NetApp Verified Architecture

Modern SAN cloud-connected flash solution

NetApp, VMware, and Broadcom Verified Architecture Design Edition with vSphere 7.0 and MongoDB Community Edition 5.0

Modernize and future-proof your enterprise SAN; implement the fastest cloud-ready solution for mission-critical tier-1 enterprise applications and workloads

NetApp: Rodrigo Nascimento, Scott Lane
VMware: Naveen Krishnamurthy, David Morse
Broadcom: Jayamohan Kallickal, Dr. Naem Saafein, Jim Zucchero

February 2022 | NVA-1163-DESIGN | Version 1.0

Abstract

This NetApp® Verified Architecture was jointly designed and verified by NetApp, VMware, and Broadcom. It uses the latest Brocade, Emulex, and VMware vSphere technology solutions and NetApp all-flash storage, which sets a new standard for enterprise SAN storage and data protection that drives superior business value.

In partnership with

BROADCOM
EMULEX | BROCADE
FIBRE CHANNEL SOLUTIONS

vmware®

Foreword: Thoughts from Broadcom and VMware

When Nonvolatile Memory Express (NVMe) was introduced, customers invested in development of applications and/or deployed NVMe for isolated applications to see if the performance lived up to the claims being made regarding NVMe performance. There were also limited use cases and limited supported technologies where NVMe solutions could be deployed.

Although the performance has been proven time and time again (see the Demartek validation at https://www.netapp.com/us/media/ar-demartek-nvme.pdf), customers have implemented limited use case NVMe deployments in which results have met and usually exceeded expectations. Customers are now locking in NVMe as part of their go forward IT strategy, and NVMe is becoming the ubiquitous storage protocol of the future.

In modern data centers, it is rare to find dedicated solutions on isolated hardware deployments. The sheer physical size and total cost of that model makes it a very impractical scheme. The dedicated deployment model issue was resolved well over a decade ago with virtual machines (VMs). So why bring up the history lesson? Because we are seeing history repeat itself. With VMware supporting NVMe over Fibre Channel (NVMe/FC), the complete deployment model based on NVMe is now available (with Hypervisor, Guest OS on the compute front, Broadcom-Brocade Gen6/7 FC Fabrics and Broadcom Emulex host bus adapters (HBAs) on the protocol front, and the NetApp A-Series storage arrays on the target front). NVMe/FC solutions offer the performance and robustness of FC transport along with the ability to run FCP and FC-NVMe protocols concurrently on the same infrastructure. Customers can unlock the performance benefits of NVMe/FC for their mission-critical SAN applications in the deployment model to which they are accustomed, enabling NVMe to become the default approach rather than for isolated applications. This NetApp Verified Architecture addresses virtualized workloads with VMware vSphere by using NetApp NVMe/FC SAN arrays. This NetApp Verified Architecture covers and compares FCP and NVMe/FC.

It is important to note that NVMe, NVMe-oF is a quantum leap in storage technology, architected and designed to meet the performance and latency demands of business-critical applications and cloud native workloads. To realize the benefits of faster storage and the protocol, it is essential to upgrade to the NVMe optimized storage stack on both the host side (ESX) and the target side (Broadcom, NetApp). Without this, the bottleneck shifts from storage device to a layer within the I/O stack, which is not desirable.

NetApp continues to lead the market in delivering superior storage solutions for mission-critical enterprise SAN applications. Broadcom is proud to partner with NetApp, a company that continues to demonstrate the highest degree of excellence in its future-forward vision and technology, which will take customers into the next decade of enterprise SANs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Executive summary

NetApp Verified Architectures describe systems and solutions that are designed, tested, and documented to facilitate and improve customer deployments. These designs incorporate a wide range of technologies and products into a portfolio of solutions that NetApp has developed to help meet the business needs of customers.

This NetApp Verified Architecture provides a solution that modernizes your VMware vSphere SAN storage with a 32Gb NVMe-oF solution based on FC and cloud connectivity options, giving your company the best-in-class cloud-ready solution for mission-critical virtualized workloads. This report addresses the following:

- The challenge that organizations face today with data assets and infrastructure
- The solution to leverage disruptive future technology nondisruptively for your business today
- Ten good reasons to modernize your traditional SAN infrastructure
- A world-class modern SAN verified reference architecture
- NetApp recommended data protection solutions for this architecture

You can leverage NVMe/FC support now available from VMware vSphere with your current or new investments in NetApp ONTAP® modern SAN FC-based architectures. NVMe/FC offers a very simple and easy migration path to upgrade applications and workflows to use NVMe/FC instead of SCSI FCP. The rest of this report focuses on requirements, testing, and benefits of adopting the NVMe/FC protocol for your MongoDB VM workloads. You can easily extrapolate the information provided to other operating system and application VMs as well.

The challenge

The challenge today is how to rapidly and nondisruptively transform, modernize, and streamline critical data and IT services to scale and adapt to continuously evolving customer and business needs. At the same time, these services must be future-proof and cloud-ready so that an organization can maintain a competitive edge. This is particularly important for modern applications and workloads such as MongoDB when deployed in a highly scaled VMware infrastructure.

Background: According to IDC, by 2020, 50% of Forbes Global 2,000 companies will see most of their business depend on their ability to create digitally enhanced products, services, and experiences. Data is the lifeblood of future-thinking companies. The consequence of ever-increasing reliance on data will be a never-ending expansion in the size of the Global DataSphere. IDC forecasts the Global DataSphere to grow to 175ZB by 2025. As businesses contend with the perpetual growth of data, they must rethink how data is captured, preserved, and processed. Performance, economics, and endurance of data at scale are paramount. It is essential to provide a competitive platform to assist businesses that are dealing with data at scale. However, for many, their current IT infrastructure is not up to the task. The growing stress on the entire IT infrastructure to manage this overload of data interferes with the ability to quickly capitalize on the inherent value of the data.

The solution

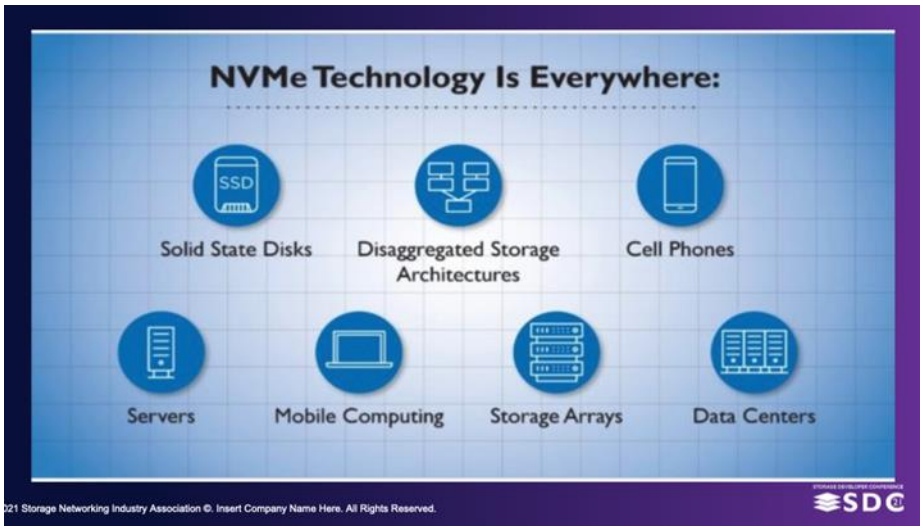
The good news is that just as flash transformed enterprise storage a few years ago, a new emerging technology, NVMe, is poised to transform enterprise storage again. NVMe is an emergent set of storage access and transport protocols that deliver the fastest response times yet for business-critical enterprise applications.

NVMe is a rich protocol optimized for nonvolatile memory media directly connected to CPU through the PCIe interface. The protocol capitalizes on multiple parallel and low latency data paths to flash devices, similar to the parallelism in CPUs, which reduces I/O overheads and results in higher performance. NVMe also consumes fewer CPU cycles than SCSI- and SATA-based protocols. The NVMe protocol is designed

to transport signals over a PCIe bus and removes the need for an I/O controller between the server CPU and the flash drives.

Ideally, there needs to be a way to extend the benefits of NVMe across the data center to multiple servers and multiple applications—actively and dynamically scaling NVMe to meet demand (Figure 1). This would provide the benefits of high performance NVMe along with the best features of centralized, shared storage we typically associate with SANs, but without the performance compromises we have all had to accept. NVMe-oF is a transformational technology because it impacts the data center strategies of today and tomorrow. It consumes significantly less resources than the legacy protocols on both the initiator and the target side, allowing better scalability.

Figure 1) The ubiquity of the NVMe protocol.*



*Source: https://nvmexpress.org/wp-content/uploads/NVMe-2.0-Infographic_Email-File-1.pdf

NVMe is about to provide a major performance boost for enterprise data storage systems. But this time, the transformative effect could be greater still because NVMe is not just a storage specification. The broader NVMe-oF protocol specifies how to encapsulate NVMe inside a variety of network and fabric protocols, such as remote direct memory access (RDMA), Ethernet, TCP, and InfiniBand. These protocols define and enable the use of NVMe across the entire data path, from server to storage system, enabling superior performance and lower latency than traditional technologies can deliver. NVMe/FC replaces the traditional SCSI commands with NVMe commands inside the FC frame—no changes are required for the application to use NVMe. From the applications perspective, NVMe/FC is just another block protocol and can be adopted nondisruptively. Additionally, both NVMe/FC and FC can use the same FC components (HBAs, fiber-optic cables, switches, and storage target HBAs) concurrently, which makes the migration from FC to NVMe/FC very simple. It does not require a cutover like moving from FC to another non-FC protocol, such as iSCSI or RDMA over Converged Ethernet (RoCE), does. This makes migrating from FC to NVMe-oF a quantum leap in the storage technology; it is architected and designed to meet the performance and latency demands of business-critical applications. To realize the benefits of faster storage and the associated protocol, VMware supports NVMe-FC, starting with vSphere 7.0.

As a result, CxOs now have the opportunity, and the challenge, to harness the power of data through digital transformation and modernization. They can also use these emerging best-in-class technologies from world-class industry leaders NetApp, VMware, and Broadcom's Brocade and Emulex divisions to do the following:

- Rapidly deliver and monetize vital digital data services
- Accelerate the pace of innovation

- Acquire, grow, and retain market share
- Improve customer service and experience
- Maximize return on investment
- Protect and secure customers and critical data
- Increase agility and response to changing business needs

Ten good reasons to modernize your SAN with NetApp, VMware, and Broadcom

This report describes a verified, unified modern SAN solution reference architecture that is designed by the industry leaders NetApp, VMware, and Broadcom, with a first-to-market enterprise NVMe/FC solution. NetApp, VMware, and Broadcom provide an end-to-end NVMe solution, from host to storage controller, that can help you realize the promise and the benefits of NVMe technology right now. With a system that yields the fastest access, management, and utilization of critical data, you can accelerate your time to innovation and leverage the following benefits:

- **Digitally transform critical business applications**. Enable the existing and next generation of your critical applications, ready for analytics, artificial intelligence (AI), and machine learning (ML) capabilities.
- **Harness the power of the hybrid cloud**. Cloud-enable your IT services to get the benefits of on-premises storage with the flexibility of public cloud. The NVMe solution is an ideal choice for hybrid cloud. Hybrid cloud offers the benefits of both public and private clouds and takes advantage of existing architecture in a data center.
- **Get a best-in-class solution for enterprise SAN**. Strengthen your competitive advantage by partnering with the best-in-class flash, virtualization, SAN, fabric, and HBA leaders.
- **Significantly simplify operations**. Improve IT responsiveness through simplification of SAN management while ensuring predictable performance.
- **Modernize and get significant cost savings**. Improve shareholder value by attaining a 30% reduction in database licensing costs, 80% to 90% or more reduction in data center floor space, 50% to 90% or more reduction in power and cooling, and 50% to 80% reduction in labor costs, as detailed in NVA-1136-DESIGN: NetApp and Broadcom Modern SAN Cloud-Connected Flash Solution. Additionally, VMs will be more efficient with higher VM densities per server, reducing infrastructure capital expenditure (capex), operational expenditure (opex) costs, and simplifying IT
- **Future-proof your SAN environment**. Nondisruptively adopt disruptive performance and technology advancements when you are ready.
- **Rapidly deliver core IT services**. Take advantage of an open platform that supports leading DevOps toolsets to vastly reduce the time to value for development.
- **Don’t compromise on availability**. Get 99.9999% availability (backed by several IDC audits, discussed here) and enterprise-grade disaster recovery capabilities.
- **Improve the customer experience**. Best-in-class data protection, the most efficient and scalable storage, and the most flexible IT infrastructure. Accelerate performance, enable instant application cloning, and enable granular data recovery to improve the user experience.
- **Get next-generation enterprise data management**. Bring the value of industry-leading innovation together with enterprise availability to deliver the next generation of your SAN environment.

