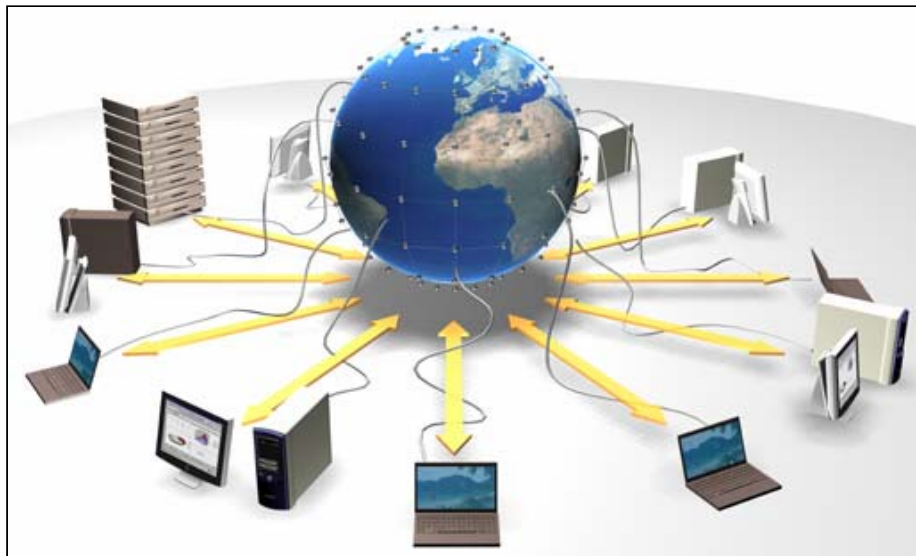


Seminararbeit zum Seminar
Semantic Grid
WS2004/2005

THEMA
**I.1 Idee und Entwicklung des
Grid Computing**
05. Januar 2005



CERN artists diagram demonstrating Grid concept.(Copyright CERN)

Inhaltsverzeichnis

1. Erklärung des Grid Begriffs	3
2. Arten von Grid Systemen	6
2.1 Computational Grid	6
2.1.1 Distributed Supercomputing.....	6
2.1.2 High Throughput Computing	7
2.2 Data Grid	7
2.2.1 Data Intensive Computing	7
2.3 Service Grid.....	7
2.3.1 On Demand Computing	8
2.3.2 Collaborative Computing	8
2.3.3 Multimedia Computing	8
3. Grid Benutzer	9
3.1 Projekte	9
3.2 CERN	11
4. Grid Architektur	13
4.1 Management System	13
4.2 Donor Software	13
4.3 Submission Software.....	14
4.4 Distributed Grid Management	14
4.5 Schedulers	15
4.6 Communications.....	16
4.7 Observation, Management, Measurement.....	16
5. Aktueller Stand / Entwicklung	17
5.1 Aktueller Stand.....	17
5.2 Entwicklung	17
6. Literatur- und Quellenverzeichnis	18

1. Erklärung des Grid Begriffs

Auf der ganzen Welt existieren viele Millionen von Computern. Neben Desktoprechnern, Laptops und Supercomputern gehört aber auch ein Vielzahl von Instrumenten wie z.B. meteorologische Sensoren oder auch Satelliten dazu. Eine Menge dieser Geräte sind über das Internet miteinander zu einem großen Netzwerk verbunden. Durch das World Wide Web können diese Computer Informationen in Form von Webseiten und anderem miteinander teilen.

Ein Grid geht noch einen Schritt weiter in diese Richtung. Die darin enthaltenen Computer und Instrumente teilen nicht nur Informationen miteinander, sondern auch Rechenleistung und Ressourcen wie z.B. Speicherplatz, Datenbanken oder Softwareanwendungen.

Während ein Computer Tage für eine komplexe Berechnung benötigt, wird dies vom Grid durch den Verbund hunderter Computer viel schneller und effizienter erledigt. Damit ermöglicht ein Grid als Infrastruktur, den Zugang zu Computerressourcen auch im high-end-Bereich.

Das Wort „Grid“ (engl.) bedeutet im Deutschen soviel wie „Gitter“ oder auch „Versorgungsnetz“ im Bereich der Elektrizität. Aber auch „Gitternetz“ im Zusammenhang mit geografischen Karten.

Der Begriff wurde als Analogie zum Stromnetz (engl.: „Electric Power Grid“) gewählt, um damit die Funktionsweise eines Grids zu verdeutlichen. Ein Grid soll so einfach zu Nutzen sein, wie es das Stromnetz ist. Der Nutzer verbindet sich mit dem Grid und bekommt ganz einfach Ressourcen zur Verfügung gestellt. Dabei muss sich der Nutzer keine Gedanken über die Herkunft der Ressourcen machen.

