

Der Urknall im Serverraum

BladeCenter erobern festen Platz in der Grundlagenforschung



Überblick

■ Die Aufgabe

Wenn für einzelne MPI-Projekte bewilligte Kapazitäten im Rechenzentrum gemeinsam betrieben werden, wird die Leistung flexibler verteilt und die Auslastung besser optimiert. Dies erfordert eine leicht erweiterbare, mit weiteren Clustern integrierbare Lösung mit geringem Managementaufwand.

■ Die Lösung

Prozessoren der jeweils neuesten Generation (jetzt Quad-Core Intel Xeon), IBM HS21 Blades, BladeCenter H, angebunden über Infini-Band Netzwerk durch IBM Partner Computacenter an DS4700 Storage Server mit performantem Dateisystem GPFS.

■ Die Vorteile

Dank einfacher Steuerung und Erweiterbarkeit sind Rechenressourcen flexibel und wirtschaftlich einsetzbar.

Was ist Masse? Die Ausgangsfrage Einsteins formulieren Grundlagenforscher heute so: Gibt es kleinere Teilchen als Quarks? Aufschluss darüber soll in den kommenden Jahren das Atlas-Experiment des Forschungszentrums CERN in Genf liefern. Der titanische Teilchenzertrümmerer liefert hierzu die Messdaten. Auf deren Basis wird berechnet, ob die Resultate die vorhandenen Modelle vom Aufbau der Materie stützen bzw. welche neuen Modelle sich aus ihnen ergeben. Dies geschieht an Standorten auf der ganzen Welt, auch im Rechenzentrum Garching (RZG), das von der Max-Planck-Gesellschaft und dem Max-Planck-Institut (MPI) für Plasmaphysik betrieben wird. Das RZG stellt Rechenleistung für viele MPIs bereit, deren Schwerpunkte in der Material- und Fusionsforschung sowie der Astrophysik liegen.

Traditionell betreibt das RZG seit fast fünf Jahrzehnten als einen Schwerpunkt Hochleistungscomputer. Zusätzlich betreibt es Computersysteme für

mittlere Anforderungen, die nicht auf den teuren Superrechnern bearbeitet werden sollen. Gleich mit dem Aufkommen der Blade-Systeme wandte man sich am RZG dieser extrem kompakten Technologie zu. „Deren Architektur kommt unseren Anforderungen sehr entgegen“, erklärt Hermann Lederer, Leiter der Anwendungsunterstützung am RZG. „Und durch den gemeinsamen Betrieb von Rechenressourcen können in Engpasssituationen nach Absprache einzelne Gruppen ‘fremde’ Ressourcen zusätzlich nutzen. Eine solche Poolbildung betreiben einige Institute schon länger – zum Vorteil aller Beteiligten“, erläutert Lederer.

Abteilungscluster leistungsfähig wie Hochleistungsrechner vor 5 Jahren

Die einzelnen MPIs sind für ihre Ausstattung mit Rechenressourcen selbst verantwortlich. Einige Institute lassen ihre Rechner ganz oder teilweise am RZG installieren und betreiben. Seit Ende 2002 entschieden sich Institute immer wieder für den Aufbau von Linux-Clustern, vor allem für Standardaufgaben der Astro- und Plasmaphysik, der Quantenoptik, der Materialforschung sowie bei Beteiligungen an internationalen Großprojekten. So werten Wissenschaftler des MPI für Physik seit 2004 Daten des MAGIC-Teleskops aus.

Durch mehrere neue Projekte wuchs die Zahl der in Garching betriebenen BladeCenter. Dabei kam stets die neueste Prozessorgeneration zum Einsatz: Die ersten 32-Bit XEON- wurden zwischenzeitlich durch 64-Bit AMD-

Prozessoren verdrängt, während seit Ende 2006 wieder XEON Dual Core bzw. seit Mitte 2007 bereits XEON Quad Core verbaut wurden. Dabei stieg nicht nur die Rechenleistung, sondern aufgrund der schnellen Vernetzung mit InfiniBand auch der Umfang der Cluster. Die neuesten Cluster unterstützen mit ungefähr 20-40 Nutzern ganze Abteilungen und können mit einer Spitzenleistung von 5 Teraflop/Sekunde für Gruppenaufgaben eingesetzt werden. „Wir haben nach den ersten guten Erfahrungen konsequent auf BladeCenter gesetzt, weil diese ein durchgängiges Konzept und hohe Skalierbarkeit bei der Clusterbetreuung gewährleisten“, erläutert Lederer. „Das Management eines Pools von Ressourcen ist dann besonders einfach, wenn diese gleichartig sind. Die BladeCenter bieten dank besonderer Hardwareausstattung eine hohe Stabilität und hohen Komfort beim Einsatz der entsprechenden Management-Software.“

