

IBM b-type Gen 7 Installation, Migration, and Best Practices Guide

Brian Larsen

Kameswara Bhaskarabhatla

Matt Brennan

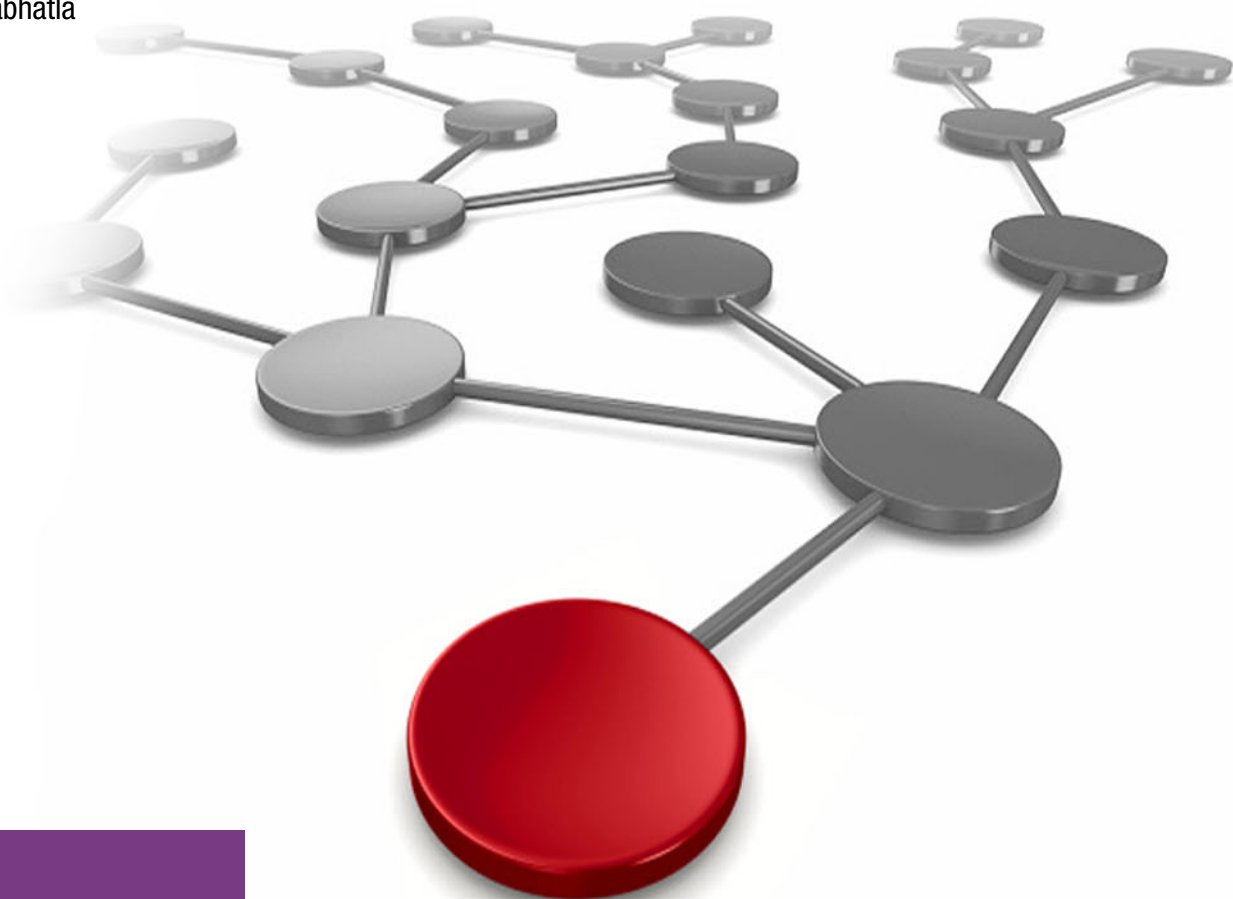
AJ Casamento

Jack Consoli

Mark Detrick

Johnny Fonseca

Tim Jeka



Storage



IBM Redbooks

IBM b-type Gen 7 Installation, Migration, and Best Practices Guide

March 2022

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

Second Edition (March 2022)

This edition applies IBM b-type Gen 7 SAN products.

© Copyright International Business Machines Corporation 2021, 2022. 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

Notices	ix
Trademarks	x
Preface	xi
Authors	xi
Now you can become a published author, too!	xiii
Comments welcome	xiii
Stay connected to IBM Redbooks	xiii
Summary of changes	xv
March 2022, Second Edition	xv
Chapter 1. Introduction to IBM b-type Gen 7	1
1.1 IBM b-type Gen 7 platforms overview	2
1.1.1 Self-learning	2
1.1.2 Self-optimizing	3
1.1.3 Self-healing	3
1.2 IBM b-type Gen 7 platforms and naming conventions	4
Chapter 2. IBM b-type Gen 7 product review	5
2.1 IBM b-type Gen 7 SAN director overview	6
2.2 IBM b-type Gen 7 SAN512B-7 (8961-F78) director overview	7
2.2.1 Hardware components	9
2.2.2 Port-side view of the device	10
2.2.3 Port-side slot numbering	11
2.2.4 Non-port side view of the device	12
2.2.5 More information	12
2.3 IBM b-type Gen 7 SAN256B-7 (8961-F74) director overview	13
2.3.1 Hardware components	14
2.3.2 Port-side view of the device	16
2.3.3 Port-side slot numbering	17
2.3.4 Non-port-side view of the device	17
2.3.5 More information	18
2.4 IBM b-type Gen 7 SAN64B-7 (8960-P64/R64) switch overview	18
2.4.1 Software license options	20
2.4.2 Port-side view of the device	20
2.4.3 Nonport-side view of the device	21
2.4.4 More information	21
2.5 Upgrading Gen 6 to Gen 7 supported platform overview	22
2.5.1 Migration steps	22
2.6 IBM b-type Gen7 transceiver module overview	23
Chapter 3. Gen 7 Fabric Operating System overview	25
3.1 Autonomous SAN	26
3.1.1 Self-learning	26
3.1.2 Self-optimizing	27
3.1.3 Self-healing	28
3.2 Automation	28
3.2.1 Why organizations should embrace automation	29

3.3 FOS 9.x features	30
3.3.1 Port identifiers	30
3.3.2 Unified addressing mode	30
3.3.3 Fabric zone locking	32
3.3.4 FICON logical switch	34
3.4 Management tools	34
3.4.1 WebTools	34
3.5 Security	36
3.6 Traffic Optimizer	36
3.6.1 Speed-based classification	38
3.6.2 Slow drain device quarantine isolation	38
3.6.3 Compatibility with QoS zones and CS_CTL	38
3.6.4 Compatibility with previous platforms	39
3.6.5 Traffic Optimizer and QoS Mode	39
3.7 Fabric Notifications	39
3.8 IBM b-type switch Fabric Notifications solution	41
3.8.1 IBM b-type Fabric Performance Impact Notification	41
3.8.2 Congestion Signal primitives	42
3.8.3 Extended link service notifications	42
3.8.4 Exchange Diagnostic Capabilities and Register Diagnostic Functions	42
3.8.5 Congestion Signal	43
3.8.6 Congestion notification	44
3.8.7 Congestion peer notification	44
3.8.8 Link Integrity	47
3.8.9 SCSI Command Delivery	50
Chapter 4. Virtual Fabrics overview	53
4.1 Logical switch	54
4.1.1 Default logical switch	54
4.1.2 Logical switches and fabric IDs	55
4.1.3 Port assignment in logical switches	56
4.1.4 Logical switches and connected devices	58
4.1.5 Management model for logical switches	59
4.2 Logical fabric	60
4.2.1 Logical fabric and ISLs	60
4.2.2 Base switch and extended ISLs	61
4.2.3 Account management and Virtual Fabrics	64
4.2.4 Setting up IP addresses for a logical switch	64
4.2.5 Logical switch log in context	65
4.2.6 Supported platforms for Virtual Fabrics	67
4.3 Device sharing	70
Chapter 5. Fibre Channel SAN design, implementation, and migration	73
5.1 Modernizing SAN infrastructures	74
5.1.1 IBM b-type Gen 7 platforms: The intelligent foundation of modern data centers	74
5.1.2 Modernize with NVMe to achieve lower latency and higher bandwidth	74
5.2 Fibre Channel SAN design	75
5.2.1 SAN design basics	75
5.2.2 Capabilities to consider for your storage architecture	77
5.2.3 Designing a SAN	79
5.2.4 High-performance workloads	81
5.2.5 Redundancy and resiliency	82
5.2.6 Switch interconnections	82

5.2.7	UltraScale ICL connectivity for IBM b-type Gen 7 platforms	84
5.2.8	Best practices for IBM b-type Gen 7 UltraScale ICL connections	84
5.2.9	Mesh topology	85
5.3	Implementation planning for a Fibre Channel SAN	86
5.3.1	Device placement and traffic locality	86
5.3.2	Scalability and performance	90
5.3.3	SAN design for critical workloads	92
5.3.4	Capacity planning	93
5.4	Migration strategies of Fibre Channel SAN infrastructure	95
5.4.1	IBM b-type Gen 7 provide infrastructure for today and tomorrow	95
5.4.2	Tools to help migration planning	97
5.4.3	SAN migration factors	100
5.4.4	SAN migration process overview	101
5.5	References	103
Chapter 6.	FICON design, implementation, and migration	105
6.1	Important support notes for Gen 7 platforms	106
6.1.1	Product support dates	106
6.1.2	Software	106
6.1.3	Link address refresher	108
6.2	FICON SAN design	109
6.2.1	Typical redundant pathing	111
6.2.2	Cascaded performance and virtual channels	112
6.2.3	Ultra-low latency	113
6.2.4	Channel paths: 32 versus 48 port cards	113
6.2.5	Chassis configuration options using 48 port cards	114
6.2.6	Two logical FICON switches per chassis	115
6.2.7	ISL B/W sharing with logical switches	116
6.3	FICON SAN implementation	117
6.3.1	Chassis-level management	117
6.3.2	FICON display mode setup	117
6.3.3	Configuring FICON fabric	119
6.3.4	Managing FICON fabrics	129
6.3.5	Converting a production switch to a FICON logical switch	130
6.4	FICON SAN migration	134
6.4.1	Firmware	134
6.4.2	FICON logical switch	134
6.4.3	Inter-Chassis Links	135
6.4.4	Cabling	135
6.4.5	Patch room and fiber plant	136
6.4.6	SFPs	136
Chapter 7.	IBM b-type Gen 7 extension design, implementation, and migration	141
7.1	Extension platform overview	142
7.1.1	IBM b-type Gen 7 extension platforms	142
7.2	Extension design and architectures	145
7.2.1	FCIP design and architectures	145
7.2.2	IP Extension design and architectures	148
7.3	Extension implementation	157
7.3.1	WAN side	157
7.3.2	FC/FICON side	190
7.3.3	LAN side	192
7.4	Extension migration strategy	199

7.4.1 IBM Extension platforms	199
7.4.2 IBM Extension solution designs	200
7.4.3 Migration methodology and techniques	200
7.4.4 Migration summary	202
7.5 Extension validation	203
7.6 Extension Monitoring	209
7.6.1 IBM SANnav SAN Management Tool	209
7.6.2 Fabric operating system CLI	215
7.6.3 REST API	215
7.6.4 SNMP	215
Chapter 8. NVMe over Fibre Channel SAN design, implementation, and migration strategies	217
8.1 NVMe overview	218
8.1.1 Comparing mapping of NVMe versus FCP protocols	220
8.1.2 NVMe over Fibre Channel	223
8.2 SAN design considerations for NVMe over Fibre Channel	225
8.3 Implementation overview of NVMe over Fibre Channel	226
8.4 Migration example of NVMe over Fibre Channel	227
8.5 Summary	231
Chapter 9. Brocade Fabric OS administrative tools and diagnostics	233
9.1 Command-line interface overview	234
9.1.1 Connecting to Brocade Fabric OS through the serial port	234
9.1.2 Connecting to Brocade Fabric OS through SSH or telnet connection	235
9.2 WebTools overview	235
9.3 Fabric Vision overview	237
9.3.1 Monitoring, alerting, and performance suite	237
9.3.2 Flow Vision	239
9.3.3 ClearLink Diagnostics	242
9.3.4 Fabric Vision SANnav managed features	247
9.3.5 I/O Insight	248
9.3.6 VM Insight	248
9.3.7 Fabric Vision diagnostics	249
Chapter 10. IBM b-type Gen 7 SAN automation	251
10.1 Motivation to automate	252
10.2 REST API overview	253
10.3 Simple automation example	255
10.4 Ansible as an alternative	257
10.5 SANnav REST API	258
10.6 Conclusion	259
Chapter 11. IBM SANnav Management Suite overview	261
11.1 IBM SANnav Management Suite 2.1 release overview	262
11.1.1 IBM SANnav Management Portal	262
11.1.2 IBM SANnav Global View	263
11.1.3 New hardware platforms supported in SANnav 2.1	263
11.2 IBM SANnav Management Portal 2.1 release overview	264
11.2.1 Browser requirements	264
11.2.2 New SANnav Management Portal server platform support and infrastructure	265
11.2.3 Server requirements	265
11.2.4 SANnav Management Portal scalability 2.1.0	266
11.2.5 Pre-installation requirements	266

11.2.6	Downloading the SANnav Management Portal software	266
11.2.7	SANnav installation and migration QuickStart checklists	277
11.2.8	Installation overview	278
11.2.9	Deploying SANnav Management Portal	279
11.2.10	System and server requirements for SANnav Management Portal	281
11.2.11	Installation prerequisites for SANnav Management Portal	282
11.2.12	Installing SANnav Management Portal	283
11.2.13	Firewall requirements for SANnav Management Portal	284
11.2.14	SANnav Management Console commands	286
11.2.15	Checking the server health	286
11.2.16	Starting SANnav Management Portal	288
11.2.17	Changing your password	289
11.2.18	User interface overview	290
11.2.19	Initial setup and configuration	293
11.3	IBM SANnav Global View 2.1 release overview	293
11.3.1	SANnav Global View browser requirements	294
11.3.2	SANnav Global View system and server requirements	294
11.3.3	SANnav Global View installation prerequisites	295
11.3.4	Installing SANnav Global View	296
11.3.5	Logging in to SANnav Global View	297
11.3.6	SANnav Global View overview	298
11.3.7	Generating a license for SANnav Global View	300
11.3.8	Global dashboard overview	303
11.3.9	Monitoring fabric health	304
11.3.10	Monitoring switch health across Management Portal instances	305
11.4	IBM SANnav WebTools	308
11.4.1	WebTools user interface overview	310
11.5	Reference guides	310
Chapter 12	Storage area network security	311
12.1	Fibre Channel SAN security layers	312
12.2	Management interfaces	312
12.3	Displaying the IP filter	313
12.4	Role-based access controls	314
12.5	Default accounts	315
12.6	User accounts	315
12.7	Security protocols	315
12.8	Access Control Lists	316
12.9	Securing end points in SAN infrastructure	317
12.10	ISL in-flight encryption and compression	318
12.11	Policy Database Distribution	322
Appendix A	Fabric details	325
A.1	Current fabrics	326
A.2	Individual fabrics	326
A.3	Device	327
A.4	Metrics and effect on SAN design and performance	327
A.5	Consolidated SAN requirements	328
A.6	Application-specific information	329
Appendix B	Brocade Fabric OS CLI monitoring commands	331
B.1	portshow (tunnel and circuits)	332
B.2	portshow (LAN Stats)	332
B.3	portshow (TCP Stats)	333

B.4	portStatsShow	338
B.5	gePortPerfShow	340
B.6	switchShow	341
B.7	portPerfShow	342
B.8	portErrShow	342

Notices

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

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®
FICON®
IBM®
IBM FlashSystem®

IBM Spectrum®
IBM Z®
Insight®
Redbooks®

Redbooks (logo) ®
z/OS®
z15™

The following terms are trademarks of other companies:

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

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

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

Ansible, Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

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

Preface

The IBM® b-type Gen 7 Fibre Channel directors and switches provide reliable, scalable, and secure high-performance foundations for high-density server virtualization, cloud architectures, and next generation flash and SSD storage. They are designed to meet the demands of highly virtualized private cloud storage and data center environments.

This IBM Redbooks® publication helps administrators understand how to implement or migrate to an IBM b-type Gen 7 SAN. It provides an overview of the key hardware and software products and explains how to install, monitor, tune, and troubleshoot your storage area network (SAN).

Read this publication to learn about fabric design, and managing and monitoring your SAN environment.

Authors

This book was produced by a team of specialists from around the world working at IBM Redbooks, Poughkeepsie Center.

Brian Larsen joined Broadcom in July 1991 and has more than 35 years of professional experience in high-end processing, storage, disaster recovery, cloud, virtualization, and networking environments. Larsen is the Director of Partner Business Development and is responsible for solution and business development within all IBM divisions. In addition to his 14 years in Business Development, Larsen has held positions in sales/consulting, product management, and solutions marketing.

Kameswara Bhaskarabhatla joined Broadcom in Oct 2019 as a Solutions Architect and has more than 20 years of professional experience. Before joining Broadcom, he worked for The Trainline and IBM as technical lead for storage. Kamesh reviews storage environments to provide the preferred best practices and architectural decisions that improve the overall availability and reliability.

Matt Brennan joined Broadcom in February 2000 and has more than 35 years of professional experience in development, design, and deployment of storage devices and networks. Matt currently supports the qualification of Broadcom products for IBM. Matt previously worked as a software developer, support engineer, and field sales engineer for storage arrays and networking products.

AJ Casamento is an IT industry veteran with over 40 years of experience working in various market segments, IT leading companies, and engineering and product development roles. In more than 23 years as a Solutions Architect at Brocade, AJ spends his time helping enterprises understand various technologies and architectural decisions, and the impact to their business. AJ started his career with Digital Equipment Corporation, and worked in various engineering and product development roles at Hewlett Packard, SUN Microsystems, IBM, and Avnet, before joining Brocade.

Jack Consoli has more than 35 years of professional experience in high-end processing, storage, disaster recovery, and networking with an emphasis on mainframe environments. Jack is a co-author of *System z End-to-End Extended Distance Guide*, SG24-8047, and a contributor to *FICON Planning and Implementation Guide*, SG24-6497. In addition to sales roles as a Field Applications Engineer and Systems Engineer, Consoli has held positions in product management, engineering management, and development engineering.

Mark Detrick is a Principal Solutions Architect, having worked 20 years for Broadcom. Involved with many Fortune 500 large-scale IBM projects worldwide, Mark plays a consultative role in storage network design and deployment. Mark is an expert in many data center technologies, including mainframe SAN design and extension; distributed systems large-scale SAN design, FC Routing, and extension; IP routing and Ethernet switching; and WAN technologies. Mark has over 30 years of experience in data center networking environments. He understands Broadcom's ASIC technology and is a consultative resource within Brocade and externally to customers, OEMs, and resellers. Mark applies his in-depth technical knowledge of fabrics, protocols, flow control, TCP/IP, and large-scale SAN/LAN architectures. Mark is adept at next-generation data center solutions, including cloud and the many types of virtualizations found in data centers.

Johnny Fonseca has over 20 years of professional experience in IP and optical communications networking, storage, and disaster recovery solutions. As Field Application Engineer, he has taken the lead in designing, implementing, and optimizing storage networks in both Open Systems and Mainframe environments around the world. Customers seek his expertise and consider him a trusted advisor. Johnny began his career in networking with the United States Marine Corps and held senior engineering positions at NTT America, Cisco, McDATA, and Brocade before joining Broadcom (via Brocade's acquisition in 2017).

Tim Jeka leads the technical and positioning effort of Brocade's storage networking business with IBM as an IBM North America Systems Engineer of Storage Networking. Tim is responsible for driving technical awareness and positioning of Brocade's storage networking products to the IBM field across the Americas. He joined Brocade in 2008, bringing over 34 years of experience in Networking, Storage, and SAN technology with IBM. Before joining Brocade, Tim was a Networking and Storage Area Networking RDS at IBM. Tim also held Technical Team leadership roles in IBM's Storage and Networking product divisions for over 34 years with IBM.

Thanks to the following people from Brocade Storage Networking (BSN), Broadcom, Inc., for their contributions to this project:

Silviano Gaona

OEM Field Application Engineer, Partner solutions

Brandon Hoff

Principal Software Architect

Marcus Thordal

Principal Solution Architect

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 IBM Redbooks residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our 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

- Send your comments in an email to:

redbooks@us.ibm.com

- Mail your comments to:

IBM Corporation, IBM Redbooks
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on LinkedIn:

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

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

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

- Stay current on recent Redbooks publications with RSS Feeds:

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

Summary of changes

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

Summary of Changes
for SG24-8497-01
for IBM b-type Gen 7 Installation, Migration, and Best Practices Guide
as created or updated on March 29, 2022.

March 2022, Second Edition

This revision includes the following new and changed information.

Changed information

Updated 11.2.6, “Downloading the SANnav Management Portal software” on page 266.



Introduction to IBM b-type Gen 7

This chapter describes the IBM b-type Gen 7 family of products. It includes an overview of the Gen 7 technical values that were designed into the products for this ever changing world. In addition, an overview of naming conventions is provided to clarify the differences between an IBM b-type Gen 7 offering and a Brocade offering.

This chapter includes the following topics:

- ▶ 1.1, “IBM b-type Gen 7 platforms overview” on page 2
- ▶ 1.2, “IBM b-type Gen 7 platforms and naming conventions” on page 4

1.1 IBM b-type Gen 7 platforms overview

As technology evolves, IT organizations are held to a higher set of standards. Not only are IT organizations responsible for delivering nonstop reliability, they are now judged on how quickly they can deploy new services to the business.

The only way to accelerate IT delivery and keep pace with ever-increasing demands on existing infrastructure is to automate. The infrastructure must monitor application performance, identify network congestion, and prioritize bandwidth. The infrastructure also must isolate configuration errors or malfunctioning devices automatically, regardless of the source of the issue in the data center.

IBM b-type Gen 7 technology delivers a collection of features that combine comprehensive data collection capabilities with powerful analytics to quickly understand the health and performance in the environment and identify potential effects or trending problems. These features are the foundation for realizing a self-learning, self-optimizing, and self-healing autonomous SAN.

IBM b-type Gen 7 technology provides unprecedented visibility and actionable intelligence across the storage network. The information that is captured is displayed in IBM SANnav Management Portal to quickly identify and isolate problems before they affect application availability. The combination of these features gives IT administrators the ability to quickly enable an autonomous infrastructure.

1.1.1 Self-learning

IBM b-type Gen 7 technology proactively monitors millions of I/O performance and behavior data points through integrated network sensors to gain deep insight into the environment. These data points are transformed into actionable intelligence that can monitor and alert when any abnormal changes occur. Through these capabilities, an administrator can identify individual applications and their performance characteristics across the fabric, and identify the performance of the various devices that comprise the fabric (the switches, hosts, and targets):

- ▶ I/O IBM Insight®: Proactively monitors I/O performance and behavior through integrated network sensors to baseline application performance and ensure operational stability. By combining the instrumentation of I/O Insight with the ability to self-learn the traffic flows, the IBM b-type Gen 7 autonomous SAN technology can generate performance and health metrics about each component and the applications to monitor for changes and alert the administrator with Monitoring and Alerting Policy Suite (MAPS).
- ▶ MAPS: Provides an easy-to-use solution for policy-based threshold monitoring and alerting. MAPS proactively monitors the health and performance of any SCSI or NVMe storage infrastructure to ensure application uptime and availability.

By using pre-built, rule-based, and policy-based templates, MAPS simplifies fabric-wide threshold configuration, monitoring, and alerting:

- ▶ **Automatic Flow Learning:** Provides automatic learning of all traffic flows from a specific host to storage across the SAN fabric. With this information, an administrator can automatically identify resource contention or congestion that is affecting application performance.
- ▶ **Fabric Performance Impact:** Uses predefined MAPS policies to automatically detect and alert administrators to different congestion severity levels and to identify credit-stalled devices (for example, misbehaving slow drain devices) or oversubscribed ports that might affect network performance. This feature pinpoints which devices are causing or are affected by the congested port, and quarantines the misbehaving devices.

1.1.2 Self-optimizing

By using actionable intelligence that is gathered from self-learning capabilities, IBM b-type Gen 7 Fibre Channel SANs automatically apply priorities for specific data traffic to help ensure performance levels and monitor for traffic pattern shifts. Learning traffic behavior enables the network to make smarter decisions on traffic prioritization, congestion avoidance, and adjustment to ensure optimal network performance for applications and storage.

When a change occurs, the IBM b-type Gen 7 autonomous SAN technology isolates the port traffic for the misbehaving device to a virtual channel in the fabric and allows all other traffic to circumvent any slow moving traffic to maintain optimal performance.

In addition, IBM SANnav Management Portal includes automated manual activities, such as infrastructure deployment and provisioning, to expedite IT services. Self-optimizing includes the following features:

- ▶ **Traffic Optimizer:** Automatically classifies and separates traffic with similar characteristics to optimize performance for most common SAN configurations. It identifies and isolates traffic flows to prevent negative effects to overall SAN performance.
- ▶ **Advanced traffic shaping:** Ensures application performance by proactively monitoring and actively shaping traffic.
- ▶ **REST APIs:** Eliminates human errors and performance effects through open DevOps.
- ▶ **SAN Automation and Ansible:** Gains cloud-like SAN orchestration for optimizing administration resources through IBM b-type Gen 7 REST API automation technology.

1.1.3 Self-healing

When potential disruptions are detected, the network automatically mitigates or resolves issues without manual intervention. This process is done by proactively monitoring the network to automatically identify abnormal or unexpected infrastructure behavior and then taking immediate action.

These actions include alerting the end devices of the issue through a notification and signaling process within a SAN. End devices can automatically adjust traffic or fail over to a healthy path to mitigate impact until a further infrastructure change solution can be implemented. In addition, these capabilities detect and automatically reconfigure fabric misconfigurations that fall outside best practices.

Self-healing includes the following features:

- ▶ Fabric congestion notification: Automatically detects congestion and notifies end devices to automatically mitigate congestion.
- ▶ Slow drain device quarantine (SDDQ): Automatically quarantines credit-stalled devices to prevent the misbehaving device from affecting the rest of traffic.
- ▶ Automatic actions: Ensures data delivery with automatic failover from physical or congestion issues, such as port decommissioning, port toggling, and port fencing.
- ▶ COMPASS: Detects and automatically reconfigures out-of-compliance configurations.
- ▶ SANnav Management Portal: Reduces troubleshooting steps with built-in best practice recommendations to quickly resolve issues.

Through the modernization process, enterprises are facing the reality that their revenue is intertwined with the success or failure of the IT organization. To succeed, IT must automate as much as possible to eliminate complexity and reduce costs.

IT organizations also must remove tedious, time-consuming, and labor-intensive tasks so that they can focus on delivering services to the business that can help deliver more revenue. IBM b-type Gen 7's autonomous SAN delivers automation and intelligence to administrators so they do not have to worry about the SAN and instead focus on initiatives that are strategic to their organization.

1.2 IBM b-type Gen 7 platforms and naming conventions

Figure 1-1 shows the IBM b-type Gen 7 platforms: 8-slot (IBM SAN512B-7), 4-slot (IBM SAN256B-7) directors, and the mid-range (IBM SAN64B-7) switch.

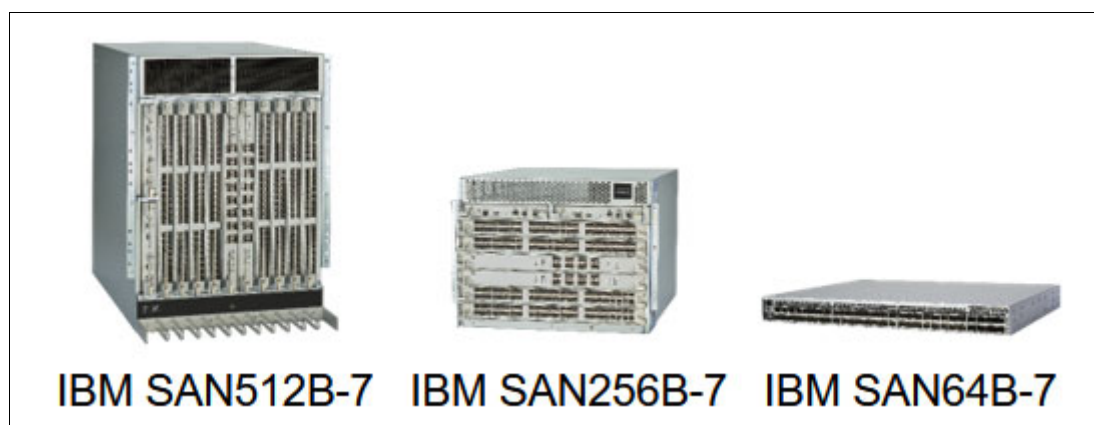


Figure 1-1 IBM b-type Gen 7 platforms

The IBM b-type Gen 7 and Brocade naming conventions for the Gen 7 platform offerings are listed in Table 1-1.

Table 1-1 IBM b-type Gen 7 platform naming conventions

IBM b-type Gen 7 name	IBM machine type and model	Brocade name
IBM Storage Networking SAN512B-7	8961-F78	X7-8
IBM Storage Networking SAN256B-7	8961-F74	X7-4
IBM Storage Networking SAN64B-7	8960-P64 and 8960-R64	G720



IBM b-type Gen 7 product review

This chapter provides a hardware review of each Gen 7 director, supported blades, and mid-range switch. New to the Gen 7 offering is the ability to upgrade from a Gen 6 Director to a Gen 7 supported director.

An overview of the Gen 6 upgrade path to a Gen 7 supported platform along with supported blades also is provided.

The last section of this chapter reviews the supported transceiver modules for the Gen 7 platforms.

This chapter includes the following topics:

- ▶ 2.1, “IBM b-type Gen 7 SAN director overview” on page 6
- ▶ 2.2, “IBM b-type Gen 7 SAN512B-7 (8961-F78) director overview” on page 7
- ▶ 2.3, “IBM b-type Gen 7 SAN256B-7 (8961-F74) director overview” on page 13
- ▶ 2.4, “IBM b-type Gen 7 SAN64B-7 (8960-P64/R64) switch overview” on page 18
- ▶ 2.5, “Upgrading Gen 6 to Gen 7 supported platform overview” on page 22
- ▶ 2.6, “IBM b-type Gen7 transceiver module overview” on page 23

2.1 IBM b-type Gen 7 SAN director overview

Technology is evolving at an incredible pace, and businesses are demanding more from their IT resources and infrastructure. Rapid adoption of flash storage and the ramp-up of NVMe-based storage unleash advancements in application design that drive new levels of performance and capacity requirements, such as advanced analytics, business intelligence, and data-intensive workloads.

With the integration of new technologies that are accelerating the delivery of data and services, the network must evolve to keep pace with innovations in storage and demands of modern applications to maximize the full value of investments in next-generation data center infrastructure.

To meet ever-increasing demands for faster, more reliable data access, it is essential for organizations to deploy a modernized infrastructure that reduces latency, increases bandwidth, and ensures continuous availability. Unprecedented performance is not enough on its own.

Powerful analytics and advanced automation capabilities are required to transform current storage networks into an autonomous SAN. This requires a network that can deliver these capabilities to maximize performance, simplify management, and reduce operational costs.

Legacy infrastructure was not designed to support the performance requirements of evolving workloads and NVMe-based storage. In fact, an aging network impedes the performance of the on-demand data center.

By modernizing the storage network with IBM b-type Gen 7, organizations enable a faster, more intelligent, and more resilient network. This maximizes the performance, productivity, and efficiency of their storage investments and resources, even as they rapidly scale their environments.

The IBM b-type Gen 7 Director provides a modular building block, which is purpose-built for scalability to accommodate growth and power large-scale storage environments. With a 50% latency reduction compared to the previous generation, IBM b-type Gen 7 Directors maximize the performance of NVMe storage and high-transaction workloads, which eliminates I/O bottlenecks and unleashes the full performance of next-generation storage.

In addition, the IBM b-type Gen 7 Director lays the foundation for the autonomous SAN. With autonomous SAN technology, the director harnesses the power of analytics and the simplicity of automation to optimize performance, ensure reliability, and simplify management. Using these capabilities enables organizations to realize a self-learning, self-optimizing, and self-healing SAN.

IBM b-type Gen 7 Directors provide up to 384 64Gbps line rate ports or up to 512 32 Gbps line rate ports, which enables organizations to scale more devices, applications, and workloads. With diverse deployment options, multiprotocol flexibility, and mixed blade capability, organizations can adapt and optimize their businesses to meet next-generation storage and server requirements.

The IBM b-type Gen 7 Director supports the concurrent use of traditional Fibre Channel and NVMe storage traffic, which allows organizations to seamlessly integrate IBM b-type Gen 7 Fibre Channel networks with next-generation NVMe-based storage without a disruptive rip-and-replace.

2.2 IBM b-type Gen 7 SAN512B-7 (8961-F78) director overview

Designed to meet continuous data growth and critical application demands, the IBM b-type Gen 7 Director is purpose-built to power large-scale storage environments that require increased capacity, greater throughput, and higher levels of resiliency and operational efficiency. This modular building block enables organizations to build the highest performing data center SAN fabric that is required for all-flash and NVMe storage environments.

The IBM b-type Gen 7 Director modular design provides flexibility with two customizable chassis that can scale on-demand for more devices, applications, and workloads. Both chassis use IBM b-type Gen 7 UltraScale ICL technology to scale out modular SANs while preserving blade ports for device connectivity. It also enables a flexible SAN design that supports core-edge or mesh topologies.

The 14U IBM b-type Gen 7 SAN512B-7 Director is built for large enterprise networks and has eight vertical blade slots to provide up to 384 64Gbps line rate ports or up to 512 32Gbps line rate ports for device connectivity. Another 32 UltraScale InterChassis Link (ICL) connections provide 128 ports for chassis-to-chassis interconnect.

The 8U IBM b-type Gen 7-4 Director is built for mid-size networks and has four horizontal blade slots to provide up to 192 64 Gbps line rate ports or up to 256 32 Gbps line rate ports for device connectivity.

An additional 16 UltraScale ICL connections provide 64 ports for chassis-to-chassis interconnect. Each blade slot within the IBM b-type Gen 7 chassis can be populated with optional port or extension blades. For device connectivity, the following blades are available:

- ▶ IBM b-type Gen 7 FC64-48 Fibre Channel port blade provides 48 x 64 Gbps Fibre Channel ports with backward-compatibility support for 8, 10, 16, and 32 Gbps Fibre Channel connectivity.
- ▶ IBM b-type Gen 7 FC32-X7-48 Fibre Channel port blade provides 48 x 32 Gbps Fibre Channel ports with backward-compatibility support for 4, 8, 10, and 16 Gbps Fibre Channel connectivity.
- ▶ IBM b-type Gen 7 FC32-64 Fibre Channel port blade provides 64 x 32 Gbps Fibre Channel ports with backward-compatibility support for 4, 8, and 16 Gbps Fibre Channel connectivity and support for 10, 25, and 40GbE FCoE connectivity.

To support disaster recovery and data protection storage solutions over long distances, the IBM b-type Gen 7 SX6 Extension Blade provides flexible Fibre Channel and IP storage replication deployment options with 16 32 Gbps Fibre Channel ports, 16 1/10-GbE ports, and 2 40 GbE ports. This blade allows organizations to seamlessly integrate extension capabilities within the IBM b-type Gen 7 Director to provide replication services for large-scale, multi-site data center environments that implement block, file, and tape data protection solutions.

The IBM b-type Gen 7 SX6 Extension Blade can be deployed with the IBM b-type Gen 7 7840 Extension Switch and the IBM b-type Gen 7 7810 Extension Switch in a data-center to-edge architecture as a cost effective option for connecting primary data centers with remote data centers and offices.

IBM b-type Gen 7 directors build upon years of innovation and use the core technology of IBM b-type Gen 7 systems to consistently deliver five-nines availability in the world's most demanding data centers. Delivering nondisruptive software upgrades, hot-pluggable components, and a no-single-point-of-failure design, the IBM b-type Gen 7 offers a highly resilient solution for today's enterprise-class storage environments.

The device includes the following key product features:

- ▶ Redundant and hot-swappable SFP+, SFP28, QSFP+, and QSFP56 transceivers; port, extension, control processor (CP), and core routing (CR) blades; power supply assemblies, fan assemblies, and WWN cards that enable a high availability platform and allow nondisruptive software upgrades for mission-critical SAN applications.
- ▶ Support for the following features when the FC32-64 blade is used:
 - Up to 16 QSFP ports on each FC32-64 port blade that provide 4x32 Gbps, 4x16 Gbps, 4x8 Gbps, 4x4 Gbps, 4x25GbE, or 40GbE operation. Port speeds for 4x32 Gbps transceivers can auto-negotiate between 4x32 Gbps and 4x16 Gbps. Port speeds for 4x16 Gbps transceivers can auto-negotiate between 4x16 Gbps, 4x8 Gbps, and 4x4 Gbps.
 - Flexport technology that allows you to configure each of the 16 QSFP ports on a blade for FC operation or Ethernet operation for FCoE connections. For more information about supported transceivers and speeds, see [this web page](#).
 - Up to four blades and all FC transceivers installed, up to 256 32 Gbps external ports are possible in a single chassis, which enables high-density SAN configurations with a reduced footprint.
 - Trunking technology that allows groups of up to eight ports to create high-performance 256 Gbps ISL trunks between switches that use 32 Gbps ports.
- ▶ Support for the following features when the FC64-48 blade is used:
 - Port blade with 48 front-end 64 Gbps FC SFP+ ports with the edge switch interconnect to the core blades.
 - Supported SFPs include 10 Gbps, 32 Gbps, and 64 Gbps FC.
 - Support for 48 Fibre Channel ports on each FC64-48 port blade that provide 64 Gbps, 32 Gbps, 16 Gbps, 10 Gbps, and 8 Gbps. The 10 Gbps transceivers can be used for any port on the FC64-48 port blades. The 10 Gbps transceivers on the FC64-48 and SX6 blade are not interchangeable.

For more information about supported transceivers and speeds, see [this web page](#).
 - Trunking technology groups up to eight ports to create high-performance 512 Gbps ISL trunks between switches using 64 Gbps ports.
- ▶ Support for the following features when using the FC32-X7-48 blade:
 - Port blade with 48 front-end 32 Gbps FC SFP+ ports.
 - Supported SFPs include 32 Gbps, 16 Gbps, and 10 Gbps FC.
 - Support for 48 Fibre Channel ports on each FC32-X7-48 port blade that provide 32 Gbps, 16 Gbps, 10 Gbps, 8 Gbps, and 4 Gbps. The 10 Gbps transceivers can be used for any port on the FC32-X7-48 port blades. The 10 Gbps transceivers on the FC32-X7-48 and SX6 blade are not interchangeable.

For more information about supported transceivers and speeds, see [this web page](#).
 - Trunking technology groups up to eight ports to create high-performance 256 Gbps ISL trunks between switches using 32 Gbps ports.
- ▶ Support for the following features when using the SX6 Blade:
 - Support for 16 Fibre Channel ports supporting 4, 8, 16, and 32 Gbps.
 - 16 GbE ports supporting 1 or 10 Gbps.
 - Two GbE ports supporting 40 Gbps on SX6 extension blades.

Note: The 10 Gbps transceivers on the FC64-48 and the FC32-X7-48 are not interchangeable with the IBM b-type Gen 7 SX6 Blade.

For more information about supported transceivers and speeds, see [this web page](#).

- ▶ Considerations for the IBM b-type Gen 7 SX6 Blade:
 - The SX6 extension blades perform as extension platforms to support Fibre Channel (FC) and IBM FICON® data flows and IP-based storage data flows over an IP WAN.
 - Universal ports that self-configure as E_Ports, F_Ports, EX_Ports, M_Ports (mirror ports), and FICON ports.

Note: The 10 Gbps ports on the FC64-48 and FC32-X7-48 port blades can function as E_Ports only.

- ▶ ClearLink Diagnostic port (D_Port) functionality on Fibre Channel ports.
- ▶ Data compression capabilities through the port blades when ports are configured as ISLs.

2.2.1 Hardware components

The device has a modular and scalable mechanical construction that allows a wide range of flexibility in installation, fabric design, and maintenance. The device can be mounted with the cables facing the front or the rear of the equipment rack. It consists of the following hardware components:

- ▶ Up to eight slots for hot-swappable port blade assemblies, which provides up to 384 64 Gbps Fibre Channel ports if FC64-48 port blades are installed or 512 32 Gbps FC ports if FC32-64 port blades are installed.

Flexport technology allows QSFP ports on this blade to be configured for FCoE operation at 4x10 GbE, 4x25 GbE, and 40 GbE speeds when using an FC32-64 port blade.

For a list of supported transceivers for these blades, see [this web page](#).

- ▶ Two half-size slots for control processor (CP) blades:
 - A single active CP blade can control all the ports in the device.
 - The standby CP blade assumes control of the device if the active CP blade fails.
- ▶ Two slots for core routing (CR) blades:
 - A CR blade interconnects all port blades.
 - The blades provide up to 32 Gen 7 QSFP56 (ICL) ports.
 - ICL ports allow interconnection with neighboring director chassis.
 - Both CR blades are active and can be hot-swapped.

For a list of supported transceivers for these blades, see [this web page](#).

- ▶ Up to four slots for modular, hot-swappable 34-port SX6 extension blades.

Blades provide 16 32 Gbps Fibre Channel (FC) ports supporting 8, 16, and 32 Gbps or 16 32 Gbps FC ports supporting 4, 8, 16, and 32 Gbps FC ports. Extension blades enable long-distance communication over an IP infrastructure.

For a list of supported transceivers for these blades, see [this web page](#).

- ▶ Modular, hot-swappable field-replaceable units (FRUs):
 - Three fan assemblies, available with NPI or NPE airflow.
 - Up to four power supplies, available with NPI or NPE airflow. For more information, see [this web page](#).

See the “Power Supply Specifications (per PSU)” in the IBM b-type Gen 7 Director Technical Specifications for the following information:

- The maximum output power, input voltage, input line frequency, and other specifications for your power supply model.
- The minimum number of power supplies required for operation and redundancy when different input voltages are applied, such as low line and high line AC.

See the “Power Consumption” sections in the IBM b-type Gen 7 Director Technical Specifications for power output data and the minimum number of power supplies for supported input voltages.

Redundant primary power connections ensure high availability. Each power supply assembly has its own connector, so the number of primary power connections is four for optimum efficiency and redundancy.

- ▶ Two World Wide Name (WWN) cards on the non-port side of the device behind the WWN card bezel.
- ▶ Port blades use small form-factor pluggable (SFP+, QSFP+, and QSFP28) optical transceivers. For more information about supported transceivers per blade type, see [this web page](#).
- ▶ Core routing blades use QSFP28 or QSFP56 optical transceivers. For a list of supported transceivers for these blades, see [this web page](#).
- ▶ Chassis door, which must be installed to meet EMI compliance certification.
- ▶ A cable management comb. These combs install on the chassis below the blades for cable management.

Note: Device control processors and management modules include batteries for RTC/NVRAM backup. Do *not* attempt to replace these batteries. Dispose of hardware components that contain these batteries as required by local ordinances and regulations.

2.2.2 Port-side view of the device

The port-side view that is shown in Figure 2-1 on page 11 shows the IBM SAN512B-7 Director with installed blades. The following components are identified:

1. Air Vent
2. Core Routing Blades (CR64-8)
3. Port Blades (FC64-48 or FC32-X7-48)
4. Cable Management Comb
5. Control Processor Blades (CPX7) – Slot 1 (upper), Slot 2 (lower)

Note: The SX6 extension blades is not shown in Figure 2-1 on page 11, but is installed in the same slots as the FC32-X7-48 or FC64-64 port blades. A maximum of four SX6 blades are supported.

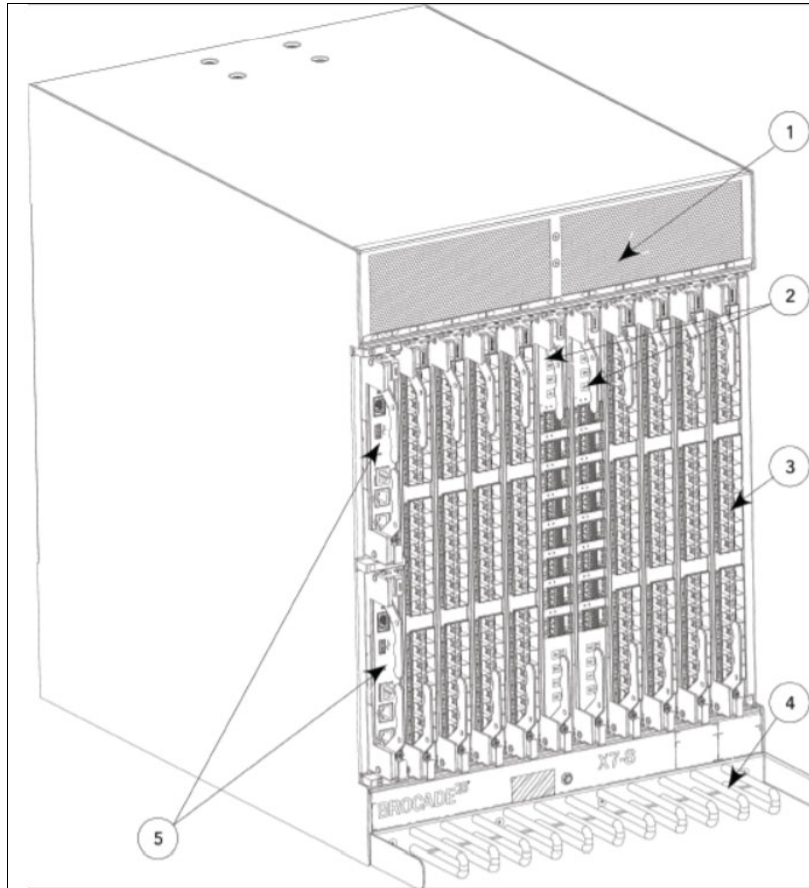


Figure 2-1 Port-Side View of IBM SAN512B-7 Director

2.2.3 Port-side slot numbering

The IBM B-TYPE SAN512B-7 Director contains 10 full-height slots and two half-height slots, for a total of 12 slots. Facing the port side of the device, the half-height slots are on the left, numbered 1 (top slot) and 2 (bottom slot). The remaining full-height slots are numbered 3 - 12, counting from left to right of the chassis. Slots contain guide pins and connectors designed for specific blade types.

Install only the control processor (CP), core routing (CR), port, and extension blades into slot numbers as follows:

- ▶ Slots 1 - 2 are restricted to CP blades. The blade that is installed in slot 1 is designated as CP0; the blade in slot 2 is designated as CP1 in CLI command and message output.
- ▶ Slots 3 - 6 and slots 9 - 12 are restricted to port and extension blades.
- ▶ Slots 7 and 8 are restricted to CR64-8 core blades.

2.2.4 Non-port side view of the device

Figure 2-2 shows the non-port-side view of the IBM SAN512B-7 Director with all fan and power supply assemblies installed. The following components are identified:

1. WWN Bezel (Logo Plate - WWN Cards Behind)
2. Power Supply Assembly
3. Fan Assembly
4. 2AWG Panduit LCD2-14AF Lug for Building Ground Connection

Note: Depending on the fans and power supplies that are installed, airflow can be from the port side to the non-port side of the chassis or from the non-port side to the port side of the chassis.

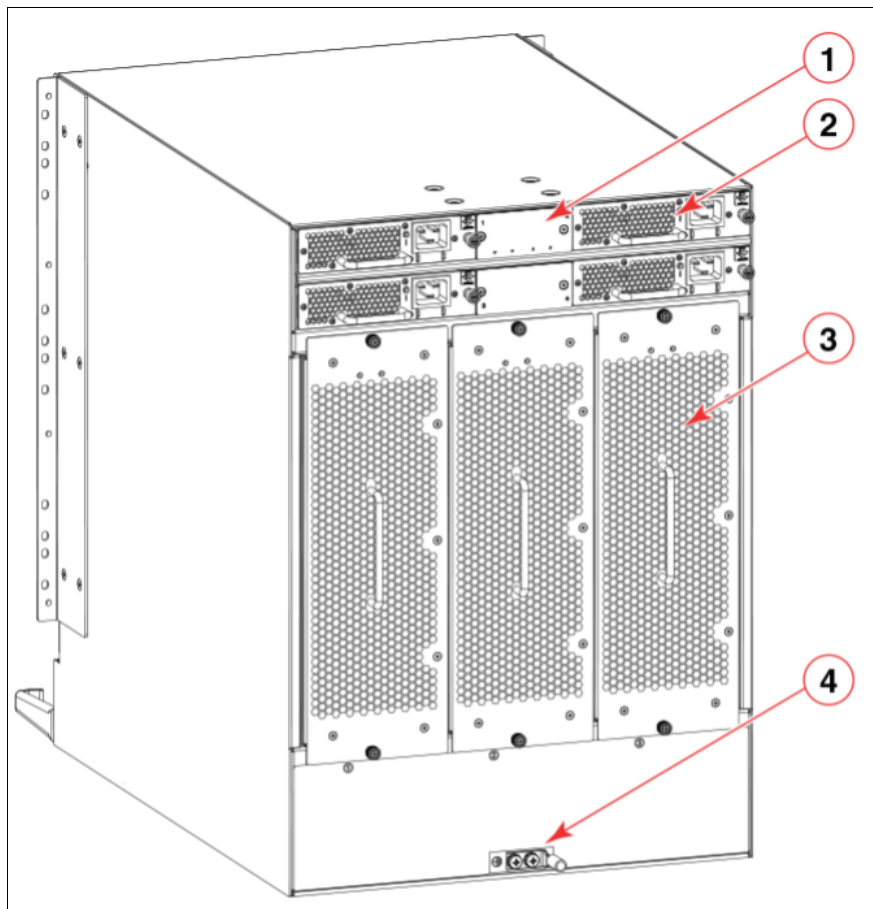


Figure 2-2 Non-port side of the IBM SAN512B-7 Director (example configuration)

Although not shown, the chassis label that contains the serial number, SKU, and WWN is on the lower portion of the chassis, which is below the fan assemblies.

2.2.5 More information

All documentation for the IBM b-type Gen 7 SAN512B-7 (8961-F8) is available at [this web page](#).

2.3 IBM b-type Gen 7 SAN256B-7 (8961-F74) director overview

This device includes the following key product features:

- ▶ Redundant and hot-swappable SFP+, SFP28, QSFP+, and Gen 7 QSFP transceivers; port, extension, control processor (CP), and core routing (CR) blades; power supply assemblies, fan assemblies, and WWN cards that enable a high-availability platform and allow nondisruptive software upgrades for mission-critical SAN applications.
- ▶ Support for the following features when using the FC32-64 blade:
 - Up to 16 QSFP ports on each FC32-64 port blade that provide 4x32 Gbps, 4x16 Gbps, 4x8 Gbps, 4x4 Gbps, 4x25 GbE, or 40 GbE operation. Port speeds for 4x32 Gbps transceivers can auto-negotiate between 4x32 Gbps and 4x16 Gbps. Port speeds for 4x16 Gbps transceivers can auto-negotiate between 4x16 Gbps, 4x8 Gbps, and 4x4 Gbps.
 - Flexport technology that allows you to configure each of the 16 QSFP ports on a blade for FC operation or Ethernet operation for FCoE connections. For more information about supported transceivers and speeds, see [this web page](#).
 - Up to four blades and all FC transceivers installed, up to 256 32 Gbps external ports are possible in a single chassis, enabling high-density SAN configurations with a reduced footprint.
 - Trunking technology that allows groups of up to eight ports to create high-performance 256 Gbps ISL trunks between switches using 32 Gbps ports.
- ▶ Support for the following features when using the FC64-48 blade:
 - Port blade with 48 front-end 64 Gbps FC SFP+ ports with the edge switch interconnect to the core blades.
 - Supported SFPs include 10 Gbps, 32 Gbps, and 64 Gbps FC.
 - Support for 48 Fibre Channel ports on each FC64-48 port blade that provide 64 Gbps, 32 Gbps, 16 Gbps, 10Gbps, and 8 Gbps. The 10 Gbps transceivers can be used for any port on the FC64-48 port blades. The 10 Gbps transceivers on the FC64-48 and SX6 blade are not interchangeable.
For more information about supported transceivers and speeds, see [this web page](#).
 - Trunking technology groups up to eight ports to create high-performance 512 Gbps ISL trunks between switches using 64 Gbps ports.
- ▶ Support for the following features when the FC32-X7-48 blade is used:
 - Port blade with 48 front-end 32 Gbps FC SFP+ ports with edge switch interconnect to the core blades.
 - Supported SFPs include 32 Gbps, 16 Gbps, and 10 Gbps FC.
 - Support for 48 Fibre Channel ports on each FC32-X7-48 port blade that provide 32 Gbps, 16 Gbps, 10 Gbps, 8 Gbps, and 4 Gbps. The 10 Gbps transceivers can be used for any port on the FC32-X7-48 port blades.
The 10 Gbps transceivers on the FC32-X7-48 and SX6 blade are not interchangeable. For more information about supported transceivers and speeds, see [this web page](#).
 - Trunking technology groups up to eight ports to create high-performance 256 Gbps ISL trunks between switches using 32 Gbps ports.
- ▶ Support for the following features when the SX6 extension blade is used:
 - 16 Fibre Channel ports supporting 4, 8, 16, and 32 Gbps.
 - 16 GbE ports supporting 1 or 10 Gbps.

- Two GbE ports supporting 40 Gbps on SX6 extension blades.
The 10 Gbps transceivers on the FC64-48 and the FC32-X7-48 are not interchangeable with the IBM b-type Gen 7 SX6 Blade. For more information about supported transceivers and speeds, see [this web page](#).
- ▶ Consider the following points regarding the IBM b-type Gen 7 SX6 Blade:
 - The SX6 extension blades perform as extension platforms to support FC and FICON data flows and IP-based storage data flows over an IP WAN.
 - Universal ports that self-configure as E_Ports, F_Ports, EX_Ports, M_Ports (mirror ports), and FICON ports.
The 10 Gbps ports on the FC64-48 port blade can function as E_Ports only.
- ▶ ClearLink Diagnostic port (D_Port) functions on FC ports.
- ▶ Data compression capabilities through the port blades when ports are configured as ISLs.

2.3.1 Hardware components

The device features a modular and scalable mechanical construction that allows a wide range of flexibility in installation, fabric design, and maintenance. The device can be mounted with the cables facing the front or rear of the equipment rack. It consists of the following hardware components:

- ▶ Up to four slots for hot-swappable port blade assemblies, providing up to 192 64 Gbps or 256 32 Gbps FC ports for device or ISL connectivity. Flexport technology allows QSFP ports on this blade to be configured for FCoE operation at 4x10 GbE, 4x25 GbE, and 40 GbE speeds when using an FC32-64 port blade. For a list of supported transceivers for these blades, see [this web page](#).
- ▶ Two half-size slots for control processor (CP) blades:
 - A single active CP blade can control all the ports in the device.
 - The standby CP blade assumes control of the device if the active CP blade fails.
- ▶ Two slots for core routing (CR) blades:
 - The CR blade interconnects all port blades.
 - The blades support up to 16 Gen 7 QSFP56 (ICL) ports.
 - ICL ports allow interconnection with neighboring director chassis.
 - Both CR blades are active and can be hot-swapped.

For a list of supported transceivers for these blades, see [this web page](#).
- ▶ Up to four slots for modular, hot-swappable 34-port SX6 extension blades. Blades provide 16 32 Gbps Fibre Channel (FC) ports supporting 4, 8, 16, and 32 Gbps FC ports. Extension blades enable long-distance communication over an IP infrastructure. For a list of supported transceivers for these blades, see [this web page](#).
- ▶ Modular, hot-swappable field-replaceable units (FRUs):
 - Two fan assemblies, available with NPI or NPE airflow.
 - Up to two power supplies, available with the NPI or NPE airflow.

For more information, see [this web page](#).

Also refer to the following documentation:

- See the “Power Supply Specifications (per PSU)” in the IBM b-type Gen 7 Director Technical Specifications for maximum output power, input voltage, input line frequency, and other specifications for your power supply model.
- See the “Power Supply Requirements” section in the IBM b-type Gen 7 Director Technical Specifications for the minimum number of power supplies required for operation and redundancy when different input voltages are used, such as low line and high line AC.
- See the “Power Consumption” sections in the IBM b-type Gen 7 Director Technical Specifications for power output data and the minimum number of power supplies for supported input voltages.

Redundant primary power connections ensure high availability. Each power supply assembly has its own connector; therefore, the number of primary power connections is two for optimum efficiency and redundancy.

- Two World Wide Name (WWN) cards on the non-port side of the device behind the WWN card bezel.
- Port blades use small form-factor pluggable (SFP+, QSFP+, and QSFP28) optical transceivers. For more information about supported transceivers per blade type, see [this web page](#).
- Core routing blades use QSFP56 optical transceivers. For a list of supported transceivers for these blades, see [this web page](#).
- ▶ Chassis door, which door must be installed to meet EMI compliance certification.
- ▶ Two vertical cable management finger assemblies. These combs install on the equipment rack for cable management. For more information, see [this web page](#).

Note: Device control processors and management modules include batteries for RTC/NVRAM backup. Do *not* attempt to replace these batteries. Dispose of hardware components that contain these batteries as required by local ordinances and regulations.

2.3.2 Port-side view of the device

Figure 2-3 shows the port-side view of the IBM SAN256B-7 Director with installed blades.

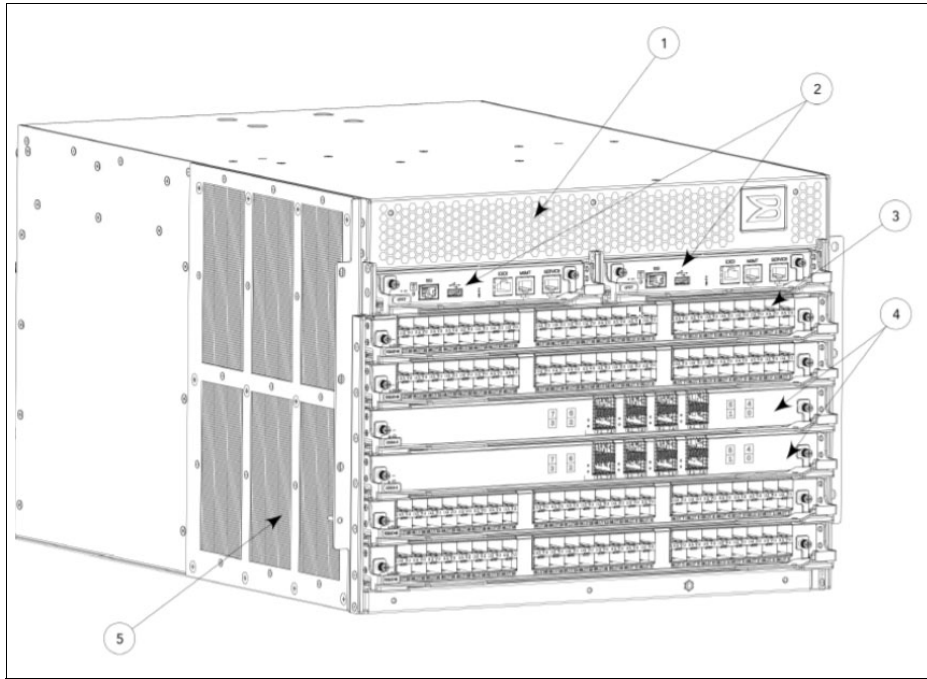


Figure 2-3 Port-side view of the IBM SAN256B-7 Director (example configuration)

The following components are identified:

1. Front Air Vent
2. Control Processor Blades (CPX7) :
 - a. Slot 1 (left)
 - b. Slot 2 (right)
3. Port Blades (FC64-48 or FC32-X7-48)
4. Core Routing Blades (CR64-4)
5. Side Air Vent

Note: The SX6 extension blade is not shown in the figure, but is installed in the same slots as the FC32-X7-48 or FC64-48 port blades. A maximum of four SX6 blades is supported.

Note: When non-port-side exhaust (NPE) fan and power supply assemblies are installed, air flows through the side vent and exhausts to the non-port side. When non-port-side intake (NPI) fan and power supply assemblies are installed, air flows from the non-port-side and exhausts through the side vent.

When installing the chassis in a four-post rack with the airflow diversion rack mount kit, airflow is redirected to or from the port side of the chassis. Therefore, if NPE fan and power supply assemblies are installed, air flows into port-side air vents and exhausts to the non-port side. If NPI fan and power supply assemblies are installed, air flows into non-port-side air vents and exhausts through to the port-side.

2.3.3 Port-side slot numbering

The IBM SAN256B-7 Director contains 6 full-width slots and two half-width slots, for a total of 8 slots. Facing the port side of the device, the half-width slots are on the top, numbered 1 and 2 from left to right. The remaining full-width slots are numbered 3 - 8, counting from top to bottom of the chassis.

Slots contain guide pins and connectors that are designed for specific blade types. Only install the control processor (CP), core routing (CR), port, and extension blades into slot numbers as follows:

- ▶ Slots 1 - 2 are restricted to CP blades. The blade that is installed in slot 1 is designated as CP0; the blade in slot 2 is designated as CP1 in CLI command and message output.
- ▶ Slots 3 and 4 and slots 7 and 8 are restricted to port and extension blades.
- ▶ Slots 5 and 6 are restricted to CR64-4 blades.

2.3.4 Non-port-side view of the device

Figure 2-4 shows the non-port-side view of the IBM SAN256B-7 Director with all fan and power supply assemblies installed.

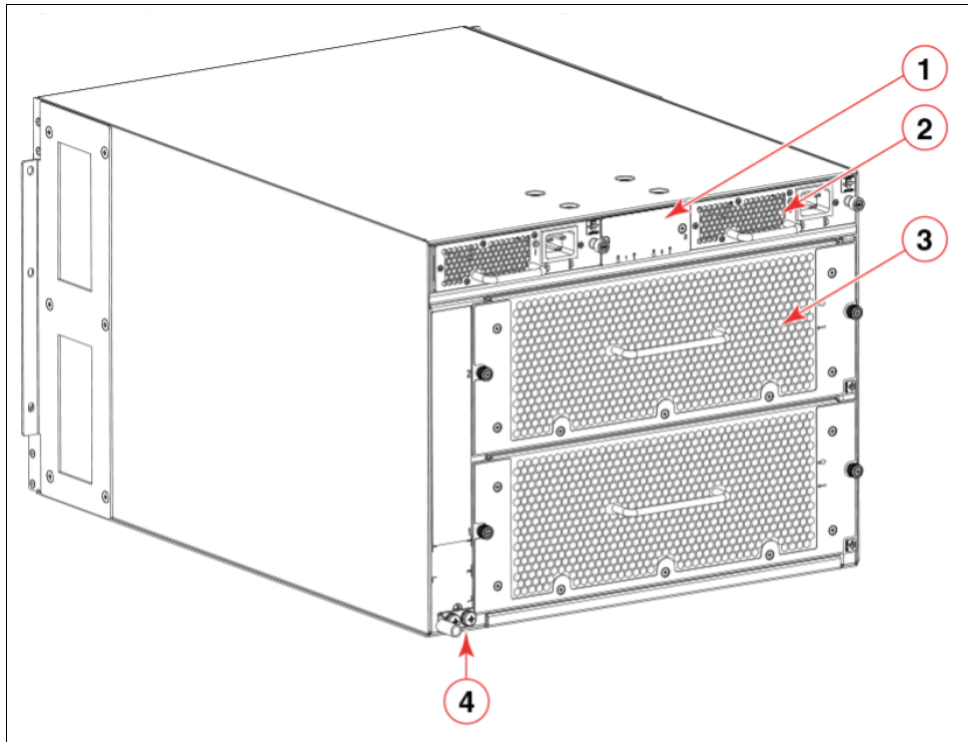


Figure 2-4 Non-port side of the IBM SAN256B-7 Director (example configuration)

1. WWN Bezel (logo plate - WWN cards behind)
2. Power Supply Assembly
3. Fan Assembly
4. 2AWG Panduit LCD2-14AF Lug for Building Ground Connection

Depending on the fans and the power supplies that are installed, airflow can be from the port side to the non-port side or from the non-port side to the port side of the chassis through the side air vent.

Use the Chassis Airflow Diversion and Port Side Exhaust Kit to divert airflow so that air fully exhausts to the port side of the chassis or is fully drawn from the port side of the chassis while mounted in a four-post rack.

Although not shown, the chassis label that contains the serial number, SKU, and WWN is on the upper portion of the chassis to the left of the fan assemblies.

2.3.5 More information

For more information about the IBM b-type Gen 7 SAN256B-7 (8961-F4), see [this web page](#).

2.4 IBM b-type Gen 7 SAN64B-7 (8960-P64/R64) switch overview

The IBM SAN64B-7 Switch offers the following features and capabilities.

Note: The IBM SAN64B-7 switch is secure-booted.

- ▶ Enterprise class 56-port Gen 7 Fibre Channel switch supported with Fabric OS (FOS) v9.0.0 that offers 56 x 64 Gbps SFP+ ports in a 1U form-factor.
- ▶ Each of the 56 SFP+ ports supports 8, 10, 16, 32, and 64 Gbps Fibre Channel speeds.
- ▶ 64 Gbps auto-sensing Fibre Channel switch and router ports.:
 - A 64 Gbps optical transceiver can auto-negotiate to 64 Gbps, 32 Gbps, or 16 Gbps.
 - A 32 Gbps optical transceiver can auto-negotiate to 32 Gbps, 16 Gbps, or 8 Gbps.
 - A 10 Gbps optical transceiver can auto-negotiate to 10 Gbps.

Note: The port speed is determined by the maximum speed that is supported by the optical transceiver at the other end of the link.

- ▶ 10 Gbps manually set capability on FC ports:
 - 10 Gbps performance is enabled by 10 Gbps SFP+ transceivers.
 - Ports can be configured for 10 Gbps for metro connectivity.
- ▶ Dynamic Ports on Demand (Dynamic-POD) scaling from a base configuration of 24 ports to 56 ports (four 8-port SFP+ PODs).
- ▶ Universal ports self-configure as a E_Ports, F_Ports, N_Ports, or D_Ports. EX_Ports can be activated on a per-port basis with the optional Integrated Routing license.

A Diagnostic Port (D_Port) provides diagnostics, troubleshooting, and verification services for the physical media.
- ▶ In-flight 64 Gbps data compression and encryption provide efficient link utilization and security. The number of ports that can be enabled with compression and encryption is listed in Table 2-1 on page 19.

Table 2-1 IBM SAN64B-7 encryption and compression: number of supported ports

Port speed	Encryption-only	Compression-only	Compression and encryption
64 Gbps ^a	4 ports	4 ports	4 ports
32 Gbps ^a	4 ports	4 ports	4 ports
16 Gbps ^a	4 ports	4 ports	4 ports
10 Gbps ^a	4 ports	4 ports	4 ports
8 Gbps	4 ports	4 ports	4 ports

a. Supported on SFP ports

- ▶ Support for port-side exhaust or non-port-side exhaust airflow for cooling.
- ▶ Hardware-enabled input and output (I/O) latency statistics collection.
- ▶ Hardware-enabled VM support.
- ▶ IBM b-type Gen 7 small form-factor pluggable plus (SFP+) optical transceivers support any combination of Short Wavelength (SWL) and Long Wavelength (LWL) optical media among the switch ports.
- ▶ Extended distance Fibre Channel to support long-distance native FC connectivity.
- ▶ 10 Gbps Fibre Channel integration on any selected port provides DWDM metro connectivity on the same switch.
- ▶ Port-to-port latency is minimized to 460 ns (including FEC) is minimized by using cut-through frame switching.
- ▶ High performance T1022E processor with two cores operating at 1.2 GHz delivers high performance, scalability, and advanced Fabric Vision functions.
- ▶ One 1000/100 Mbps RJ-45 connector for the Ethernet management connection. With EZSwitchSetup, this port supports switch IP address discovery and configuration, which eliminates the need to attach a serial cable to configure the switch IP address.
- ▶ One internal e-USB module provides 2 GB of persistent storage, increased serviceability, and error logging functionality by facilitating easier firmware upgrades and downloads of the system log files.
- ▶ One external USB connector.
- ▶ Two hot-swappable redundant integrated 350 W AC power supplies and fan assembly field-replaceable units.
- ▶ A total of 56 hot-pluggable SFP+ optical transceiver slots.
- ▶ A total of 48 LEDs (green/amber) for the first 48 SFP+ ports and 8 LEDs (green/amber/light-blue) for the last 8 SFP+ ports.
- ▶ One green LED to indicate valid system power.
- ▶ One bi-color (green/amber) LED to indicate the system status.
- ▶ Two Ethernet LEDs: one green LED to indicate link speed of 1000/100/10 Mbps and one green LED to indicate activity.
- ▶ EEPROM for switch identification.
- ▶ Console port. One RS-232 (UART), for terminal access by way of mini-USB.
- ▶ Encryptable EEPROM.
- ▶ Secure encryptable EEPROM.

Note: The License ID is protected by the secure encryptable SEEPROM.

- ▶ Secure boot.
- ▶ Real-time:
 - Power monitoring.
 - Voltage monitoring.
 - Fan monitoring, including airflow direction.
 - Digital thermometers for temperature monitoring.
- ▶ Real-time clock (RTC) with battery.

2.4.1 Software license options

The IBM SAN64B-7 Switch uses a capacity-based Ports on Demand (POD) license method. An Integrated Routing (IR) license is required to enable EX_Ports on this device. A license for software release is also required.

For more information, see [this web page](#).

2.4.2 Port-side view of the device

Figure 2-5 on page 21 shows the port-side view of the IBM SAN64B-6 Fibre Channel switch. The following components are identified:

1. System Power LED
2. System Status LED
3. Management Ethernet Port
4. SFP+ FC (upper and lower) Ports 0–7
5. SFP+ FC (upper and lower) Ports 8–15
6. SFP+ FC (upper and lower) Ports 16–23
7. SFP+ FC (upper and lower) Ports 24–31
8. SFP+ FC (upper and lower) Ports 32–39
9. SFP+ FC (upper and lower) Ports 40–47
10. SFP+ upper Port 48
11. SFP+ upper Port 50
12. SFP+ lower Port 52
13. SFP+ lower Port 54
14. SFP+ upper Port 56
15. SFP+ upper Port 58
16. SFP+ lower Port 60
17. SFP+ lower Port 62
18. SFP+ (lower) Port 14 Status LED
19. SFP+ (upper) Port 10 Status LED
20. USB Port 21. UART mini-USB Serial Console Port

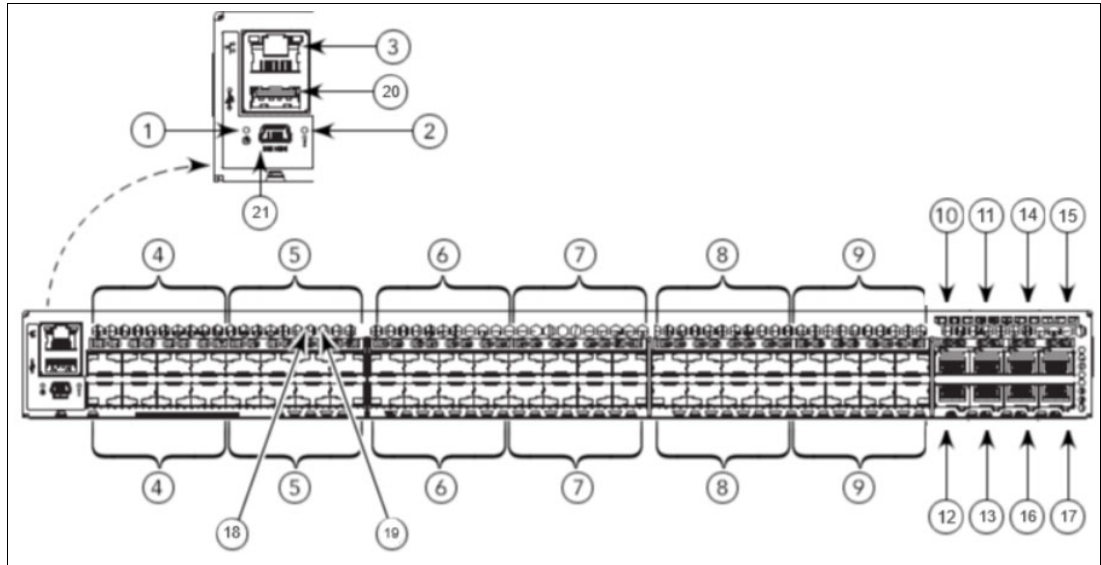


Figure 2-5 Port-Side View of the IBM SAN64B-6

Note: All the ports are connected to a single ASIC.

2.4.3 Nonport-side view of the device

Figure 2-6 shows the nonport-side view of the IBM SAN64B-7 Fibre Channel switch. The following components are identified:

1. #6-32 for Screw Mounting of the Ground Cable
2. Ground Marking
3. Power Supply and Fan Assembly 2
4. Power Supply and Fan Assembly 1
5. Captive Screw
6. Handle
7. Power Supply and Fan Assembly Status LED
8. Power-On Switch
9. Power Supply Receptacle

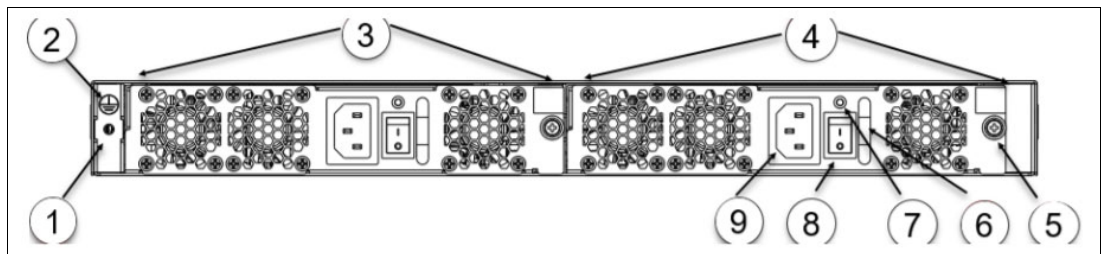


Figure 2-6 Nonport-side view with AC power supply and fan assembly units

2.4.4 More information

For more information about the documentation for the IBM b-type Gen 7 SAN64B-7 (8960-P64/R64), see [this web page](#).

2.5 Upgrading Gen 6 to Gen 7 supported platform overview

The IBM b-type Gen 7 X6 chassis was designed to be field upgradeable from a Gen 6 director to a Gen 7 director. This section provides a high-level overview of the steps to complete a conversion from Gen 6 to Gen 7 platforms.

Note: Upgrading to a Gen 7 director is permanent on the CPs. Therefore, you cannot downgrade to Gen 6 after the upgrade to Gen 7 and you cannot use the same CP blades in a Gen 6 chassis.

2.5.1 Migration steps

The process to migrate your Gen 6 director (X6-4 and X6-8) running FOS 8.x or 9.x to a Gen 7 director running FOS 9.x. includes the following tasks:

- ▶ Upgrade to FOS 9.0.0a.
- ▶ Take a snapshot of the existing configuration.
- ▶ Convert the CP blades from Gen 6 to Gen 7.
- ▶ Swap the core blades.
- ▶ Restore the original configuration or default configuration.

Note: A terminal console connection must be established to both CPs when migrating.

Figure 2-7 shows the functional differences between the three IBM b-type Gen 7 chassis that support FOS 9.0.0a.

Software and Hardware	X6 Chassis	X6 → X7 Upgrade	X7 Chassis Shipped from the Factory
Fabric OS	Fabric OS v8.0 shipped from the factory. Can upgrade to Fabric OS v9.0.0a.	Fabric OS upgrade to v9.0.0a or later.	Fabric OS v9.0.0a shipped from the factory.
Control processor blade	CPX6 blade	CPX6+ (a CPX6 blade with new Gen 7 FPGA firmware). No backward compatibility with an X6 chassis or Fabric OS 8.x.	New CPX7 blade has up-to-date Gen 7 FPGA firmware. No backward compatibility with an X6 chassis or Fabric OS 8.x.
Core routing blade	CR32	CR64	CR64
SWBD (4 or 8 slot) Use the <code>sin root</code> command to determine the SWBD value or the <code>slot.show</code> command with no parameters.	165 (4-slot) or 166 (8-slot)	179 (4-slot) or 180 (8-slot)	179 (4-slot) or 180 (8-slot)
Secure Boot feature	Not supported.	Not supported.	Supported.

Figure 2-7 Functional differences between the chassis supporting Fabric OS 9.0.0a

Figure 2-8 on page 23 shows the Gen 6 and Gen 7 blade support.

Blade	Blade ID	X6 Chassis	X6 → X7 Upgrade	X7 Chassis
FC32-48	178	✓	✓	✓
FC32-64	204	✓	✓	✓
SX6	186	✓	✓	✓
CR32-4/CR32-8	176/177	✓	Not supported	Not supported
FC32-X7-48	218	Not supported	✓	✓
FC64-48	216	Not supported	✓	✓
CR64-4/CR64-8	214/215	Not supported	✓	✓
CPX6	175	✓	Not supported	Not supported
CPX6+	175	Not supported	✓	Not supported
CPX7	220	Not supported	Not supported	✓

Figure 2-8 Blade support matrix

For documentation about Gen 6 to Gen 7 migration procedures, contact your IBM technical support representative.

2.6 IBM b-type Gen7 transceiver module overview

The IBM SAN512B-7, IBM SAN256B-7, and IBM SAN64B-7 platforms are based on the Brocade Condor 5 ASIC. Brocade Condor 5 ASIC platforms (Gen 7) must use specific transceivers that are listed in Table 2-2. Previous versions of SFP+ transceivers are not recognized by a Condor 5-based port blades or switch ports.

Table 2-2 IBM SAN512B-7, IBM SAN256B-7 and IBM SAN64B-7 supported transceivers

	Description	IBM SKU	IBM Part Number	IBM Feature Code	SAN512B-7 & SAN256B-7	FC64-8 Port Blade	FC32-X7 48 Port Blade	SAN64B-7
Gen 7 FC QSFP+	SWL	IB-000420	03GM315	2624	X			
10G FC SFP+	SWL	IB-000418	01JC462	2621		X	X	X
	LWL	IB-000417	01JC459	2620		X	X	X
16G FC SFP+	SWL	IB-000492 IB-000493	01JC494 01JC496	2615 2616			X	
	LWL	IB-000498 IB-000499	03GM300 03GM303	2617 2618			X	
	ELWL - 25 km	IB-000458	03GM306	2619			X	
32G FC SFP+	SWL	IB-000412 IB-000413	01JC449 01JC451	2625 2626		X	X	X
	LWL - 10 km	IB-000438 IB-000439	IB-000439 01JC456	2627 2628		X	X	X
4x 32G FC QSFP+	SWL	IB-000475	03GM309	2622	X			

For more information about supported transceivers for all IBM b-type Gen 7 platforms, see [this web page](#).



Gen 7 Fabric Operating System overview

Gen 7 Fabric Operating System (FOS) v9.0 introduces a broad set of new features and enhancements. It is a modern operating system that provides a solid foundation to simplify management, improve performance, increase automation, and strengthen security.

This chapter includes the following topics:

- ▶ 3.1, “Autonomous SAN” on page 26
- ▶ 3.2, “Automation” on page 28
- ▶ 3.3, “FOS 9.x features” on page 30
- ▶ 3.4, “Management tools” on page 34
- ▶ 3.5, “Security” on page 36
- ▶ 3.6, “Traffic Optimizer” on page 36
- ▶ 3.7, “Fabric Notifications” on page 39
- ▶ 3.8, “IBM b-type switch Fabric Notifications solution” on page 41

3.1 Autonomous SAN

Enterprise data centers continue to push virtualization at higher density and support a much broader set of workloads. More virtual machines (VMs) are being deployed on the server end, and more storage LUNs are provisioned on storage arrays.

At the same time, adoption of all flash storage has accelerated. Customers are starting to adopt next generation NVMe-based storage that offers much lower latency access to data. The use of virtualization, flash storage, and automation tools allows applications and services to be deployed faster while shattering performance barriers.

The unprecedented number of application and service interactions also increased the complexity, risk, and instability of mission-critical operations. For example, increasing workload virtualization leads to a higher number of application flows, but reduces the visibility of the flows.

At the same time, storage virtualization increases the number of short Fibre Channel (FC) frames and the potential for congestion. As a result, IT organizations require storage networks that deliver ultra-low latency, higher-capacity bandwidth, and greater reliability.

To meet service-level agreements (SLAs) or achieve service-level objectives (SLOs), IT administrators also need new tools that can help ensure nonstop operations, quickly identify potential points of congestion, and maximize application performance, while simplifying administration.

In the past, the process of problem detection and mitigation relied on the diligence, skill, and experience of the fabric administrator. Usually, a manual task of searching and tracing error conditions was the regimen. The operational cost and potential disruptive impact to critical applications are simply acceptable in today's data center.

To help administrative automation and improve network uptime, Broadcom developed Autonomous SAN technology that offers self-learning, self-optimizing, and self-healing capabilities. Autonomous SAN is realized through Fabric Vision technology that combines built-in capabilities in IBM b-type Gen 6 and Gen 7 platforms, Brocade Fabric OS (FOS), and Brocade SANnav Management Portal features. Fabric Vision provides detailed monitoring and alerting and response and mitigation, which vastly improves the fabric administrator's insight and response time.

3.1.1 Self-learning

Brocade technology proactively monitors millions of I/O performance and behavior data points through integrated network sensors to gain deep insight into the environment. These data points are transformed into actionable intelligence that can monitor and alert when there are any abnormal changes.

Through the following capabilities, an administrator can identify individual applications and their performance characteristics across the fabric. They also can identify the performance of the various devices that comprise the fabric: the switches, hosts, and targets:

- ▶ **I/O Insight**

Proactively monitors I/O performance and behavior through integrated network sensors to baseline application performance and ensure operational stability. By combining the instrumentation of I/O Insight with the ability to self-learn the traffic flows, the Brocade autonomous SAN technology can generate performance and health metrics on each component and the applications to monitor for changes and alert the administrator with MAPS.

- ▶ **Monitoring and Alerting Policy Suite (MAPS)**

Provides an easy-to-use solution for policy-based threshold monitoring and alerting. MAPS proactively monitors the health and performance of any SCSI or NVMe storage infrastructure to ensure application uptime and availability. By using pre-built, rule-based, and policy-based templates, MAPS simplifies fabric-wide threshold configuration, monitoring, and alerting.

- ▶ **Automatic Flow Learning**

Provides automatic learning of all traffic flows from a specific host to storage across the SAN fabric. With this information, an administrator can automatically identify resource contention or congestion that is affecting application performance.

- ▶ **Fabric Performance Impact**

Uses predefined MAPS policies to automatically detect and alert administrators to different congestion severity levels and to identify credit-stalled devices (for example, misbehaving slow drain devices) or oversubscribed ports that might affect network performance. This feature pinpoints which devices are causing or are affected by the congested port, and quarantines the misbehaving devices.

3.1.2 Self-optimizing

By using actionable intelligence gathered from self-learning capabilities, Brocade Fibre Channel SANs automatically apply priorities for specific data traffic to help ensure performance levels and monitor for traffic pattern shifts.

Learning traffic behavior enables the network to make smarter decisions about traffic prioritization, congestion avoidance, and adjustment to ensure optimal network performance for applications and storage. When something does change, the Brocade autonomous SAN technology isolates the port traffic for the misbehaving device to a virtual channel in the fabric and allow all other traffic to go around to maintain optimal performance.

In addition, Brocade SANnav Management Portal includes automated manual activities, such as infrastructure deployment and provisioning, to expedite IT services. Self-optimizing includes the following features:

- ▶ **Traffic Optimizer:** Automatically classifies and separates traffic with similar characteristics to optimize performance for most common SAN configurations. It identifies and isolates traffic flows to prevent negative effects to overall SAN performance.
- ▶ **Advanced traffic shaping:** Ensures application performance by proactively monitoring and actively shaping traffic.
- ▶ **REST APIs:** Eliminates human errors and performance effects through open DevOps.
- ▶ **SAN Automation and Ansible:** Gain cloud-like SAN orchestration for optimizing administration resources through Brocade REST API automation technology.

3.1.3 Self-healing

When potential disruptions are detected, the network automatically mitigates or resolves issues without manual intervention. This process is done by proactively monitoring the network to automatically identify abnormal or unexpected infrastructure behavior and then taking immediate action. These actions include alerting the end devices of the issue through a notification and signaling process within a SAN.

End devices can automatically adjust traffic or fail over to a healthy path to mitigate impact until a further infrastructure change solution can be implemented. In addition, these capabilities detect and automatically reconfigure fabric mis-configurations that fall outside best practices. Self-healing includes the following features:

- ▶ Fabric congestion notification: Automatically detects congestion and notifies end devices to automatically mitigate congestion.
- ▶ Slow drain device quarantine (SDDQ): Automatically quarantines credit-stalled devices to prevent the misbehaving device from affecting the rest of traffic.
- ▶ Automatic actions: Ensure data delivery with automatic failover from physical or congestion issues, such as port decommissioning, port toggling, and port fencing.
- ▶ COMPASS: Detects and automatically reconfigures out-of-compliance configurations.
- ▶ SANnav Management Portal: Reduces troubleshooting steps with built-in best practice recommendations to quickly resolve issues.

For more information about how self-healing SAN capabilities minimize the application performance impact, see [this web page](#).

3.2 Automation

Most, if not all, IT administrators have first-hand experience in the growing complexity of managing enterprise infrastructure, including SANs.

IBM and Brocade are in the unique position to spot and understand the effect of automation, which helps organizations get more from their SAN infrastructure.

IBM and Brocade offers a combination of SAN automation with RESTful APIs and a SAN management platform to help organizations drive greater efficiency from their SANs. This is accomplished by using various means, including the following examples:

- ▶ IBM and Brocade SAN automation is provided with a multilayer architecture.
- ▶ RESTful API support on switches and management tools.
- ▶ Brocade's Ansible management framework, which eliminates repetitive tasks, simplifies management, and orchestrates across the full SAN infrastructure.

3.2.1 Why organizations should embrace automation

Organizations should embrace SAN automation for the following reasons:

- ▶ Reducing human error and streamlining operational processes has never been more crucial. As organization's move to digitize and adapt to new workloads, data availability, processing time, and agility to provision applications on-demand become the life-blood of the business. This demands a more efficient and expedient infrastructure management approach, leaving no room for human error. As a result, storage administrators must be freed from repetitive manual tasks, such as configuration management, reporting, documenting inventory, and troubleshooting. Instead, IT organizations need SAN automation to help them automate and orchestrate repetitive tasks, significantly improve efficiency, and decrease the risk of operational mistakes.

- ▶ Demand for more accurate and frequent infrastructure reports is on the rise. It is not only IT managers who need information about storage performance, use, and forecasting; business stakeholders also are asking for and expecting this data on-demand.

No one wants to wait for a slot when storage administrators can allocate time to produce a report. This information should be available as frequently as business demands dictate, all at the click of a button. Automation not only provides this kind of responsiveness that traditional manual storage management processes cannot deliver, but it also can be customized so that each stakeholder is receiving more accurate data that is aligned to their responsibility.

- ▶ SAN configuration management must be streamlined. With more enterprise applications demanding access to more data and with more virtual machines, deploying and configuring servers, storage, and the network has become more time-consuming and more complex than ever.

By streamlining SAN operations through automation, application provisioning workflows are simplified across hypervisor, network, and storage. As a result, they deliver agility and responsiveness to meet dynamic business demands.

- ▶ IT service delivery is not always as responsive as the business demands. As organizations become increasingly reliant on having world-class IT services available to them for proactive, agile business decision-making, bottlenecks to IT service delivery must be identified and eliminated immediately. This must be done without hiring more storage administrators or boosting SAN-related Capex spending, which makes SAN automation the only viable option to drive increased IT agility. It also more tightly aligns IT service delivery with fast-changing business needs.
- ▶ Consistent configuration validation is a must. Manual configuration changes are occurring with greater frequency as enterprises diversify their IT architectures in general, and their storage architectures specifically. SAN automation ensures validation of consistent configuration parameters across the different SAN fabrics to facilitate troubleshooting of frequent alerts without reliance on manual intervention.

Brocade automation solutions use RESTful APIs to facilitate solutions architecture, share best practices, and get to production status faster. Brocade's SAN automation brings automation to the switch level that helps reduce the cost and complexity of managing storage systems.

For more information about how to set up and manage IBM b-type switches by using REST API, see the [Brocade Fabric OS REST API Reference Manual](#).

For more information about choosing your approach by using RESTful, Ansible, or Python, and to create an automation plan, see [SAN Automation for Dummies](#). This document also guides you through the process of how to clone and use the GitHub repositories.

3.3 FOS 9.x features

FOS 9.0 introduced new features and many improvements to the IBM b-type Gen 6 and Gen 7 SAN switches, out of which Unified Addressing Mode, Fabric Zone Lock and Mandatory FICON Logical Switch are some of the important changes.

In addition to these features, new features were added, including the following example:

- ▶ Added Java Less WebTools that are based on HTML 5.0
- ▶ Increased zone database size to 4 MB
- ▶ Updated SupportSave to run parallel to save time during troubleshooting
- ▶ Added the **reset** command to factory reset and support for the Generic USB drives

In this section, we describe some of these features. For a full list of enhancements, see the [Brocade Fabric OS Administration guide](#).

3.3.1 Port identifiers

Port identifiers (PIDs, also called *fabric addresses*) are used by the routing and zoning services in FC fabrics to identify ports in the network. All devices in a fabric must use the same PID format. When you add equipment to the SAN, you might need to change the PID format on older equipment.

Many scenarios cause a device to receive a new PID; for example, unplugging the device from one port and plugging it into a different port as part of fabric maintenance; or changing the domain ID of a switch, which might be necessary when merging fabrics; or changing compatibility mode settings.

Some device drivers use the PID to map logical disk drives to physical FC counterparts. Most drivers can change PID mappings dynamically (also called *dynamic PID binding*) or use the WWN of the FC disk for mapping (also called *WWN binding*).

Some older device drivers behave as though a PID uniquely identifies a device; they use *static PID binding*. These device drivers must be updated, if possible, to use WWN binding or dynamic PID binding instead because static PID binding creates problems in many routine maintenance scenarios.

Fortunately, few device drivers still behave this way. Many current device drivers enable you to select static PID binding and WWN binding. Select static PID binding only if there is a compelling reason and only after you evaluated the effect of doing so.

3.3.2 Unified addressing mode

Unified addressing mode (UAM) provides a single addressing mode for an 8-bit, 10-bit, or shared area. It also provides other addressing modes for different combinations of platform and logical switch type. UAM simplifies these addressing mode combinations and features into a single unified addressing mode.

UAM is introduced in FOS 9.0.0 and is the default address allocation mode for next generation platforms, such as the IBM SAN512B-7 Director, IBM SAN256B-7 Director, and IBM SAN64B-7 Switch.

Table 3-1 describes how UAM works for FOS features.

Table 3-1 Difference between UAM and pre-UAM Mode

Feature	UAM mode	Pre-UAM mode
F-Port	A port is assigned only a base area.	A port is assigned only a base area.
F_Port with NPIV	For F_Ports with NPIV fewer than 63 devices, port is assigned only a base area. If the 64th NPIV device is received in F_Port, a new area is assigned to the existing base area. Whenever a port exhausts an area after 64 extra devices log in, another area is assigned. The maximum number of NPIVs supported is 255 devices.	The 10-bit addressing mode uses the 8-bit area ID and the borrowed upper 2 bits from the AL_PA portion of the PID. Areas 0x00 - 0x8F use only 8 bits for the port address and support up to 255 NPIV devices. Areas 0x90 - 0xFF use an extra 2 bits from the AL_PA for the port address. Therefore, these ports support only 64 NPIV devices per port.
F_Port Trunking	Same as F_Port and F_Port with NPIV. In FOS 9.0, UAM allocates more areas to F_Port trunking and Emulex HBA trunking.	F_Port trunking uses an 8-bit area in VF mode enabled logical switch. Default switch in VF or non-VF mode uses existing port address.
EX_Port	EX_Port in a base switch is assigned a 10-bit area.	EX_Port in a base switch uses a port address per the current switch addressing mode.
EX_Port on ICL Ports	When an ICL port is configured as an EX_Port, a new area is assigned to the ICL port from the pool. This is not applicable for ICL ports that are not configured as an EX_Port. From the FOS 9.0 release, when an ICL port is configured as an EX_Port, the new areas are assigned in both VFmode and non-VF mode.	When an ICL is configured as an EX_Port, the switch is assigned an 8-bit area in VF. EX_Port over ICL is not supported in non-VF mode.
EX_Port Trunking	In UAM mode, an EX_Port trunk behaves in a masterless manner both in the base switch and in non-VF mode.	In pre-UAM mode, an EX_Port trunk behaves in a masterless manner only in VF mode.
FCoE	In UAM mode, FCoE address assignment is simple. Whenever Ethernet Flexport requires an area, it is assigned an area from the pool of unused areas.	In pre-UAM mode, whenever Ethernet Flexport requires an area, it is assigned an area from the pool of unused areas. For default switches in VF mode or non-VF mode, the switches are converted to a 10-bit area mode when FCoE enodes are configured.
Port Address Binding	In platforms that are enabled in UAM mode, the port address binding feature is supported in VF and non-VF mode.	In pre-UAM mode, the port address binding feature is supported in only logical switches in VF enabled mode.
Persistent PID	In platforms that are enabled in UAM mode, the persistent PID feature is supported in VF and non-VF mode.	In pre-UAM mode, the persistent PID feature is supported in only logical switches in VF enabled mode.

For more information about the benefits and restrictions with the Unified Addressing Mode, see Chapter 6, “FICON design, implementation, and migration” on page 105.

3.3.3 Fabric zone locking

Before FOS 9.0.0, the level of zoning transaction protection against multiple concurrent clients existed at the local switch level only. For example, if a user on a switch starts a zone transaction by performing a zone create or edit operation, and if another user attempts to perform a zone create or edit or commit operation on the same switch, the second user is locked out and prevented from performing their operation (see Figure 3-1). This lock out occurs because the first user opened the transaction first.

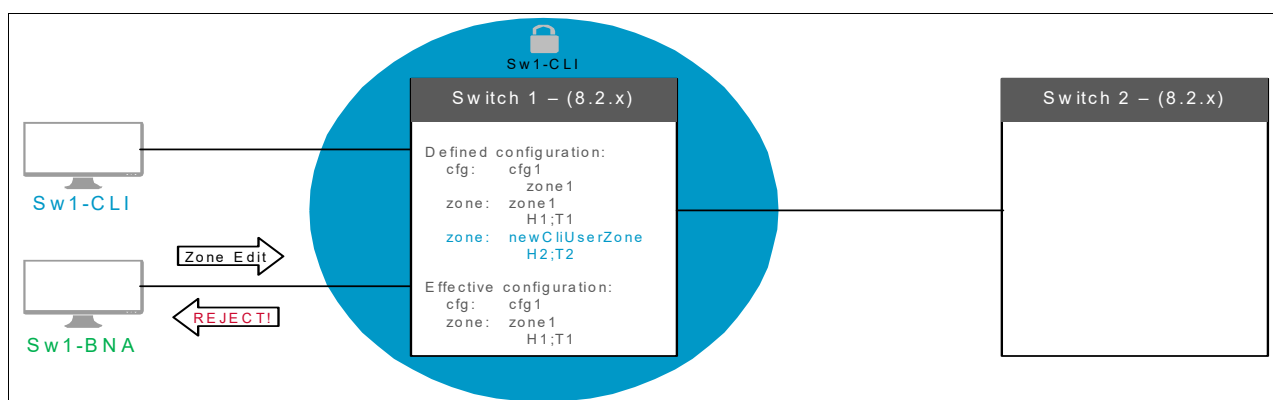


Figure 3-1 Single switch locking example

Although local switch protection exists, no fabric-wide locking mechanism is available. For versions earlier than FOS 9.0.0, the fabric-level detection support (Concurrent Transaction Detection) consists of only a warning mechanism that warns users if a transaction is open on a remote switch. However, it does not prevent users from committing changes if concurrent transactions exist across the fabric.

Furthermore, this fabric-wide detection is limited in scope to CLI zoning operations only. For example, if a user on one switch opens a zone transaction, and if a user on another switch in the fabric issues a zone edit or commit command, a warning is received that a transaction is occurring in the fabric. However, a zone commit operation can still be issued from either switch, which overwrites any pending zone transaction on other switches (see Figure 3-2).

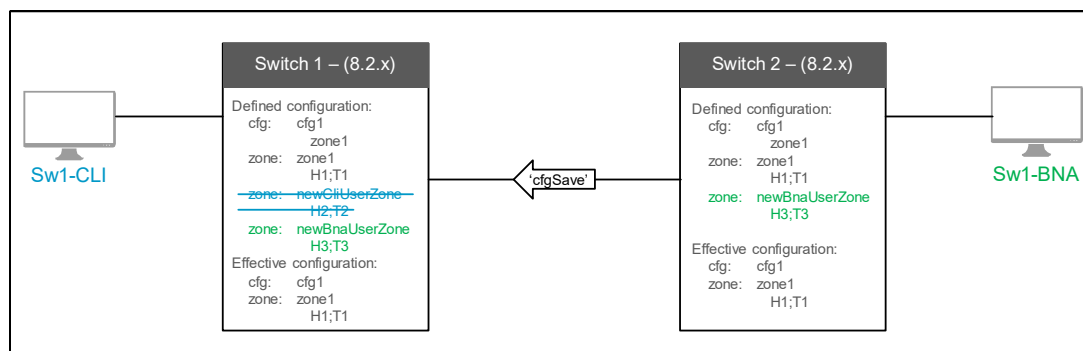


Figure 3-2 Fabric OS 8.2.x and earlier fabric-wide commit overwrite example

From FOS 9.0.0 version onward, the Zone Fabric Locking feature provides zone transaction protection for a single switch and also fabric wide (see Figure 3-3). This feature is enabled by default. If a zone edit or commit command is occurring in fabric, you cannot perform a zone edit or commit on the same or another switch for a default timeout period of 5 minutes. A lock request is sent at the beginning of a zone edit operation. The Fabric Lock Failsafe Timer is configurable and it is a fabric-wide setting. When a zone fabric lock is active, a failsafe timer is started on all remote switches.

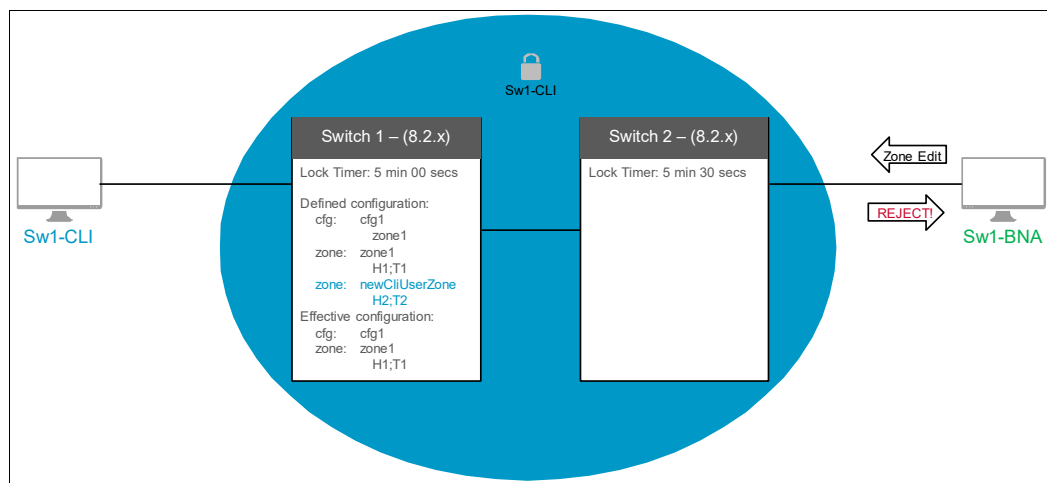


Figure 3-3 Fabric OS 9.0.0 zone fabric locking example

When the failsafe timer expires, the open zone transaction is not stopped. If the same user attempts to resume the transaction by performing another edit or commit operation after the zone fabric lock expires, the transaction is allowed and the fabric lock is restarted.

If a different user attempts to start a new transaction after the first user's transaction timer expires, the transaction is allowed and the first user's transaction is stopped before the second user's transaction starts.

Note: Consider the following points:

- ▶ If a fabric-lock is active in the fabric, the Fabric Lock Failsafe Timer timeout value cannot be modified.
- ▶ The fabric zone lock is a fabric-wide setting. Therefore, the fabric must be stable before performing zone edit, commit, and abort operations.
- ▶ If a zone transaction is stopped, the fabric zone lock is cleared.
- ▶ When the timer expires, the lock is cleared locally and on each individual switch in the fabric.
- ▶ When the virtual fabric function is used, the lock timeout value is independent of each logical switch and can have different values.
- ▶ Fabric zone locks can be preempted by zone merges.
- ▶ The Zone Fabric Locking feature is supported on the switches with the FOS version 9.0.0 or later. Down-level switches do not participate. However, FOS 9.0.0 switches that own the zone fabric lock prevent updates from down-level switches.
- ▶ On remote switches, another 30-second buffer is added to the configured lock timeout value of the lock principal switch.

The fabric-wide locking mechanism ensures that any zone transaction that is started on any switch by any interface (CLI, REST, WebTools, SANnav, or Target Driven Zone) is not overwritten or stopped by another interface's transaction (except for zone merges) while the lock timer is active.

3.3.4 FICON logical switch

With the FICON logical switch feature, you can create a FICON logical switch. You can also modify a logical switch as a FICON logical switch. The FICON logical switch is supported on all platforms that support Virtual Fabrics in FOS 8.1.0 or later. A FICON logical switch can be created independent of enabling FMS mode. A FICON logical switch cannot be the default switch in the chassis, nor can it be the base switch of the chassis.

FICON logical switch provides auto-configuration of the following components, which simplifies the process of establishing a logical switch in the FICON environment:

- ▶ Insistent Domain ID (IDID) mode is set to on.
- ▶ Routing policy is set to device-based routing.
- ▶ Area mode is set to zero-based addressing.
- ▶ SCC security policy is auto-defined with only its own WWN.
- ▶ The created SCC policy is activated.
- ▶ The fabric-wide consistency policy is set to strict SCC policy.
- ▶ High Integrity Fabric (HIF) mode is enabled.
- ▶ FMS mode is enabled if FICON control unit port (CUP) license is installed on the chassis.

For more information about creating FICON Logical switch and best practices, see Chapter 6, "FICON design, implementation, and migration" on page 105.

3.4 Management tools

The available management tools are described in this section.

3.4.1 WebTools

Brocade WebTools is a graphical user interface (GUI) that is embedded in the FOS firmware. It enables administrators to monitor and manage single or small fabrics, switches, and ports. New WebTools are Java-Free and are designed based on modern technology HTML-5. WebTools is started directly from a web browser or from SANnav Management Portal. The system requirements are described next.

WebTools system requirements

Before starting WebTools, verify that your workstation uses a supported operating system and web browser.

WebTools does not require a license. It is installed on the switch when you install FOS.

Supported operating systems

WebTools supports the following operating systems:

- ▶ Red Hat 8.0 and 8.1
- ▶ Windows 10 Pro
- ▶ Windows 2019

Supported web browsers

The following browsers can be used to access WebTools:

- ▶ Chrome
- ▶ Firefox

Starting WebTools

WebTools is started directly from a web browser or from SANnav Management Portal. You can start WebTools on any workstation with a compatible operating system and Web browser installed.

If the switch is configured with logical fabrics, you can log in to any of the logical fabrics for which you have permission.

To start directly from a web browser, open your browser, enter the IP address of the switch, and press **Enter**. You can use HTTP or HTTPS; for example, `http://10.77.77.77` or `https://10.77.77.77` (see Figure 3-4).

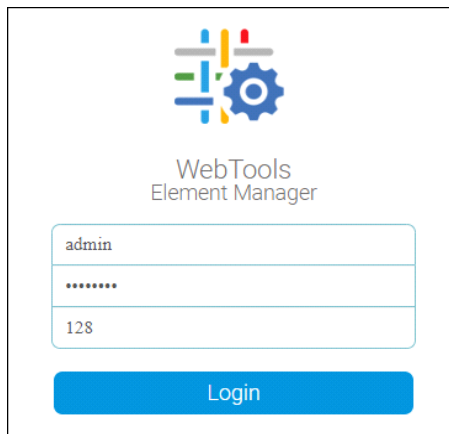
The image shows a web browser window displaying the 'WebTools Element Manager' login page. At the top center is a logo consisting of a blue gear with a white 'X' inside, surrounded by four colored vertical bars (blue, green, yellow, red). Below the logo, the text 'WebTools Element Manager' is displayed. Underneath, there are three input fields: the first contains the text 'admin', the second contains a series of dots representing a password, and the third contains the text '128'. Below these fields is a blue rectangular button with the word 'Login' in white text.

Figure 3-4 Starting WebTools

Enter the user name, password, and logical switch name or fabric ID (FID).

For the first switch login, the default username is `admin` and the default password is `password`. WebTools prompts you to change the default password.

If you are logging in to a Virtual Fabrics-enabled platform and you do not specify a logical switch, you are logged in to the default logical switch, which uses fabric ID 128. For non-VF platforms, the FID option is not displayed.

Note: If you start from SANnav Management Portal, you might not need to log in, depending on the SANnav single sign-on configuration.

For information about how to use WebTools, see [this web page](#).

3.5 Security

Security entails authentication and encryption to restrict access and protect data from unauthorized access. Brocade products support a wide range of authentication, encryption, and management tools to protect fabrics and data from unauthorized access, including the following examples:

- ▶ Authentication: Authentication protocol support includes CHAP, DH-CHAP, FCAP, IKE, IPsec, RADIUS, TACACS+, and P-EAP/MS-CHAP for RADIUS.
- ▶ Encryption (AES/3-DES): Brocade provides AES-128 and AES-256 encryption and 168-bit 3-DES encryption for IP links on extension products and management connections. Brocade also supports AES and 3-DES with IPsec. These solutions provide high-performance encryption and compression.
- ▶ In-flight encryption over ISLs: Brocade X6 with Gen 6 FC port blades, Brocade X7 with Gen 6 or Gen 7 FC port blades, and Brocade G720, G630, G620 switches support in-flight encryption for traffic over ISLs to minimize the risk of unauthorized access to data within the data center and over long-distance links. Data-at-rest and data-in-flight encryption are complementary technologies that serve different purposes, and each might be required to achieve regulatory compliance.
- ▶ Secure Boot: A switch validates the integrity and authenticity of the FOS boot image to establish a hardware-based root of trust through the manufacturing supply chain.

For more information, see Chapter 12, “Storage area network security” on page 311.

