DEISA and the Supercomputing View

Status of the European HPC Infrastructure

Vision of the European DEISA/eDEISA projects in FP6

To enhance Europe's capability computing and science by the integration of Europe's most powerful supercomputing systems Tier-1 in a European HPC e-infrastructure

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

Vision of the European Strategy Forum on Research Infrastructures (ESFRI)

Delivering a turnkey operational solution for a future persistent European HPC ecosystem integrating national Tier-1 centres and the new European Tier-0 centres

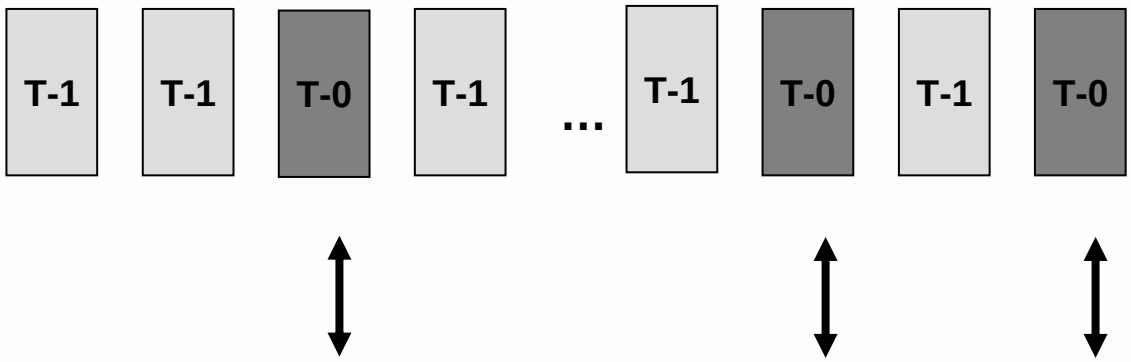
Vision of a European HPC Ecosystem in FP7

DEISA2 - The Infrastructure for the European HPC Ecosystem

Deep operational and technological integration of European HPC (T-0 and T-1) centres and systems providing efficient seamless access to shared HPC resources and large data repositories designing and approving an operational model for a large European Virtual HPC Centre.

Providing scientists access to a large distributed HPC environment via integrated services.

