

Grid-Anwendungen in Deutschland

Klaus-Peter Mickel
Forschungszentrum Karlsruhe / DGI
mickel@kit.edu



Gliederung

- Grid-Computing "in a nutshell"
- Der Weg zu D-Grid
- Acht D-Grid-1-Projekte
- Elf D-Grid-2-Projekte
- Zusammenfassung und Ausblick

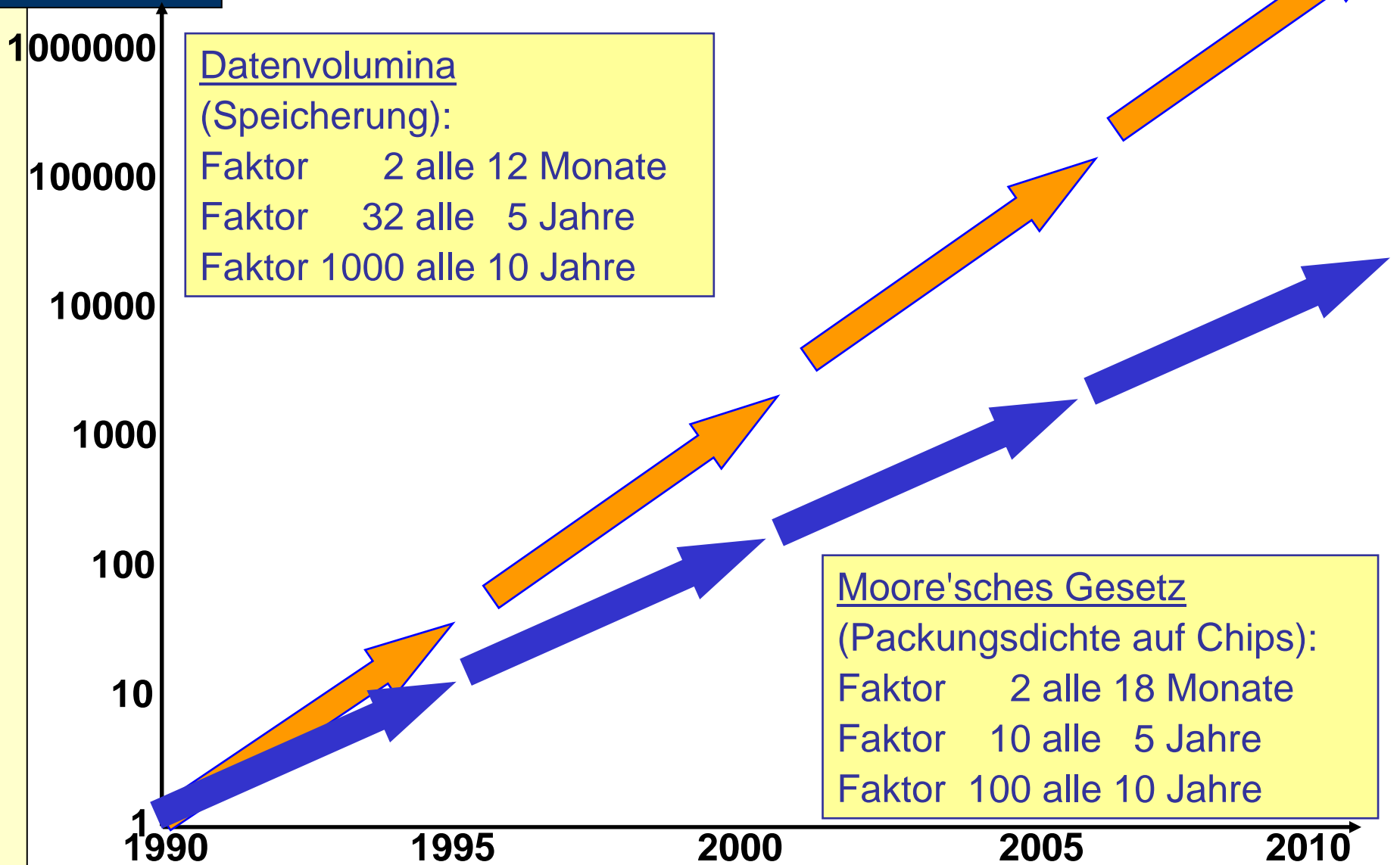


Der Ursprung der Grid-Idee:

- **Die wachsende Informationsmenge ...
... und der immer schwierigere Zugriff darauf.**
 - Die in Wissenschaft und Technik sowie im Internet verfügbare Informationsmenge verdoppelt sich jährlich
 - Das intelligente Nutzen (=Bearbeiten) dieser enormen Datenmengen mit herkömmlichen Methoden wird immer aufwendiger und immer kostspieliger
 - Das bisherige Modell großer zentraler Rechenzentren ist dafür i. a. nicht mehr wirtschaftlich



Das Wachstum der Daten und das Moore'sche Gesetz

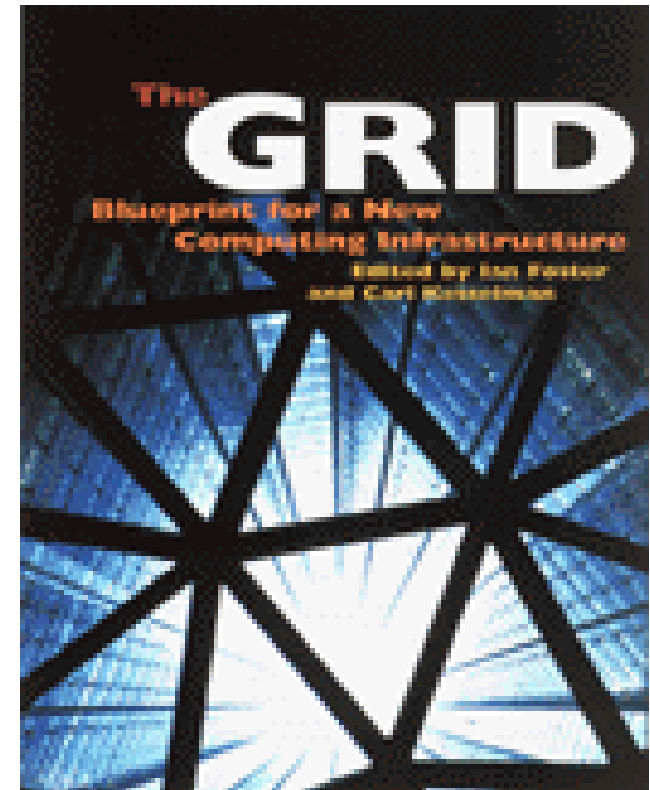


So entstand die Vision des Grid-Computing:

Der nahtlose Zugriff für "jedermann" auf beliebige IT-Ressourcen in beliebiger Menge an beliebigem Ort

