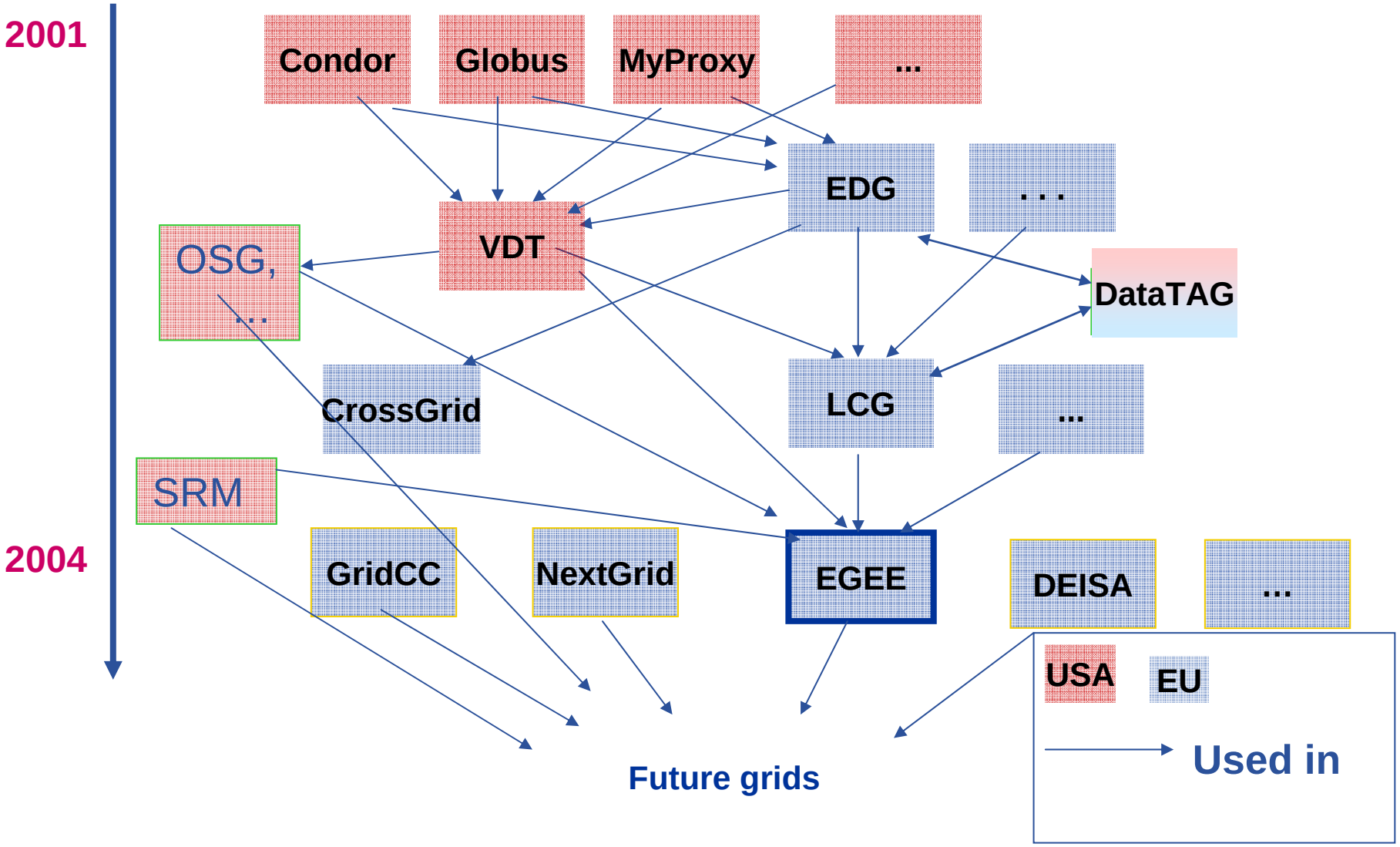
Introduction to gLite

George Goulas,
Computer Systems Laboratory,
University of Patras

gLite

- **Service-Oriented middleware**
  - Uses Web services standards
- **Official name: EGEE Middleware**
- **Based on previous middleware efforts**
- **Re-engineering / redesign to offer**
  - Scalability
  - Performance
  - Interoperability
  - Modularity
  - (…) the perfect grid middleware :-)
- **User requirements input from various user communities, mostly HEP & BioMed**

gLite in grid middleware “history”



New features

- **Increased modularity**
  - Services can be deployed independently
- **XML based configuration**
- **Finer grained security (VOMS)**
- **Pull mode option for job scheduling (lazy scheduling)**
- **POSIX IO for files located in grid storage**
- **User friendly file names (LFN)**
- **File Transfer Services (Data Management jobs)**

Site Services

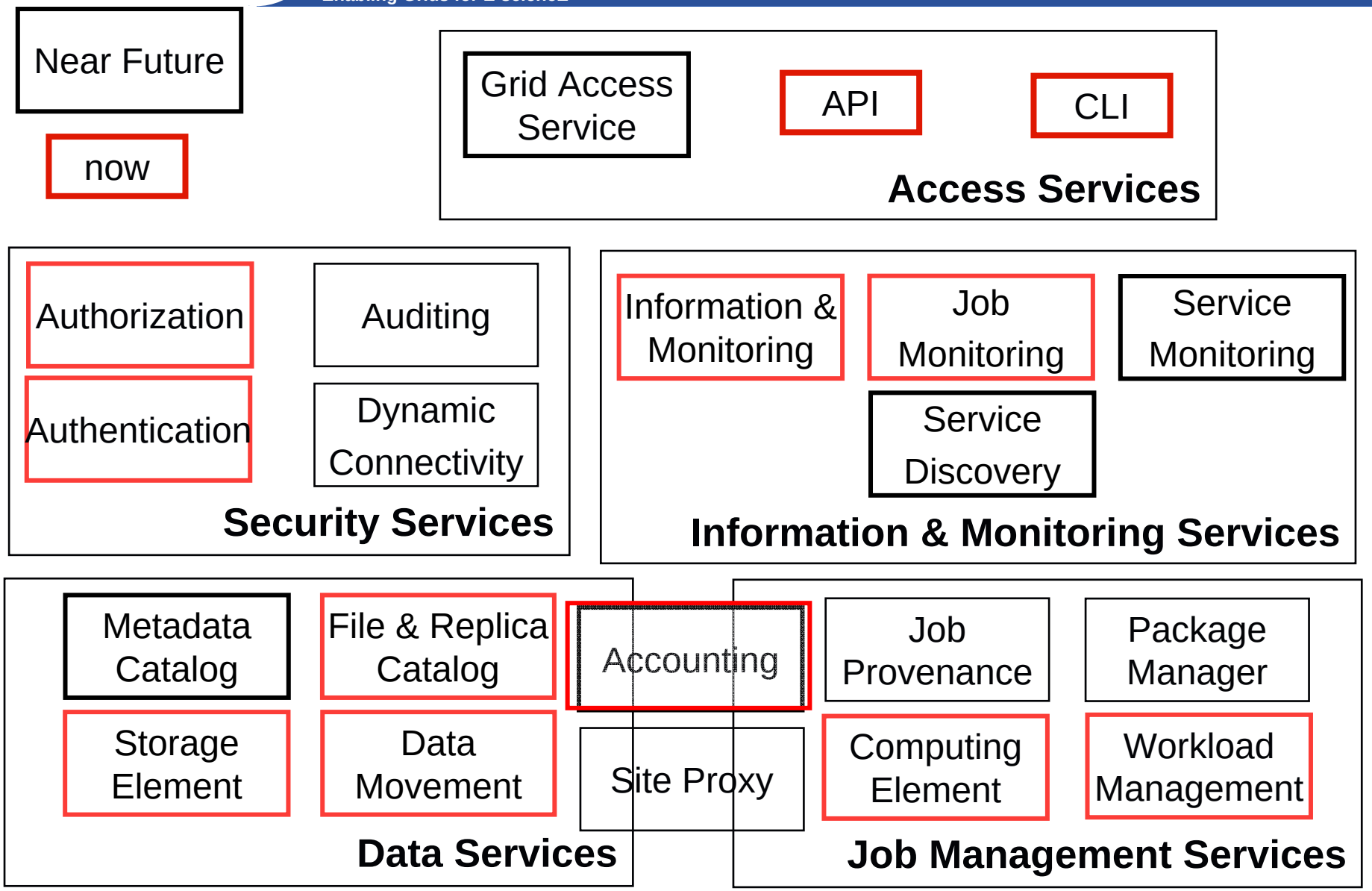
- **Computing Element (CE)**
  - Gateway to local computing resources (worker nodes cluster)
  - Local Resource Management System (LRMS)
- **Worker Nodes (WN)**
- **Storage Element**
  - Gateway to local storage
  - Disk or tape based storage
  - POSIX like Interface (through IO Server)
- **User Interface (UI)**
  - User access point to the grid
  - A set of client programs to grid services

