JOINT SOLUTION BRIEF

Building Scalable, Ultra-low Latency NVMe™ Storage Solutions

ATTO
The Power Behind the Storage

Highlights

- Low-latency storage is achieved by using low-latency JBOF controllers in conjunction with NVMe
- ATTO XstreamCORE 9400 JBOF controllers are paired with Ultrastar DC SN640 NVMe SSDs
- The ATTO XstreamCORE 9400 adds a very low amount of latency to I/O operations, typically less than 2% even under load, allowing the storage to achieve its maximum performance
- Ultrastar DC SN640 provides high IOPS with consistently low latency

Applications

- Artificial Intelligence (AI) & Machine Learning (ML)
- Cloud-based DevOps
- Online Transactional Processing (OLTP) for High-Frequency Trading
- Finance and E-commerce
- Real-Time Analytics

Certain applications require extremely low latency response across the network—from host to storage—to achieve top performance. All NVMe storage solutions provide new opportunities to lower overall latency. However, limitations in the JBOF controllers that are used to aggregate storage into external NVMe arrays may degrade the native latency characteristics of NVMe SSDs.

Using low latency JBOF controllers in conjunction with NVMe provides an opportunity to achieve very low latency storage that supports latency-sensitive applications such as AI & MI, cloud-based DevOps, OLTP, finance & e-commerce and real-time analytics. ATTO XstreamCORE® 9400, in conjunction with Western Digital Ultrastar® NVMe drives, offers one such unique storage solution.

System Configuration Overview

This enterprise-class storage array was developed with ATTO XstreamCORE 9400 JBOF controllers as a front-end interface for 24 Western Digital Ultrastar NVMe SSDs. It provides a large capacity storage pool that can be located almost anywhere across the Ethernet fabric. Using standard NVMe-oF™ communication protocols, it delivers high-throughput, low-latency, shareable storage.

While this storage array is capable of efficiently handling large file transfers, this testing simulated typical database operations with blended READ and WRITE operations. Multiple hosts, connected to the array via a simple Ethernet fabric configuration, drove large amounts of I/O load to the NVMe array to simulate a high-demand application environment.

A Linux® FIO benchmark tool was used to generate and manage all disk I/O workloads.

Testing Methodology

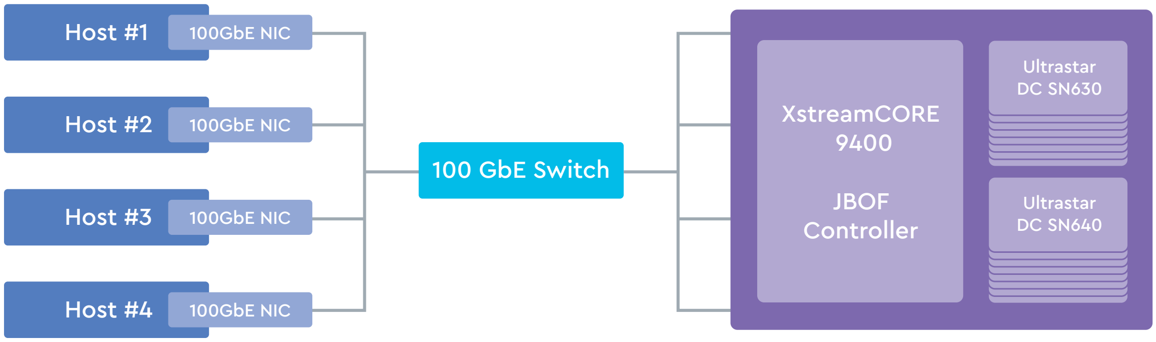
Testing focused predominantly on evaluating the data performance, measured in I/Os per second (IOPS), and latency added by each of the components in the storage array, measured in microseconds (µsec). Parameters included READ, WRITE and blended (70R/30W) data transfers.

We measured latency independently between the NVMe drives and JBOF controller to mitigate the performance impacts of the network and other, non-storage related latency. Our aim was to represent the performance of the storage array rather than the system as a whole since switches, host and other network variables would introduce unpredictable variations unnecessarily affecting our testing. Additional, application-specific testing is recommended to characterize overall system performance for a particular customer application environment.

