NVIDIA DATA CENTER GRAPHICS

Linecard

Accelerate your most demanding HPC and hyperscale data center workloads with NVIDIA® Data Center GPUs.

NVIDIA's accelerators also deliver the horsepower needed to run bigger simulations faster than ever before. Plus, NVIDIA GPUs deliver the highest performance and user density for virtual desktops, applications, and workstations.

TRAINING

Training increasingly complex models faster is key to improving productivity for data scientists and delivering AI services more quickly. Servers powered by NVIDIA® GPUs use the performance of accelerated computing to cut deep learning training time from months to hours or minutes.

INFERENCE

Inference is where a trained neural network really goes to work. As new data points come in such as images, speech, visual and video search, inference is what gives the answers and recommendations at the heart of many AI services. A server with a single GPU can deliver 27X higher inference throughput than a single-socket CPU-only server resulting in dramatic cost savings.

AN ENTERPRISE-READY PLATFORM FOR PRODUCTION AI

NVIDIA AI Enterprise, an end-to-end, secure, cloud-native suite of AI software, accelerates the data science pipeline and streamlines the development and deployment of production AI.

PART #	GPU SPECIFICATIONS			PERFORMANCE			DISPLAY TECHNOLOGY						OPTIONS					
	Part Number	CUDA® Cores	GPU Memory	Memory Bandwidth	Tensor Cores	Ray Tracing Cores	Fvvloating Point Performance, Single Precision (TFLOPS)	DisplayPort 2/3	Maximum Active Displays	Maximum Resolution Support @ 60 Hz	Dimensions	Mosaic Technology	Maximum Power Consumption	H.264 and HEVC Encode/Decode Engines	Vulkan Support	GPUDirect® for Video	Graphics Syncronization	NVLink Bridge Support	NVIDIA AI Enterprise Support
NVIDIA FOR DATA CENTERS																			
NVIDIA H200 NVL	NVH200NVLTCGPU-KIT	16896	94 GB	3938 GB/s	528	-	62 DP 31 SP	-	-	-	4.4 H" x 10.5" FH | DS	-	600W	-	-	-	-	√	√
NVIDIA H100 NVL	NVH100NVLTCGPU-KIT	16896	94 GB	3938 GB/s	528	-	62 DP 31 SP	-	-	-	4.4 H" x 10.5" FH | DS	-	400W	-	-	-	-	√	√
NVIDIA H100	NVH100TCGPU-KIT	14592	80 GB	2000 GB/s	456	-	51 DP 26 SP	-	-	-	4.4 H" x 10.5" FH | DS	-	350W	-	-	-	-	√	√
NVIDIA L40S	NVL40STCGPU-KIT	18176	48 GB	864 GB/s	568	142	91.6 SP	4	4	7680x4320	4.4 H" x 10.5" FH | DS	√	350W	√	√	√	√	-	-
NVIDIA L40	NVL40TCGPU-KIT	18176	48 GB	864 GB/s	568	142	90.5 SP	4	4	7680x4320	4.4 H" x 10.5" FH | DS	√	300W	√	√	√	√	-	-
NVIDIA L4	NVL4TCGPU-KIT	7680	24 GB	300 GB/s	240	60	30.3 SP	4	4	-	HH | LP	-	72W	-	-	-	-	-	-

For more information, contact a PNY Account Manager or email: GOPNY@PNY.COM or visit WWW.PNY.COM/NVIDIA-DATA-CENTER-SOLUTIONS

PNY Technologies, Inc. 100 Jefferson Road, Parsippany, NJ 07054 | Tel 973-515-9700 | Fax 973-560-5590 | WWW.PNY.COM

February 2025 09:34 AM