Redbooks
ibm.com/redbooks

Modernizing Your IT Infrastructure with IBM b-type Gen 6 Storage Networking and IBM Spectrum Storage Products

Simone Bosisio

Lee Cockrell

Steve Guendert

Abilio Oliveira

Rodrigo Jungi Suzuki

Storage

IBM

Redbooks

International Technical Support Organization

Modernizing Your IT Infrastructure with IBM b-type Gen 6 Storage Networking and IBM Spectrum Storage Products

February 2018

SG24-8415-00

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

First Edition (February 2018)

This edition applies to the products in the IBM Storage Area Networks (SAN) portfolio.

Contents

Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

Bluemix®
Cloudant®
DB2®
Dispersed Storage®
DS8000®
Easy Tier®
FICON®
FlashCopy®
Global Technology Services®
GPFS™
HyperSwap®
IBM®
IBM Cloud Managed Services®
IBM FlashCore®
IBM FlashSystem®
IBM LinuxONE™
IBM LinuxONE Emperor™
IBM Spectrum™
IBM Spectrum Accelerate™
IBM Spectrum Archive™
IBM Spectrum Control™
IBM Spectrum Protect™
IBM Spectrum Scale™
IBM Spectrum Storage™
IBM Spectrum Virtualize™
IBM Watson®
IBM z®
IBM zHyperWrite™
Informix®
Linear Tape File System™
MicroLatency®
Power Systems™
POWER8®
Real-time Compression™
Redbooks®
Redpaper™
Redpapers™
Redbooks (logo) ®
Storwize®
System Storage®
Tivoli®
Variable Stripe RAID™
Watson™
XIV®
z/OS®
z13®

The following terms are trademarks of other companies:

SoftLayer, and The Planet are trademarks or registered trademarks of SoftLayer, Inc., an IBM Company.

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.

ITIL is a Registered Trade Mark of AXELOS Limited.

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

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

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

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

Preface

This IBM® Redbooks® publication describes the challenge that most data centers face when updating and modernizing their IT infrastructure.

New business demands are driving new applications, joining, and creating in the digital world. A rich, meaningful digital experience is the key to effective engagement in today's integrated digital world. Companies are able to customize digital experiences for their employees with personalized, targeted content for fully connecting with customers, co-workers, and business partners in the most powerful and productive ways.

To achieve this, a robust infrastructure is required. Speed of access to data is one of the most important factors. The development of the flash storage devices helped with the insatiable desire for data access speed, but even that is not enough for the most demanding uses. The needs of SAN switches, servers, and software defined infrastructure (SDI) technologies are all requiring more; therefore, the bigger picture needs to be wholly analyzed to build a balanced ecosystem.

This publication can help you with planning for growth in your IT infrastructure. This publication explores the concept of modernization and considers important aspects of IT, such as SAN switches, storage systems, and software defined storage.

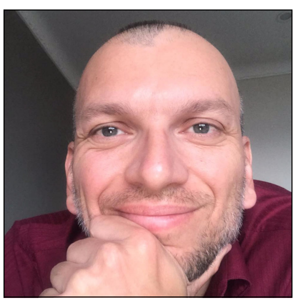
Authors

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

Simone Bosisio is a Systems Services Representative working for IBM Global Technology Services® in Switzerland. He joined IBM in 2000 and he has more than 7 years of experience in delivery of SAN and storage services. His areas of expertise include storage hardware, SAN, and storage virtualization design, solutions, and implementation. Simone holds a degree in Electronics from the Technical School (SSST Bellinzona), Switzerland.

Lee Cockrell is an IBM Technical Sales Specialist, covering several United States Federal Civilian agencies and Native American Tribal governments for IBM Federal in the Washington D.C. area. He received his B.S. in computer science from the University of Virginia, Charlottesville, VA. His current interests are in storage, cloud, and computer security. Lee joined IBM in 2010 selling storage, primarily IBM Spectrum™ Virtualize, SAN Volume Controller, IBM Storwize®, IBM XIV®, FlashSystem, and IBM DS8000®. He has worked in the storage industry since 2001, installing and configuring countless storage arrays and co-authoring expert level performance certification tests. Prior to that he was a UNIX and firewall administrator in both the public and private sectors.

Steve Guendert is a PhD. “Dr Steve” is Brocade’s Principal Director of Product Management for IBM Z products and technology. He is an industry recognized worldwide expert on ESCON, IBM FICON®, and mainframe I/O and is a published author of over 50 papers on those subjects. Steve was inducted into the Mainframe Hall of Fame in 2017. He also is an author of *Brocade Mainframe Connectivity Solutions* (from Brocade Communications Systems), and a contributing author to *Handbook of Fibre Optic Data Communication*, 4th edition. Steve has a PhD and a MS in Management Information Systems (MIS), and an MBA. He also holds the distinction of having served on both the SHARE and Computer Measurement Group (CMG) Boards of Directors.



Abilio Oliveira is an IBM Certified Expert IT Specialist and works as a Client Technical Specialist in Storage at IBM Asia Pacific. He has 23 years of IT experience. He holds a Bachelor degree in Computer Science with a Masters in Data Information Security. He has expertise in designing storage solutions with OpenStack and storage for cloud technologies. He also is a Subject Matter Expert (SME) in Storage Infrastructure Optimization studies, XIV, and DS8000 Storage Systems.

Rodrigo Jungi Suzuki is a SAN Storage specialist at IBM Brazil Global Delivery Center in Hortolandia. Currently, Rodrigo is a Subject Matter Expert (SME) account focal, and has been working with projects and support for international customers. He has 23 years of IT Industry experience with the last 8 years in the SAN Storage area. He also has a background in UNIX and IBM Informix® databases. He holds a bachelor degree in Computer Science from Universidade Paulista in Sao Paulo, Brazil. He is an IBM Certified IT Specialist, and NetApp NCDA, IBM Storwize V7000 Technical Solutions V2, and ITIL certified.

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

Deana Coble, Project Leader for this publication
International Technical Support Organization, Raleigh Center

Jon Tate, Project Executive
IBM ITSO, UK

Special thanks to Tim Werts and Silviano Gaona of Brocade Communications Systems, LLC, in the Brocade Business Unit in San Jose. Their time and technical expertise were essential to the success of this publication and their help essential in so many areas of writing and understanding.

Special thanks to David Green, of Raleigh, NC with SAN Troubleshooting and Performance, IBM Solution Central Support. His time, willingness to help, and expertise as a content contributor for this publication was essential to our success.

Future contributors welcome

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

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

ibm.com/redbooks/residencies.html

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

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806
- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm
- Stay current on recent Redbooks publications with RSS Feeds:

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

1

Introduction

An up-to-date infrastructure is an indispensable component of the modern industry. Regardless of an operation's size, the profitability of any business has a critical dependency on proper functioning of the computer systems supporting that business.

Maintaining a balance between components, lifecycles, and performance is a challenge. The benefits provided by the data centers are attainable only when they are operating flawlessly, and electronic components develop at different speeds.

To create a logical thinking path for guidance on modernizing data centers, this IBM Redbooks publication is divided into five chapters. This book describes the main components of an IT infrastructure and how they interact with each other to deliver the highest standard of quality in this ever-growing environment. This book is divided into the following chapters:

- Chapter 1, "Introduction" on page 1
- Chapter 2, "Product introduction" on page 7
- Chapter 3, "Cloud impact in the modern data center" on page 23
- Chapter 4, "Storage network modernization" on page 47
- Chapter 5, "Modernizing IBM Z storage networking" on page 91

This chapter introduces and discusses today's data centers, and how they have become the foundation for organizations to accomplish their goals. Advances of both server and flash technology have led to the digital transformation of businesses. New generations of technology have moved at different paces because not each element advances with the same timeline. This differing pace of evolving technology leads to performance bottlenecks appearing in different locations as time passes. The network is a critical component for infrastructure modernization, as SAN bottlenecks become more common due to virtualization of servers, storage, and networks. Modernizing the SAN infrastructure to match the evolving technology's needs is a critical element of today's data centers, and the need for insight into that network infrastructure continues to grow.

The following topics are covered in this chapter:

- 1.1, "Today's data center" on page 2
- 1.2, "Data center component evolution" on page 3
- 1.3, "SAN infrastructure in the data center" on page 4
- 1.4, "Network monitoring" on page 5

1.1 Today's data center

The data center environment is a compilation of servers, storage, network systems, mechanical and electrical systems, applications and tools, governance procedures, and staff. With applications such as cloud computing, analytics, high-performance computing, and cognitive computing, the data center is now the foundation of every business. These applications no longer reside on a single server, storage array, or network, but in a holistic, interwoven system residing in the modern data center.

Data center technology continues to evolve, improving the reliability, cost efficiency, and ease of management of its physical and virtual components. Recurring themes for improving data centers include: consolidation and virtualization of servers, storage, and networking; energy efficiency; improving business continuity; and meeting differing application requirements for resilience and performance.

Improving data center efficiency and management can yield tangible benefits to the organization. For example, better data center management can yield the following improvements:

- Greater investment on strategic initiatives. Staff spend more time on new projects versus maintaining infrastructure.
- Greater efficiency with more hardware managed per administrator.

There are four distinguishing characteristics of a more strategic approach to data center management:

- Optimizing the server, storage, network, and facilities assets to maximize capacity and availability
- Designing for flexibility to support changing business needs
- Using automation tools to improve service levels and availability
- Planning that aligns with the business goals and keeps those goals current

As data centers become more efficient, many have already taken out a significant portion of hard cost using consolidation and virtualization. The refocused goal of information technology and data center efficiency is on providing greater business benefits. These benefits include application availability, performance, and the ability to respond rapidly to business changes. This focus on business outcomes can result in huge payoffs for organizations where revenue generation, innovation or competitive advantage is the goal. In comparison, efficiency and cost containment are typically foundational elements of data center planning.

As cloud computing becomes more mature, the data center is an essential component of the cloud. *Cloud* is often used to refer to several similar concepts of virtual provisioning of storage, servers, network, and applications. Cloud is an amorphous term that describes some useful capabilities of new IT technologies, while glossing over the types of cloud implementations. The focus of this book is on the IBM hybrid and private cloud offerings.

IBM builds virtualization, energy efficiency, and resilience into its products, and by extension into its public cloud offerings hosted at IBM data centers. IBM Cloud (formerly IBM Bluemix®) offers both public cloud or infrastructure as a service (IaaS) offerings. In the June 2017 Gartner article ("Magic Quadrant for Data Center Outsourcing and Infrastructure Utility Services, North America"), IBM was positioned as a Leader. In the Leaders Quadrant, IBM was positioned the furthest for Ability to Execute and Completeness of Vision.

Note: Gartner, Inc. does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Data center management is often a complicated and intensive endeavor. Because of this, organizations can outsource their data center management to get better expertise or cost efficiency. These organizations should make an honest assessment of their applications and business needs, and organizational strengths. The following questions are pertinent to making that assessment:

- Does the organization have particular requirements for data privacy or security?
- Does the organization need the capacity and ownership of its own data center or would an offering like IaaS work better?
- Does the organization have the expertise in house to manage all the complexities?

These are only some of the questions to consider before deciding whether to manage your own data center environment, use an IaaS offering, or implement a public or hybrid cloud.

The evolution of the data center is one in which the destination can change as the business needs change.

1.2 Data center component evolution

The primary computing components in the data center are the servers, storage, and network. As technology has progressed over the years, incredibly rapid rates of change have improved the performance of each of the components. However, the rate of change has varied over time and component type.

Server power, as measured by CPU benchmarks, has increased in accordance with Moore's Law. Moore's Law states that the number of transistors on an integrated circuit (such as a CPU) will double every 18 - 24 months.

Storage capacities and performance have increased massively over the years, though unevenly with the mechanical nature of spinning disks. Hard drive speed and bandwidth generally have increased more slowly than hard drive capacity. This has led to the precipitous dropping over the years of the number of I/O operations per second per unit of storage (IOPS/TB ratio) of many storage environments.

The introduction of flash drives has improved this lack of performance in back-end storage. While flash-based solid state drives (SSDs) have existed for nearly 30 years, for much of that time small capacity and high cost made them uncommon in enterprise environments. Rapid adoption and technological progress continue to reduce flash cost. SSDs with SAS, SATA, or Fibre Channel (FC) protocol interfaces became more common in the early 2000s. Flash-card-based appliances such as the IBM FlashSystem line became popular in the 2010s due to the high quality, capacity, and performance even compared to SSDs.

As a result, implementing all-flash arrays today is a top priority for data center administrators for the performance, power, cooling, and footprint benefits. Also, Non-Volatile Memory Express (NVMe) over fabric introduces another major shared storage evolutionary step, further reducing latency and increasing bandwidth capabilities.

Network capabilities and transmission rates are two elements often overlooked in the data center environment. Network bandwidth over glass fiber has increased exponentially since the 1970s. Fibre Channel protocol started at 1 Gbs in 1997 and increased to 32 Gbps by 2016.

More information about fiber bandwidth is available at the IEEE Spectrum news website.

1.3 SAN infrastructure in the data center

The network has become an even more critical component for infrastructure modernization. This critical nature is related to the following demand in today's business:

- Application stability, performance, and agility
- Always on, always fast, always flexible infrastructure

Network traffic has increased vastly because of the following contributors:

- Networked application architectures
- Increasing number of networked storage and computing architectures
- Mobility of virtualized servers and storage

Growth in network capacity has accommodated this vast increase. The transmission rate of light through fiber optic glass has grown exponentially since the 1970s. However, this means that organizations must continually evaluate their network bandwidth requirements, as the ever-increasing exponential growth requirements for latency and bandwidth will overwhelm older network directors.

The storage network or storage area network (SAN) must accommodate this explosion in data capacities. The SAN infrastructure is often an overlooked piece of network bandwidth considerations. Network cables attached to a PC, or a wireless network, are seen and understood, but the SAN is buried behind layers of applications and servers. Invisible to most regular users, the SAN is crucial to bandwidth and latency for all applications at the enterprise level.

Storage networking investments are becoming a critical top priority due to the adoption of high-throughput, solid-state and flash storage. Storage performance bottlenecks are moving out of arrays and into the storage network. The "Future of Storage Protocols" report (from Gartner, Inc.) notes the following information:

> *"IT leaders planning to upgrade their storage networking need to use 16 Gbs/sec FC for block access and 25/50 Gbps for file and block access today, with a migration plan for moving to 32 Gbs/sec FC and 100 Gigabit Ethernet (GbE) in the next 5 years."*

Gartner clients can read the report at the following web page:

`https://www.gartner.com/document/3357441`

Most SAN infrastructures use Fibre Channel protocol and fiber transmission media as a purpose-built network specifically for the demands of storage traffic. The throughput, latency, bandwidth, and ease of administration of Fibre Channel make it a beneficial choice for the modern data center and enterprise storage. These same reasons make Fibre Channel SAN a natural choice for virtualized infrastructures, and private and hybrid clouds. Today Fibre Channel continues a nearly 20 year run of being a popular choice for enterprise storage. Fibre Channel is expected to continue as the primary storage protocol in the data center for the next ten years. Other storage protocols (including iSCSI, FCoE, NFS, various object storage) and physical media (copper, InfiniBand) all have use cases where they fit best based on scalability, cost, availability requirements, performance, and accessibility.

Figure 1-1 compares storage protocols from Gartner, Inc.

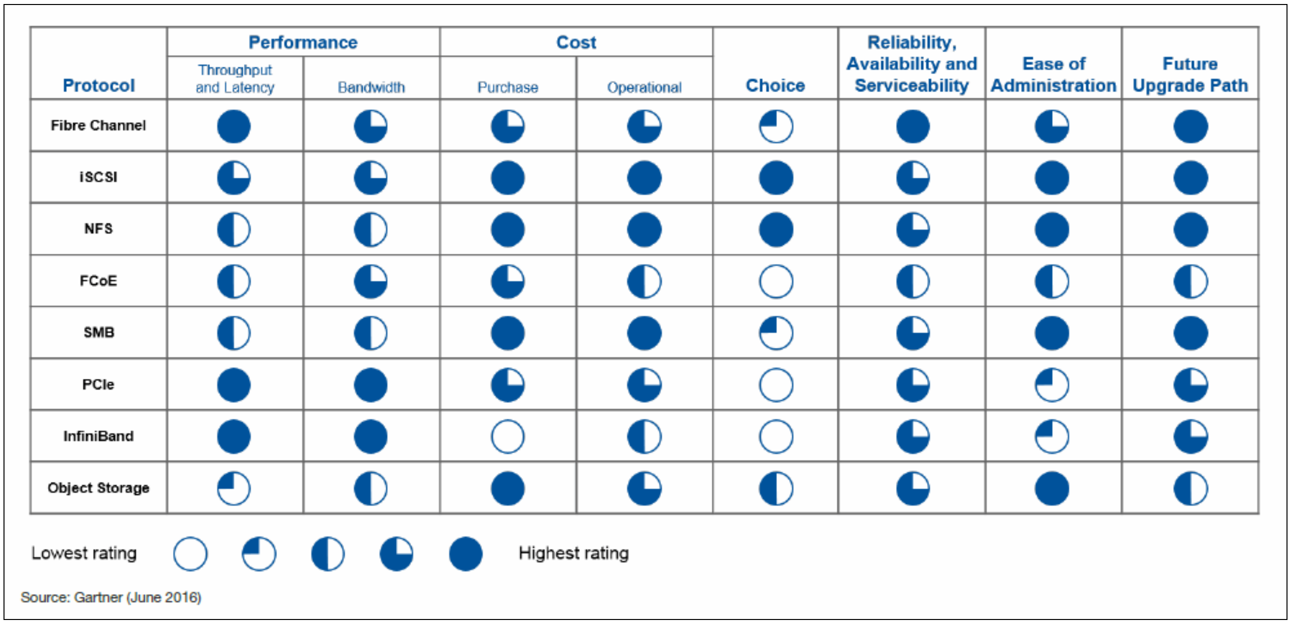


Figure 1-1 Gartner, Inc., "The Future of Storage Protocols" report from July 2016

Figure 1-1 shows the key attributes and relative comparisons between the most common storage protocols used in the data center. A high purchase and operational cost rating denotes low purchase and operational costs, compared with a low cost rating, which denotes high purchase and operational costs.

Note: FICON also runs over Fibre Channel (FC) protocol, uses the same FC directors, and has the same characteristics as SCSI over FC. The difference is that FICON primarily connects storage to servers running mainframe operating systems such as IBM Z. The two protocols differ, but are generally similar enough that much of what is said of Fibre Channel in this book can be applied separately to FICON.

1.4 Network monitoring

With ever-increasing requirements for more bandwidth, lower latency, and high reliability, organizations need intelligent software for visibility into the SAN network. A key reason for this visibility need is the ever-increasing server virtualization density and low latency of all-flash-arrays. The high number of virtual machines (VMs) per physical server has greatly increased the number of virtual machines running over the same SAN fabrics. By increasing visibility into the network traffic, administrators can delineate between the virtual machines, and thus the applications. This allows administrators to better monitor and troubleshoot application issues.

Similarly, faster application response times due to lower latency from flash arrays also increases the demand on network monitoring and troubleshooting tools. For example, the demand required to address any slowing down in application performance to meet service level agreements (SLAs). The combined impact of increased network size, higher virtual machine densities, greater configuration complexity, and elevated expectations for application performance means that the human administrators need help.

Analytic software to monitor the network, and to alert or take action if thresholds are exceeded, is a piece of this necessary support. Storage network fabric visibility tools must monitor bandwidth, health, and latency. The monitoring needs to reach down to the VM granularity to ensure SLAs are met and to help troubleshoot issues in a proactive rather than reactive fashion.

This monitoring can contribute to knowledge about the health and efficiency of the network, and by extension, the data center. The most effective means to measure the efficiency of data center operations is to take a holistic approach that considers multiple measures across all elements.

2

Product introduction

This chapter introduces products and advanced storage options related to the concepts of modernization presented in this book, such as the new generation of Storage Area Network (SAN) b-type Gen 6 switches and directors. The products introduced are designed for high performance and speed of delivery to support revolutionizing and modernizing data centers. This chapter also provides a high-level overview of SAN devices and products encountered in SAN environments.

The following topics are covered in this chapter:

2.1 IBM b-type Gen 6 storage

The IBM b-type Gen 6 device connects both servers and storage in the same network. This device is dedicated to moving storage data traffic. IBM b-type Gen 6 is the latest generation Fibre Channel technology that is revolutionizing data centers.

It was designed for network infrastructures that need to deliver high performance, virtualization, scalability, cloud architectures, and all-flash based storage for the data center.

Some of the qualities and capabilities of the IBM b-type Gen 6 SAN product family are included in the following list:

- Building scalable storage networks designed to serve mission-critical business applications
- Delivering large amounts of data faster
- Enabling high demand workloads for hyper-scale virtualization and large cloud infrastructures
- Providing optimum performance from all-flash-array storage

2.1.1 IBM b-type Gen 6 Fibre Channel

The b-type Gen 6 Fibre Channel goes beyond high speeds as new technological innovations such as IO Insight and VM Insight capabilities are added to the larger Fabric Vision tool set. Together these tools and features increase network visibility through monitoring and provide both diagnostics and proactivity improvements that help in avoiding performance degradation.

IBM b-type Gen 6 Fibre Channel with 32/128 Gbps links increases I/O performance and increases the speed of business operations eliminating I/O bottlenecks. This innovation can support double the VM density of the Gen 5 Fibre Channel, supporting large scale VM deployments.

The Gen 6 Fibre Channel has these features, among others:

- 32 Gbps performance, ideal for all flash-based storage, and 128 Gbps for high speed inter-chassis and inter-switch links.
- Forward error correction (FEC) ensures higher reliability when running on high speed networks.
- Energy Efficient Fibre Channel reduces power consumption.
- N_Port ID Virtualization (NPIV) provides scalability for the SAN and increases the numbers of server virtualization deployments.
- Investment protection with support for both legacy SAN infrastructures and ready for NVMe storage arrays.
- Enhanced security with hardware-enforced in-flight encryption.
- Supports multiple protocols including SCSI for open systems, FICON for mainframe systems, and NVMe for future flash storage arrays.

2.1.2 Product Overview

The IBM Storage Networking b-type portfolio starts with the SAN24B-6 for entry to midrange storage solutions. It grows on demand from 8 to 24 ports, offers simple deployment, high performance from 4 to 32 Gbps, and is built-in NVMe ready. This switch is ideal for small solutions needing growth capacity. The efficient 1U form factor design provides lower cost leading space utilization for simple scalability and consolidation.

Figure 2-1 shows the IBM Storage Networking SAN24B-6 switch.

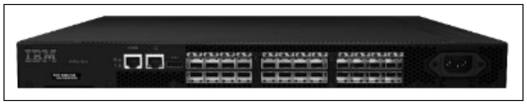


Figure 2-1 SAN24B-6

For more information, see the IBM Storage Networking SAN24B-6 web page.

The next class of switch is the SAN64B-6 that delivers performance across 48 ports supporting up to 32 Gbps and four Q-Flex ports that support up to 128 Gbps. Each Q-Flex port can also be used as four 32 Gbps ports by using breakout cables which then enables a total of 64 ports on the SAN64B-6 in a 1U form factor. The switch processes up to 100 million input/output operations per second (IOPS) and is ideal for I/O-intensive flash storage and workloads. The SAN64B-6 also support fully redundant and hot-swap power and cooling to support enterprise level, mission-critical applications.

Figure 2-2 shows the IBM Storage Networking SAN64B-6 switch.

Figure 2-2 SAN64B-6

For more information, see the IBM Storage Networking SAN64B-6 web page.

The last unit in the family of switches are the SAN256B-6 and SAN512B-6 directors. These directors are purpose-built for the Gen 6 Fibre Channel. They support high performance workloads (1 billion input/output operations per second (IOPS)). Designed with nondisruptive, hot-pluggable components and no single-point-of-failure (SPOF), the SAN512B-6 and SAN256B-6 are enterprise class directors. These directors are made for mission-critical application demands, highly virtualized data centers, and non-stop environments.

As of the publication of this book, the SAN directors can have the following configurations:

- SAN256B-6 offers up to 192 x 32 Gbps Fibre Channel ports and 16 additional 128 Gbps UltraScale ICL ports, designated for midsize enterprise systems
- The SAN512B-6 offers up to 384 x 32 Gbps Fibre Channel ports, 32 additional 128 Gbps IBM UltraScale ICL ports, 16 Tbps total chassis bandwidth, and is designated for large enterprise systems

Blades with different densities and capabilities are planned for the future.

Figure 2-3 shows the IBM Storage Networking SAN256B-6 and SAN512B-6 directors.

Figure 2-3 SAN256B-6 and SAN512B-6 Directors

For more information about SAN256B-6 and SAN512B-6, see the IBM Storage Networking b-type Gen 6 directors web page.

2.1.3 Fabric OS features

Fabric OS (known as FOS) is the operating system and the core capabilities for all Brocade Fibre Channel switches and the Brocade Fibre Channel directors.

To get the latest FOS V8.X information, such as software features, list of supported devices, compatibility, interoperability, upgrades, and important notes, read the IBM SAN b-type Firmware Version 8.x Release Notes.

2.1.4 Fabric Vision technology

Fabric Vision is a tool set that provides advanced diagnostics, monitoring, and high visibility across the storage network. By using Fabric Vision, organizations can operate in the following beneficial ways:

- Simplify monitoring for mission-critical applications
- Increase operational stability
- Reduce costs
- Streamline SAN administration
- Increase fabric resiliency
- Maximize investments

Fabric Vision quickly detects bottlenecks on ports, inter-switch link (ISL) congestion, and high latency in the fabric level. It also provides real-time bandwidth consumption data by server and storage devices, and integrates with Brocade Network Advisor to streamline monitoring.

Fabric Vision, when combined with VM Insight and IO Insight, becomes more powerful and innovative. These features are included with b-type Gen 6 and provide through Fabric Vision greater abilities for administrators. Administrators, through monitoring and diagnostic capabilities, can proactively detect abnormal behaviors in storage networks and gain high visibility across the storage network.

VM Insight

VM Insight monitors virtual machines (VM) across the storage fabric, identifies potential VM issues, and provides insights to tune VM infrastructure and the physical device. VM Insight enables administrators to set application performance thresholds and identify which layer is responsible for the degraded performance (VM or the physical device). Compliance with service level agreements and objectives compliance is further supported with VM monitoring.

IO Insight

IO Insight is an application-specific integrated circuit (ASIC) built-in capability that gathers I/O statistics. IO Insight complements Fabric Vision technology with deep visibility on storage I/O performance. This helps ensure operational stability, identify the source of performance degradation, monitor I/O workloads, and identify latency and congestion. IO Insight provides these metrics from device ports on Gen 6 platforms.

Note: Although many Fabric Vision features are offered on Gen 5 products, the IO Insight and VM Insight features are available only with Gen 6 Fibre Channel switches. Additionally, VM Insight will require support for specific hypervisor, host bus adaptor (HBA) driver, and storage array levels.

2.1.5 SAN IBM b-type family and name convention

This section provides SAN b-type Gen 6 product names and the corresponding product names or models within Brocade Communications Systems, LLC.

Table 2-1 lists the b-type family products and Brocade models.

Table 2-1 Product names and models

IBM name	IBM machine type and model	Brocade model
SAN512B-6 Director	8961-F08	Brocade X6-8
SAN256B-6 Director	8961-F04	Brocade X6-4
SAN64B-6 Switch	8960-F64-N64	Brocade G620
SAN24B-6 Switch	8960-F24	Brocade G610

For the latest information, see the IBM Storage Networking SAN b-type (Brocade) family web page.

2.2 IBM DS8000 family

The IBM System Storage® DS8000 series is designed for high performance, reliability, consistent system uptime, and high-capacity disk storage. The DS800 works with a wide range of server operating environments and the IBM Z mainframe. The DS8000 offers specialized advanced functions optimized for IBM Power Systems™ and IBM Z servers.

2.2.1 Overview of the product

The IBM DS8000 family is a high-performance, high-capacity, highly secure, and resilient series of disk storage systems. The DS8884 and DS8886 are members of the DS8880 family and can be configured as hybrid models or as all-flash storage. As of the time of this publication, DS8880 models are the latest and most advanced model of the DS8000 family.

The IBM DS8880 family includes the following three high-performance models:

- DS8884: Business class
- DS8886: Enterprise class
- DS8888: Analytics class

Figure 2-4 shows the IBM DS8884 40U Rack and the expanded DS8886/DS8888 46U Rack.

Figure 2-4 The IBM DS8000 family

The DS8880 architecture is server-based, the engines behind this advanced storage are dual powerful IBM POWER8® processor-based servers. These servers manage the cache to streamline disk I/O, thus providing high performance and throughput. The DS8880 models come standard with encryption-capable disk drives, encryption-capable solid-state drives (SSDs), and high-performance flash cards.

Figure 2-5 shows the DS8 all-flash family.

Figure 2-5 DS8 all-flash family

The DS8880 is available in several POWER8 processor options, from DS8884 single 6-core to a DS8888 48-core. Memory configurations are available from 64 GB to 2 TB.

The IBM DS8880 family offers enhanced connectivity with four-port 16 Gbps and four or eight-port 8 Gbps Fibre Channel/IBM FICON host adapters. Each host adapter port can be configured independently, based on the host adapter speed. The 16 Gbps can use Fibre Channel Protocol (FCP) or FICON protocol and 8 Gbps can use Fibre Channel Arbitrated Loop (FC-AL), FCP, or FICON protocol.

High-Performance FICON for IBM z® Systems (zHPF) is an enhancement to the IBM FICON architecture. zHPF unloads I/O management processing from the IBM Z channel subsystem to the storage host adapter and controller. zHPF is an optional feature of the IBM Z server and of the DS8000 model. Another option, IBM zHyperWrite™ technology, is designed to help accelerate IBM DB2® log-writes in synchronous data replication technologies. IBM zHyperWrite technology can reduce DB2 commit latency by up to 66%.

All capabilities are further improved with the second generation of High-performance Flash Enclosures (HPFE). The models DS8884 and DS8886 accept HPFEs for installation in the base frame and the first expansion frame. The model DS8888F accepts HPFEs for installation in the base frame and the first and second expansion frames. One HPFE (Gen 2) pair can hold 16, 32, or 48 flash cards, with capacities of 400 GB, 800 GB, 1.6 TB, or 3.2 TB.

Figure 2-6 shows the IBM High-performance Flash Enclosure Gen 2.

Figure 2-6 IBM High-performance Flash Enclosure Gen 2

Note: HPFE Gen 2 is described in *DS8880 High-Performance Flash Enclosure Gen2*, REDP-5422.

2.2.2 IBM DS8000 models and expansions

Table 2-2 shows the models and expansions that are available in the DS8880 family.

Table 2-2 Shows the DS8880 models and expansions.

IBM DS8880 family	Base/single or three-phase power	Expansion	All-flash model
DS8888F	Model 988 (three-phase power)	Model 88E	DS8888F
DS8886For Hybrid	Model 986 (three-phase power)	Model 86E	DS8886F
DS8886	Model 985 (single-phase power)	Model 85E	DS8886F
DS8884	Model 984 (single-phase power)	Model 84E	DS8884F

For more information, see the IBM DS8880 hybrid storage web page.

2.2.3 Supported environments

The DS8000 supports several platforms such as IBM Power Systems, IBM Power i, IBM System p, IBM Z servers, Intel, and so on. The current list of supported platforms is available on the IBM System Storage Interoperation Center (SSIC) web page.

2.3 IBM Flash Storage

IBM Flash Storage is able to bring flash technology to bear from the entry space all the way through mainframe environments. The IBM Flash Storage offering includes the following products:

- IBM FlashSystem® 900: Purpose-built flash hardware using IBM MicroLatency® Modules with FlashCore Technology.
- IBM FlashSystem A9000 and A9000R: Based on the IBM Spectrum Accelerate software stack, they both include FlashSystem 900.
- IBM FlashSystem V9000: A combination of IBM FlashSystem 900 and IBM SAN Volume Controller. It includes the IBM Spectrum Virtualize software stack and is able to virtualize external storage.
- IBM DS8880: The flagship product of IBM storage handles multiple environments including mainframe. The DS8886 and DS8884 arrays may contain a large number of SSDs and HPFE Flash Drawers. The DS8888, DS8886F, and DS8884F are HPFE-only arrays.
- IBM Storwize V7000 and V5000: These also include the IBM Spectrum Virtualize software stack and can include SSDs or hybrid SSD/HDD drive configurations. The Storwize V7000F and V5030F are all-flash arrays.

Figure 2-7 shows the IBM Flash Storage offerings.

Figure 2-7 The IBM Flash Storage options

2.3.1 The right flash for the right task

The IBM System Storage family provides flash technology to all environments (entry level to mainframe). The chart in Figure 2-8 is can help you choose the IBM FlashSystem that best fits the needs and the expectations in terms of general requirements.

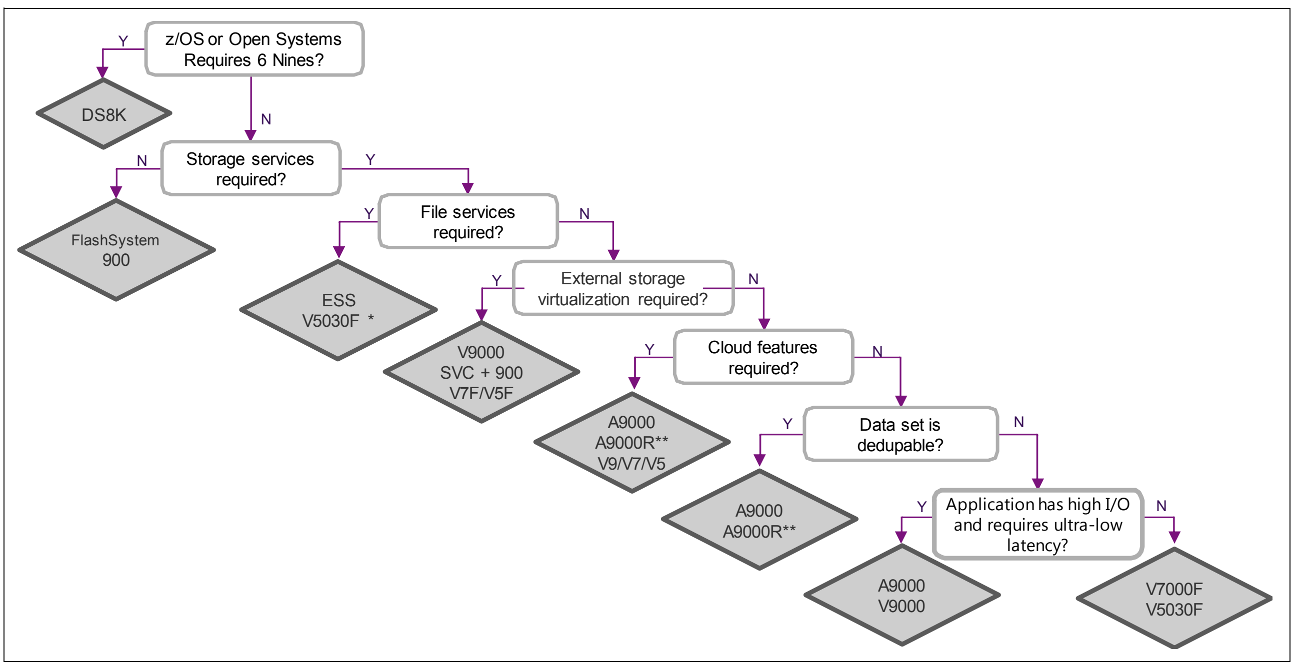


Figure 2-8 Flash decision tree

2.3.2 3D NAND technology

The actual commercial size of the 2D Planar NAND is 15nm and it is reaching its physical limits. An innovative approach, that allows for cost reduction and increases in capacity and density, is 3D implementation. Implemented in three dimensions, this technology takes advantage of stacking to increase the density. Another benefit of 3D multilevel cell (MLC) and triple-level cell (TLC) is the use of a less dense 19 - 20 nm lithography. This advancement reduces cell-to-cell interference enabling new programs and read algorithms for faster programming speeds with higher endurance.

