



A Dell oferece um portfólio mais amplo de opções de soluções, serviços e consumo para oferecer suporte às implantações de IA dos clientes



Desempenho até 37% melhor na metade do espaço do rack para processamento de linguagem natural (GPT-J), de acordo com os dados do MLPerf



Os Dell Validated Designs incluem mais arquiteturas de referência centradas em IA do que as soluções validadas da Supermicro

Encontrando o caminho para o sucesso da IA com o portfólio de IA da Dell

Uma comparação do portfólio de IA da Dell com ofertas semelhantes da Supermicro

Para implementar a inteligência artificial (IA) e colher benefícios, você deve considerar todos os aspectos de uma solução. É crucial encontrar uma parceria que ofereça suporte de ponta a ponta, desde o planejamento, a preparação de dados, a seleção do hardware, o design do modelo de IA, o teste da prova de conceito e as arquiteturas de referência até o suporte de ponta a ponta. Neste resumo, destacamos as vantagens de desempenho e espaço dos servidores nos portfólios de IA da Dell™ em relação aos da Supermicro, com base nos resultados do teste MLPerf® Benchmark disponíveis publicamente. O MLPerf® testa o desempenho tanto para treinamento quanto para inferência em vários modelos de IA. Os dados neste resumo vêm dos resultados do MLPerf® v3.1 Inference Datacenter publicados no site da MLCommons® em novembro de 2023.¹ Também comparamos informações disponíveis ao público sobre a variedade de produtos e soluções que cada fornecedor oferece para dar suporte às implantações de IA.

Servidores de quatro GPUs: os servidores Dell PowerEdge ofereceram um desempenho de MLPerf melhor do que os servidores Supermicro

Com até quatro GPUs NVIDIA H100 SXM, o Dell PowerEdge XE9640 2U oferece metade da potência computacional da GPU do PowerEdge XE9680, em um terço do espaço.² O servidor de quatro GPUs NVIDIA HGX H100 da Supermicro nas apresentações do MLPerf® é o SYS-421GU-TN XR, o qual é um servidor 4U. Nos resultados publicados do modelo de IA gptj-99,9, o PowerEdge XE9640 superou o servidor Supermicro nos testes off-line, alcançando até 1,37 vez a pontuação (consulte a Figura 1).

Também comparamos os resultados do servidor Supermicro SYS-421GU-TN XR com o Dell PowerEdge XE8640, um servidor de quatro GPUs 4U que também suporta GPUs NVIDIA H100 HGX. O Dell PowerEdge XE8640 apresenta os mais recentes processadores escaláveis Intel Xeon de 4ª geração e até 4 TB de memória para lidar com grandes conjuntos de dados e cálculos complexos, comuns em IA e lógica analítica de dados.³ Com o formato 4U, o PowerEdge XE8640 é semelhante ao Supermicro SYS-421GU-TN XR em capacidades de densidade e de GPU. No entanto, como observamos com o PowerEdge XE9640, o Dell PowerEdge XE8640 obteve melhores pontuações gptj-99 nos testes off-line do que o servidor Supermicro (consulte a Figura 2).

Resultados padronizados do MLPerf®: Dell PowerEdge XE9640 com H100 SXM5 versus Supermicro SYS-421GU-TNXR com H100 SXM5 (quanto maior, melhor)

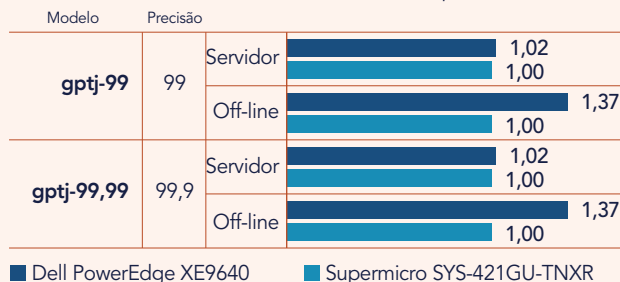


Figura 1: Resultados do MLPerf® publicados para o Dell PowerEdge XE9640 e o Supermicro SYS-421GU-TNXR em 29/11/23. Os dois sistemas usam GPUs NVIDIA H100 no formato SXM. Fonte: Principled Technologies usando os dados da MLCommons®.*

