



Enabling Grids for E-science

EGEE Middleware



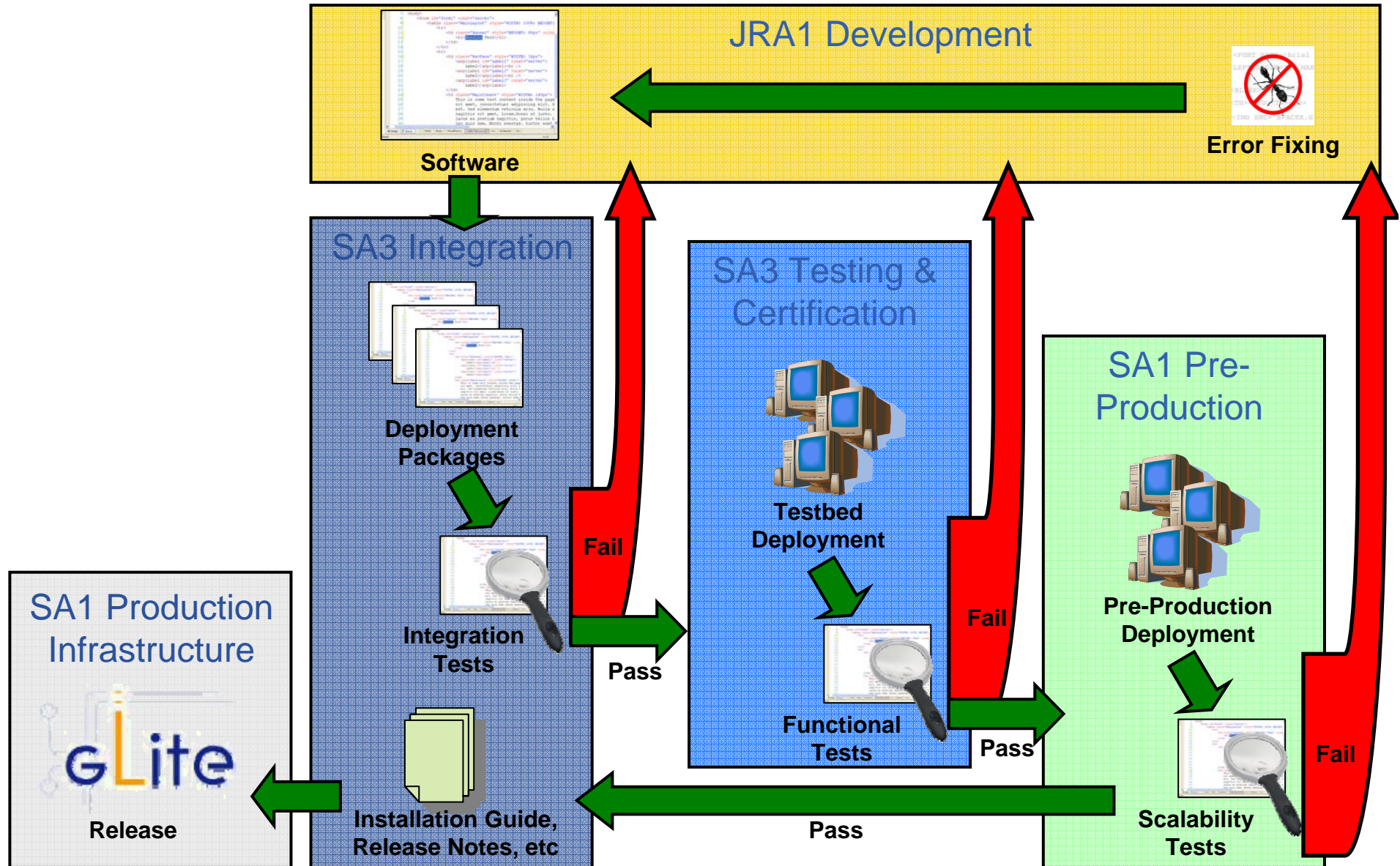
*Claudio Grandi
(INFN – Bologna)*

*EGEE Industry Day
April 27th, 2006
Paris, France*

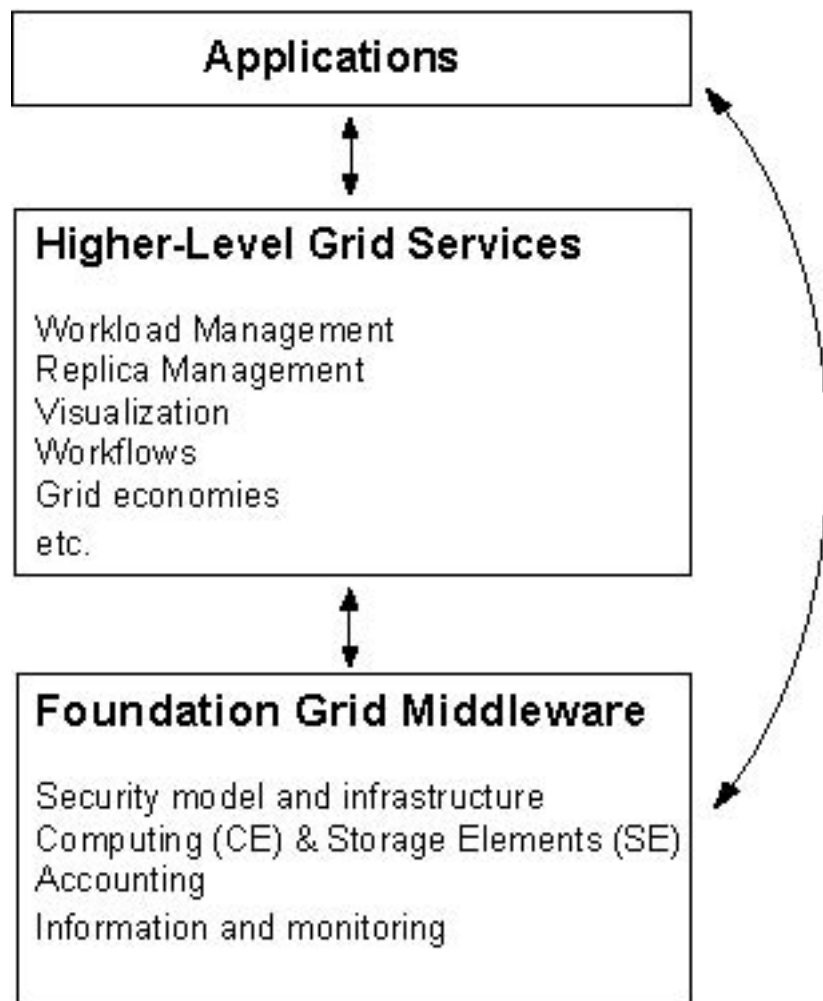
www.eu-egee.org
www.glite.org



-  **GLite Processes and Releases**
Lightweight Middleware for Grid Computing
-  **GLite Subsystems**
Lightweight Middleware for Grid Computing
 - Information System and Monitoring
 - Security Infrastructure
 - Workload Management
 - Data Management
- **Summary**