The 3D NAND technology has the following advantages:

- More density
- More reliability due to lower cell-to-cell interference
- More performance with faster programming algorithms
- More power efficiency driven by the programming reduction

As of this publication, the 256 Gb multilevel cell and the 384 Gb triple-level cell 3D NAND are available.

Figure 2-9 shows 2D planar NAND and 3D NAND.

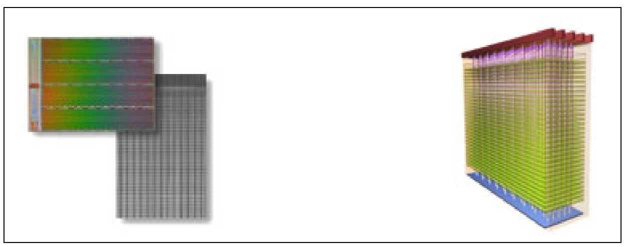


Figure 2-9 The 2D planar NAND and 3D NAND

2.3.3 IBM FlashCore technology

IBM FlashCore® technology, used in building the Flash System 900 model AE3, is the first array in the industry with the custom triple level cell (3D TLC) flash module. The FlashSystem 900 model AE3 is the core of IBM Systems Flash Storage offerings portfolio.

Figure 2-10 shows IBM FlashCore technology.

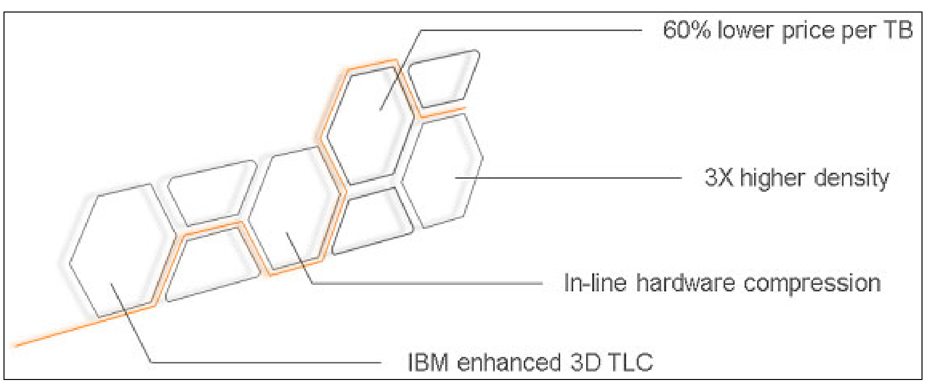


Figure 2-10 IBM FlashCore technology

The IBM FlashCore technology has the following advantages:

- Optimized 3D TLC flash card design.
- Hardware accelerated I/O:
  - Inline high-speed hardware compression.
  - Improved economics with no performance impact.
- Advanced flash management:
  - Asymmetrical wear leveling and sub-chip tiering:
    - Health binning: Up to 57% improvement in endurance.
    - Heat segregation: Up to 45% reduction in write amplification.
  - Variable voltage levels; the ideal voltage levels are set in order to minimize errors.
  - IBM Enhanced ECC.
- Maximum level of system protection:
  - IBM Patented Variable Stripe RAID™ protects from flash chip or sub-chip issues.
  - System-level RAID protects against abrupt module failure and controller failure.

2.3.4 IBM FlashSystem 900 model AE3

This section provides information about the IBM FlashSystem 900 AE3 and the key differences from the previous models (900 AE2 and 840 AE1).

FlashSystem MT-M 9840 and 9843

Table 2-3 shows the different models of IBM FlashSystem 840 and 900 and the related warranty periods. MT in the table refers to machine type.

Table 2-3 FlashSystem 9840/9843 models

Description	MT-M	Warranty
IBM FlashSystem 840	9840-AE1 9843-AE1	One year limited warranty, 24x7 Three year limited warranty, 24x7
IBM FlashSystem 900	9840-AE2 9840-AE3 9843-AE2 9843-AE3 9843-UF3	One year limited warranty, 24x7 One year limited warranty, 24x7 Three year limited warranty, 24x7 Three year limited warranty, 24x7 Three year limited warranty, 24x7

The FlashSystems AE1, AE2, AE3, and UF3 storage enclosures are similar. All of these models use the same chassis and use many of the same hardware components. Also, they can run the same software version.

Tip: No upgrade path is available to convert a FlashSystem enclosure to a different model.

Table 2-4 shows the models of IBM FlashSystem 900 and related flash module capacity. The AE1 flash module is used in the FlashSystem 840 and not included in the table.

Table 2-4 Comparison chart between Flash module capacity

FlashSystem model	Flash technology	Flash module capacity
AE2	MLC	1.2/2.9/5.7 TB
AE3	TLC	3.6/8.5/18 TB

The model EA3

The IBM FlashSystem 900 model AE3 is the next generation of IBM Flash Storage bringing a greater level of efficiency. The AE3 model is the first array in the industry with a 3D TLC NAND flash module and has three times greater capacity density compared to the previous system. The system has a 60% lower price per TB (with a compression ratio of 2:1) and up to 11 times the native chip endurance.

Figure 2-11 shows the IBM FlashSystem 900 AE3.

Figure 2-11 IBM FlashSystem 900 AE3

The FlashSystem 900 model AE3 is configurable with capacity points as low as 14.44 TB usable after RAID 5 protection to a maximum of 180 TB (effective 220 TB with embedded inline compression).

IBM will continue to sell the AE2 models with 1.2 TB MLC MicroLatency cards for clients who need smaller amounts of flash (up to 57 TB and does not include embedded inline compression).

Important new model characteristics

The FlashSystem 900 model AE3 has these characteristics:

- The new model *AE3* is configurable with the 2.6/5.8/18 TB IBM MicroLatency module type with usable capacity for scalability and flexibility.
- The new model *EA3* offers local key management or Security Key Lifecycle Manager (SKLM) hardware-based AES 256 data-at-rest encryption with no performance impact.

- The new model *EAD* has embedded inline compression (EiC) with virtually no added latency.
- The new model *EA3* has a new graphical user interface (GUI) and the IBM Spectrum Control™ Base Edition.

Specifications

Table 2-5 lists IBM FlashSystems models and their main feature and performance data.

Table 2-5 Comparison chart between model AE2 and model AE3

Specification	9840-AE2, 9843-AE2	9840-AE3, 9843-AE3
Form factor	2U rack-mounted enclosure	2U rack-mounted enclosure
IBM MicroLatency module quantity	Up to 12 IBM MicroLatency modules in increments of 4, 6, 8, 10, or 12 for the 1.2 TB module, and 6, 8, 10, or 12 for the 2.9 TB or 5.7 TB modules. Modules of either 1.2 TB, 2.9 TB, or 5.7 TB cannot be intermixed.	Up to 12 IBM MicroLatency modules in increments of 6, 8, 10, or 12 for the 3.6 TB module, and 8, 10, or 12 for the 8.5 TB or 18 TB modules. Modules of either 3.6 TB, 8.5 TB, or 18 TB cannot be intermixed.
Flash Type	IBM-enhanced MLC	IBM-enhanced 3D TLC
Minimum and maximum capacity configurable	2.2 TB / 57 TB	14.44 TB / 180 TB (effective 220 TB with Embedded inline Compression)
IBM MicroLatency module protection	ECC error correction, variable stripe RAID data protection, overprovisioning, and IBM two-dimensional Flash RAID	ECC error correction, variable stripe RAID data protection, overprovisioning, and IBM two-dimensional Flash RAID
RAID support	RAID 5	RAID 5
Connectivity options	► 16 x 8 Gbps Fibre Channel ► 8 x 16 Gbps Fibre Channel ► 16 x 10 Gbps FCoE ► 16 x 10 Gbps iSCSI ► 8 x 40 Gbps quad data rate (QDR) InfiniBand	► 16 x 8 Gbps Fibre Channel ► 8 x 16 Gbps Fibre Channel ► 16 x 10 Gbps FCoE ► 16 x 10 Gbps iSCSI ► 8 x 40 Gbps quad data rate (QDR) InfiniBand
Maximum bandwidth	10 GBps Read (100%, sequential), 4.5 GBps Write (100%, sequential)	10 GBps Read (100%, sequential), 4.5 GBps Write (100%, sequential)
Read IOPS	1,100,000	1,100,000
Write IOPS	600,000	600,000
Read latency	155 micro seconds	155 micro seconds
Write latency	90 micro seconds	95 micro seconds
Encryption	Local key Manager (AES-XTS 256) V1.5 or later: Local key and IBM Security Key Lifecycle Manager (AES-XTS 256)	Local key and IBM Security Key Lifecycle Manager (AES-XTS 256)

Related information

For more information about the IBM FlashSystem 900 product and implementation, see the following publications:

- *FlashSystem 900 Product Guide*, TIPS1261
- *Implementing IBM FlashSystem 900*, SG24-8271
- *IBM FlashSystem A9000 Product Guide*, REDP-5474
- *IBM FlashSystem A9000R Product Guide*, REDP-5475
- *IBM FlashSystem A9000 and A9000R Business Continuity Solutions*, REDP-5401
- *IBM Hyper-Scale Manager for IBM Spectrum Accelerate Family: IBM XIV, IBM FlashSystem A9000 and A9000R, and IBM Spectrum Accelerate*, SG24-8376
- *IBM FlashSystem A9000 and IBM FlashSystem A9000R Architecture and Implementation*, SG24-8345
- *IBM FlashSystem A9000, IBM FlashSystem A9000R, and IBM XIV Storage System: Host Attachment and Interoperability*, SG24-8368

2.4 IBM Spectrum Storage software-defined family

The IBM Spectrum Storage software-defined family allows migration of storage technologies to enterprise-grade, open, and scalable hybrid cloud environments. The IBM Spectrum Storage software-defined family portfolio includes the following solutions to manage, control, and orchestrate data types:

- IBM Spectrum Control: Analytics-driven data management that helps to reduce storage costs by up to 50%. Based on technology from IBM Tivoli® Storage Productivity Center and the management layer of Virtual Storage Center (VSC).
- IBM Spectrum Protect™: Optimized data protection that helps to reduce backup costs by up to 38 percent. Based on technology from Tivoli Storage Manager.
- IBM Spectrum Archive™: Supports fast data retention to reduce total cost of ownership (TCO) for archive data by up to 90%. Based on technology from IBM Linear Tape File System™ (LTFS).
- IBM Spectrum Virtualize™: Enhances the virtualization of mixed environments to store up to 5x more data. Based on technology from SAN Volume Controller.
- IBM Spectrum Accelerate™: Accelerates deployment time of Enterprise storage for cloud from months to minutes. Based on technology from Software from XIV System.
- IBM Spectrum Scale™: Supports high-performance and high scale storage for unstructured data. Based on technology from Elastic Storage Server (IBM GPFS™).

More information about these products and software-defined storage is in 3.6, “Software-defined storage (SDS)” on page 34.

2.4.1 IBM Spectrum Virtualize

IBM Spectrum Virtualize software-defined storage is the heart of the IBM Spectrum Storage™ family, which consists of these products:

- IBM SAN Volume Controller
- IBM Storwize family
- IBM FlashSystem V9000

IBM Spectrum Virtualize supports more than 400 IBM and non-IBM storage systems. It enables changes to the underlying storage without disruption to hosts or applications, and it provides continuous data availability in case of hardware failure, power failure, connectivity failure, or other disaster recovery scenarios.

Important characteristics

Some of the most powerful IBM Spectrum Virtualize characteristics that boost capacity and performance are as follows:

- IBM Spectrum Virtualize Real-time Compression™ can store up to 5 times more data, reducing the cost per TB.
- The IBM Spectrum Virtualize Easy Tier® algorithm can boost performance up to 3 times with only 5% flash capacity.

Specification

The IBM Spectrum Virtualize solution is available as part of SAN Volume Controller family system or as a software solution only. Table 2-6 shows the differences between both solutions based on various attributes.

Table 2-6 IBM Spectrum Virtualize at a glance

Specification	IBM Spectrum Virtualize as part of SAN Volume Controller or Storwize family systems	IBM Spectrum Virtualize software only
Optional functions	Purchased separately	All-inclusive
Licensing approach	Tiered cost per TB (SAN Volume Controller) or per enclosure (Storwize family) Perpetual and monthly licensing options	Simple, flat cost per capacity Perpetual and monthly licensing options
Platforms	Installed on SAN Volume Controller or Storwize family platforms	Installed on third-party x86 platforms
Hardware	Runs as an integrated system	Runs on supported x86 servers
License portability	Transportable among SAN Volume Controller engines Tied to Storwize hardware	Transportable among different supported x86 servers
Reliability, availability, and serviceability (RAS)	Integrated RAS capabilities	Flexible RAS: server hardware diagnostics; software RAS
Service	IBM support for hardware and software	IBM support for software Hardware serviced by hardware vendor
Third-party hardware, supported	Not applicable	Lenovo System x3650 M5 Supermicro SuperServer 2028U-TRTP+
Supported storage systems	More than 400 IBM and non-IBM	More than 400 IBM and non-IBM

Related information

For more information about IBM software-defined storage and implementation, see the following publications:

- *IBM Software-Defined Storage Guide*, REDP-5121
- *Implementing IBM Spectrum Virtualize software only*, REDP-5392
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V7.8*, SG24-7933

3

Cloud impact in the modern data center

Before focusing on SAN infrastructure modernization, having an understanding of cloud computing is helpful.

The following topics are covered in this chapter:

3.1 Cloud computing key characteristics

As an emerging innovative business model for information technology, cloud computing is attracting the attention of organizations and infrastructure architects. This attention is due to the benefits and flexibility that cloud computing enables.

The National Institute of Standards and Technology (NIST) code SP 800-145 provides the following definition:

> *"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (such as networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."*

For a service to be considered a cloud service, NIST describes the following essential characteristics:

- On-demand self-service: A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.
- Broad network access: Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (such as mobile phones, tablets, notebooks, and workstations).
- Resource pooling: The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with various physical and virtual resources dynamically assigned and reassigned according to consumer demand. A sense of location-independence exists in that the customer generally has no control or knowledge over the exact location of the provided resources but might be able to specify location at a higher level of abstraction (such as country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth.
- Rapid elasticity: Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
- Measured service: Cloud systems automatically control and optimize resource use by leveraging a metering capability, typically this is done on a pay-per-use or charge-per-use basis, at some level of abstraction appropriate to the type of service (such as storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

For more information about these key characteristics, see The NIST Definition of Cloud Computing publication.

3.2 Cloud computing service models

This section describes capabilities of the following cloud computing service models:

- Infrastructure as a service (IaaS)
- Platform as a service (PaaS)
- Software as a service (SaaS)

3.2.1 Infrastructure as a service (IaaS)

The IaaS model is the simplest for service providers to provision. IaaS can include many elements, such as compute, storage, and network.

Each of these elements is provisioned in an elastic fashion. An IaaS user can deploy and run their chosen software including operating systems and applications. Users do not need to manage or control the underlying cloud infrastructure to have control over the operating systems, storage, and deployed application. Also limited control exists over select networking components, such as host firewalls.

Examples of commercial implementations of IaaS include IBM SoftLayer®, IBM Cloud Managed Services®, and IBM Cloud Managed Backup.

3.2.2 Platform as a service (PaaS)

The PaaS model includes services that build on IaaS services. PaaS adds value to the IaaS services by providing a platform on which the cloud users can provision their own applications, or conduct application development activities. The user does not need to manage the underlying cloud infrastructure (network, storage, operating systems), but can control configuration of the provisioned platform services. The following services are provisioned in PaaS models:

- Middleware
- Application servers
- Database servers
- Portal servers
- Development runtime environments

Examples of commercial implementations of PaaS environments include IBM Cloud (formerly IBM Bluemix) and IBM Cloudant®.

3.2.3 Software as a service (SaaS)

The SaaS model provides software services that are complete applications that are ready to use. The cloud user simply connects to the application, which is running at a remote location. The user might not know where the system is located. The cloud service provider is responsible for managing the cloud infrastructure, the system on which the application is running, and the application itself. This approach eliminates the need for the users to install and run the application on their own computers; significantly reducing the need for maintenance and support. SaaS can include online services, such as video on demand, sales CRM models, and email services.

Examples of commercial implementations of SaaS environments include IBM Watson® Analytics, IBM API Management on Cloud, and IBM Payment Systems.

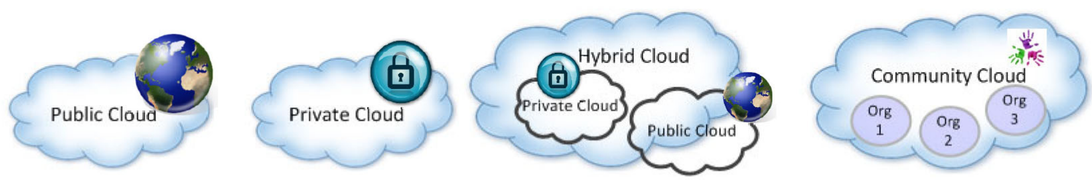
3.3 Cloud computing delivery models

From the data-location perspective, cloud computing supports multiple models that can deliver as many capabilities as required by the solution that is being designed.

This section describes the following cloud computing delivery models:

- Public cloud
- Private cloud
- Hybrid cloud

These delivery models are illustrated in Figure 3-1.



Host	Provider	Enterprise, third Party	Enterprise, third Party	Community, third Party
Owner	Provider	Enterprise	Enterprise	Community
Access	Internet	Intranet, VPN	Intranet, VPN	Intranet, VPN
Users	Public Individuals Organizations	Business Units	Business Units	Community Members

Figure 3-1 Cloud delivery models

3.3.1 Public cloud

A public cloud is one in which the cloud infrastructure is made available to the general public or a large industry group over the Internet. The infrastructure is not owned by the user, but by an organization that provides cloud services. Services can be provided either at no cost, as a subscription, or as a pay-as-you-go model.

An example of public cloud is IBM SoftLayer.

3.3.2 Private cloud

A private cloud refers to a cloud solution where the infrastructure is provisioned for the exclusive use of a single organization. The organization often acts as a cloud service provider to internal business units. These business units obtain all the benefits of a cloud without having to provision their own infrastructure. By consolidating and centralizing services into a cloud, the organization benefits from centralized service management and economies of scale.

A private cloud provides an organization with some advantages over a public cloud. The organization gains greater control over the resources that make up the cloud. In addition, a private cloud is ideal when the type of work being done is not practical for a public cloud because of network latency, security, or regulatory concerns.

A private cloud can be owned, managed, and operated by the organization, a third party, or a combination of sources. The private cloud infrastructure is usually provisioned on the organization's premises, but it can also be hosted in a data center that is owned by a third party.

IBM uses the term *local* when referring to on-premises private clouds that are owned, managed, and operated by the organization. The term *dedicated* is used by IBM when referring to off-premises third-party-managed private clouds.

Hybrid cloud

A hybrid cloud, as the name implies, is a combination of various cloud types (public, private, and community). Each cloud in the hybrid mix remains a unique entity, but is bound to the mix by technology that enables data and application portability.

The hybrid approach allows a business to take advantage of the scalability and cost-effectiveness of off-premises third-party resources without exposing applications and data beyond the corporate intranet. A well-constructed hybrid cloud can service secure, mission-critical processes, such as receiving customer payments (a private cloud service), and secondary processes, such as employee payroll processing (a public cloud service).

3.3.3 Community cloud

A community cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (such as mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them. A community cloud can exist on or off the premises.

3.4 Cloud storage use cases

Cloud storage is a cloud computing service designed to store data. It can be public, private, or hybrid. Cloud storage can be considered a type of *SaaS* because the hardware and platform underlying the service are usually not visible to the user. Some examples of public cloud storage services include Box and Dropbox. Private cloud storage services can be as simple as a NAS share, or a more purpose built infrastructure including SAN and a user-friendly front end.

Some organizations choose to outsource the data and entire workload to a cloud, as typical to *IaaS* solutions. Though simple to implement, such a solution can incur high costs, based on resource allocation and is most vulnerable to data leakage. In addition, this strategy might not even be optimal from the performance perspective.

An alternate strategy can be to replicate data on both the private and public sides, and to split the workload between the two sides. Simple queries can be computed on the private side and the complex ones can be performed over the public infrastructure.

That strategy uses local resources, and thereby reduces the cost of the required cloud services. However, the resource allocation cost and the amount of sensitive data that is exposed to the public cloud is maximum in this case. Another possibility might be to replicate only some part of the data to the public side. This scenario enables the distribution of the computation while limiting the disclosure risks and resource allocation costs to the desired thresholds.

Today, companies are realizing that all these models of computing are becoming the foundation of the next generation platform. Business disruption is at the heart of the requirement for the hybrid cloud.

Therefore, the hybrid cloud for storage is becoming the architectural framework that allows companies to use whichever deployment model best serves their business needs. Flexibility in the hybrid computing model comes from the capability to change deployment models whenever the business needs to change.

Hybrid storage cloud offers a path to IT optimization by implementing common key practices such as virtualization, standardization, and automation (as illustrated in Figure 3-2).

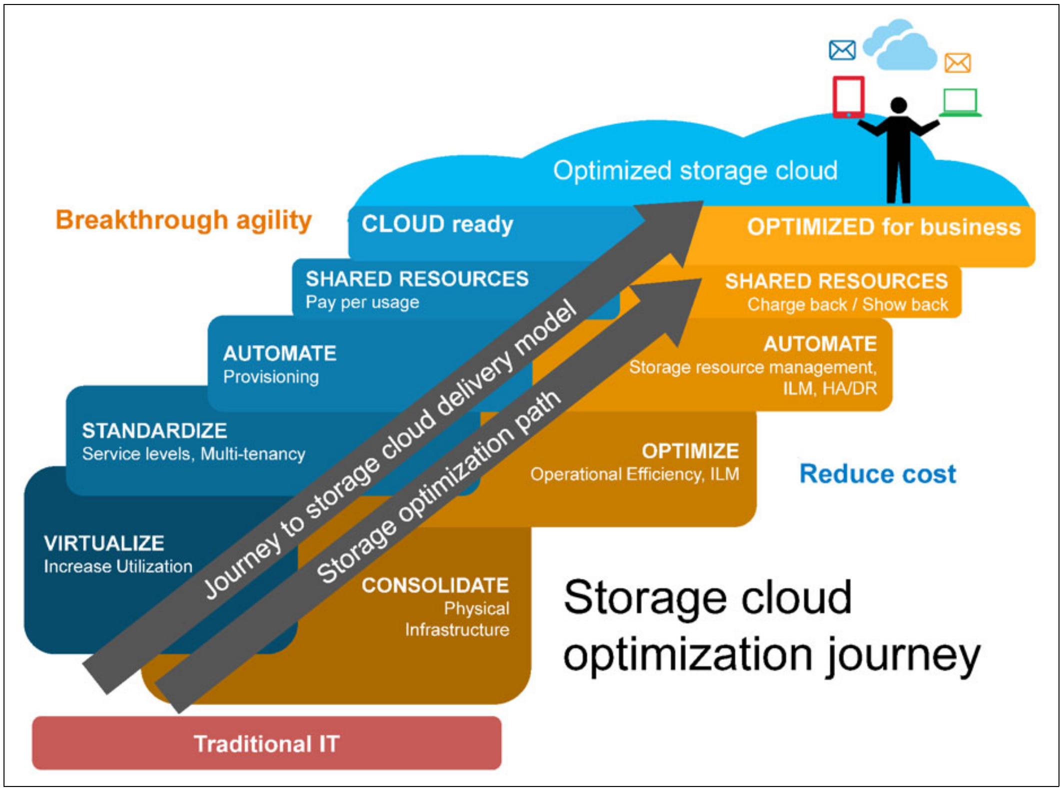


Figure 3-2 Cloud journey from traditional IT to storage cloud

An optimized storage infrastructure aligns IT resources to business requirements through managed service levels. These service levels are usually defined in a service catalog, which is supported within a storage cloud implementation.

The journey to storage cloud starts at different places for different organizations. Hybrid cloud can be an effective path for transitioning from traditional IT infrastructure to a cloud-based storage infrastructure.

The following use cases describe hybrid cloud in that manner:

- External storage virtualization
- Boosting the backup performance
- Backup repository

3.4.1 External storage virtualization

This use case focuses on the need to engage external storage and continue using the existing storage.

Many organizations are searching affordable and efficient ways to store, use, protect, and manage the data. An emphasis is put on the IBM Cognitive era of businesses of clients and their dynamic infrastructure. Therefore, a storage environment requires an easy-to-manage interface and flexibility to support many applications, servers, and mobility requirements, such as these examples:

- How to support growth of existing business critical applications
- How to stay current with vital new technologies such as encryption
- How to implement hybrid cloud strategies
- How to stay within limited budgets

Replacing the existing storage is not an option because of the following circumstances:

- High costs of new storage system
- Migration complexity from a legacy infrastructure to a new one
- Increased management effort
- Limited options because of a mix of various capabilities on heterogeneous storage systems

According to Storage Network Industry Association (SNIA), the best way to address these intrinsic complexities of managing storage is to create an abstraction layer between the host and the storage. This can be done by decoupling the storage from the storage functions that are required in the storage area network (SAN) environment. IBM implements this using IBM Spectrum Virtualize.

IBM Spectrum Virtualize decouples by abstracting the physical location of data from the logical representation of the data. The virtualization engine presents logical entities to the user and internally manages the process of mapping these entities to the actual location of the physical storage. This process is described in more detail in 3.6, “Software-defined storage (SDS)” on page 34.

By using IBM storage virtualization solutions built with IBM Spectrum Virtualize and SAN technologies, businesses have a viable solution. They can enhance the storage they already have in order to meet new demands while keeping new investment under control.

3.4.2 Boosting the backup performance

In this use case, a backup job is failing to be resolved within the backup window. This use case shows how SDS technology combined with Fibre Channel connectivity addresses the problem for this environment.

IBM Spectrum Protect for Storage Area Networks (SAN) can improve performance by routing backup data over the SAN instead of the local area network (LAN). LAN-free data movement can make LAN bandwidth available for other uses and decrease the load on the IBM Spectrum Protect server, allowing it to support a greater number of concurrent client connections.

Backing up data to tape or disk over the SAN, or restoring data from tape or disk over the SAN, has advantages over using the LAN. The following list notes those advantages over equivalent operations that are performed only over the LAN:

- Metadata is sent to the server over the LAN; sending metadata over the LAN has negligible impact on LAN performance. Client data bypasses the potentially busy and slower LAN and is sent over the faster SAN. Backing up or restoring data over a SAN is generally faster than the same operation over a LAN.
- Sending client data over the SAN frees the IBM Spectrum Protect server from the task of handling data, which leads to more efficient use of server resources because the data goes directly to storage.

This process is efficient when protecting large databases, because LAN-free data transfer allows a server to perform a backup over a Fibre Channel interface. The data moves from the SAN attached disk, through the application server, and directly to a SAN attached storage device. This methodology takes full advantage of the Fibre Channel conduit and provides high throughput, up to 32 Gbps.

3.4.3 Backup repository

A backup repository is a storage system configured to store backup data. Traditionally disk, file, and tape systems are used for this purpose. Managing these involves capacity management, and managing space reclamation and data replication.

Significant SAN/IP bandwidth is typically required because of replication requirements inherent to most traditional storage systems. In some cases, manual vaulting using couriers and physical tape shipping is still used for offsite data storage.

With the explosion of unstructured data, companies are facing the conundrum. Do they choose operational expenditure (OPEX) or capital expenditure (CAPEX) to determine what infrastructure requirements are necessary to successfully operate the current and future backup environments (especially for the less accessed data).

To build a financial balance (between frequently accessed data and archived data), and to be compliant with business requirements (such as encryption and retention) organizations are making a change. They are moving from an on-premises backup infrastructure to a cloud backup infrastructure.

IBM Cloud Object Storage is a new approach to address these concerns. It combines the best aspects of Cloud, on-premises, and Hybrid deployment models through a flexible, scalable and simple storage interface.

In this use case, illustrated in Figure 3-3 on page 31, a backup server ingests data and then processes it for movement to IBM Cloud Object Storage (COS).

This process eliminates the necessity of having a secondary site containing a full copy of the media assets. The site is not needed because COS uses a dispersed storage mechanism that uses a cluster of storages nodes to store pieces of the data across the available nodes. IBM Cloud Object Storage uses an Information Dispersal Algorithm (IDA) to break files into unrecognizable slices that are then distributed to the storage nodes.

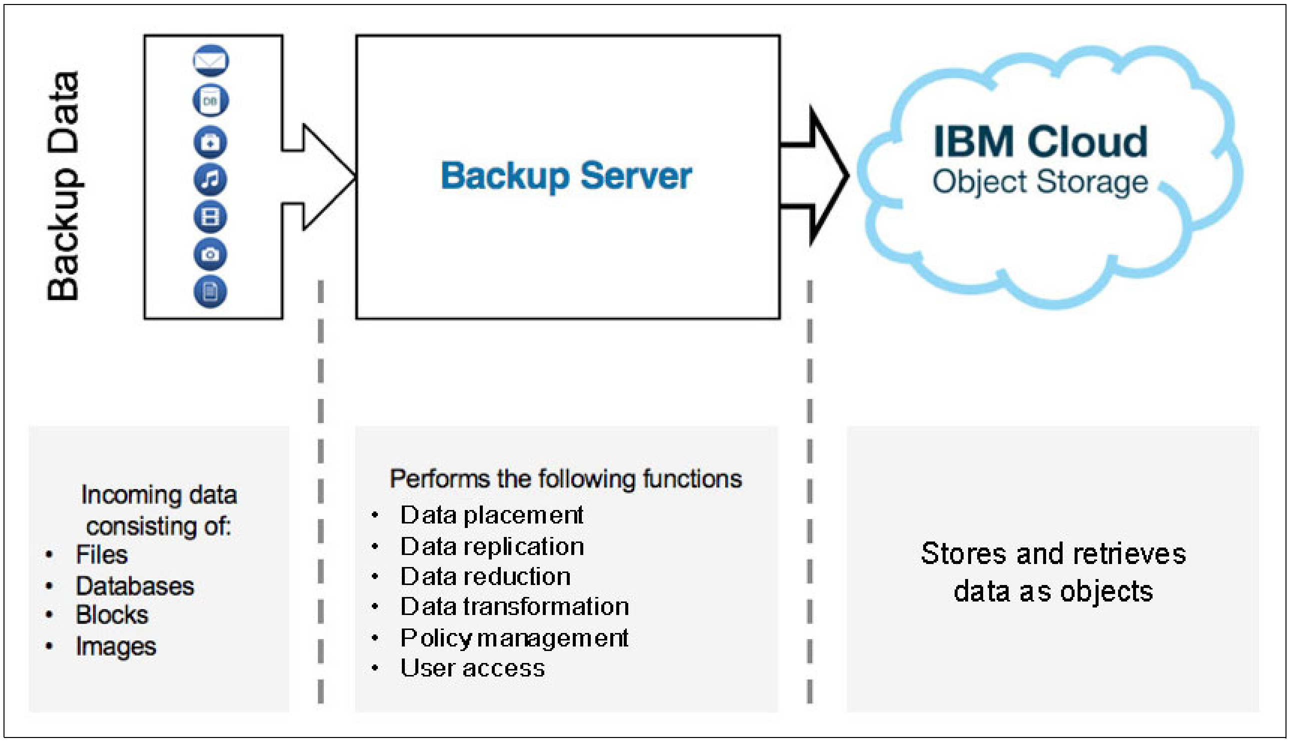


Figure 3-3 Backup server using IBM COS as a storage repository

No single node has all the data, which makes it safe and less susceptible to data breaches while needing only a subset of the storage nodes to be available to fully retrieve the stored data. This ability to reassemble all the data from a subset of the chunks dramatically increases the tolerance to node and disk failures.

3.5 Connectivity impact

In almost every IT infrastructure, speed of access to data is one of the most important factors. The development of the solid-state drive (SSD) helped with the insatiable desire for data access speed. Even so, that was not enough for the most demanding uses such as financial transaction applications or big data analytics.

To address these concerns, IBM uses Non-Volatile Memory Express (NVMe) devices as a base for IBM Flash Systems.

NVMe started as one of a handful of protocols to enable a faster connection of a flash drive to a PC using the Peripheral Component Interconnect Express (PCIe) bus. NVMe is becoming the standard in modern data centers because these devices are becoming denser and faster. NVMe delivers lower latency by using triple level cell (3D TLC) flash cards, combined with faster internal interconnects and new communications protocols.

The NVMe protocol evolved rapidly to address high latency problems. Direct memory access over the PCIe bus addressed interaction issues and is a well-tested method for moving data. NVMe changed the transfer mechanism from a *push* system that required transfer requests and acknowledgments to a system that allowed the receiving node to *pull* data when ready. Experience shows that this approach reduces CPU overhead to only a few percent.

The desire for speed and the various ways technology has risen to the challenge impacts the whole infrastructure balance. With capacity and performance increasing at such speed, the transaction bottlenecks have had a cascade effect. The bottlenecks have shifted to other areas such as switches, links, and servers.

Even before the latest technology evolutions, existing improvements were already causing bottlenecks in the present day. For example, predominantly installed 10-Gigabit Ethernet (10GbE) and 8 Gb per second Fibre Channel (8 Gbps FC) networking protocols. In this situation, lower-latency storage creates a demand for higher-throughput protocols and a redesign of storage networks. Storage performance requires deterministic, nonblocking performance, and extremely low-latency performance characteristics. Therefore, even 40GbE struggles to provide sufficient low latency for the fastest solid-state arrays that use NVMe and the fastest SSDs.

Fibre Channel proved to be the most efficient way of carrying the data, taking advantage of the speed of NVMe without creating a bottleneck in the storage network.

All SANs must perform the same functions such as these:

- Host-to-storage mapping
- Zoning, virtual LANs (VLANs) or any other type of virtualized network device segregation
- Path failover
- Congestion management
- Device registration/deregistration
- Buffer management
- Routing

Note: These functions are true whether the SAN is Ethernet-based, PCIe, Serial Attached SCSI (SAS), NVMe over fabric, FC, Fibre Channel over Ethernet (FCoE), or InfiniBand.

Taking this into consideration, there is no intrinsic administrative complexity advantage or disadvantage that one storage protocol has over another. Familiarity, availability of skills, and a mature ecosystem of support and monitoring products make some protocols simpler to implement, administer, and manage than others.

However, protocols, such as Fibre Channel, and new designs, such as Internet Wide Area RDMA Protocol (iWARP), have offloaded storage protocol processing. They have moved this role from the server to the host bus adaptor (HBA) card or network interface card (NIC). This storage offload reduces server operating system or hypervisor CPU and memory requirements. Therefore, this process requires fewer CPU resources and enables faster data transfers. Nevertheless, there are added protocol complications when multiple network layers need to be managed in a single network.

This section describes the most used protocols:

- Fibre Channel (FC)
- Internet Small Computer Systems Interface (iSCSI)
- Fibre Channel over Ethernet (FCoE)
- Determine SAN requirements

3.5.1 Fibre Channel (FC)

FC is a high-speed network technology that deliver low-latency, high-bandwidth, and high-throughput. These advantages are due to its deterministic nonblocking design. This high-link efficiency makes FC well-suited for storage traffic.