DEISA is paving the way to the efficient operation of the T-0 and T-1 ecosystem

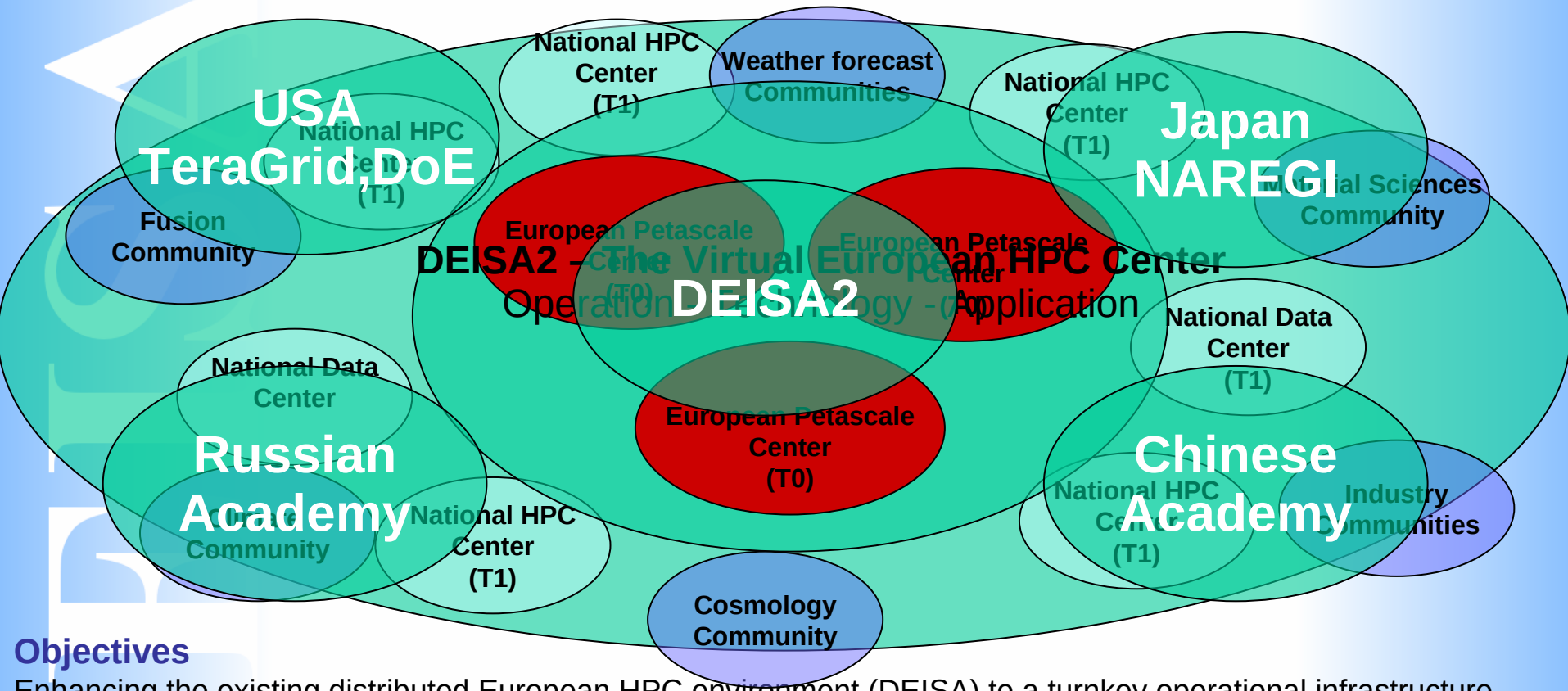


Enhancing the performance

PRACE

Building a world-class pan-European High Performance HPC Ecosystem
which is operated under the umbrella of an European Legal Entity adopting operational and technological concepts and services designed and approved by DEISA2.

Towards a European HPC Infrastructure – DEISA2



Objectives

Enhancing the existing distributed European HPC environment (DEISA) to a turnkey operational infrastructure

Advancing the computational sciences in Europe by supporting user communities and extreme computing projects

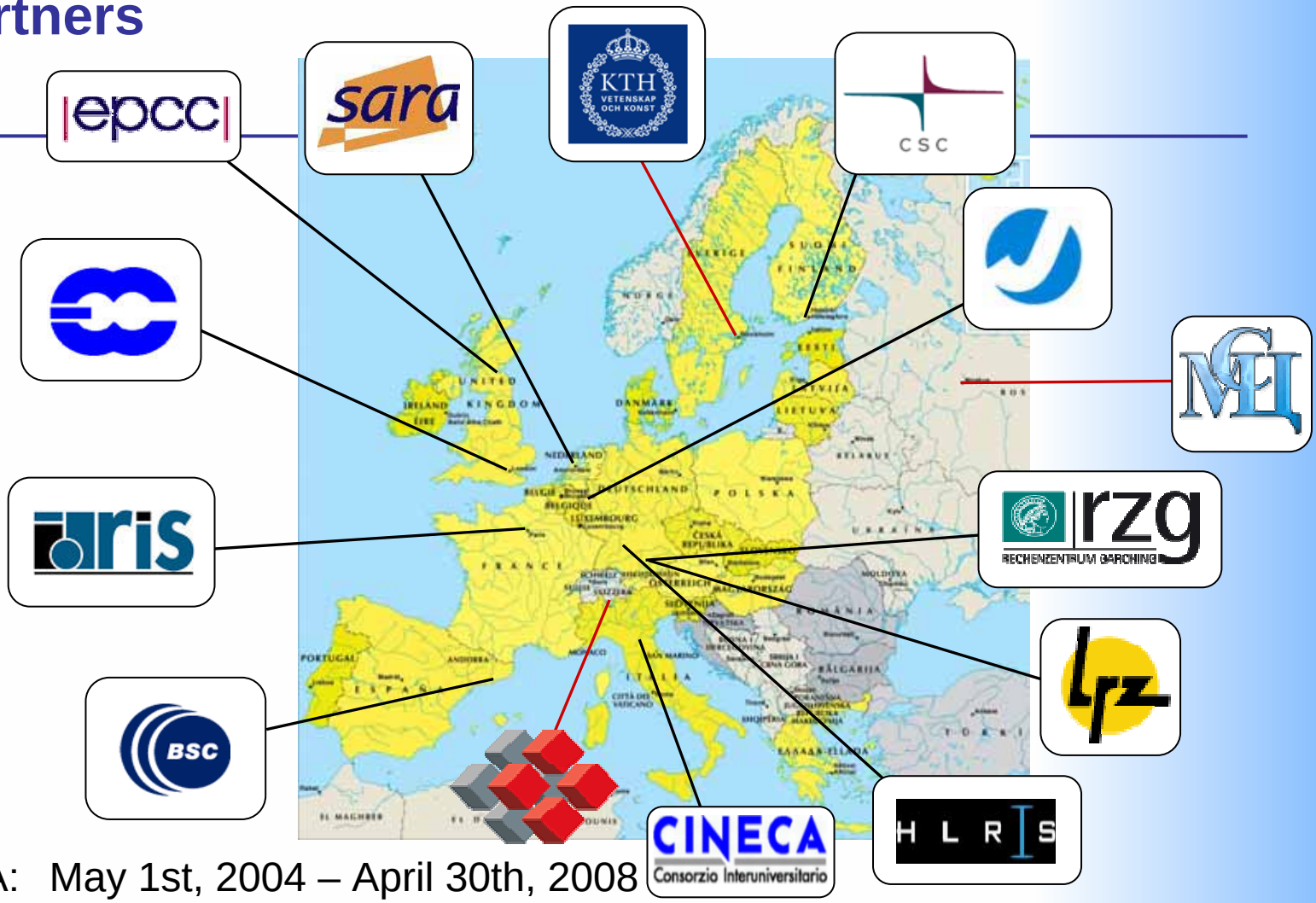
- Enhancing the service provision by offering a complete variety of options of interaction with computational resources

