Technical Report

FlexPod with All-Flash SAN Array

Availability & Integrity with NetApp ASA

Bobby Oommen, Kamini Singh, Kavyashree Mahadevaiah, Roney Daniel, NetApp
May 2024 | TR-4991

In partnership with

Abstract

This Technical Report dives into the design and implementation of a FlexPod data center solution featuring NetApp's All-Flash SAN Array (ASA). Optimized for peak performance and efficiency, it outlines best practices to ensure maximum reliability and focuses on the two principles of NetApp ASA, six-nines (99.9999%) availability & Data integrity. By leveraging the pre-validated and integrated nature of FlexPod, you gain a scalable, flexible storage solution designed to meet your future needs.

TABLE OF Contents

LIST OF TABLES

LIST OF FIGURES

FlexPod Introduction

FlexPod overview

FlexPod ® is a converged infrastructure data center architecture that includes the following components from Cisco® and NetApp ®:

- Cisco Unified Computing System (Cisco UCS)
- Cisco Nexus and MDS families of switches
- NetApp Fabric Attached Storage (FAS), All Flash FAS (AFF), and All-Flash SAN Array (ASA) systems

Shown in Figure 1 are some of the components utilized for creating FlexPod ® solutions. These components are connected and configured according to the best practices of both Cisco and NetApp to provide an ideal platform for running a variety of enterprise workloads with confidence. This solution primarily focuses on NetApp All SAN Array (ASA) which is a block only solution giving an availability of six 9s addressing the key pillars of data availability and data integrity.

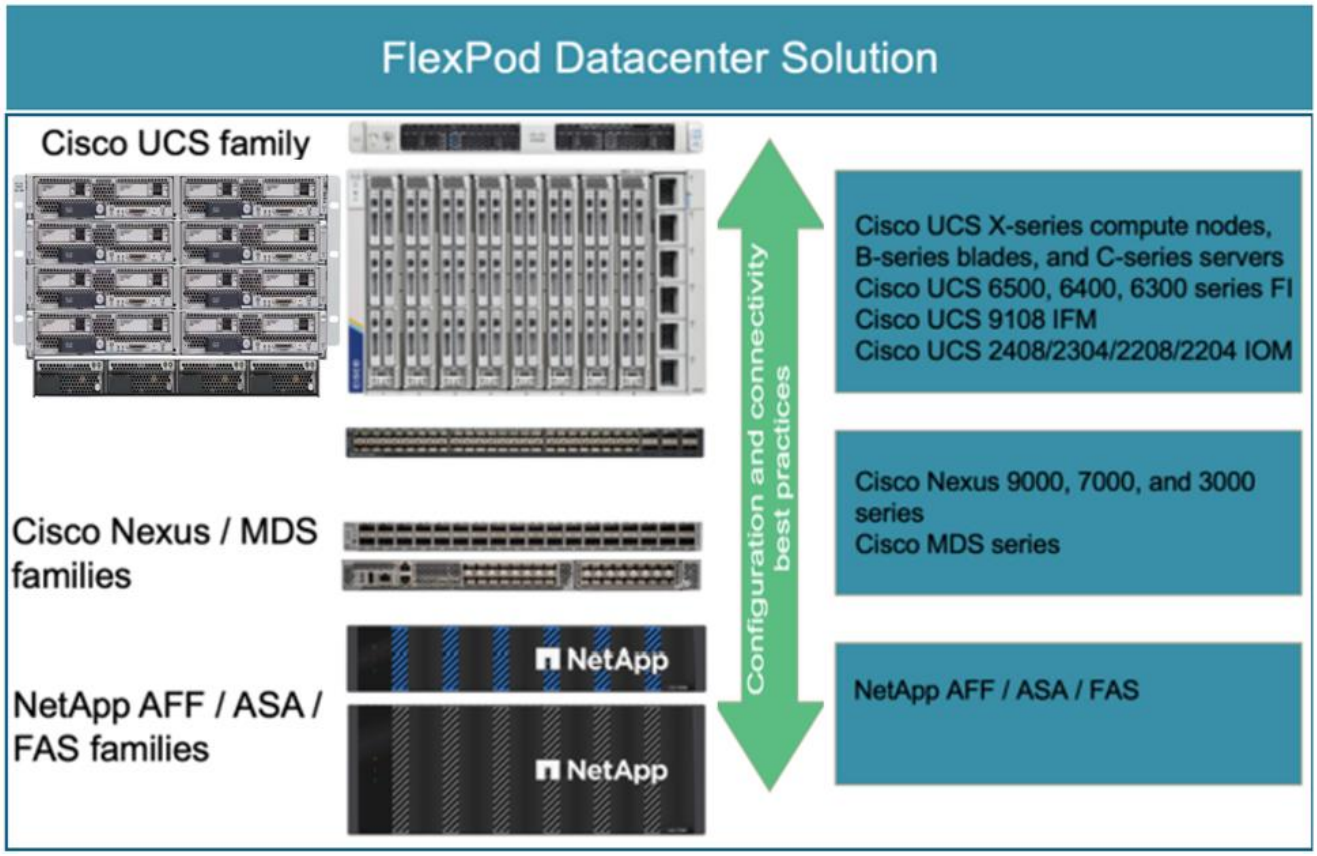


Figure 1 FlexPod Datacenter Solution

Each of the FlexPod component families shown (Cisco UCS, Cisco Nexus/MDS switches, and NetApp storage) provides platform and resource options to scale the infrastructure up or down as per application requirement, while supporting the features and functionalities that are required under the configuration and connectivity best practices of FlexPod ensuring resiliency through the infrastructure stack. FlexPod can also scale out for environments that require multiple consistent deployments by rolling out additional FlexPod stacks.

Audience

This document is intended for NetApp and partner solutions engineers (SEs) and professional services personnel. NetApp assumes that the reader has the following background knowledge:

- A solid understanding of Storage Area Network (SAN) concepts
- Technical familiarity with the configuration and administration of FlexPod systems

NetApp All-Flash SAN Array

Each NetApp All-Flash SAN Array (ASA) system is built on a NetApp AFF HA pair running ONTAP® software and provide an enterprise-class SAN solution for customers who want to consolidate and share storage resources for multiple workloads leading with two principles – Data Availability & Integrity.

ASA systems deliver the following:

- Industry-leading >99.9999% availability
- Massive clusters that scale both up and out
- The best enterprise performance in the industry (based on audited SPC-1 results)
- Industry-leading storage efficiency
- Among the most complete cloud-enabled connectivity available
- Cost-effective seamless data protection

NetApp ASA systems are built on the AFF platform to deliver continuous SAN availability with uninterrupted access to data during a planned or unplanned storage failover and deliver streamlined implementation, configuration, and management through a solution that's dedicated only to running SAN workloads. NetApp recommends ASA configurations when your requirements include the following:

- Mission-critical workloads such as databases that must have symmetric active-active paths from hosts to storage.
- Dedicated system to isolate SAN workloads.

NetApp ASA can be broadly categorized into two models: ASA A-Series and ASA C-Series. These models provide a full range of performance and capacity to meet the demands of the customers.

- ASA A-Series: optimized for performance-demanding workloads.
- ASA C-Series: optimized for capacity-intense workloads.

Refer to the following documents for detailed information about ASA A-Series and C-Series systems.

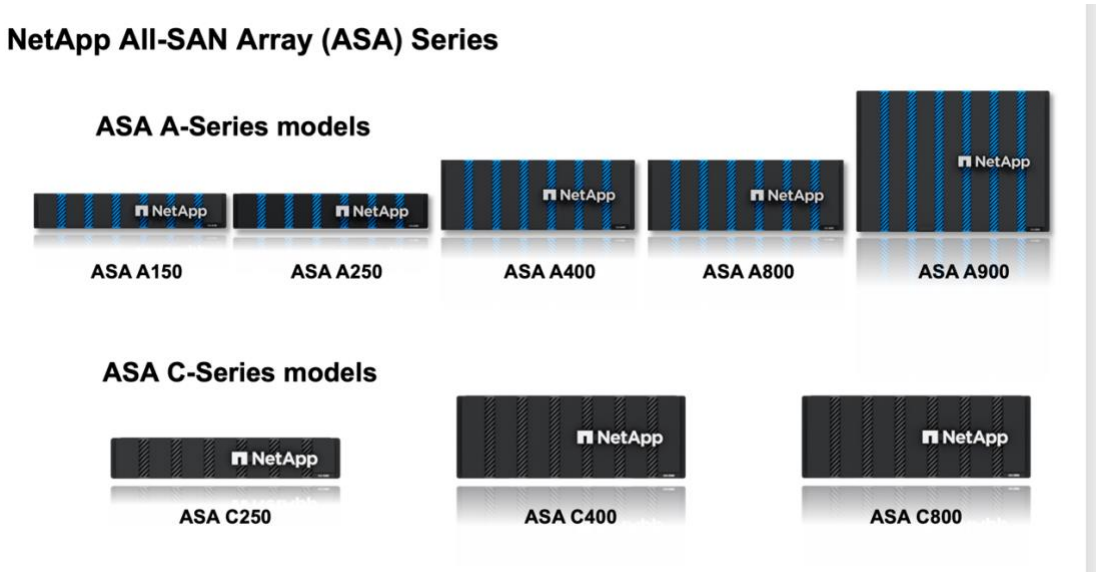


Figure 2 NetApp ASA A-Series and C-Series family

In this solution, ASA A400 and ASA C400 were used for validation. Refer the NetApp ASA Datasheet for more technical details and specifications.

Users can verify that the NetApp system is an ASA by executing the below commands on ONTAP CLI.

The output below is an example from ASA A400.

```
perf-asa-a400::> san config show
All SAN Array: true


perf-asa-a400::> system show
Node      Health Eligibility Uptime        Model       Owner    Location
--------- ------ ----------- ------------- ----------- -------- ---------------
perf-asa-a400-01
          true   true         2 days 18:43 ASA-A400
perf-asa-a400-02
          true   true         2 days 19:04 ASA-A400


perf-asa-a400::> system node run -node perf-asa-a400-01 -command sysconfig 0
        slot 0: System Board
                Model Name:         ASA-A400
```

NetApp ASA Architecture

There are two important requirements for any storage system: data should be always available and protected.

Data Availability and Integrity

During a planned or unplanned failure of a controller node in a HA pair, the IO still will be served through the block network interfaces from the surviving controller since all the block devices are presented to both controllers in the HA pairs at any given time. The interfaces on the failed controller supporting block protocols are brought online on the partner controller either by moving the IP address (iSCSI LIF failover)

or relocating the Host Bus Adapter (HBA) World Wide Name (WWN) (FCP and NVMe/FC Persistent Ports) so the host does no longer see any failed paths to storage.

Since, NetApp ASA uses active-active paths through both controllers, host OS will see optimized paths from both storage controllers, and applications will have continuous access to the underlying storage devices during storage failovers.

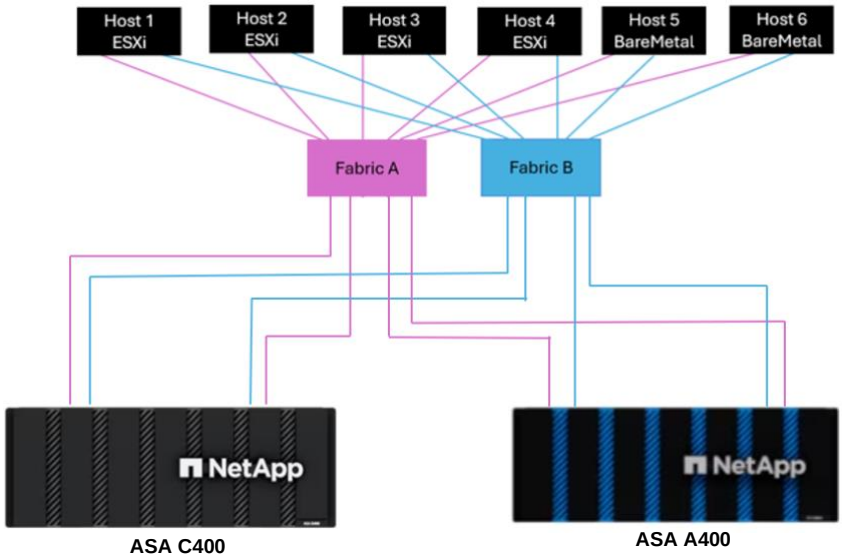


Figure 3 NetApp ASA A-Series & C-Series active-active multipathing

Table 1 Failover times at storage

Failover Type	IO Resumption Time
Planned takeover	2 - 3 seconds
Unplanned takeover	2 - 3 seconds

Apart from redundant components like power, cabling, RAID groups, disks, and controllers, each NetApp HA pair contains redundant links to support replication of NVRAM data which is a battery-backed journaling feature before any writes to the underlying disks. The NVRAM is replicated between storage controllers in an HA pair and will ensure there is zero data loss during a controller failure within a HA pair.

Data Protection

An equally important aspect of data availability and integrity is the ability to recover from inevitable user and application errors. Any enterprise that demands 99.9999% uptime from storage arrays should also plan a backup and recovery strategy that ensures increasingly larger datasets can be rapidly and reliably recovered. Customers using NetApp storage arrays can leverage the following key technologies:

- Data Protection with NetApp Snapshot™ copies
- Data Restoration with ONTAP SnapRestore® and FlexClone®
- Data Replication with NetApp SnapMirror® unified replication
- Business Continuity solution for MetroCluster® using SnapMirror® Active Sync
- Data Protection with Encryption drives and MetroCluster® & SnapMirror® Active Sync / Business Continuity Solution
- Data Preservation with NetApp SnapLock® Compliance Software
- Data Protection with Encryption Drives (NetApp Storage Encryption)

Zero Trust with FlexPod

The FlexPod Zero Trust Framework helps customers ensure the infrastructure stack is fully secured along with implementing the hardening guide which covers industry best practices for securing all components of the FlexPod solution, including VMware vSphere, Cisco UCS, Cisco Nexus/MDS, and NetApp ONTAP storage. The broader security features which are available with the base ONTAP architecture guarantees protection against ransomware attacks by providing ransomware protection, detection, and application recovery through immutable snapshots.

Workload guidelines for ASA A-Series and C-Series

NetApp recommends A-Series for latency sensitive databases or IO-intensive VMware workloads while C-Series for less latency-sensitive and capacity-centric workloads, but as part of the generic guidance following should be considered.

1. Enterprise Application workloads like EHR or any workloads which need consistent sub-millisecond reads times for their production use-cases should be deployed on ASA A-Series.
2. Any workloads which don't have the requirement of consistent sub-millisecond latency are considerations of ASA C-Series.

