Compliments of :

HITACHI
Inspire the Next

BROCADE
A Broadcom Company

Networking Next-Gen Storage

for dummies®
A Wiley Brand

Understand storage network basics

Evaluate storage networking options

Plan your next-gen storage fabric

AJ Casamento

Brocade Special Edition

Brocade and Hitachi - Alone, Exceptional. Together, Unrivaled.

Driven by one of the closest, longest-standing alliances in the industry – Hitachi and Brocade have been collaborating for more than 20 years, focusing on equipping today's customers with the technology to become the dynamic enterprise of tomorrow.

Hitachi Vantara and Brocade set the gold standard in future-proof storage networking performance, availability, resilience, simplicity, and security for the evolving organization. Integrating automation architecture, lowest power consumption, and the most advanced SAN monitoring and diagnostics in the industry.

About Brocade

Brocade, a Broadcom Inc. Company, is the proven leader in Fibre Channel storage networks that serve as the foundation for virtualized, all-flash data centers. Brocade Fibre Channel solutions deliver innovative, high-performance networks that are highly resilient and easier to deploy, manage, and scale for the most demanding environments. Brocade Fibre Channel storage networking solutions are the most trusted, widely deployed network infrastructure for enterprise storage.

www.broadcom.com

About Hitachi Vantara

Hitachi Vantara, a wholly-owned subsidiary of Hitachi, Ltd., guides its customers to what's next by solving their digital challenges. Hitachi Vantara applies its unmatched industrial and digital capabilities to its data and applications to benefit both business and society. More than 80 percent of the Fortune 100 trust Hitachi Vantara to help them develop new revenue streams, unlock competitive advantages, lower costs, enhance customer experiences, and deliver value.

www.hitachivantara.com

Networking Next-Gen Storage

Brocade Special Edition

by AJ Casamento

for dummies®
A Wiley Brand

Networking Next-Gen Storage For Dummies®, Brocade Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

ISBN 978-1-394-15980-2 (pbk); ISBN 978-1-394-15981-9 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

We're proud of this book and of the people who worked on it. Some of the people who helped bring this book to market include the following:

Development Editor: Faithe Wempen

Project Editor: Martin V. Minner

Editorial Manager: Rev Mengle

Associate Publisher: Katie Mohr

Business Development Representative: Karen Hattan

Production Editor: Tamilmani Varadharaj

Brocade Contributors:
Michelle LeMieux-Dimas, David Schmeichel, Scott Shimomura, Juan Tarrio, Marcus Thordal

Table of Contents

Introduction

You've likely noticed the rising tide of data in your life. Everywhere you turn, more and more data is being generated — by you, for you, and about you. In addition, you likely find yourself with a level of personal storage capacity at your disposal that would have stunned large computer administrators just a few decades ago. It's easy these days to carry the equivalent of an entire reference library around in your pocket.

With more data come more concerns. How is your organization going to manage all that data across multiple locations? Protect it from hardware failures, malware, and hackers? Move it swiftly and silently from place to place as needed? The most common way that businesses cope with these concerns today is by using a *storage area network (SAN)* that connects storage devices in multiple locations with high-speed hardware and can be configured to enforce data management policies. The hardware involved in connecting servers and workstations to storage is commonly referred to as a *fabric.* Fabrics that use the latest and best technologies are known as *next-generation.*

About This Book

This book, *Networking Next-Gen Storage For Dummies*, provides an easy-to-understand introduction to storage fabrics for the non-technical person. It explains what next-generation storage fabrics are, how they are commonly implemented, and what advantages they offer compared to less modern and less robust technologies. Armed with this information, you'll be able to make smart decisions about how storage fabrics fit into your business's IT plan.

If you're new to storage network technology, don't panic! The business and technical considerations involved are many, but this book will guide you through the process of making the important decisions. You'll see that modern storage fabric is not the arcane and complex mystery that you've been led to expect. In fact, depending on the type of storage fabric you select, it may be an order of magnitude easier to work with than a traditional network.

Obviously, no book can recommend a single storage solution that will work for all businesses, because each business has its own requirements and priorities, but this book will teach you the right questions to ask to find the perfect solution for your business's needs.

This book will likely not be the only reference you consult as you consider a next-generation storage fabric. It provides a great starting point for both the technical and the non-technical person; depending on your role in the organization, you may want learn more about the technologies behind storage fabrics or consult an implementation guide designed for IT professionals.

Icons Used in This Book

To make it easy to navigate to the most useful information, these icons highlight key text:

TIP

Follow the target for tips that can save you time and effort.

WARNING

Watch out for these potential pitfalls on the road ahead.

REMEMBER

Take careful note of these key takeaway points.

TECHNICAL STUFF

Read these optional passages if you crave a more technical explanation.

IN THIS CHAPTER

- » **Defining storage networks**
- » **Reviewing the history of storage networking**
- » **Learning why storage network type matters**
- » **Understanding what you're getting**
- » **Thinking about what might go wrong**

Chapter 1 Introducing Storage Networks

A *storage network* (also called a *storage area network,* or *SAN)* is an environment that allows data storage resources to be shared between multiple systems over a network.

Storage networks can vary dramatically in the ways they are used, in the types of protocols they employ, and in the business requirements that dictate how the storage is configured and managed in terms of scale, reliability, or performance.

This chapter introduces storage networks and explains some ways that they differ from one another.

A Quick History

Business requirements for storage today look very different from those of previous generations. Early storage solutions were simple, local, and slow. The technologies and methodologies have evolved over time in response to administrative hassles and bottlenecks encountered along the way.

In the beginning, there was *local storage* — that is, storage directly attached only to an individual server or workstation. This type of storage is also sometimes called *direct-attached storage*, or *DAS*.

Direct-attached storage had some operational issues, including the following:

- **Forecasting:** How do you know how much storage a server or application is going to use during its lifetime? Do you pre-pay all of the expected storage capacity up front?
- **Maintenance:** How do you do maintenance or repair on the storage environment without taking the application server offline?
- **Availability:** How do you guarantee to keep mission critical applications online with only a single system?

REMEMBER

Murphy's Law (the popular version) observes: "Whatever can go wrong, will go wrong." The corollary for this law in the storage world adds: "And it will happen in the manner that causes the maximum data loss."

Because of the inability to predict the future amount of storage needed for a system, some systems would run out of storage, while others would have excess storage that they couldn't share. Too little storage could result in downtime while more storage was added. Too much storage meant hardware going to waste (sometimes called *stranded capacity)* at a time when storage was still relatively expensive.

An early solution to these problems was the creation of the *Network File System (NFS)* by the folks at SUN Microsystems back in 1984. NFS enabled one server to offer its excess storage capacity to another server in the same network as an extension of its file system. NFS rapidly gained popularity as a means of using stranded capacity. However, the other servers getting their storage capacity externally were dependent upon the uptime of the server hosting the storage for their success. The hosting server could inflict outages on any devices that relied on its storage.

Next came the development of external storage arrays that did not rely solely on any one server. They provided resilience, redundancy, and scalability that IT departments were clamoring for. Any of the servers in the network could be updated without affecting others. Storage could be dynamically added to an application server without downtime. Redundant copies of the data could be created for reliability and security without consuming application server performance. Instead of managing many small instances of storage across all the servers, IT professionals had to manage only a few large storage arrays.

Evaluating Your Priorities

Storage networks are not all the same! You have a lot of choices, and what you choose makes a big difference. The ideal network type varies greatly based on what kind of data you will keep in the storage network and how it will be used.

One size doesn't always fit all, even within a single organization. You may find that you have multiple levels of needs that can best be met by different technologies. For example, there are significant security requirement differences between storage used for corporate e-mail versus storage used to hold payment card information. Likewise, the performance requirements are very different between storage used for archiving and storage that has a service-level agreement to support transactional databases.

TIP

In some cases, choosing different technologies for different areas of your organization can make sense, but keep in mind that it increases the administration burden. It may be more cost effective in the end to deploy the same storage environment across more connections, so you can amortize the cost of a single system across more connections and streamline its support.

To begin thinking about your storage network needs, start with the most demanding storage need you have. For example, what about your business-critical data? What are your priorities for the primary system that stores your most important data?

Necessary elements to consider in the design of such a network might include:

- Security
- Availability
- Performance
- Scalability
- Manageability

Security, obviously, is a top concern for business-critical data — not only the data needed to conduct financial business, but the data collected from customers and partners. No business can afford to lose control of sensitive data such as payment information, medical records, and so on. Not only would it cripple business operations and alienate customers, but it might also result in legal action against the company for violating consumer protection laws.

For business-critical data, you need maximum availability. Losing access to critical business data, even for a few minutes, can be disastrous. The storage system should be so reliable as to *never* go offline.