Integration of T1 and T0 centres

- The Petascale Systems need a transparent access from and into the national data repositories

Bridging worldwide HPC projects

DEISA Partners



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

Three new partners joined June 2005 (eDEISA)

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

Partners/ Associate Partners

BSC	***Barcelona Supercomputing Centre***	**Spain**
CINECA	***Consortio Interuniversitario per il Calcolo Automatico***	**Italy**
CSC	***Finnish Information Technology Centre for Science***	**Finland**
EPCC	***University of Edinburgh and CCLRC***	**UK**
ECMWF	***European Centre for Medium-Range Weather Forecast***	**UK (int)**
FZJ	***Research Centre Juelich***	**Germany**
HLRS	***High Performance Computing Centre Stuttgart***	**Germany**
IDRIS	***Institut du Développement et des Ressources en Informatique Scientifique - CNRS***	**France**
LRZ	***Leibniz Rechenzentrum Munich***	**Germany**
RZG	***Rechenzentrum Garching of the Max Planck Society***	**Germany**
SARA	***Dutch National High Performance Computing***	**Netherlands**
KTH	**Kungliga Tekniska Högskolan**	**Sweden**
CSCS	**Swiss National Supercomputing Centre**	**Switzerland**
JSCC	**Joint Supercomputer Center of the Russian Academy of Sciences**	**Russia**

The basic DEISA infrastructures and services

Dedicated high speed network infrastructure

Common AAA infrastructure

Global data management infrastructure

Integrating distributed data with distributed computing platforms, including hierarchical storage management and databases. Major highlights are:

- High performace remote I/O and data sharing with global file systems, using full network badwidth
- High performance tranfers of large data sets, using full network bandwidth

DCPE (DEISA Common Production Environment)