Additionally, while it is presumed that dual, redundant JBOF controllers would be used in a failover configuration as part of any enterprise configuration, these features were not part of the scope of this testing.

Test Environment

The following line diagram depicts our test environment:



Details of the components used for this testing were as follows:

Component	Specifications	Comments
Host #1, Host #2	SuperMicro X10DRU-i+, Intel® Xeon® E-2643 v4 @ 3.4Ghz, 256G DDR4 2133Mhz	CentOS 8.2
Host #3, Host #4	Dell® PowerEdge® R630, Intel Xeon E5-2640 v3 @ 2.6Ghz, 64G DRR4 2133 MHz	CentOS 8.2
ATTO 100GbE NICs	ATTO FastFrame3 FFRM-N312	Only used single 100G port per card
Ethernet Network Switch	Mellanox® SN2100	
ATTO XstreamCORE 9400 Controller	4 × 100GbE network ports, x16 PCIe Gen 3 xCORE acceleration engine	
Ultrastar DC SN630	16 × 6.4TB NVMe SSSD	Direct connect – no backplane
Ultrastar DC SN640	8 × 3.2TB NVMe SSD	Direct connect – no backplane
CI Design Custom Enclosure		

Results

To establish a baseline comparison, the DC SN640 drives were tested in a direct-attach configuration as a local drive with varying queue depths. This established a maximum performance result for each model drive in our test configuration.

Results for Ultrastar DC SN640 in Steady State

	Random 4K READ		Random 4K WRITE	
	IOPS	Latency (µsec)	IOPS	Latency (µsec)
QD32	289,811	109	252,986	114
QD64	467,259	130	243,314	250
QD128	638,134	194	293,514	431

	Random 4K 70% READ/30% WRITE Random 4K READ		Random 4K 70% READ/30% WRITE Random 4K WRITE	
	IOPS	Latency (µsec)	IOPS	Latency (µsec)
QD32	235,430	116	101,114	35
QD64	225,807	256	94,251	47
QD128	263,810	458	112,025	64

System test results:

	Random 4K READ			Random 4K WRITE		
	IOPS	Total Latency (µsec)	Controller Latency (µsec)*	IOPS	Latency (µsec)	Controller Latency (µsec)*
QD32	3,421,239	111	2.7	1,655,172	230	1.1
QD64	3,410,432	149	2.8	1,553,714	332	1.1

	Random 4K 70% READ/30% WRITE READ			Random 4K 70% READ/30% WRITE WRITE		
	IOPS	Total Latency (µsec)	Controller Latency (µsec)*	IOPS	Latency (µsec)	Controller Latency (µsec)*
QD32	2,446,803	122	3.0	1,048,645	78	2.6
QD64	1,964,375	242	2.0	841,635	40	1.6

Recommendations and Conclusions

- Using the ATTO XstreamCORE 9400 JBOF controller with Ultrastar DC SN640 NVMe SSDs is a good technology choice for database and other applications that require extremely low latency for optimal performance.
- The ATTO XstreamCORE 9400 adds a very low amount of latency to I/O operations, typically less than 2% even under load, allowing the storage to achieve its maximum performance.
  - Ultrastar DC SN640 provides high IOPS with consistently low latency.
  - Maximum number of IOPs @4k READ achieved was 3,421,239.
- I/O was statistically equal across all four XstreamCORE 9400 100GbE ports, demonstrating a well-balanced controller design able to handle very high workloads without creating performance-limiting resource constraints.
  - During 100% random READ testing, average controller latency was 2.15% of the total latency with maximum latency of 2.4% (2.7 µsec).
  - During 100% random WRITE testing, average controller latency was 0.4% of the total latency with a maximum latency of 0.5% (1.1 µsec).
  - During blended 70/30 READ/WRITE testing, average READ controller latency was 1.65% of the total latency with a maximum latency of 2.46% (3.0 µsec). Average WRITE controller latency was 3.65% of the total latency with a maximum latency of 4.0% (1.6 µsec).

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com