



# ThinkSystem NVIDIA H100 PCIe Gen5 GPUs Product Guide

The ThinkSystem NVIDIA H100 GPU delivers unprecedented performance, scalability, and security for every workload. The GPUs use breakthrough innovations in the NVIDIA Hopper™ architecture to deliver industry-leading conversational AI, speeding up large language models by 30X over the previous generation.

The NVIDIA H100 GPU features fourth-generation Tensor Cores and the Transformer Engine with FP8 precision, further extending NVIDIA's market-leading AI leadership with up to 9X faster training and an incredible 30X inference speedup on large language models. For high-performance computing (HPC) applications, The GPUs triple the floating-point operations per second (FLOPS) of FP64 and add dynamic programming (DPX) instructions to deliver up to 7X higher performance.

The following figure shows the ThinkSystem NVIDIA H100 GPU in the double-width PCIe adapter form factor.



Figure 1. ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU

## Did you know?

The NVIDIA H100 family is available in both double-wide PCIe adapter form factor and in SXM form factor. The latter is used in Lenovo's Neptune direct-water-cooled ThinkSystem SD665-N V3 server for the ultimate in GPU performance and heat management.

The NVIDIA H100 NVL Tensor Core GPU is optimized for Large Language Model (LLM) Inferences, with its high compute density, high memory bandwidth, high energy efficiency, and unique NVLink architecture.

#### Part number information

The following table shows the part numbers for the ThinkSystem NVIDIA H100 GPU.

Not available in China, Hong Kong and Macau: The H100 GPUs are not available in China, Hong Kong and Macau.

Table 1. Ordering information

Part number	Feature code	Description
Double-wide PCI	e adapter form fa	ctor
4X67A89325	BXAK	ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU
SXM form factor		
CTO only	C1HL	ThinkSystem NVIDIA HGX H100 80GB 700W 8-GPU Board
CTO only	BQQV	ThinkSystem NVIDIA H100 SXM5 700W 80G GPU Board
CTO only	BUBB	ThinkSystem NVIDIA H100 SXM5 700W 94G HBM2e GPU Board
NVLink bridge (fo	or PCIe adapters	only, not SXM)
4X67A71309	BG3F	ThinkSystem NVIDIA Ampere NVLink 2-Slot Bridge (3 required per pair of GPUs)

The PCIe option part numbers includes the following:

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

The following figure shows the NVIDIA H100 SXM5 8-GPU Board with heatsinks installed in the ThinkSystem SR680a V3 and ThinkSystem SR685a V3 servers.



Figure 2. NVIDIA H100 SXM5 8-GPU Board in the ThinkSystem SR680a V3 and SR685a V3 servers

#### **Features**

The ThinkSystem NVIDIA H100 GPU delivers high performance, scalability, and security for every workload. The GPU uses breakthrough innovations in the NVIDIA Hopper™ architecture to deliver industry-leading conversational AI, speeding up large language models (LLMs) by 30X over the previous generation.

The PCIe versions of the NVIDIA H100 GPUs include a five-year software subscription, with enterprise support, to the NVIDIA AI Enterprise software suite, simplifying AI adoption with the highest performance. This ensures organizations have access to the AI frameworks and tools they need to build accelerated AI workflows such as AI chatbots, recommendation engines, vision AI, and more.

The NVIDIA H100 GPU features fourth-generation Tensor Cores and the Transformer Engine with FP8 precision, further extending NVIDIA's market-leading AI leadership with up to 9X faster training and an incredible 30X inference speedup on large language models. For high-performance computing (HPC) applications, the GPU triples the floating-point operations per second (FLOPS) of FP64 and adds dynamic programming (DPX) instructions to deliver up to 7X higher performance. With second-generation Multi-Instance GPU (MIG), built-in NVIDIA confidential computing, and NVIDIA NVLink Switch System, the NVIDIA H100 GPU securely accelerates all workloads for every data center from enterprise to exascale.

#### Key features of the NVIDIA H100 GPU:

#### NVIDIA H100 Tensor Core GPU

Built with 80 billion transistors using a cutting-edge TSMC 4N process custom tailored for NVIDIA's accelerated compute needs, H100 is the world's most advanced chip ever built. It features major advances to accelerate AI, HPC, memory bandwidth, interconnect, and communication at data center scale.

#### • Transformer Engine

The Transformer Engine uses software and Hopper Tensor Core technology designed to accelerate training for models built from the world's most important AI model building block, the transformer. Hopper Tensor Cores can apply mixed FP8 and FP16 precisions to dramatically accelerate AI calculations for transformers.

#### NVLink Switch System

The NVLink Switch System enables the scaling of multi-GPU input/output (IO) across multiple servers. The system delivers up to 9X higher bandwidth than InfiniBand HDR on the NVIDIA Ampere architecture.

#### NVIDIA Confidential Computing

NVIDIA Confidential Computing is a built-in security feature of Hopper that makes NVIDIA H100 the world's first accelerator with confidential computing capabilities. Users can protect the confidentiality and integrity of their data and applications in use while accessing the unsurpassed acceleration of H100 GPUs.

#### • Second-Generation Multi-Instance GPU (MIG)

The Hopper architecture's second-generation MIG supports multi-tenant, multi-user configurations in virtualized environments, securely partitioning the GPU into isolated, right-size instances to maximize quality of service (QoS) for 7X more secured tenants.

