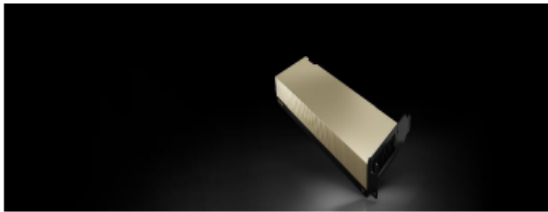# GPU SPECIFICATION

NVIDIA L4 Tensor Core GPU

The breakthrough universal accelerator for efficient video, AI, and graphics.

Accelerate Video, AI, and Graphics Workloads

The NVIDIA L4 Tensor Core GPU powered by the NVIDIA Ada Lovelace architecture delivers universal, energy-efficient acceleration for video, AI, visual computing, graphics, virtualization, and more. Packaged in a low-profile form factor, L4 is a cost-effective, energy-efficient solution for high throughput and low latency in every server, from the edge to the data center to the cloud.

Product Specifications

Form Factor	L4
FP32	30.3 teraFLOPs
TF32 Tensor Core	120 teraFLOPS*
FP16 Tensor Core	242 teraFLOPS*
BFLOAT16 Tensor Core	242 teraFLOPS*
FP8 Tensor Core	485 teraFLOPs*
INT8 Tensor Core	485 TOPs*
GPU memory	24GB

Form Factor	L4
GPU memory bandwidth	300GB/s
NVENC	NVDEC	JPEG decoders	2	4	4
Max thermal design power (TDP)	72W
Form factor	1-slot low-profile, PCIe
Interconnect	PCIe Gen4 x16 64GB/s
Server options	Partner and NVIDIA-Certified Systems with 1–8 GPUs