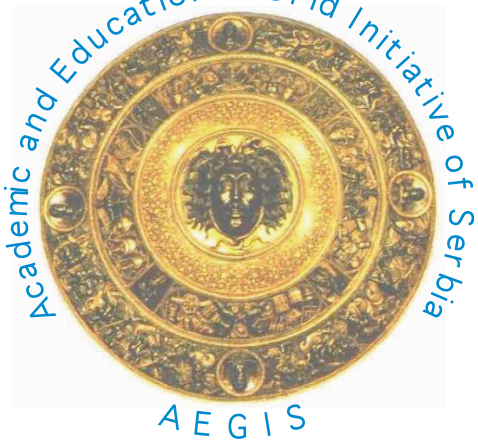
The 2nd workshop on High Performance Computing
IPM and Shahid Beheshti University, Tehran, Iran

Introduction to gLite

Antun Balaz
Scientific Computing Laboratory
Institute of Physics Belgrade, Serbia
http://www.scl.rs/



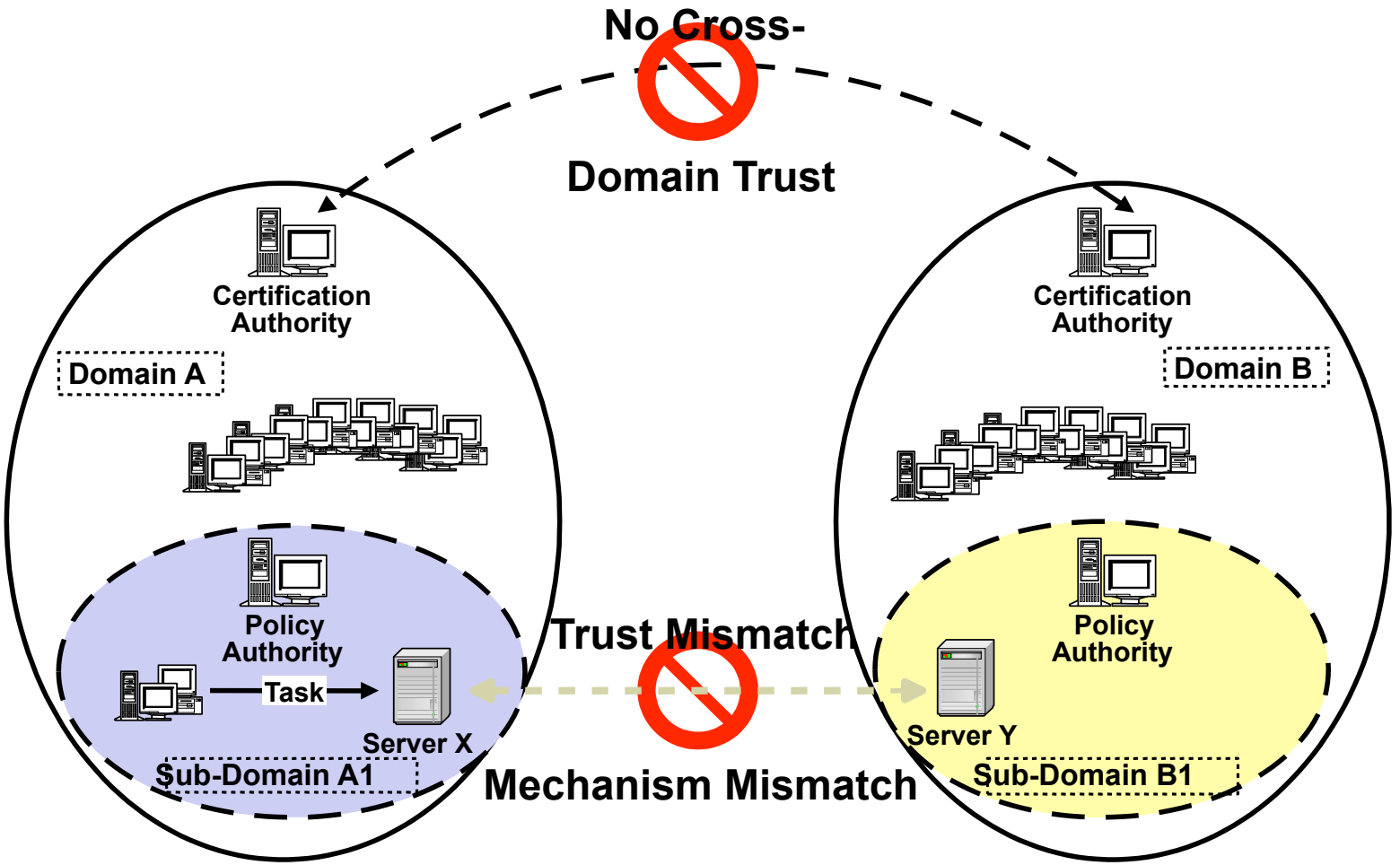
SEE-GRID-SCI
SEE-GRID eInfrastructure for regional eScience

21 Jan 2009 –
01 Feb 2009

Set of basic Grid services

- Job submission/management
- File transfer (individual, queued)
- Database access
- Data management (replication, metadata)
- Monitoring/Indexing system information

