eGee

Enabling Grids for E-sciencE

gLite Overview

Mike Mineter
National e-Science Centre, Edinburgh

Seoul, August 2005

www.eu-egee.org

Information Society

**Enabling Grids for E-sciencE**

- **Grid concepts**
- **Background to gLite**
- **gLite services**
- **gLite status**

Additional information is found in hidden slides in the file that is available from the agenda page, http://agenda.cern.ch/fullAgenda.php?ida=a054582

Grid concepts

egee

- **EGEE is establishing a production grid infrastructure to support multiple, diverse VO's**

User Interface

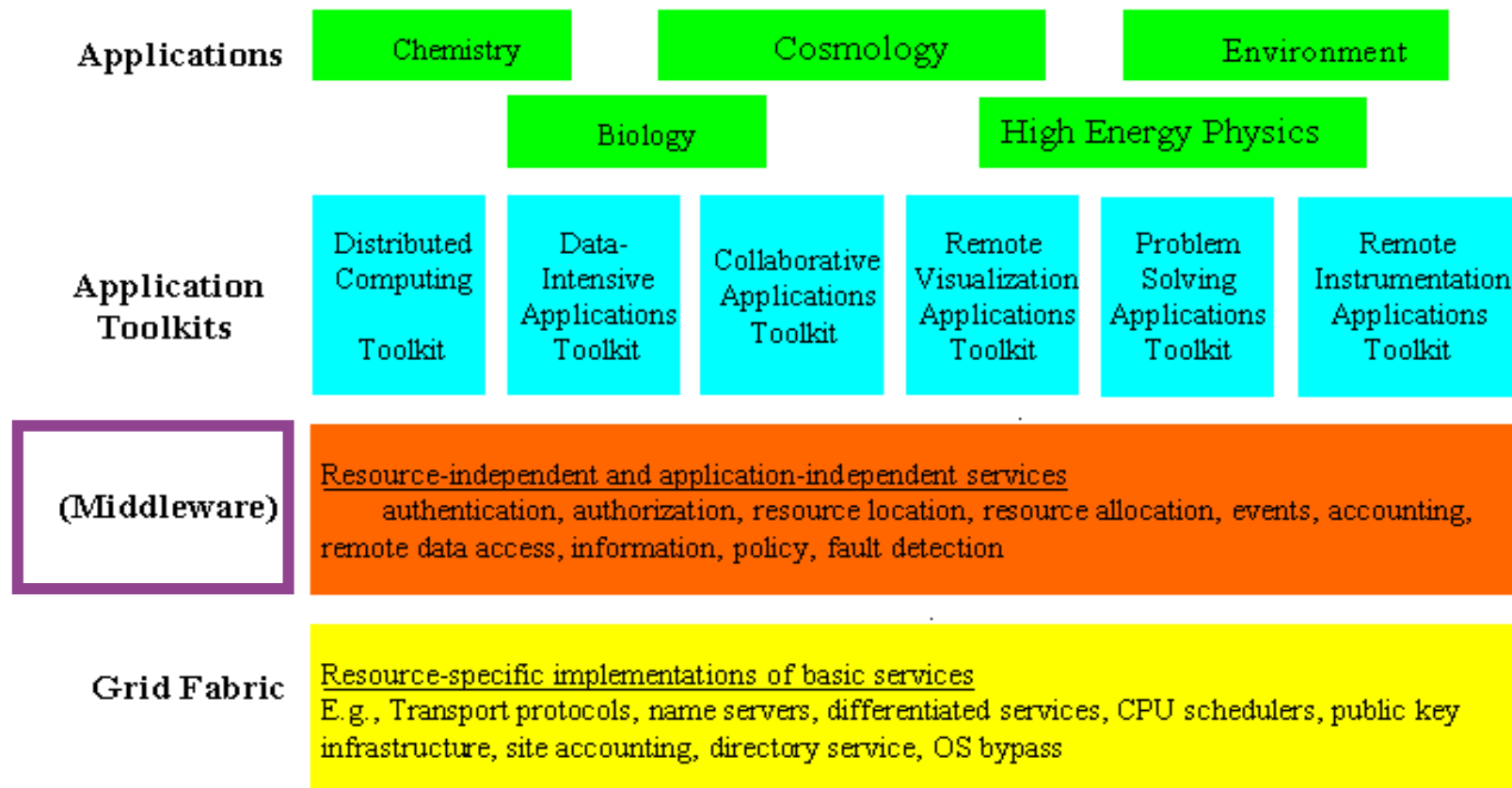**User Interface**
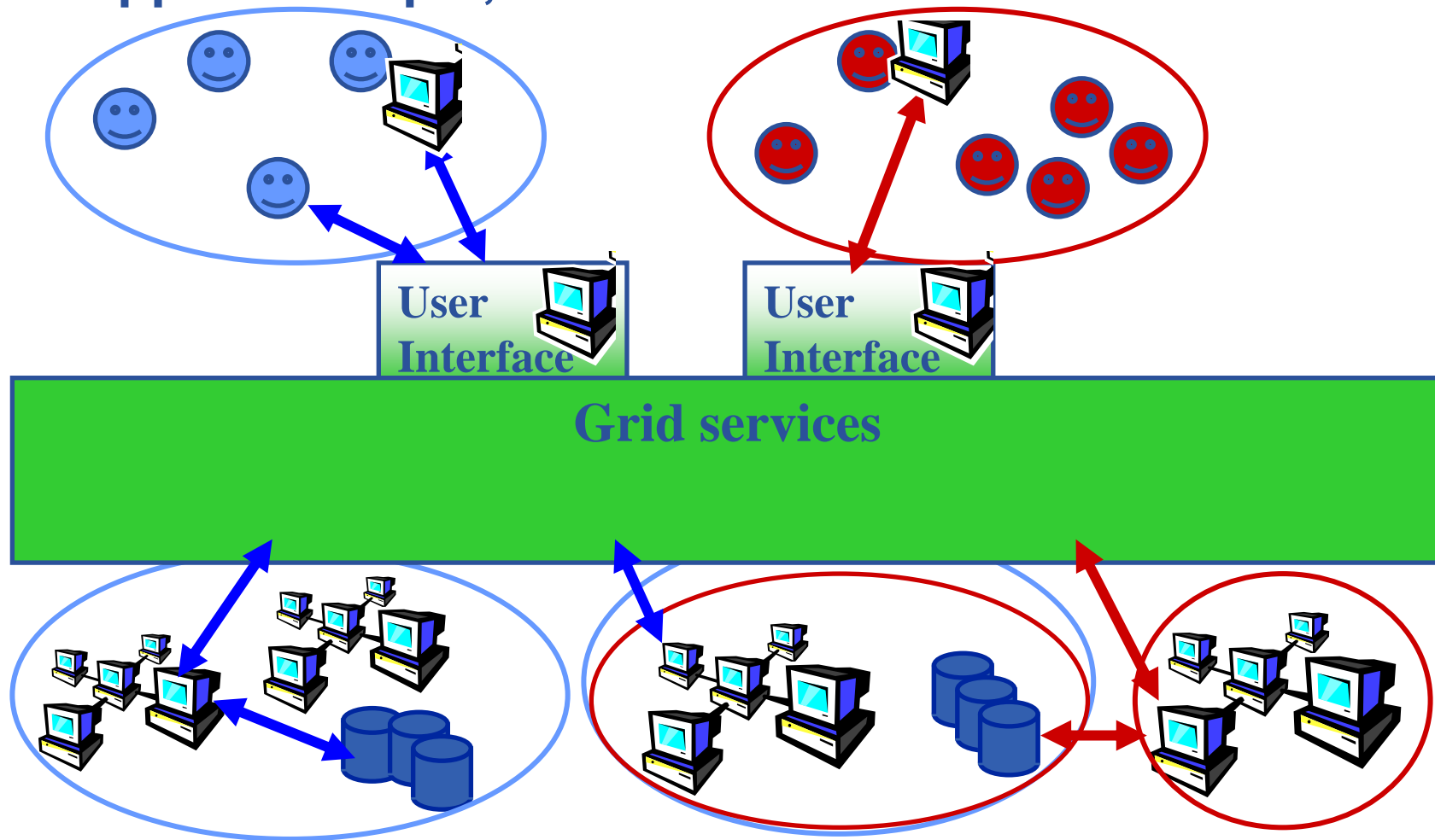
Grid services

