

Cisco UCS X210c M7 Compute Node Disk I/O Characterization

August 2023

Contents

Executive summary	3
Introduction	3
Scope of this document	6
Solution components	6
Workload characterization	7
Test configuration	7
NVMe SSD performance results	8
Performance summary	12
For more information	12
Appendix: Test environment	12

Executive summary

The Cisco UCS® X210c M7 Compute Node can integrate into the Cisco UCS X-Series Modular System. Up to eight computing nodes can reside in the Seven-Rack-Unit (7RU) Cisco UCS X9508 Chassis, offering one of the highest densities of computing, I/O, and storage resources per rack unit in the industry.

This document summarizes the Non-Volatile Memory Express (NVMe) I/O performance characteristics of the Cisco UCS X210c M7 Compute Node using NVMe Solid-State Disks (SSDs). The goal of this document is to help customers make well-informed decisions so that they can choose the right NVMe drives to meet their I/O workload needs.

Performance data was obtained using the Flexible I/O (Fio) measurement tool, with analysis based on the number of I/O Operations Per Second (IOPS) for random I/O workloads and Megabytes-per-second (MBps) throughput for sequential I/O workloads. From this analysis, specific recommendations are made for storage configuration parameters.

Introduction

The Cisco UCS X-Series Modular System simplifies your data center, adapting to the unpredictable needs of modern applications while also providing for traditional scale-out and enterprise workloads. It reduces the number of server types to maintain, helping to improve operational efficiency and agility as it helps reduce complexity. Powered by the Cisco Intersight® cloud-operations platform, it shifts your focus from administrative details to business outcomes—with hybrid-cloud infrastructure that is assembled from the cloud, shaped to your workloads, and continuously optimized.

Cisco UCS X-Series Modular System

The Cisco UCS X-Series Modular System begins with the Cisco UCS X9508 Chassis (Figure 1), engineered to be adaptable and ready for future requirements. The Cisco UCS 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 compute nodes that intersect with horizontally oriented I/O connectivity modules in the rear of the chassis. A unified Ethernet fabric is supplied with the Cisco UCS 9108 Intelligent Fabric Modules. In the future, Cisco UCS X-Fabric Technology interconnects will supply other industry-standard protocols as standards emerge. Interconnections can easily be updated with new modules.

Cisco UCS X-Series is powered by the [Cisco Intersight](#) platform, making it simple to deploy and manage at scale.

The Cisco UCS X9508 Chassis provides these features and benefits:

- The 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 Intelligent Fabric Modules at the top of the chassis connect the chassis to upstream Cisco UCS 6400 Series Fabric Interconnects. Each intelligent fabric module offers these features:
 - Up to 100 Gbps of unified fabric connectivity per compute node.
 - 8x 25-Gbps SFP28 uplink ports. The unified fabric carries management traffic to the Cisco Intersight cloud-operations platform, Fibre Channel over Ethernet (FCoE) traffic, and production Ethernet traffic to the fabric interconnects.
- 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 calls this connectivity Cisco UCS X-Fabric technology, because “X” is commonly used as a variable, signifying a system that can evolve with new technology developments.
- Six 2800-watt (W) Power Supply Units (PSUs) provide 54 volts (V) of power to the chassis with N, N+1, and N+N redundancy. A higher voltage allows efficient power delivery with less copper wiring needed and reduced power loss.
- Efficient, 4 x 100-mm, dual counter-rotating fans deliver industry-leading airflow and power efficiency. Optimized thermal algorithms enable different cooling modes to best support the network environment. Cooling is modular, so future enhancements can potentially handle open- or closed- loop liquid cooling to support even higher-power processors.



