Grid@Asia, Seoul, December 12, 2006

D-Grid

The German Grid Initiative in International Context

Wolfgang Gentzsch
Coordinator D-Grid

Ulrich Sax
MediGRID

Grid@Asia, Seoul, December 12, 2006

D-Grid

The German Grid Initiative in International Context

Wolfgang Gentzsch
Coordinator D-Grid

Ulrich Sax
MediGRID

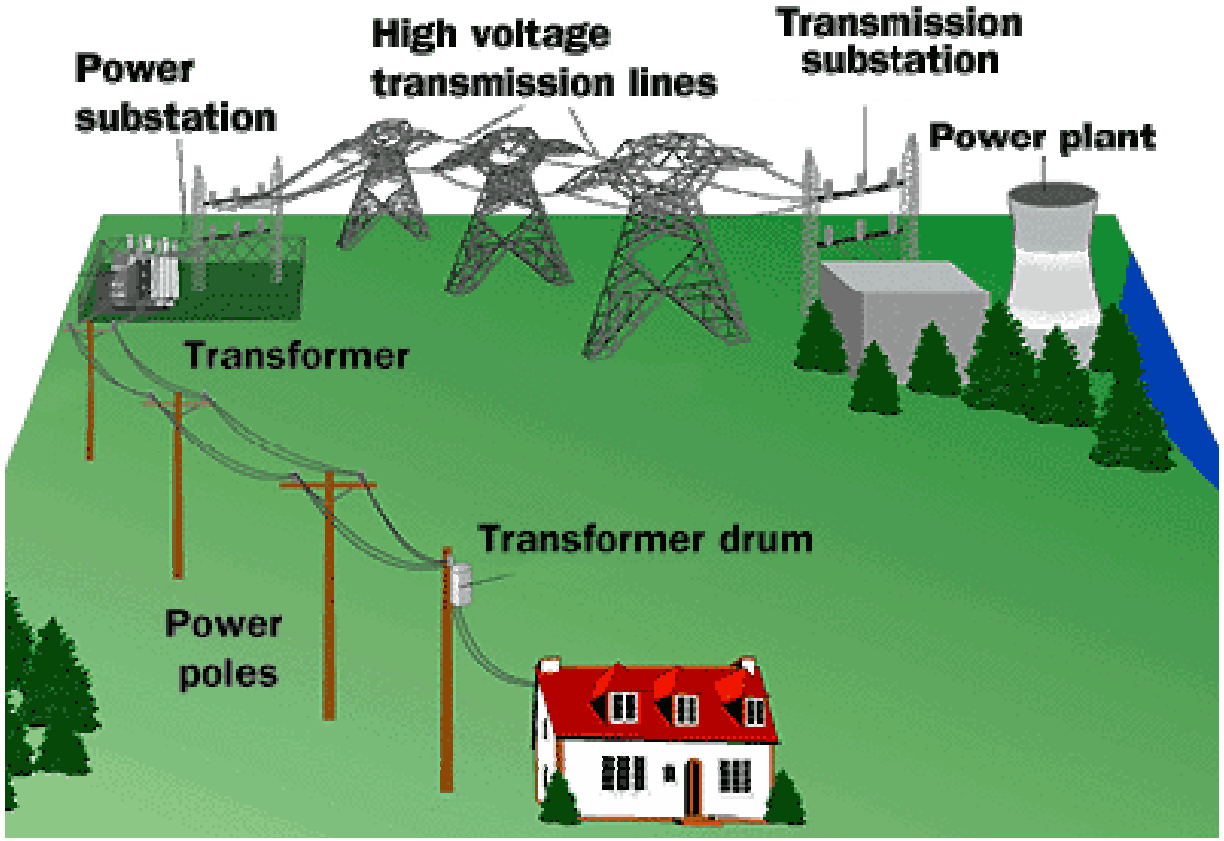
Thanks to:

Tony Hey, Satoshi Matsuoka, Hai Jin, Bob Jones, Charlie Catlett, Dane Skow
and the Renaissance Computing Institute at UNC Chapel Hill, North Carolina
Thomas Steinke, Jürgen Falkner

Our Topics this Morning

- My favorite Grid definition
- Example: The German D-Grid Initiative
- The international ‘Grid’scape
- 10 * Grid in Context
- Grid challenges still to be solved
- What’s next ?

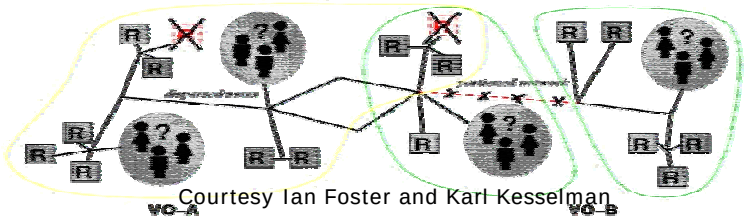
The Electrical Power Grid



What is a Grid ?