Some storage system types require you to make trade-offs among the qualities you want. For example, some solutions might be able to scale to large environments but not have fast enough performance. Others might have very high performance but not the mission critical availability. And some might be difficult to modify and maintain. All of these are things to be taken into consideration when determining which technology version of storage networking is best suited to your needs.

Understanding What You're Getting

In today's world, most people just want the results. They don't know — or want to know — how an infrastructure functions or why certain components are chosen over others.

Fair enough. You don't want to learn how a sports car's engine works — you just want to drive fast.

WARNING

People should learn at least the basics, though, for their own good, and the good of their businesses. Understanding the factors involved in setting up a storage network, and the consequences of the choices made, can save significant time and money in both the short and long terms. To go back to that car analogy, if you have a high-performance sports car, you need to know how fast the car can go before it becomes difficult to control. You need to know how often to change the oil and rotate the tires. You need to know that a sports car is not your best choice if your primary goal is to haul heavy equipment or transport a lot of people at once.

Similarly, with a storage network type, you need to understand the pros and cons of the available technologies so you can determine whether a particular technology is the best fit for the application or data that is involved.

There are plenty of stories to exemplify in detail how horribly wrong things can go if there is a mismatch between what a business's leadership team *thinks* it needs and what it *actually* needs.

For example, some types of data are subject to regulatory control, and some storage network types are better at managing that than others. Perhaps the data must be archived for a given number of years, or there must be strict controls on who accesses it. The way the data and its application are managed might also have implications for other groups in the business who are downstream of the system. All of these would constitute requirements that have to be taken into consideration when looking at what storage networking technology to deploy.

What Could Go Wrong?

What's the worst that could happen, and how will your organization recover from it? It's important to ask that question — and get a good answer — as you are planning your storage network.

REMEMBER

Keep in mind that the impact of a storage outage is very different from a temporary network performance issue. If the network is slow, people might be frustrated because they have to wait a few more seconds for their data. But if the storage isn't available at all, everything grinds to a screeching halt.

Any person who owns a smartphone has at some point experienced the frustration of the network not being available. However, the smartphone can still be used for other functions and is still operational when offline. When the network connectivity is restored, you can immediately resume use with the Internet, mail and phone capabilities. On the other hand, if the *storage* inside the smartphone (which houses your photos, personal data, e-mail or the actual operating system of the smartphone) stops functioning, the device becomes a useless brick. The two types of outcomes are completely different.

Now imagine applying this scenario to a business environment. The issue might be that the deployed network can't handle the transactional workload, or that the data availability isn't up to the business needs.

Consider the modern hospital. Virtually all of the data is now electronic, including medical imaging, physician notes, prescribed medications, medical history, and so on. Applications automate checking prescribed medications against medical records to determine whether a medication is safe for the patient to take. One health care customer described the potential impact of a data outage as a total nightmare because "No one knows how to go back to the paper system. We'd be shut down!"

TECHNICAL STUFF

Here are a couple of key measurements to consider as you think about "the worst that could happen."

Recovery point objective (RPO): *How much data can your application permanently lose and you'll still be happy?* Five minutes' worth? Thirty seconds? Or zero — none at all? Different answers to this question require different technology selections.

Recovery time objective (RTO): *How long can the application be unavailable and you'll still be happy?* To put it another way, what is the direct business impact or opportunity cost of the application being offline? The answer may vary. For example, think about a company that sells tickets online. Imagine the difference in lost revenue between an outage during ticket sales for a local county fair versus ticket sales for a major music concert on the opening day of sales.

Each organization — and each application run within it — has its own set of uptime and data integrity considerations that affect the choice of a network storage technology.

IN THIS CHAPTER

- » Understanding the differences in data, storage, and networks
- » Differentiating between fabric and non-fabric networks
- » Avoiding outages
- » Keeping customers satisfied

Chapter 2 Eliminating Common Misunderstandings

People tend to believe that if they understand one instance of a thing, all other instances are pretty much the same. Data is data, right? Storage is storage, and networks are networks?

Nope. Not even close. There are some pretty dramatic differences between data types, storage media, and network types, and those differences affect the choices an organization makes in its data storage networks. This chapter explains some of these variables and the misunderstandings surrounding them.

All Data Isn't the Same

This may sound like a very simple statement to make. It's a critical statement to understand, though, because its effect on the decisions that IT staff make is profound. Incredible variations exist in the types of data in almost any working environment, and it can be difficult for IT staff to precisely delineate which data is which.

For example, how do you define "mission critical" data? It's an easy call to say that an online order entry system that accepts payment card information is mission critical, but other important data might not stand out so obviously. For example, an e-mail from a customer outlining the requirements for a design or a major quote might be just as mission critical in terms of long-term financial impact. To make life even more interesting for IT staff, that e-mail for the design requirements sits right next to an e-mail from one team member to another about where they want to go to lunch.

WARNING

Various data sets can have different levels of business impact. For example, in a pharmaceutical manufacturing facility, the data from the manufacturing floor concerning how the medications were made is subject to government regulation. This data must be retained so that if a batch of that medicine shows a problem, the history of how it was manufactured is available for review. Simple, right? Except that a loss of even 30 minutes of that data may mean that the entire production (potentially millions of dollars' worth of product) must be scrapped because it cannot be sold without that data being on file.

Data can also vary as to its security requirements. How protected does the data have to be from unauthorized views? An organization's handling of medical records or payment history is not only a customer relationship issue, but also a regulatory issue. There are hundreds of laws that have different — and many times conflicting — requirements on how such data must be treated. Further, these laws may vary by the country or state in which the business or the customer resides.

There is also the issue of application or data availability. How important is it that a particular data set be available all the time? In years past, there were "quiet periods" in an IT day. Business hours had a night shift where housekeeping activities like data backup could be done without harming business performance. Similarly, system maintenance could be handled in weekend or holiday service windows. Today, with an online presence, almost all businesses must be able to conduct business 24x7. This increases the availability and retention needs to "24x7xforever" with no idle periods.

All Storage Isn't the Same

There have always been differences in the types of storage available to an IT environment. Some of the main differences include:

- The type of media (traditional magnetic hard disk drives, solid-state storage)
- The interface that the media uses (serial attached SCSI, serial ATA)
- How the data is presented to the application (block storage, file storage, object storage)
- The network interface that is used to connect to the server hosting the application (Direct Attach SCSI, Ethernet, InfiniBand, Fibre Channel)

There are also differences in how efficient the storage is at using the space. Early on, it was sufficient to simply have an exact full copy of the data — a mirror. To speed up data access, data was also written across multiple disks (striped) so that access wasn't sequential. Striping and/or mirroring data storage across multiple disks is known as Redundant Arrays of Independent Disks (RAID). Different types of arrays implement the system in various ways, which can have implications for services and performance.

There are also differences in the reliability of the storage systems themselves. The common understanding in IT is that a single copy of any data set — a single point of failure — is disaster waiting to happen.

And, of course, it's inevitable to have to ask how much all this will cost. Some might think the cheapest solution possible is the best, but in reality, you get what you pay for. Even though putting in a single disk drive into a server may seem cheaper, the data protection capabilities of that are virtually non-existent, and it exposes the business to more costs for application downtime, data loss, and liability.

Even the most cost-sensitive customers realize that you need to copy or replicate data outside of the server to create a secure copy. This realization was the genesis of external storage. One of the fastest alternatives for creating secure copies of data is a simple mirror. For every drive, you have a second physical drive that duplicates it in real time. But this means that every storage volume is double the cost (and double the equipment storage and maintenance too). Other alternatives might seem cheaper because they use less storage (such as parity or erasure encoding), but they require a trade-off in terms of performance. Not every business need can afford to give up performance or reliability.

WARNING

Don't make the mistake of assuming that every product in the market has the same capabilities. It is not unusual for IT organizations to have tiers of storage where each tier has different priorities and requirements. A common description might be that Tier 0 is uncompromised performance and reliability, Tier 1 might be good performance and reliability, and Tier 2 might be reliability.

Be thoughtful about what the actual application needs are and what the realistic capabilities of the storage platform are.

All Networks Aren't the Same

This may seem like an obvious conclusion, but it catches many people unprepared. There are three commonly discussed network types in the storage market: Ethernet, InfiniBand, and Fibre Channel. Each has its own unique blend of pros and cons.

REMEMBER

By volume and adoption, the two most prevalent networks in the IT market space are Ethernet and Fibre Channel, while InfiniBand is frequently considered as a niche player in the high-performance technical computing market. Fibre Channel is generally seen as the high-performance, mission critical solution set for storage. Ethernet's main benefit is pervasive connectivity because Ethernet is the dominant networking technology worldwide.

However, the catch is that different technologies have different characteristics, and the choices are not always simple. Some things to consider include:

- Shared network or dedicated
- Reliability
- Performance
- Scale

- Cost
- Ease of use

To understand these factors, think about roads and traffic. What types of cargo are being moved? How time critical is it (for example, is it furniture or an ambulance)? What is the mix of traffic (freight or passengers)? What is the volume of the traffic (neighborhood streets or eight-lane freeway)?