Multi-institution issues



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 –
01 Feb 2009

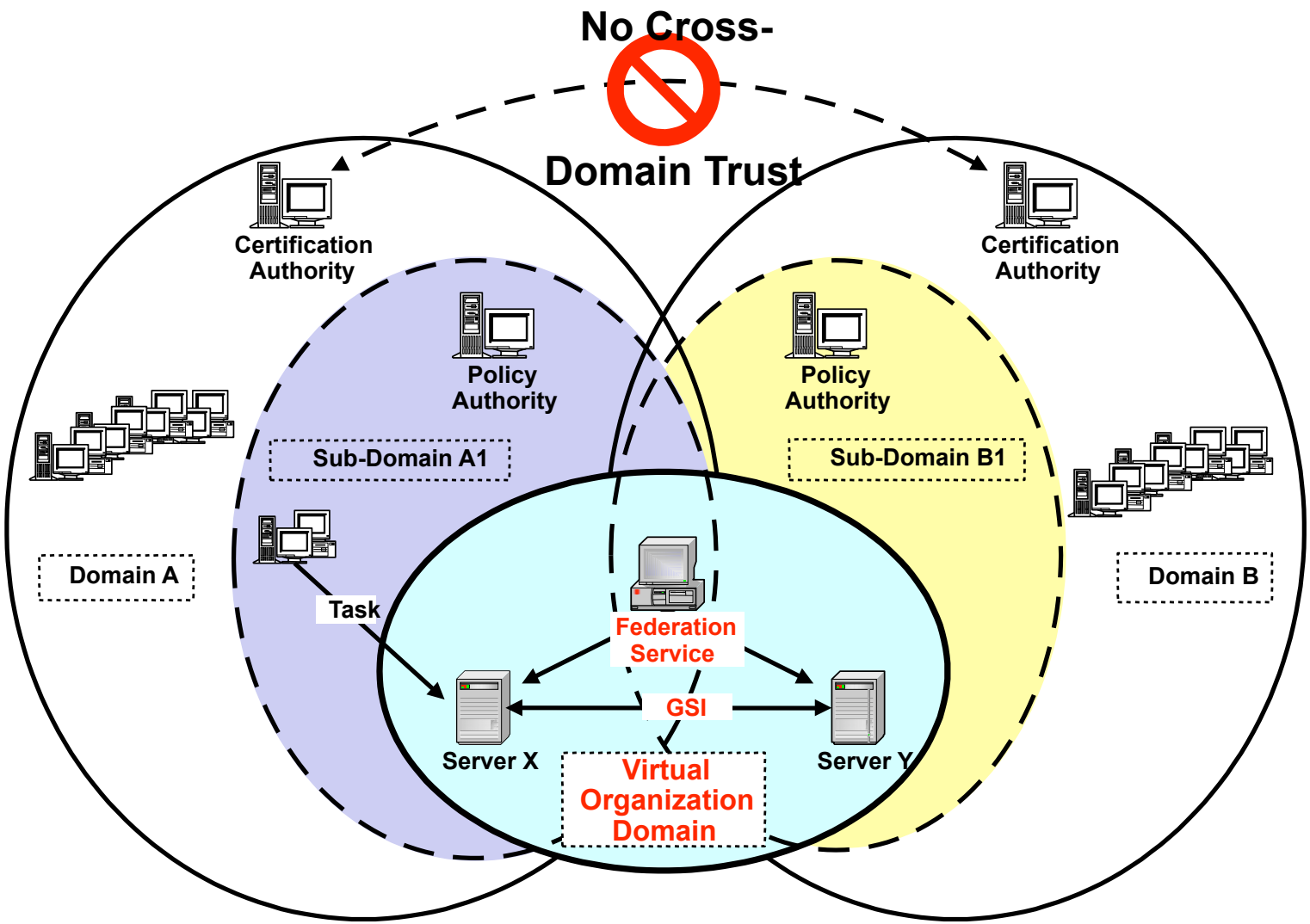
Why Grid security is hard (1)

- Resources being used may be valuable & the problems being solved sensitive
  - Both users and resources need to be careful
- Dynamic formation and management of user groups
  - Large, dynamic, unpredictable…
- Resources and users are often located in distinct administrative domains
  - Cannot assume cross-organizational trust agreements
  - Different mechanisms & credentials

Why Grid security is hard (2)

- Interactions are not just client/server, but service-to-service on behalf of user
  - Requires delegation of rights user ➔ service
  - Services may be dynamically instantiated
- Standardization of interfaces to allow for discovery, negotiation and use
- Implementation must be broadly available & applicable
  - Standard, well-tested, well-understood protocols; integrated with wide variety of tools
- Policy from sites, user communities and users need to be combined
  - Varying formats
- Want to hide as much as possible from applications!

