



NVIDIA



SUPERMICRO, AMD, WEKA, AND NVIDIA DELIVER HIGH PERFORMANCE ANSYS TURNKEY SOLUTION

Tested Solution Demonstrates That Supermicro Servers Powered by AMD, Excel with WEKA Storage and NVIDIA Networking to Reduce Time To Completion of Complex Mechanical Design Challenges



Supermicro AS-1125HS-TNR

Executive Summary

Fast time to market product development through simulation is critical to ensure stability and high quality for the manufacturing industry. Supermicro has teamed up with Ansys, AMD, NVIDIA, and WEKA to deliver the state-of-the-art pre-integrated rack solution for such purposes. The Supermicro Rack solution provides pre-tested and integrated computing, networking, and storage solutions with job schedulers so customers can quickly move on to production in days.

TABLE OF CONTENTS

Executive Summary	1
How Do AMD CPUs Improve Ansys Performance?.....	2
Supermicro Ansys Compute Node Description	3
Ansys Fluent and Ansys Mechanical Software	4
Supermicro SuperCloud Composer	5
Configuring the Optimal Server for your Workload	5
NVIDIA InfiniBand & Unified Fabric Manager (UFM) Application	9
Supermicro Rack Scale Integrated Solutions	10
Summary	11
More Information	11



How Do AMD EPYC™ CPUs Improve Ansys Performance?

Compared to the prior generation, the new 4th Gen AMD EPYC™ processors achieve better performance for Ansys applications with up to 50% more cores, higher frequencies, support for AVX-512 instructions, more memory bandwidth, and faster PCIe® and Infinity Fabric™ data transfer rate. In addition, optimizing Ansys applications with AMD compilers and libraries can help enhance performance even further.

Supermicro chose the AMD EPYC™ 9554 Processor for dual socket AMD based servers for Ansys Fluent and Ansys Mechanical workloads. The AMD EPYC CPU has a high 64-core count per socket and a high clock rate of 3.1GHz. The Ansys software is licensed based on CPU core counts. Customers can enjoy the same high-performance core performance with more cores per server while using less data center footprint. That can result in significant energy savings as well.

AMD EPYC™ 9554 Processor	
# of CPU Cores	64
Base Clock	3.1GHz
All Core Boost Speed	3.75GHz
L3 Cache	256MB
Default TDP	360W

Table: Tech spec of AMD Processor

The Ansys workload can be distributed to many servers to accelerate the workload, the network performance between the servers within the cluster plays a crucial role in delivering the overall best performance. NVIDIA QM8700 200Gbps InfiniBand Switch is used with a non-blocking network design to provide state-of-the-art high performance and low latency network performance when distributing the Ansys workload with job schedulers, such as Slurm. The shared WEKAFS File System can also take advantage of the low latency, high-performance InfiniBand network to ensure the storage I/O can deliver the optimal performance expected by the Ansys workload.

The solution consists of Supermicro Hyper A+ Server AS -1125HS-TNR 1U 2 socket AMD server for Ansys Compute Node, Supermicro Hyper A+ Server AS -2025HS-TNR 2U 2socket AMD server for Ansys management servers, Supermicro WIO A+ Server AS -1114S-WN10RT 1U 1socket AMD server for WEKA Home server, WEKA Storage server, NVIDIA Unified Fabric Manager and Supermicro SuperCloud Composer.