This is one of the places where bandwidth, IOs per second (IOPs), and latency all come into play. Think of bandwidth as the number of lanes on the highway, IOPs as the number of vehicles per second passing a given point on the highway, and latency as the length of time it takes to make the trip. These are related characteristics, but they are not the same measurement.

In the Ethernet world, the most common transport mechanism is something called TCP/IP (which stands for Transmission Control Protocol/Internet Protocol). There have been a number of modifications since its publication in 1974, but the basic premise was to provide reliable delivery of data in an unreliable physical network.

To continue the transportation metaphor, say you have a number of people (data packets) who need to go from one location to another. You assign them numbers and send them on their way. They may or may not follow the same path, and if they arrive out of order, the receiving location puts them back in order by number. If one or more of them is lost to an accident, or if there is insufficient parking at the destination and they take too long to get there, the system marks them as failed and sends a replacement. If there are a lot of failures, large volumes of retransmission can create more congestion, loss, and retransmission. A common fix for this is that TCP "backs off," which is to say it may slow or stop the traffic while the congestion clears and then ramp back up. Although this method works for reliability, the network's utilization and performance are compromised.

Fibre Channel and InfiniBand (as well as other protocols such as Intel's PCI) use something called a *buffer-to-buffer credit mechanism.* An easy way to think of this is that every person going from one location to another has an assigned parking space at the destination before they are put on the road. Additionally, barring a network outage, Fibre Channel sends the entire transmission on the same path so all the passengers (packets) arrive in the original order.

Now imagine that each of these passengers is carrying part of a document. The normal TCP/IP implementation for Ethernet has the passenger deliver his or her document to the front desk of the building, where it is copied and sent to the floor (and potentially copied again before going to the meeting room where it is needed). Fibre Channel, in contrast, uses a mechanism called Zero Copy that sends the person and the document directly to the room where the document is needed.

Performance also pertains to the types and volume of data to be moved. Individual cars (Ethernet packets) have a certain efficiency. Small buses (Ethernet jumbo frames) have a better efficiency but a different impact on the network carrying them. Mixing the two types of traffic can present a balance issue as well. Adding freight trucks (storage traffic) introduces more complexity.

Attempts can be made to isolate certain traffic in virtual networks. Think of this as a "trucks prohibited from the left lane" designation. But the traffic is still there, and still affecting scale and performance. To this point, several implementations of Ethernet-based storage (including data center bridging or DCB Ethernet switch products) have begun promoting dedicated networks, either as a best practice or in some instances as an absolute requirement. Isolating the storage traffic avoids many of the traffic mix issues, but at an additional cost.

Fibre Channel is always deployed in dual-redundant, hardware-isolated fabrics. A *fabric* is more than a simple network interconnect; it is a network interconnect that has an overlay of intelligent services. This is done for multiple reasons.

First, isolating the storage traffic to Fibre Channel means that the traffic mix is simpler to handle. By largely freeing up the Ethernet network applications from the need to deal with the storage traffic — or at least the heaviest, most mission critical storage traffic — this practice provides much better performance as well.

Second, the dual-fabric approach means that the system has an extremely high availability. On a single combined/converged Ethernet network, any significant network problem (including human error) becomes a problem for everybody and creates the potential for outages. In Fibre Channel, an outage in one of the redundant fabrics does not take the application down. Because the fabrics are isolated hardware, a single human error in a network command doesn't drop both fabrics simultaneously the way it does in Ethernet.

Third, by isolating the mix of storage traffic and application/client traffic, as done in a Fibre Channel SAN, the scale of the storage network can be much larger. If you are not having to deal with the mix of passenger cars and freight trucks, but are only dealing with freight trucks, you can move considerably more cargo in a much broader network.

This doesn't remove the need for small neighborhood streets or single-lane gravel roads, of course. It just means that you need to consider how much data is being moved, what type it is, how critical it is, and how easily you can recover the traffic.

You may very well find that no one environment solves every use case in your data center. You will probably need to balance the needs for performance, reliability, and scale against the cost of the solutions that will meet those demands.

WARNING

As discussed previously, the costs are not only (or even mostly) the initial purchase of the technology (although it frequently seems as if the finance teams are *only* able to see that item). In fact, the operational costs can dwarf the capital costs, and the cost of an outage may not only exceed the project cost but might bankrupt the company, a scenario that has been reenacted far too many times.

Not Every Network Is a Fabric

The term *fabric* generally refers to the network hardware used to create the storage network, but not every collection of network hardware constitutes a fabric. What differentiates a fabric from a typical network is the automated services that it includes, such as:

- Fabric configuration of switches
- Address assignment of devices
- Registration of devices to name services
- Recovery of services after an outage
- Updates for network configuration changes

Is a certain network a fabric, or not? The answer can vary depending upon which IT culture you are speaking with: server, storage, or network. This conversation can be widely divergent.

A fabric can be identified by the services the network provides and how automated those services are. Given the scale of today's data centers in terms of the sheer volume of applications (whether tied directly to a server, a virtual machine, or even a container), automation and services become a paramount consideration.

Configuring a network

In order to provide a fabric level service, configuration of a network should be automatic. Most networks don't handle this automation cleanly, but a Fibre Channel fabric automates the process well.

If you wanted to create a ten-switch environment in Fibre Channel, you could do this in three simple steps:

- Unbox the switches and put them in a rack.
- Connect the switch ports together in the topology you choose.
- Power the switches on.

In contrast, that same ten-switch environment using Ethernet is a full day's worth of work. Hundreds of configuration commands would be necessary to complete this task, including making certain that you haven't created any illegal connections between switches by accidentally configuring loops.

Connecting servers and storage

Once you have a network in place, the next step is connecting the servers and storage to the environment that you want to use. Again, Fibre Channel is much easier. Ethernet does offer a mechanism for automatically configuring a network address Dynamic Host Configuration Protocol (DHCP), but part of this effort involves mapping the physical device to the network Internet Protocol (IP) address using the Address Resolution Protocol (ARP) to make certain that it is unique.

A Fibre Channel fabric, however, registers all server or storage nodes via a Fibre Channel Name Service as part of the Fibre Channel standard.

Automatic Name Server registration is an integral part of the fabric services in any Fibre Channel fabric. It also allows for dynamic recovery if a temporary outage causes a connectivity issue. Contrast this with, for example, an iSCSI network where all of the devices would have to be manually reconnected to the storage services after a network outage.

Avoiding Outages

When an outage occurs in a storage fabric, the impacts are potentially significant and noticeable. And as a rule in IT, the goal is *not* to be noticed. Most managers don't pay attention to IT until something doesn't work — and then there's hell to pay. So for an IT professional, not being noticed usually means not being fired.

An outage in a storage fabric can cause a degradation of application performance or a temporary unavailability of the app. It might even mean a loss of actual data. None of these scenarios is a happy circumstance for the storage administrator.

Any hardware in any infrastructure eventually fails. It's inevitable. The goal of a well-designed infrastructure is for the end-user not to *notice* that some piece of the infrastructure failed. This requirement becomes increasingly more important as the criticality of the application being discussed increases. For example, when hospital IT systems go offline, the results can be life threatening. So, although a particular storage element may have failed, it is critical for the storage services and data to remain accessible.

One of the reasons storage fabrics are increasingly popular in these environments is that humans are no longer fast enough to cope. The environments are too big and too widespread. The interaction of the applications and the impact on every aspect of the business is too entwined to wait for human intervention at every step. IT administrators need additional functionality and automation in order to cope with this continuing scale of involvement. They need fabric services and intelligence to mitigate issues in an automated fashion as they occur. The administrators can then do follow-up analysis and problem solving when they have time to respond.

Keeping the Customer Satisfied

The uncomfortable reality is that application developers/owners and end-user consumers of IT do *not* care about infrastructure. They think of infrastructure the way most people think about microwave ovens. None of your friends and neighbors likely understand the science behind their microwave. All they want is to take last night's lasagna, put it in the microwave, push some buttons, and get today's lunch.

REMEMBER

This is how consumers view IT, and particularly the storage infrastructure. They don't know how it works or its limitations, nor do they want to know. If they are the developer, they want the application deployed. They want it to work at peak performance, they want it to scale on demand, and they want it to be secure and available. They don't want to know the technology by which these results are achieved, nor do they want to have to monitor or fix the infrastructure. They just want it to work. Therefore, the storage fabric must deliver:

- Availability
- Scalability
- Performance
- Security

These characteristics are a minimum for keeping consumers happy.

IN THIS CHAPTER

» Ethernet versus Fibre Channel

» Protecting the data

» Balancing measurement and performance

» Avoiding disaster

Chapter 3 Operating Your Storage Network

One of the common comments bandied about the halls of IT (or at least by blog writers) is "We need the EASY button." It's a way of saying that the IT environment should be both responsive and simple.