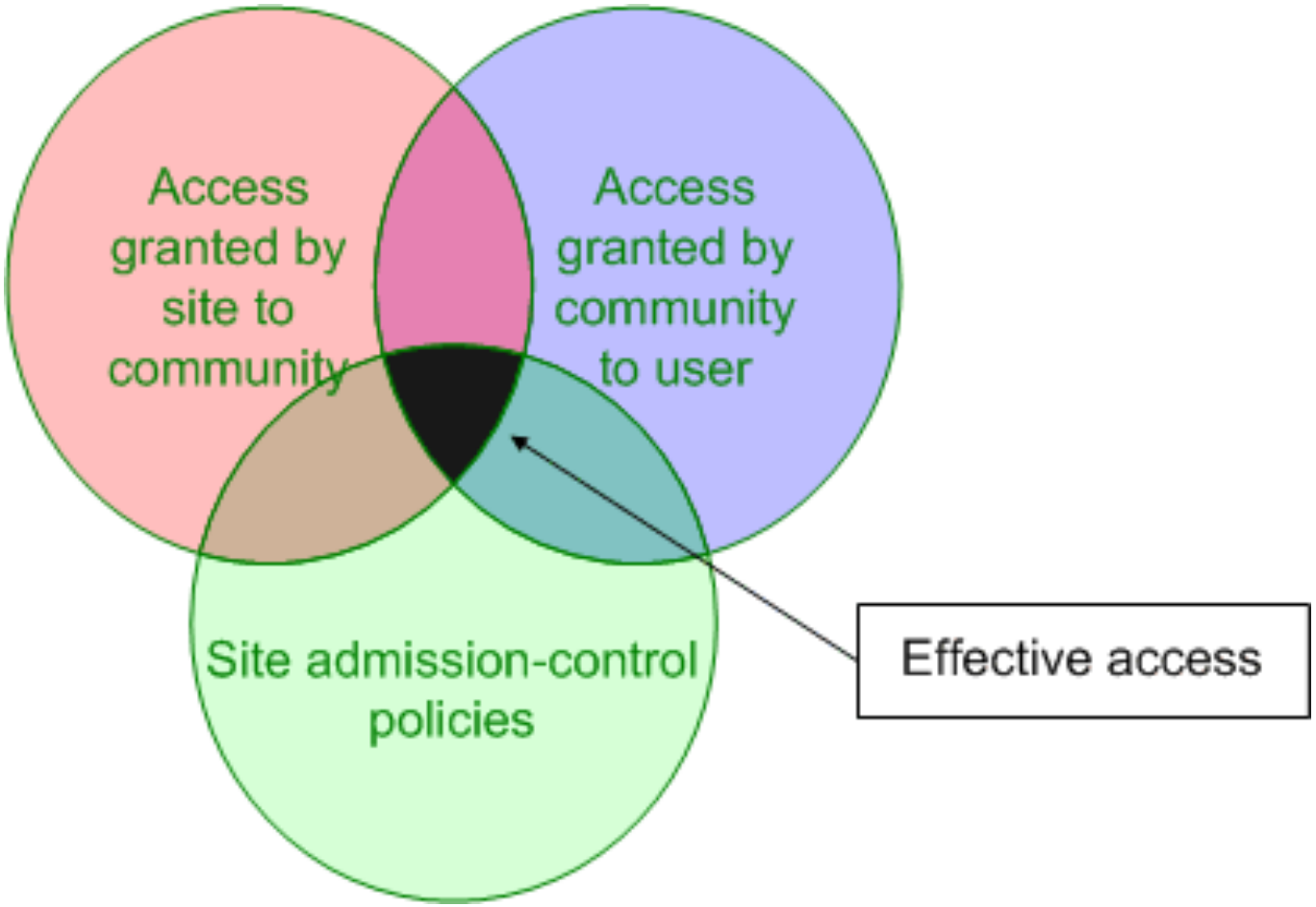
Grid solution: use of VOs



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 –
01 Feb 2009

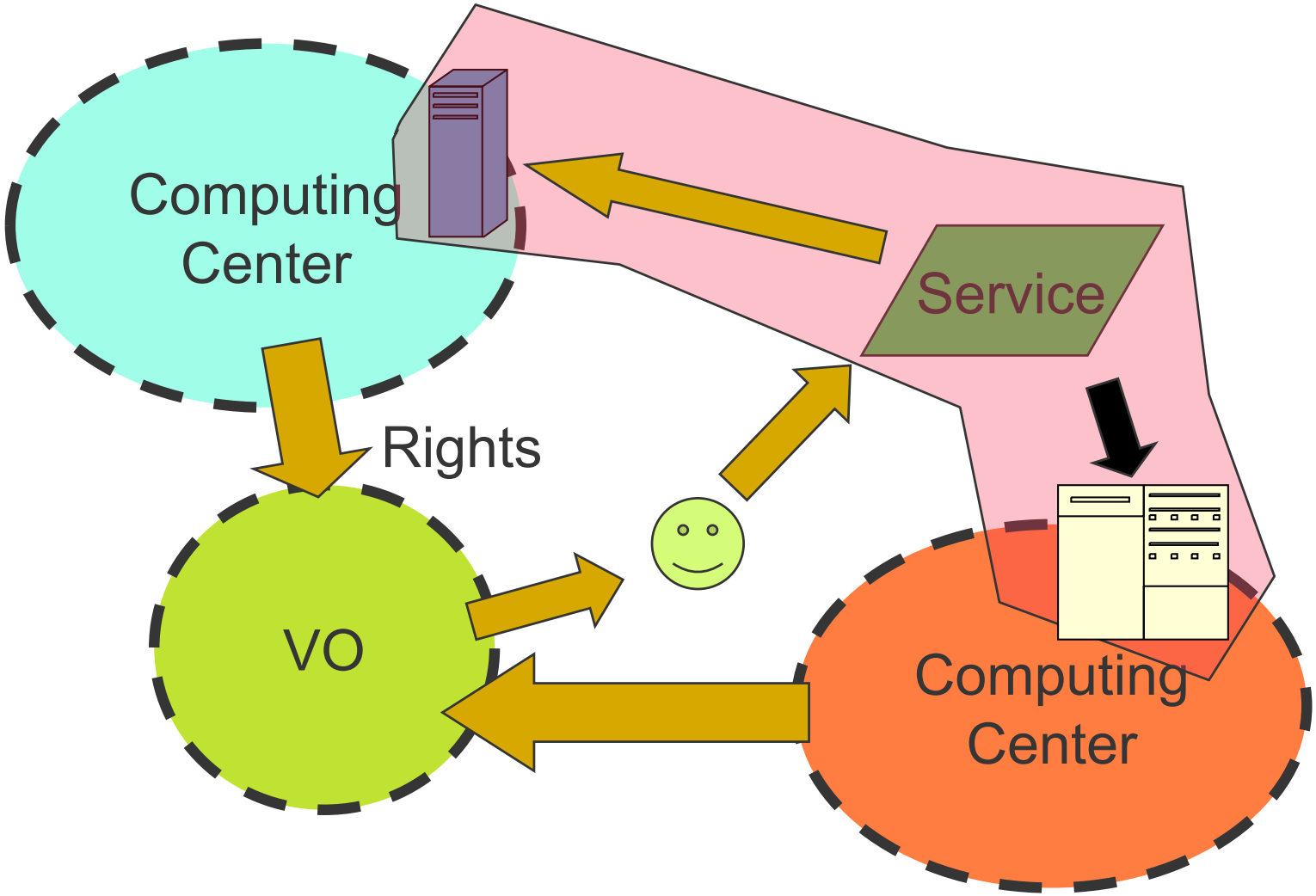
Effective policy governing access within a collaboration