Figure 1.
Cisco UCS X9508 Chassis, front (left) and back (right)

Since Cisco first delivered the Cisco Unified Computing System™ (Cisco UCS) in 2009, our goal has been to simplify the data center. We pulled management out of servers and into the network. We simplified multiple networks into a single unified fabric. And we eliminated network layers in favor of a flat topology wrapped into a single unified system. With the Cisco UCS X-Series Modular System, the simplicity is extended even further:

- Simplify with cloud-operated infrastructure. We move management from the network into the cloud so that you can respond at the speed and scale of your business and manage all your infrastructure. You can shape Cisco UCS X-Series Modular System resources to workload requirements with the Cisco Intersight cloud-operations platform. You can integrate third-party devices, including storage from NetApp, Pure Storage, and Hitachi. In addition, you gain intelligent visualization, optimization, and orchestration for all your applications and infrastructure.

-
- Simplify with an adaptable system designed for modern applications. Today's cloud-native, hybrid applications are inherently unpredictable. They are deployed and redeployed as part of an iterative DevOps practice. Requirements change often, and you need a system that does not lock you into one set of resources when you find you need a different set. For hybrid applications, and for a range of traditional data-center applications, you can consolidate your resources on a single platform that combines the density and efficiency of blade servers with the expandability of rack servers. The result is better performance, automation, and efficiency.
 - Simplify with a system engineered for the future. Embrace emerging technology and reduce risk with a modular system designed to support future generations of processors, storage, nonvolatile memory, accelerators, and interconnects. Gone is the need to purchase, configure, maintain, power, and cool discrete management modules and servers. Cloud-based management is kept up-to-date automatically with a constant stream of new capabilities delivered by the Cisco Intersight Software-as-a-Service (SaaS) model.
 - Support a broader range of workloads. A single server type supporting a broader range of workloads means fewer different products to support, reduced training costs, and increased flexibility.

Cisco UCS X210c M7 Compute Node

The Cisco UCS X210c M7 Compute Node is the second generation of compute node to integrate into the Cisco UCS X-Series Modular System. It delivers performance, flexibility, and optimization for deployments in data centers, in the cloud, and at remote sites. This enterprise-class server offers market-leading performance, versatility, and density without compromise for workloads. Up to eight compute nodes can reside in the 7-Rack-Unit (7RU) Cisco UCS X9508 Chassis, offering one of the highest densities of compute, I/O, and storage per rack unit in the industry.

The Cisco UCS X210c M7 provides these main features:

- CPU: up to 2x 4th Gen Intel® Xeon® Scalable Processors with up to 60 cores per processor and up to 2.625 MB Level-3 cache per core and up to 112.5 MB per CPU.
- Memory: up to 8TB of main memory with 32x 256 GB DDR5-4800 DIMMs.
- Storage: up to six hot-pluggable, Solid-State Drives (SSDs), or Non-volatile Memory Express (NVMe) 2.5-inch drives with a choice of enterprise-class Redundant Array of Independent Disks (RAIDs) or passthrough controllers, up to two M.2 SATA drives with optional hardware RAID.
- Optional front mezzanine GPU module: the Cisco UCS front mezzanine GPU module is a passive PCIe Gen 4 front- mezzanine option with support for up to two U.2 NVMe drives and two HHHL GPUs.

mLOM virtual interface cards

- Cisco UCS Virtual Interface Card (VIC) 15420 occupies the server's modular LAN on motherboard (mLOM) slot, enabling up to 50 Gbps of unified fabric connectivity to each of the chassis Intelligent Fabric Modules (IFMs) for 100 Gbps connectivity per server.
- Cisco UCS Virtual Interface Card (VIC) 15231 occupies the server's modular LAN on motherboard (mLOM) slot, enabling up to 100 Gbps of unified fabric connectivity to each of the chassis Intelligent Fabric Modules (IFMs) for 100 Gbps connectivity per server.

Optional mezzanine card

- Cisco UCS 5th Gen Virtual Interface Card (VIC) 15422 can occupy the server's mezzanine slot at the bottom rear of the chassis. This card's I/O connectors link to Cisco UCS X-Fabric technology. An included bridge card extends this VIC's 2x 50 Gbps of network connections through IFM connectors, bringing the total bandwidth to 100 Gbps per fabric (for a total of 200 Gbps per server).
- Cisco UCS PCI Mezz card for the Cisco UCS X-Fabric can occupy the server's mezzanine slot at the bottom rear of the chassis. This card's I/O connectors link to Cisco UCS X-Fabric modules and enable connectivity to the Cisco UCS X440p PCIe Node.
- All VIC mezzanine cards also provide I/O connections from the X210c M7 compute node to the Cisco UCS X440p PCIe Node.
- Security: the server supports an optional Trusted Platform Module (TPM). Additional features include a secure boot FPGA and ACT2 anti-counterfeit provisions.



