

Compliments of:



Hewlett Packard
Enterprise

BROCADE[®]
A Broadcom Inc. Company

NVMe over Fibre Channel

for
dummies[®]
A Wiley Brand

Achieve low latency
and high reliability

Maintain mission-critical
storage SLAs

Reduce risk with
concurrent SCSI/NVMe



HPE/Brocade Special Edition

Rupin Mohan
Curt Beckmann
Kip Hanson

About Hewlett Packard Enterprise

Hewlett Packard Enterprise (HPE) is an industry leading technology company that enables customers to go further, faster. With the industry's most comprehensive portfolio, spanning the cloud to the data center to workplace applications, our technology and services help customers around the world make IT more efficient, more productive and more secure.

To find out more: [HPE.com/storage](https://www.hpe.com/storage)

About Brocade

Brocade, a Broadcom Inc. Company, is the leading provider of Fibre Channel backbones and switches. Brocade's commitment to innovation has enabled it to deliver the industry's first Gen 6 (32G) Fibre Channel fabric, as well as advanced silicon-based features like Fabric Vision, providing deep insight into networked storage performance. Today's Fibre Channel delivers robust, high-performance connectivity across the datacenter while also extending across the globe.

www.broadcom.com/brocade



NVMe over Fibre Channel

HPE/Brocade Special Edition

**by Rupin Mohan, Curt Beckmann
and Kip Hanson**

**for
dummies®**
A Wiley Brand

NVMe over Fibre Channel For Dummies®, HPE/Brocade Special Edition

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

Copyright © 2018 by John Wiley & Sons, Inc., Hoboken, New Jersey

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Trademarks: Wiley, For Dummies, the Dummies Man logo, Dummies.com, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. Brocade and the stylized B logo are among the trademarks of Brocade Communications Systems LLC. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

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 978-1-119-54112-7 (pbk), ISBN 978-1-119-54115-8 (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:

Project Editor: Martin V. Minner
Editorial Manager: Rev Mengle
Executive Editor: Katie Mohr
Business Development Representative: Karen Hattan

Production Editor:
Mohammed Zafar Ali
Brocade Contributor:
Marc Angelinovich

Table of Contents

INTRODUCTION	1
About This Book	1
Foolish Assumptions	2
Icons Used in This Book.....	2
CHAPTER 1: Exploring NVMe over Fibre Channel	3
Picking Sides: Is it Storage, or Is it Memory?	6
Mapping the dichotomy	6
On errors.....	8
Accelerating Access to Flash	9
Understanding How NVMe Relates to SCSI.....	9
Anticipating Future Benefits of NVMe over Fibre Channel	11
Invigorating Your Fabric	12
CHAPTER 2: Delivering Speed and Reliability with NVMe over Fibre Channel	13
Reviewing Fibre Channel's Place in the Storage Ecosystem.....	14
Evaluating Performance Metrics in Storage Context	14
Storage Metrics	16
Souping up Device Metrics	19
Achieving High Performance	20
Making Use of Enhanced Queuing.....	21
Realizing Reliability.....	22
Redundant networks and multipath IO	22
Features of a lossless network.....	23
Security.....	23
Good tools matter.....	24
CHAPTER 3: Adopting and Deploying NVMe over Fibre Channel	25
Identifying Your Situation.....	25
Considering Your Adoption Strategy	26
Protecting high-value assets.....	26
Allowing for a marathon shift.....	27

Exploiting Dual-protocol FCP and FC-NVMe.....	27
Zoning and name services	30
Discovery and NVMe over Fibre Channel	31
Familiarizing Yourself with NVMe over Fibre Channel.....	32
Experimenting in your lab.....	32
Migrating your LUN to a namespace.....	33
Transitioning to production.....	34
CHAPTER 4: Comparing Alternatives to NVMe over Fibre Channel.....	35
The Long and Short of RDMA	35
InfiniBand.....	37
iWARP	37
Yo, Rocky	39
Evaluating Ethernet-based NVMe.....	40
Commodity or Premium?.....	41
Smart shopping.....	42
CHAPTER 5: Ten NVMe over Fibre Channel Takeaways	43

Introduction

Unless you've been holed up in Siberia herding reindeer since the launch of Sputnik 1, you probably know that your kid's hand-me-down iPhone 3 could whup the Apollo 13 in both computing and storage capacity. And you likely are aware that the intense pace of innovation is continuing. Today's network speeds are not only millions of times faster, but carry millions of times more data per second. Processing speed has grown exponentially. As for storage, the contents of the entire Library of Congress can be squeezed into an affordable disk array. Times have indeed changed.

About This Book

This book, *NVMe over Fibre Channel for Dummies*, Brocade Special Edition, focuses on a relatively small but very important aspect of information technology. NVMe (Non-Volatile Memory Express) over Fibre Channel is a technology that touches computer memory, storage, and networking.

If you're a hardened computer geek, you've probably heard of it. If not, this won't be the last time. Like most of the IT world, NVMe over Fibre Channel enjoys a rich history and has evolved at break-neck speed, building on the capabilities of preceding technologies while avoiding the shortcomings of competitive ones. And for the uninitiated, this book introduces you to a technology that may very well be the next big thing in networked storage.

Simply put, NVMe over Fibre Channel has it all. It has the ultra-low latency needed for working memory applications, with the reliability that's critical to enterprise storage. Because Fibre Channel, as all network geeks know, is a premium datacenter network standard, NVMe over Fibre Channel is able to leverage fabric-based zoning and name services. Best of all, NVMe over Fibre Channel plays well with established Fibre Channel upper-layer protocols, enabling a low-risk transition from SCSI (short for Small Computer System Interface) to NVMe without the need to invest in experimental infrastructure.

Foolish Assumptions

Your knowledge of network and storage technology is likely well beyond that of your Aunt Mary, who calls every weekend asking for help with her Facebook account. If not, this book offers plenty of reminders and sidebars to guide you through the more difficult parts, and to help decipher the endless, confusing acronyms lurking around every corner of the IT world.

Icons Used in This Book

A number of helpful icons are scattered throughout *NVMe over Fibre Channel For Dummies*. These will reinforce and further explain important concepts, and keep you out of trouble with your boss.



TIP

Pay special attention to the Tips icons. They contain small bits of information that will make your job lots easier, and prevent having to order late-night pizza delivery because you're stuck in the server room reconfiguring a storage array.



REMEMBER

If you're one of those who spends his or her days reading thick hardware manuals, you're bound to forget things along the way. If so, the Remember icon might just be your new best friend.



WARNING

Computer hardware and networking equipment is expensive. Replacing it because you made the wrong decision is even more so. Heed the Warning icons if you want to avoid costly mistakes in your IT strategy.



TECHNICAL
STUFF

For you dedicated hardware professionals with framed photographs of Jack Kilby and Robert Metcalfe hanging in your cubicles, keep an eye out for the Technical Stuff icons; they're chock full of additional details on esoteric subjects.

- » Classic storage versus classic memory
- » Accelerating access to flash
- » How NVMe relates to SCSI
- » The future benefits of NVMe over Fibre Channel
- » Invigorating your Fabric

Chapter 1

Exploring NVMe over Fibre Channel

This chapter offers a brief tutorial on NVMe (Non-Volatile Memory Express) over Fibre Channel. We introduce (or reintroduce, for you veterans) a bunch of confusing acronyms, discuss the merits of solid-state storage and see how it compares to memory, and break down the component parts of this exciting, relatively new technology. We'd like you to understand why it might just be the best thing for your organization since pocket protectors.

NVMe over Fibre Channel is a full-featured, high-performance technology for NVMe-based fabric-attached enterprise storage, but is a no-compromise solution for NVMe working memory use cases as well. (We'll talk about how those use cases differ.) NVMe over Fibre Channel is relatively new, even though its component parts are not. Fibre Channel (FC) has been the leading enterprise storage networking technology since the mid-1990s. Speeds of 16Gbps (called Gen 5) are widespread. Gen 6 Fibre Channel became available in 2016, delivering twice the speed of Gen 5 and even offers a staggering 8x bandwidth on 128GFC links, and it is selling like hotcakes. Gen 7, which will double speeds yet again, is already on the horizon. FC is primarily used to carry the SCSI

(Small Computer System Interface) protocol, the historic leader for direct attached PC or server storage. SCSI on Fibre Channel is called, boldly enough, Fibre Channel Protocol (FCP).

NVMe refers to several related things:

- » An open collection of standards for accessing and managing nonvolatile memory (NVM), especially high-performance solid-state memories such as flash or 3D XPoint
- » That collection's primary specification (NVMe 1.3), which provides a common, high-performance interface for accessing NVM directly over PCI Express (see Figure 1-1)
- » A nonprofit corporation (www.nvmexpress.org) that works to develop and promote the standard, and is supported by a wide range of technology companies such as Intel, Dell/EMC, Samsung, Micron, and others

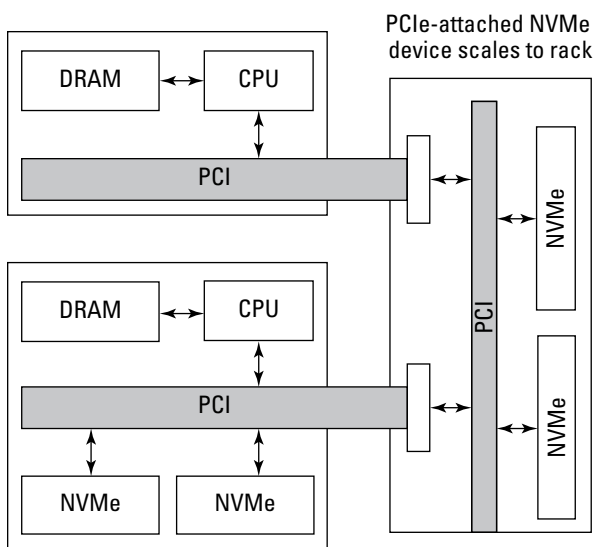


FIGURE 1-1: NVMe connects to a server PCIe bus internally or externally.

