# Project Architecture, Middleware and Delivery Schedule

Bob Jones

Technical Coordinator, WP12, CERN

Bob.Jones@cern.ch

# Outline

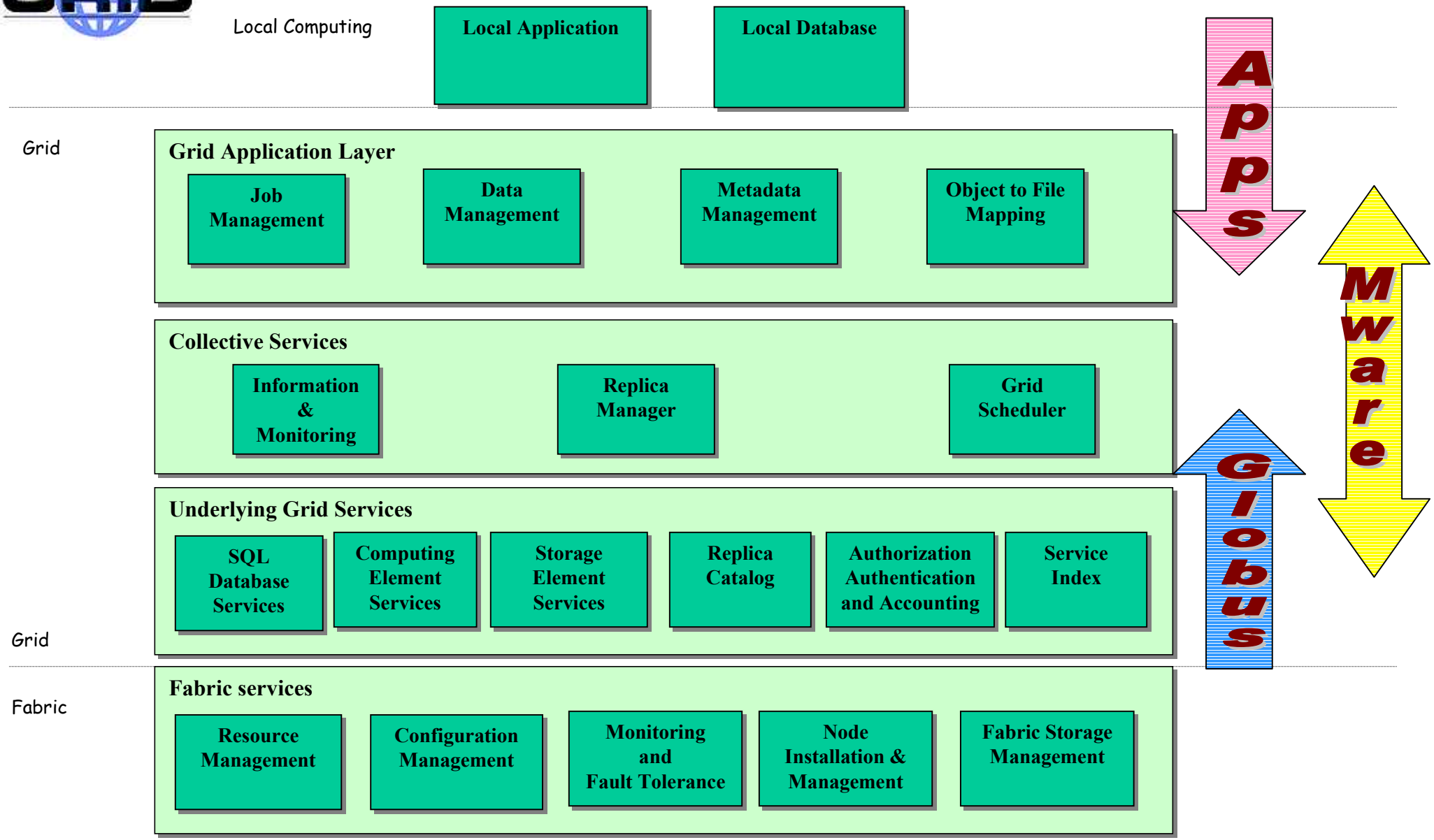
- First year objectives
- Architecture overview
- Details of each middleware Work Package
- Job submission example
- Architecture issues and actions
- Interaction with PPDG/GriPhyN/iVDGL
- Interaction with Globus & Globus Grid Forum (GGF)
- Plans for 2002
- Summary