- **Users join VO**
- **Virtual organisation contributes resources & negotiates access**

- **Grid middleware runs on each resource**
  - "Storage elements"
  - "Compute elements"

- **Additional services (both people and grid middleware) enable the grid**

- **Effect: "virtual computing" across administrative domains empowering collaboration**

INTERNET

Enabling Grids for E-sciencE
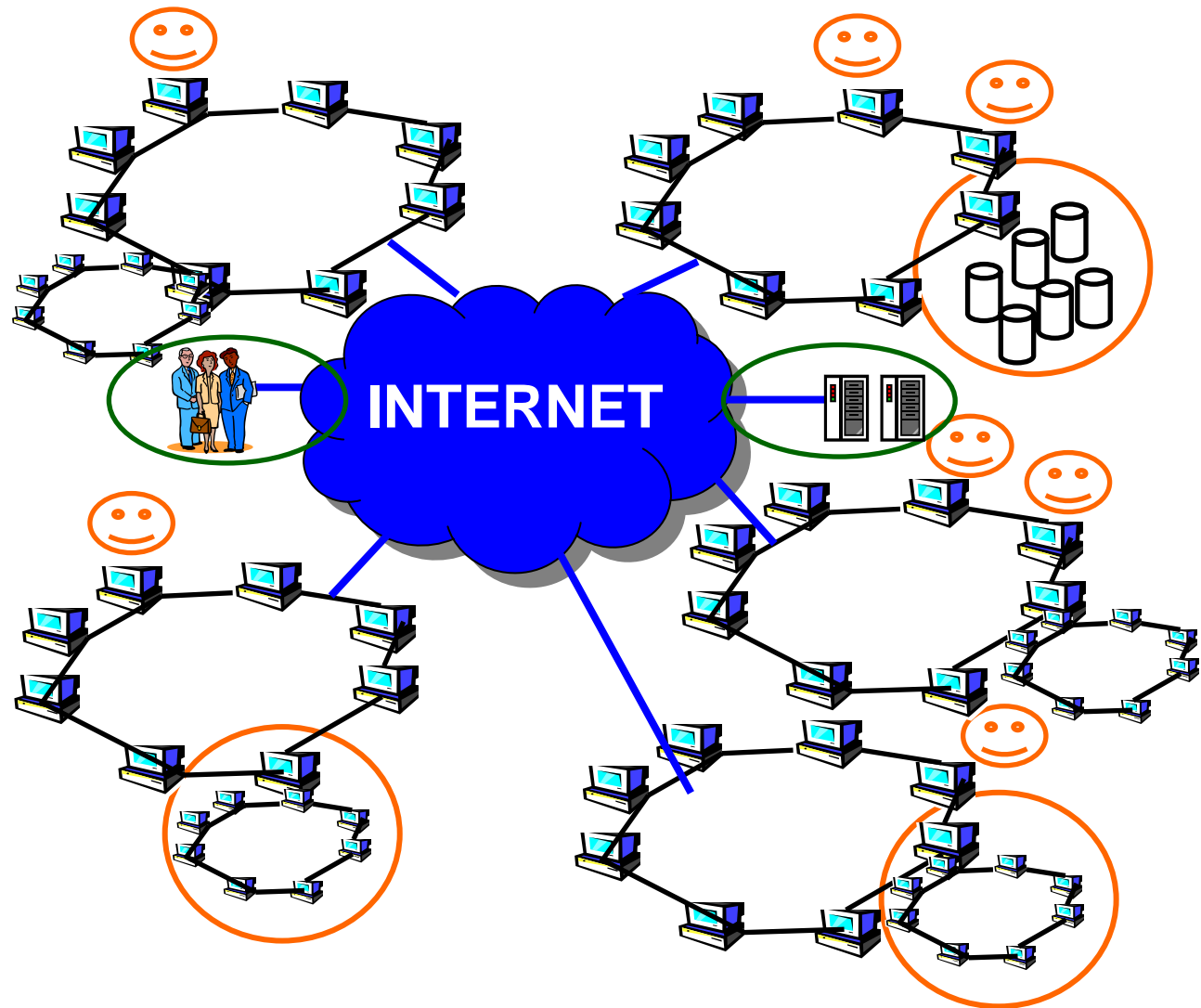
- **Virtual organisation: people who collaborate by sharing resources – e.g. data, storage, CPU's, programs - across administrative and organisational boundaries**

- **Single sign-on**
  - I connect to one machine – some sort of "digital credential" is passed on to any other resource I use, basis of:
    - *Authentication*: How do I identify myself to a resource without username/password for each resource I use?
    - *Authorisation*: what can I do? Determined by
      - *My membership of a VO*
      - *VO negotiations with resource providers*

- **Grid middleware – "the operating system of a grid"**
  - on each resource
  - services that enable the grid

- **User just perceives "shared resources" with no concern for location or owning organisation**

The Grid from a Services View

Applications	Chemistry		Cosmology		Environment	
		Biology		High Energy Physics		

Application Toolkits

Distributed Computing Toolkit	Data-Intensive Applications Toolkit	Collaborative Applications Toolkit	Remote Visualization Applications Toolkit	Problem Solving Applications Toolkit	Remote Instrumentation Applications Toolkit

(Middleware)

Resource-independent and application-independent services
authentication, authorization, resource location, resource allocation, events, accounting, remote data access, information, policy, fault detection

Grid Fabric

Resource-specific implementations of basic services
E.g., Transport protocols, name servers, differentiated services, CPU schedulers, public key infrastructure, site accounting, directory service, OS bypass

Application

Application toolkits, standards

Middleware:
"collective services"

Basic Grid services:
AA, job submission, info, …

- **The tools, services used by the VO's applications**
- **Community-specific standards**
- **Application development environment, portals, semantics, workflow**
- **In EGEE-1: Mainly VO-specific**

