Your Innovation Engine

Accelerate key workloads and visualize data into intelligence with GPU-enhanced servers from Dell Technologies.
Dell Technologies is helping customers address their graphics requirements.

Today’s applications have become much more user-interface and user-experience oriented, which helps businesses and employees do more. But delivering visualization from real-time data demands a new level of performance from applications and the hardware that enables it — graphics processing unit (GPU) accelerators — which boost computational performance from the data center to the cloud and edge.

Accelerator technology is, indeed, gaining momentum. It doesn’t matter where you are on that journey. You might be starting to think about GPUs, you’re weighing your options or you’ve already begun deploying accelerators for key workloads. The fact of the matter is that more businesses are recognizing the distinct performance advantages that GPU technology can bring to today’s most important workloads — artificial intelligence and machine learning (AI/ML), high performance computing (HPC), database and data analytics, virtual desktop infrastructure (VDI) or enhanced graphics for media delivery and streaming.

In the not too distant past, the process of incorporating accelerator technology into your infrastructure involved a significant commitment and investment in supercomputers. This approach still continues for extreme workload performance applications, but businesses have more options today to deliver faster visual outcomes across their infrastructure. Dell Technologies is helping customers address their graphics requirements with solutions that bridge their workload challenges with GPUs brought into the mainstream through proven Dell EMC PowerEdge servers. From purpose-built servers designed specifically for accelerated workloads to powerful GPUs available for upgrading their existing servers, there are multiple ways to harness the power of GPUs and deliver faster time-to-value.

Dell Technologies is uniquely positioned to bring accelerator technology into your data center — and business — and support your journey along the way. We enable this through our expansive portfolio of Dell EMC PowerEdge servers, solid partnerships with key accelerator vendors like NVIDIA and our unique technology features, services, and support advantages.
Why GPUs now?

Most people are familiar with GPU technology in the context of gaming and entertainment. What might not be known, however, is that GPUs are quickly gaining traction as a means to power other key workloads.

Businesses that have already begun adopting GPUs are recognizing the urgency to extract greater value from their data, or they are running more demanding apps that require graphics horsepower. They are continuing to invest in GPUs because either they have seen tremendous value in terms of decision-making or they have reduced time, costs, increased utilization, and are seeing greater efficiencies in business operations.

These businesses are also turning to mainstream platforms to accelerate these workloads as the cost to value equation can now deliver higher ROI, and compute power has increased.

Key workloads driving the need for GPU accelerators include:

- **AI/ML/DL** — More businesses turn to AI for different uses, including chatbots and digital assistants, facial recognition systems, smart home devices, search algorithms, smart recommendation systems and more.

- **HPC** — HPC has been around for a while, but GPU accelerators are driving that performance to another level entirely with faster processing speeds and faster time to answers or results that can help solve some of the most pressing challenges or problems.

- **Database and data analytics** — Faster processing in the data analytics realm enables organizations to extract real-time insights from information assets. Faster insights enable better, more timely decisions when they matter most.

- **VDI** — VDI initiatives continue to grow and expand with a rapidly changing workforce. As more people work remotely, from home or on the road — and especially if part of their job requires working with multiple and/or high-resolution monitors, graphics, visuals or imaging, and even office productivity apps — GPU accelerators combined with GPU virtualization software can speed collaboration and worker productivity.

**CPUs, GPUs or Both?**

While CPU technology continues to advance and many applications operate seamlessly with an all-CPU environment, the inclusion of GPUs in the mainstream workloads in the data center is gaining momentum. NVIDIA® GPU acceleration can provide easier access to the applications that will drive your business further, while freeing up valuable CPU resources for other tasks. And Dell Technologies makes it easier than ever to integrate GPU technology into your existing data center environment on your terms — when you are ready and as you need it.
Turbocharge your workloads.

The need for acceleration is a function of the amount of data to be processed, the strategy needed to train and re-train AI/ML models, the real-time requirements to make actionable decisions from data analytics and the demand from remote workforces to deliver a better user experience and the highest quality graphics visualization over VDI.

All of these applications benefit greatly from optimal performance with accelerators in a production environment, as less optimized all-CPU-based systems have difficulty handling unique hardware, software and application requirements.

Drive AI and ML — Deliver better outcomes with solutions featuring the processing power, capacity, throughput and scale to power even the most advanced AI projects.

Harness data and visualize it using your database and data analytics — Capitalize on graphics to deliver scalable data analytics platforms, turning data into actionable business insights and faster decision-making.

Accelerate HPC and entertainment — HPC systems are designed for workloads from life sciences to research, digital manufacturing and other industry solutions, and experience dedicated infrastructure for content production and delivery.