Courtesy CERN

- Distributed, networked computing & data resources
- The underlying IT infrastructure for global HPC
- Networking and computing infrastructure for utility computing
- Distributed platform for sharing scientific experiments and instruments
- The next generation of enterprise IT architecture
- The next generation of the Internet and the WWW
- Computing from the wall socket
- … and more …



Courtesy Ian Foster and Karl Kesselman

Benefits of Grid Computing

- Resource Utilization: increase from 20% to 80+%
- Productivity: more work done in shorter time
- Business Agility: flexible actions and re-actions
- On Demand: get resources, when you need them
- Easy Access: transparent, remote, secure
- Sharing: enable collaboration over the network
- Failover: migrate/restart applications automatically
- Resource Virtualization: access compute services, not servers
- Heterogeneity: platforms, OSs, devices, software
- Virtual Organizations: build & dismantle on the fly

Community Grids are all about:

- Sharing Resources:
  - Small, medium, large enterprises share networks, computers, storage, software, data, . . .
  - Researchers share ditto and large experiments, instruments, sensor networks, etc.
- Collaboration:
  - Enterprise departments with its suppliers and peers (e.g. design)
  - Research teams distributed around the world (HEP, Astro, Climate)
- Doing things which have not been possible before:
  - Grand Challenges needing huge amount of computing and data
  - Combining distributed datasets into on virtual data pool (Genome)
  - “Mass Grids” for the people (distributed digital libraries; digital school laboratories; etc)

The German D-Grid Initiative *)

