



ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU

Product Guide

The NVIDIA A10 Tensor Core GPU, combined with NVIDIA RTX Virtual Workstation (vWS) software, brings mainstream graphics and video with AI services to mainstream enterprise servers, delivering the solutions that designers, engineers, artists, and scientists need to meet today's challenges.

Built on the latest NVIDIA Ampere architecture, the A10 combines second-generation RT Cores, third-generation Tensor Cores, and new streaming microprocessors with 24 GB of GDDR6 memory for versatile graphics, rendering, AI, and compute performance. From virtual workstations, accessible anywhere in the world, to render nodes to the data centers running a variety of workloads, A10 is built to deliver optimal performance in a single-wide, full-height, full-length PCIe form factor.



Figure 1. ThinkSystem NVIDIA A10 24GB PCle Gen4 Passive GPU

Did you know?

NVIDIA A10 Tensor Core GPU delivers up to 2.5X faster virtual workstation performance, and up to 2.5X more inference performance, compared to the previous generation NVIDIA T4.

A10 is ideal for mainstream graphics and video with Al. 2nd Gen RT Cores and 3rd Gen Tensor Cores enrich graphics and video applications with powerful Al in 150W TDP for mainstream servers.

Part number information

The following table shows the part numbers for the NVIDIA A10 GPU.

Table 1. Ordering information

| Part number | Feature code | Description | Controlled GPU status |
|----------------|--------------|---|-----------------------|
| 4X67A71311 | BFTZ | ThinkSystem NVIDIA A10 24GB PCle Gen4 Passive GPU | No |
| CTO only | BQZS | ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU w/o CEC | No |

The NVIDIA A10 GPU is not Controlled which means the GPU is unrestricted and is available in all markets.

The PCIe option part numbers includes the following:

- One NVIDIA A10 GPU with full-height (3U) adapter bracket attached
- Documentation

GPUs without a CEC chip: The NVIDIA A10 GPU is offered without a CEC chip (look for "w/o CEC" in the name). The CEC is a secondary Hardware Root of Trust (RoT) module that provides an additional layer of security, which can be used by customers who have high regulatory requirements or high security standards. NVIDIA uses a multi-layered security model and hence the protection offered by the primary Root of Trust embedded in the GPU is expected to be sufficient for most customers. The CEC defeatured products still offer Secure Boot, Secure Firmware Update, Firmware Rollback Protection, and In-Band Firmware Update Disable. Specifically, without the CEC chip, the GPU does not support Key Revocation or Firmware Attestation. CEC and non-CEC GPUs of the same type of GPU can be mixed in field upgrades.

Features

The ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU offers the following features:

- **Designed For Accelerated VDI:** Optimized for user density, and combined with NVIDIA vPC software, enables graphics-rich virtual PCs to be accessible from anywhere.
- Affordable Virtual Workstations: Large frame buffer per user for entry-level virtual workstations, with NVIDIA RTX vWS software, running workloads such as computer-aided design (CAD).
- Flexibly Support Diverse User Types: Unique quad-GPU board design enables the provisioning of mixed user profile sizes and user types, such as virtual PCs and virtual workstations, on a single board.
- **Superior User Experience:** Provides increased frame rate and lower end-user latency, versus CPU-only VDI, resulting in more responsive applications and a user experience that is indistinguishable from a native PC or workstation.
- **Double The User Density:** Purpose-built for graphics-rich VDI, with support for up to 64 concurrent users per board, in a dual-slot form factor.
- **High-Resolution Display:** Supports multiple, high resolution monitors to enable maximum productivity and photorealistic quality in a VDI environment.
- More Than 2X The Encoder Throughput: More than double the encoder throughput versus previous generation M10, providing high-performance transcoding and the multiuser performance required for multi-stream video and multimedia.
- **Highest Quality Video:** Support for the latest codecs, including H.265 encode/decode, VP9, and AV1 decode for the highest-quality video experiences.
- NVIDIA Ampere Architecture: NVIDIA Ampere architecture based CUDA cores, second generation

RT-Cores, and third- generation Tensor-Cores provide the flexibility to host virtual workstations powered by NVIDIA RTX vWS software, or to leverage unused VDI resources to run compute workloads with NVIDIA AI Enterprise software.

