

PowerServe Uniti HGX HPC/AI Server

Unique platform built to reach new heights in the AI and ML performance spaces using up to eight NVIDIA A100 GPUs connected via SXM4.

KEY FEATURES

- NVIDIA A100 with 40GB or 80GB: 40GB of VRAM with 1.6 TB/s of memory bandwidth or 80GB of VRAM with 2.0TB/s bandwidth for high-level computational throughput.
- Excellent GPU-to-GPU communication via 3rd gen NVIDIA® NVLink® with 600GB/s bandwidth.
- Reduction in latency and CPU utilization with Mellanox Socket Direct® technology. In this dual-socket server, a single CPU can access the network by bypassing the inter-processor communication bus and adjacent CPU.
- PCIe 4.0 allows for faster interconnect (compared to PCIe 3.0) and low latency for NICs and NVMe drives via PCIe switch fabric.
- Ultra-fast 200Gbps access of GPUs on other servers with RDMA and HDR InfiniBand.



TECH SPECS

GPU Support

- Up to 8 NVIDIA® A100 Tensor Core GPUs connected via SMX4
- 40 GB or 80 GB Memory per GPU

CPU Support

- Dual AMD EYPC 7xxx Series Processors
- Up to 128 Total Cores

Memory Support

- Up to 4TB DDR4 3200 MHz ECC Memory
- 8 Channel Memory Per Processor Socket/32 Total Dimm Sockets

Disk Storage Capacity

- Maximum 8 x SSDs, NVME Storage Devices

Network Support

- 10 GigE, 100 GigE, 200 GigE
- Mellanox Infiniband Option Available
- IPMI Management Network Included

Operating System Support

- Red Hat Enterprise Linux
- CentOS Linux
- Ubuntu Linux
- Windows Server