

Cisco UCS C240 M7 Rack Server Disk I/O Characterization

April 2024

Contents

Executive summary	3
Introduction	4
I/O challenges	4
Scope of this document	4
Cisco UCS C240 M7 server models	5
Workload characterization	9
Test configuration	11
SSD performance results	12
NVMe performance data with a tri-mode storage controller	23
Performance summary	26
For more information	27
Appendix: Test environment	27

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 Unified Computing System™ (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 outlines the I/O performance characteristics of the Cisco UCS C240 M7 Rack Server using the new Cisco 24G Tri-Mode SAS RAID (Redundant Array of Independent Disks) Controller with cache backup to control SAS/SATA/NVMe drives, Cisco 12G RAID controller with cache backup to control Serial Attached SCSI/Serial Advanced Technology Attachment (SAS/SATA) drives, and the Cisco 12G SAS pass-through Host Bust Adapters (HBAs) to control SAS/SATA drives. Performance comparisons of solid-state disks (SSDs) with RAID and HBA controllers are presented. The goal of this document is to help customers make well-informed decisions in choosing the right internal disk types and configuring the right controller options to meet their individual I/O workload needs.

Performance data was obtained using the FIO (Flexible I/O Tester) 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.

Many combinations of drive types and RAID levels are possible. For these characterization tests, performance evaluations were limited to small-form-factor (SFF) SSDs with RAID 0, 5, and 10 virtual drives and JBOD (Just a Bunch of Disks) configurations.

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.

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. They also require that data centers fully use available resources so that they can support the newest requirements of the data center and reduce the performance gap observed industrywide.

The following are the major factors leading to an I/O crisis:

- **Increasing CPU usage and I/O:** Multicore processors with virtualized server and desktop architectures increase processor usage, 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.
- **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 I/O performance characterization tests, performance was evaluated using SSDs with configurations of RAID 0, 5, 10 virtual drives and JBOD config to achieve maximum performance for NVME, SAS, and SATA SSDs for the Cisco UCS C240 M7 Rack server. The Cisco UCS C240 M7SX server used for the I/O performance characterization tests supports up to 28 SSDs configured on dual storage controllers (14 drives on each controller). The performance tests described here were performed with a 28-disk configuration for SFF NVME/SAS/SATA SSDs with dual Cisco 24G Tri-Mode SAS RAID Controllers and JBOD 28-disk configuration with dual Cisco 12G SAS pass-through HBA controllers.

Cisco UCS C240 M7 server models

Overview

- The Cisco UCS C240 M7 SFF server extends the capabilities of the Cisco Unified Computing System portfolio in a 2U form factor with the addition of the 4th Gen Intel Xeon Scalable Processors (code-named Sapphire Rapids), 16 DIMM slots per CPU for DDR5-4800 DIMMs with DIMM capacity points up to 256GB.
- The Cisco UCS C240 M7 SFF server harnesses the power of the latest 4th Gen Intel Xeon Scalable Processors (codenamed Sapphire Rapids), and offers the following:
 - **CPU:** Up to 2x 4th Gen Intel Xeon Scalable Processors (codenamed Sapphire Rapids) with up to 60 cores per processor.
 - **Memory:** Up to 8TB with 32 x 256GB DDR5-4800 DIMMs, in a 2-socket configuration

There are several options to choose from:

Option 1 UCSC-C240-M7SX

- Up to 24 front facing SFF SAS/SATA HDDs or SAS/SATA SSDs or NVMe (Non-Volatile Memory Express) SSDs (optionally up to 4 of the slots can be direct-attach NVMe)
- I/O-centric option provides up to 8 PCIe slots using all three rear risers
- Storage-centric option provides three rear risers with a total of up to 4 SFF drives slot (SAS/SATA or NVMe PCIe Gen4) and 3 PCIe slots

Option 2:

- Up to 24 front-facing SFF NVMe SSDs (drives are direct-attach to PCIe Gen4 x2)
- I/O-centric option provides up to 8 PCIe slots using all three rear risers
- Storage-centric option provides 3 rear risers with a total of up to 4 SFF drives slot (NVMe PCIe Gen4) and 3 PCIe slots

Storage controllers:

The following list summarizes how drives are controlled on the server:

- Up to 28 SAS/SATA drives are controlled through a Cisco 12G SAS RAID controller
- Up to 14 SAS/SATA/NVMe drives are controlled through a Cisco 24G tri-mode RAID controller
- Up to 14 SAS/SATA drives are controlled through a Cisco 12G SAS pass-through HBA

Note: Performance testing data published in this document is done using dual Cisco 24G tri-mode RAID controller and dual Cisco 12G SAS pass-through HBA.

The Cisco UCS C240 M7 SFF server can be used standalone, or as part of the Cisco Unified Computing System, which unifies computing, networking, management, virtualization, and storage access into a single integrated architecture enabling end-to-end server visibility, management, and control in both bare metal and virtualized environments.

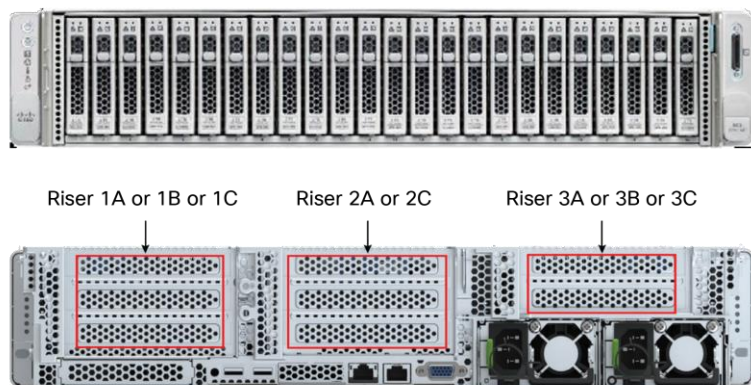


Figure 1.
Front view of server

A specifications sheet for the C240 M7 is available at:

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

SSDs and NVMe drives used

The SSDs selected for the tests were based on the following factors:

- 1.9 TB enterprise value 12G SAS SSD
- 960 GB enterprise value 6G SATA SSD
- 3.2 TB U.3 high performance high endurance NVMe SSD

Performance

When considering price, it is important to differentiate between enterprise performance SSDs and enterprise value SSDs. The differences between the two in performance, cost, reliability, and targeted applications are significant. Although it can be appealing to integrate SSDs with NAND flash technology in an enterprise storage solution to improve performance, the cost of doing so on a large scale may be prohibitive.

Virtual disk options

You can configure the following controller options with virtual disks to accelerate write and read performance and provide data integrity:

- RAID level
- Strip (block) size
- Access policy
- Disk cache policy
- Read policy
- Write policy
- RAID levels

Table 1 summarizes the supported RAID levels and their characteristics.

Table 1. RAID levels and characteristics

RAID level	Characteristics	Parity	Redundancy
RAID 0	Striping of two or more disks to achieve optimal performance	No	No
RAID 1	Data mirroring on two disks for redundancy with slight performance improvement	No	Yes
RAID 5	Data striping with distributed parity for improved fault tolerance	Yes	Yes
RAID 6	Data striping with dual parity with dual fault tolerance	Yes	Yes
RAID 10	Data mirroring and striping for redundancy and performance improvement	No	Yes
RAID 50	Block striping with distributed parity for high fault tolerance	Yes	Yes
RAID 60	Block striping with dual parity for performance improvement	Yes	Yes

To avoid poor write performance, full initialization is always recommended when creating a RAID 5 or RAID 6 virtual drive. Depending on the virtual disk size, the full initialization process can take a long time based on the drive capacity. Fast initialization is not recommended for RAID 5 and RAID 6 virtual disks.

Stripe size

Stripe size specifies the length of the data segments that the controller writes across multiple drives, not including parity drives. You can configure stripe size as 64, 128, 256, or 512 KB, or 1 MB. The default stripe size is 64 KB. The performance characterization tests in this paper used a stripe size of 64 KB. With random I/O workloads, no significant difference in I/O performance was observed by varying the stripe size. Stripe size of 64 KB was used for all the random and sequential workloads in the tests.

Strip (block) size versus stripe size

A virtual disk consists of two or more physical drives that are configured together through a RAID controller to appear as a single logical drive. To improve overall performance, RAID controllers break data into discrete chunks called strips that are distributed one after the other across the physical drives in a virtual disk. A stripe is the collection of one set of strips across the physical drives in a virtual disk.

Access policy

You can set access policy as follows:

- **RW:** Read and write access is permitted.
- **Read Only:** Read access is permitted, but write access is denied.
- **Blocked:** No access is permitted.

Disk cache policy

You can set disk cache policy as follows:

- **Disabled:** The disk cache is disabled. The drive sends a data transfer completion signal to the controller when the disk media has actually received all the data in a transaction. This process helps ensure data integrity if a power failure occurs.
- **Enabled:** The disk cache is enabled. The drive sends a data transfer completion signal to the controller when the drive cache has received all the data in a transaction. However, the data has not actually been transferred to the disk media, so data may be permanently lost if a power failure occurs. Although disk caching can accelerate I/O performance, it is not recommended for enterprise deployments.

