

Grid-Computing

François Bry¹, Wolfgang E. Nagel² und Michael Schroeder³

¹ Institut für Informatik, Ludwig-Maximilians-Universität München

² Zentrum für Hochleistungsrechnen und Fakultät für Informatik,
Technische Universität Dresden

³ Biotec (Biotechnologisches Zentrum) und Fakultät für Informatik,
Technische Universität Dresden

„Grid-Computing“, ein Mitte der 90er Jahre eingeführter Begriff [1], bezeichnet eine Architektur für verteilte Systeme, die auf dem World Wide Web aufbaut und die Web-Vision erweitert. Mit dem Grid-Computing werden die Ressourcen einer Gemeinschaft von sogenannten “virtuellen Organisation” (siehe unten) integriert. Die Hoffnung ist, dass hierdurch rechen- und/oder datenintensiven Aufgaben, die eine einzelne Organisation nicht lösen kann, handhabbar werden. Ein “Grid” bezeichnet eine nach dem Grid-Computing-Ansatz aufgebaute Rechner-, Netzwerk- und Software-Infrastruktur, die eine virtuelle Organisation zur Erledigung ihrer Aufgaben verwendet. Zu Beginn war die Möglichkeit, ungenutzte CPU-Ressourcen an anderen Stellen für die eigenen Aufgaben einzusetzen, die wesentlich treibende Kraft für erste Experimente. Internet-Computing-Projekte wie SETI@Home, distributed.net u.a., bei denen die unbenutzten Rechenzyklen von weltweit verteilten privaten PCs verwendet werden, illustrieren das Potential des Grid-Computing. Allerdings gehen die heutigen Grid-Konzepte und auch die ersten Prototypen weit über diese Anfänge hinaus und versprechen die transparente Bereitstellung von Diensten unabhängig von der räumlichen Nähe. Damit gibt es die Erwartung, dass das Grid-Computing die Nutzung von Rechnern und Rechnernetzen so grundlegend verändern wird, wie das Web den den Datenaustausch bereits verändert hat.

Vom Web zum Grid: Virtualisierung von Rechendiensten zur kooperativen Problemlösung

Das heutige Web ermöglicht die Virtualisierung von Daten, d.h. eine Übertragung der Daten von einzelnen Rechner auf vernetzte Rechnern. Die derzeit entstehenden Web-Dienste und das Semantische Web stellen zwei komplementäre Formen der Virtualisierung von Anwendungssoftware dar. Das Grid-Computing erzielt eine Virtualisierung von netzwerk-bezogenen Diensten zur kontrollierten Ressourcenteilung.

Virtuelle Organisationen

Das Grid-Computing soll so genannten “Virtuellen Organisationen” ermöglichen Ressourcen wie Daten, Software, Sensoren, Rechner und Rechnernetzwerke, mit dem Ziel zu teilen, Probleme kooperativ zu lösen. Eine virtuelle Organisation ist ein Zusammenschluss von Personen, Unternehmen und realen Organisationen. Eine virtuelle Organisation kann nur vorübergehend bestehen und/oder veränderbar sein. Virtuelle Organisationen werden über dynamische, d.h. veränderbare, “Relationen” wie etwa Rollen und Zugangsrechte definiert, die Beteiligungen von (realen oder virtuellen) Akteuren festlegen. So kann z.B. eine Web-basierte Lehr- und Lernumgebung die Hörschaft und Dozenten einer Lehrveranstaltung als virtuelle Organisation definieren und verwalten, deren Teilnehmer unterschiedliche Rollen und Rechte (z.B. Zugangsrechte zu Datenbestände, Lehrmaterial, Software und Rechnern) haben. Ein Autohersteller und seine Zulieferer können z.B. eine virtuelle Organisation bilden, um die Daten, Software sowie Rechenkraft auszutauschen, die zur Entwicklung des Modells eines Prototyps notwendig

sind. In der Bioinformatik kann eine virtuelle Organisation über ein Grid Ressourcen für rechenintensive Aufgaben wie Strukturvorhersage oder molekulardynamische Simulationen zur Verfügung stellen.

