

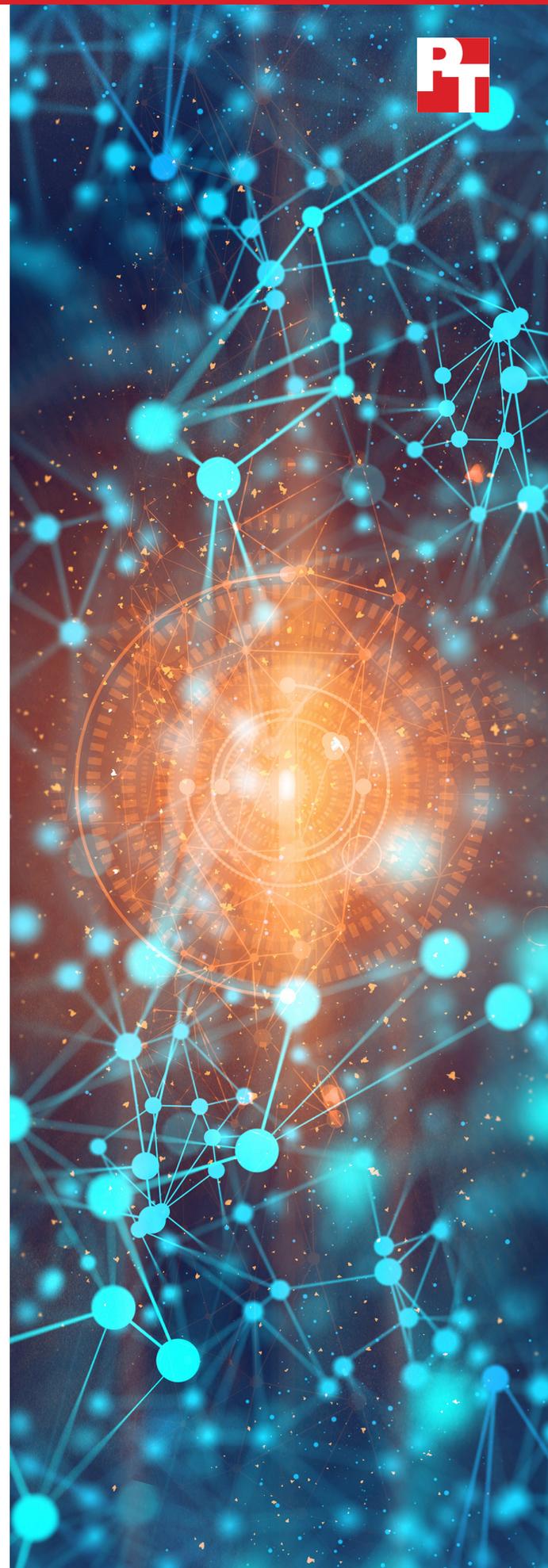


Encontrando o caminho para o sucesso da IA com o portfólio de IA da Dell

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

A inteligência artificial (IA) é a nova fronteira, pronta para remodelar as operações de negócios em todos os setores. Embora organizações de todo o espectro explorem como aproveitar a IA para aprimorar as operações de negócios, é importante lembrar que não é da noite para o dia que se implementa a IA e se colhe seus benefícios. Como cada empresa é única, cada uma deve avaliar seus dados e metas de negócios para ver como o uso da IA com seus dados pode gerar os resultados específicos desejados. Envolver-se com uma empresa como a Dell, que oferece um portfólio abrangente de IA, incluindo planejamento, preparação de dados, seleção de hardware apropriada, design de modelo de IA, teste de prova de conceito, arquiteturas de referência e suporte ponta a ponta, pode se converter em projetos de IA bem-sucedidos.

Com as inúmeras opções disponíveis no mercado, encontrar um parceiro que possa ajudar com todas essas decisões pode significar a diferença entre uma implementação bem-sucedida de IA e um erro dispendioso. Neste documento, analisamos em detalhes os portfólios de IA da Dell e da Supermicro com o objetivo de informar aos leitores as vantagens que a Dell pode oferecer durante toda a jornada de IA dos clientes. Primeiro, focaremos nas opções de servidor e computação, em que cada empresa tem ofertas grandes e variadas para os clientes. Em seguida, exploraremos como a Dell vai além das considerações de hardware para empresas que buscam ensino, serviços de planejamento, ecossistemas de parceiros e muito mais.



Resultados de servidores e desempenho para as cargas de trabalho de IA

Os servidores, que são a base da infraestrutura de TI para as cargas de trabalho de IA, podem usar CPUs, GPUs ou ambas como recursos computacionais, dependendo do tamanho ou tipo de carga de trabalho. Para cargas de trabalho maiores ou com mais demanda, como HPC ou IA, as GPUs oferecem desempenho superior. As GPUs vêm em uma variedade de formatos, incluindo PCIe universal, Open Compute Project Accelerator Module (OAM) e arquitetura NVIDIA SXM exclusiva, que atualmente oferece um desempenho superior.¹ Grandes capacidades de memória e recursos de design de servidor, como arquitetura de resfriamento e eficiência no consumo de energia, também afetam o desempenho. A maioria dos data centers ainda usa resfriamento a ar, ou seja, as cargas de trabalho de IA precisam de servidores projetados para resfriar com ar da maneira mais eficaz possível. Abaixo, destacamos as ofertas de servidor Dell PowerEdge em termos de componentes e opções de resfriamento, entre outros, além das pontuações publicadas pelo MLPerf[®] da MLCommons[®].

Resultados do teste

O MLPerf[®] é um conjunto de referências de desempenho que testa o desempenho de IA, tanto para treinamento quanto para inferência. Para que uma organização publique resultados oficiais do MLPerf[®], os resultados precisam estar em conformidade com condições específicas, definidas pela desenvolvedora da referência de desempenho, a MLCommons[®].² Essas diretrizes de conformidade apresentam padrões que facilitam a comparação do desempenho. Para os testes de inferência, o MLPerf[®] usa os conjuntos de dados "Datacenter", "Edge", "Mobile" e "Tiny" e relata as pontuações de IA e os watts de energia consumida durante os testes. O conjunto de referências de desempenho de inferência inclui testes para diversos modelos comuns de IA, ML e DL (consulte a Tabela 1).

Tabela 1: Modelos de IA, ML e DL incluídos no MLPerf[®], com testes e casos de uso típicos para cada um. Fonte: Principled Technologies.

Modelos comuns de IA	Casos de uso comuns
ResNet	Um modelo de classificação de imagens que ajuda os computadores a aprender, lembrar e identificar imagens diferentes para casos de uso como imagiologia médica, moderação de conteúdo de mídias sociais e reconhecimento facial
RetinaNet	Um tipo de detecção de objeto que pode lidar com mais complexidades que o ResNet. Ele ajuda os computadores a identificar e localizar objetos dentro de imagens ou frames de vídeo e pode classificá-los por importância. Para casos de uso como direção autônoma, tecnologia de assistência automática para veículos, monitoramento de segurança e reconhecimento facial
3D-UNet	Específico para segmentação de imagens médicas
RNN-T	Reconhecimento de voz para casos de uso como tradução automática de idiomas
BERT	Processamento de linguagem natural para casos de uso como resumo de texto, tradução de idiomas e realização automática de tarefas
DLRM-v2-99.9	Modelo de recomendação para casos de uso como anúncios direcionados e recomendações personalizadas de produtos
GPTJ-99 e 99.9	LLM para processamento de linguagem natural, que se destaca na geração de texto. Para casos de uso como chatbots e ferramentas de IA baseadas em chat



Sobre o MLPerf

Os resultados do MLPerf® incluem vários parâmetros, além dos próprios modelos de IA, o que pode fazer com que muitos dados sejam analisados em um só gráfico ou em uma só tabela. Veja abaixo uma referência rápida sobre esses parâmetros:

- 99,0 e 99,9: esses números referem-se à precisão com a qual o modelo foi treinado. Quanto maior a necessidade de precisão do resultado, mais complexo será o modelo e mais tempo ele poderá levar para processar os dados.
- Offline samples/sec: modo em que a referência de desempenho envia todas as consultas no início do teste, simulando dados já presentes no sistema.
- Server queries/sec: modo em que a referência de desempenho envia consultas durante todo o teste, simulando a análise de um fluxo de dados em tempo real.

