Enabling Grids for E-science in Europe

www.eu-egee.org

PTF Meeting, 17-June-2004

Middleware Re-engineering
Status and Plans

Erwin Laure
EGEE Deputy Middleware Manager

EGEE is a project funded by the European Union under contract IST-2003-508833

Contents

Enabling Grids for E-science in Europe

- The middleware prototype work
- The gLite architecture

EGEE Middleware Work Breakdown Structure

Enabling Grids for E-science in Europe

- Main components:
  - Middleware Re-engineering
    - Workload Management, CE
    - Data Management
    - Information Services
    - Authentication/Authorization
    - Accounting
  - Integration
  - Testing

Integration Plans

Enabling Grids for E-science in Europe

- A master SCM Plan is being finalized now
- It contains basic principles and rules about the various areas of SCM and Integration (version control, release management, build systems, bug tracking, etc)
  - Use CVS@Cern
  - Automated build system is being set up (CruiseControl, Maven)
  - Development & User's guides being produced
  - Use Savannah portal (including bug-tracking)
- Change Control Board being established
  - Assess/approve/track public changes

Integration Team @ Work

Enabling Grids for E-science in Europe

- For the duration of the project, the integration team will:
  - Enforce and support all procedures by means of automated tools
  - Run the automated and manual operations of the build and release management processes
  - Deliver integrated installation packages for all supported platforms for further testing and deployment
  - Collect further input from all parties and introduce changes in the procedures and tools as necessary
- Supported platforms
  - Being discussed with SA1 (NA4 pending)
  - Most likely
    - Red Hat Enterprise AS 3.0 with gcc 3.2.3 and icc 8.0
    - Windows XP/2003 with cygwin+ vc++ 7.1
    - On both 32 and 64-bit processors.

Objectives of the testing activity

Enabling Grids for E-science in Europe

- Test all middleware components to ensure a production quality release that fulfils applications' requirements,
- Assess that all software requirements have been correctly and completely implemented and are traceable to system requirements,
- Test deployment, installation, configuration and documentation of all services,
- Test integrated software for scalability, platform independence and stress resilience,
- Automate as much as is possible all middleware testing,
- Establish a continuous integration and testing process
- Testsuites and test reports to be delivered as a part of a release.

Development Clusters

Enabling Grids for E-science in Europe

- 4 development clusters:
  - UK (Steve Fisher)
  - CERN/DM (Peter Kunszt)
  - IT/CZ (Francesco Prelz)
  - Nordic (security) ➔ JRA3
- Clusters have a reasonable sized (distributed) development testbed
  - Taken over from EDG
  - Nordic cluster to be finalized
  - Milestone MJRA1.2 (PM3) almost reached now
- Collaboration with integration & tools cluster established
- Clusters up and running!

American Involvement in JRA1

Enabling Grids for E-science in Europe

- UWisc
  - Miron Livny part of the design Team
  - Condor Team actively involved in reengineering for resource access (CE)
    - In collaboration with Italian Cluster and AliEn
- ISI
  - Identification of potential contributions started (e.g. RLS)
    - Focused discussions being planned
- Argonne
  - Collaboration on Testing started
  - Support for key Globus Components enhancements being discussed

High-Level Strategy for Middleware

Enabling Grids for E-science in Europe

- LCG-2 middleware package
  - Front line support by LCG
  - Second line support with original developers (most now EGEE)
- EGEE Middleware –
  - Re-engineered *generic* middleware package
  - Incorporating experience from AliEn, EDG, …., VDT
  - Architected for scale and performance requirements of LCG and other applications
- Fast prototyping approach – with clear end-to-end goals
  - Short update cycles to give applications the chance to influence and give feedback

Design Team

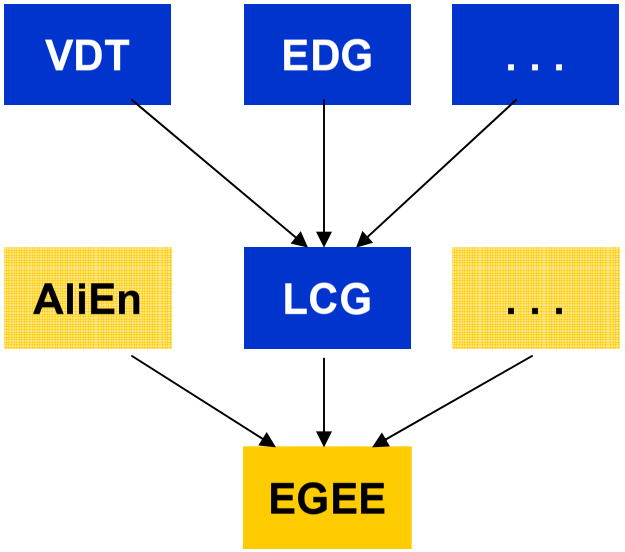
Enabling Grids for E-science in Europe

- Formed in December 2003
- Current members:
  - UK: Steve Fisher
  - IT/CZ: Francesco Prelz
  - Nordic: David Groep
  - VDT: Miron Livny
  - CERN: Predrag Buncic, Peter Kunszt, Frederic Hemmer, Erwin Laure
- Started service design based on component breakdown defined by ARDA
- Leverage experiences and existing components from AliEn, VDT, and EDG.
- A *working* document
  - Overall design & API's
  - **https://edms.cern.ch/document/458972**
- Draft of Middleware Architecture (DJRA1.1)
  - **https://edms.cern.ch/document/476451**

Guiding Principles

Enabling Grids for E-science in Europe

- Lightweight (existing) services
  - Easily and quickly deployable
- Interoperability
  - Allow for multiple implementations
- Resilience and Fault Tolerance
- Co-existence with deployed infrastructure
  - Run as an application (e.g. on LCG-2; Grid3)
  - Co-existence (and convergence) with LCG-2 are essential for the EGEE Grid service
- Service oriented approach
  - Follow WSRF standardization
  - No mature WSRF implementations exist to date, hence: start with plain WS – WSRF compliance is not an immediate goal