Empower your workforce — Transform the workplace with enhanced mobility that boosts user productivity. With remote access to your company’s data center, access to graphics and access to other critical information, you can increase collaboration among teams while helping users visualize data for improved decision-making.

49% of organizations report that the complexity of integrating AI technologies with existing infrastructure and toolsets is a major barrier to AI implementation.¹

81% of HPC users reported running/supporting AI workloads in a recent Intersect360 survey.²

By 2025, augmented consumerization functionality will drive adoption of analytics and business intelligence capabilities beyond 50% for the first time, influencing more business processes and decisions.³

We do this with GPU-ready servers, supported by a complementing family of NVIDIA GPU accelerator cards as well as purpose-designed servers with GPU acceleration built in.

Envision your advantage with the right innovation engine

Dell Technologies helps advance accelerated compute to drive enhanced workload outcomes with greater insights, inferencing and visualization.

Flexible
Elevated visualization
Optimize your infrastructure strategy to accelerate mainstream workloads.

Powerful
Scalable insights
Derive better, faster insights—at scale

Accessible
Simplified digitalization
Effectively manage your infrastructure and performance workloads

The confidence to move forward

NVIDIA-Certified Dell EMC PowerEdge servers have been designed and tested to work together to meet modern application and workload demands. Each certified server delivers the technology and solutions required to help you quickly and flexibly evolve — regardless of industry or size of business.

Visualize your insights with Dell EMC PowerEdge servers.

The engine you use to drive your business forward should help you innovate, adapt and grow. The Dell EMC PowerEdge server portfolio brings enhanced visualization capabilities, the performance needed to drive insights at scale and innovative technologies and services that simplify your digitalization efforts.

Partnerships That Matter

Dell Technologies works closely with NVIDIA to develop solutions for businesses that help you realize your objectives with today’s accelerator-optimized workloads. Based on your organization’s workloads and specific needs, NVIDIA accelerators and Dell EMC PowerEdge servers — both standard and purpose-built servers — can get you where you need to be.

Transform business outcomes with PowerEdge and NVIDIA

Watch the Video
Cool design innovations

There are a number of innovations in PowerEdge server technology that enable performance improvements especially with GPU accelerators. From architectures specifically designed to support acceleration to thermally optimized designs, today’s workloads demand higher-quality components and subsystems to flawlessly drive workload operations. Dell Technologies delivers:

A focus on acceleration with support for the most complete stack of NVIDIA GPUs in the PowerEdge portfolio, delivering maximized performance for your applications.

Thoughtful thermal designs and solutions engineered to specifically address dense heat-producing components.

Dell Technologies multi-vector cooling is an example of an advanced thermal design that streamlines the airflow pathways within the server, directing the appropriate volume of air to where it is needed within the chassis.

Dell Technologies direct liquid cooling (DLC) is being extended to more PowerEdge servers. DLC uses the exceptional heat-removal capability of liquid to absorb and remove heat.

Multi-Vector Cooling 2.0 design enables support for higher performance configurations.

Direct liquid cooling CPU cold plate solution reduces energy costs by 56%.

Security is built into all PowerEdge servers

Security is top of mind for all organizations, which is why Dell Technologies equips all PowerEdge servers with proactive resilience, which helps you protect, detect and recover from incidents. And with GPU accelerators also validated in PowerEdge servers (for example, NVIDIA accelerators also feature a root of trust ability), you can be assured of this additional level of resilience and complementing protection from partners across numerous operating environments.

Protect
- Secure and granular access control
- Cryptographic authentication
- Data protection: at rest and in flight

Recover
- BIOS and OS recovery
- Easy restore
- Rapid response to new vulnerabilities

Detect
- Audit logging and secure alerting
- Intrusion detection
- Drift and change detection

---

4 Based on Dell analysis of publicly available data, March 2021.
Dell EMC recommends the following PowerEdge servers for training. Learn why these servers are ideal.

**AI/ML/DL — training**

NVIDIA GPU acceleration can speed the processing of large data sets in ML environments such as Kinetica, MapD, BlazingDB, Brytylyt, OmniSci and SQream analyzing massive amounts of information and returning results in milliseconds. When it comes to GPU acceleration for ML, memory and storage are top priorities, and it is important to have a well-matched CPU-to-GPU ratio. The main difference between training and inferencing is that training uses forward and backward propagation, while inferencing mostly uses forward propagation. Training usually involves more CPUs or GPUs to accelerate learning performance.

NVIDIA AI Enterprise is an end-to-end, cloud-native suite of AI and data analytics software, optimized, certified and supported by Dell Technologies and NVIDIA to run on VMware® vSphere®. Deploy mainstream AI into your virtualized data centers with greater agility and confidence.

