



InfoSys 2009  
Valencia, April 20 – 25, 2009



# e-Infrastructures for Compute and Data Intensive Applications

## Solving the Grand Challenges in Big Science

Wolfgang Gentsch  
The DEISA Project & Board of Directors of OGF  
gentsch at rzg.mpg.de



# InfoSys 2009



S  
E  
D

**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.



## Components of an e-Infrastructure:

### HPC Centers, Grids and Clouds

DELS



# 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



- 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

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

- 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





DEI

# 10 Examples of e-Infrastructures

Courtesy: Dennis Gannon





# NEESGrid



Realtime access to earthquake  
Shake table experiments at remote sites.

CHEF (dev-local): Worksite - Microsoft Internet Explorer  
Address: http://neespop.ce.unr.edu:9271/chef/portal/group/NEESgridUNR/page/default.psm/js\_pane/P-f16ab04bfe-10006  
NEESgrid  
My Workspace NEESgrid UNR NEESgrid Support NEESgrid All  
Video Cameras  
Home  
Schedule  
Announcements  
Resources  
Discussion  
Telepresence Server  
Video Cameras  
ENotebook  
Data Browser/Viewer  
NEESgrid Repository  
Chat  
Browse Testbed  
TeleRobotic Video Camera 1  
UNRCamera1 Thu Nov 14 17:43:19 2002  
Select Camera: 1 2 3 4 5

NEESgrid  
My Workspace NEESgrid UNR NEESgrid Support NEESgrid All  
NEES Data Browser  
NEES Repository @ UNR  
Symbol key:  
The folder is open (click to close).  
The folder is closed (click to open).  
Download this file using GridFTP.  
Upload a file to this folder using GridFTP.  
Configure data streaming and recording for this event.  
Make this event viewable with the NEES data viewer.  
Follow a link between objects.  
New Object Delete Object Move Object Copy Object Edit Object  
[Server Root]  
UNR Demo  
Bridge Shake Table Experiment  
Data Acquisition  
White Noise System Identification  
White Noise System Identification Channel Configuration  
sensor data  
white noise  
sensor data  
white noise  
sensor data  
sensor data  
1940 Imperial Valley-EI Centro 100%  
1940 Imperial Valley-EI Centro 100% Channel Configuration  
People  
Folder "Data Acquisition"  
name (string)  
Data\_Acquisition  
lastAccessedTimestamp (timestamp)  
2002-11-12 13:15:06.055  
lastModifiedTimestamp (timestamp)  
2002-11-12 13:15:06.055  
originalVersionID (object)  
Data Acquisition  
versionNumber (integer)  
0  
creationTimestamp (timestamp)  
2002-11-12 13:15:06.055  
lockTimestamp (timestamp)  
1969-12-31 18:00:00.0

NEESgrid  
My Workspace NEESgrid UNR NEESgrid Support NEESgrid All  
Data Viewer  
Event: "core: ex2 sine1-4"  
Home  
Schedule  
Announcements  
Resources  
Discussion  
Telepresence Server  
Video Cameras  
ENotebook  
Data Browser/Viewer  
NEESgrid Repository  
sine1: 0.062791  
sine2: 0.24869  
sine3: 0.187381  
sine4: 0.125333  
sine1  
sine2  
sine1  
sine2



# BIRN – Biomedical Information



The screenshot shows the BIRN Portal interface. At the top left is the BIRN logo (Biomedical Informatics Research Network). To the right is a login section with fields for 'Username:' and 'Password:' and a 'login' button. Below the logo is a navigation bar with links: 'Portal Home', 'BIRN Website', 'Account Request', 'Style', and 'Help'.

The main content area is divided into two columns:

- Left Column:**
  - Login Information:** A section titled 'BIRN Portal Login' with the instruction 'Enter your username/password'. It contains 'Username:' and 'Password:' input fields, a 'Login' button, and a list of links: 'Request a BIRN account (must be a BIRN participant)' and 'Email BIRN Portal admins'.
  - Portal Requirements:** A section stating: '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. Older versions of Safari will experience'.
- Right Column:**
  - Welcome to the BIRN Portal:** A text block stating: '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.'
  - BIRN Portal Biomedical Informatics Research Network:** A large graphic featuring the text 'BIRN Portal' and 'Biomedical Informatics Research Network' over a background of a human brain and a DNA helix.



# Geological Information Grid Portal



**GEONgrid Portal**

Welcome Dennis Gannon: [Logout](#)

PortalHome GEONsearch myGEON GEONscience System UserProfile MapIntegration

GEON Search GEON Ontology GEON Resource Registration

**GEON Search**

**1 Metadata Related:**  
Choose resource type:  
<All Resource Types>  
Choose subjects:  
<All Subjects>  
Optional keywords:

**2 Spatial Coverage:**  
Type a place name:  
 [GO](#)  
or select an area on the map:

**Select a Subject to Show Resources**

Biological oceanography	Chemical oceanography	Cryology
Ecology	Education	Environmental science
Forestry	Geochemistry	Geologic time
Geology	Geophysics	Human geography
Hydrology	Mineralogy or petrology	Natural hazards
Other	Paleontology	Physical geography
Physical oceanography	Soil science	Structural geology
Technology		

(These subjects will be reorganized soon by something similar to the classification from [GeoRef](#).)

**Resources in Geology** 1-5 of 17 files ▶

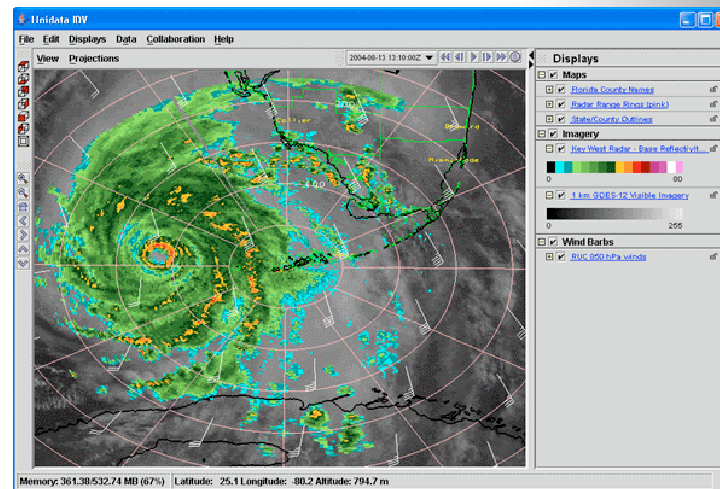
**Title:** Arizona Geology Map  
**Format:** shapefile  
**Dataset Id:** GEON-25dfb3db-e710-11d8-b226-ab22ed7681c0  
**Spatial Coverage:** North: 37 East: -109.04 South: 31.33 West: -114.82  
**Temporal Coverage:** any  
**Description:** This is a geology map of Arizona in USA.  
**Semantic Annotations:** [see details](#)

# Mesoscale Meteorology



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

- Data exploration and Grid workflow



Welcome to the **LEAD PORTAL** Linked Environments for Atmospheric Discovery  
Sponsored by the National Science Foundation

Portal Home | Geo GUI | Education and Outreach | Weather | Links | About LEAD | Help

Home

To view a local radar, select area of interest and click on the image below.