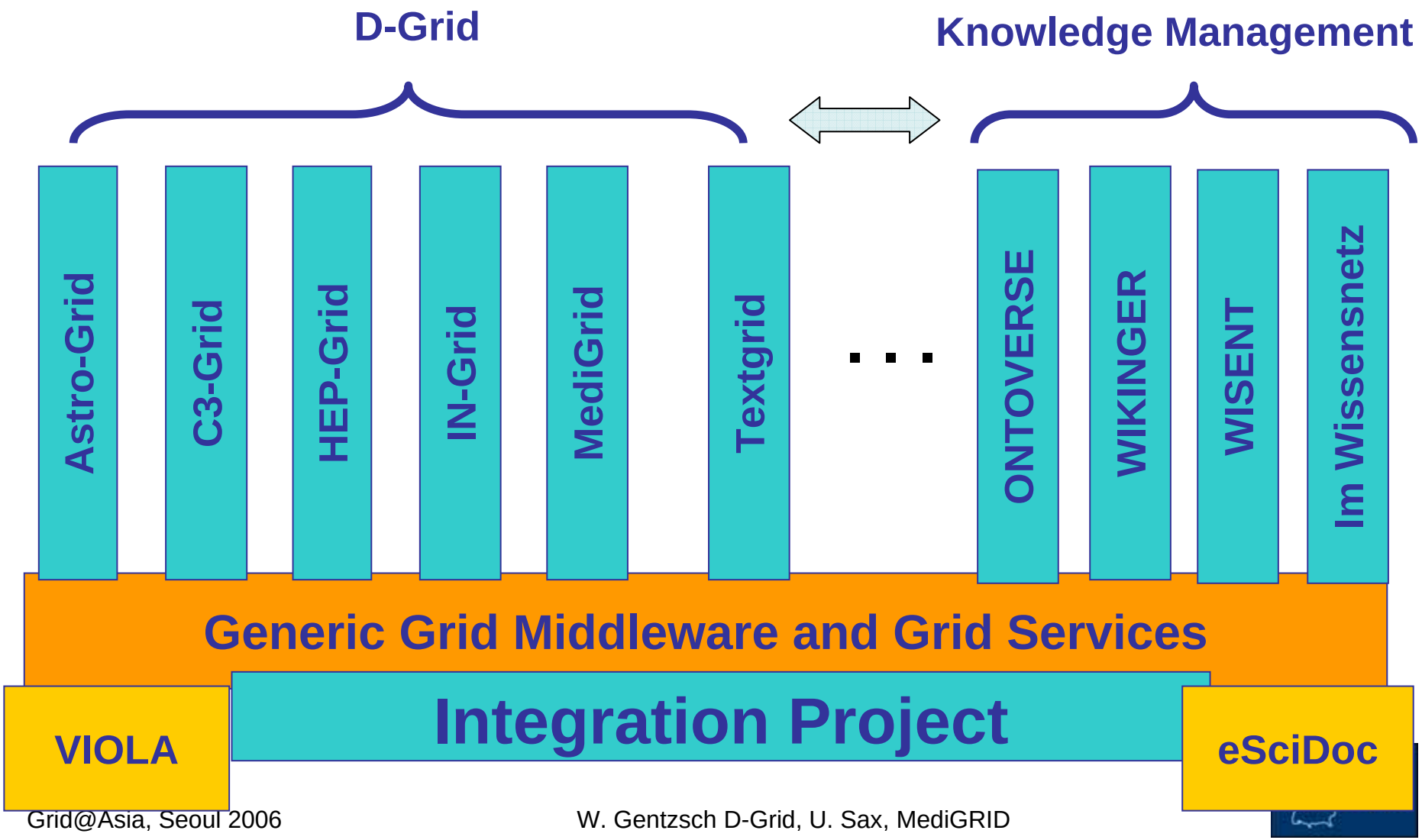
D-Grid-1
Services for Scientists

*) funded by the German Ministry for Education and Research

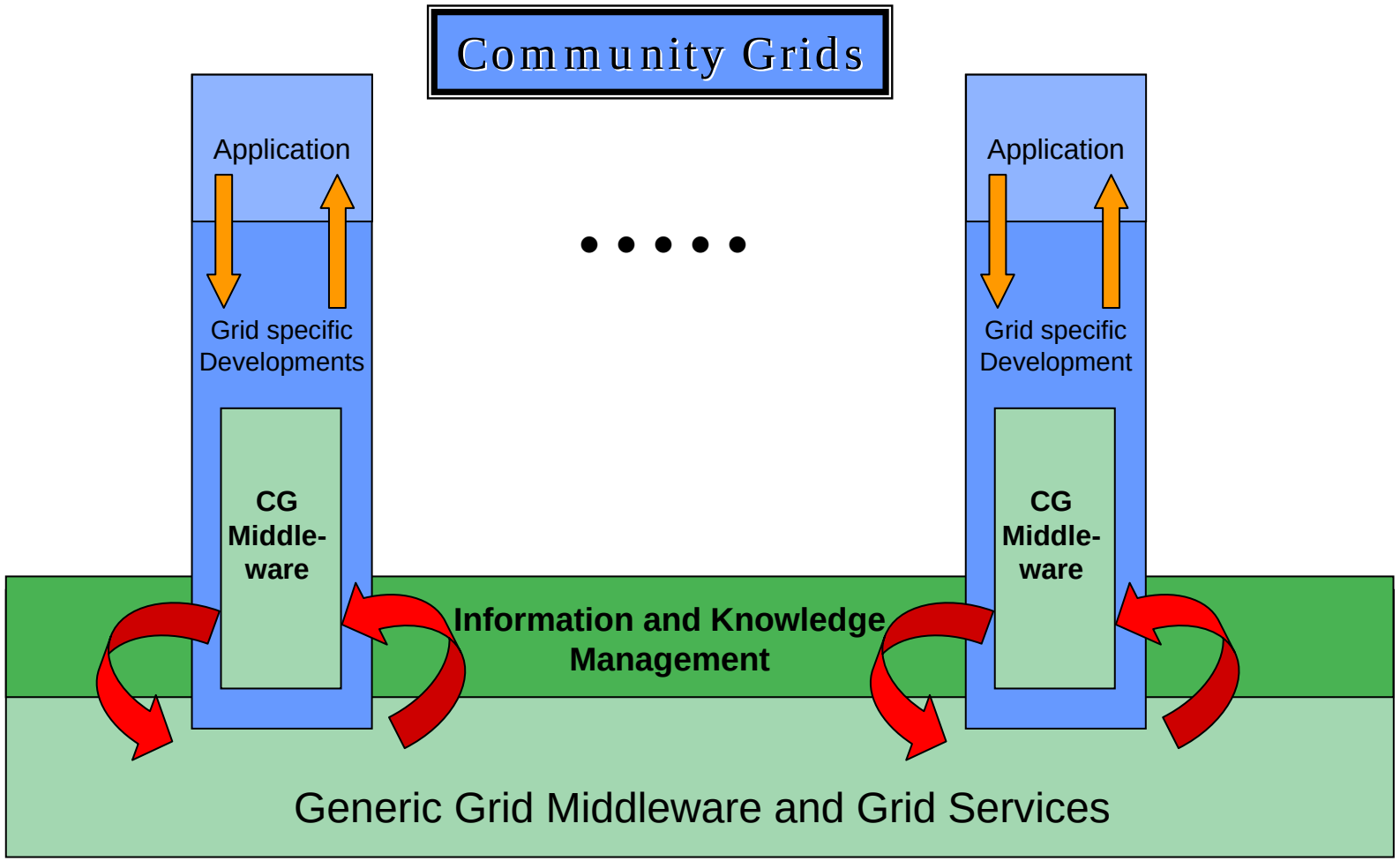
German e-Science Initiative, Key Objectives

- **Building a Grid Infrastructure in Germany**
  - Combine the existing German grid activities for infrastructure, middleware, and applications
  - Integration of the middleware components developed in the Community Grids
- **Development of e-science services for the research community**
  - Science Service Grid
- **Important:**
  - Continuing sustainable production grid infrastructure after the end of the funding period
  - Integration of new grid communities (2. generation)
  - Business models for grid services

D-Grid Projects



D-Grid Structure



Courtesy Dr. Krahl PT/BMBF

Integration Project

DGI Infrastructure Project

WP 1: D-Grid basic software components, sharing resources, large storage, data interfaces, virtual organizations, management

