InfoSys 2009
Valencia, April 20 – 25, 2009

e-Infrastructures
for Compute and Data Intensive Applications

Solving the Grand Challenges in Big Science

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

Distributed European Infrastructure for Supercomputing Applications

InfoSys 2009

ICNS 2009: Networking Services

ICAS 2009: Autonomic and Autonomous Systems

INTENSIVE 2009: Intensive Applications and Services

all deal with components important for e-Infrastructures
e.g. Grids, Clouds, Internet, sensor nets, agent networks, etc.

DEISA

Distributed
European
Infrastructure for
Supercomputing
Applications

Components of an e-Infrastructure:

HPC Centers,
Grids and Clouds

Terminology

Distributed Computing
- Loosely coupled
- Heterogeneous
- Central management

Cluster
- Tightly coupled
- Homogeneous
- Cooperative working

Grid Computing
- Large scale
- Multi-organizational
- Cross-geography
- Distributed management

Utility Computing
- Cloud services
- Pay per use
- Grid Techn. & Virtualization

HPC Centers

Distributed European Infrastructure for Supercomputing Applications

- HPC Centers are **service providers**, for past 35 years
- Computing, storage, applications, data, etc IT services
- Serve (local) research, education, and industry
- Very professional: to end-users, they look (almost) like Cloud services (Amazon Cloud definition: easy, secure, flexible, on demand, pay per use, self serve)
- But: no virtualization, semi-automatic, not flexible
- HPC centers can become a Cloud customer for dynamic scaling and adopting to changing user demands

Grids

Distributed European Infrastructure for Supercomputing Applications

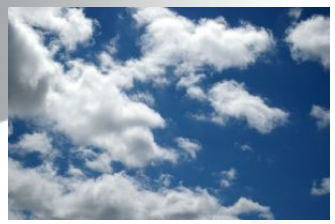
1998: The Grid: Blueprint for a New Computing Infrastructure:

"... hardware and software infrastructure ... **dependable, consistent, pervasive, inexpensive access** to high-end computational capabilities."

2002: The Anatomy of the Grid:

"... coordinated resource **sharing** and problem solving in dynamic, multi-institutional **virtual organizations**."

Quotes: Ian Foster, Carl Kesselman, Steve Tuecke



Clouds

Distributed European Infrastructure for Supercomputing Applications

- IT resources provisioned outside corporate data center
- Resources accessed over the Internet
- A virtual computing environment (Vmware, Xen,...)
- Abstraction of the hardware from the service
- Service oriented: SaaS, PaaS, IaaS, HaaS
- Variable cost of services (QoS)
- Build and deliver, always-on, pay-per-use IT services
- Scaling up/down: computing, storage, database, services, users

Distributed
European
Infrastructure for
Supercomputing
Applications

DEISA

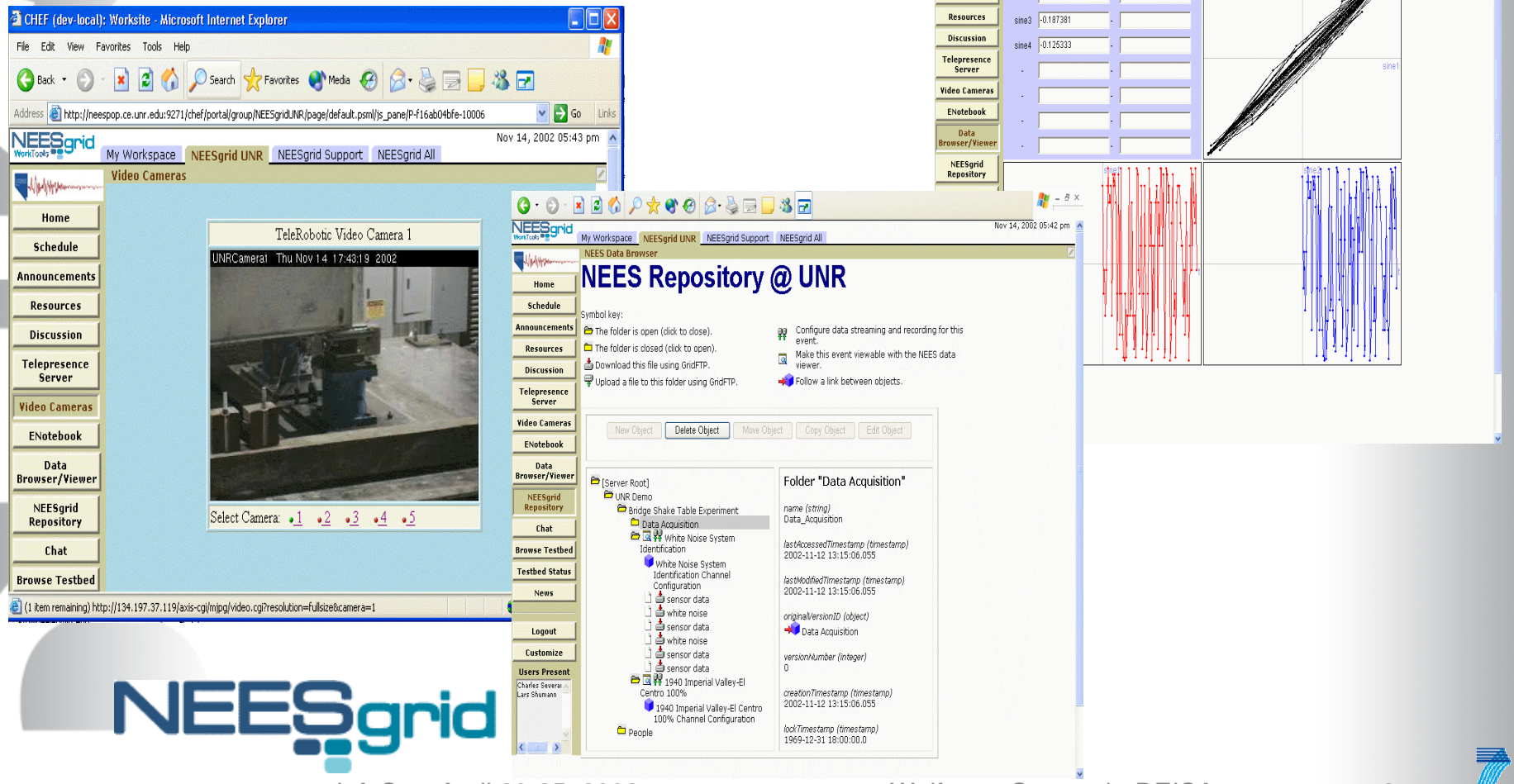
10 Examples of e-Infrastructures

Courtesy: Dennis Gannon

NEESGrid

Distributed European Infrastructure for Supercomputing Applications

Realtime access to earthquake
Shake table experiments at remote sites.



