

Lenovo ThinkSystem SR675 V3

From Exascale to Everscale™ – a modular platform tailored to your enterprise AI requirements

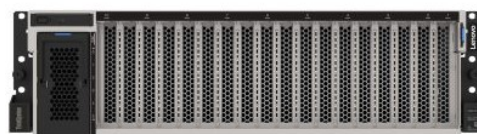
Designed for Your Enterprise

Lenovo ThinkSystem SR675 V3 delivers optimal performance for Artificial Intelligence (AI), High Performance Computing (HPC) and graphical workloads across an array of industries.

Retail, manufacturing, financial services and healthcare industries are leveraging GPUs to extract greater insights and drive innovation utilizing machine learning (ML) and deep learning (DL). Here are a few ways accelerated computing leverages GPUs in different organizations:

- Computer vision for retail customer experience
- Natural language processing (NLP) for call centers
- Large-scale Omniverse™ digital twins
- In-silico trials and immunology in Life Sciences
- Ray-traced rendering for photo-realistic graphics
- Remote visualization for work-from-home teams
- Powerful video encoding and decoding
- Automatic optical inspection (AOI) for quality control

As more workloads leverage the capabilities of accelerators, the demand for GPUs increases. The ThinkSystem SR675 V3 delivers an optimized enterprise-grade solution for deploying accelerated HPC and AI workloads in production, maximizing system performance.



EveryScale Platform means: Versatility

The SR675 V3 features a modular design for ultimate flexibility. With multiple different front compartment options, configurations include:

- One or two 4th Generation AMD EPYC™ Processors
- Up to eight double-width GPUs with NVLink Bridge
- NVIDIA HGX™ H100 4-GPU with NVLink and Lenovo Neptune™ hybrid liquid cooling performance
- AMD Instinct™ MI Series Accelerators
- Choice of front or rear high-speed networking
- Choice of local high speed 2.5 SAS/SATA/NVMe storage

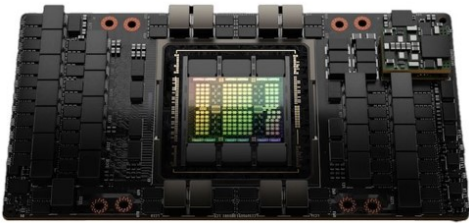
Lenovo

The ThinkSystem SR675 V3 is built on one or two 4th Generation AMD EPYC™ Processors and is designed to support the vast NVIDIA Hopper, Lovelace and Ampere datacenter portfolio and AMD Instinct™ MI Series Accelerators.

The ThinkSystem SR675 V3 delivers performance optimized for your workload, be it visualization, rendering or computationally intensive HPC and AI.

Most Powerful Compute Platform

The NVIDIA H100 Tensor Core GPU delivers unprecedented acceleration—at every scale—to power the world's highest performing elastic data centers for AI, data analytics, and HPC applications. The H100 can efficiently scale up or be partitioned into seven isolated GPU instances, with Second-Generation Multi-Instance GPU (MIG) providing a unified platform that enables elastic data centers to dynamically adjust to shifting workload demands. A rack of 13 ThinkSystem SR675 V3 servers can generate up to 5 PFLOPS of compute power.



Cutting Edge Cooling Capability

Traditional air-cooling methods are reaching critical limits. Increases in component power especially on CPUs and GPUs have resulted in higher energy and infrastructure costs, extremely loud systems and heightened carbon footprints.

To combat these challenges and dissipate heat quickly, some models of the SR675 V3 employ Lenovo Neptune™ liquid-to-air hybrid cooling technology.

The heat of the NVIDIA HGX™ H100 GPUs is removed through a unique closed loop liquid-to-air heat exchanger that delivers the benefits of liquid cooling such as lower power consumption, quiet operation and higher performance without adding plumbing.

Solutions That Scale

Whether you're just starting with AI or moving into production, your solution must scale with your organization's needs. The ThinkSystem SR675 V3 can be used in a cluster environment using high-speed fabric to scale out as your workload demands increase.

Enabled with Lenovo HPC & AI Software Stack, you can support multiple users and scale within a single cluster environment. Lenovo HPC & AI Software Stack provides our HPC customers with a fully tested and supported open-source software stack to enable their administrators and users for the most effective and environmentally sustainable consumption of Lenovo Supercomputing capabilities.

Our Confluent management system and Lenovo Intelligent Computing Orchestration (LiCO) Web portal provide an interface designed to abstract the users from the complexity of HPC cluster orchestration and AI workloads management, making open-source HPC software consumable for every customer.