#### DPX Instructions

Hopper's DPX instructions accelerate dynamic programming algorithms by 40X compared to CPUs and 7X compared to NVIDIA Ampere architecture GPUs. This leads to dramatically faster times in disease diagnosis, real-time routing optimizations, and graph analytics.

The following figure shows the NVIDIA H100 SXM5 4-GPU Board installed in the ThinkSystem SD665-N V3 server

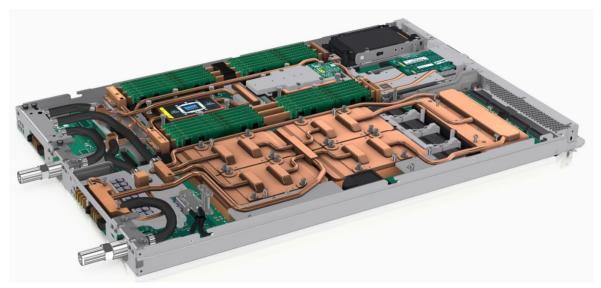


Figure 3. NVIDIA H100 SXM5 4-GPU Board in the ThinkSystem SD665-N V3 server

## **Technical specifications**

The following table lists the GPU processing specifications and performance of the NVIDIA H100 GPU.

Table 2. Specifications of the NVIDIA H100 GPU

Feature	H100 NVL 94GB PCle adapter	H100 80GB SXM board	H100 94GB SXM board
GPU Architecture	NVIDIA Hopper	NVIDIA Hopper	NVIDIA Hopper
Part number	4X67A89325	BQQV or C1HL	BUBB
GPUs per part number	1	BQQV: 4 C1HL: 8	4
NVIDIA Tensor Cores	456 fourth-generation Tensor Cores	528 fourth-generation Tensor Cores	528 fourth-generation Tensor Cores
NVIDIA CUDA Cores (shading units)	14,592 FP32 CUDA Cores	16,896 FP32 CUDA Cores	16,896 FP32 CUDA Cores
Peak FP64 performance	34 TFLOPS	34 TFLOPS	34 TFLOPS
Peak FP64 Tensor Core performance	67 TFLOPS	67 TFLOPS	67 TFLOPS
Peak FP32 performance	67 TFLOPS	67 TFLOPS	67 TFLOPS
Peak Tensor Float 32 (TF32) performance	990 TFLOPS*	989 TFLOPS*	989 TFLOPS*
Peak FP16 performance	1,980 TFOPS*	1,979 TFLOPS*	1,979 TFLOPS*
Peak Bfloat16 (BF16) performance	1,980 TFOPS*	1,979 TFLOPS*	1,979 TFLOPS*
Peak FP8 performance	3,960 TFOPS*		
INT8 Integer Performance	3,960 TOPS*	3,958 TOPS*	3,958 TOPS*
GPU Memory	94 GB HBM3	80GB board (feature BQQV): 80 GB HBM3	94GB board (feature BUBB): 90GB HBM2e
Memory Bandwidth	3.9 TB/s	80GB board (feature BQQV): 3.35 TB/sec	94GB board (feature BUBB): 2.4 TB/sec
ECC	Yes	Yes	Yes
Interconnect Bandwidth	NVLink: 600 GB/sec PCIe Gen5: 128 GB/sec	NVLink: 900 GB/sec PCle Gen5: 128 GB/sec	NVLink: 900 GB/sec PCle Gen5: 128 GB/sec
System Interface	PCle Gen 5.0, x16 lanes	PCle Gen 5.0, x16 lanes	PCIe Gen 5.0, x16 lanes
Form Factor	PCIe full height/length, double width	SXM5	SXM5
NVLink support	Yes; 3 NVLink Bridge supported per pair of GPUs (all 3 required)	Yes, integrated	Yes, integrated
Multi-Instance GPU (MIG)	Up to 7 GPU instances, 12GB each	Up to 7 GPU instances, 10GB each	Up to 7 GPU instances, 10GB each
Max Power Consumption	400 W	700 W	700 W
Thermal Solution	Passive	Water cooled	Water cooled
Compute APIs	CUDA, DirectCompute, OpenCL, OpenACC	CUDA, DirectCompute, OpenCL, OpenACC	CUDA, DirectCompute, OpenCL, OpenACC

<sup>\*</sup> With structural sparsity enabled

## **Server support**

The following tables list the ThinkSystem servers that are compatible.

**NVLink server support**: The NVLink Ampere bridge is supported with additional NVIDIA A-series and H-series GPUs. As a result, there are additional servers listed as supporting the bridge that don't support the H100 GPU.