Figure 2.
Front view of Cisco UCS X210c M7 Compute Node

A specifications sheet for the Cisco UCS X210c M7 Compute Node is available at <https://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-x-series-modular-system/x210cm7-specsheet.pdf>.

Scope of this document

For the NVMe I/O performance characterization tests, performance was evaluated using NVMe SSDs for random and sequential access patterns with Cisco UCS X210c M7 Compute Node servers. This server supports up to six NVMe SSDs, and each drive is connected directly by PCIe Gen 4 x4 lanes to the CPU.

Solution components

The performance-tested solution used these components:

- Cisco UCS X210c M7 Compute Node.
- Cisco UCS X9508 Chassis.
- Solidigm D7-P5620 6.4 TB NVMe SSD.

Workload characterization

This section provides an overview of the specific access patterns used in the performance tests.

Tables 1 and 2 list the I/O mix ratios chosen for the sequential access and random access patterns, respectively.

Table 1. I/O mix ratio for sequential access pattern

I/O mode	I/O mix ratio (read:write)	
Sequential	100:0	0:100

Table 2. I/O mix ratio for random access pattern

I/O mode	I/O mix ratio (read:write)		
Random	100:0	0:100	70:30

Note: NVMe is configured in JBOD mode on all blades.

Test configuration

The test configuration was as follows:

- NVMe: six Solidigm D7-P5620 NVMe SFF SSDs were used on each X210c M7 compute node.
- Eight x Cisco UCS X210c M7 compute-node blades were installed on a Cisco UCS X9508 Chassis.
- Random workload tests were performed using NVMe SSDs for
 - 100-percent random read for 4- and 8-KB block sizes.
 - 100-percent random write for 4- and 8-KB block sizes.
 - 70:30-percent random read:write for 4- and 8-KB block sizes.
- Sequential workload tests were performed using NVMe SSDs for
 - 100-percent sequential read for 256-KB and 1-MB block sizes.
 - 100-percent sequential write for 256-KB and 1-MB block sizes.

Table 3 lists the recommended Fio settings.

Table 3. Recommended Fio settings

Name	Value
Fio version	Fio-3.19
File name	Device name on which Fio tests should run
Direct	For direct I/O, page cache is bypassed.
Type of test	Random I/O or sequential I/O, read, write, or mix of read and write

Name	Value
Block size	I/O block size: 4-, 8-, or 256-KB or 1-MB
I/O engine	Fio engine: libaio
I/O depth	Number of outstanding I/O instances
Number of jobs	Number of parallel threads to be run
Run time	Test run time
Name	Name for the test
Ramp-up time	Ramp-up time before the test starts
Time based	To limit the run time of the test

Note: The NVMe SSDs were tested with various combinations of outstanding I/O and numbers of jobs to get the best performance within an acceptable response time.

NVMe SSD performance results

Performance data was obtained using the Fio measurement tool, with analysis based on the IOPS rate for random I/O workloads and on MBps throughput for sequential I/O workloads. From this analysis, specific recommendations can be made for storage configuration parameters.

The server specifications and BIOS settings used in these performance characterization tests and Fio settings are detailed in the appendix, [Test environment](#).

The I/O performance test results capture the maximum read IOPS rate and bandwidth achieved with the NVMe SSDs within the acceptable response time (average latency). Latency is the time taken to complete a single I/O request from the viewpoint of the application. NVMe drives on these blades are directly managed by CPUs using PCIe Gen 4 x4 lanes.