FC is transparent to the applications, as the interface is provided by the server operating system or hypervisor. FC has a low footprint and low server requirements because its processing is offloaded from the server to the HBA, leaving more CPU cycles and memory for the application services. Upgrade paths for FC customers are well designed, because each new version has backward support to at least the previous version, if not two prior versions.

For example, if a business implements a 32 Gbps FC network, they can still use the existing 8 Gbps and 16 Gbps with 32 Gbps equipment (such as SAN switches, HBAs and storage array port connections). Both 64 Gbps and 128 Gbps products are on the FC roadmap.

3.5.2 Internet Small Computer Systems Interface (iSCSI)

iSCSI uses Transmission Control Protocol/Internet Protocol (TCP/IP). TCP/IP is an industry standard providing a large number of vendors, products, and components from which to choose.Therefore, iSCSI is a relatively low-cost solution to purchase. The management is also simple.

To achieve SANs with no single point of failure, and to provide resiliency and serviceability, customers implement physically separate Ethernet networks from their existing Ethernet networks. These are similar to FC network designs, because they use dual iSCSI SANs. The iSCSI SANs use dual networks to enable dynamic upgrades and to avoid catastrophic outages due to operator error. Although a lot of standardization exists, a necessary task is to check and adhere to vendor hardware compatibility lists (HCLs) between products from iSCSI vendors. The performance of iSCSI is determined by the underlying Ethernet infrastructure.

As a matter of comparison, 10GbE roughly provides IOPS and throughput comparable to 8 Gbps FC. Although 10GbE provides more bandwidth, it is less efficient as a storage protocol than FC because of its architecture.

3.5.3 Fibre Channel over Ethernet (FCoE)

FCoE is a computer network technology that encapsulates Fibre Channel over Ethernet networks.

Although FCoE can potentially improve economics, it does not improve availability or performance relative to FC. It does add management complexity. By retaining the native Fibre Channel constructs, FCoE is meant to integrate with existing Fibre Channel networks and management software.

With FCoE, Fibre Channel becomes another network protocol running on Ethernet, alongside traditional Internet Protocol (IP) traffic. FCoE operates directly above Ethernet in the network protocol stack, in contrast to iSCSI, which runs on top of TCP and IP. As a consequence, FCoE is not routable at the IP layer, and will not work across routed IP networks.

3.5.4 Determine SAN requirements

Because one size does not fit all, determining exactly which SAN type will meet the requirements of each customer is difficult—needs will differ subtly from customer to customer. So, the choice of which block network protocol that will be in use is an application-oriented decision. Before designing any solution, consider several questions to help understand requirements:

- Availability: What level of availability is required for the SAN? Must the SAN be available at all times, or is downtime for maintenance or equipment failure tolerable?
- Performance: What level of performance is required from the SAN? Is data throughput by the host an important factor? Is total throughput of the SAN an important factor?
- Applicability: What type of data is the SAN being used for? Is it mission-critical production data or a development environment?

One of the biggest challenges for organizations is to ensure that creating predictability when managing different workloads across a huge variety of environments is possible.

In a hybrid cloud architecture, the idea is to provide the capability to dynamically change as deployment models change and adjust the environment to new workloads and new services. The clear benefit here is the flexibility gained to select the correct environment for the workload.

That is possible because a hybrid cloud model is built in open standards such as Linux, OpenStack, and Hadoop. This allows systems to talk to each other even if they are from different vendors or if they are in different places, thus, providing interoperability between on-premises systems with systems in cloud.

Different workloads have different characteristics, and the platform for them to run in an optimized manner will also differ.

Matching workloads with the right environment allows organizations to optimize efficiency while also budgeting only for the resources that are required.

Although the communication between different clouds are based on Ethernet protocols, a strong suggestion is that the connections within a cloud use Fibre Channel. Using Fibre Channel guarantees the maximum performance and reliability.

3.6 Software-defined storage (SDS)

This section provides an overview of how IT storage infrastructures can be designed with SDS technologies. This design enables dynamic assignment of IT resources. Assignment can be based on application workload requirements with best-available resources aligned to business-requirements-based service level policies.

Additionally, this section covers the following areas:

- How the IBM Spectrum Storage family is organized and their functions
- How IBM delivers a real end-to-end SDS Architecture through either of these planes:
  - IBM SDS control plane
  - SDS data plane

3.6.1 Defining SDS

Software-defined storage, as the name implies, can easily be misinterpreted as storage that is defined in software. Although it is, this characterization lacks several essential aspects of the currently intended meaning. Most IT storage systems today are already based in either software, microcode, or a mix of both options.

SDS, in today's business context, refers to IT storage that goes beyond typical array interfaces (command line interface (CLI) and graphical user interface (GUI)) to operate within a higher architectural construct. SDS supports overall IT architectural definition, configuration, and operations. These aspects are often referred to as a software-defined infrastructure (SDI) within a software-defined environment (SDE), or the more restrictive software-defined data center (SDDC).

This approach is possible (and probably easier to implement) in a homogeneous, single vendor implementation. However, its greatest value and versatility is as standardized programming interfaces applied across a heterogeneous multivendor IT infrastructure.

IBM created a smarter interface between the storage layer and the compute layer called *software-defined storage (SDS)*. Software-defined storage is divided into two main layers:

- SDS control plane

  This software layer manages the virtualized storage resources. It provides all of the high-level functions that are needed by the customer to run the business workload and enable optimized, flexible, scalable, and rapid provisioning storage infrastructure capacity. These capabilities span functions like storage virtualization, policies automation, analytics and optimization, backup and copy management, security, and integration with the API services, including other cloud provider services.

- SDS data plane

  This layer encompasses the infrastructure where data is processed. It consists of all basic storage management functions such as virtualization, RAID protection, tiering, copy services (remote, local, synchronous, asynchronous, and point-in-time), encryption, and data deduplication that can be started and managed by the control plane. The data plane is the interface to the hardware infrastructure where the data is stored. It provides a complete range of data access possibilities, spanning traditional access methods like block I/O (for example, iSCSI) or file I/O (POSIX compliant), to object-storage or Hadoop Distributed File System (HDFS).

Figure 3-4 on page 36 shows the IBM SDS architecture with a mapping of the IBM Spectrum Storage family of products across the SDS control plane and data plane.

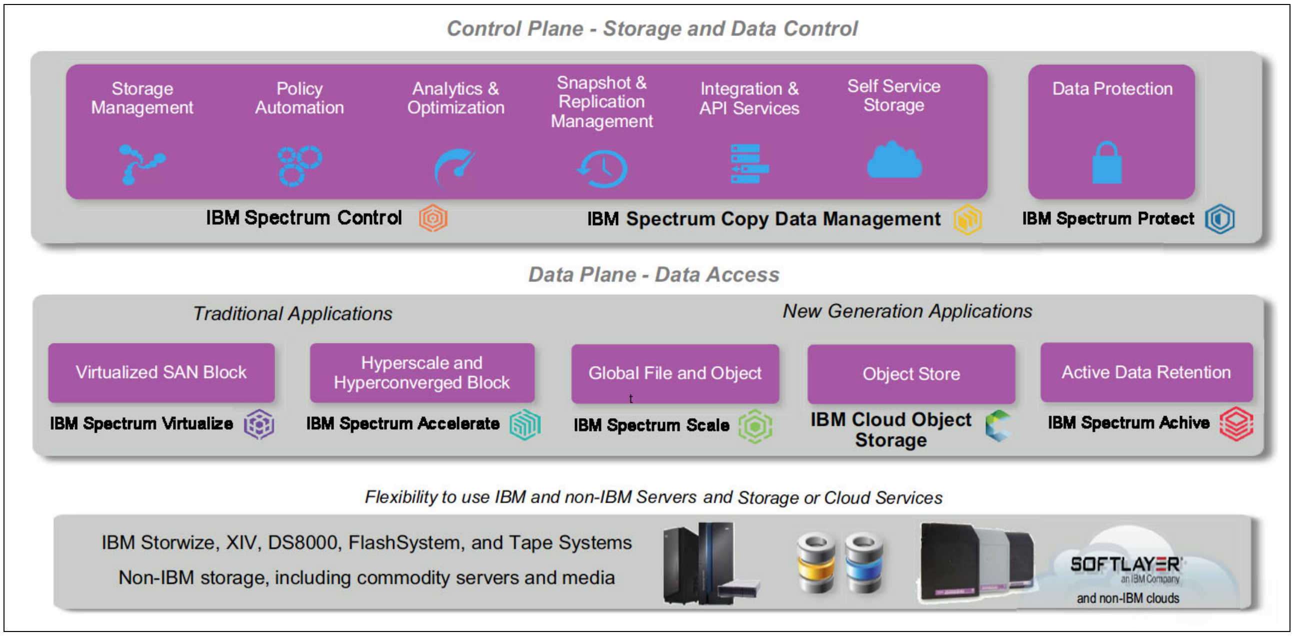


Figure 3-4 IBM Spectrum Storage family mapped to SDS control plane and data plane

3.6.2 Components in the control plane and data plane

This section describes the following components:

- Control plane components:
  - IBM Spectrum Control

    Automates control and optimization of storage and data infrastructure.
  - IBM Spectrum Copy Data Management

    Automates creation and use of copy data snapshots, vaults, clones, and replicas on existing storage infrastructure.
  - IBM Spectrum Protect

    Optimizes data protection for client data through backup and restore capabilities.
- Data plane components:
  - IBM Spectrum Virtualize

    A core SAN Volume Controller function is virtualization that frees client data from IT boundaries.
  - IBM Spectrum Accelerate

    Enterprise storage for cloud that is deployed in minutes instead of months (Storage as a Service).
  - IBM Spectrum Scale

    Storage scalability to yottabytes and across geographical boundaries.
  - IBM Cloud Object Storage

    Erasure-coding based object cloud storage.
  - IBM Spectrum Archive

    Enables long-term storage of low activity data.

IBM Spectrum Control

IBM Spectrum Control provides efficient infrastructure management for virtualized, cloud, and software-defined storage. This is done by reducing the complexity associated with managing multi-vendor infrastructures. IBM Spectrum Control helps businesses optimize provisioning, capacity, availability, protection, reporting, and management for today's business applications without having to replace existing storage infrastructure. With support for block, file, and object workloads, IBM Spectrum Control enables administrators to provide efficient management for heterogeneous storage environments.

IBM Spectrum Control helps organizations transition to new workloads and updated storage infrastructures by providing these advantages to significantly reduce total cost of ownership:

- A single management console that supports IBM Spectrum Virtualize, IBM Spectrum Accelerate, IBM Cloud Object Storage, and IBM Spectrum Scale environments. Together these systems enable holistic management of physical and virtual block, file, and object systems storage environments.
- Insights that offer advanced, detailed metrics for storage configurations, performance, and tiered capacity. This content is shared in an intuitive web-based user interface with customizable dashboards so that the most important information is always accessible.
- Performance monitoring views that enable quick and efficient troubleshooting during an issue with simple threshold configuration and fault alerting for high availability.

Figure 3-5 shows the IBM Spectrum Control dashboard where all the managed resources in your data server are presented in an aggregated view.

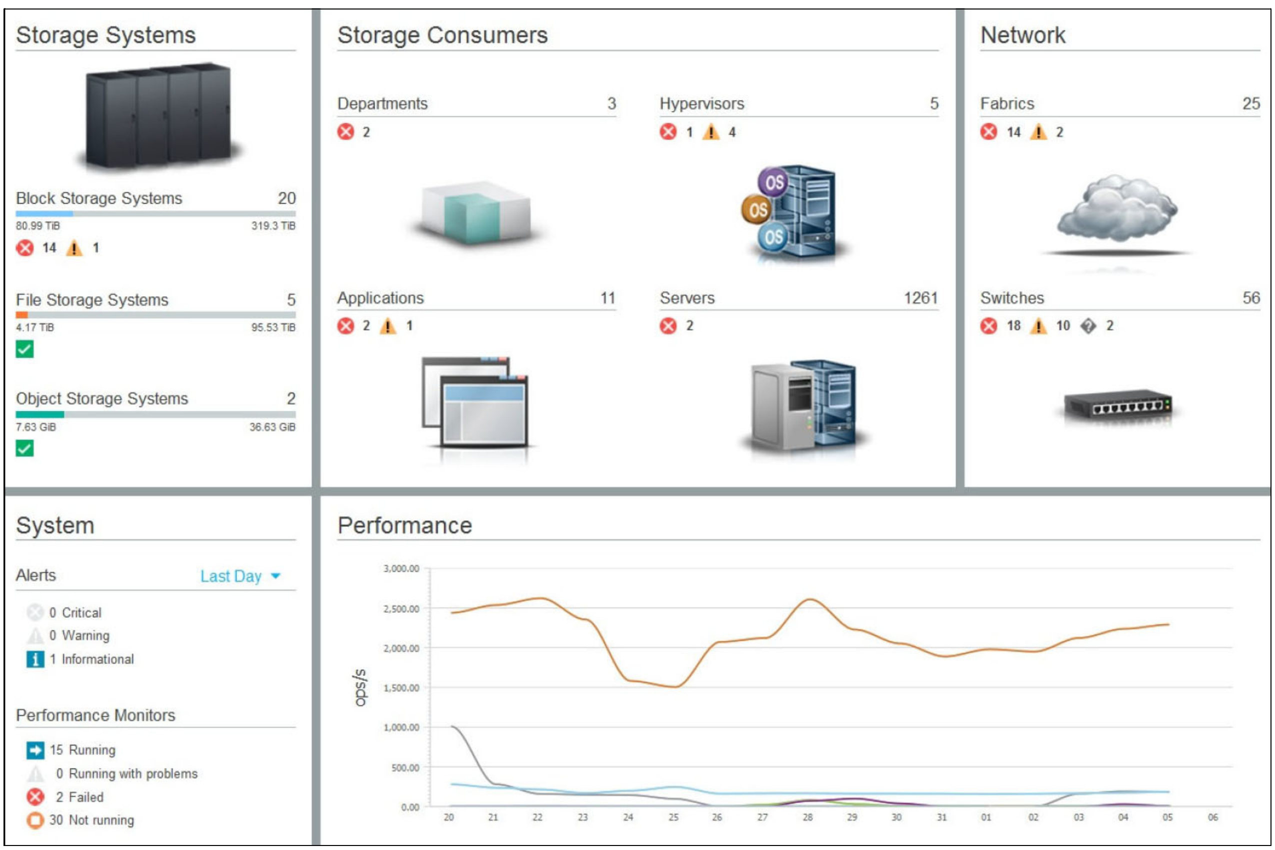


Figure 3-5 Single dashboard for monitoring all storage components

IBM Spectrum Copy Data Management

IBM Spectrum Copy Data Management makes copies available to data consumers when and where they need them. This function is done without creating unnecessary copies or leaving unused copies on valuable storage. It catalogs copy data from across local, hybrid cloud, and off-site cloud infrastructure, identifies duplicates, and compares copy requests to existing copies. This process ensures that the minimum number of copies are created to service business requirements.

Data consumers can use the self-service portal to create the copies they need when they need them, enabling business agility. Copy processes and work flows are automated to ensure consistency and reduce complexity. IBM Spectrum Copy Data Management rapidly deploys as an agentless VM. This product helps manage snapshot and IBM FlashCopy® images made to support DevOps, data protection, disaster recovery, and hybrid cloud computing environments.

This member of the IBM Spectrum Storage family automates the creation and catalogs the copy data on existing storage infrastructure (such as snapshots, vaults, clones, and replicas). One of the key use cases for this IBM product centers around its use with Oracle, Microsoft SQL server, and other databases. These servers and databases are often copied to support application development, testing, and data protection.

The IBM Spectrum Copy Data Management software is an IT modernization technology that focuses on using existing data in a manner that is efficient, automated, scalable, and easy to use. IBM Spectrum Copy Data Management, with IBM storage arrays, delivers in-place copy data management that modernizes IT processes and enables key use cases with existing infrastructure. See Figure 3-6.

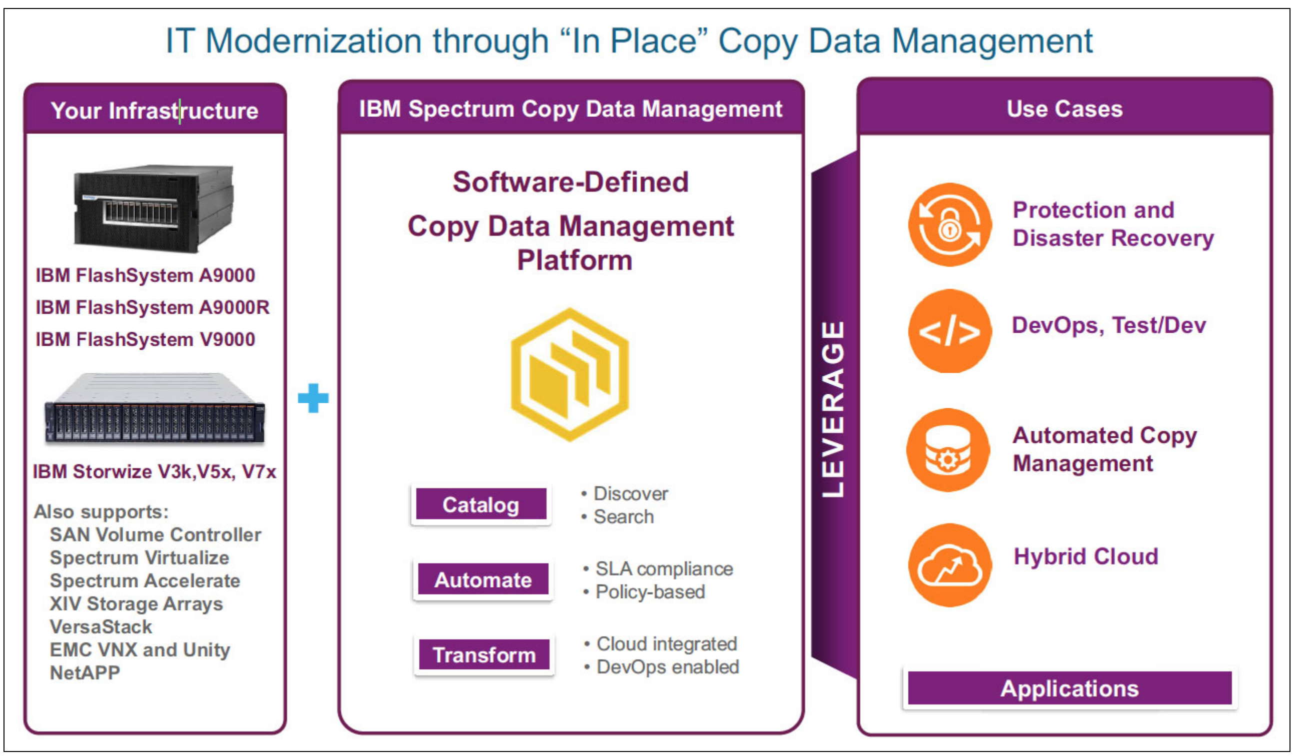


Figure 3-6 Software-defined IBM Spectrum Copy Data Management Platform

IBM Spectrum Protect

IBM Spectrum Protect is an intuitive, intelligent, and transparent software that provides a set of product features that allow you to design adaptive and comprehensive data protection solutions. It is a comprehensive data protection and recovery solution for virtual, physical, and cloud data. IBM Spectrum Protect provides backup, snapshot, archive, recovery, space management, bare machine recovery, and disaster recovery capabilities.

IBM Spectrum Protect offers these capability highlights:

- Protects virtual, physical, and cloud data with one solution
- Reduces backup and recovery infrastructure costs
- Delivers greater visualization and administrator productivity
- Simplifies backups by consolidating administration tasks
- Space management moves less active data to less expensive storage, such as tape or cloud
- Provides long-term data archive for data retention, such as for compliance with government regulations

IBM Spectrum Protect Operations Center is a graphical user interface (GUI), with new features. See Figure 3-7. It provides an advanced visualization dashboard, built-in analytics, and integrated workflow automation features that dramatically simplify backup administration.

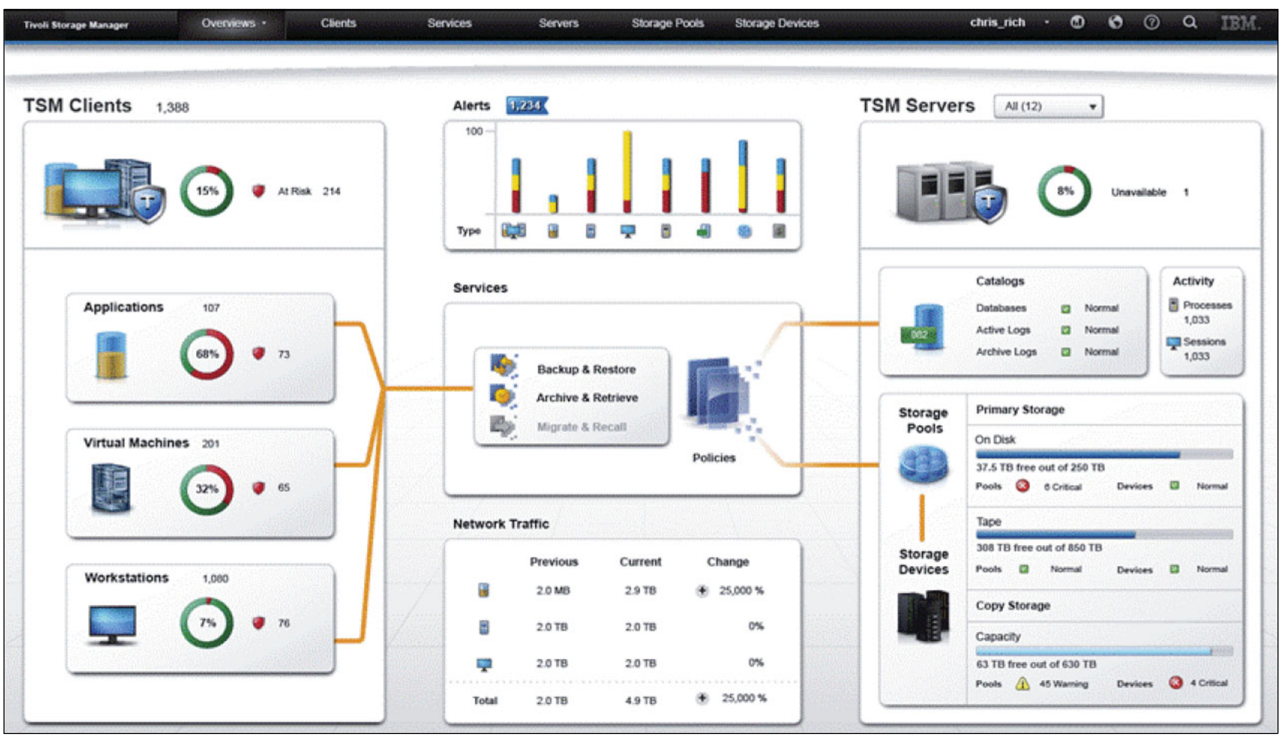


Figure 3-7 IBM Spectrum Protect graphical user interface

IBM Spectrum Virtualize

IBM Spectrum Virtualize software is at the heart of IBM SAN Volume Controller, IBM Storwize family, and IBM FlashSystem V9000. It enables these systems to deliver better data value, security, and simplicity through industry-leading virtualization. This virtualization transforms existing and new storage and streamlines deployment for a simpler, more responsive, scalable, and cost efficient IT infrastructure.

IBM Spectrum Virtualize systems provide management of storage from entry and midrange up to enterprise disk systems, and enable hosts to attach through SAN, FCoE, or iSCSI to existing Ethernet networks. IBM Spectrum Virtualize is easy to use, enabling existing staff to start working with it rapidly. IBM Spectrum Virtualize uses virtualization, thin provisioning, and compression technologies to improve storage utilization and meet changing needs quickly and easily. In this way, IBM Spectrum Virtualize products are the ideal complement to server virtualization strategies.

IBM Spectrum Virtualize software capabilities are offered across various platforms, including SAN Volume Controller, Storwize V7000 (Unified), Storwize V5000, and FlashSystem V9000. IBM Spectrum Virtualize products are designed to deliver the benefits of storage virtualization and advanced storage capabilities in environments from large enterprises to small businesses and midmarket companies. Those products are noted in the following list:

- IBM Real-time Compression for inline, real-time compression
- Stretched Cluster and IBM HyperSwap® for a high-availability solution
- IBM Easy Tier for automatic and dynamic data tiering
- Distributed RAID for better availability and faster rebuild times
- Encryption for internal and external virtualized capacities
- FlashCopy snapshots
- Remote data replication

The SAN Volume Controller combines software and hardware into a comprehensive, modular appliance that uses symmetric virtualization.

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

Figure 3-8 shows the IBM Spectrum Virtualize functions.

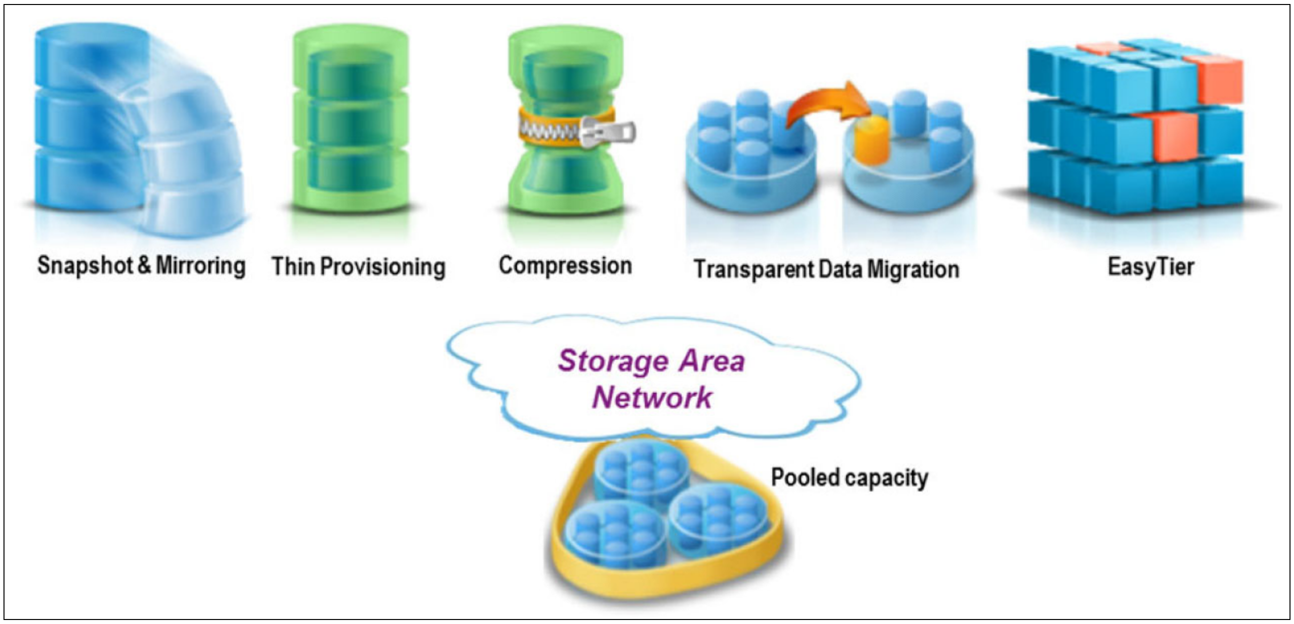


Figure 3-8 IBM Spectrum Virtualize functions

IBM Spectrum Accelerate

IBM Spectrum Accelerate is a highly flexible, software-defined storage solution that enables rapid deployment of block data storage services for new and traditional workloads (both on and off premises). It is a key member of the IBM Spectrum Storage portfolio. IBM Spectrum Accelerate allows you to run hotspot-free, grid-scale software that runs on the XIV Storage System Gen3 enterprise storage platform (within an existing data center infrastructure or in a cloud provider such as IBM SoftLayer). It offers proven grid- scale technology, mature features, and ease of use, and is already deployed on over 100,000 servers worldwide.

IBM Spectrum Accelerate delivers predictable and consistent storage performance for management scaling to more than 68 petabytes usable. It has a rich feature set that includes remote mirroring and granular multi-tenancy. IBM Spectrum Accelerate deploys on-premises on x86 commodity servers and on the optimized XIV Storage System, and off-premises as a public cloud service on SoftLayer. You can manage all your IBM Spectrum Accelerate instances, wherever they are deployed, in a single, intuitive interface. Hardware-independent, transferable licensing offers superb operational flexibility and cost benefits.

With IBM Spectrum Accelerate, customers can deploy a hyper-converged solution across their on-premises and off-premises deployments to help meet the unpredictability of today's cloud world. It runs as a virtual machine on the VMware vSphere ESXi hypervisor. It converges compute and storage, enabling customer-built, hyper-converged solutions based on proven XIV technology. The ability to run application workload VMs on the same servers as the storage, enables customers to rapidly provision and decommission workloads in a dynamic fashion.

IBM Spectrum Accelerate delivers a single management experience across software-defined storage infrastructure using IBM HyperScale Manager. The manager handles IBM Spectrum Accelerate instances, IBM XIV, and the IBM A9000 all-flash solution. This combination helps cut costs in the following ways:

- Reducing administration effort and training
- Reducing procurement costs
- Standardizing data center storage hardware operations and services
- Providing licensing flexibility that enables cost-efficient cloud building

Figure 3-9 shows how straightforward scaling is, by building a storage grid with IBM Spectrum Accelerate.

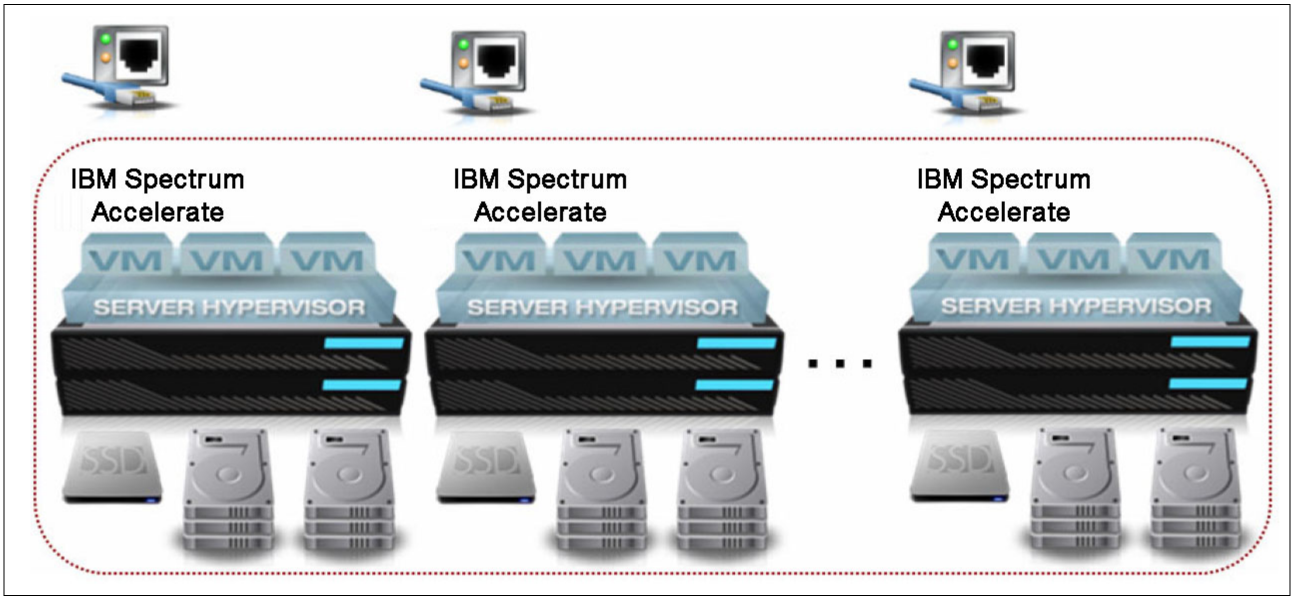


Figure 3-9 IBM Spectrum Accelerate iSCSI storage grid

IBM Spectrum Accelerate gives organizations these capabilities:

- Enterprise cloud storage in minutes, using commodity hardware
- Hotspot-free performance and QoS without any manual or background tuning needed
- Advanced remote replication, role-based security, and multi-tenancy
- Deploy on-premises or on the cloud (also as a service on SoftLayer)
- Hyper-scale management of dozens of petabytes
- Best in class VMware and OpenStack integration
- Run IBM Spectrum Accelerate and other application virtual machines on the same server

IBM Spectrum Scale

IBM Spectrum Scale is a proven, scalable, high-performance file management solution that is based on IBM General Parallel File System (GPFS). IBM Spectrum Scale provides world-class storage management with extreme scalability, flash accelerated performance, and automatic policy-based storage tiering from flash to disk, then to tape. IBM Spectrum Scale reduces storage costs up to 90% while improving security and management efficiency in cloud, big data, and analytics environments.

First introduced in 1998, this mature technology enables a maximum volume size of 8 YB, a maximum file size of 8 EB, and up to 18.4 quintillion (two to the 64th power) files per file system. IBM Spectrum Scale provides simplified data management and integrated information lifecycle tools such as software-defined storage for cloud, big data, and analytics.

It introduces enhanced security, flash accelerated performance, and improved usability. It also provides capacity quotas, access control lists (ACLs), and a powerful snapshot function.

IBM Spectrum Scale adds elasticity with the following capabilities:

- Global namespace with high-performance access scales from departmental to global
- Automated tiering, data lifecycle management from flash (6x acceleration) to tape (10x savings)
- Enterprise ready with data security (encryption), availability, reliability, and large scale
- POSIX compliant
- Integrated with OpenStack components and Hadoop

Figure 3-10 shows an example of the IBM Spectrum Scale architecture.

