

ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU

Product Guide

The ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU delivers unprecedented visual computing performance for the data center and provides revolutionary neural graphics, compute, and AI capabilities to accelerate the most demanding visual computing workloads. With up to twice the performance of the previous generation at the same power, the NVIDIA L40 is uniquely suited to provide the visual computing power and performance required by the modern data center.

The following figure shows the ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU.



Figure 1. ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU

Did you know?

The NVIDIA L40, based on the NVIDIA Ada Lovelace GPU architecture features new generation RT cores and Tensor cores, delivering in combination over a petaflop of inferencing performance. These new features are combined with the latest generation CUDA Cores and 48GB of graphics memory to accelerate visual computing workloads from high-performance virtual workstation instances to large-scale digital twins in NVIDIA Omniverse.

Part number information

The following table shows the part numbers for the GPU.

Table 1. Ordering information

Part number	Feature code	Description
4X67A84823	BT87	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU

The option part number includes the following:

- One NVIDIA L40 GPU with full-height (3U) adapter bracket attached
- Documentation

Features

From virtual workstation application to large-scale modeling and simulation, modern visual computing and scientific workflows are growing in both complexity and quantity.

Enterprises need data center technology that can deliver extreme performance and scale with versatile capabilities to conquer the diverse computing demands of these increasingly complex workloads.

The NVIDIA® L40 GPU delivers unprecedented visual computing performance for the data center, providing next-generation graphics, compute, and AI capabilities. Built on the revolutionary NVIDIA Ada Lovelace architecture, the NVIDIA L40 harnesses the power of the latest generation RT, Tensor, and CUDA cores to deliver groundbreaking visualization and compute performance for the most demanding data center workloads.

Next-Generation Graphics and Compute Performance

- 142 third-generation RT Cores and industry-leading 48 GB of GDDR6 memory: The NVIDIA L40 delivers up to twice the real-time ray-tracing performance of the previous Ampere generation to enable high-performance ray tracing of complex geometry and the creation of beautifully detailed, photorealistic models and scenes. Accelerate high-fidelity creative workflows, including real-time, full-fidelity, interactive rendering, 3D design, video streaming, and virtual
- Faster rendering and smoother frame rates with NVIDIA Deep Learning Super Sampling 3 (DLSS 3): This breakthrough frame generation technology leverages deep learning and the latest hardware innovations within the Ada Lovelace Architecture and L40 GPU, including 4th-Generation Tensor Cores and a new Optical Flow Accelerator, to boost rendering performance, deliver higher FPS, and significantly improve
- The NVIDIA L40 includes 18,176 Ada Lovelace GPU architecture CUDA cores delivering over 2X the single-precision floating-point (FP32) performance to accelerate 3D graphics, simulation and physics
- Breakthrough media acceleration capabilities: The NVIDIA L40 takes streaming and video content workloads to the next level with three video encode and three video decode engines. With the addition of AV1 encoding, the L40 delivers up to twice the performance and improved TCO for broadcast streaming, video production, and transcription
- Incredible training and inference capabilities with 90.5 TFLOPs of FP32 performance. The NVIDIA L40 includes 568 fourth-generation Tensor Cores providing up to 2X faster AI training performance than the previous generation with FP16 precision. Support for the new FP8 data format for inferencing provides more than 4X faster performance than the previous generation and reduces data memory usage by half (compared to FP16 data format)

Data Center-Ready

- The NVIDIA L40 is optimized for 24x7 enterprise data center operations and is designed, built, extensively tested, and supported exclusively by NVIDIA to ensure maximum

- Secure boot with root of trust (RoT) technology within the GPU provide an additional layer of security for data centers. The NVIDIA L40 meets the latest data center standards and is NEBS Level 3
- Passively cooled, full-height, full-length (FHFL), dual-slot design
- The NVIDIA L40 is capable of 300W maximum board power and driven by power-efficient hardware and components selected for optimum performance, durability, and

Technical specifications

The NVIDIA L40 GPU has the following specifications:

- Form factor
 - PCIe Full Height Full Length adapter (4.4-in x 10.5-in), Double-width (dual slot)
 - NVIDIA Form Factor 5.5
- Host interface:
 - PCIe 4.0 x16
 - MSI-X interrupt messaging protocol (MSI not supported)
 - PCIe Lane Polarity Inversion and Lane Reversal
- Single Root I/O Virtualization (SR-IOV) support
 - 256 virtual functions (VFs)
 - ARI Forwarding
- Hardware Root of Trust
 - Secure boot
 - Secure firmware upgrade
 - Firmware rollback protection
 - Support for in-band firmware update disable (established after each GPU reset)
 - Secure application processor recovery

The following table lists the GPU processing specifications and performance of the NVIDIA L40.