I/O cache policy

You can set I/O cache policy as follows:

- **Direct:** All read data is transferred directly to host memory, bypassing the RAID controller cache. Any read-ahead data is cached. All write data is transferred directly from host memory, bypassing the RAID controller cache if write-through cache mode is set. The direct policy is recommended for all configurations.

Read policy

You can set the read policy as follows:

- **No read ahead (normal read):** Only the requested data is read, and the controller does not read any data ahead.
- **Always read ahead:** The controller reads sequentially ahead of requested data and stores the additional data in cache memory, anticipating that the data will be needed soon.

Write policy

You can set the write policy as follows:

- **Write through:** Data is written directly to the disks. The controller sends a data transfer completion signal to the host when the drive subsystem has received all the data in a transaction.
- **Write back:** Data is first written to the controller cache memory, and then the acknowledgment is sent to the host. Data is written to the disks when the commit operation occurs at the controller cache. The controller sends a data transfer completion signal to the host when the controller cache has received all the data in a transaction.
- **Write back with battery backup unit (BBU):** Battery backup is used to provide data integrity protection if a power failure occurs. Battery backup is always recommended for enterprise deployments.

Workload characterization

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

Table 2 lists the workload types tested.

Table 2. 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
VDI: Boot process	5	Random	80:20
VDI: Steady state	5	Random	20:80

Tables 3 and 4 list the I/O mix ratios chosen for the sequential access and random access patterns, respectively.

Table 3. I/O mix ratio for sequential access pattern

I/O mode	I/O mix ratio (read:write)	
Sequential	100:0	0:100
	RAID 0, 10	RAID 0, 10
	JBOD*	JBOD*

*JBOD is configured using HBA

Table 4. I/O mix ratio for random access pattern

I/O mode	I/O mix ratio (read:write)			
Random	100:0	0:100	70:30	50:50
	RAID 0	RAID 0	RAID 0, 5	RAID 5
	JBOD*	JBOD*	JBOD*	NA

*JBOD is configured using HBA

Tables 5 and 6 show the recommended virtual drive configuration for SSDs and HDDs, respectively.

Table 5. Recommended virtual drive configuration for SSDs

Access pattern	RAID level	Strip size	Disk cache policy	Read policy	Write policy
Random I/O	RAID 0	64 KB	Disabled	No read ahead	Write back good BBU
Random I/O	RAID 5	64 KB	Disabled	No read ahead	Write back good BBU
Sequential I/O	RAID 0	64 KB	Disabled	No read ahead	Write back good BBU
Sequential I/O	RAID 10	64 KB	Disabled	No read ahead	Write back good BBU

Table 6. Recommended virtual drive configuration for NVMe SSDs

Access pattern	RAID level	Strip size	Disk cache policy	Read policy	Write policy
Random I/O	RAID 0	64 KB	Disabled	No read ahead	Write through
Random I/O	RAID 5	64 KB	Disabled	No read ahead	Write through
Sequential I/O	RAID 0	64 KB	Disabled	No read ahead	Write through
Sequential I/O	RAID 10	64 KB	Disabled	No read ahead	Write through

Table 7. Recommended virtual drive configuration for HDDs

Access pattern	RAID level	Strip size	Disk cache policy	Read policy	Write policy
Random I/O	RAID 0	1 MB	Disabled	Always read ahead	Write back good BBU
Random I/O	RAID 5	1 MB	Disabled	Always read ahead	Write back good BBU
Sequential I/O	RAID 0	1 MB	Disabled	Always read ahead	Write back good BBU
Sequential I/O	RAID 10	1 MB	Disabled	Always read ahead	Write back good BBU

Test configuration

The test configuration was as follows for SAS/SATA/NVMe SSDs:

- 28 RAID 0 virtual drives were created with 28 disks for the SFF SSDs (14 virtual drives configured on each tri-mode controllers).
- 4 RAID 5 virtual drives were created with 7 disks in each virtual drive (2 virtual drives configured on each tri-mode controllers).
- 4 RAID 10 virtual drives were created with 8 disks in one virtual drive and 6 disks in second virtual drive (2 virtual drives configured on each tri-mode controllers).
- The RAID configuration was tested with the Cisco 24G tri-mode RAID controller with SuperCap and a 4 GB cache module with 28 disks for the SFF SSDs. The server is configured with dual RAID controllers, and 14 disks are configured on each controller (28 disks on two controllers).
- The JBOD configuration was tested with a Cisco 12G SAS HBA that supports up to 16 drives for SAS and SATA SSDs. The server is configured with dual HBA controllers, and 14 disks are configured on each HBA controller (28 disks on two controllers).
- Random workload tests were performed using SAS/SATA/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 SAS/SATA/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 8 lists the recommended FIO settings.

Table 8. Recommended FIO settings

Name	Value
FIO version	Fio-3.19
Filename	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:write
Block size	I/O block size: 4, 8, or 256K, or 1M
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

Name	Value
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 test

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

SSD performance results

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

The I/O performance test results capture the maximum IOPS and bandwidth achieved with the SAS and SATA SSDs within the acceptable response time (latency) of 2 milliseconds (ms).

The RAID controller (UCSC-RAID-HP) used in this testing is Broadcom 9670W-16i controller. The performance specifications for this controller is around 6+ million IOPS for Random Read I/O and 28GB for Sequential Read I/O when used with high-performing SAS/NVMe SSDs. Refer to Broadcom 9670W-16i product brief (<https://www.broadcom.com/products/storage/raid-controllers/megaraid-9670w-16i>) for more information.

During our testing, results showed that for random I/O (reads) performance with SAS drives, 5+ million IOPS was achieved with 14 SAS drives populated on single controller of the server. This was achieved as per the maximum performance of SAS SSDs that was configured on the server. However, given that the C240 M7SX server allows 28 disks populated on two controllers in the server, performance will be cumulative of two storage controllers, and random read was measured at 10+ million IOPS.

Similarly, for random I/O (reads) performance with SATA drives, 2.3M IOPS was achieved with 28 SATA drives, a fully populated disk scenario with two storage controllers. This result aligns with the IOPS specifications of the SATA drives configured in the server (refer to Figure 8 later in this document).

For sequential I/O (reads) workloads (both with SAS and SATA drives), results showed that an adapter maximum of 58,000 MBps was achieved in the case of SAS with 28 drives and 15,000 MBps was achieved in the case of SATA as per the drive maximum with 28 drives, and captured in this document (refer to Figure 3 for SAS and Figure 9 for SATA).

There is a significant improvement in performance on Cisco UCS C240 M7SX compared to earlier generations of servers and this is observed for both random I/O and sequential I/O workloads. This is mainly due to improved server architecture, storage controller design and SAS4 drives.

This dual adapter configuration is for performance, not for any fault tolerance between adapters.

Note: These conclusions are not applicable for SAS HBA adaptors (UCSC-SAS-T-D) because the C240 M7SX's design for two HBA adaptor uses x8 PCIe lanes, and thus overall server performance matches with total drives multiplied by expected single-drive performance or as per the PCIe maximum. This SAS HBA supports up to 14 SAS HDDs and SAS/SATA SSDs operating at 3Gbps, 6Gbps, and 12Gbps.

SAS SSD (1.9 TB enterprise value 12G SAS SSD) performance for 28-disk configuration

SAS SSD RAID 0 performance

Figure 2 shows the performance of the SSDs under test for RAID 0 configuration with a 100-percent random read, 100-percent random write, and 70:30-percent random read:write access pattern. The graph shows the comparative performance values achieved for enterprise value SAS SSD drives to help customers understand the performance. The graph shows that the 1.9 TB SAS drive provides performance of more than 10M IOPS with latency of 0.3 milliseconds for random read, performance of close to 2.4M IOPS with latency of 0.3 milliseconds for random write, and performance of 5.3M IOPS with latency of 0.4 milliseconds for random read:write 70:30 percent with a 4 KB block size. The result shows the maximum performance achievable from the C240 M7 RAID controller. Latency is the time taken to complete a single I/O request from the viewpoint of the application.

