

Seminar: Mikroskopie des Universums
WS 2008/09

Web & Grid

**Wie weltweite Forschungsk Kooperation
ermöglicht wird**

Ausarbeitung des Vortrags

Andreas Herten

Betreuer: Dr. Andreas Nowack

Vortrag: 11.11.2008

Inhaltsverzeichnis

1	Einleitung	3
2	Web	3
2.1	Vom der Idee zur Umsetzung	3
2.2	Geschichte	3
2.3	Einfluss des Webs	4
3	Grid	5
3.1	Motivation für Grid-Computing	5
3.1.1	Vergleich mit Supercomputern	5
3.2	Ideen hinter Grid-Computing	6
3.2.1	Vergleich zwischen Web und Grid	6
3.3	Geschichte	7
3.4	Verschiedene Arten von Grids	7
3.4.1	Private und öffentliche Grids	8
3.4.2	Kommerzielle Grids	8
3.4.3	Wissenschaftliche Grids und Projekt-Grids	9
3.4.3.1	EGEE	9
3.5	Worldwide LHC Computing Grid	10
3.5.1	Geschichte des LCGs	10
3.5.2	Strukturelle Grundlagen des WLCGs	11
3.5.3	Technische Strukturen im Grid	11
3.5.3.1	Virtuelle Organisationen	11
3.5.3.2	Middleware	11
3.5.4	Zahlen und Daten zum WLCG	12
3.5.5	Funktionsweise und Aufbau des WLCGs	13
3.5.5.1	TIER –1: Intelligente Detektoren und Counting Room	13
3.5.5.2	TIER 0: CERN Rechenzentrum	14
3.5.5.3	TIER 1: 12 Zentren in der ganzen Welt	14
3.5.5.4	TIER 2: Universitäten und Forschungseinrichtungen	15
3.5.5.5	TIER 3, TIER 4	17
3.5.6	Jobverteilung mit Fairshare	17
3.5.7	Weitere Kooperationsmöglichkeiten	17
4	Ausblick	17
4.1	WLCG	17
4.2	Andere Ansätze	19
5	Zusammenfassung	19
6	Quellen	19
6.1	Literaturquellen	19
6.2	Bildquellen	20

1 Einleitung

An immer größeren und komplexeren Experimenten arbeiten immer mehr Wissenschaftler, die zusätzlich auf der ganzen Welt verstreut angesiedelt sind. Die Daten der Experimente müssen zu den Wissenschaftlern gelangen und dort zeitnah ausgewertet werden.

Das World Wide Web setzte dafür vor 15 Jahren den Grundstein und bereitete den Weg für die Grid-Technologie – in gewisser Weise ein Nachfolger, letztendlich aber etwas ganz anderes.

In dieser Ausarbeitung wird es zuerst kurz um das World Wide Web gehen. Der weitaus größere Teil wird sich mit Grid-Computing beschäftigen, wobei das Worldwide LHC Computing Grid als interessantes Beispiel näher erläutert wird. Schließlich wird es am Ende noch einen kurzen Ausblick geben.

2 Web

2.1 Vom der Idee zur Umsetzung

Tim Berners-Lee, ein ehemaliger Angestellter am CERN, unterbreitete die Idee des World Wide Webs bereits 1989 als Proposal. Er suchte nach einer Möglichkeit, Forschungsergebnisse auf internationaler Ebene auszutauschen – dies möglichst automatisiert. Um die Illustration von wissenschaftlichen Arbeiten zu ermöglichen sollten Bild und Text im Verbund übertragbar sein. Das Web sollte dabei für alle eine gemeinsame Basis darstellen.

Es waren ihm die ersten beiden Ws des Namens besonders wichtig: Von überall aus der Welt kann man auf das neu geschaffene Netz zugreifen.

Den Prinzipien des CERNs sollte das WWW treu sein, Berners-Lee baute in seiner späteren Ausführung das Web komplett mit offenen Standards auf. Ein jeder, ob Forscher oder Privatperson, hat so die Möglichkeit nicht nur Browser zum Anzeigen von Webseiten zu programmieren, auch die Webseiten selbst können von jedermann erstellt werden. Man stelle sich nur vor, wie das Web heute aussähe, bestünde nicht die Möglichkeit, ohne Gebühren darauf zu zugreifen und Inhalte zu kreieren.

Bereits fest in seiner Idee des Webs verankert waren Hyperlinks. Einzelne Webseiten können sich unter einander verbinden. Dazu müssen sie allerdings nicht das Ziel dieses Links erst um Erlaubnis bitten; die Hyperlinks sind unidirektional.

Das Web stellte damit eine Reihe neuer fortschrittlicher Ideen vor – die vielleicht wichtigste ist der Verbund von Text und Bild. Dies unterschied es vom damals ebenfalls verfügbaren *Gopher*.

2.2 Geschichte

Die Entwicklung des Webs ging rasant von statten: Berners-Lee machte **1989** seinen Vorschlag, bereits ein Jahr später stellte er einen ersten Prototypen eines Webbrowsers vor. Er nannte diesen, auf NeXT-Systemen laufenden Browser erst *WorldWideWeb*, später dann *Nexus*, zur Vermeidung von Verwechslungen. **1991** wurde der erste Webserver in den USA am SLAC¹ in Betrieb genommen. **1992** hatten sich schon mehrere andere Webbrowsers

¹SLAC = **S**tanford **L**inear **A**ccelerator **C**enter

entwickelt, da sich der ursprüngliche Prototyp von Berners-Lee nur sehr eingeschränkt benutzen ließ. In dieser Phase wurde auch der bekannte Mosaic-Browser entwickelt, aus dem erst der Netscape Communicator und schließlich der Firefox-Webbrowser hervorgegangen sind.

1993 stellte man das Web der Öffentlichkeit vor. Ein Jahr später gab es bereits mehr als 10.000 Server weltweit.

2.3 Einfluss des Webs

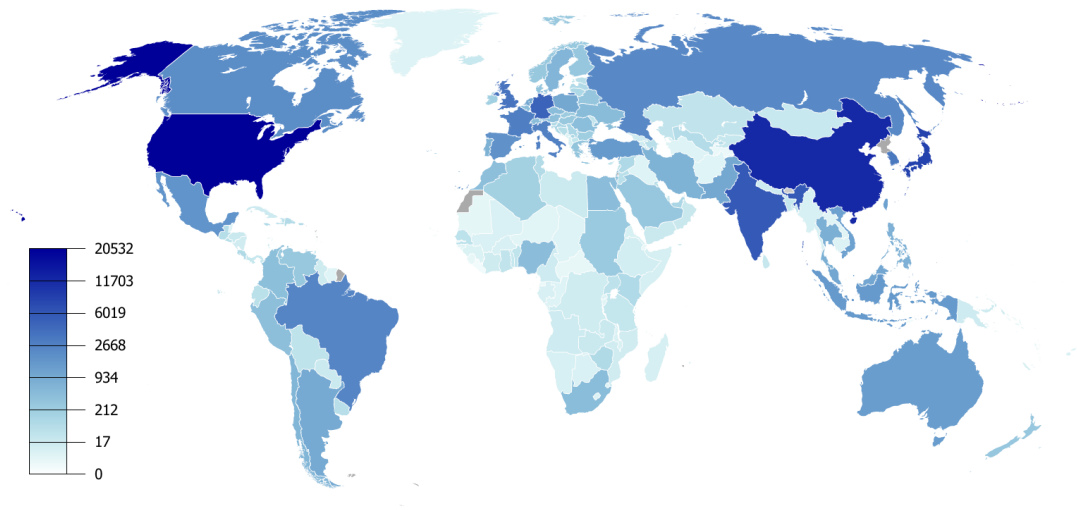


Abbildung 1: Anzahl der weltweiten Internetnutzer. In Tausend.