- **Emphasis of EGEE -1 middleware**
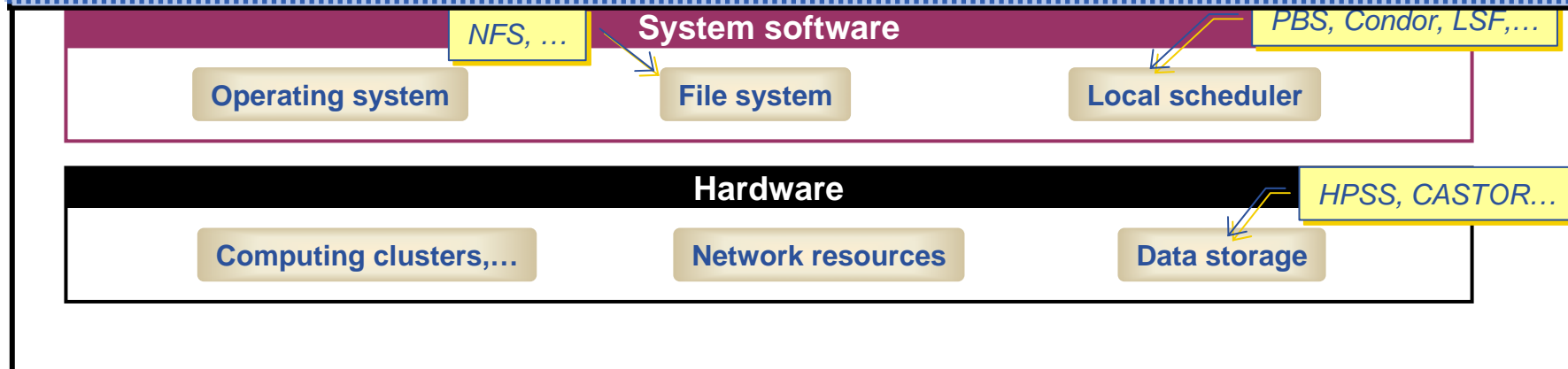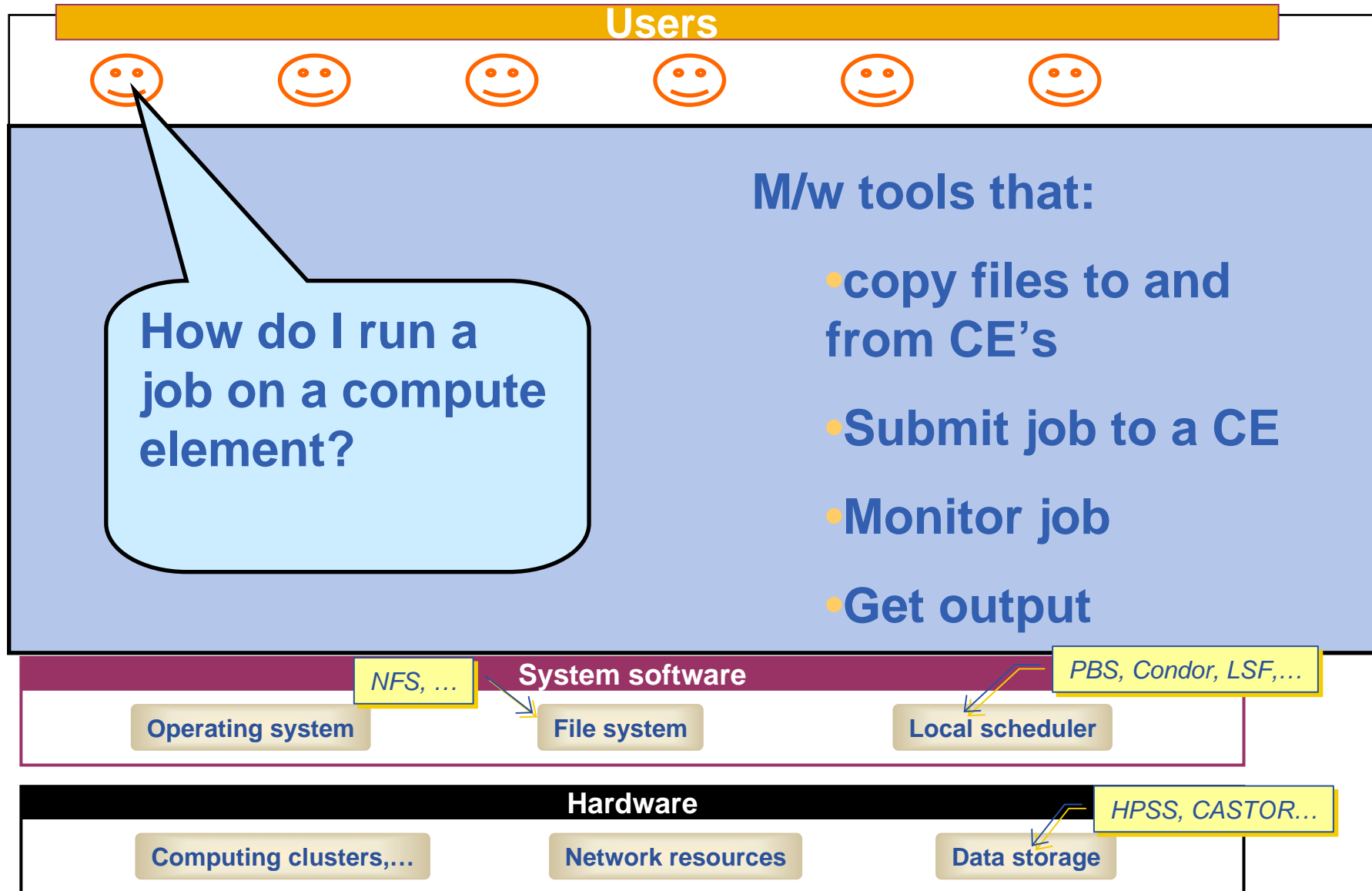
Users in many locations and organisations