Layer of Abstraction (CE, SE): Local peculiarities irrelevant

Grid / VO-wide Services

- **Security Services**
  - Virtual Organization Membership Service (VOMS)
  - MyProxy Server
- **Information System (IS)**
  - R-GMA Registry Server
  - R-GMA Schema Server
  - R-GMA Server
- **Job Management**
  - Workload Management System (WMS)
  - Logging & Bookkeeping (LB)
  - Usually co-located on a machine, WMS+LB
- **Data Management**
  - File Catalog (FiReMAN)
  - File Transfer Service (FTS)
  - File Placement Service (FPS)

gLite Components Overview



Virtual Organization Membership Service

- **Multiple VOs**
- **Multiple roles in a VO**
  - X509 compatible extensions
  - Signed by the VOMS server
- **Web admin interface**
- **Supports MyProxy**
- **Resource Providers grant access to a VO or a VO role**
- **Sites map VO members / roles to local auth mechanism**
  - Allows for local policy
  - Remember that the grid should not alter local security policies

Layer of Abstraction: Individual members irrelevant

MyProxy

- **Grid Vulnerability: A user proxy can be stolen from a UI**
- **Users should not sign long lived proxies**

MyProxy

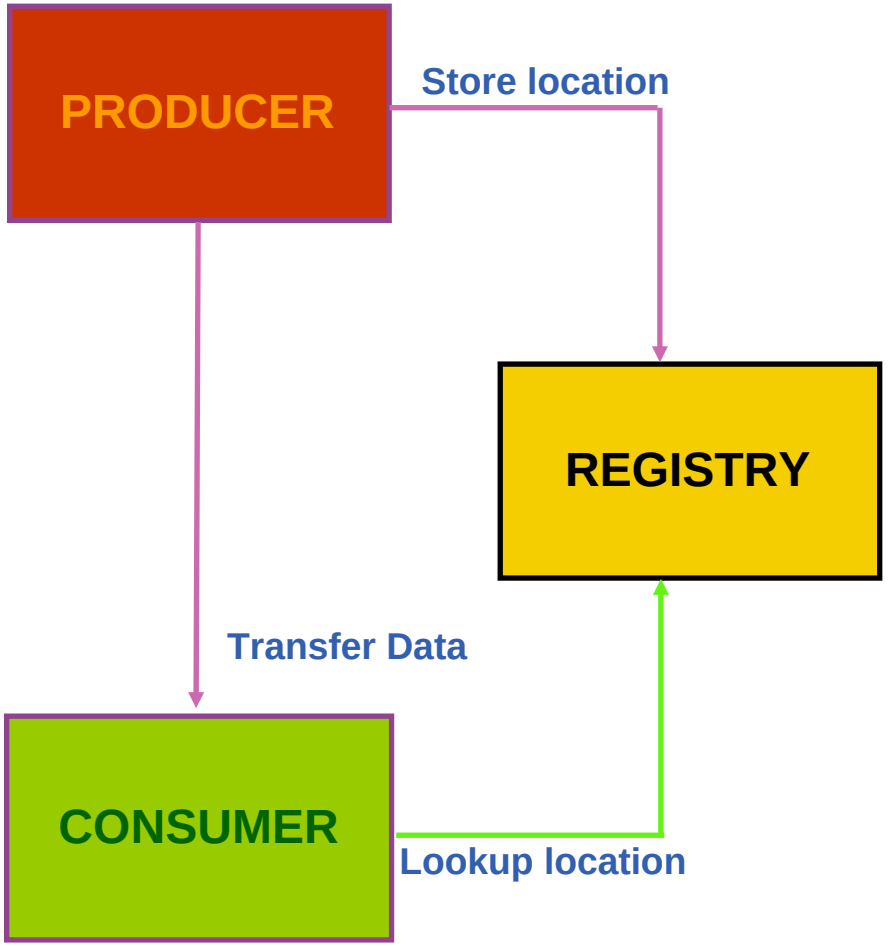
- **Provides a certificate store for medium-lived proxies (days ~ weeks)**
- **Allows for secure user mobility**
  - No need to copy globus-keys on different UI machines

Information System

- **Based on R-GMA**
  - Implements the GGF Grid Monitoring Architecture (GMA)
  - Relational – GMA
  - Like a distributed database
- **Aggregates service information from multiple grid sites**
  - Hosts, resources (CPU, storage)
  - VO information
- **Used by WMS to collect information on sites**
  - Defines the WMS view of the Grid
- **Generic Service Discovery API**
  - Used by replica management tools to locate SEs, Catalogs