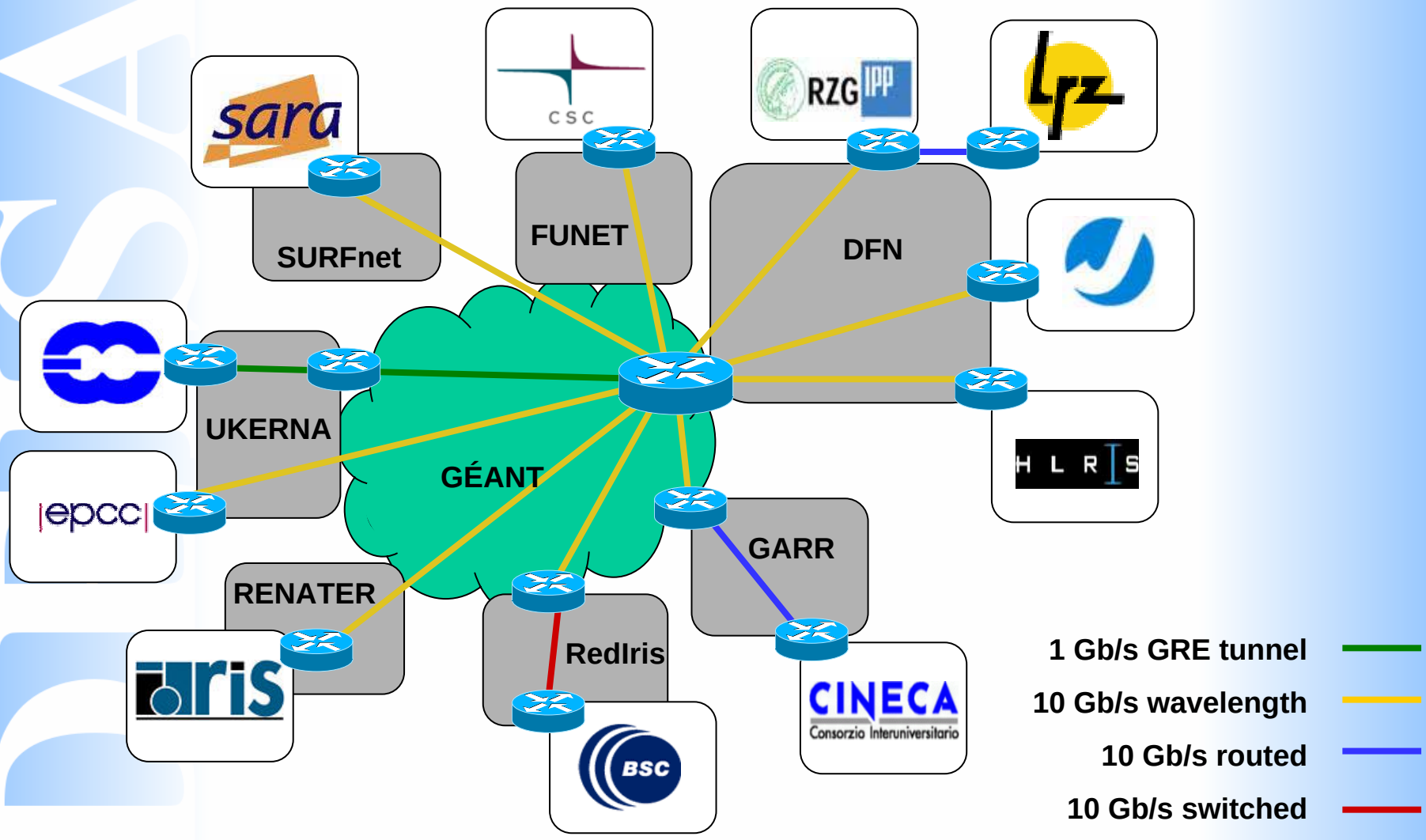
- The job management service
- The science gateways (portals) to supercomputing resources

