



LHC Computing Grid Project - LCG

Towards an ARDA Prototype Generic Middleware

ARDA Workshop

21-22 January 2004

Miron Livny (Univ. Wisconsin)

Francesco Prelz (INFN), Steve Fisher (RAL)

Frederic Hemmer, Predrag Buncic, Peter Kunszt, Erwin Laure (CERN)

Torre Wenaus (LCG AA)

AliEn EDG VDT



“Leaving religion where it belongs!” (© fca 2004)



<http://www.arदा.тm/>

AliEn EDG VDT



Middleware Contributors

- AliEn - Is the distributed computing environment of the Alice experiment.
- EDG - components deployed in LCG-1 and LCG-2 and others
 - Atlas/CMS/D0 Production tests; evaluation by Alice, LHCb, BaBar
- VDT - Provides the grid middleware foundations to Grid2003 (supports Atlas & CMS Data Challenges) and to LCG.



The Message of the ARDA Report

Deliver end-to-end capabilities (*from user to fabric*) and stability (*deployable*) at the price of services offered (*functionality*)

Services provide a natural abstraction and powerful software engineering constructs

AliEn provides a useful and stable suite of services as it meets the expectations of the Alice experiment



The ARDA Prototype Challenge

- How to meet the objectives of the ARDA report while leveraging existing middleware developed and hardened by the different Grid projects and the experience of the different groups (*users, application developers, deployment teams, support teams, fabric & site managers, ...*)
- How to exploit the independent approach of AliEn and the mandate of the ARDA report to reconsider "sacred cows" in common grid technology (*group accounts, load control, user protection, ...*)
- How to extend the AliEn experience to diverse software stacks and diverse software engineering approaches
- Do no harm!



ARDA Goals

Deliver by the set deadline a **prototype** of a generic middleware stack capable of supporting four (or more) **prototypes** of distributed analysis environments each of which representing one (or a subset of one) of the four LHC experiments.

- Powerful enough to support end-to-end capabilities required by a distributed analysis prototype
- Stable enough to support a “small” group of “friendly” users
- Flexible enough to support experiment specific meta data
 - **Boundary issue between generic and application specific MW**
- Secure enough to be accepted by “friendly” fabric managers

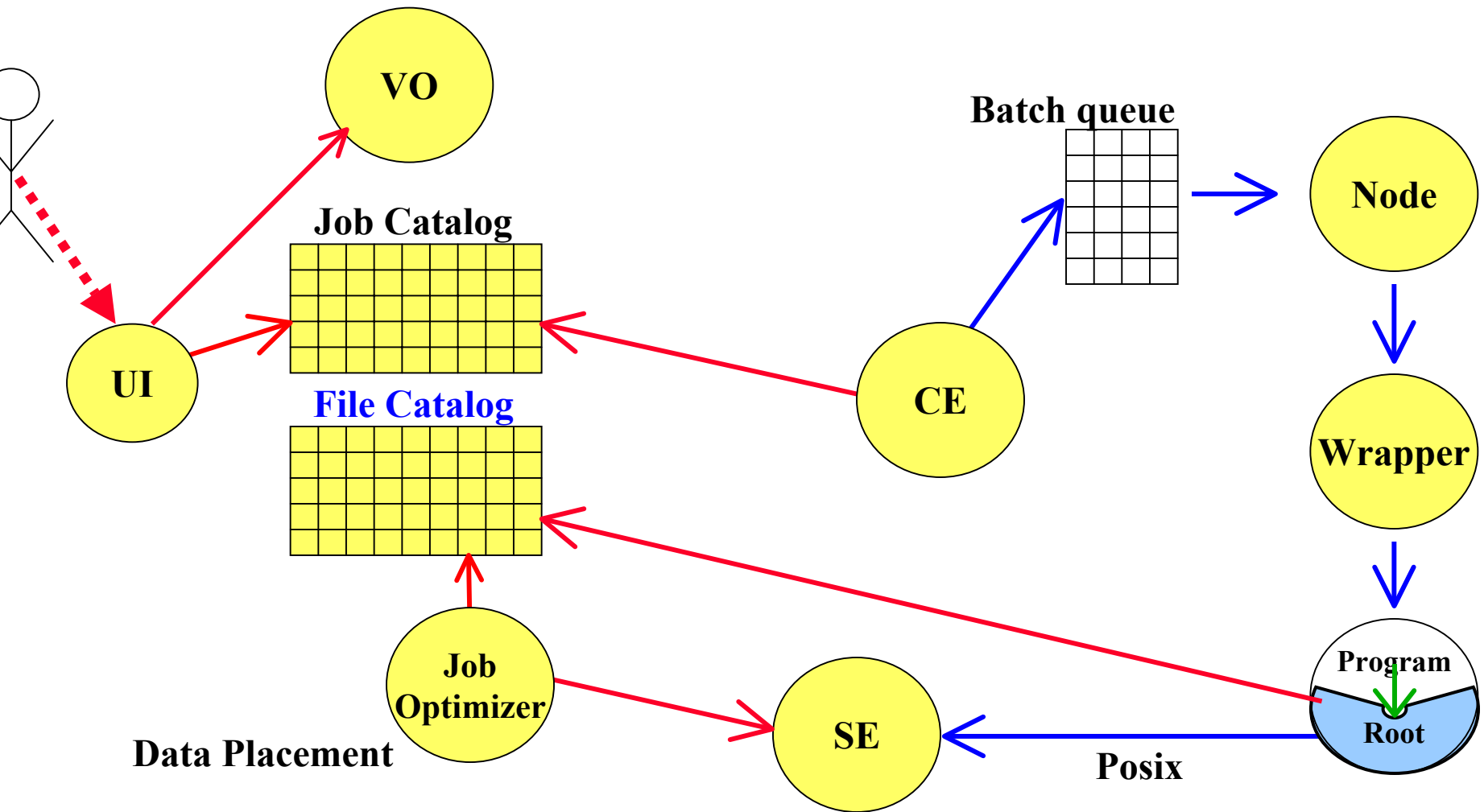
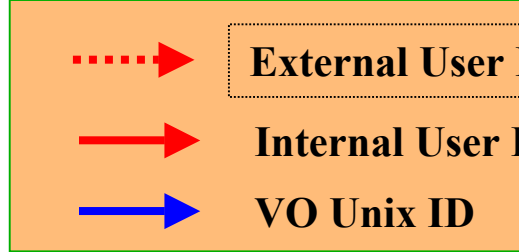