Solution Overview

Introduction

NetApp offers a simplified and comprehensive data management solution for all enterprise applications toward their digital transformation. This solution document primarily discusses the design and deployment best practices that the customers need to follow while setting up FlexPod with NetApp ASA system. This document also showcases the best practices for the design, implementation, and management of a SAN network to achieve maximum reliability for any enterprise applications or workloads like SQL, Oracle, EHR etc.

Solution Benefits

FlexPod systems with NetApp ASA offers the following benefits:

- **Data Access protocol support:**
  - Over Ethernet: iSCSI, NVMe/TCP
  - Over Fibre Channel: FCP, NVMe/FC
- **High Availability:**
  - Active-active controller architecture
  - Symmetric active-active FCP and iSCSI multipathing
  - Nondisruptive maintenance, upgrade, and scale-out clustering
  - Multi-site resiliency for continuous data access
- **Security and Compliance**
  - Multifactor admin access
  - Secure multitenant shared storage
  - In-flight and data-at-rest encryption
  - Regulatory-compliant data retention

- Multi-admin verification before executing sensitive commands

- **Storage Efficiency:**
  - Inline data compression, deduplication, and compaction
  - Space-efficient LUN, and volume cloning
- **Data Protection:**
  - Application-consistent NetApp Snapshot™ copies and restore
  - Integrated remote backup/disaster recovery
  - Synchronous zero-data-loss replication
  - Tamper-proof Snapshot
  - Business continuity across sites by synchronously replicating the data via NetApp MetroCluster or SnapMirror Active Sync
- **Cloud Integration**
  - Seamlessly tier, backup and replicate data between on-premises and private, public clouds

Solution Components

NetApp ASA A400

NetApp All-Flash SAN Array (ASA) delivers a simplified and dedicated SAN experience that provides continuous data availability for enterprise mission critical databases and other SAN workloads via block storage access protocols like FC, iSCSI, NVMe/FC, and NVMe/TCP protocols. It offers market-leading performance, with inline storage efficiency, encryption, and active data protection. The ASA features a symmetric active-active controller architecture to ensure continuous access to data during planned and unplanned outages, including upgrades and maintenance. With best-in-class data services integration with Oracle, SAP, and Microsoft SQL Server databases, plus VMware and other leading hypervisors, ASA delivers accelerated time to value for enterprise database applications. The ASA is an HA pair with both active controllers equally capable of communicating to a LUN. This symmetric active-active configuration provides uninterrupted access to your LUNs, with virtually instantaneous failover recovery. With Cloud Volumes ONTAP, customers can seamlessly extend the ASA to the hybrid cloud with leading cloud providers.

NetApp ASA A400 model comes with 4U form factor, 10 PCIe expansion slots, 24 FC target ports of 32Gb speed, and 32 FC target ports of 16Gb speed (with FC mezzanine card). The platform supports NVMe/TCP, NVMe/FC, FC, and iSCSI storage networking protocols that support a maximum effective capacity of 351PB. Refer to the following link for AFF ASA technical specifications.

Note: As the NetApp ASA system is a dedicated block storage system, file access storage protocols such as NFS and SMB are not supported on an ASA system.

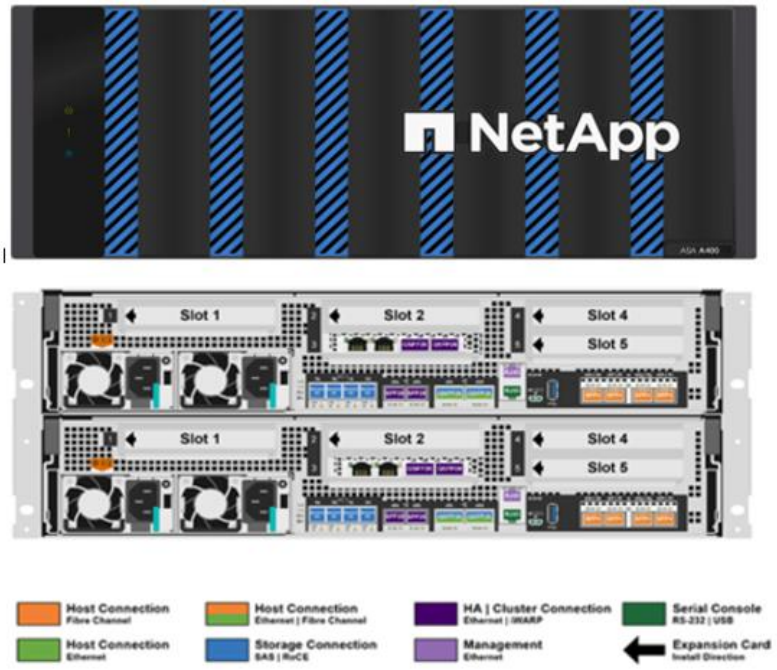


Figure 4 NetApp ASA A400 Front and Rear View

NetApp ASA C400

NetApp ASA C400 is a dedicated block storage capacity flash that is best suited for enterprise applications that require the best balance of performance and capacity. ASA C-Series systems deliver high capacity, slashing costs for intensive workloads. The model comes also in a 4U Controller Chassis form factor, supporting integration of up to 576 SSD drives, 10 PCIe expansion slots, 40 FC target ports of 32Gb speed, and 8 FC target of 16Gb speed with bundle option. The platform supports all the major block storage data access protocols offering maximum effective capacity of 35.5 PB.

Refer to the following link for detailed information on ASA C400.

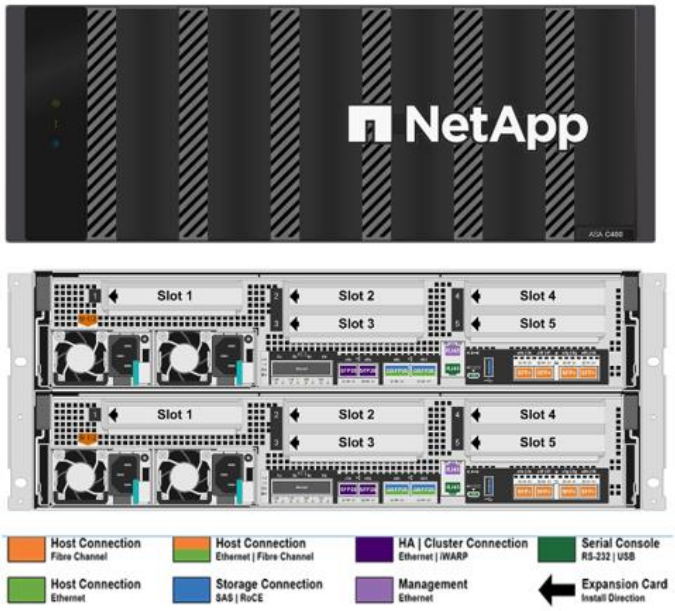


Figure 5 NetApp ASA C400 Front and Rear View

Cisco UCS X-series X210C M6 Blade Server

Figure 6 Cisco UCS X210c M6 Blade Server

Integrated into the Cisco UCS X- Series Modular System, the Cisco UCS X210c M6 Compute Node harnesses the power of the latest Third-Generation (3rd Gen) Intel Xeon Scalable processors (Ice Lake). It includes the following features:

- Up to two 3rd Gen Intel Xeon Scalable processors with up to 40 cores per processor and 1.5 MB of Level 3 cache per core.
- Up to thirty-two 256-GB DDR4 3200-MHz DIMMs for up to 8 TB of main memory. Configuring up to sixteen 512-GB Intel Optane persistent-memory DIMMs can yield up to 12 TB of memory.
- Up to six hot-pluggable Solid-State Disks (SSDs) or Non-Volatile Memory Express (NVMe) 2.5-inch drives with a choice of enterprise-class RAID or pass-through controllers with four lanes each of PCIe Gen 4 connectivity and up to two M.2 SATA drives for flexible boot and local storage capabilities.
- Modular LAN-on-motherboard (mLOM) Virtual Interface Card (VIC): The Cisco UCS VIC is connected to mLOM slot enabling up to 100-Gbps unified fabric connectivity to each of the chassis Intelligent Fabric Modules which was an IFM-100G for 200-Gbps connectivity per server.

Note: Security: The server supports an optional Trusted Platform Module (TPM). Additional features include a secure boot field-programmable gateway (FPGA) and Anti-Counterfeit Technology 2 (ACT2) provisions.

Cisco VIC 15231

The Cisco UCS VIC 15231 (Figure 7) are 2x100-Gbps Ethernet/FCoE-capable modular LAN on motherboard (mLOM) adapters designed exclusively for the Cisco UCS X210c Compute Node. The Cisco UCS VIC 15231 adapters enable a policy-based, stateless, agile server infrastructure that can present to the host PCIe standards-compliant interfaces that can be dynamically configured as either NICs or HBAs. In this solution, we use 2 VIC cards per UCS Server.

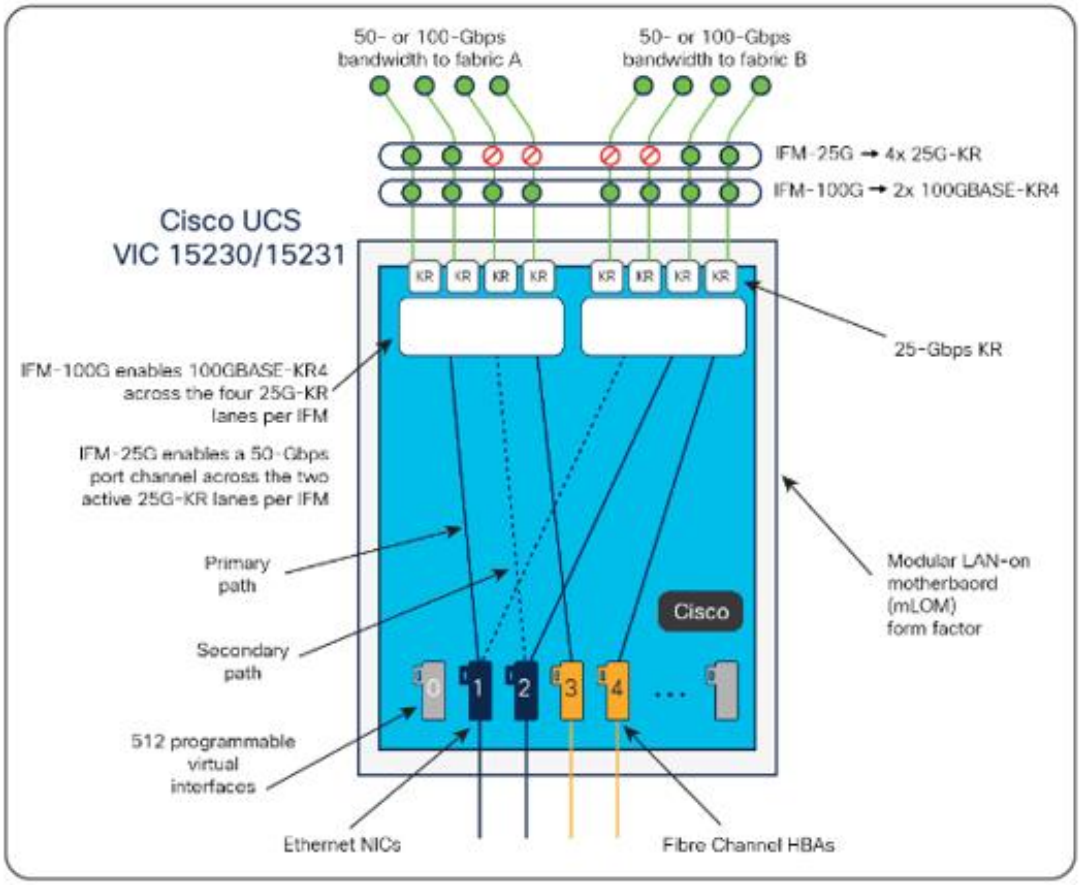


Figure 7 Cisco VIC 15231

Cisco UCS X9508 Chassis

Figure 8 Cisco UCS X9508 Chassis

The Cisco UCS X-Series Modular System begins with the Cisco UCS X9508 Chassis, engineered to be adaptable and future-ready. The X-Series is a standards-based open system designed to be deployed and automated quickly in a hybrid cloud environment.

With a midplane-free design, I/O connectivity for the X9508 Chassis is accomplished with front-loading vertically oriented computing nodes that intersect with horizontally oriented I/O connectivity modules in the rear of the chassis. Cisco UCS X-Series is powered by Cisco Intersight, making it simple to deploy and manage at scale.

The Cisco UCS X9508 Chassis provides these features and benefits:

- The Seven-Rack-Unit (7RU) chassis has eight front-facing flexible slots. These can house a combination of computing nodes and a pool of future I/O resources, which may include Graphics Processing Unit (GPU) accelerators, disk storage, and nonvolatile memory.
- Two Cisco UCS 9108 100G Intelligent Fabric Modules at the top of the chassis connect the chassis to upstream Cisco UCS 6536 Series Fabric Interconnects. Each Intelligent Fabric Module offers these features:
  - The module provides up to 100 Gbps of unified fabric connectivity per computing node.
  - The module provides eight 100-Gbps QSFP uplink ports.
- At the bottom of the chassis are slots ready to house future I/O modules that can flexibly connect the computing modules with I/O devices.

Cisco Nexus 9336C-FX2 Switches

Figure 9 Cisco Nexus 9336C-FX2 Switch

The Cisco Nexus 9336C-FX2 offers flexible port speeds supporting 1/10/25/40/100 Gbps in a compact 1 RU form factor with cloud-scale technology. It is designed to meet the changing needs of data centers, big data applications, and automated cloud environments.

- All 36 ports support 10/25/40/100 Gbps QSFP28 and wire-rate MACsec encryption.
- Supports 7.2 Tbps of bandwidth and over 2.8 bpps.
- Enhanced Cisco NX-OS Software designed for performance, resiliency, scalability, manageability, and programmability.
- Real-time buffer utilization per port and per queue, for monitoring traffic micro-bursts and application traffic patterns.

Cisco MDS 9132T Switches

Figure 10 Cisco MDS 9132T Switches

The next-generation Cisco MDS 9132T 32-Gbps 32-Port Fibre Channel Switch provides high-speed Fibre Channel connectivity from the server rack to the SAN core. It empowers small, midsize, and large enterprises that are rapidly deploying cloud-scale applications using extremely dense virtualized servers, providing the dual benefits of greater bandwidth and consolidation. Small-scale SAN architectures can be built from the foundation using this low-cost, low-power, non-blocking, line-rate, and low-latency, bi-directional airflow-capable, fixed standalone SAN switch connecting both storage and host ports. Medium-size to large-scale SAN architectures built with SAN core directors can expand 32-Gbps connectivity to the server rack using these switches either in switch mode or Network Port Virtualization (NPV) mode.