The fantasy goes something like this: Some internal or external customer is in the midst of uttering the phrase, "We need . . ." and before they can finish hitting the button, the solution is there! Just like magic! Reality doesn't work that way, unfortunately, but Fibre Channel fabrics can make things a lot easier.

Of course, the reality is that applications require infrastructure, and the mission critical applications require very solid infrastructure. *How* solid is dependent upon the types of applications and data that are being hosted in and transacted across that infrastructure. When you have conversations with the line of business owners about the reliability requirements of the application, it is possible that the response will be that zero downtime is allowed and zero data loss is allowed. Zero is a valid service-level agreement (SLA). But zero has infrastructure requirements. For example, there can be no single point of failure.

TAKE IT TO THE CLOUD?

Some folks throw up their hands at the complexity of an in-house solution and say "Take it to the cloud!" Outsourcing (which is what "cloud" translates to) may be the right thing for some implementations or businesses, but the responsibility still belongs with the IT team initiating the consumer contract. Making certain that "somebody else's server" meets the real criteria or SLA that your consumer expects is an important consideration. It's also one of the reasons companies repatriate some of their critical workloads back from the cloud based on both SLA and cost.

How does this apply to the storage network? Think about it this way: A traditional data center Ethernet network is frequently referred to as a "peer-to-peer" network — effectively a conversation between intelligent equals. A shared storage topology (of any description) is not a conversation between intelligent equals, but rather a collective of intelligent equals attempting to time-slice a shared resource. The behavior of one server can adversely affect the shared environment and the connectivity of other servers.

Fibre Channel SAN is perhaps the most prominent technology in use in data center storage for requirements of performance, availability, and scalability. It is estimated that more than 90 percent of the Global 2000 companies make some use of Fibre Channel SAN for their mission critical application environments.

Fibre Channel SANs are, in many ways, the "easy" choice. They provide an environment that you don't have to worry about, that continues to run without constant human interaction or "care and feeding." This doesn't mean you don't have to interact with it at all; you will add or remove devices from any network in the normal course of activity. But you won't have to piece things together painstakingly one connection at a time.

Ethernet versus Fibre Channel Setup

One way Fibre Channel makes things easier is in the initial setup of the network. This section offers a couple of scenarios to compare Ethernet and Fibre Channel. How does each one stack up in terms of ease of setup? Fair warning: This section is a bit technical, but it's necessary in order to explain the concepts.

Ethernet setup

Suppose you are connecting a new Ethernet switch to an existing network through multiple ports. You have to configure the ports individually on both ends of each connection. This configuration uses 40 to 90 commands to establish. And you have to take particular care not to make a mistake in the configuration because that can cause the network to hang and fail applications.

Okay, but it's no big deal right? You only have to do it once. Well, yes, unless you add something new or a link fails between switches. The network then needs to "converge" or reform, which can take seconds to achieve even though the storage normally answers in a sub-millisecond time frame. That extra delay can cause applications to perform poorly or crash. And that's just for the physical connection. There are also the software configurations and services elements to be configured separately.

Now imagine that you're creating a ten-switch Ethernet environment. These same commands have to be done for every link made between switches and for configuring the switches themselves. In order to avoid major failures, you also have to concern yourself with the topology that the interconnects form.

Fibre Channel fabric setup

In a Fibre Channel fabric, the level of setup automation available is much greater. Think again about the scenario of establishing a ten-switch Fibre Channel fabric. After you put the switches into racks, you can freely connect the switches in any topology you choose: a ring, a spine leaf, or a full mesh, with multiple links between any or all switches.

Then you power the switches on. As the switches come up, the switch with the lowest World Wide Name becomes the principal switch in the configuration. This switch now has the responsibility for assigning the domain ID for all of the other switches. The entire fabric forms automatically. All paths between switches are up and available, and nobody spent hours configuring every switch and connection.

Next you add servers and storage to the fabric. So that's going to take a lot of time and effort, right? Well, yes and no. In the Fibre Channel world, as a device gets connected, the host bus adapter (HBA) — the equivalent of a network interface card or NIC in Ethernet — goes through an automated handshake with the switch. It achieves a fabric login and is assigned a fabric address (which is like a phone number) that is guaranteed to be unique to the fabric. More importantly, that device will be automatically registered to the name server in the switch (you can think of the name server as a phone directory). Conversely, in an Ethernet environment, that normally includes registering the storage elements to the upper-level storage environment — for example, Internet Small Computer Systems Interface (iSCSI).

REMEMBER

The big benefit with that setup method is that there is no broadcast traffic. The name server lives in the switch, and every switch keeps a copy of the name server. In other words, the switches in a Fibre Channel fabric maintain distributed copies of the phone book.

When two devices are initially connected, an algorithm called Fabric Shortest Path First (FSPF) figures out the best route. It is extremely fast and doesn't generate a lot of network activity. The route between the devices, once established, is used for all traffic between them from that point forward. Having a single route guarantees the "in-order delivery" of the data, which is critical because applications do not cope well with data that shows up out of order.

Adding new switches is similarly automated. As the link comes up, the new switch gets its area code (domain ID) and its copy of the phone book (name service) from the existing fabric, along with other elements not discussed yet, all without human intervention.

In a Fibre Channel fabric, the fabric owns the responsibility for the traffic to get to its destination. In contrast, in an equivalent Ethernet network, it's the responsibility of the end devices to care if data arrived and to correct the situation if it didn't.

In a Fibre Channel fabric, the connection level protocol uses something called a *buffer-to-buffer credit* (BB_credit). The sending device does not put any data on the wire unless it knows that the other end of the link has space to take it. In comparison, Ethernet is a "drop and run" technology, where the traffic is sent without knowing whether or not the receiver (or any intermediate waypoint) has the capacity to accept it.

Delivering the data

Here's a metaphor to help you put all that technical detail you just read in perspective. Imagine that you have 50 people at your work location who need to attend a training event in another city 100 kilometers away. It's important that they all attend, because each of them represents a different group in your company. As they leave your building, you give each person a sequential number as they go to their cars. However, each person is responsible for driving himself or herself, and they may end up taking different routes.

Because of different traffic levels or congestion and the timing of traffic lights, the attendees don't show up in the same order, so the receptionist has to put them back in order by number. Worse, three of them couldn't find parking at the training location, so they never arrived. The training facility then has to have someone drive back to your company, and you send a replacement attendee from each of those departments to the training center before the class can begin. That's effectively how TCP/IP handles data delivery.

In contrast, in a Fibre Channel fabric, none of the people would be sent to the training before determining that they had space to park at the training center. The FSPF routing software would ensure that all of the attendees followed the same route to the training center and therefore arrived in order.

Should a failure cause a connectivity issue, because the Fibre Channel fabric knows all of the topology, there is no equivalent "convergence" time. If a known good path exists, the traffic is immediately routed over it, and path recovery is much faster. Although TCP/IP over Ethernet recovers any data that Ethernet drops, Fibre Channel is designed not to drop the data in the first place.

This scenario also has implications at the application level when data delivery must be both reliable and consistent for processing time. For example, in some financial applications, it is important for all of the transactions to have the same "processing time" within an allowed range. The idea that some transactions may "eventually" be completed outside of that allowed range may or may not be acceptable as a performance characteristic or at a regulatory level. Some data center designs specifically over-provision their environments not to exceed 50 percent utilization of the resources. Although this method may or may not be acceptable financially, it begs the question of what happens when a network failure causes traffic to significantly exceed those plans.

REMEMBER

These are just some of the items behind the idea that the "easy" alternative is the one where:

- The infrastructure is not problematic.
- Fabric services handles auto-configuration without a high degree of human intervention.
- Traffic is handled in a consistent fashion and failures are recovered rapidly.

Protecting the Data

One of the long-standing concepts in the data center is that it's "all about the data." After all, it isn't called the server center or the network center. Data is paramount. The ability to share it, process it, protect it, and retrieve it are all considered basic to the environment. This is also true in a Fibre Channel SAN.

Double the fabric, double the protection

REMEMBER

Fibre Channel fabrics are always deployed in dual-redundant, hardware-isolated fabrics. That statement often causes an immediate reaction in some people: "That's *double* the cost of an Ethernet network!"

What they don't understand is that in data center Ethernet, almost all servers have dual Ethernet connections (NIC teaming) where each NIC is connected to a different switch at the top of the rack. So you're doubling up on hardware anyway. The primary difference is that although both of the top-of-rack switches that Ethernet connects to will connect to the same overall Ethernet network, the dual HBAs in the server will connect to different Fibre Channel fabrics.

