eGee

Enabling Grids for E-sciencE

gLite Middleware Architecture

Dusan Vudragovic

dusan@phy.bg.ac.yu

Scientific Computing Laboratory

Institute of Physics Belgrade, Serbia

Nov. 18, 2008

www.eu-egee.org

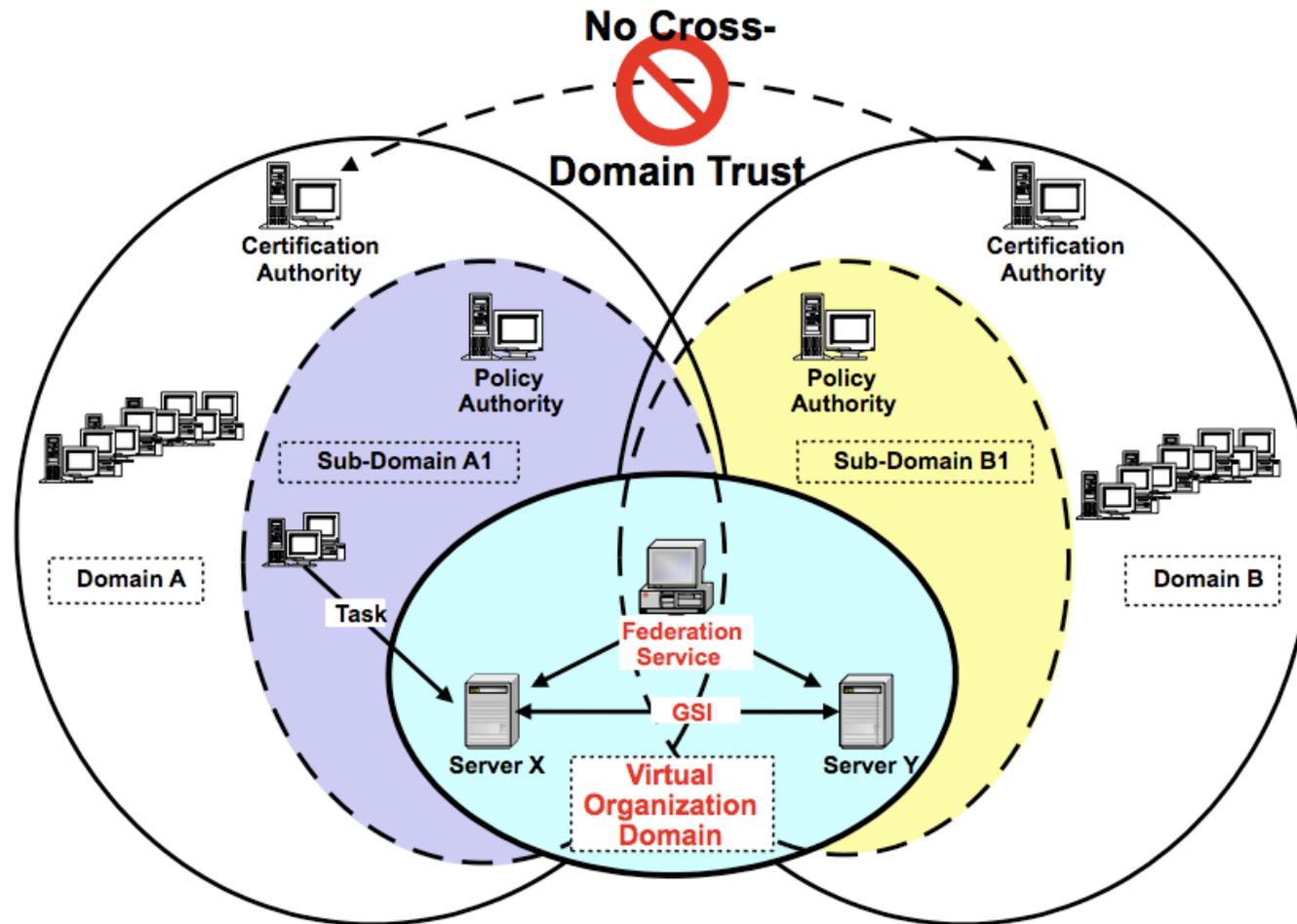# Set of basic Grid services

- **Job submission/management**