Cisco UCS 6536 Fabric Interconnect

Figure 11 Cisco UCS 6536 Fabric Interconnect

The Cisco UCS 6536 Fabric Interconnect (FI) is a core part of the Cisco Unified Computing System, providing both network connectivity and management capabilities for the system. The Cisco UCS 6536 Fabric Interconnect offers line-rate, low-latency, lossless 10/25/40/100 Gigabit Ethernet, Fibre Channel, NVMe over Fabric, and Fibre Channel over Ethernet (FCoE) functions. In addition, by supporting a unified

fabric, Cisco UCS 6536 Fabric Interconnect provides both LAN and SAN connectivity for all servers within its domain.

From a networking perspective, the Cisco UCS 6536 uses a cut-through architecture, supporting deterministic, low-latency, line-rate 10/25/40/100 Gigabit Ethernet ports, a switching capacity of 7.42 Tbps per FI and 14.84 Tbps per unified fabric domain, independent of packet size and enabled services.

NetApp ONTAP Tools for VMware vSphere

The ONTAP tools for VMware vSphere provide end-to-end life cycle management for virtual machines in VMware environments that use NetApp storage systems. It simplifies storage and data management for VMware environment by enabling administrators to directly manage storage within the vCenter Server. Refer the following link for OTV documentation.

NetApp Active IQ Unified Manager

NetApp Active IQ Unified Manager (AIQUM) is a comprehensive monitoring and proactive management tool for NetApp ONTAP systems to help manage the availability, capacity, protection, and performance risks of your storage systems and virtual infrastructure. The Unified Manager can be deployed on a Linux server, on a Windows server, or as a virtual appliance on a VMware host.

Refer the following link for AIQUM documentation.

Solution Topology

This section describes the physical topology of the FlexPod Data Center architecture, which was used for solution validation.

- Two 4-port 32G FC SFP+ (X1135A) NVMe/FC for data, 2 * 2-port 40/100G Smart IO QSFP28 (X1151A) for cluster interconnect and disk shelf access, and 2 * 2-port 100GbE RoCE QSFP28 (X1148A) for Ethernet data, 4-port 16G FC adapter comes bundled on NetApp ASA A400.
- Cisco VIC 15231 adapters with two 100G ports connect to the IFM on each X210c M6 server. The IFM-100G enables 100GBASE-KR4 across the four 25G-KR lanes per IFM.
- 8 * 100G links connected between Cisco UCS Fabric Interconnects (UCS-FI-6536) and Intelligent Fabric Module (IFM) to exploit the full bandwidth of both IFMs when the chassis is populated with eight compute nodes.
- 8 * 100G links connected between Cisco UCS Fabric Interconnects (UCS-FI-6536) and Cisco Nexus 9336C-FX2 switches.
- Four 100G links connected between NetApp ASA A400 and Cisco Nexus 9336C-FX2 to get full bandwidth on the storage.
- Four 32G FC links between Fabric Interconnects and Cisco MDS switches.

Topology Details

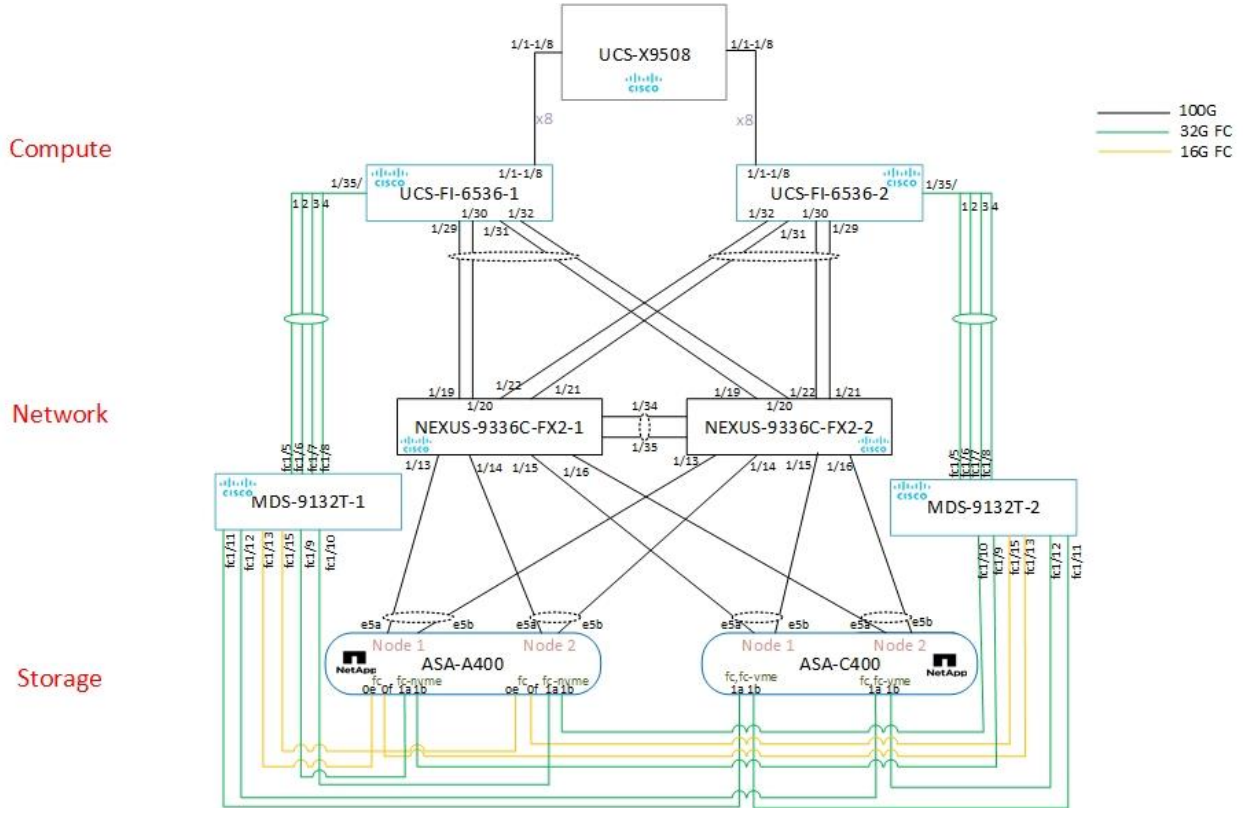


Figure 12 Solution Topology

- Four 16G FC links and four 32G FC links between Cisco MDS switches and NetApp ASA A400.
- Four 32G FC links between Cisco MDS switches and NetApp ASA C400.
- Port Channel is configured for the 4 FC links between Fabric Interconnect and MDS components.
- Port Channel configuration is set up to enable high-availability and load-sharing. One pair of Port-channel between Storage nodes and Nexus switches and one pair between Nexus and Fabric Interconnects and one port-channel between the Nexus peer switches with channel-mode active are configured.
- Virtual Port Channel (vPC) is enabled on all ethernet port-channel interfaces to provide hardware redundancy and loop free topology.
- Note that the storage failover tests for ASA A400 and C400 were performed independently to compare the functionality of both systems.

Cabling Information

Refer the following document for complete details on installation and setup of ASA A400 system.

The reference validation cabling details are shown in Figure 12 and Table 2 through Table 9. For this technical report, the console and management network of the FlexPod components are connected to the existing console and management network and are not documented in the cabling information below.

Table 2 Cabling information for Cisco Nexus 9336C-FX2 switch A

Local Device	Local Port	Remote Device	Remote Port
Cisco Nexus 9336C-FX2 A	e0M	Remote switch for in-band management network uplink	deployment specific
	1/13	NetApp ASA A400 01	e5a
	1/14	NetApp ASA A400 02	e5a
	1/15	NetApp ASA C400 01	e5a
	1/16	NetApp ASA C400 02	e5a
	1/19	Cisco UCS FI 6536 A	1/29
	1/20	Cisco UCS FI 6536 A	1/30
	1/21	Cisco UCS FI 6536 B	1/31
	1/22	Cisco UCS FI 6536 B	1/32
	1/35	Cisco Nexus 9336C-FX2 B	1/35
	1/36	Cisco Nexus 9336C-FX2 B	1/36

Table 3 Cabling information for Cisco Nexus 9336C-FX2 switch B

Local Device	Local Port	Remote Device	Remote Port
Cisco Nexus 9336C-FX2 B	e0M	Remote switch for in-band management network uplink	deployment specific
	1/13	NetApp ASA A400 01	e5b
	1/14	NetApp ASA A400 02	e5b
	1/15	NetApp ASA C400 01	e5b
	1/16	NetApp ASA C400 02	e5b
	1/19	Cisco FI 6536 A	1/31
	1/20	Cisco FI 6536 A	1/32
	1/21	Cisco FI 6536 B	1/29
	1/22	Cisco FI 6536 B	1/30
	1/35	Cisco Nexus 9336C-FX2 A	1/35
	1/36	Cisco Nexus 9336C-FX2 A	1/36

Table 4 Cabling information for NetApp ASA

Local Device	Local Port	Remote Device	Remote Port
NetApp ASA A400 01	e0M	management network	deployment specific
	e0a	NetApp ASA A400 02	e0a (25G HA interconnect port)
	e0b	NetApp ASA A400 02	e0b (25G HA interconnect port)

Local Device	Local Port	Remote Device	Remote Port
	e3a	NetApp ASA A400 02	cluster port
	e3b	NetApp ASA A400 02	cluster port
	e5a	Cisco Nexus 9336C-FX2 A	Eth 1/13
	e5b	Cisco Nexus 9336C-FX2 B	Eth 1/13
	0e	Cisco MDS 9132T A	fc1/13
	0f	Cisco MDS 9132T B	fc1/13
	1a	Cisco MDS 9132T A	fc1/9
	1b	Cisco MDS 9132T B	fc1/9
NetApp ASA C400 01	e0M	management network	
	e5a	Cisco Nexus 9336C-FX2 A	Eth 1/15
	e5b	Cisco Nexus 9336C-FX2 B	Eth 1/15
	1a	Cisco MDS 9132T A	fc 1/11
	1b	Cisco MDS 9132T B	fc 1/11

Table 5 Cabling information for NetApp ASA

Local Device	Local Port	Remote Device	Remote Port
NetApp ASA A400 02	e0M	management network	deployment specific
	e0a	NetApp ASA A400 01	e0a (25G HA interconnect port)
	e0b	NetApp ASA A400 01	e0b (25G HA interconnect port)
	e3a	NetApp ASA A400 01	cluster port
	e3b	NetApp ASA A400 01	cluster port
	e5a	Cisco Nexus 9336C-FX2 A	Eth 1/14
	e5b	Cisco Nexus 9336C-FX2 B	Eth 1/14
	0e	Cisco MDS 9132T A	fc1/15
	0f	Cisco MDS 9132T B	fc1/15
	1a	Cisco MDS 9132T A	fc1/10
	1b	Cisco MDS 9132T B	fc1/10
NetApp ASA C400 02	e0M	management network	
	e5a	Cisco Nexus 9336C-FX2 A	Eth 1/16
	e5b	Cisco Nexus 9336C-FX2 B	Eth 1/16
	1a	Cisco MDS 9132T A	fc 1/12
	1b	Cisco MDS 9132T B	fc 1/12

Table 6 Cabling information for Cisco UCS FI 6536 A

Local Device	Local Port	Remote Device	Remote Port
Cisco UCS FI 6536 A	1/29	Cisco Nexus 9336C-FX2 A	1/19
	1/30	Cisco Nexus 9336C-FX2 A	1/20
	1/31	Cisco Nexus 9336C-FX2 B	1/19
	1/32	Cisco Nexus 9336C-FX2 B	1/20
	1/35/1 – 1/35/4	Cisco MDS 9132T A	fc1/5-fc1/8
	1/1 – 1/8	Cisco UCS IFM	1/1-1/8

Table 7 Cabling information for Cisco UCS FI 6536 B

Local Device	Local Port	Remote Device	Remote Port
Cisco UCS FI 6536 B	1/29	Cisco Nexus 9336C-FX2 B	1/21
	1/30	Cisco Nexus 9336C-FX2 B	1/22
	1/31	Cisco Nexus 9336C-FX2 A	1/21
	1/32	Cisco Nexus 9336C-FX2 A	1/22
	1/35/1 – 1/35/4	Cisco MDS 9132T B	fc1/5-fc1/8
	1/1 – 1/8	Cisco UCS IFM	1/1-1/8

Table 8 Cabling information for Cisco MDS 9132T A

Local Device	Local Port	Remote Device	Remote Port
Cisco MDS 9132T A	fc1/5	Cisco UCS FI 6536 A	1/35/1
	fc1/6	Cisco UCS FI 6536 A	1/35/2
	fc1/7	Cisco UCS FI 6536 A	1/35/3
	fc1/8	Cisco UCS FI 6536 A	1/35/4
	fc1/13	NetApp ASA A400 01	0e
	fc1/9	NetApp ASA A400 01	1a
	fc1/15	NetApp ASA A400 02	0e
	fc1/10	NetApp ASA A400 02	1a
	fc1/11	NetApp ASA C400 01	1a
	fc1/12	NetApp ASA C400 02	1a

Table 9 Cabling information for Cisco MDS 9132T B

Local Device	Local Port	Remote Device	Remote Port
Cisco MDS 9132T B	fc1/5	Cisco UCS FI 6536 B	1/35/1
	fc1/6	Cisco UCS FI 6536 B	1/35/2
	fc1/7	Cisco UCS FI 6536 B	1/35/3

Local Device	Local Port	Remote Device	Remote Port
	fc1/8	Cisco UCS FI 6536 B	1/35/4
	fc1/13	NetApp ASA A400 01	0f
	fc1/9	NetApp ASA A400 01	1b
	fc1/15	NetApp ASA A400 02	0f
	fc1/10	NetApp ASA A400 02	1b
	fc1/11	NetApp ASA C400 01	1b
	fc1/12	NetApp ASA C400 02	1b

Network and Traffic segmentation

VLANs

A VLAN is a logical segment in a switched network by function or application. Each VLAN is considered a logical network consisting of the right ports and port-channel interfaces, and packets must be forwarded through a router for destinations that do not belong to the VLAN.

In FlexPod solutions, VLANs are utilized for network and traffic segmentations. Based on the specific FlexPod solution, there are various VLANs being utilized, including in-band management networks, iSCSI, NVMe/TCP protocol networks, VMware vMotion, and VM traffic etc.

Table 10 lists VLANs configured for the solution validation environment of this project.

