

Cisco UCS C220 and C240 M7 Rack Server NVMe Disk I/O Characterization

June 2023

Contents

Executive summary	3
Introduction	3
NVMe storage solutions	4
I/O challenges	4
Scope of this document	4
Solution components	5
Workload characterization	7
Test configuration	8
NVMe SSD performance results	9
Performance summary	14
For more information	14
Appendix: Test environment	14

Executive summary

The Cisco UCS® C220 M7 Rack Server is a versatile general-purpose infrastructure and application server. This high-density, 1RU, 2-socket rack server delivers industry-leading performance and efficiency for a wide range of workloads, including virtualization, collaboration, and bare-metal applications. The Cisco UCS C240 M7 Rack Server is well suited for a wide range of storage and I/O-intensive applications such as big data analytics, databases, collaboration, virtualization, consolidation, and high-performance computing in its 2-socket, 2RU form factor.

Cisco UCS C220 M7 and UCS C240 M7 rack servers extend the capabilities of the Cisco UCS rack server portfolio. They incorporate 4th Gen Intel® Xeon® Scalable Processors with 50 percent more cores per socket, advanced features such as Intel Advanced Matrix Extensions (AMX), Data Streaming Accelerator (DSA), In-Memory Analytics Accelerator (IAA), and QuickAssist Technology (QAT). Many applications will see significant performance improvements on these new platforms. You can deploy the Cisco UCS C-Series rack servers as standalone servers or as part of the Cisco Unified Computing System™ managed by Cisco Intersight™ or Cisco UCS Manager to take advantage of Cisco® standards-based unified computing innovations that can help reduce your Total Cost of Ownership (TCO) and increase your business agility.

This document summarizes the Non-Volatile Memory express (NVMe) I/O performance characteristics of Cisco UCS C220 M7 and C240 M7 rack servers 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 platform with supported NVMe drives to meet their I/O workload needs. Performance data for the M7 rack servers with the supported number of NVMe SSDs was obtained using the 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 widespread adoption of virtualization and data-center consolidation technologies has profoundly affected the efficiency of the data center. Virtualization brings new challenges for storage technology, requiring the multiplexing of distinct I/O workloads across a single I/O “pipe.” From a storage perspective, this approach results in a sharp increase in random IOPS. For spinning-media disks, random I/O operations are the most difficult to handle, requiring costly seek operations and rotations between microsecond transfers, and they also constitute the critical performance components in the server environment. Therefore, it is important that data centers bundle the performance of these components through intelligent technology so that they do not cause a system bottleneck as well as compensate for any failure of an individual component. RAID technology offers a solution by arranging several hard disks in an array so that any disk failure can be accommodated.

Data-center I/O workloads are either random (many concurrent accesses to relatively small blocks of data) or sequential (a modest number of large sequential data transfers). Currently, data centers are dominated by random and sequential workloads resulting from the scale-out architecture requirements in the data center. Historically, random access has been associated with transactional workloads, which are the most common workload types for an enterprise.

NVMe storage solutions

NVMe storage solutions on Cisco rack platforms offer the following main benefits:

- Strategic partnerships: Cisco tests a broad set of NVMe storage technologies and focuses on major vendors. With each partnership, devices are built exclusively in conjunction with Cisco engineering, so customers have the flexibility of a variety of endurance and capacity levels and the most relevant form factors, as well as the powerful management features and robust quality benefits that are unique to Cisco.
- Reduced TCO: NVMe storage can be used to eliminate the need for SANs and Network-Attached Storage (NAS) or to augment existing shared-array infrastructure. With significant performance improvements available in both cases, Cisco customers can reduce the amount of physical infrastructure they need to deploy, increase the number of virtual machines they can place on a single physical server, and improve overall system efficiency. These improvements provide savings in Capital Expenditures (CapEx) and Operating Expenses (OpEx), including reduced application licensing fees and savings related to space, cooling, and energy use.

I/O challenges

The rise of technologies such as virtualization, cloud computing, and data consolidation poses new challenges for the data center and requires enhanced I/O requests. These enhanced requests lead to increased I/O performance requirements. The following two major factors are leading to an I/O crisis:

- Increasing CPU use and I/O operations: multicore processors with virtualized server and desktop architectures increase processor use, thereby increasing the I/O demand per server. In a virtualized data center, it is the I/O performance that limits the server consolidation ratio, not the CPU or memory size.
- Randomization: virtualization has the effect of multiplexing multiple logical workloads across a single physical I/O path. The greater the degree of virtualization achieved, the more random the physical I/O requests.