Technical specifications

The following table lists the specifications of the NVIDIA A10 PCIe GPU.

Table 2. A10 specifications

| Feature | Specification |
|-------------------------------------|---|
| GPU Architecture | NVIDIA Ampere |
| NVIDIA Tensor Cores | Third-generation Tensor Cores |
| NVIDIA CUDA Cores (shading units) | 9216 FP32 CUDA Cores |
| NVIDIA RT Cores | 72 RT Cores |
| Double-Precision Performance (FP64) | Not applicable |
| Single-Precision Performance | FP32: 31.2 TFLOPS Tensor Float 32 (TF32): 62.5 TFLOPS, 125 TFLOPS* |
| Half-Precision Performance | FP16: 125 TFLOPS, 250 TFLOPS* |
| Bfloat16 | 125 TFLOPS, 250 TFLOPS* |
| Integer Performance | INT8: 250 TOPS, 500 TOPS* INT4: 500 TOPS, 1000 TOPS* |
| GPU Memory | 24 GB GDDR6 |
| Memory Bandwidth | 600 GB/sec |
| ECC | Yes |
| Interconnect Bandwidth | Not applicable |
| System Interface | PCIe Gen 4, x16 lanes |
| Form Factor | PCIe full height/length, single width |
| Multi-Instance GPU (MIG) | Not applicable |
| Max Power Consumption | 150 W |
| Thermal Solution | Passive |
| Compute APIs | CUDA, DirectCompute, OpenCL, OpenACC |

^{*} With structural sparsity enabled

Server support

The following tables list the ThinkSystem servers that are compatible.

Table 3. Server support (Part 1 of 4)

| | | | AMD V3 | | 2S Intel V3/V4 | | | 4S 8S Intel V3 | | | Multi Node V3/V4 | | | | GPU Rich | | | | | | |
|----------------|--|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------|----|------------------|
| Part Number | Description | SR635 V3 (7D9H / 7D9G) | SR655 V3 (7D9F / 7D9E) | SR645 V3 (7D9D / 7D9C) | SR665 V3 (7D9B / 7D9A) | ST650 V3 (7D7B / 7D7A) | SR630 V3 (7D72 / 7D73) | SR650 V3 (7D75 / 7D76) | SR630 V4 (7DG8 / 7DG9) | SR850 V3 (7D97 / 7D96) | SR860 V3 (7D94 / 7D93) | SR950 V3 (7DC5 / 7DC4) | SD535 V3 (7DD8 / 7DD1) | SD530 V3 (7DDA / 7DD3) | (2007 / 8007) EV 0350S | SD520 V4 (7DFZ / 7DFY) | SR670 V2 (7Z22 / 7Z23) | SR675 V3 (7D9Q / 7D9R) | a V3 | ٨3 | SR780a V3 (7DJ5) |
| 4X67A71311 | ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU | N | N | N | N | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | Ν | N | N |

Table 4. Server support (Part 2 of 4)

| | | | | /3 | | E | Edg | e | | | | Su omp | | | | | In V2 | | _ | Int V2 | - |
|----------------|---|-----------------------|------------------------|------------------------|---------------------|-----------------|-----------------|--------------|-----------------|-----------------|-----------------|-------------------|-----------------|-------------------|-------------------|--------------------------|------------------------|----------|---------------|--------------|------------------------|
| Part Number | Description | ST50 V3 (7DF4 / 7DF3) | ST250 V3 (7DCF / 7DCE) | SR250 V3 (7DCM / 7DCL) | SE350 (7Z46 / 7D1X) | SE350 V2 (7DA9) | SE360 V2 (7DAM) | SE450 (7D8T) | SE455 V3 (7DBY) | SC750 V4 (7DDJ) | SD665 V3 (7D9P) | SD665-N V3 (7DAZ) | SD650 V3 (7D7M) | SD650-I V3 (7D7L) | SD650-N V3 (7D7N) | ⁷ 2 (7D8K / 7 | ST250 V2 (7D8G / 7D8F) | V2 (7D7R | V2 (7Z75 / 7. | V2 (7Z70 / 7 | SR650 V2 (7Z72 / 7Z73) |
| 4X67A71311 | ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU | N | N | N | Ν | N | Ν | Ν | Ν | Ν | Ζ | Ν | Z | Ν | Ν | N | Ζ | Ν | Ν | Ζ | 4 |

