

## 13.2 TPC

IBM Tivoli Productivity Center has been enhanced to include the Storwize V7000 Unified. TPC will collect and track data from the cluster and is able to provide details on demand. Some areas that are able to be drilled down to are General Cluster Information, Nodes, Pools, Network Share Disks (NSD), File Systems, File Sets, Exports and selected directories.

### 13.2.1 Configuration

TPC does not use a CIM agent, but is able to natively connect to the Storwize V7000 Unified. To configure the cluster to communicate with TPS requires one simple step, create a userid with full authority, for TPC to use. Assign a password and give the userid and password details to the TPC administrator. All the remaining configuration is done from TPC and the configuration details will be retrieved and built automatically.

Additionally, if specific file system scanning is required, you might need to create and export a share to host running the TPC agent.





# 14

## Backup and Recovery

In this chapter we look at the backup and recovery of the Storwize V7000 Unified configuration data, and the host user data written to the Storwize V7000 Unified. These are two distinctly different and unrelated areas and we will cover each one separately. We will also look at the recovery processes to restore from backup.

## 14.1 Cluster backup

In the topics that follow we discuss cluster backup.

### 14.1.1 Philosophy for file and block

The storage cluster is made up of two major components, the Storwize V7000 storage subsystem and the V7000 Unified file modules. Because each system has its own existing and established backup processes which are quite different in requirements, they are backed up independently.

#### Storwize V7000 (block)

The primary level of backup is the knowledge that the most current copy of the configuration and status is held by the other operational storage nodes in the cluster, so should a node fail or at any time need to join the cluster, then it will receive all the required configuration and status information from the config node.

Additionally, the config node is saving regular checkpoint data on the quorum disks. These saves occur about every 10 minutes. The three quorum disks have several functions, including tie-breaking in the event of a split cluster.

A full config XML file is generated every day at 01:00. The is saved on the hard disk drive of the config node at that time and also written to the quorum. This file contains the hardware and storage configuration, including MDisk, volume and host mapping details, but it does not include the storage extent mapping tables. This backup can be manually triggered at anytime through the CLI.

#### File modules (file)

Again, like the storage, the V7000 Unified component relies on the knowledge that as long as one file module is operational and online, the other can be restored, even from a complete loss of all system data. The difference here is that the file modules are sufficiently unique that the backup is only usable on the particular module it was created from.

A backup is taken automatically at 02:22 every day on each file module. This is then packaged into a single file and copied to the other file module and stored on its hard disk drive, so that each file module has a copy of the others backup. Previous generations of backup are kept in case of corruption or a need to step back. This backup can be manually triggered at anytime through the CLI.

### 14.1.2 Storage enclosure backup (Storwize V7000)

The storage component is backed up independent of the file modules and uses the inherent processes of the Storwize V7000. The backup process is designed to backup configuration information and current status of your system, such as cluster setup, userid setup, volumes, local Metro Mirror information, local Global Mirror information, managed disk (MDisk) groups, hosts, hosts mappings and nodes. 3 files are created

svc.config.backup.xml	This file contains your current configuration data and status.
svc.config.backup.sh	This file contains a record of the commands issued by the backup process.
svc.config.backup.log	This file contains the backup command output log.

If an immediate backup is desired, such as before/after a critical or complex change or before a support data collection, then issue this CLI command

**svcconfig backup**

This will start an immediate backup process. Successful completion is indicated by a return to the prompt without an error message. Note that while the backup is running, any commands or processes that might change the configuration are blocked.

A manual backup is not normally required as the task runs daily at 01:00. Manual backup is only needed if changes are made or if the current status needs to be reflected in the support data.

There are high speed paths between all the nodes that form the cluster. These can be over the SAN and also with Storwize V7000, there is a high speed internal bus in the enclosure between the node canisters that also carries these paths. This enclosure link removes the need for SAN paths between the nodes and allows the cluster to operate without connection to SAN if desired.

These links are used to maintain data integrity between the 2 nodes in an IO group and are also used to ensure the cluster status and configuration is known to all nodes at all times. Any changes to the configuration, including the extent maps, are shared over these paths. This means that any node can assume the config role and will have fully up to date configuration at any time. Should a node be disconnected from the cluster (i.e from the other nodes) for any amount of time, it cannot continue handling data. It must therefore leave and re-join to re-learn the “current” configuration, thereby maintaining complete integrity of the cluster.

The code will always try to use three MDisks as quorum drives, provided at least three are available. One will be marked as the active quorum. The Storwize V7000 will reserve an area of each MDisk to contain the quorum data. Which MDisks are used and which one is the active can be altered if desired. These MDisks are seen by all nodes in the cluster. The config node is periodically writing checkpoint data to the quorums (about every 10 minutes). The quorums also serve other functions including being a means of communication between nodes that have become isolated so that tie breaking can be resolved.

If a node is properly shutdown, it will “save hardened data” and gracefully leave the cluster to be offline. When an offline node starts up, it will search for the active cluster (i.e. if any other node(s) from this cluster are already active and have formed a cluster). If so, then it will request and be granted permission to rejoin. If no cluster exists (i.e. no node responds) and the node can see the quorums to confirm that the cluster is not active, this node will assume it is the first alive and form the cluster.

Should a node fail, or for any reason be unable to save hardened data, then it has exited the cluster and can never rejoin. Any node starting up that was a cluster member but has no valid hardened data, will fail and post a permanent error, typically 578. This node must now have its definition in the cluster deleted and be added again as a new node.

Should all nodes fail without a proper shutdown, i.e. all nodes have failed from the cluster and no active nodes remain, then the quorum data can be used to rebuild the cluster. This is a rare situation and is known as a tier 3 recovery.

### 14.1.3 File module backup

A pair of 1Gb ethernet connections form the physical path over which a link exists between the two file modules. This link is used for communications and to maintain integrity of the file

systems across the cluster. It also is used to resolve tie-break situations, for recovery actions, software loading and for transfer of backup data.

A backup process runs at 02:22 each day, which creates a packaged single file. A copy of the backup file is then stored on the other node, so each file module has a copy of the other nodes backups. The file is given a unique name and stored on the file module in directory */var/sonas/managementnodebackup*

This backup can be manually triggered at any time through the CLI.

**backupmanagementnode -v**

## 14.2 Cluster recovery

In the topics that follow we describe cluster recovery.

### 14.2.1 Storage enclosure recovery (V7000)

As with the other products in this family (SVC and Storwize V7000) recovery of a node or cluster from backup is defined in 4 levels known as tiers. These are numbered in increasing level of impact and severity. For the Storwize V7000 Unified the tiers are summarized as follows:

Tier 1 (T1)	recovers from single failures of hardware or software without loss of availability or data.
Tier 2 (T2)	recovers from software failures occurring on nodes with loss of availability but no loss of data.
Tier 3 (T3)	recovers from some double hardware failures but does potentially involve some loss of customer data.
Tier 4 (T4)	assumes the loss of all data managed by the cluster and provides a mechanism to restore the clusters configuration to a point where it is ready to be restored from an off-cluster backup (e.g. tape backup).

The following descriptions are to assist the reader to understand the effect of each recovery and an overview of the processes involved. In all cases where action is required, IBM documentation and/or IBM Technical Support guidance should be closely followed.

#### Tier 1

A tier 1 recovery is where the cause of the problem can be resolved by warm starting the node. The most common trigger for this is a hardware or software error being detected by the storage node's software, which triggers an assert. An assert is a software initiated warm start of the node. It does not reboot the operating system, but restarts services and resumes its previous operational state. It also takes a dump for later analysis.

The warm start occurs quickly enough that no connectivity is lost and paths are maintained. No data is lost and cache will destage once the node is recovered. There should be no effect on the file modules, all IO will eventually be honoured. Most asserts are only observed in the event log or by an alert being posted. Any assert should be notified to IBM support for investigation.

## Tier 2

The tier 2 recovery level is much the same as a tier 1 but the node has failed and must be re-added. Again this process is automated and once the storage node has rebooted the cluster will perform recovery procedures to rebuild the nodes configuration and add it back into the cluster. During this process most configuration tasks are blocked. This recovery is normally indicated by a 1001 error code. This recovery will take about 10 minutes.

It is important to follow the procedures on the Information Center and as directed by the maintenance procedures and IBM support very carefully.

The Storwize V7000 storage component will recover and become fully operational again without intervention. No data will have been lost, cache is recovered and destaged from the partner node. Any IO being directed to the node at the time of failure will fail. Depending on multipath driver support on the hosts, this IO will either be hard failed or retry to the other node in the IO group.

File services will need manual intervention to recover. Wait for the Storwize V7000 to complete its recovery and a message to be displayed on the GUI console. Call IBM support for assistance and guidance. Start with the event log and perform maintenance on the error. Ensure you follow the directions carefully and also the Information Center.

Recovery of the file services will likely entail reboots of the file modules one at a time and checking of the status of file systems. Follow the guidance of IBM support and the online maintenance procedures. Once the recovery is complete, support will ask for logs to confirm health before advising to release the change locks.

## Tier 3

Tier 3 recovery is required if there are no storage nodes remaining in the cluster. The cluster is then rebuilt from the last good checkpoint, as stored on the quorum disks. This is a rare situation but should it be required, direct assistance from IBM support is needed.

The Storwize V7000 storage component can be recovered by the user by following the procedures in the Information Centre. IBM support can provide assistance if needed.

Once the Storwize V7000 storage component has been recovered, direct IBM support is needed to recover the file services. This will require collection and transmission of log files and remote access to the CLI, which can be achieved using IBM AOS or any acceptable process that allows support access. IBM support will investigate the status of various components in the SONAS file modules and repair as needed. It is important that no actions are done onsite without support guidance. This process can take some time.

There is potential for data loss if data existed in a node's cache at the time of failure. All access to data on this cluster is lost until the recovery is complete.

## Tier 4

This rare process is very unlikely and is not automatically invoked. It will be directly driven by IBM support after all attempts to recover data have failed or been ruled out. All user data is considered lost and the Storwize V7000 Unified will be reinitialized and restored to the last known configuration.

On recovery completion, the storage component will have the MDisks, pools and volumes defined. All previously mapped volumes will be mapped to the host definitions. No user data in volumes will be recovered and must be restored from backup.

The file modules will be reloaded and reinitialized. The user will need to provide all the configuration data used to initially build the system, which will be used to rebuild the

configuration to the same point. All file system configuration will also be reentered, including file systems, file sets and shares.

Once file services are resumed, user data can then be restored from backup. The process used will depend on the backup system in use. This is covered in 14.3, “Data backup” on page 227.

## 14.3 Data backup

The file data stored on the cluster can be backed up using conventional means by the servers or a server based backup system. Alternatively, the file systems mounted in the GPFS file system on the cluster can be backed up directly. Currently two methods of data backup are supported for backing up the file systems on the Storwize V7000 Unified, TSM and NDMP. TSM (Tivoli Storage Manager) uses IBM's proven backup and restore tools. NDMP (Network Data Management Protocol) is an open standard protocol for NAS backups.

**Warning:** Do not attempt to use NDMP with HSM or TSM. This is not supported and could cause an outage.

### 14.3.1 Data backup philosophy

#### Server

This method reads the data from the cluster using the same methodology that the servers read and write the data. The data may be read by the server that wrote it, or by a standalone server dedicated to backup processes that has share access to the data. These methods are outside the intended scope of this book.

#### TSM

The storage agent is included with the cluster software and when enabled will run on the file modules. Once configured, TSM can be scheduled to backup file systems from the GPFS system running on the file modules to external disk or tape devices. This data backup and management is controlled by TSM.

#### NDMP

A Data Management Application (DMA) is installed on an external server. When enabled and configured, the NDMP agent runs on the file modules and communicates with the DMA. Backup scheduling and management is controlled in the DMA. Data is prepared by the cluster and sent to the DMA server, which will then store the backup data onto external disk or tape.

### 14.3.2 TSM

TSM is able to utilize the Active Cloud Engine (ACE) which is incorporated in the Storwize V7000 Unified.

To configure TSM, we must enable TSM as the backup protocol and create a TSM server definition on the Storwize V7000 Unified.

First select the protocol by navigating to Files → Services and click on Backup Selection. Confirm that the TSM option is selected with a solid dot, if not use the edit function to change it as shown in Figure 14-1 on page 228.

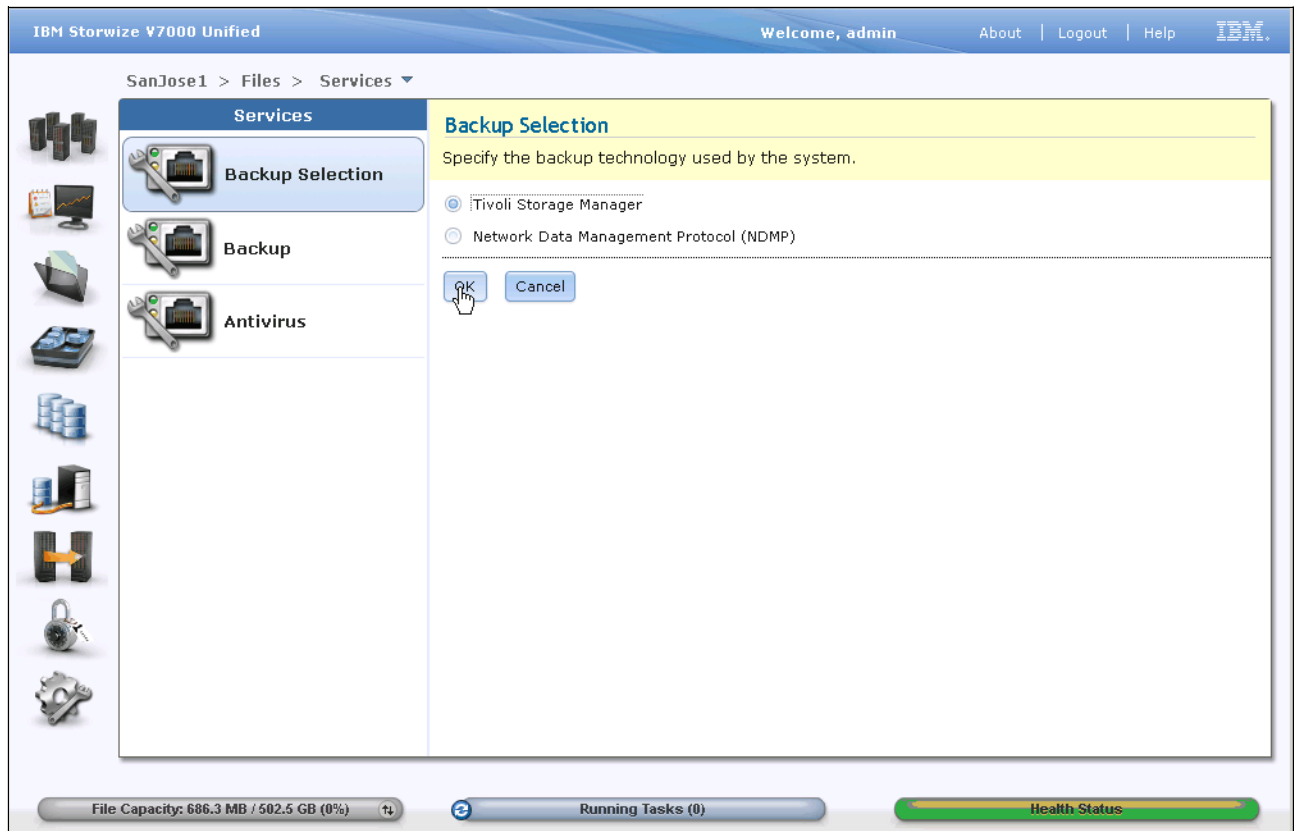


Figure 14-1 Select TSM protocol

Now click on the backup icon and this will display a message indicating that TSM is currently not configured. Click the configure button. This will launch the TSM configuration page which will show that there are definitions found. From the actions pull-down, create a new definition as shown in Figure 14-2.

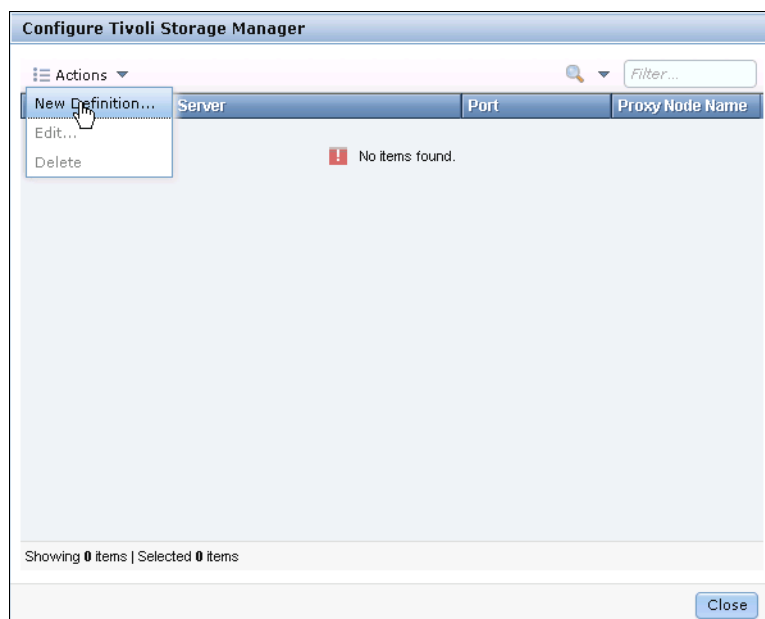


Figure 14-2 Add new definition



The new definition window is displayed. This has 4 tabs, the general tab will be displayed first. Complete the TSM server and proxy details as shown in Figure 14-3.

**New Definition**

General

\* Tivoli Storage Manager server alias:  
IBMTSM

\* Server (IP or host name): 10.18.228.25      \* Port: 1500

\* Proxy node name:  
SanJose1

Node Pairing

Script

Summary

OK Cancel

Figure 14-3 New definition - general

Now click on the node pairing tab. The two file modules will be shown. Add a prefix as required by typing the prefix in the nodes prefix field and clicking apply. Also add the password by entering in the common password field and clicking apply as shown in Figure 14-4.

**New Definition**

General

Node Pairing

Common password: [ ] Apply

Nodes prefix: [ ] Apply

Interface node and Tivoli Storage Manager node pairings:

<input checked="" type="checkbox"/>	Interface Node	Node	Password
<input checked="" type="checkbox"/>	mgmt001st001	SanJose-mgmt001st001	.....
<input checked="" type="checkbox"/>	mgmt002st001	SanJose-mgmt002st001	.....

Script

Summary

OK Cancel

Figure 14-4 New definition - node pairs

Clicking on the summary tab will display the new configuration.

Before clicking OK, you must configure this definition to the TSM server. Click on the script tab to see the commands required to configure TSM. These have been provided for your

convenience and can be copy/pasted to the TSM console as shown in Figure 14-5 on page 230.

**New Definition**

General

Node Pairing

**Script**

**Configuration text (can copy and paste to session):**

```
register node SanJose1 [replace with your own password]
register node SanJose-mgmt001st001 password
register node SanJose-mgmt002st001 password

grant proxynode target=SanJose1 agent=SanJose-mgmt001st001
grant proxynode target=SanJose1 agent=SanJose-mgmt002st001
```

Summary

OK Cancel

Figure 14-5 New definition - script

When the commands have completed successfully, then click the OK button to build the definition.

You are now able to perform backups from TSM of the data on this cluster. Management and usage of TSM and the processes required to backup the data on the cluster are handled by TSM and not within the intended scope of this book. Consult your TSM technical support for assistance with TSM.

### 14.3.3 NDMP

To configure the NDMP agent on the Storwize V7000 Unified requires three basic steps.

1. Create the network group definition
2. Set the configuration values, note while defaults are normally fine, three of the values need to be set.
3. Activate the node group

To create the group, first select the protocol being used. Navigate to Files → Services and click on Backup Selection. Confirm that the NDMP option is selected with a solid dot, if not use the edit function to change it as shown in Figure 14-6.

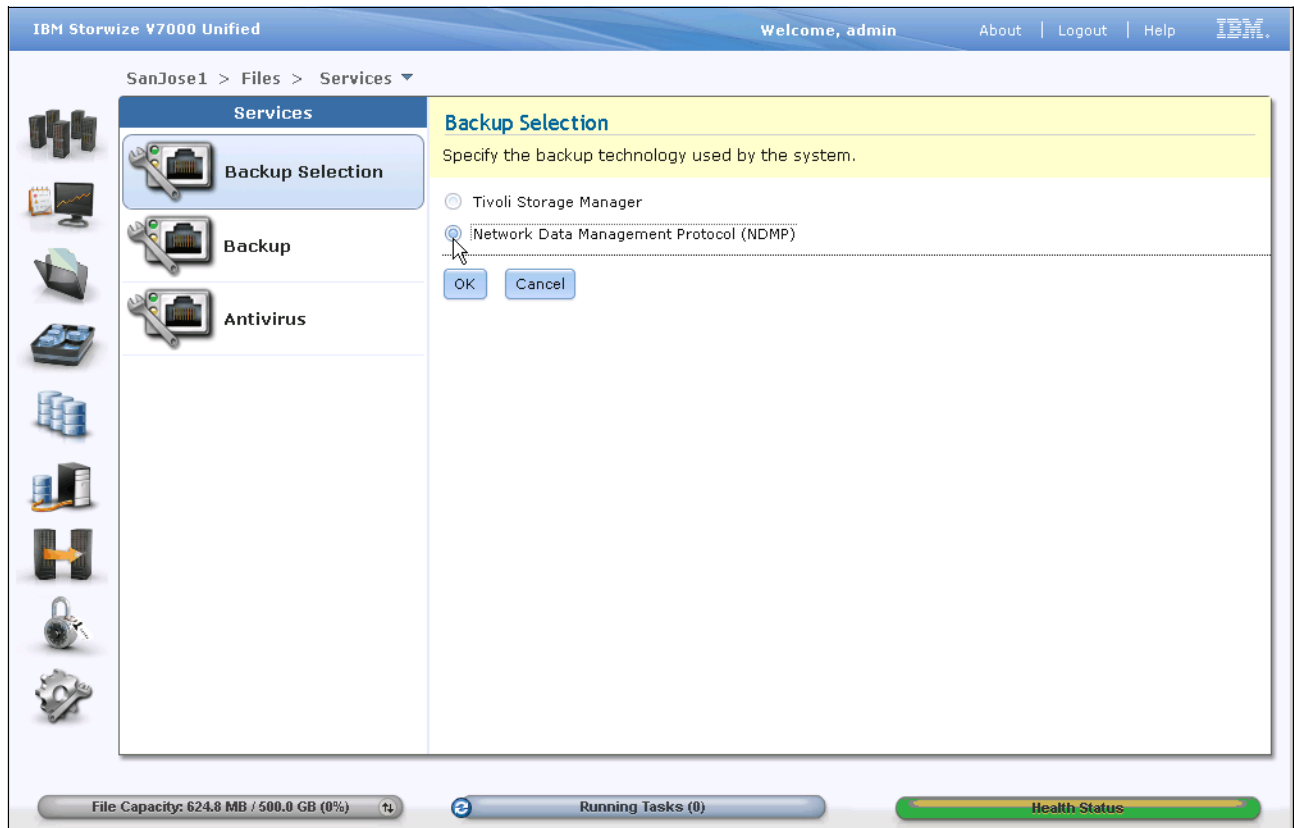
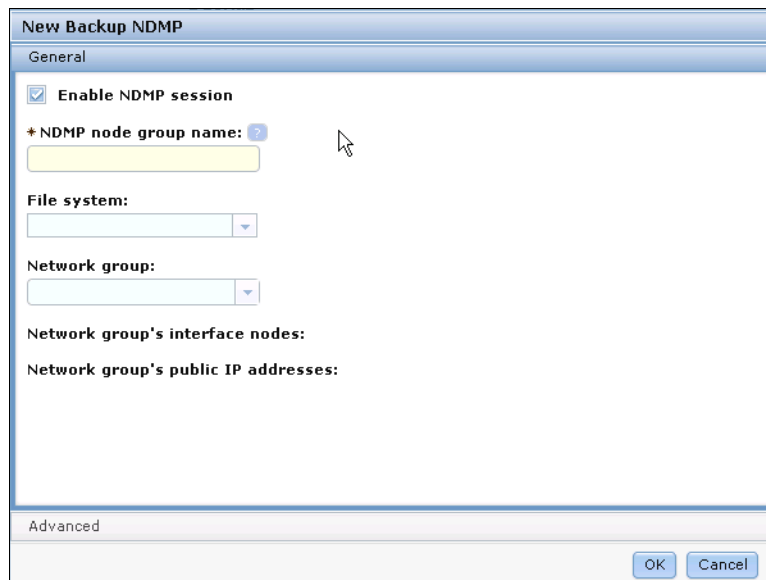


Figure 14-6 Backup selection - NDMP

Next click on the Backup icon to display the backup management window. Click on “New Backup NDMP” to create the network group. Ensure the enable button is ticked and enter a name for the group. Now select the file systems that will be included in this group and the default network group as shown in Figure 14-7 on page 232. The CLI command to create the group is:

```
cfgndmp ndmpg1 --create
```

“ndmpg1” is the name of the group in our example.



**New Backup NDMP**

General

☒ **Enable NDMP session**

\* **NDMP node group name:** ?

**File system:**

**Network group:**

**Network group's interface nodes:**

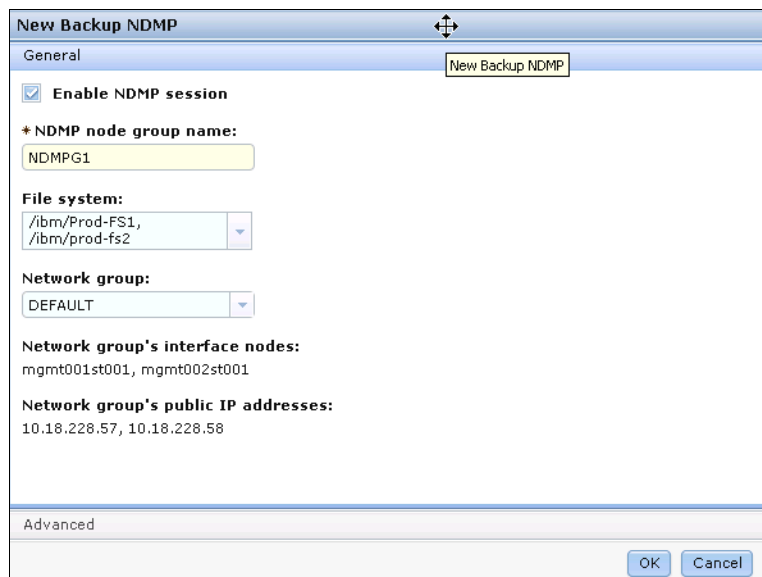
**Network group's public IP addresses:**

Advanced

OK Cancel

Figure 14-7 New group

Now click on the advanced tab to show the remaining configuration parameters. Set these parameters as detailed in the following table and shown in Figure 14-8 and Figure 14-9. Note that most of these should be left as default.



**New Backup NDMP**

General

☒ **Enable NDMP session**

\* **NDMP node group name:**

NDMPG1

**File system:**

/ibm/Prod-FS1,  
/ibm/prod-fs2

**Network group:**

DEFAULT

**Network group's interface nodes:**

mgmt001st001, mgmt002st001

**Network group's public IP addresses:**

10.18.228.57, 10.18.228.58

Advanced

OK Cancel

Figure 14-8 NDMP backup setting

**New Backup NDMP**

General

Advanced

**Protocol version:**  
Ver. 4

**Port:**  
10000

**Data-transfer IP address:**  
10.18.228.0/24 View...

**User name:**  
ibmuser

☒ **Prefetch activated**

**Prefetch application limit:**  
4

**Authorize NDMP client:**  
[Empty field]

**Data-transfer port range:**  
From: 1025 To: 65535

**TCP window size:**  
160

**Log file level:**  
0

**Password:**  
\*\*\*\*\*

**Prefetch thread limit:**  
100

OK Cancel

Figure 14-9 NDMP backup settings - advanced

Once complete, click OK to create the group. A popup will show progress, close when complete.

In Table 14-1 we show the NDMP group configuration.

Table 14-1 NDMP group configuration

Definition	Default	CLI command	Comment
NETWORK_GROUP_ATTACHED	""	cfndmp ndmpg1 --networkGroup <xxx>	The group of nodes to be attached with the NDMP group <optional>.
DEFAULT_PROTOCOL_VERSION	4	cfndmp ndmpg1 --protocol 4	Don't Change this setting
AUTHORIZED_NDMP_CLIENTS	""	cfndmp ndmpg1 --limitDMA 192.168.0.3 cfndmp ndmpg1 --freeDMA	Default = null, no restrictions
NDMP_PORT	1000	cfndmp ndmpg1 --dmaPort 10000	
DATA_TX_PORT_RANGE		cfndmp ndmpg1 --dataTransferPortRange 10020-10025	
DATA_TX_IP_ADDRS	""	cfndmp ndmpg1 --limitDataIP 17.0.0.0/24 cfndmp ndmpg1 --freeDataIP	Default = null, no restrictions
NDMP_TCP_WND_SZ	160	cfndmp ndmpg1 --tcpSize 160	
LOG_FILE_TRACE_LEVEL	0	cfndmp ndmpg1 --logLevel 3	
NDMP_USER_NAME	ndmp	cfndmp ndmpg1 --userCredentials ibmuser%mypass%mypass	Specifies username of "ibmuser" and password of "mypass" (repeated)
NDMP_PASSWORD	ndmp		

Definition	Default	CLI command	Comment
FILESYSTEM_PATHS		cfgndmp ndmpg1 --addPaths /ibm/gpfs1  cfgndmp ndmpg1 --removePaths /ibm/gpfs1	Define the paths to each file system that will be included in this backup group.
ENABLE_NEW_SESSIONS	allow	cfgndmp ndmpg1 --allowNewSessions  cfgndmp ndmpg1 --denyNewSessions	
<b>Prefetch settings</b>			
		cfgndmpprefetch ndmpg1 --activate	Enable prefetch if desired. NDMP must be deactivated to change prefetch status.
PF_APP_LIMIT	4	cfgndmpprefetch ndmpg1 --applimit 4	Maximum sessions using prefetch at one time. (1-10) NDMP must be deactivated to change prefetch status.
PF_NUM_THREADS	100	cfgndmpprefetch ndmpg1 --numThreads 180	Threads per node to be used for prefetching. (50-180) NDMP must be deactivated to change prefetch status.

This will return to the backup sessions screen and now there is a line entry on the screen for this new group. Highlight the line, then use the actions pull-down or right click to manage the group as seen in Figure 14-10.

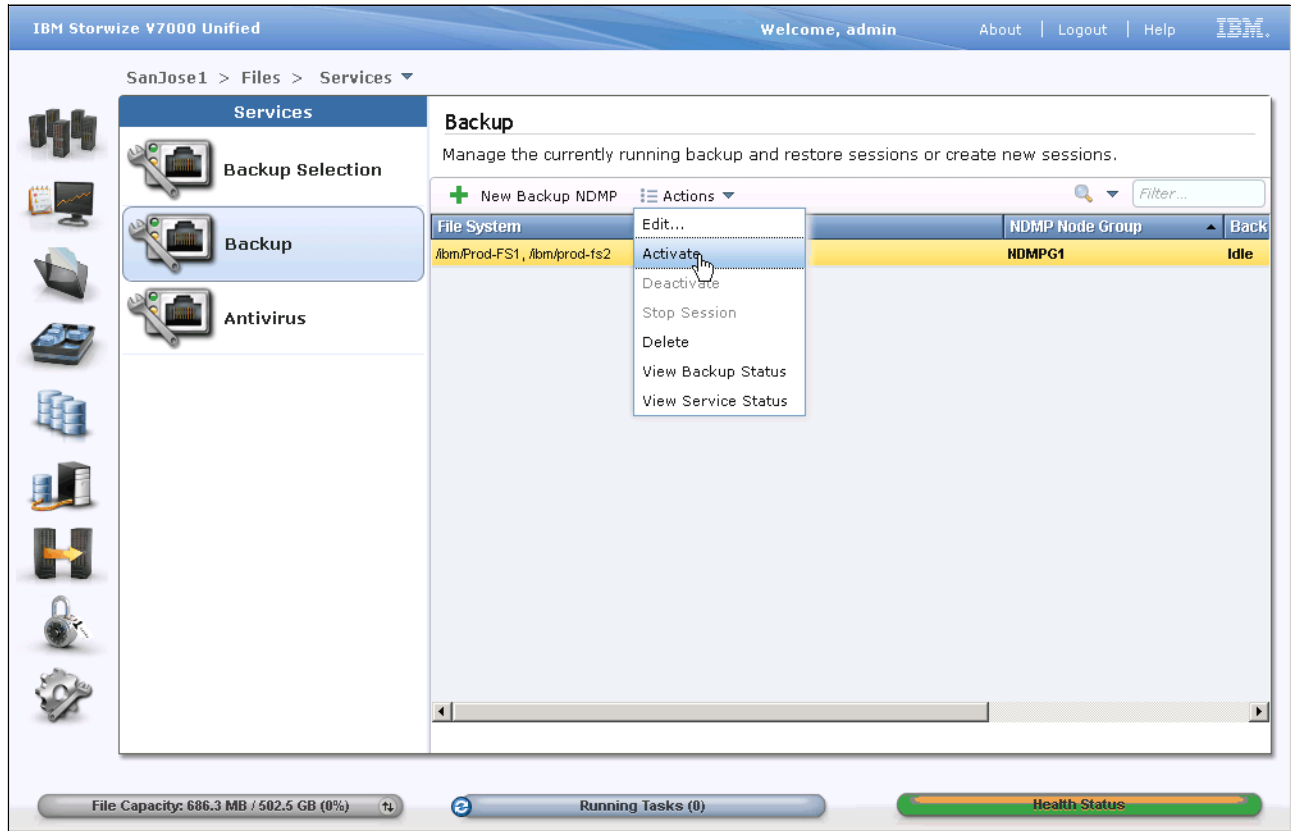


Figure 14-10 Manage NDMP backups

You can edit the groups parameters, this will take you back through the same panels as the create. Activate and deactivate the NDMP group. Also stop a session. Using the view backup status, you can show all currently running sessions. The view service status option will show you the status of the NDMP service on each node.