RI-222919

BIRN – Biomedical Information

Geological Information Grid Portal

Distributed European Infrastructure for Supercomputing Applications

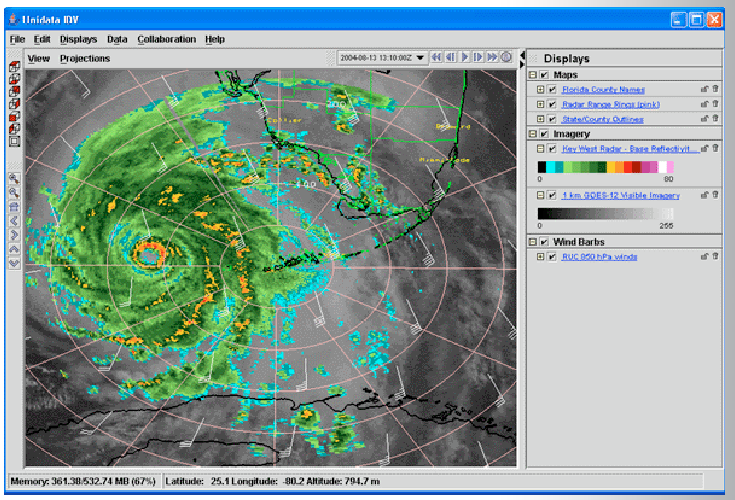
RI-222919

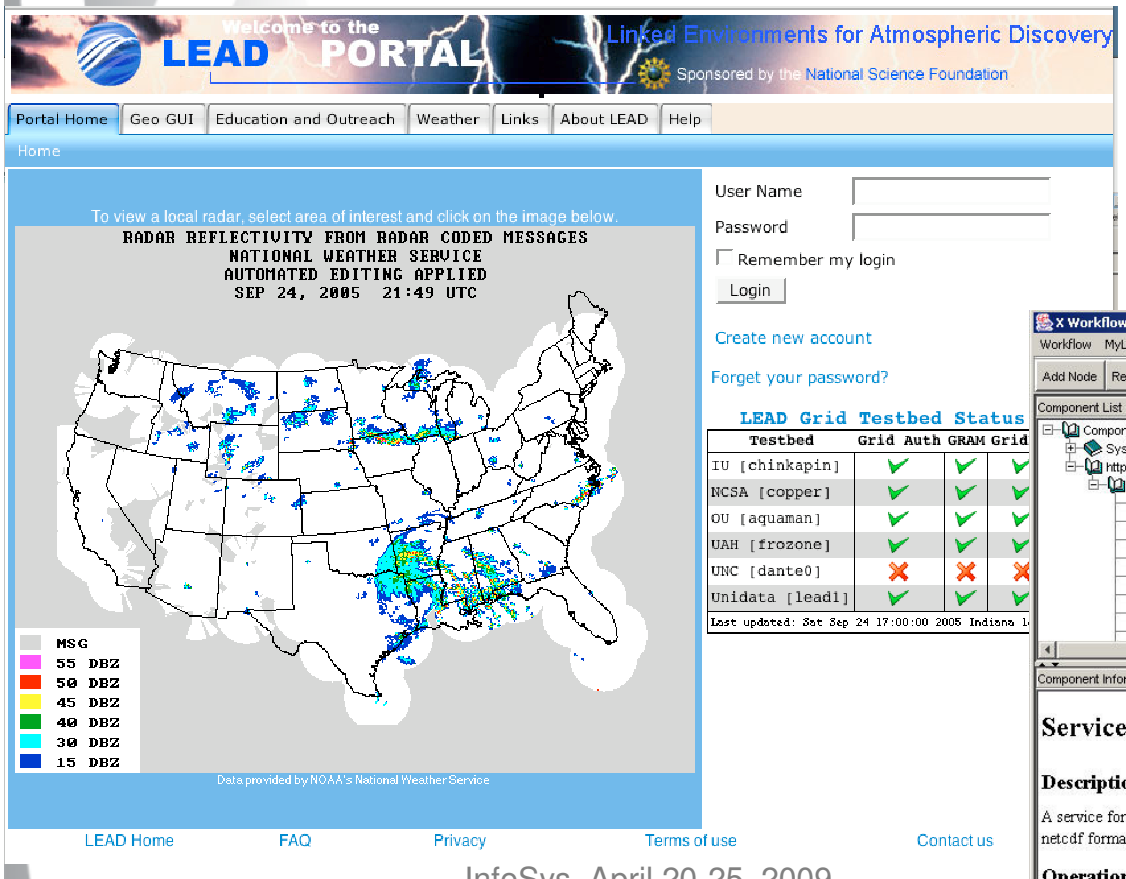
Mesoscale Meteorology

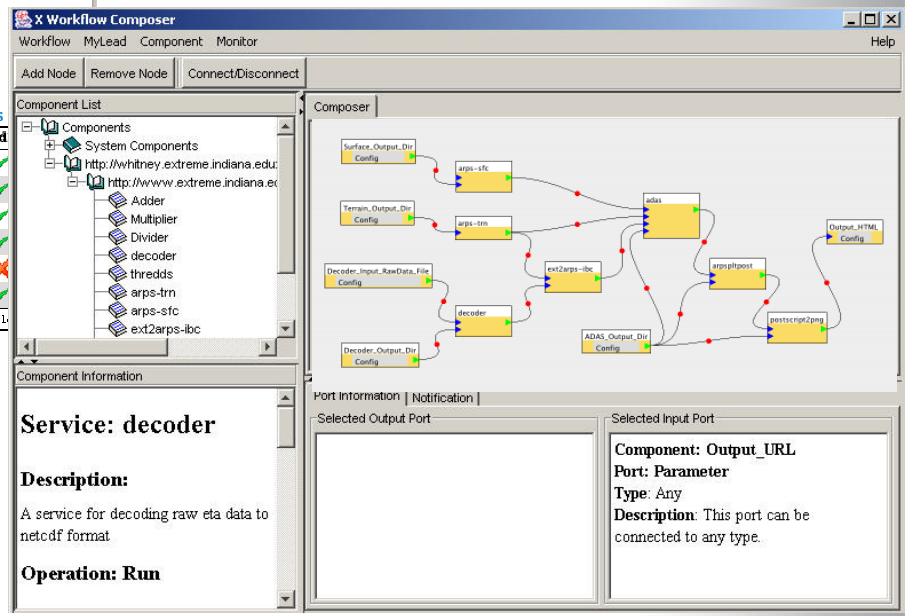
Distributed European Infrastructure for Supercomputing Applications

NSF LEAD project - making the tools that are needed to make accurate predictions of tornados and hurricanes.

