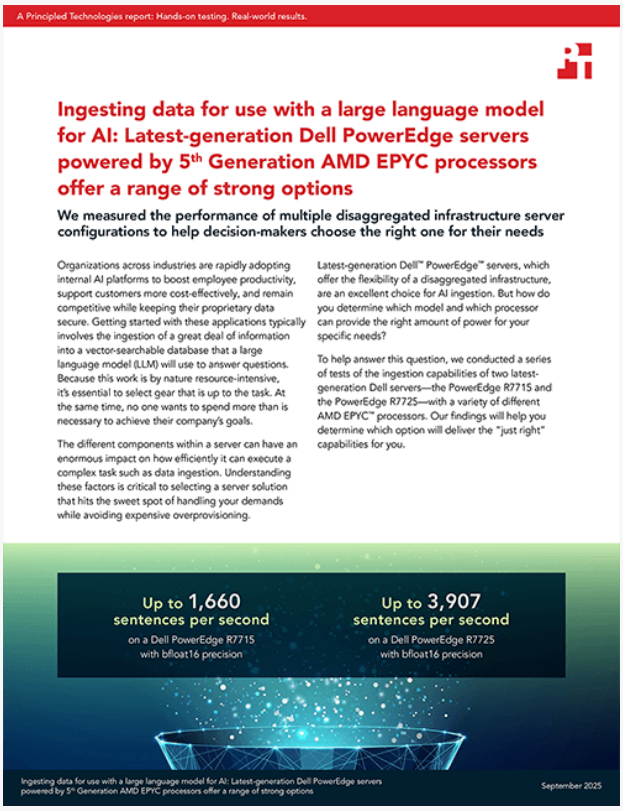
New study finds that Dell PowerEdge servers with AMD EPYC processors can provide strong AI ingestion performance

The third-party report highlights the performance of multiple server configurations to help decision-makers choose the right one for their needs

ROUND ROCK, TX, UNITED STATES, September 16, 2025 / EINPresswire.com/ -- Organizations in all sectors are implementing internal artificial intelligence (AI) platforms to enhance employee efficiency, provide more cost-effective customer support, and maintain a competitive edge. This often means ingesting large volumes of data into a vector-searchable database that a large language model (LLM) will utilize to respond to queries. When choosing the technologies that will support this process, companies must balance getting the resources they need without unnecessary overprovisioning. A disaggregated architecture, as opposed to a hyperconverged infrastructure (HCI), can enable companies to scale compute separately from storage and potentially save costs.

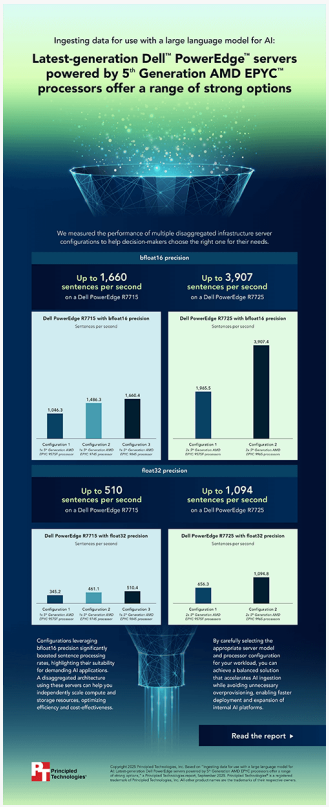


Ingesting data for use with a large language model for AI: Latest-generation Dell™ PowerEdge™ servers powered by 5th Generation AMD EPYC™ processors offer a range of strong options

Third party Principled Technologies (PT) assessed AI ingestion speeds on five configurations of Dell PowerEdge servers powered by 5th Generation AMD EPYC processors, which offer the flexibility of a disaggregated infrastructure. Their study

covered the Dell PowerEdge R7715 server and the Dell PowerEdge R7725 server, and they tested all configurations at two different precision levels—float32 and bfloat16—to highlight performance in multiple contexts.

According to the report, “In our testing, latest-generation Dell PowerEdge R7725 and R7715 servers powered by 5th Generation AMD EPYC processors demonstrated strong performance for ingesting information into vector-searchable databases for use by LLMs in AI applications. Configurations leveraging bfloat16 precision significantly boosted sentence processing rates, with the dual-socket PowerEdge R7725 models delivering up to 3,907 sentences per second, highlighting their suitability for demanding AI applications. A disaggregated architecture using these servers allows organizations to independently scale compute and storage resources, optimizing infrastructure efficiency and cost-effectiveness. By carefully selecting the appropriate server model and processor configuration based on workload needs, companies can achieve a balanced solution that accelerates AI ingestion while avoiding unnecessary overprovisioning, enabling faster deployment and expansion of internal AI platforms.”



An infographic accompanying the report

To see all the results and learn more about the Principled Technologies study, read the test report at https://facts.pt/qIC8rv2 or see the infographic at https://facts.pt/pJK7rBf.

About Principled Technologies, Inc.
Principled Technologies, Inc. is the leading provider of technology marketing and learning & development services.

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

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.