**Caution:** Do not change NDMP settings or prefetch configuration while a backup is being performed. This will cause current backups to fail when NDMP then re-cycles.

The CLI command to activate the group is:

```
cfgndmp ndmpg1 --activateSnapshot
```

**Caution:** Activate NDMP well before a backup is scheduled to run. Activating NDMP immediately before a backup may affect the backup while NDMP does startup housekeeping. If snapshots have expired and require deletion, this may prevent a new snapshot being created.

## Prefetching

This feature is designed to greatly improve backup performance of NDMP with small files. The process predicts the file sequence for the NDMP group being backed up and reads the files into the nodes cache. This gives the effect of a very high read cache hit rate.

Prefetching can be turned on for up to 10 NDMP sessions and the threads setting is provided to allow tuning of the backup workload across the nodes based on the nodes cache memory and performance. Prefetch is designed to work with very small files, under 1MB. Files over

1MB will not be prefetched, so turning on this feature on groups with predominantly large files will have little effect.

Prefetching is disabled by default. Activating causes any currently running NDMP sessions to cease.

### Taking backups

Once the NDMP sessions are configured and NDMP is active, you are able to perform the backups. This is done using your NDMP compliant DMA on an external server. Configuration and management of the DMA is beyond the intended scope of this book.

More information about Managing NDMP can be found in the IBM Storwize V7000 Unified 1.3 Information Center:

<http://ibm.biz/BdxFDG>

## 14.4 Data recovery

Although we do not show a data recovery scenario, we describe the steps required to perform a recovery.

For more information on recovery procedures in general, including:

- ▶ User ID and system access
- ▶ File module-related issues
- ▶ Control enclosure-related issues
- ▶ Restoring data
- ▶ Upgrade recovery

Refer to IBM Storwize V7000 1.3 Information Center:

<http://ibm.biz/BdxFD2>

### 14.4.1 TSM

As the Storwize V7000 Unified software contains a TSM client, restore of a file system must be performed on the cluster to recover from your TSM server. As the specific configuration and implementation of TSM differs from site to site, we will describe the process generically. You will need to consult the relevant manuals and your TSM administrator before proceeding with a restore operation.

The basic steps are as follows.

Ensure there are no backups or restores currently running by issuing the command:

**lsbackups**

4. You can restore a single file system or part of using the command:

**startrestore**

A file pattern needs to be given. Using wildcards, the whole or part of a filesystem can be restored, even to a specific file. It is also possible to filter to a timestamp which will restore files as they were at that time based on the backups available. Overwrite can be enabled



or disabled. Review the command description in the information center and the examples given before using this command.

5. The **1sbackupfs** command can again be used to monitor progress.

Refer to IBM Storwize V7000 1.3 Information Center for a full example:

<http://ibm.biz/BdxFDz>

## 14.4.2 Asynchronous data recovery

Recovering a file system with asynchronous replication requires that you configure and start a replication relationship from the target site to the source site.

For more information and an example refer to the IBM Storwize V7000 1.3 Information Center:

<http://ibm.biz/BdxFDq>

## 14.4.3 NDMP

Like the backup process, restoring data with NDMP is performed on the DMA server and is outside the intended scope of this book. This of course requires that the NDMP agent running on the cluster has been defined. In a running system requiring backup of a filesystem or part thereof, the agent will already be ready. In the case of a full rebuild, it may be necessary to define and configure the agent to proceed.

More information about Managing NDMP can be found in the IBM Storwize V7000 Unified 1.3 Information Center:

<http://ibm.biz/BdxFDG>





# Troubleshooting and Maintenance

In this chapter we will look at how errors and events are reported and logged, and what tools are available to analyze and recover from a problem. Also the procedures to follow, and the methodologies and tools provided with the Storwize V7000 Unified to enable IBM Support to assist.

We will also cover common maintenance procedures, including software upgrading and also address compression related recovery such as recovering from a volume going offline.

## 15.1 Maintenance philosophy

Many events or problems that occur in your Storwize V7000 Unified environment will require little or no user action. This is because the system employs a “self healing” philosophy so that where possible, automatic recovery is triggered for many events. Also, when the cause of a problem has abated, recovery procedures automatically run and the event warning will be closed, such as when a storage or host path is lost and then later recovered.

When a problem occurs, an entry describing the problem, using errors codes, is entered into the event log. If the event required action and cannot be automatically resolved it is marked as “unfixed”. Only unfixed events require action. Alerting can be configured to send an e-mail or SNMP alert and this can be filtered by the type of event. Recovery actions are taken by running the built in guided maintenance procedures which the user launches from the event display.

If the problem cannot be resolved using guided procedures, the user will be prompted to call IBM for support. This is achieved by calling IBM service using the local procedure. Depending on how you have set up your call home, the Storwize V7000 Unified may also have sent an alert to IBM and IBM Support may even call you first. The primary purpose of this “call home” function is to get information about a possible problem to IBM in a timely manner and serves as a backup for problem notification, but it remains the clients responsibility to be aware of the status and health of their systems and raise a call with IBM if service is required, unless special arrangements have been made beyond the standard maintenance.

All user and technical manuals are incorporated in a single interactive repository called the “Information Center”. This online web based system is discussed in 15.4, “Information Center” on page 257

Support of the Storwize V7000 Unified is primarily using IBM's Remote Support model, where the first contact is with IBM's Remote Technical Support (RTS) desks in each region. This proven process gives the fastest response and immediate engagement of specialized support. If required, higher levels of support can be engaged quickly. An IBM Support Services Representative (SSR) will only be dispatched to the client site if there are actions needed that require an IBM representative.

Most of the parts in the Storwize V7000 Unified are designated “Client Replace”. Should such a part need replacement, then the Client Replaceable Unit (CRU) will be couriered to the site and instruction given on the replacement process by the RTS specialist.

If there is a requirement for IBM Support to observe behavior or login to perform complex procedures, then the specialist will use IBM's Assist-on-site product. When configured, this will connect directly to the Storwize V7000 Unified, or a client workstation. This tool is web based and using secure authorization provides a simple remote KVM function to the cluster or to a workstation in the client's environment. It only uses normal html based ports that are typically not blocked and also allows selected authorized specialists within IBM to jointly access the session if needed. The client maintains control of their workstation and can observe all activity. This tool greatly speeds up resolution time by getting a specialist onto the system very quickly.

## 15.2 Event logs

Logging in the Storwize V7000 Unified is done in the event log. An event could be critical failures of a component that affect the data access or could be a record of a minor

configuration change. The two major components of the Storwize V7000 Unified both maintain an event log, but these logs are stored and handled independently. The format of the data in the log and the tools available to interrogate and act on the log entries also differs significantly, so they will be discussed separately.

All log entries are tagged with a status indicating the impact or severity of the event. This quickly highlights the importance of an event and allows for sorting and filtering for display.

A large number of trace and debug logs are also recorded at a low level within both components. Many of these logs wrap and must be collected close to a problem event. The logs are unformatted and not visible to the user. Many are collected by the data collection process “Download Support Package”. If additional logs need to be collected or generated, IBM technical support will provide instructions or connect to the device.

There are two places to look for these Events. The first item to look at is the lower right bar which shows the overall “Health Status” of the system as green, yellow, or red. As seen in Figure 15-1, there are tabs for “File” and “Block” events. For “File” type events you also need to check the status for each File Module using Monitoring Events → System Details → <select a File Module> → Status as seen in Figure 15-2.

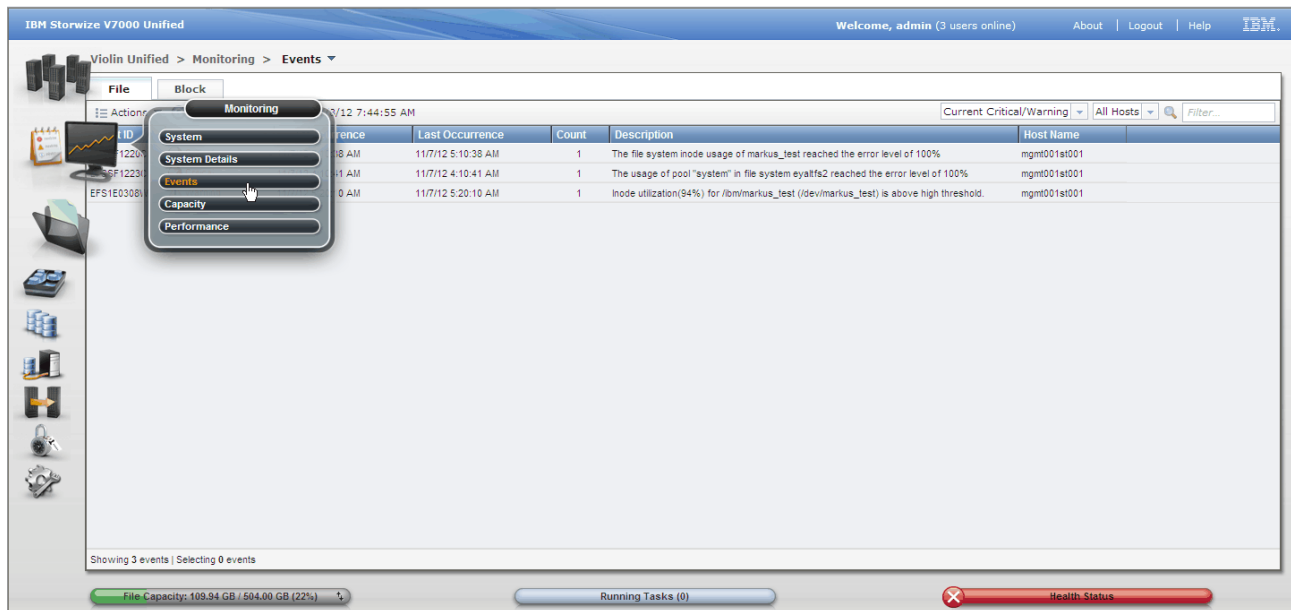


Figure 15-1 Block and File Events

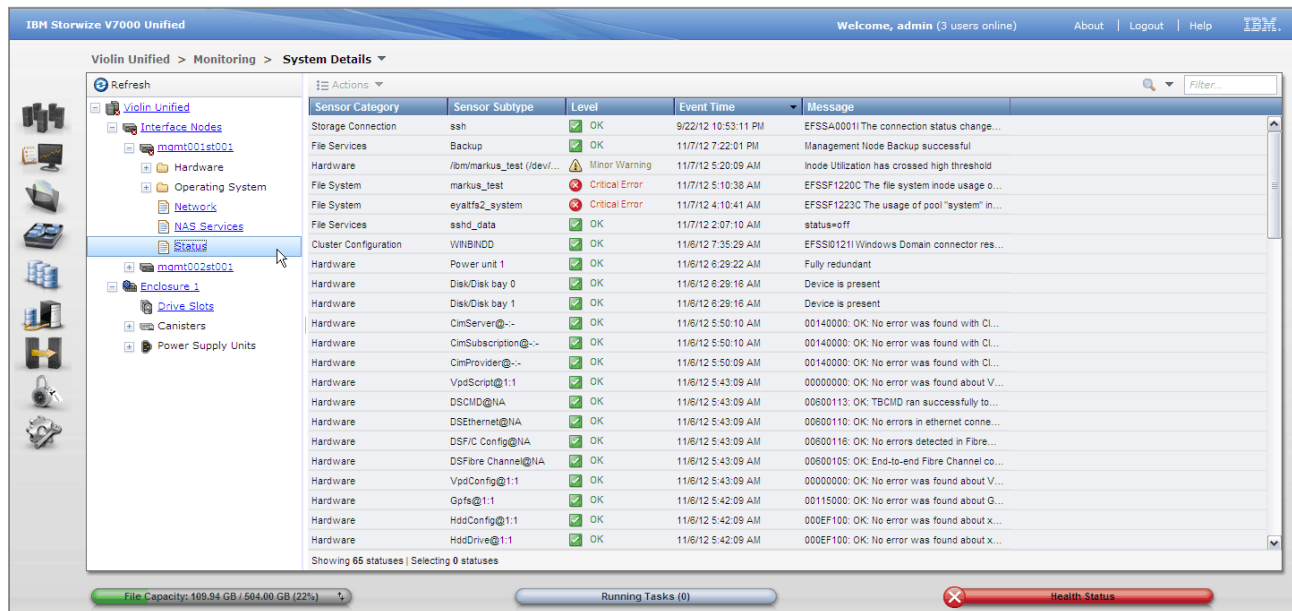


Figure 15-2 File Module Events

## 15.2.1 Storwize V7000 Storage Controller Event log (block)

To display the events, navigate to Monitoring → Events. Click on the Block tab. This will bring up the screen as shown in Figure 15-3.

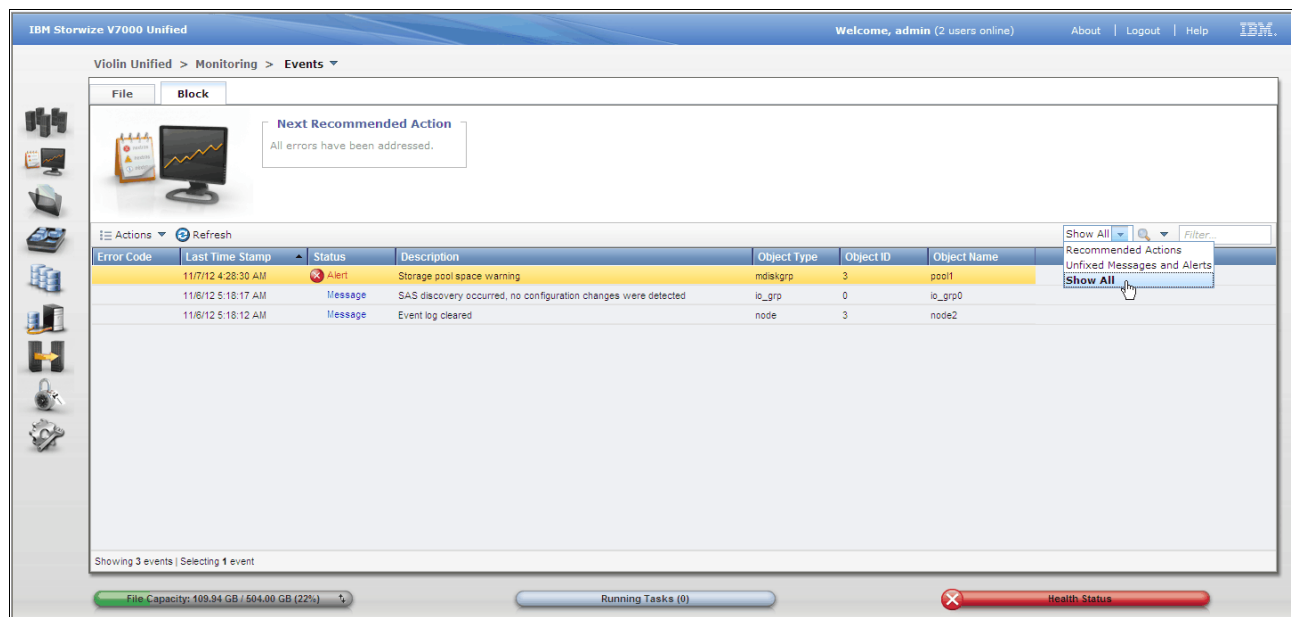


Figure 15-3 Event logs - Block

There are some options to make this screen more meaningful. We can sort on the event status. The pull down gives a choice of Recommended actions, Unfixed Messages and alerts, and Show All.

Recommended Actions

Events that the need actions performed against them, usually by running guided maintenance.

Unfixed Messages and Alerts      This list all events that are marked as “unfixed”.

Show All      List all entries in the event log.

A filter option is also provided, although it would only be needed if the log has become cluttered with a high volume of events. Next to the magnifying glass is a pull down arrow. Select the field to filter on. Then in the filter box, enter the value of your filter.

## Details

To display the event details, use the actions to select properties. Look at this example of a minor event in Figure 15-4.

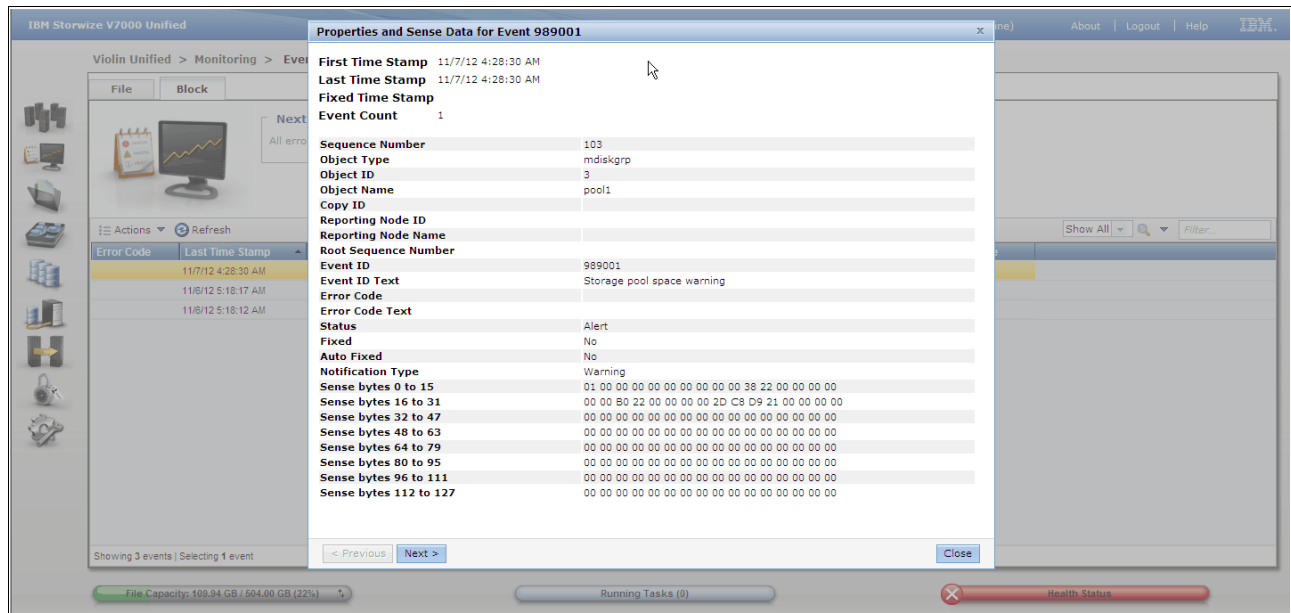


Figure 15-4 Block event - properties

Two timestamps are logged, being the first occurrence and the last. These are read in conjunction with the event count. If the event count is 1, then the timestamps will be the same, so the event occurred only once. Should the exact same event occur more than once, then the count will increase and the last timestamp will show when the last one occurred.

The sequence number uniquely identifies this event. This number increments for every new event for the life of the cluster.

The object type, ID and name identify the resource that this event refers to.

The reporting node is the node that detected and logged this event.

If this event is related to another event, then the root sequence number will be the event that triggered this one. These event would therefore be considered together in problem determination.

The event ID is a number that uniquely relates to the event being logged and can be searched on in the manuals or support knowledge bases to obtain more detail on the event. Associated with the event ID may be some text to give a verbose description of the event.

An error code may also be displayed and may have associated text. This identifies the error that caused this event to be logged. All events requiring guided maintenance will have an error code.

Status gives the type of event and its severity.

The notification type gives the priority of the event and the level of notification. This is used by the alerting tasks in the cluster to determine based on the configured rules what alerts will be generated.

The fixed field is very important. Unfixed events can cause recovery routines to not be run if locks exist on processes as a result of an unfixed problem. Unfixed problems require action and there should be no unfixed events in the log unless they are for a current issue.

Additional information may also be supplied including sense data which will vary based on the event.

**Important:** Regularly check for unfixed problems in the event log. While high priority events will cause alerts based on your configuration, it is possible to miss important events. Ensure that all unfixed problems are actioned. If thin provisioning or compression is used then it is very important that any out of space warnings be corrected immediately before it causes issues with volumes going offline due to lack of resources.

## Actions

Highlight a log entry by left clicking. It will backlight yellow. Then either right click or use the Actions pull down to display the choices. The actions are as follows:

Run Fix Procedures	If the event is an error or qualifies for guided maintenance (and therefore has an error code logged), then this option will be available, otherwise it is greyed out. This is the correct action for any unfixed problem needing maintenance. This action will launch Guided Maintenance Procedures as detailed in , “Properties This gives additional information on the event. Also in this window is a hyperlink to the Storwize V7000 Unified Information Center which will take you directly to a description of the error code as seen in Figure 15-8 on page 249.”
Mark as Fixed	This option changes the fix status of an event to Fixed = yes. This is useful when multiple events have been logged as a result of a problem and the problem is being fixed using another event. It is also useful when you are confident the cause has been resolved and there is no need to run maintenance, but do so with caution as this action may bypass cleanup routines. If a problem is informational and fix procedures are not available, then use this option to remove the event from the unfixed list.
Mark as Unfixed	If an event was incorrectly marked as fixed, this will mark it as Unfixed.
Filter by Date	Gives a choice of start and end dates to customise the display.
Show entries with...	Expands to give a choice of minutes, hours or days previous to limit the displayed events.
Reset Date Filter	Resets the above choice.
Clear Log	Use the Clear Log with great caution. This will delete the event contents from the log which may compromise IBM Support’s ability to diagnose problems. There is no need to clear the log, there is plenty of space to store the data and the cluster performs its own housekeeping periodically to purge old unimportant entries. The refresh button refreshes the display.



Properties                      Displays the details of the event and the control fields as discussed in , “Details” above.

**Suggestion:** Do not clear the event log, there is no practical gain in doing so, other than in special cases for site security reasons.

## 15.2.2 V7000 Unified File Module Event log file

The “*File*” event log is cluster wide and is common for both of the file modules. To display the events, navigate to Monitoring → Events and click on the File tab to see it as shown in Figure 15-5.

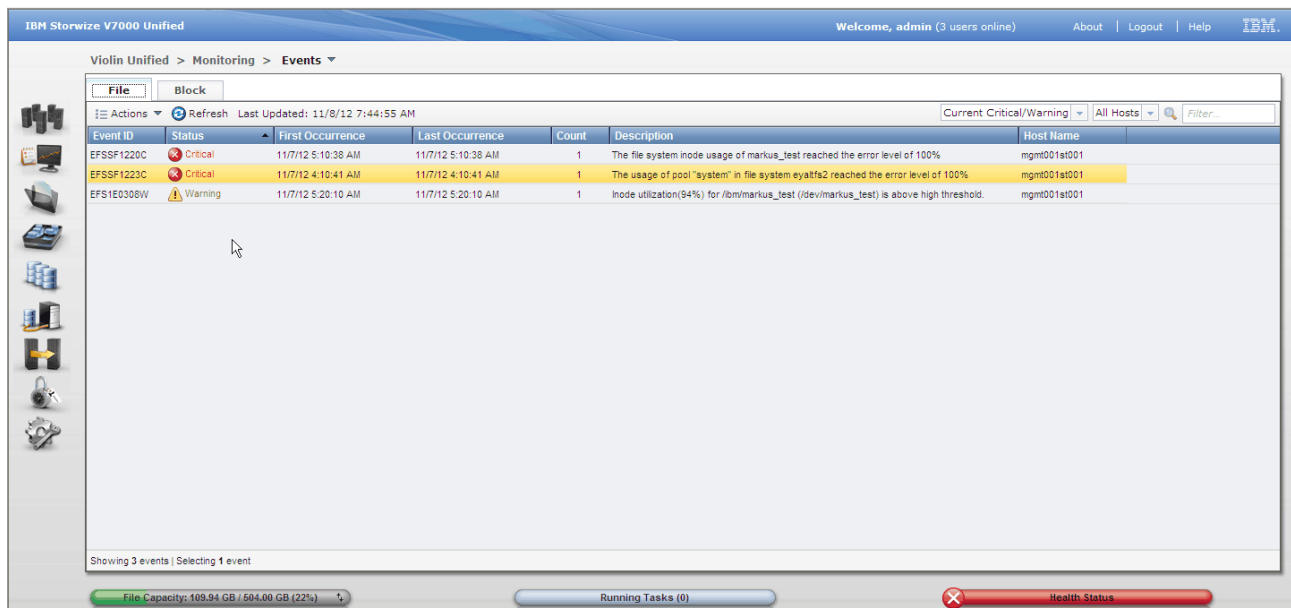


Figure 15-5 File Event

Like the block display, there are some options to make this screen more meaningful. We can sort on the event status and currency.

Current Critical/Warning                      Unresolved events that are status critical or warning.

Critical/Warning                                All events in the log that are status critical or warning.

Show All    List all entries in the event log.

Additionally you can filter the display to show the events generated by either or both file modules

A filter option is also provided that allows you to further reduce the display by including a search word or phrase. Simply type the word and press the enter key. A reset option clears the filter.

### Actions

Unlike the block event log, there are no actions to fix or clear a specific log, this display is for viewing only. Repairing events is covered in , “Properties This gives additional information on the event. Also in this window is a hyperlink to the Storwize V7000 Unified Information Center which will take you directly to a description of the error code as seen in Figure 15-8 on page 249.” check Figure 15-8 on page 249. Highlight a log entry by left clicking. It will

backlight yellow. Then either right click or use the Actions pull down to display the choices. The actions are as follows:

Filter by Date	Gives a choice of start and end dates to customise the display.
Show entries within...	Expands to give a choice of minutes, hours or days previous to limit the displayed events.
Reset Date Filter	Resets the above choice.
Clear log	Clears all entries from the event log. Again, use with caution.
Properties	This gives additional information on the event. Also in this window is a hyperlink to the Storwize V7000 Unified Information Center which will take you directly to a description of the error code as seen in Figure 15-8 on page 249.

In the topics that follow we discuss guided maintenance and the procedures therein.

### 15.2.3 Block

Any event that needs maintenance action has the option to run Guided Maintenance enabled and is set to unfixed. To display unfixed problems use the procedure detailed in 15.2.1, “Storwize V7000 Storage Controller Event log (block)” on page 242 above.

It is important to complete maintenance on all unfixed events in the cluster. Unfixed events can prevent maintenance routines from running and create contention for maintenance resources which can prevent problems from auto recovering. Make regular checks of the event log to ensure that all unfixed events are actioned. Of particular importance are “out of space” conditions when compression is being used and which need to be addressed before physical storage is exhausted.

There are several ways of actioning an unfixed event:

1. Use the *Run Fix Procedures* action to perform guided maintenance on the event. This is the preferred option.
2. Mark the event as fixed. This would be done if the event was informational and fix procedures were not enabled. In this case, the event is intended to be read in conjunction with another event or is purely for your information and is marked as unfixed to gain your awareness.

Events can also be marked as fixed when there are a number of similar events and the problem has been resolved by running maintenance on another entry. Use this option with caution as this bypasses the Guided Maintenance routines which includes cleanup and discovery processes, and this may leave a resource in a degraded or unusable state.

3. The Storwize V7000 uses a self healing philosophy where possible. Many events, known to be transient, will trigger an autorun of the maintenance procedures and will recover. Subsequent events that clearly negate an earlier one will automatically mark the earlier event as fixed. For example, a node offline message will remain in the log, but get marked as fixed when a Node Online message for the same node is logged 10 minutes later.

#### Run Fix Procedure

To display the block event log Select Recommended Actions from the filter pull down. If there are any events requiring guided maintenance to be run, they will be listed in the window. The software will have also made a suggestion on which is the first event to work on, which can be seen in the section above the event list. This is generally based on the lowest numbered error code. The rule of thumb is to start with the lowest number. The shortcut button in this window will launch you straight into the guided maintenance of the recommended event.

An example of this is shown in Figure 15-6.

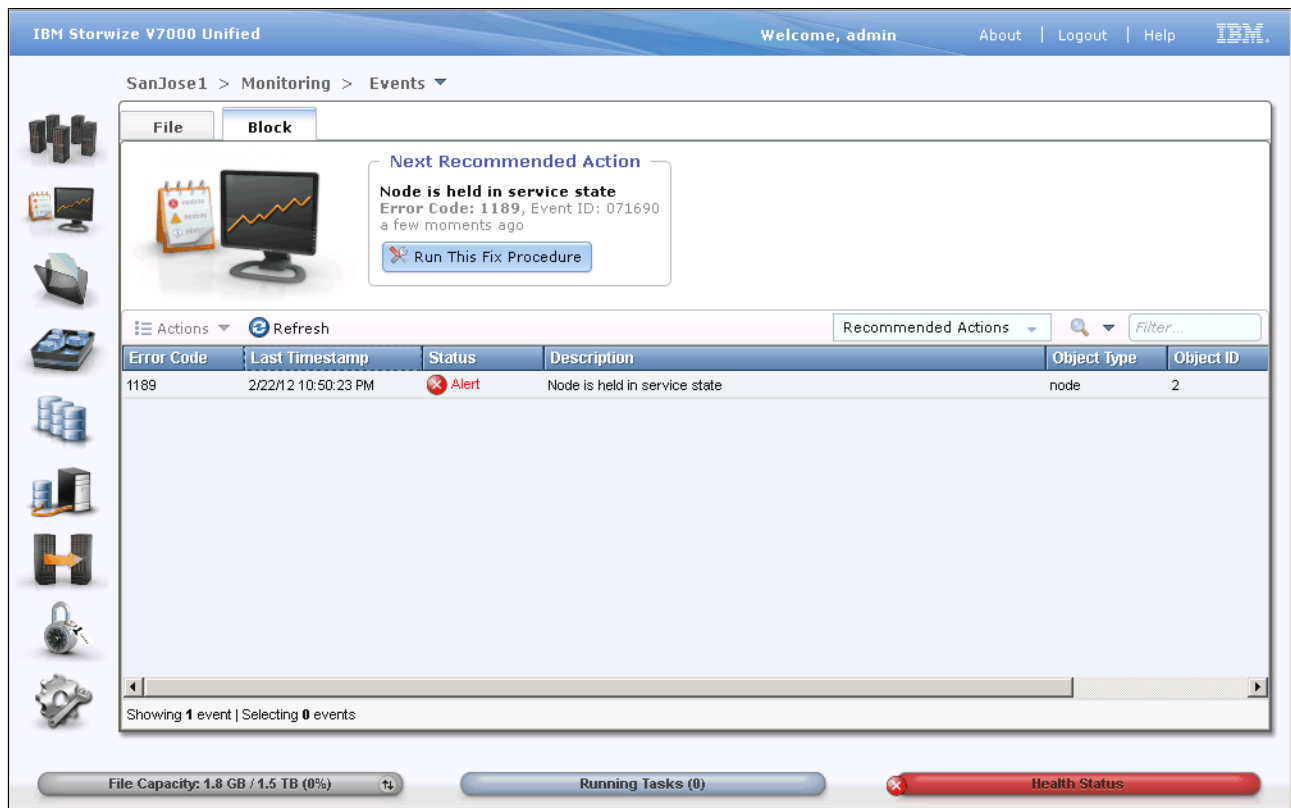


Figure 15-6 Block events - error

To manually select the event, highlight the event in the list with a left click. Then use the action pull down or right click to see the action list and select *Run Fix Procedure*.

This will launch the guided maintenance. The panels that display are unique to the error and will vary from a single window giving information on how to resolve the error to a series of questions requiring responses and actions to perform. For a part failure, the guided maintenance will step through diagnosing and identifying the failed part, preparing the cluster for its isolation, powering off components if required, guiding you through replacement and testing. The procedure will confirm that the original error is now cleared and put the appropriate resources online. Finally it will mark the event as fixed, closing the problem.

For most errors the panels are quite interactive and will ask for confirmation of status or configuration. Be careful answering these, ensure that the requested state is definitely correct before confirming. For example, for a pathing error, you might be asked if all zoning and pathing is currently operational and as intended. If you answer yes, even though one path has failed, then the process assumes that previously missing paths are no longer in the configuration and will not look for them again.

