Executive summary

Finding the path to AI success with the Dell AI portfolio

A comparison of the Dell AI portfolio vs. similar offerings from Supermicro

To implement artificial intelligence (AI) and reap its benefits, you must consider all aspects of a solution. Finding a partner that can support end to end ,from planning, data prep, hardware selection, AI model design, proof-of-concept testing, reference architectures, and end-to-end support is crucial. In this summary, we highlight performance and space advantages of servers in the Dell™ AI portfolios over those from Supermicro based on publicly available MLPerf® Benchmark test results. MLPerf® tests performance for both training and inferencing on several AI models. The data in this summary comes from MLPerf® v3.1 Inference Datacenter results published on the MLCommons® website from November 2023.[1] We also compare publicly available information about the range of products and solutions each vendor offers to support AI deployments.

Four-GPU servers: Dell PowerEdge servers offered better MLPerf performance than Supermicro servers

With up to four NVIDIA H100 SXM GPUs, the 2U Dell PowerEdge XE9640 offers half the GPU computational power of the PowerEdge XE9680 in one-third the space.[2] The four-GPU server with NVIDIA HGX H100 GPUs from Supermicro in the MLPerf® submissions is the SYS-421GU-TNXR, which is a 4U server. In published gptj-99.9 AI model results, the PowerEdge XE9640 outperformed the Supermicro server in the offline tests, achieving up to 1.37 times the score (see Figure 1).

We also compared the Supermicro SYS-421GU-TNXR server results to the Dell PowerEdge XE8640, a 4U four-GPU server that also supports NVIDIA H100 HGX GPUs. The Dell PowerEdge XE8640 features the latest 4th Generation Intel Xeon Scalable processors and up to 4 TB of memory to handle the large datasets and complex computations common in AI and data analytics.[3] With its 4U form factor, the PowerEdge XE8640 is similar to the Supermicro SYS-421GU-TNXR in both density and GPU capabilities. However, much as we saw with the PowerEdge XE9640, the Dell PowerEdge XE8640 got better gptj-99 scores in the offline tests than the Supermicro server (see Figure 2).

Dell offers a broader portfolio of solutions, services and consumption options to support customers' AI deployments

Up to 37% better performance in half the rack space
for natural language processing (GPT-J) per MLperf data

Dell Validated Designs include more AI-centric reference architectures than Supermicro validated solutions

Normalized MLPerf® Results: Dell PowerEdge XE9640 with H100 SXM5 vs. Supermicro SYS-421GU-TNXR with H100 SXM5 *(Larger is better)*

Model	Accuracy		Dell PowerEdge XE9640	Supermicro SYS-421GU-TNXR
gptj-99	99	Server	1.02	1.00
		Offline	1.37	1.00
gptj-99.99	99.9	Server	1.02	1.00
		Offline	1.37	1.00

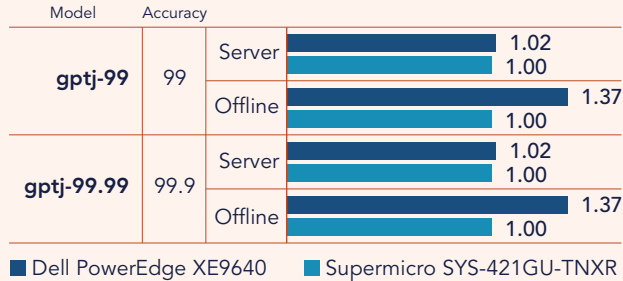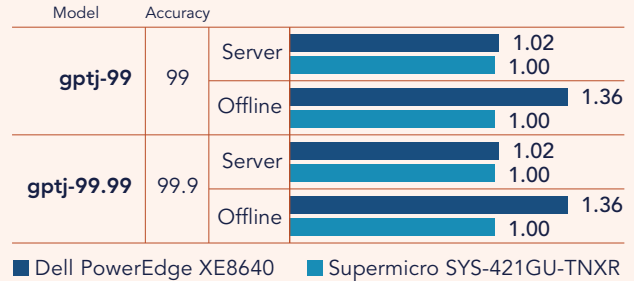■ Dell PowerEdge XE9640 ■ Supermicro SYS-421GU-TNXR

Figure 1: Published MLPerf® results for the Dell PowerEdge XE9640 and Supermicro SYS-421GU-TNXR as of 11/29/23. Both systems equip the SXM form factor of the NVIDIA H100 GPU. Source: Principled Technologies using data from MLCommons®.*

Normalized MLPerf® Results: Dell PowerEdge XE8640 with H100 SXM5 vs. Supermicro SYS-421GU-TNXR with H100 SXM5 *(Larger is better)*

Model	Accuracy		Dell PowerEdge XE8640	Supermicro SYS-421GU-TNXR
gptj-99	99	Server	1.02	1.00
		Offline	1.36	1.00
gptj-99.99	99.9	Server	1.02	1.00
		Offline	1.36	1.00

■ Dell PowerEdge XE8640 ■ Supermicro SYS-421GU-TNXR

Figure 2: Published MLPerf® results for the Dell PowerEdge XE8640 and Supermicro SYS-421GU-TNXR as of 11/29/23. Both systems equip the SXM form factor of the NVIDIA H100 GPU. Source: Principled Technologies using data from MLCommons®.*

Beyond hardware: Looking at the bigger picture

Organizations looking for flexibility in GPU workstation offerings can find a wider range of options in the Dell AI portfolio. The greater Dell AI portfolio encompasses more than just hardware; it gives customers a single services and support solution for every stage of their deployment. In contrast, the Supermicro AI portfolio offers services limited to manuals, troubleshooting, and product returns. In our research, we were unable to find design, implementation, management, or educational services in the Supermicro AI portfolio. Because AI adoption is so complex, the services that Dell offers make it a more promising training partner for deploying these workloads. Plus, we found that Dell Validated Designs include more AI-centric reference architectures than Supermicro validated solutions and offered more in-depth guidance.

Read our full report, which explores in depth the benefits that the Dell AI portfolio offers in compute, storage, and networking options; professional services for AI; third-party partnerships for AI workloads; Dell Validated Designs, which help take the guesswork out of AI hardware solutions; and management services and iDRAC.

*Verified MLPerf. score of v3.1 Inference Closed. Retrieved from https://mlcommons.org/benchmarks/inference-datacenter/ 5 December 2023, entries 3.1-0133, 3.1-0066, and 3.1-0067. The MLPerf. name and logo are registered and unregistered trademarks of MLCommons. Association in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited. See www.mlcommons.org for more information.

1. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results," accessed February 7, 2024, https://mlcommons.org/en/inference-datacenter-31/.

2. Dell, "PowerEdge XE9640 Rack Server," accessed January 5, 2024, https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640.

3. Dell, "PowerEdge XE8640 Rack Server," accessed January 5, 2024, https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640.

Read the report at https://facts.pt/q9p46K9 ▶

Principled Technologies®

Facts matter.®