End-to-End NVMe for FlexPod with Cisco UCSM, VMware vSphere 7.0, and NetApp ONTAP 9

FlexPod

NetApp
March 25, 2024

Table of Contents

End-to-End NVMe for FlexPod with Cisco UCSM, VMware vSphere 7.0, and NetApp ONTAP 9

TR-4914: End-to-End NVMe for FlexPod with Cisco UCSM, VMware vSphere 7.0, and NetApp ONTAP 9

Chris Schmitt and Kamini Singh, NetApp

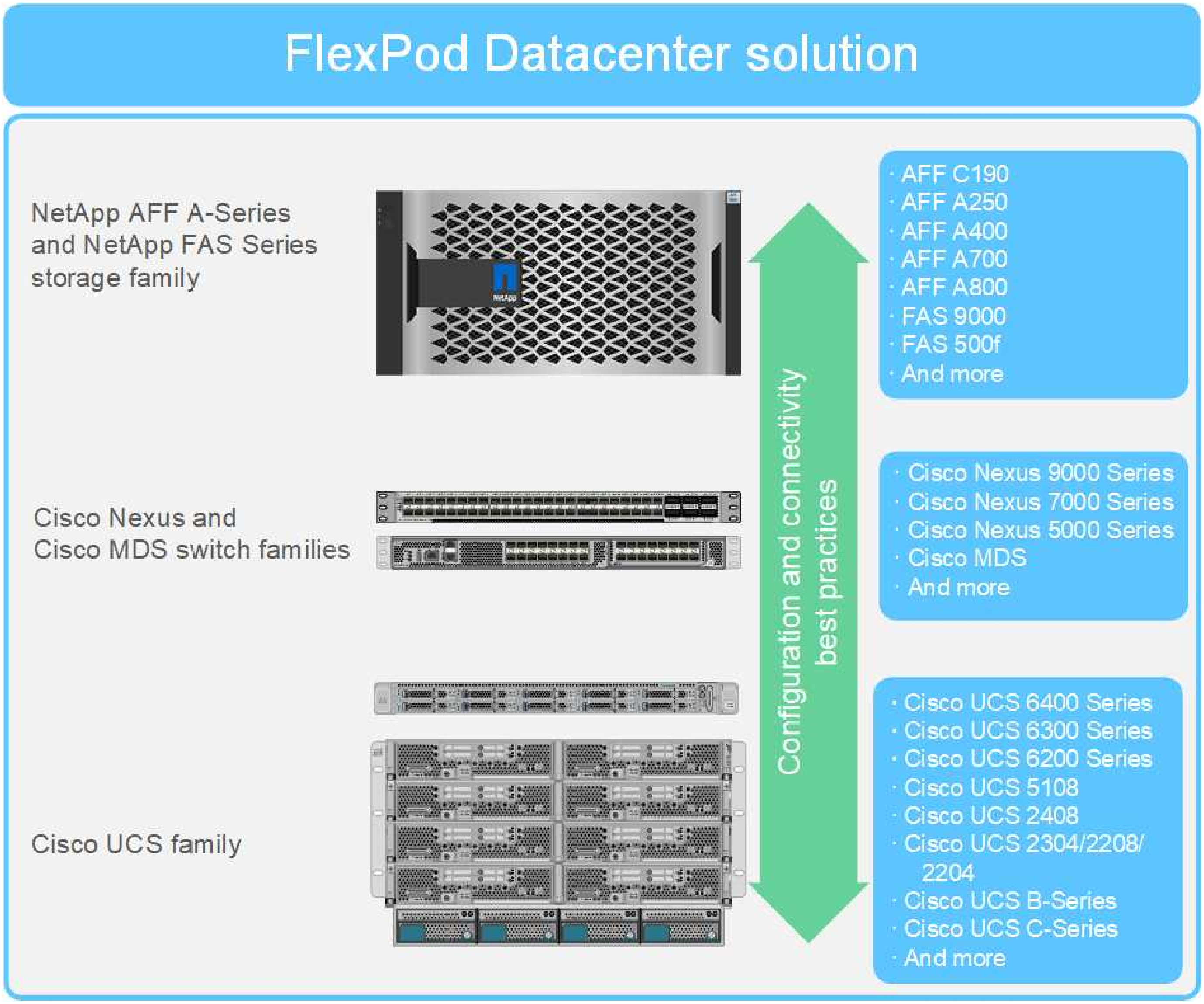
In partnership with:

The NVMe data-storage standard, an emerging core technology, is transforming enterprise data storage access and transport by delivering very high bandwidth and very low latency storage access for current and future memory technologies. NVMe replaces the SCSI command set with the NVMe command set.