Ein Grid ermöglicht die Interoperabilität innerhalb einer dynamischen und heterogenen Gemeinschaft von virtuellen Organisationen. Jede an einem Grid beteiligte virtuelle Organisation kann Ressourcen wie Daten, Software und Rechenleistungskraft anbieten und über die Nutzung fremder Ressourcen mit anderen virtuellen Organisation des Grids verhandeln. Eine Grid-Gemeinschaft ist heterogen, da die an ihr beteiligten virtuellen Organisationen unterschiedliche Hardware-, Netzwerk-, Basis- und Anwendungssoftware verwenden können. Desweiteren ist die Beziehung zwischen virtuellen Organisationen von unterschiedlichem Vertrauen geprägt, und der Zugang zu Ressourcen unterliegt unterschiedlichen Bedingungen (z.B. Kosten).

Innerhalb eines Grids ist die Kontrolle über lokale Ressourcen lokal: Eine virtuelle Organisation legt selber fest, wann, wo, wofür und unter welchen Bedingungen (z.B. zu welchen Kosten) Ressourcen der Grid-Gemeinschaft zur Verfügung gestellt werden. Folglich muss ein Grid zum einen unterschiedliche lokale Sicherheitslösungen (Authentifizierung (authentication) und Autorisierung) ermöglichen, zum anderen Lösungen zum Ressourcenaufspüren (resource discovery) bieten. Ein Grid unterstützt unterschiedliche Formen der Ressourcennutzung: Einzelbenutzer- oder Mehrbenutzernutzung, performance-sichernde oder kostensparende Nutzung. Folglich muss ein Grid Lösungen für Dienstqualität (quality-of-service, Scheduling, gleichzeitige mehrfache Ressourcenzuweisung (co-allocation) und Abrechnung (accounting) anbieten. Die Verwaltung der veränderbaren Regelwerke (policies), die die Bedingungen der Ressourcennutzung spezifizieren, ist eine weitere Kernaufgabe eines Grids.

Dem Grid-Computing liegt die Annahme zugrunde, dass dynamische und heterogene virtuelle Organisationen und deren Verwaltung, die Ressourcenvermittlung und die koordinierte Ressourcenteilung nicht nur zur Lösung von daten- und/oder rechenintensive Aufgaben, also nicht nur für das wissenschaftlichen Rechnen, sondern auch allgemein in der angewandten Informatik zunehmend wichtig sein werden. Wie das Web sich weit über den Austausch von wissenschaftlichen Daten und Dokumenten – hinaus etabliert hat, so wird erwartet, dass das Grid-Computing eine sehr breite Anwendung finden wird.

Grid-Architektur: Von Internet-Protokollen zu Grid- und Intergrid-Protokollen

Um die Interoperabilität von heterogenen virtuellen Organisationen zu ermöglichen, setzt das Grid-Computing auf Protokolle. In einer verteilten und heterogenen Umgebung sind Protokolle deswegen unabdingbar, weil sie es ermöglichen festzulegen, wie Komponenten des verteilten Systems kommunizieren können, um ein bestimmtes Verhalten zu erreichen, ohne dabei vorauszusetzen, wie das gewünschte Verhalten realisiert wird.

Das Grid-Computing baut auf das Internet auf. Es erweitert die Internet- und TCP/IP-Protokoll-Architektur mit zusätzlichen Protokollen und mit auf Internet-Protokollen basierenden Diensten, die in der standardisierten Open Grid Services Architecture (OGSA) wie folgt definiert sind.

