



Presseinformation

TU Dresden erhält landesweit stärkste HPC-Umgebung für „Datenintensives Rechnen“

Technische Universität investiert 15 Mio in neuartige Infrastruktur für Wissenschaftliches Computing - Generalunternehmer SGI installiert großes Shared-Memory-System Altix mit 6 TeraByte Hauptspeicher

Dresden, 14 Juli 2005 – Die Technische Universität Dresden wird sich im Laufe der nächsten 15 Monate mit einer erheblichen Erweiterung ihrer Rechnerkapazitäten rüsten, was der landesweiten Gemeinde des Wissenschaftlichen Computings zu Gute kommen und ganz neue Möglichkeiten für datenintensive Anwendungen schaffen soll. Im Rahmen einer feierlichen Vertragsunterzeichnung in Gegenwart der Landesregierung Sachsen hat die TU Dresden heute offiziell die Firma SGI beauftragt, als Generalunternehmer eine zukunftsorientierte Hochleistungs-Umgebung zu installieren, welche eine Investition von 15 Mio darstellen und die TU in der Landschaft der Rechenzentren differenzieren wird.

Am Campus, im „Zentrum für Informationsdienste und Hochleistungsrechnen“ (ZIH), wird SGI in 2 Projektstufen eine ultramoderne, vielseitig nutzbare Infrastruktur mit 1000en von Prozessorkernen implementieren - die dann Sachsens leistungsfähigste Rechnerinfrastruktur bilden und Anwender von Physik, Materialforschung und Engineering bis hin zu Bioinformatik und Nanotechnologie unterstützen soll.

Die neue Umgebung wird in Form eines HPC-Komplexes mit fortschrittlichster Shared-Memory-Technologie auf Basis eines SGI@Altix®-Systems einzigartige Möglichkeiten des „Capability-Computing“ schaffen und das Lösen extrem anspruchsvoller Problemstellungen erlauben. Ergänzend integriert SGI eine PC-Farm, die als Plattform für heterogenes Computing dienen und gleichzeitig 100e von einzelnen Anwendungen unterstützen wird. Über die beiden Komponenten ist das ZIH auch in der Lage, Software für unterschiedliche Prozessor-Typen zu unterstützen: Der HPC-Server SGI Altix stellt neueste Intel®-Itanium®-Power, die PC-Farm, geliefert von Hersteller LinuxNetworx, aktuelle AMD-Opteron-Leistung zur Verfügung.

Die Neuanschaffung - sie trägt den Titel "Hochleistungsrechner/Speicherkomplex 'Datenintensives Rechnen'" - wird einen Ressourcenkomplex bilden, der neben hoher Rechnerleistung gezielt auch Top-Bandbreiten-Fähigkeiten bietet, so dass durchgängiges, beispielhaft schnelles Bewegen und Speichern großer und größter Datenmengen bis hin zur Archivierung möglich ist.

Die Beschaffung ist eine der größten, die in diesem Jahr in Europa im HPC-Umfeld vergeben wird. Das Projekt wird im Rahmen eines HFBG-Verfahren realisiert.

Prof. Hermann Kokenge, Rektor der TU: "Wir freuen uns auf eine ausgewogene und mit exzellenten Komponenten ausgerüstete Rechner- und Speicher-Umgebung,



welche die Innovationskraft der Universität, des Standorts Dresden und der ganzen Region wirkungsvoll stärken dürfte. Mit ihr können wir unseren Instituten und der weiteren akademischen Gemeinde eine „kritische Masse“ mehr an „Power“ und neuartigen Arbeitsmöglichkeiten bieten - und sie damit in etlichen traditionell repräsentierten wie auch aufstrebenden neuen Forschungsgebieten beim Erringen vieler wertvoller und bahnbrechender Erkenntnisse unterstützen."