NVMe was designed to work with nonvolatile flash drives, multicore CPUs, and gigabytes of memory. It also takes advantage of the significant advances in computer science since the 1970s, enabling streamlined command sets that more efficiently parse and manipulate data. An end-to-end NVMe architecture also enables data center administrators to rethink the extent to which they can push their virtualized and containerized environments and the amount of scalability that their transaction-oriented databases can support.

FlexPod is a best-practice data center architecture that includes the Cisco Unified Computing System (Cisco UCS), Cisco Nexus switches, Cisco MDS switches, and NetApp AFF systems. These components are connected and configured according to the best practices of both Cisco and NetApp to provide an excellent platform for running a variety of enterprise workloads with confidence. FlexPod can scale up for greater performance and capacity (adding compute, network, or storage resources individually as needed), or it can scale out for environments that require multiple consistent deployments (such as rolling out of additional FlexPod stacks).

The following figure presents the FlexPod component families.



FlexPod is the ideal platform for introducing FC-NVMe. It can be supported with the addition of the Cisco UCS VIC 1400 Series and Port Expander in existing Cisco UCS B200 M5 or M6 servers or Cisco UCS C-Series M5 or M6 Rack Servers and simple, nondisruptive software upgrades to the Cisco UCS system, the Cisco MDS 32Gbps switches, and the NetApp AFF storage arrays. After the supported hardware and software are in place, the configuration of FC-NVMe is similar to the FCP configuration.

NetApp ONTAP 9.5 and later provides a complete FC-NVMe solution. A nondisruptive ONTAP software update for AFF A300, AFF A400, AFF A700, AFF A700s, and AFF A800 arrays allow these devices to support an end-to-end NVMe storage stack. Therefore, servers with sixth-generation host bus adapters (HBAs) and NVMe driver support can communicate with these arrays using native NVMe.

Objective

This solution provides a high-level summary of the FC-NVMe performance with VMware vSphere 7 on FlexPod. The solution was verified to successfully pass FC-NVMe traffic, and performance metrices were captured for FC-NVMe with various data block sizes.

Solution benefits

End-to-end NVMe for FlexPod delivers exceptional value for customers with the following solution benefits:

- NVMe relies on PCIe, a high-speed and high-bandwidth hardware protocol that is substantially faster than older standards such as SCSI, SAS, and SATA. High-bandwidth, ultra-low latency connectivity between the Cisco UCS Server and NetApp storage array for most of the demanding applications.
- An FC-NVMe solution is lossless and can handle the scalability requirements of next-generation applications. These new technologies include artificial intelligence (AI), machine learning (ML), deep learning (DL), real-time analytics, and other mission-critical applications.
- Reduces the cost of IT by efficiently using all resources throughout the stack.
- Dramatically reduces response times and boosts application performance, which corresponds to improved IOPS and throughput with reduced latency. The solution delivers ~60% more performance and reduces latency by ~50% for existing workloads.
- FC-NVMe is a streamlined protocol with excellent queuing capabilities, especially in situations with more I/O operations per second (IOPS; that is, more transactions) and parallel activities.
- Offers nondisruptive software upgrades to the FlexPod components such as Cisco UCS, Cisco MDS, and the NetApp AFF storage arrays. Requires no modification to applications.

Next: Testing approach.

Testing approach

Previous: Introduction.

This section provides a high-level summary of the FC-NVMe on FlexPod validation testing. It includes both the test environment/configuration and the test plan adopted to perform the workload testing with respect to FC-NVMe for FlexPod with VMware vSphere 7.

