Empowering CSPs to deliver Al and HPC Innovation



Growing demand for AI and HPC workloads is creating new opportunities for differentiation and growth for cloud service providers (CSPs). Success hinges on providing best-in-class delivery while maintaining cost efficiency and operational flexibility.

To capitalize on this market opportunity, CSPs require infrastructure that is high performing, scalable, reliable, and secure. By adopting solutions optimized for AI acceleration and HPC frameworks, highly specialized CSPs can minimize performance bottlenecks, manage energy costs, and streamline deployment to stay ahead of the competition in AI innovation.

Lenovo and AMD deliver the right set of solutions to help CSPs drive maximum business value from the most demanding and complex Enterprise AI and HPC workloads.



Smarter technology for all





Al as a Service, at scale

Lenovo ThinkSystem® servers, powered by AMD EPYC™ processors and AMD Instinct™ GPUs, provide the power and reliability CSPs need.

Ideal for training, fine tuning and inferencing large AI models and HPC workloads, AMD Instinct GPUs are powered by AMD Compute DNA (CDNA) architecture, offering performance with high memory capacity and bandwidth, scalability, and energy efficiency. AMD Instinct MI300 Series GPUs are designed to hold today's large AI models in fewer GPUs, allowing for lower overall cost.

Lenovo ThinkSystem servers leverage AMD Infinity Fabric™ which provides interconnects for seamless workload scaling in cloud environments. These systems are engineered with advanced thermal management that enables you to remain flexible for evolving workloads.

Performance meets security

Lenovo ThinkSystem servers powered by AMD Instinct GPUs are built on a foundation of security, reliability, and performance. A trusted global supply chain and manufacturing processes help enable CSPs to deploy AI and HPC workloads with confidence.



Key challenges for CSPs in Al and HPC

- Scaling service offerings to accommodate growth
- Optimizing workload acceleration for performance and efficiency
- Ensuring data integrity and security
- Simplifying infrastructure management and deployment
- Delivering complete ecosystem with support for industry AI and HPC models and frameworks

Lenovo has been recognized for security and reliability



#1

x86 server reliability for 10 consecutive years¹



#1

x86 security for 5 consecutive years²

> Smarter technology for all

Built to optimize resource utilization

Lenovo and AMD enable cloud service providers with solutions that help accelerate their clients' AI workloads while optimizing resource utilization, helping CSPs meet growing demand without compromising performance or energy efficiency.

Cirrascale case study

Creating specialized IT services for AI innovation

See how Cirrascale built a brand-new GPU-as-a-Service cluster using Lenovo ThinkSystem servers with AMD EPYC™ processors and AMD Instinct™ GPUs — cutting time to market.

Read more

Open-source flexibility

AMD ROCm™ open software is part of a robust opensource ecosystem, delivering drivers, development tools, and APIs designed for AMD Instinct GPUs. Lenovo and AMD support industry-standard AI and machine learning frameworks and models such as Bloom, Falcon, Llama, Mistral, Mixtral, PyTorch, and TensorFlow, empowering CSPs to integrate, manage, and scale AI workloads with flexibility and seamless interoperability across diverse cloud environments.

Management simplicity

Lenovo XClarity® System Management enables streamlined operations with centralized, automated infrastructure management. Simplify cloud operations with an integrated management console that enables you to maximize efficiency and availability and flexibly scale up customers. Lenovo XClarity System Management seamlessly integrates into a wide range of IT applications.

Fast and efficient

The high-performance capabilities of Lenovo ThinkSystem servers and AMD Instinct GPUs enable CSPs to help customers achieve faster insights. Greater efficiency features enable the ability to consolidate workloads on fewer servers, helping to lower overall operating costs for CSPs. And because AMD ROCm software is an open source software ecosystem, CSPs avoid software licensing fees, potentially further reducing operational costs for more efficient operations.



The powerhouse server CSPs need

The **Lenovo ThinkSystem SR685a V3 server**, supporting up to 8 AMD Instinct™ GPUs, is built for AI training, high-performance inferencing, complex data analysis, large-scale simulations, and model development while helping maximize computational efficiency.

Engineered to support high-performance workloads and optimized for industry-standard AI models and HPC frameworks, the ThinkSystem SR685a V3 simplifies scaling, improves resource efficiency, and accelerates time-to-insight for AI as a Service and other cloud-based applications. A 2-socket, 8U system, this server fits into industry standard 19-inch racks and features an advanced air-cooled design.



ThinkSystem SR685a V3

A team you can trust

Enterprises around the world rely on Lenovo for reliability, security, and performance. Lenovo ThinkSystem servers powered by AMD EPYC™ processors and AMD Instinct™ GPUs provide CSPs the performance and optimized efficiency they need to help support AI and HPC applications in a highly competitive marketplace.

Learn more, or contact your Lenovo sales representative.

Sources

- Information Technology Intelligence Consulting, "ITIC 2024 Global Server Hardware, Server OS Reliability Report," November 2024
 Information Technology Intelligence Consulting, "ITIC 2023 Global Server Hardware,
- 2 Information Technology Intelligence Consulting, "ITIC 2023 Global Server Hardware. Server OS Reliability Report," September 2023

© 2025 Lenovo. All rights reserved. Lenovo and the Lenovo logo are trademarks of Lenovo. AMD, the AMD Arrow logo, EPYC, AMD Instinct, ROCm, AMD Compute DNA (CDNA). AMD Infinity Fabric, and combinations thereof are trademarks of Advanced Micro Devices, Inc. All other trademarks are the property of their respective owners. v1.00 April 2025.

Smarter technology for all