**RADAR REFLECTIVITY FROM RADAR CODED MESSAGES**  
NATIONAL WEATHER SERVICE  
AUTOMATED EDITING APPLIED  
SEP 24, 2005 21:49 UTC

MSG  
55 DBZ  
50 DBZ  
45 DBZ  
40 DBZ  
30 DBZ  
15 DBZ

Data provided by NOAA's National Weather Service

User Name: \_\_\_\_\_  
Password: \_\_\_\_\_  
 Remember my login

[Create new account](#)  
[Forgot your password?](#)

**LEAD Grid Testbed Status**

Testbed	Grid Auth	GRAM	Grid
IU [chinkapin]	✓	✓	✓
NCSA [copper]	✓	✓	✓
OU [aquaman]	✓	✓	✓
UAH [frozone]	✓	✓	✓
UNC [dante0]	✗	✗	✗
Unidata [lead1]	✓	✓	✓

Last updated: Sat Sep 24 17:00:00 2005 Indiana 1

LEAD Home | FAQ | Privacy | Terms of use | Contact us

Workflow Composer

Workflow MyLead Component Monitor

Add Node Remove Node Connect/Disconnect

Component List

- System Components
- http://whitney.extreme.indiana.edu
- http://www.extreme.indiana.edu
- Adaptor
- Multiplier
- Decoder
- Decoder
- Arps-trm
- Arps-sfc
- Ext2arps-ibc

Composer

Component Information

**Service: decoder**

**Description:**  
A service for decoding raw eta data to netcdf format

**Operation: Run**

Port Information | Notification |

Selected Output Port

Selected Input Port

**Component: Output\_URL**  
**Port: Parameter**  
**Type: Any**  
**Description:** This port can be connected to any type.

InfoSys, April 20-25, 2009

# 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



InfoSys



# X-Ray Crystallography



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

Welcome to the  
Crystallography Portal

Username:   
 Password:    
 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](#)

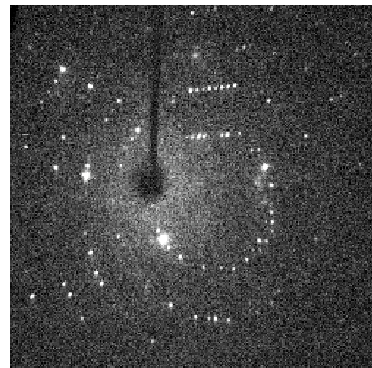
[Other  
collaborators](#)

[NCS  
Southampton,  
UK](#)

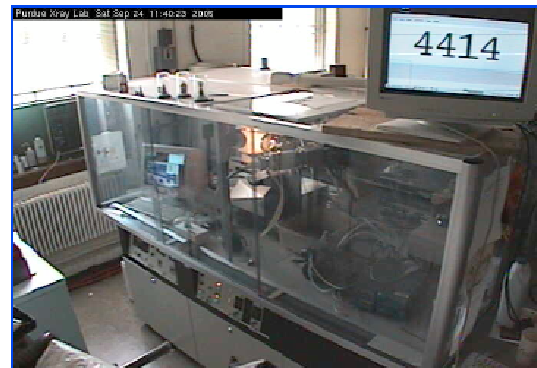
The Purdue Chemistry Crystallography Center

## 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](#)



<O>

[Streaming video from the lab showing the Nonius instrument](#)



<O>

[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.

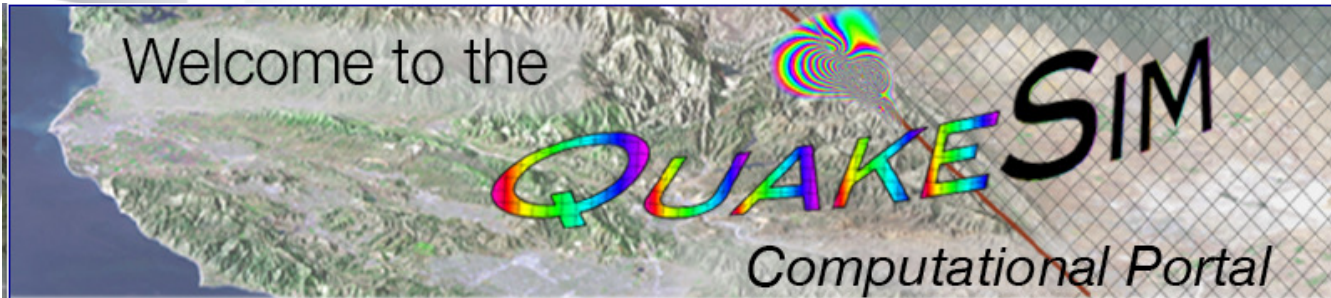
Times in UTC

### 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: <a href="#">All previous voltages</a>	3.42 X-ray Generator is: <b>OFF</b>	2005-09-24 16:36:48



# ServoGrid Portal



Username:

Password:

[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



The Queen's University of Belfast  
**BeSC**  
Belfast e-Science Centre

**GeneGrid**

Queen's University Belfast

English

GridSphere  
Home

**fusion antibodies**  
from genes to proteins to antibodies

**amtec**

**Login**

User Name

Password

Remember my login

Login

[Forget your password?](#)

**Welcome to the GeneGrid Prototype - Release 0.6**

This is the **GeneGrid Test Bed** release 0.6 managed by the [Belfast e-Science Centre](#), utilising resources in BeSC, Queen's University of Belfast, Melbourne University, BT and the San Diego Super Computing Centre.

Users are limited to selected staff of both commercial partners - [Fusion Antibodies](#), [Amtec Medical](#) - and the [Belfast e-Science Centre](#). To obtain a user account, please contact the appropriate representative - [P.V. Jithesh](#) (BeSC), [Mark McCurley](#) (Fusion) or [Dr. Shane McKee](#) (Amtec). Authorized users will be provided with a username and password by BeSC.

All users are requested to subscribe to the **GeneGrid mailing list** and to use it for directing queries etc. Mail [GeneGrid](#), and place the word "subscribe" (without the quotes) in the message body.

For more on the GeneGrid project, please click [here](#).