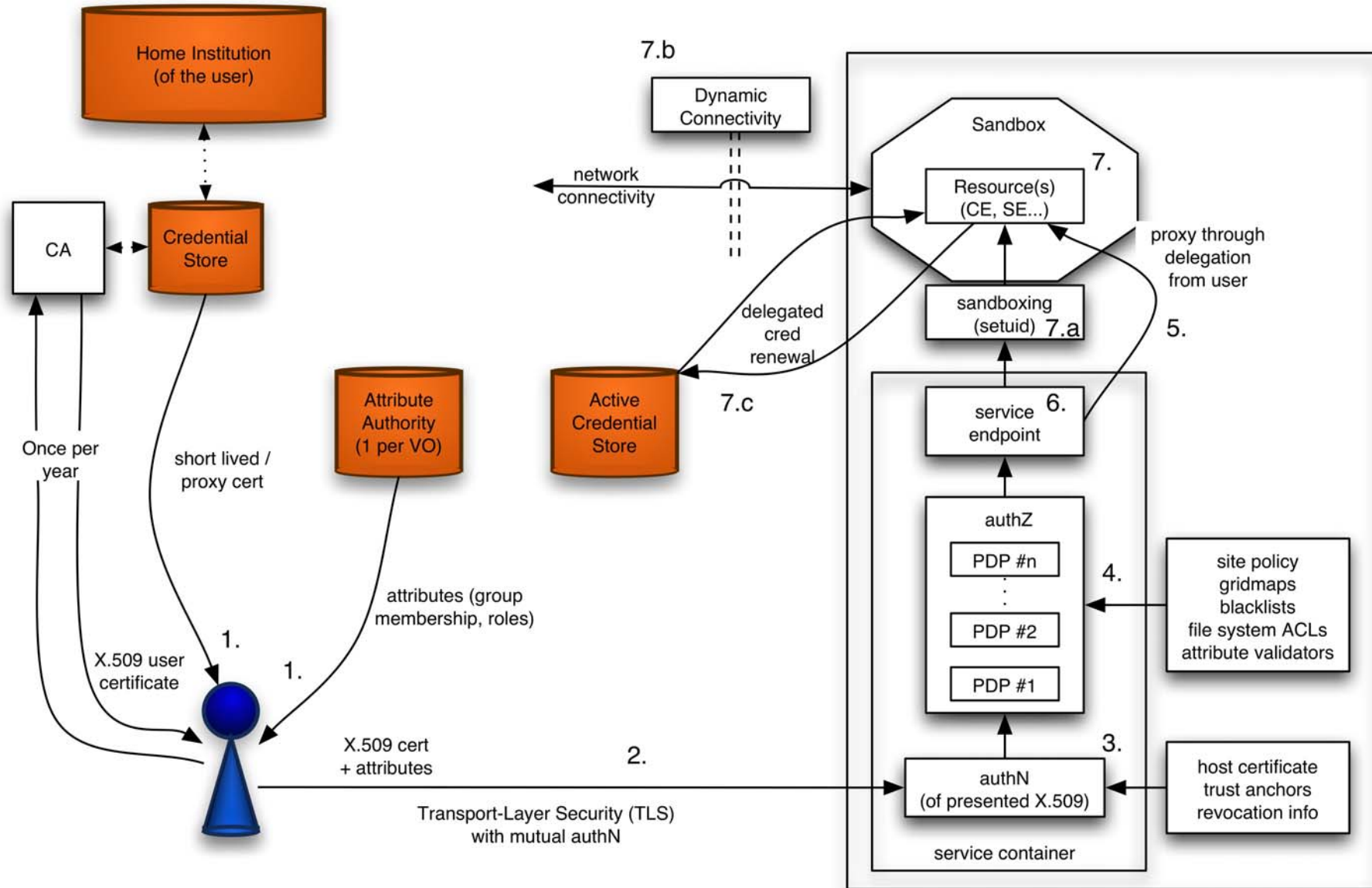


- **Converge from LCG and gLite to a single middleware stack called gLite. The first version is gLite 3**
 - Process controlled by the **Technical Coordination Group**
- **Deployment on the Production System will start in May**
 - Now on the Pre-Production System
- **After gLite 3.0:**
 - Continuous release of single components
 - as needed by users and as made available by developers
 - Major releases provide a “check-point”
 - in general in coincidence with major application challenges