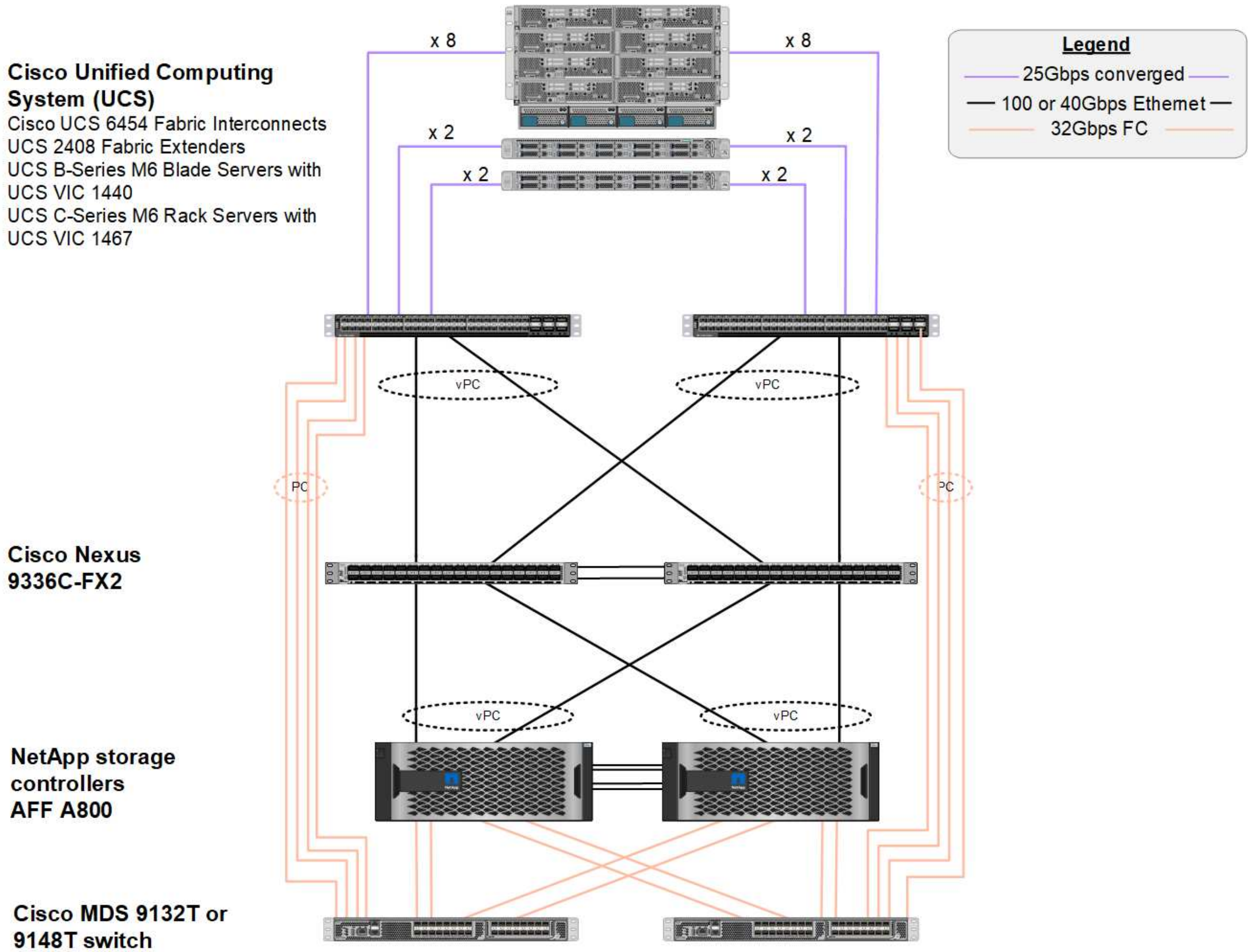
Test environment

The Cisco Nexus 9000 Series Switches support two modes of operation:

- NX-OS standalone mode, using Cisco NX-OS software
- ACI fabric mode, using the Cisco Application Centric Infrastructure (Cisco ACI) platform

In standalone mode, the switch performs like a typical Cisco Nexus switch, with increased port density, low latency, and 40GbE and 100GbE connectivity.

FlexPod with NX-OS is designed to be fully redundant in the computing, network, and storage layers. There is no single point of failure from a device or traffic path perspective. The figure below shows the connection of the various elements of the latest FlexPod design used in this validation of FC-NVMe.



From an FC SAN perspective, this design uses the latest fourth-generation Cisco UCS 6454 fabric interconnects and the Cisco UCS VICs 1400 platform with port expander in the servers. The Cisco UCS B200 M6 Blade Servers in the Cisco UCS chassis use the Cisco UCS VIC 1440 with Port Expander connected to the Cisco UCS 2408 Fabric Extender IOM, and each Fibre Channel over Ethernet (FCoE) virtual host bus adapter (vHBA) has a speed of 40Gbps. The Cisco UCS C220 M5 Rack Servers managed by Cisco UCS use the Cisco UCS VIC 1457 with two 25Gbps interfaces to each Fabric Interconnect. Each C220 M5 FCoE vHBA has a speed of 50Gbps.

The fabric interconnects connect through 32Gbps SAN port channels to the latest-generation Cisco MDS 9148T or 9132T FC switches. The connectivity between the Cisco MDS switches and the NetApp AFF A800 storage cluster is also 32Gbps FC. This configuration supports 32Gbps FC, for Fibre Channel Protocol (FCP), and FC-NVMe storage between the storage cluster and Cisco UCS. For this validation, four FC connections to each storage controller are used. On each storage controller, the four FC ports are used for both FCP and FC-NVMe protocols.

Connectivity between the Cisco Nexus switches and the latest-generation NetApp AFF A800 storage cluster is also 100Gbps with port channels on the storage controllers and vPCs on the switches. The NetApp AFF A800 storage controllers are equipped with NVMe disks on the higher-speed Peripheral Connect Interface Express (PCIe) bus.

The FlexPod implementation used in this validation is based on FlexPod Datacenter with Cisco UCS 4.2(1) in UCS Managed Mode, VMware vSphere 7.0U2, and NetApp ONTAP 9.9.

