e-Infrastructures for data and compute intensive applications

Internet, Web, Grids, Clouds, how do they come together?

Wolfgang Gentzsch
The DEISA Project & Board of Directors of OGF
gentzsch at rzg.mpg.de
Content

- Service Infrastructures
- Components: HPC, Grids, Clouds, Internet, Web,...
- Example: DEISA Ecosystem for HPC Applications
- Example: The Telecoms and their Potential
- Next-Generation e-Infrastructure: Digital City
Service Infrastructures

Ancient Rome: ten aqueducts, some 150,000 m³ of water each day

Electrical Power Grid Infrastructure

Transportation Grids

EGEE – Enabling Grid in E-SciencE
HPC Centers
(in the light of Grids and Clouds)

- HPC Centers are service providers, for past 35 years
- Computing, storage, applications, data, etc IT services
- Serve (local) research, education, industry
- Very professional: to end-users, they look (almost) like Grid or Cloud services

(Amazon Cloud definition: easy, secure, flexible, on demand, pay per use, self serve)
Cloud... as a Service

Cloud: dynamically scalable and virtualized resources provided as a service over the Internet

Infrastructure (IaaS)
- Accessible online, anytime, anywhere
- Pay for what you use
- Available on demand
- Service Level Agreements
- Automated:
  - Scalability
  - Failover
  - Concurrency management

Platform (PaaS)

Software (SaaS)
The Cloud of Cloud Companies

- Amazon
- Google
- Salesforce
- Microsoft
- Sun
- IBM
- Oracle
- EMC
- Cloudera
- Cloudsoft
- Akamai
- Areti Internet
- Enki
- Fortress ITX
- Joyent
- Layered Technologies
- Rackspace
- Terremark
- Xcalibre
- Manjrasoft/ANEKA
- ...

DEISA

InfoSys, April 20-25, 2009

Wolfgang Gentzsch, DEISA
e-Infrastructure Challenges (Inhibitors)

- Sensitive data, sensitive applications (med.patient records)
- Different organizations have different ROI
- Accounting, who pays for what (sharing!)
- Security policies: consistent and enforced across the grid!
- Interoperability of components and grids (standards ?)
- 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 little 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)
Examples of a successful Research e-Infrastructures
BIRN – Biomedical Information

The Biomedical Informatics Research Network (BIRN) Portal provides BIRN members with a single sign on web portal to access data grid files, computation grid resources, and a variety of collaboration tools to facilitate the scientific needs of BIRN researchers. Non-BIRN participants may access the portal through a guest registration.

Portal Requirements

- You must have cookies enabled to login to the BIRN Portal, in addition, Javascript is highly recommended but not required.
- The latest version of Java will be required to access some of the applications.
- For optimal browsing please use a Mozilla based browser.
- Other versions of Safari will support a basic experience.
myGrid is a collection of services and components that allows the high level integration of biological applications. The architecture provides the infrastructure necessary, in a web service environment, for an e-science workbench that actively supports the scientific lifecycle. Each component or service contributes to a system that allows the e-scientist to perform complex in-silico experiments across distributed bioinformatics resources.
NEESGrid

Realtime access to earthquake Shake table experiments at remote sites.
Renci Bio Portal

Providing access to biotechnology tools running on a back-end Grid.

- leverage state-wide investment in bioinformatics
- undergraduate & graduate education, faculty research
- another portal soon: national evolutionary synthesis center
ServoGrid Portal

Welcome to the QUAKE SIM Computational Portal

SERVO Grid
Solid Earth Research Virtual Observatory Grid

- QuakeSim home page
- Old GEM General Earthquake Modeling Web Site
- SLIDE Distributed File System for NASA Computational Technology Project
- Report from the Earth Science Enterprise Computational Technology Requirements Workshop April 30-May 1 2002 where SERVO concept first introduced
- Discover the Grid at the Grid Forum or at this collection of papers
- Other collected papers and presentations on SERVOGrid and related topics are available from the Community Grids Lab publications page.

Participating Institutions:
IU CGL | NASA JPL | UC Davis | UC Irvine | USC
Example of an e-Infrastructure:

The DEISA Ecosystem for HPC Grand-Challenge Applications

Distributed European Infrastructure for Supercomputing Applications
new “petaflop” supercomputers

PRACE
petaflop supercomputers

EU

DEISA
virtual supercomputer

Local and regional
supercomputers

Mario Campolargo
European Commission
OGF23, June 2008
DEISA: Vision and Mission

Vision:
Persistent European HPC ecosystem integrating Tier-1 (Tflop/s) centres and European Tier-0 (Pflop/s) centres.