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 – 01 Feb 2009

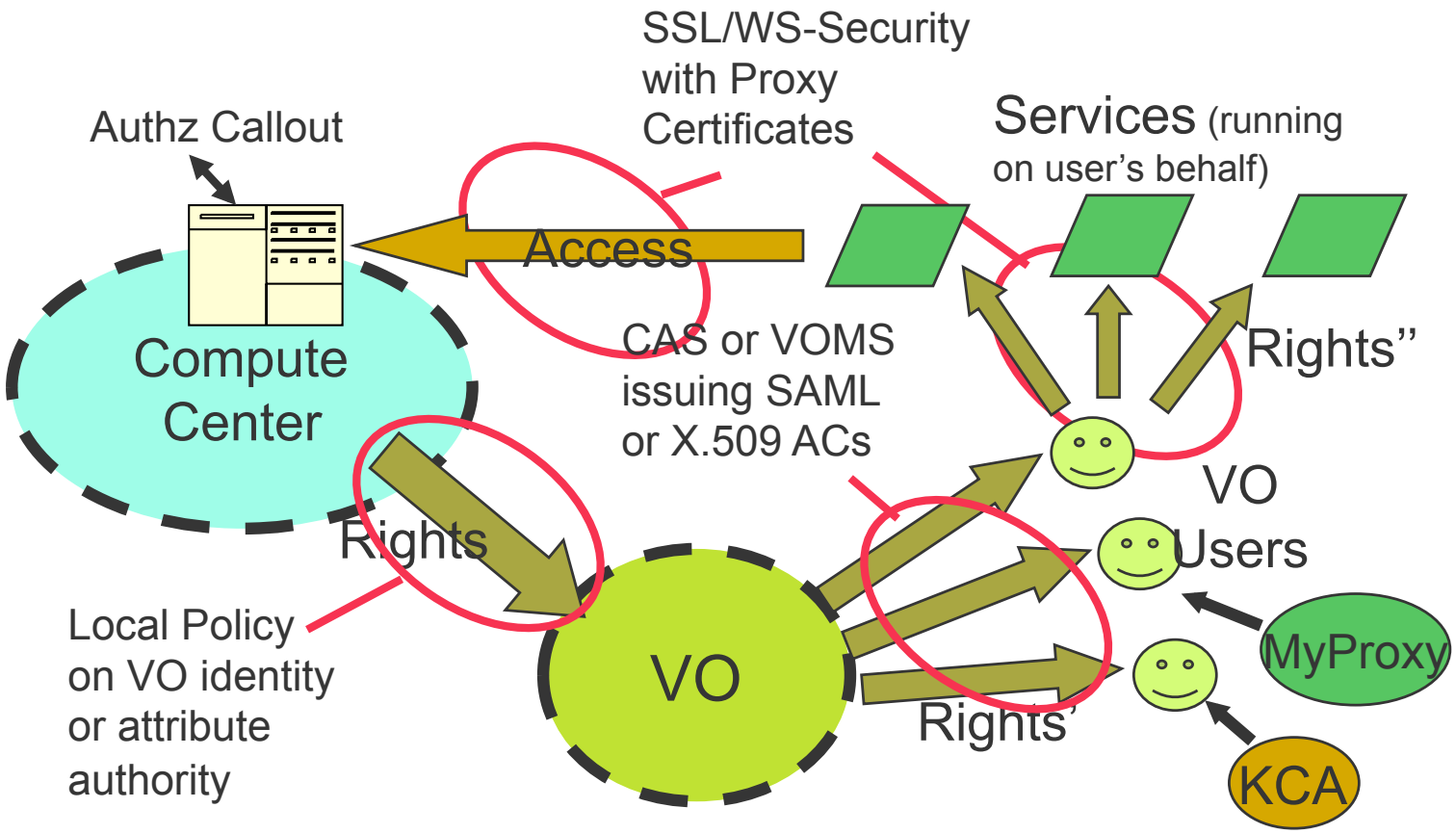
Use delegation to establish dynamic distributed system



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 –
01 Feb 2009

GSI implementation



Grids and VOs (1)

- Virtual organizations (VOs) are groups of Grid users (authenticated through digital certificates)
- VO Management Service (VOMS) serves as a central repository for user authorization information, providing support for sorting users into a general group hierarchy, keeping track of their roles,etc.
- VO Manager, according to VO policies and rules, authorizes authenticated users to become VO members

Grids and VOs (2)

- Resource centers (RCs) may support one or more VOs, and this is how users are authorized to use computing, storage and other Grid resources
- VOMS allows flexible approach to A&A on the Grid

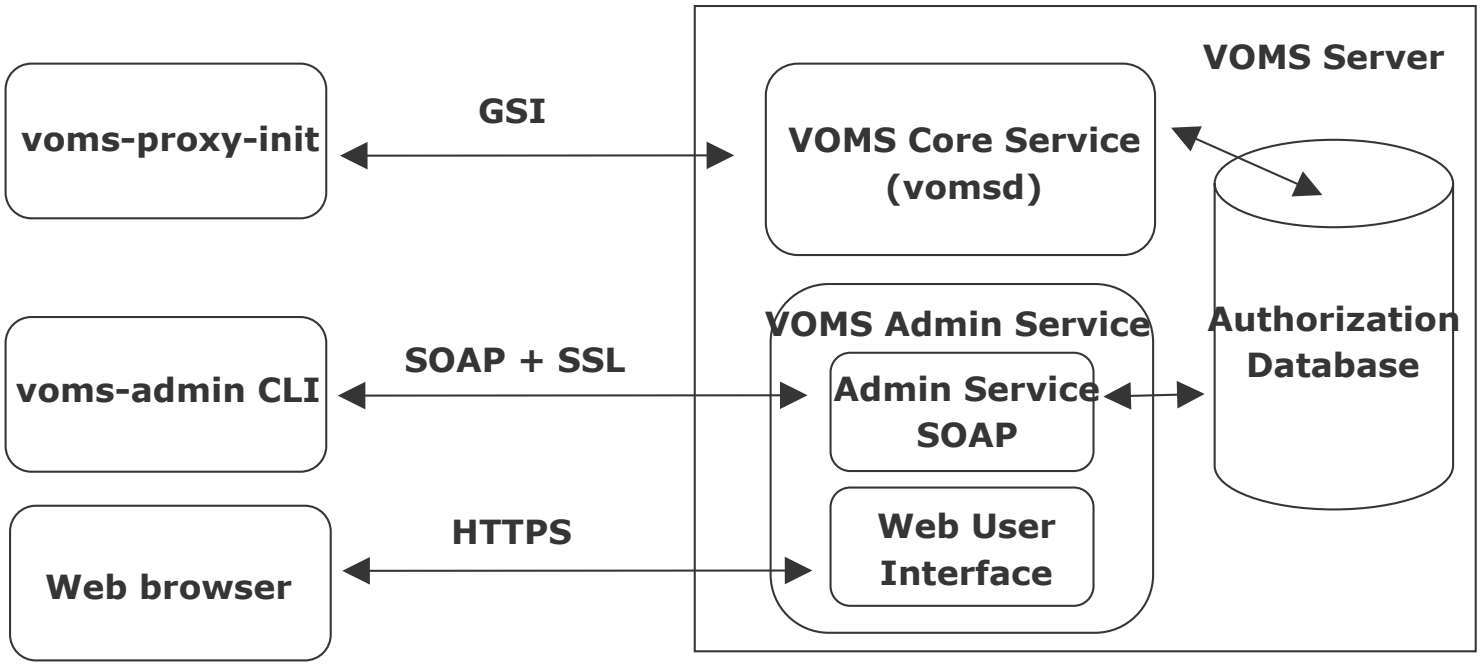
VOMS Ingredients (1)