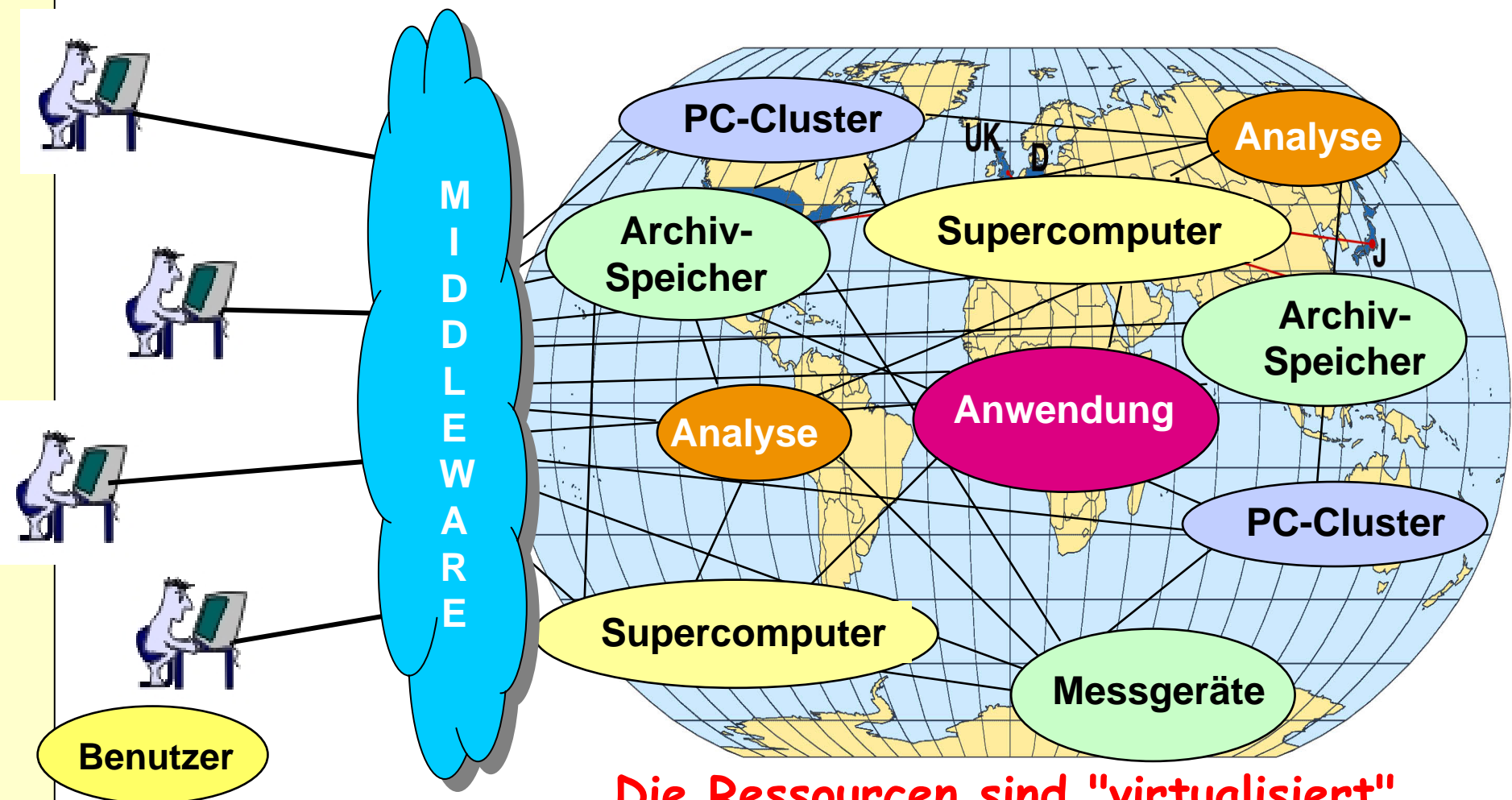
Kernpunkte eines Grid:

- * Verteilte Hardware
- * Verteilte Daten
- * Verbindende Netze
- * Grid-Middleware
- * Verteilte Anwender
- * Gemeinsame Projekte
- * Gemeinsame Lösungen



Ian Foster,
Carl Kesselman (1998)

Charakteristika eines Grid aus vernetzten IT-Ressourcen



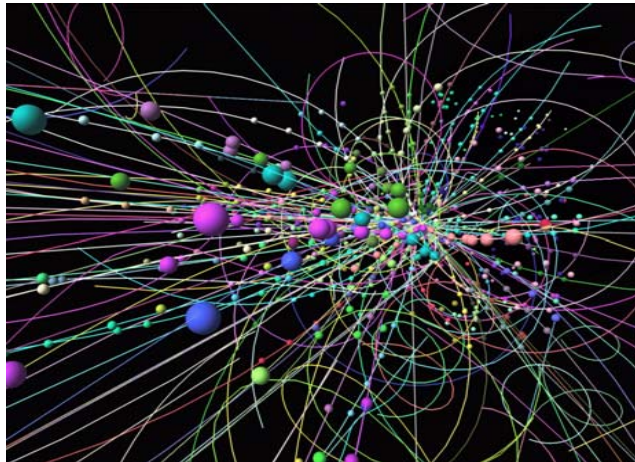
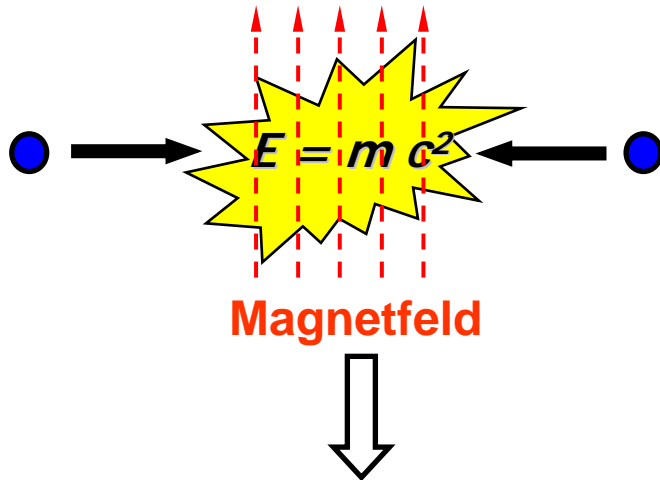
Die Ressourcen sind "virtualisiert",
d.h. nicht mehr einzeln sichtbar



Middleware: Der Kern des Grid-Computing

- **Anforderungen an die Middleware:**
 - Verwaltung Virtueller Organisationen (Individuen, Ressourcen)
 - Verwaltung von Zugangs- und Nutzungsrechten
 - Verwaltung verfügbarer Ressourcen:
 - "Wo im (globalen) Grid sind wann welche Ressourcen verfügbar?"
 - Wo im (globalen) Grid befinden sich die von einem Nutzer benötigten Daten?
 - Gewährleistung der Sicherheit

Teilchenkollisionen – Lupe ins Universum



Fundamentale Fragen der Physik

- Welches Teilchen ist verantwortlich für die Masse der Materie? (Higgs?)
- Warum bestehen wir aus Materie und nicht aus Antimaterie?
- 25% des Universums bestehen aus Dunkler Materie und 70% aus Dunkler Energie. Was verbirgt sich dahinter?
- Was geschah in der ersten Nanosekunde nach dem Urknall?
- Gibt es eine Universalwechselwirkung, die die vier bekannten fundamentalen Kräfte vereint?

Das CERN-Projekt "LHC" (Large Hadron Collider)

Beschleunigerring:

- * 27 km Umfang
- * Beschleunigung und Kollision von Protonen mit 7 TeV/Teilchen
- * Beginn der Experimente: 2008

Datenvolumen:

- * 4 Experimente
 - * ca. 10^6 Bytes / Ereignis / Experiment
 - * ca. $3 \cdot 10^2$ Ereignisse / Sekunde
 - * ca. 10^5 Sekunden / Tag
 - * ca. $2 \cdot 10^2$ Messtage / Jahr
- $24 \cdot 10^{15}$ Bytes / Jahr = **24 Petabytes / Jahr**

Benutzer: Etwa 8.000 Kernphysiker weltweit.



Hierarchie der Grid-Zentren:

1 * Tier-0 / 10 * Tier-1 / 130 * Tier-2

10 "Tier-1-Zentren", in denen jeweils 30 % aller Daten gespeichert werden:

An jedes Tier-1-Zentrum sind je 10 bis 15 Tier-2-Zentren angeschlossen, die jeweils einen kleineren Teil der Daten übernehmen.