- **File transfer (individual, queued)**

- **Database access**

- **Data management (replication, metadata)**

- **Monitoring/Indexing system information**

Enabling Grids for E-sciencE

- ## Virtual Organization Membership Service
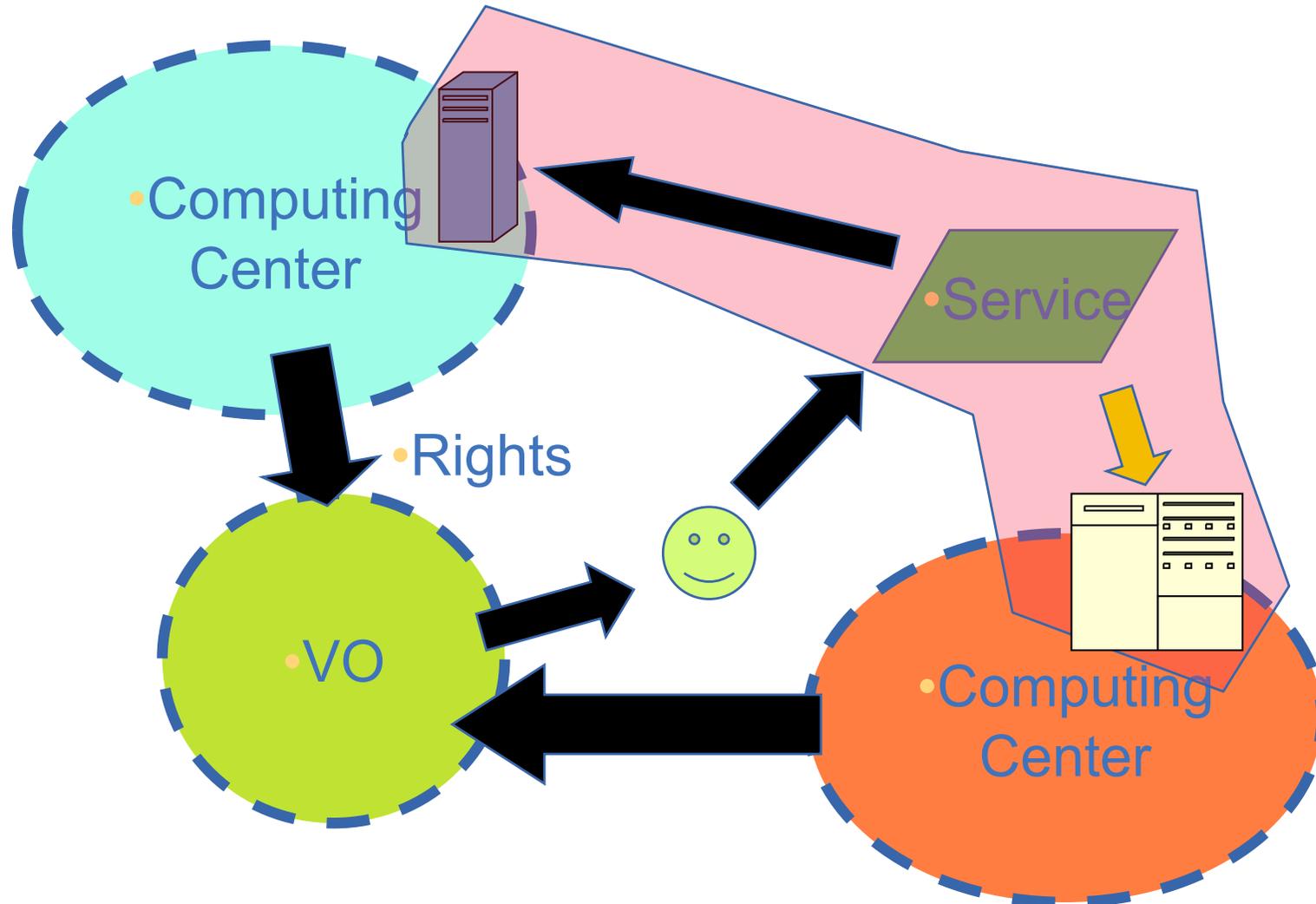
eGee

