



SUPERMICRO TURNKEY AI WORKLOAD SOLUTION

Pre-configured, ready-to-deploy platform for AI & Generative AI



Supermicro Superworksation AS -531AW-TC

Table of Contents

Executive Summary	1
Solution Highlights	2
Technical Specs	3
Use Cases	3
System Hardware Config	5
Getting Started	6
For More Information.	6

Executive Summary

Adoption of AI and Generative AI is now critical across all industries, yet effectively building and managing secure AI environments remains a challenge for IT and DevOps teams. To successfully adopt AI and accelerate innovation, companies must address:

- **Infrastructure Readiness:** Ensure that existing IT environments can support AI workloads, including storage, networking, and compute resources.
- **Data & Workload Management:** Effectively managing the associated data and image workloads is crucial.
- **Operational Costs:** Factor in hardware, software, power consumption, and maintenance expenses when deploying AI solutions.

- **Orchestration layer:** An orchestration layer is needed to enable seamless communication and coordination between the infrastructure and the workload, yielding desired results.
- **Performance Optimization:** Choose hardware and software configurations that maximize AI training and inference efficiency.
- **Skill & Expertise:** Ensure that IT teams and developers have the necessary expertise to manage and optimize AI workloads effectively.

Supermicro and RAVEL have collaborated to develop an industry-leading turnkey AI workload solution that delivers a seamless, ready-to-deploy AI infrastructure designed to eliminate the complexity of AI adoption. By integrating the latest NVIDIA GPUs with RAVEL Orchestrate software, this pre-configured system enables businesses and developers to implement AI



and GenAI workloads with minimal setup and efficiently. Offering optimized performance, enterprise-grade reliability, and scalable architecture, this solution empowers organizations to accelerate innovation while reducing time to value. Whether for machine learning, generative AI, or edge AI applications, Supermicro's AI solution provides the necessary hardware, software, and tools to drive success in AI-powered initiatives.

This solution addresses the needs of several key market segments:

1. Enterprises Scaling AI Operations – Companies looking to integrate AI for automation, predictive analytics, and decision-making without the burden of infrastructure complexity.
2. AI & Machine Learning Developers – Startups and research teams needing high-performance computing for training and deploying AI models efficiently.
3. Data Centers & Cloud Providers – Organizations offering AI-as-a-Service (AlaaS) requiring scalable, high-performance AI infrastructure.
4. Healthcare & Life Sciences – Businesses utilizing AI for medical imaging, drug discovery, and diagnostics that need reliable AI processing power.
5. Manufacturing & Industry 4.0 – Companies adopting AI for quality control, predictive maintenance, and smart automation in production lines.
6. Financial Services & FinTech – Firms leveraging AI for fraud detection, algorithmic trading, and risk analysis.
7. Retail & eCommerce – Businesses enhancing customer experiences with AI-driven recommendation systems and demand forecasting.
8. Media & Entertainment – Organizations using AI for content generation, video analytics, and personalization.
9. Government & Defense – Agencies requiring AI-powered cybersecurity, surveillance, and intelligence analysis tools.

Solution Highlights

Supermicro's Turnkey AI Workload Solution simplifies AI deployment by integrating cutting-edge hardware and software into a pre-configured, ready-to-deploy system. Powered by RAVEL Orchestrate software and the latest generation of NVIDIA GPUs, this solution enables businesses and developers to seamlessly integrate AI capabilities into their operations with minimal effort.

Key Benefits

- Plug-and-Play AI Deployment: Eliminate the complexity of setting up AI infrastructure with a pre-configured solution that is ready to use out of the box.
- Optimized Performance: Leverage NVIDIA's latest GPUs to accelerate AI and GenAI workloads for superior efficiency and scalability.