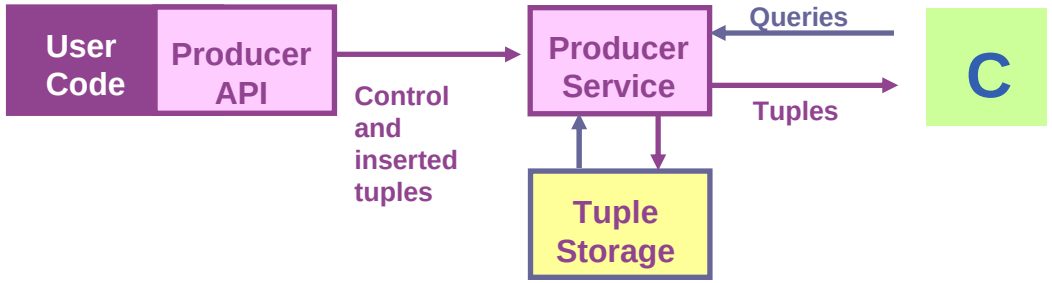
Grid Monitoring Architecture

- The Producer stores its location (URL) in the Registry.
- The Consumer looks up producer URLs in the Registry.
- The Consumer contacts the Producer to get all the data or the Consumer can listen to the Producer for new data.

