

Neue Computer-Architektur: Zeitraffer für die Demenz-Forschung

Bonn, 28. März 2018 - Das Deutsche Zentrum für Neurodegenerative Erkrankungen (DZNE) startet in Bonn gerade einen neuen Hochleistungsrechner. Dieser soll die Auswertung biomedizinischer Daten enorm beschleunigen und zu schnelleren Fortschritten in der Demenz-Forschung führen. Dazu nutzt der Rechner die Prinzipien der neuartigen Computer-Architektur „Memory-Driven Computing“.

Die Zeit drängt: Denn Demenz-Erkrankungen wie Alzheimer zählen bereits heute zu den größten Herausforderungen in der Medizin. Das Problem wird sich durch die wachsende ältere Bevölkerung noch verstärken. Neue Ansätze für die Prävention und Therapie könnten sich aus der Analyse von Gendaten und von Hirnscans ergeben. Doch deren Auswertung erfordert enorme Rechenpower. Deshalb haben das DZNE und Hewlett Packard Enterprise (HPE) eine Kooperation gestartet, um das Potenzial des „Memory-Driven Computing“ für die medizinische Forschung zu verwirklichen. Das DZNE ist weltweit das erste Institut, das diese radikal neue Computer-Architektur in der biomedizinischen Forschung einsetzt.

Nach einer gemeinsamen Machbarkeitsstudie folgt nun der nächste Schritt: Gerade wurde ein neuer Hochleistungsrechner vom Typ „HPE Superdome Flex“ im Bonner DZNE-Rechenzentrum in Betrieb genommen. Bisher hatten die Wissenschaftler des DZNE ihre Algorithmen auf Rechnern von HPE in den USA getestet. Bereits dort gelang es, die Zeit eines Rechenprozesses für die Auswertung von Gendaten von 22 Minuten auf 13 Sekunden zu verkürzen. Nun steht den Bonner Forschern ein eigenes System zur Verfügung.

Eine völlig neue Architektur

Das von HPE entwickelte Memory-Driven Computing bricht radikal mit der Tradition aller bisherigen Computer, indem es nicht den Prozessor, sondern den Arbeitsspeicher ins Zentrum der Architektur stellt: „Idealerweise liegen sämtliche Daten gleichzeitig im riesigen Arbeitsspeicher vor“, so Prof. Joachim Schultze, Genomforscher und Arbeitsgruppenleiter am DZNE. „Sie müssen also nicht erst aus externen Speichermedien eingelesen werden, sondern der Prozessor kann unmittelbar darauf zugreifen.“ Das spart viel Zeit und Energie. Schultze weiter: „Die Arbeit mit externen Speichermedien gleicht einem Puzzlespieler, der die einzelnen Puzzleteile in einer großen Menge von Schachteln verteilt hat, die er erst nacheinander öffnen müsste, um nach passenden Teilen zu suchen. Hat man dagegen alle Teile vor sich ausgebreitet, gelangt man wesentlich schneller ans Ziel.“ Ähnlich sei das beim Memory-Driven Computing, so Schultze.

Der Superdome-Flex-Rechner wurde von HPE auf der Grundlage der Memory-Driven-Computing-Architektur entwickelt. Außerdem ist die Hardware auf extrem schnellen Datenaustausch ausgelegt. Damit ist es möglich, den DZNE-Forschern das Memory-Driven Computing fachübergreifend zur Verfügung zu stellen. „Derzeit arbeiten wir an den Spezifikationen der Datenverarbeitung. Wir möchten aber möglichst bald mit der Auswertung von Daten aus Experimenten und Studien beginnen. Unser Ziel ist, Gendaten von Tausenden von Personen innerhalb weniger Minuten zu vergleichen“, betont Schultze.

Denn ein Kernproblem der Demenzforschung sind die riesigen Datenmengen. Das gilt beispielsweise für die Analyse von Abschnitten des Erbguts, von denen vermutet wird, dass sie für eine Erkrankung von Bedeutung sind. Solche Sequenzen können hunderte Millionen genetischer Bausteine umfassen. Deren Auswertung und Abgleich mit Referenzdaten kann mit konventionellen Hochleistungsrechnern einige Wochen beanspruchen. Mit Hilfe des Rechners, der nun in Bonn zur Verfügung steht, wollen die Wissenschaftler des DZNE ihre Auswertungs-Algorithmen erheblich optimieren.

Ein weiteres Anwendungsbeispiel sind Studien mit tausenden Probanden. Dabei müssen die erhobenen Werte immer wieder mit Referenzdaten anderer Studienteilnehmer verglichen werden. Schon bei einem einzigen Probanden können dabei hunderte Gigabytes an Daten zusammenkommen – also hunderttausende Millionen von digitalen Informationseinheiten. Dies gilt insbesondere, wenn verschiedene Arten von Daten miteinander verknüpft werden, etwa Daten aus einer Genomanalyse mit denen von Hirnscans. Und diese Datenpakete müssen mit einer jeweils ebenso riesigen Datenmenge tausender anderer Probanden verglichen werden, um beispielsweise Alzheimer-spezifische Veränderungen zu finden. Die Aufgabe, in der Größenordnung von Petabytes (Millionen mal Milliarden) an Daten nach Symptomen zu suchen, ist nicht sequenziell lösbar. Vielmehr müssen die DZNE-Forscher möglichst alle Daten gleichzeitig im unmittelbaren Zugriff, also im Arbeitsspeicher, haben.

10.000mal schneller

HPE antwortet mit seinem neuen Ansatz auf die Tatsache, dass eine weitere Steigerung der Prozessorleistung an physikalische Grenzen stößt. Dies ist ein großes Problem angesichts der exponentiell steigenden Datenmengen, die Computer zu verarbeiten haben. Schon jetzt hat das Unternehmen jedoch nachgewiesen, dass sich Berechnung mittels Memory-Driven Computing viele tausend Mal beschleunigen lassen. Seit Mai 2017 betreibt HPE einen Prototyp in seinem Labor in Fort Collins (USA) mit 160 Terabyte Arbeitsspeicher – dem größten einheitliche Arbeitsspeicher in einem Rechner weltweit. Die Architektur erlaubt grundsätzlich eine Skalierung des Arbeitsspeichers auf 4.096 Yottabytes. Zum Vergleich: Dies ist das 250.000-fache der Datenmenge, die auf der Erde heute vorliegt.