

Web & Grid

Wie weltweite Forschungskooperation ermöglicht wird

Seminarvortrag von
Andreas Herten

andreas.herten@rwth-undso

Betreut von A. Nowack

Übersicht

- Web
 - Idee
 - Entwicklung
- Grid
 - Idee
 - Arten von Grids
 - LHC Computing Grid (LCG)

Web

- Tim Berners-Lee, Proposal 1989
- CERN
- Motivation:
 - Austausch von Forschungsergebnisse
 - International
 - Automatisch
 - Bilder/Grafiken
 - Gemeinsame Basis

Web

- Idee
 - Zugänglich von überall
 - *World Wide Web*
 - Offene Standards
 - Jeder kann Beitragen (Browser, Websites)
 - Wie sähe das Web sonst heute aus?
 - Hyperlinks (unidirektional)
 - **Bild und Text im Verbund**

Web

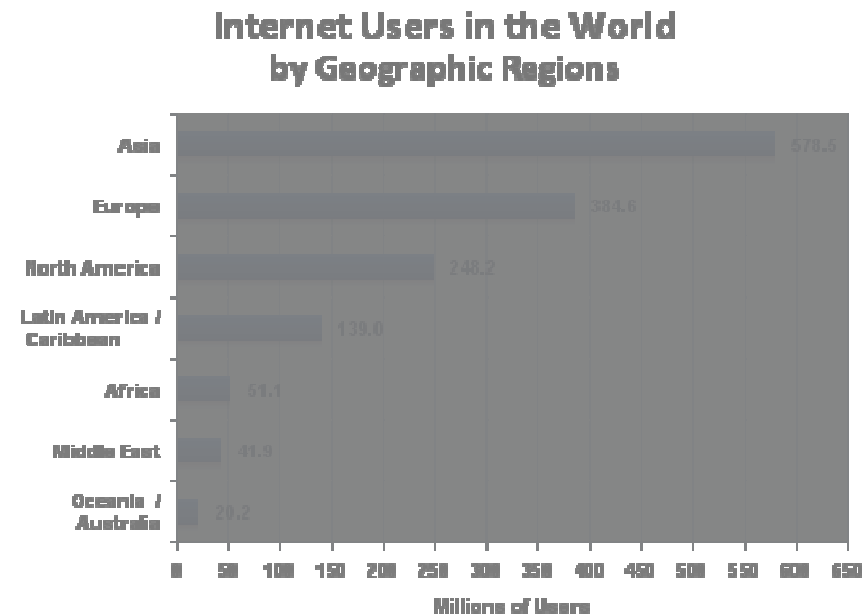
- Timeline
 - 1989: Anfang: Berners-Lee
 - 1990: Softwareprototyp
 - 1991: Teilchenphysik
 - 1. Webserver in US (SLAC, Stanford)
 - 1992: Browserentwicklung
 - Ursprünglicher Browser nur eingeschränkt nutzbar
 - 1993: Öffentlichkeit
 - 1994: 10.000 Server

Web

- Seitdem?
 - Leben ohne fast nicht mehr möglich
 - Wissenschaft, Privat, **Wirtschaft**
 - Beispiel: Google (*1998)
 - Umsatz 2007: 16 Mrd. Dollar
 - Marktwert 2007: 220 Mrd. Dollar (vor P&G und Toyota)
 - Markenwert 2008: 86 Mrd. Dollar (vor GE, MS, Coca-Cola)
 - Mitarbeiter 2008: 20.000

Web

- Seitdem?
 - Leben ohne fast nicht mehr möglich
 - **Wissenschaft, Privat, Wirtschaft**
 - Welt: 1,6 Billion Internetnutzer 2008
 - DE: 70% online



Source: Internet World Stats - www.internetworldstats.com/stats.htm
Estimated Internet users is 1,483,832,381 for Q2 2008
Copyright © 2008, Minihwatta Marketing Group

Web

- Ortsunabhängig
- Informationen

Grid

Grid

- Motivation
 - Große Datenmengen
 - LHC: 15 Petabyte Daten pro Jahr
 - 25 km hoher CD-Stapel
 - 1% der Informationen der Menschheit p. a.
 - Aufwändige Berechnungen

