Complete data analysis faster with Google Cloud C3 high CPU instances enabled by 4th Gen Intel Xeon Scalable processors

Complete queries 1.88 times as fast

Compared to N2 high CPU instances with previous-gen processors, these instances sped up query completion times

The underlying hardware of cloud instances running database analytics workloads is typically the biggest factor in how long it takes to complete those queries. Completing data analysis faster reduces the time it takes for decision makers to spot trends and ultimately make choices that can steer a business to success.

To show the benefits that newer technologies can have on database analytics performance, we compared the SQL Server TPROC-H performance of two Google Cloud™ instance types: C3 high CPU instances with 4th Gen Intel® Xeon® Scalable processors and N2 high CPU instances with previous-gen processors. Both instance types used Google Cloud Hyperdisk Extreme storage, a new network block storage service. In our analytics tests using multiple simultaneous data streams, the C3 high CPU instances completed queries nearly twice as fast as similarly configured N2 high CPU instances with previous-gen processors.

Getting faster query speeds using C3 high CPU instances means that your organization can reduce the time it takes to gain business insights, finish analytics within predetermined windows, and ultimately save funds by reducing the instance uptime you have to pay for.

This project was commissioned by Intel.

How we tested

We compared two similarly configured high CPU instances with processor technology from different generations:

Table 1: Key configuration details for the two instances we tested. Source: Principled Technologies.

	N2 (previous-gen)	C3 (current gen)
Instance	n2-highcpu-80	c3-highcpu-88
Region	us-central-1a	us-central-1a
CPU cores/ vCPU	80	88
Memory (GB)	80	176
Processor	Intel Xeon Platinum 8373C	Intel Xeon Platinum 8481C
Storage	1TB Hyperdisk Extreme, 100K IOPS (data) 60GB SSD persistent (log)	1TB Hyperdisk Extreme, 100K IOPS (data) 60GB SSD persistent (log)
DB size (WH/GB)	300/300	300/300

Google Cloud has many options for instance sizing, and performance varies significantly among them. At the time of our testing, the only C3 instance type available for general use was the C3 high CPU. While we would typically choose an instance with a higher memory-to-core ratio for a database workload, the standard and high-memory C3 instances were still in private preview at the time of testing. For comparison, we chose the N2 high CPU instance type and configured it as similarly as possible. Hyperdisk Extreme disks require at least 64 vCPUs, so we configured the C3 instance with 88 vCPUs and selected the closest available N2 configuration, which was 80 vCPUs.

We chose the TPROC-H 300 scale database (~300 GB) to ensure that it would not fit entirely into RAM and would engage the disks. We ran the recommended number of simultaneous streams for a 300-scale database according to HammerDB documentation and sized the disks to provide enough performance to avoid bottlenecks. Note: At the time of our testing, Google Cloud reported a known issue with Hyperdisks of limited performance on Windows instances. With this issue in mind, we used subsystem metrics to confirm that the storage was not bottlenecking performance during our runs.

To learn more about our testing, including step-by-step instructions for completing these tests, read the science behind the report.

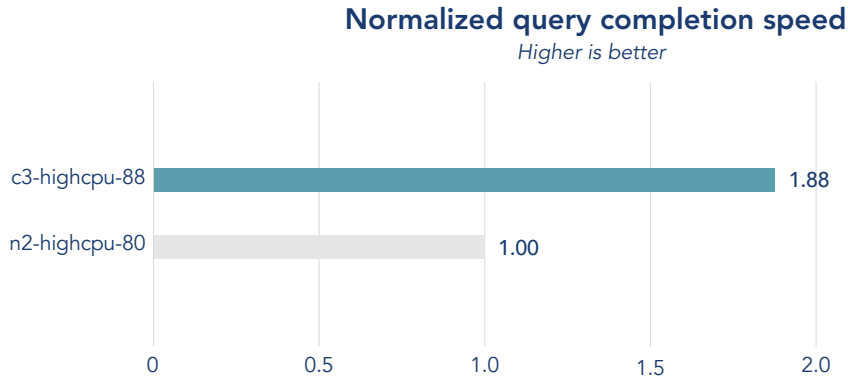
About the HammerDB workload

The TPROC-H workload from the HammerDB benchmark suite is an online analytics processing (OLAP) workload. This test measures the time required for instances to analyze streams of database queries, where one stream comprises 22 serialized database queries. While the HammerDB developers derived TPROC-H from the TPC-H specification, the workload is not a full implementation of the TPC-H standard. As such, the results in this paper are not directly comparable to officially published TPC-H results.

