Complete data warehouse queries up to **1.54x as fast**

Scale up your work without drastically increasing query times

Get consistently better performance across instance sizes

Improve your Microsoft SQL Server query times with new Microsoft Azure instances

Powered by Intel Xeon Platinum 8272CL processors, the new instances completed queries in less time than older instances based on Intel Xeon E5-2673 v4 processors

If your company is considering how best to migrate your database analytics work to the cloud, the wide variety of options available may make it difficult to decide which specific cloud instance is right for you. One important factor to consider is the processor that powers the instance. Newer processors, such as the most recent offerings from Intel, often come with technological advances that can help you take your work further and finish it in less time.

At Principled Technologies, we tested a data warehouse workload on two memory-optimized series of Microsoft SQL Server instances for Microsoft Azure: new Eds_v4 instances powered by Intel® Xeon® Platinum 8272CL processors and older Es_v3 instances powered by Intel Xeon E5-2673 v4 processors. In our tests, the new Eds_v4 instances completed queries in the data warehouse workload faster than the older Es_v3 instances, regardless of VM size or the number of data streams running simultaneously. In this report, we examine the results of our data warehouse testing in three different ways to paint a detailed picture of what they can mean for your company.

How we tested

We used a data warehouse workload from the HammerDB suite to test Microsoft Azure SQL Server instances in the memory-optimized Eds_v4 and Es_v3 series: new instances powered by Intel Xeon Platinum 8272CL processors, and older instances powered by Intel Xeon E5-2673 v4 processors. (Microsoft Azure SQL Server Es_v3 instances are available in a few different CPU configurations, but we used only Intel Xeon E5-2673 v4 processors for our testing.) The HammerDB developers derived the data warehouse workload from TPC-H benchmark specifications. Because the HammerDB test does not totally comply with the original TPC-H benchmark, our results are not directly comparable to published TPC-H results.

We chose three instance sizes for this testing to demonstrate the performance of the new Eds_v4 series compared to the same instance size in the Es_v3 series at different workload and performance levels. As with most other cloud services, Microsoft Azure defines an instance's size by its vCPU count. We chose to test a 4-vCPU pair, a 16-vCPU pair and a 64-vCPU pair.

Microsoft Azure scales an instance's memory to better fit its compute power. In the same vein, we also chose to scale the drive performance and database size to represent a likely workload scenario for each compute level. To ensure the CPU was under more workload pressure than the storage drives, we sized each database to fit within the RAM allocated in each instance. Here are the instance specifics we tested:

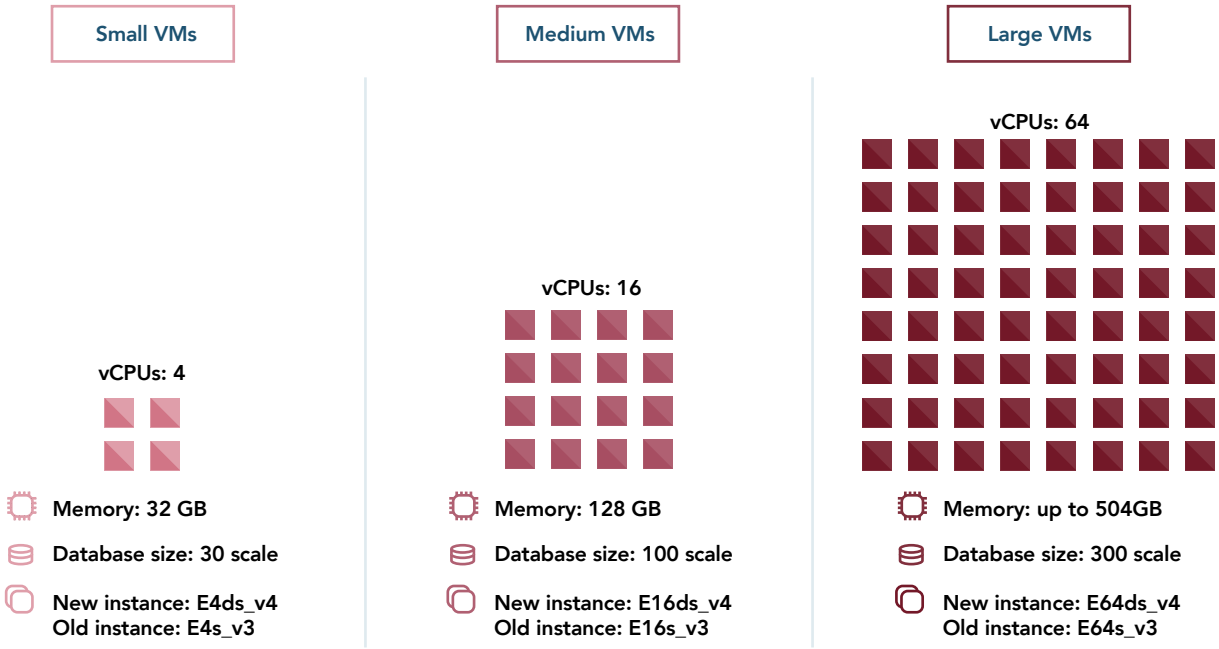


Figure 1: Specifications for the Microsoft Azure SQL Server VMs we used for testing.
Source: Principled Technologies.

The HammerDB data warehouse workload measures the time required to complete sets (or "streams") of 22 serialized database queries. For each database size, we tested up to the maximum number of streams from Transaction Processing Performance Council (TPC) recommendations.[1] We ensured that each VM's processors were fully saturated at the maximum stream count.

In the following sections, we present our findings from three different perspectives: performance comparisons between the Eds_v4 and Es_v3 instances across stream counts, consistency of the performance difference between the old and new instance pairs with increasing VM size, and performance consistency of the Eds_v4 instances with increasing VM size.

Complete database analytics query streams faster

We began our tests with a single stream of data warehouse queries. To determine how the increasing amounts of work impacted query times for each instance, we then scaled up to the maximum number of streams that the TPC-H specification recommends for each size of database. Figures 2 through 4 show that regardless of VM size, the new Eds_v4 instances achieved better performance than the older Es_v3 series instances: Up to 1.39 times as fast for small instances, up to 1.54 times as fast for medium instances, and up to 1.53 times as fast for large instances.

The more data streams you can support concurrently, the more data you can analyze at once. Plus, executing the queries faster means you can use that analysis sooner to gain insights that will help you take action to improve your organization. Whether your work requires processing single query streams or multiple streams concurrently, you should see a significant performance increase with the new Eds_v4 instances.

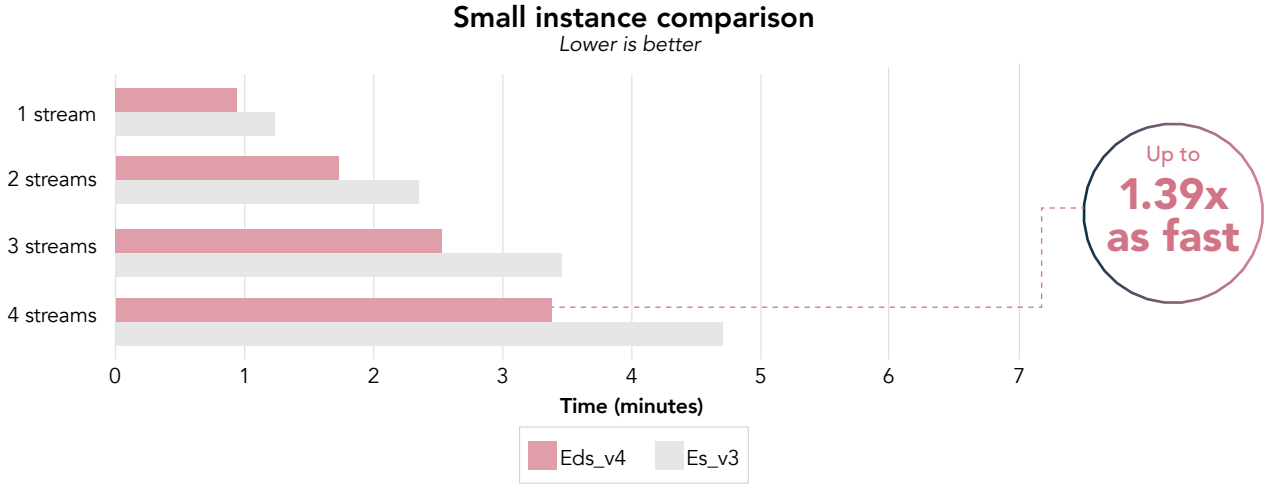


Figure 2: Time (in minutes) to complete concurrent query set streams from the HammerDB benchmarking suite. Less time is better. Source: Principled Technologies.