Table 5. Server support (Part 3 of 4)

| | | | Αľ | ИD | V1 | | D | ens | se V | /2 | 4 V | - | 88 | 4 | s v | 1 | 18 | int | tel \ | V1 |
|----------------|---|---------------------|---------------------|-----------------|---------------------|---------------------|-----------------|-----------------|---------|-----------|--------------|---------|---------------------|---------------------|----------------------|---------------------|--------------------|-----|--------|---------------------|
| Part Number | Description | SR635 (7Y98 / 7Y99) | SR655 (7Y00 / 7Z01) | SR655 Client OS | SR645 (7D2Y / 7D2X) | SR665 (7D2W / 7D2V) | SD630 V2 (7D1K) | SD650 V2 (7D1M) | \ - | V2 (7Z69) | V2 (7D31 / 7 | V2 (7Z5 | SR950 (7X11 / 7X12) | SR850 (7X18 / 7X19) | SR850P (7D2F / 2D2G) | SR860 (7X69 / 7X70) | ST50 (7Y48 / 7Y50) | (7 | (7Y54) | SR250 (7Y52 / 7Y51) |
| 4X67A71311 | ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU | N | N | N | N | 3 | N | N | Ν | N | N | N | N | N | Ζ | N | Ν | N | N | N |

Table 6. Server support (Part 4 of 4)

| | | | | 28 | Int | tel \ | V1 | | | D | ens | e V | ′1 |
|-------------|---|---------------------|-----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---|--------------|------|-----|--------------|
| Part Number | Description | ST550 (7X09 / 7X10) | 530 (7X07 / 7X0 | SR550 (7X03 / 7X04) | SR570 (7Y02 / 7Y03) | SR590 (7X98 / 7X99) | SR630 (7X01 / 7X02) | SR650 (7X05 / 7X06) | | SD530 (7X21) | (7X5 | × I | SN850 (7X15) |
| 4X67A71311 | ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU | N | N | Ν | Ν | Ν | Ν | 4 | Ν | Ν | N | N | N |

Operating system support

The following table lists the supported operating systems:

Tip: These tables are automatically generated based on data from Lenovo ServerProven.

Table 7. Operating system support for ThinkSystem NVIDIA A10 24GB PCIe Gen4 Passive GPU, 4X67A71311

| Operating systems | SR650 V2 | SR670 V2 | SR665 | SR650 (Xeon Gen 2) | SR650 (Xeon Gen 1) |
|-------------------------------|----------|----------|-------|--------------------|--------------------|
| Microsoft Windows Server 2016 | Υ | Υ | Υ | Υ | Υ |
| Microsoft Windows Server 2019 | Υ | Υ | Υ | Υ | Υ |
| Microsoft Windows Server 2022 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 7.7 | Ν | Ν | Y 1 | Υ | Υ |
| Red Hat Enterprise Linux 7.8 | Ν | Ν | Y 1 | Υ | Υ |
| Red Hat Enterprise Linux 7.9 | Υ | Υ | Υ1 | Υ | Υ |
| Red Hat Enterprise Linux 8.0 | N | N | N | Υ | Υ |
| Red Hat Enterprise Linux 8.1 | N | N | Υ1 | Υ | Υ |
| Red Hat Enterprise Linux 8.2 | Υ | Υ | Υ1 | Υ | Υ |
| Red Hat Enterprise Linux 8.3 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.4 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.5 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.6 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.7 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.8 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 8.9 | Υ | Υ | Υ | Υ | Ν |
| Red Hat Enterprise Linux 9.0 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 9.1 | Υ | Υ | Υ | Υ | Υ |

