NVIDIA-CERTIFIED SYSTEMS

Enabling the Enterprise Accelerated Data Center

NVIDIA-Certified Systems bring together NVIDIA® GPUs and NVIDIA high-speed, secure networking to systems from leading partners in configurations validated for optimum performance, reliability, and scale for a diverse range of workloads. A wide variety of system types are available, including popular data center and edge server models, as well as desktop and mobile workstations. With NVIDIA-Certified Systems, enterprises can confidently choose performance-optimized hardware solutions to power accelerated computing workloads—from the desktop to the data center and edge.

The NVIDIA-Certified Systems certification test suite is designed to exercise the performance and functionality of the system by running a set of software that represents a wide range of real-world applications. This includes deep learning (DL) training, AI inference, end-to-end AI frameworks, including NVIDIA Riva and NVIDIA Clara™, data science using Spark, intelligent video analytics (IVA), high performance computing (HPC) and CUDA® functions, and rendering. Security features like TPM are also tested.

KEY BENEFITS TAILORED FOR ENTERPRISE IT

- Optimized Designs: NVIDIA-Certified Systems have been validated for running a diverse range of accelerated workloads with optimum performance, reliability, and scale.
- Customer Choice: Certified servers and workstations are available in a wide variety of models from a broad range of systems providers.
- Enterprise Support: NVIDIA-Certified Systems provide the best foundation for supported enterprise solutions from NVIDIA and partners.

RESOURCES

- **NVIDIA-Certified Systems**
- **Qualified System Catalog**
- **Accelerate Compute-Intensive Workloads with NVIDIA-Certified Systems [White Paper]**

Servers

Performance	NVIDIA-Certified servers deliver excellent performance for a diverse range of workloads. Customers can run most accelerated applications on these systems and be confident that they will perform well.
Manageability	NVIDIA-Certified servers are validated for proper functionality of remote administration features. Certain servers are also tested to ensure full compatibility with enterprise software suites, such as NVIDIA AI Enterprise, enabling administrators to manage these systems using enterprise IT management frameworks.
Scalability	NVIDIA-Certified servers are tested on single-node and multi-node configurations to validate cluster-level features and performance. They enable IT to scale out accelerated infrastructure to meet future workload demands.
Security	NVIDIA-Certified servers secure workflows by protecting data at the platform, network, and application layers. Whether deployed in a data center or at the edge, customers can be assured that they do not have to compromise on security features when running accelerated applications.

Workstations

Performance	NVIDIA workstation GPUs deliver breakthrough performance to professional designers, creators, engineers, researchers, and data scientists. GPUs in NVIDIA-Certified workstations are based on the NVIDIA Ampere architecture built on third generation Tensor core technology, and enable rapid AI model prototyping. These GPUs also introduce next generation RTX technology for even faster photorealistic, ray-traced rendering, graphics, compute, and AI-accelerated workflows.
Reliability	Certified workstations have been validated to make optimal use of GPUs. Well-balanced designs that account for issues, such as thermal management and PCIe configuration, ensure that system resources are used to their fullest potential.
Interoperability	Testing with a wide variety of libraries and applications ensures a consistent experience for end-to-end application workflows across mobile workstations, desktop workstations, and servers. Models, code, and other assets developed on workstations are preserved when the workload is transitioned to intermediate or production servers

Categories of NVIDIA-Certified Systems

NVIDIA and partners have a well-rounded, end-to-end portfolio of certified systems for accelerated computing, no matter where the deployment. Different categories of certification help you choose the best systems for your particular workload.

Category	Use Cases	Key Capabilities
Computer Server	**AI Training and Inferencing, Data Analytics, HPC**	> Large GPU memory to support large AI models and data analytic workloads > Third-Generation Tensor Cores for accelerated AI > High speed GPU memory bandwidth and PCI Gen4 or NVIDIA NVLink® interconnect to support multi-GPU workloads > MIG to support multiple inference jobs > FP64 to support HPC
General Purpose Server	**Visualization, Rendering, Virtual Workstation**	> Second-Generation RT Cores for high-throughput rendering > Large GPU memory to support large models > Latest generation CUDA Cores for accelerated compute > Third-Generation Tensor Cores for accelerated AI > High speed GPU memory bandwidth and PCI Gen4 or NVLink interconnect to support large data sets
High Density Virtualization	**Virtual Desktop, Virtual Workstation**	> Support for multiple, high-resolution monitors to enable maximum productivity and photorealistic quality. > Support for the latest video codecs to enable the highest quality video experiences. > Flexibility to host mixed user profile sizes and mixed user types on the same board.
Enterprise Edge	**Edge Inferencing in industrial or rugged environments**	> Third-Generation Tensor Cores for accelerated AI > MIG to support multiple inference jobs > TPM for secure boot > Redfish for secure remote management capabilities
Industrial Edge	**Edge Inferencing in controlled environments**	> NEBS, IPC rated and/or otherwise ruggedized servers or industrial PCs > Small form factor, low-power and wide temperature ranges > Third-Generation Tensor Cores for accelerated AI > MIG to support multiple inference jobs > TPM for secure boot
Desktop Workstation	**Design, Content Creation, Data Science**	> Second-Generation RT Cores for high-throughput rendering > Third-Generation Tensor Cores for accelerated AI and data analytics workloads > Large GPU memory to support large models and data sets > NVLink interconnect to support largest models and data sets
Mobile Workstation Design	**Content Creation, Data Science, Software Development**	> Second-Generation RT Cores for high-throughput rendering > Third-Generation Tensor Cores for accelerated AI and data analytics workloads > Large GPU memory to support large models and data sets

Learn more

To learn more about NVIDIA-Certified Systems, visit
nvidia.com/certified-systems

NVIDIA