Grid

- Motivation
 - Große Datenmengen
 - Aufwändige Berechnungen
 - Vorteile gegenüber Supercomputern
 - Geldgeber: Lokale Systeme
 - Dynamische Nutzung über Projekt hinaus
 - Infrastruktur vor Ort
 - Skalierbarkeit
 - Ausfallsicherheit / Backups

Grid

- Idee
 - Ressourcenteilung
 - Rechenleistung
 - Speicherplatz
 - Sicherheit (/Abschottung)
 - Sicherer Zugang
 - Intelligente Jobverteilung
 - Job wird zu Daten geschickt
 - Offene Standards

Grid

Ian Foster:

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

Web

- Ortsunabhängig
- Informationen

Web | Grid

- Ortsunabhängig
- Informationen
- (Datenspeicher)

- Ortsunabhängig
- Rechenleistung
Datenspeicher
- → Informationen

Grid

- Geschichte I
 - Ursprüngliche Zentralcomputer / Großrechner
 - Grid-Vorgängerprojekte (1995)
 - FAFNER
 - I-WAY
 - 1998: **Ian Foster** & Carl Kesselman:
The Grid: Blueprint for a New Computing Infrastructure
 - Analogie zum Power Grid

Grid

- Geschichte II
 - 1998: „Bibel“ von Foster & Kesselman
 - 1999: *Globus Project* von Foster & Kesselman
 - 1999: SETI@home
 - Proof of concept
 - Verschiedene Systeme
 - 2008: Stanfords Folding@home mit 4 PFLOPS (top500 #1: 1 PFLOPS)

Grid

- **Verschiedene Arten von Grids**
 - Nicht das Grid
 - *Unterschied zum Web!*
 - Nationale Grids
 - Institutions-Grids
 - Projekt-Grids
 - Goodwill-Grids / Volunteer-Grids

Grid

- **Verschiedene Arten von Grids**
 - Private / öffentliche Grids
 - Jeder kann, wenn er will... (*Middleware*)
 - United Devices → World Community Grid
 - distributed.net
 - SETI@home
 - Folding@home
 - FightAIDS@home
 - *LHC*@home

Grid

- **Verschiedene Arten von Grids**
 - **Kommerzielle Grids**
 - **Wirtschaft & Forschung**
 - Z.B. D-Grid-Pilotprojekt mit BMW
 - » Crash-Tests
 - » Windkanal-Tests
 - » Produktionssysteme
 - **Herausforderungen**
 - Preisgestaltung / Abrechnung
 - Vertraulichkeit
 - » Wer rechnet potenziell?
 - » Wer rechnet gerade?
 - Sicherheit

Grid

- **Verschiedene Arten von Grids**
 - **Wissenschaftliche Grids / Projekt-Grids**
 - D-Grid (DE)
 - National Grid Service (UK)
 - EUFORIA project
 - EU Fusion fOR Iter Applications
 - Open Science Grid (US)
 - GridPP (UK), NorduGrid (Skandinavien)
 - EGEE / EGEE II (EU, Welt)
 - Enabling Grids for E-science
 - ICEAGE
 - WISDOM: Malaria, H5N1



Grid

- Verschiedene Arten von Grids
 - Nicht das Grid
 - Private / öffentliche Grids
 - Kommerzielle Grids
 - Wissenschaftliche Grids / Projekt-Grids
 - Und noch viele, viele mehr
 - http://en.wikipedia.org/wiki/List_of_distributed_computing_projects
 - <http://gridcafe.web.cern.ch/gridcafe/gridprojects/projects.html>

Grid

Das wichtigste Grid nicht erwähnt:

WLCG

Worldwide LHC Computing Grid



Grid » LCG

- Geschichte des LCG
 - 1998-2000: MONARC-Studie
 - 2001: European DataGrid (EDG)
 - Ab 2004: EGEE
 - 2002: Gründung LCG
 - 2005: LHC@home
 - 2008 (Oktober): Start WLCG (offiziell)