PCI-based NVMe has low latency, but it has important limitations relative to fabric-based media. The benefits of fabric connectivity include greater capacity, better capacity utilization, shared access, and advanced services (like deduplication, snapshots, replication

and backups). Using a fabric also eliminates single points of failure and simplifies management. To bring all these benefits to the NVMe ecosystem, NVMeExpress.org developed NVMe over Fabrics (aka NVMe-oF), which defines how NVMe commands can be transported across different fabrics in a consistent, fabric-independent way. (See Figure 1-2.) And that makes life a lot easier for software developers!

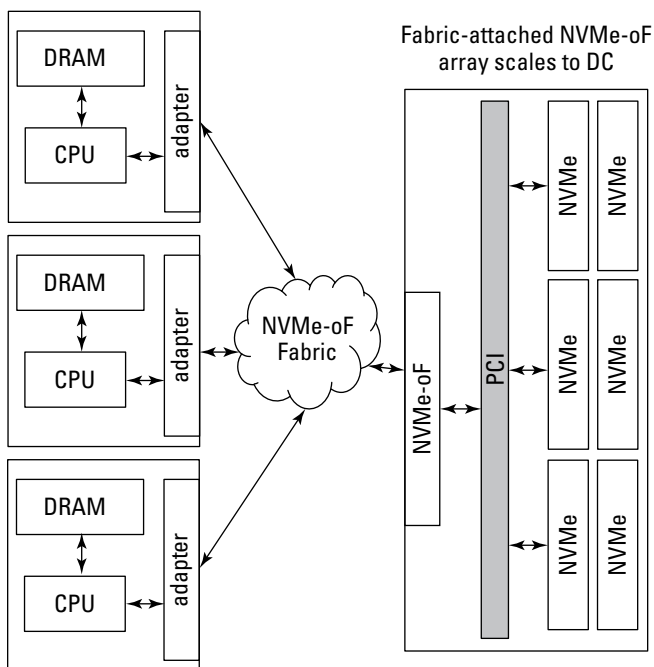


FIGURE 1-2: Using NVMe over Fabrics to increase the scale of NVMe.

The NVMe-oF 1.0 specification, released in 2016, described two fabric categories, Fibre Channel fabrics and Remote Direct Memory Access (RDMA) fabrics.

Fibre Channel fabrics – NVMeExpress.org chose the T11 standards body, which handles all Fibre Channel standards, to define the new “FC-NVMe” protocol. In mapping NVMe onto Fibre Channel, the T11 committee members followed in the footsteps of SCSI, making it straightforward to carry both SCSI and NVMe traffic on the same infrastructure. The T11 committee finished its work in October of 2017.

RDMA fabrics – RDMA is an established protocol that has run for years on InfiniBand, RoCE (pronounced “rocky”) and iWARP (we warned you about acronyms). Building on RDMA allowed NVMe-Express.org to target three existing fabric transports in one effort.

Note: In early 2017 (after NVMe-oF 1.0), a group at NVMeExpress.org advanced an effort to map NVMe over TCP (without RDMA) so NVMe-oF can run in existing data centers that lack RDMA support.



REMEMBER

Fibre Channel is capable of transporting multiple higher-level protocols concurrently, such as FCP and FC-NVMe (the label for the specific frame type of NVMe-over-Fibre Channel traffic), as well as FICON, a mainframe storage protocol. That bears repeating: FC-NVMe can coexist on your FC SAN and HBAs right along with your existing FCP or FICON traffic.

As you will see, NVMe over Fibre Channel offers robust interoperability, darned fast performance, and extremely scalable architecture. Whether you’re faced with a legacy storage network in need of an upgrade or a spanking new memory-centric implementation, NVMe over Fibre Channel offers a best-of-both-worlds solution while allowing a smooth transition for traditional users.

Picking Sides: Is it Storage, or Is it Memory?

Some may detect a hint of tension built into the phrase *NVMe over Fibre Channel*. That’s because FC is a storage-oriented technology while the term *NVM* is obviously memory-oriented. By contrast, the other NVMe fabrics (InfiniBand, RoCE, and iWARP) are memory-oriented (they support remote direct memory access, or RDMA). Indeed, recent conferences covering flash and other persistent memory technologies have gushed over the arrival of the “storage/memory convergence.”

Mapping the dichotomy

Wait a minute . . . memory and storage converging? *What?* For decades, memory and storage have represented a dichotomy. Both could hold information, but memory was built into the server, while storage has largely been separate, holding data independent

of a server or application. To some degree that dichotomy has been self-reinforcing:

- » Classic enterprise storage is relatively slow compared to dynamic random-access memory (DRAM), with extensive error checking and read/writes that are often sequential. Memory is designed to be fast but transient.
- » Hard disk drives (HDD) and solid state drives (SSD) scale far beyond normal memories. They are cheaper per bit, and can persist their data when powered down, which is important for archiving.
- » Enterprise storage also supports a range of service-level agreements (SLAs), with cool features like redundant array of independent disks (RAID), replication, deduplication, and compression. Try that with memory.

Table 1-1 offers a comparison of memory and storage characteristics.

TABLE 1-1 **Memory and storage characteristics**

Feature	“Ideal memory” priority	Flash memory is like. . .	NVMe protocol is aimed at. . .	“Ideal storage” priority
Read Bandwidth	Very high	Memory	Memory	Medium
Write Bandwidth	Very high	Storage	Memory	Medium
Read Latency	Very high	Memory	50/50	Medium
Write Latency	Very high	Storage	50/50	Medium
Read Granularity	High	Memory	Storage	Low
Write Granularity	High	Storage	Storage	Low
Scale	GB to TB	GB to PB	Storage	TB to EB
Random Access	Very High	Memory	Memory	Low
Persistence	Low	Storage	Storage	Very high
Rewritable	High	Storage	Both	Low to Med
Reliability	High	Memory	Storage	Very high
Density	Medium	Storage	Storage	High

OTHER DATA ABOUT DATA

For decades, magnetic recording technology in the form of disk and tape were the dominant storage technologies, while silicon (mostly DRAM) has been the dominant memory technology. The main technology driver behind the NVMe protocol has been the steady improvement in density and performance of flash memory. Flash has displaced disk-based storage over a number of years, and received a big boost with the high-profile replacement of the disk-based iPod mini with the flash-based iPod Nano in late 2005.

So, memory and storage still look different and that's likely to continue. To paraphrase Mark Twain, reports of the convergence of memory and storage may be somewhat exaggerated. Perhaps we can say that there is something of a trend toward convergence, rather than an imminent event. We can also say that there's an emerging convergence of protocols for shared memory and shared storage.

On errors

While understandable, it's somewhat ironic how memory errors are largely tolerated (or at least not eliminated) in computing, while storage errors are not. This is true at a variety of tiers. Laptops (user grade compute) typically have no error correction on DRAM, but have CRC error detection built into drives. Servers (enterprise-grade compute) have ECC DRAM, which corrects single-bit errors but simply aborts or shuts down on dual-bit errors. By contrast, enterprise-grade storage includes redundancy in some form, such as RAID or erasure coding.

Instead of fixing every memory error, the industry approach is to abort and restart a computation using the same stored data. That's because storage has a higher level of guarantee, or service-level agreement (SLA), which is the contract between storage customers and their storage provider. This terminology is often used even within organizations between consumers of storage and their IT departments.



WARNING

Solving the problems of working memory would not fully solve the rare but meaningful untrustworthiness of computation. Viruses, loss of network connectivity, and power failures are just a few of the events that often disrupt inflight computation. This is why overinvesting in working memory rarely makes sense. By

contrast, long-term storage must be recoverable because no “do-over” mechanism exists. The moral? Don’t cheap out on storage. We return to this point in upcoming chapters.

Accelerating Access to Flash

Flash is disrupting the storage world in a big way, but it’s hardly a first for the industry. The storage market is such a poster child for disruption that way back in 1995, Harvard’s Clayton Christensen used it as a prime example in his classic HBS paper on disruption. The early disruptions were more about size and cost, while flash is more about performance.

Flash has always offered much faster read capability (especially for random access) than spinning disk drives. But flash’s early density was much lower than disk drives. In addition, writing to flash is trickier than writing to DRAM or magnetic storage media. Flash has relatively low write endurance, in the neighborhood of one million erase/write cycles for each flash block. Repeated writes to a block of flash can also degrade the reliability of adjacent blocks. So, despite its speed, flash’s early shortcomings limited it to niche uses.

But over time, the density of flash increased dramatically, and effective software algorithms were developed to mask its write challenges. The combination of speed, density, and tolerable write endurance have brought flash to the point where it is poised to displace spinning disks.

Indeed, the killer speed benefit of flash, as well as other solid-state memory technologies, has highlighted that the old tried-and-true storage protocols had a weakness: performance.

Understanding How NVMe Relates to SCSI

The basis of most of today’s storage-oriented protocols, including FCP, is the SCSI standard established in the 1980s. SCSI was originally built around hard disk drives, but has been extended a number of times to include other storage devices while maintaining backward compatibility. SCSI currently supports well over 100 commands. Besides hauling a lot of baggage, SCSI also lacks deep command queues.

SCSI's numerous extensions and extended support for legacy applications have resulted in a protocol stack that is sluggish in comparison to the NVMe stack, which has dramatically enhanced queuing and has been simplified and optimized for both semiconductor memory and today's operating systems (see Figure 1-3).

Here are some important things to understand for those who may, at some point, be migrating SCSI-based storage assets into an NVMe environment:

- » **Mapping legacy SCSI to NVMe:** The NVMe community has recognized the importance of the storage market and the prominence of SCSI within that market. That's why the NVMe standards groups invested time and energy to ensure that NVMe could implement the functionality needed by legacy storage-dependent applications.
- » **LUNs and namespace IDs:** LUN stands for logical unit numbers, the SCSI mechanism for identifying different volumes within a single storage target. In other words, each volume is a LUN. NVMe uses the term "namespace ID" in a similar way. "Namespace" is a curious term, considering that each one is treated as a set of logical block addresses (LBAs), not as a set of names.