Das Web hat seitdem eine lange Entwicklung hinter sich gebracht. In vielen Bereichen hat es das Arbeiten und auch das Leben teils drastisch verändert. Die moderne Gesellschaft sähe ohne das Internet heute stark anders aus.

Im **wirtschaftlichen Sektor** sei hier das Beispiel *Google* angeführt. 1998 gründete sich dieser Suchmaschinendienst und entwickelte sich in den Jahren darauf zum Marktführer auf dem Gebiet. 2007 hatte Google einen Umsatz von 16 Milliarden US-Dollar. Der Marktwert Googles im gleichen Zeitraum entspricht 220 Milliarden US-Dollar – damit ist Google vor dem Konsumgütergiganten *Procter & Gamble*, sowie vor dem weltweit größten Automobilhersteller *Toyota* angesiedelt. Der Wert der Marke Google entspricht 86 Milliarden US-Dollar. Dies ist ebenfalls mehr als der Wert der Marken *General Electric*, *Microsoft* oder *Coca-Cola*. Google ist sowohl bei Markt-, wie auch bei Markenwert Spitzenreiter. 2008 beschäftigt Google über 20.000 Mitarbeiter.

Weltweit greifen momentan ca. 1,6 Milliarden Menschen auf das Web zu. Bevölkerungsstarke Länder wie Indien oder China befinden sich dabei noch im Anfangsstadium der Internetnutzung, hier werden in der Zukunft noch besonders viele Menschen das Internet verwenden.

In **Deutschland** sind 70% aller Erwachsenen online, dies entspricht ca. 42 Millionen Nutzern.

Das Web ist also eine wichtige Möglichkeit, Informationen ortsunabhängig auszutauschen.

3 Grid

3.1 Motivation für Grid-Computing

Wissenschaftliche Experimente liefern immer mehr Daten. So wird beispielsweise der LHC nach seinem Start jährlich ca. 15 Petabyte Daten liefern. Dies entspricht einem Stapel von hüllenlosen CDs, der sich 25 km in den Himmel streckt². 15 Petabyte sind 1% der jährlichen produzierten Informationen der kompletten Menschheit.

Neben der Verwaltung dieser riesigen Datenmenge muss auch sicher gestellt sein, dass sie ausgewertet werden können. Aufwändige Berechnungen sind dazu nötig. Diese Berechnungen finden nicht nur in Bereichen der Physik statt, sondern erstrecken sich auch auf andere Bereiche, wie z.B. Crashtests oder Simulationen von Wetter.

3.1.1 Vergleich mit Supercomputern

Generell sind diese Problemstellungen auch von großen Supercomputern lösbar, die zentral an einem bestimmten Ort stehen. Verteiltes, delokalisiertes Rechnen mit einem Grid-System bietet aber einige Vorteile.

- Bei einem internationalen Projekt, wie es das LHC ist, nehmen viele Nationen teil. Diese Nationen wollen ihre finanziellen Ressourcen natürlich lieber auf lokalem Boden ausgeben und nicht in der Schweiz für ein Rechenzentrum des CERNs.
- Das technische Know-How bleibt im jeweiligen Land vorhanden; die angeschafften Ressourcen können dynamisch über das Projekt hinaus genutzt werden.
- Für eine große Zahl von Rechenmaschinen muss auch die nötige Infrastruktur gebaut werden. Dies fängt bei einfachen Serverschränken an und geht bis zu ganzen Gebäudekomplexen für große Rechenzentren. Auch diese Infrastruktur kann über das spezielle Projekt hinaus genutzt werden und wird von lokalen Unternehmen vor Ort erstellt.
- Ein Computer-Cluster ist gegenüber eines Supercomputers wesentlich besser skalierbar. Es können einzelne Teile hinzugefügt werden, ohne auf die teils starken Beschränkungen eines Großcomputers achten zu müssen. Mit entsprechender Software kann fast jede Art Computer und Computerbauteil eingegliedert werden.
- Wenn sich nicht auf ein zentrales, großes Rechenzentrum verlassen werden muss, sondern viele kleine Rechenzentren zur Verfügung stehen, dann ist auch die Ausfallsicherheit als Gesamtes betrachtet erhöht. Backups können ebenfalls im kleinen Maßstab leichter durchgeführt werden.

²Bei 700 MB pro CD und einer Dicke einer Plasticscheibe von 0,12 cm. 15 PB wurden zur Vereinfachung als 15.000.000.000 MB angenähert.

Natürlich besitzt ein Grid-Projekt auch Nachteile gegenüber einem Supercomputer. Grid-Bestandteile möchten verbunden werden, was nicht nur finanzieller und technischer Mehraufwand bedeutet, sondern auch zusätzliches Störpotenzial beinhaltet. Die Strecken zwischen Grid-Zentren sind ausgedehnt, was zu Verzögerungen im Datenverkehr führen könnte.

Da in einem Supercomputer die Komponenten genau aufeinander abgestimmt sind, ist hier meist höhere Rechenleistung (bei vergleichbaren Hardwarekomponenten) vorhanden. Backups bei einem Supercomputer durchzuführen bereitet zwar mehr Aufwand, dafür sind sie danach in einer zentralen Position vorhanden. Die Ausfallsicherheit kann ebenfalls mit entsprechendem Aufwand bis nahe an die 100% getrieben werden.

3.2 Ideen hinter Grid-Computing

Grid-Computing greift auf eine Anzahl Ideen zurück:

- Es findet Ressourcenteilung statt – egal ob es sich um die Teilung der Ressource „Rechenleistung“, oder der Ressource „Speicherplatz“ handelt.
- Ein sicherer Zugang zum Grid ist gewährleistet; das Grid selbst ist ebenfalls nach außen hin abgeschottet und damit sicher.
- Rechenanforderungen, so genannte Jobs, werden intelligent verteilt – bei der Notwendigkeit des Zugriffs auf große Datenmengen beispielsweise werden die Jobs zu den Daten geschickt, nicht anders herum. Diese Jobverteilung findet im Idealfall völlig automatisiert statt.
- Das Grid beruht auf offenen Standards. Dies macht es, ganz analog zum Web, erweiterbar und flexibel.