Grid » LCG

- Strukturelle Grundlagen des WLCGs
 - LCG kein eigenes, großes Grid, nutzt bestehende Strukturen
 - EDG → EGEE
 - Größtes, europäisches Grid
 - EU-finanziert (3*30 Millionen Euro)
 - Open Science Grid (OSG)
 - GridPP
 - NorduGrid
 - CERN openlab

Grid » LCG

- Virtuelle Organisationen
 - Einteilung der Benutzergruppen
 - Experiment = eigene virtuelle Organisation
 - Benutzerverwaltung
 - Authentifizierung
 - Autorisierung

Grid » LCG

- Middleware » Grundlagen
 - Softwareschnittstelle zwischen Ressourcen eines Grids
 - → Zugang zum Grid
 - Verbindet unterschiedliche Rechnerarten (Betriebssysteme, Speichersysteme, CPUs)
 - Macht Computerhaufen zu strukturiertem Grid

Grid » LCG

- Middleware » Auswahl
 - **UNICORE**: Für Supercomputer, aus Jülich, seit 1997
 - **Globus Toolkit**: Open-Source-Baukasten der Globus Alliance (Foster/Kesselman), Quasi-Standard zum Grid-Bau
 - **Virtual Data Toolkit**
 - **EDG → LCG → gLite**
 - **BOINC**: Aus Berkeley, Consumer-Client-orientiert

Grid » LCG

- Middleware » gLite



- Globus Toolkit

- Kann alles, bis auf automatisches Job-Routing:
 - Manuelles bestimmen, wo im Grid Job ausgeführt werden soll

- EDG: EDG-Middleware

- Mit Resourcebroker!

- → gLite (Open Source)

Grid » LCG

- Middleware » WLCG
 - Unterschiedliche Middlewares:
 - Europa: gLite
 - USA: Virtual Data Toolkit
 - → Gateways zwischen den Sub-Grids

Grid » LCG

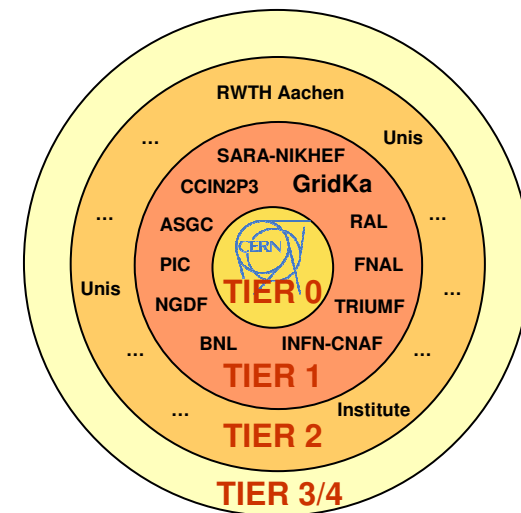
- Zahlen
 - 150.000.000 Sensoren in Experimenten
 - 100.000 CPUs in Rechenzentren
 - 20% davon im CERN
 - Speicherplatz: 60 PB HDD, 50 PB Band
 - Kosten Rechenzentrum CERN:
46 Millionen Euro (2008: 10 Mio. Euro)
 - Ziel 2012: Rechenleistung insg.
340.000 kSI2K
 - Pentium 4 @ 3 GHz: 1 kSI2k („kilo SPECint2000“)

Grid » LCG



Grid » LCG

- Funktionsweise des WLCGs
 - Verschiedene Stufen der Datenverteilung
 - TIERS
 - Jede Ebene gleich viel Rechenleistung
 - Über Highspeed-Kabel in die ganze Welt



Grid » LCG

- *TIER -1*: ‚intelligente‘ Detektoren, Counting Room

- ATLAS: Event-Selection, Ereignisrate von 1 GHz → ~200 Hz reduzieren

- Detektoren: 300 GB/s
- → Counting Room
 - Mehrere Filterungsmechanismen
 - Insgesamt: 300 GB/s auf 320 MB/s Rohdaten reduzieren

| Exp. | Datenrate |
|-------|-----------|
| ATLAS | 320 MB/s |
| CMS | 220 MB/s |
| ALICE | 100 MB/s |
| LHCb | 50 MB/s |