Scope of this document

For the NVMe I/O performance characterization tests, performance was evaluated using Solidigm D7 -P5620 NVMe SSDs for random and sequential access patterns for Cisco UCS C220 M7N, C240 M7SX, and C240 M7SN servers.

- Cisco UCS C220 M7N is an all-NVMe server in which all 10 front NVMe SSDs are connected to PCIe Gen4 x4 lanes managed by the CPU. Drives 5, 8, 9, and 10 are managed by CPU1, and drives 1, 2, 3, 4, 6, and 7 are managed by CPU2.
- Cisco UCS C240 M7SN is all-NVMe server in which 24 front-facing NVMe SSDs are connected to PCIe Gen4 x2 lanes managed by the CPU. The UCS C240 M7SN server also supports four NVMe SSDs at the rear, connected to PCIe x4 lanes. Drives 1 to 12 are managed by CPU1, and 13 to 24 are managed by CPU2.
- The Cisco C240 M7SX server supports four front and four rear NVMe SSDs, and all drives use PCIe Gen4 x4 lanes and are managed by the CPU. Drives 1, 2, 101 and 102 are managed by CPU1, and drives 3, 4, 103 and 104 are managed by CPU2.

Solution components

The solution tested used these components:

- Cisco UCS C220 M7N with 10 x 6.4 TB NVMe SSDs
- Cisco UCS C240 M7SN with 24 x 6.4 TB NVMe SSDs
- Cisco UCS C240 M7SX with 8 x 6.4 TB NVMe SSDs

Cisco UCS C220 M7 server models

The Cisco UCS C220 M7 Rack Server brings many new innovations to the UCS rack server portfolio. With the introduction of PCIe Gen 5.0 for high-speed I/O, a DDR5 memory bus, and expanded storage capabilities, the server delivers significant performance and efficiency gains that will improve your application performance.

- Supports up to two 4th Gen Intel Xeon Scalable CPUs, with up to 52 cores per socket
- Up to 32 DDR5 DIMMs for up to 4 TB of capacity using 128 GB DIMMs (16 DIMMs per socket)
- 4800 MT/s DDR5 memory plus other speeds depending on the CPU installed
- Up to 3 PCIe Gen 4.0 slots or up to 2 PCIe Gen 5.0 slots, plus a modular LAN on motherboard (mLOM) slot
- Support for Cisco UCS VIC 15000 Series adapters as well as third-party options
- Up to 10 SAS/SATA or NVMe disk drives
 - New tri-mode RAID controller supports SAS4 RAID or NVMe hardware RAID with, optionally, up to four direct-attach NVMe drives
 - Option for 10 direct-attach NVMe drives at PCIe Gen4x4 each
- M.2 boot options
 - Up to two 960GB SATA M.2 drives with hardware RAID, or
 - Up to two 960GB NVMe M.2 drives with NVMe hardware RAID
- Up to three GPUs supported
- Hybrid modular LOM/OCP 3.0
 - One dedicated PCIe Gen4 x16 slot that can be used to add an mLOM or OCP 3.0 card for additional rear-panel connectivity
 - mLOM allows for Cisco UCS Virtual Interface Cards (VICs) without consuming a PCIe slot, supporting quad port 10/25/50 Gbps or dual port 40/100/200 Gbps network connectivity
 - OCP 3.0 slot features full out-of-band management for selected adapters.

See the following data sheet (which has links to specification sheets and installation guides) for more information about the C220 M7 server.

<https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/ucs-c220-m7-rack-server-ds.html>.



Figure 1.
Front and rear view of Cisco UCS C220 M7 Rack Server

Cisco UCS C240 M7 server models

The Cisco UCS C240 M7 Rack Server brings many new innovations to the UCS rack server portfolio. With the introduction of PCIe Gen 5.0 expansion slots for high-speed I/O, a DDR5 memory bus, and expanded storage capabilities, the server delivers significant performance and efficiency gains that will improve your application performance. Its features including the following:

- Supports up to two 4th Gen Intel Xeon Scalable CPUs, with up to 60 cores per socket
- Up to 32 DDR5 DIMMs for up to 8 TB of capacity using 256 GB1 DIMMs (16 DIMMs per socket)
- 4800 MT/s DDR5 memory plus other speeds depending on the CPU installed
- Up to 8 PCIe 4.0 slots or up to 4 PCIe 5.0 slots, plus a hybrid modular LAN on motherboard (mLOM) /OCP 3.0 slot
- Support for Cisco UCS VIC 15000 Series adapters as well as third-party options
- Up to 28 hot-swappable Small-Form-Factor (SFF) SAS/SATA or NVMe drives (with up to 8 direct-attach NVMe drives)
 - New tri-mode RAID controller supports SAS4 plus NVMe hardware RAID
 - Option for 28 NVMe drives at PCIe Gen4 x2 for front 24 drives and Gen4 x4 for rear 4 drives
- M.2 boot options
 - Up to two 960GB SATA M.2 drives with hardware RAID, or
 - Up to two 960GB NVMe M.2 drives with NVMe hardware RAID
- Up to five GPUs supported
- Modular LOM/OCP 3.0
 - One dedicated PCIe Gen4x16 slot that can be used to add an mLOM or OCP 3.0 card for additional rear-panel connectivity.
 - mLOM slot that can be used to install a Cisco UCS Virtual Interface Card (VIC) without consuming a PCIe slot, supporting quad port 10/25/50 Gbps or dual port 40/100/200 Gbps network connectivity
 - OCP 3.0 slot features full out-of-band management for selected adapters.

See the following data sheet (which has links to specification sheets and installation guides) for more information about the C240 M7 server:

<https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/ucs-c240-m7-rack-server-ds.html>.

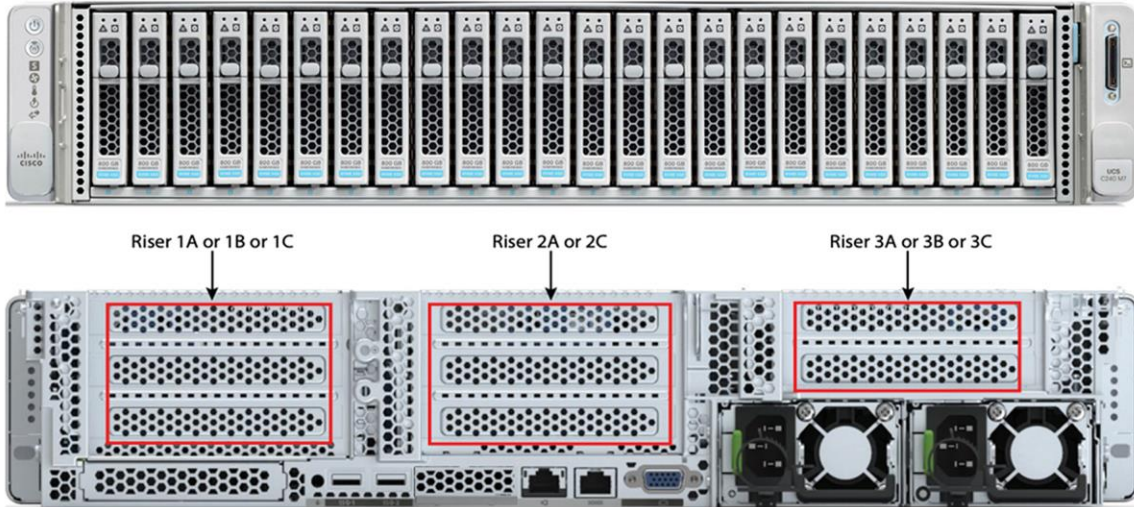


Figure 2.
Front and rear view of Cisco UCS C240 M7 Rack Server

Workload characterization

Table 1 provides an overview of the specific access patterns used for industry-standard workloads.

Table 1. Workload types

Workload type	RAID type	Access pattern type	Read: write (%)
Online transaction processing (OLTP)	5	Random	70:30
Decision-support system (DSS), business intelligence, and video on demand (VoD)	5	Sequential	100:0
Database logging	10	Sequential	0:100
High-performance computing (HPC)	5	Random and sequential	50:50
Digital video surveillance	10	Sequential	10:90
Big data: Hadoop	0	Sequential	90:10
Apache Cassandra	0	Sequential	60:40
Virtual desktop infrastructure (VDI): boot process	5	Random	80:20
VDI: steady state	5	Random	20:80

Tables 2 and 3 list the I/O mix ratios chosen for sequential-access and random-access patterns, respectively.

Table 2. I/O mix ratio for sequential-access pattern

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

Table 3. 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 servers. The performance results for RAID with NVMe drives will be available in future.

Test configuration

The test configuration was as follows:

- Ten NVMe SFF SSDs on the Cisco UCS C220 M7N server (all front-facing NVMe PCIe SSDs)
- Eight NVMe SFF SSDs on the Cisco UCS C240 M7SX server (four front-facing and four rear-facing SFF NVMe PCIe SSDs)
- Twenty-four NVMe SFF SSDs on the C240 M7SN server (24 front-facing SFF NVMe PCIe SSDs, which are x2 PCIe lane)
- 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 4 lists the recommended Fio settings.