**Important Note: Current GeneGrid Users please continue to use the Release 0.5 available [here](#).**

**e-Science**

powered by gridsphere



# MyGrid - Bioinformatics

myGrid



## Navigate

- [Home](#)
- [About](#)
- [Downloads](#)
- [Components](#)
  - [Component Overview](#)
  - [Research Components](#)
- [Using myGrid](#)
- [Research Using myGrid](#)
- [Links](#)
- [Publications](#)
- [Contact](#)

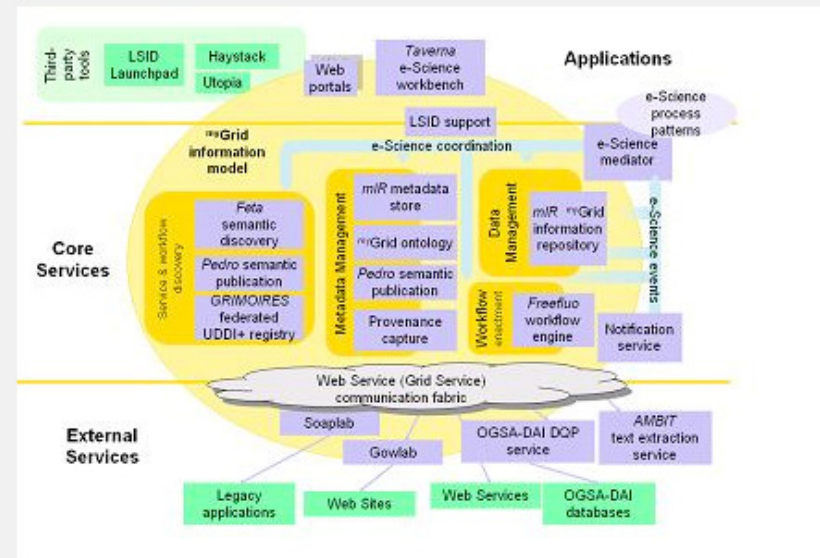
## Log In

Username

Password

[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, to support 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.



# The DEISA Ecosystem for HPC Grand-Challenge Applications

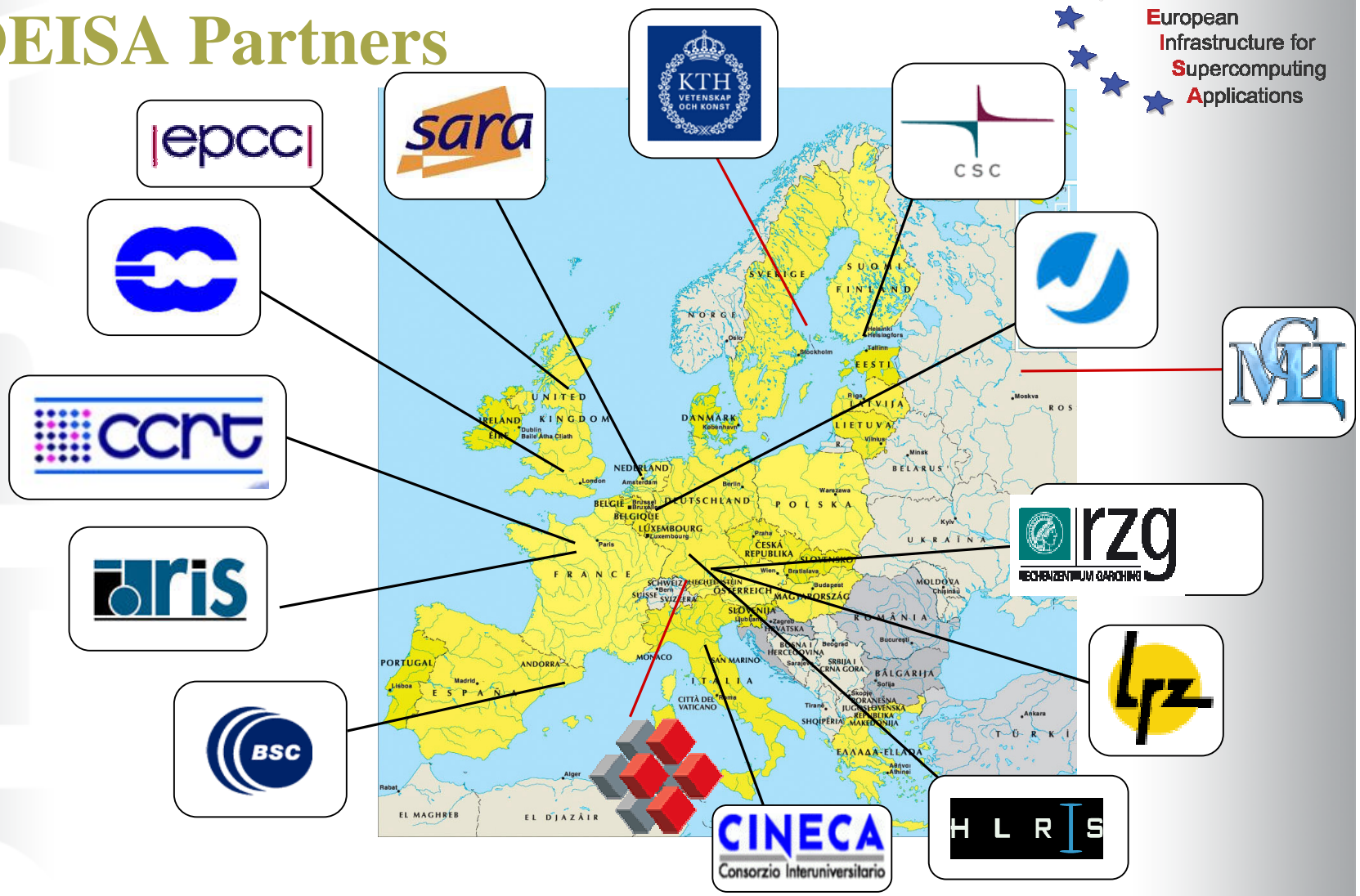
## Distributed European Infrastructure for Supercomputing Applications