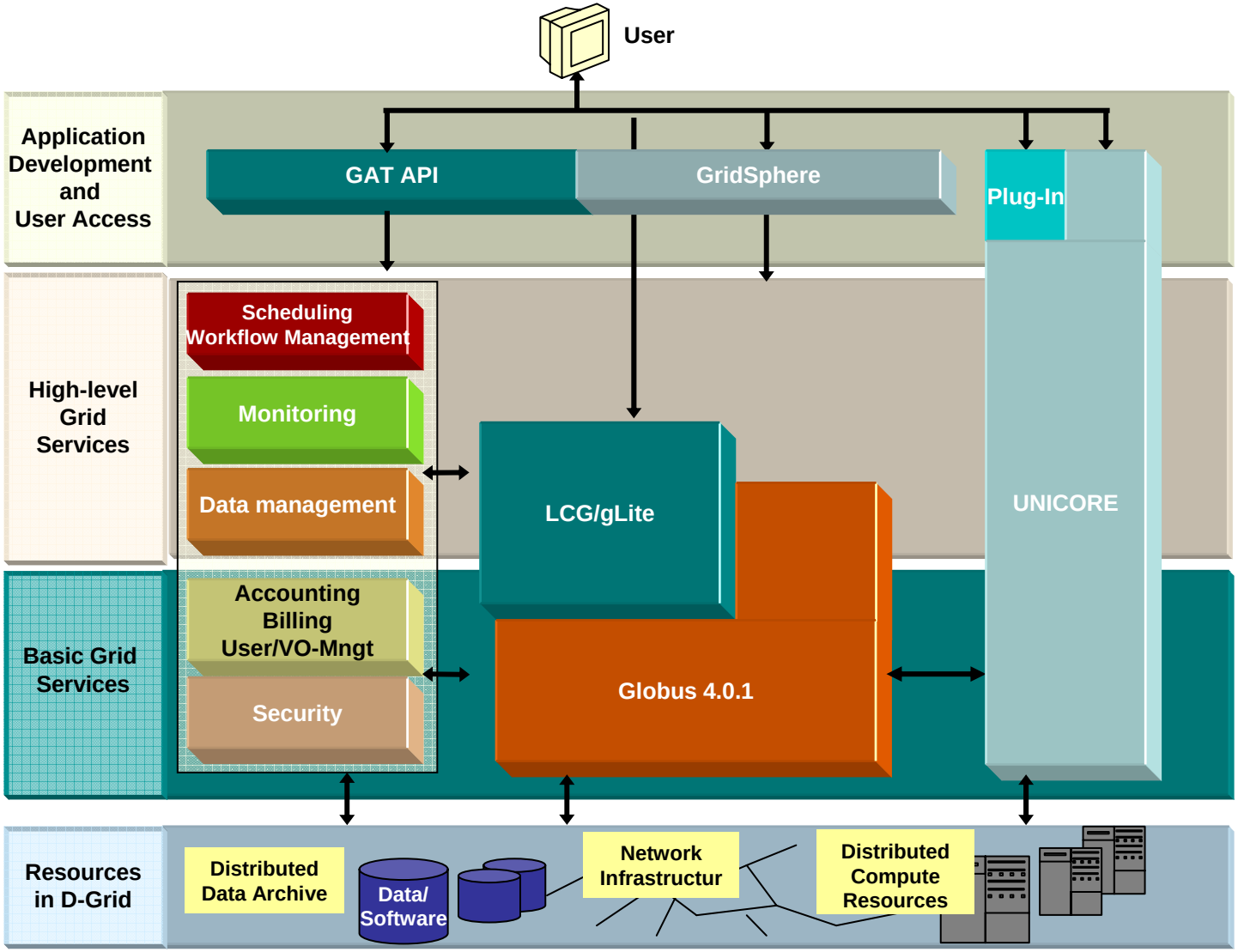
WP 2: Develop, operate and support robust core grid infrastructure, resource description, monitoring, accounting, and billing

WP 3: Network (transport protocols, VPN), Security (AAI, CAs, Firewalls)

WP 4: Business platform and sustainability, project management, communication and coordination

- **Scalable, extensible, generic grid platform for future**
- **Longterm, sustainable grid operation, SLAs based**

D-Grid Middleware



DGI Services, Available Dec 2006

- Sustainable grid operation environment with a set of core D-Grid middleware services for all grid communities
- Central registration and information management for all resources
- Packaged middleware components for gLite, Globus and Unicore and for data management systems SRB, dCache and OGSA-DAI
- D-Grid support infrastructure for new communities with installation and integration of new grid resources into D-Grid Help-Desk, Monitoring System and central Information Portal

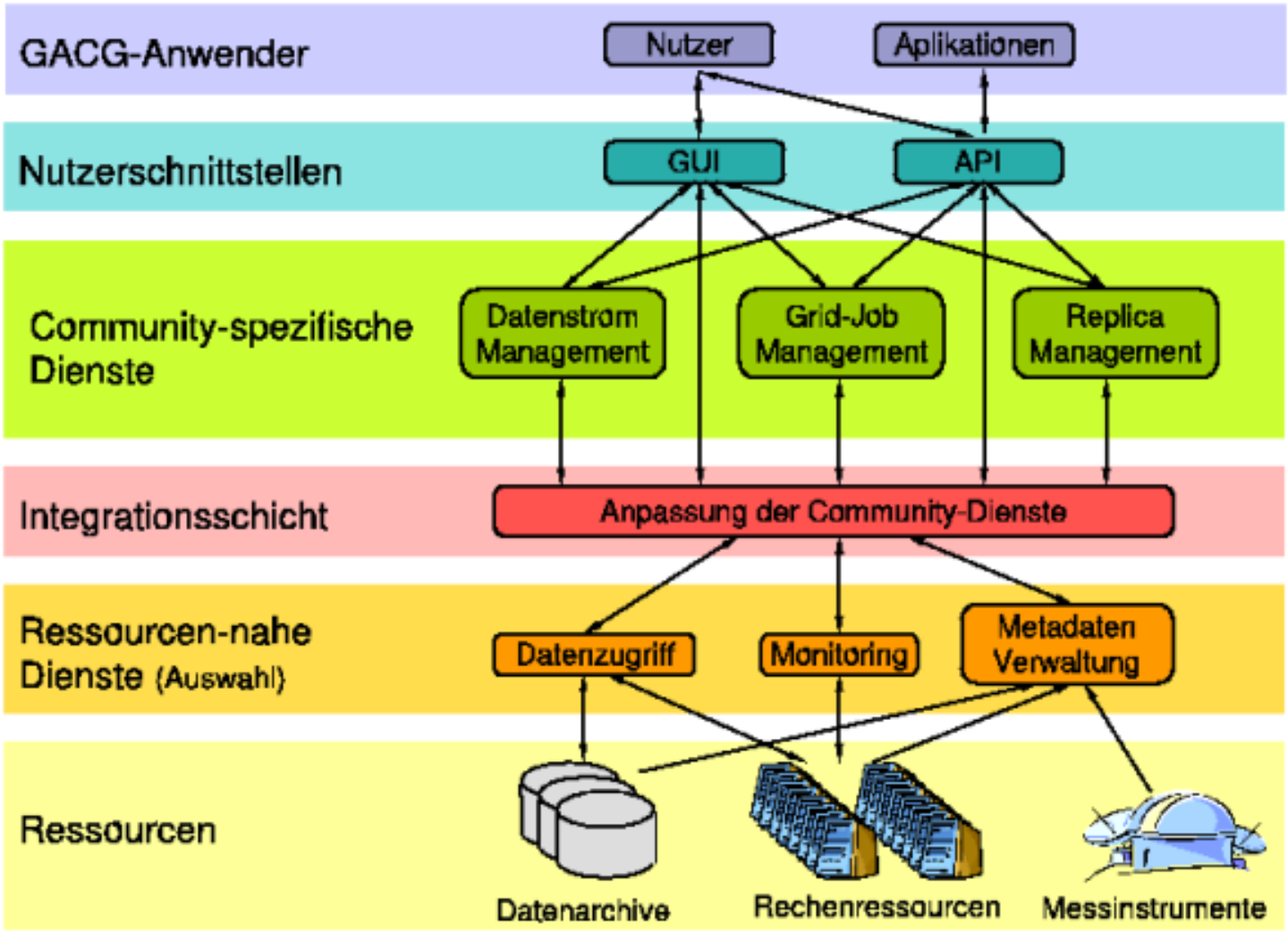
DGI Services, Dec 2006, cont.

