5 REASONS AMD EPYC™ 9005 PROCESSORS ADVANCE ENTERPRISE AI IN GPU-ACCELERATED SOLUTIONS

AT A GLANCE

GPU accelerators have become the workhorse for modern AI, excelling in training large, complex models and real-time inference at scale. However, maximizing the potential of your GPU investment requires a powerful CPU partner.

By choosing an AMD EPYC 9005 CPU, you can be confident that your AI workload will benefit from its high performance, efficiency, and compatibility. AMD EPYC CPUs are designed to accelerate AI training and inference, enabling you to achieve faster results and optimize your AI project outcomes.

1

HIGH FREQUENCY & CORE COUNT

High frequency allows the CPU to process instructions faster, crucial for handling extensive data tasks. Additionally, it facilitates faster post-processing of results generated by the AI model.

5th Gen EPYC 9575F (64C) is purpose-built to be a high performing AI host-node processor running at speeds up to 5GHz.[1] As a GPU host node, 9575F (64C) runs the GPU system performance ~20% faster than an Intel Xeon 8592 (64C) in a Llama3.1-70B Inference Benchmark at FP8 precision (2P servers each hosting 8 GPUs) [9xx5-014]

2

LARGE CACHE SIZE

A large cache size reduces the need to fetch data from slower main memory, improving performance. This is essential for facilitating fast access to massive datasets in training and running AI models and helps minimize latency, leading to quicker training times and faster execution.

AMD EPYC 9005 CPUs offer large L3 cache ranging from 64MB to 512MB.

3

HIGH MEMORY BANDWIDTH & HIGH PERFORMANCE I/O

CPUs with high memory bandwidth and high-performance I/O interfaces are well-suited for AI tasks, enable efficient data flow, enabling rapid data transfer between the CPU and GPU, and helping minimize bottlenecks during AI model training and inference.

With support for up to 12 channels of DDR5 memory and 128 PCIe® 5 I/O lanes per CPU, AMD EPYC 9005 processors deliver exceptional memory and I/O bandwidth.

4 ENERGY-EFFICIENT CORES

AMD EPYC processors have energy-efficient cores that enable efficient GPU-accelerated systems, optimizing high performance, energy efficiency, memory, and IO scalability. AMD EPYC processors power energy-efficient servers, delivering exceptional performance and helping reduce energy costs.

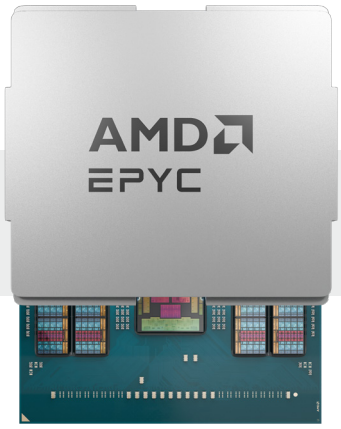
5 GPU ECOSYSTEM COMPATIBILITY

GPU compatibility enables seamless communication and optimized performance between different components. This allows for efficient utilization of GPU resources and facilitates smooth software integration, leading to faster training times and better model performance.

AMD EPYC CPUs are compatible with a wide range of GPUs, including those from Nvidia and AMD. This allows for flexibility in choosing the best GPU for your specific AI needs. Additionally, AMD collaborates closely with major AI software vendors to optimize their software for AMD EPYC processors, enabling smooth integration and efficient utilization of resources.

5TH GEN AMD EPYC™ PROCESSORS ADVANCE DATA CENTER AI

TOGETHER WE ADVANCE_AI



1 - EPYC-018: Max boost for AMD EPYC processors is the maximum frequency achievable by any single core on the processor under normal operating conditions for server systems

PID: 242908532-A October 2024