- Attribute Certificates: AC is a PKI container, defined in RFC 3281, capable of containing a set of attributes tied to a specific identity. It is the system used by VOMS to issue its attributes.
- VOMS groups: /aegis/scl
- VOMS roles: /Role=VO-Admin
  - Roles can be defined for groups as well

VOMS Ingredients (2)

- FQAN (Fully Qualified Attribute Name) is a compact way to represent user's membership in a group, along with its role holdership, if any
  - Syntax: <groupname>/Role=<rolename>/Capability=NULL where the /Capability=NULL may be omitted, since it refers to a deprecated feature of VOMS
  - /aegis/scl/Role=NULL/Capability=NULL

VOMS Architecture



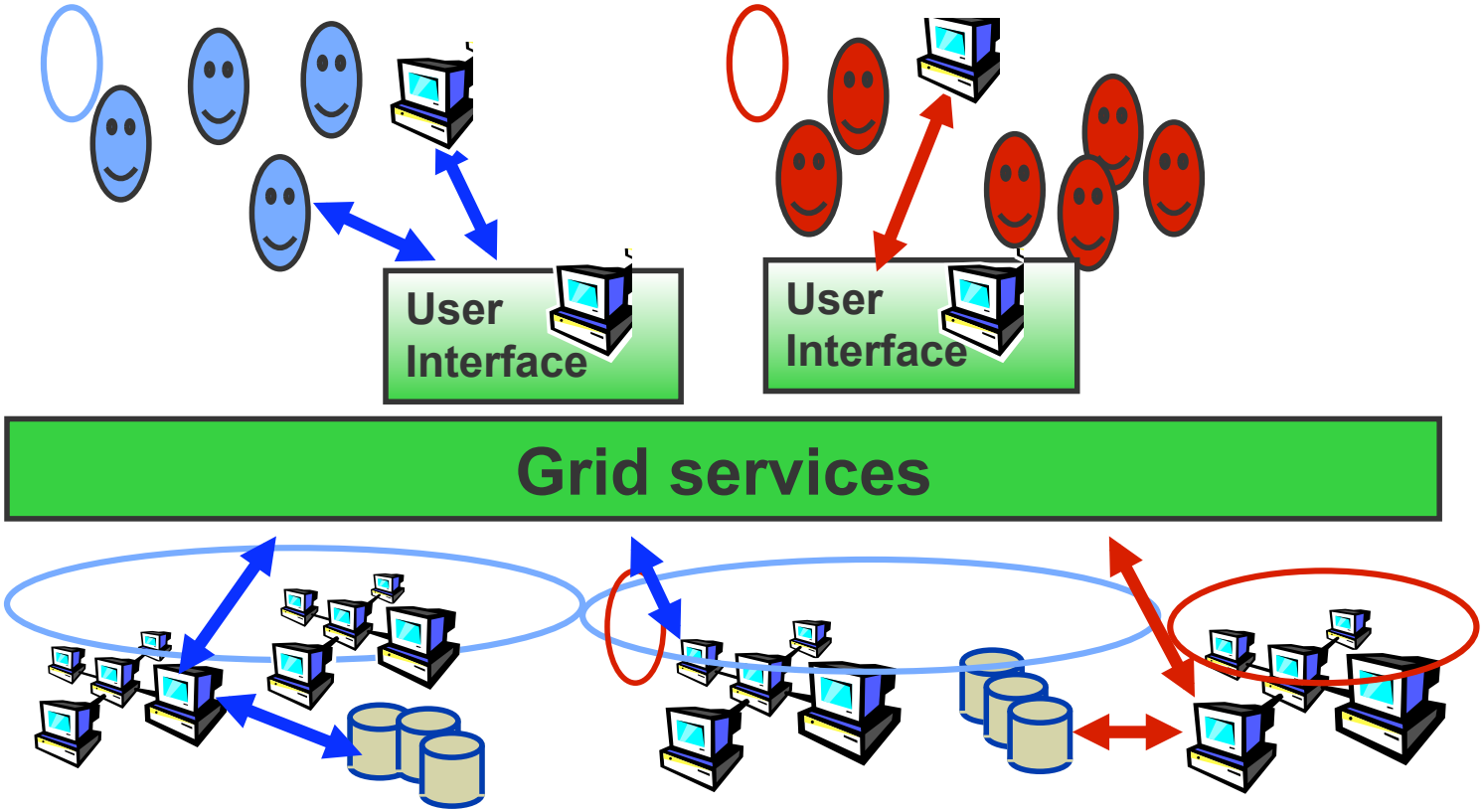
“Logging on” to the Grid

- To run programs, authenticate to Grid:

  voms-proxy-init –voms VONAME

  Enter PEM pass phrase: **********
- Creates a temporary, local, short-lived proxy credential for use by our computations
- Delegation = remote creation of a (second level) proxy credential, which allows remote process to authenticate on behalf of the user