In our example as shown in Figure 15-7, the guided maintenance will not take any action on the status as it does not know the reason why it is set. Service mode would typically be set manually by the user and therefore is set for a reason. It can also be set by a failed node, and therefore there would be an accompanying event of higher importance. The panel displayed gives detail on what the state is and where to go to clear it. Note that for this particular error, the process leaves the event marked as unfixed, advising that it will automatically be marked fixed, when the state clears.

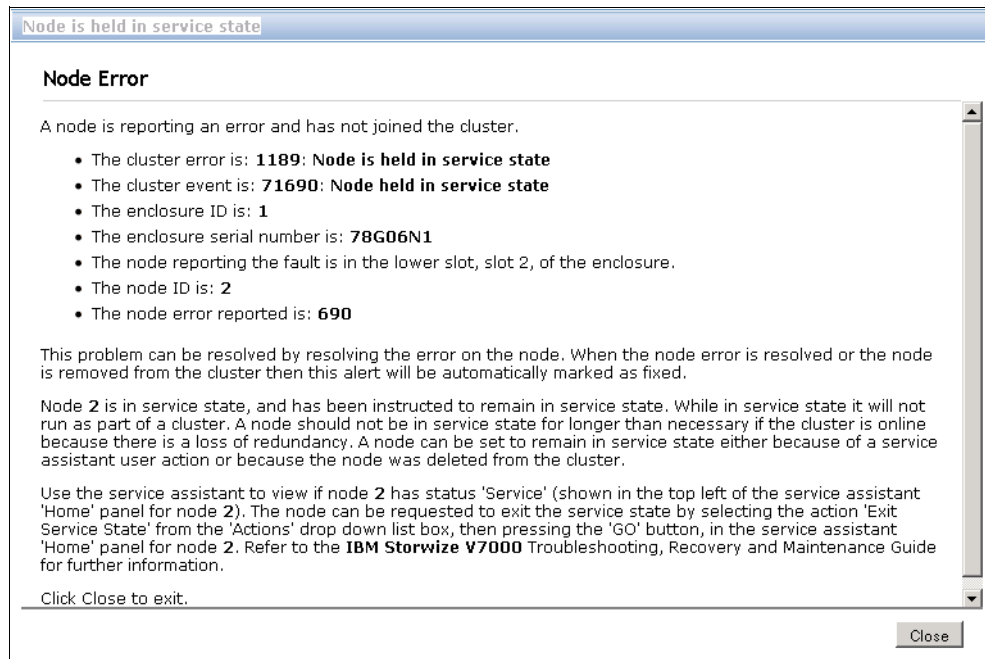


Figure 15-7 Guided maintenance - 1189 error

Note in this example the three distinctly different codes:

Event code	This is the reference code (5 digits) of the event and uniquely identifies the type of event that has occurred and the verbose detail given.
Error code	An error code (4 digits) is only posted if an error has occurred and this code is used by the guided maintenance and support personal to repair the fault.
Node error code	This code (3 digits) is related to the node module itself, not the cluster software.

## 15.2.4 File

The code running in the File Modules does not incorporate a guided maintenance system. All maintenance and recovery procedures are guided from the Information Center. Using the error code as a starting point and search the code in the information center. The Information Center will give appropriate actions for each error.

Events in the system are logged in the event log as we discussed earlier. By viewing the event log and using the action pull down (or right click) to display the *properties*, we can see the event details shown in this example in Figure 15-8. This gives additional information on the event. Also in this window is a hyperlink to the Storwize V7000 Unified Information Center which will take you directly to a description of the error code.

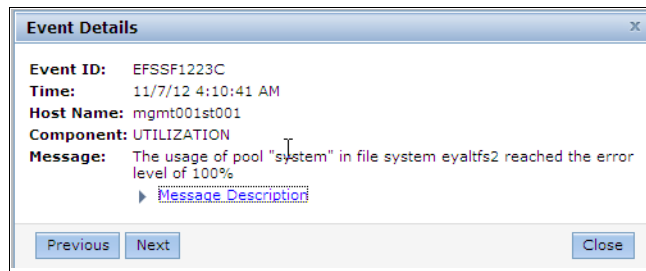


Figure 15-8 File event - details

In our example, this will launch page as shown in Figure 15-9.

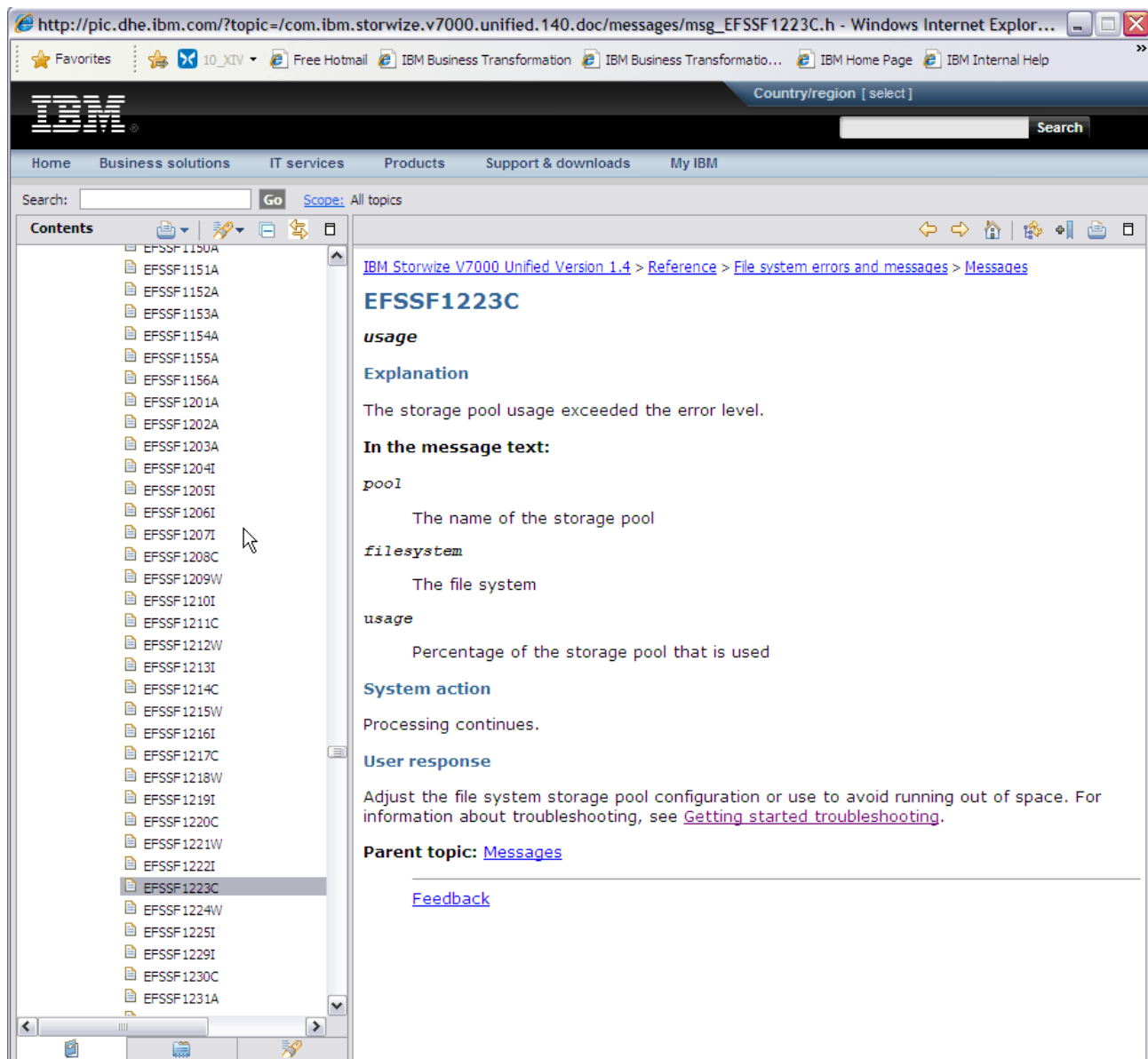


Figure 15-9 Information Center

Use the procedures outlined in the Information Center to resolve the problem. Once the issue has been resolved, then you need to “Mark Event as resolved” using the following procedure.

For “File” type events you need to check the status for each File Module using Monitoring → System Details → <select a File Module> → Status. This will be the screen that is used to mark the error as being fixed as shown in Figure 15-10.

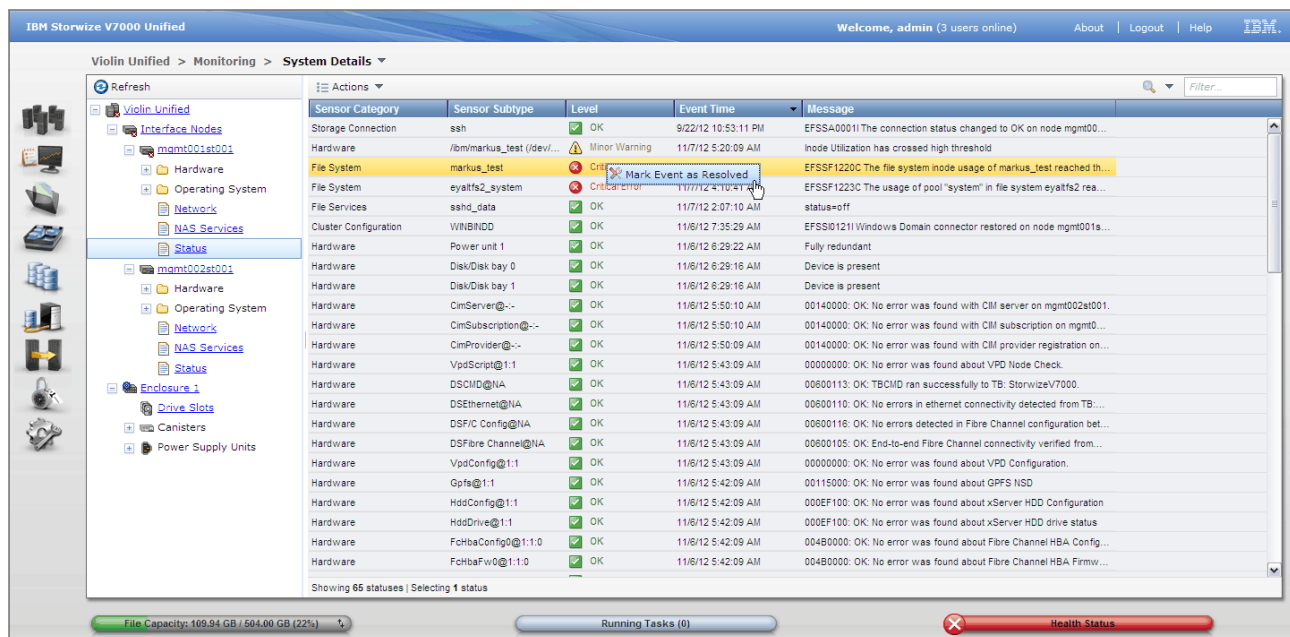


Figure 15-10 Mark Event as resolved

A specific example for resolving a volume out of space condition and marking the event as resolved is covered in the 15.2.5, “Working with compressed volumes out of space conditions”.

## 15.2.5 Working with compressed volumes out of space conditions

The most important utilization metric to monitor is the utilization of the physical space currently used in a storage pool. This metric relates to the actual physical storage capacity already used for storing compressed data written to the pool. It is important to make sure that physical allocation does not go over the suggested threshold (the default is set to 80%). In order to reduce the utilization of the used space, the storage pool size needs to be increased. Adding physical capacity to the pool reduces its utilization. Whenever a threshold is passed and an alert is generated, the system will point you to the procedures outlined in the Information Center to resolve the problem. If a corrective action to reduce utilization is not performed before the storage pool reaches 100% utilization, NSDs can go offline, which can cause the filesystem to go offline.

**Note:** The steps that follow were verified during the publication of this book, however we suggest that the Information Center be used as it may contain updated steps to resolve volume “out of space” conditions.

The overall steps required to correct an unmounted file system as a result of a compressed volume (NSD) going offline is as follows:

1. Block Event. Run fix procedure for the NSD that went offline and follow the steps provided by the fix procedure for the block event. The NSD offline condition must be addressed before the filesystem can be mounted.
2. Start the NSD once space is available.

3. Determine if there are any stale NFS file handles.
4. Perform single node reboots to clear the stale NFS file handles.
5. Mount the filesystem

Proceed with the following detailed steps:

### Run Fix Procedure (Block)

To display the block event log Select Recommended Actions from the filter pull down. If there are any events requiring guided maintenance to be run, they will be listed in the window. The error for a volume copy offline due to insufficient space is “Error 1865”. The software will have also made a recommendation on which is the first event to work on, which can be seen in the section above the event list. This is generally based on the lowest numbered error code. The rule of thumb is to start with the lowest number. The shortcut button in this window will launch you straight into the guided maintenance of the recommended event.

An example of this is shown in Figure 15-11.

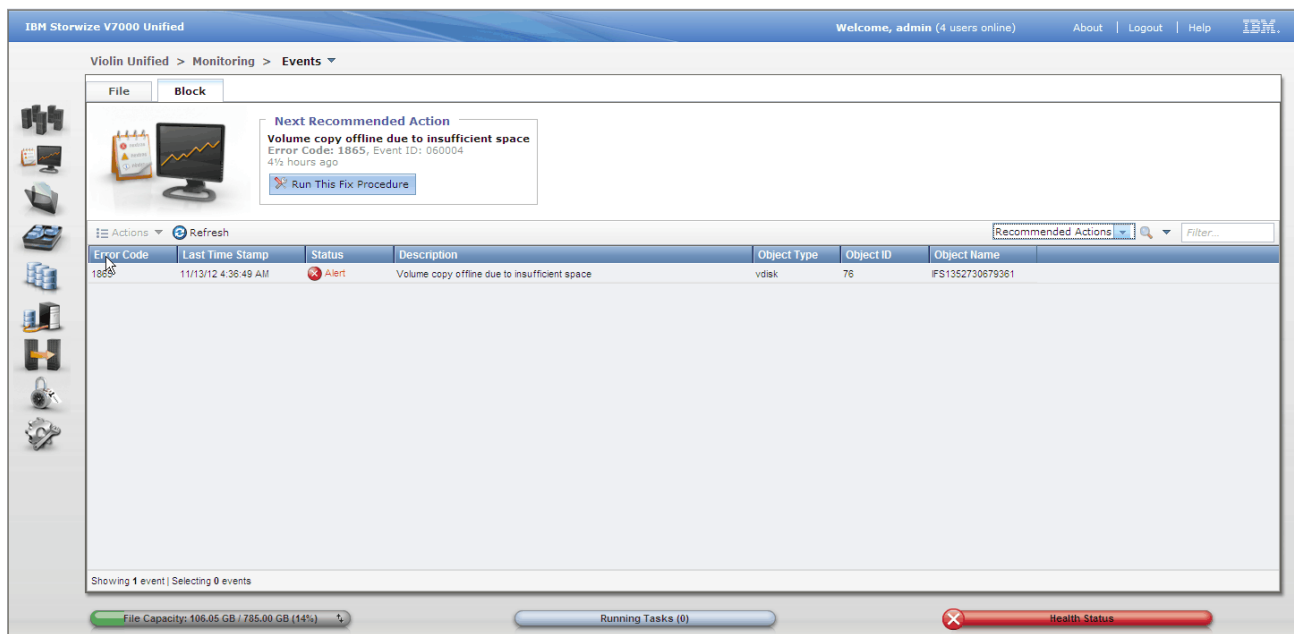


Figure 15-11 Run Fix Procedure for NSD offline

The guided maintenance procedure will provide guidance as to how to resolve the volume offline condition as seen in Figure 15-12 on page 252.



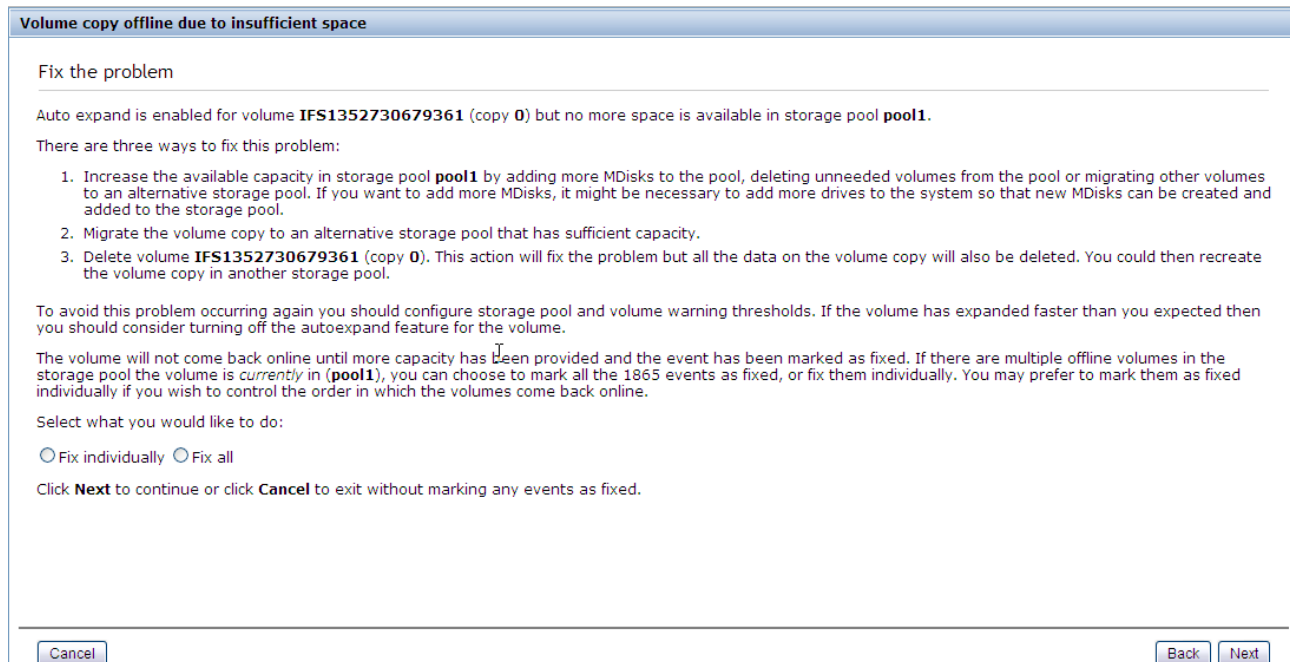


Figure 15-12 Fix the problem.

Once the issue that caused the NSD to go offline has been resolved, you have to specify that the problem has been fixed which will cause the NSD to come back online and the 1865 error will be cleared.

After the NSD is back online, the filesystem remains unmounted and now you must follow the steps that will be provided by the File Events.

The first step is to verify the state of the filesystem as shown in Figure 15-13 which will indicate that the filesystem is not mounted.

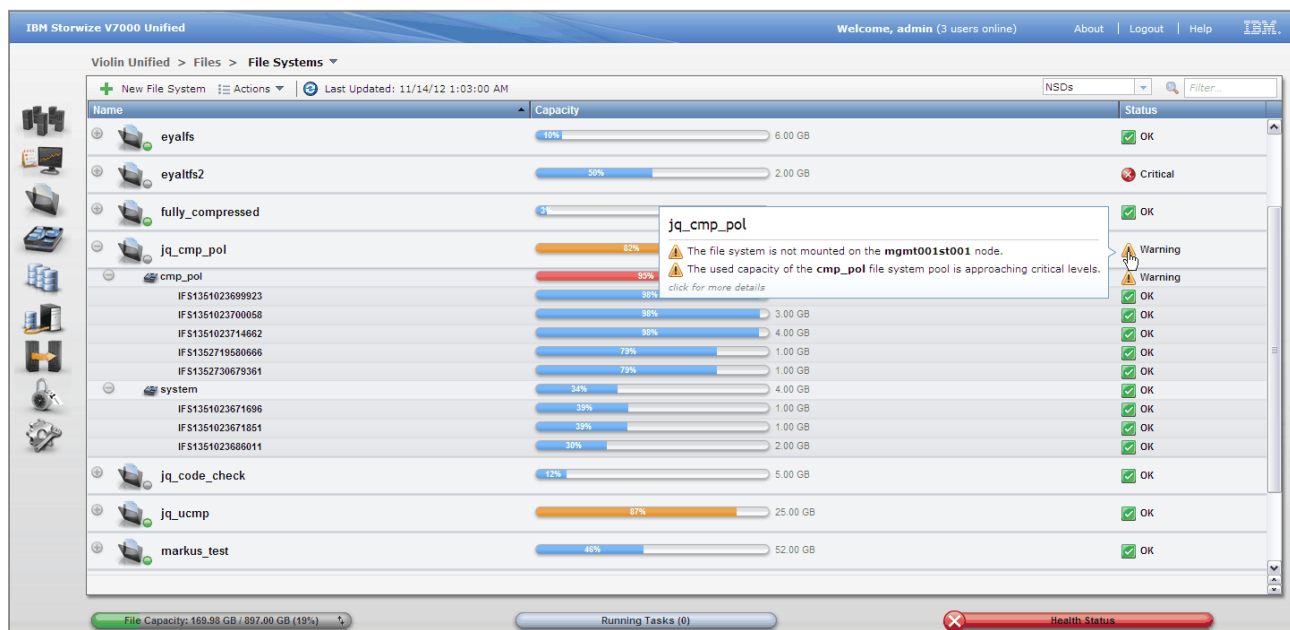


Figure 15-13 Filesystem not mounted



By viewing the File Event Log for the filesystem unmounted error and using the action pull down (or right click) to display the *properties*, we can see the event details shown in Figure 15-14. There will be a hyperlink “Message Description” provided that will connect to the Storwize V7000 Unified Information Center which will take you directly to a description of the error code and instructions to get the filesystem checked and remounted..

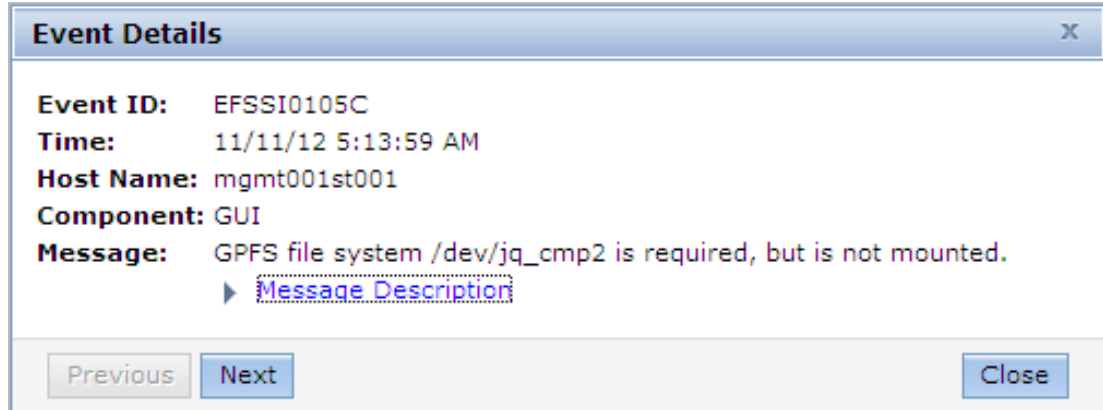


Figure 15-14 Filesystem not mounted

**Note:** The steps that follow were verified during the publication of this book, however we suggest that the Information Center be used as it may contain updated steps to resolve a filesystem unmounted condition.

The first page that the hyperlink will show is shown in Figure 15-15.

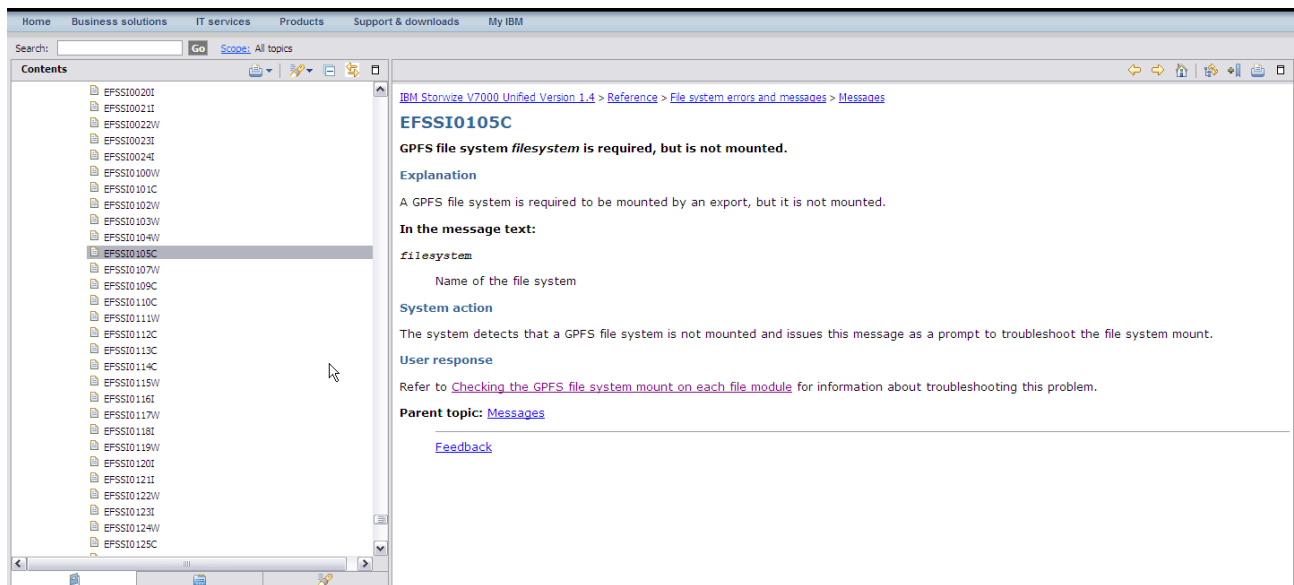


Figure 15-15 Event EFSSI0105C

Select the “User response” link to obtain details and steps needed to bring the filesystem online again. Make sure that the steps are performed in sequence and completed. The high level steps are shown in Figure 15-16 on page 254.

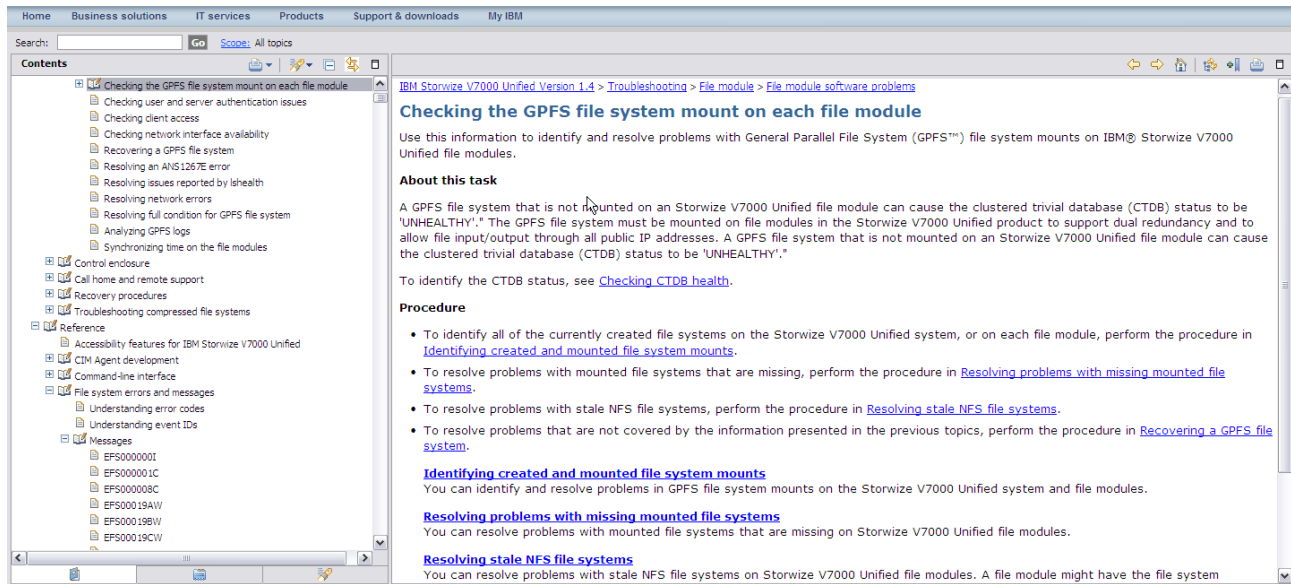


Figure 15-16 Checking the GPFS filesystem

## 15.3 Collect Support Package

**Note:** The workstation you are using to connect to the cluster will need to have internet access to be able to display the Information Center.

The three modules of the Storwize V7000 Unified each have their own data collection processes. Each file module collects its own log and configuration files and the storage control enclosure also collects its own snap file. The cluster triggers these collects and combines them into a single file so only the one operation is required.

This is triggered from the cluster GUI. Navigate to Settings → Support → Download logs.

This will present the download logs window as shown in Figure 15-17.

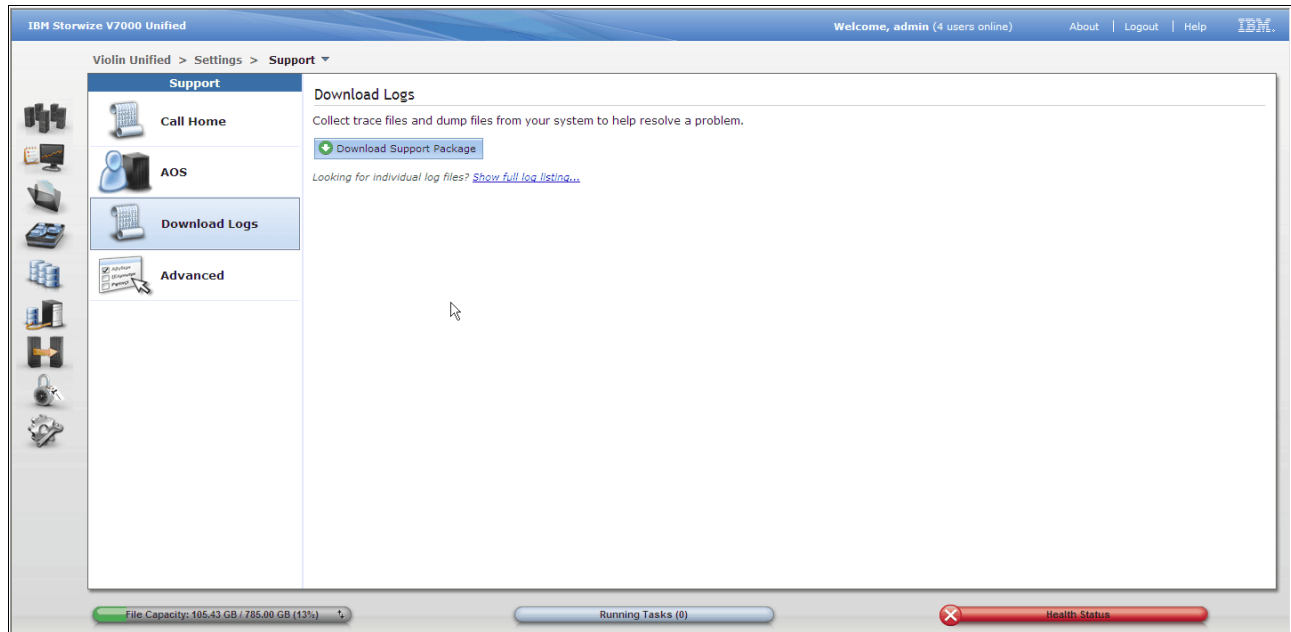


Figure 15-17 Download logs

Here you can show a listing of the packages currently stored on the cluster and you can initiate a current data collection. If the file list is not displayed, select *Show full log listing...* to see the file list. Then select the node to view from the pull down.

**Caution:** Each module only stores the data it collects. The data collect file is collected and stored by the current management node. Ensure you are looking at the listing for the correct node.

To collect the logs, click on the *Download Support Package* button. This will then ask for the type of logs required as seen in Figure 15-18. Choose “full logs” and then click download to obtain all the logs. A progress window will then display. Wait for the entire task to complete and confirm it was successful. Then close the window.

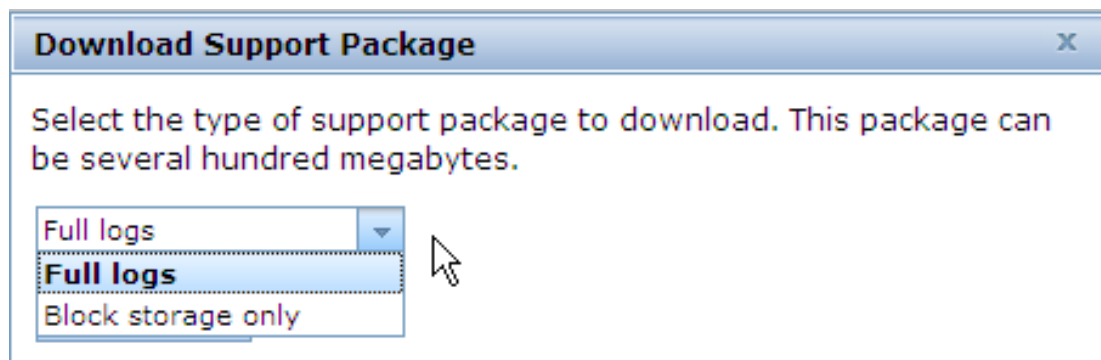


Figure 15-18 Download logs - full logs

Locate the resultant file in the file list. The filename includes a date-time stamp, use this to identify the recent log. Highlight the file and either right click or use the action pull down, then select download. This file can then be saved to your workstation and is ready to be uploaded to IBM Support. Avoid renaming this file.

