AIVRES

KR4268V2

4U 8 GPU PCIe 5.0 AI Server with Flexible Expansion

KR4268V2 is a powerful 4U AI server with the latest PCIe Gen 5.0 architecture, supporting up to 8 FHFL PCIe GPU cards. Its flexible configurations and topologies support a variety of advanced AI applications including AI cloud, IVA, video processing, and more.

- 8 FHFL double-width GPU in 4U
- Supports additional 2x double-width PCIe GPU, 1x FL single-width PCIe GPU
- Supports NVIDIA L40S, H100, and A100 GPUs
- 2x AMD EPYC™ 9004 processors or 2x 4th Gen Intel® Xeon® Scalable processors
- Front/back I/O, supports flexible operating scenario
- Supports CPU+GPU+DIMM liquid cooling

Advanced Performance

- PCIe 5.0 architecture supports up to 600W power consumption
- Built-in 8TB system memory and 300TB local high-speed storage for improved data access speed

Enhanced Architecture

- Latest CXL1.1 interconnect technology supports storage-level memory space expansion
- Supports latest 400G NDR infiniband network and 400G OCP 3.0

Flexible Configuration

- Modular design concept supports flexible component-level configuration
- 4 CPU-GPU topologies for various AI application scenarios

Technical Specifications

Model	KR4268-E2-A0-R0-00	KR4268-X2-A0-R0-00
Form Factor	4U rack server	
Processor	2x AMD EPYC™ 9004 processors, up to 400W	2x 4th Gen Intel® Xeon® Scalable Processors
GPU	8x FHFL double-width PCIe 5.0 GPU cards Supports additional 2x FL double-width PCIe GPU, 1x FL single-width PCIe GPU Supports L40S, H100, A100	
Memory	24x DDR5 4800MT/s RDIMMs	32x DDR5 4800MT/s RDIMMs
Storage	24x 2.5" SSD, up to 16x NVMe U.2 2x NVMe/SATA M.2	
Expansion Slots	11x PCIe 5.0 x16 slots	
I/O	**Front:** 1x USB 3.0, 1x USB 2.0, 1x VGA, 1x Type-C **Rear:** 2x USB3.0, 1x RJ45, 1x MicroUSB, 1x VGA	
OCP	1x OCP 3.0 SFF, Support NCSI	
Management	Aspeed 2600	
Security	TPM 2.0 (Trusted Platform Module)	
PSU	4x 12V 3000W, 2+2	
Dimensions	447mm (W) x 174.5mm (H) x 850mm (D)	