Ganz einfach Ausgedrückt kann man die Funktion eines Grids auch als „Computer aus der Steckdose“ bezeichnen.

Das Grid gilt mittlerweile auch als die nächste Generation des Internets.

Da die Funktion eines Clusters, der des Grids sehr ähnlich ist, soll in Tabelle 1.1 einmal der Unterschied beider Techniken aufgezeigt werden.

Cluster	Grid
<ul style="list-style-type: none"> • Verbund von mehreren Rechnern • Kommunikation über LAN • gemeinsamer lokaler Standort • zentraler Admin • Existenz des Clusters wegen Berechnungen • Cluster ist Hauptarbeit der Rechner 	<ul style="list-style-type: none"> • Verbund von mehreren Ressourcen • Kommunikation über WAN • verteilte Standorte • Jede Maschine eigener Admin • Existenz des Grids aus verschiedenen Gründen • Grid ist Nebenarbeit der Ressourcen

Tabelle 1.1: Unterschiede zwischen einem Cluster und einem Grid

Ein Cluster besteht aus einem Verbund von mehreren Rechnern, die alle an einem Standort stehen und über ein lokales Netz miteinander verbunden sind (z.B. in einem Unternehmen). Der Cluster wird von einem zentralen Administrator verwaltet. Der Hauptgrund für den Betrieb jedes einzelnen Rechners ist der Cluster, mit dem eine Vielzahl von Berechnungen durchgeführt werden können.

Im Gegensatz dazu ist ein Grid ein Verbund mehrerer Ressourcen, also nicht nur aus Rechnern. Diese können über die ganze Welt verteilt sein und sind über ein WAN (z.B. das Internet) miteinander verbunden. Daraus ergibt sich, dass jede Maschine einen eigenen Administrator besitzt, der bei privaten Ressourcen natürlich oft gleichzeitig der Besitzer ist. Die Arbeit, die im Gridverbund erledigt wird, ist nur Nebenaufgabe für die Ressourcen. Die eigentliche Hauptaufgabe variiert von Ressource zu Ressource.

Der Begriff des „Grid Computing“ wurde 1998 von Ian Foster und Carl Kesselmann, in ihrer Veröffentlichung „The Grid: Blueprint for a New Computing Infrastructure“ geprägt. Das Grid Computing ist eine aktuelle, immer wichtiger werdende Basistechnologie zur Nutzung verteilter Rechen- und Speicherressourcen. Darunter versteht man das Zusammenfassen der

Rechenleistung innerhalb eines Netzwerkes, wobei jeder Computer eine gleichgestellte Einheit darstellt.

Ein großer Vorteil des Grid Computing, sind die geringeren Kosten und die höhere Kapazität des Netzes, die sogar die eines Supercomputers leicht übertreffen kann.

Eine wichtige Eigenschaft von Grid Systemen ist deren Skalierbarkeit. Das bedeutet, dass zu jedem Zeitpunkt Rechner zum Grid hinzugefügt oder auch weggenommen werden können, ohne dass die Funktion des Grids in irgendeiner Art beeinträchtigt wird. Des Weiteren können Grids problemlos zu einer größeren Einheit, so genannten „Über-Grids“ zusammengefasst werden.

Eine Infrastruktur zum Aufbau eines Grids wird u.a. im Forschung und Entwicklungsprojekt „Globus“ entwickelt (www.globus.org). Das Ziel dieser Gruppe ist die Erstellung der notwendigen Basis-Technologien als Grundlage für das Grid Computing.

2. Arten von Grid Systemen

Abbildung 2.1 zeigt eine hierarchische Einteilung von Grid Systemen, die im folgenden Kapitel genauer erläutert werden.

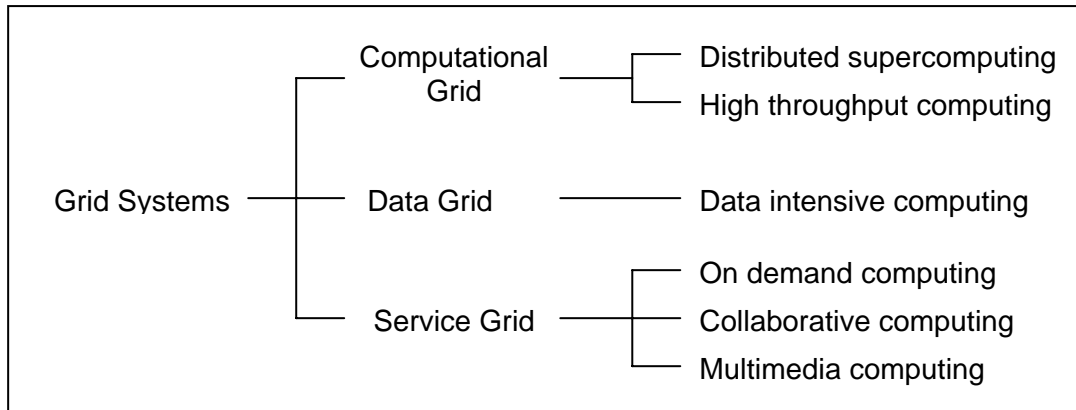


Abbildung 2.1: Hierarchische Einteilung von Grid Systemen

Grid Systeme können in *Computational Grids*, *Data Grids* und *Service Grids* unterteilt werden.