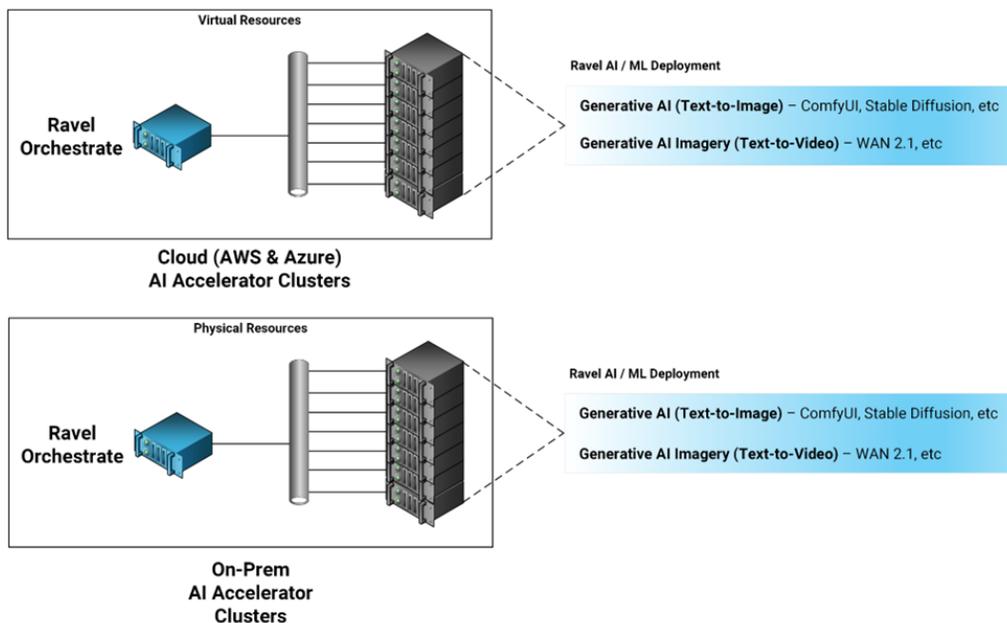
- Comprehensive Software Integration: Utilize RAVEL Orchestrate software to simplify AI workload management and orchestration.
- Enterprise-Grade Reliability: Built with Supermicro's industry-leading hardware to ensure high performance, reliability, and scalability.
- Reduced Time to Value: Quickly implement AI-driven applications without the need for extensive configuration or technical expertise.

Technical Specifications

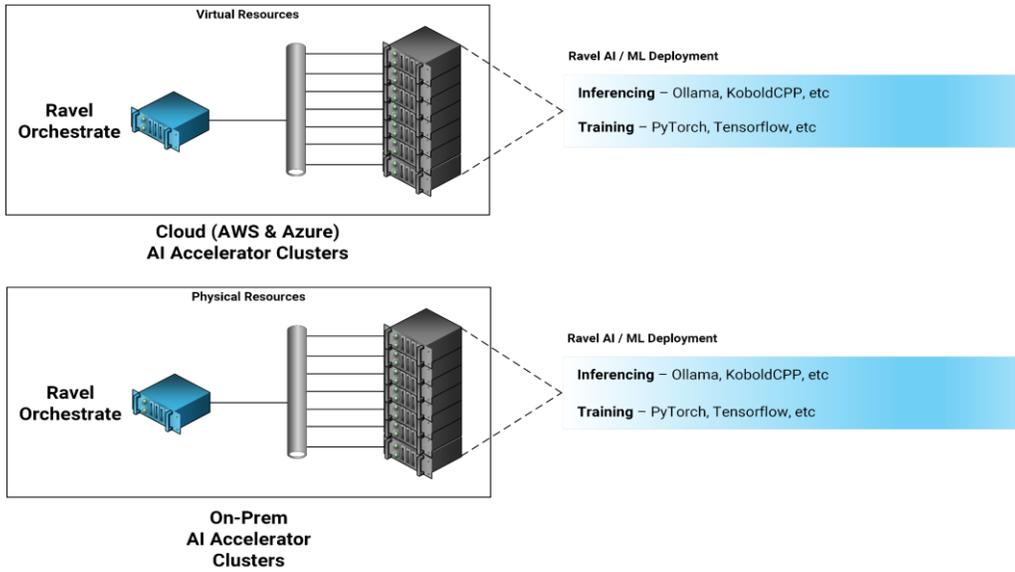
- Hardware: Latest-generation Supermicro AI servers with NVIDIA GPUs for high-speed AI computation.
- Software: RAVEL Orchestrate for streamlined AI workload management and deployment.
- Scalability: Flexible architecture that supports growing AI demands and can be customized for specific workloads.
- Security: Enterprise-grade security measures to protect sensitive AI data and workloads.