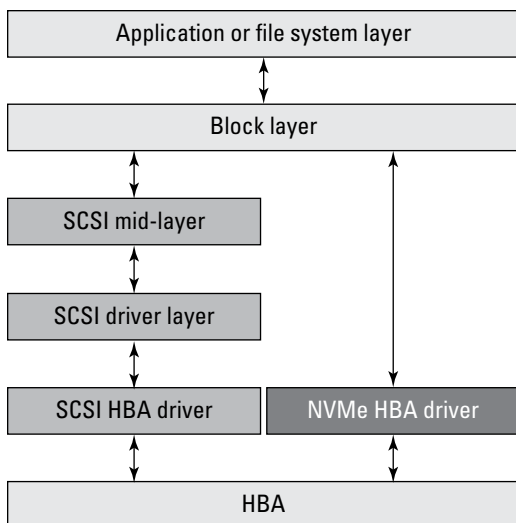


FIGURE 1-3: Comparing the SCSI and NVMe software stacks.

- » **Enhanced command queuing:** NVMe over Fibre Channel exposes the enhanced queuing capabilities of NVMe, allowing thousands of parallel requests across a single connection. With today's multithreaded servers and virtual machines running dozens to hundreds of applications, the benefits of parallelism are massive.
- » **Fibre Channel protocol (FCP):** Fibre Channel protocol is, like namespace, an odd moniker for what is really "SCSI over Fibre Channel." FCP is not about basic Fibre Channel; it is about the way SCSI features were implemented on top of Fibre Channel.
- » **Leveraging FCP:** Without going into the gory details, you should know that NVMe over Fibre Channel uses a new FC-NVMe frame type for non-I/O communications while reusing the FCP frame type for I/O operations. So, if you capture all the frames running across an NVMe-over-FC interface, you will see FCP in the mix.
- » **Other protocols:** The name FCP for SCSI-on-FC might be partly responsible for a perception that Fibre Channel is limited to SCSI, but it's important to remember that SCSI is not the only popular protocol used with FC. The mainframe storage protocol FICON runs over FC, and NVMe is another protocol that now runs on top of FC (which is the whole point of this book).



WARNING

SCSI's sluggishness is a characteristic of the protocol stack, not the Fibre Channel (FC) transport. Some NVMe advocates have pointed to the implementation of SCSI over FC and incorrectly blamed its relatively poky performance on Fibre Channel. This is an incorrect assertion: FC-NVMe is faster than SCSI over FC.

Anticipating Future Benefits of NVMe over Fibre Channel

One of the key benefits of NVMe over Fibre Channel is its scalability. Built from the ground up with non-volatile memory in mind, it also leverages the speed and robustness of Fibre Channel. (We say more about the inherent benefits of Fibre Channel in Chapter 2.) Leveraging Fibre Channel as a transport gives users easy access to all the speed and parallelism of NVMe over Fabrics with none of the disruption entailed in building parallel infrastructure.

Here are some additional considerations as NVMe over Fibre Channel increases its lead on competing technologies:

- » As flash becomes more storage-oriented, the dominant storage protocol (SCSI) has hampered one of the key advantages of flash, its speed. That will change as more storage vendors embrace NVMe over Fibre Channel.
- » New semiconductor memories such as Intel/Micron's 3D XPoint (see sidebar) are just coming onto the scene, and hold out the promise of much faster writes as well as 100x or even 1000x write endurance.

Invigorating Your Fabric

As the underlying technology of storage arrays moves from spinning disk to flash, and from flash to even faster technologies, the increasing speed will generate increased pressure to save those few precious microseconds that NVMe offers over SCSI.

In addition, many applications will have mixed needs, requiring some storage-oriented volumes and some memory-oriented volumes. Even more compelling, there will be times when you want to do both with the same information. That is, you'll want to maintain a master copy of some data asset, enabling all the high reliability features of enterprise storage.

At the same time, other consumers of that data asset may only need high-speed read access to that data. It may make sense to publish (using the dual-protocol concurrency of your Fibre Channel fabric) your master storage volume to a "working reference memory" on an NVMe-over-Fibre Channel drive. Such an image could be read-only and would have no need for features that could drive up latency or cost. By using memory-oriented NVMe-over-Fibre Channel arrays, you may save money as well.

- » Reviewing Fibre Channel's place in the storage ecosystem
- » Evaluating performance metrics in the storage context
- » Recognizing the advantage of improved queuing
- » Achieving high performance
- » Realizing reliability

Chapter 2

Delivering Speed and Reliability with NVMe over Fibre Channel

Ethernet was already the mainstream network of choice when Fibre Channel (FC) first started shipping in 1997. Ethernet was well on its way to overrunning protocols like Token Ring and asynchronous transfer mode (ATM), and many in the networking community doubted whether FC had any future at all.

Boy, were they wrong. In the face of strong Ethernet headwinds, Fibre Channel networking managed to grow to the point where nearly all enterprises rely on it for their mission-critical storage needs. FC's success was not magic. FC was and remains different from Ethernet in important ways. Understanding why FC has been so successful in an Ethernet-dominant world is a necessary first step toward deciding whether to adopt this robust networking technology.

Reviewing Fibre Channel's Place in the Storage Ecosystem

Ethernet remains the primary transport for communication between servers. There are, however, distinct advantages to using Fibre Channel in a storage-centric environment. Traditional Ethernet, including most Ethernet deployed today, doesn't do much about network congestion. If you pretend for a moment that your network is like a freeway at rush hour, Ethernet allows unlimited cars (frames) to enter the onramps, but then steers them into the ditch (drops the frames) when the freeway runs out of space. For those cars that don't reach their destination, no biggie: Transmission Control Protocol (TCP) patiently, hesitantly, tries again, sending as many cars as necessary, even if there's already a car crash up ahead.

Fibre Channel, on the other hand, was designed for well-ordered transport of data. It has a feature similar to the entrance ramp lights most of us grumble about on our daily commute. No frames are lost, and each is delivered in the proper order. The Fibre Channel freeway doesn't have traffic jams or fender benders. This makes Fibre Channel the ideal solution for mission-critical storage requirements.

Evaluating Performance Metrics in Storage Context

You can't improve what you can't measure. That's why it's important to establish performance metrics before embarking on any improvement project, even for something as trivial as planting a garden — if the seeds aren't deep enough or the wrong fertilizer used, you might just go hungry.

Storage is no different. Without a clear understanding of how your company's megabuck hardware investment is performing, or whether data is being lost or users are complaining over slow access to corporate data, you might end up with a lot more time on your hands. Maybe you can take up gardening?

This section focuses on storage metrics; they are what the storage consumer ultimately cares most about, and they better enable apples-to-apples comparisons between different NVMe over

Fabrics options. (For a peek at fabric metrics, see the sidebar, “Fabric Metrics.”) The storage community has long used three key metrics for measuring performance:

- » Latency
- » Throughput
- » Input/output operations per second (IOPS)

The relative importance of these performance indicators depends a great deal on the user application. Systems focused on minimizing response time value latency above the other metrics. Streaming high definition video requires huge amounts of data, so good throughput will be paramount. And the heavy read/write activity seen in databases calls for high numbers of IOPS (pronounced *eye-ops*). Even the experts vary. One calls latency the “king” of storage metrics while another refers to IOPS as the “grandfather.”

FABRIC METRICS

Fabric metrics can often resemble storage metrics, and can sometimes cause confusion if you’re not familiar with their distinct meanings. Both storage and networking folks should be alert to the subtleties to avoid embarrassing apples-versus-oranges fruit salad episodes.

While storage latency measures a full storage operation from start to finish, fabric latency tells you how much incremental latency a fabric device would add to a connection relative to a direct connection. It is measured as the time between the arrival of the first bits of a frame and the time those same bits are first transmitted (“first in, first out” — the FIFO model). The fabric latency is the same for both read and write operations and for different I/O sizes.

Fabric throughput is usually viewed as how much data can be pushed through the fabric when all ports are running maximum speed. A 64-port, 10G device usually has a throughput of 640 Gbps (double that if input and output are separately counted). But some low-end devices have internal “oversubscription” and cannot forward at full speed on all ports at once, so check the fine print.

No fabric metric corresponds to IOPS. However, the IOPS metric exists because the latency metric does not tell the full story when I/O operations overlap. Similarly, when multiple traffic flows overlap in a network and create congestion, no simplistic metric exists to capture behavior.

Of course, there's more to overall performance than the speed of the network. If your hard drives are dogs, why bother with Fibre Channel? Similarly, if the server is still using Pentium Pro processors, you'd better address your server performance first before attacking any network upgrades. As with everything in life, some things have greater priority. And sure enough, the whole NVMe conversation itself was started by high-speed solid-state drives (SSDs).

Storage Metrics

Here's a quick introduction to the key storage metrics:

- » **Latency**, especially read latency, is the main claim to fame of flash-based systems, and it's the benefit most often touted for NVMe-based data transfers (normally in comparison to SAS, or serial attached SCSI).

Storage users care about overall operation latency, which is the time from when a read or write operation begins until that operation has fully completed. Storage latency depends on the size and direction of the I/O operation, whether I/Os are random or sequential, as well as the connection speed, and the relevant values should be included when mentioning a particular latency metric value.

- » **Throughput** describes how fast, in megabytes or gigabytes per second, a storage device can read or write data. These metrics are most often better for large I/Os. As with latency measurements, throughput measurements should state the I/O direction (read or write) and access type (sequential or random). I/O size is good to include for completeness, but if it isn't mentioned, you can assume the number applies for larger I/Os.

- » **IOPS** tells us how many individual read and/or write operations a device can handle per second. Like latency and throughput, IOPS metrics vary by the size, direction, and access type (sequential versus random) of the I/Os.

Typically, a device's IOPS metric is higher for smaller I/Os. That's why quoted IOPS metrics are often for an I/O size such as 4KB. However, many applications that demand high IOPS use larger I/O sizes, such as 64KB. You need to look closely to ensure that your device IOPS metric is aligned with what your application needs.

In a simple case, IOPS are closely related to latency. If you imagine a connection where a 4KB read operation takes 1 ms, you might expect to be able to perform 1,000 of those I/Os in a second, and indeed that might be the case. But because this simple picture doesn't always hold true (more on this later in the chapter), it's useful to check both latency and IOPS metrics.