| Operating systems | SR650 V2 | SR670 V2 | SR665 | SR650 (Xeon Gen 2) | SR650 (Xeon Gen 1) |
|---|----------|----------|-------|--------------------|--------------------|
| Red Hat Enterprise Linux 9.2 | Υ | Υ | Υ | Υ | Υ |
| Red Hat Enterprise Linux 9.3 | Υ | Υ | Υ | Υ | Ν |
| SUSE Linux Enterprise Server 15 SP1 | N | N | Y 1 | Υ | Υ |
| SUSE Linux Enterprise Server 15 SP2 | Υ | Υ | Υ | Υ | Υ |
| SUSE Linux Enterprise Server 15 SP3 | Υ | Υ | Υ | Υ | Υ |
| SUSE Linux Enterprise Server 15 SP4 | Υ | Υ | Υ | Υ | Υ |
| SUSE Linux Enterprise Server 15 SP5 | Υ | Υ | Υ | Υ | Υ |
| Ubuntu 18.04.5 LTS | Υ | Υ | N | Ν | Ν |
| Ubuntu 20.04 LTS | Υ | N | N | Ν | Ν |
| Ubuntu 22.04 LTS | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 6.5 | N | N | N | Ν | Υ |
| VMware vSphere Hypervisor (ESXi) 6.5 U1 | Ν | Ν | N | N | Υ |
| VMware vSphere Hypervisor (ESXi) 6.5 U2 | Ν | Ν | N | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 6.5 U3 | N | Ν | N | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 6.7 | N | Ν | N | N | Υ |
| VMware vSphere Hypervisor (ESXi) 6.7 U1 | N | Ν | N | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 6.7 U2 | N | Ν | N | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 6.7 U3 | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 7.0 | N | Ν | Y 1 | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 7.0 U1 | N | Ν | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 7.0 U2 | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 7.0 U3 | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 8.0 | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 8.0 U1 | Υ | Υ | Υ | Υ | Υ |
| VMware vSphere Hypervisor (ESXi) 8.0 U2 | Υ | Υ | Υ | Υ | Υ |

¹ The OS is not supported with EPYC 7003 processors.

NVIDIA GPU software

This section lists the NVIDIA software that is available from Lenovo.

- NVIDIA vGPU Software (vApps, vPC, RTX vWS)
- NVIDIA Omniverse Software (OVE)
- NVIDIA AI Enterprise Software
- NVIDIA HPC Compiler Software

NVIDIA vGPU Software (vApps, vPC, RTX vWS)

Lenovo offers the following virtualization software for NVIDIA GPUs:

Virtual Applications (vApps)

For organizations deploying Citrix XenApp, VMware Horizon RDSH or other RDSH solutions. Designed to deliver PC Windows applications at full performance. NVIDIA Virtual Applications allows users to access any Windows application at full performance on any device, anywhere. This edition is suited for users who would like to virtualize applications using XenApp or other RDSH solutions. Windows Server hosted RDSH desktops are also supported by vApps.

Virtual PC (vPC)

This product is ideal for users who want a virtual desktop but need great user experience leveraging PC Windows® applications, browsers and high-definition video. NVIDIA Virtual PC delivers a native experience to users in a virtual environment, allowing them to run all their PC applications at full performance.

NVIDIA RTX Virtual Workstation (RTX vWS)

NVIDIA RTX vWS is the only virtual workstation that supports NVIDIA RTX technology, bringing advanced features like ray tracing, Al-denoising, and Deep Learning Super Sampling (DLSS) to a virtual environment. Supporting the latest generation of NVIDIA GPUs unlocks the best performance possible, so designers and engineers can create their best work faster. IT can virtualize any application from the data center with an experience that is indistinguishable from a physical workstation — enabling workstation performance from any device.

The following license types are offered:

• Perpetual license

A non-expiring, permanent software license that can be used on a perpetual basis without the need to renew. Each Lenovo part number includes a fixed number of years of Support, Upgrade and Maintenance (SUMS).

• Annual subscription

A software license that is active for a fixed period as defined by the terms of the subscription license, typically yearly. The subscription includes Support, Upgrade and Maintenance (SUMS) for the duration of the license term.

• Concurrent User (CCU)

A method of counting licenses based on active user VMs. If the VM is active and the NVIDIA vGPU software is running, then this counts as one CCU. A vGPU CCU is independent of the connection to the VM.

The following table lists the ordering part numbers and feature codes.