Table 10 Configured VLANs and their usage

Name	VLAN ID	Usage
Native-VLAN	2	VLAN 2 used as native VLAN instead of default VLAN (1)
IB-MGMT-VLAN	2214	In-band management VLAN for ESXi hosts, VM management, etc.
iSCSI-A-VLAN	411	iSCSI-A fabric VLAN for iSCSI traffic
iSCSI-B-VLAN	412	iSCSI-B fabric VLAN for iSCSI traffic
vMotion-VLAN	405	VMware vMotion traffic VLAN for vCenter vMotion
VM-Traffic-VLAN	2217	VMware VM traffic VLAN for ESXi VMs
NVMe-TCP-A-VLAN	501	NVMe-A fabric VLAN for NVMe/TCP traffic
NVMe-TCP-B-VLAN	502	NVMe-B fabric VLAN for NVMe/TCP traffic

Virtual Networking

In this solution, NetApp ASA controller failover scenarios were tested for both bare metal and VMware virtualized deployments.

VMware

Each host in the cluster is deployed with identical virtual networking. The design separates the different traffic types using vSphere Standard Switch (vSwitch) and Distributed Switches (vDS). Dedicated

vSwitches for in-band management, iSCSI-A/B and NVMe-TCP-A/B networks and a vDS for Virtual Machine and vMotion network.

The virtual switches (vSwitch, vDS) are deployed with two uplinks per virtual switch for availability and load-sharing; the uplinks at the ESXi hypervisor level are referred to as vmnics and virtual NICs (vNICs) on Cisco UCS Software. The vNICs are created on the Cisco UCS VIC adapter in each server using Cisco UCS service profiles in Intersight.

The standard design option for the customers is to create a second vDS switch in the vCenter and migrate the iSCSI-A, iSCSI-B, NVMe-TCP-A and NVMe-TCP-B port-groups onto this vDS as depicted in Figure 13.

Note: In this case, Native VLAN must be set to iSCSI VLAN and both iSCSI and NVMe/TCP vNICs are pinned to either Fabric-A or Fabric-B.

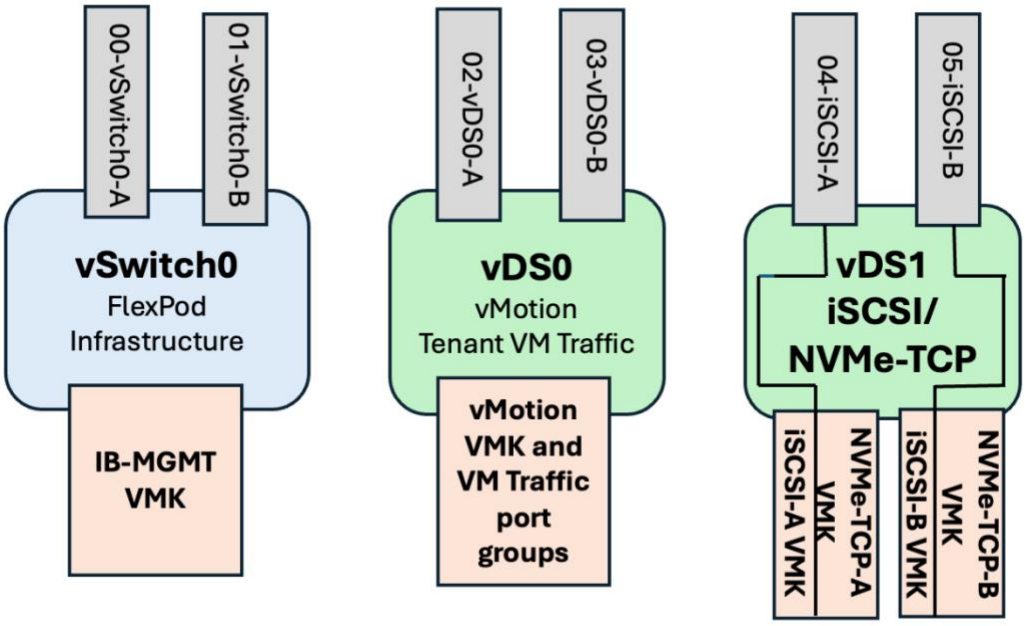


Figure 13 VMware Virtual Networking design with iSCSI Boot

Table 11 through Table 12 lists the details of vNIC Ethernet LIFs and FC LIFs creation on UCS Server.

Table 11 vNIC placement on Cisco UCS Server (Virtualization)

CDN Name	Switch ID	PCI Order
00-vSwitch-A	A	0
01-vSwitch-B	B	1
02-vDS0-A	A	2
03-vDS0-B	B	3
04-iSCSI-A	A	4
05-iSCSI-B	B	5

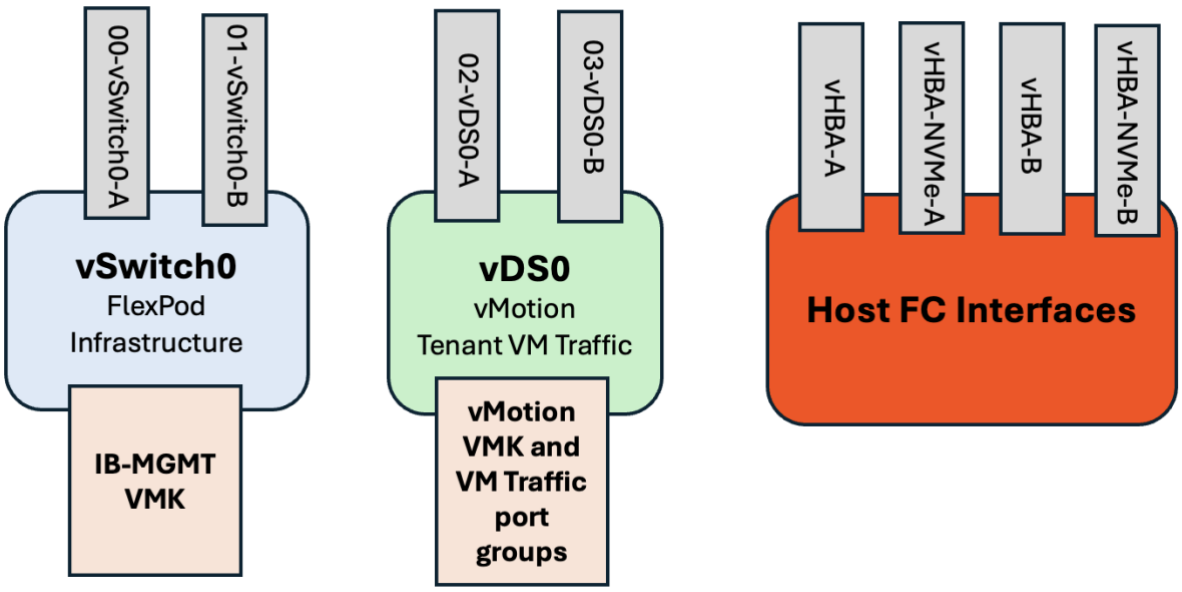


Figure 14 VMware Virtual Networking design with FC Boot

Table 12 vNICs and FC LIFs for Cisco UCS Server (Virtualization)

CDN Name	Switch ID	PCI Order
00-vSwitch-A	A	2
01-vSwitch-B	B	3
02-vDS0-A	A	4
03-vDS0-B	B	5
vHBA-A	A	0
vHBA-B	B	1
vHBA-NVME-A	A	6
vHBA-NVME-B	B	7

vSwitch0 is defined during VMware ESXi host configuration, and it contains the IB-Mgmt and the ESXi host VMkernel (VMK) ports for management. An infrastructure management virtual machine port group is also placed on vSwitch0 for any critical infrastructure management virtual machines that are needed.

A vSphere distributed switch is configured on the vCenter Server system, and the configuration is propagated to all hosts that are associated with the switch. The vDS contains port groups for tenant VM traffic and vMotion VMKs. Placing the vMotion VMK on vDS, provides ability to modify QoS to the vMotion port group, if necessary, in the future.

Two iSCSI vSwitches are used in this validation. Two vNICs are configured in Cisco UCS for iSCSI boot. These vNICs use iSCSI VLAN of the appropriate fabric as the native VLAN and are attached to the appropriate iSCSI vSwitch. Network segmentation is achieved through standard vSwitches and vDS for virtualized traffic separation, as well as VLAN configurations to isolate compute, network, and storage layers.

Figure 15 below depicts the logical connectivity and data path for Ethernet traffic in a FlexPod deployment.

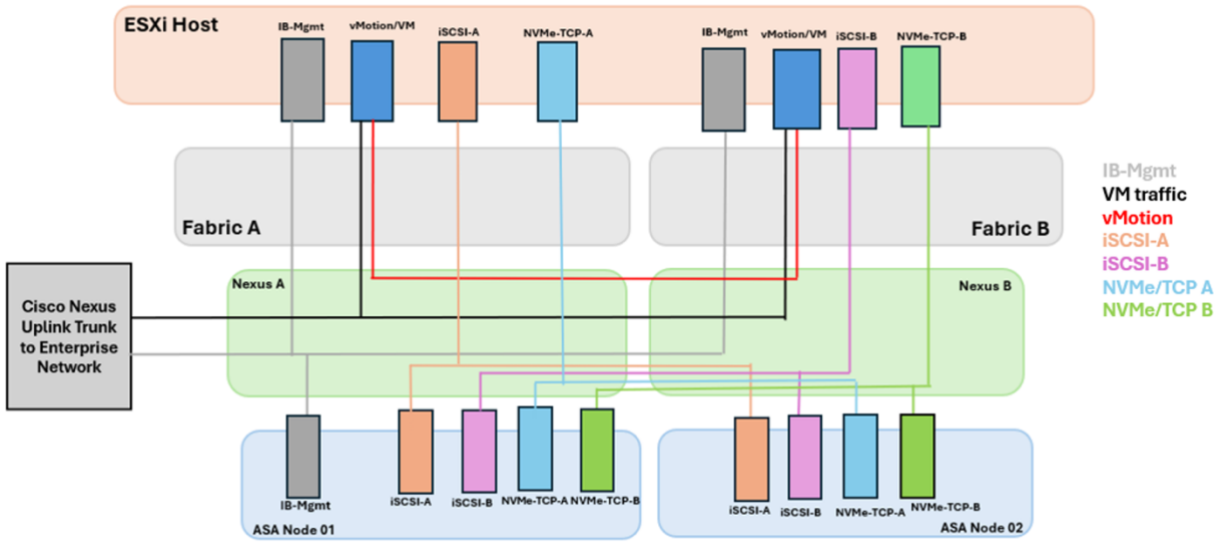


Figure 15 Data Path for Ethernet Traffic

Data Path for Fibre Channel traffic

For FCP and NVMe/FC protocols, the traffic between the compute and storage nodes flows via Cisco MDS switches over FC cables. Traffic isolation in MDS is achieved via VSAN and Zoning configurations.

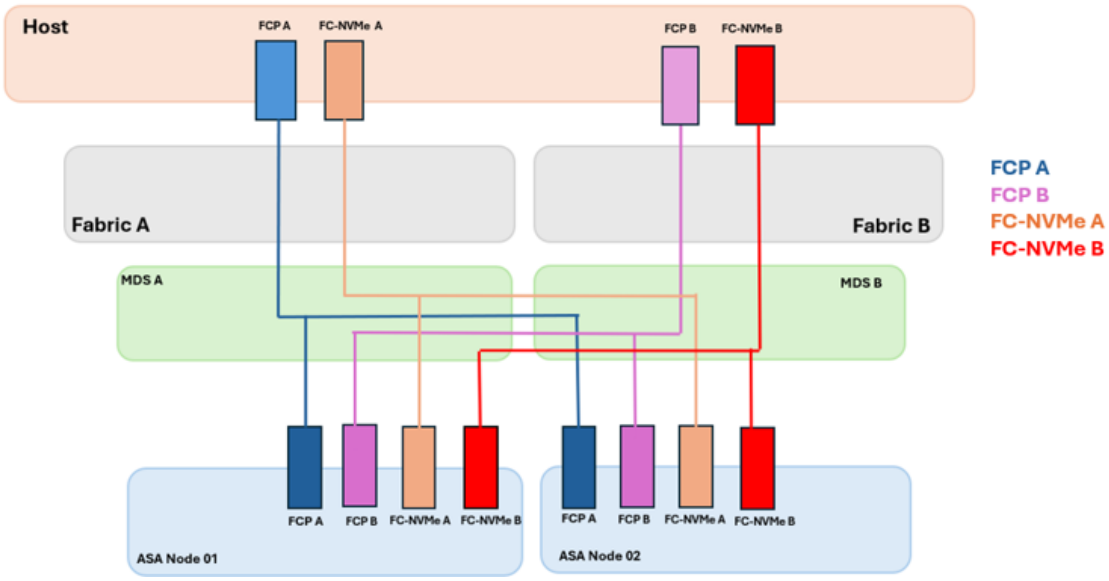


Figure 16 Data Path for Fibre Channel Traffic

Figure 16 shows the end-to-end data path for traffic over Fibre Channel media from compute to storage.

Bare Metal

We have used RHEL OS for Bare Metal installation. Table 13 highlights Ethernet and FC logical interfaces info for Bare Metal deployment.

Table 13 Ethernet and FC LIFs placement for Bare Metal

vNIC/vHBA Name	Switch ID	PCI Order
vHBA-A	A	0
vHBA-B	B	1
vHBA-NVMe-A	A	2
vHBA-NVMe-B	B	3
Eth0 (Failover Enabled)	A	4
04-iSCSI-A	A	5
05-iSCSI-B	B	6
06-NVMe-TCP-A	A	7
07-NVMe-TCP-B	B	8

Zoning on the Fabric switch

VSANs are created prior to creating the zoning configuration to provide isolation.

Zoning for NVMe setup

The zoning was implemented to provide isolation between the virtualized and bare metal environment.

- Two vHBAs were used from each VIC adapter port for all the workloads.
- A total of four NVMe/FC logical interfaces [LIFs] were created, two on each controller node for each corresponding fabrics A and B.
- Two separate zones were created on each MDS, one for the Bare Metal servers and another for ESXi hosts.