Why this isolation? Because it only takes one human to take down a network. Studies show that more than 30 percent of unplanned downtime in any data center is due to human error. Although it is certainly possible for a single network administrator to cause both Fibre Channel fabrics to go down at the same time, the standard practice is to make changes and updates to one fabric at a time. This process generally ensures that the administrator will see the issues caused by the mistake made in the first fabric and not create the same failure in the alternate fabric, whereas a single bad network configuration can cause outages to both of the network connections from the server.

This same premise is true for software driver updates, as well as firmware updates to the cards in the servers or the switches in the fabric. The administrator makes changes to one of the two interfaces and then waits to verify clean operation before continuing to make the change to the alternate interface. As a result, Fibre Channel network updates carry much less risk of operational impact.

This benefit also applies when updating the fabric elements. In normal data center Ethernet, administrators commonly implement the next generation of new switch technology when the next rack of servers is implemented. Thus, network hardware upgrades usually occur only when server infrastructure is added or replaced. This is one of the reasons why so much of data center Ethernet is still using old technology. Fibre Channel fabrics, in contrast, use the redundant fabric design to update one fabric while the other continues running production. This allows the Fibre Channel environment to stay current in technology with little or no impact to production.

TECHNICAL STUFF

What is "old technology"? Most customers and market analysts report that less than 30 percent of their current networks support line rate 10 Gigabit Ethernet (10GE), almost 19 years after 10GE first appeared in the IT market. In that same intervening time frame, Fibre Channel fabrics have gone from 1 gigabit per second (Gb/s) to 2 Gb/s, then 4 Gb/s, 8 Gb/s, 16 Gb/s, 32 Gb/s, and soon to 64 Gb/s.

Virtual LANs: Do not disturb

Ethernet and Fibre Channel share the idea that not every device in the network should be able to see, talk to, or disturb every other device in the network. When networks were small, this was not a major issue, but as network sizes grew, the amount of broadcast traffic and disruption by bad network citizens became a serious problem. From this need, Ethernet first created virtual local area networks (VLANs), a mechanism where you can allow smaller numbers of devices to see each other.

In Fibre Channel, this mechanism is known as *zoning.* In a Fibre Channel fabric, a zoning service is distributed across all switches. Every new switch automatically learns this security definition for the fabric when it joins before any traffic is allowed to transit the switch.

Zoning comes in very handy because not every operating system supports the same features and functions, and because some early operating systems' environments are aggressive about storage. Anything they can see, they attempt to own. It's like having a 2-year-old wandering around the house; the breakables begin to move to higher and higher "safe" locations.

Storage as a shared resource

In a storage network, the goal is to have storage as a shared resource. Any storage that is not attached through a network becomes a stranded resource and a long-term data management issue. At a minimum, you need some remote storage to keep a "secure copy" of your data for recovery. Anybody who has ever dropped their smartphone in water and lost pictures that hadn't been backed up knows why this is critical. However, some environments are not quite so resilient. For example, if a network outage occurs and is recovered in an iSCSI environment, all of the storage elements must be manually re-registered. In the Fibre Channel SAN, however, if a link is lost and then recovered, there is no need to re-register devices. The route and the storage port simply resume their business.

Peer zones: The best of both

How then do you arrange for the storage element to be shared but for the elements doing the sharing not to know about each other?

In Ethernet this is generally done in small groups. (There may be networks where VLANs are done on a one-to-one basis, but you would hit the limit of 4K VLANs pretty quickly.)

In Fibre Channel, the best practice had traditionally been to create a private connection between every server and its storage. However, a few years ago this practice was superseded by the development of peer zones. For example, say a storage port on an all-flash array is being used to provision storage to 30 servers. Instead of a single zone definition for each server-to-storage pair, there would be a single peer zone containing all 31 devices.

REMEMBER

The difference between this configuration and an Ethernet VLAN with all 31 devices is that in Ethernet all of the devices would see each other; broadcast traffic from any one device would hit them all. In a Fibre Channel peer zone, however, the storage port(s) are defined as a principal. The principal sees all of the elements in the peer zone. However, the peers can only see the principal. No peer sees another peer, and no traffic from any peer if visible to any other peer, not even broadcast traffic.

Another element available in Fibre Channel fabric is Target Driven Zoning. In Target Driven Zoning, the same storage configuration used to provision storage to each server is also used to create a peer zone. There is no need for the storage administrator to access the Fibre Channel management console at all. One could reasonably make an argument that not having to log into the management of the network constitutes "easy."

Balancing Measurement and Performance

Network management without the ability to measure network performance is like driving with a blindfold on. Do you speed up? Slow down? You have no notion of what you should be doing. If you can't measure, you can't manage.

Networks have traditionally had to balance inspection and measurement against the possibility of reducing performance. Because the control processor of the switch handled the measurement, too much load could have unfortunate consequences (including having the switch hang and reboot). As a result, network administrators would most frequently resort to taking small samples and hoping their samples were representative of the overall trends. On a high-end core platform, the default mechanism might be to measure 1 packet in 8,000. An aggressive setting might be 1 packet in 2,000.

Granted, the expectation is that one can model the network behavior based on this statistical sample rate over a period of time. That is basically true, but the data is historical, not immediate. It's true in the same way that the Department of Transportation can model traffic for a highway over the course of a year. It's useful, but in my morning commute I'm more concerned with individual incidents. How do I see what's going on right now, and avoid problems?

REMEMBER

In Fibre Channel, a fabric measurement is granular. Every frame on every port is counted — and not just for total throughput. Because every frame header tells you the source and destination, it is possible to track many thousands of "flows" at a frame level. Fibre Channel technology is capable of collecting performance data on every port. Every frame is accounted for, allowing instant detection and granular identification of issues in the fabric.

This kind of visibility and granularity is vital to modern storage fabrics. As the latencies of the storage devices and servers continue to drop with flash and NVMe, it's going to become even more critical.

Problem? What problem?

A running joke among data center server and storage admins is that whenever something goes wrong and they ask the network team if there is a problem in the network, the response they get is, "We don't see a problem in the network." That may very well be a true statement. They aren't saying that there isn't a problem in the network, but that *they* don't see the problem in the network. If the inspection window is wider or broader than the event duration, how will you ever see the event?

Think about attempting to take a picture of a lightning bolt with a standard camera. Will your camera be pointed in the right direction and will the shutter be open when the lightning bolt strikes? Now apply this to your network. Will your monitor see the issue in that exact second? This concern becomes more and more critical as application and storage element performances increase and latencies decline. If you don't know when a problem is going to hit, how can you plan to measure it? And — just as critically — automated remediation of these issues becomes paramount. As the technologies progress, humans are simply not fast enough to respond any longer.

Gathering deeper info

Another development in the last few years is the capability for hardware-level deep packet inspection.

Deep packet inspection goes beyond handling the forwarding of traffic based on addressing. It looks further into the packet to see what some of the other traffic characteristics are.

As with measurement, implementing deep packet inspection has traditionally been balanced against performance. In Ethernet, some percentage of traffic may get forwarded to a separate processor for such inspection, but this is not routinely done for all traffic because of the overhead and latency involved. Additionally, if the entire packet is mirrored, it poses the security risk that the data being mirrored is available to people who are not permitted to see it.

However, in a Fibre Channel fabric, it is possible to look at the operational commands of the traffic, not the actual data transported. This capability allows statistical analysis of the read and write commands on a flow. Which storage target is the recipient? What was the response time to the first request? What was the completion time? How many outstanding or pending IOs are there?

Sometimes one element's traffic or behavior affects another element that is not obviously related to it. A common term used to describe this effect is "slow drain." The implication is that one slower device is potentially causing a network congestion. This might be a temporary issue on the part of the device or software, or a temporary change in the workload. It also might be a longer-term impact of having a mismatch of technologies in the same shared network (a much older and slower platform, for example). It is like having somebody in the high-speed lane of the freeway driving at a slower rate than the traffic is capable of.

The main mechanism that most network technologies use to deal with this is to drop the slow drain device traffic. That approach presents two problems:

- How does one know that the traffic being dropped is not mission critical? The loss of such traffic might be catastrophic.
- The behavior of the device in question has not changed. As soon as it starts transmitting data again, the same behavior will likely reoccur.

REMEMBER

Fibre Channel fabrics can use the granular latency measurement previously discussed to identify this kind of slow traffic behavior. Once this flow is identified, the SAN can use a function called Slow Drain Device Quarantine (SDDQ) to solve the issue.

Here's how that works. The links between switches in the Fibre Channel SAN, which are known as Inter-Switch Links (ISLs), have multiple virtual channels (VCs) on them. Imagine this as multiple-lane bridges over a river. When a slower flow is detected, the fabric software moves the flow to a low-priority lane, allowing the faster performing traffic to continue to flow unimpeded, and without dropping any data. This type of non-disruptive solution is another example of the difference between a network and a fabric in terms of the types of functions and exception handling that one would expect to find implemented.