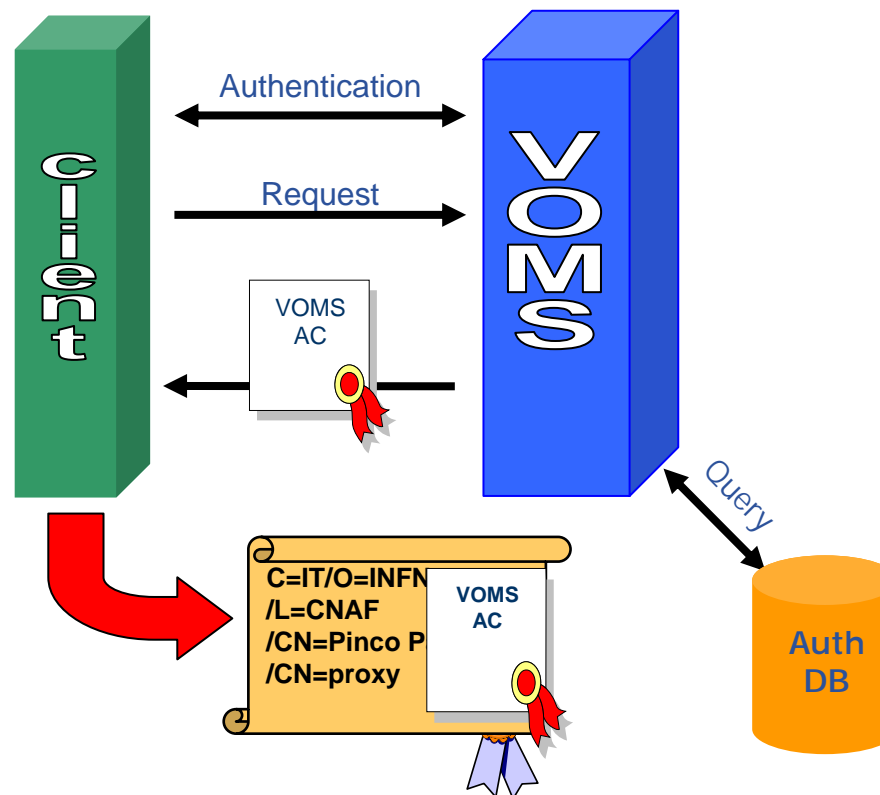


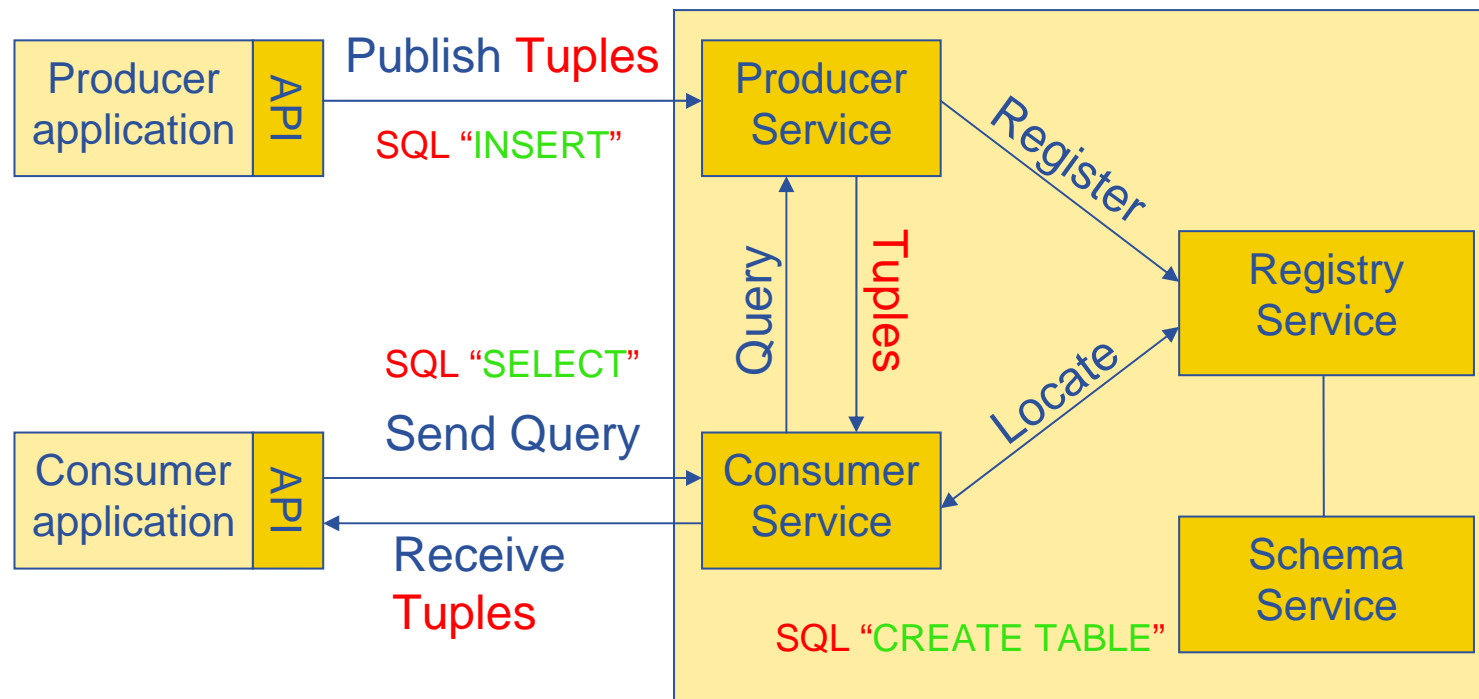
- **Foundation Grid Middleware is going to be deployed on the infrastructure**
 - should not assume the use of Higher-Level Grid Services
 - must be complete and robust
 - should allow interoperation with other major grid infrastructures
- **Higher-Level Grid Services may or may not be used by the applications**
 - should help them but not be mandatory

- **Authentication based on X.509 PKI infrastructure**
 - **Certificate Authorities** (CA) issue (long lived) **certificates** identifying individuals (much like a passport)
 - Commonly used in web browsers to authenticate to sites
 - Trust between CAs and sites is established (offline)
 - In order to reduce vulnerability, on the Grid user identification is done by using (short lived) **proxies** of their certificates
- **Proxies can**
 - Be **delegated** to a service such that it can act on the user's behalf
 - Include **additional attributes** (like VO information via the VO Membership Service VOMS)
 - Be stored in an **external proxy store** (myProxy)
 - Be **renewed** (in case they are about to expire)



- Bare certificates are not enough for defining user capabilities on the Grid
- Users belong to VO's, to groups inside a VO and may have special roles
- **VOMS** provides a way to add attributes to a certificate proxy:
 - mutual authentication of client and server
 - VOMS produces a signed **Attribute Certificate (AC)**
 - the client produces a new proxy that contains the attributes
- The attributes are used to provide the user with additional capabilities according to the VO policies.

