

BUILT FROM THE GROUND UP FOR THE INTERSECTION OF AI & HPC

NVIDIA® GPU Solutions



NumberSmasher® 4U Tower/GPU Server

Balanced CPU:GPU Configuration

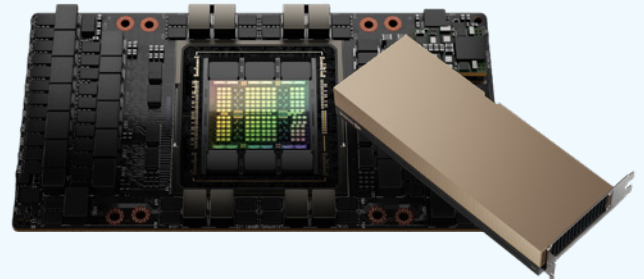
A balanced, versatile configuration with 2 CPUs + 4 GPUs split evenly between them suitable for many HPC applications. These 4U GPU servers offer additional PCI-E slots and optional hardware RAID support for up to 8x drives. Support for a wide array of GPUs and PCI-E devices is available.



NumberSmasher, Navion 4U HGX H100 4-GPU

Fast GPU:GPU Transfers with NVLink Interconnect

Tackle your most data-intensive AI training or HPC workloads with NVIDIA HGX™ H100 4-GPU. With two Intel or AMD CPUs + four NVIDIA H100 GPUs interconnected with 4th Gen NVIDIA NVLink* (300GB/s between any two GPUs), you can deploy the latest in CPU and GPU technology.



NVIDIA H100 Tensor Core GPUs

Unprecedented Performance, Scalability & Security for Every Datacenter

- Max 34 TFLOPS FP64 | 67 TFLOPS FP32 | 989 TFLOPS TF32 | 1979 TFLOPS FP16 Tensor Core/BFLOAT16
- 4th Gen NVIDIA NVLink interconnect: max 900GB/s
- 80GB memory, max 3.35TB/s Memory Bandwidth
- SXM5 or PCI-E form factor, with PCI-E Gen5 interface
- Leadership HPC, AI Training, or Inference Performance



Navion® 4U 8/10 GPU, Octoputer™ GPU Servers

Highest GPU Capacity

Scale up to 8 or 10 NVIDIA H100 Tensor Core GPUs with 260 FP64 TFLOPS for HPC, 14.78 PFLOPS of BFLOAT16 for AI. Ideal for highly accelerated GPU applications & GPUDirect® P2P transfers over PCI-E Gen5. Optional InfiniBand for clustering, RAID, NVMe storage, NVIDIA RTX™ professional graphics.



NumberSmasher HGX H100 8-GPU Servers

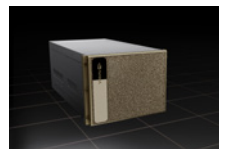
Optimized for AI Training + Fully Configurable

Deploy NVIDIA's HGX H100 8-GPU platform for superior AI Training—up to 15.46 PFLOPS FP16 Tensor Core/BFLOAT16. Four 3rd Gen NVIDIA NVSwitches provide fully non-blocking bandwidth of 900GB/s between any 2 GPUs & 7.2TB/s of total GPU fabric BW. A fully configurable platform based on the underlying architecture used in DGX H100.

NVIDIA DGX H100™

The Gold Standard for AI Infrastructure

- Complete AI appliance with 15.46 PFLOPS of AI training performance
- End-to-end software stack: from data prep to cluster management
- 8 NVIDIA H100 with NVLink GPUs
- Easy Shared Usage: 2nd Gen Multi Instance GPU for up to 7 confidential partitions per GPU
- Available in NVIDIA DGX POD™ and SuperPOD™ scale-out deployments to build your AI Center of Excellence



Software Integration Experts

- NVIDIA-Certified System™ HW + SW configurations available for many offerings
- NVIDIA AI Enterprise, vGPU, or Omniverse™ integration
- NVIDIA CUDA® SDK or HPC SDK installed and configured
- Microway Cluster Management Software (MCMS™) with Open OnDemand support for clusters
- Legendary Service + Technical Support



Microway Scalable NVIDIA® GPU Clusters



NVIDIA GPU Clusters

Microway's robust, NVIDIA GPU-based clusters offer high-speed interconnects, NVIDIA Datacenter GPUs, NVIDIA GPU Software, MCMS Remote Cluster Management and Monitoring Tools, plus the NVIDIA CUDA® and HPC SDKs.

Microway provides these fully integrated Linux clusters at very competitive prices. Users worldwide pushing the limits of technology in life sciences, universities, commercial and government research count on our expertise and attention to detail.

Delivering Innovative HPC Solutions Since 1982



	NVIDIA Datacenter GPU Cluster Options
Processor	Intel® Xeon® Scalable Processors AMD® EPYC™ 9000 Series Processors
Processor Sockets	1P, 2P
Compute Nodes	NumberSmasher® (Xeon) & Navion® (EPYC) 2U, 4U GPU Servers NVIDIA DGX™ AI Systems
GPU Options	NVIDIA H100, NVIDIA A100, or NVIDIA RTX™ Professional
GPU Companion Software	NVIDIA AI Enterprise, NGC™, RAPIDS™, vGPU, Omniverse™
Visualization Node Option	NumberSmasher or Navion 1U, 2U, 4U GPU Servers
Memory per Node	2P: up to 6 TB DDR5
Disk/Media Bays	4U: 8x3.5" Hard Drives; OctoPuter: 24x 2.5" Hard Drives Diskless compute node configurations available
Cluster Interconnect	NVIDIA ConnectX®-7 NDR 400Gb or ConnectX®-6 HDR InfiniBand Omni-Path; 200G, 100G, 50/25G, 10G, or Gigabit Ethernet
Management Interfaces	IPMI v 2.0, OpenBMC, and Redfish
Storage Options	Ethernet or InfiniBand-attached NFS Storage Servers DDN EXAScaler® Lustre, VAST Data, BeeGFS®, or Panasas®
Operating Systems	Linux: Rocky, Red Hat, Ubuntu, more; Windows® Server
Compilers	NVIDIA CUDA, NVIDIA HPC SDK, Intel and GNU
Cluster Software	MVAPICH2, OFED, and OpenMPI; SLURM, PBS Professional, or IBM Spectrum LSF
Monitoring & Management Software	Microway Cluster Management Software (MCMS™) MPI Link-Checker™ and InfiniScope™ NVIDIA Bright Cluster Manager or IBM Spectrum Cluster Manager
Cabinets and Infrastructure	APC® NetShelter™ Cabinets, APC PDUs and UPS Protection
Services	Optional onsite installation, Factory Pre-Installation Service (Power & network pre-wiring, pre-installation of rails) Optional complete-rack air-ride shipment (roll-off, power-on)
Hardware Warranty	2 years with advanced replacement parts or return-to-factory Optional extended warranty term

www.microway.com/nvidia

Call our technical sales team at **508.746.7341**