**"User interface" machines :
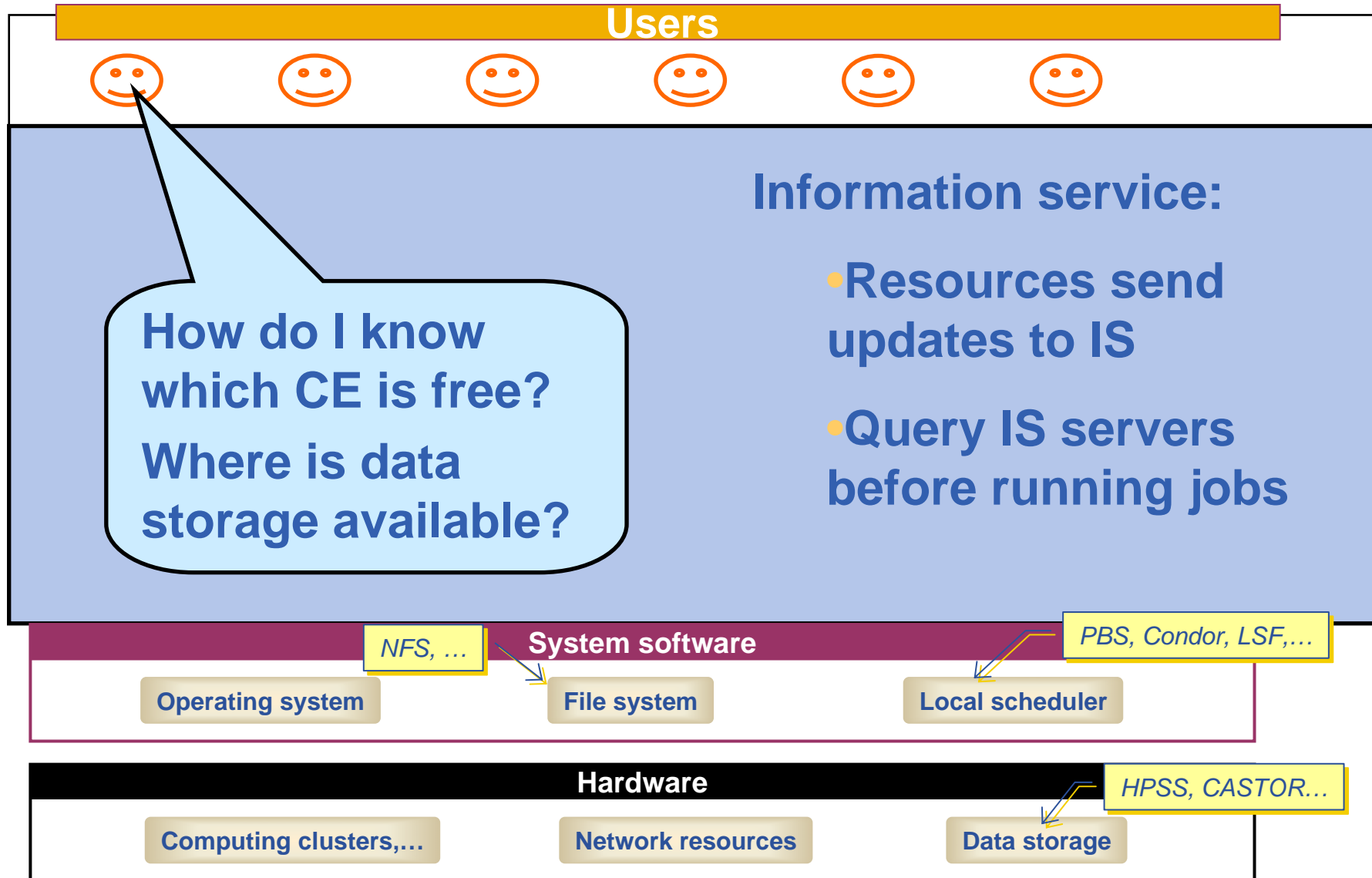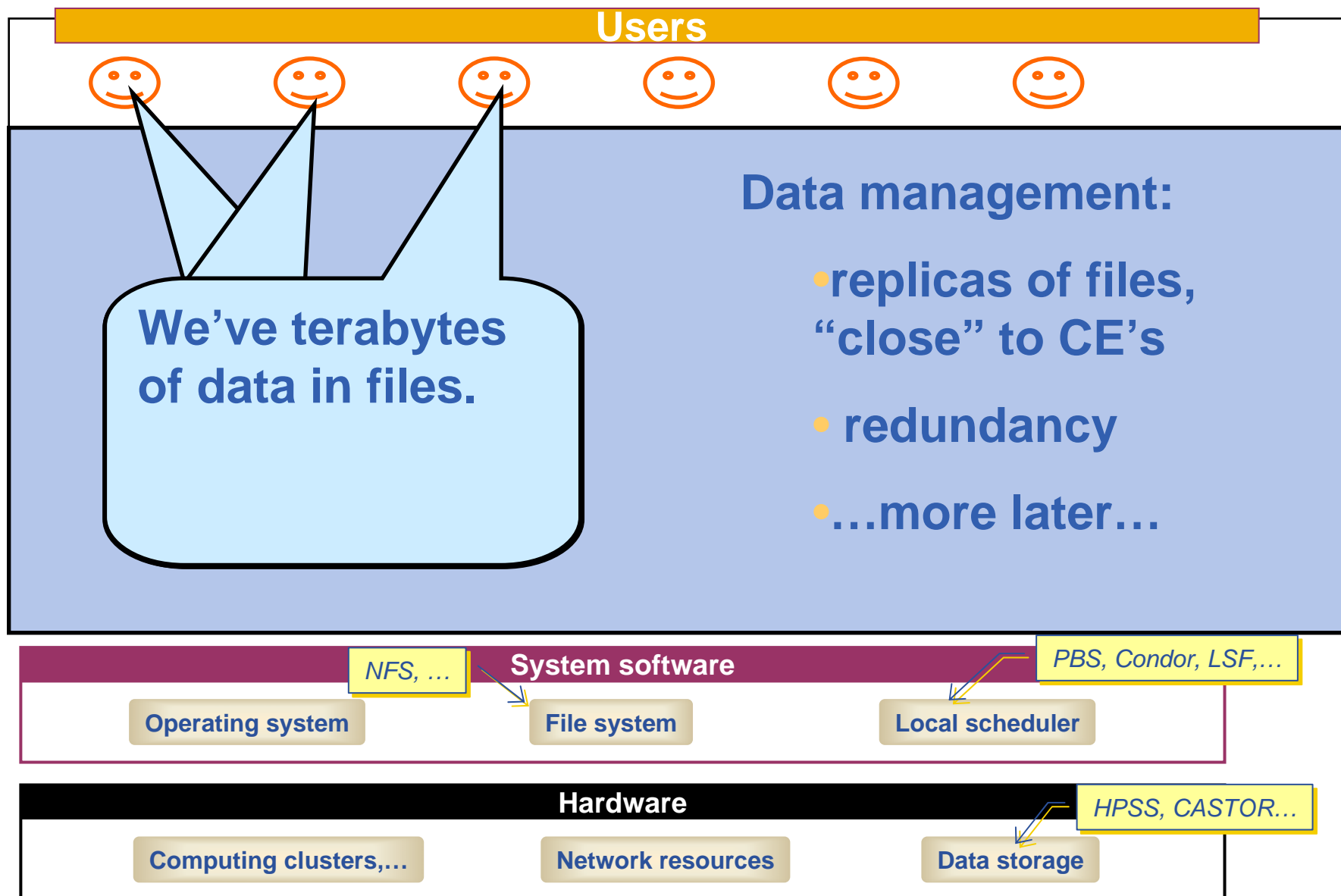logon, upload credentials, run m/w commands**

MIDDLEWARE

Built on Grid Security Infrastructure

**"Gate keeping": middleware on resources
map user's credential to local user id / account**

System software

NFS, …

PBS, Condor, LSF,…

Operating system   **File system**   **Local scheduler**

Hardware

HPSS, CASTOR…

Computing clusters,…   **Network resources**   **Data storage**

egee

Enabling Grids for E-sciencE

Users

How do I run a job on a compute element?

M/w tools that:

- **copy files to and from CE's**

- **Submit job to a CE**

- **Monitor job**

- **Get output**

System software

NFS, …

| Operating system | File system | Local scheduler |

PBS, Condor, LSF,…

Hardware

HPSS, CASTOR…

| Computing clusters,… | Network resources | Data storage |

eGee

Enabling Grids for E-sciencE

Users

Information service:

- **Resources send updates to IS**

- **Query IS servers before running jobs**

How do I know which CE is free?

Where is data storage available?

NFS, …

System software

PBS, Condor, LSF,…

Operating system   **File system**   **Local scheduler**

Hardware

HPSS, CASTOR…

Computing clusters,…   **Network resources**   **Data storage**

eGee

Enabling Grids for E-sciencE

Users

We've terabytes of data in files.

