

Hintergrund

HPI-Spitzenforschung revolutioniert Hochleistungs-Rechenverfahren

Juni 2010

Potsdam. Ähnlich wie vor knapp 20 Jahren, als die Großcomputer fast flächendeckend durch verteilte Klient-/Server-Systeme abgelöst wurden, erlebt die Informatik derzeit nun **gleich zwei revolutionäre Veränderungen** auf dem Gebiet der Rechnerarchitektur – sowohl im Wachstum der Rechenleistung als auch in der Art der schnellen Verarbeitung riesiger Datenmengen.

- Da aufgrund physikalischer Grenzen die zentrale Verarbeitungseinheit eines Computers (CPU) durch Erhöhen der Taktfrequenz kaum noch schneller zu machen ist, wird die Leistungsfähigkeit von Rechnern dadurch vervielfacht, dass **Mehrkernprozessoren** zum Einsatz kommen. Das sind Mikrochips, die in einer zentralen Verarbeitungseinheit mehrere Rechenkerne (MultiCore) integrieren. Die volle Rechenleistung dieser mehrkernigen Prozessoren kommt allerdings nur dann zum Tragen, wenn mehrere Befehlsströme gleichzeitig verarbeitet werden.
- Hinzu kommt, dass die Speichersysteme moderner Rechner über **immense Hauptspeicher** verfügen, die zudem durch Halbleiterlaufwerke (SSD – Solid State Drive) ergänzt werden, so dass umfangreiche Datenmengen komplett in den Speicher solcher Rechner geladen werden können (In-Memory-Technologie). Gerade der Zugriff auf externe Speichermedien (über das Ein- und Ausgabesystem) stellte bislang einen Flaschenhals in Rechnern dar und limitierte die Verarbeitungsgeschwindigkeit.

Beide Entwicklungen führen dazu, dass im Bereich von Unternehmenssoftware schon bald die meisten **Verarbeitungsschritte um mehrere Größenordnungen (Faktor 10-100) beschleunigt** werden können. Wichtige Beiträge dazu leisten auch eine **veränderte Art der Organisation von Datenbanken**, die **Virtualisierung** - also Methoden zur Zusammenfassung oder Aufteilung von Computer-Ressourcen - und das Cloud Computing. Beim **Cloud Computing** rufen Anwender akut benötigte Rechenleistung aus externen Rechenzentren ab, die über das Internet verfügbar sind.

Unternehmenssoftware so blitzschnell wie das iPhone?

Mit diesen Zusammenhängen beschäftigen sich die Forscher am Potsdamer Hasso-Plattner-Institut. Ihre Ambition: Wie der Verbraucher bei seinem Smartphone sofort nach dem Starten einer Anwendung erlebt, dass sich etwas tut, so sollen auch bei den Großrechnern der Unternehmen die Antwortzeiten im Bereich weniger Sekunden liegen. Detaillierte betriebswirtschaftliche Analysen, die bisher nur mit stunden- oder tagelanger Verzögerung verfügbar gemacht werden konnten, stehen künftig auf einen Mausklick hin blitzschnell bereit – gewissermaßen in echter Echtzeit. Das dürfte die Strategieplanung und die Prozesse in den Führungszentralen der Konzerne revolutionieren.

Wie die Potenziale neuester Rechner-Architekturen mit vielen Prozessor-Kernen künftig im Zusammenspiel mit massiv parallel arbeitender Software genutzt werden können, wird im neuen **Spitzenforschungslabor des HPI** untersucht, das am 16. Juni 2010 seine Arbeit aufnimmt. Auf modernster Hard- und Software, die gerade aus den Entwicklungsabteilungen der IT-Konzerne kommt, erforschen die Informatiker des HPI und eingeladene Experten anderer akademischer Einrichtungen realitätsnah neue Konzepte, die für das **Service-Oriented Computing (SOC)** der Zukunft wichtig sind. Dabei geht es unter anderem darum, dass Software-Anwendungen nicht mehr im firmeneigenen Rechenzentrum oder vom Endbenutzer selbst, sondern von Anbietern entsprechender Dienste (Services) ausgeführt würden.

Das Thema Service-Oriented Systems Engineering vereint auch die Doktoranden der HPI Research School. Im internationalen Graduiertenkolleg des Instituts forschen derzeit rund 40 Nachwuchswissenschaftler aus Potsdam, Haifa (Technion) und Kapstadt (University of Cape Town). Auf seinem jährlich stattfindenden Symposium beschäftigt sich das Forschungskolleg vom 16. bis 18. Juni nun schon zum fünften Mal mit künftigen Trends im Service-Oriented Computing. Dabei werden führende Wissenschaftler und Industrievertreter unter anderem auch die Forschungsansätze präsentieren, die im neuen Potsdamer Spitzenforschungslabor des HPI untersucht werden sollen. Zum Beispiel wird es dort auch um die **Hochverfügbarkeit und Zuverlässigkeit künftiger Industriestandardserver** gehen. Dabei wird etwa erforscht: Was passiert bei Prozessorausfällen? Wie toleriert die Software solche Hardware-Probleme? Wie können verschiedenen Nutzern vorab vereinbarte Teile der Rechenleistung garantiert zur Verfügung gestellt werden?