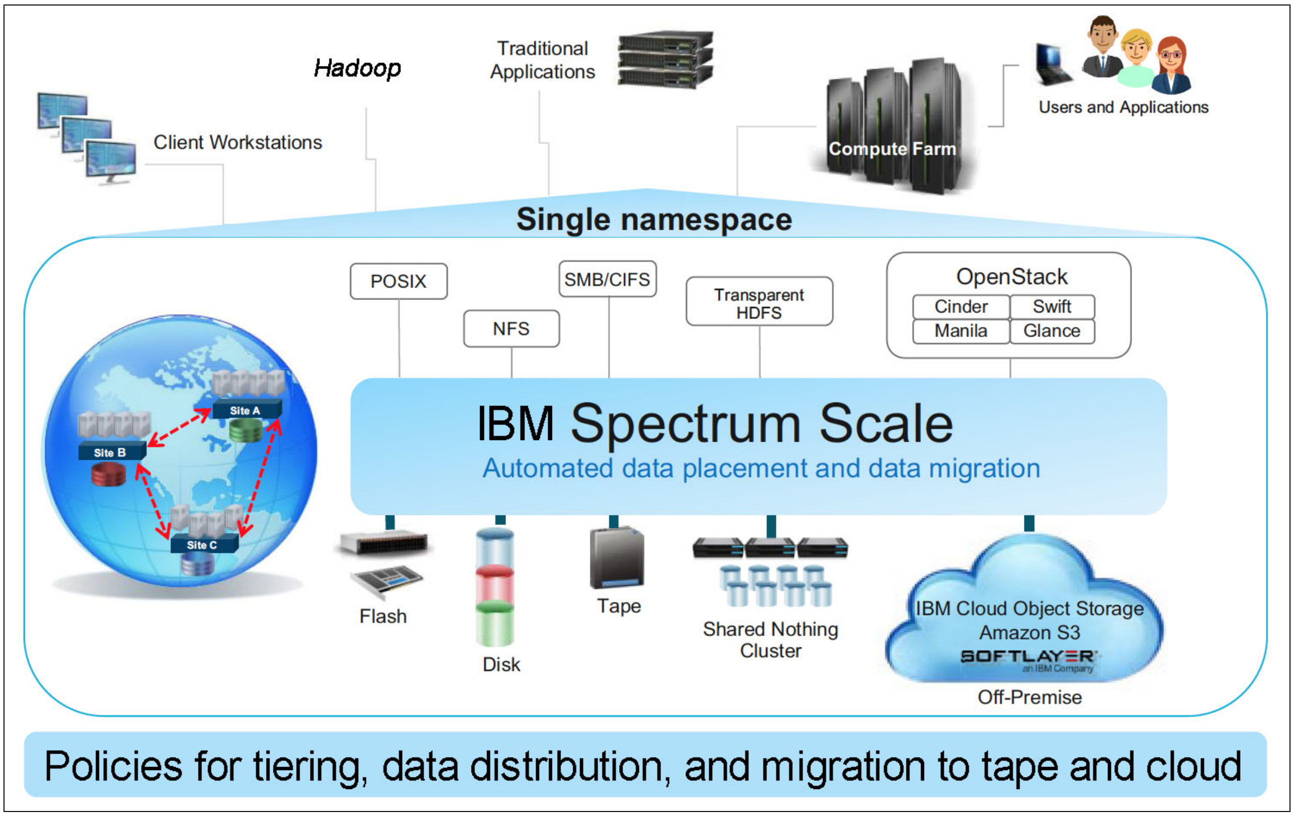


Figure 3-10 IBM Spectrum Scale architecture

IBM Cloud Object Storage

The IBM Cloud Object Storage (COS) system is a breakthrough cloud platform that helps solve petabyte and beyond storage challenges for companies worldwide. Clients across multiple industries use IBM Cloud Object Storage for large-scale content repository, backup, archive, collaboration, and software as a service (SaaS).

The Internet of Things (IoT) allows every aspect of life to be instrumented through millions of devices that create, collect, and send data every second. These trends are causing an unprecedented growth in the volume of data being generated. IT organizations are now tasked with finding ways to efficiently preserve, protect, analyze, and maximize the value of their unstructured data as it grows to petabytes and beyond. And object storage is designed to handle unstructured data at web-scale.

The IBM Cloud Object Storage portfolio gives clients strategic data flexibility, simplified management, and consistency with on-premises, cloud, and hybrid cloud deployment options. See Figure 3-11.

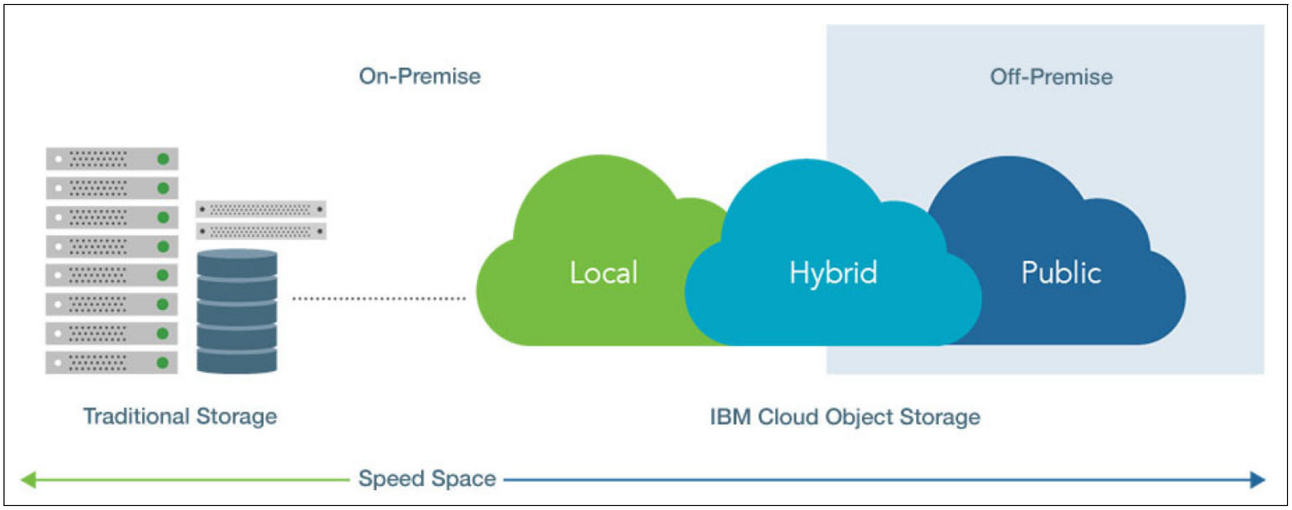


Figure 3-11 Cloud Object Storage offers flexibility for on-premises, cloud, and hybrid cloud deployment options

IBM Cloud Object Storage's Dispersed Storage® Network (dsNet) solutions enhance on-premises storage options for clients and service providers with low-cost, large-scale active archives and unstructured data content stores. The solutions complement the IBM software defined IBM Spectrum Storage portfolio for data protection and backup, tape archive, and a high-performance file and object solution where the focus is on response time.

IBM Cloud Object Storage can be deployed as an on-premises, public cloud, or hybrid solution, providing you unprecedented choice, control, and efficiency:

- On-premises solutions: Deploy IBM Cloud Object Storage on premises for optimal scalability, reliability, and security. The software runs on industry standard hardware for flexibility and simplified management.
- Cloud solutions: Easily deploy IBM Cloud Object Storage on the IBM SoftLayer public cloud.
- Hybrid solutions: For optimal flexibility, deploy IBM Cloud Object Storage as a hybrid solution to support multiple sites across your enterprise (on-premises and in the public cloud) for agility and efficiency.

IBM Spectrum Archive

IBM Spectrum Archive enables direct, intuitive, and graphical access to data stored in IBM tape drives and libraries. This is done by incorporating the Linear Tape File System (LTFS) format standard for reading, writing, and exchanging descriptive metadata on formatted tape cartridges. It is a member of the IBM Spectrum Storage family. IBM Spectrum Archive eliminates the need for extra tape management and software to access data.

IBM Spectrum Archive offers three software solutions for managing your digital files with the LTFS format:

- Single Drive Edition (SDE)
- Library Edition (LE)
- Enterprise Edition (EE)

With IBM Spectrum Archive Enterprise Edition and IBM Spectrum Scale, tape can now add savings as a low-cost storage tape tier. Being able to use a tier of tape for active but *cold* (infrequently used) data enables enterprises to look at new ways to cost optimize their unstructured data storage. They are able to match the value of the data, or the value of the copies of data to the most appropriate storage media.

In addition, the capability to store the data at the cost of tape storage has allowed customers to build their cloud environments to take advantage of this new cost structure. IBM Spectrum Archive provides enterprises with the ability to store cold data at costs that can be cheaper than some public cloud provider options. To understand the potential costs with large-scale cold data storage and retention, IBM created a Tape TCO Calculator.

Network attached unstructured data storage with native tape support using LTFS delivers the best mix of performance and lowest cost storage.

IBM Spectrum Archive options can support small, medium, and enterprise businesses with the following capabilities:

- Seamless virtualization of storage tiers
- Policy-based placement of data
- Single universal namespace for all file data
- Security and protection of assets
- Open, non-proprietary, cross platform interchange
- Integrated functionality with IBM Spectrum Scale

4

Storage network modernization

This chapter provides an overview of a storage network modernization process. This process is essential to meet the current issues and future flash based storage requirements. This process can be daunting but there are some practices and helpful product features that can make this process smoother. This chapter describes some of those SAN b-type design practices and Fabric Vision features utilization. Many Fabric Vision features are useful during all modernization phases and the last section of this chapter shows some useful implementations.

This following topics are covered in this chapter:

4.1 Modernizing SAN infrastructures is essential

Enterprises have started using low-latency, high-throughput flash arrays such as the IBM FlashSystem 900 for demanding and performance sensitive workloads. The IBM SAN B-type Gen 6 Fibre Channel switch with IO Insight and VM Insight enhancements is perfectly suited to these types of workloads. The sub-microsecond latency through the switch and the increased bandwidth offered by 32 Gbps throughput speed supports demanding workloads. A faster storage network also improves overall performance. Performance testing has shown that 16 Gbps and even 8 Gbps attached all-flash arrays can realize dramatic benefits by connecting to a Gen 6 SAN and host adapter (offering gains up to 2x over Gen 5 SANs).

Figure 4-1 shows the impact of SAN infrastructure modernization in relation to performance.

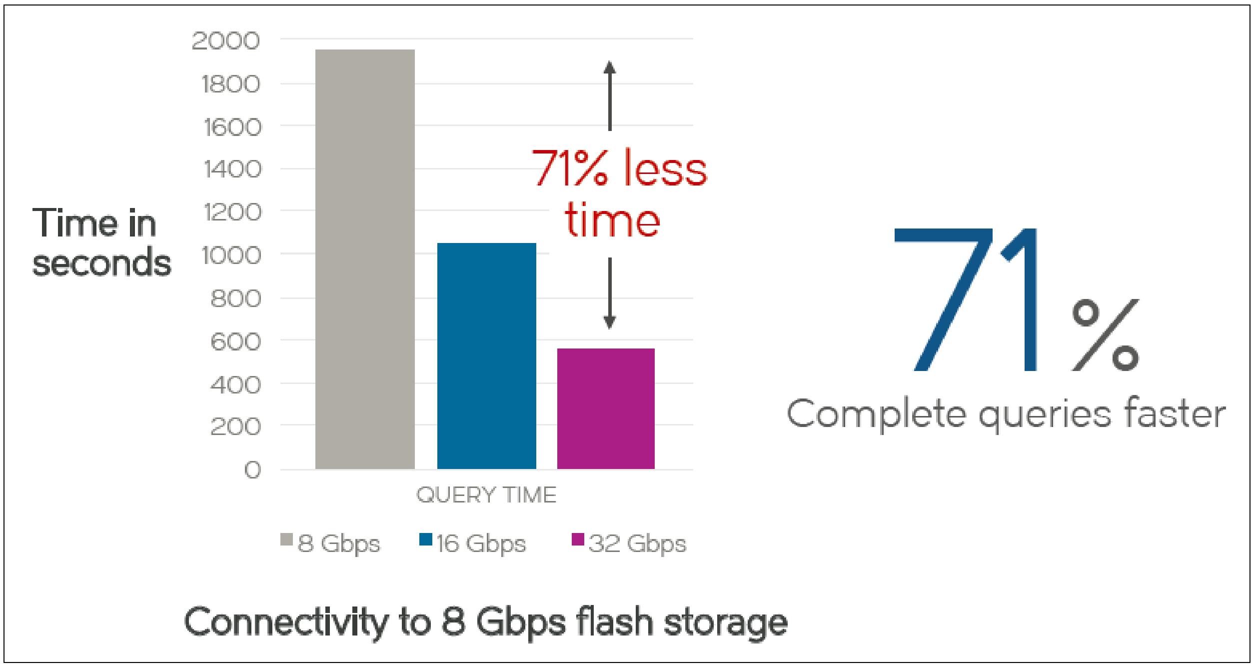


Figure 4-1 Impact of SAN infrastructure modernization in relation to performance

In addition to offering fast throughput speeds, Fibre Channel fabrics ensure a highly reliable network with deterministic data delivery. To ensure consistency of data transfer with predictable performance, forward error correction (FEC) is integral to Gen 6 and its standards of sustaining 32 Gbps throughput. FEC provides data transmission error control and can correct one burst, of up to 11-bit errors, in every 2112-bit transmission.

IT administrators understand that deploying an environment is not simply the deployment process and then all is complete. Administrators must be able to monitor commitments to a service level agreement (SLA) to show that they are meeting the customer's performance requirements. Fabric Vision technology is a suite of tools and capabilities designed to pre-validate, monitor, troubleshoot, optimize, and tune storage environments to prevent issues from occurring and to mitigate their impact when they do occur. Gen 6 Fibre Channel extends the Flow Vision capabilities with the introduction of the IO Insight and VM Insight.

4.1.1 Critical capabilities in a storage architecture

Usually a SAN is not renewed as often as the other IT infrastructure. After the SAN is in place, it is not easy to architect again. Because of that, sizing should not be based on actual workloads. A critical aspect is to consider business objectives and future requirements.

The process of a SAN design should account for the following parameters:

- Availability: The b-type Gen 6 Fibre Channel switch can perform buffer credit recovery, forward error correction (FEC), and ClearLink Diagnostic Port (D_Port).
- Scalability: The b-type Gen 6 Fibre Channel switch 128 Gbps UltraScale Inter-Chassis Link enable support for up to 12 chassis using a 4x8 core/edge design. This allows connectivity for up to 6,144 device ports. A full-mesh topology, using up to 9 directors allows connectivity for up to 4,608 device ports.
- Performance: The b-type Gen 6 Fibre Channel switch is running at speeds of 8 Gbps, 16 Gbps, and 32 Gbps. The next planned steps for availability are 64 Gbps and even 128 Gbps.
- Extensibility: The b-type Gen 6 Fibre Channel switch is taking advantage of newer and older technologies within the same environment. It has the ability to run multiple vendors within the same architecture to meet evolving storage requirements. Three generations of backward-compatibility, future-proof investments with a Gen 7-ready storage networking platform and NVMe over Fabrics as new protocol for solid-state storage.
- Manageability: The b-type Gen 6 Fibre Channel switch integrated to the IBM Network Advisor is a user-friendly interface and useful utility for management, troubleshooting, and tuning the SAN environment.
- NVMe Over Fabrics ready: The b-type Gen 6 Fibre Channel switch is NVMe ready. NVMe and SCSI protocols will coexist for years to come. This coexistence simplifies transitioning and avoids the planning, building, and maintenance of a separate infrastructure.

4.2 Basic preferred practices for SAN design and RAS

This chapter describes a high-level preferred practice for SAN design and a reliability, availability, and serviceability (RAS) guide based on IBM b-type Gen 6 products and features. This section focuses on planning, implementing, and monitoring of the SAN infrastructure during a modernization process and covers the following topics:

- Architecting a SAN
- Developing technical specifications
- Zoning
- Predeployment cabling and optics validation
- Monitoring, troubleshooting and optimization
- Summary of SAN Design and implementation preferred practices

4.2.1 Architecting a SAN

Before starting with the modernization process, consider these important points in planning:

- Investigation:
  - Gather requirements based on actual and estimated future workloads
  - Gather requirements on business objective and availability
- Collect information:
  - Hardware platform that will be connected to the SAN
  - Software solution and visualization that will be used
- Scalability:
  - Number of needed ports
  - Number of inter-switch links (ISL) to minimize congestion
  - Evaluate the possibility of using UltraScale inter-chassis links (ICL)
- Backup system:
  - Latency and peak loads
- Diagnostics and Manageability:
  - Command line interface (CLI)
  - Graphical user interface (GUI)
  - IBM Network Advisor
- Investment Protection:
  - What device interoperability is required
  - Is NVMe over Fabric support needed

A critical aspect of designing a new SAN or expanding an existing one is collecting all the relevant information and selecting which criteria to meet. After the architectural aspect is defined, an organization can develop the technical specifications that reach the goal criteria that was selected.

4.2.2 Developing technical specifications

The architectural design section (4.2.1, "Architecting a SAN" on page 50) defined some technical aspects and, even more important, the business requirements that the SAN must meet. When developing the technical specification for a new SAN, several main topics must be considered in order to fulfill the requirements defined through that previous process. This section looks at the following topics that emerged:

- Topologies
- Redundancy
- Resiliency

Topologies

The suggested SAN topology to optimize performance, management, and scalability is the core-edge topology. This topology has only one hop from server to storage and uses ICL ports for intra-switching. These attributes assure minimal fabric latency.

High performance and high throughput flash-array storage need solution designs that are low latency. A full-mesh topology allows servers and storage to be placed anywhere. The use of the UltraScale ICL ports for interconnectivity is essential to ensure the bandwidth, minimize the oversubscription, and ensure maximum device port availability and utilization.

Redundancy

The objective of redundancy is to remove any single point of failure. A redundant SAN design has two completely separate fabrics. Servers and storage devices must be connected to both fabrics by using a multipath I/O (MPIO) device driver. MPIO ensures that at least one path is always available. Four paths by volume is the number of paths for which multipathing software, such as SDDPCM and SDDDSM, and the SAN Volume Controller or Storwize, are optimized to work with for performance.

> **Tip:** A large number of paths by volume, eight or even more, does not significantly improve reliability or availability. In some circumstances, a larger number can possibly reduce performance.

Resiliency

Resiliency is built into the Fabric OS with that target of functionality during failure and even recovery from a failure. To work in this manner, resiliency needs special redundancy considerations (to be designed and implemented). For example, the resiliency of a link failover can only be used if a redundant link is available.

Several common types of abnormal behavior originating from fabric components or attached devices can adversely affect SAN health:

- Faulty media (fiber optic cables and small form-factor pluggable optics): These elements can cause frame loss due to excessive cyclic redundancy check (CRC) errors, invalid transmission words, and other conditions. These faults can result in I/O failure and application performance degradation.
- Misbehaving devices, links, or switches: Occasionally, a condition arises where a device (server or storage array) or link (inter-switch link) behaves erratically and causes disruptions in the fabric. If not immediately addressed, this situation might result in severe stress on the fabric.
- Congestion: When caused by latencies or insufficient link bandwidth, some end devices do not respond as quickly as expected. This situation can cause the fabric to hold frames for excessive periods, which can result in application performance degradation or, in extreme cases, I/O failure.
- Credit loss: This issue can happen when the receiving end of a link fails to acknowledge a request to transmit a frame because no buffers are available to receive the frame.

Maintaining an optimal FC SAN environment is essential. Fabric Vision features can improve the reliability and resiliency of the solution and assist with monitoring, protecting, and troubleshooting fabrics. Fabric Vision technology is a set of tools that should be considered to ensure that the fabric is resilient by design.

Data flow considerations

Congestion in fabric can lead to poor performance. These are major types of congestion:

- Traffic-based congestion: Occurs when the link throughput reaches its capacity
- Frame-based congestion: Occurs when the link has run out of buffer credits

With current link speeds, consistently saturating a link is difficult. The more likely issue is frame-based congestion relating to the high number of VMs running on a physical server. Frame-based congestion is caused by latencies on the SAN. The introduction of FlashSystem storage amplifies this effect by accelerating the response times from storage and increasing the rate of exchanges between VMs and storage. This situation can result in credit starvation that is backing up across the fabric. This condition is called *back pressure*. To mitigate frame congestion, the source of poor performance must be addressed.

The Edge Hold Time (EHT) mechanism can help to minimize the effect of some latencies in the fabric. EHT is a time-out value that can cause some blocked frames to be discarded. The EHT default setting for F_Ports is 220 milliseconds and the default EHT setting for E_Ports is 500 milliseconds.

The most important aspect of data flow is the oversubscription. The *oversubscription ratio* is the number of device ports that share a single port. Sizing the oversubscription is not an easy task. Be sure to have a proper balance between source port and target port. The source port's sum of peak bandwidth should not exceed the supported bandwidth of the target port. Monitoring I/O and bandwidth is the key to analyze and tune port traffic. Fabric Vision technology has powerful tools for administrators to help in analyzing and tuning port traffic.

The preferred practice is to enable Fabric Performance Impact (FPI) monitoring on all switches in the fabric and to leave it on to continuously gather statistics. FPI is a Fabric Vision Tool.

Workloads

Workload is the amount of data transmitted. Workload is the product of the number of I/O and block-size (the size of the data transmitted).

Various operating systems and applications generate various types of workloads, and some are noted in the following list:

- Transaction-based applications:
  - Generate high volumes of small block
  - Do not consume a lot of network bandwidth
  - Can lead to frame-based congestion
- I/O-intensive applications:
  - Generate a lot of long block or sequential I/O
  - Typically consume bandwidth
  - Generate latencies and traffic-based congestion
- Host high availability (HA) clustering:
  - Continuously check their connected storage for data integrity reasons
  - Lead to frame-based congestion
- Host-based replication:
  - Generates much traffic typically at the ISLs level
- Array-based replication:
  - Generates much traffic typically at the ISLs level (to mitigate this problem IBM SAN Volume Controller solution requires, if needed, dedicated ISL connectivity)

With IO Insight, monitoring storage or device response latency is possible. VM Insight goes even further and monitors response time at the single VM level. The integration in Gen 6 switches of IO Insight and VM Insight with Flow Vision and Monitoring and Alerting Policy Suite (MAPS) allows for monitoring and alerting about traffic levels. Further, this technology allows monitoring of host and storage device I/O workloads and behaviors. This monitoring can identify and isolate the source of performance degradation or congestion and reduce troubleshooting time. The use of I/O statistics is helpful to optimize the performance of the SAN infrastructure.

Device placement

Device placement refers to where physically the subsystems are connected into the SAN infrastructure. Device placement is a critical point in avoiding congestion. It can allow a balance between traffic isolation, scalability, manageability, and serviceability. To minimize fabric latency to sub-microsecond levels, the best solution is either of the following options:

- A no-hop fabric connection through a single ASIC switch
- Local switching on a single director port blade

When using switches that contain multiple switching ASICs, configuring host and target connections on the ports that share a common ASIC can minimize latency. This minimizing is achieved by avoiding the need to move data across multiple ASICs/port groups or across a director backplane to a different blade. However, this approach does not scale well and other topology must be adopted. In this case, congestion must be mitigated with proper provisioning of ISLs/UltraScale ICLs.

Switch interconnections

Switch interconnections are the physical connections between two or more SAN switches. When planning for how many inter-switch links are needed in order to avoid frame-based congestion, resources must be understood. For example, if not enough ISLs or trunk resources exist, consider data flow.

In addition, to avoid frame-based congestion, redundant links should be considered in environments where there are several data flows between switches. A better approach is to create several two-link trunks rather than one large trunk with multiple links. Trunking provides protection from credit loss on ISLs and optimizes the use of bandwidth by merging a group of ISLs into a single logical link.

The following suggestions might help with optimal performance of ISL trunks:

- Ports used must belong to the same ASIC
- No more than 30m in cable length difference
- Redundant links should be placed on different blades, different ASICs, or at least different port groups whenever possible

In addition to the 32 Gbps connectivity, Gen 6 provides 128 Gbps UltraScale ICL ports using Q-Flex QSFP connectors. The use of UltraScale ICL ports reduce oversubscription because of its huge bandwidth. This type of port also increases SAN infrastructure scalability.

> **Tip:** Only b-type directors offer UltraScale ICL ports. The IBM SAN24B-6 switch does not have any 128 Gbps ports. The IBM SAN64B-6 has four Q-Flex ports where each port can be used as a single 128 Gbps port or four independent 32 Gbps connections using a breakout cable.

4.2.3 Zoning

Zoning plays a key role in a SAN environment and it is used to isolate and segregate traffic. Isolating and segregating is used to specify the devices in the fabric that should be allowed to communicate with each other in an effective, efficient, and secure way.

Preferred practices for zoning are as follows:

- Implement some form of zoning and include all devices
- Draw a zoning diagram layout
- Implement Single Initiator zoning or Peer Zoning

- Develop a naming convention
- Monitor the zoning database size
- Clean up the zoning configuration every time a change happens
- Back up the configuration
- Periodically run SAN Health

Peer zoning

A single-initiator zone can have defined only one Fibre Channel port with initiator functionality (host). However, it can have one or multiple Fibre Channel ports with target functionality (storage). Single-initiator zoning is considered a preferred practice. It is considered the more efficient zoning method in terms of hardware resources and registered state change notification (RSCN) volume. The RSCN is a notification frame that is sent to all devices that are zoned together to update about fabric changes. Adding a unique zone for each host and target can rapidly lead to exceeding zone database size limits and to difficult zone configuration and maintenance.

Figure 4-2 depicts single-initiator zoning.

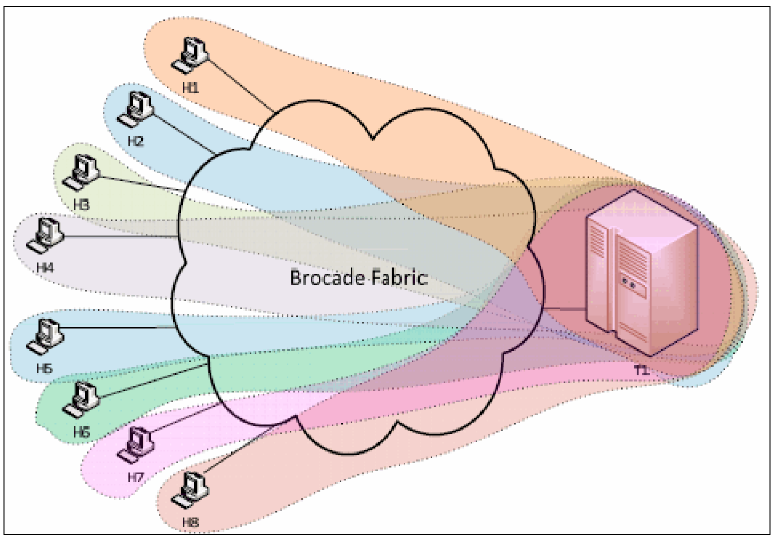


Figure 4-2 Single-initiator zoning

Example 4-1 shows a single-initiator zoning set. All the zones defined are single initiator. They all contain only one intiator port (the port from the host) and, in this case, a single target port (the port from the storage).

Example 4-1 Single-initiator set

```
zone: zone_h1_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:01
zone: zone_h2_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:02
zone: zone_h3_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:03
zone: zone_h4_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:04
zone: zone_h5_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:05
zone: zone_h6_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:06
zone: zone_h7_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:07
zone: zone_h8_stg1 20:00:00:00:00:00:00:00; 10:00:00:00:00:00.00:08
```

With the introduction of peer zoning (FOS 7.4, defined in FC-SW-6 and FC-GS-7standards), the zone membership is differentiated into principal members and non-principal or peer members. Peer zoning configuration allows communication between a principal member and a peer member. However, it does not allow communication between two principal or peer members. By configuring the targets as principal members and hosts as peer members, peer zoning has advantages. For example, one advantage is that peer zoning has fewer zones compared to multiple single-initiator zones and therefore less RSCN distribution.

Figure 4-3 depicts peer zoning.

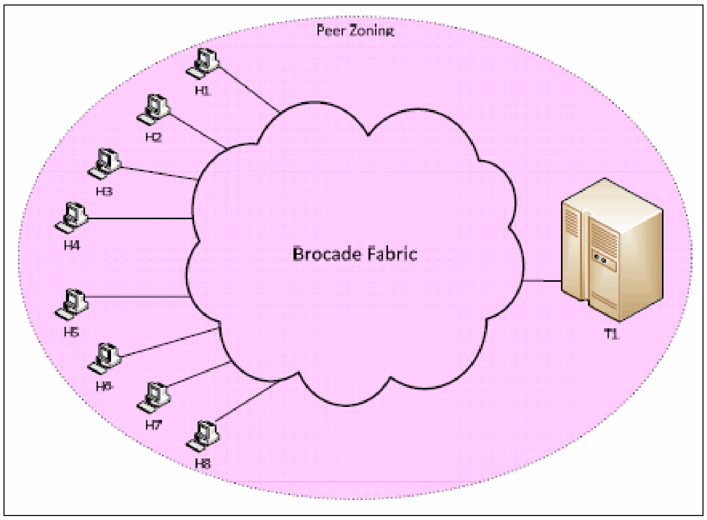


Figure 4-3 Peer zoning

With peer zoning, the result in only one zone. Peer zoning can be completed using a GUI, CLI, and FOS 8.x. Using FOS 8.x, it can also contain an alias. Example 4-2 shows how to implement peer zoning using the CLI.

Example 4-2 Peer zone implementation to create the zone

```
zonecrete –peerzone “peer_h1_h8_stg1” –principal “20:00:00:00:00:00:00:00”
–members “10:00:00:00:00:00.00:01; 10:00:00:00:00:00.00:02;
10:00:00:00:00:00.00:03; 10:00:00:00:00:00.00:04; 10:00:00:00:00:00.00:05;
10:00:00:00:00:00.00:06; 10:00:00:00:00:00.00:07; 10:00:00:00:00:00.00:08”
```

Example 4-3 shows how the peer zone is defined using the CLI and the **zoneshow** or **cfgshow** commands.

Example 4-3 Peer zoning definition

```
zone: peer_h1_h8_stg1
00:02:00:00:00:03:00:01*; 20:00:00:00:00:00:00:00
10:00:00:00:00:00.00:01; 10:00:00:00:00:00.00:02;
10:00:00:00:00:00.00:03; 10:00:00:00:00:00.00:04;
10:00:00:00:00:00.00:05; 10:00:00:00:00:00.00:06;
10:00:00:00:00:00.00:07; 10:00:00:00:00:00.00:08
```

Example 4-3 shows that the first octet has an asterisk beside it. This setting is for the property member of peer zoning and is not a part of intiator or target worldwide port name (WWPN). Property members set signature, creator ID, zone type, and variable data field. This feature did not exist before in zoning.

When using peer zoning, consider the following rules for planning:

- Active configuration can contain both regular zones and peer zones.
- In a single peer zone, a device can be only a principal member or a peer member but not both.
- Peer (non-principal) members are allowed to communicate with a principal member. This means peer members are allowed to communicate with all principal members present in the defined peer zoning.
- Peer (non-principal) members are not allowed to communicate among themselves.
- For peer zones with more than one principal member, the principal members are not allowed to communicate among themselves.
- If a regular zone exists with peer zone members in an active zone set, the non-principal members in the peer zone can still communicate with each other, if they are permitted by the regular zone.

Tip: Using alias names as principal and non-principal members is possible only when all the switches in the fabric are running Fabric OS 8.1.0 or a later version.

4.2.4 Predeployment cabling and optics validation

From the SAN point of view, modernization of the data center requires reviewing Fibre Channel cabling. Cabling that has been in place for some time might not be able to facilitate new requirements that are presently available or future infrastructure planning. Either opportunity could be beyond the cable's capabilities. Moving to the new Gen 6 technology requires a careful check of the Fibre Channel cabling infrastructure. To know if an upgrade is needed, both cable run length and data rate information is required.

In a modernization process, selecting the correct optical transceiver is also important. Gen 6 switches support multiple options. to gather matching information for Fibre Channel cable and optical transceiver, see "Optical transceiver and Fibre Channel cable association" on page 57. After the physical cabling is in place, ClearLink Diagnostics is a useful tool for predeployment cabling and optics validation. ClearLink Diagnostics is part of Fabric Vision tools and it uses diagnostic port (D_Port) to validate configurations prior to deployment.

ClearLink Diagnostics can works with Gen 5 and Gen 6 Fibre Channel optics and cables. This tool is useful to test and validate a single link (in case of troubleshooting) or to perform a complete optical, electrical, and link saturation test to ensure reliable connections.

Note: Testing optics and cabling that is active is disruptive for the data traffic.

For more information about how ClearLink works, see "ClearLink Diagnostics" on page 69.

Optical transceivers

As mentioned previously, Gen 6 supports various optical transceiver options. This section describes the available 32 Gbps features.

The two transceivers that optimize connectivity for servers, storage, and switches within 32 Gbps Fibre Channel fabrics are described in the following list:

- The Brocade 32 Gbps Short Wavelength (SWL) optical transceiver industry-standard small form-factor pluggable (SFP+). Associated with the OM4 Fibre Channel cable for distance up to 100 m.
- The Brocade 32 Gbps Long Wavelength (SLWL) 10 Km optical transceiver industry-standard Small Form-Factor Pluggable (SFP+). Associated with the Singe Mode Fibre Channel cable for greater distance.

Fibre Channel cable

During the modernization process, facilities might continue working with cable that is already in place. To understand if cabling can be used, needs to be replaced, or is a part of future planning, administrators need to verify the ability of cables to meet technical requirements. Prior cabling and new Fibre Channel cabling have different qualities and lengths.

Color is used to identify the properties of the cables. Table 4-1 shows the Telecommunication Industry Association's approved color codification for Fiber Channel cable.

Table 4-1 Color codification

Color	Type of cable
Aqua	OM4 (see note)
Aqua	OM3
Orange	OM1 or OM2
Yellow	Single Mode

Note: The Telecommunication Industry Association in 2009 approved the color codification Aqua for OM4 Fibre Channel cable (Aqua is also used for the previous OM3 Fibre Channel cable). The European market introduced the color violet for OM4. In several countries, you might still find OM4 aqua Fiber Channel cabling.

Optical transceiver and Fibre Channel cable association

The optical transceiver (short wavelength and long wavelength) is associated with the IBM products they leverage, cable quality, and operating distance. The Brocade Hot-pluggable 32 Gbps Short Wavelength (SWL) optical transceivers industry-standard Small Form-Factor Pluggable (SFP+), LC connector is optimized to fully leverage the IBM 32 Gbps backbone, director, and switch products.

The main characteristics of the 32 Gbps SWL SFP+ optical transceiver are as follows:

- 850 nm multimode VCSEL transmitter
- FC-PI-6 compliance for auto-sensing 28.05/14.025/8.5 Gbps
- Diagnostic features per SFF-8472
- Industry-standard LC duplex connector
- 100 m link lengths at 28.05 Gbps on OM4 fiber
- IEC 60825-1 Class 1/CDRH Class 1 laser, eye-safe
- Compliance with the Restriction on Hazardous Substances (RoHS) directive

Table 4-2 shows the range of usable distance for the 32 Gbps SWL SFP+ optical transceiver in relation to Fiber Channel cable quality and the speed used:

Table 4-2 Cable quality and operating distance for 32 Gbps SWL SFP+

Cable	8 Gbps	16 Gbps	32 Gbps
OM2	0.5 - 50 m	0.5 - 35 m	0.5 - 20 m
OM3	0.5 - 150 m	0.5 - 100 m	0.5 - 70 m
OM4	0.5 - 190 m	0.5 - 125 m	0.5 - 100 m

The Brocade Hot-pluggable 16 Gbps SWL optical transceivers industry-standard SFP+, LC connector is optimized to fully leverage the IBM 16 Gbps backbone, director, and switch products.

The main characteristics of the 16 Gbps SWL SFP+ optical transceiver are as follows:

- 850 nm multimode VCSEL transmitter
- FC-PI-5 compliance for auto-sensing 14.025/8.5/4.25 Gbps
- Diagnostic features per SFF-8472
- Industry-standard LC duplex connector
- 100 m link lengths at 14.025 Gbps OM3 fiber
- IEC 60825-1 Class 1/CDRH Class 1 laser, eye-safe
- Compliance with the Restriction on Hazardous Substances (RoHS) directive

Table 4-3 shows the range of usable distance for the 16 Gbps SWL SFP+ optical transceiver in relation to Fiber Channel cable quality and the speed used.

Table 4-3 Cable quality and operating distance for 16 Gbps SWL SFP+

Cable	4 Gbps	8 Gbps	16 Gbps
OM1	0.5 - 70m	0.5 - 21m	0.5 - 15m
OM2	0.5 - 150 m	0.5 - 50 m	0.5 - 35 m
OM3	0.5 - 380 m	0.5 - 150 m	0.5 - 100 m
OM4	0.5 - 400 m	0.5 - 190 m	0.5 - 125 m

The Brocade Hot-pluggable 32 Gbps Long Wavelength (LWL) optical transceivers industry-standard Small Form-Factor Pluggable (SFP+), LC connector is optimized to fully leverage the IBM 32 Gbps backbone, director, and switch products.