Table 3. Server support (Part 1 of 4)

			AMD V3		2S Intel V3/V4				4	4S 8S Intel V3			Multi Node V3/V4								
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)	SR650 V4 (7DGC / 7DGD)	SR650a V4 (7DGC / 7DGD)	SR850 V3 (7D97 / 7D96)	SR860 V3 (7D94 / 7D93)	SR950 V3 (7DC5 / 7DC4)	SD535 V3 (7DD8 / 7DD1)	SD530 V3 (7DDA / 7DD3)	SD550 V3 (7DD9 / 7DD2)	ST45 V3 (7DH4 / 7DH5)	3 (7DF4 / 7	ST250 V3 (7DCF / 7DCE)	SR250 V3 (7DCM / 7DCL)
Double-wide	PCle adapter form factor																				
4X67A89325	ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU	Z	3	Ν	3	Ν	Ν	3	Ν	2	4	Z	Z	Ν	Ν	Ν	Ν	Ν	Ζ	Ν	N
SXM form fac	ctor																				
C1HL	ThinkSystem NVIDIA HGX H100 80GB 700W 8-GPU Board	N	Ν	N	N	Ν	N	Ν	N	N	Ν	Ν	Ν	Ν	Ν	N	N	N	N	N	N
BQQV	ThinkSystem NVIDIA H100 SXM5 700W 80G HBM3 GPU Board	N	Ν	N	N	Ν	N	Ν	N	N	Ζ	Ζ	Ζ	Ν	Ν	N	N	N	N	Ν	N
BUBB	ThinkSystem NVIDIA H100 SXM5 700W 94G HBM2e GPU Board	Ν	Ν	N	N	Ν	N	Ν	N	N	Ζ	Ζ	Ζ	Ν	Ν	N	N	N	N	Ν	N
NVLink bridg	e (for PCle adapters only, not SXM	l; o	rde	r 3	per	· pa	ir o	f G	PU	s)											
4X67A71309	ThinkSystem NVIDIA Ampere NVLink 2-Slot Bridge	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

Table 4. Server support (Part 2 of 4)

			GP	U R	ich				Ed	ge			,	Sup	er	Cor	npu	tin	g
Part Number	Description	SR670 V2 (7Z22 / 7Z23)	SR675 V3 (7D9Q / 7D9R)	SR680a V3 (7DHE)	SR685a V3 (7DHC)	SR780a V3 (7DJ5)	SE100 (7DGR)	SE350 (7Z46 / 7D1X)	SE350 V2 (7DA9)	SE360 V2 (7DAM)	SE450 (7D8T)	SE455 V3 (7DBY)	SC750 V4 (7DDJ)	SC777 V4 (7DKA)	SD665 V3 (7D9P)	SD665-N V3 (7DAZ)	SD650 V3 (7D7M)	D650-I V3 (	SD650-N V3 (7D7N)
Double-wide	PCle adapter form factor																		
4X67A89325	ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU	N	8	N	N	N	N	Ν	Ν	Ν	N	Ν	N	N	N	N	N	Ν	N
SXM form fac	ctor					•	•											,	
C1HL	ThinkSystem NVIDIA HGX H100 80GB 700W 8-GPU Board	N	N	1 <sup>1</sup>	1 <sup>1</sup>	N	N	Ν	Ν	Ν	N	Ν	N	N	N	N	N	N	N
BQQV	ThinkSystem NVIDIA H100 SXM5 700W 80G HBM3 GPU Board	N	1 <sup>2</sup>	N	N	N	N	Ζ	Ζ	Ν	Ν	Ν	Ν	Ν	Ν	12	N	Z	12
BUBB	ThinkSystem NVIDIA H100 SXM5 700W 94G HBM2e GPU Board	N	N	N	N	N	N	Z	Z	Ζ	Ν	Z	Ζ	Ν	Ζ	1 <sup>2</sup>	N	Z	1 <sup>2</sup>
NVLink bridg	e (for PCle adapters only, not SXM; o	rde	r 3 p	er p	oair	of (	GPU	Js)											
4X67A71309	ThinkSystem NVIDIA Ampere NVLink 2-Slot Bridge	Υ	Υ	N	N	N	N	Ν	Ν	Ν	N	N	N	N	N	N	N	Ν	N

<sup>1.</sup> Contains 8 separate GPUs connected via high-speed interconnects

<sup>2.</sup> Contains 4 separate GPUs connected via high-speed interconnects

Table 5. Server support (Part 3 of 4)

			S In V2		2S Intel V2				ΑN	ΙD	<b>V</b> 1		Dense V2				_	S '2	88
Part Number	Description	ST50 V2 (7D8K / 7D8J)	ST250 V2 (7D8G / 7D8F)	SR250 V2 (7D7R / 7D7Q)	ST650 V2 (7Z75 / 7Z74)	SR630 V2 (7Z70 / 7Z71)	SR650 V2 (7Z72 / 7Z73)	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)
Double-wide	PCle adapter form factor																		
4X67A89325	ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU	N	N	N	N	N	N	N	Ν	N	Ν	N	N	N	N	N	Ν	N	N
SXM form fac	ctor																		
C1HL	ThinkSystem NVIDIA HGX H100 80GB 700W 8-GPU Board	N	N	N	N	Ν	N	N	Ν	N	Ν	N	N	N	N	N	Ν	N	N
BQQV	ThinkSystem NVIDIA H100 SXM5 700W 80G HBM3 GPU Board	N	N	N	N	Ν	N	Ν	Z	Ν	Z	Ν	Ν	Ν	Ζ	Ν	Ν	N	N
BUBB	ThinkSystem NVIDIA H100 SXM5 700W 94G HBM2e GPU Board	N	N	N	N	Ν	N	N	Z	Ζ	Z	Z	Ζ	Ν	Z	Ζ	Ζ	N	N
NVLink bridg	e (for PCle adapters only, not SXM; ord	er 3	3 pe	er p	air	of (	GPU	Js)											
4X67A71309	ThinkSystem NVIDIA Ampere NVLink 2-Slot Bridge	N	N	N	N	Ν	N	N	Ν	N	Ν	Ν	N	N	Ν	N	Ν	N	N