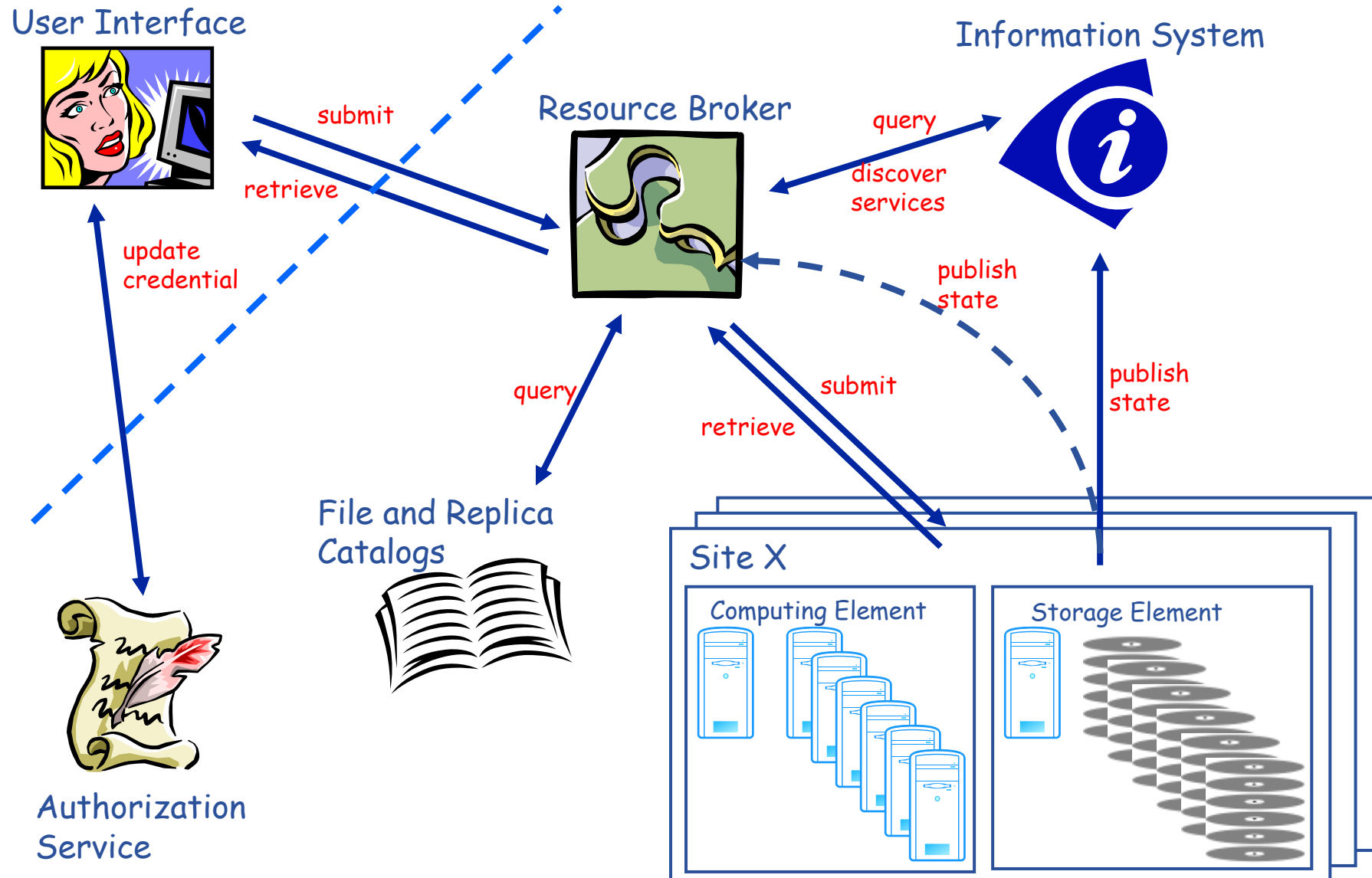




- The Relational Grid Monitoring Architecture (R-GMA) provides a uniform method to access and publish both information and monitoring data.
- From a user's perspective, an R-GMA installation currently appears similar to a single relational database.
- Relational implementation of the GGF's Grid Monitoring Architecture (GMA)

- The gLite **Service Discovery** provides a standard set of methods for locating Grid services
- hides underlying information system
- plug-ins for R-GMA, BDII and XML files (others could be developed if required)
- API available for Java and C/C++
- command line version also available
- Used by WMS and Data Management clients

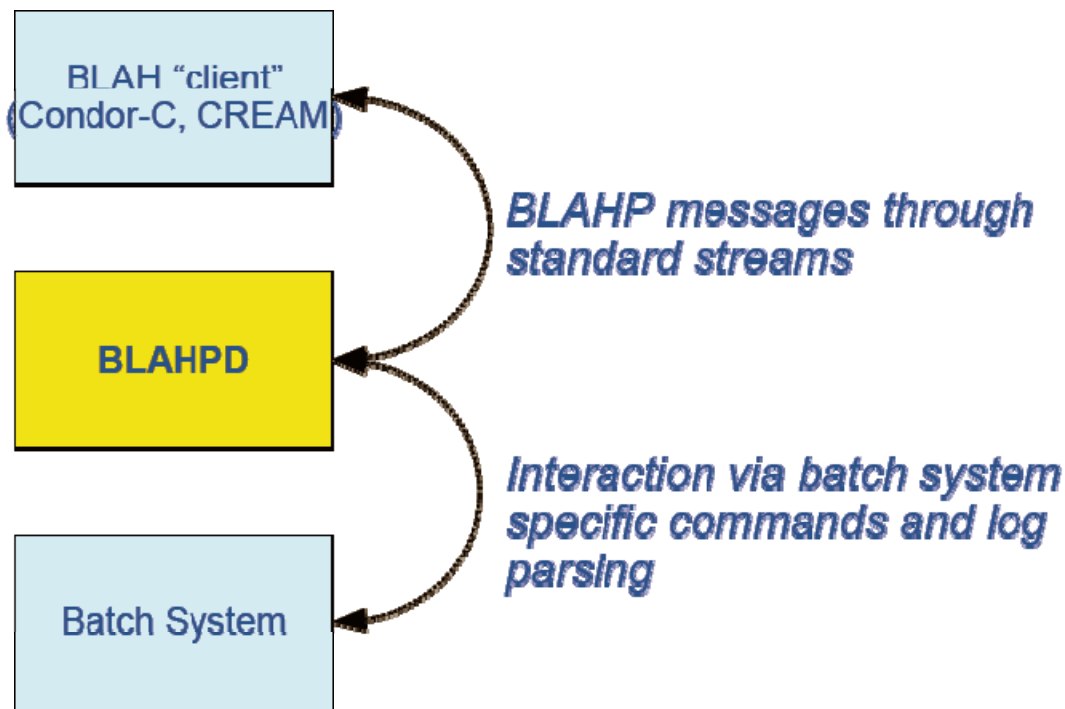
- Production Services uses **BDII** (by the LCG project) as the Information System
 - based on LDAP