Einer der Köpfe hinter der Idee des Grid-Computing ist **Ian Foster**. Er fasst die Idee des Grids in einem Satz zusammen:

Grid computing is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations.³

3.2.1 Vergleich zwischen Web und Grid

Ist das Web eine ortsunabhängige Möglichkeit, Informationen auszutauschen, so lässt sich analog für das Grid sagen, dass es eine Möglichkeit ist, ortsunabhängig Rechenleistung und Datenspeicher auszutauschen. Letzteres ist natürlich indirekt auch wieder Information. Heutzutage, in Zeiten immer schnellerer Internetverbindungen und immer leistungsfähiger Arbeitsplatz-Computer ist das Web in gewisser Weise auch zum Datenspeicher geworden. Man denke nur an YouTube.

³Aus: Ian Foster & Karl Kesselman: „The Grid“, S.23

Web	Grid
Ortsunabhängig	Ortsunabhängig
Information	Rechenleistung & Datenspeicher
Datenspeicher	→ Information

Tabelle 1: Vergleich zwischen Web und Grid.

3.3 Geschichte

Nimmt man es genau, so fängt die Geschichte des Grid-Computings bereits sehr früh an, denn die Idee des verteilten Rechnens ist keine Neue. Die aller ersten Rechner waren Zentralcomputer, auf die mit kleinen Terminals zugegriffen werden konnte. Diese Großrechner hatten gewissermaßen auch gridähnliche Strukturen.

Direkte Vorgängerprojekte des Grids sind die **1995** ins Leben gerufenen US-Projekte FAFNER und I-WAY. In Letzterem war bereits erwähnter Ian Foster involviert, der schließlich mit seinem Kollegen Carl Kesselman **1998** die Idee des Grids in einem Buch publizierte: *The Grid: Blueprint for a New Computing Infrastructure*. Hier tauchte erstmalig der Begriff des Grids auf. Foster und Kesselman wählten ihn in Analogie zum Power Grid, also zum Stromnetz, um die Idee der delokalisierten Rechenleistung „aus der Steckdose“ zu verdeutlichen. Das Buch wird heute als die „Bibel“ des Grid-Computings bezeichnet.

Aus ihrer Idee heraus gründeten Foster und Kesselman **1999** das *Globus Project*. Unter diesem Dach wird seitdem Software das Grid betreffend weiterentwickelt und die Idee des Grids propagiert. Unter der Software ist zum Beispiel das *Globus Toolkit*, ein Open-Source Framework zum Aufbau eines eigenen Grid-Netzwerks.

Ebenfalls **1999** ging das erste öffentliche Grid-Projekt an den Start. *SETI@home* startete zur durch heimische Rechner unterstützen Suche nach außerirdischem Leben. Dies war der erste große Test der Grid-Technologie, ein „proof of concept“. Dieser Test verlief sehr gut, so dass heutzutage eine Vielzahl verschiedener Grid-Systeme vorhanden sind. Als Beispiel sei das Projekt *Folding@home* erwähnt, das von der Stanford University (Kalifornien / USA) aus verwaltet wird und sich mit biochemischen Problemen der Proteinfaltung beschäftigt. Im Jahr **2008** konnte Folding@home über eine gesamte Rechenleistung von 4 PFLOPS⁴ verfügen. Zum Vergleich: 2008 durchbrach der erste Supercomputer in der top500.org-Liste die 1-PFLOPS-Grenze und steht mit eben diesem Wert an der Spitze. Folding@home greift dabei nicht nur auf die Ressource Heimcomputer zurück, sondern benutzt auch Grafikchips („GPU“) und Spielekonsolen (z.B. PlayStation 3) zum Berechnen der Faltungen.

3.4 Verschiedene Arten von Grids

Im Gegensatz zum Web gibt es nicht das – also ein einzelnes – Grid. Es gibt eine große Anzahl verschiedener Arten von Grids. Sie lassen sich auf viele Weisen klassifizieren und einordnen. Es gibt nationale Grids, Institutions-Grids, Projekt-Grids und Goodwill- bzw. Volunteer-Grids – aber dies ist nur eine Möglichkeit der Einordnung.

⁴(P)FLOPS = (Peta) Floating Point Operations Per Second = (10¹⁵) Gleitkommaoperationen (Multiplikation, Addition) pro Sekunde.

3.4.1 Private und öffentliche Grids

Bei privaten oder öffentlichen Grids kann ein jeder beitreten, wenn er denn will. Der Schlüssel hier zu ist öffentlich zugängliche, einfach verwendbare *Middleware* (hierzu in 3.5.3.2 mehr). Ein paar der bekannteren öffentlichen Grid-Projekte sind:

- **World Community Grid:** Ein Grid-Projekt, das verschiedene non-profit-Projekte berechnet. Der Hauptsponsor ist *IBM*, wo auch die nötige Software auf Grundlage der Technik des Vorgängerprojekts *United Devices* entwickelt wird. *United Devices* engagierte sich von 1999 bis 2007 im Bereich der Krebsforschung.
- **distributed.net:** Ein großes Projekt, dessen erklärtes Ziel es ist, verteiltes Rechnen weiterzuentwickeln und den Einsatz entsprechender Technologien zu fördern. Momentan gibt es zwei große Rechenprojekte: Im Bereich der Kryptografie werden RSA-Schlüssel⁵ immer höherer Komplexität versucht zu entschlüsseln; im Bereich mathematischer Theorien wird ein Golomb-Lineal⁶ der Ordnung 25 berechnet.
- **SETI@home:** Dieses bereits erwähnte Projekt ist das wohl bekannteste Projekt des verteilten Rechnens. Hier wird versucht, aus Daten von Radioteleskopen Anzeichen für extraterrestisches Leben zu finden.
- **Folding@home:** Ebenfalls bereits erwähnt, ist dieses das momentan größte Grid-Projekt und beschäftigt sich mit biochemischen Problemen der Proteinfaltung.
- **FightAIDS@home:** Ein Forschungs-Grid, was nach Medikamenten zur AIDS-Behandlung sucht.
- **LHC@home:** Der Ableger des *LHC Computing Grids*, der sich an den Anwender am heimischen PC richtet. Die Aufgabe bestand in der Vergangenheit darin, Simulationsrechnungen für das große LHC Computing Grid durchzuführen. Mittlerweile gibt es keinen direkten Nutzen mehr für LHC@home, es wird nach einer neuen Möglichkeit gesucht, Jobs zum Rechnen zu erhalten.