The main characteristics of the 32 Gbps LWL SFP+ optical transceiver are as follows:

- 1310 nm multimode VCSEL transmitter
- FC-PI-6 compliance for auto-sensing 28.05/14.025/8.5 Gbps
- Diagnostic features per SFF-8472
- Industry-standard LC duplex connector
- 10 km link lengths on 9 micrometers single-mode fiber
- EC 60825-1 Class 1/CDRH Class 1 laser, eye-safe 1310 nm DFB laser
- Compliance with the Restriction on Hazardous Substances (RoHS) directive

Attention: Do not look through the optical port because it is a potential eye hazard.

Note: Single-mode Fibre Channel cable does not have quality classifications. Multi-mode Fibre Channel cable does have quality classifications.

4.2.5 Monitoring, troubleshooting and optimization

This section reviews some of the main monitoring, troubleshooting, and optimization tools available to manage SAN.

Availability and health monitoring

After a SAN is in place, it is important to monitor that is functioning properly. To monitor a SAN, a GUI management application is not necessary but is useful. Some of the benefits of monitoring a SAN are tuning, proactive actions preventing congestion, and rapid troubleshooting.

IBM Network Advisor

IBM Network Advisor is an optional software management application that provides easy, centralized, and comprehensive management for storage networks. This single tool can deliver quick access to all product configuration applications, end-to-end visibility, and insights across various network types. IBM Network Advisor supports Fibre Channel SANs (including Gen 5 and Gen 6 Fibre Channel platforms, IBM FICON and IBM b-type SAN extension solutions). Using this application, you can configure, manage, and monitor your networks with ease. The management application's main window contains useful tabs, logs, and organized information. The following list notes some of those areas:

- Dashboard tab: Provides a high-level overview of the network managed by widgets, and displays operational status, inventory status, event summary, and overall network/fabric status
- SAN tab: Provides the Product List and Connectivity Map (topology).
- Master Log: Lists the events and alerts that have occurred on the SAN.

IBM Network Advisor is available in the following versions:

- SAN Professional: up to 1000 switch ports, one fabric
- SAN Professional Plus: supports up to 2560 switch ports and 36 fabrics
- SAN Enterprise: supports up to 9000 switch ports and 36 fabrics

Monitoring and Alerting Policy Suite (MAPS)

Brocade MAPS is the Fabric Vision feature for monitoring and alerting. MAPS has embedded 20 years of experience building, delivering, and monitoring SAN infrastructure. MAPS works with already predefined policies for monitoring the SAN environment.

Some of the predefined policies that MAPS offers are as follows:

- Port Health
- Traffic Performance
- Fabric Performance Impact

MAPS is automatically enabled during the installation of Fabric OS 7.4.0 or later versions. However, without the Fabric Vision license, only a limited set of MAPS functions are available. A Fabric Vision license enables the full functionality of MAPS. The following advanced technologies and capabilities are available with the optional Fabric Vision technology license:

- MAPS
- Flow Vision
  - IO Insight
  - VM Insight

Tools for gathering data

Data collection takes a snapshot of the infrastructure and includes information about the connected devices, zoning, and statistics. Data collection is critical for inventory analysis and troubleshooting. This section describes several possibilities of gathering data with the following tools:

- SAN Health
- Supportsave

SAN Health

SAN Health is a web downloadable program from Brocade Communications Systems, LLC. SAN Health is available at no cost and helps you gather an accurate view of the SAN environment. The SAN Health Diagnostics Capture utility is a powerful tool to collect data, identify potential issues, and check the results over time. After capturing switch diagnostic data, it automatically generates a Microsoft Visio topology diagram and a detailed report of SAN fabrics, switches, and individual ports.

The SAN Health Diagnostics Capture utility that works with Brocade is available to download. The utility helps to collect the data that must be provided to Brocade to get the SAN report. With this report the SAN administrator will be able to take the following actions:

- Inventory devices, switches, firmware versions, and SAN fabrics
- Capture and display historical performance data
- Compare zoning and switch configurations against preferred practices
- Assess performance statistics and error conditions
- Produce detailed graphical reports and diagrams

For more information, see the Brocade SAN Health web page.

Supportsave

To be useful, statistical data inside an IBM SAN Switch log needs to have a baseline. A baseline setup is done by clearing the SAN switch statistics.

The **`supportsave`** Fabric OS (FOS) command can be run by using the CLI. This command collects all needed data for IBM to perform troubleshooting. Be aware that **`supportsave`** needs an FTP sever running on a workstation reachable by the switch in order to save the output.

The following preferred process collects a log that will be sent to IBM for analysis:

1. Run **`supportsave`** to get the current status of the switch.
2. Clear counters by using the **`supportinfoclear –clear`** CLI command.
3. Wait 3 hours.
4. Run **`supportsave`** again to collect a new log with a baseline.
5. Compress all collected files before sending them to IBM.

> **Tip:** Do not collect data using the web interface. Use **`supportsave`** to collect data with Network advisor or CLI.

4.2.6 Summary of SAN Design and implementation preferred practices

The key to a useful SAN design is to keep it simple and deploy the features that are really needed. The following list highlights the main preferred practices for a SAN design:

- Starting pre-deployment:
  - Use ClearLink Diagnostics to check cables and optics.
- Managing Fabric Behavior:
  - Verify Edge Hold Time.
  - Enable Buffer Credit recovery tools.
  - Enable Fabric Performance Impact (bottleneck monitoring).
  - Enable Slow Drain Device Quarantine (SDDQ). To do so, quality of service (QoS) must be enabled on all switches (which is the factory default).
- Managing ports:
  - Set all ports set with fixed speed. This fixed setting helps in troubleshooting. When ports are set to auto-negotiate port speed, it can lead to connections being at a different speed than expected.
  - Use Dynamic Portname. This helps populate information in various fields, such as switch name, port type, port index, and alias name.
  - Enable the Diagnostic Port, D_Port setting. It is not automatically enabled and must be enabled in order to function.
  - For security reasons, when not in use, set the ports to `persistent disable`.
- Managing long-distance fabrics:
  - Set the ISL to LS mode. This setting permits buffer credits to be based on a user-defined desired distance value.
- Zoning. Use preferred practice. Peer or single-initiator zoning can save zone-database size.
- Monitoring, troubleshooting, and optimization:
  - Use Fabric Vision features for monitoring and alerting.
  - SAN Health for inventory and topology.

4.3 When, where, and how to use b-type products

Currently, with data increasingly growing, hyperscale virtualization, high demand workloads, and more business operations, the IT storage infrastructure needs addressing. Infrastructures can need restructuring to continue to better serve customers and increasing demands.

Many IT organizations are already modernizing their structures for current and future workloads. Gen 6 is available to provide fast access to data, high-performance, high-security, low-latency, so on, all predicates that the applications need. Legacy infrastructure simply was not designed to support these new technologies. Data centers that do not restructure will lose demand, will lose field and, finally, the customer.

The b-type Gen 6 Fibre Channel combines hardware innovation, software, and integrated network sensors, ensuring high levels of operational stability and application performance. In view of unpredictable virtualized workloads, data growth, and all-flash storage environments, organizations need to ensure that the network does not become a bottleneck. Following this new era, the b-type Gen 6 can be used in many types of environments and meets the most diverse of data centers requirements.

Application response time performance can be achieved by upgrading the network to 32 Gbps with Gen 6 Fibre Channel without changes being made in storage systems.

Looking to the future, organizations can perfectly integrate b-type Gen 6 Fibre Channel networks with next-generation NVMe flash storage over fabrics. The efficiency of NVMe plus the high performance and low latency of Gen 6, enable the administrators to accelerate IOPS. With greater IOPS, systems can deliver the high performance, better response time, and scalability needed for next-generation data centers.

The b-type Gen 6 has a portfolio to attend all demands, from environments that require entry level switches to environments that require high-end enterprise directors. Their use depends where the user is and where they want to go. Due to the extensive features of Gen 6 it can be integrated with legacy environments normally. At the same time, looking for investment protection, b-type Gen 6 directors offer three generations of backward-compatibility support for connectivity to 4, 8, and 16 Gbps Fibre Channel products. This product allows for connectivity between older and newer generations.

4.3.1 The differences between Gen 5 and Gen 6

The use of Gen 5 or Gen 6 technology depends on user needs and requirements. Gen 6 technology is the most advanced of today, bringing significant innovations and significant modifications to current needs. Also, this product already glimpses the new era of increasingly powerful equipment (for example with NVMe technology). At the same time, to protect your investment, Gen 6 maintains connectivity with previous generations.

Gen 5 is a solution that aims for high performance, high demand, fast deliveries, and scalability. Compared to Gen 6, it is a more conservative option in investment and technology. Gen 6 with NMVe flash storage over fabrics can provide the latest advanced performance capability and, with Fabric Vision IO Insight and VM Insight, can change the whole scenario on fabric network management.

The following differences exist between Gen 5 and Gen 6:

- The Gen 6 ASIC is capable of handling 33% more IOPS than the prior generation ASIC used in Gen 5 SAN technology.
- Gen 6 Fibre Channel with 32/128 Gbps links increases I/O performance and throughput to complete workloads faster. At the same time, providing a highly scalable infrastructure that supports massive application growth. The higher throughput delivered by Gen 6 supports double the VM density of Gen 5, providing larger server utilization and ensuring optimized performance for high VM deployments.
- With 128 Gbps links, Gen 6 provides up to 8x the bandwidth compared to Gen 5. This surpasses today's most advanced server and infrastructure environments and also provides scalability for VM environment growth.
- Gen 5 FICON technology is capable of outstanding performance in an IBM z13® mainframe solution. Gen 6 FICON technology takes this performance to a whole new level and, when combined with the z14 mainframe and all-flash arrays, enables the solution's full performance potential.
- Fabric Vision VM Insight and IO Insight features are available only with Gen 6. Gen 5 does not have these Fabric Vision features.
- Gen 5 and Gen 6 Fibre Channel switches already support NVMe over Fabrics, with no changes required.
- Gen 6 runs only on FOS 8.X versions. Gen 5 runs on FOS 7.X versions.

Table 4-4 lists differences between SAN b-type Gen 5 and Gen 6 entry and midsize switches.

Table 4-4 Gen 5 and Gen 6 storage network product differences in entry and midsize switches

	IBM Gen 5			**IBM Gen 6**	
IBM model	**IBM SAN24B-5**	**IBM SAN48B-5**	**IBM SAN96B-5**	**IBM SAN24B-6**	**IBM SAN64-6**
Brocade model	Brocade 6505 Switch	Brocade 6510 Switch	Brocade 6520 Switch	Brocade G610	Brocade G620 Switch
IBM machine type and model number (MTM)	2498-X24	2498-F48	2498-F96 / 2498-N96 (airflow)	8960-F24	8960-F64/N64 (airflow)
Ports	12, 24	24, 36, 48	48, 72, 96	8, 16, 24	24, 36, 48 SFP+ 4 QSFP for up to 64 total ports
Speed & protocol	16/10/8/4/2 Gbps FC	16/10/8/4/2 Gbps FC, Fibre Channel Routing (FCR)	16/10/8/4/2 Gbps FC, FCR	32/16/8/4 Gbps FC	128/32/16/8/4 Gbps FC, FCR
Height	1 RU	1 RU	2 RU	1 RU	1 RU
FICON support	No	Yes	No	No	Yes
Hot Code Activation	Yes	Yes	Yes	Yes	Yes
Expandability	Ports on Demand: 12 Ports (FC 7212)	Ports on Demand: 12 Ports (FC 7411)	Ports on Demand: 24 Ports (FC 7812)	• Ports on Demand: 8 Ports Ports on Demand (FC 7490) • Eight 16 Gbps SFP+ (FC 7498) 12 SFP+ Ports with 32 Gbps SFP+ • Zero (0) SFP+ (FC 7499)	• Ports on Demand (FC 7490) • 12 SFP+ Ports with 32 Gbps SFP+ • Ports on Demand (FC 7496) No SFP+ • QFlex Ports on Demand (FC 7495) with 4 SWL QSFP with 4x32Gb breakout • QFlex Ports on Demand (FC 7497) with No QSFP
Airflow	Non-port to port side	Non-port to port side	• F96: Non-port to port side • N96: Port to non-port side	Non-port side to port side	• F64: Non-port side to port side • N64: Port side to non-port side
Hardware features	• 1 Hot-swappable, redundant, integrated power supply/fan unit • Fixed rack-mount kit included • Mid-mount rack kit optional (FC 5911) *Optional Hot-swappable, redundant, integrated power supply/fan unit (FC 5998)	• 2 Hot-swappable, redundant, integrated power supply/fan units. • Fixed rack-mount kit included • Mid-mount rack kit (FC 5911)	• 2 hot-swap, redundant power supplies • 3 hot-swap, redundant fan units • Fixed rack-mount kit included • Mid-mount rack kit option (FC 5910)	• Single fixed power supply • Fixed redundant fans • Single switch FRU • Fixed rack-mount kit included • Mid-mount rack kit option (FC 5910)	24 SFP+ ports enabled, 2 hot-swappable / redundant integrated power supplies/fan units, Rack mount kit
Standard software	• Web Tools • Adaptive Networking (FOS 7.2.0c)	• Web Tools • Adaptive Networking (FOS 7.2.0c) • Fabric Watch	• Web Tools • Adaptive Networking (FOS 7.2.0c)	• Advanced Zoning • Full Fabric • Web Tools • Adaptive Networking	• Web Tools • Adaptive Networking

	IBM Gen 5			IBM Gen 6	
IBM model	**IBM SAN24B-5**	**IBM SAN48B-5**	**IBM SAN96B-5**	**IBM SAN24B-6**	**IBM SAN64-6**
Options	F24 Enterprise Package (FC 7216) includes all the individual options below except Extended Fabrics. Individual Optional Features: • Adaptive Networking (FC 7206) • Advanced Performance Monitoring (FC 7204) • Extended Fabrics (FC 7203) • Fabric Watch (FC 7202) • ISL Trunking (FC 7205) • Fabric Vision (FC 7219) **Note:** If you install Advance Performance Monitoring and Fabric Watch licenses you are Fabric Vision capable without the Fabric Vision license.	F48 Enterprise Package (FC 7416) includes all the individual options below except FICON CUP, Integrated Routing, and Integrated 10G. Individual Optional Features: • Adaptive Networking (FC 7406) • Advanced Performance Monitoring (FC 7404) • Extended Fabrics (FC 7403) • FICON CUP (FC 7409) • Integrated 10G FC License (FC 7413) • Integrated Routing (FC 7407) • ISL Trunking (FC 7405) **Note**: Fabric Vision license not offered a la carte. If you install APM (FC 7404) or Enterprise Package (FC 7416) with the standard installed Fabric Watch license then Fabric Vision is enabled.	F96/N96 Enterprise Package (FC 7816) includes all the individual options below except Integrated Routing and Integrated 10G. Individual Optional Features: • Advanced Performance Monitoring (FC 7804) • Extended Fabrics (FC 7803) • Fabric Watch (FC 7802) • Integrated 10G FC License (FC 7813) • Integrated Routing (FC 7807) • Integrated Routing (FC 7805) • Fabric Vision (FC 7819) **Note**: If you install Advance Performance Monitoring and Fabric Watch licenses you are Fabric Vision capable without the Fabric Vision License.	F24 Base Switch (FC 2445): • 8 ports and 0 SFP+ F24 Base Switch (FC 2446): • 8 ports and 16 Gbps SFP+ F24 Enterprise Switch Bundle (FC 2447): • 24 ports and 16 Gbps SFP+ • ISL Trunking • Fabric Vision • Extended Fabric F24 Enterprise Software Package (FC 7422): • Extended Fabrics • ISL Trunking • Fabric Vision	F64 Enterprise Package (FC 7420) includes: • Extended Fabrics • ISL Trunking • Fabric Vision with IO and VM Insight Mainframe Enterprise Package (FC 7421) includes: • Extended Fabrics • ISL Trunking • Fabric Vision with IO and VM Insight • FICON CUP Integrated Routing (FC 7407)
Other base features	• Advanced Diagnostic Tools • Access Gateway • Fabric Performance Impact (FOS 8.0+) • Buffer credit recovery (ASIC) • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • EZ Switch Setup Wizard • Forward error correction • NPIV Support • Security • Traffic Isolation Zones	• Advanced Diagnostic Tools • Access Gateway • Fabric Performance Impact (FOS 8.0+) • Buffer credit recovery (ASIC) • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • EZ Switch Setup Wizard • Forward error correction • NPIV Support • Security • Traffic Isolation Zones • Virtual Fabrics • In-Flight Encryption and Compression	• Advanced Diagnostic Tools • Fabric Performance Impact (FOS 8.0+) • Buffer credit recovery (ASIC) • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • Forward error correction • In-Flight Encryption and Compression • NPIV Support • Redundant Core Switching & Control Processor blades • Security • Traffic Isolation Zones	• Advanced Diagnostic Tools • Bottleneck Detection • Fabric Performance Impact • Buffer credit recovery • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • Forward error correction • NPIV Support • Security • Traffic Isolation Zones	• Advanced Diagnostic Tools • Access Gateway • Fabric Performance Impact • Buffer credit recovery • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • Forward error correction • In-Flight Compression • In-flight Encryption (planned FOS v8.2) • NPIV Support • Security • Traffic Isolation Zones • Virtual Fabrics
Fabric OS	FOS 7.0.1 or later	FOS 7.0.0a or later	FOS 7.1.0c or later	FOS 8.1.0 or later	FOS v8.0.1b or later

Table 4-5 notes some of the differences in SAN b-type Gen 5 and Gen 6 directors.

Table 4-5 SAN B-type Gen 5 and Gen 6 (storage network products) directors

	IBM Gen 5	**IBM Gen 6**
IBM model	**IBM SAN768B-2 / SAN384B-2**	**IBM SAN512-6/SAN256-6**
Brocade model	Brocade DCX 8510-8 / 4	Brocade X6-8 / X6-4
IBM machine type and model number (MTM)	2499-816 / 416	8961-F08 / F04 (airflow)
Ports	• Up to 512 SFP+ and 32 UltraScale Interchassis Links • Up to 256 SFP+ and 16 UltraScale Interchassis Links	• Up to 384 SFP+ and 32 UltraScale Interchassis Links • Up to 192 SFP+ and 16 UltraScale Interchassis Links
Speed & protocol	16/8/4 Gbps FC, FCR, 1/10 GbE FCIP	128/32/16/8/4 Gbps FC, FCR, 1/10/40 GbE FCIP
Height	14 RU / 8RU+1RU shelf	14 RU / 8RU+1RU shelf
FICON support	Yes	Yes
Hot Code Activation	Yes	Yes
Expandability	• Scale to 8 blades for SAN768 • Scale to 4 blades for SAN384B • Gen 5 ICL POD (FC 7871) • (16) 64 Gbps ICL ports w/QSFPs for 50m, Gen 5 ICL POD (FC 7874) • (16) 64 Gbps ICL ports w/QSFPs for 100m	• Scale to 8 blades for SAN512B • Scale to 4 blades for SAN256B • Gen 6 ICL POD (FC 7982) • (16) 128 Gbps ICL ports w/QSFPs
Airflow	Non-port to port side	• 1 FAN tray non-port side intake (FC 8002) • 1 FAN tray non-port side exhaust (FC 8003)
Hardware features	• SAN768B: 2 Control Processors, 2 Core Blades, 2 Power Supplies, 3 Fan Blowers, rack mount. • SAN384B: 2 Control Processors, 2 Core Blades, 2 Power Supplies, 2 Fan Blowers, rack mount. No port blade are included	• SAN512B: 2 Control Processors, 2 Core Blades, 3 Power Supplies, 3 Fan Blowers, cable comb, rack mount. •SAN256B: 2 Control Processors, 2 Core Blades,2 Power Supplies, 2 Fan Blowers, rack mount. No port blades are included
Standard software	• Web Tools • Extended Fabrics • Fabric Vision • ISL Trunking • Adaptive Networking	• Fabric Vision with IO and VM Insight • Web Tools • Extended Fabrics • ISL Trunking • FICON CUP • Adaptive Networking
Options	• Gen 5 ICL Port On Demand (FC 7871) Sixteen (16) 4x16 Gbps ICL ports. Includes 50 m QSFP transceivers, no cables • Gen 5 ICL Port On Demand (FC 7874) Sixteen (16) 4x16 Gbps ICL ports. Includes 100 m QSFP transceivers, no cables • Gen 5 ICL Port On Demand (FC 7875) Eight (8) 4x 16 Gbps ICL ports. Includes 8x 2 Km QSFP transceivers, no cables. • Enterprise ICL for >5 chassis (FC 7873/72) • Integrated Routing (FC 7889/90) • Integrated 10 Gbps License (FC 7892) • Advanced Extension (FC 7891) • FICON Accelerator (FC 7893) • FICON CUP (FC 7886/84)	• Integrated Routing (FC 7897) • Gen 5 ICL Port On Demand: Sixteen (16) 4x 16 Gbps ICL ports. Includes qty 16x 16 Gbps 100m QSFP transceivers, no cables (FC 7980) • Gen 5 ICL Port On Demand: Eight (8) 4x 16 Gbps ICL ports. Includes qty 8x 16 Gbps 2km QSFP transceivers, no cables (FC 7981) • Gen 6 ICL Port On Demand: Sixteen (16) 4x 32 Gbps ICL ports. Includes qty 16x 32 Gbps 100m QSFP transceivers, no cables (FC 7982) • Gen 6 ICL Port On Demand: Sixteen (16) 4x 32 Gbps ICL ports. Includes qty 8x 32 Gbps 2km QSFP transceivers, no cables (FC 7984) • Gen 6 ICL Port on Demand: Eight (8) 4x 32 Gbps ICL ports. Includes qty 8x 32 Gbps 2km QSFP transceivers, no cables (FC 7944)

	IBM Gen 5	IBM Gen 6
IBM model	**IBM SAN768B-2 / SAN384B-2**	**IBM SAN512-6/SAN256-6**
Other base features	• Advanced diagnostic tools • Fabric Performance Impact (FOS 8.0+) • Buffer credit recovery (ASIC) • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • Forward error correction • In-Flight Encryption and Compression • NPIV Support • Redundant Core Switching & Control Processor blades • Security • Traffic Isolation Zones • Virtual Fabrics	• Advanced diagnostic tools • Fabric Performance Impact • Buffer credit recovery • ClearLink (D_Port) Diagnostics • Dynamic Fabric Provisioning • Dynamic Path Selection • Forward error correction • In-Flight Compression • In-flight Encryption (planned FOS v8.1) • NPIV Support • Redundant Core Switching & Control Processor blades • Security • Traffic Isolation Zones • Virtual Fabrics
Fabric OS	FOS 7.0.0a or later	FOS v8.0.1b or later

4.3.2 Importance of Gen 6

With technology demanding high performance and fast connectivity, IBM has responded. The IBM b-type Gen 6 combined with an all flash storage design achieves the objective to reach the maximum for low latency and high-throughput. Investments are dedicated to infrastructure to ensure that IT organizations can achieve SLAs. Benchmark tests show that faster network speeds improve the overall ability of flash storage to maximize IOPS, even if the storage connectivity speed is less than the maximum network throughput.

IT organizations are investing significantly to obtain the maximum performance in flash storage. However, they are not only looking at flash storage but also into connectivity to the servers. To take maximum advantage of flash based storage, the network needs advancements. Organizations are tailoring applications with flash storage and networks are requiring low latency, higher bandwidth capacity, and greater reliability.

No matter the maximum network throughput and independent of connectivity speed, flash storage has the ability to maximize IOPS. Integrating environments with an IBM b-type Gen 6 network fabric with 32 Gbps connectivity to the server and 8 Gbps Fibre Channel connectivity to flash storage, enables greater processing. Within the same time span, the system is able to perform up to four times more queries. With Gen 6 connectivity, the application response time is enhanced significantly. The following list notes some examples of the differences:

- On average, 50% faster than 6 Gbps storage networks with host bus adapters (HBAs)
- On average, 71% reduction in the application response time of an 8 Gbps storage network and HBAs

Figure 4-4 shows the Demartek test results, comparing the response times for various application workloads (lower results are better)[1].

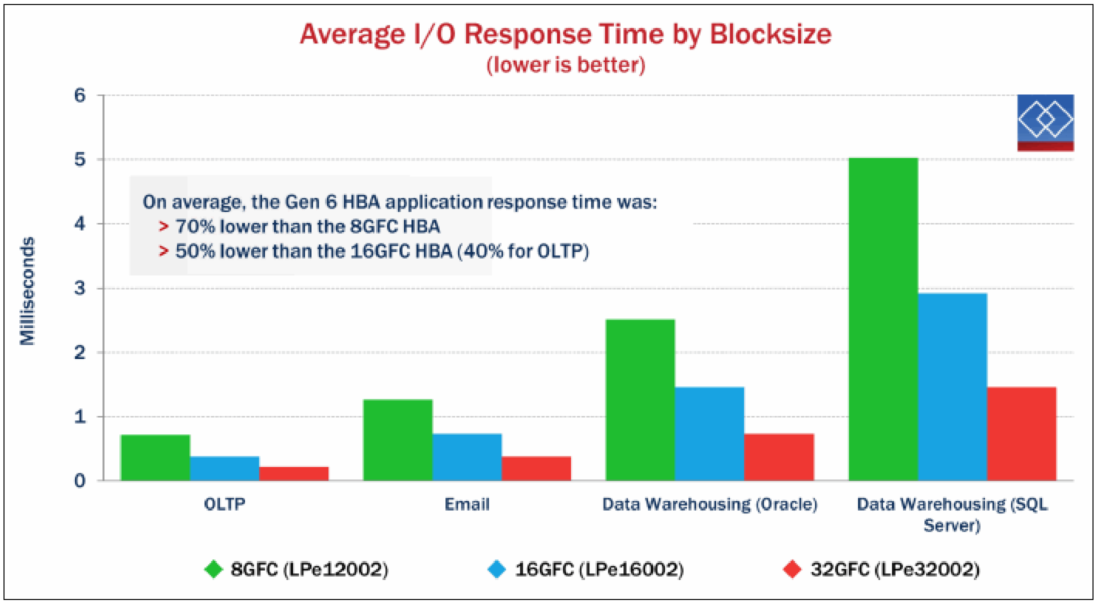


Figure 4-4 Average I/O response time by block size (lower is better)

Gen 6 is important for flash to be able to meet all the demands that are required and to extract the maximum of its technological capacity and features. Gen 6 is the interface that will make all business operations happen faster and safer. Gen 6 Fiber Channel maximizes the performance of flash arrays even when at speeds less than 32 Gbps. With the need to increase the environment, and add more flash storage into an environment with fast connectivity, Gen 6 can support growth. With Gen 6, this growth can happen without concern of bottlenecks with application performance.

4.3.3 NVMe over Fibre Channel

Non-Volatile Memory Express (NVMe) is a communication interface and protocol developed to reduce overhead from drivers, applications and operating systems (OS). NVM Express defines an interface for hosts to communicate with non-volatile memory subsystems over PCI Express (NVMe over PCIe).

The next wave of all flash arrays will provide flash media connectivity to NVMe over Fibre Channel, providing reduced latency and performance gains.

To avoid storage response latency, flash media connected to NVMe over Fibre Channel allows direct communication from the host to the storage. This is done by using the Component Interconnect Express (PCIe) peripheral commands, bypassing the Small Computer System Interface (SCSI) translation layer.

In a traditional disk-based system storage device, this translation time is not significant. However, translation time is relevant in flash-based array where response time is much faster.

NVMe over Fibre Channel can be integrated with the Gen 6 Fiber Channel. Doing so enables storage architects to take advantage of the storage media performance connected to NVMe while using the existing Fiber Channel network. Fiber Channel networks are ready for use

[1] This figure is reprinted here with permission from Demartek, LLC. See the full article Accelerating Oracle Data Warehousing with Emulex and Brocade Gen 6 Fibre Channel, July 2016.

with NVMe over fabrics. These networks provide the same design features as fast throughput, resiliency, and high reliability. This ensures that organizations integrating Gen 6 Fiber Channel networks with the new NVMe generation will be fully supported.

NVMe over Fibre Channel provides some of the highest performance and lowest latency storage. With Gen 6 Fiber Channel, organizations can increase IOPS to deliver fast, secure data, and still have scalability for next generation data centers.

4.4 When, where and how to use Fabric Vision features

Data center and SAN infrastructures are more than speed. Fabric Vision features are a suite of monitoring, management, and diagnostic tools. These tools are predefined with threshold base rules and policies, dashboard views, and live monitoring. Gen 6 brings storage I/O level monitoring (IO Insight) and VM host level monitoring (VM Insight) to the Fabric Vision Technology. These additional aspects monitor the flow level of a device port or a VM host, and identify and alert on matters of latency and congestion.

4.4.1 Fabric Vision tools overview

This section provides an overview of Fabric Vision features:

- ClearLink Diagnostics
- Forward error correction (FEC)
- Buffer credit recovery
- Monitoring and Alerting Policy Suite (MAPS)
- Dashboards
- COMPASS
- Fabric Performance Impact (FPI) monitoring
- Flow Vision
- IO Insight
- VM Insight

Figure 4-5 shows an overview of the Fabric Vision capabilities and system design.

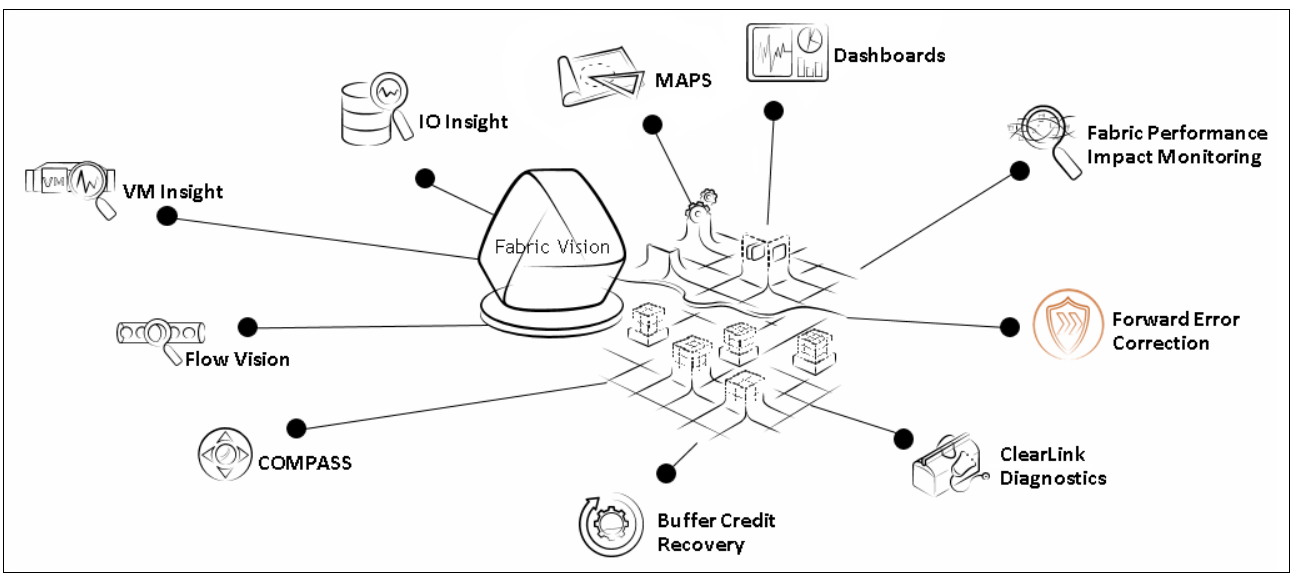


Figure 4-5 Fabric Vision features and capabilities

ClearLink Diagnostics

ClearLink Diagnostics is a cable and optic diagnostics program. It simplifies the deployment and support of large fabrics and uses the Diagnostic Port (D_Port). Brocade ClearLink Diagnostic Port helps to validate configurations prior to deployment. It does so by performing complete optical, electrical, and link saturation testing. This testing ensures reliable connections avoiding downstream problems, reducing fabric deployment time, supporting ISLs, and host-to-switch links.

Features highlight

These are the main features of ClearLink Diagnostics:

- Ensures optical and signal integrity for Gen 5 and Gen 6 Fibre Channel optics and cables
- Measures latency and distance across ISLs to ensure predictable application performance
- Provides link power loss information

How ClearLink Diagnostics works

ClearLink Diagnostics works in the following ways; the numbers indicated in parentheses correspond to the functions shown in Figure 4-6:

- Electrical loopback: The electrical loopback phase of the diagnostic suite isolates issues associated with the seating of the SFP (1) into the port receptacle.
- Optical loopback: The optical loopback phase of the diagnostic suite isolates issues associated with contamination or seating of the fiber optic cable (2) into the SFP.
- Link traffic: The link saturation phase of the diagnostic suite isolates issues associated with marginal component performance: cable (3), SFP (4), and port (5).
- Link latency and distance measurement: At the end of the link saturation phase of the diagnostic suite, ClearLink Diagnostics reports the link latency, link distance, and power loss (dB) measurements. These measurements provide input for administrators to properly design cable layout and configure buffer-to-buffer credits.

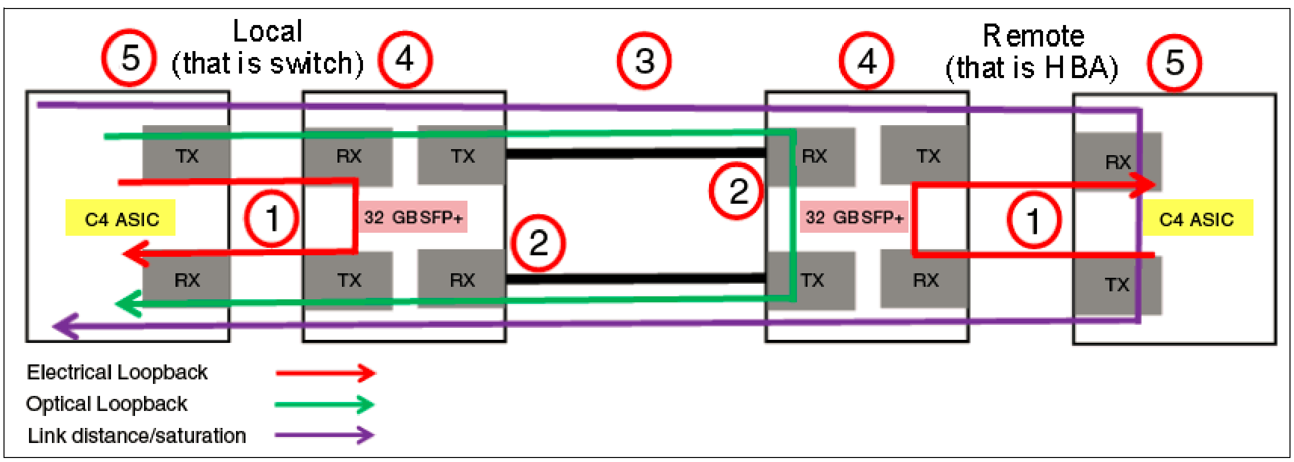


Figure 4-6 ClearLink Diagnostic functions

Brocade ClearLink Diagnostics has a full MAPS integration and an enhanced CLI output for clarity.

> **Tip:** To run D_Port tests between a switch and a non-Brocade HBA, the Fabric Vision license is required and the HBA vendor must have implemented Brocade HBA D_Port support.