Table 6. Server support (Part 4 of 4)

		4	s v	1	18	ln	tel '	V1			28	In	tel \	<b>V</b> 1			D	ens	se V	/1
Part Number	Description	SR850 (7X18 / 7X19)	SR850P (7D2F / 2D2G)	SR860 (7X69 / 7X70)	ST50 (7Y48 / 7Y50)	ST250 (7Y45 / 7Y46)	SR150 (7Y54)	SR250 (7Y52 / 7Y51)	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)
Double-wide	PCle adapter form factor																			
4X67A89325	ThinkSystem NVIDIA H100 NVL 94GB PCle Gen5 Passive GPU	N	N	N	Ν	N	N	Ν	Ν	Ν	N	N	Z	N	N	Ν	Ζ	Ν	Ν	N
SXM form fac	ctor		•			•	•								•					
C1HL	ThinkSystem NVIDIA HGX H100 80GB 700W 8-GPU Board	N	N	N	N	N	N	N	N	N	Ν	Ν	Ν	Ν	N	N	Ν	N	N	N
BQQV	ThinkSystem NVIDIA H100 SXM5 700W 80G HBM3 GPU Board	N	N	N	Ν	N	N	N	Ν	N	Ν	Ν	Ζ	Ν	N	N	Ν	Ν	N	N
BUBB	ThinkSystem NVIDIA H100 SXM5 700W 94G HBM2e GPU Board	N	Ν	N	Ν	N	N	N	Ν	N	Ν	Ν	Z	Ν	N	N	Ν	Ν	N	N
NVLink bridg	e (for PCle adapters only, not SXM; o	rde	r 3	peı	· pa	ir c	f G	PU	s)											
4X67A71309	ThinkSystem NVIDIA Ampere NVLink 2-Slot Bridge	N	N	N	N	N	N	N	N	N	Ν	N	Ν	N	N	N	Ν	Ν	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 H100 80GB PCIe Gen5 Passive GPU, 4X67A82257