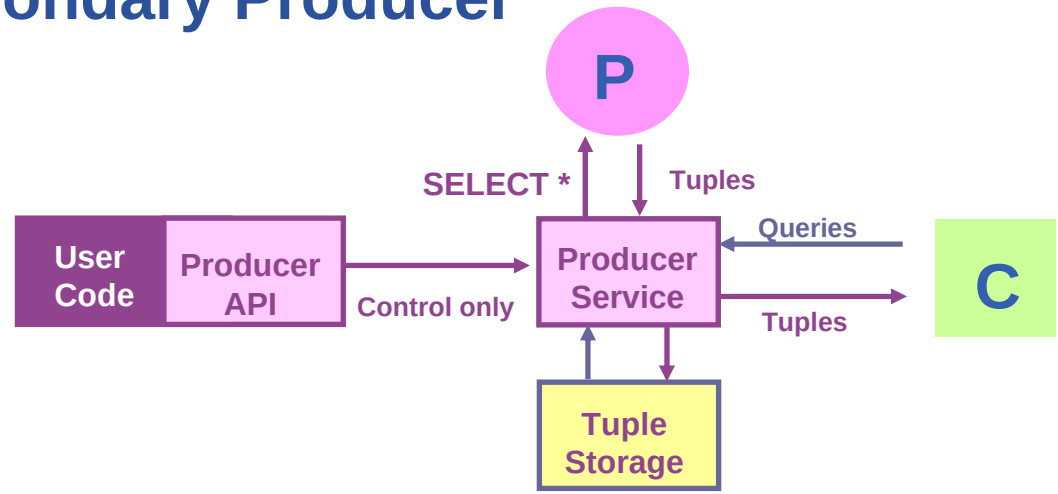


Producer Types

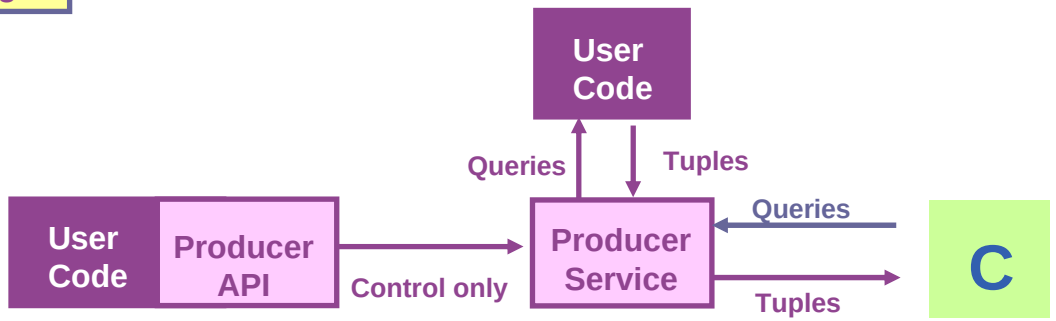
- **Primary Producer**



- **Secondary Producer**

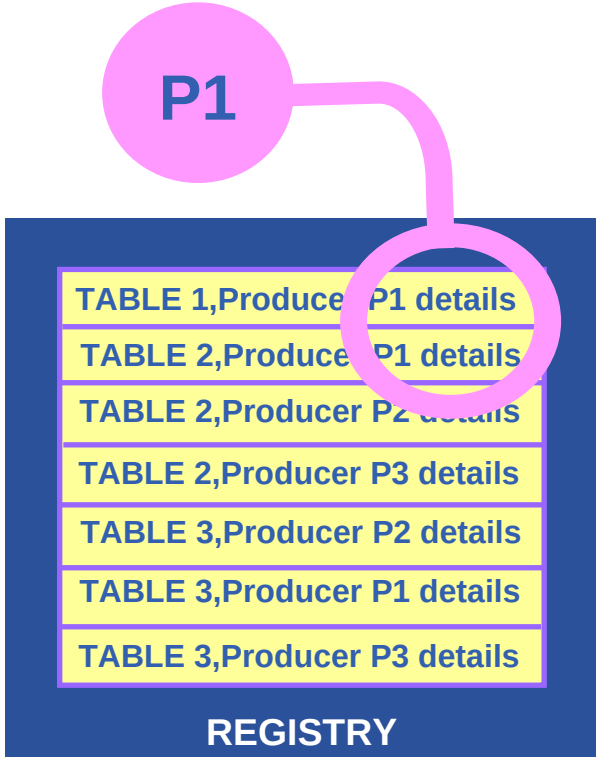


- **On-Demand Producer**

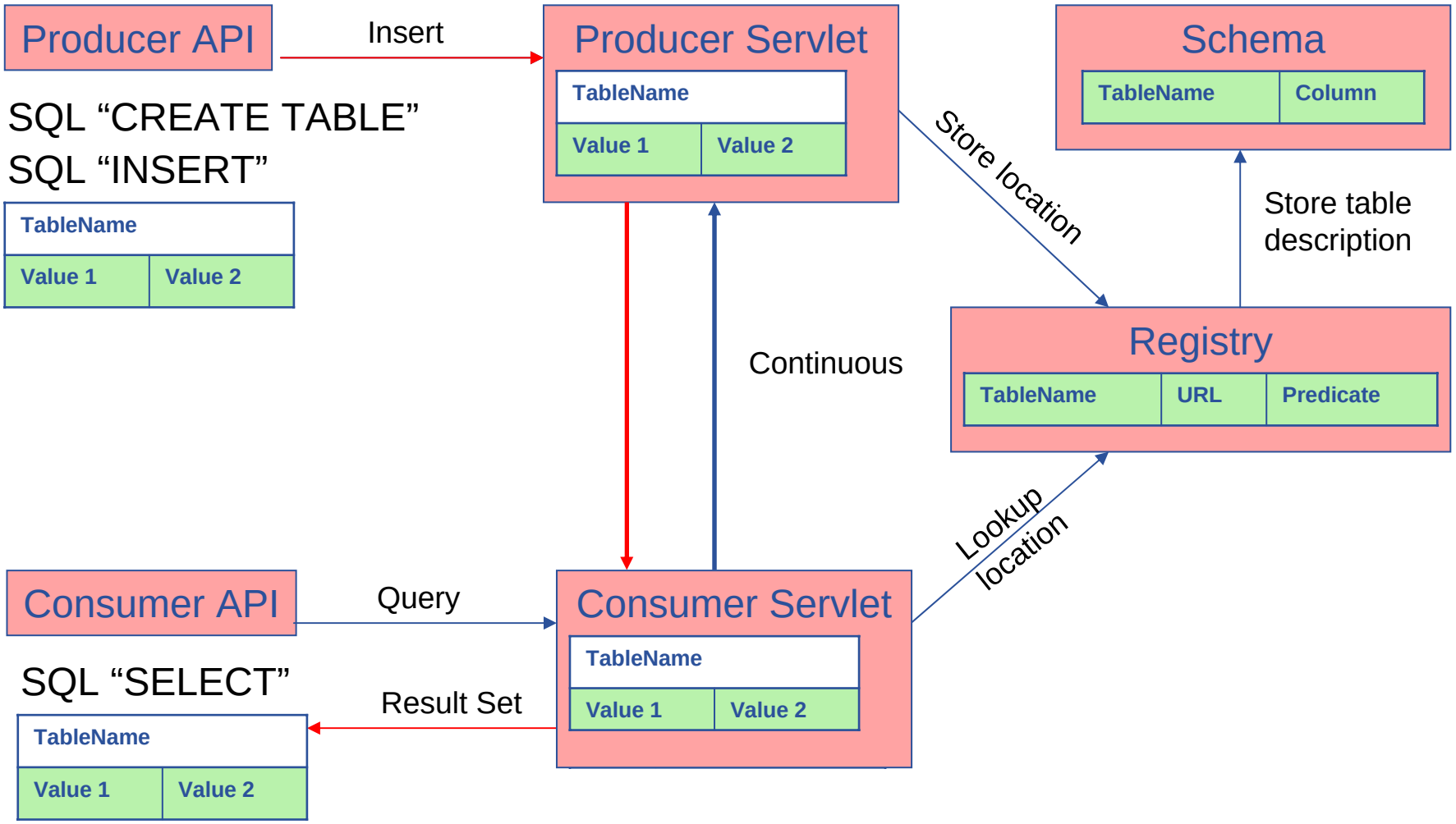


Query Types

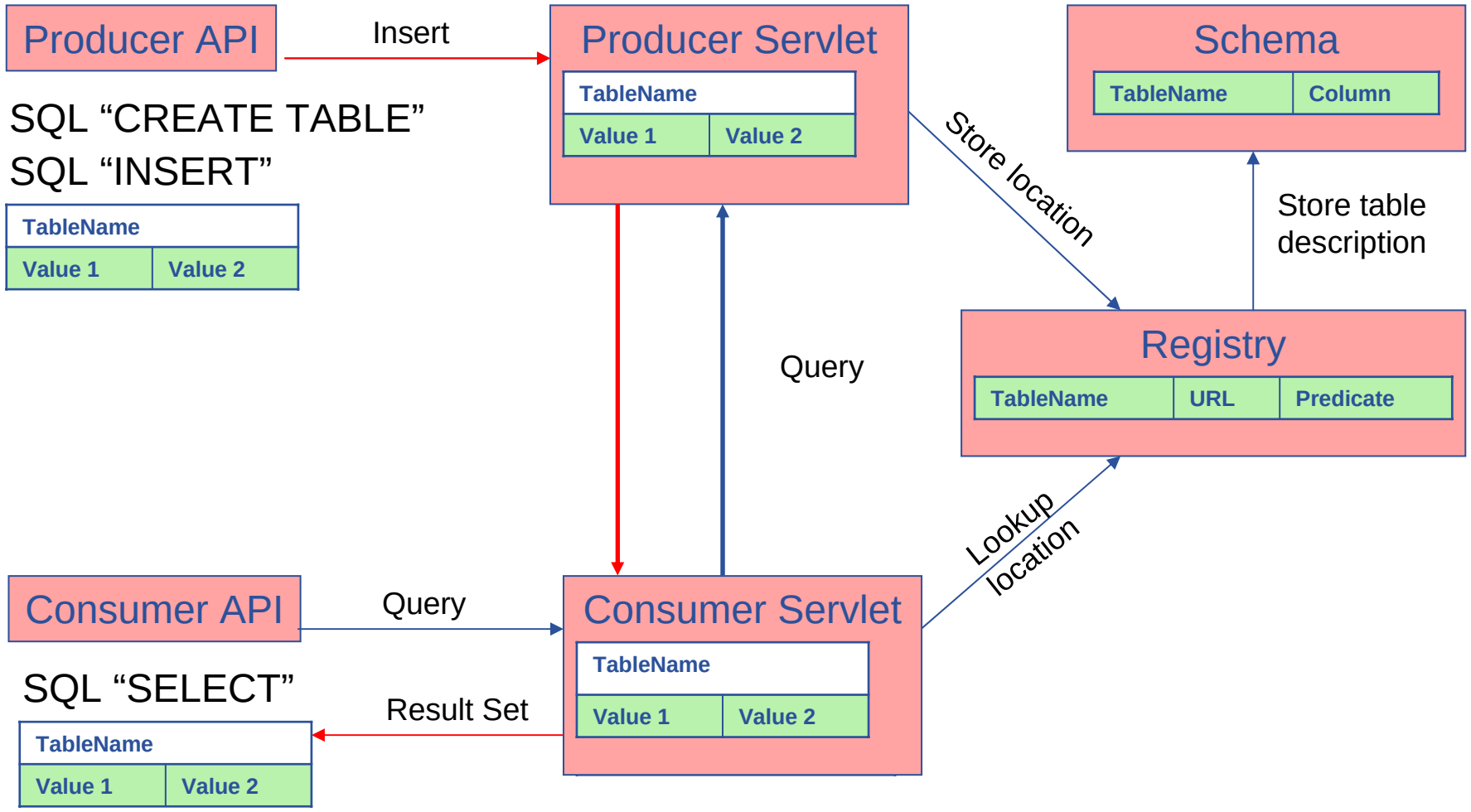
- Continuous
- Latest
- History
- Static