Data management:

- replicas of files, "close" to CE's
- redundancy
- …more later…

System software

NFS, …

Operating system

File system

PBS, Condor, LSF,…

Local scheduler

Hardware

HPSS, CASTOR…

Computing clusters,…

Network resources

Data storage

- **A software toolkit: a modular "bag of technologies"**
  - Made available under liberal open source license
- ***Not* turnkey solutions, but *building blocks* and *tools* for application developers and system integrators**
- **Tools built on Grid Security Infrastructure to include:**
  - Job submission: run a job on a remote computer
  - Information services: So I know which computer to use
  - File transfer: so large data files can be transferred
  - Replica management: so I can have multiple versions of a file "close" to the computers where I want to run jobs
- **Production grids are (currently) based on the Globus Toolkit release 2 … so is gLite**
- **Globus Alliance: http://www.globus.org/**

- **Command line interface to the tool for job submission – need to know name of a Compute Element**

globus-job-submit grid-data.rl.ac.uk/jobmanager-pbs /bin/hostname -f

https://grid-data.rl.ac.uk:64001/1415/1110129853/

globus-job-status https://grid-data.rl.ac.uk:64001/1415/1110129853/

DONE

globus-job-get-output https://grid-data.rl.ac.uk:64001/1415/1110129853/

grid-data12.rl.ac.uk

- **Build on this tool to support job submission _to the grid_, not just to a named CE – and to make more friendly interfaces for users**
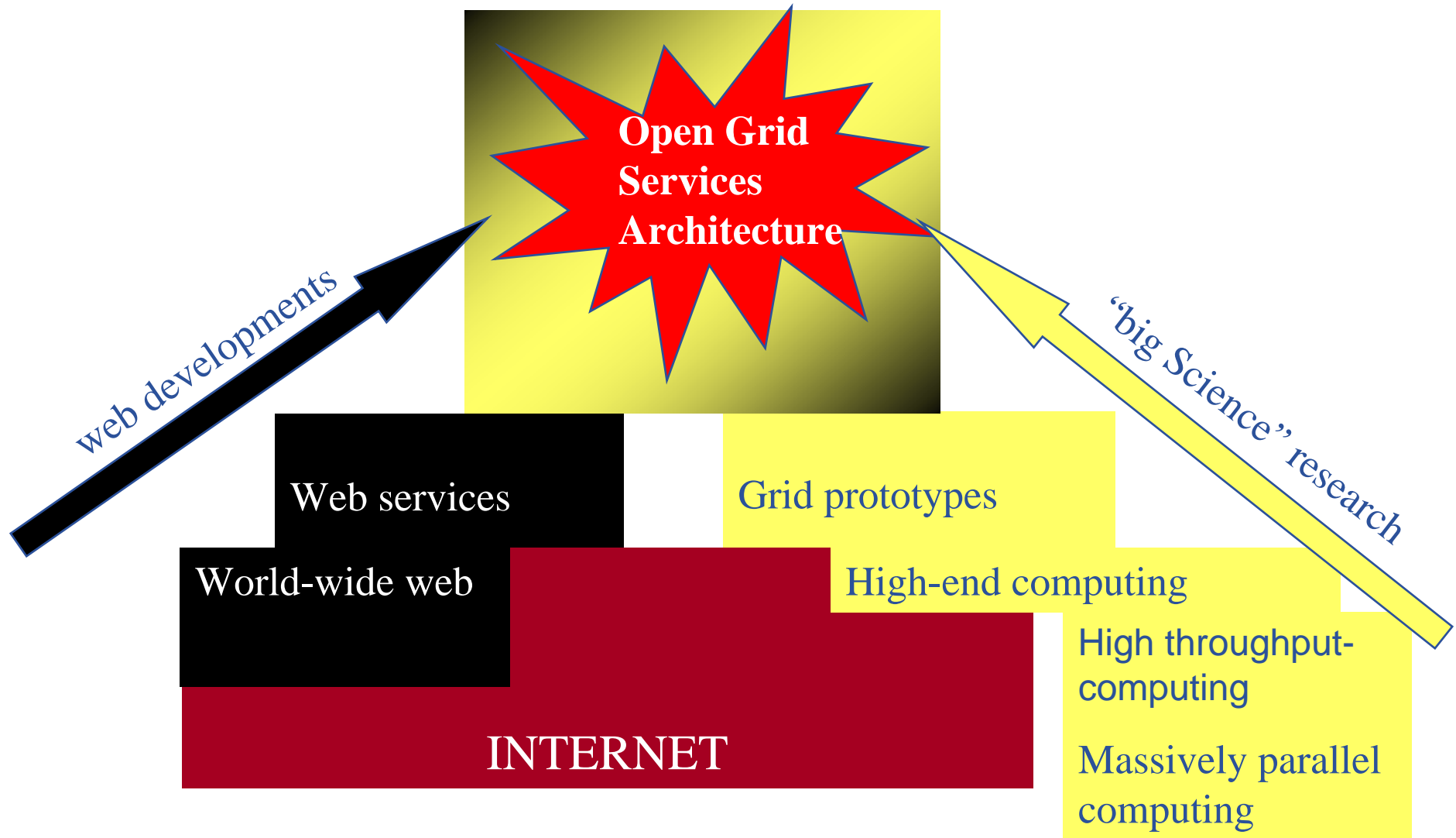
Background to gLite

- **Many key concepts identified and known**
- **Many grid projects have tested, and benefit from, these**
- **Major efforts now on establishing:**
  - Standards (a slow process)
    (e.g. Global Grid Forum, http://www.gridforum.org/ ,
    OASIS, W3C, IETF)
  - Production Grids *for multiple VO's*
    - "Production" = Reliable, sustainable, with commitments to quality of service
    - One stack of middleware that serves many communities
    - Operational procedures and services (people, policy,..)
  - New user communities
- **… whilst research & development continues**
- **"Service orientation" widely seen as *the* way to build grids**

