

21.09.2025 – 07:10 Uhr

Huawei Cloud: Förderung des fruchtbaren Bodens für Datenverarbeitung, Stärkung von KI-Pionieren für die Industrie

Shanghai (ots/PRNewswire) -

Am zweiten Tag der HUAWEI CONNECT 2025 hielt Zhang Ping'an, Vorstandsmitglied von Huawei und Geschäftsführer von Huawei Cloud, eine Grundsatzrede mit dem Titel „All Intelligence: Empowering AI Pioneers for Industries“ (All Intelligence: Stärkung von KI-Pionieren für die Industrie). Er berichtete über die Innovationen und Praktiken der Huawei Cloud in den Bereichen KI-Computing-Services, Gründungsmodelle, verkörperte KI, KI-Agenten und vieles mehr.

Ständige Innovation im Bereich KI-Computing-Services: leistungsstarkes Computing im Zeitalter der Intelligenz

In diesem Jahr hat Huawei Cloud seinen KI-Computing-Service angekündigt, der auf CloudMatrix384 basiert. Die Spezifikationen des Huawei CloudMatrix-Supernodes werden von 384 Karten auf 8.192 Karten erweitert. Die Supernodes können einen Hyperscale-Cluster mit 500.000 bis 1 Million Karten unterstützen und bieten somit robuste KI-Rechenleistung, eine unschätzbare Ressource im Zeitalter der Intelligenz. Huawei Cloud kündigte außerdem mit seinem Elastic Memory Service (EMS) eine innovative Speicherlösung an, die als Branchenneuheit den Video-RAM mit Speicher erweitert. Dadurch wird die Latenzzeit von Konversationen mit mehreren Runden auf Foundation-Modellen drastisch reduziert und die Benutzerfreundlichkeit erheblich verbessert.

Huawei Cloud hat vollständig flüssigkeitsgekühlte KI-Rechenzentren in den chinesischen Regionen Guizhou, Innere Mongolei und Anhui eingerichtet. Diese KI-Rechenzentren unterstützen eine Wärmeableitung von 80 kW pro Schrank, reduzieren die Stromverbrauchseffektivität (PUE) auf 1,1 und bieten KI-gestützte Betriebs- und Wartungsfunktionen. Das bedeutet, dass Unternehmen keine traditionellen Rechenzentren umbauen oder neue bauen müssen. Stattdessen benötigen sie lediglich ein Paar Glasfaserkabel, um sich mit dem Rechenzentrum zu verbinden und auf leistungsstarke KI-Rechenleistung sowie dedizierte Full-Stack-KI-Cloud-Services auf Huawei Cloud zuzugreifen.

Zhang Ping'an wies darauf hin, dass der KI-Token-Service der Huawei Cloud die zugrunde liegende technische Komplexität abstrahiert und den Nutzern direkt die endgültigen KI-Berechnungsergebnisse liefert. Auf diese Weise können die Benutzer die Rechenleistung der Inferenzen so effizient wie möglich nutzen. Der Supernode CloudMatrix384 ermöglicht das vollständige Pooling von Rechen-, Speicher- und Speicherressourcen, entkoppelt Rechenaufgaben, Speicheraufgaben und KI-Expertensysteme und wandelt serielle Aufgaben in verteilte parallele Aufgaben um, wodurch die Inferenzleistung des Systems erheblich verbessert wird. In Szenarien mit Inferenzaufgaben mit unterschiedlichen Latenzanforderungen, wie Online-, Nearline- und Offline-Inferenz, liefert CloudMatrix384 eine durchschnittliche Inferenzleistung pro Karte, die drei- bis viermal so hoch ist wie die von H20.

Auf der Konferenz kündigte Zhang Ping'an den offiziellen Start des AI Token Service powered by CloudMatrix384 an. Der Dienst bietet den Kunden überlegene Leistung, Service und Qualität.

Herausforderungen offensiv angehen: Unterstützung von Unternehmen beim Aufbau eigener Modelle

Huawei Cloud hat seine Pangu-Modelle durch die Erforschung branchenspezifischer Szenarien verfeinert und mit seinen Kunden zusammengearbeitet, um deren dringlichste Herausforderungen zu bewältigen und die Möglichkeiten in diesen Branchen neu zu definieren. Huawei nutzt openPangu, um Best Practices für KI-Training und -Inferenz bereitzustellen, die es Entwicklern erleichtern, die KI-Rechenleistung effizient zu nutzen. Zhang Ping'an merkte an, dass Huawei zur gleichen Zeit das quelloffene Pangu-Modell entwickelt. Huawei wird die Investitionen in Pangu-Modelle kontinuierlich erhöhen, Branchenszenarien untersuchen, um die Kundenanforderungen besser zu verstehen, und Kunden bei der Entwicklung ihrer eigenen branchenspezifischen Modelle unterstützen, um so die intelligente Transformation in allen Branchen zu beschleunigen.

Die Pangu-Modelle wurden bereits in mehr als 500 Szenarien in über 30 Branchen eingesetzt. Sie haben eine wichtige Rolle in Bereichen wie Regierungsdienste, Finanzen, Fertigung, Gesundheitswesen, Kohlebergbau, Stahl, Eisenbahn, autonomes Fahren und Meteorologie gespielt.