Use Cases

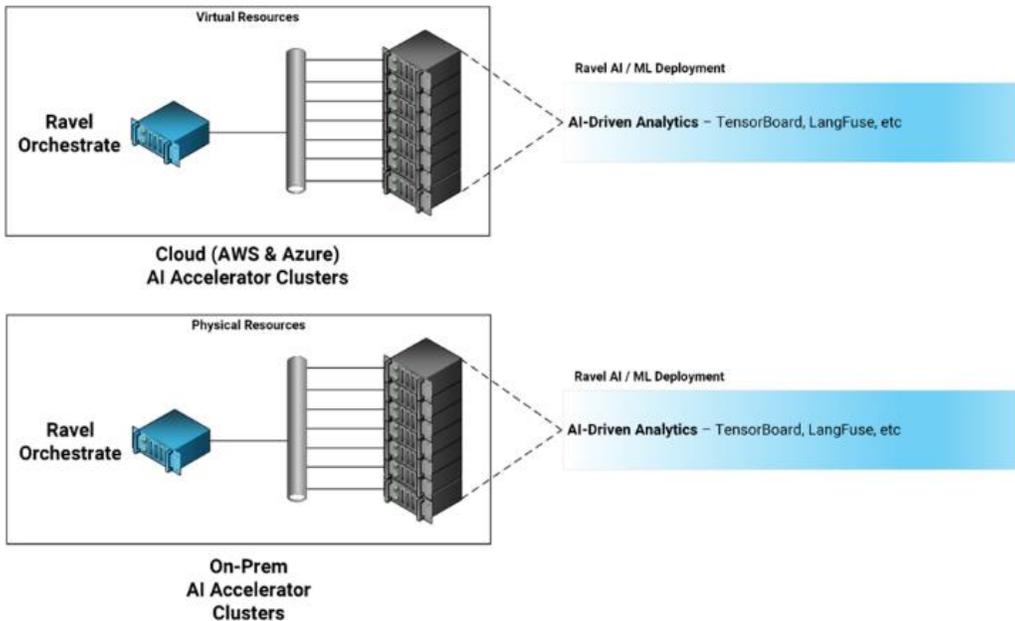
- Generative AI (GenAI): Deploy and run generative AI models efficiently for applications in content creation, image generation, and language modeling.



- Machine Learning & Deep Learning: Accelerate model training and inference for a wide range of AI applications.



- AI-Driven Analytics: Enhance data-driven decision-making with powerful AI insights.



System Hardware Configuration

Processor	AMD Ryzen™ Threadripper™ PRO 7985WX 64C/128T
GPU	NVIDIA RTX PRO™ 6000 Blackwell Workstation Edition, 96GB, 600W TDP
System Memory	256GB DDR5 ECC RDIMM (4x 64GB)
Storage	4TB PCIe NVMe M.2 PCIe 4.0 Enterprise grade drive
On-Board Devices	Network Connectivity: 2 RJ45 10GbE with Broadcom® BCM57416 1 RJ45 1GbE with ASPEED AST2600 IPMI: Support for Intelligent Platform Management Interface v.2.0 IPMI 2.0 with virtual media over LAN and KVM-over-LAN support
Input / Output	LAN: 2 RJ45 10 GBASE-T LAN ports (Broadcom® BCM57416) 1 RJ45 1 GbE Dedicated BMC LAN port (ASPEED AST2600) USB: 2 USB 3.2 Gen1 Type-A ports(front) 1 USB 3.2 Gen2 Type-C port(Rear) 1 USB 3.2 Gen2 Type-C port(Front) 4 USB 3.2 Gen2 Type-A ports(Rear) Video: 1 VGA port Serial: 1 COM port(Rear) Audio: 7.1 HD Audio TPM: 1 TPM header
OS	Microsoft Windows 11 Pro
Power Supply	2000W 80Plus Platinum
Form Factor	Mid-Tower
Dimensions	205 x 450 x 470mm (8.07" x 17.72" x 18.5")

The Supermicro AS-531AW-TC workstation with NVIDIA RTX PRO™ 6000 Blackwell Workstation Edition GPUs. Specifically, the layout is detailed below:



Supermicro AS -531AW-TC Workstation



NVIDIA RTX PRO™ 6000 Blackwell Workstation Edition

How to Get Started

Supermicro is a trusted leader in high-performance computing solutions, delivering industry-leading hardware optimized for AI workloads. By combining Supermicro's expertise with NVIDIA's cutting-edge GPU technology and RAVEL Orchestrate software, this turnkey solution provides businesses with a reliable, efficient, and scalable AI infrastructure that accelerates innovation.

Discover how Supermicro's turnkey AI workload solution can revolutionize your AI initiatives. Contact us today to learn more or request a demo.

For Integrators & Resellers Interested in partnering to host your own AI Workload POC, please get in touch with us at <https://ravelinc.com/partner/#>.

Further Information

<https://www.supermicro.com/en/>

<https://ravelinc.com/>

SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

Learn more at www.supermicro.com

RAVEL

RAVEL is a leading provider of DevOps tools for the orchestration of cloud, on-prem, and data center infrastructure for complex workloads and workflows. Our flagship product RAVEL Orchestrate™ provides teams with a no-code, DevOps & IT solution that automates & rapidly assembles complex software and hardware environments, including managing infrastructure, virtualization technology, identity management, and customized software images. From digital content creation studios to architecture and engineering environments, to large enterprises in the energy, computing, and healthcare sectors, our solutions help our customers elevate their productivity and improve their workflow.

Learn more at: www.ravelinc.com