Planning

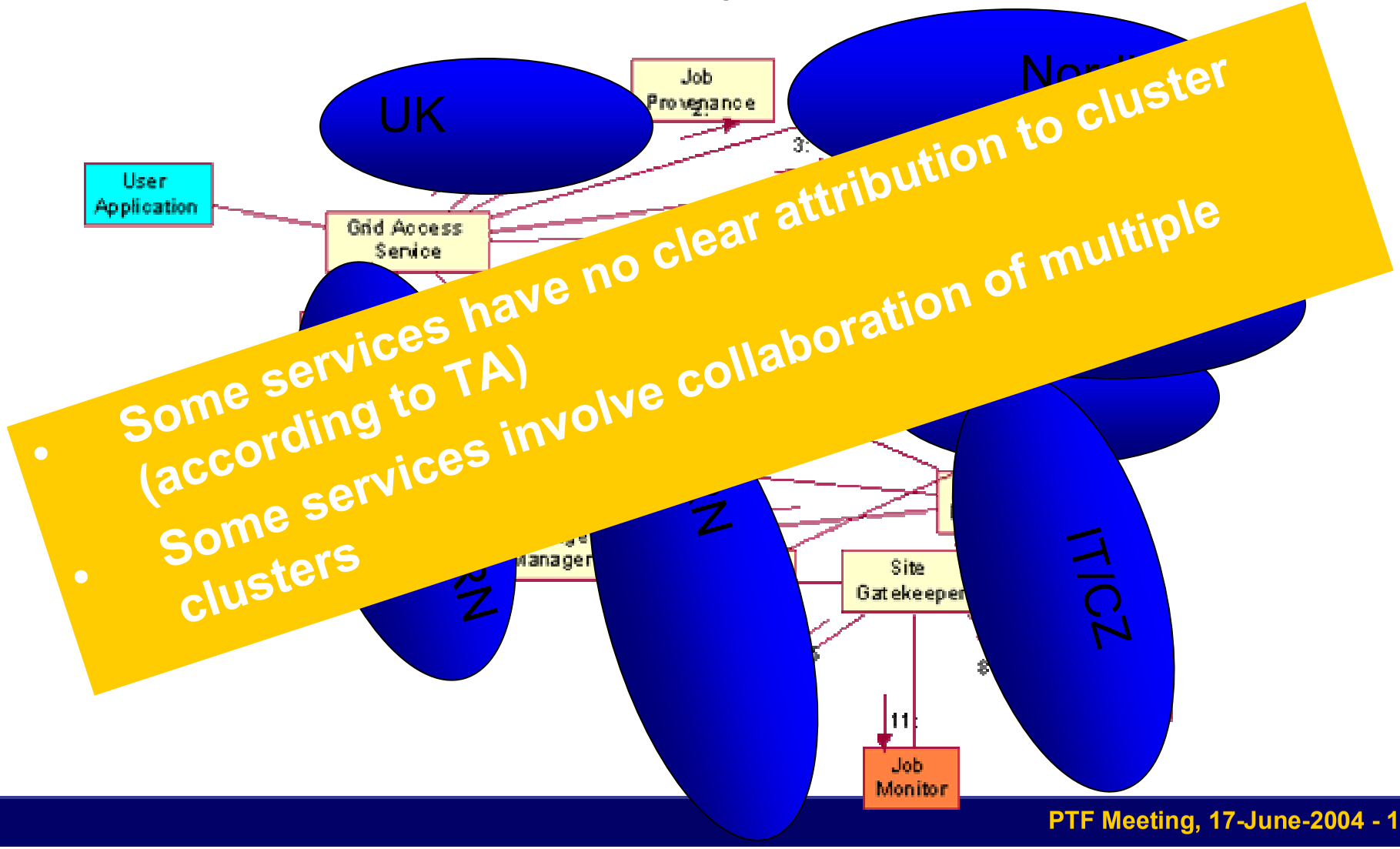
Enabling Grids for E-science in Europe

- Evolution of the prototype
  - Envisaged status at end of 2004:
    - Key services need to fulfill all requirements (application, operation, quality, security, …) and form a deployable release
    - Remaining services available as prototype
  - Developing a roadmap
    - Incremental changes to prototype (where possible)
    - Early user feedback through ARDA (and other sciences) and early deployment on SA1 pre-production service
    - Detailed release plan being produced (https://edms.cern.ch/document/468699)
  - Converge prototype work with integration & testing activities
    - Need to get rolling now!
    - First components started using SCM in May

High Level Service Decomposition

Enabling Grids for E-science in Europe

- Taken from the ARDA blueprint



- Some services have no clear attribution to cluster (according to TA)
- Some services involve collaboration of multiple clusters

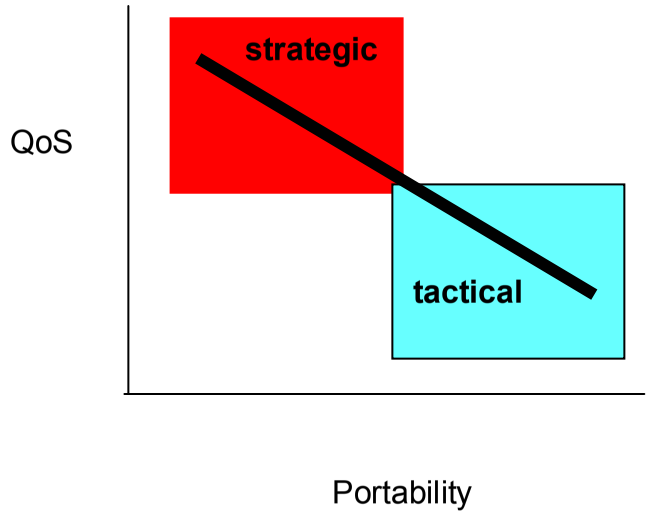
Initial focus in prototype work

Enabling Grids for E-science in Europe

- Data management
  - Storage Element
    - SRM based; allow POSIX-like access
- Workload management
  - Computing Element
    - Allow pull and push mode
    - Leverage CondorG for managing jobs on CE
- Information and monitoring
- Security
  - Need to integrate components with quite different security models
  - Start with a minimalist approach based on VOMS and myProxy