3.6 Traffic Optimizer

Traffic Optimizer is a new feature that was introduced in FOS 9.0 on Brocade Gen 7 platforms that optimizes SAN traffic. Traffic Optimizer automatically groups traffic flows with similar predefined attributes, such as flow destination speed and priority, as a Performance Group.

Each Performance Group uses a dedicated resource path within the fabric. By organizing flows by Performance Groups and preventing mixed traffic, Traffic Optimizer prevents slow traffic from obstructing the flow of higher-speed or higher-priority traffic on that path.

With Gen 7 platforms, Traffic Optimizer provides a way to optimize shared resource allocation for each Performance Group. It also automatically maps traffic flows to a predefined Performance Group.

FOS 9.0 supports the Performance Groups that are listed in Table 3-2 on page 37. These predefined performance groups provide the following benefits:

- ▶ Support of speed-based classification
- ▶ Credit-stalled device isolation by way of SDDQ
- ▶ Compatibility with:
 - Adaptive Networking (QoS Zone and CSCTL) feature
 - Previous platforms

Traffic Optimizer is enabled by default on all Gen 7 platforms and Gen 7 blades.

Table 3-2 Supported Performance Groups in Fabric OS 9.0

Flow classification	Performance Group name	Speed range (Gbps)	Assigned priority
Speed Based	PG_64G	32 < speed <= 64	Medium
	PG_32G	16 < speed <= 32	Medium
	PG_16G	0 < speed <= 16	Medium
SDDQ	PG_low_qos_1	0 < speed <= 64	Low
	PG_low_qos_2	0 < speed <= 64	Low
QoS Zones	PG_high_qos_1	0 < speed <= 64	High
	PG_high_qos_2	0 < speed <= 64	High
	PG_high_qos_3	0 < speed <= 64	High
	PG_high_qos_4	0 < speed <= 64	High
	PG_high_qos_5	0 < speed <= 64	High

Gen 6 platforms in FOS 9.0 continue the established behavior and maintain compatibility with Gen 7 platforms in the same fabric. If a Gen 7 switch or blade is connected to a port on a Gen 6 platform and the traffic flow source port is on the Gen 7 switch, the traffic flow is speed-based that is classified within the Gen 7 switch until reaching the port on the Gen 6 platform. Then, the traffic flow is handled with established behavior starting from the Gen 6 port.

Many years ago, Brocade Virtual Channels (VC) addressed this issue with great success and significant performance gains. Resources that affected performance include bandwidth, egress queues, ingress buffers, and QoS.

The implementation features many distinct egress queues that established virtual channels across the physical link, each channel having its own set of autonomous buffer credits. The bandwidth is shared, but flows are independent. When traffic characteristics are uniform, VCs offer an expedited pathway. Great benefit is gained by sorting traffic into VCs that are based the most pertinent characteristic, which is speed.

Normally, VCs service multiple FC flows because a limited number of VCs are available per classification of traffic. Users do not configure or setup VCs; instead, they are a fundamental component of the Fabric Operating System (FOS). FC wire speeds continue to increase, now at 64 Gbps (Gen 7) and bandwidth features a wide spread from 8, 16, 32, to 64 gigabits (32 - 56 Gbps). Wide variations in speed can impose flow-control problems within a network.

For years, Brocade VCs operated within a relatively narrow bit rate of 1, 2, 4, 8, and 16 gigabit (8 - 15 Gbps) by use many logical independent paths. VCs are assigned by a hash called PID2VC, which results in a pseudorandom assignment of each flow to a VC. To a degree, this statistically mitigated the interference of slower flows impeding faster flows. Also, optionally faster flows can be manually assigned to a high QoS VC and slower flows to a low QoS VC to prevent interference.

Technology evolves and Brocade optimized VC efficiency by enhancing effectiveness through targeted flow characteristics. Demands on an enterprise SAN have never been greater because of other complimentary technology evolutions, such as NVMe, AFA, and host virtualization. Flows now compete and head of line blocking is not an option. Blocking must be efficaciously dealt with, which is where Traffic Optimization comes in.

Brocade Traffic Optimization technology takes VCs to the next level. Traffic Optimization organizes and manages traffic flows by using performance groups (PG). Fabric resources are allocated based on performance groups. Instead, flows are assigned to a VC based on the destination port speed and potentially by way of protocol (for example, NVMe). Brocade fabrics know the destination port speed and protocol for every flow.

Brocade Gen 7 hardware added VCs, E_Ports, and EX_Ports. They use 16 new Traffic Optimization VCs, which are assigned: Four VCs per speed (<16, 32, 64) for a total of 12 VCs, plus another four reserved VCs for future use. QoS VCs did not change, there are 2 for low, 4 for medium (default), and 5 for high. QoS VCs are used when traffic is designated to a high or low priority. SDDQ uses the QoS low VCs.

Pre-Gen 7 and existing fabric coexistence is fully supported. E_Port connections between Gen 7 and pre-Gen 7 platforms results in Traffic Optimized VC being remapped by using the established classification. QoS VCs pass-through unchanged because they remain consistent with Traffic Optimization.

Traffic Optimization is new and unique function that is delivered in Brocade's latest Gen 7 FC platforms. It is a powerful enhancement to the existing VCs feature and is incumbent to past generations of Brocade switching equipment. Traffic Optimization requires no configuration and works on all Gen 7 devices. Traffic is classified by speed or NVMe Performance Groups and assigned to a specific VC. Traffic Optimization is interoperable with FC cohort technologies (FICON, Extension, Access-Gateway, Long Distance Links, Security, and NPV) and Brocade Gen 6 fabrics.

3.6.1 Speed-based classification

Flows are classified into one of several Performance Groups based on the destination speed or protocol. Automatic speed classification isolates traffic flows into different speed-based Performance Groups and eliminates speed mismatches that cause performance interference when lower speed devices collide with higher speed devices that are connected in the same fabric.

3.6.2 Slow drain device quarantine isolation

A slow-drain device, also sometimes referred to as a *credit-stalled device*, is a misbehaving device that promptly stops returning R_RDY signals (buffer credits) to the switch, which causes the switch to stop sending frames to the device. A credit-stalled device that stops returning credits to the switch for hundreds of milliseconds or more causes frame drops and link resets. This issue typically occurs when the end device is not working per its specification. An impaired flow that is associated with credit-stalled devices is automatically mapped to a low-priority Performance Group: PG_low_qos_1 or PG_low_qos_2.

3.6.3 Compatibility with QoS zones and CS_CTL

Gen 7 platforms maintain compatibility if the QoS feature is enabled. If QoS Low or QoS High priority are defined, traffic flows are assigned to one of the QoS Performance Groups that are defined as listed in Table 3-2 on page 37.

If CS_CTL is enabled on a port, traffic flows on that port retain the established CS_CTL-based classification and are not classified into any Performance Group.

3.6.4 Compatibility with previous platforms

Previous platforms that participated in FOS 9.0 fabrics continue the established behavior and maintain compatibility with Gen 7 platforms in the same fabric. If a Gen 7 switch or blade is connected to a port on a Gen 6 platform and the traffic flow source port is on the Gen 7 switch, the traffic flow is speed-based classified within the Gen 7 switch until reaching the port on the Gen 6 platform. Then, the traffic flow is handled with established behavior, starting from the Gen 6 port.

3.6.5 Traffic Optimizer and QoS Mode

The Traffic Optimizer and QoS mode are automatically enabled by default on all Gen 7 platforms. Traffic Optimizer's advance classification (for example, map traffic flow to predefined Performance Group) requires QoS mode to be enabled on all the ports of the supported platforms.

It is *not* recommended to disable QoS mode on any ports. If QoS mode is disabled on a port, Traffic Optimizer's advanced classification do not work on the port, and the port's traffic classification falls back to older PID to VC mapping.

FCR considerations for Traffic Optimizer

Traffic Optimizer is supported on only an FCR edge-to-edge fabric topology. It is *not* supported on a backbone-to-edge fabric topology. For Traffic Optimizer to run over FCR, all switches and FC routers in edge fabrics and backbone fabrics must be running FOS 9.0 or later.

Traffic Optimization over FC routers is supported only if the Virtual Fabrics is disabled in the Backbone fabric. Traffic Optimization over FC routers is not supported if the Virtual Fabrics is also enabled in the Backbone fabric.

For more information about the operating modes of Backbone fabrics, see [Brocade Fabric OS Administration Guide, 9.0.x](#).

If the traffic flow in the metaSAN encounters a switch platform that does not support Traffic Optimization, the traffic flow is mapped to the older behavior and continues to retain the established behavior all the way.

3.7 Fabric Notifications

Fabric Notifications is a facility for proactively notifying devices in a fabric of performance impacting behaviors that affect the normal flow of I/O. Fabric impacting behaviors (FPI), such as signal integrity degradation, frame loss, or congestion, are detected by the fabric or participating devices. Signaling by way of FC Primitives, and notifications by way of Extended Link Services (ELS) are then used to communicate the type and severity of the FPI to the affected members of the fabric.

Fabric Notifications include the following components:

- ▶ Registered capabilities

A set of ELS commands that are exchanged between two ports arbitrate the range of Fabric Notification capabilities that are supported by each.

- ▶ Congestion detection signaling

Hardware-generated primitive signals, supported by Gen 7 platforms only that are sent toward the point of congestion indicate that the queue of frames to be transmitted reached a level where fabric performance might be affected.

- ▶ FPI notifications (FPINs)

FPINs are software that is generated by using ELS commands. If a device registered to use them, administrators can be notified of events that might affect fabric performance.

These components allow members of the fabric to indicate their ability to interpret and react to signals and notifications. Participating devices can respond by taking actions to mitigate and reduce the effects of congestion, frame loss, and signal integrity degradation. The FOS is the primary generator of signals and notifications because of its awareness of I/O flows, physical layer impacts, and device responses.

The relationship between the event detection, signaling, and notification operations is shown in Figure 3-5.

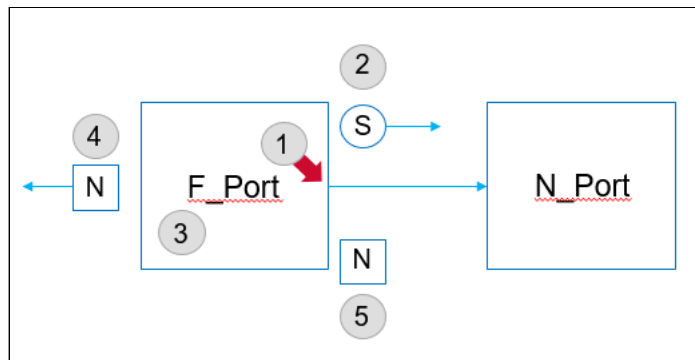


Figure 3-5 Signaling or notification operations

When an event is detected (1), the fabric starts the suitable response. If frame transmission is stalled, signal events are generated (2) toward the point of congestion (that is credit-stalled or oversubscribed device). If a transmission failure or sustained congestion event occurs, notification events are generated (3) and distributed to peer devices (4).

The fabric determines the distribution of the notification event from the zone member information of the triggering N_Port. If a transmission failure or software-based congestion signaling occurs, notification events are sent to the attached N_Port (5) to engage the device in mitigating activities.

3.8 IBM b-type switch Fabric Notifications solution

Fabric Notifications was developed to optimize I/O behavior and avoid impaired paths by notifying devices of current fabric conditions. The IBM b-type Fabric Notifications solution combines hardware (Brocade FC switching ASIC), software (Fabric Operating System), and Management (MAPS and SANnav) to form a SAN-wide solution for impairment detection, notification, and remediation. The problem is logged and operational information concerning affected flows (ITL), location, timestamp, and reason are provided for easy debugging and troubleshooting.

3.8.1 IBM b-type Fabric Performance Impact Notification

IBM b-type Fabric Performance Impact Notification (FPIN) is a new feature for Brocade FOS v9.0. It is available on Brocade Gen 6 and Gen 7 switches.

This feature enables the switch to detect issues on a fabric such as congestion or physical link issues and then notify the affected devices that registered for these notifications. The devices that receive these notifications can then proactively take steps, such as path failover, rather than react to a path being down.

FPIN provides a means to notify devices of link and other issues with a connection or possibly a path through the fabric. For FPIN, a device must register with fabric services to receive these notifications. The new IBM b-type Gen 7 hardware can send hardware and software signal notifications. Gen 6 can send software notifications only. The hardware and software notifications require FOS v9.0 on the Switches and Directors.

Hardware signals can be sent from the switch to the adapter in the device. The adapter can then decide what to do about the notification. Software signals are sent higher up in the Fibre Channel stack, and the adapter driver then decides how to handle the notifications.

One advantage to notifications in hardware is reaction time; the adapter can process the notifications and react more quickly than can the driver. Another advantage is that the hardware-based notification is an FC primitive, which means that even if buffer credits are depleted, the signal can still be delivered to the device on the other end of the link.

Primitives are not frames; therefore, they do not need buffer credits to be sent. The software layer signal is an ELS frame, so they can be affected by buffer credit depletion and other link congestion.

Whether the signal that is sent is hardware or software, how the devices handle the notifications is up to the vendor of the adapter. Some might log the notification, while others might take action. The action that an adapter takes is also vendor-specific. FPIN can alert devices about end device congestion, device link integrity (CRC), and frame drops.

At the time of this writing, the following vendors support FPIN:

- ▶ Linux Multi-Path in RHEL 8.2
- ▶ IBM AIX® (registers for Link Integrity and Congestion notifications)
- ▶ Emulex (supports Congestion and Link Integrity notifications on Linux)

Support for more HBA and Storage Controller vendors for FPIN is expected.

FPIN can be sent from any device that it supports. Potentially, the storage, host, or switch can share this information. Having end-to-end capability to act on the data based on the FPIN notifications, the SAN self-heals, which results an end-to-end autonomous SAN infrastructure.

3.8.2 Congestion Signal primitives

Congestion Signal primitives are ASIC-based and supported on IBM b-type Gen 7 platforms only. Congestion Signal primitives are signals that are sent within link encoding to directly connected FC devices.

Congestion Signal primitives notify peer Gen 7 platforms of degraded performance that are caused by over-subscription or credit stall conditions. Fabric Notifications hardware intelligently detects sudden congestion and reacts instantly by signaling the attached physically connected port.

Before receiving a Congestion Signal primitive, Gen 7 ports must notify peer ports that it is capable. For example, IBM b-type Gen 6 platforms do not support Congestion Signal primitives, but are capable of FOS-level ELS-based FPIN Congestion notifications.

3.8.3 Extended link service notifications

Extended Link Service (ELS) is a communication protocol that is used between Brocade Fabric Operating System (FOS) platforms for exchanging capabilities, service registration, notifications, and other vital control traffic. The Fabric Notifications solution augments hardware Congestion Signal primitives with software FPIN notifications for Congestion, Peer Congestion, Link Integrity, and SCSI Command Delivery. Fabric Notifications are FOS-based and supported on Gen 6 and Gen 7 platforms running FOS 9.0 or later.

Hosts, switches, and storage can generate ELS notifications upon detecting an impacting event. A notification can be sent by any fabric member detecting a problem, be it on a switching platform (F_Ports and E_Ports) or an end-device (N_Ports); it is processed the same.

Fabric Notifications and MAPS work in tandem. MAPS rules for congestion, SCSI command discard, and link integrity trigger corresponding FPIN ELS notifications.

FPIN notifications are sent to all in-zone end-devices that registered to receive that specific notification type. Notifications are not dependent on zoning type, whether traditional or peer; the operation is the same with no advantage or disadvantage. Also, Fabric Notifications send FPIN notifications to all peer switches and the peer switches forward notifications to their registered in-zone end-devices. FPIN notifications are restricted to devices:

- ▶ Supporting FPIN notifications
- ▶ Registered to receive a particular notification type
- ▶ Experiencing the notification condition
- ▶ Are an in-zone peer device

3.8.4 Exchange Diagnostic Capabilities and Register Diagnostic Functions

Exchange Diagnostic Capabilities (EDC) and Register Diagnostic Functions (RDF) are ELS services that are associated with Fabric Notifications (see Figure 3-6 on page 43). EDC is used between devices to exchange diagnostic capabilities. RDF is used by devices to register for specific notifications. The fabric does not transmit any notifications (Primitive or ELS) until a device is registered to receive them. Devices cannot deal with every type of problem; therefore, they cannot register for every available notification.

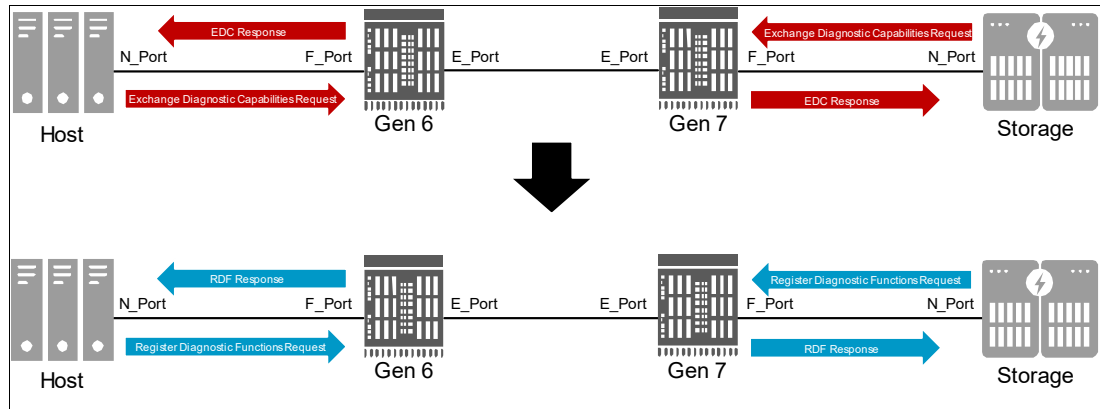


Figure 3-6 ELS: EDC and RDF

FPIN Notifications a device can register to receive the following information:

- ▶ Congestion notification descriptors are sent to devices on congested ports. Causes are credit loss, credit stall, and over-subscription. Hosts and targets are equal offenders.
- ▶ Peer Congestion notification descriptors are sent upstream from congested ports to in-zone registered devices. The FPIN notification informs the receiving device about the reason for congestion.
- ▶ Link Integrity notifications are sent to the attached device and other in-zone end-devices when a port experiences link integrity problems. CRC and ITW MAPS alerts are translated into a Link Integrity FPIN notification and sent to registered devices. Link Integrity notifications are used by end-devices to troubleshoot link problems. MPIO can prioritize a healthier path over an impaired path while continuing operation.
- ▶ Delivery (SCSI Command frame discard), is when a notification is sent to the registered originator of a dropped SCSI command or status frame. When a command frame is dropped, the originator is notified, and immediately the SCSI exchange is failed. Rather than waiting for an upper-level protocol timeout, which can take tens of seconds, the driver retries immediately.

3.8.5 Congestion Signal

Congestion Signals are an immediate feedback mechanism between the devices of a link that indicates that transmission resources are being used. Congestion Signal primitive signals are immediately sent upon detection, within microseconds (μ s). Congestion Signals are a link-by-link feature only.

Brocade Gen 7 FC ASICs support direct congestion signaling between platforms (see Figure 3-7). In turn, if the condition causes MAPS to trigger, the receiving switch uses FPIN Peer Congestion notification to other registered in-zone end-devices for potential I/O queuing optimization and remediation.

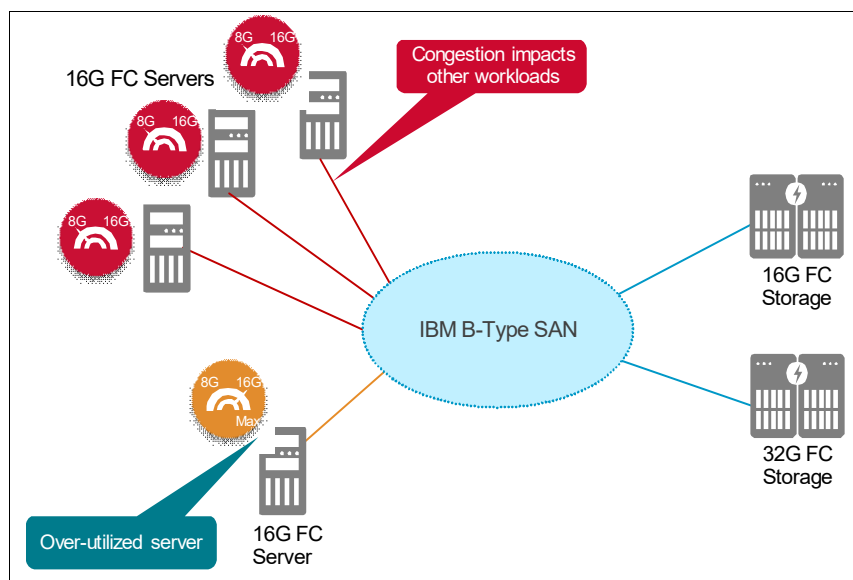


Figure 3-7 Congestion signal and notification for over-subscription

3.8.6 Congestion notification

FPIN Congestion notifications are valuable information for end-devices, which facilitates optimized I/O scheduling; that is, consider slower transfer rates or issue serial read I/Os.

The ELS Congestion Notification is the software equivalent of the Congestion Signal primitive and sent to end-devices that support notifications. An ELS congestion notification for over-subscription means that too much data was requested, and that the fabric was unable to accommodate the request without congestion. An ELS congestion notification for credit stall indicates to the end-device “you’re stuck - fix the situation”.

In general, congestion notifications indicate why exchange completion times might exist.

3.8.7 Congestion peer notification

FPIN Congestion Peer Notifications are sent to registered in-zone peers of end-devices who are experiencing congestion. Various remedies are available that peers can use to relieve congestion.

With speed matching, the peer’s port might need to auto-negotiate to a faster speed than the destination port; the peer can update its speed to match. Peers can choose alternative pathways to potentially circumvent a flow causing congestion.

Paths can be weighted based on peer congestion notifications to proportionally deliver traffic over multiple paths in a weighted fashion.

An example of how the Congestion Notification works in real world, in collaboration with Emulex HBA on the IBM AIX server, with the IBM b-type Gen 7 SAN, IBM SANnav Management software, and Emulex SAN Manager, a management tool to manage host HBAs in the environment is shown in Figure 3-8.

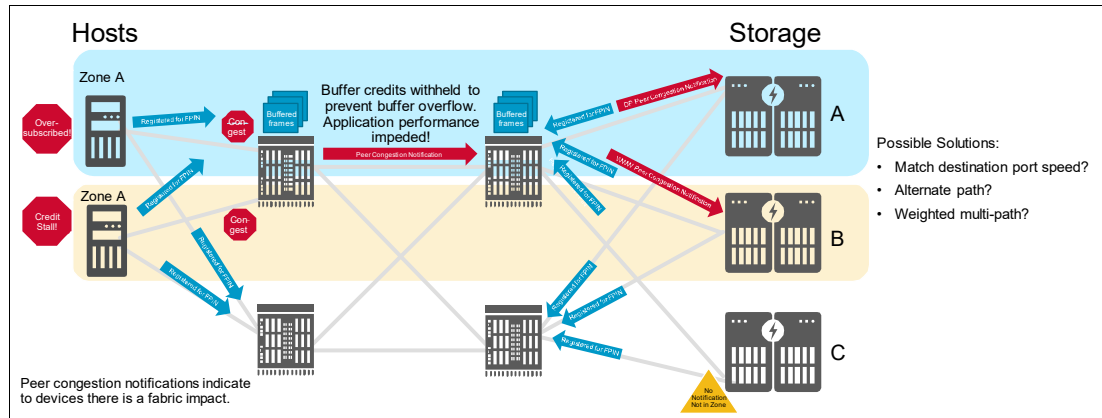


Figure 3-8 Peer Congestion Notification

Under normal operations, as shown in Figure 3-9, IBM SANnav management tool and the Emulex SAN Manager report minimal to no congestion.

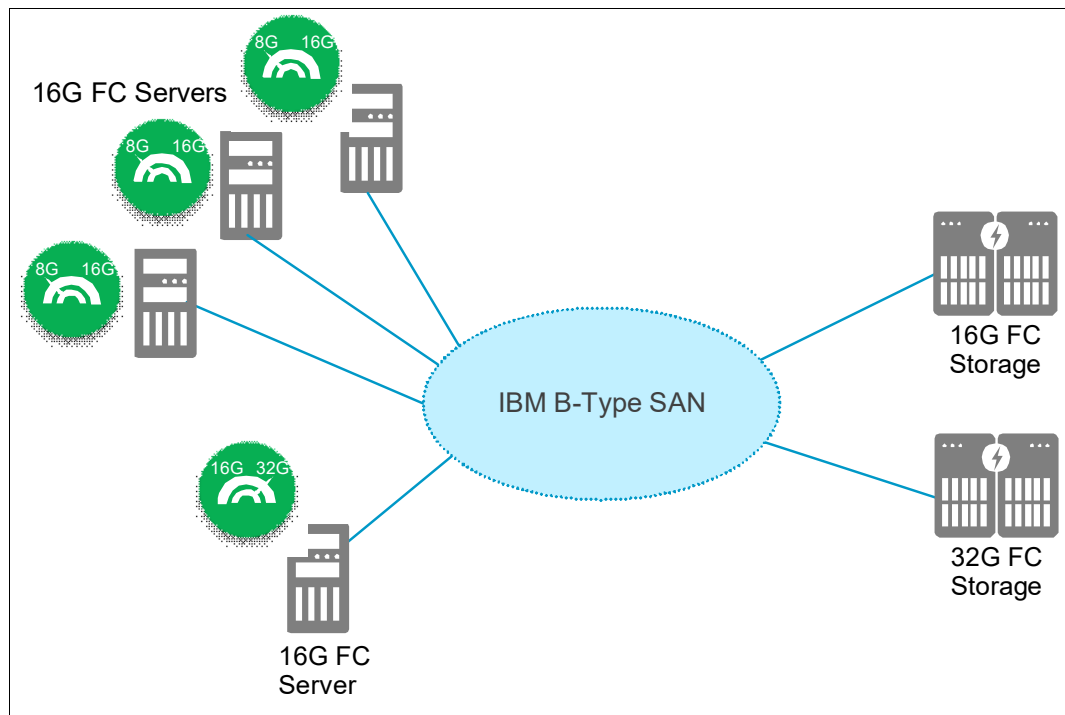


Figure 3-9 Normal operations

The workloads generally grow over time and becomes larger than the hardware can support. As shown in Figure 3-10, this issue has a ripple effect on the entire infrastructure, which causes the congestion and results in major performance issues.

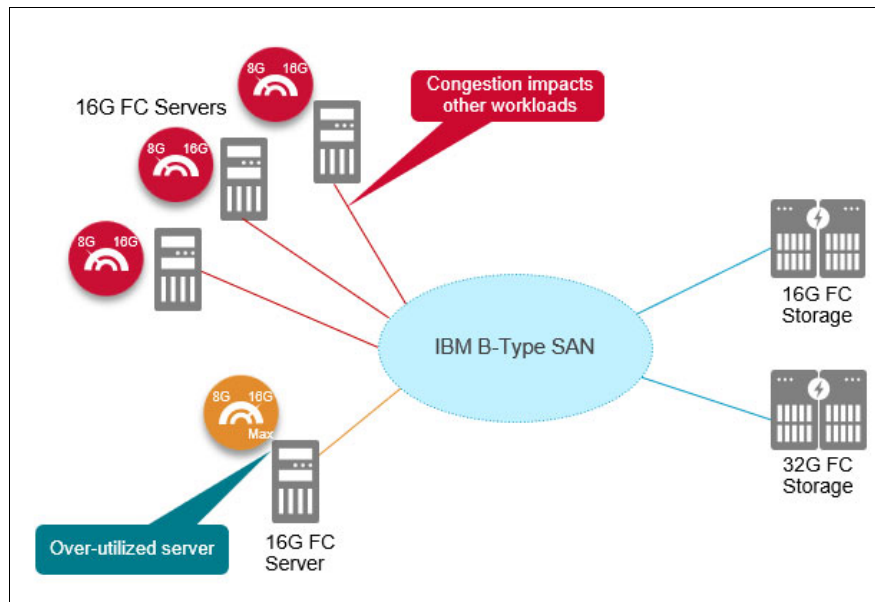


Figure 3-10 Over-subscription

IBM b-type Gen 7 SAN detects these types of F-Port congestion issues and sends FPIN Congestion Notification to the Emulex HBA on the AIX server. The Emulex SAN Manager (the management software to manage HBAs on the AIX server) is enabled to set congestion policy.

As shown in Figure 3-11 on page 47, depending on the choice of the policy between conservative, moderate, and aggressive, the action is taken to address the endpoint congestion issue. For the bully server, the workload is still too large; therefore, administrative attention is required.

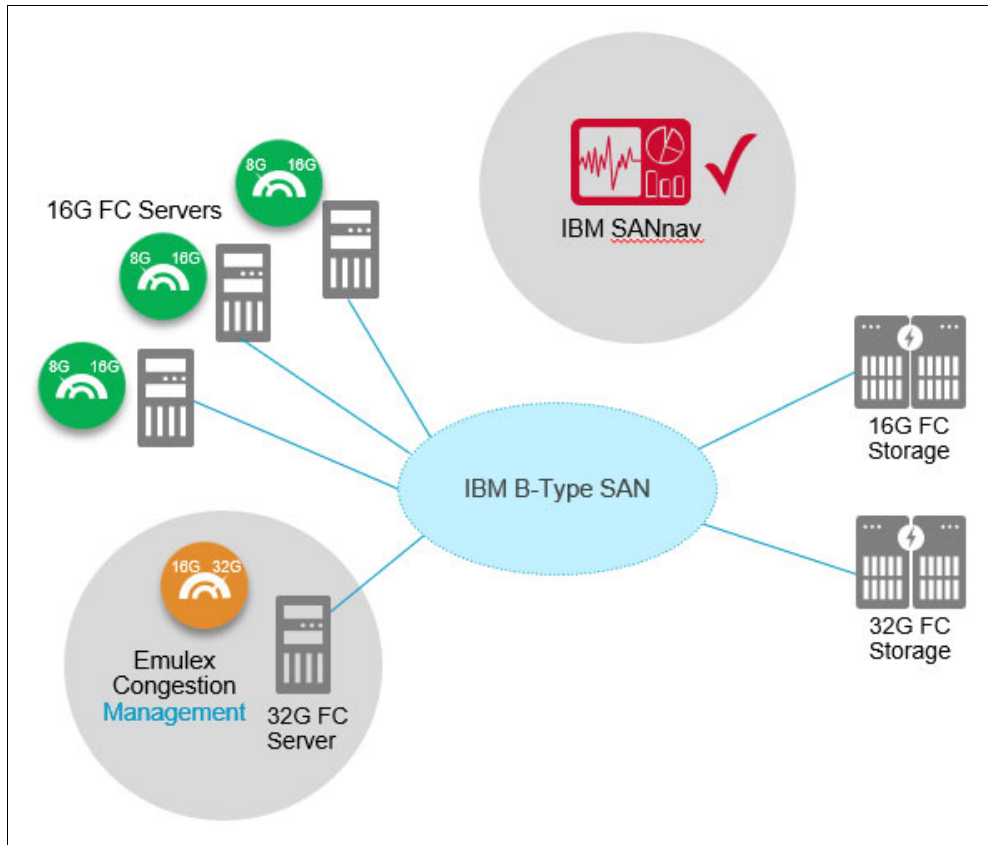


Figure 3-11 Addressing the congestion issue

3.8.8 Link Integrity

A complete solution requires end-devices (N_Port) to notify the fabric (F_Port) of receive-side errors of which the fabric otherwise is unaware. A link is a two-way street and FPIN Link Integrity notifications enable the identification of degraded paths in each direction, from N_Port → F_Port and F_Port → N_Port.

In notification F_Port → N_Port, HBAs register to receive FPIN notifications of interest; for example, Link Integrity. When F_Port experiences incoming errors because of faulty electronics, marginal optic, dirty connector, damaged fiber cable, and so on, a Link Integrity notification is sent to the HBA letting the host know of the transmission problem. In turn, the HBA forwards notifications to the MPIO driver, which can distinguish between a device failure and physical path failure. Link Integrity notifications are used by end-devices for troubleshooting links and by MPIO for selecting the best path.

In notification N_Port → F_Port, the HBA registers with the fabric to receive and send FPIN notifications. If the HBA sees incoming errors, it notifies the fabric that its connection to the end-device is erratic. An FPIN Link Integrity notification that is sent by N_Port to F_Port (specifically, the Fabric Controller) is distributed to the in-zone devices as well. The fabric logs the notification information so that the SAN administrator can take action.

Registered in-zone end-devices must be notified of link integrity problems. MAPS F_Port alerts for Cyclic Redundancy Check (CRC) and Invalid Transmission Word (ITW) are translated into Link Integrity FPINs and sent to the end-devices. Notifications integrate with the operating system device manager and notifying initiators and targets of link integrity issues along a data path enhances debugging and recovery efforts from multiple perspectives.

MPIO handles FPIN events and manages path selection. On a new pathway, MPIO drivers communicate with the appropriate service to manage outstanding and pending I/Os. End-devices determine their best course of remediation, such as immediate failover to an alternative path within the MPIO environment. Link Integrity problems include the following examples:

- ▶ Slow host application performance can be traced to CRC errors on a storage port
- ▶ Slow array performance can be traced to host port ITWs

How the Link Integrity works in real world, in collaboration with Emulex HBA on the AIX server and with the IBM b-type SAN, is shown in Figure 3-12.

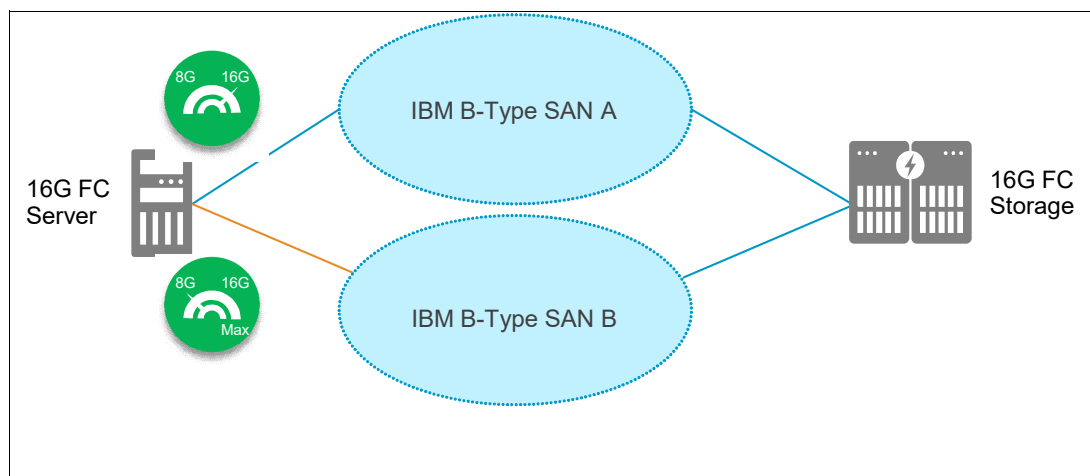


Figure 3-12 FC normal operations

Under normal operations, as shown in Figure 3-12, with the FC links working, data is balanced between the different paths to the storage system. Over time, links might not work as well as they should.

When links are “sick” but not dead, I/Os that go over the bad link reduce the data that is available to the application (see Figure 3-13). Quick detection and action are essential to improve the performance and in some cases availability of the application.

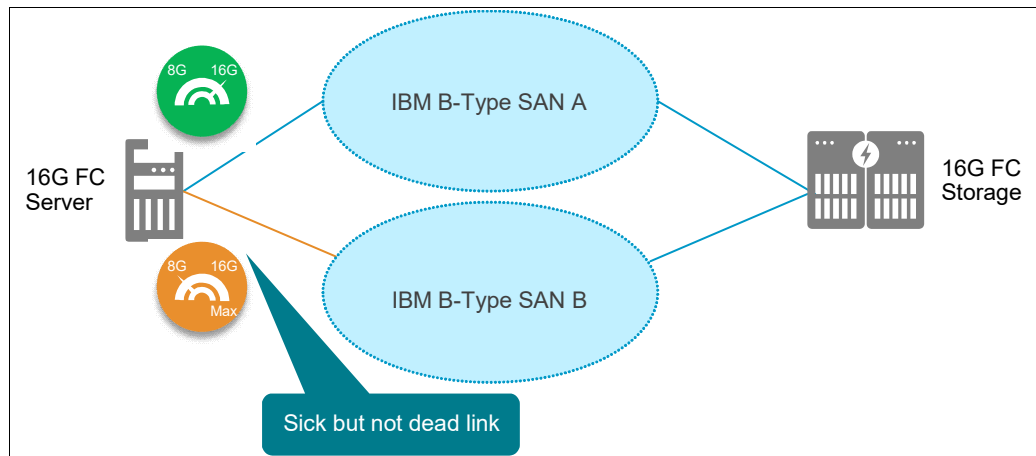


Figure 3-13 Effect of a bad link

With the Link Integrity feature, as shown in Figure 3-14, when a “sick” but not dead path is detected, a Fabric Performance Impact Notification – Link Integrity (FPIN-LI) notification is sent to the host HBA. Then, the Emulex HBA passes the FPIN up the stack to the AIX multi-pathing software to address the issue by taking corrective actions.

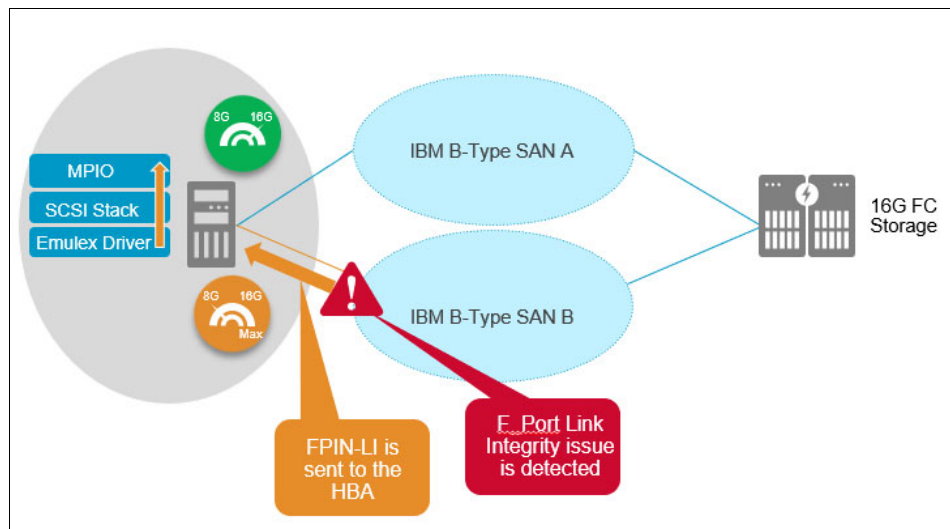


Figure 3-14 Fabric detected link integrity issues

As shown in Figure 3-15, after the notification is received by the host HBA, the multi-pathing software makes intelligent decisions based on the data from the IBM b-type Gen 7 SAN fabric to rebalance the traffic away from the bad link, which results in improved availability and minimizes performance problems.

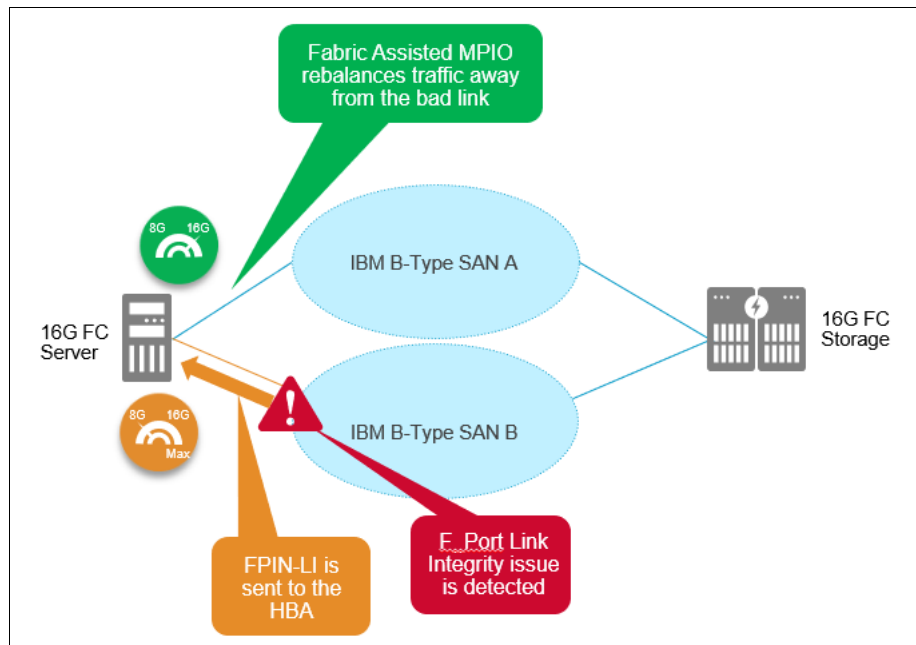


Figure 3-15 Addressing Link Integrity issues by host HBA

3.8.9 SCSI Command Delivery

When the fabric discards a command or status, Fabric Notifications notifies the originator of the failure to deliver by sending an FPIN Delivery notification (see Figure 3-16). No matter if the command is dropped by an E_Port or F_Port, the originator is notified. The notification starts an immediate SCSI I/O failure without waiting for a timeout. This ability allows the end-device to immediately retry the discarded command to avoid long wait times and application disruption.

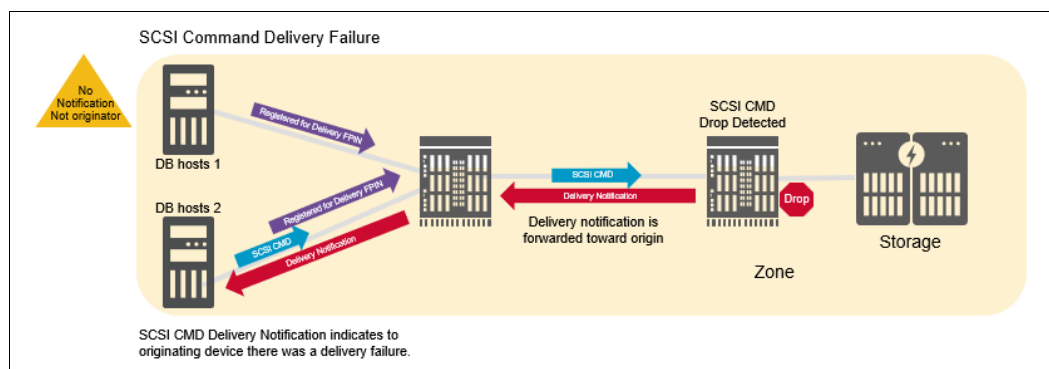


Figure 3-16 FPIN SCSI CMD delivery notification

The Brocade's Fabric Notifications solution addresses each of these issues through a technology system of innovation. Various notifications with problem-specific descriptors are sent to or received from attached end-devices. In some cases, they are propagated to other directors or switches to notify peer end-devices.

Notifications are limited to registered in-zone devices and the device must register for each notification type it wants to receive.

Fabric Notifications compliments MAPS technology and was introduced by Broadcom BSN in FOS 9.0.



Virtual Fabrics overview

Virtual Fabrics is an architecture that is used to virtualize hardware boundaries. Traditionally, SAN design and management is done at the granularity of a physical switch. Virtual Fabrics allows SAN design and management to be done at the granularity of a port.

Virtual Fabrics is a suite of related features that can be customized based on your needs. The Virtual Fabrics suite consists of the following specific features:

- ▶ Logical switch
- ▶ Logical fabric
- ▶ Device sharing

Terminology: *Virtual Fabrics* is the name of the suite of features. A *logical fabric* is a type of fabric that you can create by using the Virtual Fabrics suite of features.

This chapter describes a logical fabric and logical switch types, and device sharing with Virtual Fabrics.

Important: SNMPv3 is required to manage Virtual Fabrics. If you use SNMPGET to poll the switches (for example HiTrack), the SNMPv3 user must be configured based on the logical switch that is being monitored. If SNMP monitoring is for the default switch, you do not need the SNMP user to be configured as a user on the switch.

When SNMP monitoring is VF-specific, you must create a user with permission for the corresponding VF and add the created user into the SNMP database as a SNMPv3 user. When the SNMPv3 user is successfully added, you can use the same username in management applications for SNMP queries to the corresponding logical switch.

This chapter includes the following topics:

- ▶ 4.1, “Logical switch” on page 54
- ▶ 4.2, “Logical fabric” on page 60
- ▶ 4.3, “Device sharing” on page 70

4.1 Logical switch

Traditionally, each switch and all the ports in the switch act as a single Fibre Channel (FC) switch that participates in a single fabric. The logical switch feature allows you to divide a physical chassis into multiple fabric elements. Each of these fabric elements is referred to as a *logical switch*. Each logical switch functions as an independent, self-contained FC switch.

4.1.1 Default logical switch

To use the Virtual Fabrics features, you must first enable Virtual Fabrics on the switch, which creates a single logical switch in the physical chassis. This logical switch is called the *default logical switch*.

The default logical switch initially contains all of the ports in the physical chassis.

Figure 4-1 shows a switch before and after enabling Virtual Fabrics. In this example, the switch has 10 ports, which are labeled P0 - P9.

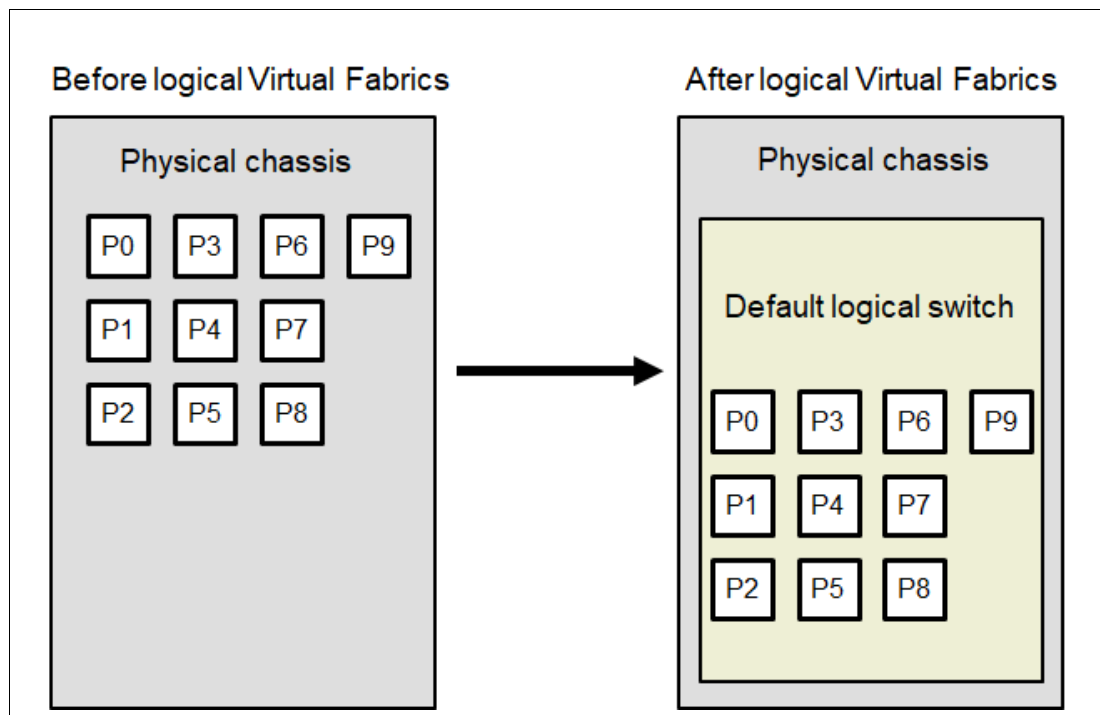


Figure 4-1 Switch before and after enabling virtual fabrics

Figure 4-2 shows a Virtual Fabrics-enabled switch before and after it is divided into logical switches. Before you create logical switches, the chassis appears as a single switch (default logical switch). After you create logical switches, the chassis appears as multiple independent logical switches. All of the ports continue to belong to the default logical switch until you move them to other logical switches.

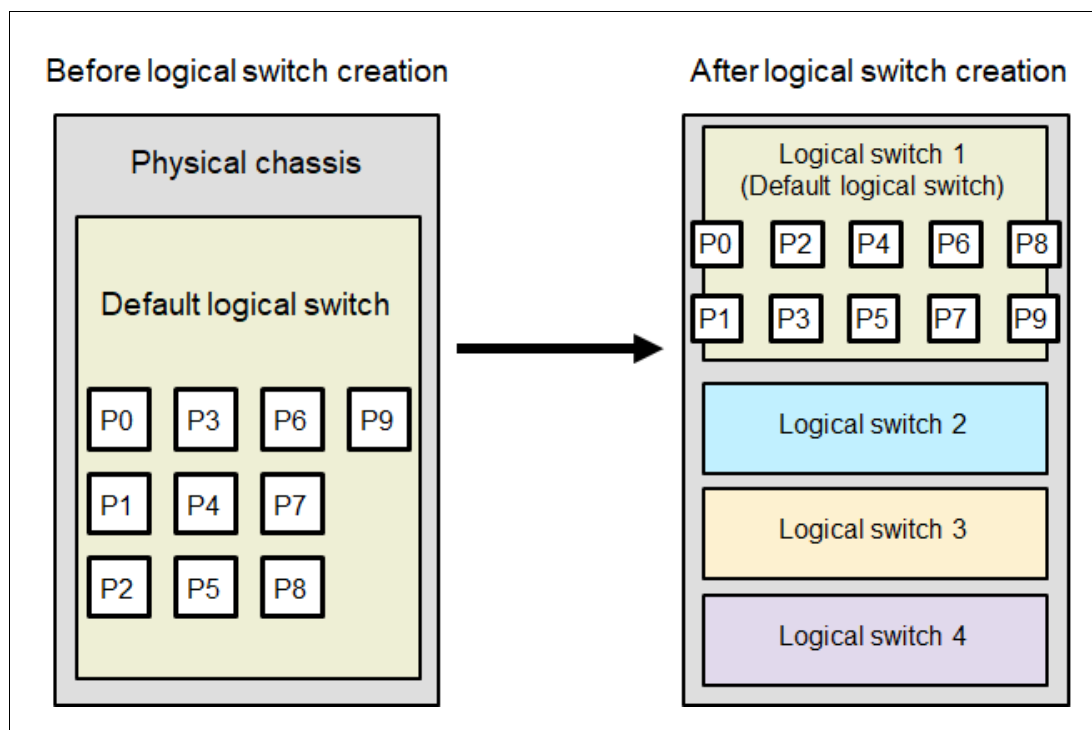


Figure 4-2 Switch before and after creating logical switches

The default logical switch always exists. You can add and delete other logical switches, but you cannot delete the default logical switch unless you disable Virtual Fabrics.

Note: A FICON logical switch or an existing logical switch that was modified to become a FICON logical switch cannot serve as the default logical switch.

4.1.2 Logical switches and fabric IDs

When you create a logical switch, you must assign it a fabric ID (FID). The fabric ID uniquely identifies each logical switch within a chassis and indicates to which fabric the logical switch belongs. You cannot define multiple logical switches with the same fabric ID within the chassis.

As shown in Figure 4-3, logical switches 2, 3, 4, and 5 are assigned FIDs of 1, 15, 8, and 20, respectively. These logical switches belong to different fabrics, even though they are in the same physical chassis. For example, you cannot assign logical switch 5 a fabric ID of 15 because logical switch 3 is assigned FID 15 in the chassis. The default logical switch is initially assigned FID 128. You can change this value later.

Note: Each logical switch is assigned only one FID. The FID identifies the logical fabric to which the logical switch belongs.

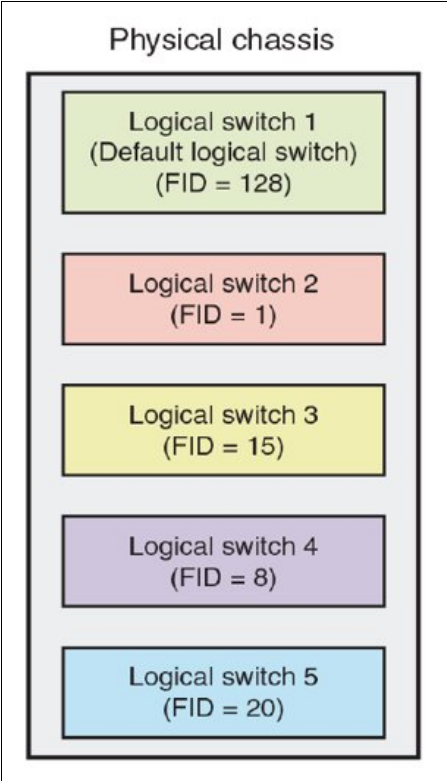


Figure 4-3 Fabric IDs assigned to logical switches

4.1.3 Port assignment in logical switches

Initially, all ports belong to the default logical switch. When you create logical switches, they are empty, and you must assign ports to those logical switches. As you assign ports to a logical switch, the ports are moved from the default logical switch to the newly created logical switch. A specific port can be in only one logical switch.

As shown in Figure 4-4 on page 57, the default logical switch initially has 10 ports, which are labeled P0 - P9. After logical switches are created, the ports are assigned to specific logical switches. Ports 0, 1, 7, and 8 are not assigned to a logical switch; therefore, they remain assigned to the default logical switch.

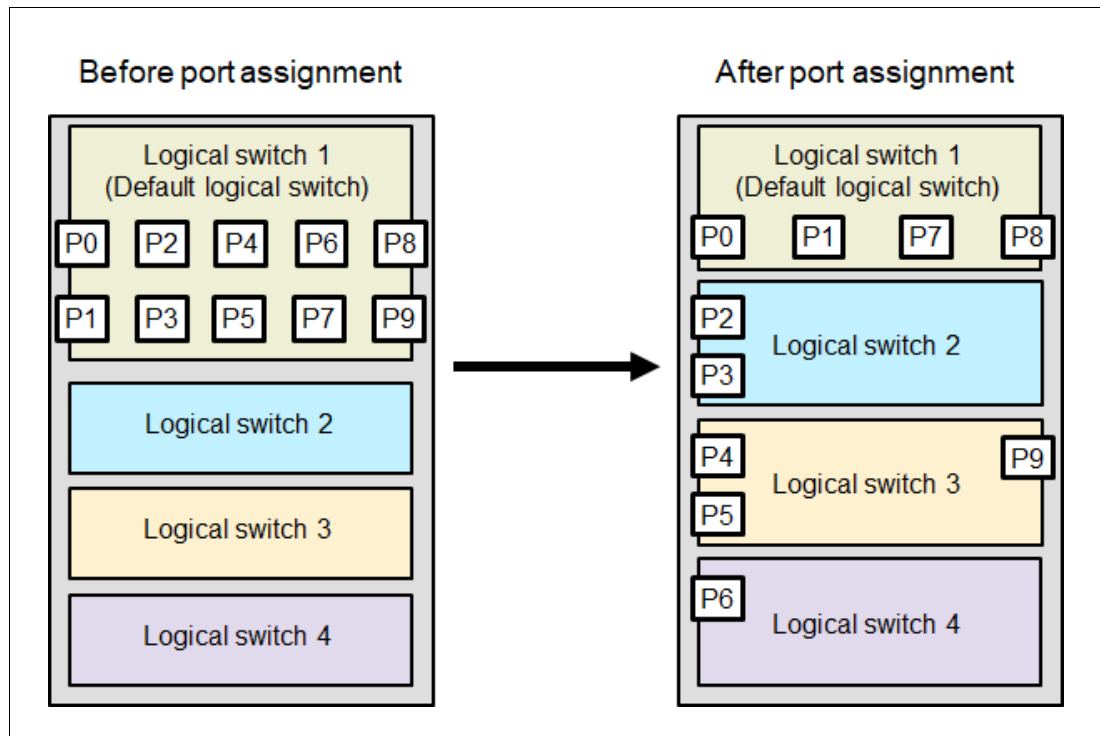


Figure 4-4 Assigning ports to logical switches

A specific port is always in only one logical switch. The following scenarios refer to the chassis after port assignment that is shown in Figure 4-4:

- ▶ If you assign P2 to logical switch 2, you cannot assign P2 to any other logical switch.
- ▶ If you want to remove a port from a logical switch, you cannot delete it from the logical switch, but must move it to a different logical switch. For example, if you want to remove P4 from logical switch 3, you must assign it to a different logical switch: logical switch 2, logical switch 4, or logical switch 1 (the default logical switch).
- ▶ If you assign a port to a logical switch, it is removed automatically from the logical switch it is in. If you assign P3 to Logical switch 3, P3 is automatically removed from logical switch 2.
- ▶ If you do not assign a port to any logical switch, it remains in the default logical switch, as is the case with ports 0, 1, 7, and 8.

A logical switch can have as many ports as are available in the chassis. As shown in Figure 4-4, the chassis has 10 ports. You can assign all 10 ports to a single logical switch, such as logical switch 2; however, if you made this assignment, no ports are available for logical switches 3 and 4.

You can move only F_Ports and E_Ports from one logical switch to another. If you want to configure a different type of port, such as a VE_Port or EX_Port, you must configure them after you move them. Some types of ports cannot be moved from the default logical switch.

4.1.4 Logical switches and connected devices

You can connect devices to logical switches, as shown in Figure 4-5. In logical switch 2, P2 is an F_Port that is connected to H1. In logical switch 3, P4 is an F_Port that is connected to D1. H1 and D1 cannot communicate with each other because they are in different fabrics, even though they are both connected to the same physical chassis.

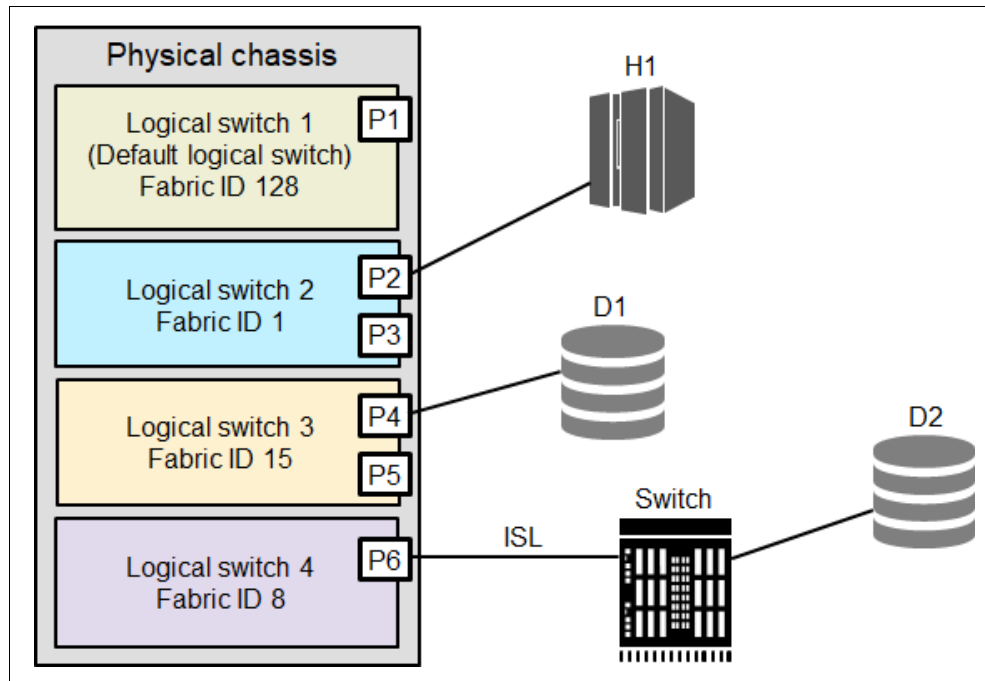


Figure 4-5 Logical switches that are connected to devices and non-Virtual Fabrics switches

You can also connect other switches to logical switches. In Figure 4-5, P6 is an E_Port that forms an inter-switch link (ISL) between logical switch 4 and the non-Virtual Fabrics switch. Logical switch 4 is the only logical switch that can communicate with the non-Virtual Fabrics switch and D2 because the other logical switches are in different fabrics.

Figure 4-6 on page 59 shows the logical switch view of a chassis where virtual fabrics are enabled and the devices are connected. The logical switch that includes Fabric 128 (see Figure 4-6 on page 59) is the Default switch when virtual fabrics are enabled. Logical switches Fabric 1 (connected to Host) and Fabric 15 (connected to Storage), and logical switch Fabric 8 is connected to a non-Virtual Fabric switch and storage.

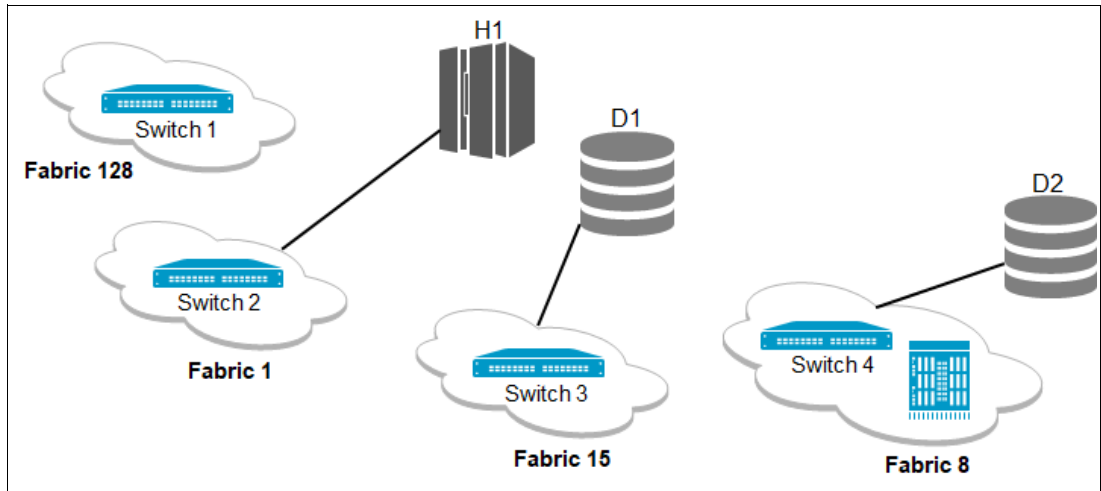


Figure 4-6 Logical switches in a Single Chassis belong to separate fabrics

4.1.5 Management model for logical switches

The operations that you can perform on a logical switch depend on the context you are in. Some operations affect only a single logical switch, and some operations affect the entire physical chassis.

All user operations are classified into one of the following chassis management operations that span logical switch boundaries:

- ▶ Logical switch configuration (creating, deleting, or modifying logical switches)
- ▶ Account management (determining which accounts can access which logical switches)
- ▶ Field-replaceable unit (FRU) management (`slot` commands, such as `slotShow`)
- ▶ Firmware management (firmware upgrade, high availability failover)

These operations are limited to the logical switch, such as displaying or changing port states. Logical switch operations include all operations that are not covered in the chassis management operations.

When you log in, you are assigned an active context, or active logical switch. This context filters the view that you receive, and determines which ports you can see. You can change the active context. For example, if you are working with logical switch 1, you can change the context to logical switch 5. When you change the context to logical switch 5, you see only the ports that are assigned to that logical switch. You do not see any of the other ports in the chassis.

The scope of logical switch operations is defined by the active context. When you are in the context of a logical switch, you can perform port, switch, and fabric-level operations, subject to role-based access control (RBAC) rules.

If you have permission to run chassis-level commands, you can do so, regardless of which logical switch context you are in.

4.2 Logical fabric

A *logical fabric* is a fabric that contains at least one logical switch.

The four fabrics that are shown in Figure 4-5 on page 58 and Figure 4-6 on page 59 are logical fabrics because they each have at least one logical switch. You can connect logical switches to non-Virtual Fabrics switches and to other logical switches. You connect logical switches to non-Virtual Fabrics switches that use an ISL, as shown in Figure 4-5 on page 58.

You connect logical switches to other logical switches in the following ways:

- ▶ Using ISLs
- ▶ Using base switches and extended ISLs (XISLs)

4.2.1 Logical fabric and ISLs

Figure 4-7 shows two physical chassis that are divided into logical switches. ISLs are used to connect the logical switches with FID 1 and the logical switches with FID 15. The logical switches with FID 8 are each connected to a non-Virtual Fabrics switch. The two logical switches and the non-Virtual Fabrics switch are all in the same fabric, with FID 8.

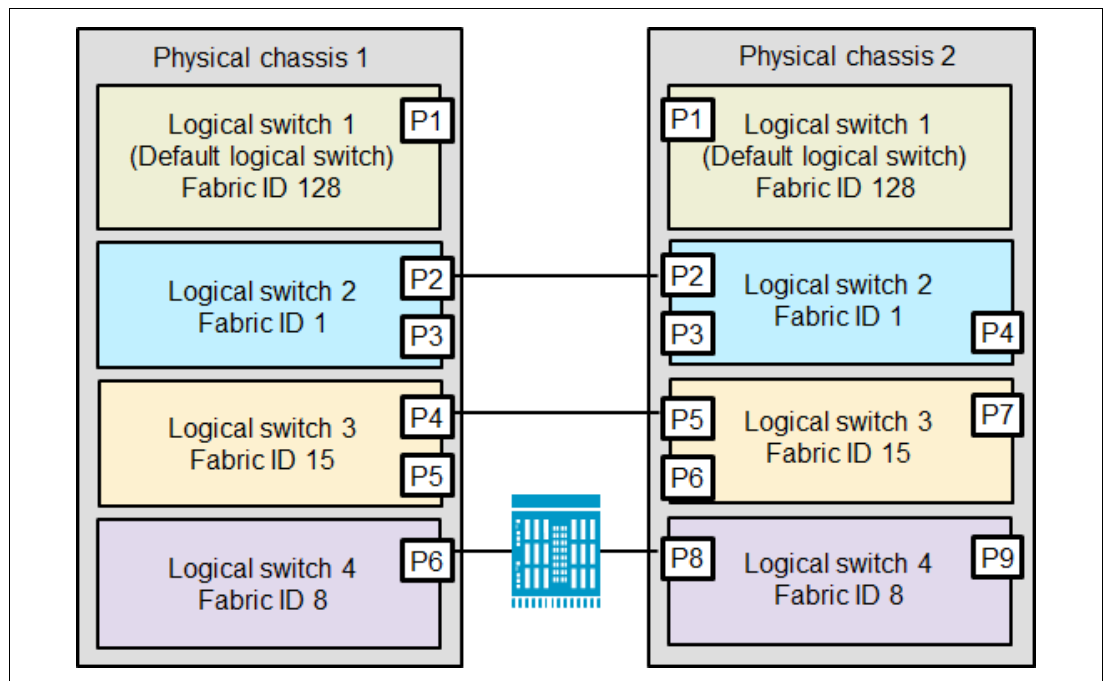


Figure 4-7 Logical switches that are connected to other logical switches through physical ISLs

The ISLs between the logical switches are *dedicated ISLs* because they carry traffic only for a single logical fabric. In Figure 4-8 on page 61, Fabric 128 has two switches (the default logical switches), but they cannot communicate with each other because they have no ISLs between them and they cannot use the ISLs between the other logical switches.

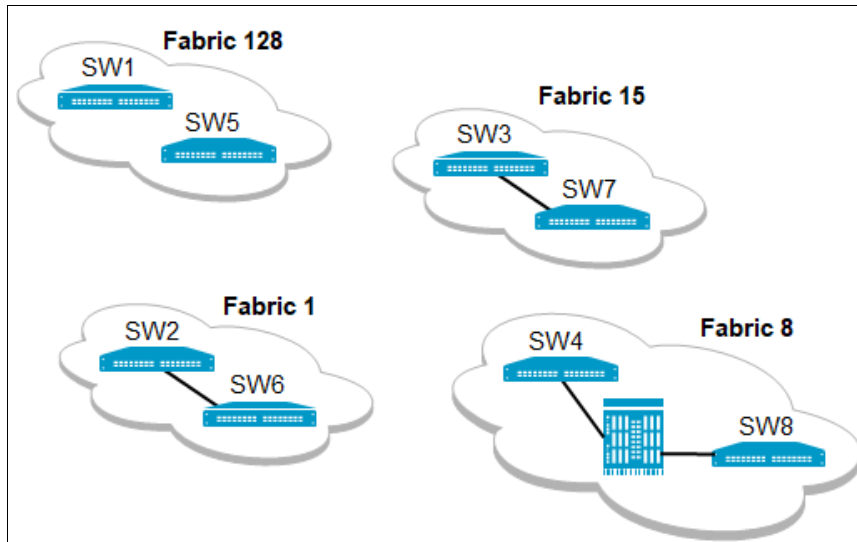


Figure 4-8 Logical switches connected to form logical fabrics

Note: Only logical switches with the same FID can form a fabric. If you connect two logical switches with different FIDs, the link between the switches segments.

4.2.2 Base switch and extended ISLs

One method to connect logical switches is to use extended ISLs and base switches.

When you divide a chassis into logical switches, you can designate one of the switches to be a base switch. A *base switch* is a special logical switch that is used for interconnecting the physical chassis.

A base switch features the following properties:

- ▶ ISLs that are connected through the base switch can be used for communication among the other logical switches.
- ▶ Base switches do not support direct device connectivity. A base switch can have only E_Ports, VE_Ports, or EX_Ports, but no F_Ports.
- ▶ The base switch provides a common address space for communication between different logical fabrics.
- ▶ A base switch can be configured for the preferred domain ID, just like a non-Virtual Fabrics switch.
- ▶ You can have only one base switch in a physical chassis.

A base switch can be connected to other base switches through a special ISL, which is called a *shared ISL* or *extended ISL* (XISL). An extended ISL connects base switches. The XISL shares traffic among different logical fabrics.

Note: A FICON logical switch or a logical switch that was modified to become a FICON logical switch cannot serve as the base switch.

Fabric formation across an XISL is based on the FIDs of the logical switches.

Figure 4-9 shows two physical chassis that are divided into logical switches. Each chassis has one base switch. An ISL connects the two base switches. The ISL is an extended ISL (XISL) because it connects base switches.

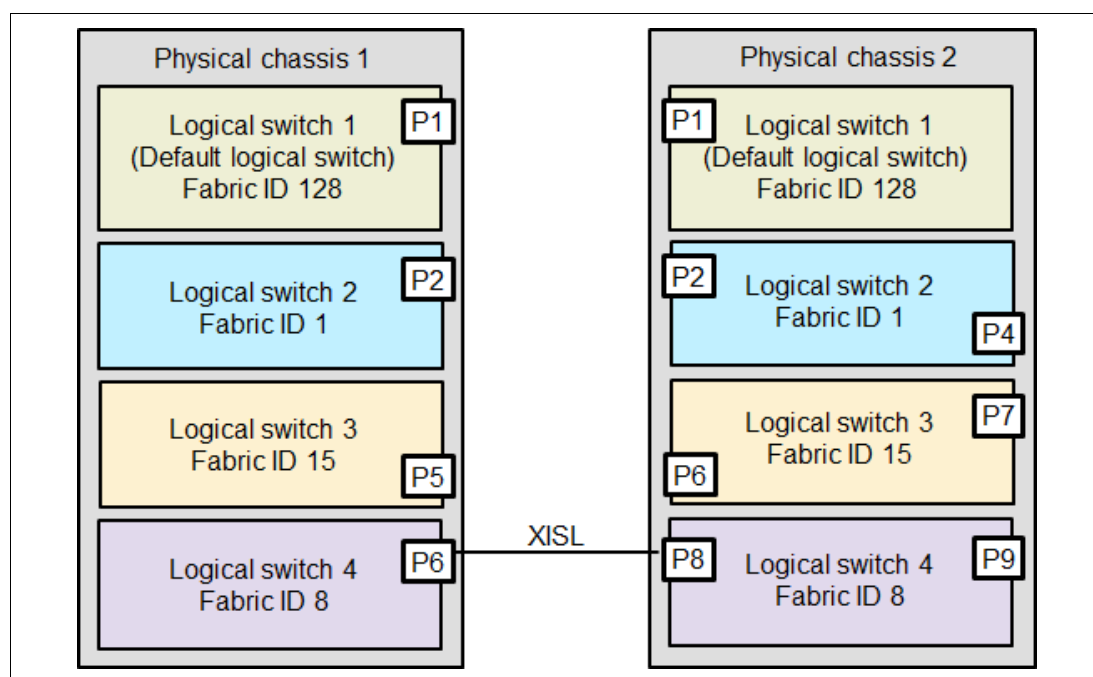


Figure 4-9 Base switches connected by an XISL

Traffic between the logical switches can now flow across this XISL. The traffic can flow only between logical switches with the same fabric ID. For example, traffic can flow between logical switch 2 in chassis 1 and logical switch 6 in chassis 2 because they both have FID 1. Traffic cannot flow between logical switch 2 and logical switch 7 because they have different fabric IDs (and are, therefore, in different fabrics).

It is useful to think of the logical switches as being connected with logical ISLs, as shown in Figure 4-10 on page 63. The logical ISLs are not connected to ports because they are not physical cables. They are a logical representation of the switch connections that are allowed by the XISL.

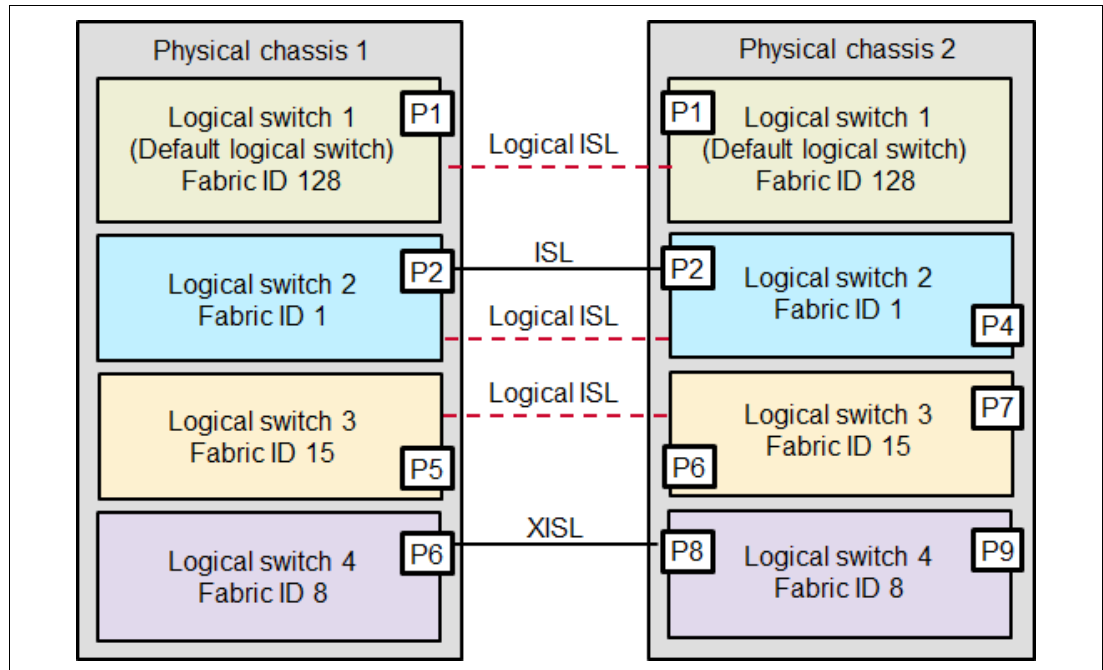


Figure 4-10 Logical fabric that uses ISLs and XISLs

You can also connect logical switches by using a combination of ISLs and XISLs, as shown in Figure 4-10. In this example, traffic between the logical switches in FID 1 can travel over the ISL or the XISL. Traffic between the other logical switches travels over the XISL only.

By default, the physical ISL path is favored over the logical path (over the XISL) because the physical path has a lower cost. This behavior can be changed by configuring the cost of the dedicated physical ISL to match the cost of the logical ISL.

Attention: If you disable a base switch, all of the logical ISLs are disconnected and the logical switches cannot communicate with each other unless they are connected by a physical ISL.

Base fabric

Base switch ports on different chassis can be connected to form a fabric, called a *base fabric*. Similar to other logical switches, the base switches must have the same FID to be connected. If the base switches have different FIDs, the link between the switches is disabled.

The base fabric follows normal routing policies. If physical connectivity is available, the base fabric maintains connectivity for the logical fabrics.

Logical ports

Logical ISLs are formed to connect logical switches. A *logical port* represents the ports at each end of a logical ISL. A logical port is a software construct only and does not correspond to any physical port.

Most port commands are not supported on logical ports. For example, you cannot change the state or configuration of a logical port. However, state change on a logical link is permitted by using the `lfcfg --lislEnable` command.

The worldwide name (WWN) for logical ports is in NAA=5 format, and uses the following syntax:

5n:nn:nn:nz:zz:zz:zx:xx

The NAA=5 syntax uses the following variables:

- ▶ *nnnnnn* is the Organizationally Unique Identifier (OUI).
- ▶ *zzzzzz* is the logical fabric serial number.
- ▶ *xxx* is the logical port number, in the range 0 - FFF.

Logical fabric formation

Fabric formation is not based on connectivity; rather, it is based on the FIDs of the logical switches. The fabric formation uses the following basic order:

- ▶ Base fabric forms.
- ▶ Logical fabrics form when the base fabric is stable.
- ▶ Devices begin recognizing one another.
- ▶ Traffic is started between the logical switches.

4.2.3 Account management and Virtual Fabrics

When user accounts are created, they are assigned a list of logical fabrics to which they can log in and a home logical fabric (home FID). When you connect to a physical chassis, the home FID defines the logical switch to which you are logged in by default. You can change to a different logical switch context by using **setcontext** command.

When you are logged in to a logical switch, the system prompts changes to display the FID of that switch. The following example prompts for when you are logged in to the default logical switch (FID = 128) and a user-defined logical switch (FID = 15):

```
switch:FID128:admin>
switch:FID15:admin>
```

4.2.4 Setting up IP addresses for a logical switch

Each physical chassis has one common IP address that is shared by all of the logical switches in the chassis. You can set up individual IPv4 addresses for each logical switch.

IPv4 addresses that are assigned to individual logical switches are assigned to IP over Fibre Channel (IPFC) network interfaces. In Virtual Fabrics environments, a single chassis can be assigned to multiple fabrics, each of which is logically distinct and separate from one another. Each IPFC point of connection to a specific chassis needs a separate IPv4 address and prefix to be accessible to a management host.

For a management host to access logical switches, the host bus adapter (HBA) must connect with the common, shared IP address and the individual IPv4 addresses that are configured for each logical switch.

Note: IPv6 is not supported when setting the IPFC interface for Virtual Fabrics.

IPFC addresses are not handled by configupload or configdownload. The IPFC address of the default logical switch or a non-VF switch is stored on the WWN card or compact flash. This address does not display in a configshow. Non-default logical switch IPFC addresses display in a configshow. The **ipaddrshow** command displays all switch addresses and IPFC addresses configured in the chassis.

Use the following procedure to set up IP addresses for a logical switch:

- ▶ Connect to the switch and log in by using an account with admin permissions.
- ▶ Enter the **ipAddrSet -ls** command. Consider the following points:
 - To add an IPv4 address, use the **--add** parameter. Specify the network information in dotted-decimal notation for the Ethernet IPv4 address with a Classless Inter-Domain Routing (CIDR) prefix.
 - To delete an IPv4 address, use the **--delete** parameter.
- ▶ Enter the **ipaddrshow** command to verify the result.

Example 4-1 shows how to sets IP addresses with the CIDR prefix for logical switches with FID 1, 2, and 128 (default logical switch).

Example 4-1 Setting IP addresses

```
switch:FID128:admin> ipaddrset -ls 1 --add 192.0.2.11/24
IP address is being changed...Done.
switch:FID128:admin> ipaddrset -ls 2 --add 192.0.2.22/24
IP address is being changed...Done.
switch:FID128:admin> ipaddrset -ls 128 --add 192.0.2.0/24
IP address is being changed...Done.
```

```
switch:FID128:admin> ipaddrshow
```

```
SWITCH
Ethernet IP Address: 198.51.100.0
Ethernet Subnetmask: 255.255.255.0
Gateway IP Address: 198.51.100.1
DHCP: Off
IPFC address for virtual fabric ID 128: 192.0.2.0/124
IPFC address for virtual fabric ID 1: 192.0.2.11/24
IPFC address for virtual fabric ID 2: 192.0.2.22/24
```

Example 4-2 show how to delete an IP address for the logical switch with FID 1.

Example 4-2 Deleting an IP address

```
switch:FID128:admin> ipaddrset -ls 1 --delete
```

4.2.5 Logical switch log in context

When you log in to a logical switch with Telnet or SSH, the login shell uses the destination IPFC address to look up and set the corresponding logical switch context. When you log in from the serial console, the login shell always sets your default home logical switch context.

The IPFC that is assigned to a logical switch is managed by using the **ipAddrSet** command. When you use Telnet or SSH from a remote host, the IPFC address must share the network (subnet mask) as the management interface.

The values of FCIP addresses and subnet masks are stored in the WWN card. For example, if the management interface IP address is 10.38.18.183 with subnet mask 255.255.240.0, you can create the IPFC address (see Example 4-3).

Example 4-3 Adding an IP address

```
ipaddrset -ls 8 --add 10.38.18.183/20
```

Note: The logical switch login context supersedes the default FID context of a user. The CIDR subnet mask of the IPFC address must match the CIDR of the assigned management IP address to avoid routing issues.

To delete the IPFC address, you can use the **ipaddrset -ls 8 -delete** command.

If you log in to a switch with the management interface IP address, the home virtual fabric is set for the context. You can log in to the logical switch context by using the telnet or SSH service with the configured IPFC address. You can manage the permissions by using the **userConfig** command (see Example 4-4).

Example 4-4 Using the userConfig command

```
switch> userconfig --show admin
Account name: admin
Description: Administrator
Enabled: Yes
Password Last Change Date: Unknown (UTC)
Password Expiration Date: Not Applicable (UTC)
Locked: No
Home LF Role: admin
Role-LF List: admin: 1-128
Chassis Role: admin
Home LF: 128
Day Time Access: N/A
```

Role-LF is the list of logical switch contexts for which you have permission to log in over the IPFC address. *Home LF Role* is the default logical switch context when you do not have permission to log in to a specific logical switch context or over management interface. Example 4-5 shows the result of the **ipAddrShow** command. Only one IPFC is mapped to FID 8.

Example 4-5 Using the ipAddrShow command

```
switch> ipaddrshow
SWITCH
Ethernet IP Address: 10.38.18.183
Ethernet Subnetmask: 255.255.240.0
Gateway IP Address: 10.38.16.1
DHCP: Off
IPFC address for virtual fabric ID 8: 10.38.18.183/20
```

The following conditions apply for default users:

- ▶ For admin and root users, the login is successful to the specific logical switch with the corresponding logical switch context set (FID 1 to 128). When you telnet to IPFC 10.38.18.183, the context is set to FID 8.

Note: If the home VF of a user is not present on the switch, the user is logged in to default VF if the default VF is part of the user's VFs list. If not, the user is logged in to the lowest VF that is available on the switch that is part of the user's VFs list.

The following conditions apply for non-default users:

- ▶ For non-default users that are configured in the switch with or without the logical switch permission, if the user logs in with the IPFC address without the permission for the corresponding logical switch, only the default home virtual fabric of the user is set for the context.
- ▶ For non-default users that are configured in the switch with or without the logical switch permission, if the user logs in with the IPFC address with the permission for the corresponding logical switch, login is successful to the specific logical switch with the corresponding logical switch context set.

4.2.6 Supported platforms for Virtual Fabrics

The following platforms support the Virtual Fabrics:

- ▶ IBM Storage Networking SAN512B-7
- ▶ IBM Storage Networking SAN256B-7
- ▶ IBM Storage Networking SAN64B-7

Some restrictions apply to the ports, depending on the port type and blade type. The following sections explain these restrictions.

Supported port configurations in the fixed-port switch

No restrictions exist on the ports in IBM SAN64B-7 switches.

The following rules apply:

- ▶ Any port can belong to any logical switch (including the base switch and default logical switch), with the exception that F_Ports cannot belong to the base switch.
- ▶ The default logical switch can use XISLs.
- ▶ The default logical switch can also be a base switch, except on Gen 7 directors and Gen 6 directors.

Supported port configuration in directors

Some of the ports in the Gen 7 directors are not supported on all types of logical switches. Table 4-1 lists the blades and ports that are supported on each type of logical switch.

Table 4-1 Blade and port types that are supported on logical switches

Blade type	Default logical switch	User-defined logical switch	Base switch
FC64-48: FC Port	Yes (F, E)	Yes (F, E)	Yes (F, E)
FC32-X7-48: FC Port	Yes (F, E)	Yes (F, E)	Yes (F, E)
FC32-64: FC Port FC32-64: Eth Port	Yes (F, E) Yes	Yes (F, E) Yes	Yes (F, E) Yes
SX6: FC ports GE ports	Yes (F, E) Yes (VE)	Yes (F, E) Yes (VE)	Yes (F, E) Yes (VE)
ICL ports	Yes (E)	Yes (E)	Yes (E,EX)

Virtual Fabrics interaction with other Fabric OS features

Table 4-2 lists some Fabric OS (FOS) features and considerations that apply when Virtual Fabrics are used.

Table 4-2 Virtual Fabrics interaction with Fabric OS features supported logical switch creation limits

Fabric OS features	Virtual Fabrics Interaction
Access Gateway	Virtual Fabrics is not supported on a switch if AG mode is enabled.
Configuration upload and download	Virtual Fabrics uses a configuration file that is different from the configuration file that is used to download system configuration parameters.
FC-FC routing service	All EX_Ports must reside in a base switch. You cannot attach EX_Ports to a logical switch that has XISL use enabled. You must use ISLs to connect the logical switches in an edge fabric.
FICON	Up to two logical switches per chassis can run FICON Management Server (CUP), but the FICON logical switch can use ISLs and XISLs.
ISL R_RDY mode	ISL R_RDY mode is supported in the base switch and logical switch.
Licensing	Licenses are applicable for all logical switches in a chassis.
Network time protocol (NTP)	The clock information for a chassis is maintained by the default logical switch and not from the principal or primary FCS switch.
QoS	QoS VCs are maintained across the base fabric.
Traffic isolation	Traffic Isolation with Failover-Disabled (TI FD) zoning is not supported over LISLs. Traffic Isolation zones with Failover-Enabled (TI FE) is supported. TI FD zoning is supported over DISLs (physical links). TI FD zoning is not supported in the base fabric.

Limitations and restrictions of Virtual Fabrics

Before you use the Virtual Fabrics feature, be aware of the restrictions and limitations regarding QSFP ports and the maximum number of logical switches per chassis.

In the core blades, the four ports that belong to the same QSFP can be assigned to different logical switches. However, if one of the four ports belongs to the base switch, the remaining three ports cannot be assigned to any other logical switch.

The maximum number of logical switches per chassis varies depending on the switch model. Table 4-3 lists the supported platforms and the maximum number of logical switches (including the default logical switch) supported on each.

Table 4-3 Supported logical switch creation limits

Platform	Maximum logical switches per chassis ^a
IBM Storage Networking SAN512B-7	16
IBM Storage Networking SAN256B-7	16
IBM Storage Networking SAN64B-7	4

a. Numbers include the default switch and base switch.

Note: If a blade slot is being decommissioned and has ports that are configured in logical switches, the logical port assignments must be removed from that blade before removing the blade. This action ensures a seamless transition for any new port or AP blade that might occupy that slot in the future and does not apply if you are replacing a blade of the same type.

Restrictions on the use of XISLs

The **Allow XISL Use** option of the configure command allows a logical switch to use XISLs in the base switch and any standard ISLs that are connected to that logical switch.

The following restrictions apply when XISLs are used. XISL use is not permitted when the logical switch:

- ▶ Has ICL ports
- ▶ Is the default logical switch
- ▶ Is a base switch
- ▶ Is an edge switch for an FC router

When the logical switch is an edge switch for an FC router, if the logical switch is enabled, you cannot allow XISL use. If the logical switch is disabled or has not yet joined the edge fabric, you *can* allow XISL use; however, fabric segmentation occurs when the logical switch is enabled or connected to an edge fabric.

Note: Using XISL and fmsmode at the same time is permitted, but this combination works in a one-hop topology only.

Restrictions on moving ports

FC ports cannot be moved if any one of the following features is enabled:

- ▶ Long distance
- ▶ QoS
- ▶ F-Port buffers

- ▶ F_Port trunking
- ▶ E_Port credits
- ▶ SIM_Port
- ▶ Encryption

Before moving VE_Ports, you must remove the VE_Port tunnel configuration.

VE_Ports on the SX6 blade can be moved to any logical switch that is independent of the location of the physical GE port.

For more information about how to run a command in a different logical switch, change the fabric ID of a logical switch, or change a logical switch to base switch, see [Brocade Fabric OS Administration Guide, 8.1.x](#).

4.3 Device sharing

The behavior of device sharing in the fabric depends on whether Virtual Fabrics is enabled or not enabled in the base switch. If the Virtual Fabrics is not enabled, FC-FC routing behavior is unchanged. If Virtual Fabrics is enabled, a base switch is like a backbone switch and a base fabric is like a backbone fabric in the FC-FC routing context.

FC-FC routing and Virtual Fabrics

If Virtual Fabrics is enabled, the following rules apply:

- ▶ EX_Ports can be configured on the base switch only. When you enable Virtual Fabrics, the chassis is automatically rebooted. When the switch comes up, only one default logical switch is present, with the default fabric ID (FID) of 128. All previously configured EX_Ports are persistently disabled with the reason ExPort in non-base switch.

You must specifically create a base switch, move the EX_Ports to the base switch, and then enable the ports. If you move EX_Ports to any logical switch other than the base switch, these ports are automatically disabled.

If you want to change an EX_Port on the logical switch to be a non-EX_Port, you must use the **portCfgDefault** command. You cannot use the **portCfgExPort** command because that command is allowed on the base switch only.

- ▶ From FOS v9.0.x onwards, EX_Ports in a base switch receives a 10-bit dynamic area address instead of a zero-bit address. An EX_Port in the base switch is assigned any 10-bit area from 0x0000 - 0xFFC0, with the following exceptions as these areas are reserved for internal lookups:
 - Areas 0x8000-8300
 - 0x8040-8340
 - Areas 0x8080-8380
 - Areas 0x80C0-83C0
 - 0xC500
 - 0xC540
 - 0xC580
 - 0xC5C0

- ▶ EX_Ports can connect to a logical switch that is in the same chassis or in a different chassis. However, the following configuration rules apply:
 - If the logical switch is on the same chassis, the EX_Port FID must be set to a different value than the FID of the logical switch to which it is connecting.
 - If the logical switch is on a different chassis, no FID for any logical switch in the FC router backbone fabric can be the same as the FID of the logical switch to which the EX_Port is connecting.
- ▶ EX_Ports in FC routers and those in a base switch cannot connect to any edge fabric with logical switches that are configured to use XISLs. If you connect an EX_Port to an edge fabric, you must ensure that no logical switches with XISL use are enabled in that edge fabric. If any logical switch in the edge fabric allows XISL use, the EX_Port is disabled. Because XISL use is disallowed, dedicated links must be configured to route traffic across switches in the same logical fabric.
- ▶ Backbone-to-edge routing is not supported in the base switch (see Figure 4-7 on page 60 for an example of established FC routers that allow backbone-to-edge routing with Virtual Fabrics).
- ▶ All FC router commands can be run in the base switch context only.
- ▶ The **fcrconfigure --bbfid** command *cannot* be used when Virtual Fabrics is enabled. Instead, use the **lscfg** command to configure the FID.
- ▶ From FOS v9.0.x onwards, if a port is configured as an EX_Port and if the administrator moves the port from the base switch to another logical switch, the operation fails. You must remove the EX_Port configuration before moving it to another logical switch.
- ▶ When a Virtual Fabrics-enabled switch is part of the backbone, only the base switch with active EX_Ports can route frames between edge fabrics. If other base switches are participating in the fabric and do not have active EX_Ports, those switches participate as transient switches to forward frames to the next hop until they can be routed to the intended remote edge fabric.

If any other default or logical switch is introduced as a hop between edge fabrics, the newly introduced switch that is not the base switch does not merge with the base switch fabric. This issue causes frames to be dropped.

For more information about device sharing and FC-FC routing, see [Brocade Fabric OS Administration Guide, 9.0.x](#).



Fibre Channel SAN design, implementation, and migration

This publication is a high-level storage area networking (SAN) design and best-practices guide that is based on IBM b-type Gen 7 products and features, focusing on Fibre Channel (FC) Storage Area Network (SAN) design. The storage landscape continues to modernize, and many choices must be made to design the correct FC architecture. Covered topics include the early planning phase, understanding possible operational challenges, and monitoring and improving an existing SAN infrastructure.

Modern applications have an insatiable need for data. The rise of NVMe flash and storage-class memory drastically accelerate the performance of storage systems that are tasked with holding that data. This means that having the right network technology in place is essential.

Fortunately, the IBM b-type Gen 7 networking solution delivers an autonomous SAN with next-level performance and integrated intelligence. IBM b-type Gen 7 provides the accelerated foundation that modern data centers need to fully realize the benefits of the latest storage technology today and possibly for the next decade.

This chapter focuses on FC SAN design and best practices. It covers planning, topologies, workload monitoring, and detecting server and storage latencies, which help with decisions that are required for successful IBM b-type Gen 7 SAN design.

This chapter includes the following topics:

- ▶ 5.1, “Modernizing SAN infrastructures” on page 74
- ▶ 5.2, “Fibre Channel SAN design” on page 75
- ▶ 5.3, “Implementation planning for a Fibre Channel SAN” on page 86
- ▶ 5.4, “Migration strategies of Fibre Channel SAN infrastructure” on page 95
- ▶ 5.5, “References” on page 103

5.1 Modernizing SAN infrastructures

Technology is evolving at an incredible pace, and businesses are demanding more from their IT resources and infrastructure. Rapid adoption of flash storage and the ramp-up of NVMe-based storage unleash advancements in application design that drive new levels of performance and capacity requirements, such as advanced analytics, business intelligence, and data-intensive workloads.

With the integration of new technologies that are accelerating the delivery of data and services, the network must evolve to keep pace with innovations in storage and the demands of modern applications to maximize the full value of investments in next-generation data center infrastructure. To meet ever-increasing demands for faster, more reliable data access, it is essential for organizations to deploy a modernized infrastructure that reduces latency, increases bandwidth, and ensures continuous availability.

Unprecedented performance is not enough on its own. Powerful analytics and advanced automation capabilities are required to transform current storage networks into an autonomous SAN. This effort requires a network that can deliver these capabilities to maximize performance, simplify management, and reduce operational costs. Older infrastructure was not designed to support the performance requirements of evolving workloads and NVMe-based storage.

In fact, an aging network impedes the performance of the on-demand data center. By modernizing the storage network with IBM b-type Gen 7, organizations enable a faster, more intelligent, and more resilient network. This maximizes the performance, productivity, and efficiency of their storage investments and resources, even as they rapidly scale their environments.

5.1.1 IBM b-type Gen 7 platforms: The intelligent foundation of modern data centers

The products that make up IBM b-type Gen 7 portfolio are the new Brocade X7 Director and the G720 Switch. These solutions are equipped with higher-performing hardware to unleash NVMe technology. They also can discover and produce comprehensive telemetry data across the fabric. They analyze and take actions that are based on that data to optimize the storage network automatically.

5.1.2 Modernize with NVMe to achieve lower latency and higher bandwidth

According to IBM b-type Gen 7, technology substantially accelerates a data center environment, offering 50% lower latency than the previous generation, while increasing bandwidth with 64 Gbps links. By using NVMe-ready FC technology, IBM b-type Gen 7 not only maximizes the performance of NVMe storage and high-transaction workloads, it simplifies the process of integrating higher-performing NVMe storage into the environment without a rip-and-replace process.

IBM b-type Gen 7 simplified IT modernization efforts. The NVMe storage integrates seamlessly with existing ports; Brocade supports the concurrent use of SCSI and NVMe technologies. In fact, with the release of vSphere 7 from VMware, an organization can perform a live migration of an individual virtual machine from SCSI to NVMe in the same Gen 6 or Gen 7 fabric without stopping the application. This result constitutes one of the most risk averse technology transitions in years.

IBM b-type Gen 7 provides investment protection by enabling organizations to upgrade their Brocade Gen 6 director chassis to Gen 7. Organizations can use Gen 6 blades in Gen 7 directors, mixing Gen 6 and Gen 7 blades in the same chassis. In addition, Brocade supports three generations of compatibility with an earlier version within the fabric. This feature is important because it reduces risk and maximizes the value of the organization's IT investments.

5.2 Fibre Channel SAN design

NVMe-ready FC technologies shift leads to a set of demands on the design of storage area networks. This is true regardless of the technology that is used. If the expectation is to use the capacity and performance of these platforms, serious consideration must be given to the design.

Dedicated storage network infrastructure that provides lossless, low-latency, deterministic, scalable, and performant storage services to the applications becomes critical. Storage administrators talk about “fan-in” or “fan-out” ratios for storage platforms, which are the number of servers or applications in the network that use a specific array or array port for their storage access.

Depending on the types of applications and their performance needs, that ratio might range anywhere from low single digits to 40 - 50 servers (hundreds of virtual machines). As with any provisioning scenario, the storage administrator is dealing with projections of how much capacity and performance any server or application is going to use. But, application performance is variable; time of day, week, month, and seasonal or event-driven events cause spikes or drops in application demand.

Another consideration is that the entire application base does not refresh at the same time. In fact, the more common scenario is that multiple generations of performance exist in the environment simultaneously. This is driven by the ability to take a service window to replatform existing environments to new server and storage acquisitions and the fact that some established applications might not feature an environment that runs in current operating system releases or application versions.

You might have 10-year-old or older operating systems that are running connected with a host bus adapter (HBA) that is two or more older generations of technology behind. How then do you balance that still critical application against the needs and performance of the newer machines?

The answer to this question is a combination of topology, balanced provisioning, granular monitoring, and automated mitigation.

5.2.1 SAN design basics

This section provides high-level guidelines for installing a typical SAN. The focus is on best practices for core-edge or edge-core-edge fabrics. The discussion starts at the highest level, the data center, and works down to the port level, and provides recommendations at each point along the way.

Topologies

A typical SAN design comprises devices on the edge of the network, switches in the core of the network, and the cabling that connects all of the devices. Topology often is described in terms of how the switches are interconnected, such as collapsed core, core-edge, and fully meshed.

The recommended SAN topology to optimize performance, management, and scalability is a tiered, core-edge topology. This approach provides good performance without unnecessary interconnections. At a high level, the tiered topology includes many edge switches that are used for device connectivity and a few core switches that are used for routing traffic between the edge switches, as shown in Figure 5-1.

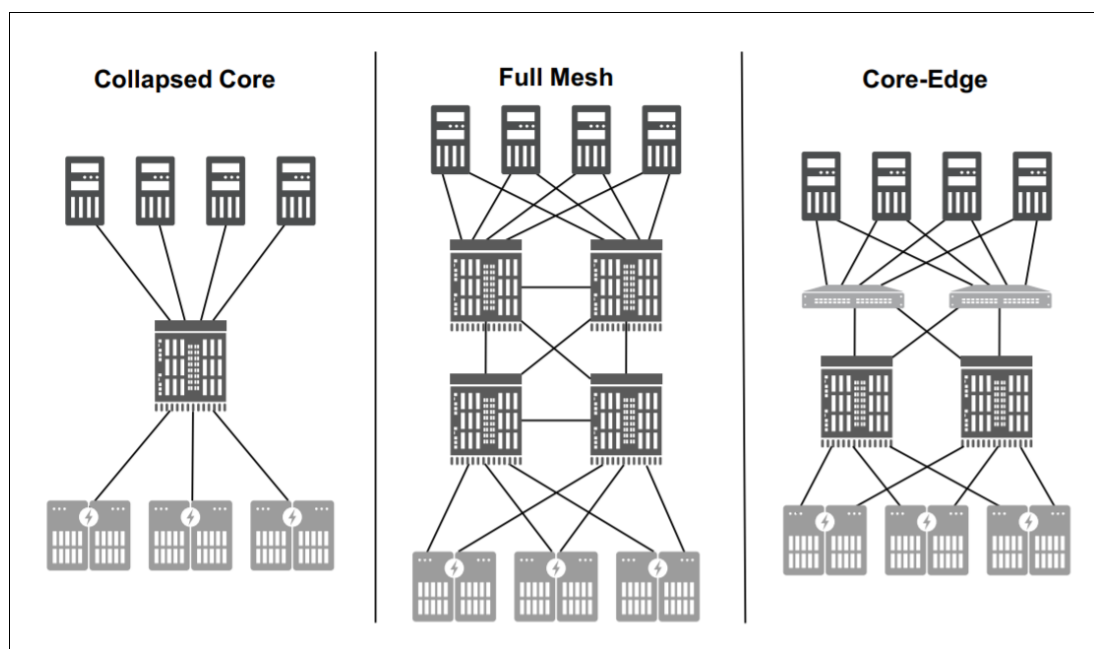


Figure 5-1 Tiered network topologies

The difference between the following scenarios is device placement (where devices are attached to the network) and the associated traffic flow:

- Scenario A, with a collapsed core, has localized traffic, which can sit on the same ASIC, on the same blade, or between blades. Collapsed core can have small performance advantages for performance-optimized workloads, but does not provide ease of scalability and, depending on how many fabrics you need, determines the effect on manageability.
- Scenario B, core-edge, separates the storage and servers, which provides ease of management and higher scalability. An core-edge topology has only one hop from server to storage, and provides similar performance benefits as full mesh while allowing higher scalability.
- Scenario C is a full-mesh topology, and server-to-storage is no more than one hop. Designing fabrics with UltraScale ICLs is an efficient way to save front-end ports. Users can easily build a large (for example, 3456 ports or larger) fabric with minimal SAN design considerations.

Collapsed core

The collapsed core topology places initiators (servers) and storage (targets) on the same ASIC, line card, or chassis. This configuration realizes several benefits, depending on the size of the environment. This is used frequently when customers are migrating from multiple switches to a single dual core architecture where they can fit all initiators and storage into the same chassis. If future design requirements are needed, the use of a core-edge might be more beneficial in the long term regarding scale and ease of management.

Core-edge

The core-edge topology places initiators (servers) on the edge tier and storage (targets) on the core tier. Because the servers and storage are on different switches, this topology provides ease of management and good performance with minimal fabric latency, with most traffic traversing only one hop from the edge to the core. (Storage-to-storage traffic is two hops if the second storage is on another core switch, but the two cores can be connected if fabrics are redundant.)

The disadvantage to this design is that the storage and core connections are in contention for expansion as they scale; however, the use of modular platforms allows for flexibility while allowing the use of ICL ports for intraswitch connectivity to free up device ports.

Full mesh

A full-mesh topology allows you to place servers and storage anywhere because communication between source and destination is no more than one hop. The use of director-class switches with UltraScale ICL ports for interconnectivity is essential to this design to ensure maximum device port availability and utilization. Design this architecture with a minimum of two switches up to nine switches in a full mesh.

5.2.2 Capabilities to consider for your storage architecture

A considerable number of critical capabilities must be considered when choosing the right storage strategy. Some of the more important considerations are availability, scalability, performance, extensibility, and manageability.

Availability

The key to high availability and enterprise-class installation is redundancy. By eliminating a single point of failure, business continuance can be provided through most foreseeable and even unforeseeable events. At the highest level of fabric design, the complete network should be redundant, with two separate fabrics that do not share any network equipment (routers or switches).

Recommendations for the SAN design are to ensure application availability and resiliency by way of the following guidelines:

- ▶ Redundancy that is built into fabrics to avoid a single point of failure
- ▶ Servers that are connected to storage by way of redundant fabrics
- ▶ MPIO-based failover from server to storage
- ▶ Redundant fabrics that are based on similar architectures
- ▶ Redundant ISLs/ICLs for interswitch connectivity
- ▶ Separate storage and server tiers for independent expansion
- ▶ Core switches of equal or higher performance compared to the edges
- ▶ The highest performance switch in the fabric that is defined to be the principal switch

Scalability

IBM b-type Gen 7 products are designed with scalability in mind, knowing that most installations continue to expand and that growth is supported with few restrictions. However, following the same basic principles that were outlined in previously as the network grows ensures that the level of performance and availability continues.

Evaluate the effect on topology, data flow, workload, performance, and most importantly, redundancy and resiliency of the entire fabric when one of the following actions is performed:

- ▶ Adding or removing initiators:
 - Changes in workload
 - Changes in provisioning
- ▶ Adding or removing storage:
 - Changes in provisioning
 - Changes in storage media type (for example, increased deployment of flash-based storage)
- ▶ Adding or removing switches.
- ▶ Adding or removing ISLs and ICLs.
- ▶ Change in virtualization (workload and storage) strategies and traffic flow pattern. If these design best practices are followed when the network is deployed, small incremental changes should not adversely affect the availability and performance of the network. However, if changes are ongoing and the fabric is not correctly evaluated and updated, performance and availability can be jeopardized.

Consider the following points when looking at the status of a production FC network when reviewing redundancy and resiliency:

- ▶ Are there at least two physically independent paths between each source and destination pair?
- ▶ Are there two redundant fabrics?
- ▶ Does each host connect to two different edge switches?
- ▶ Are edge switches connected to at least two different core switches?
- ▶ Are inter-switch connections composed of two trunks of at least two ISLs?
- ▶ Does each storage device connect to at least two different edge switches or separate port blades?
- ▶ Are storage ports provisioned such that every host has at least two ports through which it can access LUNs?

Performance

IBM b-type Gen 7 technology substantially accelerates a data center environment, offering 50% lower latency than the previous generation, while increasing bandwidth with 64 Gbps links. By using NVMe-ready FC technology, IBM b-type Gen 7 maximizes the performance of NVMe storage and high-transaction workloads, and simplifies the process of integrating higher-performing NVMe storage into the environment without a rip-and-replace process.

Reviewing performance requirements:

- ▶ Host-to-storage port fan-in/out ratios
- ▶ Over-subscription ratios:
 - Host to ISL
 - Edge switch to core switch
 - Storage to ISL
- ▶ Size of trunks
- ▶ Routing policy and currently assigned routes; evaluate use for potential imbalances
- ▶ Use of FEC for all ISLs and connections to Gen 5 (if supported) and Gen 6 devices

Watching for effects, such as the following examples:

- ▶ Poor storage performance
- ▶ Overloaded hosts or applications
- ▶ Distance issues over constrained long distance links that result from changes in usage, such as adding mirroring, or too many workloads
- ▶ Deteriorating optics that result in declining signal strength and increased error rate

Extensibility

The IBM b-type Gen 7 FC switch is taking advantage of newer and older technologies within the same infrastructure. It can run multiple vendors within the same architecture to meet evolving storage requirements.