# Objectives for the first year of the project

- Collect requirements for middleware
  - Take into account requirements from application groups
- Survey current technology
  - For all middleware
- Core Services testbed
  - Testbed 0: Globus (no EDG middleware)
- First Grid testbed release
  - Testbed 1: first release of EDG middleware

- WP1: workload
  - Job resource specification & scheduling
- WP2: data management
  - Data access, migration & replication
- WP3: grid monitoring services
  - Monitoring infrastructure, directories & presentation tools
- WP4: fabric management
  - Framework for fabric configuration management & automatic sw installation
- WP5: mass storage management
  - Common interface for Mass Storage Sys.
- WP7: network services
  - Network services and monitoring

# DataGrid Architecture



# EDG Interfaces

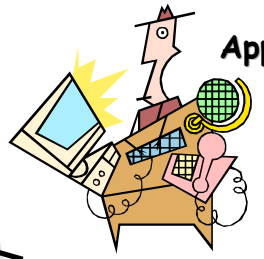
Application Developers

Local Application

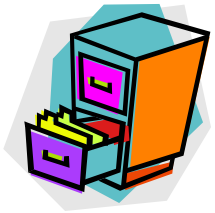
Local Database

Scientists

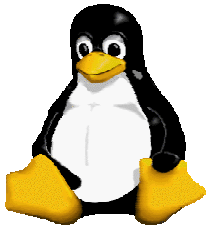
Certificate Authorities



System Managers



File Systems



Operating Systems

User Accounts

Grid Application Layer

Job Management | Data Management | Metadata Management | Object to File Mapping

Collective Services

Information & Monitoring | Replica Manager | Grid Scheduler

Underlying Grid Services

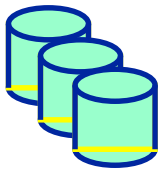
SQL Database Services | Computing Element Services | Storage Element Services | Replica Catalog | Authorization Authentication and Accounting | Service Index

Fabric services

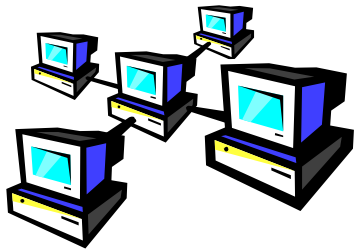
Resource Management | Configuration Management | Monitoring and Fault Tolerance | Node Installation & Management | Fabric Storage Management

Batch Systems

Mass Storage Systems
HPSS, Castor

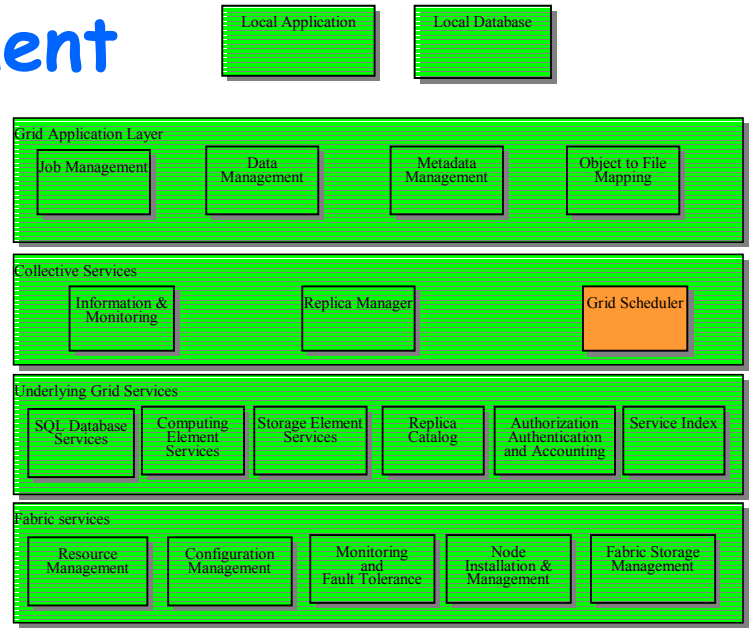


Storage Elements



Computing Elements

# WP1: WorkLoad Management




- Achievements
  - Analysis of work-load management system requirements & survey of existing mature implementations Globus & Condor (D1.1)
  - Definition of architecture for scheduling & res. mgmt. (D1.2)
  - Development of "super scheduling" component using application data and computing elements requirements
- Issues
  - Distributed nature of WP1 development groups