Leicht zu steuern

Für das mehr auf Supercomputing ausgerichtete RZG sind Hosting und Pflege der Linux-Cluster eher Zusatzaufgaben. Für das RZG ist es von hoher Bedeutung, den Personalaufwand bei immer mehr zu betreuenden Clustern so gering wie möglich zu halten. Deshalb betreibt es auf allen BladeCentern identische Linux-Systemversionen. „Das einheitliche Betriebssystem und die zentrale Konfiguration der BladeCenter machen die Steuerung besonders einfach“, sagt Christian Guggenberger. Als Verantwortlicher für das Linux-Cluster-Management gestaltet er die Steuerung konzeptionell und setzt sie größtenteils selbst um. „Ich kann von jedem Ort der Welt aus alle Dinge überwachen, einschalten oder darauf zugreifen.“ Für die System-Steuerung nutzt Guggenberger das IBM Cluster Management System (CSM) sowie eigenentwickelte Softwaremanagement-Tools, mit denen sich Fehler ein-

fach identifizieren und Ressourcen übersichtlich zuordnen lassen. Zudem erleichtert auch die Hardware-Ausstattung der BladeCenter den Betrieb, etwa durch stets vorhandene Managementinfrastruktur, ihre redundante Stromversorgung und unaufwändige Verkabelung. „Das macht Umzüge innerhalb des Rechenzentrums leichter – und die gibt es öfters, wenn wir Platz für neue Systeme schaffen müssen“, berichtet Guggenberger.

Eingespielte Partnerschaft

Die Hardware-Installation der BladeCenter übernimmt seit 2 Jahren der IBM Business Partner Computacenter, ein international tätiger IT-Dienstleister mit rund 4000 Mitarbeitern in Deutschland. Die Software hingegen installiert das RZG weitgehend selbst, damit diese stets einheitlich bleibt und den Bedürfnissen der Nutzer entsprechend konfiguriert ist. Bei Hardware-Konfiguration und Systemintegration greift es auf die Beratung von Computacenter zurück. „Dieses Zusammenspiel ist sehr fruchtbar, weil wir unsere Kompetenzen optimal ergänzen und die Rechner den Instituten schnell zur Verfügung stellen können“, so Guggenberger.

Im Ergebnis hat das RZG mit Hilfe der Blade-Technologie flexibel einsetzbare, leicht zu steuernde Rechenkapazitäten bereitgestellt, die ihrem derzeit noch in Betrieb befindlichen Supercomputer aus dem Jahr 2002/3 entsprechen oder diesen sogar übertreffen. „Vor vier bis fünf Jahren simulierten bei uns Wissenschaftler vom Max-Planck-Institut für Astrophysik die Materieverteilung seit dem Urknall mit 10 Milliarden Teilchen. Das war damals die größte derartige Berechnung aller Zeiten“, erzählt Lederer. „Heute können sie diese mit ihrem Linux-Cluster wiederholen – und zwar schneller.“ Wenn dieses Jahr die ersten Daten vom Atlas-Projekt des CERN kommen, werden sie gleichfalls mit einem solchen Cluster ausgewertet.

Technische Daten

BladeCenter Cluster HS21 mit 2*Intel 5160 (bzw. HS21 XM mit 2*Intel 5345), 2 GB pro Core, IB 4xSDR, Voltaire IB Fabric, I/O Knoten (x3650) mit DS4700 Disk Subsystem und Dateisystem GPFS. BladeCenter Cluster HS21 XM mit 2*Intel 5345, 2GB pro Core, Ethernet only, 10 GB Uplink pro Chassis.



Weitere Informationen

Dieter Oehmigen
IBM Deep Computing
E-Mail: dieter_oehmigen@de.ibm.com



IBM Deutschland GmbH
70548 Stuttgart
ibm.com/de

Die IBM Homepage finden Sie unter:
ibm.com

IBM, das IBM Logo und ibm.com sind eingetragene Marken der IBM Corporation.

BladeCenter und DS4000 sind Marken der IBM Corporation in den USA und/oder anderen Ländern.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium sind Marken der Intel Corporation in den USA und/oder anderen Ländern.

Weitere Unternehmens-, Produkt- oder Service-namen können Marken anderer Hersteller sein.

Diese Erfolgsgeschichte verdeutlicht, wie ein bestimmter IBM Kunde Technologien/Services von IBM und/oder einem IBM Business Partner einsetzt. Die hier beschriebenen Resultate und Vorteile wurden von zahlreichen Faktoren beeinflusst. IBM übernimmt keine Gewährleistung dafür, dass in anderen Kundensituationen ein vergleichbares Ergebnis erreicht werden kann. Alle hierin enthaltenen Informationen wurden vom jeweiligen Kunden und/oder IBM Business Partner bereitgestellt. IBM übernimmt keine Gewähr für die Richtigkeit dieser Informationen.

Gedruckt in Deutschland.

© Copyright IBM Corporation 2008
Alle Rechte vorbehalten.

IBM Form GK12-4315-00 (08/2008)