By supporting three generations of backward-compatibility, the Gen 7 storage networking platform provides investment protection by future-proofing support of new protocols, such as NVMe over Fabrics. You want to ensure you do not need to rip and replace.

Manageability

The IBM b-type Gen 7 FC switches are integrated into IBM SANnav Management Suite as the SAN Management Interface. IBM SANnav 2.1.0 is a major new release of IBM's b-type Gen 7 two FC SAN management software products: IBM SANnav Management Portal and IBM SANnav Global View.

SANnav Management Portal 2.1.0 supports the introduction of IBM's new Gen 7 platforms with Fabric OS (FOS) 9.0.0. It adds new features and capabilities that make managing IBM SAN environments easier than ever before.

5.2.3 Designing a SAN

The SAN planning process is similar to any type of project planning and includes the following phases:

- ▶ Phase I: Gathering requirements
- ▶ Phase II: Developing technical specifications
- ▶ Phase III: Estimating project costs
- ▶ Phase IV: Analyzing Return on Investment (ROI) or Total Cost of Ownership (TCO), if necessary
- ▶ Phase V: Creating a detailed SAN design and implementation plan

When selecting which criteria to meet, engage users, server, and storage subject matter experts (SMEs), and other relevant experts to understand the role of the fabric.

Because most SANs tend to operate for a long time before they are renewed, consider future growth because SANs are difficult to redesign. Deploying new SANs or expanding existing ones to meet more workloads in the fabrics requires a critical assessment of business and technology requirements.

A focus on planning ensures that the SAN, once deployed, meets all current and future business objectives, including availability, deployment simplicity, performance, future business growth, and cost.

Design worksheets (see Appendix A, “Fabric details” on page 325) that assist in designing a SAN environment are provided as a reference for documenting assets and metrics for SAN projects.

A critical aspect for successful implementation that is often overlooked is the ongoing management of the fabric. Identifying systems-level SMEs for all components that make up the SAN, and adequate and up-to-date training on those components, is critical for efficient design and operational management of the fabric. When designing a new SAN or expanding an existing SAN, consider the following parameters:

- ▶ Application virtualization
 - Which applications will run under a virtual machine (VM) environment?
 - How many VMs will run on a physical server?
 - Under what conditions will the VMs be migrated (business and nonbusiness hours; is more CPU or memory needed to maintain response times)?
 - Is there a need for solid-state storage media to improve read response times?
- ▶ Homogeneous and heterogeneous server and storage platforms:
 - Is auto-tiering in place?
 - Which FOS versions are supported in a multivendor storage environment?
 - What is the planned refresh cycle of servers and storage platforms (2 years or 3 years)?
- ▶ Sizing
 - How many user ports are needed now?
 - How many devices will connect through an access gateway?
 - How many inter-switch links (ISLs)/Brocade UltraScale inter-chassis links (ICLs) are required to minimize congestion in the fabric?
 - What distances for ISL/ICL connections need to be supported?
 - Does the fabric scale out at the edge or the core?
- ▶ Backup and disaster tolerance:
 - Is there a centralized backup? (This parameter determines the number of ISLs that are needed to minimize congestion at peak loads.)
 - What is the impact of backup on latency-sensitive applications?
 - Is the disaster solution based on a metro FC or Fibre Channel over IP (FCIP) solution?

- Diagnostics and manageability
 - What is the primary management interface to the SAN (command-line interface [CLI], Brocade SANnav, or third-party tool)?
 - How often will Brocade FOS and SANnav be updated?
 - How is cable and optics integrity validated?
 - Is support needed for adding Gen 7 switches into a Gen 6 fabric?
 - Is support needed for storage technologies, such as NVMe over Fabrics?
 - What device interoperability support is required?
 - Is interoperability required for other technologies, such as UCS?

5.2.4 High-performance workloads

Over the last few years, enterprises use low-latency, high-throughput flash arrays for demanding, performance-sensitive workloads. Brocade's Gen 7 FC is perfectly suited to these types of workloads because of the sub-microsecond latency through the switch and the increased bandwidth that is offered by 32 GB and 64 Gb throughput speeds while providing accurate I/O latency instrumentation.

Performance testing shows that 32 Gb and even 8 Gb all-flash arrays can realize dramatic benefits by connecting to a Gen 7 SAN and host adapter, offering gains up to 2x over Gen 5 and Gen 6 SANs.

The Gen 6 and Gen 7 standard includes the use of forward error correction (FEC) to ensure transmission reliability and a highly deterministic data flow. FEC corrects up to 140 corrupt bits per 5280-bit frame at the receiving end of the link, which avoids the need to retransmit frames when bit errors are detected.

For these demanding workloads, a no-hop fabric connection through a single ASIC switch, such as the Brocade G720 or locally switched on a single director port blade minimizes SAN fabric latency to sub-microsecond speeds. *Local switching* is the ability to switch data traffic through a single ASIC by using ingress and egress switching ports in a common port group.

When switches are used that contain multiple switching ASICs, such as the X6 /X7 director, configuring host and target connections on the ports that share a common ASIC minimizes latency by avoiding the need to move data across multiple ASICs/port groups or across a director backplane to a different blade.

For more information about port groups and local switching configuration, see the [Brocade Fabric OS Administration Guide](#) and the hardware installation guide for the suitable product.

Because Gen 7 FC is backwards compatible with Gen 6 networking, a Gen 7 edge switch or director can be added to a Gen 6 fabric. This compatibility allows for local all-flash connectivity to a Gen 7 switch to gain the performance advantages of Gen 7 while preserving the investment in Gen 6 networks.

For more information, see the [Brocade Fabric OS Administration Guide](#).

5.2.5 Redundancy and resiliency

An important aspect of SAN topology is the resiliency and redundancy of the fabric. The main objective is to remove any single point of failure.

Resiliency is the ability of the network to continue to function or recover from a failure; *redundancy* describes duplication of components, even an entire fabric, to eliminate a single point of failure in the network.

Brocade fabrics feature resiliency that is built into FOS, the software that runs on all Brocade B-Series switches, that can quickly “repair” the network to overcome most failures. For example, when a link between switches fails, FSPF quickly recalculates all traffic flows. This assumes that a second route is available, which is when redundancy in the fabric becomes important.

The key to high availability and enterprise-class installation is redundancy. By eliminating a single point of failure, business continuance can be provided through most foreseeable and even unforeseeable events.

At the highest level of fabric design, the complete network should be redundant, with two separate fabrics that do not share any network equipment (routers or switches). Servers and storage devices should be connected to both networks by using some form of multipath I/O (MPIO) solution, such that data can flow across both networks seamlessly in an active/active or active/passive mode.

MPIO ensures that if one path fails, an alternative path is readily available. Ideally, the networks are identical; however, at a minimum, they should be based on the same switch architecture to ensure consistency of performance. In some cases, these networks are in the same location. However, to provide for disaster recovery (DR), two separate locations are often used for each complete network or for sections of each network. Regardless of the physical geography, two separate networks are available for complete redundancy.

5.2.6 Switch interconnections

At least two of every element must be used in the SAN to provide redundancy and improve resiliency. The number of available ports and device locality (server/storage tiered design) determine the number of ISLs that is needed to meet performance requirements. Therefore, a minimum of two trunks, with at least two ISLs per trunk, are used. Each source switch must be connected to at least two other switches, and so on.

Each of the connection lines in Figure 5-2 on page 83 shows at least two physical cable connections.

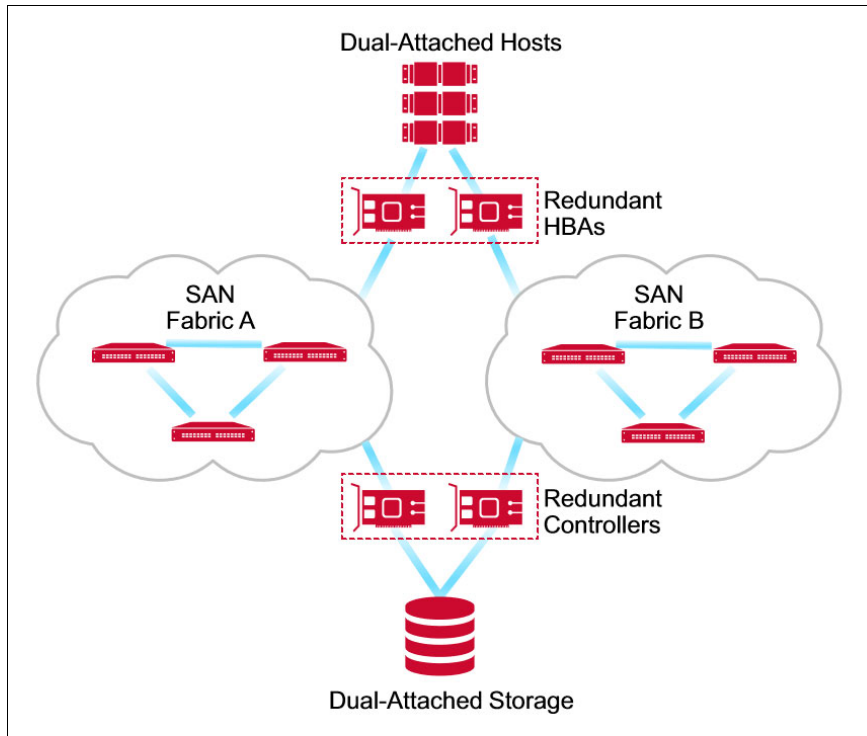


Figure 5-2 Connecting devices through redundant fabrics

In addition to redundant fabrics, redundant links are placed on different blades, different ASICs, or at least different port groups whenever possible, as shown in Figure 5-3.

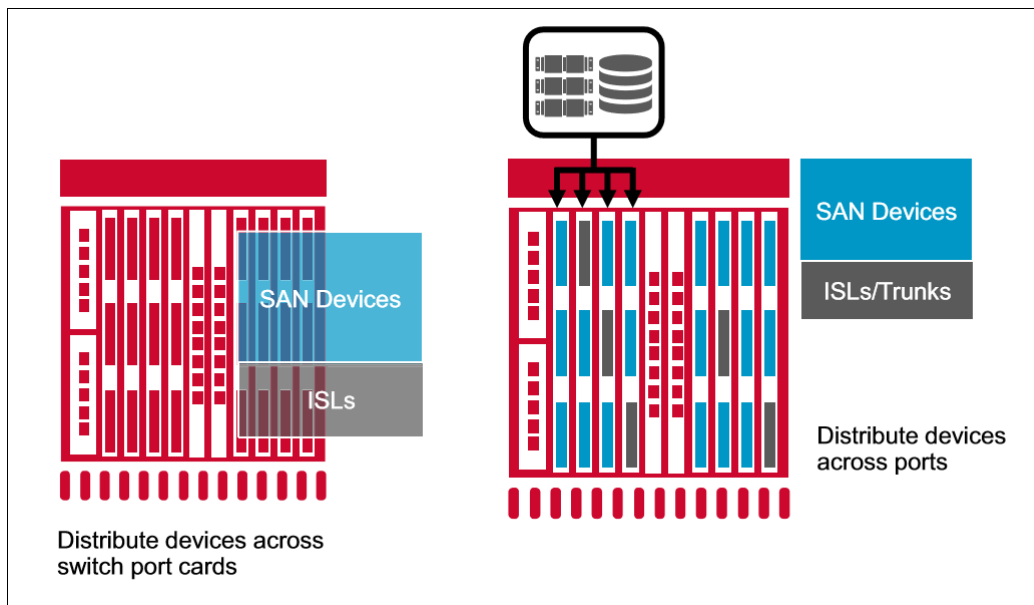


Figure 5-3 Examples of distributed ISL placement for redundancy

(For more information about determining trunk groups for the various port blades, see the suitable hardware manual or the For more information, see the [Brocade Fabric OS Administration Guide](#).)

Whatever method is used, it is important to be consistent across the fabric. For example, do not place ISLs on lower port numbers in one chassis (as shown in the left diagram in Figure 5-3 on page 83) and stagger them in another chassis (as shown in the right diagram in Figure 5-3 on page 83). Doing so results in mismatched ISL placement.

Note: In Figure 5-3, ISL trunks are placed on separate ASICs or port groups. It is important to match ISL placement between devices and across fabrics to ensure simplicity in design and assist in problem determination.

5.2.7 UltraScale ICL connectivity for IBM b-type Gen 7 platforms

The IBM b-type Gen 7 DCX 8510, X6, and X7 platforms use second-generation UltraScale ICL technology with optical QSFPs. The IBM b-type Gen 7 DCX 8510-8, X6-8, and X7-8 support up to 32 QSFP ports per chassis (see Figure 5-4), and the IBM b-type Gen 7 DCX 8510-4, X6-4, and X7-4 support up to 16 QSFP ports to help preserve director ports for connections to end devices. Each QSFP port features four independent links, each of which terminates on a different ASIC within the core blade.

Note: Consider the following points:

- ▶ X6 ICL ports can also connect to DCX 8510 ICL ports at 16-Gb speeds with the port speed on the X6 ICL configured to 16 Gb and use of the 16-Gb QSFP on the attached X6 port.
- ▶ X7 ICL ports can also connect to DCX 8510 ICL ports at 16-Gb speeds with the port speed on the X7 ICL configured to 16 Gb and use of the 16-Gb QSFP on the attached X7 port.



Figure 5-4 A 12-chassis UltraScale ICL-based core-edge design

5.2.8 Best practices for IBM b-type Gen 7 UltraScale ICL connections

Each core blade in a chassis must be connected to each of the two core blades in the destination chassis to achieve full redundancy, as shown in Figure 5-5 on page 85.

Note: For redundancy, use at least one pair of links between two core blades.

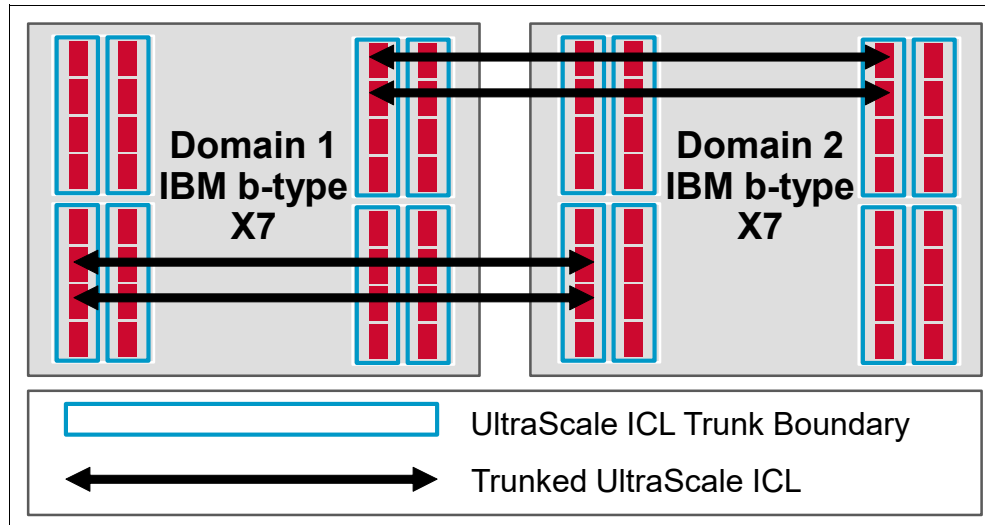


Figure 5-5 Minimum ICL connections needed between IBM b-scale X7 chassis

5.2.9 Mesh topology

A mesh design provides a single hop between source and destination. Beginning with FOS 7.3.0x, IBM supports a 9-chassis mesh design with up to 100-meter distances by using select QSFPs and OM4 fiber.

In the configuration that is shown in Figure 5-6, up to 1152 16 Gbps ports are supported by using UltraScale ICLs with a 12:1 over-subscription. As more UltraScale ICLs are added, over-subscription can be reduced to 3:1.

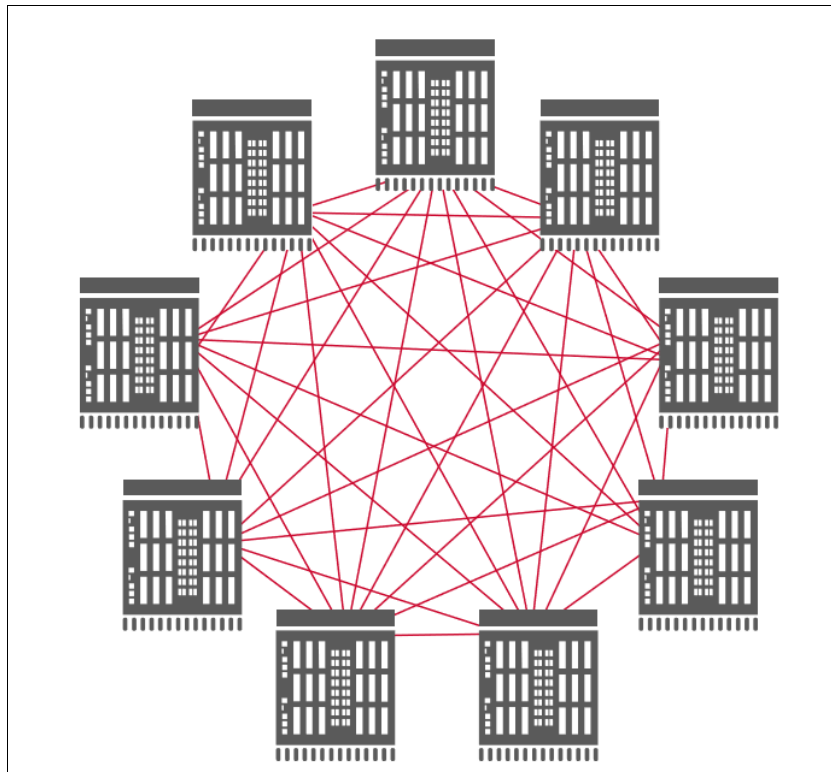


Figure 5-6 A 9-chassis UltraScale ICL-based full-mesh topology

Note: For more information, see the [Scale-Out Architecture with IBM \(Brocade\) UltraScale Inter-Chassis Links Design Guide](#).

UltraScale ICL connections are considered a “hop of no concern” in a FICON fabric. When core-edge SAN design methods are used, edge switches must connect to at least two core switches with trunks of at least two ISLs each. Each of those trunks are attached to a different blade or port group. To be fully redundant, a completely mirrored second fabric and devices must be connected to both fabrics by using MPIO.

Recommendations for switch ISL/UltraScale ICL connectivity include the following considerations:

- ▶ At least two core switches are used.
- ▶ Every edge switch has at least two trunks to each core switch.
- ▶ Small trunk groups are selected (keep trunks to two ISLs), unless you anticipate high traffic volumes. This selection ensures that you can lose a trunk member without losing ISL connectivity.
- ▶ Redundant links are placed on separate blades.
- ▶ Trunks are in a port group (ports within an ASIC boundary).
- ▶ The same cable length is used for all UltraScale ICL connections.
- ▶ Use ISL or UltraScale ICL connectivity into the same domain. Mixing the two types of connections is *not* supported.
- ▶ The same type of optics is used on both sides of the trunks: Short Wavelength (SWL), Long Wavelength (LWL), or Extended Long Wavelength (ELWL).

5.3 Implementation planning for a Fibre Channel SAN

This publication is for technical IT architects who are directly or indirectly responsible for SAN design that is based on IBM b-type Gen 6 and Gen 7 FC SAN platforms. It describes many of the challenges that face SAN designers in greenfield and established storage environments.

Although not intended as a definitive design document, this publication introduces concepts and guidelines to help you avoid potential issues that can result from poor design practices.

This section describes best practices guidelines.

5.3.1 Device placement and traffic locality

Device placement is a balance between traffic isolation, scalability, manageability, and serviceability. With the growth of virtualization and multi-node clustering on the UNIX platform, frame congestion can become a serious concern in the fabric if interoperability issues exist with the end devices.

Designing device connectivity depends largely on the expected data flow between devices. For simplicity, communicating hosts and targets can be attached to the same switch (see Figure 5-7 on page 87).

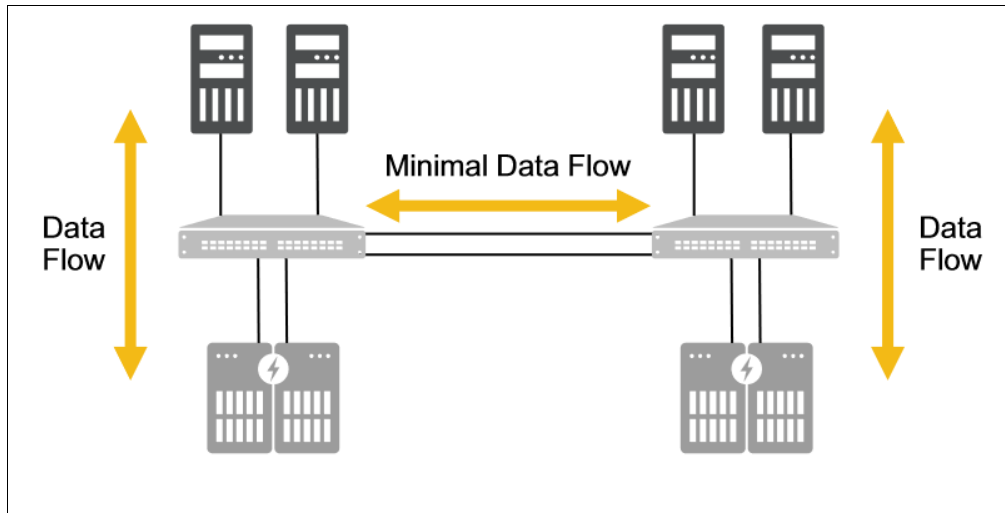


Figure 5-7 Hosts and targets attached to the same switch to maximize locality of data flow

However, this approach does not scale well. Because of the high-speed, low-latency nature of FC, attaching these host-target pairs on different switches does not mean that performance is adversely affected for common workloads.

Although traffic congestion is possible, it can be mitigated with suitable provisioning of ISLs/UltraScale ICLs. With current generation switches, locality is not required for performance or to reduce latencies.

For mission-critical applications that depend on extremely fast response times, architects might want to localize the traffic when flash storage is used or in exceptional cases, particularly if the number of ISLs available is restricted or a concern exists for resiliency in a multi-hop environment (see Figure 5-8).

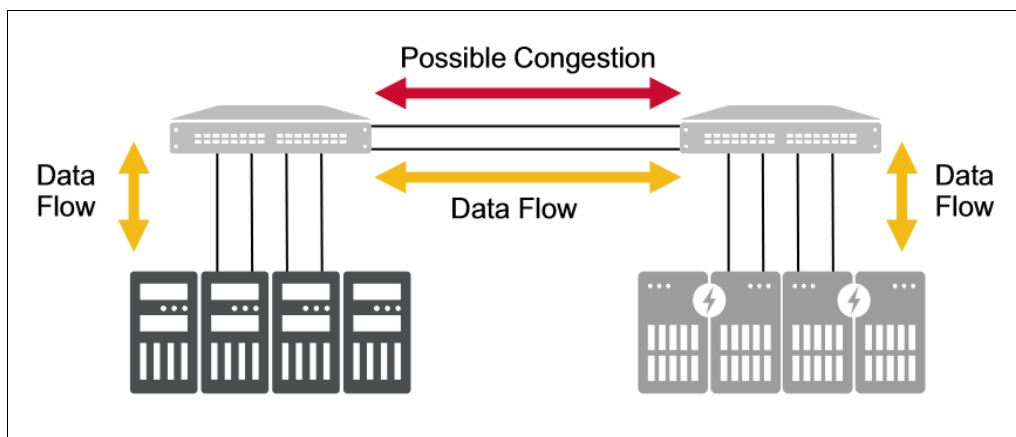


Figure 5-8 Hosts and targets attached to different switches for ease of management and expansion

One common scheme for scaling a core-edge topology is dividing the edge switches into a storage tier and a host/initiator tier. This approach lends itself to ease of management and ease of expansion. In addition, host and storage devices generally have different performance requirements, cost structures, and other factors that can be readily accommodated by placing initiators and targets in different tiers.

Planning for storage

Storage arrays have evolved dramatically over the last few years. Performance increased exponentially, capacities exploded, and more LUNs are supported than ever before. The performance and capacity of low-end arrays also improved.

Today's storage arrays can include the following characteristics:

- ▶ Tiers comprising all-flash or flash-enhanced storage media became commonplace for mid-range and enterprise-class environments.
- ▶ Some arrays time out and reset their ports if they do not receive acknowledgments from the connected host after specific intervals.
- ▶ Use of in-band FC for control purposes has increased, putting extra stress on FC port buffer usage. The use of flash-enhanced storage not only reduces the response times and increases the I/O level by up to 10x, it also increases the efficiency of the host platforms, allowing them to run at a higher level of utilization and execute more I/O commands rather than waiting on responses from slow disk media.

Adding a layer of flash can open bottlenecks and relieve congestion that result from long wait times. Still, storage array performance can degrade over time, which can be attributed to factors that include the following examples:

- ▶ Proliferation of VMs and the "I/O blender" effect, whereby I/O processes become increasingly non-sequential and increase reads and writes to disk.
- ▶ Insufficient controller cache.
- ▶ Misaligned LUN configurations where operating system block sizes do not match well to those on block storage.
- ▶ Mis-configured control values, such as edge hold time and queue depths.
- ▶ Provisioning strategies can favor capacity over usage.

Design guidelines

Consider the following design guidelines:

- ▶ Ensure that network throughput capacity matches or exceeds the capabilities of storage devices.
- ▶ Use care if you deploy mixed arrays with different performance characteristics. Experience shows that depending on how it is used, it is easy for a Tier 3 storage array to affect the performance of arrays in the same fabric. Troubleshooting in these situations is difficult.
- ▶ Limit network switch and optic speeds to two generations at most, with a best practice of uniform speed to avoid target ports from slowing to accommodate lower-speed ports and creating a back-pressure situation.
- ▶ Enable Forward Error Correction for ISL connections and connectivity to Gen 5 devices that support it. FEC is on by default for Gen 6 networks and always should be used to ensure consistency of network transit times.
- ▶ Control the number of LUNs behind each storage port based on the type of usage they receive.
- ▶ Check on any special short-frame traffic to avoid frame congestion at array ports. It might be necessary to increase the number of buffers at the array port to accommodate the extra control traffic.
- ▶ Use advance Brocade FOS threshold timers to monitor hosts and storage arrays to ensure that array ports do not reset because of a high-latency host, and therefore do not adversely affect other connected hosts.

NVMe

The next consideration in terms of why the storage network must be rethought and redesigned is Non-Volatile Memory Express (NVMe).

At a device level, consider the following characteristics:

- ▶ Density: Current NVMe devices have 8 - 10 times the density of DRAM.
- ▶ Latency: Current NVMe devices have a sub-20-microsecond latency.
- ▶ Bandwidth: Current NVMe devices use 4 PCIe Gen 3 lanes (32 Gbps).
- ▶ Streamlined software: Current NVMe software includes 13 required and 25 optional commands.

For more information about NVMe over FC fabrics, see Chapter 8, “NVMe over Fibre Channel SAN design, implementation, and migration strategies” on page 217.

Storage virtualization

Storage virtualization enables LUNs that are accessed by servers to be extracted into a logical layer of LUNs from the physical storage (typically storage arrays) on which they are stored. (These LUNs are not the same as traditional storage array LUN allocations, which can also be viewed as a form of virtualization.) Virtualized LUNs that are disassociated from their storage allow for more flexible storage provisioning processes. Performance can also improve because the virtual LUNs can be striped across multiple storage arrays.

Two general types of storage virtualization are available: one uses an external controller (called *in-line virtualization*), and in the other, the virtualization occurs inside a storage array. In-line solutions are slightly more flexible because they can use physical storage from a variety of sources and vendors.

Figure 5-9 shows a typical implementation of an in-line virtualized storage solution.

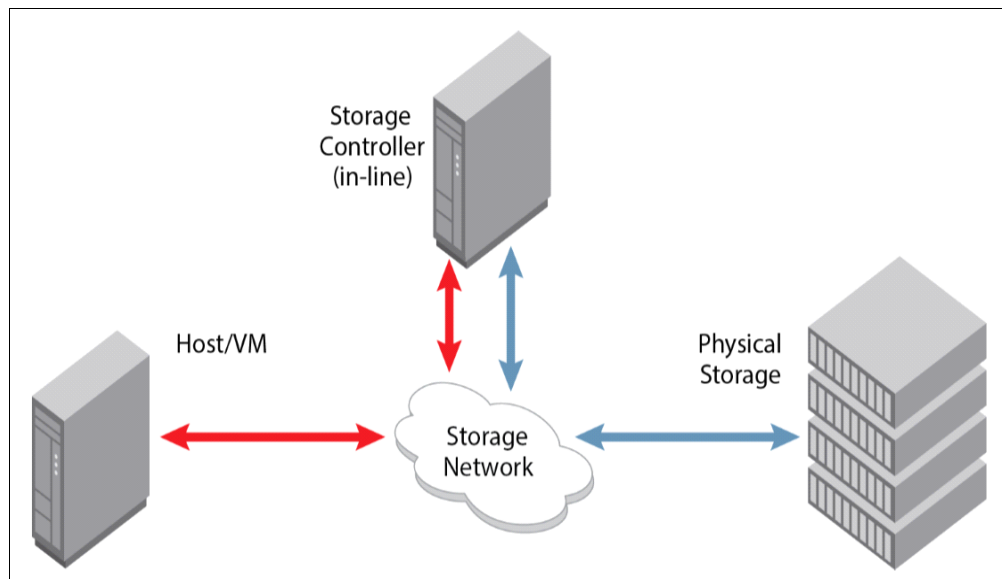


Figure 5-9 Typical implementation of an in-line virtualization storage solution

The host or VM accesses storage by way of a storage controller (shown at the top of Figure 5-9 on page 89) through the storage network. The red arrows indicate data access to and from the storage controller. The storage controller typically controls all access to the physical storage, shown on the right of Figure 5-9 on page 89 (and indicated by the blue arrows). This configuration creates a flexible storage solution because logical LUNs can be striped across several physical arrays to improve performance. Also, logical LUNs can be manipulated transparently to the host or VM.

The major benefit of this type of storage virtualization is that storage can now be provisioned in units of capacity (500 GB or 1 TB) rather than physical LUNs. This first step is toward viewing storage as a service instead of as physical units. VM provisioning now becomes less complex and easier to automate.

Design guidelines

Consider the following design guidelines:

- ▶ Each storage controller in an in-line solution serves as an initiator and a target.
- ▶ ISL use increases with in-line virtualized storage. Ensure that enough ISL bandwidth exists to manage the increased load.
- ▶ The in-line storage heads might communicate through FC or generate more SCSI control frames to manage their attached storage, which can contribute to frame congestion. You might need to increase the number of buffers at the ports that connect to the storage controller to accommodate this behavior.
- ▶ It is much more difficult to determine initiators and targets with in-line virtualized storage. Because they are on the same switch, be careful about deploying tools, such as Port Fencing.

Monitoring

Consider the following points regarding monitoring:

- ▶ Use MAPS to track thresholds of critical parameters.
- ▶ FPI monitoring is useful in determining latencies that are associated with virtualized storage.
- ▶ Use Flow Vision to review high-traffic-usage flows and identify the source of bottlenecks and congestion.
- ▶ Consider running Flow Vision constantly for the most mission-critical applications to keep a historical record of application performance profile and intermittent irregularities.
- ▶ For “frequent-caller” applications, run Flow Vision on a regular basis where time permits to verify good health.

5.3.2 Scalability and performance

IBM b-type Gen 7 products are designed with scalability in mind, knowing that most installations continue to expand and that growth is supported with few restrictions. However, following the same basic principles that were outlined in previous sections of this chapter as the network grows ensures that the levels of performance and availability continue.

Evaluate the affect on topology, data flow, workload, performance, and perhaps most importantly, redundancy and resiliency of the entire fabric any time one of the following actions is performed:

- ▶ Adding or removing:
 - Changes in provisioning
 - Storage
 - Switches
 - AISLs and ICLs
- ▶ Change in virtualization (workload and storage) strategies and traffic flow pattern if these design best practices are followed when the network is deployed, small incremental changes should not adversely affect the availability and performance of the network. However, if changes are ongoing and the fabric is not suitably evaluated and updated, performance and availability can be jeopardized.

Consider the following points when reviewing the status of a production FC network include:

- ▶ Reviewing redundancy and resilience:
 - Are there at least two physically independent paths between each source and destination pair?
 - Are there two redundant fabrics?
 - Does each host connect to two different edge switches?
 - Are edge switches connected to at least two different core switches?
 - Are interswitch connections composed of two trunks of at least two ISLs?
 - Does each storage device connect to at least two different edge switches or separate port blades?
 - Are storage ports provisioned such that every host has at least two ports through which it can access LUNs?
 - Are redundant power supplies attached to different power sources?
 - Are zoning and security policies configured to allow for patch or device failover?
- ▶ Reviewing performance requirements:
 - Host-to-storage port fan:
 - Host to ISL
 - Edge switch to core switch
 - Routing policy and currently assigned routes (evaluate actual utilization for potential imbalances)
 - Use of FEC for all ISLs and connections to Gen 5 (if supported) and Gen 6 devices.
- ▶ Watching for latencies because of the following issues:
 - Poor storage performance.
 - Overloaded hosts or applications.
 - Distance issues over constrained long-distance links, which result from changes in usage, such as adding mirroring or too many workloads.
 - Deteriorating optics resulting in declining signal strength and increased error rate.
 - In Gen 6 and Gen 7 networks, storage response latency can be base-lined and monitored continuously by using I/O Insight with MAPS. Correct latencies immediately; they can have a profound effect on the fabric.

Although IBM b-type Gen 7 SANs are designed to allow for any-to-any connectivity and they support provision-anywhere implementations, these practices can have an adverse effect on the performance and availability of the SAN if left unchecked. The network must be monitored for changes and routinely evaluated for how well it meets wanted redundancy and resiliency requirements.

5.3.3 SAN design for critical workloads

With all-flash arrays (AFAs) being the standard in enterprise data centers today and the transition to FC-NVMe AFAs beginning, critical business applications are increasingly dependent on consistent low-latency, high-throughput storage performance for demanding, performance-sensitive workloads.

When designing the SAN, it is important to consider the placement of critical workloads (in relation to the placement of storage), the fan-in ratio to storage ports, and the fan-in ratio to ISLs/trunks.

Whereas the IBM b-type Gen 7 SAN technology provides a suite of built-in measures to avoid interference of workloads that expose any kind of congestion behavior, such as Traffic Optimizer, FPIN, MAPS, and SDDQ, protecting critical workloads is crucial. Ideally, the most demanding and critical workloads include dedicated storage array target ports (or even dedicated arrays) and the shortest possible path through the SAN. The purpose for this configuration is to avoid any interference of other less critical workloads that might expose behavior that results in congestion or back pressure, which can interfere and adversely affect the performance of critical workloads.

Placement of servers with business-critical workloads

With core-edge SAN designs, it is often advantageous to connect servers with critical workloads directly on the core with the storage array ports. This configuration works well when the number of business-critical workloads is easily defined to a specific subset of servers in the SAN and enough of ports are available on the core to connect both storage and business-critical servers.

If the number of business-critical servers exceeds the number of available ports on the core for server connections, the business-critical servers must be connected on edge switches. The most common model is to use dedicated edge switches for business-critical servers and decrease the fan-in ratio of servers to ISLs.

An alternative model is to attempt to evenly distribute business-critical servers across edge switches with the assumption that the workload pressure evens out with the less critical and demanding workloads on other servers. Although approach might seem logical, experience shows that throughout the lifetime of applications and the SAN, it is operationally more complex to ensure optimal performance for business-critical workloads by using this model.

Business-critical VMs

In today's data centers, it is not uncommon for some or all business-critical workloads to be running on VMs. Combining this with the digital business environment, the value (business criticality) and performance requirements for an application often change throughout the life cycle of the application. Therefore, it can be difficult (or impossible) to predict the requirements for an application and plan placement accordingly from the beginning; thankfully, hypervisors provide the ability to move VMs between hosts without downtime and to migrate storage when necessary.

To apply the same principle of server placement of (bare metal installed) business-critical applications to VMs, deploy dedicated hypervisor clusters that are connected on the core or on high-performance edge switches. Use these hypervisor clusters for business-critical and performance-sensitive VM workloads only.

Visibility into each VM's workload on the same data store (backed by LUN or NSID) can be achieved with VM Insight. VM Insight enables storage administrators to monitor VM-level application performance and set baseline workload behavior. This information can be used to quickly determine whether a storage fabric is the source of performance anomalies for applications running on VMs. Based on VM Insight, storage administrators can provision and plan placement in the SAN on application requirements and can fine-tune the infrastructure to meet service-level objectives.

5.3.4 Capacity planning

One way to avoid performance issues and downtime is with suitable capacity planning, which helps you get an overall view of your IT infrastructure's overall capacity.

Gathering requirements

The SAN project team must interview all stakeholders (IT application owners, finance, corporate facilities, IT lab administrators, storage and network administrators, and users) who have a vested interest in the project. This interview process applies equally to planning for new and updated SANs.

Application owners

As critical stakeholders, application owners care because everyone is measured on application uptime. Application outages are something that users notice, and they can have severe financial effect on a business. With a redundant or a resilient infrastructure, hardware outages are transparent to the user, and only SAN administrators must pay attention.

To better understand their requirements, questions can be asked of the application owners including, the following examples:

- ▶ What is the business goal for this application? (Is it a database that multiple applications rely on for business transactions?)
- ▶ What are the availability requirements?
- ▶ Is the application latency sensitive?
- ▶ Are there peak periods of utilization or other traffic patterns?
- ▶ What are the IOPS requirements in terms of read/writes?
- ▶ What is the worst-case response time before an outage?
- ▶ Is the application running on a cluster?
- ▶ Is there application downtime that can be used for applying patches, software upgrades, and maintenance?
- ▶ Can the application run on a VM? If so, how many other VMs can coexist on the same physical hardware?

The business criticality of the application determines the SAN design and the DR strategy, including backup and recovery. If the application is mission critical, the infrastructure must be fully redundant, with no single point of failure for mainframe or distributed open systems architectures.

Server and storage administrators

After the application requirements are defined, identify physical server and storage on which the application and data are stored to determine the overall high-level architecture of the SAN, especially if this includes existing and new equipment.

Consider the following points:

- ▶ Gather information about the servers on which the applications are running (blade or rack, CPU, memory, HBA/embedded FC switch, operating system level, operating system patch level, or HBA driver version)?
- ▶ How many HBAs are in the rack servers?
- ▶ Does each server have single or multiple port HBAs?
- ▶ What is the wanted storage response time (latency)? Do longer latency times that result from a disk hit severely affect performance?
- ▶ For hybrid storage, how much storage is allocated to SSD?
- ▶ For HDD-based storage, what is the current cache use? Is there enough cache to meet required response times? What is the average drive use (the greater the use, the longer the response times) for HDDs and SSDs? Contact your drive vendor to identify response times based on use for sizing workloads.
- ▶ How much storage capacity is allocated to flash (more cache allows for more consistent response time) if using a hybrid array?
- ▶ Are storage tiers used in the environment? What is the policy that is used for migrating data? Are different tiers used for online storage? What is the effect?
- ▶ What is the raid level used? This factor determines available disk space and performance for the application.
- ▶ How many FC ports exist in the array?
- ▶ Are the arrays front-ended by a storage virtualization controller? If so, what is the additional latency?
- ▶ What are the recommended fan-in and fan-out ratios for the arrays that are used for this application? What are the limits?
- ▶ Does a Disaster Recovery (DR) site exist? If so, how is it connected (dark fiber, FCIP, and so on)?
- ▶ What is the available or required bandwidth between the intra-site for DR? Can the storage infrastructure support DR with the extra load?
- ▶ What tools are used for mirroring and replication (host-based or array-based)? If host-based, was the failover tested? If so, was application uptime affected? If storage-based, was the failover tested? Did the LUNs appear on the active ports?
- ▶ Was application uptime affected?

SAN administrator: General

A SAN administrator is responsible for the daily operation of the network. The SAN design must be easy to monitor, manage, and maintain.

If the current SAN is being expanded, adequate performance metrics must be collected to ensure that the existing design can be expanded to address new workloads. Consider the following points:

- ▶ Are there performance (bandwidth) or latency issues in the existing SAN?
- ▶ Are procedures in place to address redistribution of capacity when switch port use exceeds 75%?

- ▶ Is the current design two-tier (core-edge) or three-tier (edge-core-edge)?
- ▶ Is the SAN centrally managed by a tool, such as IBM Storage Insights or IBM Spectrum® Control?
- ▶ Does a SAN exist? If so, how is it managed (CLI, IBM Network Advisor, and so on)? Is there a separate network for SAN management?
- ▶ Are access control policies in place for change management (zoning)? Is there a zoning policy? Are there devices in the zone database that no longer exist? What type of zoning is used (port or WWN)?
- ▶ Is the current SAN a redundant configuration?
- ▶ Is there an identified server to capture logs from the fabric?
- ▶ Is the traffic equally distributed across the ISLs or the trunks?
- ▶ Is historical performance data available for initiators, targets, and ISLs?
- ▶ How many unused ports are available per switch?

5.4 Migration strategies of Fibre Channel SAN infrastructure

This section provides an overview of SAN migration planning and strategies. It helps you understand how a Gen 6 platform can be upgraded to a Gen 7 platform. Migrating to the Gen 7 platform helps protect the investments that you made and gives you the added benefit of newer technologies.

5.4.1 IBM b-type Gen 7 provide infrastructure for today and tomorrow

The IBM b-type Gen 7 platform's breakthrough 64 Gbps performance accelerates application response time by up to 71%, which eliminates I/O bottlenecks and unleashing the full performance of an all-flash data center.

These systems seamlessly integrate next-generation NVMe over Fabrics with Gen 7 FC networks without a disruptive "rip and replace". In addition, these platforms also seamlessly connect older generation devices and storage networking equipment with three generations of compatibility with an earlier version.

Investment protection with IBM b-type Gen 7 Fibre Channel solutions

For investment protection, the IBM b-type Gen 7 family offers three generations of backward-compatibility support for connectivity to 4, 8, and 16 Gbps FC (and FICON) products. This support enables seamless connectivity between the older generation of devices and storage networking equipment.

This connectivity enables an older storage infrastructure to continue to serve an organization's needs. An organization can also strategically plan to introduce IBM b-type Gen 7 Directors into their existing fabrics when devices are added to their fabrics, whether because of new requirements or ongoing growth.

Further protecting future investments, the IBM b-type Gen 7 Director family supports future FC generations as a Gen 7-ready storage networking platform. The IBM b-type Gen 7 backplane is designed with more links to accommodate a doubling and quadrupling of bandwidth in the future. This ability enables platform reusability for Gen 7 FC, which also enables organizations to maximize their IBM b-type Gen 7 investments and optimize the performance of their networking infrastructures.

IBM b-type Gen 7 platforms also enable customers to future-proof their storage networks by accelerating new technology deployments, such as NVMe over Fabrics. NVMe is a new protocol for solid-state storage devices that are built with nonvolatile memory technologies.

NVMe provides dramatically lower latency for storage I/O operations and significantly higher IOPS per device. In the future, organizations will be able to transition their high-performance, latency-sensitive workloads to flash-based storage with NVMe over Fabrics. By doing so, they can realize significant performance gains and reap the full benefits of flash.

NVMe over Fabrics scale up the number of devices it can address by adopting NVMe over Fabrics technology. NVMe over Fabrics enables users to achieve faster application response times from network-attached storage and harness the performance of hundreds of solid-state drives for better scalability across virtual data centers.

FC is one of the fabric technologies that are supported by NVMe over Fabrics. IBM b-type Gen 5 and Gen 6 FC switches support NVMe over Fabrics, with no changes required. Customers can seamlessly integrate IBM b-type Gen 7 FC networks with next-generation NVMe over Fabrics without a disruptive “rip and replace” for the increased performance, application response time, and scalability that their next-generation data centers require.

Upgrading a Gen 6 Director to a Gen 7 Director

IBM offers a simple upgrade option to extend the life of the Brocade Gen 6 chassis and gain the extra value of Gen 7 technology. Because the IBM b-type Gen 7 uses the same chassis, only a simple upgrade is required to transform the Gen 6 director into a Gen 7 director. Although the upgrade requires taking the chassis offline to replace the core routing blades, it is nondisruptive to the applications in a dual fabric because traffic moves to the other fabric while the upgrade is occurring.

After the Gen 6 director is upgraded, organizations can choose to keep their existing port blades and start to use the new Gen 7 blades within the same chassis.

A detailed procedure is available that provides the materials and instructions that are required to migrate your IBM b-type Gen 6 director (SAN512B-6 and SAN256B-6) running FOS 8.x or 9.x to an IBM b-type Gen 7 director (SAN512B-7 and SAN256B-7).

For more information about Gen6 to Gen7 migration procedures, contact IBM Support for the full documentation.

You see how to perform the following tasks:

- ▶ Take a snapshot of the configuration.
- ▶ Prepare the configuration for conversion of the Control Processor (CP) blades.
- ▶ Swap the Core Routing (CR) blades.
- ▶ (Optional) Restore the original customer configuration (or default configuration).

The result is a one-way upgrade to a Brocade Gen 7 director.

Note: Upgrading to a Gen 7 director is permanent on the CPs. Therefore, you cannot downgrade to Gen 6. After the upgrade to Gen 7, you cannot use the same CP blades in a Gen 6 chassis.

The Gen 7 director supports mix-and-match blades, which allows Gen 6 and Gen 7 blades to be installed within the same chassis. The following optional port or extension blades are available:

- ▶ 48-port Gen 6 IBM b-type Gen 7 FC32-X7-48 port blade with 32 Gbps line rate ports:
 - Added value over FC32-48 (reduced latency, congestion avoidance, and Traffic Optimizer)
 - Non-upgradeable to 64 Gbps speeds
- ▶ IBM b-type Gen 7 FC32-64 high-density port blade: Increases scalability and maximizes space use.
- ▶ IBM b-type Gen 7 SX6 Extension blade: Disaster recovery solutions over long distances.

Technical migration guides

The [IBM \(Brocade\) X6 Field Migration Guide](#) is available to walk through the process on the Broadcom Customer Support Portal (CSP).

Required time and items for Gen 6 to Gen 7 upgrades

As a guideline, the migration procedure might require at least 1 hour to over 2 hours, depending on the number of upgrade tasks.

You must factor in the time to complete the following tasks:

- ▶ Install power supplies.
- ▶ Install supported blades.
- ▶ Remove the Gen 6 core blades and replace them with Gen 7 core blades. Except for the core blades, all other blades that are installed in the Gen 6 chassis are still supported after the upgrade.
- ▶ The chassis must be powered off before the core blades are replaced. The power-on process takes approximately 20 - 30 minutes.
- ▶ Cabling new blades or relocating existing cables.
- ▶ The Migrations steps are documented in the [IBM \(Brocade\) X6 Field Migration Guide](#).

5.4.2 Tools to help migration planning

In this section, we provide information about tools that can assist with SAN design, planning, and migration.

Brocade SAN health

Migrations are done by using a planned process. Understanding the state of your infrastructure is a good place to begin. To start, SAN Health is run to provide users the overall health of a fabric and to capture the details of a known working environment.

Brocade SAN Health is a free tool that allows SAN administrators to securely capture, analyze, and report comprehensive information about fabrics with switches that are running FOS and Cisco MDS fabrics that are running SANOS/NXOS.

SAN Health can perform tasks, such as the following examples:

- ▶ Taking inventory of devices, switches, firmware versions, and SAN fabrics.
- ▶ Capturing and displaying historical performance data.
- ▶ Comparing zoning and switch configurations against best practices.
- ▶ Assessing performance statistics and error conditions.
- ▶ Alerting the administrator of relevant technical support bulletins (TSBs).
- ▶ Producing detailed reports (in Microsoft Excel) and diagrams (in Microsoft Visio).

The SAN Health user interface is shown in Figure 5-10.

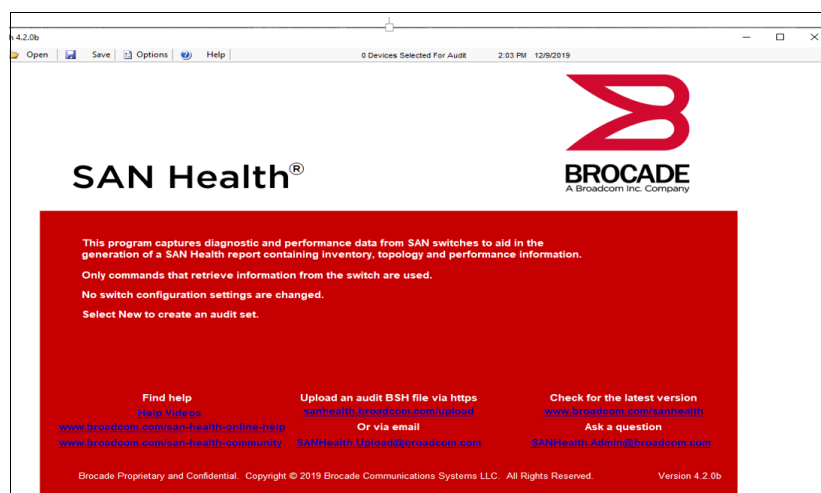


Figure 5-10 SAN Health user interface

For more information about downloading Brocade SAN Health and how to use it, see [this web page](#).

SAN Health helps to capture information to analyze the current health of your fabric. If issues with the health of your fabric are found, resolve them before moving forward with any migration. In addition to the health status, information about the physical and logical layout can be used to transition to newer SAN platforms while maintaining consistency in your design.

Brocade SAN Health data capture

SAN Health uses a data capture application and a back-end report processing engine. After capturing switch diagnostic data, it automatically generates a Visio topology diagram and a detailed report of SAN fabrics, switches, and individual ports. Other useful information includes alerts, historical performance graphs, and recommended best practices.

Because it provides a point-in-time snapshot of your SAN, the SAN Health Diagnostics Capture is invaluable to your change-tracking process.

Download Brocade SAN Health Diagnostic Capture at [this web page](#).

By using Brocade's free SAN Health tool, you can generate personalized storage network performance and inventory reports (see Figure 5-11 - Figure 5-14 on page 100) to help you prevent issues and optimize operations.

SAN SUMMARY DETAILS FOR SAN_EXAMPLE														Table Of Contents		
SWITCHES IN SAN SAN_Example																
Fabric Name	Switch Name	Domain	IP Address	World Wide Name	Model	Speed	OS Ver	Ports	Unused							
Storage_Edge	sw3200_32	32	192.168.163.32	10:00:00:60:69:c0:06:55	3200	2G	3.2.1a	8	1							
Storage_Edge	sw4100-41	41	192.168.163.41	10:00:00:05:1e:34:56:5e	4100	4G	5.1.0d	32	24							
Storage_Edge	sw3850-50	50	192.168.163.50	10:00:00:05:1e:34:12:20	3850	2G	5.0.1a	16	10							
Server_Edge	sw3800_38	38	192.168.163.38	10:00:00:60:69:50:08:7e	3800	2G	3.2.0a	16	4							
Server_Edge	sw48000-48	48	192.168.163.48	10:00:00:60:69:e4:25:18	48000	4G	5.1.0d	48	39							
Server_Edge	sw24000-24	24	192.168.163.24	10:00:00:60:69:e2:03:b0	24000	2G	5.1.0d	32	21							
Server_Edge	sw3900-39	39	192.168.163.39	10:00:00:60:69:90:0c:a3	3900	2G	5.1.0d	32	23							
HEALTH AND MONITORING STATUS FOR SAN_Example																
Fabric Name	Switch State		Power Supplies		Fans		Temp Sensors			Errors		SNMP		SysLog		
	Maing	OK	Bad	Marg	OK	Bad	Low	OK	High	Lvl1	Lvl2	No	Yes	No	Yes	
Storage_Edge	0	3	2	0	2	0	0	12	0	12	0	0	3	0	3	0
Server_Edge	3	2	1	0	7	0	0	19	0	17	0	0	5	0	5	0
TOTALS	4	5	4	0	10	0	0	34	0	35	0	0	9	0	9	0
PORT USE																
Fabric Name	Port Use					Fan Out Ratios				Port Long Distance Modes						
	Disk	Tape	Host	ISL	Free	Total	Host:Disk	Port:ISL	Device:ISL	10km	25km	50km	100km	Auto		
Storage_Edge	5	0	0	16	35	56	0.5	2.5:1	0.31:1	56	0	0	0	0		
ZONING METRICS																
Fabric Name	Zone	Aliases Statistics				Zone Statistics				Config Statistics						
	Database Use	Aliases	AvlMem	MaxMem	Hanging	Zones	AvlMem	MaxMem	Hanging	Configs	AvlMem	MaxMem	Hanging			
Storage_Edge	0.8% of 258k	24	1	1	19	11	4.5	15	1	1	10	10	1			
Server_Edge	0.9% of 258k	30	1	1	11	11	4.9	20	1	1	11	11	1			
TOTALS		54	0.7	1	30	22	3.1	20	2	2	7	11	2			

Figure 5-11 Comprehensive reporting with health and best practice checks

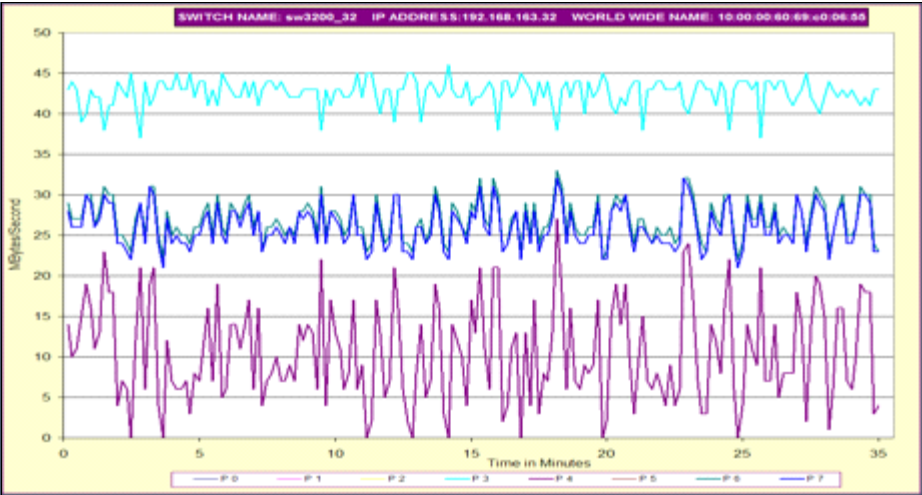


Figure 5-12 Performance graphs for visualizing activity

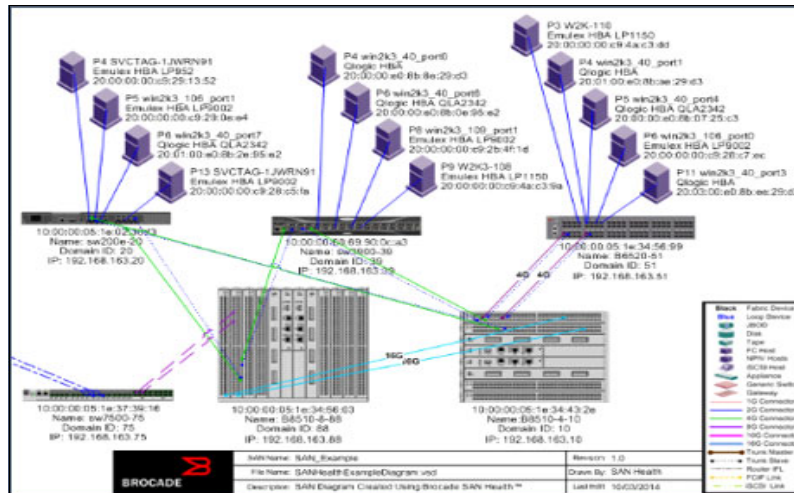


Figure 5-13 Useful topology diagrams to view physical layout

 DeviceSummary.csv	 PortCounters.csv	 SwitchPortUsage.csv	 ZoneInfo.csv
 FabricSummary.csv	 SwitchFRUs.csv	 SwitchSummary.csv	 Summary.xml
 PortConfig.csv	 SwitchSLS.csv	 ZoneAliasInfo.csv	

Figure 5-14 Search, filter filter, and sort the audit data

5.4.3 SAN migration factors

Today's SAN administrators are faced with the need for more storage capacity and speed. They require high performance and redundant SAN Networks that can meet their demands and scale for growth.

To accommodate these new requirements, SAN administrators often must migrate from and upgrade their storage network. Migrating from older SANs to newer SAN technology requires a plan that includes design, configuration, and implementation processes along with a post migration analysis.

This section describes those migration factors that must be considered for the migration plan. The use of SAN Health allows you to gather information to help make the decisions for a smooth and seamless transition.

In addition to knowing about the infrastructure, the SAN design must be reviewed to help incorporate future growth, backup and recovery or business continuity projects, and server or storage upgrades.

Many factors must be considered, including the following examples:

- ▶ Assess the fabric topology and take inventory of the network.
- ▶ Application failover considerations.
- ▶ Storage failover considerations.
- ▶ Review topology changes.
- ▶ Identify critical servers, storage arrays, and applications.
- ▶ Develop a zone configuration export and import strategy.
- ▶ Ensure that correct licensing sets are included.
- ▶ Plan for software interoperability.
- ▶ Prepare and update SAN or storage diagram to meet new requirements.

5.4.4 SAN migration process overview

This section provides a high-level overview of the steps are needed to migrate a SAN infrastructure to a new SAN infrastructure. This information is not a “how to guide”; instead, it provides the concepts that a user should plan for in their migration strategy. The scope of this migration reviews replacing an infrastructure with a one for one replacement of SAN Directors.

Preparing for the migration

There are several steps to a migration, but one of the most important steps is to capture the information about your infrastructure so that a baseline of success can be established.

The following guidelines provide a path to a new SAN infrastructure:

- ▶ Preliminary work: Run the IBM b-type Gen 7 SAN Health tool to capture a snapshot of your environment and determine whether you have any current “network health” problems that need attention before the migration (see Figure 5-15 for migration example).

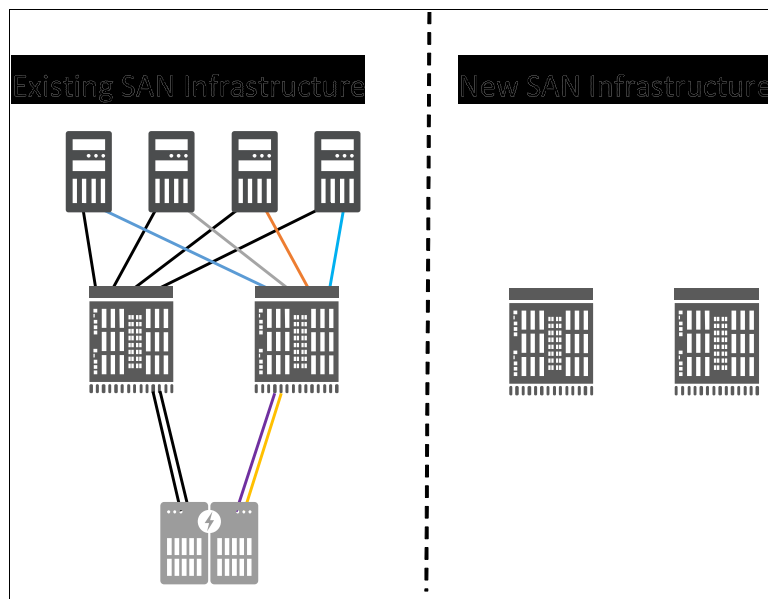


Figure 5-15 SAN Director migration example (setting up new SAN)

- ▶ Install and apply power to the new IBM b-type Gen 7 SAN Directors, which leaves them disconnected from all storage or network cables.
- ▶ Gather information from the IBM b-type Gen 7 SAN Directors that are required for building and configuring the infrastructure.

- ▶ Use SAN Health to model the new SAN infrastructure topology:
 - Modeling is done by using the Visio diagram of the SAN infrastructure that was created by SAN Health.
 - Rearrange the switches as needed. Insert newly proposed server and arrays into the SAN Health report; then, optimize the configuration for any-to-any connectivity.
 - If flaws exist in the overall design of your data fabric, the topology diagram helps bring them to the forefront, which makes it easier for the infrastructure design team to improve on designs.
- ▶ Configure the IBM b-type Gen 7 SAN Directors:
 - Create baseline configurations for all Directors and Switches.
 - Import or purge zoning sets.
 - Create zones for new devices.
- ▶ Validate the configuration

Run SAN Health and address any errors that are found before moving forward with the migration.
- ▶ Attach all cables to server and storage as required, as shown in Figure 5-16.

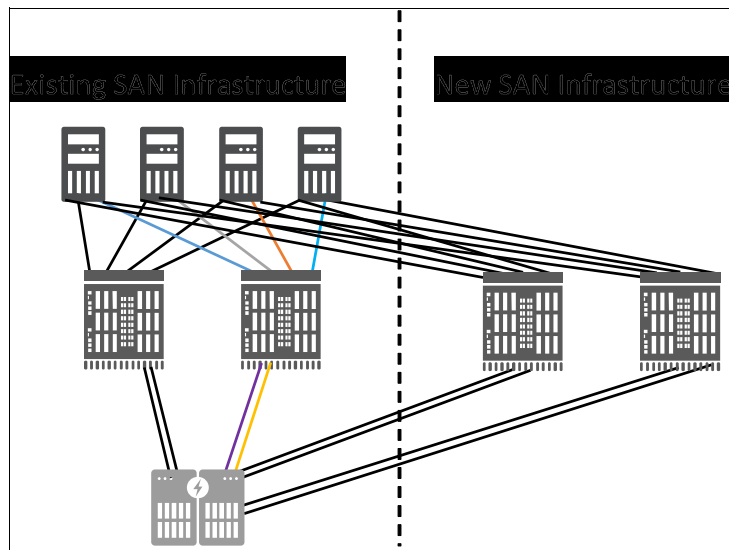


Figure 5-16 SAN Director migration example (integrating new SAN Directors/Switches)

- ▶ Validate application operations.
- ▶ Back up the new SAN configurations.
- ▶ Retire old SAN infrastructure (see Figure 5-17 on page 103).

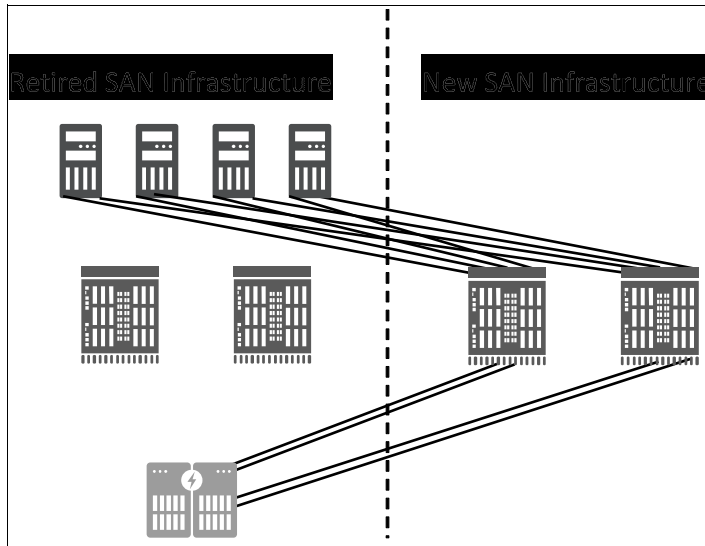


Figure 5-17 SAN director migration example (retiring old SAN)

Ongoing infrastructure management by using SAN health

Savvy IT leaders learned the value of running SAN Health regularly, especially before and after any hardware or software changes in the storage infrastructure. Beyond the insights into the architecture and performance, they find that the information that is gained increases their confidence and understanding of the SAN infrastructure and documents the fabric, including the attached servers and storage devices.

SAN Health improves capacity planning and productivity, and reduces the time that is required to troubleshoot problems when they arise. Consider adding SAN Health to your toolkit because it is free of charge.

5.5 References

Software and Hardware Product FOS documentation is available in two locations within Broadcom for IBM b-type Gen 7 products and software. The publicly facing content is available at [Broadcom.com](https://www.broadcom.com). The nonpublic documentation is on the Broadcom Customer Support Portal (CSP).

For more information about your specific Fabric operating system release, see the following Broadcom web pages:

- ▶ [Fibre Channel Networking](#)
- ▶ [Brocade X7 Directors](#)

These web pages include all user guides, reference manuals, white papers, ebook, product briefs, administrative guides, compatibility guides, and case studies, including the following publications:

- ▶ *Brocade Fabric OS Administration Guide*
- ▶ *Brocade Fabric OS Command Reference Manual*
- ▶ *Brocade Fabric OS MAPS User Guide*
- ▶ *Brocade Fabric OS Flow Vision User Guide*
- ▶ *Brocade Fabric OS Access Gateway User Guide*

- ▶ *Brocade Fabric OS Extension User Guide*
- ▶ *Brocade X6-4 Director Hardware Installation Guide*
- ▶ *Brocade X6-8 Director Hardware Installation Guide*
- ▶ *Brocade X7-4 Director Hardware Installation Guide*
- ▶ *Brocade X7-8 Director Hardware Installation Guide*
- ▶ *SAN Fabric Resiliency and Administration Best Practices User Guide*
- ▶ *Brocade SANnav Flow Management User Guide*
- ▶ *Brocade SANnav Management Portal Installation and Migration Guide*
- ▶ *Brocade SANnav Management Portal REST API and Northbound Streaming Reference Manual*
- ▶ *Brocade SANnav Management Portal User Guide*
- ▶ *Brocade SANnav Global View User Guide*



FICON design, implementation, and migration

Mainframes are typically used for mission critical applications. To ensure that all components in a mainframe SAN meet the reliability standards and consistent behavior that most mission critical applications demand, more testing and qualification programs were implemented for all devices that are connected to an IBM Z® processor.

Because of test limitations, not all Fibre Channel (FC) configurations are supported in IBM Z environments. For more information, see [this web page](#) (log in required).

FICON is a protocol that is transported over FC. In addition to meeting FICON requirements, all FC requirements also must be met. Because supported FICON configurations are a subset of supported FC configurations, you must consider only cascaded environments.

For more information about enterprises with cascaded environments, see 5.2.6, “Switch interconnections” on page 82, and 5.2.7, “UltraScale ICL connectivity for IBM b-type Gen 7 platforms” on page 84, which discuss Inter-Switch Links (ISLs) and Inter-Chassis Links (ICLs).

Most SAN upgrades skip a generation. Therefore, most upgrades to Gen 7 are from Gen 5. The most significant change when upgrading from Gen 5 is that a 32-port card is no longer supported on the director class (bladed) switches. The new blades are 48 ports. Also, the control processor (CP) cards were moved, which resulted in new slot numbering on the bladed (director class) switches.

Gen 5 and Gen 7 IBM b-type SAN Directors feature the following significant differences that must be considered for a Gen 5 to Gen 7 migration:

- ▶ CHPID path mapping
- ▶ Documentation templates
- ▶ Patch room layout

IBM Network Advisor is nearing end of support and is being replaced by SANnav. IBM Network Advisor does not support Fabric OS (FOS) 9.0 or any Gen 7 product. SANnav supports Gen 5 products on the 7.4 and 8.2 code streams. To minimize change activity when upgrading to Gen 7, upgrading to SANnav first is recommended.

Although Gen 7 offers data center managers, higher performance, and valuable new analytics, none of these features or any other Gen 7 feature affects FICON planning and implementation.

This chapter includes the following topics:

- ▶ 6.1, “Important support notes for Gen 7 platforms” on page 106
- ▶ 6.2, “FICON SAN design” on page 109
- ▶ 6.3, “FICON SAN implementation” on page 117
- ▶ 6.4, “FICON SAN migration” on page 134

6.1 Important support notes for Gen 7 platforms

Remember to plan the timing of your upgrade projects to ensure adequate time before the end of support dates.

6.1.1 Product support dates

Table 6-1 FICON support summary

Product	EOS	Comments
IBM z15™ on Gen 5	31 Oct 2021	Although support for these hardware platforms continues through 30 April 2025, the last day of support is 31 Oct 2021 for z15 connections.
Gen 5	30 April 2025	These are the SAN768B-2 (2499-816), SAN384B-2 (2499-416) and SAN48B-5 (2498-F48) products.
IBM Network Advisor	8 Feb 2023	IBM Network Advisor FOS support stops at FOS v8.2.x SANnav is the replacement product. Most customers plan the management system upgrade to coincide a few months before the end of their IBM Network Advisor support contract.

6.1.2 Software

The last code stream that is supported on Gen 5 is FOS 8.x. Gen 6 supports the FOS 8.x and 9.x code streams.

All SAN switches in a fabric must be no more than one major code revision apart. All switches that are operating with FOS 8.x and 9.x can participate in the same fabric, but switches that are operating with FOS 7.x cannot join the same fabric that includes switches that are operating with FOS 9.x.

Deprecating the prohibit or allow matrix (which is known as *PDCM*) is planned for FOS v9.1. Customers that use the PDCM should plan on considering other methods (typically zoning) to affect the blocking of specific paths. The PDCM is being deprecated because all IBM software products that used the PDCM are no longer supported by IBM.

All Gen 7 switches require FOS v9.0.0 or higher.

FICON logical switch

Configuring virtual fabrics (VF) is analogous to logical partitions (LPARs) on a IBM Z CEC. Mainframe customers often disable logical fabrics when configuring switches or leave VF enabled but use the default logical switch and configure the switch to meet the FICON requirements. With FOS 9.0 and above, having this feature enabled is required, and you must create a logical switch that is configured specifically for FICON.

The FICON logical switch was introduced in FOS 8.1.0. The difference between a logical switch that is defined for FICON and configured for FICON is that the logical switch that is configured for FICON can be reconfigured such that it no longer meets the FICON requirements. A logical switch that is defined as a FICON logical switch cannot be reconfigured to remove the FICON requirements.

A FICON logical switch provides the following advantages:

- ▶ Improves security
- ▶ Ensures properly configured FICON SAN fabrics
- ▶ Simplifies configuration

In earlier versions of FOS, determining whether the port address was bound was challenging for non-expert users. A “bound” address means that the FC address for the port cannot change. This is important because the IBM Z channel subsystem builds FC addresses based on the link addresses that are defined in the Hardware Configuration Definition (HCD) tool.

Another reason for the FICON logical switch is that it is possible to configure a switch to meet all of the FICON requirements and then back out specific changes after the channels come online. Because a mainframe channel checks only the suitable security attributes at login (when the PCHID is enabled, not when the CHPIDs are configured online), there is the potential for a channel not to come online the next time it goes off and back online. For most customers, this is the result of a CEC IML or channel path maintenance.

In addition to addressing potential problems, several addressing mode restrictions depend on the switch type, which causes considerable field confusion and in some cases, limits options for customers.

Although circumstances for changing the FC address that is assigned to a port on the switch is unlikely, if it does occur, the IODF is no longer valid, which results in device or channel errors. This potential problem is fixed with the defined FICON logical switch.

Control Unit Port in a logical switch

An independent channel that is connected to the default switch with a Control Unit Port (CUP) defined is not necessary. Chassis-related issues are reported on all logical switches through the CUP.

Other than common chassis components, each logical switch behaves as independent hardware. No information is shared between one logical switch and another. When tools are used, such as System Automation or Disk Magic, to manage and monitor switches through CUP, a separate channel with CUP defined must be attached to each logical switch.

If a separate logical fabric is created for disk or tape mirroring is put into its own logical fabric for mirroring, a separate channel is required to monitor mirroring ports. However, because that CHPID is a FICON channel, the logical switch must be defined as a FICON logical switch. Defining a logical switch as a FICON logical does not affect other protocols, such as FCP, which is used for mirroring. The FICON logical switch only adds secure features and supports CUP.

Since TS7700, tape mirroring when done with FC and Metro Mirror use FCP protocol, which means a dedicated FICON channel just for CUP is required for CUP support. An alternative, keep mirroring ports in the same logical switch with FICON CHPID to control unit paths and use zoning to isolate the traffic. This configuration avoids the need for dedicating a channel for CUP only.

6.1.3 Link address refresher

Link addressing did not change; however, because of the change from 32 to 48 port cards, link addresses land on different port cards. As a result, much of the design work is dedicated to CHPID path mapping, which requires an in-depth understanding of how link addresses relate to FC addressing (see Figure 6-1).

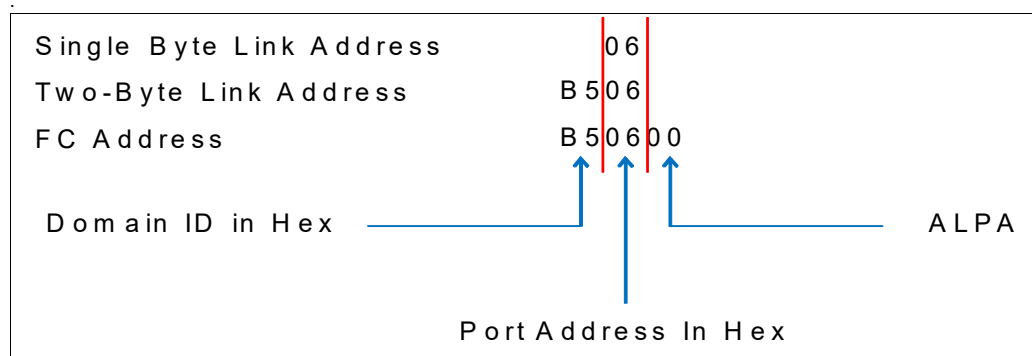


Figure 6-1 Link address and FC address

Domain ID and fabric security

For single-byte addressing, when the channel logs in, the domain area that is returned from the switch is used for the domain area in the FC address. Therefore, the channel and control unit must be in the same logical switch. With 2-byte link addressing, the most significant byte of the link address is used for the domain area of the FC address.

After 2-byte link addressing is defined for a channel, all link addresses for that channel must use 2-byte link addressing. After 2-byte link addresses are defined (whether cascading or not), all the fabric security settings must be set on the switch. For FOS 9.0 and above (and all Gen 7 switches), the channel must be connected to a FICON logical switch.

If the high integrity settings are not configured on the switch, zOS sets the channel offline and moves it to the “invalid attach” state. After a channel is in the invalid attach state, the misconfiguration on the switch must be correct and the channel re-enabled. It is the PCHID, *not* the CHPID, that is in the invalid attach state. You cannot configure the CHPID offline and back online. The PCHID must be toggled from the System Element (SE).

Because the fabric security features (HIF) are not checked by the channel when single-single-byte addressing is used, it is not recommended. Configuring 2-byte link addresses is the most secure method and required in some cases.

Port address

The hexadecimal port address is the single byte link address or the least significant byte of a 2-byte link address.

The ALPA was originally used in FC for arbitrated loop devices. Arbitrated loop is no longer supported and was never supported for FICON. Today, the ALPA is used for an extension of the port addressing and Node Port Identification Virtualization (NPIV).

The bits for NPIV allow multiple WWNs to log in to the same switch port. These bits determine to which logical entity frames belong. This is how virtual servers that use Linux, which use FCP protocol, can share a channel. FICON does not use these bits; however, the ALPA must be the same for all FICON-attached control units and channels.

Because the eight-lot director class products can exceed 256 ports in a single domain (logical switch), the upper 2 bits of the ALPA are used in specific addressing modes.

A FICON channel builds the ALPA portion of the FC address by using the same ALPA where the channel logged in. Therefore, the ALPA portion of the FC address of control units must have the same ALPA as the channel. This is why specific addressing modes do not work for FICON. In a logical switch that is defined for FICON, the ALPA is always 00.

6.2 FICON SAN design

Before starting a design project, review 6.4, “FICON SAN migration” on page 134 for more information about the switch differences and to determine how those differences affect the SAN design. This publication focuses on the SAN, but a redesign of the cable plant is likely and possibly patch room redesign.

Often, storage and processor upgrades are aligned with SAN upgrades. Make sure that all SAN requirements are well known before you begin any design planning. Also, remember to account for swing ports when upgrading storage and processors.

Consider the following areas as you design your FICON:

- ▶ Management:
 - Although FICON SANs can be managed by using the command-line interface (CLI), doing so is unusual. Nearly all FICON SANs use management software.
 - Network Advisor does not support FOS 9.x. Therefore, Gen 7 switches are not supported by Network Advisor.
 - The SANnav Management Portal is the replacement product for Network Advisor. The enterprise version is required whenever managing a director class product or environments with more than 600 ports. Otherwise, the base version is adequate.
 - FICON environments nearly always require the Enterprise version of SANnav.
 - SANnav 2.1 or higher is required when managing Gen 7 products.
 - SANnav 2.2 or higher is recommended for managing FICON fabrics.
 - SANnav Global View is a new software package that can be useful when managing multiple independent production sites or active-active backup sites. SANnav Global View provides a single pane of glass management for up to 20 instances of the SANnav Management Portal.
 - Typically, customers plan their upgrade from Network Advisor to SANnav to coincide with the end of their Network Advisor support contract. It is highly recommended to start the transition 90 days in advance. A free of charge, fully functional 90-day trial version of SANnav is available that customers can use for this purpose.

- ▶ Routing method:
 - All switches in a fabric must use the same routing method; therefore, an outage must be planned if the routing method is changed.
 - Confirm whether all control units in your environment support FiDR.
 - Exchange-based routing is the most efficient and resilient routing method; therefore, if FiDR is supported on all channels and control units, exchange-based routing is recommended.
- ▶ Use 48 port cards and high availability:

Host Channel Definition (HCD) changes and cable plant layout must be considered during a SAN migration.

Most FICON SANs use director class products with 32 port blades. Furthermore, the CHPID mapping tool and other tools that are used to aid in channel path planning are built around 32 port cards. CHPID mapping that is based on 32 port cards likely does not work with 48 port cards.
- ▶ Cascaded performance:
 - HCD changes and cable plant layout must be considered during a SAN migration.
 - Virtual channels always were a part of Brocade's FC switching architecture. In cascaded environments, the virtual channel that is used for a path is determined by the port address, which is often referred to as the *port ID* (or PID). The placement of where control units are connected can affect channel I/O rate performance.
 - Multiple zones with different quality of service can be deployed. This issue is an especially important consideration for cascaded channel-to-channel (CTC) traffic.
- ▶ Ultra low latency:
 - HCD changes and cable plant layout must be considered during a SAN migration.
 - Brocade FC switches are built on ASICs that support local switching. *Local switching* refers to when the input and output are on the same ASIC. For most applications, the added design considerations to take advantage of local switching is not necessary. Applications that demand low latency are high transaction rate processing, such as high frequency trading.
- ▶ Naming convention:
 - Before the mandatory FICON logical switch is used for a Gen 7 platform, the difference between a chassis, switch, and fabric were subtle and of little or no consequence to mainframe administrators. Although the technical differences mean little to a mainframe administrator, there are times when you must differentiate between a chassis and a logical switch, especially when creating multiple logical switches in the same chassis. Similarly, when cascading, you must differentiate the logical switches from the fabric.
 - Naming conventions are preferences that vary for each organization. The only recommendation here is that you make each name unique and easily identifiable. This can be done by appending `_fab` at the end of fabric names and `_chassis` at the end of chassis names.
 - Chassis, fabric, and switch names can be as long as 64 characters. Packing too much information into a name can make it difficult to easily identify the item it is associated with and take up too much screen space. The SANnav database has a description field where information, such as rack and aisle location, can be added so there is no need for overly complex names. Although there is no specific recommendation for naming conventions, the general advice is to keep names simple.

- ▶ Domain ID and Switch ID:
 - HCD changes and cable plant layout must be considered during a SAN migration.
 - The recommended best practice is to use the hex value of the domain ID of the switch to be the Switch ID in HCD. This recommendation also implies that all FICON logical switches have a unique domain ID and that the Switch ID is always two characters. Valid Domain IDs are 0x01 – 0xEF.
- ▶ Supported hardware and software:
 - Complete an inventory check to ensure that all hardware and software is supported and that interoperability between new hardware and software that is introduced to the fabric is supported in the environment.
 - The recommended best practice is to complete all required changes to the environment before other upgrades are made to compartmentalize changes.
- ▶ Implementation items

Review 6.3, “FICON SAN implementation” on page 117, especially user configurable items.

6.2.1 Typical redundant pathing

The typical configuration uses 8-way pathing (see Figure 6-2) to DASD spread across four directors. Although two paths are connected to the same director, the paths should be planned such that the channel and control units that are associated with a path do not share port cards.

A common design approach is to plan the first path group to use the cards to the left of the core blades, which are in the center of the director, and the second path group to use the cards to the right of the core blades.

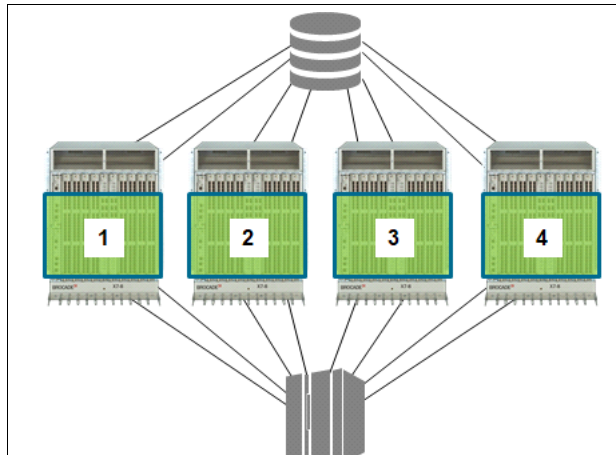


Figure 6-2 Typical 8-way pathing

Despite this separation, most FICON implementations that use directors with 32 port cards place both channel paths on the same switch. Therefore, both channel paths used the same domain ID so the high byte of the link addresses was the same for both paths. Only the lower byte of the link address was different.

6.2.3 Ultra-low latency

Planning for ultra-low latency takes more planning; therefore, it is done only when the application demands the highest speed switching with the lowest latency.

Brocade ASICs are designed to use cut-through routing. When the header of a frame is read, the frame is forwarded to the next port. Local switching occurs when the channel and control unit are connected to ports that share the ASIC.

The switching latency within the Gen 7 ASIC is 460 ns. This latency is additive when traffic is traversing multiple ASICs. All ASIC to ASIC switching is performed by way of a connection through the core blade, even if the ASICs are on the same port card. This process involves another ASIC and backplane traffic (see Figure 6-4).

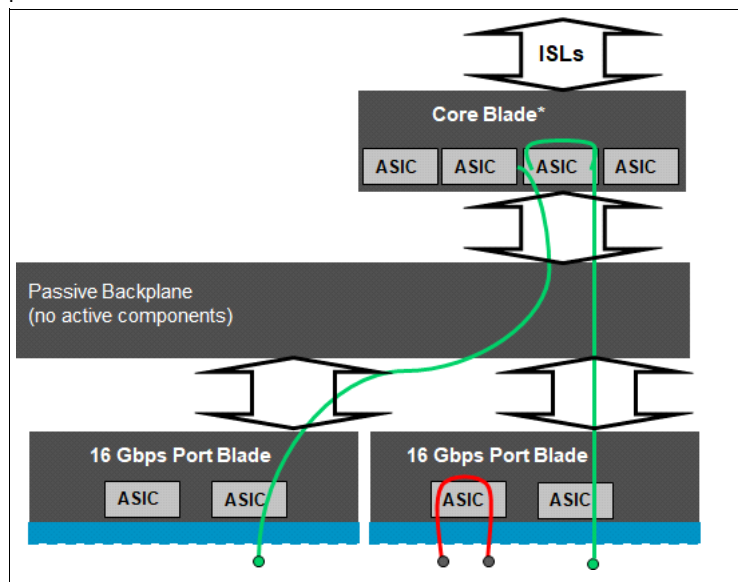


Figure 6-4 Local switching versus switching through the backplane

6.2.4 Channel paths: 32 versus 48 port cards

Because 48 ports are used per blade instead of 32, channel paths that were once planned to land on the same FRU (blade) might no longer land on the same FRU. Conversely, channel paths that use different FRUs might now be on the same FRU.

Table 6-2 Chassis configuration options with 48 port cards

Option	Considerations
8-slot chassis, one FICON logical switch	▶ Although this configuration is not recommended, it serves as a useful discussion about how 48 port cards alter the chassis configuration. ▶ A logical switch can be created on each chassis by using only the lower 16 ports on each column of ports per blade, which effectively makes it a 32-port card. Most customers who use this order some empty blades and move SFPs from the upper unused ports. Instead of leaving the upper 16 ports vacant, they can be put into another logical switch; however, few (if any) mainframe customers have a use for such a logical switch. ▶ An 8-slot chassis can also be ordered with only five cards for a total of 240 ports or six cards for a total of 288 ports, 32 of which are left unused or used for another purposes, such as mirroring.
4-slot Director	The maximum port count is limited to 192 ports. However, if FRUs were considered in the channel path planning, a new CHIPD map likely is required.
8-slot Director, 2 FICON logical switches per chassis	▶ Creating two FICON logical switches per chassis allows for the greatest density. Instead of two paths per switch in each chassis, use one path per logical switch with two logical switches per chassis. ▶ Instead of 254 usable ports per chassis, as many as 384 usable ports can be available per chassis, which results in greater than 50% increase in port density for the same foot print. If a fully populated director is not necessary, this approach can still be used. The recommended layout is to start populating the first logical switch with port cards on the left half of the chassis and populate the right side of the switch with port cards for the second logical switch. This layout allows for symmetric growth.

6.2.6 Two logical FICON switches per chassis

Although the only aspect that is new in logical switching is the defined FICON logical switch, for those users who are new to the idea of the use of logical switches, logical switching provides some extra options to fully use all ports of a 48 port card.

For FICON environments, when one logical switch is used per chassis, the maximum number of ports per logical switch is 256 ports (254 usable ports if CUP is enabled). Because 256 is not evenly divisible by 48, five port cards (which yields a total of 240 ports) or six port cards (which yields a total of 288 ports) can be used.

In addition to losing port count with only five port cards, the use of an odd number of port cards creates some planning challenges with redundant pathing. Using six port cards wastes 32 ports. In either case, two or three slots cannot be used for FICON unless a second FICON logical switch is in the chassis.

Although partitioning a chassis into two symmetrical logical FICON switches results in a maximum logical switch size of 192 ports, by splitting redundant paths to their own logical switch, another 64 link addresses are available per path group. With two logical FICON switches per chassis, this results in a total of 128 more link addresses that are available per chassis.

When CUP is enabled, link addresses xxFE and xxFF are logical addresses. If no ports are associated with those addresses, those ports must be disabled. Because link addresses xxFE and xxFF do not need to be associated in a 192 port switch, those physical ports are also available, which results in a slightly better than 50% port density improvement with this approach.

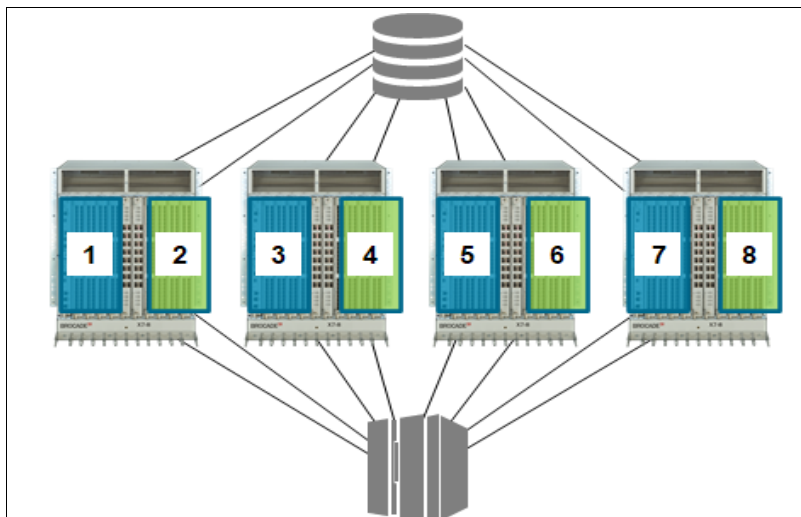


Figure 6-6 Eight logical switches for 8-way pathing option to fully use all ports

6.2.7 ISL B/W sharing with logical switches

Logical switches can be used for several other purposes (see Figure 6-7 on page 117). The idea of the base switch allows multiple logical fabrics to share ISLs. When a base switch is used for transport, these switches are referred to as *XISLs*.

XISLs feature the following limitations:

- ▶ They are not visible through CUP.
- ▶ System automation tools cannot be used to manage or monitor the links.
- ▶ RMF or Disk Magic cannot access the 74-7 records on these links.
- ▶ The quality of service (QoS) feature cannot be used.
- ▶ Traffic Isolation Zones cannot be used.
- ▶ More bandwidth planning and congestion management considerations are required.
- ▶ Over subscription of XISL B/W in one logical switch affects available B/W for other logical switches.

Complete the following steps:

1. In the upper right, click the person icon. In the drop-down menu, select **User Preferences**, as shown in Figure 6-8.

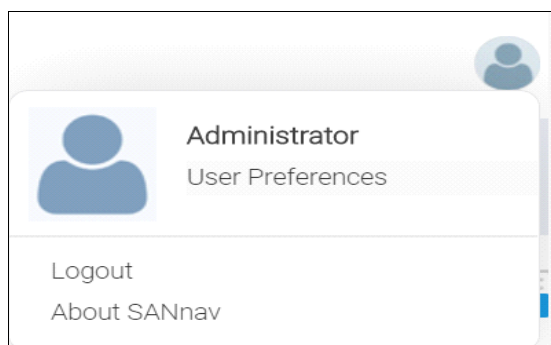


Figure 6-8 Configuring user preferences

2. Scroll to the bottom of the next window and find FICON Display under the Tables section. If the FICON Display is Disabled, click **Edit**, as shown in Figure 6-9.

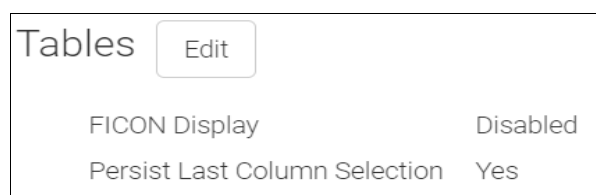


Figure 6-9 Editing FICON Display

3. Select the **FICON Display** option. Most users select **Yes** for Persist Last Column Selection. When set to Yes, display customization persists after logging out and logging back in. When done, click **Save**, as shown in Figure 6-10.

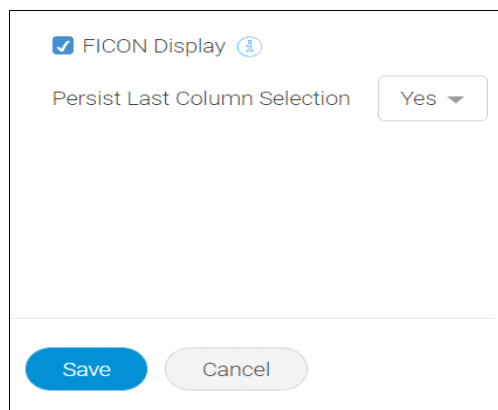


Figure 6-10 Selecting user preferences

6.3.3 Configuring FICON fabric

Even when configuring a single switch fabric, it is a FICON fabric that is configured. The individual switches are part of the fabric and take on the properties of the fabric, except for individual switch parameters, such as the switch domain ID.

Complete the following steps to configure FICON fabric:

1. Navigate to the SAN configuration window.
2. Click **SANnav**. Click **SAN Configuration**. Then, click **Logical Fabric Management** (see Figure 6-11).

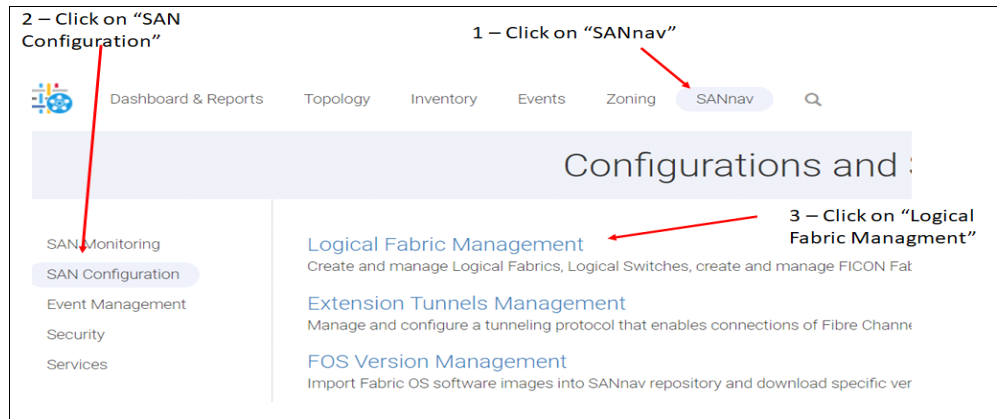


Figure 6-11 Switch configuration

3. Create a FICON fabric (see Figure 6-12):
 - a. Click **FICON Fabrics**.
 - b. Click **+** to create a new FICON fabric.

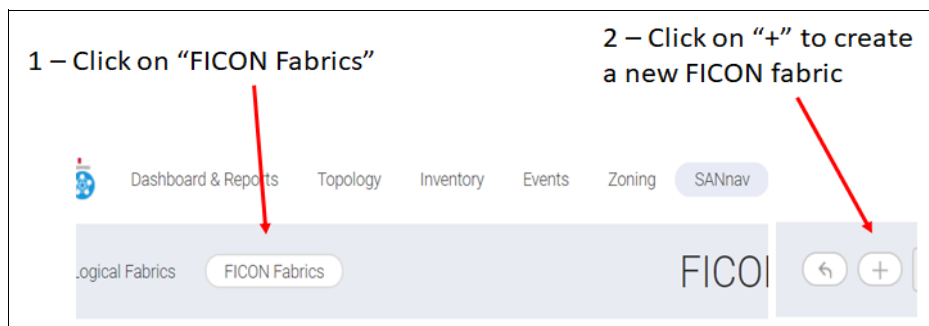


Figure 6-12 Selecting FICON fabrics

4. Configure fabric parameters (see Figure 6-13):

- a. Enter a name for the fabric.
- b. Enter a description for the fabric (optional).
- c. Enter an FID number for the fabric.
- d. Click **Add**.

The screenshot shows a web-based configuration interface for fabric parameters. It includes input fields for Name, Description, Tags, and Fabric ID. Below these fields is a table with columns: Name, WWN, IP Address, Dom..., Model, VF St..., and Status. An 'Add' button is located at the bottom right. Red arrows and numbers indicate the following steps:

- 1 – Give it a fabric name: Points to the Name input field containing 'FICON_FAB_10'.
- 2 – (Optional) Give it a description: Points to the Description input field containing 'FICON Logical fabric 10'.
- 3 – Give it a FID number: Points to the Fabric ID input field containing '10'.
- 4 – Click "Add": Points to the Add button.

Figure 6-13 Configuring fabric parameters

The fabric parameters are listed in Table 6-3.

Table 6-3 Fabric parameters

Parameter	Description
Name	The name of the fabric, not the logical switch.
Description	The description of the fabric, not the logical switch. With single switch fabrics, this field often is left empty; however, when cascading, the recommendation is to add a meaningful description, such as "tape backup for ..." or "disk mirroring for ...".
Tags	Tags are a new SANnav feature with which you can assign a tag that can be used for generating filtered reports. This tag is a database search tag in SANnav. It is unrelated to the RNID tag. Creating a tag is optional and can always be added later. Tags in mainframe environments are typically used by large environments only or when the same instance of SANnav is used to manage mainframe and open systems environments.
Fabric ID	▶ A fabric ID (FID) is a required integer 1-128 that is used to identify logical fabrics. ▶ By default, the FID for the default switch is 128. Although this ID can be changed, the recommended best practice is to continue to use 128 for the default FID and create a unique FID for other logical fabrics. ▶ It is also a common practice to use FID 127 for the base switch for ISL sharing. Choosing a FID number in the range of 1-126 is recommended. ▶ This ID is strictly an FC requirement. It has no effect on HCD or any other references to a FICON switch. The Domain ID, not the Fabric ID, is what is used for the high-order byte of the link address in HCD. ▶ The FID must be unique for each logical switch in a chassis and must be the same for all logical switches in the same fabric. It does not need to be unique for each fabric on different chassis. ▶ Although the specific FID number does not need to be unique, having an FID numbering convention is recommended. For example: – All FICON fabrics use FID 1. – When a chassis is divided into two FICON fabrics, slots 3, 4, 5, and 6 use FID 1; slots 9, 10, 11, and 12 use FID 2. – Logical fabrics for mirroring use FID 10.

5. Complete the following steps to create the FICON logical switches for the fabric:

a. Select the chassis to create FICON switches for this fabric (see Figure 6-14).

In this window, you select the chassis where the FICON logical switch to be created. For non-cascaded environments, you choose only one chassis. Although switches can be created individually and merged later, you save some configuration steps later if you create all the logical switches for a single cascaded fabric at creation time.

The 'Select Chassis' dialog box displays a table with the following data:

☐	Name ^	IP Address *	Logical Switch Count *	VF Status *	Firmware *
☐	Brocade6510	10.124.5.186	1	Enabled	v8.2.1
☐	Brocade_DCX4S	10.124.5.183	1	Enabled	v8.2.1c
☒	BrocadeG620	10.124.5.187	1	Enabled	v8.2.1a
☒	fcs01enwdco_X6	10.124.5.180	1	Enabled	v8.2.1b

At the bottom of the dialog, there are 'Next' and 'Cancel' buttons. A red arrow points from the text '1 – Select chassis to create FICON switch(es) for this fabric.' to the 'fcs01enwdco_X6' row in the table.

Figure 6-14 Selecting the logical switch

b. Set the Domain IDs for each switch in the fabric (see Figure 6-16 on page 122). The domain ID (DID) is the highest byte of the 2-byte link address. The recommended best practice is to make the DID match Switch ID in HCD.

The 'Select Chassis' dialog box shows the configuration for the selected chassis. It has two columns: 'Name ^' and 'Domain ID *'.

Name ^	Domain ID *
BrocadeG620	0x01
fcs01enwdco_X6	0x02

At the bottom of the dialog, there are 'Back', 'Next', and 'Cancel' buttons.

Figure 6-15 Setting the switch domain ID

- c. Add the ports. Most administrators click the **Slot/Port #** column to sort by port number. The right side of this set up page (which is not shown in Figure 6-16), allows you to manually edit the port address (low byte of the 2-byte link address). By default, port addresses are assigned sequentially; therefore, the recommendation is to select the ports in the order that you want the link addresses to be. Typically, this order is the same order in which the slot or port is listed.

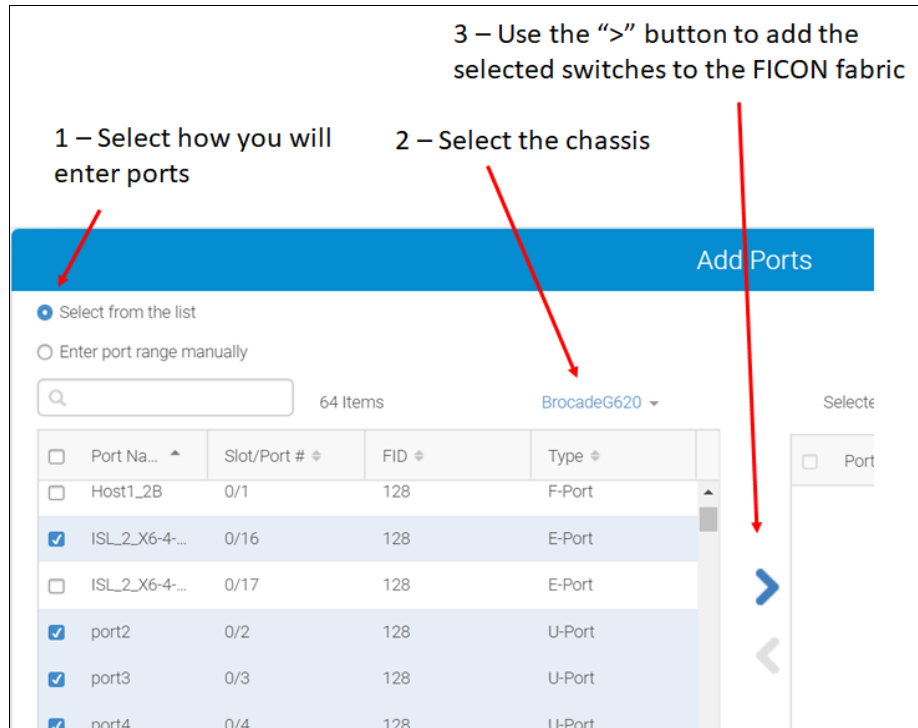


Figure 6-16 Selecting the port to add

- d. Click **OK** to finish adding the ports (see Figure 6-17).

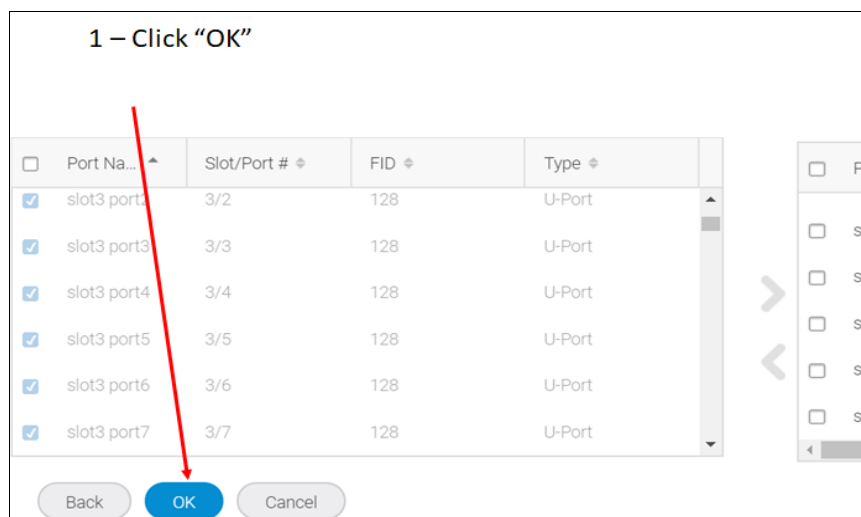


Figure 6-17 Completing the addition of the ports

- e. Configure the remaining fabric parameters. These parameters are applied to all logical switches in the fabric.

- f. Select **Fabric Parameters** to display parameter options.
- g. Enter the wanted parameter options, as listed in Table 6-4, and click **Save** (see Figure 6-18).

Click on the pull down arrow next to “Fabric Parameters” to display these options.

▼ Fabric Parameters

RA TOV: 10000

ED TOV: 2000

Maximum Hops: 7

BB Credit: 16

Data Field Size: 2112

Routing Policy: Exchange Based ▼

☐ Allow XISL use
☐ Enable device probing
☐ Per-frame routing priority
☐ Suppress Class F traffic
☐ Long distance fabric

☒ Activate

Save Cancel

It doesn't matter what order you select the options in. Once all options are selected, click “Save”

Figure 6-18 Configuring the remaining fabric parameters

Table 6-4 Fabric parameter definitions

Parameter	Description
RA TOV	Do not change.
ED TOV	Do not change.
Maximum Hops	Do not change.
BB Credit	Do not change. This parameter is not how BB credits are changed on a per-port basis.
Data Field Size	Do not change.
Routing Policy	For more information, see “Routing policy” on page 132.
Allow XISL	Is selected only if the ISLs are in a base switch to be shared with this logical switch in the same chassis.
Enable Device Probing	Should be cleared for FICON fabrics.
Per-frame routing priority	Should be cleared for FICON fabrics.
Suppress Class F traffic	Should be cleared for FICON fabrics.

Parameter	Description
Long-distance fabric	Because selecting this option adds buffer credits to the ISL ports (E-Port), it should be used in cascaded environments. The number of buffer credits depends on the speed of the link and FC generation. Adding buffer credits allows more frames to be in flight, which can result in unintended effects on congestion with short distance links. It also means that more frames must be recovered if a network failure occurs. Select this option only when more buffer credits are needed. For normal tape and DASD traffic, setting a long-distance mode often is not necessary until the distance exceeds 1 km; therefore, this option is typically not set for cascading within a campus environment. For distances that exceed 10 km, other methods are used for providing more buffer credits; therefore, this option typically is set for distances of 1 - 10 km. The exceptions are channel to channel, CTC, or extension through DWDM. Block storage always uses high-performance FICON (zHPF), which means the Extended Count Key Data (ECKD) frames are packaged with the data so that the average FC frame size for FICON traffic approaches the full 2 K byte FC frame. CTC traffic often is used for transporting small messages; therefore, the average frame size often is small. Because each FC frame uses a buffer credit, more buffer credits are required for CTC traffic. Some active DWDM equipment supplies buffer credits; therefore, buffer credits do not need to be increased. When cascading by using DWDM, it is important to discuss whether the added buffer credits must be configured on the SAN switch or the DWDM interface with the DWDM vendor. Details about how and when to configure more buffer credits for extended distance links is beyond the scope of this publication. Check with your IBM technical representative before configuring any cascaded environment.
Activate	When selected, the logical switches are enabled as are the ports after the creation process completes. Typically, the cabling is done before the logical switch creation, so it is common to select the Activate option. It is not safe to perform cabling duties with enabled SFPs; therefore, this box is not checked whether the cabling operations are incomplete.

6. Discover the logical switches.

Logical switches are discovered by default. The system administrator might disable automatic discovery or a network error might prevent discovery. It is always prudent to validate that the switch was discovered.

7. Rename the logical switches.

The switch names by default are `Switch_` followed by the FID number. In our example, both switches are named `Switch_10`. The names of these switches can be changed in the Inventory window.

8. Disable unused ports.

It is not safe to have lasers that are enabled on SFPs to which nothing is connected. In some instances, the open connection is at the patch panel, not the switch. As a safety precaution, all unused ports should be disabled.

9. Zone the FICON ports.

FICON zoning includes the following recommendations:

- Put all ports into a single zone that is based on the port domain and index.
- Do not put ICL ports in the zone.
- Although it is ideal to not have any E-Ports in the zone when cascading, having the E-Ports in the zone does not cause problems.
- Consider more zoning granularity. For example, separate mirrored ports with zoning.
- Domain, port index zoning is always used for FICON.

- To simplify zoning algorithms and to accommodate various switch form factors, a port index rather than a port number is used for zoning. This is a subtle difference few customers notice. If the ports are in the zone, the internal zoning mechanism has no effect on how paths are configured with HCD.
- Create a zone configuration with the zone.
- Activate the zone configuration.
- Disable the All Access option.

Complete the following steps to create a zone (see Figure 6-19):

- Select **Zoning**.
- Select **Zones**.
- Click **+**.



Figure 6-19 Creating a zone

- Select the fabric (see Figure 6-20). Zoning is applied to a fabric, *not* an individual switch. Because most FICON fabrics are single switch fabrics, this distinction is subtle; however, you are looking for the fabric name, not the switch name, when zoning. In this example, FICOn-XRC-z15 is the name of the fabric created.