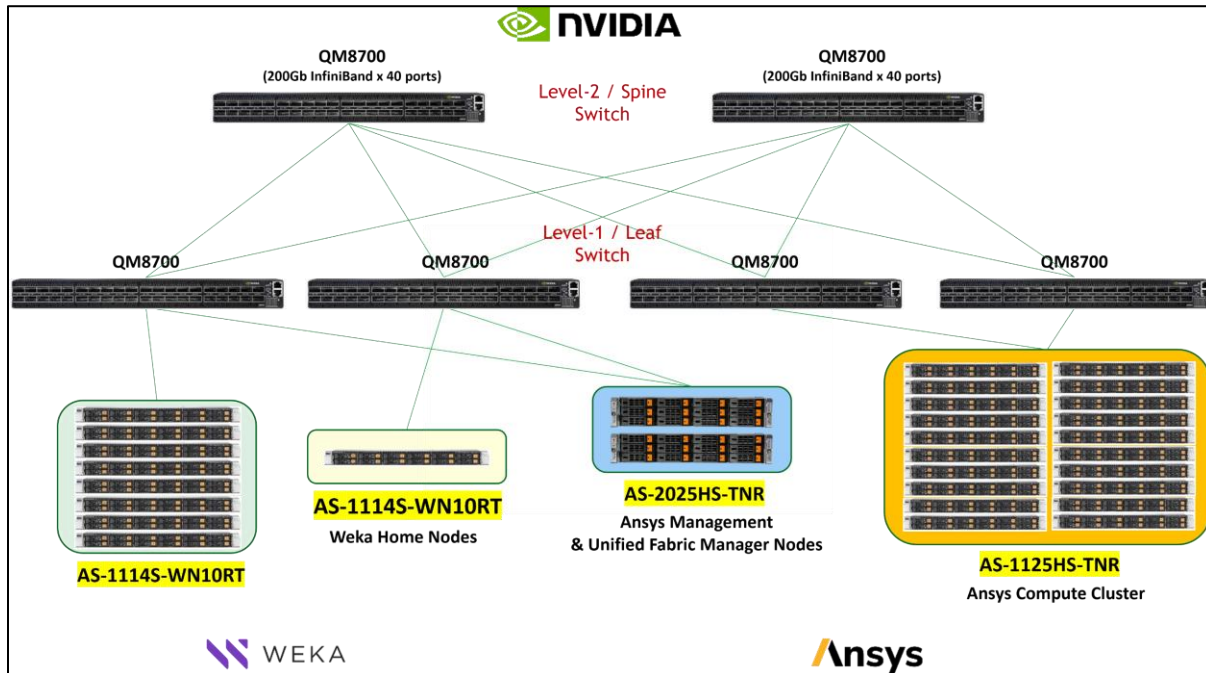


Figure 1 - Overall Architecture of Supermicro and Weka Storage Solutions

Operating System	Ubuntu 22.04
Ansys Fluent / Mechanical	2024 R1
Job Scheduler	Slurm
MPI	Intel MPI 2021.10
MLNX OFED	5.8-1.1.2.1:
WEKA	v4.2.4 with Rocky Linux 8.6

Table: Software Version

BIOS Options	Settings
NPS (NUMA per socket)	4
SMT (Symmetric Multithreading)	OFF

Table: BIOS Optimized Settings for Ansys

NPS (NUMA Per Socket) controls the number of NUMA (Non-uniform Memory Access) nodes, which enables fine-tuning of the transfer speed between specific CPU cores and the closest (fastest) available memory channels. SMT (Symmetric Multithreading) determines whether each physical core can appear to the operating system as two “logical” cores, which can boost performance significantly in applications that perform many smaller or random tasks. However, the tested Ansys applications perform best when each physical core is fully dedicated to the compute-intensive and memory bandwidth

intensive simulation workload. Therefore, Supermicro recommends using NPS=4 and SMT=OFF when running Ansys Fluent and Ansys Mechanical on systems with AMD EPYC 9004 Processors.

Supermicro servers with 4th Gen AMD EPYC processors and 12 memory channels per processor, which also support AVX-512 instructions, can deliver high throughput per node for computational applications such as Ansys Fluent and Ansys Mechanical since they benefit from multicore parallelism and greater memory bandwidth.

Supermicro Ansys Compute Node

AS -1125HS-TNR
AMD EPYC 9554 64C 3.1GHz CPU x2
64GB 4800MHz x24 (1.5TB)
Micron NVMe M.2 800GB NVMe SSD x2
3.2TB NVMe 3DWPDP SSD x2
NVIDIA 200Gb CX6 1P QSFP x2
Broadcom 25Gb 2P SFP28 AIOM x1



Table: Supermicro 1U 2 socket Server with the AMD EPYC 9004 Processor for Ansys Computing Node

AS -2025HS-TNR
AMD EPYC 9554 64C 3.1GHz CPU x2
96GB 4800MHz x24 (2.3TB)
Micron NVMe M.2 800GB NVMe SSD x2
3.2TB NVMe 3DWPDP SSD x2
Broadcom 25Gb 2P SFP28 AIOM x1
NVIDIA 200Gb CX6 1P QSFP x2



Table: Supermicro 2U 2 socket Server with the AMD EPYC 9004 Processor for Ansys Management Nodes