3.4.2 Kommerzielle Grids

Kommerzielle Grids sind eine denkbare und sehr gute Möglichkeit, die Rechenleistung des verteilten Rechnens zu nutzen. Momentan ist dieser Zweig der Anwendung aber noch in der Phase der Forschung – es gibt erst einige Pilotprojekte. Eines dieser Pilotprojekte wird vom deutschen D-Grid mit dem Automobilhersteller *BMW* zusammen betrieben. Als Forschung-Wirtschaft-Kooperation wird hier an der Simulation von Crash- und Windkanaltests gearbeitet, außerdem werden komplexe Produktionsstraßen simuliert.

Bei kommerziellen Grids gibt es eine Reihe außergewöhnlicher Herausforderung in Bezug auf öffentliche, aber auch auf Forschungs-Grids.

Da ist zum Einen die Preisgestaltung und die Abrechnung. Wird die genutzte CPU-Leistung bezahlt? Wird ein Zeitraum abgerechnet? Wird verbrauchter Festplattenspeicherplatz in Rechnung gestellt?

⁵Wikipedia: RSA-Kryptosystem (<http://de.wikipedia.org/wiki/RSA-Kryptosystem>)

⁶Wikipedia: Golomb-Lineal (<http://de.wikipedia.org/wiki/Golomb-Lineal>)

Besonders sensibel ist das Thema Vertraulichkeit: Nicht nur die Frage „Wer rechnet potenziell?“, sondern auch die Frage „Wer rechnet gerade?“ ist von entscheidender Bedeutung. Ein hypothetischen Szenario sei hier angedacht: Ein Grid-System hat sich auf das Berechnen von Windkanal-Simulationen spezialisiert. Zwei konkurrierende Automobilhersteller A und B haben beide Verträge zur Nutzung dieses. Kurz vor Markteinführung eines neuen Automobils von Hersteller A bindet dieser viel Rechenleistung auf dem Grid-System. Hersteller B vermutet Probleme bei der Markteinführung und denkt darüber nach, sein neues Automobil etwas schneller fertig zu stellen und es noch vor A auf den Markt zu bringen. Allein die Information, das gerechnet wird könnte so sehr viel wert sein. Allgemein ist die Sicherheit der Jobs und der Daten bei industriellen Interessen von sehr hoher Bedeutung.

3.4.3 Wissenschaftliche Grids und Projekt-Grids

Viele Anwendungen von Grid-Systemen finden sich im Bereich der wissenschaftlichen Grids. Dies sind meist projektgebundene Grids.

- In Deutschland ist ein Vertreter das **D-Grid**, welches einer Vielzahl von wissenschaftlichen Projekten die Basis für eine Grid-Berechnung bietet.
- Im Vereinigten Königreich gibt es den **National Grid Service**, ein rein akademischer Grid-Dienst.
- Für den internationalen Kernfusionsreaktor ITER in Frankreich ist man gerade dabei, die notwendige Recheninfrastruktur zu planen. Im **EUFORIA project**⁷ plant man auch schon fest mit einem Grid zur Datenbewältigung.
- Das größte amerikanische Grid-Projekt ist das 2004 gegründete **Open Science Grid**.
- Ein spezielles Grid im Vereinigten Königreich ist das **GridPP**, ein Grid der Teilchenphysik (**P**article **P**hysics) des Vereinigten Königreichs.
- In Skandinavien gibt es das **NorduGrid**, ein Zusammenschluss von fünf Universitäten aus Dänemark, Finnland, Norwegen und Schweden.
- In ganz Europa ist das **EGEE** verbreitet.

3.4.3.1 EGEE

Das größte Grid Europas ist das EGEE. Genauer EGEE III, da es seit Mitte diesen Jahres in seiner dritten von drei Projektphasen ist. EGEE steht für **E**nabling **G**rids for **E**-science**E**. EGEE wird von der EU finanziert. Dementsprechend war es auch zu Beginn ein rein europäisches Grid (und das letzte E im Namen stand für Europa), wurde aber sehr schnell internationaler.

Bei EGEE machen 300 Computerfarmen mit, die in 50 Ländern organisiert sind. Über 10.000 Nutzer können bei EGEE auf über 80.000 CPU-Kerne zugreifen.

Eigens zum Erlernen des Umgangs mit EGEE, ohne dabei den Betriebsfluss im Netz stören

⁷EUFORIA = **EU** Fusion **fOR** Iter **A**pplications

zu müssen und Ressourcen zu binden, wurde eine Lernumgebung namens ICEAGE⁸ ins Leben gerufen.

Ein Beispielprojekt im EGEE-Grid ist WISDOM. Dieses Projekt berechnet Medikamentenzusammensetzungen, so dass beim teuren Testen im Labor bereits eine Vorauswahl auf Funktionalität durch den Computer getroffen wurde. 2006 wurde hier ein Impfstoff zu H5N1 („Vogelgrippe“) erforscht.

Über die hier beispielhaft erwähnten Projekte hinaus gibt es noch viele, viele mehr. Im Wikipedia-Artikel [List of distributed computing projects](#)⁹ oder im [GridCafé des CERNs](#)¹⁰ gibt es einen Überblick über Projekte.

3.5 Worldwide LHC Computing Grid

Das vielleicht wichtigste Grid wurde bisher nicht erwähnt. Das WLCG, kurz für **Worldwide LHC Computing Grid**, ist das Grid, was die gewaltigen Datenmengen des Large Hadron Colliders speichern und verarbeiten soll.

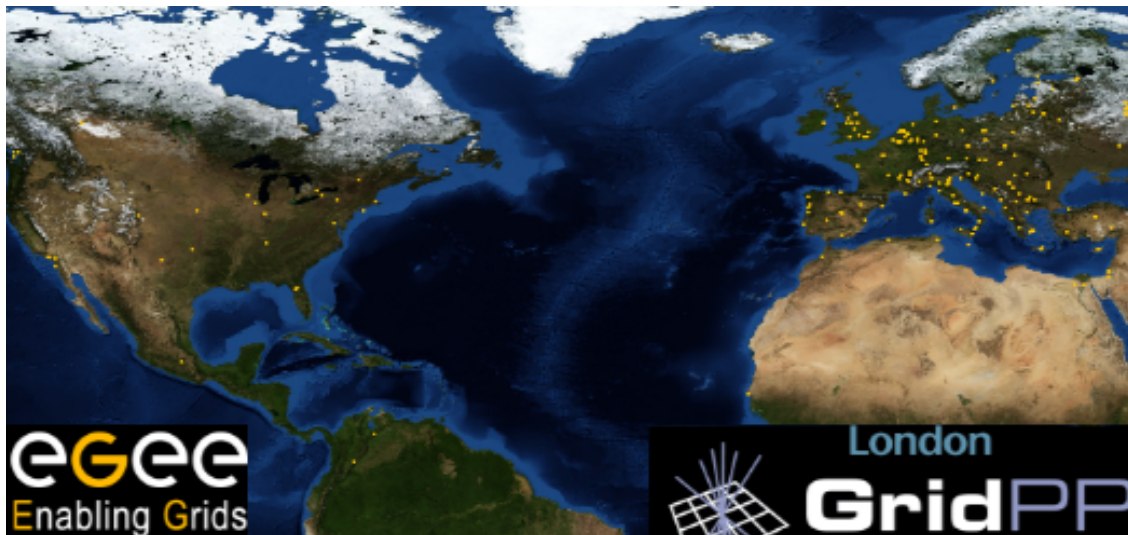


Abbildung 2: Ein Auszug aus dem GridPP-LCG-Monitor. Die gelben Punkte repräsentieren WLCG-Sites.