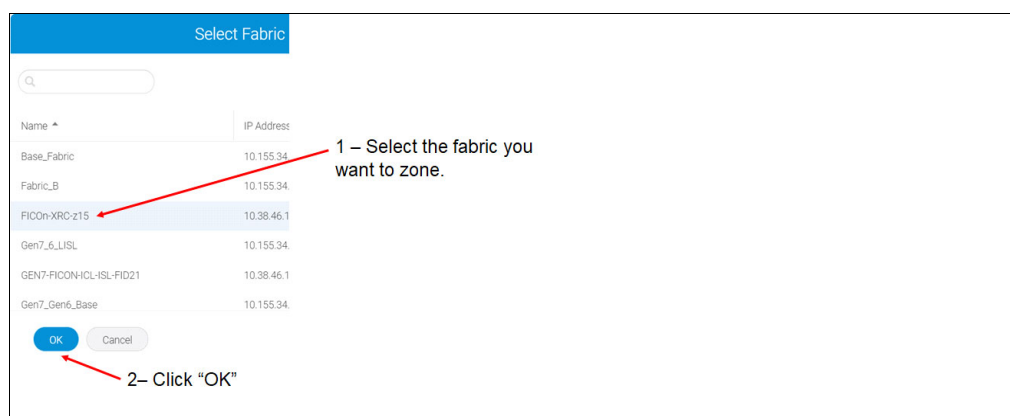


Figure 6-20 Selecting the fabric to zone

- e. Name the zone and configure the type of zone (see Figure 6-21).

1 – Name the zone

2 – Optional – Add a description of the zone

Name: FICON_Zone

Description: All ports based on d.i

Tags:

Zone Type: Standard Zone

Fabric: FICON-XRC-z15

4 – Click “Add”

Members

Zone AL...	Member Count	Type	Status
0 Items			

Add

Figure 6-21 Name zone and configure zone type

- f. Add the ports to the zone (see Figure 6-22). Typically, all but the ISL ports are added to the zone.

1– By default zone by “Alias” will be selected.
For a FICON zone, select “Domain, Port Index”

2– Select the ports to add to the zone. For most FICON fabrics, all ports will be selected which can be accomplished with a single check in the upper box by “Domain, Port”

3– Click the right arrow key to add the ports to the zone.

Select discovered Devices/Ports

Enter manually

Alias: Domain, Port Index

Domain, Port Index

Domain, Port ...	P...	F...	S...	Zone Ali...
0x65, 0x74	–	6500...	sw14...	–
0x65, 0x75	sw14...	6515...	sw14...	–
0x65, 0x76	sw14...	6517...	sw14...	–
0x65, 0x77	–	6503...	sw14...	–
0x65, 0x78	–	6511...	sw14...	–

Select

Figure 6-22 Adding the ports to the zone

- g. Complete the zone creation and save it (see Figure 6-23).

1– Click “OK” at the bottom of the page after adding ports to the zone.

2– Click “Save” at the bottom of the page after completing the zone definition.

OK

Save

Figure 6-23 Completing the zone creation

- h. Add the zone (or zones) to a zone configuration (see Figure 6-24 on page 127). A zone configuration is a collection of zones.

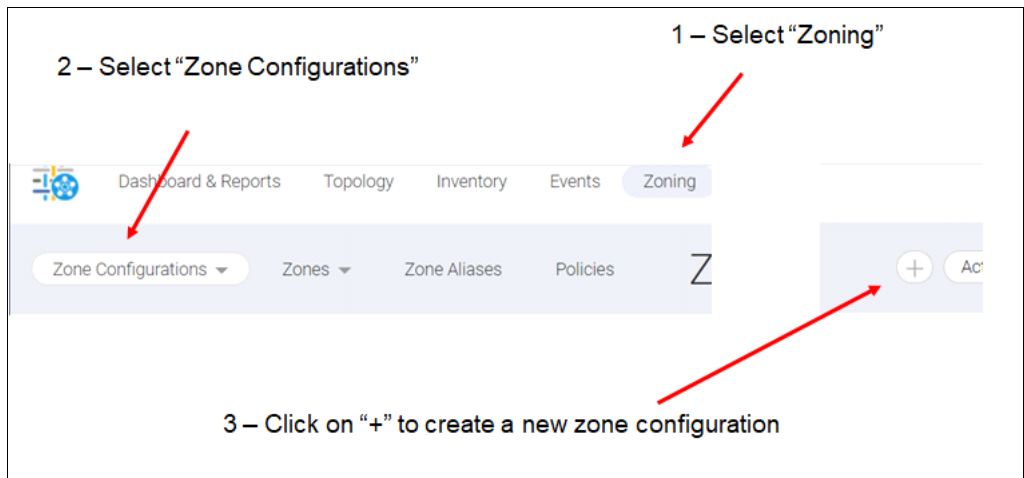


Figure 6-24 Creating a zone configuration

For FICON users who defined only one zone (see Figure 6-25), this addition can seem redundant.

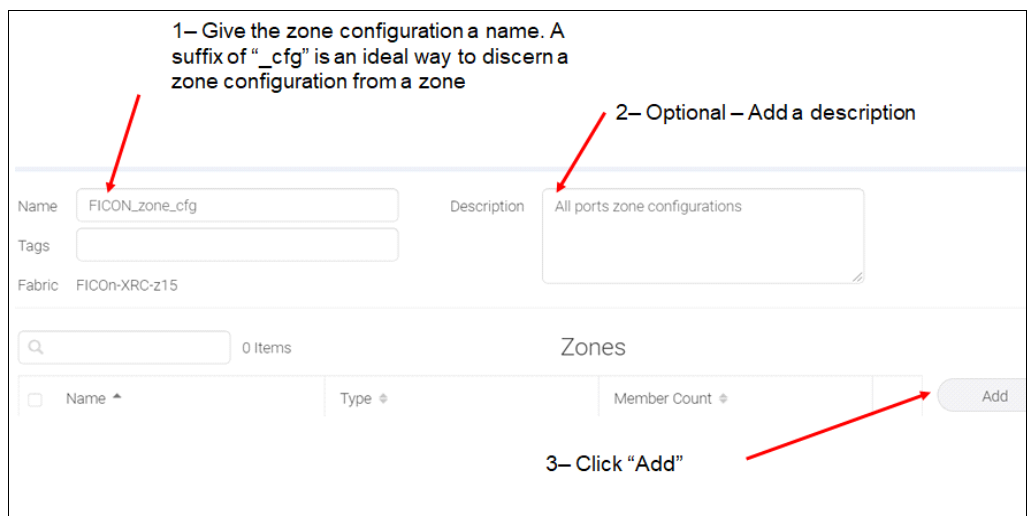


Figure 6-25 Defining the zone configuration

- i. Add the zones to the zone configuration. In open systems environments, it is common to have multiple zone configurations; therefore, activating the zone configuration is optional. For most FICON environments, only one zone configuration is used, which is often activated when the zone configuration is created (see Figure 6-26).

The screenshot shows the 'Add Zones' dialog box. It has a search bar at the top. Below it is a table with columns 'Name' and 'Type'. The table contains two rows: 'FICON_Zone' (checked) and 'gen7-xrc'. Below the table are buttons 'OK', 'Create New', and 'Cancel'. At the bottom, there is a checkbox labeled 'Activate' (checked) and buttons 'Save' and 'Cancel'. Red arrows point from text annotations to these elements.

1— Typically only one zone with all ports is added to the zone configuration but if a more granular zoning approach was used, select all zones that should be in the zone configuration.

2— Click “OK”

3— Check “Activate” to activate the zone configuration.

4— Click “Save”

Figure 6-26 Saving and activating the zone configuration

- j. For added security, disable the All Access zone (see Figure 6-27).

The screenshot shows the 'Policies' tab in a configuration interface. It has a table with columns 'Name' and 'WWN'. The table contains four rows: 'Base_Fabric', 'Fabric_B', 'FICON-XRC-z15' (checked), and 'All Access'. To the right of the table is a dropdown menu for the 'All Access' row, showing 'All Access' and 'Disable (No Access)'. Red arrows point from text annotations to these elements.

1 — Select “Policies”

2 — Select the fabric

3 — Use the pull down arrow by the fabric to select “Disable (No Access)”

Figure 6-27 Disabling the All Access policy

6.3.4 Managing FICON fabrics

As of this writing, the available version of SANnav features the following limitations:

- ▶ The equivalent of the Network Advisor connectivity display is split into multiple windows, primarily one for channel connections and another for storage connections.
- ▶ The RNID type for the connected device was missing.

Both of these issues are scheduled to be addressed. It is likely the corrections will be made in SANnav before most enterprises deploy SANnav.

Customizing the display

This customization is optional. Setting the FICON display mode defaults all column displays to the typical displays with RNID data that most mainframe administrators prefer to see, but the displays can be customized (see Figure 6-28).

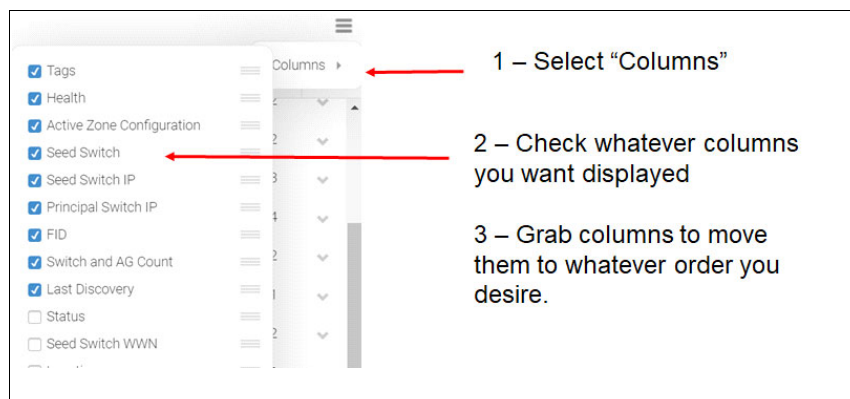


Figure 6-28 Customizing the display

Viewing connections

Complete the following steps to view the connections (see Figure 6-29):

1. Select **Inventory**.
2. Select **Host Ports** to see the channels, or **Storage Ports** to view the control units.

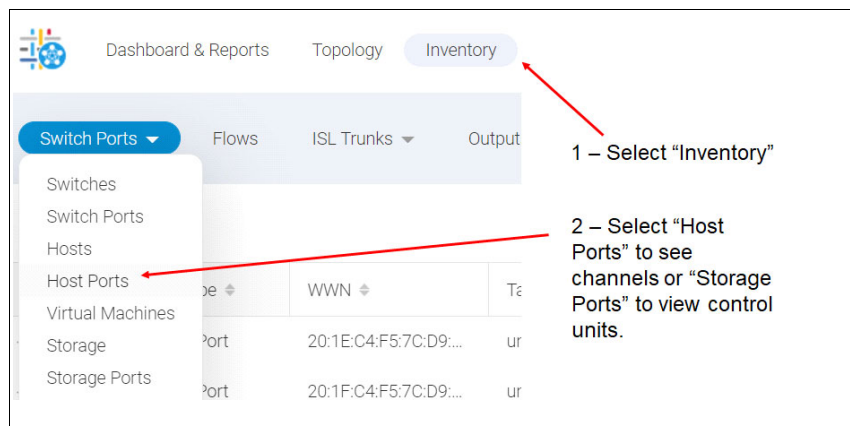
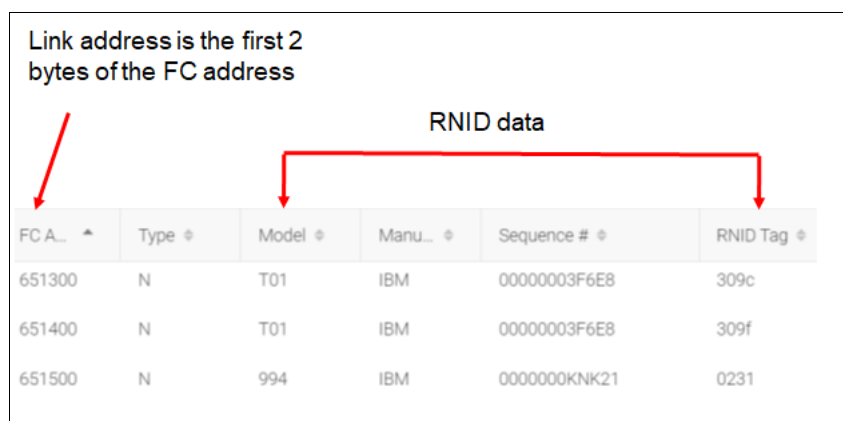


Figure 6-29 Viewing channel connections

The resulting display is shown in Figure 6-30.



Link address is the first 2 bytes of the FC address

RNID data

FC A... ^	Type *	Model *	Manu... *	Sequence # *	RNID Tag *
651300	N	T01	IBM	00000003F6E8	309c
651400	N	T01	IBM	00000003F6E8	309f
651500	N	994	IBM	0000000KNK21	0231

Figure 6-30 CHPID connectivity display

Similarly, control units can be monitored by clicking **Storage Ports** instead of **Host Ports**.

6.3.5 Converting a production switch to a FICON logical switch

This section describes how to check whether an outage is required for customers who are considering converting a switch that is not configured for a defined FICON logical switch to use a defined FICON logical switch. It is expected that most customers do not need to make this change in an environment.

The easiest way to check is by using the CLI. Save the session output so that you can email it to your Brocade Business Partner or provide it to your Brocade representative if you need assistance.

Checking high integrity fabric and address mode

Validate that the address mode is mode 1 (zero based) and that the high integrity fabric (HIF) mode is set (see Figure 6-31).

switch_30:FID30:admin> switchshow	
switchName	switch_30
switchType	166.0
switchState	Online
switchMode	Native
switchRole	1
switchDomain	Principal
switchId	fffc01
switchWwn	10:00:c4:f5:7c:64:5b:62
zoning	OFF
switchBeacon	OFF
FC Router:	OFF
HIF Mode	ON
Allow XISL Use	OFF
LS Attributes	FID: 30, Base Switch: No, Default Switch: No, Ficon Switch: No, Address Mode 1]

Figure 6-31 Validating the address mode

If HIF Mode is not ON or the Address Mode is not 1, an outage likely is needed.

Validating address binding

Validating the address binding is shown in Example 6-1.

Example 6-1 Validating address binding

```
switch_30:FID30:admin> portaddress --show
Partition Address Mode :1
Index Slot Port Area Mode User_bound
=====
72 9 8 0x0000 8 bit Y
73 9 9 0x0200 8 bit Y
74 9 10 0x0300 8 bit Y
75 9 11 0x0400 8 bit Y
76 9 12 0x0500 8 bit Y
```

A Y must appear in the column for all ports.

Validating FICON logical switch

In response to the **switchshow** command, HIF mode must be ON and the FICON switch must be Yes, as shown in Figure 6-32.

switch_30:FID30:admin> switchshow	
switchName	switch_30
switchType	162.0
switchState	Online
switchMode	Native
switchRole	Principal
switchDomain	1
switchId	fffc01
switchWwn	10:00:c4:f5:7c:16:8b:3e
zoning	OFF
switchBeacon	OFF
FC Router:	OFF
HIF Mode	ON
Allow XISL Use	OFF
LS Attributes	FID: 30, Base Switch: No, Default Switch: No, Ficon Switch: No, Address Mode 1]

Figure 6-32 Validating FICON logical switch

Routing policy

This policy primarily effects load balancing in cascaded environments. Most FICON environments today support FICON Dynamic Routing (FiDR).

The routing policy must be the same for all switches in a SAN, but each SAN is independent from one another. Therefore, each SAN might have different routing policies. Workload Manager continues to operate as expected, even though paths are connected to switches with different routing methods. Upgrades can include a routing method change and each SAN upgraded in separate change control window.

A route is the set of switches and ports a path takes to deliver FC frames from the channel to the control unit. As a practical matter, this is only significant in cascaded environments. The routing methods are listed in Table 6-5 on page 133.

Table 6-5 Routing methods and descriptions

Routing method	Description
Port Based Routing (PBR)	▶ Deprecated. Many older FICON SAN deployments are configured this way. ▶ Routes are determined by a predefined map for each possible route in the SAN. The channel path always takes the same route (same ISL or ISL trunk group). ▶ Today, the only reason to configuration a switch for PBR is if a new switch must be connected to an older switch that is still configured for PBR. Setting PBR is not an option in SANnav but can be done by using the CLI. (How this process is done is beyond the scope of this publication.) Contact your technical account team if configuring PBR is required.
Device-Based Routing (DBR)	Default. Most FICON SAN deployments are configured this way. Routes are determined by usage. A round robin algorithm is used to select the ISL or ISL trunk group. Routes are established dynamically as a channel to control unit path is needed, but after that path is selected, the channel path always takes the same route. This has a better statistical probability of balancing work load across ISLs than PBR.
Exchange-Based Routing (EBR).	▶ Analogous to FICON Dynamic Routing (FiDR). Although EBR is the preferred routing method, DBR is the default because many FICON SANs still have older control units that do not support FiDR. Typically, these are older tape control units. Check with your storage vendor to be sure FiDR is supported before selecting EBR for the routing method. Some control units can require a driver upgrade. ▶ EBR selects routes dynamically on a per exchange basis. An FC “exchange” is analogous to a CCW chain. The least-used ISL or ISL trunk group for a route is selected at the time a switch receives a new CCW chain from the channel. The route remains intact for the life of the CCW chain. This method results in the most efficient and dynamic use of ISL bandwidth.

Exchange-based routing and FiDR background

Although this level of detail is not required knowledge, questions about FiDR and EBR are common.

Mainframe channel protocol was developed for buffered control units that managed physical storage devices. Channel programs were built around channel command words (CCWs). Write commands included data. CCWs can be (and often are) strung together in a channel program.

Status is a bit flag in a byte returned by the control unit. The only status bits pertinent to this discussion are Channel End (CE) and Device End (DE). CE indicates that the control unit accepted the channel program. Device End (DE) indicates that the operation with the physical storage completed. So that a channel can perform other work while a control unit was working on a channel program, the status that was split in that CE was returned to the channel and DE after the control unit performed the action with the physical storage.

In a serialized bit stream across FC, CCWs are divided into FC packets and the packets are sent to their destination. The equivalent to a bus connection in FC is an exchange.

All CCWs, and all FC packets in a channel program are sent on the same exchange so all packets arrive in order.

When split status is returned to the channel, DE is returned on a different exchange. Although the protocol did not change, the devices are now often in memory so the control unit can complete the operation much more quickly. This means DE can be available almost immediately after returning CE.

Because DE is on a different FC exchange, the FC packet with DE can be routed across a shorter or less congested ISL thereby arriving before CE. Because it is not possible for a device to complete an operation before the control unit completed the operation, the channel program never had to account for receiving DE before CE. If this issue occurred before FiDR, a channel detected error is reported and the job might end abnormally (ABEND).

FiDR recognizes this potential race condition and waits for CE instead of returning an error. The channel and control unit must support FiDR.

6.4 FICON SAN migration

Getting a SAN Health report is the best way to answer to determine migration tasks. It is also the best documentation for your account team to become familiar with your SAN before giving you advice. Because SAN Health reports are confidential, remember to authorize your IBM and Brocade representatives to receive a copy.

It is anticipated that most migrations are from Gen 5 to Gen 7.

6.4.1 Firmware

FICON switch configurations (with the suitable FICON supported FOS levels) that are supported today are also supported by switches that include a FICON logical switch configuration. All upgrade scenarios that include a mix of logical switches that are defined as FICON and FCP switches are supported; however, all switches in a fabric cannot be more than one major FOS revision apart. Any switches that are on FOS 7.x must be upgraded to 8.x before any other switch in the fabric can be upgraded to 9.x.

6.4.2 FICON logical switch

Upgrades from FOS 8.x on Gen 6 switches to FOS 9.x are blocked for switches that do not meet the FOS 9.0 requirements. To avoid the need for an outage, it is highly recommended that all new Gen 6 switch configurations include virtual fabrics that are enabled and a logical switch that is defined as a FICON logical switch.

As of this writing, the FOS v8.x code stream was still actively supported; however, eventually, it reaches end of support before FOS 9.x and likely before a hardware upgrade of the Gen 6 switch. At some future date, Gen 6 switches must be upgraded to FOS 9.x for CVE patches and other changes.

Enterprises with Gen 6 switches operating on FOS 8.x firmware that were not configured for a FICON logical switch should plan on upgrading the configuration so that a FICON logical switch is used for all FICON connections. The addressing does not to be changed; therefore, the channel paths do not need to be redesigned.

Use Table 6-6 to determine whether an outage is required to configure a FICON logical switch.

Table 6-6 Requirements for a nondisruptive reconfiguration to a FICON logical switch

Parameter	Requirement for nondisruptive conversion	Comments
Switch type	SAN512B-6 (8961-F08) or SAN256B-6 (8961-F04)	The SAN42B-R (2498-R42) and SAN64B-6 (8960-F64/N64) always requires an outage.
Logical switch	Non-default logical switch	VF must be enabled and a logical switch other than the default logical switch must be correctly configured for FICON.
Address mode	Zero based (mode 1)	Zero-based address mode is the addressing mode that ensures that the FC address that is assigned to a port is in compliance with the channel path link addressing. It is not always set today because specific configurations were inherently in compliance; therefore, you must check.
High Integrity Fabric (HIF)	Must be enabled	Setting the HIF ensures that the security policy is set and distributed throughout the fabric, and the insistent domain ID is set. All correctly configured FICON SAN fabrics should have HIF enabled today.
Port Address Binding	Must be in effect	All correctly configured FICON SAN fabrics should have port address binding in effect today.

6.4.3 Inter-Chassis Links

Inter-Chassis Links (ICLs) are the Quad Small Form-factor Pluggable (QSFP) fiber links between core routing blades that eliminate the need for Inter-Switch Links (ISLs) in cascaded environments. Although downward compatibility to Gen 6 and Gen 5 switches is supported, the QSFPs that are required for connectivity are limited to 32 G.

Therefore, to get the full speed capabilities of Gen 7 ICLs, it is recommended that both directors that are cascaded with ICLs are upgraded at the same time to avoid the need for purchasing QSFPs twice.

ICLs are available in 100 m and 2 Km options. The 100 m option requires OM4 or OM5 multi-mode fiber and an MPO/MTP connector. The 2 KM option requires single mode cable and uses an LC (9 micron) connector. As of this writing, the 2 Km option for 64 G (Gen 7) ICLs was not available. No plans are in place for a 2 Km Gen 5 to Gen 7 option. Also, because attenuation is not required with the 2 Km option, it can be used instead of the 100 m option.

6.4.4 Cabling

Although single-mode cable replacement is generally not necessary in mainframe environments, it is not uncommon to need to replace a small percentage of cables. The most common reason for this is that a lens is scratched during cable cleaning.

In some cases, degraded cables might not exhibit problems until higher speeds are attempted. Although multi-mode cable is rare in mainframe environments, some exceptions exist. The most common exception for using multi-mode cable is for DWDM connections. All multi-mode cable should be at least OM-4 rated.

As with any repurposing of cabling, good cable hygiene is recommended, which must include cleaning all connection points. Even when performing a push-pull upgrade of a switch, connector cleaning should include the connections at the patch panel and the channel or control unit in addition to the connection at the switch port.

6.4.5 Patch room and fiber plant

Migrating from products with different form factors or changing port count can affect documentation and patch room design. This issue is especially so when migrating from director class products with 32 port blades to 48 port blades.

6.4.6 SFPs

When considering re-purposing SFPs, remember that Gen 7 switches and switch blades require SFPs that are designed with the secure optics feature. This means that although Gen 7 can accept a 32G optic, you cannot re-purpose 32G optics from Gen 6 switches in a Gen 7 switch or switch blade.

As with all previous FC switching products from Brocade, the latest generation product supports SFPs of the same generation and one generation below. A Gen 7 switch port is capable of 64G; therefore, it can accept a 64G SFP or a 32G SFP.

A Gen 6 switch port is capable of 32G; therefore, it can accept a 32G or 16G SFP. Each SFP supports three speeds: the current generation, one generation back, and two generations back. Consider the following points:

- ▶ A 64G optic can negotiate to 64G, 32G, and 16G
- ▶ A 32G optic can negotiate to 32G, 16G, and 8G
- ▶ A 16G optic can negotiate to 16G, 8G, and 4G

If a 4G connection is required (typically for older tape or DWDM), you need a 16G SFP. Because 16G SFPs are not supported on Gen 7 products, you need a Gen 6 switch or Gen 6 switch blade to accommodate older connections that require 4G. You can use the Brocade branded 16G SFPs from your Gen 5 products in a Gen 6 switch or switch blade, but you should only do so if 4G support is required.

Most FICON SANs are single switch (single domain) fabrics. Channel paths can be striped across SANs by using a mix of Gen 5, Gen 6, and Gen 7 technology if they are configured correctly for FICON. If they are not in the same fabric, they can be on any version of FOS that is supported for FICON, and can have different addressing modes and different routing methods. Therefore, switches in a channel path can be upgraded without the need to be concerned with redundant channel paths (see Figure 6-33 on page 137).

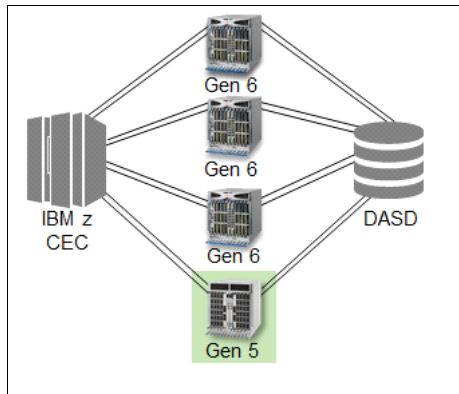


Figure 6-33 Mixed Gen 5 and Gen 7 Single Switch Fabrics

In cascaded environments, a non-VF enabled switch can be connected to a logical switch that has VF enabled if every switch in the fabric that has VF enabled has the same Fabric ID (FID).

To minimize outages, most customers upgrade both switches in the channel path. However, if the need arises, mixing switch generations in the same fabric is supported but, with the restrictions that are described next. A SAN42B-R (2498-R42) extension switch is a Gen 5 switch.

Configuration requirements

The following configuration requirements must be met:

- ▶ The FOS code must be at v9.0 or higher in the Gen 7 switch and a supported version of v8.x or higher in the Gen 5 or Gen 6 switch.
- ▶ All the fabric parameters must match.
- ▶ For FICON environments, the routing policy is the most likely parameter that is different because most Gen 5 implementations used port-based routing or device-based routing and new installations are likely to use exchange-based routing (FiDR).
- ▶ It is unlikely that other fabric parameters are different; however, if switch generations must be mixed in the same fabric (see Figure 6-34), contact your technical account team to validate all fabric settings.

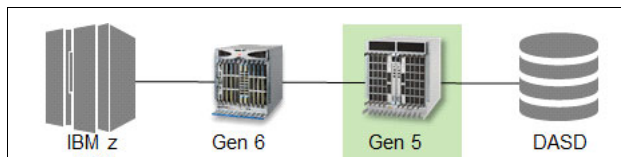


Figure 6-34 Sample mixed Gen 5 and Gen 7 cascaded fabrics

FMS

Enabling the FICON Management Server (FMS) allows administrators to define and configure the Control Unit Port (CUP) device online. Because all chassis level alerts are reported to the CUP device on all logical switches, a channel does not need to be dedicated to the default switch.

All other actions, such as display matrix commands and 74-7 record capturing, are limited to the logical switch where the channel with CUP is connected. If multiple logical switches are used, you must define CUP on at least one channel that is connected to each logical switch where CUP management is wanted.

Gen 5 to Gen 7 fixed port switches

The Gen 7 SAN64B-7 (8960-P64/R64) has an extra eight ports, but otherwise has the same physical format as the Gen 6 SAN64B-6.

Assuming slower login speeds are not an issue, a simple push-pull can be done for physical installation when upgrading from a SAN48B-5 to a SAN64B-7.

It is not uncommon for enterprises to make CHPID path changes as part of an upgrade, but there are no changes in the SAN64B-7 that require changes to the IODF.

Gen 6 to Gen 7 fixed port switches

To use some available space on the front panel and to make full use of all ports on the ASIC, the Gen 6 switch also includes 4 QSFP ports. The QSFP ports were only qualified as ISL ports when used in FICON environments. They cannot be used with breakout cables for FICON connectivity. They can be used with breakout cables in a logical switch for non-FICON connectivity.

The QSFP ports on the SAN64B-6 were replaced with dual capable SFP ports, so the only change is to environments that used the QSFPs for inter-switch links (ISLs) or for non-FICON applications.

Few, if any, enterprises fully used all QSFP. If more than 56 ports were used with the SAN64B-6, another switch is required. The extra eight ports on the SAN64B-7 are designed to accept a dual SFP. However, as of this writing, only standards SFPs were available.

Because the use of QSFPs is uncommon for FC, the equivalent Gen 7 switch was designed to use the space that was used for QSFP for SFP ports that support double density SFPs. As of this writing, double density SFPs were not available, but it is anticipated that the double density SFPs is supported for ISLs only, not FICON connections. Because those ports support standard SFPs as well, the SAN64B-7 can be configured as a 56 port switch with three 8-port licenses.

For environments that used the QSFPs as ISLs with a total port count of 56 or less, the ISL connections must be moved off of the QSFPs to the SFP ports. Replacing the SAN64B-6 can involve some recabling, but otherwise, a simple push-pull is possible.

It is not uncommon for enterprises to make CHPID path changes as part of an upgrade but since the SAN64B-7 supports more FICON ports than the SAN64B-6, there are no required changes to the IODF when upgrading Gen 6 switches to Gen 7 technology.

Gen 6 to Gen 7 directors

The Gen 6 directors, SAN256B-6 and SAN512B-6, have the same physical card layout and form factor as the Gen 7 directors, SAN256B-7 and SAN512B-7.

In most cases, a simple push-pull of the Gen 6 director for a Gen 7 director is possible. Exceptions are when 4G connections must be supported because these connections require a 16G SFP, which is not available for Gen 7.

The options to accommodate older 4G connections is to move them to a fixed port switch or add a Gen 6 blade to the Gen 7 chassis. Gen 6 blades are available for purchase or can be migrated from an existing Gen 6 director to a Gen 7 director.

It is not uncommon for enterprises to make CHPID path changes as part of an upgrade, but no changes are in the SAN256B-7 or SAN512B-7 that require changes to the IODF.

Gen 5 to Gen 7 directors

Most, if not all, FICON implementations on Gen 5 directors, SAN768B-2 and SAN384B-2, were with 32 port blades. SAN upgrades typically skip a generation, so Gen 5 to Gen 7 is the most common upgrade path to Gen 7.

Because redundant channel paths usually are planned to be on different FRUs (port cards), the same channel path mapping that is used with 32 port blades likely do not work on 48 port cards. This issue is the primary consideration when upgrading from Gen 5 to Gen 7 directors.

The Control Processor (CP) blades are now half height blades that occupy the first physical slot in the chassis. Reducing the total number of physical slots from 12 to 11 was done to allow extra air flow to improve cooling for the port cards. This change was necessary because the higher speed optics produce more heat.

These changes affect the following components:

- ▶ CHPID path mapping
- ▶ Cable diagrams and other documentation that includes chassis slot numbering
- ▶ Supported SAN switch addressing modes
- ▶ Patch room design
- ▶ Automation such as automatic service ticket generation

The control processor (CP) cards in the 8510-4 and 8510-8 chassis occupied full slots in the middle of the chassis. With the Gen 6 and Gen 7 chassis, the CP blades are half height blades, both installed in the first chassis slot. The slot numbering is different.

To distinguish the physical location from one CP to the other, they are referred to as slots 1 and 2, even though they both occupy the same physical space as a single slot occupies (see Figure 6-35).

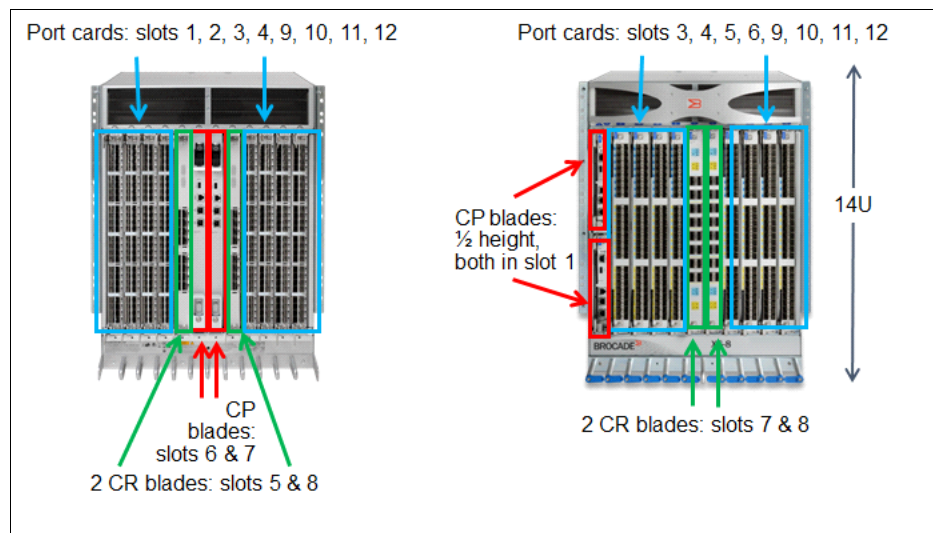


Figure 6-35 Gen 5 versus Gen 6 and Gen 7 card slots

Because most Gen 5 directors that are deployed for FICON used the 32-port card, it is likely that several cabling changes must be made. Those changes depend on the design approach that is used. For patch room designs that were planned to match the symmetry of the blades and channel paths, changes (at least in documentation) might be required for the patch room as well.

Consider the following points regarding physical upgrades:

- ▶ All documentation templates must be updated.
- ▶ Custom scripts and software to automatically generate service tickets or other monitoring activity might need to be updated.
- ▶ The patch room layout might need to be redesigned.
- ▶ Cabling methodologies likely will change.

In most deployments, significant changes to the IODF are required. Much of the design section is dedicated to resolving channel pathing that arises as a result of architectures that are build around 32 port cards to 48 port cards.

For more information, see the following resources:

- ▶ [Brocade FICON Resource page](#)
- ▶ [SANnav Introduction](#)
- ▶ [SANnav Announcement](#)
- ▶ [SANnav Data Sheet](#)



IBM b-type Gen 7 extension design, implementation, and migration

Extension is the most common and least expensive form of Fibre Channel (FC) and FICON storage connectivity across a distance IP network. Fibre Channel over Internet Protocol (FCIP) is a technology that allows block data replication over long distances by using the TCP/IP protocol.

IP Extension is a technology that accelerates, secures, and monitors IP storage replication flows over long-distance IP networks. In most cases, Extension is implemented in Disaster Recovery (DR) scenarios with some kind of data replication between the primary and secondary sites.

Extension is a tunneling technology, which means FC frames (FCIP) or IP datagrams (IP Extension) are encapsulated into TCP segments and IP datagrams (TCP/IP). As such, encapsulation is not apparent to devices that are connecting through by way of Extension. To use Extension, a tunneling device is required on both ends of the TCP/IP link that integrates multiprotocol connectivity. Extension platforms are stand-alone devices (IBM SAN42B-R and IBM SAN18B-6) or blades that are combined with a director-class product (IBM SX6 Extension Blade). Storage arrays, such as the IBM DS8900 and IBM FlashSystem® 9200, support IBM b-type Gen 7 Extension connectivity.

For more information, see *IBM FlashSystem 9200 and 9100 Best Practices and Performance Guidelines*, SG24-8448.

An essential aspect of the Extension scenario is the IP link quality. With Extension, dedicated bandwidth is prudent if the link is shared with other traffic. Because the connection between two sites is low-traffic or used only for email, do not assume that this is always the case. Storage is sensitive to latency and congestion; a spyware problem or a DDOS attack on the IP network can be disruptive. Furthermore, data that is in-flight should always be encrypted, which is easily done on IBM b-type Gen 7 Extension platforms.

Also, when you are communicating with your organization’s networking architects, distinguish between megabytes per second (MBps) and megabits per second (Mbps). In the storage world, bandwidth is often specified in MBps, but network engineers specify bandwidth in Mbps. If you want 1 Gbps and you fail to designate GB explicitly, you can end up with 1 Gbps, which supplies approximately 100 MBps. If you include safety margins, this link is not as fast as you might need, so ensure the correct terminology.

This chapter describes the IBM b-type Gen 7 (Brocade) Extension family of products. It includes an overview of the physical platforms, the different sides of an Extension platform, features, terminology, implementation, technical insights, and architectures.

This chapter includes the following topics:

- ▶ 7.1, “Extension platform overview” on page 142
- ▶ 7.2, “Extension design and architectures” on page 145
- ▶ 7.3, “Extension implementation” on page 157
- ▶ 7.4, “Extension migration strategy” on page 199
- ▶ 7.5, “Extension validation” on page 203
- ▶ 7.6, “Extension Monitoring” on page 209

7.1 Extension platform overview

IBM Extension offers advanced data transport features that accommodate various environments, from small to medium businesses to the largest most demanding enterprises. Extension is a high-speed, high-volume data transport technology for moving storage between data centers over an IP network. Extension provides exceptional data transport efficiency by incorporating sophisticated high-efficiency encapsulation, transport, security, compression, and data recovery mechanisms.

Moving FC data over an IP network is referred to as *FCIP*. IP storage data can be transported by IP Extension, which benefits from the advanced transport technology and encryption.

7.1.1 IBM b-type Gen 7 extension platforms

At a high level, this section familiarizes you with each physical Extension platform. For more information, see the data sheets.

IBM SAN18B-6

The IBM SAN18B-6 is a 1RU chassis of standard width, as shown in Figure 7-1. The front has access to the serial port, Management Ethernet port, USB port, 12 32G FC ports (Gen 6), and six application Ethernet interfaces (GE0/GE1 and GE2/GE3 are mutually exclusive).

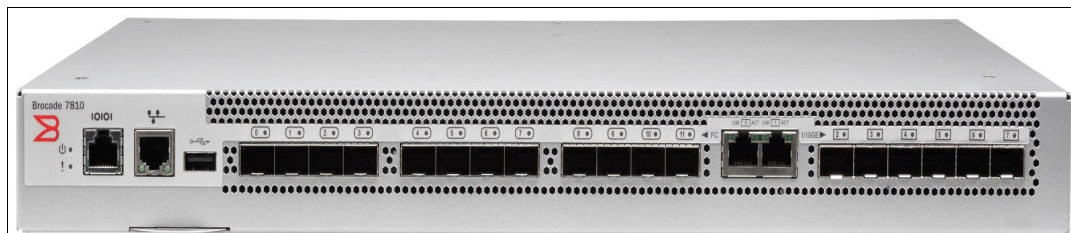


Figure 7-1 IBM SAN18B-6/Brocade 7810 front view

The IBM SAN18B-6 has two license modes: base and full. The full unit has a WAN side maximum of 2.5 Gbps. The base unit has a WAN side maximum of 1 Gbps and is enforced by disallowing circuits to have more than a cumulative maximum rate of 1 Gbps. There is one data processor (DP).

The IBM SAN18B-6 does not support Virtual Fabrics or FICON.

The IBM SAN18B-6 combines the power supply and fans into a single hot-swappable FRU, as shown in Figure 7-2. There are two redundant FRUs in the rear of the chassis. An offline power supply does not prevent power to any fan; a single online power supply can power all fans. Airflow is unidirectional rear to front.

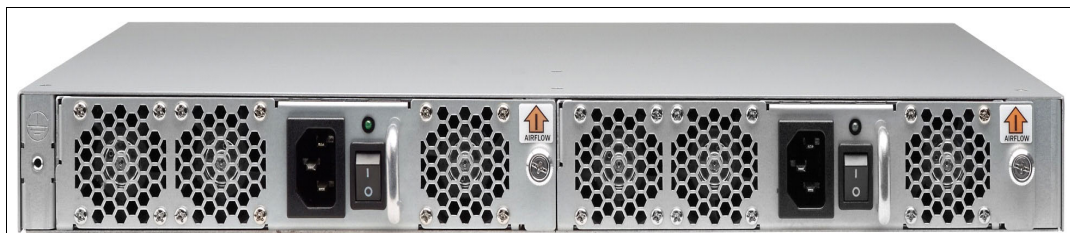


Figure 7-2 IBM SAN18B-6/Brocade 7810 rear view

For more information about IBM Storage Networking SAN18B-6, see [this web page](#).

IBM SAN42B-R

The IBM SAN42B-R is a 2RU chassis of standard width, as shown in Figure 7-3. The front has access to the serial port, management Ethernet port, USB port, 24 16G FC/FICON ports (Gen 5), 16 application GE/10GE Ethernet interfaces, and two 40GE interfaces.

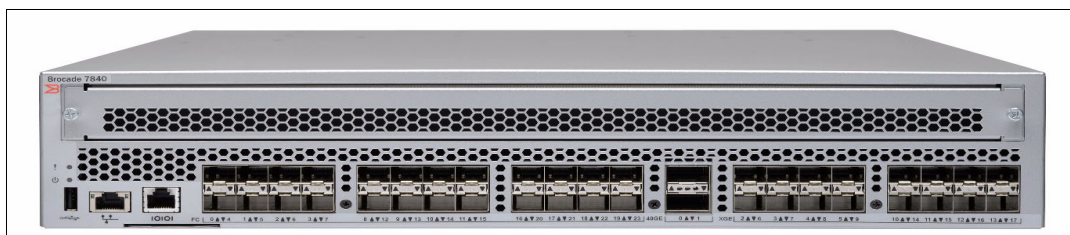


Figure 7-3 IBM SAN42B-R/Brocade 7840 front view

The IBM SAN42B-R has three license modes: 5 Gbps (base unit), 10 Gbps (Upgrade-1), and 40 Gbps. Upgrade-2 requires Upgrade-1, and Upgrade-2 is necessary to enable the 40GE interfaces.

There are two DPs, each capable of 20 Gbps. The maximum WAN side bandwidth is 40 Gbps. The 5 and 10 Gbps modes are limited by disallowing cumulative circuit bandwidth that is greater than the licensed limit. The full 40 Gbps licensing has no restriction on the cumulative circuit rate; nonetheless, 40 Gbps is the maximum hardware capacity.

The IBM SAN42B-R supports Virtual Fabrics and FICON.

Note: Extra optional hardware is not available for the empty slot.

The IBM SAN42B-R has a separate power supply and fan FRUs, as shown in Figure 7-4. There are two redundant power supplies and three fan trays in the rear of the chassis. An offline power supply does not prevent power to any fan; any online power supply can power all fans. Airflow is unidirectional rear to front.

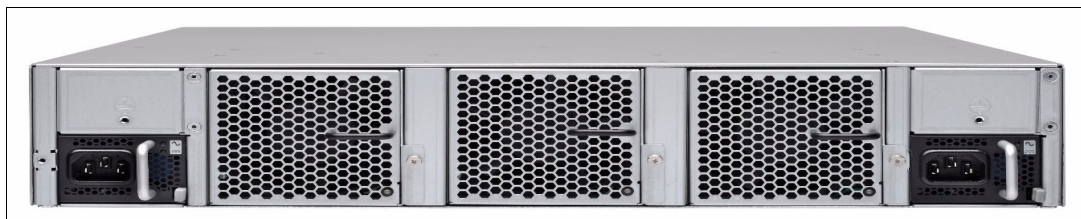


Figure 7-4 IBM SAN42B-R/Brocade 7840 rear view

For more information about IBM System Storage SAN42B-R Extension Switch, see [this web page](#).

SX6 extension blade

Up to four IBM SX6 Extension Blades are supported in IBM SAN256B-7/SAN512B-7 Directors, as shown in Figure 7-5.

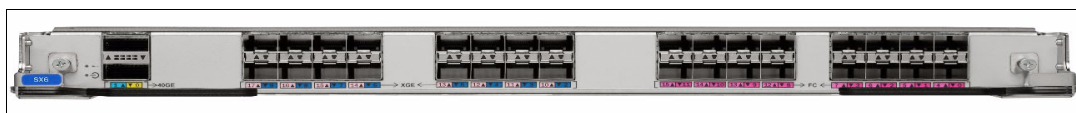


Figure 7-5 Brocade SX6 extension blade

Two classes of the SX6 Extension Blade are available:

- ▶ **SX6 for Open Systems:**
 - Includes all licenses, except those licenses that are specific to the mainframe.
 - No circuit bandwidth limits.
 - Includes SWL MMF optics.
- ▶ **SX6 for Mainframe:**
 - Includes all licenses, including FICON Accelerator and FICON Management Server (CUP).
 - No circuit bandwidth limits.
 - Includes LWL SMF optics.

The front has access to 16 FC/FICON ports (Gen 6 - supported in Gen 7 Directors), 16 application GE/10GE Ethernet interfaces, and two 40GE interfaces. The Ethernet SFP/SFP+ optics are not included and must be ordered as needed. Brocade branded optics are required to come online.

There are two DPs, each capable of 20 Gbps. The maximum WAN side bandwidth is 40 Gbps. The SX6 Extension Blades have no restriction on cumulative circuit rate; nonetheless, 40 Gbps is the maximum hardware capacity.

For information about the IBM SX6 Extension Blade, see [this web page](#).

7.2 Extension design and architectures

Most commonly, Extension is used for Business Continuance by way of DR. Protect critical data by replicating it outside the area of a potentially catastrophic event. Remote Data Replication (RDR) and remote tape are used for this purpose. Data replication systems with Extension permits organizations to recover operations in a timelier manner.

The next sections describe the most common Extension designs and architectures.

7.2.1 FCIP design and architectures

Typically, RDR is a disk-based, array-to-array replication. The local production site storage array sends data to the backup site array. Often, Native FC is used when the recovery site is within a metropolitan distance, and DWDM or dark fiber exists between the sites; however, frequently, cost-sensitive ubiquitous IP infrastructure is available.

IBM Extension is high-speed and adds approximately 75 μ s of propagation delay per pass through the platform. Four passes round trip = 0.3 ms total added propagation delay, which is added to the network latency. In some cases, the added propagation delay might be acceptable for synchronous RDR applications.

Deployment best practice connects array N_Ports to Extension F_Ports directly and does not connect through the production fabric. Storage replication ports are dedicated to RDR and should not pass through fabrics with host traffic, which can cause head-of-line-blocking issues for hosts.

Nevertheless, valid reasons exist for connecting by way of the production fabric, such as tape applications and too many arrays ports for the Extension platforms. If more ports are required and connecting to the production fabric is necessary, it is prudent to separate these ports into a logical replication fabric by using Virtual Fabrics.

An architecture best practice uses a separate autonomous replication fabric and does not combine the replication fabric with the production fabric.

Two-box architecture

A single Extension platform can be directly connected to different storage controllers and is referred to as a Two-Box solution, one at each DC, as shown in Figure 7-6.

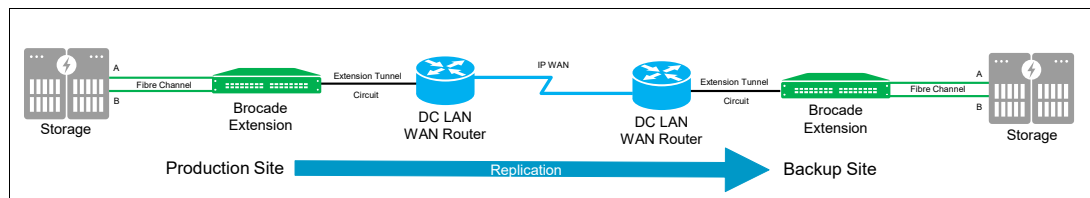


Figure 7-6 FCIP two-box architecture

A two-box solution is common with the IBM SAN18B-6 base unit, which supports one circuit per VE_Port. It supports two VE_Ports (tunnels), and those tunnels can run in parallel. However, this configuration is not the same as two circuits that are part of a single tunnel because no coordination or data recovery exists between different VE_Ports.

Not all storage applications support parallel FCIP connections. When parallel Extension tunnels are used with exchange-based routing (EBR), exchanges can be delivered out of order. Device-based routing or port-based routing (flow-based) also can be used to avoid issues with out of order Exchanges.

Four-box architecture

Alternatively, an Extension platform is dedicated to fabric and controller “A.” A physically different Extension platform is dedicated to fabric and controller “B.”

Two Extension platforms in each data center is referred to as a *four-box solution* (see Figure 7-7). A single WAN link for both tunnels, all circuits, can be used.

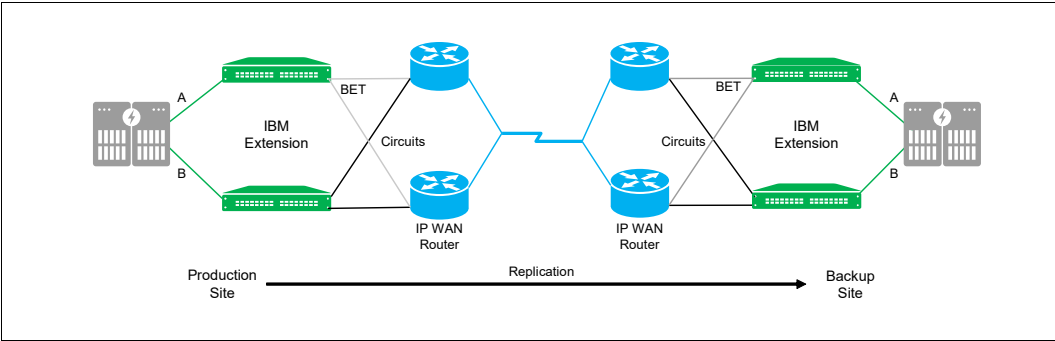


Figure 7-7 FCIP four-box architecture

Alternatively, two WAN links from different service providers can be used, with the circuits symmetrically spread across the WAN connections. The number of WAN connections depends on the availability requirements, cost, and recovery tolerance. With two WAN connections, Cir0 from each tunnel are pinned to WAN-A, and Cir1 from each tunnel is pinned to WAN-B.

Extension + FCR architecture

A production fabric can be extended by using Extension with or without FC Routing (FCR), as shown in Figure 7-8. Compelling reasons might exist to attach Extension to a production fabric; however, if no such reasons exist, the best practice is to connect end devices directly to the Extension platforms.

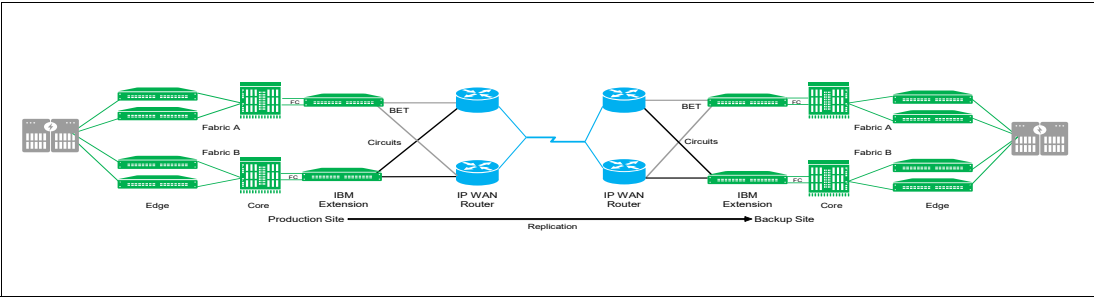


Figure 7-8 FCR+FCIP production edge fabrics extension architecture

Direct connectivity of end devices provides straightforward operations, simplicity, and easier troubleshooting for the Extension network and the production network. Therefore, this configuration is recommended only in instances that require fabric connectivity.

A common b-type Extension use case is sending distributed (Open Systems) tape traffic over distance to a remote data center. Distributed tape uses the SAN to connect many local servers over distance through b-type Extension to tape infrastructure in a remote location. A feature that is known as Open Systems Tape Pipelining (OSTP) within b-type Extension platforms is enabled to maintain performance over distance.

Common device: Not a recommended architecture

In environments that require Extension that is attached to a production SAN, it is *not* a best practice to interconnect the same Extension platform to the “A” and “B” fabrics.

Instead, a best practice is to use two separate and redundant fabrics in a production environment, especially if the organization might suffer financial losses during a SAN outage. Even a momentary SAN outage can “blue screen” or hang servers, which forces a reboot and database reconciliation that in most situations takes significant time.

For maximum availability, it is best practice to divide a SAN into “A” and “B” fabrics, which implies an “air gap” exists between the two autonomous fabrics from the server to the storage. Intentionally, no physical links are used between the two fabrics. Servers, storage, and VMs are equipped with drivers that monitor pathways and send data accordingly. When a path is detected as down, the driver fails traffic over to a remaining path.

Connecting Extension by using FCR to both production edge fabrics (see Figure 7-9) is not recommended. Without FCR, the fabrics merge into one large fabric, which destroys any notion of redundant autonomous fabrics.

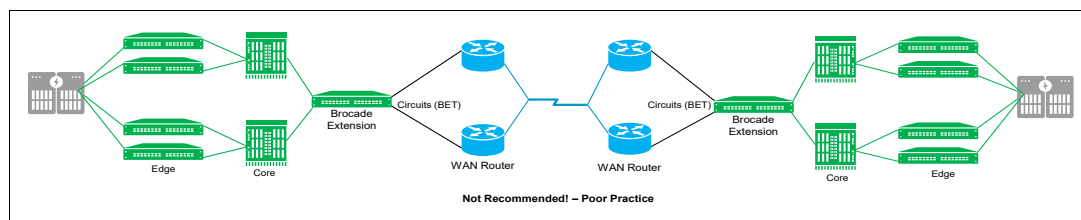


Figure 7-9 Common extension to production edge fabrics (not recommended)

If FCR is used, the fabrics do not merge; however, a common Linux kernel exists on a device that is attached to both fabrics. If maximum availability is the goal in a critical environment, this architecture is not acceptable and considered high risk.

An architecture with a common platform to both fabrics is susceptible to an operating system defect, driver defects, unexpected hardware failures, environmental conditions, and human error, which can bring down the entire SAN.

When connecting Extension to production fabrics, each production fabric must be designed by using best practice core-edge concepts. Because storage attaches directly to the core in a core-edge design, Extension switches connect directly to the core, or an Extension blade is placed in the Core Directors. A stand-alone Extension platform must be connected to a fabric with at least two ISLs (Inter-Switch Links) for the port, optic, and cable redundancy.

In a four-box architecture, it is unsuitable to make ISL cross-connections between the two Extension platforms and the “A” and “B” fabrics.

Cross-connecting circuits from a tunnel to Ethernet DC LAN switches or network devices is encouraged. Circuits traversing the IP network are point-to-point and can take alternative resilient and redundant paths without merging fabrics.

7.2.2 IP Extension design and architectures

In this section, we discuss various architectures that are supported with IP Extension. IBM Extension plays a large role in IBM TS7700 Grid and TCT architectures. The other architectures are generic and can be used in mixed or managed environments.

IBM TS7700 grid and TCT

IP Extension is often used in IBM TS7700 Grid environments. This section addresses design considerations that are specific to these environments.

A consideration in deploying IP Extension with IBM TS7700 Grid is how to direct grid traffic to the IP Extension gateway; the following methods are available:

- ▶ The default gateway is used on the EN (Ethernet Network) interfaces to send traffic directly to the IP Extension gateway.
- ▶ Traffic is sent to the traditional gateway and routes on the gateway router forwards traffic to the IP Extension gateway.

Four subnets must be used: one for each EN# interface on the TS7700 system. Ethernet Network (EN) interfaces of the same number must communicate with each other. EN interfaces of different numbers do not communicate with each other; therefore, as shown in the following examples, each subnet must be in its own VLAN and have its own IP Extension gateway:

- ▶ EN0:
 - IP addresses:10.10.0.100-199
 - Gateway (IPEX):10.10.0.254
 - Subnet Mask:255.255.255.0 (/24)
- ▶ EN1:
 - IP addresses:10.10.1.100-199
 - Gateway (IPEX):10.10.1.254
 - Subnet Mask:255.255.255.0 (/24)

Example subnets and IPEX_GWs:

- ▶ Subnet used for all EN0 in the data center: 10.10.0.0/24 IPEX-GW 10.10.0.254 VLAN 100
- ▶ Subnet used for all EN1 in the data center: 10.10.1.0/24 IPEX-GW 10.10.1.254 VLAN 101
- ▶ Subnet used for all EN2 in the data center: 10.10.2.0/24 IPEX-GW 10.10.2.254 VLAN 102
- ▶ Subnet used for all EN3 in the data center: 10.10.3.0/24 IPEX-GW 10.10.3.254 VLAN 103

Depending on the architecture that is built, IP Extension gateways can be consolidated or distributed:

- ▶ All IP Extension gateways on the same DP: Highest risk
- ▶ IP Extension gateways that are divided across two DPs:
 - On the same blade: High risk
 - On different blades but the same chassis: Medium risk
- ▶ IP Extension gateways that are divided across four DPs:
 - All different blades but two chassis: Medium risk
 - All different blades but four chassis: Lowest risk

The second DP can be kept in reserve for eHCL that is moving circuits into or used by different applications, such as RDR.

It is imperative to understand that TCL rules live only on the DP onto which the target VE_Port exists. VE_Ports exist only in one DP (DP0 or DP1). Refer to the TCL section for IP Extension gateways that are created on DP0 or DP1.

Ingress traffic that matches a TCL rule is directed to a specific target (VE_Port). If the VE_Port target that is designated in a TCL rule is not congruent with the IP Extension gateway's DP, the rule is not evaluated. That is, the IP Extension gateway *and* target VE_Port used must be on the *same* DP for the rule to be evaluated. If the rule is expected to be assessed but is not, check to determine whether traffic is hitting an IP Extension gateway that is configured on the other DP.

Method 1: Default gateway is used on the EN

The IBM TS7700 Grid EN interfaces are configured with their default gateway set to the IP Extension gateway, as shown in Figure 7-10. The IP Extension gateway becomes the default gateway for the EN interfaces.

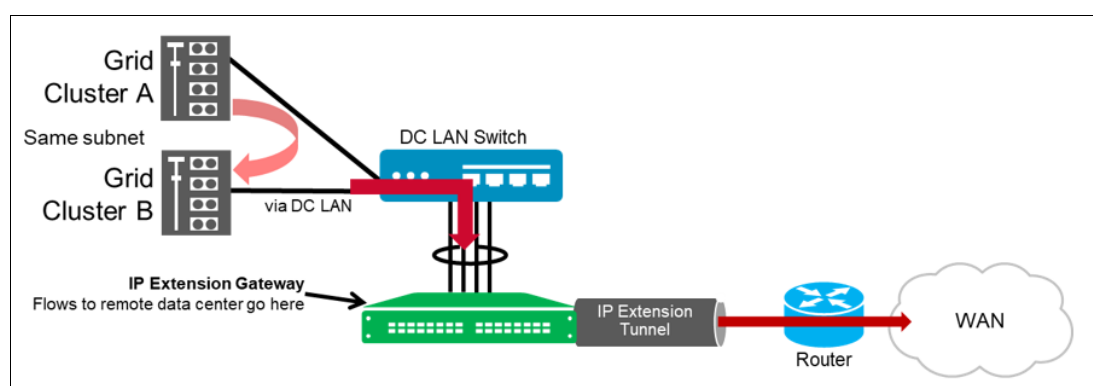


Figure 7-10 Remote grid connectivity by way of IP extension gateway

IBM TS7700 Grid EN interfaces do not support independent routing tables; each interface cannot specify an IP route for destination subnet specific gateways. The IP Extension gateway is required for transporting grid traffic to the remote data center by using IP Extension.

The remote grid subnet is the only destination that is outside of the local grid subnet. In this method, the EN interfaces cannot be configured with the traditional router gateway.

A LAN switch that switches traffic without a gateway router requires EN interfaces to be on the same subnet and VLAN, which is Layer-2 switching (L2 switching). The Extension platforms are not Ethernet switches and cannot facilitate communication between different clusters; the data center LAN switch must do this.

Communication with other grid nodes or clusters must be done at layer 2 (the Ethernet layer) through their common subnet, which does not involve a gateway router. Traffic on the same subnet is sent directly to the other device by way of Ethernet.

If all the local grid EN# interfaces live on the same subnet and VLAN, they can communicate by way of L2 switching, as shown in Figure 7-11. Grid traffic that is destined for a remote data center by way of IP Extension is sent to the IP Extension gateway by using the gateway that is configured on the EN# interfaces.

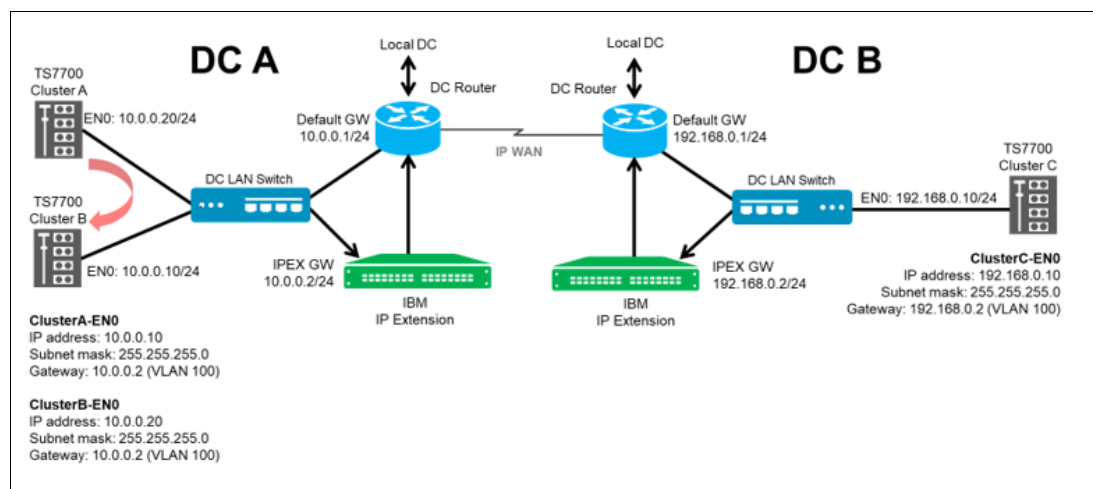


Figure 7-11 EN Interfaces with IP Extension Gateway

Method 2: Traffic is sent to the traditional gateway

In this method, the IBM TS7700 EN interfaces are configured to use the traditional router as the default gateway. Grid traffic that is destined for various subnets is sent to a conventional gateway router. The router forwards the traffic towards the destination, which can be another node or cluster within the local data center. It can also be a remote cluster by way of the IP Extension gateway, as shown in Figure 7-12.

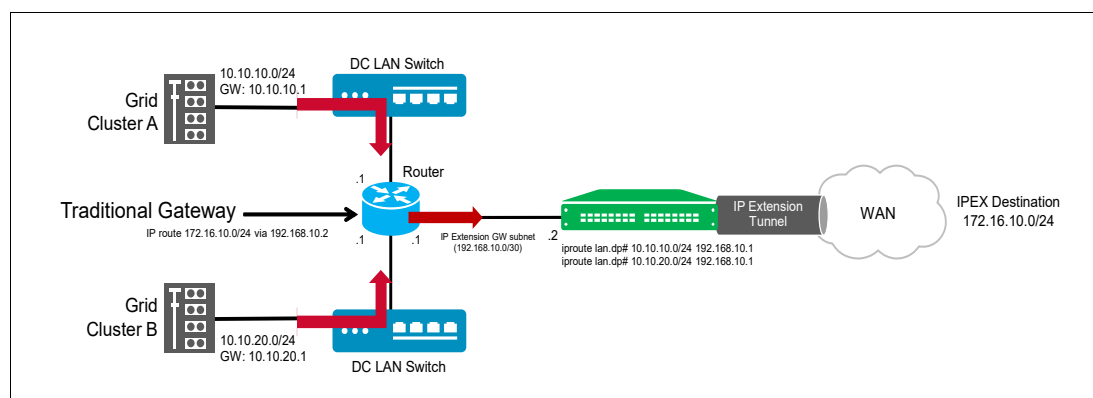


Figure 7-12 Local and remote grid connectivity by way of a traditional gateway

This method enables communications between multiple clusters by using different subnets; however, it does require the IP network to route traffic that originates from the local grid subnet and is destined for the remote grid subnet by way of the IP Extension gateway.

The traditional router can be configured to route to the IP Extension gateway in various ways, such as connected interface routing protocol participation, static routes, or Policy Based Routing (PBR). IBM b-type Gen 7 Extension platforms are not IP routers and do not support routing protocols. Different grid subnets between nodes and clusters use the default gateway that is designated on the local data center's EN interfaces.

IBM TS7700 grid three DC architecture

The three data center IBM TS7700 Grid architecture spans three data centers, and IP Extension is used between two metro data centers and the remote DR site. This three data center grid architecture uses the SAN42B-R or SX6 Blades. Both DPs are used on each Extension platform (DP0 and DP1), although not necessarily both for each EN interface.

Each DP has a capacity of 20 Gbps when licensed. The DR data center features a pair of Extension platforms for each incoming site. If Sx6 Extension Blades are used, a second blade is required in each chassis. If the SAN42B-R is used, another box is required. The diagram and Extension icons are generic; two Directors with two SX6 blades in each or four SAN42B-R boxes.

IBM TS7700 Grid traffic is compressed; therefore, compression is disabled. With no compression, the ingress and egress data rates are equal at 20 Gbps for IP Extension. Encryption is enabled and causes no degradation in throughput or performance.

A different subnet is used for each EN interface. Each cluster has four EN interfaces (EN0, EN1, EN2, and EN3). DC-A and DC-B are metro data centers that are connected by DWDM at L2, the Ethernet layer.

The respective EN interfaces from clusters (4, 6, 5, and 7) are on the same subnets across these two data centers. All EN0 interfaces can communicate without routing (L3), as with EN1, EN2, and EN3. In DC-A and DC-B, all like numbered EN interfaces can communicate with each other without Extension. DC-C is further away and requires a routed L3 IP network and Extension.

Figure 7-13 shows the schema for addressing the EN interfaces, which creates gateways and connects the various DPs. The gateways and VE_Ports are associated with one specific DP. EN interface IP addresses were selected to readily designate the data center (third byte), EN interface (third byte), and cluster (fourth byte) within and across wanted broadcast (L2) domains. The subnets are /24 (255.255.255.0). The layout depends on the number of DPs in each data center and how best to disperse EN interfaces across the DPs to gain bandwidth and high availability.

CL/EN	Source IP	Destination IP	IPEX GW	DP	Target	SW	SW	Target	DP	IPEX GW	Source IP	Destination IP	CL/EN	# of Cirs	Min-Rate Each (Gbps)	Max-Rate Each (Gbps)
CL4																
0	10.0.0.40	10.0.10.10	10.0.0.250/24-dp0	DP0	12/16	DCA-sw0	DCC-sw0	12/16	DP0	10.0.10.250/24-dp0	10.0.10.10	10.0.0.40	0	2	5	5
1	10.0.1.40	10.0.11.10	10.0.1.250/24-dp0	DP0	12/16	DCA-sw1	DCC-sw1	12/16	DP0	10.0.11.250/24-dp0	10.0.11.10	10.0.1.40	1	2	5	5
2	10.0.2.40	10.0.12.10	10.0.2.251/24-dp1	DP1	12/26	DCA-sw0	DCC-sw2	12/16	DP0	10.0.12.250/24-dp0	10.0.12.10	10.0.2.40	2	2	5	5
3	10.0.3.40	10.0.13.10	10.0.3.251/24-dp1	DP1	12/26	DCA-sw1	DCC-sw3	12/16	DP0	10.0.13.250/24-dp0	10.0.13.10	10.0.3.40	3	2	5	5
CL5																
0	10.0.0.50	10.0.10.10	10.0.0.252/24-dp0	DP0	12/16	DCB-sw0	DCC-sw0	12/17	DP0	10.0.10.250/24-dp0	10.0.10.10	10.0.0.50	0	2	5	5
1	10.0.1.50	10.0.11.10	10.0.1.252/24-dp0	DP0	12/16	DCB-sw1	DCC-sw1	12/17	DP0	10.0.11.250/24-dp0	10.0.11.10	10.0.1.50	1	2	5	5
2	10.0.2.50	10.0.12.10	10.0.2.253/24-dp1	DP1	12/26	DCB-sw0	DCC-sw2	12/17	DP0	10.0.12.250/24-dp0	10.0.12.10	10.0.2.50	2	2	5	5
3	10.0.3.50	10.0.13.10	10.0.3.253/24-dp1	DP1	12/26	DCB-sw1	DCC-sw3	12/17	DP0	10.0.13.250/24-dp0	10.0.13.10	10.0.3.50	3	2	5	5
CL6																
0	10.0.0.60	10.0.10.10	10.0.0.250/24-dp0	DP0	12/16	DCA-sw0	DCC-sw0	12/16	DP0	10.0.10.250/24-dp0	10.0.10.10	10.0.0.60	0	2	5	5
1	10.0.1.60	10.0.11.10	10.0.1.250/24-dp0	DP0	12/16	DCA-sw1	DCC-sw1	12/16	DP0	10.0.11.250/24-dp0	10.0.11.10	10.0.1.60	1	2	5	5
2	10.0.2.60	10.0.12.10	10.0.2.251/24-dp1	DP1	12/26	DCA-sw0	DCC-sw2	12/16	DP0	10.0.12.250/24-dp0	10.0.12.10	10.0.2.60	2	2	5	5
3	10.0.3.60	10.0.13.10	10.0.3.251/24-dp1	DP1	12/26	DCA-sw1	DCC-sw3	12/16	DP0	10.0.13.250/24-dp0	10.0.13.10	10.0.3.60	3	2	5	5
CL7																
0	10.0.0.70	10.0.10.10	10.0.0.252/24-dp0	DP0	12/16	DCB-sw0	DCC-sw0	12/17	DP0	10.0.10.250/24-dp0	10.0.10.10	10.0.0.70	0	2	5	5
1	10.0.1.70	10.0.11.10	10.0.1.252/24-dp0	DP0	12/16	DCB-sw1	DCC-sw1	12/17	DP0	10.0.11.250/24-dp0	10.0.11.10	10.0.1.70	1	2	5	5
2	10.0.2.70	10.0.12.10	10.0.2.253/24-dp1	DP1	12/26	DCB-sw0	DCC-sw2	12/17	DP0	10.0.12.250/24-dp0	10.0.12.10	10.0.2.70	2	2	5	5
3	10.0.3.70	10.0.13.10	10.0.3.253/24-dp1	DP1	12/26	DCB-sw1	DCC-sw3	12/17	DP0	10.0.13.250/24-dp0	10.0.13.10	10.0.3.70	3	2	5	5
CL4																
0	10.0.0.40	10.0.10.30	10.0.0.250/24-dp0	DP0	12/17	DCA-sw0	DCC-sw0	12/26	DP1	10.0.10.251/24-dp1	10.0.10.30	10.0.0.40	0	2	5	5
1	10.0.1.40	10.0.11.30	10.0.1.250/24-dp0	DP0	12/17	DCA-sw1	DCC-sw1	12/26	DP1	10.0.11.251/24-dp1	10.0.11.30	10.0.1.40	1	2	5	5
2	10.0.2.40	10.0.12.30	10.0.2.251/24-dp1	DP1	12/27	DCA-sw0	DCC-sw2	12/26	DP1	10.0.12.251/24-dp1	10.0.12.30	10.0.2.40	2	2	5	5
3	10.0.3.40	10.0.13.30	10.0.3.251/24-dp1	DP1	12/27	DCA-sw1	DCC-sw3	12/26	DP1	10.0.13.251/24-dp1	10.0.13.30	10.0.3.40	3	2	5	5
CL5																
0	10.0.0.50	10.0.10.30	10.0.0.252/24-dp0	DP0	12/17	DCB-sw0	DCC-sw0	12/27	DP1	10.0.10.251/24-dp1	10.0.10.30	10.0.0.50	0	2	5	5
1	10.0.1.50	10.0.11.30	10.0.1.252/24-dp0	DP0	12/17	DCB-sw1	DCC-sw1	12/27	DP1	10.0.11.251/24-dp1	10.0.11.30	10.0.1.50	1	2	5	5
2	10.0.2.50	10.0.12.30	10.0.2.253/24-dp1	DP1	12/27	DCB-sw0	DCC-sw2	12/27	DP1	10.0.12.251/24-dp1	10.0.12.30	10.0.2.50	2	2	5	5
3	10.0.3.50	10.0.13.30	10.0.3.253/24-dp1	DP1	12/27	DCB-sw1	DCC-sw3	12/27	DP1	10.0.13.251/24-dp1	10.0.13.30	10.0.3.50	3	2	5	5
CL6																
0	10.0.0.60	10.0.10.30	10.0.0.250/24-dp0	DP0	12/17	DCA-sw0	DCC-sw0	12/26	DP1	10.0.10.251/24-dp1	10.0.10.30	10.0.0.60	0	2	5	5
1	10.0.1.60	10.0.11.30	10.0.1.250/24-dp0	DP0	12/17	DCA-sw1	DCC-sw1	12/26	DP1	10.0.11.251/24-dp1	10.0.11.30	10.0.1.60	1	2	5	5
2	10.0.2.60	10.0.12.30	10.0.2.251/24-dp1	DP1	12/27	DCA-sw0	DCC-sw2	12/26	DP1	10.0.12.251/24-dp1	10.0.12.30	10.0.2.60	2	2	5	5
3	10.0.3.60	10.0.13.30	10.0.3.251/24-dp1	DP1	12/27	DCA-sw1	DCC-sw3	12/26	DP1	10.0.13.251/24-dp1	10.0.13.30	10.0.3.60	3	2	5	5
CL7																
0	10.0.0.70	10.0.10.30	10.0.0.252/24-dp0	DP0	12/17	DCB-sw0	DCC-sw0	12/27	DP1	10.0.10.251/24-dp1	10.0.10.30	10.0.0.70	0	2	5	5
1	10.0.1.70	10.0.11.30	10.0.1.252/24-dp0	DP0	12/17	DCB-sw1	DCC-sw1	12/27	DP1	10.0.11.251/24-dp1	10.0.11.30	10.0.1.70	1	2	5	5
2	10.0.2.70	10.0.12.30	10.0.2.253/24-dp1	DP1	12/27	DCB-sw0	DCC-sw2	12/27	DP1	10.0.12.251/24-dp1	10.0.12.30	10.0.2.70	2	2	5	5
3	10.0.3.70	10.0.13.30	10.0.3.253/24-dp1	DP1	12/27	DCB-sw1	DCC-sw3	12/27	DP1	10.0.13.251/24-dp1	10.0.13.30	10.0.3.70	3	2	5	5

Figure 7-13 Cluster IP addresses, IP Extension Gateways, Tunnels, and DP Assignments

On each switch in DC-C (sw0, sw1, sw2, and sw3), there are two IP Extension gateways on each subnet, one for each DP (.250 and .251). There are four IP Extension gateways on each subnet in DC-A and DC-B, one for each DP in each data center (.250, .251, .252, and .253). You do not want to traverse the DWDM network to reach an IP Extension gateway.

In DC-A and DC-B, EN0 and EN2 use DP0, and EN1 and EN3 use DP1. In DC-A and DC-B, EN0 and EN1 also use different switches; EN2 and EN3 use separate switches, for example, DCA-sw0 and DCA-sw1. In DC-C, each EN interface uses a different switch because four platforms are used. In DC-C, cluster 1 uses DP0, and cluster 3 uses DP1.

Two tunnels are used per DP. Each tunnel has two circuits, one circuit for each data center LAN switch. If one of the data center LAN switches goes offline, the remaining circuit continues to operate without disrupting the grid.

Four circuits are available per DP. The maximum capacity of a DP is approximately 20 Gbps; therefore, each circuit should be configured at 5 Gbps, which is 10 Gbps per tunnel. The **--min-comm-rate** and the **--max-comm-rate** must be set equal at 5 Gbps, which is called the *committed information rate (CIR)*.

When the minimum and maximum rates are different, it is referred to as *ARL*. In a grid environment with well-balanced traffic across all paths, disabling ARL and using CIR can improve performance. Moreover, if metric 1 backup circuits and failover groups are used, the best practice is CIR over ARL.

Figure 7-14 shows a high-level connectivity architecture with ports.

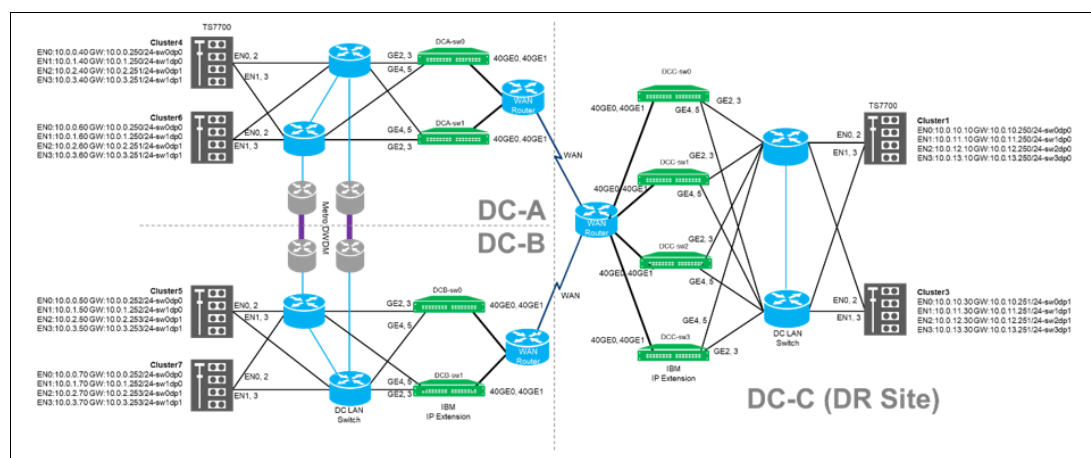


Figure 7-14 TS7700 grid three DC architecture

Consider tunnels as connecting a local IP Extension gateway to a remote IP Extension gateway. IP Extension gateways belong to a specific DP (DP0 or DP1). Think of the local and remote IP Extension gateways as endpoints of a tunnel.

Considering that the TCL directs traffic into a specific tunnel, an IP Extension gateway can be the endpoint of more than one tunnel within the same DP. Traffic must traverse the same path outbound and return; therefore, the same gateway pair must be deterministic. The TCL ensures a deterministic path (gateways and tunnel) is used when outbound and return traffic consistently arrives at the lan.dpx ipif.

It helps for planning purposes to make a diagram of how the tunnels connect from end-to-end; that is, IP Extension gateway to IP Extension gateway (see Figure 7-15).

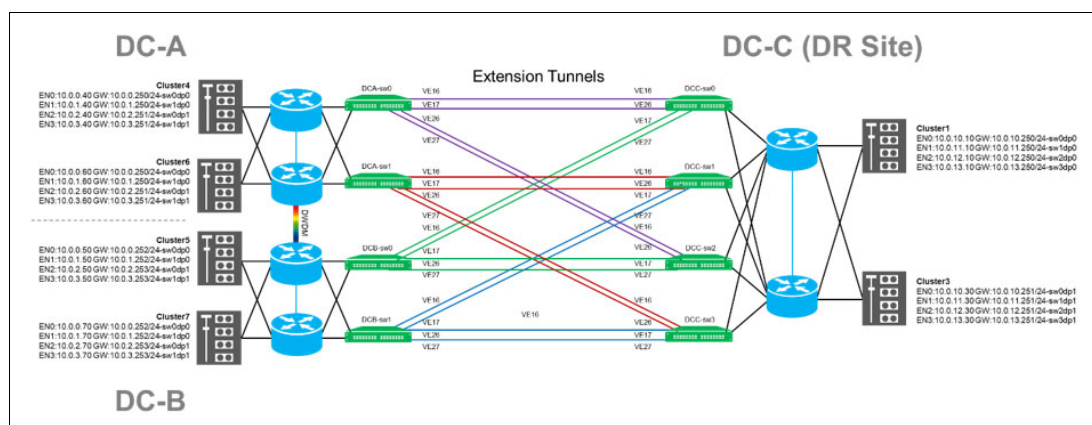


Figure 7-15 S7700 grid three DC tunnel schema

The Extension tunnels between DC-A and DC-B to DC-C complete a full mesh architecture in which every EN interface of the same number can communicate. Clusters 4, 6, 5, and 7 can communicate through Layer2 because their EN interfaces are on the same subnet and connected by DWDM.

Tunnel connectivity is shown in Figure 7-15. Notice the VE_Port range for DP0 and DP1 are different. On the SX6 Blade, DP0 is VE16 to VE20, and DP1 is VE26 to VE30.

The same gateway can direct traffic into multiple tunnels by way of the TCL by matching the source/destination IP addresses. For example, on the Extension platform that is connected to Cluster1-EN0, if the rule matches src:10.0.10.10 and dst:10.0.0.40 and designates target 16, that specific traffic uses tunnel VE16 to Cluster4-EN0. A mirrored TCL must be configured on the Extension platform that is connected to Cluster4 to return traffic.

Multiple gateways can direct traffic into the same tunnel by way of the TCL. For example, an Extension platform uses the gateway 10.0.0.250 for EN0 and directs traffic into VE16. A different gateway for EN2 (10.0.2.250) can direct traffic into the same tunnel (VE16). If Upgrade1 + Upgrade2 licenses are installed, it is possible to create parallel tunnels for each EN interface; however, this configuration assumes that enough VE_Ports are available.

The DP maximum capacity is no more than 20 Gbps on the WAN side, no matter how much circuit bandwidth is configured. With WAN Upgrade 1+2 licenses, bandwidth is not restricted on creating circuits. The best practice is to use the same tunnels for various gateways when a tunnel is between the same domains and DPs.

DC-C dedicates a DP to each EN interface. DC-C has four switches (sw0, sw1, sw2, and sw3) because it consolidates incoming connections from two remote data centers. In DC-A and DC-B, a DP is dedicated to EN0 and EN2, and a DP is dedicated to EN1 and EN3.

A balance is struck between the two metro sites that are going into the one DR site. Adding SX6 Blades or SAN42B-R boxes increases the number of DPs that is available to EN interfaces; for example, if more bandwidth was needed. One DP can accommodate 20 Gbps, which is twice the bandwidth of a single EN interface. More than one DP per EN interface is not beneficial.

Two IBM SAN18B-6 (base unit) architecture

The IBM SAN18B-6 base unit supports one circuit per VE_Port (tunnel). The base unit does not support BET. Two VE_Ports are supported; therefore, two parallel tunnels between the same domains is an option. Two parallel tunnels are different than one tunnel with two circuits. Two autonomous tunnels have no coordination or bandwidth management between them.

The TCL cannot match IP Extension traffic and direct it into more than one target (tunnel). If two VE_Ports were used, the traffic across each requires a rule match from different TCL rules.

Two IP Extension fabrics are shown in Figure 7-16.

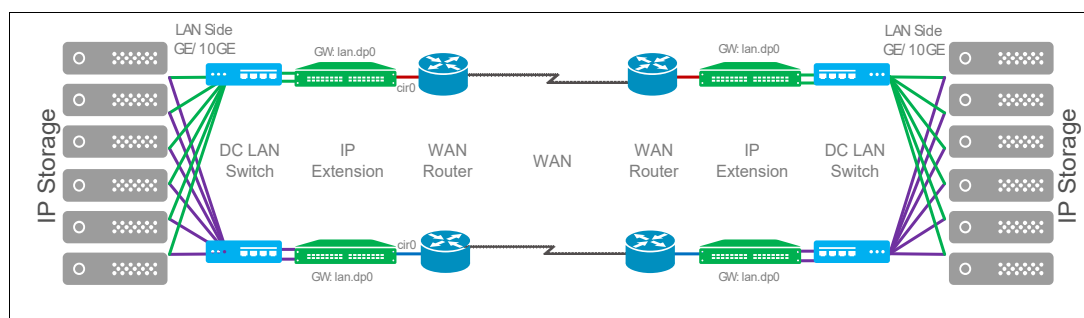


Figure 7-16 IBM SAN18B-6 base unit architecture

Each fabric has one tunnel with one circuit. The maximum bandwidth that is supported by the IBM SAN18B-6 base unit is 1 Gbps on the WAN side. Two WAN connections also are shown; however, a single WAN connection also works.

The GE interfaces can be GE or 10GE. Two different IP Extension gateways are required, and the end devices must configure the different gateways per interface or per set of interfaces.

Two IBM SAN18B-6 (full unit) architecture

The IBM SAN18B-6 full unit supports four VE_Ports, 2.5 Gbps maximum WAN side bandwidth, and BET. In Figure 7-17, two circuits are shown (red and blue). Each circuit takes a different path through the IP network. Each circuit has its settings for bandwidth, failover, QoS, and KATOV.

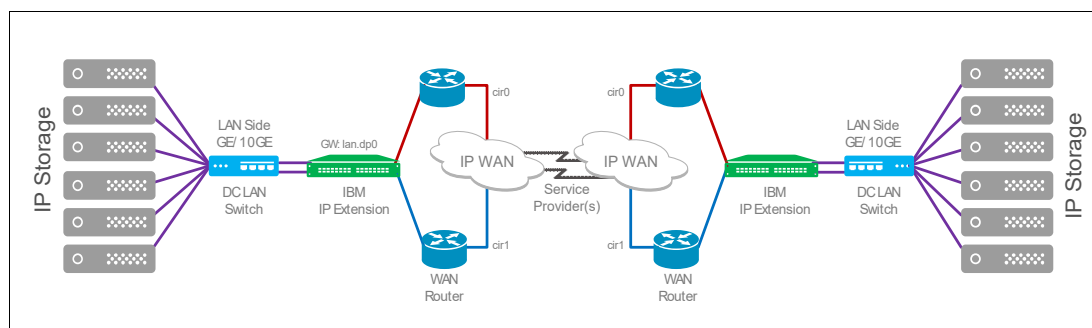


Figure 7-17 Two IBM SAN18B-6 full unit architecture

Two IBM IP extension with two non-virtualized DC LAN switches

Not all data center LAN switches are configured to form a single logical switch. The architecture that is shown in Figure 7-18 shows IP storage that is dual connected to each data center LAN switch; there are two LANs. A single IP Extension connection to each LAN (broadcast domain), which can be a LAG. Each LAN must have its own subnet and IP Extension gateway.

Two circuits (red and blue) make up a single tunnel between Extension platforms. Each circuit takes a different path through the IP network.

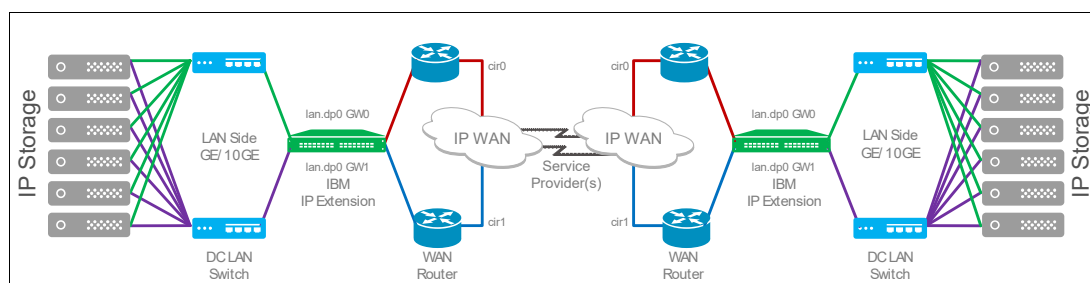


Figure 7-18 Two IP extension platforms with non-virtualized DC LAN switches

Two IBM Extension platforms with virtualized DC LAN switches

The data center LAN switches in this architecture are virtualized, meaning they form a single logical switch (see Figure 7-19).

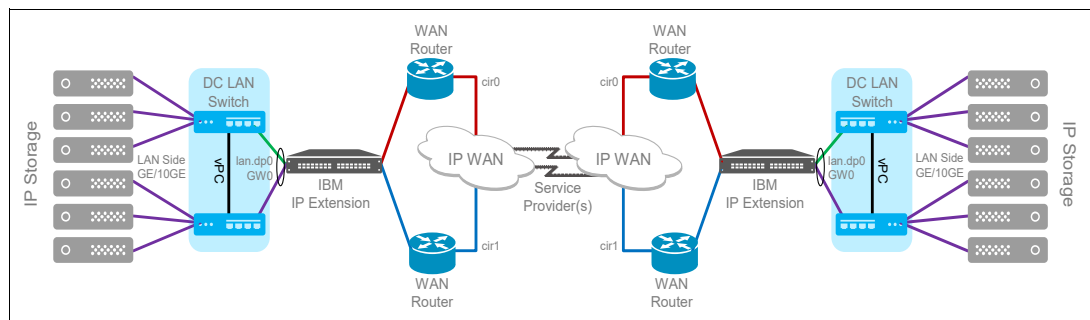


Figure 7-19 Two IP Extension Platforms with Virtualized DC LAN Switches

Two switches creating one virtual switch is often referred to as vPC or VSS in the Cisco nomenclature. There are connections from the IP Extension platform to both switches because the switches are logically one. It is perceived as a single Link Aggregated (LAG) connection, as indicated by the circles in Figure 7-19.

The links can be tagged or untagged and must be consistent on each end. There is a single IP Extension gateway for the IP storage VLAN. The end devices can connect to either one of the data center LAN switches or dual connect to both.

On the WAN side, there are two circuits (red and blue) comprising one tunnel between the Extension platforms. The circuits take different paths for redundancy.

Four IBM Extension platforms with non-virtualized DC LAN switches

For a higher degree of availability, resiliency, and bandwidth capacity, a four Extension platform architecture could be deployed (see Figure 7-20).

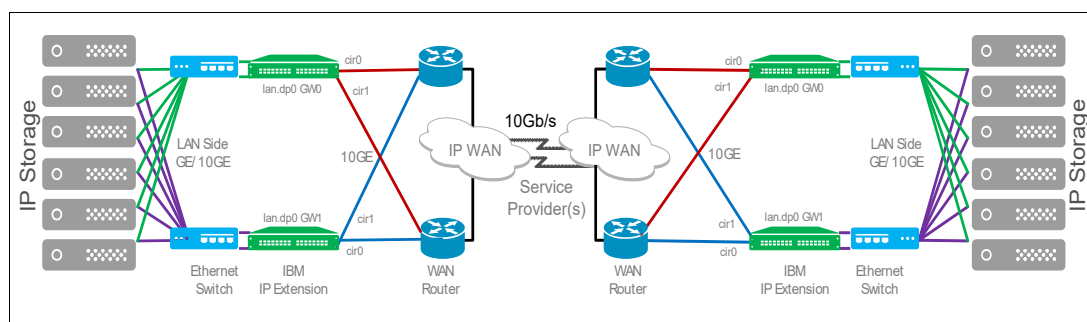


Figure 7-20 Four IP extension platforms with non-virtualized DC LAN switches

The end devices are dual connected to separate LAN switches; each is a LAN (broadcast domain). Each connection (purple and green) requires its own unique subnet and IP Extension gateway. The end devices must independently designate a gateway per interface or per set of interfaces.

The Extension platforms are LAG connected to a corresponding data center LAN switch. The links can be tagged or untagged and must be consistent on both ends.

There is a single tunnel between the top two Extension platforms (red tunnel) and one between the bottom two (blue tunnel). Each tunnel has two circuits (cir0 and cir1), each taking a different path.

Four IBM Extension platforms with virtualized DC LAN switches

For a higher degree of availability, resiliency, and bandwidth capacity, a four Extension platform architecture can be deployed (see Figure 7-21).

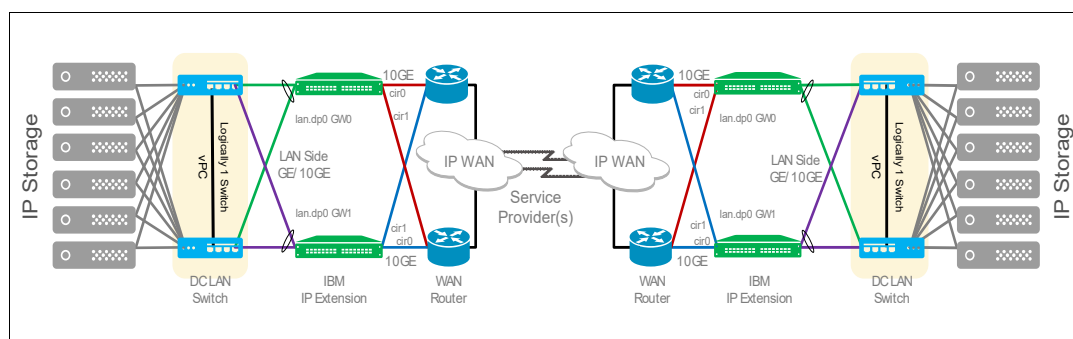


Figure 7-21 Four IP extension platforms with virtualized DC LAN switches

Each end device is dual connected to a single logically virtualized LAN switch, which has two VLANs (two broadcast domains). Each connection (purple and green) requires its own unique subnet and IP Extension gateway. The same IP Extension gateway cannot be put on two Extension platforms, and Virtual Router Redundancy Protocol (VRRP) is not supported. The end devices must independently designate a gateway per interface or per set of interfaces.

The Extension platforms are LAG connected across the two virtualized data center LAN switches. The links can be tagged or untagged and must be consistent on both ends.

There is a single tunnel between the top two Extension platforms (red tunnel) and a tunnel between the bottom two Extension platforms (blue tunnel). Each tunnel has two circuits (cir0 and cir1), each taking a different path.

7.3 Extension implementation

Extension platforms are considered to have three sides: the WAN side, FC/FICON side, and LAN side. This section discusses the various extension platforms, best practices, features and functions, monitoring, and architectures.

7.3.1 WAN side

Although the FC/FICON side and LAN side can be used, the WAN side must be used when doing Extension. The WAN side connects to an IP network and forms an Extension tunnel between data centers. The WAN side includes features, functions, and configurations that are specific to creating the tunnel, routing traffic, encryption, compression, and protocol optimization. Also, the WAN side must manage with the IP network and WAN.

IP network

The IP network must transport Extension datagrams with adequate performance and reliability to meet the service level goal. The wanted service level depends on the policy that is set forth by the organization. Acceptable performance depends on the following factors:

- ▶ Amount of data that must be sent within an interim period.
- ▶ The WAN's capacity to accommodate the sent data.
- ▶ The IP network's ability to deliver the data without loss.

Reliability depends on the stability of the IP network, its overall architecture, redundancy and resiliency, and its ability to manage and deliver data across the WAN.

IBM does *not* support IP network devices with oversubscribed or blocking switch ports. Some Ethernet switches include oversubscribed or blocking host access ports, which can cause excessive drop packets and degrade storage throughput.

The WAN

Any type and size of WAN connection can be used for Extension if it meets the application and extension platform requirements. The storage application might not have more stringent requirements than the Extension platforms.

Most commonly, Extension is deployed between data centers, separated by a distance of 2 ms RTT or more. The supported propagation delay across an IP network is up to 200 ms RTT, which in most cases, reaches anywhere in the world.

IBM supports no more than 0.1% packet loss between the Extension Ethernet ports. By modern standards, 0.1% is considered significant packet loss. Packet loss beyond this level indicates an evaluation and potential remediation of the IP WAN.

It is important to understand that the effective bandwidth of a WAN link is highly dependent on latency and link quality in terms of packet loss. Even a small amount of packet loss with resulting retransmits significantly degrades overall link effective bandwidth.

All IBM Extension platforms have a minimum bandwidth of 20 Mbps per circuit.

The IBM SAN18B-6 can have a maximum circuit size of 2.5 Gbps if licensed as a full unit; otherwise, 1 Gbps as a base unit. VE_Ports can accommodate up to 2.5 Gbps of cumulative circuits as a full unit; otherwise, 1 Gbps as a base unit. On the full unit, a tunnel can be composed of multiple circuits of varying sizes up to six circuits; otherwise, only a single circuit per VE_Port as a base unit.

The IBM SAN42B-R has a maximum circuit bandwidth of 10 Gbps if the Upgrade 1 license is applied; otherwise, 5 Gbps as a base unit. VE_Ports can accommodate up to 20 Gbps of cumulative circuits if the Upgrade1 + Upgrade2 licenses are applied; otherwise, 10 Gbps for Upgrade1 and 5 Gbps as a base unit. A tunnel is composed of multiple circuits of varying sizes, up to 10 circuits.

IBM SX6 Extension Blade has a maximum circuit bandwidth of 10 Gbps. VE_Ports can accommodate up to 20 Gbps of cumulative circuits. Upgrade1 and Upgrade2 licenses do not apply to this blade. A tunnel can be composed of multiple circuits of varying sizes, up to 10 circuits.

A best practice is to connect Extension directly to the WAN routers or switches, which avoids intermediate devices, such as firewalls. If this connection is not possible, every effort should be made to connect the Extension Ethernet ports as close to the WAN as possible to prevent upstream over-subscription and performance degradation. Extension traffic drives large amounts of data, and uplink over-subscription often bodes poorly.

Consider the following points when you are planning for your Extension links:

- ▶ For redundancy, use at least two WAN links between sites; however, in critical environments deploying two or more replication fabrics, consider a WAN connection for each fabric. Generally, WAN connections are the weakest and least reliable part of replication fabrics.
- ▶ Consider dedicated WAN connections for storage interconnections. Separate storage and non-storage traffic to better manage bandwidth, latency, and prevent congestion.
- ▶ Ensure that you have a service level agreement (SLA) with the WAN service provider that meets the enterprise's needs and expectations.
- ▶ If Global Mirror with Change Volumes (GMCV) is not used, make sure to size the WAN connection to sustain peak workloads. For replication over long distances, the best practice is to use GMCV.

Bandwidth

For critical Remote Data Replication (RDR), the best practice is to use a separate and dedicated IP connection between the production data center and the backup site. Nonetheless, often a dedicated IP connection between data centers is not practical. In this case, bandwidth must be, at a minimum, logically dedicated to Extension, which can be done by using one of the following methods:

- ▶ Use QoS, which gives Extension a higher priority and logically gives shared bandwidth to Extension over other competing traffic.
- ▶ Use Committed Access Rate (CAR) to identify and rate-limit certain traffic types. Use CAR on non-Extension traffic to apportion and limit that traffic to a maximum bandwidth, which leaves the remaining bandwidth to Extension. Set the aggregate of the circuit min-comm-rates to use the remaining dedicated bandwidth. Restricting non-Extension traffic results in logically allocating bandwidth for Extension.
- ▶ Use massive over-provisioning to provide bandwidth for all traffic to coexist without congestion. This method is the easiest, most costly, and most common practice seen in Extension deployments.

Brocade Extension uses an aggressive TCP stack that is called WAN Optimized TCP (WO-TCP), which dominates other TCP flows on the path and causes them to dramatically back off. UDP-based flows can result in considerable congestion and excessive packet drops for all traffic.

The best practice is to rate-limit Extension traffic on the platforms at the source and destination and not rate limit Extension in the IP network. Rate limiting Extension in the IP network often leads to performance problems and difficult troubleshooting issues. IBM Extension rate limiting is advanced, accurate, and consistent; there is no need to double the rate limit.

To determine the amount of network bandwidth that is needed, record the number of bytes that are written over at least a month. A granular recording with the ability to go back and calculate rolling averages of varying lengths is most helpful. It is essential to understand the number of bytes that are written to volumes during interims of the day and night. Write profiles show the quantities of data that must be replicated to the maintain Recovery Point Objective (RPO). Expediting data transmission is essential.

If replicating asynchronous RDR (RDR/A) across a tunnel, calculate the average value over a finite number of minutes throughout the day and night to determine the maximum average. Keep in mind, business transactions can increase, and replication might require significantly more bandwidth during quarter-end, fiscal-end, and specific holidays.

RDR/A needs only enough bandwidth to accommodate the high averages that are discovered over a finite period. RDR/A essentially performs traffic shaping, which moves peaks into troughs and averages over interims that are too long causes a backlog of writes when the troughs do not adequately relieve the peaks. A replication backlog, which is data journaling, can result in extended periods of recovery.

For RDR/S (synchronous replication), the best practice is to dedicate the network without sharing bandwidth. A 10 Gbps DWDM lambda is an example of dedicated bandwidth that is often used with synchronous replication.

Dedicated bandwidth eliminates packet drops and timely TCP retransmits. If you plan to replicate RDR/S across an Extension tunnel, record the peak traffic rates. RDR/S must have enough bandwidth to send writes immediately, which accommodates the entire demand at any time.

Plot the recorded values into a histogram. Suppose you calculated 5-minute rolling averages from your recorded data. The number of bytes that is written in each period is an integer value. The x-axis contains the number of bytes written during each 5-minute average. The x-axis starts at zero bytes on the left and goes up to the largest number of bytes that is written during an interim. Many interims include the exact number of bytes that form a Gaussian curve.

The y-axis is the number of times a specific interim occurred having the exact value. The smallest size occurred relatively rarely, and the largest size occurred relatively infrequently. A total of 68% of the averages fell into the middle section, as shown in Figure 7-22, which is the largest number of occurrences.

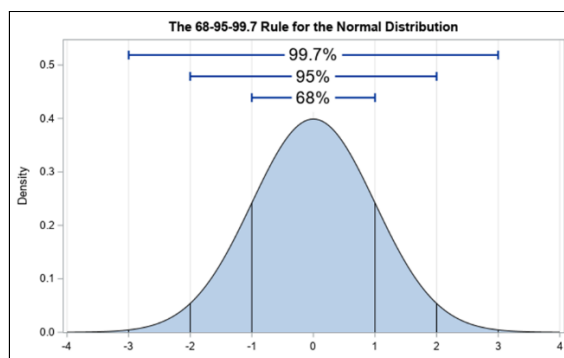


Figure 7-22 Gaussian standard normal distribution

Based on cost, plan your WAN bandwidth to accommodate at least the first standard deviation (68% $\mu \pm 1\sigma$), and if possible, include the second standard deviation (95% $\mu \pm 2\sigma$). In most cases, beyond the second deviation, occurrences are rarer and potentially disregarded in bandwidth planning.

You can plan for a specific amount of compression, such as 2:1 if your data is compressible; however, data compressibility is a moving target. The best practice is to use compression as a safety margin to address unusual and unforeseen situations. Also, achieving some amount of compression can provide headroom for potential future growth. From the start, pushing the limit might not be fortuitous, mainly if no margin accommodates periodic demand increases.

For remote tape, Extension connects VTLs that are in various data centers. It is vital to measure tape volume in MBph and the number of simultaneous drives in use. If enough drives are available, bandwidth might become the limiting factor for backup windows.

GE interfaces

The IBM Extension platforms have various gigabit Ethernet (GE) interface connectivity to the IP network for transporting application data to and from the LAN and WAN sides. GE interfaces are used by application data to communicate to the LAN side or from the WAN side to the long-distance IP network.

All Extension Ethernet interfaces are referred to as *GE interfaces*.

WAN-side Ethernet interfaces do not support LAG; instead, Brocade Extension Trunking (BET) is used. BET accomplishes the same as LAG, plus it has the benefit of not requiring the destination switches to be virtually connected, lossless link loss, and guaranteed in-order delivery.

IBM Extension GE interfaces cannot enable loops within a data center LAN. Extension GE interfaces are akin to server NIC interfaces; they are not Ethernet switch ports, nor are they router ports.

GE interfaces do not participate in routing protocols and do not run spanning tree protocol because there is no need. Ethernet frames are not permitted to “hairpin,” meaning to come back out of an incoming interface to prevent loops. Ingress traffic cannot go out of any other GE interface of the same type; it must go to a different side of the Extension platform, or it is dropped.

Table 7-1 shows the various Ethernet interfaces and VE_Ports that are available on IBM b-type Gen 7 Extension platforms.

Table 7-1 Ethernet interfaces on IBM b-type Gen 7 extension platforms

Item	Brocade 7810	Brocade 7840		SX6 Extension Blade	
Data Processor (DP)	DP0	DP0	DP1	DP0	DP1
GE	6 GE/10GE (2 can be GE RJ-45)	16 GE/10GE shared by both DPs		16 GE/10GE shared by both DPs	
10 GE					
40 GE	n/a	2 shared by both DPs		2 shared by both DPs	
Max VE Bandwidth	2.5 Gbps	20 Gbps	20 Gbps	20 Gbps	20 Gbps
Number of VEs					
Platform max WAN side Bandwidth	2.5 Gbps	40 Gbps		40 Gbps	

Note: II Extension platform Ethernet interfaces are referred to as GE, regardless of their type and speed. The management Ethernet interface is not referred to as a GE port.

All commands include **--help**, as shown in the following example:

```
help portcfgge or portcfgge --help
```

Setting GE Interface Mode

On the IBM SAN18B-6 Extension Switch, two 1 Gbps RJ-45 copper interfaces (GE0 and GE1) are available. These interfaces and the first two optical bays (GE2 and GE3) are mutually exclusive. They enable and disable as a set.

GE0 and GE1 or GE2 and GE3 are enabled, but *not* both. No platform reboot is required to make the change; however, no configuration can be tied to the ports, as shown in the following CLI configuration:

The IBM SAN18B-6 enables the copper (GE0 and GE1) or the optical (GE2 and GE3):

```
extncfg --help
extncfg --ge-mode optical
```

Setting GE Interface Type

When deploying IP Extension, at least one GE interface must be configured in LAN mode, preferably more than one for redundancy. The LAN side GE interfaces are associated with IP storage that is communicating to the IP Extension gateway. The IP Extension gateway is not assigned to specific LAN GE interfaces; instead, the IP Extension gateway is accessed from any LAN interface. If you are planning to use a PortChannel (LAG), multiple LAN interfaces are required.

Two types of GE interfaces are available: WAN and LAN. The default type is WAN, which is shown as “FCIP” in switchshow.

Setting a GE interface to be of type LAN is shown in the following example:

```
7840-Side-A:FD128:admin> portcfgge ge6 --set -lan
```

The configuration of the GE interfaces is shown in the following example:

```
Local7810:admin> portcfgge --show
```

Port	Speed	Flags	Channel	Lag Name
ge0	1G	AC--	N/A	-
ge1	1G	AC--	N/A	-
ge2	10G	--LG	N/A	MYLAG1
ge3	10G	--LG	N/A	MYLAG1
ge4	10G	----	N/A	-
ge5	10G	----	N/A	-
ge6	10G	----	N/A	-
ge7	10G	----	N/A	-

Flags: A:Auto-Negotiation Enabled C:Copper Media Type
L:LAN Port G: LAG Member

Set a GE interface to LAN

```
portcfgge ge2 --set -lan
```

Setting GE Interface Speed

Configure the speed to match both the data center LAN switch ports and the SFP (1 Gbps) or SFP+ (10 Gbps) that is populated in the Extension platform. SFP and SFP+ optics cannot change speed. No speed auto-negotiation exists in GE/10GE ports; this setting is a manual configuration setting.