Table 4. 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 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 best possible response time (latency) measured in microseconds.

NVMe drives on these servers are directly managed by CPU using PCIe lanes. For Cisco UCS C220 M7N (10 drives) and Cisco UCS 240 M7SX (8 drives) the PCIe lanes are “Gen4 x4” lanes, and for Cisco UCS C240 M7SN (24 front drives) the PCIe lanes are “Gen4 x2” lanes.

This x2 lane design of the C240 M7SN server meets the expected performance of drives (based on the drive specifications), and most of the supported drives meet this expectations. For those high-end NVMe drives, whose drive specifications exceed the PCIe Gen4 x2 lane maximum, performance will be capped to the PCIe lane limits.

Also, note that these NVMe drives are directly connected to the CPU, and hence heavy -I/O operations can consume significant CPU cycles, which may need to be balanced with application demands for CPU cycles.

NVMe performance on Cisco UCS C220 M7N with 10-disk configuration

Figure 3 shows the performance with 10 NVMe SSDs (6.4 TB D7 -Intel P5620 NVMe high-performance high-endurance drive) in the front slots. The graph shows the comparative performance achieved for NVMe SSDs to help customers understand the performance. The graph shows that the 6.4 TB high-performance high-endurance NVMe drive provides performance (for 4-KB block size) of 11 million IOPS for 100-percent random read operations, 4 million IOPS for 100-percent random write operations, and 5.2 million IOPS for 70:30-percent random read:write operations. This is achieved with a latency (latency is the time taken to complete a single I/O request) of less than 260 milliseconds.

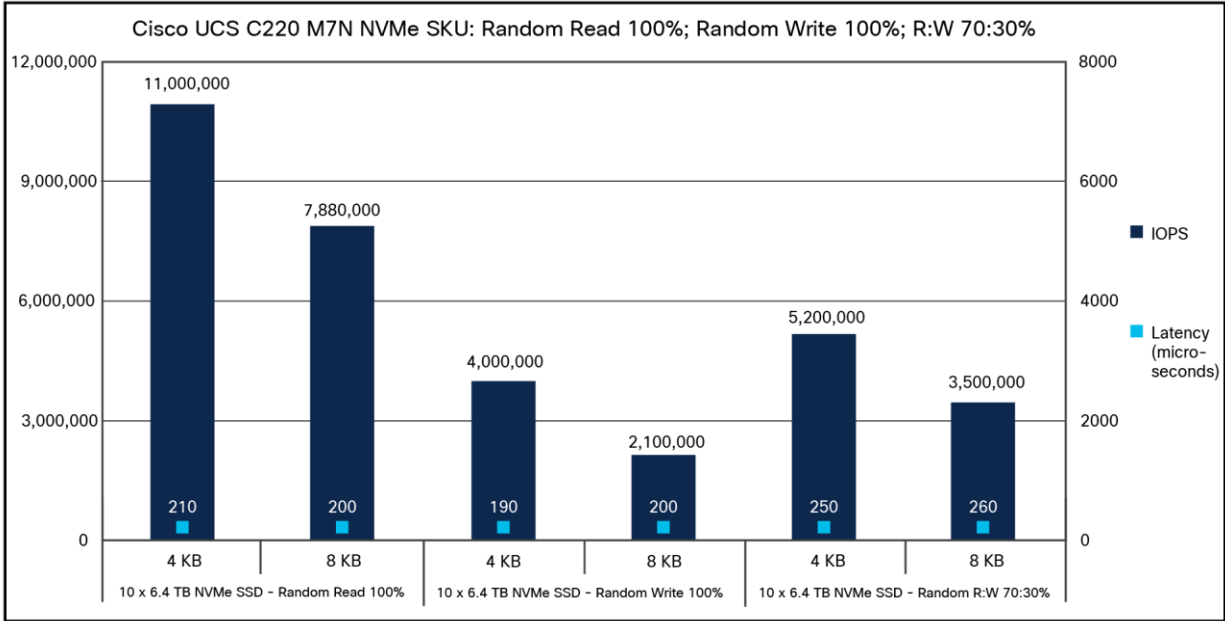


Figure 3. Cisco UCS C220 M7N 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 4 shows the performance with 10 NVMe SSDs for a 100-percent sequential read and sequential write access pattern. The graph shows the comparative performance values achieved for NVMe SSDs to help customers understand the performance. The graph shows that the 6.4 TB high-performance, high-endurance NVMe drive provides performance of 71,000 MBps with latency of 710 microseconds for 100-percent sequential read operations and 42,000 MBps with latency of 260 microseconds for 100-percent sequential write operations with a 256-KB block size. Performance is similar for a 1-MB block size with a slight increase in latency.