DEISA



# DEISA Partners

**Distributed European Infrastructure for Supercomputing Applications**

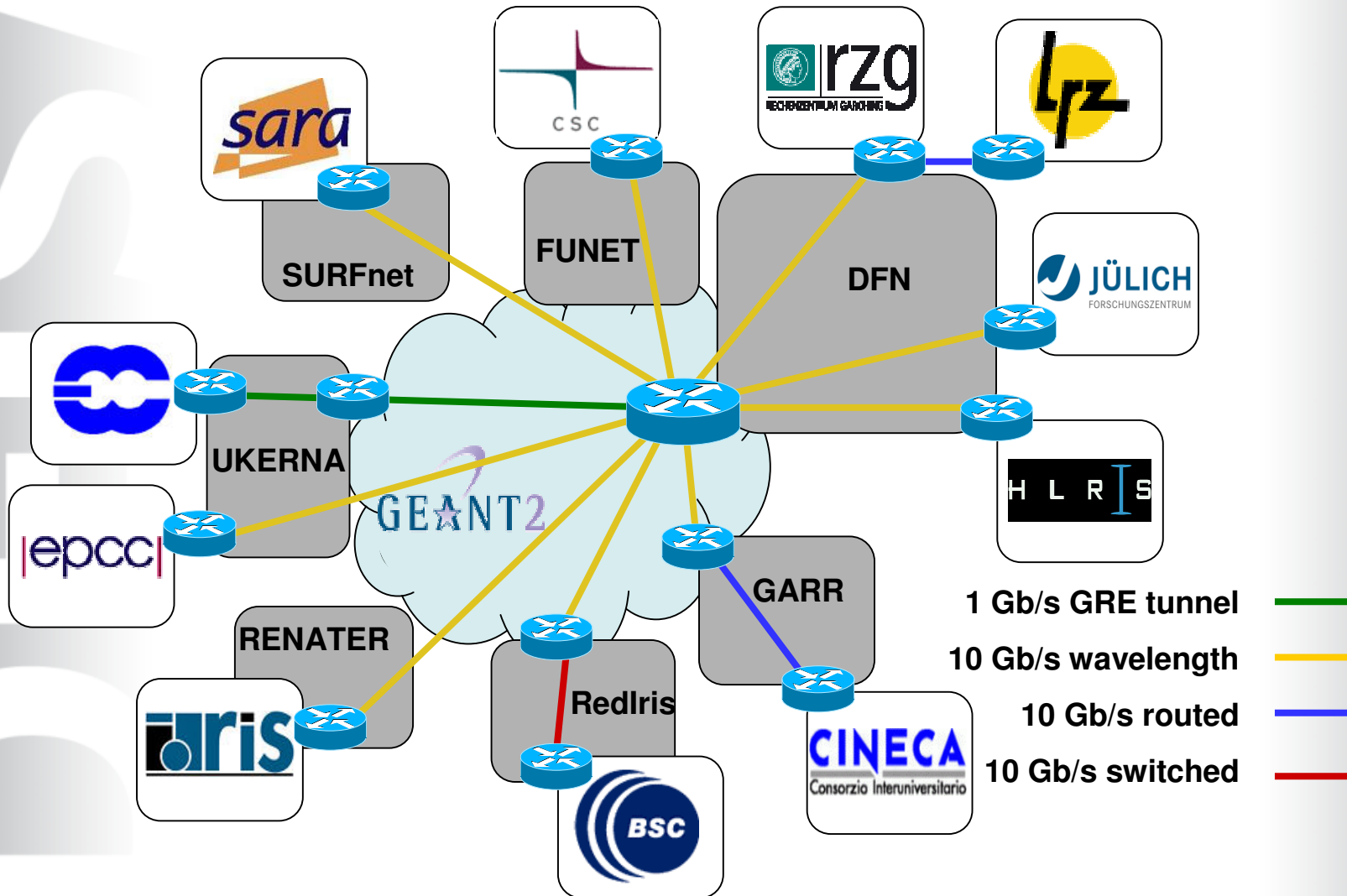


DEISA1: May 1st, 2004 – April 30th, 2008

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



# DEISA dedicated high speed network on GEANT2 and the NRENs



# DEISA: Vision - Mission - Strategy



## Vision:

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

## Mission:

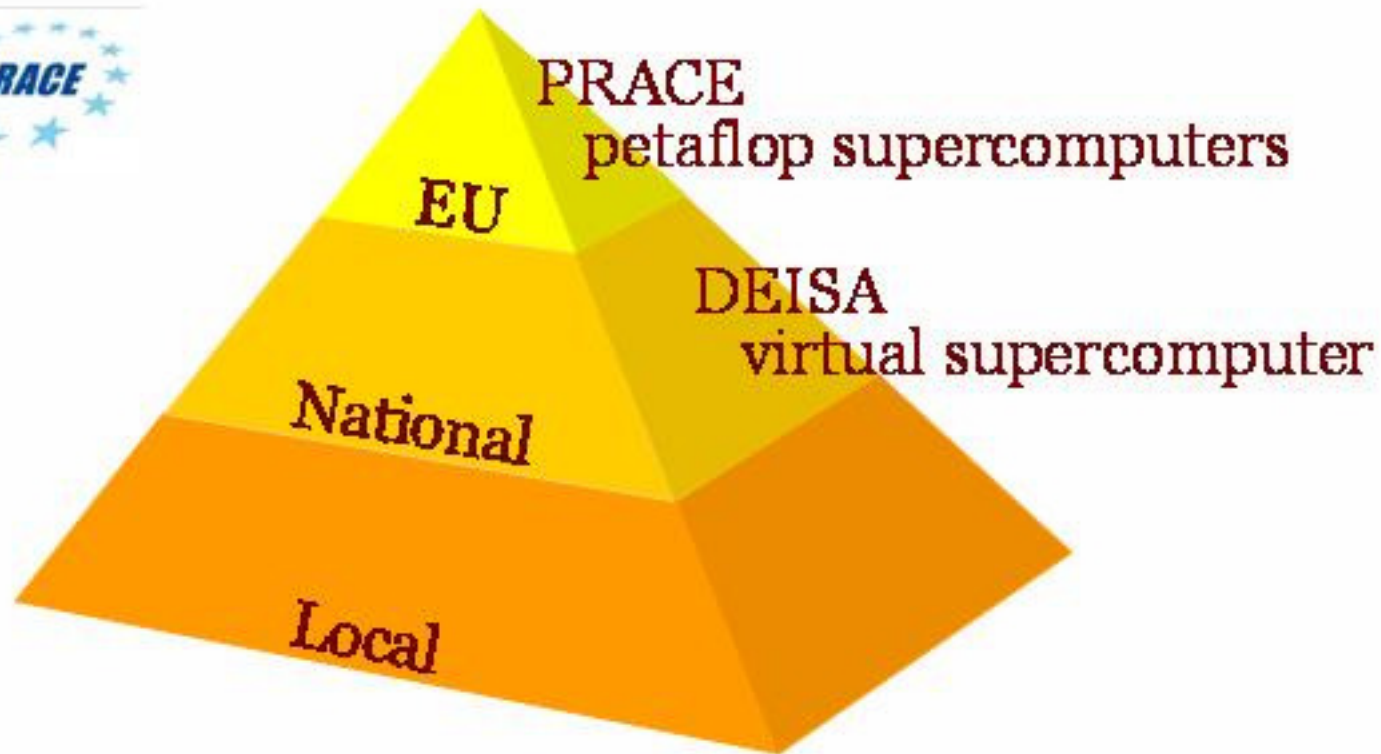
Enhance Europe's capability in computing and science by integrating most powerful supercomputers into a European HPC e-infrastructure.