So entstand ein weltweites Netz mit ~ 150 Computerzentren

Deutschland:	GridKa, FZK, Karlsruhe
Frankreich:	IN2P3, Lyon
Großbritannien:	Rutherford Laboratory, Chilton
Italien:	CNAF, Bologna
Niederlande:	NIKHEF, Amsterdam
Schweiz:	CERN, Genf
Spanien:	PIC, Barcelona
Taiwan:	Academia Sinica, Taipei
USA:	Fermi National Lab, Batavia, IL
USA:	Brookhaven National Lab, NY

Tier-2-Zentren an GridKa:

- DESY, Hamburg
- GSI, Darmstadt
- MPI für Physik, München
- Universität Aachen
- Universität Freiburg
- Universität Karlsruhe
- Universität Wuppertal
- Universität Innsbruck (Österreich)
- Universität Krakau (Polen)
- Universität Prag (Tschechien)
- Nationales wiss. RZ Manno (Schweiz)

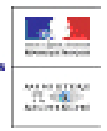


- Grid-Computing "in a nutshell"
- **Der Weg zu D-Grid**
- Acht D-Grid-1-Projekte
- Elf D-Grid-2-Projekte
- Zusammenfassung und Ausblick

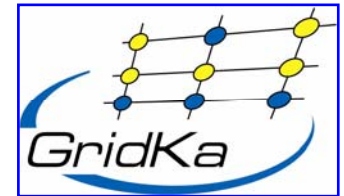


Nationale Grid-Initiativen in Europa

- Zunächst erhebliche Förderung einiger nationaler Grid-Initiativen (USA, Japan)
- Dann wachte auch Europa auf:
 - Austrian Grid Initiative
 - DutchGrid
 - France: e-Toile; ACI Grid
 - Germany: D-Grid; Unicore
 - Greece: HellasGrid
 - Grid Ireland
 - Italy: INFNGrid; GRID.IT
 - NorduGrid
 - UK e-Science: National Grid Service; OMII; GridPP



- Einige Grid-Aktivitäten in Deutschland:
 - 2001: Im Forschungszentrum Karlsruhe beginnt der Aufbau von **GridKa** als Tier-1-Zentrum für das globale Grid für die Hochenergiephysik
www.gridka.de
 - 2002: Das BMBF veröffentlicht "IT-2006", ein Programm zur Förderung der Informationstechnologie in Deutschland
www.it2006.de
 - 2003: Die Idee einer nationalen Grid-Initiative in Deutschland entsteht; binnen kurzem schließen sich Dutzende von Einrichtungen und Hunderte von Wissenschaftlern an
www.d-grid.de
 - 2004: Das BMBF veröffentlicht die erste Grid-Ausschreibung ("D-Grid")





D-Grid – Die erste Ausschreibung ('04)

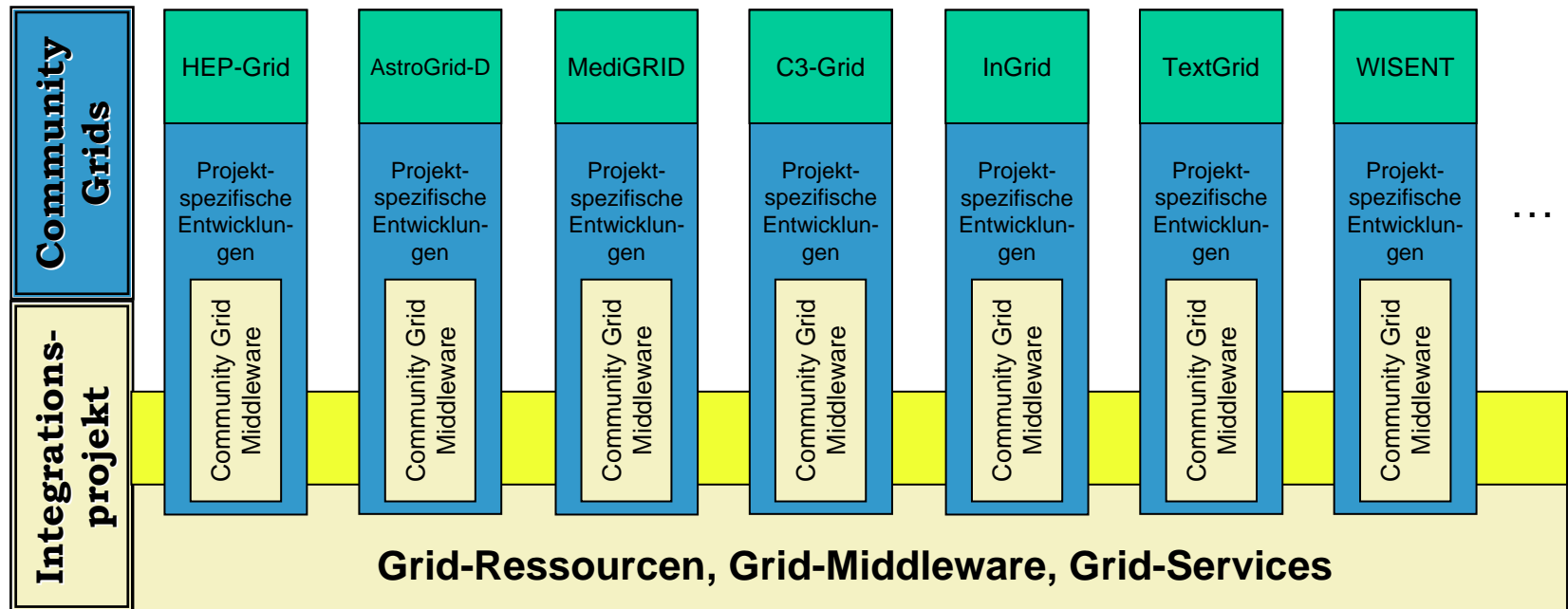


Bundesministerium
für Bildung
und Forschung

- Etablierung einer flexiblen, nationalen Grid-Struktur als technische Grundlage für die Realisierung wissenschaftlicher Arbeitsumgebungen
- Grundlegende Ziele sind dabei:
 - Entwicklung von e-Science-Diensten für die Wissenschaft
 - Offenheit des Ansatzes
 - Perspektive einer späteren professionellen Verwertung
 - Robustheit und Flexibilität der technischen Infrastruktur
- Die Communities sind die eigentlichen Impulsgeber der e-Science-Entwicklung. Kriterien zur Messbarkeit des Erfolgs des Forschungsprogramms sind daher:
 - Steigerung der wissenschaftlichen Leistungsfähigkeit dieser Communities
 - Bildung und Integration weiterer Communities

Ergebnis der ersten D-Grid-Ausschreibung ('05)

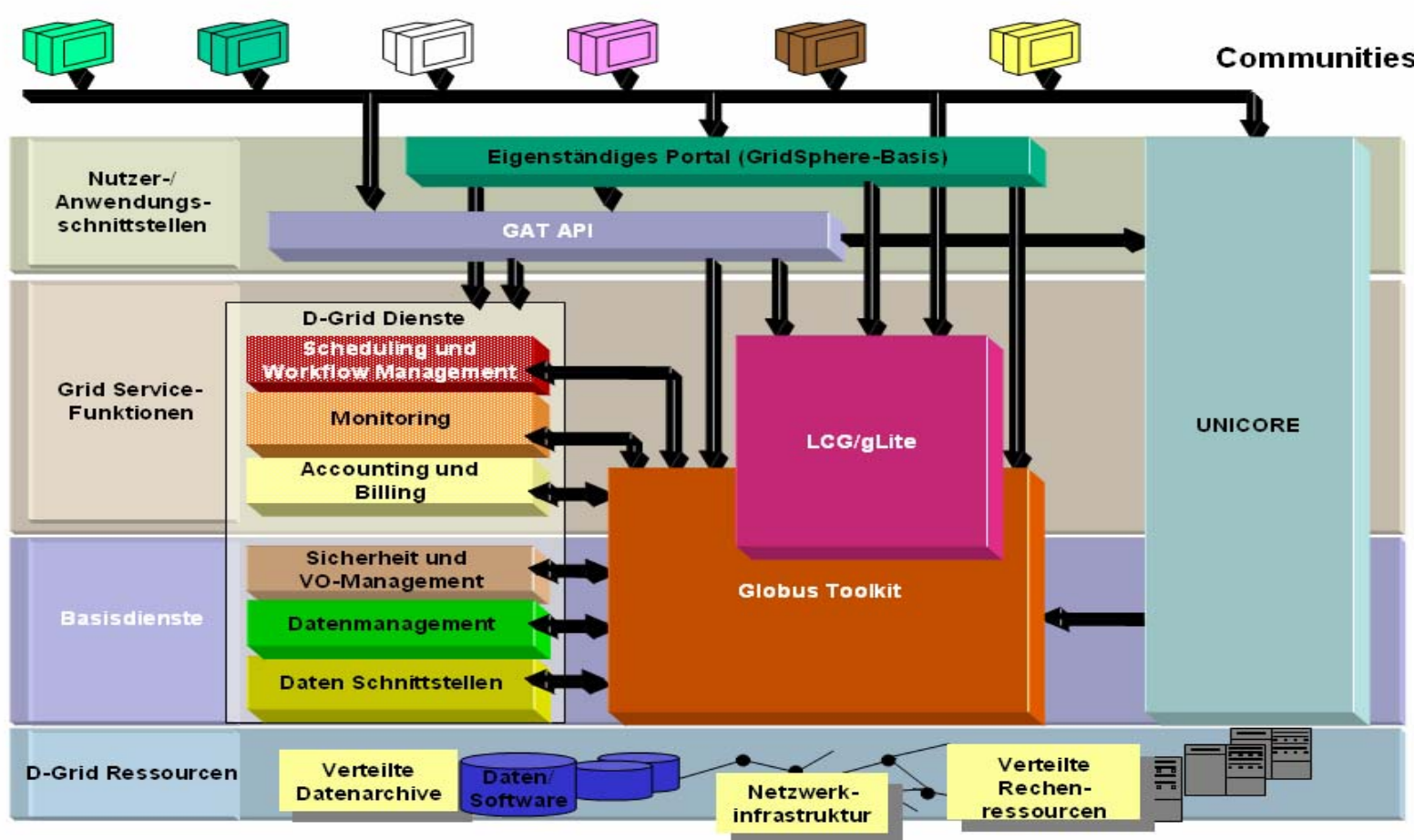
- Ein D-Grid Integrationsprojekt (DGI) – 4 Teilprojekte
 - Etablierung der allgemein nutzbaren Plattform ("Kern-D-Grid")
- Sieben Community Projekte:
 - Hochenergiephysik, Astrophysik, Medizin, Klimaforschung, Ingenieurwissenschaften, Geisteswissenschaften, Energiemeteorologie
- Ca. 100 Partner, Förderung: 01.09.05 bis 31.08.08





