

Status and Plans of gLite Middleware

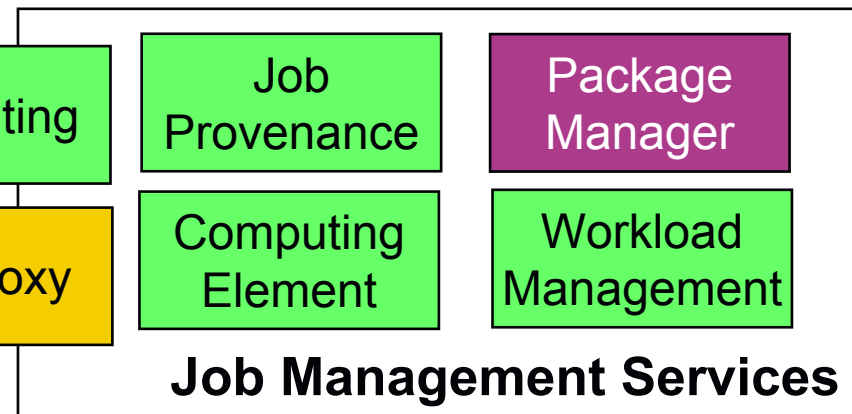
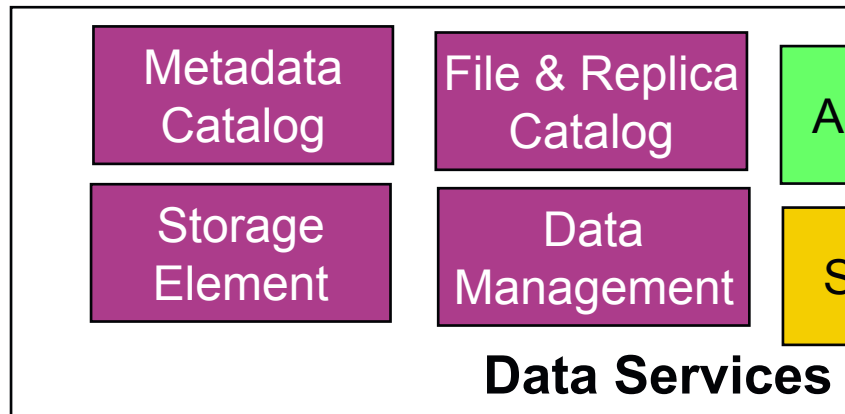
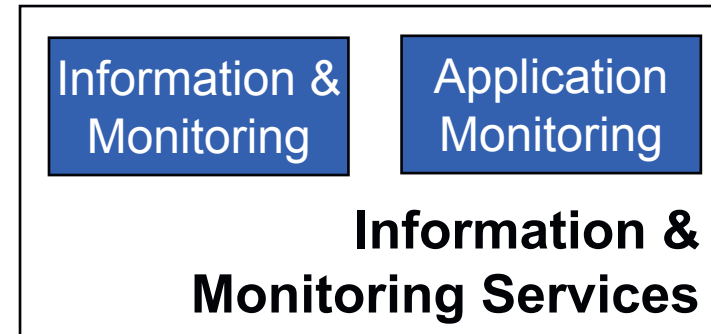
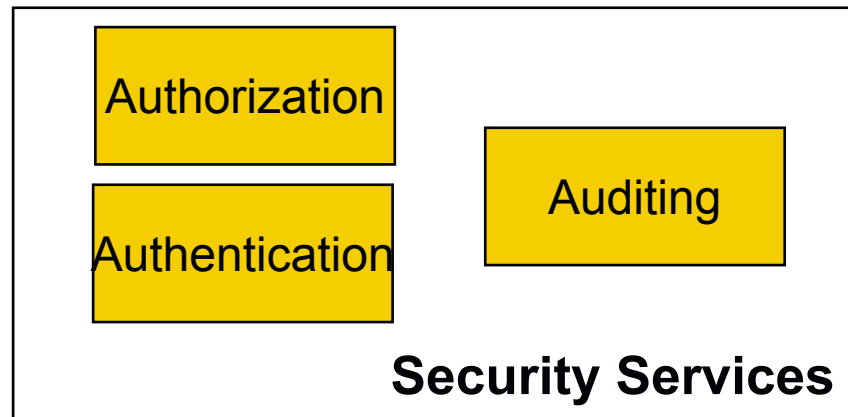
Erwin Laure