Prof. Wolfgang E. Nagel, Direktor des ZIH: „Wesentlich für eine solche Großanschaffung ist, dass wir ein hohes Maß an Flexibilität erhalten. Flexibilität in puncto vielseitiger, variabler und innovativer Nutzbarkeit, Flexibilität auch hinsichtlich zukünftiger Veränderungen. Wir haben uns zum Ziel gesetzt, mutige, hochaufwändige Projekte zu ermöglichen, zugleich aber auch die Interessen und wachsenden Anforderungen 100er anderer Anwender mit kleineren Projekten zu unterstützen. Die Dynamik, mit der sich heute Forschungslandschaften entwickeln, lässt nicht genau vorhersehen, welche Themen und Methoden in einigen Jahren die Akzente setzen und Unterstützung verlangen. Mit der zukunfts-offenen Umgebung, die wir bei SGI in Auftrag geben, sehen wir uns optimal abgesichert, um auch mittelfristig veränderliche Herausforderungen abdecken zu können.“

Fragen, denen die Wissenschaftler mit Hilfe der neuen TUD-Rechnerumgebung beispielsweise in Zukunft intensiv nachgehen können:

- Wie lassen sich organische Stoffe finden, die bei knochen-chirurgischen Eingriffen metallische Legierungen ersetzen, jedoch weniger Unverträglichkeiten hervorrufen? - ein Problem der Materialforschung mit Analyse komplexer biomolekularer Reaktionen.
- Wie lassen sich neuartige und perfekter gewachsene Kristalle züchten? - wenn man sie als elektrisch leitfähige Fluide unter Einfluss von Magnetfeldern betrachtet und mit Methoden der Computational Fluid Dynamics (CFD) turbulente Strömungen des Fluids analysieren kann.
- Wie lassen sich Störgeräusche im Innenraum eines Autos beeinflussen? - ein ausgesprochen komplexes Akustik-Problem, das beispielhaft für die strategische Bedeutung von CFD an der TUD steht.
- Wie lassen sich mit der Automatischen Mikroskopie Zellwachstumsprozesse verfolgen und verstehen? - ein Problem, bei dem es in 1000en von kurzfristig anfallenden Filmen Strukturen zu erkennen und Zustände zu beschreiben gilt.
- Wie lassen sich beim Protein-Docking die chemischen Reaktionen besser verstehen? - wenn man mit Rechnerhilfe versucht, die 3D-Struktur (Faltung) der äußerst komplexen Moleküle zu modellieren.
- Wie lassen sich genetische Ursachen von Krankheiten verstehen und beeinflussen? - wenn man mit neuen rechnergestützten Verfahren jene enormen Datenmengen analysieren kann, welche Biologen und Mediziner bei ihren Versuchen, die Aktivität von Genen zu messen und auszuschalten, generieren. --- (Details im Anhang)



Pressestelle

HPC-Komplex - Geballte Ressource für Einzelne

Den HPC-Komplex wird in Dresden eine Shared-Memory-Plattform SGI® Altix® der nächsten Generation bilden. Das System (Codename „Tollhouse“) wird in seiner letztlichen Ausbaustufe mit weit mehr als 1000 der aktuellsten Intel®-Itanium®-2-Prozessorkernen sowie mit 6 TeraByte Hauptspeicher ausgestattet sein. Hunderte von parallel arbeitenden Prozessorkernen können dann das Memory als eine einzige große, zusammenhängend adressierbare Systemressource nutzen, in den riesigen Hauptspeicher enorme Datensätze laden und diese in unzerstückelter Form effizient bearbeiten oder auf Muster und Ähnlichkeiten untersuchen.

Die SGI-Plattform wird im ZIH als „Capability-Computing“-Maschine fungieren und die Ressource zur Deckung des „homogenen“ Computing-Bedarfs bilden: Die geballte Verarbeitungs- und Speicherkapazität des SGI-Systems (oder zumindest eines 2/3-Teils) soll zu *einem* Zeitpunkt *einer* einzelnen Anwendung bzw *einer* einzelnen Benutzergruppe zur Verfügung stehen. Wissenschaftler werden damit in der Lage sein, erstmals bestimmte, äußerst umfangreiche, hochkomplexe Aufgaben anzugreifen.

Die HPC-Ressource liefert ihnen ein Maximum an Performanz. Gleichgültig, ob ihr hochgradig parallelisierter Code für eine Shared-Memory-Architektur oder (nach dem MPI-Programmier-Paradigma) für eine Distributed-Memory-Plattform geschrieben ist - die SGI-Plattform mit ihrer extrem skalierenden ccNUMA-Architektur und der bandbreiten-starken *NUMAlink*-Interconnect-Technologie hält hochschnelles Processing für beide Anwendungsarten bereit.