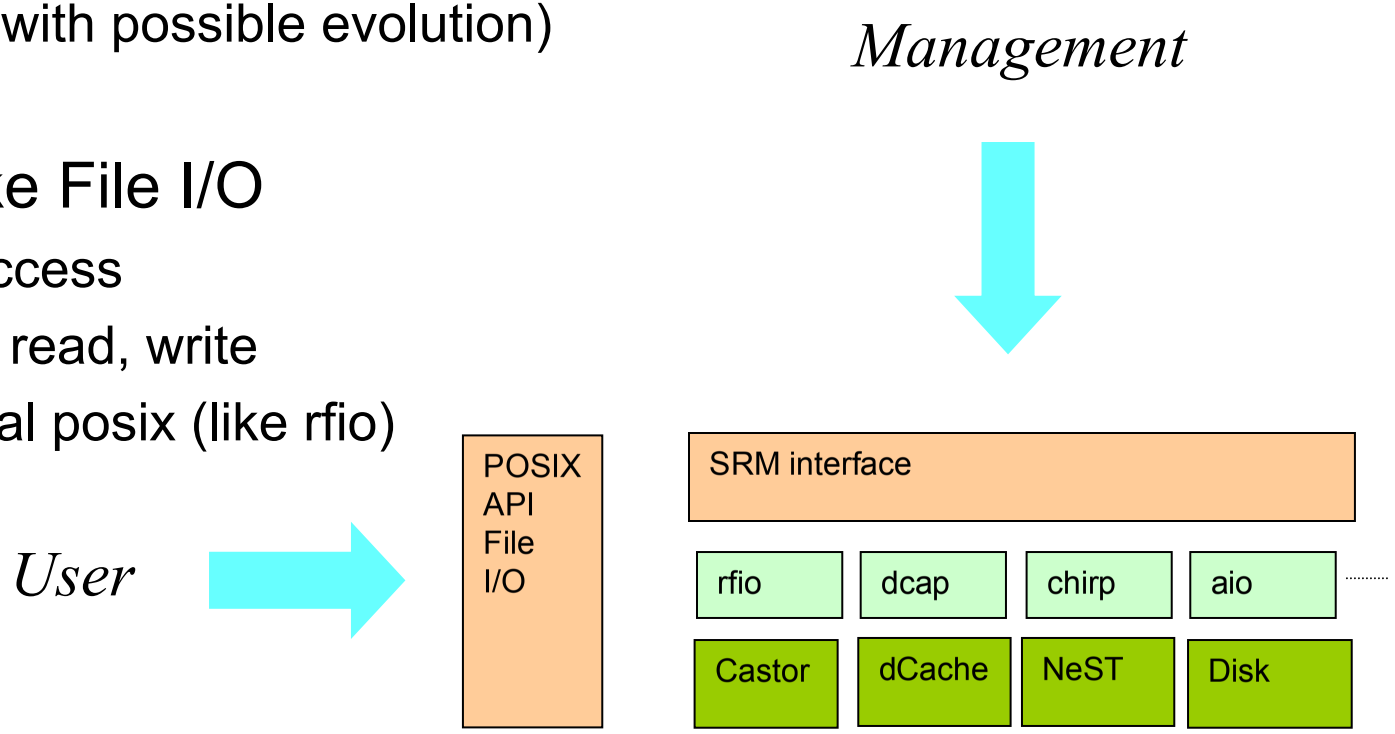
Storage Element

- ‘Strategic’ SE
  - High QoS: reliable, safe..
  - Has usually an MSS
  - Place to keep important data
  - Needs people to keep running
  - Heavyweight
- ‘Tactical’ SE
  - Volatile, ‘lightweight’ space
  - Enables sites to participate in an opportunistic manner
  - Lower QoS



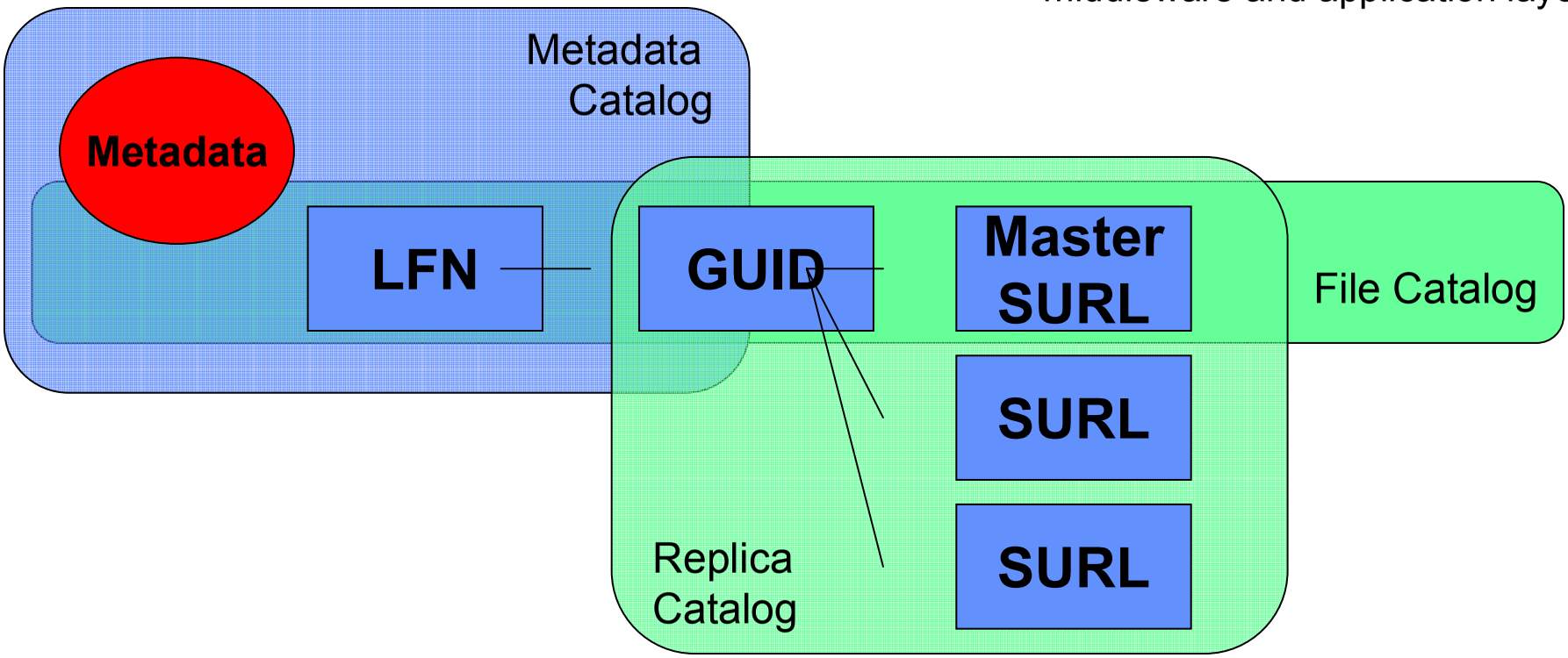
Storage Element Interfaces

- SRM interface
  - Management and control
  - SRM (with possible evolution)
- Posix-like File I/O
  - File Access
  - Open, read, write
  - Not real posix (like rfio)