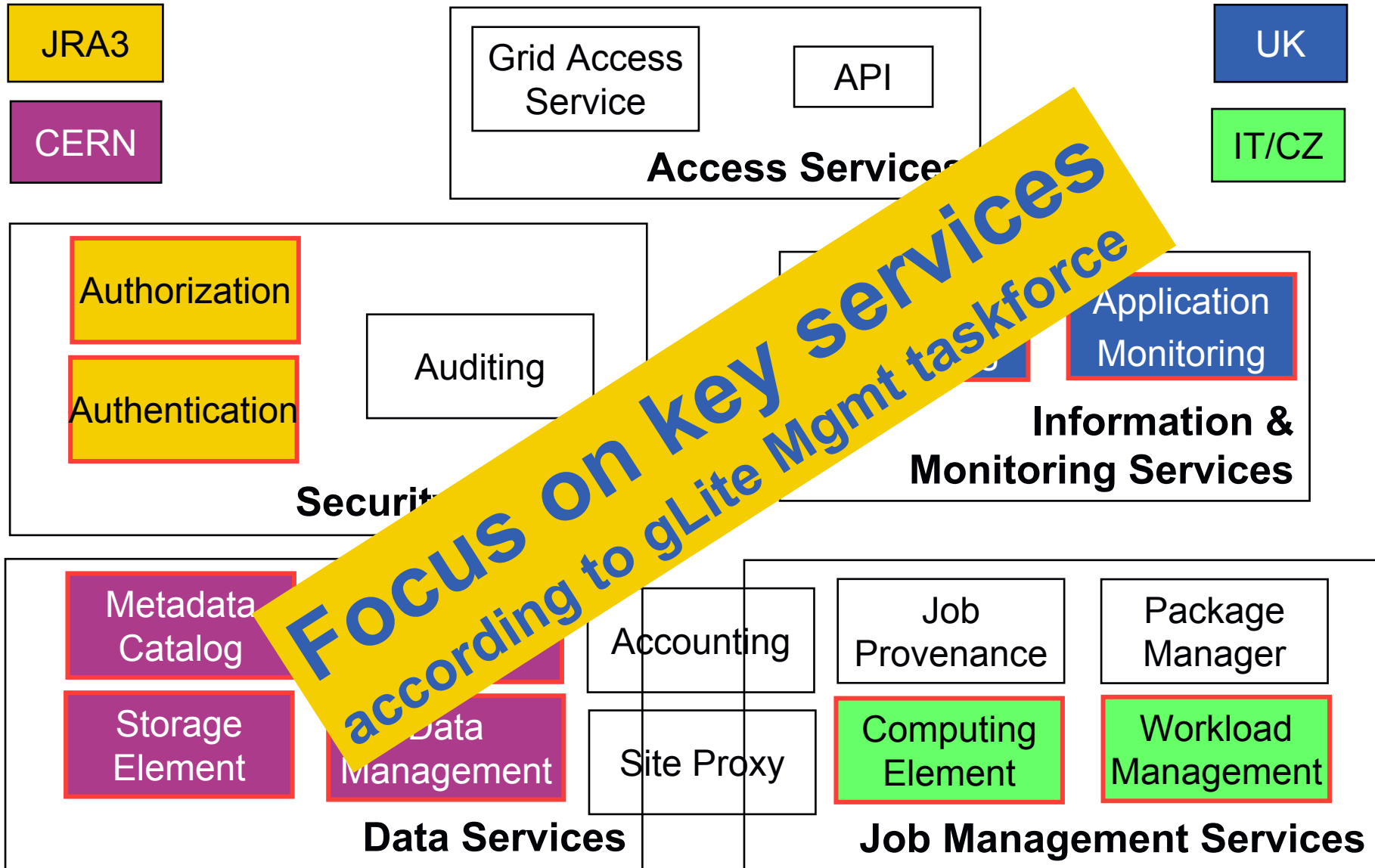
4th ARDA Workshop

7-8 March 2005



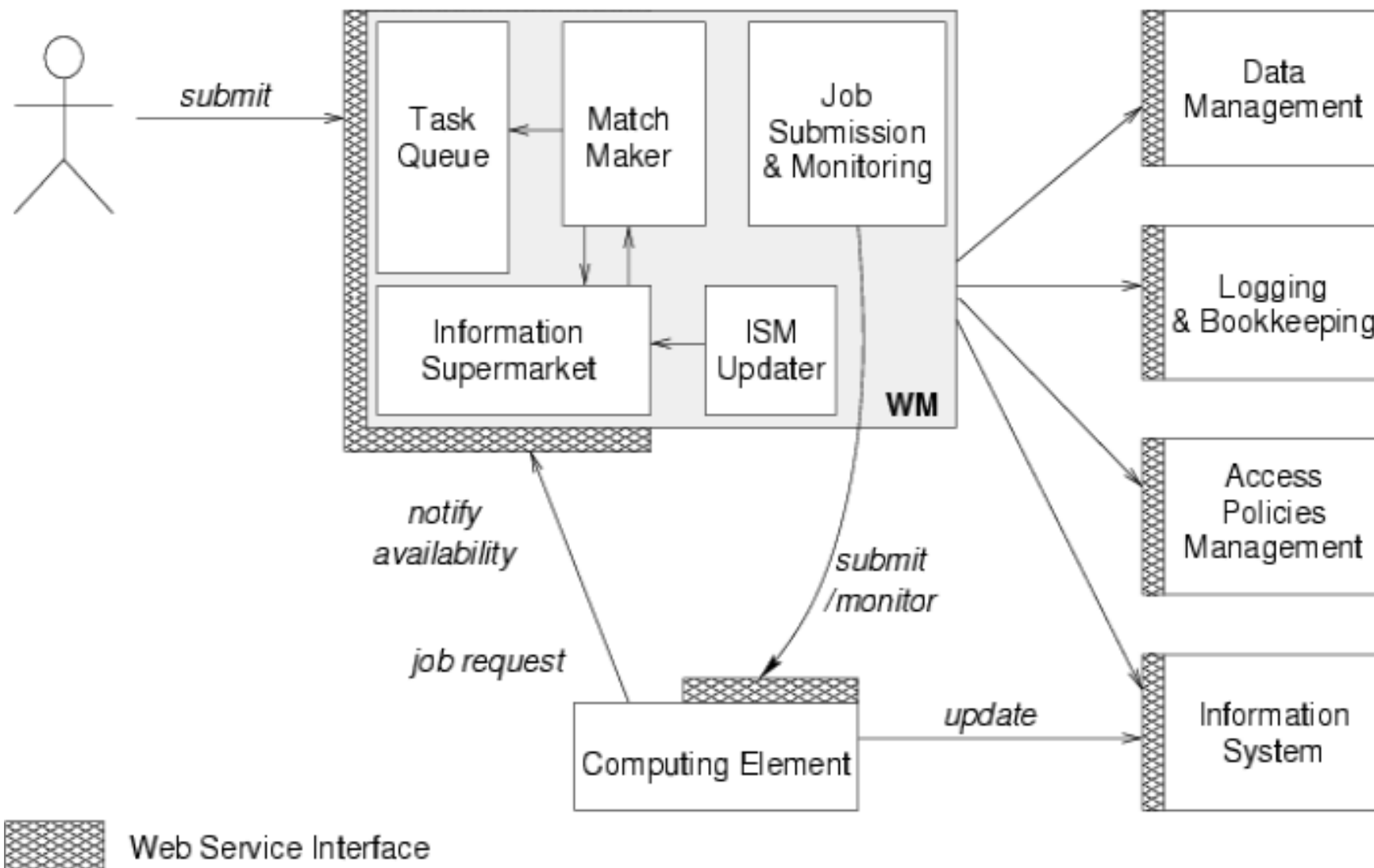
- **Design team** including representatives from Middleware providers (AliEn, Condor, EDG, Globus,...) including US partners produced middleware architecture and design.
- Takes into account **input and experiences** from applications, operations, and related projects
- **DJRA1.1 – EGEE Middleware Architecture (June 2004)**
 - <https://edms.cern.ch/document/476451/>
- **DJRA1.2 – EGEE Middleware Design (August 2004)**
 - <https://edms.cern.ch/document/487871/>
- Much **feedback** from within the project (operation & applications) and from related projects
 - Being used and actively discussed by OSG, GridLab, etc. Input to various GGF groups