The architecture

This NetApp, VMware, and Broadcom modern SAN NetApp verified reference architecture for VMware vSphere includes the following key NetApp, VMware, and Broadcom technologies:

- FCP
- NVMe/FC
- Sixth and seventh generation host and fabric technology

The performance benefits accrue as you adopt these technologies. Adopt all of them and get game-changing performance benefits with end-to-end visibility through Fabric Vision technology. In the future, you will be able to add storage-class memory and persistent memory so that you can realize further increased performance.

Program summary

This report is part of the modern SAN best practices program that provides test and validated design and configuration recommendations for next-generation NVMe-powered fabrics. This report is part of a series that covers the deployment of popular enterprise applications. This program is a collaboration between NetApp, VMware, and Broadcom's Brocade and Emulex divisions. The information is designed to support IT organizations that want to upgrade their existing SAN architectures to next-generation NVMe-based fabrics to meet the low-latency, high-performance requirements of modern and future enterprise apps.

This report describes the system and solution that were designed, tested, and documented to facilitate modern SAN deployments in a virtualized infrastructure with tier-1 mission-critical enterprise applications and workloads. These designs incorporate a wide range of technologies and products into a portfolio of solutions that NetApp has developed to meet the business needs of customers like you. This report also describes the design choices and the best practices for this shared infrastructure platform. These design considerations and recommendations are not limited to the specific components that are described in this report; they also apply to other versions of components.

Solution overview

NetApp, VMware, and Broadcom modern SAN solution benefits

This solution comprises Brocade Gen 7 Fibre Channel Switches, Emulex Gen 7 FC HBAs, VMware vSphere 7.0, and NetApp AFF storage systems. It is a predesigned, best practice configuration that is built on FC protocol but compares the performance benefits of NVMe/FC versus SCSI FCP (FCP—FC using SCSI command sets) on the latest NetApp, VMware, and Broadcom technologies.

This solution delivers a baseline configuration and can also be sized and optimized to accommodate many different use cases and requirements. It supports tight integration with virtualized and cloud infrastructures and data protection, making it the logical choice for long-term investment.

The solution delivers operational efficiency and consistency with the versatility to meet various SLAs and IT initiatives, including the following:

- Application rollouts or migrations
- Business continuity
- Cloud delivery models (public, private, and hybrid) and service models (infrastructure as a service [IaaS], platform as a service [PaaS], and software as a service [SaaS])
- Asset consolidation and virtualization
- Data center consolidation and footprint reduction

NetApp, VMware, and Broadcom have thoroughly validated and verified this solution architecture and its many use cases. They have also created a portfolio of detailed documentation, information, presale and post-sale services, and references to assist you in transforming your data center to this shared infrastructure model. This portfolio includes, but is not limited to, the following items:

- Best practice architectural design
- Workload sizing and scaling guidance
- Implementation and deployment instructions

- Technical specifications (rules for what is and what is not a reference architecture)
- Frequently asked questions (FAQs)
- NetApp, VMware, and Broadcom jointly validated designs that focus on various use cases

Target audience

The target audience for this NetApp Verified Architecture document includes the following groups:

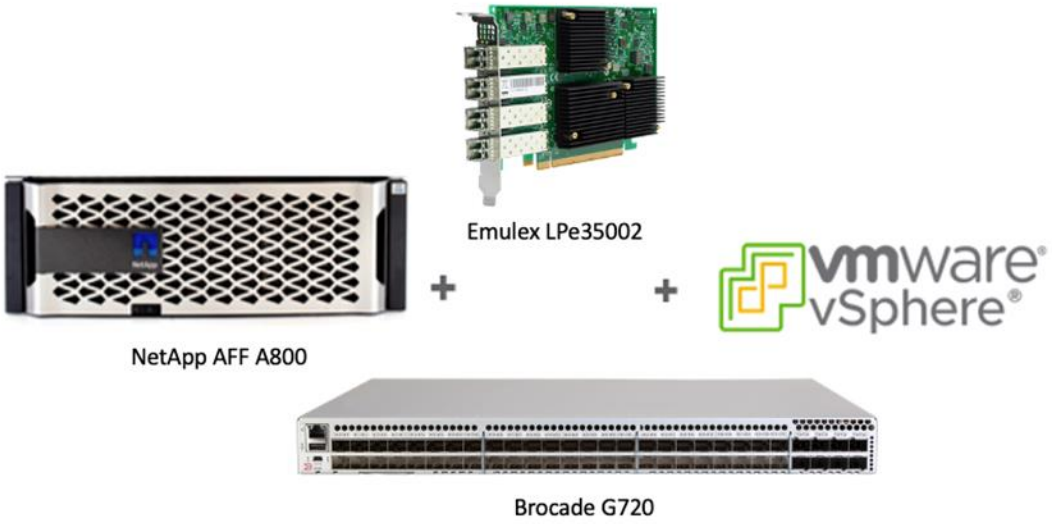
- The CIO, CTO, and CFO, who can benefit from the executive summary, use case examples, ROI and TCO information, and information about future strategies
- Business information officers, who can learn new ways to serve line-of-business owners with benefits from modern technologies
- Architects, administrators, and solutions engineers who are responsible for designing and deploying infrastructure for enterprise mission-critical applications
- Database administrators, who require new data management capabilities and performance to serve evolving data requirements
- Application owners, who need real-time, lower-latency data to feed current and newer generations of applications
- Virtualization architects and administrators who are responsible for designing, deploying, and managing virtualized enterprise mission-critical environments
- Data architects, who require platforms that are designed to enable more real-time analytics and to serve the AI and ML requirements that new workloads need
- Cloud architects, who must harness the power of the hybrid cloud and leverage core and cloud-native solutions
- Backup administrators, who must protect data and leverage new innovations to make data protection seamless and nondisruptive to the business
- Service delivery managers, who must meet SLAs and service-level objectives (SLOs) that require IT infrastructure and solutions that promote consistent and predictable results

Solution technology

In this report, we focus on virtualized workloads. We assume some numbers for typical inefficient utilization rates that we see on legacy storage. We also factor in our 2:1 to 4:1 storage efficiency and workload multitenancy benefits when consolidating multiple traditional SAN storage systems into a NetApp AFF A800 configuration.

Figure 2 shows the component families of the architecture. Implementation of this solution should reduce the footprint, management overhead, maintenance spending, and power and cooling, and it should improve service availability and performance.

Figure 2) Component families of the NetApp, VMware, and Broadcom joint architecture.



Most of today's all-flash arrays are deployed on low-risk, multiqueue-capable, deep-queue-rich, and proven FC-based storage networks, with their robust scalable fabric services and credit-based flow control. Because of their reliability and deterministic performance, FC fabrics serve as the most widely implemented storage network infrastructure for mission-critical applications. Because little change is required in the standards to implement NVMe/FC, the introduction of NVMe/FC along with existing storage is easy, seamless, and noninvasive. And because NVMe/FC can use the same infrastructure components concurrently with other FC traffic, it is easy to migrate workloads at the pace that works for your organization. NVMe/FC also allows the efficient transfer of NVMe commands and structures end to end with no translations.

The world's first end-to-end virtualization enterprise NVMe/FC solution with vSphere 7.0, NetApp All-Flash array, Broadcom's Emulex LPe35002 HBAs, and Brocade Gen 7 Fibre Channel network is purpose-built for tomorrow's mission-critical workloads by leveraging today's infrastructure.

New innovations in storage technology are disrupting the data center industry. The introduction of faster media types and more efficient mechanisms to access those media across well-defined various infrastructures is unlocking unprecedented speeds, lower latencies, and dramatic improvements in system and application efficiency and performance.

The current testing uses available data center solutions, specifically with Broadcom NVMe/FC (and other hardware). You can also use Gen 6 switches and other NetApp controllers, such as the AFF A250, A300, A320, A400, A700, and A700s configurations.

NVMe

The NVMe specification is designed to leverage NVMEM in all kinds of compute environments, from mobile phones to web scale service providers. It adds massive I/O path parallelization (65,535 I/O queues, each with a queue depth of up to 64Kb outstanding I/O operations), making communication with storage systems massively parallel. Because of lower protocol overhead and lower-latency connectivity between servers and storage devices, this parallelization provides greater bandwidth.

The massive number of queues and the huge queue depths that each can support enable today's storage and servers to use the increasingly large numbers of cores and memory they have. This capability accelerates processing of I/O threads by spreading the processing across multiple CPU cores. This

attribute is critical to bring together traditional enterprise applications with real-time analytics workloads, enabling new digital services for the modern enterprise.

NetApp technology is built for the future. With the industry's only unified data management platform that supports SAN or NAS, all-flash, software-defined, hybrid, and cloud storage, it supports both existing (traditional) and emerging applications (for example, NoSQL databases and AI). These features and capabilities are all part of the data fabric powered by NetApp. NetApp systems support scaling (up and out) dynamically in seconds or minutes, instead of taking hours or days. And you can allocate applications to where they run best across your data fabric delivered by NetApp, whether it's on the premises or in the cloud. And to maximize performance and reduce overall storage cost, NetApp FabricPool enables you to move data automatically between AFF storage solutions and Simple Storage Service (S3) and cloud storage tiers.

Along with the Broadcom's Brocade and Emulex divisions, which are leaders in the SAN fabric space, NetApp is the first to market with an end-to-end enterprise NVMe/FC solution over a 32Gbps FC fabric. With this joint solution, you can enable and accelerate this digital transformation for your enterprise—now.

Brocade Gen 7 Fibre Channel platforms

Broadcom's Brocade has been the leading provider of storage networking solutions worldwide for more than 20 years, supporting the mission-critical systems and business-critical applications of most large enterprises. Brocade networking solutions help organizations achieve their critical business initiatives as they transition to a world where applications and information can reside anywhere. Today, Brocade is extending its proven data center expertise across the entire network with open, application-optimized, and efficient solutions that are built for consolidation and unmatched business agility.

Brocade Gen 7 provides the accelerated foundation that modern data centers need to fully realize the benefits of the latest storage technology today and possibly for the next decade.

The seventh generation of FC is aimed at satisfying the needs of growing deployments of flash storage and hyperscale virtualization, and new high-speed data center architectures run twice as fast as the current Gen 6 standard 32Gbitps and have 50% lower latency.