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.

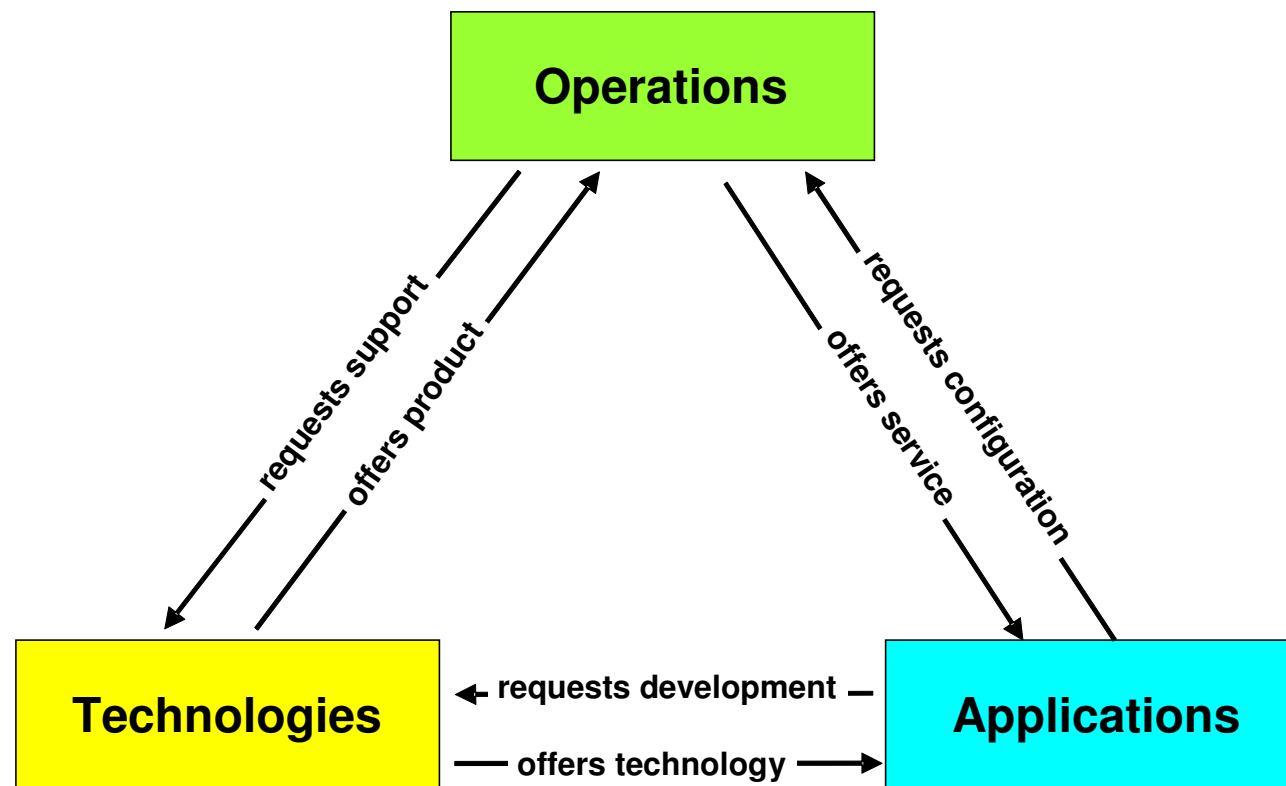
# new "petaflop" supercomputers



*Mario Campolargo  
European Commission  
OGF23, June 2008*

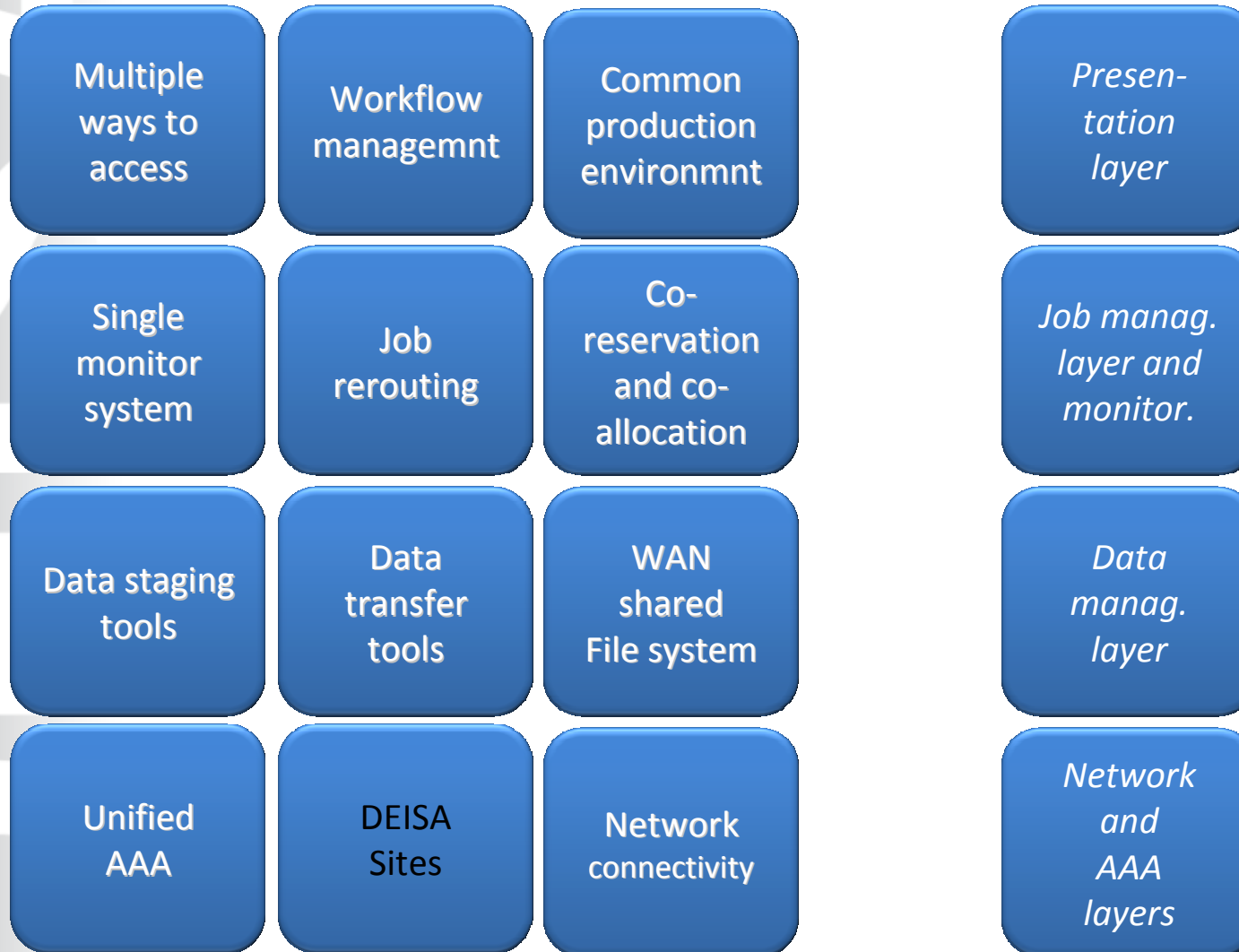