2.1 Computational Grid

Computational Grids zeichnet die Eigenschaft aus, dass das Gridsystem eine höhere Rechenkapazität für eine einzelne Anwendung besitzt, als jeder einzelne Rechner im System.

Abhängig von der Nutzung solcher Systeme, werden diese in *Distributed Supercomputing* und *High throughput computing* unterteilt.

2.1.1 Distributed Supercomputing

Distributed Supercomputing (engl.) bedeutet übersetzt soviel wie "verteiltes Superrechnen". Dabei wird ein Verbund von Supercomputern genutzt um die Berechnungszeit für große Aufgaben zu verkürzen. Dies wird vor allem bei Berechnungen genutzt, bei denen ein großer Bedarf an CPU-Leistung und Speicherplatz besteht.

Eine Beispielanwendung hierfür wäre die Wettervorhersage, bei der hochkomplexe Berechnungen durchgeführt und eine Menge an Daten berücksichtigt werden müssen.

2.1.2 High-Throughput Computing

Bei *Hight-Throughput Computing* werden Berechnungen mit einem hohen Datendurchsatz durchgeführt. Hierfür werden u.a. ungenutzte Ressourcen in schwach gekoppelten Systemen zusammengeführt und für Berechnungen genutzt, die selbst für Supercomputer nicht in angemessener Zeit zu lösen wären.

Das zu lösende Problem muss dabei jedoch der Anforderung entsprechen, dass es in kleinere Teilaufgaben unterteilt werden kann.

Ein Beispiel für eine solche Aufgabe ist das Projekt „Seti@Home“ (Siehe Kapitel 3.1).

2.2 Data Grid

Als *Data Grid* werden Systeme bezeichnet, die auf vorhandene Datenbestände (z.B. digitale Büchereien) die im WAN angeboten werden zugreifen und neue Informationen aus diesen zusammenstellen.

Eine Untergruppe von *Data Grids* ist das *Data intensive computing*.

2.2.1 Data Intensive Computing

Darunter fallen Datenintensive Berechnungen, die bei der Synthese von neuen Informationen aus geographisch verteilten Datenquellen und Datenbanken anfallen. Dabei werden oft hohe Ansprüche sowohl an die Rechenleistung, als auch an die Kommunikationsleistung des Systems gestellt.

In diesem Zusammenhang sei auch das Semantic Grid erwähnt, welches u.a. diese Kommunikationsleistung ermöglichen soll.

2.3 Service Grid

Systeme, die zur Kategorie der *Service Grids* zählen, stellen Dienste zur Verfügung, die nicht von einem einzelnen Rechner angeboten werden.

Diese Kategorie wird in *on-demand computing*, *collaborative computing* und *multimedia computing* unterteilt.

2.3.1 On-Demand Computing

Beim on-demand computing können nach Bedarf Ressourcen hinzugefügt werden. On-demand computing wird oft als Dienstleistung angeboten. Dabei können kurzzeitig auftretende Ressourcenbedürfnisse, die lokal nicht kosteneffektiv oder bequem und schnell erfüllt werden können, durch Anmietung von Ressourcen von Anbietern dieser Dienste gedeckt werden.

Eine Beispielsituation hierfür wäre die eines Wissenschaftlers, der die Genauigkeit einer Simulation steigern möchte, indem er weitere über das Grid angeforderte Rechenleistung in die Berechnung der Simulation mit einbezieht.

2.3.2 Collaborative Computing

Ein collaborative Grid verbindet Benutzer und Applikationen zu Arbeitsgruppen. Diese Systeme ermöglichen eine Echtzeit-Interaktion zwischen Benutzern und Applikationen auf einem virtuellen Arbeitsfeld.

2.3.3 Multimedia Computing

Ein Multimedia Grid stellt die Infrastruktur für Echtzeit Multimedia-Applikationen zu Verfügung. Dies verlangt die Unterstützung von QoS über verschiedene Rechner hinweg.

3. Grid Benutzer

Zu den Nutzern von Grid Systemen gehören unter anderem Regierungen, Gesundheitsorganisationen und wissenschaftliche Einrichtungen.