Resultados padronizados do MLPerf®: Dell PowerEdge XE8640 com H100 SXM5 versus Supermicro SYS-421GU-TNX com H100 SXM5 (quanto maior, melhor)

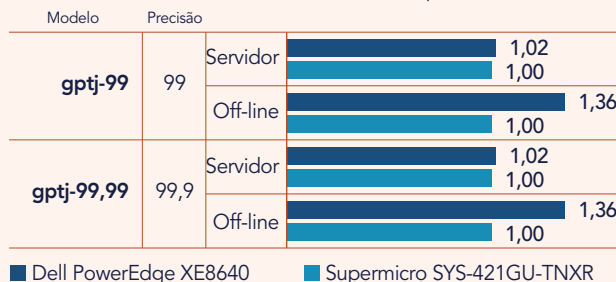


Figura 2: Resultados do MLPerf® publicados para o Dell PowerEdge XE8640 e o Supermicro SYS-421GU-TNXR em 29/11/23. Os dois sistemas usam GPUs NVIDIA H100 no formato SXM. Fonte: Principled Technologies usando os dados da MLCommons®.*

Além do hardware: exame do panorama geral

As organizações que buscam flexibilidade nas ofertas de estações de trabalho de GPU podem encontrar uma variedade maior de opções no portfólio de IA da Dell. O maior portfólio de IA da Dell abrange mais do que apenas hardware; ele oferece aos clientes uma solução única de serviços e suporte para cada estágio da implantação. Em contrapartida, o portfólio de IA da Supermicro oferece serviços limitados a manuais, solução de problemas e devoluções de produtos. Na nossa pesquisa, não conseguimos encontrar serviços de design, implementação, gerenciamento ou ensino no portfólio de IA da Supermicro. Uma vez que a adoção da IA é tão complexa, os serviços que a Dell oferece a tornam uma parceira de treinamento mais promissora para a implantação dessas cargas de trabalho. Além disso, descobrimos que os Dell Validated Designs incluem mais arquiteturas de referência centradas em IA do que as soluções validadas da Supermicro e oferecem orientações mais detalhadas.

Leia nosso relatório completo, que explora detalhadamente os benefícios que o portfólio de IA da Dell oferece em opções de computação, armazenamento e sistema de rede; serviços profissionais para IA; parcerias com terceiros para cargas de trabalho de IA; Dell Validated Designs, que ajudam a eliminar as suposições das soluções de hardware de IA; e serviços de gerenciamento e iDRAC.

*Pontuação encerrada da inferência v3.1 verificada pelo MLPerf. Recuperado de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 de dezembro de 2023, entradas 3.1-0133, 3.1-0066 e 3.1-0067. O nome e logotipo MLPerf são marcas registradas e não registradas da MLCommons. Association nos Estados Unidos e em outros países. Todos os direitos reservados. Uso não autorizado estritamente proibido. Acesse www.mlcommons.org para obter mais informações.

1. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results", acessado em 7 de fevereiro de 2024, <https://mlcommons.org/en/inference-datacenter-31/>.
2. Dell, "PowerEdge XE9640 Rack Server", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
3. Dell, "PowerEdge XE8640 Rack Server", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.

► Visualize a versão original em inglês do relatório em <https://facts.pt/kRTwky3>

Clique aqui para ler o relatório ►



Facts matter.®

Principled Technologies é uma marca registrada da Principled Technologies, Inc. Todos os outros nomes de produtos são marcas comerciais de seus respectivos proprietários. Para obter mais informações, consulte o relatório.