RACK-SCALE SOLUTION

AMAX ENGINEERING

With 40 years of engineering expertise, our team specializes in transforming standard IT components into high-performance computing solutions with optimized thermal, electrical, mechanical, and networking design.

AMAX RackScale 32 with NVIDIA HGX B200

AMAX RackScale 32 with NVIDIA HGX™ B200 is a rack-scale solution built on the NVIDIA reference architecture, delivering high-performance AI computing with NVIDIA Blackwell GPUs, high-speed interconnects, and offering up to 180GB HBM3e memory per GPU.

KEY FEATURES

- Up to 5.76 terabytes of HBM3e GPU memory per rack
- Direct GPU-to-GPU interconnect via NVIDIA NVLink™ delivers 1.8TB/s bandwidth
- Dedicated one-GPU-to-one-NIC topology

Real-Time Inference and Next-Level Training

AMAX's RackScale 32 delivers up to 15X higher inference performance over the previous NVIDIA Hopper™ generation for massive AI models and the second-generation Transformer Engine, enables up to 3X faster training for large language models.
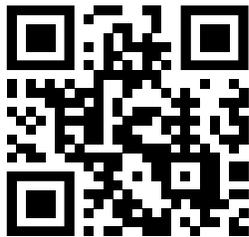
SPECIFICATIONS

AMAX RackScale 32 with NVIDIA HGX™ B200 GPU Specifications	
Networking	2x NVIDIA NDR 400Gbps InfiniBand switches
Storage and Data Management	NVIDIA Certified High Performance Storage Appliance
Compute Nodes	4x NVIDIA HGX B200 8-GPU systems per rack (32x NVIDIA Blackwell GPUs per rack)
Compute Fabric	NVIDIA NDR 400 Gbps InfiniBand
Software	NVIDIA AI Enterprise & Advanced Cluster Management

NVIDIA HGX™ B200 Server Specifications	
Processor	Dual Socket Intel® Xeon®
Memory	16+16 DDR5 DIMM slots (2DPC)
GPU	NVIDIA HGX B200 (1.4TB HBM3e Total GPU Memory), 1.8TB/s NVLINK GPU to GPU Interconnect w/ NVSWITCH
Networking	8x NVIDIA Single-port 400Gbps/NDR OSFP NICs 2x NVIDIA Dual-port 200Gbps/NDR200 QSFP112 NICs 1:1 networking to each GPU to enable NVIDIA GPUDirect RDMA and Storage

DISCOVER AMAX POD SOLUTIONS AND TRANSFORM YOUR DATA CENTER TODAY

High-Speed Networking for AI and HPC

AMAX integrates NVIDIA Quantum InfiniBand for ultra-fast, low-latency networking in AI and HPC workloads. With 400Gbps NDR InfiniBand switches, this high-performance fabric accelerates complex simulations, massive datasets, and parallelized algorithms. Leveraging In-Network Computing offload, AMAX optimizes data movement, reducing bottlenecks and enhancing efficiency while lowering cost and complexity.

Design to Deployment

AMAX provides AI infrastructure solutions built for performance and scalability. We partner with you from initial site assessment to final deployment, handling infrastructure design, cooling and network optimization, GPU and software integration, rigorous testing, and ongoing support. Our expert engineers ensure a reliable, high-performance foundation for AI and HPC workloads.

AMAX // DATASHEET