- Tools for managing VOs based on VOMS and Shibboleth
- Test implementation for Monitoring & Accounting for Grid resources, and first concept for a billing system
- Network and security support for Communities (firewalls in grids, alternative network protocols,...)
- DGI operates „Registration Authorities“, with internationally accepted Grid certificates of DFN & GridKa Karlsruhe
- Partners support new D-Grid members with building their own „Registration Authorities“

DGI Services, Dec 2006, cont.

- DGI will offer resources to other Communities, with access via gLite, Globus Toolkit 4, and UNICORE
- Portal-Framework Gridsphere can be used by future users as a graphical user interface
- For administration and management of large scientific datasets, DGI will offer dCache for testing
- New users can use the D-Grid resources of the core grid infrastructure upon request

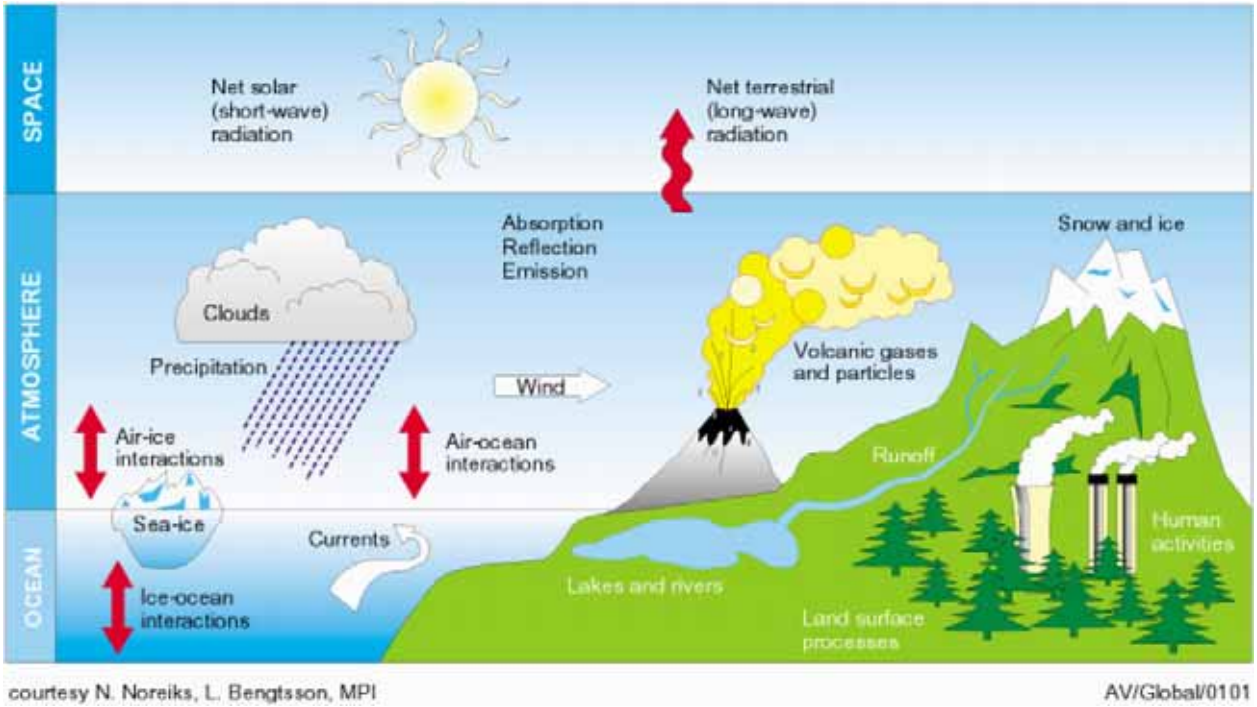
AstroGrid


GACG-Anwender
Nutzer
Aplikationen
Nutzerschnittstellen
GUI
API
Community-spezifische Dienste
Datenstrom Management
Grid-Job Management
Replica Management
Integrationsschicht
Anpassung der Community-Dienste
Ressourcen-nahe Dienste (Auswahl)
Datenzugriff
Monitoring
Metadaten Verwaltung
Ressourcen
Datenarchive
Rechenressourcen
Messinstrumente