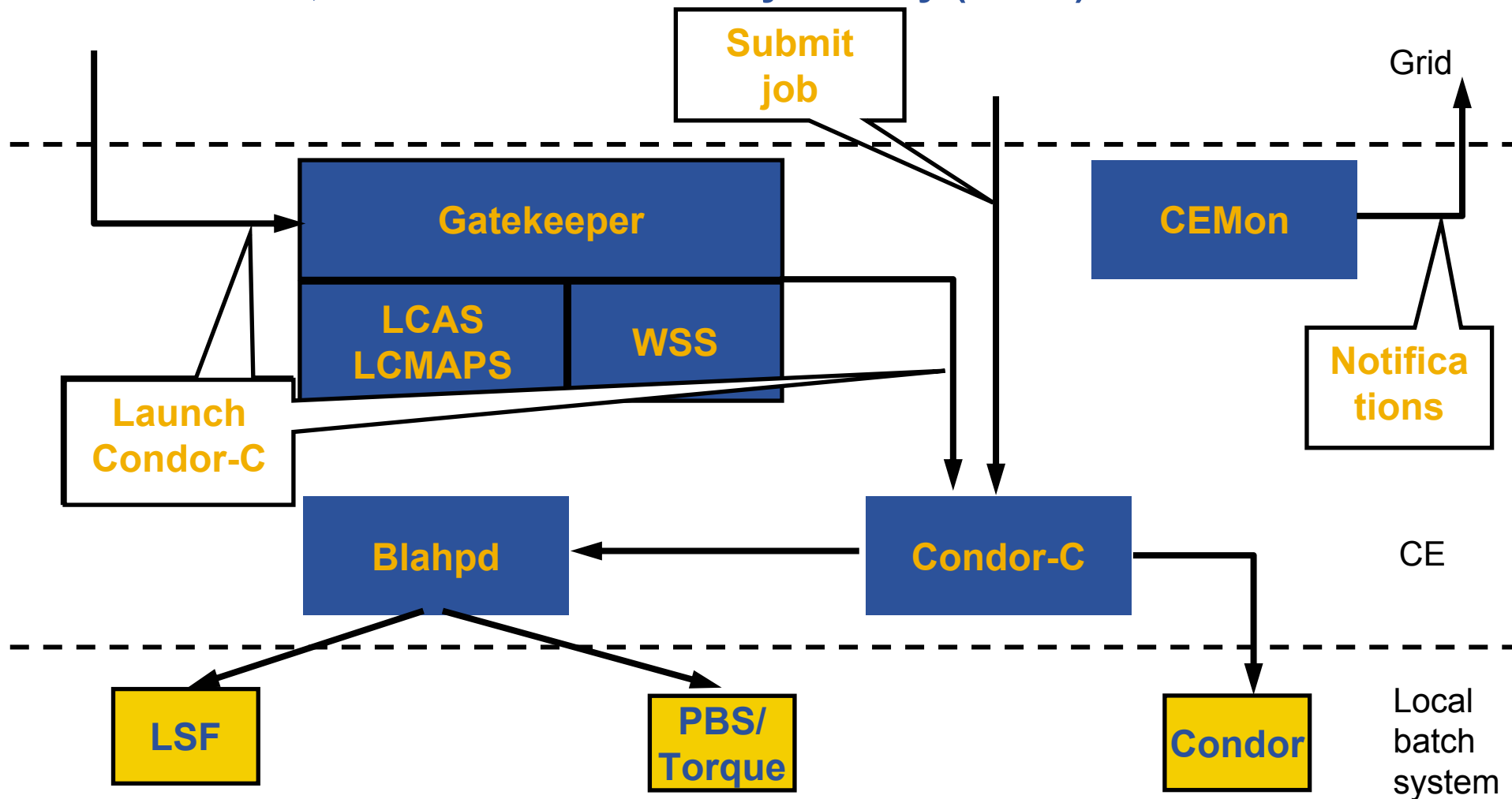


- **Computing Element**
 - Gatekeeper (Globus)
 - Condor-C (Condor)
 - CE Monitor (EGEE)
 - Local batch system (PBS, LSF, Condor)
- **Workload Management**
 - WMS (EDG)
 - Logging and bookkeeping (EDG)
 - Condor-C (Condor)
- **Storage Element**
 - File Transfer/Placement (EGEE)
 - glite-I/O (AliEn)
 - GridFTP (Globus)
 - SRM: Castor (CERN), dCache (FNAL, DESY), other SRMs
- **Catalog**
 - File and Replica Catalog (EGEE)
 - Metadata Catalog (EGEE)
- **Information and Monitoring**
 - R-GMA (EDG)
- **Security**
 - VOMS (DataTAG, EDG)
 - GSI (Globus)
 - Authentication for C and Java based (web) services (EDG)

- **Efficient and reliable scheduling of computational tasks on the available infrastructure**
- **Started with LCG-2 Workload Management System (WMS)**
 - Inherited from EDG
 - Support **partitioned** jobs and jobs with **dependencies**
 - Support for different **replica catalogs** for data based scheduling
 - Modification of internal structure of WMS
 - **Task queue**: queue of pending submission requests
