

# Study shows small businesses can run in-house AI chatbots on an AMD EPYC processor-powered Dell PowerEdge R6615 server

*Small organizations and departments can use this server as an AI entry point and upgrade to a CPU + GPU configuration to support more users as needs increase*

ROUND ROCK, TX, UNITED STATES, February 10, 2025 /EINPresswire.com/ -- Businesses of all sizes are eager to explore the potential of the in-house AI chatbot. By using a publicly available large language model (LLM) along with an organization's own data, these AI chatbots can improve customer support and boost employee productivity with search and content-generation capabilities—all while keeping data private and secure.

Decision-makers in smaller companies, believing that GPUs are a requirement for AI, might assume that the price of entry for such chatbots would be too high for them. In fact, benchmark testing performed by Principled Technologies (PT) revealed that a single-socket Dell PowerEdge R6615 server with only a CPU can be a great AI entry point for small businesses and departments within larger organizations, and can give them the option of scaling up by adding GPUs should their needs grow over time.

To investigate how different hardware options might support an in-house AI chatbot, PT used the PTChatterly benchmark tool to test two configurations of a single-socket Dell PowerEdge R6615



**Principled Technologies®**

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

### Run your in-house AI chatbot on an AMD EPYC 9534 processor-powered Dell PowerEdge R6615 server

Testing revealed that this server can be a great AI entry point for small organizations or departments, while upgrading to a CPU + GPU configuration could help them scale to support more users

Generative artificial intelligence (GenAI) is today's hottest topic, with decision-makers across industries wanting to take advantage of new and rapidly evolving capabilities. Many organizations are intrigued by one GenAI use case in particular: the in-house AI chatbot. Combining a publicly available large language model (LLM) with an organization's own private corpus of data, these AI chatbots can provide accurate and specific customer support, help increase staff productivity with AI-powered search and content-generation capabilities, and supply valuable insights from an organization's existing data.

One of the first steps in building an in-house AI chatbot is selecting an appropriate computing solution to back it. Much of the discussion around AI in the media emphasizes how resource-intensive it can be, and some might assume that GPUs are a necessity. In fact, if only a small number of people will use your organization's chatbot simultaneously, a server platform with a powerful CPU can get the job done. Such a solution also gives you the option of scaling up by adding GPUs if your needs grow over time.

To investigate how different hardware options might support an in-house AI chatbot utilizing an LLM sized for this configuration and retrieval augmented generation (RAG), we tested two configurations of a single-socket Dell™ PowerEdge™ R6615 server. One had only a 64-core AMD EPYC™ 9534 processor and the other had this same processor plus an NVIDIA® L4 GPU.

Our testing revealed that a small organization or a department within an enterprise could use this solution in a CPU-only configuration to adopt a chatbot that supports a minimum of 9 simultaneous users at an affordable entry price. And because it is unlikely that all employees of an organization would be using the chatbot at once, this solution would effectively support larger departments comfortably. We also found that by adding a GPU to the server, the organization could scale up to support 23 or more simultaneous users as its requirements expanded.

**Deploy an in-house chatbot on a server with only a CPU**  
Strong GenAI performance without a GPU

**Support up to 9 simultaneous chatbot users**  
with the majority getting complete responses in under 5 seconds

**Support more users with a clear and easy upgrade path**  
Sustain up to 23 simultaneous users by adding a GPU

Run your in-house AI chatbot on an AMD EPYC 9534 processor-powered Dell PowerEdge R6615 server February 2025

server. In the first phase of testing, the PowerEdge R6615 with only a 64-core AMD EPYC 9534 processor could support up to nine simultaneous chatbot users with a 5-second response time. In the second phase of testing, PT added a single NVIDIA L4 GPU to the Dell PowerEdge R6615 server and re-ran PTChatterly to measure any improvement the GPU might make. This configuration supported 23 simultaneous users with the same response time.

The report states, "Our testing revealed that a small organization or a department within an enterprise could use this solution in a CPU-only configuration to adopt a chatbot that supports a minimum of 9 simultaneous users at an affordable entry price. And because it is unlikely that all employees of an organization would be using the chatbot at once, this solution would effectively support larger departments comfortably. We also found that by adding a GPU to the server, the organization could scale up to support 23 or more simultaneous users as its requirements expanded."

To learn more, read the report at <https://facts.pt/MxYZe83>.

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](http://www.principledtechnologies.com).

Sharon Horton

Principled Technologies, Inc.

[press@principledtechnologies.com](mailto:press@principledtechnologies.com)

Visit us on social media:

[Facebook](#)

[X](#)

[LinkedIn](#)

[YouTube](#)

---

This press release can be viewed online at: <https://www.einpresswire.com/article/784748025>

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.

© 1995-2025 Newsmatics Inc. All Right Reserved.