Lenovo Launches New and Updated Servers with 5th Gen Intel Xeon Scalable Processors

Article

Lenovo has expanded its portfolio with new ThinkSystem servers and ThinkAgile hyperconverged solutions that advance cloud deployment, hybrid connectivity and AI capabilities, powered by the 5th Gen Intel® Xeon® Scalable Processors.

New ThinkSystem servers featuring the 5th Gen Intel Xeon Scalable processors:

- ThinkSystem SD530 V3
- ThinkSystem SD550 V3
- ThinkSystem SD650-N V3

ThinkAgile systems and ThinkSystem servers now updated with 5th Gen Intel Xeon Scalable processors:

- ThinkAgile HX with Nutanix software
- ThinkAgile VX with VMware vSAN software
- ThinkAgile MX with Microsoft Azure software
- ThinkSystem SR650 V3
- ThinkSystem SR630 V3
- ThinkSystem SD650 V3
- ThinkSystem SD650-I V3

New ThinkSystem servers featuring the Intel Xeon E-2400 processor:

- ThinkSystem ST250 V3
- ThinkSystem SR250 V3

In addition to its advanced hybrid cloud platform, Lenovo is helping customers accelerate AI implementation through its new Professional Services for AI. Powered by Lenovo's vast knowledge base and global team of industry experts, the service is a single vendor solution that helps organizations make sense of AI and guides them through the most effective, end-to-end IT deployments that are tailored to achieve desired business outcomes quickly.

Visit https://www.lenovo.com/us/en/servers-storage/hybrid-cloud-ai-solutions/ to learn more.

Introducing the new 5th Gen Intel Xeon Scalable Processor

Get impressive performance per watt gains across all your workloads, plus outsized performance and TCO in AI, database, networking, and HPC.(2) 5th Gen Intel Xeon Scalable processors deliver more compute and faster memory(3) at the same TDP as the previous generation. 5th Gen Xeon processors are software-compatible and platform-compatible with the previous generation, so you can minimize testing and validation when deploying new systems.

With AI acceleration in every core, 5th Gen Intel Xeon Scalable processors are ready to handle your demanding AI workloads—including inference and fine tuning on models up to 20 billion parameters(4)—before you need to add discrete accelerators.

Benefits of the new processors include:

- Performance benefits:
  - 21% average performance gain at the same TDP as 4th Gen Intel Xeon Scalable processors(1)
  - Up to 16% memory bandwidth improvement(3) and 2.7x increased last level cache(5) vs. 4th Gen Intel Xeon processors
  - 87% average performance gain over 3rd Gen Intel Xeon Scalable processors(6)
  - Out-of-box software performance using accelerators
- AI benefits:
  - Less than 100 ms second-token latency on LLMs under 20 billion parameters(4)
  - Up to 2.7x better AI inference performance vs. 4th Gen AMD EPYC processors(7)
  - Up to 14x better AI training and inference performance vs. 3rd Gen Intel Xeon processors(8)
  - Software tools and ecosystem to accelerate AI
  - Confidential computing to help protect AI models
- Efficiency benefits:
  - Up to 10x higher performance per watt using built-in accelerators(9)
  - Up to 3x higher performance per watt with built-in accelerators vs. 4th Gen AMD EPYC processors(10)
- Quality and security benefits:
  - Software- and pin-compatible with 4th Gen Intel Xeon processors
  - Silicon-based security features, confidential computing, and trust services
  - Leading quality and enhanced telemetry
  - Largest ecosystem of hardware and software vendors

To see the technical specifications for each 5th Gen processor model, see the Intel Xeon Scalable processor comparison reference: https://lenovopress.lenovo.com/lp1262-intel-xeon-scalable-processor-comparison#generation=5

Notes:

1. Average performance gain as measured by the geomean of SPEC CPU rate, STREAM Triad, and LINPACK compared to 4th Gen Intel Xeon processor. See G1 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

2. As measured by performance per watt on a range of AI, database, networking, and HPC workloads compared to 4th Gen Intel Xeon processor. See A2, A19-A25, D1, D2, D5, H1, N16 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

3. See G12 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

4. Based on Intel internal modelling as of December 2023.

5. See G11 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

6. Average performance gain as measured by the geomean of SPEC CPU rate, STREAM Triad, and LINPACK compared to 3rd Gen Intel Xeon processor. See G3 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

7. Based on performance gains of 1.19x to 2.69x with Intel Advanced Matrix Extensions (Intel AMX) for inference on GPT-J, LLaMA-2 13B, DLRM, DistilBERT, BERT-Large, and ResNet50v1.5 compared to AMD EYPC 9654 and 9754. See A201, A202, A208-A211 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

8. Based on performance gains of 4.4x to 14.2x for training (ResNet50v1.5, BERT-Large, SSD-ResNet34, RNN-T, MaskRCNN, and DLRM) and 2.9x to14x for inference (ResNet50v1.5, BERTLarge, SSD-ResNet34, RNN-T (BF16 only), Resnext101 32x16d, MaskRCNN (BF16 only), DistilBERT) compared to 3rd Gen Intel Xeon processor. See A15-A16 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

9. Based on performance per watt gains of 1.46x to 10.6x with built-in accelerators on a range of AI, database, and networking workloads. See A19-A25, D1, D2, D5, N16 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

10. Based on performance per watt gains of 1.11x to 2.96x with built-in accelerators on a range of AI, database, networking, and HPC workloads compared to AMD EYPC 9554, 9654, and 9754. See A208-A211, D201-D204, H201, N201 at https://intel.com/processorclaims: 5th Gen Intel Xeon Scalable processors. Results may vary.

