



# Aceleradores de servidores

Turbine seus aplicativos



## 7 em cada 10

executivos da tecnologia classificaram a IA como uma das três tecnologias mais importantes dos próximos 5 a 10 anos<sup>1</sup>

## até 100 vezes

CPUs em uma única GPU<sup>2</sup>

## > 600 aplicações

aceleradas e todos os frameworks de deep learning<sup>3</sup>

### Acelere a percepção e a inovação

Para a empresa digital, o sucesso depende do aproveitamento de dados grandes e rápidos. Porém, à medida que os conjuntos de dados crescem, os data centers tradicionais estão começando a esbarrar em limitações de desempenho e escala, especialmente quando se trata de incluir e consultar fontes de dados em tempo real.

Embora alguns tenham aproveitado os aceleradores para agilizar a visualização, modelagem e simulação, atualmente muitos aplicativos conhecidos podem aproveitar os aceleradores para impulsionar insights e inovações. Aceleradores, como unidades de processamento gráfico (GPUs) e matrizes de gate programáveis em campo (FPGAs) complementam e aceleram CPUs, usando processamento paralelo para fazer a análise de grandes volumes de dados com mais rapidez e eficiência. Os data centers acelerados também podem oferecer uma maior economia, proporcionando desempenho inovador com menos servidores, resultando em percepções mais rápidas e menores custos.

As organizações de vários setores estão adotando aceleradores de servidores para superar a concorrência — aprimorando ofertas de produtos e serviços com insights obtidos por dados, aumentando a produtividade com melhor desempenho de aplicativos, otimizando operações com lógica analítica rápida e avançada e reduzindo o time-to-market, de uma forma mais rápida do que nunca.

A Dell EMC oferece aceleradores de servidor compatíveis com os servidores Dell EMC PowerEdge para que você possa turbinar seus aplicativos.

### Casos de uso do acelerador

- **Aprendizagem automática e deep learning** — os aceleradores trouxeram a IA da teoria para a prática, permitindo a capacidade de processamento paralelo necessário para acelerar as cargas de trabalho de treinamento e de inferência.
- **A lógica analítica preditiva** — a IA, habilitada pelos aceleradores, pode impulsionar massivamente a lógica analítica, possibilitando a correlação dinâmica e oferecendo resultados preditivos com velocidade, precisão e dimensionamento impressionantes.
- **Bancos de dados acelerados** — os aceleradores podem ajudar a trazer velocidade aos agregados de dados, a operações de classificação e agrupamento, para resolver lógicas analíticas complexas que sobrecarregam os bancos de dados tradicionais.
- **Dados de streaming** — a Internet das coisas (IoT) criou um fluxo gigantesco de dados. Os aceleradores permitem inclusão simultânea, exploração e visualização de dados de streaming para análise em tempo real.

<sup>1</sup> "Pesquisa em IA de 2019 da Edelman", março de 2019.

<sup>2</sup> "GPU com núcleo Tensor NVIDIA Tesla V100", maio de 2019.

<sup>3</sup> "NVIDIA Tesla P100", maio de 2019.

## GPUs

As GPUs (Graphics Processing Units, unidades de processamento gráfico) são co-processadores criados para acelerar o desempenho de computação. Geralmente, uma GPU tem milhares de núcleos projetados para a execução eficiente de funções matemáticas. Partes de uma carga de trabalho são transferidas da CPU para a GPU, enquanto o restante do código é executado na CPU, melhorando o desempenho geral do aplicativo.

A Dell EMC oferece uma variedade de GPUs na forma de placas PCIe que encaixam nos slots PCIe do servidor e de módulos SXM2 montados na placa-mãe do servidor.

O servidor PowerEdge C4140 com módulos SXM2 pode aproveitar a interconexão de alta velocidade do NVLINK entre GPUs.

## FPGAs

As matrizes de gate programáveis em campo (field programmable gate arrays, FPGAs) são aceleradores com características técnicas muito específicas para a execução de certos tipos de algoritmos até 1.000 vezes mais rápidas do que as soluções de softwares tradicionais. A FPGA pode ser configurada para atender aos requisitos de tarefas ou aplicativos específicos com muita precisão.