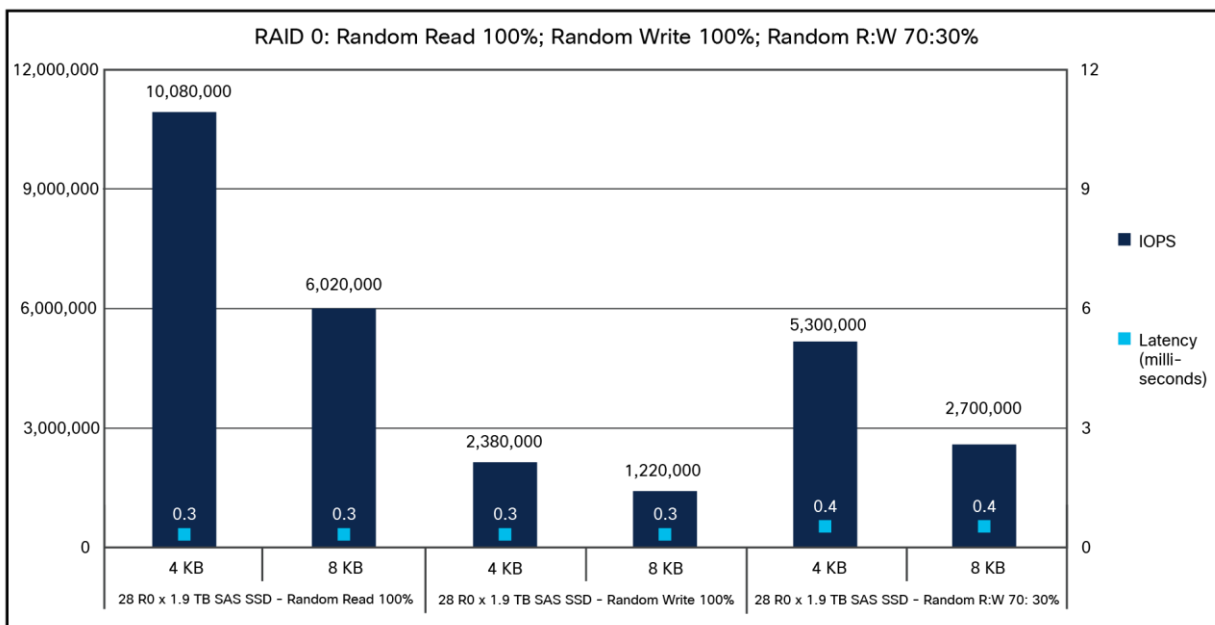


Figure 2.

Cisco UCS C240 M7SX RAID 0 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 3 shows the performance of the SSDs under test for the RAID 0 configuration with a 100-percent sequential read and write access pattern. The graph shows that the 1.9 TB enterprise value SAS SSD drive provides performance of more than 58,000 MBps for sequential read and more than 46,000 MBps for sequential write, with 256 KB and 1 MB block sizes within a latency of 1 millisecond.

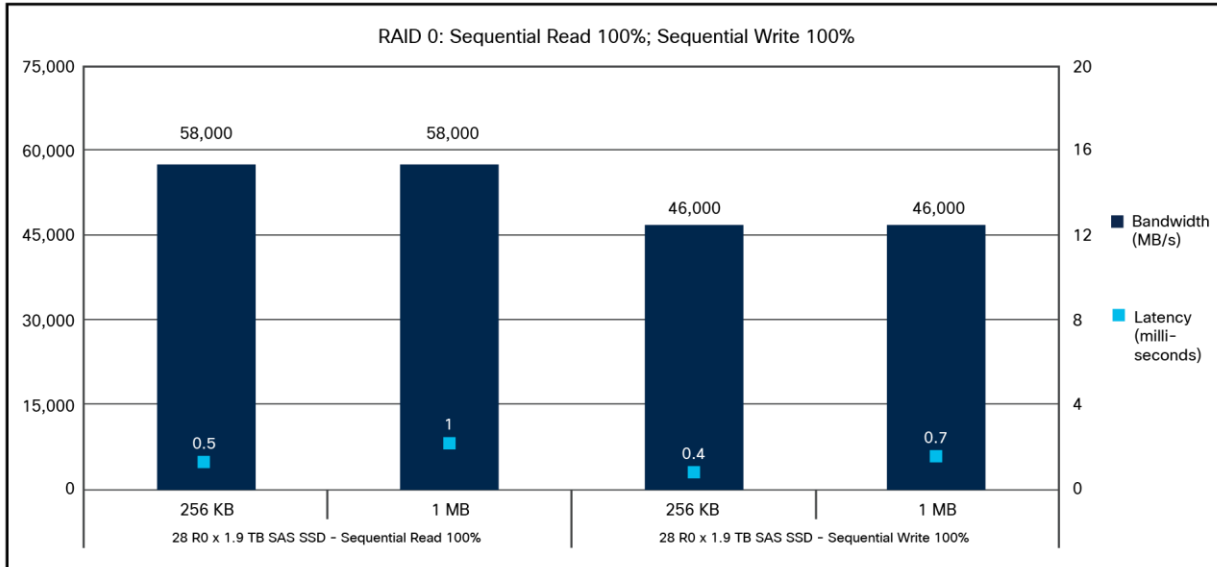


Figure 3. Cisco UCS C240 M7SX RAID 0 100-percent sequential read; 100-percent sequential write

SAS SSD RAID 5 performance

Figure 4 shows the performance of the SSDs under test for the RAID 5 configuration with a 70:30- and 50:50-percent random read:write access pattern. The graph shows the comparative performance values achieved for enterprise value SAS SSD drives to help customers understand the performance for the RAID 5 configuration. The graph shows that the 1.9 TB SAS drive provides performance of close to 2.1M IOPS with latency of 0.3 milliseconds for random read:write 70:30 percent and 2M IOPS with latency of 0.4 milliseconds for random read:write 50:50 percent with a 4 KB block size.

Each RAID 5 virtual drive was carved out using 7 disks, and a total of 4 virtual drives were created using 28 disks on dual RAID controllers (14 disks are configured on each controller). The performance data shown in the graph is cumulative data by running tests on all four virtual drives in parallel.

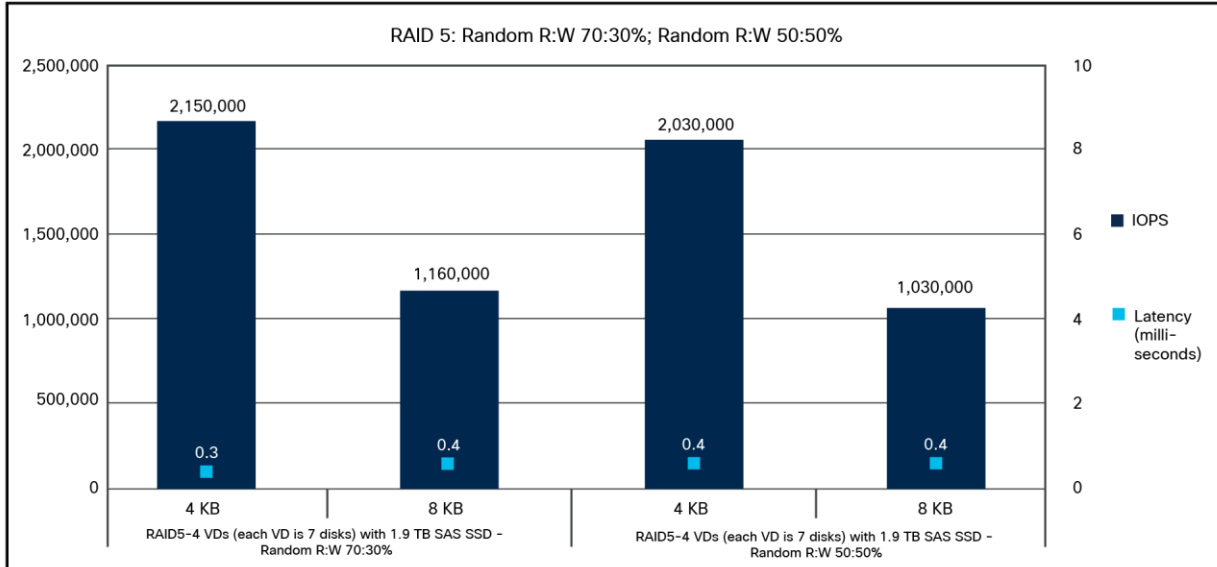


Figure 4. Cisco UCS C240 M7SX RAID 5 70:30-percent random read:write; 50:50-percent random read:write

SAS SSD RAID 10 performance

The C240 M7SX server has two RAID controllers and each controller manages 14 drives. For the purpose of this RAID 10 testing, for each RAID controller, 2 RAID VD (VD - virtual drive) were carved out, with one VD configured using 8 drives and second VD with 6 drives. RAID 10 volume requires even number of drives to be configured on each VD and to manage this requirement, we have taken this approach (from 14 drives on a single controller). The performance data shown in the graphs, is gathered from all 4 VD across both the RAID controllers.

Figure 5 shows the performance of the SSDs under test for the RAID 10 configuration with a 100-percent sequential read and write access pattern. The graph shows the comparative performance values achieved for enterprise value SAS SSD drives to help customers understand the performance for the RAID 10 configuration. The graph shows that the 1.9 TB SAS drive provides performance of close to 57,000 MBps with latency of 0.5 milliseconds for sequential read and 26,000 MBps with latency of 1 milliseconds for sequential write with 256 KB and a 1 MB block size.