# Categories of DEISA services



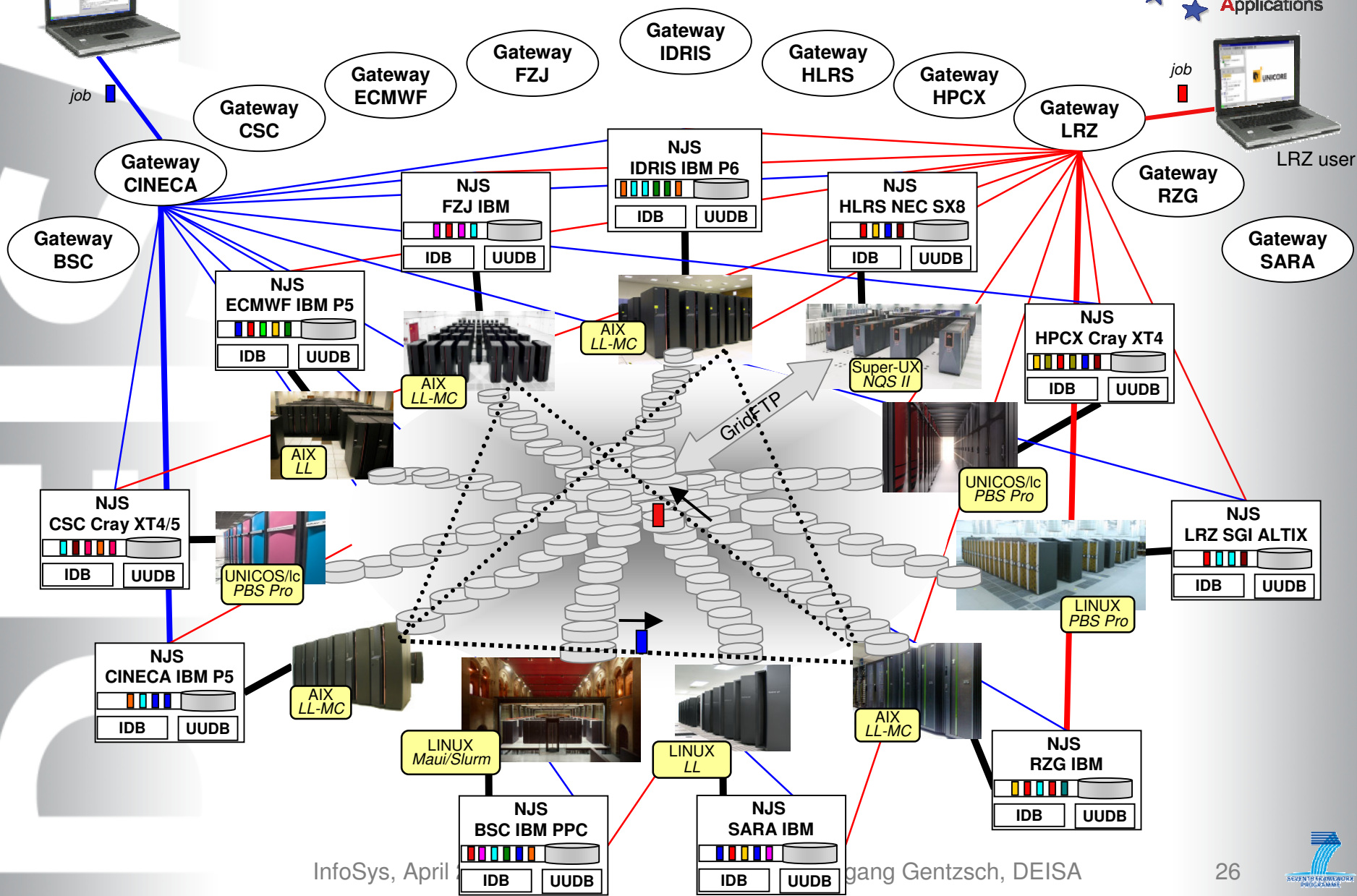


# DEISA Service Layers

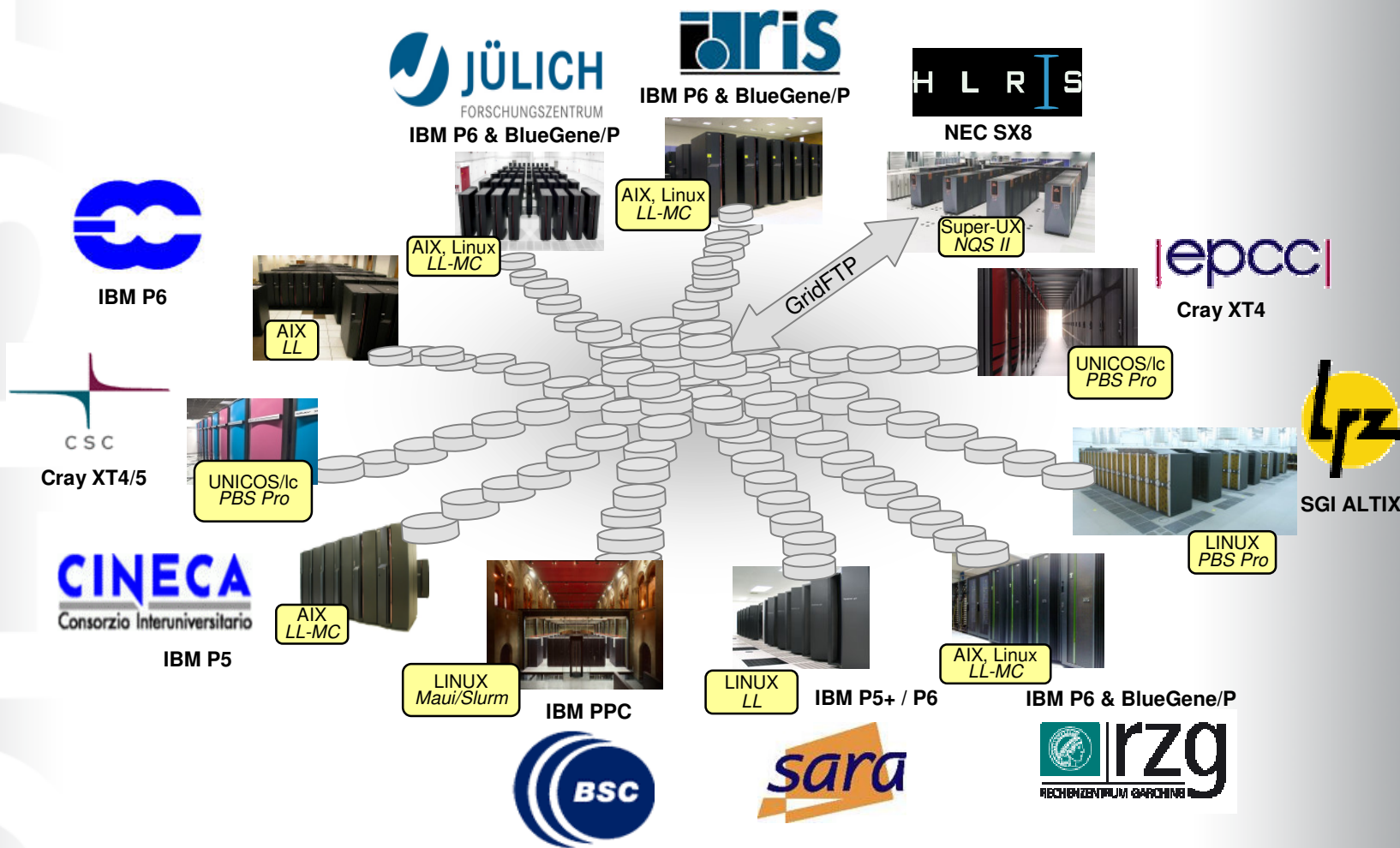


# DEISA UNICORE Infrastructure

**Distributed European Infrastructure