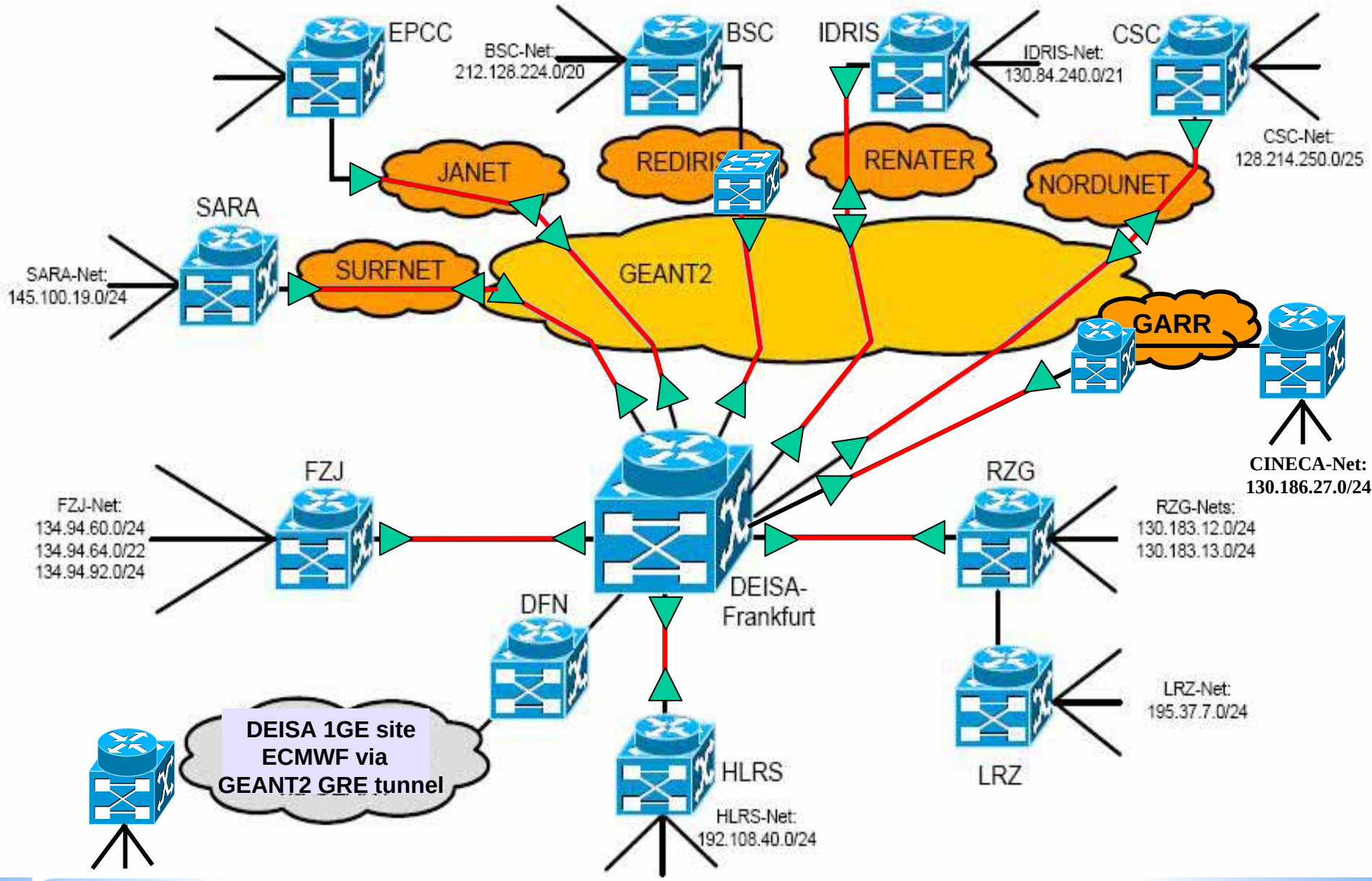
Common Operation Environment

- Common monitoring and Information systems
- Common system operation
- Common help desk