- Data exploration and Grid workflow





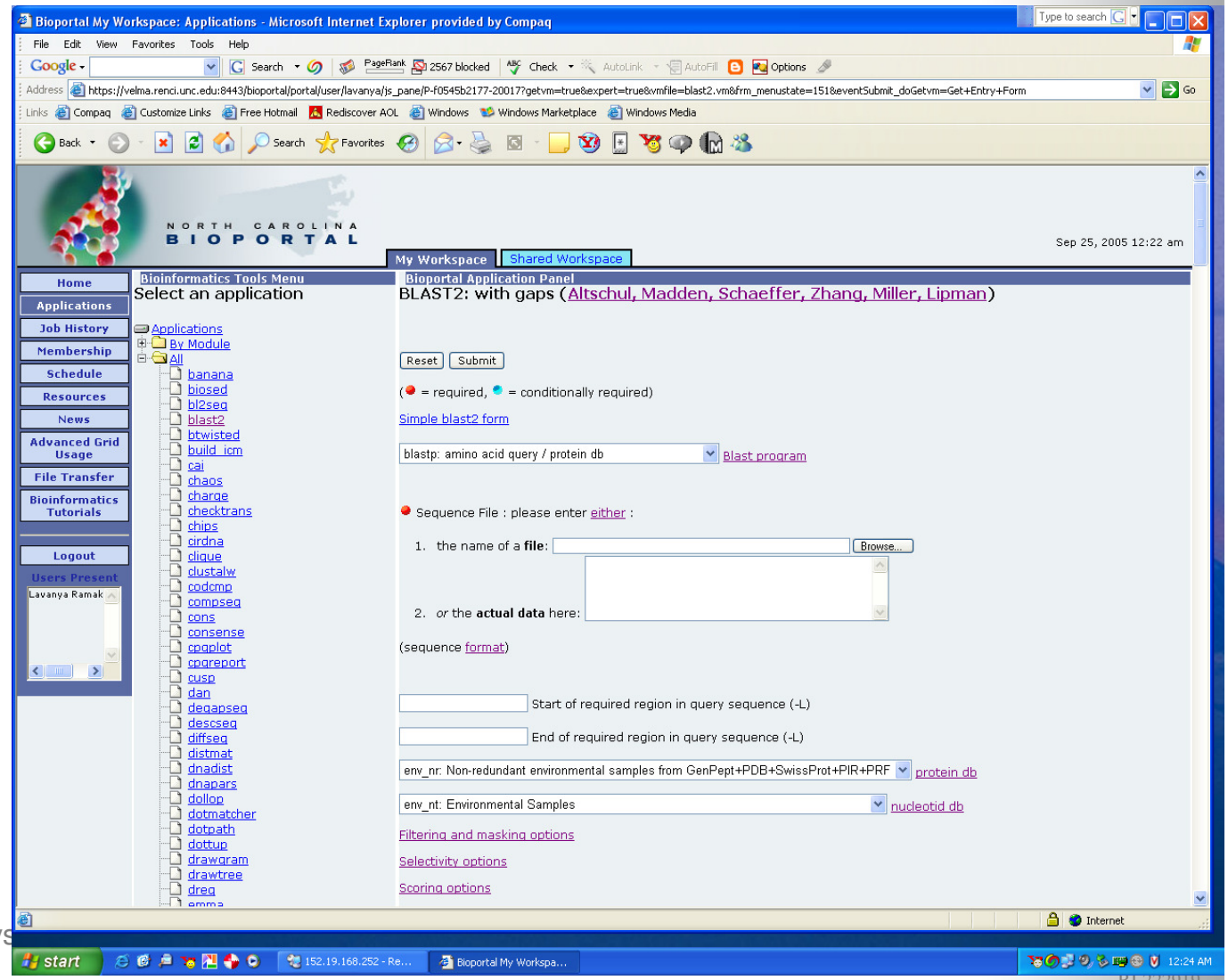


Renci Bio Portal

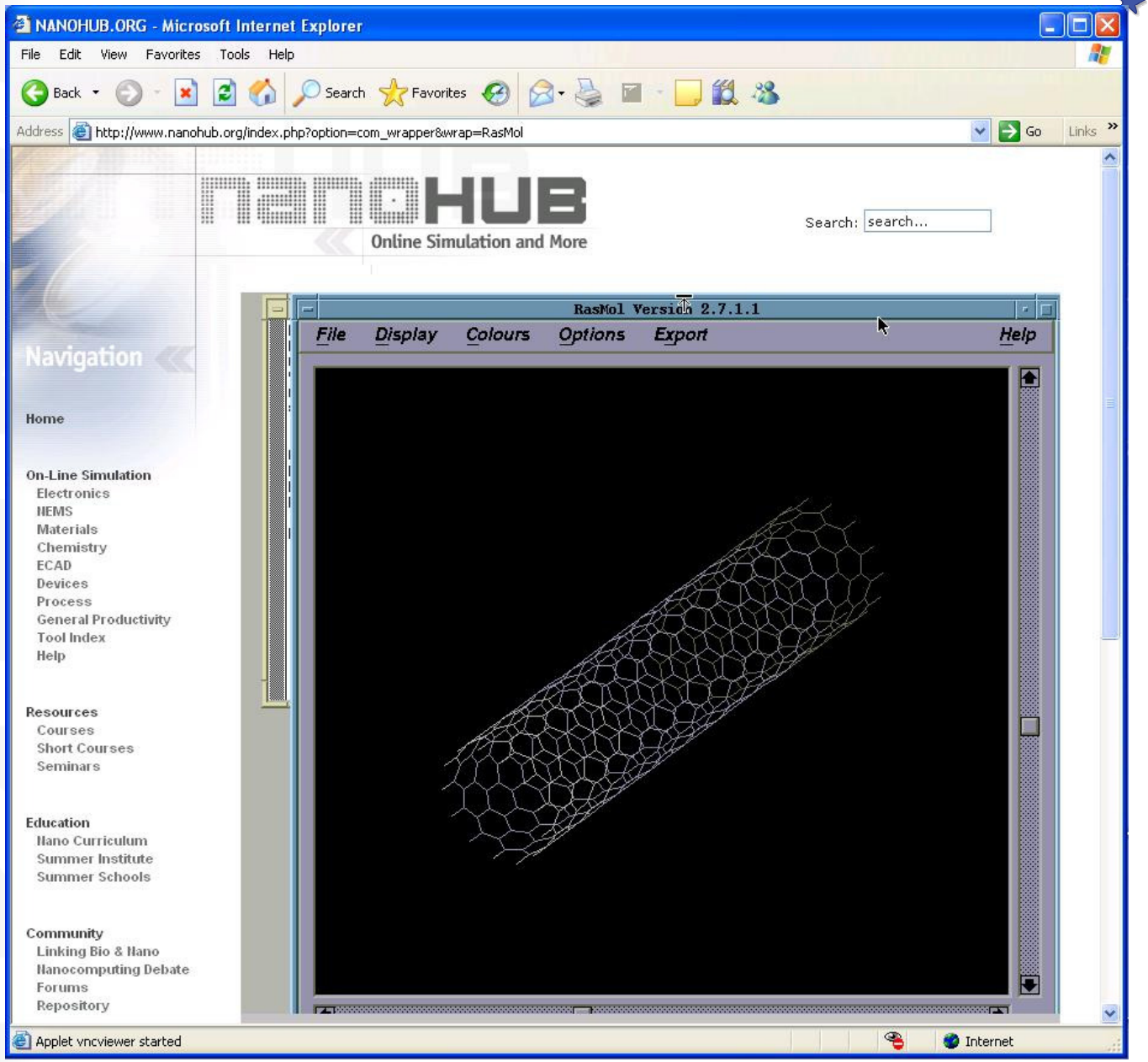
Distributed European Infrastructure for Supercomputing Applications

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



Nanohub - nanotechnology



Distributed European Infrastructure for Supercomputing Applications

RI-222919

X-Ray Crystallography

Distributed European Infrastructure for Supercomputing Applications

C.I.M.A. Common Instrument Middleware Architecture

Welcome to the Crystallography Portal

Username: Password: Login

Remember me on this computer

Login Help

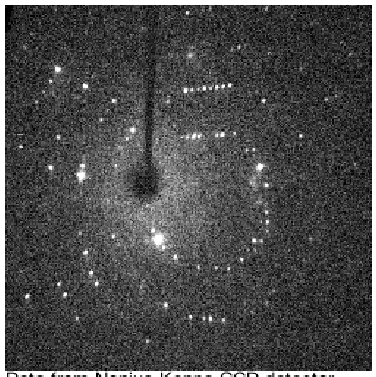
Home | Current Status | Data Repository | About

- IUB IUMSC
- IUB Myers Hall
- Purdue Crystallography Center
- CSAF Sydney, Australia
- Minnesota X-ray Lab
- ChemMatCARS - Univ. of Chicago at APS
- Other collaborators
- NCS Southampton, UK

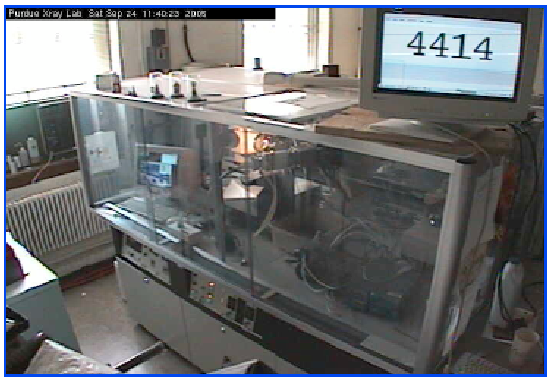
The Purdue Chemistry Crystallography Center

The Purdue Chemistry Crystallography Center

Disable your browser's cache to get the live stream!



Data from Nonius Kappa CCD detector
(Under development!)
Total Number of jpg: 10
Frame: s01f0010.jpg
All available jpg images
Browse the 20 latest jpg images



<□>

Streaming video from the lab showing the Nonius instrument

<□>

Streaming video from the crystal microscope on the Nonius diffractometer

Local date/time: 2005-09-24 11:36:54

These values are updated approx. every 60 sec. Times in UTC

LabJack U12

Instrument Enclosure Temp. & Humidity:	23.4 C Rel. Humid. 43.1 %	2005-09-24 16:35:59
Chill Water In:	16.4 C	2005-09-24 16:36:25
Chill Water Out:	19.3 C	2005-09-24 16:36:25
Generator Relay Voltage: All previous voltages	3.42 X-ray Generator is: OFF	2005-09-24 16:36:48

ServoGrid Portal

Distributed European Infrastructure for Supercomputing Applications

Username: Password: Login

Create New Account | Login Help

SERVOGrid FaultDB Search QuakeTables Portal Search

SERVOGrid

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.