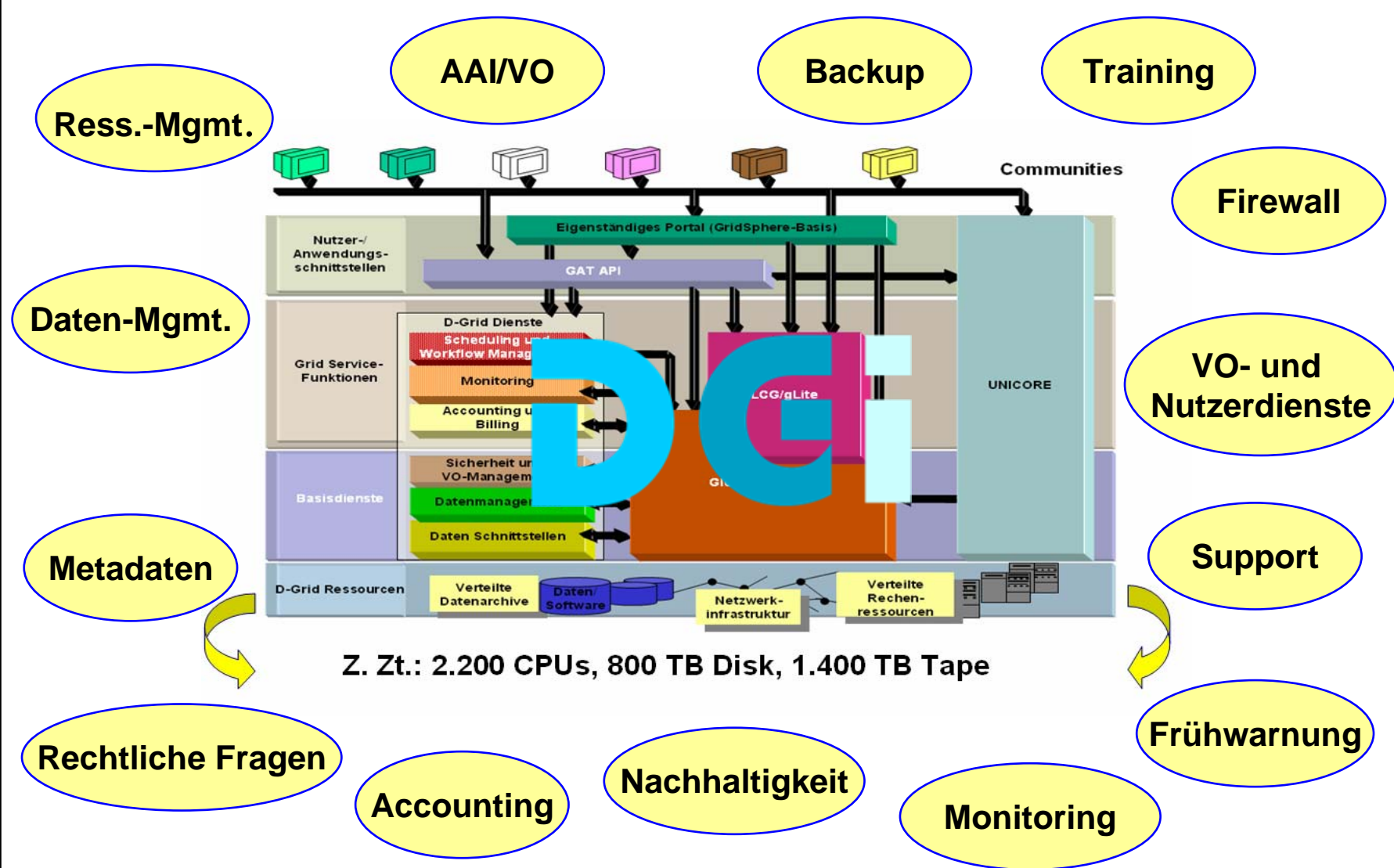
- Grid-Computing "in a nutshell"
- Der Weg zu D-Grid
- **Acht D-Grid-1-Projekte**
- Elf D-Grid-2-Projekte
- Zusammenfassung und Ausblick

Aufgabe des DGI: Etablierung der allg. Grid-Plattform (1)