Para obter mais informações sobre a MLCommons® e os resultados do MLPerf®, consulte <https://mlcommons.org/benchmarks/inference-datacenter/>.

As informações neste relatório são provenientes dos resultados de novembro de 2023 do MLPerf® v3.1 Inference Datacenter, publicados no site da MLCommons®.³ Esses resultados incluem envios de fabricantes de tecnologia e provedores de serviços em nuvem e abrangem diversas configurações. Em comparação com as apresentações publicamente disponíveis da Supermicro, os servidores Dell PowerEdge produziram resultados comparáveis. A Tabela 2 fornece os detalhes do servidor.

Tabela 2: Servidores Dell e Supermicro incluídos nos resultados do MLPerf® 3.1 da MLCommons® publicados em novembro de 2023. Fonte: Principled Technologies.

Remetente	Modelo de servidor	Nº e modelo de GPUs	Descrição
Dell ⁴	PowerEdge XE9680	8 NVIDIA H100 SXM	Para inferência e treinamento de IA com amplas cargas de trabalho, como grandes modelos de linguagem
	PowerEdge XE9640	4 NVIDIA H100 SXM	Para treinamento de grandes modelos de IA em datacenters de alta densidade e resfriados por líquido
	PowerEdge XE8640	4 NVIDIA H100 SXM	Para orientar aplicativos tradicionais de treinamento de IA, HPC e lógica analítica de dados, em um formato 4U, para data centers resfriados a ar
Supermicro ⁵	AS-8125GS-TNHR	8 NVIDIA H100 SXM	Para treinamento em IA de larga escala e cargas de trabalho HPC com processadores AMD
	SYS-821GE-TNHR	8 NVIDIA H100 SXM	Para treinamento em IA em larga escala e cargas de trabalho de HPC com processadores Intel
	SYS-421GU-TNXR	4 NVIDIA H100 SXM	Design modular para flexibilidade para suportar cargas de trabalho de HPC e IA

Como a Dell e a Supermicro apresentaram resultados com configurações de GPU semelhantes, a comparação do desempenho é simples. Como mostram as Figuras 1 e 2, a compatibilidade entre as configurações de ambos os fornecedores leva a resultados que são, em grande parte, iguais. Para outras configurações, como as mostradas nas Figuras 3 e 4, a Dell superou a Supermicro no modelo gptj-99,9 nos testes de 4 GPUs. Observe que, embora a Dell tenha apresentado resultados para todos os modelos disponíveis com todos os três servidores, a Supermicro não. Comparamos apenas os modelos para os quais ambos os servidores têm resultados. Para ver o conjunto completo de resultados da Dell, visite os resultados do MLPerf® da MLCommons®.

Resultados do servidor para oito GPUs

O Dell PowerEdge XE9680 oferece suporte para até oito GPUs NVIDIA H100 SXM5 para aceleração de IA e para até dois processadores escaláveis Intel® Xeon® de 4ª geração. A família de produtos PowerEdge XE tem uma arquitetura modular que suporta GPUs NVIDIA SXM4 ou SXM5 ou os conjuntos de GPU Open Compute Project Accelerator Module (OAM), que podem aumentar o desempenho em comparação com uma GPU PCIe padrão. O Dell PowerEdge XE9680 também oferece o acelerador AMD Instinct™ MI300X.⁶ Ocupando apenas 6U de espaço em rack, o PowerEdge XE9680 é um compacto servidor NVIDIA H100 SXM5 de oito vias.

Por outro lado, o servidor Supermicro de 8 GPU requer 33% mais espaço em rack em 8U, ao mesmo tempo em que oferece o formato SXM para GPUs NVIDIA. A diferença de tamanho significa que você pode encaixar sete servidores Dell PowerEdge XE9680 em um rack, em comparação com apenas cinco servidores Supermicro. Nas Figuras 1 e 2, comparamos os resultados do Dell PowerEdge XE9680 com os de duas configurações do servidor Supermicro 8-GPU: o SYS-821GE-TNHR com processadores Intel e o AS-8125GS-TNHR com processadores AMD. Observe que a Supermicro não apresentou resultados para RNN-T no SYS-821GE-TNHR, portanto excluímos esse modelo do gráfico na Figura 1.

Resultados padronizados do MLPerf® Dell PowerEdge XE9680 com H100 SXM5 versus Supermicro SYS-821GE-TNHR com H100 (quanto maior, melhor)

