NVIDIA HGX B300

Enterprise AI compute. Connected by RoCE.

PEGATRON / NETWEB PLATFORM OPTIONS | BILL OF MATERIALS | 16 SEPTEMBER 2026

Build your AI infrastructure around eight Blackwell Ultra GPUs, high-capacity memory and fast local NVMe storage. Designed for large-model training, fine-tuning and inference.

8 x B300	2 TB	30.72 TB	RoCEv2
BLACKWELL ULTRA GPUs	SYSTEM MEMORY	RAW NVMe CAPACITY	ETHERNET GPU FABRIC

Compute, memory and storage

#	Component	Configuration / specification	Qty / server
1	**Server platform**	Pegatron AS801-2T1-8H2 reference / Netweb Tyrone HGX B300 platform option. 8U rackmount, dual Intel sockets, air cooling and 12 NVMe bays. Final OEM model subject to quotation.	**1** system
2	**GPU assembly**	Eight NVIDIA Blackwell Ultra SXM GPUs with HGX baseboard, fifth-generation NVLink and NVLink Switch; 2.1 TB total GPU memory.	**1** 8-GPU set
3	**Processors**	Intel Xeon 6760P, 64 cores each, 2.20 GHz base / up to 3.80 GHz turbo, 320 MB cache, 330 W TDP. 128 CPU cores per server.	**2** CPU
4	**System memory**	32 x 64 GB ECC DDR5 RDIMM = 2,048 GB. 6400 MT/s-rated modules; 5200 MT/s operation at 2DPC on the Pegatron reference.	**32** DIMM
5	**Data storage**	4 x 7.68 TB enterprise U.2 NVMe SSDs with power-loss protection and hot-swap carriers. 30.72 TB raw capacity before RAID / formatting.	**4** SSD
6	**Boot storage**	2 x 960 GB enterprise PCIe 4.0 NVMe M.2 SSDs for an OEM-supported mirrored boot solution.	**2** SSD
7	**Boot integration kit**	OEM-approved carrier / controller and cabling for two boot SSDs. Required for the Pegatron reference, which lists one onboard M.2 slot.	**1** kit
8	**Power supplies**	12 x 3000 W hot-swap PSUs, 6+6 redundancy. Titanium CRPS on the Pegatron reference; final OEM power design to be confirmed.	**12** PSU
9	**Power cords**	C19-to-C20, 1.5 m, 16 A, 14 AWG or OEM-approved equivalent matched to server inlets and rack PDUs.	**12** cord
10	**Cooling and integration**	OEM fans, heatsinks, internal cables, integration, factory burn-in and protective packaging.	**1** set
11	**Rack installation kit**	OEM-approved rack rails, mounting hardware and cable management for the populated 8U chassis.	**1** kit

Configuration basis. One complete server. Netweb model and all OEM-specific configuration details remain subject to the final supplier quotation.

Connected for AI

ConnectX-8 networking with Ethernet RoCEv2.

16 x 400GbE GPU-FABRIC LINKS PER SERVER | 6.4 Tb/s AGGREGATE ONE-WAY LINE RATE

NVLink connects the GPUs within the server. RoCEv2 provides RDMA over Ethernet for the external GPU fabric, with separate connections for storage, services and management.

#	Component	Configuration / specification	Qty / server
12	**Onboard LAN**	Two onboard 1GbE RJ45 ports; Intel i350-AM2 or OEM equivalent.	**2** port
13	**Dedicated management**	Dedicated RJ45 management port with IPMI / Redfish. ASPEED AST2600 on the Pegatron reference.	**1** port
14	**GPU fabric adapters**	8 x NVIDIA ConnectX-8 C8180, PN 900-9X81E-00EX-ST0 or qualified integrated equivalent. Each configured as 2 x 400GbE RoCEv2 with GPUDirect RDMA.	**8** adapter
15	**GPU fabric optics**	Eight server-side OSFP optical assemblies qualified for dual-400GbE operation, providing 16 links. Switch optics and fibers are listed on page 3.	**8** assembly
16	**Auxiliary DPU**	One NVIDIA BlueField-3 DPU with dual 200GbE connectivity for storage / service networking. Exact SKU and media to be confirmed.	**1** DPU
17	**Additional LAN adapters**	Two dual-port 10/25GbE Ethernet adapters, providing four additional LAN ports.	**2** adapter
18	**25GbE server optics**	Four 25GbE server-side transceivers matched to the selected adapters and link media.	**4** module
19	**Trusted platform module**	TPM 2.0, integrated or OEM-supported module.	**1** unit
20	**NVMe RAID licence**	Intel VROC Premium for supported NVMe RAID 0/1/5/10. Motherboard, VMD routing, SSD and boot-solution compatibility to be confirmed.	**1** licence
21	**Software and validation**	OEM-qualified Linux, NVIDIA GPU drivers, CUDA, NCCL and NIC software integration, with GPU, storage, RDMA and link validation.	**1** service
22	**Warranty**	Three-year onsite parts and labour support to be specified in the quotation, including response SLA and support location.	**1** support lot