Continuous

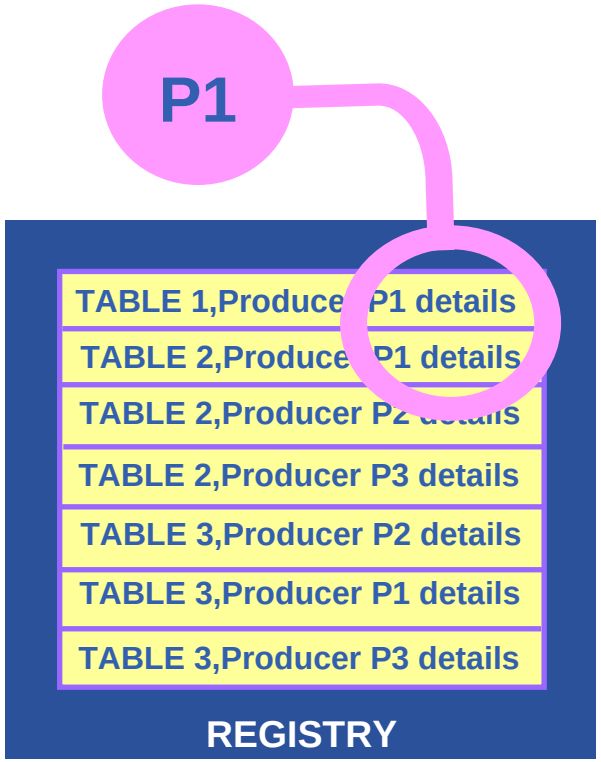

Producer API
Insert
Producer Servlet
TableName
Value 1
Value 2
Schema
TableName
Column
SQL “CREATE TABLE”
SQL “INSERT”
TableName
Value 1
Value 2
Store location
Store table description
Registry
TableName
URL
Predicate
Continuous
Lookup location
Consumer API
Query
Consumer Servlet
TableName
Value 1
Value 2
SQL “SELECT”
Result Set
TableName
Value 1
Value 2