Setting the GE interface speed is shown in the following example:

```
portcfgge ge2 --set -speed 10G
```

Showing the GE interface types is shown in the following example:

```
7840-Side-A:FID128:admin> portcfgge --show
```

Port	Speed	Flags	Channel	Lag Name
ge0	40G	----	N/A	-
ge1	40G	----	N/A	-
ge2	10G	----	N/A	-
ge3	10G	----	N/A	-
ge4	10G	----	N/A	-
ge5	10G	----	N/A	-
ge6	10G	--L-	N/A	-
ge7	10G	--L-	N/A	-
ge8	10G	--L-	N/A	-
ge9	10G	--L-	N/A	-
ge10	10G	----	N/A	-
ge11	10G	----	N/A	-
ge12	10G	----	N/A	-
ge13	10G	----	N/A	-
ge14	10G	----	N/A	-
ge15	10G	----	N/A	-
ge16	10G	----	N/A	-
ge17	10G	----	N/A	-

Flags: A:Auto-Negotiation Enabled C:Copper Media Type
L:LAN Port G: LAG Member

Switchshow is useful for showing the GE interface settings as well.

On the IBM SAN42B-R and IBM SX6 Extension Blade, there are two 40GE interfaces (GE0 and GE1), which are 40 Gbps only. The IBM SAN42B-R requires Upgrade1 + Upgrade2 license to enable the 40GE interfaces. The IBM SX6 Extension Blade does not require an optional license for the 40GE interfaces; the 40GE interfaces are included and enabled. The 40GE optics are an optional purchase.

GE Interface Speeds

Depending on the platform model, various interface speeds and types are available:

- ▶ GE (RJ-45)
- ▶ GE (SFP)
- ▶ 10GE
- ▶ 40GE

The IBM SAN42B-R requires Upgrade1 + Upgrade2 to enable the 40GE interfaces.

The SX6 Extension Blade includes Upgrade1 and Upgrade2, and the 40GE interfaces are enabled.

All GE/10GE interfaces are enabled on the base IBM SAN42B-R and SX6 Blade. The GE/10GE interfaces are in Ethernet switching ASIC port-groups, and within a group, the speed must be consistent; otherwise, blocking occurs.

The following ports block each other when their speed is not set the same. The best practice is first to use ports 2, 3, 4, 5, 10, 11, 12, and 13; then, move double up within a port group:

- ▶ 2 and 6
- ▶ 3 and 7
- ▶ 4 and 8
- ▶ 5 and 9
- ▶ 10 and 14
- ▶ 11 and 15
- ▶ 12 and 16
- ▶ 13 and 17

For more information, see [*Brocade Fabric OS Extension User Guide, 9.0x*](#).

The IBM SAN18B-6 has two RJ-45 interfaces, which are 1 Gbps only. When they are enabled, SFP bays GE2 and GE3 are disabled, and vice-versa. SFP bays GE2 and GE3 are capable of GE/10GE.

Although no reboot of the IBM SAN18B-6 is required to change the active interfaces, no configuration (ipif or iproute) can be attached to the currently active ports.

GE SFP / 10GE SFP+

GE interfaces do not support auto-negotiate 1 - 10 Gbps port speeds. The speed of a GE interface is manually configured. Use `switchshow` to check the current configuration.

SFP (GE) is not capable of 10 Gbps, and SFP+ (10GE) is not capable of 1 Gbps. You must use the correct optic for the wanted port speed.

LLDP

A best practice is to use LLDP as a validator of Ethernet connectivity between the Extension application Ethernet interfaces and the directly connected DC switch. LLDP is beneficial to both Extension administrators and network administrators because network administrators see what is connected to their switches; storage administrators see to what they connected.

At both ends of an Ethernet link, LLDP logs connectivity information. For troubleshooting, adds, moves, and changes, LLDP is a handy and convenient tool for referencing Ethernet interfaces that must be investigated.

LLDP CLI output is shown in the following example:

```
lldp --show -nbr
Local7810:admin> lldp --show -nbr
Local-Int  DeadInt  Remain  Remote-Int  ChassisID          Tx          Rx          System
ge2        120      106     Ethernet1/16  0062.ecbb.f14f  88928  88922  torsw9
ge3        120      102     Ethernet1/16  0062.ecbb.eea5  88927  100    torsw8
ge6        120      106     Ethernet1/17  0062.ecbb.f150  86675  812    torsw9
ge7        120      102     Ethernet1/17  0062.ecbb.eea5  88901  109    torsw8
```

LLDP is enabled by default and operates on all application data GE interfaces. It does not operate on the management interface. LLDP is supported and enabled by default on most data center LAN switches.

There are specific TLVs (type, length, and value) that must be enabled or disabled because they are used or unused on IBM Extension platforms, as listed in Table 7-2.

Table 7-2 LLDP TLVs

Enable (used)	Disable (unused)
chassis ID	dcbx
mgmt-addr	dot1
port ID	dot3
port-desc	fcoe-app
sys-capabilities	fcoe-lls
sys-desc	
sys-name	

LLDP before (defaults) is shown in the following example:

```
7840-Side-A:FID128:admin> lldp --show
LLDP Global Information
  system-name: 7840-Side-A
  system-description: Brocade 7840, Fabric OS Version 8.2.2
  State: Enabled
  Mode: Receive/Transmit
  Advertise Transmitted: 30 seconds
  Hold time for advertise: 120 seconds
  Tx Delay Timer: 1 seconds
  Transmit TLVs: Chassis ID      Port ID
                  TTL           Port Description
                  System Name   System Description
                  System Capabilities Management Address
                  IEEE 8021      IEEE 8023
                  IEEE DCBx      DCBx FCoE App
                  DCBx FCoE Logical Link
  DCBx FCoE Priority Values: 3
```

LLDP after is shown in the following example:

```
7840-Side-A:FID128:admin> lldp --disable -tlv dcbx
7840-Side-A:FID128:admin> lldp --disable -tlv dot1
7840-Side-A:FID128:admin> lldp --disable -tlv dot3
7840-Side-A:FID128:admin> lldp --disable -tlv fcoe-app
7840-Side-A:FID128:admin> lldp --disable -tlv fcoe-lls
7840-Side-A:FID128:admin> lldp --show
LLDP Global Information
  system-name: 7840-Side-A
  system-description: Brocade 7840, Fabric OS Version 8.2.2
  State: Enabled
  Mode: Receive/Transmit
  Advertise Transmitted: 30 seconds
  Hold time for advertise: 120 seconds
  Tx Delay Timer: 1 seconds
  Transmit TLVs: Chassis ID Port ID
                  TTL Port Description
                  System Name System Description
                  System Capabilities Management Address
DCBx FCoE Priority Values: 3
```

Tunnels

Tunnels are logical entities, each represented by a VE_Port. A tunnel is made up of one or more circuits. Circuits are made up of multiple TCP sessions. Within the scope of a tunnel is encryption (IPsec), compression, protocol optimization (FCIP-FastWrite, FCIP-OSTP, and FICON-Accelerator), and enabling IP Extension. These settings apply to the entire tunnel, not per circuit.

VE_Ports and tunnels are synonymous. The following number of tunnels are available per platform:

- ▶ IBM SAN18B-6 Extension Switch2 or 4 (base or full)
- ▶ IBM SAN42B-R Extension Switch10 or 20 (two modes of operation)
- ▶ SX6 Extension Blade10 or 20 (two modes of operation)

The IBM SAN256B-7 and SAN512B-7 support up to four SX6 Extension Blades per chassis.

A tunnel configuration mismatch is indicated when the OpStatus is not Up. In this case, one side of the tunnel is Degrade, and the other side is UpWarn.

Some mismatched settings prevent the tunnel from coming up; for example, IPsec policy or compression mismatches. Some mismatched settings give warnings; for example, IP Extension enabled or disabled and inconsistent local or remote minimum and maximum communication rates, as shown in the following example:

Tunnel	Circuit	OpStatus	Flags	Uptime	TxMBps	RxMBps	ConnCnt	CommRt	Met/G
24	-	Degrade	--i---DI	1m13s	0.00	0.00	2	-	-
24	0 ge2	Degrade	----a--i4	1m13s	0.00	0.00	2	100/150	0/-
24	1 ge4	Degrade	----a--i4	1m13s	0.00	0.00	2	100/150	0/-

Flags (tunnel): l=Legacy QoS Mode
 i=IPSec f=Fastwrite T=TapePipelining F=FICON r=ReservedBW
 a=FastDeflate d=Deflate D=AggrDeflate P=Protocol
 I=IP-Ext
 (circuit): h=HA-Configured v=VLAN-Tagged p=PMTU i=IPSec 4=IPv4 6=IPv6

```

ARL a=Auto r=Reset s=StepDown t=TimedStepDown S=SLA
7840-Side-B:FID128:admin> portshow fcip tunnel -c --summary
Tunnel Circuit OpStatus Flags      Uptime TxMBps RxMBps ConnCnt CommRt Met/G
-----
24      -      UpWarn  --i---D-    7s    0.00   0.00   2      -      -
24      0 ge3    UpWarn  ----a--i4   7s    0.00   0.00   2    100/150  0/-
24      1 ge5    UpWarn  ----a--i4   7s    0.00   0.00   2    100/150  0/-
-----
Flags (tunnel): l=Legacy QoS Mode
                 i=IPSec f=Fastwrite T=TapePipelining F=FICON r=ReservedBW
                 a=FastDeflate d=Deflate D=AggrDeflate P=Protocol
                 I=IP-Ext
(circuit): h=HA-Configured v=VLAN-Tagged p=PMTU i=IPSec 4=IPv4 6=IPv6
ARL a=Auto r=Reset s=StepDown t=TimedStepDown S=SLA

```

The portshow fcip tunnel --detail output from both Extension switches is shown in the following example. In this case, IP Extension was enabled on one side and not the other. Configuration must match on both ends:

```

7840-Side-A:FID128:admin> portshow fcip tunnel --detail
Tunnel: VE-Port:24 (idx:0, DP0) Broadcom Extension
=====
Oper State      : Degraded
TID             : 24
Flags           : 0x00000000
IP-Extension    : Enabled
Compression     : Aggressive Deflate
FC-Compression  : Aggressive Deflate (Inherited)
IP-Compression  : Aggressive Deflate (Inherited)
QoS Distribution : Protocol (FC:50% / IP:50%)
FC QoS BW Ratio : 50% / 30% / 20%
IP QoS BW Ratio : 50% / 30% / 20%
Fastwrite       : Disabled
Tape Pipelining : Disabled
IPSec           : Enabled
IPSec-Policy    : MyTunnel
Legacy QoS Mode : Disabled
Load-Level (Cfg/Peer): Failover (Failover / Failover)
Local WWN       : 10:00:c4:f5:7c:3b:24:86
Peer WWN        : 00:00:00:00:00:00:00:00
RemWWN (config) : 00:00:00:00:00:00:00:00
cfgmask         : 0x0000001f 0x4001034e
Uncomp/Comp Bytes : 11048 / 11048 / 1.00 : 1
Uncomp/Comp Byte(30s): 0 / 0 / 1.00 : 1
Configuration Warnings:
  Tunnel Type (FCIP|HYBIRD)
Flow Status      : 0
ConCount/Duration : 2 / 2m31s
Uptime           : 1m30s
Stats Duration    : 1m30s
Receiver Stats    : 41214 bytes / 212 pkts / 55.00 Bps Avg
Sender Stats      : 31216 bytes / 160 pkts / 55.00 Bps Avg
TCP Bytes In/Out  : 532370 / 638344
ReTx/000/SloSt/DupAck: 8 / 6 / 6 / 0
RTT (min/avg/max) : 0 / 11 / 21 ms
Wan Util         : 0.7%

```

```

TxQ Util          : 0.0%
7840-Side-B:FD128:admin> portshow fcipunnel --detail
Tunnel: VE-Port:24 (idx:0, DP0) Broadcom Extension
=====
Oper State        : Online Warning
TID               : 24
Flags             : 0x00000000
IP-Extension      : Disabled
Compression       : Aggressive Deflate
QoS BW Ratio      : 50% / 30% / 20%
Fastwrite         : Disabled
Tape Pipelining   : Disabled
IPSec             : Enabled
IPSec-Policy      : MyTunnel
Legacy QoS Mode   : Disabled
Load-Level (Cfg/Peer): Failover (Failover / Failover)
Local WWN         : 10:00:c4:f5:7c:3b:25:06
Peer WWN          : 10:00:c4:f5:7c:3b:24:86
RemWWN (config)   : 00:00:00:00:00:00:00:00
cfgmask           : 0x0000001f 0x4000030e
Uncomp/Comp Bytes : 22514 / 22514 / 1.00 : 1
Uncomp/Comp Byte(30s): 0 / 0 / 1.00 : 1
Configuration Warnings:
  Tunnel Type (FCIP|HYBIRD)
Flow Status       : 0
ConCount/Duration : 2 / 2m35s
Uptime           : 1m51s
Stats Duration    : 1m51s
Receiver Stats    : 33176 bytes / 167 pkts / 65.00 Bps Avg
Sender Stats      : 42970 bytes / 219 pkts / 58.00 Bps Avg
TCP Bytes In/Out  : 639720 / 800940
ReTx/000/SloSt/DupAck: 9 / 4 / 6 / 0
RTT (min/avg/max) : 20 / 20 / 21 ms
Wan Util         : 0.5%
TxQ Util         : 0.0%

```

VE_Ports

VE_Ports are Virtual Expansion Ports. They are a type of E_Port (Expansion Port), which is commonly used in FC fabrics. VE_Ports connect to VE_Ports and are used to connect a domain to a domain (a switch to a switch) the same as E_Ports; the difference is VE_Ports are WAN side facing.

VE_Ports are tied to a DP. All Extension platforms have a DP0; also, the IBM SAN42B-R and IBM SX6 Extension Blade have a DP1.

A best practice is to start at the first VE_Port on DP0. When a second tunnel is required on the SAN42B-R and SX6 Blade, use the first VE_Port on DP1. On the SAN18B-6, move to the next VE_Port on DP0:

- IBM SAN42B-R:
 - VE_Port max bandwidth 20 Gbps (applicable license required)
 - DP0: VE 24 to 28
 - DP1: VE 34 to 38

- ▶ IBM SX6 Extension Blade:
 - VE_Port max bandwidth 20 Gbps (applicable license required)
 - DP0: VE 16 to 20
 - DP1: VE 26 to 30
- ▶ IBM SAN18B-6:
 - VE_Port max bandwidth 2.5 Gbps (applicable license required)
 - DP0: VE 12 to 15

It is imperative to understand TCL rules only live on the DP onto which the target VE_Port exists. VE_Ports exist only in one DP (DP0 or DP1). For more information, see “Traffic Control List” on page 186.

IP Extension gateways are created on a specific DP (DP0 or DP1). Ingress traffic that matches a TCL rule is directed to a specific target (VE_Port). If the VE_Port target that is designated in a TCL rule is not congruent with the IP Extension gateway’s DP, the rule is not evaluated. That is, the IP Extension gateway and target VE_Port used must be on the same DP for the rule to be evaluated. If the rule is expected to be assessed but is not, check whether traffic is hitting an IP Extension gateway that is configured on the other DP.

A VE_Port (tunnel) cannot be shared across different VF LS; the VE_Port must live within a single LS. GE interfaces must stay in the Default Switch where they can be used by circuits that are originating from any LS. For more information, see “Ethernet interface sharing” on page 191.

The best practice is to persistently disable unused ports.

The maximum cumulative circuit bandwidth of a VE_Port cannot exceed (license permitting) the following limits:

- ▶ 2.5 Gbps on the IBM SAN18B-6
- ▶ 20 Gbps on the IBM SAN42B-R
- ▶ 20 Gbps on the IBM SX6 Extension Blade

Interoperability

From the perspective of tunnel connectivity and features (VE_Port to VE_Port), the IBM SAN18B-6, SAN42B-R, and SX6 Extension Blade interoperate together.

The SAN42B-R and SX6 Blade have equal bandwidth capabilities when licensed correctly.

When a SAN42B-R or SX6 Blade is connected to a SAN18B-6, the maximum bandwidth to a SAN18B-6 base unit is 1 Gbps and 2.5 Gbps to a full unit.

The IBM SAN42B-R and SX6 Blade support performing eHCL; the IBM SAN18B-6 does not. However, the IBM SAN18B-6 supports the required eHCL connectivity that is needed by the IBM SAN42B-R and SX6 Blade.

Virtual Fabrics

Virtual Fabrics (VF) are composed of one or more logical switches within the physical chassis. Connected logical switches form Logical Fabrics within the overall physical fabric.

VF plays a primary role in achieving deterministic paths for Extension protocol optimization. Also, the implementation of FICON and FC Extension often requires unique configurations and management in mainframe environments, which is achievable only within the same chassis when different logical switches are used.

Protocol optimization requires that an Exchange's sequences and frames pass through the same VE_Ports for outbound *and* return. This determinant path means that only a single VE_Port can exist within a VF LS. Putting a single VE_Port in a logical switch defines a single path between the two logical switches. A logical switch sufficiently scales FC port connectivity, is operationally simplistic, and creates a stable environment. The IBM SAN42B-R and IBM SX6 Extension Blade support VF. The IBM SAN18B-6 does not support VF.

IP Extension

IP Extension is enabled within the tunnel scope (`fciptunnel`) by using `--ipext enable`. The IBM SAN18B-6 Extension platform does not have different app-modes for FCIP Only and Hybrid; essentially, it is always in Hybrid mode.

The IBM SAN42B-R and IBM SX6 Extension Blade must be put into the Hybrid app-mode before IP Extension can be enabled on a tunnel. Changing the app-mode requires a reboot of the platform or slot. Use the `extncfg` command to check the status or perform this app-mode task. Only change the app-mode to Hybrid if, at some point, IP Extension will be used.

Compression

Compression is recommended with Remote Data Replication (RDR) applications because it improves the efficient use of WAN bandwidth. Fast-Deflate compression is ultra-fast and ultra-low latency. It can be used with RDR/S (sync) to improve response time by reducing data serialization time across the wire, which essentially moves more data faster.

Commonly, tape traffic is compressed, and compressing data is not helpful. Therefore, if the traffic is compressed, leaving compression disabled might be prudent.

Note: IBM makes no guarantees, warranties, or claims about the compression ratio that is achieved with customer-specific data.

The following modes of compression are available (not including disabled):

- **Fast-Deflate**

Fast-Deflate typically gets about a 2:1 compression ratio. Fast-Deflate is a hardware-implemented (FPGA = Field Programmable Gate Array) compression algorithm that is suitable for synchronous applications. It adds a mere 10 μ s of propagation delay and accommodates large bandwidth rates.

In hybrid mode on the IBM SAN42B-R and IBM SX6 Extension Blade at 2:1 compression, fast-deflate can accommodate 20 Gbps of FC traffic per DP. In FCIP mode, the IBM SAN42B-R and IBM SX6 Extension Blade can accommodate 40 Gbps of FC traffic per DP.

- **Deflate**

Deflate typically gets about a 3:1 compression ratio. Deflate accommodates up to 16 Gbps ingress from the FC side per DP. Deflate works efficiently with circuits up to 5 Gbps per DP. Deflate is a software processed algorithm with a hardware assist. Software-based algorithms take longer to process and might not be suitable for synchronous applications.

- **Aggressive-Deflate**

Aggressive-Deflate takes further the trade-off between the compression ratio and the compression rate. Aggressive-Deflate typically gets about a 4:1 compression ratio. The maximum rate per DP is 10 Gbps ingress from the FC side. Aggressive-Deflate works efficiently with circuits up to 2.5 Gbps per DP. As with Deflate, Aggressive-Deflate is a software-based algorithm and might not be suitable for synchronous applications.

IPsec and encryption

A best practice is to use encryption to protect data end-to-end. It is prudent to enable IPsec when IBM Extension is implemented. All data that is leaving the secure confines of the data center into public infrastructure guarantees no security, and service providers do not guarantee unencrypted data while in flight.

Links must be authenticated and data encrypted to prevent attacks and eavesdropping. IBM IPsec is easy and practical to deploy. Would your company operate WiFi with no encryption? Of course not, and Extension is no different.

Encryption is a hardware implementation that operates at line-rate. IPsec is included in all base Extension platforms, and no licenses are available to enable IPsec. Encryption adds a minor propagation delay (~5 μ s). Preshared Key (PSK) configuration is simple. CA-signed or self-signed certificates also are supported.

IPsec is Suite B compliant and implements the latest encryption technologies, such as AES 256, SHA-512 HMAC, IKEv2, and Diffie-Hellman. Rekeying occurs in the background approximately every 2 billion frames or every 4 hours, and the process is nondisruptive.

Firewalls are not considered best practice for the following reasons:

- ▶ IPsec operates up to 20 Gbps per DP, which is the WAN side line-rate (40 Gbps if implementing both DPs). Most firewalls cannot meet this throughput requirement.
- ▶ Storage traffic requires minimal propagation delay.
- ▶ IPsec encrypts data closer to the data source and destination, which is considered a best practice.
- ▶ Firewalls that proxy TCP sessions, like some WAN optimization devices, result in remote TCP segments not being identical, which is not supported. Such WAN optimization devices are also not supported for this same reason.

From end to end, the IP network must allow specific protocols to pass. When IPsec is not used, Brocade Extension uses TCP destination ports 3225 and 3226. The Extension TCP implementation selects a random source port 49152 - 65535 as the ephemeral (or initiating) port to open to destination ports 3225 and 3226. TCP URG flags must not be dropped or modified.

When IPsec is used, IKEv2 uses UDP port 500. Management connectivity protocols (ssh, snmp, https, and so on) are not covered here.

Before creating the tunnel, first, create the IPsec policy, as shown in the following example:

```
portcfg ipsec-policy <name> <option> [<args>]
portcfg ipsec-policy MYPOLICY create --preshared-key "mykeymykeymykeymykey"
```

Protocol acceleration

Protocol acceleration is used to expedite data without the time that is needed to traverse the WAN to request a transfer ready. Protocol optimization effectively mitigates the latency effects on SCSI/FCP disk writes and tape read/writes. Also, protocol optimization exists for FICON tape and IBM z/OS® GM (XRC).

FastWrite, Open Systems Tape Pipelining (OSTP), and FICON Acceleration are features of IBM Extension platforms. IBM protocol acceleration requires no extra hardware or hardware licenses. The same Data Processor (DP) that forms tunnels and trunks also performs protocol acceleration for a higher level of integration, efficiency, resource utilization, and reduced assets.

Flows through Extension are monitored for protocol optimization viability. Not including FICON, IBM storage products do not benefit from or engage FCIP-FastWrite. FastWrite is not enabled if Extension is used for Global Mirror or Metro Mirror alone.

If Extension is deployed with another application that is known to benefit from and engage FastWrite, FastWrite should be enabled. When FastWrite is enabled, it works with the other application while ignoring Global Mirror and Metro Mirror. IBM FICON tape, Teradata, and z/OS GM (XRC) can be optimized with IBM Advanced FICON Accelerator, which is included in the mainframe version of the IBM SX6 Extension Blade and an optional license on the IBM SAN42B-R.

Brocade Extension Trunking (BET) is essential for protocol optimization techniques, such as FastWrite, Open Systems Tape Pipelining (OSTP), and FICON Acceleration. BET uses a single VE_Port and is essentially an ISL.

Taking a different VE_Port outbound versus inbound breaks protocol optimization by creating an inconsistent state machine. Individual tunnels, each with their own VE_Port, leads to indeterminate VE_Ports for sequences. It is not possible to perform protocol acceleration with multiple parallel VE_Ports between the same domains. Regardless of the routing mode (EBR, DBR, and PBR), multiple VE_Ports prevent a bidirectional deterministic path. For example, with PBR, the outgoing sequence might consistently take the same VE_Port; however, the return sequence might always take the same incorrect VE_Port.

When protocol acceleration is used, it is necessary to confine traffic to a single bidirectional path, which can be accomplished by using one of the following methods:

- ▶ Use only a single pair of VE_Ports between the domains.
- ▶ If more than one VE_Port is required, use VF-LS (Virtual Fabrics Logical Switch) to separate each VE_Port pair.

Note: Protocol optimization techniques prevent the use of load balancing and failover between VE_Ports.

Protocol acceleration uses state machines to perform optimization and requires traffic to be confined to a single VE_Ports pair. State machines must know every sequence that passed during the exchange.

State machines facilitate the correct handling of an exchange until it is complete, after which the state machine is discarded. State machines are created and destroyed with every exchange.

IBM Extension can accommodate tens of thousands of simultaneous state machines and an equivalent number of flows. Ingress flows are verified if they can be optimized; if not, the flow does not engage protocol acceleration.

If a flow cannot be optimized, it merely passes without optimization. A state machine cannot exist across multiple VE_Ports, and no communication occurs between different state machines.

If an exchange starts on VE24 and returns on VE25, the exchange passes through different state machines, as shown in Figure 7-23. One is knowing nothing about the other. This situation causes protocol acceleration to break and an error condition.

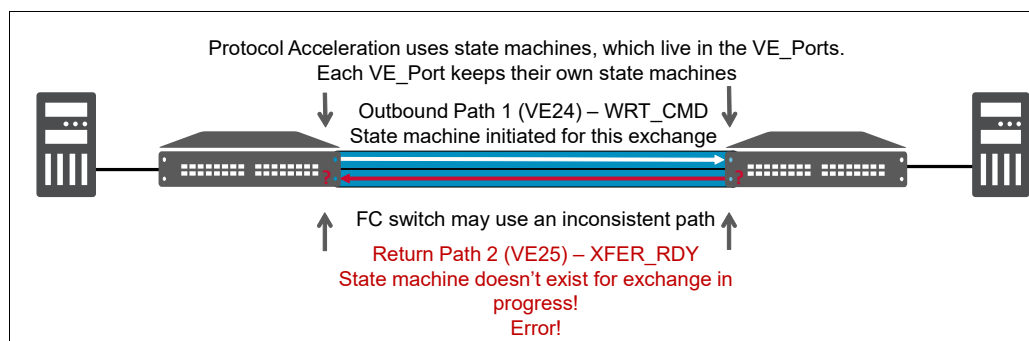


Figure 7-23 No Deterministic Path for Protocol Optimization

One advantage of the use of BET is to have a single tunnel, regardless of how many circuits make up the tunnel. The state machines exist at the TCP Supervisor endpoints and remain consistent no matter which circuit is used (see Figure 7-24). BET permits protocol optimization to function across load-balanced links with nondisruptive failover or failback.

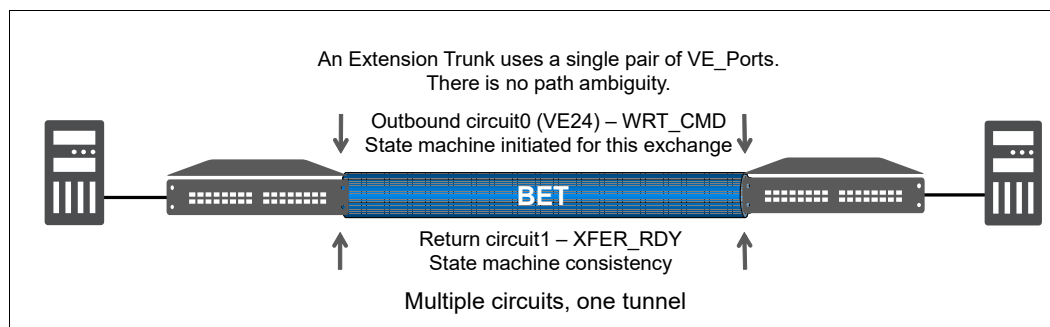


Figure 7-24 Deterministic Path for Protocol Optimization

In situations where protocol acceleration is not enabled, data flows can be routed across multiple VE_Ports between domains. The routing mode is based on the APTpolicy. One of three methods is used internally to route FC flows between equal-cost VE_Ports from the FC switching ASIC:

- ▶ Exchange Based Routing (EBR), which is the default: Originator Exchange ID/Source ID/Destination ID (OXID/SID/DID)
- ▶ Device-Based Routing (DBR): SID/DID
- ▶ Port-Based Routing (PBR): Source Port

FastWrite (FCIP-FW)

FCIP-FastWrite is a protocol optimization technique that intercepts an FCP exchange and spoofs a local response back to the initiator to expedite data out to the target. The initiator's request does not need to first traverse the WAN, be responded to by the target, and travel the WAN again before data starts being sent.

IBM GM and MM do not benefit from or engage FastWrite on IBM b-type Gen 7 Extension platforms. Although optimization is built into these replication applications, FastWrite might still be needed and can be used with cohort replication applications without detriment to GM or MM.

Open Systems Tape Pipelining

FCIP-Open Systems Tape Pipelining (OSTP) is a protocol optimization technique that intercepts FCP exchanges and spoofs a local response. OSTP supports read and write tape optimization. Read optimization involves spoofed responses at the target side to expedite sequential block reads. Write optimization involves spoofed responses at the initiator side to expedite sequential block data out. Sequential block tape requests do not need to traverse the WAN, target response, and traverse the WAN again.

FICON Accelerator

Advanced FICON Accelerator is an FCIP protocol optimization technique that works with FICON tape, Teradata, and z/OS GM (XRC) over distances 2 ms RTT or greater.

Two versions of the IBM SX6 Extension Blade are available: distributed systems and a mainframe version. The IBM SX6 Extension Blade mainframe version includes the FICON Accelerator License (FAL). FAL is an optional license on the SAN42B-R. The SAN18B-6 does not support FICON features and functions.

Circuits

Circuits make up a tunnel. A tunnel consists of one or more circuits. A tunnel that consists of multiple circuits does not rely on a single port, optic, cable, switch, router, WAN connection, or service provider. In essence, a single point of failure in the IP network can be avoided.

The tunnel (VE_Port) manages circuits. Circuits contain several WAN Optimized TCP sessions. Different TCP sessions are created for each QoS priority (H/M/L) within each protocol plus control traffic. Two Extension protocols are available: IP Extension and FCIP.

Within the scope of circuits is ARL (minimum and maximum committed rate), local and remote IP, local and remote HA IP (not on the IBM SAN18B-6), metrics and group, KATOV, and QoS (DSCP and L2CoS).

IP interfaces

A WAN side GE interface can have multiple IP interfaces (ipif) assigned to it; each is a circuit endpoint. A circuit is assigned to one Ethernet interface; for example, ge2.dp0. You cannot split a circuit across multiple Ethernet interfaces. An IP interface (ipif) is configured with an IP address and GE interface. A circuit is created with a local IP address and bound to a GE interface, which is the ipif with which the IP address was configured.

IP routes

IP routes (iproute) are used to forward traffic to the next-hop gateway. IP routes apply to the LAN side and WAN side. For the LAN side, an IP route is used to send only IP Extension traffic out to an end device, and traffic that comes in from the WAN side. IP routes are not used to route ingress LAN traffic into an egress tunnel; only TCL is used for this purpose.

IP routes are specific to a GE interface and specified when created. Every GE interface features its own routing table. Therefore, two circuits that use the same source subnet, go to the same destination subnet, and are assigned to different GE interfaces require the following ipif and iproute creation:

- ▶ portcfg ipif ge2.dp0 create 10.10.10.10/24
- ▶ portcfg ipif ge3.dp0 create 10.10.10.11/24
- ▶ portcfg iproute ge2.dp0 create 10.20.20.0/24 10.10.10.1
- ▶ portcfg iproute ge3.dp0 create 10.20.20.0/24 10.10.10.1

IP routes are used on the WAN side to specify the gateway when the destination subnet is not the same as the source subnet (Layer3 network). For example, the local ipif subnet is 10.10.10.0/24, and the remote ipif subnet is 10.20.20.0/24. Because these subnets are different, an IP route is required. IP routes are not required if the local and remote IP addresses are on the same subnet (Layer2 network).

Continuing with this example, assume the local router gateway was 10.10.10.1. The IP route gateway 10.10.10.1 is used to get to subnet 10.20.20.0/24. The gateway is an IP address on the same subnet as the IP address on the ipif.

Diverse pathways

Each circuit is intended to take a different pathway to gain resiliency and redundancy across the IP network. Therefore, circuits connect by using various physical Ethernet links to the different physical data center LAN switches or WAN routers.

When more than one WAN connection is available, the circuits must be statically assigned to those WAN connections. BET failover or failback is automatic, lossless, and ensures in-order delivery; therefore, it is best to have the circuit go offline, which stops the keepalives and forces traffic to failover at the FC/FICON level. In the case of two WAN links, four circuits often are used, as shown in the following example:

- ▶ Cir0: DC LAN Switch A - WAN 1
- ▶ Cir1: DC LAN Switch A - WAN 2
- ▶ Cir2: DC LAN Switch B - WAN 1
- ▶ Cir3: DC LAN Switch B - WAN 2

The effective RTT for all circuits in a trunk is the longest RTT that is experienced by a circuit. By effectively simulating the same RTT for all circuits, data can be driven at line-rate without significant data arriving out of order. Wait times climb with excessive out of order data.

Bandwidth scheduling is weighted per circuit based on the comm-rate. A typical example is a tunnel that is deployed across a ring architecture, in which one span is a relatively short distance and the other span is much longer. It remains practical to trunk two circuits across each span of the ring. Asynchronous RDR is not affected when both paths experience longer latency, yet the data rate delivery rate is optimized.

Brocade Extension Trunking

BET is available on all Extension platforms. The base units of the IBM SAN42B-R and SX6 Extension Blade include BET. The IBM SAN18B-6 requires the full unit license for this feature. BET is an advanced extension feature that provides the following benefits:

- ▶ Single logical tunnel that is composed of one or more individual circuits
- ▶ Efficient use of VE_Ports
- ▶ Aggregation of circuit bandwidth
- ▶ Failover and failback
- ▶ Failover groups and metrics
- ▶ Use of disparate WAN paths with differing characteristics
- ▶ Lossless Link Loss (LLL)
- ▶ In-order delivery
- ▶ Non-disruptive link loss
- ▶ A deterministic path for protocol acceleration

BET provides a single logical tunnel that is composed of multiple circuits. A tunnel with multiple circuits is referred to as a *trunk*. Brocade Extension uses a VE_Port to form an Extension ISL that consists of associated circuits.

If each circuit were a tunnel, a VE_Port is required for each; however, each circuit makes up a single tunnel with BET. A tunnel is an ISL and treated as such in fabric design. Circuits are individual connections within the tunnel, each with a unique source and destination IP address and other settings.

Figure 7-25 shows an example of a tunnel trunking four circuits.

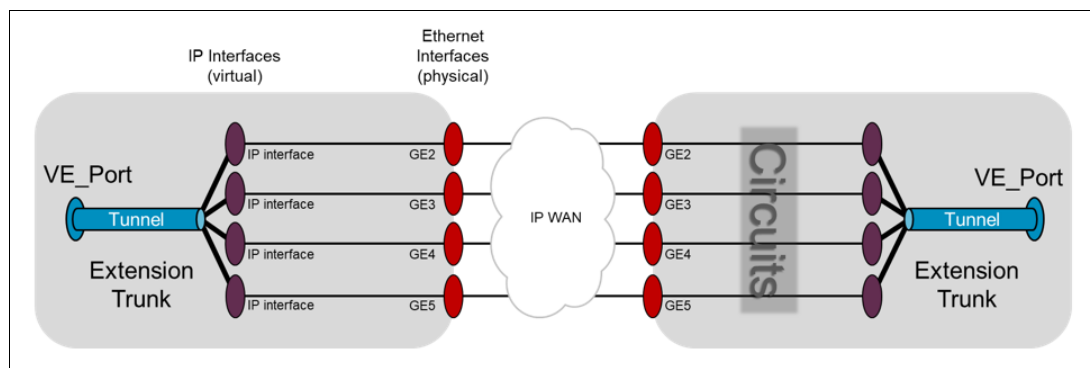


Figure 7-25 Brocade extension trunking

Each circuit endpoint is assigned a unique IP address by way of IP interfaces (`portcfg ipif`). An IP interface is bound to an Ethernet interface when the `ipif` is configured. In this case, each IP interface is assigned to a different Ethernet interface; however, this assignment is not required.

Depending on the environment's needs, Ethernet interface assignments are flexible and can be made as needed. For example, multiple IP interfaces can be assigned to a single Ethernet interface. No subnet restrictions exist. Circuits connect from the local IP interface to the remote IP interface by way of the bound Ethernet interfaces.

VE_Ports are the demarcation point between the FC world and the TCP/IP world. Brocade Extension uses FC switching Application Specific Integrated Circuits (ASICs), which know only FC protocol. VE_Ports are not an included part of these ASICs and are logical representations of FC ports. Internally, multiple backend FC ports feed a VE_Port to gain high data rates into the Extension DP.

Because one VE_Port represents the endpoint of multiple trunked circuits, the following benefits are realized:

- ▶ Fewer VE_Ports are needed, and the remaining VE_Ports are available for other destinations. Typically, only one VE_Port is required per remote location.
- ▶ Bandwidth aggregation is achieved by adding circuits to a trunk. Circuits do not have to be configured identically and do not have to take the same or similar paths. The Round Trip Time (RTT) does not have to be the same.
- ▶ Link failover is fully automated with BET. The use of circuit metrics and failover groups enables active and passive circuits to coexist within the same trunk.

When a connection is lost, data that is in-flight also is lost. Any frames that are transmitted at the instance of the outage are lost because of partial transmission.

Consider multiple circuits in which Out-Of-Order-Segments (OOOS) occur when some segments arrive, one or more are lost in-flight, and then segments continue to arrive. OOOS are problematic for some devices, particularly mainframes, which results in an Interface Control Check (IFCC).

For example, frames 1 - 4 are sent. Frames 1 and 2 are received. Frame 3 is lost in transit, and frame 4 is received. The receiving side detects an OOOS because it was expecting frame 3, but received frame 4. BET solves this problem with Lossless Link Loss (LLL).

For BET, the DP that owns the VE_Port controls all circuits. Distributed processing, load sharing, or LLL does not occur across DPs. Failover of FC traffic from one DP to another (for example, a VE on DP0 gets rerouted to a VE on DP1) is done at the FC level if the configuration permits it. The only shared components might be various circuits that are bound to the Ethernet interfaces.

When a link is lost, the lost frames are retransmitted by LLL (see Figure 7-26). Usually, when a TCP segment is lost because of a dirty link, bit error, congestion, or any other issue, TCP is aware and retransmits the segment. In the case of an offline circuit, TCP has no way to retransmit lost segments because TCP is no longer operating on that connection. Recovery must occur at a higher level in the stack.

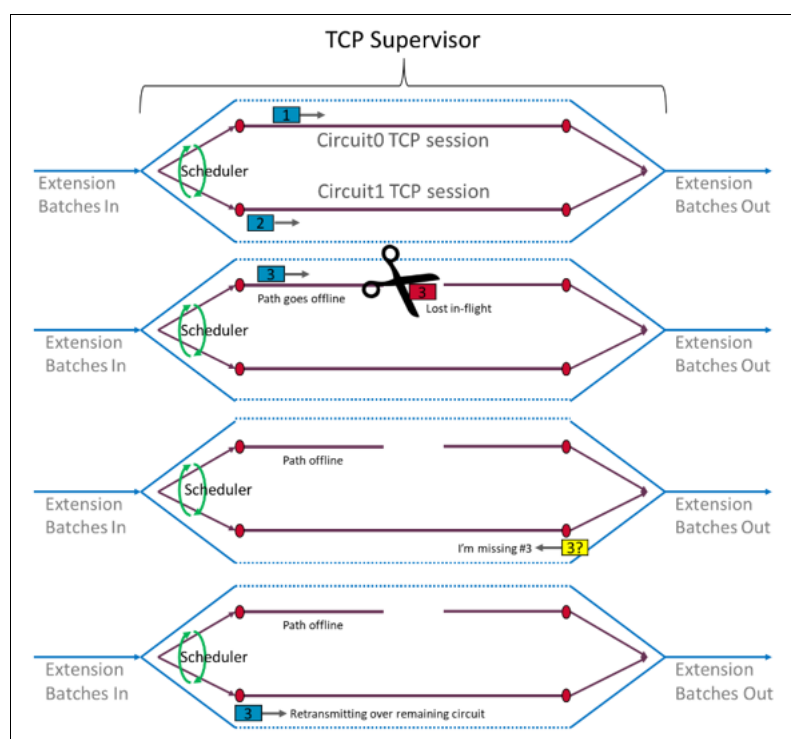


Figure 7-26 Lossless link loss process

TCP sessions are managed by a “TCP supervisor,” which feeds each circuit’s TCP sessions through a load balancer and a sophisticated egress queuing and scheduling mechanism that compensates for different bandwidths, latencies, and congestion events.

The TCP Supervisor is a Brocade proprietary technology, purpose-designed, and a special function algorithm. It operates at the presentation level (Level 6) of the Open Systems Interconnection (OSI) model and does not affect any network devices that are monitoring or providing security at the TCP level (L4).

Suppose a link goes offline, and data can continue over remaining connections. In that case, missing frames that are indicated by noncontiguous sequence numbers trigger a duplicate acknowledgment back to the TCP supervisor to retransmit the missing segment, even though that segment was initially sent by a different, now defunct, TCP session. Data that is lost in flight requires all segments to be held in memory for possible retransmission until acknowledged at the supervisor level.

Failover, failback, and spillover

Circuits can be active or passive within a trunk and configured with metrics. Within a failover group, all online circuits with the lowest metric value are active; for example, a value of 0. Circuits with a value of 1 are not active and become active only after *all* metric 0 circuits are offline within the group. In an example of two groups, one-to-one pairings of a metric 1 circuit with a metric 0 circuit is shown in Figure 7-27.

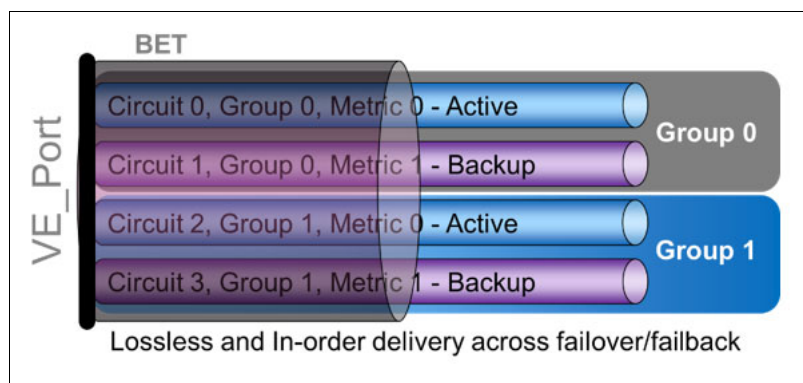


Figure 7-27 Failover Metrics and Groups

Within a group, the metric 1 circuits losslessly resume the metric 0 traffic upon the last metric 0 going offline. Metric 1 circuits should take different routes, considering network outages are the most likely cause the offline metric 0 circuits.

Metrics with failover groups permit passive circuits to become active over a less desirable path when the regular production path becomes unavailable. In a three-site triangular architecture, the primary circuit has a metric of 0 in which regular production traffic takes a single hop, shortest distance, lowest latency path.

Sending traffic through the secondary path, which has two hops and is typically longer in distance and latency, is prevented unless the primary path is interrupted. Setting a circuit to metric 1 makes it passive until all metric 0 circuits within the group have gone offline.

Nonetheless, dormant circuits are members of the trunk. Convergence over to the secondary path is lossless, and no out-of-order segments exist. No IFCC is generated on a mainframe. Extension Trunking is required to assign metrics to circuits.

ARL

ARL is an integral part of most Extension designs. When more than one circuit is feeding into the same WAN link or when the WAN is shared with other traffic, ARL is essential. ARL is used to manage data sent into the IP network based on minimum and maximum rate settings and the unconsumed bandwidth between minimum and maximum (see Figure 7-28).

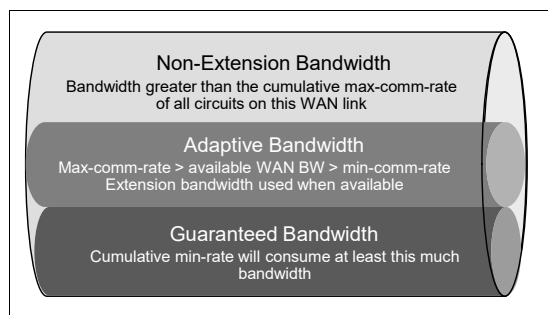


Figure 7-28 Guaranteed, adaptive, and non-extension bandwidth by using ARL

One or multiple WAN connections might exist. The cumulative circuit bandwidth within a specific WAN connection is what matters. Each WAN connection is evaluated independently.

In a shared WAN, consider bandwidth as separated into the following distinct sections:

- ▶ Extension guaranteed bandwidth (0 to min-comm-rate)
- ▶ Extension adaptive bandwidth (min-comm-rate to max-comm-rate, used when available)
- ▶ Non-Extension bandwidth (max-comm-rate to WAN bandwidth).

ARL manages the bandwidth into these sections. The Extension guaranteed section is the minimum bandwidth that is available for Extension on the WAN. The Extension guaranteed section is the amount of bandwidth that is reserved exclusively for Extension, and ARL aggressively pushes at least this amount.

Note: The minimum bandwidth is the aggregate of *all* circuit minimums on the WAN connection.

Adaptive is the area between the top of the guaranteed section and the bottom of the non-Extension section. The bottom of the non-Extension section is the aggregate of the max-comm-rates for all WAN connection circuits. Extension circuits can use the adaptive bandwidth when other non-Extension applications are not using it.

The non-Extension bandwidth section is reserved exclusively for applications that are sharing the WAN. Extension traffic does not exceed the aggregate of the max-comm-rates for all circuits on the WAN connection.

Assume a 1 Gbps WAN as an example. The ARL max-comm-rate is set to the line-rate GE interface or the maximum available WAN BW, whichever is lowest. In this case, the max-comm-rate is set to 1 Gbps.

Each circuit is configured with a floor (--min-comm-rate) and ceiling (--max-comm-rate) bandwidth value in kilobits per second (kbps). Assume adequate bandwidth demand from the storage application; the minimum circuit bandwidth is never less than the min-comm-rate and is never more than the max-comm-rate.

The circuit bandwidth adjusts automatically between the minimum and maximum based on WAN conditions at that moment. A congestion event causes the rate-limit to read only towards the minimum. An absence of congestion events causes it to rise towards the maximum.

If the current rate is not maximum, ARL attempts to adjust upward; however, the rate remains stable if repeated congestion events are detected. When constant congestion exists, the minimum comm-rate becomes the prevailing rate.

The ARL min-comm-rate is set to the max-comm-rate divided by the number of circuits that is feeding the WAN connection. In this example, $1 \text{ Gbps} \div 4 = 250 \text{ Mbps}$. The min-comm-rate is set to 250 Mbps. When all the circuits are up, each runs at 250 Mbps.

In an extreme case in which three circuits go offline, the remaining circuit runs at 1 Gbps. At 1 Gbps, all the WAN bandwidth continues to be used, and the replication application remains satisfied.

When more than one circuit with the same ARL settings feeds a WAN link, the two circuits equally use the available bandwidth. When an interface or the entire platform goes offline, ARL readjusts to use available bandwidth that is no longer being used by offline circuits. Adaptivity and readjustment of bandwidth maintain optimal usage during periods of maintenance and failures.

Note: Changing minimum and maximum comm-rates on circuits is a nondisruptive operation.

There are many ways in which ARL can be used.

Extension Hot Code Load

Extension Hot Code Load (eHCL) is unique to IBM b-type Gen 7 Extension platforms. It enables Extension platforms to non-disruptively update firmware without the tunnels going offline. Moreover, the process maintains all data that is in-flight and ensures in-order delivery. No data is lost and no data is out of order.

eHCL works by moving the circuits from their native DP to the opposite DP. Firmware upgrades occur sequentially one DP at a time. When a DP upgrade is complete, the circuits move back to their original DP. Firmware updates can be started only on one tunnel endpoint at a time. A stable fabric with no device adds, moves, or changes occurring is strongly suggested.

To enable eHCL, another set of high availability (HA) ipif, iproute, and IP addresses must be added to each tunnel on both ends. Assume a tunnel and circuits are on DP0-VE24:

- ▶ The HA ipifs are created on DP1 (the opposite DP); for example, ipif ge2.dp0 10.10.10.10/24, the complimentary HA ipif ge2.dp1 10.10.10.11/24. Typically, the same subnet is used for convenience, but is not required.
- ▶ The HA iproutes are created on DP1; for example, iproute ge2.dp0 192.168.10.0/24 10.10.10.1, the complimentary HA iproute ge2.dp1 192.168.10.0/24 10.10.10.1.
- ▶ Add to fcipcircuit --local-ha-ip 10.10.10.11, and the --remote-ha-ip 192.168.10.11.

Nothing else needs to be configured or enabled. The eHCL process occurs automatically with each firmware update if the HA status is “HA Online,” as shown in the following example CLI output in which the status of the HA configuration is checked:

```
ASH-7840-A:FID128:admin> portshow fcip tunnel -c --hcl-status
Checking FCIP Tunnel HA Status.
Current Status           : Ready
CP Version               : v8.2.1c1
DP0 Status:
  State                  : Online - Inactive
  Version                : v8.2.1c1
  Current FC HA Stage    : IDLE
  Current IP HA Stage    : IDLE
  IP SVI Swapped         : NO
  DP COMM Status         : UP
DP1 Status:
  State                  : Online - Inactive
  Version                : v8.2.1c1
  Current FC HA Stage    : IDLE
  Current IP HA Stage    : IDLE
  IP SVI Swapped         : NO
  DP COMM Status         : UP
Tunnel 24 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.
Tunnel 25 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.
Tunnel 26 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.
Tunnel 27 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.
Tunnel 28 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.
```

Tunnel 34 (FID:128) FC:HA Ready IP:HA Ready - FC and IP traffic will be disrupted.

Tunnel 35 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.

Tunnel 36 (FID:128) FC:HA Ready IP:HA Ready - FC and IP traffic will be disrupted.

Tunnel 37 (FID:128) FC:HA Ready IP:HA Ready - FC and IP traffic will be disrupted.

Tunnel 38 (FID:128) FC:HA Online IP:HA Online - Traffic will not be disrupted.

KATOV

Each circuit uses keepalives to reset the keepalive timer. The time interval between keepalives is a configurable setting. If the keepalive timer expires, the circuit is deemed to be down and is removed from the trunk.

Also, when an Ethernet interface loses light, those circuits are immediately considered down. When a circuit goes offline, the egress queue for that circuit is removed from the load-balancing algorithm. Traffic continues across the remaining circuits, albeit at the reduced bandwidth due to removing the link. Consider the following points:

1. How does the keepalive mechanism work on a circuit?

A Keepalive Timeout Value (KATOV) is configured for each circuit in the tunnel. BET uses the value to determine how frequently a keepalive is sent. The math for that determination is not entirely straightforward and is beyond the scope of this document.

The algorithm ensures that multiple keepalives are sent within the timeout period. If the receiving side does not receive any keepalives within the timeout period, the circuit is deemed down.

2. How are keepalives queued and handled through the network layers?

Keepalives are treated as normal data to WAN-Optimized TCP (WO-TCP). They are not sent on any unique path through the network stack; therefore, they intermix with the other data that is passing through TCP, which ensures keepalive transmission. Keepalives cannot be lost unless the circuit is down.

Moreover, the keepalive process determines the delay through an IP network. The mechanism considers latency, reorder, retransmits, and any other network condition.

TCP is the preferred transport for keepalives because it keeps tight control of the time that data is permitted on the WAN. If getting through the network exceeds the KATOV, the circuit is torn down, and pending data is re-queued into the remaining circuits.

It is recommended to configure the KATOV based on the storage application's timeout values. The exception is circuits that are configured with the FICON option, which have a 1-second KATOV and must not be altered. The default value for non-FICON tunnels is 6 seconds and can often be reduced depending on the network characteristics. This reduction is evaluated on a case-by-case basis. When a tunnel contains multiple circuits, the KATOV must be less than the storage application timeout; this facilitates LLL before the application times out.

3. Can too much data through a circuit bring it down?

Yes and no. The answer is no if no issues exist in the IP network, and the tunnel is not large enough to handle the driven workload. No network issues means that no congestion or packet loss is occurring, and ARL is suitably configured; the bandwidth merely cannot accommodate the application's needs. When all the available bandwidth is used, buffer credit flow control applies backpressure to the incoming flows to prevent buffer overflow. No circuits go down because of keepalive timeout.

Keepalives are not negatively affected by large amounts of queued data. The amount of data that can be queued but not sent to the network is relatively small (no more than a few milliseconds). A few milliseconds is not enough time to cause a keepalive timeout because of saturation.

However, mis-configured ARL bandwidth, network issues, congestion, and packet loss are conditions that can lead to keepalive timer expiration, and subsequently, a dropped circuit. Improper rate limiting allows more data to enter the network than bandwidth is available, which results in network errors. Severe buffer overruns, excessive packet loss, and time exhausting retransmits can result in enough lost keepalives to bring down a circuit.

QoS

IBM Extension includes a comprehensive quality of service (QoS) suite:

- ▶ Protocol distribution (enforcement)
- ▶ Priority high/medium/low (enforcement)
- ▶ DSCP (marking)
- ▶ 802.1P (marking)
- ▶ PTQ (isolation)

Note: The best practice is to *not* alter class-F (control traffic) QoS markings unless required to differentiate and expedite control traffic across the IP network. Control traffic that fails to arrive promptly causes fabric instability.

The following CLI output example shows a tunnel (VE24) with one circuit (0 on GE2) and its individual QoS components (control, high, medium, and low). The CommRt shows bandwidth distribution to each priority out of 100% (these are not rate-limits). Control is a strict priority; that is, if there is something to send, the control traffic is sent before any other:

```
LVN-7840-B_CA:FID1:admin> portshow fcipunnel -c --qos
Tunnel Circuit OpStatus Flags Uptime TxMBps RxMBps ConnCnt CommRt Met/G
-----
24 - Up c----F-D- 28d22h 0.00 0.00 1 - -
24 0 ge2 Up ----r---4 28d22h 0.00 0.00 1 0/100 0/-
24 - Up h----F-D- 28d22h 0.00 0.00 1 - -
24 0 ge2 Up ----r---4 28d22h 0.00 0.00 1 50/100 0/-
24 - Up m----F-D- 28d22h 0.00 0.00 1 - -
24 0 ge2 Up ----r---4 28d22h 0.00 0.00 1 30/100 0/-
24 - Up l----F-D- 28d22h 0.00 0.00 1 - -
24 0 ge2 Up ----r---4 28d22h 0.00 0.00 1 20/100 0/-
-----
Flags (tunnel): c=Control h=HighPri m=MedPri l=LowPri
                i=IPSec f=Fastwrite T=TapePipelining F=FICON r=ReservedBW
                a=FastDeflate d=Deflate D=AggrDeflate P=Protocol
                I=IP-Ext
(circuit): h=HA-Configured v=VLAN-Tagged p=PMTU i=IPSec 4=IPv4 6=IPv6
           ARL a=Auto r=Reset s=StepDown t=TimedStepDown S=SLA
```

Protocol distribution

Protocol distribution is the highest tier of QoS enforcement in Brocade Extension. Enforcement occurs only when contention exists for WAN side bandwidth. Two protocols are available: FCIP and IP Extension.

The question of which protocol gets the bandwidth is answered by protocol distribution (which is user-configurable). The default is 50/50; that is, each protocol receives 50% of the tunnel's cumulative circuit bandwidth during contention.

When egress contention exists, no limit is imposed by protocol distribution. For example, FCIP receives 100% of the bandwidth if no IP Extension traffic exists at that moment. Protocol distribution does not need to be configured when only one of the two protocols is implemented because no contention exists between protocols.

A frequently used case is a tape grid (IP Extension) with disk replication (FCIP). The best practice is to manage the overall bandwidth by using the same VE_Port (tunnel).

Often, disk replication is more critical, and the desire is to permit disk replication over the tape. Assign disk replication the bandwidth that it needs; the tape grid receives what disk does not use.

Setting protocol distribution to 90/10 FCIP/IP, respectively, the disk gets 90% of the bandwidth if and when it needs that much. Worst case scenario, the tape grid gets 10% of the bandwidth.

If the WAN is sized correctly for disk plus tape, rarely does disk replication require 90% of the bandwidth. Unused bandwidth is used by IP Extension for the tape grid. During periods of low disk bandwidth usage, the tape grid receives most of the bandwidth. Protocol distribution is a safety mechanism for disk replication on a shared WAN connection.

Priority levels

IBM Extension supports three levels of priority (High, Medium, and Low). The default amount of BW that the scheduler apportions during times of contention is 50/30/20 percent. QoS apportioning of BW occurs only during contention; otherwise, the BW is shared equally across all priorities. It is possible to change the default portions to any quantity needed if the amounts are 10% or greater and add up to 100%.

Seven outbound priorities are available per circuit. The following default percentages are used:

- ▶ Class-F (strict)
- ▶ FCIP (50%):
 - High (50%)
 - Medium (30%)
 - Low (20%)
- ▶ IP Extension (50%):
 - High (50%)
 - Medium (30%)
 - Low (20%)

Class-F (control) uses strict queuing; that is, if any data must be sent, it is transmitted first. Few control traffic exists and it does not interfere with data. Each QoS priority has an autonomous TCP session and no reliance on other sessions.

Each priority TCP session can be configured with its own DSCP, VLAN tagging, and 802.1P values. The ability to set each circuit uniquely permits sessions to be treated independently that is based on priority SLA between sites.

QoS exists within the SAN fabrics and across FC ISLs by using virtual channels (VCs). Different VCs are used for H/M/L and Class-F. Each VC has its own Buffer-to-Buffer Credits and flow control.

Devices are assigned to QoS VCs by enabling QoS in the fabric and putting QOSH_ or QOSL_ as the prefix to a zone name. The default is QOSM_; therefore, medium zones do not need to be specifically designated; although, the prefix is supported.

Devices use VCs throughout the fabric, including Extension. If data ingresses by way of a “QOSH_” prefixed zone, the data uses High VCs within the fabric and is automatically assigned to a High TCP session in the Extension circuits.

Traversing the fabric is not required to set traffic to Extension TCP priorities. Any device that is directly connected to an Extension platform is assigned a priority TCP session based on the zone name prefix. The default TCP session is medium.

DSCP

Differentiated Services Code Point (DSCP) is an IP based (L3) QoS marking; therefore, it is an end-to-end QoS marking protocol. DSCP has 64 values; however, the values 0 - 63 do not denote lowest to highest priority; the marking schema is different.

All odd values are for private use and can be used in any way an enterprise sees fit, which is similar to how RFC 1918 IP addresses are for private use. For example, 192.168.10.1 and DSCP 3 are private values that are used in any way an enterprise wants.

Value 46, a non-private DSCP value, is referred to as *Expedited Forwarding* and is the highest DSCP priority. Zero is the default and the lowest priority. Four groups of High/Medium/Low values are referred to as *Assured Forwarding*. Another group of numbers is used for backward compatibility with legacy Type of Service (ToS).

Using DSCP in an IP network is the network administrators' responsibility. Without their buy-in and configuration of the Per-Hop Behavior (PHB) that is associated with QoS, no QoS can occur. Ethernet switches' default behavior is to ignore QoS values by replacing any ingress value with zero; unless data that is arriving on that interface is configured as trusted. Specifically configuring ingress data as trusted prevents users from setting QoS values unannounced to their Network Administrators.

DSCP marking is configured per-circuit + per-protocol + per-priority (H/M/L).

802.1P

802.1P is a data link-based (L2) QoS marking; therefore, the scope extends from the Extension platform's GE interface to the Ethernet interface of the directly attached switch. Devices that are enforcing 802.1P provide QoS across the data link.

The 802.1P header is within the 802.1Q VLAN header; therefore, VLAN tagging is required to get 802.1P QoS marking. Brocade Fabric Operating System refers to 802.1P as L2CoS. (Layer 2 Class of Service). There are only eight 802.1P values 0 - 7. Zero is the default and lowest priority; seven is the highest priority. The IP network must be configured to acknowledge QoS markings before being useful; otherwise, the Extension markings are ignored.

L2CoS is configured per-circuit + per-protocol + per-priority (H/M/L).

Priority TCP QoS

Priority TCP QoS (PTQ) is not configurable; however, it is a crucial QoS concept to understand when extending QoS across Extension. Any FC zone with a prefix that indicates QoS uses priority virtual circuits within the fabric and automatically maps traffic to a corresponding TCP priority in the Extension circuits.

One TCP hallmark is ensuring in-order delivery. By nature, QoS indicates flows that require expediency and designates such handling by the network. TCP and QoS work against each other if QoS expedited a flow, and TCP returned it to the original sent order. Therefore, PTQ applies autonomous TCP sessions for each QoS priority. With PTQ, expediting flows across different priorities is not an issue because all traffic within each TCP flow has the same priority (H/M/L).

Virtual fabrics, logical switches, and Ethernet interface sharing

Moving a VE_Port into a VF-LS has no functional effect on the LAN side; however, for administration and role-based access, putting an IP Extension VE_Port into a logical switch can be done to reach those objectives.

Ethernet interface sharing applies only to circuits on the WAN side. The IP Extension gateways sit behind the LAN interfaces, and LAN interfaces must remain in the Default LS.

Note: IP Extension GE interfaces must exist in the Default LS and cannot be moved to a different context.

PortChannels

LAN side interfaces support IEEE 802.1AX Link Aggregation (LAG or PortChannel) to the data center LAN switches. Only a single “connection” per VLAN from the Extension platform to the data center LAN is supported; however, multiple links can be considered one logical connection. If the data center switches are a single virtual switch, which is often referred to as vPC, a LAN side PortChannel that spans the switches is considered a single connection.

As shown in Figure 7-29, four links form one LAG. From the viewpoint of the Extension platform, this configuration is a single connection.

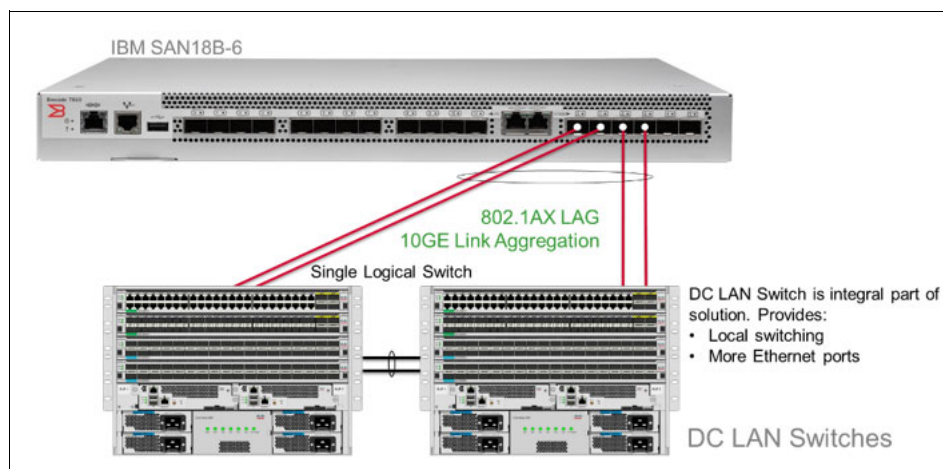


Figure 7-29 IP extension LAN side connectivity to two switches

Figure 7-30 shows a PortChannel that consists of two links to one data center LAN switch. Forming a LAG to a single data center switch is an alternative means of connectivity when the data center switches are not joined into a single virtual switch.

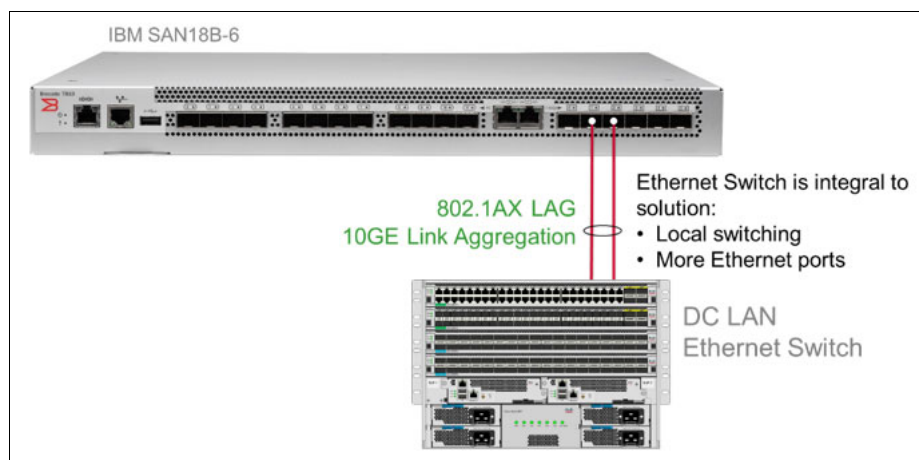


Figure 7-30 P Extension LAN Side Connectivity to a Switch

Two PortChannel modes are available: static and dynamic. The dynamic mode uses Link Aggregation Control Protocol (LACP) and requires a unique system priority (sysprio) that is configured on the data center LAN switches. A DC LAN switch can have multiple Link Aggregation (LAG) connections from various other systems; therefore, it is important to assign a unique sysprio (which is also known as *key*) to your dynamic LAG. Assigning a sysprio is not required when configuring a static LAG:

```
lACP --config -sysprio <priority>
lACP --sysprio 781
portchannel --create <po_name> -type <static|dynamic> [-key <po-number>]
portchannel --create MYLAG1 -type dynamic -key 781
portchannel --add <po_name> -port <port-num|port-range>
portchannel --add MYLAG1 -port ge2,ge3
```

IP extension compression

Compression operates within the scope of a tunnel. Compression cannot be configured per circuit. Compression must be configured identically at both tunnel ends; asymmetrical compression is not supported.

Fast-Deflate is not available for IP Extension on the IBM SAN42B-R and IBM SX6 Extension Blade; IP Extension compression is limited to Deflate and Aggressive-Deflate. The IBM SAN18B-6 does not have Fast-Deflate because the platform does not support the data rates that require Fast-Deflate.

Compression can be configured specifically for each protocol (FCIP and IP Extension). For example, on the IBM SAN42B-R or SX6 Blade, Fast-Deflate for FCIP and Deflate for IP Extension can be configured together. Configuring fast-deflate for FCIP is considered best practice in a hybrid environment because the Fast-Deflate and Deflate compression engines are different resources. Fast-Deflate uses a hardware engine; Deflate uses the processor. A total of 20 Gbps of FC ingress to Fast-Deflate does not use any IP Extension capacity that is available for Deflate or Aggressive-Deflate.

Table 7-3 provides a reference guide that helps determine which compression algorithm should be used by Extension platform in a hybrid mode environment

Table 7-3 Compression algorithm in hybrid modes

Hybrid mode	Brocade 7810	Brocade 7840	Brocade SX6 Blade
Disabled (1:1)	2.5 Gbps (VE max egress)	20 Gbps (VE max egress)	20 Gbps (VE max egress)
Fast Deflate (2:1)	N/A (not on the platform)	N/A (No IP Extension)	N/A (No IP Extension)
Deflate (3:1)	1.5 Gbps - 2.5 Gbps	10 Gbps - 15 Gbps	10 Gbps - 15 Gbps
Aggressive Deflate (4:1)	20 Mbps - 1.5 Gbps	20 Mbps - 10 Gbps	20 Mbps - 10 Gbps

Note: Compression ratios are approximate. IBM makes no warranties, guarantees, or claims about the actual compression ratio achieved with customer-specific data.

Traffic Control List

A Traffic Control List (TCL) is used to match LAN side ingress traffic to a specific outgoing tunnel. Because each TCL rule is assigned a name, give TCL rules meaningful human-readable names. IP Extension traffic does not flow until an active (`--admin enable`) TCL priority exists that matches the incoming traffic that directs that traffic into a specified target (tunnel/VE_Port). Multiple tunnels can exist, and specific traffic can be directed into a specific tunnel by matching the rule with the wanted target.

The priority number order of TCL rules matters because rule processing stops upon the first match. Rules are processed from lowest to highest priority number. The best practice is to pick priority numbers by leaving adequate space between if another rule must be added. For example, start at 10 and count by 10s when creating multiple rules.

Traffic cannot be matched to more than one tunnel because only a single tunnel can be specified in a rule matching a flow. Traffic that does not match any configured rule eventually reaches the bottom. The last rule has the highest priority number (65535); it is a “deny all” rule and is not modifiable or able to be deleted. Any traffic reaching the bottom is dropped.

Note: For TCL rules, the default admin mode is `--admin disabled`; therefore, ensure that the rule’s priority number has an “*” next to it, which indicates that it is active.

The hit counter for each TCL rule identifies matching incoming LAN traffic. The hit counter increases by 1 for every new TCP session formed (3-way handshake). Additionally, if the rule matches a UDP, ICMP, or broadcast, unknown, multicast (BUM) packet, the hit counter counts by 1 for every datagram. In addition to some BUM traffic, for example, broadcasts, no traffic should unintentionally arrive at the IP Extension gateway.

The TCL rules that are described next are named PINGS and MYTCL; use human-readable nomenclature that is meaningful to your environment. Priority 10 gives some space before (1 - 9) to add rules if needed. For example, a temporary rule can be added for counting ICMP pings that are arriving at the IP Extension gateway. Such a rule is useful for troubleshooting communications between end-devices and the IP Extension gateway.

The following below shows production rules (priority 10) and a temporary rule for troubleshooting by capturing pings (proto 1). The TCL hit counter registers any matched pings. You can ping local LAN side IP addresses from the LAN, but you cannot ping local WAN side IP addresses from the LAN. You can ping local WAN side IP addresses from the WAN, but you cannot ping the WAN's local LAN side IP address. The temporary PINGS rule does not need to be deleted; however, it should be *--admin disable* when not being used:

```
portcfg tcl <name> <option> [<args>]
Example: Captures ICMP pings to see on hit counter
portcfg tcl PINGS create --priority 8 --target 12 --src-addr 10.10.10.0/24
--dst-addr 192.168.10.0/24 --admin enable --l4proto icmp
Example: Local-SAN18B-6
portcfg tcl MYTCL create --priority 10 --target 12 --src-addr 10.10.10.0/24
--dst-addr 192.168.10.0/24 --admin enable
Example: Remote-SAN18B-6
portcfg tcl MYTCL create --priority 10 --target 12 --src-addr 192.168.10.0/24
--dst-addr 10.10.10.0/24 --admin enable
Local7810:admin> portshow tcl
```

Pri	Name	Flgs	Target Src-Addr	L2COS	VLAN	DSCP	Proto	Port	Hit
*8	PINGS	AI---	12-Med 10.10.10.0/24	ANY	ANY	ANY	1	ANY	29
*10	MYTCL	AI---	12-Med 10.10.10.0/24	ANY	ANY	ANY	ANY	ANY	248223
*65535	default	D----	- ANY	ANY	ANY	ANY	ANY	ANY	2949

Flags: *=Enabled ..=Name Truncated (see --detail for full name)

A=Allow D=Deny I=IP-Ext P=Segment Preservation

R=End-to-End RST Propagation N=Non Terminated.

Active TCL Limits: Cur / Max

DP0	3 / 32

Configured Total: 3 / 256

Do not create an open TCL in which everything can communicate with everything, including broadcast and multicasts. Sending broadcasts across the WAN is not considered best practice. However, it is not required to close communications to only the devices' exact IP addresses. Typically, specifying the entire end-device source and destination subnets (that is, -S 192.168.10.0/24 -D 10.10.10.0/24) works well.

It is imperative to understand TCL rules live on only the DP unto which the target VE_Port exists. VE_Ports exist only in one DP (DP0 or DP1). For more information, see "Traffic Control List" on page 186.

IP Extension gateways are created on a specific DP (DP0 or DP1). Ingress traffic that matches a TCL rule is directed to a specific target (VE_Port). If the VE_Port target that is designated in a TCL rule is not congruent with the IP Extension gateway's DP, the rule is not evaluated. That is, the IP Extension gateway and target VE_Port that are used must be on the same DP for the rule to be evaluated. If the rule is expected to be assessed but is not, check if traffic is hitting an IP Extension gateway configured on the other DP.

TCL rules are evaluated once when the TCP session first connects during the TCP 3-way handshake. After removing, adding, or making changes to TCL rules, the VE_Port for the tunnel must be `portdisable` and `portenable` to cycle the tunnel.

Cycling the VE_Port forces the proxied TCP sessions to reform their connections and be evaluated by the updated TCL. If bringing down the entire tunnel is too disruptive, for example, FCIP replication coexists within the same tunnel, it is possible to disable all the LAN side GE interfaces. Wait until all the TCP sessions time-out; then, re-enable the GE interfaces.

Another alternative is to disable the IP storage replication ports on the end-device. The ultimate goal is to get the TCP sessions to start new connections to engage the updated TCL.

The TCL is specific to IP Extension and has no bearing on WAN side IP routing or FC/FICON side functions. WAN-side IP interfaces (`ipif`) starting with `ge#.dp#` use `iproute` to direct outgoing traffic toward a specific router gateway. IP interfaces starting with `lan.dp#` use TCL to match incoming LAN side traffic to one particular tunnel for IP Extension.

FICON

Native FICON, such as FICON tape and z/OS GM (XRC), is supported by the IBM SAN42B-R and the SX6 Extension Blade. The Extension of FICON tape alongside TS7700 Grid is referred to as the high availability architecture. The host on either side can access the opposite side VTL if some portion of the infrastructure has failed.

When creating a tunnel for carrying FICON traffic, the FICON attribute must be specified within the tunnel scope. For example, `portcfg fcipunnel 24 create -F` creates a FICON tunnel on VE_Port 24.

It is essential to understand that array to array replication of mainframe volumes; for example, Global Mirror between two DS8900 is not FICON. Although the volumes are FICON, the GM replication of those volumes is FCP. Tunnels do not need to be specified as FICON when the replication protocol is FCP.

FICON connectivity requires the creation of a FICON LS (logical switch); for example, `1scfg --create 1 -ficon`. This CLI command creates a FICON logical switch with the Fabric ID (FID) of 1. The VE_Port must be moved into this context. The GE interfaces remains in the Default LS.

Note: Consider the following points:

- ▶ KATOV for FICON tunnels is automatically set to 1 second.
- ▶ FICON is not supported on the IBM SAN18B-R.

For more information about FICON Administration, see [Brocade Fabric OS FICON Configuration Guide, 8.1.0](#).

Virtual Fabrics

Virtual Fabrics (VF) are composed of one or more Logical Switches (LS) within the physical chassis. Connected Logical Switches form Logical Fabrics within the overall physical fabric. VF plays a primary role in achieving deterministic paths for Extension protocol optimization. Additionally, in mainframe environments, the implementation of FICON and FC Extension often requires unique configurations and management, which is achievable only within the same chassis when different LSs are used.

Protocol optimization requires all of an exchange's sequences and frames pass through the same VE_Port pair both outbound and return. This deterministic path means that only a single VE_Port can exist within a VF LS when enabling protocol optimization. Putting a single VE_Port in an LS defines a single path between the two LS. An LS can sufficiently scale FC port connectivity; they are operationally simplistic and create a stable environment. The IBM SAN42B-R and SX6 Extension Blade support VF; the SAN18B-6 does not.

Extension + FCR

An Extension tunnel nearly always traverses an IP WAN, which features different characteristics than an FC network. An Extension tunnel crossing a WAN is essentially an FC ISL over a WAN. An FCIP link is considered an ISL.

A WAN link that is experiencing flapping can disrupt a connected fabric if permitted to merge across the WAN. With each flap, disruption from fabric services that are attempting to converge then converge again and converge yet again across a WAN. The more devices that are logged in to the fabric, the more difficult convergence becomes. Convergence requires CPU and excessive convergence leads to an overload. If the CPU can no longer process tasks promptly, instability can occur.

Limiting fabric services within a local edge-fabric and not permitting services to span the WAN prevents large scale and taxing convergence. FCR is not required unless a production SAN must be isolated from potential WAN instability. The best practice is to construct separate production and replication SANs. FCR provides no benefit when only array replication ports are directly connected to a dedicated replication SAN; there are no production end devices to disrupt.

Isolated fabrics that are connected to FCR EX_Ports are referred to as *edge-fabrics*. EX_Ports are the demarcation points in which fabric services do not extend. Fabric services do not pass beyond an EX_Port, which forms an "edge." FCR constrains fabric services to within either an edge-fabric or the backbone.

The following basic architectures are available with and without FCR:

- ▶ The simplest example is no FCR. One fabric is the most common Extension implementation for both distributed systems and mainframe. Directly connecting storage array replication ports to the Extension boxes is highly reliable and easy to implement, manage, and troubleshoot.
- ▶ Edge-backbone-Edge by using FCR and Extension in which edge-fabrics bookend a transit Extension backbone fabric, as shown in Figure 7-31. The backbone fabric has EX_Ports that are outward facing to the edge-fabrics. Extension is within the backbone segment by using VE_Ports.

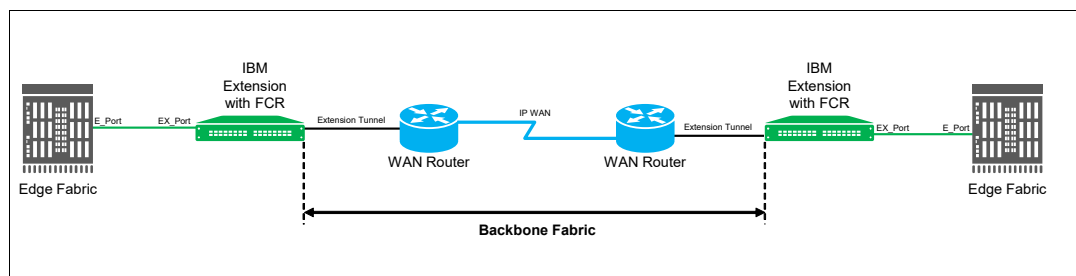


Figure 7-31 Edge-backbone-edge extension architecture

Note: Consider the following points:

- ▶ Mainframe FICON environments do not support FCR.
- ▶ When a mainframe host writes FICON to a direct access storage device (DASD) volume and subsequently performs RDR to a remote DASD, the array-to-array replication (Global Mirror - PPRC) does not use FICON. IBM Global Mirror replication is FC-based.

Compression

Three compression algorithms are available (for more information about compression, see “IP extension compression” on page 185). Specific to this section about FC/FICON is the Fast-Deflate algorithm. Fast-Deflate is on only the IBM SAN42B-R and SX6 Extension Blade and available to FC/FICON only. Fast-Deflate is not available for IP Extension or on the IBM SAN18B-6.

Fast-Deflate is an ultra-high-rate, ultra-low-latency hardware (FPGA) implemented compression algorithm. It can accommodate 40 Gbps FC/FICON ingress from the application and 20 Gbps egress to the WAN. These numbers can vary based on the compression ratio your data can achieve. The algorithm design is for an average of 2:1.

FC/FICON can use the Deflate (16 Gbps maximum ingress) and Aggressive-Deflate (10 Gbps maximum ingress) algorithms. The compression ingress data rates must be accommodating to the application and the wanted WAN side bandwidth:

- ▶ If Deflate gets 3:1 compression, the WAN side egress needs approximately 5 Gbps.
- ▶ If Aggressive-Deflate gets 4:1 compression, the WAN side egress needs approximately 2.5 Gbps.

Note: The incorrect assignment of a compression algorithm can result in a bottleneck.

7.3.2 FC/FICON side

Another side of Extension is the FC/FICON side, and it is associated with FCIP. Not all sides must be used; for example, IP Extension can be used without FC/FICON (which is supported). The FC/FICON switch in the Extension platforms is a full-fledged switch with the same Fabric OS (FOS), features, and functions of other IBM b-type Gen 7 SAN switches.

FCIP Only and Hybrid modes

Extension app-mode (`extncfg --show`) on the IBM SAN42B-R and SX6 Extension Blade can be set to FCIP Only or Hybrid (FCIP + IP Extension). If the implementation of IP Extension is intended, these platforms are put into Hybrid mode, which requires restarting the platform or slot.

If IP Extension is not used, the default is FCIP Only mode, which must be maintained. Internal resources are partitioned between FCIP and IP Extension when in Hybrid mode; therefore, it is best to remain in FCIP Only mode when not intending to use IP Extension, as shown in the following example:

```
LVN-7840-B_CA:FID1:admin> extncfg --show
APP Mode is FCIP
VE-Mode: configured for 10VE mode.
GE-Mode: Not Applicable.
```

Ethernet interface sharing

Any circuit from any VE_Port in any Logical Switch can use GE interfaces in the Default Switch.

It is possible to assign multiple IP addresses and circuits to the same Ethernet interface by assigning multiple ipif to the same interface, each with its unique IP address. The same IP address cannot be configured more than once in the same box. An IP address, ipif, and circuit can be associated with only one Ethernet interface. When more than one Ethernet interface is needed, multiple circuits must be created.

Brocade Extension Trunk (BET) uses a single VE_Port with multiple circuits. Frequently, circuits are assigned to different GE interfaces for the port, optic, and cable redundancy. Ethernet interfaces do not need to be dedicated to one circuit.

Sharing an Ethernet interface with multiple circuits that belong to different VF LSs is possible. The VE_Port is moved into the LS context by using the `lscfg` CLI command. The tunnel and circuits also reside in the LS context and are configured there.

The Ethernet interfaces, ipif (IP interface), and iproute (IP routes) must remain and be configured in the Default Switch (context 128). An ipif is configured by designating the IP address, subnet mask, VLAN, MTU, and GE interface.

The creation of an ipif assigns the IP address to an Ethernet interface. When the circuit is configured, the `--local-ip` binds the circuit to the Ethernet interface. Ethernet Interface Sharing facilitates efficient use of the 40GE interfaces by allowing multiple VE_Ports in multiple Logical Switches to use the large bandwidth connection.

Figure 7-32 shows an example of two VF LS (10 and 20).

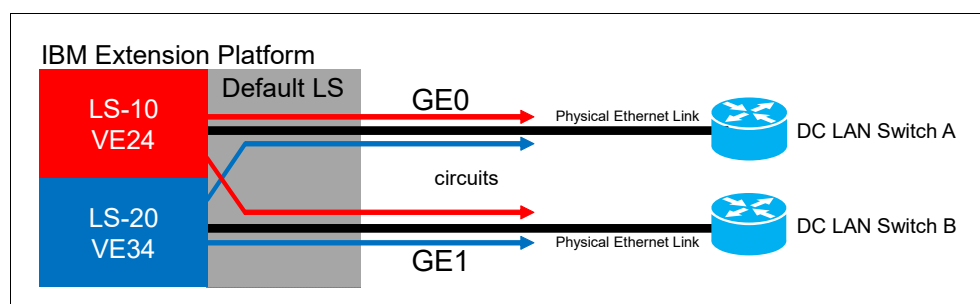


Figure 7-32 Virtual fabric logical switches Ethernet interface sharing

GE0 and GE1 are 40 Gbps interfaces and in the Default Switch. There are two circuits per BET (the red BET and blue BET as shown in Figure 7-32). VE24 (in DP0) has one circuit that goes to GE0 and one circuit that goes to GE1. VE34 (in DP1) has one circuit that goes to GE0 and one circuit that goes to GE1.

Two 10 Gbps circuits are used within each 40GE interface, each from a different LS. The 40GE interfaces are physically connected to respective A and B DC LAN switches. Enough 40GE bandwidth remains to create \ backup circuits (metric 1 circuits) if a 40GE link goes offline.

As shown in Figure 7-33, Ethernet Interface Sharing is used in an architecture where the 40GE interfaces are used for physical connectivity to the data center LAN. Circuits from three tunnels pass through the physical Ethernet links. Because the traffic on each circuit is VLAN tagged, the data center LAN switches can forward the traffic into specific DWDM circuits.

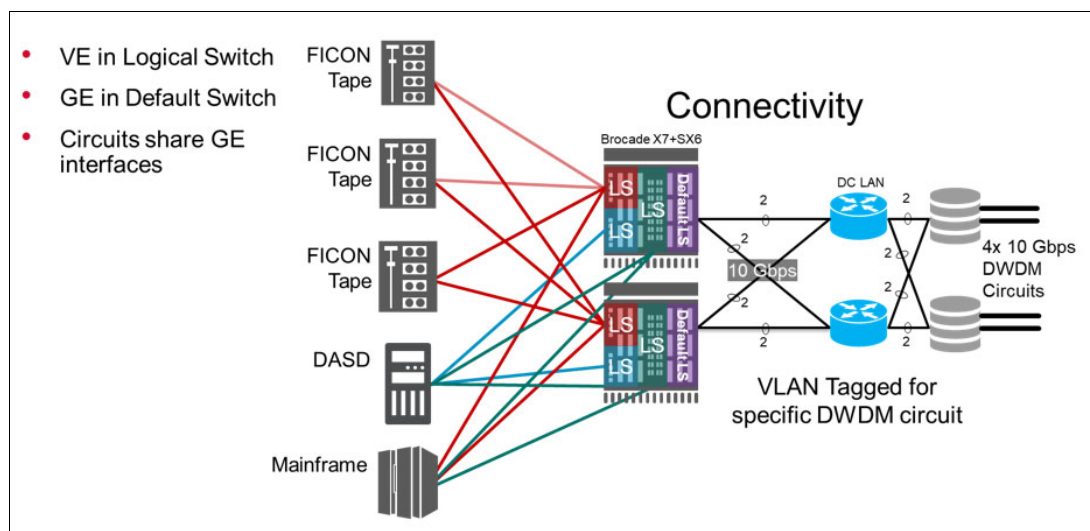


Figure 7-33 Ethernet interface sharing used in practice

7.3.3 LAN side

The LAN side is one of the three sides that is intrinsic to IBM Extension. The LAN side is synonymous with IP Extension and is only used to connect IP storage. By using IBM Extension technology, IP Extension accelerates, secures, and manages the bandwidth of supported IP storage applications.

FCIP and IP Extension protocols use the same tunnel technology, and sharing a common tunnel is considered a best practice. Latency and packet loss are detrimental to TCP/IP distance replication. IP Extension overcomes throughput degradation that is caused by these inherent characteristics, and it provides encryption.

This section covers points that are unique to IP Extension design, best practice, and architectures.

Use cases and purposes

IP Extension use cases include acceleration, encryption, bandwidth management, data migration, network visibility, troubleshooting tools, and congestion avoidance.

Figure 7-34 on page 193 shows a high availability architecture in an IBM Z with TS7700 environment.

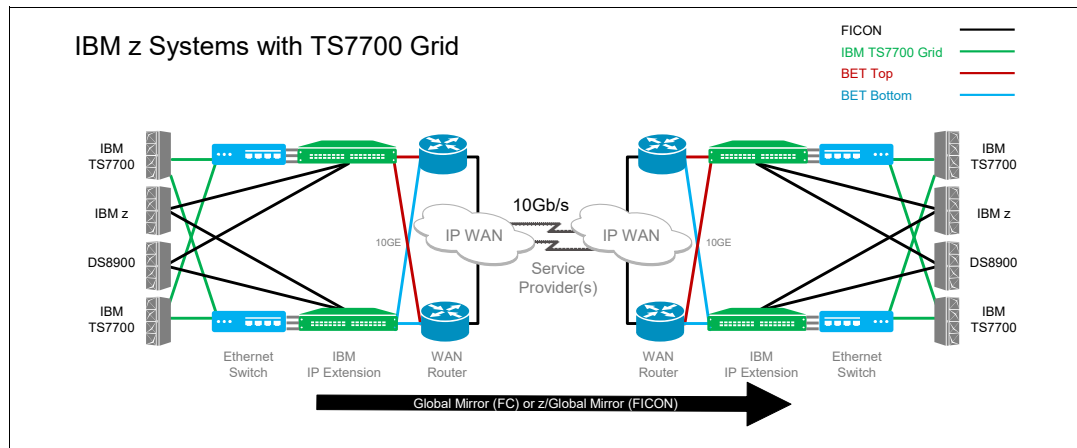


Figure 7-34 IP extension high availability architecture for TS7700

Another IP Extension solution is IBM Transparent Cloud Tiering (TCT). Data is transferred between VTLs in different data centers. Data is replicated quickly and securely to maintain a coherent Recovery Point Objective (RPO).

FCIP Only and Hybrid modes

Extension app-mode (`extncfg --show`) on the IBM SAN42B-R and SX6 Extension Blade can be set to FCIP Only or Hybrid (FCIP + IP Extension). If the implementation of IP Extension is intended, these platforms are put into Hybrid mode, which requires restarting the platform or slot. There is no IP Extension Only mode.

If IP Extension is not used, the default is FCIP Only mode, which must be maintained. Internal resources are partitioned between FCIP and IP Extension when in Hybrid mode; therefore, it is best to remain in FCIP Only mode when not intending to use IP Extension, as shown in the following example:

```
7840-Side-A:FID128:admin> extncfg --show
APP Mode is HYBRID (FCIP with IPEXT)
VE-Mode: configured for 10VE mode.
GE-Mode: Not Applicable.
```

VLANs

Ethernet frames can be tagged (IEEE 802.1Q) or untagged. Ethernet ISLs with tagging enabled are called *trunks*, which incorporates traffic from multiple VLANs. Respective VLAN traffic is differentiated by using the tags.

If IP Extension traffic is across multiple VLANs, only a single set of trunked links must be connected. All the VLANs can communicate across a single PortChannel of trunks. An IP Extension gateway is configured specifically for each subnet in each VLAN; which are the same subnets that the end device replication ports are on. Many IP Extension gateways can be used, depending on how many subnets are being used. The maximum number of IP Extension gateways that can be used per DP is eight.

If tagging is configured on the data center LAN switch, it must be configured on the Extension platform. A tagging mismatch causes Ethernet links not to come online. Although VLAN ID mismatches cause communication to fail, the links come online.

On the Extension platform, configuring VLAN tagging is accomplished by adding the VLAN ID when creating the ipif (portcfg ipif). As shown in Example 7-1, an IP Extension Gateway (10.0.0.2/24) is created and receives incoming and sends outgoing traffic on VLAN 10 only. When creating the ipif, the MTU can be specified as well; the default MTU is 1500 bytes.

Example 7-1 Creating an IP Extension Gateway on VLAN 10

```
portcfg ipif [<slot>/]<port> <option> [<args>]
portcfg ipif lan.dp0 create 10.0.0.2/24 vlan 10
```

In Example 7-2, the configured IP Extension gateways are shown.

Example 7-2 Showing configured IP Extension gateways

```
Local7810:admin> portshow ipif lan
Port          IP Address          / Pfx  MTU   VLAN  Flags
-----
lan.dp0       10.0.0.2            / 24   1500  10    U R M
-----
Flags: U=Up B=Broadcast D=Debug L=Loopback P=Point2Point R=Running I=InUse
       N=NoArp PR=Promisc M=Multicast S=StaticArp LU=LinkUp X=Crossport
```

If the IP storage traffic is on a single VLAN, VLAN tagging does not need to be configured (see Example 7-3). The data center LAN switch ports should be members of the VLAN without tagging. The Ethernet links should not be trunks. A LAG of untagged Ethernet links can be established between the LAN switches and the Extension platform.

Example 7-3 Creating an IP Extension Gateway without VLAN tagging

```
portcfg ipif [<slot>/]<port> <option> [<args>]
portcfg ipif lan.dp0 create 10.0.0.2/24
```

The configured IP Extension gateways are shown in Example 7-4.

Example 7-4 Showing configured IP Extension gateways

```
Local7810:admin> portshow ipif lan
Port          IP Address          / Pfx  MTU   VLAN  Flags
-----
lan.dp0       10.0.0.2            / 24   1500  0     U R M
-----
Flags: U=Up B=Broadcast D=Debug L=Loopback P=Point2Point R=Running I=InUse
       N=NoArp PR=Promisc M=Multicast S=StaticArp LU=LinkUp X=Crossport
```

IP Extension Gateway

IP Extension is the gateway for data crossing the WAN between data centers. If IP storage traffic is not forwarded to the IP Extension gateway, it does not use IP Extension.

Figure 7-35 on page 195 shows an example of traditional data center routing. Traffic comes from the IP storage cluster to the data center LAN switch. The LAN switch forwards it to the traditional gateway. The router can be an integral part of the LAN switch, or not. The router forwards the traffic onto the WAN towards the destination.

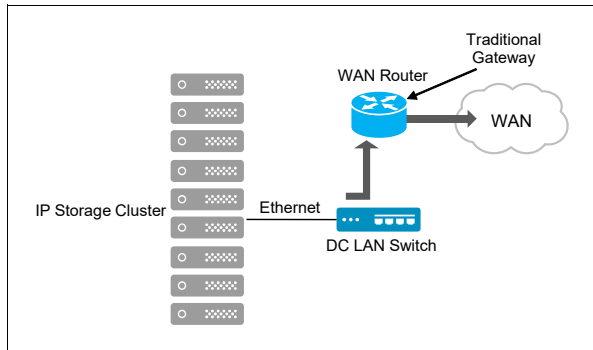


Figure 7-35 Data center traditional gateway implementation

IP Extension requires the end-device to direct traffic to its gateway, usually by static route or a gateway setting. Static routes are more specific than a default route; therefore, traffic for that specific destination subnet is forwarded to the IP Extension gateway instead of the default gateway. The destination subnet and IP Extension gateway are specified in the static route. The default route stays pointed to the traditional router gateway and is used when a more specific route does not exist.

Inserting IP Extension in the path and removing it from the path merely involves activating or deactivating these end-device static routes. With static routes, the traffic goes to the IP Extension gateway (IP Extension in the path). Without static routes, traffic goes to the traditional router (IP Extension out of the path).

IP Extension supports the connectivity of multiple data center LANs by using VLAN tagging (802.1Q). Tagging frames on Ethernet links that are connecting the Extension platform and the data center LAN switches is called *trunking*. An IP Extension gateway (`ipif lan.dp#` with `vlan <ID>`) must be configured for each subnet trunked to the Extension platform.

In Example 7-5, two `ipif` are created (one for each DP). The same subnet can be used; however, the gateways are different. The Ethernet trunks that are connected to the LAN interfaces must be tagging frames for VLAN 100 and VLAN 200. IP Extension gateways 10.10.10.2 and 10.10.10.3 are listening to VLAN 100 traffic only; and 192.168.10.2 and 192.168.10.3 are listening to VLAN 200 only. Of course, in your environment, the VLANs will be different.

Example 7-5 Creating IP Extension gateways

```
7840-Side-A:FID128:admin> portcfg ipif lan.dp0 create 10.10.10.2/24 vlan 100
Operation Succeeded.
7840-Side-A:FID128:admin> portcfg ipif lan.dp1 create 10.10.10.3/24 vlan 100
Operation Succeeded.
7840-Side-A:FID128:admin> portcfg ipif lan.dp1 create 192.168.10.3/24 vlan 200
Operation Succeeded.
7840-Side-A:FID128:admin> portcfg ipif lan.dp0 create 192.168.10.2/24 vlan 200
Operation Succeeded.
```

The IP Extension gateways are shown in Example 7-6.

Example 7-6 Showing the IP Extension gateways

```
7840-Side-A:FID128:admin> portshow ipif lan
```

Port	IP Address	/ Pfx	MTU	VLAN	Flags
lan.dp0	10.10.10.2	/ 24	1500	100	U R M

lan.dp0	192.168.10.2	/ 24	1500	200	U R M
lan.dp1	10.10.10.3	/ 24	1500	100	U R M
lan.dp1	192.168.10.3	/ 24	1500	200	U R M

Flags: U=Up B=Broadcast D=Debug L=Loopback P=Point2Point R=Running I=InUse
N=NoArp PR=Promisc M=Multicast S=StaticArp LU=LinkUp X=Crossport

Two DPs are on the IBM SAN42B-R and SX6 Extension Blade. The IP Extension gateway on each DP must be unique; the same gateway cannot be used on both DPs. Virtual Router Redundancy Protocol (VRRP) is not supported on the IP Extension platforms.

Only one IP Extension gateway can be configured per subnet on each DP. If you are planning to use both DPs, the storage application end devices must be segregated into groups that use each IP Extension gateway, or replication ports on the end devices must be capable of independent gateway settings.

It is imperative to understand TCL rules live on only the DP unto which the target VE_Port exists. VE_Ports exist in only one DP (DP0 or DP1). For more information, see “Traffic Control List” on page 186.

IP Extension gateways are created on a specific DP (DP0 or DP1). Ingress traffic that matches a TCL rule is directed to a specific target (VE_Port). If the VE_Port target that is designated in a TCL rule is not congruent with the IP Extension gateway’s DP, the rule is not evaluated. That is, the IP Extension gateway and target VE_Port used must be on the same DP for the rule to be evaluated. If the rule is expected to be assessed but is not, check whether traffic is hitting an IP Extension gateway configured on the other DP.

IP storage traffic that is forwarded to the IP Extension gateway is first evaluated by the Traffic Control List (TCL), as shown in Figure 7-36. If a flow matches an allow rule, the traffic is forwarded to the specified tunnel (target). The data is handled by the WAN side of the Extension platform for delivery to the remote data center.

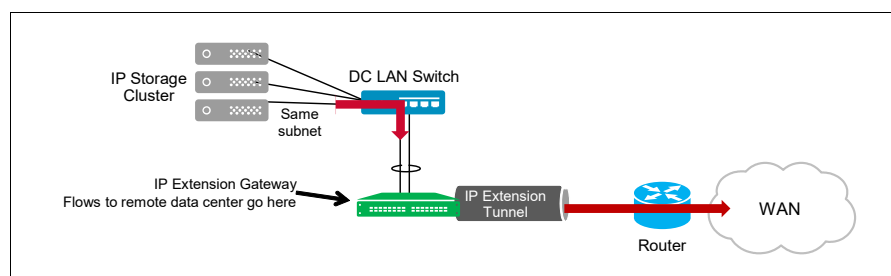


Figure 7-36 IP extension gateway

Note: The LAN side IP subnets are used on the IP storage end devices cannot be the same in the local and remote data centers. Different subnets for the end devices in each data center is required.

TCP flows, sessions, and streams

IP Extension operates entirely at layer 4 by using TCP. *Flows*, *sessions*, *connections*, and *streams* are synonymous TCP terms; each represents a connection that is established by a successful TCP 3-way handshake.

- ▶ IBM SAN18B-6256 TCP sessions
- ▶ IBM SAN42B-R512 TCP sessions per DP (two DPs)
- ▶ IBM SX6 Extension Blade512 TCP sessions per DP (two DPs)

WAN Optimized TCP, an aggressive purpose-built Extension TCP stack, is used to transport large storage quantities across long-distance WAN connections. Data is repackaged, optimized, and overhead is lowered by IBM b-type Gen 7 High-Efficiency Encapsulation for WAN side delivery.

The IBM TS7700 grid uses many IP Extension TCP sessions and can exhaust all of them. IBM TS7700 tape grid generates a few TCP sessions for transporting tape data, but generates hundreds of control TCP sessions in a cluster.

```
7840-Side-A:FID128:admin> portcfg tcl MyTS7700_CTL create --priority 100 --target
24 --src-addr 10.10.10.0/24 --dst-addr 192.168.10.0/24 --proto-port 1415-1416
--non-terminated enabled --admin enable
Operation Succeeded.
```

```
7840-Side-A:FID128:admin> portcfg tcl MyTS7700_TAPE create --priority 200 --target
24 --src-addr 10.10.10.0/24 --dst-addr 192.168.10.0/24 --admin enable
Operation Succeeded.
```

Pri	Name	Flgs	Target	L2COS	VLAN	DSCP	Proto	Port
Hit		Src-Addr			Dst-Addr			
*100	MyTS7700_CTL	AI--N	24-Med	ANY	ANY	ANY	ANY	1415-1416 0
		10.10.10.0/24			192.168.10.0/24			
*200	MyTS7700_TAPE	AI---	24-Med	ANY	ANY	ANY	ANY	0
		10.10.10.0/24			192.168.10.0/24			
*65535	default	D----	-	ANY	ANY	ANY	ANY	0
		ANY			ANY			

Chapter 7. IBM b-type Gen 7 extension design, implementation, and migration 197

A=Allow D=Deny I=IP-Ext P=Segment Preservation
 R=End-to-End RST Propagation N=Non Terminated.
 Active TCL Limits: Cur / Max

DP0	3 / 128
DP1	1 / 128

Configured Total:	3 / 1024

Note: Consider the following points:

- ▶ Beyond the supported number of flows, newly attempted TCP sessions are denied when exceeded; they are not merely passed through un-optimized. End devices must not produce an excessive number of flows, be configured to limit the maximum number of flows, or operate without fault when flows are not permitted to connect.
- ▶ IP storage applications must reconcile remote data delivery at the application layer and not rely on TCP acknowledgments. IP Extension operates by using TCP proxies, which provide local ACKs. Local TCP acknowledgments are not indicators of ultimate data arrival at the remote end device. An Extension platform or network failure that occurs after the local ACK and before the remote ACK results in data loss for which the sending end device saw an ACK. Then, it is up to the storage application to reconcile the data that was lost in-flight. IBM supported IP storage applications for IP Extension were tested to ensure data reconciliation at the application layer of data is lost in-flight.
- ▶ UDP and Broadcast, Unknown, Multicast (BUM) traffic is supported in small quantities. No optimization of these traffic types is performed. The handling of these traffic types is not done in hardware and requires processor intervention for every datagram that is received. It is not recommended to purposely send these traffic types to the IP Extension gateway or permit these traffic types in the TCL to cross the tunnel.

VE_Ports

VE_Ports for IP Extension are merely representative of the tunnel ID. Unlike FCIP, no IP Extension data flows through a VE_Port. Although, disabling a VE_Port disables the entire tunnel, even if configured for IP Extension only. The VE_Port must be up for the tunnel to be up.

The best practice is to use the same VE_Port for FCIP and IP Extension between the same two domains; tunnels are point-to-point. You want the VE_Port to manage the bandwidth of both protocols across the tunnel. Bandwidth management cannot be accomplished when different VE_Ports are used (one for each protocol).

It is imperative to understand TCL rules live on only the DP unto which the target VE_Port exists. VE_Ports exist in one DP only (DP0 or DP1). For more information, see "Traffic Control List" on page 186.

IP Extension gateways are created on a specific DP (DP0 or DP1). Ingress traffic that matches a TCL rule is directed to a specific target (VE_Port). If the VE_Port target that is designated in a TCL rule is not congruent with the IP Extension gateway's DP, the rule is not evaluated. That is, the IP Extension gateway and target VE_Port used must be on the same DP for the rule to be evaluated. If the rule is expected to be assessed but is not, check whether traffic is hitting an IP Extension gateway that is configured on the other DP.

7.4 Extension migration strategy

The primary technology that is used to deliver performance and reliable data replication is Fibre Channel over IP (FCIP). FCIP uses high-efficiency encapsulation to form TCP/IP datagrams from FC frames. WAN Optimized TCP transports data over long-distance connections from anywhere to anywhere: across the country to around the globe.

IBM continues to deliver the industry's most advanced Extension products to achieve the most predictable, stable, and performant replication networks. These products give clients the ability to protect their most critical asset reliably: data.

The most common data protection scheme that is used by enterprise customers is disk replication or mirroring. Other methods are available for protecting data, but array-based mirroring is one of the most commonly used methods in the enterprise market today. The IBM Extension platforms are designed to excel in this environment. They deliver the most outstanding data protection and performance, regardless of distance.

With the ever-increasing amount of data that must be protected and the increase in WAN speed from GE to 10 GE to 100 GE, many older IBM Extension platforms are reaching the end of their product support lifecycle. For those charged with business continuance (BC) or disaster recovery (DR) for their enterprise data, it is a mistake not to review the current BC/DR architecture to avoid an unsupported platform situation.

This section describes the general principles of migrating and refreshing a current environment to a new, modernized Extension platform.

7.4.1 IBM Extension platforms

Today, IBM offers three platforms to help clients meet their long-distance replication requirements. To assist our largest enterprise clients with director-class SAN switching platforms, IBM offers a "blade" (SX6 blade) that integrates directly into those director platforms. The SX6 Extension Blade transports FC frames through the IP WAN to remote locations where secondary copies are kept. The SX6 Extension Blade is compatible with the IBM SAN256B-6, SAN512B-6, SAN256B-7, and SAN512B-7 Director platforms to help design flexibility and more cost-effective solutions. SX6 Extension tunnels are interoperable with the IBM SAN18B-6 and SAN42B-R.

For clients that require high port density for FC connectivity and Extension tunnel bandwidth but might not be deploying directors, IBM offers a 2RU platform called the IBM SAN42B-R Extension Switch. The IBM SAN42B-R and SX6 Extension Blade are considered enterprise-class systems and support all of the solutions for open systems and mainframe (FICON) environments.

The third offered Extension platform is the IBM SAN18B-6 Extension Switch; its predecessor was the IBM SAN06B-R. The IBM SAN18B-6 is a 1RU system that incorporates many of the advanced functions that are found in the SX6 and SAN42B-R. The SAN18B-6 can be integrated into enterprise-class solutions, but it also supports clients who do not need nor require high port densities and large Extension tunnel bandwidths. The IBM SAN18B-6 supports all disk replication solutions but does not support specific mainframe solutions that implement FICON or emulation techniques.

The IBM SAN06B-R platform was introduced over a decade ago and now entered End of Life (EOL). Thousands of production units are used today, and those units must be refreshed before the End of Service (EOS) date on October 31, 2024. The remainder of this section focuses on the methods and strategies for refreshing the IBM SAN06B-R infrastructure.

7.4.2 IBM Extension solution designs

IBM Extension platforms give users the flexibility to create a highly effective infrastructure to meet their current and future data replication needs. Extension technology and its ability to intermix with an FC SAN allows for endless possibilities in infrastructure design. To narrow this section's scope, we focus on the most common network designs for BC/DR, emphasizing migration from EOL Extension platforms.

The most common Extension deployments contain redundant fabrics, dual paths that connect the storage array through an IP network to a remote site. Having two paths between data centers offsets the impact if one path fails. Redundancy and data path protection are intrinsic to building the most resilient infrastructure.

Users must have different IP WAN service providers when redundant paths are needed. Different WAN service providers create a tremendous amount of protection by having wholly separate and disparate routes between sites. History shows that users who do not confirm that the paths are entirely separate and disparate likely are affected during an outage event, such as a backhoe digging for a new sewer line. Figure 7-37 shows the “dual fabric” configuration.

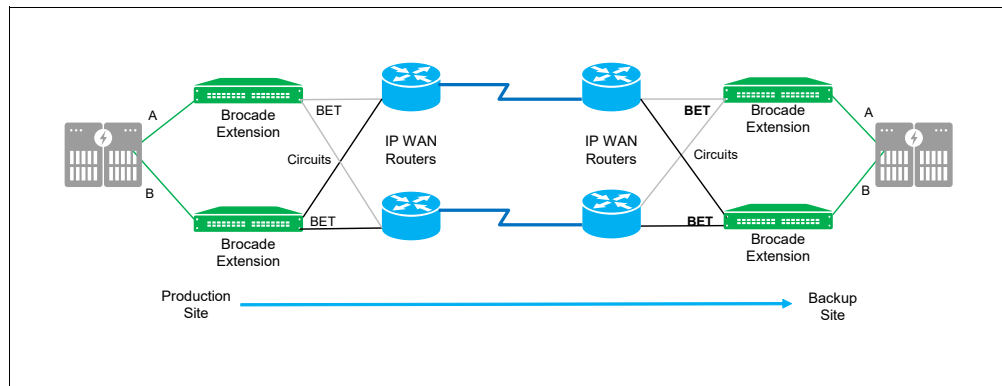


Figure 7-37 Dual Replication SAN

The example dual fabric replication SAN shows how a storage array on the left side of Figure 7-37 features two separate and isolated paths. Each IBM SAN18B-6 makes up replication Fabric A and Fabric B. Those paths are “paired” with the remote storage array on the right side.