Getting virtual visibility

In Ethernet network environments, visibility into the performance of an individual virtual machine is, for the most part, left to the hypervisor platform. There is no low-overhead mechanism for tagging all the traffic of individual virtual machines because the mechanism requires forwarding traffic to a separate service processor. Over the past 10 to 15 years, server virtualization has complicated the issue of visibility. When a single server hardware port supports 24 or 48 virtual machines (VMs) or more, how do you monitor traffic?

Hypervisor administrators frequently say that the first two or three application owners to complain about a performance problem are not usually the problem children; they are the sensitive neighbors in the apartment complex who can't stand the loud music. How do you determine which application is the victim and which application is the bully? In many environments, the default behavior has become to migrate the VM that complained in the hope that you are moving him away from the bully.

However, the Gen 6 Fibre Channel (32Gb/128Gb FC) standard includes a new mechanism called *VMID tagging.* It tags individual Fibre Channel frames with the ID of the VM actually creating the traffic so you can tell who is causing the problem.

Avoiding Disaster

Everybody wants to avoid a disaster. But as discussed early on, different applications and different data sets have different requirements for how their disasters should be treated. How much, if any, data can be lost (recovery point objective or RPO)? How much, if any, downtime can be tolerated (recovery time objective or RTO)? Are any regulatory or compliance issues involved? Are there data retention periods? Disaster recovery is a far more complex question than most people realize. And that's just understanding what your data issues are, let alone fixing them.

Disaster avoidance is a term that one has heard more frequently in the past several years. One used to hear *disaster recovery* or even *disaster tolerance* much more often. Those are still topics. But more and more commonly, infrastructure and application teams talk about avoiding a disaster in the first place, not just moving past it when it occurs. Although avoidance isn't always feasible, good planning can make it simpler and more potentially achievable.

REMEMBER

Disaster planning, like so many things, has always been a balancing act among equipment, people, and money, based on how big a disaster you were trying to plan for. One size does not necessarily fit all needs. Consider these questions:

- Are you planning to avoid a simple human error? If so, you may only need a smaller level of redundancy in applications and power with multiple instances in the same building.
- Are you trying to avoid a building event (a power outage or fire, for example)? If so, perhaps you need two data center rooms in different buildings on the same campus.
- Are you more concerned about a local proximity situation such as a power outage to your entire campus? If so, consider separating your environments across a metropolitan area.
- Are you seeking to protect against a large-scale weather event or natural disaster (a hurricane, tornado, earthquake, blizzard, or flood)? If so, you may need a multi-state or cross-country implementation.
- Is your corporation multinational? If so, you might have data centers in varying countries that have different compliance issues or requirements.

As you can imagine, the range of implementations can vary drastically.

One advantage that networked or shared storage has in some of these scenarios is the ability to connect two geographically separate data centers with the same functional environment. In the Fibre Channel fabric, it is not uncommon for infrastructure teams to stretch fabrics across a metropolitan region anywhere from one to ten miles or more.

Comparing replication types

Because the speed of light in fiber optic networks is not the speed of light in a vacuum, keep in mind that it adds roughly 5 microseconds (millionths of a second) for every kilometer of fiber that the light travels. For example, a 20-kilometer distance adds roughly 100 microseconds to the latency — which many applications can tolerate. It is therefore possible to provide synchronous replication across such a network.

The term *synchronous replication* refers to the idea that a data write is not acknowledged to the local application until the write to the remote location is verified. In many banking environments, for example, a transaction should not commit to one side if it doesn't commit to the other. Synchronous replication guarantees consistency between the two sites. There will be no data discrepancy or loss if one site is lost and the other survives.

For less mission critical applications, *asynchronous replication* may be sufficient. With asynchronous replication, the local write is acknowledged to the application as soon as it's made, and the remote write is made as quickly as possible afterward. An upside is that the local application will likely be somewhat faster. A downside is that if something happens to the primary site before the data is replicated to the remote site, that data may be lost. Some businesses and applications can sustain such a loss with little or no issue, but many cannot. Therefore, the choice of maintaining a synchronous mirror or an asynchronous mirror depends on the business criticality of the data. Not all data has the same business value.

Scrutinizing security

An advantage of Fibre Channel has to do with the security domains used in its deployment. First, the server-to-storage traffic is in dedicated zones that are hardware-enforced, ensuring that other devices cannot connect to those ports. Second, when storage is assigned to a server, the server can only see the storage element/data that is assigned to it, as compared with a hyperconverged infrastructure (HCI) solution on Ethernet, where the server is also a storage node for other devices and has data local to it that doesn't belong to it. As a consequence, if a server in Fibre Channel is compromised by a hacker, only its data is at risk, while the same hacker has access to more than just the compromised server in HCI.

Additionally, because the alternative network technologies invariably deploy in a shared use case where storage traffic and application access are on the same ports, the environment is susceptible to all of the Ethernet hacks (including eavesdropping) that may hit the main network. For better or worse, Ethernet has the dubious distinction of being the most hacked protocol on the planet. Vast libraries and entire websites are dedicated to Ethernet and TCP/IP hacking suites. Fibre Channel has no such burden to bear.

REMEMBER THE RULES

For a bit of educational fun, this book's author, AJ Casamento, would like to share some rules he has learned from his IT customers over the past 40 years. Here are "AJ's Rules for IT" (they may save your job!):

1. **The most powerful force in IT is inertia.** The business's first choice is always "do nothing."
2. **Software is the hard part.** Software should never have been called "software." It should have been called "difficult-ware" or "not-as-easy-as-you-think-ware" or "greatly-underestimated-resource-and-schedule-ware."
3. **IT people hate to touch a running system.** Why? It's running. If you touch a running system, of the three potential outcomes, two of them are not as good as running.
4. **If you can't measure, you can't manage.** Without measurements, you're driving with a blindfold on. Do you speed up? Slow down? Turn right? Turn left?

(continued)

(continued)

5. **All hardware eventually breaks, and most software eventually works.** No matter who makes the product, it eventually ages and fails. No matter who writes the software, it always gets patched.
6. **No matter how good you are, you are not faster than the Enter key.** Studies have shown that the largest segment of unplanned downtime in any data center is due to human error. Even if you realize the error as soon as you make it, the command has already been sent.
7. **Applications expand to consume all the performance that the infrastructure provides.** With every improvement in technology, speed, capacity, or performance, applications evolve to consume it all.
8. **Failures fall into three categories:**
   - No immediate reaction to the failure
   - Yelling and screaming
   - Losing your job (pay special attention to this category!)
9. **Management by magazine is a problem.** You can substitute "blog," "vBlog," or "article" for "magazine." When an implementation is mandated by someone who doesn't understand the details, things can go horribly wrong.
10. **Application users have many positive attributes; patience isn't one of them.** Whether the customer is internal or external, try to architect so that infrastructure failures don't cause application outages.
11. **There is no good way to become famous in IT.** As the old saying goes, "Everybody gets their 15 minutes of fame." However, nobody ever promised you that it would be *good* fame, did they?

IN THIS CHAPTER

» **Understanding storage needs**

» **Planning your next-gen storage fabric**

» **Finding balance**

Chapter 4
Putting It All Together

In this chapter, you get a closer look at how to plan for your storage needs and the impact of new developments such as Non-Volatile Memory Express (NVMe), the up-and-coming storage technology. You also learn how to start planning your next-generation storage fabric and get a few tips for finding the appropriate balance for your company's needs.

Understanding Storage Needs

It's important not to take storage services and design for granted. You likely deal with large amounts of stored data (even if it's just gigabytes of photos on your smartphone), and you can easily begin to think of it as background noise. Nothing could be further from the truth. Losing *network* connectivity to your laptop, tablet, or phone is annoying, but you can continue to work with local content and your device doesn't reboot. However, if you lose connectivity to your *storage*, the device becomes unusable and reboots when it comes back — meaning you lose whatever you were working on.

REMEMBER

Those are two very different impacts. The same is true with your data center and its applications. A temporary network outage may be inconvenient and expensive, but an outage in your storage connectivity can be disastrous.

Historically, this issue has been a focus of the Fibre Channel SAN designs. "Uptime all the time!" is a mantra in the industry. The Redundant Array of Independent Disks (RAID) market has worked hard to eliminate downtime although, to a degree, RAID has been a victim of its own success because many people have lost track of or forgotten the elements built into the array technologies over the years.

For example, when a hot-spot traffic pattern appears as a result of different applications attempting to access data from the same disks, the array rebalances the workloads in the background by separating those data sets to different disks. Or, when a device is seeing a high number of *soft errors* (meaning that the error occurs but is recovered without the application knowing), which can be an indicator that the disk is about to fail, the array rebuilds the data set in the background on a spare disk and marks the failing storage media for replacement.

