H12 2U 2-Node Multi-GPU

Multi-Node Design for Compute and GPU-Acceleration Density

A+ Server 2114GT-DNR

2 Nodes in 2 Rack Units with 3–6 GPUs per Node

Power GPU-accelerated workloads with a high-density solution:

- Resource-efficient hot-pluggable 2-node architecture in 2 rack units
- Single 2nd or 3rd Gen AMD EPYC™ Processor per node
- Up to 3 double-width or 6 single-width GPUs per node
- Up to 8 DIMMs per node for up to 2 TB of DDR4-3200 memory
- Flexible internal and front-panel storage options
- Flexible networking with OCP 3.0 interfaces (AIOM)
- Redundant 2600W Titanium Level power supplies

If your workload calls for GPU acceleration, our 2U 2-Node Multi-GPU server combines the computing power of AMD EPYC™ processors with your choice of up to six double-width or twelve single-width GPUs—all in a dense 2RU form factor.

Redefine Datacenter and Edge Computing

This Supermicro system redefines how you propel graphics-intensive workloads regardless of whether in the datacenter or at the edge. You gain high efficiency and lower costs compared to discrete servers through the platform's shared power and cooling systems. Further reducing cost is the single AMD EPYC processor that delivers up to 64 cores of processing power, delivering in a single chip the performance that not long ago required two CPUs to attain. The AMD EPYC processor uses a system-on-chip design that eliminates the need for chip sets and external disk controllers. For powering GPU-intensive workloads the AMD EPYC CPU connects every accelerator with 16 lanes of PCI-E 4.0 bandwidth with room left for 200 Gpbs of network connectivity.

Deploy for Virtual Desktop Infrastructure

When your power users demand high-end workstation-class computing, whether they are on site or remote, the A+ Server 2114GT-DNR can deliver a dedicated GPU for up to six users. Use virtualization software to partition the CPU, giving each of your users dedicated datacenter-class CPU power, more than they would typically have at their desktop.

Power Machine Learning and AI Inferencing

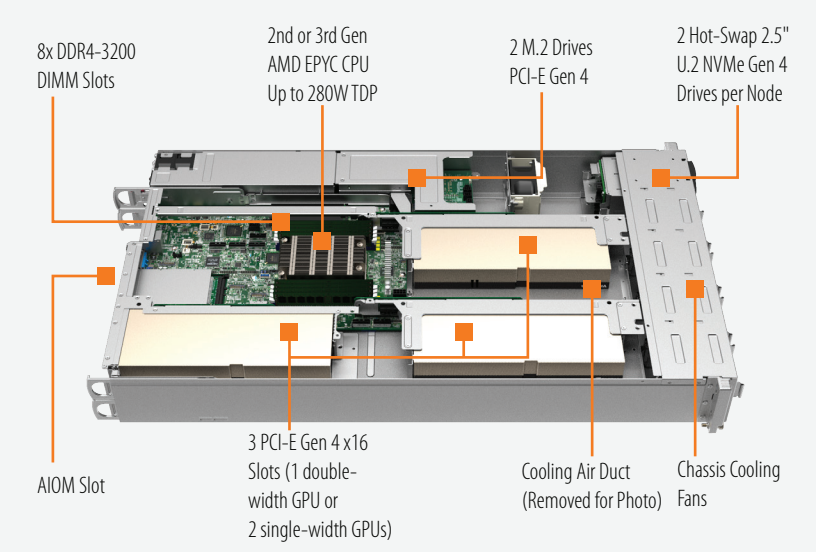
The AMD EPYC CPU's record-setting integer and floating-point performance is matched with 128 lanes of PCI-E 4.0 bandwidth. The processor excels at moving data quickly between disk and network and the GPUs, whether AMD Instinct™ or NVIDIA® Ampere. At edge locations, where artificial intelligence has to perform, the server has more than enough GPU power for AI inferencing.