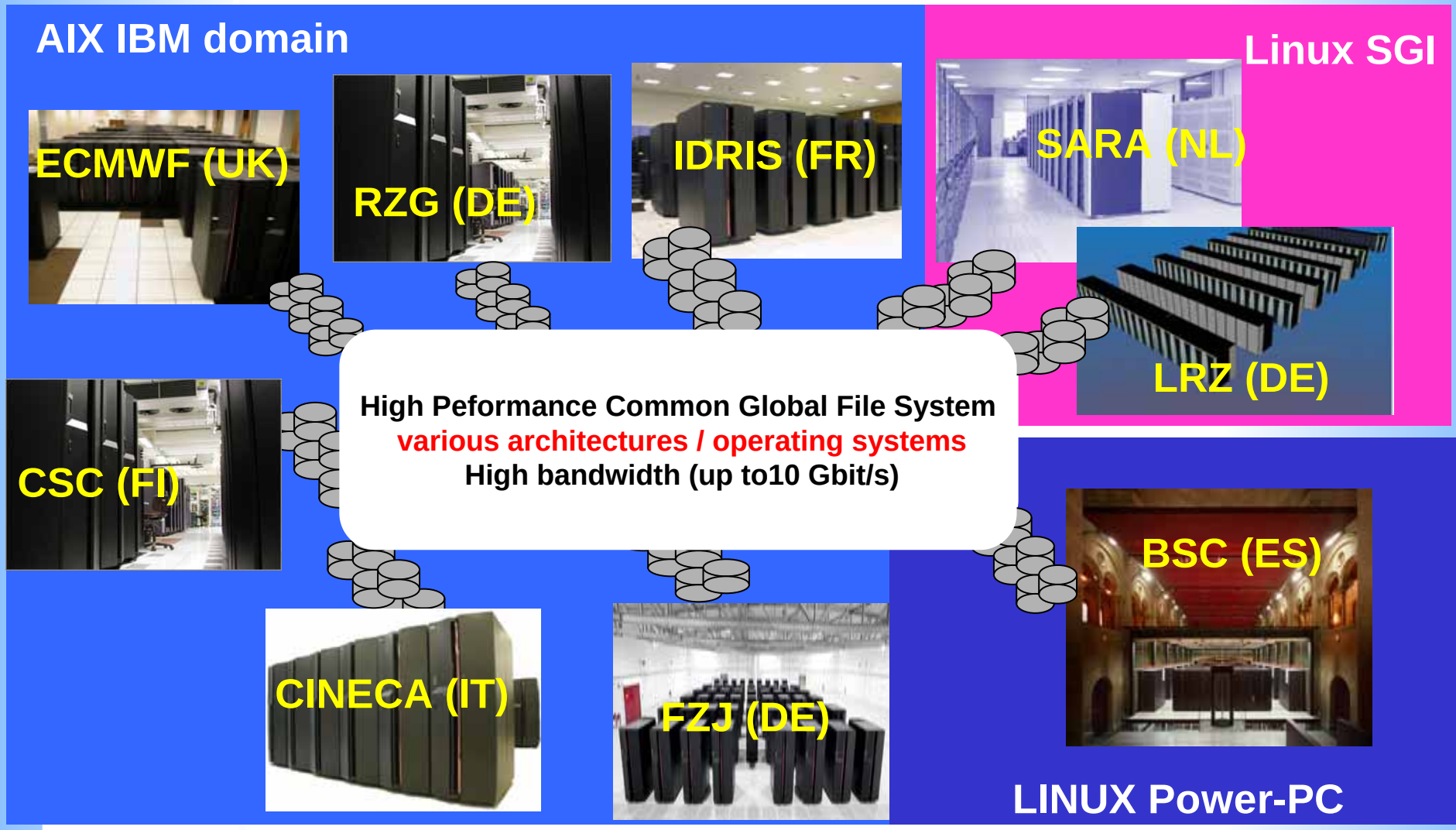
Global Application Support

DEISA network infrastructure

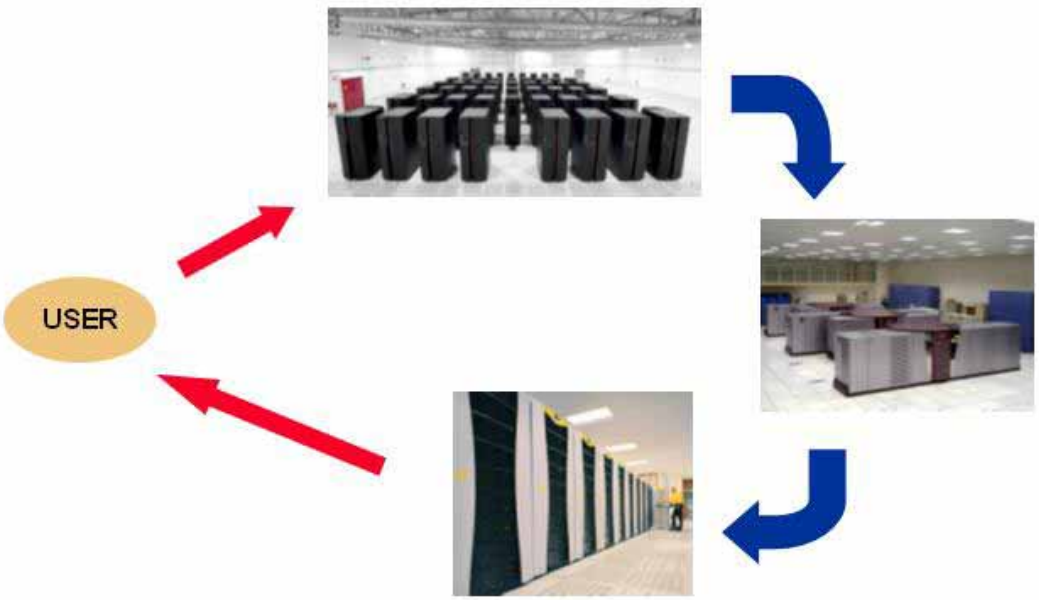



EPCC
BSC
BSC-Net:
212.128.224.0/20
IDRIS
IDRIS-Net:
130.84.240.0/21
CSC
CSC-Net:
128.214.250.0/25
JANET
REDIRIS
RENATER
NORDUNET
SARA
SARA-Net:
145.100.19.0/24
SURFNET
GEANT2
GARR
CINECA-Net:
130.186.27.0/24
FZJ
FZJ-Net:
134.94.60.0/24
134.94.64.0/22
134.94.92.0/24
RZG
RZG-Nets:
130.183.12.0/24
130.183.13.0/24
DEISA-
Frankfurt
DFN
LRZ-Net:
195.37.7.0/24
LRZ
DEISA 1GE site
ECMWF via
GEANT2 GRE tunnel
HLRS
HLRS-Net:
192.108.40.0/24
ECMWF-Net:
136.156.48.0/24

DEISA Data Services - Global File System
(based on IBM's GPFS)



Workflow simulations using UNICORE



UNICORE supports complex simulations that are pipelined over several heterogeneous platforms (workflows).

UNICORE handles workflows as a unique job and transparently moves the output – input data along the pipeline.

UNICORE clients that monitor the application can run in laptops.

UNICORE has a user friendly graphical interface. DEISA has developed a command line interface for UNICORE.

Enabling Science

The basic service providing model for scientific users is the Extreme Computing Initiative (see www.deisa.org)

Identification, deployment and operation of a number of « flagship » applications requiring the infrastructure services, in selected areas of science and technology.

European Call for proposals in May-June every year. Applications are selected on the basis of scientific excellence, innovation potential and relevance criteria, with the collaboration of the national HPC evaluation committees.