Grid » LCG

- TIER 0: CERN Rechenzentrum
 - Über dedizierte 10Gbit/s-Leitungen an Counting-Rooms angeschlossen
 - Aufgaben
 - Speicherung Rohdaten auf Festplatten und Bändern
 - Bänderkapazität 17 PB
 - Festplattenkapazität 5,5 PB
 - ca. 27 TB/Tag
 - Event Summary Data erstellen
 - Erstes Rekonstruieren und Kalibrieren
 - in der CERN Analysis Facility
 - zusätzliche 140 MB/s (ca. 10 TB/Tag)
 - Speicherung auf Bändern
 - 4700 kSI2k
 - Weiterverteilung an TIER 1

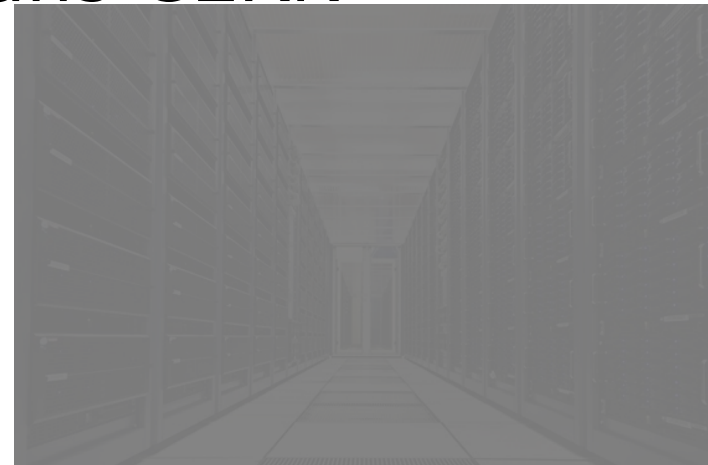
Grid » LCG

- TIER 1: 12 Zentren
 - In Deutschland, Frankreich, Italien, Kanada, Niederlande, Skandinavien, Spanien, Taiwan, UK, USA (2*) und am CERN
 - Dedizierte 10-GBit/s-Leitungen zu TIER 0
 - Untereinander verbunden
 - Pro TIER 1 eine eigene Auswahl an Daten, insgesamt kompletter Satz
 - Aufgaben
 - Backup der Daten von TIER 0 & TIER 2 (Band & HDD)
 - Rekalibrierung
 - Reprozessierung (ESD2)
 - Insg. 38.000 kSI2k, 20 PB HDD / 21 PB Band
 - Unterverteilung an TIER 2

Grid » LCG

- TIER 1: GridKa

- Deutscher LCG-Knotenpunkt im Forschungszentrum Karlsruhe
- 20-GBit/s-Anbindung ans CERN
- 2008:
 - 11.000 kSI2k
 - 4 PB HDD / 5 PB Band
 - 6500 CPU-Kerne
- Ausbau
- Unterstützt alle vier Experimente und vier nicht-LHC-Experimente