C3 Grid: Collaborative Climate Community Data and Processing Grid

Climate research moves towards new levels of complexity:

Stepping from Climate (=Atmosphere+Ocean) to Earth System Modelling



Increased Computational demand factor: O(1000 -10000)

Earth system model wishlist:

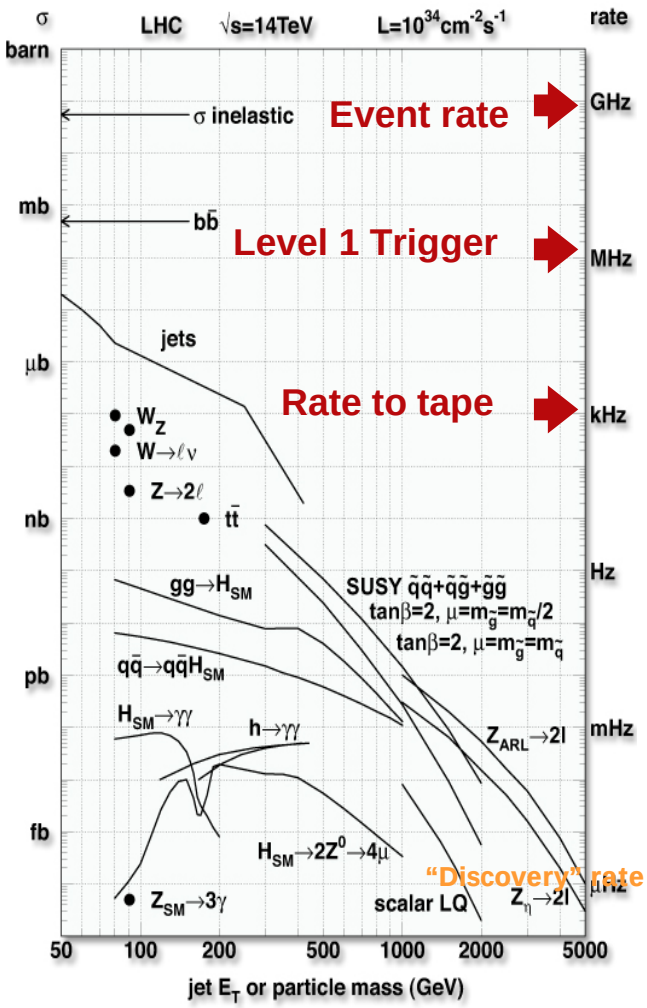
Higher spatial and temporal resolution

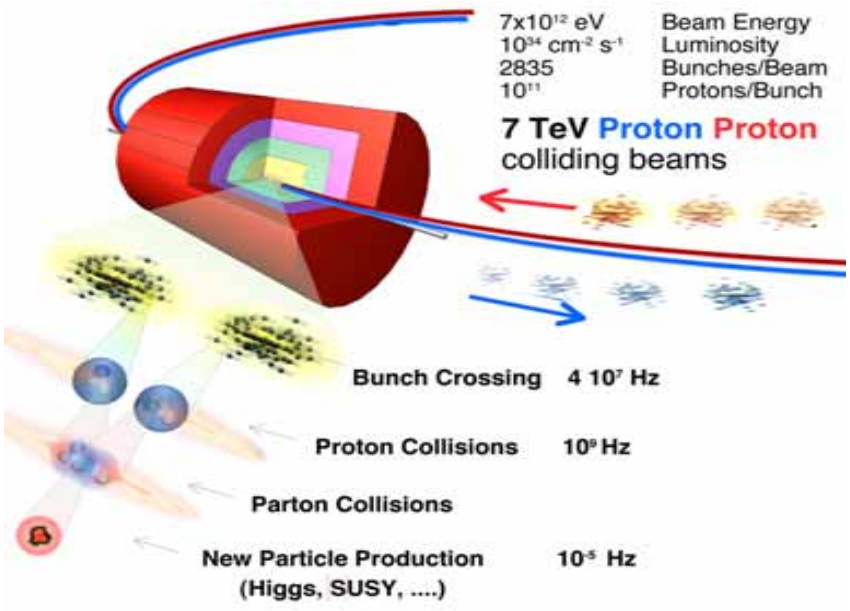
Quality: Improved subsystem models

Atmospheric chemistry (ozone, sulfates,..)

Bio-geochemistry (Carbon cycle, ecosystem dynamics,..)