Open Grid Services Architecture

web developments

"big Science" research

Web services

World-wide web

Grid prototypes

High-end computing

High throughput-computing

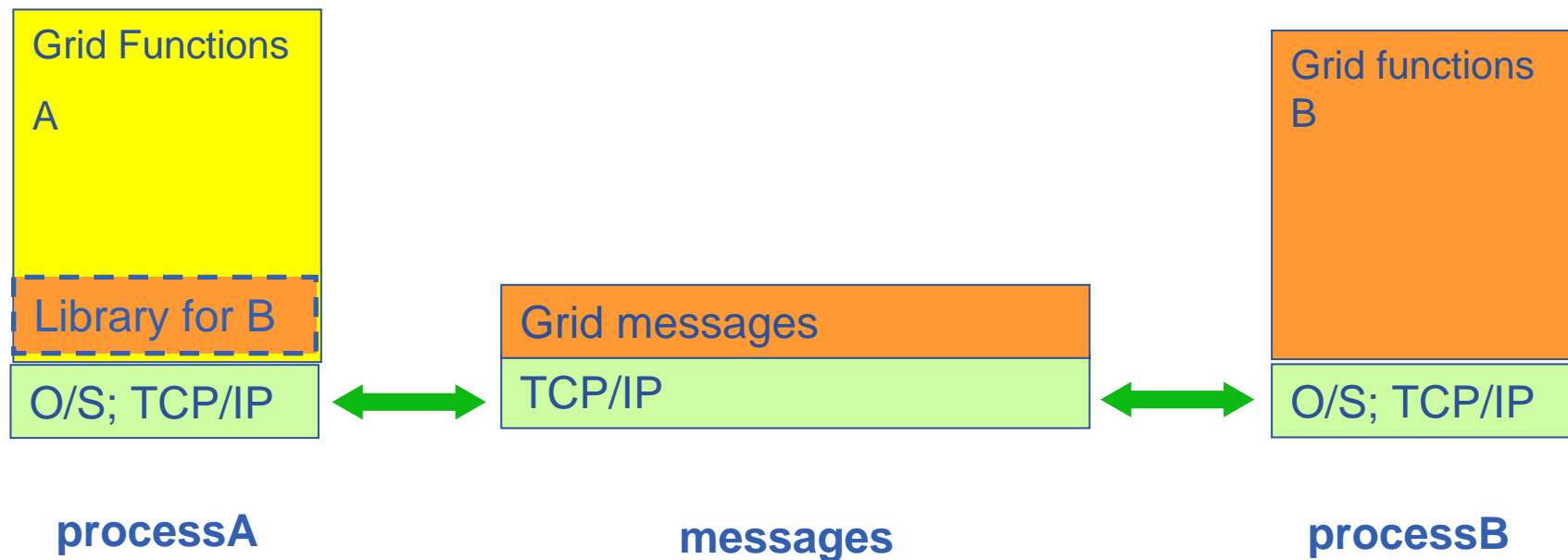Massively parallel computing

INTERNET

- **Service Oriented Architecture**
  - Components are loosely coupled by messages
  - Facilitates interoperability
  - Allows easier compliance with upcoming standards, hosting environments, toolkits
  - Architecture is not bound to specific implementations
  - Heterogeneous resources (storage, computation…)
- **Flexibility in configuration**
  - services can be deployed and used independently
- **Facilitates development of clients for different architectures**
- **The gLite service decomposition has been largely influenced by the work performed in the LCG project**
  - Follow WSRF standardization
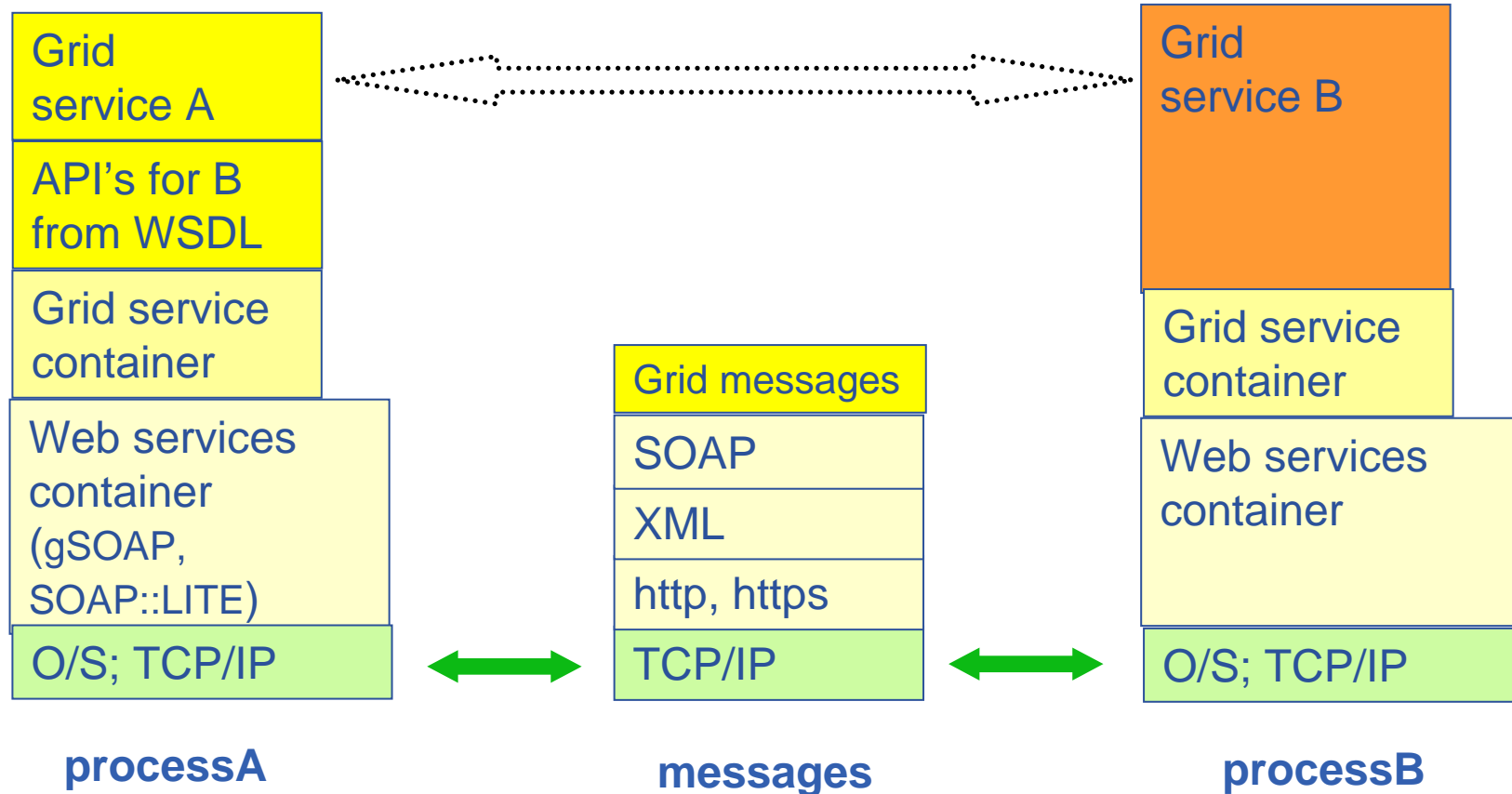 - Start with plain WS (WS-I)

- **Components that are loosely coupled by messages**
  - Accessible across network; modular and self-contained; clean modes of failure
  - So can change implementation without changing interfaces
  - Can be developed in anticipation of new uses
- **… and are based on standards.**

- **Opens EGEE to:**
  - New middleware (plethora of tools now available)
  - Heterogeneous resources (storage, computation…)
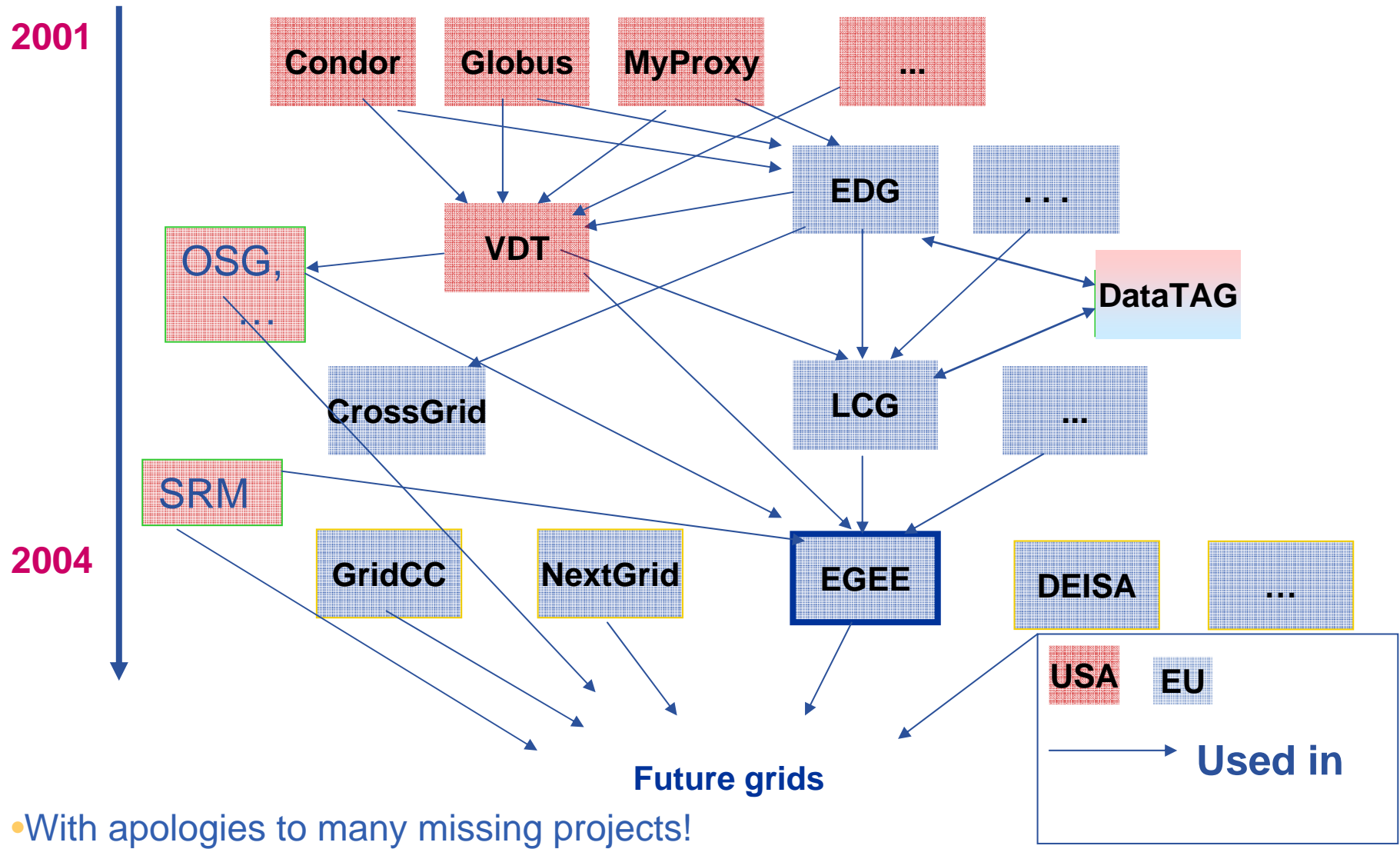 - Interact with other Grids (international, regional and national)