Table 14 Zoning for NVMe setup

Fabric Name	Zone Name	Host vHBA (initiator)	Storage LIFs (Target)
Fabric A	zone1 (VMware virtualized)	Host-Infra-FC-NVME-01-A Host-Infra-FC-NVME-02-A	ASA-Infra-SVM-FC-NVMe-lif-01a ASA-Infra-SVM-FC-NVMe-lif-02a
	zone2 (bare metal)	BM-Host-1-5-FC-NVMe-A BM-Host-1-6-FC-NVMe-A	ASA-Infra-SVM-FC-NVMe-lif-01a ASA-Infra-SVM-FC-NVMe-lif-02a
Fabric B	zone1 (VMware virtualized)	Host-Infra-FC-NVME-01-B Host-Infra-FC-NVME-02-B	ASA-Infra-SVM-FC-NVMe-lif-01b ASA-Infra-SVM-FC-NVMe-lif-02b
	zone2 (bare metal)	BM-Host-1-5-FC-NVMe-B BM-Host-1-6-FC-NVMe-B	ASA-Infra-SVM-FC-NVMe-lif-01b ASA-Infra-SVM-FC-NVMe-lif-02b

Zoning for FCP setup

The zoning was implemented to provide isolation between the virtualized and bare metal environment.

- Two vHBAs were used from each VIC adapter port for all the workloads.
- A total of four FC logical interfaces [LIFs] were created, two on each controller node for each corresponding fabrics A and B.
- Two separate zones were created on each MDS, one for the Bare Metal servers and another for ESXi hosts.

Table 15 Zoning for FCP setup

Fabric Name	Zone Name	Host vHBA (initiator)	Storage LIFs (Target)
Fabric A	zone1 (VMware virtualized)	Host-Infra-FC-01-A Host-Infra-FC-02-A	ASA-Infra-SVM-fcp-lif-01a ASA-Infra-SVM-fcp-lif-02a
	zone2 (bare metal)	BM-Host-1-5-FC-A BM-Host-1-6-FC-A	ASA-Infra-SVM-fcp-lif-01a ASA-Infra-SVM-fcp-lif-02a
Fabric B	zone1 (VMware virtualized)	Host-Infra-FC-01-B Host-Infra-FC-02-B	ASA-Infra-SVM-fcp-lif-01b ASA-Infra-SVM-fcp-lif-02b
	zone2 (bare metal)	BM-Host-1-5-FC-B BM-Host-1-6-FC-B	ASA-Infra-SVM-fcp-lif-01b ASA-Infra-SVM-fcp-lif-02b

Hardware and Software Revisions

This solution can be extended to any FlexPod environment running supported software, firmware, and hardware versions as defined in the NetApp Interoperability Matrix Tool, UCS Hardware and Software Compatibility, and VMware Compatibility Guide. The following table shows the FlexPod hardware and software revisions used in this solution.

Table 16 Hardware and Software versions

Component	Product	Version
Compute	Cisco UCSX-210C-M6	5.2(0.230092)
	Cisco UCS Fabric Interconnects 6536	4.3(2.240002)
	Cisco VIC 15231	5.3(2.40)
	CPU	Intel(R) Xeon(R) Gold 6338N CPU @ 2.20GHz
Network	Cisco Nexus 9336C-FX2 NX-OS	10.2(6)
	Cisco MDS 9132T	9.3(2a) 9.4.2 *
Storage	NetApp ASA - A400 and C400	ONTAP 9.14.1P2
Software	NetApp ONTAP Tools for VMware vSphere	9.13P1 [9.13P1-9828.OVA]
	NetApp Active IQ Unified Manager	9.14RC1
	VMware vSphere	8.0
	VMware vCenter Appliance	8.0
	VMware ESXi nfnic Fibre Channel driver	5.0.0.37
	VMware ESXi nenic ethernet driver	1.0.45.0
	Cisco Intersight Assist Virtual Appliance	1.0.11-4206
	RHEL	8.9
	Microsoft Windows Server	2022

Component	Product	Version
	Microsoft SQL Server	2022
	SQL Server Management Studio	19.3
	HammerDB workload for SQL	4.10
	GenIO Tool for EHR	1.17.3
	Oracle Database	21c

***** If you are using the onboard 16G FC ports on storage controllers and the FC port on MDS is configured for **switchport mode auto**, the interfaces on MDS may get stuck in offline state when the controller is rebooted. A shut/no shut on the MDS port is required to bring up the link. The workaround is to configure the switch port for **switchport mode F**. For more information, refer to the following bug [CSCwi36075 : Bug Search Tool (cisco.com)](). This defect is fixed in NX-OS 9.4.2 release.

NetApp ASA Integration with Cisco Intersight

To integrate NetApp ASA A400 and C400 with Cisco Intersight, you need to deploy and configure:

- Cisco Intersight Assist virtual appliance
- NetApp Active IQ Unified Manager virtual appliance

Using the Cisco Intersight Assist, NetApp Active IQ Unified Manager (AIQUM) is claimed as a target in Cisco Intersight. When NetApp AIQUM is claimed, the NetApp storage clusters configured in AIQUM will appear in Intersight and can be monitored and orchestrated.

Obtain Storage-level Information

After successfully claiming the NetApp Active IQ Unified Manager as a target, you can view storage-level information in Cisco Intersight if you have already added NetApp ASA A400 to the NetApp Active IQ Unified Manager.

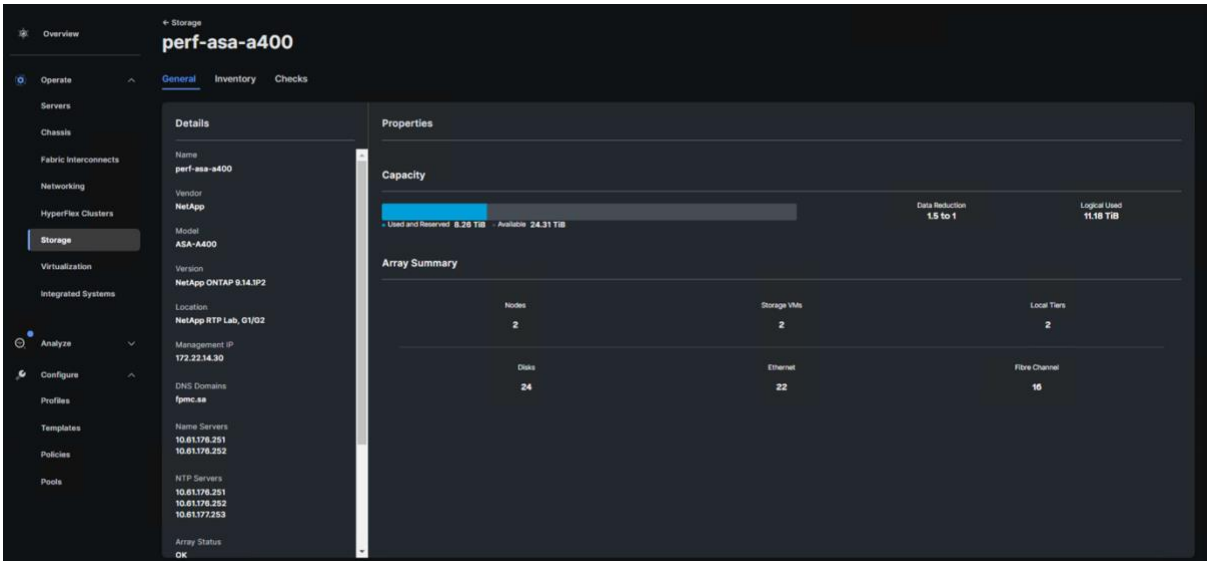


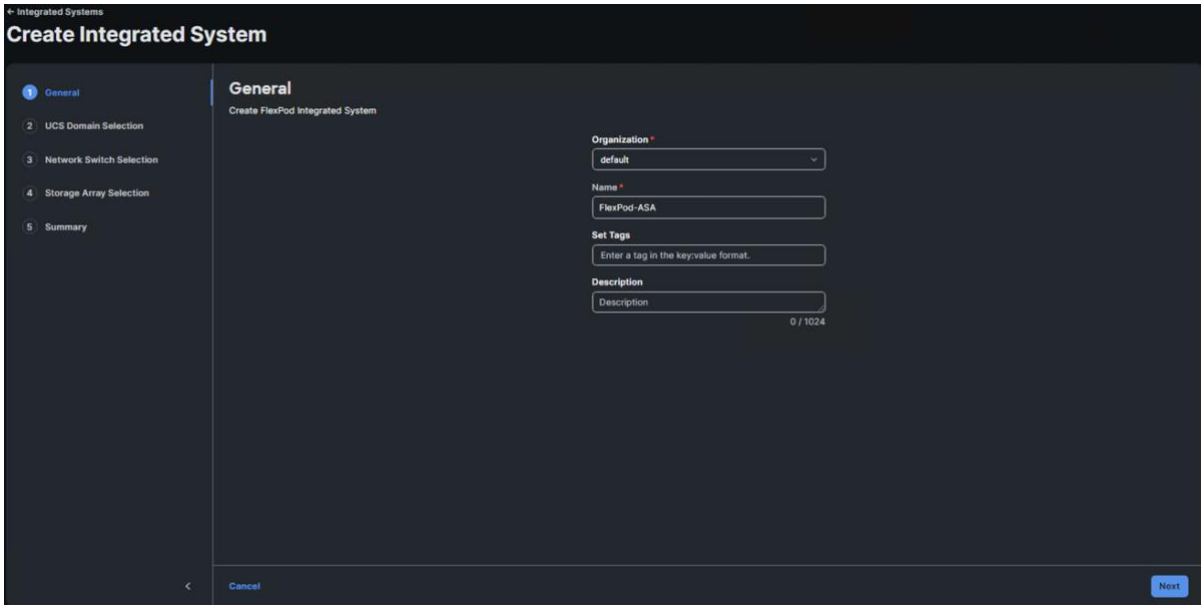
Figure 17 NetApp ASA A400 Information in Cisco Intersight

Cisco Intersight Integrated Systems with NetApp ASA

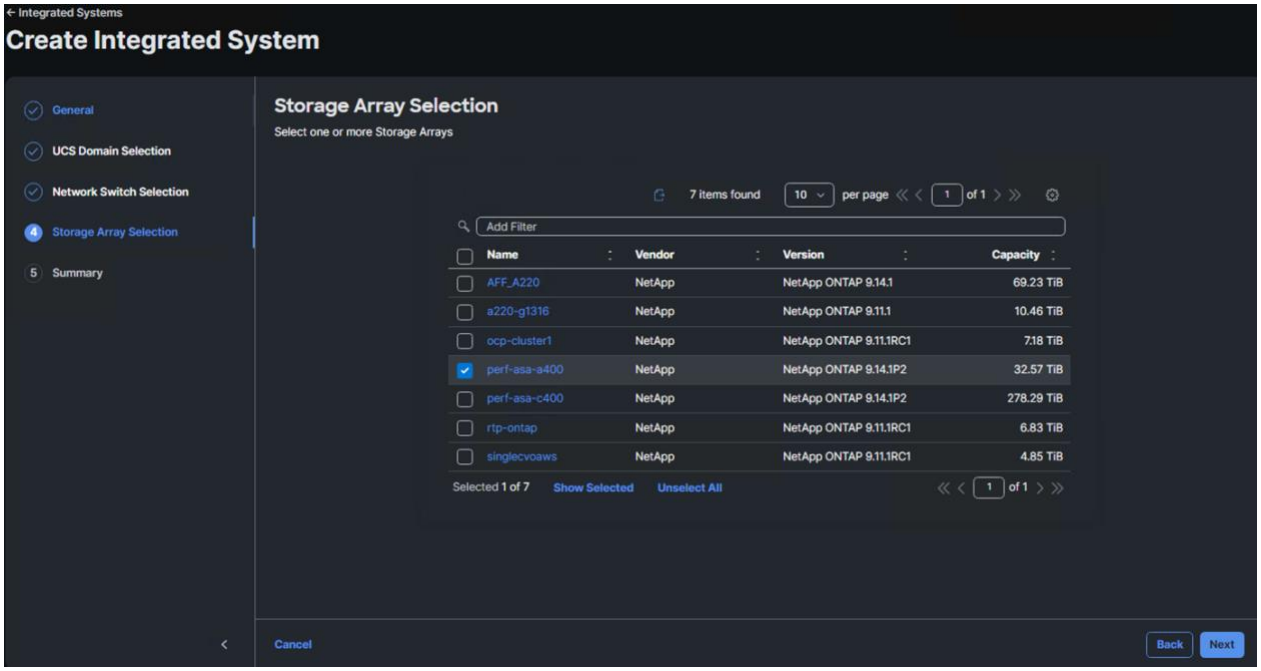
Cisco Intersight Integrated Systems logically group compute, networking, storage, and hypervisor entities of FlexPod to have a single converged infrastructure view of the FlexPod stack.

Follow the below steps to create a FlexPod Integrated System with NetApp ASA:

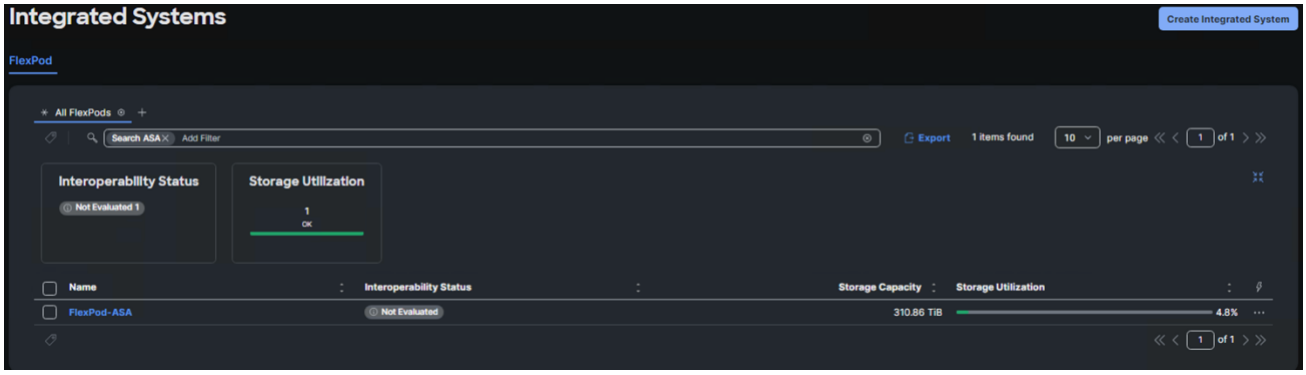
- Log into Cisco Intersight and connect to the account for your FlexPod.
- From Cisco Intersight, click Infrastructure Service > Operate > Integrated Systems.
- Click Create Integrated System. In the center pane, select FlexPod and click Start.
- Select the correct Organization, provide a suitable name, and optionally any Tags or a Description and click Next.