Table 2. Specifications of the ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU

Feature	Specification
GPU Architecture	NVIDIA Ada Lovelace
NVIDIA CUDA Parallel Processing Cores	18,176
NVIDIA Tensor Cores (4th gen)	568
NVIDIA RT Cores (3rd Gen)	142
Peak FP32 performance (non-Tensor)	90.5 TFLOPS
Peak FP16 Tensor performance with FP16 Accumulate	181.05 TFLOPS, 362.1 TFLOPS*
Peak Tensor Float 32 (TF32) performance	90.5 TFLOPS, 181 TFLOPS*
Peak Bfloat16 (BF16) performance with FP32 Accumulate	181.05 TFLOPS, 362.1 TFLOPS*
Peak FP8 Tensor performance	362 TFLOPS, 724 TFLOPS*
Peak Integer Performance	INT8: 362 TOPS, 724 TOPS* INT4: 724 TOPS, 1448 TOPS*
RT Core performance	209 TFLOPS
GPU Memory	48 GB GDDR6
Memory Bandwidth	864 GB/s
ECC	Yes
NVIDIA NVLink	No support
System Interface	PCIe Gen 4, x16 lanes
Form Factor	PCIe full height/length, double width (10.5" x 4.4")
Multi-Instance GPU (MIG)	No support
Max Power Consumption	300 W
Thermal Solution	Passive
vGPU Software Support	NVIDIA vPC/vApps, NVIDIA RTX Virtual Workstation (vWS)
Display connectors	4x DisplayPort 1.4a
Max Simultaneous Displays	Up to four 5K Monitors at 60Hz per card or dual 8K displays @ 60Hz (requires DisplayPort 1.4 DSC); Each display port can support 4K at 120 Hz with 30-bit color
Graphics APIs	DirectX 12 Ultimate, Shader Model 6.6, OpenGL 4.6, Vulkan 1.3
Compute APIs	CUDA 12.0, Direct Compute, OpenCL 3.0

* With structural sparsity enabled

Server support

The following tables list the ThinkSystem servers that are compatible.

Table 3. Server support (Part 1 of 4)

Part Number	Description	2S AMD V3		2S Intel V3		4S 8S Intel V3		Multi Node		GPU Rich		1S V3					
		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)	SR850 V3 (7D97 / 7D96)	SR860 V3 (7D94 / 7D93)	SR950 V3 (7DC5 / 7DC4)	SD535 V3 (7DD8 / 7DD1)	SD530 V3 (7DDA / 7DD3)	SD550 V3 (7DD9 / 7DD2)	SR670 V2 (7Z22 / 7Z23)	SR675 V3 (7D9Q / 7D9R)	ST250 V3 (7DCF / 7DCE)
4X67A84823	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU	N	3	N	3	N	3	2	4	N	N	N	N	8	8	N	N

Table 4. Server support (Part 2 of 4)

Part Number	Description	Edge				Super Computing				1S Intel V2		2S Intel V2					
		SE350 (7Z46 / 7D1X)	SE350 V2 (7DA9)	SE360 V2 (7DAM)	SE450 (7D8T)	SE455 V3 (7DBY)	SD665 V3 (7D9P)	SD665-N V3 (7DAZ)	SD650 V3 (7D7M)	SD650-I V3 (7D7L)	SD650-N V3 (7D7N)	ST50 V2 (7D8K / 7D8J)	ST250 V2 (7D8G / 7D8F)	SR250 V2 (7D7R / 7D7Q)	ST650 V2 (7Z75 / 7Z74)	SR630 V2 (7Z70 / 7Z71)	SR650 V2 (7Z72 / 7Z73)
4X67A84823	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU	N	N	N	2	2	N	N	N	N	N	N	N	N	N	N	3

Table 5. Server support (Part 3 of 4)

Part Number	Description	AMD V1				Dense V2				4S V2		8S		4S V1		1S Intel V1			
		SR635 (7Y98 / 7Y99)	SR655 (7Y00 / 7Z01)	SR655 Client OS	SR645 (7D2Y / 7D2X)	SR665 (7D2W / 7D2V)	SD630 V2 (7D1K)	SD650 V2 (7D1M)	SD650-N V2 (7D1N)	SN550 V2 (7Z69)	SR850 V2 (7D31 / 7D32)	SR860 V2 (7Z59 / 7Z60)	SR950 (7X11 / 7X12)	SR850 (7X18 / 7X19)	SR850P (7D2F / 2D2G)	SR860 (7X69 / 7X70)	ST50 (7Y48 / 7Y50)	ST250 (7Y45 / 7Y46)	SR150 (7Y54)
4X67A84823	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU	N	N	N	N	3	N	N	N	N	N	N	N	N	N	N	N	N	N