Brocade's Gen 7 platforms are equipped with higher-performing hardware to unleash NVMe technology, and they can discover and produce comprehensive telemetry data across the fabric. They analyze and take actions based on that data to optimize the storage network automatically.

The self-learning and self-optimizing capabilities of Brocade Gen 7 automatically ensure that storage performance and other SAN-related activities are optimized. The collected telemetry data is composed of millions of data points. The Brocade technology leverages it automatically to learn application flows, and then creates a baseline of each application's performance from end to end across the fabric to detect when something is abnormal. With Virtual Machine Identification (VMID) tagging, telemetry data is collected on a frame-by-frame basis to the VM. The Brocade systems then provide I/O profiles at the application level, rather than simply on a port basis. The analysis happens within the ASIC integrated circuit to ensure that performance is not impaired.

Gen 7 Fabric Notifications is a mechanism that provides end devices (hosts and storage) with more information about events in the fabric to automatically remediate congestion issues and response time diagnosis. They include notifications regarding link integrity issues, delivery notification issues, and congestion issues. The Brocade Fabric Notifications solution combines hardware (Brocade FC switching ASIC), software (Fabric Operating System), and management (MAPS and SANnav) to form a SAN-wide solution for impairment detection, notification, and remediation. Problems are logged, including operational information concerning affected flows, location, timestamp, and reason for easy debugging and troubleshooting.

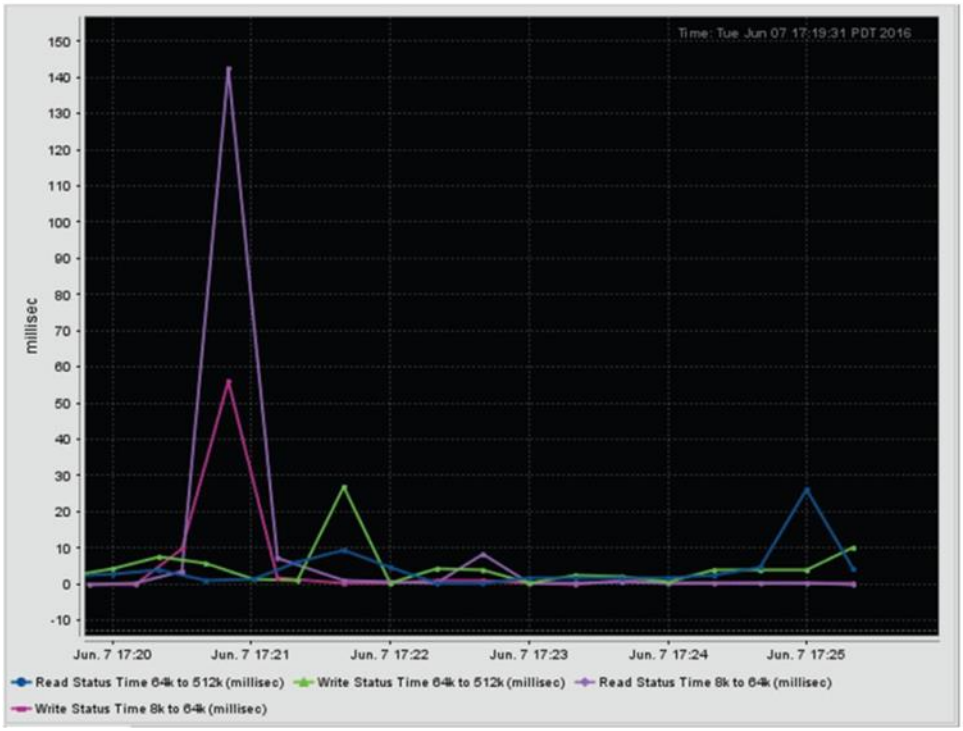
In partnership with Brocade, NetApp now supports Fabric Notifications to optimize storage traffic throughput by preventing congestion and application performance degradation anomalies such as marginal paths, link integrity issues, oversubscription, and credit stall.

The NVMe/FC feature supports both NVMe-oF and SCSI over FC protocols concurrently. Your organization can seamlessly integrate Brocade FC networks with the next generation of low-latency flash storage, without a disruptive rip and replace.

Brocade's IO Insight is the industry's first integrated network sensor tool that proactively and nonintrusively monitors real-time storage I/O health and performance statistics for both SCSI and NVMe traffic from any device port on a Gen 6 or Gen 7 FC platform. IO Insight then applies this information within an intuitive, policy-based monitoring and alerting suite to quickly identify the root cause of problems at the storage or at the VM tier. This level of granularity enables quick identification of degraded application performance at the host VM and storage tiers, reducing time to resolution. Furthermore, Brocade VM Insight further extends IO Insight by enabling visibility of I/O statistics at the level of the individual VM. By applying the interpretation of FC standards based VM identification to IO Insight analysis, VM Insight enables the storage administrator to pinpoint issues not just on a physical server, but down to the specific VM workload. VM Insight further enables I/O profiles for individual VMs to be baselined and monitored over time with Fabric Vision MAPS for continual performance monitoring and optimization.

IO Insight proactively monitors individual host and storage devices to gain deeper insight into the performance of the network to maintain SLA compliance. It also obtains total I/O operations, first response time, I/O latency (Exchange Completion Time [ECT]) and outstanding I/O performance metrics for a specific host or storage device, in order to diagnose I/O operational issues. Lastly, it enables tuning of device configurations with integrated I/O metrics to optimize storage performance. Preventative actions, such as admin notifications and port fencing, can be defined to avoid greater negative impact, as shown in Figure 3.

Figure 3) IO Insight metrics displayed in a Brocade Network Advisor real-time performance graph.



Also, to meet the security requirements challenge, Brocade is accelerating the secure digital transformation of data centers. Brocade products support a wide range of authentication, encryption, and management tools to protect fabrics and data from unauthorized access:

- **Authentication**. Authentication protocol support includes CHAP, DH-CHAP, FCAP, IKE, IPsec, RADIUS, TACACS+, and P-EAP/MS-CHAP for RADIUS.

- **Encryption (AES/3-DES**). Brocade provides AES-128 and AES-256 encryption and 168-bit 3-DES encryption for IP links on extension products and management connections. Brocade also supports AES and 3-DES with IPsec. These solutions provide high-performance encryption and compression.
- In-flight encryption over ISLs: Brocade Gen 7 Fibre Channel directors and Brocade G720 switches support in-flight encryption for traffic over ISLs to minimize the risk of unauthorized access to data within the data center and over long-distance links. Data-at-rest and data-in-flight encryption are complementary technologies that serve different purposes, and you might require each to achieve regulatory compliance.
- **Secure Boot**. A switch validates the integrity and authenticity of the FOS boot image to establish a hardware-based root of trust through the manufacturing supply chain.
- Multiple layers of challenge to an attacker to provide the best protection against all types of threats.

The NVMe/FC feature supports both NVMe-oF and SCSI over FC protocols concurrently. Your organization can seamlessly integrate Brocade Gen 7 Fibre Channel networks with the next generation of low-latency flash storage, without a disruptive rip and replace.

VMware

The popularity of NVMe continues to increase, and with good reason. With its low latency and high throughput, NVMe offers the industry additional benefits over traditional storage. VMware is always striving to increase storage performance and efficiency and has been working on NVMe technologies from the start. With the announcement of vSphere 7, VMware has added support for NVMe-oF, enabling customers to connect to external NVMe arrays over the wire. With the initial release, FC and RDMA using RoCEv2 protocols are supported with vSphere 7.

This is an exciting announcement and many of VMware's storage partners have been part of the development of NVMe-oF. Being able to access external NVMe arrays essentially as DAS with the performance and throughput benefits of NVMe solid-state drives (SSDs), gives infrastructures numerous advantages.

Emulex Gen 7 FC HBAs

Emulex FC HBAs by Broadcom are designed to meet the demanding performance, reliability, and management requirements of modern networked storage systems that use high-performance and low-latency SSDs.

The Emulex LPe35000/36000 FC HBAs with Dynamic Multi-core Architecture delivers unparalleled performance and more efficient port usage than other HBAs by applying all ASIC resources to any port that needs it, providing industry-leading 32Gb FC performance of over 5 million IOPs and over 11 million IOPS for 64Gb FC. The LPe35002/36002 series delivers 12800MBps (two 32Gb FC ports) full duplex, and three times better hardware latency than previous generation adapters. Emulex Gen 7 HBAs also support NVMe/FC and SCSI FCP concurrently, providing investment protection and enabling data centers to transition to end-to-end NVMe over FC SANs at their own pace.

The secure firmware update feature protects and ensures the authenticity of device firmware. Emulex Gen 7 FC HBAs are NVMe/FC-enabled, delivering up to 55% lower insertion latency for NVMe/FC than SCSI over FC. And for investment protection, these FC HBAs also support both NVMe/FC and SCSI over FC protocols concurrently.

ONTAP tools for VMware vSphere

NetApp offers several standalone software tools that you can use together with ONTAP and vSphere to manage your virtualized environment. The Virtual Storage Console (VSC) is a vCenter plug-in that simplifies storage management and efficiency features, enhances availability, and reduces storage costs and operational overhead, whether using SAN or NAS. VSC uses best practices for provisioning datastores and optimizes ESXi host settings for NFS and block storage environments. For all these

benefits, it is recommended as a best practice when using vSphere with systems running ONTAP software. It includes both a VSC server appliance and UI extensions for vCenter.

Other ONTAP tools not specifically used in this NVA include the NetApp NFS Plug-In for VMware VAAI to use VAAI offload features with NFS storage, the VASA Provider for ONTAP to enable VMware Virtual Volumes (vVols) support, and the Storage Replication Adapter used together with VMware Site Recovery Manager (SRM) to manage data replication between production and disaster recovery sites.

Brocade SAN Health

Your storage architecture is critical for your business agility and success. Brocade's free SAN Health tool delivers clear insights into performance, inventory, and bottlenecks to optimize your SAN infrastructure and to align it with your business needs. This hardware-agnostic and easy-to-run tool generates personalized storage network performance and inventory reports to help you prevent issues, avoid application downtime, reduce troubleshooting time to resolution, and improve capacity planning and productivity. Figure 4 shows the components of the SAN Health tool, and Figure 5 shows how to use it.

Contact your NetApp account team (link) to sign up for a SAN Health check or to get a copy of the NetApp-branded SAN Health tool.

Figure 4) SAN Health report title page and table of contents.