Modernes Labor mit neuester Hard- und Software

Das HPI-Spitzenforschungslabor ist mit neuesten Serversystemen ausgestattet. Jedes dieser Systeme integriert vier bis acht Mehrkernprozessoren mit bis zu acht Prozessorkernen. Von denen lässt sich zudem jeder in zwei logische Prozessorkerne aufteilen. Die **bis zu 128**

logischen Rechenkerne – jeder davon einer herkömmlichen CPU vergleichbar – können gemeinsam auf bis zu zwei Terabyte Hauptspeicher zugreifen. Mehrere dieser Server können dann zu Clustern zusammengeschaltet werden, so dass die Leistungsfähigkeit im Prinzip beliebig vervielfältigt werden kann. Darauf muss sich Software allerdings optimal einstellen.

Neben der reinen Rechenleistung ist die Geschwindigkeit des Ein- und Ausgabesystems (Input-/Output-System) ein limitierender Faktor in Serversystemen. Mit wachsender Zahl der Rechenkerne wird es immer schwieriger, allen Verarbeitungseinheiten gleich schnellen Zugriff auf alle Daten zu ermöglichen. Wachsende Hauptspeichergrößen sowie neue Halbleiterlaufwerke (SSD – Solid State Disk) helfen, viele bislang benötigte Zugriffe auf das Input-/Output-System zu vermeiden. Werden zudem Anwendungsdaten passend reorganisiert, so wird der schnelle parallele Zugriff auf diese Daten durch viele Prozessorkerne möglich.

Was in dieser Richtung machbar ist, haben Studenten und Wissenschaftler am HPI bereits aufgezeigt. Sie drückten zum Beispiel die für einen Mahnlauf benötigte Zeit bei einem Großunternehmen mit 100.000 Kunden auf eine Minute. Vorher hatte die Beschaffung der Daten aller Kunden, die mit ihrer Zahlung im Verzug sind, rund 20 Minuten gedauert. Weitergehende Untersuchungen der Experten des HPI-Fachgebiets Enterprise Platform and Integration Concepts deuten an, dass die Laufzeit auf nur noch fünf Sekunden sinkt, wenn bestimmte Engpässe in der Speicherverwaltung beseitigt werden können. Und auf massiv parallel arbeitender Hardware konnte ein solcher Vorgang sogar in gut einer Sekunde abgewickelt werden. Das ist eine Beschleunigung um rund das 1000-fache.

Forschung mit riesigen Mengen echter Daten aus der Wirtschaft

Geforscht wird am HPI mit echten, sehr umfangreichen und hoch komplexen Datensätzen aus Unternehmen. Genau wie Mediziner, die sich mit echten klinischen Fällen beschäftigen und wie Ingenieure, die mit realen Teststrecken oder Testverfahren arbeiten, begnügen sich die Software-Spezialisten am HPI nicht mit frei erfundenen Geschäftsfällen und konstruierten Zufallsdaten im Kleinmaßstab. Ihnen stehen vielmehr **300 Millionen echte Vertriebsdatensätze** aus dem Bereich der Konsumgüterindustrie zur Verfügung. Sind sie erst einmal in den Hauptspeicher geladen, können die HPI-Forscher heute diese 300 Millionen Sätze in 700 Millisekunden durchsuchen. Und mit der nächsten Version des unterliegenden Datenbanksystems wird diese Operation in weniger als 100 Millisekunden möglich sein.

Selbstverständlich wurden die Daten für die Zwecke der HPI-Forscher anonymisiert, sodass man keine Rückschlüsse auf den Konzern und sein Geschäft ziehen kann. Produktnamen und auch Werte wurden dazu leicht

verändert. Aber die Verteilungen und die Ausprägungen sind so nah wie möglich an der Realität geblieben. Und alles ist am HPI eingeschlossen und nicht öffentlich zugänglich. Ferner müssen von den Beteiligten für bestimmte Forschungsarbeiten auch Geheimhaltungsvereinbarungen unterschrieben werden.

Pressekontakt HPI:

Hans-Joachim Allgaier, AllgaierCommunication, Tel.: 0331 55 09-119,
06081 57 76 30, Mobil: 0179 267 54 66, Fax: 06081 96 25 17,
Mail: hansjoachim.allgaier@hpi.uni-potsdam.de,
info@allgaiercommunication.de

Kontakt für Fotos, Illustrationen und Logos:

Katrin Augustin, Hasso-Plattner-Institut, Tel.: 0331 55 09-150,
Fax: 0331 55 09-169, Mail: katrin.augustin@hpi.uni-potsdam.de