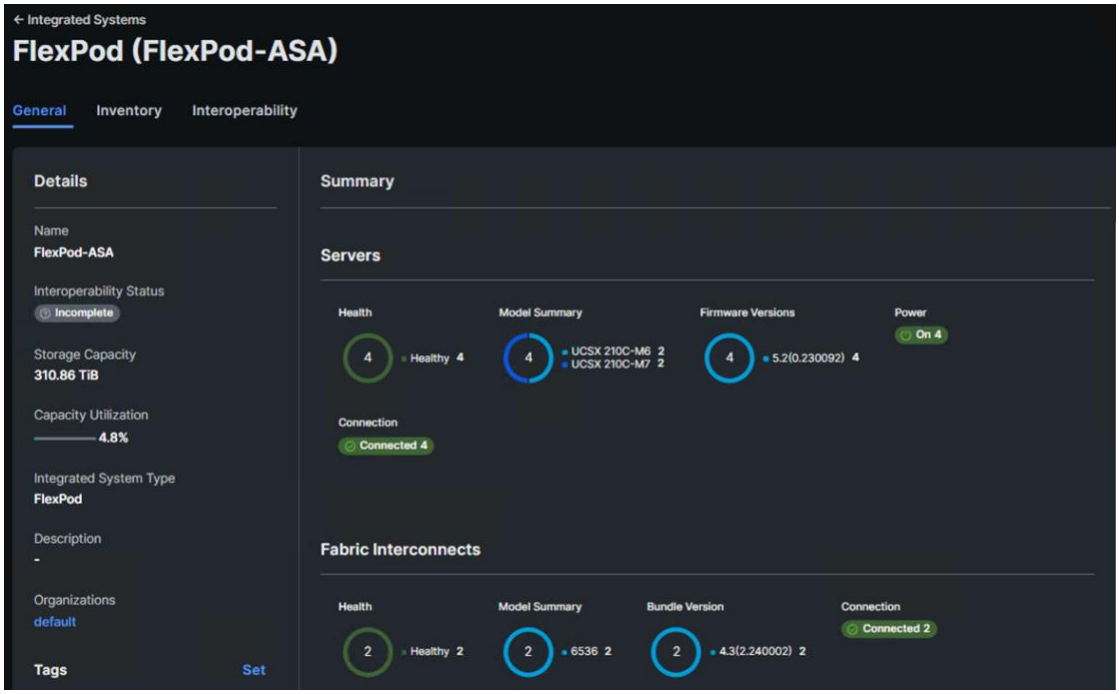
- Select the UCS Domain used in this FlexPod and click Next.
- Select the two Cisco Nexus switches used in this FlexPod and click Next.
- Select the NetApp ASA A400 storage used in this FlexPod and click Next.

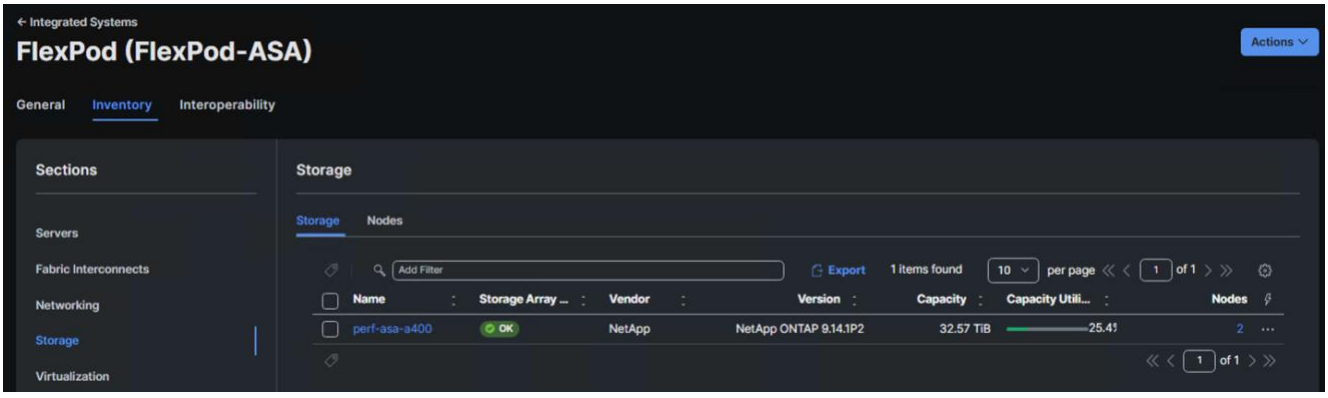


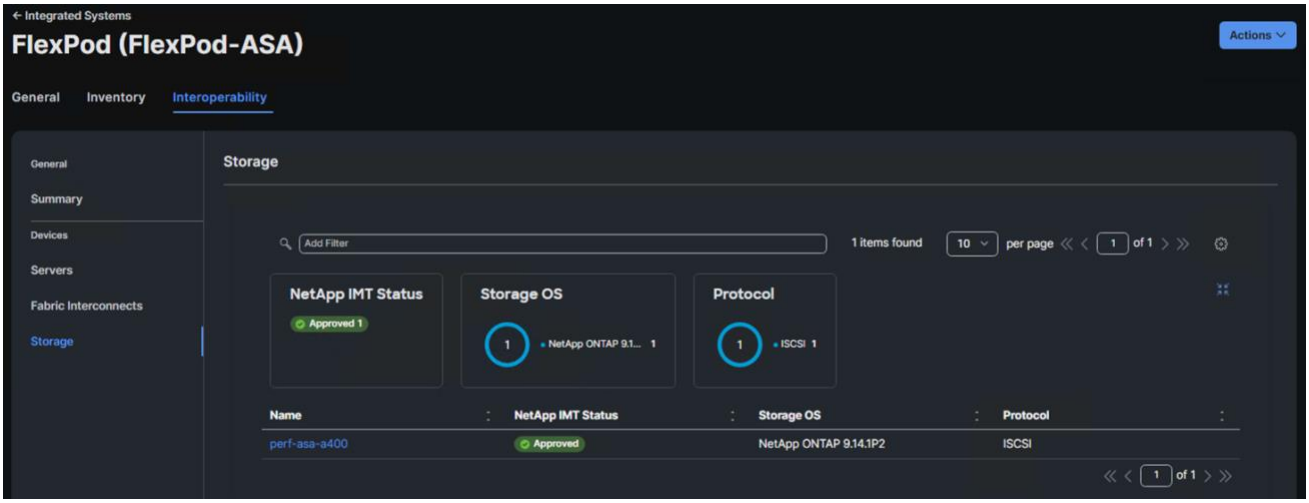
- Review the Summary information and click Create. After a few minutes, the FlexPod Integrated System will appear under Integrated Systems.



- Click the FlexPod name to see detailed General, Inventory, and Interoperability data on the FlexPod Integrated System.







Solution Validation

For FlexPod infrastructure setup, refer to the deployment guide for manual deployment. In this solution, UCS was configured in Intersight Managed Mode (IMM).

Refer to the IaC document if you want to proceed with automated deployment of FlexPod.

In this solution, NetApp ASA controller failover scenarios are tested for both bare metal and VMware virtualized deployments. Table 17 below lists the server details used for virtualization and bare metal validations. For the VMware virtualized configuration, VMFS and vVols datastores were provisioned for applications. For bare metal configuration, storage LUNs and NVMe namespaces were provisioned and mounted onto the server.

The workloads were simulated using HammerDB for SQL, GenIO for EHR and SLOB for Oracle applications, each having corresponding read and write characteristics.

To show the IO availability and data integrity of the solution following steps were ran for each use-case:

1. While the workloads were ramped up, a “halt” was issued on one of the controllers which would migrate the interfaces to the surviving controller
2. During the failover process application were issuing the IO in terms of reads and writes
3. Once the controller was fully brought down and, a “node” reboot was issued to bring the failed controller back while the workloads still running
4. While the failed node was brought up, the interfaces were automatically migrated back from the surviving nodes

Observation

1. During this process, all the paths to the storage devices were active and the whole failover process was transparent to the application workloads
2. The combined throughput from all three workloads was lesser when only one controller node was running which was expected.

The complete Cisco UCS compute platform comprises of the chassis, IFM module, servers, the attached VIC adapters, and the Cisco Fabric Interconnects are intelligently managed and configured using Cisco Intersight.

Note: The size of the storage devices [LUNS and Namespaces] could vary depending on the application deployment.

Table 17 Servers used for bare metal and VMware Virtualization configuration

Platform	UCS Servers
Bare Metal	fpsa-perf-g02-6536-1-5
Bare Metal	fpsa-perf-g02-6536-1-6
ESXi (FC SAN Boot)	fpsa-perf-g02-6536-1-1
ESXi (FC SAN Boot)	fpsa-perf-g02-6536-1-2
ESXi (iSCSI Boot)	fpsa-perf-g02-6536-1-3
ESXi (iSCSI Boot)	fpsa-perf-g02-6536-1-4

FlexPod solution is designed in such a way that customers would be able to create repeatable building blocks with the same exact configurations and recommendations. This is achieved through FlexPod IaC (Infrastructure as Code) process building the entire base FlexPod stack through D0 automation playbooks.

Note: As NFS protocol is not supported on an ASA system, the procedures related to NFS should be skipped.

Below is a detailed overview of failover test scenarios and their results.

VMware

1. VMFS

Storage layout

The table 18 below details the storage layout for VMFS with FCP & iSCSI.

Table 18 Storage layout for VMFS with FCP & iSCSI

Workload	LUNs			Datastores
	Category	**Number of LUNs**	**Each LUN Size**	**VMFS6 FC & iSCSI datastore mapped to the host** **Example:** ds:///vmfs/volumes/xyz
GenIO for EHR	data	8	800G	800G
	journal	2	300G	300G
SLOB for Oracle	data	8	500G	500G
	log	2	150G	150G
	grid	1	5G	5G
HammerDB for SQL	data	4	500G	500G
	log	2	200G	200G

a. FCP

- A volume for ESXi boot LUN was created to provide FC SAN boot.
- For all the workloads, equal distribution of volumes on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four FC logical interfaces [LIFs] were created, two on each storage node.
- After the storage configuration, the FC LUNs were mapped to the ESXi hosts in the VMware vCenter and then VMFS FC datastores were created using those fibre channel disks. Then, these datastores were mapped to the VMs (one VM each for Oracle, SQL, and EHR).
- The corresponding workloads were run on each VM and the datastore mounted to the VM was populated with workload data.
- When the storage failover test was performed by bringing down one controller, the workloads were still working fine, and no data loss was observed. This verifies the data availability and integrity of NetApp ASA in the event of failover.

Figure 18 represents the ASA system state during failover scenario which was taken from ONTAP System Manager.

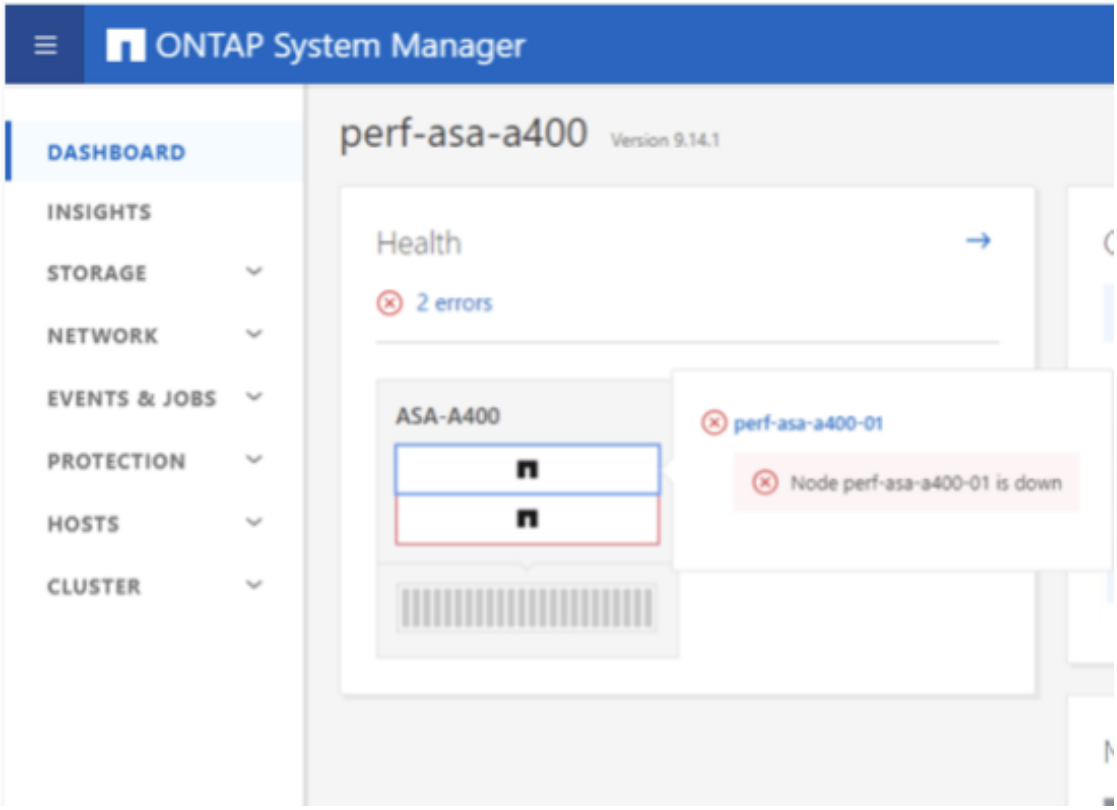


Figure 18 ASA Status on System Manager during storage failover

The VMFS datastore created on an FC disk and the multipathing feature of the ASA system for the datastore on an FC disk are depicted in Figure 19 and 20.