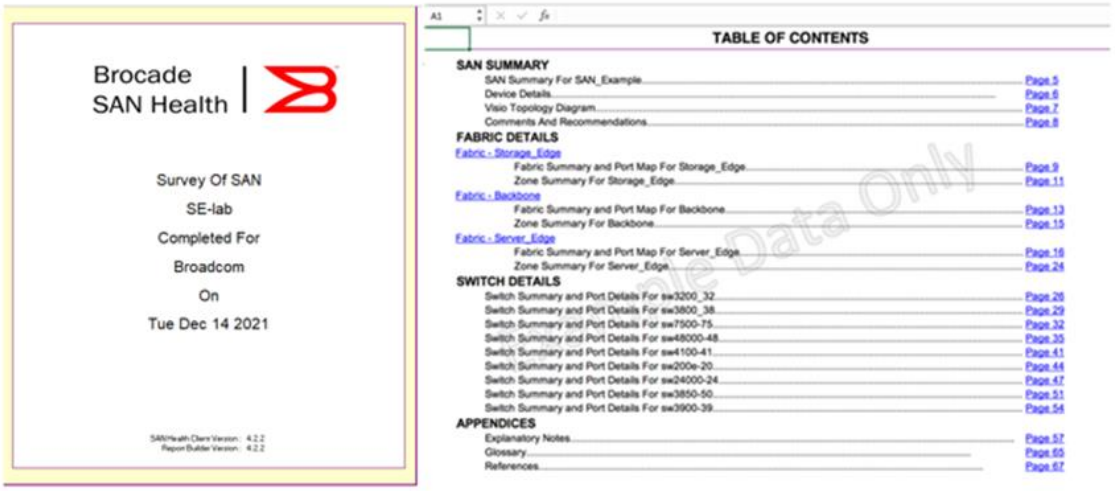
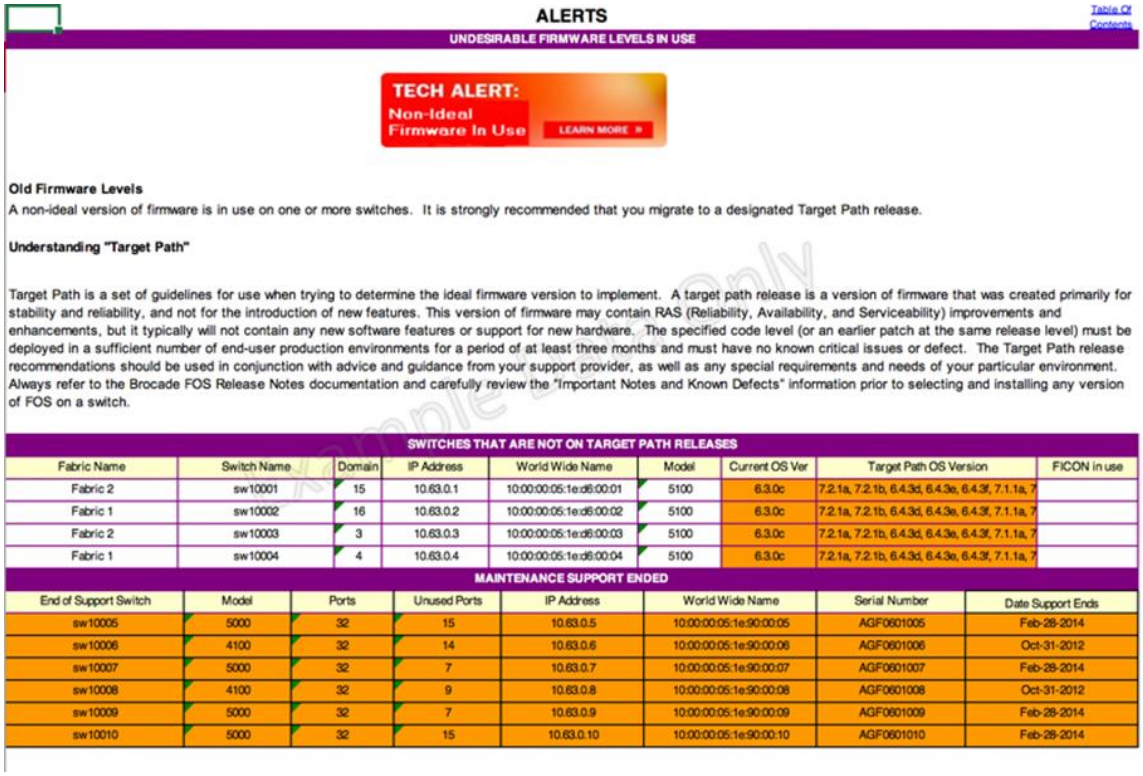


Figure 5) SAN Health color-coded alerts and warnings.



ALERTS

Table Of Contents

UNDESIRABLE FIRMWARE LEVELS IN USE

TECH ALERT: Non-Ideal Firmware In Use LEARN MORE »

Old Firmware Levels

A non-ideal version of firmware is in use on one or more switches. It is strongly recommended that you migrate to a designated Target Path release.

Understanding "Target Path"

Target Path is a set of guidelines for use when trying to determine the ideal firmware version to implement. A target path release is a version of firmware that was created primarily for stability and reliability, and not for the introduction of new features. This version of firmware may contain RAS (Reliability, Availability, and Serviceability) improvements and enhancements, but it typically will not contain any new software features or support for new hardware. The specified code level (or an earlier patch at the same release level) must be deployed in a sufficient number of end-user production environments for a period of at least three months and must have no known critical issues or defect. The Target Path release recommendations should be used in conjunction with advice and guidance from your support provider, as well as any special requirements and needs of your particular environment. Always refer to the Brocade FOS Release Notes documentation and carefully review the "Important Notes and Known Defects" information prior to selecting and installing any version of FOS on a switch.

SWITCHES THAT ARE NOT ON TARGET PATH RELEASES

Fabric Name	Switch Name	Domain	IP Address	World Wide Name	Model	Current OS Ver	Target Path OS Version	FICON in use
Fabric 2	sw10001	15	10.63.0.1	10:00:00:05:1e:d6:00:01	5100	6.3.0c	7.2.1a, 7.2.1b, 6.4.3d, 6.4.3e, 6.4.3f, 7.1.1a, 7	
Fabric 1	sw10002	16	10.63.0.2	10:00:00:05:1e:d6:00:02	5100	6.3.0c	7.2.1a, 7.2.1b, 6.4.3d, 6.4.3e, 6.4.3f, 7.1.1a, 7	
Fabric 2	sw10003	3	10.63.0.3	10:00:00:05:1e:d6:00:03	5100	6.3.0c	7.2.1a, 7.2.1b, 6.4.3d, 6.4.3e, 6.4.3f, 7.1.1a, 7	
Fabric 1	sw10004	4	10.63.0.4	10:00:00:05:1e:d6:00:04	5100	6.3.0c	7.2.1a, 7.2.1b, 6.4.3d, 6.4.3e, 6.4.3f, 7.1.1a, 7	

MAINTENANCE SUPPORT ENDED

End of Support Switch	Model	Ports	Unused Ports	IP Address	World Wide Name	Serial Number	Date Support Ends
sw10005	5000	32	15	10.63.0.5	10:00:00:05:1e:90:00:05	AGF0601005	Feb-28-2014
sw10006	4100	32	14	10.63.0.6	10:00:00:05:1e:90:00:06	AGF0601006	Oct-31-2012
sw10007	5000	32	7	10.63.0.7	10:00:00:05:1e:90:00:07	AGF0601007	Feb-28-2014
sw10008	4100	32	9	10.63.0.8	10:00:00:05:1e:90:00:08	AGF0601008	Oct-31-2012
sw10009	5000	32	7	10.63.0.9	10:00:00:05:1e:90:00:09	AGF0601009	Feb-28-2014
sw10010	5000	32	15	10.63.0.10	10:00:00:05:1e:90:00:10	AGF0601010	Feb-28-2014

Using the SAN Health tool eliminates whiteboard-base and gives users clear high quality reporting that they can use as the basis for SAN documentation and for management reporting. This report, and the Visio topology diagram (Figure 6), accelerate discovery and enable users to proactively fix problems.

Figure 6) SAN Health Visio fabric layout.

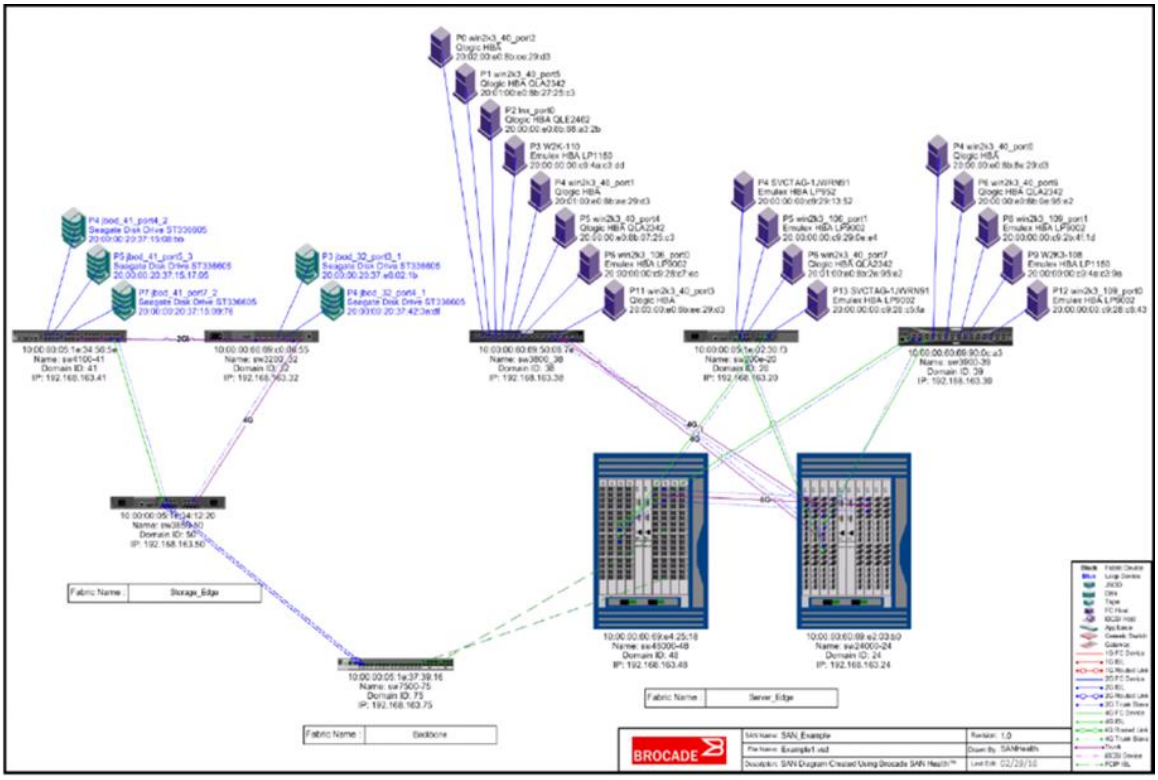
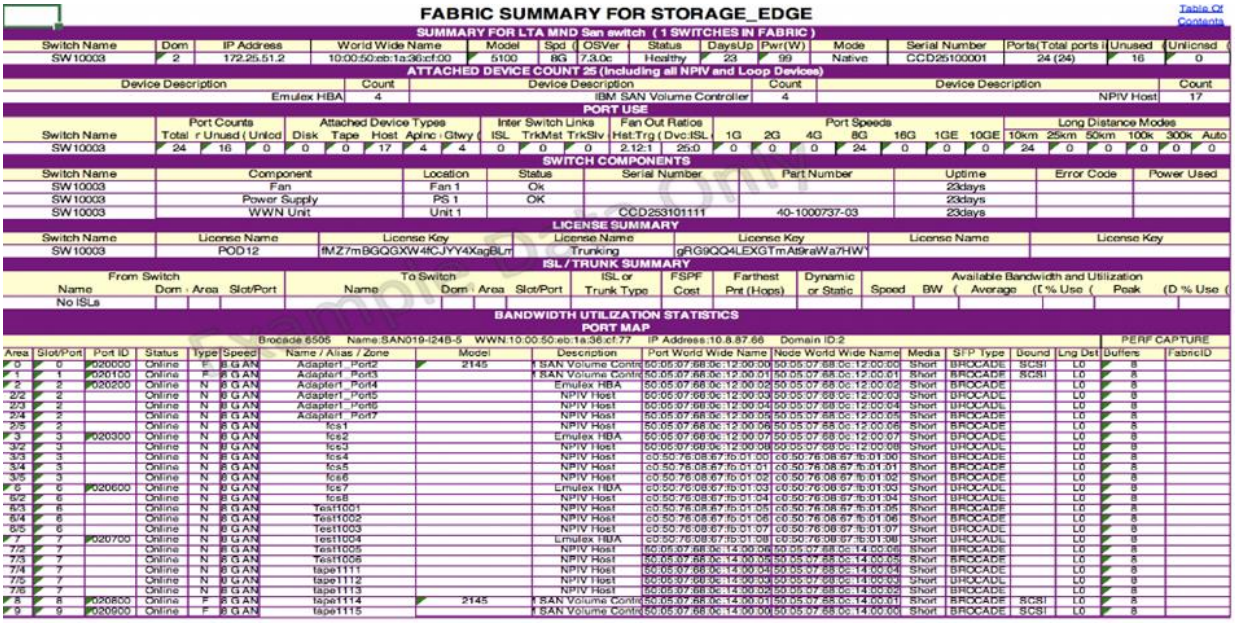


Figure 7 illustrates SAN Summary and how it offers an overview of the audited switches and helps to identify environmental factors where NVMe can enhance workload performance, such as high traffic, producing an insight into balanced traffic.