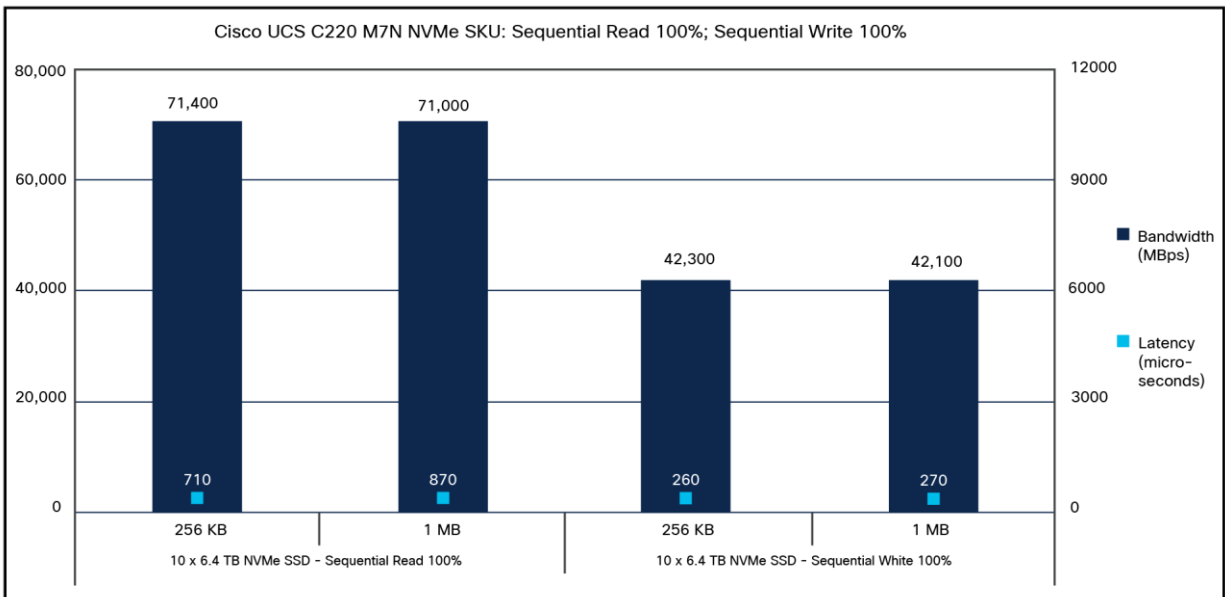


Figure 4. Cisco UCS C220 M7N 100-percent sequential read; 100-percent sequential write

NVMe performance on Cisco UCS C240 M7SX with 8-disk configuration

Figure 5 shows the performance with eight NVMe SSDs (6.4 TB Intel D7 -P5620 NVMe high-performance, high-endurance drive), with four in the front slots and four in the rear slots, for 100-percent random read and write operations with a 70:30-percent random read and write access pattern. The graph shows that eight of 6.4 TB NVMe drives provide performance (for 4-KB block size) of 8.8 million IOPS for 100-percent random read operations, performance of 3.2 million IOPS for 100-percent random write operations, and performance of 4.2 million IOPS for 70:30-percent random read and write operations. This is achieved with a latency (latency is the time taken to complete a single I/O request) of less than 260 milliseconds.