WARNING

Storage metrics can't tell you everything. In order to allow for apples-to-apples comparisons, performance benchmarking is done in controlled situations, such as a single tester connected to a single device. The upshot is that, in the real world, “your mileage may vary.”

Although the “controlled situation” aspect of performance benchmarks is legitimate, it also creates natural pressure to tune devices so they excel in the test situation, although they may not always perform as well in familiar real-world environments:

» **Example 1:** Some devices can have excellent throughput or IOPS numbers for sequential accesses, but that performance may apply only when you have a small number of storage clients issuing operations. A large number of requesters can overload the device resources, causing the sequential performance benefit to be lost.

» **Example 2:** Some flash arrays keep a pool of erased flash blocks to allow for faster writing. For a write operation, the controller “remaps” the relevant logical block addresses (LBAs) to a block from the pool, writes the new data there, and marks the old flash block to be erased in the background. If the device runs low on erased blocks, the “garbage collection” process can lurch into the foreground and dramatically slow normal operations until the process finishes.

Because the real world is not a controlled situation, IT architects who care about consistent high performance demand that their environments include tools that enable rapid and deep investigation into system behavior. Brocade, a Broadcom Inc. Company, has long recognized that the expectations of Fibre Channel customers (unlike the commodity-biased Ethernet market) justify an investment in analytics tools. Those customers who are seeking the performance benefits of NVMe technology are especially likely to find themselves in need of tools for optimizing their environments. (For more information, see the sidebar on Brocade Fabric Vision Features.)

BROCADE FABRIC VISION FEATURES

Brocade's Fabric Vision technology provides numerous tools for analyzing and optimizing Fibre Channel fabric performance and reliability. Here is a brief summary of some of its more important features.

IO Insight: On supported products, IO Insight proactively and non-intrusively monitors storage device IO latency and behavior through integrated network sensors, providing deep insight into problems and ensuring service levels.

VM Insight: Seamlessly monitors virtual machine (VM) performance throughout a storage fabric with standards-based, end-to-end VM tagging. Administrators can quickly determine the source of VM/application performance anomalies, as well as provision and fine-tune the infrastructure based on VM/application requirements to meet service-level objectives.

Monitoring and Alerting Policy Suite (MAPS): Leverages prebuilt, rule/policy-based templates within MAPS to simplify fabric-wide threshold configuration, monitoring, and alerting. Administrators can configure the entire fabric (or multiple fabrics) at one time using common rules and policies, or customize policies for specific ports or switch elements. With Flow Vision and VM Insight, administrators set thresholds for VM flow metrics in MAPS policies in order to be notified of VM performance degradation.

Flow Vision: A set of flow-oriented tools that enables administrators to identify, monitor, and analyze specific application and data flows in order to simplify troubleshooting, maximize performance, avoid congestion, and optimize resources. Two of the Flow Vision tools are listed next.

Flow Monitor: Provides comprehensive visibility into flows within the fabric, including the ability to automatically learn flows and non-disruptively monitor flow performance.

Flow Mirroring: Provides the ability to non-disruptively create copies of specific application and data flows or frame types that can be captured by Brocade's Analytics Monitoring Platform (AMP) for in-depth analysis.

Souping up Device Metrics

Read caching can help hard disk drives (HDDs) when accessing data sequentially. That's because the disk drive can read an entire disk track and cache the extra data blocks. Read caches don't help SSDs as much (they handle any reads pretty quickly), but sometimes a read cache can be helpful for SSDs that are otherwise unable to read during write operations.

Write caching can be helpful for flash in two ways:

- » **Write speed:** Writing to a dynamic random-access memory (DRAM) is faster than writing to the flash device, which must be erased before they can be written.
- » **Write endurance:** An application may write the same block multiple times in a short period. The write cache can wait a bit, then transform multiple cache writes into a single write to flash.

Parallelism is a simple matter of using a large number of underlying devices to deliver higher throughput, and often higher IOPs.

Pipelining describes a system that performs different functions in parallel. For example, read operations may be broken into different functional stages: command pre-processing, physical access of backend devices, error correction, and sending. The functional stages are like different sections of a pipeline, and you can see the read operation “moving through the pipe.” With pipelining, the system can be working on different stages of two read operations at the same time. Separate pipelines are normally used for read and write.

Pipelining allows a device's IOPS metric to exceed what you might expect from its latency metric. For example, you might expect a device with a 1 ms latency metric for 256KB reads to have a 1,000 IOPS metric. But when reads are overlapped (later reads are sent before earlier reads have completed), the device may deliver a 1,500 IOPS metric for 256KB reads.

Achieving overlapped reads or writes is tricky for “single threaded” applications, but most performance-sensitive apps are now “multi-threaded” to generate the overlapping I/Os and take advantage of pipelined device performance. In addition, virtualized servers running lots of apps in parallel also issue many

overlapping I/Os. (An interesting side effect is that minor latency changes may not change overall IOPS, depending on where the system bottlenecks are. More on this later in the chapter.)

As you'll see, the NVMe standard includes architectural enhancements that can greatly accelerate overlapping I/Os.



TIP

If you're considering using a device with write caching, ensure that the device's power-fail write-back behavior is aligned with the application needs. If the application requires that all writes are made persistent, the device must guarantee that the cache contents are saved in the event of a power failure.



WARNING

Be aware that architectural performance enhancements can only go so far. Write caches can help with write bursts, but if the long-term requested write throughput exceeds what the hardware behind the cache can swallow, the cache fills up, and the requested writes get throttled to match the underlying device throughput. Pipelines may lose performance benefits when operations to the same underlying device overlap. You'll need to test specific products with your applications.

Achieving High Performance

High performance is relatively straightforward in NVMe over Fibre Channel, just as it is with Fibre Channel itself. Fibre Channel vendors have pushed the performance edge from the beginning, with top speeds and features like direct placement of storage payloads to reduce memory copy overhead. Customers contributed as well, by opting for optical connections more often than mainstream networking did. Fibre Channel provided a much simpler networking stack, with a single network layer, simplified addressing, and topology-agnostic routing that allows for all links to be used in parallel. In addition, Fibre Channel fabrics are typically implemented as parallel, redundant, active-active fabrics, which offers both reliability and added performance. Similarly, Fibre Channel's built-in credit-based flow control offers reliability while also improving performance.

All these optimizations make perfect sense in a data center architecture, even if some of these choices would be out of place in a campus or Internet context. NVMe over Fibre Channel leverages all the traditional benefits of Fibre Channel, while

providing additional performance benefits inherent in NVMe. The streamlined protocol stack is one benefit that we've mentioned already. The other major improvement is enhanced queuing.

Making Use of Enhanced Queuing

For decades, storage vendors have competed for leadership on each of three of the metrics described in this chapter, and have long made use of techniques such as caching, parallelism, and pipelining to boost their performance metrics.

We've already mentioned how flash drives have a big advantage in latency over disk drives because they don't have to twiddle their thumbs while the media spins or the read/write head lumbers out to the appropriate track. Another advantage is that flash chips are much smaller than the smallest disk drives. This means using thousands of flash chips in parallel is much easier than using thousands of disk drives.

And, just as the storage targets have been moving toward extreme parallelism, so have the storage initiators. Servers are running vastly more threads, cores, and virtual machines. As a result, the number of parallel I/Os that can be generated has risen steeply.

SCSI-based devices offered parallelism, allowing storage initiators to "queue up" multiple commands in parallel. But the SCSI queue depth has been limited for both individual LUNs (logical unit numbers, or volumes) and for target ports, which typically support several LUNs. SCSI queue depth limits vary from 8 to 32 commands per LUN, and 512 commands per port. Historically, these seemed more than adequate, but as today's environments scale out, SCSI is feeling the squeeze.

The designers of NVMe were aware of the trends, and defined the protocol's queuing depths accordingly. NVMe supports a vast 64,000 queues with a depth of 64,000 commands each.



REMEMBER

Enhanced queuing allows for a far greater parallelism. This doesn't mean you'll immediately see 100x improvements with NVMe, but it isn't hard to imagine a significant boost of 2x or more. The advanced analytics features available with Brocade's Fibre Channel fabrics enable you to track the queue depth across all the devices and applications in your environment.

Realizing Reliability

Everyone wants reliability: reliable cars, reliable employees, reliable Internet. But without reliable data, the preceding examples might be difficult to attain.

Of course, everyone in IT knows that users want reliable computing. And yet, there are different ways to deal with errors, and some are more expensive than others. Companies tolerate the occasional crash in their laptops rather than spend the money (and battery life) on the error correction circuitry (ECC) required to minimize bit errors. After all, laptops have many exposures, so users seek to protect them in other ways, such as background backup software.



REMEMBER

Many companies' servers have ECC to fix single-bit errors, but crash on double-bit errors. Servers, like laptops, are also vulnerable to viruses and power failures, so companies don't spend more for killer ECC that fixes double-bit errors. They live with the occasional server crash because they can rerun the app to get the results. That's only true because their enterprise storage guarantees availability of golden copies of the critical data assets. How would life be different if you couldn't count on those assets?

Redundant networks and multipath IO

When Apollo 13 was launched, it was chock-full of redundant systems designed to make the spacecraft function properly in the event of an isolated failure. Your data assets may not be quite as precious as the lives of astronauts, but you do have critical assets, and you've had enough experience to know where redundancy makes sense. For enterprise storage customers, it's a matter of economics; the right amount of redundancy enables you to deliver "five 9's" reliability, keeping your customers happy.



TIP

Cutting corners is penny wise but pound foolish if you disappoint your customers and they toddle down the road to a competitor.

Enterprise storage vendors understand all that, and have invested significant amounts of R&D to make robust systems, both in their storage targets and in their storage networks, that keep operating in the presence of the inevitable rare glitch. (The technology is more mature than a Saturn V, and fortunately, the glitches are rarer.)

The vendors have also worked hard to provide multipath I/Os so that you have no single point of failure, and you get the added benefit of tiny or even zero service upgrade windows. Customers have effectively forced the storage vendors to do this by voting with their wallets. And all the vendors must be prepared to deliver 24/7 enterprise support for their own products as well as other products that they sell.