3.5.1 Geschichte des LCGs

- 1998 – 2000: Die MONARC¹¹-Studie wird in Auftrag gegeben. Diese untersucht, wie die benötigte Rechenleistung für die kommenden LHC-Experimente aufgebracht werden kann. In dieser Phase wird bereits die Aufteilung in verschiedene, verteilte Ebene („TIERs“) nahegelegt.

⁸ICEAGE = International Collaboration to **E**xtend and **A**dvan**C**e **G**rid **E**ducation

⁹Wikipedia (Englisch): List of distributed computing projects

(http://en.wikipedia.org/wiki/List_of_distributed_computing_projects).

¹⁰<http://gridcafe.web.cern.ch/gridcafe/gridprojects/projects.html>.

¹¹MONARC = **M**odels **O**f **N**etworked **A**nalysis at **R**egional **C**entres for LHC Experiments

- 2001: Das European DataGrid (EDG) wird gegründet. Es ist das erste europäische Grid-Projekt und der Vorläufer zum EGEE, welches 2004 aus dem EDG hervorgeht.
- 2002: Am CERN wird das LCG gegründet.
- 2005: Zu Simulationen wird LHC@home ins Leben gerufen.
- 2008 (Oktober): Das WLCG geht offiziell an den Start. Da mittlerweile mit dem amerikanischen Open Science Grid ein großer, nicht-europäischer Partner am LCG teilnimmt, hat man es in *Worldwide LCG* umbenannt.

3.5.2 Strukturelle Grundlagen des WLCGs

Beim LCG handelt es sich nicht um ein eigenes, großes Grid, welches komplett aus eigenen Strukturen besteht. Vielmehr greift das LCG zu großen Teilen auf bereits bestehende Strukturen zurück.

Dazu nimmt es die bereits bekannten Grid-Projekte, die weiter oben angeführt wurden. Das EU-finanzierte EGEE bildet die größte Grundlage, das Open Science Grid hat ebenfalls einen großen Anteil. Darüber hinaus nehmen das GridPP und das NorduGrid am LCG teil.

Innerhalb des CERNs wird auf den Strukturen des *CERN openlabs* aufgebaut. Das CERN openlab ist eine Zusammenarbeit zwischen dem CERN-Rechenzentrum und der Industrie, die dem CERN neueste, eventuell noch nicht marktreife Technologie zur Verfügung stellt. Im CERN wird die Technik dann bis an ihr Leistungsmaximum gebracht und auf „Herz und Nieren“ getestet.

3.5.3 Technische Strukturen im Grid

3.5.3.1 Virtuelle Organisationen

Innerhalb eines Grids werden Nutzer in virtuelle Organisationen eingeteilt. Diese Benutzergruppen können z.B. ein Experiment sein, an dem Benutzer verschiedener Arbeitsgruppen teilnehmen. Die virtuelle Organisation regelt dabei die Benutzerverwaltung. Welcher Benutzer innerhalb der virtuellen Organisation darf welche Aufgaben ausführen (*Autorisierung*) und wie meldet er sich an (*Authentifizierung*).

3.5.3.2 Middleware

Ein zentraler Begriff, wenn es um Grid-Computing geht, ist die *Middleware*. Sie ist die Softwareschnittstelle zwischen den einzelnen Ressourcen eines Grids und bildet damit den Zugang dazu.

Die Middleware verbindet unterschiedliche Rechnerarten und Rechnersysteme, zum Beispiel verschiedene Betriebssysteme, Speichersysteme und CPUs. Etwas überspitzt gesagt macht sie einen losen Computerhaufen zu einem strukturierten Grid.

Eine Auswahl verschiedener, bedeutender Middlewares:

- **UNICORE:** Eine speziell auf Supercomputer zugeschnittene Middleware aus dem Forschungszentrum Jülich. Seit 1997 wird diese entwickelt.

- **Globus Toolkit:** Der Quasi-Standard der Globus Alliance von Foster und Kesselman. Ein Open-Source-Baukasten, aus dem man seine eigene Grid-Middleware zusammenstellen kann.
- **Virtual Data Toolkit:** Diese amerikanische Middleware wurde aus dem Globus Toolkit heraus entwickelt.
- **gLite:** Das EGEE-Vorgängerprojekt EDG erstellte seiner Zeit eine gleichbenannte Middleware, die weiterentwickelt wurde und heute den Namen gLite trägt.
- **BOINC¹²:** Aus Berkeley kommt diese consumer-client-orientierte Middleware. Sie besitzt eine aufwändige, grafische Oberfläche und ist die Plattform für viele Volunteer-Grids, also meist der Grids mit „@home“ im Namen.

gLite

Das Globus Toolkit bietet für die aller meisten Problemstellungen eines Grids Lösungen in Form der Middleware. Einzig allein ein automatisches Job-Routing fehlt: Ein Benutzer, der einen Job ins Grid eincheckt, muss sich selbst darüber Gedanken machen, wo der Job berechnet werden soll. Das ist allerdings eine Arbeit, die man gerade beim Grid vermeiden will.

Die EDG-Middleware löste dieses Problem mit einem Resourcebroker. Dieser hält automatisch Ausschau nach freien Ressourcen im Grid-Netzwerk und schickt den Job dort hin. Aus dieser EDG-Middleware hat sich die Open-Source-Middleware gLite entwickelt.

Middleware im WLCG

Im WLCG kommt ein Mischsystem von Middlewares zum Einsatz.

Auf europäischer Seite wird gLite verwendet, in den USA das Virtual Data Toolkit. Letzteres besitzt ebenfalls einen Resourcebroker, baut aber nicht auf der EDG-Middleware, sondern auf dem Globus Toolkit auf. Zwischen Virtual Data Toolkit und gLite gibt es Gateways, die gewissermaßen zwischen den Middleware-Sprachen übersetzen und so Sub-Grids zu einem großen Grid, dem WLCG, machen.