**Note:** The files will remain on the file module indefinitely. It is your job to maintain these and delete old copies. There is no concern about how many remain on the module other than confusion. Our suggestion is to delete files only after the problem is resolved and always keep the last one.

If using compression, the RACE module maintains internal diagnostics information, which is kept in memory in each of the V7000 nodes in a compression I/O Group. This information is available as part of the Support Package for the Block Storage. The “Full Logs” will contain the **Standard logs plus most recent statesave from each node** which is fine if the system observed a failure. For all other diagnostic purposes such as a performance issue then the **Standard logs plus new statesaves** is required.

To obtain the **Standard logs plus new statesaves** select **Block storage only** as shown in Figure 15-19 and then select **Standard logs plus new statesaves** as in Figure 15-20.

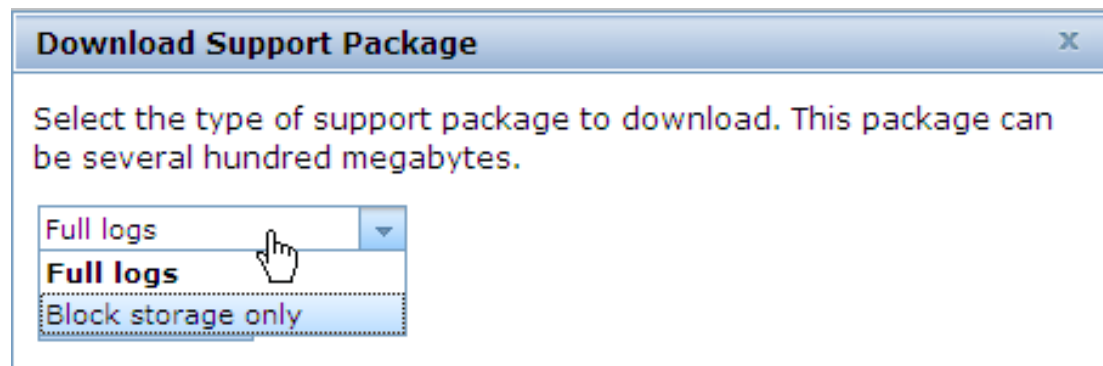


Figure 15-19 Block storage only

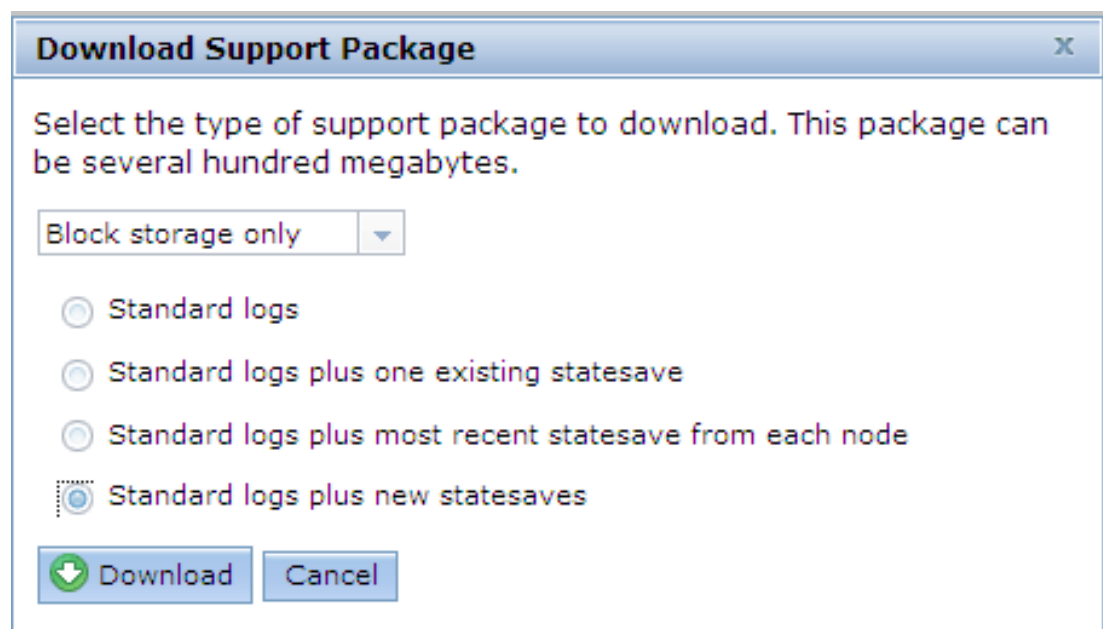


Figure 15-20 Standard logs plus new statesaves

## 15.4 Information Center

All documentation for the Storwize V7000 Unified is compiled into the online *Information Center*. This includes installation and configuration, troubleshooting and administration. The information center home page is shown in Figure 15-21 on page 258, and can be found at this URL:

[http://publib.boulder.ibm.com/infocenter/storwize/unified\\_ic/index.jsp](http://publib.boulder.ibm.com/infocenter/storwize/unified_ic/index.jsp)

As seen in Figure 15-21 on page 258, the page is broken up into three main areas, the search bar, left pane and right pane. The left pane has three tabs at the bottom to choose the view displayed. Either contents (displayed as default), index, and search results. The right pane is used to display the selected detail page. The search bar can be used to do a context search of the entire Information Center using rich search arguments.

### 15.4.1 Contents

This tab in the left pane displays an expandable map view of the headings and subheadings of the information available. This view can be used to navigate quickly to the information and to see all information available under a subject heading.

### 15.4.2 Index

In this tab we see an alphabetical listing of keywords that can be scrolled through or searched using the word find box at the top of the panel.

### 15.4.3 Search results

The results of a search performed in the search tool bar at the top of the page will be listed in this tab. For each result the section heading matching the search argument is given along with an abbreviated summary of the text in that section. By click on the heading, which is a URL, the page will be displayed in the right panel.

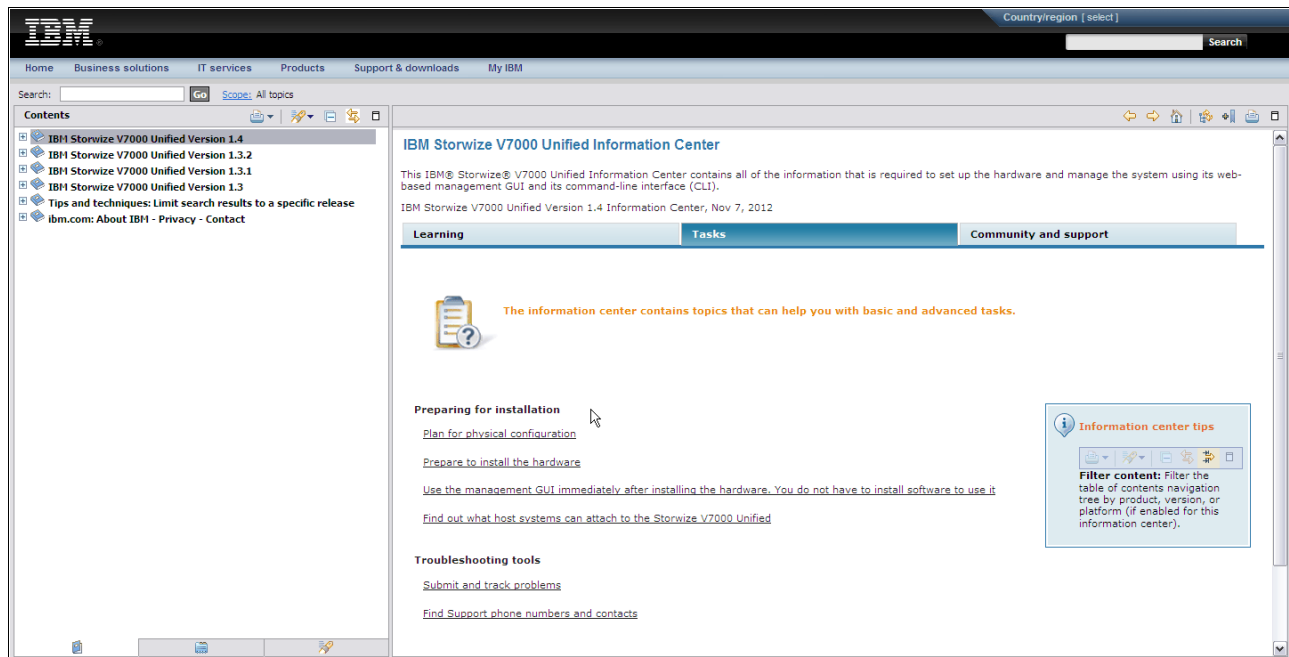


Figure 15-21 Information Center - home

The most successful way to locate information in the Information Center is to construct a good search argument for the information you want to see, enter that into the search field in the search bar at the top of the page. Scan the results for the section that best meets your requirements and display each one in the right window by clicking on the URL. If you want to look at several, use your browser's functions to open the links in a new tab or window. Note, by hovering over the url a popup displays showing the source of the item. Be careful of multiple similar hits where some are marked *previous version*.

#### 15.4.4 Offline Information Center

For situations where internet access is not possible or poor, a downloadable version of the Information Center is available. Navigate to the download section of the support page to locate the link to the files. The current download can be found at

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S4001035>

**Note:** The download is very large, over 1 GB. When installed, it will require a lot of storage resource.

There are two modes the offline Information Center can be installed in, either on your workstation for personal use or on a server for use by anyone. The readme file gives details on each and how to set up.

**Caution:** Once downloaded, the information in the Information Center is not updated and may not reflect the most current detail. In general, the information provided will be correct and valid, but if in doubt, particularly with recovery procedures, consult the online system to confirm. To maintain concurrency, the file needs to be downloaded and the offline Information Center reinstalled on a regular basis.

The only way to learn about what the Information Center offers, is to use it. Spend time looking at what information is available and how to find it. Practice with searches and locating procedures and references. In a short time you will be comfortable with the layout of the information and able to quickly locate the topics you are looking for.

## 15.5 Call home and alerting

While the cluster will automatically detect events and errors and log these, alerting is needed to ensure the user is made aware of an issue in a timely manner. The cluster incorporates three methods of alerting. For the configuration of these services, refer to 11.8.3, “Support” on page 172.

### 15.5.1 SNMP

This agent utilizes the well known SNMP protocols which are in common use. Alerts are sent to the designated SNMP server. System health reporting and error alerting are determined by the individual implementation of the SNMP process within each enterprise and is beyond the scope of this book.

The cluster is configured with the address of the SNMP server. Additionally, the severity of each alert type can be individually set to filter the alerts being presented. Multiple server targets, each individually defined can be configured.

### 15.5.2 Email

This is a very versatile method of alerting which uses a simple and well known process (email), to send an alert. This can be to an individual’s mailbox, a group mailbox or even a server.

The cluster is first configured to allow email alerts, and the SMTP server it will send the emails to, needs to be defined. Next each recipient needs to be defined. As many email recipients as desired can be configured. Each definition needs the email address and can be individually set up for the type of event to alert and the severity of each type. This gives good flexibility for each user to receive only the alerts they are interested in.

**Caution:** Resist the urge to enable all levels of alerting. This will generate a high volume of alerts as every minor change in the system will be notified. We suggest “Critical only” for most users and “Critical / Warning” for normal monitoring.

### 15.5.3 Call Home

Call home is one of the IBM Support tools included to assist IBM in supporting your system. Using the same process as the email alerting, an email is sent directly to an IBM server on the web, which includes a brief log of the error. Provided that your system has been properly setup in IBM’s support system, this will generate a Problem Management Record (PMR). A call will then be queued to the support team for your region.

If you have not already called IBM, depending on the nature and severity of the error, IBM may call you. This call home process gives IBM Support timely awareness of the problem and allows them to begin analysis. It also serves as a backup for the client alerting.

**Important:** While IBM may receive direct and early alerting for a error, this does not constitute a service call. It is the client's responsibility to ensure alerting within their enterprise is effective and critical events reacted to. The client should always place a service call, unless IBM have already called them. In most cases the call home improves IBM's responsiveness and leads to faster analysis and recovery.

## 15.6 IBM Support Remote Access

There may be times when during an outage IBM Support need to connect to the cluster. This may be to speed up problem determination and recovery by allowing IBM to issue commands and see responses directly, or there may be commands and access required that can only be performed by IBM.

IBM uses an existing tool called Assist-on-Site (AOS) which has been used by IBM Support for many years in support of x-series and storage products. This involves a client daemon loaded on the client's workstation or system and this is connected to a secure server within IBM. IBM Support specialists can access this server internally and are able to view the console and use the keyboard and mouse. Access within IBM is strictly controlled and restricted. If required, once a session is established, other approved IBM specialists can also monitor the session allowing a cooperative approach ultimately speeding up recovery.

For a Windows workstation, the client accesses a web page and registers. A small temporary agent is downloaded which provides the client side daemon. Access is via well known ports 80 and 433 and is secure. A session unique pass key must be entered to complete the connection which is passed verbally. The client and the IBM specialist share the video, mouse and keyboard and the client maintains full control.

With Storwize V7000 Unified a permanent client has been included in the software package. This is disabled by default and must be enabled and configured by the user. There are 2 modes of operation, lights on and lights out. Lights on implies that the site is manned at all times that a connection will be needed. Any connection attempt will require authorisation by the client connecting locally to the cluster GUI and approving the connection. Light out implies the cluster is in an unmanned (at least for some of the time) location. Connection attempts will therefore complete without the need for local approval. Use this setting to define how you wish IBM Support to connect based on your local security policies. We suggest lights on as providing a level of protection and awareness, provided that if needed, approval via the GUI can be given in a timely manner.

Setup of the AOS connection is covered in the Chapter 11, "Implementation" on page 133.

## 15.7 Changing parts

If a part or component of the Storwize V7000 Unified fails, your first indication will be an entry in the event log and most likely an alert, depending on your alert settings. Depending on the nature of the failure and the impact, you may have run Guided Maintenance Procedures against the event (in the case of block storage) or you may have researched the error code in the Information Center. If the procedures indicate that a part should be changed, then you need to place a call to IBM using your local procedures. If at any time during your analysis of the problem you feel uncomfortable or the problem is pervasive, immediately place a call to IBM for assistance.



IBM Support will almost certainly ask you for a Data Collection so they can review the logs, interrogate other related information in the data and research the problem against internal problem databases. IBM Support may ask you to perform other recovery actions, data gathering or execute commands. If IBM determine that a part needs replacement, they will arrange the shipment of the part and depending on which part, give you instructions on the replacement procedures or send a service representative to perform the actions.

The Storwize V7000 Unified is designed so that many parts can be replaced by the client. This speeds up recovery and reduces overall cost. Almost all parts in the Storwize V7000 storage enclosure can be replaced by the client. These are typically done under the guidance of the Guided Maintenance Procedures which are launched against the error log. IBM term these parts as CRUs.

If the part is not a CRU, or if more technical problem determination is required, IBM will dispatch a service representative.

It is important to ensure you are familiar with the safety procedures outlined in the Information Center and have a good understanding of the process before changing any parts. If at any time you feel unsure, call IBM Support for guidance.

CRU parts are typically couriered to your location. Depending on local country procedures, this shipment will include a process to return the removed part which you must ensure is completed promptly.

## **15.8 Preparing for recovery**

When a disaster strikes, time is always the most important aspect. Getting the problem diagnosed and systems recovered is generally against the clock. Many tasks and procedures are only ever encountered or needed at the worst time and if any of these are unprepared or untested, this frequently leads to an outage being longer than it needed to be.

The following are a number of key items that if prepared beforehand will greatly improve the diagnosis and recovery time.

### **15.8.1 Superuser password**

This is the password for the userid “*superuser*” which is used to log on to the Storwize V7000 storage enclosure directly. In the unified cluster, this password is seldom, if ever used as all functions are performed from the Unified GUI and CLI. But if certain problems occur on the storage component, then this password will become essential and if it is not handy, will delay service actions.

Store the password securely, but quickly accessed.

### **15.8.2 admin password**

While this userid may be in regular use, the password needs to be made available at short notice when service is required. If the userid and password are held aside and all users have their own IDs, a userid and password with full authority needs to be available for service use. Alternate methods of producing a valid ID/password need to be in place in the event that the authorised person is not present or available.

### 15.8.3 Root password

As discussed in the implementation section, currently the root password to the file modules is widely known to IBM Support personnel. If this poses a risk, then we suggest the client asks IBM to change it at install time. If this is done, then the password needs to be secured and able to be produced on demand. Currently a number of service and recovery procedures require root access. If it is not available, IBM Support's ability to recover a problem will be compromised. IBM intends to negate the need for field support to use this level of access in the future. When this happens, this action will no longer be required.

### 15.8.4 Service ip addresses

An often overlooked step during implementation is to set the IP addresses for service access to the storage nodes and the file modules. The storage node service IP addresses have default settings, but these are unlikely to match the clients networks, needing local and direct physical access to gain a connection. They should be set to assigned IPs in the local LAN so they are accessible from the client's network. These addresses need to be documented and easily accessed.

Considerable time is wasted if these addresses are not known or not in the local networks.

### 15.8.5 Test GUI connection to the Storwize V7000 Storage enclosure

Ensure that as part of the install and on a regular basis, you have confirmed that you can connect to the Storwize V7000 storage nodes directly and that this is tested regularly. You should be able to connect to and logon to the storage system GUI over the management address and to both the Service Assistant GUIs over the two service IP addresses.

Also set up both the CLI connections and test.

### 15.8.6 Assist-on-site

IBM has included a tool in the software that allows simple access to the Storwize V7000 Unified. AOS needs to be configured, so ensure that this has been done and tested before it is needed to ensure no time is lost if there is a need for IBM Support to connect to the cluster. This includes resolving any firewall and security issues. We encourage you to do this configuration and testing during the implementation.

### 15.8.7 Backup config saves

Prior to any maintenance activity, do a config backup and then off-load it either as files or by doing a normal data collection. This will greatly assist support in the event of a problem and may increase the likelihood of a successful recovery.

For the storage units, issue the CLI command:

```
svcconfig backup
```

This will gather and save a current version of the config.xml file. Then do a data collect as normal, offload and save the file safely. It is wise to perform this activity at regular intervals (e.g. monthly) as a matter of housekeeping.

## 15.9 Software

All the software used in the Storwize V7000 Unified is bundled into a single package. Each bundle is uniquely named using the IBM standard for software name and version control. The bundle is identified by its Version.Release.Modification.Fix (VRMF) number (for example, GA release was 1.4.0.0). These numbers increase as new code is released. The meaning of each field is as follows:

Version	Only increments for a major change to the functions and features of the software product.
Release	Increments as new functions or features are released
Modification	Increments when a function or feature is modified
Fix	This number increments for each group of PTFs (fixes) that are released. Note, PTFs do not change function, only fix code problems.

### 15.9.1 Software package

There are five different software packages used with the Storwize V7000 Unified.

- ▶ Full software install DVD. This is only needed in the case of a full rebuild of a cluster. It is a bootable DVD which will reload all the File Module software, management software, update HW BIOS and utilities if needed and update Storwize V7000 software if required. This process initializes the system, destroying previous configurations. This is an ISO image and can only be obtained from IBM Support for a specific situation.
- ▶ Software update. This file is used to concurrently upgrade the File Module software, management software, update HW BIOS and utilities if required and update Storwize V7000 software if required. This file can be downloaded from IBM and is installed using the upgrade tools in the Storwize V7000 Unified cluster.
- ▶ Storwize V7000 control enclosure software update. This update file is included in the Storwize V7000 Unified package and is managed and installed by the File Module management software. It will be automatically updated if required. Do not attempt to upgrade the storage control enclosure manually unless directed to do so by IBM Support.
- ▶ File Module HW BIOS and Utilities. This firmware is managed by the Storwize V7000 Unified software and will be automatically upgraded if required. Do not attempt to manually upgrade this firmware.
- ▶ Install test utility. This small package is loaded in the same way as the normal software update. It install a test utility that then checks for known issues or conditions that could potentially cause the software apply to fail. It is important to run this utility first.

**Warning:** Do not use the DVD to upgrade an operational system. The DVD is only used to install and initialise a new or recovering system. Installing the software from the DVD *will* destroy the existing configuration.

### Download

If your Storwize V7000 Unified has a connection to the internet and sufficient bandwidth is available to perform the download, then log on to the management GUI and from the upgrade software option, download the upgrade file. If connectivity is not available or direct download is not desired, then download the files from the web as we describe here.

### Web download

All software packages are downloaded from IBM Fix Central. To access this site you need an IBM ID. If you don't already have one, there is a link on the front page to register.

Select two packages from the list of available software. First select and download the latest UpgradeTestUtility. This is a small file and contains the test utility that is installed first to check for known problems.

Then select and download the latest Storwize V7000 Unified software bundle. This includes the main upgrade file. Also included is the license agreement document, the release notes for this release and a small file with the MD5 checksum value. It is important to read the release notes carefully. This file constitutes the “read me” and special instructions for this release.

To access the latest software go to:

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

Select:

Product Group → System Storage  
 Product Family → Storage Software  
 Product Type → Storage Virtualisation  
 Product → All others  
 Then continue

If not already in your product list then add Storwize V7000 Unified to your list and select it.

This will show a list of downloads for the Storwize V7000 Unified. Click on the latest version to see the downloadable files page. Scroll down and review as required the information given on this page, which includes details and links for compatibility tables, interoperability and restrictions. At the bottom is the package download links. Choose the applicable line and select the download option, HTTP (or FTP if available). Read and accept the terms and conditions popup. This will then take you to the download page as shown in Figure 15-22.

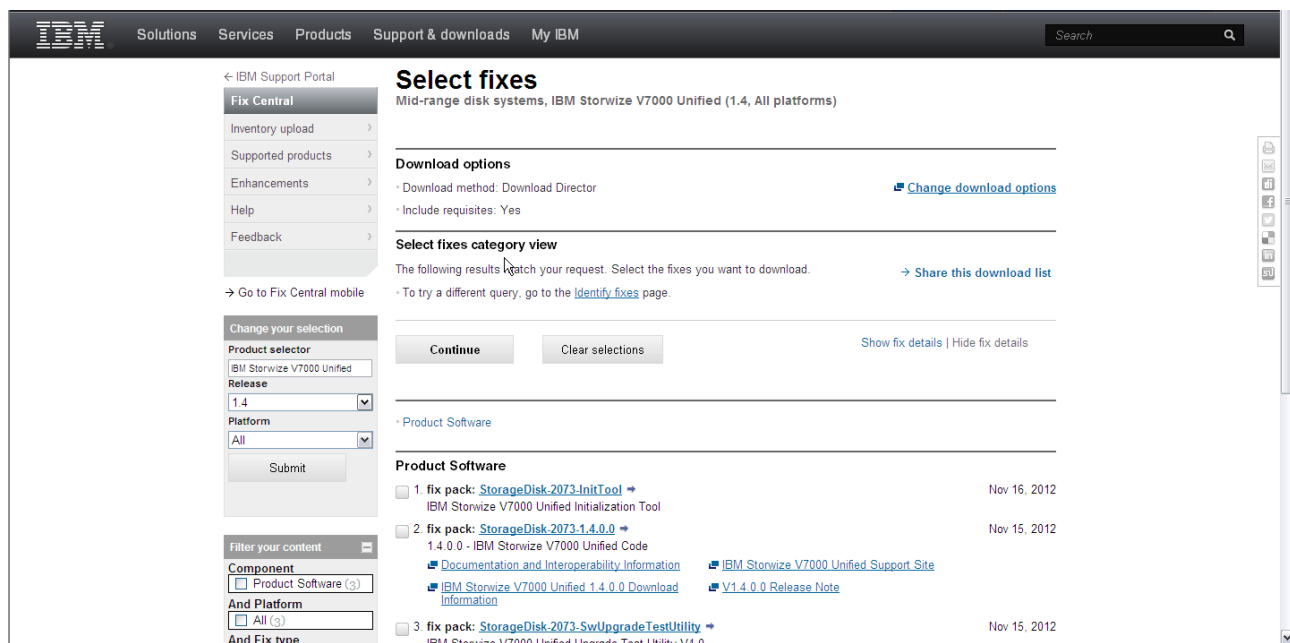


Figure 15-22 Software download

Make sure the package you need is selected and click continue. Follow the prompts to save the package.

## 15.9.2 Software upgrade

**Important:** Always read the release notes before continuing.

There are two methods of installing software on the Storwize V7000 Unified. The normal process is to perform an upgrade which is applied concurrently. This involves a stepped process of stopping and upgrading one file module, then bringing it back online. After a settling time, the other file module is done in the same way. If required, the control enclosure is also upgraded. This involves stopping and upgrading one node, waiting 30 minutes, then doing the other node. All this is concurrent with the Storwize V7000 Unified operation and client access is not interrupted.

File system access will switch between the modules and host access will continue uninterrupted provided that the hosts have network connectivity to both file modules. For block access over the SAN, all hosts must have operational multipathing and at least one path to each node canister.

**Important:** Ensure all LAN connected file system hosts have confirmed access to both file modules and all SAN attached fibre channel hosts have paths to both node canisters.

The second method involves restoring all software from DVD. This is disruptive and by design will destroy the Storwize V7000 Unified configuration and any data stored on the system will be lost. Unless this is a new install or rebuild where no data exists on the system, do *not* perform this procedure without direct IBM Support direction.

### Install new software from DVD

This procedure is described in the installation procedures 11.5, “Install latest software” on page 147 and should only be used on a new system or under the direction of support during a recovery action.

### Concurrent software upgrade

The following procedures are run from the “Upgrade Software” option on the management GUI of the Storwize V7000 Unified. Select:

Settings ==> General ==> Upgrade Software.

You will get a window as shown in Figure 15-23.

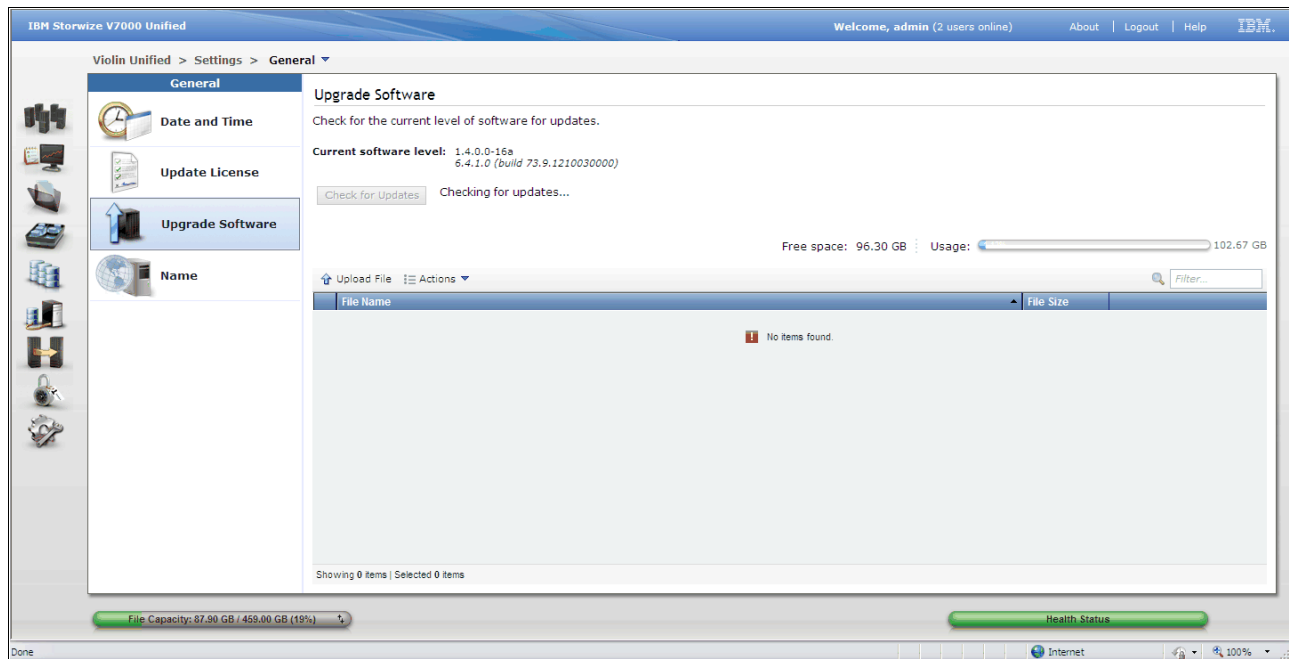


Figure 15-23 Upgrade software - home

If your cluster has good web access, you can click “Check for updates” to connect to IBM and download the latest upgrade file to the cluster. Note that this file is over 1GB and may take some time to download using this method.

Alternatively, using the procedure detailed above, connect to IBM Fix Central and download the upgrade file to your workstation. You will then need to upload the file to the cluster. Click the “Upload File” button on the page and this will display a popup window as shown in Figure 15-24 on page 267.

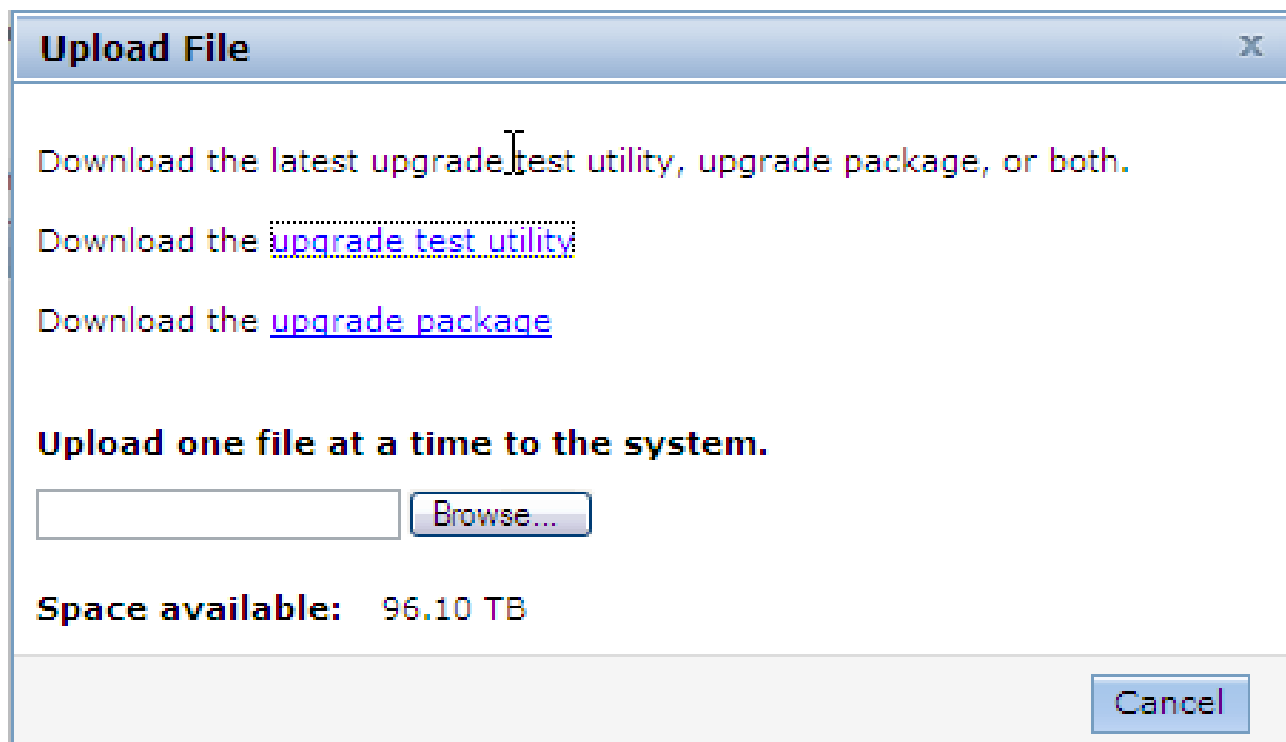


Figure 15-24 Upload File

**Hint:** Once the upgrade process has completed (success or fail), the file is deleted. For this reason, you might find it beneficial to download the file to your workstation, then upload it to the cluster, rather than use the check for updates button. This way you always have a copy of the file and only need to upload it again in the event of a failure. This also saves time and bandwidth if you have multiple clusters. The likelihood of needing the file again on the same machine after the upgrade has been successfully applied is very low, but this method does offer some piece of mind.

Browse to the upgrade file that you have downloaded on your workstation, then click OK.

Repeat the upload step for the test utility that you also downloaded with the code file.