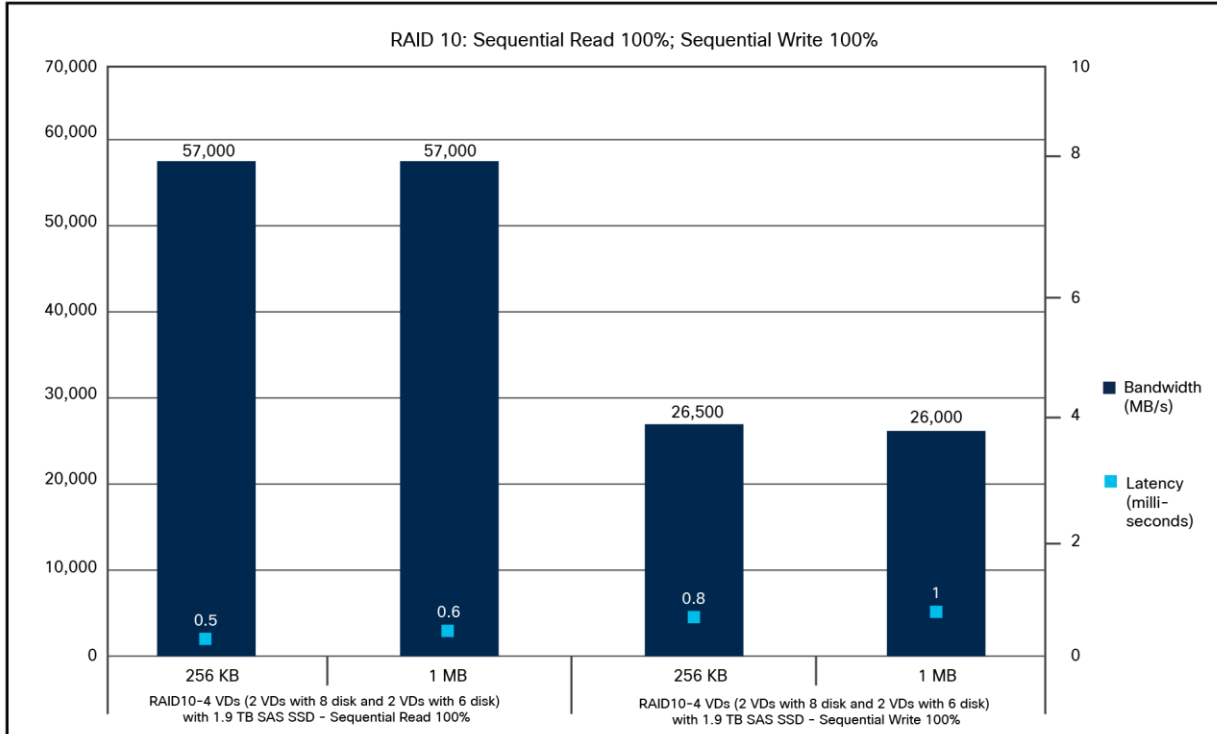


Figure 5.
Cisco UCS C240 M7SX RAID 10 100-percent sequential read; 100-percent sequential write

SAS SSD HBA performance

The C240 M7SX server supports the dual 12-Gbps modular SAS HBA adapters. Each adapter manages 14 disks on a 28-drive system. This configuration helps to scale performance from individual adapters, based on the performance specification of the drives. This dual adapter configuration is for performance, not for any fault tolerance between adapters.

Figure 6 shows the performance of the SSDs under test for the JBOD configuration with the HBA adapter for the 100-percent random read and random write and 70:30-percent random read:write access patterns. The graph shows that the 1.9 TB enterprise value SAS drive provides performance of 6.1M IOPS with latency of 0.3 milliseconds for random read, 2.1M IOPS with latency of 0.2 milliseconds for random write, and 5.1M IOPS with latency of 0.5 milliseconds for random read:write 70:30 percent with a 4 KB block size.

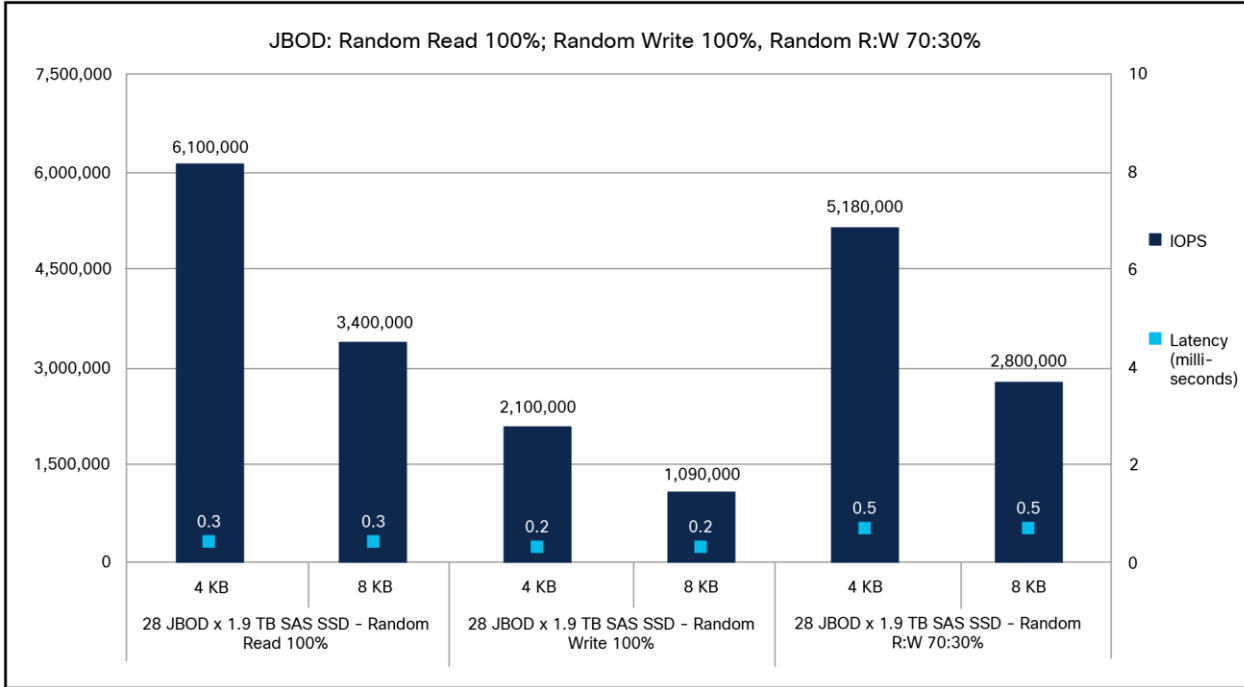


Figure 6. Cisco UCS C240 M7SX JBOD 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 7 shows the performance of the SSDs under test for the JBOD configuration with a 100-percent sequential read and write access pattern. The graph shows that the 1.9 TB enterprise value SAS SSD drive provides performance of close to 29,000 MBps for sequential read and write with 256 KB and 1 MB block sizes within latency of 1 milliseconds.

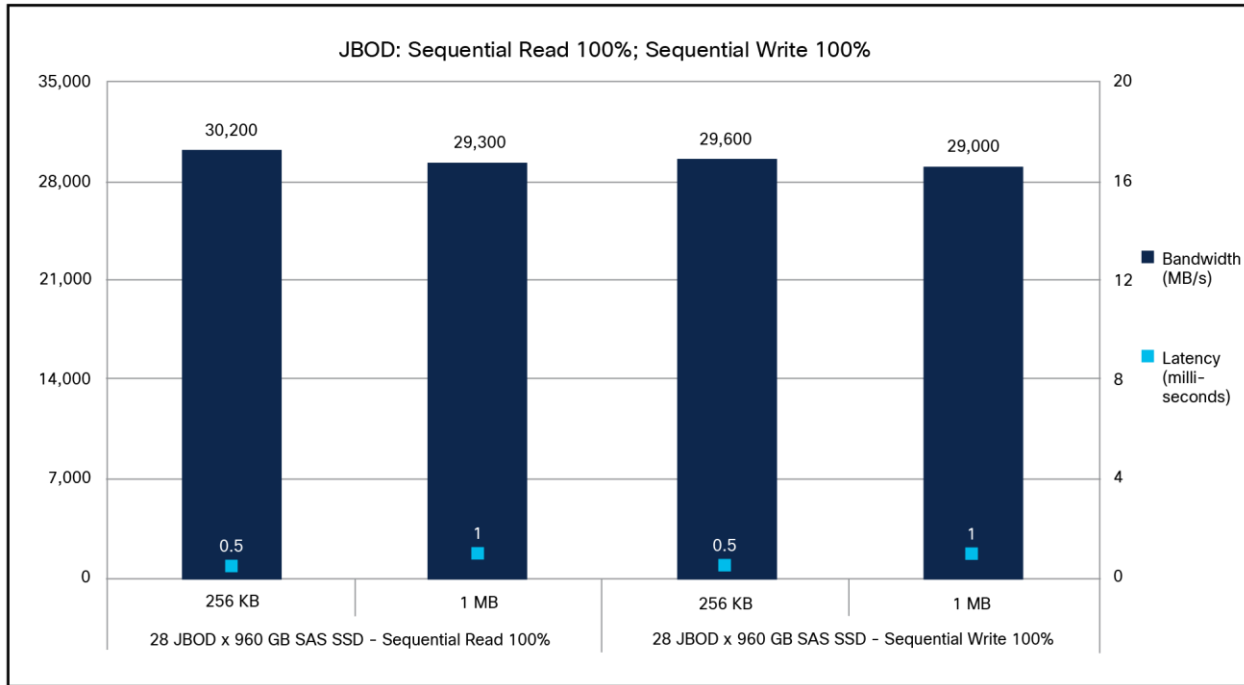


Figure 7. Cisco UCS C240 M7SX JBOD 100-percent sequential read; 100-percent sequential write

SATA SSD (960 GB enterprise value 6G SATA SSD) performance for 28-disk configuration

SATA SSD RAID 0 performance

Figure 8 shows the performance of the SSDs under test for the RAID 0 configuration with a 100-percent random read, 100-percent random write and 70:30-percent random read:write access pattern. The graph shows the comparative performance values achieved for enterprise value SATA SSD drives to help customers understand the performance for the RAID 0 configuration. The graph shows that the 960 GB enterprise value SATA drive provides performance of more than 2.3M IOPS with latency of 0.4 milliseconds for random read, performance of 880K IOPS with latency of 0.3 milliseconds for random write, and more than 2M IOPS for random read:write 70:30 percent with a 4 KB block size. Latency is the time taken to complete a single I/O request from the viewpoint of the application.