Ansys Fluent and Ansys Mechanical Software

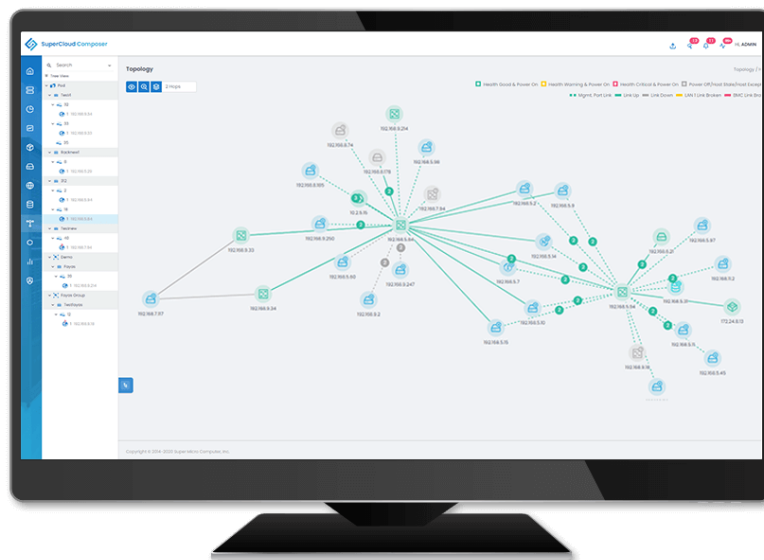
Ansys develops the most widely used multiphysics engineering simulation software solutions for product design, testing, and operation. With the Supermicro systems, designers and engineers can run complex simulations on the compute clusters in the early stage of the design process to test and validate design concepts before production and reduce the cost and risk of quality failures. Ansys Fluent is a general-purpose computational fluid dynamics (CFD) application that models fluid flow, heat, mass transfer, chemical reactions, and more.

Ansys Mechanical enables you to solve complex structural engineering problems and make better, faster design decisions. With the finite element analysis (FEA) solvers available in the suite, you can customize and automate solutions for your structural mechanics problems and parameterize them to analyze multiple design scenarios.

Supermicro SuperCloud Composer - Cloud Management Software & Gateway to Composable Disaggregated Infrastructure

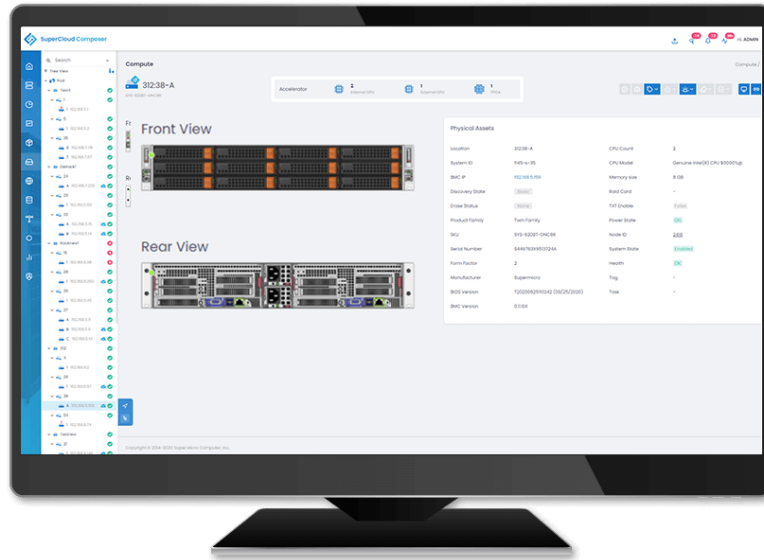
Integrating Supermicros' SuperCloud Composer with WEKA enables unprecedented deployment possibilities, where WEKA cluster members can be deployed across multiple systems simultaneously. Data center operators achieve a clear vision of their WEKA performance metrics through SuperCloud Composer's customizable dashboard. Administrators can also enforce configurable Event Management Systems with the flexibility to create rule-based alerts for their WEKA cluster.

Respond and adapt quickly to dynamic business needs with storage, compute, and networking flexibility that accommodates ever-changing workload requirements. Leveraging composable disaggregated infrastructure technology, SuperCloud Composer makes your data center software-defined with the ability to meet and support rapidly changing workload requirements.