The enterprise storage vendors understood their market, even in the mid-1990s when Ethernet was beating up on other protocols. Despite the traction of Ethernet, the vendors chose Fibre Channel. And for many good reasons.

Features of a lossless network

Losing a dog, losing your car keys. Loss of something you value is a terrible thing, and so you take care of those things. Other things, like soda straws — not so much. If you lose it, you'll get another one.

Fibre Channel has always been a lossless networking technology. It treats a payload like a loved one. Every Fibre Channel link is governed by buffer credits that the receiver shares with the sender. The sender knows how many buffers are available at receiving side and does not send a frame that the receiver can't handle. By contrast, Ethernet has been lossy network for decades, treating packets like soda straws, dropping packets in a whole variety of situations and relying on TCP or other mechanisms to replace the lost straws. Data Center Bridging (DCB) is an Ethernet variant that uses PAUSE frames (rather than buffer credits) to avoid packet loss, but DCB still has some notable interoperability challenges.

Security

Obviously, if we are treating our frames like loved ones, we need to protect against inappropriate human actions as well, whether those are simple mistakes or something more troubling. Fibre Channel offers extra security in a number of ways. Fibre Channel is already trusted to keep the world's most important data secure, and Fibre Channel brings this security model to NVMe over Fabrics.



One key advantage of Fibre Channel comes from its specialized nature. Datacenter-centric Fibre Channel is not the protocol of the Internet, so hackers cannot shove Fibre Channel frames across the Internet to sneak past your firewall and into your datacenter.

Fibre Channel also provides a zoning service that integrates storage access controls into the network. This tried-and-true service works across all enterprise storage vendors, even in multivendor environments.

Good tools matter

To err is human, and wise IT administrators know that it's a good career move to use robust tools to reduce the opportunities for, shall we say, demonstrating their humanity. HPE recognizes this as well, and developed Smart Fabric to improve storage reliability and security. Check out the HPE Smart Fabric and Smart SAN sidebar for more info.

HPE SMART FABRIC AND SMART SAN

HPE Smart SAN for 3PAR makes SAN configuration and management simple and error-free through intelligent automation. HPE Smart Fabric is a protocol agnostic application embedded in storage networking components that enables 3PAR storage arrays to automate and orchestrate end-to-end configuration, settings and policies. HPE Smart SAN's Target Driven Peer Zoning enables you to configure zones accurately in minutes instead of hours. Combined with automatic discovery mechanisms, this powerful platform offers real-time diagnostics and SAN analysis to significantly enhance the resiliency of your environment.

- » Identifying your situation
- » Considering your adoption strategy
- » Exploiting dual-protocol FCP and FC-NVMe
- » Familiarizing yourself with NVMe over Fibre Channel

Chapter 3

Adopting and Deploying NVMe over Fibre Channel

You've done all your homework. You've read a bunch of manuals, gone to a few seminars, and talked to colleagues. Everyone on your team agrees it's time to move forward with NVMe-over-Fibre Channel adoption in your organization. Only one big question remains: Where do you start?

Identifying Your Situation

Several factors work in your favor when you implement NVMe over Fibre Channel. You don't need to divide your budget and invest in parallel infrastructure. Nor do you have any worries about multivendor interoperability of equipment, or new protocols and discovery algorithms to grapple with. And you don't have to risk education funds on uncertain IP/Ethernet NVMe protocols, as you might with competing technologies. That being said, you should check a few legacy infrastructure boxes before going further:

- » You can deploy NVMe over Fibre Channel on an existing Brocade Fibre Channel infrastructure, provided it is relatively up to date. Be sure to check with your hardware supplier on interoperability.

- » The Fibre Channel fabric must be Gen 5 (16 Gbps) or better yet Gen 6 (32 Gbps).
- » Servers using NVMe over Fibre Channel need Gen 6 host bus adapters (HBAs); Gen 6 HBAs also work with Gen 5 fabrics.
- » You need a storage device that supports FC-NVMe frame types. This can be an FC-NVMe-enabled array, or, for early familiarization, a server with an FC-NVMe HBA running in target mode can fill the role of storage target.



TIP

None of these requirements are unreasonable. For example, if you have servers running performance-sensitive applications and you're still on Gen 4, it's definitely time for an upgrade, regardless of any NVMe-over-Fibre Channel undertaking. This presents an excellent opportunity to do so.

Considering Your Adoption Strategy

Setting aside the glorious predictions of the NVMe hype-masters, few in the storage and networking communities would dispute the notion that the move to NVMe over Fabrics will be gradual, lasting several years. Unfortunately, this slow transition would make it painful to have a classic Fibre Channel (FC) network sitting side by side with your brand new Ethernet-based NVMe network. When you buy more storage, what kind do you buy? As you build new apps, which environment do you connect them to? And if you go with Ethernet, which kind: iWARP or RoCEv2? Neither is the obvious winner, and neither has much of a historical adoption record. However, adopting a dual-protocol Fibre Channel fabric (running concurrent FCP and FC-NVMe traffic) eliminates or simplifies those questions. Fibre Channel is widely adopted and enjoys excellent support by storage vendors and other technology providers, making NVMe over Fibre Channel a relatively low-risk proposition.

Protecting high-value assets

Although NVMe over Fibre Channel is new technology, for most storage-oriented usage (as opposed to “working memory”), the goal is to apply the technology to an existing storage asset. With concepts like Big Data, data mining, data lakes, and so on gaining traction, the value of everyone's data assets is climbing rapidly. This makes a top-notch, high-performing storage solution even

more crucial than in days past, and at the same time highlights the importance of keeping risk to a minimum.



TIP

If an application is merely using a copy of a data asset (not modifying or updating the master copy), the application is effectively treating this asset as working memory. Conversely, if the application will be maintaining the master copy of the data, maintaining integrity and availability of the data asset is vital.



REMEMBER

Most often, when an existing data asset is involved, you are looking at a “brownfield” (as opposed to a “greenfield”) scenario. Migration from legacy architecture, which is typically built on a SCSI infrastructure, to one based on NVMe should be done in an incremental fashion after lab validation, and allow for the option of rolling back architectural changes. The ideal adoption strategy includes a process and infrastructure that allow for such a model.

Allowing for a marathon shift

With all the buzz about the speed of NVMe and the rapid rise of all-flash arrays, won’t the whole world go to NVMe storage by the end of next week? Probably not. Consider all the claimed advantages of IPv6 and the length of time required for significant adoption of that technology. The transition is still moving very slowly, many years after most infrastructures were IPv6 ready, with full support available on both hardware and software. Even if the adoption of NVMe is substantially faster, the transition will take years.

Realistically, some firms or some departments will adopt NVMe rapidly when they are interested only in working memory, and have no urgent need to maintain high value data assets. Often such use cases will be handled with direct-attached NVMe products only, and fabrics will not be required during this phase. Other firms that do not have compelling need to accelerate their working memory may first adopt NVMe for storage use cases. And of course, there will be some combinations. The point is, there will be different adoption rates for different kinds of usage.

Exploiting Dual-protocol FCP and FC-NVMe

IT people are responsible for company data and productivity, and therefore rightly concerned about risk reduction (“de-risking,” in geek speak).

As IT departments plan to deploy NVMe-based arrays in production, they have two main options. They can either create some new NVMe infrastructure (see Figure 3-1), or they can leverage an existing infrastructure (see Figure 3-2). If they build a new, separate infrastructure, every array purchase after that becomes a gamble because they must choose which infrastructure will access the array. That’s why dual-protocol concurrency is the big win for NVMe over Fibre Channel.

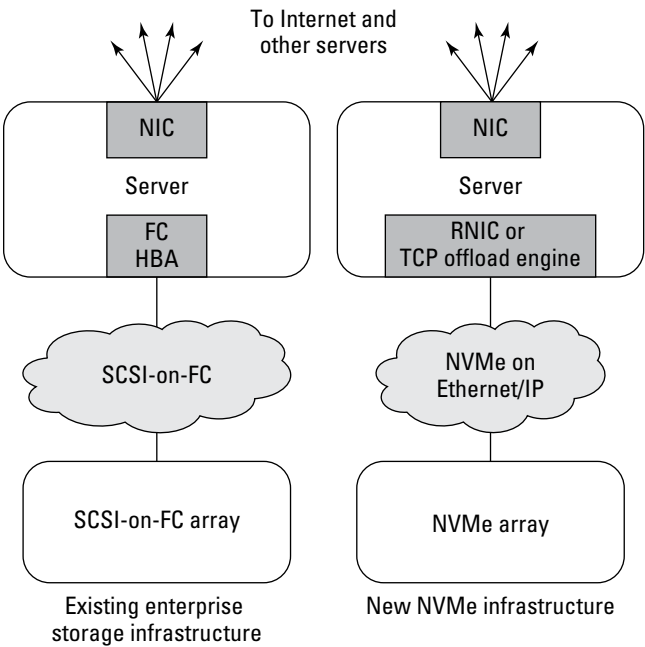


FIGURE 3-1: New separate infrastructure (not recommended).

By leveraging a “known quantity” such as a Fibre Channel SAN (storage area network), companies can easily support dual protocols and remove any risk associated with the inevitable question, “How long will the transition take?”

Nor is a dual-protocol approach unreasonable. A strong precedent for mixing protocols on top of Fibre Channel exists. Since 2001, it’s been an established practice to support both FCP and FICON traffic simultaneously on the same Fibre Channel fabric (see the sidebar, “FICON and FCP”).