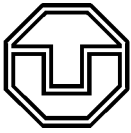
„In-Memory“-Computing ist nur eine der Innovationen, welche das ZIH Wissenschaftlern und Forschern über die HPC-Plattform bieten kann: Erstmals wird z.B. die Möglichkeit bestehen, mehrere komplette wissenschaftliche Datenbanken in den Hauptspeicher zu laden und sie dort von der Anwendung mit unvergleichlich hoher Geschwindigkeit nach bestimmten Korrelationen durchsuchen zu lassen. Was solche Aufgaben bisher extrem bis unzumutbar langwierig gestaltet, entfällt – „In-Memory“-Computing kommt ohne die üblich vielen Zeit verschlingenden Festplattenzugriffe (I/O-Prozesse) aus.

Neues Instrument für neue Herausforderungen

„Wir wollten nicht ein weiteres Rechenzentrum einrichten, das im wesentlichen nur Numbercrunching unterstützt“, erklärt Nagel. „Unser Fokus ist, der Wissenschaftlichen Computing-Gemeinde ein neuartiges Instrument, einen neuen Typ von HPC-Tool zur Verfügung zu stellen. Nicht die üblichen Simulationsprobleme stehen hier im Vordergrund, es geht uns mehr darum, den Anwendern eine Plattform zu liefern, mit der sie aus Unmengen von strukturierten oder unstrukturierten Daten, welche viel verstecktes Wissen in sich tragen, neues Wissen, mehr Wissen, dichteres Wissen extrahieren und herausarbeiten können.“

HPC-SAN - Archivierung - Daten-Management

Wesentliche weitere Komponente ist eine mehrstufige, intelligent organisierte



Pressestelle

Massenspeicherung mit extrem schnellem Zugriff auf große Datenmengen. Hierzu installiert SGI ein HPC-SAN (Storage Area Network), das mit einer Bandbreite von 8 GB/s an das Altix-System angeschlossen ist, 60 TeraByte Online-Diskspeicher umfasst und mit ebenfalls hoher Bandbreite ein PetaByte großes Archiv-Bandroboter-System anbindet.

Der 60 TB Online-Festplattenspeicher wird mit einem *SGI® InfiniteStorage SAN* realisiert, mit FibreChannel-Disk-Array-Systemen des Hersteller Data Direct Network (DDN); als Archiv-Roboter wird ein PetaByte-Tape-Library-System des Herstellers Storage Technology Corp dienen. Für die durchgängige hierarchische Speicherverwaltung mit Daten-Lifecycle-Management und Datensicherung/-Wiederherstellung sorgt die Software *SGI® InfiniteStorage DMF (Data Migration Facility)*. Als Dateisystem kommt das technologieführende Shared-Filesystem *SGI® InfiniteStorage CXFS* zum Einsatz, das ein highschnelles, gleichzeitiges Zugreifen vieler Anwendungen oder Prozesse auf dieselben Daten im SAN erlaubt, ohne dass Dateien deshalb kopiert werden müssten. CXFS kann ein Dateisystem mit bis zu 18 Mio TeraByte Dateivolumen und eine Einzeldateigröße von bis zu 9 Mio TeraByte managen.

4 TB Memory kompakt laden - in 10 Minuten

„Hohe Bandbreite für schnelles Bewegen großer Datenmengen ist top wichtig für unseren Capability-Computing-Ansatz mit der neuartigen Umgebung“, sagt ZIH-Direktor Nagel. "Für den Datentransfer von den SAN-Disks ins Memory der HPC-Maschine forderten wir, dass 4 TB in nicht mehr als 10 min in den Hauptspeicher ladbar sind. Und beim Transfer zum/vom Archiv-System werden wir 25 TB in nur 4 h bewegen können - das ist herausragend. Zusammen mit dem 6 TB großen Shared-Memory des Altix-Systems wird dies unsere Lösung am ZIH ganz wesentlich differenzieren.“

Dank der bandbreitenstarken, durchgängigen Anbindung des HPC-Systems an die Massenspeicherung wird das ZIH in der Lage sein, den HPC-Komplex in kürzester Zeit einem nächsten Nutzer oder nächsten Projekt zur Verfügung zu stellen - d. h. die für das Projekt benötigten Daten sehr schnell vom Archiv über das SAN in den Altix-Hauptspeicher laden, den Anwender exklusiv die Aufgabe bearbeiten lassen, und nach Beendigung des Jobs den Datensatz löschen, oder, falls sich die Daten verändert haben, das Ergebnis wieder zurück ins Archiv schreiben.