Supported by the Applications Task Force (ATASKF)

Hyperscaling of huge parallel applications, data oriented applications

Workflows and coupled applications

Production of an European Benchmark Suite for HPC systems

DEISA Extreme Computing Initiative

DECI call 2005

51 proposals, 12 European countries involved
30 mio cpu-h requested
29 proposals accepted, 12 mio cpu-h
(standardized to P4+ at FZJ)

DECI call 2006

41 proposals, 12 European countries involved
28 mio cpu-h requested
23 proposals accepted, 12 mio cpu-h

DECI call 2007

63 proposals, 14 European countries involved
(US, Canada, Brazil, Israel)
70 mio cpu-h requested
45 proposals accepted, 30 mio cpu-h

DEISA Extreme Computing Initiative

15 European countries so far involved in DECI:

Austria	Finland	France	Germany
Hungary	Italy	Netherlands	Poland
Portugal	Romania	Russia	Spain
Sweden	Switzerland	UK	

With collaborators from four other continents

Asia	Australia	North America	South America

The purpose of the DEISA2

Consolidation of the existing DEISA infrastructure

Guaranteeing the continuity of those activities and services that currently contribute to the effective support of world-leading computational science in Europe

DEISA2 providing a lean and reliable turnkey operational solution for a persistent European HPC ecosystem

Evolvement this European infrastructure towards a robust and persistent European HPC ecosystem