For more information, please visit https://www.hammerdb.com/docs/ch11s01.html.

Speed up data analytics by choosing Google Cloud C3 instances with 4th Gen Intel Xeon Scalable processors

We ran our OLAP workload with six simultaneous streams and used the longest completion time, in seconds, for our comparison. Figure 1 shows that on our HammerDB TPROC-H workload, new C3 instances with 4th Gen Intel Xeon Scalable processors completed analytics queries 1.88 times as fast compared to the previous-gen N2 instances with 3rd Gen Intel Xeon Scalable processors. Both instance types used new Google Cloud Hyperdisk Extreme storage.



Complete queries 1.88 times as fast

Figure 1: HammerDB TPROC-H workload results, in normalized query completion speed, that the N2 and C3 instances achieved. Higher numbers are better. Source: Principled Technologies.

These results show that selecting the C3 high CPU instance type over an instance using previous-generation hardware can help organizations complete their data analytics in significantly less time, which can shrink the time to insight and allow companies to save money by reducing the instance uptime they have to pay for.

About Google Cloud Hyperdisk

Google Cloud now offers Hyperdisk, a network block storage service for NVMe and SCSI interfaces. According to Google Cloud, Hyperdisk "offers a scalable, high-performance storage service with a comprehensive suite of data persistence and management capabilities. With Hyperdisk you can easily provision, manage, and scale your Compute Engine workloads without the cost and complexity of a typical on-premises storage area network (SAN)."[1] There are two types of Hyperdisk: Hyperdisk Extreme, for maximum performance for the most demanding workloads, and Hyperdisk Throughput, for flexibility and efficiency of scale-out workloads. In our testing, we used 1TB Hyperdisk Extreme storage.

To learn more about Google Cloud Hyperdisk, visit https://cloud.google.com/compute/docs/disks/hyperdisks.

About 4th Generation Intel Xeon Scalable processors

The Intel strategy for 4th Gen Intel Xeon Scalable processors "aligns CPU cores with built-in accelerators optimized for specific workloads and delivers increased performance at higher efficiency for optimal total cost of ownership."[2] Intel claims the processors deliver "a range of features for managing power and performance, making the best use of CPU resources to achieve key sustainability goals. In addition, the Xeon CPU Max and the Max Series GPU add high-bandwidth memory and maximum compute density to solve the world's most challenging problems faster."[3]

To learn more about the latest processor technology from Intel, visit https://www.intel.com/content/www/us/en/products/docs/processors/xeon-accelerated/4th-gen-xeon-scalable-processors.html.

Conclusion

Data collected, but not analyzed, offers only potential, not insights, to an organization. Sorting and analyzing that data can turn rows and tables into meaningful insights about what customers want or which business initiatives should proceed. The faster an organization gets insights from data, the quicker its leaders can make decisions. In our tests, where both instance types used Google Cloud Hyperdisk Extreme storage and were configured similarly, Google Cloud C3 high CPU instances with 4th Gen Intel Xeon Scalable processors delivered 1.88x the query completion speeds of N2 high CPU instances with previous-gen processors. Selecting these newer C3 high CPU instances with Hyperdisk can help ensure that data analysis completes within targeted windows and give decision makers the insights they need to quickly adjust to the changing tides of business and realize success.

---

1. Google, "Google Cloud Hyperdisk," accessed July 27, 2023, https://cloud.google.com/compute/docs/disks/hyperdisks.
2. Intel, "4th Gen Intel Xeon Scalable Processors," accessed July 27, 2023, https://www.intel.com/content/www/us/en/newsroom/resources/press-kit-4th-gen-intel-xeon-scalable-processors.html.
3. Intel, "4th Gen Intel Xeon Scalable Processors."

Read the science behind this report at https://facts.pt/PxqDj5w ▶

Principled Technologies®

Facts matter.®

This project was commissioned by Intel.