

Google Cloud™ N2 VM Instances Executed 2.82x the Wide & Deep Inference Work and Delivered 2.75x the Performance per Dollar than N1 VM Instances



HiBench



**Complete 2.82x
the Wide & Deep
Inference Work**

than on N1 instances



**Reap 2.75x
the Inference
Performance
per Dollar**

than on N1 instances

Get a better return on your Google Cloud investment by selecting new N2 VM instances Featuring 2nd Gen Intel® Xeon® Scalable Processors

As the amount of data online explodes, companies increasingly rely on inference tools that help users cut through this abundance to find meaningful relationships in the data. One such tool is the Wide & Deep Learning Recommender System. It uses a deep learning topology that brings together wide linear models and deep neural networks.

Like other deep learning workloads, recommender systems are, by nature, compute-intensive and take advantage of powerful processors. To carry out your Wide & Deep workloads quickly and deliver recommendations sooner, select a new Google Cloud N2 instance type enabled by 2nd Gen Intel® Xeon® Scalable processors.

In TensorFlow Wide & Deep tests comparing 32-vCPU Google Cloud VM instances, new N2 VM instances enabled by 2nd Gen Intel Xeon Scalable processors completed 2.82x the work of older N1 VM instances. With a cost that is less than 3 percent higher than the older instances, the new instances achieved 2.75x the performance per dollar.

For your deep learning recommender system workloads, choose a new N2 instance enabled by 2nd Gen Intel Xeon Scalable processors.

Wide & Deep INT8 Inference Relative Performance (Higher is better)

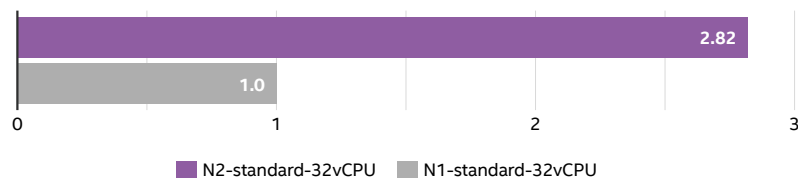


Figure 1. TensorFlow test results comparing the Wide & Deep performance of the 32-vCPU N2 Standard VM instance type to the 32-vCPU N1 Standard VM instance type.



More than twice the performance means you need half as many VM instances

You've decided to run your Wide & Deep Learning Recommender System workloads on Google Cloud N-series instances. Did you know that by selecting new N2 instances enabled by 2nd Gen Intel® Xeon® Scalable processors you could cut the number of instances you need in half?

As Figure 1 shows, in TensorFlow tests comparing the Wide & Deep performance of VM instances with 32 vCPUs Google Cloud N2 VM instances enabled by Intel Xeon Scalable processors completed 2.82x the work of an N1 instance using older processors. This means a single N2 VM instance could perform more work than two N1 VM instances.

Another way to understand the benefits of choosing new Google Cloud N2 VM instances enabled by Intel Xeon Scalable processors is to consider pricing. Despite their much greater performance, these new instances cost only 2.2 percent more than their older counterparts. As Figure 2 shows, when we combine pricing and performance, the new N2 VM instances deliver 2.75x the work per dollar than the older N1 VM instances.

Wide & Deep INT8 Inference Relative Performance per Dollar (Higher is better)

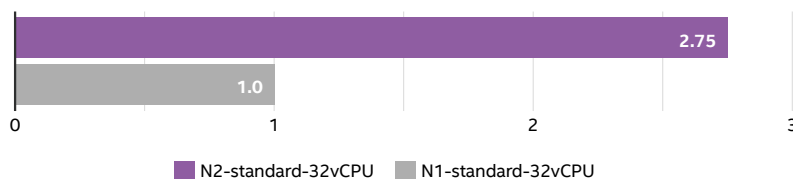


Figure 2. Relative Wide & Deep inference performance per dollar of the 32-vCPU N2 Standard VM instance type and the 32-vCPU N1 Standard VM instance type.

If you chose the new N2 VM instances enabled by Intel Xeon Scalable processors, you would get greater performance per dollar and save because of the fact that you'd be able to replace two or three older N1 VM instances with a single new one to meet the a given performance threshold.

Learn More

To begin running your deep learning applications on Google Cloud Platform N2 VM instances with 2nd Gen Intel Xeon Scalable processors, visit <http://intel.com/googlecloud>.



Performance varies by use, configuration and other factors. Learn more at <https://intel.com/benchmarks>.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy. Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others

Printed in USA 0321/JO/PT/PDF US001

Please Recycle