



Supermicro X14 Gaudi 3 AI Training Platform

8x Gaudi 3 Gaudi 3 rack 2024

OVERVIEW

—	1.0 TB	—	—
Peak Dense	Aggregate Memory	Aggregate Bandwidth	Aggregate Power

CONFIGURATION

Accelerator	Gaudi 3 128GB
Count	8
Memory per GPU	128 GB
Bandwidth per GPU	3.7 TB/s

POWER & EFFICIENCY

Aggregate Power	—
TDP per GPU	900 W
Efficiency	—
Aggregate Bandwidth	—

SYSTEM DETAILS

Vendor: Supermicro	Family: X14	Architecture: Gaudi 3
Form Factor: rack	Release Year: 2024	Process Node: 5nm

INTERCONNECT

Eight Gaudi 3 accelerators all-to-all on an OAM 2.0 universal baseboard, with six on-board OSFP 800GbE ports for scale-out over standard Ethernet

NOTES

Supermicro's X14 platform for Intel Gaudi 3, and by Supermicro's account the first AI server built on the accelerator. Eight Gaudi 3 cards sit on an OAM 2.0 universal baseboard driven by two Intel Xeon 6 processors, with 24 DIMM slots taking up to 6 TB of host DDR5 in one-DIMM-per-channel, up to eight hot-swap PCIe 5.0 NVMe drives, and four further PCIe 5.0 slots. Scale-out is six on-board OSFP 800GbE ports over ordinary Ethernet rather than a proprietary fabric, which is the Gaudi platform's central pitch alongside an open software stack with no licensing cost. Power is eight 3,000 W Titanium supplies in a 4+4 redundant arrangement, which is supply capacity rather than a draw figure, so no system power is recorded. Air-cooled builds use the HL-325L accelerator and liquid-cooled builds the HL-335, which is a part this catalogue does not yet carry. Supermicro publishes no throughput figure for the system, so none is shown; per-accelerator compute is on the Gaudi 3 chip entry. Supermicro describes the baseboard only as an OAM 2.0 or OAM universal baseboard and never names Intel's HLB-325, though that is what the description matches. Search results give the part number as SYS-822GA-NGR3 in an 8U chassis, but neither Supermicro page carrying that detail could be loaded, so it is reported here rather than recorded as a specification.