The paths are paired A-to-A and B-to-B. Each Extension switch has two Ethernet connections to the IP network. The IP network includes redundant switches, routers, and WAN connections. By designing a network in this way, you can protect your replication scheme from single points of failure. This example of a replication architecture is most commonly deployed.

7.4.3 Migration methodology and techniques

Replacing any part of your BC/DR infrastructure can be stressful, especially if a working infrastructure exists and processes are in place. Fortunately, changing from an existing Brocade Extension platform to a newer model limits migration-related anxiety. You can achieve greater performance and enhanced functions when working with a familiar platform that is built upon proven generations of IBM Extension technology.

One of the most stressful parts of migrating to a new Brocade Extension platform is the issue of whether an outage is necessary and if so, how long it lasts. The short answer is an outage might be needed. Users can minimize replication downtime when following proper preparation and methods.

To remove nearly all the risk of taking a full outage, plan to stage and configure your new platforms to mimic as close as possible the same configurations that are used today. Suppose you are going through a consolidation effort, increasing the number of WAN connections, or upgrading GE to 10GE Ethernet. In that case, more steps are needed to integrate those changes. For more information about the changes and upgrades, see [Brocade Fabric OS Extension User Guide, 9.0.x](#).

In general, the following migration methodology and techniques can be applied to any platform refresh strategy. However, this section focuses on migrating from the IBM SAN06B-R to the SAN18B-6 platform. This migration section is not intended to be a step-by-step technical guide; instead, it should be used as a planning guideline.

For a more information about migrating from the IBM SAN16B-R to the IBM SAN18B-6, see [Brocade 7800 Extension to Brocade 7810 Extension Migration User Guide](#).

This migration example assumes a one-for-one replacement of the IBM SAN06B-R to SAN18B-6 with no FC connection changes or Extension circuit changes. It also assumes the same subnets and bandwidth settings are used for the Extension circuits. However, the IP addresses on the subnet can be the same or different. By using Figure 7-37 on page 200 as our example network, we walk through some guidelines for migrating platforms.

A benefit of having a “dual fabric” architecture is to provide a seamless transition from one platform to another. By using one of the fabrics (or paths), replication can be maintained without taking a full outage.

The following migration tasks are used as a guideline to achieve a seamless transition from an older extension platform to a new extension platform. Specific migration plans and steps must be developed for specific situations:

- ▶ Preliminary work: Run Brocade SAN Health to capture a snapshot of your environment to see if any current “network health” problems exist that require attention before the migration. For more information, see [Brocade SAN Health Frequently Asked Questions Version 4.2.2a](#).
- ▶ Install and apply power to the IBM SAN18B-6 at both sites. Leave the storage and network cables disconnected.
- ▶ Gather required information from the SAN06B-R for building and configuring the IBM SAN18B-6. Record the output from the following CLI commands.
 - Local IP addresses and subnet masks (`portshow ipif`).
 - Gateways and remote subnets (`portshow iproutes`).
 - Min and Max bandwidth rates (`portshow fciptunnel -c`).
- ▶ Configure the IBM SAN18B-6 at both sites with the configuration information that is gathered from the IBM SAN18B-6. The Brocade 7810s can be configured by using one of the following methods:
 - IBM SANnav Management Tool.
 - CLI.

- ▶ Now, the Extension tunnel on the IBM SAN18B-R should have a status of “Up.” If not, do not proceed. The tunnel must be operational before proceeding. For more information about trouble-shooting the tunnel and circuits, see 7.5, “Extension validation” on page 203.
- ▶ Select which fabric you intend to migrate first, and, if required by the storage vendor’s process, suspend the path on the storage array’s replication interfaces:
 - Review your storage vendor’s administration guide for the proper process to take down one or more replication paths, and so on.
 - With a dual fabric design, the other fabric or path continues replication; however, some backlog in replication might occur with half the bandwidth if a heavy workload exists.
- ▶ For the first path to be migrated, disable power to the IBM SAN06B-R at both sites.
- ▶ Remove storage and network cables and reattach the cables to the IBM SAN18B-R.
- ▶ Check for Extension tunnel status and error messages.

Note: The new path should be online and can replicate. You can skip the WAN Test Tool validation if you feel confident that the IP network’s load testing is unnecessary.

- ▶ If you want to confirm that the IP WAN can support your bandwidth SLA, disable the circuits, create Wtool sessions, and run Wtool at the bit rate you expecting on the WAN. The Wtool must be run for at least 10 minutes. Review the results.
- ▶ After the path has been verified, disable the Wtool sessions, and enable the circuits that were disabled.
- ▶ After the Extension tunnel is operational, enable storage replication interfaces.
- ▶ Monitor replication and the network for errors, and confirm everything is operating as expected.
- ▶ After the initial fabric or path is working as expected, repeat these steps on the other fabric.

As you can see, by taking down paths individually, the need for a complete outage can be reduced or eliminated.

The IBM Network Advisor, the legacy SAN Management tool for managing the IBM SAN06-R, is no longer available. A newer SAN Management tool is available, which features a new look and feel, and was redesigned for the next generation of SAN infrastructure. IBM SANnav comprises two management portals for local and global views to monitor and manage small to large environments.

For more information about IBM SANnav, see [Brocade SANnav Management Portal User Guide, 2.1.0x](#).

7.4.4 Migration summary

As the digital world explodes with new workloads and the risk of losing your most critical data becomes increasingly intense, having the assurances of securing your data with the most current and flexible infrastructure becomes a necessity to ensure business sustainability. IBM Extension switching platforms are relevant with the increased use of Flash and new technologies, such as NVMe over Fabrics (FC). Keeping the network one step ahead of servers and storage is critical to ensuring any investment in those new endpoints gives you the expected performance.

It is also essential to understand the need for product lifecycle management and the eventual migration to newer supported technologies. As you continue to deploy new workloads and enhance protection requirements with solutions, such as Remote Disk Replication (RDR), avoid explaining sudden budget requirements for refreshing replication infrastructure when it is imminent (or beyond) the ended support date.

When migrating from the IBM SAN16B-R (Brocade 7800) to the IBM SAN18B-6 (Brocade 7810), a few considerations are important. Extension technology evolved; features, functions, and the CLI have changed. Notable changes include IPsec and encryption, IP Extension, licensing, WAN side bandwidth, and various CLI syntax and references.

7.5 Extension validation

After setting up new or troubleshooting existing circuits, validating connectivity is the next logical step. The following tools are useful for validating WAN side connectivity:

► **portcmd --wtool**

The WAN Test Tool (Wtool) is a base feature that is included with all IBM Extension platforms. W-tool is the most comprehensive and thorough of the validation tools. The Wtool utility can generate line-rate traffic between the local and remote Extension platforms by way of one or more of the tunnel's circuits. Wtool simulates circuits by using the same IP addresses, rate limiting, IPsec, QoS, TCP sessions, and TCP ports to ensure realistic measurements over pathways. Wtool can be run for any length of time. The following Wtool output shows measurements taken during the test interval:

```
Local7810:admin> portcmd --wtool show --peer
```

WTool Session (0)	(Local)	(Remote)

Admin / Oper State	: Up	: Up
Up Time	: 5m57s	: 5m57s
Run Time	: 15m0s	: 0s
Time Remaining	: -	: 15m0s
Port	: ge2	: -
IP Addr	: 172.16.1.3	: 172.16.2.3
IP-Sec Policy	: Enabled	: Enabled
Configured Comm Rate	: 1000000 kbps	: 1000000 kbps
Actual Comm Rate	: 1000000 kbps	: 1000000 kbps
PMTU Discovery (MTU)	: disabled (1500)	: disabled (1500)
L2cos /DSCP	: (none) / (none)	: (none) / (none)
Tx Rate	: 963893.01 Kbps	: 2073.04 Kbps
Rx Rate	: 2013.07 Kbps	: 963112.38 Kbps
Tx Utilization	: 98.03%	: 4.15%
Rx Utilization	: 4.03%	: 98.47%
RTT (Min/Max)	: 19 ms/204 ms	: 20 ms/249 ms
RTT VAR (Min/Max)	: 1 ms/337 ms	: 1 ms/372 ms
Tx pkts	: 12348259	: 0
Rx pkts	: 0	: 12348245
Ooo pkts	: 0	: 0
Drop pkts	: 0	: 14
Drop% (Overall / 5s)	: 0.00% / 0.00%	: 0.00% / -

Best practice is to perform a comprehensive evaluation and baseline of the IP network before placing an Extension tunnel into production. The evaluation can run for a few minutes, days, or longer, depending on the important period of interest for testing the IP network; for example, during end of quarter processing. The IP network often is not what is expected.

An existing circuit must be used with Wtool; however, the circuit must be taken offline before the associated Wtool session can be enabled. Other circuits can remain in production. Disable only the circuits that you want to test at one time.

Wtool is configured on both ends, as with any circuit. The Wtool session bit-rate must be configured; the circuit ARL settings are not used.

After completing the Wtool evaluation, the sessions can be disabled, and the circuits enabled again. Sessions can be used again later; they do not need to be deleted. Upon re-enabling the circuits, they are back in production.

Wtool reports the amount of local data sent and remote data that is received to determine the IP network's characteristics and reliability, including packet drops, latency (RTT), jitter (RTT Var), and out-of-order segments plus more. Typically, one end is the data source, and the other end is the data sink; however, bidirectional traffic is supported. The process continues for the user-specified duration or until keyboard terminated (by pressing Ctrl+C).

For more information about setting up and using Wtool, see the [Brocade 7810 IP Extension Switch User Guide](#).

► **portcmd --ping**

The **portcmd --ping** command tests simple connectivity between a local IP address (ipif) on a specific GE interface and a destination IP address. ICMP ECHO protocol is used. The **portcmd --ping** command requires designating which GE interface the ping is sent from, which can be `lan.dpx` or `ge.dpx` interfaces. The source IP address for the ping must exist on the specified LAN or GE interface.

A LAN side ping to an end device must be on the same subnet as the IP Extension gateway. Pinging a different subnet requires a configured iproute on the Extension platform. The destination device might also require a complimenting return route.

In the following example, a ping was done from the ipif `lan.dp0` IP Extension gateway to the traditional router gateway on the same VLAN. The ping verifies connectivity and communications from the IP Extension gateway through the data center LAN switch:

```
LVN-7840-A:FID128:admin> portcmd --ping lan.dp0 -s 10.175.1.248 -d 10.175.1.1
PING 10.175.1.1 (10.175.1.248) with 64 bytes of data.
64 bytes from 10.175.1.1: icmp_seq=1 ttl=255 time=37 ms
64 bytes from 10.175.1.1: icmp_seq=2 ttl=255 time=7 ms
64 bytes from 10.175.1.1: icmp_seq=3 ttl=255 time=1 ms
64 bytes from 10.175.1.1: icmp_seq=4 ttl=255 time=100 ms

-- 10.175.1.1 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 887 ms
rtt min/avg/max = 1/36/100 ms
```

If no local or return IP route exists, ping a destination within the subnet, such as the local gateway (typically, 1). A native subnet ping demonstrates if L2 (Ethernet level) connectivity is operational through the connected data center switch. If local and return IP routes exist, you can ping to a foreign subnet destination; however, the remote device must have a complimenting IP return route to reply. A foreign subnet ping demonstrates if L3 (IP level) connectivity is operational:

```
Portcmd --ping [<sx6 slot>/]<port.dp#> <options>
```

Example ping:

```
Local7810:admin> portcmd --ping ge6.dp0 -s 172.16.1.3 -d 172.16.2.3
PING 172.16.2.3 (172.16.1.3) with 64 bytes of data.
64 bytes from 172.16.2.3: icmp_seq=1 ttl=250 time=19 ms
64 bytes from 172.16.2.3: icmp_seq=2 ttl=250 time=19 ms
64 bytes from 172.16.2.3: icmp_seq=3 ttl=250 time=19 ms
64 bytes from 172.16.2.3: icmp_seq=4 ttl=250 time=19 ms
-- 172.16.2.3 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 789 ms
rtt min/avg/max = 19/19/19 ms
```

Note: The **ping** command is used for the management interface only. The **portcmd --ping** command is used for the GE interfaces.

Traceroute should be used on L3 routed networks with two or more hops. The following example includes two hops: first is to the router, the second is to the Extension platform. The intermediate and remote devices must have a complementing IP route to return the traceroute response. It is not unusual for intermediate devices that are owned by service providers to not return traceroute responses:

```
7840-A:FID128:admin> portcmd --traceroute ge3.dp0 -s 10.2.0.10 -d 10.1.0.10
traceroute to 10.1.0.10, 64 hops max, 64 byte packets
 1 10.2.0.1 59 ms 59 ms 59 ms
 2 10.1.0.10 59 ms 59 ms 59 ms
```

► nsShow

When connecting end devices to an FC fabric, they log into the fabric, the **nsshow** CLI command shows the devices that logged into the fabric and information about those devices. If the device successfully logged into the fabric, it is connected and communicating with the switch:

```
7840-Side-A:FID128:admin> nsshow
{
  Type Pid      COS      PortName      NodeName      SCR
  N      0a0000;    3;10:00:00:10:9b:34:6e:e0;20:00:00:10:9b:34:6e:e0;
0x00000003
  SCR: Fabric-Detected Nx-Port-Detected
  FC4s: FCP
  PortSymb: [34] "Emulex PPN-10:00:00:10:9b:34:6e:e0"
  NodeSymb: [104] "Emulex LPe32002-M2 FV11.4.204.20 DV11.4.33.1
HN:Lenx3650M5-212999.dhcp.broadcom.net OS:VMware ESXi 6.7.0"
  Fabric Port Name: 20:00:c4:f5:7c:3b:24:86
  Permanent Port Name: 10:00:00:10:9b:34:6e:e0
  Device type: Physical Initiator
  Port Index: 0
  Share Area: No
  Redirect: No
  Partial: No
  LSan: No
  Slow Drain Device: No
  Device link speed: 16G
  Connected through AG: No
  Real device behind AG: No
  FCoE: No
  FC4 Features [FCP]: Initiator
```

```

N    0a0100;      3;10:00:00:10:9b:34:6e:e1;20:00:00:10:9b:34:6e:e1;
0x00000003
SCR: Fabric-Detected Nx-Port-Detected
FC4s: FCP
PortSymb: [34] "Emulex PPN-10:00:00:10:9b:34:6e:e1"
NodeSymb: [104] "Emulex LPe32002-M2 FV11.4.204.20 DV11.4.33.1
HN:Lenx3650M5-212999.dhcp.broadcom.net OS:VMware ESXi 6.7.0"
Fabric Port Name: 20:01:c4:f5:7c:3b:24:86
Permanent Port Name: 10:00:00:10:9b:34:6e:e1
Device type: Physical Initiator
Port Index: 1
Share Area: No
Redirect: No
Partial: No
LSAN: No
Slow Drain Device: No
Device link speed: 16G
Connected through AG: No
Real device behind AG: No
FCoE: No
FC4 Features [FCP]: Initiator
The Local Name Server has 2 entries }

```

► switchshow

Switchshow is useful for an overall summary of the fabric, domain, FC/FICON ports, state, connected devices, VE_Ports, GE interfaces, and logical switch context, as shown in the following example:

```

7840-Side-A:FID128:admin> switchshow
switchName:      7840-Side-A
switchType:      148.0
switchState:     Online
switchMode:      Native
switchRole:      Principal
switchDomain:    10
switchId:        fffc0a
switchWwn:       10:00:c4:f5:7c:3b:24:86
zoning:          OFF
switchBeacon:    OFF
FC Router:       OFF
Fabric Name:     REPLICATION
HIF Mode:        OFF
Allow XISL Use:  OFF
LS Attributes:   [FID: 128, Base Switch: No, Default Switch: Yes, Ficon Switch:
No, Address Mode 0]

```

Index	Port	Address	Media	Speed	State	Proto
0	0	0a0000	id	N16	Online	FC F-Port
10:00:00:10:9b:34:6e:e0						
1	1	0a0100	id	N16	Online	FC F-Port
10:00:00:10:9b:34:6e:e1						
2	2	0a0200	id	N8	No_Light	FC
3	3	0a0300	id	N16	No_Light	FC
4	4	0a0400	id	N16	No_Light	FC
5	5	0a0500	id	N16	No_Light	FC
6	6	0a0600	id	N16	No_Light	FC

7	7	0a0700	id	N16	No_Light	FC	
8	8	0a0800	id	N16	No_Light	FC	
9	9	0a0900	id	N16	No_Light	FC	
10	10	0a0a00	id	N16	No_Light	FC	
11	11	0a0b00	id	N16	No_Light	FC	
12	12	0a0c00	id	N16	No_Light	FC	
13	13	0a0d00	id	N16	No_Light	FC	
14	14	0a0e00	id	N16	No_Light	FC	
15	15	0a0f00	id	N16	No_Light	FC	
16	16	0a1000	id	N16	No_Light	FC	
17	17	0a1100	id	N16	No_Light	FC	
18	18	0a1200	id	N16	No_Light	FC	
19	19	0a1300	--	N16	No_Module	FC	
20	20	0a1400	id	N16	No_Light	FC	
21	21	0a1500	id	N16	No_Light	FC	
22	22	0a1600	id	N16	No_Light	FC	
23	23	0a1700	--	N16	No_Module	FC	
24	24	0a1800	--	--	Online	VE	VE-Port
10:00:c4:f5:7c:3b:25:06 "7840-Side-B" (downstream)							
25	25	0a1900	--	--	Offline	VE	
26	26	0a1a00	--	--	Offline	VE	
27	27	0a1b00	--	--	Offline	VE	
28	28	0a1c00	--	--	Offline	VE	
29	29	0a1d00	--	--	Offline	VE	Disabled (10VE Mode)
30	30	0a1e00	--	--	Offline	VE	Disabled (10VE Mode)
31	31	0a1f00	--	--	Offline	VE	Disabled (10VE Mode)
32	32	0a2000	--	--	Offline	VE	Disabled (10VE Mode)
33	33	0a2100	--	--	Offline	VE	Disabled (10VE Mode)
34	34	0a2200	--	--	Offline	VE	
35	35	0a2300	--	--	Offline	VE	
36	36	0a2400	--	--	Offline	VE	
37	37	0a2500	--	--	Offline	VE	
38	38	0a2600	--	--	Offline	VE	
39	39	0a2700	--	--	Offline	VE	Disabled (10VE Mode)
40	40	0a2800	--	--	Offline	VE	Disabled (10VE Mode)
41	41	0a2900	--	--	Offline	VE	Disabled (10VE Mode)
42	42	0a2a00	--	--	Offline	VE	Disabled (10VE Mode)
43	43	0a2b00	--	--	Offline	VE	Disabled (10VE Mode)
	ge0		--	40G	No_Module	FCIP	
	ge1		--	40G	No_Module	FCIP	
	ge2		id	10G	Online	FCIP	
	ge3		--	10G	No_Module	FCIP	
	ge4		id	10G	Online	FCIP	
	ge5		--	10G	No_Module	FCIP	
	ge6		--	10G	No_Module	LAN	
	ge7		--	10G	No_Module	LAN	
	ge8		--	10G	No_Module	LAN	
	ge9		--	10G	No_Module	LAN	
	ge10		--	10G	No_Module	FCIP	
	ge11		--	10G	No_Module	FCIP	
	ge12		--	10G	No_Module	FCIP	
	ge13		--	10G	No_Module	FCIP	
	ge14		id	10G	No_Light	FCIP	
	ge15		id	10G	No_Light	FCIP	
	ge16		--	10G	No_Module	FCIP	

```
ge17      --      10G      No_Module  FCIP
```

Example of switchshow with a FICON context used for Extension:

```
LVN-7840-B_CA:FID1:admin> switchshow
switchName:    LVN-7840-B_CA
switchType:    148.0
switchState:   Online
switchMode:    Native
switchRole:    Principal
switchDomain:  16
switchId:      fffc10
switchWwn:     10:00:50:eb:1a:69:ee:93
zoning:        OFF
switchBeacon:  OFF
FC Router:     OFF
HIF Mode:      ON
Allow XISL Use: ON
LS Attributes: [FID: 1, Base Switch: No, Default Switch: No, Ficon Switch:
Yes, Address Mode 1]
Index Port Address Media Speed State Proto
=====
0 0 100000 id N16 Online FC F-Port
c0:50:76:cf:ed:08:1c:d1
1 1 100100 id N16 Online FC F-Port
c0:50:76:cf:ed:08:18:11
2 2 100200 id N16 No_Light FC Disabled (Persistent)
3 3 100300 id N16 No_Light FC Disabled (Persistent)
4 4 100400 -- N16 No_Module FC Disabled (Persistent)
5 5 100500 -- N16 No_Module FC Disabled (Persistent)
24 24 101800 -- -- Online VE VE-Port
10:00:50:eb:1a:fc:e0:da "BRM-7840_CA" (downstream)
```

► **fabricshow**

When making connections between domains, such as FCIP connections, fabricshow displays the merged switches that joined the fabric. In this example, two switches with domains 10 and 20 joined fabric REPLICATION. The management IP address of each switch also are shown. Before FCIP traffic can flow between connected end devices (end-to-end), the two domains that are connected by FCIP must first form a fabric:

```
7840-Side-A:FID128:admin> fabricshow
Switch ID Worldwide Name Enet IP Addr FC IP Addr Name
-----
10: fffc0a 10:00:c4:f5:7c:3b:24:86 10.155.2.221 0.0.0.0
>"7840-Side-A"
20: fffc14 10:00:c4:f5:7c:3b:25:06 10.155.2.220 0.0.0.0
"7840-Side-B"
The Fabric has 2 switches
Fabric Name: REPLICATION
```

► **portshow**

Portshow displays detailed information about FC/FICON ports. In the example below, FC port 0 is shown:

```
7840-Side-A:FID128:admin> portshow 0
portIndex: 0
portName: port0
portHealth: HEALTHY
```

```

Authentication: None
portDisableReason: None
portCFlags: 0x1
portFlags: 0x20b03      PRESENT ACTIVE F_PORT G_PORT U_PORT LOGICAL_ONLINE
LOGIN NOELP ACCEPT FLOGI
LocalSwcFlags: 0x0
portType: 24.0
portState: 1      Online
Protocol: FC
portPhys: 6      In_Sync      portScn: 32      F_Port
port generation number: 0
state transition count: 1
portId: 0a0000
portIfId: 43020028
portWwn: 20:00:c4:f5:7c:3b:24:86
portWwn of device(s) connected:
      10:00:00:10:9b:34:6e:e0
16b Area list:
Distance: normal
portSpeed: N16Gbps
FEC: Inactive
Credit Recovery: Inactive
Aoq: Inactive
FAA: Inactive
F_Trunk: Inactive
LE domain: 0
Peer beacon: Off
FC Fastwrite: OFF
Interrupts:      22      Link_failure: 0      Frjt:      0
Unknown:      2      Loss_of_sync: 0      Fbsy:      0
Lli:      22      Loss_of_sig: 2
Proc_rqrd:      33      Protocol_err: 0
Timed_out:      0      Invalid_word: 0
Tx_unavail:      0      Invalid_crc: 0
Delim_err:      0      Address_err: 0
Lr_in:      2      Ols_in: 0
Lr_out:      0      Ols_out: 2

```

7.6 Extension Monitoring

Monitoring of Extension tunnels and circuits can be done in various ways: IBM SANnav SAN Management tool, CLI, Rest API automation, and SNMP.

7.6.1 IBM SANnav SAN Management Tool

IBM SANnav SAN Management Tool is a comprehensive tool for configuring, troubleshooting, monitoring, and performing various operational tasks. For more information about the use of IBM SANnav, see Chapter 11, “IBM SANnav Management Suite overview” on page 261.

Various dashboards can be constructed to fit your environment’s needs and. Top port traffic, top port utilization, top buffer-credit at zero, REPLICATION fabric health, configuration drifts, and other top fabric events are monitored, as shown in Figure 7-38. The widgets feature a pull-down menu, and various other statistics and errors can be monitored. More rows and widgets also can be added.

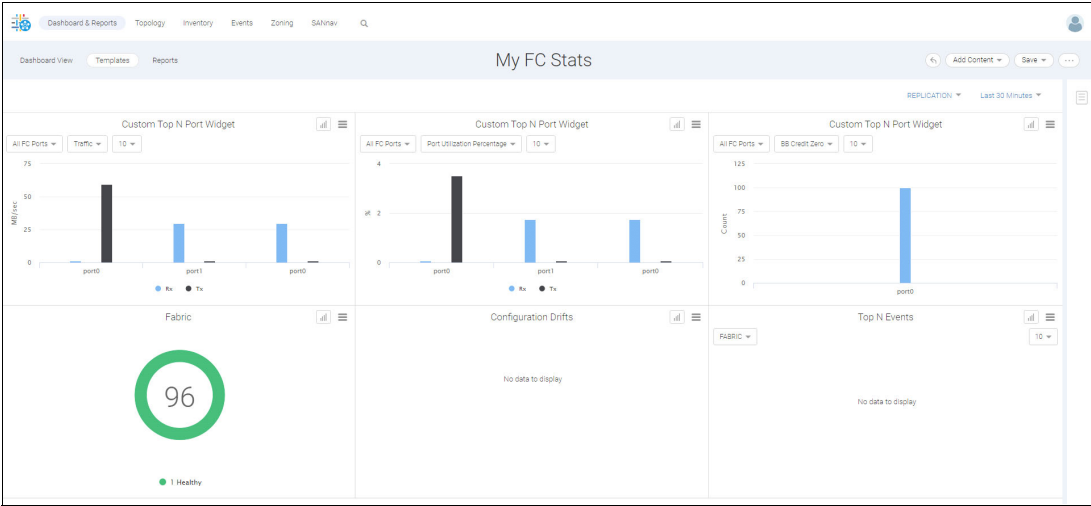


Figure 7-38 SANnav FC replication dashboard

SANnav Health Summary is shown in Figure 7-39. The overall status of the fabrics, switches, hosts, and storage is indicated. In this case, one replication fabric is used; typically, two or more replication fabrics are used (for example, A and B). If SANnav also monitors the production SAN, those fabrics and devices are included.

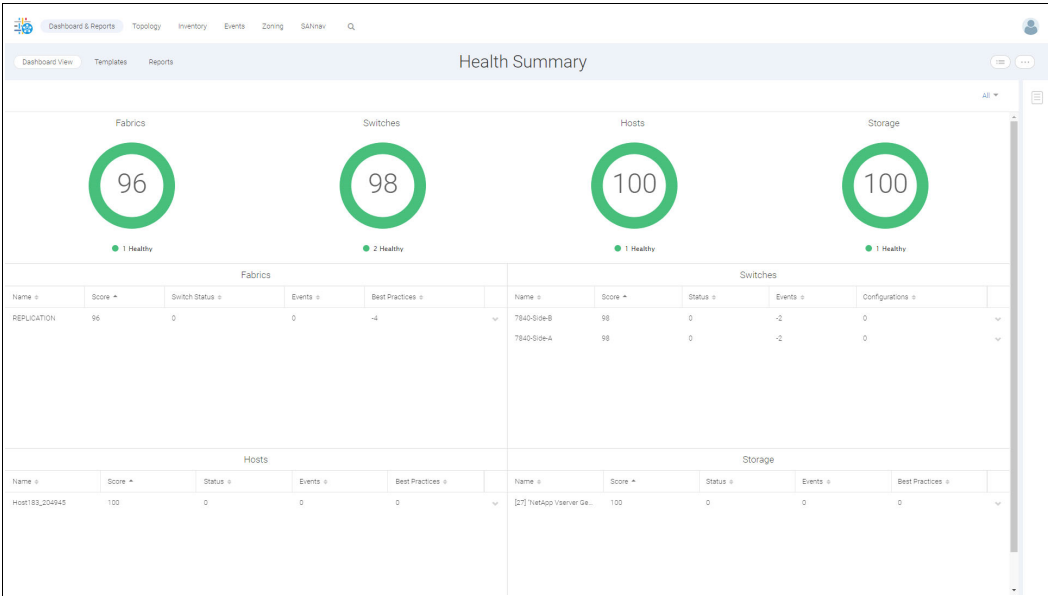


Figure 7-39 SANnav health summary

In Figure 7-40, a simple two-site replication SAN is shown. In many practical instances, the replication SAN is autonomous from the production SAN. The topology view can narrow down to a specific fabric for simplicity and investigation. The FOS fabricname was configured as “REPLICATION.” In SANnav, this topology also was saved as REPLICATION. A single tunnel uses VE24.

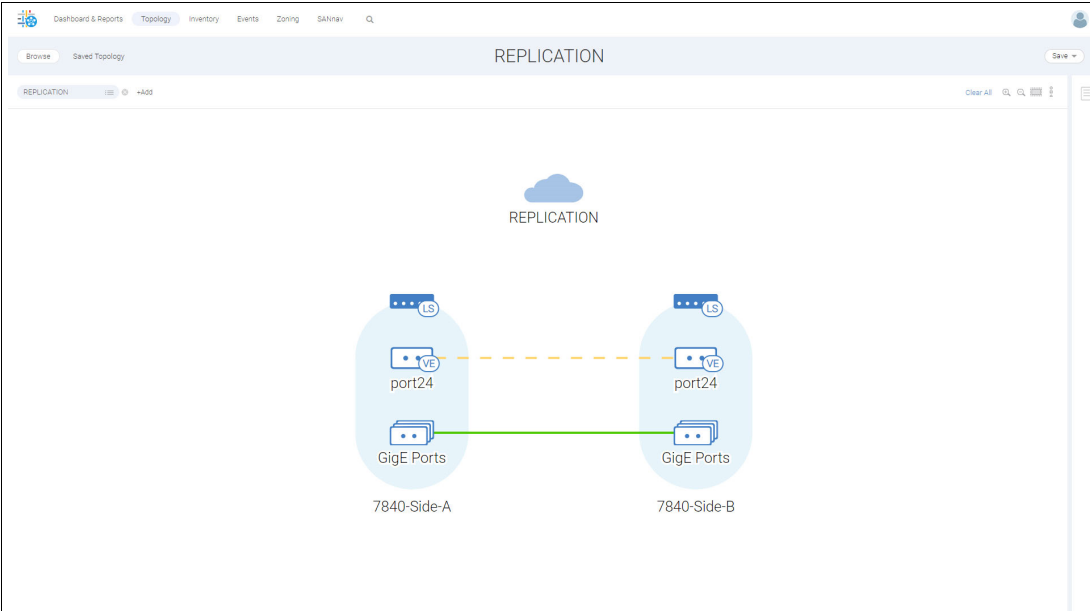


Figure 7-40 SANnav Topology of Replication Environment

Clicking the GE ports displays information that is specific to the interfaces, as shown in Figure 7-41.

GigE Ports for 7840-Side-A						
Name	Slot/P...	Speed	Status	Rx Util...	Tx Util...	
ge2	0/ge2	10 Gb/s	Online	0.48	0.02	▼
ge4	0/ge4	10 Gb/s	Online	0.49	0.02	▼
Close						

Figure 7-41 SANnav extension GE interfaces in topology

In Figure 7-42, IBM SANnav is monitoring the traffic going through physical GE interfaces. The GE interfaces show what quantity of traffic is sent to the directly connected data center switches. It does not specifically show tunnel or circuit traffic statistics. The GE interfaces are abstracted portals through which circuits pass; zero to many circuits might be included per GE interface.

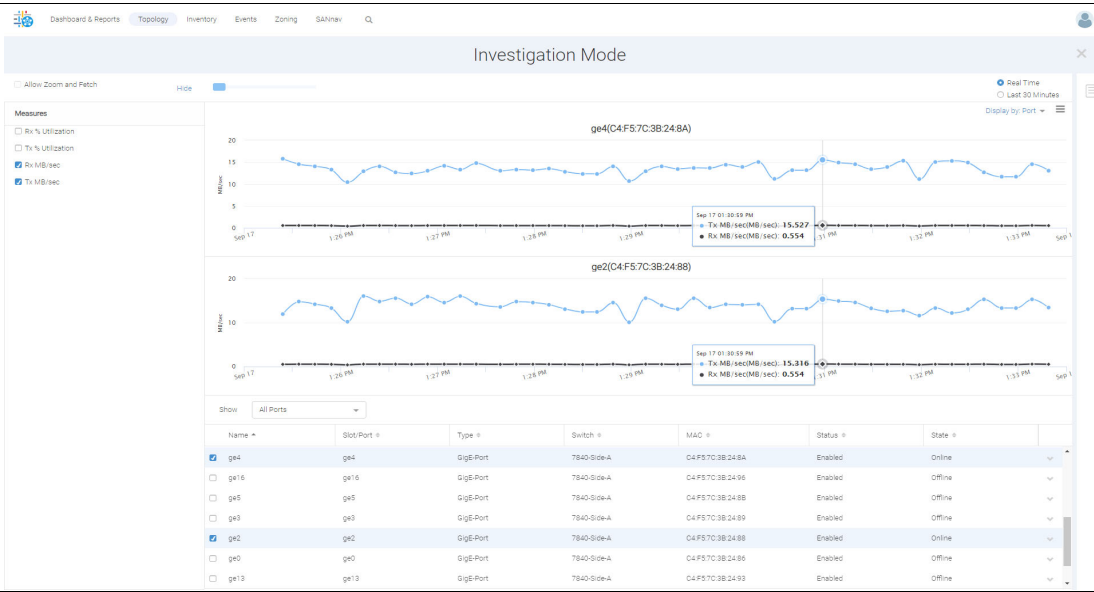


Figure 7-42 SANnav extension GE interface monitoring

Figure 7-43 shows a customizable Dashboard, although a pre-constructed Extension Dashboard is available. MyDashboard shows tunnel and circuit Rx and Tx activity, dropped packets, RTT, jitter, and out-of-order packets. Other available network characteristics are not shown.

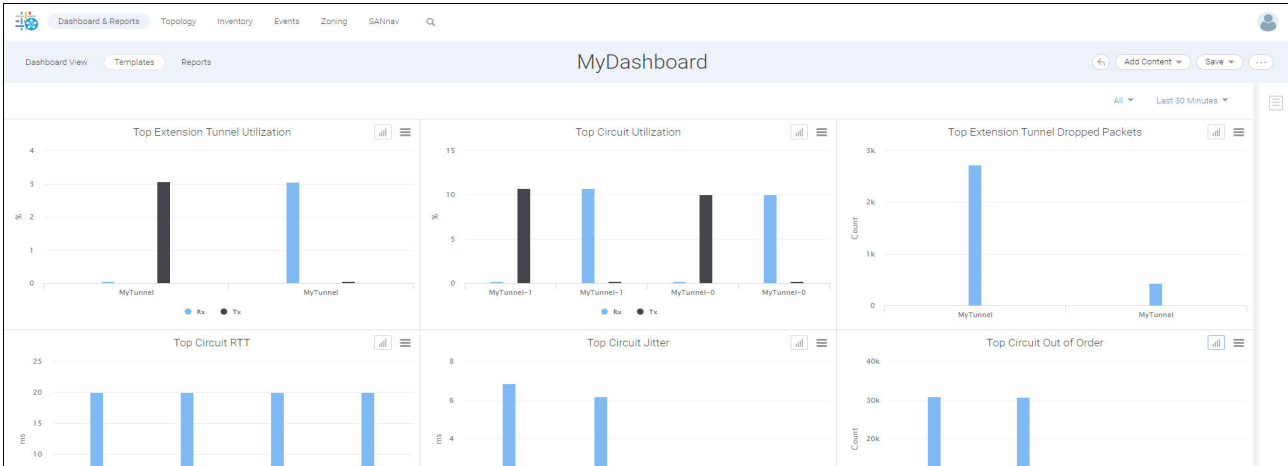


Figure 7-43 SANnav extension dashboard

IBM SANnav is monitoring the Extension Tunnel with selected characteristics, as shown in Figure 7-44. Compression for this specific data is shown to be approximately 2.3:1. Latency is being shown at 20 ms RTT. The Extension platform is detecting out-of-order segments, which is indicative of packet loss on the network. Transmit utilization and MBps are monitored. Traffic is being sent from 7840-Side-A (switch1) to 7840-Side-B (switch2). The Receive traffic consists of acknowledgments and was chosen not to be monitored here.

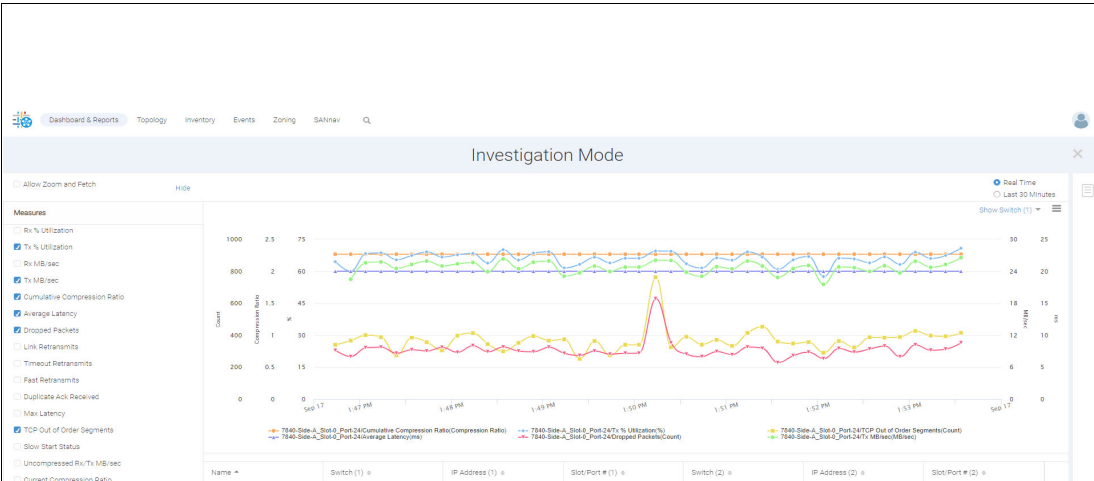


Figure 7-44 SANnav investigative mode for an extension tunnel

IBM SANnav is monitoring the Extension Circuits with selected characteristics, as shown in Figure 7-45. Transmission in MBps is shown. Rx MBps was not selected because it shows little traffic that consists of acknowledgments; data is replicated from switch1 to switch2. Four measurements can be enabled at once, while two circuits are monitored. Mousing over the charts displays the legend.

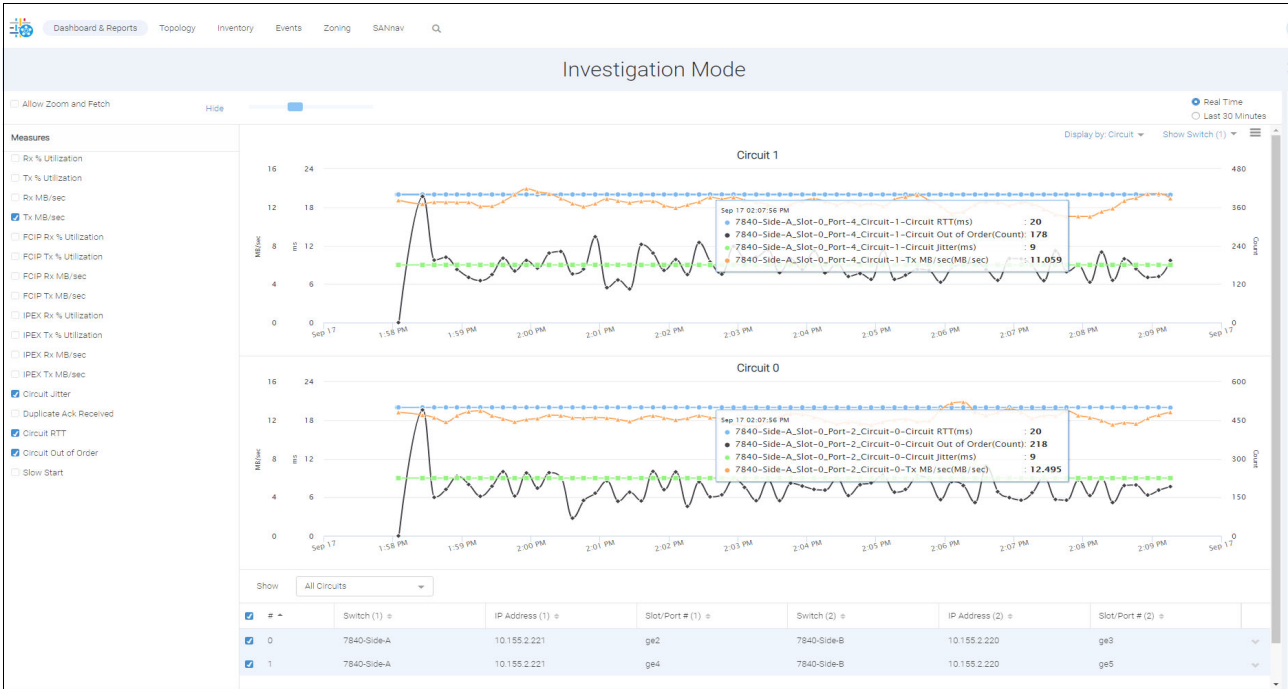


Figure 7-45 SANnav investigative mode for extension circuits

A customized Extension report (see Figure 7-46) can be constructed from the most interesting widgets for your environment.

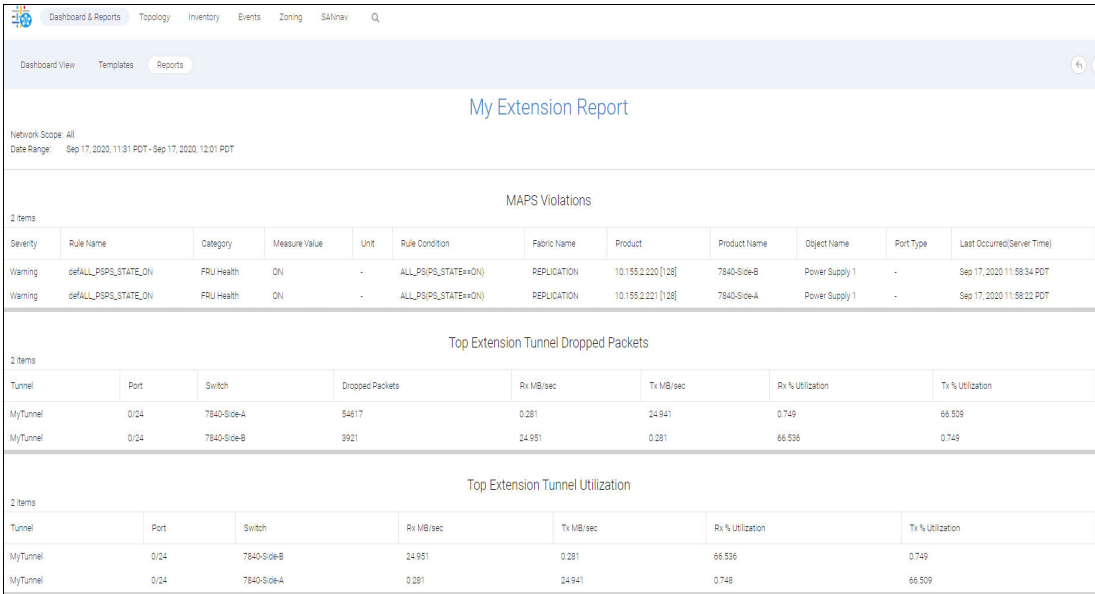


Figure 7-46 SANnav extension report

The Monitoring Alerting Policy Suite (MAPS) monitors every aspect of the Extension platform, and it is prudent to be aware of MAPS violations. Over long operational periods, a WAN often experiences interims of low quality. By monitoring packet drop rates, RTT stability, increases in jitter, and reduced tunnel use, operations can drive inquiries about observed changes with the network administrators and potentially preempt further degradation or an entire outage.

Your customized Extension Report with selected statistics can be generated at regular intervals, as shown in Figure 7-47. For example, you can configure the report to be generated daily at midnight so that it is waiting for you in the morning. The formats that can be delivered are HTML, PDF, and CSV.

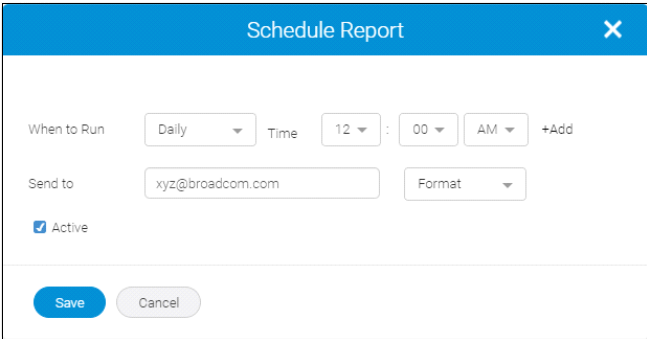


Figure 7-47 SANnav schedule report

7.6.2 Fabric operating system CLI

Many CLI commands show statistical output that is based on traffic, interfaces, circuits, tunnels, TCP sessions. For more information about the following most commonly used commands with their output for monitoring Extension, see Appendix B, “Brocade Fabric OS CLI monitoring commands” on page 331:

- ▶ **portshow** (tunnels and circuits)
- ▶ **portshow** (LAN stats)
- ▶ **portshow** (TCP stats)
- ▶ **portstatsshow**
- ▶ **geportperfshow**
- ▶ **switchshow**
- ▶ **portperfshow**
- ▶ **porterrshow**

7.6.3 REST API

IBM Extension platforms include a fully capable REST API in which configuration and monitoring can be performed. For more information, see Chapter 10, “IBM b-type Gen 7 SAN automation” on page 251.

Typically, Ansible Playbooks are used in environments that use this type of automation.

7.6.4 SNMP

SNMPv1 and SNMPv3 can be configured to send traps or have statistics read by other management tools. The output of `snmpconfig` for `snmpv1` traps is shown in the following example:

```
BCM-7840-A:FID128:admin> snmpconfig --show snmpv1
SNMPv1 community and trap recipient configuration:
Community 1: Secret C0de (rw)
    No trap recipient configured yet
Community 2: OrigEquipMfr (rw)
    No trap recipient configured yet
Community 3: private (rw)
    No trap recipient configured yet
Community 4: mySecret1 (ro)
    Trap recipient: 10.10.10.10
    Trap port: 162
    Trap recipient Severity level: 4
Community 5: mySecret2 (ro)
    Trap recipient: 192.168.10.10
    Trap port: 162
    Trap recipient Severity level: 4
Community 6: mySecret3 (ro)
    Trap recipient: 172.16.10.10
    Trap port: 162
    Trap recipient Severity level: 4
SNMPv1:Enabled
```

Below is a list of the trap MIBs available.

7840-Side-A:FID128:admin> snmptraps --show

#	Mib Name	Supported Traps
----- -----		
001	SW-MIB	sw-fc-port-scn sw-ip-v6-change-trap sw-pmgr-event-trap sw-event-trap sw-fabric-reconfig-trap sw-fabric-segment-trap sw-state-change-trap sw-zone-config-change-trap sw-port-move-trap sw-brcd-generic-trap sw-device-status-trap
002	FICON-MIB	link-rnid-device-registration link-rnid-device-deregistration link-lirr-listener-added link-lirr-listener-removed link-rlir-failure-incident
003	FA-MIB	conn-unit-status-change conn-unit-port-status-change conn-unit-event-trap
004	MIB-2	cold-restart-trap warm-restart-trap
005	IF-MIB	if-link-up-trap if-link-down-trap
006	RFC1157	snmp-authetication-trap
007	HA-MIB	fru-status-change-trap fru-history-trap cp-status-change-trap
008	MAPS-MIB	maps-trap maps-quiet-time-trap
009	T11-FC-ZONE-SERVER-MIB	t11ZsRequestRejectNotify t11ZsMergeSuccessNotify t11ZsMergeFailureNotify t11ZsDefZoneChangeNotify t11ZsActivateNotify

For more information, see [Brocade Fabric OS MIB Reference Manual, 9.0.x](#).



NVMe over Fibre Channel SAN design, implementation, and migration strategies

The combination of Gen 7 and Fabric OS (FOS) 9.0 bring a new level of support for NVM Express (NVMe) in the product set. This support includes the capability to carry the traffic and the specific visibility to monitor the traffic. Although previous generations carried the traffic, they had little or no ability to monitor it.

This chapter includes the following topics:

- ▶ 8.1, “NVMe overview” on page 218
- ▶ 8.2, “SAN design considerations for NVMe over Fibre Channel” on page 225
- ▶ 8.3, “Implementation overview of NVMe over Fibre Channel” on page 226
- ▶ 8.4, “Migration example of NVMe over Fibre Channel” on page 227
- ▶ 8.5, “Summary” on page 231

8.1 NVMe overview

At base, it is important to understand that NVMe is an industry standard that deals with how software accesses non-volatile memory devices by using a PCI Express (PCIe) bus interface to the device.

The organization driving the industry standard is the [NVM Express consortium](#) and is responsible for the definition, management, and promotion of the protocol in the industry. The organization is also responsible for the specification concerning Fabrics (NVMe-oF), which addresses the use of NVMe commands over networked fabrics, such as Fibre Channel (FC), InfiniBand, or Ethernet (this specification can be found on the NVM Express website as well).

The reason that it is important to understand that NVMe is a protocol that was developed for accessing memory devices is that most implementations for storage to date dealt with protocols (for example, SCSI, SAS, and SATA) that expected rotational media or even non-storage related devices, such as printers or scanners.

As such, NVMe is a much more streamlined and less “chatty” protocol than Small Computer System Interconnect (SCSI), which was written more than 30 years ago and accommodated such a broad range of devices. In fact, one of the earliest significant benefits of running NVMe-oF is the improved performance on the server stack.

As we discuss later, protocol comparisons show a 30 - 60% reduction in CPU overhead for the I/O. There is much excitement for this change because application owners are expecting performance gains, even on existing hardware.

For more information about the software stack, see Performance Benchmarking for PCIe and NVMe Enterprise Solid-State Drives, at [this web page](#).

The CPU cycles for IOPS that uses ACHI/SCSI is shown in Figure 8-1.

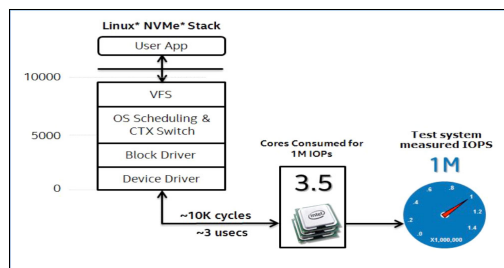


Figure 8-1 CPU cycles for IOPS using ACHI/SCSI

The CPU cycles for IOPS that uses NVMe is shown in Figure 8-2.

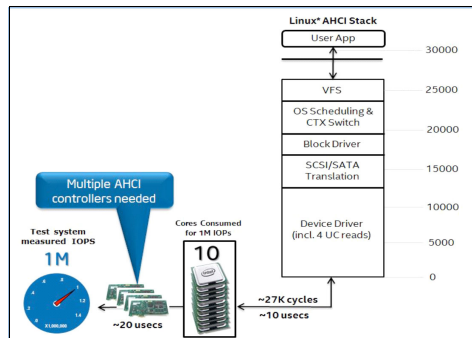


Figure 8-2 CPU cycles for IOPS using NVMe

In addition to the fact that the NVMe software stack is designed for memory devices and therefore avoids all of the older software that is necessary for dealing with rotational media, disk drive geometries, and so on, the definition includes a level of parallelism that is different than the traditional SCSI, Serial Attach SCSI (SAS), or Serial ATA (SATA) environments for storage.

Those older protocols traditionally feature a single command queue with a more limited queue depth (in the case of SCSI, a maximum of 64 commands). In the NVMe standard, support exists for 64 K (65,535) I/O queues, each with 64 K (65,535) commands, which is commonly referred to a *queue depth* or the number of outstanding commands.

It is important to remember that the use of this parallelism continues to be developed by operating systems and applications. A view of how the parallelism might be viewed is shown in Figure 8-3.

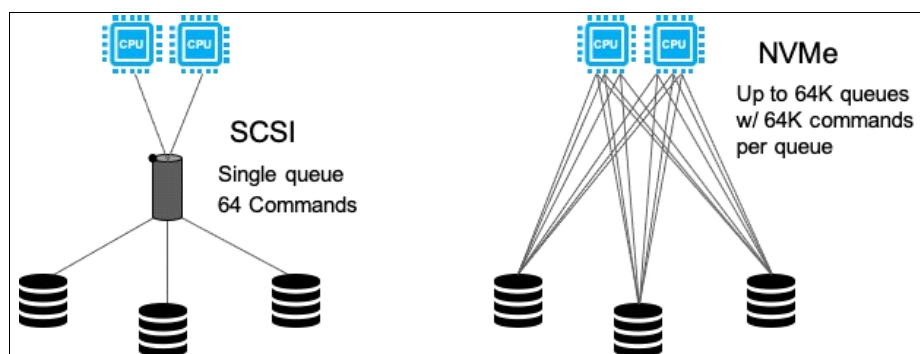


Figure 8-3 SCSI and NVMe queues

NVMe queues are mapped to CPU cores, which allows a scalable performance model. Over time, this parallelism can enable changes where virtualization platforms can assign an NVMe queue to each virtual machine.

When NVMe was first brought to market, it was almost exclusively implemented as an embedded device directly on the PCIe bus internal to the server. It was intended as an expansion of system memory that traded off the higher latency that NVMe features compared to DRAM for a lower cost versus DRAM.

At a device level, common NVMe products use 4 x Gen 3 PCIe lanes (4 x 8 Gbps = 32 Gb) at a device latency of sub 20 microseconds. Eventually, the market determined that the following characteristics made the need for NVMe over Fabric a wanted goal:

- ▶ The ability to scale memory capacity past that which can be achieved on a single card.
- ▶ The need to not to burden every server with the cost of an NVMe device.
- ▶ Create a shared resource pool of NVMe that can be accessed by multiple servers.
- ▶ The need for platform reliability where single-server failure or restart does not affect the application.

These requirements fed a migration, as shown in Figure 8-4.

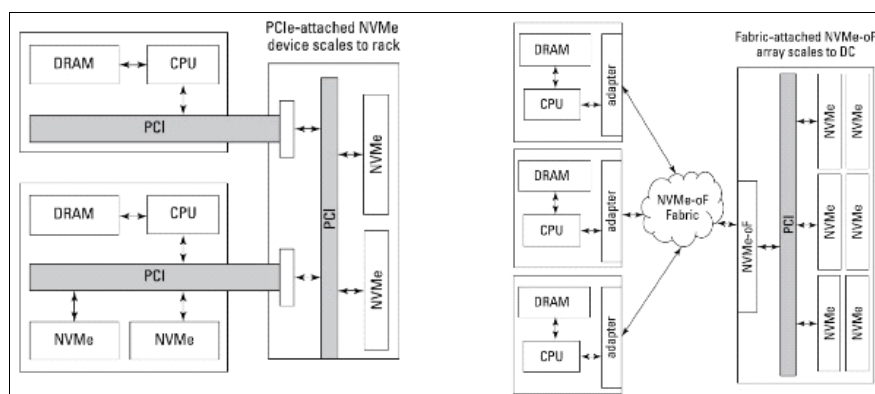


Figure 8-4 Migration to NVMe over Fabric

The NVMe Consortium is also responsible for the efforts around NVMe over Fabrics. A part of that specification is the NVMe Transport binding specification. The NVMe Transport binding specification for FC is defined in *INCITS 540 Fibre Channel – Non-Volatile Memory Express (FC-NVMe)*, which is available at [this website](#).

NVMe is a protocol that is designed to speak to memory devices over a PCI interface. As a result, it has a much lighter weight stack than SCSI. In NVMe over FC, the FC acts as the transport protocol for the NVMe protocol. NVMe frames are encapsulated in FC and transmitted over the SAN. Table 8-1 lists the mappings between the NVMe protocol and the FC protocol.

8.1.1 Comparing mapping of NVMe versus FCP protocols

Table 8-1 Mappings between the NVMe and FCP protocols

NVMe entity	FCP mapping
Namespace/NSID	LUN
I/Os	Exchanges
Command	Command routing-control (RCTL) = 6
Status	Status RCTL = 7
Extended/ intermediate status	E_Status RCTL = 7
NVMe subsystem reset	Unsupported
No Recovery, only retries	Sequence retransmission request (SRR) / read exchange concise (REC)
Association creation	Extended link service (ELS)
I/O Q creation	ELS
Controller	Target
Namespace	LUN
Rd	Read
Wr	Write

When a networked environment is designed for NVMe, it is also important to remember that the technology is still expanding in terms of function and support. Many of these capabilities will become available over time in the market and some will be made available through software updates to the existing releases. Some of these capabilities were released, such as Asymmetric Namespace Access (ANA) for multipathing (the equivalent of the ALUA in SCSI) and Sequence Level Error Recovery (SLER), while others take some time to be implemented.

At the time of this writing, NVMe includes 13 required commands and 25 optional. Although it is to be expected that augmentation of functions over time increases the number of commands, it provides an exceedingly streamlined interface as compared to the SCSI software stack. For over 30 years it accommodated storage devices and any other peripherals (for example, printers and scanners).

Figure 8-5 shows a simplified view of the comparison between the SCSI software stack and the NVMe software stack. In addition to the fact that NVMe is written to access memory devices instead of traditional SCSI disk drive elements, the software layer was streamlined to address the block layer, which resulted in a significant performance increase.

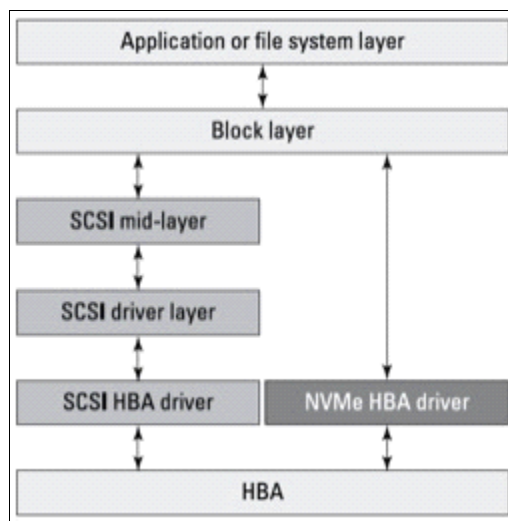


Figure 8-5 SCSI software stack versus NVMe software stack

It is in part this simplified software stack that produces the reduction in I/O overhead that is observed in the Intel white paper. It is expected that over time the NVMe interface becomes the preferred software interface for enterprise storage. Although this process does not occur overnight because of the millions of lines of code and product implementations for SCSI that are in use, the benefits of the NVMe software performance \ eventually become dominant.

This realization is especially true for the All Flash Array environments in the enterprise as organizations seek to get a full use of the performance of the assets. This consideration is also one aspect in choosing the type of fabric to use for the storage.

It is easy in the planning stages to assume that “all bandwidth is the same”. That is, 100 Gb is 100 Gb in performance, regardless of the software stack or protocol that is used. However, this view is flawed.

The overhead and services that are encompassed in every protocol always matter. It is in fact that point that caused the decision to use User Datagram Protocol (UDP) in the Remote Direct Memory over Converged Ethernet version 2 (RoCE v2) standard for NVMe.

The choice was made because the latency of the Transmission Control Protocol (TCP) was considered to be too high to allow maximum NVMe performance.

FC uses a Zero Copy mechanism; that is, when an application requests data from an FC storage platform, it provides a logical address range in the application buffer to which the data to be written. Effectively, the data that is returned from the request is written directly into the application.

FC provided this function and its associated performance gain for more than 20 years. As the storage landscape moves to use more flash storage, the need to achieve a maximum Fan-in or Fan-out ratio (these terms are used to describe multiple server connections sharing access to a single port on a storage array) for those environments has made FC an optimal choice.

Part of the discussion here is the shift in the technology. Although solid-state drives (SSD) were packaged for years to be the same form factor, same connectors, and same software protocol for access as hard-disk drives (HDD), no disk drive is involved. SSD is composed most frequently of Flash memory. It is persistent or non-volatile memory, yet memory all the same.

The term *non-volatile* is used to distinguish this type of memory from the Double Data Rate Synchronous Dynamic-Access Memory (DDR SDRAM) that is used for server memory. Although SDRAM retains only data while power is applied to the device, Flash memory is persistent in that it requires power only to change state, not to maintain state.

Almost all USB and SSD devices today make use of NAND style memory chips that are non-volatile memory. As the name suggests, the same applies to drives that are composed of NVMe devices.

What makes this issue interesting is that Moore's Law (or equivalent variations) now applies to enterprise storage for the first time in the history of enterprise storage. Density and latency of 3D NAND is continuing to improve with several 3D NAND suppliers moving to 176-layer technology. This fact drives most considerations about the types of SAN fabrics that are suitable to carry such performance and to accommodate multiple iterations of new technologies within a normal (for example, 5-year) depreciation schedule of capital assets.

In the current data center infrastructure, the need for NVMe and SCSI to coexist for a significant period is readily apparent.

8.1.2 NVMe over Fibre Channel

It is important to remember that FC, although long associated predominately with SCSI, always was a protocol that was meant to carry multiple upper layer protocols, as shown in Figure 8-6.

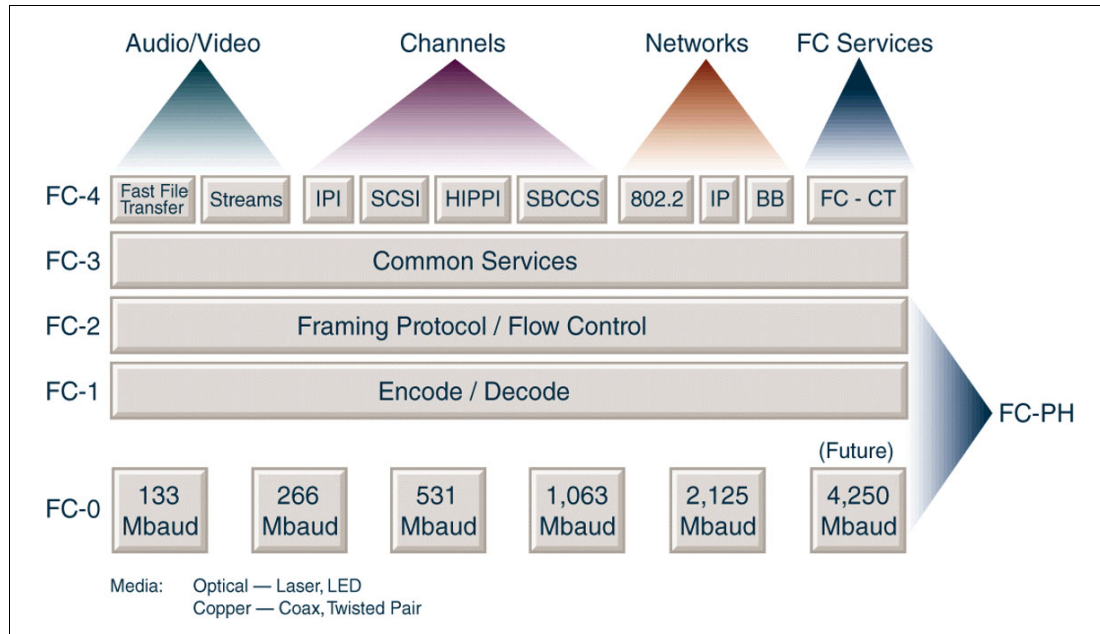


Figure 8-6 Fibre Channel protocol

However, the predominance of uses cases for FC today falls into the scenario that is shown in Figure 8-7.

FC - 4	Upper Layer Protocol SCSI / FICON / NVMe
FC - 3	General Services
FC - 2	Framing Protocol and Flow Control
FC - 1	Encode/Decode 64b/66b
FC - 0	Transmission Rate 4Gb/8Gb/10Gb/16Gb/32Gb/64Gb

Figure 8-7 Fibre Channel protocol stack

As a market dynamic, when a new FC transmission rate is enabled, the oldest speed is decremented from the product sets. This normally meant that in Gen 7, the 4 Gb transmission rate no longer is supported.

However, in the instance of the Gen 7 platforms, the IBM b-type Gen 7 SAN retains the 4 Gb transmission rate in the Brocade FC32-X7-48 blade. The blade can be incorporated into the IBM Storage Networking SAN512B-7 or the IBM Storage Networking SAN256B-7 chassis-based platforms.

NVMe is the first new Upper Layer Protocol (ULP) mapping in several years. What makes this interesting is that the infrastructure that IBM sold for years can “carry” NVMe traffic. Transmission Rate is a concern only in regard to performance. All of the current shipping transceivers/optics can carry the NVMe traffic if they are in a compatible FC product.

Consider the following points:

- ▶ **NVMe capable adapter:** In the case of FC, this adapter is a Host Bus Adapter (HBA). Any Gen 6 or newer HBA with the correct driver can carry NVMe traffic. Because operating system support varies for drivers, it is important to verify supported versions with your operating system vendor.
- ▶ **NVMe capable fabric:** In the case of FC, any IBM b-type SAN platform that is Gen 6 or Gen 7 can carry NVMe traffic. The required operating system level Gen 6 is FOS v8.1 or higher; for Gen 7, it is FOS v9.0 or higher.

Note: The ability to carry the NVMe traffic is not the same as visibility of the NVMe traffic. Gen 5 has no visibility to NVMe traffic in I/O Insight. In the Gen 6 platforms only the SAN128B-6 and the 64-port blade in the SAN512B-6 or SAN256B-6 have I/O Insight visibility to NVMe traffic. However, ALL Gen 7 platforms have I/O Insight visibility to NVMe traffic.

- ▶ **NVMe capable array:** In the case of FC in the IBM storage product line, any members of the IBM FlashSystem family of IBM FlashSystem 5000, IBM FlashSystem 7200, or the IBM FlashSystem 9200 support NVMe over FC fabric.

It is historically true that over time as various elements in the IT infrastructure go through generational development (for example, 16 Gb FC to 32 Gb FC), bottlenecks move around in the environment. One of the common scenarios in design consideration is balance.

Extremely fast server platforms that are communicating with slow storage platforms potentially experience congestion. Similarly, the mix of older platforms can be problematic, especially in the interaction in the network. Why is this a discussion point regarding NVMe? Because the latency of NVMe is such that the “inter I/O-Gap” is shrinking. This issue can be thought of as the “white space” between I/Os on the cable, as shown in Figure 8-8.

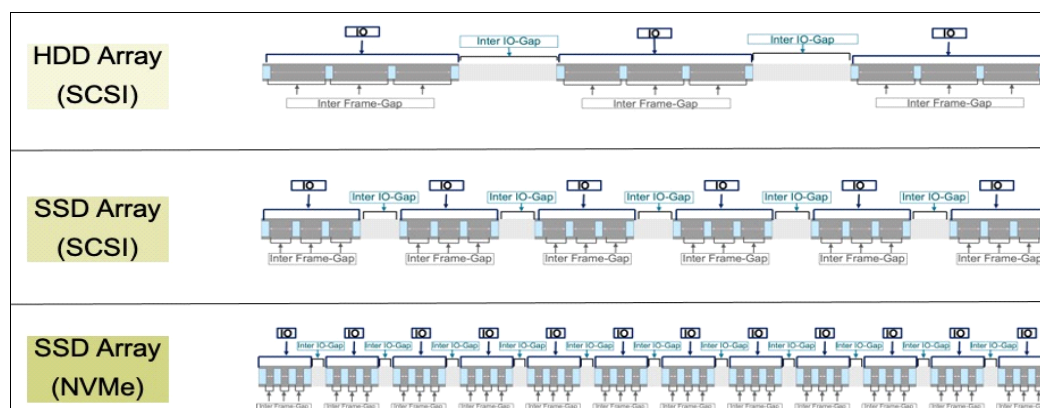


Figure 8-8 Inter I/O-gap reduction as technology improves

A frequent misunderstanding in SAN design is to consider only the bandwidth requirements of a specific device or devices. The reason that this issue causes challenges is that it is easy to forget that although bandwidth, I/Os per second (IOPs), and latency are related characteristics, they are not the same characteristic.

Think of bandwidth as the number of lanes on a highway. IOPs is how many vehicles are going past a point on the highway per second and latency is how long it takes to get from home to work. Performance applications are becoming more latency and IOP driven.

The environment might not drive a full traffic load on the highway, but it might frequently peak for brief periods (for example, during database analytics) or an application set that routinely requires low latency; for example, payment card data.

It is always important to have a grasp of the range of application workloads that are going to be supported by any SAN design. It is an understanding of the types of workloads to be supported that provides the “fit for purpose” decision data about which network technology meets those needs.

The IBM FlashSystem platforms can achieve a less than 100-microsecond latency off the front end of the array controller by using FC-NVMe as the fabric. The fact that the SCSI driver stack is frequently in the millisecond to seconds range for latency is what is driving much of the interest in NVMe over FC.

The NVMe over fabric standards continue to develop. One of the most recent updates includes the function of Sequence Level Error Recovery (SLER). The SLER mechanism allows a much faster recovery of errors. Instead of waiting for an application timeout the SLER implementation uses a new set of commands (FLUSH and Responder Error Detected [RED]) between the initiator and target ports to recover from any sequence level errors. FC uses Frames.

Multiple Frames make up a Sequence and multiple Sequences make up an Exchange. An easy way to think of it is word, sentence, and paragraph. Rather than wait for the Exchange to be retried, recovery is possible at the Sequence level by way of these new signaling commands.

For more information about the current FC-NVMe-2 standard, see [this web page](#).

8.2 SAN design considerations for NVMe over Fibre Channel

For most customers, the implementation of NVMe over FC is not an “all or nothing” event. In fact, part of the SAN design must consider that a mix of involved technologies is used.

The use of SCSI-based infrastructure is not removed overnight and for most implementations. Therefore, the need exists to allow for concurrent operations of the SCSI over FC and NVMe over FC traffic without being disruptive to either traffic or hampering the wanted performance of the NVMe systems.

Fortunately, the two protocols can run concurrently in the same fabric. The most common scenario encounters some portion of the environment being equipment that can participate in an NVMe over FC deployment, such as the following examples:

- ▶ Gen 6 HBAs (32 Gb) with a supported NVMe driver stack
- ▶ Gen 6 or Gen 7 IBM b-type Gen 7 SAN platforms with a minimum of FOS v8.1
- ▶ An IBM Storage array that supports NVMe over Fabric

Because the HBAs and SAN switches and directors were shipping for more than six years, it is likely that some portion of most SAN infrastructure is in place to support NVMe over FC.

Further, the most common SAN topology of Core-Edge is suited to the “in-place” integration of NVMe technologies into those SANs. Other server elements or storage elements that do support NVMe over FC can readily be added into such a topology in a nondisruptive manner. If the core or the edge SAN switches or directors must be refreshed, that process is readily accomplished with little to no risk of downtime because of the dual redundant hardware infrastructure design of SAN implementations.

Because of the intermix of the technologies, the ability of the Gen 7 IBM b-type Gen 7 SAN platforms to provide functions, such as Traffic Optimizer and I/O Insight, is of critical help to the SAN administrator. As we discuss later in this chapter, the migration of applications or virtual machines from SCSI-based storage to NVMe-based storage proved to be a risk averse technology migration.

As you will see, the combination of the suitable topology, along with balanced provisioning, granular monitoring, and automated mitigation provide today's SAN administrator a straightforward window into implementing NVMe over FC fabrics.

For more information about this type of implementation, see [Brocade SAN Design and Best Practices Design Guide](#).

8.3 Implementation overview of NVMe over Fibre Channel

One of the interesting aspects of NVMe over FC is how much of the environment is the same as the existing SAN environment for SCSI or FICON. The enterprise services that are expected of an FC SAN are readily available. However, so too are some of the newer NVMe functions so that the NVMe services become a superset of the normal SAN services, including the following examples:

- ▶ Centralized discovery
- ▶ Sharing of discovery data across the fabric
- ▶ Zoning and Security
- ▶ Registered State Change Notification
- ▶ Sequence Level Error Recovery

The normal functions of Zone creation, inclusive of Peer Zoning or Target Driven Zoning, are the same for NVMe devices as it is for SCSI. For more information about the complete environment, see [Brocade Fabric OS Administration Guide, 9.0.x](#).

Also, all of the Fabric Vision functions of I/O Insight, Fabric Performance Impact Monitoring, and so on, are supported.

However, some things do change. In an NVMe subsystem, the naming conventions are different. In the SCSI world, a Logical Unit Number (LUN) is how individual volumes that are behind the SCSI target device are configured. NVMe uses the terminology of a namespace ID for that same function. A namespace is composed of a set of Logical Block Addresses (LBA) and not a set of names.

It is true that because FC is a standard link level network transport that supports multiple protocols, it needs only the new upper layer protocols (ULPs) to support NVMe over FC. Therefore, although no changes are necessary to the FC switching logic, another element to understand is that while the data is carried as standard FC Protocol (FCP) traffic, the NVMe commands are carried as a new frame type FC-NVMe.

Therefore, when an NVMe device logs in on the fabric and the generic name server entry is created, it also updates the other information, which includes but is not limited to, the following examples:

- ▶ FC4 type: NVMe device (0x28)
- ▶ Device type: None, host, target, or both host and target
- ▶ Discovery services: Absent or present (specific to NVMe devices only)
- ▶ Port symbolic name
- ▶ Node symbolic name

Figure 8-9 shows the namespace and namespace IDs of an NVMe subsystem. The logical equivalent of a SCSI target is the NVMe controller in this figure.

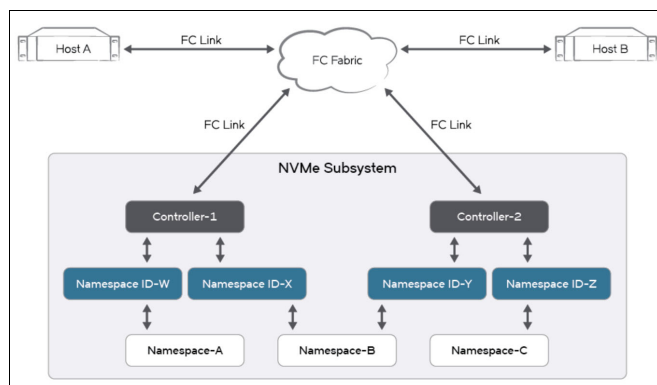


Figure 8-9 NVMe subsystem

The logical equivalent of the SCSI LUN (address range) is the namespace ID that is behind the controller. The physical namespace or range of Logical Block Addresses is then presented by way of the namespace ID.

An individual namespace can be presented back to the application environment by way of multiple controllers for purposes of performance or design redundancy or application resilience by using Multipath I/O (MPIO) driver support.

As shown in Figure 8-9, the Namespace-B is presented to two separate controllers by way of namespace ID-X and namespace ID-Y. Therefore, although addressing the same range of LBA, the Namespace ID (NSID) is different, depending upon whether it is addressed through Controller-1 or Controller-2.

In Figure 8-9, the MPIO driver can access Namespace-B through either controller.

8.4 Migration example of NVMe over Fibre Channel

It is likely that much of the infrastructure that is running today can support NVMe over FC with a simple software update. To demonstrate this concept, we use an example in which VMWARE vSphere 7.1 for a migration from SCSI to NVMe. (For more information, see this [VMware Dcos web page](#).)

The configuration in our example is Red Hat Linux v8.2 that is running as a guest on top of ESXi 7.1 in the server.

Figure 8-10 shows a topology view of the host and storage.

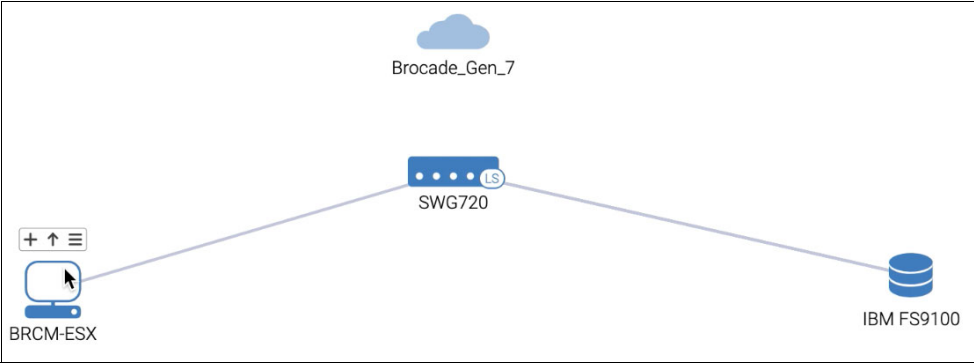


Figure 8-10 Host and storage topology view

By using IBM SANnav Management Portal, you can see the ESXi server platform that is connecting across a SAN64B-7 switch to an IBM FlashSystem 9100.

As shown in Figure 8-11, the FS9100 is expanded to show the port worldwide names (WWN) of each of the multi-ported controllers.

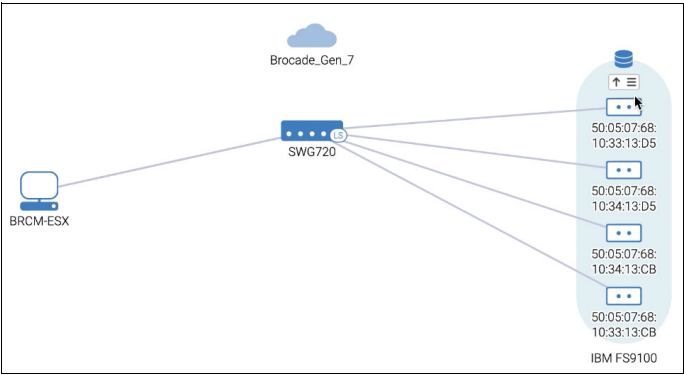


Figure 8-11 FS9100 multi-ported controllers

Now, by selecting and examining the top controller port, we can see the worldwide names of the Virtual Ports that are behind the controller port, as shown in Figure 8-12.

Virtual Ports for 50:05:07:68:10:33:13:D5		
WWN	Zone Alias	FC Address
50:05:07:68:10:37:13:D5	FS9100_6_SCSI_all	010801
50:05:07:68:10:3B:13:D5	FS9100_6_NVME_all	010802
Close		

Figure 8-12 Virtual ports

In this fashion, the Zone alias designations for the SCSI LUN target and NVMe target can be seen. By selecting the ESXi server, we can see more information about the two VMware HBA ports, as shown in Figure 8-13.

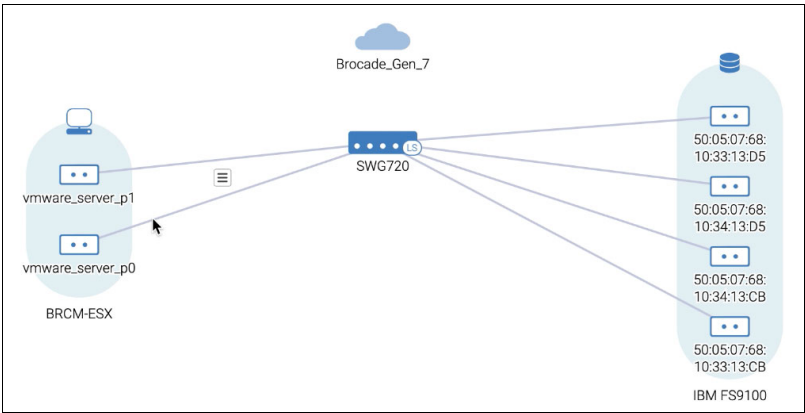


Figure 8-13 BRCM-ESX VMware HBA ports

In Figure 8-14, we can see that the NVMe device is now assigned as a storage element to the ESXi platform. You also see that the transport is FC and the owner is something that VMware refers to as *HPP*.

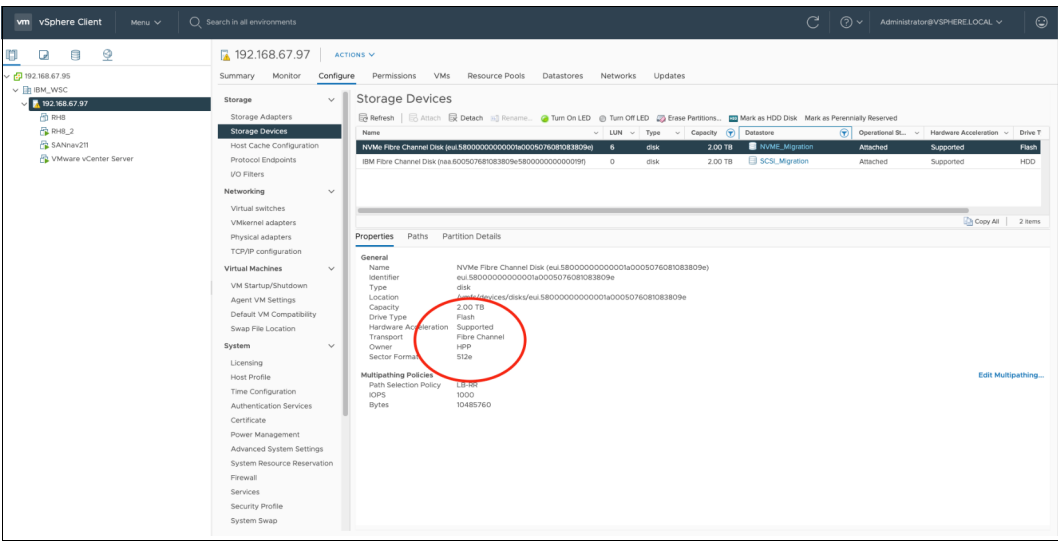


Figure 8-14 ESX HPP ownership

The VMware High Performance Plug-In (HPP) replaces the VMware Native Multipathing Plug-In for high-speed devices, such as NVMe. In fact, in NVMe-oF environments, HPP is the default plug-in that helps to eliminate the higher latency of the SCSI implementations.

Within ESXi, the NVMe-oF targets are emulated and presented to users as SCSI targets for simplicity. For more information about the HPP environments, see this [VMware Docs web page](#).

Notice that the FC SCSI device in this case is still owned by the NMP driver, as shown in Figure 8-15.

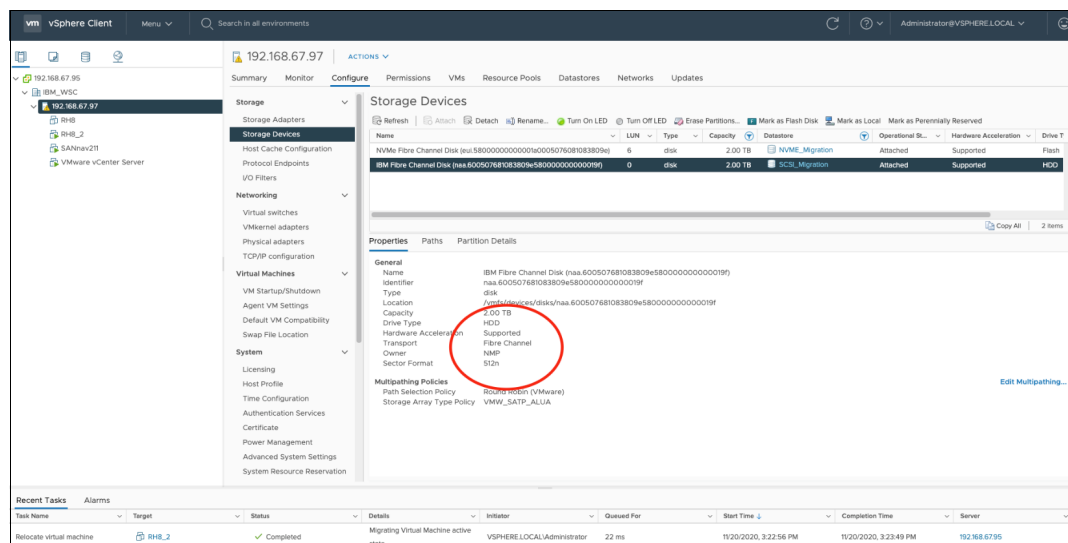


Figure 8-15 ESX NMP ownership

The next step is to select the storage element that is to be mirrored. In this example, the FS9100 SCSI target (identified as RH8) is selected to perform a storage-only migration. First, we must identify the SCSI target that is the source of the migration. The data store source to be migrated is SCSI_Migration.

Next, we select the NVMe device that is the target of the migration where the name of the target data store is NVME_Migration (see Figure 8-16).

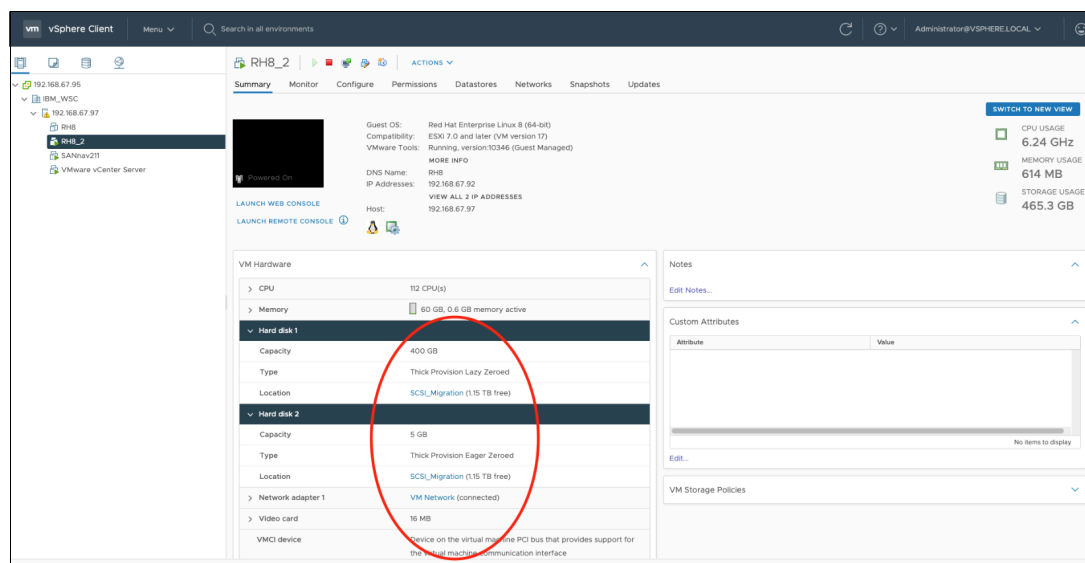


Figure 8-16 Migration source and destination selection

The next step is to start the per VM migration of the data store that is named SCSI_Migration.

After the migration is completed, we can see that the NVMe device is now assigned as a storage element to the ESXi platform, as shown in Figure 8-17 on page 231. Now, both disks (VDisks) are on namespace ID (NSID) backed data stores.

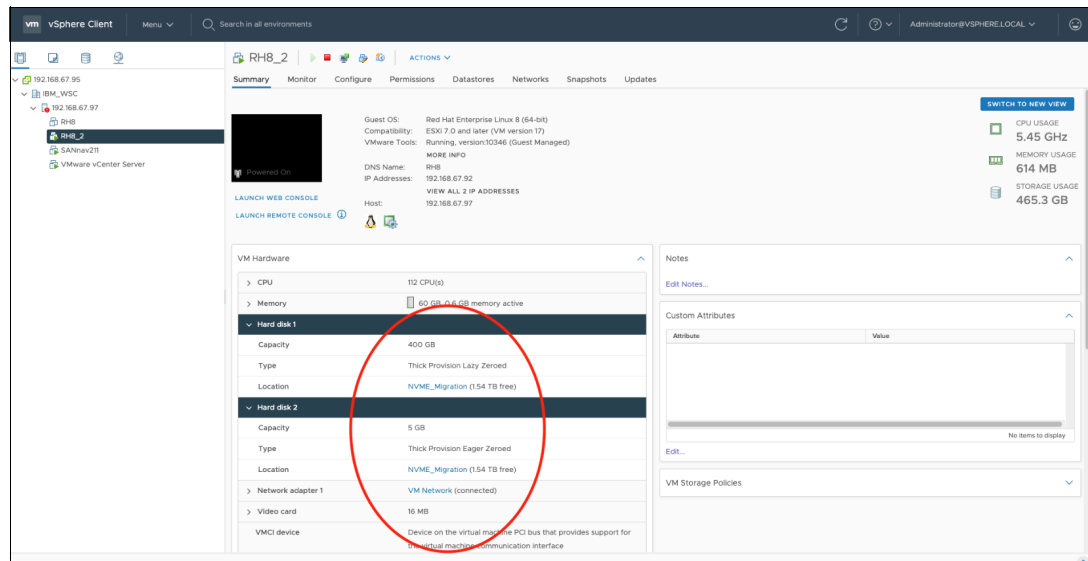


Figure 8-17 NSID-backed data stores

Finally, by selecting the data store source and reviewing the Recent Tasks section that is at the bottom of the window (see Figure 8-18), we see that the VM migration to the new data store completed successfully.

Task Name	Target	Status	Details	Initiator	Queued For	Start Time	Completion Time	Server
Relocate virtual machine	RH8_2	Completed	Migrating Virtual Machine active state	VSPHERE.LOCAL\Administrator	22 ms	11/20/2020, 3:22:56 PM	11/20/2020, 3:23:49 PM	192.168.67.95

Figure 8-18 Recent Tasks

This migration can be accomplished with a live Storage vMotion of the application in question with no downtime. That is, from the same FC HBA across the same IBM b-type Gen 7 SAN switch to the same IBM FS9100 array. This migration is one of the most risk averse technology migrations possible today.

In this migration, the protocol that is used to address the storage is changed. Benchmark comparisons for SCSI versus NVMe with ESXi v7.1 are showing as much as an 80% increase in IOPS, which provides an enormous benefit to the application performance.

8.5 Summary

In this chapter, we discussed the following points:

- ▶ NVMe is more than just a “new drive hardware”. It is a new language for talking to memory-based storage.
- ▶ NVMe is an incredibly efficient protocol in terms of overhead as compared to SCSI.
- ▶ Much of the IBM b-type SAN infrastructure can carry NVMe traffic.
- ▶ The implementation of NVMe is supported in standard SAN topologies, whether we are discussing “core-edge” or “full mesh” designs.
- ▶ The ability to drop NVMe-capable systems into an existing IBM b-type SAN without disruption is key.

- ▶ Migration of VMware workloads from SCSI-based data stores to NVMe based data stores without application downtime makes FC SAN a clear choice for NVMe implementations
- ▶ Implementing NVMe is a straightforward and risk averse operation in an IBM FC SAN.

Because of the benefits of NVMe-based storage platforms, the ability of Gen 7 platforms to readily support NVMe technology makes IBM b-type Gen 7 SAN the logical evolutionary choice for the adoption of the technology. It provides the following benefits:

- ▶ A stable infrastructure that is understood.
- ▶ Lossless, low latency, deterministic network characteristics that are critical for storage.
- ▶ The ability to provide scale and performance simultaneously.
- ▶ I/O Insight auto learning up to 20,000 SID/DID flows per switch or director (including NVMe).
- ▶ The ability to manage NVMe as part of a standard FC environment.
- ▶ At 64 Gb line rate, the IBM SAN64B-7 provides a 460-nanosecond latency for NVMe.
- ▶ The same SAN management (for example, zoning) that is understood by IT.



Brocade Fabric OS administrative tools and diagnostics

This chapter provides a review of the embedded management tools that are available in Brocade Fabric OS (FOS) 9.0.x. The external SAN management tool SANnav is covered in Chapter 11, “IBM SANnav Management Suite overview” on page 261.

This chapter includes the following topics:

- ▶ 9.1, “Command-line interface overview” on page 234
- ▶ 9.2, “WebTools overview” on page 235
- ▶ 9.3, “Fabric Vision overview” on page 237

9.1 Command-line interface overview

The command-line interface (CLI) enables an administrator to monitor and manage individual switches, ports, and entire fabrics from a standard workstation. It is accessed through the Ethernet network through Telnet or SSH connections or physically through the serial console on the switch.

FOS uses RBAC to control access to all FOS operations. Each feature is associated with an RBAC role. You must know which role can run a command, modify the switch, or view the output of the command. Consider these requirements when you create users.

To determine which RBAC role you must run a command, review the “Role-Based Access Control” section in [Brocade Fabric OS Administration Guide, 9.0.x](#).

When you use the CLI, the commands and options are case-sensitive. For more information about the available CLI commands, and their options and uses, see the “Fabric OS command reference” section of [Brocade Fabric OS Administration Guide, 9.0.x](#).

You can connect to the FOS through a SSH or telnet connection or by using a console session on the serial port.

9.1.1 Connecting to Brocade Fabric OS through the serial port

Be aware of the following behaviors for serial connections:

- ▶ Some procedures require that you connect through the serial port; for example, setting the IP address or setting the boot PROM password.
- ▶ Brocade X7 and X6 Director families: You can connect to CP0 or CP1 by using either of the two serial ports.

Complete the following steps to connect to the FOS by using the serial port:

Note: If you connect the serial console port to a terminal server, ensure that flow control is disabled.

1. Connect the serial cable to the serial port on the switch and to an RS-232 serial port on the workstation. If the serial port on the workstation is an RJ-45 port instead of RS-232, remove the adapter on the end of the serial cable and insert the exposed RJ-45 connector into the RJ-45 serial port on the workstation.
2. Open a terminal emulator application (such as HyperTerminal in a Windows environment, or TERM, TIP, or Kermit in a UNIX environment), and configure the application:
 - In a UNIX environment, enter the following string at the prompt:

```
tip /dev/ttyb -9600
```

If ttyb is in use, you can use ttya at the prompt:

```
tip /dev/ttya -9600
```
 - In a Windows environment, enter the parameters that are listed in Table 9-1 on page 235.

Table 9-1 Parameters to configure a Windows Terminal Emulator application

Parameter	Value
Bits per second	9600
Data bits	8
Parity	None
Stop bits	1
Flow control	None

9.1.2 Connecting to Brocade Fabric OS through SSH or telnet connection

The switch must be physically connected to the network. If the switch network interface is not configured or the switch was disconnected from the network, use a console session on the serial port to set up the IP address.

Complete the following steps to connect to the FOS by using the management interface:

1. From a management station, open an SSH or telnet connection that uses the IP address of the switch to which you want to connect. The login prompt is displayed when the SSH or telnet connection finds the switch in the network.

2. Enter the account ID at the login prompt.

3. Enter the password.

If the default system password was not changed, you are prompted to change it upon login.

4. Verify that the login was successful.

The prompt displays the switch name and user ID to which you are connected:

```
login: admin
password: xxxxxxx
```

Beginning in FOS 9.0.0, several CLI commands were consolidated and the redundant command were deprecated. For more information about the deprecated commands, see [Brocade Fabric OS Command Reference Manual, 9.0.x](#).

9.2 WebTools overview

Brocade WebTools is a graphical user interface (GUI) that is embedded in the FOS firmware. It enables administrators to monitor and manage single or small fabrics, switches, and ports. WebTools is started directly from a Web browser or from SANnav Management Portal.

WebTools supports the following operating systems:

- ▶ Red Hat 8.0 and 8.1
- ▶ Windows 10 Pro
- ▶ Windows 2019

The following browsers can be used to access WebTools:

- ▶ Chrome
- ▶ Firefox

If the switch is configured with logical fabrics, you can log in to any of the logical fabrics for which you have permission.

WebTools does not require a license. It is installed on the switch when you install FOS.

WebTools can be started from a browser or through SANnav Management Portal. Consider the following points:

- ▶ To start directly from a Web browser, open your browser, enter the IP address of the switch, and press **Enter**. You can use HTTP or HTTPS, for example: `http://10.77.77.77` or `https://10.77.77.77`.
- ▶ To start from SANnav Management Portal, locate the switch on the SANnav Inventory page, click the down arrow to the right of the switch, and select **View in WebTools** from the action menu (see Figure 9-1).

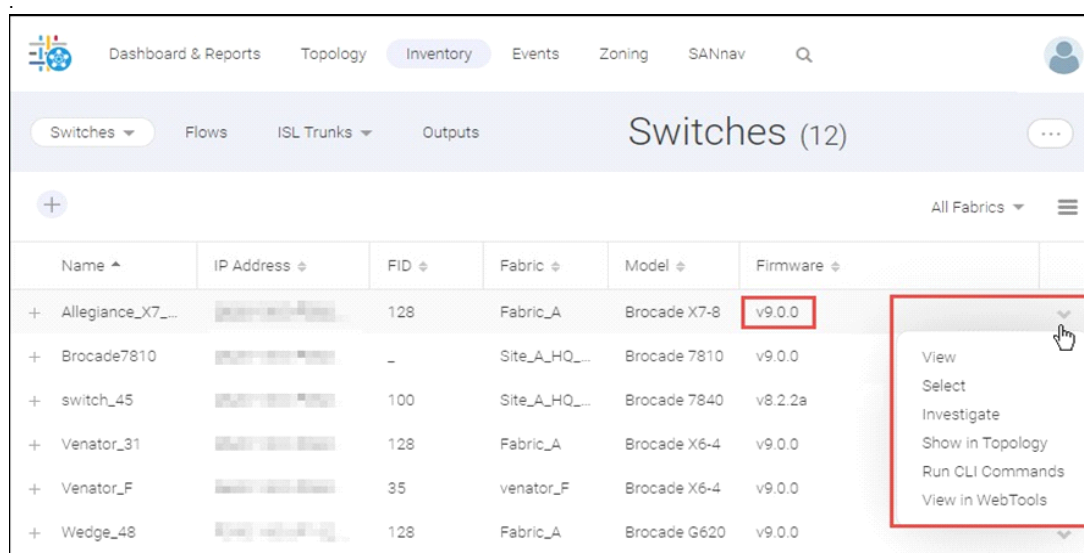


Figure 9-1 Starting WebTools from SANnav Management portal

When the WebTools Element Manager is started, you see a log in page (see Figure 9-2 on page 237.) Consider the following points:

- ▶ Enter the username, password, and logical switch name or fabric ID (FID).
- ▶ For the first switch login, the default username is `admin` and the default password is `password`. WebTools prompts you to change the default password.
- ▶ If you are logging in to a Virtual Fabrics-enabled platform and you do not specify a logical switch, you are logged in to the default logical switch, which uses fabric ID 128. For non-VF platforms, the FID option is not displayed.
- ▶ If you start from the SANnav Management Portal, you might not need to log in, depending on the SANnav single sign-on configuration.

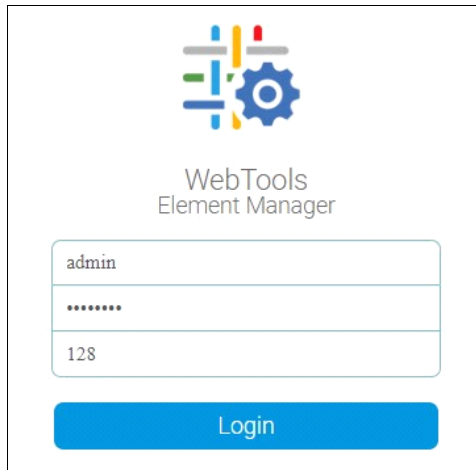


Figure 9-2 WebTools Element log in window

For more information about working with WebTools, see [Brocade Fabric OS WebTools User Guide, 9.0.x](#).

9.3 Fabric Vision overview

Fabric Vision technology is an advanced hardware and software architecture that combines capabilities from FOS, b-type Gen 7 devices, and SANnav. This suite of tools can be used to help administrators address problems before operations are affected and accelerate new application deployments.

Fabric Vision technology features are described next.

9.3.1 Monitoring, alerting, and performance suite

The Monitoring and Alerting Policy Suite (MAPS) is a storage area network (SAN) health monitoring and alerting application that features early fault detection and isolation. MAPS helps the SAN administrators to perform the following tasks:

- ▶ Proactively monitor the health and performance of the SAN infrastructure to ensure application uptime and availability by using predefined and custom policies and rules.
- ▶ Automate the policy-based real-time monitoring and alert thresholds on a port or group basis.
- ▶ Configure each MAPS-enabled switch to constantly monitor itself for potential faults and automatically alert you to problems before they become failures.
- ▶ Configure an entire fabric at once by using common rules and policies, or custom policies for specific ports or switch elements with the simplified fabric-wide threshold configuration, monitoring, and alerting.
- ▶ Simplify the configuration of a fabric with a common policy and remediate the conditions that require attention.
- ▶ Automatically fence, toggle, or quarantine ports without manual intervention.
- ▶ Monitor fabric-wide events, ports, FRUs, environmental parameters, security, traffic flows, and performance impacts.

MAPS is a licensed feature and requires a Fabric Vision license. This publication assumes that the administrator has activated (enabled) the Fabric Vision license is activated (enabled) in the FOS software to use the full set of MAPS options.

For more information about licensing, see [Brocade Fabric OS Software Licensing Guide, 9.0.x](#).

MAPS includes predefined rules, groups, and policies. However, you can create custom rules, groups, and policies.

MAPS enables cloning predefined rules, groups, and policies to facilitate monitoring. Predefined rules and groups are system-specific, and each system has its own rules and groups. All MAPS configuration is persistent, is retained across restarts or HA failovers, and can be uploaded and downloaded.

MAPS features the following predefined policies:

- ▶ `dflt_conservative_policy`: Contains rules with more lenient thresholds that allow a buffer and do not immediately trigger actions. Use this policy in environments where the elements are resilient and can accommodate errors.
- ▶ `dflt_moderate_policy`: Contains rules with threshold values between aggressive and conservative policies.
- ▶ `dflt_aggressive_policy`: Contains rules with stringent thresholds.
- ▶ `dflt_base_policy`: Contains rules based on the features that can be monitored without a Fabric Vision license.

MAPS can be configured and managed by using the CLI.

For more information about configuring MAPS, download [Brocade Fabric OS MAPS User Guide, 9.0.x](#).

MAPS can also be configured and managed through SANnav MAPS policy management. This feature allows an administrator to create MAPS policies, rules, and custom groups, and apply and activate MAPS policies across one or more switches of a fabric.

You can import and modify MAPS policies. You can also view MAPS-enabled switches, policies and rules, fabrics, and groups to which the switches belong, and switch configuration actions.

Complete the following steps to manage MAPS through SANnav:

1. Click **SANnav** in the navigation bar, and then, select **SAN Monitoring** → **MAPS Policy Management**. The Switches window opens (see Figure 9-3 on page 239).

Switch	Active	Policy	Fabric	Switch Actions
DCX18_29_...	faultySFP	Custom	LF_ode_Fab...	RAS Log Event, SNMP Trap, Switch Status C
DCX8510-8...	dflt_conser...	System	LF_ode_Fab...	Switch Status Critical, Switch Status Margin
Sw18_29	dflt_conser...	System	FabricReset...	Switch Status Critical, Switch Status Margin
FID010	dflt_conser...	System	DND-FID10_...	-
switch_10_...	dflt_conser...	System	DND-FID10_...	Switch Status Critical, Switch Status Margin
DT_Awing_...	dflt_conser...	System	PM_Gen6_F...	Switch Status Critical, Switch Status Margin

Figure 9-3 WebTools Switches window

2. Click the **Switches** tab to view the list of MAPS-enabled switches with an active policy name, policy type, and the fabric to which they belong, and switch configuration actions.
3. Click the **Policies** tab to view the policies that are applied to the switch.
4. Click the **Groups** tab to view the preconfigured groups available on the switch. You can create user-defined custom groups only for ports, SFPs, and circuits.

For more information about working with SANnav to manage MAPS, see the “MAPS management” section of [Brocade SANnav Management Portal User Guide, 2.1.0x](#).

9.3.2 Flow Vision

Flow Vision provides a comprehensive vision of and deep insight into fabric traffic flows, along with the ability to nondisruptively create and capture copies of traffic flows for analysis of traffic flows, bottlenecks, bandwidth utilization, and similar fabric connectivity functions.

Flow Vision also provides a test flow generation capability that you can use to pretest a SAN infrastructure for robustness. This test flow generation capability is also useful for testing the internal connections on a switch before deploying the switch into a production environment. In addition, Flow Vision allows you to test for fabric connectivity issues, such as slow drain, bandwidth utilization, and similar issues.

Automatic monitoring (introduced with FOS 9.0.0) provides I/O monitoring to the SANnav administrator. With these I/O Insight capabilities, Flow Vision provides proactive, non-intrusive, real-time monitoring, and alerting of storage I/O health and performance. The extra visibility delivers deep insights into problems and concerns that affect the ability to maintain service levels.

Integration with MAPS allows the automatically detected flows to be monitored by using the predefined rules. Hardware real-time violations are flagged at the I/O (ECT and FRT) and FC level (frame latency).

For more information about MAPS, see [Brocade Fabric OS WebTools User Guide, 9.0.x](#).

Automated monitoring is supported by SANnav. For details about SANnav, see [Brocade SANnav Management Portal User Guide, 2.1.0x](#).

Flow Vision includes the following features

► Flow Monitor

Flow Monitor running FOS 9.0.0 provides automated monitoring, which enables you to learn SCSI and NVMe traffic and obtain Read/Write Frame Count and Read/Write Data statistics. A cohesive view of metrics provides you with all of the necessary information in one output.

To provide automated non-intrusive monitoring in FOS 9.0.0, a new predefined flow command was introduced. When you upgrade to FOS 9.0.0, the `sys_flow_monitor` flow is activated.

After it is activated, the flow enables automatic learning for Gen 6 and 7 switches or blades running FOS 9.0.0. It also automatically clears the associated flow statistics and force deletes all user-defined flows.

The predefined flow automatically discovers application flows that are based on the direction and port parameters. With ingress or egress direction, this flow scan gathers DID, SID, SFID, and Protocol details.

► Flow Generator

Flow Generator is a test traffic generator with which you can pretest a SAN infrastructure (including internal connections) for robustness before deploying it.

Flow Generator enables you to perform the following tasks:

- Configure a 32-Gbps or 64-Gbps, FC-capable port as a simulated device that can transmit frames at a full 16 Gbps, 32 Gbps, or 64 Gbps line rate.
- Emulate a 32-Gbps or 64-Gbps SAN without having any 32-Gbps or 64-Gbps hosts or targets or SAN testers.
- Pretest the entire SAN fabric at the full line rate, including optics and cables on ISLs and internal connections within a switch.

Flow Generator uses simulation mode (SIM) ports. SIM ports behave like normal F_Ports, but are used for testing only. By using SIM ports, Flow Generator traffic is stopped at the destination port and does not leave the fabric.

Flow Generator can generate standard frames or create custom frames with sizes and patterns that you specify. A sample use case is creating a traffic flow from a Source ID (SID) to a Destination ID (DID) to validate routing and throughput. Creating a Flow From a Specific Source ID to a Specific Destination ID provides an example of the command and the results for this use case.