 - **Information supermarket**: repository of information on resources
 - Better reliability, better performance, better interoperability, support push and pull mode
 - *Under development*
 - Web Services interface supporting bulk submission (after V1.0)
 - *Bulk submission supported now by use of DAGs*

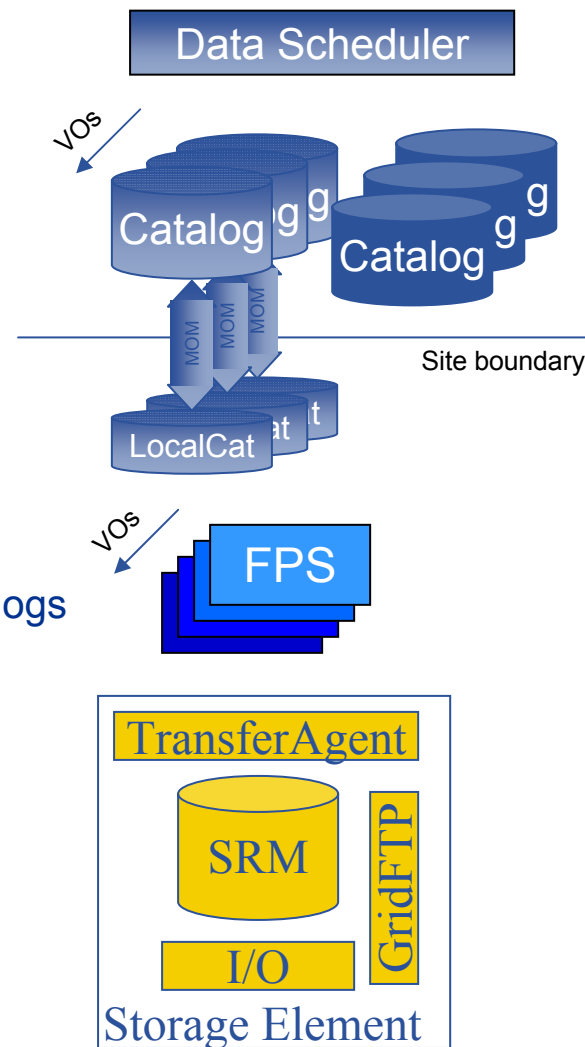


- Collaboration of INFN, Univ. of Chicago, Univ. of Wisconsin-Madison, and the EGEE security activity (JRA3)



- **Essentially addressing shortcoming of LCG-2 in the area of**
 - Robustness
 - Efficiency
- **Read-only Information System cache (ISM)**
 - Reduce reliance of an always working Information System
 - Updated by
 - Active polling of resources (CE in push mode)
 - Notification of available resources (CE in pull mode)
 - Combination of both
- **Task Queue**
 - Holding job requests when matchmaking is not possible
 - Persistent queue
- **Condor-C**
 - Reliable job submission between the WM and the CE
- **Bulk job submission**
 - Currently through DAGs without dependencies

