

DGX B200 Hardware Specification

DGX B200 Physical Specifications



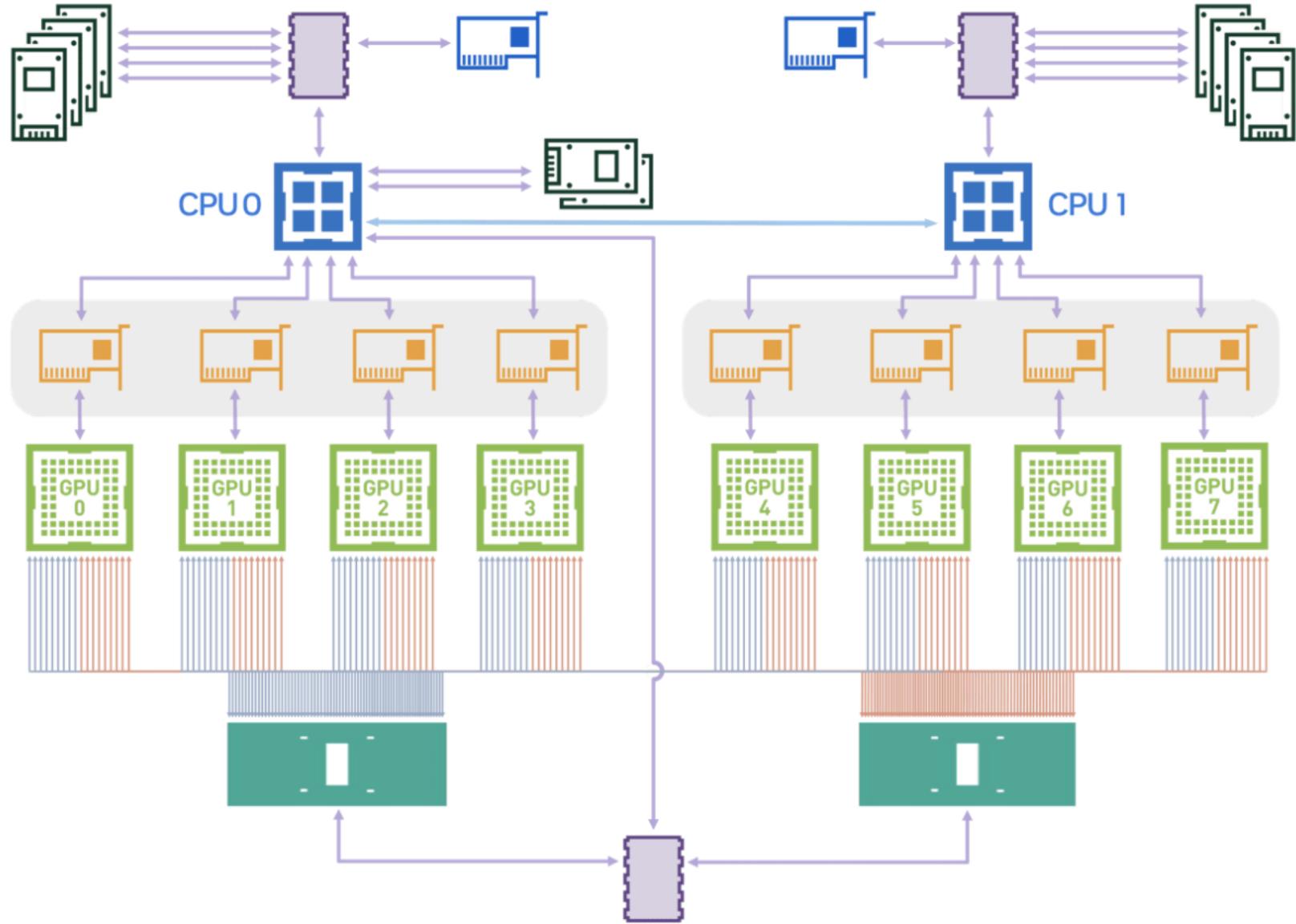
System dimensions (HxWxL)	444 x 482.2 x 897.1 mm 17.5 x 19 x 35.3 in
Rack Units	10 RU
Power Inlets	6
Voltage	200-240 VAC
Max current per outlet	16 A
Power Consumption	14.3 kW Max
Power supplies	6 x 3.3 kW
Heat output	48,794 Max

DGX B200 Hardware Feature Summary



GPU	8 x NVIDIA B200 Tensor Core GPUs
GPU memory	8 x 180 GB = 1440 GB total
Performance	72 petaFLOPS FP8 training 144 petaFLOPS FP4 inference
NVswitches	2 x 5th generation NVlink
CPUs	2 x Intel® Xeon® Platinum 8570 Processors
System Memory	2 TB default (up to 4TB)
Networking connectivity and speed	4 x OSFP ports provide connectivity to 8 x NVIDIA ConnectX-7s (Cluster network) 8 x 400Gb/s InfiniBand/Ethernet 2 x dual-port NVIDIA BlueField 3 DPU (Storage or system management network) 2 x 400Gb/s InfiniBand/Ethernet per card
Cache Storage	8 x U.2 3.84TB NVMe Self-Encrypting Drives
Boot Storage	2 x 1.92TB M.2 NVMe (software encryptable)
Host Management	On-board 10Gb/s RJ-45 Ethernet
Remote System Management	Baseboard Management Controller (BMC) 1Gb/s RJ-45 network connectivity Remote Keyboard, Video, Mouse (KVM) Remote Storage RedFish, IPMI management
Operating System	DGX OS 7 based on Ubuntu 24.04 LTS Ubuntu & Red Hat Enterprise Linux 8 & 9 Rocky Linux

DGX B200 Block Diagram



BlueField-3

ConnectX-7 Network Module

NVMe

PCIe Switches

NVSwitch

PCIe

CPU communication

NVIDIA