A Dell EMC oferece uma gama de FPGAs na forma de placas PCIe que encaixam nos slots PCIe do servidor.

## Processamento paralelo

Processamento paralelo é um método de dividir e executar tarefas do programa simultaneamente em vários microprocessadores, reduzindo, assim, o tempo de processamento. O processamento paralelo pode ser feito por meio de um computador com dois ou mais processadores ou por meio de uma rede de computadores. O processamento paralelo também é chamado de computação paralela.

## Otimize o código

Para aproveitar todas as vantagens dos aceleradores de servidores, otimize o código do software.

A arquitetura de dispositivos unificados de computação (Compute Unified Device Architecture, CUDA) dá acesso direto ao conjunto de instruções virtuais das GPUs e aos elementos computacionais paralelos para a execução de kernels de computação.

A configuração da FPGA geralmente é especificada usando uma linguagem de descrição de hardware (HDL). Por meio do HDL, os conceitos fundamentais podem ser configurados para atender aos requisitos de tarefas ou aplicativos específicos, em essência, imitando circuitos integrados específicos de aplicativo (ASICs).

- **Visualização** — os aceleradores melhoram o desempenho de aplicativos de visualização 3D, como um projeto computadorizado, permitindo que o software desenhe modelos em tempo real à medida que o usuário os movimenta.
- **Modelagem e simulação** — os aceleradores podem fornecer uma simulação de modelagem mais rápida do que aquela em tempo real para avaliação antecipada, testes rápidos de modificações nos projetos e mais iterações em projetos.
- **Modelagem financeira** — as soluções de AI (inteligência artificial) e HPC estão revolucionando as ferramentas tradicionais de lógica analítica, permitindo que o setor aproveite dos amplos pontos de dados a fim de entender melhor os riscos e retornos.
- **Processamento sísmico** — as empresas de petróleo e gás estão encontrando maneiras novas e melhores para extrair informações de grandes data stores sísmicos, aproveitando-se dos aceleradores para agilizar os resultados e diminuir custos.
- **Processamento de sinais** — os aceleradores permitem que os prestadores de serviço modelem e analisem enormes fluxos de dados de sinal oriundos de computadores, rádios, vídeos e celulares em tempo real.

## GPUs e FPGAs para os servidores Dell EMC PowerEdge

Turbine seus aplicativos com os aceleradores de desempenho disponíveis em alguns servidores Dell EMC PowerEdge em rack e em torre. A tabela a seguir mostra o número e os modelos de aceleradores (GPUs e FPGAs) compatíveis com os [servidores PowerEdge](#). Xilinx® Alveo™ U200 e a Placa de Aceleração Programável Intel® (PAC) com Arria® 10 GX são FPGAs e o restante são GPUs NVIDIA®.

Torre 2S	Rack 1U, 2S		Rack 2U, 2S		Rack 4U, 4S	
<a href="#">T640</a>	<a href="#">R640</a>	<a href="#">C4140</a>	<a href="#">R740</a>	<a href="#">R7425</a>	<a href="#">R840</a>	<a href="#">R940xa</a>
4 aceleradores	1 acelerador	4 aceleradores	3 aceleradores	3 aceleradores	2 aceleradores	4 aceleradores
			P4			
	T4		T4	T4		
P40		P40	P40	P40	P40	
M10			M10	M10	M10	
V100 16 GB ou 32 GB		V100 16 GB ou 32 GB PCIe ou SXM2	V100 16 GB ou 32 GB			
			U200	U200	U200	U200
	10GX		10GX		10GX	10GX

U = unidades de rack, S = soquetes ou CPUs

O número de aceleradores listados na tabela refere-se a placas de largura dupla e altura completa. O T4 é uma placa de meia altura e baixo perfil.