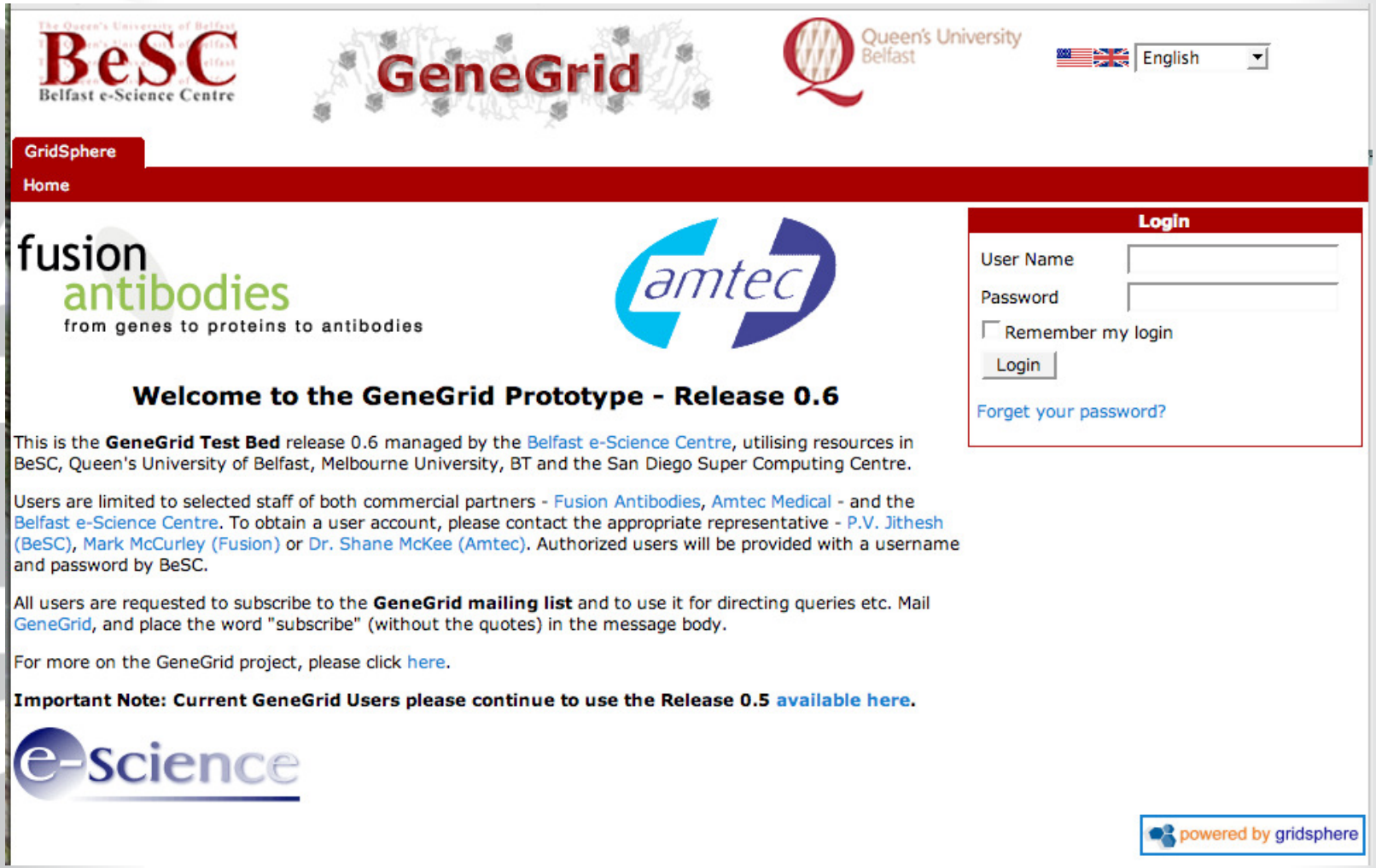
QuakeSim Web Portal | User Manual | Support | Report Bugs | QuakeSim Web Site

Participating Institutions:

IU CGL | NASA JPL | UC Davis | UC Irvine | USC

Belfast Gene Grid Portal

Distributed European Infrastructure for Supercomputing Applications



MyGrid - Bioinformatics

Distributed European Infrastructure for Supercomputing Applications

myGrid

Navigate

- Home
- About
- Downloads
- Components
  - Component Overview
  - Research Components
- Using myGrid
- Research Using myGrid
- Links
- Publications
- Contact

Log In

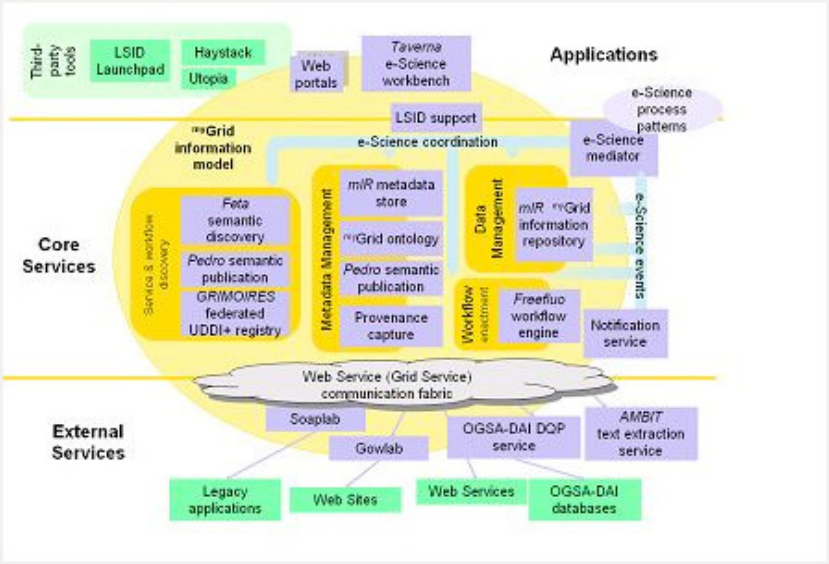
Username

Password

Log In

New Account Signup
Forgot Password

myGrid Architecture



myGrid components - overview

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, t e-science workbench that actively supports the scientific lifecycle. Each component or service contri system that allows the e-scientist to perform complex in-silico experiments across distributed bioinfo resources.