NVMe SSD 6.4 TB high-performance, high-endurance performance test

For the NVMe SSD 6.4 TB high-performance and high-endurance performance test, the Cisco UCS X9508 Chassis is configured with eight Cisco UCS X210c M7 compute-node blades. Each blade is populated with six Solidigm 6.4 TB D7-Intel P5620 NVMe high-performance, high-endurance SSDs, and a total of 48 NVMe SSDs are configured on eight blades. The graphs in the following figures show the scaling of performance from one blade to eight blades with tests run in parallel for random and sequential I/O patterns.

Figure 3 shows the performance of the NVMe SSDs under test for X210c M7 compute-node blades with the X9508 chassis with a 100-percent random-read access pattern with both 4- and 8-KB block sizes. The graph shows the performance (for the 4-KB block size) of 6.6 million IOPS with an average latency of 200 microseconds when six NVMe SSDs are populated on a single X210c M7 compute node.

The aggregate performance of approximately 52 million IOPS with an average latency of 205 microseconds is achieved with all eight X210c M7 compute nodes populated with forty-eight NVMe drives in the chassis. This result demonstrates the linear scaling of IOPS without any variation in latency. A similar trend is observed for the 8-KB block size, as can be seen in the graph in Figure 3.

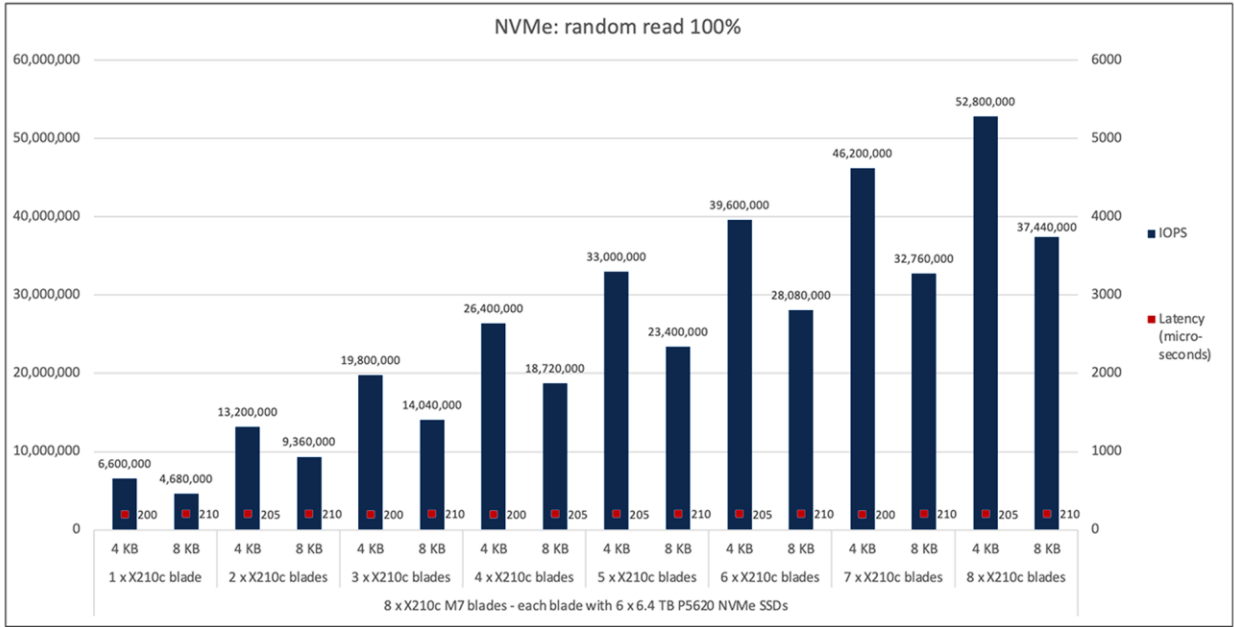


Figure 3.
Cisco UCS X210c M7 100-percent random read

Figure 4 shows the performance of the NVMe SSDs under test for eight X210c M7 compute-node blades with the X9508 chassis with a 100-percent random-write access pattern with 4- and 8-KB block sizes. The graph shows the performance of 2.4 million IOPS with an average latency of 100 microseconds with six NVMe SSDs on a single X210c M7 compute node for the 4-KB block size.