Application (z.B. HTTP, FTP, SMTP)	Application
	Collective
	Resource
Host-to-Host Transport (z.B. TCP)	Connectivity
Internet (z.B. IP, ARP, ICMP, IGMP)	
Network Interface	Fabric

Internet-Protokoll-Architektur und Open Grid Services Architecture in Beziehung zueinander

Die Fabric-Schicht ermöglicht den Zugang zu den Ressourcen wie zum Beispiel Programme, Speicher, Kataloge (z.B. Datenbanken), Sensoren, Rechner und Netzwerke, die über ein Grid vermittelt werden. Die Fabric-Schicht besteht aus Protokollen, um z.B. Information über Ressourcen (wie z.B. Betriebssystem, Hardware, Last) zu erhalten, Programme aufzurufen, Dateien zu erhalten und abzulegen,

Versionen von Programmcode zu verwalten, Kataloge und Datenbanken anzufragen oder zu verändern. Eine Fabric-Schicht kann unterschiedliche Grade der Ressourcenteilung, z.B. Co-Scheduling, anbieten. Fabric-Schichten mit in diesem Kontext beschränkten Möglichkeiten sind einfacher zu realisieren und werden deswegen oft vorgezogen.

Die Connectivity-Schicht besteht aus den Kommunikations- und Authentifizierungsprotokollen, die den vom Grid angebotenen Transaktionen zugrunde liegen. Die Kommunikationsprotokolle ermöglichen den Austausch von Daten zwischen den Ressourcen, die über die Fabric-Schicht zugänglich sind. Dies umfasst die Funktionen Transport, Routing, und Naming, die über die Internet-Protokolle der Internet-, Transport- und Application-Schichten realisiert werden können. Die Authentifizierungsprotokolle bieten sichere (auf kryptographischen Methoden beruhenden) Verfahren zur Feststellung der Identität von Benutzern (d.h. virtuellen Organisationen) wie Ressourcen (z.B. Sensor oder Rechner). Die Authentifizierungsprotokolle sollen folgendes ermöglichen:

- Nach einer einziger Authentifizierung soll ein Benutzer Zugang zu allen (über die Fabric-Schicht erreichbaren) Ressourcen haben.
- Ein Benutzer soll seine Nutzungsrechte einem Dienst übertragen, so dass dieser Dienst Zugang zu Ressourcen anstelle des Benutzers hat. Man spricht von *Delegation*.
- Unterschiedliche (z.B. Windows-, Unix-, Linux-, oder Kerberos-basierte) Sicherheitslösungen statt einer einheitlicher Sicherheitslösung sollen für die an einem Grid beteiligten Ressourcenanbieter möglich sein.
- Wird einem Benutzer Zugang zu Ressourcen von verschiedenen Ressourcenvermittler gewährt, so ist dieser Zugang möglich, ohne dass die Ressourcenvermittler Vertrauensbeziehungen zueinander aufbauen. Man spricht von Benutzer-zentrierten Vertrauensbeziehungen.

Die Ressource-Schicht besteht aus Protokollen, die sicheren Zugang und Verwaltung sowie Abrechnung (accounting) und Bezahlung von Ressourcen regeln. Die Ressource-Schicht baut auf die Kommunikations- und Authentifizierungsprotokolle der Connectivity-Schicht auf. Man unterscheidet zwischen Informationsprotokollen, die Auskünfte über die Struktur und Zustand einer Ressource geben, und Verwaltungsprotokollen, die Zugang zu Ressourcen ermöglichen.

Die Collective-Schicht besteht aus Protokollen zur koordinierten Nutzung von mehreren Ressourcen z.B. zur gleichzeitigen mehrfachen Ressourcenzuweisung (co-allocation), zum Scheduling und zum Workload Management.

Die Application-Schicht bezieht sich auf die Anwendungen, die ein Grid zur Verfügung stellt. Diese Anwendungen werden unter Verwendung der Protokolle aller niedrigeren Schichten implementiert.

Intergrid-Protokolle ermöglichen, dass auch Grids miteinander interagieren können, die auf unterschiedlichen Grid-Protokollen beruhen.