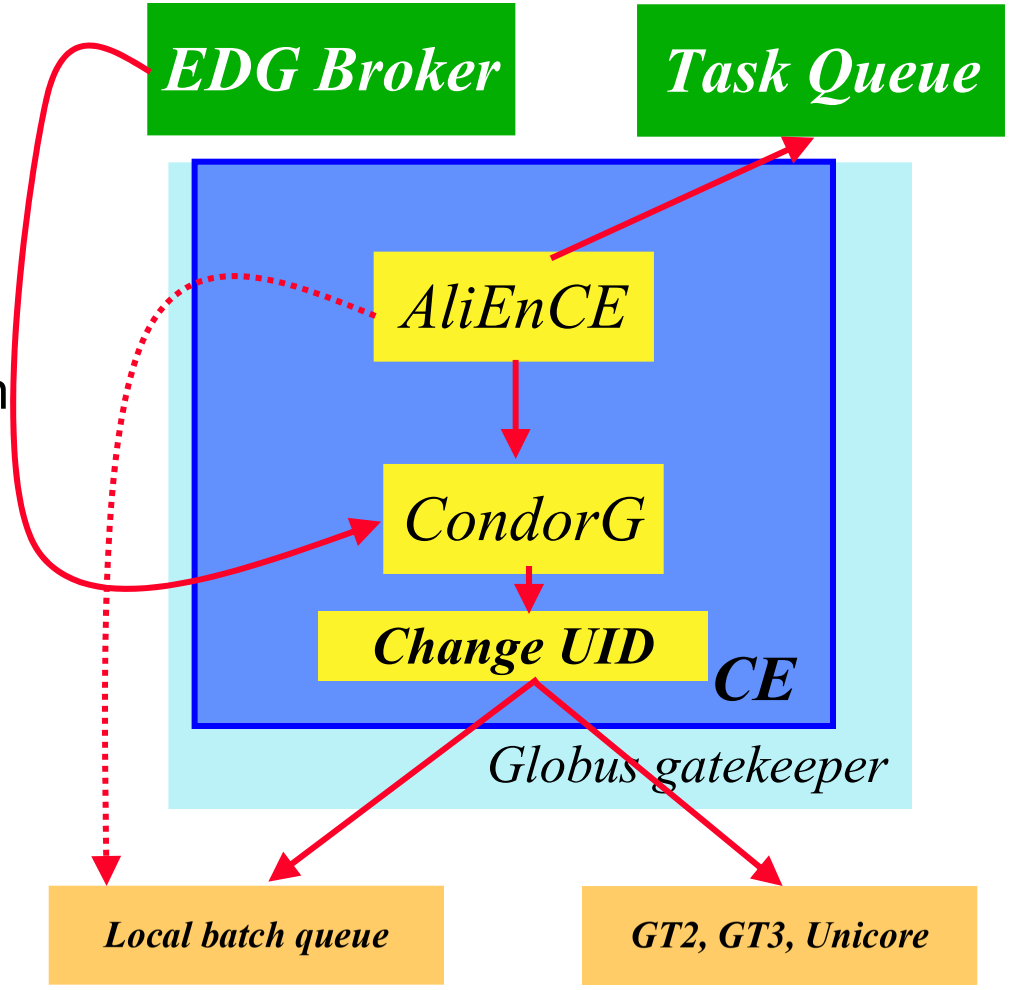
Catalogs

- File Catalog
  - Filesystem-like view on logical file names
- Replica Catalog
  - Keep track of replicas of the same file
- (Meta Data Catalog)
  - Attributes of files on the logical level
  - Boundary between generic middleware and application layer



Computing Element

- Layered service interfacing
  - various batch systems (LSF, PBS, Condor)
  - Grid systems like GT2, GT3, and Unicore
- CondorG as queuing system on the CE
  - Allows CE to be used in push and pull mode
- Call-out module to change job ownership (security)
- Lightweight service
  - should be possible to dynamically install e.g. within an existing globus gatekeeper



Information Service

Enabling Grids for E-science in Europe

- Adopt a common approach to information and monitoring infrastructure.
- There may be a need for specialised information services
  - e.g. accounting, package management, grid information, monitoring, provenance, logging
  - these may be built on an underlying information service
- A range of visualisation tools may be used
- Using R-GMA

Authentication/Authorization

Enabling Grids for E-science in Europe

- Different models and mechanisms
- Authentication based on Globus/GSI, AFS, SSH, X509, tokens
- Authorization
  - AliEn: exploits mechanism of RDBMS backend
  - EDG: gridmap file; VOMS credentials and LCAS/LCMAPS
  - VDT: gridmap file; CAS, VOMS (client)

Security and protection at a level acceptable by fabric managers and end users needs to be discussed and "blessed" in advance.

A minimalist approach to security

Enabling Grids for E-science in Europe

- Need to integrate components with quite different security model
- Start with a minimalist approach
  - Based on VOMS (proxy issuing) and myProxy (proxy store)
  - User stores proxy in myProxy from where it can be retrieved by access services and sent to other services
  - Credential chain needs to be preserved
    - Allow service to authenticate client
  - Local authorization could be done via LCAS if required
  - User is mapped to group accounts or components like LCMAPS are used to assign local user identity

Towards a prototype

Enabling Grids for E-science in Europe

- Focus on key services discussed; ex
- Initially an ad-hoc installation at Ce
- **First instance ready May 18**
  - Open only to a small user commu
  - Expect frequent changes (also A… …ation of further services
- Enter a rapid fe…
  - C… …g servi…
  - …vices based on …er-feedback
- Progre… …on weekly basis – http… …ns.cern.ch/document/457150/

Initial prototype compon… f… …(e.g. WMS)

This is not a release!
It's purely an ad-hoc installation

- Access service:
  - AliEn shell, APIs
- Information & Monitoring:
  - R-GMA
- CE:
  - AliEn CE, Globus gatekeeper, CondorG (with LSF/PBS backend via blahpd)
- Security:
  - VOMS, myProxy
- Workload mgmt:
  - AliEn task queue
- SE:
  - SRM (Castor; dCache), GridFTP, rfio
- File Transfer Service:
  - AliEn FTD
- File and Replica Catalog:
  - AliEn File Catalog, RLS

