

VIENNA SCIENTIFIC CLUSTER

Der neue Supercomputer geht in Betrieb

Wie bereits in den letzten comment-Ausgaben¹⁾ berichtet, wird an der TU Wien ein neuer Supercomputer errichtet. Die Universität Wien hat die ursprünglichen Pläne eines Ausbaus bzw. einer Erneuerung des bereits etwas betagten Schrödinger III zugunsten einer Zusammenarbeit mit der TU Wien aufgegeben: Die Universität Wien beteiligt sich zu 50% an der Finanzierung und erhält dafür auch die Hälfte der Ressourcen. Mit dem Vienna Scientific Cluster (VSC) sind die Wiener Universitäten im Bereich High Performance Computing (HPC) nach vielen Jahren wieder international einigermaßen konkurrenzfähig.



Die dicht gepackten Rechenknoten des Vienna Scientific Cluster: Jeder der insgesamt 436 Knoten verfügt über 24 GB Hauptspeicher und eine Festplatte mit 500 GB.

Was bisher geschah

Im Sommer 2008 gab es erste Sondierungen zwischen der Universität Wien und TU Wien über eine Zusammenarbeit im HPC-Bereich, im Dezember wurde die entsprechende Vereinbarung unterzeichnet. Vom ZID der TU Wien wurde – unter Berücksichtigung der Anforderungen der BenutzerInnen der Universität Wien – eine Ausschreibung vorbereitet und am 27. Jänner 2009 veröffentlicht, die Abgabefrist endete am 31. März. Von zehn Anbietern wurden insgesamt 13 Angebote abgegeben. Die Bewertung der Angebote war bis zum 21. April abgeschlossen: Unter etlichen attraktiven Angeboten war das der Wiener Firma IPS²⁾ zwar knapp, aber eindeutig vorne.

Das von IPS angebotene System ist ein Cluster, der hauptsächlich aus Komponenten von Sun Microsystems³⁾ besteht. Neben den üblichen zentralen Servern (Fileserver, Zugangsknoten, Management-Server) besteht der Cluster aus folgenden Komponenten⁴⁾:

- Die 436 Rechenknoten (Compute Nodes) haben je zwei Quadcore-Prozessoren (Intel X5550 „Nehalem“) mit einer Taktfrequenz von 2,66 GHz. Die Nehalem-Prozessoren, die erst seit wenigen Monaten verfügbar sind, liefern vor allem für HPC-Anwendungen eine wesentlich höhere Rechenleistung als ihre Vorgängermodelle: Auf Nehalem-Rechnern können auf acht Cores acht Prozesse fast die volle Rechenleistung ausnutzen, ohne sich gegenseitig zu behindern. Jeder Rechenknoten hat 24 GB Hauptspeicher (1,333 GHz DDR3) und eine Festplatte (S-ATA) mit 500 GB.
- Alle Knoten sind mit InfiniBand⁵⁾ vernetzt. InfiniBand ist heute mit Abstand die bei Parallelrechnern weitestverbreitete Netzwerktechnologie. Für Parallelanwendungen ist neben hoher Bandbreite auch eine geringe Latenzzeit essentiell: In der Konfiguration des VSC beträgt die Bandbreite 40 Gbit/s, die Latenzzeit für eine Verbindung zwischen beliebigen Knoten liegt bei ca. eineinhalb Mikrosekunden. Alle InfiniBand-Komponenten stammen von QLogic.⁶⁾
- Als Betriebssystem wird CentOS⁷⁾ eingesetzt, eine für den Server- und HPC-Betrieb optimierte Linux-Variante.
- Die Softwareausstattung unterscheidet sich nicht wesentlich vom Schrödinger III: Als Batchsystem kommt die Sun Grid Engine zum Einsatz, parallelisiert wird mit MPI, Compiler (Fortran, C/C++) und numerische Bibliotheken stammen von Intel. Es ist nicht vorgesehen, kommerzielle Programmpakete in den Bereichen Finite Elemente und Computational Fluid Dynamics (CFD) einzusetzen.

Der Aufbau des Clusters erfolgte über die Sommermonate und verlief weitgehend nach Plan: Bis August 2009 war der Cluster fertig aufgebaut und verkabelt, dann starteten umfangreiche Tests. Dabei waren einige Schwierigkeiten zu überwinden: Es ist keine triviale Aufgabe, fast 500 Rechner mit Strom zu versorgen und die von den dicht gepackten Knoten – jeder Knoten ist nur eine Höheneinheit (4,45 cm) hoch – erzeugte Wärme abzuführen. Nach einem vierzehntägigen Dauertest, bei dem alle Komponenten des Clusters ihre Verlässlichkeit unter Volllast beweisen müssen, beginnt voraussichtlich Mitte Oktober der Benutzerbetrieb.

Personal

Im Rahmen des Projektbudgets werden zwei Personen für Betrieb und Benutzerbetreuung angestellt: Eine wird von der Universität Wien finanziert, die andere von der Universität für Bodenkultur, die dafür einen Anteil an den Ressourcen erhält. Es konnten zwei Experten gewonnen

werden, die an der Universität Wien keine Unbekannten sind: **Jan Zablouil** war in der Arbeitsgruppe Computer-gestützte Materialphysik um Prof. Hafner und Prof. Kresse tätig und konnte dort viele Erfahrungen im Supercomputing aus Anwendersicht sammeln. **Bernhard Hermann** war bei init.at angestellt und dort für die Betreuung des Schrödinger III zuständig; Zahlreiche Benutzeranfragen hat er rasch und kompetent beantwortet.