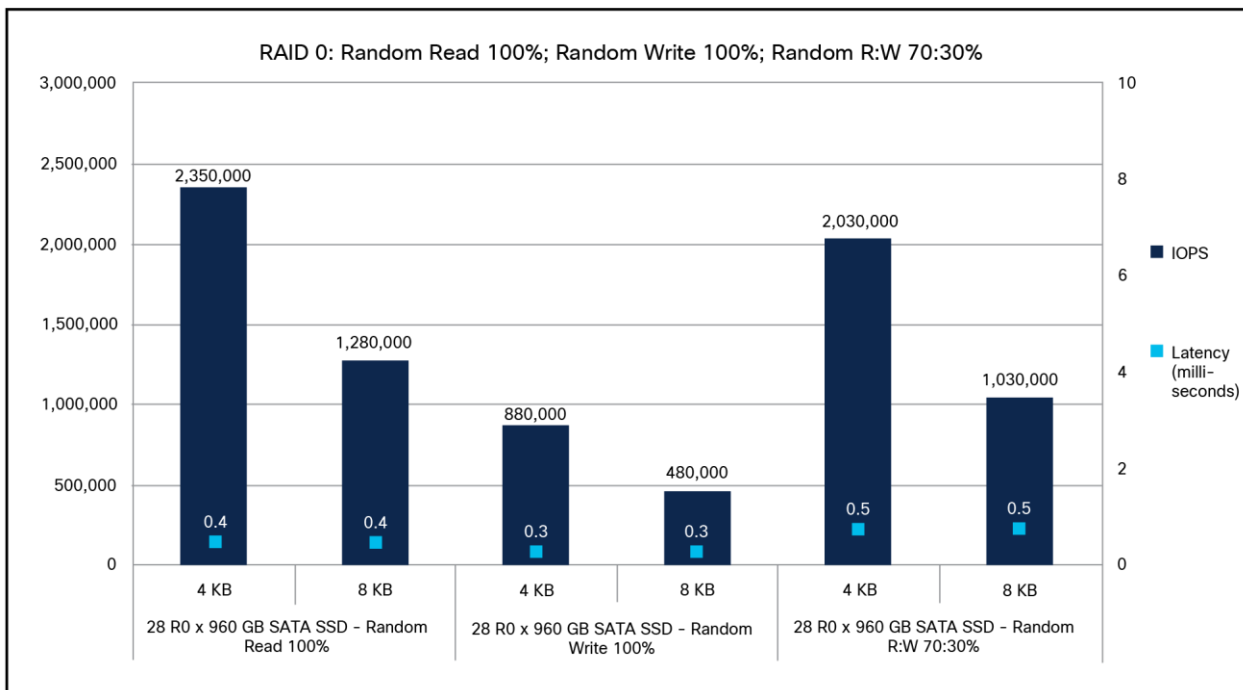


Figure 8.

Cisco UCS C240 M7SX RAID 0 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 9 shows the performance of the SSDs under test for the RAID 0 configuration with a 100-percent sequential read and 100-percent sequential write access pattern. The graph shows that the 960 GB enterprise value SATA SSD drive provides performance of more than 15,000 MBps for sequential read and more than 14,000 MBps for sequential write within latency of 2 milliseconds for 256 KB and 1 MB block sizes.

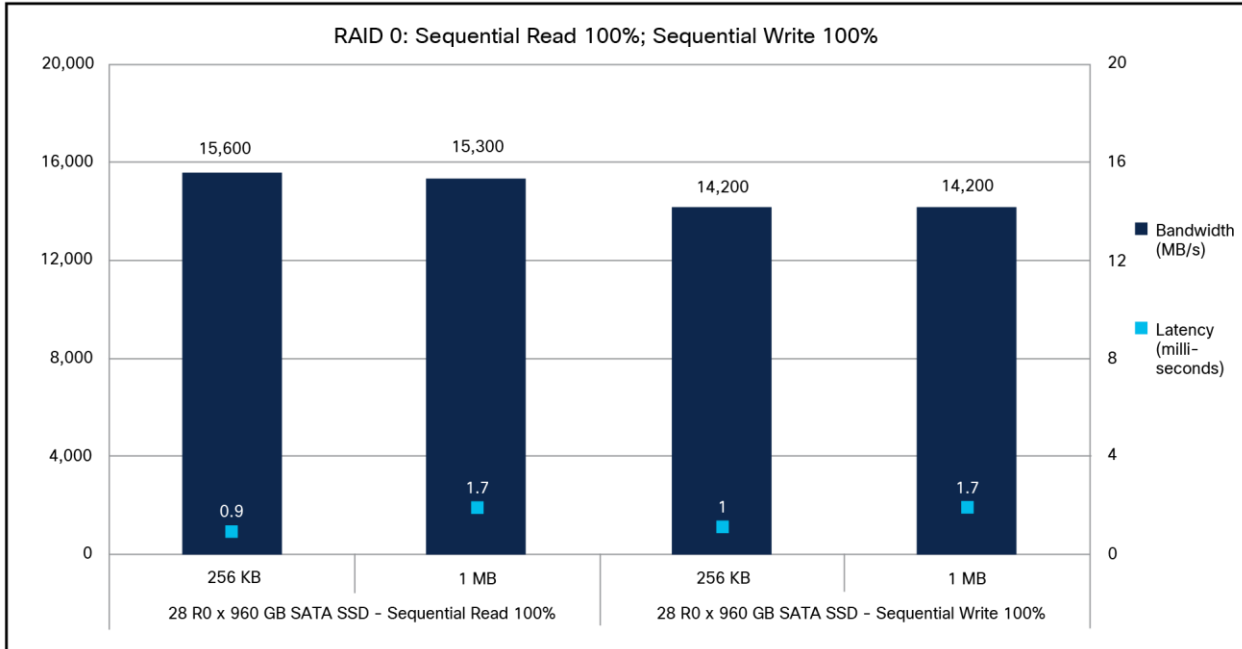


Figure 9.
Cisco UCS C240 M7SX RAID 0 100-percent sequential read; 100-percent sequential write

SATA SSD RAID 5 performance

Figure 10 shows the performance of the SSDs under test for the RAID 5 configuration with 70:30- and 50:50-percent random read:write access patterns. The graph shows the comparative performance values achieved for enterprise value SATA SSD drives to help customers understand the performance for the RAID 5 configuration. The graph shows that the 960 GB SATA drive provides performance of 1.1M IOPS with latency of 0.7 milliseconds for random read:write 70:30 percent and 570K IOPS with latency of 0.8 milliseconds for random read:write 50:50 percent with a 4 KB block size.

Each RAID 5 virtual drive was carved out using 7 disks, and a total of 4 virtual drives were created using 28 disks on dual RAID controllers (14 disks are configured on each controller). The performance data shown in the graph is cumulative data by running tests on all 4 virtual drives in parallel.