Figure 7) SAN Health Summary showing current configuration, best practice, and health and configuration checks.



FABRIC SUMMARY FOR STORAGE_EDGE

Table Of Contents

SUMMARY FOR LTA MND San switch (1 SWITCHES IN FABRIC)

Switch Name	Dom	IP Address	World Wide Name	Model	Spd	OSVer	Status	DaysUp	Pwr(W)	Mode	Serial Number	Ports(Total ports i	Unused	Unlicnsd
SW10003	2	172.25.51.2	10:00:50:eb:1a:36:cf:00	5100	8G	7.3.0c	Healthy	23	99	Native	CCD2510000 1	24 (24)	16	0

ATTACHED DEVICE COUNT 25 (Including all NPIV and Loop Devices)

Device Description	Count	Device Description	Count	Device Description	Count
Emulex HBA	4	IBM SAN Volume Controller	4	NPIV Host	17

PORT USE

Switch Name	Port Counts Total	r Unused (	Unlcd	Attached Device Types Disk	Tape	Host	Aplnc	Gtwy (	Inter Switch Links ISL	TrkMst	TrkSlv	Fan Out Ratios Hst:Trg (	Dvc:ISL (	Port Speeds 1G	2G	4G	8G	16G	1GE	10GE	Long Distance Modes 10km	25km	50km	100k	300k	Auto
SW10003	24	16	0	0	0	17	4	4	0	0	0	2.12:1	25:0	0	0	0	24	0	0	0	24	0	0	0	0	0

SWITCH COMPONENTS

Switch Name	Component	Location	Status	Serial Number	Part Number	Uptime	Error Code	Power Used
SW10003	Fan	Fan 1	Ok			23days		
SW10003	Power Supply	PS 1	OK			23days		
SW10003	WWN Unit	Unit 1		CCD2531011 11	40-1000737-03	23days		

LICENSE SUMMARY

Switch Name	License Name	License Key	License Name	License Key	License Name	License Key
SW10003	POD12	fMZ7mBGQGXW4fCJYY4XagBLm	Trunking	gRG9QQ4LEXGTmAt9raWa7HWY		

ISL / TRUNK SUMMARY

From Switch Name	Dom	Area	Slot/Port	To Switch Name	Dom	Area	Slot/Port	ISL or Trunk Type	FSPF Cost	Farthest Pnt (Hops)	Dynamic or Static	Available Bandwidth and Utilization Speed	BW (	Average ([% Use (	Peak (D % Use (
No ISLs															

BANDWIDTH UTILIZATION STATISTICS

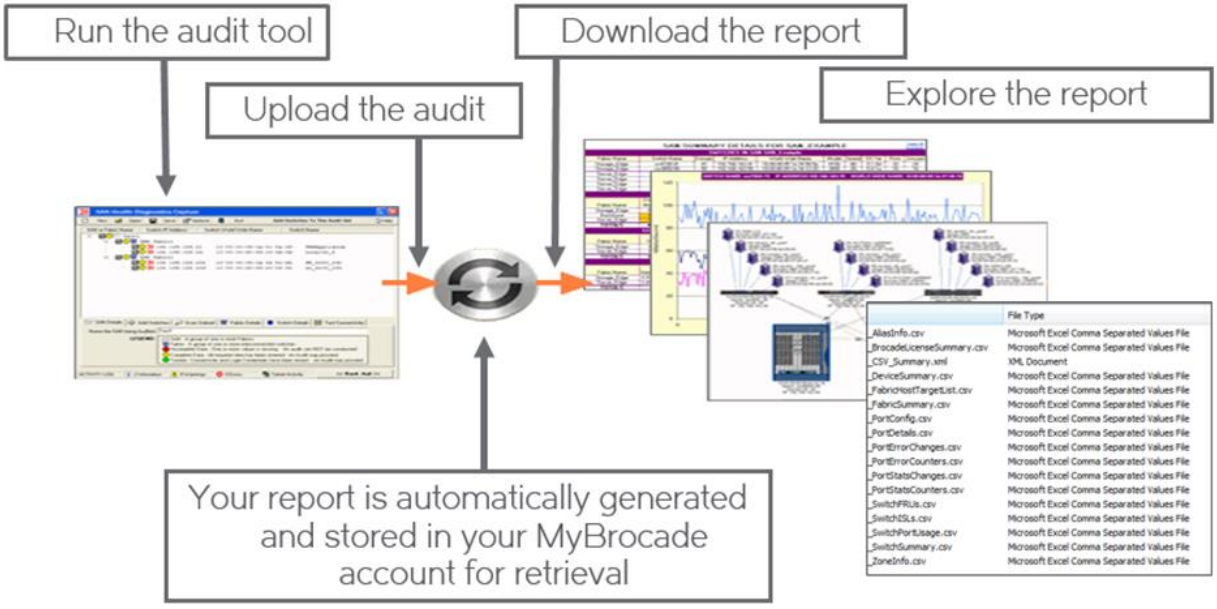
PORT MAP

Brocade 6505 Name:SAN019-I24B-5 WWN:10:00:50:eb:1a:36:cf:77 IP Address:10.8.87.66 Domain ID:2 PERF CAPTURE

Area	Slot/Port	Port ID	Status	Type	Speed	Name / Alias / Zone	Model	Description	Port World Wide Name	Node World Wide Name	Media	SFP Type	Bound	Lng Dst	Buffers	FabricID
0	0	020000	Online	F	8 GAN	Adapter1_Port2	2145	SAN Volume Contro	50:05:07:68:0c:12:00:00	50:05:07:68:0c:12:00:00	Short	BROCADE	SCSI	L0	8	
1	1	020100	Online	F	8 GAN	Adapter1_Port3		SAN Volume Contro	50:05:07:68:0c:12:00:01	50:05:07:68:0c:12:00:01	Short	BROCADE	SCSI	L0	8	
2	2	020200	Online	N	8 GAN	Adapter1_Port4		Emulex HBA	50:05:07:68:0c:12:00:02	50:05:07:68:0c:12:00:02	Short	BROCADE		L0	8	
2/2	2		Online	N	8 GAN	Adapter1_Port5		NPIV Host	50:05:07:68:0c:12:00:03	50:05:07:68:0c:12:00:03	Short	BROCADE		L0	8	
2/3	2		Online	N	8 GAN	Adapter1_Port6		NPIV Host	50:05:07:68:0c:12:00:04	50:05:07:68:0c:12:00:04	Short	BROCADE		L0	8	
2/4	2		Online	N	8 GAN	Adapter1_Port7		NPIV Host	50:05:07:68:0c:12:00:05	50:05:07:68:0c:12:00:05	Short	BROCADE		L0	8	
2/5	2		Online	N	8 GAN	fcs1		NPIV Host	50:05:07:68:0c:12:00:06	50:05:07:68:0c:12:00:06	Short	BROCADE		L0	8	
3	3	020300	Online	N	8 GAN	fcs2		Emulex HBA	50:05:07:68:0c:12:00:07	50:05:07:68:0c:12:00:07	Short	BROCADE		L0	8	
3/2	3		Online	N	8 GAN	fcs3		NPIV Host	50:05:07:68:0c:12:00:08	50:05:07:68:0c:12:00:08	Short	BROCADE		L0	8	
3/3	3		Online	N	8 GAN	fcs4		NPIV Host	c0:50:76:08:67:fb:01:00	c0:50:76:08:67:fb:01:00	Short	BROCADE		L0	8	
3/4	3		Online	N	8 GAN	fcs5		NPIV Host	c0:50:76:08:67:fb:01:01	c0:50:76:08:67:fb:01:01	Short	BROCADE		L0	8	
3/5	3		Online	N	8 GAN	fcs6		NPIV Host	c0:50:76:08:67:fb:01:02	c0:50:76:08:67:fb:01:02	Short	BROCADE		L0	8	
6	6	020600	Online	N	8 GAN	fcs7		Emulex HBA	c0:50:76:08:67:fb:01:03	c0:50:76:08:67:fb:01:03	Short	BROCADE		L0	8	
6/2	6		Online	N	8 GAN	fcs8		NPIV Host	c0:50:76:08:67:fb:01:04	c0:50:76:08:67:fb:01:04	Short	BROCADE		L0	8	
6/3	6		Online	N	8 GAN	Test1001		NPIV Host	c0:50:76:08:67:fb:01:05	c0:50:76:08:67:fb:01:05	Short	BROCADE		L0	8	
6/4	6		Online	N	8 GAN	Test1002		NPIV Host	c0:50:76:08:67:fb:01:06	c0:50:76:08:67:fb:01:06	Short	BROCADE		L0	8	
6/5	6		Online	N	8 GAN	Test1003		NPIV Host	c0:50:76:08:67:fb:01:07	c0:50:76:08:67:fb:01:07	Short	BROCADE		L0	8	
7	7	020700	Online	N	8 GAN	Test1004		Emulex HBA	c0:50:76:08:67:fb:01:08	c0:50:76:08:67:fb:01:08	Short	BROCADE		L0	8	
7/2	7		Online	N	8 GAN	Test1005		NPIV Host	50:05:07:68:0c:14:00:06	50:05:07:68:0c:14:00:06	Short	BROCADE		L0	8	
7/3	7		Online	N	8 GAN	Test1006		NPIV Host	50:05:07:68:0c:14:00:05	50:05:07:68:0c:14:00:05	Short	BROCADE		L0	8	
7/4	7		Online	N	8 GAN	tape1111		NPIV Host	50:05:07:68:0c:14:00:04	50:05:07:68:0c:14:00:04	Short	BROCADE		L0	8	
7/5	7		Online	N	8 GAN	tape1112		NPIV Host	50:05:07:68:0c:14:00:03	50:05:07:68:0c:14:00:03	Short	BROCADE		L0	8	
7/6	7		Online	N	8 GAN	tape1113		NPIV Host	50:05:07:68:0c:14:00:02	50:05:07:68:0c:14:00:02	Short	BROCADE		L0	8	
8	8	020800	Online	F	8 GAN	tape1114	2145	SAN Volume Contro	50:05:07:68:0c:14:00:01	50:05:07:68:0c:14:00:01	Short	BROCADE	SCSI	L0	8	
9	9	020900	Online	F	8 GAN	tape1115		SAN Volume Contro	50:05:07:68:0c:14:00:00	50:05:07:68:0c:14:00:00	Short	BROCADE	SCSI	L0	8	

SAN Health is quick to install and run, as shown in Figure 8.

Figure 8) Steps required to run and use SAN Health.