TIP

Contrast this with some of the more recent marketing attempts to replace SAN technology, such as hyperconverged infrastructure (HCI). Although marketing documents describe how the platform has "dedicated processor, memory, storage and network resources" in an attempt to make this environment sound like something new, it is architecturally no different from any stand-alone server that you would have found in the data center in 1995. All of the standard caveats exist around eventually needing to scale either processor/memory or storage capacity independently of one another. The challenge is that when you need more storage, you have to add the entire HCI platform, which means you add more processor capacity and memory capacity even if you don't need it. Additionally, that extra storage is now accessed across the network and typically through a more congested connection.

HCI or software-defined storage (SDS) generally expects all participating nodes to be uniform. This issue is important because the software that distributes the data or the workload treats all of the devices the same. It expects each device to have the same storage capacity, storage performance, processor performance, and so on. It doesn't have the ability to treat the devices differently. In fact, the vendors of these platforms usually recommend that if you want to add newer, bigger, higher-performing systems, you run them as separate clusters. Even if they are the same configuration, adding resources causes a rebalance of capacity and workload across the network where the current work is running.

This also affects your ability to troubleshoot the environment. Which of the distributed storage is the one creating the problem? The node where the primary copy is? The node where the secondary copy is? Is it a network issue in between? The problem now has many more potential locations than in consolidated shared storage.

Virtually every HCI or SDS offering has a best practice (when performance is the goal) to isolate the storage traffic from the standard network traffic. This is in part because otherwise the non-storage traffic tends to affect the performance of the storage traffic, but also because when something must be recovered in the storage environment, the spike in traffic can be problematic for the non-storage application traffic.

TIP

Contrast the performance and the problem isolation issues that HCI and SDS are both faced with. The Fibre Channel SAN solution does not have this issue because the storage traffic is already isolated into a separate network from the application traffic. The standard dual-redundant hardware-isolated SAN fabrics not only guarantee better performance but also easier problem identification.

With the advent of Non-Volatile Memory Express (NVMe) these issues become even more difficult because the speed of the new technology is such that there is even less forgiveness for network problems. Imagine that a standard hard disk drive (HDD) or spinning disk is a line of cars traveling at 60 mph with a 30-foot gap between vehicles. Moving to solid state drives (SSD) reduces the gap to about 3 feet. Moving to NVMe reduces that gap still further to about 6 inches. Your network needs to be better, monitoring needs to be better, and control needs to be better, because reaction time for problems is shrinking.

Another consideration regarding the speed of NVMe is that as performance increases, you'll want to consolidate more applications to the same hardware. You're going to have the ability to host even more virtual machines (VMs) on every server. But remember that you're working with an ecosystem. If you don't have the performance connection to the server for your network, you can't take advantage of the additional performance.

The hardware itself is not all of the NVMe story, nor even the part with the biggest potential impact overall. NVMe is also a language interface. For the past 30 years or so, the IT industry has been using Small Computer System Interface (SCSI), or some form of it, to talk to storage. Now, for the first time, that is changing.

The older SCSI interfaces will still be around for years. Certainly, most companies don't have the luxury of simply throwing away existing assets. But moving forward, you will see more and more NVMe coming into production storage environments. Therefore, it makes sense that the two technologies will coexist in your data center. At the same time, you'll want to achieve the benefits of the newer technology. That requires careful planning.

TECHNICAL STUFF

Where SCSI has more than 200 commands in its lexicon, NVMe has just 13 required commands and 25 optional ones at the moment. This makes for a much lighter-weight software stack. Early testing shows a 30–35 percent reduction in overhead on servers when using the NVMe protocol. This is a potentially huge win for IT given the ratio of servers to storage in the environment (that is, many servers to few storage elements).

Planning Your Next-Gen Storage Fabric

Now that you have some idea of the scope of the technology change that the storage market is about to undergo, how do you plan for it? What are the characteristics that are going to get you the most benefit for your money in supporting the business and the application base? Those are questions that drive the intelligent design of the next-gen storage fabric.

Easy deployment

For a start, you don't have unlimited time, people, and money. Therefore, a storage fabric should be easy for you to deploy. A Fibre Channel SAN fabric has several advantages in this regard over an Ethernet network.

The first advantage is building the network of switches. With a Fibre Channel SAN you put the switches into their racks, you connect the cables between them, and you power them on — only three steps. Contrast this to Ethernet where you need tens to hundreds of commands on each switch to do the same thing. The process takes minutes with Fibre Channel versus hours with Ethernet.

A second advantage is the addition and registration of devices to the network. In a Fibre Channel SAN fabric, as you connect servers and storage elements to the network, they are automatically registered to the distributed name server and given fabric addresses. Every switch knows every device in the fabric. If one switch fails, as long as there is still a valid path from the server to its storage, the path is recovered immediately. This behavior offers an advantage over Ethernet, where the failure of a switch or a link can cause the entire network to reconfigure. In addition, for certain Ethernet storage protocols such as Internet Small Computer System Interface (iSCSI), you have to manually re-register the devices when they reconnect.

A third advantage is that if you're running either a Gen 5 (16 Gb/s) or Gen 6 (32 Gb/s) Fibre Channel SAN fabric, you can simply connect NVMe devices to the existing SAN and run. The infrastructure is already "good to go" to carry NVMe traffic without a rip-and-replace of the switch elements. Just as importantly, you can seamlessly run both the existing SCSI traffic and the new NVMe traffic on the same server, and even on the same port of the network, without any issues, and still take full advantage of the performance. Contrast that with Ethernet, where the performance implementation of NVMe requires new switches and network interface cards (NICs), necessitating a disruptive upgrade.

Reliability and availability

Any storage outage can become a major issue. Therefore, the reliability of the storage fabric is critical. Ethernet attempts to assure reliability by using dual NICs in the server. However, these ports both connect to the same overall network, though generally not on the first switch. The Fibre Channel SAN fabric, on the other hand, uses dual, redundant, hardware-isolated SAN fabrics. The advantage is that a single error in a configuration doesn't drop both connections to the application server as it might in Ethernet.

Another advantage is that while Ethernet depends upon the end devices to know if data was lost in transmission and to recover it, Fibre Channel depends upon the network itself. The Fibre Channel SAN fabric never forwards data on a link without knowing in advance that there is space for the data to be received. Contrast that with Ethernet — there is no awareness of available space for the data; therefore you are dependent upon the receiving device to realize that data has been dropped and to ask for a retransmission. The retransmitted data adds to traffic congestion, potentially causing additional data loss and retransmission. This problem is so prevalent in Ethernet that an entire software algorithm has been created to avoid it (TCP Slow Start). The side effect is that by avoiding this issue, you don't get full link speed out of the Ethernet connection. Thus, you lose out on performance.

Another issue exists in HCI and SDS environments where the server administrators are not accustomed to being the storage provider for other servers. If a server administrator needs to apply a security patch to the operating system, he or she simply does it and reboots the server. However, if that server contains a storage element that is serving another system, that has additional implications. The Fibre Channel SAN fabric, through the use of shared and consolidated storage services, doesn't face the same rate of service windows as the distributed storage systems. As a consequence, the availability of the system is much higher.

Performance

The one constant in any IT environment is that new application development consumes the resources that the infrastructure deploys. This may take some time to happen, but historically it is inevitable. New applications have more functionality, become more feature-rich, and ever-higher performing. However, it's also true that every generation of servers and storage increases performance. With the inclusion of NVMe in the storage mix, both the capacity of the storage devices and the performance of the storage devices is increasing drastically.

An outcome of this pattern is the desire to consolidate even more applications to the storage element. That scenario is problematic in an HCI or SDS deployment because the storage elements scattered among a number of servers don't readily allow full utilization. The applications on the single server may be able to get local performance to a degree, but then when data is sent off the platform, either to provide a secure copy, to rebalance workloads, or to acquire additional storage capacity, the Ethernet network connectivity doesn't provide the ability to sustain line rate on the port. In contrast, the Fibre Channel SAN fabric allows line rate utilization, which means that you can achieve even greater density of applications and get the full performance out of your new storage technologies.

Manageability

If you can't measure, you can't manage! That doesn't imply that Ethernet doesn't have management utilities; hundreds of choices are available for Ethernet management. And in both environments, you can use application programming interfaces (APIs) to help create automated processes.

One of the big differences is that while Ethernet traditionally relies upon sampling the traffic performance, the Fibre Channel SAN fabric measures every frame on every port. It also can measure the latency of the traffic on every port at a rate of 400,000 times per second. This capability provides a level of granularity and visibility that becomes crucial to the next-gen storage fabric.

As the scale and the performance of the storage continue to increase, regardless of which technology you choose, the scope of the potential problems becomes greater too. As a customer recently phrased it, "When you have a few children calmly walking about it's easy to spot the problems. When you have a lot of children running about it's an entirely different scale problem!" How do you quickly identify the problem? How do you automate the correction of the problem or, at a minimum, the mitigation of the problem? Where Ethernet struggles to find the element that is causing a slow drain or congestion problem, and then drops the traffic from the congested device to clear the traffic, the Fibre Channel SAN fabric moves the slow drain or congested device to a low-priority lane on the link between switches without losing any data. Isn't that the kind of automated correction that today's performance data center needs?