Grids before service orientation

For A to use B, build process with library for B

Tight-coupling: to program A need code and detailed information about B.
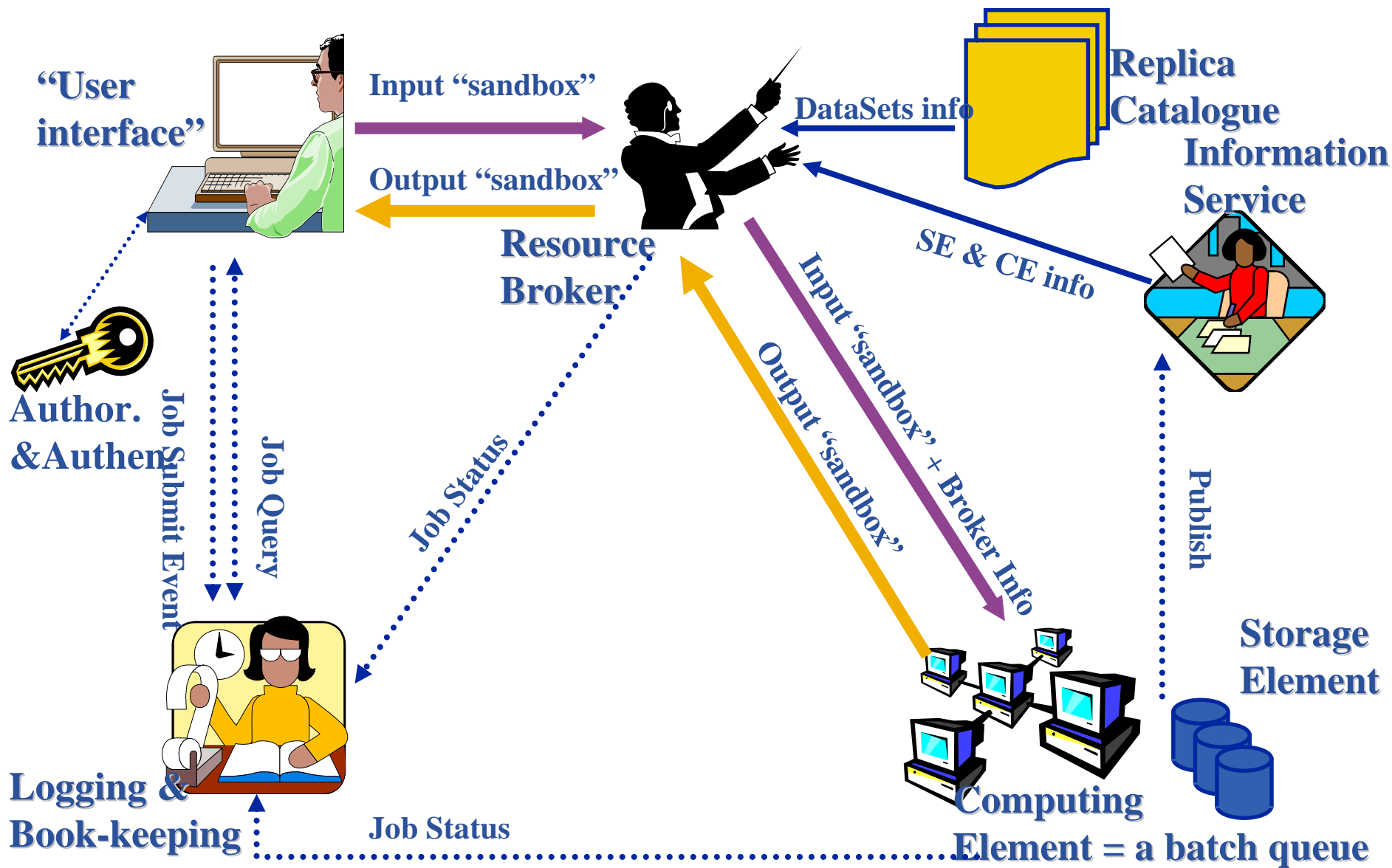
Grid Functions A		Grid functions B
Library for B	Grid messages	
O/S; TCP/IP	TCP/IP	O/S; TCP/IP

processA   **messages**   **processB**

- **Using service B from service A:**
  - From WSDL build APIs to use service
  - (Usually) use SOAP to access service

Grid service A		Grid service B
API's for B from WSDL		
Grid service container	Grid messages	Grid service container
Web services container (gSOAP, SOAP::LITE)	SOAP	Web services container
	XML	
	http, https	
O/S; TCP/IP	TCP/IP	O/S; TCP/IP
processA	**messages**	**processB**

2001

Condor Globus MyProxy ...

EDG ...

VDT

OSG, ..

DataTAG

CrossGrid

LCG ...

SRM

2004

GridCC NextGrid EGEE DEISA ...

USA EU

Used in

Future grids

• With apologies to many missing projects!

"User interface"

Input "sandbox"

Output "sandbox"

Resource Broker

DataSets info

Replica Catalogue

Information Service

SE & CE info

Author. &Authen.

Job Submit Event

Job Query

Job Status

Input "sandbox" + Broker Info

Output "sandbox"

Publish

Storage Element

Logging & Book-keeping

Job Status

Computing Element = a batch queue

- **Real-time monitor**
  - http://www.hep.ph.ic.ac.uk/e-science/projects/demo/index.html
- **Current status**
  - http://goc.grid-support.ac.uk/gridsite/monitoring/

12 August 2005

Total Sites: 188

Total CPUs: 16564

Total Storage (PB): 4

Grid3

gLite

- **Jobs are:**
  - (as in LCG) run from batch queues, termed "computing elements" CE's
  - Described in "Job Description Language"

- **gLite also supports**
  - Interactive jobs
    - Jobs run in batch mode – "listener" receives messages from CE
  - Parallelism using MPI
    - *MPI jobs can run on CE's that support MPI not across administrative domains (not MPICH-G)*
  - Workflow (DAGs, from Condor)
  - Checkpointing
  - Partitioned jobs (soon) – e.g. Monte-Carlo

Enabling Grids for E-sciencE

Simple data

– Files

- **Requires**
  - **Replica files**
    - Move data to computation
  - **Virtual filesystems**
  - **Metadata for files**
  - **File transfer**