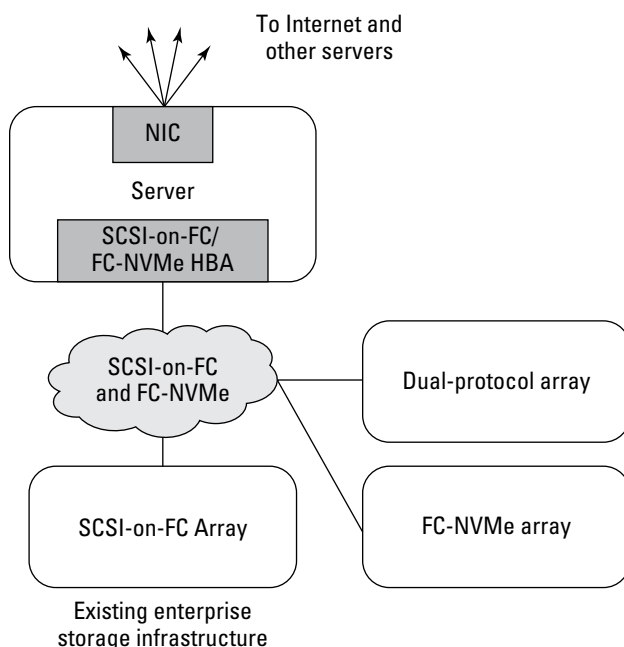


FIGURE 3-2: Dual-protocol infrastructure (recommended).

Other considerations include:

- » **Incremental migration:** Often a single application uses many storage volumes and may have different requirements on those volumes. In a dual-protocol environment, individual volumes can be migrated as appropriate. High value, risk-intolerant assets can remain on trusted infrastructure. Lower value latency-sensitive volumes can move to the hottest new targets. Changes can be rolled back easily. In a dual-protocol environment, these changes can be achieved administratively without making disruptive hardware or cabling changes.
- » **Dual protocol publishing:** Master copies of high value data assets may be maintained on trusted legacy arrays, and regularly “published” to the latest rocket-fast NVMe-over-Fibre Channel arrays, allowing other applications to use the data as “working memory.” Again, this can be achieved with existing infrastructure.

FICON AND FCP

FCP is the SCSI-on-FC protocol used in open systems such as Windows and Linux. FICON is the Fibre Channel version of a mainframe storage protocol. Many firms that use mainframes have a recurring need to “publish” mainframe data assets so that they can be consumed by open systems. At other times — for example, when consolidating data assets after a merger — they may have a need to “ingest” an FCP data asset into the FICON world.

To achieve the transfer of assets, special migration servers are configured with dual protocol (FCP and FICON) HBAs connected to a dual-protocol SAN that connects to both FCP storage arrays and FICON storage arrays.

The reality of the FCP/FICON dual protocol SANs is that they are not aimed at a long-term transition from one to the other, but offer an enduring “bridge” between the two realms. Because the mainframe and open systems are architecturally different, there is no notion of a gradual migration of applications from one side to the other, and so normal application servers are not configured to use both protocols.

Concurrent dual-protocol FCP/FC-NVMe is different because both protocols are designed for open systems, and consequently you can readily plan for gradual, incremental, non-disruptive migration of an application from one protocol to another. Some apps may benefit from an ability to operate in a dual-protocol mode for an extended period of time, consuming (reading) assets using one protocol and publishing (writing) assets on the other. You may choose to transition other applications relatively quickly, within a matter of days or weeks, and those may spend relatively little time in dual-protocol mode. Either way, the ability to leverage dual-protocol FC gives you tremendous flexibility as you move to deploy NVMe in production.



REMEMBER

Fibre Channel lets IT administrators leverage already familiar elements such as zoning and name services. In addition, dual-protocol concurrency offers opportunities that are otherwise difficult to achieve.

Zoning and name services

Network administrators use *zoning* to improve security. Two approaches to zoning are available:

- » **Port zoning** restricts access based on the specific port of the Fibre Channel device to which a node is attached.
- » **Name zoning** restricts access based on a device's World Wide Name (WWN).

Each of these methods restricts devices from accessing network areas they should not be visiting. Which approach makes the most sense for you depends on your usage. Name zoning usually reduces maintenance effort, except in cases where a sequence of different devices will be connected on one port.



TIP

You may come across references to *hard* and *soft zoning*. Modern FC fabrics use hard zoning, in which robust silicon-based logic blocks traffic between nodes that are not allowed to reach each other. In contrast, early products used software to do soft zoning. Unable to block traffic, the software hid information. It was like covering your address instead of locking your door. Sensibly, soft zoning is no longer used.



REMEMBER

The point is that FCP and NVMe over Fibre Channel can both leverage FC zoning. Fibre Channel's zone services are implemented in the fabric, which is a different approach from those used by competing technologies. Consequently, the results are more predictable and easier to manage, reducing the opportunity for security holes.

Name services, on the other hand, translate obscure computer and device addresses into human-friendly names. They are similar to DNS, but are network resident, greatly simplifying interoperability and management. This has long been the case for FCP, and remains true for NVMe over Fibre Channel.

Discovery and NVMe over Fibre Channel

The NVMe over Fabrics specification described a discovery mechanism, but left many details up to the specific implementation. For non-FC fabrics, this leaves a gaping interoperability chasm, which will likely be as slow to close as previous interoperability challenges such as Priority Flow Control (PFC) and Data Center Bridging Exchange (DCBX).

Fibre Channel's ability to deliver a dual-protocol FCP/FCNVMe fabric provides a clearly mapped forward path to interoperability. HBA vendors are creating drivers that leverage FCP for device discovery, then check those devices for FC-NVMe traffic support. Enterprise storage vendors who offer SCSI-over-FC arrays are

motivated to support this two-step approach as well. Newcomer NVMe array vendors interested in the FC market can leverage the NVMe Express standard mapping from SCSI to NVMe, and they will most likely follow the model as well in order to appeal to existing FC customers.

Familiarizing Yourself with NVMe over Fibre Channel

Someday NVMe over Fibre Channel will be an old friend. For now, at least, it's more like a first date. You're not sure how it's going to react to certain stimuli, what level of attention it requires, or whether it's going to unexpectedly overreact to a stressful situation.



TIP

Go easy. You've already learned all you can about this exciting technology, so now it's time to roll up your sleeves and start playing. Set aside time to become familiar with the nuances of NVMe over Fibre Channel. Go through the various commands and determine how your particular setup is going to work. And then look beyond your test environment, giving thought to how NVMe over Fibre Channel will fit into your production system.

Experimenting in your lab

Like Dr. Frankenstein, you might be tempted to shout, "It's ALIVE!" the first time you fire up your NVMe-over-Fibre Channel test environment. Try to resist this level of enthusiasm, because it might give people the idea that you were surprised by your success. Instead, calmly nod your head and say, "Yep, that's what I'm talking about," and go get yourself another Red Bull.

Not sure where to start? Here are some steps a typical IT department might take when establishing an NVMe-over-Fibre Channel test lab:

- » Set up a single server with an internal drive, a single switch, and a single FC-NVMe-enabled array. You might choose to add a small FCP array later.
- » Connect the server to a lab IP network for access to other lab servers.
- » Explore the HBA config and storage array options you've read so much about.

- » Configure a volume to use the NVMe array (details depend on your NVMe array management tools).
- » Copy a file from the internal drive to the NVMe volume and back again.
- » Run your preferred performance testing application (such as Iometer) to benchmark your NVMe volume. Compare it to your internal volume.
- » Lather, rinse, and repeat until your paranoid nature is quelled.

Migrating your LUN to a namespace

Moving massive quantities of data from one storage technology to another is never much fun. This is especially true when you aren't completely sure of the process. Start small. Migrate a volume or two from SCSI to NVMe (from LUN to namespace ID) to become confident in the process.

You should also run some applications on the NVMe volumes (namespaces) in the lab to build up that muscle memory and be certain you haven't forgotten any steps. At this point, you should have a warm fuzzy feeling that you understand how it all works.

Next, let's expand that comfort level. You've done the basics, so go ahead — open all the storage management apps, SAN management applications, and analytics tools, such as Brocade Network Advisor, Brocade SAN Health, and Brocade IO Insight. Ensure that these tools have been upgraded appropriately and are FC-NVMe-enabled. You should also make sure you're familiar with the alterations that NVMe over Fibre Channel brings to these applications. Lastly, give some thought to which of these tools and features you'll need as you bring NVMe over Fibre Channel online in your production environment.



TIP

As with any high-profile production change, you should think about doing “baseline” performance measurements. Bear in mind that a few hours after flipping the switch to production, you're likely to get a support request because some traditional application is having issues. Was it just a routine human error, or an actual hiccup? If it's a hiccup, did your flipping of the switch cause it? The Boy Scout motto applies here: Be prepared! You should have enough information by now to recognize whether the

support request is linked to the recent change or not. The base-lines you made earlier will help.

Even in the absence of a support call, you'll want to understand how much performance has improved by moving from SCSI to NVMe, because that information will help you evaluate and prioritize other moves. Plus, when you get all those emails from jubilant users thrilled with the speed of their applications, you'll be able to tell them precisely how much faster the system is running. Okay, we know that isn't going to happen, but it might earn you a "Job well done!" from your boss. If you don't know the initial performance, you will not be in a good position to pump your hands overhead, thus broadcasting your NVMe-over-Fibre Channel rock star status.



TIP

It's all well and good that you're now an NVMe-over-Fibre Channel Jedi Knight, but what happens when you go on vacation, or problems arise during the night? Unless you like receiving phone calls while skiing the slopes of Aspen, you'd better make sure the rest of your crew is up to speed and good documentation is in place before going live.

Transitioning to production

Hold on tight, you're going live! Setting aside all the nail biting and late night pizza parties, moving a test system to production is an exciting time. Because your breast now swells with confidence and understanding, it's time to start selecting and prioritizing which applications and volumes are the best place to begin the rollout.



WARNING

Just as you did in the lab, make sure that all your management tools have been properly refreshed for NVMe over Fibre Channel. You don't want to start a cross-country journey and then realize that you've forgotten the map, the gas tank is empty, and the rear tires are low on tread. You're certainly going to be anxious to share the fruits of your labor, but be sure not to shortchange this last — and in many ways most important — part of the process.

Finally, schedule the cutover for a time that won't interfere with your customers (those people whose eyes glaze over when you tell them to reboot), and make certain everyone in the company knows about the looming transition. There should be no surprises for anyone (especially you) and all should enjoy the journey down the NVMe-over-Fibre Channel superhighway.