**Components**

Job Description Language

Resource Broker

Job Submission Service

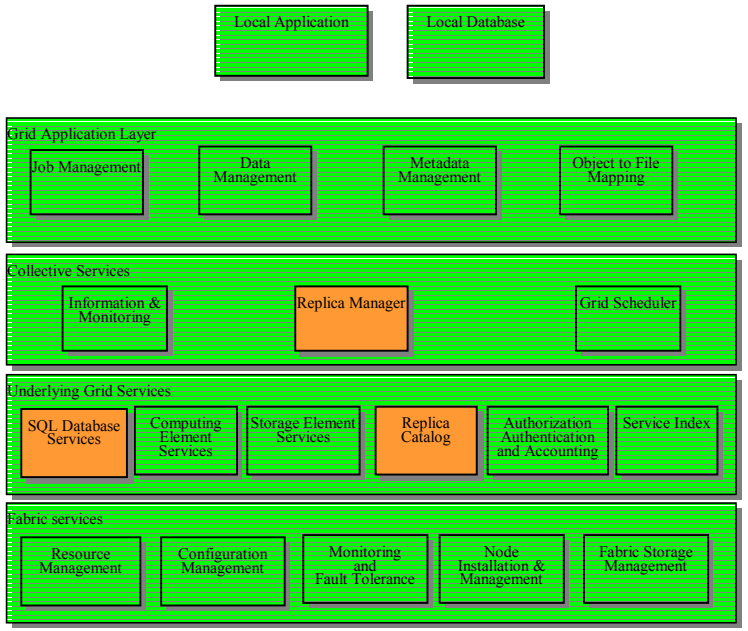
Information Index

User Interface

Logging & Bookkeeping Service

# WP2: Data Management

- Achievements
  - Survey of existing tools and technologies data access and mass storage systems (D2.1)
  - Definition of architecture for data management (D2.2)
  - Close collaboration with Globus, PPDG/GriPhyN & Condor
  - Working with GGF on standards
  - Deployment of GDMP in testbed1
- Issues
  - Security: clear methods handling authentication and authorization