3.5.4 Zahlen und Daten zum WLCG

- Der LHC verfügt über insgesamt 150.000.000 Sensoren in allen Experimenten.
- In allen Rechenzentren sind insgesamt 100.000 CPUs verfügbar.
 - 20% davon allein im CERN.
- Es sind 60 PB an Festplatten- und 40 PB an Bandlaufwerksspeicherplatz vorhanden.
- Das Rechenzentrum am CERN hat bisher ca. 46 Millionen Euro gekostet, 2008 waren es allein 10 Millionen Euro.
- Für 2012 ist das erklärte Ziel, eine Rechenleistung von insgesamt 340.000 kSI2k¹³.

¹²BOINC = **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing

¹³kSI2k = kilo SPECint2000. Ein Benchmark, zur Analyse der Rechenleistung einer CPU, auf den sich geeinigt wurde. Als Richtwerte: Ein etwas älterer 3-GHz-schneller Pentium-4-Prozessor kommt ungefähr auf einen Wert von 1 kSI2k, bei einem modernen Intel Quad-Core-Prozessor mit 3 GHz kommt jeder Kern auf 3 kSI2k. Siehe <http://www.spec.org/>.

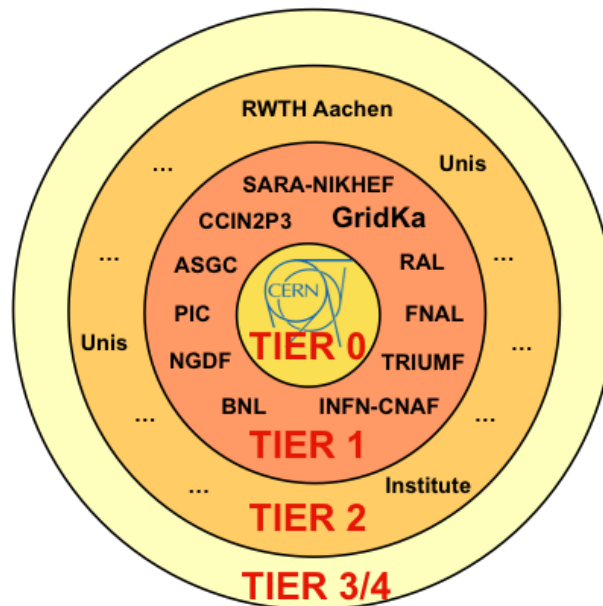


Abbildung 3: Der Aufbau des WLCG in die einzelnen TIERS im Zwiebelschalenprinzip. Die TIER 0 im CERN steht im Zentrum. Die Aachener TIER 2 ist exemplarisch eingezeichnet.

3.5.5 Funktionsweise und Aufbau des WLCGs

Die Datenverteilung und -verarbeitung findet beim WLCG auf verschiedenen Stufen statt. Diese Ebenen werden nach der englischen Übersetzung „TIER“ genannt.

Die TIERS sind über Highspeed-Netzwerke in der ganzen Welt verbunden und können so konstanten und schnellen Datenaustausch miteinander betreiben. Jede TIER-Ebene hat dabei insgesamt gleich viel Rechenleistung.

Im Zentrum dieser Verteilung im Zwiebelschalenprinzip ([Abbildung 3](#)) steht das CERN, was nach außen hin an die externen Rechenzentren verteilt.

3.5.5.1 TIER –1: Intelligente Detektoren und Counting Room

Der Begriff des TIER –1 gibt es in dieser Form nicht. Er wurde von mir gewählt, um deutlich zu machen, dass die Datenverteilung bereits vor dem eigentlichen LCG stattfindet.

Als Beispiel sei hier das ATLAS-Experiment angeführt. Hier wird durch Ereignis Selektion („Event Selection“) die Ereignisrate von 1 GHz auf ca. 200 Hz reduziert.

Von den Detektoren kommen ca. 300 GB/s an Datenstrom. Dieser geht in den Counting Room, wo er durch mehrere Hard- und Softwarefiltermechanismen auf 320 MB/s-Rohdaten reduziert wird.

Die Mechanismen unterscheiden sich von Experiment zu Experiment und finden teilweise schon direkt an den Detektoren im Experiment statt.

Aus den einzelnen Experimenten des LHCs kommt damit ein Datenstrom, wie er in [Tabelle 2](#) aufgeschlüsselt ist. Die Daten werden von hier zum TIER-0-Zentrum weitergeleitet.

Experiment	Datenrate
ATLAS	320 MB/s
CMS	220 MB/s
ALICE	100 MB/s
LHCb	50 MB/s

Tabelle 2: Datenstrom der einzelnen Experimente des LHCs.

3.5.5.2 TIER 0: CERN Rechenzentrum

Das eigentliche Zentrum des LHC Computing Grids ist das Rechenzentrum des CERNs. Über dedizierte 10 GBit/s-Leitungen sind die Counting Rooms an dorthin angeschlossen. Das TIER-0-Zentrum hat hauptsächlich drei Aufgaben:

- Speicherung der Rohdaten auf Festplatten und Bändern.
 - Bänderkapazität: 17 PB¹⁴.
 - Festplattenkapazität: 5,5 PB.
 - Es werden ca. 27 TB pro Tag abgespeichert.
- Event Summary Data¹⁵ (ESD) erstellen.
 - Das Erstellen der ESD ist ein erstes Rekonstruieren und Kalibrieren.
 - Das Rechnen findet in einem eigenen Teil des Rechenzentrum statt, in der CERN Analysis Facility. Hier sind 4700 kSI2k an Rechenleistung vorhanden.
 - Es fallen zusätzlich 140 MB/s, bzw. 10 TB/Tag an, die ebenfalls gespeichert werden müssen.
- Weiterverteilung an die TIER-1-Zentren.

Das TIER-0-Zentrum speichert also und verteilt die Daten der Experimente, samt Zusatzdaten weiter an die unteren TIERS.

3.5.5.3 TIER 1: 12 Zentren in der ganzen Welt

Die TIER-1-Zentren befinden sich in der ganzen Welt verstreut. In Deutschland, Frankreich, Italien, Kanada, Niederlande, Skandinavien, Spanien, Taiwan, UK und zwei Mal in den USA. Am CERN selbst ist ebenfalls ein TIER-1-Zentrum angesiedelt.

Über dedizierte, also extra dafür vorhandene 10-GBit/s-Leitungen sind die TIER-1-Zentren mit dem TIER-0-Zentrum in der Schweiz vernetzt. Untereinander sind die TIER-1-Zentren ebenfalls verbunden.

Ein TIER-1-Zentrum erhält keinen kompletten Datensatz aller Experimente, sondern nur eine vorher festgelegte Auswahl, abhängig davon, welche Forschergruppen auf die jeweilige TIER 1 zugreifen. Insgesamt, wenn man alle TIER 1 nimmt, ist ein vollständiger Datensatz vorhanden.

Die Aufgaben der TIER-1-Zentren sind:

¹⁴Die hier erwähnten Daten stammen von Anfang 2008.

¹⁵Dieser Begriff ist von Experiment zu Experiment unterschiedlich.