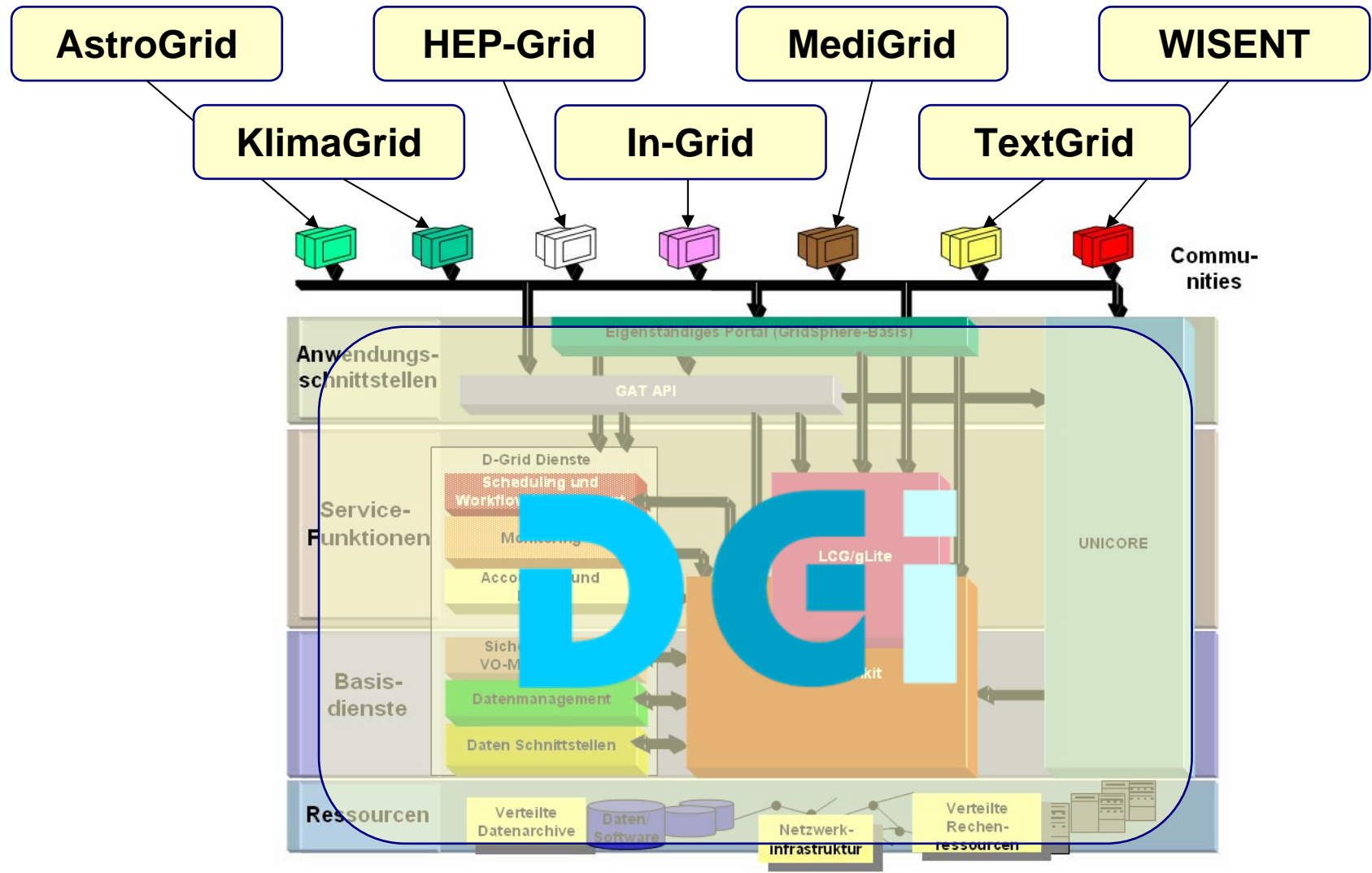


**Z. Zt.: 2.200 CPUs, 800 TB Disk, 1.400 TB Tape
(25 Standorte, vernetzt via DFN)**

Aufgabe des DGI: Etablierung der allg. Grid-Plattform (2)

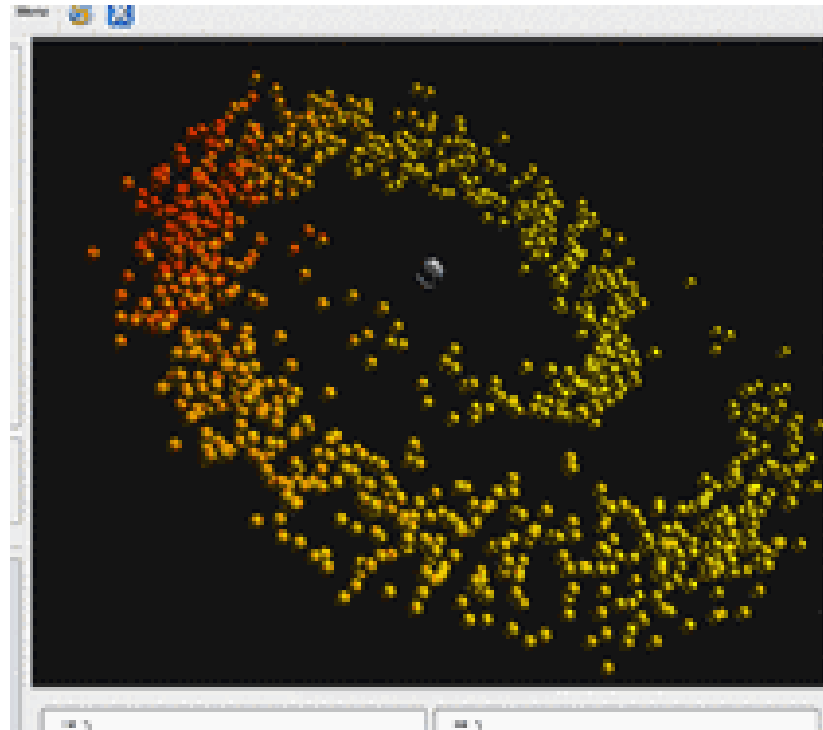


DGI und Community-Grids (CGs)

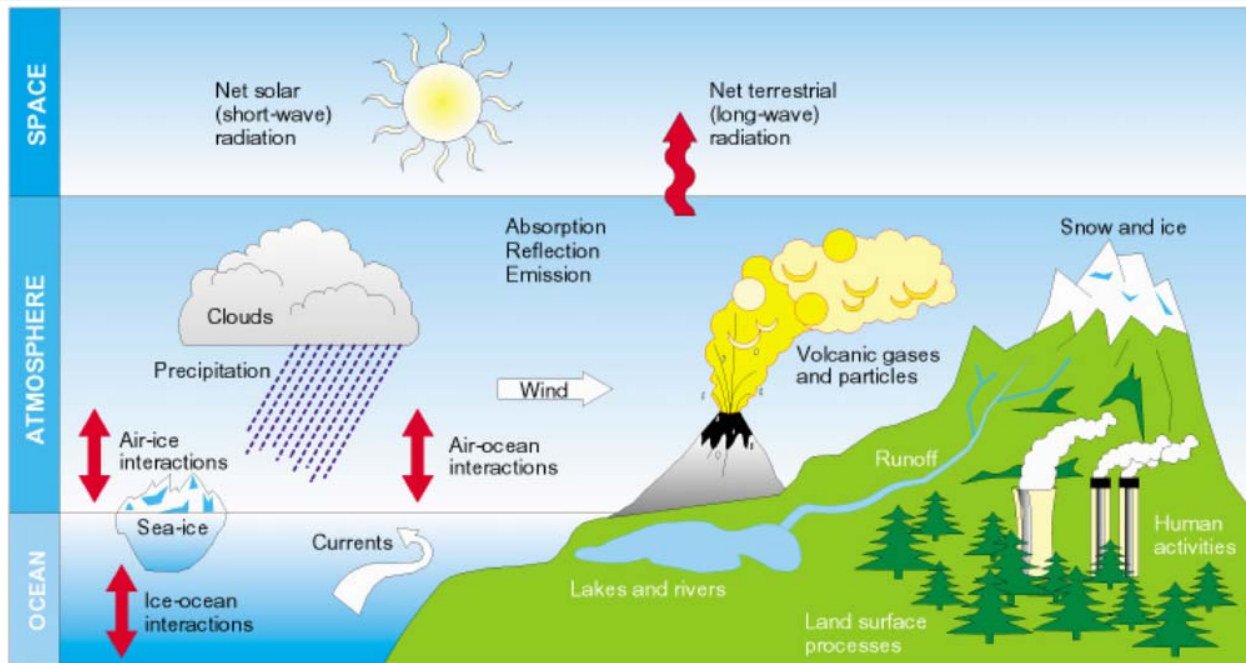




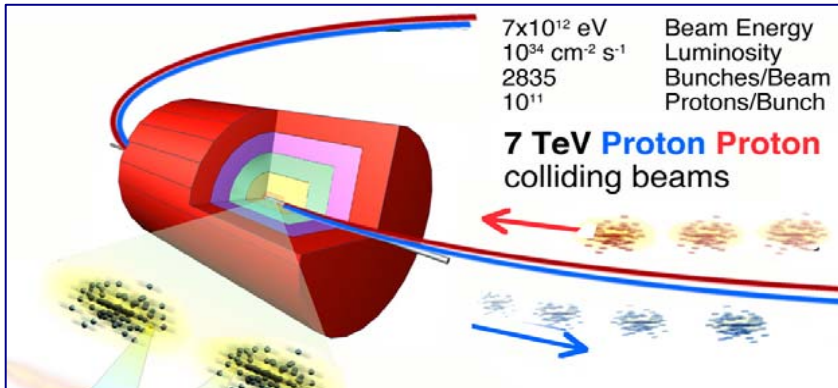
- Schaffung einer Grid-basierten Infrastruktur für Daten- und Computere Ressourcen
- Einbindung verteilter astronomischer Datenarchive
- Erhöhung der Reaktionsgeschwindigkeit bei interessanten astronomischen Events
- Einbindung kleinerer Institutionen an die Grid-Ressourcen
- Ziel u. a.: Experimenteller Nachweis von Gravitationswellen