LiCO Web portal provides workflows for both AI and HPC, and supports multiple AI frameworks, allowing you to leverage a single cluster for diverse workload requirements.

Get a Hands-on, Concierge Experience

Getting started is easy. At the Lenovo AI Innovation Centers, you can test your own Proof of Concept (PoC) on different hardware and software platforms, including the ThinkSystem SR675 V3. Lenovo data scientists and AI solution architects are available to help you along the way. Lenovo can work with you to develop an end-to-end solution with professional services and deep industry partnerships to ensure your success.

Specifications

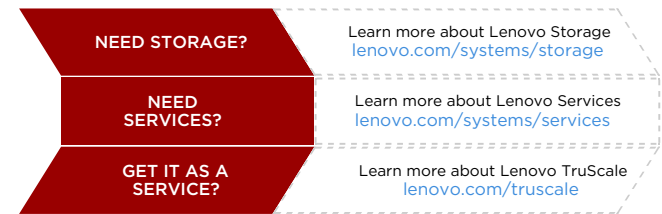
Form Factor	3U rack
Processor	1x or 2x 4th Generation AMD EPYC™ Processors per node
Memory	Up to 3TB using 24x DDR5 DIMMs with max frequency 4800 MHz 12 channels per CPU with 1DPC Capacities: Up to 128GB
Base Module	Up to 4x double-wide, full-height, full-length; FHFL GPUs; PCIe Gen5 x16 Up to 8x 2.5" Hot Swap SAS/SATA/NVMe
Dense Module	Up to 8x double-wide, full-height, full-length GPUs each PCIe Gen5 x16 on PCIe switch Up to 6x EDSFF E1.S NVMe SSDs or up to 4x EDSFF E3.S 1T NVMe HS SSDs
HGX Module	NVIDIA HGX™ H100 4-GPU with 4x NVLink connected SXM5 GPUs Up to 4x 2.5" Hot Swap NVMe SSDs or up to 4x EDSFF E3.S 1T NVMe HS SSDs
RAID Support	Software RAID is not supported. Only RAID Controllers and HBA
I/O Expansion	Up to 6x PCIe Gen5 x16 adapters (2 front, 4 rear) and 1x OCP NIC 3.0 (x16/x8/x4) (rear) depending on the configuration
Power and Cooling	Four N+N redundant hot-swap PSUs (up to 2600W Titanium) Full ASHRAE A2 support with internal fans and Lenovo Neptune™ liquid-to-air hybrid cooling on NVIDIA HGX™ H100
Management	Lenovo XClarity Controller 2 (XCC2), Confluent and Lenovo HPC & AI Software Stack
OS Support	Red Hat Enterprise Linux, SUSE Linux Enterprise Server, Microsoft Windows Server, VMware ESXi, Alma Linux, Rocky Linux Tested on Canonical Ubuntu

About Lenovo

Lenovo (HKSE: 992) (ADR: LNVGY) is a US\$62 billion revenue global technology powerhouse, ranked #171 in the Fortune Global 500, employing 77,000 people around the world, and serving millions of customers every day in 180 markets. Focused on a bold vision to deliver smarter technology for all, Lenovo is expanding into new growth areas of infrastructure, mobile, solutions and services. This transformation is building a more inclusive, trustworthy, and sustainable digital society for everyone, everywhere.

For More Information

To learn more about the Lenovo ThinkSystem SR675 V3, contact your Lenovo representative or Business Partner or visit lenovo.com/thinksystem. For detailed specifications consult the [SR675 V3 Product Guide](#).



© 2024 Lenovo. All rights reserved.

Availability: Offers, prices, specifications and availability may change without notice. Lenovo is not responsible for photographic or typographic errors. **Warranty:** For a copy of applicable warranties, write to: Lenovo Warranty Information, 1009 Think Place, Morrisville, NC, 27560. Lenovo makes no representation or warranty regarding third-party products or services. **Trademarks:** Lenovo, the Lenovo logo, From Exascale to Everscale, Lenovo Neptune®, ThinkSystem®, and XClarity® are trademarks or registered trademarks of Lenovo. Linux® is the trademark of Linus Torvalds in the U.S. and other countries. Microsoft®, Windows Server®, and Windows® are trademarks of Microsoft Corporation in the United States, other countries, or both. Other company, product, or service names may be trademarks or service marks of others. Document number DS0151, published November 10, 2022. For the latest version, go to lenovopress.lenovo.com/ds0151.