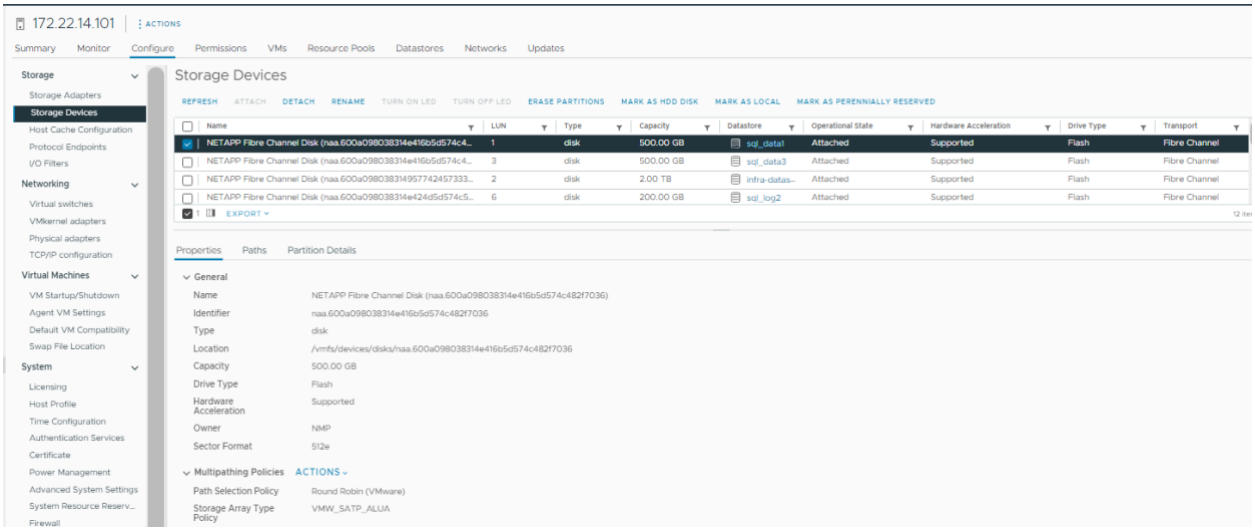


Figure 19 VMFS FCP datastore created on FC disk

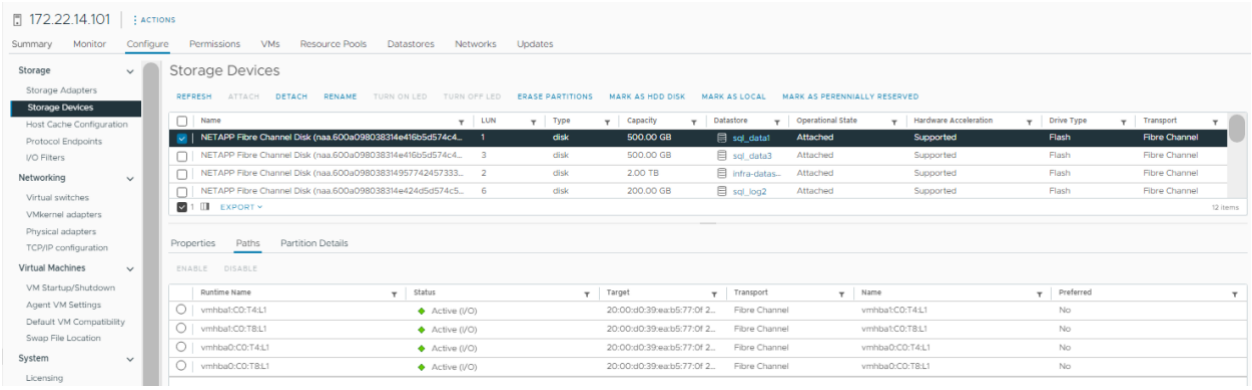


Figure 20 Multipathing for VMFS with FCP

b. iSCSI

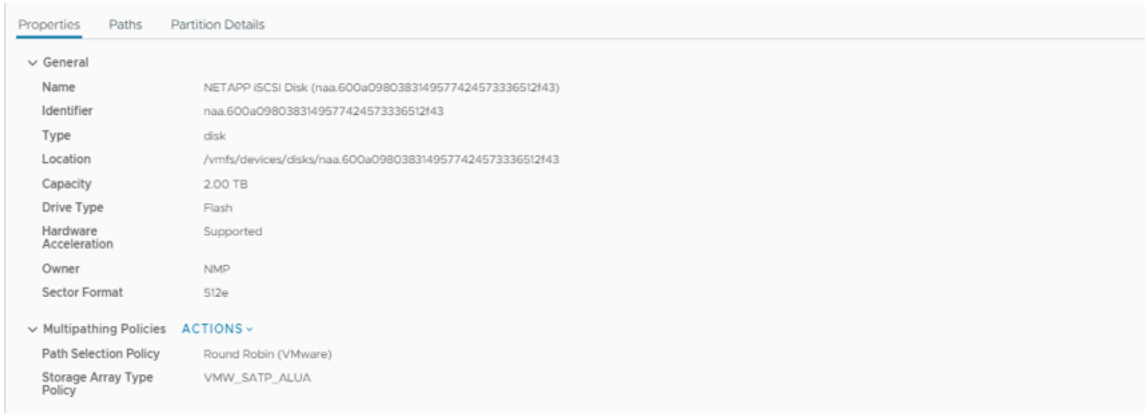
- A volume for ESXi boot LUN was created to provide iSCSI SAN boot.
- For all the workloads, equal distribution of volumes on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four iSCSI logical interfaces [LIFs] were created, two on each storage node.
- The required volume and LUN configuration for workloads using iSCSI was done in a similar way as FC, as mentioned above.

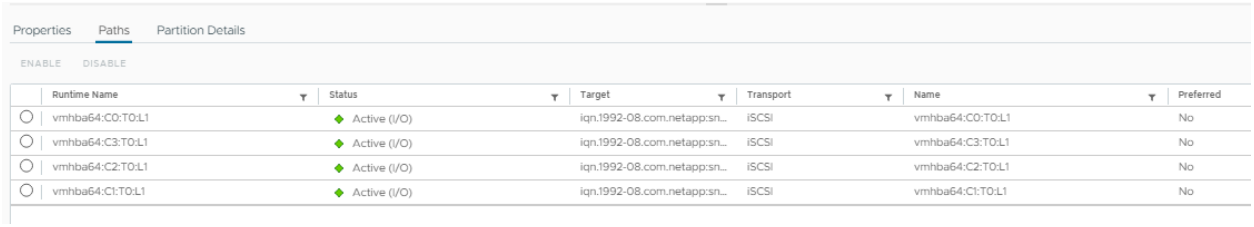
Note: With NetApp ASA, if there are iSCSI LIFs on the SVM that are not enabled for automatic LIF failover, the newly created LIFs will not be enabled for automatic LIF failover either. If automatic LIF failover is not enabled and there is a failover event, iSCSI LIFs will not migrate. Enable automatic failover for an iSCSI LIF using the below command:

```
network interface modify -vserver SVM_name -lif iSCSI_lif -failover-policy sfo-partner-only -
auto-revert true
```

```
perf-asa-a400::> net int modify -vserver Infra-SVM -lif iscsi-lif-* -failover-policy sfo-partner-only
  (network interface modify)
4 entries were modified.
```

- When the storage failover was performed by bringing down one controller, the workloads were still working fine, and no data loss was observed. Because of automatic failover being enabled, all the iSCSI LIFs on the node that was down were migrated to the active node. This verifies the data availability and integrity of NetApp ASA in the event of failover.





c. NVMe/FC

The table 19 below details the storage layout for VMFS with NVMe/FC & NVMe/TCP.

Table 19 Storage Layout for VMFS with NVMe/FC & NVMe/TCP

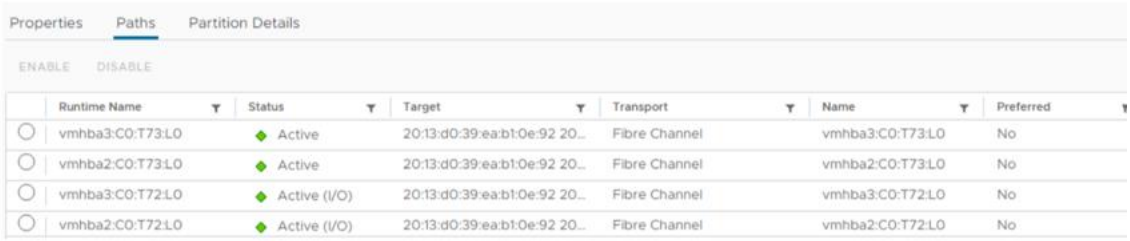
Workload	Namespace		Datastores	
	Category	**Number of namespaces**	**Each namespace Size**	**VMFS6 NVMe/FC & NVMe/TCP datastore mapped to the host** Example: ds:///vmfs/volumes/xyz
GenIO for EHR	data	8	800G	800G
	journal	2	300G	300G
SLOB for Oracle	data	8	500G	500G
	log	2	150G	150G
	grid	1	5G	5G
HammerDB for SQL	data	4	500G	500G
	log	2	200G	200G

- A volume for ESXi boot LUN was created to provide FC SAN boot.
- For all the workloads, equal distribution of namespaces on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four NVMe/FC logical interfaces [LIFs] were created, two on each storage node.
- One subsystem (fc_nvme_subsystem) was created and all the host NQNs were added.
- The required namespaces were mapped to this subsystem.
- NVMe datastores were mapped to all ESXi hosts and were also presented to VMs to run the workloads. The workloads on the VM used the disks presented via the NVMe/FC protocol. When the failover test was triggered with one node down, all the three applications and its data were intact.

Note: As this is NVMe, during the failover scenario only two LIFs that were present on the functioning node were active from the storage side.

```
perf-asa-a400::> network interface show -vserver Infra-SVM -data-protocol fc-nvme
            Logical    Status     Network            Current       Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node          Port    Home
----------- ---------- ---------- ------------------ ------------- ------- ----
Infra-SVM
            fc-nvme-lif-01a up/- 20:14:d0:39:ea:b1:0e:92 perf-asa-a400-01 1a true
            fc-nvme-lif-01b up/- 20:15:d0:39:ea:b1:0e:92 perf-asa-a400-01 1b true
            fc-nvme-lif-02a up/up 20:16:d0:39:ea:b1:0e:92 perf-asa-a400-02 1a true
            fc-nvme-lif-02b up/up 20:17:d0:39:ea:b1:0e:92 perf-asa-a400-02 1b true
```

From the host side, two paths for I/O were active which is an expected behavior for NVMe.



d. NVMe/TCP

- A volume for ESXi boot LUN was created to provide iSCSI SAN boot.
- For all the workloads, equal distribution of namespaces on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four NVMe/TCP logical interfaces [LIFs] were created, two on each storage node.
- One subsystem (nvme_tcp_subsystem) was created and all the host NQNs were added.
- The required namespaces were mapped to this subsystem.
- On the host side, NVMe drives were presented to the host via vHBA interfaces over the NVMe/TCP protocol. These datastores were then presented to VMs to run the workloads.
- The namespace configuration and datastore provisioning steps were similar to the NVMe/FC section mentioned above.
- In the event of storage failover, all the three workloads and its data were intact.

Figure 21 shows the pathing at the vSphere level for NVMe/TCP when one controller is down for SQL workload.

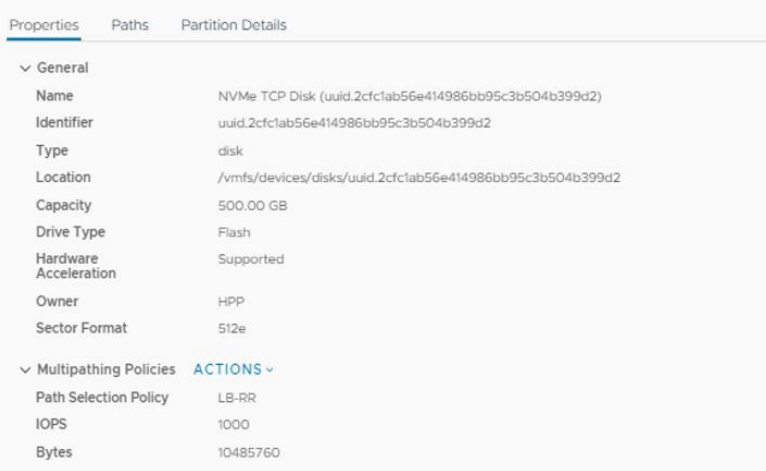


Figure 21 Multipathing for VMFS with NVMe/TCP

Two paths for I/O were active on the host side, which is an expected behavior for NVMe as shown in Figure 22.

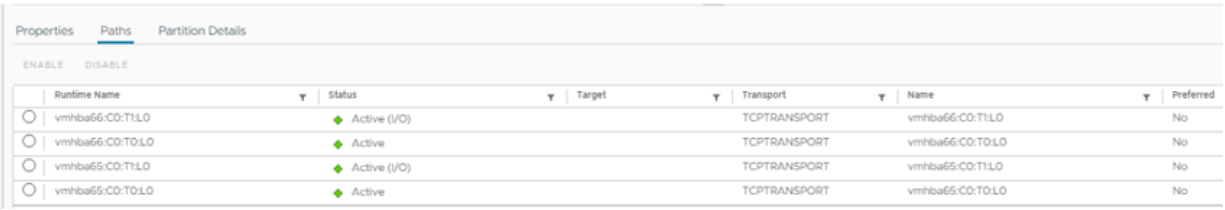


Figure 22 NVMe/TCP showing optimized and non-optimized paths

2. vVols

In this solution, NetApp ONTAP Tools for VMware vSphere (OTV) is used to provision vVols to the hosts and VMs. The detailed installation procedure of OTV can be referred to here.

Note: ONTAP Tools for VMware (OTV) is a requirement for vVols and it should be deployed in the vCenter. The provisioning of storage is done at the vCenter level through OTV, and underlying storage objects will be created automatically based on the datastore specifications.

Note: The multipathing depiction for vVols is the same as VMFS datastores in vCenter.

Storage layout

The table 20 below details the datastore layout for vVols with FCP & iSCSI.

Table 20 Storage Layout for vVols with FCP & iSCSI

Workload	vVols		
	Category	Number of Datastores	Size of Datastores
GenIO for EHR	data	8	800G
	journal	2	300G
SLOB for Oracle	data	8	500G
	log	2	150G

Workload	vVols		
	grid	1	5G
HammerDB for SQL	data	4	500G
	log	2	200G

a. FCP

- A volume for ESXi boot LUN was created to provide FC SAN boot.
- For all the workloads, equal distribution of vVols on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four FC logical interfaces [LIFs] were created, two on each storage node.
- After the storage configuration, the FCP vVols datastores were provisioned using ONTAP tools in vCenter. These datastores were then automatically mounted on the hosts and then manually mapped to the VMs. (one VM each for Oracle, SQL, and EHR workloads)
- The corresponding workloads were run on each VM and the vVols datastore mounted to the VMs were populated with workload data.
- One of the storage controllers was brought down to trigger a failover scenario. The workloads were migrated onto the active storage controller node and all the data was intact and no data loss was observed. This illustrates the data availability and integrity of NetApp ASA in the event of failover.

b. iSCSI

- One volume (esxi_boot) was created on one of the controllers with one LUN for iSCSI SAN boot.
- For all the workloads, equal distribution of vVols on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four iSCSI logical interfaces [LIFs] were created, two on each storage node.
- The storage configurations and vVols datastore provisioning for the host procedures were similar to the FCP section described above.
- In the event of storage failover, the workloads were migrated to the active controller node due to the automatic iSCSI LIF failover property of ASA. All the applications and data were intact, and no data loss was observed.

c. NVMe/FC

The table 21 below details the storage layout for vVols with NVMe/FC.