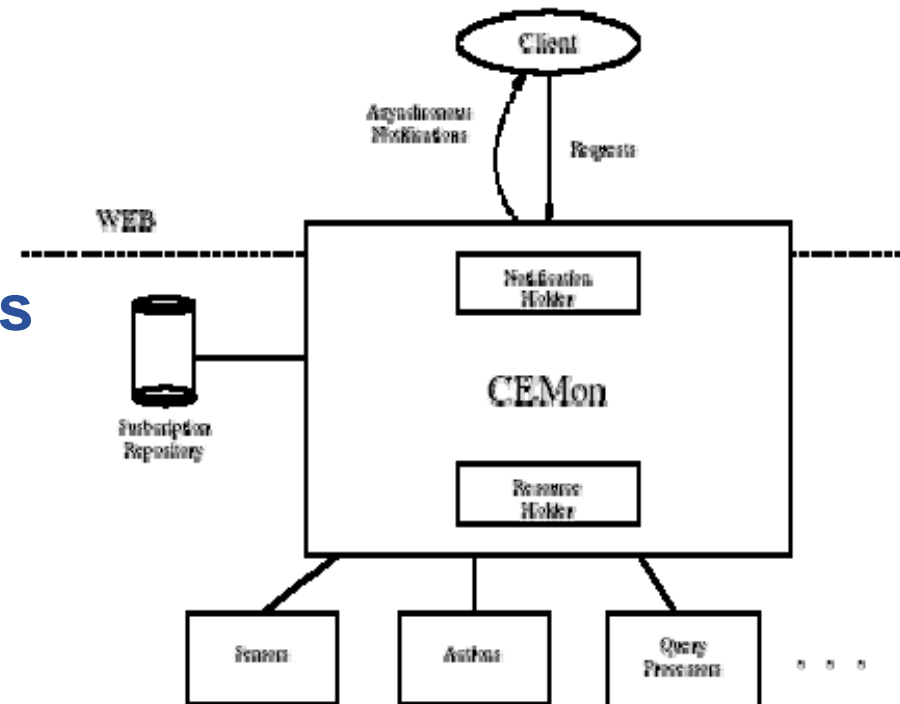
- **Service representing a computing resource**
 - **LCG-CE**: GT2 GRAM (to be dismissed when other CEs prove reliability)
 - **gLite CE** (Condor-C GSI enabled)
 - **CREAM**: new lightweight web service CE *(Not in gLite 3 release)*

- **Batch Local ASCII Helper (BLAH)**

- Interface between the CE and the Local Resource ManagerSystem
- submit, cancel and query
- Development work on *information pass-through*
 - parameters from the user to the batch system, e.g.
 - selecting machines with enough memory
 - reserving an appropriate number of nodes for MPI jobs

