X14 Max Performance Servers

Game-Changing Performance at Scale for AI, HPC, and Media

New and upgraded architectures supporting a broad range of next-generation technologies

- **Maximum performance systems including GPU-optimized, high-density multi-node, and high performance rackmount**
- **Multi form-factor GPU support including NVIDIA H100, B200, and Intel Gaudi® 3**
- **Single or dual Intel® Xeon® 6900 series processors with P-cores; will also support Intel Xeon 6900 series processors with E-cores in 1Q'25**
- **Support for the latest industry technologies including PCIe 5.0, DDR5, CXL 2.0, OCP 3.0, as well as EDSFF E1.S and E3.S storage drives**
- **Supermicro rack scale plug-and-play service to deliver complete, validated solutions within weeks, not months**
- **Optimized for Supermicro's complete liquid cooling solutions including CPU/GPU cold plates, Cooling Distribution Unit, Cooling Distribution Manifolds, connectors, tubing, and cooling tower**

Engineered for Performance

Introducing the new range of high performance X14 servers from Supermicro. The new X14 range has been fully redesigned to leverage next-generation technologies and maximize performance, including support for new CPUs, next-gen GPUs, upgraded memory, 400G networking, and new high density storage options. Powered by the upcoming Intel Xeon 6900 series processors with P-cores, Supermicro X14 performance-optimized systems will deliver maximum I/O flexibility, storage throughput, and performance-per-core to accelerate even the most demanding workloads.

Latest Generation GPU Support

Whether training complex AI models or undertaking advanced simulations and calculations, GPU acceleration can significantly enhance performance, reduce time-to-result, and lower costs. Supermicro's X14 family supports leading GPU technologies and architectures from NVIDIA, AMD, and Intel, in a range of form factors including SXM5/SXM6, OAM, and PCIe. X14 GPU-optimized systems can be configured to run as a single server, a fully integrated rack, or a multi-rack cluster complete with the networking and storage required to enable maximum GPU utilization.

Ready for Rack-Scale

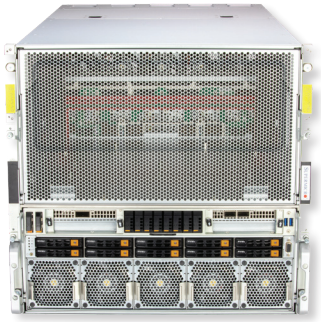
Not only is Supermicro X14 optimized for performance, but X14 GPU-optimized systems are also the foundation for creating rack-scale AI and HPC clusters to address the world's most complex computational challenges. With a production capacity of 5,000 racks per month globally, including 1,350 liquid-cooled 100kW racks, and lead time as short as 2 weeks, Supermicro is unmatched in its ability to design, build, validate, and deliver fully customized, workload-optimized rack-scale solutions to customers, leveraging the most advanced AI hardware available today.

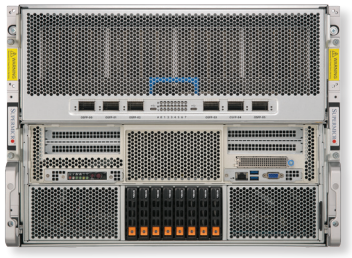
Complete Liquid Cooling Solutions

The new generation of CPUs, GPUs, and other components supported by Supermicro X14 systems produce unprecedented amounts of heat, which must be dissipated effectively to ensure maximum performance and efficiency. Supermicro offers complete, in-house developed liquid cooling solutions including CPU and GPU cold plates, Cooling Distribution Unit (CDU), Cooling Distribution Manifolds (CDM), tubing, connectors and even cooling towers which can be integrated at rack-scale to reduce TCO and TCE. With Supermicro X14 and complete liquid cooling solutions, data center liquid cooling can be free, with a bonus.

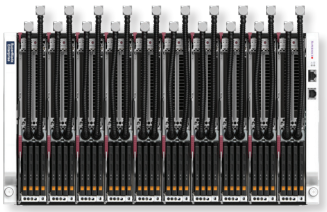
Powered by Upcoming Intel® Xeon® 6900 Series Processors

Supermicro X14 takes performance to the next level, with future support for the new generation of Intel Xeon 6900 series processors with P-cores that will deliver the highest performance-per-core of any Intel Xeon processor ever. Designed for maximum performance and ideal for the most demanding AI, HPC, and cloud environments, Intel Xeon 6900 series processors with P-cores will feature up to 128 cores per socket, include new FP16 instructions on the built-in Intel AMX accelerator to further enhance AI workload performance, and bring new support for MRDIMMs up to 8800MT/s for up to 37% faster memory bandwidth than standard RDIMMs. P-cores are optimized for high performance-per-core and excel at the widest range of workloads, including better AI performance than any other general-purpose CPU.