VMware vSphere

There are many reasons why more than 50,000 customers have selected ONTAP as their storage solution for vSphere, as a unified storage system supporting both SAN and NAS protocols, robust data protection capabilities by using space-efficient NetApp Snapshot™ copies, and a wealth of tools to help you manage application data. Using a storage system separate from the hypervisor enables you to offload many functions and maximize your investment in vSphere host systems. This approach not only makes sure your host resources are focused on application workloads, but also avoids random performance impacts to applications from storage operations.

Using ONTAP together with vSphere is a great combination that lets you reduce host hardware and VMware software expenses, make sure data is protected at lower cost, and provide consistent high performance. And because virtualized workloads are mobile, you can explore different storage approaches by using Storage vMotion to move VMs across VMware Virtual Machine File System (VMFS), NFS, or vVol datastores, all on the same storage system.

As noted, vSphere 7.x supports NVMe-oF, enabling connectivity to external NVMe arrays by using FC protocol. As NVMe continues to grow and become the preferred storage, being able to connect to external NVMe arrays through the vSphere infrastructure is critical.

Data protection

Backing up your VMs and quickly recovering them are among the great strengths of ONTAP for vSphere; it is easy to manage this ability inside vCenter with the VSC and NetApp Snapshot technology. Use Snapshot copies to make quick copies of your VM or datastore without affecting performance. This approach minimizes storage space and network bandwidth by only storing changed information.

The power of ONTAP Snapshot technology is extended further with FabricPool. This data fabric technology enables cold Snapshot blocks to automatically move to a separate object storage tier to increase the number of Snapshot copies that can be maintained (to as many as 1,023) while reducing the

cost of storage. This object tier can be in the form of a private cloud (for example, NetApp StorageGRID®) or a public cloud (such as Amazon Web Services (AWS) or Azure). The solution moves cold data to the cloud as the blocks age, yet they are recalled automatically if the Snapshot copy is needed for recovering a VM or entire datastore.

Use case summary

The use case for virtualizing enterprise applications is well known. Not only does virtualization reduce costs through physical consolidation of servers and storage, which increases asset utilization, it also provides business flexibility. You can provision new server instances in moments to address urgent business needs. Furthermore, you can apply automation (the software-defined data center) to enable greater consistency, reducing problems that impact availability and data security.

Yet enterprises today face new imperatives that a modern SAN approach can address simply and quickly. Here are some of the ways NetApp, VMware, and Broadcom customers are adding value with ONTAP:

- **Cloud.** A broad array of hybrid cloud options enabled through the data fabric powered by NetApp supports the goals of most enterprise organizations to use a combination of public and private clouds to add more flexibility and reduce their infrastructure management overhead. Use cloud offerings from Azure, AWS, IBM, Google, and others with integrated ONTAP offerings for data protection, cloud computing, and business continuance while avoiding provider lock in.
- **Data protection.** Integrated data protection by using Snapshot copies and cloning speeds virtual storage provisioning and better protects critical data than relying on external protection systems.
- **Cost efficiency.** Integrated storage efficiency enables ONTAP to significantly reduce storage costs over legacy SAN systems. NetApp AFF systems can run all storage efficiency capabilities in production with no performance impact, something most other SAN arrays cannot do. The broad array of ONTAP storage efficiency features has resulted in customers seeing up to 5:1 savings for virtual server infrastructure and up to 30:1 for virtual desktop infrastructure. NetApp makes it simple to plan for these efficiency benefits with the most effective guarantee available.
- **Security.** ONTAP offers a range of features to meet an organization's security needs. You can enable NetApp Volume Encryption (NVE) quickly on any ONTAP volume and it does not require an external key server. You can also use it to enable digital shredding of data. Or use NetApp Storage Encryption (NSE) with self-encrypting disks for full disk encryption. Snapshot copies are used by many customers to protect against malware and ransomware and can be further strengthened with immutable Snapshot copies by using NetApp SnapLock® software. For more information, see TR-4572: The NetApp Solution for Ransomware.
- **Performance.** As described throughout this document, a modern SAN solution leveraging 32Gb FC SAN or NVMe/FC can meet the ever-faster performance requirements demanded by today's global, always-on enterprise.
- **Flexibility.** Needs change quickly in today's organization, and ONTAP is quick to adapt. Most of these capabilities are included with an ONTAP system at no additional charge or can be enabled with a license key. And while the focus of this NetApp Verified Architecture is SAN, the unified storage capabilities of ONTAP make it simple to add NAS protocols to support other applications and file sharing.

Technology requirements

This section covers the minimal technology requirements for this NetApp, VMware, and Broadcom NVMe/FC verified architecture.

Hardware requirements

Table 1) Supported hardware requirements for the joint solution.

Hardware	
Storage controllers	NetApp AFF A250/A300/A320/A400A700/A700s/A800 high-availability (HA) pair with 32Gb FC target ports and at least 24 SAS 960GB or larger SSDs
Switches	G720, X7 Directors, X6 Directors, G630, G620, G610 Switches, 8510 Directors, 6520, 6510 & 6505 Switches
FC HBAs	Emulex LPe35002-M2 32Gb FC
x86 servers	Fujitsu Primergy RX2540 M4

Software requirements

Table 2) Supported software requirements for the joint solution.

Software	Version
NetApp	ONTAP 9.7 or later
Brocade Fabric OS (FOS)	8.1.0a or later
Emulex Firmware	FW:12.6.234.3 DRVR:12.6.228.4 or later
VMware	vSphere 7.0 or later

Technology used during testing

This section covers the technology used in our lab for this NetApp, VMware, and Broadcom NVMe/FC verified architecture.

Table 3) Hardware used for the joint solution.

Hardware	Quantity
Storage NetApp AFF A800 HA pair with four 32Gb FC target ports and 24 SAS 1.9TB SSDs	1
Switches Brocade G720 64Gb FC switches 10Gb Ethernet switches	2 2
FC HBAs Emulex LPe36002-M2 32Gb FC	3
x86 servers Fujitsu Primergy RX2540 M4	3

Table 4) Software used for the joint solution.

Software	Version
NetApp ONTAP	9.9.1
Brocade Fabric OS (FOS)	9.0.1b
Emulex Firmware	12.2.299.23
MongoDB Community Edition	5.0
Ubuntu	20.04
VMware	vSphere 7.0 Update 2

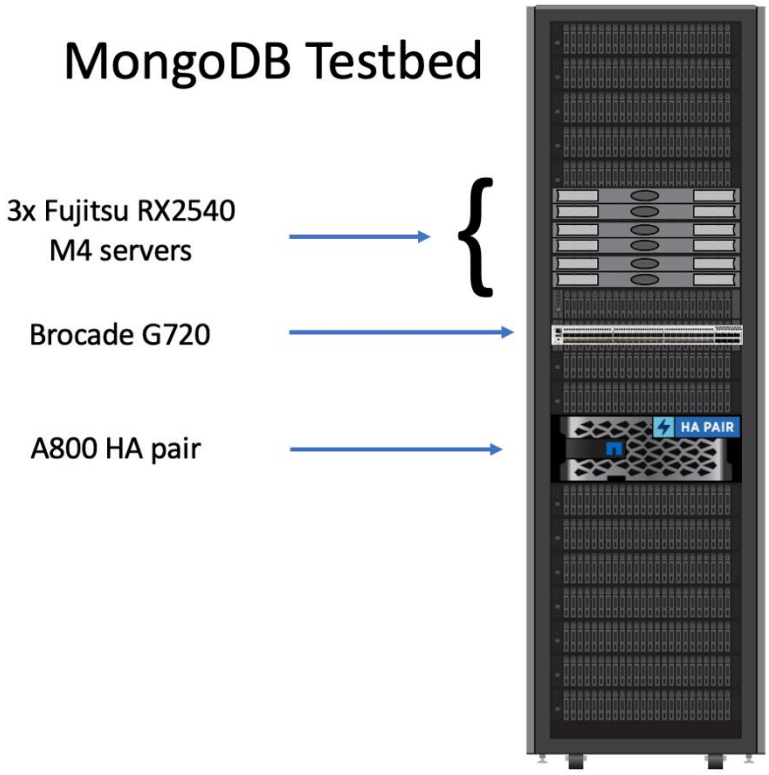
Testbed design

This section provides details for the tested configurations and an overview of the hardware that was used for the performance results.

As shown in Figure 9, we deployed three Fujitsu Primergy RX2540 M4 servers. We installed VMware vSphere 7.0 Update 2 on each server and installed three Ubuntu 20.04 VMs per server for a total of nine VMs. Each set of three VMs composed a MongoDB replica set.

The diagram in Figure 9 shows the rack layout of our solution used to generate the workload.

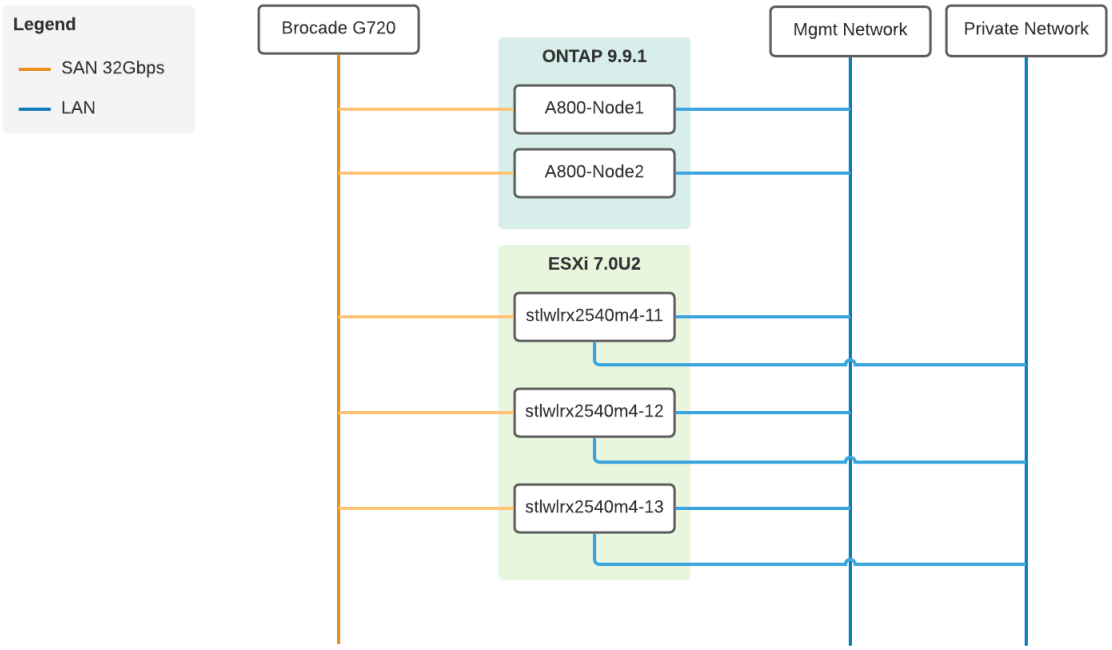
Figure 9) NetApp, VMware, and Broadcom validated architecture testbed layout.



We configured each ESXi host to have two 32Gb FC ports each connected to a two G720 Brocade switches. The AFF A800 had four 32Gb FC connections on each of two storage nodes, resulting in a total of four 32Gb connections to each G720 switch. We configured zoning on the Brocade director by using single initiator zoning and worldwide port names (WWPNs) to identify zone members.

Figure 10 shows a block diagram of the SAN testbed environment.

Figure 10) SAN testbed block.