The new files will be opened and interrogated by the cluster code and if acceptable, will appear in the list. First run the upgrade test, always using the latest version available. Select the file, then right click to select install, or use the actions pull down as shown in Figure 15-25.

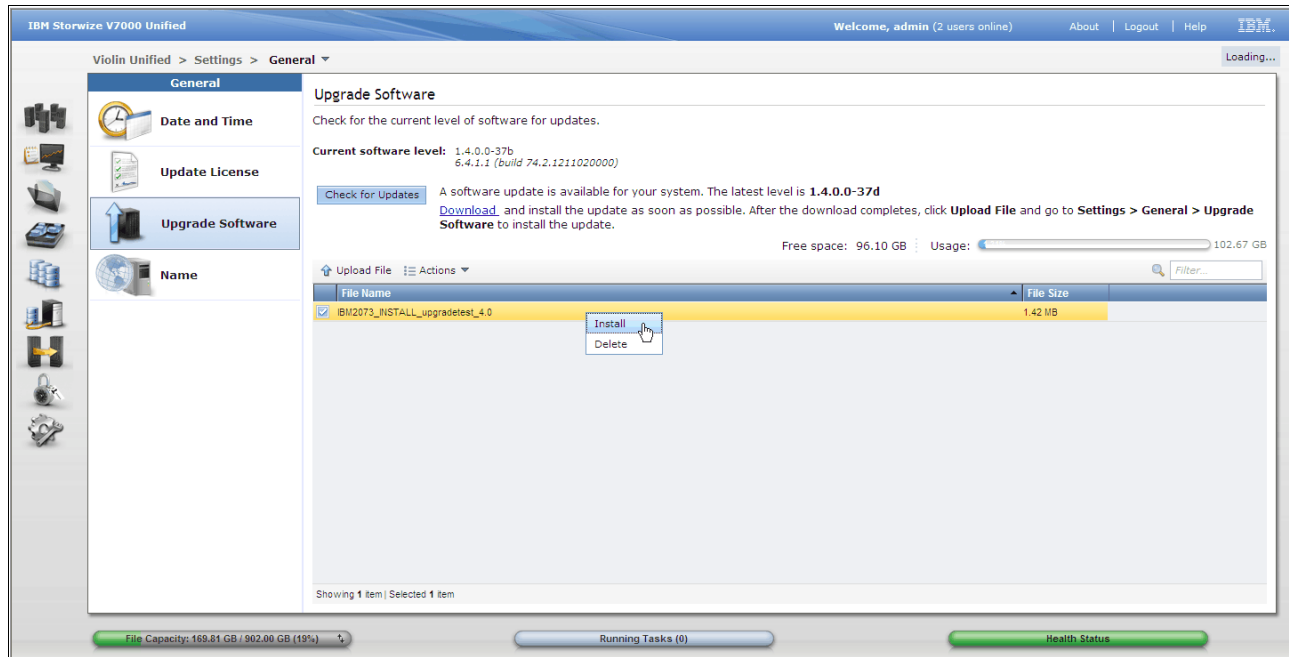


Figure 15-25 Upgrade software - install test utility

**Hint:** You need to select the file before performing an action, shown as highlighted yellow. then the actions are available.

Wait for the utility to run and when complete, review the result in the window before closing it. An example of a successful run is shown in Figure 15-26 on page 268.

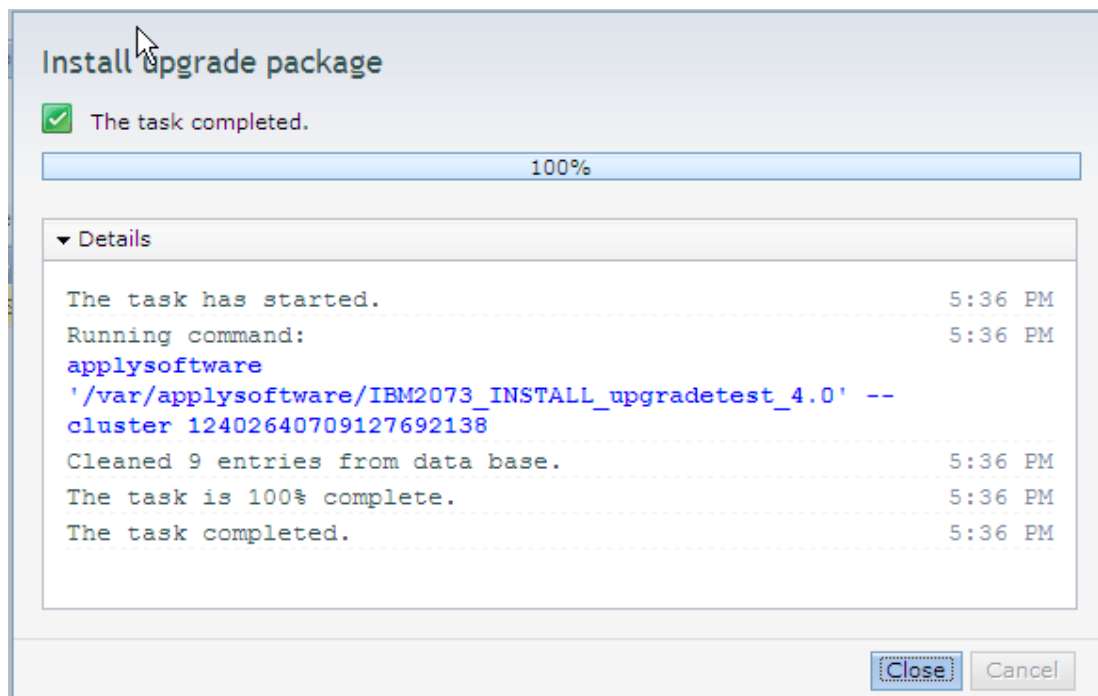


Figure 15-26 Upgrade software - test results



If there were any problems reported these need to be resolved before continuing with the software upgrade.

**Important:** While code upgrades are being applied, all data access will continue. Block access continues to be provided over the SAN provided that hosts have multipathing correctly configured with active paths to both nodes and that hosts with file access have IP connectivity to both file modules and will support a fail over of the IP addresses between the file modules.

However, do not make any configuration changes to the cluster during the software upgrade. Many tools and functions will be temporarily disabled by the apply process to prevent changes from occurring.

Once you have confirmed the test utility has shown no problems, the next step is to click on the upgrade file to select it and then take the action to install, as seen in Figure 15-27 on page 269.

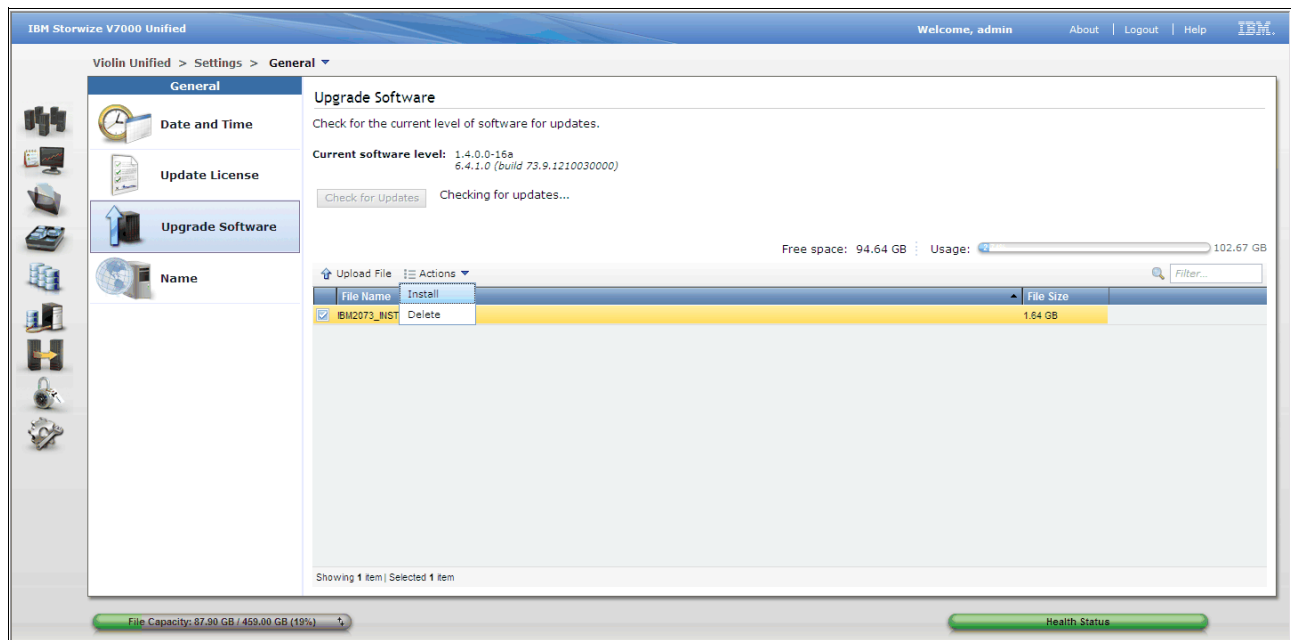


Figure 15-27 Upgrade software - install

A popup progress window will display. Wait for the task to complete, then close. This does not indicate that the upgrade is applied, but that it has been successfully started. The upgrade software window will now show a status screen with three progress bars. The top one is the overall progress, and the other two are for the file modules. If the new upgrade includes a code level that is higher than that currently on the storage control enclosure, then it will automatically perform an upgrade of the control enclosure first. This step will take about 90 minutes and can be monitored by the overall progress bar. It is complete when the bar is at 33%.

When the storage is complete, then the process moves on to the first file module as shown in Figure 15-28 on page 270.

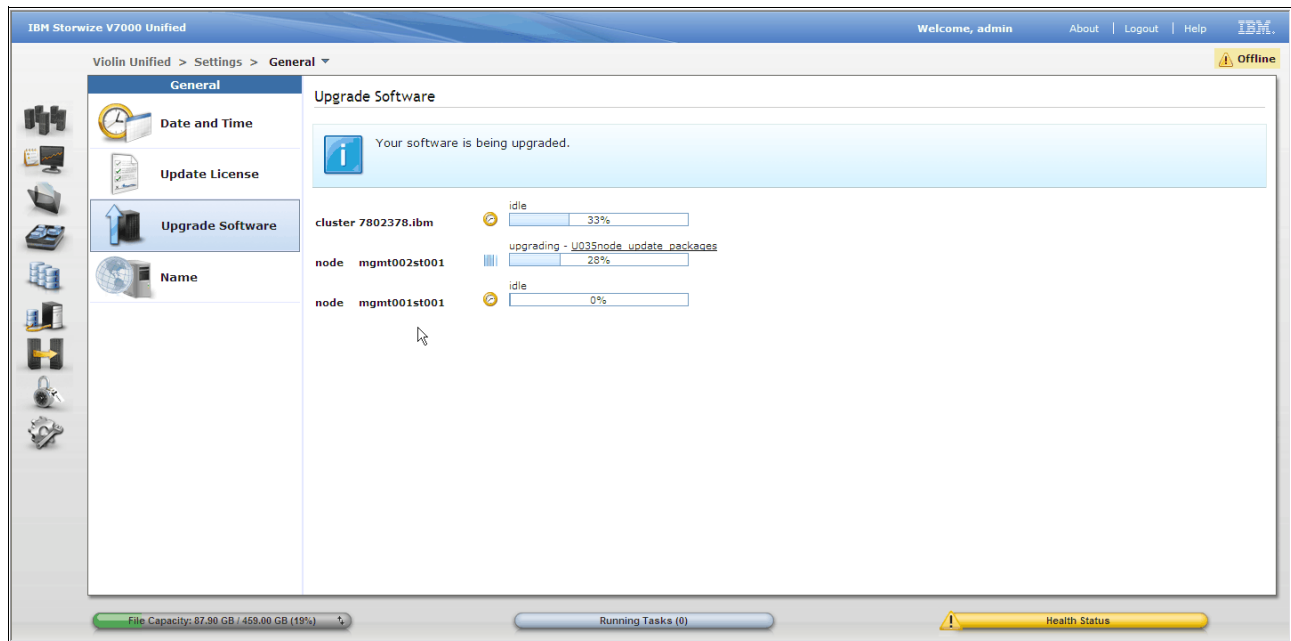


Figure 15-28 Upgrade software - progress

The file modules take about 30 minutes to upgrade each. Once the first one is complete, then the process will move on and upgrade the second one. At this time the GUI connection will be lost due to the active management module now being taken down and swapped over. You will need to log back on to continue monitoring the progress.

When all three modules have been successfully upgraded, then the process upgrade cluster common modules and processes begins. This take about another 30 minutes. The upgrade is complete when this process finishes as shown in Figure 15-29 on page 271.

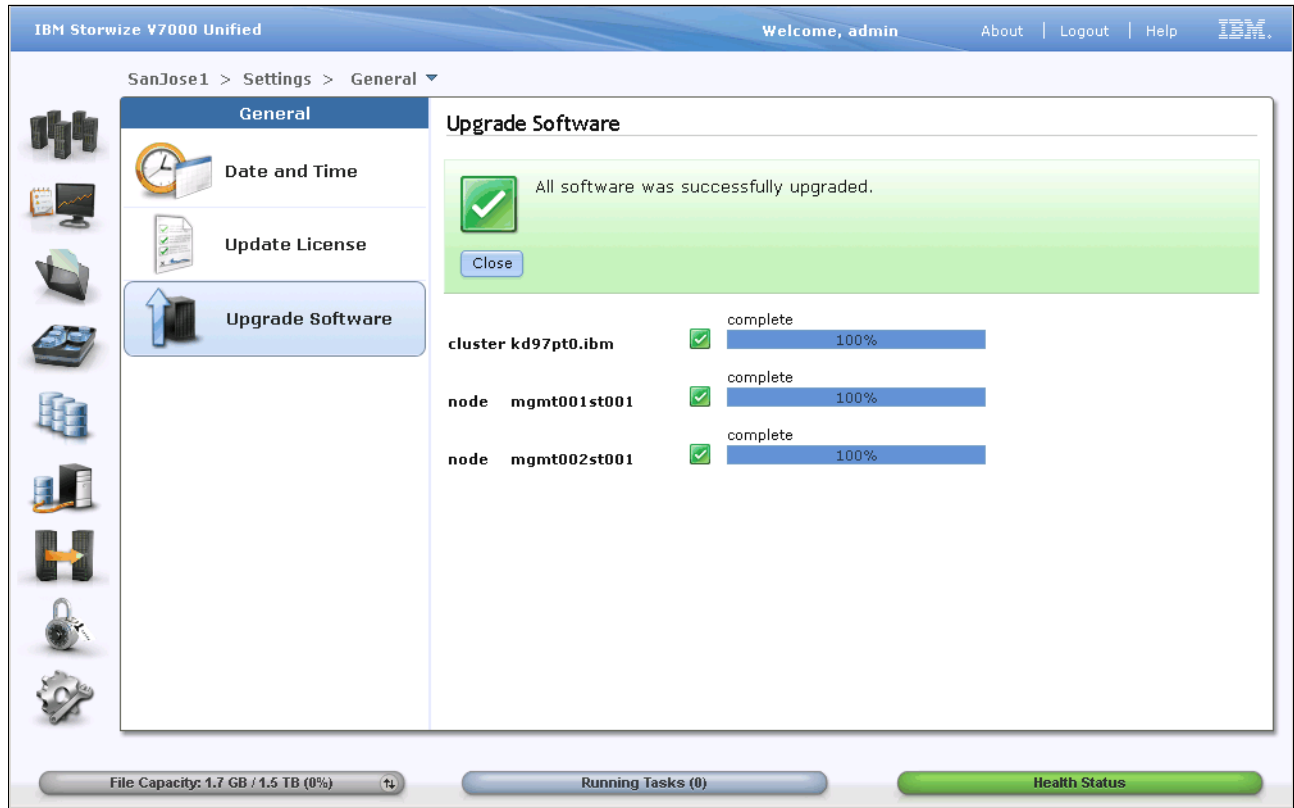


Figure 15-29 Upgrade software - complete

Your software is now upgraded.





**16**

# **Real-time Compression in the IBM Storwize V7000 Unified**

## 16.1 General Considerations for compression and block volume compression use cases

*Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859 describes the base technology for real time compression in great detail and describes the compression use cases for compressed block volumes. The following sections will cover specific considerations of compression in the file system context of the IBM Storwize V7000 Unified.

## 16.2 Compressed File System Pool Configurations

Using the feature to compress volumes together with the NAS feature of the Storwize V7000 Unified provides some benefits over using compression with an external host that creates a file system on a compressed volume:

- ▶ The policy language capabilities of the Storwize V7000 Unified allows a flexible way to create compression rules based on file types and the path where files are placed.
- ▶ The file systems GUI panel allows us to monitor the capacity of file systems, file system pools and the related mdiskgroups in a single view.
- ▶ The creation of file systems using the Storwize V7000 Unified GUI will ensure that best practices are followed.

Compression in the Storwize V7000 Unified Storage system is defined per volume. Not all the degrees of freedom resulting from this implementation should be utilized. This section describes some example configurations that are applicable to different use cases.

### **Important:**

File system metadata should not be compressed.

It is not recommended to mix compressed and uncompressed data volumes in a single file system pool.

The file system pools should be regarded as the smallest entity where compression is enabled. So although compression is implemented at a volume level, we refer to compressed pools or compressed file systems meaning that all data volumes of a file system, or all data volumes of a file system pool are compressed.

### 16.2.1 Selectively compressed file system with two pools

This configuration is considered as best practice to create file systems that make use of compression. The creation of two pools together with a placement policy allows to place only compressible files to a compressed pool. This approach provides high compression savings with best performance. This is shown in Figure 16-1 on page 275.

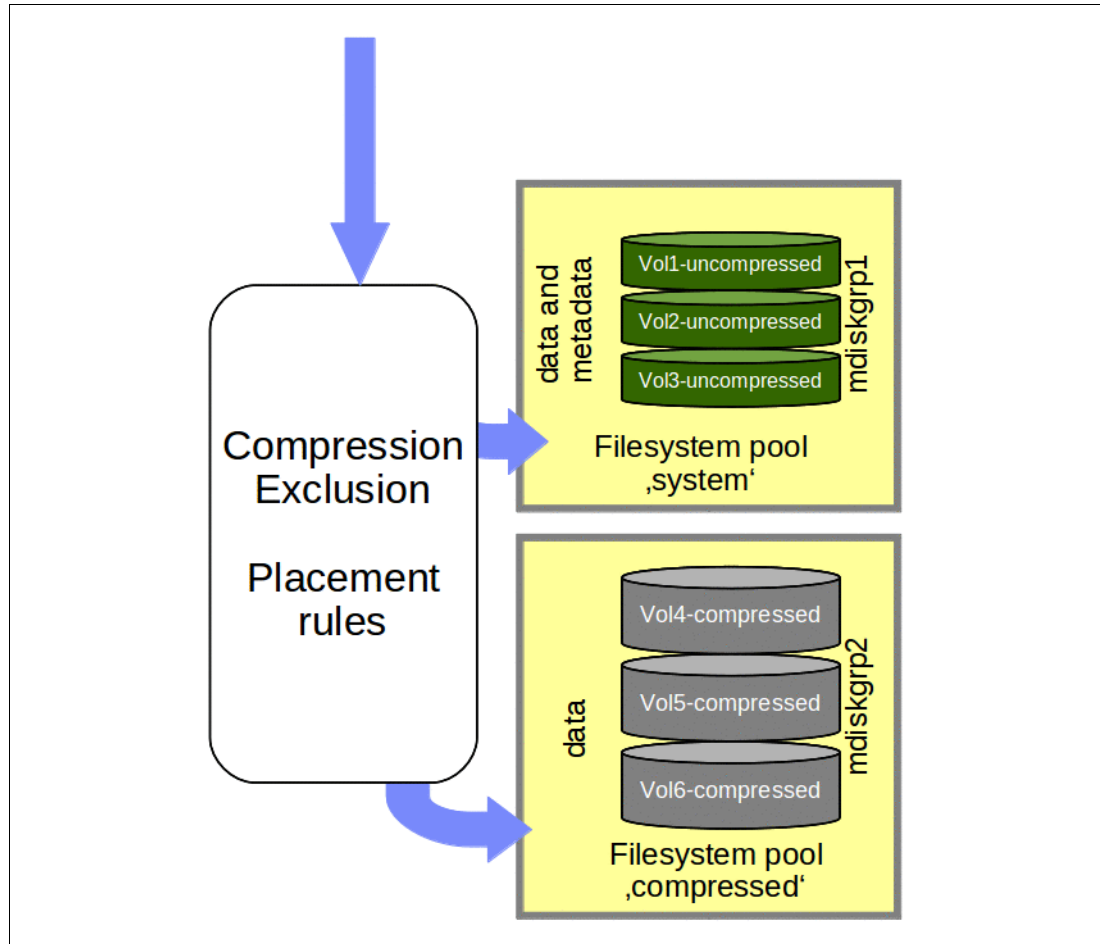


Figure 16-1 Selective compressed configuration with two file system pools

This filesystem consists of two file system pools. The 'system' file system pool is based on uncompressed volumes, while the compressed pool consists of compressed volumes only. A placement policy decides where files are placed during creation.

The placement rules that will be employed will direct files that are known to not compress well to the 'system' file system pool, whereas all other files are placed to the 'compressed' file system pool. The placement rules are only evaluated while a file is created. Updating a file will not change the placement. A change in a placement rule will also not work retrospectively. In order to correct a faulty placement, a migration rule can be run. In 16.5, "Managing compressed file systems", some examples of migration rules are described.

**Pools:** File system storage pools are different from V7000 storage pools.

The Storwize V7000 storage pools are synonymous with mdiskgroups. In order to avoid confusion, the following sections will use the term 'mdiskgroup' rather than 'storage pool'.

The file system storage pools are parts of a file system. The initial file system pool of a file system is called 'system'. Only the 'system' pool contains the metadata of the file system, but it can also contain data. All other filesystem pools contain data only. When a file volume is created, the usage type of the volume is either data only, metadata only or data and metadata. Compressed volumes should be data only volumes.

## 16.2.2 Configuring a selective compressed file system

This section describes how to create a selective compressed file system using the GUI.

### 1. Define a file system with two file system pools

A selective compressed file system can be configured in the “New File System” dialog which is invoked by clicking “+ New File System” in the Files → File Systems view. Choose the custom preset and add a second file system pool by clicking the green + sign. The “Compressed” flag needs to be enabled for this second file system pool. A meaningful name should be given to the filesystem pool. File system pool names are only unique per file system, so choosing ‘compressed’ is always a good option for a compressed file system pool. See Figure 16-2.

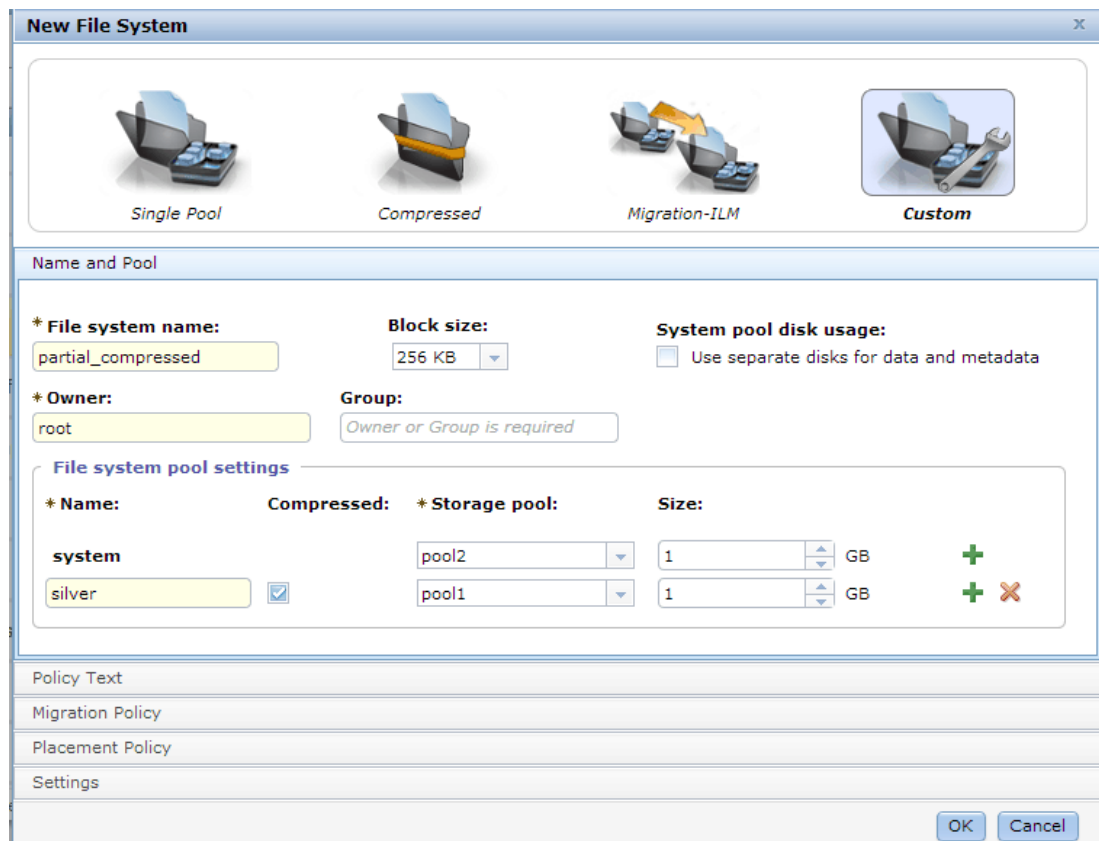


Figure 16-2 Selective Compressed File system configuration: Two file system pools will be created, only the “compressed” file system pool will be compressed.

### 2. Select mdiskgroups

The mdiskgroup(s) must be chosen. For improved performance, the system pool which contains metadata should reside on a separate mdiskgroup.

It is also suggested to separate compressed and uncompressed volumes into separate mdiskgroups.

As a variation of our described configuration, the file system metadata can be placed on separate dedicated volumes by choosing the “Use separate disks for data and metadata” option. On systems that contain solid state disks (SSDs), the mdiskgroups that contain the SSDs can then be used to provide metadata-only volumes.



When only one mdiskgroup is chosen for the 'system' filesystem pool, the metadata replication feature has to be turned off in the 'Settings' tab. If sufficient mdiskgroups are available, it is suggested to define two mdiskgroups for the system pool and enable metadata replication.

### 3. Define the size of the file system pools

The uncompressed size of the two file system pools has to be provided now.

### 4. Define placement policy to exclude uncompressible file from the compressed pool

Example 16-1 lists file types which, at the time of writing, are known to not compress well because they either contain compressed or encrypted data.

*Example 16-1 Extensions of file types that are known to not compress well.*

---

```
7z, 7z.001, 7z.002, 7z.003, 7zip, a00, a01, a02, a03, a04, a05, ace, arj, bkf,
bz2, c00, c01, c02, c03, cab, cbz, cpgz, gz, nbh, r00, r01, r02, r03, r04, r05,
r06, r07, r08, r09, r10, rar, sisx, sit, sitx, tar.gz, tgz, wba, z01, z02, z03,
z04, z05, zip, zix, aac, cda, dvf, flac, gp5, gpx, logic, m4a, m4b, m4p, mp3, mts,
ogg, wma, wv, bin, img, iso, docm, pps, pptx, acsm, menc, emz, gif, jpeg, jpg,
png, htm, swf, application, exe, ipa, part1.exe, crw, cso, mdi, odg, rpm, dcr,
jad, pak, rem, 3g2, 3gp, asx, flv, m2t, m2ts, m4v, mkv, mov, mp4, mpg, tod, ts,
vob, wmv, hqx, docx, ppt, pptm, thmx, djvu, dt2, mrw, wbmp, abr, ai, icon, ofx,
pzl, tif, u3d, msi, xls, scr, wav, idx, abw, azw, contact, dot, dotm, dotx, epub,
keynote, mobi, mswmm, odt, one, otf, pages, pdf, ppsx, prproj, pwi, onepkg, potx,
tiff, !ut, atom, bc!, opml, torrent, xhtml, jar, xlsx, fnt, sc2replay, lst, air,
apk, cbr, daa, isz, m3u8, rmvb, sxw, tga, uax, crx, safariextz, xpi, theme,
themepack, 3dr, dic, dlc, lng, ncs, pcd, pmd, rss, sng, svp, swp, thm, uif, upg,
avi, fla, pcm, bbb, bik, nba, nbu, nco, wbc, dao, dmg, tao, toast, pub, fpx,
prg, cpt, eml, nvram, vmsd, vmxf, vswp
```

---

This list of extensions in this example can be used to define an exclusion rules for a placement policy. The GUI provides an editor to define the exclusion rules for file types that should not be compressed, as shown in Figure 16-3 for the placement policy. The GUI placement editor defines the placement rules in a case insensitive way. The file attribute "Extension" has to be chosen and the operator "NOT IN" for the exclusion placement policy. The list of extensions shown in this section can be added to the GUI dialog using cut and paste, the GUI entry field will accept extensions with various delimiters and will format it automatically. See Figure 16-3 on page 278.

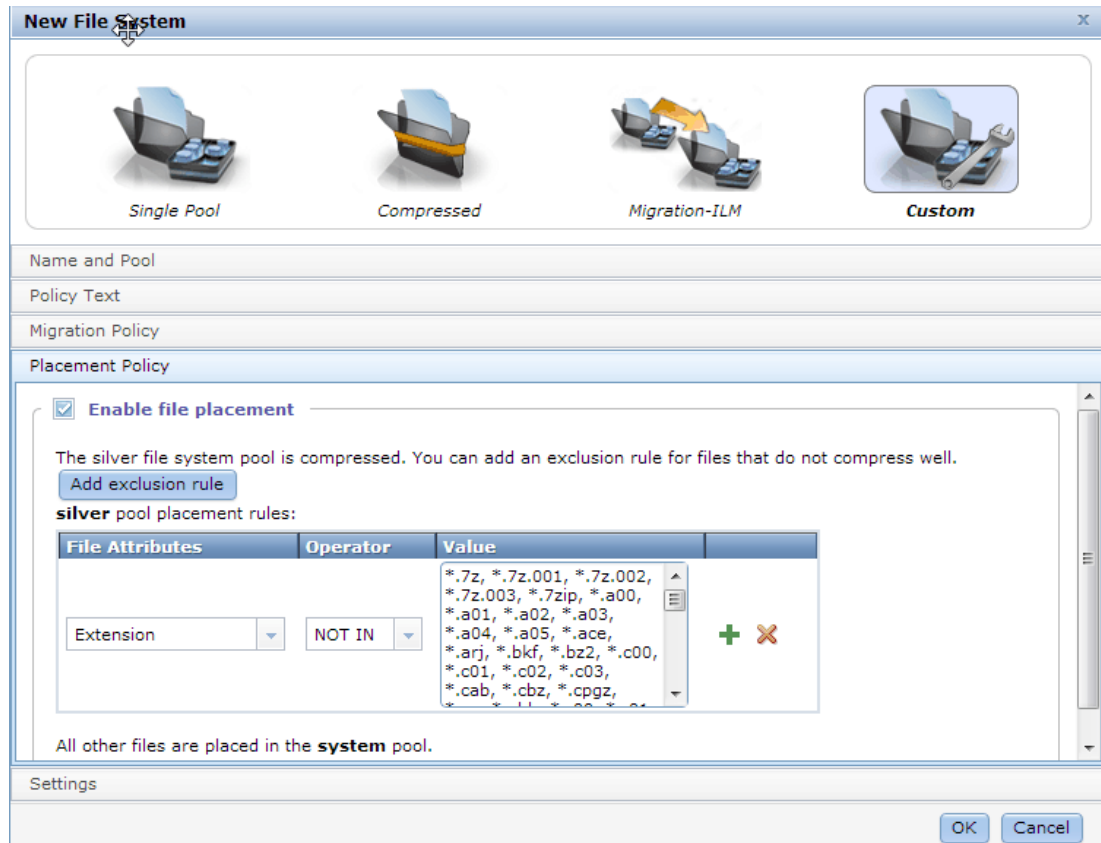


Figure 16-3 Defining an exclusion list of files that should not be placed on the compressed pool

## 5. Adapt placement policy to individual configuration

The list of file extensions that should be excluded from compression may need adapting for different reasons:

- Some file types are not compressible, but they are not listed in the compression exclusion list.

If you know that your environment contains files that are not compressible, because they are already compressed or they are encrypted, they should be part of the compression exclusion list. The Edit dialog of the Files → File System panel for the File system in question can be used to add the missing file types of the exclusion list. Adding the missing file types will prevent the system to try to compress file types which can not be compressed further.

- Some files are compressible, but are listed in the compression exclusion list

Some extensions may be used by different applications with different content. Possibly there are compressible files with an extension that is part of the compression exclusion list. In this case, the extension in question could be removed from the exclusion list using the Edit dialog of the Files → File System panel for the File system in question.

Saving the adapted placement policy will not move files to other pools. Moving files that had been previously misplaced is possible by a migration policy.

## 6. Complete the file system creation

Finally, the file system creation process is triggered by clicking OK and confirming the summary message. The file system creation dialog will provide the CLI commands which are called while the file system is created.