HEP-Grid: p-p collisions at LHC at CERN
(from 2007 on)





Crossing rate	40 MHz
Event Rates:	~10^9 Hz
Max LV1 Trigger	100 kHz
Event size	~1 Mbyte
Readout network	1 Terabit/s
Filter Farm	~10^7 Si2K
Trigger levels	2
Online rejection	99.9997 % (100 Hz from 50 MHz)
System dead time	~ %
Event Selection:	~1/10^{13}

Luminosity
Low 2x10^{33} cm^{-2} s^{-1}
High 10^{34} cm^{-2} s^{-1}
Data analysis:
~1PB/year

Courtesy David Stickland

InGrid: Virtual Prototyping & Modeling in Industry

Molding | **Metal Forming** | **Fluid Processes** | **Groundwater Transportation** | **Fluid-Structur / Magneto-Hydro-dynamic Interaction**

Methods and models for solving engineering problems in Grids

- Knowledge-based support for engineering-specific decision support
- Support for engineering-specific Workflows
- Distributed simulations-based product & process optimization

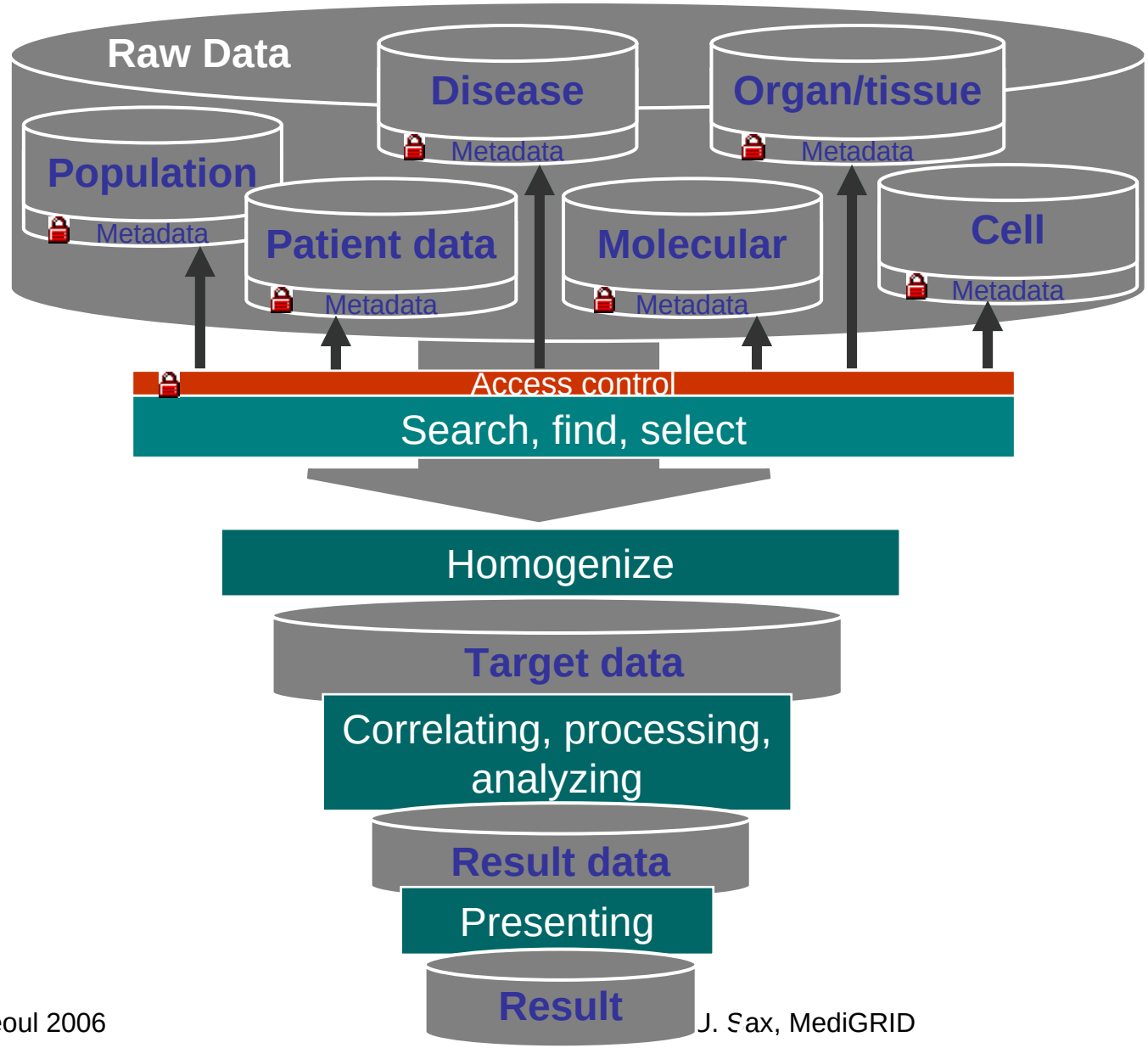
Grid-specific developments

- Cooperation and business models
- Security and trust models

Integration project

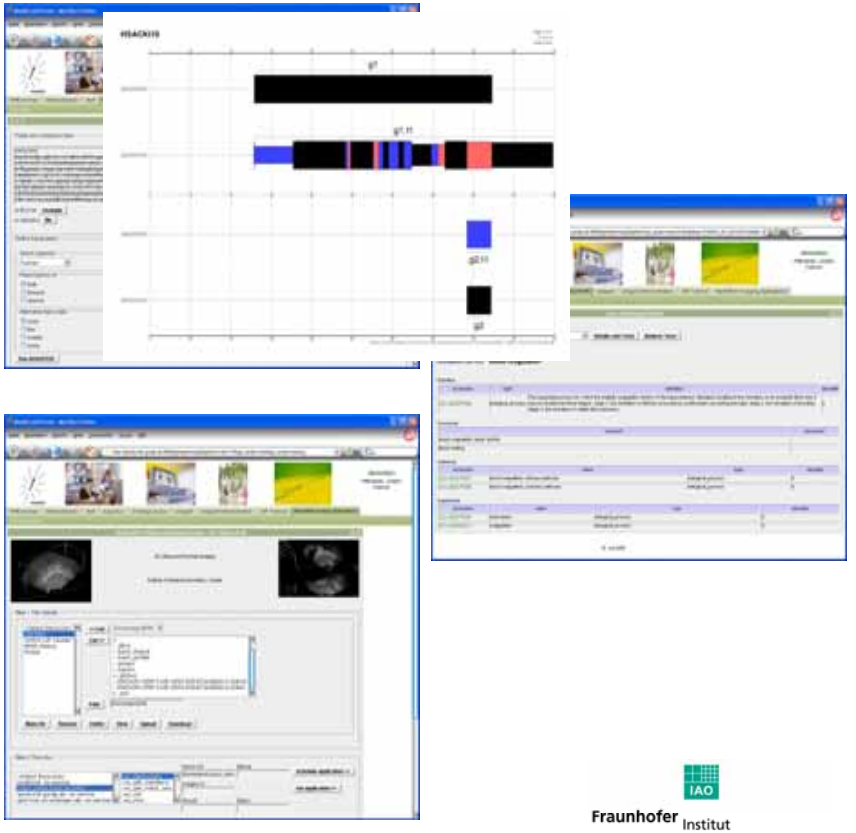
AP 2 AP 3 AP 4

MediGRID


Raw Data
Disease
Metadata
Organ/tissue
Metadata
Population
Metadata
Patient data
Metadata
Molecular
Metadata
Cell
Metadata
Access control
Search, find, select
Homogenize
Target data
Correlating, processing, analyzing
Result data
Presenting
Result

Pilot Applications

- Integrated in the MediGRID Portal:
  - AUGUSTUS:
    Genome sequence analysis
  - Ontology-Access:
    with OGSA-DAI-Service
  - Medical Imaging
    - 3D US Prostate biopsy
    - Virtual vascular surgery
  - In the pipeline:
    - clinical studies Neurology



Fraunhofer Institut Arbeitswirtschaft und Organisation

MediGRID US-Workshop November 11-15, 2006

v.l.n.r. Yannick Legré (HealthGrid EU, Port-au-Chevau), Otto Rienhoff (MediGRID, Univ. Göttingen), Peter Covitz (caBIG, NCICB, Washington), Berit Hamer (Univ. Göttingen), Dagmar Krefting (MediGRID, Charité Berlin), Howard Bilofsky (US HealthGrid, Univ. of Pennsylvania, Philadelphia), Parvati Dev (US HealthGrid, University School of Medicine, Stanford), Michael Hartung (MediGRID, Univ. Leipzig), Anette Weisbecker (MediGRID, Fraunhofer IAO, Stuttgart), Jochen Hampe (MediGRID, UKSH Kiel), Sebastian Claudius Semler (MediGRID, TMF, Berlin), Thomas Steinke (MediGRID, Zuse Institute Berlin).

D-Grid-2 Call