Enabling Grids for E-sciencE
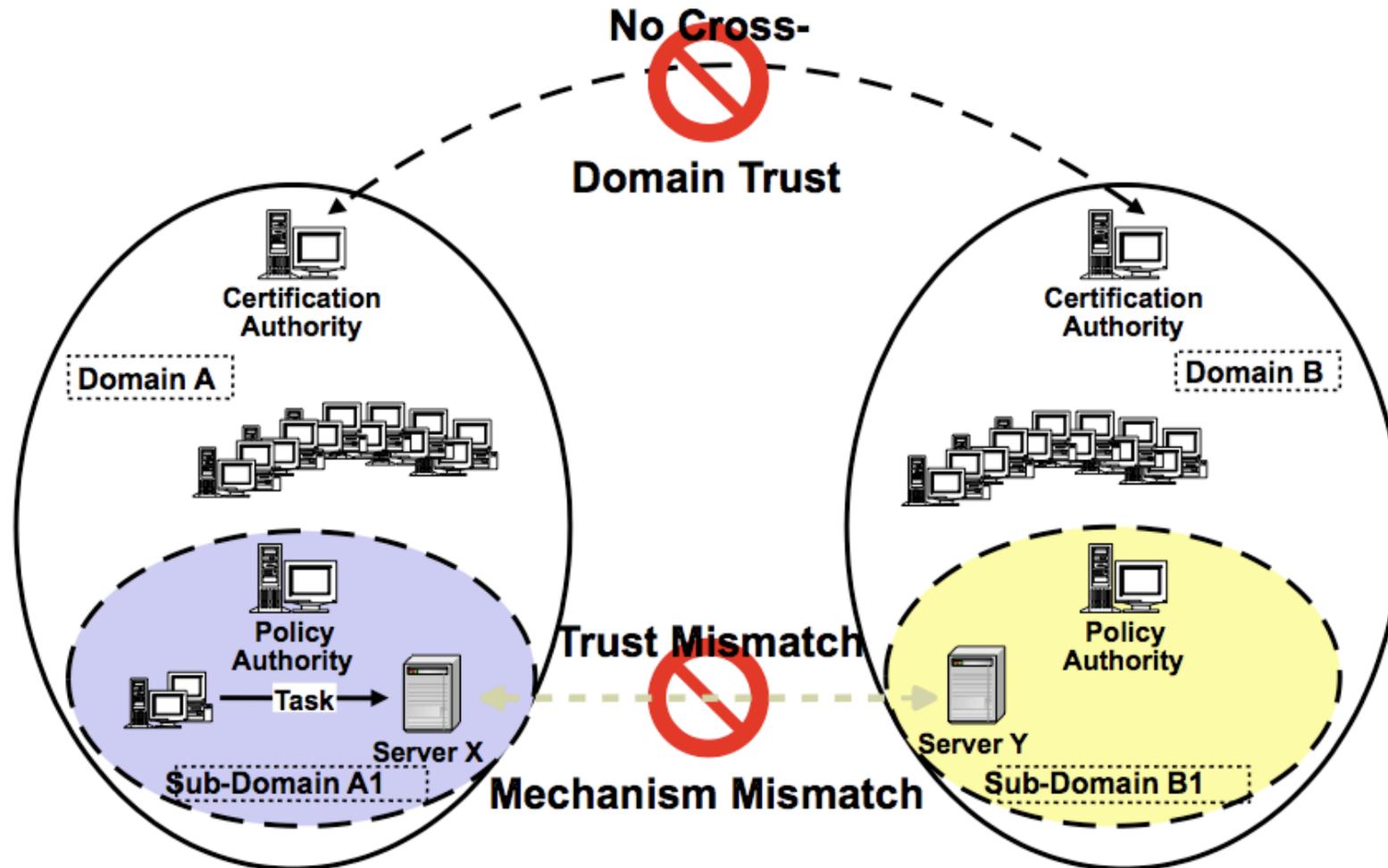
- Grid Security Infrastructure (GSI) enables secure authentication and communication over an open network
  - Public key encryption
  - Digital X.509 certificate
  - Secure Sockets Layer (SSL) communication protocol
- Certification Authority (CA)
- grid-mapfile mechanism
- LCAS/LCMAPS mechanism allow for a more detailed definition of user privilege
- VOMS server example