ThinkSystem V3: Now with 5th Gen Intel Xeon CPUs

ThinkSystem servers now updated with 5th Gen Intel Xeon Scalable processors.

ThinkSystem SR650 V3

Our mainstream 2U rack server, the SR650 V3, now supports two 5th Gen Intel Xeon Scalable processors, up to 64 cores and 350W TDP. The server also supports the new 5600 MHz memory for a total of 8TB of high-performance memory using 16x 256GB 3DS RDIMMs. Also just announced are three new Gen5 riser cards for high-performance PCIe adapters.

For details, see these updated resources:

- Product guide
- Datasheet
- Interactive 3D Tour

ThinkSystem SR630 V3

Our mainstream 1U rack server, the SR630 V3, also now supports two 5th Gen Intel Xeon Scalable processors, up to 64 cores and 350W TDP. The server also supports the new 5600 MHz memory for a total of 8TB of high-performance memory using 16x 256GB 3DS RDIMMs.

For details, see these updated resources:

- Product guide
- Datasheet
- Interactive 3D Tour

ThinkSystem SD650 V3

The SD650 V3 is our dual-node water-cooled server tray, with two dual-processor nodes in a 1U direct-water-cooled (DWC) tray design. The server now supports two 5th Gen Intel Xeon Scalable processors and 5600 MHz DDR5 memory. The server supports processors up to 385W, including the Intel Xeon Platinum 8593Q processor model.

For details, see these updated resources:

- Product guide
- Datasheet
- Interactive 3D Tour

ThinkSystem SD650-I V3

The SD650-I V3 is a water-cooled server tray that combines two Intel processors and four powerful Intel Data Center Max Series GPUs in a single 1U direct-water-cooled (DWC) tray design. The server now supports two 5th Gen Intel Xeon Scalable processors and 5600 MHz DDR5 memory. The server supports processors up to 385W, including the Intel Xeon Platinum 8593Q processor model.

For details, see these updated resources:

- Product guide
- Datasheet
- Interactive 3D Tour

SD530 V3 and SD550 V3 Multi-Node Servers

The Lenovo ThinkSystem SD530 V3 and SD550 V3 are economical two-socket server nodes in half-width rack form factors. Combining the efficiency and density of blades with the value and simplicity of rack-based servers, the SD530 V3 and SD550 V3 deliver a cost-efficient scale out platform that is thermally designed to deliver maximum performance in the smallest footprint.

Each SD530 V3 node supports two high-performance 5th Gen Intel Xeon Platinum processors, up to 2TB of DDR5 memory, and two hot-swap E3.S NVMe SSD drives per node. At the rear of the server are a low-profile PCIe Gen5 x16 slot for a GPU or other adapter, and an OCP 3.0 PCIe Gen5 x16 slot for networking.

Figure 1. ThinkSystem SD530 V3

Each SD550 V3 node also supports high-performance 5th Gen Intel Xeon Platinum processors, up to 2TB of DDR5 memory and six hot-swap SAS/SATA/NVMe SSD drives per node for workloads requiring more internal storage. The node also offers two low-profile PCIe x16 expansion slots (1x PCIe Gen5, 1x PCIe Gen4) for 2x GPUs or other adapters, and one OCP 3.0 PCIe Gen5 x16 slot for networking.

Figure 2. ThinkSystem SD550 V3

The SD530 V3 and SD550 V3 are installed in the ThinkSystem D3 Chassis, which holds either 4x SD530 V3 nodes, or 2x SD550 V3 nodes or a mix of the two. The D3 Chassis holds the 3x hot-swap power supplies used to power all installed nodes.

Figure 3. ThinkSystem D3 Chassis with 4x SD530 V3 nodes installed

For more information, see these resources:

- SD530 V3:
  - Article, Five Highlights of the Lenovo ThinkSystem SD530 V3 Server
  - Datasheet
  - Interactive 3D Tour
- SD550 V3
  - Article, Five Highlights of the Lenovo ThinkSystem SD550 V3 Server
  - Datasheet
  - Interactive 3D Tour

Note: The systems are planned to be configurable in DCSC in February 2024. Product Guides will be available at the same time.

SD650-N V3 Neptune DWC Server

The ThinkSystem SD650-N V3 Neptune DWC server is a high-performance server based on the fifth generation Lenovo Neptune™ direct water cooling platform.

With two 5th Gen Intel Xeon Scalable processors or two Intel CPU Max Series processors, along with four NVIDIA H100 SXM5 GPUs, the ThinkSystem SD650-N V3 server features the latest technology from Intel and NVIDIA, combined with Lenovo's market-leading water-cooling solution, which results in extreme performance in an extreme dense packaging.