Development Roadmap

Enabling Grids for E-science in Europe

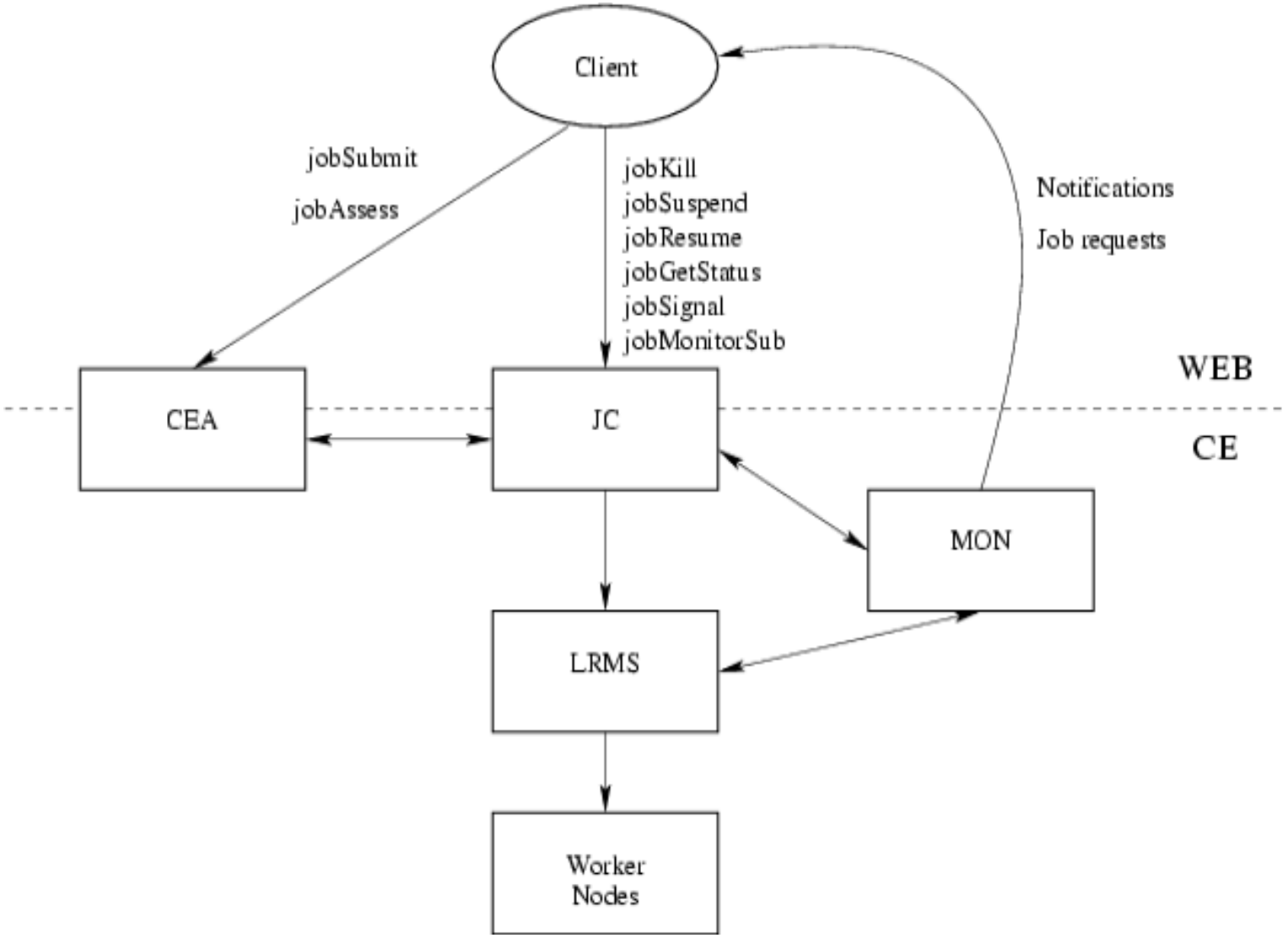
- Prototype work as starting point
- Priorities need to be adjusted based on user feedback
  - HEP and
  - Biomed
- Incremental, frequent releases
- **All discussions and decisions take place in the design team**
  - PTF to oversee this activity
  - Components and issue per cluster identified
    - Detailed release plan being worked out
      - First draft available: **https://edms.cern.ch/document/468699**
- EGEE boundary conditions:
  - Architecture document due end of Month 3 (June)
    - Draft available at: **https://edms.cern.ch/document/476451**
  - Design document due end of Month 5 (August)

Details from architecture document

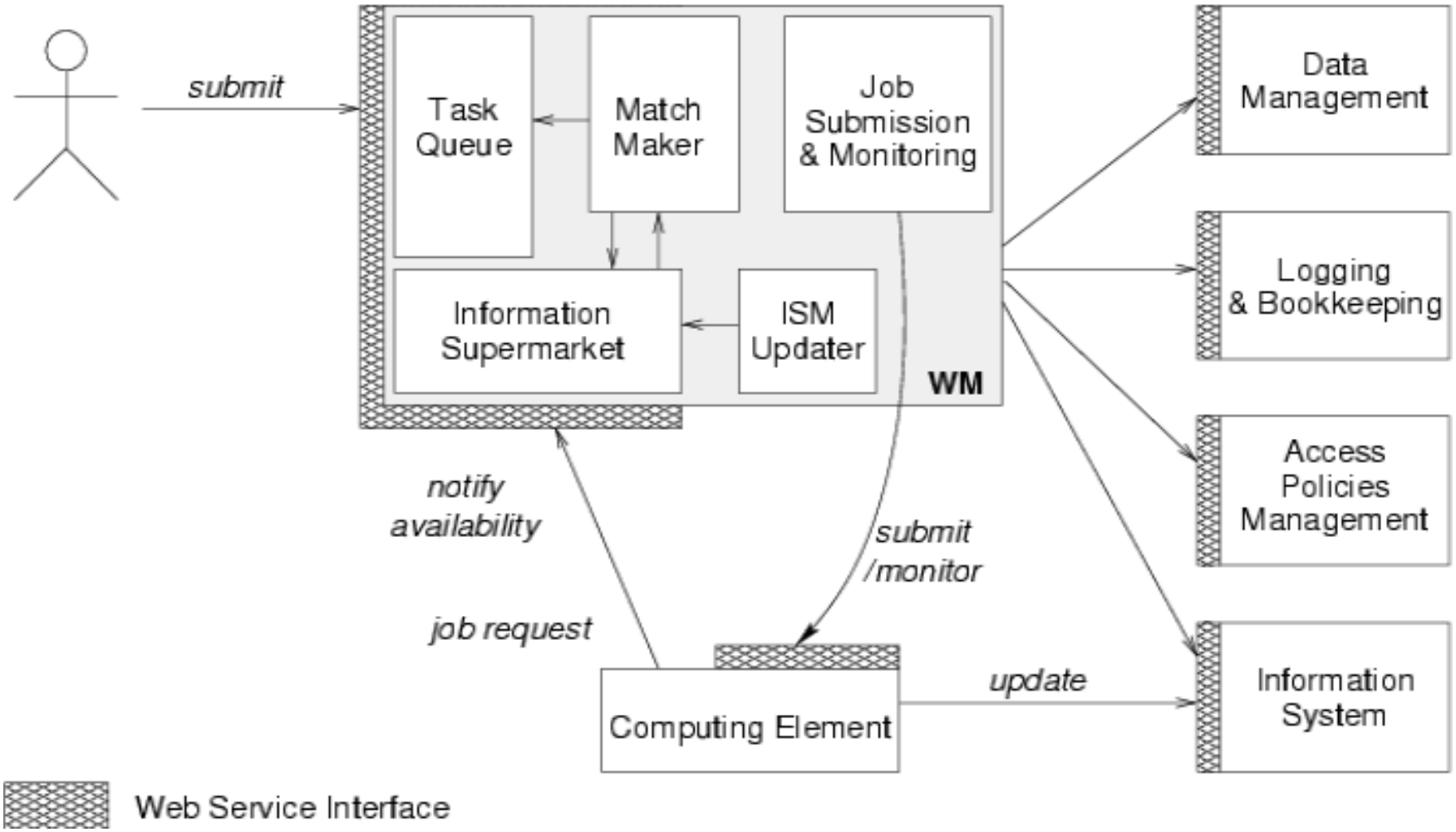
Enabling Grids for E-science in Europe

- CE
- Workload management
- Data management
- Policy enforcement
- Use Cases
  - Grid login
  - Production job
  - Analysis job

CE


Client
jobSubmit
jobAssess
jobKill
jobSuspend
jobResume
jobGetStatus
jobSignal
jobMonitorSub
Notifications
Job requests
WEB
CE
CEA
JC
MON
LRMS
Worker
Nodes

WMS


submit
Task Queue
Match Maker
Job Submission & Monitoring
Information Supermarket
ISM Updater
WM
Data Management
Logging & Bookkeeping
Access Policies Management
Information System
notify availability
job request
submit /monitor
update
Computing Element
Web Service Interface