Validated hardware and software

The following table lists the hardware and software versions used during the solution validation process. Note that Cisco and NetApp have interoperability matrixes that should be referenced to determine support for any specific implementation of FlexPod. For more information, see the following resources:

- NetApp Interoperability Matrix Tool
- Cisco UCS Hardware and Software Interoperability Tool

Layer	Device	Image	Comments
Computing	• Two Cisco UCS 6454 Fabric Interconnects • One Cisco UCS 5108 blade chassis with two Cisco UCS 2408 I/O modules • Four Cisco UCS B200 M6 blades, each with one Cisco UCS VIC 1440 adapter and port expander card	Release 4.2(1f)	Includes Cisco UCS Manager, Cisco UCS VIC 1440, and port expander
CPU	Two Intel Xeon Gold 6330 CPUs at 2.0 GHz, with 42-MB Layer 3 cache and 28 cores per CPU	–	–
Memory	1024GB (16x 64GB DIMMS operating at 3200MHz)	–	–
Network	Two Cisco Nexus 9336C-FX2 switches in NX-OS standalone mode	Release 9.3(8)	–
Storage network	Two Cisco MDS 9132T 32Gbps 32-port FC switches	Release 8.4(2c)	Supports FC-NVMe SAN analytics
Storage	Two NetApp AFF A800 storage controllers with 24x 1.8TB NVMe SSDs	NetApp ONTAP 9.9.1P1	–
Software	Cisco UCS Manager	Release 4.2(1f)	–
	VMware vSphere	7.0U2	–
	VMware ESXi	7.0.2	–
	VMware ESXi native Fibre Channel NIC driver (NFNIC)	5.0.0.12	Supports FC-NVMe on VMware

Layer	Device	Image	Comments
	VMware ESXi native Ethernet NIC driver (NENIC)	1.0.35.0	–
Testing tool	FIO	3.19	–

Test plan

We developed a performance test plan to validate NVMe on FlexPod using a synthetic workload. This workload allowed us to execute 8KB random reads and writes as well as 64KB reads and writes. We used VMware ESXi hosts to run our test cases against the AFF A800 storage.

We used FIO, an open-source synthetic I/O tool that can be used for performance measurement, to generate our synthetic workload.

To complete our performance testing, we conducted several configuration steps on both the storage and servers. Below are the detailed steps for the implementation:

1. On the storage side, we created four storage virtual machines (SVMs, formerly known as Vservers), eight volumes per SVM, and one namespace per volume. We created 1TB volumes and 960GB namespaces. We created four LIFs per SVM as well as one subsystem per SVM. The SVM LIFs were evenly spread across the eight available FC ports on the cluster.
2. On the server side, we created a single virtual machine (VM) on each of our ESXi hosts, for a total of four VMs. We installed FIO on our servers to run the synthetic workloads.
3. After the storage and the VMs were configured, we were able to connect to the storage namespaces from the ESXi hosts. This allowed us to create datastores based on our namespace and then create Virtual Machine Disks (VMDKs) based on those datastores.

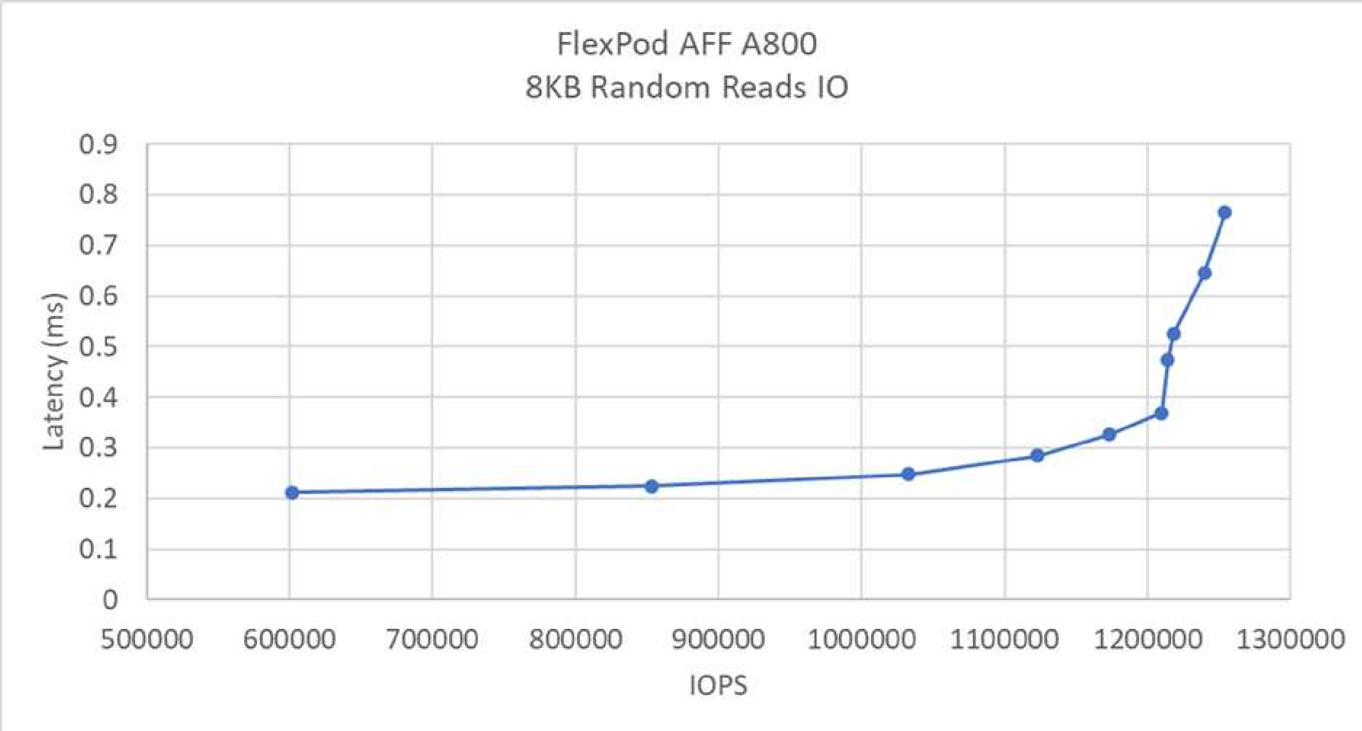
Next: Test results.