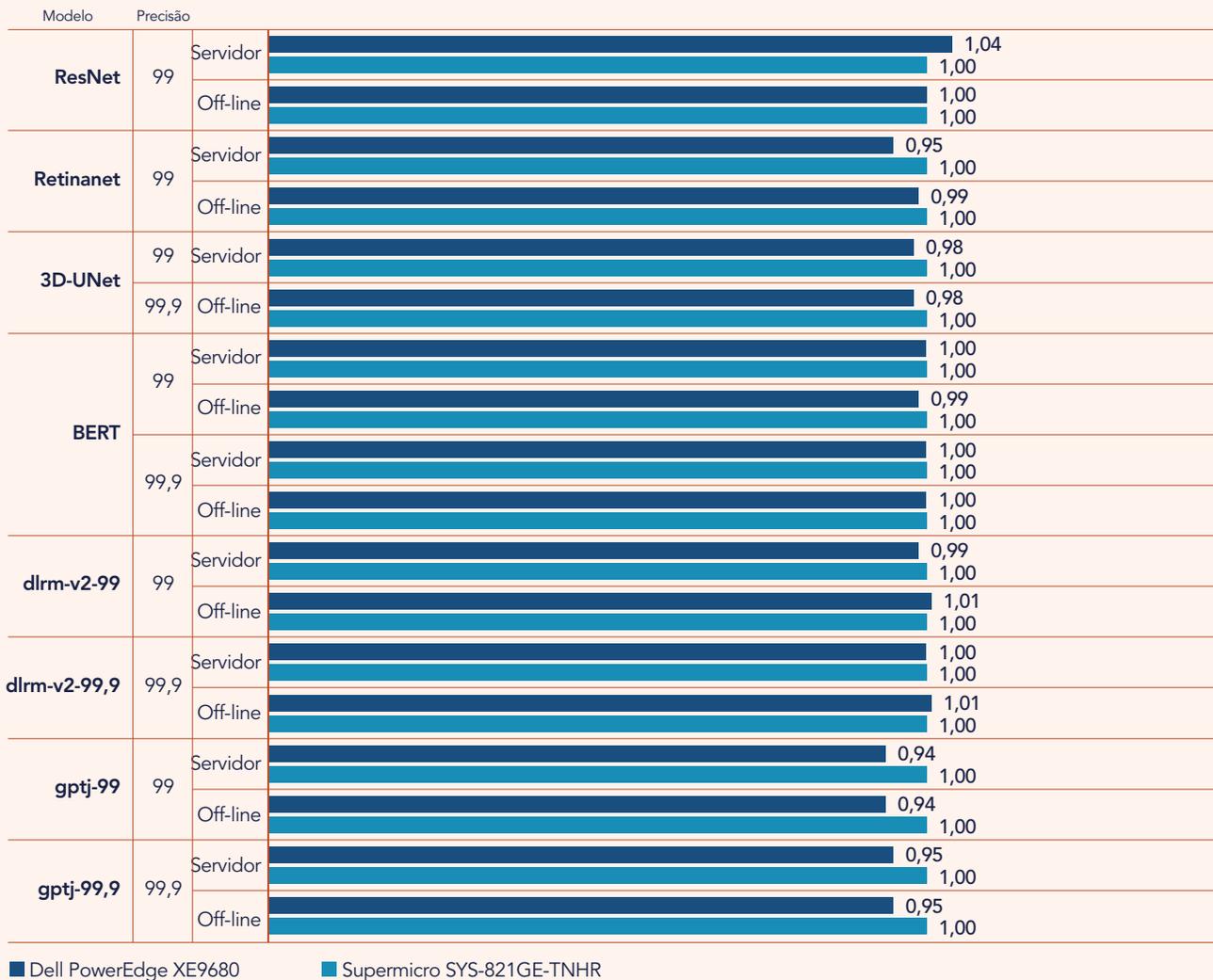


Figura 1: Resultados do MLPerf® publicados para o Dell PowerEdge XE9680 e o Supermicro SYS-821GE-TNHR em 29/11/23. Os dois sistemas usam GPUs NVIDIA H100 no formato SXM. Fonte: Principled Technologies com dados da MLCommons®.^{7,8}



Resultados padronizados do MLPerf® Dell PowerEdge XE9680 com H100 SXM5 versus Supermicro AS-8125GS-TNHR com H100 SXM5 (quanto maior, melhor)

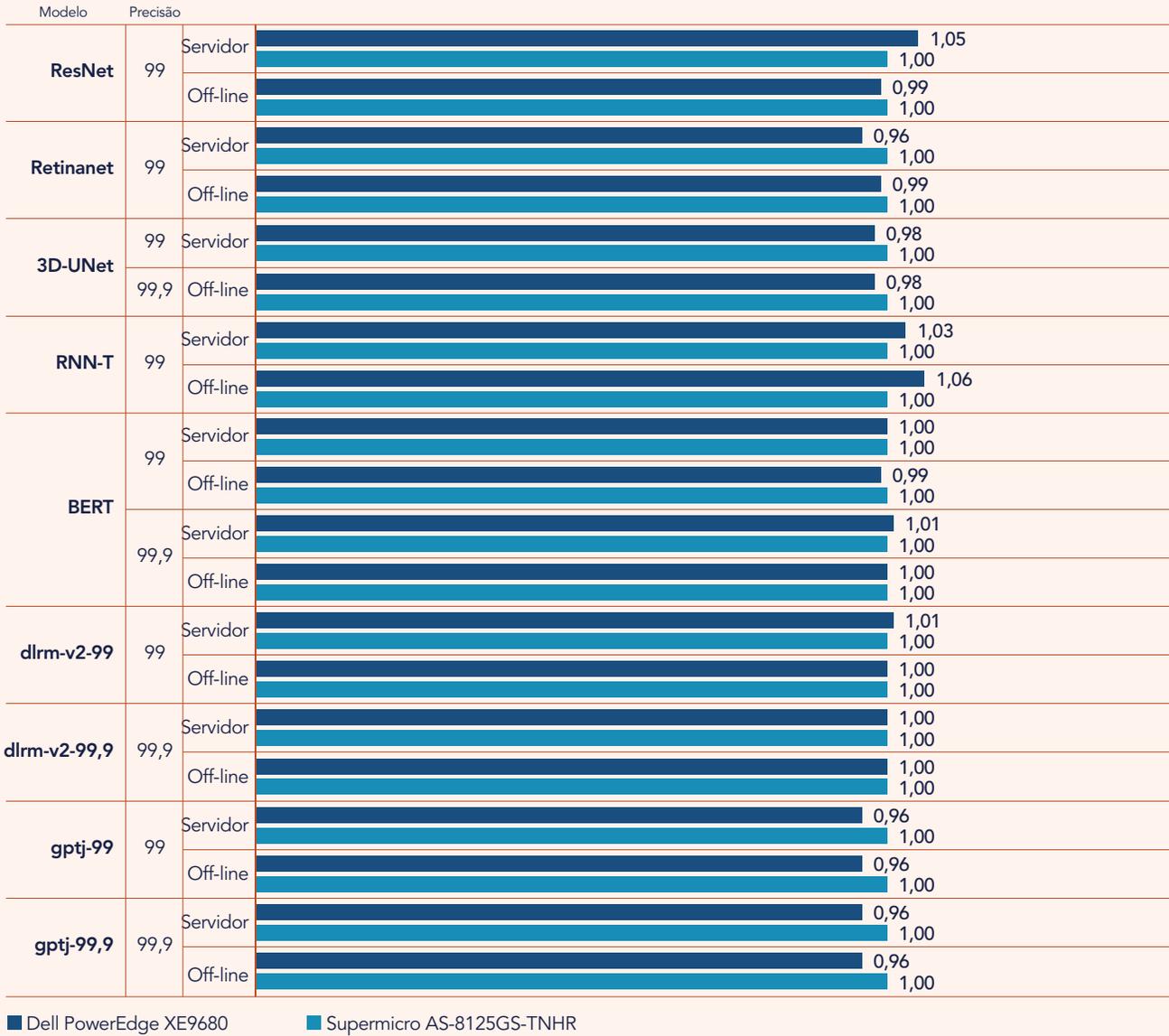


Figura 2: Resultados do MLPerf® publicados para o Dell PowerEdge XE9680 e o Supermicro AS-8125GS-TNHR em 29/11/23. Os dois sistemas usam GPUs NVIDIA H100 no formato SXM. Fonte: Principled Technologies com dados da MLCommons® 9,10

Como esses resultados mostram, a escolha do Dell PowerEdge XE9680 proporcionará um desempenho semelhante em uma infinidade de cargas de trabalho de inferência de IA, ao mesmo tempo em que ocupa menos espaço no data center.



