

IBM stellt On-Chip-beschleunigten Prozessor für künstliche Intelligenz vor

Die nächste im IBM Telum-Prozessor verbaute Chip-Generation ermöglicht die Nutzung von künstlicher Intelligenz in Transaktionen und verbessert damit z.B. die Möglichkeiten zur Betrugsbekämpfung erheblich

Ehningen/ARMONK, N.Y., 23. August 2021 -- Auf der jährlichen Hot Chips-Konferenz hat IBM heute Details zum neuen [IBM Telum-Prozessor](#) vorgestellt. Telum mit seiner On-Chip-Beschleunigung für KI ist branchenweit der erste Prozessor, der Kunden – basierend auf Deep Learning – die Nutzung von KI-Inferenz innerhalb zeitkritischer Transaktionen erlaubt. Telum wartet im Vergleich zu anderen führenden Serverprozessoren mit der zehnfachen Rechenleistung pro Core für KI-Aufgaben auf und ermöglicht so große Fortschritte in der Betrugsbekämpfung, bei Entscheidungen über Kreditvergaben, in der Schadensregulierung und im Finanzhandel.

Telum wurde in den letzten drei Jahren entwickelt und ist ein großer Fortschritt, um KI in Echtzeit einzusetzen, basierend auf Hunderten von Patenten. Ein IBM System, das die Vorteile von Telum bietet, wird voraussichtlich in der ersten Hälfte des Jahres 2022 verfügbar sein.

In einer kürzlich von IBM in Zusammenarbeit mit Morning Consult in Auftrag gegebenen Studie gaben 90 % der Befragten an, dass die Nutzung von Transaktionsdaten sowie die Möglichkeit, KI-Projekte dort zu erstellen und durchzuführen, wo sich diese Daten befinden, entscheidend sind, um Echtzeit-Einblicke in Geschäftsprozesse zu erhalten. IBM Telum ist so konzipiert, dass Anwendungen effizient dort ausgeführt

werden können, wo sich die Daten befinden. Zudem hilft der Prozessor dabei, herkömmliche KI-Ansätze in Unternehmen zu überwinden, die in der Regel erhebliche Speicher- und Datenverschiebungskapazitäten erfordern, um Inferenzen zu verarbeiten. Mit Telum befindet sich der Beschleuniger in unmittelbarer Nähe zu geschäftskritischen Daten und Anwendungen, was bedeutet, dass Unternehmen ein hohes Inferenzvolumen für sensible Echtzeit-Transaktionen durchführen können, ohne auf KI-Lösungen außerhalb der Plattform zurückgreifen zu müssen, was die Leistung beeinträchtigen kann. Kunden können auch KI-Modelle außerhalb der Plattform erstellen und trainieren und auf einem Telum-fähigen IBM-System zur Analyse bereitstellen und inferieren.

Im „[2020 Consumer Sentinel Network Databook](#)“ der [Federal Trade Commission](#) gaben Verbraucher an, im Jahr 2020 mehr als 3,3 Milliarden Dollar durch Betrug verloren zu haben, gegenüber 1,8 Milliarden Dollar im Jahr 2019.

Der neue Chip zeichnet sich durch ein einzigartiges, zentralisiertes Design aus. Dadurch können Kunden die volle Leistung des KI-Prozessors für KI-spezifische Workloads nutzen. Das prädestiniert ihn für zahlreiche Aufgaben im Finanzdienstleistungsbereich, wie etwa Betrugserkennung, Kreditverarbeitung, Clearing, Geschäftsabwicklung, Geldwäschebekämpfung und Risikoanalyse: KI-Algorithmen inklusive Machine Learning und Deep Learning versetzen Kunden in die Lage, die bestehende Betrugserkennung weiterzuentwickeln, Kreditgenehmigungsprozesse zu beschleunigen, den Kundenservice und die Rentabilität zu verbessern, vorherzusagen, welche Geschäfte oder Transaktionen fehlschlagen werden, und Lösungen vorzuschlagen, um Abwicklungsprozesse effizienter zu gestalten. Darüber hinaus können Kunden KI-Modelle auch überall erstellen und trainieren.