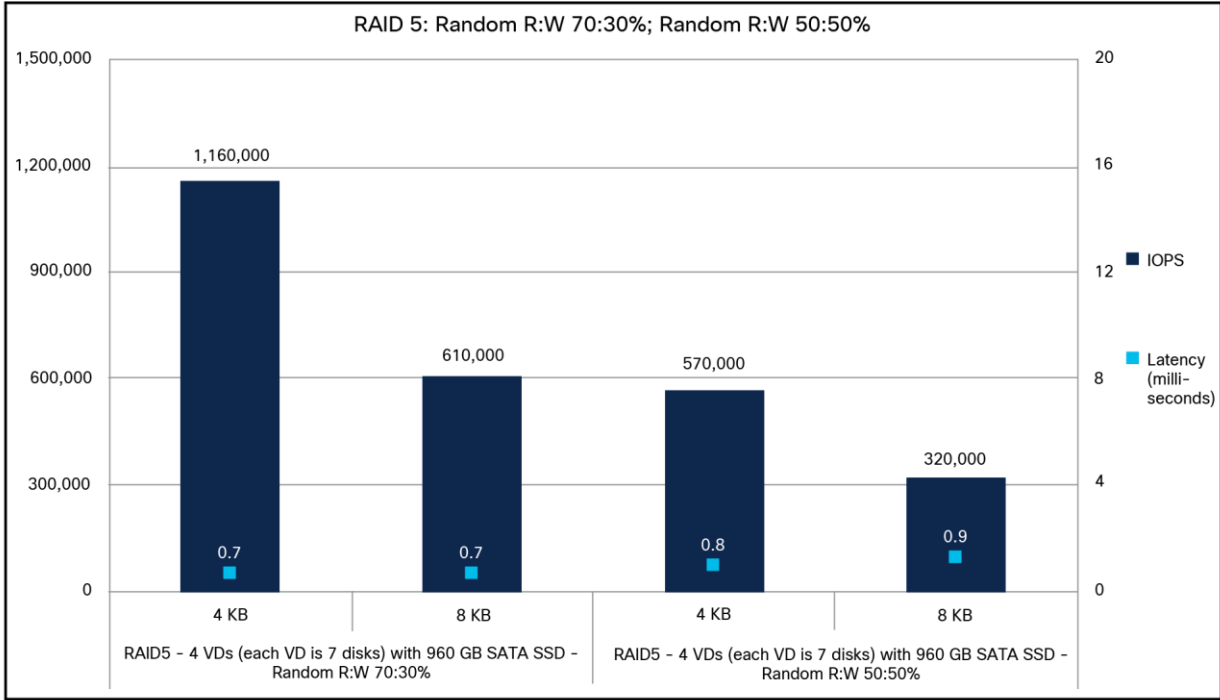


Figure 10.
Cisco UCS C240 M7SX RAID 5 70:30-percent random write; 50:50-percent random read:write

SATA SSD RAID 10 performance

The C240 M7SX server has two RAID controllers and each manages 14 drives. For the purposes of this RAID 10 testing, for each RAID controller, 2 RAID VDs (VD - virtual drive) were carved out, with one VD configured using 8 drives and second VD with 6 drives. RAID 10 volume requires even number of drives to be configured on each VD and to manage this requirement, we have taken this approach (from 14 drives on a single controller). The performance data shown in the graphs, is gathered from all 4 VDs across both the RAID controllers.

Figure 11 shows the performance of the SSDs under test for the RAID 10 configuration with a 100-percent sequential read and write access pattern. The graph shows the comparative performance values achieved for enterprise value SATA SSD drives to help customers understand the performance for the RAID 10 configuration. The graph shows that the 960 GB SATA drive provides performance of close to 12,000 MBps for sequential read and close to 6,200 MBps for sequential write within latency of 1 milliseconds for 256 KB and 1 MB block size.

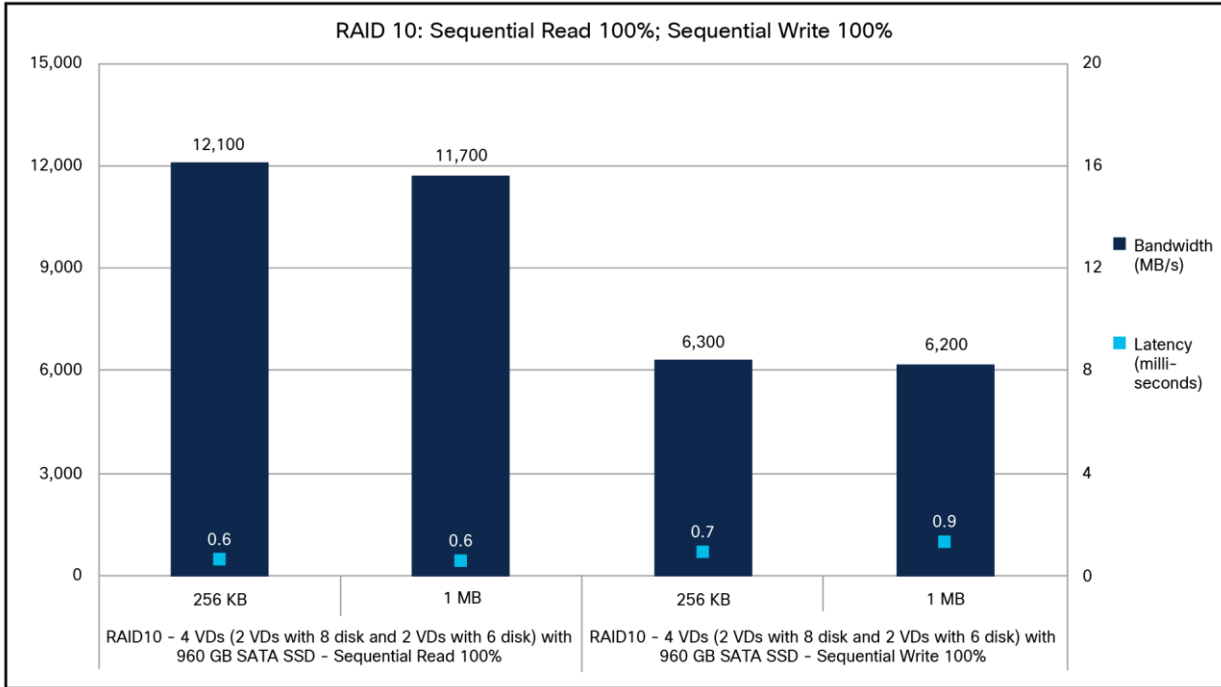


Figure 11.
Cisco UCS C240 M7SX RAID 10 100-percent sequential read; 100-percent sequential write

SATA SSD HBA performance

The C240 M7SX server supports dual 12G SAS HBA adapters. Each adapter manages 14 disks on a 28-disk system. This configuration helps in scaling of performance from individual adapters, based on the performance specification of drives. This dual adapter configuration is for performance, not for any fault tolerance between adapters.

Figure 12 shows the performance of the SSDs under test for the JBOD configuration with a 100-percent random read, 100-percent random write and 70:30-percent random read:write access pattern. The graph shows that the 960 GB enterprise value SATA drive provides performance of 2.3M IOPS with latency of 0.4 milliseconds for random read, 870K IOPS with latency of 0.3 milliseconds for random write, and 2M IOPS for random read:write 70:30 percent with a 4 KB block size. Latency is the time taken to complete a single I/O request from the viewpoint of the application.

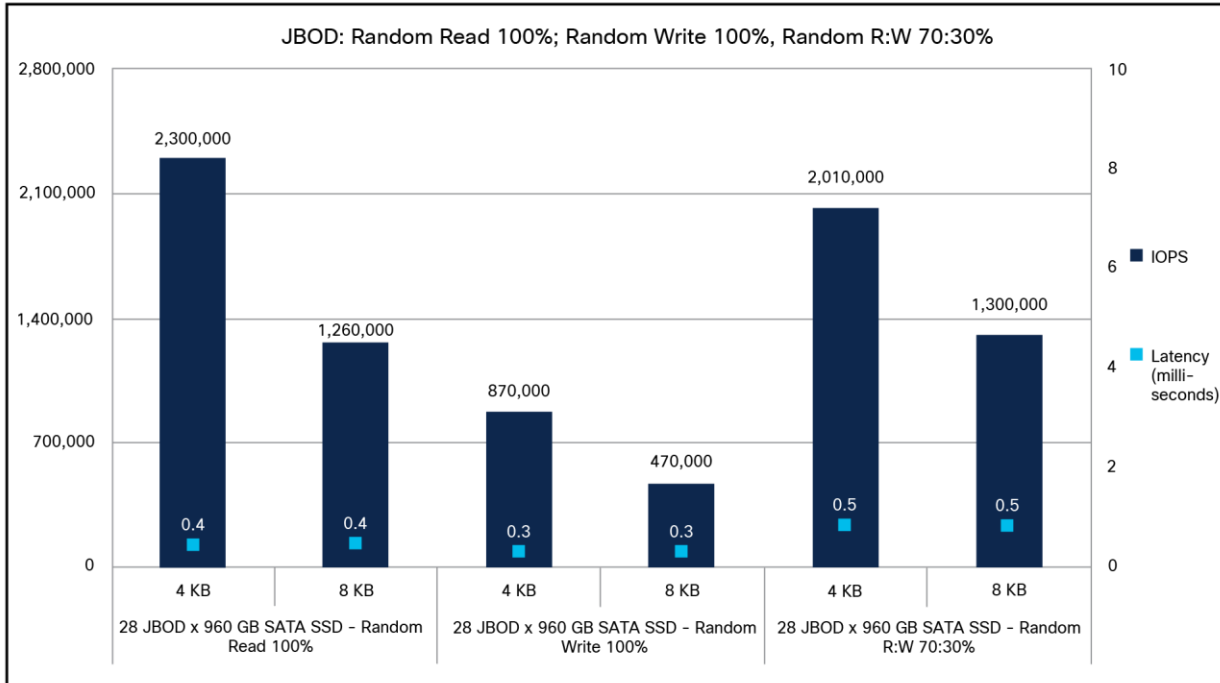


Figure 12. Cisco UCS C240 M7SX JBOD 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 13 shows the performance of the SSDs under test for the JBOD configuration with a 100-percent sequential read and write access pattern. The graph shows that the 960 GB enterprise value SATA SSD drive provides performance of more than 15,000 MBps with latency of 0.5 milliseconds for sequential read and more than 14,000 MBps for sequential write with the 256 KB and 1 MB block sizes.