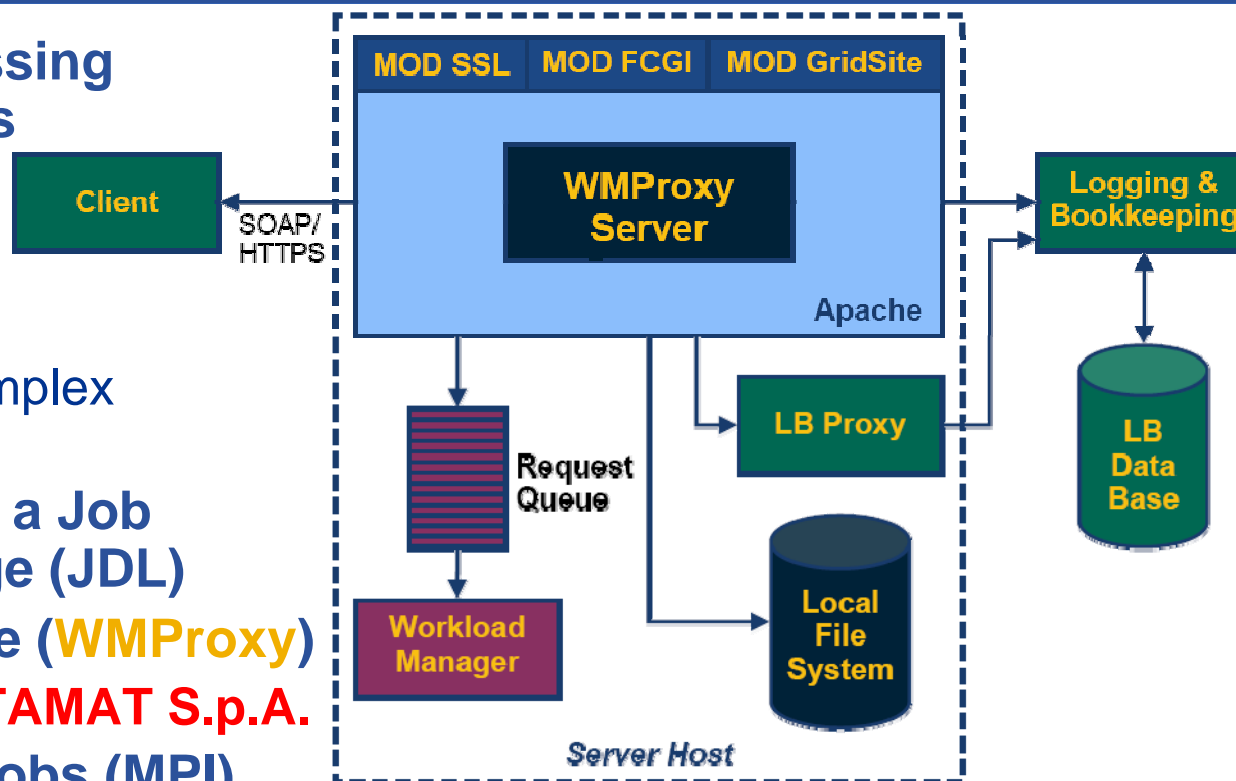


- **Web service to publish status of a computing resource to clients**
 - WMS or individual clients
- **Supports both synchronous queries or asynchronous notification of events**
 - Clients subscribe and are notified according to user defined policies
 - on job status
 - on CE characteristics and status
 - May be used to pull jobs
- **In gLite 3 CEMon will be available but the baseline is that the WMS queries the bdl**

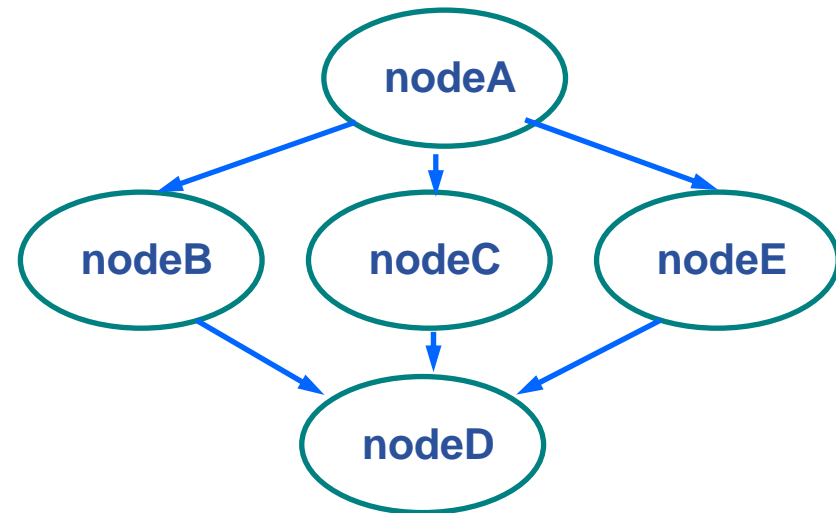


- **Logging and Bookkeeping service**
 - Tracks jobs during their lifetime (in terms of events)
 - L&B Proxy provides faster, synchronous and more efficient access to L&B services to Workload Management Services
 - Support for “CE reputability ranking”
 - Maintains recent statistics of job failures at CE’s
 - Feeds back to WMS to aid planning
 - Working on inclusion of L&B in the VDT
- **Job Provenance (*not in gLite 3 release*)**
 - Long term job information storage
 - Useful for debugging, post-mortem analysis, comparison of job executions in different environments
 - Useful for statistical analysis

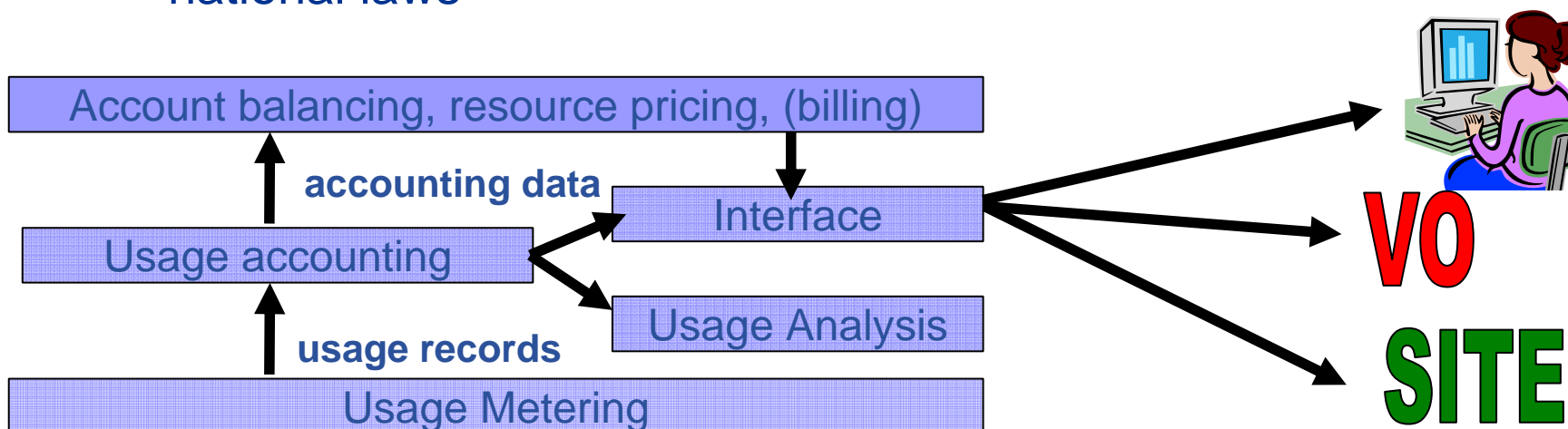
- **Helps the user accessing computing resources**
 - resource brokering
 - management of input and output
 - management of complex workflows
- **Job specification via a Job Description Language (JDL)**
- **Web service interface (WMPProxy)**
 - **Developed by DATAMAT S.p.A.**
- **Support for parallel jobs (MPI)**
 - even without a shared file system
- **Support collection of information from many sources**
 - CEMon, bdll, R-GMA
- **Support for different Data management interfaces**
 - DLI and StorageIndex
- **Job File Perusal: file peeking during job execution**