Es werden mittlerweile eine Vielzahl von öffentlichen Projekten durchgeführt, an denen jeder teilnehmen kann. Darüber hinaus gibt es auch einige große Projekte, die in geschlossenen Grid Systemen durchgeführt werden.

Einen kurzen Einblick in die Funktionsweise von Grids in Projekten soll im folgenden anhand des offenen Projektes „Seti@Home“ und eines geschlossenen Projektes am CERN gegeben werden.

3.1 Grid Computing Projekte:

Der bekannteste Vertreter eines öffentlichen Grid Projektes, ist mit mehr als 5 Millionen Teilnehmern „Seti@Home“ der Universität Berkeley in Californien (<http://setiathome.ssl.berkeley.edu/>). Bei diesem Projekt stellen weltweit Nutzer ihre Rechner zur Verfügung, um die über ein Radioteleskop in Arecibo, Puerto Rico empfangenen Signale aus dem Weltraum zu untersuchen.

Da die bei der Untersuchung anfallenden Datenmengen immens sind, werden so viele Rechner wie nur möglich gebraucht, die gemeinsam diese Datenmenge auswerten können.

Um an dem Projekt teilnehmen zu können, muss man als Erstes einen Client herunterladen und auf seinem Rechner installieren.

Startet man dieses Programm, wird bei bestehender Internetverbindung ein Datenpaket von den Servern der Universität Berkeley heruntergeladen. Danach wird bis zum Versandt der Ergebnisse keine Internetverbindung mehr benötigt. Sobald der Rechner nun nicht anderes zu tun hat, startet der Client als Bildschirmschoner und beginnt mit den Berechnungen. Dies kann, je nach Rechenleistung bis zu 20 Stunden dauern.

Ist das Paket abgearbeitet, werden die berechneten Daten über das Internet zurück an den Server geschickt und ein neues Datenpaket wird angefordert.

