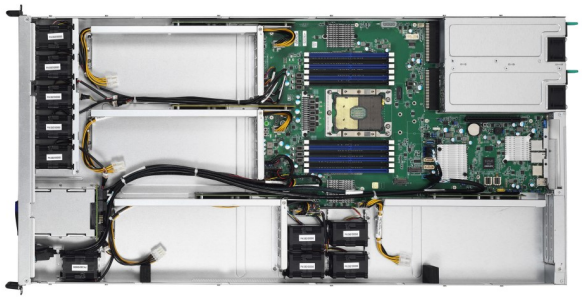
BRESSNER
A ONE STOP SYSTEMS COMPANY

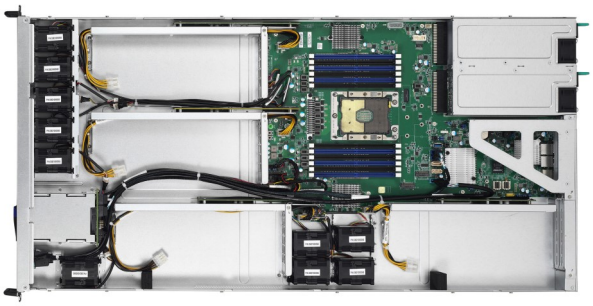
SHARK A.I. 4-GPU

High Performance 4-GPU Server for Machine Learning, Deep Learning and Artificial Intelligence

Features

- 4x PCIe x16 Gen3 GPU
- 1U Rackmount
- Single Socket Xeon® Scalable Processor
- 12 DDR4 DIMM Slots (up to 384GB)
- 2x SATA 6Gb/s Hot-swappable 2.5" Drives
- 1600W AC/DC Platinum PSU, 1+1 redundancy
- 2x 10GbE Ports





System	
CPU	Single Socket Intel® Xeon® Scalable Processor
Chipset	Intel® C621
GPU	Supports Single & Dual root complex Design
Memory	Up to 384GB DDR4-2666 RDIMM
Storage	2x SATA 6Gb/s hot-swappable 2.5" Drives
TPM	optional
GPGPU	4x PCIe x16 Gen3 GPU Slots
Expansion Slot	1x PCIe Gen3 x16 (Full Height / Half Length)
Interface	
USB	Front: 2x USB 3.0, Rear: 1x USB 3.0
Ethernet	2x 10GbE Ports, 1x GbE dedicated for IPMI
Video	1x VGA Port
COM	1x 2x5 Pin Header
Mechanical	
Power Supply	1600W AC/DC Platinum Power Supply 1+1 redundancy Input: 1000W: 100-127 Vac / 12A 1600W: 220-240 Vac / 9,48A
Cooling	10x 4cm fans with redundancy
Dimensions	438mm x 43,5mm x 885mm (W x H x D)
Form Factor	1U Rackmount
Certificates	
Regulation	FCC (DoC), CE (DoC), CB/LVD, RCM, VCCI
RoHS	RoHS 6/6 compliant

BRESSNER
A ONE STOP SYSTEMS COMPANY

SHARK A.I. 4-GPU

Introduction

Artificial Intelligence (AI) nowadays is not only an academic subject, but is moving fast towards the real world with applications in facial recognition, robotics, revolutionary analytics, disease prevention and smart city constructions. All the groundbreaking scientific progress calls for acceleration in machine-learning (ML) and deep -learning (DL) training, and the increasing adoption of GPUs will satisfy the thirst for tremendous computing power.

SHARK A.I. is a carrier-grade, multi-purpose platform designed for edge applications. Combining a server node, a PCIe expansion box with PCIe switching, the DEVKIT has a capacity to support a combination of up to four NVIDIA® GPUs depending on the needs of the application. A configurable edge platform that supports different workloads and demands, the SHARK A.I. 1U Server can support multiple topologies and bandwidths between GPUs and CPUs with simple cable routing adjustments. Moreover, Infiniband support allows it to be easily scaled up to multiple GPU clusters.

Framework Flexibility for Various AI Applications The SHARK A.I. 1U Server supports both single and dual root complexes for various AI applications. For deep learning applications, a single root complex can utilize all the GPU clusters to focus on large-size data training jobs and the use CPU to handle small tasks; For machine learning applications, a dual root complex can allocate more tasks to the CPUs, and arrange fewer distributed data training jobs among GPUs. The flexible framework of the AI SHARK A.I. makes it an extremely flexible AI platform. Enabling a switching option to configure specific PCIe lanes of GPU's to specific I/O and CPU cores enhances the ability to improve the overall flow of information to and from multiple virtualized applications. This provides the developer a broad range of options for configurability and manageability without the need to rack and stack systems, eating up valuable space, power and cooling.

Flexible System Built for Edge High Performance Computing (EHPC) Moving high compute systems to the edge for GPU acceleration is critical for specific solutions needing to optimize high performance computing (HPC) applications and remote virtualization. The SHARK A.I. Server is able to increase cloud-scale flexibility and agility at the edge. It provides the flexibility to implement different head-nodes and the freedom to choose the numbers of GPUs per virtual machine (VM). It is an ideal hardware system that can support a wide variety of configurations via software implementation.

Rugged and Carrier-Grade for Reliability and Serviceability Designed for reduced OPEX and system reliability, the hardware structure of the SHARK A.I. 1U Server is all hot-swappable, has redundant fan modules and redundant hot-swappable 1+1 power supplies. GPU cards can be easily installed after removal of the top cover. The SHARK A.I. promotes efficient serviceability while delivering optimal performance.