Test results

Previous: Testing approach.

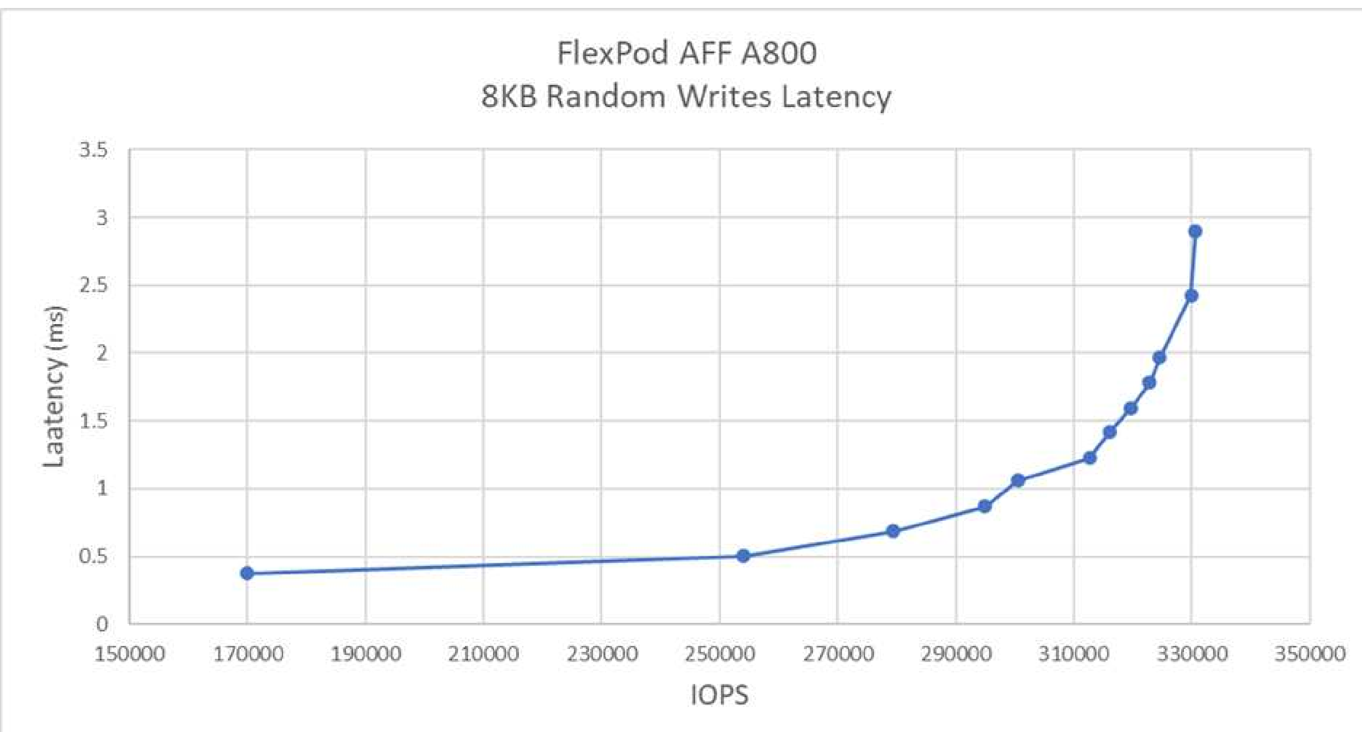
Testing consisted of running the FIO workloads to measure the FC-NVMe performance in terms of IOPS and latency.

