



04 · AI ACCELERATION

X8 RTX PRO 6000 Blackwell

Enterprise inference.
Visual AI. Dense PCIe
acceleration.

A 4U dual-socket AMD EPYC platform with
eight NVIDIA RTX PRO 6000 Blackwell
Server Edition GPUs, 2 TiB of host memory
and 400GbE-class fabric connectivity.



PRODUCT VIEW · X8 RTX PRO 6000 BLACKWELL

4U

AIR-COOLED PLATFORM

8 ×

RTX PRO 6000 BLACKWELL

96 GB

GDDR7 ECC / GPU

2 TiB

HOST MEMORY

DESIGNED FOR

enterprise inference · governed RAG · multimodal and visual AI · digital twins · private model hosting

01

Consolidate eight Blackwell GPUs

Eight 96 GB GDDR7 ECC accelerators provide
768 GB of nominal device memory across the
node for inference, visual AI and digital-twin
workloads.

02

Balance GPU and host resources

Two EPYC 9455 processors and 2 TiB of
DDR5 host memory support model serving,
data preparation and orchestration close to
the accelerators.

03

Connect at 400GbE class

Eight 400GbE-capable east-west ports
support Standard and Performance fabric
profiles; BlueField-3 provides separate
service connectivity.

Product imagery is illustrative. Final chassis details and installed options are defined by the order configuration.

ENTERPRISE OPERATIONS

Operate through separated service paths.

Server BMC and BlueField out-of-band access separate platform administration
from workload traffic, while mirrored boot and redundant power support planned
service continuity.

SERVER BMC
DPU OOB
3+1 POWER

Eight Blackwell GPUs. A balanced enterprise AI node.

X8 RTX PRO 6000 Blackwell combines dense PCIe acceleration, dual-socket host compute, local NVMe and high-speed networking for enterprise inference and visual AI.

CUSTOMER BENEFITS

- 01 Consolidate eight Blackwell GPUs**
Eight 96 GB GDDR7 ECC accelerators provide 768 GB of nominal device memory across the node for inference, visual AI and digital-twin workloads.
- 02 Balance GPU and host resources**
Two EPYC 9455 processors and 2 TiB of DDR5 host memory support model serving, data preparation and orchestration close to the accelerators.
- 03 Connect at 400GbE class**
Eight 400GbE-capable east-west ports support Standard and Performance fabric profiles; BlueField-3 provides separate service connectivity.
- 04 Design for continuous service**
Mirrored boot devices, local NVMe, a 3+1 power arrangement and front-to-rear cooling support data-center operation and maintenance.

IDEAL WORKLOADS

- 01 Enterprise inference**
- 02 Governed RAG**
- 03 Multimodal and visual AI**
- 04 Digital twins**
- 05 Private model hosting**

CONFIGURATION FLEXIBILITY

- Standard or Performance fabric
- 2 TiB or 4 TiB host-memory profile
- EPYC 9455 or order-defined 9555 option
- Local NVMe population defined by order

TECHNICAL HIGHLIGHTS

FORM FACTOR

4U air-cooled rack server; 438 × 176 × 800 mm (W × H × D)

ACCELERATORS

8 × NVIDIA RTX PRO 6000 Blackwell Server Edition; 96 GB nominal GDDR7 ECC each; configurable up to 600 W/GPU

PROCESSORS

2 × AMD EPYC 9455; 48 cores / 96 threads and 300 W default TDP each

HOST MEMORY

16 × 128 GiB DDR5 RDIMMs = 2 TiB standard; 32 × 128 GiB = 4 TiB capacity option, subject to ordered DIMMs

BOOT / LOCAL DATA

2 × 1.92 TB-class mirrored M.2 boot SSDs; 4 × 7.68 TB E1.S local data SSDs in the reference configuration

PCIe / NUMA

Delivered root complexes, link widths, peer access, IOMMU settings and GPU-to-NIC affinity are validated for the ordered server

EAST-WEST FABRIC

ConnectX-8 switch-board subsystem; 8 × 400GbE-capable ports, with 4 active in Standard Fabric

SERVICE / MANAGEMENT

BlueField-3 in a PCIe 5.0 x16 slot; 2 × 200GbE service ports; separate server BMC and DPU out-of-band access

POWER

4 × 3,200 W power supplies in a 3+1 reference arrangement

COOLING

Front-to-rear air cooling; final fan, power-cap and thermal profile follows the ordered configuration

CONFIGURATION NOTE

The eight GPUs are separate PCIe devices; their 768 GB nominal memory total is not one coherent pool. This configuration is not an NVLink/NVSwitch scale-up system. Final PCIe topology, active fabric ports, GPU power limits, firmware, operating system and thermal profile are fixed in the signed order.

DEFINE THE RIGHT SYSTEM

Align compute, memory, storage, network and acceleration around the workload.

DISCUSS A CONFIGURATION

www.sidera-is.com