- » The long and short of Remote Direct Memory Access
- » InfiniBand
- » iWARP
- » RoCEv2
- » Evaluating Ethernet-based NVMe

Chapter 4

Comparing Alternatives to NVMe over Fibre Channel

In most cases, having alternatives is a good thing. This is true whether you're deciding which color carpeting to install in the master bedroom or which way to get home at rush hour. NVMe over Fabrics also offers alternatives, although some might not be to your liking. It can be run over fabrics such as iWARP, RoCEv2, or InfiniBand. This chapter looks at a few of the pros and cons of each, and examines such performance considerations as wire speed, architecture, virtualization, and which special features are supported. Of course, performance is meaningless in the face of risk, so this chapter also evaluates predictability and potential disruption.

The Long and Short of RDMA

RDMA stands for remote direct memory access. It is a protocol designed years ago for use in “tightly coupled” server environments, especially those that fall into the high-performance computing (HPC) category. If humans used RDMA to communicate, there would be no need for speech or body language — thoughts

and emotions would be shared directly, brain to brain, greatly increasing communication speed and eliminating any chance of misinterpretation.

Fortunately, our brains aren't computers, and we can all keep our thoughts to ourselves. For clustered server applications, however, RDMA is a great way to share dynamic information. One server effectively gives "ownership" of some part of its memory to a remote server. For many multi-server applications, especially those involving dynamically changing data, this method offers significant performance advantages.

ZERO COPY FANFARE

When the TCP stack was being developed during the 1980s, a good variety of networking technologies existed. The stack was therefore designed to work with whatever was available, whether it was Token Ring or a phone line. Including clean networking layers made perfect sense for interoperability, and one way to achieve that was the use of intermediate buffering, thus making buffer copies commonplace. As speeds increased, however, most buffer copies were optimized away, except in cases where that practice would break backward compatibility.

In the mid-1990s, a good networking stack could claim single-copy efficiency. A network adaptor received frames and wrote them (using DMA) into DRAM buffers associated with the networking stack. (The unavoidable DMA step is not a DRAM-to-DRAM copy, so it is not counted.) The networking stack would first process the frame, and then copy the "payload" to the memory location desired by the high-level application. For a period of time, this single-copy architecture seemed fully optimized.

Yet by the time FC was being "productized," the game had begun to change. Fibre Channel's main claim was speed, so pressure to optimize was high. Chip technology allowed for more complexity, and the Fibre Channel/SCSI stack was not constrained by the same backward compatibility challenges that IP stacks faced. FC was focused on one "application" (storage), and had a simpler layer structure than TCP/IP/Ethernet. For all these reasons, Fibre Channel was more motivated, and more able to implement a network adaptor/driver/stack architecture that eliminated the single copy. And that's precisely what happened. Fibre Channel has been quietly delivering "zero copy" for the past two decades.

NVMeExpress.org hosts a white paper that describes two types of fabric transports for NVMe: NVMe over Fabrics using RDMA, and NVMe over Fabrics using Fibre Channel. Despite this explicit recognition of FC as an NVMe fabric, some RDMA advocates will claim that because NVMe over Fibre Channel does not use RDMA, it somehow is not an NVMe fabric, despite the fact that NVMe over Fibre Channel does not need RDMA, as we review later in this chapter. FC uses native direct placement capabilities while also enabling the dual-protocol support that allows for a low-risk transition from SCSI to NVMe. If you have any reservations about those claims, hang on tight: We're about to defuse them.

InfiniBand

InfiniBand (IB) came on the scene more recently than Ethernet or Fibre Channel. Focused on server cluster communication, IB is designed to deliver RDMA natively. IB is focused on speed rather than mainstream adoption. Special adapters and switches are required to use IB, which is one reason it never reached broad acceptance or partner compatibility except in specialized HPC applications. In fact, only one IB chip supplier at the time of this writing offers InfiniBand products, a factor that discourages new adopters and raises questions about the cost of switching to the InfiniBand protocol. That sole IB chip vendor seems to have recognized this reality, and has therefore focused its NVMe over Fabrics marketing efforts on promoting RoCE.

iWARP

Though not widely deployed, iWARP is nearly ten years old. The acronym stands for “Internet wide-area RDMA protocol,” an IETF (Internet Engineering Task Force) standard described in five RFCs in 2007, and then updated by three more RFCs as recently as 2014. iWARP is designed to run on top of TCP, which is categorized as a reliable streaming transport protocol because it includes various techniques to ensure that every byte that is sent has been received by the sender. But iWARP's TCP basis is not ideal for storage because TCP normally ramps up transmission speeds slowly. Because many storage applications have traffic patterns that include so-called “bursty elephant flows,” the slow-start behavior of TCP leads to latency challenges and reduced IOPS metrics.

TCP has been extended a number of times, and newer versions of TCP stacks have various configurable (and sometimes negotiable) features that older versions lack. Early TCP implemented a “slow

start” mechanism that began transmissions slowly in order to avoid overflowing network buffers. If protocol timeouts or receiver ACK messages indicate that some transmissions were not received, traditional TCP retransmits and backs off on transmission bandwidth (it “collapses the TCP window”). Some newer versions of TCP, such as Data Center TCP (DCTCP), have features that work better in a datacenter environment, but these features are not compatible with WAN usage. This is why datacenter architects and implementers face a challenge if they want to use TCP for high-performance use cases, and are faced with one of three options:

- » Choose a single complex TCP stack that can be configured differently for different use cases, and mandate this complex stack across all operating system images
- » Choose two or more TCP stacks and manage which image gets which TCP stack
- » Have some OS / hypervisors images that get dual (or more) TCP stacks, mapped internally (probably by IP address) to the desired usage

Each of these three options is problematic, increasing complexity and placing additional burdens on network administrators.

One final drawback comes with performance. TCP was designed to work across a wide range of networks, which of course includes wide area networks (as referenced in the iWARP acronym). But in order to be effective across a variety of networks, TCP has been designed to attempt to minimize lost messages that can occur when the sender sends too fast, which generally means “slow down.” It’s for these reasons among others that iWARP has not been well adopted by the networking community.



TECHNICAL
STUFF

As a reliable streaming protocol, standalone TCP was never intended to guarantee packet or frame alignment. That’s because TCP does not send a set of packets, but rather a stream of individual bytes, removing any certainty that commands placed at, say, byte 8 of the packet will be processed in a timely manner, as the entire stream must first be decoded. TCP is a complex, software-based stack and thus the alignment question would have stumped the hardware processing of iWARP. To address the alignment problem, one of the iWARP RFCs (RFC 5044) created a fix (“Marker PDU Aligned Framing,” or iWARP-MPA) that allows for packet alignment at the cost of increased stack complexity. This is strongly reminiscent of the burdensome SCSI stack that NVMe was created to replace.

Yo, Rocky

RoCEv2 is short for RDMA over Converged Ethernet, version two. It's pronounced "Rocky vee two," and if you call it anything else you'll be laughed out of the server room. It is an odd standard in that it was developed by the InfiniBand Trade Association rather than IETF or IEEE, where most IP and Ethernet standards are developed and maintained.

The RoCEv2 name says "RDMA over Converged Ethernet," but it is a slight misnomer with Version 2, which runs over UDP and thus is no longer directly coupled to Ethernet. For performance and reliability, however, RoCEv2 recommends Converged Ethernet, a dated term for a lossless Ethernet network. Lossless Ethernet is now more formally referred to as Data Center Bridging (DCB) which includes such interdependent features as Priority Flow Control (PFC), Enhanced Transmission Selection (ETS), and Data Center Bridging Capabilities Exchange (DCBX).

DCB is an ongoing effort to enhance Ethernet with the features designed into Fibre Channel during the mid-1990s. Although this is a laudable goal, the interoperability of real-world DCB deployments remains rather low, which interacts problematically with the forgiving nature of Ethernet/IP. A misconfigured DCB network can easily go undetected for long periods, operating as a classic best-effort Ethernet network and suffering random packet loss during traffic spikes.

For a convenient comparison of the differences between Ethernet-based options for NVMe and Fibre Channel, see Table 4.1.

TABLE 4-1

Ethernet-based options for NVMe	NVMe over Fibre Channel
Developing a new fabric protocol standard	Built on T11-standardized Fibre Channel fabric protocol
Standards group dealing with challenges of scaling I/O commands, status and data to the datacenter.	Fibre Channel solved these problems when FCP protocol was developed for SCSI
Transport options: iWARP, RoCEv2 and (soon) TCP	Transport is Fibre Channel. Runs on existing ASICs

(continued)

TABLE 4-1 (continued)

Ethernet-based options for NVMe	NVMe over Fibre Channel
iWARP and RoCEv2 use RDMA (TCP does not)	Does not need RDMA, leverages FCP
Complex network configuration if RDMA is enabled	Fibre Channel is well understood
New I/O protocol, New transport	New I/O protocol, Existing reliable transport
Low latency if RDMA used with RNICs	Same zero copy performance as with FCP
Discovery and zoning services are still in proposal phase	Leverages tried-and-true fabric services

Evaluating Ethernet-based NVMe

Choosing Ethernet for your storage network’s physical layer is challenging, for several reasons. First, the industry is looking at four Ethernet-based NVMe protocols, iWARP, RoCEv2, NVMe over FCoE and soon NVMe over TCP. To get low latency with iWARP or RoCEv2, you need to install RDMA-enabled NICs (RNICs). Big operators like Facebook will choose the commodity-oriented NVMe over TCP (parallel to today’s iSCSI), even if it’s slower. And, of course, you must buy NVMe arrays that support your fabric choice, but who knows which will win? The upshot is that, until the dust settles, choosing any Ethernet-based NVMe protocol is risky.

Second, contrary to the recommendations of the NVMeExpress.org white paper on NVMe over Fabrics, Ethernet flow control does not use the reliable credit-based flow control mechanisms found in Fibre Channel, PCI Express, and InfiniBand transports.

Third, whether you choose iWARP or RoCEv2, you are choosing a multilayered network, with an associated increase in stack complexity (see Figure 4-1) to transport NVMe. Ethernet advocates tout benefits like jumbo frames, even though authorities such as Demartek recommend disabling jumbo frames when using RoCEv2. Some datacenters are using VXLAN, which adds extra Ethernet and IP headers and requires extra management of the “maximum Protocol Data Unit” (MaxPDU) setting for every network port. MaxPDU affects IP fragmentation, which in turn affects IPv4 and IPv6 differently. All of these layers are required in part for legacy reasons, and in part because Ethernet/IP is designed for

Internet scale rather than the datacenter scale of Fibre Channel. Opting for a complex multilayer transport is rather an odd choice when the key benefits of NVMe derive from its simplified stack.