C3-Grid – Collaborative Climate Community



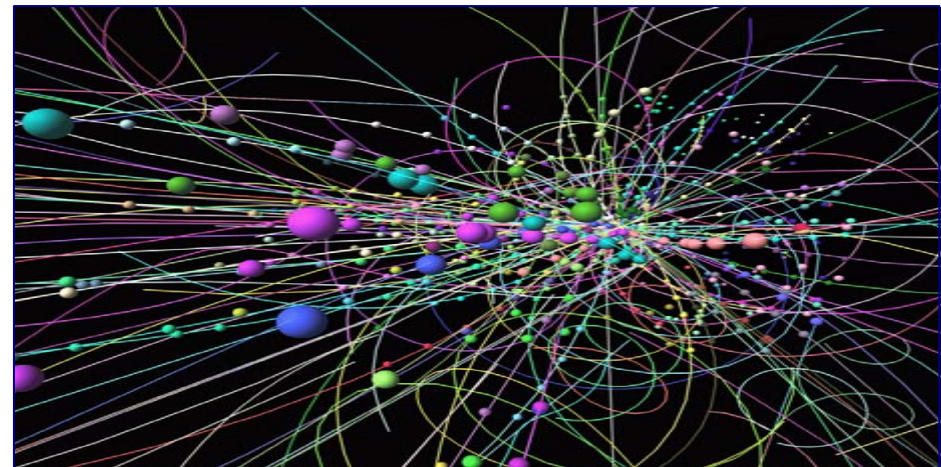
- Bereitstellung von Speicher- und Prozessierungsressourcen für die Wissenschaftler
- Schaffung einer Grid-basierten Infrastruktur für Daten- und Computere Ressourcen
- Schaffung eines einheitlichen Zugriffs auf alle verteilten Datenquellen
- Unterstützung der wissenschaftlichen Anwender bei der Analyse der Daten durch standardisierte Workflows



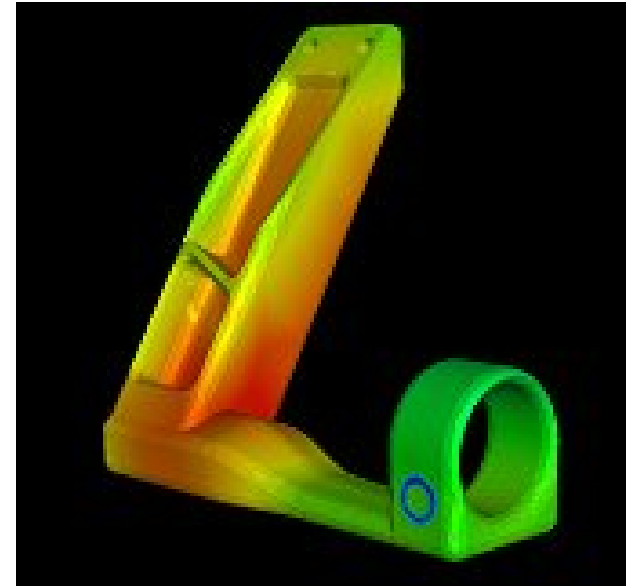
- HEP = Hochenergie- und Elementarteilchenphysik
- Auswertung riesiger Datenmengen aus Elementarteilchenbeschleunigern (CERN, Fermilab, Stanford Univ.)
- Zahlreiche Universitäten und Forschungseinrichtungen als Projektpartner beteiligt

Ziele:

- Schaffung einer Grid-basierten Infrastruktur für Daten- und Compute-Ressourcen
- Nachweis von dunkler Materie und dunkler Energie im Weltall
- Nachweis des "Higgs-Teilchens" (wird als Ursache der Masse von Materie vermutet)

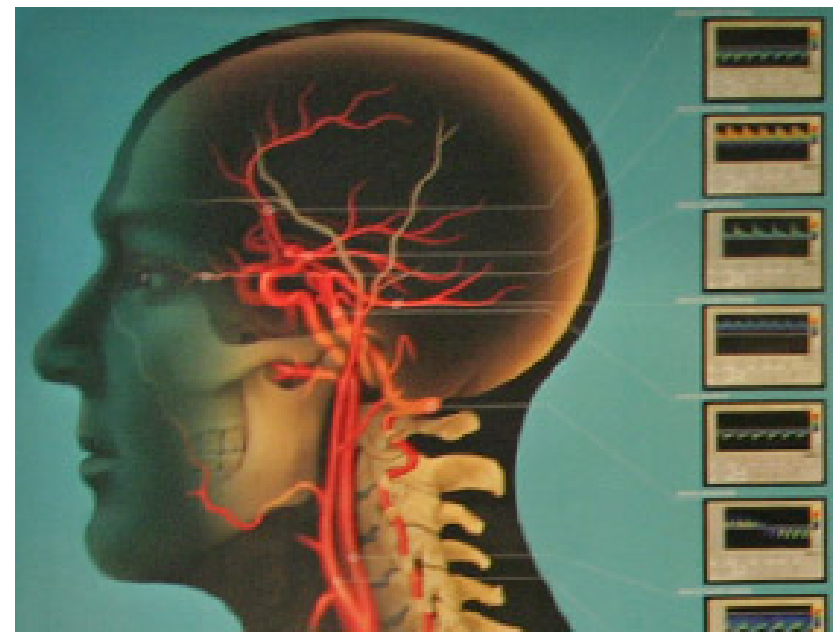
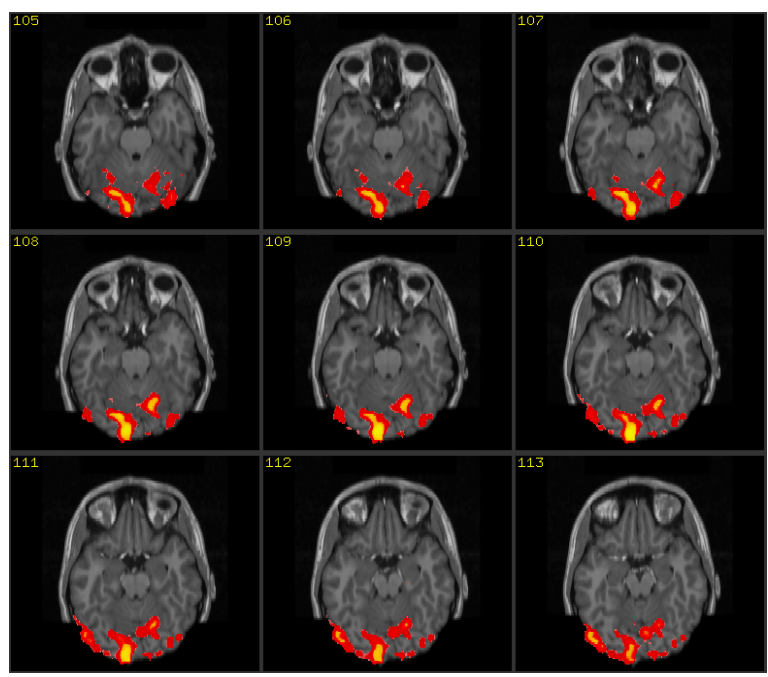


- Methoden und Modelle zur Lösung ingenieurspezifischer Probleme im Grid:
 - Wissensbasierte Unterstützung von ingenieurspezifischen Entscheidungsprozessen
 - Unterstützung von ingenieur-spezifischen Workflows
 - Verteilte simulations-basierte Produkt- und Prozessoptimierung



- Virtuelle Prototypen und Modellierung in der Industrie mit Grid-Methoden
 - Gießprozesse
 - Umformprozesse
 - Strömungsprozesse
 - Grundwassertransportprozesse
 - Magneto-hydrodynamische Kopplung von Fluidstrukturen