Table 8. NVIDIA vGPU Software

| Part number | Feature code 7S02CTO1WW | Description |
|--------------|----------------------------|--|
| NVIDIA vApps | | |
| 7S020003WW | B1MP | NVIDIA vApps SUMS ONLY 5Yr, 1 CCU |
| 7S020004WW | B1MQ | NVIDIA vApps Subscription License 1 Year, 1 CCU |
| 7S020005WW | B1MR | NVIDIA vApps Subscription License 3 Years, 1 CCU |
| 7S02003DWW | S832 | NVIDIA vApps Subscription License 4 Years, 1 CCU |
| 7S02003EWW | S833 | NVIDIA vApps Subscription License 5 Years, 1 CCU |
| NVIDIA vPC | | |

| Part number | Feature code 7S02CTO1WW | Description |
|---------------|-------------------------|---|
| | | |
| 7S020009WW | B1MV | NVIDIA vPC SUMS 5Yr ONLY, 1 CCU |
| 7S02000AWW | B1MW | NVIDIA vPC Subscription License 1 Year, 1 CCU |
| 7S02000BWW | B1MX | NVIDIA vPC Subscription License 3 Years, 1 CCU |
| 7S02003FWW | S834 | NVIDIA vPC Subscription License 4 Years, 1 CCU |
| 7S02003GWW | S835 | NVIDIA vPC Subscription License 5 Years, 1 CCU |
| NVIDIA RTX vW | /S | |
| 7S02000FWW | B1N1 | NVIDIA RTX vWS SUMS ONLY 5Yr, 1 CCU |
| 7S02000GWW | B1N2 | NVIDIA RTX vWS Subsc Lic 1Yr 1 CCU |
| 7S02000HWW | B1N3 | NVIDIA RTX vWS Subscription License 3 Years, 1 CCU |
| 7S02000XWW | S6YJ | NVIDIA RTX vWS Subscription License 4 Years, 1 CCU |
| 7S02000YWW | S6YK | NVIDIA RTX vWS Subscription License 5 Years, 1 CCU |
| 7S02000LWW | B1N6 | NVIDIA RTX vWS EDU SUMS ONLY 5Y, 1CCU |
| 7S02000MWW | B1N7 | NVIDIA RTX vWS EDU Subscription License 1 Year, 1 CCU |

NVIDIA Omniverse Software (OVE)

B1N8

S830

S831

7S02000NWW

7S02003BWW

7S02003CWW

NVIDIA Omniverse™ Enterprise is an end-to-end collaboration and simulation platform that fundamentally transforms complex design workflows, creating a more harmonious environment for creative teams.

NVIDIA RTX vWS EDU Subscription License 3 Years, 1 CCU

NVIDIA RTX vWS EDU Subscription License 4 Years, 1 CCU

NVIDIA RTX vWS EDU Subscription License 5 Years, 1 CCU

NVIDIA and Lenovo offer a robust, scalable solution for deploying Omniverse Enterprise, accommodating a wide range of professional needs. This document details the critical components, deployment options, and support available, ensuring an efficient and effective Omniverse experience.

Deployment options cater to varying team sizes and workloads. Using Lenovo NVIDIA-Certified Systems™ and Lenovo OVX nodes which are meticulously designed to manage scale and complexity, ensures optimal performance for Omniverse tasks.

Deployment options include:

- Workstations: NVIDIA-Certified Workstations with RTX 6000 Ada GPUs for desktop environments.
- Data Center Solutions: Deployment with Lenovo OVX nodes or NVIDIA-Certified Servers equipped with L40, L40S or A40 GPUs for centralized, high-capacity needs.

NVIDIA Omniverse Enterprise includes the following components and features:

- Platform Components: Kit, Connect, Nucleus, Simulation, RTX Renderer.
- Foundation Applications: USD Composer, USD Presenter.
- Omniverse Extensions: Connect Sample & SDK.
- Integrated Development Environment (IDE)
- Nucleus Configuration: Workstation, Enterprise Nucleus Server (supports up to 8 editors per scene);
 Self-Service Public Cloud Hosting using Containers.
- Omniverse Farm: Supports batch workloads up to 8 GPUs.
- Enterprise Services: Authentication (SSO/SSL), Navigator Microservice, Large File Transfer, User

Accounts SAML/Account Directory.

- User Interface: Workstation & IT Managed Launcher.
- Support: NVIDIA Enterprise Support.
- Deployment Scenarios: Desktop to Data Center: Workstation deployment for building and designing, with options for physical or virtual desktops. For batch tasks, rendering, and SDG workloads that require headless compute, Lenovo OVX nodes are recommended.