As GPUs variam de acordo com o número de núcleos CUDA, quantidade de memória e requisitos de energia e refrigeração. Por exemplo, a GPU T4 NVIDIA tem 2560 núcleos de CUDA, memória de 16 GB e uso de até 70 Watts. O NVIDIA Tesla® V100 tem 5120 núcleos de CUDA, memória de até 32 GB e uso de até 300 Watts. As especificações do FPGA Intel e Xilinx incluem capacidade de memória incorporada e fora do chip, que varia de 35 a 64 MB usando de 10 a 225 Watts.



## 47 vezes mais

throughput do que uma CPU na inferência de deep learning<sup>4</sup>



## 1,5 TFLOPs

(tera operações de ponto flutuante por segundo) e eficiência no consumo de energia de até 20 GFLOPs/Watts<sup>5</sup>



## até 90 vezes

mais desempenho do que em CPUs para inferência de aprendizagem automática, transcodificação de vídeo, pesquisa de banco de dados e lógica analítica<sup>6</sup>

### GPUs NVIDIA Tesla

As GPUs NVIDIA Tesla oferecem a potência necessária para executar visualizações e outras cargas de trabalho com mais rapidez do que nunca. Além disso, o Tesla oferece alto desempenho e densidade de usuário para o VDI (Virtual Desktop Infrastructure, infraestrutura de desktop virtual).

- [GPU de núcleo Tensor de T4](#)
- [Acelerador de GPU Tesla V100](#)
- Interconexões [de fabrics NVIDIA NVLink™](#)
- Contêineres [NVIDIA GPU CLOUD™](#)
- [Catálogo de aplicativos de software](#)
- Recursos [para desenvolvedores NVIDIA](#)

As GPUs NVIDIA estão disponíveis nos [servidores Dell EMC PowerEdge](#), incluindo R640, R740, R740xd, R7425, R840, R940xa, C4140, e em [Dell EMC Ready Solutions para a IA HPCand](#).

### FPGAs Intel

Os FPGAs Intel podem ser reprogramados dinamicamente com um caminho de dados que corresponde exatamente às suas cargas de trabalho, como lógica de dados, inferência de imagem, criptografia e compactação.

- [Intel PAC com FPGA Intel Arria 10 GX](#)
- [Eixo de aceleração de software](#)
- [Hub de aceleração de FPGA Intel](#)

Disponível nos [servidores Dell EMC PowerEdge](#) incluindo R640, R740, R740xd, R840 e R940xa e nas [soluções de HPC e IA](#). Faça download do [guia de configuração](#).

### FPGAs Xilinx

Baseadas na arquitetura do Xilinx® 16nm UltraScale™, as placas aceleradoras Xilinx Alveo podem ser adaptadas para alterar os requisitos de aceleração e os padrões de algoritmo, além de serem capazes de acelerar qualquer carga de trabalho sem alterar hardware, e reduzir o custo total de propriedade.

- [Alveo U200](#)
- [Ecossistema de software](#)
- [Ferramentas do desenvolvedor](#)

Disponível nos [servidores Dell EMC PowerEdge](#), incluindo R740, R740xd, R7425, R840 e R940xa, e nas [soluções de HPC e de IA](#).

## Soluções aceleradas da Dell EMC

Economize tempo com as soluções da Dell EMC e dos parceiros com GPUs e FPGAs.

### Dell EMC Ready Solutions

[A Dell EMC Ready Solutions para HPC](#) torna a adoção de computação avançada mais rápida e mais simples. A Dell EMC oferece uma gama de soluções de computação de alto desempenho flexíveis e dimensionáveis, com servidores, sistemas de rede, armazenamento, soluções e serviços otimizados em conjunto para abordar os casos de uso em uma variedade de setores.

[As Dell EMC Ready Solutions para IA](#) incluem tudo o que você precisa para acelerar suas iniciativas de IA. Tornando a inteligência artificial mais simples, esses sistemas integrados são ideais para aprendizagem automática e deep learning, de modo que você consegue obter insights mais rápidos e profundos sobre seus clientes e o seu negócio.