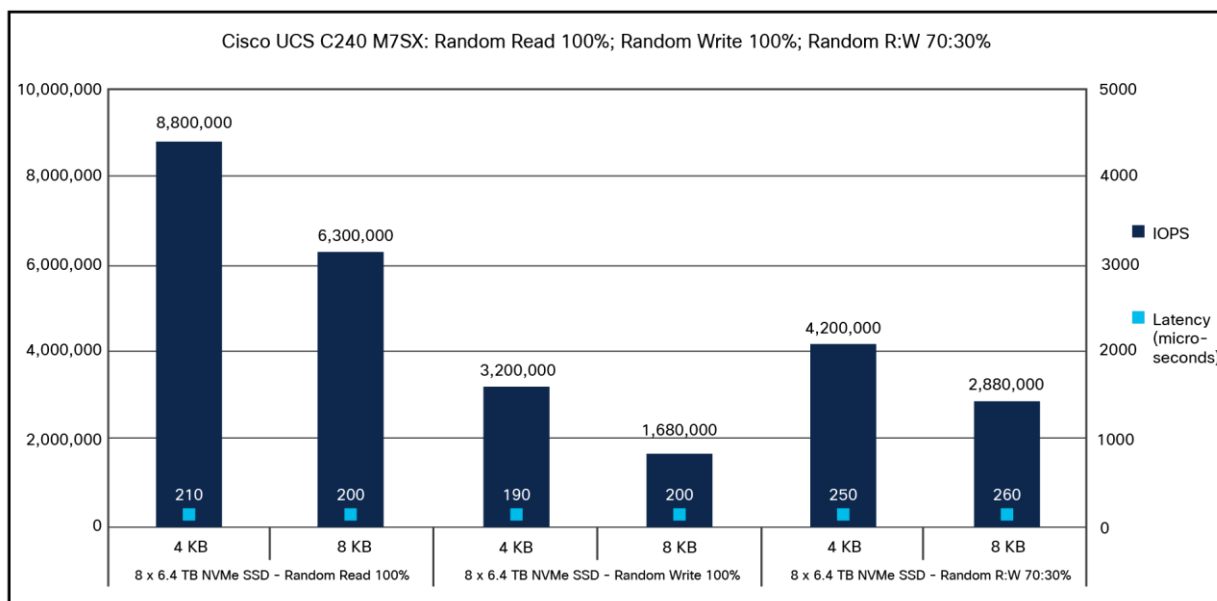


Figure 5. Cisco UCS C240 M7SX 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 6 shows the performance with eight NVMe SSDs for a 100-percent sequential read and write access pattern. The graph shows that the 6.4 TB high-performance, high-endurance NVMe drive provides performance of 57,000 MBps with latency of 710 microseconds for 100-percent sequential read operations and 34,000 MBps with latency of 260 microseconds for 100-percent sequential write operations with a 256-KB block size. Performance is similar for a 1-MB block size.

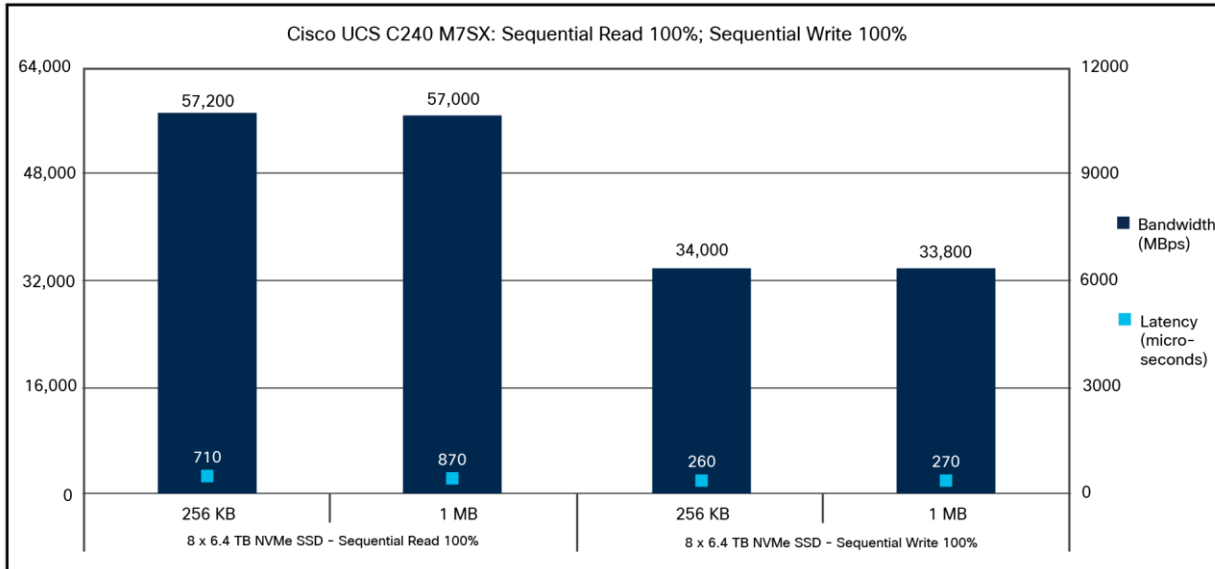


Figure 6. Cisco UCS C240 M7SX 100-percent sequential read; 100-percent sequential write

Note: This Cisco UCS C240 M7SX server has an additional 20 slots in the front panel that can accommodate SAS/SATA SFF SSDs (managed by a storage adapter), and overall server performance will be the sum of the performance of 20 SAS/SATA SSDs and 8 NVMe SSDs.

