Image provided by Dell[†]

Process more data

with more YCSB operations per second*

Analyze data from Cassandra databases more quickly: Select Dell PowerEdge C6620 servers with Dell PowerEdge RAID controllers (PERC 12)

This new Dell PERC 12 solution delivered stronger Apache Cassandra distributed database performance than a legacy solution

Strategic business decisions are based on data that an organization collects, including fraud detection, compliance, or customer preference data. It is important to be able to quickly process that data and turn it into actionable insights. If your company uses Apache® Cassandra® distributed database workloads, improving performance by upgrading to new technologies could give you an edge by allowing for faster data analysis.

At Principled Technologies, we used Yahoo Cloud Serving Benchmark (YCSB) workload B to measure read-intensive Apache Cassandra performance of two server solutions in virtualized environments:

- A Dell™ PowerEdge™ C6620 server with a Dell PowerEdge RAID controller (PERC 12)
- A previous-generation Dell PowerEdge C6520 server with a Dell PERC 11 controller

The PowerEdge C6620 server featuring the PERC 12 RAID controller delivered more YCSB operations per second than the older solution. Handling more operations helps drive critical services that rely on data patterns, such as compliance tracking or fraud detection, which can improve the quality of your business decisions.

Dell PowerEdge C6620 server with a PERC 12 RAID controller vs. Dell PowerEdge C6520 server with a PERC 11 RAID controller

Our test approach

According to the Apache organization, Cassandra is "an open source NoSQL distributed database trusted by thousands of companies for scalability and high availability without compromising performance. Linear scalability and proven fault-tolerance on commodity hardware or cloud infrastructure make it the perfect platform for mission-critical data."[2]

We deployed two Cassandra environments, one with the Dell PowerEdge C6620 server with a Dell PERC 12 RAID controller, and one with the PowerEdge C6520 server with a PERC 11 RAID controller. The Dell PowerEdge C6620 server used two Dell U2 Gen4 3.84TB drives, while the Dell PowerEdge C6520 server used six 960GB mixed-use SAS 12Gbps SFF drives. Both environments used six Apache Cassandra database VMs joined in a cluster configuration.

To generate the test load and measure performance, we used YCSB workload B on each solution from a group of YCSB driver VMs. Improvements in operations per second processed and lower latency means an organization could catch compliance issues or other data anomalies sooner. It could also allow them to lower costs by using less hardware to perform a given amount of work.

Image provided by Dell[3]

About the Dell PowerEdge C6620 server

Part of the Dell modular infrastructure PowerEdge C-Series, the Dell PowerEdge C6620 features up to two 4th Generation Intel® Xeon® Scalable processors, with up to 56 cores per processor.

The PowerEdge C6620 offers memory speeds of up to 4,800 MT/s and supports up to 16 NVMe drives for workload acceleration. According to initial specification sheets we received from Dell, the PowerEdge C6620 provides flexible I/O options, including several PCIe slots and SNAP I/O support.

To learn more about the Dell PowerEdge C6620, check out information about the PowerEdge C-Series at https://www.dell.com/en-us/shop/modular-infrastructure/sf/modular-infrastructure.

About the Yahoo Cloud Serving Benchmark

According to Yahoo!, "the goal of the Yahoo Cloud Serving Benchmark (YCSB) project is to develop a framework and common set of workloads for evaluating the performance of different 'key-value' and 'cloud' serving stores."[4] The benchmark serves many databases including Apache HBase and Cassandra, two NoSQL databases that can handle large datasets.

For more information on YCSB, visit **https://research.yahoo.com/news/yahoo-cloud-serving-benchmark**.

Retrieve more detailed data insights sooner

In our testing, the Dell PowerEdge C6620 server with a Dell PERC 12 RAID controller delivered 1.25 times as many operations per second on YCSB workload B and lower application latency than the previous-generation PowerEdge C6520 with a PERC 11 RAID controller. This indicates that for this type of Apache Cassandra workload, the newer solution is faster, which can translate to catching anomalies in data earlier.

Total operations per second on YCSB workload B
Higher is better

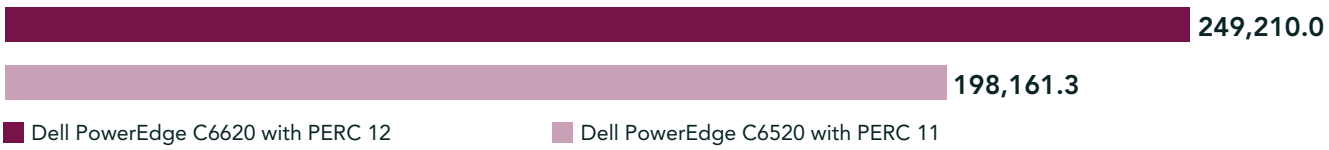


Figure 1: Total operations per second the two solutions achieved on YCSB workload B. Higher is better. Source: Principled Technologies.