<sup>4</sup> ["GPU com núcleo Tensor NVIDIA Tesla V100"](#), maio de 2019.

<sup>5</sup> ["Recursos das FPGAs Intel Arria 10"](#), maio de 2019.

<sup>6</sup> ["Placa aceleradora de data center Alveo U200"](#), maio de 2019.



### Soluções disponíveis com parceiros da Dell EMC

[As soluções Amulet Hotkey® para desktops virtuais](#) combinam servidores de nível empresarial com aceleradores de GPU virtuais para oferecer soluções otimizadas para data center e de alta densidade, a fim de simplificar a transição para o Windows® 10. Além disso, as GPUs virtuais ajudam a atender à crescente demanda por virtualização acelerada por gráficos dos programas do dia a dia, como Windows 10, Microsoft® Office 365®, YouTube® e muito mais, para uma experiência excepcional de desktop virtual. [Leia sobre os sucessos dos clientes da Amulet Hotkey.](#)

[O node de computação acelerada do FPGA baseado no BittWare Stratix® 10](#) permite executar as cargas de trabalho mais exigentes de data center, usando um servidor de montagem em rack de alta densidade otimizado para os aceleradores FPGA, e até quatro FPGAs Intel Stratix 10 por 1U. Os sistemas podem ser adquiridos diretamente com a BittWare ou a Dell EMC. [Assista ao vídeo.](#)

[Kinetica®](#) é um mecanismo de percepção que inclui um banco de dados acelerado por GPU, recursos visuais de detecção e aprendizagem automática e computação paralela acelerada. Em execução em servidores Dell PowerEdge e GPUs NVIDIA, o Kinetica ajuda as organizações a enfrentar os desafios que vêm com enormes quantidades de dados complexos e imprevisíveis. Leia o artigo: [Explicando GPUs para seu CEO: O poder da produção.](#)

[A Tracewell Systems®](#) oferece tecnologia de computação avançada e pronta para uso para empresas, agências governamentais e OEMs em lugares onde os fatores ambientais criam desafios exclusivos de computação, seja no ar, no mar ou em terra, em instalações fixas e móveis, ou em situações em que a integração com hardware especializado ou software seja necessária. [Obtenha data sheets, vídeos e recursos.](#)

### Parceiros de software de aceleração da Dell EMC

[Os aplicativos virtuais NVIDIA GRID™](#) Virtual PC e GRID® aprimoram desktops e aplicativos virtuais, com desempenho comprovado baseado nos GPUs NVIDIA® Tesla®.

[O software Kinetica](#) acelera drasticamente as cargas de trabalho tradicionais de processamento de lógica analítica on-line (OLAP) usando GPUs para computação paralela.

[O software BitFusion](#) desagrega GPUs, FPGAs e/ou ASICs e os conecta dinamicamente em qualquer lugar no data center.

[O data warehousing acelerado por GPU da SQream Technologies®](#) são capazes de fazer o dimensionamento de terabytes para petabytes, adaptando-se a qualquer escala e carga de trabalho.

[O software nativo de GPU FASTDATA.io](#) PlasmaENGINE® permite o processamento em tempo real de dados infinitos em movimento, em vários nós, com várias GPUs.

[O RAPIDS](#) é um conjunto de bibliotecas de dados baseado no NVIDIA CUDA-X para executar pipelines de treinamento científico de ponta a ponta para dados em GPUs NVIDIA.

### Torne-se um parceiro da Dell Technologies

Ao participar do [Dell Technologies Partner Program](#), você contribui para uma rede de parceiros que, juntos, estão tornando real a transformação do digital, da TI, da força de trabalho e da segurança de organizações de todo o mundo, todos os dias. Complementando o portfólio mais completo do setor, com soluções da borda ao núcleo e até a nuvem, o Dell Technology Partners Program foi criado para ser simples. Previsíveis. Lucrativos.



## Recursos

Prepare seu data center com os servidores PowerEdge para lidar com qualquer carga de trabalho.