Increased Data Center Operations Efficiency

Reduce manual intervention by automating device management to deploy, provision, monitor, and repurpose infrastructure assets through a single pane of glass. Consolidate servers and racks seamlessly into groups of IT assets that share common workloads and can be managed effortlessly simultaneously.

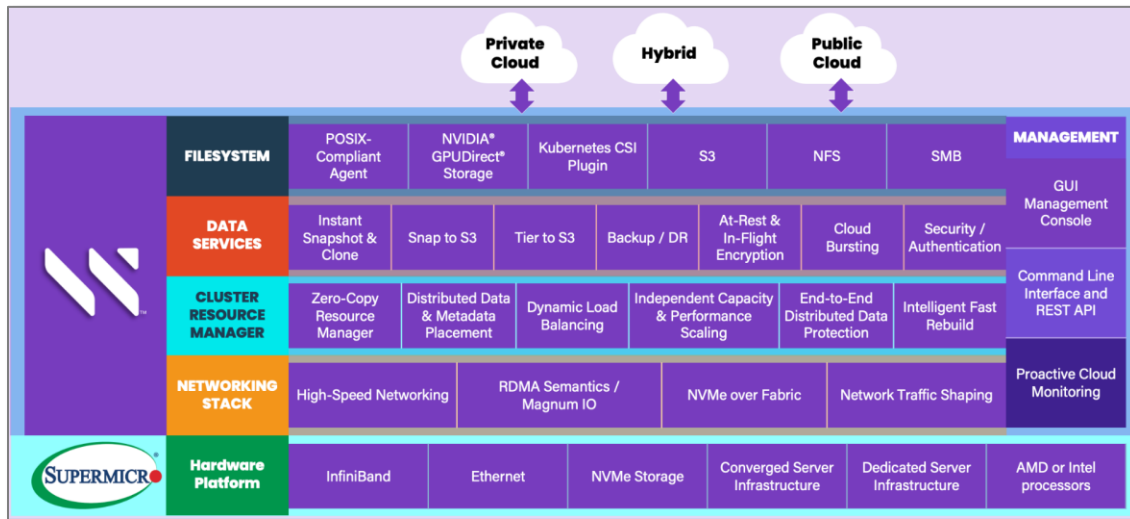


Data-driven Control and Compliance

Manage customer's entire infrastructure, including 3rd party devices, at a glance with state-of-the-art dashboards that provide a comprehensive view of the customer data center. SuperCloud Composer's data-rich analytics provide customers with predictive analytics and complete insight into the performance of customer's data centers with live statistics, historical information, and IT telemetry.



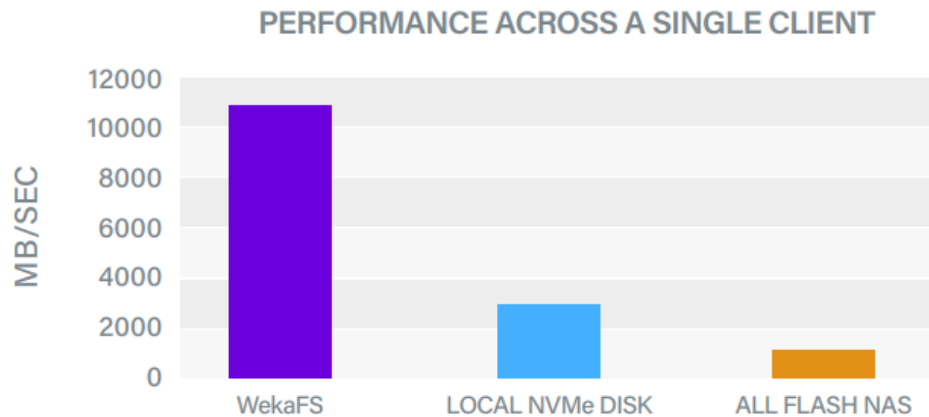
Integrated Dashboard for WEKA storage monitoring



The System Architecture of Supermicro WEKA Storage

WEKA has built a software-defined, high-performance, file-based storage solution that is highly scalable and easy to deploy, configure, manage, and expand. The design philosophy behind the WEKA file system (WEKAFS) was to create a single storage architecture that runs on-premises with the performance of all-flash arrays, the simplicity and feature set of Network-Attached Storage (NAS), and the scalability and economics of the cloud.