The following part numbers are for a subscription license which is active for a fixed period as noted in the description. The license is for a named user which means the license is for named authorized users who may not re-assign or share the license with any other person.

Table 9. NVIDIA Omniverse Software (OVE)

| Part number | Feature 7S02CTO1WW | NVIDIA part number | Description |
|-------------|-----------------------|--------------------|--|
| 7S02003ZWW | SCX0 | 721-OV7006+P3CMI12 | NVIDIA Omniverse Enterprise Subscription per GPU, 1 Year |
| 7S020042WW | SCX3 | 721-OV7006+P3CMI36 | NVIDIA Omniverse Enterprise Subscription per GPU, 3 Years |
| 7S020044WW | SD5T | 721-OV7006+P3CMI60 | NVIDIA Omniverse Enterprise Subscription per GPU, 5 Year |
| 7S020041WW | SCX2 | 721-OV7006+P3INI12 | NVIDIA Omniverse Enterprise Subscription per GPU, INC, 1 Year |
| 7S020040WW | SCX1 | 721-OV7006+P3EDI12 | NVIDIA Omniverse Enterprise Subscription per GPU, EDU, 1 Year |
| 7S020043WW | SCX4 | 721-OV7006+P3EDI36 | NVIDIA Omniverse Enterprise Subscription per GPU, EDU, 3 Years |
| 7S020045WW | SD5U | 721-OV7006+P3EDI60 | NVIDIA Omniverse Enterprise Subscription per GPU EDU, 5 Year |

NVIDIA AI Enterprise Software

Lenovo offers the NVIDIA AI Enterprise (NVAIE) cloud-native enterprise software. NVIDIA AI Enterprise is an end-to-end, cloud-native suite of AI and data analytics software, optimized, certified, and supported by NVIDIA to run on VMware vSphere and bare-metal with NVIDIA-Certified Systems™. It includes key enabling technologies from NVIDIA for rapid deployment, management, and scaling of AI workloads in the modern hybrid cloud.

NVIDIA AI Enterprise is licensed on a per-GPU basis. NVIDIA AI Enterprise products can be purchased as either a perpetual license with support services, or as an annual or multi-year subscription.

- The perpetual license provides the right to use the NVIDIA AI Enterprise software indefinitely, with no expiration. NVIDIA AI Enterprise with perpetual licenses must be purchased in conjunction with one-year, three-year, or five-year support services. A one-year support service is also available for renewals.
- The subscription offerings are an affordable option to allow IT departments to better manage the flexibility of license volumes. NVIDIA AI Enterprise software products with subscription includes support services for the duration of the software's subscription license

The features of NVIDIA AI Enterprise Software are listed in the following table.

Table 10. Features of NVIDIA AI Enterprise Software (NVAIE)

| Features | Supported in NVIDIA AI Enterprise |
|---|-----------------------------------|
| Per GPU Licensing | Yes |
| Compute Virtualization | Supported |
| Windows Guest OS Support | No support |
| Linux Guest OS Support | Supported |
| Maximum Displays | 1 |
| Maximum Resolution | 4096 x 2160 (4K) |
| OpenGL and Vulkan | In-situ Graphics only |
| CUDA and OpenCL Support | Supported |
| ECC and Page Retirement | Supported |
| MIG GPU Support | Supported |
| Multi-vGPU | Supported |
| NVIDIA GPUDirect | Supported |
| Peer-to-Peer over NVLink | Supported |
| GPU Pass Through Support | Supported |
| Baremetal Support | Supported |
| Al and Data Science applications and Frameworks | Supported |
| Cloud Native ready | Supported |

Note: Maximum 10 concurrent VMs per product license

The following table lists the ordering part numbers and feature codes.