The aggregate performance of approximately 19.2 million IOPS with a latency of 100 microseconds is achieved with all eight X210c M7 compute nodes in the chassis. As with the random-read results, these results show linear IOPS scaling without any variation in latency. A similar trend is observed for the 8-KB block size, as can be seen from the graph in Figure 4.

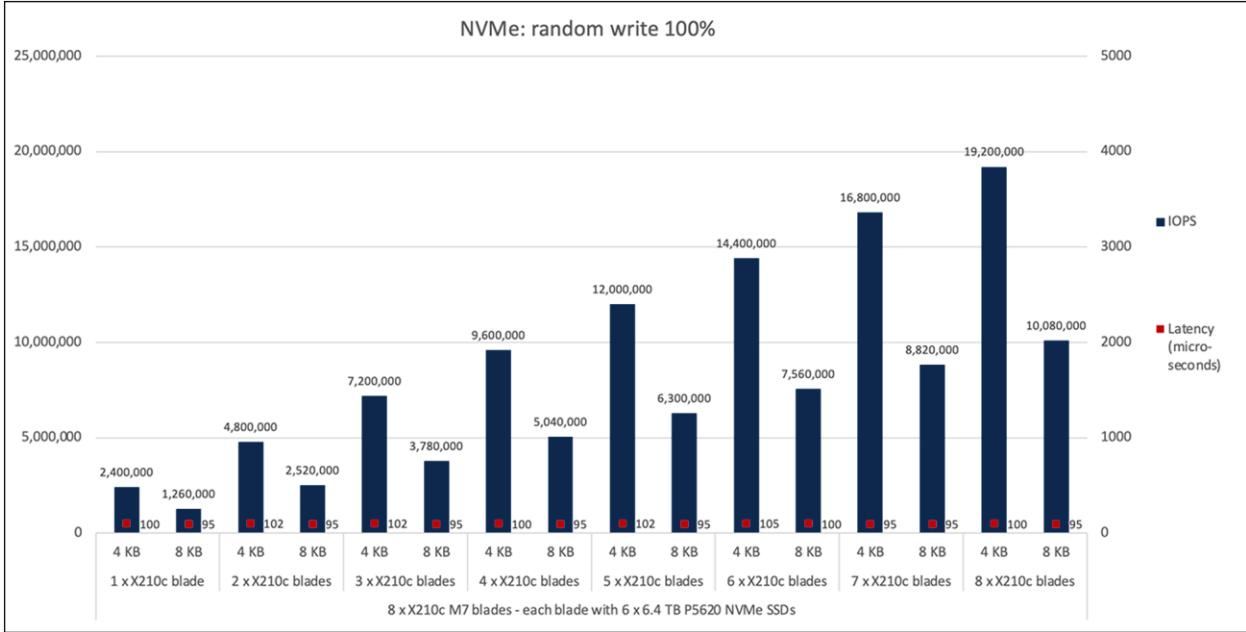


Figure 4.
Cisco UCS X210c M7 100-percent random write

Figure 5 shows the performance of the NVMe SSDs under test for eight X210c M7 compute-node blades with the X9508 chassis with a 70:30 percent random read:write access pattern. The graph shows the performance of approximately 3.2 million IOPS (combined read-and-write IOPS) with a read latency of 205 microseconds and a write latency of 100 microseconds with six NVMe SSDs on a single X210c M7 compute node for a 4-KB block size.