From Report to Prototype

- Understand AliEn (including services)
- Identify potential contributions from existing middleware
- Understand requirements (how does analysis differ from “production”?)
- Develop a plan - what, how, when, who
 - Semantic of exposed services
 - Authentication/protection model
 - Integration/testing/deployment procedures
 - Documentation
 - ...
- Execute the plan!



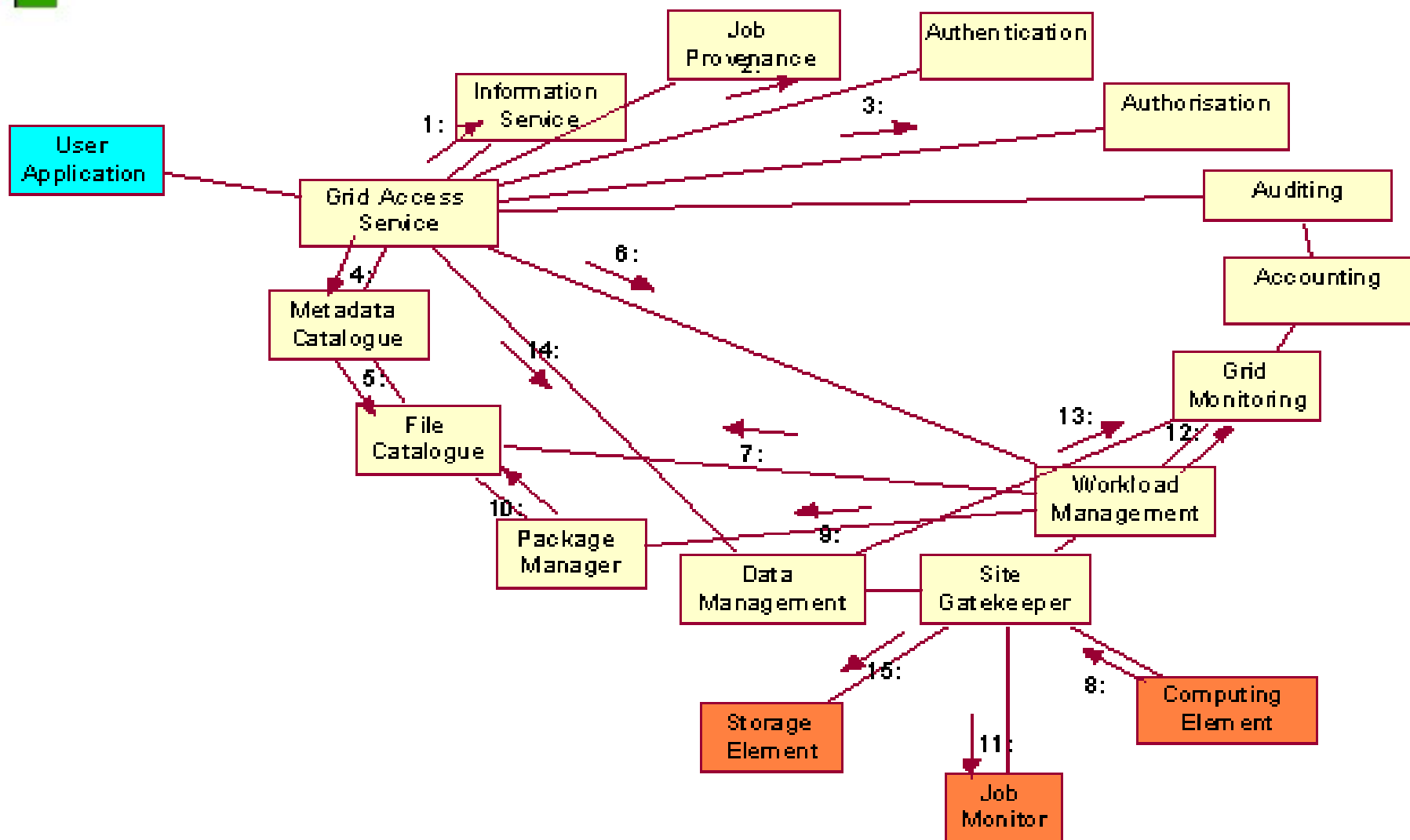
A simplified AliEn Functional Decomposition