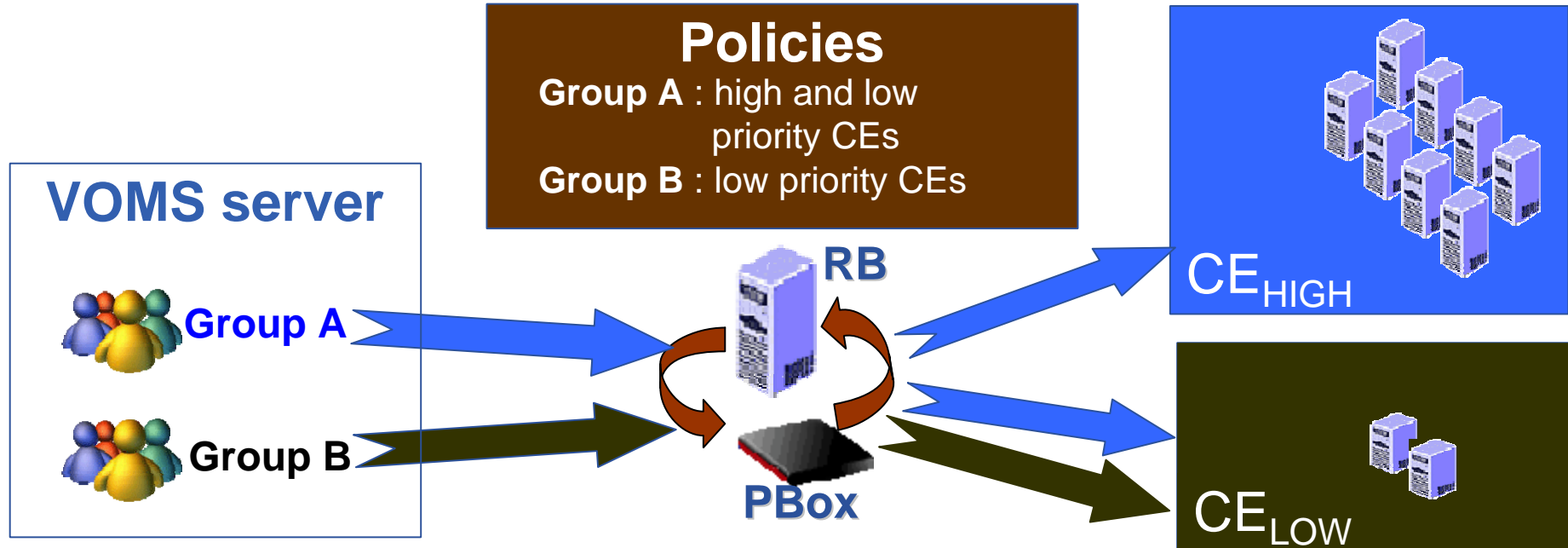
- **Direct Acyclic Graph (DAG)** is a set of jobs where the input, output, or execution of one or more jobs depends on one or more other jobs
- **A Collection** is a group of jobs with no dependencies
 - basically a collection of JDL's
- **A Parametric job** is a job having one or more attributes in the JDL that vary their values according to parameters
- **Using compound jobs** it is possible to have one shot submission of a (possibly very large, up to thousands) group of jobs
 - Submission time reduction
 - Single call to WMPProxy server
 - Single Authentication and Authorization process
 - Sharing of files between jobs
 - Availability of both a single Job Id to manage the group as a whole and an Id for each single job in the group



- **DGAS (Not in gLite 3 release):**
 - Collects Grid accounting information (*User, JobId, user VO, VOMS FQAN(role,capabilities), system usage (cpuTime, ...), ...*)
 - Secure storage and transfer of accounting records
- **APEL:**
 - propagate and display accounting information for infrastructure monitoring
- **Compliant with privacy requirements**
 - Sites may keep personal information local if requested by national laws

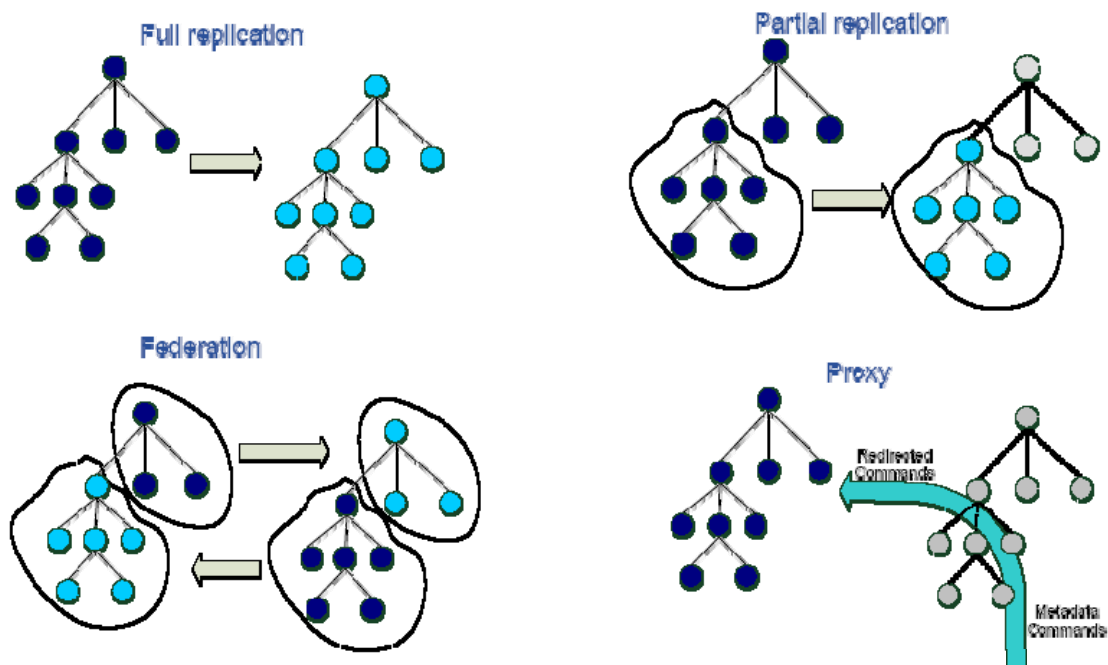


- **GPBOX (Not in gLite 3 release):**
 - Interface to define, store and propagate fine-grained VO policies
 - based on VOMS groups and roles
 - enforcement of policies done at sites – sites may accept/reject policies
 - May be interfaced to dynamic sources of information
 - e.g. an accounting system to provide fair share
 - Standards Compliant (RBAC, XACML, GSI)