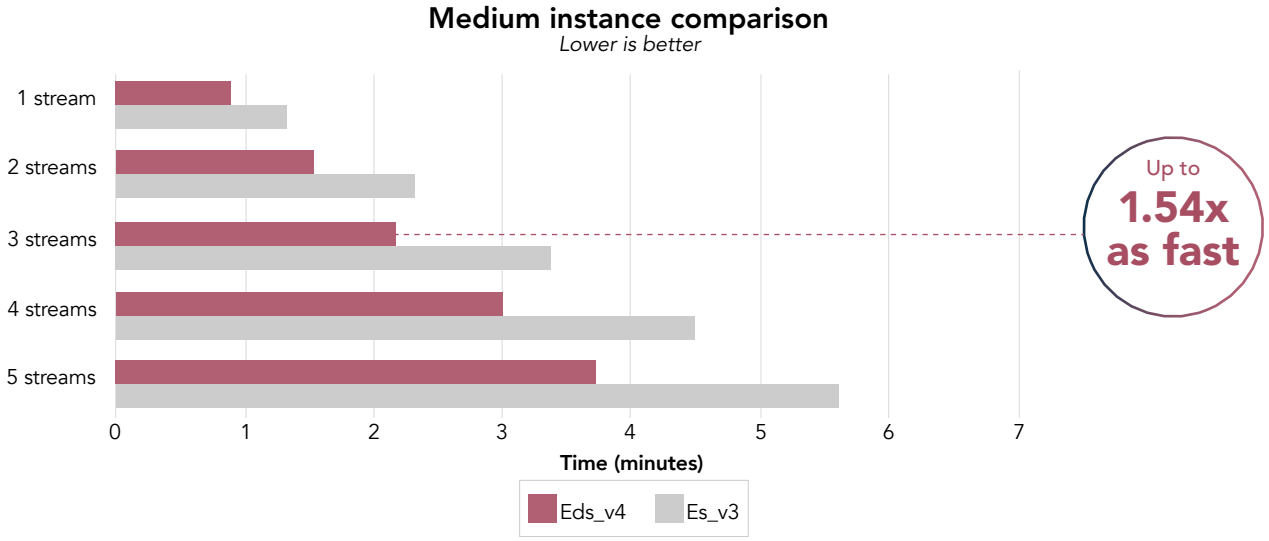


Figure 3: Time (in minutes) to process concurrent query set streams from the HammerDB benchmarking suite. Less time is better. Source: Principled Technologies.