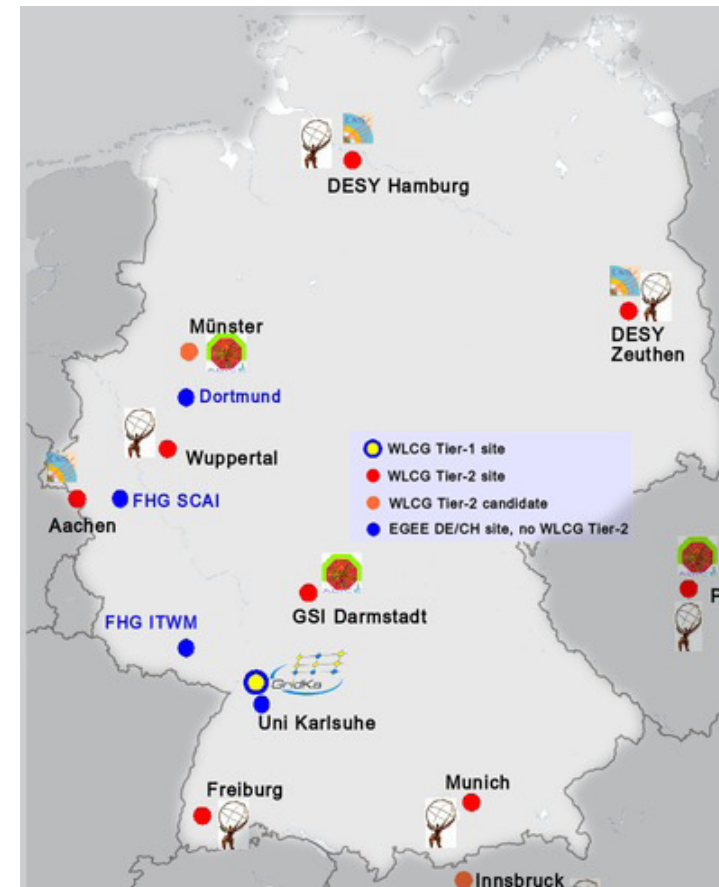


Grid » LCG

- TIER 2: Universitäten
 - 140 Standpunkte in 38 Ländern
 - Untereinander in Wissenschaftsnetzen verbunden
 - Zugriff auf Daten der TIER 1
 - Keine Archivierung bei TIER 2!
 - Ergebnisse von Berechnungen nach TIER 1
 - Aufgaben
 - Zugang für ~7000 Forscher
 - Analysen und eigentliche Berechnungen
 - MC-Simulationen → TIER 1
 - Insg. 46.000 kSI2k, 13 PB HDD

Grid » LCG

- TIER 2: Standorte in Deutschland
 - ATLAS
 - DESY Hamburg
 - Freiburg & Wuppertal
 - München
 - ALICE
 - GSI Darmstadt
 - CMS
 - DESY Hamburg & RWTH Aachen
 - 600 & 450 kSI2k CPU
 - 170 & 100 TB HDD
 - Insg. 3500 kSI2k



Grid » LCG

- TIER 3/4
 - Server, gegenüber Grid nicht zugesichert
 - Client-Rechner
 - Zugangspunkte allg.

Grid » LCG

- Verteilung von Jobs
 - Erhalt von Rechenleistung in „kSI2k“
 - Fairshare
 - Vorher definierte Anteile („Shares“) für
 - Arbeitsgruppen
 - User in Arbeitsgruppen
 - Dynamische Rechenleistungsverteilung (bedarfsorientiert)
 - Abhängig von bereitgestellter CPU-Leistung und verbrauchter CPU-Leistung

Grid » LCG

- Weitere Kooperationsmöglichkeiten
 - Über Grid hinausgehende Informationen
 - Videokonferenzen
 - Mailing Listen
 - Wikis
 - Webseiten
 - Status
 - Grid- / Jobstatus via GridView
<http://gridview.cern.ch/>
 - GridPP Real Time Monitor
 - Dokumentation

Ausblick

Ausblick

- LCG
 - Erhöhung der Kapazitäten
 - Viele Investitionen
 - In allen Ebenen
 - z.B. TIER 1 GridKa (2008 vs. 2011)
 - CPU: 11.000 kSI2k vs. 27.000 kSI2k
 - HDD: 4 PB vs. 14 PB
 - Erhöhung der Anbindung

Ausblick

- Grid Computing
 - Stabilere, flexiblere und funktionsreichere Middleware
- Virtualisierung
 - Dedizierte Umgebungen auf einem Großrechner
 - Betriebssysteme
- Cloud Computing
 - Keine Rechenmaschine lokal vorhanden, kompletter Einkauf der Rechenleistung von externem Anbieter
 - z.B. Amazon S3



- **Grid:**

- Ian Foster
- Teilung von Ressourcen (HDD, CPU)

- **WLCG:**

- Zusammenschluss Grids
- TIERs
 - GridKa

- **Begriffe:**

- Middleware
- Virtuelle Organisation

Danke!
...und Alaaf!