NVMe performance on Cisco UCS C240 M7SN with 24-disk configuration

Figure 9 shows the performance with 24 NVMe SSDs (6.4 TB Intel D7 -P5620 NVMe high-performance, high-endurance drive), populated in the front 24 slots for 100-percent random read and write operations with a 70:30-percent random read and write access pattern. These 24 NVMe drives are connected to Gen4 x2 PCIe lane direct-attach CPU-managed slots. The graph shows that all 24 of the 6.4 TB NVMe drives provide performance of 20 million IOPS with latency of 210 microseconds for 100-percent random read operations, 9.3 million IOPS with latency of 200 microseconds for 100-percent random write operations, and 12.6 million IOPS with latency of 250 microseconds for a 70:30-percent random read:write pattern with a 4-KB block size.

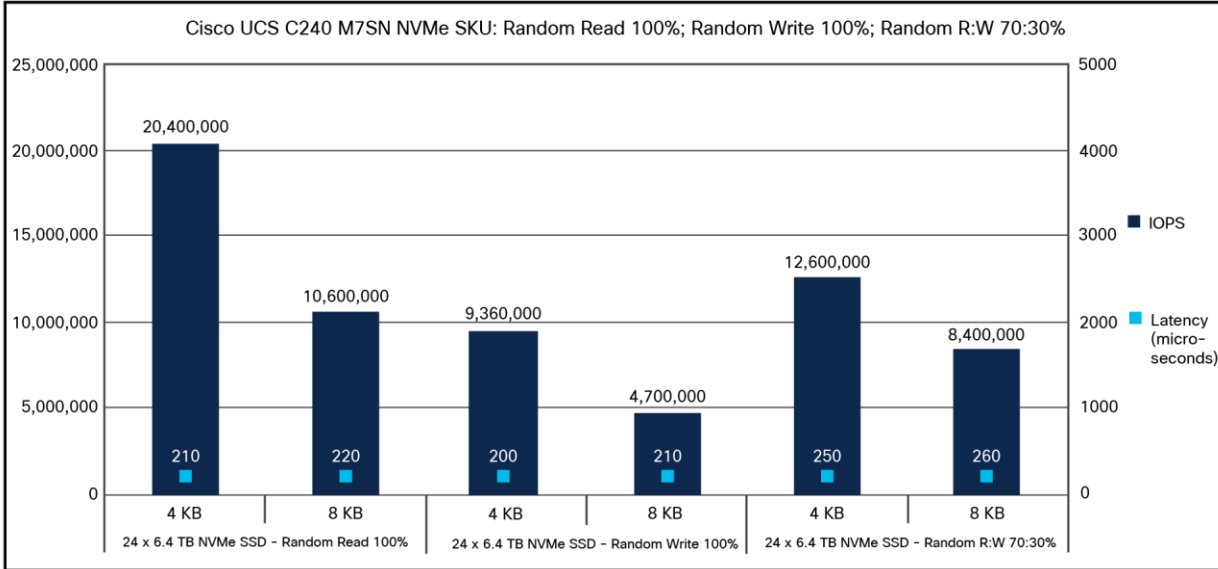


Figure 7. Cisco UCS C240 M7SN 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 8 shows the performance of 24 NVMe SSDs for a 100-percent sequential read and write access pattern. The graph shows that the 6.4 TB high-performance, high-endurance NVMe drive provides performance of 84,000 MBps with latency of 850 microseconds for 100-percent sequential read operations and 84,000 MBps with latency of 380 microseconds for 100-percent sequential write operations with a 256-KB block size. The performance is similar for a 1-MB block size with higher latency.