**Components**

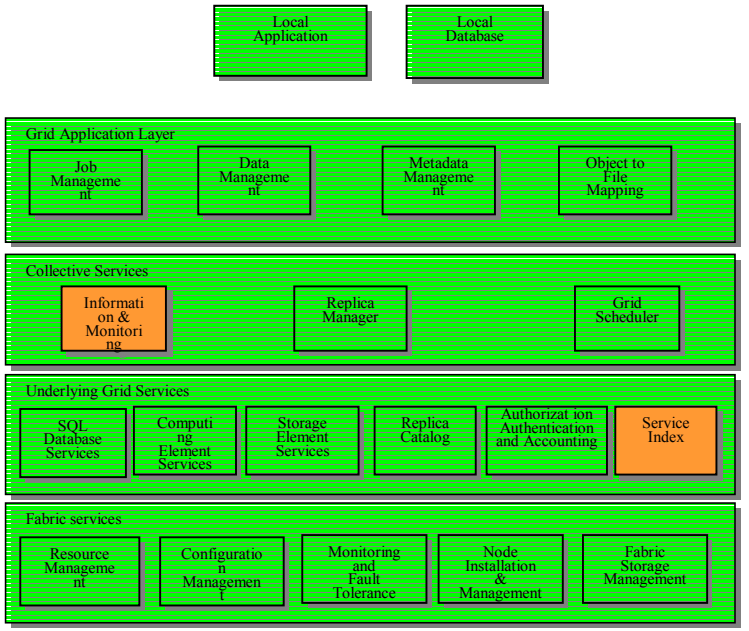
SpitFire

GDMP

Replica Catalog

# WP3: Grid Monitoring Services

- Achievements
  - Survey of current technologies for info & monitoring in grid environments (D3.1)
  - Coordination of schemas in testbed 1
  - Development of Ftree caching backend based on OpenLDAP to address shortcoming in MDS v1
  - Design of Relational Grid Monitoring Architecture (R-GMA) (D3.2) – to be further developed within the context of GGF
  - Collaboration with Globus for Ftree and PPDG/GriPhyN for res. discovery and schemas
  - GRM and PROVE adapted to grid environments to support end-user application monitoring
- Issues
  - R-GMA development



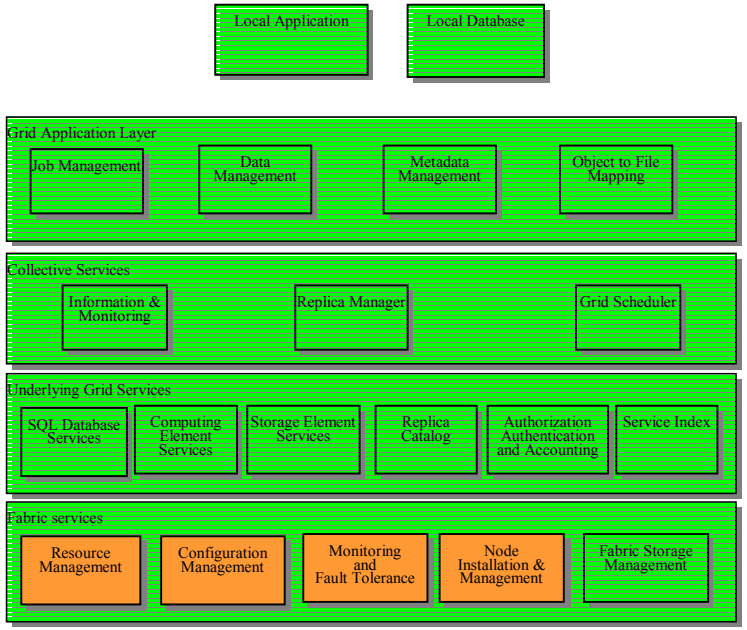

**Components**

MDS/Ftree

R-GMA

GRM/PROVE

# WP4: Fabric Management




- Achievements
  - Survey of existing tools, techniques and protocols for resource specification, configuration and management, as well as integrated cluster management suites (D4.1)
  - Defined an agreed architecture for fabric management (D4.2)
  - Initial implementations deployed at several sites in testbed 1
- Issues
  - Image Installation and Configuration Cache Manager components remain to be integrated