Resultados do servidor para quatro GPUs

Quando a principal preocupação é minimizar o uso de energia ou o espaço do data center, o Dell PowerEdge XE9640 2U pode ser a resposta. Com até quatro GPUs NVIDIA H100 SXM, o PowerEdge XE9640 oferece metade da potência computacional da GPU do PowerEdge XE9680, em dois terços do espaço.¹¹ O servidor Dell PowerEdge XE9640, densamente compactado, incorpora a tecnologia Dell Smart Cooling, disponibilizando uma variedade de tecnologias térmicas, incluindo o resfriamento líquido direto para CPUs e GPUs.¹² O chassi 2U do PowerEdge XE9640 acomoda mecanismos aprimorados de fluxo de ar, incluindo dissipadores de calor e ventiladores maiores, para ajudar a resfriar os outros componentes vitais, como placas PCIe e memória.¹³

A Supermicro oferece um servidor mais antigo, o SYS-220GQ-TNAR+, que fornece quatro GPUs NVIDIA A100 HGX em um formato de 2U, mas não foi possível encontrar um servidor Supermicro de 2U com quatro das mais recentes GPUs HGX H100 que corresponderiam ao PowerEdge XE9640.¹⁴ O servidor 4-GPU com GPUs NVIDIA HGX H100 da Supermicro nas apresentações do MLPerf[®] é o SYS-421GU-TNXR, que é um servidor de 4U. Como mencionamos anteriormente, a Supermicro apresentou os resultados do MLPerf[®] 3.1 apenas para o SYS-421GU-TNXR no modelo de IA gptj-99,9; portanto, não podemos compará-los com o PowerEdge XE9640 nos outros modelos. No entanto, nos resultados publicados, o PowerEdge XE9640 superou o servidor Supermicro nos testes off-line, alcançando até 1,37 vez a pontuação (consulte a Figura 3).

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



Figura 3: 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 com dados da MLCommons[®].^{15,16}

Também podemos comparar os resultados do servidor Supermicro SYS-421GU-TNXR com o Dell PowerEdge XE8640, um servidor de 4U 4-GPU que também suporta GPUs NVIDIA H100 HGX. Embora seja maior do que o PowerEdge XE9640, o PowerEdge XE8640 não requer resfriamento líquido direto, tornando-se uma transigência entre as tecnologias de densidade e resfriamento para data centers que não têm acesso ao resfriamento com água. O PowerEdge XE8640 tem resfriamento a ar para processadores e um radiador de resfriamento a ar auxiliado por líquido para GPUs, que não exige disponibilidade de água para rack nas instalações.¹⁷ 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.¹⁸ Com o formato 4U, o PowerEdge XE8640 é semelhante ao Supermicro SYS-421GU-TNXR em capacidades de densidade e de GPU. No entanto, como observamos com o PowerEdge XE9640, o Dell PowerEdge XE8640 obtém melhores pontuações gptj-99 nos testes off-line do que o servidor Supermicro (Figura 4).

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

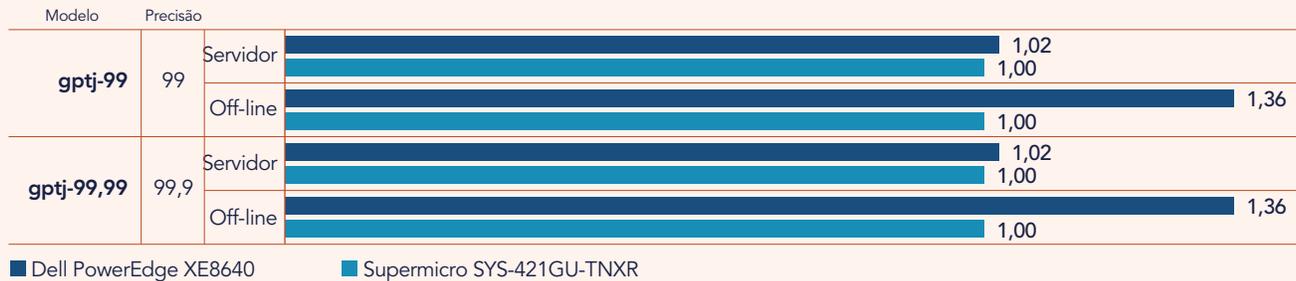


Figura 4: Resultados do MLPerf® publicados para o Dell PowerEdge XE8640 e o Supermicro SYS-421GU-TNXR em 29/11/23. Os dois sistemas usam GPUs NVIDIA H100 no formato SXM. Fonte: Principled Technologies com dados da MLCommons®.^{19,20}

Como observamos, os resultados de desempenho do MLPerf® para as ofertas de servidores baseados em GPU da Supermicro e da Dell são semelhantes em geral, com uma vantagem notável para os servidores Dell PowerEdge XE8640 e PowerEdge XE9640 em um modelo de IA.

Uma vez que o desempenho é apenas um aspecto da jornada de implementação da IA, também analisamos as demais ofertas de IA da Dell e da Supermicro, desde estações de trabalho, armazenamento e sistema de rede até serviços, suporte, ensino e muito mais. Descobrimos que o portfólio de IA da Dell é mais amplo do que a oferta da Supermicro e oferece soluções para muitos dos obstáculos que as empresas enfrentam além do desempenho de computação.

Estações de trabalho do cliente e ofertas de armazenamento

Opções de computação adicionais com estações de trabalho

Alguns casos de uso de IA exigem uma abordagem de computação diferente, e nem todos que precisam de acesso ao hardware com capacidade de IA podem permanecer vinculados ao data center. Cientistas que trabalham em laboratórios podem não ter espaço para um rack de servidores, e é impraticável que as pessoas que trabalham na borda se comprometam com um sistema de desktop grande e pesado. É aqui que as estações de trabalho entram em ação. O portfólio de IA da Dell inclui várias estações de trabalho Precision prontas para IA, incluindo configurações de torre, móvel e rack, para atender a uma variedade de necessidades.²¹ Em contrapartida, a Supermicro não oferece estações de trabalho móveis para acesso fácil e móvel. Suas ofertas de estação de trabalho de GPU consistem em várias configurações de torre diferentes, algumas das quais a Supermicro alega que os usuários podem montar em rack.²² Nossa pesquisa revelou duas estações de trabalho montáveis em rack, mas elas parecem mais antigas, disponíveis apenas com GPUs NVIDIA A100 e possivelmente não estão mais à venda.²³ Se a sua empresa precisar de flexibilidade no tipo e na mobilidade das estações de trabalho GPU que implanta, o portfólio de IA da Dell é mais adequado para essas necessidades.

Considerações sobre armazenamento

O armazenamento pode ser tão vital quanto a computação ao executar cargas de trabalho de IA. A maior quantidade de dados melhora a precisão do modelo de IA. No entanto, o armazenamento e o gerenciamento de grandes conjuntos de dados podem desafiar as capacidades de muitos data centers. Além disso, como os modelos geralmente treinam usando dados não estruturados, os sistemas de armazenamento prontos para IA precisam manusear muitos tipos de dados diferentes com facilidade.²⁴ Para oferecer capacidade e dimensionamento para conjuntos de dados de IA, ML e DL, a Dell oferece a série PowerScale™ para armazenamento em arquivo e o Elastic Cloud Storage (ECS) ou ObjectScale definido por software para armazenamento em objeto.