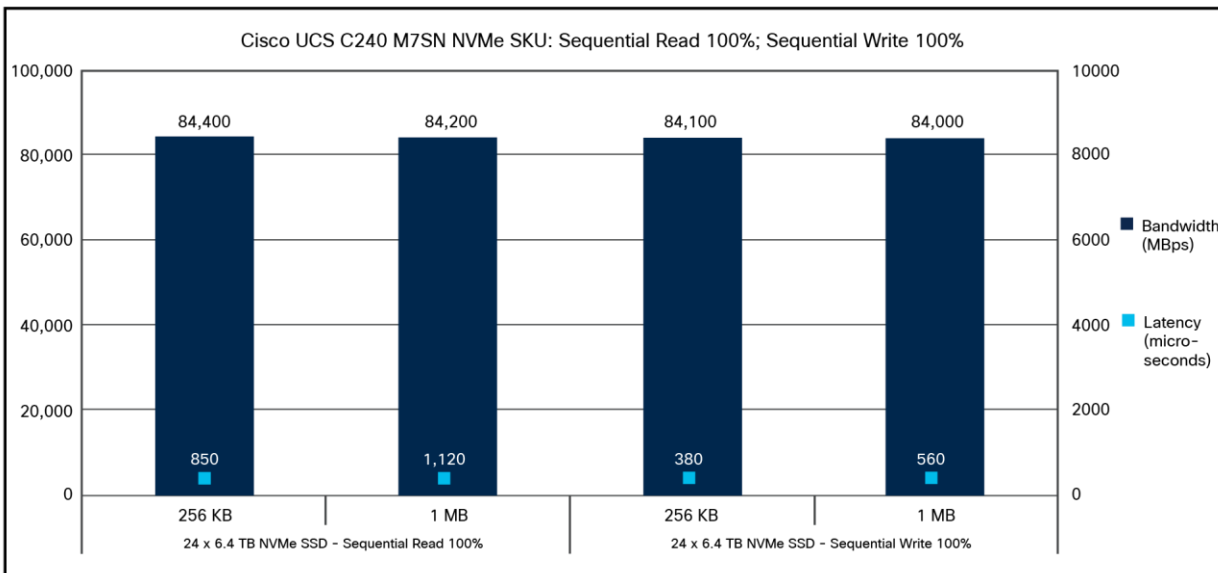


Figure 8. Cisco UCS C240 M7SN 100-percent sequential read; 100-percent sequential write

Note: The performance observed from 6.4 TB Solidigm D7 -P5620 NVMe on the Cisco UCS C240 M7SN SKU is different compared to other server SKUs mentioned in this document. This is due to the C240 M7SN PCIe lane design and the drive specifications exceeding the PCIe Gen4 x2 lane maximum.

Performance summary

The Cisco UCS C240 M7SX server is ideal for applications that demand large capacity with tiered storage benefits, where front and rear NVMe drives can be configured as low latency tier and storage -adapter- managed SSD drives for capacity tier. The overall storage performance will be a combination of NVMe drives and storage-adapter configured SSD drives maximums.

The Cisco UCS C240 M7SN server is an All NVMe server which is a good fit for applications that need higher NVMe storage with lower latency. Since all front 24 NVMe slots are Gen4 x2 PCIe lanes, performance will be as per the PCIe lane maximums.

The Cisco UCS C220 M7N 1RU server gives the best performance with high-performing NVMe SSDs on all slots because of Gen4 x4 PCIe lane design thus giving the best storage performance combining of high performance, low latency with decent number of drive slots and capacity.

For more information

- <https://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/c220m7-sff-specsheet.pdf>.
- <https://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/c240m7-sff-specsheet.pdf>.

Appendix: Test environment

Table 5 lists the details of the server under test.

Table 5. Server properties

Name	Value
Product names	Cisco UCS C220 M7 and C240 M7 rack servers
CPUs	CPU: two 1.9-GHz Intel Xeon Gold 8490H processors
Number of cores	60
Number of threads	120
Total memory	512 GB
Memory DIMMs (16)	32 GB DDR5 x 16 DIMMs
Memory speed	4800 MHz
Virtual interface card (VIC) adapter	Cisco UCS VIC 15238 modular LOM (mLOM) 40/100/200 Gbps dual-port small form-factor pluggable
SFF NVMe SSDs	6.4-TB 2.5-inch Intel P5620 NVMe high-performance, high-endurance (UCS-NVME4-6400) (6.4-TB drive specification)

Table 6 lists the recommended server BIOS settings for a standalone Cisco UCS C-Series rack server for NVMe performance tests.

Table 6. BIOS settings for standalone rack server

Name	Value
BIOS version	Release 4.3.1
Cisco Integrated Management Controller (IMC) version	Release 4.3(1)
Cores enabled	All
Hyper-threading (All)	Enable
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	1x refresh
Energy-efficient turbo	Enable
Turbo mode	Enable
Energy performance preference (EPP) profile	Performance
CPU C6 report	Enable
Package C-state	C0/C1 state
Power performance tuning	OS controls EPB
Workload configuration	I/O sensitive

Note: The rest of the BIOS settings were kept at platform default values. These settings are used for the purposes of testing and may require additional tuning for application workloads. For additional information, refer to the [BIOS tuning guide](#).

Note: The RHEL OS settings are configured on the Cisco UCS C240 M7SN SKU to get maximum performance on all twenty-four NVMe SSDs when Fio tests are run in parallel on all drives. An **'nvme set-feature /dev/nvme0 -f=8 -value=0x108'** value was set on all available NVMe drives for our internal testing.

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)