- The world's fastest shared file system was validated on SPEC SFS 2014, SPEC Storage 2020, IO-500, and STAC benchmarks.
- Flexible application storage access, including POSIX, NFS, SMB, S3, and NVIDIA GPUDirect storage.
- Zero performance tuning to support small and large files simultaneously, with both mixed random and sequential I/O patterns.
- Application-level 4K I/O, consistent sub-250 microsecond latency on high-speed networks, unlimited random IOPs performance—linear scaling in line with the size of the cluster
- Automated built-in tiering to expand the namespace from fast flash to object storage either on-premises or in the cloud.
- Distributed resilience that eliminates the bottlenecks of traditional built-in data protection
- Filesystem cloning for rapid development and testing workflows.
- Non-disruptive upgrades for storage services
- Full GUI, CLI, and API Management



The figure above outlines the actual performance of WEKAFS in a production environment with 100Gbit networking versus All-flash NAS and direct-attached local storage. WEKAFS was 3x faster than local storage and 10x faster than NFS-based all-flash NAS. Doubling the network bandwidth to 200Gbit would result in 2x the client throughput, delivering 6x improvement in the local file system and 20x the performance of NFS. The data is provided based on the [WEKA Architecture White Paper](#).

The following shows the sample configuration of the WEKAFS storage node base Supermicro AS-1114S-WN10RT A+ Servers. Depending on capacity/performance requirements, the number and capacity of the NVMe SSD drives can be sized based on the requirements. The chart below shows a typical 8-node configuration with over 200GB/s throughput (136 GB/s read throughput performance and 65 GB/s write throughput) performance with a 200G non-blocking InfiniBand network, showing excellent performance based on these specific hardware configuration combinations.

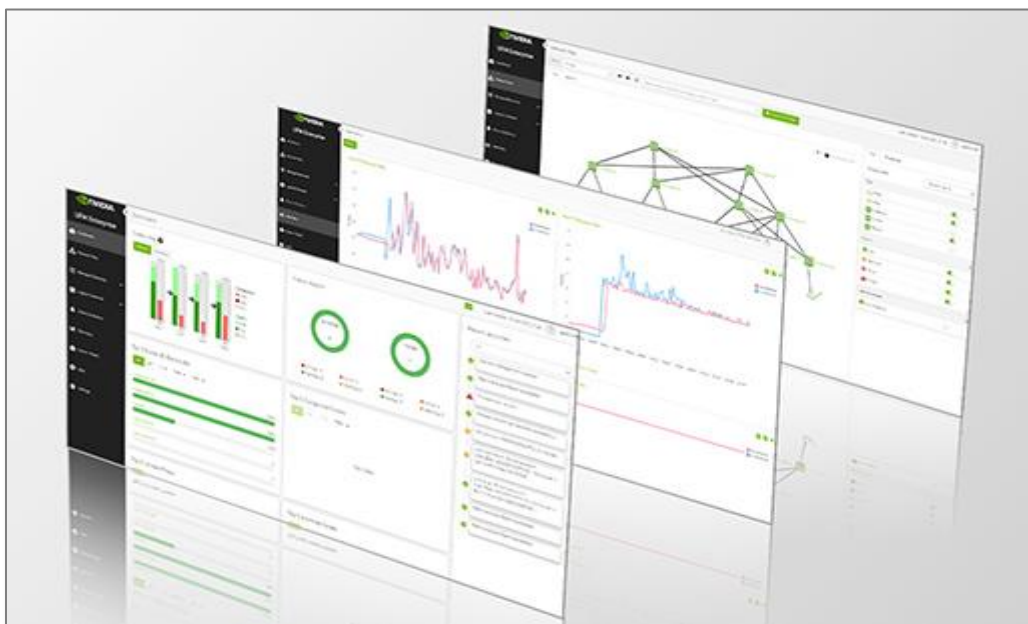
AS -1114S-WN10RT
AMD EPYC 7543P 32C 2.8GHz x1
32GB 3200MHz x8
960GB M.2 NVMe SSD x2
3.2TB NVMe SSD x4 (support x10)
NVIDIA 200Gb IB x2ports
25Gb SFP28 x2ports



Table: Supermicro 1U 1 socket Server for High-Performance WEKA Storage

NVIDIA InfiniBand & Unified Fabric Manager (UFM)

NVIDIA® UFM® platforms revolutionize InfiniBand network management. By combining enhanced and real-time network telemetry with AI-powered cyber intelligence and analytics, the UFM platforms empower customers to discover operation anomalies and predict network failures for preventive maintenance. UFM platforms comprise multiple levels of solutions and capabilities to suit customer data centers' needs and requirements. At the basic level, the UFM Telemetry platform provides network validation tools and monitors the network performance and conditions. It captures, for example, rich real-time network telemetry information, workload usage data, and system configuration and streams it to a defined on-premises or cloud-based database for further analysis.



