Lenovo Unleashes ThinkSystem Server Portfolio with 4th Gen AMD EPYC Processors on HPC and AI Workloads

Article

In November 2022, Lenovo announced multiple ThinkSystem V3 servers using 4th Gen AMD™ EPYC processors that are ideal for High-Performance Computing (HPC) and Artificial Intelligence (AI) applications. This announcement was a big leap for HPC and AI, and we are especially excited at the impact on these data and compute-intensive workloads.

The new AMD EPYC processors provide more cores, higher frequencies, and significantly more bandwidth than their predecessors. Further, the 4th Gen AMD EPYC processors include up to 96 cores, clock speeds as high as 4.4GHz, and support for the DDR5, PCIe 5.0, and Compute Express Link (CXL) standards. These performant CPUs deliver outstanding raw horsepower to traditional codes, support today's rapidly growing data and analytics requirements, and help reduce bottlenecks for accelerated workloads.

Let's take a closer look at the AMD EPYC processors powering many of the new ThinkSystem V3 servers.
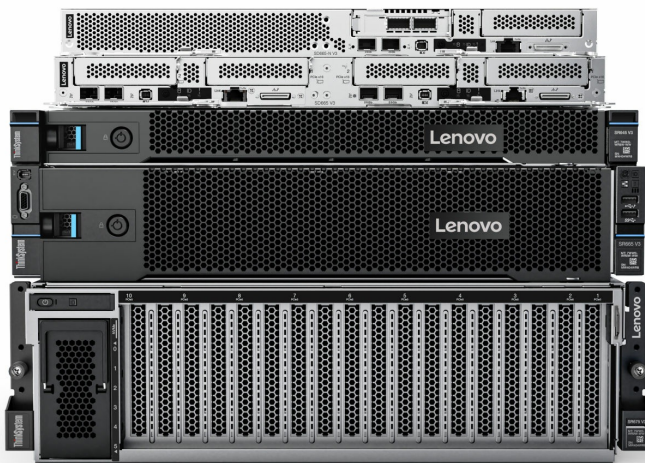

Figure 1. The family of ThinkSystem V3 servers with AMD processors

Monster Performance for traditional HPC codes

Today, most HPC applications are CPU-based. The processing of these compute-intensive codes benefits greatly from more cores and higher frequencies. With up to 96 cores running at up to 4.4GHz, the new 4th Gen AMD EPYC processors take the "Cores per CPU" crown. These processors also offer organizations running HPC workloads significant performance increases with up to 14 percent improvement in instructions per clock over the prior generation (footnote 1).

Improved performance for Analytics, AI, and Accelerated Workloads

While CPU cores and frequencies are important, they don't tell the whole story for many of today's accelerated HPC, AI, and analytics workloads. These data-hungry use cases require bandwidth to move huge volumes of data to the CPU for processing or feed GPUs and other accelerators. Utilizing PCIe Gen 5.0, the ThinkSystem servers with 4th Gen AMD EPYC processors deliver up to 2X the bandwidth of their PCIe Gen 4.0 predecessors (footnote 2). The extreme I/O performance of these new systems powered by AMD processors is ideal for running Big Data and analytics workloads, such as ETL (Extraction Transformation and Loading), and the PCIe Gen 5.0 bus supports the massive data movement required to support those operations.

In many cases, expensive GPUs and other accelerators have low utilization rates, as they're largely idle, waiting for data to process. PCIe Gen 5.0 increases this utilization by delivering twice the bandwidth, improving overall performance, and allowing organizations to generate a greater return from their accelerator investments. This increased bandwidth is especially important for accelerated workloads increasingly found in HPC and typically found in AI.

Unleashing next-Gen AMD processors with Lenovo Neptune

Lenovo ThinkSystem V3 servers allow organizations to get outstanding performance, density, and energy efficiency from AMD EPYC processors. By collaborating with AMD engineering years before the release of the latest generation of processors, Lenovo has integrated Neptune cooling technologies in the following performance-optimized servers:

- **ThinkSystem SR645 V3** is a powerful 2-socket 1U server. Ideal for Cloud Service Providers and high-performance computing workloads like Computer-Aided Engineering (CAE) or Electronic Design Automation (EDA) and Bioinformatics with its dense design and option for liquid cooling.

- **ThinkSystem SR665 V3** is a 2-socket 2U rack server with the performance and flexibility to manage complex workloads like data management, analytics, virtualization, cloud, and AI. The two-processor configuration provides customers with up to 192 cores from the 4th Gen AMD EPYC processors and up to 6TB of DDR5 memory to maximize the performance of this 2U server. Neptune cooling on the SR665 V3 is planned for a future date.

- **ThinkSystem SR675 V3** is a 2-socket 3U GPU-rich powerhouse server optimized for performance. It provides eight full-size GPUs with up to 160 PCIe lanes to maximize operations, making it ideal for Artificial intelligence and Machine learning. It features a model with liquid-cooled GPUs in its air-cooled chassis to increase the efficiency of the ThinkSystem SR675 V3 without compromising performance.

- **ThinkSystem SD665 V3** combines the latest AMD processors and our Lenovo Neptune full system direct water-cooling solution for outstanding CPU performance and DDR5 memory bandwidth in an ultra-dense 6U, 12-node form factor chassis ideally suited for HPC clusters for research in areas like weather and climate, or seismic exploration, simulations.

- **ThinkSystem SD665-N V3** combines the latest AMD processors, powerful accelerators, and our Lenovo Neptune full system direct water-cooling ultra-dense 6U chassis, for best-in-class compute performance as needed in complex workloads from Fintech to Molecular dynamics to astrophysics and nuclear research.

Summary

These HPC and AI-optimized servers unleash the power and performance of the 4th Gen AMD EPYC processors while providing improved density and helping to reduce energy costs. Working with AMD before the launch, we've already garnered 24 world record benchmark results on the SR665 V3 server powered by 4th Gen AMD EPYC processors (footnote 3).

Learn more about our new ThinkSystem servers with AMD EPYC processors from the Lenovo-AMD alliance page.

About the author

Patrick Moakley is the Director of Marketing for HPC & AI at Lenovo Infrastructure Solutions Group. He has over 30 years in IT sales and marketing of x86 systems at Lenovo and IBM and has a Bachelor's degree in Economics from the University of Wisconsin-Madison.

Footnotes

[1] EPYC-038: AMD EPYC 9004 Series delivers up to a ~14% geomean IPC single thread uplift generationally based on AMD internal testing as of 09/19/2022, performance improvement at the same fixed-frequency on a 4th Gen AMD EPYC 9554 CPU compared to a 3rd Gen AMD EPYC 7763 CPU using a select set of workloads (33) including est. SPECrate®2017_int_base, est. SPECrate®2017_fp_base, and representative server workloads. LINK

[2] 4TH GEN AMD EPYC PROCESSOR ARCHITECTURE' White Paper, published by AMD. LINK

[3] Lenovo internal summary document. ThinkSystem Performance and Benchmarks. The summary document references all benchmark attributions. Nathan Pham, Principal Engineer. As of Dec 1, 2022

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
8001 Development Drive
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document, LP1686, was created or updated on January 17, 2023.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1686
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1686.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at https://www.lenovo.com/us/en/legal/copytrade/.

The following terms are trademarks of Lenovo in the United States, other countries, or both:
Lenovo®
Neptune®
ThinkSystem®

The following terms are trademarks of other companies:

AMD and AMD EPYC™ are trademarks of Advanced Micro Devices, Inc.

SPECrate® is a trademark of the Standard Performance Evaluation Corporation (SPEC).

IBM® is a trademark of IBM in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.