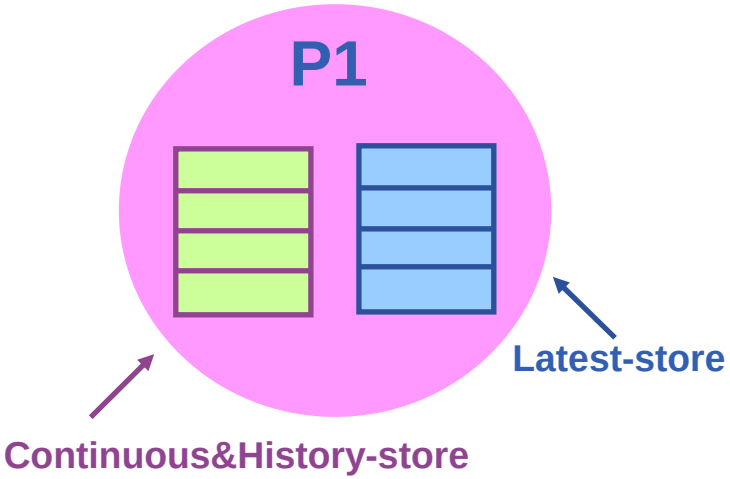
History or Latest



Query Types

- Continuous
- Latest
- History
- Static





Latest Retention Period

History Retention Period

Job submission

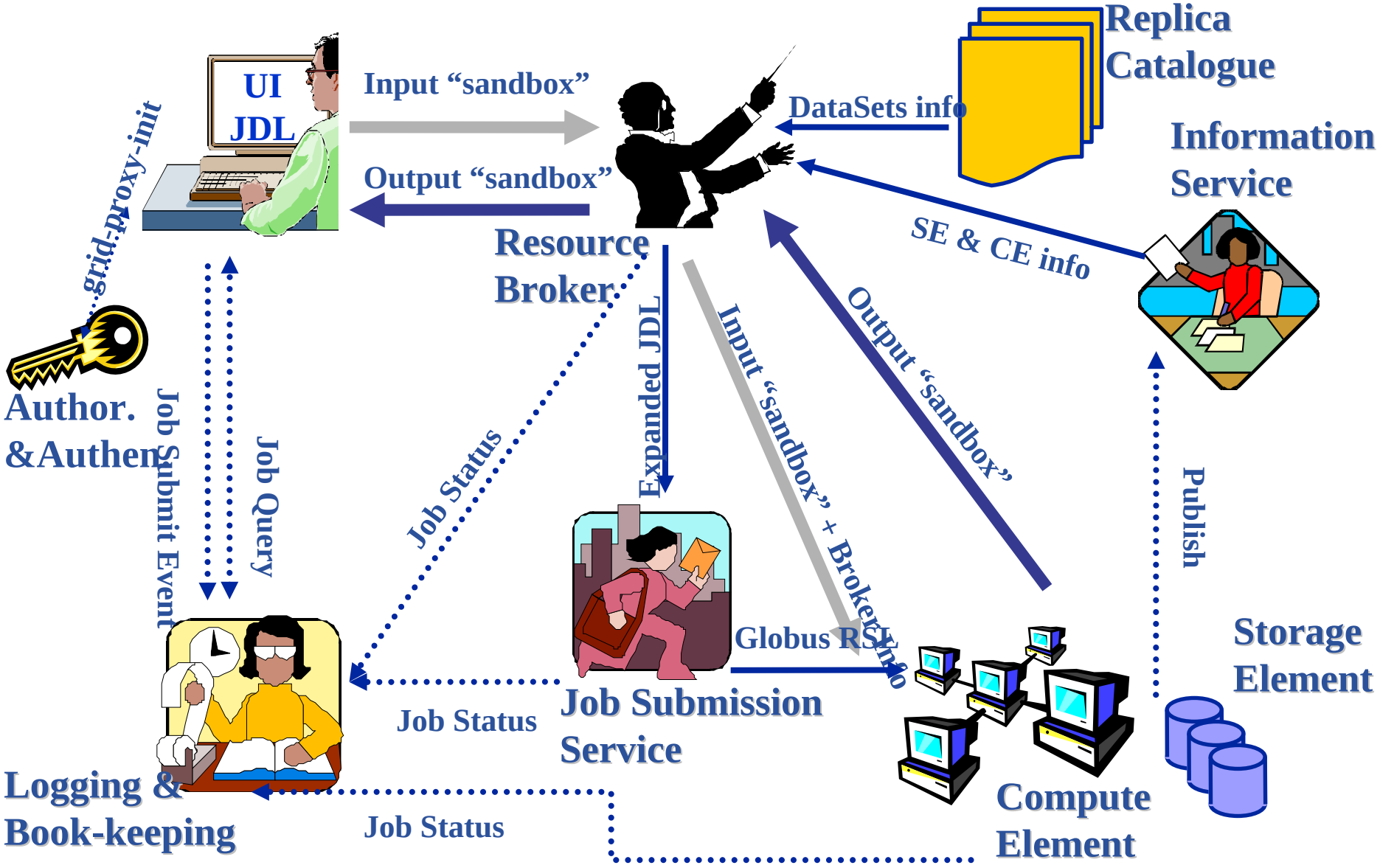
WMS

User submits a job from the UI to the WMS

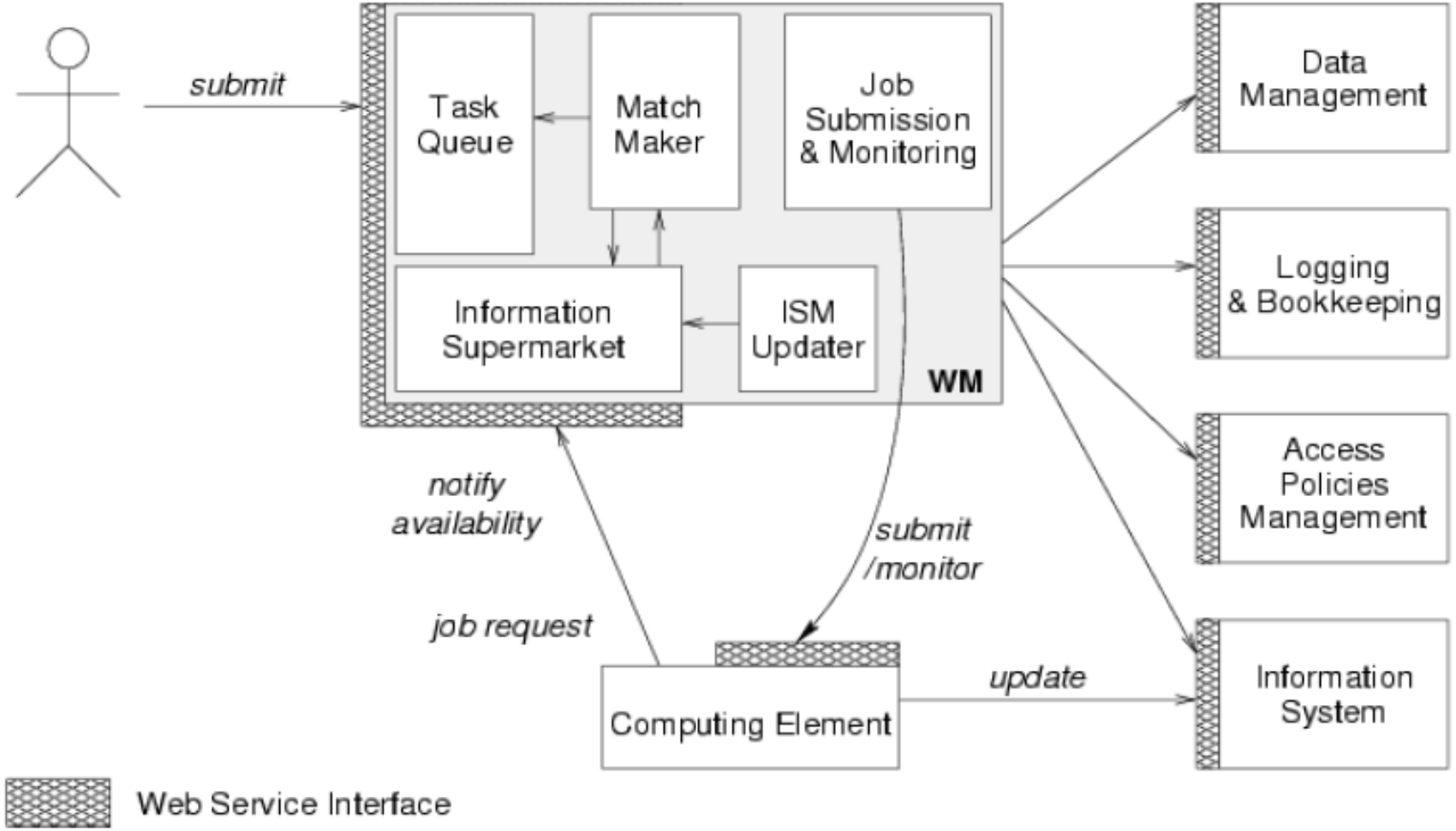
- **WMS finds the best location of a job**
  - Considering job requirements and available resources
- **PUSH operation mode**
  - WMS finds the CEs from the Information System
  - Full scheduling model
- **PULL operation mode**
  - The CEs ask the WMS for jobs
  - Lazy scheduling model
- **Logging & Bookkeeping keeps track of job status**

Layer of abstraction: Sites irrelevant

A typical job workflow



Workload Management System



Data Management Services

- **Storage Element**
  - Storage Resource Manager (not part of gLite)
  - POSIX-IO (gLite IO Server)
  - Access protocols

- **Catalogs**
  - File Catalog
  - Replica Catalog
  - File Authorization service
  - Metadata catalog

gLite FiReMAN catalog (for MySQL and Oracle)

gLite standalone metadata catalog

FiReMAN: File & Replica MANager

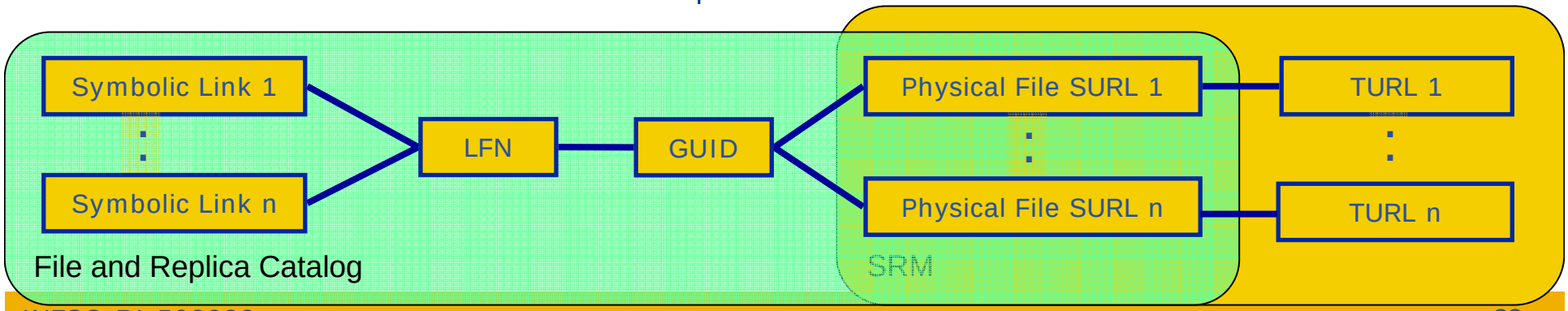
- **File Transfer**
  - File Transfer Service / File Placement Service
  - No separate services, as of gLite 1.4