RI-222919

Distributed
European
Infrastructure for
Supercomputing
Applications

The DEISA Ecosystem for HPC Grand-Challenge Applications

Distributed European Infrastructure for Supercomputing Applications

DEISA Partners

Distributed European Infrastructure for Supercomputing Applications

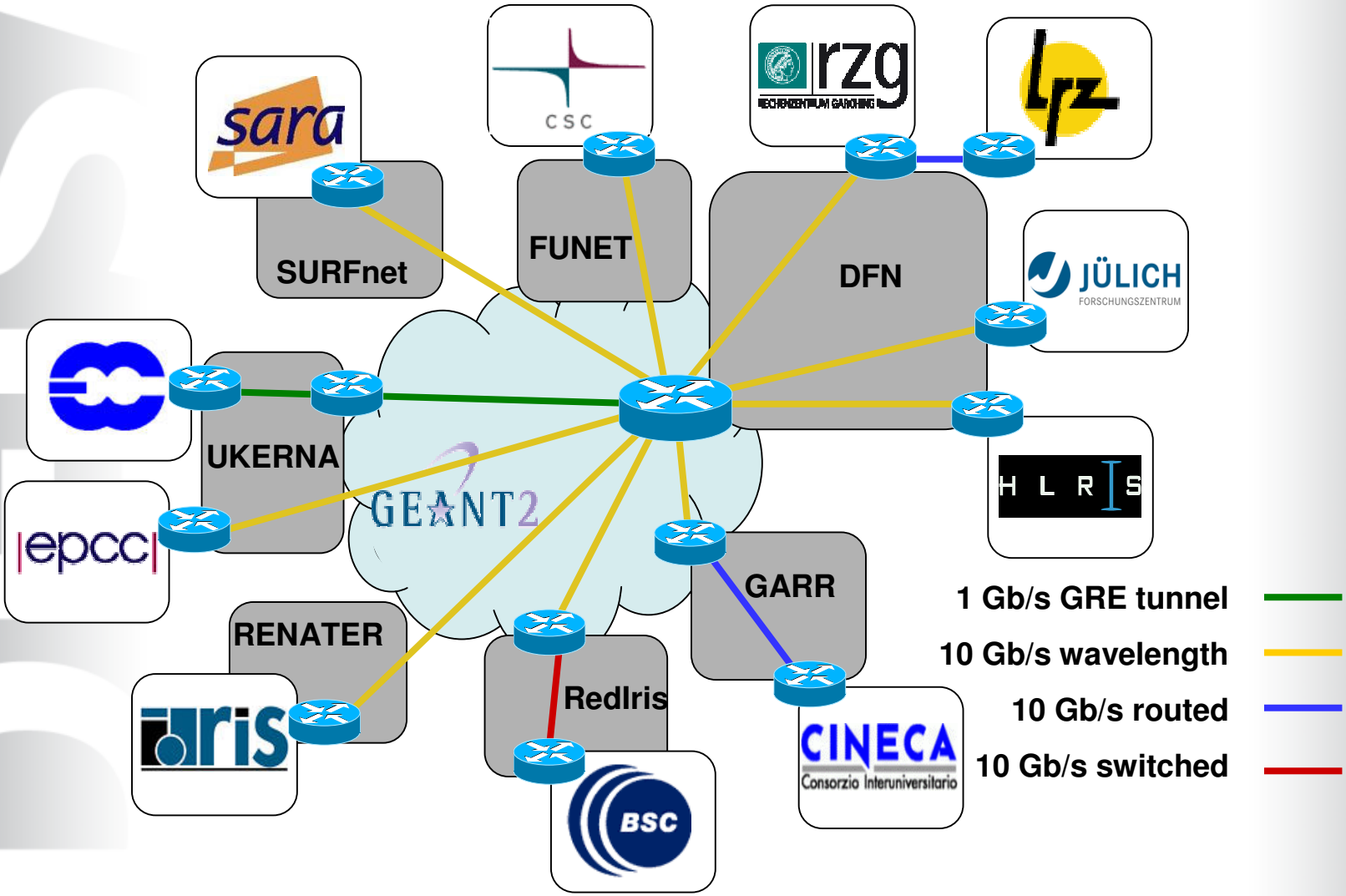
DEISA1: May 1st, 2004 – April 30th, 2008

DEISA2: May 1st, 2008 – April 30th, 2011

SEVENTH FRAMEWORK PROGRAMME

DEISA dedicated high speed network on GEANT2 and the NRENs

Distributed European Infrastructure for Supercomputing Applications



RI-222919

DEISA: Vision - Mission - Strategy

Distributed European Infrastructure for Supercomputing Applications

Vision:

Persistent European HPC ecosystem integrating Tier-1 (Tflop/s) centres and the new 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.

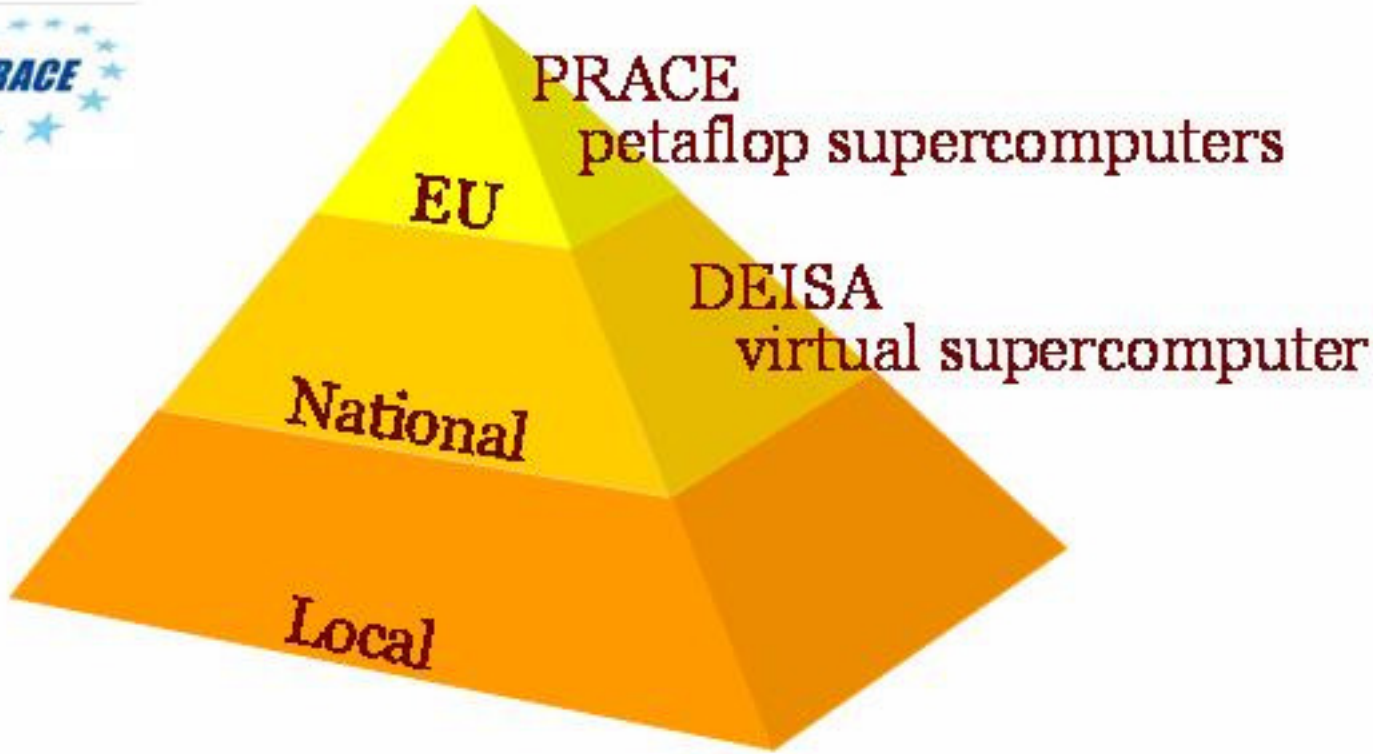
Built European Supercomputing Service on top of existing national services, based on the deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope.

Strategy:

Consolidate the existing DEISA1 HPC infrastructure and services.