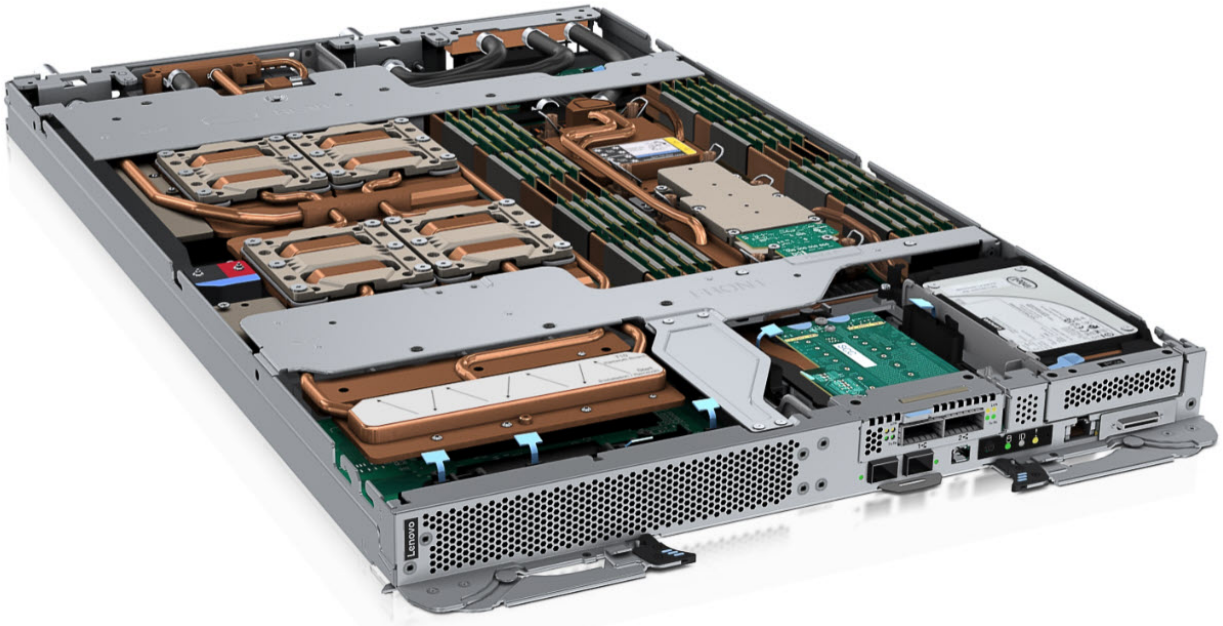

Figure 4. ThinkSystem SD650-N V3

For more information on the SD650-N V3, see these resources:

- Datasheet
- Interactive 3D Tour

Note: The systems are planned to be configurable in DCSC in early 2024. The Product Guide will be available at the same time.

SR250 V3 and ST250 V3 Servers

The Lenovo ThinkSystem SR250 V3 and ST250 V3 are new high-value single-socket servers that provide enterprise-class reliability, management, and security. These servers are suitable for multi-site locations including retail, and for growing businesses that need optimized performance. The SR250 V3 is a 1U rack server supporting 10x 2.5-inch drives or 4x 3.5-inch drives, and the ST250 V3 is a 4U tower server supporting 16x 2.5-inch drives or 8x 3.5-inch drives.

The SR250 V3 and ST250 V3 support one Intel Xeon E-2400 Series processor (formerly codenamed "Raptor Lake-E") and up to 128 GB of 4800 MHz TruDDR5 ECC memory. Both servers use the next-generation Lenovo XClarity Controller 2 (XCC2) management processor with provides advanced control, monitoring, and alerting functions.

Figure 5. Lenovo ThinkSystem SR250 V3 with 10x 2.5-inch hot-swap drives

Figure 6. Lenovo ThinkSystem ST250 V3 with 16x 2.5-inch hot-swap drives

For more information, see these resources:

- SR250 V3:
  - Datasheet
  - Interactive 3D Tour
- ST250 V3:
  - Datasheet
  - Interactive 3D Tour

Note: The systems are planned to be configurable in DCSC in January 2024. Product Guides will be available at the same time.

Related product families

Product families related to this document are the following:

- Processors
- ThinkSystem SD530 V3 Server
- ThinkSystem SD550 V3 Server
- ThinkSystem SD650 V3 server
- ThinkSystem SD650-I V3 server
- ThinkSystem SD650-N V3 Server
- ThinkSystem SR250 V3 Server
- ThinkSystem SR630 V3 Server
- ThinkSystem SR650 V3 Server
- ThinkSystem ST250 V3 Server

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, LP1866, was created or updated on December 14, 2023.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1866
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1866.

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®
Lenovo Neptune®
ThinkAgile®
ThinkSystem®
XClarity®

The following terms are trademarks of other companies:

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Microsoft® and Azure® are trademarks of Microsoft Corporation in the United States, other countries, or both.

SPEC® and SPEC CPU® are trademarks of the Standard Performance Evaluation Corporation (SPEC).

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