AliEn EDG VDT



ARDA Services



AliEn EDG VDT



Ranking Services (and functionality)

Reality: virtually all the services identified by the ARDA are essential. However, only a subset of them can be (fully) developed and implemented in the first phase.

- Which one can be given low priority in the first round?
- Which one can be only partially addressed?
 - Functionality
 - Centralized or distributed
 - Robustness, stability, performability
 - Ease of use, deployment and management
- Which one unify our middleware base?
- Which one are ready "out of the box"?



Middleware Working Document

Abstract: This working document is used to break down the high level services defined by ARDA to actual components and tries to map these components to existing implementations coming from AliEn, EDG, and VDT. The structure and initial AliEn input is taken from Chapter 5 of Draft v0.2 of the ARDA document (unpublished)

- Started after the December meeting as a vehicle to exchange and record information and ideas among the middleware providers.
 - Identification of services
 - Service interplay and semantics
 - Understand how existing MW could implement these services
 - Input from AliEn, EDG, VDT, commercial, (others?)
 - Specify interfaces to applications

AliEn EDG VDT



Access Services (and APIs)

The Access services and APIs expose the capabilities of the computational environment to the end user.

- User entry point to distributed analysis services
- Stateful service authenticating and authorizing the user
- One instance per analysis session
- Instantiates the proper UI according to user roles
- Service calls can be routed through these access services

AliEn offers a rich set of such services and APIs.



Authentication/Authorization

- 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.



Workload Management Services

- Distributes the workload on the Grid resources
- AliEn uses a pull model; EDG a push model, VDT supports both (uses mainly push)
- Site must control (access and priority) consumption of ALL local resources - head node, worker nodes, storage resources, network bandwidth
- All resources must be protected by a claim and leasing protocol
- Requires a hierarchy of optimizers, planners, matchmaker, ... capable of dynamic (eager, lazy, just in time) workload management
- Requires flexible local resource management policies



Data Management

Central to analysis and requires a broad suite of services. Initial focus on:

- Replica Location/File Catalog
- File Transfer Services
- Storage Element (SE)
 - SRM
 - GridFTP
 - Posix I/O
- Interaction with Pool
- Metadata catalogs
 - Application specific
 - Borderline between application area and generic MW



Replica Location/File Catalog

- Associations between logical (LFN) and physical (PFN) file names
- Provides small amounts of metadata (file size, checksum, etc.)
- Supports a file-system like view



File Transfer Service

- Scheduled file transfers (treat them as “first class” jobs) allow planning of file transfers
- Transfer services needs to be reliable (i.e. involving re-tries)
- Allocation of network resources and storage resources should be also treated as a job.



Information Service

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



Immediate Next Steps

- Identify and propose application level APIs
- Get experiments involved in assessing the APIs
- Start deploying the prototype based on existing components

Experiments need to be closely involved and committed to use the prototype of the middleware early and provide timely feedback



'We should not miss this
opportunity to put it all
together!" (© fca 2004)