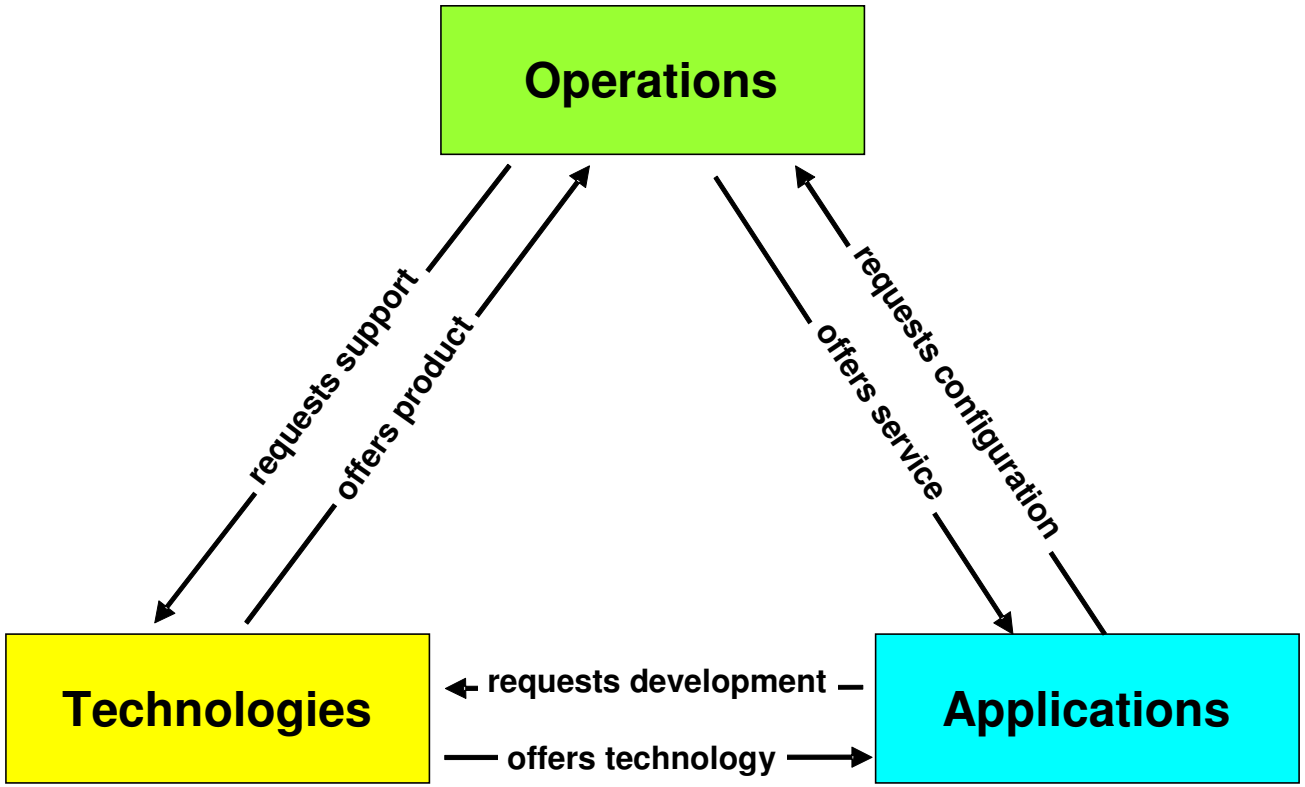
Deliver a turnkey operational solution for the future persistent European HPC ecosystem.

new "petaflop" supercomputers



Mario Campolargo
European Commission
OGF23, June 2008

Categories of DEISA services



RI-222919

DEISA Service Layers

Distributed European Infrastructure for Supercomputing Applications

Multiple ways to access	Workflow managemnt	Common production environmnt	*Presen-tation layer*
Single monitor system	Job rerouting	Co-reservation and co-allocation	*Job manag. layer and monitor.*
Data staging tools	Data transfer tools	WAN shared File system	*Data manag. layer*
Unified AAA	DEISA Sites	Network connectivity	*Network and AAA layers*