The following graph illustrates our findings when running a 100% random read workload using 8KB block sizes.



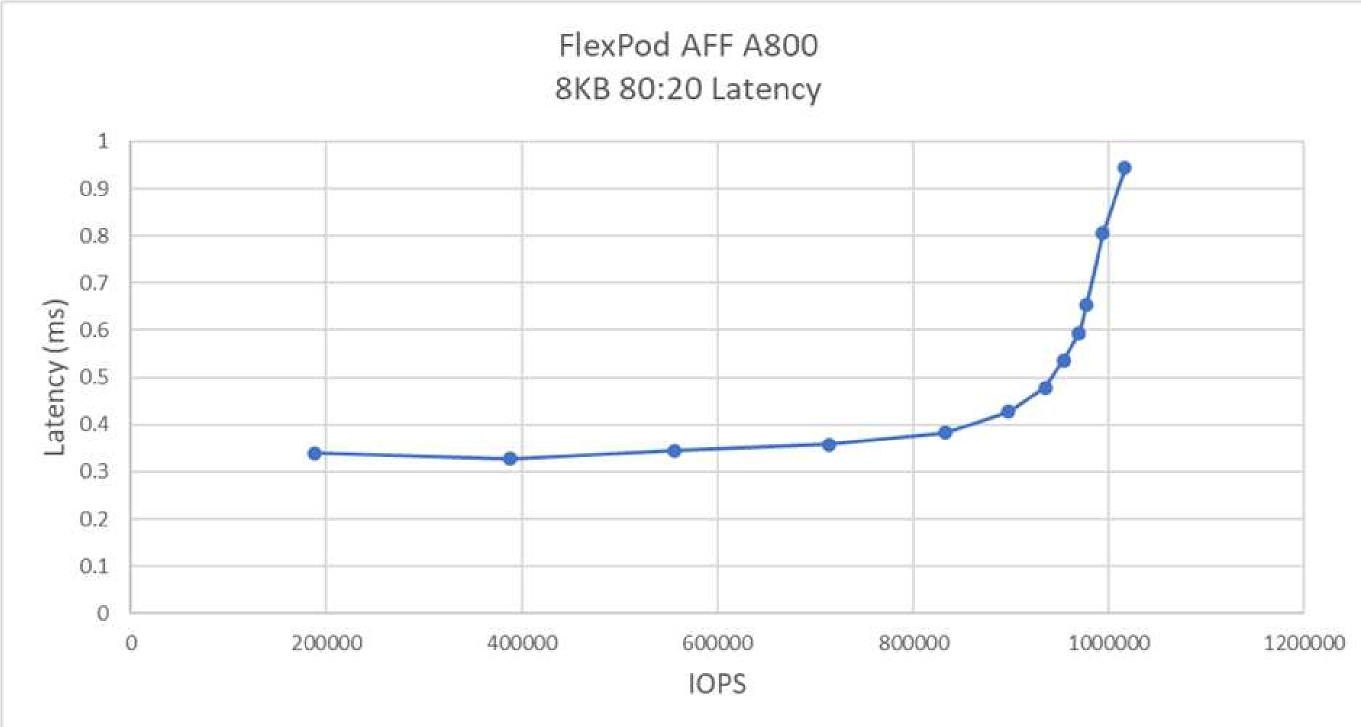
In our testing, we found that the system achieved over 1.2M IOPS while maintaining just under 0.35ms of server-side latency.

The following graph illustrates our findings when running a 100% random write workload using 8KB block sizes.