Ressourcenvergabe

Die Modalitäten der Ressourcenvergabe wurden vom *Steering Committee*⁸⁾ festgesetzt. Der Zugang zum Vienna Scientific Cluster erfolgt grundsätzlich auf Basis von Projekten, welche einen Peer-Review-Prozess erfolgreich durchlaufen haben und neben wissenschaftlicher Exzellenz den Bedarf an extrem hoher Rechenleistung nachweisen. Sofern es sich um Projekte handelt, die bereits vom FWF, der EU usw. positiv begutachtet wurden, ist kein weiterer Review-Prozess erforderlich. Zum Einreichen von Projekten und Beantragen von Ressourcen steht die Webmaske <http://service.zid.tuwien.ac.at/vsc/> zur Verfügung, die Authentifizierung und Autorisierung erfolgt mittels SAML2 über die ACONet-AAI⁹⁾. Zusätzlich werden noch rasch und unbürokratisch Testaccounts vergeben (befristet und mit begrenzten Ressourcen).

Für BenutzerInnen der Universität Wien gibt es einen eigenen Zugangsknoten vsc.univie.ac.at, der sich logisch im Datennetz der Universität Wien befindet. Physikalisch wurde die Verbindung zwischen Universität Wien und TU Wien auf 10 Gbit/s aufgestockt, sodass auch der Transfer großer Datenmengen zwischen dem VSC und Servern an Instituten der Universität Wien problemlos möglich ist. Die Home-Verzeichnisse werden auf den Backup-Servern der Universität Wien gesichert.

Abschaltung des Schrödinger III

Aufgrund der hohen Betriebskosten ist ein weiterer Betrieb des Schrödinger III – dessen Rechenleistung nur wenige Prozent des VSC beträgt – nicht mehr zu rechtfertigen. Der Betrieb wird daher mit Ende Oktober 2009 eingestellt: Am 27. Oktober um 9:00 Uhr werden alle noch laufenden Jobs beendet und der Benutzerbetrieb eingestellt. Danach erfolgt eine letzte Datensicherung, und am Nachmittag wird mit dem Abbau begonnen. Spätestens Freitag nachmittags (am 30. Oktober 2009) muss dieser beendet sein, denn am 31. Oktober läuft der Mietvertrag mit der Firma Interxion aus, bei der der Schrödinger aufgestellt ist. Der Zugangsknoten mit allen Daten und der gesamten Softwareumgebung des Schrödinger III wird als virtueller Server mit einzelnen Rechenknoten noch mehrere Monate weiter betrieben: Dieser Mini-Schrödinger, der unverändert unter der Adresse schroedinger.univie.ac.at zu erreichen sein wird, dient ausschließlich dazu, die Migration auf den VSC und andere Systeme zu erleichtern und ist nicht für einen Produktionsbetrieb geeignet.

Rechenleistung

Mit dem VSC wird Österreich höchstwahrscheinlich wieder in der **Liste der 500 schnellsten Supercomputer der Welt**¹⁰⁾ vertreten sein – seit dem Platz 344 für den Schrödinger II im November 2003 war dort kein österreichischer Rechner zu finden. Der VSC schafft mehr als 35 TFlop/s (Billionen Rechenoperationen pro Sekunde), was im Juni noch für einen Platz um 120 gereicht hätte. Zumindest ein Platz in den hinteren Rängen sollte sich aber auch in der nächsten Liste ausgehen – Mitte November werden wir es erfahren, wenn auf der Supercomputing-Konferenz SC09¹¹⁾ in Portland (Oregon) die aktuelle Liste präsentiert werden wird.

Wie schnell die Rechenleistung von Supercomputern ansteigt und wie groß die ständigen Anstrengungen sein müssen, um nicht den Anschluss zu verlieren, zeigt ein Blick auf die TOP500-Listen vergangener Jahre: Im November 2001 schaffte der Spitzenreiter 7,2 TFlop/s – vor acht Jahren wäre der VSC also mit Abstand schnellster Supercomputer der Welt gewesen. Schon im Juni 2002 wäre er jedoch vom japanischen Earth Simulator mit 35,8 TFlop/s knapp geschlagen worden. Im Juni 1993 war eine Connection Machine im Los Alamos National Laboratory mit 59,7 GFlop/s Spitzenreiter der TOP500-Liste – das ist weniger als die Rechenleistung eines einzigen Knotens des VSC.

Peter Marksteiner ■



Von außen betrachtet kaum mehr als ein großer Kasten: Der Supercomputer an seinem Standort im Freihaus-Gebäude der TU Wien.

- 1) *Doch kein Schrödinger IV – Ein HPC-Cluster für Wiener Universitäten*, <http://comment.univie.ac.at/08-3/22/>
- 2) www.ips.at
- 3) www.sun.com
- 4) *Vienna Scientific Cluster: Der gemeinsame Hochleistungsrechner von Universität Wien, Universität für Bodenkultur und TU Wien*. ZIDline 20, www.zid.tuwien.ac.at/zidline/zl20/vsc/
- 5) www.infinibandta.org
- 6) www.qlogic.com
- 7) www.centos.org
- 8) Steering Committee = gemeinsamer Beirat, in dem die Universität Wien durch Vizerektor Prof. Engl, Prof. Dellago (Dekan der Fakultät für Physik) und Dr. Rastl vom ZID vertreten ist.
- 9) siehe dazu auch: *AAI in Aktion*, <http://comment.univie.ac.at/07-2/21/> bzw. *Weblogin mit Shibboleth*, <http://comment.univie.ac.at/08-3/20/>
- 10) www.top500.org
- 11) <http://sc09.supercomputing.org/>