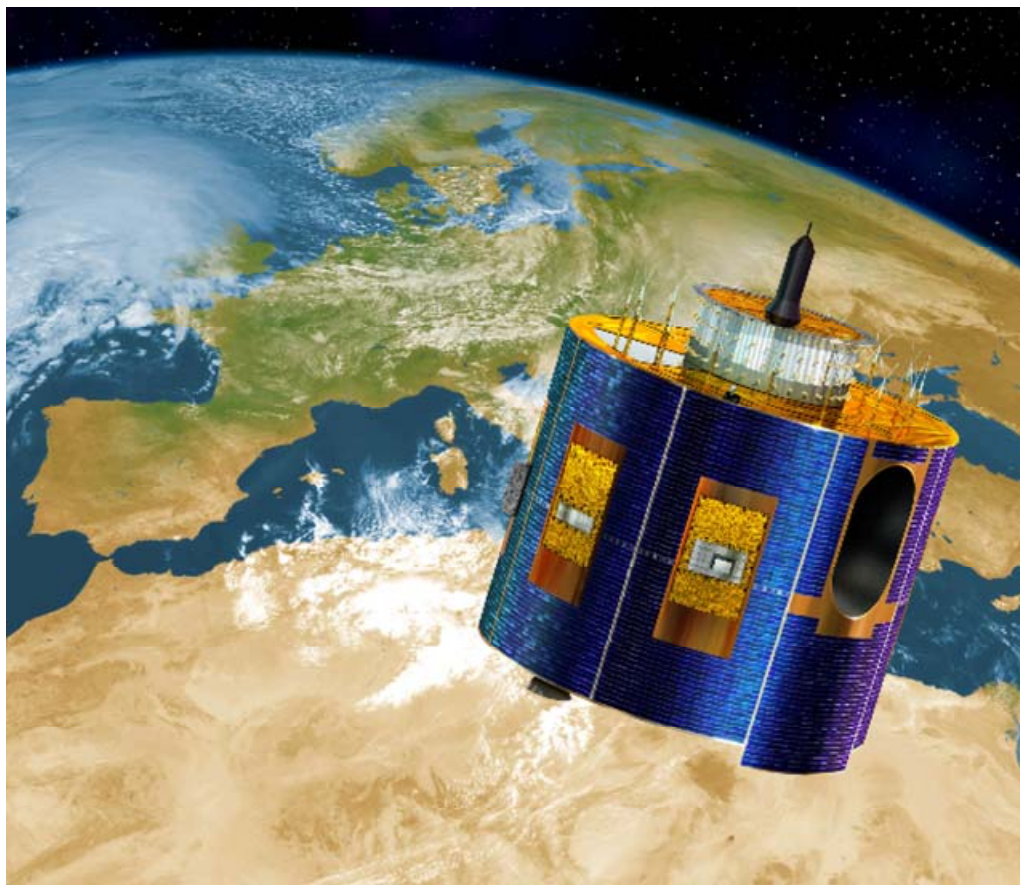
- MediGrid zeigt am Beispiel biomedizinischer Forschung mit hochdimensionalen Daten und der Verknüpfung von vielfältigen genotypischen und phänotypischen Daten die Anwendbarkeit von Grid-Diensten in der Medizin



- Zentrales Ziel des Verbundvorhabens MediGRID ist eine GRID-unterstützte Verbesserung der interdisziplinären, weitgehend standortunabhängigen Zusammenarbeit in der biomedizinischen Forschung.

- Modulare Grid-Plattform für verteilte und kooperative wissenschaftliche Textdatenverarbeitung.
- Ein Community-Grid für die Geisteswissenschaften.
- Ziele: Die gemeinschaftliche philologische Bearbeitung, Analyse, Annotation, Edition und Publikation von Textdaten für die Philologie, Linguistik und angrenzende Wissenschaften.
- Schaffung einer interdisziplinären, internationalen und vernetzten virtuellen Forschungsplattform
- Aufzeigen und Nutzen von Synergien in der wissensch. Textdatenverarbeitung.
- Rationalisierung des wissenschaftlichen Arbeitens unter anderem durch optimierten Zugriff auf Primärquellen und Werkzeuge.





Wettersatellit MSG-1: Die Forscher der Universität Oldenburg und des DLR können damit alle 15 Minuten flächendeckend und präzise den verfügbaren "Brennstoff" für Solarenergieanlagen bestimmen.

Quelle: EUMETSAT

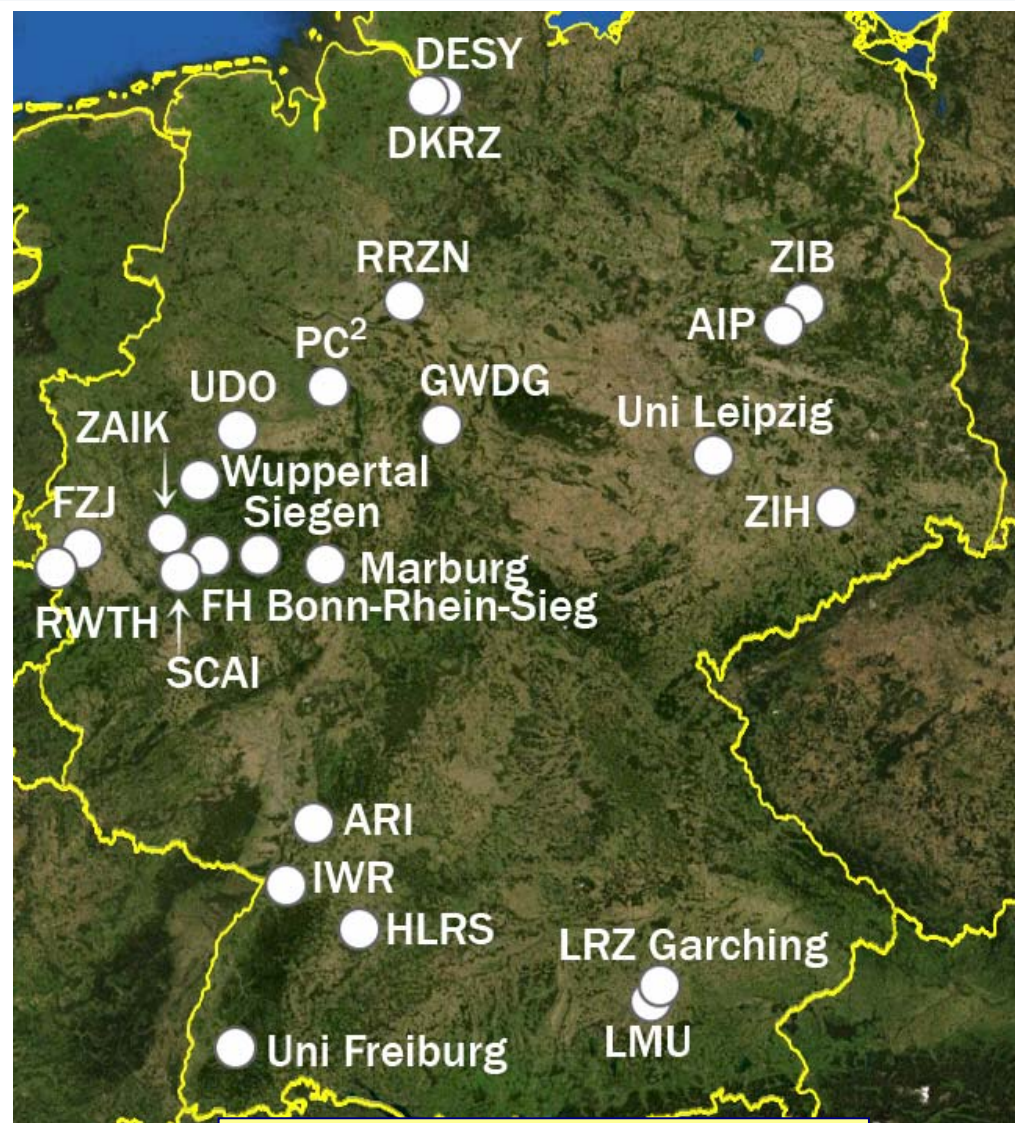
- Gewinnung von Informationen zur Charakterisierung der fluktuierenden Energieerzeugung aus Solar- und Windenergie
- Erkenntnisgewinn durch interdisziplinäre Kooperation
- Klima und Wetter (Meteorologie)
- Umwandlung von Wind und Sonnenenergie (Physik)
- Leistungsfähige, flexible verteilte Systeme (Informatik)



Die DGI-Infrastruktur (09/2007)