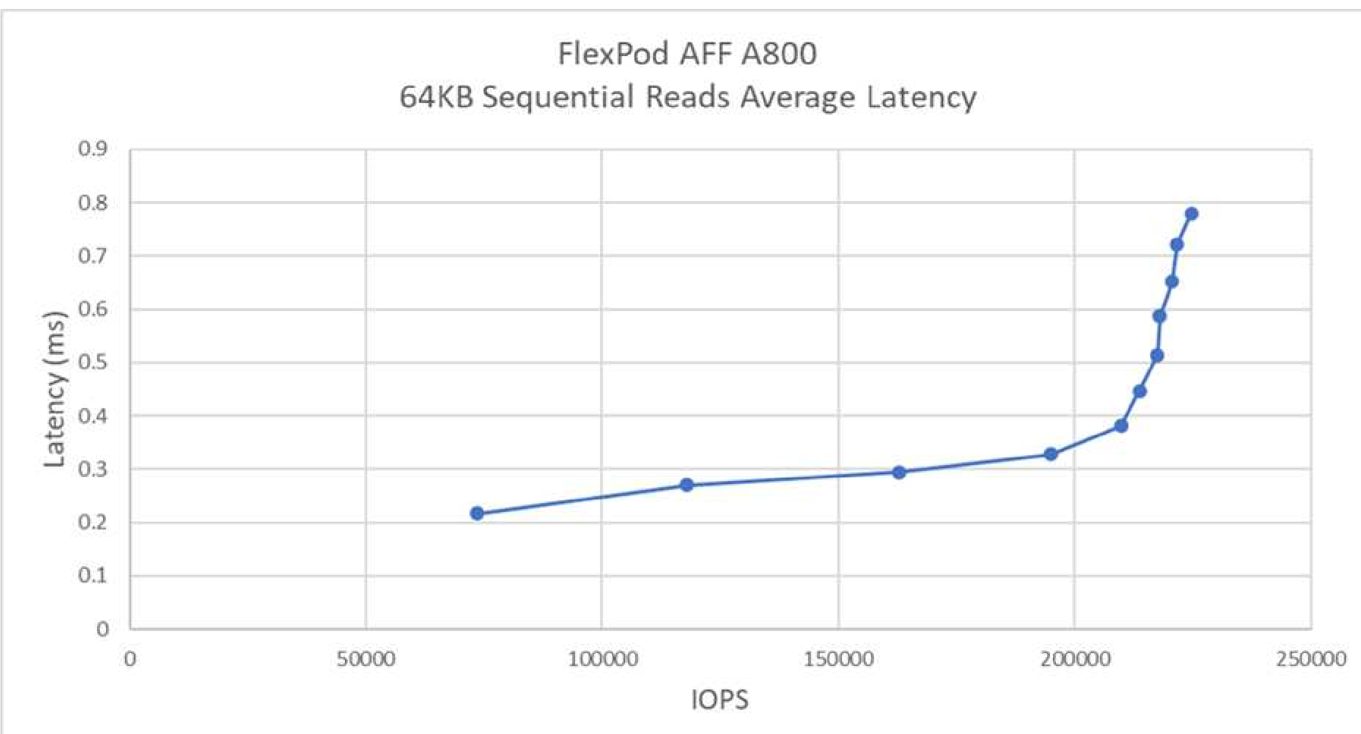
In our testing, we found that the system achieved close to 300k IOPS while maintaining just under 1ms of server-side latency.

For 8KB block size with 80% random reads and 20% writes, we observed the following results:



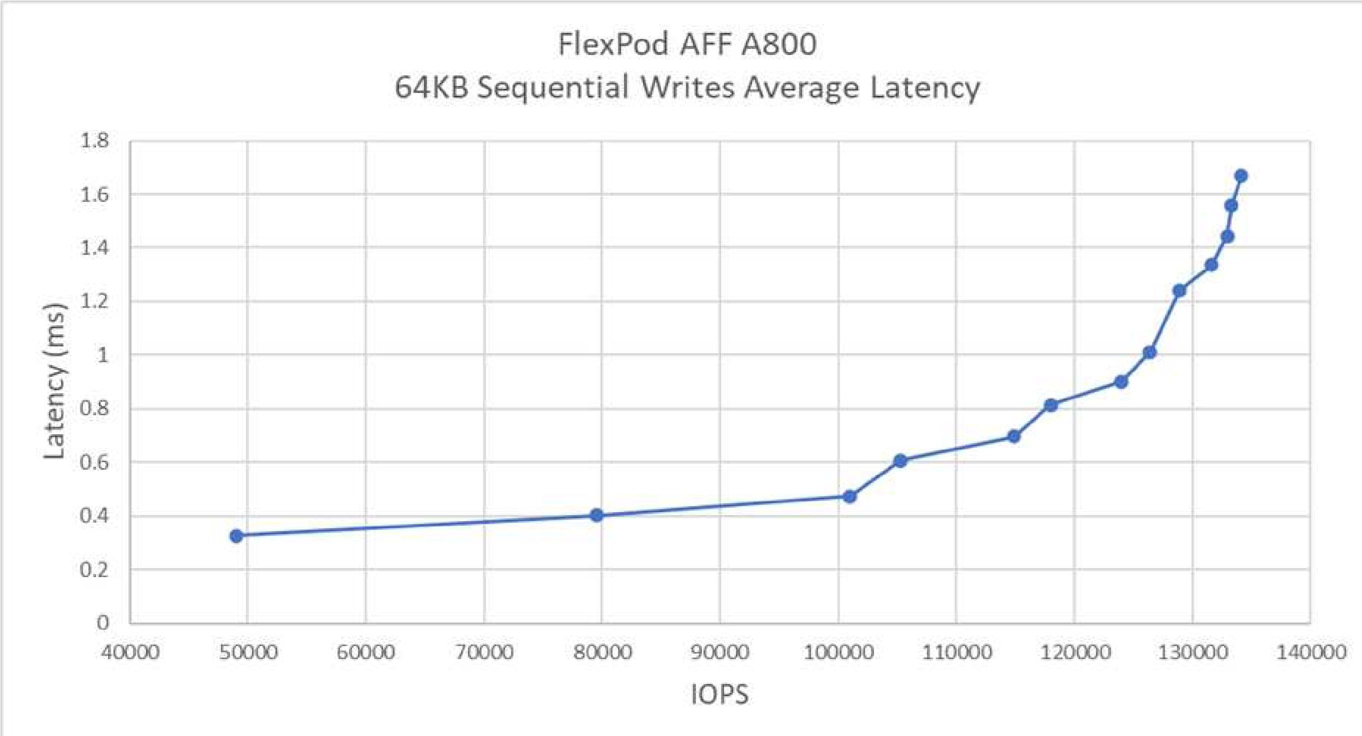
In our testing, we found that the system achieved over 1M IOPS while maintaining just under 1ms of server-side latency.