https://voms.phy.bg.ac.yu:8443/voms/aegis

Enabling Grids for E-sciencE

Enabling Grids for E-sciencE

- Use delegation to establish dynamic distributed system

egee

Enabling Grids for E-sciencE

- **Berkely Database Information Index**
  - Information Service (IS)
    - Globus Monitoring and Discovery Service (MDS)
      - *GLUE Schema*
      - *Lightweight Directory Access Protocol (LDAP)*
      - *Distinguished Name (DN)*
      - *Directory Information Tree (DIT)*
      - *Grid Resource Information Server (GRIS)*
    - Relational Grid Monitoring Architecture

- **GStat**

  `http://goc.grid.sinica.edu.tw/gstat/`

- **GOC DB**

  `https://goc.gridops.org/`
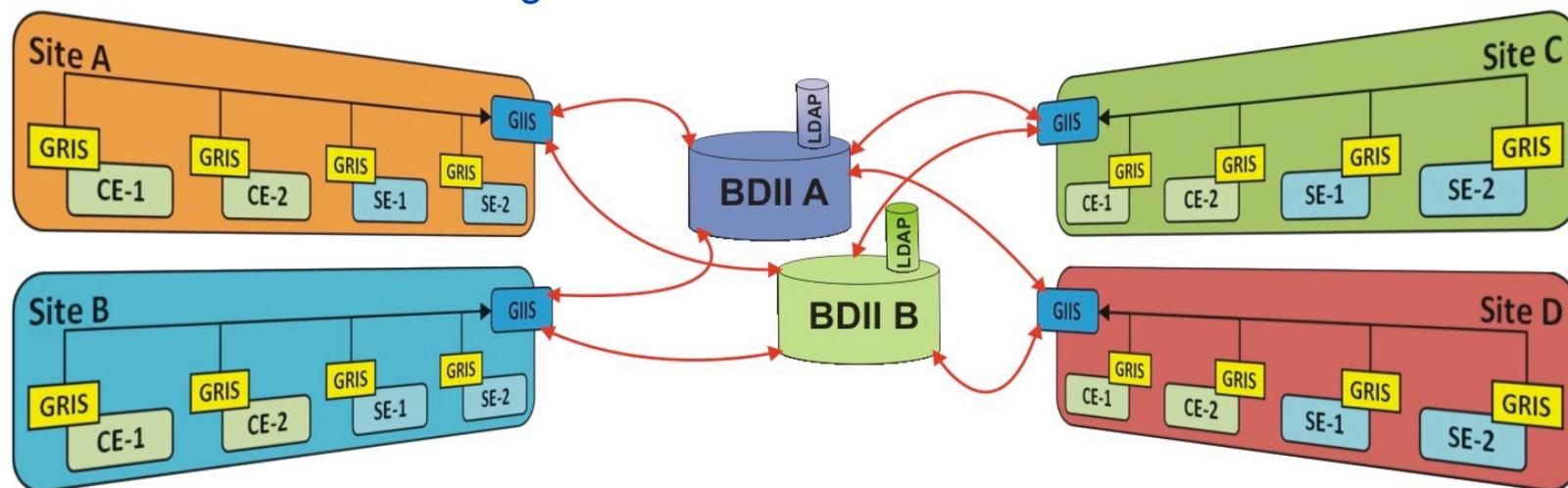
- **Relational Grid Monitoring Architecture**
  - Global distributed relational database
  - Model is more powerful than the LDAP-based one
  - Advance query operations
  - Architecture consists of three major components
    - Producers - provide the information
    - Consumers - request the information
    - Registry - mediates the communication between the Producers and the Consumers

- **EGEE Accounting Portal**

  `http://www3.egee.cesga.es/gridsite/accounting/CESGA/`
  `egee_view.php`

Enabling Grids for E-sciencE

- **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**

- **Logging and Bookkeeping service**
  - Who did what and when?
  - What's happening to my job?
  - Usually runs on WMS node