X14	GPU-Optimized Maximum acceleration for AI Training, LLMs, and Generative AI	PCIe GPU Flexible configurations for AI training, Media, 3D Design, and Simulation	Gaudi® 3 Purpose-Built AI Training and Inference Platform
Family Highlights	• Next-generation architecture for the most intensive AI workloads • Dual Intel® Xeon® 6900 series processors with P-cores • Up to 8 NVIDIA SXM GPUs • Up to 10 PCIe 5.0 slots • Support for DDR5-6400 and 8800MT/s MRDIMMs • Up to 10 hot-swap 2.5" NVMe + 8 E1.S EDSFF drives • Direct-to-chip CPU and GPU liquid cooling options	• Dual Intel® Xeon® 6900 series processors with P-cores • Up to 10 double-width NVIDIA, AMD, or Intel PCIe GPUs • Up to 13 PCIe 5.0 slots • Support for DDR5-6400 and 8800MT/s MRDIMMs • Up to 24 NVMe drives • Enhanced thermal design to support up to 10 GPUs with free-air cooling • Direct-to-chip CPU and GPU liquid cooling options	• 8 Gaudi 3 HL-325L (air-cooled) or HL-335 (liquid-cooled) accelerators on OAM 2.0 baseboard • Industry's only Gaudi 3 platform powered by Intel® Xeon® 6900 series processors with P-cores • Support for DDR5-6400 and 8800MT/s MRDIMMs • Up to 8 hot-swap PCIe 5.0 NVMe drives • 8 high efficiency 3000W fully redundant (4+4) Titanium Level power supplies • 6 on-board OSFP 800GbE ports for scale-out networking up to 100+ nodes • 2 PCIe 5.0 x16 (FHHL) + 2 PCIe 5.0 x8 (FHHL) expansion slots
Key Applications	• Large-scale AI Training • Large Language Models • AI/Deep Learning Training • Industrial Automation • Conversational AI • Drug Discovery • Climate and Weather Modeling • Finance & Economics	• AI Model Training • Digital Twins • 3D Simulation • Real-time Ray-tracing • Animation and Modeling • Cloud Gaming • Design & Visualization • 3D Rendering • VDI • Media/Video Streaming • Diagnostic Imaging	• AI Training • AI Inference
Form Factor	10U air cooled rackmount 4U liquid cooled rackmount	5U rackmount	8U rackmount
Socket Count/Type	Dual Intel® Xeon® 6900 series processors with P-cores	Dual Intel® Xeon® 6900 series processors with P-cores	Dual Intel® Xeon® 6900 series processors with P-cores
GPU Compatibility	8 SXM6 GPUs; NVIDIA H100, H200, B200	Up to 10 PCIe GPUs	8 Intel® Gaudi® 3 OAM 2.0 accelerators
Storage	Up to 10 hot-swap 2.5" NVMe	Up to 10 hot-swap 2.5" NVMe	Up to 8 hot-swap 2.5" NVMe drives



X14	**6U SuperBlade®** Memory-Optimized Multi-Node Architecture for AI and HPC	**FlexTwin™** Purpose-Built HPC-at-Scale Solution	**Hyper** Flagship Performance Rackmount Architecture
Family Highlights	• 100 servers per rack (Up to 12,800 high performance CPU cores) • 6U enclosure with 10 single-wide or 5 double-wide servers, sharing power supplies, cooling fans, CMMs, and Ethernet switches • Single Intel® Xeon® 6900 series processor with P-cores per node • Support for DDR5-6400 and 8800MT/s MRDIMMs • Up to 4 E3.S and 2 M.2 NVMe devices per server • Up to 4 GPUs or network cards per server • 400G Infiniband or Ethernet (PCIe 5.0 x16 slots), and up to 4x 25G Ethernet switches with 100G uplinks • Reusable enclosure, power supplies, cooling fans, CMMs, and switches for future generation servers • 96% efficiency, (N+N / N+1) Titanium Level redundant power supplies • Direct-to-chip liquid cooling option	• 2U 4-node architecture • Front-accessible nodes with front I/O and storage for cold-aisle servicing • Dual socket per node supporting Intel® Xeon® 6900 series processors with P-cores and direct-to-chip liquid cooling • Support for DDR5-6400 and 8800MT/s MRDIMMs • EDSFF E1.S drive support • Flexible networking options with AIOM or PCIe AOC • Integrated front I/O management module • N+N redundant power supplies	• Single or dual socket configurations supporting Intel® Xeon® 6900 series processors with P-cores and direct-to-chip liquid cooling • Support for DDR5-6400 and 8800MT/s MRDIMMs • All-hybrid hot-swappable NVMe/SAS/SATA; up to 24 drive bays • Flexible networking options with up to 2 AIOM networking slots (OCP NIC 3.0 compatible) • Optional PCIe slot configurations up to 8 PCIe 5.0 x8 or 4 PCIe 5.0 x16 slots with support for double-width GPU/Accelerator cards
Key Applications	• Artificial Intelligence • Machine Learning • HPC • Big Data Analytics • Financial Services	• HPC Data Center • Financial Services • Manufacturing • Climate & Weather Modeling • Oil & Gas • Scientific Research	• Artificial Intelligence (AI) • Data Analytics • High Performance Computing • Virtualization
Form Factor	6U enclosure with single or double-wide blades	2U4N rackmount	1U or 2U rackmount
Socket Count/Type	Single Intel® Xeon® 6900 series processor with P-cores	Dual Intel® Xeon® 6900 series processors with P-cores per node	Single or Dual Intel® Xeon® 6900 series processors with P-cores
Storage	4 hot-swap E3.S NVMe SSDs 1 M.2 NVMe SSD Additional 4 M.2 NVMe SSDs with optional mezzanine card	Up to 2 hot-swap E1.S NVMe SSDs per node	Up to 24 hot-swap 2.5" NVMe SSDs