**Accelerator options**

- **IDEAL**
  - NVIDIA A100 Tensor Core GPU
  - NVIDIA AI Enterprise

- **BEST**
  - Dell EMC PowerEdge XE8545
  - Dell EMC PowerEdge R750xa

- **BETTER**
  - Dell EMC PowerEdge R7525
  - Dell EMC PowerEdge DSS8440

- **GOOD**
  - Dell EMC PowerEdge R750

**Training solution with PowerEdge and NVIDIA**

Watch the Video
Dell EMC recommends the following PowerEdge servers for inferencing. Learn why these servers are ideal.

### Servers and accelerator options for key workloads

**AI/ML/DL — inferencing**

One of the key testing suites used to determine performance is the **MLPerf Inference**, a benchmark suite for measuring how fast ML and DL systems can process inputs and produce results using a trained model. The benchmarks belong to a very diversified set of ML use cases that are popular in the industry and provide a need for competitive hardware to perform ML-specific tasks. Hence, good performance under these benchmarks signifies a hardware setup that is well optimized for real-world ML inferencing use cases. Users can compare overall system performance in AI use cases of natural language processing, medical imaging, recommendation systems and speech recognition as well as different use cases in computer vision.

The solutions here represent the latest combination of industry technologies to deliver high-performance, accelerated computing for demanding inferencing.

---

**Accelerator options**

**IDEAL**

NVIDIA A100
Tensor Core GPU

NVIDIA A30
Tensor Core GPU

**NVIDIA AI Enterprise**

**BEST**

Dell EMC PowerEdge R7525
Dell EMC PowerEdge R750xa

**BETTER**

Dell EMC PowerEdge R750
Dell EMC PowerEdge R7515

**GOOD**

Dell EMC PowerEdge R650
Dell EMC PowerEdge R6525
Dell EMC PowerEdge R6515

Dell EMC PowerEdge DSS8440

Dell EMC PowerEdge XE2420

---

Supercharge AI workloads with PowerEdge and NVIDIA
Servers and accelerator options for key workloads

HPC

GPU computing is the most accessible and energy-efficient path to HPC in the data center. High performance computing targets complex computing tasks across a broad array of specialized outcomes industries, including financial services, life sciences and research, CAD/CAE, Oil/gas and other engineering simulations, as well as AI and other graphics-demanding workloads. Performance and rapid results are key, as well as other aspects of the infrastructure including solution scalability and a framework of software solutions to take advantage of the hardware for multiple tasks or programming. The NVIDIA CUDA® programming model for general computing on GPUs offers a language-based solution for programmers who want to fine-tune their applications for the best possible performance. CUDA supports more than 600 GPU-accelerated applications, including the top 15 HPC applications. When it comes to HPC, be sure to consider high GPU density, high network bandwidth I/O (IB/OPA/100Gb or above), and peer-to-peer transfers between GPUs.
By 2024, more than 25% of data management vendors will provide a complete framework for data fabric support through a combination of their own products and partners, up from less than 5% today.  

---

**Database and data analytics**

Database and data analytics technologies help companies examine data, draw conclusions and make decisions. Data analytics workloads have a wide variety of business uses across industries — for example, fraud detection in financial services, biomedical analysis in healthcare, customer insights in retail and demand forecasting in manufacturing. Data analytics may involve structured data in databases or unstructured data—or both. The top considerations for data analytics include memory bandwidth, high I/O storage (NVMe™/SSDs), and CPU sizing for an accelerated solution.

Dell EMC recommends the following PowerEdge servers for database and data analytics. Learn why these servers are ideal.

**BEST**

Dell EMC PowerEdge R750xa

**BETTER**

Dell EMC PowerEdge servers

**IDEAL**

NVIDIA A100 Tensor Core GPU

---

Dell EMC recommends the following PowerEdge servers for VDI.

Learn why these servers are ideal.

**Accelerator options**

**IDEAL**

- NVIDIA A16 GPU

**GOOD**

- Dell EMC PowerEdge T550

**BETTER**

- Dell EMC PowerEdge R750

**BEST**

- Dell EMC PowerEdge R7525

- Dell EMC PowerEdge R750xa

Solutions to enable VDI include NVIDIA vPC software and NVIDIA RTX vWS software. NVIDIA vPC for VDI enables consistent performance and better productivity for knowledge workers across multiple devices and 4K resolution monitors.

VDI with NVIDIA RTX vWS helps enable new, powerful virtual workstations quickly, and certified compatibility with industry-leading visualization and analytic applications means you can deploy with confidence.