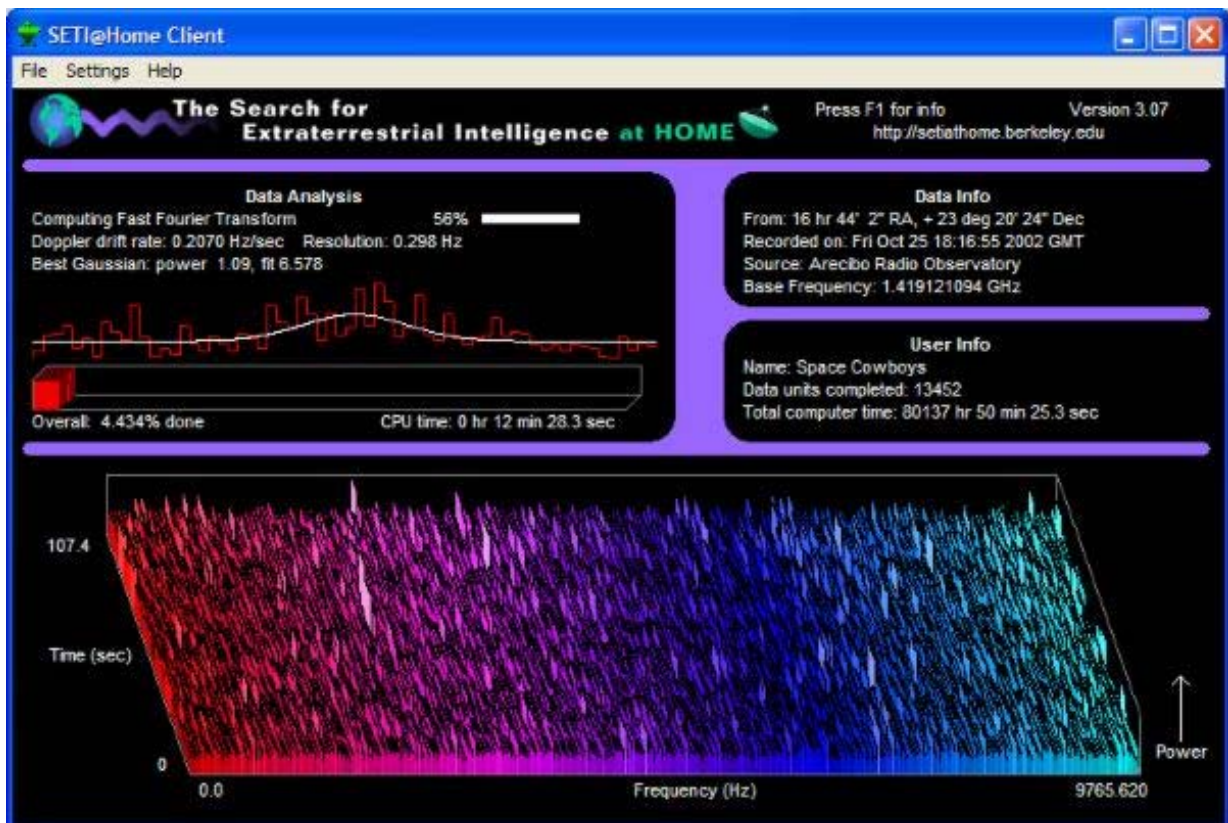


Abbildung 3.1: Clientoberfläche des Seti@Home Projektes

Ein Anreiz zur Teilnahme an diesem Projekt wird durch Rubriken auf der Webseite, wie z.B. der *User of the Day* oder auch herausgegebene Urkunden für eine Bestimmte Anzahl bearbeiteter Pakete gegeben und natürlich mit dem Ausblick, der Erste zu sein, der Außerirdische Signale empfängt.

Bis zum 19.Januar 2004 wurden ca. 2.195.136 Jahre Rechenzeit geleistet (diese Berechnungszeit ergibt sich aus der Summe der Berechnungszeit der teilnehmenden Nutzer, die natürlich parallel rechnen).

Weitere interessante Projekte sind z.B.

- Folding@Home <http://folding.stanford.edu/>
Simulation von Faltungsvorgängen bei Proteinen
- Internet Movie Project <http://www.imp.org/>
Modelling und Rendering eines kompletten Films
- ZetaGrid <http://www.zetagrid.net/>
Verifizierung der Riemann'schen Hypothese (Mathematik)

Derzeit existieren über 50 kommerzielle und nicht-kommerzielle Projekte. Einen Überblick über aktuelle Projekte kann man sich auf <http://distributedcomputing.info/> ansehen. Eine gute deutsche Übersicht ist unter <http://www.rechenkraft.net> zu finden.

3.2 CERN

Auch im Kernforschungszentrum CERN (Centre Européenne pour la Recherche Nucléaire, <http://www.cern.ch>) hat man die Vorteile eines Grid erkannt. Im Jahr 2007 wird dort der neue Large Hadron Collider (LHC) in Betrieb genommen, ein 27 Kilometer langer Teilchenbeschleuniger (Ringtunnel), in dem Teilchen mit Geschwindigkeiten nahe der des Lichts circa 40 Millionen Mal die Sekunde auf riesige Detektoren treffen. Daraus entsteht ein derart hohes Datenaufkommen (ca 10 Petabyte = 10.000.000 GB pro Jahr), dass dies nicht mehr vom CERN und seinen angeschlossenen Partnern bewältigt werden kann.

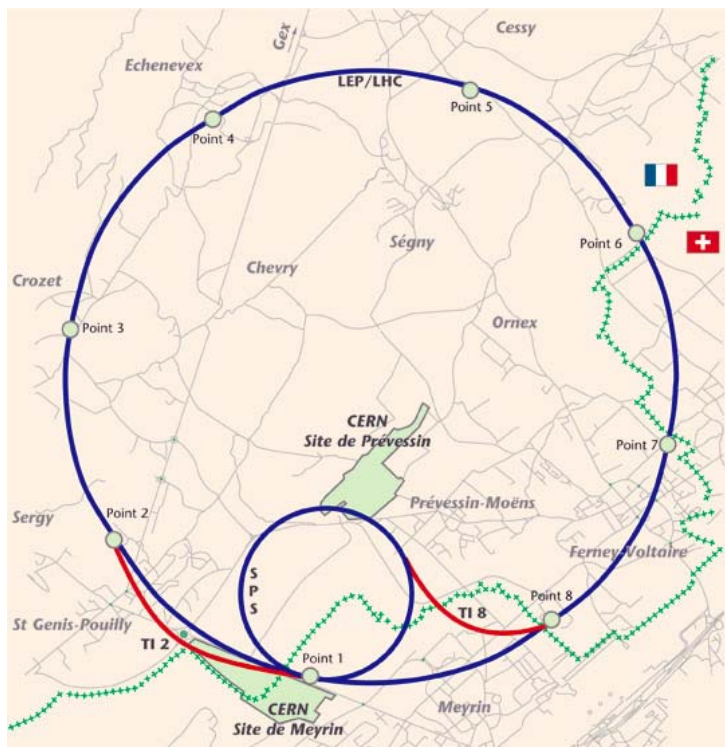


Abbildung 3.2:
Plan des LHC Ringes

Darüber hinaus arbeitet die Hälfte aller Teilchenphysiker der Welt regelmäßig an CERN-Rechnern (siehe auch Abbildung 3.3). Aus diesen Gründen hat das

Forschungsinstitut einen großen Teil zur Schaffung des Grids beigetragen und wird auch weiterhin darauf setzen.

