A MICRON **TECHNICAL MARKETING** BRIEF

Running Away From Spin:
Cassandra Cluster Performance With Micron's M500DC Enterprise SSD – Up to 6.5X Better Performance in 1/6th the Space

Overview

Apache Cassandra is available as an open source or branded, bundled NoSQL database product. Both implementations are designed from the ground up to support a distributed (node distributed or geographically distributed) deployment model. Because Cassandra is easily scalable and does not require high-performance processing to reach very high I/O levels, it is widely adopted in big data analytics, automatic product recommendation systems, online catalog display, messaging platforms, query analytics, and a host of other real-time and near real-time applications.

Traditionally, Cassandra cluster performance has been scaled by adding nodes, but this increases complexity, cost, and space requirements. Micron's M500DC SSD improves per-node performance without additional requirements.

Speed and Spin: Oil and Water?

As Cassandra's breadth of deployment has grown, it has become clear: Speed is the engine for more growth and M500DC SSDs can fuel that engine.

Inherent HDD design limitations, together with Cassandra's internal read I/O process (reads are highly random) make even enterprise-grade HDDs a poor fit for Cassandra. They limit performance and require larger, more expensive and complex clusters, which can be both more costly and more difficult to manage.

Analyzing Single-Node Performance

Sizing an entire Cassandra cluster in one step is a challenge. Each node has to be designed, the number of nodes estimated, and the entire system deployed. However, analyzing single-node performance gives insight into overall cluster sizing requirements, as well as performance for a given workload.

Using YCSB to Measure Performance

A node performance sizing metric must support a variety of workloads that are closely aligned with common Cassandra use.

The Yahoo! Cloud Serving Benchmark (YCSB) is a widely adopted performance measurement framework for measuring key value pair databases.

YCSB's built-in workloads closely match several common Cassandra workloads.

WORKLOAD	READ %	WRITE %	USAGE EXAMPLE
A	50	50	Session/action recording
B	95	5	Adding tags to photos
C	100	0	Static data cache
D	95	5 (insert)	Recent stats tracking
F	50	50 (R/M/W)	User record changes

TABLE 1: Common Cassandra Workloads in YCSB

HDD vs. SSD Head-to-Head Comparison

For this comparison, 12 enterprise-grade HDDs were tested against 2 Micron M500DC SSDs using YCSB to measure common Cassandra workload performance. In each configuration, only the storage configuration was changed.

Twelve 10K RPM HDDs were split into two mdadm RAID 0 (stripe) sets of six drives each. (RAID 0 was chosen for optimal performance.)

One RAID set was mounted at /var/lib/cassandra/data, and the other set was mounted at /var/lib/cassandra/commitlog.

A pair of 800GB M500DC SSDs were also configured in Linux; mdadm RAID 0 was mounted at /var/lib/Cassandra. (The SSD design was simpler because the entire database and the commitlog were on the same RAID set.)

A test database was initially created using YCSB Workload A's load parameter to generate a 517GB dataset. A snapshot of this database was created within Cassandra and was backed up to a spare location. Prior to each test, the M500DC SSDs were restored to their FOB state and were properly preconditioned. For each configuration under test, the database was restored using this snapshot in order to start every test from a consistent state. Figure 1 shows the performance for the 12-drive HDD configuration compared to the 2-drive SSD setup by thread count.

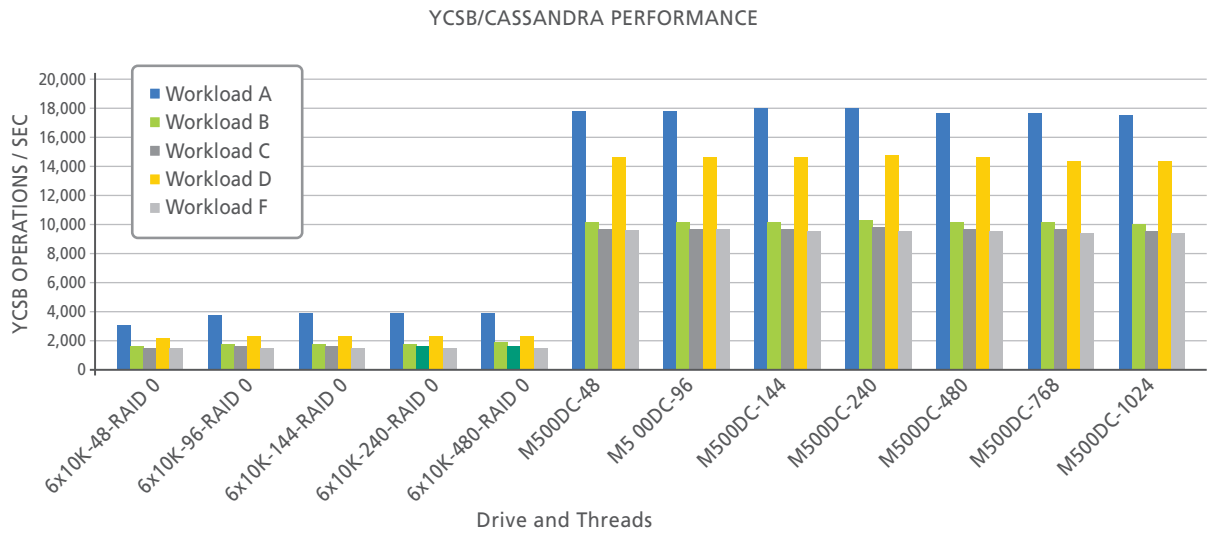

FIGURE 1: YCSB Performance by Workload and Thread Count – 12X HDD RAID Set and 2X M500DC RAID Set

M500DC SSDs Eclipse HDDs

» **Performance Gains.** The M500DC pair measured between 4.8X and 6.5X faster than the 12x HDD configuration across thread counts ranging from 48 to 480. (See Table 2.)

WORKLOAD	THREADS	M500DC ADVANTAGE	M500DC AVERAGE ADVANTAGE
A: Update Heavy	48	5.7	4.8X
	96	4.8	
	144	4.6	
	240	4.7	
	480	4.5	
B: Read Mostly	48	6.3	5.7X
	96	5.6	
	144	5.6	
	240	5.7	
	480	5.5	
C: Read Only	48	6.6	6.1X
	96	6.0	
	144	6.1	
	240	6.1	
	480	5.9	
D: Read Latest	48	6.9	6.5X
	96	6.3	
	144	6.3	
	240	6.4	
	480	6.3	
F: Read Modify Write	48	6.7	6.4X
	96	6.4	
	144	6.4	
	240	6.4	
	480	6.3	

TABLE 2: M500DC Performance Advantage

» **Space Efficiency.** While the HDD configuration required the commitlogs and the data to be on separate stripes (for best performance), the M500DC pair enabled a far simpler configuration with both the data and the commitlogs placed on the same stripe. The M500DC configuration occupied a total of one-sixth the drive bay space compared to the HDD configuration (2 bays versus 12 bays).

» **Smaller, Simpler Clusters.** Single-node performance can be a good indicator of expected cluster performance. Characterizing a single node can save time and money by providing a performance foundation. If each node performs better, fewer are needed.

The Bottom Line

Based on comparative performance by node and on recommendations from leading Cassandra vendors, SSDs are the logical choice when building new Cassandra clusters. Micron's M500DC enables far better per-node performance for simpler and cheaper clusters, while requiring far less storage space in each node.

micron.com