For 64KB block size and 100% sequential reads, we observed the following results:



In our testing, we found that the system achieved around 250k IOPS while maintaining just under 1ms of server-side latency.

For 64KB block size and 100% sequential writes, we observed the following results:



In our testing, we found that the system achieved around 120k IOPS while maintaining under 1ms of server-side latency.

Next: Conclusion.

Conclusion

Previous: Test results.

The observed throughput for this solution was 14GBps and 220k IOPS for a sequential read workload under 1ms latency. For random read workloads, we reached a throughput of 9.5GBps and 1.25M IOPS. The ability of FlexPod to provide this performance with FC-NVMe can address the needs of any mission-critical applications.

FlexPod Datacenter with VMware vSphere 7.0 U2 is the optimal shared infrastructure foundation to deploy FC-NVMe for a variety of IT workloads thereby providing high-performance storage access to applications that require it. As FC-NVMe evolves to include high availability, multipathing, and additional operating system support, FlexPod is well suited as the platform of choice, providing the scalability and reliability needed to support these capabilities.

With FlexPod, Cisco and NetApp have created a platform that is both flexible and scalable for multiple use cases and applications. With FC-NVMe, FlexPod adds another feature to help organizations efficiently and effectively support business-critical applications running simultaneously from the same shared infrastructure. The flexibility and scalability of FlexPod also enables customers to start with a right-sized infrastructure that can grow with and adapt to their evolving business requirements.

Additional information

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

- Cisco Unified Computing System (UCS)

  http://www.cisco.com/en/US/products/ps10265/index.html

- Cisco UCS 6400 Series Fabric Interconnects Data Sheet

  https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/datasheet-c78-741116.html

- Cisco UCS 5100 Series Blade Server Chassis

  http://www.cisco.com/en/US/products/ps10279/index.html

- Cisco UCS B-Series Blade Servers

  http://www.cisco.com/en/US/partner/products/ps10280/index.html

- Cisco UCS C-Series Rack Servers

  http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c-series-rack-servers/index.html

- Cisco Unified Computing System Adapters

  http://www.cisco.com/en/US/products/ps10277/prod_module_series_home.html

- Cisco UCS Manager

  http://www.cisco.com/en/US/products/ps10281/index.html

- Cisco Nexus 9000 Series Switches

  http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

- Cisco MDS 9000 Multilayer Fabric Switches

  http://www.cisco.com/c/en/us/products/storage-networking/mds-9000-series-multilayer-switches/index.html

- Cisco MDS 9132T 32-Gbps 32-Port Fibre Channel Switch

  https://www.cisco.com/c/en/us/products/collateral/storage-networking/mds-9100-series-multilayer-fabric-switches/datasheet-c78-739613.html

- NetApp ONTAP 9

  http://www.netapp.com/us/products/platform-os/ontap/index.aspx

- NetApp AFF A-Series

  http://www.netapp.com/us/products/storage-systems/all-flash-array/aff-a-series.aspx

- VMware vSphere

  https://www.vmware.com/products/vsphere

- VMware vCenter Server

  http://www.vmware.com/products/vcenter-server/overview.html

- Best Practices for modern SAN

  https://www.netapp.com/us/media/tr-4080.pdf

- Introducing End-to-End NVMe for FlexPod

  https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-b-series-blade-servers/whitepaper-c11-741907.html

Interoperability matrixes

- NetApp Interoperability Matrix Tool

  http://support.netapp.com/matrix/

- Cisco UCS Hardware Compatibility Matrix

  https://ucshcltool.cloudapps.cisco.com/public/

- VMware Compatibility Guide

  http://www.vmware.com/resources/compatibility

Acknowledgements

The authors would like to thank John George from Cisco and Scott Lane and Bobby Oommen from NetApp for the assistance and guidance offered during this project execution.