RI-222919

DEISA UNICORE Infrastructure

Distributed European Infrastructure for Supercomputing Applications

CINECA user

job

Gateway CINECA

Gateway BSC

Gateway CSC

Gateway ECMWF

Gateway FZJ

Gateway IDRIS

Gateway HLRS

Gateway HPCX

Gateway LRZ

Gateway RZG

Gateway SARA

job

LRZ user

NJS ECMWF IBM P5
IDB UUDB

NJS FZJ IBM
IDB UUDB

NJS IDRIS IBM P6
IDB UUDB

NJS HLRS NEC SX8
IDB UUDB

NJS HPCX Cray XT4
IDB UUDB

NJS CSC Cray XT4/5
IDB UUDB

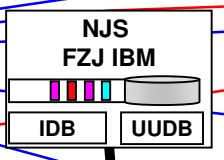
NJS LRZ SGI ALTIX
IDB UUDB

NJS CINECA IBM P5
IDB UUDB

NJS RZG IBM
IDB UUDB

NJS BSC IBM PPC
IDB UUDB

NJS SARA IBM
IDB UUDB



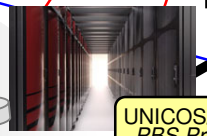
AIX LL-MC

AIX LL-MC

Super-UX NQS II

AIX LL

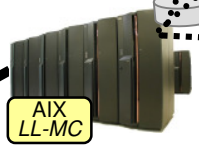
GridFTP



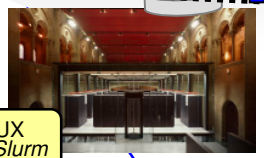
UNICOS/lc PBS Pro

UNICOS/lc PBS Pro

LINUX PBS Pro



AIX LL-MC



LINUX Maui/Slurm

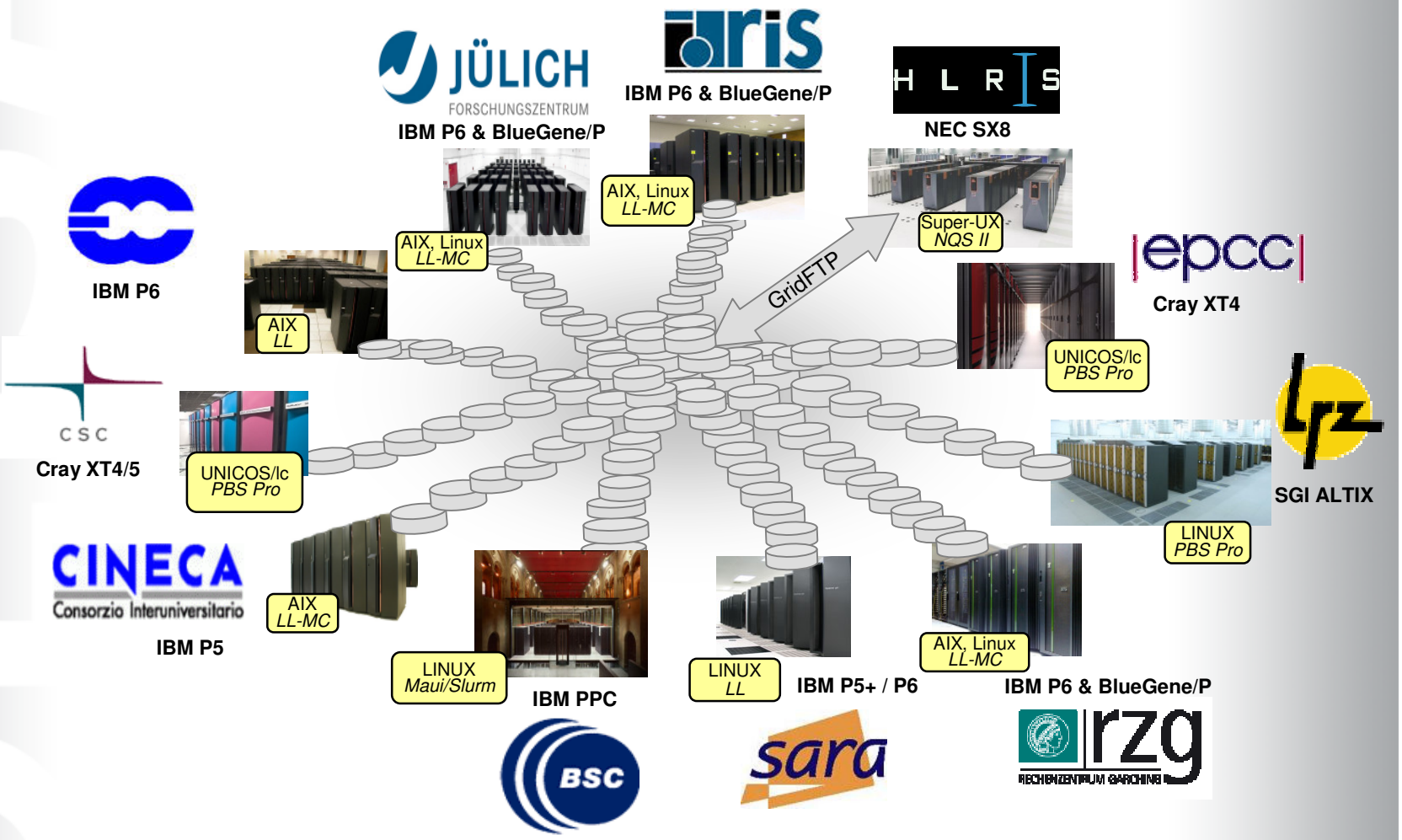
LINUX LL

AIX LL-MC

RI-222919

DEISA Global File System

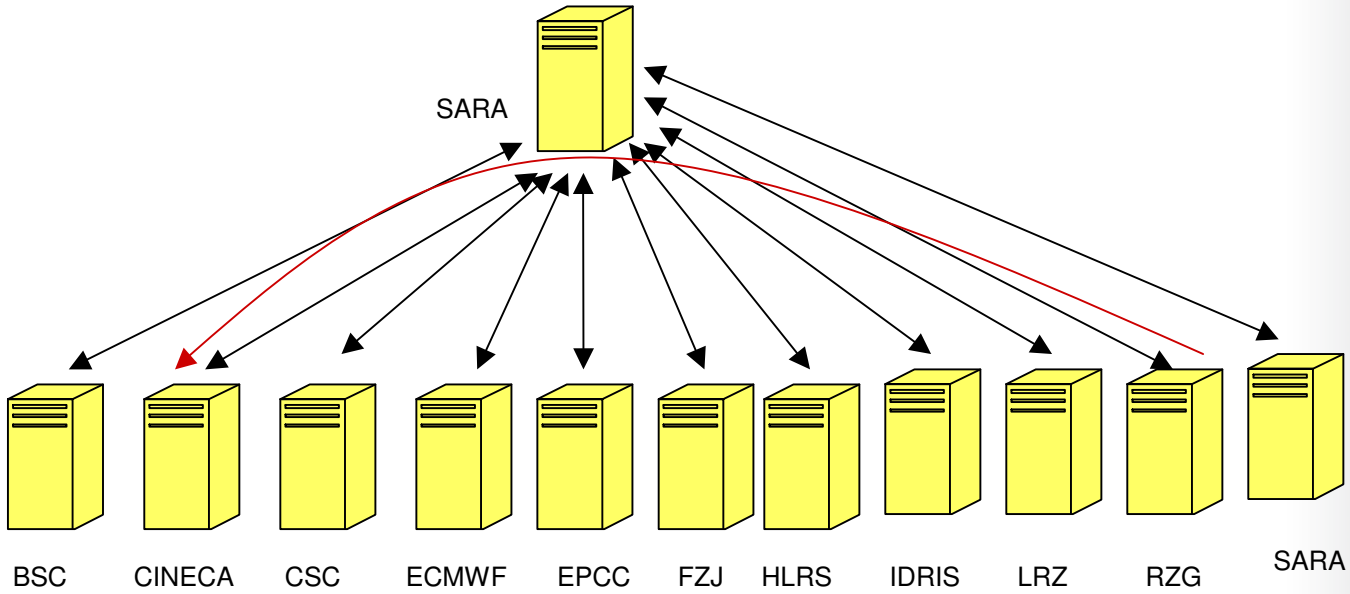
Distributed European Infrastructure for Supercomputing Applications



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

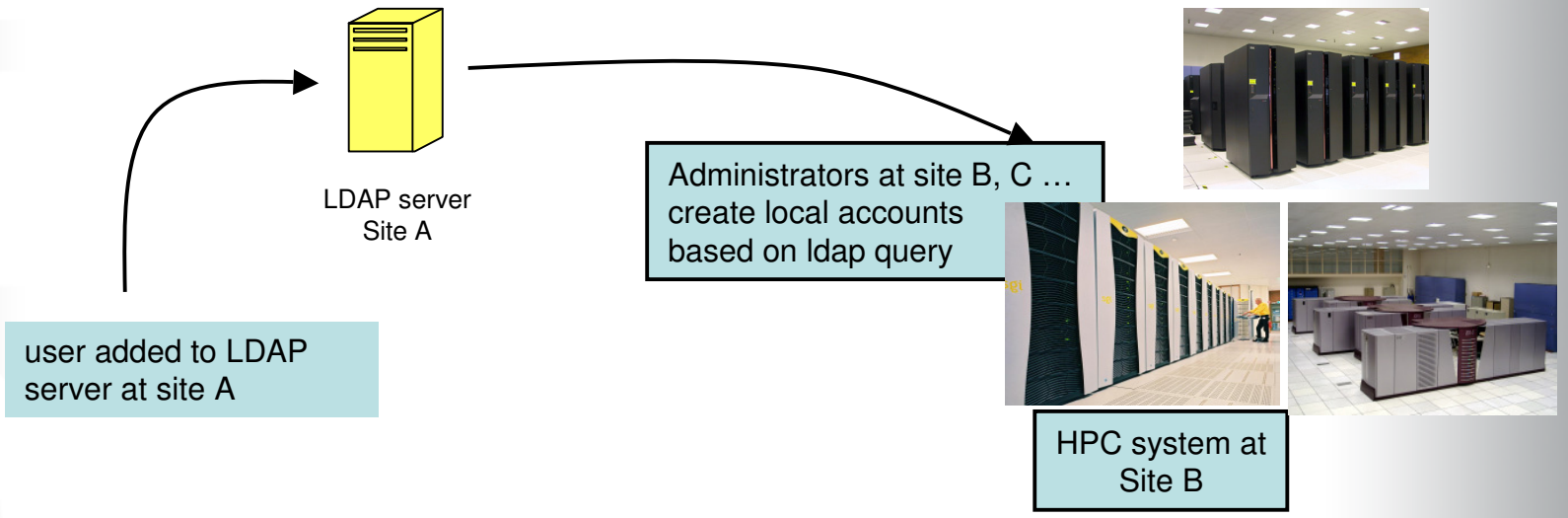
Management of users in DEISA

- 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



Common User Administration

- Each partner is responsible for the registration of users affiliated to the partner (home organization)
- Other partners update local user administration (LDAP, NIS, /etc/passwd) with data from other sites on a daily basis. Based on trust between partners!



Distributed
European
Infrastructure for
Supercomputing
Applications

DEISA

Next-Generation e-Infrastructures

A Peek at Intel's Digital City Vision

Courtesy Robert Fogel, Intel

Facets of the Digital City

Distributed European Infrastructure for Supercomputing Applications

Serving Citizens

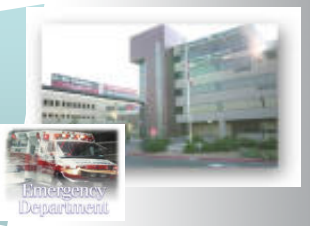
Digital Education

Economic Vitality
(Digital Office)

Digital Govt
GAPP Programs

Digital City

Digital Healthcare



Bridging the
Digital Divide

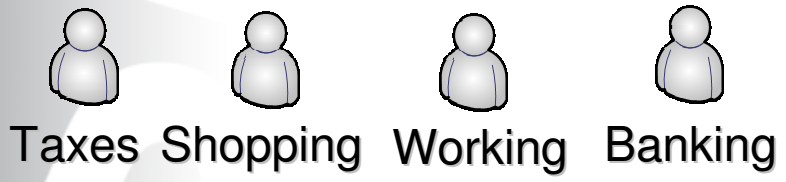
Digital Home

Safety & Security

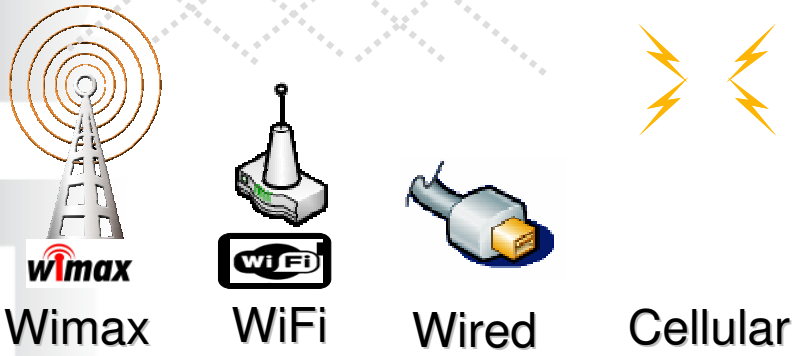
The Digital City is the Fabric that Connects the Community