Microsoft Windows Server 2022	Operating systems	SR650 V3 (4th Gen Xeon)	SR650 V3 (5th Gen Xeon)	SR655 V3	SR665 V3	SR675 V3	SR850 V3	SR860 V3		SR670 V2	SR665
Microsoft Windows Server 2019  Microsoft Windows Server 2022  Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	Microsoft Windows 10	N	Υ	Υ	Υ	Ν	Ν	N	Ν	Ν	N
Microsoft Windows Server 2022	Microsoft Windows 11	N	Υ	Υ		Ν				N	
Microsoft Windows Server 2025	Microsoft Windows Server 2019	Υ	Υ	Υ	Υ	Υ	Y 3	Y 3	Υ	Υ	Y 4
Red Hat Enterprise Linux 7.9  Red Hat Enterprise Linux 8.3  N N N N N N N N N N N N N N N N N N N	Microsoft Windows Server 2022	Υ	Υ	Y 2	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 8.3  N N N N N N N N N N N N N N N N N N N	Microsoft Windows Server 2025	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ
Red Hat Enterprise Linux 8.4  Red Hat Enterprise Linux 8.5  Red Hat Enterprise Linux 8.6  Red Hat Enterprise Linux 8.6  Red Hat Enterprise Linux 8.6  Red Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  SUSE Linux Enterprise Server 15 SP3  SUSE Linux Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP5  Yeld Nover Nover 15 SP4  SUSE Linux Enterprise Server 15 SP5  Yeld Nover Nover 15 SP5  Yeld Nover Nover 15 SP6  Nover	Red Hat Enterprise Linux 7.9	N	N	N	Ν	Ν	Ν	Ν	Υ	Υ	
Red Hat Enterprise Linux 8.5  Rod Hat Enterprise Linux 8.6  Rod Hat Enterprise Linux 8.6  Rod Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  N N N N N N N N N N N N N N N N N N N	Red Hat Enterprise Linux 8.3	N	N	N	Ν	Ν	Ν	Ν	Υ	Υ	Y 4
Red Hat Enterprise Linux 8.6  Red Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP6  Red Hat Enterprise Server 15 SP7  SUSE Linux Enterprise Server 15 SP7  SUSE Linux Enterprise Server 15 SP8  Red Hat Enterprise Server 15 SP8  Red Hat Enterprise Server 15 SP9  Re	Red Hat Enterprise Linux 8.4	N	N	N	Ν	Ν	Ν	N	Υ	Υ	Y 4
Red Hat Enterprise Linux 8.7  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  N N N N N N N N N N N N N N N N N N N	Red Hat Enterprise Linux 8.5	N	N	N	Ν	Ν	Ν	N	Υ	Υ	Y 4
Red Hat Enterprise Linux 8.8  Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  N N N N N N N N N N N N N N N N N N N	Red Hat Enterprise Linux 8.6	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 8.9  Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  SUSE Linux Enterprise Server 15 SP3  N N N N N N N N N N N N N N N N N N N	Red Hat Enterprise Linux 8.7	Υ	N	Υ	Υ	Υ	Y	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 9.0  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP4  Red Hat Enterprise Server 15 SP4  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.2  Red H	Red Hat Enterprise Linux 8.8	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 9.1  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP4  RUSE Linux Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP4  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux	Red Hat Enterprise Linux 8.9	Υ	Υ	Υ	Υ	Ζ	Υ	Υ	Υ	Υ	Υ
Red Hat Enterprise Linux 9.2  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Ha	Red Hat Enterprise Linux 9.0	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 9.3  Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP4  SUSE Linux Enterprise Server 15 SP5  Red Hat Enterprise Server 15 SP3  Red Hat Enterprise Linux 9.4  Red Hat Enterpri	Red Hat Enterprise Linux 9.1	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 9.4  Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  N N N N N N N N N N N N N N N N N N N	Red Hat Enterprise Linux 9.2	Υ	Υ	Υ	Υ	Υ	Y	Υ	Υ	Υ	Y 4
Red Hat Enterprise Linux 9.5  Red Hat Enterprise Server 15 SP3  Red Hat En	Red Hat Enterprise Linux 9.3	Υ	Υ	Υ	Υ	Ν	Υ	Υ	Υ	Υ	Υ
SUSE Linux Enterprise Server 15 SP3  N N N N N N N N N N Y Y Y Y SUSE Linux Enterprise Server 15 SP4  Y N Y² Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	Red Hat Enterprise Linux 9.4	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ
SUSE Linux Enterprise Server 15 SP4  Y N Y <sup>2</sup> Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	Red Hat Enterprise Linux 9.5	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ
SUSE Linux Enterprise Server 15 SP5  Y Y Y N Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	SUSE Linux Enterprise Server 15 SP3	N	N		Ν	Ν	Ν	N	Υ	Υ	Y 4
Ubuntu 18.04.5 LTS       N N N N N N N N N N N N N N N N N N N	SUSE Linux Enterprise Server 15 SP4	Υ	N	Y 2	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
Ubuntu 20.04 LTS       N       Y       Y       Y       Y       Y	SUSE Linux Enterprise Server 15 SP5	Υ	Υ	Υ	Υ	Ν	Υ	Υ	Υ	Υ	Y 4
Ubuntu 20.04.5 LTS       N       N       Y	Ubuntu 18.04.5 LTS	N	N	N	Ν	Ν	Ν	N	Υ	Υ	N
Ubuntu 22.04 LTS       Y	Ubuntu 20.04 LTS	N	N	N	Ν	Ν	Ν	N	Υ	Ν	N
Ubuntu 22.04.5 LTS       N N N N Y N N N N N N N N N N N N N N N	Ubuntu 20.04.5 LTS	N	N	Υ	Υ	Υ	Υ	Υ	N	Ν	N
Ubuntu 24.04 LTS       Y       N       Y	Ubuntu 22.04 LTS	Υ	Υ1	Y 2	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
VMware vSphere Hypervisor (ESXi) 7.0 U3       Y <td>Ubuntu 22.04.5 LTS</td> <td>N</td> <td>N</td> <td>N</td> <td>Υ</td> <td>Ν</td> <td>Ν</td> <td>N</td> <td>Ν</td> <td>Ν</td> <td>N</td>	Ubuntu 22.04.5 LTS	N	N	N	Υ	Ν	Ν	N	Ν	Ν	N
VMware vSphere Hypervisor (ESXi) 8.0       Y       N       Y <sup>2</sup> Y       N       N       Y <td>Ubuntu 24.04 LTS</td> <td>Υ</td> <td>N</td> <td>Υ</td> <td>Υ</td> <td>Υ</td> <td>Υ</td> <td>Υ</td> <td>Υ</td> <td>Υ</td> <td>Υ</td>	Ubuntu 24.04 LTS	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ
VMware vSphere Hypervisor (ESXi) 8.0 U1 Y N Y Y Y Y Y Y Y Y Y Y Y Y	VMware vSphere Hypervisor (ESXi) 7.0 U3	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
	VMware vSphere Hypervisor (ESXi) 8.0	Y	N	Y 2	Υ	Ν	Ν	N	Υ	Υ	Y 4
VMware vSphere Hypervisor (ESXi) 8.0 U2 YYYYYYYYY	VMware vSphere Hypervisor (ESXi) 8.0 U1	Υ	N	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4
	VMware vSphere Hypervisor (ESXi) 8.0 U2	Y	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Y 4

<sup>&</sup>lt;sup>1</sup> Ubuntu 22.04.3 LTS/Ubuntu 22.04.4 LTS

-

#### **NVIDIA GPU software**

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

- NVIDIA AI Enterprise Software
- NVIDIA HPC Compiler Software

The PCIe adapter H100 GPUs include a five-year software subscription, including enterprise support, to the NVIDIA AI Enterprise software suite:

- ThinkSystem NVIDIA H100 NVL 94GB PCIe Gen5 Passive GPU, 4X67A89325
- ThinkSystem NVIDIA H100 80GB PCIe Gen5 Passive GPU, 4X67A82257

This license is equivalent to part number 7S02001HWW listed in the NVIDIA AI Enterprise Software section below.

To activate the NVIDIA AI Enterprise license, see the following page:

https://www.nvidia.com/en-us/data-center/activate-license/

**SXM GPUs**: The NVIDIA AI Enterprise software suite is not included with the SXM H100 GPUs and will need to ordered separately if needed.