**Components**

LCFG

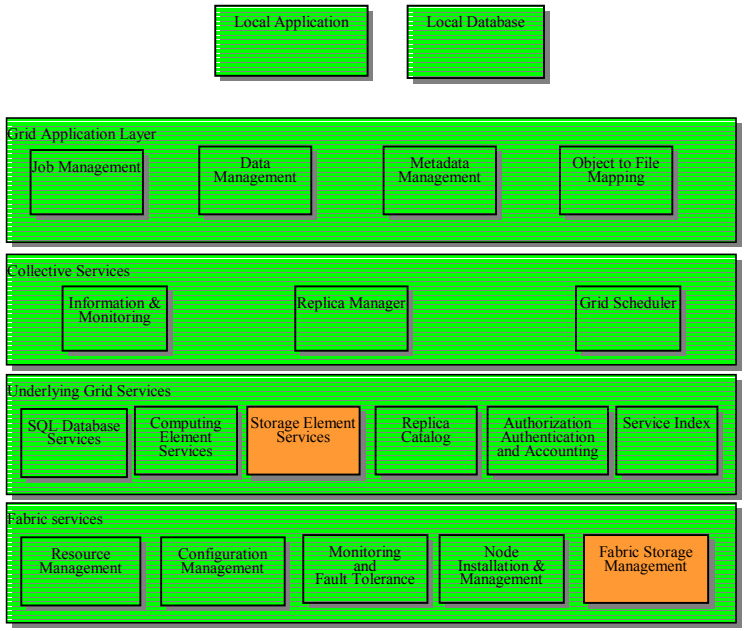
PBS & LSF info providers

Imagine installation

Config. Cache Mgr

# WP5: Mass Storage Management

- Achievements
  - Review of Grid data systems, tape and disk storage systems and local filesystems (D5.1)
  - Definition of Architecture and Design for DataGrid storage Element (D5.2)
  - Collaboration with Globus on GridFTP/RFIO
  - Collaboration with PPDG on control API
  - First attempt at exchanging HSM tapes
- Issues
  - Scope and requirements for storage element
  - Interworking with other Grids




**Components**

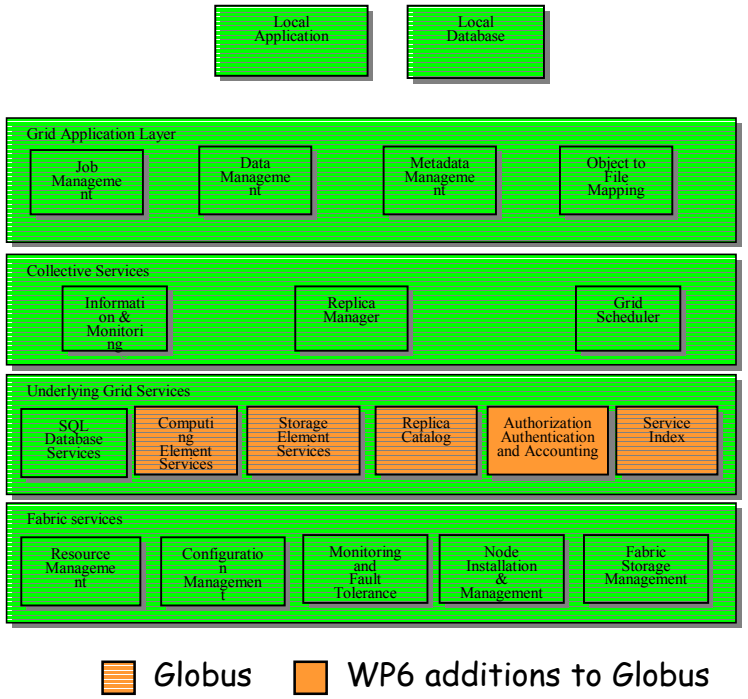
SE info providers

RFIO

MSS staging

# WP6: TestBed Integration

- Achievements
  - Integration of EDG sw release 1.0 and deployment
  - Working implementation of multiple VOs & basic security infrastructure
  - Definition of acceptable usage contracts and creation of Certification Authorities group
- Issues
  - Procedures for software integration
  - Test plan for software release
  - Support for production-style usage of the testbed