- **These services are amongst those provided in gLite**

Structured data

– RDBMS, XML databases

- Require *__extendable__* middleware tools to support
  - *computation near to data*
  - easy access, controlled by AA
  - integration and federation
- Hence **OGSA-DAI**
  DAI: Data Access and Integration

- **OGSA-DAI is NOT currently being ported to gLite**

EGEE middlewares face to face
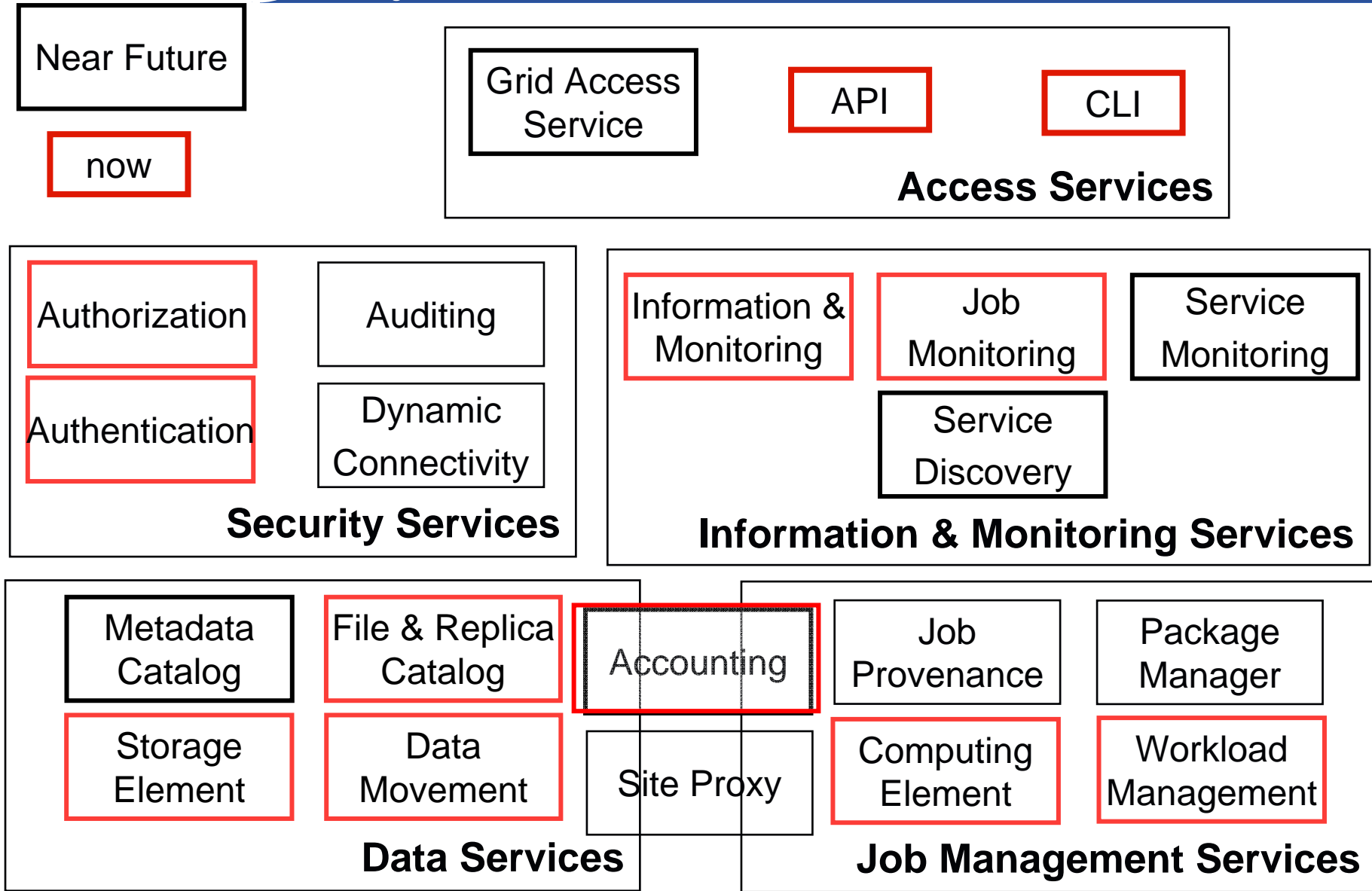
LCG

- **Security**
  - GSI
- **Job Management**
  - Condor + Globus
  - CE, WN
  - Logging & Bookkeeping
- **Data Management**
  - LCG services
- **Information & Monitoring**
  - BDII (evolution of MDS)
- **Grid Access**
  - CLI + API
- **Operating system**
  - Red Hat 7.3

gLite

- **Security**
  - GSI and VOMS
- **Job Management**
  - Condor + Globus + blahp
  - CE, WN
  - Logging & Bookkeeping
  - Job Provenance
  - Package management
- **Data Management**
  - LFC
  - gLite-I/O + FiReMan
- **Information & Monitoring**
  - BDII
  - R-GMA + Service Discovery
- **Grid Access**
  - CLI + API + Web Services
- **Easier installation / configuration**
- **Currently Scientific LINUX, will be available on others, incl. Windows**

Enabling Grids for E-sciencE

- **http://hepix.fzk.de/upload/lectures/BLAH_batch_system _intface.pdf**

- **The protocol**
  - **The BLAHP (Batch Local ASCII Helper Protocol) provides a set of plain ASCII commands used by Condor-C (and CREAM) to manage jobs on the batch systems.**

- **The daemon**
  - **BLAHPD implements the helper daemon responsible for converting BLAHP commands into batch system actions, interpreting their results and reporting them in BLAHP format.**

gLite components overview

eGee
Enabling Grids for E-sciencE

Near Future

now

Access Services

Grid Access Service

API

CLI

Security Services

Authorization

Authentication

Auditing

Dynamic Connectivity

Information & Monitoring Services

Information & Monitoring

Job Monitoring

Service Monitoring

Service Discovery

Data Services

Metadata Catalog

File & Replica Catalog

Accounting

Storage Element

Data Movement

Site Proxy

Job Management Services

Job Provenance

Package Manager

Computing Element

Workload Management

- **Computing element**
  - A queue in a Local Resource Management System
    - Batch jobs that run on site's cluster

- **Storage element**
  - Implements SRM interfaces (+... See later)
  - gLite has been tested with
    - CASTOR
    - dCache
  - Being tested for DPM

Enabling Grids for E-sciencE

- **Used in "pre-production" mode**
- **gLite v1.3 released 05/08/2005**
  - File Placement Service, File Placement Service clients added to UI and WNs modules
  - new data transfer agents including architecture refactoring to allow proper inter-VO scheduling
- **gLite v1.2 released 22/07/2005**
  - File Transfer Service and the File Transfer Agents
  - improvements in all modules.
- **gLite v. 1.1 released 13/05/2005**
  - File Transfer Service and the Metadata Catalog
- **gLite v. 1.0 released 05/04/2005**

- **http://www.glite.org/**

Enabling Grids for E-sciencE

- **gLite, the EGEE middleware:**
  - Is exiting prototyping phase and entering real production phase
    - LHC first real data are only 2 years away from now!
  - Implements a full and complete stack of grid services

- **Service orientation allows**
  - Use gLite services all together or separately
  - Can migrate from LCG to gLite incrementally

- **Is seeking to balance**
  - Conforming to (emerging) standards
  - Need to deliver a production service that demands efficiency, speed

Enabling Grids for E-sciencE

- **EGEE http://public.eu-egee.org/**

- **gLite http://www.glite.org/**

- **EGEE Middleware Architecture https://edms.cern.ch/document/594698/**