Security

The need for data security in today's world is clear. Data breaches and losses are almost a daily occurrence. And certainly you have many choices for security tools and technologies to deploy in an Ethernet network. However, other considerations are present in an Ethernet storage environment.

A storage device inside a server platform is owned, in a hierarchical sense, by the server. This arrangement potentially provides back-door access to data that the administrator of that server shouldn't have because the physical system and management of the device belong to the server administrator, and the "sharing" mechanism of HCI or SDS is a higher-level application.

In contrast, the Fibre Channel SAN fabric assigns a portion of storage capacity to a server, and only that server can access that data. Another server in the same network has no access to any part of the storage not specifically assigned to it. Although this may seem like a basic premise, it is critical.

Finding Balance

Finding balance is a valid consideration. Any IT environment is an ecosystem, and nothing operates in a vacuum. The biggest, fastest CPU in the world waits for its data at exactly the same speed as any other system. In order for a server to perform at its best, all the characteristics must be in balance. CPU, memory, storage, and network all must perform in concert with one another. You may use multiple storage technologies in your solution sets, but it's important to make certain that the technology you choose is "fit for purpose" for the application base you are serving with it. And *that* is the balance that you need.

REMEMBER

Over the past 40 years, the industry has moved the bottleneck or choke point in this balance from one location to another. In the next-generation storage fabric, the correct choice is to make certain that the fabric is *not* the next choke point. That is what the modern Fibre Channel fabric does. It becomes the storage fabric of choice for reasons of reliability, scalability, manageability, and performance.

IN THIS CHAPTER

» Reviewing key facts about next-generation storage fabrics

» Understanding the technology factors involved

Chapter 5 Ten Takeaways

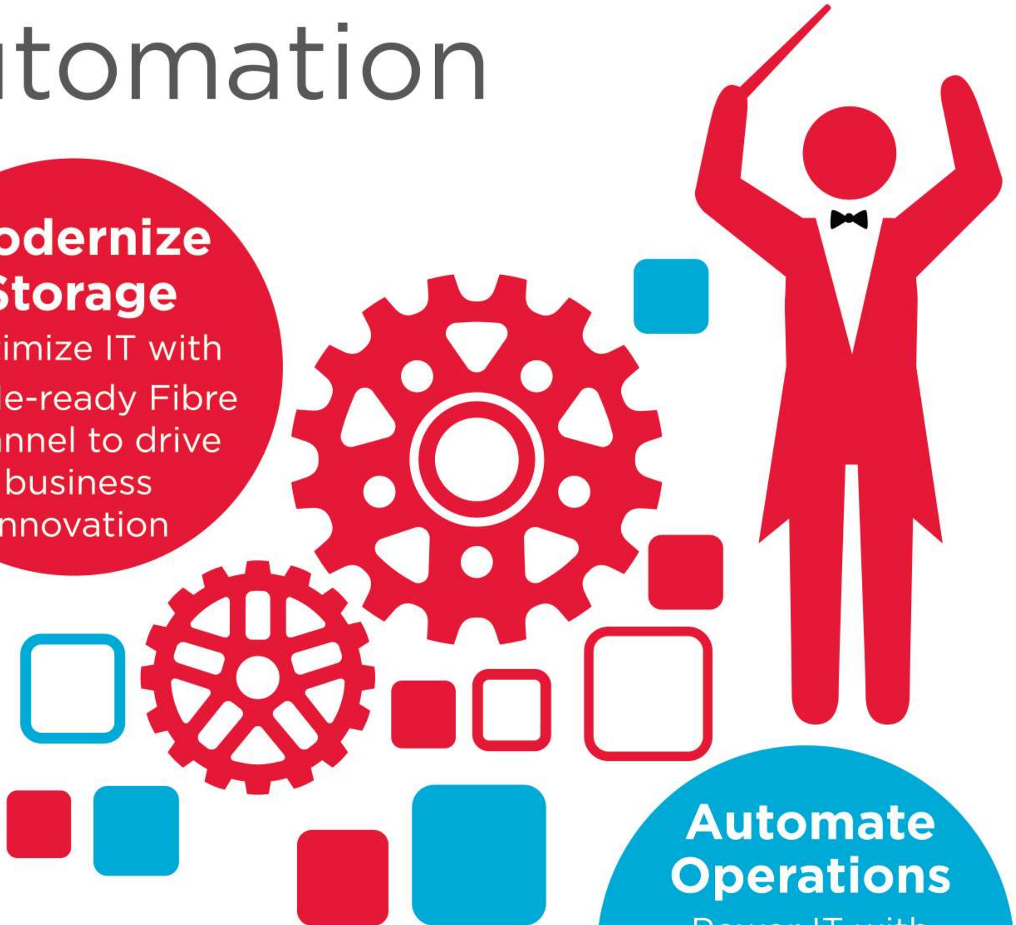
By reading this book, you've learned a lot in a very short time! Here's a quick review of the most important points to remember about next-generation storage fabrics:

- **Not all networks are fabrics.** Networks provide connectivity. Fabrics add services to the connectivity. Automated configuration, automatic registration, and adding bandwidth between switches without manual configuration are examples of why services matter.
- **Data loss is never okay.** No matter how difficult times get, no storage team was ever told that it's okay to lose some data. The current market saying is "Data is the new oil," meaning data is a precious resource with a quantifiable value. Your infrastructure must reflect that. When data is stored or moved in your organization, you should be able to rely on it being safe from loss.
- **Uptime should be all the time.** Windows are periods of time where something occurred that could potentially affect normal use. Decades ago, the industry used to talk about backup windows, service windows, and so on. Almost every environment today deals with some version of 24x7 for access time. There is never a good time for availability to be disrupted. This is one of the advantages of the dual-redundant hardware-isolated design style. You have a much higher availability and ability to do updates.

- **Technology innovation cycles are short.** The storage plan you make has to be based on silicon innovation cycles. How will the solutions you choose hold up as new technologies become available? The traditional four- or five-year technology cycle between storage speeds or network speeds is dropping. The new cycle is 18-24 months, which means you need to design to support multiple generations of storage arrays in the life cycle of a storage fabric.
- **Measurement is a must.** There is no way to manage something that you aren't measuring. As the performance of the infrastructure increases, the management needs to be even more comprehensive. Measurement needs to be granular enough to not only immediately recognize when an issue occurs, but to help avoid it to begin with.
- **Storage has needs.** When you're dealing with storage, you need to remember that the data being kept has requirements. There are expectations of the business line owners (such as performance) and can also include regulatory expectations. Know what your data's needs are and make certain that your fabric can meet them.
- **Mixed use is always a compromise.** Storage traffic and application traffic perform better when isolated.
- **Agility and scalability are key.** A next-generation storage fabric must be agile. The rate of change is going to continue in cycles of less than two years, so a fabric must accommodate multiple updates. It must be able to dynamically scale so that business needs can be met without downtime.
- **Security is a must.** Every new data breach proves how dire the consequences are when data is not secure.
- **Great networks are lossless, low latency, and deterministic.** A next-generation storage fabric should be: lossless to deliver it right the first time, low latency since applications hate waiting, deterministic to provide predictable performance, and *fast*!

With this information in mind, you're ready to make smart decisions about how storage fabrics fit into your business's IT plan.

Brocade® Modernizes Fibre Channel with NVMe and Automation



Modernize Storage

Optimize IT with NVMe-ready Fibre Channel to drive business innovation

Automate Operations

Power IT with simple and open automation to increase productivity

Accelerate business operations, respond to dynamic demands, and eliminate complexity

Learn More

go.broadcom.com/fc-networking

BROCADE
A Broadcom Company

Your guide to networking next-gen storage

This book provides an easy-to-understand introduction to storage networking. It explains what next-generation storage fabrics are, how they are commonly implemented, and what advantages they offer compared to less modern and less robust technologies. Armed with this information, you'll be able to make smart decisions about how storage fabrics fit into your business's IT plan.

Inside...

- Learn about data, storage, and networks
- Explore storage network types
- Increase customer satisfaction
- Balance measurement and performance
- Understand how Non-Volatile Memory Express (NVMe) improves performance
- Plan your next-gen storage fabric

HITACHI Inspire the Next BROCADE A Broadcom Company

AJ Casamento is a 40-year IT veteran with experience in a variety of market segments and IT companies as well as engineering and product development roles. In more than 21 years as a Solutions Architect at Brocade, a Broadcom Inc. Company, AJ helps enterprises understand technologies, architectural decisions, and their impact on the business.

Not For Resale

ISBN 978-1-394-15980-2
90000
9 781394 159802

WILEY END USER LICENSE AGREEMENT