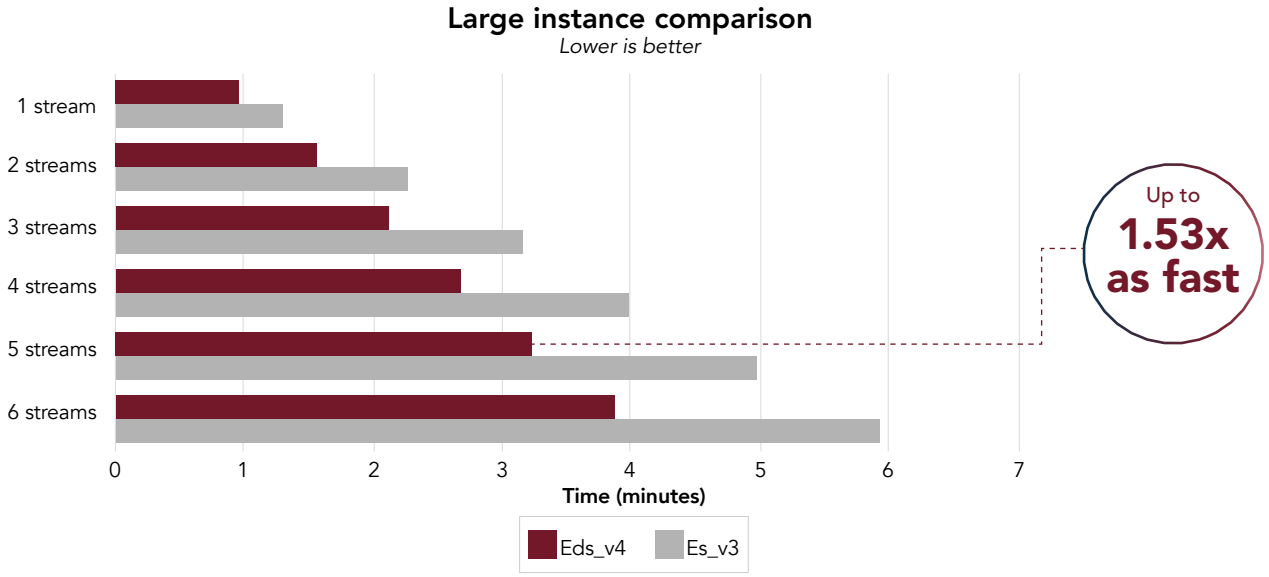


Figure 4: Time (in minutes) to process concurrent query set streams from the HammerDB benchmarking suite. Less time is better. Source: Principled Technologies.

To get a better idea of how our results can relate to the real world, consider the following hypothetical scenario. Each night, a company has a four-hour window in which to analyze data on a 300GB database. The company uses this analysis to generate reports for executives each morning and to drive other business processes throughout the day.

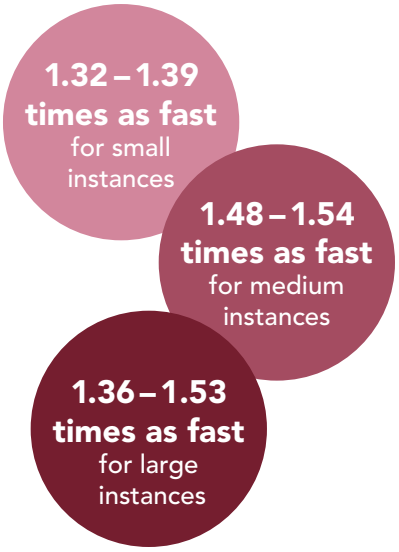
Based on the results of our single-stream tests, we calculate that a new E64ds_v4 VM would enable this hypothetical company to run 252 query sets within their analysis window each night. By contrast, the E64s_v3 VM would complete just 184 query sets in the same time frame (27% fewer).

If this company only needed to complete 184 query sets each night, the new E64ds_v4 VM would enable them to finish their work in just 2.9 hours, shrinking the required analysis window by 1.1 hours compared to the E64s_v3 VM. Over the course of a year, this hypothetical company would save 401.5 hours—or, 16.7 full days' worth—of analysis time per year while maintaining the same rate of analysis they could get with the E64s_v3 VM, enabling them to save money on VM uptime.

Get consistently better database analytics speed

At every data point, the new Eds_v4 SQL Server instances powered by Intel Xeon Platinum 8272CL processors were faster than the older Es_v3 instances powered by Intel Xeon E5-2673 v4 processors. For each VM size, as we increased the number of concurrent data streams up to four, the relative performance advantage remained similar and in many cases improved.

This means that regardless of your VM sizing needs or the number of concurrent streams you need to support, you can complete your database analytics work faster on the new instances than on the older instances. In other words, you can get consistently better performance with the new Eds_v4 series, even as your workload scales to larger VM instances.



Process more database analytics work without drastically longer wait times

Every business has differing database, workload size, and performance needs. Choosing the right Azure instance for your data warehouse databases is important to ensure you can meet those needs. To show how the Intel Xeon Platinum 8272CL processor-based instances scale in performance, we measured the time each Eds_v4 instance took to complete a range of simultaneous streams from a single stream up to four, the maximum number of streams the 4 vCPU VM supported.