Table 6. Server support (Part 4 of 4)

Part Number	Description	2S Intel V1								Dense V1			
		ST550 (7X09 / 7X10)	SR530 (7X07 / 7X08)	SR550 (7X03 / 7X04)	SR570 (7Y02 / 7Y03)	SR590 (7X98 / 7X99)	SR630 (7X01 / 7X02)	SR650 (7X05 / 7X06)	SR670 (7Y36 / 7Y37)	SD530 (7X21)	SD650 (7X58)	SN550 (7X16)	SN850 (7X15)
4X67A84823	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU	N	N	N	N	N	N	N	N	N	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 L40 48GB PCIe Gen4 Passive GPU, 4X67A84823

Operating systems	SE450	SE455 V3	SR650 V3 (4th Gen Xeon)	SR650 V3 (5th Gen Xeon)	SR655 V3	SR665 V3	SR675 V3	SR850 V3	SR860 V3	SR650 V2	SR670 V2	SR665
Microsoft Windows 10	N	N	N	Y	Y	Y	N	N	N	N	N	N
Microsoft Windows 11	N	N	N	Y	Y	Y	N	N	N	N	N	N
Microsoft Windows Server 2016	N	N	N	N	N	N	N	N	N	Y	Y	Y
Microsoft Windows Server 2019	Y	Y	Y	Y	Y	Y	Y	Y ¹	Y ¹	Y	Y	Y
Microsoft Windows Server 2022	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 7.9	Y	N	N	N	N	N	N	N	N	Y	Y	Y ²
Red Hat Enterprise Linux 8.3	N	N	N	N	N	N	N	N	N	Y	Y	Y
Red Hat Enterprise Linux 8.4	Y	N	N	N	N	N	N	N	N	Y	Y	Y
Red Hat Enterprise Linux 8.5	Y	N	N	N	N	N	N	N	N	Y	Y	Y
Red Hat Enterprise Linux 8.6	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 8.7	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 8.8	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 9.0	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 9.1	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Red Hat Enterprise Linux 9.2	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
SUSE Linux Enterprise Server 15 SP3	N	N	N	N	N	N	N	N	N	Y	Y	Y
SUSE Linux Enterprise Server 15 SP4	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
SUSE Linux Enterprise Server 15 SP5	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Ubuntu 18.04.5 LTS	N	N	N	N	N	N	N	N	N	Y	Y	N
Ubuntu 18.04.6 LTS	Y	N	N	N	N	N	N	N	N	N	N	N
Ubuntu 20.04 LTS	N	N	N	N	N	N	N	N	N	Y	N	N
Ubuntu 20.04.5 LTS	Y	Y	N	N	Y	Y	Y	Y	Y	N	N	N
Ubuntu 22.04 LTS	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
VMware vSphere Hypervisor (ESXi) 6.7 U3	N	N	N	N	N	N	N	N	N	Y	Y	Y
VMware vSphere Hypervisor (ESXi) 7.0 U3	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
VMware vSphere Hypervisor (ESXi) 8.0	Y	N	Y	N	Y	Y	N	N	N	Y	Y	Y
VMware vSphere Hypervisor (ESXi) 8.0 U1	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
VMware vSphere Hypervisor (ESXi) 8.0 U2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

¹ For limitation, please refer [Support Tip TT1591](#)

² 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 \(vApps, vPC, RTX vWS, and vCS\)](#)
- [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, AI-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	Description
NVIDIA vApps		
7S020003WW	B1MP	NVIDIA vApps Perpetual License and SUMS 5Yr, 1 CCU

Part number	Feature code 7S02CTO1WW	Description
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		
7S020009WW	B1MV	NVIDIA vPC Perpetual License and SUMS 5Yr, 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 vWS		
7S02000FWW	B1N1	NVIDIA RTX vWS Perpetual License and SUMS 5Yr, 1 CCU
7S02000GWW	B1N2	NVIDIA RTX vWS Subscription License 1 Year, 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 Perpetual License and SUMS 5Yr, 1 CCU
7S02000MWW	B1N7	NVIDIA RTX vWS EDU Subscription License 1 Year, 1 CCU
7S02000NWW	B1N8	NVIDIA RTX vWS EDU Subscription License 3 Years, 1 CCU
7S02003BWW	S830	NVIDIA RTX vWS EDU Subscription License 4 Years, 1 CCU
7S02003CWW	S831	NVIDIA RTX vWS EDU Subscription License 5 Years, 1 CCU