Platform integration. OEM to confirm GPU-to-NIC paths, SocketDirect / integrated topology and full dual-400GbE operation. A single PCIe 5.0 x16 host link cannot sustain 800 Gb/s. Aggregate link rate is not guaranteed application throughput.

Commercial scope. Final pricing, delivery, onsite support SLA and paid software subscriptions are to be stated in the quotation. External fabric and site power / cooling must be sized for the deployment.

Plan your AI deployment with Cyfuture

Discuss configuration, deployment and support: **cyfuture.ai/contact-us** | +91 8800766220

Complete the fabric

RoCEv2 network scope for deployment planning.

EXTERNAL NETWORK COMPONENTS | QUANTITIES PER SERVER | QUOTED SEPARATELY

The optical baseline below completes the server connections. Final switch quantities, uplinks, redundancy and link lengths follow the cluster size and site design.

#	Component	Configuration / specification	Qty / server
1	**GPU fabric access ports**	16 x 400GbE switch access ports per server. Native 400GbE ports or validated 800GbE-to-2x400GbE breakout. Uplinks and spare ports are additional.	**16** port
2	**GPU switch-end optics**	400GbE optical modules matched to the eight server-side dual-400GbE OSFP assemblies. Baseline assumes one 400GbE switch module per logical link.	**16** module
3	**GPU optical links**	Matched fiber links for 16 x 400GbE connections. OEM to specify fiber type, connector, polarity and reach after confirming the exact optical modules.	**16** link
4	**GPU fabric switches**	RoCEv2 Ethernet fabric supporting PFC, ECN and suitable congestion control / QoS. Switch count, uplinks and oversubscription depend on node count and the final topology.	**TBD** switch
5	**DPU connectivity**	Two 200GbE switch ports plus compatible optics at both ends and two links, or complete qualified cables. Exact port speed and media to match the chosen BlueField-3 SKU.	**1** 2-link set
6	**25GbE switch optics**	Switch-end 25GbE optical modules for the four additional LAN ports. Matching server-end modules are included in the server BOM.	**4** module
7	**25GbE optical links**	Fiber links for the four additional LAN ports. Final length and connectors to match the selected optical modules.	**4** link
8	**LAN / BMC cabling**	RJ45 copper patch cables for two onboard LAN ports and one dedicated management port; three corresponding switch ports required.	**3** cable
9	**Service / management switches**	Ethernet switches with sufficient 200GbE, 25GbE and 1GbE ports, as applicable. Quantity and redundancy to follow the site network design.	**TBD** switch

Optical compatibility. Match optics, fiber, FEC, lane mapping and reach end to end. A qualified complete DAC / AOC breakout assembly may replace the corresponding optics and fibers; quote only the selected media option.

Commissioning. Configure Ethernet mode, RoCEv2 priorities, PFC / ECN, QoS, congestion control and consistent MTU. Validate multi-node RDMA and NCCL on the intended fabric using at least two servers.

Technical references

Pegatron reference platform: AS801-2T1-8H2 specifications
NVIDIA Ethernet port mapping: ConnectX-8 port configurations
NVIDIA adapter part number: ConnectX-8 ordering information
NVIDIA GPU platform: HGX B300 specifications
Intel processor: Xeon 6760P specifications
Brand and sales contact: Cyfuture media center | Cyfuture AI contact