Über Terminals hinaus: unbegrenzte Intelligenzentwicklung in der Cloud ermöglichen

In diesem Jahr hat Huawei Cloud die CloudRobo Embodied AI Platform eingeführt, die komplexe Algorithmen und intelligente Logik in der Cloud einsetzt, um leichtere Roboter zu realisieren. Durch die Nutzung der enormen Rechenleistung und der fortschrittlichen KI-Modelle in der Cloud macht die Plattform die Roboterausführung intelligenter. Die Cloud-Intelligenz überwindet die Einschränkungen, die Roboter bisher behindert haben, und macht sie für mehr Szenarien einsetzbar.

Um einen einheitlichen, offenen und sicheren Kommunikationskanal zwischen Robotern und der Cloud aufzubauen, hat Huawei Cloud das R2C-Protokoll (Robot to Cloud) eingeführt. Zhang Ping'an gab bekannt, dass die ersten 20 Partner des R2C-Protokolls offiziell an Bord sind.

Kunpeng Cloud Services: Förderung von Brancheninnovationen durch Software-Hardware-Synergien und ein offenes Ökosystem

Eine der wichtigsten Strategien von Huawei Cloud ist die Entwicklung von Kunpeng-gestützten ARM-Cloud-Services, die Leistung, Sicherheit und Zuverlässigkeit bieten. Im vergangenen Jahr ist die Anzahl der Kunpeng-Rechenkerne in der Huawei Cloud von 9 Millionen auf 15 Millionen gestiegen, was einer Steigerung von 67 % entspricht. Darüber hinaus wurde die Kunpeng-Plattform kontinuierlich verbessert, um die Kompatibilität mit Mainstream-Software zu unterstützen, und wurde an mehr als 25.000 Anwendungen angepasst. Die Kunpeng-Plattform bietet eine solide Unterstützung für Kunpeng Cloud Services, die zusätzlich zu Transcoding, Datenbanken, Webanwendungen und Cloud-Telefonen auch für allgemeinere Computerszenarien eingesetzt werden können.

GaussDB: Aufbau einer effizienten und zuverlässigen Datenbasis auf Basis von Supernodes und vollständigem Pooling

Die GaussDB-Datenbanken der Huawei Cloud basieren auf Allzweck-Computing-Supernodes und ermöglichen das schichtweise Pooling von Rechen-, Speicher- und Storage-Ressourcen sowie das gleichzeitige Lesen und Schreiben mehrerer Daten auf einem beliebigen Node, wodurch die Einschränkungen der herkömmlichen Architektur, bei der nur der primäre Node das Lesen und Schreiben von Daten unterstützt, aufgehoben werden. GaussDB-Datenbanken unterstützen auch die dynamische Lastplanung, wodurch die Leistung der parallelen Transaktionsverarbeitung erheblich verbessert wird. Ein GaussDB-Cluster auf der Basis von Supernodes kann 5,4 Millionen Transaktionen pro Minute verarbeiten, was einer 2,9-fachen Leistungssteigerung gegenüber Clustern ohne Supernodes entspricht.

Verteilte Cloud für alle Szenarien: universelle und optimale Rechenleistung mit lokalem Zugriff

Huawei Cloud hat eine verteilte Cloud-Lösung entwickelt, die alle Szenarien abdeckt, einschließlich CloudOcean, CloudSea, CloudLake und CloudPond. Dies deckt zentrale Regionen, Hotspot-Gebiete und Edge-Standorte ab und sorgt für eine konsistente Huawei-Cloud-Erfahrung, wo auch immer sich das Unternehmen des Kunden befindet.

Aufbau einer benutzerfreundlichen, effektiven und offenen Plattform für die Entwicklung und den Betrieb von Agenten

Huawei Cloud hat Versatile auf den Markt gebracht, eine Agentenplattform der Unternehmensklasse. Es soll als einfach zu bedienende, effektive und offene Plattform für die Entwicklung und den Betrieb von KI-Agenten dienen. Mit dieser Plattform werden Kunden besser in der Lage sein, schnell KI-Agenten zu entwickeln, die ihren Anwendungsszenarien entsprechen.

Auf der Grundlage von Versatile müssen die Benutzer lediglich Dokumente zur Geschäftsbeschreibung und Flussdiagramme vorbereiten und eingeben. Nach einer einfachen Bestätigung kann der Agent in zwei Schritten generiert werden, was die Effizienz der Generierung erheblich steigert.

Zusätzlich zu den Hauptvorträgen auf der Veranstaltung bot Huawei Cloud eine Reihe von Veranstaltungen an, wie z. B. Gipfelforen und Rundtischgespräche. Huawei Cloud arbeitet auch mit Kunden und Partnern zusammen, um eine breite Palette innovativer Technologien und Verfahren in Bereichen wie Cloud-Infrastruktur, große Modelle, Datenbanken, KI-Agenten und verkörperte KI zu präsentieren und zu demonstrieren, wie Technologie die digitale und intelligente Transformation von Branchen erleichtern kann.

Foto –

https://mma.prnewswire.com/media/2777251/Zhang_Ping_an_Huawei_s_Executive_Director_Board_CEO_Huawei_Cloud.jpg

View original content: <https://www.prnewswire.com/news-releases/huawei-cloud-forderung-des-fruchtbaren-bodens-fur-datenverarbeitung-starkung-von-ki-pionieren-fur-die-industrie-302562279.html>

Pressekontakt:

Liang Xu,
cyril.xuliang@huawei.com

Diese Meldung kann unter <https://www.presseportal.ch/de/pm/100090258/100935295> abgerufen werden.