The aggregate performance of approximately 25 million IOPS (combined read-and-write IOPS) with a read latency of 205 microseconds and a write latency of 100 microseconds is achieved with a total of forty-eight NVMe SSDs on eight X210c M7 compute-node blades for the 4-KB block size, maintaining the linearity

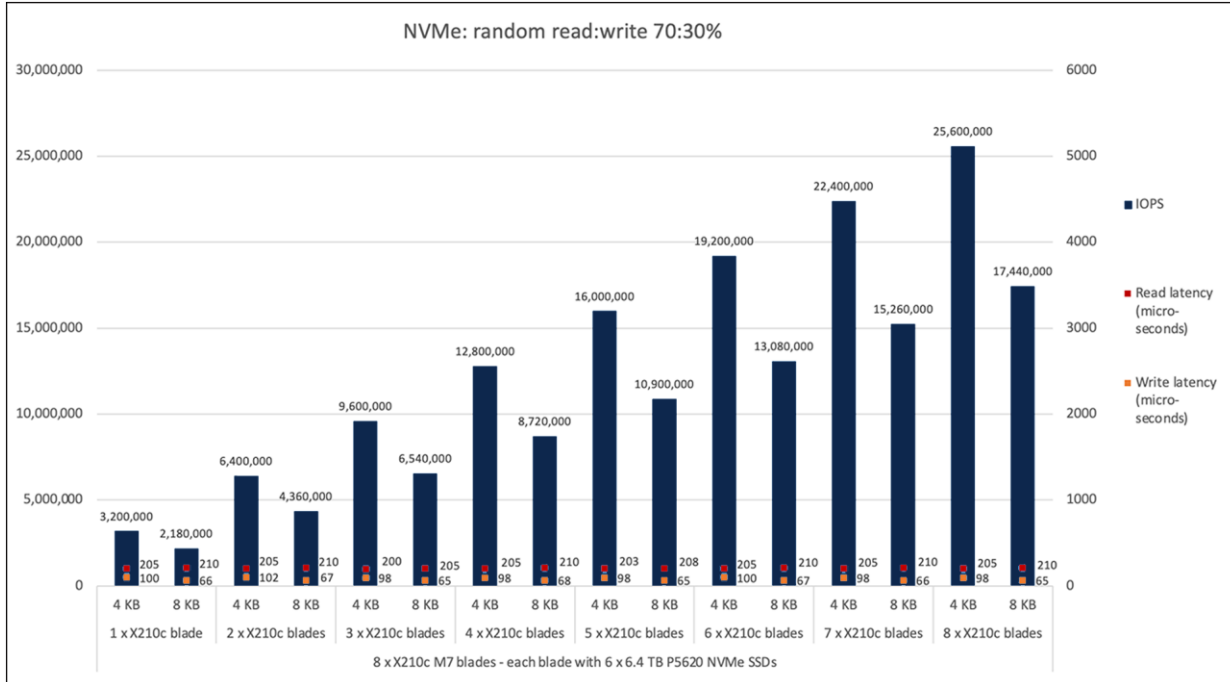


Figure 5. Cisco UCS X210c M7 70:30-percent random read:write

Figure 6 shows the performance of the NVMe SSDs under test for eight X210c M7 compute-node blades with the X9508 chassis with a 100-percent sequential-read access pattern. The graph shows the performance of 42,800 MBps with a latency of 650 microseconds with six NVMe SSDs on a single X210c M7 compute node for a 256-KB block size.

The aggregate performance of approximately 342,000 MBps with a latency of 650 microseconds is achieved with a total of forty-eight NVMe SSDs for a 256-KB block size. A similar linear trend (scaling from one blade to eight blades) is observed for a 1-MB block size as well.

