

AI Server

AGS8200

The logo for Edgecore AI, featuring the word "Edgecore" in black with a red swoosh above the "e", and "AI" in a larger, bold black font to the right.

The Edgecore AGS8200 is a high-performance, scalable, GPU-based server suitable for AI (Artificial Intelligence)/ML (Machine Learning) applications. The server is ideal for training large language models, automation, object classification and recognition use cases. The system is designed with eight Intel® Habana® Gaudi® 2 processors and dual Xeon® Sapphire-Rapids processors. The Gaudi® 2 processor integrates 96GB HBM2E memory and 24 NICs of 100GbE RoCEv2 RDMA. The 24 100G NICs offer all-to-all connectivity and scale-out internally and externally for training, fine-tuning, and other deep learning processing.

Each Gaudi® 2 processor has 21 x 100GbE non-blocking, all-to-all connectivity to other Gaudi® 2 processors within the server, allowing training across all eight Intel Gaudi® 2 processors without requiring an external Ethernet switch.

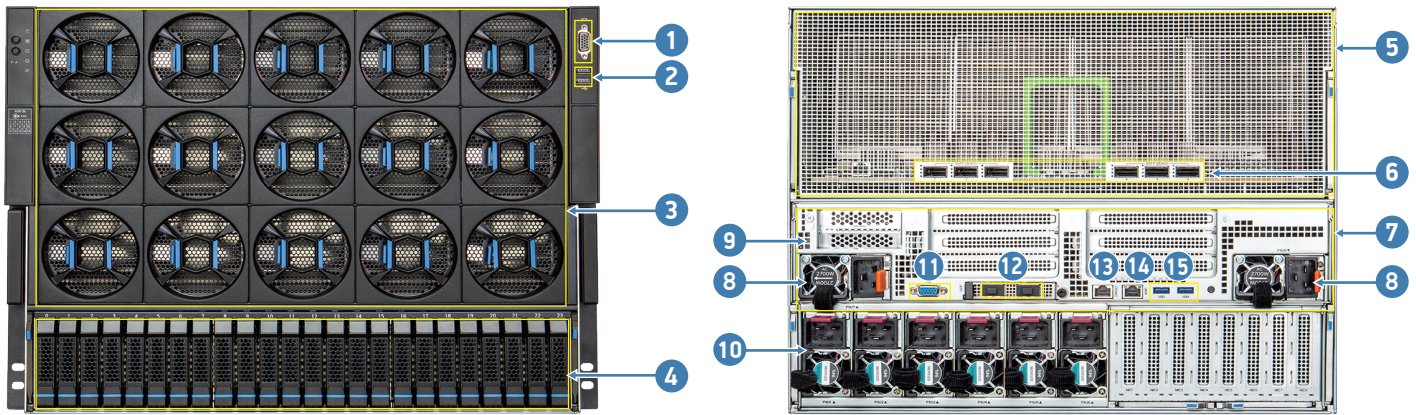
Each AGS8200 supports 6 QSFP-DD ports for scale-out. The 400G ports can be connected to 400G switches, or 100G switches via breakout cables, in racks and clusters of Intel® Gaudi® 2-based nodes.

The server is able to include 16 x HDD/SSD + 8 x NVMe or 8 x HDD/SSD + 16 x NVMe for storage that supports RAID HBA with 0/1/10/5/6/50/60.

Key Features and Benefits

- Featuring eight Intel® Gaudi® 2 AI accelerators that deliver leading training and inferencing performance
- Various cluster architectures can be built using the same servers to scale the AI training and inference cluster utilizing thousands of Intel Gaudi® 2s
- Dual 4th Gen Intel® Xeon® scalable processors
- Expanded networking capacity with 24 x 100GE RoCE ports integrated into every Gaudi® 2
- 21 RoCEv2 100GbE connectivity allows training across all 8 Intel Gaudi® 2 processors no external Ethernet switch required
- Maximum 24 x 100GbE, 3 ports per Intel Gaudi® 2 accelerator, 6 400GbE QSFP-DD per server scale out bandwidth
- Ease of system build or migration with Habana® SynapseAI® software stack
- Standardized architecture and Ethernet instead of proprietary Infiniband and NVLINK