- **Efficient and reliable data storage, movement, and retrieval on the infrastructure**
- **Storage Element**
 - Reliable file storage (SRM based storage systems)
 - Posix-like file access (gLite I/O)
 - Transfer (gridFTP)
- **File and Replica Catalog**
 - Resolves logical filenames (LFN) to physical location of files (URL understood by SRM) and storage elements
 - Hierarchical File system like view in LFN space
 - Single catalog or distributed catalog (*under development*) deployment possibilities
- **File Transfer and Placement Service**
 - Reliable file transfer and transactional interactions with catalogs
- **Data Scheduler**
 - Scheduled data transfer in the same spirit as jobs are being scheduled taking into account e.g. network characteristics (collaboration with JRA4)
 - *Under development*
- **Metadata Catalog**
 - Limited metadata can be attached to the File and Replica Catalog
 - Interface to application specific catalogs have been defined



- **Similar to lcg-utils**

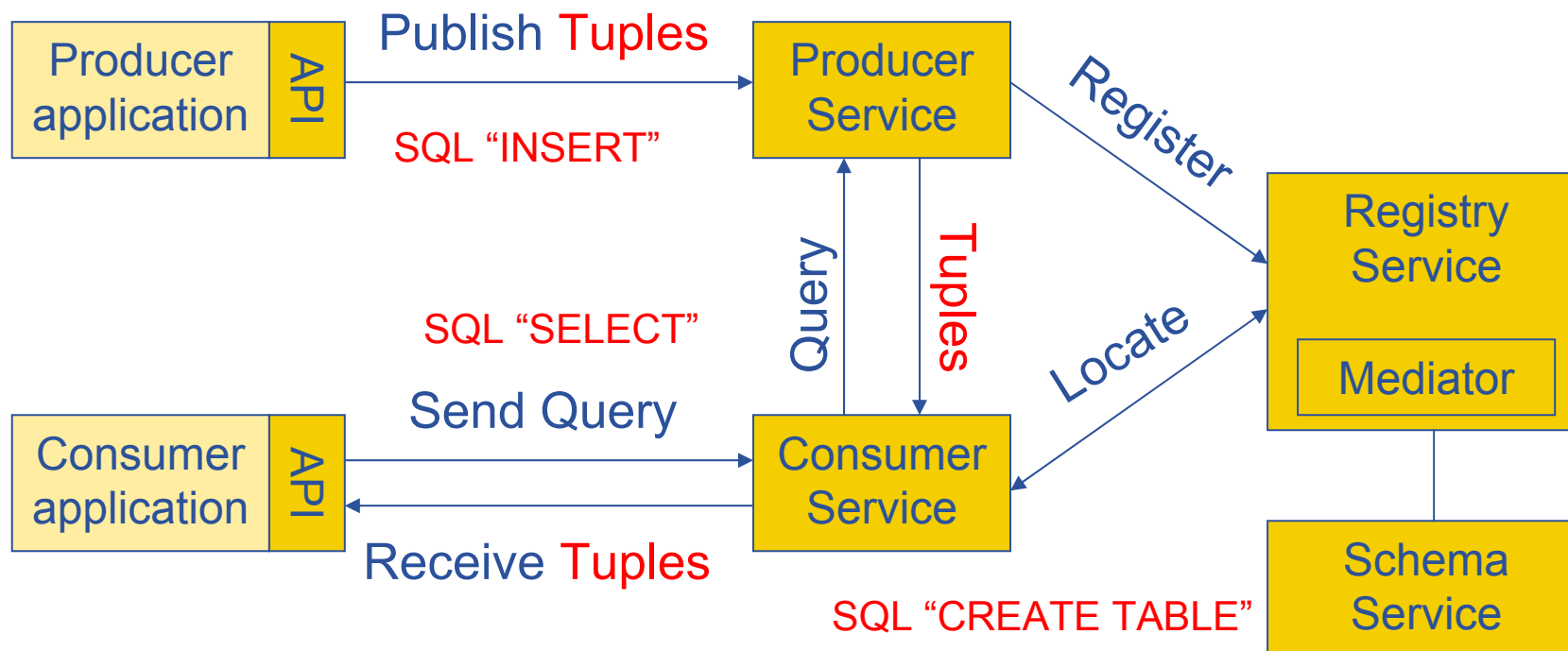
- glite-put and glite-transfer-submit (lcg-cr)
- glite-transfer-submit (lcg-rep)
- glite-rm (lcg-del)
- glite-catalog-get-replica (lcg-lr)
- glite-catalog-create (lcg-rf)
- ...

