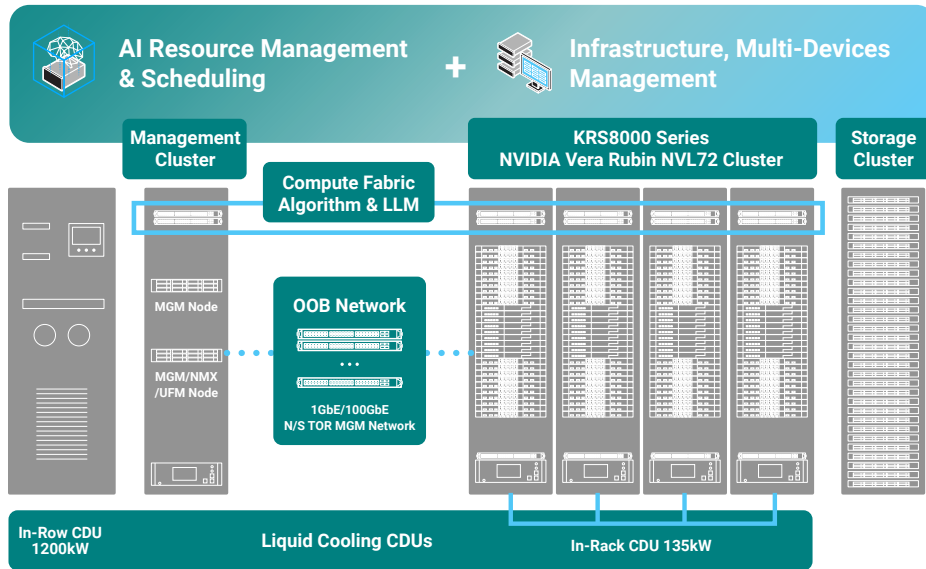


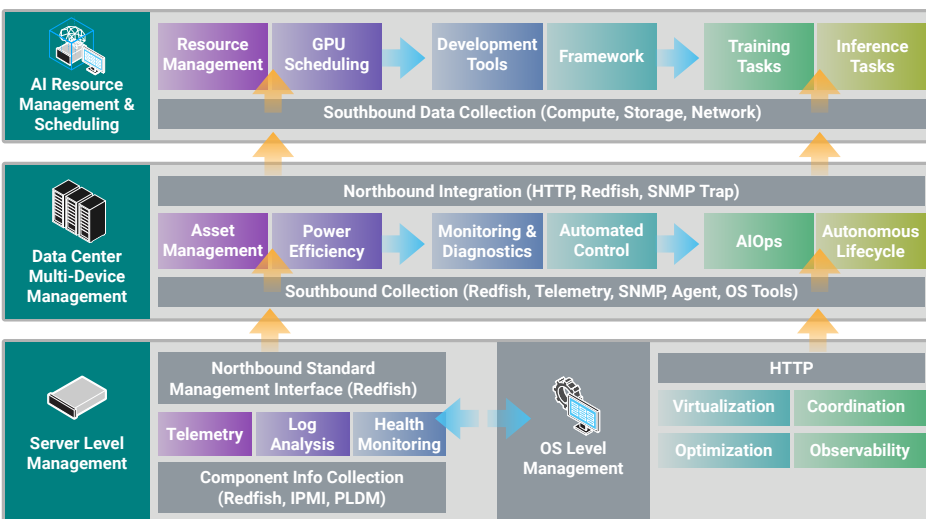
Total AI Data Center Solution

Complete Liquid-Cooled Cluster Architecture + Data Center Management Software

Aivres End-to-End Liquid-Cooled Rack Scale Solution



Fullstack Data Center Management Software



Experience the Aivres Difference

We enable customers to build AIDCs with expertise, agility, and reliability at every stage from design to deployment.



Total AI Infrastructure Solutions



Hardware Design Innovation

With our product design experience, engineering expertise, and rigorous R&D, we are constantly pushing the boundaries of technology to deliver products with groundbreaking performance features.



Empowering Intelligence

Our unique capabilities in deep learning architectures, advanced hardware platforms, and infrastructure expertise enable us to create turnkey solutions to help customers harness the power of AI.



Sustainable Computing

Sustainability is critical to continuous technology innovation, and Aivres is committed to the development and adoption of energy-efficient solutions like liquid cooling to reduce the climate impact of computing.



Open Collaboration

Aivres collaborates actively with various partners and communities to drive leading-edge product design and new technology adoption, so that our solutions can best serve the customer's evolving needs.



Customer-First Approach

We work closely with the customer to bring about their product vision, combining our design, manufacturing, and deployment capabilities to deliver a custom-made solution that is right for their business.



Discover unlimited possibilities with Aivres total AI solutions accelerated by NVIDIA.