Telum und IBMs Full-Stack-Ansatz für das Chipdesign

Telum steht in der langjährigen Tradition von IBM, einen einzigartigen, ganzheitlichen Ansatz für das Chipdesign zu verfolgen, beginnend auf der Siliziumebene über Hardware, Firmware, Betriebssystem bis hin zur Middleware optimiert für die kritischsten Workloads unserer Kunden. Durch sein neuartiges Design unterstützt der in Telum verbaute KI-Beschleuniger eine Vielzahl von Modelltypen und erreicht dabei für alle eine führende Leistungseffizienz.

Mit Telum können Kunden Zehntausende von KI-integrierten Transaktionen pro Sekunde verarbeiten. Der Chip enthält acht Prozessorkerne mit einer tiefen superskalaren Out-of-Order-Befehlspipeline, läuft mit einer Taktfrequenz von mehr als 5 GHz und ist für die Anforderungen heterogener Workloads unserer Kunden optimiert. Die völlig neu gestaltete Cache- und Chip-Verbindungsinfrastruktur bietet 32 MB Cache pro Kern und ermöglicht den Aufbau von Systemen mit bis zu 256 physischen Kernen. Der Chip enthält 22 Milliarden Transistoren und 30km an Leiterbahnen verteilt auf 17 Metallschichten.

Führend bei Halbleitern

Telum ist der erste IBM-Chip, der mit einer vom [IBM Research AI Hardware Center entwickelten Technologie](#) arbeitet. Darüber hinaus ist Samsung der Technologieentwicklungspartner von IBM für den Telum-Prozessor, der im 7-nm-EUV-Technologieknoten entwickelt wurde.

Telum ist ein weiteres Beispiel für die führende Rolle von IBM in der Hardwaretechnologie. IBM Research, eine der weltweit größten industriellen Forschungsorganisationen, [kündigte vor kurzem](#) die Skalierung auf den 2-nm-Knoten an, den neuesten Meilenstein in IBMs Erbe an Beiträgen zu Silizium- und Halbleiterinnovationen. In Albany, New York, wo sich das IBM AI Hardware Center und der Albany Nanotech Complex befinden, hat IBM Research das führende Ökosystem für die Zusammenarbeit mit öffentlichen und privaten Akteuren aus der Halbleiter-Industrie geschaffen, um die Fortschritte in der Forschung voranzutreiben und dazu beizutragen, den globalen Fertigungsbedarf zu decken und das Wachstum der Chipindustrie zu beschleunigen.

###

Für weitere Informationen besuchen Sie bitte www.ibm.com/it-infrastructure/z/capabilities/real-time-analytics

Aussagen über die künftige Ausrichtung und Absicht von IBM können ohne Vorankündigung geändert oder zurückgezogen werden und stellen lediglich Ziele und Absichten dar.

Über IBM

IBM ist ein weltweit führender Anbieter von hybriden Cloud- und KI-Lösungen sowie Geschäftsdienstleistungen, der Kunden in mehr als 175 Ländern dabei hilft, Erkenntnisse aus ihren Daten zu gewinnen, Geschäftsprozesse zu optimieren, Kosten zu senken und Wettbewerbsvorteile in ihren Branchen zu erzielen. Fast 3.000 Regierungs- und Unternehmenseinheiten in kritischen Infrastrukturbereichen wie Finanzdienstleistungen, Telekommunikation und Gesundheitswesen verlassen sich auf die IBM Hybrid-Cloud-Plattform und Red Hat OpenShift, um ihre digitale Transformation schnell, effizient und sicher zu gestalten. Die bahnbrechenden Innovationen von IBM in den Bereichen KI, Quantencomputing, branchenspezifische Cloud-Lösungen und Business Services bieten unseren Kunden offene und flexible Optionen. All dies wird durch das legendäre Engagement von IBM für Vertrauen, Transparenz, Verantwortung, Inklusivität und Service unterstützt. Weitere Informationen finden Sie unter www.ibm.com.

###

Weitere Informationen für Journalisten:

Svetlana Stavreva

IBM Unternehmenskommunikation

Email: stavreva@at.ibm.com

Mobil: +43 664 618 6851