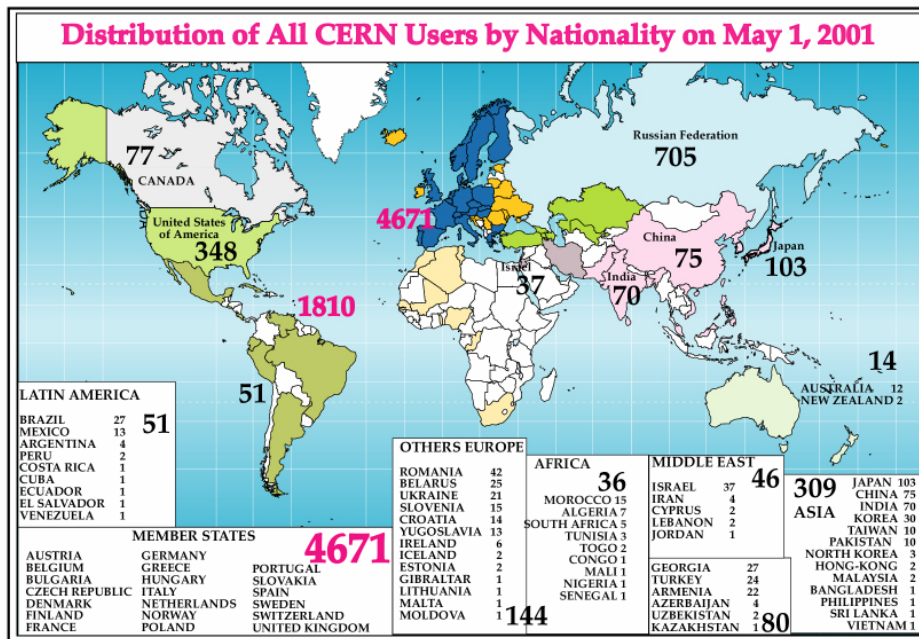


Abbildung 3.3:
Verteilung der
CERN Nutzer

Um die Organisation der Datenverteilung zu bewältigen, wurde eine mehrstufige Hierarchie (Tiers) entwickelt.

Das CERN (Tier 0) verteilt die Daten an einzelne Rechenzentren (Tier 1). Von dort werden ausgewählte Institute (Tier 2) mit den Daten versorgt. Diese Daten stehen den Forschern dann lokal zur Verfügung.

Für den Datenzugriff und die Verteilung wird hierbei Grid-Technologie eingesetzt. Dabei stellt sich für die anschließende Auswertung und Bearbeitung der Daten durch die Forscher die Frage, ob hierzu Daten lokal benötigt und deshalb transferiert werden müssen oder ob es für einzelne Rechenaufgaben sinnvoller ist, die Berechnungsaufgabe an einen Rechner zu schicken, der den Datensatz schon besitzt. Über das Grid kann ermittelt werden, wo die benötigten Daten und die geeigneten Rechnerkapazitäten zur Verfügung stehen. Langfristig können Daten bei Bedarf an freie Rechner geschickt werden, um eine schnellere Bearbeitung zu ermöglichen.

4. Grid Architektur

Ein einfaches Grid kann von ein paar Programmierern in ihrer Freizeit eingerichtet werden. Bei einem komplexeren Gridsystem, mit dem die Nutzer wichtige Berechnungen durchführen und das minimal Fehleranfällig sein darf, ist eine durchdachte Planung und Realisierung unerlässlich.

Im Folgenden werden einige Schlüsselfunktionen vorgestellt, die bei der Einrichtung eines Grids beachtet werden sollten.

4.1 Management System

Jedes Grid System besitzt einige Management Komponenten. Diese haben unterschiedliche Aufgaben.

Zum Beispiel die Überwachung der für das Grid verfügbaren Ressourcen und der Mitglieder des Grids. Diese Informationen werden dann als Hauptgrundlage für die Jobvergabe herangezogen.

Darüber hinaus wird die Verfügbare Leistung und die Auslastung eines jeden Knotens vom Management System überwacht. Anhand dessen erfolgt die Jobplanung im Grid und es kann so der allgemeine Zustand des Grids beurteilt werden. Weiterhin werden diese Messungen für die Erstellung von Statistiken sowie die Aufzeichnung von accounts über die Nutzung von Grid Ressourcen verwendet.

Weiterentwickelte Grid Management Software besitzt außerdem die Fähigkeit des „automatic computing“ / „recovery oriented computing“. Das bedeutet, dass diese Software eine automatische Fehlerbehebung beherrscht, um die im Grid auftretende Probleme wie z.B. der Ausfall eines Gridbereiches, zu kompensieren und so alternative Wege finden zu können, auf denen die Arbeit dennoch erledigt werden kann.