Table 21 Storage Layout for vVols with NVMe/FC

Workload	vVols		
	Category	**Number of Datastores**	**Size of Datastores**
GenIO for EHR	data	8	800G
	journal	2	300G
SLOB for Oracle	data	8	500G
	log	2	150G
	grid	1	5G
HammerDB for SQL	data	4	500G
	log	2	200G

- A volume for ESXi boot LUN was created to provide FC SAN boot.
- Four NVMe/FC logical interfaces [LIFs] were created, two on each storage node.
- On the host side, NVMe/FC vVols were provisioned for each VM using ONTAP Tools. These vVols (used the NVMe Fibre Channel disks) were then mapped to VMs to run the workloads.
- In the event of storage failover, the workloads were migrated to the active controller node. All the applications and data were intact, and no data loss was observed.

Bare Metal

In this solution, we have used the RHEL OS for bare metal installation.

Use the below command on the RHEL servers to view the current multipath configurations.

```
[root@Perf-BM ~]# multipath -ll
3600a0980383149577424573336516144 dm-10 NETAPP,LUN C-Mode
size=500G features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua' wp=rw
`-+- policy='service-time 0' prio=50 status=active
  |- 0:0:0:4  sde  8:64   active ready running
  |- 0:0:1:4  sdq  65:0   active ready running
  |- 1:0:0:4  sdac 65:192 active ready running
  `- 1:0:1:4  sdao 66:128 active ready running
3600a0980383149577424573336516145 dm-11 NETAPP,LUN C-Mode
size=500G features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua' wp=rw
`-+- policy='service-time 0' prio=50 status=active
  |- 0:0:0:6  sdg  8:96   active ready running
  |- 0:0:1:6  sds  65:32  active ready running
  |- 1:0:0:6  sdae 65:224 active ready running
  `- 1:0:1:6  sdaq 66:160 active ready running
```

Storage layout

The table 22 below details the storage layout for bare metal with FCP and iSCSI.

Table 22 Storage Layout for bare metal with FCP/iSCSI

Workload	LUNs		
	Category	**Number of LUNs**	**Each LUN Size**
GenIO for EHR	data	8	800G
	journal	2	300G
SLOB for Oracle	data	8	500G
	log	2	150G
	grid	1	5G

a. FCP

- A volume for RHEL boot LUN was created to provide FC SAN boot to the Bare Metal.
- For all the workloads, equal distribution of volumes on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four logical interfaces [LIFs] were created, two on each storage node.
- After the storage configuration, the LUNs were mounted on the hosts.
- The corresponding workloads were run on each host and were populated with workload data.

- When the storage failover was performed by bringing down one controller, the workloads were still working fine, and no data loss was observed. This verifies the data availability and integrity of NetApp ASA in the event of failover.

The LIF migration onto the active controller node during the failover process is depicted in the below image.

```
perf-asa-a400::> net interface  show -vserver Infra-SVM -lif fcp-lif-*
  (network interface show)
            Logical    Status     Network            Current       Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node          Port    Home
----------- ---------- ---------- ------------------ ------------- ------- ----
Infra-SVM
            fcp-lif-01a  up/up    20:0f:d0:39:ea:b1:0e:92 perf-asa-a400-02 1a false
            fcp-lif-01b  up/up    20:10:d0:39:ea:b1:0e:92 perf-asa-a400-02 1b false
            fcp-lif-02a  up/up    20:11:d0:39:ea:b1:0e:92 perf-asa-a400-02 1a true
            fcp-lif-02b  up/up    20:12:d0:39:ea:b1:0e:92 perf-asa-a400-02 1b true
4 entries were displayed.
```

b. iSCSI

- A volume for RHEL boot LUN was created to provide iSCSI boot to the Bare Metal.
- For all the workloads, equal distribution of volumes on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four iSCSI logical interfaces [LIFs] were created, two on each storage node.
- The storage configurations and LUN mounting for the host procedures were similar to the FCP sections described above. As iSCSI LIFs failover is enabled, the LIFs on the passive node gets migrated onto the active node during failover.

```
perf-asa-a400::> network interface show -vserver Infra-SVM -service-policy default-data-iscsi
             Logical       Status        Network              Current          Current    Is
Vserver      Interface     Admin/Oper    Address/Mask         Node             Port       Home
-----------  ----------    ----------    -----------------    -------------    -------    ----
Infra-SVM
             iscsi-lif-01a   up/up      192.168.111.21/24    perf-asa-a400-02   a0a-411    false
             iscsi-lif-01b   up/up      192.168.112.21/24    perf-asa-a400-02   a0a-412    false
             iscsi-lif-02a   up/up      192.168.111.22/24    perf-asa-a400-02   a0a-411    true
             iscsi-lif-02b   up/up      192.168.112.22/24    perf-asa-a400-02   a0a-412    true
```

- In the event of storage failover, the workloads were migrated to the active controller node due to the automatic iSCSI LIF failover property of ASA. All the applications and data were intact, and no data loss was observed.

c. NVMe/FC

- A volume for RHEL boot LUN was created to provide FC SAN boot to the Bare Metal.
- For all the workloads, equal distribution of namespaces on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four NVMe/FC logical interfaces [LIFs] were created, two on each storage node.

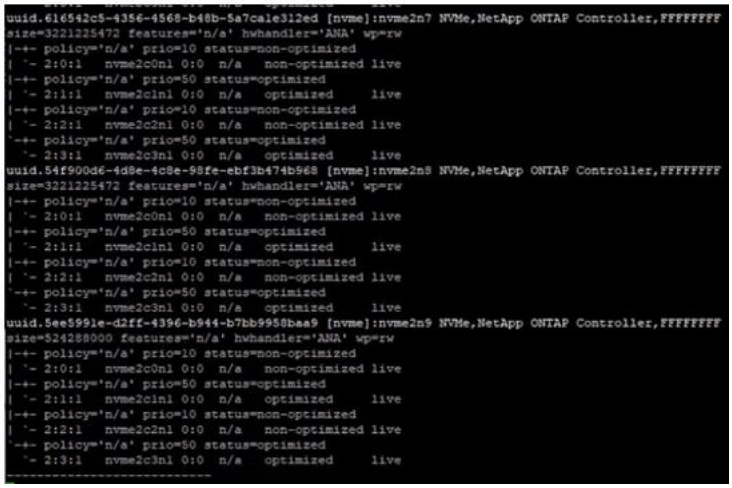
Table 23 Storage Layout for bare metal with NVMe/FC & NVMe/TCP

Workload	Namespaces		
	Category	**Number of Namespaces**	**Each Namespace Size**
GenIO for EHR	data	8	800G
	journal	2	300G

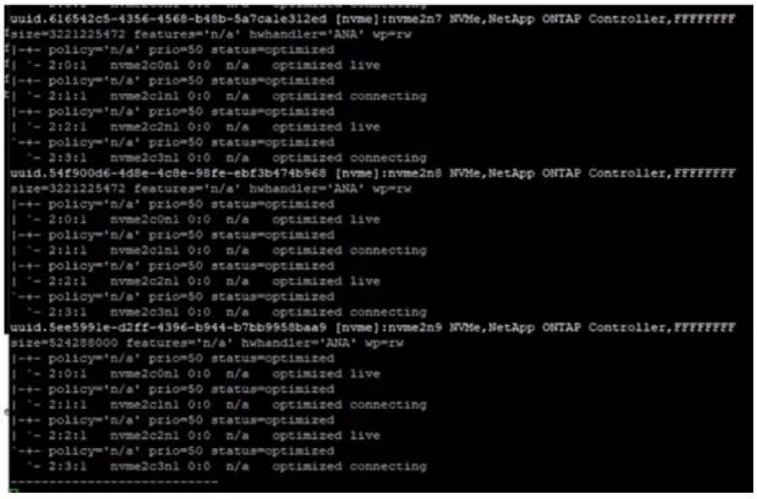
Workload	Namespaces		
SLOB for Oracle	data	8	500G
	log	2	150G
	grid	1	5G

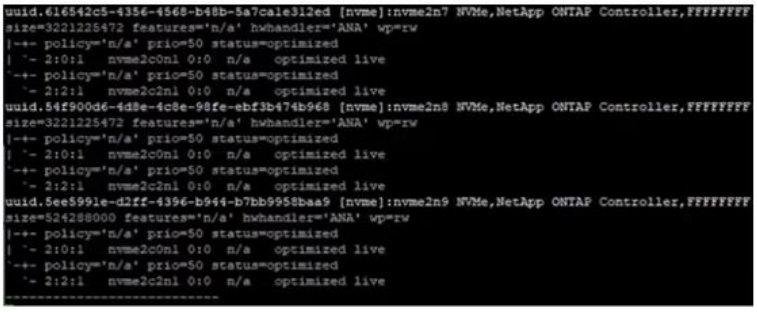
- One subsystem (fc_nvme_subsystem) was created and all the host NQNs were added.
- The required namespaces were mapped to this subsystem.
- On the host side, NVMe Fibre Channel disks were presented to the servers via vHBA interfaces to run the workloads via NVMe/FC protocol.
- When the failover test was triggered with one node down, all the three applications and its data were intact.

Below image depicts the status of disks mapped over NVMe/FC before failover.



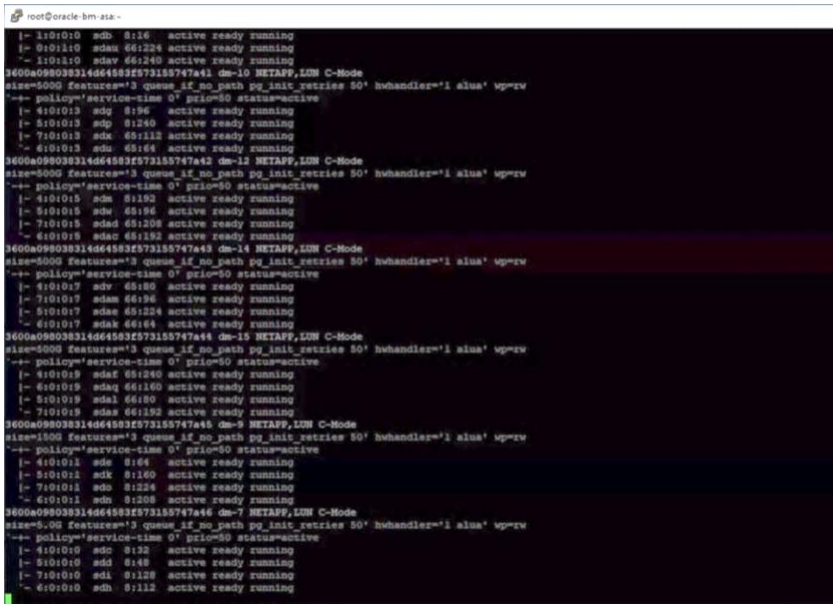
Below images depicts the status of disks mapped over NVMe/FC during the failover operation. As explained earlier, for NVMe use-cases there will be optimized and non-optimized paths and during a failover process only optimized paths will be available until the failed controller node is brought back up.





d. NVMe/TCP

- One volume (esxi_boot) was created on one of the controllers with one LUN for iSCSI SAN boot.
- For all the workloads, equal distribution of namespaces on each of the ASA A400 controllers to distribute the data traffic and the optimal CPU parallelization on the storage.
- Four NVMe/TCP logical interfaces [LIFs] were created, two on each storage node.
- One subsystem (nvme_tcp_subsystem) was created and all the host NQNs were added.
- The required namespaces were mapped to this subsystem.
- On the host side, NVMe drives are presented via vHBA interfaces over NVMe/TCP protocol to run the workloads.
- The namespace configuration and disks provisioning are similar to the NVMe/FC section mentioned above.
- In the event of storage failover, all the three workloads and its data were intact.



Summary

NetApp ASA avoids delays during failover with the active-active design. With both controllers constantly active and accessible by the host, there's no need to switch paths during a failover. ASA takes this a step further by continuously syncing crucial LUN metadata between controllers. This means the surviving controller has the information it needs to take over the failed controller's drives immediately, minimizing disruption to data access. The key metric here is the total time it takes for applications to resume normal I/O operations. While the storage system itself might recover quickly, there might be a slight delay for the host to fully utilize the failed controller's resources. This additional time is factored into the total I/O resumption timeframe.

Conclusion

FlexPod Data center is a predesigned, best-practice data center architecture that is built on the cloud-scale technologies offered by Cisco and NetApp, enabling management synergies across the complete IT infrastructure environment. FlexPod is an optimal platform for both Bare metal and Virtualization deployments of enterprise workloads. NetApp ASA systems deliver SAN-optimized storage for both capacity-intensive (ASA C-Series) and performance-intensive (ASA A-Series) workloads. The efficiency of the ASA system rendered through the active-active multipathing coupled with the reliable design of FlexPod architecture guarantees optimal customer experience and simplifies Hybrid Cloud integration.

The solution validation tests and results documented include both bare metal and VMware virtualized datastores (VMFS and vVols) implementing block storage protocol configurations covering enterprise workloads such as Epic EHR for Hospitals, Oracle, and SQL for businesses. The resiliency tests carried out validate the reliability of the solution design for mission critical workloads which are looking for six 9’s of availability and data integrity. This document serves as the future-proof reference for subsequent block storage architecture designs built on FlexPod with the NetApp ASA platform.

Acknowledgment

Abhinav Singh - Sr. Technical Marketing Engr, NetApp

Jyh-shing Chen - Sr. Technical Marketing Engr, NetApp

Anand Louis - Principal Product Manager, NetApp

Chance Bingen - Technical Marketing Engr, NetApp

Jeffrey Steiner - Principal Technical Marketing Engr, NetApp

Dr. Neto from Brazil – CPOC, NetApp

Where to find additional information

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

- FlexPod Home Page: https://www.flexpod.com
- NetApp Interoperability Matrix Tool: http://support.netapp.com/matrix/
- Cisco UCS Hardware and Software Interoperability Tool: http://www.cisco.com/web/techdoc/ucs/interoperability/matrix/matrix.html
- NetApp Product Documentation: https://www.netapp.com/support-and-training/documentation/
- Cisco Validated Design and deployment guides for FlexPod: https://www.cisco.com/c/en/us/solutions/design-zone/data-center-design-guides/flexpod-design-guides.html
- NetApp All Flash SAN Storage Array: https://www.netapp.com/data-storage/all-flash-san-storage-array/
- FlexPod Datacenter Zero Trust Framework Design Guide: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_zero_trust_design.html

Version history

Version	Date	Document version history
Version 1.0	May 2024	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

NetApp