The NVIDIA UFM platform features robust graphical user interfaces (GUI)

The mid-tier UFM Enterprise platform enhances network monitoring, management, workload optimizations, and periodic configuration checks. In addition to including all of the UFM Telemetry services, it provides network setup, connectivity validation, secure cable management, automated network discovery and network provisioning, traffic monitoring, and congestion discovery. UFM Enterprise also enables job scheduler provisioning and integration with Slurm and Platform LSF, as well as network provisioning and integration with OpenStack, Azure Cloud, and VMware.

Supermicro Rack Integrated Solutions

Supermicro's L11 rack scale integration process not only includes the design, assembly, configuration, and testing of each server at the node and rack level but also involves simultaneous testing of all nodes to ensure that they have the exact configuration and are communicating with one another as they were designed to.

The manufacturing process at L12 involves networking from the rack to the multi-rack level, then loading the customer's application software (Ansys Fluent and Mechanical, for this case) for complete, real-world testing at rack scale. Supermicro can even provide benchmarking data to ensure that rack-scale performance meets requirements.

Zero Failure Goal - Failures are costly, impacting revenue and reputation. A rack-scale solution validated and proven before shipping to the customer reduces the likelihood of issues post-deployment.

Rack Optimization - Adopting a standardized design and modular approach optimized for power and cooling efficiency.

Cable Management - Power, networking, and liquid cooling cabling are organized and managed to reduce clutter, improve airflow, and simplify system troubleshooting and maintenance.

Burn-In Testing - Best-in-class quality assurance for servers and racks with automated L11 cluster testing to validate integration quality and hardware reliability.

Delivery - Racks are professionally packed in shock-resistant and reusable crates for rapid shipping worldwide; they are cabled, labeled, thoroughly tested, and ready for operation upon arrival at your data center.

Full Plug and Play - Supermicro's L12 solutions are ready to deploy on day one; simply connect power and networking to start running your applications immediately.

Knowledge Transfer Best practice documentation is critical for long-term success and collaboration across multiple teams from design to delivery of your L12 Solution.

Summary

Without simulation, there are no autonomous vehicles. No 5G networks. No space exploration. Ansys Multiphysics software solutions and digital mission engineering help companies innovate and validate like never before. The Supermicro Solution Architect team can architect, fully test, and validate the Ansys solution design at the cluster level so customers can have the Ansys CAE solution operational in days. As a result, engineers, designers, artists, and scientists can focus on fewer system implementation complexities and more technological breakthroughs.

Deploy Supermicro Rack Scale Solutions of:

- Multi-rack plug-and-play design. Easily grow the cluster as an organization's workloads increase.

- AS -1125HS-TNR and AS -2025HS-TNR servers provide the best-in-class performance and flexible computing architecture with dual 4th Gen AMD EYPC processors for Ansys simulation workload.
- High-performance NVIDIA InfiniBand networking enables high bandwidth and low latency for workloads simultaneously utilizing multiple systems. NVIDIA Unified Fabric Manager easily manages it with graphical user interfaces.
- High-performance and highly scalable WEKA File System quickly provides hundreds of GB/s throughput performance in a small cluster with very rich storage access protocols, including POSIX, NFS, SMB, S3, GPUDirect, Kubernetes CSI Plugin, and flexible software defined storage resiliency features.

For More Information:

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

<https://www.nvidia.com/en-us/networking/infiniband/ufm>

<https://www.supermicro.com/products/system/hyper/1u/as-1125hs-tnr>

<https://www.supermicro.com/products/system/mainstream/2u/as-2024s-tr>

<https://www.supermicro.com/en/solutions/management-software/supercloud-composer>

<https://www.supermicro.com/manuals/brochure/Brochure-Rack-Services.pdf>

<https://www.weka.io/wp-content/uploads/resources/2023/03/weka-architecture-white-paper.pdf>

<https://www.amd.com/en/products/processors/server/epyc/4th-generation-9004-and-8004-series.html>

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.