Deploying AI at the Edge with Lenovo ThinkEdge Series Servers

Solution Brief

Edge computing challenge and a solution

The benefits of AI In today's business environment are becoming more obvious as companies strive to take advantage of the new technology. Although typically most of the AI processing tends to occur in the data center, there are many opportunities at edge or remote locations where the data is created or collected. The Lenovo ThinkEdge series provides an ideal combination of compact size with abundant storage and powerful processing. Placing ThinkEdge servers at remote locations can help offload the AI data and processing demands on the datacenter.

The ThinkEdge series servers are ideal for deployment in tight spaces. Depending on the model, they can be mounted on a wall, or ceiling, placed on a desk or mounted in a rack. These edge servers place increased processing power, storage, and network closer to where data is generated, allowing actions resulting from the analysis of that data to take place more quickly. The ThinkEdge servers are designed to handle a wide range of workloads, such as: Augmented Reality, CDN, Edge AI & MRP, Machine Learning, NFV, Online Gaming, Realtime processing, Smart surveillance, Video Analytics & Streaming.

Figure 1. Lenovo ThinkEdge SE455 V3

Lenovo ThinkEdge offerings are ideal for running applications at remote locations because of their low cost and high-performance capabilities. They are industry-standard servers providing cost effective computing and fast high-density local storage at the edge. With their smaller footprint, the ThinkEdge series servers are ideal for deployment in remote office and edge locations.

AI Edge Use Cases

A key benefit of AI and edge computing is the ability of devices to locally compute, process, and analyze data with high quality and minimal network latency from typical datacenter connections. There are many examples of edge computing use cases but here are a few popular ones.

- Healthcare – Patient and equipment monitoring
- Retail and grocery stores – Cashier-less payments, real-time inventory, security
- Financial Services – Kiosk safety and fraud prevention
- Manufacturing – Automation, smart sensors, quality assurance
- Gas stations – Payments, security, inventory
- Transportation – Trucking, shipping, autonomous vehicles

Figure 2. Lenovo ThinkEdge SE450

Microsoft Azure SQL Edge on ThinkEdge Systems

SQL Server is a well-known, mission-critical database system used in many customer datacenters. It is also available in an edge-sized version that runs efficiently on edge servers as a containerized application. Azure SQL Edge is a small-footprint, robust IoT database product that runs on Lenovo Edge servers to process the data locally for increased security and reduced latency.

The combination of Azure SQL Edge and Lenovo's ThinkEdge series servers provides the ideal solution for processing data aggregated from scores of IoT devices in typical edge use cases like retail, transportation, and manufacturing. Azure SQL Edge also has Machine Learning (ML) inferencing capabilities. ML Models can be trained on-premises and deployed to the edge where AI inferencing can be done.

Figure 3. Lenovo ThinkEdge SE360 V2

Lenovo ThinkEdge Portfolio

Below is a summary of current Lenovo ThinkEdge servers and configurations available. Below are GPUs supported on the edge servers.

GPUs are important in any AI solution due to their parallel processing capability. They contain thousands of smaller cores designed for handling multiple tasks simultaneously. This is useful for AI tasks that require a large number of matrix multiplications and other parallel computations.

GPUs can also manage large volumes of data more efficiently than traditional CPUs due to their parallel architecture. After training, AI models need to make predictions or inferences. GPUs are also used to speed up the inference process, better enabling real-time AI applications.

System	Processor	GPU	Memory	Storage
SE455 V3 (2U)	1-socket AMD EPYC™ 8004 Series (4th Gen AMD EPYC), up to 64 cores and 200W TDP	Up to 6x single-width GPUs or 2x double-width GPUs	6x TruDDR5 slots; Maximum 576GB using 6x 96GB RDIMMs	4x 2.5-inch 15mm front drives 4x 2.5-inch 15mm internal 2x M.2 boot drives
SE450 (2U)	1-socket 3rd Gen Intel Xeon Platinum, up to 36 cores, up to 225W TDP	Up to 4x single-width GPUs or 2x double-width GPUs	10x DDR4 slots; Maximum 1TB using 8x 128GB 3DS RDIMMs	6x 2.5-inch 7mm drives; Up to 6x NVMe drives 2x M.2 boot drives
SE360 V2 (2U)	1-socket Intel Xeon D-2700, up to 16 cores	Support for NVIDIA A2, NVIDIA L4, Intel Data Center GPU Flex 140, Qualcomm Cloud AI 100	Up to 256GB in 4x slots, using 64GB DIMMs; 3200MHz TruDDR4	2x SATA/NVMe 2.5" 7mm drives; or 8x M.2 using NVMe, 2x M.2 2280 NVME boot drives

For More Information

To learn more about this Lenovo solution contact your Lenovo Business Partner or visit: https://www.lenovo.com/us/en/solutions/ai/

References

Lenovo ThinkEdge servers

ThinkEdge SE455 V3

ThinkEdge SE450

ThinkEdge SE360 V2

Lenovo Edge AI solutions

Lenovo and Azure SQL Edge solution

Related product families

Product families related to this document are the following:

- Edge Servers
- Microsoft Alliance

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, LP1981, was created or updated on June 27, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1981
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1981.

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®
ThinkEdge®

The following terms are trademarks of other companies:

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

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

Microsoft®, Azure®, and SQL Server® 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.