**Components**

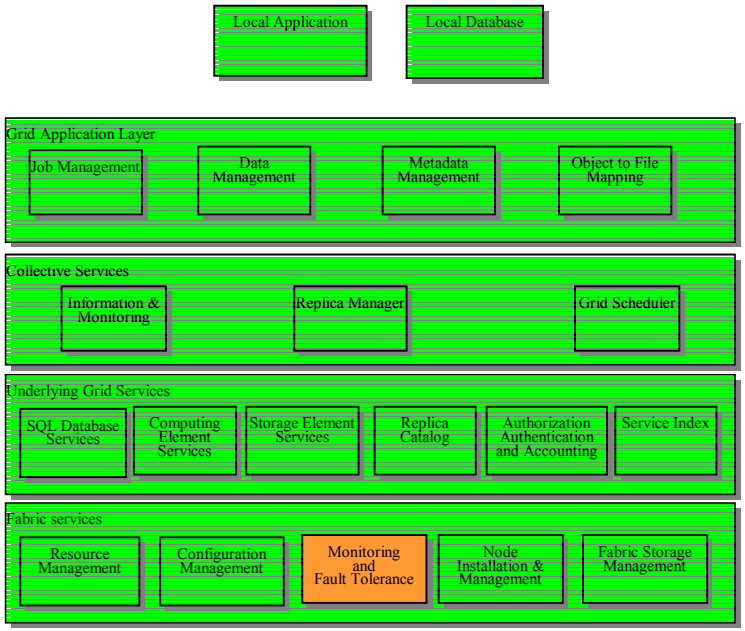
Globus packaging & EDG config

Build tools

End-user documents

# WP7: Network Services

- Achievements
  - Analysis of network requirements for testbed 1 & study of available network physical infrastructure (D7.1)
  - Collaboration with Dante & DataTAG
  - Working with GGF (GHPN) & Globus (monitoring/MDS)
  - Use of European backbone GEANT since Dec. 2001
  - Initial network monitoring architecture defined(D7.2) and first tools deployed in testbed 1
- Issues
  - Resources for study of security issues
  - End-to-end performance for applications depend on a complex combination of components




**Components**

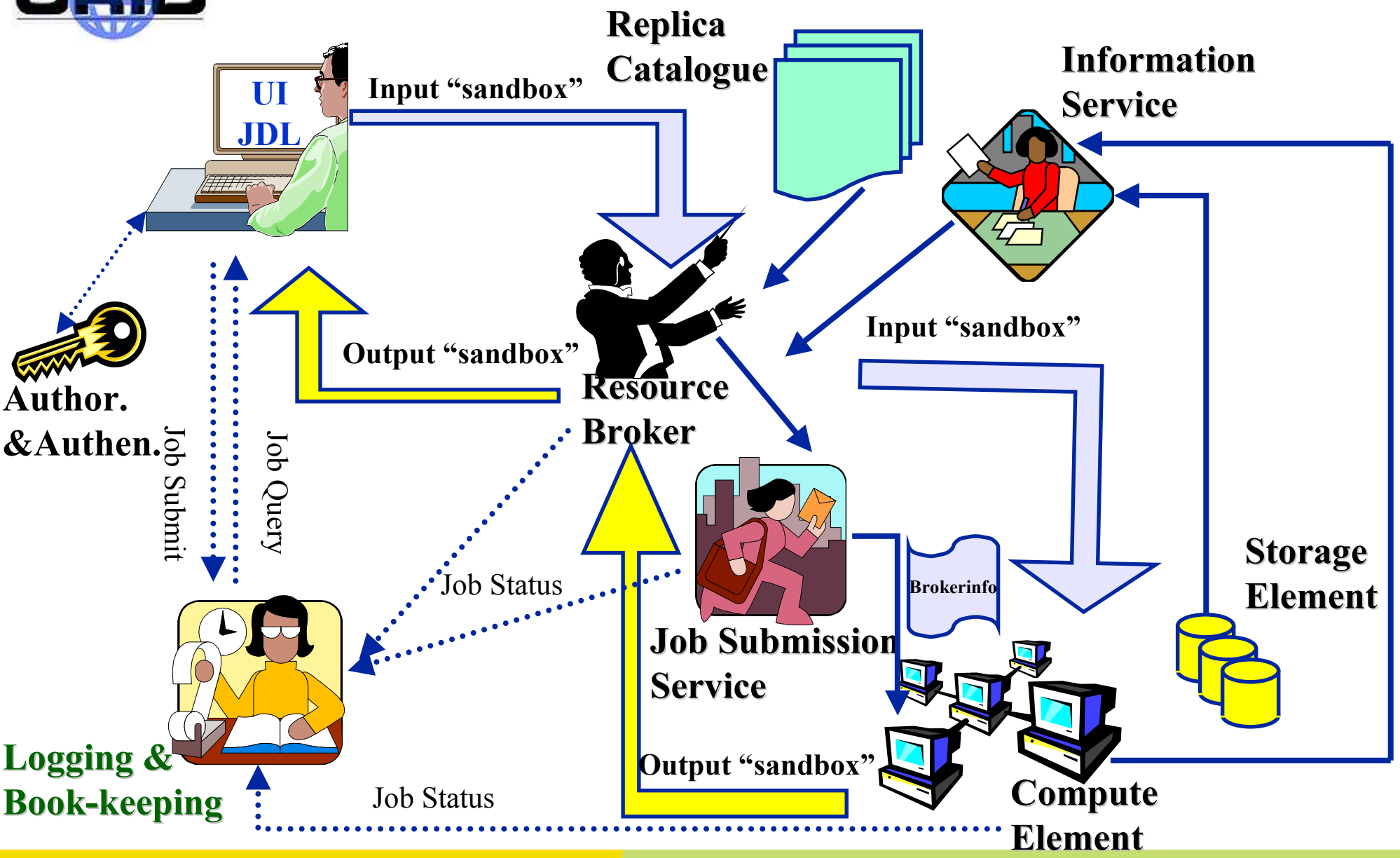
network monitoring tools:

PingER

Udpmon

Iperf

# A Job Submission Example



# Architecture Issues and Actions

- Some concepts remain vague
  - e.g. interactive jobs
- Some boundaries are unclear
  - e.g scope/functionality of a Storage Element
- Some requirements are not yet addressed
  - e.g. anonymous users
- The various software components are not yet fully integrated
  - e.g. Storage Element, Computing Elements & Info. Sys.
- Short term/ Long term trade-offs
- Impact of Open Grid Services Architecture
  - Forthcoming developments by Globus/IBM/GGF
- Convergence with US Grid project activities (PPDG/GriPhyN)

- The new **architecture group** will address these points taking into account our experience from testbed1 and further requirements
- Implementation of **iterative releases, nightly builds** and separation of development testbed from production testbed

# Plans for 2002

- Extension of testbed
  - more users, sites & nodes-per-site
  - Split testbed into development and production sites
  - Investigate inter-operability with US grids
- Iterative releases up to testbed 2
  - incrementally extend functionality provided via each Work Package
  - better integrate the components
  - improve stability
- Testbed 2 (fall 2002) extra requirements
  - Interactive jobs
  - Job partitioning for parallel execution
  - Advance reservation
  - Accounting & Query optimization
  - Security design (D7.6)
  - . . .

- Planned intermediate release schedule
  - **TestBed 1:** November 2001
  - **Release 1.1:** January 2002 ⇐ demos
  - **Release 1.2:** March 2002
  - **Release 1.3:** May 2002
  - **Release 1.4:** July 2002
  - **TestBed 2:** September 2002
- Similar schedule will be made for 2003
- Each release includes
  - feedback from use of previous release by application groups
  - planned improvements/extension by middle-ware WPs
  - use of WP6 *software infrastructure*
  - feeds into architecture group

# Interaction with PPDG/GriPhyN/iVDGL

- Work with dataTAG via the InterGrid to investigate inter-operability of US and EU grids
  1. **Authentication infrastructure** –perform cross organizational authentication
  2. **Unified service discovery and information infrastructure** – discover the existence and configuration of service offered by the testbeds.
  3. **Data movement infrastructure** – move data from storage services operated by one organization to another
  4. **Authorization services** – perform some level of cross organization, community based authorization
  5. **Computational services** – coordinate computation across organizations – to allow submission of jobs in EU to run on US sites and vice versa

# Interaction with Globus & GGF

- **Software Licensing** – open source agreements for all components
- **WP1**
  - Advance Reservation Infrastructure
- **WP2**
  - GDMP joint-development and overlap with plans for future Globus Replica Manager
  - GridFTP/NetLogger integration
- **WP3**
  - MDS/Ftree integration
  - Relationship between R-GMA, GGF and OGSA
- **WP4**
  - Authorization capabilities of the Globus gatekeeper
  - Use resource mgmt subsystem instead of Globus job manager
- **WP5**
  - Thread-safe GSI API
  - RFIO using GridFTP
- **WP6**
  - Packaging
  - Community Authorization Service
- **WP7**
  - Integration of MapCentre with MDS
  - Network message publication to Info. Services

# Summary

- Application groups requirements defined and analysed
- Extensive survey of relevant technologies completed and used as a basis for EDG developments
- First release of the testbed successfully deployed
- Excellent collaborative environment developed with key players in Grid arena
- Project can be judged by:
  - level of "buy-in" by the application groups
  - wide-spread usage of EDG software
  - number and quality of EDG sw releases
  - positive influence on developments of GGF standards & Globus toolkit