4.2 Donor Software

Jeder Rechner, der Ressourcen innerhalb des Grids zur Verfügung stellt, muss als Mitglied dieses Grids eingetragen sein und eine Software, die Donor

Software installiert haben, die die Verwaltung der Ressourcen dem Grid gegenüber erledigt.

In der Regel ist eine Autorisierung notwendig, bevor der Rechner in ein Grid eingetragen wird. Diese Autorisierung kann auf verschiedenen Wegen erfolgen, z.B. über das Betriebssystem oder über ein eigenes System im Grid. Einige Grid Systeme erlauben es jedem, sich dem System anzuschließen. In anderen Systemen ist es wiederum für jeden Teilnehmer möglich, Aufträge zur Bearbeitung an das Grid zu senden.

Solche Systeme sind jedoch bezüglich des Sicherheitsaspektes sehr zweifelhaft, denn eine Kontrolle der in Auftrag gegebenen Berechnungen ist sehr wichtig. So könnte jede Person bei Bedarf ihre Atombombenberechnungen an das öffentlich Grid senden, was durch eine Kontrolle jedoch verhindert werden kann.

Die Donor Software überwacht die Auslastung und die Menge der zur Verfügung gestellten Ressourcen eines Rechners. Diese Information wird hierarchisch an die Management Software des Grid nach oben weitergereicht und dazu verwendet, die Ressource entsprechend zu nutzen.

Am wichtigsten ist jedoch, dass die Software für die Annahme und Ausführung eines Auftrags des Grid Management Systems zuständig ist. Wenn ein Nutzer einen Bearbeitungsauftrag an das Grid sendet, muss die Grid Management Software diesen Auftrag an die Grid Donor Software weiterreichen. Der Auftrag wird verarbeitet und das Ergebnis zurück an den Sender geschickt.

4.3 Submission Software

In der Regel kann jeder Rechner eines Grids dazu genutzt werden, Aufgaben an das Grid zu senden. In einigen Grid Systemen ist diese Funktion in einem separaten Software enthalten, der Submission Software. Diese Software ist auf „submission nodes“ oder „submission clients“ installiert.

4.4 Distributed grid management

Größere Grids sind in einem hierarchischen Aufbau organisiert. Der Vorteil dieser Organisation liegt in der Steigerung der Flexibilität des Grids.

Einzelne Rechner werden in dieser Hierarchie in „Cluster“ zusammengefasst, die wiederum in größeren Clustern zusammengefasst sein können.

Die Aufgaben, die von einem Grid bearbeitet werden sollen, werden von einem zentralen job scheduler an die einzelnen Cluster geschickt. Von dort aus werden die Aufträge von Hierarchiestufe zu Hierarchiestufe bis zu den Rechnern nach unten weitergegeben. Das Ergebnis gelangt am Ende über den gleichen Weg zurück zum job scheduler und von dort aus zum Management System.

4.5 Schedulers

Die meisten Systeme enthalten eine Art von job scheduling Software. Diese Software ermittelt die Rechner, auf denen der von einem Nutzer eingesandte Auftrag abgearbeitet werden kann.

Im einfachsten Fall werden die Aufträge im Round- Robin Verfahren an die Ressourcen verteilt. Stehen zum Beispiel die Ressourcen 1,2,3 und 4 zur Verfügung, so werden diese in einem Zyklus von 1,2,3,4,1,2,3,4 usw. beauftragt.

Einige scheduler beinhalten ein job priority System, mit dem die Wichtigkeit eines Auftrags berücksichtigt werden kann.

In einem solchen System können manchmal auch policies enthalten sein, mit denen Berechtigungen bezüglich der Aufgaben oder Ressourcen eingeschränkt werden können. Z.B. kann mit einer policy festgelegt werden, dass Aufträge nur zu bestimmten Tageszeiten ausgeführt werden sollen.

Scheduler können auf die aktuelle Grid Auslastungen reagieren, indem Sie Informationen über die aktuelle Nutzung der Ressourcen auswerten und so feststellen, welche Ressourcen auf neue Aufträge warten.

Auch scheduler können hierarchisch organisiert sein. Z.B. kann ein meta-scheduler seinen Auftrag erst an einen cluster scheduler weiterreichen anstelle dies direkt an einen Rechner zu tun.

Höherentwickelte scheduler können Aufgaben überwachen und so, z.B. bei Verlust einer Aufgabe, diese erneut an einen anderen Rechner zur Bearbeitung schicken.

4.6 Communications

Ein Grid System kann auch Software enthalten, die die Kommunikation innerhalb von Aufträgen ermöglicht.