### 16.2.3 Fully Compressed File system

The simplest scenario using compression is a filesystem with only one file system pool. The data volumes of this single pool are compressed, while the metadata volumes are not compressed. This configuration should only be used when the type of data for this file system will compress well. Therefore it should only be used in very specific situations after careful data analysis in order to avoid unnecessary CPU cost trying to compress data that is not compressible. See Figure .

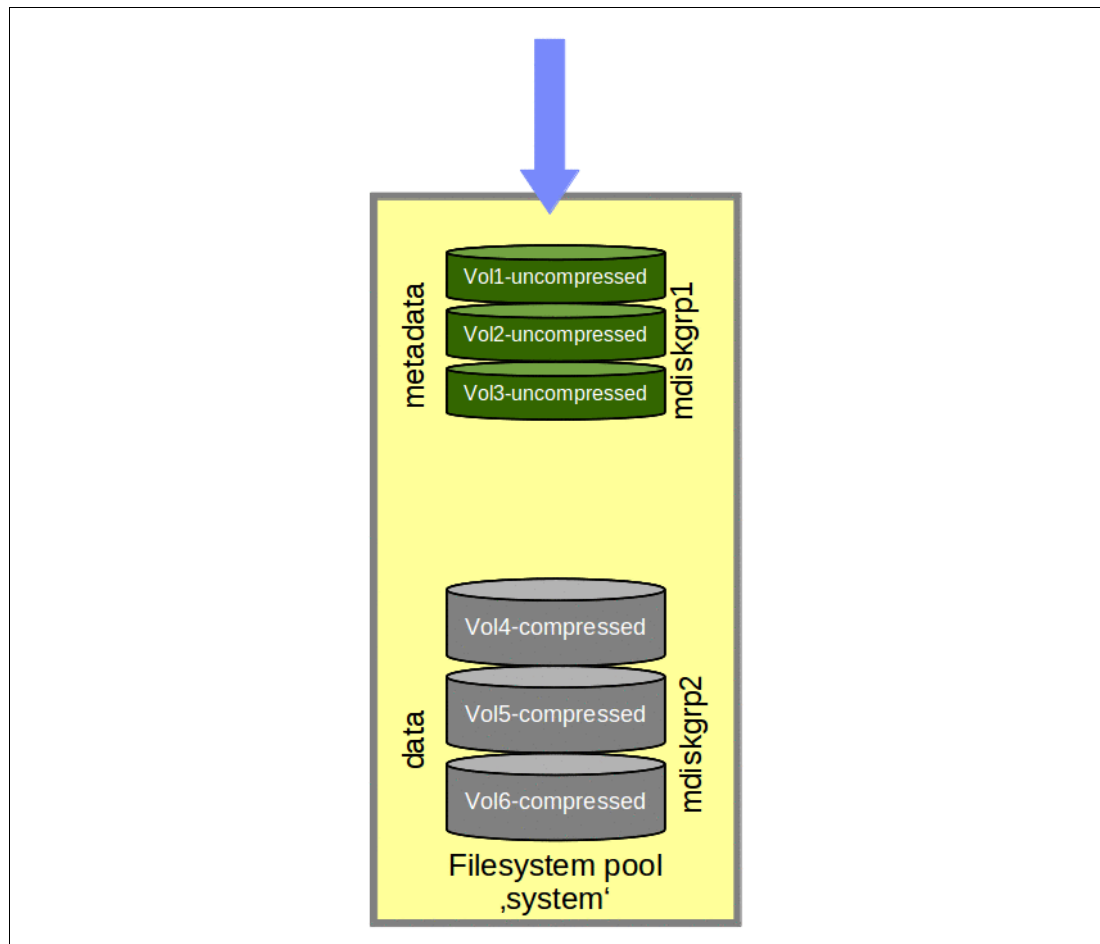


Figure 16-4 Single Pool scenario with compressed data volumes

While this is the simplest configuration, uncompressible files will be processed through the RtC compression engine.

This type of file configuration can be configured in the GUI “New File System → Compressed” panel, as shown in Figure 16-5 on page 280.

**New File System**

Single Pool      **Compressed**      Migration-ILM      Custom

\* **File system name:**  
fully\_compressedfs

\* **Owner:**  
root

**File system pool settings**

* Name:	* Storage pool:	Size:
system (not compressed)	pool2	50 GB
compressed	pool1	60 GB

⚠ Metadata replication is not configured. In order to enable metadata replication, two or more storage pools must be selected for the **system** file system pool.

OK Cancel

Figure 16-5 Using the compressed preset

A fully compressed file system can be created and only the metadata volumes are *not* compressed. The placement policy is created automatically with the extensions mentioned above.

## 16.2.4 Compression Rules by File Set

Placement rules can be applied to a whole file system or to a file set. A file set is a subtree of a file system namespace that in many respects behaves like a separate file system. File sets provide the ability to partition a file system to allow administrative operations at a finer granularity than the entire file system.

Placement rules per file set will allow different rules for parts of the file system. Shares that are created in the directory tree below these file sets will be subject to these special rules.

In the following complex scenario, we will have to create placement rules by directly entering placement rule text.

In order to define placement rules which are specific to a file set, the file sets can be dependent or independent.

### Example scenario

A placement rule for the whole file system like in the selected compressed configuration will ensure uncompressible files are excluded from the compressed pool.

The shares in fileset “uncompressed” do not use compression at all.

The shares in the fileset “special\_application” is used by an application that only writes two types of files. The “\*.cfg” files are control files and will be placed in the system pool, while the “\*.pak” files are highly compressible and are placed in the compressed pool. The file extensions used by this application are different than the ones in the built-in file placement policy in the Compressed preset, and special rules are needed to place the .cfg files in the compressed pool, and .pak files in the non-compressed pool.

***Steps to create the Example scenario:***

1. Create a compressed file systems with two pools

Use configuration as described in 16.2.1, “Selectively compressed file system with two pools” to create a file system and the default placement policy.

2. Create two file sets

Navigate to Files → File Sets and invoke the New Fileset dialog by clicking “+ New Fileset” in the filesets grid. A dialog as shown in Figure 16-6 will allow us to create file sets. Choose the file system mount point as parent path. The Owner should be a user that will own the path where the file set is linked in the file system tree. For CIFS shares this could be the domain administrator. The fileset name must be unique per file system. Placement rules can reference both dependent and independent file sets, in this example a dependent file set was chosen.

**New File Set**

**Basic** **Custom**

**\* Parent path:**

**\* Name:**

**Junction path:**

**\* Owner:**

**Comment:**

**Type:**  **Independent parent file set that dependent file set belongs to:**

Figure 16-6 Create filesets in the Files → File Sets → New File Set dialog.

### 3. Define placement rules

Use the “Edit” action on the File system in the Files → File System panel to open the file system edit dialog which allows to edit the placement rules.

The custom rules that we want to add can not be configured using the placement pool wizard. The Policy Text editor in the GUI allows to define the additional rules. The three additional rules have to be added *before* the default placement rule text which are already in place as we defined them in step 1). Once custom rules have been created using the policy text editor in the GUI, the placement policy editor can not be used any more. In our case, we will refer to the system pool as target pool for a special rule which is not supported in the placement policy editor. However, we created the default placement policy using the graphic editor first, which is very useful as it processes the list of known file type extension that do not compress well into a correct placement policy string. See Example 16-2 on page 283.



- ▶ All files that belong to the exclusion list as specified in the generatedPlacementRule1 will be written to the system pool. This means that the .img files which are not written to the fileset special application will not be compressed.
- ▶ All other files will be written to the compressed pool, due to the “NOT” expression in the generatedPlacementRule1.
- ▶ The final default placement rule for the system pool will never be evaluated, all remaining files are already covered by the generatedPlacementRule1

**Note:** A high number of file set specific placement rules could impact performance. While it is possible to create thousands of file sets per file system, the management of separate placement rules for each of them could become problematic. The first placement rule that matches to the file that is written will be applied, which means that all rules must be traversed for files that only match the default rule.

In case we want to start over with a new set of exclusion rules and need the placement policy editor back, we can apply a default placement policy to the system pool in the Policy Text editor as in Example 16-3 and the placement policy editor comes back.

*Example 16-3 Default placement policy*

---

```
RULE 'default' SET POOL 'system';
```

---



## 16.3 Capacity Planning

The capacity planning approach is determined by the file system storage pool architecture that has been chosen. The capacity planning approaches that are discussed here will be a fully compressed file system and a selectively compressed file system

### 16.3.1 Planning Capacity with the Comprestimator tool

The Comprestimator tool allows to assess the compressibility of data by scanning existing data on block storage level. The comprestimator will provide an average compressibility estimation of data. This information will provide insight on how much physical space is needed overall to store compressed data. The average compressibility of the data will provide the information on how to size the and configure a fully compressed file system. The Comprestimator tool is described in *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

### 16.3.2 Capacity planning for selectively compressed file systems

The sizing for the overall physical disk capacity that is needed for a selectively compressed file system can also be done with the output of the Comprestimator tool. However, in order to configure a selectively compressed file system additional analysis steps are needed.

In order to configure a selectively compressed file system, not only the average compression rate must be known, but also the distribution of files into the compressed and uncompressed pools. For best performance, files that yield compression savings below 40% would not be compressed. This rule is also applied by the exclusion list as described in 16.2.1, “Selectively compressed file system with two pools”.

The average compression rate can be gathered by analyzing external storage systems with the comprestimator tool, as discussed above. The final step needed to do a prediction on the correct file system pool sizes can be done by creating a small file system with two pools and the appropriate placement rules. This file system can then be populated with sample data, and the Storwize V7000 Unified reporting will provide the required information that is necessary to plan for the final file system. When following this procedure, files should not be deleted from the sample file system before the estimation is complete.

The Monitor → Capacity → File system pools screen as shown in Figure 16-8 provides the metrics needed to complete the planning:

- ▶ The *File System Used* capacity provides the information how much uncompressed data was placed into the two pools by the placement policy. The ratio of the two values for the compressed and system pool should be used to scale the file system pool sizes in the New File System creation dialog.
- ▶ The *Real Capacity* value provides the information how much capacity on real disk was allocated. For the uncompressed pool which is fully provisioned this value is identical to the file system pool size. However, for the compressed pool this value together with the *File System Used* value indicates how much real data on disk was used for the compressed pool. The Compression savings metric ignores zeros which are not accounted. So the ratio of Real Capacity to File system used provides a value that will determine how much space will be required on real disk relative to the file system size.
- ▶ The *Used* block volume capacity only accounts disk capacity that has been taken up by data already, while the *Real Capacity* block volume capacity accounts for the already allocated blocks on disk which belong to the volume. The *Used* capacity is only displayed

per volume and not per file system pool and can be viewed in the File > File Systems grid by viewing the NSD Properties. Because the “Used” capacity is smaller than the “Real” capacity and it will provide more accurate and even better reduction rates. The rsize value which determines the Real Capacity that is allocated for a given used capacity is set to 2% for volumes created in the GUI. Therefore the difference is small, and for practical sizing purposes the *Real Capacity* value is good enough and gives a small contingency, unless the rsize has been updated per CLI and is set to a much higher value.

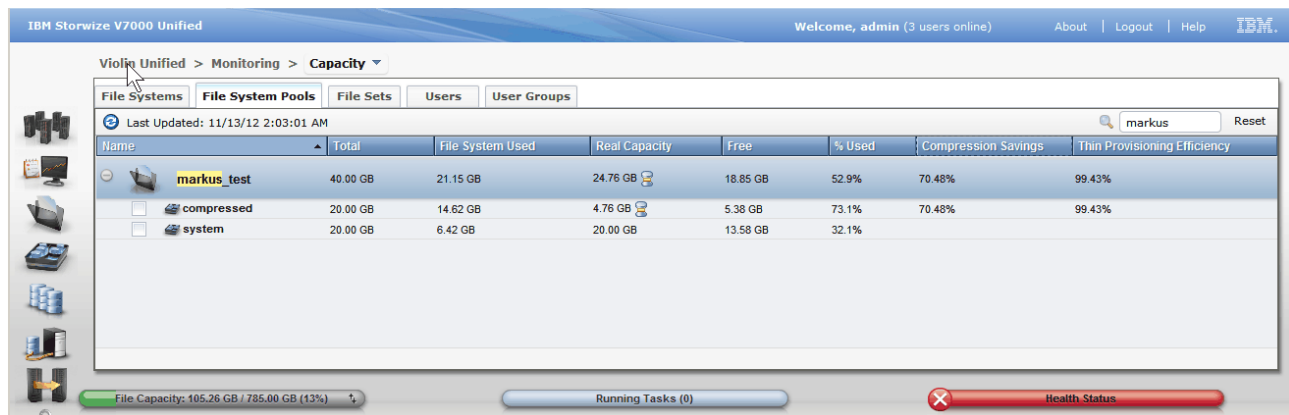


Figure 16-8 The Monitor → Capacity → File System Pools view can help with capacity planning based on a data sample

Example 16-4 shows the calculation based on Figure 16-8.

#### Example 16-4 Calculation

Planned file system size: 2000 GB

Sample file system overall used capacity: 21.15 GB

‘System’ file system pool:

Sample ‘system’ pool used capacity: 6.42 GB

File system pool capacity fraction: 6.42 GB/21.15 GB = 30.3 %

**Planned ‘system’ file system pool capacity: 2 TB \* 30.3 % = 606 GB**

Planned Physical disk for the system pool: 606 GB

‘Compressed’ file system pool:

Sample file system compressed pool used capacity: 14.62 GB

File System pool capacity fraction = 14.62 GB/21.15 GB = 69.1 %

**Planned ‘compressed’ file system pool capacity: 2 TB \* 69.1 % = 1382 GB**

Sample compressed pool Real Capacity: 4.76 GB

Reduction ratio (compression and thin provisioning): 4.76 GB/14.62 GB = 32.5 %

Physical disk for the compressed pool: 2000 GB \* 69.1 % \* 32.5 % = 449 GB

With the data as calculated in the example, finally the file system can be configured as shown in Figure 16-9. The expected uncompressed file system pool capacities are entered in the pool size entry fields. For the fully provisioned mdiskgroup the GUI will check if the capacity requirements are sufficient. For the compressed pool, the GUI will allow us to create the compressed file system pool if at least 25%% of the entered uncompressed file size is available. This assumes a compression savings of 75%, and for most cases this value is too optimistic. It is therefore recommended to ensure that available space in the mdiskgroup that will host the compressed pool is at least as large as the estimated needed physical disk capacity of the compressed pool. In the above estimation this is 449 GB.

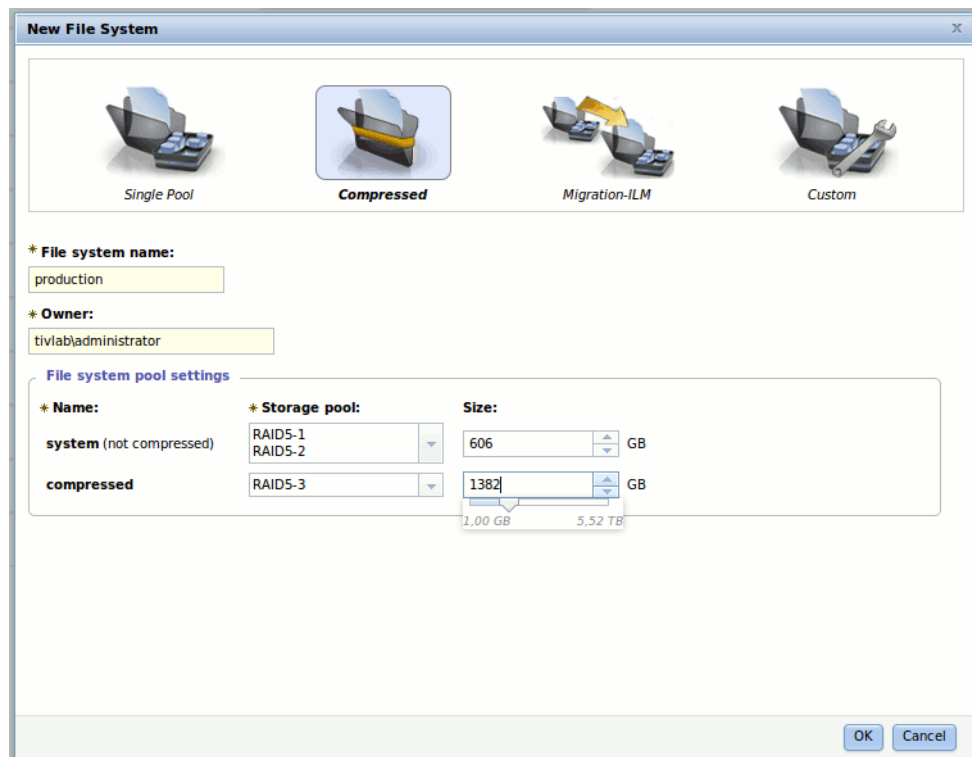


Figure 16-9 The evaluated file system pool sizes are entered to create a new file system

Notice the slider below the compressed file system pool size entry box. The slider maximum value allows us to configure four times the available physical size of the chosen storage pool.

## 16.4 Compression Metrics for File Systems

The Storwize V7000 Unified provides several metrics that are specific to file systems and compression. As the compression feature introduces the thin provisioning concept, viewing these metrics provide insight both on the compression aspects and on thin provisioning related to file systems usage data.

In Figure 16-10, the perspectives of the file system, the block device thin provisioning and the compression engine on a file system with a few files are compared.

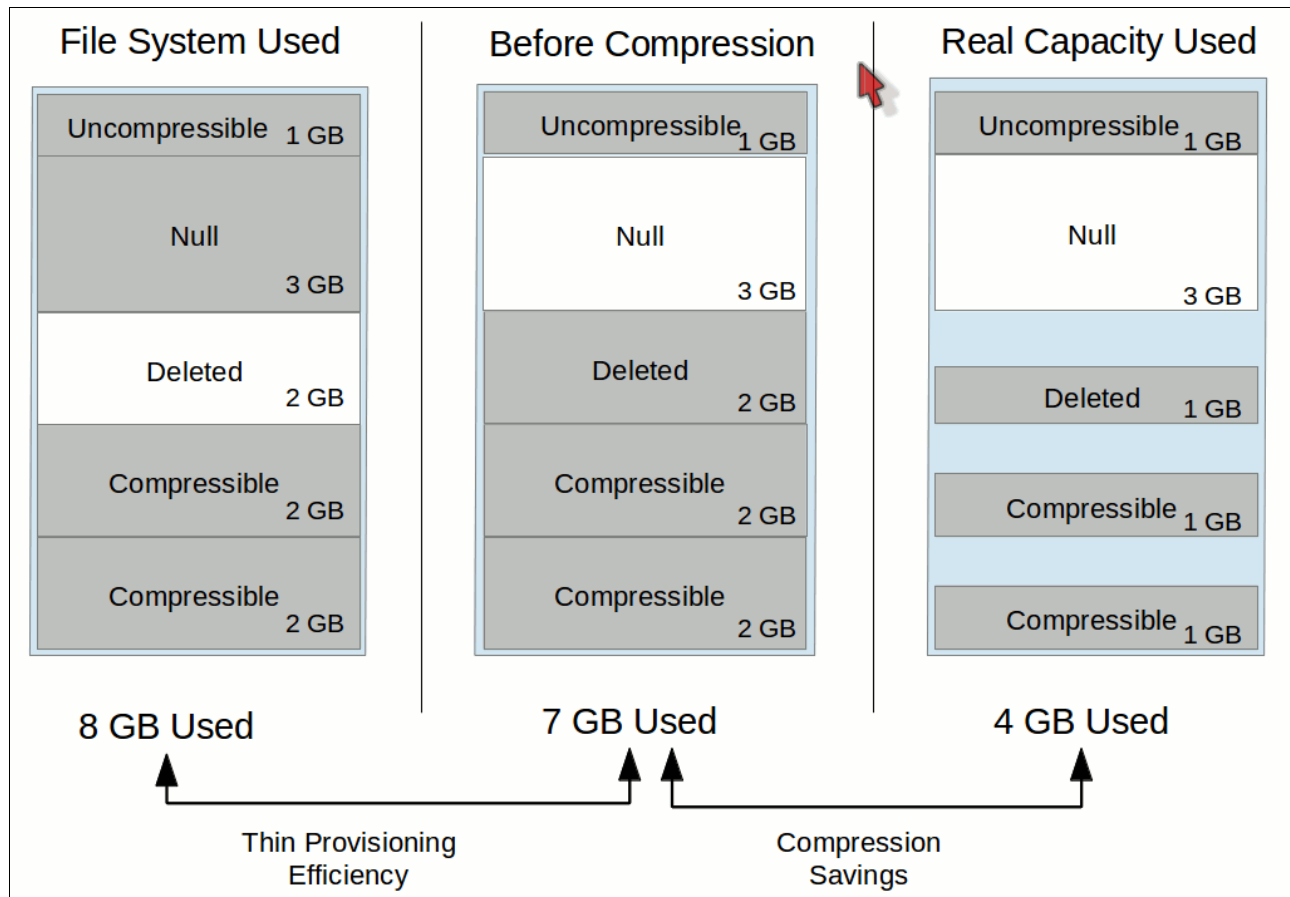


Figure 16-10 Capacity views from the file system, block device and compression engine perspective

The file system is aware of the uncompressed files as well as snapshots, clones and metadata it handles. The compressibility of files and the fact that files may contain zeros is not accounted for by the file system.

The block storage device layer handles zeros inside files in a space efficient way, as zeros are not written to disk and also not accounted for otherwise. The thin provisioning approach assumes all blocks that are not specifically written with non zero data contain zeros. Therefore, files that contain zeros like virtual machine images or empty database containers are not accounted on the block device.

The Thin Provisioning Efficiency that is shown on the Storwize V7000 Unified GUI in the Monitor → Capacity → File systems by Pools view shows per file system pool the ratio of the file system used capacity divided by the block device uncompressed used capacity. The following examples provide scenarios with expected ratios and potential user actions:

- ▶ A file system that is used as archive where no files have been deleted yet and files don't contain zeros would show a Thin Provisioning Efficiency of 100%
- ▶ A file system that has been just created and was filled with newly formatted database container files will show a thin provisioning efficiency of 100%. The file system fully accounts the file size, whereas the block devices saves space on the zeros which are also not accounted for. Theoretically the ratio between file system used capacity and before compression capacity is even higher than 100%, but the efficiency value is limited to 100%.

- A file system was chosen much larger than its current file system used capacity. Files are regularly created and deleted. The thin provisioning capacity will be much lower than 100%, because the block device still accounts for the files that have been deleted. Potentially, this file system was chosen too large and the file system size should be reduced to save space. Similar to a fully provisioned file system, the compressed pools of a file system should not be oversized to avoid wasted space.

## 16.5 Managing compressed file systems

In the following sections we discuss areas of interest for managing compressed file systems.

### 16.5.1 Adding a compressed pool to a file system

When a file system that is not using compression is to be converted to a file system that uses compression, several steps are needed after analysis of the data in the not compressed file system provided insight if the file system is a good candidate for compression. See the Planning section of this chapter of approaches to analyze data.

#### ***Create a compressed pool***

Another pool which has the compression flag activate is created using the GUI Edit dialog in Files → File Systems. The pool sizes should match the expected uncompressed capacity. At this point, it is not yet necessary to define a placement policy

#### ***Create and run a migration policy***

The compressible files in the system pool should be compressed. Therefore, a migration policy which moves the compressible files to the compressed pool has to be defined and executed. The screenshot on how to define a migration policy is shown in Figure 16-11 on page 290. The default list of non compressible file types or an adapted version that matches the system environment, as listed in Example 16-1 on page 277, can be entered in the exclusion rule for this migration policy. In this example the policy is executed automatically due to the low thresholds, as an upper threshold of 1% and a lower threshold of 0 was chosen.

The GUI policy editor was used to create the migration policy as it nicely formats the long list of extensions to exclude from migration into the correct policy language statements. Alternatively, the CLI can be used to create such a policy (which is more cumbersome), but also to run the policy using the **runpolicy** command which will execute the migration policy only once and at the time the command is started. In 16.5.2, “Making a selectively compressed file system uncompressed”, the creation and CLI based execution of a migration run is described.

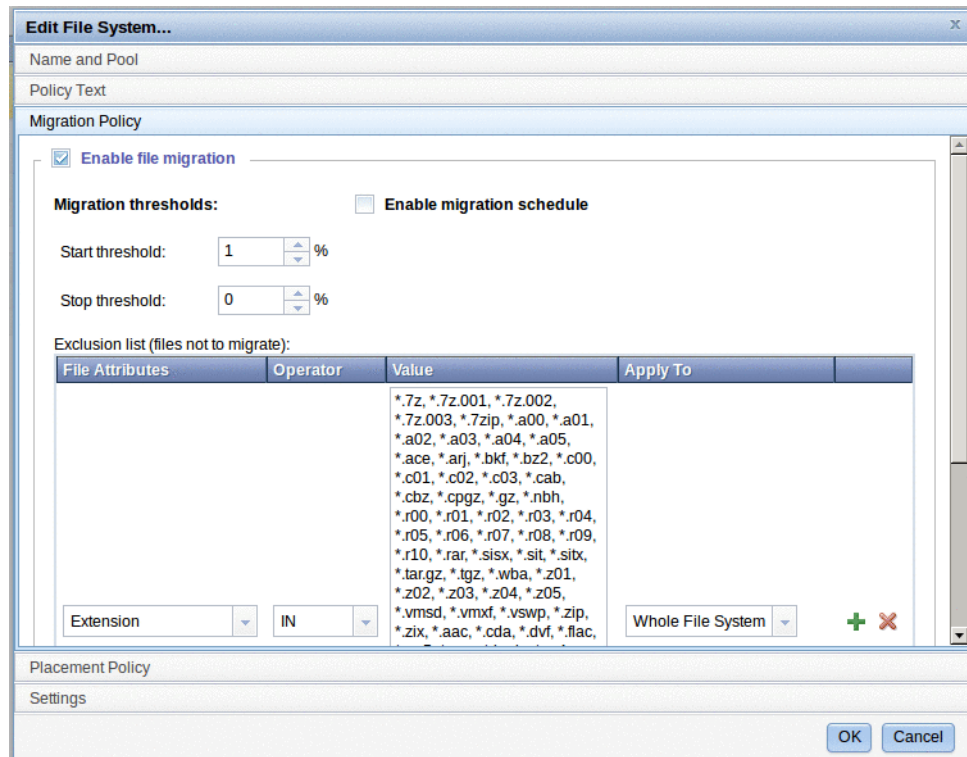


Figure 16-11 The GUI migration policy editor can be used to define which files should not be migrated to the compressed pool

As the threshold for the migration was set low, the system starts the migration job automatically. The CLI allows us to view the active jobs with the command `lsjobstatus` and also provides a view on the migration job results using the `showlog` command. See Example 16-5.

*Example 16-5 View the migration policy and monitor the migration job using the CLI*

```
# ***** View the GUI generated policy rule for migration
# ***** The exclusion list is truncated in the example
[7802378.ibm]$ lspolicy -D converttocompressed
RULE 'generatedMigrationRule0'
MIGRATE
  FROM POOL 'system'
  THRESHOLD(1,0)
  TO POOL 'compressed'
  WHERE NOT (
    (LOWER(NAME) LIKE '%.7z' OR LOWER(NAME) LIKE '%.7z.001'
    ... OR LOWER(NAME) LIKE '%.zip' OR LOWER(NAME) OR LOWER(NAME) LIKE '%.docx'))
RULE 'default' SET POOL 'system'
EFSSG1000I The command completed successfully.

# ***** List current active jobs, for all jobs used the --all keyword
[7802378.ibm]$ lsjobstatus
File system      Job              Job id Status  Start time      End
time/Progress RC Message
converttocompressed auto_migration 42   running 11/13/12 3:19:45 PM IST
EFSSG1000I The command completed successfully.
```

```
# ***** Show the detailed log of the job, the log can already be viewed
# ***** while it is still running. Log truncated in this example
[7802378.ibm]$ showlog 42
Primary node: mgmt002st001
Job ID : 42
[I] GPFS Current Data Pool Utilization in KB and %
compressed69121048576000.006592%
system79142401048576007.547607%
[I] 238008 of 12336640 inodes used: 1.929277%.
[I] Loaded policy rules from /var/mmfs/tmp/tspolicyFile.mmapplypolicy.943899.
Evaluating MIGRATE/DELETE/EXCLUDE rules with CURRENT_TIMESTAMP =
2012-11-13@13:19:47 UTC
parsed 1 Placement Rules, 0 Restore Rules, 1 Migrate/Delete/Exclude Rules,
  0 List Rules, 0 External Pool/List Rules
RULE 'generatedMigrationRule0'
  MIGRATE
    FROM POOL 'system'
    THRESHOLD(1,0)
    TO POOL 'compressed'
    WHERE NOT (
      (LOWER(NAME) LIKE '%.7z' OR LOWER(NAME) LIKE '%.7z.001' OR LOWER(NAME) LIKE
'%.hqx' OR LOWER(NAME) LIKE '%.docx'))
RULE 'default' SET POOL 'system'
[I] Directories scan: 223268 files, 10751 directories, 0 other objects, 0
'skipped' files and/or errors.
[I] Summary of Rule Applicability and File Choices:
Rule#Hit_CntKB_HitChosenKB_ChosenKB_IllRule
  0134646490876013464649087600RULE 'generatedMigrationRule0' MIGRATE FROM POOL
'system' THRESHOLD(1,0) TO POOL 'compressed' WHERE(.)
[I] Filesystem objects with no applicable rules: 99362.
[I] GPFS Policy Decisions and File Choice Totals:
  Chose to migrate 4908760KB: 134646 of 134646 candidates;
  Chose to premigrate 0KB: 0 candidates;
  Already co-managed 0KB: 0 candidates;
  Chose to delete 0KB: 0 of 0 candidates;
  Chose to list 0KB: 0 of 0 candidates;
  OKB of chosen data is illplaced or illreplicated;
Predicted Data Pool Utilization in KB and %:
compressed49156721048576004.687950%
system30234001048576002.883339%
[I] A total of 134646 files have been migrated, deleted or processed by an
EXTERNAL EXEC/script;
  11 'skipped' files and/or errors.
```

```
-----
End of log - auto_migration job ended
-----
```

```
EFSSG1000I The command completed successfully.
```

---

### ***Remove the migration policy and enable a placement policy***

The migration policy can now be disabled by unchecking the placement and a placement policy can be defined. For compressed file system pool, a placement policy can be created just like during initial creation of a file system with a compressed pool as described above.



The placement policy that will be enabled will ensure that only compressible files will be added to the new pool in future.

## 16.5.2 Making a selectively compressed file system uncompressed

It is possible to revert a selective compressed file system to become uncompressed while there is full data access. The following example shows which steps are needed to migrate a file system that has an uncompressed system pool and a second compressed filesystem pool to a fully uncompressed file system pool.

The steps needed to make a selectively compressed file system uncompressed are:

1. Analyze how much physical space is needed to move all data to the system pool. The conversion can be done most easily when the uncompressed space needed is available as free disk space in addition to the compressed pool. If this is not the case, a step by step approach must be taken and volumes from the compressed pool that is emptied must be removed.
2. Create a default placement rule for the file system. See Example 16-6.

*Example 16-6 The policy text to define a default placement rule for the file system that is converted to uncompressed*

---

```
RULE 'default' SET POOL 'system';
```

---

3. Create and run a migration rule that will move data to the system pool.

In order to migrate data back to the system pool, a simple policy rule that will move all data to the system pool from the compressed pool is created. A migration policy can be applied to a file system which will ensure the migration starts automatically and is regularly executed. This will happen if the policy is created in the GUI policy text editor. Because the migration rule should only run once, another approach is chosen. The CLI will be used to create a policy, start a policy run and monitor the migration of data.

Example 16-7 shows the simple policy rule.

*Example 16-7 Simple policy rule*

---

```
# Create a policy rule - this policy is not applied to any file system
[7802378.ibm]# mkpolicy migratetosystem -R "RULE 'migratetosystem' migrate from
pool 'compressed' to pool 'system';"
EFSSG1000I The command completed successfully.
```

```
# The runpolicy command will only run the policy once on our file system
[7802378.ibm]$ runpolicy markus_test -P migratetosystem
EFSSA0184I The policy is started on markus_test with JobID 806.
# The showlog command provides the output of policy runs
[7802378.ibm]$ showlog 806
Primary node: mgmt001st001
Job ID : 806
[I] GPFS Current Data Pool Utilization in KB and %
compressed 6531584 39845888 16.392116%
system 20494848 49283072 41.585979%
[I] 1000507 of 1200640 inodes used: 83.331140%.
[I] Loaded policy rules from
/var/opt/IBM/sofs/PolicyFiles/policy4252244384126468096.
Evaluating MIGRATE/DELETE/EXCLUDE rules with CURRENT_TIMESTAMP =
2012-11-15@07:39:05 UTC
```



```

parsed 0 Placement Rules, 0 Restore Rules, 1 Migrate/Delete/Exclude Rules,
  0 List Rules, 0 External Pool/List Rules
RULE 'migratetosystem' migrate from pool 'compressed' to pool 'system'
[I] Directories scan: 923603 files, 72924 directories, 0 other objects, 0
'skipped' files and/or errors.
[I] Summary of Rule Applicability and File Choices:
Rule# Hit_Cnt KB_Hit ChosenKB Chosen nKB_Ill Rule
  0574246 49452407 424649452400RULE 'migratetosystem' MIGRATE FROM POOL
'compressed' TO POOL 'system'
[I] Filesystem objects with no applicable rules: 422280.
[I] GPFS Policy Decisions and File Choice Totals:
Chose to migrate 4945240KB: 574246 of 574246 candidates;
Chose to premigrate 0KB: 0 candidates;
Already co-managed 0KB: 0 candidates;
Chose to delete 0KB: 0 of 0 candidates;
Chose to list 0KB: 0 of 0 candidates;
0KB of chosen data is illplaced or illreplicated;
Predicted Data Pool Utilization in KB and %:
compressed1586344398458883.981199%
system254728564928307251.686827%
[I] A total of 574246 files have been migrated, deleted or processed by an
EXTERNAL EXEC/script;
  0 'skipped' files and/or errors.