Accelerate Media Transcoding

GPU acceleration makes media transcoding in real time a snap. With the GPU density offered by our 2U 2-Node Multi-GPU servers combined with maximum PCI-E 4.0 bandwidth to each GPU, and I/O expansion for high-speed networking, the AS -2114GT-DNR is an excellent choice.

Drive Cloud Gaming and 3D Rendering

The GPU density of the A+ Server 2114GT-DNR makes it an excellent choice for supporting cloud gaming back-end software and 3D rendering for both gaming and motion picture development. When you host your 3D intensive applications on this server, you can set a higher bar for realism and responsiveness



H12 Generation	Single-Socket AS -2114GT-DNR Node
Processor Support	• Single SP3 socket for one AMD EPYC™ 7002 or 7003 Series processor • Up to 64 cores, up to 280W TDP†
Memory Slots & Capacity	• 8 DIMM slots for DDR4-3200 MHz RDIMM/LRDIMM; up to 2 TB registered ECC • 8GB, 16GB, 32GB, 64GB, 128GB, and 256GB DIMM sizes supported
On-Board Devices	• System on Chip • 6-Gbps SATA3 storage interface via AMD EPYC processor' • IMPI 2.0 with virtual-media-over-LAN and KVM-over-LAN support • ASPEED AST2600 BMC graphics
I/O Ports	• Flexible networking via AIOM, 1 per node • 1 RJ45 dedicated IPMI LAN port • 2 USB 3.0 ports (rear) • 1 VGA, 1 COM port • 1 TPM 2.0 header
Expansion Slots	• 6 PCI-E 4.0 x16 (4 internal, 2 external) slots • 1 PCI-E 4.0 x8 AIOM networking slot
Drive Bays	• 2 M.2 NVMe/SATA3 slots per node, PCI-E 4.0 (x4) • M.2 Key
BIOS	• 256 Mb SPI Flash ROM with AMI BIOS
System Management	• Integrated IPMI 2.0 plus KVM with dedicated LAN • Supermicro Server Manager (SSM) and Supermicro Platform Manager (SPM), and Supermicro Update Manager (SUM) • Supermicro SuperDoctor® 5 and Watch Dog
Chassis	
Form Factor	• 2U rackmount
Drive Bays	• 4 hot-swap 2.5" U.2 NVMe Gen 4 drives (front panel, 2 per node)
Shared Power & Cooling	• 4 hot-swap heavy-duty 80 mm PWM fans • Redundant 2600W Titanium Level power supplies with PM Bus

†Certain CPUs with high TDP may be supported only under specific conditions. Please contact Supermicro Technical Support for additional information about specialized system optimization

Ready for Your Choice of Compatible GPU

The AS -2114GT-DNR server supports the GPU most suitable for your workload including the most popular accelerators from AMD and NVIDIA. The maximum number of GPUs per node are determined by compatibility and the power and cooling envelope of the server.

AMD Instinct	Max per Node
MI100	3
NVIDIA	
A2	6
A10	5
A16	3
A30	3
A40	3
A100	3
A4000	5
RTX A5000	3
RTS A6000	3
T4	6

Flexible Networking

Each node includes a single Advanced I/O Module (AIOM) slot that is Open Compute Project (OCP) 3.0 compliant. This means that you can select industry-standard interfaces — such as dual 100 Gigabit Ethernet and 100G InfiniBand EDR — from the vendors you prefer, including Broadcom, Intel, NVIDIA, and Mellanox.

With OCP 3.0-compliant I/O connectivity, you enjoy the benefits of this industry standard, including:

- **Better thermal characteristics** from increased airflow
- **Easy serviceability** through tool-less installation
- **TCO optimization** by decreasing time to service and minimizing down time.

Designed for Flexibility

This system is designed to flexibly meet the needs of your applications. You can choose from the entire range of AMD EPYC 7002 and 7003 Series processors to match CPU power to your workload. You can mix and match NVMe and SATA storage to balance performance and capacity to best power your workloads. And you can use the optional internal M.2 slots to extend storage capacity even further.