Figure 5 shows that even when deploying larger databases and running more concurrent streams on the new Eds_v4 instances, the time to complete the queries remained relatively close. In our tests, we increased the size of the database with each VM size increase. The large VM we used hosted a 300-scale database—10 times larger than the 30-scale database on the small VM. The consistency we saw in testing is thanks in large part to the vCPU scaling at each VM level.

With the new Eds_v4 instances, you shouldn't have to worry that increasing the database size or number of data streams will dramatically increase the time it takes you to complete queries. The large VM we tested required about a minute to process a single query stream, and only 2.6 minutes to process four streams concurrently.

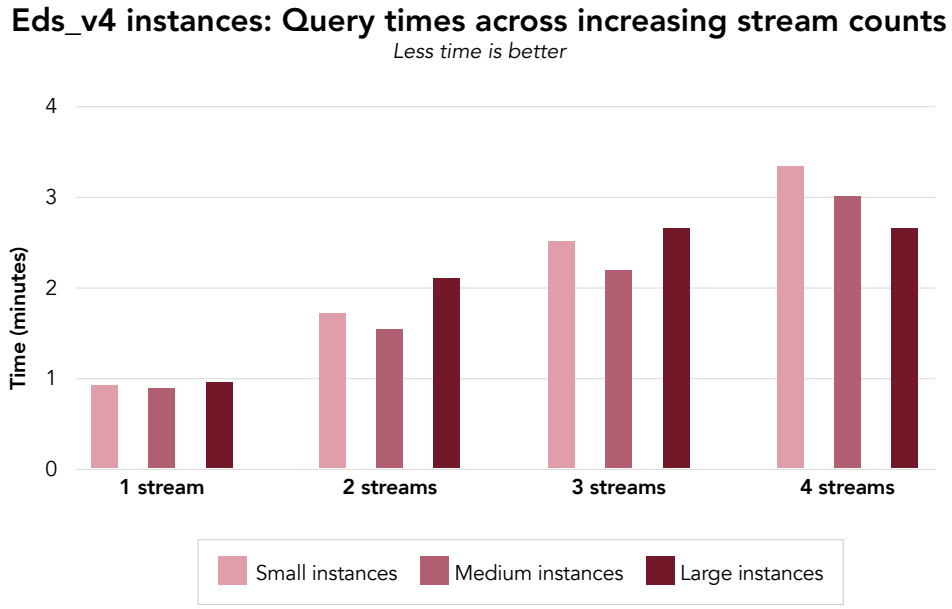


Figure 5: Consistency in time to complete data warehouse query set streams across small, medium, and large Eds_v4 instances. Less time is better. This graph shows that, despite larger databases on the larger VMs, the time to complete a given number of query streams remains similar across VM size. Source: Principled Technologies.

Performance that more than justifies the cost

Our performance findings suggest the Eds_v4 instances can complete data warehouse workloads anywhere from 1.32 to 1.49 times as fast as the Es_v3 instances. Yet, at the time of this writing, all sizes and specifications of the Eds_v4 instances cost just 1.17 times as much as their Es_v3 counterparts.2 By investing in Eds_v4 instances as opposed to Es_v3 instances, you could be getting better performance for your money.

Conclusion

When moving database analytics work to the cloud, it's important to consider the type of processor that will be driving your workloads. In our tests comparing two versions of Microsoft Azure SQL Server instances, the newer Eds_v4 instances, powered by Intel Xeon Platinum 8272CL processors, were faster at processing data warehouse query streams compared to Es_v3 instances powered by Intel Xeon E5-2673 v4 processors. This trend held for small, medium, and large VMs.

By choosing these new Eds_v4 instances, your organization could more quickly complete work that can lead to key decisions that can improve your business and help it grow. And, because the new Eds_v4 instances save big on time, it means choosing to invest in them can ensure your dollar goes further.

1 TPC Benchmark H Standard Specification Revision 2.18.0," accessed August 6, 2020, http://www.tpc.org/tpc_documents_current_versions/pdf/tpc-h_v2.18.0.pdf.

2 "Windows Virtual Machine Pricing," accessed August 13, 2020, https://azure.microsoft.com/en-us/pricing/details/virtual-machines/windows/.

Read the science behind this report at http://facts.pt/ccqc4xe ►

Principled Technologies®

Facts matter.®

This project was commissioned by Intel.