Flow Generator supports predefined flows to generate traffic between all configured SIM ports.

After you activate a Flow Generator flow, the flow stays active until you deactivate the flow. The flow stays active, even though the SID and DID SIM-ports are offline. When SID and DID SIM-ports are online, traffic starts.

Flow Generator also supports generating simulated traffic to a routed fabric. To generate traffic, you specify the imported proxy ID or the real PWWN of the device. Flow Generator support is similar to Flow Monitor support, except Flow Generator supports only flows from edge to edge but not from backbone to edge or from edge to backbone.

The Brocade MAPS can be used to monitor SIM port traffic thresholds while Flow Generator is running. MAPS treats SIM ports as F_Ports, so MAPS can issue warnings on these ports if threshold values are triggered. If you do not want to see MAPS warnings for SIM ports, you must disable MAPS monitoring for those ports.

Important: Do not use Flow Generator in an active production environment. Traffic that is created by Flow Generator can saturate the links, which might negatively affect any other traffic that is sharing the links.

► Flow Mirror

As storage networks get larger and more complicated, non-intrusive diagnostic tools are becoming increasingly important to help identify problems without disturbing the operating fabric. Flow mirroring is a diagnostic feature within Flow Vision that addresses this need.

With Flow Mirror, you can perform the following tasks:

- Non-disruptively create copies of application flows that optionally can be captured for deeper analysis.
- Define a traffic pattern and create a real-time copy of this traffic, with which you can analyze a live system without disturbing existing connections. You can also use this feature to view traffic passing through a port.
- Use a MAPS logical port group for an ingress or egress port. Logical port groups are not supported on CPU flow mirroring (CFM) or local flow mirroring (LFM).
- Use a predefined flow with the flow **--show** command to mirror SCSI command, first response, status, and ABTS frames from all Gen 6 F_Ports on a switch; for example:
flow --show sys_analytics_vtap.

Flow Mirror duplicates the specified frames in a user-defined flow and sends them to a sink. This sink can be one of the types:

- The local switch control central processor unit (CPU). This form is called *CPU flow mirroring* (CFM). It has a limit of 256 frames per second.
- An analyzer/packet sniffer that is connected through a port in the metaSAN. The bandwidth limit for a flow that uses this type of link is the bandwidth of the mirror destination port. This form is called *Local flow mirroring* (LFM) and mirrors the flow to a port on the same physical switch. This process requires that a loopback SFP be plugged in at the other end of the analyzer or on the port that is configured as a mirror port, which must be in the same domain.

All forms of mirroring are mutually exclusive. Only one form of mirroring (LFM, CFM, or RFM) can be active at a time.

Flow Mirror flows can be in an active or inactive state. If the mirror flow is “active”, mirroring starts immediately; if the flow is “inactive”, the flow must be activated (by using the flow **--activate** command) for mirroring to start. Mirrored flows can be unidirectional or bidirectional. A sample use case is to mirror the traffic flow from a slow-draining F_Port to see what is causing this condition.

For more information about the use of SANnav to create and manage Flow Vision flows, see [Brocade SANnav Flow Management User Guide, 2.1.0x](#).

9.3.3 ClearLink Diagnostics

Brocade ClearLink Diagnostic Port (D_Port) mode allows you to convert an FC port into a diagnostic port for testing traffic and running electrical loopback and optical loopback tests. The test results can be useful in diagnosing various port and link problems.

ClearLink D_Port testing can be used to exercise the hardware before deploying ports into the fabric. For example, you might install new SFPs to link two switches or to link a host or target to a switch, and you want to exercise the hardware before adding them to the fabric. If the test passes, you can add the ports to the fabric. If the test fails, you can fix the problem.

When D_Port testing is in progress, SANnav Management Portal changes the port type for all selected ports and associated attached ports to D-Port during the test. A port in D_Port mode does not carry any user traffic and is designed to run only specific diagnostics tests for identifying link-level faults or failures. When the test is complete, SANnav changes the port type back to the original port type.

The following tests are performed:

- ▶ Electrical loopback test
- ▶ Optical loopback test
- ▶ Link traffic test
- ▶ Link latency and distance measurement

Test results are presented after each test completes.

D_Port testing is supported on the following port types:

- ▶ E_Ports
- ▶ F_Ports
- ▶ N_Ports

Starting D_Port testing

D_Port testing is useful if you want to diagnose port and link problems or if you want to exercise hardware before deploying ports into the fabric.

Requirements:

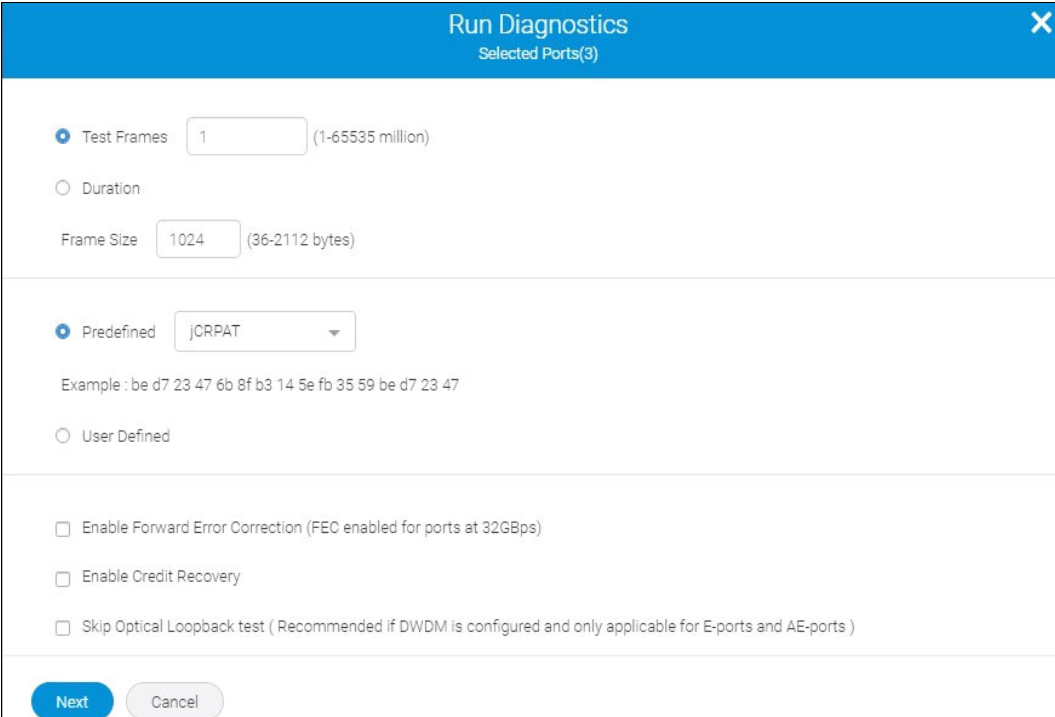
- ▶ You have Troubleshooting privilege with read/write permission.
- ▶ The source and destination ports are managed by SANnav Management Portal.

Note: A D_Port test cannot be stopped after it is started.

When you run a D_Port test, SANnav Management Portal changes the port type for all selected ports and associated attached ports to D_Port during the test. This issue might cause the fabric to segment. When the test is complete, SANnav changes the port type back to the original port type.

Complete the following steps:

1. Click **Inventory** in the navigation bar, and then, select **Switch Ports** from the drop-down list.
 2. Select the ports on which you want to perform the diagnostic tests:
 - To select a single port, click the down arrow in the rightmost column for the switch port, and select **Run Diagnostics**.
 - Alternatively, to select multiple ports, use the bulk select option:
 - i. Click the more button () , and select **Bulk Select**.
A column of check boxes is displayed on the left side of the table.
 - ii. Select the check boxes for the ports on which you want to run diagnostic tests, and then, click **Actions** → **Run Diagnostics** in the upper right of the window.
- If Run Diagnostics cannot be selected, one or more of the selected ports are not supported for running diagnostic tests.
3. Click **OK** in the confirmation dialog.
 4. Configure the test parameters, or accept the default values (see Figure 9-4).



The image shows a 'Run Diagnostics' dialog box with a blue header bar containing the title 'Run Diagnostics' and a close button 'X'. Below the header, it says 'Selected Ports(3)'. The main area is divided into sections for configuring test parameters. The first section has two radio buttons: 'Test Frames' (selected) and 'Duration'. The 'Test Frames' section has a text input field with '1' and a range '(1-65535 million)'. The 'Duration' section is currently empty. The second section has two radio buttons: 'Predefined' (selected) and 'User Defined'. The 'Predefined' section has a dropdown menu showing 'jCRPAT' and an example of a hexadecimal string: 'Example : be d7 23 47 6b 8f b3 14 5e fb 35 59 be d7 23 47'. The 'User Defined' section is currently empty. The third section contains three checkboxes: 'Enable Forward Error Correction (FEC enabled for ports at 32GBps)', 'Enable Credit Recovery', and 'Skip Optical Loopback test (Recommended if DWDM is configured and only applicable for E-ports and AE-ports)'. At the bottom, there are two buttons: 'Next' (blue) and 'Cancel' (grey).

Figure 9-4 Run Diagnostics: Test parameters

- a. Specify the number of frames to send (in millions) or the time duration for which the frame traffic test runs. The default is to send one million test frames.
- b. In the Frame Size field, enter the size of the test frames that are generated to run the test. The default value is 1024.

Note: For FOS 9.0.x and higher, the test uses a test frame size that is a multiple of four. If you specify a size that is not a multiple of four, the nearest higher multiple of four is used as the test frame size when the test is run. For example, if you specify a test frame size of 37, the tests are run with a test frame size of 40.

- c. Select a predefined pattern to be used in the payload, or enter a user-defined pattern in decimal. The default is to use the pattern jCRPAT.

You can optionally select the check boxes at the bottom of the window. FEC is not supported on D_Ports that are configured with dense wavelength division multiplexing (DWDM). For 32-Gbps ports, FEC is automatically enabled and cannot be disabled.

5. Click **Next**.

The diagnostic test starts running. A dialog displays the status of the test (see Figure 9-5).

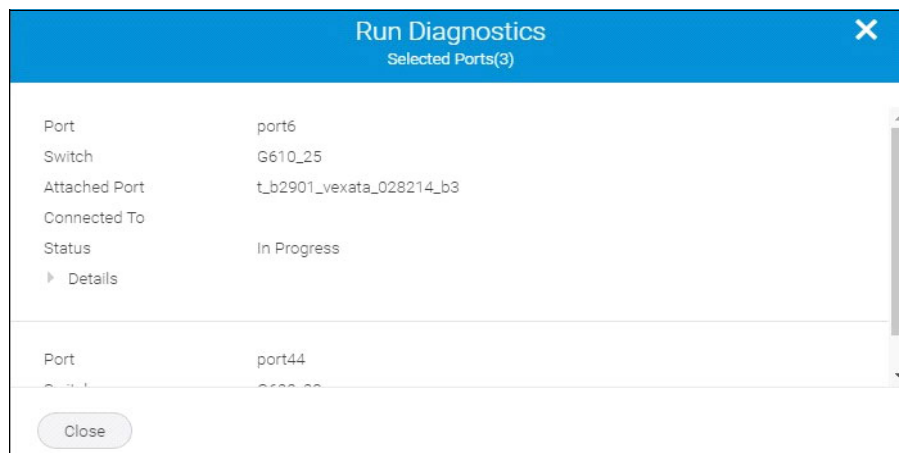


Figure 9-5 Run Diagnostics: Test status

Testing might take a few minutes to several hours, depending on the selected parameters.

6. Click **Close**. You can view this status on the Outputs page.
7. Click **Done** in the confirmation dialog to return to the Inventory page. The port type for the selected ports are changed to D_Port.

To view the test results, click **Outputs** in the subnavigation bar, and select **Diagnostics** from the drop-down list.

Scheduling D_Port testing

You can schedule D_Port testing to run at some time in the future. Because D_Port testing can cause the fabric to segment, you might want to schedule the test at a time when traffic is lighter, such as in the middle of the night.

The following requirements must be met:

- ▶ You have Troubleshooting privilege with read/write permission.
- ▶ The source and destination ports are managed by SANnav Management Portal.

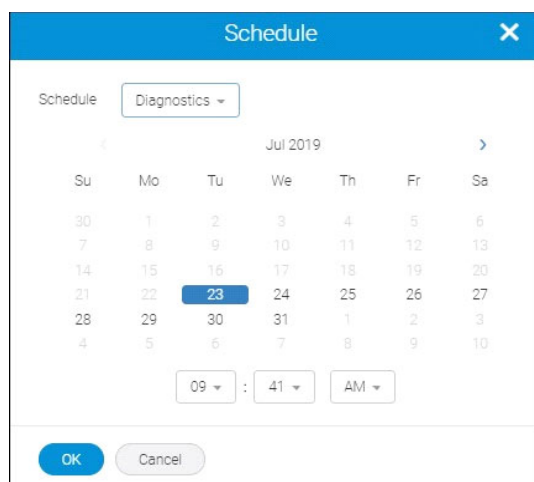
Complete the following steps:

1. Click **Inventory** in the navigation bar, and then, select **Switch Ports** from the drop-down list.
2. Select the ports on which you want to perform the diagnostic tests:
 - To select a single port, click the down arrow in the column on the right side for the switch port, and select **Schedule**.
 - Alternatively, to select multiple ports, use the bulk select option:
 - i. Click the More button () , and select **Bulk Select**. A column of check boxes displays on the left side of the table.
 - ii. Select the check boxes for the ports on which you want to run diagnostic tests, and then, click **Actions** → **Schedule** in the upper right of the window.

If Schedule is grayed out, one or more of the selected ports are not supported for running diagnostic tests.

3. Select **Diagnostics** from the drop-down list, select the date and time for the test to start, and then, click **OK** (see Figure 9-6).

If Diagnostics is unavailable, one or more of the selected ports are not supported for running diagnostics tests.



The image shows a 'Schedule' dialog box with a blue header and a close button (X). Inside, there's a 'Schedule' label and a dropdown menu currently showing 'Diagnostics'. Below this is a calendar for July 2019. The days of the week are listed as Su, Mo, Tu, We, Th, Fr, Sa. The dates are arranged in a grid. The 23rd is highlighted in blue. Below the calendar, there's a time selection area with two dropdowns for hours (09) and minutes (41), followed by an AM/PM dropdown (AM). At the bottom of the dialog are two buttons: 'OK' (blue) and 'Cancel' (gray).

Figure 9-6 Diagnostics

4. Click **OK** in the confirmation dialog.
5. Configure the test parameters or accept the default values and click **Next**.

6. Click **Done** in the confirmation dialog to return to the Inventory page.

You can view the scheduled test by clicking **Outputs** in the subnavigation bar and then selecting **Diagnostics** from the drop-down in the upper right of the page (see Figure 9-7).

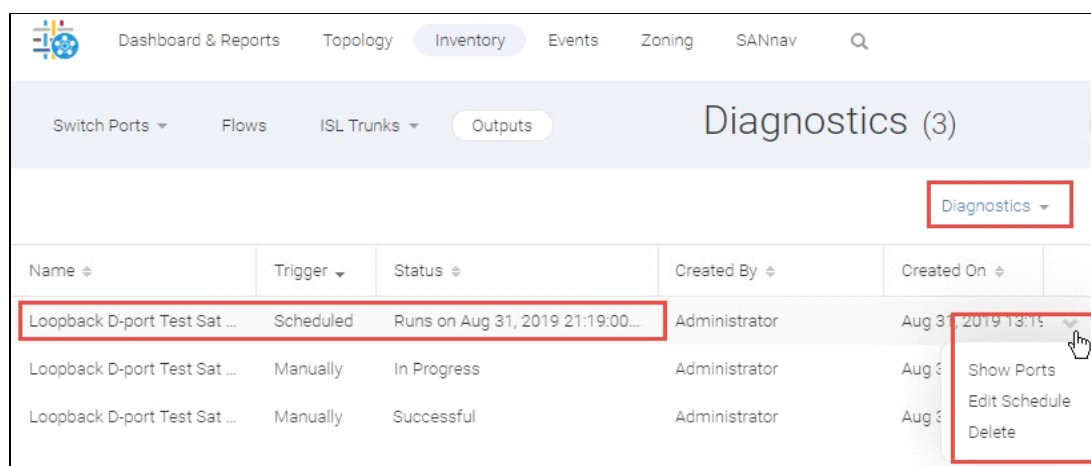


Figure 9-7 Inventory output

From the action menu of the scheduled test, you can display the ports that are selected for the test, edit the schedule, and delete the test.

Viewing D_Port test results

While D_Port tests are running and after the tests are complete, you can view the results from the Diagnostics section of the Outputs tab.

Only the user who ran or scheduled the test can view, delete, or rerun the test. Complete the following steps:

1. Click **Inventory** in the navigation bar, and then, click **Outputs**.
2. Select **Diagnostics** from the drop-down list in the upper right.
3. Locate the test that you want to view, and click **View** from the action menu in the rightmost column.

The View option is available for only tests that are completed or in progress. It is not available for scheduled tests.

If you used the **Bulk Select** option to run diagnostics on several ports at the same time, only one entry exists in the Outputs page for those tests (see Figure 9-8).

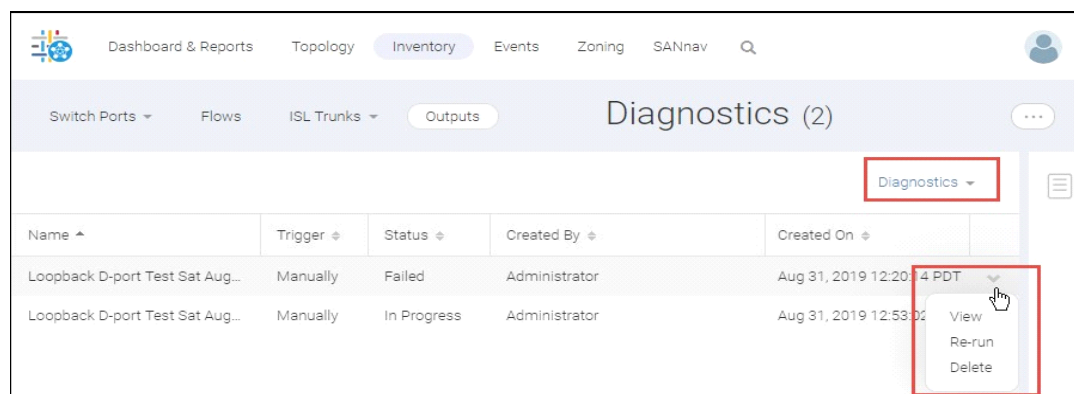


Figure 9-8 Inventory output with Bulk Select option

4. Click **Details** to see the detailed results of the test.

Figure 9-9 shows the test results when the **Bulk Select** option is used to run diagnostics on multiple ports.

You can rerun a test on the same port by selecting **Re-run** for a completed test in the Outputs page.

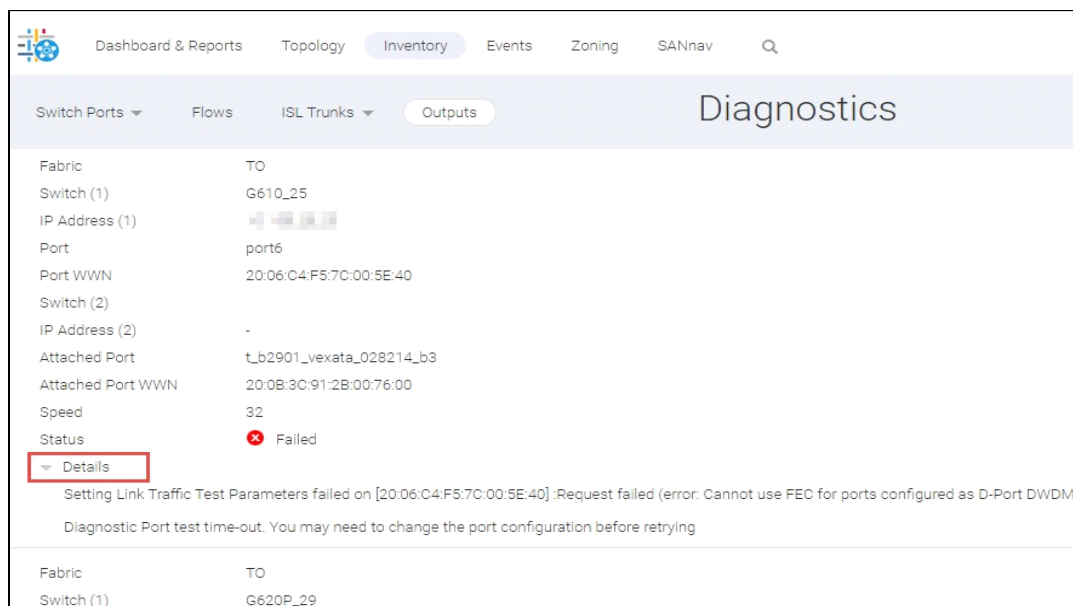


Figure 9-9 Inventory Output Details

For more information about the use of SANnav to manage ClearLink diagnostics, see the “Troubleshooting and Diagnostics” section of [Brocade SANnav Management Portal User Guide, 2.1.0x](#).

9.3.4 Fabric Vision SANnav managed features

In this section, we discuss the features that are managed by Fabric Vision SANnav.

Dashboards

Customizable health and performance dashboards provide at-a-glance views of switch status and various conditions that are contributing to performance issues. This feature enables administrators to get instant visibility into any hot spots at a switch level and take corrective actions.

Dashboards feature the following views:

- Overall status of the switch health and the status of each monitoring category, including any out-of-range conditions and the rules that were triggered.
- Historical information about the switch status for up to the last seven days. It automatically provides raw counter information for various error counters. This integrated dashboard view also provides a single collection point for all dashboard data from a fabric for a specific application flow.

Configuration and Operational Monitoring Policy Automation Services Suite

Configuration and Operational Monitoring Policy Automation Services Suite (COMPASS) simplifies deployment, safeguards consistency, and increases operational efficiencies of larger environments with automated switch and fabric configuration services. Administrators can configure a template or adopt an existing configuration as a template and seamlessly deploy the configuration across the fabric. In addition, they can ensure that settings do not drift over time with COMPASS configuration and policy violation monitoring within SANnav dashboards.

Fabric Performance Impact

Fabric Performance Impact (FPI) uses predefined MAPS policies to automatically detect and alert administrators to different congestion severity levels and to identify creditstalled devices (also known as *slow-drain devices*) or oversubscribed ports that can affect network performance. This feature pinpoints which devices are causing or are affected by congested ports. It also automatically quarantines credit-stalled devices to prevent buffer credit starvation.

By using FPI, you can monitor the congestion-related issues on all physical E_Ports and F_Ports at all times. Specifically, FPI monitors abnormal device latency and oversubscription issues.

9.3.5 I/O Insight

Proactively monitors I/O performance and behavior through integrated network sensors to gain deep insight into problems and ensure service levels. This capability automatically learns the traffic flows to a fabric non-disruptively and non-intrusively gathers I/O statistics for both SCSI and NVMe traffic from any device port on a Gen 6 and Gen 7 Fibre Channel platform. Then, it applies this information within an intuitive, policy-based monitoring and alerting suite to configure thresholds and alarms.

Integrated application-level and device-level I/O latency and IOPS monitoring provide the ability to baseline application performance and detect degraded performance. This feature enables administrators to proactively control performance and availability to ensure operational stability.

The following key capabilities are featured:

- ▶ Monitors individual host or storage devices to gain deeper insight into the performance of the network to maintain SLA compliance.
- ▶ Obtains total I/Os, first response time max/average, I/O latency (Exchange Completion Time, or ECT) max/ average, and outstanding I/Os total/ average performance metrics for a specific host or storage for a LUN or NSID to diagnose I/O operational issues.
- ▶ Enables tuning of device configurations with integrated I/O metrics to optimize storage performance.

9.3.6 VM Insight

Seamlessly monitors VM performance throughout a storage fabric with standards based, end-to-end VM tagging. Administrators can quickly determine the source of VM/application performance anomalies, and provision and fine-tune the infrastructure based on VM/application requirements to meet service level objectives.

9.3.7 Fabric Vision diagnostics

FOS 9.0.x includes the diagnostics that are discussed in this section. These diagnostics do not require configuration or management.

Forward Error Correction

Forward Error Correction (FEC) provides a data transmission error control method by including redundant data (error correcting code) to ensure error-free transmission on a specified port or port range. This feature enables recovery from bit errors in Fibre Channel links, which negates the need for retransmission.

Consider the following points regarding FEC:

- ▶ Always active on ports operating at 64 Gbps and 32 Gbps speed based on the standard, regardless of the FEC settings on the switch.
- ▶ Supported on:
 - E_Ports on 16 Gbps-capable switches at 10G/16G speeds.
 - The N_Ports and F_Ports of an Access Gateway using RDY, Normal (R_RDY), or Virtual Channel (VC_RDY) flow control modes.
 - F_Ports if the device attached to it supports FEC.
- ▶ Is enabled by default.
- ▶ Enables automatically when negotiation with a switch detects FEC capability.
- ▶ Persists after driver reloads and system restarts.
- ▶ Functions with features, such as QoS, trunking, and BB_Credit recovery.

Buffer credit recovery

Buffer credit recovery allows links to recover after buffer credits are lost when the buffer credit recovery logic is enabled. The buffer credit recovery feature also maintains performance. If a credit is lost, a recovery attempt is started. During link reset, the frame and credit loss counters are reset without performance degradation.

Consider the following points:

- ▶ Credit recovery is supported on E_Ports, F_Ports, and EX_Ports. Buffer credit recovery is enabled automatically across any long-distance connection for which the E_Port, F_Port, or EX_Port buffer credit recovery mechanism is supported
- ▶ For 64 Gbps (Gen 7) and 32 Gbps (Gen 6) FC switches and blades, use the **portCfgCreditRecovery** command to disable or enable buffer credit recovery on a port.

For more information about Forward Error Correction (FEC) and Buffer Credit Recovery, download see [Brocade Fabric OS Administration Guide, 9.0.x](#).



IBM b-type Gen 7 SAN automation

Most, if not all, IT administrators have first-hand experience in the growing complexity of managing enterprise infrastructure, including SANs.

IBM b-type Gen 7 is in the unique position to spot and understand the effect of automation helping organizations get more from their SAN infrastructure.

IBM b-type Gen 7 offers a combination of SAN automation with RESTful APIs and a SAN management platform to help organizations drive greater efficiency from their SANs.

This greater efficiency is realized accomplished by various means:

- ▶ IBM b-type Gen 7 SAN automation is provided with a multilayer architecture.
- ▶ RESTful API support on directors and switches, and on management tools.
- ▶ IBM b-type Gen 7's Ansible management framework, which is designed to eliminate repetitive tasks, simplify management and orchestrate across the full SAN infrastructure.

This chapter includes the following topics:

- ▶ 10.1, "Motivation to automate" on page 252
- ▶ 10.2, "REST API overview" on page 253
- ▶ 10.3, "Simple automation example" on page 255
- ▶ 10.4, "Ansible as an alternative" on page 257
- ▶ 10.5, "SANnav REST API" on page 258
- ▶ 10.6, "Conclusion" on page 259

10.1 Motivation to automate

Organizations should embrace SAN automation for the following reasons:

- Reducing human error and streamlining operational processes has never been more crucial.

As organizations move to digitize and adapt to new workloads, data availability, processing time, and agility to provision applications on-demand become the life-blood of the business. Achieving these business objectives demands a more efficient and expedient infrastructure management approach, which leaves no room for human error.

As a result, storage administrators must be freed from repetitive manual tasks, such as configuration management, reporting, documenting inventory, and troubleshooting. Instead, IT organizations need SAN automation to help them automate and orchestrate repetitive tasks, which significantly improves efficiency, and decreases the risk of operational mistakes.

- Demand for more accurate and more frequent infrastructure reports is on the rise.

It is not only IT managers who crave information about storage performance, usage, and forecasting. Business stakeholders also are asking for and expecting this data on demand.

No one wants to wait for a slot when storage administrators can allocate time to produce a report. This information must be available as frequently as business demands dictate, all at the click of a button.

Automation not only provides this kind of responsiveness that traditional manual storage management processes cannot deliver, but it also can be customized so that each stakeholder is getting more accurate data that is aligned to their responsibility.

- SAN configuration management must be streamlined.

With more enterprise applications demanding access to more data and with more virtual machines, deploying and configuring servers, storage, and the network became more time-consuming and more complex than ever. By streamlining SAN operations through automation, application provisioning workflows are simplified across hypervisor, network, and storage, which delivers agility and responsiveness to meet dynamic business demands.

- IT service delivery is not always as responsive as the business demands.

As organizations become increasingly reliant on having world-class IT services available to them for proactive, agile business decision-making, bottlenecks to IT service delivery must be identified and eliminated immediately. This identification must be done without hiring more storage administrators or boosting SAN-related Capex spending, which makes SAN automation the only viable option to drive increased IT agility and more tightly align IT service delivery with fast-changing business needs.

- Consistent configuration validation is a must.

Manual configuration changes are occurring with greater frequency as enterprises diversify their IT architectures in general, and their storage architectures specifically. SAN automation ensures validation of consistent configuration parameters across the different SAN fabrics in to facilitate troubleshooting of frequent alerts without reliance on manual intervention.

IBM b-type Gen 7 automation solutions use RESTful APIs to facilitate solutions architecture, share best practices, and get to production status faster.

10.2 REST API overview

The Brocade Fabric OS (FOS) REST API is a programmable web service interface for IBM b-type Gen 7 FOS that can manage IBM b-type Gen 7 SAN switches across a fabric. This API uses standard HTTP methods to perform Create, Read, Update, and Delete (CRUD) operations on the fabric configuration data. It also provides an interface for provisioning, status, and validation operations by using the YANG data model that is described in the YANG 1.1 RFC (but not the data store that is managed with NETCONF). An Apache web server that is embedded in FOS is used to serve the API.

The RESTful API approach helps you think of a network device as a web server. By using standard web-based tools, an automation can send and receive transactions to or from a network device just as it sends transactions to and from a website. Therefore, transactions occur over a secure socket by using HTTP rules to handle the exchange. The data appears in the form of XML or JSON, depending on the RESTful API services that are implemented on the networking device.

To interact with a SAN (or other) device, review its RESTful API reference to learn, among other things, what “Uniform Resource Identifiers” (URIs) you must use. (URIs are identifiers that can be used as part of a web address.) According to the documentation, the following URL is used for accessing a listing of the zones in the active configuration:

```
GET <base_URI>/running/zoning/defined-configuration/
```

The model that is used to represent state and configuration information is expressed in a modeling language called Yang. Yang describes the structure of the different elements that are inside the model. It is used to describe whether each element is read-only or read/write.

It also describes the type of data that the element can hold, such as string or integer, and it shows the relationship among various elements, the other nested elements they contain, their peer elements, and the parent elements that contain them. A segment of the description of a zone in Yang is shown in the following example:

```
list zone {
  key "zone-name";
  description
    "List of the members in the zone. The members can only be identified as a
    WWN, domain, index, or a zone alias.";
  leaf zone-name {
    type zoning-name-type;
    description
      "The zone name.";
  }
  leaf zone-type {
    type zone-type-type;
    description
      "The zone type. Not that target zone types cannot be created or modified
      (only deleted).";
  }
  container member-entry {
    description
      "The zone member.";
    leaf-list entry-name {
      type zone-member-type;
      min-elements 1;
      description
```

```

        "List of the members in the zone. The members can only be identified as a
        WWN, domain, index, or zone alias.";
    }
    leaf-list principal-entry-name {
        when "../..zone-type=1 or ../..zone-type=2";
        type zone-member-type;
        min-elements 1;
        description
            "List of the principal members in the peer zone. The members can only be
            identified as a WWN, domain, index, or zone alias.";
    }
}
}
}

```

Ordinarily, more information goes into a Yang module, such as revisioning and governance information; this listing omits them for brevity. Therefore, the Yang description is complete, but it is also verbose. Although this precision is necessary when interacting with the model programmatically, it is sometimes useful to get a global view of the abstraction that is provided by the model to see how the data is structured.

An open source tool that is called `pyang` can parse the Yang model and produce a tree that represents the elements in the model. The listing includes information about each element, such as whether it is read-only or read/write, a list, optional, or nested.

A representation of the zoning model in tree form is shown in the following example:

```

module: IBM b-type-zone
+--rw IBM b-type-zone
+--rw defined-configuration
|   +--rw cfg* [cfg-name]
|   |   +--rw cfg-name      zoning-name-type
|   |   +--rw member-zone
|   |       +--rw zone-name* zoning-name-type
|   +--rw zone* [zone-name]
|   |   +--rw zone-name      zoning-name-type
|   |   +--rw zone-type?     zone-type-type
|   |   +--rw member-entry
|   |       +--rw entry-name* zone-member-type
|   |       +--rw principal-entry-name* zone-member-type
|   +--rw alias* [alias-name]
|   |   +--rw alias-name     zoning-name-type
|   |   +--rw member-entry
|   |       +--rw alias-entry-name* union
+--rw effective-configuration
+--rw cfg-name?          zoning-name-type
+--rw checksum?          string
+--rw cfg-action?        uint8
+--rw default-zone-access? uint8
+--ro db-max?            uint32
+--ro db-avail?          uint32
+--ro db-committed?      uint32
+--ro db-transaction?    uint32
+--ro transaction-token? uint32
+--ro db-chassis-wide-committed? uint32
+--ro enabled-zone* [zone-name]
+--ro zone-name          zoning-name-type

```


Next, you perform a GET of the URI to return the current configuration by using the Custom Basic value that is returned from the login for authentication:

```
curl -v -H "Authorization: Custom_Basic
YWRtaW46eHh4OjNkYTl1ZmM3NzMxYjk4OGU2ODg1YzZkMGRjNWJlM
zMyNjBhZDYxZThkOWQ2MWMxNzNiMGVlMjU3YmM2OTcyYjA="
http://10.18.254.37/rest/running/zoning/defined-configuration
```

By default, curl uses the GET method, so you do not need to specify it. -H “Authorization: Custom_Basic YWR...jA=” is the authentication and session identifying string that is returned in the previous command. -H places the string into the GET request header, as shown in the following trace:

```
* Trying 10.18.254.37...
* Connected to 10.18.254.37 (10.18.254.37) port 80
(#0)
> GET /rest/running/zoning/defined-configuration
HTTP/1.1
> Host: 10.18.254.37
> User-Agent: curl/7.47.0
> Accept: */*
> Authorization: Custom_Basic
YWRtaW46eHh4OjNkYTl1ZmM3NzMxYjk4OGU2ODg1YzZkMGRjNWJlM
zMyNjBhZDYxZThkOWQ2MWMxNzNiMGVlMjU3YmM2OTcyYjA=
>
< HTTP/1.1 200 OK
< Date: Wed, 31 Jan 2018 16:09:39 GMT
< Server: Apache
< Cache-Control: no-cache
< X-Frame-Options: DENY
< Content-Security-Policy: default-src 'self'
< X-Content-Type-Options: nosniff
< X-XSS-Protection: 1; mode=block
< Connection: close
< Transfer-Encoding: chunked
< Content-Type: application/yang-data+xml <
<?xml version="1.0"?>
<Response>
<defined-configuration>
<cfg>
<cfg-name>CFG_FABRIC_A</cfg-name> <member-zone> <zone-name>CLUSTER1</zone-name>
<zone-name>Z_AIXHOST_FCS2_VMAX01_SN1234_9F0 </zone-name>
...
<alias> <alias-name>esx66_5d3d00</alias-name> <member-entry>
<alias-entry-name>10:00:8c:7c:ff:5d:3d:00 </alias-entry-name>
</member-entry>
</alias>
</defined-configuration>
</Response>
* Closing connection 0
```

The results appear as an XML data segment that is structured according to the description in the Yang model. Therefore, so it is important to access that model along with the RESTful API manual. The models can be found on GitHub as a repository among those in IBM b-type’s repositories at <http://github.com/IBM-b-type/yang>.

The RESTful API manual can be retrieved from the Broadcom website. Having retrieved the zoning information from the fabric, you close the session by using the CLI command (the results are omitted to save space):

```
curl -v -H "Authorization: Custom_Basic  
YWRtaW46eHh4OjNkYT1lZmM3NmMxYjk4OGU2ODg1YzZkMGRjNWJlM  
zMyNjBhZDYxZThkOWQ2MWMxNzNiMGVlMjU3YmM2OTcyYjA=" http://10.18.254.37/rest/logout
```

In version 8.2.2 or later of FOS, the REST API session-less operation allows you to provide authentication credentials directly for each GET request. Essentially, FOS REST API session-less operation completes the login, GET operation, and logout as one complete request. You can use only basic authentication formats for REST API session-less operation, which includes plain text or Base64.

The following example shows a GET request that uses plain text authentication:

```
curl -u admin:password http://10.155.2.190/rest/running/IBM  
b-type-media/media-rdp>
```

10.4 Ansible as an alternative

The previous section showed an example of an approach that uses a procedural methodology. The workflow starts at the beginning, runs a series of steps, and then ends. Most traditional programs work this way.

Ansible takes a declarative approach. Rather than provide steps, Ansible describes each of the hosts in an inventory. The description appears in a document that is called a *playbook*.

For example, rather than provide steps to install a specific application, Ansible describes a host state where the application is installed. When you run the playbook, Ansible takes no action if the application is installed. If the application is not installed, Ansible calls installation routines so the host is brought into the wanted state without requiring the administrator to write any specific steps.

In the realm of storage networks, the use of a declarative language means you can describe switches and fabrics where, for example, a zone is configured with the suitable hosts and storage arrays. When you run the Ansible playbook, those zones are defined as needed, and the hosts and storage arrays are added to them, if necessary.

With some other declarative automation utilities, it is necessary to install an agent on each host that the utility manages. This agent retrieves the commands from a command center and runs them on the local host.

Ansible is unique in that it does not require agents. To make switch state changes, Ansible establishes a Secure Shell session to a proxy and sends it a small Python script. The script carries out the necessary operations by using the switch API and removes itself from the host.

You need two different skill sets to successfully implement an Ansible solution. First, you must understand the most common playbook operations. These operations are coded and installed for use by the playbooks. As vendors announce support for Ansible, they also provide script libraries for the most common tasks. If a task exists that is required but is not available in the official Ansible distribution, the open source community might provide code for that task in publicly available repositories.

Second, you must understand your business needs to provide ongoing playbook development. The person that is maintaining the playbooks does not need to be a programmer and does not need to know how remote system operations occur. That person only needs to know the wanted outcomes and must construct playbooks in YAML, which is the markup language that is used by Ansible. An Ansible playbook is shown in the following example:

```
---
- hosts: edgeSwitches
  vars_files:
  ../fos_passwords.yml
  gather_facts: False
  tasks:
  - name: run fos commands
    IBM b-type_fos_command:
      switch_login: "{{switch_admin_account}}"
      switch_password: "{{switch_password}}"
      switch_address: "{{switch_ip_address}}"
      command_set:
        - command: alicreate "SampleAlias1", "10:23:45:67:76:54:32:10"
```

The three dashes at the beginning are part of the YAML specification. The hosts section identifies automation target switches. You can keep sensitive information in a separate file, as demonstrated by the `fos_passwords.yml` line. The name of a variable in double braces, such as `{{switch_password}}`, indicates variable substitution. The variable file that is specified in `vars_files` indicates where to find external variables.

10.5 SANnav REST API

SANnav Management Portal and SANnav Global View provide end-to-end visibility into enterprise SANs. These tools detect, analyze and take action based on SAN behavior and performance, which helps administrators get to the root of problems faster and remediate them fully. Storage administrators can troubleshoot across the storage fabric in as little as 30 seconds. This ability is unprecedented with any other shared storage infrastructure architectures.

The SANnav REST API provides functions that complement what is found in the FOS REST APIs. The SANnav REST API includes support for the following SANnav features:

- ▶ List the fabrics that are managed by the SANnav server.
- ▶ List the seed and principal switch data for each fabric.
- ▶ List the switches (members) in the fabric.
- ▶ Display the FCR topology information for routing topologies, such as edge-to-edge, backbone-to-edge, and edge-to- backbone.
- ▶ Configure event-forwarding.
- ▶ Acknowledge or unacknowledge events.
- ▶ Display a list of filtered events.
- ▶ Perform an inventory search.

10.6 Conclusion

SAN automation is a critical element in IT modernization and digital transformation because it helps organizations manage storage-related processes more efficiently without hiring more administrators or adding to the storage infrastructure Capex budget. SAN automation is a high-use approach to turning network storage into a strategic asset.

IBM b-type Gen 7's commitment to SAN automation, which is combined with its long-standing leadership in storage fabrics and technical innovation, makes it an ideal candidate for your IT infrastructure automation strategy.

For more information about automation, see [this web page](#).



IBM SANnav Management Suite overview

This chapter describes how to deploy IBM SANnav Management Suite 2.1 in your SAN Infrastructure. It also includes detailed steps for installing the IBM SANnav Management Portal and SANnav Global View 2.1.

This chapter also describes all of the system and server requirements before you start the IBM SANnav Management suite installation. The system and server requirements are included for deployment of SANnav Management Portal and SANnav Global View.

Note: In this chapter, SANnav Management Portal might also be referred to as *SANnav*.

This chapter includes the following topics:

- ▶ 11.1, “IBM SANnav Management Suite 2.1 release overview” on page 262
- ▶ 11.2, “IBM SANnav Management Portal 2.1 release overview” on page 264
- ▶ 11.3, “IBM SANnav Global View 2.1 release overview” on page 293
- ▶ 11.4, “IBM SANnav WebTools” on page 308
- ▶ 11.5, “Reference guides” on page 310

11.1 IBM SANnav Management Suite 2.1 release overview

SANnav Overview SANnav is the next generation SAN management application suite for IBM b-type Gen 7 SAN environments. SANnav allows you to efficiently manage your SAN infrastructure through various easy-to-use functions.

SANnav implements a highly scalable client/server architecture for SAN management. With a modern browser-based UI, SANnav eliminates the need for a Java-based thick client.

The user interface of SANnav is designed based on real-world use cases and user workflows, which provide a highly intuitive user experience.

SANnav also uses a micro-services-based architecture that is based on Docker container technology that allows it to scale to meet the management needs of small and large SAN environments and those environments that can change over time. This scalable architecture also allows SANnav to support new functions in the future without causing degradation to the performance of the application.

To address the management needs of large-scale SAN environments or those environments that are distributed by function or location, SANnav supports a hierarchical management model. In this model, a higher-level “global” application (SANnav Global View) provides comprehensive visibility, summarization, and seamless navigation across multiple instances of the SANnav Management Portal application.

The following SANnav product offerings are available:

- ▶ IBM SANnav Management Portal
- ▶ IBM SANnav Global View

11.1.1 IBM SANnav Management Portal

SANnav allows you to efficiently manage your SAN infrastructure through various easy-to-use functions. SANnav implements a highly scalable client/server architecture for SAN management.

With a modern, browser-based UI, SANnav eliminates the need for a Java-based thick client. The SANnav user interface is designed based on real-world use cases and user workflows, which provide a highly intuitive user experience.

SANnav Management Portal allows management of one or more SAN fabrics that are in the same or different geographical locations, and it supports a maximum of 15,000 physical SAN ports. For environments that are larger than 15,000 ports, you can deploy multiple SANnav Management Portal instances. SANnav Management Portal does not replace Brocade WebTools or the Fabric OS (FOS) command-line interface.

For more information, see [this web page](#).

11.1.2 IBM SANnav Global View

You can get a comprehensive view of a SAN environment that spans multiple SANnav Management Portal instances through the SANnav Global View product. SANnav Global View allows you to navigate seamlessly across multiple SANnav Management Portal instances and drill down to any individual instance to perform detailed monitoring, investigation, and troubleshooting.

For more information, see [this web page](#).

11.1.3 New hardware platforms supported in SANnav 2.1

SANnav 2.1.0 is a major new release of IBM b-type Gen 7 two Fibre Channel SAN management software products: IBM SANnav Management Portal and IBM SANnav Global View. SANnav Management Portal 2.1.x supports the introduction of IBM b-type Gen 7 platforms with FOS 9.0.0, and adds features and capabilities that make managing IBM b-type Gen 7 SAN environments easier than ever before.

Supported hardware and software

SANnav Management Portal 2.1.x supports the FOS software versions and hardware platforms that are listed in this section.

- ▶ FOS software support:
 - FOS 9.0 or later
 - FOS 8.0 or later
 - FOS 7.4 or later
- ▶ Supported SAN switches:
 - IBM b-type Gen 7 (64 Gbps) Fixed-Port Switches
 - IBM SAN64B-7 Gen 7 56-port Switch
 - IBM b-type Gen 7 (64 Gbps) Directors:
 - IBM SAN256B-7 Gen 7 4-slot Director
 - IBM SAN512B-7 Gen 7 8-slot Director
 - IBM b-type Gen 6 (32 Gbps) Fixed-Port Switches:
 - IBM SAN24B-6 Switch
 - IBM SAN64B-6 Switch
 - IBM SAN128B-6 Switch
 - IBM SAN18B-6 Extension Switch
 - IBM b-type Gen 6 (32 Gbps) Directors:
 - IBM SAN256B-6 Director
 - IBM SAN512B-6 Director
 - IBM b-type Gen 5 (16 Gbps) Fixed-Port Switches:
 - IBM 24B-5 Switch
 - IBM SAN48B-5 Switch
 - IBM SAN96B-5 Switch
 - IBM SAN42B-R Extension Switch
 - IBM b-type Gen 5 (16 Gbps) Directors:
 - IBM SAN384B-2 Director
 - IBMSAN768B-2 Director

- IBM b-type Gen 4 (8 Gbps) Fixed-Port Switches:
 - IBM SAN24B-4 Switch
 - IBM SAN06B-R Extension Switch
- IBM b-type Gen 4 (8 Gbps) Directors:
 - IBM SAN768B Director
 - IBM SAN384B Director

Switches that support data streaming

SANnav uses Kafka to provide real-time data streaming from switches that are running FOS 8.2.1a or higher. Data streaming provides collection of high-frequency performance statistics, FCIP performance and error statistics, and flow statistics and violations.

The following Gen 7 platforms support data streaming:

- ▶ IBM SAN64B-7 Switch
- ▶ IBM SAN256B-7 and SAN512B-7 Directors

The following IBM b-type Gen 6 platforms support data streaming:

- ▶ IBM SAN24B-6, SAN64B-6, and SAN128B-6 Switches
- ▶ IBM SAN18B-6 Switch
- ▶ IBM SAN256B-6 and SAN512B-6 Directors

11.2 IBM SANnav Management Portal 2.1 release overview

SANnav Management Portal allows management of one or more SAN fabrics that are in the same or different geographical locations. It also supports managing a maximum of 15,000 physical SAN ports.

For environments that are larger than 15,000 ports, you can deploy multiple SANnav Management Portal instances, which in turn can be managed by SANnav Global View. Use SANnav Management Portal to monitor and manage fabrics, switches, switch ports, and other elements of your SAN.

Dashboards provide summary status and performance information, from which you can drill down to get detailed views. By using filters and tags, you can sort and search your inventory to find the information that you want.

A highly flexible reporting infrastructure allows you to generate custom graphical or tabular reports. SANnav Management Portal does not replace Brocade WebTools or the FOS command-line interface.

11.2.1 Browser requirements

Any notebook or machine that starts web applications can be used to start SANnav Management Portal. For optimal performance, have at least 16 GB of memory. The following browsers can be used to access the SANnav server:

- ▶ Chrome
- ▶ Firefox

Note: If you access the client from Remote Desktop, the user interface might degrade performance. Starting Brocade WebTools 9.0.0 and later from a SANnav client is supported on Chrome and Firefox. Starting Brocade WebTools versions earlier than 9.0.0 is supported on Firefox only.

11.2.2 New SANnav Management Portal server platform support and infrastructure

SANnav Management Portal 2.1.0 or 2.1.0a can be deployed on a single bare-metal host or virtual machine (VM) or on a cluster of bare-metal servers or VMs. The tables that are included in this section provide more information about server requirements.

For specific information about installing SANnav Management Portal on a VM, see [this web page](#).

11.2.3 Server requirements

Figure 11-1 shows the VM or bare metal installation requirements.

Max Switch Ports Under Management (Base or Enterprise)	Operating System	Host Type	Minimum CPU	Minimum number of CPU Sockets	Memory	Hard Disk
600 Ports (Base) 3000 (Enterprise)	RHEL 7.7 (*), 7.8 (**), 8.0 (*), 8.1 (**) CentOS 7.7 (*), 7.8 (**), 8.0 (*), 8.1 (**)	Bare metal/ VMware ESXi VM	16 cores, 2000 MHz	2	48 GB	600 GB
15000 (Enterprise)	RHEL 7.7 (*), 7.8 (**), 8.0 (*), 8.1 (**) CentOS 7.7 (*), 7.8 (**), 8.0 (*), 8.1 (**)	Bare metal/ VMware ESXi VM	24 cores, 2000 MHz	2	96 GB	1.2 TB

(*) 7.7 and 8.0 are supported in SANnav Management Portal 2.1.0 and deprecated in SANnav Management Portal 2.1.0a
 (**) 7.8 and 8.1 are supported in SANnav Management Portal 2.1.0.a

Figure 11-1 VM or bare metal installation

Figure 11-2 shows the requirements for SANnav installation on a bare metal server or VM.

Product/Edition	Max. Switch Ports/Instances under Management	Operating System	Host Type	CPU	Memory
Brocade SANnav Global View	Up to 20 SANnav Management Portal instances	RHEL 7.7, 8.0, & 8.1 CentOS 7.7, 8.0, & 8.1	Bare Metal, ESXi VM	16 Cores	32 GB

Figure 11-2 Requirements for SANnav installation on a bare metal server or VM

11.2.4 SANnav Management Portal scalability 2.1.0

Figure 11-3 lists the scalability limits.

Feature	Scalability Limit – SANnav Management Portal <u>Base</u>	Scalability Limit – SANnav Management Portal <u>Enterprise</u>
Maximum number of SAN ports managed	600	15,000
Maximum number of end device ports managed	2000	40,000
Maximum number of end device ports per fabric	10,000	
Maximum number of events stored	10 Million	
Maximum number of MAPS violations stored	10 Million	
Port statistics stored	• 5-minute samples are stored for up to 30 days. • 1-hour data is stored for 30 days. • 1-day aggregated data is stored for 30 days. • 2-second samples are collected for up to 3 days for a maximum of 100 user-selected Gen 6 or Gen 7 ports. These ports can be on the same switch or across multiple Gen 6 or Gen 7 switches. Once data collection is complete, the data is retained for 14 more days.	
Extension Tunnel Statistics stored	• 5-minute samples are stored for up to 30 days. • 1-hour data is stored for 30 days. • 1-day aggregated data is stored for 30 days. • 5-second samples are collected for up to 3 days for a maximum of 100 circuits (only supported for the SX6 Blade and 7810 switch). These circuits can be on the same switch or across multiple switches. Once data collection is complete, the data is retained for 14 more days.	
Maximum number of Flows Supported	• Enterprise Edition (15K ports) platform: 400K Flows, 4 AMPs • Base Edition (600 ports) platform and Enterprise (3K ports) platform: 8K Flows from 4 Gen 6 chassis or 16 fixed-port switches.	

Figure 11-3 Scalability limits

11.2.5 Pre-installation requirements

Before you install IBM SANnav Management Suite, ensure that the following requirements are met:

- ▶ The SANnav Management Portal software is downloaded.
- ▶ SANnav Installation Pre-Reqs and migration Quicklist checklists are completed.

For more information about downloading the suitable documents, see 11.2.7, “SANnav installation and migration QuickStart checklists” on page 277.

11.2.6 Downloading the SANnav Management Portal software

Download the following software and documentation from the Brocade IBM Assist website:

- ▶ IBM b-type Software and Firmware Download Instructions (IBM Assist Site for IBM Brocade Software downloads)
- ▶ Step-by-step instructions to access the FOS firmware, SANnav, and b-type IBM Network Advisor (INA).

Important:

- ▶ You must have active IBM Support to access Brocade Software for any device.
- ▶ IBM is responsible for supporting IBM b-type (Brocade) Storage Networking solutions.

Brocade FOS and SAN Management software can be accessed at [IBM Support Fix Central](#) or by calling +1 800 426 7378.

Complete the following steps:

1. Go to [IBM Support Fix Central](#).
2. Click the **Select product** tab.
3. Select **Product Group** → **System Storage**.
4. Under **Select from System Storage**, select **Storage area network (SAN)**.
5. Under **Storage area network (SAN)**, select **SAN b-type** for Brocade switches or **SAN management software** for IBM Network Advisor and SANnav.
6. If you select **SAN b-type**, complete the following steps:
 - a. Under **Select from SAN b-type**, select the correct model.
 - b. Under **Installed Version**, select the correct FOS version.
 - c. Click **Continue**.

Figure 11-4 shows the **Select product** tab with the completed fields.

Fix Central provides fixes and updates for your system's software, hardware, and operating system. Not looking for fixes or updates? Please visit [Passport Advantage](#) to download most purchased software products, or [My Entitled Systems Support](#) to download system software.

For additional information, click on the following link.
[Getting started with Fix Central](#)

Find product | **Select product**

Select the product below.

When using the keyboard to navigate the page, use the **Alt** and **down** arrow keys to navigate the selection lists.

Product Group*
System Storage

Select from System Storage*
Storage area network (SAN)

Select from Storage area network (SAN)*
SAN b-type

Select from SAN b-type*
SAN128B-6 Switch (8960-F96-N96)

Installed Version*
8.x

Continue

Figure 11-4 Select product tab

7. On the Select fixes page, select the suitable fix pack (FOS version) for your product, and then select the download method.

8. If you select **SAN management software**, complete the following steps:
- Under “Select from SAN management software”, select the correct management software.
 - For Product Name, enter IBM SANnav.
 - Type the product name to access a list of product choices. When using the keyboard to navigate the page, use the Tab or down arrow keys to navigate the results list.
- Here are the results for “IBM SANnav” (Figure 11-5):
- IBM SANnav Global Edition
 - IBM SANnav Management Portal Base Edition
 - IBM SANnav Management Portal Enterprise Edition

The screenshot shows a web interface with two tabs: "Find product" (active) and "Select product". Below the tabs, there is instructional text: "Type the product name to access a list of product choices." and "When using the keyboard to navigate the page, use the Tab or down arrow keys to navigate the results list." A "Product selector*" field contains the text "ibm sannav" with a right-pointing arrow button. A dropdown menu is open, displaying three options: "IBM SANnav Global View", "IBM SANnav Management Portal Base Edition", and "IBM SANnav Management Portal Enterprise Edition".

Figure 11-5 Results for “IBM SANnav”

9. In this example, we choose **IBM SANnav Management Portal Enterprise Edition**. Select the correct version in the **Installed Version** drop-down menu (in our example, we select **2.x**), and click **Continue** (Figure 11-6).

The screenshot shows the same web interface as Figure 11-5. The "Product selector*" field now contains "IBM SANnav Management Portal Enterprise Edition" with a right-pointing arrow button. Below this, the "Installed Version*" field is a drop-down menu showing "2.x" with a downward arrow. At the bottom left, there is a black button labeled "Continue".

Figure 11-6 IBM SANnav Management Portal Enterprise Edition

10.You must provide the correct information for the FOS entitlement check (see Figure 11-7).

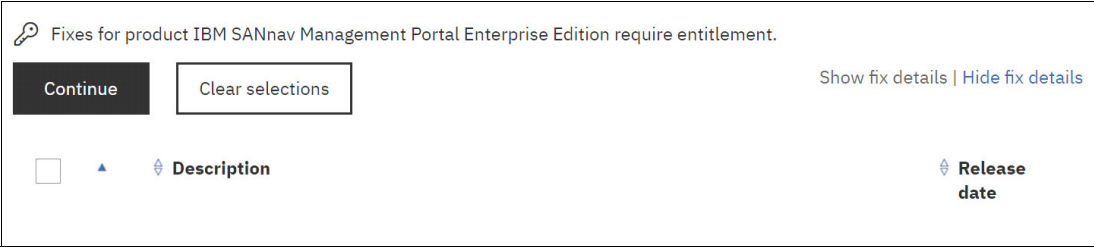


Figure 11-7 Fix Central: Quick order

11.Select the checkbox at the left. The Show Fix Details for SAN Software Networking b type window opens (Figure 11-8).

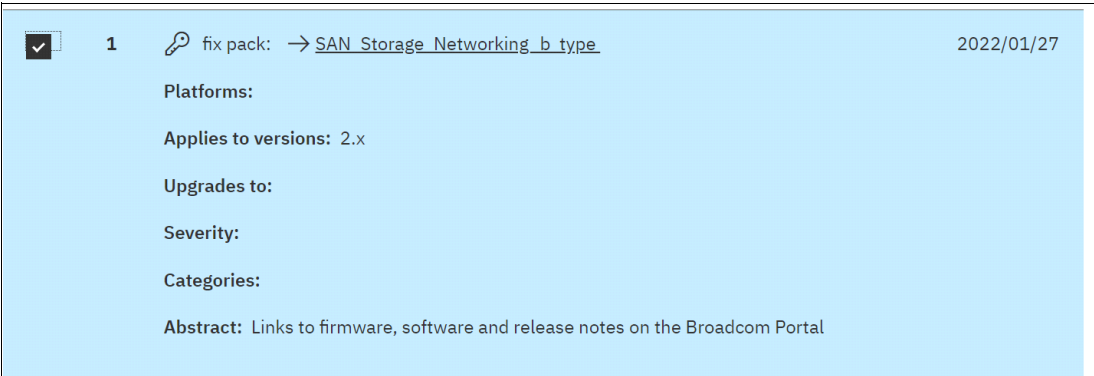


Figure 11-8 Show Fix Details for SAN Software Networking b type window

12.Click **Continue**.

13. Log in to IBM with your IBMid (if requested) and create an IBMid account (if necessary) (Figure 11-9).

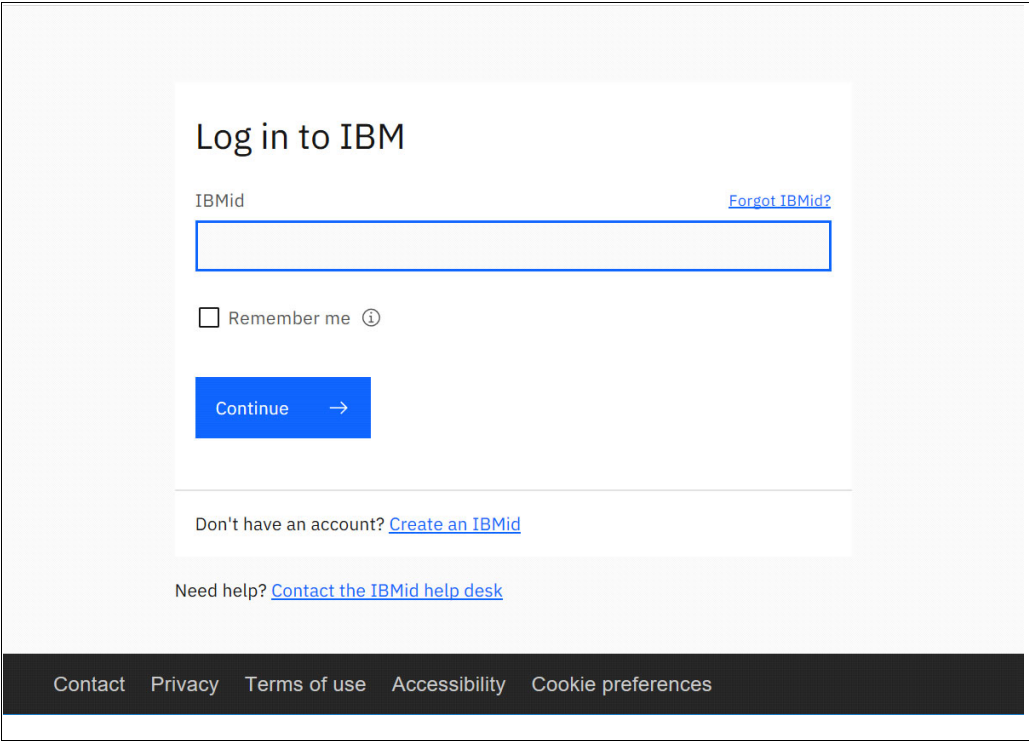
The image shows the IBM login page. At the top, it says "Log in to IBM". Below that is a text input field for the IBMid, with a "Forgot IBMid?" link to its right. Under the input field is a checkbox labeled "Remember me" with an information icon. Below the checkbox is a blue "Continue" button with a right-pointing arrow. At the bottom of the main content area, it says "Don't have an account? [Create an IBMid](#)". Below that is a link: "Need help? [Contact the IBMid help desk](#)". At the very bottom, there is a dark footer bar with links: "Contact", "Privacy", "Terms of use", "Accessibility", and "Cookie preferences".

Figure 11-9 Logging in to IBM

A successful login opens the Fix Central Download Files page for SAN Management software (Figure 11-10).

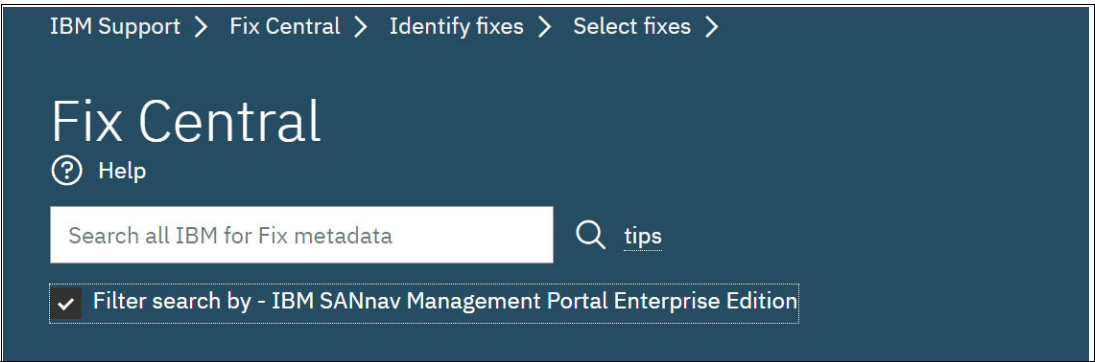
The image shows the Fix Central interface. At the top is a breadcrumb trail: "IBM Support > Fix Central > Identify fixes > Select fixes >". Below this is the "Fix Central" title in large white text on a dark blue background. Under the title is a "Help" link with a question mark icon. Below that is a search bar with the placeholder text "Search all IBM for Fix metadata". To the right of the search bar is a magnifying glass icon and the word "tips". Below the search bar is a filter dropdown menu that is currently set to "Filter search by - IBM SANnav Management Portal Enterprise Edition".

Figure 11-10 Fix Central Download files page for SAN Management software

14. The IBM Quick Order for SANnav Management Portal Edition window opens, as shown in Figure 11-11 on page 271. Enter the IBM Machine Type and 7-digit Machine Serial Number. Click **Continue**.

Note: IBM employees may bypass the Machine type and Machine Serial Number fields.

Fix Central
Quick order

[I am an IBMer - Skip this step](#)

Please provide the serial number of the machines for which Machine Code update(s) are designated and will be installed (each a "Target Machine").

The **Type Number** is a 4-digit number (usually followed by a 3-character Model Identifier) printed on the exterior of your IBM system. It may be the first part of an ID labeled "Model" or "System Model" ID.

The **Serial Number** is a 7 digit ID labeled "S/N" on the exterior of your IBM system. Dash ("-") characters may be omitted.

The **Country** selection is based on the location of your IBM system.

See [more information](#) for details about this page, and the actions available below.

Country
United States

	Machine type	Machine Serial Number
1.	2499	

[+ Add another](#)

Upload machine type and serial number data

[Browse](#)

[Save to file](#)

If you are a third party representative of an IBM customer who has been duly authorized by the IBM customer to download Machine Code update(s), then by downloading Machine Code update(s), you agree to comply with all obligations of the IBM customer with respect to any Machine Code or Machine Code updates. Any copying, reproduction, distribution or installation of Machine Code updates, other than as expressly authorized by IBM, is prohibited.

[Continue](#) [Clear settings](#) [Back](#)

[Contact and feedback](#)

Figure 11-11 Entering the Machine Type and Serial Number

15. Download the file by using Download Director and clicking **Download now** (Figure 11-12). Then, select the link to BroadcomSWPortal.

IBM Support > Fix Central > Identify fixes > Select fixes

Fix Central
Help

Search all IBM for Fix metadata [tips](#)

Filter search by - IBM SANnav Management Portal Base Edition

Download files using Download Director
SAN management software, IBM SANnav Management Portal Base Edition (2.x, All platforms)

Order number: 427327510

✓ 1. fix pack: SAN_Storage_Networking_b_type	Jan 27, 2022

SAN_Storage_Networking_b_type

The remaining files for this fix can be downloaded from a third party location.

[LinkToBroadcomSWPortal](#)

[Download now](#) [Back](#)

IBM recommends only the most current service level for this product be downloaded. IBM makes available older service levels for archival purposes. Older service levels may not contain all available fixes, including fixes that address security vulnerabilities. As you may expose yourself to security vulnerabilities during installation, IBM strongly advises you to review Security Bulletins prior to each installation of the product(s). IBM strongly advises you to review the support pages in the IBM Support Portal for information, including Security Bulletins, about service levels for this product. IBM also recommends that you subscribe to My Notifications through the IBM Support Portal.

Download options
[Change download options](#)

Quick order
[Share this download list](#)

[Contact and feedback](#)

Figure 11-12 Download files using Download Director window

16.Leave the IBM website by clicking **Continue** (Figure 11-13).

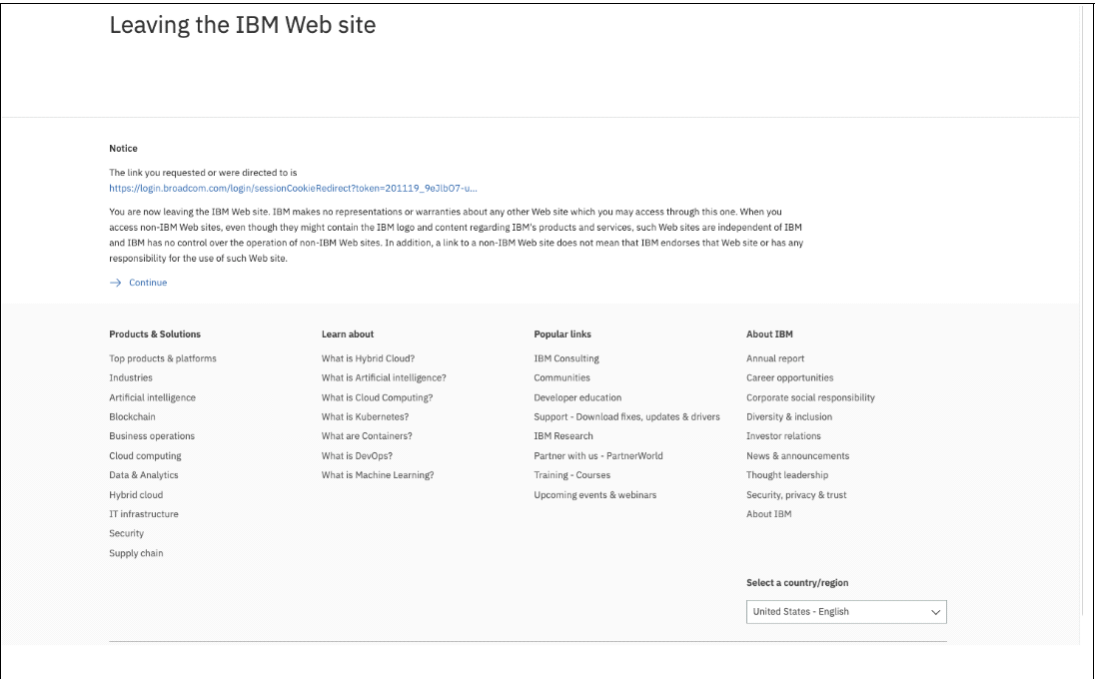


Figure 11-13 Leaving the IBM Web Site window

The welcome window of the Broadcom Assist Portal opens (Figure 11-14).

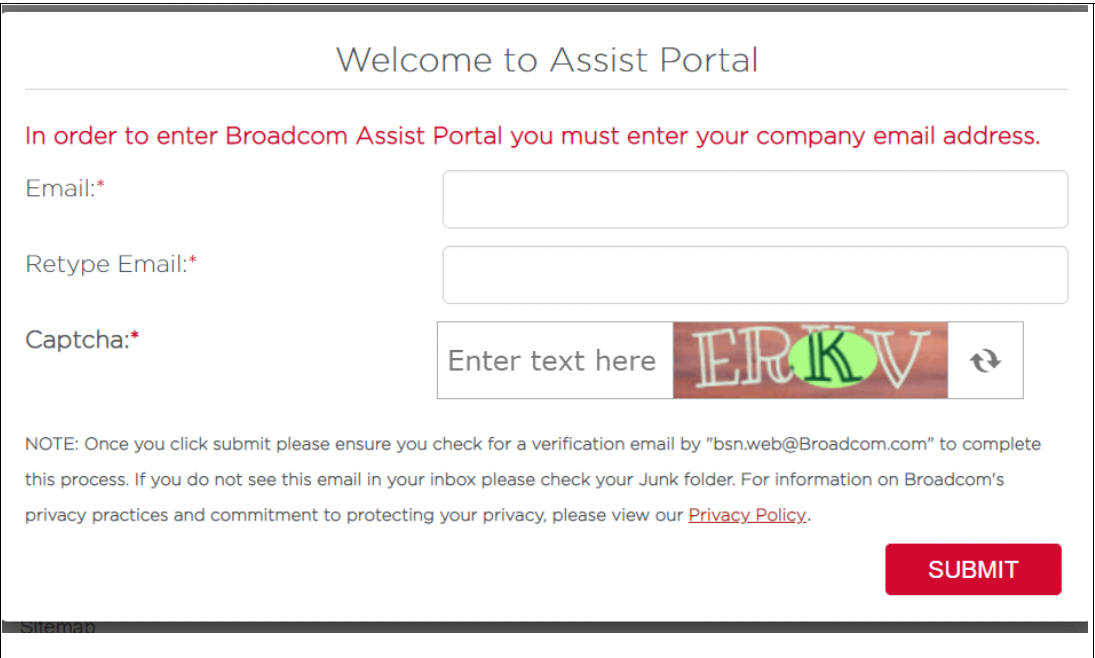


Figure 11-14 Broadcom Assist Portal welcome window

At the time of writing, a few significant changes have been implemented within the Assist Portal, as shown in Figure 11-15:

- ▶ There is a new CAPTCHA login window.
- ▶ A valid serial number is now required to perform software downloads. The portal validates the serial numbers that are entered.
- ▶ Platform Specific Downloads (PSDs). Unlike traditional FOS release images that can be installed on any supported Brocade switch and director, FOS PSD images must be downloaded separately for each platform on which the FOS release will be used. The full list of unique FOS PSD images is noted in the PSDs. When users access the portal, they are asked to enter an email address. Use the email address that is associated with the PSD.

Welcome to Assist Portal

In order to enter Broadcom Assist Portal you must enter your company email address.

- The Email ID you are using does not currently have access to the Brocade Assist Portal for additional assistance, email bsn-webcustomercare@broadcom.com. If you feel you reached this message in error, please contact your Support provider or Broadcom Fibre Channel Networking Technical Support team at <https://www.broadcom.com/support/fibre-channel-networking>. When contacting Broadcom Support team make sure you have the serial number of the device you are trying to download software or documentation for.

Email:*

Retype Email:*

Captcha:*  

NOTE: Once you click submit please ensure you check for a verification email by "bsn.web@Broadcom.com" to complete this process. If you do not see this email in your inbox please check your Junk folder. For information on Broadcom's privacy practices and commitment to protecting your privacy, please view our [Privacy Policy](#).

SUBMIT

Figure 11-15 Entering an email address

After you log in to the Assist Portal, you see the Brocade End User License Agreement (EULA), as shown in Figure 11-16.

The screenshot shows the 'Assist Portal' interface. At the top right, there are two tabs: 'Software Downloads' and 'Documentation'. A red rectangular box highlights a warning message in the center of the page. The warning text reads: 'Brocade wants to remind you that it is a violation of the Brocade EULA to install Fabric OS on equipment not under active entitlement or past its End of Support (EOS) Date. Excessive software downloads may result in your companies account being suspended. If you are having any issues finding what you are looking for, contact your service provider for assistance.' Below the warning, there is a 'Serial Number' search section with a text input field and a 'Search Serial' button. A small note below the input field says: 'Please enter a single serial number or multiple serials (Max 10) for each platform software is needed for. Example: (ARC323456/8903043567) no space'. Below this is a section with two radio buttons: 'Browse' (selected) and 'Search Product/Document Title'. Under 'Browse', there is a 'Select*' dropdown menu showing 'SANnav' and a 'Search' button. Below the dropdown is a large text area displaying 'SANnav 2.x' with a dropdown arrow on the right. At the bottom, there is a 'Related Links' section with a link to 'Assist Portal User Guide'.

Figure 11-16 Brocade EULA

The Assist Portal is divided in to two sections:

- Software Downloads
- Documentation

To go to either Software Downloads or Documentation, click the appropriate tab at the upper right (Figure 11-17).

This screenshot is identical to the one in Figure 11-16, showing the 'Assist Portal' interface. However, in this version, the 'Software Downloads' and 'Documentation' tabs at the top right are highlighted with a red rectangular box, indicating they are the focus of the action described in the caption.

Figure 11-17 Clicking the Software Downloads or Documentation tabs

- If you click **Software Downloads**, click **SANnav** in the **Select** list, and then click the carrots at the right to go to the option that you need, as shown in Figure 11-18.

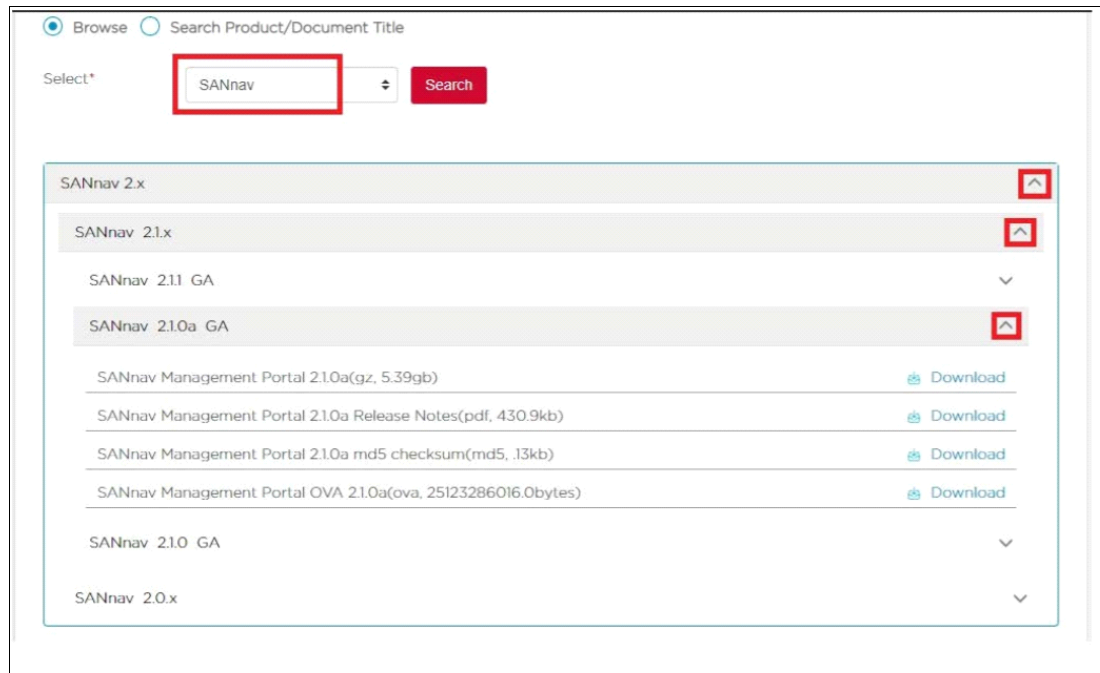


Figure 11-18 Selecting SANnav in the Software Downloads tab

- If you need to download the relevant documentation for your FOS versions and software products, click the **Documentation** tab at the upper right, as shown in Figure 11-19.

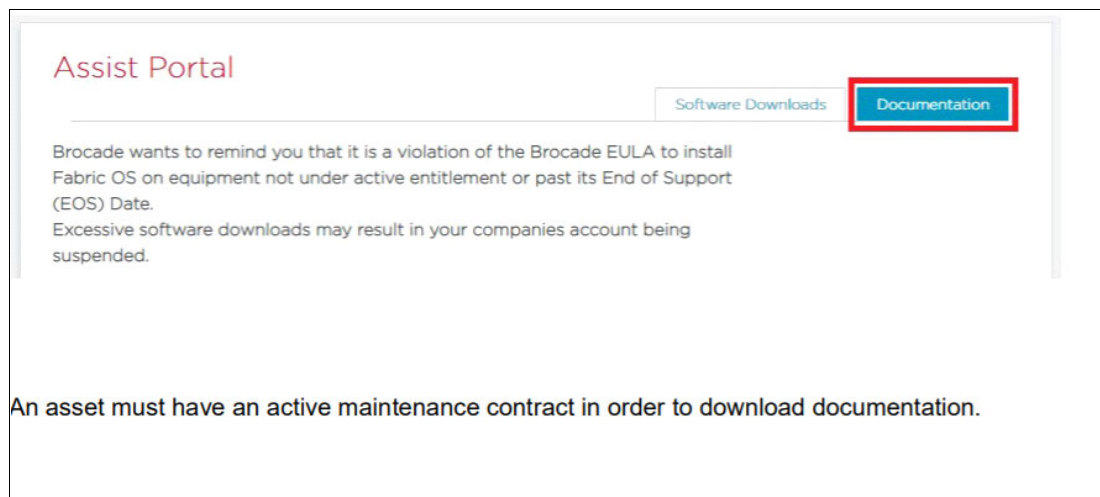


Figure 11-19 Clicking the Documentation tab

The window that is shown in Figure 11-20 opens. To find the documentation for a specific product, enter the serial number of the product into the Serial Number field and click **Search Serial**.

Enter the serial number of a product with active maintenance, and click the **Search Serial** button.

Assist Portal

Brocade wants to remind you that it is a violation of the Brocade EULA to install Fabric OS on equipment not under active entitlement or past its End of Support (EOS) Date.
Excessive software downloads may result in your companies account being suspended.
If you are having any issues finding what you are looking for, contact your service provider for assistance.

Serial Number **Search Serial**

serials (Max 10) for each platform software is needed for. Example: (ABC12345678901234567)
no space

The serial number entered is verified for product type and maintenance status. Once active maintenance is confirmed, documentation options are presented based on the product type. In the following example, documentation options for a Brocade X6 Director are provided.

Brocade X6 Director 8.x

Brocade X6 Director 9.x

SANnav 2.x

Figure 11-20 Searching for a product by using a serial number

After the documentation options are displayed, click the carrots at the right to find the document that you need. Click **Download** to download the file (Figure 11-21).

Brocade X6 Director 8.x

Brocade X6 Director 9.x

SANnav 2.x

SANnav 2.1.x

SANnav 2.1.1 GA

SANnav 2.1.0a GA

SANnav Management Portal 2.1.0a Release Notes(pdf, 430.9kb)

SANnav 2.1.0 GA

SANnav 2.0.x

⬆

⬆

⬆

⬆

⬆

⬆

⬆

⬆

⬆

Download

Figure 11-21 Downloading the documentation

11.2.7 SANnav installation and migration QuickStart checklists

These checklists are provided for experienced users who are familiar with SANnav installation. For all other users, start with Installation Overview.

For more information, see [this web page](#).

Installation checklist

Table 11-1 shows a quick checklist for installing SANnav.

Table 11-1 Installation checklist

#	Item	Description
1	Ensure that your server meets the requirements for SANnav installation. Upgrade the operating system if you are running an unsupported version.	For more information, see 11.3.2, “SANnav Global View system and server requirements” on page 294
2	Review and comply with the installation prerequisites.	For more information, see 11.2.11, “Installation prerequisites for SANnav Management Portal” on page 282
3	Ensure that the required ports are open in the firewall.	
4	Configure the <code>firewalld</code> back end if you use CentOS/RHEL 8.1 or higher.	
5	Download the SANnav software package to the folder where you want to install the application.	Do not create the SANnav installation folder with a space in the name; otherwise, installation fails.
6	Untar the <code>.gz</code> file.	<code>tar -xvzf <package_name>.gz</code> The resulting directory is referred to as <code><install_home></code> throughout the rest of the checklists.
7	Install SANnav.	<code><install_home>/bin/install-sannav.sh</code>
8	Check the server status.	<code><install_home>/bin/check-sannav-status.sh</code>

Migration checklist

11.2.8, “Installation overview” on page 278 provides a quick checklist for migrating from an earlier version of SANnav.

Table 11-2 Migration checklist

#	Item	Description
1	Back up the current SANnav installation before you start the migration process.	
2	▶ Ensure that your server meets the requirements for SANnav installation. ▶ Upgrade the operating system if you are running an unsupported version.	For more information, see 11.3.2, “SANnav Global View system and server requirements” on page 294
3	Review and comply with the installation prerequisites.	11.2.11, “Installation prerequisites for SANnav Management Portal” on page 282
4	Ensure that the required ports are open in the firewall.	
5	Configure the <code>firewalld</code> back end if you use CentOS/RHEL 8.1 or higher.	

#	Item	Description
6	Download the SANnav software package to the folder where you want to install the application.	Note: Do not create the SANnav installation folder with a space in the name; otherwise, the installation fails.
7	Extract the .gz file.	<code>tar -xvzf <package_name>.gz</code>
8	Install SANnav.	<code><install_home>/bin/install-sannav.sh</code>
9	Check the server status.	<code><install_home>/bin/check-sannav-status.sh</code>
10	Clear the browser cache and restart the SANnav client (browser).	Close the previous version of the SANnav client (browser) and clear the browser cache before starting the new version of SANnav.

11.2.8 Installation overview

The SANnav application uses a script-based installation. You must run the scripts that are provided in the `<install_home>` directory to install the application. All the scripts for the SANnav application must be run in the bash shell.

Note: SANnav Management Portal and SANnav Global View are two different software products. You cannot install both software products on the same physical host or virtual machine (VM). However, you can install Management Portal and Global View on different VMs in the same host, if the host has enough resources.

For switches that are running FOS versions lower than 8.2.2, port 22 is required for SANnav Management Portal to use the internal firmware repository and SCP and SFTP servers.

For more information, see the Installation Prerequisites in *Brocade SANnav Management Portal Installation and Migration Guide, 2.1.0x.*, which is available at the Assist Portal.

Assist Portal

The Assist Portal is divided into the following sections:

- Documentation
- Software Downloads

Navigating between the two sections is done by clicking the corresponding tab in the upper right of the page.

The Documentation Tab allows users to download relevant documentation for FOS versions and software products. Click the Documentation tab on the upper right.

Documentation

In the Documentation tab, users can download relevant documentation for FOS versions and software products. Click the **Documentation** tab in the upper right (see Figure 11-22 on page 279).

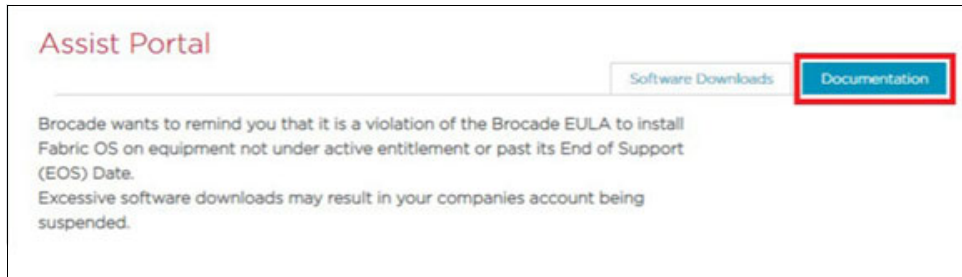


Figure 11-22 Documentation tab

Users must have an active maintenance to download any documentation.

Firewalls

If a firewall exists between the client and server or between the server and SAN, you must open a set of ports for SANnav to function correctly. The list of ports is provided in Firewall Requirements for SANnav Management Portal. If the installation script detects that an earlier version of SANnav is running, you are prompted whether you want to migrate your data to the new version.

11.2.9 Deploying SANnav Management Portal

During SANnav Management Portal installation, you are prompted several times to accept default values or provide customized values for various settings. If you are migrating from an earlier version of SANnav, you are not prompted for these customizations, and the settings from the previous installation remain in effect.

Figure 11-23 lists the installation customization options.

Item	Description
Docker installation directory	The default home directory for installing Docker is <code>/var/lib/docker</code> , but you can change this to another directory.
Swap space	SANnav requires 16GB swap space. - If there is not enough swap space, the installer prompts you to provide a location in which to create the remainder of the swap space. - If there is no swap space, the installer prompts you to provide a location in which to create the full 16GB of swap space..
IPv6 capability	The default is IPv4 communication between SANnav and the SAN switches. If you have IPv6-capable switches in your data center, you can configure SANnav to use IPv4 and IPv6 communication.
HTTP port 80 to HTTPS redirection	Choose to allow or disallow port 80 to be redirected to port 443 (default) or to another port that you can customize. If you disallow port 80 redirection, the web browser times out when pointed to port 80 and must be explicitly pointed to port 443 or the customized port to be able to log in to SANnav.
Server-to-switch communication protocol	Select an option to configure HTTP or HTTPS connections between SANnav and the SAN switches: 0 for HTTP (Insecure communication.) 1 for HTTPS (Secure communication. Requires that you have an IP-provided SSL certificate or self-signed certificate and that your switches are configured for HTTPS.) 2 for HTTPS then HTTP (First HTTPS is tried, and if that fails, HTTP is used.)
Single sign-on (SSO) options when launching Web Tools	If you launch Web Tools from the SANnav application, SANnav prompts you to provide switch login credentials. You can configure SANnav to automatically log in to the switch when launching Web Tools for switches running Fabric OS 9.0.0 or higher. 0 for always log in manually. SANnav prompts you for switch login credentials. 1 to log in with switch credentials. SANnav does not prompt you, but attempts to log in to the switch using the credentials that SANnav used when discovering the switch. 2 to log in with user credentials. SANnav does not prompt you, but attempts to log in to the switch using the credentials that the user used when logging in to SANnav. For switches running Fabric OS versions earlier than 9.0.0, SANnav always prompts you to log in to the switch when launching Web Tools, regardless of the SSO settings.
Port customization	You can customize ports when installing SANnav Management Portal. To use a default port, that port must be unused and available. The following is the list of default values: - SSH server port is 22. - Client-to-server HTTPS port: Default HTTPS port is 443. - SNMP trap: Default SNMP trap port is 162. - Syslog port: Default syslog port is 514. - Secure syslog port: Default secure syslog port is 6514. Note: If you are installing SANnav in IPv6 mode, the following ports are reserved for internal communication. Do not use any of these ports for customization: 5432, 6060, 6061, 7021, 7022, 7051, 7052, 7053, 7054, 7055, 7056, 7060, 7072, 7080, 7082, 7087, 7088, 7089, 7090, 7096, 7099, 7890, 7997, 8021, 8022, 8081, 8082, 8083, 8085, 8200, 9022, 9090, 9091, 9100, 9101, 9200, 9443, 12181, 18081, 19093, 19094, and 19095.
Database password	You can change the default SANnav database (Postgres database) password. You are given the option to proceed with the installation using the default password or to choose a new password for the SANnav database. The default password is "passw0rd" (where 0 is a zero).
SCP/SFTP password	You can change the default SANnav Management Portal SCP/SFTP password. You are given the option to proceed with the installation using the default password or to choose a new password for the SANnav internal SCP/SFTP server. The default password is "passw0rd" (where 0 is a zero).
License auto-renewal	By default, SANnav is configured to automatically retrieve and activate a renewal license when the license expires. You can deactivate automatic license renewal, in which case you must manually apply the license yourself.
Allowing data collection to be sent to Broadcom	SANnav collects usage data for the application. You can decide whether SANnav sends the data to Broadcom to improve user experience in the future.

Figure 11-23 SANnav installation customization

11.2.10 System and server requirements for SANnav Management Portal

You must meet all the system and server requirements before you start the SANnav Management Portal installation. The table that is shown in Figure 11-24 lists the system and server requirements for deployment of SANnav Management Portal.

Note: The disk space requirement that is listed in the table in Figure 11-24 is for SANnav only. Be sure to account for required extra space by the operating system, for saving files, and for SANnav TAR files and extracted files.

The disk space can be from a direct-attached disk or through a network-mounted disk:

- ▶ The default home directory for installing Docker is `/var/lib/docker`; however, you can choose another location during installation. Docker must be installed on a local disk.
- ▶ The default swap space directory is the `/` directory. If the directory does not have enough space, you can choose a different location during installation. The required number of CPU cores must be equally distributed over the sockets.

Requirement	Base License or Enterprise License with up to 3000 Ports	Enterprise License with up to 15,000 Ports
Operating system	• Red Hat Enterprise Linux (RHEL): 7.7, 7.8, 8.0, and 8.1 • CentOS 7.7, 7.8, 8.0, and 8.1 Language = English, Locale = US Although RHEL and CentOS 7.7 and 8.0 continue to be supported, it is recommended that you upgrade to RHEL/CentOS 7.8 or 8.1. Note: RHEL and CentOS 7.8 are supported only in SANnav 2.1.0a and higher.	• RHEL: 7.7, 7.8, 8.0, and 8.1 • CentOS 7.7, 7.8, 8.0, and 8.1 Language = English, Locale = US Although RHEL and CentOS 7.7 and 8.0 continue to be supported, it is recommended that you upgrade to RHEL/CentOS 7.8 or 8.1. Note: RHEL and CentOS 7.8 are supported in SANnav 2.1.0a and higher.
Processor architecture	x86	x86
Host type	• Bare metal server • VMware ESXi virtual machine	• Bare metal server • VMware ESXi virtual machine
CPU	16 cores	24 cores
CPU sockets (minimum)	2	2
CPU speed (minimum)	2000 MHz	2000 MHz
Memory (RAM)	48 GB	96 GB
Hard disk space (minimum)	600 GB, distributed as follows: • 450 GB — Installation directory • 120 GB — Docker installation directory • 16 GB of swap space	1.2 TB, distributed as follows: • 1050 GB — Installation directory • 120 GB — Docker installation directory • 16 GB of swap space

Figure 11-24 System and server requirements for SANnav management portal installation

11.2.11 Installation prerequisites for SANnav Management Portal

Review all SANnav Management Portal installation prerequisite tasks that are listed in Table 11-3 before you extract the installation file.

Note: Use the latest generation processors for better SANnav performance.

Table 11-3 Installation prerequisites

Task	Description
Gather necessary information and components	Ensure that the following information is available: ▶ Root user credentials. You must log in to the SANnav server as the root user or a user with root privilege. ▶ The SANnav Management Portalserver IP address.
Uninstall other applications	SANnav is expected to be installed and run on a dedicated host. If any other application is installed on the host, uninstall it before starting SANnav installation. If you are migrating SANnav, do not uninstall the current SANnav instance.
Uninstall Docker, if already installed	The SANnav installation installs Docker. If you have a Docker installed other than the Docker that SANnav installs, you must remove it before starting the installation.
Ensure that IP network addresses do not conflict with Docker addresses	▶ SANnav comes with Docker preinstalled. By default, Docker uses an IP address range of 192.168.255.240/28. ▶ If you use IPv4, do not use an address in these ranges when choosing your VM IP address and gateway. If you do, the IP address is unreachable even though the deployment may be successful. ▶ IPv6 connectivity is not affected. ▶ The installation script allows you to change the default Docker address range to a different address range.
Check operating system requirements	▶ Ensure that the operating system can be loaded through a bootable disk or PiXe server. ▶ Ensure that the <code>lsf</code> and <code>nslookup</code> packages are installed on the operating system machine. If they are not installed, run the following commands to install them: – <code>yum install lsf</code> – <code>yum install bind-utils</code>
Format the XFS file system	If you use XFS as the file system, make sure that you set <code>d_type=true</code> while creating the disk. You can verify this information by running the <code>xfs_info <docker-installation-directory></code> command and verifying that <code>ftype=1</code> . The default Docker installation directory is <code>/var/lib</code> .
Set umask and ulimit	▶ "umask" for the root user must be set to 0022. Enter the following command to set the umask: umask 0022 You must set the umask before you unzip the installation files. If you extracted the installation files before setting the umask, you must delete the installation folder, run umask 0022, and unzip the files again. ▶ Ensure that the ulimit is set correctly. To set the ulimit, edit the <code>/etc/security/limits.conf</code> file and add the following limit at the end of the file: <code>elasticsearch - nofile 65536</code>

Task	Description
Check port 22 availability	By default, port 22 is used for the internal firmware repository, but you can change this port number during the installation process. If the port is unavailable, you must use an external FTP, SCP, or SFTP server for switch supportsave and firmware download functions. For switches that are running FOS versions earlier than 8.2.2, you must always use an external FTP, SCP, or SFTP server for switch supportsave and firmware download functions if you change to a port number other than 22. To free port 22 for SANnav Management Portal, complete the following steps: 1. Edit the <code>/etc/ssh/sshd_config</code> file: a. Locate the following line: <pre>#port 22</pre> b. Uncomment the line and change the port number to another, unused port, such as 6022: <pre>port 6022</pre> Note that whatever port you select must be available and allowed in the firewall. 2. Restart the SSHD using the following command: <pre>systemctl restart sshd</pre> The current SSH session remains logged in, but any new sessions must now use port 6022.
Check port 80 availability	Port 80 must be available if you allow HTTP port 80 to HTTPS redirection; otherwise, the installation fails. After installation, port 80 must continue to be available always; otherwise, you cannot start (or restart) SANnav.
Check additional port requirements	For more information about other ports that must be open, see 11.2.12, “Installing SANnav Management Portal” on page 283.

11.2.12 Installing SANnav Management Portal

In this section, we describe how to download and install SANnav Management Portal on the server.

Before you extract the installation file, be sure to review and comply with the system and server requirements and the installation prerequisites that are listed in the following sections:

- System and Server Requirements for SANnav Management Portal
- Installation Prerequisites

Note: If the scripts fail during the installation or start, uninstall SANnav, restart the server, and then reinstall SANnav. Do not attempt to fix the issue and rerun the installation script without first uninstalling the application.

Download and copy the SANnav Management Portal software package to the server. The package contains the SANnav Management Portal .tar file.

Complete the following steps to download and install SANnav Management Portal on the server:

1. Download the SANnav Management Portal .tar file (for example, Portal_<version>-distribution.tar.gz) to the folder where you want to install the application.

Note: Do not create the SANnav Management Portal installation folder with a space in the name; otherwise, the installation fails.

2. Extract the .gz file to extract the file to the current location. tar -xvzf Portal_<version>-distribution.tar.gz This step creates a directory with a name similar to Portal_<version>_bldxx. (This directory is referred to as the <install_home> directory in this document.)
3. Go to the <install_home>/bin directory:

```
[root@RHEL7-10-100 home]# cd Portal_<version>_bldxx/bin
```
4. Run the install-sannav.sh script to install SANnav Management Portal:

```
[root@RHEL7-10-100 bin]# ./install-sannav.sh
```

If an earlier instance of SANnav Management Portal is installed, the installation script prompts whether you want to continue with migration or exit the installation.
5. If you are prompted about migrating SANnav, enter one of the following options:
 - To proceed with migration, press Enter. You are prompted to enter the location of the existing SANnav installation.
 - To exit the installation, press Ctrl+C. The script ends. At this point, you can back up the current SANnav instance and restart the installation script. You also can uninstall the current SANnav instance, and restart the installation script without migrating.

As the installation proceeds, the script runs a preinstallation requirements test. If any test fails, the installation exits with error messages. You must fix the reported issues, uninstall the application, and repeat the process beginning with step 1. After the diagnostics pass, installation of SANnav Management Portal software continues. On successful installation of the software, the SANnav Management Portal server starts, which can take up to 20 minutes.

11.2.13 Firewall requirements for SANnav Management Portal

If your network uses a firewall between the SANnav Management Portal client and the server or between the server and the SAN, a set of ports must be open in the firewall to ensure suitable communication.

The table that is shown in Figure 11-25 lists the ports that must be open in the firewall. These ports are added to the IP tables by default when the SANnav server is running; therefore, you do not need to open them in the firewall if it is enabled and running on the SANnav server.

Port Number	Transport	Inbound/ Outbound	Communication Path	Description
22	TCP	Both	Client --> Server Server <--> Switch	Internal SSH server
80	TCP	Both	Client --> Server Server --> Switch	HTTP port for access from browser to server HTTP port for access from server to switch
161	UDP	Outbound	Server --> Switch	SNMP port
162	UDP	Inbound	Switch --> Server	SNMP trap port
443	TCP	Both	Client --> Server Server --> Switch Server --> vCenter	HTTPS port for secure access from browser to server HTTPS port for secure access from server to switch HTTPS port for secure access from server to vCenter
514	UDP	Inbound	Switch --> Server	Syslog port
6514	UDP	Inbound	Switch --> Server	Secure syslog port
8081	TCP	Inbound	Switch --> Server	Avro schema registry port
19094	TCP	Inbound	Switch --> Server	Secured Kafka port
19095	TCP	Inbound	Switch --> Server	Secured Kafka port

Figure 11-25 Ports that must be open in the firewall

If `firewalld` is enabled, you must add the SSH service to the trusted zone in `firewalld` for the firmware download feature to work.

For more information about how to configure `firewalld`, see “Configuring a firewall for SANnav” on page 287.

If you configure an external authentication server (LDAP, RADIUS, or TACACS+) or an email server (SMTP), ensure that the SANnav Management Portal server can access the ports that are listed in the table in Figure 11-26. The default ports that are listed can change the default.

Port Number	Transport	Inbound/ Outbound	Communication Path	Description
25	TCP	Outbound	Server --> SMTP Server	SMTP server port for email communication if you use email notifications without SSL or TLS
49	TCP	Outbound	Server --> TACACS+ Server	TACACS+ server port for authentication if you use TACACS+ for external authentication
389	TCP	Outbound	Server --> LDAP Server	LDAP server port for authentication if you use LDAP for external authentication and SSL is not enabled
465	TCP	Outbound	Server --> SMTP Server	SMTP server port for email communication if you use email notifications with SSL
587	TCP	Outbound	Server --> SMTP Server	SMTP server port for email communication if you use email notifications with TLS
636	TCP	Outbound	Server --> LDAP Server	LDAP server port for authentication if you use LDAP for external authentication and SSL is enabled
1812	UDP	Outbound	Server --> RADIUS Server	RADIUS server port for authentication if you use RADIUS for external authentication

Figure 11-26 Ports that the server must be able to access

11.2.14 SANnav Management Console commands

The sannav-management-console.sh script allows you to perform several actions on the SANnav server without the need to run individual scripts.

By using this one script, you can perform the following actions:

- ▶ Check SANnav status.
- ▶ Restart SANnav.
- ▶ Stop SANnav.
- ▶ Start SANnav.
- ▶ Uninstall SANnav.
- ▶ Show the SANnav port and server configuration. Go to the <install_home>/bin folder and run the following script:

```
./sannav-management-console.sh
```

You are presented with a list of options from which to choose.

11.2.15 Checking the server health

After the installation is complete, you can check the health of the SANnav server by using the check-sannav-status.sh script. Any service that is down is listed in the script output.

To check the health of the server, go to the <install_home>/bin folder and run the ./check-sannav-status.sh script.

Note: Any service that is down while the server health status is checked is automatically started by systemmonitor within 20 minutes.

The following example output is from a healthy server:

```
-bash-4.2# sh ./check-sannav-status.sh
SANnav server is healthy.
All the services are currently in running state.
```

The following example is output from an unhealthy server:

```
-bash-4.2# sh ./check-sannav-status.sh
Following services are currently down or starting filters-middleware
topology-middleware SANnav Disk Usage Alert monitors the disk space used for every
ten minutes.
```

There are three threshold levels: 70%, 80%, and 90%. An event is sent when the threshold levels exceed or drop below the defined level. The following event severity thresholds are used:

- ▶ Warning: 70%
- ▶ Error: 80%
- ▶ Critical: 90%

Changing the self-signed certificates for client and server communication

You can replace the SSL self-signed certificates with third-party signed certificates. Make sure that the SSL certificate and key files are copied to this host or VM.

Go to the `<install_home>/bin` folder and run the following script:

```
./replace-server-certificates.sh
```

When you run this script, SANnav is automatically restarted for the new certificates to take effect. After the server is back up, you must rediscover or stop monitoring and then monitor all switches that are registered for telemetry data; otherwise, the new certificates do not take effect.

Configuring a firewall for SANnav

Perform the following steps to set up a firewall by using **firewalld**. This example uses Red Hat Enterprise Linux (RHEL).

1. Start the firewall by using the **sudo systemctl start firewalld** command.
2. Check that the firewall is running by using the **sudo systemctl status firewalld** command.
3. Enable the firewall automatically after a system restart by using the **sudo systemctl enable firewalld** command.
4. Add the SSH service to the trusted zone:

```
sudo firewall-cmd --zone=public --permanent --add-service=ssh
```

If any other default ports are customized, add the services for those ports as well. For example, if you are using HTTPS port 443, enter the following command:

```
sudo firewall-cmd --zone=public --permanent --add-service=https
```

5. Add ports by using the following commands. In these commands, public is the default zone. If your default zone is different, use your default zone for the ports:

```
sudo firewall-cmd --zone=public --add-port=2377/tcp --permanent  
sudo firewall-cmd --zone=public --add-port=7946/tcp --permanent  
sudo firewall-cmd --zone=public --add-port=7946/udp --permanent  
sudo firewall-cmd --zone=public --add-port=4789/udp --permanent
```
6. If you are installing on CentOS or RHEL, version 8.0 or higher, enable masquerading on the firewall:

```
sudo firewall-cmd --zone=public --add-masquerade --permanent
```
7. Associate the interface (if this is not done) with the default profile:

```
sudo firewall-cmd --permanent --zone=public --change-interface=<interface_name>
```
8. After the ports are added, use the **sudo firewall-cmd --reload** command to reload the firewall configuration.
9. Verify whether the configuration is correct by using the **sudo firewall-cmd --list-all** command.

11.2.16 Starting SANnav Management Portal

Complete the following steps to start SANnav Management Portal to monitor and manage your fabrics:

1. Open your browser and enter the IP address or fully qualified domain name (FQDN) of the SANnav Management Portal server. You can use http or https; for example:
`http://192.0.2.0` or `https://192.0.2.0`.

The SANnav Management Portal login window opens (see Figure 11-27).

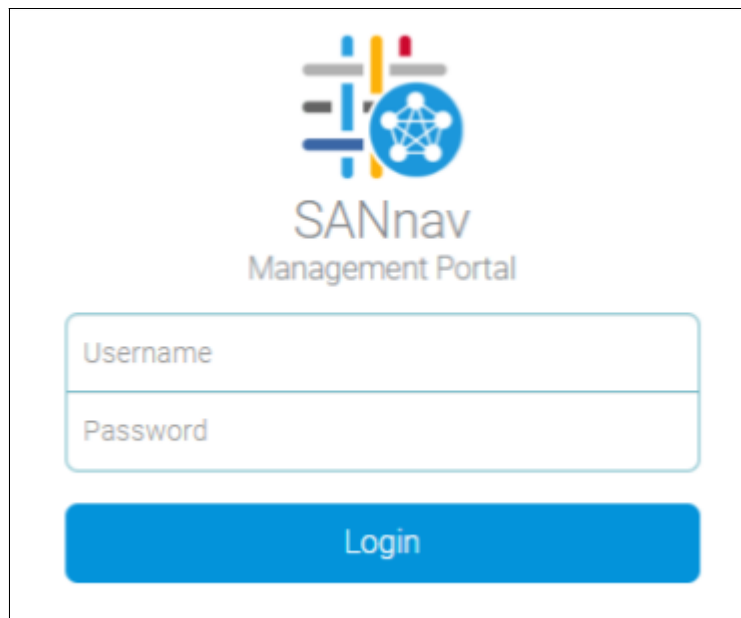


Figure 11-27 SANnav Management Portal - Login

If the login window does not open, try the following fixes:

- Ping the SANnav server to ensure that it is up.
- Check whether SANnav services are running. For more information, see 11.2.15, “Checking the server health” on page 286.
- Check the firewall settings.

2. Enter your SANnav username and password. Click **Login**.

For the first SANnav log in, the default username is Administrator and the default password is password.

SANnav starts with the default dashboard displayed. If SANnav displays the message Login Failed. Service is not available at this time instead of starting, SANnav is in the process of starting. Wait a few minutes and attempt to log in again. The first time that you log in, change the default password.

11.2.17 Changing your password

The first time that you log in to SANnav, change your password. After you change your password, you are automatically logged out, and you must log in again by using the new password.

Complete the following steps:

1. Click the user icon in the upper right of the window, and then, click **User Preferences** (see Figure 11-28).

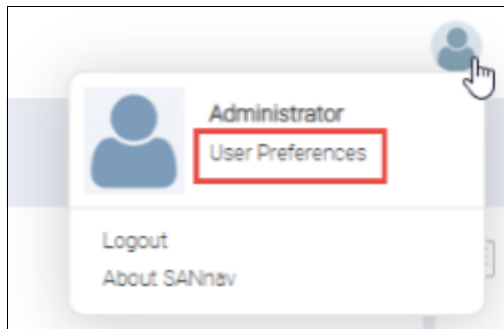


Figure 11-28 User Preferences

2. Click **Edit** next to the “Logging in” text (see Figure 11-29).

The image shows the 'Preferences' page in the SANnav Management Portal. The page has a navigation bar at the top with links: Dashboard & Reports, Topology, Inventory, Events, Zoning, and SANnav. Below the navigation bar is a section titled 'Preferences'. Under 'Personal Info', there is an 'Edit' button. Below that, the user's details are listed: Username: Administrator, Phone Number: -, Time Zone: PDT (America/Los_Angeles). Below this is the 'Logging in' section, which is highlighted with a red box. It contains an 'Edit' button. Below the 'Logging in' section, there are two more items: Inactive Logout Time: 30 Minutes and Password: Last Updated On Mar 29, 2019.

Figure 11-29 Logging in

3. Click **Change** on the Logging in page.
4. Complete the Change Password fields and then, click **OK**. You are automatically logged out of SANnav.
5. Log in to SANnav again by using your new password.

11.2.18 User interface overview

By using the SANnav Management Portal, you can manage one or more SAN fabrics in multiple locations. After you are familiar with the basic components of SANnav, you can quickly start monitoring and managing your fabrics.