```

```

-----
End of log - runpolicy job ended
-----

```

EFSSG1000I The command completed successfully.

---

#### 4. Remove all volumes of the compressed pool

After the policy run has completed, all data was moved to the system pool and the volumes of the compressed pool can be removed. The GUI can be used to perform this task in the edit file system dialog. Setting the capacity of the compressed pool to 0 will remove all compressed volumes and finally remove the file system pool.

Figure 16-12 on page 294 shows how to remove them.

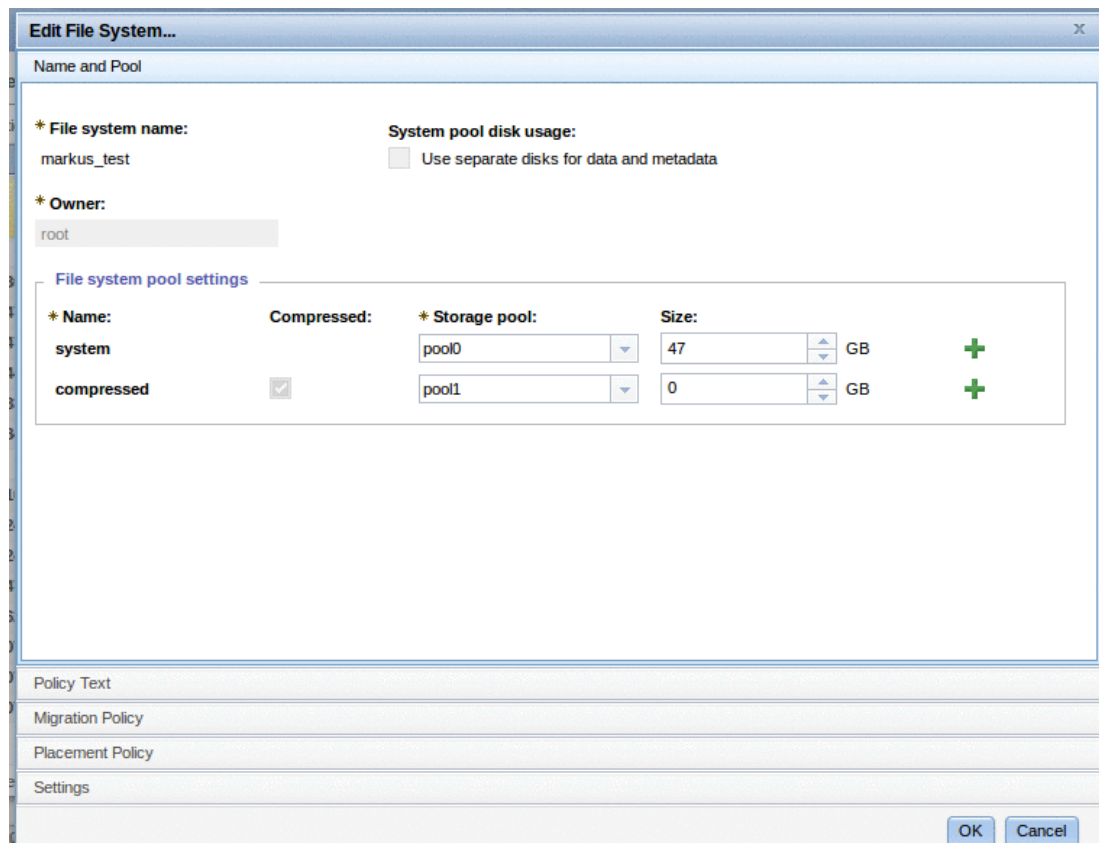


Figure 16-12 The compressed file system pool size is set to 0 which will trigger the removal of all the related compressed volumes

## 16.6 Compression Saving Reporting

Understanding capacity in a virtualized storage system can be a challenge. Probably the best approach is to review the various layers in storing data, from different points of view:

- ▶ Data: The data itself as stored by the host OS onto the volume.
- ▶ Shares: A logical unit that the user write the data into.
- ▶ File system: Logical unit in the Storwize V7000 Unified system. Composed from file System Pools.
- ▶ Pools: Composed from MDisk.
- ▶ File system pool: The storage container of compressed volumes

By design, the compression technology implemented in the Storwize V7000 Unified is transparent to the host. It compresses the client data before writing to the disk, and extracts the data as it is read by the host. The data size looks different from a different point of view. For example, the compressed size is not reflected to the user, and can be seen only from the Storwize V7000 Unified points of view.

### 16.6.1 Reporting basic overview

The following are basic concepts of the Storwize V7000 Unified system reporting of compressed data:

**Real Capacity** Real size is the amount of space from the storage pool allocated to allow the volume data to be stored. A compressed volume is by default a thin-provisioned volume that allows you to allocate space on demand. When a new volume is created there is an option to define the actual size it creates as a percentage of the original volume size. By default it is 2%. The volume expands automatically according to the usage.

**Used Capacity** The amount of real size that is used to store data for the volume, which is sometimes called *compressed size*.

**Before Compression size** The size of all the data that has been written to the volume calculated as though it was written without compression. This size is reflected on the host operating system.

**Virtual capacity** Virtual capacity is the volume storage capacity that is available to a host. Real capacity is the storage capacity that is allocated to a volume copy from a storage pool. In a fully allocated volume, the virtual capacity and real capacity are the same. In a thin-provisioned or compressed volumes, however, the virtual capacity can be much larger than the real capacity.

**Compression Savings** The ratio between the compressed size and the uncompressed size.

## 16.6.2 Reporting of compression in the GUI

Reporting of compressed data is available in the GUI in the following screens:

**System level** Monitoring → System

**File system level** Monitoring → Capacity and in files and file systems

**Storage pool level** Pools--> MDisk by pools

**Volume level** Volumes by pool or File system page.

### System level

To view the file system level, click on Monitoring → System.

Figure 16-13 on page 296 shows three areas of the system disks:

- ▶ Compressed size
- ▶ Before compression size
- ▶ Virtual capacity

As you can see in this figure by looking at the tube to the left of the Storwize V7000 Unified, the compression ratio is approximately 50%, the second area of the “before compression capacity” is greater than the lower section of the compressed capacity.

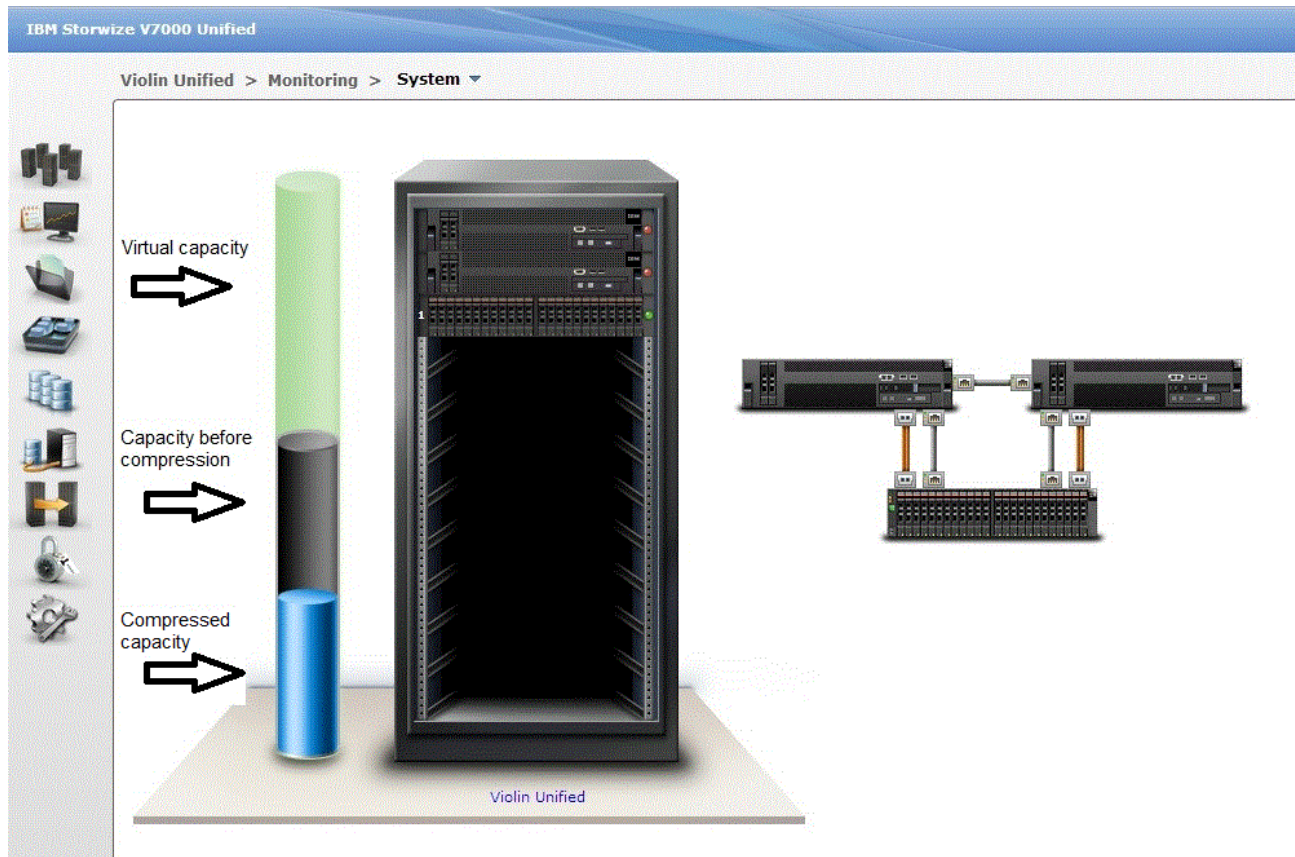


Figure 16-13 Areas of the system disks

**Note:** The reporting view from the IBM Storwize V7000 Unified system is different than the reporting from the Storwize V7000. For more information about the Storwize V7000 reporting refer to *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

<http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/redp4859.html?Open>

### File system level

To get information at the file system level, click on the Monitoring icon and choose Capacity option. Choose File system Pools tab and you can see the compression saving of each file system and the amount of used space from each file system pool.

In Figure 16-14 you can see the real capacity, filesystem used capacity and the compression saving capacity of each filesystem. By expanding the plus sign you can see it per file system pool.

In this example in the 'compressed\_fs' file system you can see that the *Real capacity* of 'system' and the Total capacity is equal. This is due to the fact that un-compressed file system are fully allocated. But these values of the compressed pool are different. The *Real capacity* is the allocated size from the pool, in this case 22.19GB.



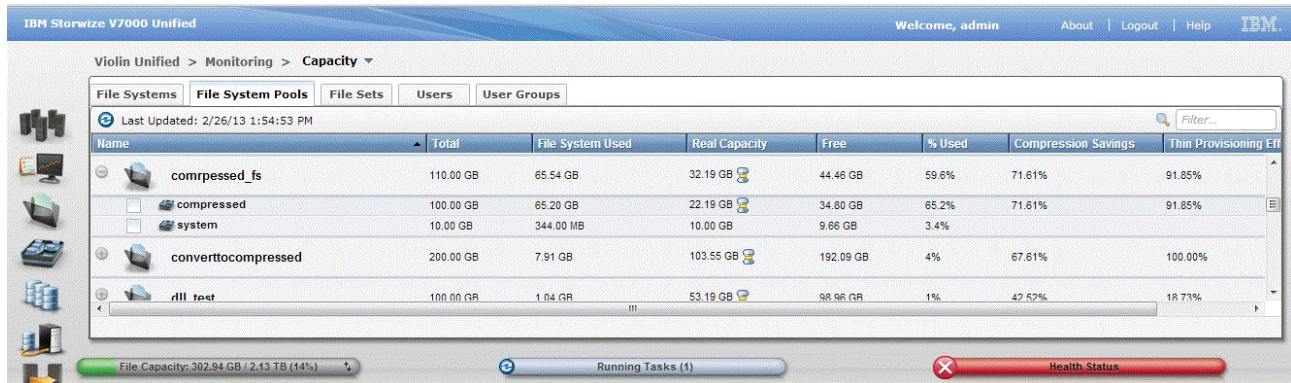


Figure 16-14 Real capacity, filesystem used capacity and the compression saving capacity of each filesystem

## Storage pool level

To get information at the storage pool level, click Mdisks by pool in the pools icon menu. Clicking this provides the compression ratio benefits of the entire pool.

**Notice:** The pool contains volumes that are used by the file system of the IBM Storwize V7000 Unified system and VDisks of the Storwize V7000 block devices.

Figure 16-15 shows the amount of used capacity and the virtual capacity.

**Used size:** The percentages in the shaded blue part are the amount of used space from the allocated capacity. In this example 1TB/2.98TB is used, which means 47% of pool0 is used.

**Virtual Capacity:** The numbers are also present the virtual capacity. This is the sum of all volume sizes in the pool. In this example, pool0 has 4.08 of virtual capacity.

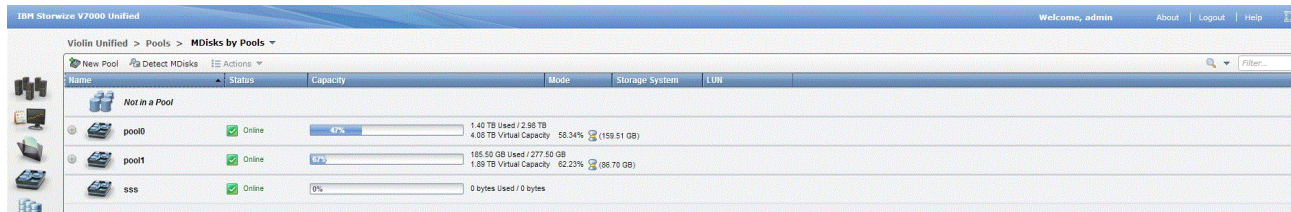


Figure 16-15 Amount of used capacity and the virtual capacity

**Note:** The file system is composed of volumes. Therefore this view includes the files system volumes and the block device volumes that are using the same storage pool.

## Volume level

There are two methods to report at a volume level.

**Method 1:** Click on **Files** icon and navigate to **File systems**.

Expand the file system and the file system pools in order to see the volumes it is composed of. See Figure 16-16.

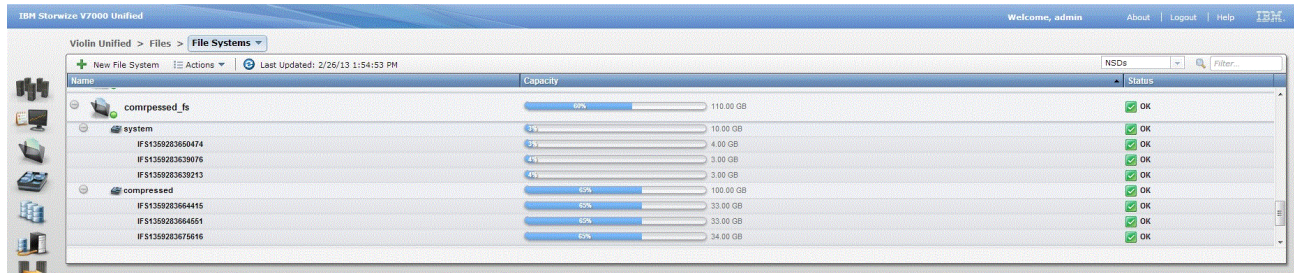


Figure 16-16 Volumes used capacity

The view in Figure 16-16 shows the used capacity of each volume in the file system per file system pool.

In order to see more details, right click on a volume and choose *properties*.

In Figure 16-17 you can see the before compression size, real size and the compression saving of the volume. In this example, the original data size is 23.45GB and it got compressed to 6.67GB. Therefore the compression saving is 71.57%.

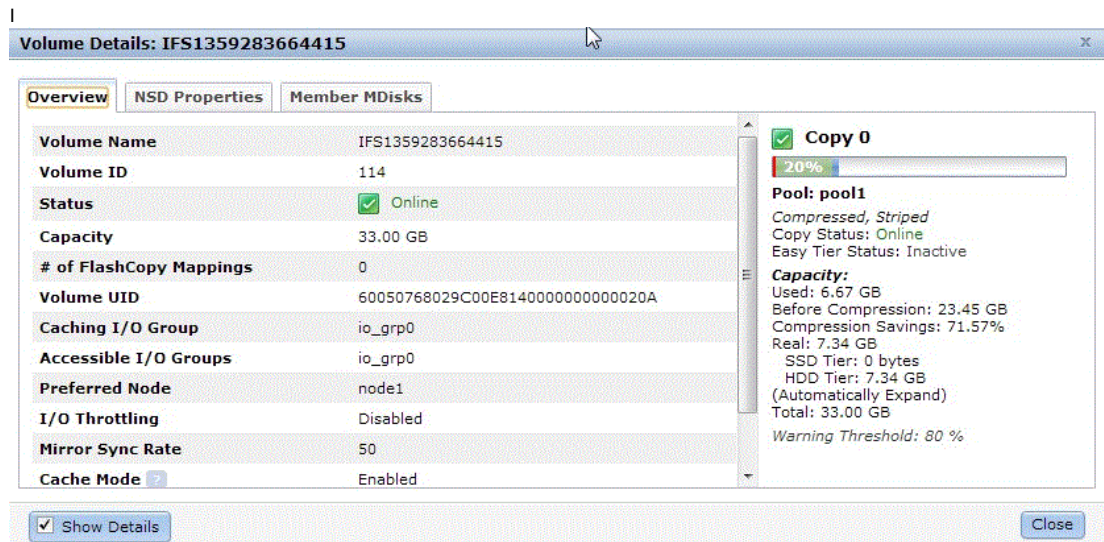


Figure 16-17 Before compression size, real size and the compression saving of the volume

Make sure the *Show Details* check box is checked.

**Notice:** The compressed data is divided equally between the volumes in the same file system pool so the compression saving shown in one volume is reflected in the entire file system pool.

**Method 2:** Click on the **Pools** icon and navigate to **Volumes by pool**. The table provides the compression saving per volume in a specific pool. This view is more useful for block device volumes rather than file systems.

In the graph at the right upper side of Figure 16-18 you can see the volume capacity - 5.41TB and the virtual capacity - 10.82TB. The volumes that configured in this pool are over allocated and the total size of them is 10.82TB.

The second bar reports the compression savings. The compressed size of the data is 4.3GB and the original size of the data is the compressed size + saved size - 4.3 + 5.3 = 9.6GB. The



compression ratio of the entire pool0 is 44% and calculated with this formula: compressed size/total size - 4.3/9.6. = 44% The saving ratio is ~55%.

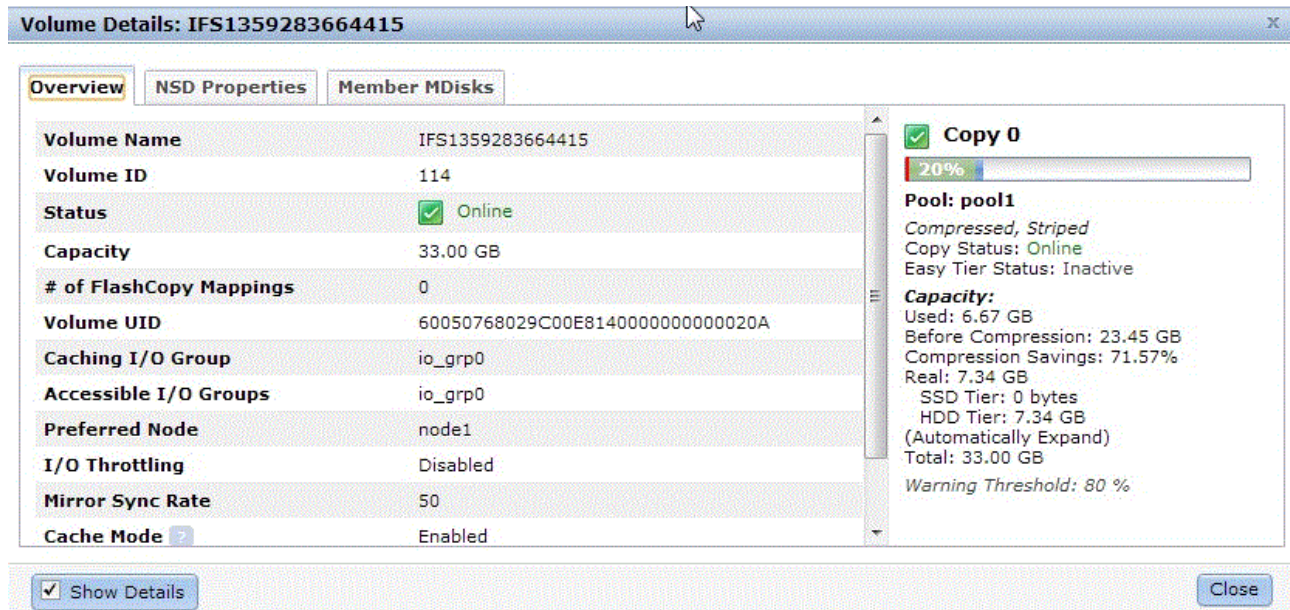


Figure 16-18 Volume capacity - 5.41TB and the virtual capacity - 10.82TB

**Consideration:** The reported numbers are dynamically updated as data is sent from the Storwize V7000 cache to the RACE component. This updating causes a slight delay (typically a few seconds) between the host writes and the updated reporting

### 16.6.3 Compression reporting using the CLI

The CLI commands provides information in a volume and storage pool level. Therefore, in order to see the capacity reporting of a file system through the CLI, you should know the VDisks name the file system is composed of.

You can generate reports on compressed data by using the following CLI commands :

**System level:** `lssystem`

**Specific storage pool:** `lsmdiskgrp [<mdisk_grp_name>]`

**All storage pools:** `lsmdiskgrp`

**Specific volume:** `lssevdiskcopy [<vdisk_name>]`

**All volumes in a pool:** `lssevdiskcopy -filtervalue  
mdisk_grp_name=[<mdisk_grp_name>]`

**All volumes in a system:** `lssevdiskcopy`

Example 16-8 shows the output of the `lssystem` command.

*Example 16-8 lssystem output*

```
[7802378.ibm]$ lssystem
id 00000200A7003A05
name Violin Unified
```

```

location local
partnership
bandwidth
total_mdisk_capacity 3.3TB
space_in_mdisk_grps 3.3TB
space_allocated_to_vdisks 1.57TB
total_free_space 1.7TB
total_vdiskcopy_capacity 5.97TB
total_used_capacity 1.48TB
total_overallocation 183
total_vdisk_capacity 5.95TB
total_allocated_extent_capacity 1.58TB
[...]
compression_active yes
compression_virtual_capacity 4.58TB
compression_compressed_capacity 166.52GB
compression_uncompressed_capacity 412.73GB
[...]

```

---

Example 16-9 shows the output of the **lsmdiskgrp** command for a specific pool:

*Example 16-9 lsmdiskgrp output*

---

```

[7802378.ibm]$ lsmdiskgrp pool0
id 0
name pool0
status online
mdisk_count 2
vdisk_count 129
capacity 2.98TB
extent_size 256
free_capacity 1.58TB
virtual_capacity 4.08TB
used_capacity 1.34TB
real_capacity 1.40TB
overallocation 136
warning 80
easy_tier auto
easy_tier_status inactive
tier generic_ssd
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier generic_hdd
tier_mdisk_count 2
tier_capacity 2.98TB
tier_free_capacity 1.58TB
compression_active yes
compression_virtual_capacity 2.78TB
compression_compressed_capacity 113.89GB
compression_uncompressed_capacity 273.40GB

```

---

Example 16-10 shows the output of the **lsmdiskgrp** command for the entire node.



*Example 16-10 lsmdiskgrp output for entire node*


---

```
[7802378.ibm]$ lsmdiskgrp
0 pool0 online 2 129 2.98TB 256 1.58TB 4.08TB 1.34TB 1.40TB 136 80 auto inactive yes 2.78TB 113.89GB
```

---

Example 16-11 shows the lssevdiskcopy output.

*Example 16-11 lssevdiskcopy -nohdr output*


---

```
[7802378.ibm]$ lssevdiskcopy -nohdr -filtervalue mdisk_grp_name=pool0
0 Nitzan 0 0 pool0 400.00MB 4.63MB 24.00MB 19.38MB 1666 on 80 no yes 4.44MB
1 CSB-SYSTEM_AG_COMPRESSED 0 0 pool0 50.00GB 21.24GB 22.25GB 1.02GB 224 on 80 no yes 49.96GB
3 bosmatTest 0 0 pool0 100.00GB 15.10GB 17.11GB 2.01GB 584 on 80 no yes 28.04GB
12 VM_Compressed 0 0 pool0 34.00GB 7.09GB 7.78GB 711.68MB 436 on 80 no yes 15.56GB
18 export for esx 200 0 0 pool0 100.00GB 3.25MB 2.02GB 2.01GB 4961 on 80 no yes 0.00MB
19 New_VM_Comp 0 0 pool0 300.00GB 22.12GB 28.13GB 6.01GB 1066 on 80 no yes 48.63GB
20 New_VM_Clear 0 0 pool0 100.00GB 33.96GB 35.96GB 2.00GB 278 on 80 256 yes no 33.96GB
25 Nitzan_01 0 0 pool0 400.00MB 0.75MB 17.25MB 16.50MB 2318 on 80 256 yes no 0.75MB
26 VM_Comp_PO_0 0 0 pool0 150.00GB 5.19GB 8.20GB 3.01GB 1828 on 80 no yes 11.21GB
27 VM_Comp_PO_1 0 0 pool0 150.00GB 5.19GB 8.20GB 3.01GB 1829 on 80 no yes 11.21GB
28 VM_Comp_PO_2 0 0 pool0 150.00GB 5.20GB 8.20GB 3.01GB 1829 on 80 no yes 11.21GB
29 VM_Comp_PO_3 0 0 pool0 150.00GB 5.19GB 8.20GB 3.01GB 1829 on 80 no yes 11.21GB
30 VM_Comp_PO_4 0 0 pool0 150.00GB 5.19GB 8.20GB 3.01GB 1828 on 80 no yes 11.21GB
36 Copy Mirror for Pendulum 0 0 pool0 500.00GB 494.69MB 10.48GB 10.00GB 4768 on 80 no yes 1.31GB
50 IFS1352811948934 0 0 pool0 33.00GB 521.38MB 1.18GB 683.38MB 2804 on 80 no yes 1.57GB
59 IFS1352811949068 0 0 pool0 33.00GB 513.53MB 1.18GB 691.21MB 2804 on 80 no yes 1.56GB
60 IFS1352812042137 0 0 pool0 34.00GB 516.28MB 1.20GB 708.93MB 2841 on 80 no yes 1.55GB
85 lev_production_backup 0 0 pool0 500.00GB 17.49GB 27.50GB 10.02GB 1817 on 80 no yes 26.83GB
91 IFS1352369292204 0 0 pool0 1.00GB 3.31MB 36.48MB 33.17MB 2806 on 80 no yes 0.00MB
92 IFS1352369292340 0 0 pool0 1.00GB 3.31MB 36.48MB 33.17MB 2806 on 80 no yes 0.00MB
93 IFS1352369304265 0 0 pool0 2.00GB 3.31MB 56.96MB 53.65MB 3595 on 80 no yes 0.00MB
94 jq_vol1 0 0 pool0 5.00GB 3.25MB 118.40MB 115.15MB 4324 on 80 no yes 0.00MB
95 jq_vol2 0 0 pool0 5.00GB 8.19MB 118.40MB 110.21MB 4324 on 80 no yes 4.94MB
100 IFS1352638421928 0 0 pool0 33.00GB 972.16MB 1.61GB 681.28MB 2043 on 80 no yes 14.03GB
101 IFS1352638422076 0 0 pool0 33.00GB 972.19MB 1.61GB 681.09MB 2043 on 80 no yes 14.04GB
102 IFS1352638501197 0 0 pool0 34.00GB 988.22MB 1.65GB 701.51MB 2060 on 80 no yes 14.27GB
```

---

For further information about capacity reporting of block device disks refer to *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

<http://www.redbooks.ibm.com/redpapers/pdfs/redp4859.pdf>



# 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. Note that some publications referenced in this list might be available in softcopy only.