for Supercomputing Applications**



# DEISA Global File System

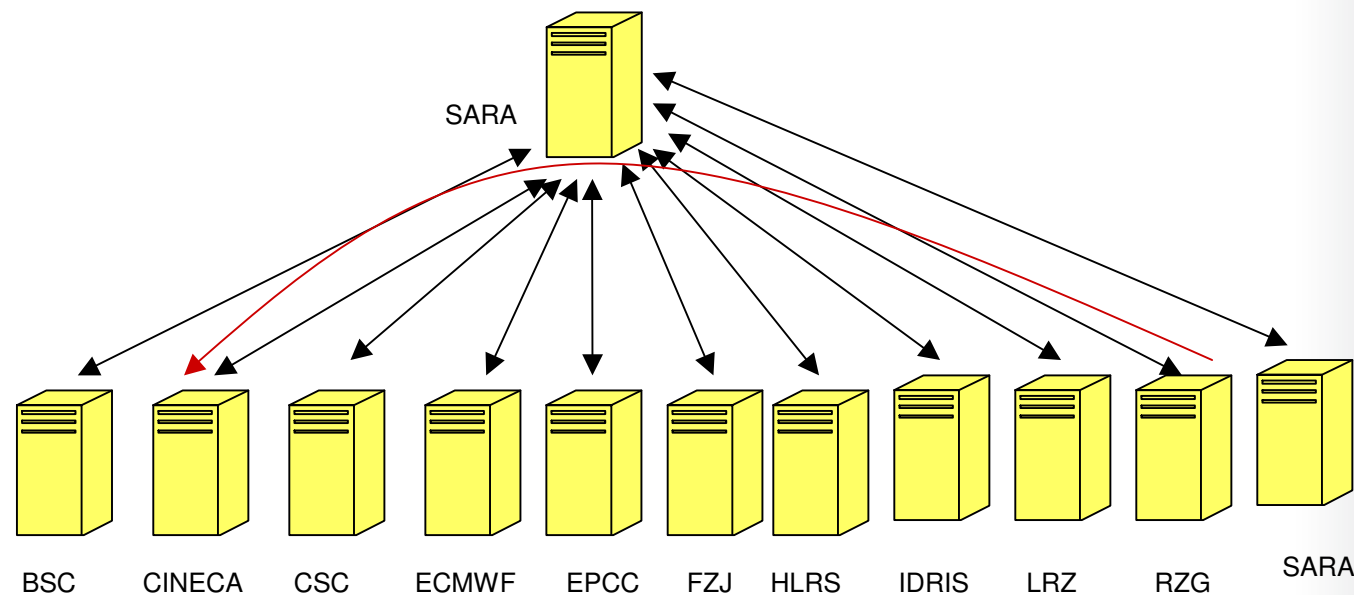


Global transparent file system based on the Multi-Cluster General Parallel File System (MC-GPFS of IBM)