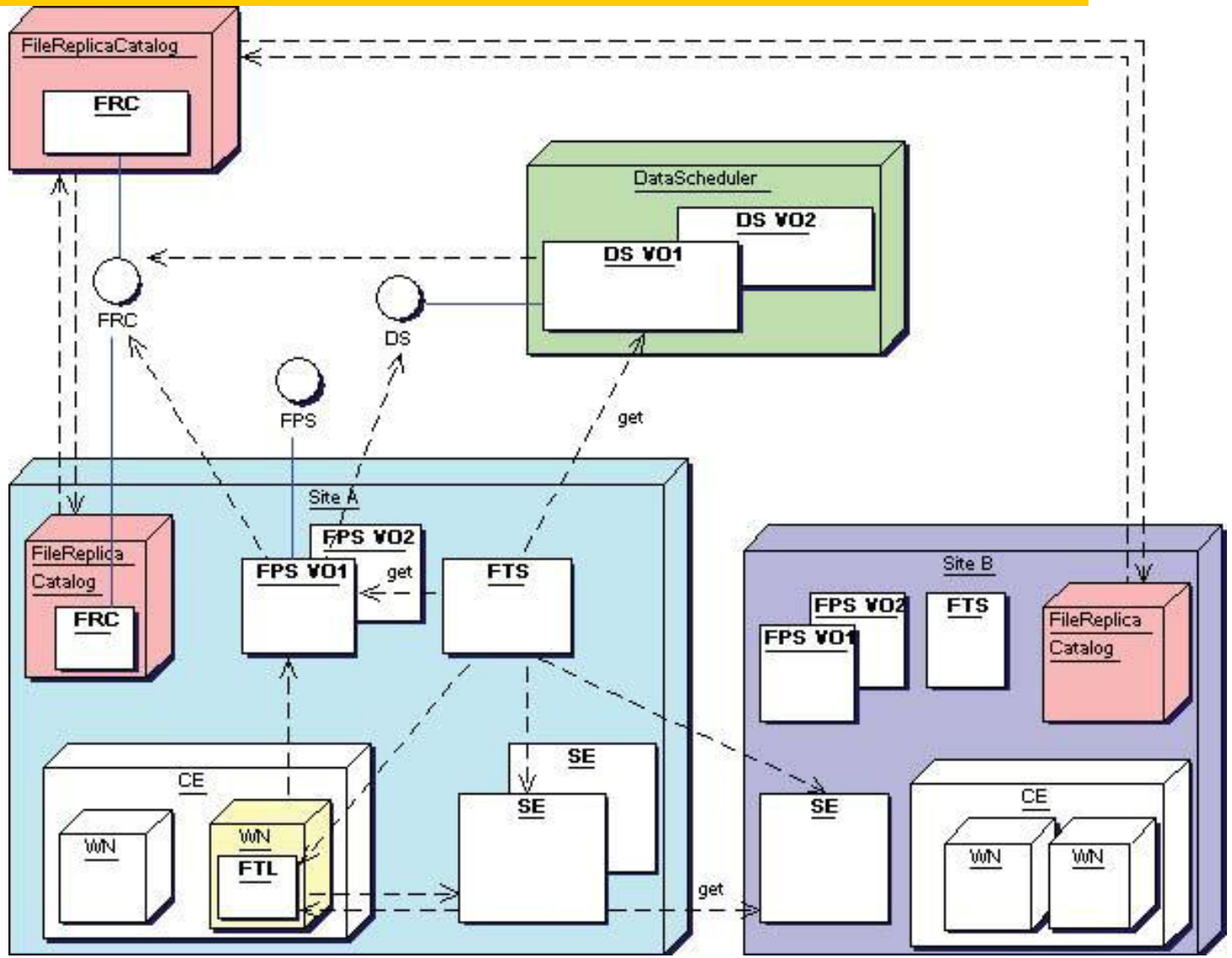
Data Management

Enabling Grids for E-science in Europe

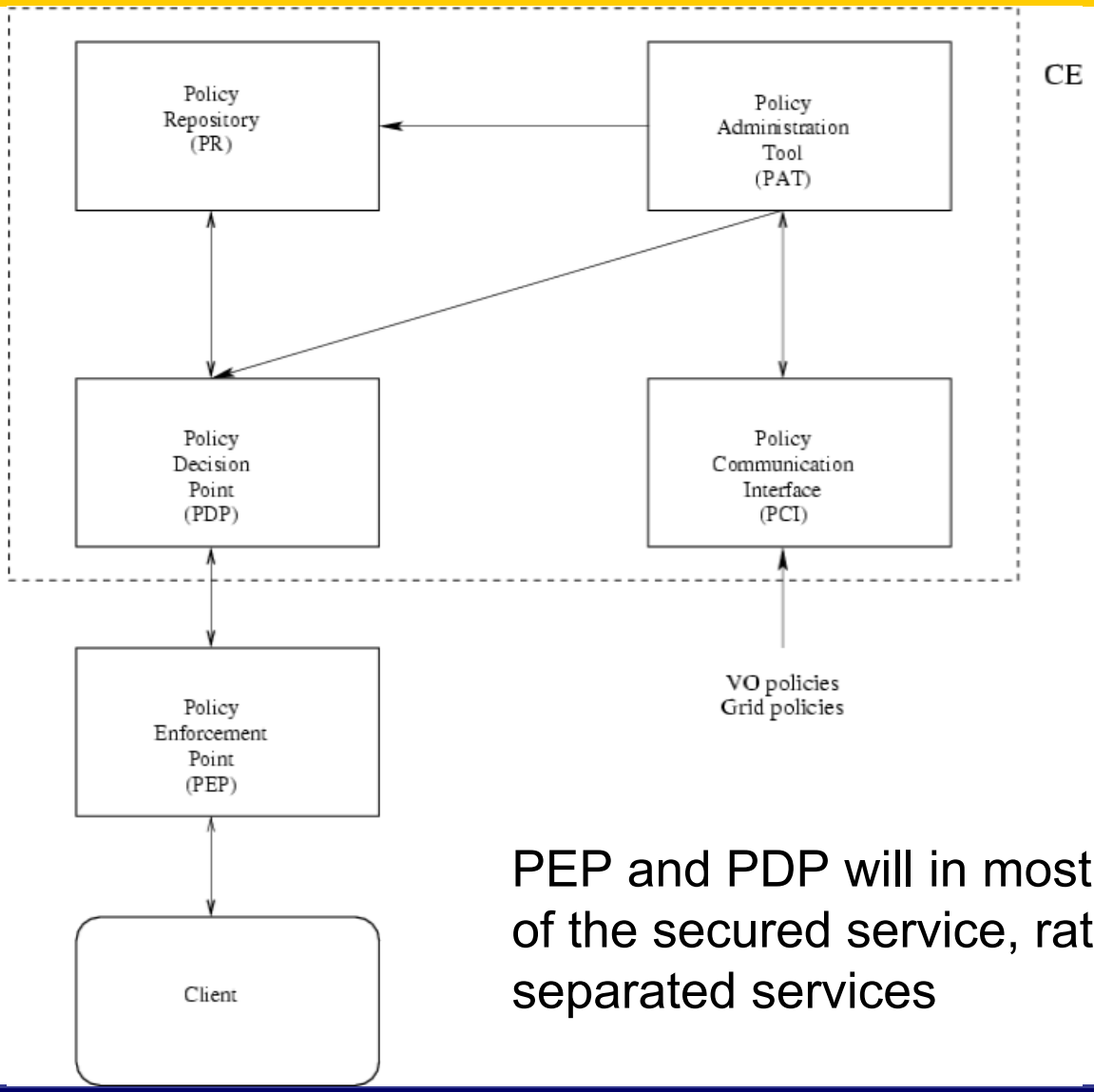
- Scheduled data transfers
  - Much like jobs
  - Investigation of Stork (Univ. of Wisconsin) as building block
- ACLs on
  - File Catalog
  - SE