- Backup der Daten von TIER-0-Zentrum und TIER-2-Zentren auf Bandlaufwerken und Festplatten.
- Daten rekalisieren...
- ... und reprozessieren. Im Terminus des ATLAS-Experiment ist dies die ESD2.
- Unterverteilung an TIER-2-Zentren.

An dieser Stelle sind insgesamt 38.000 kSI2k Rechenleistung vorhanden, sowie 20 PB Festplatten- und 21 PB Bandspeicherplatz.

TIER 1: GridKa

Der deutsche LCG-Knotenpunkt ist im Forschungszentrum Karlsruhe angesiedelt und heißt GridKa. Mit zwei 10-GBit/s-Leitungen, also insgesamt mit 20 GBit/s ist er an das TIER-0-Zentrum angeschlossen.

Im Jahr 2008 verfügt der GridKa über eine CPU-Leistung von 11.000 kSI2k aus 6500 CPU-Kernen und kann auf 4 PB Festplatten-, sowie 5 PB Bandspeicherplatz zurückgreifen. Das Zentrum befindet sich im Ausbau und erweitert seine verfügbaren Ressourcen ständig. GridKa unterstützt nicht nur alle vier LHC-Experimente, sondern ist auch für vier weitere Nicht-LHC-Experimente zuständig.

3.5.5.4 TIER 2: Universitäten und Forschungseinrichtungen

Die TIER-2-Zentren bilden Universitäten und Forschungseinrichtungen überall auf der Welt. Es gibt insgesamt 140 Standpunkte in 38 Ländern.

Untereinander sind die TIER-2-Zentren über spezielle Wissenschaftsnetze oder das Internet verbunden.

Jedes TIER-2-Zentrum greift auf die Daten der TIER-1-Zentren zu. Es findet keine Datenspeicherung, bzw. langfristige Datenarchivierung in den TIER-2-Zentren statt. Alle Ergebnisse von Berechnungen an den TIER-2-Zentren werden wieder zum Speichern an die zuständige TIER-1-Zentrum geschickt.

Die beiden zentralen Aufgaben der TIER-2-Zentren sind:

- Sie bilden Zugang für ungefähr 7000 Forscher weltweit, die hier ihre eigentlichen Analysen und Berechnungen durchführen.
- Es werden MC-Simulationen durchgeführt, deren Ergebnisse schließlich auch wieder zu TIER-1-Zentren geschickt werden.

Insgesamt sind auf dieser Ebene 46.000 kSI2k CPU-Leistung vorhanden, außerdem 13 PB Festplattenspeicherplatz.

TIER 2: Standorte in Deutschland

Nach Experimenten aufgeschlüsselt sind die Standorte in Deutschland:

- ATLAS
 - DESY Hamburg
 - Freiburg und Wuppertal
 - München
- ALICE
 - GSI Darmstadt
- CMS
 - DESY Hamburg und RWTH Aachen (treten zusammen als ein TIER-2-Standort auf)
 - * Rechenleistung: 600 kSI2k und 450 kSI2k
 - * Speicherplatz: 170 TB und 100 TB (Festplatte)

Insgesamt sind in Deutschland ca. 3500 kSI2k Rechenleistung vorhanden. Die Standorte sind in [Abbildung 4](#) dargestellt.

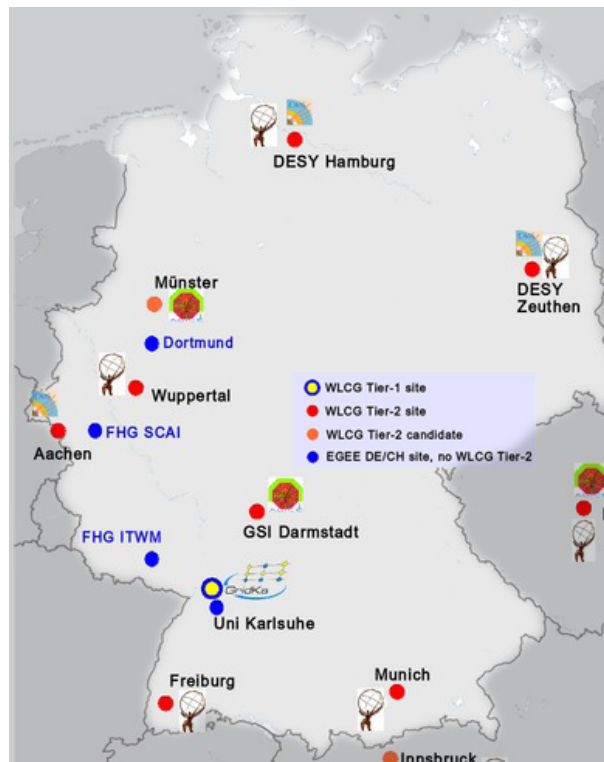


Abbildung 4: LCG-Standorte in Deutschland.

3.5.5.5 TIER 3, TIER 4

Unter dem Sammelbegriff der TIER-3- und TIER-4-Zentren fasst man all die Server und Rechner zusammen, die eine wissenschaftliche Einrichtung nicht gegenüber des Grids vertraglich zugesichert hat. Natürlich können diese Ressourcen trotzdem für Berechnungen und Analysen herangezogen werden.

TIER-3-/4-Zentren sind somit alle Zugangspunkte, die, wenn nötig, Berechnungen im Grid durchführen können.

3.5.6 Jobverteilung mit Fairshare

Der Schlüssel, mit dem entschieden wird, wann welches Experiment wieviel Rechenleistung im LCG erhält, nennt sich *Fairshare*. Die Währung, mit der das Fairshare-System kalkuliert ist die Rechenleistung in kSI2k, wie wir sie weiter oben schon gesehen haben.

Vor dem Start der Experimente werden Anteile, so genannte „Shares“ definiert. Diese Shares gibt es beispielsweise für Arbeitsgruppen und für einzelne Benutzer in Arbeitsgruppen. Abhängig von bereitgestellter CPU-Leistung und von verbrauchter CPU-Leistung wird nun bedarfsorientiert und dynamisch freie Rechenleistung verteilt. *Fair*, da es auch Möglichkeiten gibt, Rechenleistung auf Kredit zu erhalten, wenn das persönliche Kontingent bereits aufgebraucht ist, gerade aber von keiner anderen berechtigten Person Rechenleistung genutzt wird.

3.5.7 Weitere Kooperationsmöglichkeiten

Das Grid selbst ist zur Berechnung und zur Speicherung der Daten aus den LHC-Experimenten vorhanden. Es gibt aber eine Vielzahl Informationen, die über das Grid hinausgehend ausgetauscht werden müssen. An das LCG schließen sich somit direkt weitere Einrichtungen an.

Über Videokonferenzen wird direkt, von Person zu Person über neue Pläne gesprochen und aktuelle Information ausgetauscht. Es gibt Mailing-Listen und Wikis, in denen Selbiges stattfindet.