Eine sehr frühe Realisierung eines großen Teils dieser Funktionalität und der damit angebotenen Dienste ist im – vom BMBF im Zeitraum 1997 bis 2002 geförderten – UNICORE-Projekt realisiert worden und befindet sich heute zumindest im wissenschaftlichen Anwendungsbereich erfolgreich im Produktionseinsatz.

Grid- und Peer-to-Peer-Computing im Vergleich

Das Peer-to-Peer-Computing zum Datenaustausch, welches Systeme wie Napster und Gnutella bekannt gemacht hat, hat ähnliche Ziele wie das Grid-Computing. Allerdings liegt ein wesentlicher Unterschied zwischen Grid-Computing und Peer-to-Peer darin, dass letzteres keine Protokoll-Architektur besitzt.

Konvergenz von Techniken

Das Grid-Computing verlangt die Konvergenz von mehreren Techniken und Ansätzen. Gebiete, die das Grid-Computing unmittelbar betreffen, sind: Verteilte Systeme, Rechnernetze, Peer-to-Peer-Computing, Authentifizierung und Autorisierung, Abrechnung und Bezahlung.

Weitere Gebiete, von denen bei der Entwicklung des Grid-Computings Beiträge zu erwartet werden können, sind: Datenbanksysteme, insbesondere höhere (Web-)Anfragesprachen (u.a. zum Aufspüren von Ressourcen), Wissensmanagement und automatisches Schließen (u.a. für Verhandlungen, Regelwerke (policies), Modellierung und Verwaltung von virtuellen Organisationen, Ressourcenvermittlung), Semantische Web (u.a. zur Modellierung, Anfrage und Aktualisierung von Meta-Daten über virtuelle Organisationen, Ressourcen und Regelwerke), Kooperative Systeme (wie CSCW, d.h. Computer Supported Cooperative Work), und Workflow-Management.

Grid-Computing und e-Science

Unter e-Science wird die Vision einer flächendeckende Nutzung des Grid-Computings in der Wissenschaft – v.a. zum wissenschaftlichen Rechnen aber auch zur besseren Nutzung von Ressourcen wie Software und Literatur – verbreitet. Doch kann Grid-Computing zu einem ähnlichen Erfolg wie das Web werden? Das Web schlug u.a. durch, da es zu einer „Win-Win“-Situation für Informationsanbieter und -konsumenten führte. Die Anbieter konnten bei minimalen Kosten Informationen anbieten, die für Konsumenten oft frei verfügbar sind. In den meisten Fällen sind zwischen beiden Parteien keine Verträge bezüglich Qualität, Sicherheit, etc. nötig. In einem Grid ist die Situation wesentlich komplexer: Anbieter von Diensten haben Kosten, die von Nutzern gedeckt werden müssen. Anbieter und Nutzer von Diensten müssen einander vertrauen, da Daten der Nutzer und Infrastruktur der Anbieter zusammenkommen müssen und nicht hundertprozentig voreinander geschützt werden können. Die Zahl der Nutzer komplexer Grid-Dienste ist gegenwärtig wesentlich kleiner als die Massen von Nutzern, die einfachste Informationsdienste im Web abrufen. Zusammenfassend bedeutet dieses, dass viele schwierigen Probleme zu lösen sind, bevor Grid-Computing zu einem etablierten und breit akzeptierten „Phänomen“ wie das Web werden kann.

Weitere Informationen

Globus: <http://www.globus.org>

Unicore: <http://www.unicore.org>

Global Grid Forum: <http://www.gridforum.org>

Danksagung

Diese Arbeit wurde im Rahmen des Projektes REVERSE (<http://reverse.net>) von der Europäischen Kommission und dem schweizer Bundesamt für Bildung und Wissenschaft (BBW) gefördert.

Literatur

- [1] Foster I. and Kesselman C. *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann, 1999.