- **More information**

- <http://cern.ch/egee-jra1-dm/rc1-doc.htm>
- <http://egee-jra1-dm.web.cern.ch/egee-jra1-dm/lcg-utils-glite.htm>
- http://cern.ch/egee-jra1-data/glite-data_branch_1_1_1_RC1/stage/share/doc/glite-data-catalog-interface/html/index.html

- **Addressing shortcomings of LCG-2 data management**
 - RLS performance
 - A non distributed catalog
 - Although only single catalog supported in release 1
 - Lack of consistent grid-storage interfaces
 - Unreliable data transfer layer
- **Fireman Catalog**
 - Hierarchical Name Space
 - Bulk Operations
 - ACLs
 - Web Services Interface
 - Performance/scalability
- **gLite I/O**
 - Support of ACL's
 - Support of Fireman catalog in addition to RLS
- **File Transfer Service**
 - Did not exist on LCG-2

- Efficient and reliable provision of Grid information and Grid and Application monitoring data
- R-GMA (Relational Grid Monitoring Architecture)
 - Implements GGF GMA standard
 - Development started in EDG, deployed on the production infrastructure for accounting



- **Producer, Consumer, Registry and Schema services with supporting tools**
 - Registry replication
 - Simpler API – matching the next (WS) release
 - Provides smooth transition between old API and WS
 - coping with life on the Grid: poorly configured networks, firewalls, MySQL corruptions etc

- **Generic Service Discovery API**

- ***Under development***
 - Web Service version
 - File (as well as memory and RDBMS) based Producers
 - Native python interface
 - Fine grained authorization
 - Schema replication

- **Prototypes of Grid Access Service and Package Manager implemented in the AliEn framework**
- **Grid Access Service**
 - Acts on user's behalf
 - Discovers and manages Grid services for the user
- **Package Manager**
 - Provides dynamically distribution of application software needed
 - Does not install Grid middleware

- **Job Management Services**
 - WMS, LB, and CE implement authorization based on VOMS VO, groups, and user information
- **Data Services**
 - Authorization: ACL and Unix permissions
 - Fine-grained ACL on data enforced through gLite-IO and Catalogs
 - Catalog data itself is authorized through ACLs
 - Currently supported through DNs
 - VOMS integration being developed
- **Information Services**
 - Fine grained authorization based on VOMS certificates being implemented

- **Workload Management System** works in push and pull mode
- **Computing Element** moving towards a VO based scheduler guarding the jobs of the VO (reduces load on GRAM)
- **Distributed and re-factored file & replica catalogs**
- **Secure catalogs** (based on user DN; VOMS certificates being integrated)
- **Scheduled data transfers**
- **SRM based storage**
- **Information Services:**
R-GMA with improved API and **registry replication**
- **Prototypes of additional services**
 - Grid Access Service (GAS)
 - Package manager
 - DGAS based accounting system
 - Job provenance service
- **Move towards Web Services**



Here follows a list of main topics we still need to address, details and other topics need to be worked out with operations and applications