To start a port in D_Port mode, follow these basic steps within a GUI or CLI:

1. Disable the ports on both the ends of the link.
2. Enable D_Port functionality on both ends of the link.
3. Enable the ports on both the end of the link.

Forward error correction (FEC)

FEC enables ASIC to recover bit errors and enhances the reliability of transmission and thus performance. FEC is enabled by default to ensure reliability for all high-speed 32 Gbps links.

Buffer credit recovery

Buffer credit loss can lead to performance degradation. With buffer credit recovery, single buffer credit loss is detected and recovered automatically at the virtual channel (VC) level on Fibre Channel ISLs. The loss of multiple credits and stuck VC (with no credits) are recovered automatically using link reset with RASlog notification. Buffer credit recovery is supported on normal and long-distance ISLs.

Monitoring and Alerting Policy Suite (MAPS)

MAPS is automation that simplifies policy-based monitoring and alerting. It works with four predefined sets of policies: base, moderate, conservative, and aggressive. Policies can be customized and personalized. MAPS can be set up to send email notifications about threshold violations, SNMP traps, and RASlog events.

Dashboards

Dashboards are customizable health and performance overviews, with all critical information on one screen. Administrators can create a custom dashboard that includes a widget from existing Brocade Fabric Vision technology features. The widget features can include MAPS dashboards, Fabric Performance Impact (FPI) monitoring violations, and IO Insight metrics.

COMPASS

Configuration and Operational Monitoring Policy Automation Services Suite (COMPASS) simplifies SAN configuration and maintains consistency in operational behavior. COMPASS monitors configuration changes between a configured setting and the switch configuration and helps to more easily provision new switches in a fabric. It uses user-defined templates to compare the configuration settings on a switch to configuration settings in a template. This enables users to monitor subsets of switch configuration and receive notification when a drift of any of the defined configuration settings occur.

> **Tip:** COMPASS is supported in only Network Advisor Professional Plus and Enterprise editions.

Fabric Performance Impact (FPI) monitoring

FPI monitoring quickly detects and clearly alerts administrators to high levels of latency. This feature helps to identify slow-drain devices.

Benefits highlight

These are the main benefits of Fabric Performance Impact monitoring:

- Latency notifications
- Hardware counters monitoring at sub-microsecond levels
- Brief severe and sustained abnormal latency conditions detection
- Automatic slow drain device mitigation and recovery

To mitigate a device-based latency, manual actions are usually required to move slow responding hosts or arrays to a different path. These actions are disruptive for the traffic. With Gen 6 SAN switches, Fabric Vision can use two automatic actions associated with the FPI rules to address these issues:

- Slow-Drain Device Quarantine (SDDQ)

  This action is automatically triggered when FPI detects an F_Port in either the IO_PERF_IMPACT state or the IO_FRAME_LOSS state. The SDDQ action sends the port identifier (PID) that is associated with the slow draining device to all switches in a fabric. All switches move the traffic destined for the slow draining device into a low-priority virtual channel (VC). As a result, buffer credits on the regular, medium-priority VC are freed for traffic destined to other devices. This effectively removes the impact of a slow drain device to the fabric performance without disruption to traffic. The result of this action is that slow-drain devices are isolated in a quarantine but remain online. To use the SDDQ action, the switches in the fabric are required in order to enable quality of service (QoS) on all ports in the flow path.

- Port Toggle

  This action disables a port for a short and user-configurable duration and then re-enables the port. The port toggle action can recover slow draining devices such as those caused by a faulty host adapter. In addition, the Port Toggle action can induce multipath I/O (MPIO) software to trigger traffic failing over to an alternate path to prevent severe performance degradation.

By using the SDDQ or Port Toggle actions, administrators can monitor for device-based latency and automatically mitigate the problem when such conditions are detected by FPI.

> **Tip:** FPI monitoring is enabled by default on switches running Fabric OS v8.0 or later.

Flow Vision

Flow Vision identifies, monitors, and analyzes the performance of specific flows or frame types. A flow is a set of FC frames or packets and it can be defined by port, direction, or frame parameter. Flow Vision has three features: Flow Monitor, Flow Generator, and Flow Mirror.

MAPS can create policies for events related to flows and sub-flows and filter list of MAPS violations for imported flows or sub-flows.

IO Insight

IO Insight automatically detects degraded storage I/O performance with integrated device latency and IOPS monitoring.

Figure 4-7 shows IO Insight workload monitoring.

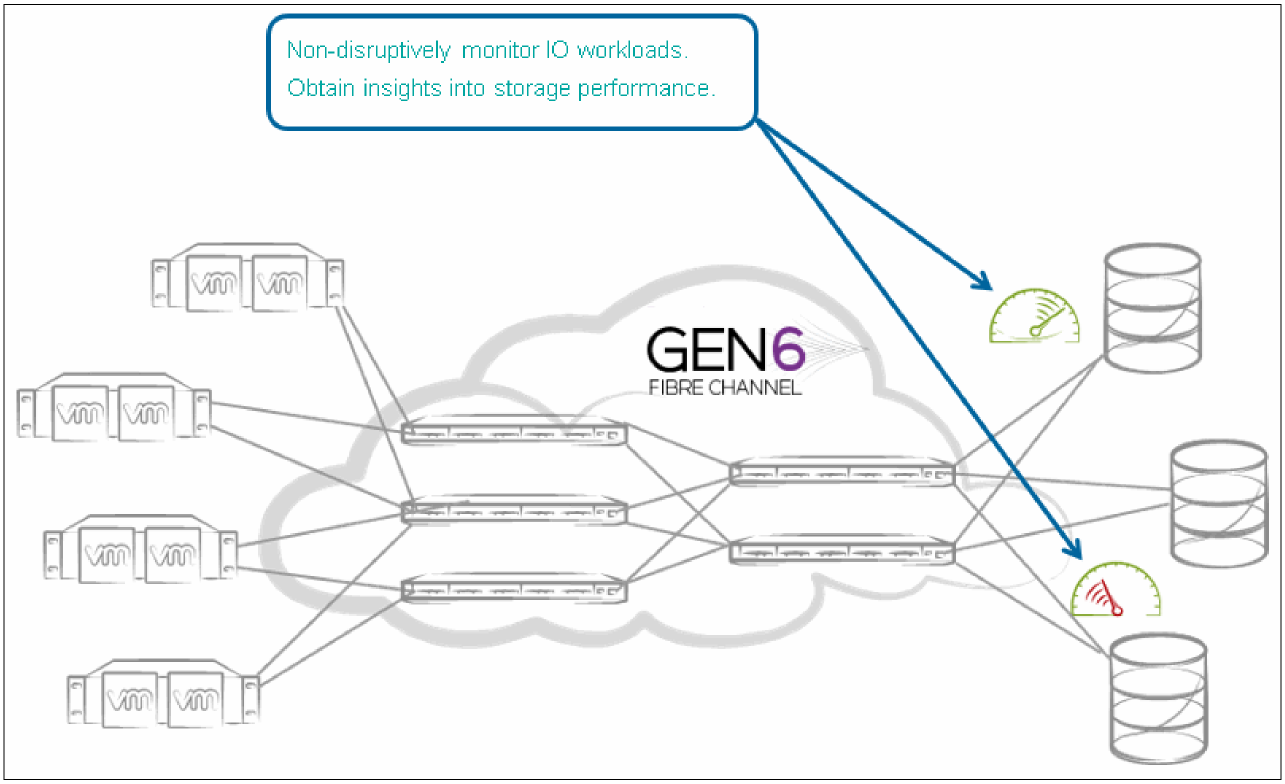


Figure 4-7 IO Insight workload monitoring

Providing critical visibility for operational stability, IO Insight complements and extends existing Fabric Vision technology and features. IO Insight uses deep visibility on storage input/output performance to ensure operational stability. IO Insight metrics include the timing of when a frame enters or leaves a switch port.

With IO Insight capabilities, the following use cases are supported:

- Storage performance SLA compliance
- Storage performance troubleshooting
- Storage performance optimization

Features highlight

These are the main features of IO Insight:

- Monitors individual host and storage devices
- Obtains total I/Os, first response time, and I/O latency
- Enables tuning of device configurations with integrated I/O metrics, to optimize storage performance

The following two metrics are usually the best indications of application performance impact:

- First Response Time
- Command Completion Time

IO Insight with Flow Vision and MAPS

IO Insight capabilities are integrated into the existing features of Fabric Vision technology. The IO Insight metrics are accessible through Flow Vision.

After IO Insight metrics are made available through a Flow Monitor flow, the flow can be imported into MAPS. After it is imported to MAPS, it can be used to configure threshold-based monitoring and alerting on the latency metrics. When latencies above the threshold are detected by MAPS, users receive MAPS notifications based on the pre-configured threshold actions. Users can then take appropriate measures to respond to the provided insights.

VM Insight

VM Insight seamlessly monitors VM performance throughout a storage fabric with standards-based, end-to-end VM tagging.

Figure 4-8 shows how VM Insight workload monitoring.

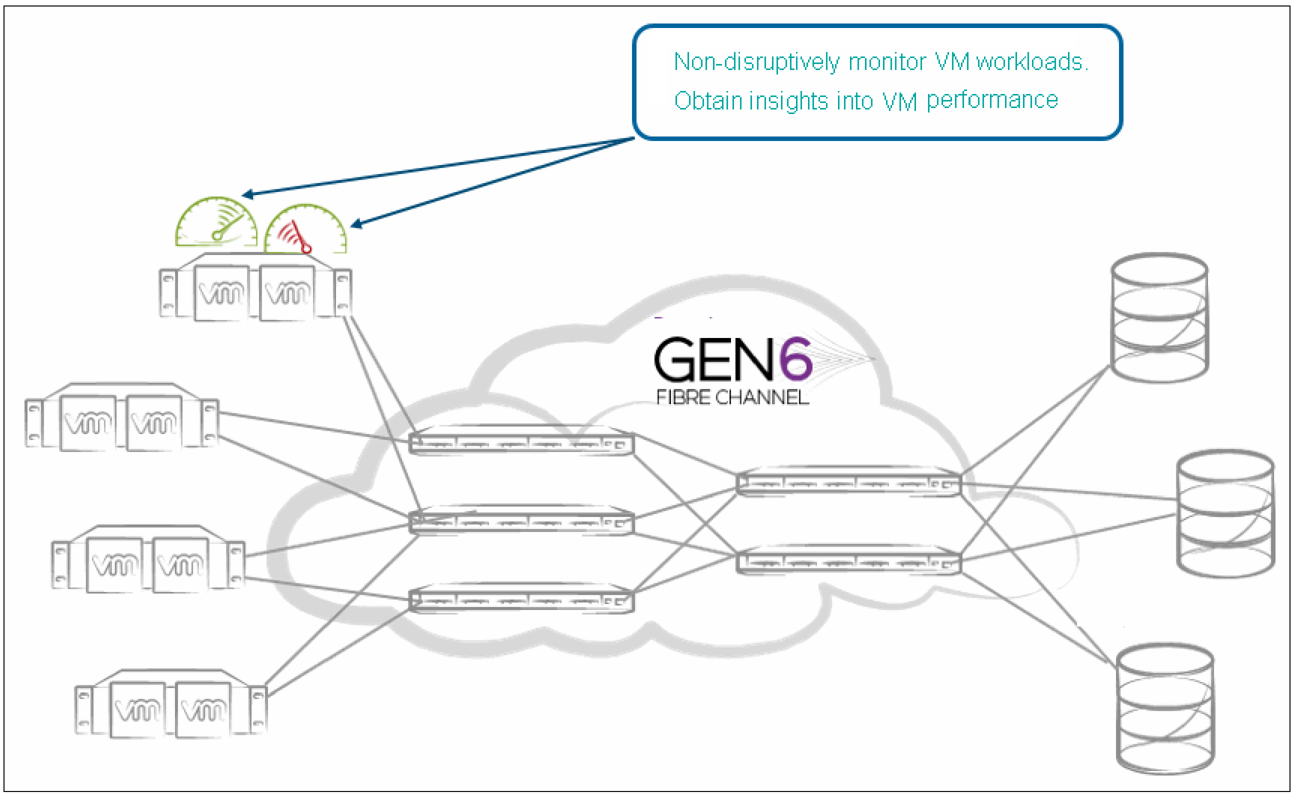


Figure 4-8 VM Insight workload monitoring

Virtualization consolidates storage access by multiple VMs on a single physical server through a single HBA to the SAN. VM Insight complements and extends existing Fabric Vision technology and features with deep visibility on VM input/output. Supported hypervisors are VMware ESX version 5.5 and later.

4.4.2 Before, during, and after the modernization effort

During the storage network modernization, Fabric Vision tools provide some helpful functionality for all modernization process phases (before, during, and after). Some helpful tools in the phases of modernization are described in this section:

- Gathering information
- Deployment
- Monitoring, troubleshooting and optimization

Gathering information

Before starting any modernization process be sure to have a picture of the actual situation. Note where the weak point are and what to do to get the best performance of the new SAN infrastructure. To help on this task the followings tools are available:

- SAN Health
- Flow Vision

SAN Health

SAN Health is a helpful utility that can give a status quo picture of the running environment. It collects all the information you need in order to understand, analyze, and design the actual SAN configuration.

Flow Vision

Flow Vision can help you analyze the actual situation of your traffic flow and bandwidth utilization. This information is helpful to develop the technical specifications for the new SAN infrastructures. Some of those specifications are noted in the following list:

- Data flow
- Workload
- Device placement

Deployment

Deployment refers to when the hardware has been installed and describes the first steps related to resiliency concepts. During the deployment phase, the following Fabric Vision tools and utilities are helpful:

- Predeployment cabling and optics validation
- Forward error correction
- Buffer credit recovery

Predeployment cabling and optics validation

ClearLink Diagnostics is a feature that can be used either as a predeployment utility or a troubleshooting test. As a predeployment utility, Cabling and Optics Validation can be used to perform complete optical, electrical, and link saturation testing to ensure reliable connections. As a troubleshooting utility, it can help to diagnose a problematic link.

Forward error correction

This function enhances transmission reliability and thus performance.
In Gen 5, this feature is on, by default, but can be switched off by the administrator.
In Gen 6, this feature is always on.

Buffer credit recovery

This function maintains application performance with buffer credit loss early detection and recovery. In Gen 5, this feature is on, by default, but can be switched off by the administrator. In Gen 6, this feature is always on.

Monitoring, troubleshooting and optimization

After the SAN infrastructure is in place, it is important to monitor the environment and take any proactive optimization actions needed. Monitoring is also fundamental for any troubleshooting analysis and performance issue. The following features are an integral part of proper maintenance monitoring, troubleshooting, and optimization:

- MAPS
- Dashboards
- COMPASS
- Fabric Performance Impact (FPI)
- Flow Vision
- IO Insight
- VM Insight

MAPS

MAPS complements existing Fabric Vision features with simple threshold-based monitoring and alerting that works with predefined policies, rules, and action. The integration with Flow Vision new Gen 6 features (such as IO Insight and VM Insight) allow latency and response-time monitoring to automatically detect impacts and alert on SLA deviations.

Dashboards

Dashboards create views that display the most relevant and critical SAN metrics to show out-of-compliance conditions, congestion, and latency. It can display real-time and historical event data, and helps quickly perform root-cause analysis.

COMPASS

COMPASS monitors the environment to ensure settings do not get changed over time ensuring consistent configuration.

Fabric Performance Impact (FPI)

The FPI simplified advanced detection feature helps the mitigation of slow-drain devices in a SAN.

Flow Vision

Flow Vision is a single tool that helps to quickly identify the source of performance issues and keep the network performing at peak. Flow Vision gains deep visibility with IO Insight to identify I/O performance degradation between host and storage. With VM Insight, Flow Vision gains unparalleled visibility into the performance of individual VMs.

IO Insight

IO Insight is an extension of Flow Vision that helps in locating IOPS and latency issues that create storage-related performance problems. Visibility into the performance of storage I/O workloads between host and storage devices allows for proactive monitoring thus ensuring consistent storage performance.

VM Insight

VM Insight is an extension of Flow Vision that gives visibility into VM-level performance metrics across the SAN. This feature quickly identifies abnormal VM behaviors and what applications are impacted.

4.4.3 Help ensure highest availability, performance, and visibility

This section describes how some of the Fabric Vision tools can be used in the modernization process. Examples include setup validation, ongoing monitoring, greater availability, highest performance, deeper visibility, optimizing queue depths, and ensuring SLAs.

This section covers the following products:

- ClearLink Diagnostics
- MAPS dashboard
- Fabric Performance Impact and Slow-Drain Device Quarantining
- Flow Vision

ClearLink Diagnostics

ClearLink Diagnostics tools allow operators to verify optics and cables.

Example 4-4 shows the result of a test between a switch D_Port and HBAs. Using a CLI, run the **`portDportTest --show all`** command to display the capabilities and test results of all the D_Ports in a switch. The result shows that port 26 failed.

Example 4-4 The portDportTest --show all command

```
switch:admin> portdporttest --show all
Port  State    SFP Capabilities   Test Result
===============================================
24    ONLINE   E,O                PASSED
26    ONLINE   E,O                FAILED
33    ONLINE   E,O                PASSED
```

However, with a different CLI command flag, having greater test detail about the failing port (26) is possible. Example 4-5 shows the **`portdporttest –show 26`** command, which gives greater insight. The examples shows an output where the electrical and optical tests pass, but the link traffic test fails. In this case, the Fibre Channel cable must be replaced and the test run again to see if the issue was resolved.

Example 4-5 The portdporttest -show 26 command

```
switch:admin> portdporttest --show 26
D-Port Information:
===================
Port:           26
Remote WWNN:    10:00:00:05:33:13:2f:b5
Remote port index:    42
Mode:           Automatic
Start time:     Wed Feb 2 01:41:43 2011
End time:       Wed Feb 2 01:43:23 2011
Status:         PASSED
================================================================================
Test                  Start time  Result       EST(secs)     Comments
================================================================================
Electrical loopback   01:42:08    PASSED        --           ----------
Optical loopback      01:42:16    PASSED        --           ----------
Link traffic test     01:43:15    FAILED        --           ----------
================================================================================
Roundtrip link latency:        1108 nano-seconds
Estimated cable distance:      20 meters
Buffers required:                       1 (for 1024 byte frames at 16Gbps speed)
Egress power:           Tx:-3.3 dBm, Rx: Not Avail
Ingress power:          Rx:-3.5 dBm, Tx: Not Avail
```

MAPS dashboard

In the MAPS dashboard, you can monitor the switch status. Using a CLI, enter the **mapsdb --show** command followed by a scope parameter: **all**, **history**, or **details**.

Example 4-6 shows a typical result of running the **mapsdb --show all** command.

Example 4-6 Result of entering mapsdb --show all

```
switch:admin> mapsdb --show all
1 Dashboard Information:
=======================
DB start time:                Mon May 11 18:38:12 2015
Active policy:                test_xy1
Configured Notifications:     RASLOG,FENCE
Fenced Ports :                None
Decommissioned Ports :        None
Fenced circuits :             38/0,38/1,38/2,38/3,38/4,38/5,38/6,38/7
Quarantined Ports :           None


2 Switch Health Report:
=======================
Current Switch Policy Status: HEALTHY


3.1 Summary Report:
===================
Category                 |Today                     |Last 7 days               |
--------------------------------------------------------------------------------
Port Health              |In operating range        |In operating range        |
BE Port Health           |No Errors                 |In operating range        |
Fru Health               |In operating range        |In operating range        |
Security Violations      |No Errors                 |No Errors                 |
Fabric State Changes     |No Errors                 |In operating range        |
Switch Resource          |In operating range        |In operating range        |
Traffic Performance      |In operating range        |In operating range        |
FCIP Health              |No Errors                 |No Errors                 |
Fabric Performance Impact|In operating range        |In operating range        |


3.2 Rules Affecting Health:
===========================
Category(Rule Count)|RepeatCount|Rule Name       |Execution Time   |Object      |Triggered Value       |
(Units)        |
-------------------------------------------------------------------------------------------------------
Switch Resource (1) |1          |defCHASSISCPU_80|03/02/15 12:10:01|Chassis     |99.00 %               |

4 History Data:
===============
Stats(Units)        Current     03/03/15    03/02/15    --/--/--    --/--/--    --/--/--    --/--/--
Port(val)   Port(val)   Port(val)
----------------------------------------------------------------------------------------------------
CRC(CRCs)           -           -           -           -           -           -           -
ITW(ITWs)           -           -           -           -           -           -           -
LOSS_SYNC(SyncLoss) -           -           -           -           -           -           -
LF                  -           -           7/13(65)    -           -           -           -
-                   -           -           7/14(1)     -           -           -           -
-                   -           -           7/15(1)     -           -           -           -
LOSS_SIGNAL(LOS)    7/4(51)     7/4(52)     7/4(44)     -           -           -           -
-                   7/5(51)     7/5(52)     7/5(44)     -           -           -           -
```

```
-                 7/6(51)     7/6(52)     7/6(44)     -           -           -           -
-                 7/7(51)     7/7(52)     7/7(44)     -           -           -           -
-                 -           -           7/14(1)     -           -           -           -
PE(Errors)        -           -           -           -           -           -           -
STATE_CHG         -           -           7/14(2)     -           -           -           -
-                 -           -           7/15(2)     -           -           -           -
-                 -           -           7/13(1)     -           -           -           -
LR                -           -           7/13(55)    -           -           -           -
-                 -           -           7/14(3)     -           -           -           -
-                 -           -           7/15(3)     -           -           -           -
C3TXTO(Timeouts)  -           -           -           -           -           -           -
RX(%)             2/11(38.46) 7/31(5.87)  2/11(7.39)  -           -           -           -
-                 7/31(37.38) -           7/31(6.82)  -           -           -           -
-                 7/0(29.11)  -           7/0(4.81)   -           -           -           -
-                 7/8(2.79)   -           -           -           -           -           -
TX(%)             2/11(38.46) 2/11(6.20)  2/11(7.68)  -           -           -           -
-                 7/31(37.39) 7/0(4.25)   7/31(6.54)  -           -           -           -
-                 7/0(29.10)  7/8(1.89)   7/0(5.00)   -           -           -           -
-                 7/8(16.74)  -           7/8(4.02)   -           -           -           -
UTIL(%)           2/11(38.46) 2/11(3.38)  2/11(7.53)  -           -           -           -
-                 7/31(37.39) 7/31(3.20)  7/31(6.68)  -           -           -           -
-                 7/0(29.11)  7/0(2.33)   7/0(4.90)   -           -           -           -
-                 7/8(9.77)   7/8(1.01)   7/8(2.21)   -           -           -           -
BN_SECS(Seconds)  -           -           -           -           -           -           -

5 History Data for Backend ports:
=================================
Stats(Units)      Current     03/03/15    03/02/15    --/--/--    --/--/--    --/--/--    --/--/--
----------------------------------------------------------------------------------------------------
CRC(CRCs)         -           -           -           -           -           -           -
ITW(ITWs)         -           -           -           -           -           -           -
LR                -           -           -           -           -           -           -
BAD_OS(Errors)    -           -           2/21(42709) -           -           -           -
-          -          2/11(41447) -           -           -           -
-          -          2/4(36958)  -           -           -           -
-          -          2/24(36175) -           -           -           -
-          -          2/18(11153) -           -           -           -
FRM_LONG(Errors)  -           -           -           -           -           -           -
FRM_TRUNC(Errors) -           -           -           -           -           -           -
```

Fabric Performance Impact and Slow-Drain Device Quarantining

In a fabric, many flows share the same link. Because of this sharing, a slow-draining device can slow down the return of credits and have a negative effect on the healthy flows through the link. To mitigate this problem, the Slow-Drain Device Quarantining (SDDQ) feature in conjunction with quality of service (QoS) monitoring allows MAPS to take action. MAPS can identify a slow-draining device and quarantine it. This feature automatically moves all traffic destined to the F_Port (that is connected to the slow-draining device) to a low-priority VC. This means the traffic in the original VC does not experience back pressure. SDDQ is enabled by having a valid Fabric Vision license. SDDQ is activated by including the SDDQ action in the configured MAPS action list. Using a CLI, enter the **`mapsconfig --actions`** command, and include **`SDDQ`** as one of the actions.

Example 4-7 shows this command and action.

Example 4-7 The mapsconfig --actions command with SDDQ as an action

```
switch:admin> mapsconfig --actions SDDQ
```

You can view a list of the quarantined ports by using the **sddquarantine --show** command. Example 4-8 shows this command and a response.

Example 4-8 View quarantined ports and response

```
switch:admin> sddquarantine --show
----------------------------------------------------------
Ports marked as Slow Drain Quarantined in the Local Switch:   2/1, 1/3
----------------------------------------------------------
Ports marked as Slow Drain Quarantined but not enforced:   1/3
----------------------------------------------------------

Online Quarantined Devices across the fabric
----------------------------------------------------------
Port Index |   PID    |             PWWN
----------------------------------------------------------
17         | 051100 | 30:10:00:05:33:ac:c6:13
17         | 051101 | 30:10:01:05:33:ac:c6:13
----------------------------------------------------------
```

You can manage other actions by using commands such as **nsshow**, **nscamshow**, and **nodefind**. These commands verify that a device is slow-draining. Example 4-9 shows the output of the **nsshow** command, showing where the slow-draining device is. The identifying line is called out in Example 4-9. If no slow-draining device is found, the line does not appear.

Example 4-9 The nsshow command with a slow-draining device

```
switch:admin> nsshow
{
Type Pid    COS     PortName                NodeName                 TTL(sec)
N    010000;    2,3;20:00:00:05:1e:92:e8:00;20:00:00:05:1e:92:e8:00; na
.   Fabric Port Name: 20:00:00:05:1e:92:e8:00
.   Permanent Port Name: 20:00:00:05:1e:92:e8:00
.   Port Index: 0
.   Share Area: No
.   Device Shared in Other AD: No
.   Redirect: No
.   Partial: No
.   LSAN: No
.   Port Properties: SIM Port
N    014a00;    2,3;30:0a:00:05:1e:84:b5:c3;10:00:00:05:1e:84:b5:c3; na
.   FC4s: FCP
.   NodeSymb: [42] "dsim:fdmi_host:sw_5300edsim[172.26.26.188]"
.   Fabric Port Name: 20:4a:00:05:1e:92:e8:00
.   Permanent Port Name: 30:0a:00:05:1e:84:b5:c3
.   Port Index: 74
.   Share Area: No
.   Device Shared in Other AD: No
.   Redirect: No
.   Partial: No
.   LSAN: No
N    015000;      3;10:00:00:00:00:01:00:01;10:00:00:00:00:00:01:01; na
.   Fabric Port Name: 20:50:00:05:1e:92:e8:00
.   Permanent Port Name: 10:00:00:00:00:01:00:01
.   Port Index: 80
.   Share Area: No
```

```
.   Device Shared in Other AD: No
.   Redirect: No
.   Partial: No
.   LSAN: No
.   Slow Drain Device: Yes     <== Slow-draining device identified
N    015100;      3;10:00:00:00:00:02:00:01;10:00:00:00:00:00:02:01; na
.   Fabric Port Name: 20:51:00:05:1e:92:e8:00
.   Permanent Port Name: 10:00:00:00:00:02:00:01
.   Port Index: 81
.   Share Area: No
.   Device Shared in Other AD: No
.   Redirect: No
.   Partial: No
.   LSAN: No
The Local Name Server has 4 entries }
```

Flow Vision

This section shows how this feature can be used and what information is provided. Some examples of information provided are noted in the following list:

- Traffic statistic reporting with size classification of I/O performance and latency counters
- Read and write first response time
- Read and write status time
- Read and write pending I/O

At the end of this section, several case studies describe how to use the data provided from the Flow Vision tool.

The first set of statistics describes the output of the standard Flow Monitor counters. Specifically, this output includes the FC frames and SCSI throughput (Example 4-10).

Example 4-10   FC frames and SCSI throughput metrics

```
=========================================================================================================
Name      : iostest   Features: mon(Activated)   noConfig: Off
Definition: EgrPort(12),ScrrDev(0x020d00),DstDev(0x020c00)
Flow Monitor (Activated):
Monitor time: |Mon Feb 22 16:26:39 UTC 2016|
---------------------------------------------------------------------------------------------------------
|Tx Frames Count     |Tx Frames per Sec.   |Tx Bytes Count  |Tx Throughput(Bps) |Avg TX Frm S(Bytes) |
---------------------------------------------------------------------------------------------------------
1.22G               |760.92k              |2.34T           |1.48G              |2100                |
---------------------------------------------------------------------------------------------------------
I/O count                |I/O Per Sec. (IOPS)  |I/O bytes Transferred     |I/O bytes Per Sec.      |
Reads/Write/Total        |Reads/Write/Total    |Reads/Write/Total         |Reads/Write/Total       |
---------------------------------------------------------------------------------------------------------
64M/ 9.65M/ 19.30M       |5.99k/ 5.99k/ 11.98k |2.29T/ 2.29T/ 4.59T       |1.46G/ 1.46G/ 2.92G     |
---------------------------------------------------------------------------------------------------------
```

The next set of statistics describes the I/O Latency Metrics, specifically Status Time. This is the time between the read or write command by the initiator and the status frame sent by the target (Example 4-11).

Example 4-11 Status Time metrics

```
|I/O Latency Metrics:                                                                                   |
---------------------------------------------------------------------------------------------------------
|Metric                        | IO SIze              | Max                      | AVG                      |
|                              |                      | 6 sec   /   All          | 6sec   / All             |
---------------------------------------------------------------------------------------------------------
|RD CMD -> Status Time         | <8K                  |         /                |        /                 |
|                              | 8K - < 64K           |         / 3.96m          |        / 1.17m           |
|                              | 64K - <512K          | 23.59m  / 27.54m         | 5.09m  / 5.02m           |
|                              | >=512K               |         / 39.86m         |        / 17.33m          |
|                              | ALL                  | 23.59m  / 39.86m         | 5.09m  / 5.02m           |
---------------------------------------------------------------------------------------------------------
|WR CMD -> Status time         | <8K                  |         / 4.50m          |        / 226u            |
|                              | 8K - < 64K           |         / 5.31m          |        / 1.10m           |
|                              | 64K - <512K          | 18.57m  / 35.37m         | 4.47m  / 4.43m           |
|                              | >=512K               |         / 70.16m         |        / 18.70m          |
|                              | ALL                  | 18.57m  / 70.16m         | 4.47m  / 4.42m           |
---------------------------------------------------------------------------------------------------------
```

The next set of statistics describes a continuation of the I/O Latency Metrics, specifically First Response. This is the time duration between the read or write command frame and the first response (Example 4-12).

Example 4-12 First Response metrics

```
|I/O Latency Metrics:                                                                                   |
---------------------------------------------------------------------------------------------------------
|Metric                        | IO SIze              | Max                      | AVG                      |
|                              |                      | 6 sec   /   All          | 6sec   / All             |
---------------------------------------------------------------------------------------------------------
|RD CMD -> 1st Time            | <8K                  |         /                |        /                 |
|                              | 8K - < 64K           |         / 218u           |        / 105u            |
|                              | 64K - <512K          | 3.88m   / 5.20m          | 124u   / 123u            |
|                              | >=512K               |         / 222u           |        / 111u            |
|                              | ALL                  | 3.88m   / 5.20m          | 124u   / 123u            |
---------------------------------------------------------------------------------------------------------
|WR CMD -> 1st XFER_RDY time   | <8K                  |         / 4.38m          |        / 24u             |
|                              | 8K - < 64K           |         / 660u           |        / 25u             |
|                              | 64K - <512K          | 3.86m   / 5.19m          | 38u    / 37u             |
|                              | >=512K               |         / 109u           |        / 33u             |
|                              | ALL                  | 3.86m   / 5.19m          | 38u    / 37u             |
---------------------------------------------------------------------------------------------------------
```

The last set of statistics also describes a continuation of I/O Latency Metrics, specifically Pending I/Os. This is the average number of outstanding I/O operations yet to be completed. This value can help to optimize the queue depth (Example 4-13).

Example 4-13 Pending I/Os

```
|I/O Latency Metrics:                                                                          |
------------------------------------------------------------------------------------------------
|Metric                      | IO SIze             | Max                    | AVG                   |
|                            |                     | 6 sec   /   All        | 6sec   / All          |
------------------------------------------------------------------------------------------------
|RD Pending IOs              | <8K                 | 0       / 0            | 0      / 0            |
|                            | 8K - < 64K          | 0       / 1            | 0      / 2            |
|                            | 64K - <512K         | 64      / 64           | 31     / 61           |
|                            | >=512K              | 0       / 1            | 0      / 2            |
|                            | ALL                 | 64      / 64           | 31     / 61           |
------------------------------------------------------------------------------------------------
|WR Pending IOs              | <8K                 | 0       / 1            | 0      / 2            |
|                            | 8K - < 64K          | 0       / 1            | 0      / 2            |
|                            | 64K - <512K         | 64      / 64           | 27     / 54           |
|                            | >=512K              | 0       / 1            | 0      / 2            |
|                            | ALL                 | 64      / 67           | 27     / 60           |
------------------------------------------------------------------------------------------------
```

Case studies

The following case studies provide an easy way to read data that is from Flow Monitor so that you can start SAN troubleshooting and ports traffic optimization.

- Case1: Delay in transmitting the frame.

  The following information is gathered from the I/O Latency Metrics:

  – Status Time shows a high value
  – First Response shows an acceptable value

  These results show that the target is responding fast but there is some delay to transmit the frame. A possible reason for this might be related to the initiator or cable/optics. To verify if the factor is related to the initiator, check the I/O Latency Metrics: Pending I/O. If the Pending I/Os show a high value, then the issue is with the initiator and a resolution might be to optimize the queue depth parameter.

- Case 2: Target is responding too slow.

  The following information is gathered from the I/O Latency Metrics:

  – Status Time shows a high value
  – First Response shows a high value

  These results show that the target is responding too slowly. The reason might be target related or related to the cable/optics. In this case, deep investigation is needed. The first step is to start checking optics and cables, then, look at the target.

4.5 Preferred practices, high level steps, and products for modernizing storage networks

Modernizing your data center components is one of the most important steps you can take to ensure your infrastructure keeps up with your evolving business needs. Modernizing components happens relatively infrequently due to the cost and labor involved. Therefore, a critically important task is to prepare and plan for these relatively large changes, and to follow general preferred practices during migrations at either hardware or software levels.

SAN infrastructure can be added to or replaced depending on expansion or growth requirements. Multiple types of fabric topologies exist, but the most common is the dual-SAN, which is two separate physical SAN networks that provide fabric redundancy in case of failure on one of them. These can be a single physical switch for each fabric, or multiple physical switches connected by ICLs or ISLs on one fabric, with (usually) an identical physical switch topology for the redundant fabric.

All of these SAN migration methods should be planned meticulously in advance with walkthroughs of the *written* plan. Also, create contingencies for backing out in case of unexpected problems. Throwing together a plan with no feedback from other parties, failing to rehearse the plan, and attempting to do too much during a limited outage window are all typical ways to experience calamity.

A preferred practice for adding switches to a fabric is to ensure the new fabric is running a compatible operating system with the current fabric, and, further, that the same compatible operating system does not have any zones, aliases, or configurations defined. This practice allows the new switch to download the configuration from the existing fabric and apply it. Any new additions of worldwide names (WWNs) to the new switch are automatically detected by the whole fabric.

Note: Consider that ISLs (and to a lesser extent, ICLs) add latency to any inter-switch hops between WWNs that are zoned together but on separate physical switches. Seek out guides for the specific storage layout and how it should be physically connected and zoned (for instance, SAN Volume Controller inter-node communication).