Table 11. NVIDIA AI Enterprise Software (NVAIE)

| Part number | Feature code 7S02CTO1WW | Description | |
|------------------------------------|----------------------------|--|--|
| Al Enterprise Perpetual License | | | |
| 7S02001BWW | S6YY | NVIDIA AI Enterprise Perpetual License and Support per GPU, 5 Years | |
| 7S02001EWW | S6Z1 | NVIDIA AI Enterprise Perpetual License and Support per GPU, EDU, 5 Years | |
| Al Enterprise Subscription License | | | |
| 7S02001FWW | S6Z2 | NVIDIA AI Enterprise Subscription License and Support per GPU, 1 Year | |
| 7S02001GWW | S6Z3 | NVIDIA AI Enterprise Subscription License and Support per GPU, 3 Years | |
| 7S02001HWW | S6Z4 | NVIDIA AI Enterprise Subscription License and Support per GPU, 5 Years | |
| 7S02001JWW | S6Z5 | NVIDIA AI Enterprise Subscription License and Support per GPU, EDU, 1 Year | |
| 7S02001KWW | S6Z6 | NVIDIA AI Enterprise Subscription License and Support per GPU, EDU, 3 Years | |
| 7S02001LWW | S6Z7 | NVIDIA AI Enterprise Subscription License and Support per GPU, EDU, 5 Years | |

Find more information in the NVIDIA AI Enterprise Sizing Guide.

NVIDIA HPC Compiler Software

Table 12. NVIDIA HPC Compiler

| Part number | Feature code 7S09CTO6WW | Description | |
|---------------------------------------|----------------------------|---|--|
| HPC Compiler Support Services | | | |
| 7S090014WW | S924 | NVIDIA HPC Compiler Support Services, 1 Year | |
| 7S090015WW | S925 | NVIDIA HPC Compiler Support Services, 3 Years | |
| 7S09002GWW | S9UQ | NVIDIA HPC Compiler Support Services, 5 Years | |
| 7S090016WW | S926 | NVIDIA HPC Compiler Support Services, EDU, 1 Year | |
| 7S090017WW | S927 | NVIDIA HPC Compiler Support Services, EDU, 3 Years | |
| 7S09002HWW | S9UR | NVIDIA HPC Compiler Support Services, EDU, 5 Years | |
| 7S090018WW | S928 | NVIDIA HPC Compiler Support Services - Additional Contact, 1 Year | |
| 7S09002JWW | S9US | NVIDIA HPC Compiler Support Services - Additional Contact, 3 Years | |
| 7S09002KWW | S9UT | NVIDIA HPC Compiler Support Services - Additional Contact, 5 Years | |
| 7S090019WW | S929 | NVIDIA HPC Compiler Support Services - Additional Contact, EDU, 1 Year | |
| 7S09002LWW | S9UU | NVIDIA HPC Compiler Support Services - Additional Contact, EDU, 3 Years | |
| 7S09002MWW | S9UV | NVIDIA HPC Compiler Support Services - Additional Contact, EDU, 5 Years | |
| HPC Compiler Premier Support Services | | | |
| 7S09001AWW | S92A | NVIDIA HPC Compiler Premier Support Services, 1 Year | |
| 7S09002NWW | S9UW | NVIDIA HPC Compiler Premier Support Services, 3 Years | |
| 7S09002PWW | S9UX | NVIDIA HPC Compiler Premier Support Services, 5 Years | |
| 7S09001BWW | S92B | NVIDIA HPC Compiler Premier Support Services, EDU, 1 Year | |
| 7S09002QWW | S9UY | NVIDIA HPC Compiler Premier Support Services, EDU, 3 Years | |
| 7S09002RWW | S9UZ | NVIDIA HPC Compiler Premier Support Services, EDU, 5 Years | |
| 7S09001CWW | S92C | NVIDIA HPC Compiler Premier Support Services - Additional Contact, 1 Year | |
| 7S09002SWW | S9V0 | NVIDIA HPC Compiler Premier Support Services - Additional Contact, 3 Years | |
| 7S09002TWW | S9V1 | NVIDIA HPC Compiler Premier Support Services - Additional Contact, 5 Years | |
| 7S09001DWW | S92D | NVIDIA HPC Compiler Premier Support Services - Additional Contact, EDU, 1 Year | |
| 7S09002UWW | S9V2 | NVIDIA HPC Compiler Premier Support Services - Additional Contact, EDU, 3 Years | |
| 7S09002VWW | S9V3 | NVIDIA HPC Compiler Premier Support Services - Additional Contact, EDU, 5 Years | |

Auxiliary power cables

The A10 option part number does not ship with auxiliary power cables. Cables are server-specific due to length requirements. For CTO orders, auxiliary power cables are derived by the configurator. For field upgrades, cables will need to be ordered separately as listed in the table below.