Ein Beschleuniger - ähnlich wie am CERN

"Auf diesem Wege kann das ZIH ein wirklich neuartiges Instrument für das Datenintensive wissenschaftliche Rechnen bereit stellen," erklärt Prof. Nagel. „Ganz ähnlich, wie die Physiker am CERN oder in Stanford seit den frühen Jahren ihrer Hochenergieforschung einen Teilchenbeschleuniger für eine bestimmte Zeit in voller Leistungsstärke alleine nutzen und dann an andere Kollegen übergeben, so wollen wir einen mächtigen Beschleuniger für wissenschaftliches Rechnen liefern. Ausgewählte Benutzer/Projekte sollen in der Lage sein, den HPC-Komplex über eine gewisse Zeit hinweg als eine mächtige Capability-Computing-Ressource exklusiv zu nutzen - und



Pressestelle

damit hochkomplexe Fragestellungen mit innovativen Analyse-Methoden und Algorithmen angehen können. - Mit SGI als Generalunternehmer, Systemintegrator und Lieferant technologieführender Lösungen für bandbreitenstarkes Shared-Memory-Computing sind wir überzeugt, einen solchen „Theorie“-Beschleuniger bald anbieten zu können."

PC-Farm - Ressource für viele gleichzeitige Benutzer

Im Rahmen der Neuanschaffung wird SGI eine PC-Farm integrieren. Der Komplex wird aus über 700 Einzelsystemen bestehen und insgesamt mehrere 1000 Prozessor-Kerne enthalten. Mit der Farm wird das ZIH die heterogenen Anforderungen an seine Umgebung abdecken und eine Vielzahl von Benutzern bei einer Vielfalt von kleineren Jobs unterstützen. Hunderte von Anwendern werden gleichzeitig rechnen können und mit Codes, die für AMD Opteron optimiert sind, auf eine Rechner-Ressource von 2 bis 16 Prozessorkernen pro Einzelsystem zurückgreifen können. Die einzelnen Einheiten sind zudem über InfiniBand Interconnect-Technologie verbunden und lassen sich im Bedarfsfall zu einer größeren Ressource für höher paralleles Distributed-Memory-(MPI)-Processing aggregieren. Lieferant der PC-Farm ist LinuxNetworx.

Auch für die PC-Farm integriert SGI einen SAN-Komplex - ebenfalls mit 60 TB RAID-Storage des Herstellers DDN sowie mit schneller Anbindung sowohl an die Farm wie auch an das HPC-SAN (und damit an das Archiv). Als Shared-Filesystem für den PCFarm-Teil wird *Lustre* fungieren, womit nun auch beim ZIH in Dresden *die* Technologie zum Einsatz, welche sich bei weltweit renommierten großen US-Labs als Standard für Linux-Cluster etabliert hat.

HPC-Komplex wie auch PC-Farm werden beide unter Novell's Betriebssystem SuSE Linux Enterprise Server arbeiten.

Wesentlich für die Effizienz einer solchen neuen großen Rechnerinfrastruktur sind auch Software-Tools. So wird die Firma Allinea Software beispielsweise den DDT-Debugger plattformübergreifend zur Verfügung stellen und damit die Nutzung des Gesamtsystems erleichtern.

Neuer Maschinenraum - für Tonnen von Hardware

Am ZIH wird eigens ein neuer 3-geschossiger Maschinenraum entstehen, um die neue Rechner- und Speichertechnik auf einer Gesamtfläche von 430 m² unterzubringen. Die Masse alleine des SGI-Tollhouse-Systems mit seinen Dutzenden von Racks wird rund 18 to betragen. Die SGI-Plattform wird mit Luft, die PC-Farm mit Wasser gekühlt. An elektrischer Energie wird die Gesamtlösung mehr als 1/2 MegaWatt verbrauchen. Die Fertigstellung des Raums ist für Anfang 2006 geplant.