O portfólio NAS All-Flash do Dell PowerScale oferece opções de capacidade bruta por nó de 3,84 TB até 720 TB e opções All-Flash em cluster com capacidade bruta de até 186 PB. A flexibilidade e a escala do PowerScale podem aceitar uma ampla variedade de clientes e casos de uso de IA.²⁵ Todos os três modelos do PowerScale All-Flash (F200, F600 e F900) incluem desduplicação e compactação de dados em linha para melhorar a eficiência de armazenamento.²⁶ Cada modelo de armazenamento PowerScale usa o sistema de arquivos Dell OneFS™, que aproveita as políticas avançadas em camadas para garantir que os dados acessados com mais frequência residam nos níveis de armazenamento de desempenho mais alto.²⁷ A Dell também oferece o software OneFS no Amazon Web Services (AWS) Marketplace com o Dell APEX File Storage for AWS. Os clientes podem aproveitar o OneFS com suas instâncias de computação do AWS para obter uma experiência consistente do usuário, com os mesmos recursos disponíveis nos arrays OneFS locais.²⁸

As ofertas de armazenamento da Supermicro consistem em servidores de armazenamento de dados montados em rack, densos e de vários tamanhos e densidades.²⁹ Para obter o armazenamento em arquivos da Supermicro, os clientes devem escolher entre várias ofertas de armazenamento definidas por software de terceiros, como WekaIO, Scarity RING ou OSNEXUS.³⁰ Embora o Scarity RING e o OSNEXUS incluam opções de armazenamento em arquivos como parte de suas descrições de plataforma, o WekaIO parece ser a principal opção para os clientes que procuram um armazenamento em arquivos básico. A Supermicro oferece várias arquiteturas de referência que abrangem uma ampla variedade de casos de uso, mas os clientes devem ter uma assinatura ou licença de software do WekaIO, que pode aumentar o custo geral da solução.³¹

As opções de armazenamento em objeto da Dell incluem o Dell ECS Enterprise Object Storage, que foi “criado especificamente para armazenar dados não estruturados em escala de nuvem pública”.³² Juntamente com a compatibilidade integrada com o armazenamento em objeto do Amazon S3 para funcionalidade de nuvem híbrida, os nós de armazenamento do ECS oferecem capacidades de até 14 PB por rack.³³ Assim como no armazenamento em arquivos, as ofertas de armazenamento em objetos da Supermicro exigem configurações de terceiros. A plataforma OSNEXUS é uma plataforma de armazenamento combinada para arquivos, blocos e objetos, enquanto a solução Scarity RING combina armazenamento em arquivos e objetos; ambas requerem licenças com fornecedores terceirizados.^{34,35} Para armazenamento apenas em objetos, os clientes podem comprar a solução da Supermicro para o Quantum ActiveScale para armazenamento em objetos em nuvem privada com uma assinatura do software Quantum.³⁶ Até o momento da redação deste documento, não conseguimos encontrar opções flexíveis de consumo/despesas operacionais oferecidas pela Supermicro.

Como as ofertas de armazenamento da Supermicro exigem que os clientes se envolvam com fornecedores de software de terceiros, é provável que eles incorram em custos adicionais de licenciamento ou assinatura e possam enfrentar dificuldades de suporte, solução de problemas e muito mais. O portfólio de IA da Dell oferece aos clientes de armazenamento de dados da Dell uma solução única e confiável de serviços e suporte em todos os aspectos da solução de armazenamento de dados.

Sobre a Dell APEX

Para clientes que desejam consumir seu armazenamento em arquivos "as a service", a Dell oferece o APEX Data Storage Services, que inclui armazenamento de arquivos, blocos e backup. Usando o Dell APEX Console, os clientes podem solicitar novas assinaturas, ajustar e monitorar a capacidade de armazenamento e muito mais. De acordo com a Dell, essa solução permite que você “obtenha a facilidade e a agilidade da experiência em nuvem com mais controle dos aplicativos e dados”.³⁷

Para saber mais sobre o Dell APEX, acesse <https://www.dell.com/en-us/dt/apex/storage/data-storage-services/index.htm>



Opções de rede

O sistema de rede é outro componente vital da infraestrutura de IA. Com muitas cargas de trabalho de IA em execução em grandes clusters de servidores, que exigem comunicação constante entre si e com o armazenamento, as cargas de trabalho de IA precisam de um sistema de rede robusto para evitar gargalos. Se o sistema de rede for insuficiente para a carga de trabalho de IA, os tempos de treinamento e inferência aumentarão, desacelerando o processamento de dados e o tempo para obter insights. A Dell oferece switches topo de rack (Tor) PowerSwitch Data Center e módulos de E/S PowerEdge MX para redes Ethernet e de fabric.³⁸ As ofertas do PowerSwitch variam de 1GbE a 400GbE para atender a uma variedade de necessidades. Além disso, os switches Dell PowerSwitch Série Z oferecem conexões de 100GbE e 400GbE otimizadas para fabrics de leaf/spine.³⁹

A Supermicro também oferece switches com portas Ethernet de até 400GbE para Tor e outras aplicações, como Data Center Spine e leaf.⁴⁰ No entanto, os serviços Dell Networking oferecem vários benefícios de facilidade de uso e flexibilidade que a Supermicro não oferece. Serviços como o Dell Fabric Design Center podem ajudar a evitar discrepâncias de sistema de rede, lacunas ou ineficiências, ajudando os clientes a planejar e implantar fabrics de rede com automação.⁴¹ Para ambientes específicos, como VMware VxRail, VMware ESXi e configurações do Dell PowerStore, a Dell oferece o SmartFabric Services, que permite a implantação da infraestrutura definida por software e gerenciamento do ciclo de vida. Com o PowerStore, os SmartFabric Services podem automatizar "até 99% das tarefas de conectividade de LAN com um fabric plug-and-play".⁴² Serviços como esses, fornecendo automação, orientação e muito mais para design e implementação do sistema de rede, dão suporte aos clientes à medida que eles navegam pelo processo de adoção da IA.

Serviços, treinamento e muito mais

O principal desafio não relacionado ao hardware na implementação da IA é a necessidade de expertise interna para estratégia, planejamento, preparação de dados e gerenciamento. O gerenciamento e a manutenção de cargas de trabalho de IA exigem um conjunto exclusivo de conhecimento, incluindo expertise em hardware mais tradicional e operações de aprendizado de máquina e ciência de dados. As pessoas que projetam e implementam a estratégia de IA também precisam ter um profundo conhecimento das metas operacionais exclusivas da empresa para garantir que as novas cargas de trabalho de IA as atendam.⁴³

Outro obstáculo significativo pode ser integrar a IA de modo ininterrupto aos sistemas operacionais existentes. Essa integração exige um alinhamento estratégico das novas tecnologias de IA com os processos de negócios atuais, garantindo que a introdução da IA mantenha os fluxos de trabalho estabelecidos intactos. A parceria com uma empresa como a Dell, que fornece várias arquiteturas de referência de solução otimizadas e validadas, cursos de treinamento, opções de gerenciamento e um grande rede de parceiros, pode facilitar a jornada de adoção da IA.