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 9. Features of NVIDIA AI Enterprise Software (NVAIE)

Features	Supported in NVIDIA AI Enterprise
Per GPU Licensing	Yes

Features	Supported in NVIDIA AI Enterprise
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
AI 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 10. NVIDIA AI Enterprise Software (NVAIE)

Part number	Feature code 7S02CTO1WW	Description
AI 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
AI 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 11. NVIDIA HPC Compiler

Part number	Feature code 7S09CTO6WW	Description
HPC Compiler Support Services		

Part number	Feature code 7S09CTO6WW	Description
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 L40 option part number does not ship with auxiliary power cables. Cables are server-specific due to length requirements and the connector on the server end of the cable. 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.

Tip: The names of the cable options below may only include the H100 GPU, however these cables are also supported with the L40.

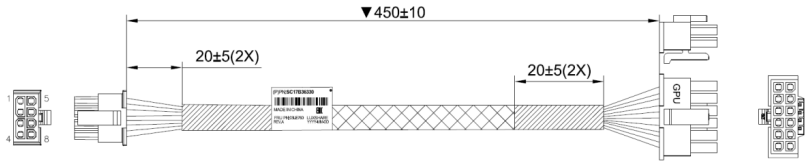
Table 12. Auxiliary power cables for NVIDIA L40

Auxiliary power cable needed with the SE455 V3 and SE450
--

450mm 16-pin (2x6+4) cable

Option:

- SE455 V3: 4X97A89288, ThinkEdge NVIDIA L40 GPU Power Cable Kit (contains 2 cables)
- SE450: 4X97A89288, ThinkEdge NVIDIA L40 GPU Power Cable Kit (contains 2 cables)



Feature: BVVG

SBB: SBB7A74096

Base: SC17B36330

FRU: 03LE760

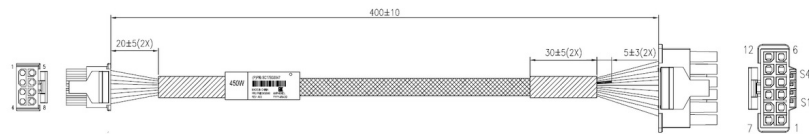
Additional components may be required when installing the GPU. See the [SE455 V3 product guide](#) for details.

Auxiliary power cable needed with the SR665, SR650 V2, SR650 V3, SR655 V3, SR665 V3

400mm 16-pin (2x6+4) cable

Option:

- SR665: 4X97A85028, ThinkSystem 400mm 2x6+4 GPU Power Cable
- SR650 V2: 4X97A85028, ThinkSystem 400mm 2x6+4 GPU Power Cable
- SR650 V3: 4X67A82883, ThinkSystem SR650 V3 GPU Full Length Thermal Option Kit*
- SR655 V3: 4X67A86438, ThinkSystem SR655 V3 GPU Enablement Kit*
- SR665 V3: 4X67A85856, ThinkSystem SR665 V3 GPU Full Length Thermal Option Kit*



Feature: BRWK

SBB: SBB7A66338

Base: SC17B33047

FRU: 03KM846

* The option part number is for the thermal kit and includes other components needed to install the GPU. See the respective server product guide details.

Auxiliary power cable needed with the SR675 V3

235mm 16-pin (2x6+4) cable

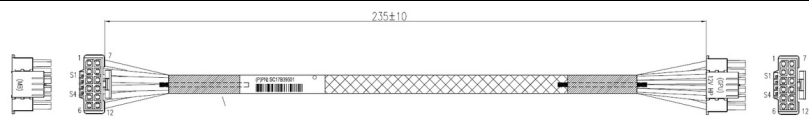
Option: 4X97A84510, ThinkSystem SR675 V3 Supplemental Power Cable for H100 GPU Option

Feature: BSD2

SBB: SBB7A65299

Base: SC17B39301

FRU: 03LE554



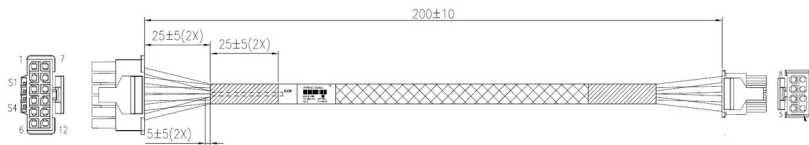
Auxiliary power cable needed with the SR670 V2

215mm 16-pin (2x6+4) cable
Option: 4X97A85027, ThinkSystem
 SR670 V2 H100/L40 GPU Option
 Power Cable
Feature: BRWL
SBB: SBB7A66339
Base: SC17B33046
FRU: 03KM845