The basic layout of the SANnav user interface is shown in Figure 11-30.

The image shows the main dashboard of the SANnav Management Portal. The dashboard has a navigation bar at the top with links: Dashboard & Reports, Topology, Inventory, Events, Zoning, and SANnav. Below the navigation bar is a section titled 'Switches (12)'. Below this section is a table with columns: Name, FID, Fabric, Health, Model, Firmwa..., and Port Count. The table contains four rows of data. The components are highlighted by red boxes and numbers: 1. Navigation bar, 2. Search icon, 3. User profile icon, 4. Switches (12) section, 5. Add button (+), 6. Table of switches.

Name	FID	Fabric	Health	Model	Firmwa...	Port Count
G610	-	TC G610	Healthy	Brocade G610	v8.2.1	24
G620_LS1	10	G620-LF1	Healthy	Brocade G620	v8.2.1	8
G620_LS2	20	G620-LF2	Healthy	Brocade G620	v8.2.1	8
S1	10	Untitled New ...	Degraded	Brocade X6-4	v8.2.1	-

Figure 11-30 SANnav Management Portal User Interface

The following components are shown in Figure 11-30 on page 290:

1. Navigation bar: Contains links to feature pages. The SANnav link displays the Configurations and Settings page.
2. Search: Click the icon to display the search box, select a context on which to search from the list, and enter the search term in the field.
3. Profile menu: Click the icon to display more links for changing user preferences, displaying the SANnav version, and logging out.
4. Subnavigation bar: Provides the page title and optional item count within parentheses. Also includes buttons and menus to take actions within the page. The subnavigation bar is the main way to navigate within a feature.
5. Filter bar: Allows you to filter the display based on columns, fabrics, network scope, and customized filters.
6. Expandable sidebar: Provides an area where you can save selected inventory items for investigation later.

Detail pages

Clicking a fabric name, switch name, or port name in a table opens a detail page for that object (see Figure 11-31). The detail page displays more information about the object and can includes tasks that you can perform.

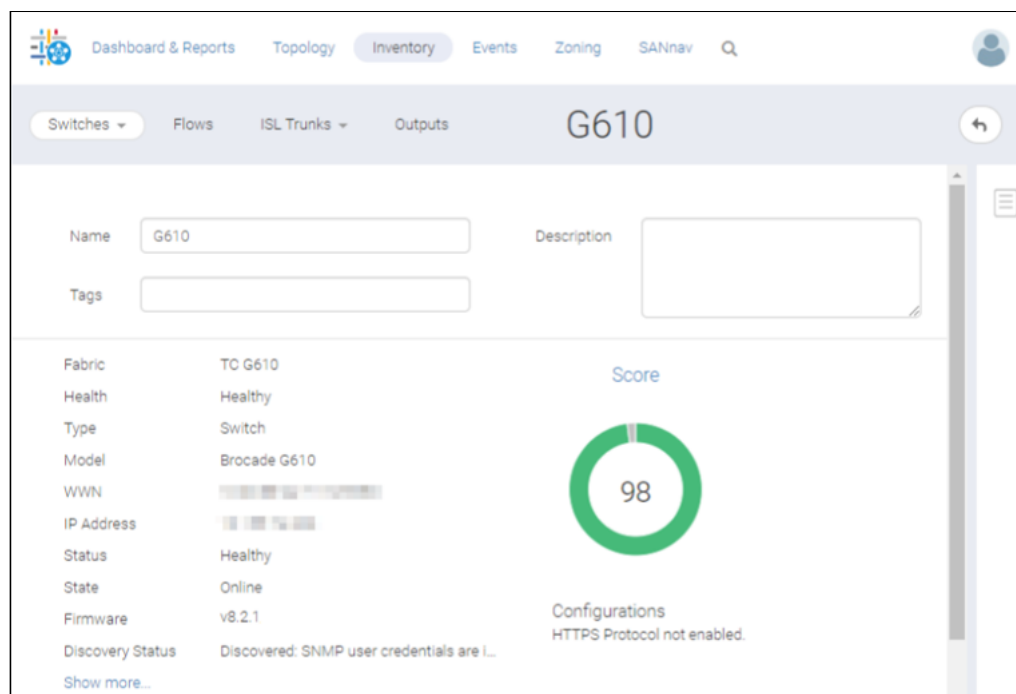


Figure 11-31 Detail page for a switch

The detail page is different, depending on the context. For example, the detail page for a fabric on the Inventory page is different from the detail page for the same fabric on the Discovery page.

Tables

In tables, if an entry is truncated (indicated by an ellipsis at the end of the entry), hover over the entry to see the full text (see Figure 11-32).

Name ^	Type ^	Description ^	Created ... ^	Sharing ^
Health Sum...	Dashboard	Health Summary	(System)	Shared
Network Po...	Dashboard	Monitors each port and ...	(System)	Shared
WAN Vision		Monitors each port and provides meaningful summary of their congestion states		Shared

Figure 11-32 Hovering the mouse over truncated table text

Many tables include an action menu that you can access by clicking the down arrow in the rightmost column. Click this arrow to display more tasks that you can perform on the associated object (see Figure 11-33).

Firmware ^	Port Count	
v8.2.0	8	▼
v8.2.0	8	▼
v8.2.0		View
v8.2.0		Select
v8.2.0		Investigate
v8.2.0		Show in Topology
v8.2.0		Open in WebTools
v8.2.0	29	▼

Figure 11-33 Using the down arrow to display more actions

On the right side of the subnavigation bar, the More button often contains a Bulk Select option (see Figure 11-34), which allows you to select and perform operations on one or more items at the same time. For example, you can turn on monitoring for multiple fabrics, or you can create a tag and apply it to multiple switches.

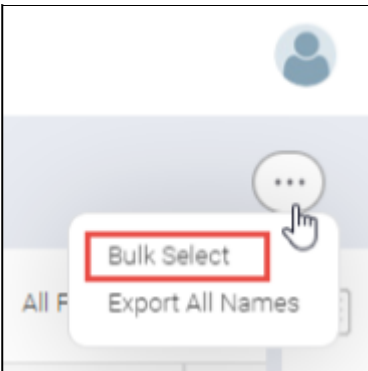


Figure 11-34 Bulk Select option

11.2.19 Initial setup and configuration

If you are familiar with SAN network management, the following tasks might need to be performed directly after the SANnav installation:

- Discover fabrics.

SANnav must connect to a seed switch per fabric to discover the items in the fabric. For more information, see *Discovering a Fabric within the Brocade SANnav Management Portal User Guide*.

- Configure service notifications.

Configure the SMTP mail server and Call Home, register the SANnav server as a syslog and SNMP trap recipient, and forward SNMP traps and syslog messages. For more information, see *Call Home and ESRS and Event Management within the Brocade SANnav Management Portal User Guide*.

- Set up user accounts.

Add users, roles, and areas of responsibility (AORs). Configure a password policy. Set up external authentication. See Security within the Brocade SANnav Management Portal User.

- Configure MAPS policies.

You can create a policy or import policies from a JSON or XML file. For example, you can import MAPS policies that were exported from Brocade Network Advisor.

- Create custom dashboards.

SANnav provides several default dashboards, but you can create custom dashboards to fit your specific needs.

For more information about Brocade SANnav setup, configuration, and navigation, see [this web page](#).

11.3 IBM SANnav Global View 2.1 release overview

SANnav Global View is a higher-level “global” management application that provides a comprehensive view of a SAN environment that spans multiple SANnav Management Portal instances. By using SANnav Global View, you can navigate seamlessly across multiple SANnav Management Portal instances and drill down to any individual instance to perform detailed monitoring, investigation, and troubleshooting tasks.

Note: SANnav Global View is a separate product from SANnav Management Portal, and it requires separate installation and licensing. You log in to the SANnav Global View application and add portals to SANnav Global View, which then uses information in the portals to build a global view of the dashboard and inventory.

SANnav Global View is compatible only with SANnav Management Portal instances that are running the same version as Global View. Older versions of SANnav Management Portal are not supported.

11.3.1 SANnav Global View browser requirements

Any notebook or machine that starts web applications can be used to start SANnav Global View. For optimal performance, have at least 16 GB memory available.

The following browsers can be used to access SANnav Global View:

- ▶ Chrome (Google)
- ▶ Firefox (Mozilla). If you access the client from the Remote Desktop, the performance of the user interface might be degraded.

11.3.2 SANnav Global View system and server requirements

All the system and server requirements must be met before you start installing SANnav Global View.

The table that is shown in Figure 11-35 lists the system and server requirements for SANnav Global View.

Note: The disk space requirements that are listed in the table is for SANnav Global View only. Be sure to account for more space that is required by the operating system and for saving files.

Requirement	Value
Operating system	• Red Hat Enterprise Linux: 7.7, 8.0, and 8.1 • CentOS 7.7, 8.0, and 8.1
Host type	• Bare metal server • VMware ESXi virtual machine
CPU	16 cores
Memory (RAM)	32 GB
Hard disk	450 GB, distributed as follows: • 300 GB - Installation folder • 90 GB - /var/lib • 20 GB - "/" directory (swap space) The disk space can be from a direct-attached disk or through a network-mounted disk.
Minimum number of CPU sockets	2
Minimum CPU speed	2000 MHz

Figure 11-35 SANnav Global View single-node deployment

11.3.3 SANnav Global View installation prerequisites

Review all SANnav Global View installation prerequisites that are shown in Figure 11-36 *before* you extract the installation files.

Note: Use the latest generation processors for better SANnav performance.

Task	Task Details or Additional Information
Gather necessary information and components.	Make sure that you have the following: • Root user credentials. You must log in to the SANnav server as the root user, or a user with root privilege. • The SANnav Global View server IP address.
Uninstall other applications.	SANnav Global View is expected to be installed and run on a dedicated host; if any other application is installed on the host, uninstall it before starting the SANnav Global View installation. If you are migrating SANnav, do not uninstall the current SANnav instance.
Uninstall Docker, if already installed.	The SANnav installation installs Docker. If you have a Docker installed other than the Docker that SANnav installs, you must remove it before starting the installation.
Check operating system requirements.	• Ensure that the entire physical server (boot, log, and data) runs on a single partition. • Ensure that the operating system can be loaded through a bootable disk or through a PXE server. • Ensure that the <code>lsuf</code> and <code>nslookup</code> packages are installed on the operating system machine.
Format the XFS file system.	If you are using XFS as the file system, make sure that you set <code>d_type=true</code> while creating the disk. You can verify this by running the command <code>xfs_info <docker-installation-directory></code> and verify that <code>ftype=1</code> . The default docker installation directory is <code>/var/lib</code> .
Set umask and ulimit.	• "umask" for the root user must be set to 0022. Enter the following command to set the umask: <pre>umask 0022</pre>You must set the umask before you unzip the installation files. If you extracted the installation files before you set the umask, you must delete the installation folder, run <code>umask 0022</code>, and then unzip the files again. • Ensure that the ulimit is set correctly. To set the ulimit, edit the <code>/etc/security/limits.conf</code> file with appropriate privileges.
Check port availability.	Port 80 must be available if you allow HTTP port 80 to HTTPS redirection; otherwise, installation fails. After installation, port 80 must continue to be available all the time; otherwise, you cannot start (or restart) SANnav Global View. If your network utilizes a firewall, there may be other ports that must be open. See the Firewall Requirements section for details.
Ensure that IPv4 IP forwarding is enabled.	To check whether IPv4 IP forwarding is enabled, enter the following command: <pre>/sbin/sysctl net.ipv4.ip_forward</pre> If the output is <code>net.ipv4.ip_forward=1</code> , forwarding is enabled, and you do not need to make any changes. If the output is <code>net.ipv4.ip_forward=0</code> , forwarding is disabled, and you must change it as follows: 1. Enter the following command to set IP forwarding for this session: <pre>/sbin/sysctl -w net.ipv4.ip_forward=1</pre> 2. Edit the <code>/etc/sysctl.conf</code> file and add the following lines: <pre># IP Forwarding is enabled for Broadcom SANnav net.ipv4.ip_forward = 1</pre>
Run additional commands.	• Ensure that the <code>hostname -i</code> command resolves to a valid IPv4 address. • The <code>nslookup</code> command must be successful for the host name of the physical host and VM.

Figure 11-36 SANnav Global View installation prerequisites

11.3.4 Installing SANnav Global View

In this section, we describe how to install SANnav Global View on the server.

Before extracting the installation file, be sure to review and comply with the following system and server requirements and installation prerequisites:

- ▶ 11.3.2, “SANnav Global View system and server requirements” on page 294
- ▶ 11.3.3, “SANnav Global View installation prerequisites” on page 295

Download and copy the SANnav Global View software package to the server. The package contains the SANnav Global View .tar file. Then, complete the following steps:

1. Download the SANnav Global View .tar file (for example, `Global_2.1.0_rc_bldxx-distribution.tar.gz`) to the folder where you want to install the application.

Note: Do not create the SANnav Global View installation folder with spaces in the name; otherwise, the installation fails.

2. Extract the .gz file to extract the file to the current location:

```
tar -xvzf Global_2.1.0_rc_bldxx-distribution.tar.gz
```

This step creates a directory with a name that is similar to `Global_2.1.0_rc_bldxx`. This directory is referred to as the `<install_home>` directory in this publication.

3. Go to the `<install_home>/bin` directory.
4. Run the `install-sannav.sh` script to install SANnav Global View.

The `./install-sannav.sh` installation script checks whether an earlier instance of SANnav Global View is installed, and if so, it prompts you to decide whether want to exit the installation and take a backup or continue the installation without backing up the server.

5. If you are prompted about migrating SANnav, enter one of the following options:
 - Press `Ctrl+C` to exit the installation procedure. Back up the server or uninstall the existing SANnav instance. Then, restart the installation script.
 - Press `Enter` to proceed with migration. You are prompted to enter the location of the existing SANnav installation.

As the installation proceeds, the script runs a pre-installation requirements test. If the test fails, the installation exits with error messages. You must fix the reported issues and re-run the installation script. After the diagnostics pass, the installation of SANnav Global View software continues. On successful installation of the software, the SANnav Global View server starts. The startup might take up to 10 minutes.

11.3.5 Logging in to SANnav Global View

To log in to SANnav Global View, complete the following steps:

1. Open your browser and enter the IP address or fully qualified domain name (FQDN) of the SANnav Global View server.

You can use HTTP or HTTPS, for example: `http://192.0.2.0` or `https://192.0.2.0`. The SANnav Global View login window opens (see Figure 11-37).

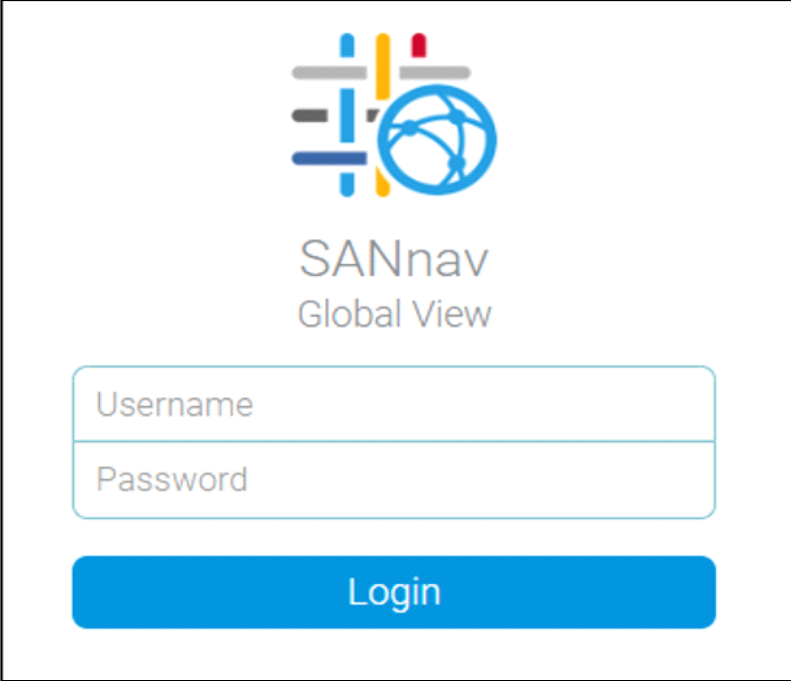
The image shows the SANnav Global View login window. At the top center is the SANnav logo, which consists of a stylized globe with blue and yellow lines, and the text "SANnav Global View" below it. Below the logo are two input fields: "Username" and "Password". Below these fields is a blue button labeled "Login".

Figure 11-37 SANnav Global View Login window

If the log in window does not appear, try the following fixes:

- Ping the SANnav server to ensure that it is running.
- Check whether SANnav services are running. For more information, see 11.2.15, “Checking the server health” on page 286.
- Check the firewall settings.

2. Enter your SANnav username and password and then, click **Login**.

For the first SANnav login, the default username is Administrator and the default password is password. SANnav starts with the default dashboard displayed. If SANnav displays the message Login Failed. Service is not available at this time. instead of starting, SANnav is in the process of start. Wait a few minutes and attempt to log in again.

11.3.6 SANnav Global View overview

After you are familiar with the basic components of SANnav Global View, you can quickly start monitoring SANnav Management Portal instances. When you first log in to SANnav Global View, you see the Summary dashboard, the single dashboard that is provided with Global View. The basic layout of the SANnav Global View user interface is shown in Figure 11-38.

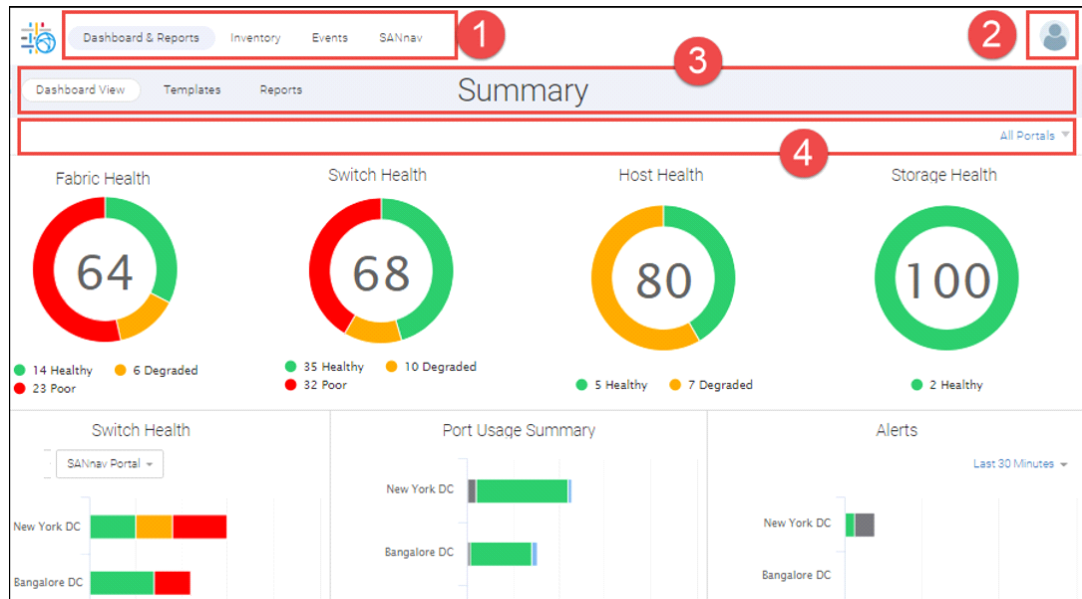


Figure 11-38 Summary dashboard

The following components are shown in Figure 11-38:

1. **Navigation bar:** Contains links to feature pages. The SANnav link displays the page where various settings and configurations can be performed.
2. **Profile menu:** Displays links for changing user preferences, displaying the SANnav version, and logging out.
3. **Subnavigation bar:** Provides the page title and optional item count within parentheses. Also includes buttons and menus to take actions within the page.
4. **Filter bar:** Allows you to filter the display based on columns, portal scope, and customized filters.

Click **Inventory** to display inventory information about fabrics, switches, switch ports, chassis, host and storage devices, and host and storage ports across all Management Portal instances (see Figure 11-39 on page 299).

Name	Type	FID	Fabric	Health	Model	Port Count	Portal
A__15643...	Switch	128	FabricTest...	Poor	Brocade X6-8	221	Munich DC
AG_17-mo...	Access Gate...	-	FCIP-104	Healthy	Brocade 6505	24	San Jose DC
AG_17-mo...	Access Gate...	-	Untitled N...	Healthy	Brocade 6505	24	Untitled New Portal

Figure 11-39 Inventory

Click **Events** to show all application events that stem from a user or system action (see Figure 11-40). A system action is triggered under specific situations, such as when a portal is disconnected. You can filter the events list by date range. Events from SANnav Management Portal instances are not listed here.

Description	Module	Category	Username	Occurred Time
Successfully authenti...	Authentication	User Action Event	-	Aug 21, 2019 11:19:37 PDT
Successfully authenti...	Authentication	User Action Event	-	Aug 21, 2019 11:15:27 PDT
Filter - Switches runni...	Filter Mgmt	User Action Event	Administrator	Aug 21, 2019 08:45:47 PDT

Figure 11-40 Events

Click **SANnav** to discover Management Portals, manage filters, and select various configuration settings across several feature categories. The profile menu is where you can add your phone number and change your password. It is also where you can enable Persist Last Column Selection, wherein the system remembers your column customization and FICON Display.

Column customization is the ability to choose which columns you want to see in a tabular view. For example, only a few of the many available columns in inventory are shown by default, but you can customize the table by adding or removing columns. When FICON Display is enabled, specific columns in the switch and switch port inventory pages are rearranged for FICON mode.

The Administrator User Preferences window is shown in Figure 11-41.

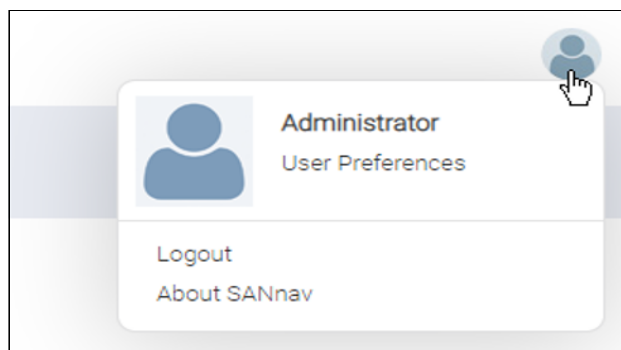


Figure 11-41 Administrator User Preferences

The profile menu is where you can add your phone number and change your password. It is also where you can enable Persist Last Column Selection, wherein the system remembers your column customization, and FICON Display.

Column customization is the ability to choose what columns you want to see in a tabular view. For example, by default, only a few of the many available columns in inventory are shown, but you can customize the table by adding or removing columns.

When FICON Display is enabled, certain columns in the switch and switch port inventory pages are rearranged for FICON mode.

11.3.7 Generating a license for SANnav Global View

Access the Broadcom licensing portal to generate a SANnav license key. Before you generate the license, make sure that you have a server unique ID (UID). After you install SANnav, you can obtain the server UID from the SANnav Licensing page.

Complete the following steps to generate and obtain a SANnav license key:

1. Obtain a transaction key from your SANnav vendor. You receive an email with the license transaction key in the form of an electronic transaction key.
Do not discard the email with the electronic key. Keep it in a safe place in case it is needed for technical support or product replacement.
2. Log in to [myBroadcom](#) to complete the software license request. If you do not have a login ID and password, request access by following the online registration instructions.

3. Click **License Management** in the Brocade Products section (see Figure 11-42).

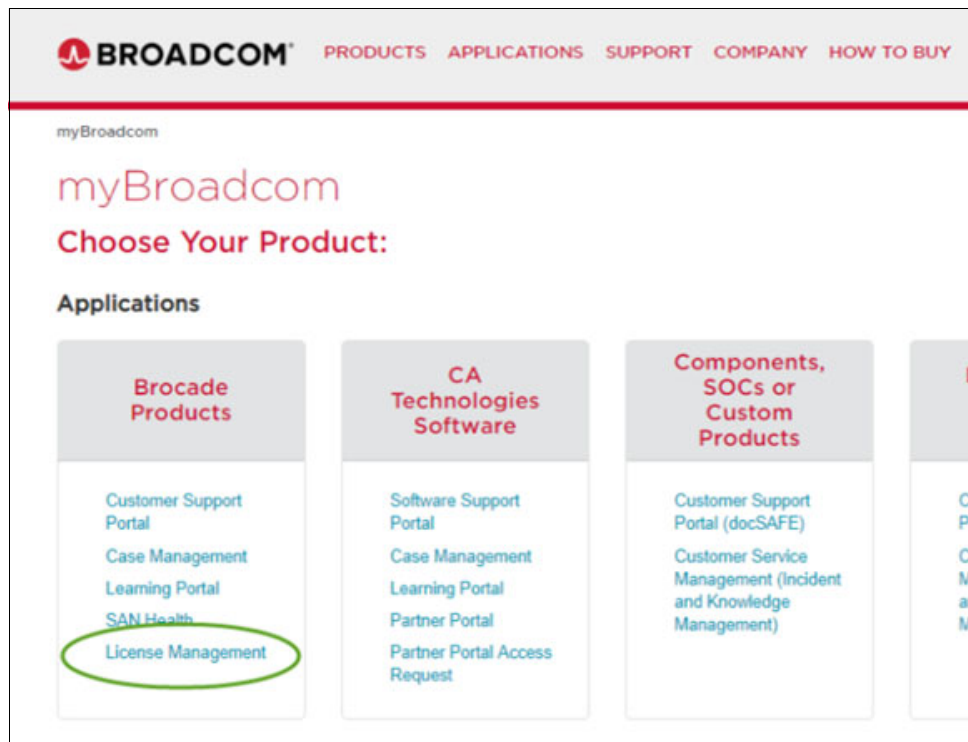


Figure 11-42 License Management option

4. Enter the licensing transaction key or rehost key in the License Generation window and click **Next** (see Figure 11-43).

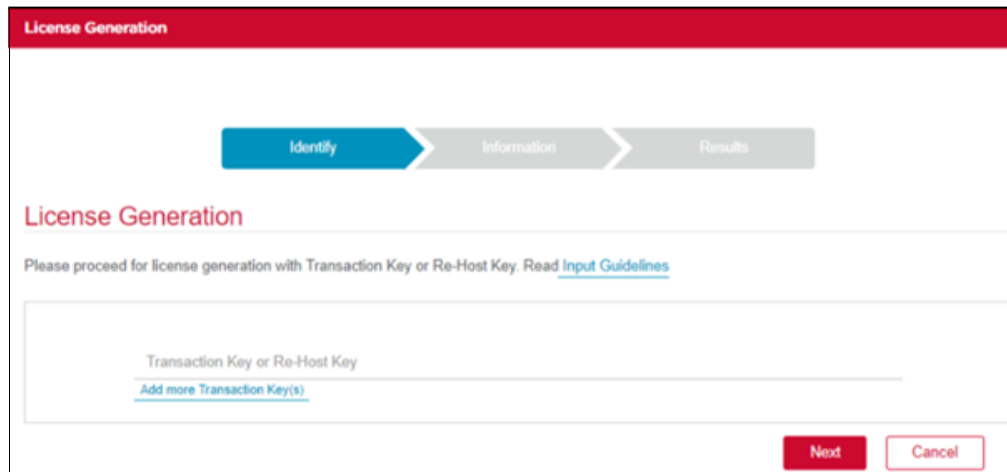


Figure 11-43 License Generation window

5. In the Product Information area, enter the server UID that you obtained from SANnav (see Figure 11-44).

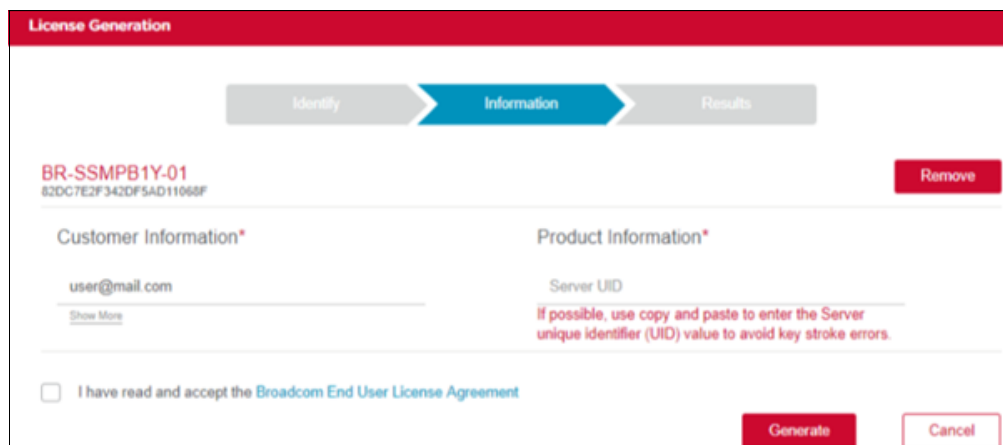


Figure 11-44 Entering the UID

6. Read the Broadcom End User License Agreement. If you agree to the terms, select **I have read and accept Broadcom End User License Agreement**.
7. Click **Generate**. The Results window displays an order summary and the results of the license request. Consider the following points:
- If the license request is successful, the License field contains a hyperlink to the generated license file. The license file is automatically emailed to the specified customer email address.
 - If the license request fails, the reason for failure and the action to take are displayed on the page.
8. Click the hyperlink in the License field to display the license key.
9. Copy the license key to a .txt file and save it (see Figure 11-45). You use this license key when you add the license to SANnav.

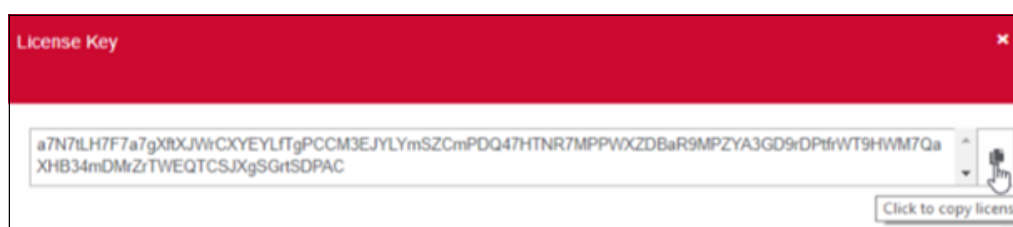


Figure 11-45 Copying license key

10. Click **Export to Excel** to export the results to a Microsoft Excel file, or click **Generate Another License** to generate a license.

11.3.8 Global dashboard overview

The SANnav Global View dashboard provides comprehensive “global” visibility across multiple SANnav Management Portal instances. The Global View dashboard is auto-refreshed every 15 minutes.

The Global View dashboard provides the following widgets:

- ▶ Health summary
- ▶ Switch Health
- ▶ Port Usage Summary
- ▶ Alerts

These widgets indicate the status of the fabric, switch, host, and storage entities that are managed by the Portals that are added in SANnav Global View. The health summary widgets are categorized by health state: Healthy, Degraded, or Poor.

Each widget that is listed in Figure 11-46 are the exact number of entities in each health state. The health state for each type of object is determined by its health score, which is determined by a set of predefined factors. For example, a health score for fabrics is influenced by whether a link is down or redundant paths are missing.

Health State	Health Score
Healthy	Greater than 90
Degraded	Between 71 and 90
Poor	70 or less

Figure 11-46 Health states and scores

SANnav dashboards allow you to gather more information by drilling down into the dashboard widgets. By default, the scope of SANnav Global View is set to All Portals (that is, all Management Portal instances), as shown in Figure 11-47.

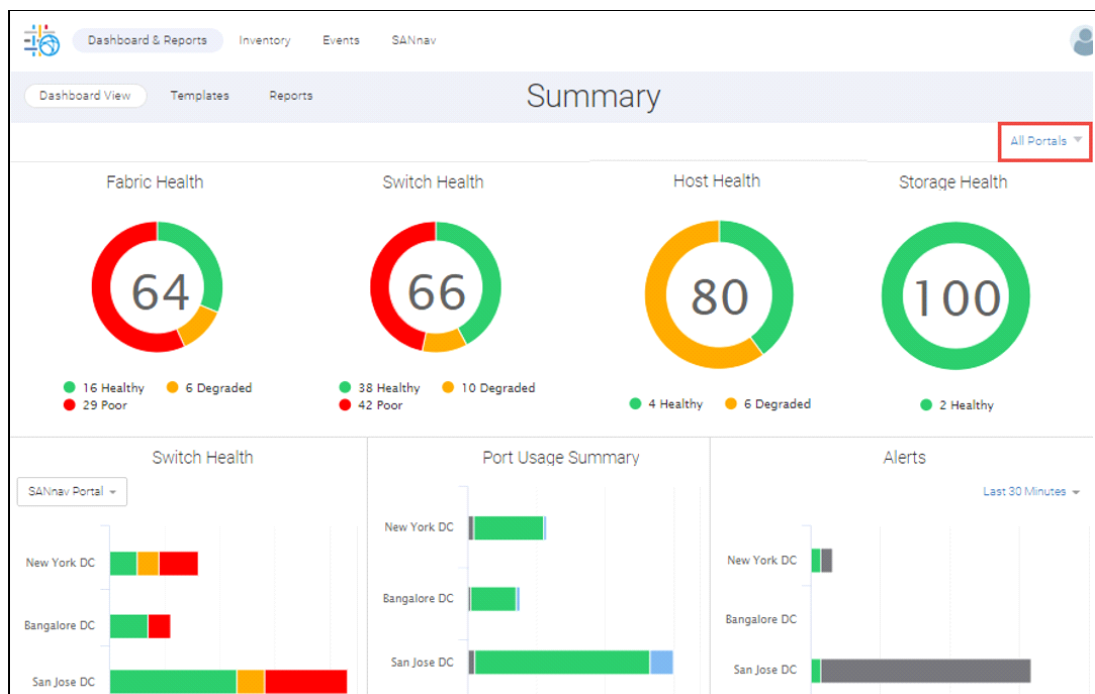


Figure 11-47 Summary

You can choose one or more Management Portal instances to change the scope.

11.3.9 Monitoring fabric health

The Fabric Health widget of SANnav Global View shows the number of fabrics in Healthy, Degraded, and Poor state. The least overall score across all objects in this group of fabrics is shown in the center of the circular widget.

Complete the following steps:

1. Click **Dashboard & Reports** in the navigation bar.
2. Hover over a colored portion of the Fabric Health widget to display details of the fabrics with that state. For example, hover over the red portion to display a list of portals and the number of fabrics in each portal having a poor state (see Figure 11-48).

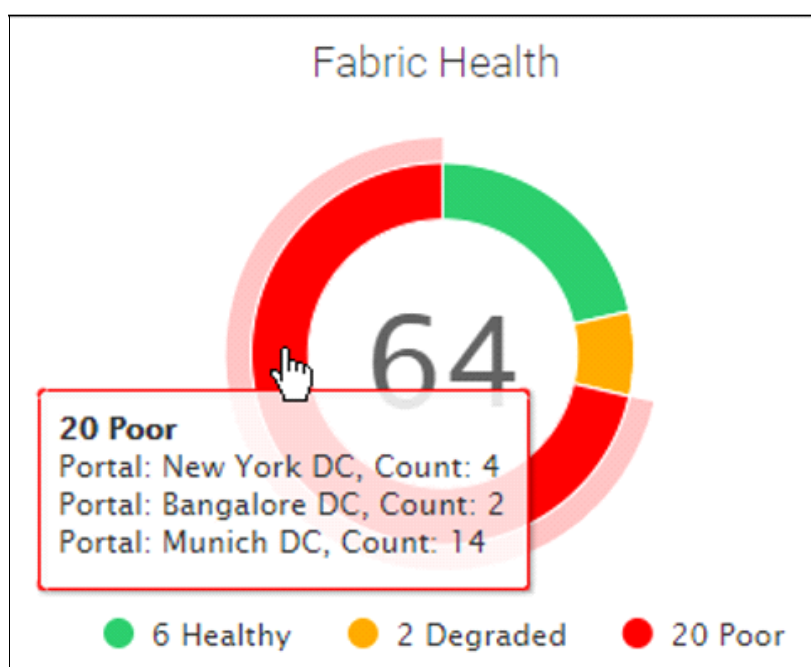


Figure 11-48 Fabric Health: 20 Poor

3. To view more details about the fabrics, click the colored segment of the widget. For example, clicking the red segment displays a table that shows the health scores for the fabrics in poor health.
4. Click **Show Details** in the drop-down menu of the fabric list (see Figure 11-49 on page 305) to see details that are associated with that health score. The Show Details option is available only if the score is less than 100.

Fabric Health: 20 Poor						
20 Items						
Name ▲	Score ◆	Switch Status ◆	Events ◆	Configurations ◆	Portal ◆	
Fabric_A	64	-30	-2	-4	New York DC	▼
FabricTest_27	70	-30	0	0	M	Show Details Show in SANnav Portal
FcoE	68	-30	0	-2	M	
FIDnine	70	-30	0	0	Munich DC	▼

Close

Figure 11-49 Show Details option

A pop-up window displays the factors that cause deductions in the health score for this fabric.

- Click **Back** to return to the list of portals.
- Select **Show in SANnav Portal** from the drop-down menu to start the Management Portal instance that is responsible for the data you are viewing. You might need to make sure that pop-ups are not blocked in your browser. If you are not logged in to this portal, credentials for this portal are required.

11.3.10 Monitoring switch health across Management Portal instances

The Switch Health widget displays details of switch health across Management Portal instances. You can get a summarized view and details of switches based on firmware version, Management Portal instance, switch model, and product category.

Complete the following steps to use the Global View dashboard to display switch health across multiple Management Portal instances:

1. Click **Dashboard & Reports** in the navigation bar. In the bottom row of the dashboard, the Switch Health widget displays (see Figure 11-50). In this window, you see switches that are grouped by SANnav Portal instances. You can select a different grouping from the drop-down list.

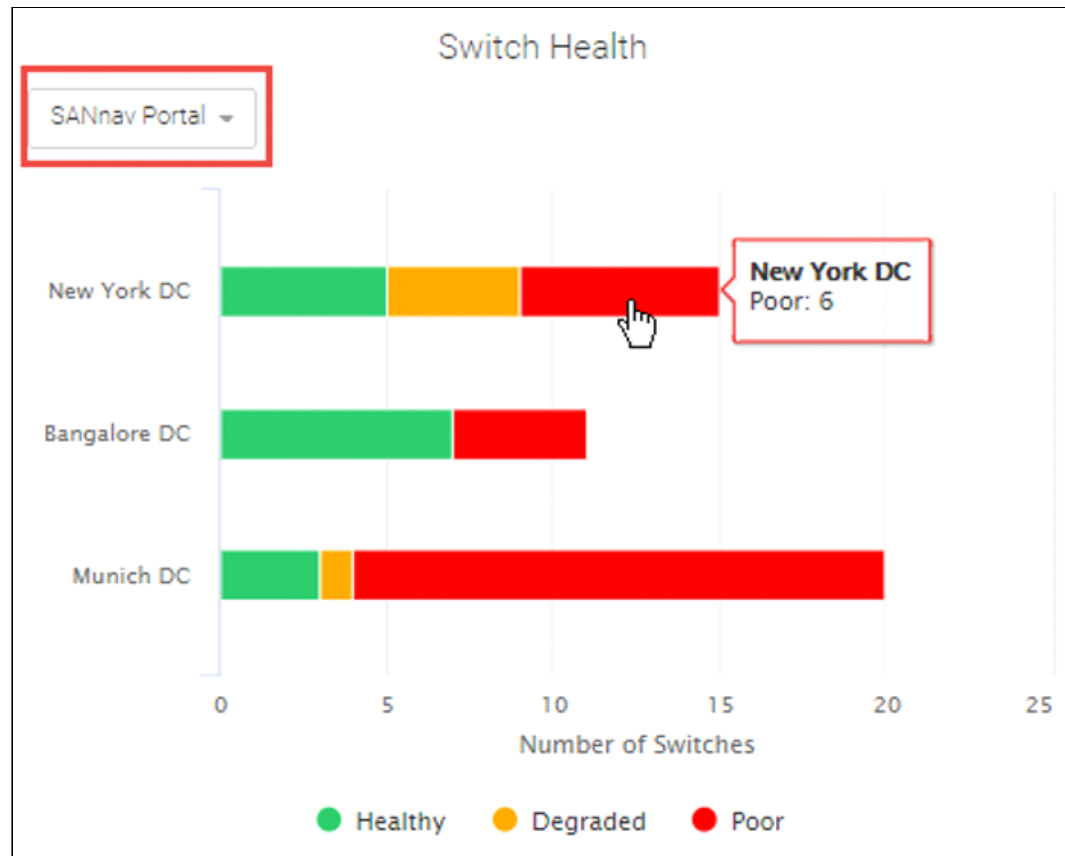


Figure 11-50 Switch Health: SANnav Portal

2. Hover over the graph to display details (see Figure 11-51). For the Firmware Version and Model categories, more bars might be displayed in the graph than are listed in the row names on the left. In this case, hover over an intermediate bar in the graph to display the details.

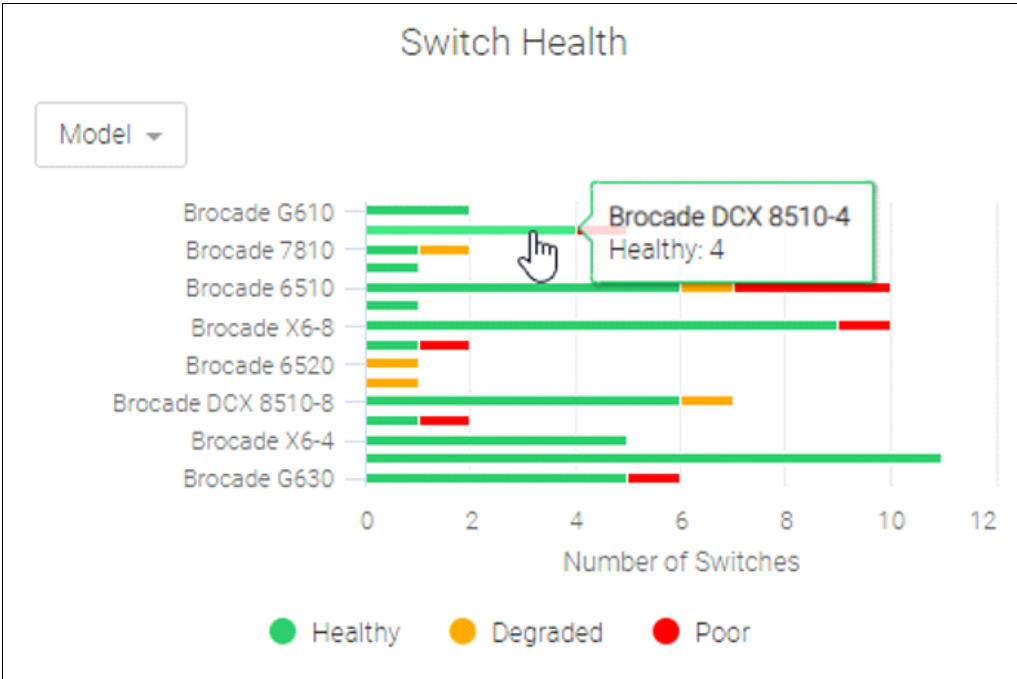


Figure 11-51 Switch Health: Model

3. To see the details of all switches in a specific category, click the bar chart of the widget, and then, select **Show Details** (see Figure 11-52).

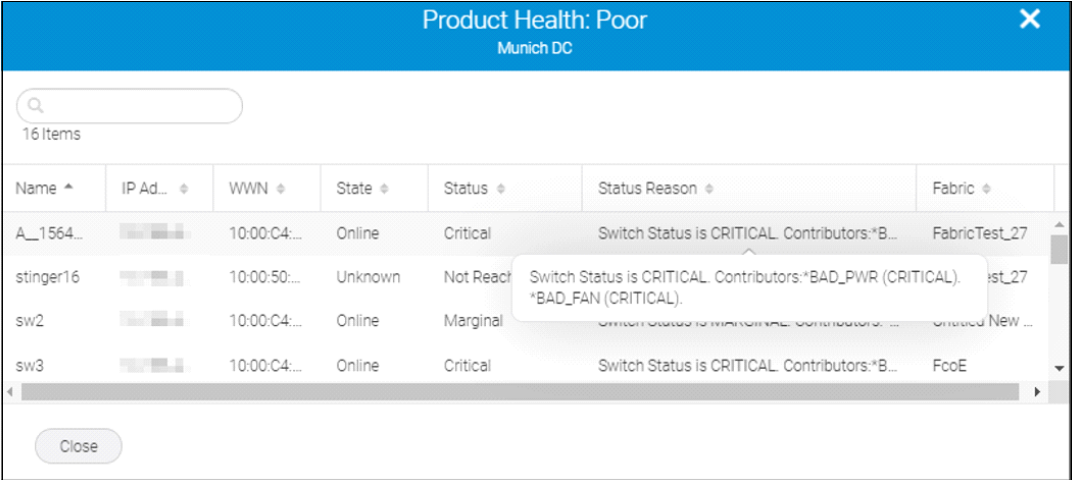


Figure 11-52 Product Health: Poor

4. Click **Close** to return to the dashboard.
5. Select **SANnav Portal** from the drop-down list in the widget, click the bar graph, and then, select **Show in Portal** to display the switch inventory in the selected SANnav Management Portal instance. If you are not logged in to this portal, credentials for this portal are required. A separate window opens for the Management Portal instance.

11.4 IBM SANnav WebTools

WebTools is started directly from a Web browser or from SANnav Management Portal. You can start WebTools on any workstation with a compatible operating system and Web browser installed.

If the switch is configured with logical fabrics, you can log in to any of the logical fabrics for which you have permission. Complete the following steps:

1. Start WebTools directly from a browser or from SANnav Management Portal:

- To start directly from a web browser, open your browser, enter the IP address of the switch, and press Enter.

You can use HTTP or HTTPS, for example: `http://10.77.77.77` or `https://10.77.77.77`.

- To start from SANnav Management Portal, locate the switch on the SANnav Inventory page, click the down arrow to the right of the switch, and select **View in WebTools** from the action menu (see Figure 11-53).

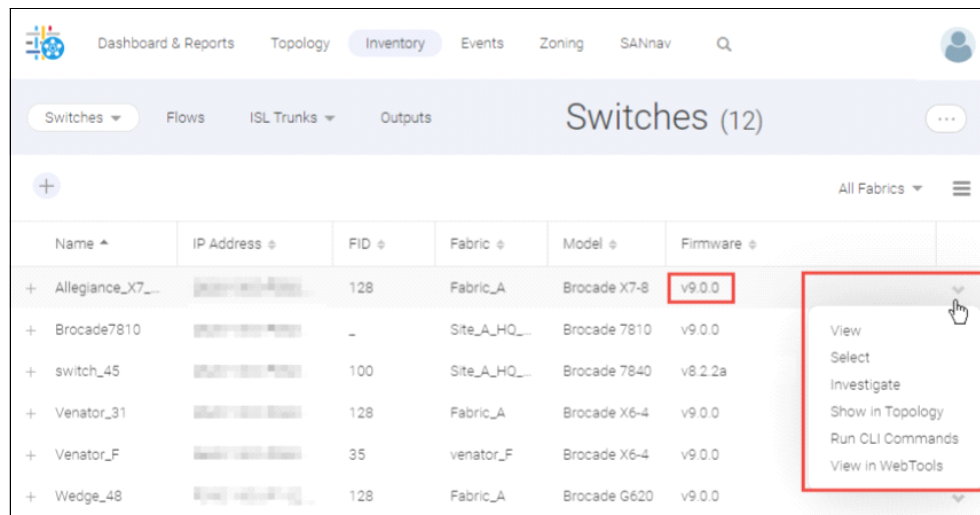


Figure 11-53 Starting WebTools from SANnav Management Portal

2. Enter the username, password, and logical switch name or fabric ID (FID). For the first switch login, the default user name is admin and the default password is password (see Figure 11-54).

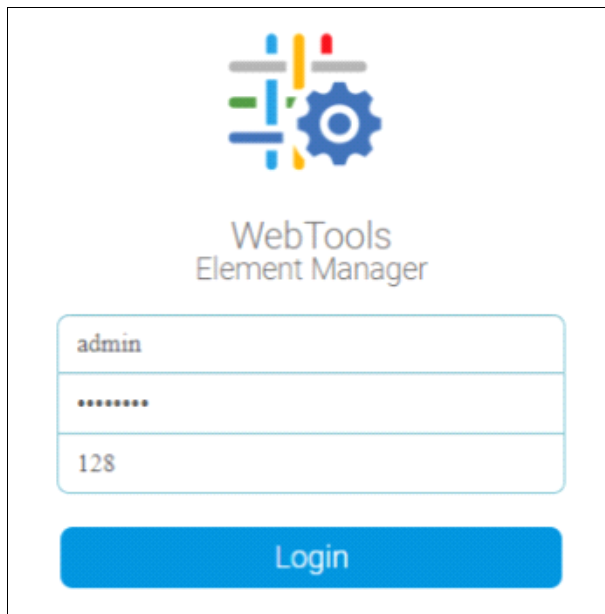
The image shows the WebTools Element Manager login interface. At the top is a logo consisting of a blue gear and several vertical bars in blue, green, yellow, and red. Below the logo, the text "WebTools" is in a larger font, and "Element Manager" is in a smaller font below it. There are three input fields stacked vertically: the first contains the text "admin", the second contains a series of dots representing a password, and the third contains the number "128". Below these fields is a large blue button with the word "Login" in white text.

Figure 11-54 WebTools Element Manager

WebTools prompts you to change the default password. If you are logging in to a Virtual Fabrics-enabled platform and you do not specify a logical switch, you are logged in to the default logical switch, which uses fabric ID 128. For non-VF platforms, the FID option is not displayed.

If you start from SANnav Management Portal, you might not need to log in, depending on the SANnav single sign-on configuration.

3. Click **Login**.

11.4.1 WebTools user interface overview

After you are familiar with the basic components of WebTools, you can quickly start monitoring and managing your switch. The basic layout of the WebTools user interface is shown in Figure 11-55.



Figure 11-55 WebTools user interface

The following components are shown in Figure 11-55:

1. Navigation bar: Contains links to feature pages.
2. Profile menu: Displays the link for logging out.
3. Subnavigation bar: Provides the page title and includes buttons and menus to take actions within the page.

For more information about the SANnav WebTools product, see [this web page](#).

11.5 Reference guides

For more information about the hardware and software that is supported for SANnav Management Portal V2.1, see the Documentation section at [this web page](#).

For more information, see the following guides:

- *Brocade SANnav Management Portal User Guide* describes how to monitor and manage your storage area network (SAN) using Brocade SANnav Management Portal.
- *Brocade SANnav Flow Management User Guide* explains how to configure and manage flows by using SANnav Management Portal.
- *Brocade SANnav Management Portal REST API and Northbound Streaming Reference Manual* contains definitions of REST APIs you can use to access SANnav Management Portal, including streaming performance and flow metrics to an external server.
- *Brocade SANnav Global View User Guide* describes how to use SANnav Global View to monitor and manage multiple Management Portal instances. SANnav Global View is a separate product.
- *Brocade SANnav Management Portal Release Notes* includes a summary of the new, unsupported, and deprecated features for this release.



Storage area network security

Security is always an advantage with Fibre Channel (FC) based storage area networks (SANs). From its air-gapped deployment (often times, these networks are physically separated from the rest of the enterprise network), to the much more difficult medium (fiber) that is used for eaves dropping, to the much reduced compromise domain where an impacted server can affect only the LUNs zoned and masked to it and no others.

However, FC SANs cannot ignore security. Mission-critical applications run on SANs and this critical network and its data must be protected.

When we discuss FC SAN Security, several challenges exist. A layered security model should be deployed to ensure it continues to be secure for years to come.

In this chapter, we review each of these layers independently, working from the outermost layers that deal with Management Security to the innermost layers that deal with the data.

This chapter includes the following topics:

- ▶ 12.2, “Management interfaces” on page 312
- ▶ 12.3, “Displaying the IP filter” on page 313
- ▶ 12.4, “Role-based access controls” on page 314
- ▶ 12.5, “Default accounts” on page 315
- ▶ 12.6, “User accounts” on page 315
- ▶ 12.7, “Security protocols” on page 315
- ▶ 12.8, “Access Control Lists” on page 316
- ▶ 12.9, “Securing end points in SAN infrastructure” on page 317
- ▶ 12.10, “ISL in-flight encryption and compression” on page 318
- ▶ 12.11, “Policy Database Distribution” on page 322

12.1 Fibre Channel SAN security layers

Figure 12-1 shows the layers of FC SAN Security.

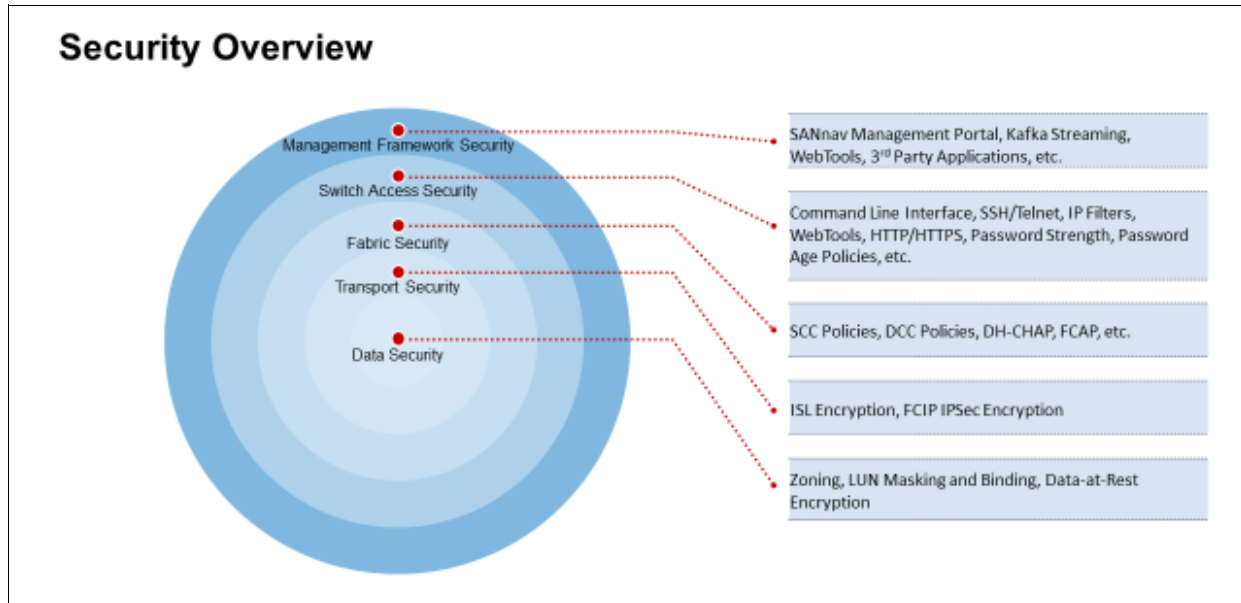


Figure 12-1 FC SAN security layers

12.2 Management interfaces

In this section, we discuss the Management Interface, which is the first and most vulnerable area we review. Several methods are available for Brocade Fabric OS (FOS) for protection.

We want to limit what IP addresses can access the management VLAN from where our devices are managed. We restrict those IP addresses by using separate VLANs, while locking down which protocols are used (disabling the least secure protocols, and closing their ports).

We use IP Filter in FOS to allow or deny the entry of traffic across these interfaces. We also want to make sure, if necessary, that we equally limit both IPv4 and IPv6 traffic. The IP Filter policy is a set of rules that are applied to the IP management interfaces and act as a packet filtering firewall. This firewall permits or denies the traffic to go through the IP management interfaces according to the IP Filter rules.

As a preferred practice, non-secure IP protocols that are used for switch management, such as Telnet and HTTP, must be blocked. SSH is enabled on the default IP Filter policy and SSL must be configured to use HTTPS for web access.

The IP Filter policy must be used in environments where you need strict control of fabric access. As with other ACL policies, it is important to regularly review and properly maintain the IP Filter policy.

12.3 Displaying the IP filter

To display the current IP Filter in use, use the `ipfilter --show` command, as shown in the following examples:

```
switch:admin> ipfilter --show
```

```
Name: default_ipv4, Type: ipv4, State: active
```

Rule	Source IP	Protocol	Dest. Port	Action
1	any	tcp	22	permit
2	any	tcp	23	permit
3	any	tcp	80	permit
4	any	tcp	443	permit
5	any	udp	161	permit
6	any	udp	123	permit
7	any	tcp	600 - 1023	permit
8	any	udp	600 - 1023	permit

```
switch:admin> ipfilter --show
```

```
Name: default_ipv4, Type: ipv4, State: active
```

Rule	Source IP	Protocol	Dest. Port	Action
1	any	tcp	22	permit
2	any	tcp	23	permit
3	any	tcp	80	permit
4	any	tcp	443	permit
5	any	udp	161	permit
6	any	udp	123	permit
7	any	tcp	600 - 1023	permit
8	any	udp	600 - 1023	permit

```
Name: default_ipv4, Type: ipv4, State: active
```

Rule	Source IP	Protoco l	Dest Port	Action
1	any	tcp	22	permit
2	any	tcp	23	permit
3	any	tcp	80	permit
4	any	tcp	443	permit
5	any	udp	161	permit
6	any	udp	123	permit
7	any	tcp	600 - 1023	permit
8	any	udp	600 - 1023	permit

A newly created IP Filter to deny an unsecured protocol, such as Telnet that uses TCP Port 23 is shown in the following example (the IPv6 filter remains unchanged):

```
ipfilter --clone BlockTelnet -from default_ipv4
ipfilter --save BlockTelnet
ipfilter --addrule BlockTelnet -rule 1 -sip any -dp 23 -proto tcp -act deny
ipfilter --save
ipfilter --activate BlockTelnet
```

```
switch:admin> ipfilter --show
```

```
Name: default_ipv4, Type: ipv4, State: active
```

Rule	Source IP	Protocol	Dest Port	Action
1	any	tcp	22	permit
2	any	tcp	23	deny
3	any	tcp	80	permit
4	any	tcp	443	permit
5	any	udp	161	permit
6	any	udp	123	permit
7	any	tcp	600 - 1023	permit
8	any	udp	600 - 1023	permit

```
Name: default_ipv6, Type: ipv6, State: active
```

Rule	Source IP	Protocol	Dest Port	Action
1	any	tcp	22	permit
2	any	tcp	23	permit
3	any	tcp	80	permit
4	any	tcp	443	permit
5	any	udp	161	permit
6	any	udp	123	permit
7	any	tcp	600 - 1023	permit
8	any	udp	600 - 1023	permit

12.4 Role-based access controls

One way to provide limited accessibility to the fabric is through user roles. FOS includes predefined user roles, each of which can access a subset of the CLI commands. These roles are known as *role-based access controls* (RBACs), which are associated with the user's login credentials.

When you log in to a switch, your user account is associated with a predefined role or a user-defined role. The role that your account is associated with determines the level of access you have on that switch and in the fabric.

12.5 Default accounts

FOS offers four predefined accounts:

- ▶ Admin
- ▶ Factory
- ▶ Root
- ▶ User

Although the root and factory accounts are reserved for development and manufacturing, the password for all default accounts must be changed and secured during the initial installation and configuration of each switch.

Recovering lost passwords might require significant effort and fabric downtime and assistance from Support resources.

12.6 User accounts

In addition to the default permissions that are assigned to the roles of admin, factory, root, and user, FOS supports up to 252 other user accounts on the chassis. These accounts expand the ability to track account access and audit administrative activities.

Each user account is associated with the following attributes:

- ▶ Virtual Fabric list: Specifies the Virtual Fabric that a user account is allowed to log in to.
- ▶ Home Virtual Fabric: Specifies the Virtual Fabric that the user is logged in to, if available. The home Virtual Fabric must be a member of the user's Virtual Fabric list. If the fabric ID is not available, the next-lower valid fabric ID is used.
- ▶ LF Permission List: Determines functional access levels within the bounds of the user's Virtual Fabrics.
- ▶ Chassis role. Similar to switch-level roles, but applies to a different subset of commands.

12.7 Security protocols

Security protocols provide endpoint authentication and communications privacy by using cryptography. Typically, the user is authenticated to the switch while the switch remains unauthenticated. This authentication means that you can be sure that you know with whom you are communicating.

The next level of security, in which both ends of the conversation are sure with whom they are communicating, is known as *two-factor authentication*. Two-factor authentication requires public key infrastructure (PKI) deployment to clients.

The security protocols that are supported by FOS are listed in Table 12-1.

Table 12-1 Security protocols that are supported by FOS

Protocol	Description
CHAP	Challenge Handshake Authentication Protocol (CHAP) uses shared secrets to authenticate switches.
HTTPS	Hyper Text Transfer Protocol over SSL (HTTPS) is a Uniform Resource Identifier (URI) scheme that is used to indicate a secure HTTP connection. WebTools supports the use of HTTPS.
IPSec	Internet Protocol Security (IPSec) is a framework of open standards for providing confidentiality, authentication, and integrity for IP data transmitted over untrusted links or networks.
LDAP	Lightweight Directory Access Protocol (LDAP) with Transport Layer Security (TLS) uses a certificate authority (CA). By default, LDAP traffic is transmitted unsecured. With imported signed certificates, you can make LDAP traffic confidential and secure by using Secure Sockets Layer (SSL)/TLS technology with LDAP.
SCP	Secure Copy Protocol (SCP) is a means of securely transferring computer files between a local and a remote host or between two remote hosts that use the Secure Shell (SSH) protocol. Configuration upload and download support the use of SCP.
Secure Syslog	Secure syslog requires importing syslog CA certificates by using the secCertMgmt command.
SFTP	Secure File Transfer Protocol (SFTP) is a network protocol for securely transferring files on a network. Configuration upload and download support the use of SFTP.
SNMP	Simple Network Management Protocol (SNMP) is used in network management systems to monitor network-attached devices for conditions that warrant administrative attention. Supports SNMPv1 and v3.
SSH	Secure Shell (SSH) is a network protocol that allows data to be exchanged over a secure channel between two computers. Encryption provides confidentiality and integrity of data. SSH uses public-key cryptography to authenticate the remote computer and allow the remote computer to authenticate the user, if necessary.
SSL	Secure Sockets Layer (SSL) is used by FOS to support HTTPS. A certificate must be generated and installed on each switch to enable SSL. This configuration supports SSLv3, 128-bit encryption by default. It also supports TLSv1.0, TLSv1.1, and TLSv1.2.

12.8 Access Control Lists

Access control lists (ACLs) are used to provide network security through policy sets. FOS provides several ACL policies, including a Switch Connection Control (SCC) policy, a Device Connection Control (DCC) policy, a Fabric Configuration Server (FCS) policy, and an IP Filter:

- SCC policy

Use the SCC policy to restrict FC switches that can join the fabric. Only switches that are specified in the policy can join the fabric. All other switches fail authentication if they attempt to connect to the fabric, which results in the respective E_Ports being segmented because of the security violation.

Use the SCC policy in environments where you need strict control of fabric switch members.

The SCC policy can prevent switches from participating in a fabric; therefore, it is important to regularly review and properly maintain the SCC policy.

- DCC policy

Use the DCC policy to restrict devices that can attach to a single FC port. The policy specifies the FC port and one or more worldwide names (WWNs) that can connect to the port. The DCC policy comprises all of the DCC policies that are defined for individual FC ports.

Not every FC port must have a DCC policy, and only ports with a DCC policy in the active policy set enforce access controls. A port that is present in the active DCC policy allows only WWNs in its respective DCC policy to connect and join the fabric. All other devices fail authentication when they attempt to connect to the fabric, which results in the respective F_Ports being disabled because of the security violation.

Use the DCC policy in environments where you need strict control of fabric device members.

Because the DCC policy can prevent devices from participating in a fabric, it is important to regularly review and properly maintain the DCC policy.

- FCS policy

Use the FCS policy to restrict the source of fabric-wide settings to one FC switch. The policy contains the WWN of one or more switches, and the first WWN (that is online) in the list is the primary FCS. If the FCS policy is active, only the primary FCS can make or propagate fabric-wide parameters. These parameters include zoning, security (ACL) policies databases, and other settings.

12.9 Securing end points in SAN infrastructure

The SAN infrastructure consists of HBAs, switches, and storage devices. Each of these components offer security features that must be enabled to maximize security.

Several options are available regarding HBAs. We can use Single Initiator Zones to limit the number of LUNs that might be compromised if the server is compromised.

On the storage device, LUN masking must be implemented to ensure only those LUNs that are zoned to the HBA are accessible to the HBA through its zoned storage port. We also can enable DH-CHAP to authenticate these devices when they join the fabric.

Similar options available regarding storage devices. Single Initiator Zones must be implemented with LUN masking to limit exposure if a device be compromised. Secure protocols, such as DH-CHAP, also are available to authenticate these devices when they join the fabric.

Even more options are available regarding FC switches. We can disable persistently ports that are not used in the switch. We can also disable the ability of the online ports from becoming E-Ports (ISLs) by locking down their port configuration to disallow that type of login.

By using a DCC policy, we can create an ACL of sorts whereby we can control what device can log in to specific ports by its WWN. By using an SCC Policy, we can also use this same ACL concept to lock down the switches that can participate in a fabric by the switch's WWN. These features are used heavily when High Integrity Fabrics are created that are required for most FICON environments.

We can also use DH-CHAP or FCAP to authenticate switches. This authentication type is required when establishing Encrypted E-Ports (ISLs) between switches, usually across long distances or connected by way of DWDM devices. It is a best practice to enable encryption on ISLs or FCIP links extending beyond the data center walls. DH-CHAP is a simpler deployment where shared secrets are exchanged between switches; FCAP uses more secure signed digital certificates for authentication.

12.10 ISL in-flight encryption and compression

IBM b-type Gen 7 FC platforms support in-flight compression and encryption at a port level for local and long-distance ISLs. In-flight data compression is a useful tool for saving money when bandwidth caps or bandwidth usage charges are in place for transferring data between fabrics. Similarly, in-flight encryption enables a further layer of security with no key management impact when transferring data between local and long-distance data centers besides the initial setup.

Enabling in-flight ISL data compression or encryption slightly increases the latency as the ASIC processes the frames for compression or encryption or both.

How to implement DH-CHAP for ISL Encryption is shown in the following example process:

1. Verify whether DH-CHAP Secrets exist in the switch database:

```
switch:admin> secauthsecret --show
Chap secret key database does not exist.
```

2. Run **secauthsecret --set** to set DH-CHAP authentication and create secrets on each switch:

```
switch:admin> secauthsecret --set
This command is used to set up secret keys for the DH-CHAP authentication.
The minimum length of a secret key is 8 characters and maximum 40
characters.
Setting up secret keys does not initiate DH-CHAP authentication.
If switch is configured to do DH-CHAP, it is performed whenever a port or a
switch is enabled.
Warning: Please use a secure channel for setting secrets.
Using an insecure channel is not safe and may compromise secrets.
Following inputs are specified for each entry.
  WWN for which secret is being set up.
  Peer secret: The secret of the peer that authenticates to peer.
  Local secret: The local secret that authenticates peer.
Press Enter to start setting up shared secrets >
Enter WWN, Domain, or switch name (Leave blank when done):
10:00:8j:94:7f:40:90:20
Enter peer secret:
Reenter peer secret:
Enter local secret:
Reenter local secret:
Warning: Keeping same peer and local secrets can compromise the security of
your network. To change, reenter WWN and secrets.
Enter peer WWN, Domain, or switch name (Leave blank when done):
Are you done? (yes, y, no, n): [no] y
Saving data to keystore... Done.
```

3. Validate that switch secrets are now stored in the switch database:

```
switch:admin> secauthsecret --show
WWN                               DId   Name
-----
10:00:8j:94:7f:40:90:20         60    switchAname_DH-CHAP_example
```

4. Enable DH-CHAP as the authentication method:

```
switch:admin> authutil --set -a dhchap
Authentication is set to dhchap.
```

5. Set the DH Group to 4:

```
switch:admin> authutil --set -g "4"
switch:admin> authutil --show
AUTH TYPE      HASH TYPE      GROUP TYPE
-----
fcap,dhchap    sha256,sha1,md5 4
```

6. Activate the configuration:

```
switch:admin> authutil --policy -sw active
Warning: Activating the authentication policy requires either DH-CHAP secrets
or PKI certificates depending on the protocol selected. Otherwise, ISLs will be
segmented during next E-port bring-up.
ARE YOU SURE (yes, y, no, n): [no] yes
Auth Policy is set to active.
```

7. Validate the authentication configuration:

```
switch:admin> authutil --show

AUTH TYPE      HASH TYPE      GROUP TYPE
-----
fcap,dhchap    sha256,sha1,md5 4
Switch Authentication Policy: ACTIVE
Device Authentication Policy: OFF
```

8. Configure ports to use encryption for ISLs on local and remote switches:

```
switch:admin> portdisable 1/23
switch:admin> portcfgencrypt --enable 1/23
switch:admin> portenable 1/23
```

9. Validate the port's configuration.:

```
switch:admin> portcfgshow 1/23
Area Number:                23
Octet Speed Combo:          1(32G|16G|8G|4G)
Speed Level:                 8G
AL_PA Offset 13:            OFF
Trunk Port                   ON
Long Distance                OFF
VC Link Init                 OFF
Locked L_Port                OFF
Locked G_Port                OFF
Disabled E_Port              OFF
Locked E_Port                OFF
ISL R_RDY Mode               OFF
RSCN Suppressed              OFF
Persistent Disable           OFF
LOS TOV mode                 0(OFF)
```

NPIV capability	ON
QOS Port	AE
Port Auto Disable:	OFF
EX Port	OFF
Mirror Port	OFF
SIM Port	OFF
Credit Recovery	ON
F_Port Buffers	OFF
E_Port Credits	OFF
Fault Delay:	0(R_A_TOV)
NPIV PP Limit:	126
NPIV FLOGI Logout:	OFF
CSCTL mode:	OFF
TDZ mode:	OFF
D-Port mode:	OFF
D-Port over DWDM:	OFF
Compression:	OFF
Encryption:	ON
10G/16G FEC:	ON
16G FEC via TTS:	OFF

The following process shows an alternative way to enable FCAP for ISL encryption:

1. Generate switch CSR files:

```
switch:admin> seccertmgmt generate -csr fcap
```

2. Export Switch CSR Files to CA for signing:

```
switch:admin> seccertmgmt export -csr fcap
```

Exact Sample (by using complete and customer supplied input):

```
seccertmgmt export -csr fcap -protocol scp -ipaddr 10.1.1.1 -login  
hostusername -password hostpassword -remotedir /home/fcap/fcap_files
```

3. Import CA Cert to each switch:

```
switch:admin> seccertmgmt import -ca -client fcap
```

The following code is an exact sample (uses complete and customer supplied input):

```
seccertmgmt import -ca -client fcap -protocol scp -ipaddr 10.1.1.1 -remotedir  
/home/fcap/fcap_files -certname ca.pem -login hostusername -password  
hostpassword
```

4. (On host) Verify switch CSRs:

```
openssl req -text -noout -verify -in FcapSw.csr
```

5. (On Host) CA signs switch CSRs:

```
openssl ca -in FcapSw.csr -out FcapSw.pem
```

6. Verify CA signed switch CSRs:

```
openssl x509 -in FcapSw.pem -text -noout
```

7. Validate filename permission on host:

```
chmod 777 FcapSw.pem
```

8. Import CA signed switch CSRs:

```
switch:admin> seccertmgmt import -cert fcap
```

The following code is an exact sample (uses complete and customer supplied input):

```
seccertmgmt import -cert fcap -protocol scp -ipaddr 10.1.1.1 -remotedir  
/home/fcap/fcap_files -certname FcapSw.pem -login hostsusername -password  
hostspassword
```

9. Validate switch database:

```
switch:admin> seccertmgmt show -all
```

10. Enable FCAP authentication:

```
switch:admin> authutil --set -a fcap  
Authentication is set to fcap.  
switch:admin> authutil --set -g "4"  
DH Group was set to 4.  
switch:admin> authutil --set -h sha256  
Hash is set to sha256.
```

11. Configure ports to use ISL Encryption on local and remote switches:

```
switch:admin> portdisable 1/23  
switch:admin> portcfgencrypt --enable 1/23  
switch:admin> portenable 1/23
```

12. Validate the port's configuration:

```
switch:admin> portcfgshow 1/23  
Area Number:          23  
Octet Speed Combo:    1(32G|16G|8G|4G)  
Speed Level:          8G  
AL_PA Offset 13:      OFF  
Trunk Port            ON  
Long Distance         OFF  
VC Link Init          OFF  
Locked L_Port         OFF  
Locked G_Port         OFF  
Disabled E_Port       OFF  
Locked E_Port         OFF  
ISL R_RDY Mode        OFF  
RSCN Suppressed       OFF  
Persistent Disable    OFF  
LOS TOV mode          0(OFF)  
NPIV capability       ON  
QOS Port              AE  
Port Auto Disable:    OFF  
EX Port               OFF  
Mirror Port           OFF  
SIM Port              OFF  
Credit Recovery       ON  
F_Port Buffers        OFF  
E_Port Credits        OFF  
Fault Delay:          0(R_A_TOV)  
NPIV PP Limit:        126  
NPIV FLOGI Logout:    OFF  
CSCTL mode:           OFF  
TDZ mode:             OFF  
D-Port mode:          OFF  
D-Port over DWDM:     OFF  
Compression:          OFF
```

Encryption:	ON
10G/16G FEC:	ON
16G FEC via TTS:	OFF

Consider the following points regarding in-flight encryption and compression:

- ▶ Encryption and compression are supported on E_Ports and EX_Ports.
- ▶ ISL ports must be set to Long-Distance (LD) mode when compression is used.
- ▶ Twice the number of buffers must be allocated if compression is enabled for long distances because frame sizes might be half the size.
- ▶ If compression and encryption are used, enable compression first.

Implementing ISL encryption by using multiple ISLs between the same switch pairs requires that all ISLs be configured for encryption, or none at all. No more than four ports on one ASIC can be configured with encryption, compression, or both when running at 64 Gbps speed. More ports can be used for data encryption, data compression, or both if the ports are running lower than 64 Gbps speeds.

Consider the following points regarding Virtual Fabric (encryption and compression):

- ▶ The E_Ports in the user-created logical switch, base switch, or default switch can support encryption and compression.
- ▶ Encryption and compression are supported on XISL ports, but not on LISL ports.
- ▶ If encryption or compression is enabled and ports are moved from one LS to another, it must be disabled before the ports are moved.

12.11 Policy Database Distribution

Security Policy Database Distribution provides a mechanism for controlling the distribution of each policy on a per-switch basis. Switches can individually configure policies to accept or reject a policy distribution from another switch in the fabric.

In addition, a fabric-wide distribution policy can be defined for the SCC and DCC policies with support for strict, tolerant, and absent modes. This setting can be used to enforce whether the SCC or DCC policy must be consistent throughout the fabric.

The Policy Database Distribution features the following modes:

- ▶ **Strict**
All updated and new policies of the specified type (SCC, DCC, or both) must be distributed to all switches in the fabric, and all switches must accept the policy distribution.
- ▶ **Tolerant**
All updated and new policies of the type specified (SCC, DCC, or both) are distributed to all switches (FOS V6.2.0 or later) in the fabric. However, the policy does not need to be accepted.
- ▶ **Absent**
Updated and new policies of the type specified (SCC, DCC, or both) are not automatically distributed to other switches in the fabric. Policies can still be manually distributed.

Together, the policy distribution and fabric-wide consistency settings provide a range of control on the security policies from little or no control to strict control. Strict mode is required for FICON environments that require High Integrity Fabrics be enabled.

For more information about other topics in this subject area that are out of the scope of this document, such as encrypting data that is at-rest and in-flight, see [Brocade Fabric OS Administration Guide, 9.0x](#).



Fabric details

This appendix provides example tables that you can use to identify dominant factors, including facilities that can affect the storage area network (SAN) design.

This appendix includes the following topics:

- ▶ A.1, “Current fabrics” on page 326
- ▶ A.2, “Individual fabrics” on page 326
- ▶ A.3, “Device” on page 327
- ▶ A.4, “Metrics and effect on SAN design and performance” on page 327
- ▶ A.5, “Consolidated SAN requirements” on page 328
- ▶ A.6, “Application-specific information” on page 329

A.1 Current fabrics

Use the table that is shown in Table A-1 to record the fabrics that are used.

Table A-1 *Current fabrics*

SAN/ fabric	Number of switches	Type of switches	Total ports	Domains	Number of servers	Number of storage devices	Location	Notes
Fabric 1								
Fabric 2								
Fabric 3								
Fabric 4								
Fabric 5								

A.2 Individual fabrics

Use the table that is shown in Table A-2 to record the individual fabrics that are used.

Table A-2 *Individual fabrics*

SAN/ fabric	Domain number	Serial number	Model	Speed	WWN	IP addresses	Brocade Fabric OS or M-EOS Version	Notes
Switch 1								
Switch 2								
Switch 3								
Switch 4								
Switch 5								

A.3 Device

Use the table that is shown in Table A-3 to record information about the device that is used.

Table A-3 Devices

Servers and Storage	Vendor	Model	WWN	Alias	Zone	Version	Application	Switches	Notes
Server 1									
Server 2									
Server 3									
Storage 1									
Storage 2									
Storage 3									

A.4 Metrics and effect on SAN design and performance

Use the table that is shown in Table A-4 to record the metrics and effects on SAN design.

Table A-4 Metrics and effect on SAN design and performance

Metric	Source	Effect
Servers in the SAN	Estimate/Brocade SAN Health	Normal operations
Host Level Mirroring ²	Estimate	▶ Distance ▶ ISL ▶ Traffic levels
Clusters (MSFT, HACMP, NetApp) Average number of nodes Workload level	Estimate Estimate: High/Med/Low	▶ In-band heartbeat ▶ Frame congestion ▶ Host fan-in ▶ Traffic isolation
Virtualization: VIO Server Number of servers Consolidation ratio	Estimate Estimate Estimate	▶ Frame congestion ▶ Edge traffic increase per port ▶ Server fan-in on target ports ▶ Device latencies
Virtualization: VMware Number of VMware servers Consolidated ratio? Shared VMFS? DRS? RDM? I/O intensive?	Estimate Estimate Yes/No Yes (%) / No Yes (%) / No High/Med/Low Yes/No	▶ Frame congestion ▶ Device latencies ▶ SCSI12 reser

A.5 Consolidated SAN requirements

Use the table that is shown in Table A-5 to record the consolidated SAN requirements.

Table A-5 Consolidated SAN requirements

SAN requirements data (complete for each SAN)	
Fabric information	
Target number of user ports per fabric	
Target number of total ports per fabric	
Target number of switches per fabric (number of switches/switch type, total switches)	
Number of fabrics	
Number of sites in environment	
Topology (core-edge, ring, mesh, other)	
Maximum hop count	
Fabric licenses	
SAN device information	
Number/types of hosts and operating system platforms	
Number/types of storage devices	
Number/types of tapes	
Number/types of HBAs	
Other devices (VTL/deduplication appliance)	
Total number of SAN devices per fabric	
Customer requirement for failover/redundancy, reliability of SAN (multipathing software utilized)	
Application details	
SAN Application (Storage Consolidation, Backup and Restore, Business Continuity) Fabric management application(s)	
Fabric management application(s)	
Performance	
Maximum latency (ms)	
Targeted ISL oversubscription ratio (3:1, 7:1, 15:1, other)	

A.6 Application-specific information

Use the table that is shown in Table A-6 to record application-specific information.

Table A-6 Application-specific information

Backup/restore infrastructure		
Servers		
System	Operating system version, patch level	HBA driver version
Server 1/HBA		
Server 2/HBA		
Server 3/HBA		
Backup software		
Vendor	Version	Patch
FC switch		
Vendor	Model	Firmware
Brocade		
Storage		
Vendor	Model	Firmware
Array 1		
Array 2		
Tape library		
Vendor	Model	Firmware
Library		

Note: Keep a similar table for each application.



Brocade Fabric OS CLI monitoring commands

Many CLI commands show statistical output based on traffic, interfaces, circuits, tunnels, and TCP sessions. In this chapter, we describe some of the most commonly used commands.

For more information about the Brocade Fabric OS (FOS) commands, see [Brocade Fabric OS Command Reference Manual, 9.0.x](#).

This appendix includes the following topics:

- ▶ B.1, “portshow (tunnel and circuits)” on page 332
- ▶ B.3, “portshow (TCP Stats)” on page 333
- ▶ B.4, “portStatsShow” on page 338
- ▶ B.5, “gePortPerfShow” on page 340
- ▶ B.6, “switchShow” on page 341
- ▶ B.7, “portPerfShow” on page 342
- ▶ B.8, “portErrShow” on page 342

B.1 portshow (tunnel and circuits)

FOS provides for the CLI monitoring of tunnels and circuits:

- ▶ State: Up
- ▶ Tunnel: VE24
- ▶ Circuits: 2 (0 and 1)
- ▶ Tx rate: 18.4 MBps (sending data)
- ▶ Rx rate: 0.4 MBps (receiving ACK)
- ▶ Load Balance: 9.2 MBps
- ▶ Compression ratio: 4.4:1, indicating 4 times the WAN rate is coming from the application
- ▶ RTT: 20 ms
- ▶ ReTx (retransmits): IP network has packet loss
- ▶ ARL BW:
 - Circuit minimum comm-rate is 100 Mbps.
 - Circuit maximum comm-rate is 150 Mbps.
- ▶ TxWAN%: 73% (transmitting into the minimum comm-rate)
- ▶ Tx Q%: 0 (Transmit Queue % is empty)
- ▶ Met: 0 (failover metric can be 0 or 1, 0 is the default)
- ▶ G: - (group number - no group is the default)

```
7840-Side-A:FID128:admin> portshow fcipunnel -c --perf --summary
```

Tunnel	Circuit	St	Flg	TxMBps	RxMBps	CmpRtio	RTTms	ReTx	TxWAN%	TxQ%/BW	Met/G
24	-	Up	---	18.4	0.4	4.4:1	-	19367	73	0	-
24	0 ge2	Up	---	9.2	0.2	-	20	9799	71	100/150	0/-
24	1 ge4	Up	---	9.2	0.2	-	20	9568	75	100/150	0/-

Flg (tunnel): I=IP-Ext, s=Spillover

St: High level state, Up or Dn

TxWAN (tunnel): Tx WAN utilization high of primary circuits (--qos for range)

(circuit): Tx WAN utilization high (--qos for range)

TxQ (tunnel): Tx data buffering utilization high (--qos for range)

B.2 portshow (LAN Stats)

LAN stats per-flow shows active flows from the end devices to the IP Extension gateways. If flows are listed, the end device is directing traffic to the IP Extension gateway. The flows are matching a TCL rule and being directed into a tunnel. The amount of Tx/Rx traffic per flow is shown in the following example:

```
LVN-7840-A:FID128:admin> portshow lan-stats --per-flow -all
```

*** Displaying 142 connections ***

DP	Idx	Src-Address	Dst-Address	Sport	Dport	Pro	Tx(B/s)	Rx(B/s)
DP0	1592	10.13.30.171	10.75.62.74	49630	6690	TCP	0	0
DP1	5599	10.75.52.130	10.89.122.129	62447	5667	TCP	10	0
DP1	8600	10.75.52.133	10.89.122.126	37638	5667	TCP	8	0

```

... <cut for brevity> ...
DP1    6567 10.75.52.130    10.89.122.126    62461 5667  TCP    8      0
DP1    7554 10.75.52.133    10.89.122.127    51029 5667  TCP    8      0
-----
Sport=Source-Port Dport=Destination-Port Pro=Protocol
DP      ActTCP  ExdTCP  TCLDeny TCLFail
-----
DP0      159     0       38928   0
DP1      254     0       53038   0
-----
ActTCP=Active TCP Conns  ExdTCP=Exceeded TCP Conn Cnt

```

B.3 portshow (TCP Stats)

Monitoring of TCP statistics shows all the sessions carried by a single circuit (VE24-Circuit0):

```

LVN-7840-B_CA:FID1:admin> portshow fcipunnel 24 0 --tcp
Tunnel: VE-Port:24 (idx:0, DP0)
=====
Oper State      : Online
TID             : 24
Flags           : 0x00000000
IP-Extension    : Disabled
Compression     : Aggressive Deflate
QoS BW Ratio    : 50% / 30% / 20%
Fastwrite       : Disabled
Tape Pipelining : Disabled
IPSec           : Disabled
Legacy QoS Mode : Disabled
Load-Level (Cfg/Peer): Failover (Failover / Failover)
Local WWN       : 10:00:50:eb:1a:69:ee:93
Peer WWN        : 10:00:50:eb:1a:fc:e0:da
RemWWN (config) : 00:00:00:00:00:00:00:00
FICON           : Enabled
  XRC            : Enabled
  Tape Pipelining : Disabled
  Read Block id   : Disabled
  Tin Tir         : Enabled
  DevAck          : Enabled
  Tera Read       : Disabled
  Tera Write      : Disabled
  Print           : Disabled
FICON Max Read Pipe : 32
FICON Max Write Pipe : 32
FICON Max Read Devs : 16
FICON Max Write Devs : 16
FICON Write Timer   : 300
FICON Write Chain   : 3200000
FICON OXID Base     : 0x9000
FICON Debug Flags   : 0x77c90000
cfgmask             : 0x003fc33f 0x40000208
Uncomp/Comp Bytes   : 5836840800 / 5836828155 / 1.00 : 1
Uncomp/Comp Byte(30s): 24169 / 24169 / 1.00 : 1
Flow Status         : 0

```

```

ConCount/Duration      : 1 / 28d22h
Uptime                 : 28d22h
Stats Duration         : 28d22h
Receiver Stats         : 9253712699 bytes / 37116200 pkts /    2.73 KBps Avg
Sender Stats           : 5997624195 bytes / 29887256 pkts /   870.00 Bps Avg
TCP Bytes In/Out       : 22185797130 / 19158427516
ReTx/000/SloSt/DupAck: 2555 / 1790 / 1519 / 1476
RTT (min/avg/max)      : 21 / 21 / 53 ms
Wan Util               : 0.1%
TxQ Util               : 0.0%
TCP Connection 24.0 HA-Type:Main Pri:Low Conn:0x0a8a941c

```

```

=====
Local / Remote Port    : 57308 / 3225
Duration               : 28d22h
MSS                   : 1460 bytes
ARL Min / Cur / Max    : 10000 / 10000 / 50000
ARL Reset Algo         : Reset
Send Window
  Size / Scale         : 1249280 / 9
  Slow Start Threshold : 16777216
  Congestion Window    : 16778676
  Pkts InFlight        : 0
Recv Window
  Size / Scale         : 1249792 (Max:1249792) / 9
SendQ Next / Min / Max : 0x3a459b50 / 0x3a459b50 / 0x3a459b50
RecvQ Next / Min / Max : 0xf20a52e9 / 0xf20a52e9 / 0xf21d6449
RecvQ Pkts             : 0
Sender Stats
  Sent Bytes / Pkts    : 1105903340 / 10103550
  Unacked Data         : 0
  Retransmits Slow / Fast : 177 / 0 (High:0)
  SlowStart            : 123
Receiver Stats
  Recv Bytes / Pkts    : 1105864704 / 10103344
  Out-of-Order         : 7 (High:18)
  Duplicate ACKs       : 0
RTT / Variance (High)  : 20 ms (52 ms) / 0 ms (32 ms)
TCP Connection 24.0 HA-Type:Main Pri:Low Conn:0x0a8a941e

```

```

=====
Local / Remote Port    : 51162 / 3226
Duration               : 28d22h
MSS                   : 1460 bytes
ARL Min / Cur / Max    : 10000 / 10000 / 50000
ARL Reset Algo         : Reset
Send Window
  Size / Scale         : 1249280 / 9
  Slow Start Threshold : 16777216
  Congestion Window    : 16778676
  Pkts InFlight        : 0
Recv Window
  Size / Scale         : 1249792 (Max:1249792) / 9
SendQ Next / Min / Max : 0x3a4555be / 0x3a4555be / 0x3a4555be
RecvQ Next / Min / Max : 0xd5999547 / 0xd5999547 / 0xd5aca747
RecvQ Pkts             : 0
Sender Stats

```

```

Sent Bytes / Pkts      : 1105897920 / 10103599
Unacked Data          : 0
Retransmits Slow / Fast : 182 / 0 (High:0)
SlowStart              : 128
Receiver Stats
Recv Bytes / Pkts      : 1105888500 / 10103382
Out-of-Order           : 7 (High:23)
Duplicate ACKs          : 0
RTT / Variance (High)  : 20 ms (54 ms) / 0 ms (48 ms)
TCP Connection 24.0 HA-Type:Main Pri:Medium Conn:0x0a8a9526
=====
Local / Remote Port    : 57309 / 3225
Duration               : 28d22h
MSS                    : 1460 bytes
ARL Min / Cur / Max    : 15000 / 15000 / 50000
ARL Reset Algo         : Reset
Send Window
Size / Scale           : 1249280 / 9
Slow Start Threshold   : 16777216
Congestion Window      : 16778676
Pkts InFlight          : 0
Recv Window
Size / Scale           : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max  : 0xfa9fa897 / 0xfa9fa897 / 0xfa9fa897
RecvQ Nxt / Min / Max  : 0x22bad4e3 / 0x22bad4e3 / 0x22cde643
RecvQ Pkts             : 0
Sender Stats
Sent Bytes / Pkts      : 5975313510 / 51406132
Unacked Data          : 0
Retransmits Slow / Fast : 462 / 207 (High:0)
SlowStart              : 327
Receiver Stats
Recv Bytes / Pkts      : 7321106136 / 51436908
Out-of-Order           : 0 (High:28)
Duplicate ACKs          : 750
RTT / Variance (High)  : 20 ms (55 ms) / 0 ms (40 ms)
TCP Connection 24.0 HA-Type:Main Pri:Medium Conn:0x0a8a9527
=====
Local / Remote Port    : 51163 / 3226
Duration               : 28d22h
MSS                    : 1460 bytes
ARL Min / Cur / Max    : 15000 / 15000 / 50000
ARL Reset Algo         : Reset
Send Window
Size / Scale           : 1249280 / 9
Slow Start Threshold   : 16777216
Congestion Window      : 16778676
Pkts InFlight          : 0
Recv Window
Size / Scale           : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max  : 0xfa9d726d / 0xfa9d71cd / 0xfa9d726d
RecvQ Nxt / Min / Max  : 0xdd8bb7a2 / 0xdd8bb7a2 / 0xdd9ec9a2
RecvQ Pkts             : 0
Sender Stats
Sent Bytes / Pkts      : 5975138146 / 51404866

```

```

Unacked Data          : 160
Retransmits Slow / Fast : 472 / 203 (High:0)
SlowStart             : 342
Receiver Stats
  Recv Bytes / Pkts    : 7320332818 / 51434434
  Out-of-Order         : 0 (High:31)
  Duplicate ACKs       : 701
  RTT / Variance (High) : 20 ms (52 ms) / 0 ms (35 ms)
TCP Connection 24.0 HA-Type:Main Pri:High Conn:0x0a8a962f
=====
Local / Remote Port    : 57310 / 3225
Duration               : 28d22h
MSS                    : 1460 bytes
ARL Min / Cur / Max    : 25000 / 25000 / 50000
ARL Reset Algo         : Reset
Send Window
  Size / Scale         : 1249280 / 9
  Slow Start Threshold : 16777216
  Congestion Window    : 16778676
  Pkts InFlight        : 0
Recv Window
  Size / Scale         : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max  : 0x3a455ef4 / 0x3a455ef4 / 0x3a455ef4
RecvQ Nxt / Min / Max  : 0x181ac577 / 0x181ac577 / 0x182dd6d7
RecvQ Pkts             : 0
Sender Stats
  Sent Bytes / Pkts    : 1105865240 / 10103399
  Unacked Data        : 0
  Retransmits Slow / Fast : 192 / 0 (High:0)
  SlowStart           : 133
Receiver Stats
  Recv Bytes / Pkts    : 1105894888 / 10103203
  Out-of-Order         : 20 (High:23)
  Duplicate ACKs       : 0
  RTT / Variance (High) : 20 ms (58 ms) / 0 ms (43 ms)
TCP Connection 24.0 HA-Type:Main Pri:High Conn:0x0a8a9631
=====
Local / Remote Port    : 51164 / 3226
Duration               : 28d22h
MSS                    : 1460 bytes
ARL Min / Cur / Max    : 25000 / 25000 / 50000
ARL Reset Algo         : Reset
Send Window
  Size / Scale         : 1249280 / 9
  Slow Start Threshold : 16777216
  Congestion Window    : 16778676
  Pkts InFlight        : 0
Recv Window
  Size / Scale         : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max  : 0x3a459e77 / 0x3a459e77 / 0x3a459e77
RecvQ Nxt / Min / Max  : 0x46b0fc20 / 0x46b0fc20 / 0x46c40e20
RecvQ Pkts             : 0
Sender Stats
  Sent Bytes / Pkts    : 1105886288 / 10103344
  Unacked Data        : 0

```

```

Retransmits Slow / Fast : 171 / 0 (High:0)
SlowStart                : 122
Receiver Stats
Recv Bytes / Pkts        : 1105864848 / 10103140
Out-of-Order             : 5 (High:29)
Duplicate ACKs           : 0
RTT / Variance (High)    : 20 ms (54 ms) / 0 ms (43 ms)
TCP Connection 24.0 HA-Type:Main Pri:Control Conn:0x0a8a9739
=====
Local / Remote Port      : 57311 / 3225
Duration                 : 28d22h
MSS                     : 1460 bytes
ARL Min / Cur / Max      : 0 / 0 / 50000
ARL Reset Algo           : Reset
Send Window
Size / Scale             : 1249280 / 9
Slow Start Threshold     : 16777216
Congestion Window        : 16778676
Pkts InFlight            : 0
Recv Window
Size / Scale             : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max    : 0x447a03c6 / 0x447a03c6 / 0x447a03c6
RecvQ Nxt / Min / Max    : 0x0f08db12 / 0x0f08db12 / 0x0f1bed12
RecvQ Pkts               : 0
Sender Stats
Sent Bytes / Pkts        : 1392172536 / 12980231
Unacked Data             : 0
Retransmits Slow / Fast : 241 / 4 (High:0)
SlowStart                : 168
Receiver Stats
Recv Bytes / Pkts        : 1560151020 / 13039059
Out-of-Order             : 2 (High:11)
Duplicate ACKs           : 12
RTT / Variance (High)    : 20 ms (63 ms) / 0 ms (43 ms)
TCP Connection 24.0 HA-Type:Main Pri:Control Conn:0x0a8a973a
=====
Local / Remote Port      : 51165 / 3226
Duration                 : 28d22h
MSS                     : 1460 bytes
ARL Min / Cur / Max      : 0 / 0 / 50000
ARL Reset Algo           : Reset
Send Window
Size / Scale             : 1249280 / 9
Slow Start Threshold     : 16777216
Congestion Window        : 16778676
Pkts InFlight            : 0
Recv Window
Size / Scale             : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max    : 0x447b4df6 / 0x447b4df6 / 0x447b4df6
RecvQ Nxt / Min / Max    : 0x8e1cf9f2 / 0x8e1cf9f2 / 0x8e300bf2
RecvQ Pkts               : 0
Sender Stats
Sent Bytes / Pkts        : 1392250536 / 12980470
Unacked Data             : 0
Retransmits Slow / Fast : 240 / 4 (High:0)

```

```

SlowStart                : 176
Receiver Stats
Recv Bytes / Pkts        : 1560694216 / 13039703
Out-of-Order             : 0 (High:14)
Duplicate ACKs           : 13
RTT / Variance (High)    : 20 ms (52 ms) / 0 ms (42 ms)

```

B.4 portStatsShow

Many port statistics are maintained. To display port traffic and error statistics, use the **portstatsshow** CLI command. To get FC, VE, and GE detailed statistics about the CLI issue, use the **portstatsshow <port#>** command.

FC and VE ports use only the number <#>. GE interfaces use ge<#>. An example of the output for FC port 0 is shown in the following example:

```
7840-Side-A:FID128:admin> portstatsshow
```

Usage:

```
portstatsshow [--detail] PortNumber
```

OR

```
portstatsshow [--detail] -i <port_index | portindex_range> [-f]
```

OR

```
portstatsshow [--detail] <-slot | -s > <slot# | slotrange>
```

-i: Confirms port swap has been disabled and to give port index as operand

-x: Confirms port swap has been disabled and to give port index

in HEX format as operand

-f: ignores non-existing indexes

portindex_range: Specifies the range of port index <portindex1-portindex2>

(example: 12-14)

OR

```
portstatsshow [<SlotNumber>/]ge<PortNumber>
```

```
7840-Side-A:FID128:admin> portstatsshow ge2
```

```

ge_stat_tx_frms          542382084  GE transmitted frames
ge_stat_tx_octets        636962488472  GE transmitted octets
ge_stat_tx_ucast_frms    542237792  GE transmitted unicast frames
ge_stat_tx_mcast_frms    0  GE transmitted multicast frames
ge_stat_tx_bcast_frms    144292  GE transmitted broadcast frames
ge_stat_tx_vlan_frms     0  GE transmitted vlan frames
ge_stat_tx_pause_frms    0  GE transmitted pause frames
ge_stat_rx_frms          258254545  GE received frames
ge_stat_rx_octets        40648834562  GE received octets
ge_stat_rx_ucast_frms    258110264  GE received unicast frames
ge_stat_rx_mcast_frms    0  GE received multicast frames
ge_stat_rx_bcast_frms    144281  GE received broadcast frames
ge_stat_rx_vlan_frms     0  GE received vlan frames
ge_stat_rx_pause_frms    0  GE received pause frames
ge_err_crc               0  GE CRC Errors
ge_err_carrier           1  GE lost carrier sense
ge_err_jabber            0  GE jabbers
ge_stat_tx_octets        636962488472  GE transmitted
octets
ge_stat_tx_pkts64octets  347  GE transmitted
64byte octets

```



```

ge_stat_tx_pkts65to127octets          62670368    GE    transmitted
65to127byte octets
ge_stat_tx_pkts128to255octets          21247343    GE    transmitted
128to255byte octets
ge_stat_tx_pkts256to511octets          34952514    GE    transmitted
256to511byte octets
ge_stat_tx_pkts512to1023octets          1495311     GE    transmitted
512to1023byte octets
ge_stat_tx_pkts1024to1518octets         422016201   GE    transmitted
1024to1518byte octets
ge_stat_tx_pkts1519to2047octets          0           GE    transmitted
1519to2047byte octets
ge_stat_tx_pkts2048to4095octets          0           GE    transmitted
2048to4095byte octets
ge_stat_tx_pkts4096to9216octets          0           GE    transmitted
4096to9216byte octets
ge_stat_rx_octets                       40648834562 GE    received
octets
ge_stat_rx_pkts64octets                  399         GE    received
64byte octets octets
ge_stat_rx_pkts65to127octets            143348534   GE    received
65to127byte octets
ge_stat_rx_pkts128to255octets            43617129    GE    received
128to255byte octets
ge_stat_rx_pkts256to511octets            70946937    GE    received
256to511byte octets
ge_stat_rx_pkts512to1023octets            140918      GE    received
512to1023byte octets
ge_stat_rx_pkts1024to1518octets           200628      GE    received
1024to1518byte octets
ge_stat_rx_pkts1519to2047octets            0           GE    received
1519to2047byte octets
ge_stat_rx_pkts2048to4095octets            0           GE    received
2048to4095byte octets
ge_stat_rx_pkts4096to9216octets            0           GE    received
4096to9216byte octets
ge_stat_rx_pfc_control_frame              0           GE    Rx PFC control
frame
ge_stat_tx_pfc_control_frame              0           GE    Tx PFC control
frame
ge_stat_rx_dvlan_tag_frame                0           GE    Rx Double VLAN
tag frame
ge_stat_tx_dvlan_tag_frame                0           GE    Tx Double VLAN
tag frame
7840-Side-A:FID128:admin> portstatsshow 0
stat_wtx          217334      4-byte words transmitted
stat_wrx          77324      4-byte words received
stat_ftx          8841       Frames transmitted
stat_frx          4537       Frames received
stat_c2_frx        0         Class 2 frames received
stat_c3_frx        4537       Class 3 frames received
stat_lc_rx         0         Link control frames received
stat_mc_rx         0         Multicast frames received
stat_mc_to         0         Multicast timeouts
stat_mc_tx         0         Multicast frames transmitted

```

tim_rdy_pri	0		Time R_RDY high priority
tim_txcrd_z	0		Time TX Credit Zero (2.5Us ticks)
tim_txcrd_z_vc 0- 3:	0	0	0 0
tim_txcrd_z_vc 4- 7:	0	0	0 0
tim_txcrd_z_vc 8-11:	0	0	0 0
tim_txcrd_z_vc 12-15:	0	0	0 0
tim_latency_vc 0- 3:	1	1	1 1
tim_latency_vc 4- 7:	1	1	1 1
tim_latency_vc 8-11:	1	1	1 1
tim_latency_vc 12-15:	1	1	1 1
fec_cor_detected	0		Count of blocks that were corrected by FEC
fec_uncor_detected	0		Count of blocks that were left uncorrected by FEC
er_enc_in	0		Encoding errors inside of frames
er_crc	0		Frames with CRC errors
er_trunc	0		Frames shorter than minimum
er_toolong	0		Frames longer than maximum
er_bad_eof	0		Frames with bad end-of-frame
er_enc_out	0		Encoding error outside of frames
er_bad_os	0		Invalid ordered set
er_pcs_blk	0		PCS block errors
er_rx_c3_timeout	0		Class 3 receive frames discarded due to timeout
er_tx_c3_timeout	0		Class 3 transmit frames discarded due to timeout
er_unroutable	0		Frames that are unroutable
er_unreachable	0		Frames with unreachable destination
er_other_discard	0		Other discards
er_type1_miss	0		frames with FTB type 1 miss
er_type2_miss	0		frames with FTB type 2 miss
er_type6_miss	0		frames with FTB type 6 miss
er_zone_miss	0		frames with hard zoning miss
er_lun_zone_miss	0		frames with LUN zoning miss
er_crc_good_eof	0		Crc error with good eof
er_inv_arb	0		Invalid ARB
er_single_credit_loss	0		Single vcrdy/frame loss on link
er_multi_credit_loss	0		Multiple vcrdy/frame loss on link
other_credit_loss	0		Link timeout/complete credit loss
phy_stats_clear_ts	0		Timestamp of phy_port stats clear
lgc_stats_clear_ts	0		Timestamp of lgc_port stats clear

B.5 gePortPerfShow

To monitor the data rates through the GE interfaces, use the **geportperfshow** CLI command:

```
7840-Side-A:FID128:admin> geportperfshow
```

Throughput of GE port

ge0	ge1	ge2	ge3	ge4	ge5	ge6	ge7	ge8	
0	0	11.3m	0	11.1m	0	0	0	0	
ge9	ge10	ge11	ge12	ge13	ge14	ge15	ge16	ge17	Total
0	0	0	0	0	0	0	0	0	22.5m

B.6 switchShow

The best way to show the status of a VE_Port is by using the **switchshow** command. In the following example, VE_Port 12 shows the port type (VE), port status (online), and the connected device Remote7810. Be aware that for the GE interfaces, the notation FCIP indicates WAN. FCIP terminology remains because of older output before IP Extension.

```
Local7810:admin> switchshow
```

```
switchName:    Local7810
switchType:    178.0
switchState:   Online
switchMode:    Native
switchRole:    Principal
switchDomain:  1
switchId:      fffc01
switchWwn:     10:00:88:94:71:5c:f7:e2
zoning:        OFF
switchBeacon:  OFF
FC Router:     OFF
HIF Mode:      OFF
```

Index	Port	Address	Media	Speed	State	Proto
0	0	010000	--	N32	No_Module	FC
1	1	010100	--	N32	No_Module	FC
2	2	010200	--	N32	No_Module	FC
3	3	010300	--	N32	No_Module	FC
4	4	010400	--	N32	No_Module	FC
5	5	010500	--	N32	No_Module	FC
6	6	010600	--	N32	No_Module	FC
7	7	010700	--	N32	No_Module	FC
8	8	010800	--	N32	No_Module	FC
9	9	010900	--	N32	No_Module	FC
10	10	010a00	--	N32	No_Module	FC
11	11	010b00	--	N32	No_Module	FC
12	12	010c00	--	--	Online	VE VE-Port 10:00:88:94:71:60:98:5e
"Remote7810" (downstream)						
13	13	010d00	--	--	Offline	VE Disabled (Persistent)
14	14	010e00	--	--	Offline	VE Disabled (Persistent)
15	15	010f00	--	--	Offline	VE Disabled (Persistent)
	ge0		cu	1G	Offline	FCIP Copper Disabled
	ge1		cu	1G	Offline	FCIP Copper Disabled
	ge2		id	10G	Online	LAN
	ge3		id	10G	Online	LAN
	ge4		id	10G	No_Sync	LAN Disabled (Persistent)
	ge5		id	10G	No_Sync	LAN Disabled (Persistent)
	ge6		id	10G	Online	FCIP
	ge7		id	10G	Online	FCIP

B.7 portPerfShow

PortPerfShow is a real-time iterative CLI output that shows the rate of traffic on the port. Only FC/FICON and VE ports are shown with PortPerfShow, use geportperfshow for GE interfaces.

In this example, the two FC ports (0 and 1) are feeding into VE_Port 24 for transmission across FCIP (see Figure B-1):

```
7840-Side-A:FID128:admin> portperfshow
```

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
41.7m	41.7m	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	0	0	0	0	0	0	85.4m	0	0	0	0	0	0	0
32	33	34	35	36	37	38	39	40	41	42	43	Total			
0	0	0	0	0	0	0	0	0	0	0	0	168.9m			

Figure B-1 portPerfShow command

B.8 portErrShow

To view the number of transmitted/received frames and errors on F_Port (FC fabric ports) and VE_Port (FCIP Virtual E_Ports), use the **porterrshow** command (see Figure B-2):

```
7840-Side-A:FID128:admin> porterrshow
```

	frames tx	frames rx	enc in	enc err	GC CRC	GC CRC	too short	too long	bad CRC	enc out	disc c3	link fail	loss sync	loss sig	coll ctn	copy err	timeout tx	pos err	wrap err
0:	40.3m	661.2m	0	0	0	0	0	0	0	0	0	2	0	5	0	0	0	0	0
1:	40.3m	662.8m	0	0	0	0	0	0	0	0	0	2	0	5	0	0	0	0	0
2:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24:	1.3g	82.0m	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure B-2 portErrShow command

Redbooks

IBM b-type Gen 7 Installation, Migration, and Best Practices Guide

SG24-8497-01

ISBN 0738460478



(0.5" spine)

0.475" <-> 0.873"

250 <-> 459 pages



SG24-8497-01

ISBN 0738460478

Printed in U.S.A.

Get connected