Mission:
Enhance Europe’s capability in computing and science by integrating most powerful supercomputers into a European HPC e-infrastructure.

Build European Supercomputing Service on top of existing national services, based on the deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope.
DEISA Service Layers

- Multiple ways to access
- Workflow management
- Common production environment

- Single monitor system
- Job rerouting
- Co-reservation and co-allocation

- Data staging tools
- Data transfer tools
- WAN shared File system

- Unified AAA
- DEISA Sites
- Network connectivity

- Presentation layer
- Job management layer and monitor
- Data management layer
- Network and AAA layers
DEISA Global File System

Global transparent file system based on the Multi-Cluster General Parallel File System (MC-GPFS of IBM)

InfoSys, April 20-25, 2009

Wolfgang Gentzsch, DEISA
Management of users in DEISA

- A dedicated LDAP-based distributed repository administers DEISA users
- Trusted LDAP servers are authorized to access each other (based on X.509 certificates) and encrypted communication is used to maintain confidentiality
One Example of Virtual Communities: Joint Research Activity “Life Sciences”

The DEISA Life Science Portal

Joint Research Activity (JRA)

Promoting parallel apps in the life science community

Running big simulations on DEISA infrastructure that couldn’t be done locally

Providing ease of access to resources

Application support for life science portal
NICE EnginFrame
Cluster/Grid/Cloud Portal

Provides remote, interactive, transparent, and secure access to applications and data on your corporate Intranet or Internet, or in the Cloud.

Users and administrators can access and control computing resources via an intuitive and standard Web interface virtually anywhere using a standard Web browser.
e-Infrastructure for Business:
Example: Telecommunications
The Core: A scalable Infrastructure

Example: GridwiseTech Momentum
New Powerful End-User Devices

Carry-along PCs (CAPS), Ultra-Mobile PCs (UMPC), Cellphones

- CAPC, from Samsung, South Korea
- T83 Tablet from Asus, Taiwan, demoed at CeBit 2007
- OLPC from MIT
- Asus R2H
- Fujitsu UMPC
- CAPC from HTC
The Symbian OS

Symbian Ltd was founded by Ericsson, Motorola, Nokia and Psion to create Symbian OS

1998

100 million phones shipped

2006

250 different models since Symbian’s creation

2008

250 million phones shipped

2009

• Nokia acquires Symbian Ltd
• The Symbian Foundation is announced by the initial board members.

Courtesy: Lars Kurth, Symbian Foundation
The Symbian Platform

(code on the device)

The platform enables different groups of developers

- **Web Applications**: WebKit, Flash Lite
- **Scripting Languages**: Python, Ruby, Perl
- **Java**: Java ME, eRCP
- **Native run-times**: Qt, Open C/C++ (=POSIX & Standard C++ & Boost), .NET
- **Native development**: Symbian C++
A Peek at Intel’s Digital City Vision

Imagine a city...

That Transforms the way we Work, Live & Play...

Courtesy Robert Fogel, Intel
Today’s Digital Challenge

- Banking
- Working
- Shopping
- Taxes
- Multiple Identities
- Incompatible Networks
- Disconnected Agencies

- DEISA
- Distributed European Infrastructure for Supercomputing Applications

- Wimax
- WiFi
- Wired
- Cellular
- Business
- Citizen
- Govt
- Employee

InfoSys, April 20-25, 2009
Wolfgang Gentzsch, DEISA
Tomorrow's Integrated Digital City

- A single, portable identity based on strong security
- An intelligent infrastructure supporting seamless access
- A gateway to integrated e-Services spanning multiple agencies

One Identity
One Network
One Face
Facets of the Digital City

Serving Citizens

Digital Govt GAPP Programs

Digital Education

Economic Vitality (Digital Office)

Digital Healthcare

Digital Home

Bridging the Digital Divide

Safety & Security

The Digital City is the Fabric that Connects the Community
Finally: connected anyone, anywhere anytime, any device

- **Integration** of new devices, data and information sources, PDAs, cell phones, smart sensors, sensor arrays, health monitors
- **Embedded** in cars, engines, roads, bridges, clothes,...
- **Real-time** analysis of huge amount of data
- **Policies**, grid & cloud economy, maintain stability and efficiency
- **Society** support, to bridge political and social boundaries . . .

... very much like any other vital infrastructure today, such as roads, telecom, water, electricity, etc...
More about DEISA in:

Thank You!

Gentzsch@rzg.mpg.de