- **When a user submits a job, JDL options are to:**
  - Specify CE
  - Allow WMS to choose CE (using optional tags to define requirements)
  - Specify SE (then WMS finds "nearest" appropriate CE, after interrogating File catalogue service)

- ## Computing Element

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

Job request

Information system

L&B

Logging

Gatekeeper

Loc. Info system

A&A

Grid gate node

Local resource management system: Condor / PBS / LSF master

Homogeneous set of worker nodes

- **Storage Element**

File transfer

Requests

L&B

Info system

Event Logging

GridFTP

Gatekeeper

Loc. Info System

A&A

Disk arrays or tapes
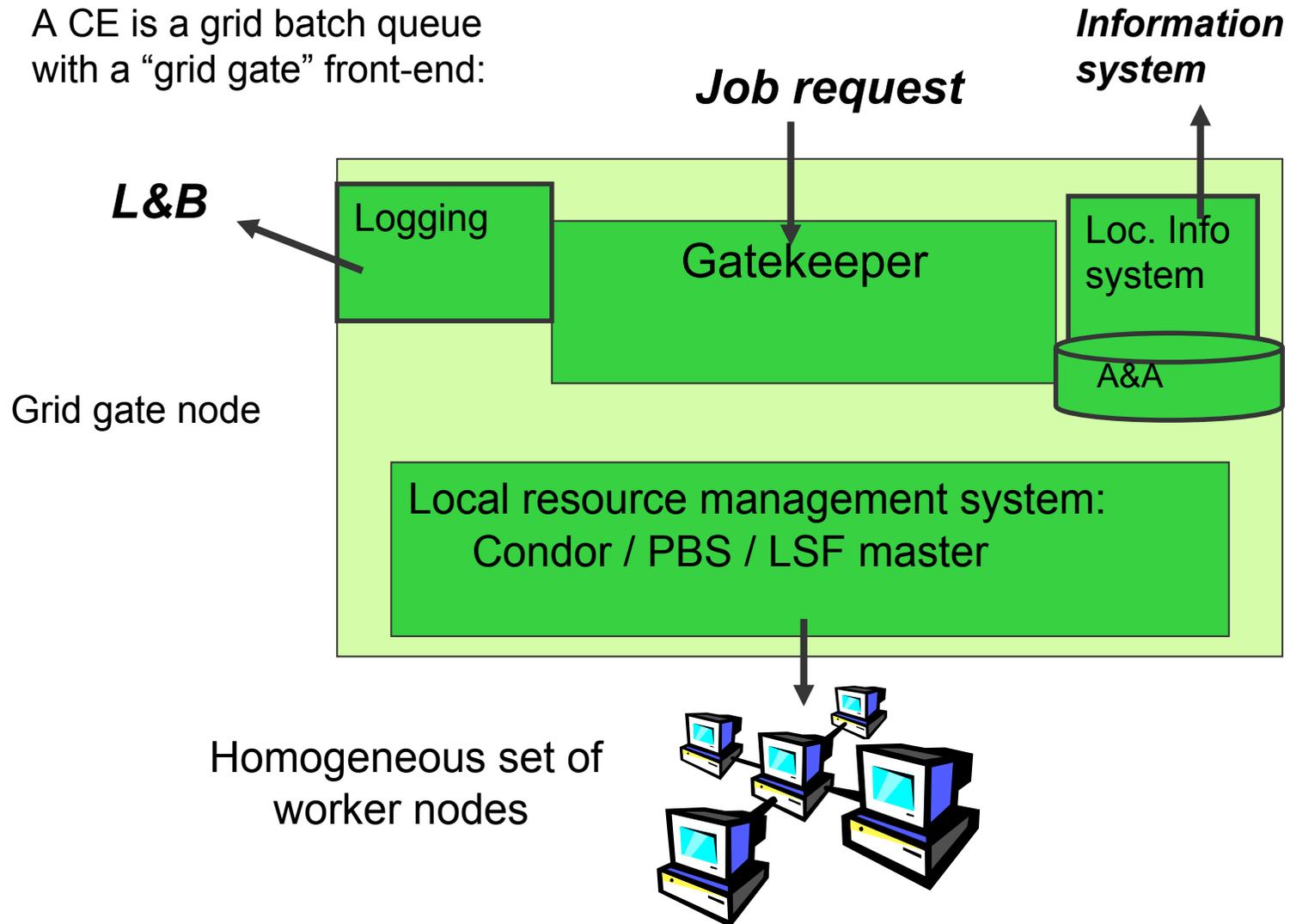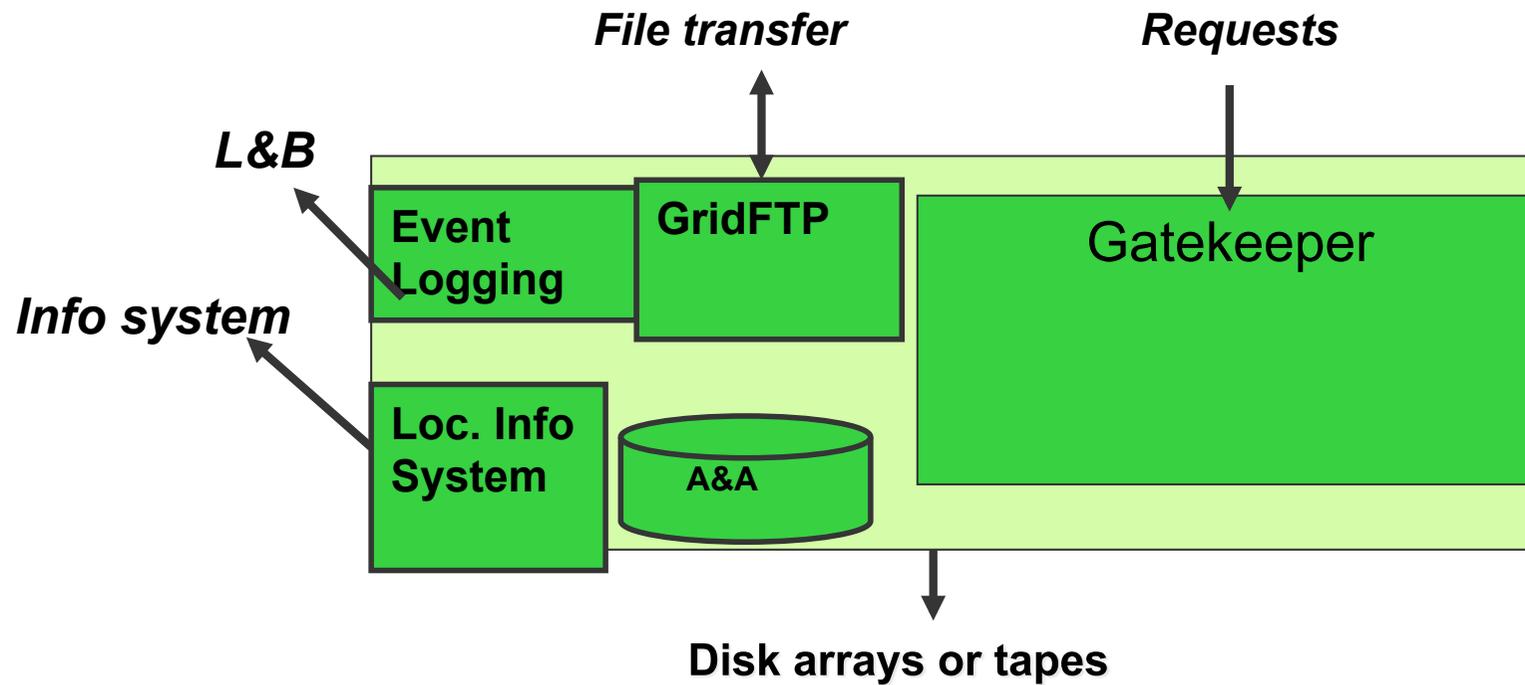
- **PX (MyProxy)**

- **FTS (File Transfer Service)**

- **LFC (Logical File Catalog)**

- **AMGA (ARDA Metadata Grid Application)**

egee

Enabling Grids for E-sciencE