Average read latency on YCSB workload B
Milliseconds | Lower is better

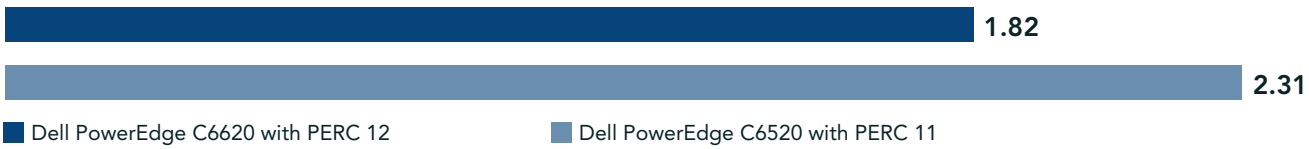


Figure 2: Average read latency achieved on YCSB workload B. Lower is better. Source: Principled Technologies.

Average update latency on YCSB workload B
Milliseconds | Lower is better

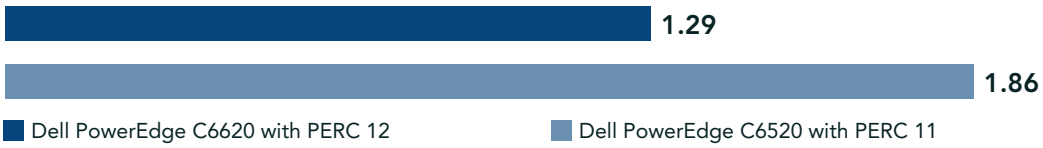
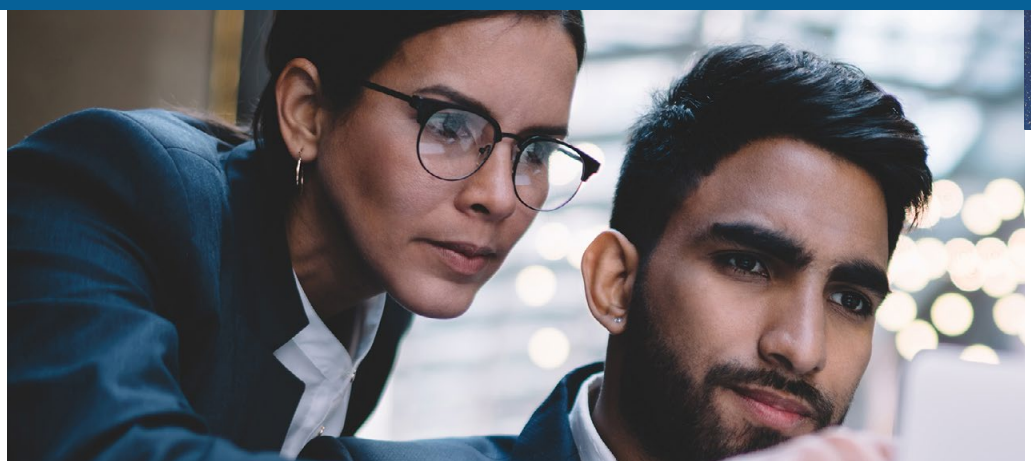


Figure 3: Average update latency achieved on YCSB workload B. Lower is better. Source: Principled Technologies.

About the Dell PERC 12 RAID controller

The latest series of Dell RAID controllers, the PERC 12 Series, offers expanded support and capabilites compared to previous versions. According to documentation we received from Dell, PERC 12 RAID controllers offer support for 24Gbps SAS drives, have 3,200 MHz cache memory speed, support the 16-lane host bus type, and support both NVMe and SAS on the front controller.

To learn more about the PERC 12, visit **https://www.dell.com/support/kbdoc/en-us/000131648/list-of-poweredge-raid-controller-perc-types-for-dell-emc-systems**.



Conclusion

When speedy, in-depth analysis is crucial in a NoSQL environment, relying on legacy hardware can hurt your company. We found that a Dell PowerEdge C6620 server with a Dell PERC 12 RAID controller handled more OPS for a Cassandra data analysis workload than a previous-generation Dell PowerEdge C6520 server with a PERC 11 RAID controller. This superior performance could help business units in your organization reap in-depth actionable intelligence from unstructured data more quickly.

---

1. Dell provided the image showing a fully populated Dell PowerEdge C6600 chassis. Our Dell PowerEdge C6600 chassis included four Dell PowerEdge C6620 blades and eight disks. We conducted our testing on one blade and two disks.
2. Apache, "Apache Cassandra Open Source NoSQL Database," accessed November 14, 2022, https://cassandra.apache.org/_/index.html.
3. Dell provided the image showing a fully populated Dell PowerEdge C6600 chassis. Our Dell PowerEdge C6600 chassis included four Dell PowerEdge C6620 blades and eight disks. We conducted our testing on one blade and two disks.
4. "Yahoo Cloud Serving Benchmark," accessed November 14, 2022, https://research.yahoo.com/news/yahoo-cloud-serving-benchmark.

Read the science behind this report at https://facts.pt/A5YAzMh ▶

Principled Technologies®

Facts matter.®

This project was commissioned by Dell Technologies.