

Dell PowerEdge XE7740 supported up to 74 concurrent AI agents in Principled Technologies testing

Test report highlights results from 4-, 6-, and 8-GPU configurations on real-world agentic AI workloads in manufacturing and financial services.

ROUND ROCK, TX, UNITED STATES, July 17, 2026 /EINPresswire.com/ -- Principled Technologies (PT) has released a new report evaluating the agentic AI performance of the Dell PowerEdge XE7740, a purpose-built enterprise AI server. PT testing measured how three GPU configurations of the server, powered by Intel Xeon 6747P processors and NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs, handled real-world agentic AI workloads in the manufacturing and financial services industries.

PT tested three configurations of the Dell PowerEdge XE7740: one with 4 GPUs, one with 6 GPUs, and one with 8 GPUs. On a financial services workload with a 30-second latency SLA, the 8-GPU configuration of the PowerEdge XE7740 supported up to 74 concurrent AI agents and sustained 1,179 tokens per second. On a manufacturing workload with a 90-second SLA, the 8-GPU configuration of the PowerEdge XE7740 supported 15 concurrent agents and delivered more than double the throughput of the 4-GPU configuration.



Gain enterprise-ready agentic AI with the Dell PowerEdge XE7740

Exploring agentic AI performance results for multiple configurations of the Dell PowerEdge XE7740

Dell™ PowerEdge™ XE7740

- Intel Xeon 6747P processors
- NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs

Product images provided by Dell Technologies

Executive summary

Agentic AI is moving from pilot to production, and with that shift comes a new and daunting set of demands on enterprise infrastructure. A single agent task can spawn hundreds of inference calls, each requiring GPU compute, CPU coordination, and tool execution working in tight concert. That requires computing power, and lots of it. For IT decision-makers planning their AI infrastructure spend, the question is what platform to choose—and which configuration of that platform—to have the best chance of meeting that demand as workloads grow. Organizations concerned about security and local data control are seeking to choose the right servers for their data centers.

We tested several GPU configurations of the Dell™ PowerEdge™ XE7740 server, powered by Intel Xeon® 6747P processors and featuring NVIDIA RTX™ PRO 6000 Blackwell Server Edition GPUs, on two agentic AI workloads. We ran each workload across multiple GPU configurations to measure how performance scaled and offer valuable sizing advice for IT buyers. We also assessed one configuration's tokenomics to look at an example of how this on-premises solution can deliver value year over year. Our results empower IT buyers to choose a solution that can grow with their workflows, reducing the risk that the infrastructure decisions you make today create problems on the horizon.

Spend as little as \$0.09/Mtok to achieve strong agentic performance

Support up to 74 AI agents with <30s latency to task completion on a financial services workload

Get up to 1,179 tokens/second on an agentic AI financial services workload

A Principled Technologies report: Hands-on testing. Real-world results.

Report: Gain enterprise-ready agentic AI with the Dell PowerEdge XE7740

Agentic AI represents a significant change in enterprise AI infrastructure demand. Where a traditional chatbot query might involve a single inference call, a single agent task can spawn hundreds. That workload increase falls on both the GPU and the CPU, making server platform choice critical for organizations planning production deployments of agentic AI.

Key takeaways from the report included:

- More GPUs mean more agents—increasing GPU count also increased the number of agents the server could support
- Agentic AI demands CPU as well as GPU power—PT saw a wide range of CPU utilization, peaking at 99% during task assignment and agent orchestration
- On-premises ownership has the potential to offer a cost advantage—compared to Amazon Bedrock, the PowerEdge XE7740 offered a lower cost per million tokens (\$/Mtok)

To test the server, PT used “a custom agentic benchmark that runs realistic AI-agent workflows end-to-end on a computing solution.” As PT notes in the report, “Because agentic AI agents can do a wide range of work, much of it industry- and company-specific, we built the benchmark to simulate a range of realistic industry-specific workflows. For this study, we used the financial services and manufacturing workflows, each of which incorporates eight workflow scenarios.”

To learn more, [read the full report](#).

About Principled Technologies, Inc.

Principled Technologies, Inc. is the leading provider of fact-based marketing and competitive analysis backed by transparent methodologies.

Principled Technologies, Inc. is located in Durham, North Carolina, USA. For more information, please visit www.principledtechnologies.com.

Sharon Horton

Principled Technologies, Inc.

press@principledtechnologies.com

Visit us on social media:

[LinkedIn](#)

[Facebook](#)

[YouTube](#)

[X](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/927218646>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