AI Data Center Building Blocks from Server Platforms to Rack Solutions

Liquid-Cooled Exascale Rack Solution based on NVIDIA Vera Rubin NVL72

KRS8000V4 Rack-Scale Solution for Next-Generation Agentic AI



Based on NVIDIA Vera Rubin NVL72 integrating 72 Rubin GPUs and 36 Vera CPUs with NVLink 6, ConnectX-9, BlueField-4, and Spectrum-X to deliver 10X increased performance per watt

Rubin-Based AI Compute Architecture

- NVIDIA Rubin GPUs with next-generation tensor cores optimized for trillion-scale LLM and MoE workloads
- Designed to maximize compute density within a single NVL72 rack

Rack-Scale & Data Center-Scale Interconnect

- NVLink™ 6 switch system enabling low-latency, rack-scale GPU communication
- NVIDIA ConnectX-9 and BlueField-4 enabling InfiniBand/Ethernet scale-out for AI factories




High-Bandwidth Memory & Data Path

- Expanded GPU memory capacity and bandwidth to support long-context inference
- 75 TB fast memory tier for checkpointing, data staging, and KV-cache expansion

Liquid-Cooled, Deployment-Ready Design

- Fully liquid-cooled compute and switch trays optimized for high power density
- Modular rack, manifold, and CDU options to support diverse data center environments

AI Data Center Rack-Scale Solutions with NVIDIA Blackwell & Blackwell Ultra Platforms

NVIDIA GB300 NVL72 KRS8500V3	NVIDIA GB200 NVL72 KRS8000V3	Liquid-Cooled NVIDIA HGX™ B200 / B300 Rack with KR5288	Air-Cooled NVIDIA HGX™ B200 / B300 Rack with KR9288
Exascale supercomputing AI solution with fully integrated rack-scale liquid cooling for trillion-parameter models		High GPU density, energy-efficient, and flexible deployment	Fast deployment in existing traditional data centers
			

Total Liquid Cooling AI Solutions

Sustainable AI development needs liquid cooling to offset the growing power and thermal needs of intensive AI workloads.

Aivres provides end-to-end, integrated liquid cooling solutions from server to rack to total data center architecture to support AI deployments at any scale.



Server Cold-Plate Cooling

Direct-to-chip liquid cooling adopted in Aivres AI server platforms support efficient operation in high performance, high heat workload scenarios.



Liquid Cooled AI Racks

Total rack-scale liquid cooling solutions for exascale AI racks supporting leading-edge data center platforms like NVIDIA Vera Rubin NVL72 and NVIDIA Blackwell.



Liquid Cooled AI Data Center

Complete data center infrastructure solutions encompassing cooling components for every business scale, in-depth architecture planning, and integrated deployment.

Next-Generation Peak Performance Servers with NVIDIA HGX™ Rubin NVL8

KR2288 2U Liquid-Cooling AI Server



- Next-generation HBM4 memory architecture enables 5x performance improvement at NVFP4 precision
- High efficiency, upgrade-ready fully liquid-cooling architecture
- Front-accessible OS drive with cold-swap support for improved serviceability
- Optimized for long-context processing, million-token inference, large-scale generative models

KR5288 5U Hybrid-Cooling AI Server



- Hybrid air and liquid cooling design compatible with traditional air-cooled data center architectures
- Flexible scale-out with easy deployment without modifying power supply, cooling, or racks
- Robust I/O expansion capabilities
- Purpose-built for LLM training, multimodal inference, and general AI workloads

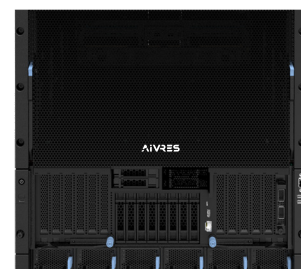
AI Training and Inference Platforms with NVIDIA HGX™ B300 / B200 NVL8

KR5288 5U Liquid-Cooled AI Server



- Over 80% server liquid cooling coverage with direct-to-chip cold-plate cooling for CPU, GPU, and NVSwitch
- Dual-input power for reliable operation
- Modular chassis and hardware design for ease of operation and maintenance

KR9288 10U Air-Cooled AI Server



- Up to 12 full-height PCIe expansions (including 10 DPUs) and up to 10 hard drives for flexible IO scalability and versatile performance
- Modular 10U chassis design provides ease of operation and maintenance
- Advanced thermal design and powerful fans for efficient air cooling

Enterprise AI and Visual Computing Platforms with NVIDIA® RTX PRO 4500 / 6000 Blackwell Server Edition

KR6278 6U Universal AI Server



Supports NVIDIA® RTX PRO 4500 Blackwell, NVIDIA RTX PRO™ 6000 Blackwell

- Based on the NVIDIA MGX modular architecture design featuring scale-out high-bandwidth interconnect enabling ultra-fast, low-latency data exchange
- Designed for cloud service providers and large enterprises deploying high performance, highly scalable AI computing clusters

KR6268 6U NVIDIA RTX PRO™ Server



Supports NVIDIA RTX PRO™ 6000 Blackwell, NVIDIA H200 NVL

- Delivers powerful computing capabilities and high-bandwidth memory
- Flexible IO and storage expansion capabilities support applications from deep learning to advanced graphics
- 4+4 redundant power supply system ensures continuous operation of critical processes