- 'Horizontal' Service Grids: professional Service Providers for heterogeneous user groups in research and industry

- 'Vertical' Community Service Grids using existing D-Grid infrastructure and services, supported by Service Providers

- D-Grid extensions, based on a D-Grid 1 gap analysis
  - Tools for operating a professional grid service
  - Adding business layer on top of D-Grid infrastructure
  - Pilot service phase with service providers and 'customers'

!! Reliable grid services require sustainable grid infrastructure !!

Challenges, Potential Grid Inhibitors

- Sensitive data, sensitive applications (medical patient records)
- Accounting, who pays for what (sharing!)
- Security policies: consistent and enforced across the grid !
- Lack of standards prevent interoperability of components
- Current IT culture is not predisposed to sharing resources
- Not all applications are grid-ready or grid-enabled
- Open source is not equal open source (read the small print)
- SLAs based on open source (liability?)
- “Static” licensing model don’t embrace grid
- Protection of intellectual property
- Legal issues (FDA, HIPAA, multi-country grids)

Our Vision :

The Three Waves of Grid Computing

2000

The Research Wave

Technology, Prototypes
Virtual Organizations
Standards
GGF, IETF

2004

The Industry Wave

Grid-Enabled Products
Enterprise Solutions
Interoperability
GGF, EGA, IETF, OASIS

2008

The Consumer Wave

Commodity
IT Utility
Integration
Legal, Ethical, Political Orgs

Grid is a Journey . . .

Old World	New World
Static	Dynamic
Silo	Shared
Physical	Virtual
Manual	Automated
Application	Service

Transitioning from **S**ilo **O**riented **A**rchitecture
to
Service **O**riented **A**rchitecture

Finally:
Grid 2.0 for Web 2.0

Anyone, anywhere, anytime, any device, connected to a Grid

- Policies, SLAs, grid economy, to maintain reliability stability and efficiency
- Integration of new devices, data and information sources: e.g. Cell phones, PDAs, smart sensors, sensor arrays, health monitors
- Devices embedded in cars, engines, roads, bridges, clothes,...
- Handle huge amount of data for real-time analysis
- Bridges political, organizational, societal boundaries

… enabling ‘equal opportunity’ for our fellow citizens

The Grid Engine

The Combustion Engine

The Steam Engine

Thank You !

Slides are available

usax@med.uni-goettingen.de
wgentzsch@d-grid.de