#### **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 oneyear, 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 8. 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

<sup>&</sup>lt;sup>2</sup> For limitation, please refer Support Tip TT1064

<sup>&</sup>lt;sup>3</sup> For limitation, please refer Support Tip TT1591

<sup>&</sup>lt;sup>4</sup> HW is not supported with EPYC 7002 processors.

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

Part number	Feature code 7S02CTO1WW	NVIDIA part number	Description
Al Enterprise Pe	rpetual License		
7S02001BWW	S6YY	731-AI7004+P3CMI60	NVIDIA AI Enterprise Perpetual License and Support per GPU Socket, 5 Years
7S02001EWW	S6Z1	731-AI7004+P3EDI60	NVIDIA AI Enterprise Perpetual License and Support per GPU Socket, EDU, 5 Years
Al Enterprise Su	bscription License		
7S02001FWW	S6Z2	731-AI7003+P3CMI12	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, 1 Year
7S02001GWW	S6Z3	731-AI7003+P3CMI36	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, 3 Years
7S02001HWW	S6Z4	731-AI7003+P3CMI60	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, 5 Years
7S02001JWW	S6Z5	731-AI7003+P3EDI12	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, EDU, 1 Year
7S02001KWW	S6Z6	731-AI7003+P3EDI36	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, EDU, 3 Years
7S02001LWW	S6Z7	731-AI7003+P3EDI60	NVIDIA AI Enterprise Subscription License and Support per GPU Socket, EDU, 5 Years

Find more information in the NVIDIA AI Enterprise Sizing Guide.

## **NVIDIA HPC Compiler Software**

Table 10. NVIDIA HPC Compiler

Part number	Feature code 7S09CTO6WW	Description
HPC Compiler S	upport 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 P	remier Support Se	ervices
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 power cables needed for the H100 SXM GPUs are included with the supported servers.

The H100 PCle GPU 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 11. Auxiliary power cables for H100

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

#### 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\*

20±5(2X)

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 numbers are for thermal kits and include other components needed to install the GPU. See the SR650 V3 product guide or SR655 V3 product guide or SR665 V3 product guide for details.

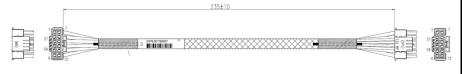
#### Auxiliary power cable needed with the SR675 V3

## 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 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



## 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



## Regulatory approvals

The NVIDIA H100 GPU has the following regulatory approvals:

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

## **Operating environment**

The NVIDIA H100 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.

## Seller training courses

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

<sup>\*</sup> A period not more than 96 hours consecutive, not to exceed 15 days per year.

## 1. Lenovo Cloud Architecture VTT: Supercharge Your Enterprise AI with NVIDIA AI Enterprise on Lenovo Hybrid AI Platform

2025-04-17 | 75 minutes | Employees Only

Join us for an in-depth webinar with Justin King, Principal Product Marketing Manager for Enterprise Al exploring the power of NVIDIA AI Enterprise, delivering Generative and Agentic AI outcomes deployed with Lenovo Hybrid AI platform environments.

In today's data-driven landscape, AI is evolving at high speed, with new techniques delivering more accurate responses. Enterprises are seeking not just an understanding but also how they can achieve AI-driven business outcomes.

With this, the demand for secure, scalable, and high-performing AI operations-and the skills to deliver them-is top of mind for many. Learn how NVIDIA AI Enterprise, a comprehensive software suite optimized for NVIDIA GPUs, provides the tools and frameworks, including NVIDIA NIM, NeMo, and Blueprints, to accelerate AI development and deployment while reducing risk-all within the control and security of your Lenovo customer's hybrid AI environment.

Tags: Artificial Intelligence (AI), Cloud, Data Management, Nvidia, Technical Sales

Published: 2025-04-17 Length: 75 minutes

#### Start the training:

Employee link: Grow@Lenovo

Course code: DVCLD221

#### 2. Al VTT: GTC Update and The Lenovo LLM Sizing Guide

2025-03-12 | 86 minutes | Employees Only

Please view this session that is two parts. Part one is Robert Daigle, Director, Global Al Solutions and Hande Sahin-Bahceci, Al Solutions Marketing Leader explaining the upcoming announcements for NVIDIA GTC. Part Two is Sachin Wani, Al Data Scientist explaining the Lenovo LLM Sizing Guide with these topics:

- Minimum GPU requirements for fine-tuning/training and inference
- Gathering requirements for the customer's use case
- · LLMs from a technical perspective

Tags: Artificial Intelligence (AI), Technical Sales

Published: 2025-03-12 Length: 86 minutes

#### Start the training:

Employee link: Grow@Lenovo

Course code: DVAI214

#### 3. Partner Technical Webinar - NVIDIA Portfolio

2024-11-06 | 60 minutes | Employees and Partners

In this 60-minute replay, Jason Knudsen of NVIDIA presented the NVIDIA Computing Platform. Jason talked about the full portfolio from GPUs to Networking to AI Enterprise and NIMs.

Tags: Artificial Intelligence (AI), Nvidia

Published: 2024-11-06 Length: 60 minutes

#### Start the training:

Employee link: Grow@Lenovo
Partner link: Lenovo Partner Learning

Course code: 110124

#### 4. NVIDIA Data Center GPU Portfolio

2024-09-26 | 11 minutes | Employees and Partners

This course equips Lenovo and partner technical sellers with the knowledge to effectively communicate the positioning of NVIDIA's data center GPU portfolio, enhancing your ability to showcase its key advantages to clients.