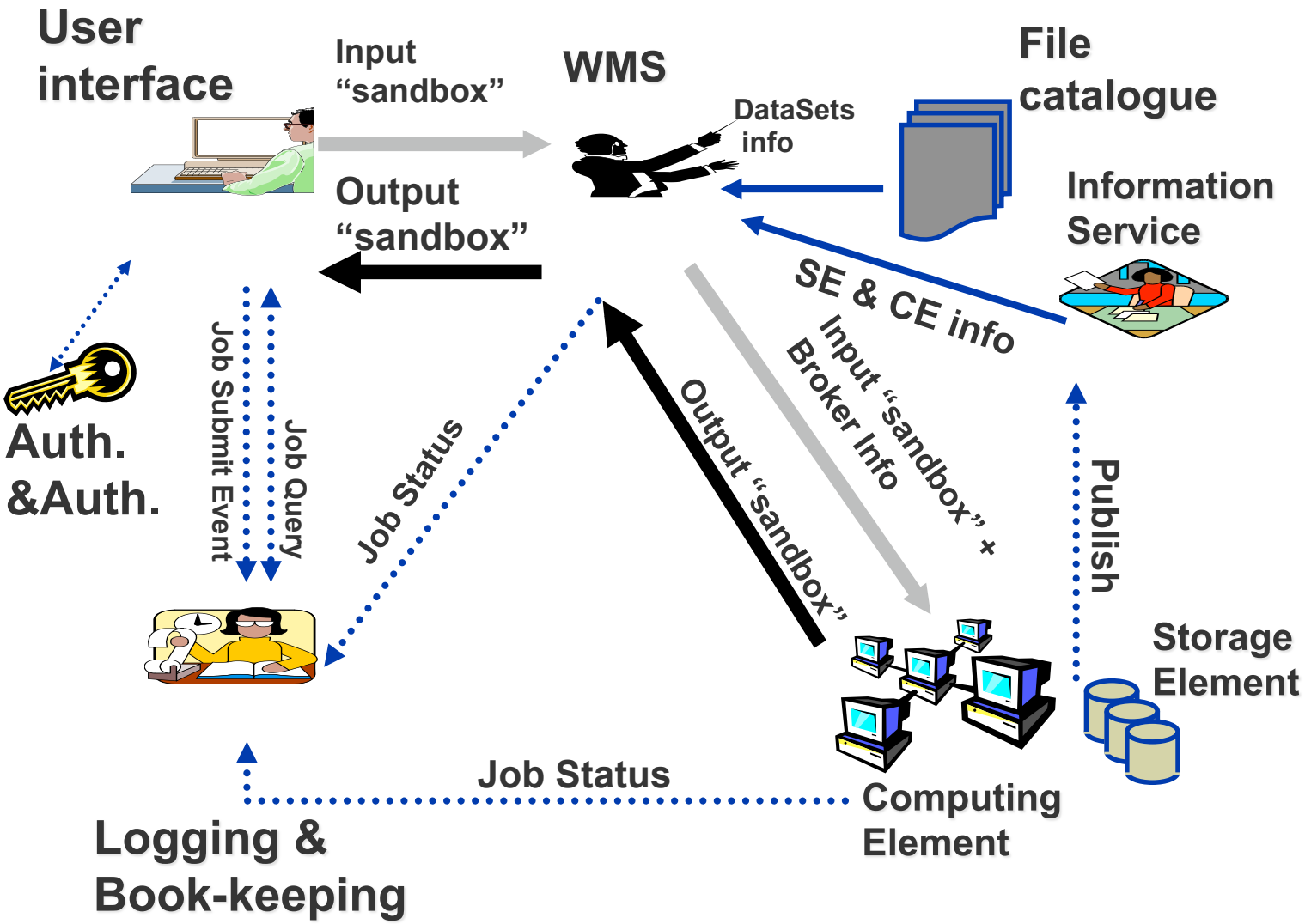
User view of the Grid



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 –
01 Feb 2009

What really happens



SCIENTIFIC COMPUTING LABORATORY

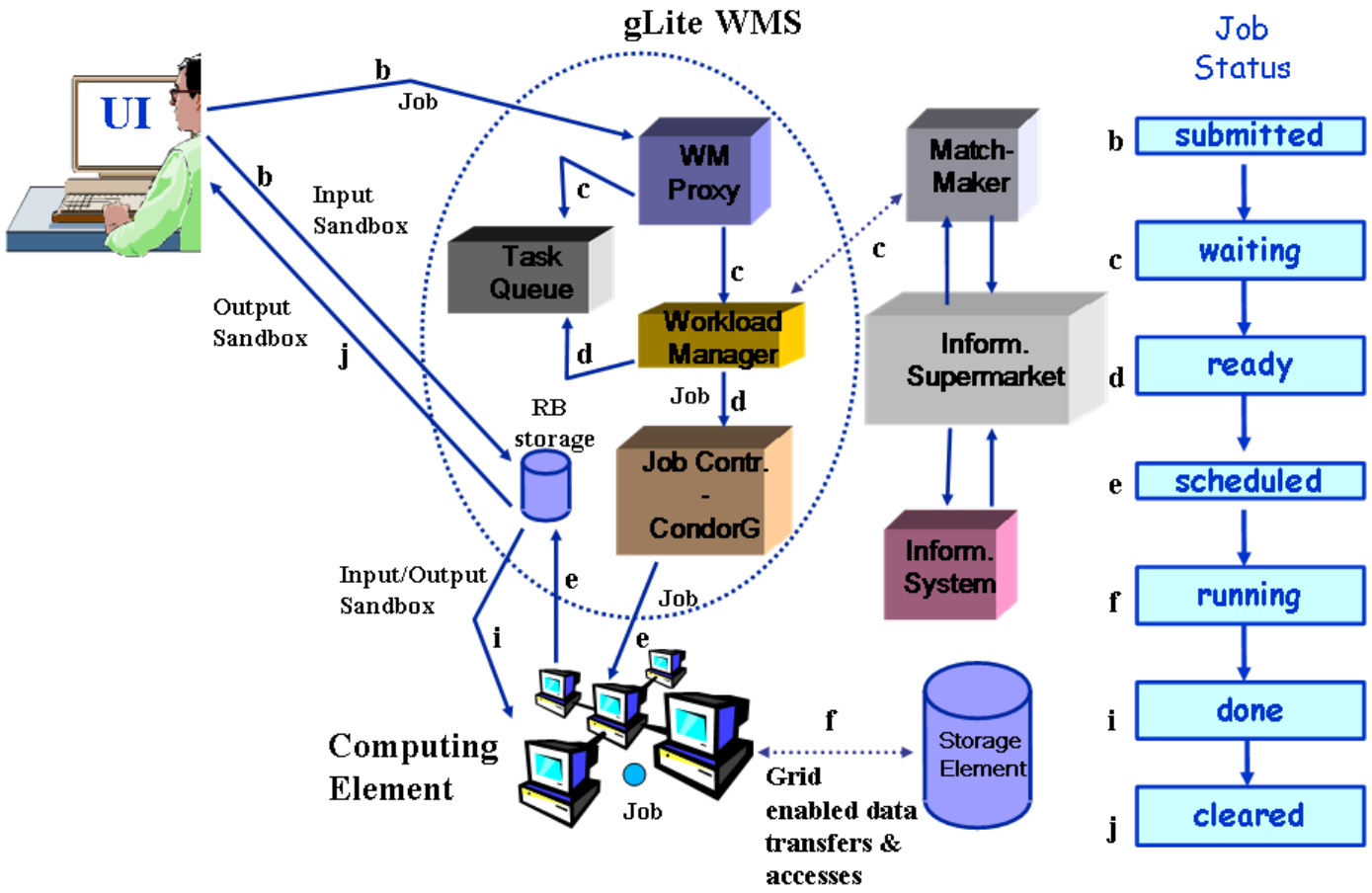
21 Jan 2009 – 01 Feb 2009

Workload Management System (WMS)

- Distributed scheduling
  - multiple UIs where you can submit your job
  - multiple WMSs from where the job can be sent to a CE
  - multiple CEs where the job can be put in a queuing system
- Distributed resource management
  - multiple information systems that monitor the state of the grid
  - Information from SE, CE, sites

SCIENTIFIC COMPUTING LABORATORY