Data Management

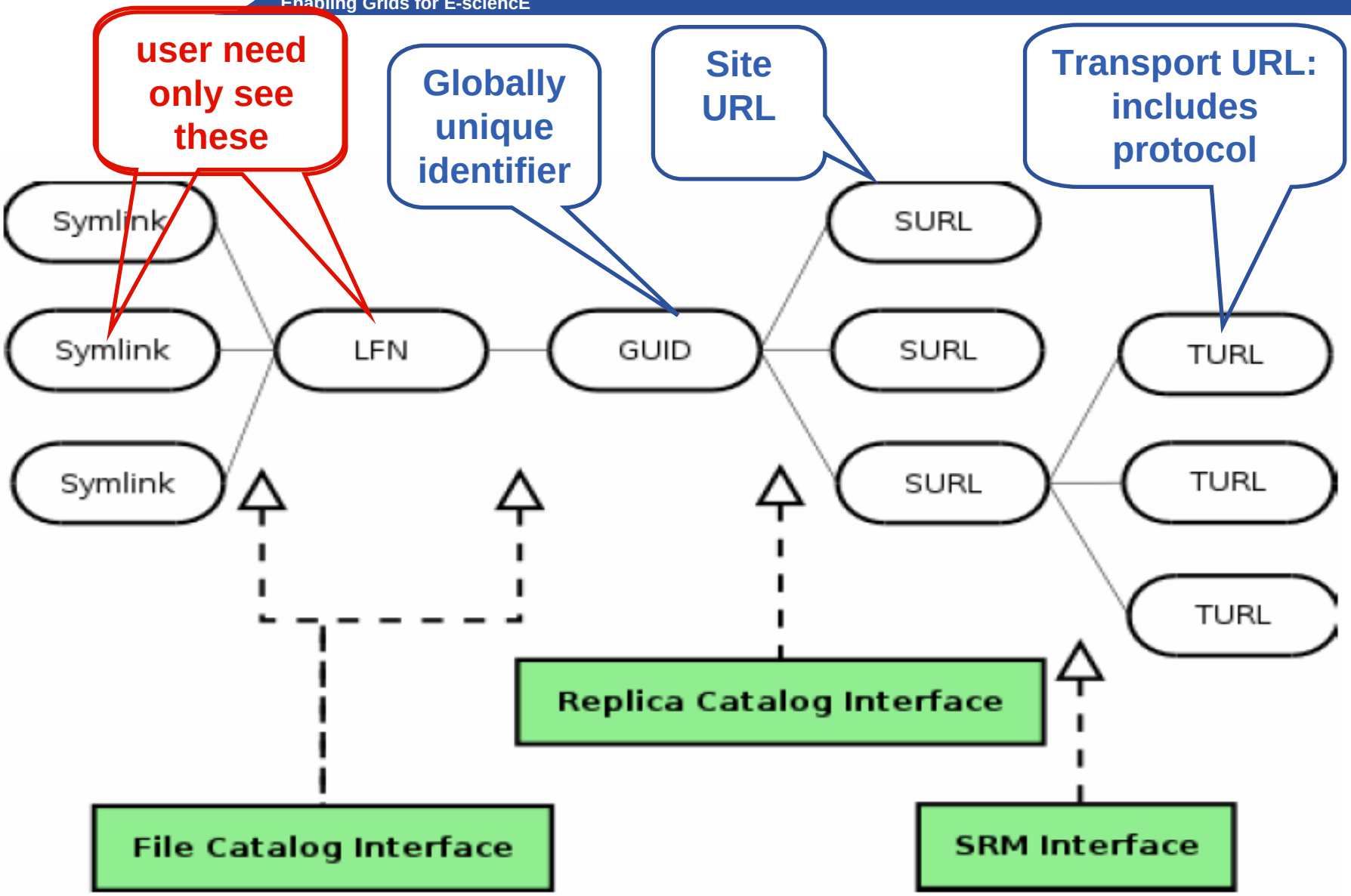
- **Catalog remembers location of files**
  - Only deals with their locations (not data, not transfers!)
  - Data transfer handled separately: PFNs point to actual storage location and access protocol
- **Files can be replicated on multiple SEs**
- **Each file registered has unique ID**
  - Same files get different ID when registered multiple times
- **LFNs are names that make sense to the user**
- **Layer of abstraction: file location irrelevant**

Files & replicas: Name Conventions

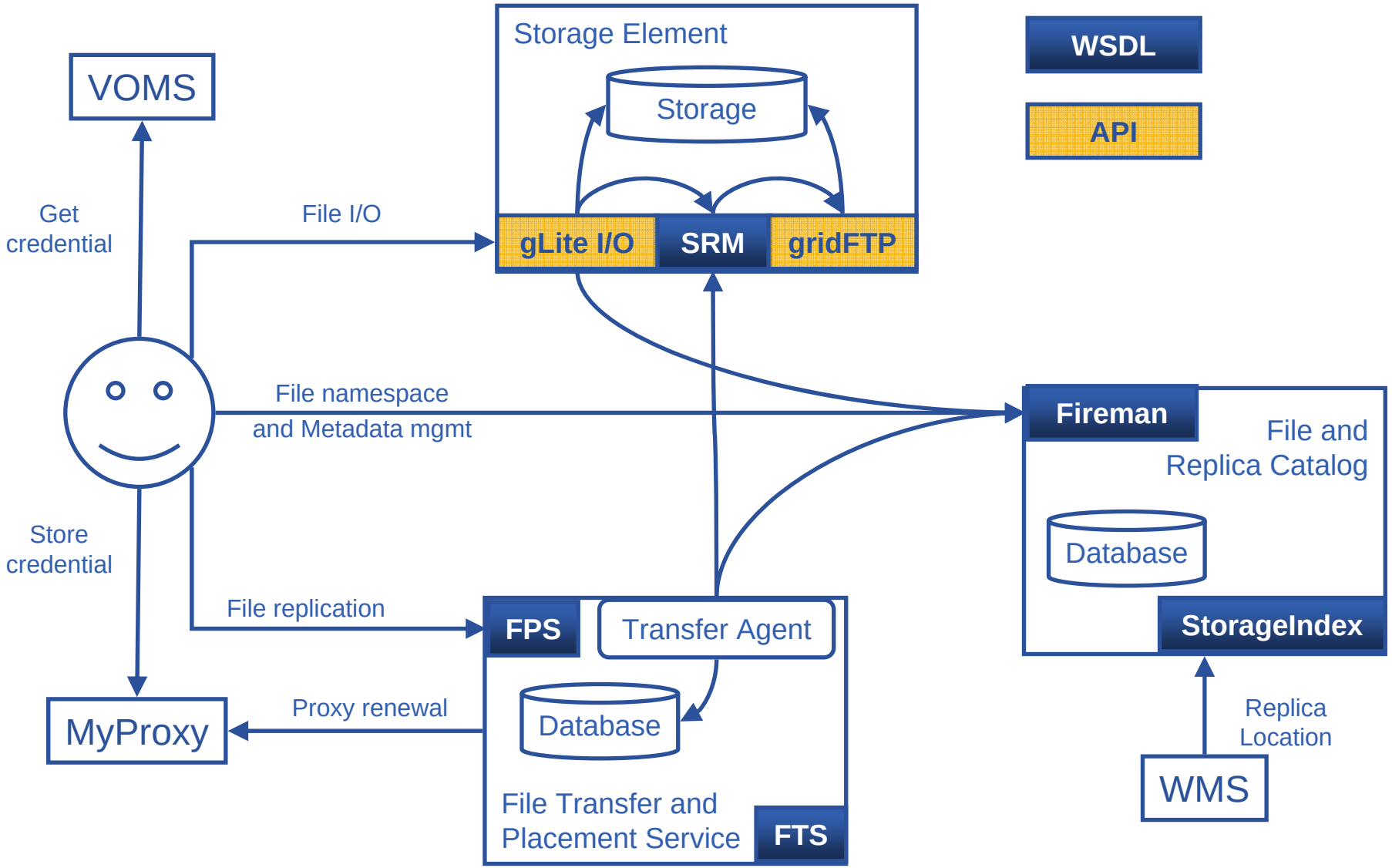
- **Symbolic Link** in logical filename space
- **Logical File Name (LFN)**
  - An alias created by a user to refer to some item of data, e.g. "lfn:cms/20030203/run2/track1"
- **Globally Unique Identifier (GUID)**
  - A non-human-readable unique identifier for an item of data, e.g. "guid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6"
- **Site URL (SURL) (or Physical File Name (PFN) or Site FN)**
  - The location of an actual piece of data on a storage system, e.g.
    "srm://pcrd24.cern.ch/flatfiles/cms/output10_1" (SRM)
    "sfn://lxshare0209.cern.ch/data/alice/ntuples.dat" (Classic SE)
- **Transport URL (TURL)**
  - Temporary locator of a replica + access protocol: understood by a SE, e.g.
    "rfio://lxshare0209.cern.ch//data/alice/ntuples.dat"