Upon completion of this training, you will be familiar with the following:

- Data Center GPUs for AI and HPC
- Data Center GPUs for Graphics
- •GPU comparisons

Tags: Artificial Intelligence (AI), High-Performance Computing (HPC), Nvidia

Published: 2024-09-26 Length: 11 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DAINVD201

## 5. **Q2 Solutions Launch TruScale GPU Next Generation Management in the AI Era Quick Hit** 2024-09-10 | 6 minutes | Employees and Partners

This Quick Hit focuses on Lenovo announcing additional ways to help you build, scale, and evolve your customer's private AI faster for improved ROI with TruScale GPU as a Service, AI-driven systems management, and infrastructure transformation services.

Tags: Artificial Intelligence (AI), Lowdown2025ep2, Services, TruScale

Published: 2024-09-10 Length: 6 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: SXXW2543a

#### 6. VTT AI: The NetApp AlPod with Lenovo for NVIDIA OVX

2024-08-13 | 38 minutes | Employees and Partners

AI, for some organizations, is out of reach, due to cost, integration complexity, and time to deployment. Previously, organizations relied on frequently retraining their LLMs with the latest data, a costly and time-consuming process. The NetApp AIPod with Lenovo for NVIDIA OVX combines NVIDIA-Certified OVX Lenovo ThinkSystem SR675 V3 servers with validated NetApp storage to create a converged infrastructure specifically designed for AI workloads. Using this solution, customers will be able to conduct AI RAG and inferencing operations for use cases like chatbots, knowledge management, and object recognition.

Topics covered in this VTT session include:

- •Where Lenovo fits in the solution
- •NetApp AlPod with Lenovo for NVIDIA OVX Solution Overview
- •Challenges/pain points that this solution solves for enterprises deploying Al
- •Solution value/benefits of the combined NetApp, Lenovo, and NVIDIA OVX-Certified Solution

Tags: Artificial Intelligence (AI), Nvidia, Sales, Technical Sales, ThinkSystem

Published: 2024-08-13 Length: 38 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DVAI206

#### 7. Introduction to Artificial Intelligence

2024-08-02 | 11 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding!

This NVIDIA course aims to answer questions such as:

- · What is AI?
- · Why are enterprises so interested in it?
- How does Al happen?
- Why are GPUs so important for it?
- · What does a good AI solution look like?

#### Course Objectives:

By the end of this training, you should be able to:

- 1. Describe AI on a high level and list a few common enterprise use cases
- 2. List how enterprises benefit from AI
- 3. Distinguish between Training and Inference
- 4. Say how GPUs address known bottlenecks in a typical AI pipeline
- 5. Tell a customer why NVIDIA's AI solutions are well-respected in the market

Tags: Nvidia

Published: 2024-08-02 Length: 11 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DAINVD104r2

#### 8. **GPU Fundamentals**

2024-08-02 | 10 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding.

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. We will also talk about the value GPUs bring to modern-day enterprise computing.

## Course Objectives:

By the end of this training, you should be able to:

- 1.Distinguish between serial and parallel processing
- 2. Explain what a GPU is and what it does at a high level
- 3. Articulate the value of GPU computing for enterprises
- 4.List three typical GPU-accelerated workloads and a few uses cases
- 5. Recommend the appropriate NVIDIA GPU for its corresponding enterprise computing workloads

Tags: Nvidia

Published: 2024-08-02 Length: 10 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DAINVD103r2

#### 9. Key NVIDIA Use Cases for Industry Verticals

2024-08-02 | 32 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding.

In this NVIDIA course, you will learn about key AI use cases driving innovation and change across Automotive, Financial Services, Energy, Healthcare, Higher Education, Manufacturing, Retail and Telco industries.

#### Course Objectives:

By the end of this training, you should be able to:

- 1. Discuss common Al use cases across a broad range of industry verticals
- 2. Explain how NVIDIA's AI software stack speeds up time to production for AI projects in multiple industry verticals

Tags: Nvidia

Published: 2024-08-02 Length: 32 minutes

#### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DAINVD108

#### 10. Generative Al Overview

2024-08-02 | 17 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding!

Since ChatGPTs debut in November of 2022, it has become clear that Generative AI has the potential to revolutionize many aspects of our personal and professional lives. This NVIDIA course aims to answer questions such as:

- What are the Generative AI market trends?
- · What is generative AI and how does it work?

#### Course Objectives:

By the end of this training, you should be able to:

- 1. Discuss the Generative AI market trends and the challenges in this space with your customers.
- 2. Explain what Generative AI is and how the technology works to help enterprises to unlock new opportunities for the business.
- 3. Present a high-level overview of the steps involved in building a Generative AI application.

Tags: Nvidia

Published: 2024-08-02 Length: 17 minutes

#### Start the training:

Employee link: Grow@Lenovo
Partner link: Lenovo Partner Learning

Course code: DAINVD106r2

## 11. Retrieval Augmented Generation

2024-08-02 | 15 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding!

In this NVIDIA course, Dave Barry, Senior Solutions Architect, talks about a technique known as Retrieval Augmented Generation (RAG). It is a powerful tool for enhancing the accuracy and reliability of Generative AI models with facts fetched from external sources.