Figure 13. Cisco UCS C240 M7SX JBOD 100-percent sequential read; 100-percent sequential write

NVMe performance data with a tri-mode storage controller

The Cisco Tri-Mode 24G SAS RAID Controller supports up to 14 SAS HDDs and SAS/SATA/NVMe SSDs operating at 3Gbps, 6Gbps, 12Gbps, and 24Gbps. It includes a SuperCap and a 4GB flash-back write cache (FBWC). Performance data was obtained using the FIO measurement tool on Red Hat Enterprise Linux (RHEL) with analysis based on the IOPS rate for random I/O workloads and MBps throughput for sequential I/O workloads. From this analysis, specific recommendations can be made for storage configuration parameters.

The I/O performance test results capture the maximum IOPS and bandwidth achieved with the NVMe SSDs within the acceptable response time (latency) of 2 milliseconds (ms).

The RAID controller (UCSC-RAID-HP) used in this testing is Broadcom 9670W-16i controller. Refer to the Broadcom 9670W-16i product brief (<https://www.broadcom.com/products/storage/raid-controllers/megaraid-9670w-16i>) for more information.

During our testing, results showed that for random I/O (reads) performance with NVMe drives, 5.5+ million IOPS were achieved with 14 NVMe drives populated on a single controller of the server. This was achieved as per the maximum performance of x1 PCIe lane for each NVMe SSD configured, and as per the controller design. However, given that the C240 M7SX server allows 28 disks to be populated on two controllers in the server, performance will be cumulative of two storage controllers, and random read was measured at 11+ million IOPS.

Cisco UCS C240 M7SX is the first rack server to support Cisco tri-mode storage controllers where NVMe drives can be configured for RAID. The dual adapter configuration is for performance, not for any fault tolerance between adapters.

Note: The section on performance data, below is for RAID configured with NVMe SSDs using the Cisco tri-mode controller. Performance data for JBOD or direct attached configuration with NVMe SSDs is detailed in different white paper. Here is a link to a M7 rack server NVMe I/O performance white paper: <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/ucs-c220-c240-m7-rack-server-wp.html>.

NVMe performance (3.2 TB U.3 NVMe SSD) for 28-disk configuration with tri-mode controller

NVMe SSD RAID 0 performance

Figure 14 shows the performance of the NVMe SSDs under test for a RAID 0 configuration with a 100-percent random read, 100-percent random write, and 70:30-percent random read:write access pattern. The graph shows the comparative performance values achieved for U.3 NVMe SSD drives to help customers understand the performance. The graph shows that the 3.2 TB NVMe drive provides performance of more than 11M IOPS with latency of 0.23 milliseconds for random read, performance of more than 10M IOPS with latency of 0.21 milliseconds for random write, and performance of 11M IOPS with latency of 0.22 milliseconds for random read:write 70:30 percent with a 4 KB block size. The result shows the maximum performance achievable from the C240 M7 tri-mode RAID controller. Latency is the time taken to complete a single I/O request from the viewpoint of the application.

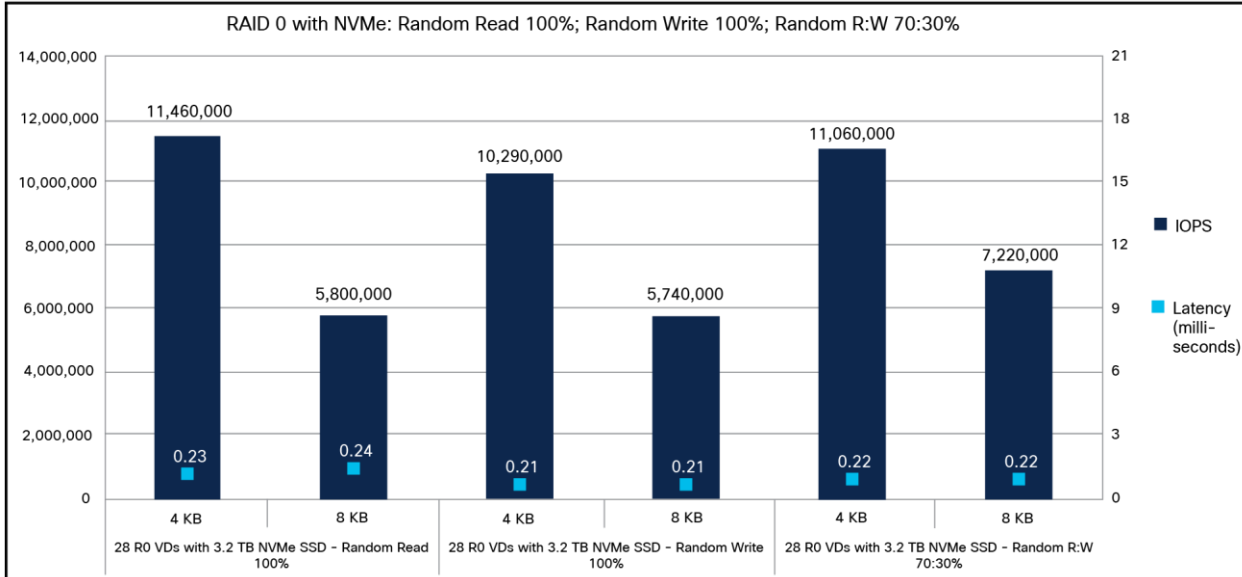


Figure 14. Cisco UCS C240 M7SX RAID 0 with NVMe 100-percent random read; 100-percent random write; 70:30-percent random read:write

Figure 15 shows the performance of the NVMe SSDs under test for a RAID 0 configuration with a 100-percent sequential read and write access pattern. The graph shows that the 3.2 TB NVMe SSD drive provides performance of more than 48,000 MBps for sequential read and sequential write, with 256 KB and 1 MB block sizes.

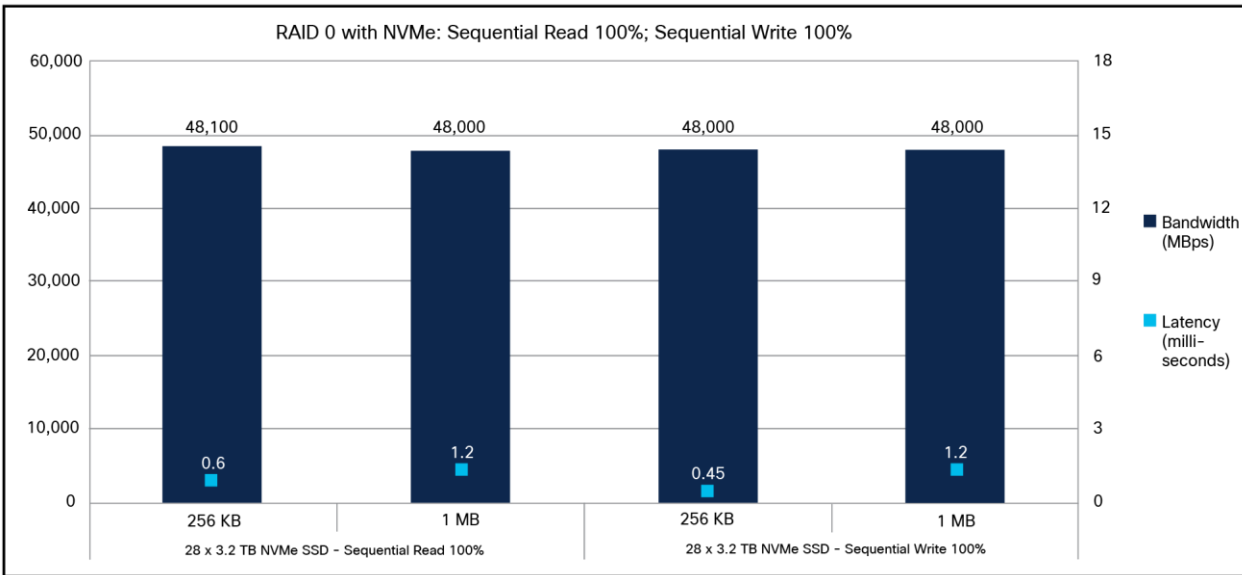


Figure 15. Cisco UCS C240 M7SX NVMe RAID 0 with NVMe 100-percent sequential read; 100-percent sequential write

NVMe SSD RAID 5 performance

Figure 16 shows the performance of the NVMe drives under test for a RAID 5 configuration with a 70:30- and 50:50-percent random read:write access pattern. The graph shows the comparative performance values achieved for NVMe SSD drives to help customers understand the performance for the RAID 5 configuration. The graph shows that the 3.2 TB U.3 NVMe drive provides performance of 3.1M IOPS with latency of 0.3 milliseconds for random read:write 70:30 percent and 1.9M IOPS with latency of 0.3 milliseconds for random read:write 50:50 percent with a 4 KB block size.

Each RAID 5 virtual drive was carved out using 7 disks, and a total of 4 virtual drives were created using 28 disks on dual RAID controllers (14 disks are configured on each controller). The performance data shown in the graph is cumulative data from running tests on all four virtual drives in parallel.

