

How big is Huawei s AI server



Overview

SoftStone Information Technology subsidiary SoftStone Huafang introduced a 6U dual-socket AI server product, "Super A860 A5," designed to support up to eight Atlas 350 accelerator cards and based on a new model of the Kunpeng 920 processor, according to statements made at the. SoftStone Information Technology subsidiary SoftStone Huafang introduced a 6U dual-socket AI server product, "Super A860 A5," designed to support up to eight Atlas 350 accelerator cards and based on a new model of the Kunpeng 920 processor, according to statements made at the. Huawei Technologies used its China Partner Conference 2026 to push its Ascend computing stack into broader commercial deployment, unveiling the Ascend 950PR-powered Atlas 350 accelerator card and positioning it as a higher-throughput alternative to Nvidia's H20 for inference workloads. At the. BEIJING, April 29 (Reuters) - Demand for Huawei's Ascend 950 AI chips has surged following the release of DeepSeek's V4 artificial intelligence model that runs on the Shenzhen-based tech firm's chips, with major Chinese internet firms rushing to secure orders, three people familiar with the matter. China's domestic AI chips took 41% of the accelerator server market in 2025. China's AI hardware landscape shifted

dramatically in 2025, with domestic chip makers claiming nearly half the country's AI accelerator server market. 9x the power of Nvidia's most powerful AI server the GB200 NVL72, Huawei's CloudMatrix 384 cluster of Ascend 910C chips delivers twice the compute performance.

How big is Huawei's AI server



Huawei unveiled Ascend 950PR-based Atlas 350 at Partner Conf 2026, claiming 2.87x Nvidia H20 compute, FP4 inference, 112GB HBM and 1.4TB/s bandwidth.



Huawei launches the Ascend 950PR with 1.56 petaflops and 112GB HBM. How the 6,000 AI chip compares to Nvidia and reshapes China's OB market.



Huawei's Ascend AI processor attracted more than 3,000 partners and 4 million developers while Kunpeng, the company's server central processing units, garnered more than 6,800 ...



With more than 3.6x aggregate memory capacity and 2.1x more memory bandwidth, Huawei and China now have AI system capabilities that can beat Nvidia's.



In addition to announcing its first AI cluster with 1 FP4 ZettaFLOPS performance, Huawei also revealed a detailed roadmap of its upcoming Ascend neural processing units (NPUs) that ...



BEIJING, April 29 (Reuters) - Demand for Huawei's Ascend 950 AI chips has surged following the release of DeepSeek's V4 artificial intelligence model that runs on the Shenzhen-based tech firm's ...



At \$8 million, the price of the Huawei cluster is 3x the price of the Nvidia GB200 NVL72, while power consumption is almost 4x more, which shows that Huawei has developed the cluster to ...



Huawei shipped 812,000 AI chips in 2025 as Chinese firms claim 41% of China's AI server market, reshaping Asia's hardware landscape.

Contact Us

For more information, pricing, or custom energy solutions, please contact us:

Website: <https://gdroofing.co.za>

Email: sales@gdroofing.co.za

Phone: +27 72 418 9365

Address: 22 Electron Avenue, Isando, Johannesburg, 1600, South Africa

This document is for informational purposes only. Specifications subject to change without notice.