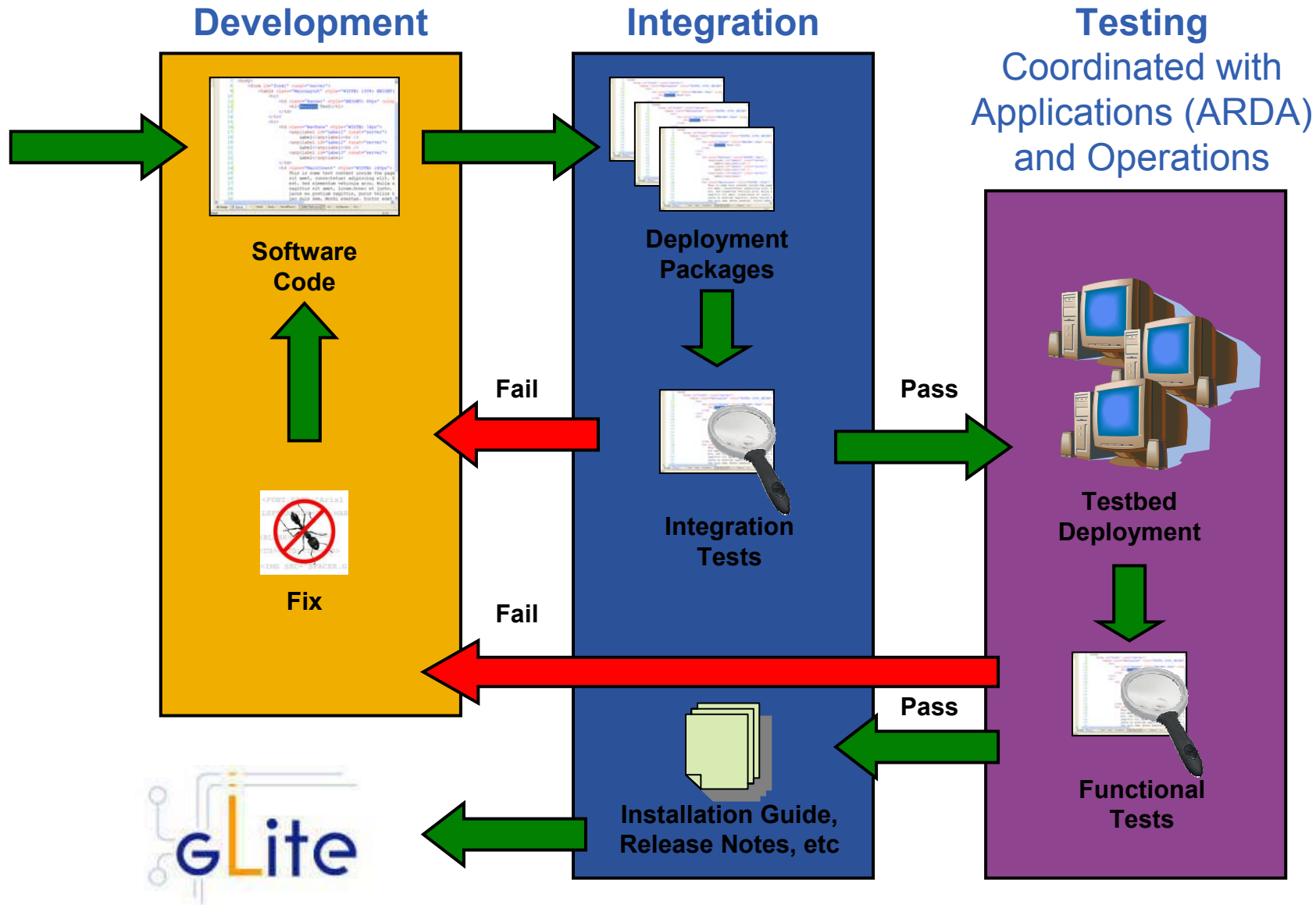
- **WMS**

- WS Interface
 - Better support for bulk job submission

- **CE**

- Head node monitoring
 - Guard and if necessary pause/resume services running on the head node
- SUDO service
 - Currently one Condor-C instance needed per user
 - Condor-C should run under a VO user and submit jobs via a sudo service to the local batch system

- **Catalogs**
 - Distributed and single deployment options
- **FTS/FPS**
 - Channel management
- **Data Scheduler**
 - “broker” for data transfer “jobs”
- **R-GMA**
 - Web service version
- **Package Manager**



- **All of the Services are available now on the development testbed**
 - User documentation currently being added
 - On a limited scale testbed
- **Most of the Services are being deployed on the LCG Preproduction Service**
 - Initially at CERN, more sites once tested/validated
 - Scheduled in April-May
- **Schedule for deployment at major sites by the end of May**
 - In time to be included in the LCG service challenge that must demonstrate full capability in July prior to operate as a stable service in 2H2005

- **Revision of architecture document in May 2005**
- **Revision of design document in June 2005**

- **bugfix releases on a weekly basis**
- **Functionality changes on a monthly basis**
- **2nd major gLite release in Dec. 2005**

- **gLite evolution will be driven by feedback from user community and operations**

<http://www.glite.org>

<http://cern.ch/egee-jra1>