---

**Virtual desktop infrastructure**

With VDI, a user desktop runs inside a virtual machine that lives on a server in the data center. VDI gives end-users the mobility and freedom to access virtual desktops anytime, anywhere, from any device. VDI can also help organizations streamline their processes and lower costs by consolidating and centralizing management. NVIDIA has the industry's highest user-density GPU-accelerated solution with support for up to 64 virtual desktops per GPU. When it comes to VDI, the most important factors are CPU-to-GPU ratio, storage capacity and memory.

41% of decision-makers say virtualization is a top IT infrastructure and IT operations priority over the next 12 months.⁶

---

Dell EMC recommends the following PowerEdge servers for media and entertainment. Learn why these servers are ideal.

**Accelerator options**

- **IDEAL**
  - NVIDIA A40 GPU
  - NVIDIA RTX Virtual Workstation (vWS)

---

**Performance graphics, media and entertainment**

Professional visualization, delivering the performance and features that can power professional graphics and computing anywhere, and media and entertainment delivery require an infrastructure that can support real-time demands for broadcasting, multiple video streams, live video encoding and decoding, as well as virtual reality, including immersive augmented reality. Most visualization tasks, including ray-tracing and engineering simulation, require high-performance cores and low latency for fast results. Digital manufacturing to content aggregation and distribution centers harness the benefits of optimized accelerators with multiple cores dedicated to parallelization and real-time video operations.

VDI with NVIDIA RTX vWS helps enable new, powerful virtual workstations quickly, and certified compatibility with industry-leading visualization and analytic applications means you can deploy with confidence.
Key benefits
Organizations of all sizes want assurances that business can run efficiently on PowerEdge servers. But you also want to know that your infrastructure can scale easily and meet future needs and that you have made the right investment choice. Accelerator solutions with servers powered by Dell Technologies drive increased agility, enable you to operate efficiently and help you unlock value. Specifically, you can:

- **Reduce TCO** with time and costs savings.
- **Drive improved business outcomes** with access to new CPUs, NVIDIA GPUs and virtual GPU software, and system innovation.
- **Gain confidence** with a strong, secure foundation for all your workloads.
- **Experience greater choice** in certified and validated configurations that optimize workload and application performance.
- **Build out your infrastructure** on your own time, investing as you grow.
- **Increase performance** and decrease latency across your operations.
- **Maximize efficiency** and enable optimal resource usage as you configure your system for specific workloads and reduce unused CPU cycles.
- **Leverage Dell Technologies partnerships** with other industry leaders to create an ecosystem that streamlines operations across your organization.

31% less time required to deploy a model using a third-party solution.¹

---

What does NVIDIA-Certified mean for me and my customer?
NVIDIA-Certified Systems™ consist of NVIDIA GPUs and networking installed in qualified enterprise-class servers that have passed a set of certification tests that validate the best system configurations for a wide range of workloads, as well as for manageability, scalability, and security. GPUs must be NVIDIA Ampere or Turing™ architecture data center GPUs.

Learn more about NVIDIA-Certified PowerEdge servers.
Achieve more, deliver quick results and maximize efficiency.

Dell EMC Technologies Validated Designs are purpose-designed with IT’s transformation journey in mind to run intelligent applications and processes in the digital business.

Along with Dell EMC PowerEdge servers, Dell Technologies partners and collaborates with industry leaders including Intel, Microsoft®, NVIDIA, and others to optimize IT for your critical business workloads together with emerging technologies such as AI, machine learning, and blockchain.

- Supporting VxRail and PowerEdge servers, as well as NVIDIA A100 and A30 Tensor Core GPUs
- Validated Designs for AI, including deep learning with NVIDIA and Cloudera®
- Validated Designs for Analytics
- Validated Designs HPC
- Validated Designs for VDI

New Validated Design Unlocks the Power of AI
Read the Blog
Read the Solution Brief
Why Dell Technologies?

It does not matter where you are on your journey. Dell Technologies can help. Whether you’re still considering GPUs or you know you want to implement more accelerators to enhance processing speed, we are at the cutting edge of that technology, ready to make those innovations available for your business.

Our PowerEdge server technology with NVIDIA accelerators is primed and ready for any stage of the accelerator journey. We have solutions that fit your data center and your workload requirements, and our unique thermal design accommodates GPU-specific heating issues. Our close partnership with NVIDIA means you get proven solutions that are tested and validated to work better together. And we are determined to make your experience as smooth and convenient as possible with our OpenManage Enterprise management suite, cyber-resilient security and quality service and support capabilities.

Ready to get started?

Learn more about how you can visualize more insights from your data today. Contact your Dell Technologies sales representative or visit us at DellTechnologies.com/Accelerators