You can also upgrade the optical transceivers on a switch to a faster speed (if supported by that physical switch). This is a relatively simple upgrade, but you should ensure that the new transceiver and Fibre Channel speeds are supported by your switch, HBA, HBA driver, and operating system. In some cases, upgraded cables (OM2 replaced by OM4 for instance) are also required.

To fully replace a physical SAN topology, a common practice is to perform *single leg* migrations. This works best for dual-SAN environments that require high uptime. Also, one can pre-stage zoning and cabling for the new physical switches to the existing hosts and storage.

Ensure that each host’s redundant connections are active and that path failover software is installed and working correctly:

1. Single connected hosts will require a short storage outage. During the first outage window, replace all cabling to the first, old fabric by connecting those hosts and storage to the new physical fabric.
2. After ensuring stability and working order on the new, first physical fabric, take a second outage window (often the next night or next week) to move each host’s and storage second connections to the second, new physical fabric.

3. After stability is ensured on both new fabrics, check the old fabrics for any straggling or unexpected connected hosts or storage. If you planned correctly, you should not find any stragglers. Sometimes, even with planning, the unexpected happens and contingency planning can go into effect.
4. After, all hosts are on the new fabrics, decommission the old physical fabrics.

New, b-type Gen 6 fabrics can help when existing fabrics experience congestion at legacy speeds. Gen 6 fabrics also support new technologies such as NVMe over Fibre Channel. Much of the rest of this publication covers the benefits of b-type Gen 6 so it is not exhaustively covered in this section.

Other migration methods include application level, physical server based, virtual machine, and storage virtualization.

Application level migration often consists of moving data from one version of application software to a higher version on different hardware. It can also consist of only upgrading the application on current hardware. An application level migration can include movement from cloud to cloud, or data center to data center. In those cases, network bandwidth is of utmost concern. You should also meticulously follow the application's instructions for upgrade, and check afterward for any discrepancies or bugs uncovered by the new application.

Physical server migrations are common when the current servers' physical capabilities can no longer be upgraded and are aged or outgrown by increasing storage, server, or network requirements. In today's data center, migrating a physical server to a virtual server is generally a good idea, except in cases where a physical server is required. Businesses can either migrate to a virtual server on a public, private, or hybrid cloud, or to new, dedicated physical infrastructure in their data center. Both options necessitate detailed requirements for the new servers (cloud or physical). The positive side of such migrations is that they are generally less complex than application migrations. The negative side is that a lot of human action is involved to ensure the physical requirements are met and compatible with the current platform.

Virtual machine migrations, for instance vMotion, are often relatively simple. However, they require much hardware and software setup prior to becoming viable. They also require much network and storage network bandwidth due to the large size of the virtual machines being moved. Ensure you have the bandwidth and failover capabilities at the virtual machine level to accommodate these moves.

Storage virtualization is a component of software-defined storage. Storage virtualization, which is a relatively easy and useful method of moving back-end storage to other storage, offers advantages:

- It allows hosts to continually access data while being moved from one hardware platform to another on the back end.
- It reduces the number of outages required, and gives you much flexibility in moving data from tier to tier to match performance requirements with the correct tier of storage.

Back end arrays may be taken out of service due to hardware issues, or aged out and replaced with newer arrays with no host outage taken. Software-defined storage often requires more storage network bandwidth due to these moves. The moves can often be throttled so as not to overwhelm the storage network's capacity.

IBM Spectrum Virtualize is the IBM storage virtualization product, and is included with SAN Volume Controller, FlashSystem v9000, and all Storwize arrays. It can be combined with IBM Spectrum Control and IBM Network Advisor for the ideal combination of software-defined storage and storage array, and storage network monitoring and insights.

IBM FlashSystem 900 AE3 array is the latest version of the IBM FlashSystem portfolio. It includes the FlashCore and MicroLatency technologies that make the FlashSystem 900 the best performing Flash in the industry. The AE3 also includes compression on all data to increase the capacity of the array on the same hardware platform. This results in 220 TB of usable flash capacity on a 2U array. The AE2 is the previous model, which has smaller flash modules and no compression. The maximum capacity of the AE2 is 57 TB.

4.6 Providing guidance for modernizing your storage network

Choosing to modernize a storage network is not a question that has a yes or no answer. It is a matter of when it must happen. Modernization in some form is required. The questions are more about when to take the step, where to make the changes, and then how best to take that step.

The guidance covered in this section is designed to flow in the following realistic order:

4.6.1 Deciding when to upgrade SAN

The digital age and market demands require service and performance. To meet these requirements, providers must have modern infrastructures and data centers. Everything that is discussed today in technological worlds (cloud, analytics, social platforms, and big data) are increasingly for more advanced systems, platform growth, and increased data storage.

Today, Fibre Channel is the standard for storage connectivity, especially with the advances in all-flash based storage. The movement to flash environments is changing the design of the data center. All infrastructure is being reassessed and pushed to the limit and legacy environments will have performance problems. The new generation of SAN b-type Gen 6 was developed to meet the all-flash-based demands and help provide high-performance, high-availability, and reliability.

The sooner businesses modernize their infrastructures, the faster the return on investment will be. The environment will be more up-to-date, easier to manage with fewer problems, and therefore provide more customer satisfaction.

4.6.2 When and where to add new switches in regards to legacy equipment

The addition of new SAN switches in a fabric network is based on the business needs. Almost always, when adding new SAN switches at fabric network, legacy equipment will exist. Exceptions include businesses with unusual budgets or superior and well-defined refresh and replace plans, or if they are creating an environment from scratch. Adding new SAN switches in a legacy environment is a normal situation but requires study and a plan to know how and where to install. The SAN administrators and architects must be up-to-date with the new technologies when planning future deployments. Strategy is important when mixing legacy and new SAN devices; replacing the legacy devices is a fact. Analyze the plan and consider stressors; an important task is to ensure that the transition is mitigated for impact and that future deployments are simpler.

4.6.3 Modernize without disruption

Modernizing without disruption to the critical services and applications on the existing SAN requires extensive planning to ensure near-perfect execution. Two approaches are available to migrate the devices on the existing SAN to a new Gen 6 switch architecture:

- Keep new SAN switches separate from the existing SAN.
- Merge new switches into existing SAN through the use of ISLs.

Keep separate fabrics

The first approach has the main advantage of protecting the existing fabrics from disruption, because the new fabric is kept separate. However, because the new fabric is separate, it requires more planning and coordination of moving devices to the new SAN. It also requires more configuration, such as zoning, be done manually on the new fabric. All of this should be done before devices are migrated.

The physical migration of devices is nearly always a multi-step, multi-day (preferred multi-week) process. During the migration, devices such as IBM SAN Volume Controller, the back-end storage, and critical hosts, will serve as bridges between the old and new fabrics. After all connections of each SAN device are migrated, the legacy environment is ready to be deactivated.

The configuration, planning and migration must then be repeated on any redundant fabrics. Do not do concurrent migrations in the event that the migration of devices causes problems.

Merge new switches into existing fabrics

The second approach is to connect the new switches into the existing SAN using ISLs. This approach allows for some of the configuration, such as zoning, to be merged automatically. However, some risk of disruption to the existing fabric exists. This design creates a dual-core or, if adding switches to existing dual-core, a multi-core design.

Moving devices from the existing switches to the new switches can happen at a slower pace. However, be careful when minimizing the traffic that flows across the ISLs between existing and new switches. Approaching the process with this method means the following procedures are adhered to:

- Storage, IBM SAN Volume Controller (if any) and critical hosts should all be migrated at the same time.
- Any edge switches connected to the existing switches in a core-edge design should also be connected to the new switches. Otherwise any hosts or devices attached to those switches will now have to traverse multiple hops to reach their storage ports.

One final consideration on which approach is best for your environment is to consider the port speeds on your existing SAN. The preferred guideline is to be sure that the fastest port speed does not exceed 2x the slowest. If this guideline is not adhered too, the slowest devices will eventually become slow-drain devices. The faster the fastest device is (and the greater the differential) the more quickly that happens.

For example, if new switches are16 Gbps capable and there is 16 Gbps capable devices in the existing SAN, then migrating any 4 Gb or slower devices is not a good approach. So, by using the formula noted previously (fastest port speed should not exceed 2x the slowest), in this case, 16 > 2(4). In this case, the better option is to maintain separate fabrics until the slow devices can either be upgraded or removed from the SAN.

4.6.4 How to ensure low latency apps and flash are in optimum position

An all-flash environment by itself means high performance and low latency; however, even with that advantage, monitoring apps and performance is still beneficial. Several available tools can help monitor performance at the fabric network and are noted in 4.2.5, “Monitoring, troubleshooting and optimization” on page 59.

Specifically considering applications and flash, the IO Insight tool has several useful features. One of the ways to ensure an application is meeting satisfactory performance values is the use of a new generation SAN b-type Gen 6 with Fabric Vision IO Insight. The Gen 6 allows Fabric Vision to increase its capabilities with the introduction of the Brocade IO Insight. It has integrated network sensors that provide deep visibility into I/O performance in the storage structure. This improvement means you can quickly identify performance degradation in the application, helping to decrease the time for resolution.

After correctly setting up SAN devices in the environment according to specification. businesses can use Fabric Vision IO Insight to proactively monitor performance and behavior of the environment. This product also informs administrators of potential problems. The Fabric Vision IO Insight integrated with Brocade Network Advisor provides performance information and ensures that the flash environment is in optimum position.

4.6.5 SAN assessment using various tools

The following two tools provide help when doing assessment of the SAN environment:

- Brocade SAN Health
- “Brocade Network Advisor”

Brocade SAN Health

Brocade SAN Health is a no-cost software tool that provides an easy view of the fabric network topology. It offers a report in Microsoft Excel format that shows information about switches, directors, servers, HBA's, firmware version, ports, and so on. Brocade SAN Health main focus is to optimize data collection about the SAN.

Brocade SAN Health performs tasks such as these:

- Creates an inventory of devices
- Captures and displays port utilization, ISL/trunk details, and bandwidth utilization statistics
- Shows both zoning and config zones
- Provides performance statistics and frame error counts
- Provides fabric topology in Microsoft Visio format

Figure 4-9 shows an example of the Brocade SAN Health spreadsheet.

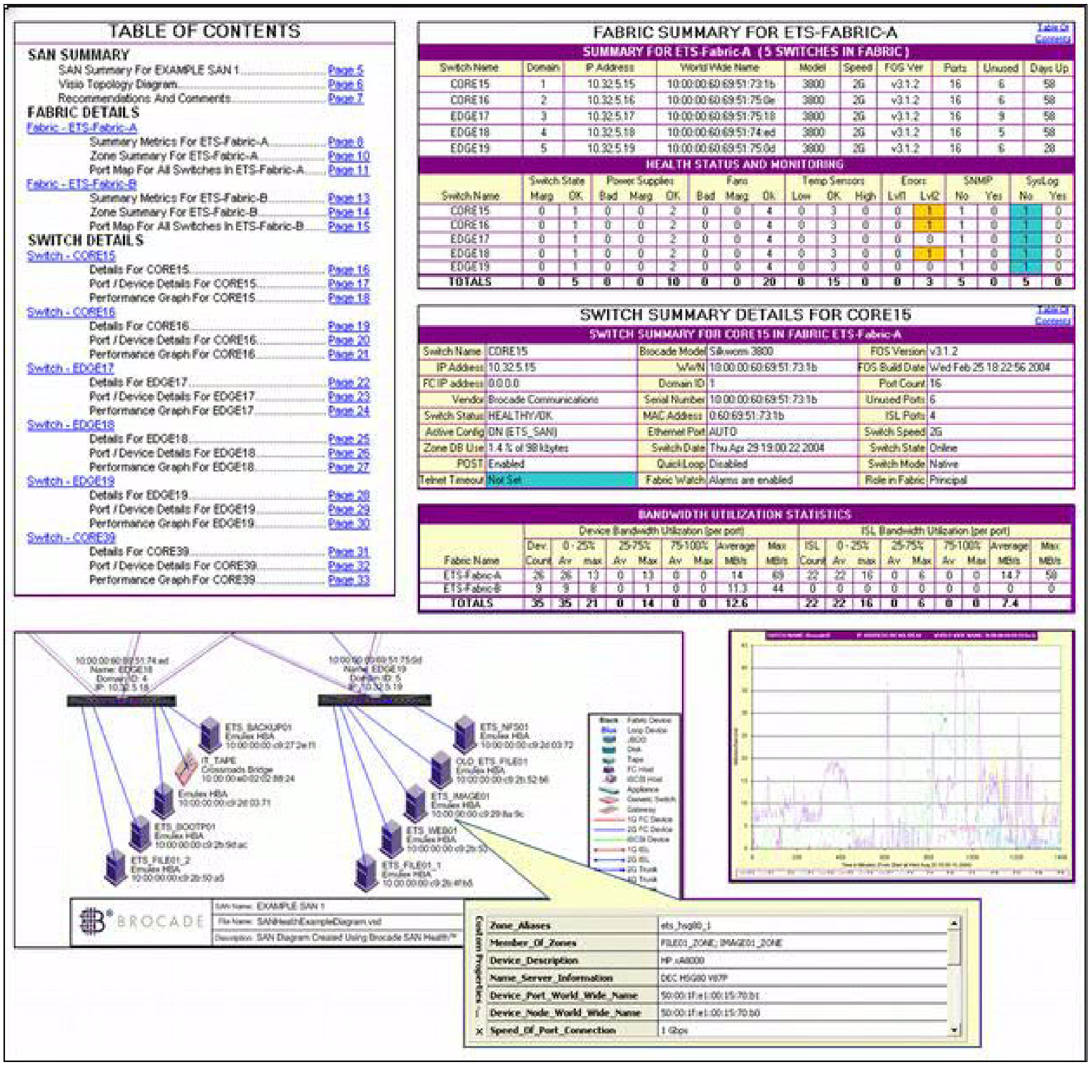


Figure 4-9 Brocade SAN Health spreadsheet

To implement this tool, see Brocade SAN Health web page.

Brocade Network Advisor

Brocade Network Advisor is a software tool that simplifies daily operations of the overall SAN. It offers dashboards that provide easy fabric network understanding, stores data to have historical network visibility, and provides alert notifications to help proactively manage the SAN network. Brocade Network Advisor integrates with storage data center automation solutions to provide end-to-end network visibility.

This visibility is provided through frameworks such as the Storage Management Initiative-Specification (SMI-S). It reduces the number of tools needed for end-to-end network visibility, management, and control. This is done by unifying management of SAN devices and integration with third-party solutions.

Brocade Network Advisor provides capabilities such as these:

- Customizable dashboards
- Performance reports
- Trend analysis
- Event management
- Call home (automatically collects data, sends notifications to support for diagnostics)
- Role-based-access-control (RBAC)

Brocade Network Advisor enables quick troubleshooting with multiple dashboard views. It provides information about abnormal behaviors that can be verified instantly so that decisions can be made. Brocade Network Advisor integrates with Brocade Fabric Vision to provide visibility and insight across storage networks. Brocade Network Advisor supports the following Brocade Fabric Vision features:

- IO Insight
- VM Insight
- Monitoring and Alerting Policy Suite (MAPS)
- ClearLink Diagnostics
- Flow Vision
- Fabric Performance Impact (FPI) monitoring
- Dashboards
- Configuration and Operational Monitoring Policy Automation Services Suite (COMPASS)
- Forward error correction (FEC)
- Credit loss recovery
- Fabric Insight Portal

Figure 4-10 shows the Brocade Network Advisor interface.

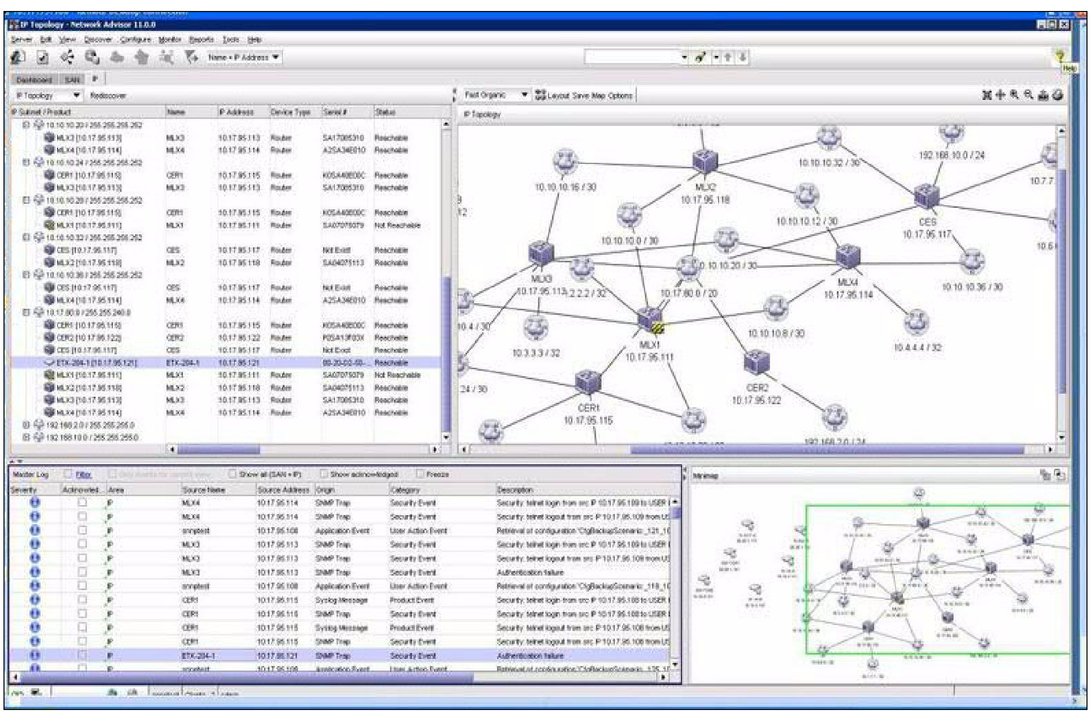


Figure 4-10 Brocade Network Advisor interface

For more information, see the Brocade Network Advisor web page.

4.6.6 Troubleshooting

Troubleshooting in a SAN environment is not simple or obvious. It has several variables such as storage, switch, director, servers, cables, and each one with its proper characteristics such as firmware version.

There is no rule to be followed. Troubleshooting uses personal expertise to detect the issue, environmental knowledge, and tools to configure and monitor the SAN environment. The following areas can be investigated during troubleshooting:

- Storage hardware failure
- Server hardware failure
- SAN switch and director hardware failure
- SAN connectivity hardware failures (SFP/cable)
- Interoperability matrix issues
- ISL bottlenecks
- Zoning configuration issues
- Host configuration issues
- Capacity limits issues
- Application issues

Teamwork is also an indispensable aspect of troubleshooting. Whether colleagues in the environment contribute or you open a case with product support, start a process with all teams involved to resolve issues with greatest speed.

4.6.7 Tools to use to streamline the process

The information in this section might help speed the troubleshooting process.

IBM Service Request

Use the IBM Service Request website to open calls for equipment under your service coverage. Service Request coverage can include hardware or software. To access the Service Request website, log in with your IBMid or create an IBMid at that website.

A second option is to access the IBM Support web page. On that page, click on a category to open a ticket.

IBM Fix Central

Fix Central is an IBM website that provides fixes and updates for system software, hardware, and operating systems.

See the Fix Central web page.

5

Modernizing IBM Z storage networking

Brocade Communications Systems, LLC, and IBM released a series of features to address the IBM Z mainframe storage networking demands of the data center. These technologies leverage the capabilities of Gen 6 Fibre Channel, and extend them to the applications that drive the world's most critical systems. This chapter summarizes these unique mainframe features, which reflect the dedication to technical excellence that both companies provide.

The following topics are covered in this chapter:

5.1 Gen 6 technology with IBM Z

The partnership between IBM and Brocade Communications Systems, LLC, allows both organizations to guide the introduction of key technologies to the market place, while the integration between the system test and qualification teams ensures the integrity of those products. Driving these efforts are the deep technical relationships between the Brocade and IBM Z I/O architecture and development teams. As a reflection of this close relationship between the two companies, IBM partnered with Brocade to release the IBM b-type products. The teams close collaboration provides the transformative technology to propel today's ever changing data center.

During the past seven years, the technology of mainframe storage networking systems has evolved to embrace new levels of scale. Business continuity and recovery requirements have stretched the limits of yesterday's technology, and the cloud movement is forcing organizations to reexamine transactional processing. These pressures led IBM to produce one of the most innovative and highly functioning processors in history—the z14. At the same time, IBM b-type products provided the storage networking necessary to harness the power of the z14.

The technological leap forward by the z14 is complemented by the inventiveness and integration of the IBM b-type Gen 6 feature set in the following ways:

- Inter-chassis links (ICLs) provide world-class scalability and flexible chassis interconnectivity.
- Multi-hop topologies provide flexibility to geographically dispersed data centers to support unique disaster recovery and business continuity configurations.
- The 16 Gb FC (GFC) FICON Express16S+ device connectivity and 32 GFC fabric connectivity provide the fastest data rates in the industry.
- FICON Dynamic Routing uses the full capacity of inter-switch link (ISL) and ICL connections in cascaded network topologies, leveraging Brocade Exchange-based Routing (EBR) technology.
- Fabric I/O Priority capabilities provide a foundation for applications, such as IBM z/OS® Workload Manager, to prioritize individual I/Os for optimal storage network performance.
- Port Decommission is used for easier infrastructure management through automated interaction between the fabric and the mainframe I/O subsystem.
- IBM Health Checker for z/OS leverages the internal control unit port (CUP) of the FICON director to evaluate single-point-of-failure conditions as well as path integrity and performance.
- Problem determination and isolation are enhanced through the Read Diagnostic Parameters (RDP) operations.
- Forward error correction (FEC) improves connection reliability and performance.

5.1.1 Gen 6 highlights

IBM b-type mainframe SAN technology was first to market with several innovations:

- Gen 6 32 Gbps Fibre Channel
- Gen 5 16 Gbps Fibre Channel with forward error correction
- FICON Dynamic Routing
- z/OS Health Checker integration
- Read Diagnostic Parameters functionality

Leveraging the IBM Z I/O team partnership, Brocade has also produced significant mainframe storage networking market exclusives such as ICLs and port decommissioning. This shared dedication to the success and growth of the mainframe storage network can also be seen in newer endeavors, including Fabric I/O Priority and FICON multi-hop topologies. These technologies are detailed in the following sections to provide insight into IBM b-type Gen 6 mainframe storage networking functionality:

- 5.1.2, “Speed, scale, and performance” on page 93
  - “Multi-hop topologies”
  - “FICON interface connectivity 16Gbps and 32Gbps” on page 95
  - “FICON Dynamic Routing” on page 96
  - “Fabric I/O Priority” on page 97
- 5.1.3, “Autonomics and Integrity” on page 97
  - “Port decommission” on page 98
  - “IBM Health Checker for z/OS Integration” on page 99
  - “Enhanced diagnostics” on page 100
  - “Forward error correction (FEC)” on page 101

5.1.2 Speed, scale, and performance

Inter-chassis links (ICLs) provide short-distance, chassis-to-chassis connectivity for Gen 5 and Gen 6 FICON directors for use with FICON and/or FCP mainframe solutions (Figure 5-1). This technology is used to build a powerful storage networking core without sacrificing device ports for inter-switch link (ISL) connectivity. ICLs minimize the latency between chassis to provide the lowest-latency switching using a backplane relative to using comparable ISLs.

Frame-based trunking is automatically enabled between four ICLs, which maximizes the load balancing and availability characteristics of the chassis. When used with FICON Dynamic Routing, the I/O exchanges are evenly distributed across the ICL infrastructure for optimal performance. Furthermore, as the ICL connection is not considered to be a *hop of concern* in cascaded topologies, it can provide configuration flexibility for large deployments.

Figure 5-1 shows the difference between ICLs and ISLs.

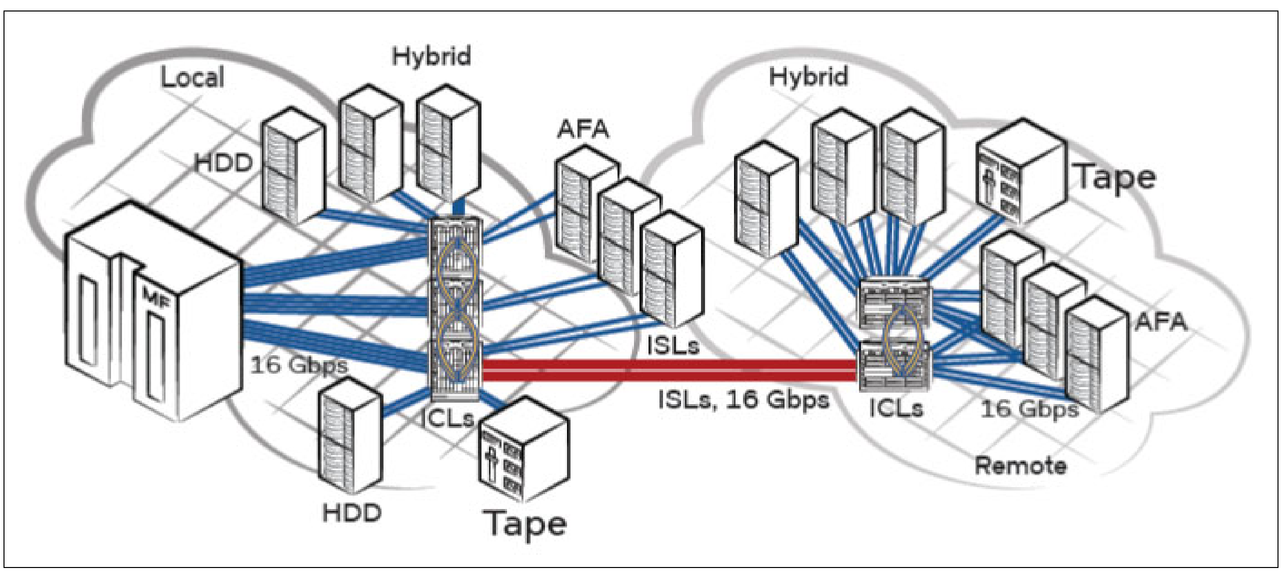


Figure 5-1 Link comparison

Multi-hop topologies

Multi-hop topologies are the evolution of classical FICON SAN topologies in response to the movement to disperse mainframe data centers geographically. Simple cascaded systems have given way to multi-hop variations that maximize Brocade ICL and FCIP channel extension technologies, while maintaining the integrity of the *hop of concern* provision. Advanced diagnostic capabilities, such as the IBM Health Checker for z/OS and the Fibre Channel Read Diagnostic Parameters function, maintain the reliability required for mainframe storage networks in extended topologies.

Multi-hop topologies include two-, three-, and four-site disaster recovery and business continuity configurations (Figure 2). Each topology is designed to address specific business requirements and regulatory provisions. In addition, the supported topologies operate with the same level of predictability and reliability found in today's cascaded solutions.

Figure 5-2 shows an example of multi-hop topology.

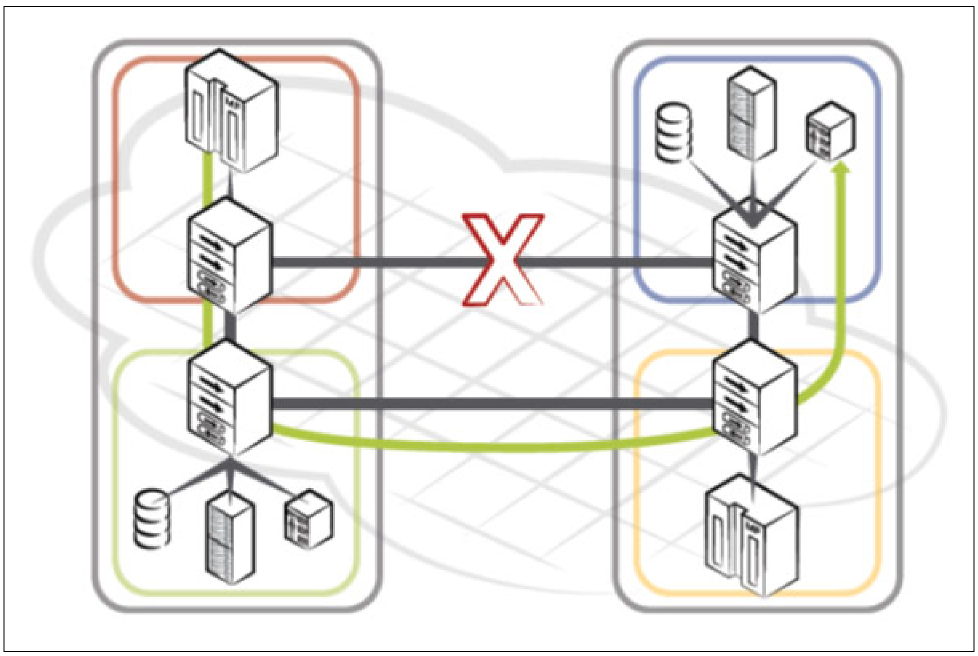


Figure 5-2 Multi-hop topology

FICON interface connectivity 16Gbps and 32Gbps

IBM b-type Gen 6 FICON directors provide optimal interface attachment rates for the FICON Express16S+ (FX16S+) channel and also storage arrays. The 16 Gbps Fibre Channel device connectivity and 32 Gbps Fibre Channel fabric connectivity provide minimal storage network latency, which enhances data center application operations and performance.

Figure 5-3 shows the FICON Express16S+ channel card.

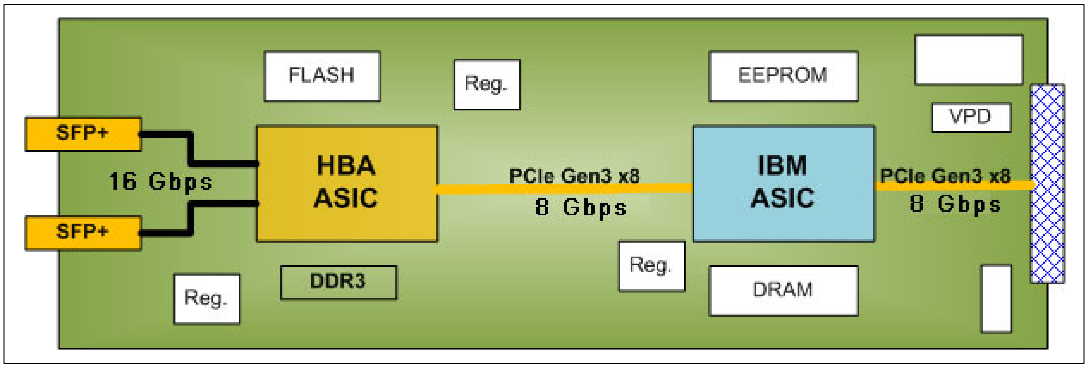


Figure 5-3 FICON channel card

The faster link speeds provide reduced I/O latency for critical middleware I/O, such as database log writing, which significantly improves DB2 transactional latency. The faster link speed also shrinks batch windows and reduces the elapsed time for I/O-bound batch jobs. Because faster link speeds are more sensitive to the quality of the cabling infrastructure, new standards are included to provide enhanced error correction on optical connections, ensuring a smooth transition to 16 Gbps and 32 Gbps Fibre Channel technologies.

FICON Dynamic Routing

FICON Dynamic Routing allows FICON devices to utilize both static and dynamic SAN routing policies. It is designed to support the dynamic routing policies provided by the FICON director, such as the Brocade Exchange-Based Routing (EBR) protocol. FICON Dynamic Routing enables organizations to use SAN dynamic routing policies across cascaded FICON directors to simplify configuration and capacity planning, and to provide persistent and repeatable performance and higher resiliency (see Figure 5-4).

In Peer-to-Peer Remote Copy configurations, the sharing of switches is simplified, and hardware costs can be reduced by allowing FICON and FCP to share the same storage network infrastructure. The introduction of the z13 server included several performance optimizations designed to improve overall throughput in the mainframe storage network.

The benefits of FICON Dynamic Routing are highlighted in FICON Dynamic Routing (FIDR): Technology and Performance Implications, a technical white paper that is co-authored by IBM and Brocade.

Figure 5-4 shows FICON Dynamic Routing.

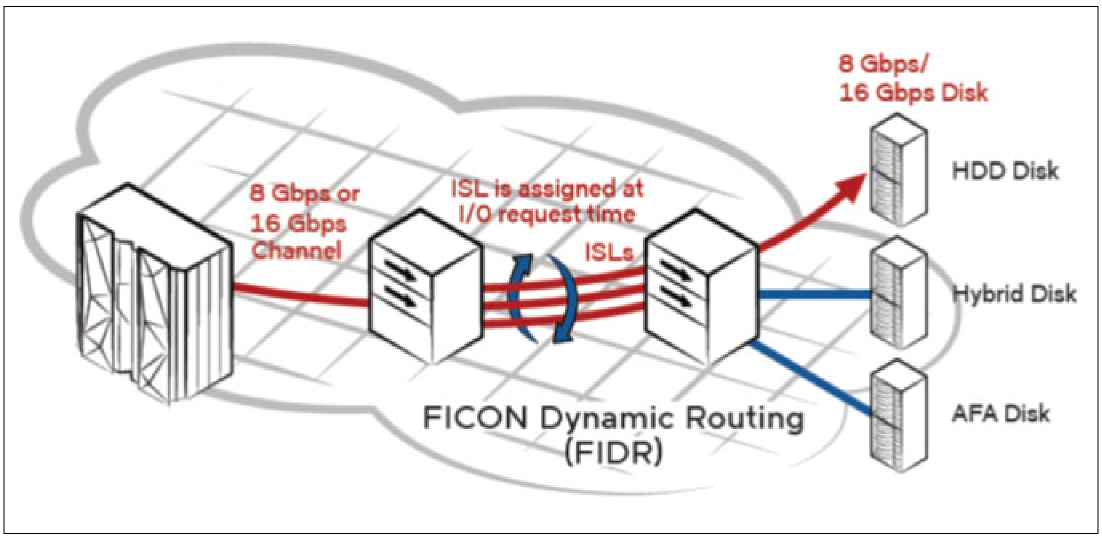


Figure 5-4 FICON Dynamic Routing

Fabric I/O Priority

SAN Fabric I/O Priority integration translates application importance into Fabric Priority (quality of service). This allows the application to maintain the desired quality of service (QoS) through the fabric. Further, it allows the priority level to be communicated between the host and the end devices for end-to-end management of application goals (Figure 5-5).

> **Availability note:** Brocade and IBM Z are exploring the full integration of Fabric I/O Priority to allow system management functions (such as z/OS Workload Manager) to maximize end-to-end, integrated quality of service characteristics.

Figure 5-5 shows Fabric I/O Priority.