Enhancing the existing services, by deploying new services including support for European Virtual Communities, and by cooperating and collaborating with new European initiatives, especially PRACE that will enable shared European PetaFlop/s supercomputer systems

DEISA2 as the vector for the integration of Tier-0 and Tier-1 systems in Europe

Integrated Tier-1 Systems

Participant short name	*Architecture*	*TF/s[1]*	*#Cores[2]*	*Memory in TB[3]*
MPG-RZG	IBM BlueGene P	27.0	8,192	4.0
	IBM Power6	120.0	6,400	20.0
BAdW-LRZ	SGI-Altix Itanium	62.3	9,728	39.0
BSC	IBM PowerPC	94.2	10240	20.0
CINECA	IBM Blade Center	26.6	5120	10.0
	IBM Power5	3.9	512	1.0
CSC	Cray XT4 dc	10.1	2,024	2.0
	Cray XT4 qc	70.0	6,736	7.8
FZJ	IBM BlueGene P	223.0	65,536	32.0
	IBM p690 cluster JUMP	8.9	1,312	5.0
	JUMP successor	> 50	n/a	n/a
IDRIS-CNRS	IBM p690+ Cluster	6.7	1800	10
	IBM p690+ successor	> 40	n/a	n/a
SARA	IBM Power5+	14.6	1920	7.7
	IBM Power6	60.2	3328	15.6
UEdin-EPCC	Cray XT4	65.0	11,520	35.0
UStutt-HLRS	NEC SX-8	13.0	576	9.0

Authentication

DEISA vision is one global federated namespace, enabling Single Sign-on facilities.

- Direct (interactive) access to all systems required for
  - Testing/debugging of code on specific architectures
  - Submission and checking of production jobs

DEISA AuthN is X.509 based PKI

- DEISA trusts IGTF accredited CAs
- Difficult to handle by users

Shibboleth technology interesting

- User doesn't have to handle authN information – handled by middleware

No interoperability issues for AuthN

- Although DEISA may need more attribute information than IdPs deliver (including attribute certificates)

IGTF
International Grid Trust Federation
AP | EU | TAG

Authorization (1)

Vision is simple authorization scheme

- Same policies everywhere

Access to systems based on subjectName (DN) from X.509 certificate

- DN mapped to UNIX uids
- No pool accounts – access to data (global file system) based on uid.

Access through UNICORE (v5) and GT4

- Authorization information through UUDB (UNICORE) and grid-mapfile (Globus).

Authorization information based on LDAP based repository

- No dynamic retrieval of authZ information – UUDB and gridmapfile are updated statically – different from VOMS model where authZ information is added dynamically (to attribute or proxy certificates).
- Basically not very different from VOMS model – both systems give attribute information about users - VOMS repository could be filled from LDAP system
  - But middleware must be able to handle the information

Authorization (2)

AuthZ information providers

- LDAP repository managed by sites, not by VOs! Within HPC/DEISA accounts are managed by sites, so would also be done for VOMS servers.
- Partners/sites are the VOs
- Policies needed for managing authZ information, like for authN through PMAs

Interoperability -

- DEISA can operate a VOMS server for use by other infrastructures
- DEISA can import data to LDAP from external VOMS server (but additional attribute information needed)

Accounting (1)

- **Vision: uniform model for usage and budget information**
- **Usage information published using OGF UR-WG format recommendation**
- **Developed our own system for publishing usage records – existing systems did have dependencies on other tools**
  - Each site publishes data locally in DB (eXist)
  - Access based on role – user, PI (principal investigator), site admin
    - Based on X.509 SubjectName
  - GUI for producing reports (DART)
  - Conversion (comparison) between systems based on CPU performance

Accounting (2)

Wishlist

- Publishing summary records
- All data in central repository (automatically)
- Budget check based on usage information
  - Planning implementation – no job submission possible if budget is exceeded

Interoperability

- Data can be exported in XML
- For importing usage information a separate DB could be set up
- But only subset of UR-WG attributes is supported

Thank you!


RZG
IPP
epcc
sara
CSC
HLRS
ECMWF
JÜLICH
FORSCHUNGSZENTRUM
CINECA
Consorzio Interuniversitario