Und natürlich gibt es viele Webseiten, die sich unter anderem mit der Dokumentation des Grids beschäftigen, oder aber Zusatzinformationen bereit halten. Zur Kategorie Letzterer gehören die beiden Statusseiten GridView¹⁶, bei der man den Grid- bzw. Jobstatus kontrollieren kann, und der GridPP Real Time Monitor¹⁷, der über die Auslastung des Grids berichtet. *Abbildung 2* war ein Auszug aus dem GridPP Real Time Monitor.

4 Ausblick

4.1 WLCG

Das WLCG ist alles andere als fertiggestellt und in seinem Finalstadium. Es wird ständig ausgebaut und erweitert.

¹⁶<http://gridview.cern.ch/GRIDVIEW/>

¹⁷<http://gridportal.hep.ph.ic.ac.uk/rtm/>

So werden die Kapazitäten erhöht und dies in allen Ebenen. Das GridKa plant beispielsweise, seine verfügbare CPU-Leistung bis zum Jahr 2011 von 11.000 kSI2k auf 27.000 kSI2k zu erhöhen. Auch der Festplattenspeicherplatz wird von 4 PB auf 17 PB wachsen.

Ein weiterer Erweiterungspunkt sind die Anbindungen der TIER-Zentren untereinander und ans CERN. Hier soll die Geschwindigkeit erhöht werden.

4.2 Andere Ansätze

Es wird sich aber in der Zukunft auch in allgemeineren Ebenen viel tun.

Grid Computing

Grid Computing im Gesamten kann sich noch an vielen Stellen weiterentwickeln. Stablere und flexiblere Middleware, die durch mehr Funktionsreichtum aufwartet ist hier ein zentraler Ansatzpunkt

Virtualisierung

Bei *Virtualisierung* handelt es sich um den Einsatz von virtuellen Computern. Auf einem Großrechner wird dafür softwareseitig eine dedizierte Umgebung erstellt, die für den Anwender wie ein eigener Rechner funktioniert.

Man hat so die Möglichkeit, auf einem Rechner verschiedene Betriebssysteme virtuell laufen zu lassen. Außerdem kann man, wenn denn gewollt, die Rechenleistung eines Großrechners auf kleine, virtuelle Sub-Rechner verteilen und somit Leistungsspitzen relativ kostengünstig abfangen.

Cloud Computing

Ein relativ neuer Ansatz ist das *Cloud Computing*, bei dem der Anwender nicht mehr über eine eigene Rechenmaschine verfügt (weder real, noch virtuell). Er kauft vielmehr die Rechenleistung bei einem kommerziellen Unternehmen ein und braucht sich nicht mehr um die äußeren Gegebenheiten eines Großrechners zu kümmern.

5 Zusammenfassung

Das World Wide Web ist in seinen 15 Jahren, in denen es existiert, zu einer nicht mehr wegzudenken, zentralen Möglichkeit des Datenaustauschs geworden.

Grid Computing bezeichnet das Teilen von verschiedenen Ressourcen, meist Rechenleistung und Festplattenspeicherplatz. Ian Foster ist hier eine zentrale Persönlichkeit.

Das interessanteste Grid-Projekt ist momentan das WLCG, welches zu großen Teilen auf bereits existierenden Grids aufbaut – darunter das große europäische EGEE. Das WLCG ist aufgeteilt in verschiedene TIERS, die nach dem Zwiebelschalenprinzip unterverteilt werden. Das deutsche TIER-1-Zentrum steht in Karlsruhe; Aachen ist zusammen mit dem DESY Hamburg ein TIER-2-Zentrum.

Wichtige Begriffe im Zusammenhang mit Grids bzw. dem WLCG sind: Middleware (Software, die Computerhaufen zu Grids macht), Virtuelle Organisation (Benutzerverwaltung innerhalb eines Grids) und Fairshare (Schlüssel, nach dem im WLCG Rechenleistung verteilt wird).

6 Quellen

6.1 Literaturquellen

Informationen zu dieser Ausarbeitung wurden aus vielerlei Quellen zusammen getragen. Nachfolgend die Wichtigsten:

- Geschichte des World Wide Webs: Wikipedia-Artikel „World Wide Web“ (http://de.wikipedia.org/wiki/World_Wide_Web) und „How the web began“-Infoseite des CERNs unter <http://public.web.cern.ch/public/en/About/WebStory-en.html>.
- Informationen zu Google: Wikipedia-Artikel zu „Google“ (<http://de.wikipedia.org/wiki/Google>) und „Google Inc.“ (http://de.wikipedia.org/wiki/Google_Inc.).
- Daten zur Internetnutzung weltweit: <http://www.internetworldstats.com/>; Daten zur Internetnutzung Deutschlands: ARD-ZDF-Onlinestudie (<http://www.ard-zdf-onlinestudie.de/>).
- Informationen zu Grid-Computing: CERN GridCafé (<http://www.gridcafe.org>) (Häufig benutzt, allerdings noch in der alten Version.). Vieles auch aus den Flash-Animationen dort.
- Geschichte des Grid-Computings: (Englischer) Wikipedia-Artikel „Grid-Computing“ (http://en.wikipedia.org/wiki/Grid_computing#History), GridCafé.
- Generelle Informationen zum LHC Computing Grid
 - (Englischer) Wikipedia-Artikel „LHC Computing Grid“ (http://en.wikipedia.org/wiki/LHC_Computing_Grid).
 - WLCG-Infoseite des WLCGs selbst unter <http://lcg.web.cern.ch/LCG/>.
 - GridCafé Infoseite zum LCG.
- Fairshare-Informationen beim GridKa (<http://www.gridka.de/cgi-bin/frame.pl?seite=/pbs/fairshare.html>).
- Kosten des CERN-Rechenzentrums aus Paper „Sizing and Costing of the CERN TO center“ (CERN-LCG-PEB-2004-21).
- Einige Informationen aus den LCG-Artikeln (S. 128ff) der Ausgabe 21/2008 der Computerzeitschrift *c't*.

6.2 Bildquellen

- Abbildung 1: Bild aus der Wikipedia „GM - Countries by Internet Users.png“ (http://commons.wikimedia.org/w/index.php?title=Image:GM_-_Countries_by_Internet_Users.png&oldid=14911771). Veröffentlicht unter Creative-Commons-Attributions-ShareAlike-2.5-Lizenz von Wikipediabnutzer *Aaaarg* (<http://en.wikipedia.org/wiki/User:Aaaarg>).

- **Abbildung 2:** Auszug eine Screenshots vom GridPP Real Time Monitor <http://gridportal.hep.ph.ic.ac.uk/rtm/>.
- **Abbildung 3:** Selbsterstellt.
- **Abbildung 4:** Auszug aus einem Bild auf der LCG-Webseite <http://lcg.web.cern.ch/LCG/>.