Serviços profissionais para AI

Para cobrir ensino e o planejamento, a Dell oferece uma variedade de serviços especificamente para IA.⁴⁴ Os serviços da Dell que dão suporte à implementação de IA abrangem consultoria, preparação de dados, implantação, suporte e ensino, cada qual direcionado a aspectos específicos da adoção da IA. Os Serviços de consultoria da Dell para IA generativa ajudam os clientes a criar um roteiro, identificando casos de uso e ajudando as empresas a simplificar processos.⁴⁵ Da mesma forma, os Adoption Services for Generative AI fornecem workshops com profissionais da Dell para analisar suas necessidades e desafios exclusivos e determinar um modelo pré-treinado para sua empresa, além de realizar sessões de transferência de conhecimento para treinar sua equipe de TI.⁴⁶ A Dell também oferece serviços de implementação, dimensionamento e gerenciamento para IA generativa que fornecem níveis variados de suporte e treinamento, até uma infraestrutura de IA totalmente gerenciada, que permite que sua equipe de TI se concentre nos modelos e dados enquanto a Dell gerencia o hardware.⁴⁷ Os serviços ProSupport garantem o desempenho ideal do sistema e fornecem assistência essencial em hardware e software para operações contínuas de IA, abordando problemas técnicos.⁴⁸

Os serviços de ensino são essenciais para promover as habilidades e o conhecimento necessários para a utilização da IA. As ofertas de treinamento da Dell incluem programas abrangentes de treinamento em ciência de dados, certificações de lógica analítica e workshops sobre tecnologias de IA específicas, como aprendizado de máquina.⁴⁹

Já os serviços da Supermicro estão, em sua maioria, limitados a solução de problemas, manuais, autorizações de devolução de mercadorias (RMAs) e garantias.⁵⁰ Não foi possível encontrar qualquer design, implementação, gerenciamento ou serviço educacional no portfólio de IA da Supermicro. Para empresas que buscam um parceiro de treinamento para lidar com as complexidades da adoção da IA, a Dell é a escolha clara entre essas duas opções.

Parcerias de terceiros para cargas de trabalho de IA

A Dell Technologies e a NVIDIA colaboram para oferecer os Dell Validated Designs, com o objetivo de fornecer uma solução abrangente para IA generativa em ambientes de negócios. Esse projeto cria uma infraestrutura escalável e de alto desempenho com base nas tecnologias e no software Dell e NVIDIA, juntamente com uma estrutura de modelo de IA que permite às empresas criar e executar modelos de IA personalizados. A solução permite que os clientes executem cargas de trabalho de IA generativa rapidamente.⁵¹ Para saber mais, leia a seção Dell Validated Designs abaixo.

A Dell fez uma parceria com várias empresas para aprimorar as aplicações da tecnologia de IA. Com o Hugging Face, a Dell facilita a configuração de modelos de linguagem grande (LLMs) no local. Essa parceria combina a expertise em IA do Hugging Face com os servidores e sistemas de armazenamento da Dell. Um portal de Hugging Face específico da Dell fornecerá ferramentas para a implantação direta e segura de modelos de IA de código aberto do Hugging Face. O objetivo contínuo é continuar aprimorando esses modelos para os sistemas Dell, aumentando o desempenho e oferecendo suporte a novas aplicações de IA.⁵²

A Dell e a Starburst estão trabalhando em um data lake escalável e de alto desempenho que integra a lógica analítica da Starburst à tecnologia de computação e armazenamento da Dell, buscando oferecer um ponto único de acesso a todas as fontes de dados para ferramentas de IA e ML. Os clientes poderão aproveitar essa parceria para ajudar a eliminar silos de dados.⁵³

De acordo com a nossa pesquisa, a Supermicro tem parcerias muito mais limitadas para IA. A SIMA.ai e a Supermicro colaboraram para desenvolver o Supermicro SYS-E300-13AD, um servidor Edge ML compacto projetado para processamento de lógica analítica de vídeo de múltiplos fluxos. Esse servidor, equipado com o pipeline SIMA.ai ML em um chip, lida eficientemente com vários canais de vídeo, reduz o custo total de propriedade e melhora a confiabilidade e a segurança. O servidor oferece uma configuração computacional projetada para o processamento e a análise de vários fluxos de vídeo, fornecendo inteligência de ponta adequada para várias aplicações empresariais.⁵⁴

Dell Validated Designs

Para eliminar as suposições das soluções de hardware de IA, a Dell oferece arquiteturas de referência validadas em laboratório otimizadas para várias cargas de trabalho de IA e outras. Esses Validated Designs incluem conceitos de arquitetura, visões gerais completas da solução, desempenho e outras validações no laboratório que comprovam os recursos da solução na carga de trabalho para a qual ela foi projetada. Essas cargas de trabalho incluem ambientes virtualizados, MLOps, aprendizado de máquina, IA conversacional, inferência de IA generativa, ajuste do modelo de IA generativa, NVIDIA Fleet Command e OpenShift AI.⁵⁵

Por exemplo, a IA para o Virtualized Environments Validated Design combina a IA pronta para VMware com NVIDIA AI Enterprise na infraestrutura da Dell, otimizando a IA nos ambientes virtuais.⁵⁶ O Guia de Validated Design inclui resultados de desempenho que mostram o treinamento do modelo ResNet, comprovando aos clientes que o design funciona e mostrando o tipo de desempenho que eles podem esperar.⁵⁷ Essas validações fornecem valor aos clientes, além de simplesmente listar o hardware que funciona em conjunto; elas explicam conceitos, recomendam configurações e mostram aos clientes as considerações e as expectativas de desempenho.⁵⁸

A Supermicro oferece soluções com base em casos de uso, mas fica aquém do nível de arquitetura de referência que observamos com os Dell Validated Designs. Em vez de recomendar uma solução específica para cargas de trabalho específicas, a Supermicro organiza os servidores e GPUs em categorias, como inferência e treinamento de IA, HPC/IA, visualização e design e muito mais.⁵⁹ Nos folhetos e fichas técnicas, essas categorias consistem em alguns servidores e GPUs que eles recomendam como mais adequados para a tarefa, vários casos de uso e listas das tecnologias importantes no escopo e sugestões de software.⁶⁰ Ao contrário dos Dell Validated Designs, eles não parecem incluir arquiteturas de rede e dados de desempenho ou validação. A Supermicro também fornece vários Reference Designs que oferecem uma arquitetura de referência mais detalhada para algumas soluções de IA, como um treinamento de IA em larga escala com um resumo da solução de resfriamento líquido Supermicro publicado com a NVIDIA.⁶¹ Os clientes podem encontrar uma arquitetura de referência mais detalhada para um cenário específico, mas no momento da nossa pesquisa, encontramos apenas três: a arquitetura de resfriamento líquido que mencionamos anteriormente, uma arquitetura de estação de trabalho de IA e uma arquitetura do RedHat OpenShift.⁶²

Em geral, descobrimos que os Dell Validated Designs cobriram mais cargas de trabalho de IA e ofereceram orientações mais detalhadas do que as ofertas da Supermicro.



Serviços de gerenciamento e iDRAC

De acordo com um [relatório de abril de 2023](#) da Principled Technologies, o Integrated Dell Remote Access Controller (iDRAC) oferece vários recursos avançados em relação à Intelligent Platform Management Interface (IPMI) da Supermicro, particularmente em automação, segurança e configuração.⁶³ A Tabela 3 apresenta uma comparação dos recursos de gerenciamento da Dell e da Supermicro retirada desse relatório, que mostra como o iDRAC pode fornecer implantação e atualização de firmware mais fáceis e mais recursos de segurança do que a Supermicro IPMI. Observe que alguns achados podem ter mudado desde a publicação original.

Tabela 3: Resumo da comparação da Principled Technologies de abril de 2023 entre as ferramentas de gerenciamento da Dell e da Supermicro. Alguns achados podem ter mudado desde a publicação. Fonte: Principled Technologies <https://facts.pt/V5fDf06>.