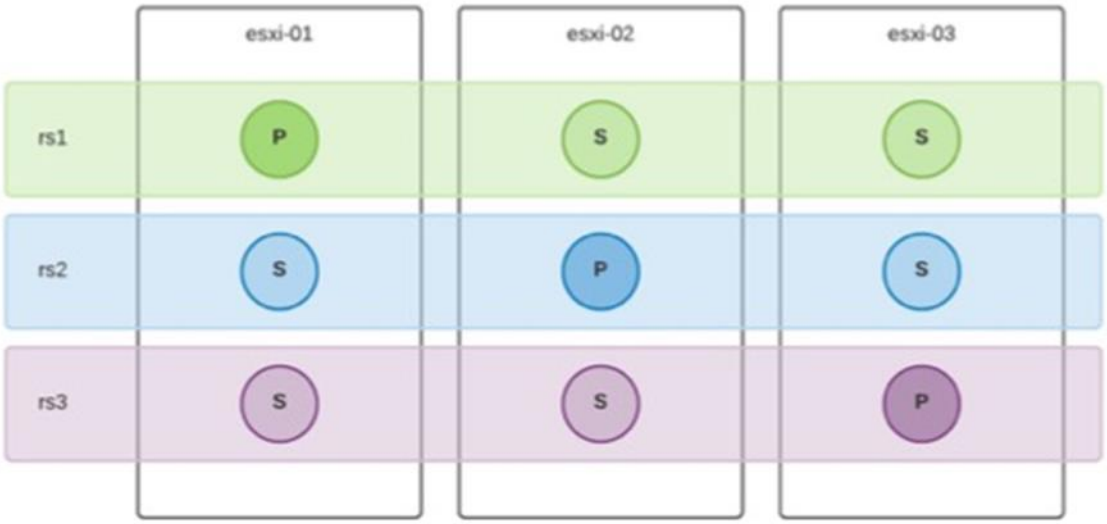
For Ethernet connectivity, each of the hosts had both 1Gb and 10Gb links for management, vMotion, and other provisioning traffic.

Defining the number of volumes, LUNs, and namespaces is a task that must consider several factors, such as application workload, performance, and data protection policies. For our FCP testing, all LUNs and LIFs were created in storage node-1 by using the aggregate named `n1_aggr1`. For our NVMe/FC testing, all namespaces and LIFs were created in storage node-2 by using the aggregate named `n2_aggr1`. Node-1 had 54 LUNs for FCP, and node-2 had 27 namespaces for NVMe/FC.

We allocated each MongoDB VM with 12 vCPUs and 24GB vRAM. For our FCP testing, we distributed each database across six Virtual Machine Disks (VMDKs) (one per datastore, six datastores per VM) using Linux LVM2. We had one Linux volume group containing three striped logical volumes: one for the database binary files and journal, one for the database oplog files, and one for the application database files. For our NVMe/FC testing, we distributed each database across six VMDKs (two per datastore, three datastores per VM) using Linux LVM2. We used the same volume group layout described in the FCP testing.

As shown in Figure 11, we used three MongoDB replica sets. Each replica set is a three-member cluster using the Primary-Secondary-Secondary (PSS) topology.

Figure 11) Three MongoDB replica sets, each one deployed as a PSS topology.



Workload design

MongoDB is a database vastly used in web applications, and there are several use cases, such as content management, personalization, product catalog, and real-time analytics, among others. From a storage perspective, the predominant workload for these use cases is the read workload.

For our testing, we used Yahoo! Cloud Service Benchmark (YCSB) to generate the workload on MongoDB. YCSB has a set of predefined workloads that can be used to compare the performance of NoSQL and cloud databases.

We used two predefined YCSB workloads: workload B, and workload C—with a couple changes regarding the number of operations and record count per workload to compare the performance of FCP versus NVMe/FC.

YCSB is not a tool for benchmarking the database storage engine but rather the whole database engine. This is important because we need to work with a large dataset to be able to force I/O down to the storage engine. A large dataset is needed to ensure that the working set size does not fit completely in RAM and allow database cache evictions and disk I/O activity.

We installed YCSB in the primary member of each replica set, for a total of three YCSB clients. For the database creation, data ingestion, and any write related operation (inserts, updates, or deletes), the MongoDB writeConcern was set to `w=1`.

We did not have the writeConcern set to `w="majority"` because we are interested in measuring the application latency without the overhead of waiting on the majority of the nodes in the replica set to acknowledge a write operation.

For each replica set, we created a 540GB database by using the following YCSB command:

```
./bin/ycsb load mongodb -s -P workloads/workloada -p
mongodb.url="mongodb://rs1m1,rs1m2,rs1m3/ycsb1?replicaSet=rs1&w=1&wtimeoutMS=30000" -p
fieldlength=380 -p mongodb.batchsize=100 -p recordcount=135000000 -p threadcount=8
```

In other words, we had a database called `ycsb1`, containing a `usertable` collection with 135,000,000 documents. Each document had 10 fields, each field 380 bytes long. The average document size was ~4KB.

Solution verification

This section describes the test methodology used to verify the architecture while we ran a suite of synthetic workloads.

Test methodology

After all three replica sets had their database in place, we executed a script on each primary to kick off the YCSB workload. The script had a loop to control the number of threads per lap per replica set. We started 8 YCSB threads and scaled up to 256 threads by a factor of two, and we characterized a full run, which consisted of six laps. We also added a 5-minute interval in between thread counts.

First, the script called workload C, and then workload B. In both cases, YCSB splits read between the secondary replica sets, and sent all the writes (at first) to the primary replica sets. In this case, because we wanted to create as much I/O activity upstream/downstream as possible, we decided to read from the secondaries, but it is important to highlight that if your application needs to scale on reads, you might be better served with a sharded cluster deployment.

It is also important to highlight that our testing was a worst-case scenario testing because we instructed YCSB to read from or write to any document in the range of one up to 135,000,000 documents, meaning that the working set size was the entire collection.

To run workload C, run the following command:

```
./bin/ycsb run mongodb -s -P ./workloads/workloadc -p
mongodb.url="mongodb://rs1m1,rs1m2,rs1m3/ycsb1?replicaSet=rs1&w=1&wtimeoutMS=30000&maxPoolSize=25
6&readPreference=secondary" -p operationcount= 200000000 -p recordcount=135000000 -p
maxexecutiontime=1800 -p threadcount=64
```

While the workload was running, we collected several application, database, and storage performance indicators for each lap.

YCSB generates a file which contains a summary of the operations executed and their latency stats. We ran mongostat from the database and collected its output, and collected storage throughput, latency, CPU utilization, disk utilization, and network traffic utilization.

We completed all the preceding steps for FCP first, then we removed the FCP VMDKs from the VMs. Next, we attached the NVME/FC VMDKs to the VMs and repeated all the steps.

In the end, we had a set of runs for workload C and workload B on FCP, and another set of runs for the same workloads on NVME/FC. In this way, the only difference between both sets is the protocol.

Test results

We realized throughput and latency improvements on both workloads by just switching the protocol. We also observed other benefits, such as a reduction in storage CPU utilization when running on NVMe/FC compared to FCP.

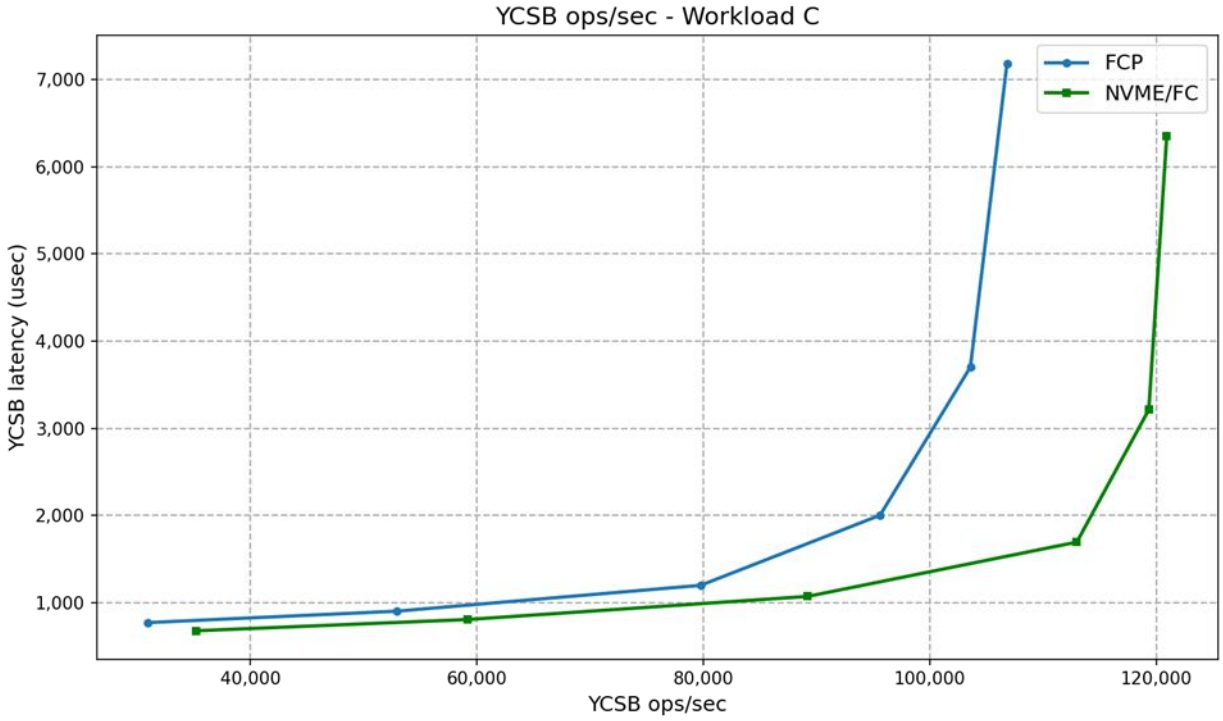
MongoDB has a query profiling feature which by default flags slow queries if they go over 100ms. Looking Figure 12, you can see that the latency is below 1ms at the lowest thread count, and no higher than 7ms at the highest thread count.

Figure 12 shows YCSB Workload C. The application’s performance on NVME/FC improved by 18% compared to FCP. Looking at the knee of the curve for FCP, YCSB workload C achieved ~94,000 operations per second at 2ms, while NVMe/FC achieved ~114,500 operations per second at 1.7ms.

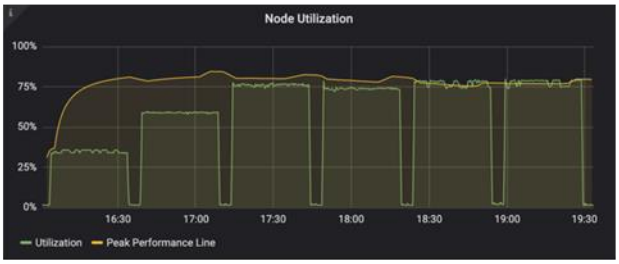
You might notice after the fourth data point in the chart that the latency starts climbing faster than the previous data points. At that point, we begin to experience high CPU utilization on the ESXi hosts. In the last data point, the average CPU utilization on our ESXi hosts was 97%.

In our test, we also observed a reduction in CPU utilization on the storage side. During our FCP testing, the average CPU utilization was ~75%, and during the NVME/FC testing, the average CPU utilization was 50%.

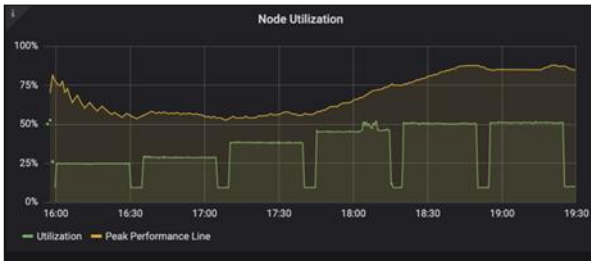
Figure 12) Workload C: YCSB operations per second versus YCSB read latency.