Data Management

Enabling Grids for E-science in Europe

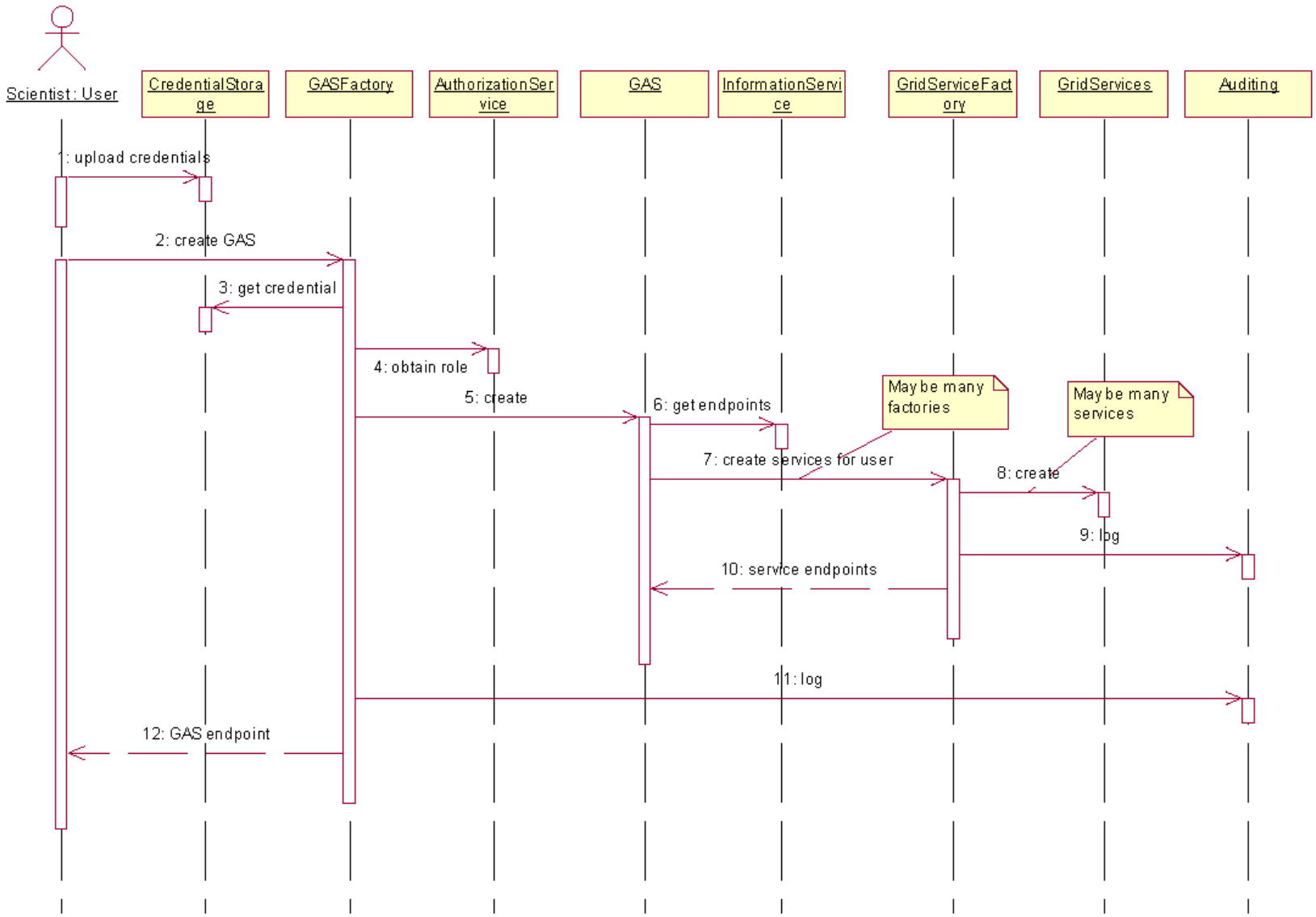


Security

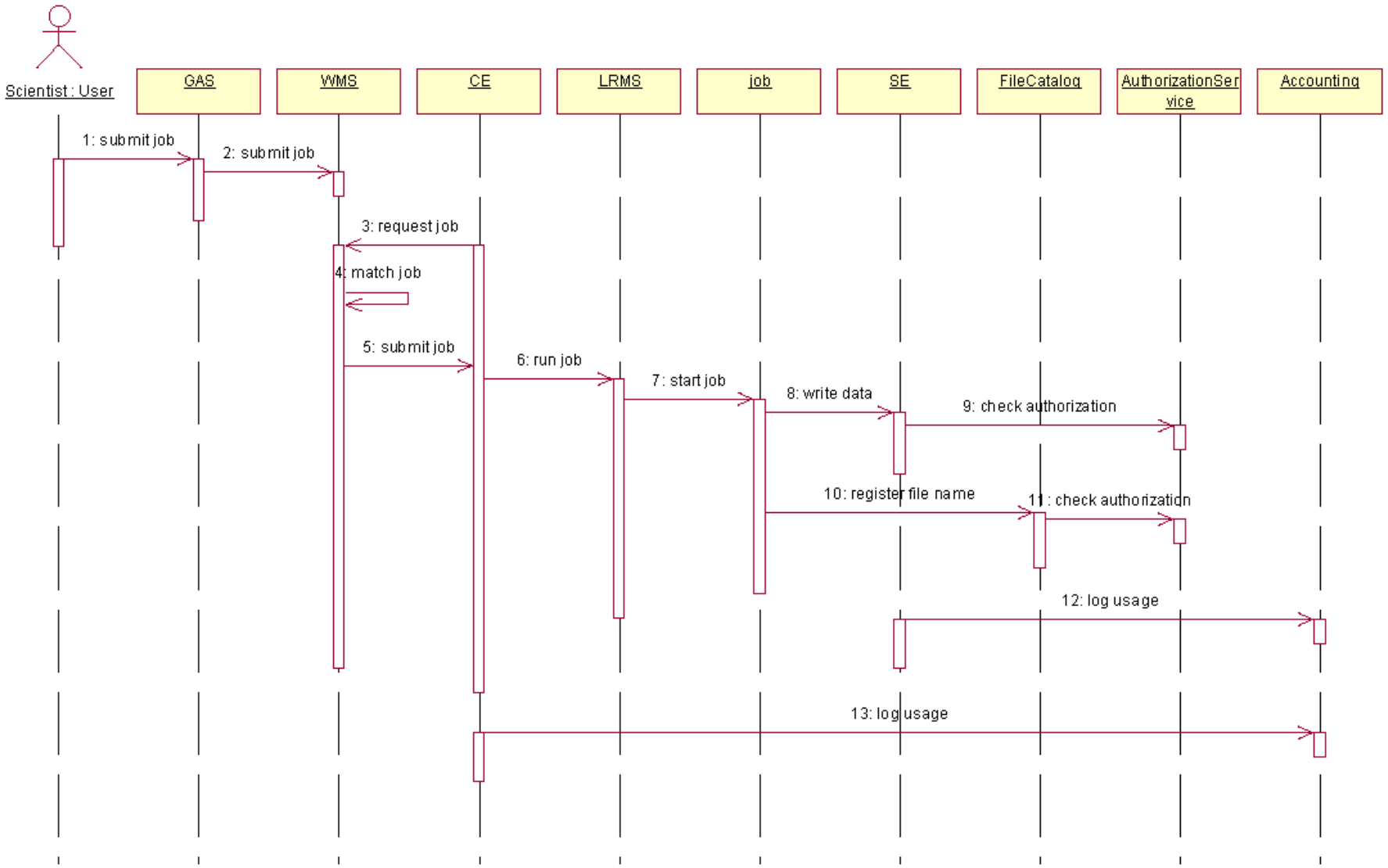


PEP and PDP will in most cases be part of the secured service, rather than separated services

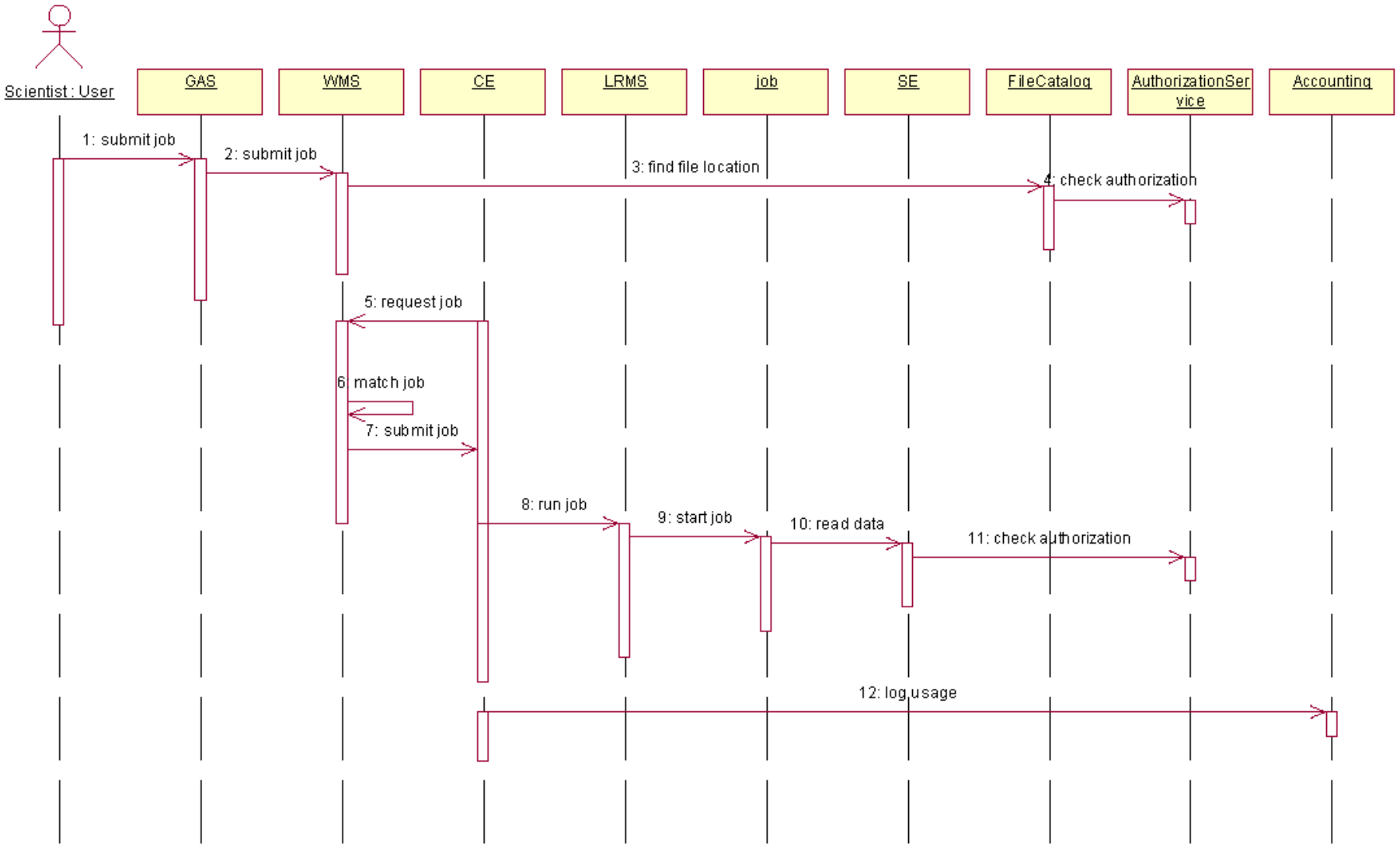
Grid Login


Scientist: User
CredentialStorage
GASFactory
AuthorizationService
GAS
InformationService
GridServiceFactory
GridServices
Auditing
1: upload credentials
2: create GAS
3: get credential
4: obtain role
5: create
6: get endpoints
7: create services for user
May be many factories
8: create
May be many services
9: log
10: service endpoints
11: log
12: GAS endpoint

Production Job



Analysis Job



JRA1/JRA4 Interactions

Enabling Grids for E-science in Europe

- Discussions with JRA4 ongoing on
  - Network monitoring
  - Network queries (status, bandwith reservations, …)
  - Ask for network reservation
- Used for
  - Data management
    - Find “good” replicas
    - Accept/deny/schedule file transfers
  - Workload mgmt
    - Advanced reservation

Summary

Enabling Grids for E-science in Europe

- Next generation middleware being designed and assembled
  - Prototype first tangible outcome
    - **BUT this is a PROTOTYPE!**
  - Architectural and design work well advanced
  - Incremental changes to prototype
    - Feedback from NA4 essential!
      - Work with HEP via ARDA started
      - Work with Biomed started June 15
- Detailed release plan being worked out
- Continuous integration and testing scheme defined and being adopted (first components in May)
- Technology Risk
  - Will WS allow for all upcoming requirements?
  - Divergence to standards