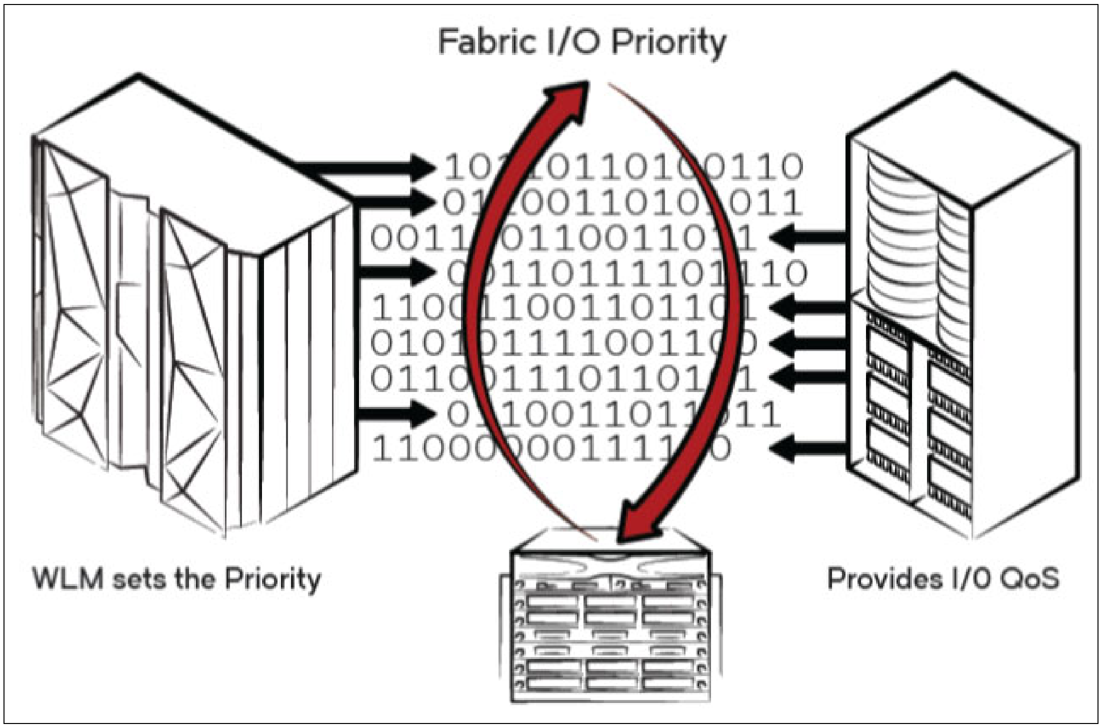


Figure 5-5 Fabric I/O Priority

5.1.3 Autonomics and Integrity

This section covers various strengths provided by IBM b-type FICON SAN products to provide autonomics and ensure integrity. Autonomic computing has the ability to self-manage and adapt to unforeseen challenges without passing on the complex maneuvers required if an operator were to manually address them. Integrity refers to methods ensuring that data is real, accurate, and protected against unauthorized manipulation. The following topics are covered in this section:

- Port decommission
- IBM Health Checker for z/OS Integration
- Enhanced diagnostics
- Forward error correction (FEC)

Port decommission

Port decommission is an autonomic, integrated function designed to simplify daily port/path maintenance. Routine port maintenance that once required significant administrative overhead and complex operations coordination can now be run seamlessly without operator intervention. Upon receiving an indication to remove a port or path from active service, Brocade Network Advisor coordinates the activities with both the host I/O subsystem and the fabric. After z/OS sends a confirmation that it migrated the active workload from the designated port/path, Brocade Network Advisor removes the port from operation and indicates it is ready for maintenance (Figure 5-6).

After the maintenance operations are completed, the port/path is returned to full operational status through the port recommission function. As with the decommission operation, Brocade Network Advisor coordinates and communicates the recommission action with z/OS to ensure the operation occurs error-free.

Figure 5-6 shows port decommission.

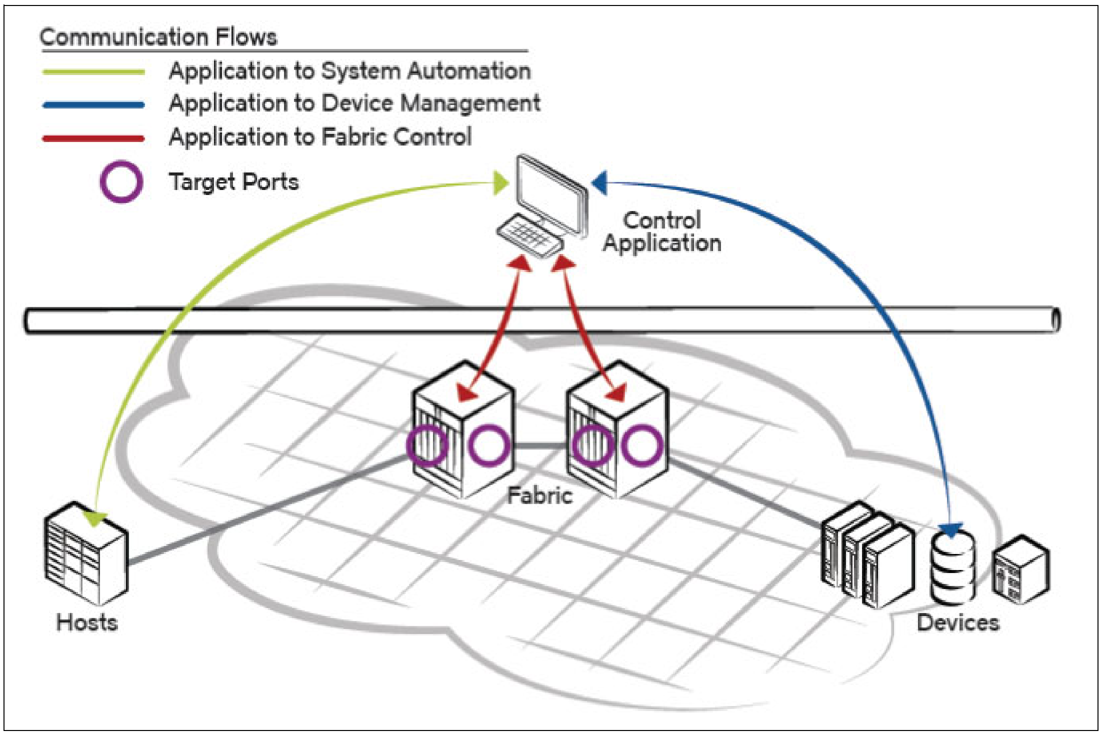


Figure 5-6 Port decommission

IBM Health Checker for z/OS Integration

The IBM Health Checker for z/OS extracts information from the FICON director through the Brocade CUP interface to the switch (see Figure 5-7). This information provides the Health Checker with the ability to identify and report the shared components of the mainframe storage network. The resulting path analysis provides system administrators with the insight necessary to ensure members of a path group are not susceptible to single-point-of-failure conditions.

In addition to the single-point-of-failure analysis, IBM Health Checker for z/OS leverages the FICON director CUP interface to assess performance characteristics of the members of a path group. The *flow* information provided includes utilization characteristics, operational conditions, and diagnostic error statistics. This unique solution also incorporates the Brocade Fabric OS (FOS) bottleneck detection mechanism, which provides alerts to z/OS when fabric anomalies are detected.

Figure 5-7 shows IBM Health Checker for z/OS integration.

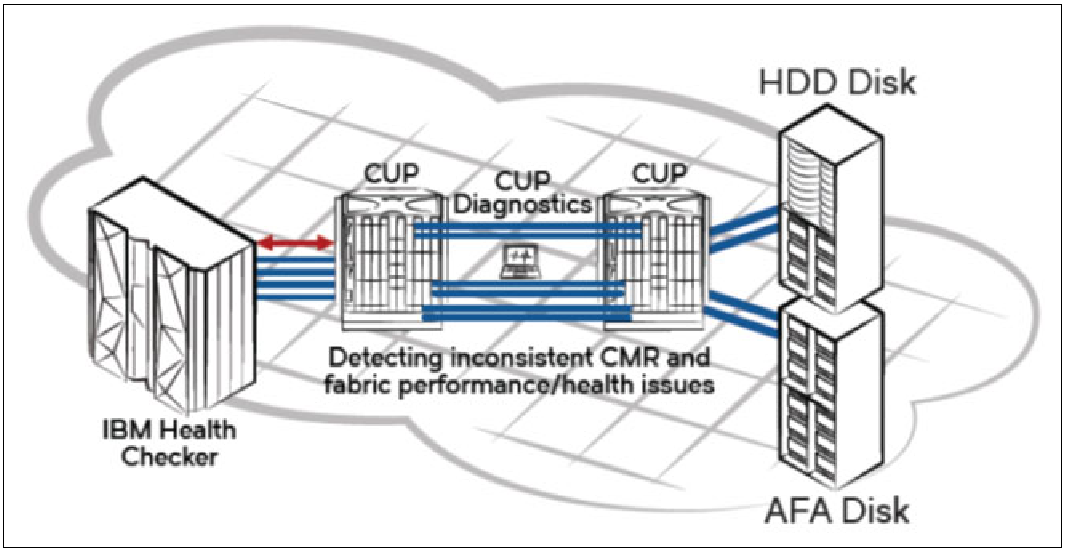


Figure 5-7 IBM Health Checker

Enhanced diagnostics

Keeping pace with the ever-growing complexity of mainframe storage networks, Brocade and IBM co-developed in-band diagnostic mechanisms that enrich the view of fabric components. The IBM Z channel and control unit solutions include functionality that allows the FICON director to read information about the optical transceiver attached to the IBM Z device (see Figure 5-8). This same functionality is used by z/OS to provide similar information about the transceivers installed in the director.

The new Fibre Channel standard Read Diagnostic Parameters (RDP) function was created to enhance path evaluation and automatically differentiate between cable hygiene errors and failing components. The RDP function provides the optical characteristics of each end of the link, including optical signal strength and environmental operating metrics, without requiring manual insertion of optical measurement devices. Furthermore, the RDP operation is performed periodically throughout the mainframe storage network to proactively detect marginal conditions that can be addressed before they affect production operations.

Figure 5-8 shows Read Diagnostic Parameters function.

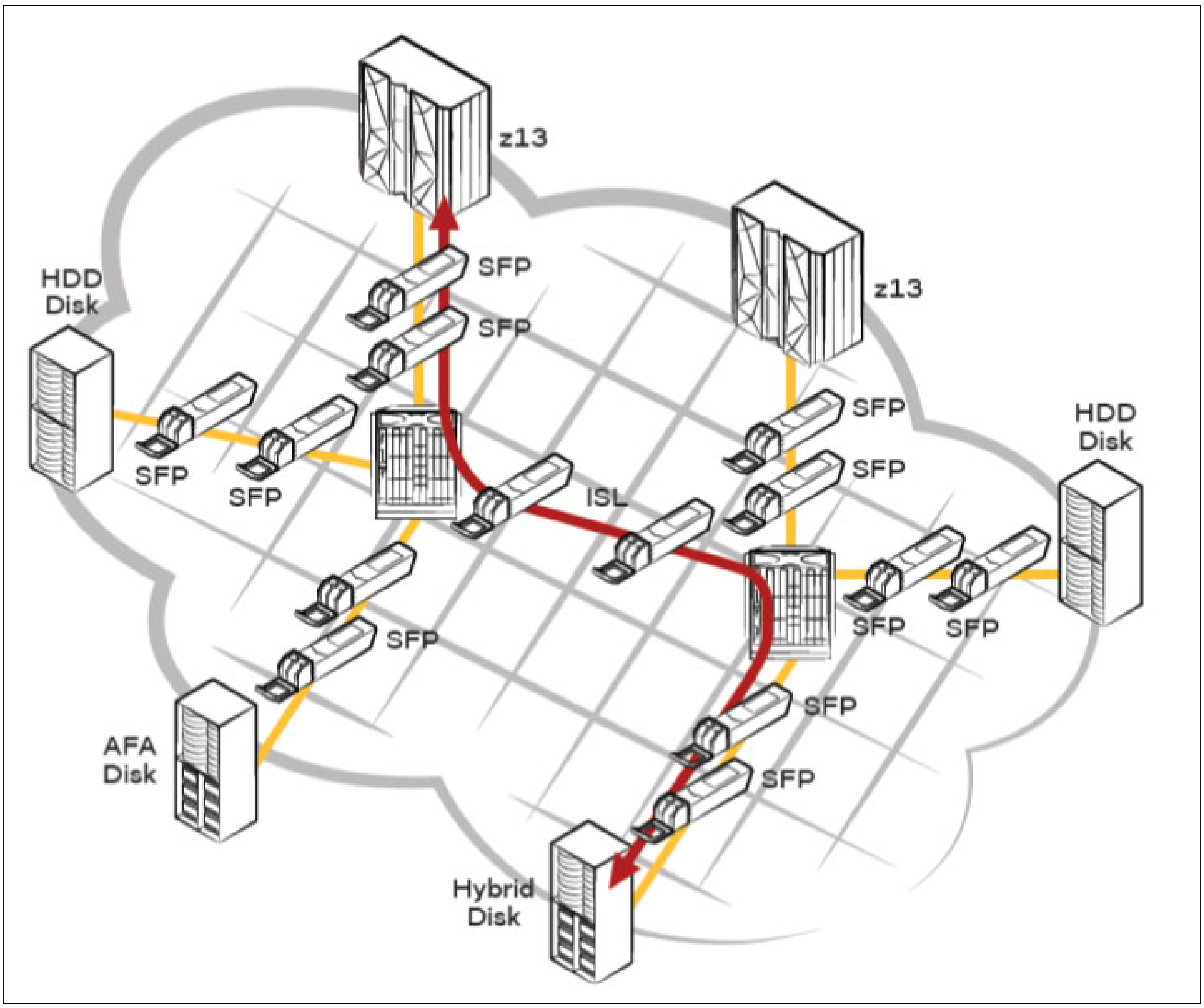


Figure 5-8 Read Diagnostic Parameters function

Forward error correction (FEC)

Gen 6 Fibre Channel includes an integrated FEC technology to provide reliable signal integrity at high data rates (such as 32 GFC/128 GFC). Gen 5 Fibre Channel offers this same functionality as an optional feature to be utilized at the solution provider's discretion. IBM b-type SAN products includes Gen 5 Fibre Channel FEC technology with all Brocade FICON director and switch solutions. The IBM FICON Express16S+ channel leverages this technology to ensure z/OS FICON solutions achieve the highest level of end-to-end reliability and performance in the industry.

Figure 5-9 shows forward error correction.

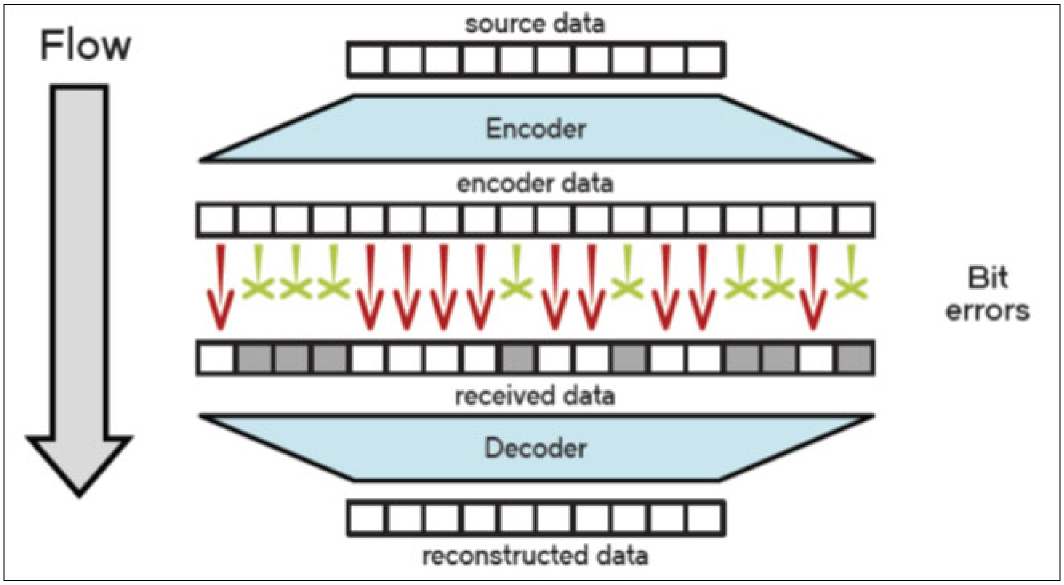


Figure 5-9 Forward error correction

5.2 Gen 6 Fibre Channel (FICON) performance benefits for IBM Z and Flash Storage

Gen 6 Fibre Channel brings three primary and immediate technology benefits to the IBM z14, IBM LinuxONE™, and flash storage architecture:

- Improved performance
- Improved virtual machine (VM) support
- Support for the upcoming NVMe over Fabrics technology

5.2.1 Improved performance

Enterprise-class storage performance has historically focused on four metrics:

- Latency
- Response time
- Throughput (bandwidth)
- I/O operations per second (IOPS)

Some common misperceptions exist, including that the terms *response time* and *latency* can be used interchangeably. Here is a review of what each of these metrics actually measures:

- *Latency* is a measure of delay in a system. In the case of storage, latency is the time taken to respond to an I/O request. In the example of a read from disk, this is the amount of time taken until the device is ready to start reading the block. Note, it does not include the time taken to complete the read. On the spinning disk (HDD), this includes examples such as the seek time (moving the actuator arm to the correct track) and the rotational latency (spinning the platter to the correct sector). Both of these actions are mechanical processes, and therefore slow.

  Latency in a storage system also includes the finite speed for light in the optical fiber, and electronic delay due to capacitive components in buses, host bus adapters (HBAs), switches, and other hardware in the path. The advantage flash storage or SSD has over spinning disk technology is the lower latency and the consistency in the latency. SSDs can exhibit latency that is lower by a factor of 100 compared to HDD technology, primarily due to the absence of the mechanical processes. Equally important, for random access I/O, the SDD is able to move data continuously with consistently low latency, while the HDD must continuously return for chunks of data, repeating the mechanical process and slowing performance.

- *Response time* is the total amount of time taken to respond to a request for service. It is the sum of the service time and wait time. The service time is the time taken to do the actual work requested. The wait time is how long the request had to wait in a queue before being serviced, and it can vary greatly depending on the amount of I/O. Transmission time gets added to response time when the request and response have to travel over a network, and varies with distance. Mainframe response time metrics for DASD include IOSQ, PEND, CONN, and DISC.

  Because SSD technology is capable of handling significantly more IOPS than HDD, the time spent waiting in a queue is typically shorter for SSD than HDD, leading to better response times for SSD. When SSD is coupled with Gen 6 Fibre Channel, studies have demonstrated a 71 percent improvement in response time.

- *Throughput* (MBps) is the data transfer speed. It is the measure of data volume transferred over time; the amount of data that can be pushed or pulled though a system per second. It is simply a product of the number of IOPS and the I/O size. As a performance metric, throughput matters most for large block-size data traffic.

  In a mainframe environment, these would be batch-type processing and data warehousing applications. Batch processing handles requests that are cached (prestored) and then executed all at the same time. An example of batch processing is the reconciliation of a bank's ATM transactions for a given day being compared against all bank records of all customers, and then reporting the ATM transaction records at a time outside normal business hours. Greater throughput capabilities, such as Gen 6 Fibre Channel's 32 Gbps, lead to shortened batch windows, which can be invaluable to large customers of IBM Z. When IBM introduced FICON Express16S channels in 2015, one of the primary cited benefits compared to FICON Express8S channels was an average 32% reduction in the elapsed clock time of the batch window.

- For a modern mainframe environment, throughput is secondary in importance to a storage device's *IOPS* capability. Flash storage/SSD technology is capable of significantly more IOPS than spinning disk technology. Mainframes are primarily used for online transaction processing (OLTP) workloads. Transaction processing consists of large numbers of small block-size data I/O transactions, often seen in banking, ATMs, order processing, reservation systems, and financial transactions. The most common block-size traffic seen in today's mainframe environment is 4,000 bytes. Coincidentally, that number is used for most storage performance benchmarking tests, and also IBM Z FICON channel performance testing.

The IBM z14 family of mainframes uses the latest industry-standard Peripheral Component Interconnect Express Generation 3 (PCIe Gen3) technology for its I/O subsystem (IOS). It also supports 50% more I/O devices per FICON and FCP channel (32,000 I/O devices per channel), with a maximum capacity of 320 FICON channels per z13. FICON Express16S+ channel is capable of over 300,000 IOPS running 4 kB block size, typical of OLTP workloads.

This increase in performance capability triples the capability of the previous Z generation (Figure 5-10 on page 104). The z14 processors are designed to handle massive increases in I/O demands brought about by CAMSS (cloud, analytics, mobile, social, and security) workloads. For many users, mobile workloads and their I/O requirements alone are driving the need for flash storage attached to z14.

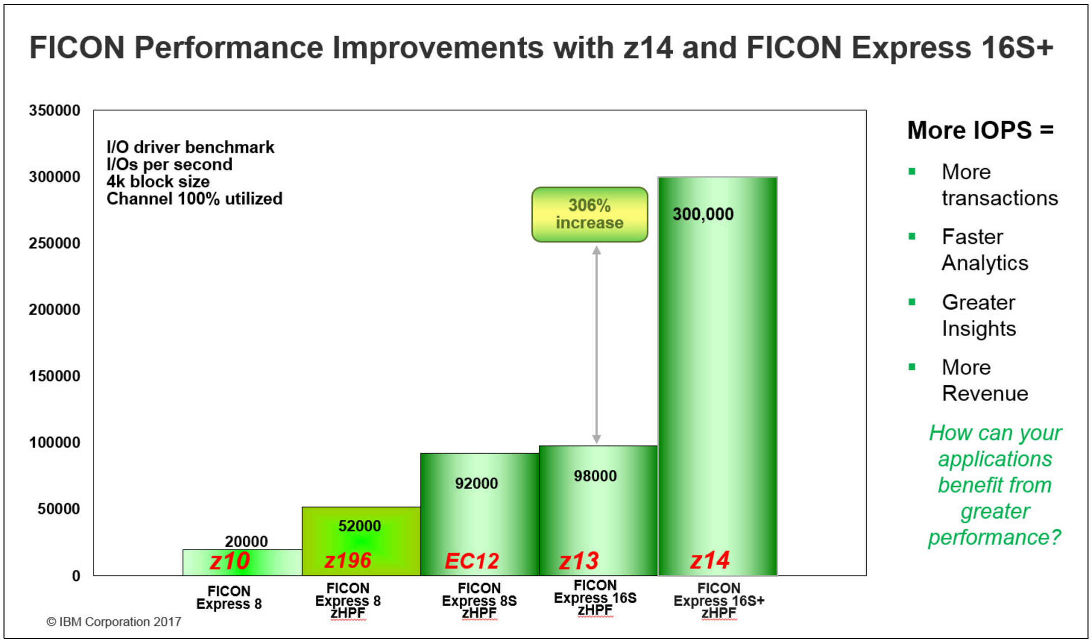


Figure 5-10 Performance improvements

IBM b-type Gen 6 technology provides a significant leap in IOPS capabilities. The Condor 4 ASIC is capable of handling 33 percent more IOPS than the prior generation Condor 3 ASIC used in Gen 5 SAN technology. Gen 5 FICON SAN technology is capable of outstanding performance in a z14 flash storage architecture. IBM b-type Gen 6 FICON SAN technology takes this performance to a whole new level and allows flash/SSD to reach its full performance potential.

The enterprise computing space has been at the forefront of the flash storage revolution. Applications that are performance-critical have been and are being moved from the traditional spinning disk DASD arrays to flash storage arrays. This movement started with hybrid arrays. However, when mainframe storage array vendors, such as EMC, HDS, and IBM, brought all-flash arrays to the market, users began migrating to the all-flash array technology. The reduction in latency and improved response times that can be realized by implementing flash storage array technology allow for faster processing of transactions, or more transactions processed for a given time interval. This can translate directly into more revenue.

For instance, see the High-Frequency Trading: The Co-Location Advantage article at the TABB Forum web page. It states the following estimate regarding performance:

> "Speed is important to high-frequency traders. It has been estimated that a one-millisecond advantage can be worth $100 million a year to such trading firms."

5.2.2 Improved VM support

IBM Z has up to 141 cores using the world's fastest commercial processor, running at 5.0 GHz, to enable high performance coupled with massive scaling. The z14 and the IBM LinuxONE Emperor™ systems are capable of supporting up to 8,000 virtual servers. When connected to flash storage via FCP channels and Brocade SAN hardware, such as Brocade X6 Directors, the z14 and the LinuxONE Emperor support Node Port ID Virtualization (NPIV).

This support of NPIV has increased significantly with z13 and LinuxONE Emperor. The increase doubled to support 64 virtual images per FCP channel from the prior support of 32 virtual images per FCP channel. The I/O subsystem of these IBM servers is built on Peripheral Component Interconnect Express (PCIe) Gen 3 technology. This combination can have up to 40 PCIe Gen 3 fanouts and integrated coupling adapters per system at 16 GBps each. The z14 also supports 32 TB of available RAM (three times its predecessor). This massive amount of memory, as well as server and I/O virtualization, allows VMs and applications to run at their full potential. Gen 6 Fibre Channel is ideal for these environments and is capable of supporting sustained high levels of performance, enabling full utilization of the z14 and LinuxONE Emperor virtualization capabilities.

Using Brocade Fabric Vision IO Insight capabilities, zLinux environments connected to storage through Brocade Fibre Channel technology can gain latency insight into application performance.

5.2.3 NVMe over fabrics

Non-Volatile Memory Express (NVMe) is an innovative method of accessing storage media. NVMe has emerged as the storage connectivity platform that will drive massive performance gains. It is ideal for flash/SSD. Applications can see better random and sequential performance by reducing latency and enabling much more parallelism through an optimized PCIe interface purpose-built for solid-state storage. The momentum behind NVMe has been increasing since it was introduced in 2011. In fact, NVMe technology is expected to improve along two dimensions over the next couple of years: improvements in latency and the scaling up of the number of NVMe devices in large solutions.

In 2014, the Fibre Channel Industry Association (FCIA) announced a new working group within the INCITS T11 committee (responsible for Fibre Channel standards) to align NVMe with Fibre Channel as part of the NVM Express over Fabrics initiative. This was an important evolution because it kept Fibre Channel at the forefront of storage innovation.

FC-NVMe defines a common architecture that supports a range of storage networking fabrics for NVMe block storage protocol. This common architecture includes the following advantages:

- Enabling a front-side interface into storage systems
- Scaling out to large numbers of NVMe devices
- Extending the distance within a data center over which NVMe devices and NVMe subsystems can be accessed

FC-NVMe offers compelling performance that is synergistic with the SSD and mainframe technologies that demand it. Gen 6 Fibre Channel is NVMe-ready and offers the ideal connectivity for high performance mainframe and flash storage.

5.3 Trends in storage array internal architecture

Spinning disk arrays are rapidly becoming a technology of the past as the cost of NAND based flash rapidly decreases. Over the past several years, hybrid arrays (combinations of spinning disk, and NAND based flash storage devices with the same form factor as a spinning drive) have become the norm. All flash arrays are becoming more and more prevalent.

The internal architecture of high end storage arrays today, in terms of the internal bus technology, is primarily PCIe Gen 3. The exception to this is the interfaces to the actual HDD and SSD devices. The hybrid arrays typically use the same device adapter/controller for connecting to both the spinning disk and flash devices. The technology used here is the same technology that has been used for spinning disk arrays for years: SATA/AHCI, SAS Gen 2/Gen 3, or Fibre Channel. The current generations of all flash arrays are still using this same device or adapter controller technology with their all flash arrays (with one exception being IBM). As faster versions of PCIe (Gen4 is now available) internal to servers and storage array bus technology become available, and faster connectivity between server and storage become available, another issue arises. These older architectures relying on SATA and SAS to connect to the end devices are rapidly going to become a bottleneck. SATA and SAS have the limitation of only one queue. Their throughput is also limited (SAS Gen 3 is 12 Gbps).

Rapidly emerging next generation non-volatile storage class memory (SCM) technology (such as 3D XPoint and resistive random access memory (RRAM)) offer 1000x latency improvements over existing NAND flash technology. They will exacerbate the device connectivity bottleneck. To solve this issue, device interface technology that bypasses or eliminates the traditional device adapter/controller is needed. These interfaces will use PCIe (aka NVMe), and take advantage of the massive parallelism offered by NVMe interface technology (64k queues).

More information about these interfaces is in the IBM Systems Technical University - Day 3 blog post.

Figure 5-11 shows PCI Express link performance.

PCI Express link performance

PCI Express Version	Line Code	Transfer Rate	Throughput* x1	x4	x8	x16
1.0	8b/10b	2.5GT/s	250 MB/s	1 GB/s	2 GB/s	4 GB/s
2.0	8b/10b	5 GT/s	500 MB/s	2 GB/s	4 GB/s	8 GB/s
3.0	128b/130b	8GT/s	984.6 MB/s	3.94 GB/s	7.9 GB/s	15.8 GB/s
4.0 (2017)***	128b/130b	16GT/s	1969 MB/s	7.9 GB/s	15.8 GB/s	31.5 GB/s
5.0 (2019)****	128b/130b	32 or 25GT/s**	3938 or 3077 MB/s	15.8 or 12.3 GB/s	31.5 or 24.6 GB/s	63.0 or 49.2 GB/s

*bi-directional throughput

**both rates are being considered for technical feasibility

*** PCIe 4.0 Announced by PCI-SIG 8 June 2017

**** PCIe 5.0 Preliminary announcement by PCI-SIG June 2017, expected to be standardized in 2019

Figure 5-11 PCI Express link performance table

5.4 Trends and directions in mainframe to storage connectivity

The year 2018 marks twenty years since FICON debuted. In this time period, IBM has shifted from the 1200 IOPS of ESCON, to the over 300,000 IOPS of zHPF on the z14 FICON Express 16S+ channels with a corresponding reduction in latency and response times for mainframe I/O. With the z14 announcement, a new mainframe I/O technology in the form of IBM zHyperLink Express was introduced. This IBM exclusive technology is positioned to be able to reduce latency on eligible (currently DB2) workloads by up to 5x. Connectivity is restricted to 150 meters distance using an ICA-SR cable that essentially connects a zHyperlink adapter port on a z14 I/O drawer to a similar port on an IBM DS8880 storage array. The DS8880 must be running code release 8.3 (or later).

IBM zHyperLink also makes a change away from the traditional mainframe interrupt driven asynchronous I/O methodology (using system assist processors (SAPs), the channel subsystem, and the IOS). Instead, zHyperlink uses a synchronous I/O methodology that is polling driven. This methodology is used when a determination is made that a called for I/O is *synch I/O* eligible. At that point, the channel subsystem and IOS are effectively bypassed, and instead are directly driven from the zHyperLink connection.

The DS8880 side of the connection uses a PEX8733 PCIe switch, which contains a direct memory access (DMA) engine. In essence, the internal PCIe Gen3 I/O bus of a z14 is connected directly to the PCIe Gen3 I/O bus of the DS8880. The latency associated with the traditional channel program/protocol/interrupt and FICON/FCP adapter cards is removed. The 2 microseconds of latency due to transiting FICON switching devices is also removed, but when compared to the latency reduction achieved by the change to synchronous I/O, those 2 microseconds are not the significant component of the reduction.

IBM zHyperLink still requires FICON channel connectivity, and FICON SAN for distance, and avoids direct attached FICON architectures. zHyperLink is restricted to 150 meters distance, IBM DS8880, z14, and specific workloads/access methods. It also functions best with read cache hits.

5.5 Summary

The enterprise computing community represents the unique sector of the storage networking market that propels the most critical systems in the world. These systems must operate cohesively and under extremely stringent conditions to provide the everyday services and operations enjoyed by billions.

Brocade Communications Systems, LLC, and IBM build on their 30 years of mainframe and storage networking leadership to deliver the industry’s highest performance and most reliable and scalable FICON infrastructure. With seamless FICON connectivity and support for innovative features that only Brocade can offer (including ClearLink D_Port, CUP diagnostics, port decommissioning, and ISL encryption) organizations can achieve the full potential from new flash storage and IBM z14 and future mainframe.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks domains and associated publications provide additional information about the topics in this document. Note that some publications might be available in softcopy only:

- IBM Storage Networking Redbooks:

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/san

- IBM Flash storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/flash

- IBM Software Defined Storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/sds

- IBM Disk storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/disk

- IBM Storage Solutions Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/storagesolutions

- IBM Tape storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/tape

You can search for, view, download or order these documents and other IBM Redbooks, Redpapers™, Web Docs, draft, and additional materials, at the following website:

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

Online resources

The following websites are also relevant as further information sources:

- IBM Storage hardware, software, and solutions:

  http://www.ibm.com/storage

- IBM Storage (for SAN):

  http://www.ibm.com/storage/san/

- Broadcom:

  http://www.broadcom.com

- Brocade:

  http://www.brocade.com

- Cisco:

  http://www.cisco.com
- Finisar:

  http://www.finisar.com
- IEEE:

  http://www.ieee.org
- Storage Networking Industry Association:

  http://www.snia.org
- Fibre Channel Industry Association:

  http://www.fibrechannel.org
- SCSI Trade Association (STA):

  http://www.scsita.org
- Internet Engineering Task Force:

  http://www.ietf.org
- American National Standards Institute (ANSI):

  http://www.ansi.org
- Technical Committee T10:

  http://www.t10.org
- Technical Committee T11:

  http://www.t11.org

The following web pages are referenced in the chapters of this book:

- IEEE Spectrum news:

  https://spectrum.ieee.org/semiconductors/optoelectronics/is-kecks-law-coming-to-an-end
- Gartner report:

  https://www.gartner.com/document/3357441
- IBM Storage Networking SAN24B-6:

  https://www.ibm.com/us-en/marketplace/san24b-6
- IBM Storage Networking SAN64B-6:

  https://www.ibm.com/us-en/marketplace/san64b-6
- IBM Storage Networking b-type Gen 6 directors (SAN256B-6 and SAN512B-6):

  https://www.ibm.com/us-en/marketplace/san512-256
- IBM SAN b-type Firmware Version 8.x Release Notes (FOS V8.X):

  http://www.ibm.com/support/docview.wss?uid=ssg1S1009577
- IBM Storage Networking SAN b-type (Brocade) family:

  https://www.ibm.com/storage/san/b-type
- IBM DS8880 hybrid storage:

  https://www.ibm.com/us-en/marketplace/ds8000#product-header-top

- IBM System Storage Interoperation Center (SSIC):

  https://www.ibm.com/systems/support/storage/ssic
- The NIST Definition of Cloud Computing publication:

  http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf
- IBM Storage Tape TCO Calculator:

  https://www.ibm.com/systems/storage/tape/tco-calculator/
- Brocade SAN Health:

  http://www.brocade.com/en/support/support-tools/brocade-san-health.html/
- Brocade Network Advisor:

  http://www.brocade.com/en/products-services/network-management/brocade-network-advisor.html
- Accelerating Oracle Data Warehousing with Emulex and Brocade Gen 6 Fibre Channel (Demartek article):

  http://www.demartek.com/Reports_Free/Demartek_Emulex_32GFC_HBA_Oracle_Data_Warehousing_Evaluation_2016-07.pdf
- IBM Service Request:

  https://www.ibm.com/support/servicerequest/
- IBM Support:

  https://www.ibm.com/support/home/?lnk=msu
- Fix Central:

  https://www.ibm.com/support/fixcentral/
- FICON Dynamic Routing (FIDR): Technology and Performance Implications:

  https://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102651
- IBM Systems Technical University - Day 3 (blog post):

  https://www.ibm.com/developerworks/community/blogs/InsideSystemStorage/entry/IBM_Systems_Technical_University_Day_3?lang=en
- TABB Forum article: “High-Frequency Trading: The Co-Location Advantage”

  http://tabbforum.com/opinions/high-frequency-trading-the-co-location-advantage

Help from IBM

IBM Support Portal and downloads:

https://www.ibm.com/support/entry/portal/support

IBM Global Technology Services:

http://ibm.co/1lyI24R

Redbooks

Modernizing Your IT Infrastructure with IBM b-type Gen 6 Storage Networking and IBM Spectrum Storage Products

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SG24-8415-00

ISBN 0738442933

Printed in U.S.A.

Get connected

ibm.com/redbooks