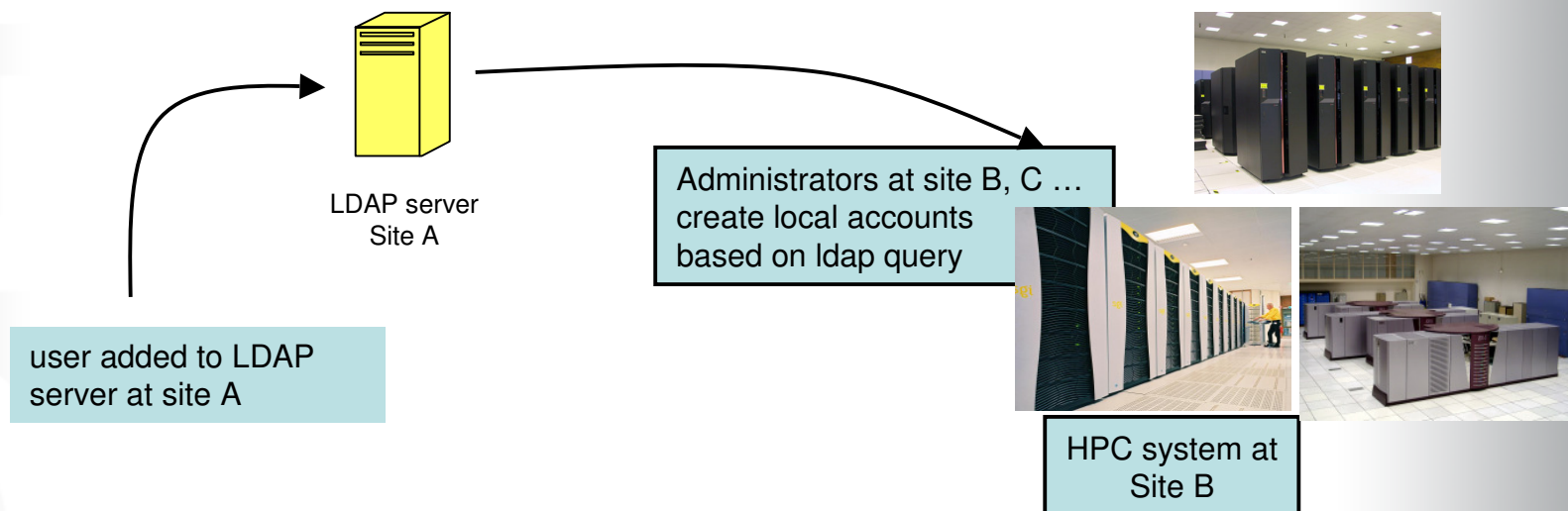
# Management of users in DEISA

- 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!





DEIS

# 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**



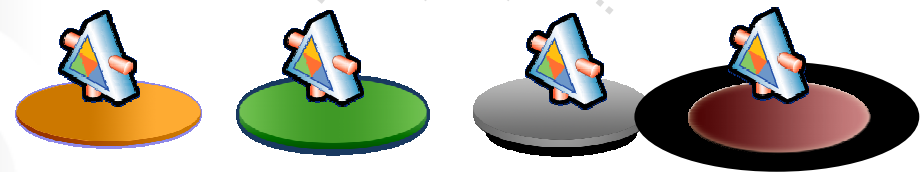
# Today's Digital Challenge

  
 Taxes Shopping Working Banking

Multiple Identities

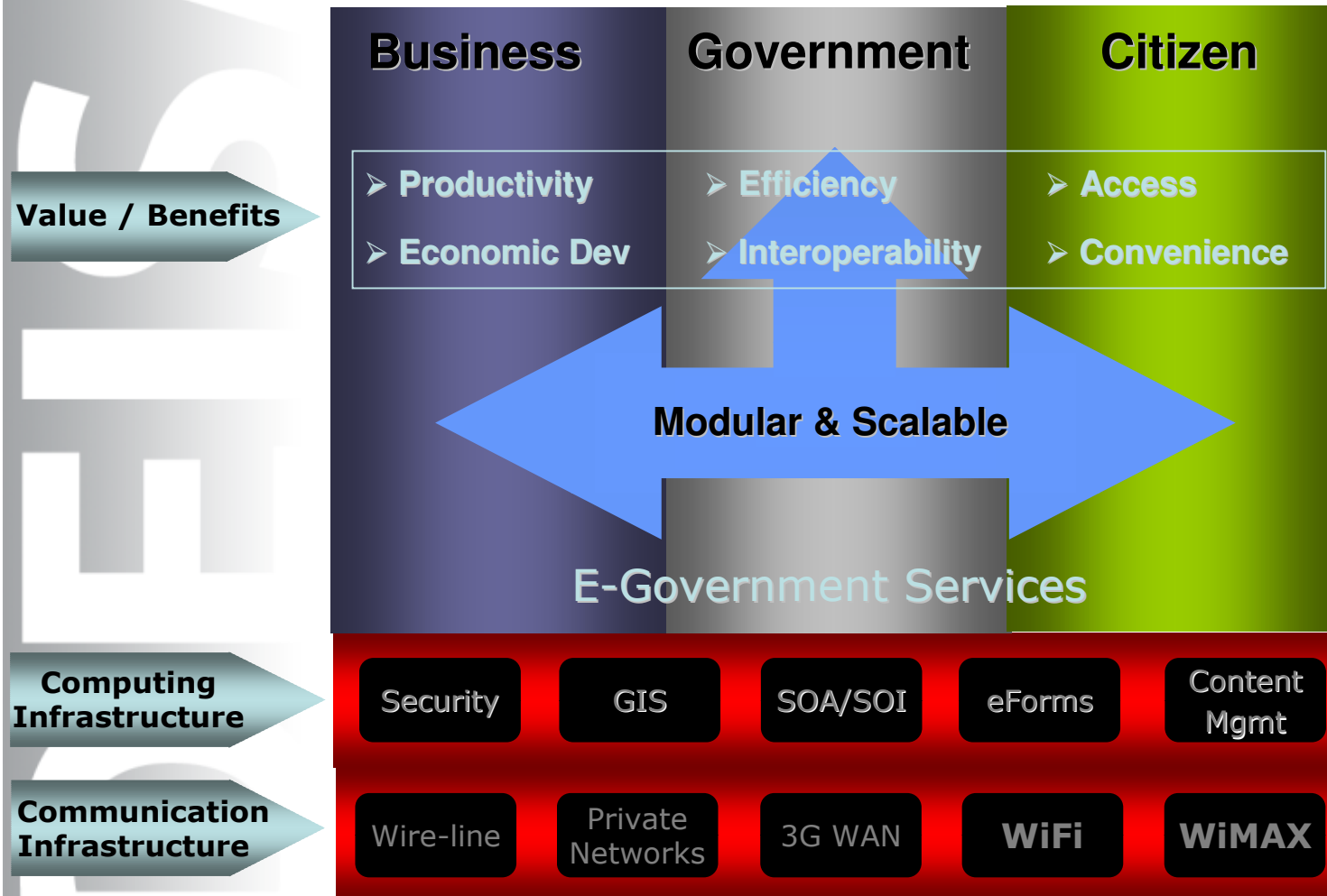
  
 Wimax WiFi Wired Cellular

Incompatible Networks

  
 Business Citizen Govt Employee

Disconnected Agencies

# Building the Digital City Today



# Tomorrow's Integrated Digital City



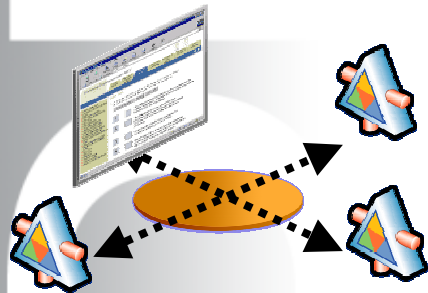
A single, portable identity based on strong security

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 )

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

- **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



- **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.





**Thank You!**  
***GRACIAS POR SU ATENCIÓN***

**Gentzsch@rzg.mpg.de**