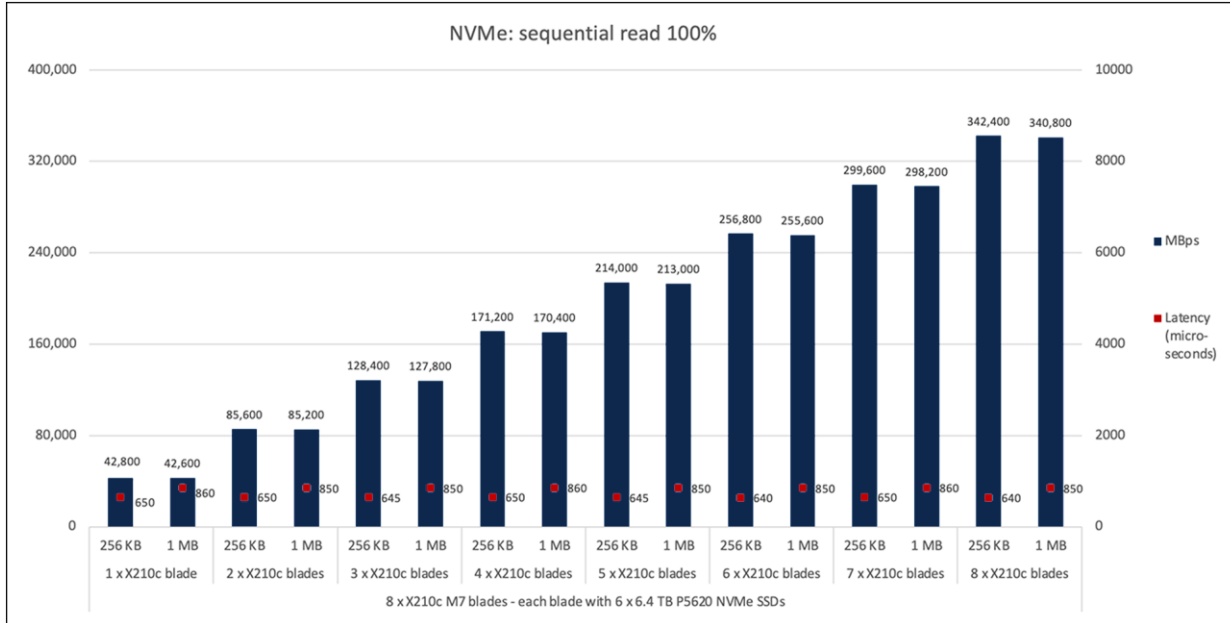


Figure 6.
Cisco UCS X210c M7 100-percent sequential read

Figure 7 shows the performance of the NVMe SSDs under test for eight X210c M7 compute-node blades with the X9508 chassis with a 100-percent sequential-write access pattern. The graph shows the performance of approximately 25,500 MBps with a latency of 150 microseconds with six NVMe SSDs on a single X210c M7 compute node for a 256-KB block size.

The aggregate performance of approximately 204,000 MBps with a latency of 150 microseconds is achieved with a total of forty-eight NVMe SSDs on eight X210c M7 compute-node blades for a 256-KB block size. A similar linear trend (scaling from one blade to eight blades) is observed for a 1-MB block size as well.