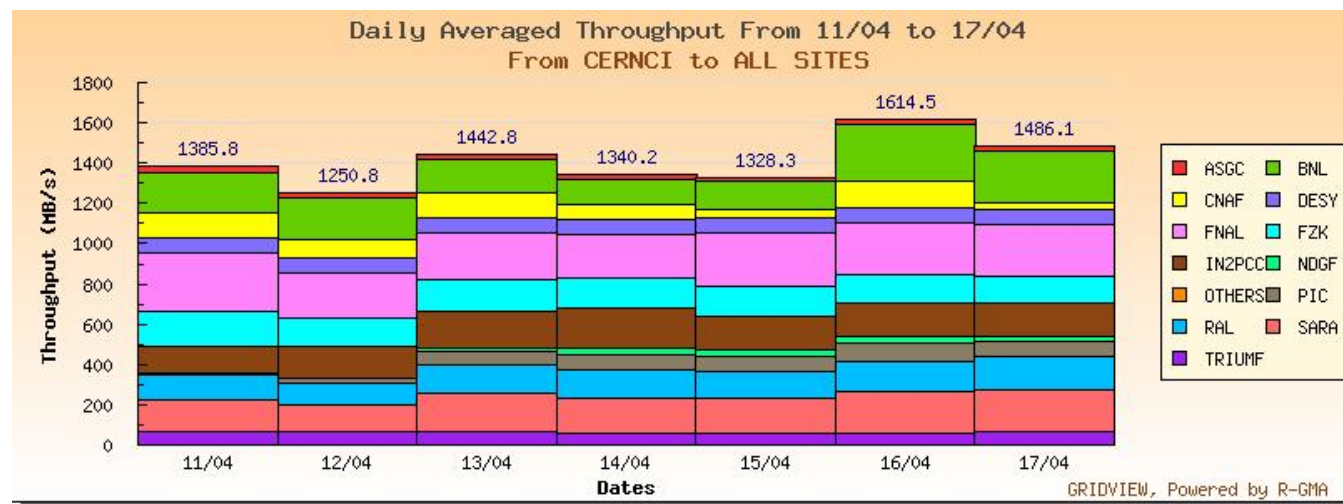
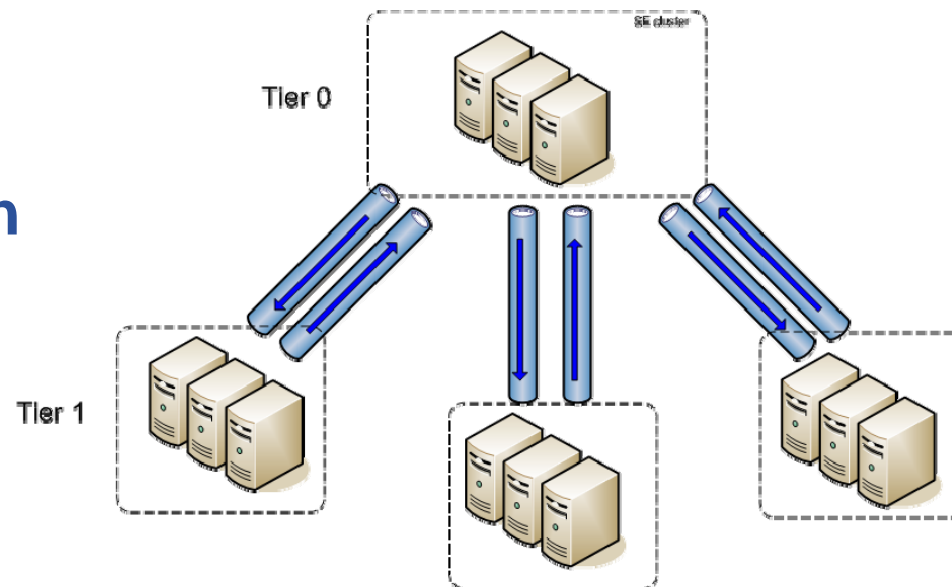


- **Storage Element**
 - Common interface: SRMv1, migrating to SRMv2
 - Various implementation from LCG and other external projects
 - disk-based: DPM, dCache
 - tape-based: Castor, dCache
 - Support for ACLs in DPM, in future in Castor and dCache
- **Posix-like file access:**
 - Grid File Access Layer (GFAL) by LCG
 - Support for ACL in the SRM layer
 - gLite I/O (to be dismissed when Biomed are fine with GFAL)
 - Support for ACLs from Fireman data catalog
- **Hydra keystore (*Limited support in gLite 3*):**
 - store keys for data encryption
 - 3 instances: at least 2 need to be available for decryption
 - Demonstrated at the SRM/DICOM demo at EGEE'04 conference in Pisa (system with gLite I/O, Fireman and AMGA)

- **Data Catalogs**
 - LFC from LCG
 - Fireman (to be dismissed when Biomed are fine with LFC)
- **AMGA Metadata Catalog (*Limited support in gLite 3*):**
 - Joint JRA1-NA4 (ARDA) development
 - Authentication based on Password, X509 Cert, Grid Proxy
 - Posix-ACLs and Unix permissions for entries and collections
 - Compatible with VOMS groups and roles
 - Support for replication
 - also via LCG-3D for Oracle backend only



- **Reliable, scalable and customizable file transfer**
- **Manages transfers through channels**
 - mono-directional network pipes between two sites
- **Support for different user and administrative roles**
- **Web service interface**
- **Sustained file transfer record (1.6 GB/s)**



- **EGEE provides a complete middleware stack**
 - security infrastructure
 - information system and monitoring
 - workload management
 - data management
- **Developed according to a well defined process**
 - Controlled by the EGEE Technical Coordination Group
- **gLite 3 will be available on the production infrastructure in June**
- **Development is continuing to provide increased robustness, usability and functionality**



Lightweight Middleware for
Grid Computing

www.glite.org