Wenn z.B. ein großer Auftrag in kleinere Unteraufträge geteilt wird, diese separaten Unteraufträge jedoch während den Berechnungen Informationen miteinander austauschen müssen, ist es notwendig die anderen Unteraufträge lokalisieren zu können, eine Verbindung aufzubauen und die Informationen auszutauschen.

Diese Kommunikation wird oft durch das *Open standard Message Passing Interface* (MPI) ermöglicht.

4.7 Observation, management, measurement

In Donor Software ist normalerweise ein „load sensor“ enthalten, der die Auslastung und Aktivität der einzelnen Rechner misst.

Einige Grid Systeme stellen die Möglichkeiten zur Verfügung, eigene „load sensors“ zu implementieren, um auch andere Ressourcen (z.B. CPU, Speicherplatz) überwachen zu können.

Diese Informationen sind für das scheduling, aber auch für eine Analyse des Nutzungsmusters des Grids nützlich. Auch können diese Informationen für Kontozwecke genutzt werden, als Basis für den Ressourcenhandel.

5. Aktueller Stand/ Entwicklung

5.1 Aktueller Stand

Grid Computing erlangt zur Zeit eine massive Unterstützung von Institutionen aller Art, wie der Forschung (CERN, NASA), der IT Industrie (SUN, Microsoft, IBM, HP) und Regierungen (EU). Diese Unterstützung wird in Zukunft noch mehr zunehmen.

Darüber hinaus machen immer mehr Unternehmen ihre technisch-wissenschaftlichen Anwendungen dem Trend nach „gridfähig“ (z.B. Microsoft, Sun, HP, Oracle, u.v.a.).

5.2 Entwicklung

Kurzfristig wird die allgemeine Unterstützung im Bereich des Grid Computing immer mehr zunehmen.

Gerade durch das große LHC Projekt des CERN kann der Begriff des Grid Computing und die damit verbundene Technik immer mehr Verbreitung erlangen.

Die Grid Technologie wird mittelfristig in allen Betriebssystemen integriert sein (oder zumindest optional verfügbar sein). Erste Implementierungen existieren hierfür bereits (z.B. Globus Toolkit, Data Grid Projekt, BOINC Software).

Langfristig könnte sich so etwas wie ein WWG (World Wide Grid) etablieren, was jedoch zum jetzigen Zeitpunkt noch in weiter Ferne liegt.

Ob der Hype anhalten wird und dieser Technologie der Durchbruch gelingt, oder ob in absehbarer Zukunft von den Thema Grid Computing niemand mehr spricht, wird sich in ein paar Jahren entscheiden.

6. Literatur- und Quellenverzeichnis

- V. Berstis (2002): “Fundamentals of Grid Computing”. IBM Corp..
<http://www.ibm.com/redbooks>
- I. Foster, C. Kesselman (1999): “Computational Grids“. In: The Grid: Blueprint for a New Computing Infrastructure. Morgan-Kaufman.
<http://www.globus.org/research/papers/chapter2.pdf>
- I. Foster, C. Kesselman, S. Tuecke (2001): “The Anatomy of the Grid: Enabling Scalable Virtual Organizations” International J. Supercomputer Applications, 15(3).
<http://www.globus.org/research/papers/anatomy.pdf>
- K.Krauter, R.Buyya, M.Maheswaran (2001): “A taxonomy and survey of grid resource management systems for distributed computing”. John Wiley and Sons Ltd..
- D. de Roure, M. A. Baker, N. R. Jennings, N. R. Shadbolt (2003): “The evolution of the grid”. In: F. Berman, G. Fox, and T. Hey. Grid Computing: Making the Global Infrastructure a Reality. John Wiley and Sons Ltd..
<http://www.ecs.soton.ac.uk/%7Enrj/download-files/evolution-of-grid.pdf>
- Christoph Stoidner (2003): „Grid Computing“ Seminararbeit, FH Giessen
<http://homepages.fh-giessen.de/~hg51/Seminar-0304/Grid-Computing.PDF>
- Ramin Yahyapour (2002): „Grids vernetzen Rechner“ In: Computer Postille, 12. Jahrgang, Nr 4, Uni Dortmund
<http://www.hrz.uni-dortmund.de/computerPostille/Dezember2002/011.html>
- Wikipedia
<http://de.wikipedia.org/>
- CERN Press Office Photo Selection
<http://info.web.cern.ch/Press/PhotoDatabase/>
- Seti@Home Projekt
<http://setiathome.ssl.berkeley.edu/>
- Überblick über Grid Projekte
Distributed Computing Info
<http://distributedcomputing.info/>
Rechenkraft.Net
<http://www.rechenkraft.net>
- The Globus Alliance
www.globus.org