FCP: Storage CPU utilization

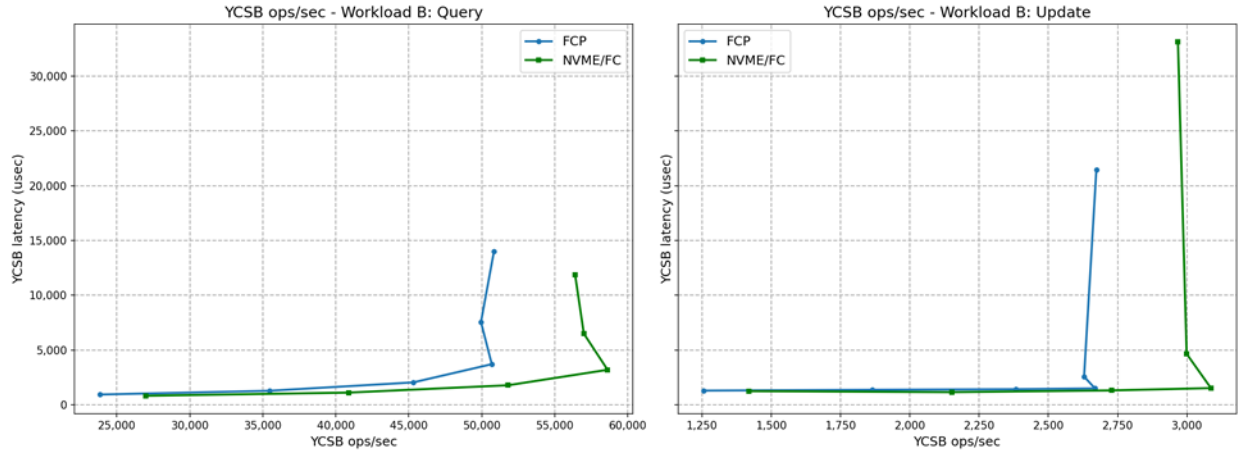


NVME/FC: Storage CPU utilization



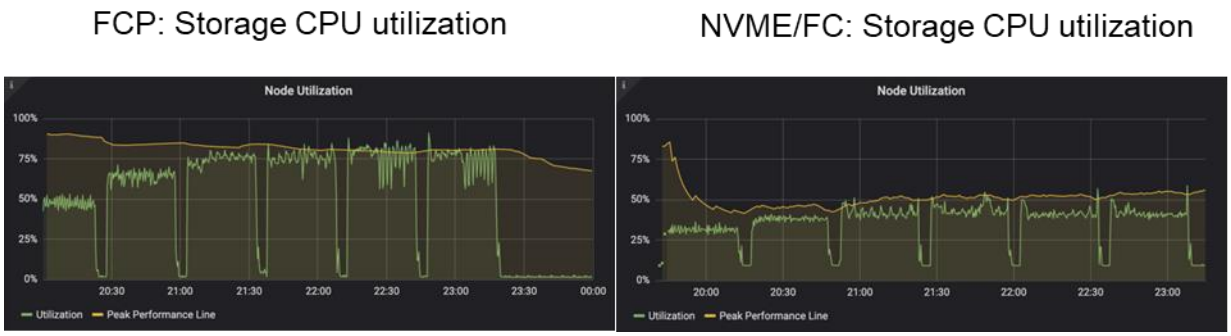
As shown in Figure 13, workload B query and update performance on NVMe/FC was 15% greater than FCP performance. Looking at the knee of the curve for FCP, YCSB workload B achieved ~53,000 operations per second at 3.7ms, while NVMe/FC achieved ~62,000 operations per second at 3.1ms.

Figure 13) Workload B: YCSB operations per second versus latency.



From a storage CPU utilization perspective, as shown previously in workload C, we realize the same benefits here, as shown in Figure 14. For FCP, the average CPU utilization is about 75%, while NVMe/FC is about 50%.

Figure 14) FCP versus NVMe/FCP storage CPU utilization.



Workload B follows the same distribution pattern for queries as Workload C, meaning that all queries were sent to secondary storage, but updates were always sent to the primary. In this workload, updates are only 5% of the total requests and latency seems to be almost flat until we hit a CPU bottleneck on the ESXi servers, where the latency climbs quickly.

In addition, note the following benefits:

- **NVMe/FC is easy to adopt**. All the performance gains that we observed were made possible by a simple software upgrade.
- **NVMe/FC protects your investment**. The benefits that we observed were with existing hardware that supports 32Gb FC.
- **NVMe/FC promotes data center consolidation**. With increased IOPS density, your system can complete more work in the same hardware footprint. Also, because NVMe/FC often reduces processor and memory loads on initiators, if you adopt NVMe/FC, your organization might be able to reduce the number of servers that you need for your workloads. This reduction translates to fewer servers and lower software licensing, footprint, and power and cooling costs.

Better performance with existing hardware

You can achieve the performance benefits described in this report by simply applying a software upgrade for the FC HBAs. By moving to NVMe/FC with the same storage hardware, you can attain dramatic increases in performance.

NVMe/FC benefits—FC HBAs

NVMe/FC brings native parallelism and efficiency to block storage that FCP (FC-SCSI) cannot. In separate testing over at least the past year, Broadcom (Emulex division) has observed performance improvements of up to two times with NVMe/FC over FC-SCSI.

NVMe/FC benefits—FC switches

Brocade Gen 7 Fibre Channel fabrics transport both NVMe and FCP (FC-SCSI) traffic concurrently with the same high bandwidth and low latency. Overall, the NVMe performance benefits are in the end nodes—initiators and targets. NVMe/FC provides the same proven security that the traditional FCP has provided for many years. FC provides full fabric services for NVMe/FC and FCP (FC-SCSI), such as discovery and zoning. Also, NVMe/FC is the first enterprise NVMe-oF transport that meets the same high bar as SCSI over FC with full-matrix testing as an enabler and as essential for enterprise-level support.

Future disruptive innovation

For the past few years, the IT industry has undergone a rapid chain of innovation that has resulted in substantial disruption to traditional IT delivery models and has rendered many legacy hardware vendors obsolete. Most architectures are unable to evolve with the changes resulting in successive waves of disruption, rearchitecture, forklift upgrades, and migration for customers that they can no longer afford from either an inefficiency or financial perspective.

At NetApp, we have pioneered the concept of nondisruptive operations (NDO) migrations and online transitions between generations of technology with heterogeneously scalable IT infrastructure. NetApp has focused on innovation in software and on the ability for you to add infrastructure as you grow, with connections between each generation of technology. The following is just a short list of recent disruptions. NetApp stands ready to take these innovations into our architectures of today and help you integrate them without forklift upgrades or disruptive migrations.

Key technology initiatives that are driving change include the following:

- HDDs replaced by flash
- Hardware appliances augmented or replaced by software-defined storage (SDS)
- NVMe-based media attached for flash
- NVMe-based host attachment
- Cloud-based IT infrastructure
- Hyper converged infrastructure
- AI, deep learning computing

As these initiatives come into the market, NetApp continues to support the evolution and revolution of IT with an agile software-defined approach. We support initiatives such as IoT, DevOps, hybrid cloud, and in-memory database server technologies, beyond what other vendors can comfortably discuss. We recently announced partnerships with three major hyperscalers for the NetApp cloud-connected flash array; our edge-to-core-to-cloud data pipeline; and the ability to mix SDS, hardware, and cloud instances of our data platform. These offerings give us a superior ability to future-proof your architecture.

As we have discussed in this report, with a simple software upgrade to the NVMe/FC protocol, you can easily future-proof your infrastructure and accelerate tier-1 mission-critical enterprise applications and workloads with an investment in NetApp.

Conclusion

In this report, we presented the NetApp, VMware, and Broadcom modern enterprise SAN verified architecture. This solution is the optimal infrastructure approach for you to leverage best-in-class, end-to-end, modern SAN and NVMe technologies to deliver business-critical IT services today while preparing for the future. As we have demonstrated, that future includes serving high-performance database, analytics, AI, and ML, and IoT requirements.

NetApp, VMware, and Broadcom have created an architecture framework that is both future-ready and usable today and is easy for you to implement within your current operational processes and procedures. One of our main objectives is to enable organizations like yours to quickly and nondisruptively streamline and modernize their traditional SAN infrastructure and the IT services that rely on it. To meet this objective, these modern platforms must do the following:

- Be high-performing to provide more real-time analysis and availability of critical data
- Adopt modern future-facing and disruptive technologies in a nondisruptive manner
- Provide agility, flexibility, and high scalability
- Fit within current operational frameworks
- Align with organizational objectives to consolidate and streamline infrastructure and operations

In this NetApp Verified Architecture, tests on a virtualized environment represent the benefits of a modern SAN architecture that is suited for multiple use cases and for critical SAN-based workloads. These benefits apply to most virtualized environments running VMware vSphere in a SAN environment.

With the flexibility and scalability of this NetApp Verified Architecture, your organization can start with a framework to modernize and right-size your infrastructure and can ultimately grow with and adapt to evolving business requirements. With these benefits, your system can serve existing workloads while streamlining infrastructure, reducing operational costs, and preparing for new workloads in the future.

Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:

- Leading the Future of Flash with NVMe
  www.netapp.com/us/info/nvme.aspx
- An Industry First: All-Flash NVMe over Fibre Channel
  https://blog.netapp.com/leading-the-industry-with-nvme-over-fibre-channel
- When You’re Implementing NVMe Over Fabrics, the Fabric Really Matters
  https://blog.netapp.com/nvme-over-fabric/
- TR-4597: VMware vSphere with ONTAP
  https://docs.netapp.com/us-en/netapp-solutions/virtualization/vsphere_ontap_ontap_for_vsphere.html
- TR-4684: Implementing and Configuring Modern SANs with NVMe/FC
  https://www.netapp.com/us/media/tr-4684.pdf
- TR-4080: Best Practices for Modern SAN ONTAP 9
  https://www.netapp.com/us/media/tr-4080.pdf
- NetApp SAN Solutions
  https://www.netapp.com/us/products/storage-systems/storage-area-network.aspx

- Brocade Fibre Channel Networking Switches
  https://www.broadcom.com/products/fibre-channel-networking/switches/
- Brocade Fibre Channel Networking Directors
  https://www.broadcom.com/products/fibre-channel-networking/directors
- Brocade/NetApp Partner Documents
  https://www.broadcom.com/company/oem-partners/fibre-channel-networking/netapp
- Fabric Notifications Technical Brief
  https://docs.broadcom.com/doc/FOS-90-Fabric-Notifications-OT
- NVMe over Fibre Channel for Dummies
  https://docs.broadcom.com/doc/nvme-over-fibre-channel-for-dummies-book
- NetApp SAN Health Check
  https://www.netapp.com/us/forms/campaign/amer-us-fy19q3-sss-san-san-health-check-inquiry-form.aspx?ref_source=smc&cid=27476
- SANnav Management Portal
  https://www.broadcom.com/products/fibre-channel-networking/software/sannav-management-portal
- Brocade Ansible automation GitHub repositories (switch)
  https://github.com/brocade/ansible
- Brocade Fabric Vision Technology
  https://www.broadcom.com/products/fibre-channel-networking/software/storage-fabrics-technology

Version history

Version	Date	Document version history
Version 1.0	February 2022	Initial release.

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

Copyright information

Trademark information

NVA-1163-DESIGN-0222

NetApp