Lieferung in 2 Stufen

Die Realisierung des Gesamtprojekts wird sich in 2 Stufen vollziehen. Bereits für diesen Herbst ist die Lieferung von Technologie geplant, die etwa 1/3 der Gesamtkapazität abdecken soll. Sie wird dem ZIH in erster Linie als eine „Preparation“-Umgebung dienen und den Benutzer mit ihren vielfältigen Interessen



Pressestelle

erlauben, Anwendungs-Codes und Algorithmen zu entwickeln und zu optimieren und sich für die neuen Möglichkeiten vorzubereiten. In Rahmen von Stufe1 wird zunächst ein SGI-Altix-Bx2-System installiert. Lieferstufe Stufe 2 soll ab dem kommenden Jahr erfolgen und im Sommer 2006 abgeschlossen sein. Wenn dann in rund 15 Monaten die letztlich komplette Installation realisiert ist, wird das Altix-Nachfolge-System „Tollhouse“ die Arbeit als HPC-Maschine übernommen haben; das Next-Generation-System wird den Vorgänger BX2 im Laufe der Stufe2 ersetzen.

Zuschlag nach hartem Wettbewerb

„Es ist dies das 3. Mal in Folge, dass wir hier in Dresden SGI als bevorzugten HPC-Partner wählen, und es ist ein 128-Prozessor-System SGI Origin3800, auf dem unsere derzeitigen HPC-Shared-Memory-Jobs laufen“, erklärt ZIH-Direktor Nagel. „Doch SGI musste sich in einem harten, äußerst herausfordernden Wettbewerb durchsetzen. Wir haben uns für SGI entschieden, weil das Unternehmen ein System mit einem dermaßen großen zusammenhängend nutzbaren Hauptspeicher liefern kann. Das ist ein Alleinstellungsfaktor, mit dem auch wir uns als Dienstleister bei unseren Kunden bei herausfordernden neuen Forschungsaktivitäten abheben können.“

„Wir werden einen hochgradig ausbalancierten und vielseitig nutzbaren Rechner/Speicher-Gesamtkomplex erhalten, der uns mit seinen exzellenten Einzelkomponenten und durchgängig hohen Bandbreiten erlaubt, beim Datenintensiven Wissenschaftlichen Rechnen in homogener wie in heterogener Hinsicht neue, spannende Herausforderungen anzunehmen.“

Anhang

Details zu einigen der herausfordernden neuen Anwendungen/Einsatzbereichen

Das Einsatzgebiet des neuen Systemkomplexes umfasst ein weites Spektrum. Um z. B. die genetischen Ursachen verschiedener Krankheiten zu verstehen, generieren Biologen und Mediziner große Datenmengen mit neuen Verfahren, die es erlauben, die Aktivität von Genen zu messen und Gene auszuschalten. Die Bioinformatikgruppe am Biotec der TU Dresden hilft, diese Daten zu interpretieren. Die Gruppe um Prof. Schroeder entwickelt z.B. eine intelligente biomedizinische Suchmaschine, GoPubMed, die automatisch relevantes Wissen aus der biomedizinischen Literatur extrahiert. Bei über 15 Millionen Artikeln reichen herkömmliche Rechner bei weitem nicht mehr aus.

Andere Probleme aus der Strukturbiologie beschäftigen sich beispielsweise mit dem Threading, das für eine Aminosäuresequenz die Struktur vorhersagt, oder mit dem Docking, das die Wechselwirkung von Proteinen vorhersagt. Beide Verfahren realisieren komplizierte mathematische Berechnungen für eine große Zahl von möglichen Lösungen und sind somit extrem rechenzeitaufwendig. Durch die Neuinstallation werden derartige Untersuchungen am Standort Dresden erst in vernünftigem Maßstab möglich.

Ein weiteres Anwendungsspektrum ist die automatische Analyse von Bildern. Das



Pressestelle

Max-Planck Institut für molekulare Zellbiologie und Genetik (MPI-CBG) unterhält eine einmalige Infrastruktur zur Ausschaltung von Genen im hohen Durchsatz. Per Videomikroskopie werden dabei Tausende von Bildern und Filmen der untersuchten Zellen generiert. Die entstehenden Daten und deren Analysen sind so umfangreich, dass sie mit herkömmlichen Methoden auf normalen Rechnern nicht mehr effektiv zu verarbeiten sind. Auch hier wird mit dem neuen Hochleistungsrechner/Speicherkomplex eine Tür zu neuem Wissen geöffnet werden.

Weitere Informationen für Journalisten: Prof. Wolfgang E. Nagel, Tel. 0351 463-35450, E-Mail: nagel@zhr.tu-dresden.de

Dresden, 14. Juli 2005
Mag – 174ZHRRechner