	Qual é o diferencial das ferramentas de gerenciamento da Dell?	O quanto é melhor?
Atualizações mais fáceis de firmware <i>iDRAC9 vs. Supermicro IPMI</i> <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> Atualizações on-line automatizadas com iDRAC9, com opções de programação O OME permite a criação de repositórios de firmware personalizados e pode atualizar o firmware do BIOS, BMC e outros componentes do servidor sem ferramentas ou agentes adicionais 	<ul style="list-style-type: none"> Configuramos as atualizações automáticas no iDRAC em apenas 74 segundos A Supermicro IPMI não tem função de atualização automática disponível, portanto, os administradores precisam atualizar manualmente O SSM suporta apenas atualizações de firmware de BIOS e BMC e requer o SUM para atualizar outros componentes
Mais recursos de segurança <i>iDRAC9 vs. Supermicro IPMI</i> <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> O iDRAC9 oferece MFA e desativação dinâmica de portas USB sem tempo de inatividade do sistema O OME oferece controle de acesso baseado em função (RBAC) e controle de acesso baseado em escopo (SBAC) para restringir o gerenciamento de dispositivos a um subgrupo de dispositivos 	<ul style="list-style-type: none"> A Supermicro IPMI não tem recursos de MFA A Supermicro IPMI requer uma reinicialização do sistema e a entrada na configuração do BIOS para desativar as portas USB O Supermicro SSM oferece RBAC, mas não o SBAC mais restritivo
Gerenciamento facilitado do ciclo de vida <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> Gerenciamento do ciclo de vida completo e sem agentes via OME para facilitar o gerenciamento e o monitoramento 	<ul style="list-style-type: none"> O SSM requer o agente SuperDoctor5 para métricas detalhadas de integridade do sistema local e o Supermicro Update Manager (SUM) para atualizar componentes adicionais
Implementação mais fácil de servidores <i>iDRAC9 vs. Supermicro IPMI</i>	<ul style="list-style-type: none"> Importe um perfil de servidor completo da Dell em apenas 12 etapas, usando iDRAC9 Opções robustas de configuração do BIOS com iDRAC9, com 52 recursos do BIOS e suporte para configuração de componentes como NIC RAID e iDRAC 	<ul style="list-style-type: none"> A Supermicro IPMI nos permitiu salvar e restaurar apenas a configuração da IPMI em vez do perfil inteiro do servidor O iDRAC9 tem 52 recursos do BIOS, enquanto a IPMI não oferece opções de configuração do BIOS
Mais opções para geração de relatórios e lógica analítica <i>iDRAC9 vs. Supermicro IPMI</i> <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> O iDRAC9 oferece streaming de telemetria, que permite que os usuários enviem facilmente dados do servidor para ferramentas de lógica analítica como o Splunk O OME envia os dados de telemetria diretamente para o CloudIQ para facilitar o monitoramento 	<ul style="list-style-type: none"> A IPMI oferece apenas um recurso SYSLOG que os administradores podem usar para enviar mensagens para agregação e análise eventual O SSM não tem uma solução de gerenciamento baseada em nuvem equivalente ao Dell CloudIQ
Mais recursos de sustentabilidade <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> Mais métricas para monitoramento no OME Power Manager, incluindo dados de pegada de carbono 	<ul style="list-style-type: none"> O SSM tem métricas de utilização menos robustas e não oferece qualquer maneira para rastrear a pegada de carbono
Mais formas de monitorar <i>OME vs. Supermicro SSM</i>	<ul style="list-style-type: none"> Gerencie servidores Dell de qualquer lugar por meio do aplicativo OpenManage Mobile Monitore dispositivos de terceiros com o OME usando IPs de servidor e credenciais com suporte para importar MIBS SNMP de terceiros 	<ul style="list-style-type: none"> O SSM não tem aplicativo móvel O SSM não permite o monitoramento de dispositivos de terceiros com IPs de servidor



Conclusão

Quando se trata de projetar, implementar, gerenciar e manter soluções de IA em uma empresa, há muitos fatores a serem considerados. Para ajudar você a investir com sabedoria e aproveitar ao máximo sua solução de IA, é recomendável procurar um fornecedor que possa ser mais do que um provedor de hardware. Nossa pesquisa indica que a Dell oferece serviços que podem ajudar você como um parceiro durante toda a jornada; portanto, pense em investir com a Dell enquanto você se dedica à IA.

1. Vipera, "NVIDIA's H100 and A100 GPU Cards: Exploring the Intricacies of SXM and PCI-E Connections", acessado em 5 de janeiro de 2024, <https://www.viperatech.com/unraveling-the-mysteries-sxm-vs-pci-e-connections-in-nvidias-high-end-h100-and-a100-gpus/>.
2. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results", acessado em 5 de janeiro de 2024, <https://mlcommons.org/en/inference-datacenter-31/>.
3. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results".
4. Dell, "PowerEdge XE Servers", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/servers/specialty-servers/poweredge-xe-servers.htm>.
5. Supermicro, "Next Leap of AI Infrastructure is Here", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/accelerators/nvidia>.
6. Dell, "PowerEdge XE9680", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9680-spec-sheet.pdf>.
7. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0069. O nome e o logotipo da 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.
8. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0135. O nome e o logotipo da 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.
9. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0069. O nome e o logotipo da 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.

-
10. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0132. O nome e o logotipo da 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.
 11. Dell, "PowerEdge XE9640 Rack Server", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
 12. Accelsius, "Enabling the AI Revolution with Liquid Cooling", acessado em 5 de janeiro de 2024, <https://www.accelsius.com/blog/enabling-the-ai-revolution-with-liquid-cooling>.
 13. Dell, "PowerEdge XE9640 Technical Guide", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9640-technical-guide.pdf>.
 14. Supermicro, "GPU Server Systems", acessado em 5 de janeiro de 2024, https://www.supermicro.com/en/products/gpu?pro=pl_grp_type%3D1.
 15. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0067. O nome e o logotipo da 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.
 16. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0133. O nome e o logotipo da 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.
 17. Dell, "PowerEdge XE8640", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe8640-spec-sheet.pdf>.
 18. Dell, "PowerEdge XE8640 Rack Server", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.
 19. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0066. O nome e o logotipo da 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.
 20. Pontuação encerrada da inferência v3.1 verificada pelo MLPerf®. Recuperada de <https://mlcommons.org/benchmarks/inference-datacenter/>, 5 de dezembro de 2023, entrada 3.1-0133. O nome e o logotipo da 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.
 21. Dell, "Artificial Intelligence (AI) technologies powered by Dell Precision workstations", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/ai-technologies/index.htm?hve=explore+dell+precision+for+ai#tab0=0>.
 22. Supermicro, "Super Workstations", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/products/superworkstation>.
 23. Supermicro, "Rackmount Workstations", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/products/rackmount-workstations>.
 24. Stephen Pritchard, "Storage requirements for AI, ML and analytics in 2022", acessado em 5 de janeiro de 2024, <https://www.computerweekly.com/feature/Storage-requirements-for-AI-ML-and-analytics-in-2022>.
 25. Dell, "PowerScale AI-Ready Data Platform", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale>.
 26. Dell, "Dell PowerScale All-Flash", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/pt-br/products/storage/technical-support/h15963-ss-powerscale-all-flash-nodes.pdf>.
 27. Dell, "Recursos de software do Dell PowerScale OneFS", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/pt-br/products/storage/technical-support/h18275-onefs-software-features-data-sheet.pdf>.
 28. Dell, "Dell ECS Enterprise Object Storage", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/storage/ecs/>.