Zuse-Institut Berlin
Universität Dortmund
ZIH Dresden
Universität Freiburg
GWDG Göttingen
DESY Hamburg
DKRZ Hamburg
RRZN Hannover
ARI Heidelberg
FZJ Jülich
FZK Karlsruhe
Universität Köln
Universität Leipzig
Universität Marburg
LRZ München
DLR Oberpfaffenhofen
OFFIS Oldenburg
AIP Potsdam
FhG SCAI St. Augustin
Universität Siegen
HLRS Stuttgart
Universität Wuppertal

2.200 CPU-Cores, 0,8 PB Disk, 1,4 PB Tape



<http://www.d-grid.de>



D-Grid, Menschen und Mittel (1)

- BMBF hat D-Grid bis 2006 mit ca. 25 Mio Euro gefördert
- D-Grid-1 umfasst acht Projekte (DGI, sieben CGs)
 - 09/2005 bis 08/2008
- BMBF hat die D-Grid-Infrastruktur erheblich gefördert
- An D-Grid-1 sind ca. 100 Einrichtungen mit ca. 250 Personen beteiligt
- Neben D-Grid-1 gibt es noch viele weitere vom BMBF geförderte "e-Science-Projekte"
 - Bsp.: Instant-Grid

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



D-Grid, Menschen und Mittel (2)

- 2006: Ausschreibung D-Grid-2-Projekte
- Schwerpunkt: Service-Grids für Wissenschaft und Wirtschaft
- D-Grid-2 umfasst elf weitere Projekte
 - Q2/2007 bis Q1/2010
- BMBF fördert D-Grid-2 mit ca. 35 Mio Euro
- Ende 2007 wird das BMBF die D-Grid-Infrastruktur abermals erheblich fördern
- An D-Grid-1 + D-Grid-2 sind ca. 150 Einrichtungen mit ca. 450 Personen beteiligt
 - 45 Industrieunternehmen

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



- Grid-Computing "in a nutshell"
- Der Weg zu D-Grid
- Acht D-Grid-1-Projekte
- **Elf D-Grid-2-Projekte**
- Zusammenfassung und Ausblick



D-Grid-2-Projekte

- **AeroGrid:** Luft- und Raumfahrt-Forschung
- **BauVOGrid:** Bauindustrie
- **BIS-Grid:** Betriebliche Informations-Systeme
- **Biz2Grid:** Betriebliche Grids ("SAP u Grid")
- **D-MON:** Monitoring von Ressourcen und Diensten
- **F&L-Grid:** Grid-basierte IT-Dienste für Forschung und Lehre
- **FinGrid:** Finanz-Sektor (Börsendaten)
- **GDI-Grid:** Geo-Daten-Infrastruktur
- **PartnerGrid:** Kooperation in industriellen Organisationen (KMUs)
- **ProGrid:** Automobilindustrie
- **SUGI:** Grid-Support für kleine Institutionen und KMUs

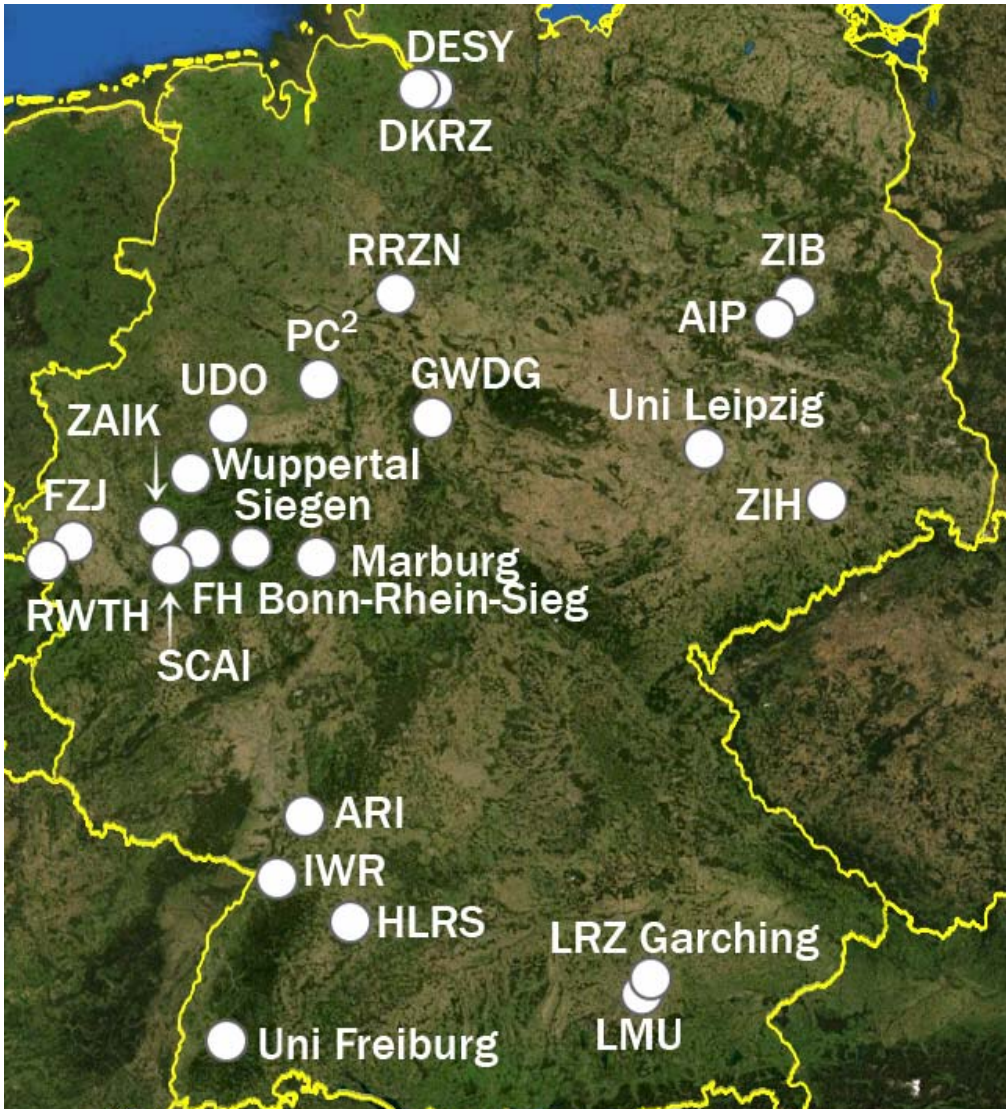


Die DGI-Infrastruktur (Q1/2008)

Zuse-Institut Berlin
Universität Dortmund
ZIH Dresden
Universität Freiburg
GWG Göttingen
DESY Hamburg
DKRZ Hamburg
RRZN Hannover
ARI Heidelberg
FZJ Jülich
FZK Karlsruhe
Universität Mainz
Universität Leipzig
Universität Marburg
LMU München
Universität Oberpfaffenhofen
FAU Oldenburg
DFK Potsdam
FhG SCAI St. Augustin
Universität Siegen
HLRS Stuttgart
Universität Wuppertal

Plus ca. 10 weitere Standorte

10.000 CPU-Cores, 3,3 PB Disk, 5,5 PB Tape



<http://www.d-grid.de>



Zusammenfassung und Ausblick

- 2003 war Deutschland auf der "Grid-Weltkarte" ein nahezu weißer Fleck
- Die e-Science-Initiative des BMBF hat diese Situation nachhaltig verändert
- Die ab 2008 verfügbare Kern-D-Grid-Infrastruktur ist international vergleichbar und konkurrenzfähig
- Die z. Zt. 19 geförderten D-Grid-Projekte decken einen großen Teil wissenschaftlicher und wirtschaftlicher Themengebiete ab
- Die nächste Grid-Ausschreibung des BMBF wird in Kürze kommen
 - weitere 25 bis 30 Mio Euro, weitere ca. 15 Grid-Projekte
 - Schwerpunkte: SLAs, Community-Service-Grids, Grid-basierte Wissensschicht
- Langfristig wird D-Grid zu einer nachhaltigen Infrastruktur für die deutsche Wissenschaft und Wirtschaft werden
- D-Grid fördert die Wettbewerbsfähigkeit der deutschen Wissenschaft und Wirtschaft

D a n k e