- ▶ *Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933
- ▶ *Implementing the IBM Storwize V7000 V6.3*, SG24-7938
- ▶ *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859
- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *IBM System Storage: Implementing an IBM SAN*, SG24-6116
- ▶ *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363
- ▶ *IBM System Storage Business Continuity: Part 1 Planning Guide*, SG24-6547
- ▶ *IBM System Storage Business Continuity: Part 2 Solutions Guide*, SG24-6548
- ▶ *Get More Out of Your SAN with IBM Tivoli Storage Manager*, SG24-6687
- ▶ *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848
- ▶ *DS8000 Performance Monitoring and Tuning*, SG24-7146
- ▶ *Monitoring Your Storage Subsystems with TotalStorage Productivity Center*, SG24-7364
- ▶ *Using the SVC for Business Continuity*, SG24-7371
- ▶ *SAN Volume Controller: Best Practices and Performance Guidelines*, SG24-7521
- ▶ *SAN Volume Controller V4.3.0 Advanced Copy Services*, SG24-7574
- ▶ *IBM XIV Storage System: Architecture, Implementation and Usage*, SG24-7659
- ▶ *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725
- ▶ *IBM SAN Volume Controller 4.2.1 Cache Partitioning*, REDP-4426

## Other publications

These publications are also relevant as further information sources:

- ▶ *IBM System Storage SAN Volume Controller: Planning Guide*, GA32-0551
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Planning Guide*, GA22-1052
- ▶ *IBM System Storage SAN Volume Controller: Service Guide*, GC26-7901
- ▶ *IBM System Storage SAN Volume Controller Model 2145-8A4 Hardware Installation Guide*, GC27-2219

- ▶ *IBM System Storage SAN Volume Controller Model 2145-8G4 Hardware Installation Guide*, GC27-2220
- ▶ *IBM System Storage SAN Volume Controller Models 2145-8F2 and 2145-8F4 Hardware Installation Guide*, GC27-2221
- ▶ *IBM SAN Volume Controller Software Installation and Configuration Guide*, GC27-2286
- ▶ *IBM System Storage SAN Volume Controller Command-Line Interface User's Guide*, GC27-2287
- ▶ *IBM System Storage Master Console: Installation and User's Guide*, GC30-4090
- ▶ *Multipath Subsystem Device Driver User's Guide*, GC52-1309
- ▶ *IBM System Storage SAN Volume Controller Model 2145-CF8 Hardware Installation Guide*, GC52-1356
- ▶ *IBM System Storage Productivity Center Software Installation and User's Guide*, SC23-8823
- ▶ *IBM System Storage Productivity Center Introduction and Planning Guide*, SC23-8824
- ▶ *Subsystem Device Driver User's Guide for the IBM TotalStorage Enterprise Storage Server and the IBM System Storage SAN Volume Controller*, SC26-7540
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Installation Guide*, SC26-7541
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Service Guide*, SC26-7542
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Configuration Guide*, SC26-7543
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Command-Line Interface User's Guide*, SC26-7544
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: CIM Agent Developers Reference*, SC26-7545
- ▶ *IBM System Storage Open Software Family SAN Volume Controller: Host Attachment Guide*, SC26-7563
- ▶ *Command-Line Interface User's Guide*, SC27-2287
- ▶ *IBM System Storage Productivity Center User's Guide Version 1 Release 4*, SC27-2336
- ▶ *IBM TotalStorage Multipath Subsystem Device Driver User's Guide*, SC30-4096
- ▶ *IBM System Storage SAN Volume Controller V5.1.0 - Host Attachment Guide*, SG26-7905
- ▶ *IBM Tivoli Storage Productivity Center IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide*, SC27-2337
- ▶ *IBM TotalStorage Multipath Subsystem Device Driver User's Guide*, SC30-4096

## Online resources

These websites are also relevant as further information sources:

- ▶ IBM TotalStorage home page  
<http://www.storage.ibm.com>
- ▶ SAN Volume Controller supported platform  
<http://www-1.ibm.com/servers/storage/support/software/sanvc/index.html>
- ▶ Download site for Windows Secure Shell (SSH) freeware  
<http://www.chiark.greenend.org.uk/~sgtatham/putty>
- ▶ IBM site to download SSH for AIX  
<http://oss.software.ibm.com/developerworks/projects/openssh>
- ▶ Open source site for SSH for Windows and Mac  
<http://www.openssh.com/windows.html>
- ▶ Cygwin Linux-like environment for Windows  
<http://www.cygwin.com>
- ▶ IBM Tivoli Storage Area Network Manager site  
<http://www-306.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageAreaNetworkManager.html>
- ▶ Microsoft Knowledge Base Article 131658  
<http://support.microsoft.com/support/kb/articles/Q131/6/58.asp>
- ▶ Microsoft Knowledge Base Article 149927  
<http://support.microsoft.com/support/kb/articles/Q149/9/27.asp>
- ▶ Sysinternals home page  
<http://www.sysinternals.com>
- ▶ Subsystem Device Driver download site  
<http://www-1.ibm.com/servers/storage/support/software/sdd/index.html>
- ▶ IBM TotalStorage Virtualization home page  
<http://www-1.ibm.com/servers/storage/software/virtualization/index.html>
- ▶ SVC support page  
<http://www-947.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=4&brandind=5000033&familyind=5329743&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1>
- ▶ SVC online documentation  
<http://publib.boulder.ibm.com/infocenter/svcic/v3r1m0/index.jsp>
- ▶ IBM Redbooks publications about SVC  
<http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=SVC>

## Help from IBM

IBM Support and downloads

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

IBM Global Services

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

# Index

## A

Access 13, 95  
 access authentication 160  
 Access control 29, 198  
 access control 16, 28, 32, 37, 39–40, 42, 59  
 access control entries 46  
 access control list 42  
 Access Control Lists 15  
 access control lists 33  
 access control mechanism 42  
 access levels 95  
 access protocol 96  
 access to files 22  
 Accessed Time stamp 30  
 ACE 46, 72, 229  
 ACL 15, 33, 39, 42  
 ACL formats 42  
 ACL mapping 42  
 ACL type 46  
 ACLs 29, 38  
 Active Cloud Engine 72, 229  
 active cluster 225  
 Active Directory 41, 43, 46, 111, 119, 161, 163, 182  
 Active Directory domain 43  
 Active Directory ID mappings 46  
 Active Directory Server 38–39  
 Active Directory server 162  
 active mode 18–19  
 active quorum 225  
 active-active 32  
 AD 31, 41  
 AD authentication 43  
 AD with SFU 44  
 adaptive 71  
 Additional Gateways 170  
 address space 13  
 admin password 263  
 air circulation 138  
 airflow 138  
 alerting 96, 178, 261  
 alerting tasks 246  
 alerts 96, 220, 244  
 analyze workloads 101  
 animated graphics 89  
 Antivirus 72, 94, 208  
 Antivirus Definition 214  
 Antivirus scanner 73  
 Antivirus vendor product 210  
 AOS 173, 227, 264  
 API 18  
 Application Programming Interface 18  
 Archive bit 29  
 Arrays 192  
 assert 226

Assist on site 173  
 Assist-on-Site 262  
 asynchronous 14, 78  
 asynchronous replication 80  
 asynchronous scheduled process 27  
 attention lights 153  
 Audit Log 95  
 authenticated 38  
 authentication 14, 29, 37, 111, 161  
 authentication method 96  
 authentication methods 39  
 authentication server 43  
 authentication support 17  
 authority 33, 188  
 authorization 29, 37  
 automated data placement 61  
 automated storage tiering 27  
 automatic extent level migration 26  
 automatic recovery 242  
 automatically detect events 261  
 auxiliary 77  
 auxiliary VDisk 77  
 AV connector 209–210  
 availability 60  
 available scan nodes 212  
 AV-Connector 208

## B

background copy 76–78  
     rate 78  
 background copy rate 194  
 Backup 194, 197  
 backup 77–78  
 backup command output log 224  
 backup file 226  
 backup process 226  
 backup workload 237  
 backups from TSM 232  
 bandwidth 209  
 bandwidth friendly 80  
 bandwidth setting 193  
 base addresses 150  
 Base configuration 155  
 base file system 197  
 base licences 104  
 Batch Scan 216  
 batch scan 209  
 battery backup 23  
 bitmaps 76  
 blank carriers 136  
 block 8  
 block commands 97  
 block devices 191  
 block remote copy services 193

- block size 68
- block storage 8, 146
- block-level changes 80
- blocks 8
- bootable DVD 148, 265
- buffer resources 50
- BypassTraversalCheck 47
- byte range locking 32
- byte range locks 32
- byte-range 71
- byte-range locking 33

## C

- cable 136
- cable management system 140
- cache 227
- cache data 16
- caching 32, 71–72, 110
- call home process 261
- candidate 152
- canister 141
- canisters 23
- Capacity 93
- capacity 103
  - thin provisioned volumes 77
- Capacity Magic 101, 103
- capacity requirements 103
- Cascaded FlashCopy 77
- case insensitive file lookup 29
- centralized management concept 62
- Certificate 163
- chain 118
- chains 141
- change rate 79
- change the cluster name 171
- change the passwords 186
- Change Volumes 76, 79
- checklist 134
- checkpoint 227
- checkpoint data 224
- child directory 46
- child file 46
- CIFS 10, 17
- CIFS ACLs 117
- CIFS share 117, 202
- CIFS Windows 198
- CKD 8
- class 94
- cleaning rate 194
- clear text 39
- CLI 210
- CLI session 97
- client access 214
- Client Replaceable Unit 242
- client server infrastructures 10
- client server protocol 15
- client side caching 30, 110
- client side firewalls 19
- client systems 71
- Clone 194–195
- clone 76
- clones 106
- close after write 209
- cluster 26, 78
- cluster admin password 187
- cluster manager 27, 32–33, 71
- cluster messages 181
- cluster nodes 68
- cluster performance graphs 222
- clustered 15, 29
- Clustered File Modules 60
- clustered file system 66
- clustered implementation 32
- Clustered Trivial Data Base 29
- clustering 25, 32
- code levels 147
- Command Line Interface 210
- command set 15
- commit request 14
- Common Internet File System 10
- common tasks 92
- Concurrency 13
- concurrent access 33
- concurrent read 33
- concurrent users 33
- config node 86, 175, 224
- config role 225
- config saves 264
- configuration 224
- configuration change 243
- configuration data 223
- configuration tasks 134
- configure alerting 178
- configure block storage 192
- configure the controllers 191
- Configure USB key 149
- configuring 134
- connecting remotely 89
- connection oriented 15
- connectivity 24
- Consistency Groups 95, 194
- consistency groups 76
- consistent 76–77
- consistent snapshot 76
- consistent state 80
- Control Enclosure 140
- Control enclosure 136, 146
- control enclosure 145
- Control Enclosures 139
- control port 19
- Controller Enclosure 86
- controller enclosures 138
- cooling 138
- copy
  - operation 78
  - rate 78
- copy operations 95
- copy process 76
- Copy Services 95
- Count Key Data 8



- Create a CIFS share 202
- Create a file set 200
- Create a file system 199
- Create an HTTP, FTP, SCP export 204
- Create an NFS export 203
- Created Time stamp 30
- credential 38
- credential verification 38–39
- credentials 37
- critical 180
- critical failures 242
- cross-node 27
- cross-platform mapping 29
- cross-protocol 27
- CRU 242, 263
- CTDB 29, 33
- current configuration data 224
- current firmware 147
- Custom 200
- cycling period 79

## D

- daemons 19, 71
- DAP 41
- data
  - production 77
- data access 59
- data blocks 73
- data collect file 257
- Data Collection 263
- data collection processes 256
- data collects 173
- data integrity 32, 225
- Data Management Application 107, 229
- data mining 77
- data organization 8
- data placement 60
- data port 19
- data protection 59–60
- data transfer 14, 33
- Database Management System 8
- DBMS 8
- DC 41
- debug 243
- decrypts 39
- default ACL 47
- Default Gateway 170
- default passwords 186
- default rule 73
- Delete 214
- delete
  - FlashCopy mappings 78
- delete after complete 194
- deletion 61
- deny access 209
- deny-mode 32
- dependent file set 198
- dependent File Sets 71
- dependent file sets 71
- destage 226

- destaged 227
- DFS 30
- dialect 15, 17
- dialects 15
- digital certificates 38
- directory 38
- Directory Access Protocol 41
- directory service 40, 43
- Directory Services 96
- disable the animation 89
- disaster recovery 80
- disaster tolerance 80
- disk 78
- Disk Magic 101, 103
- disk partitions 149
- distance 78
- distributed 62
- distributed computing 13, 15
- Distributed File System 30
- distributed file system 22
- distribution boards 139
- DMA 107, 229
- DMA server 239
- DNS 32, 41, 119
- DNS domain 96
- DNS server 160
- DNS servers 121
- Domain Controller 41
- domain controller 38
- domain name 160
- Domain Name System 32, 41, 102
- DOS attributes 29
- Download Logs 173
- Download Support Package 257
- download support package 173
- drain 81
- Drive assemblies 136
- drive bays 23
- drive class 94
- drive failures 25
- drive types 27
- dual fabric 193
- dumps 173
- duplicate transfers 31
- DVD restore 148

## E

- EA 210
- Easy Setup wizard 156
- Easy Tier 25
- e-mail 96
- emulate 31
- enclosure link 225
- enclosure management address 96
- enclosure nodes 150
- enclosures 94
- encrypt 19
- encrypted 39
- encrypted transfer 20
- encryption algorithm 15

- enforcement of ACLs 31
- error alerting 261
- error code 245
- errors 77
- Ethernet 140
- event 242
- event count 245
- event details 245
- event ID 245
- event log 227, 242–243
- event logs 92–93, 185, 220
- Event Notifications 96
- Events 93
- events 178
- exclusion list 209
- Expansion enclosure 136
- Expansion Enclosures 139
- expansion enclosures 23, 138
- export 47, 197
- exports 13–14, 197
- EXT2 149
- EXT3 149
- ext3 9
- extended attributes 209–210
- Extended NIS 161, 164
- extent maps 225
- extent sizes 25
- extents 94
- external access 159
- external directory service 38
- external disks 94
- external scan engines 108
- External storage 192
- external storage 76, 94
- EZ-Setup 155

## F

- failover 106
- Failure 13
- failure group 72
- failure groups 72
- FAT32 149
- FB 8
- FBA 8
- FC 9, 24–25
- FCP 25
- Fiber optic cables 144
- Fibre Channel 9, 96
- Fibre Channel Protocol 25
- File 139
- file 11, 145
- file access 38, 57
- file access concurrency 29
- file access IP addresses 96
- file access protocols 10, 13, 22
- file access request 38
- file clients 22
- file export 209
- File Module 145
- File module backup 225

- file module resources 222
- file module software level 147
- File Modules 22–24, 140
- file modules 25, 76, 87, 139
- file offset 14
- file open 209
- file path list 80
- file pattern 238
- file recall 108, 210
- file replication 95
- file server 22
- File Server Appliance 11
- file server subsystem 22–24
- File Service 93
- File service components 197
- File Services configuration 197
- file serving 9, 11, 26
- File sets 197
- file sets 93, 201
- file shares 174
- File sharing 29
- file sharing 9, 11, 26
- file sharing functions 29
- file sharing protocol 18
- file sharing protocols 10, 13, 22
- file signature 208
- file space 17
- file storage 11
- File System 8
- file system 9, 11, 66, 68
- file system pools 68
- file system-level snapshots 27
- File systems 197
- file systems 9, 22
- File Transfer Protocol 11, 18
- file transfer protocol 20
- file transfer services 26
- file usage by each user 93
- file versions 59
- file volumes 11
- filer 59
- files 11
- filler panel 138
- firewall 19
- firewalls 14, 19
- Fix 265
- Fixed Block Architecture 8, 11
- fixed block storage 8
- fixed field 246
- FlashCopy 76–77, 95, 194
  - create 76
  - creating 76
  - mappings 77–78
  - operations 77
  - source 76
  - target 76
- FlashCopy Mappings 95
- FM 22
- Frequency 216
- frequency 59

FS 8, 14  
 FTP 11, 13, 18, 22  
 FTP client 11  
 FTP file client 11  
 FTP file server 11  
 FTP over SSH 20  
 FTP server 11  
 FTP support 31

## G

General 96  
 General Parallel File System 22, 65  
 Generic Security Services Application Program Interface 17  
 GID 40  
 Global Mirror 76–78, 95  
 global name space 27, 33  
 global namespace 197  
 Global time-out 212  
 GNU General Public License 29  
 GPFS 22, 27–29, 31–33, 40, 42, 59–61, 65–66, 103, 197, 200, 229  
 GPFS ACLs 42  
 GPFS cluster 68  
 GPFS file sets 71  
 GPFS NFSv4 42  
 GPFS one-tier architecture 63  
 GPFS quota management 74  
 GPFS Snapshots 73  
 GPFS technical concepts 66  
 GPL 29  
 grace period 74  
 grain size 25  
 granular space management 59  
 Graphical User Interface 210  
 grid parallel architecture 66  
 group identifiers 40  
 GSSAPI 17–18  
 GUI 210  
 Guided Maintenance 248  
 Guided Maintenance Procedures 262  
 Guided Maintenance routines 248

## H

Hard Disk Drives 8  
 hard quota 59, 74  
 hard quotas 74  
 hardened data 225  
 hardware 166  
 hardware placement 118  
 hash 80  
 hashes 38  
 HDD 8, 26  
 HDDs 23  
 health 185  
 health reporting 261  
 health state 32  
 health status indicator 220  
 heterogeneous file sharing 39

high latency 27  
 high-available 32  
 HIM 24  
 historic usage 93  
 home directory 59  
 Host Interface Module 24  
 Host Mappings 95  
 host user data 223  
 Hosts 94  
 hosts 94  
 HSM 61  
 HTTP 11, 13, 19  
 HTTP aliases 31  
 HTTP client 11  
 HTTP over SSL 19  
 HTTP Secure 19  
 HTTP server 11  
 HTTP/1.1 19  
 HTTPS 19, 22  
 https connection 88  
 HTTPS support 31  
 Hypertext Transfer Protocol 11, 19

## I

I/O 8  
 I/O group 78  
 I/O patterns 101  
 IBM SAN File System 66  
 IBM Support Remote Access 262  
 ID mapping 43  
 ID mapping parameters 46  
 identity 38  
 IETF 14, 19  
 IETF standard 18  
 ILM 27, 73, 107  
 image mode 25  
 implementation 134  
 implementation checklist 134  
 in-band 33  
 inclusion list 209  
 inconsistent 77  
 increases security 209  
 incremental 194  
 incremental changes 80  
 Incremental FlashCopy 77  
 independent 198  
 independent File Sets 71  
 independent file sets 71, 73  
 individual files 209  
 infected files 214  
 infection 208  
 in-flight workloads 32  
 Information Archive 66  
 Information Lifecycle Management 27, 73, 107  
 inheritance model 209  
 init tool 150  
 initial file set 198  
 Initialise the file modules 153  
 initialization process 153  
 initialize the system 149

- initiator 25
- inode 198
- input/output 8
- install DVD 265
- install utility 149
- installation 134
- intercluster 78
- interface node 26
- interface node function 32
- Interface Nodes 61
- interface nodes 63, 80
- Interface Pool 170
- internal disk drives 94
- Internal Storage 94
- internal storage clients 22
- Internet Engineering Task Force 14
- Internet SCSI 9
- intracluster 78
- IP addresses 136
- IPv4 113
- iSCSI 9, 24–25, 96, 176
- ISO image 148

## J

- junction 198
- junction point 198

## K

- Kerberos 14–15, 17–18, 39, 41, 43
- kerberos 163
- Kerberos 5 18
- Kerberos API 18
- Kerberos name 163
- Kerberos realm 163
- Kerberos server 39
- key 39, 97
- key file 91
- keys 38, 96

## L

- layers 79
- layers from 11
- LBA 8
- LBA storage 8
- LDAP 31, 38, 41, 111, 162–163
- LDAP authentication 44
- LDAP database 41
- LDAP entry 41
- LDAP server 163
- leaf directory 47
- leases 30
- level of notification 246
- license 157
- License Agreement 156
- licensing details 96
- lights on 173
- lights out 173
- Lights-on 173

- Lights-out 173
- Lightweight Directory Access Protocol 41
- Linux 40
- LIPKEY 15
- listening port 19
- load balancing technology 219
- locking 16, 29, 33
- locking change 32
- locking characteristics 31
- locking mechanisms 15, 59, 71
- locking requests 32
- locking services 27
- locks 32
- log listing 173
- logging 96
- logging server 181
- Logical Block Addressing 8
- logical interface 169
- logical unit 8
- Logical Unit Number 8
- logical volume 8
- Logical Volume Manager 8
- logical volumes 22, 24
- logs 173
- logs directory 173
- low bandwidth 27
- Low Infrastructure Public Key Mechanism 15
- low reliability 27
- LU 8
- LUN 8
- LV 8
- LVM 8

## M

- mailbox 261
- maintenance procedures 241–242
- maintenance routines 248
- manage shares 201
- management 257
- management address 96
- management IP address 87
- management node 26
- mandatory scan 209
- manual authorisation 173
- manual backup 225
- manually triggered 224
- manually tuned 219
- mapped 76
- mapped volumes 227
- mapping 11, 76, 95, 194
- mapping details 224
- mapping logical volumes 25
- mappings 77
- Massachusetts Institute of Technology 18, 39
- MBCS 34
- McAfee 108
- MDisk Groups 25
- MDisks 94, 197
- mesh 146
- mesh pattern 145

- metadata 33, 72–73, 107
- metrics 222
- Metro Mirror 76, 78, 95
- Microsoft New Technology File System 9
- migration 27, 61
- Migration policies 61
- Migration-ILM 200
- mirroring 25
- MIT 18
- MIT Kerberos 39
- Mobility 13
- modelling tools 103
- Modification 265
- Modified Time stamp 30
- module 139
- mount 78
- mounts 13
- multipath driver 227
- multipathing 267
- multiple scan nodes 210
- Multiple Target FlashCopy 77

## N

- Name resolution 29
- namespace 59, 61
- NAS 2, 11
- NDMP 94, 232, 239
- NDMP agent 229
- NDMP topologies 107
- Nearline 27
- nested groups 42
- nested mounts 28
- NetBIOS 15, 17
- NetBios name 164
- Netgroup 45, 161
- netgroup 43
- netgroups 42
- Network 96, 174
- network adapters 174
- Network Attached Storage 2, 11
- network authentication protocol 39
- Network Data Management Protocol 94
- network definition 169
- network devices 146
- network file sharing 27
- Network File System 10, 13
- Network Information Service 41, 43
- network monitoring tools 222
- Network Protocol 96
- network retry 28
- Network Shared Disk 66
- Network Shared Disks 61
- Network Storage Devices 199
- Network Time Protocol 101, 119
- network traffic 101
- new definition 215
- New File System 199
- NFS 10, 13, 22, 38, 117
- NFS access 203
- NFS daemon 34

- NFS export 38
- NFS exports 45
- NFS file client 10
- NFS file server 10, 38
- NFS file service 10
- NFS protocol support 28
- NFS requests 14
- NFS Version 2 14
- NFS Version 3 14
- NFS Version 4 14
- NFS version 4.1 15
- nfsd 28
- NFSv2 14
- NFSv3 14
- NFSv3 advisory locking mechanism 28
- NFSv4 14, 28, 42
- NFSv4 ACL 28, 42
- NFSv4 ACLs 28
- NFSv4.1 15
- NIS 41, 43, 45
- NIS authentication 164
- NIS domain 164
- NIS ID mappings 46
- No action 214
- node canister 145
- node canisters 225
- node fail 224
- node failover 28
- node settings 214
- nodes 41, 61
- NSD 66
- NSD clients 66
- NSD grouping 72
- NSD servers 66
- NSDs 61, 72, 199
- NT LAN Manager 17
- NT4 17
- NTFS 9
- NTP 101, 119, 121, 184

## O

- offload data 73
- operational storage nodes 224
- oplock 16, 32
- oplocks 30, 71, 110
- opportunistic locking 16, 32, 71, 110
- Opportunistic locks 30
- opportunistic locks 30
- OSI 18

## P

- P2P 10
- packets 15
- packing slip 136
- parallel 209
- parallel access 71
- parallel computing 66
- parallel processing 15
- parent directory 46

- partition 149
- Partnership 193
- Partnerships 95
- parts 262
- passive mode 18–19
- password 39, 188
- passwords 38
- path 209, 214
- PDC 43, 161
- PDC server 164
- Peer-to-Peer 10
- performance 14, 32, 93, 103, 209
- performance profiles 27
- performance requirements 27
- periodic backup 171
- periodic configuration backup 171
- permissions 14, 42
- physical volume 8
- Placement policies 61
- planning 134
- planning worksheet 138
- plug locations 140
- PMR 261
- point-in-time 95
- point-in-time copy 27, 76
- policy based 60
- policy engine 73
- policy engine scan 80
- policy-based automated placement 27
- pool of scan nodes 210
- Pools 197
- port mapper 14
- Portable Operating System Interface 15
- portmap 14
- ports 14
- Ports by Host 95
- POSIX 15, 28
- POSIX attributes 29
- POSIX bits 42
- POSIX byte range locks 31
- POSIX locks 31
- POSIX-compliant 29
- power 138
- power cords 136, 139
- power outlets 118
- power strips 139
- Prefetch 237
- Prefetching 237
- presets 77
- primary 78
- Primary Domain Controller 43
- proactive scan operation 209
- Problem Management Record 261
- profiles 186
- protocol 25
- proxy authentication 159
- public IP addresses 32
- public network domain 160
- Public Networks 96
- PuTTY 90

- PV 8

## Q

- Quarantine 214
- quorum 68, 224
- quorum data 225
- quorum disk 68
- quorum disks 227
- quota 59, 93
- quotas 17, 198

## R

- rack 136
- rack space 138
- racking of the modules 138
- Rack-mounting hardware kit 136
- RAID 8, 25
- random access 8
- random access mass storage 8
- random port 19
- read-ahead 33
- ReadOnly bit 29
- record updates 32
- recovery 13, 263
- recovery of a node 226
- recovery procedures 242
- recovery routines 246
- recovery situations 97
- Redbooks website
  - Contact us xvi
- RedHat 6.1 22
- redundancy 72, 145
- Redundant Array of Independent Disks 8
- redundant communication 113
- redundant power supplies 23
- registry 38
- relationship 76, 78, 95
  - removing 76
- Release 265
- reload 148
- Remote Copy 95
- Remote copy 193
- remote copy 157
- remote copy partner 79
- Remote Copy Services 104
- Remote Procedure Call 13
- remote support functionality 171
- remote support functions 172
- Remote Technical Support 242
- Replicate File Systems 95
- replication 27
- replication layer 79
- Representational State Transfer 31
- Request For Comment 13
- reset packet 32
- REST 31
- retention 59
- retention period 107
- Reverse FlashCopy 77

RFC 13–14  
 RFC 1813 14  
 RFC 3010 14  
 RFC 5661. NFSv4 15  
 roles 71  
 root 198  
 root directory 47  
 Root password 264  
 root squash 203  
 round-robin access 32  
 RPC 13  
 rpc.portmap 14  
 rpcbind 14  
 rsa key 91  
 rsync 27, 80  
 rsync transfer 80  
 RTS 242  
 rule 73

## S

SA 86  
 SAMBA 41, 43  
 Samba 29, 209  
 Samba PDC 44, 163  
 SAMBA PDC authentication 44  
 Samba4 41  
 SAN fabrics 146  
 SAS 24–25  
 SAS cables 118, 141  
 SAS chains 118  
 SAS ports 141  
 save files 22  
 SBOD 23  
 Scalability 13  
 scalability 209  
 Scale Out Network Attached Storage 22  
 scale-out NAS 61  
 scan engine 72  
 scan nodes 208–210  
 Scan protocol 212  
 scan request 208  
 scanning process 210  
 scope 59  
 SCP 11, 13, 19, 22  
 SCP and SFTP support 31  
 scripting tool 97  
 scripts 32  
 SCSI 8  
 search tool bar 259  
 secondary 77–78  
 secret-key strong cryptography 18  
 sector size 8  
 Secure Copy Protocol 11, 19  
 Secure FTP 11, 20  
 Secure Shell 11  
 Secure Sockets Layer 19, 43  
 security 15  
 security identifier 40  
 security key 91  
 self healing 242, 248

self-healing 71  
 semantics 29  
 sequential 25  
 Server Message Block 10, 15  
 service access 264  
 Service announcement 29  
 Service Assistant 86  
 service assistant IP address 150  
 service call 262  
 Service for Unix 162  
 Service IP Addresses 96  
 Service IP addresses 150  
 Service ip addresses 264  
 service layers 8  
 service ports 175  
 services 93  
 Services for Unix 41, 43–44  
 session 15  
 session based 38  
 session credential 39  
 SFTP 11, 13, 20, 22  
 SFU 41, 43–44, 162  
 share 18, 59, 93, 197  
 share concept 31  
 share files 22, 59  
 share level 16  
 Shares 93, 201  
 shares 10, 15, 60, 197  
 Shares/Exports 198  
 SID 40  
 signature 208–209  
 Simple FTP 20  
 Simple Public Key Mechanism 15  
 simplified graphics 89  
 Single Pool 199  
 Small Computer System Interface 8  
 Smart Analytics 66  
 SMB 10, 13, 15, 22  
 SMB 2 18  
 SMB client 39  
 SMB file client 10  
 SMB file server 10, 39  
 SMB file service 10  
 SMB file sharing 32  
 SMB file sharing function 28  
 SMB protocol 29, 39  
 SMB protocol support 29  
 SMB share 39  
 SMB Version 2 18  
 SMP protocol semantics 32  
 SMTP mail server 178  
 Snapshot 194–195  
 snapshot 77, 93  
 Snapshots 59, 107  
 snapshots 27  
 SNIA 56  
 SNMP 96  
 SNMP alert 242  
 SNMP protocols 261  
 SNMP server 178, 180



- soft quota 59, 74
- Soft quotas 74
- software bundle 266
- software components 26
- software levels 147
- Software package 265
- software stack 57, 63
- software upgrades 96
- software upgrading 241
- Solid State Drives 8
- SONAS 22, 27, 29, 61, 224
- SONAS asynchronous replication 80
- SONAS cluster manager 29, 32–33
- SONAS file modules 224
- SONAS HSM 30
- SONAS Snapshots 30
- SONAS Software 79
- SONAS software 26
- source 76
- sources 77
- space efficient 73, 80
- Space Efficient FlashCopy 77
- space-efficient 79
- SPKM-3 15
- split brain 68
- SQL 73
- SSD 8, 26
- SSDs 23
- SSH 11, 90
- SSH daemon 20
- SSH File Transfer Protocol 20
- SSH2 20
- sshd 31
- SSL 19, 43
- SSL mode 163
- SSL secured communication 30
- SSL security 163
- stateful protocol 14
- stateless operation 14
- stateless protocol 15
- states 78
  - Synchronized 78
- Static data 61
- statistical counters 221
- statistical data 221
- statistics interval 221
- status 224
- status indicator 185
- status information 224
- storage agent 229
- storage client 8
- storage clients 24
- Storage controller configuration 191
- storage enclosure password 186
- storage extent mapping tables 224
- storage functions 25
- storage layer 79
- Storage Networking Industry Association 56
- storage node 26
- Storage Nodes 61

- storage nodes 63, 86
- storage pools 25, 94
- storage tier 61
- string 15
- stripe size 68
- striped 25
- stripes 68
- structure 9
- stub file 61, 73
- sub block 68
- subdirectory tree 209
- suggested tasks 92
- Sun 13
- Sun Microsystems 13
- Superuser password 263
- Support 96
- support files 173
- support notifications 158
- support package 173
- Switched Bunch Of Drives 23
- Symantec 108
- symbolic links 30
- symmetric key cryptography 39
- symmetrical 118
- synchronized 77
- Synchronized state 78
- synchronous 14
- Syslog Server 181
- System 92
- system attributes 156
- System bit 29
- System details 92
- system licences 171
- System License 157
- System Migration 94

## T

- T1 226
- T2 226
- T3 226
- T4 226
- tape restore 80
- target 25, 76–77
- target only 25
- targets 77
- TCP 13–14
- TCP/IP 15
- TDA 103
- TDB 171
- Technical Delivery Assessment 103
- terminology 8
- test e-mail 178
- test utility 265, 269, 271
- Third Extended File System 9
- threshold 27
- ticket 39
- tickle-ACK 32
- tie breaker 68
- tie-break 226
- tie-breaking 224



- tier 66
- Tier 1 226
- tier 1 recovery 226
- Tier 2 226
- tier 2 227
- Tier 3 226–227
- tier 3 recovery 225
- Tier 4 226–227
- tiering 27
- tiers 25, 107, 226
- time 77
- time and date 96
- time consistency 95
- time stamps 30
- timestamps 245
- Time-Zero 76
- Tivoli Storage Manager 94
- TLS 19, 43
- trace 243
- transfer size 14
- transferring files 11
- Transmission Control Protocol 13
- transparent file recall 61
- transport layer 13–14
- transport layer protocols 15
- Transport Layer Security 19
- Transport Level Security 43
- transport protocol 15
- traversal permission 47
- tree 13, 41
- tree view 92
- trees 18
- Trivial DataBase 171
- trusted 38
- trusted networks 14
- TSM 30
- tuning 237
- two-tier architecture 63, 66

## U

- UDP 13, 15
- UID 16, 38–39
- uncommitted data 14
- unfixed events 185, 242, 246, 248
- unique user IDs 38
- universal tool 150
- UNIX 13, 162, 198
- Unix ACLs 29
- Unix authentication 38–39
- Unix NFS client 38
- UNIX permissions 15
- unpack 136
- unstable writes 14
- upgrade 147
- upper layer protocols 11
- upper layers 8
- usage metrics 93
- USB key 149
- USB mass storage key 149
- use cases 57

- user access 95
- user authentication 18, 31
- User Datagram Protocol 13
- user groups 95
- User ID 16
- user identifiers 39
- user names 39
- User Security 186
- Users 95
- utility 265

## V

- V7000 23
- V7000 controller enclosure 23
- V7000 expansion enclosures 23
- V7000 storage pools 69
- VDisk 76–78
- Version 265
- virtual volumes 24
- virtualisation license 157
- virtualized NSDs 68
- virus signature 209
- virus signature definition 209
- VLAN ID 170
- volume 8
- volume mappings 68
- Volume Shadow Service 80
- Volume Shadow Services 30
- Volumes 95, 197
- volumes 94
  - source 78
  - target 77
  - thin-provisioned 77
- Volumes by Host 94
- Volumes by Pool 94
- Volumes The 94
- VRMF 265
- vsftpd 31
- VSS 30, 80

## W

- warm start 226
- warning levels 59
- web server 86
- Web-based Distributed Authoring and Versioning 31
- WebDAV 31
- well known port 14
- well known ports 19
- win32 share modes 29
- Windows access control semantics 29
- Windows Active Directory 39
- Windows authentication 38–39
- Windows domain controller 39
- Windows registry 39
- wizard 94, 156
- workload parameters 103
- workload-sharing 32
- write caching options 110
- write operations 14

writes 78

## **Z**

zoned 191

zoning 193

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide:>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine:fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.



# Storwize V7000 Unified

(1.5" spine)  
1.5"<-> 1.998"  
789 <-> 1051 pages



# Storwize V7000 Unified

(1.0" spine)  
0.875"<-> 1.498"  
460 <-> 788 pages



# Storwize V7000 Unified

(0.5" spine)  
0.475"<-> 0.873"  
250 <-> 459 pages



# Storwize V7000 Unified

(0.2" spine)  
0.17"<-> 0.473"  
90<->249 pages

(0.1" spine)  
0.1"<-> 0.169"  
53<->89 pages

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide:>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine:fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Draft Document for Review March 29, 2013 4:35 pm

8010spine.fm 320



# Storwize V7000 Unified

(2.5" spine)  
2.5" <-> nnn.n"  
1315<-> nnnn pages



# Storwize V7000 Unified

(2.0" spine)  
2.0" <-> 2.498"  
1052 <-> 1314 pages





# Implementing the IBM Storwize V7000 Unified



**Consolidates storage and file serving workloads into an integrated system**

**Simplifies management and reduces cost**

**Integrated support for Real-time Compression**

IBM® Storwize® V7000 Unified is a virtualized storage system designed to consolidate block and file workloads into a single storage system for simplicity of management, reduced cost, highly scalable capacity, performance and high availability. IBM Storwize V7000 Unified Storage also offers improved efficiency and flexibility through built-in solid state drive (SSD) optimization, thin provisioning, Real-time Compression, and non-disruptive migration of data from existing storage. The system can virtualize and reuse existing disk systems offering a greater potential return on investment.

## **INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION**

### **BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE**

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

**For more information:**  
**[ibm.com/redbooks](http://ibm.com/redbooks)**