Table 13. Auxiliary power cables for A10 (click images to show larger versions)

Auxiliary power cable needed with the SR665, SR650 V2

460mm 8-pin (2x4) cable Option:

SR650 V2: Part of 4H47A80491 or 4H47A38666*

SR665: Part of

4M17A80478 or 4M17A11759 Cable kit: 4X97A81933 Feature: BACW SBB: SBB7A21686 Base: SC17A60223 FRU: 02YE421

Note: Use this cable only when there is one A10 GPU per riser. For two GPUs per riser, use Y-cable

n+5 30±5(3x)

BC09/SBB7A23757.

* SR650 V2: For additional requirements, see the SR650 V2 NVIDIA A10 requirements.

460mm Y-cable 8-pin (2x4) to 2x 8-pin (2x4)

Option:

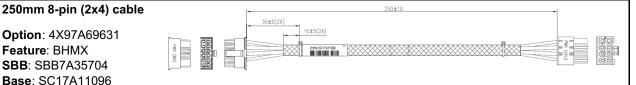
SR650 V2: Part of 4H47A80491 or 4H47A38666 SR665: Part of 4M17A80478 or 4M17A11759

Cable kit: 4X97A81933 Feature: BC09 SBB: SBB7A23757 Base: SC17A60224 FRU: 02YE422

Note: Use this Y-cable when there are two A10 GPUs in a riser.

Auxiliary power cable needed with the SR670 V2

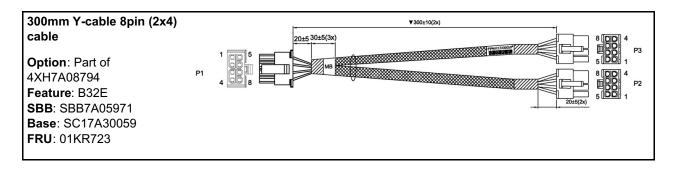
Option: 4X97A69631 Feature: BHMX **SBB**: SBB7A35704 Base: SC17A11096 FRU: 02YF205



▼460±10

Auxiliary power cable needed with the SR650





Regulatory approvals

The NVIDIA A10 GPU has the following regulatory approvals:

- RCM
- BSMI
- CE
- FCC
- ICES
- KCC
- cUL, UL
- VCCI

Operating environment

The NVIDIA A10 GPU has the following operating characteristics:

- Ambient temperature
 - Operational: 0°C to 50°C (-5°C to 55°C for short term*)
 - Storage: -40°C to 75°C
- Relative humidity:
 - Operational: 5-85% (5-93% short term*)
 - Storage: 5-95%

Warranty

One year limited warranty. When installed in a Lenovo server, the GPU assumes the server's base warranty and any warranty upgrades.

Related publications

For more information, refer to these documents:

- ThinkSystem and ThinkAgile GPU Summary: https://lenovopress.lenovo.com/lp0768-thinksystem-thinkagile-gpu-summary
- ServerProven compatibility: https://serverproven.lenovo.com/
- NVIDIA A10 product page: https://www.nvidia.com/en-us/data-center/products/a10-gpu/
- NVIDIA Ampere Architecture page https://www.nvidia.com/en-us/data-center/ampere-architecture/

^{*} A period not more than 96 hours consecutive, not to exceed 15 days per year.

Related product families

Product families related to this document are the following:

• GPU adapters

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc. 8001 Development Drive Morrisville, NC 27560 U.S.A.

Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

© Copyright Lenovo 2024. All rights reserved.

This document, LP1816, was created or updated on September 15, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at: https://lenovopress.lenovo.com/LP1816
- Send your comments in an e-mail to: comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1816.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at https://www.lenovo.com/us/en/legal/copytrade/.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo®

ServerProven®

ThinkAgile®

ThinkSystem®

The following terms are trademarks of other companies:

AMD is a trademark of Advanced Micro Devices, Inc.

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Linux® is the trademark of Linus Torvalds in the U.S. and other countries.

Microsoft®, Windows Server®, and Windows® are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.