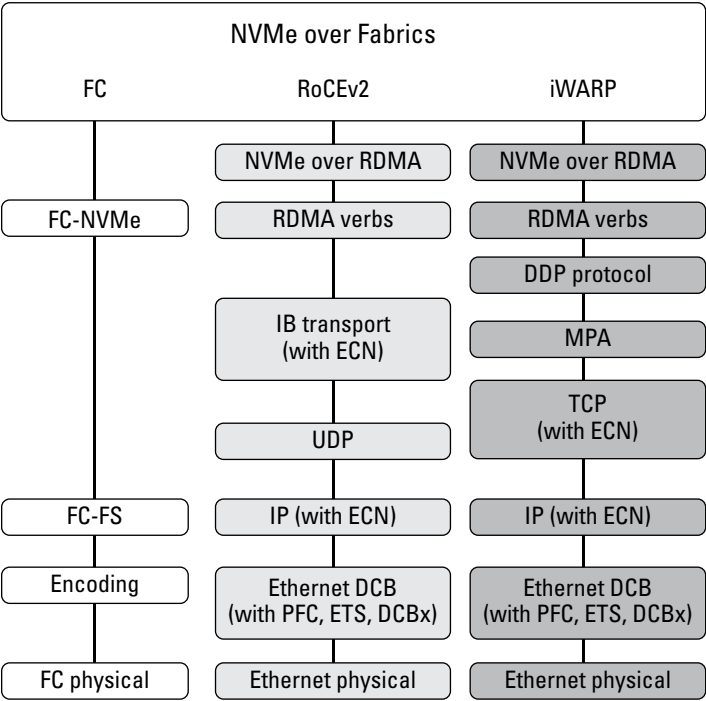


FIGURE 4-1: Relative complexity of different NVMe fabric stacks.

Commodity or Premium?

Ethernet wins as a great low-cost, best-effort technology. It is easy to deploy for common uses and supports myriad upper-layer protocols and applications. Ethernet’s plug-and-play behavior was designed for widespread use, and tens of thousands of well-informed technicians in the industry know how to manage that mainstream Ethernet configuration. Ethernet has simplistic, robust mechanisms like Spanning Tree Protocol that shut down links and guarantee there are no loops that can cause problems with broadcast or multicast storms. These issues might otherwise be commonplace, because broadcasts are a normal part of address learning.

Unfortunately, tree-based topologies are not ideal in today’s datacenter traffic flows, so companies tend to use IP routers at

the top of server racks. Much of the reason Ethernet is so easy to deploy is that its main customer, TCP/IP, is so resilient and forgiving, at the cost of modest performance. Layer 2 Ethernet doesn't scale too far, but coupled with IP, it can scale to the Internet. Ethernet and IP can also be bought anywhere. Interesting products are available on eBay and Amazon, making it a highly competitive marketplace. Fibre Channel networking products are largely available from storage vendors, and it is not always easy playing vendors against one another for lower pricing.

Smart shopping

The flip side is that most storage vendors have invested a significant amount of time in testing their arrays with the networking products they sell. They are familiar with all the enhanced analytics and visibility features that have been designed into the ASICs and the management software. The storage vendor understands the network-resident features like Fibre Channel Name Services and Fibre Channel Zoning, including target-driven zoning, features that are not yet defined for Ethernet-based NVMe fabrics. These vendors are comfortable handling support questions related to the tried and true interplay of servers, HBAs, storage arrays, and networks.



TIP

If you acquire your Ethernet/IP networks from the lowest bidder, consider what the support picture will look like when you call the storage provider about some strange issue. Where do you begin? How do you inspect the network to even begin to tackle the problem? Despite widespread recommendations that enterprise storage should be deployed on a dedicated network, IP storage is frequently connected to a shared network. In light of that, consider surveying your existing Ethernet/IP network and evaluating whether you would be able to support a storage SLA on such a network.

In summary, the widespread success of Ethernet and IP is a double-edged sword. Many of the aspects that make Ethernet and IP widespread and commoditized are problematic when you need a more specialized, premium network. These two protocol suites are obviously the leading answer for the Internet, the campus, homes, and mobile devices, places where Fibre Channel doesn't make sense. Ethernet and IP even work well in the datacenter for multiprotocol best-effort communication needs. But enterprise storage in the data center is a more demanding use case. That's why "good enough" is anything but, and investing in valuable assets makes sense.

Chapter 5

Ten NVMe over Fibre Channel Takeaways

You're ready to get serious about NVMe over Fibre Channel. Here are ten key points to get you started:

- » **Storage and memory have different needs:** You will probably want to support both. Enterprise storage is all about protecting high-value data assets while providing fast access. Working memory, on the other hand, is coupled to computation, where error elimination is secondary to minimizing latency.
- » **Enterprise storage eases handling memory errors:** Mainstream server architectures correct single-bit memory errors while regarding double-bit error correction as not worth the cost. Instead, it's simpler to rely on known-good images kept on enterprise storage and rerun an application when uncorrectable errors occur.
- » **Reduce risk with a transitional adoption strategy:** Installing an entirely new non-Fibre Channel (FC) fabric infrastructure to adopt NVMe is an all-or-nothing approach. It presents risks for your long-term high value data assets as well as your budget. A better tactic is to extend your existing infrastructure, providing an as-needed, gradual transition that protects your data and investments while leveraging existing IT skills.
- » **NVMe's input/output operations per second (IOPS) will have as much impact as its latency:** Most of the buzz over NVMe has been focused on its amazingly low latency,

because this measure is easy to benchmark and the early focus of NVMe is on memory use cases. But as architectures move toward storage and massively parallel processing, IOPS will matter more, increasing the need for robust datacenter fabrics, analytics features, and the excellent vendor support for which FC is known.

- » **Scale NVMe adoption with a tried-and-true fabric:** As NVMe adoption grows beyond direct-attached memory needs to datacenter-scale storage, IT architects will need a robust, credit-based lossless fabric to deliver predictable speed and reliability. Experimenting with immature fabrics will not help the transition to NVMe.
- » **The NVMe future will come at its own pace:** A quick glance at IPv6 reminds you that the vast installed SCSI infrastructure will continue to deliver value for years to come. That's why a sound SCSI-to-NVMe transition strategy must provide flexibility, whether the move lasts two years or two decades.
- » **Include vendor support in your adoption plan:** Reliable IT infrastructure takes far more than sexy technology; products from multiple suppliers must also work well together. Enterprise vendor support for SCSI-based storage works because of interoperability testing, so be sure to ask your vendors about their NVMe fabric interop test plans.
- » **FC and Ethernet/IP are optimized differently:** Fibre Channel was born and raised during an era of explosive Ethernet/IP growth. FC succeeded because it was optimized for one premium use case: reliable, high-speed delivery of bursty storage traffic. Conversely, Ethernet and IP won as commoditized technologies that work anywhere, but, like SCSI, they come with lots of baggage.
- » **Dedicated NVMe fabrics are recommended:** Storage experts know that reference architectures from mission-critical storage vendors call for dedicated storage fabrics. This is why an FC fabric with HBAs is the low-risk choice compared to an Ethernet/IP fabric, which requires gambling on either RDMA-capable NICs or TCP offload engine technology (TOE).
- » **Dual FCP/FC-NVMe fabrics offer big advantages:** Dual protocol FCP/FC-NVMe fabrics offer ultra-low latency for working memory needs, as well as the as-needed, low-risk SCSI-to-NVMe migration you need for high value storage assets. Such fabrics also simplify your storage purchasing decisions during the SCSI-to-NVMe transition, a process that is likely to take years.



Hewlett Packard Enterprise

Get Ready for the NVMe Future

Accelerate performance, avoid storage networking bottlenecks and simplify fabric deployments with HPE 3PAR and Smart SAN

HPE offers a seamless path to emerging technologies such as NVMe Express (NVMe) and NVMe over Fibre Channel.

HPE 3PAR Storage recently completed a technology preview of 3PAR 3D Cache for flash hyper-acceleration using Storage Class Memory over NVMe.

HPE B-series 32Gb/s Gen6 FC switches are NVMe-FC Ready to meet the extreme performance demands of All Flash Storage

End-to-end “Smart SAN” technology orchestrates SAN zoning all from the array, reducing host provisioning up to 99% with HPE Smart SAN for 3PAR

Start your storage networking transformation today

Learn more

www.HPE.com/storage/B-series

www.HPE.com/storage/smartsan

Move into the NVMe over Fibre Channel future

NVMe (Non-Volatile Memory Express) is today's new massively parallel, ultra-low latency memory protocol. NVMe over Fibre Channel extends this blockbuster new protocol to the scale of enterprise storage. Fibre Channel, the premium datacenter fabric, can transport both NVMe and SCSI concurrently, giving you a low-risk NVMe adoption path that protects your high-value data assets. This book gets you started with NVMe over Fibre Channel, helps you create an adoption strategy, and shows you the way forward.

Inside...

- Adopt NVMe at your pace with low risk
- Deliver speed and reliability
- Increase IOPS with enhanced queuing
- Protect high-value storage assets
- Leverage concurrent SCSI and NVMe
- Simplify storage purchasing decisions
- Analyze and optimize with Fabric Vision

Rupin Mohan is an MIT Sloan Fellow, Director of R&D and Chief Technologist of Storage Networking at HPE Storage with three dozen patents filed.

Curt Beckmann is Principal Architect for NVMe over Fibre Channel at Broadcom, where he has also held roles from SDN chair to ASIC director.

Kip Hanson is a one-time IT geek turned freelance writer. He spends his days making difficult subjects easy to understand, while lamenting the loss of floppy disks and DOS prompts.

Go to **Dummies.com**®
for videos, step-by-step photos,
how-to articles, or to shop!

ISBN: 978-1-119-54112-7

Not For Resale

for
dummies®
A Wiley Brand



WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.