Os servidores em torre PowerEdge são projetados para crescer junto com sua organização, em seu ritmo. Os servidores em rack PowerEdge combinam uma arquitetura altamente dimensionável e o equilíbrio ideal entre computação e memória para maximizar o desempenho na mais ampla variedade de aplicativos.

### Resultados de desempenho

Obtenha dados de benchmark por carga de trabalho, arquiteturas de referência e blogs de engenharia em [hpcatdell.com](http://hpcatdell.com) e no [suporte a recursos de soluções de servidores](#). Explore os seguintes white papers:

- [Comparação de desempenho do deep learning — Dimensionamento horizontal e dimensionamento vertical do PowerEdge C4140](#)
- [Aceleração do banco de dados da GPU no PowerEdge R940xa](#)
- [Inferência de deep learning no PowerEdge R7425](#)
- [Inferência de deep learning com a Placa Aceleradora Programável da Intel na infraestrutura Dell EMC](#)

### Acesse os serviços educacionais

Adquira as competências, o treinamento e as certificações de que você precisa em [education.emc.com](http://education.emc.com). Saiba como resolver problemas com deep learning no Deep Learning Institute da Dell EMC.

### Visite o Laboratório de inovação em IA e HPC

Engenharia de soluções de nível internacional, testes de tecnologia e colaboração: [dellemc.com/pt-br/innovationlab](http://dellemc.com/pt-br/innovationlab)

## Entre em contato conosco

Para saber mais, acesse [dellemc.com/pt-br/saplandscapes](http://dellemc.com/pt-br/saplandscapes), [dellemc.com/pt-br/hana](http://dellemc.com/pt-br/hana) ou [entre em contato](#) com seu representante local ou revendedor autorizado.

### Conecte-se a centros de excelência

Hubs globais para inovação e expertise: [dellemc.com/pt-br/coe](http://dellemc.com/pt-br/coe)

## Centre de recursos

Faça download dos [Recursos da Solução do Servidor](#)

## Documentação técnica

Veja os resultados de desempenho, as arquiteturas de referência e os blogs da engenharia em HPC em [hpcatdell.com](http://hpcatdell.com)

## Rack virtual

Consulte servidores e soluções no rack virtual em [esgvr.dell.com](http://esgvr.dell.com)

## Participe da Comunidade de HPC da Dell EMC

Um fórum técnico mundial que promove a troca de ideias [dellhpc.org](http://dellhpc.org)

© 2019 Dell Inc. ou suas subsidiárias. Todos os direitos reservados. Dell, EMC e outras marcas comerciais são marcas comerciais da Dell Inc. ou de suas subsidiárias. Outras marcas comerciais podem pertencer a seus respectivos proprietários. Número de referência: 05/19 DELL-EMC-BRO-Accelerators-USLET-101

Intel® é marca comercial da Intel Corporation nos Estados Unidos e em outros países. NVIDIA®, Tesla® e GRID™ são marcas registradas da NVIDIA Corporation. Microsoft®, Windows® e Office 365® são marcas registradas ou comerciais da Microsoft Corporation nos Estados Unidos e/ou em outros países. YouTube® é marca registrada da Google Inc. Arria® e Stratix® são marcas comerciais da Altera Corporation e registradas no Escritório de patentes e marcas comerciais dos Estados Unidos e em outros países. Xilinx®, UltraScale™ e Alveo™ são marcas comerciais da Xilinx, Inc. SQream Technologies® é marca registrada da SQream Technologies Ltd. Kinetica® é marca comercial da Kinetica e de suas subsidiárias nos Estados Unidos e em outros países. FASTDATA.io® é marca registrada da FASTDATA.io. nos Estados Unidos e em outros países. Amulet Hotkey® é marca comercial da Amulet Hotkey Ltd. e registrada no Reino Unido, Estados Unidos e/ou em outros países. Tracewell Systems é marca registrada da Tracewell Systems nos Estados Unidos e em outros países.

A Dell EMC assegura que as informações apresentadas neste documento estão corretas na data da publicação. As informações estão sujeitas a alterações sem prévio aviso.