Today's Digital Challenge

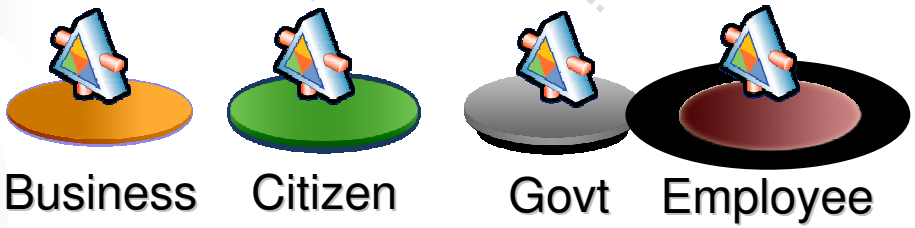
Distributed European Infrastructure for Supercomputing Applications



Multiple Identities



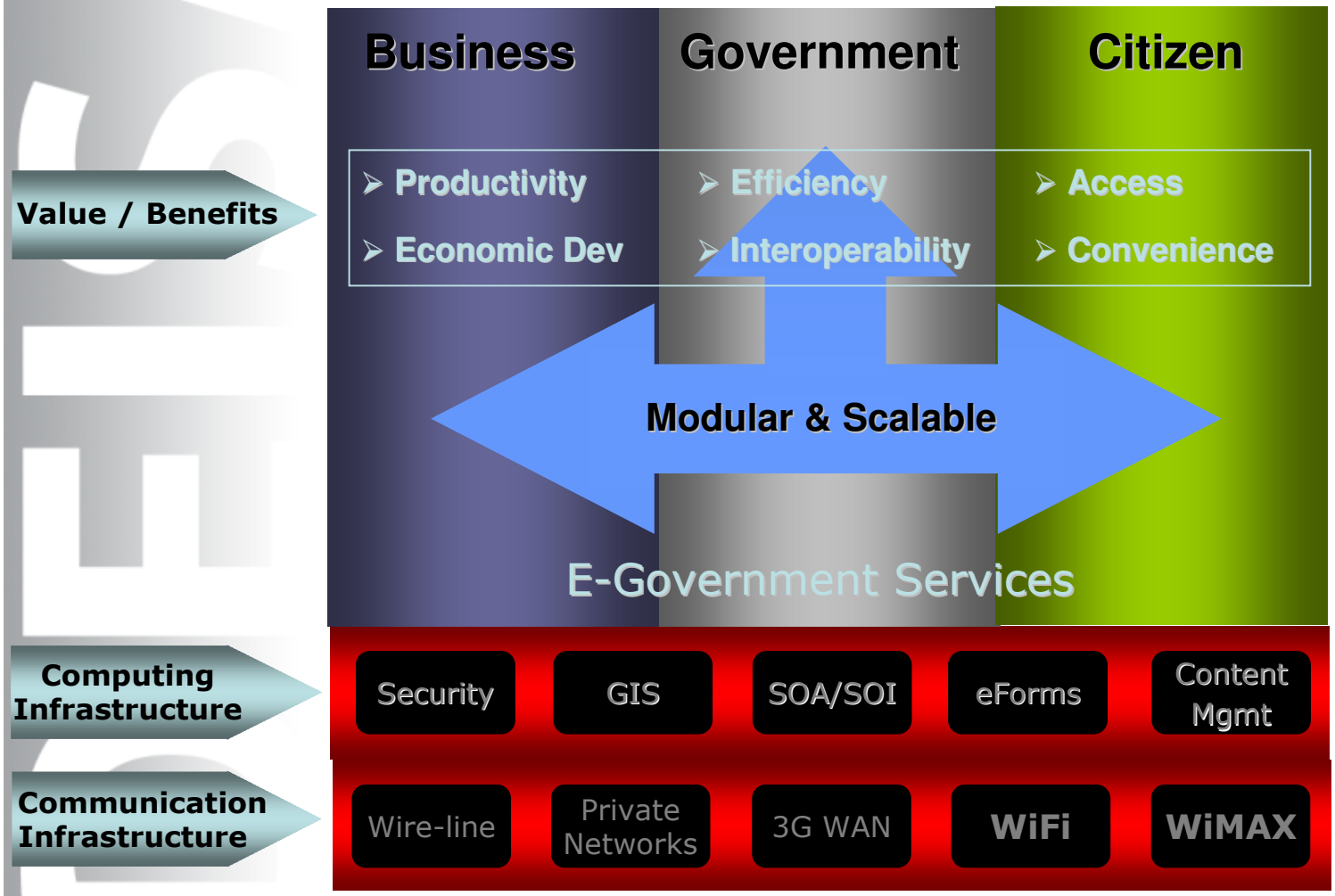
Incompatible Networks



Disconnected Agencies

Building the Digital City Today

Distributed European Infrastructure for Supercomputing Applications

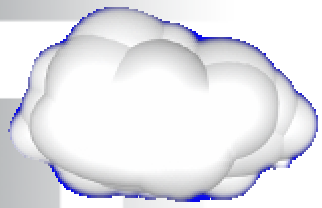


Tomorrow's Integrated Digital City

Distributed European Infrastructure for Supercomputing Applications

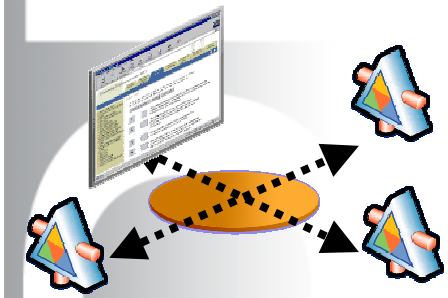
A single, portable identity based on strong security

One Identity



An intelligent infrastructure supporting seamless access

One Network



A gateway to integrated e-Services spanning multiple agencies

One Face

A Model for Sustainability (a checklist)

Distributed European Infrastructure for Supercomputing Applications

Reduce or eliminate the barriers in all the different areas such technology, culture, legal, economics and politics !

Especially, incorporate existing sustainability already achieved with individual components !

Therefore, the DEISA sustainability model is based on ensuring sustainability of every *individual* component:

- Technology and infrastructure
- Operations and services
- Expertise
- Communities
- Collaborations
- Eco-political landscape

The DEISA Model for Sustainability

Distributed European Infrastructure for Supercomputing Applications

- Technology and Infrastructure:
  - DEISA infrastructure is built on existing, proven, sustainable technology components,
  - GEANT2, NRENs, Supercomputers, HPC services, global sw environment
  - deliver and operate a European supercomputing infrastructure and related services
- Operations and Services
  - benefit from the many-years operations of the individual European supercomputers centres
  - orchestrated by the partners after the end of the funded project
  - activities relevant for applications enabling, operation, and technologies have been developed
- Expertise
  - tight collaboration of the expert groups in the different HPC centres
  - provided in the future to the wider European HPC communities.

The DEISA Model for Sustainability

Distributed European Infrastructure for Supercomputing Applications

- Communities
  - annual DEISA Extreme Computing Initiative (DECI)
  - supporting single projects, Virtual European Communities, and international science communities across existing political boundaries
- Collaboration
  - Distributed Common Production Environment (DCPE)
  - Collaboration with new European and other international initiatives.
  - contacts to research infrastructure projects established by the ESFRI, and the European HPC and Grid projects such as PRACE and EGEE
  - European & international HPC centres; initiatives in Australia, China, Japan, Russia, US, and leading HPC projects worldwide
  - Participate in evaluation and implementation of interoperation standards
- Eco-Political Landscape
  - ESFRI European Strategy Forum on Research Infrastructures
  - PRACE: preparing installation of a limited number of leadership-class Tier-0 supercomputers in Europe.

Distributed European Infrastructure for Supercomputing Applications

Thank You!

GRACIAS POR SU ATENCIÓN

Gentzsch@rzg.mpg.de