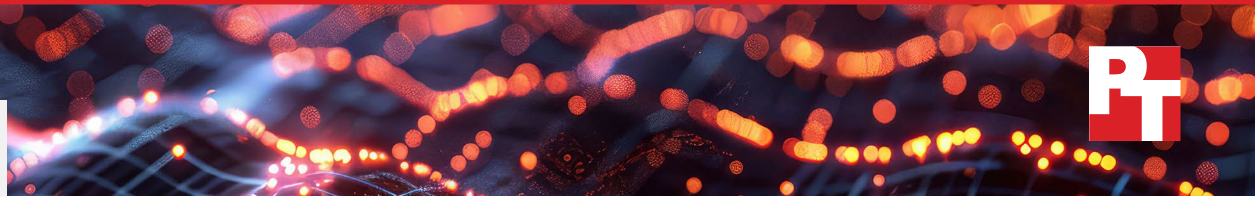

Executive summary

Dell PowerEdge R7615 servers with Broadcom BCM57508 NICs can accelerate your AI fine-tuning tasks

A cluster of Dell™ PowerEdge™ R7615 servers featuring AMD EPYC processors achieved much stronger performance on multi-GPU, multi-node operations using Broadcom 100GbE NICs than the same cluster using 10GbE NICss

Up to 83% less time to complete multi-GPU, multi-node operations*

Up to 66% lower latency on multi-GPU, multi-node operations*

Up to 6.1x the bandwidth on multi-GPU, multi-node operations*

*cluster of Dell PowerEdge R7615 servers featuring AMD EPYC 9374F processors and Broadcom 100GbE BCM57508 NetXtreme-E NICs vs. the same cluster with 10GbE NICs.

Organizations across industries are considering how they can use generative AI to improve operations. One way is to start with a public large language model (LLM) and fine-tune it with your own data to build your own in-house LLM. But what hardware should you choose for the resource-intensive task of training this model? Training an LLM typically requires the resources of many GPUs. One effective approach is to use a cluster of server nodes, each with its own set of GPUs, and spread the work across the distributed GPUs. In this environment, low latency and high bandwidth between GPUs become important.** We explored this approach by testing the performance of a two-node Dell cluster with two networking configurations: one with Broadcom® 100GbE BCM57508 NetXtreme-E network interface cards (NICs) with remote direct memory access (RDMA) over Ethernet (RoCE) support, and the other with Broadcom 10GbE BCM57414 NICs. The cluster comprised two Dell PowerEdge R7615 servers with AMD EPYC™ 9374F processors and NVIDIA® L40 GPUs.

LLM training and inference frameworks deployed on distributed GPUs use low-level algorithms to move data between GPUs, operate on that data, and share the results with other GPUs. Our testing focused on three of these algorithms as implemented in the NVIDIA Collective Communications Library (NCCL) library. This library, which many AI frameworks use, can send data over RoCE network paths or ordinary Ethernet network paths, and can perform RDMA transfers between distributed NVIDIA GPUs.

For each configuration, we studied three multi-GPU, multi-node AI computations from the NCCL test suite[1] at packet sizes ranging from 4 B to 256 MB and measured the time to complete the operation and the effective bandwidth of the network during the operation. This operational bandwidth is a combination of the very fast data transfer between GPUs on the same node, and the slower data transfer between GPUs on different nodes. Across this range of packet sizes and each of the three low-level AI operations, the cluster with 100GbE networking dramatically outperformed the cluster with 10GbE networking. Compared to the 10GbE networking configuration, the operational latency decreased by 26 percent to 67 percent, and the operational bandwidth was 3.7 to 6.1 times as high. In addition, the 100GbE cluster achieved these gains without increasing power usage.

**These tests do not send enough data between servers to overwhelm the networking link. Rather, these tests comprise a sequence of computational steps on each GPU, where a given step may require data from other GPUs. In such cases, a GPU can only start the next computational step once it has the data from those other GPUs, even if that data is as small as a single byte. The operational bandwidth depends on the timely transfer of data between GPUs on different servers.

What we found

We measured time to complete three AI fine tuning tasks—all-reduce, reduce-scatter, and send-receive—plus operational bandwidth and latency. We observed the same patterns across all three tasks: the configuration with 100GbE networking performed significantly better.

Time to complete tasks

Figure 1 shows performance results for the send-receive task. As the data packet size increased, the time the configuration with the 10GbE networking needed increased at a much faster rate than it did for the configuration with 100GbE networking. At the largest packet size, the configuration with 100GbE networking took 83 percent less time to complete. On the other two tasks, time to complete decreased by 82 percent (see the complete report).

Figure 1: Time in microseconds to complete the send-receive task on datasets of multiple sizes. Lower is better. Source: Principled Technologies.

Latency for multi-GPU, multi-node AI tasks

Using the 100GbE NIC improved latency by over 65 percent for the all-reduce and reduce-scatter tasks, and by 26.7 percent for the send-receive task (see Table 1).

Table 1: Latency for multi-GPU, multi-node tasks. Lower latency and higher improvement are better. Source: Principled Technologies.

Multi-GPU, multi-node operation	**Latency (microseconds)** Lower is better		**Percentage reduction**
	100GbE configuration	10GbE configuration	(Higher is better)
all-reduce (packet size: 4 B)	40	123	67.4%
reduce-scatter (packet size: 4 B)	29	85	65.8%
send-receive (packet size: 48 B)	41	56	26.7%

Bandwidth for multi-GPU, multi-node AI tasks

Figure 2 shows operational bandwidth for the send-receive task. The 100GbE configuration achieved 6 times the bandwidth of the 10GbE configuration. On the other two tasks, the 100GbE network configuration achieved 5 times the bandwidth (see the complete report).

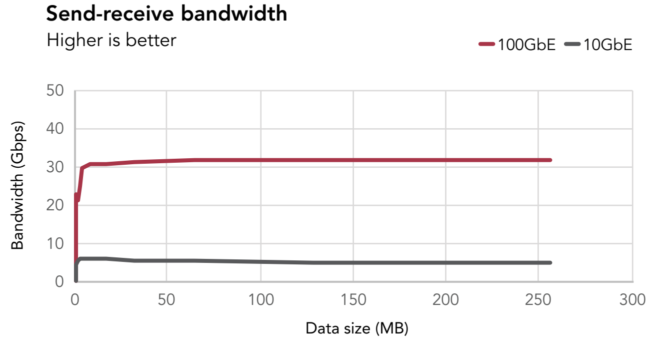

Figure 2: Bandwidth available for send-receive task. Higher is better. Source: Principled Technologies.

Power usage

AI workloads can have heavy power and cooling demands, making it important to selecting servers with increased power efficiency. We measured the power consumption of both servers during testing. Despite the great multi-GPU, multi-node AI task performance improvements the 100GbE Broadcom card enabled, power usage did not increase significantly with its use.

Conclusion

Using Broadcom 100GbE BCM57508 NICs and software in a cluster of two Dell PowerEdge R7615 servers with AMD EPYC processors and NVIDIA GPUs provided dramatically lower latency and greater bandwidth than using only 10GbE networking, with no increase in power usage.

1. NVIDIA, “NCCL Tests,” accessed November 16, 2024, https://github.com/NVIDIA/nccl-tests.

Read the report at https://facts.pt/QAauY1Y

Principled Technologies®

Facts matter.®