File names and identifiers in gLite



Data Management Interactions



gLite I/O

- **POSIX-like access to grid files**
- **Both CLI & API**
- **GUID or LFN can be used**
  - open(“/goulas/grid/myFile.txt”);

File Transfer Service

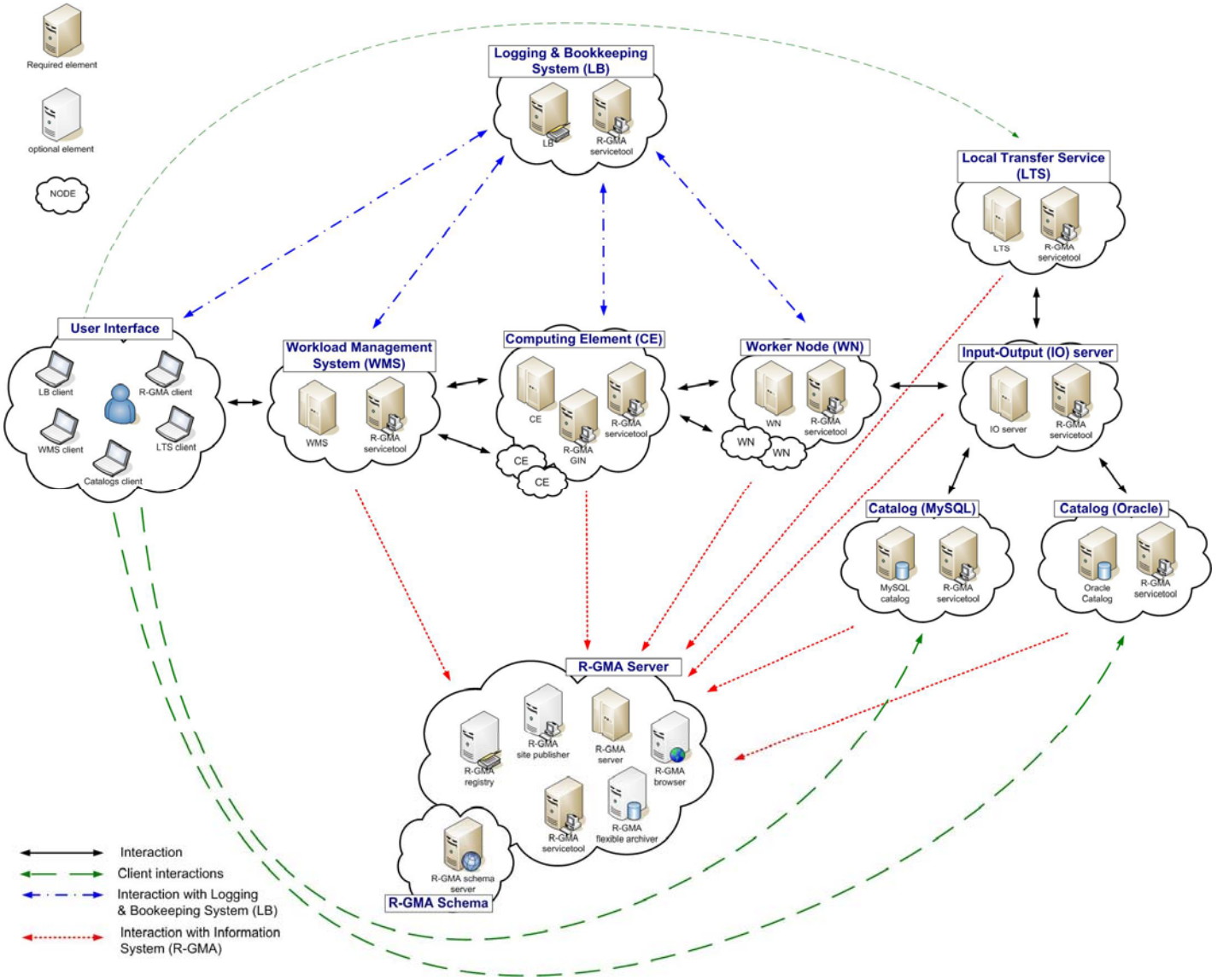
File Transfer Service

- **Handles data management jobs**
  - "Resource Broker" for data jobs
- **Responsible for reliable file transfers between grid sites**
  - Transfers (set of files) between 2 SE's
  - Endpoints with same protocol (gsiftp, ..)
- **Can be shared among VOs**
- **Channels**
  - Point to point queues
  - State
  - Bandwidth
  - Concurrent transfers

File Placement Service

- **Understands logical source files**
  - lfn:///goulas/myfile.txt
- **Understands logical destination**
  - Transfer to cern.ch
- **Updates the File Catalogs**
  - Registers new replica SURL in FiReMAN
- **Builds on FTS**

gLite services Deployment Scenario



Summary

- **More Standards compliant (WS)**
- **More security, virtualization of resources**

- **Some components evolving keeping compatibility**
- **Commands renamed in a uniform manner, same functionality**
- **New / re-architected components**

- **Still evolving**
  - Presentation based on gLite 1.4
  - Current version: gLite 1.5, released Thu, 19/1/06 (last week)