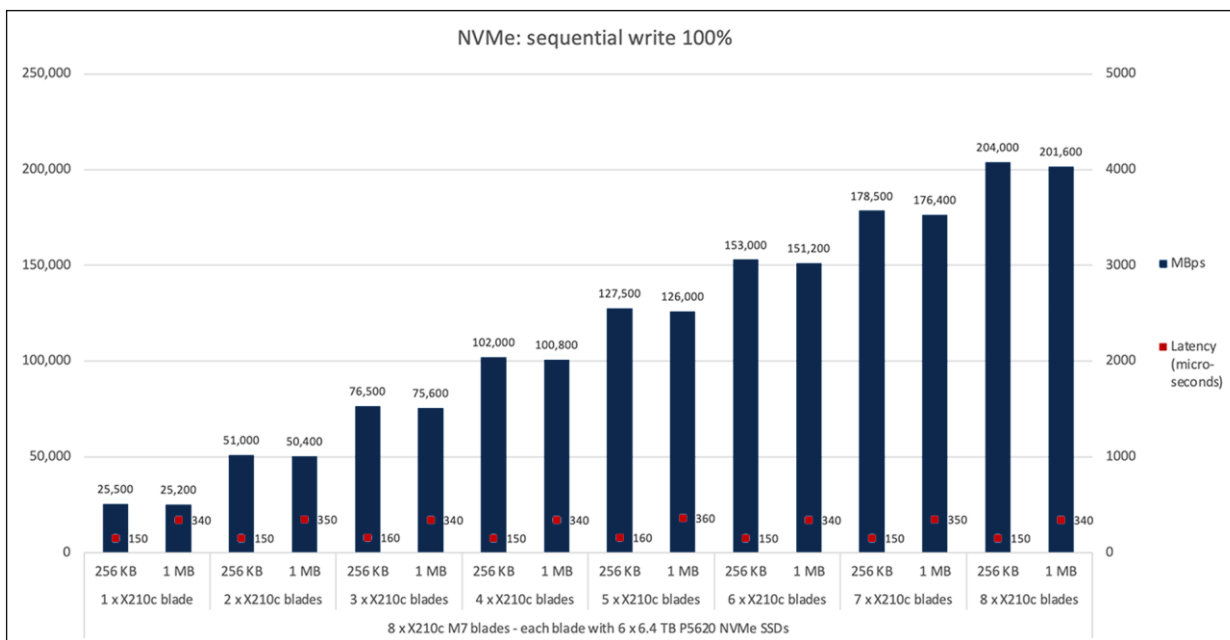


Figure 7.
Cisco UCS X210c M7 100-percent sequential write

Performance summary

The Cisco UCS X210c M7 Compute Node with all-NVMe drives configured is a good fit for applications that need higher compute and higher NVMe storage with lower latency. Given that there are PCI Gen 4 with x4 lanes for all of these NVMe slots, the overall server performance is a sum of individual drives with maximum performance and is not constrained, thus proving the effectiveness of the modular design of Cisco UCS x210c M7 compute nodes.

For more information

See the following resources for more information:

- For additional information about the Cisco UCS X-Series Modular System, refer to the following resources:
 - <https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-x-series-modular-system/index.html>.
 - <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-x-series-modular-system/solution-overview-c22-2432175.html?ccid=cc002456&oid=sowcsm025665>
- For information about the Cisco Intersight platform, refer to the Cisco Intersight configuration guide: <https://www.cisco.com/c/en/us/products/cloud-systems-management/intersight/index.html>.
- For information about BIOS tunings for various workloads, refer to the BIOS tuning guide: <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-b-series-blade-servers/ucs-m7-platforms-wp.html>.
- For information about the fio tool, refer to https://fio.readthedocs.io/en/latest/fio_doc.html.

Appendix: Test environment

Table 4 lists the details of the server under test.

Table 4. Blade properties

Name	Value
Product names	Cisco UCS X210c M7 Compute Node
CPUs	CPU: Two 2.2-GHz Intel Xeon Gold 6454S processors
Number of cores	32
Number of threads	64
Total memory	1024 GB

Name	Value
Memory DIMMs (16)	64 GB x 16 DIMMs
Memory speed	4800 MHz
VIC adapter	UCSX-ML-V5D200G-D: Cisco UCS VIC 15231 2x100/200G mLOM
SFF NVMe SSDs	6.4 TB 2.5-inch Intel D7-P5620 NVMe high-performance high-endurance SSDs (UCSX-NVME4-6400-D)

Table 5 lists the server BIOS settings applied for disk I/O testing.

Table 5. Server BIOS settings

Name	Value
Firmware version	5.2(0.230041)
BIOS version	X210M7.4.3.2a.0.0710230213
Fabric interconnect firmware version	9.3(5)I43(2a)
Cores enabled	All
Hardware prefetcher	Enable
Adjacent-cache-line prefetcher	Enable
Data Cache Unit (DCU) streamer	Enable
DCU IP prefetcher	Enable
Non-Uniform Memory Access (NUMA)	Enable
Memory refresh enable	1 x refresh
Energy-efficient turbo	Enable
Turbo mode	Enable
Extended Power Profile (EPP)	Performance
CPU C6 report	Enable
Package C-state	C0 and C1 states
Power performance tuning	OS controls Energy Performance Bias (EPB)
Workload configuration	I/O sensitive

Name	Value
UPI link enablement	Auto
UPI link power management	Enabled
UPI link speed	16 GT/s
Directory mode	Enabled

Note: The rest of the BIOS settings are platform default values.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at <https://www.cisco.com/go/offices>.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/go/trademarks>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)