This course requires prior knowledge of Generative Al concepts, such as the difference between model training and inference. Please refer to relevant courses within this curriculum.

#### Course Objectives:

By the end of this training, you should be able to:

- 1. Explain the limitations of large language models to customers
- 2. Articulate the value of RAG to enterprises
- 3. Demo an NVIDIA RAG workflow with a video
- 4. Drive TCO conversations using an authentic use case

Tags: Nvidia

Published: 2024-08-02 Length: 15 minutes

#### Start the training:

Employee link: Grow@Lenovo
Partner link: Lenovo Partner Learning

Course code: DAINVD107

#### 12. Al Industry Use Cases & Solutions

2024-08-02 | 25 minutes | Employees and Partners

IMPORTANT: If you receive the following error message:

"There is an issue with this slide content. Please contact your administrator", please change your VPN location setting and try again. We are actively working on fixing this issue. Thank you for your understanding!

This NVIDIA course aims to answer the question:

· How does NVIDIA bring AI solutions to market with and through the partner ecosystem?

Course Objectives:

By the end of this training, you should be able to:

- 1. Think of solutions in terms of an industry and use case approach
- 2. Develop solutions that address the industry-specific challenges (with FSI as the illustrative model)
- 3. Engage customers with their conversations and advance deals with stakeholder's concerns in mind
- 4. Replicate NVIDIA's best practices and ecosystem engagement strategies appropriately

Tags: Nvidia

Published: 2024-08-02 Length: 25 minutes

### Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DAINVD105r2

#### 13. Partner Technical Webinar - NVIDIA Smart Spaces

2024-07-24 | 60 minutes | Employees and Partners

In this 60-minute replay, Alex Pazos, NVIDIA BDM for Smart Spaces, reviewed the NVIDIA AI for Smart Spaces framework and use cases. Alex reviewed the Metropolus Framework and the Smart Spaces ecosystem. Then he reviewed several use cases including sports stadiums, warehouses, airports, and roadways.

Tags: Artificial Intelligence (AI)

Published: 2024-07-24 Length: 60 minutes

## Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: 071924

#### 14. Guidance for Selling NVIDIA Products at Lenovo for ISG

2024-07-01 | 25 minutes | Employees and Partners

This course gives key talking points about the Lenovo and NVIDIA partnership in the Data Center. Details are included on where to find the products that are included in the partnership and what to do if NVIDIA products are needed that are not included in the partnership. Contact information is included if help is needed in choosing which product is best for your customer. At the end of this session sellers should be able to explain the Lenovo and NVIDIA partnership, describe the products Lenovo can sell through the partnership with NVIDIA, help a customer purchase other NVIDIA product, and get assistance with choosing NVIDIA products to fit customer needs.

Tags: Artificial Intelligence (AI), Nvidia

Published: 2024-07-01 Length: 25 minutes

## Start the training:

Employee link: Grow@Lenovo

Partner link: Lenovo Partner Learning

Course code: DNVIS102

## 15. Think Al Weekly: Lenovo Al PCs & Al Workstations

2024-05-23 | 60 minutes | Employees Only

Join Mike Leach, Sr. Manager, Workstations Solutions and Pooja Sathe, Director Commercial AI PCs as they discuss why Lenovo AI Developer Workstations and AI PCs are the most powerful, where they fit into the device to cloud ecosystem, and this week's Microsoft announcement, Copilot+PC

Tags: Artificial Intelligence (AI), ThinkStation

Published: 2024-05-23 Length: 60 minutes

#### Start the training:

Employee link: Grow@Lenovo

Course code: DTAIW105

## 16. VTT Cloud Architecture: NVIDIA Using Cloud for GPUs and AI

2024-05-22 | 60 minutes | Employees Only

Join JD Dupont, NVIDIA Head of Americas Sales, Lenovo partnership and Veer Mehta, NVIDIA Solution Architect on an interactive discussion about cloud to edge, designing cloud Solutions with NVIDIA GPUs and minimizing private\hybrid cloud OPEX with GPUs. Discover how you can use what is done at big public cloud providers for your customers. We will also walk through use cases and see a demo you can use to help your customers.

Tags: Artificial Intelligence (AI), Cloud, Nvidia, Software Defined Infrastruture (SDI), Technical Sales

Published: 2024-05-22 Length: 60 minutes

#### Start the training:

Employee link: Grow@Lenovo

Course code: DVCLD212

## 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 H100 product page: https://www.nvidia.com/en-us/data-center/h100/
- NVIDIA Hopper Architecture page https://www.nvidia.com/en-us/data-center/technologies/hopper-architecture/
- ThinkSystem SD665-N V3 product guide https://lenovopress.lenovo.com/lp1613-thinksystem-sd665-n-v3-server
- ThinkSystem SR680a V3 product guide https://lenovopress.lenovo.com/lp1909-thinksystem-sr680a-v3-server
- ThinkSystem SR685a V3 product guide https://lenovopress.lenovo.com/lp1910-thinksystem-sr685a-v3-server

## 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 2025. All rights reserved.

This document, LP1732, was created or updated on December 16, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at: https://lenovopress.lenovo.com/LP1732
- Send your comments in an e-mail to: comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1732.

#### **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 <a href="https://www.lenovo.com/us/en/legal/copytrade/">https://www.lenovo.com/us/en/legal/copytrade/</a>.

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

Lenovo®

Neptune®

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.