WMS and job states



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 –
01 Feb 2009

Authentication and Authorization

- Authentication
  - User obtains certificate from CA
  - Connects to UI by ssh
  - Downloads certificate
  - Invokes Proxy certificate
  - Single logon – to UI - then Secure Socket Layer with proxy identifies user to other nodes
- Authorization - currently
  - User joins Virtual Organisation
  - VO negotiates access to Grid nodes and resources (CE, SE)
  - Authorization tested by CE, SE: VOMS (or grid-mapfile) maps user to local accounts

SCIENTIFIC COMPUTING LABORATORY

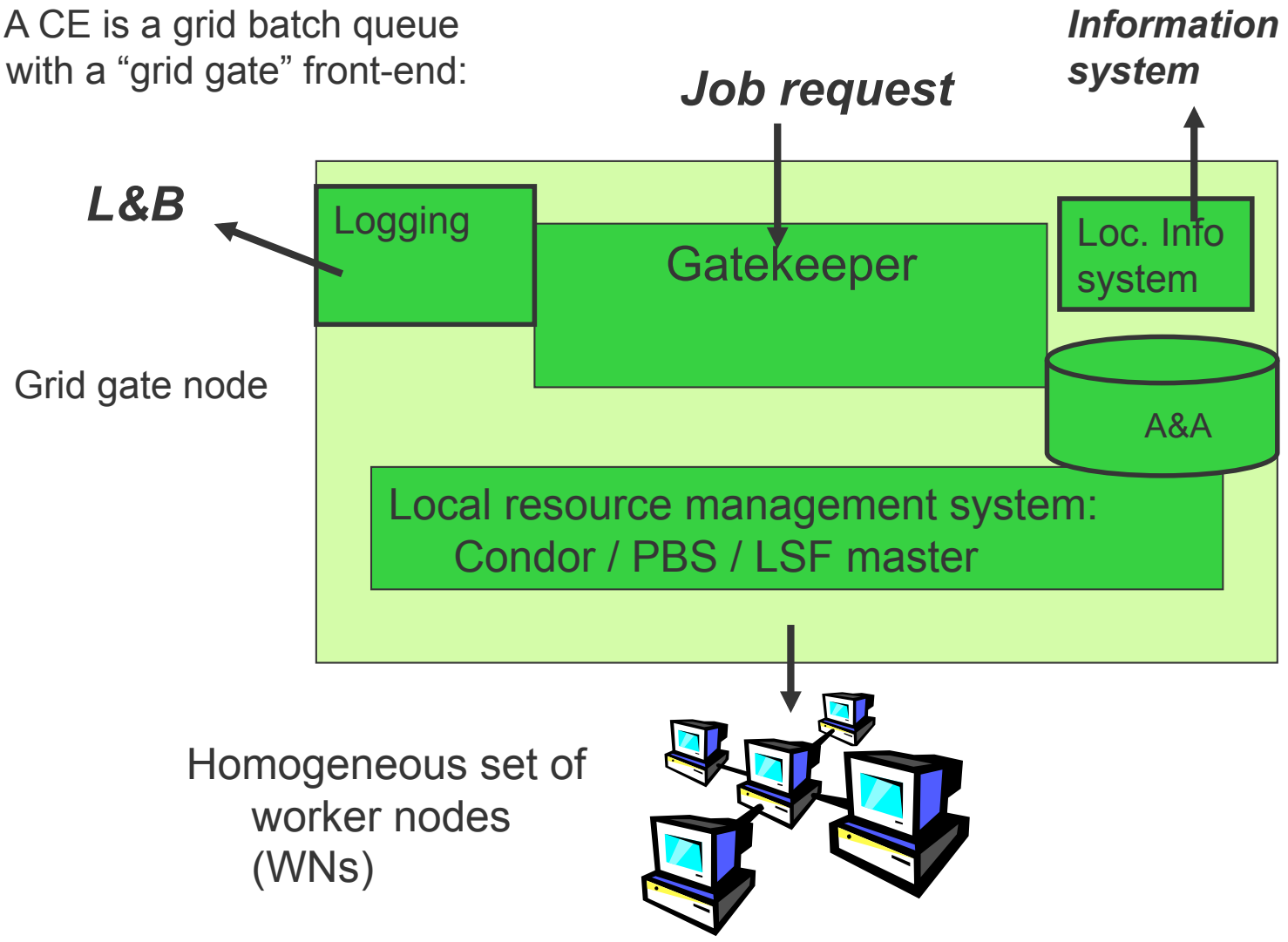
User Interface (UI)

