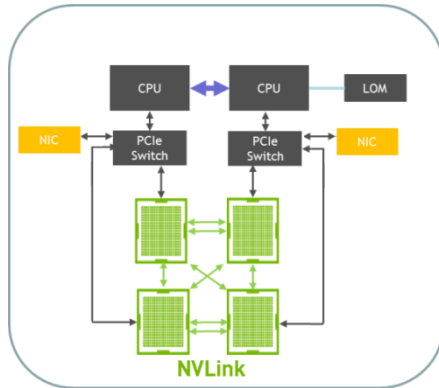


Deep Learning GPU Accelerated Server with 4 NVIDIA V100 SXM2

The 1U GPU accelerated server (OSS-VOLTA4) sets a new standard for HPC compute power in a space efficient design. Supporting four of the latest Volta-based NVIDIA GPUs, the OSS-VOLTA4 provides 31.2 TeraFLOPS of double precision performance for the most demanding HPC applications. For state-of-the-art deep learning workloads the OSS-VOLTA4 provides 62.8 TeraFLOPS of single precision performance. GPU management and monitoring and software are preinstalled on the OSS-VOLTA4. The GPU accelerated server also includes dual high-performance “Broadwell” E5 2698v4 2.2GHz processors and a base configuration of 512GB of DDR4 memory scalable to 1TB. Four PCIe Gen3 slots are available for additional expansion and for scale out creating GPUultima clusters using IB or high speed Ethernet networking. The appliance also includes two 2.5” removable 1.9TB SATA SSD drives.

Part Number: OSS-VOLTA4



Features

- 1U Chassis
- Dual Intel Xeon 3.2GHz CPUs
- Up to 1TB DDR4 LRDIMM System Memory
- Two 2.5” 1.9TB SATA SSDs
- Four Volta GPU SXM2 with 80GB/s NVLink
- Three x16 PCIe 3.0 slots
- One x8 PCIe 3.0 slot
- Two 2000W Titanium Power Supplies
- GPU Management and Monitoring pre-installed
- Software bundle pre-installed

Advantages

- All GPUs capable of peer-to-peer direct access to all other GPUs’ memory as well as direct transfer operations via NVLink at high Bandwidth
- High performance for collective communications
- PCIe bandwidth fully available for host and/or NIC communication during inter-GPU communication

Specifications

Dimensions:	1U Rack Units <ul style="list-style-type: none"> • 1.75” H x 17.2” W x 35.2” D (39.3” with rails)
CPUs:	Dual 16-core Intel® Xeon® processors E5-2698v4 3.2GHz
System Memory	Memory Capacity <ul style="list-style-type: none"> • 16x 288-pin DDR4 DIMM slots Memory Type <ul style="list-style-type: none"> • 2400 MHz ECC DDR4 SDRAM 72-bit DIMM Sizes <ul style="list-style-type: none"> • LRDIMM: 64GB, 32GB
GPUs:	4x Tesla V100 <ul style="list-style-type: none"> • 62.8 TeraFLOPS single precision; 31.2 TeraFLOPS double precision • 16 GB HBM2 Memory per GPU • 20480 NVIDIA CUDA cores
Software Bundle	Choice of Operating System <ul style="list-style-type: none"> • CentOS7 • RHEL7 • SL7 Choice of Machine Learning Framework <ul style="list-style-type: none"> • Caffe2 • Pytorch • Mxnet

Software Bundle (Cont.)	<ul style="list-style-type: none"> • Microsoft Cognitive Toolkit • Tensorflow • Theano MLPython ML Dependencies (400MB Python) cuDNN (5.0 & 5.1) DIGITS Caffe on Spark CUDA & NVIDIA driver CUB (CUDA building blocks) NCCL GPU Management from Bright Computing <ul style="list-style-type: none"> • Health Management • Workload Integration
Expansion Slots	PCI-Express <ul style="list-style-type: none"> • Three PCIe 3.0 x16 slots • One PCIe 3.0 x8 low-profile slot
On-board devices:	Chipset <ul style="list-style-type: none"> • Intel® C612 chipset SATA <ul style="list-style-type: none"> • SATA3 (6Gbps) with RAID 0, 1, 5, 10 IPMI <ul style="list-style-type: none"> • Support for Intelligent Platform Management Interface v.2.0 • IPMI 2.0 with virtual media over LAN and KVM-over-LAN support • ASPEED AST2400 BMC Network Controllers <ul style="list-style-type: none"> • Intel® X540 Dual Port 10GBase-T • Virtual Machine Device Queues reduce I/O overhead • Supports 10GBASE-T, 100BASE-TX, and 1000BASE-T, RJ45 output Graphics <ul style="list-style-type: none"> • ASPEED AST2400 BMC
Drive Bays:	Hot-swap <ul style="list-style-type: none"> • Two 2.5" Hot-swap drive bays Fixed <ul style="list-style-type: none"> • Two 2.5" internal drive bays
System BIOS:	BIOS Type <ul style="list-style-type: none"> • 128Mb SPI Flash EEPROM with AMI® BIOS
Front Panel:	Buttons <ul style="list-style-type: none"> • Power On/Off button • System Reset button LEDs <ul style="list-style-type: none"> • Power LED • Hard drive activity LED • Two Network activity LEDs • System Overheat LED / Fan fail LED / • UID LED
Cooling Fans:	<ul style="list-style-type: none"> • 8x 4cm heavy duty counter-rotating fans with air shroud & optimal fan speed control
Power Supply	2000W Redundant Power Supplies with PMBus Total Output Power <ul style="list-style-type: none"> • 1000W/1800W/1980W/2000W Dimension (W x H x L) <ul style="list-style-type: none"> • 73.5 x 40 x 265 mm Certification <ul style="list-style-type: none"> • UL/cUL/CB/BSMI/CE/CCC Titanium Level
Environment:	Operating Temperature: <ul style="list-style-type: none"> • 10°C to 30°C (50°F to 86°F) Non-operating Temperature: <ul style="list-style-type: none"> • -40°C to 70°C (-40°F to 158°F) Operating Relative Humidity: <ul style="list-style-type: none"> • 8% to 90% (non-condensing) Non-operating Relative Humidity: <ul style="list-style-type: none"> • 5% to 95% (non-condensing)
Compliance:	RoHS 6 of 6, WEEE