

Preview

In 2025, global AI chips focus on high-end HBM memory; NVIDIA's new Blackwell platform drives growth, amid geopolitical limits and steady AI server demand, with rapid HBM technology evolution toward HBM4 in 2026. io, Quantum ActiveScale, Cloudian, OSNexus and Scality. Optimal storage architecture balances both flash-based parallel file system high-performance storage and. These include GPUs, custom AI application-specific integrated circuits (ASICs) used by hyperscalers and cloud service providers (CSPs), AI-capable central processing units (CPUs), and other AI ASICs developed by both AI chip-focused startups and large vendors. High-end AI chips primarily use HBM memory; mid- to low-end rely on GDDR. NVIDIA's. Huawei OceanStor Next-Gen High-Performance Distributed File Storage for AI provides a unified storage solution for the end-to-end (E2E) AI training and inference data process. It helps enterprises overcome data silos, aggregate diverse corpus data, improve AI cluster computing power utilization. Based on our experience running AIMultiple's cloud GPU benchmark with 10 different GPU models in 4 different scenarios, these are the top AI hardware companies for data center workloads. Follow the links to see our rationale behind each selection: Revenue & volume leader. Qualcomm AI250 introduces an innovative memory architecture, offering a generational leap in effective memory bandwidth.

End-to-end solutions for integrated direct-to-chip liquid cooling, CDUs, and power infrastructure for AI factories--optimized for GPU densities above 100

Together, we are committed to delivering the same high-quality content and insights that have been the hallmark of ITPro Today, Network Computing, and IoT World Today.

Detailing and benchmarking hardware architectures for current market players and emerging AI chips for server CPUs, custom AI ASICs, and other AI chips,

Designed for your current needs and future ambitions, Marvell delivers the data infrastructure technology transforming tomorrow's

Discover the latest advancements in server chip technologies, from Nvidia's Blackwell and AMD's EPYC to Google's Axion and Intel's upcoming Granite Rapids. Learn how AI-driven innovations are shaping

Yole Group - Access daily business, market & technology updates in the semiconductor industry, our Analysts' Analysis and Presentations and more

Memory prices are high because manufacturers have shifted factories toward lucrative AI and server chips, creating an artificial squeeze on everyday

High-end storage chips for AI servers

According to TrendForce's latest findings on AI servers, NVIDIA's high-end AI chip shipment mix is expected to change in 2026. The combined share of Hopper and Rubin series in

Supply Chain & Distribution Hardware Root of Trust Essential for AI Chip Integrity

An AI ASIC accelerator is a custom chip purpose-built for AI tasks such as inference or training. It offers high energy efficiency and performance for targeted

This reliance on high-end AI servers necessitates the use of high-end AI chips, which in turn will not only drive up demand for HBM during 2023~2024,

The cause is mostly the newly booming AI industry, as AI servers require a lot of memory, both in terms of long-term storage and short-term

With increasing requirements for large-scale data storage and transfer in big data, cloud computing, and AI, there is strong demand for high

AI chips enable parallel computing capabilities are increasingly in demand. This article will information to you on 10 popular AI chip makers.

During recent years the Non-Volatile Memory express (NVMe) solid-state-drive (SSD) has become the de facto storage technology. NVMe provides

New NVIDIA Inference Context Memory Storage Platform with NVIDIA BlueField-4 storage processor to accelerate agentic AI reasoning. Microsoft's next-generation Fairwater AI superfactories

Web: <https://www.kremgroup.co.za>