- UI is the user's interface to the Grid - Command-line interface to
  - Proxy certificate
  - Job operations
    - To submit a job
    - Monitor its status
    - Retrieve output
  - Data operations
    - Upload file to SE
    - Create replica
    - Discover replicas
  - Other grid services
- To run a job user creates a JDL (Job Description Language) file

SCIENTIFIC COMPUTING LABORATORY

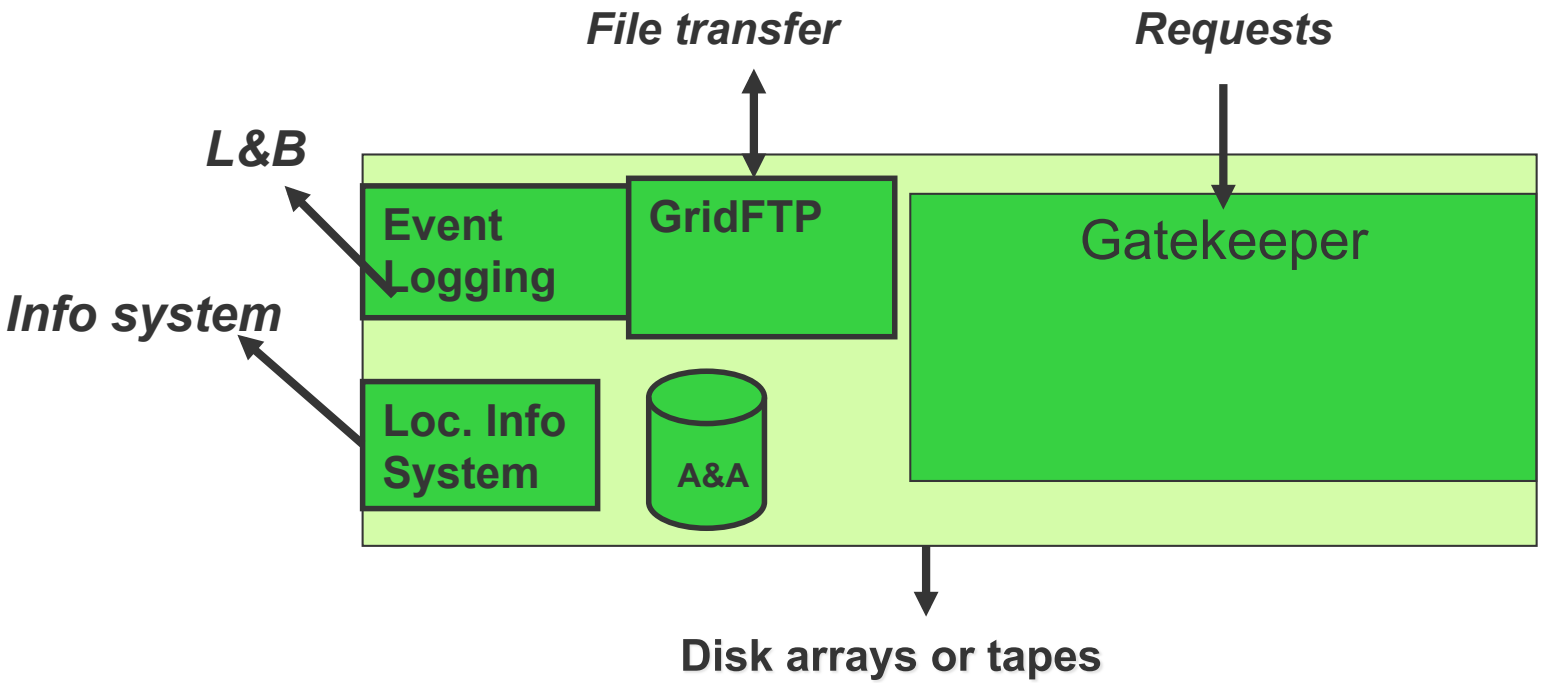
Computing Element (CE)

A CE is a grid batch queue with a “grid gate” front-end:



Storage Element (SE)

- Storage elements hold files: write once, read many
- Replica files can be held on different SE:
  - “close” to CE; share load on SE
- File Catalogue - what replicas exist for a file and where are they?



SCIENTIFIC COMPUTING LABORATORY

21 Jan 2009 – 01 Feb 2009

WMS (RB)

- Run the Workload Management System
  - To accept job submissions
  - Dispatch jobs to appropriate Compute Element (CE)
  - Allow users
    - To get information about their status
    - To retrieve their output
- A configuration file on each UI node determines which WMS node(s) will be used
- When a user submits a job, JDL options are to:
  - Specify CE
  - Allow RB to choose CE (using optional tags to define requirements)
  - Specify SE (then RB finds “nearest” appropriate CE, after interrogating File catalogue service)

Logging and Bookkeeping

- Who did what and when?
- What is happening to my job?
- Usually runs on the WMS node

Information System

- Receives periodic (~5 min) updates from CE, SE, etc.
- Used by WMS (RB) node to determine resources to be used by a job
- Currently BDII is used

Typical Grid site

- CE + batch system + set of WNs
- SE + set of disk nodes
- MON: accounting and R-GMA
- BDII_site: collects information about all elements
- Additional services (WMS+LB, PX, VOMS, etc.)

Grid in a nutshell

- Grid structure is complicated but hidden from end-users, enabling all the comfort they need
- Users just need to join the VO and obtain certificates: we already have some VOs at hand for you!
- Use of Grid is then just as easy as the use of a typical Linux cluster