Auxiliary power cable needed with the SR850 V3, SR860 V3

200mm 16-pin (2x6+4) cable
Option: 4X97A88016, ThinkSystem
 SR850 V3/SR860 V3 H100 GPU
 Power Cable Option Kit
Feature: BW28
SBB: SBB7A72759
Base: SC17B40604
FRU: 03LF915



Regulatory approvals

The NVIDIA L40 GPU has the following regulatory approvals:

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

Operating environment

The NVIDIA L40 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%

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

Warranty

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

Seller training courses

The following sales training courses are offered for employees and partners (login required). Courses are listed in date order.

1. Generative AI Overview Foundational
2024-02-16 | 17 minutes | Employees Only

It seems the whole world is excited about Generative AI, and while some of it is just hype, it has become clear that Generative AI has the potential to revolutionize many aspects of our personal and professional lives. In this brief NVIDIA course, we'll explore one aspect of the Generative AI excitement, the value you get from Generative AI technology. We will discuss what Generative AI is, how it works, and how enterprises are planning to use this technology.

By the end of this course, you will be able to discuss the Generative AI market trends and the challenges in this space with your customers. And you will be able to explain what Generative AI is and how the technology works to help enterprises unlock new opportunities for business.

Published: 2024-02-16
Length: 17 minutes
Employee link: [Grow@Lenovo](#)
Course code: DAINVD106

2. Industry Use Cases in Modern Computing Foundational
2024-02-16 | 9 minutes | Employees Only

As GPU powered computing continues to improve exponentially, applications that were once science fiction are becoming best practice. This is an introductory NVIDIA course that explores some exciting industry focused use cases that are providing companies with faster time to insight, productivity at scale and a great ROI.

By the end of this course, you will be able to explain how companies in a few key industry verticals are benefiting from a variety of accelerated compute use cases.

Published: 2024-02-16
Length: 9 minutes
Employee link: [Grow@Lenovo](#)
Course code: DAINVD105

3. Introduction to Artificial Intelligence Foundational
2024-02-16 | 10 minutes | Employees Only

This NVIDIA course aims to answer questions such as, what is AI and why are enterprises so interested in it? and how does AI happen, why are GPUs so important for it, and what does a good AI solution look like?

By the end of this training, you should be able to describe AI and relate it to some common enterprise use cases. You'll know the difference between training and inference and be able to visualize a typical AI workflow. More importantly, you'll understand the difficulties of traditional CPU-based AI and appreciate why businesses would benefit greatly by adopting GPU-accelerated workflows. Finally, you'll also understand what features contribute to an awesome AI solution and why customers respect and enjoy NVIDIA's solutions.

Published: 2024-02-16
Length: 10 minutes
Employee link: [Grow@Lenovo](#)
Course code: DAINVD104

4. **GPU Fundamentals Foundational**
2024-02-16 | 10 minutes | Employees Only

This NVIDIA course introduces you to two devices that a computer typically uses to process information, the CPU and the GPU. We'll discuss their differences and look at how the GPU overcomes the limitations of the CPU. Once you understand the power and advantages of GPU processing, we will talk about the value GPUs bring to modern-day enterprise computing.

By the end of this course, you should know the difference between serial and parallel processing. You will be able to explain what a GPU is in very simple terms and explain the value that GPUs bring to enterprises. Additionally, you'll become familiar with the typical GPU-accelerated enterprise workloads and list one or two use cases under them. By the time you exit this course, you should be able to target various GPU-accelerated computing opportunities with the right NVIDIA GPU.

Published: 2024-02-16
Length: 10 minutes
Employee link: Grow@Lenovo
Course code: DAINVD103

5. **Partner Technical Webinar – NVidia**
2023-12-11 | 60 minutes | Employees and Partners

In this 60-minute replay, Brad Davidson of Nvidia will help us recognize AI Trends, and Discuss Industry Verticals Marketing.

Published: 2023-12-11
Length: 60 minutes
Employee link: Grow@Lenovo
Partner link: [Lenovo Partner Learning](#)
Course code: 120823

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:
<http://www.lenovo.com/us/en/serverproven>
- NVIDIA L40 product page:
<https://www.nvidia.com/en-us/data-center/l40/>

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, LP1718, was created or updated on December 5, 2023.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
<https://lenovopress.lenovo.com/LP1718>
- Send your comments in an e-mail to:
comments@lenovopress.com

This document is available online at <https://lenovopress.lenovo.com/LP1718>.

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:

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®, DirectX®, 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.