-
29. Supermicro, "Accelerating AI Data Pipelines", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/products/storage>.
 30. Supermicro, "Supermicro Software-Defined Storage and Memory Solutions", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/software-defined-storage>.
 31. Supermicro, "Supermicro WEKA Distributed Storage Solution", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/wekaio>.
 32. Dell, "Dell ECS Enterprise Object Storage", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/storage/ecs/>.
 33. Dell, "Dell ECS Enterprise Object Storage".
 34. Supermicro, "Supermicro OSNEXUS Software-Defined Storage Solution", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/osnexus>.
 35. ASBIS, "Supermicro Solution for Scalify RING", acessado em 5 de janeiro de 2024, <https://news.asbis.com/news/suppliers/supermicro-renewed-the-line-of-scality-ring-solution/>.
 36. Supermicro, "Supermicro solution for Quantum ActiveScale", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/activescale>.
 37. Dell, "Scalable and elastic Storage as-a-Service", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/apex/storage/data-storage-services/>.
 38. Dell, "Flip the Switch to Open Networking with PowerSwitch", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/networking/>.
 39. Dell, "Dell PowerSwitch Data Center Switches", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/networking/data-center-switches/>.
 40. Supermicro, "SSE-T7132S - 400Gb Ethernet Switch", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/products/accessories/Networking/SSE-T7132SR.php>.
 41. Dell, "Dell EMC Networking SmartFabric Services Deployment with VxRail 4.7 — Fabric Design Center", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/dell-emc-networking-smartfabric-services-deployment-with-vxrail-4-7-1/fabric-design-center-26/>.
 42. Dell, "Dell SmartFabric Services", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/networking/smartfabric/>.
 43. Penny Madsen, "Scaling Skills for AI: Lessons from Early Adopters", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/industry-market/idc-brief-importance-of-skills-for-ai-dell.pdf>.
 44. Dell, "Design Guide — Generative AI in the Enterprise — Model Customization — Overview", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/overview-5381/>.
 45. Dell, "Design Guide — Generative AI in the Enterprise — Model Customization — Advisory Services for Generative AI", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/advisory-services-for-generative-ai-1/>.
 46. Dell, "Design Guide — Generative AI in the Enterprise — Model Customization— Adoption Services for Generative AI", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/adoption-services-for-generative-ai-1/>.
 47. Dell, "Design Guide — Generative AI in the Enterprise — Model Customization — Adoption Services for Generative AI".
 48. Dell, "Artificial Intelligence (AI) Ready Solution Services", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/services/solutions/artificial-intelligence-services.htm>.
 49. Dell, "Comprehensive AI Training Modules Tailored for You", acessado em 5 de janeiro de 2024, <https://education.dell.com/content/emc/en-us/home/training/aiml.html>.
 50. Supermicro, "Services and Support", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/support>.
 51. Travis Vigil, "Dell and NVIDIA: Bringing Generative AI to the Enterprise", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/blog/dell-and-nvidia-bringing-generative-ai-to-the-enterprise/>.
 52. Dell, "Dell Technologies and Hugging Face to Simplify Generative AI with On-Premises IT", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/corporate/newsroom/announcements/detailpage.press-releases~usa~2023~11~20231114-dell-technologies-and-hugging-face-to-simplify-generative-ai-with-on-premises-it.htm>.

53. Richard DeMare, "Starburst and Dell expand partnership to accelerate AI efforts with more intelligent data collection", acessado em 5 de janeiro de 2024, <https://www.starburst.io/blog/starburst-and-dell-expand-partnership-to-accelerate-ai-efforts-with-more-intelligent-data-collection/>.
54. Business Wire, "SiMa.ai and Supermicro Announce Partnership to Accelerate Power Efficient ML at the Edge", acessado em 5 de janeiro de 2024, <https://www.businesswire.com/news/home/20231129609794/en/SiMa.ai-and-Supermicro-Announce-Partnership-to-Accelerate-Power-Efficient-ML-at-the-Edge>.
55. Dell, "Dell AI Solutions", acessado em 5 de janeiro de 2024, <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>
56. Dell, "Unlock the power of AI in virtualized environments", acessado em 5 de janeiro de 2024, <https://www.delltechnologies.com/asset/en-us/products/ready-solutions/briefs-summaries/ai-vxrail-powerscale-brief.pdf>.
57. Dell, "Design Guide — Virtualizing GPUs for AI with VMware and NVIDIA Based on Dell Infrastructure — Performance Results", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/design-guide-virtualizing-gpus-for-ai-with-vmware-and-nvidia-based-on-dell-infrastructure-1/performance-results-15/>.
58. Dell, "Design Guide — Virtualizing GPUs for AI with VMware and NVIDIA Based on Dell Infrastructure — Design Considerations", acessado em 5 de janeiro de 2024, <https://infohub.delltechnologies.com/l/design-guide-virtualizing-gpus-for-ai-with-vmware-and-nvidia-based-on-dell-infrastructure-1/design-considerations-105/>.
59. Supermicro, "Accelerate Every Workload", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/ai-deep-learning>.
60. Supermicro, "Supermicro Enterprise AI Inference & Training", acessado em 5 de janeiro de 2024, https://www.supermicro.com/datasheet/Datasheet_AI-Workloads_Enterprise_AI_Inferencing_and_Training.pdf.
61. Supermicro, "SUPERMICRO RACK SCALE SOLUTIONS: LARGE SCALE AI TRAINING WITH LIQUID COOLING", acessado em 5 de janeiro de 2024, https://www.supermicro.com/solutions/Solution-Brief_Rack_Scale_AI.pdf.
62. Supermicro, "Accelerate Every Workload", acessado em 5 de janeiro de 2024, <https://www.supermicro.com/en/solutions/ai-deep-learning>.
63. Principled Technologies, "Dell management tools made server deployment and updates easier, offered more comprehensive security, and provided more robust infrastructure analytics", acessado em 5 de janeiro de 2024, <https://www.principledtechnologies.com/Dell/Management-tools-vs-Supermicro-0423.pdf>.

► Visualize a versão original em inglês do relatório em <https://facts.pt/q9p46K9>

Este projeto foi encomendado por Dell Technologies.



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.

ISENÇÃO DE RESPONSABILIDADE DE GARANTIAS, LIMITAÇÃO DE RESPONSABILIDADE:
A Principled Technologies, Inc. empreendeu esforços razoáveis para assegurar a precisão e a validade de seus testes; outrossim, a Principled Technologies, Inc. isenta-se especificamente de qualquer garantia, implícita ou expressa, relacionada à análise e ao resultado dos testes, à sua precisão, à sua perfeição ou à sua qualidade, incluindo qualquer garantia implícita de adequação para qualquer propósito específico. Todas as pessoas ou empresas que contam com os resultados de qualquer teste fazem isso sob seu próprio risco e concordam que a Principled Technologies, Inc., seus funcionários e seus funcionários terceirizados não têm qualquer responsabilidade sobre qualquer reclamação de perda ou danos derivados de erros ou defeitos alegados em resultados ou procedimentos de testes.

Em hipótese alguma a Principled Technologies, Inc. será responsável por quaisquer danos indiretos, especiais, incidentais ou consequentes em conexão com seus testes, mesmo que ela tenha sido alertada sobre a possibilidade de tais danos. Em hipótese alguma a responsabilidade da Principled Technologies, Inc., inclusive sobre danos diretos, deverá exceder as quantias pagas com relação aos testes da Principled Technologies, Inc. Os únicos recursos para o cliente são apenas aqueles estabelecidos na presente.