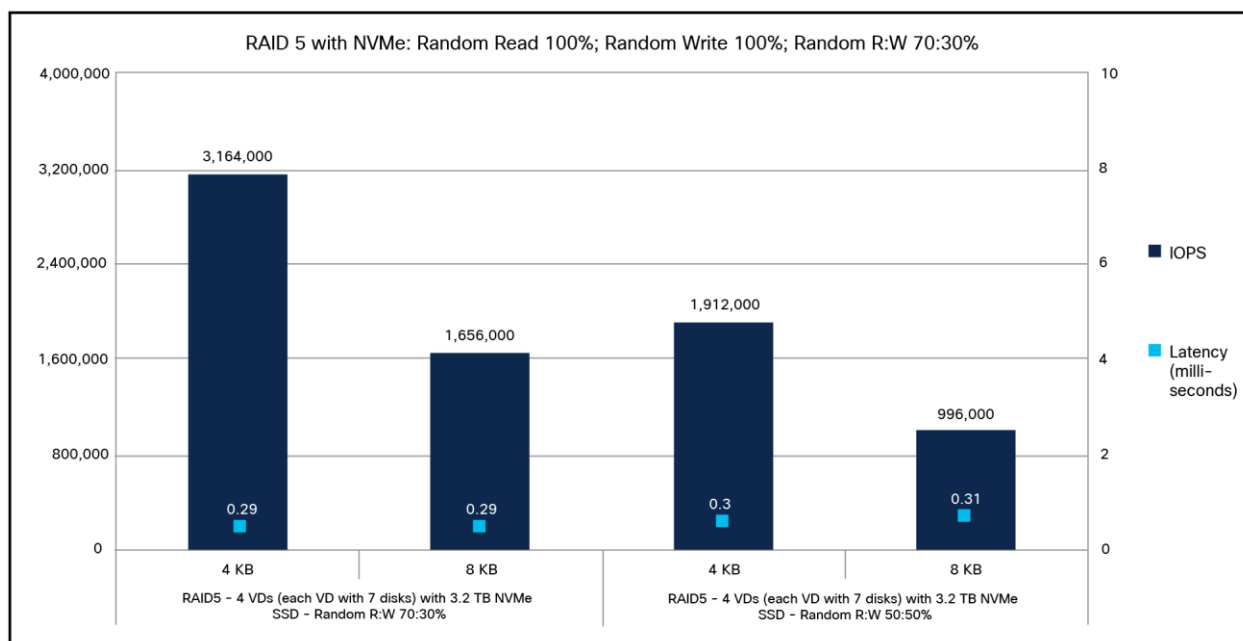


Figure 16.

Cisco UCS C240 M7SX RAID 5 with NVMe 70:30-percent random read:write; 50:50-percent random read:write

NVMe SSD RAID 10 performance

The C240 M7SX server has two RAID controllers, and each controller manages 14 drives. For this RAID 10 testing, for each RAID controller 2 RAID virtual drives (VDs) were carved out, with one VD configured using 8 drives and the second VD using 6 drives. The RAID 10 volume requires an even number of drives to be configured on each VD; to manage this requirement, we have taken this approach (from 14 drives on a single controller). The performance data shown in the graphs is gathered from all 4 VDs across both the RAID controllers.

Figure 17 shows the performance of the SSDs under test for a RAID 10 configuration with a 100-percent sequential read and write access pattern. The graph shows the comparative performance values achieved for enterprise value NVMe SSD drives to help customers understand the performance for the RAID 10 configuration. The graph shows that the 3.2 TB NVMe drive provides performance of 48,000 MBps within latency of 0.5 milliseconds for sequential read and 26,000 MBps within latency of 0.5 milliseconds for sequential write with 256 KB block size.

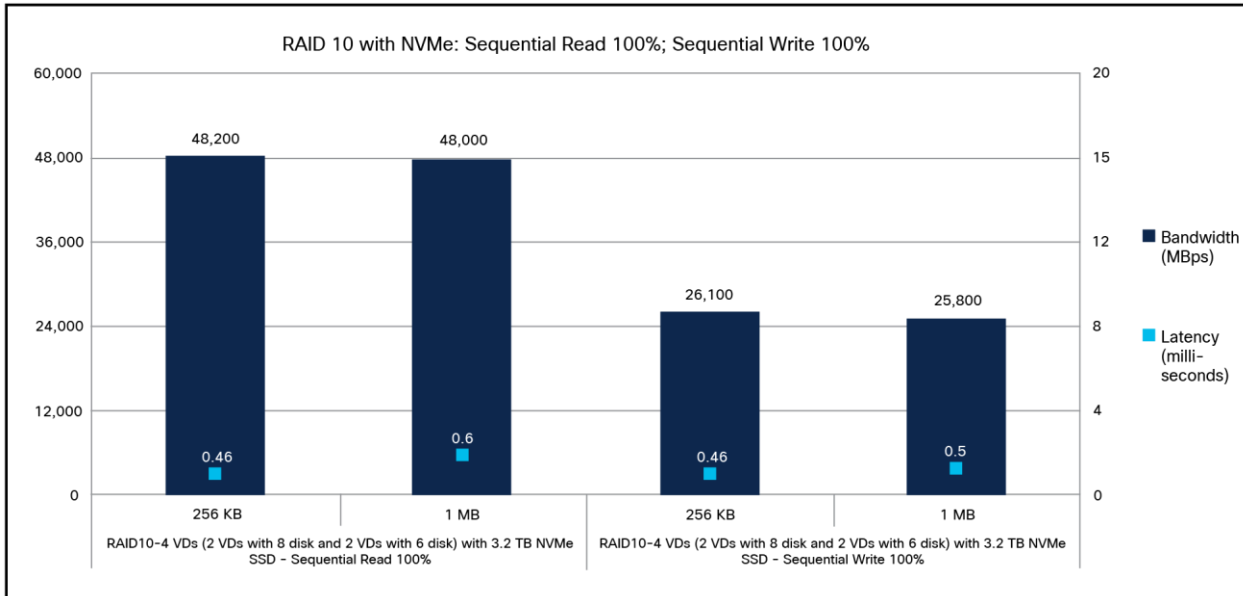


Figure 17. Cisco UCS C240 M7SX RAID 10 with NVMe 100-percent sequential read; 100-percent sequential write

Performance summary

The Cisco UCS C240 M7SX Rack Server is ideal for applications that demand large capacity with tiered storage benefits, where front and rear NVMe drives can be configured as a low-latency tier and storage-adapter-managed SSD drives for a capacity tier. With the introduction of a 24-Gbps modular tri-mode storage controller, NVMe hardware RAID can be configured on this storage controller, which was not possible before. Since the server can be configured with dual storage controllers, overall server performance will be cumulative of both controllers, depending on type of drive configured (SAS/SATA/NVMe).

There is a gain in storage I/O performance on the Cisco UCS C240 M7SX Rack Server compared to earlier generations of servers, and this is observed for both random I/O and sequential I/O workloads; this is mainly due to improved server architecture, storage controller design, support for NVMe SSDs with storage controller, new SAS4 drives.

For more information

For additional information on the server, refer to: <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/ucs-c240-m7-rack-server-ds.html>.

For information on BIOS tunings for different 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 on the FIO tool, refer to: https://fio.readthedocs.io/en/latest/fio_doc.html.

Appendix: Test environment

Table 9 lists the details of the server under test environment conditions.

Table 9. Server properties

Name	Value
Product names	Cisco UCS C240 M7SX
CPUs	CPU: Two 1.90-GHz Intel Xeon Platinum 8490H
Number of cores	60
Number of threads	120
Total memory	1024 GB
Memory DIMMs (16)	64 GB x 16 DIMMs
Memory speed	4800 MHz
VIC adapter	Cisco UCS VIC 15238 mLOM 40/100/200-Gbps Ethernet or FCoE
RAID controllers	<ul style="list-style-type: none">• Cisco Tri-Mode 24G SAS RAID Controller w/4GB Cache (UCSC-RAID-HP)• Cisco 12G SAS HBA (UCSC-SAS-T-D)
SFF SSDs	<ul style="list-style-type: none">• 1.9 TB 2.5-inch enterprise value 12G SAS SSD (UCS-SD19TKA1X-EV)• 960 GB 2.5-inch enterprise value 6G SATA SSD (UCS-SD960G6S1X-EV)• 3.2 TB 2.5-inch high performance high endurance NVMe SSD (UCS-NVMEG4-M3200)

Table 10 lists the server BIOS settings applied for disk I/O testing.

Table 10. Server BIOS settings

Name	Value
BIOS version	Release 4.3.3a
Cisco Integrated Management Controller (Cisco IMC) version	4.3 (3.240022)
Cores enabled	All
Hyper-Threading (All)	Enable
Hardware prefetcher	Enable
Adjacent-cache-line prefetcher	Enable
DCU streamer	Enable
DCU IP prefetcher	Enable
NUMA	Enable
Memory refresh enable	1x Refresh
Energy-efficient turbo	Enable
Turbo mode	Enable
EPP profile	Performance
CPU C6 report	Enable
Package C state	C0/C1 state
Power Performance Tuning	OS controls EPB
Sub NUMA Clustering	SNC4
Workload configuration	I/O sensitive

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)