Interfaces



Description

- | | |
|---|---|
| 1. VGA port | 9. 8 x PCIe slots (Half Height (left 2), Full Height (Central 6)) |
| 2. 2 x USB 2.0/3.0 | 10. 3+3 redundant/hot-swappable AC PSUs for GPU |
| 3. 14+1 Hot-swappable fans | 11. VGA port |
| 4. 16 HDD/SSD+8 NVMe | 12. OCP 3.0 |
| 5. GPU module | 13. Serial console port (UART 115200) |
| 6. 6 x QSFP-DD 400GbE ports | 14. Management port 1GbE |
| 7. Server module | 15. 2 x USB 2.0/3.0 |
| 8. 1+1 redundant/hot-swappable AC PSUs for system | |

Form Factor

- 8U

Compute Node

- CPU: Sapphire Rapids, 2 Sockets
Intel® Xeon Gold 6448H, 32c, 64 threads, 60MB, 250 W
- PCH: Emmitsburg
- Memory: Up to 2TB 16x DDR5 memory slots per CPU
32 x 32GB DDR5 (1TB)
- Operating System: Ubuntu 20.04
- BIOS: 32MB Flash

GPU

- 8 x OAM (Intel Habana HL-225H/C)

Input/Output

- Front: 2 x USB 2.0/3.0, 1 x VGA, 1 x UID, 1 x PWR
- Rear: 2 x USB 2.0/3.0, 1 x VGA, 1 x RJ-45, 1 x UID

Scale-Out Interface

- RDMA (ROCE v2)
- RDMA
- 6 x QSFP-DD

Storage

- Internal: 2 x M.2 SATA SSD (480 GB)
- Front: 16 x HDD/SSD+ 8 x NVMe or 8 x HDD/SSD+ 16 x NVMe
16 x 2.5" SATA SSD (1TB) + 8x 2.5" NVMe (2TB)
- 1 RAID card

BMC

- AST2600

TPM 2.0

CD-ROM

- Support for external USB CD-ROM

PSU

- System: 1+1 CRPS 2700 W redundant/hot-swappable AC/DC
- GPU: 3+3 CRPS 3000 W redundant/hot-swappable AC/DC

Fans

- 14+1 hot-swappable fans

Dimensions

- 900 mm x 447 mm x 352mm

Operating Temperature

- 5°C-35°C

Expansion Slots

- 1 x OCP 3.0
- 8 x PCIe slots

Software

- SynapseAI: 1.13.0
- Kernel: 5.4.0 and above
- Python: 3.10
- PyTorch: 2.1.0
- TensorFlow: 2.13.1
- Open MPI: 4.1.5
- Libfabric: 1.16.1 and above
- Transformers: >= 4.33.0, <4.35.0

Warranty

Please check www.edge-core.com for the warranty terms in your country.

For More Information

To find out more about Edgecore Networks Corporation products and solutions, visit www.edge-core.com.

About Edgecore Networks Corporation

Edgecore Networks Corporation is in the business of providing innovative network solutions. In the service provider network, in the data center or in the cloud, Edgecore Networks Corporation delivers the software and systems that transform the way the world connects. Edgecore Networks Corporation serves customers and partners worldwide. Additional information can be found at www.edge-core.com.

Edgecore Networks Corporation is a subsidiary of Accton Technology Corporation, the leading network ODM company. The Edgecore data center switches are developed and manufactured by Accton.

To purchase Edgecore Networks solutions, please contact your Edgecore Networks Corporation representatives at +886 3 563 8888 (HQ) or +1 (949)-336-6801 or authorized resellers.

© Copyright 2025 Edgecore Networks Corporation. The information contained herein is subject to change without notice. This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered by Edgecore Networks Corporation. Edgecore Networks Corporation shall not be liable for technical or editorial errors or omissions contained herein.