The D0 Grid Architecture
(Early Draft)

Introduction

This document is the beginning of the D0 Grid architecture definition. In its early stage, this architecture is driven not only by the D0 needs, but also by the availability of production-quality D0 software such as SAM, as well as by our understanding of the standard Grid technologies and their prospects. Naturally, as the Global Grid effort makes progress the architecture will evolve.

The principal novelties of this architecture, when compared to European DataGrid architecture, as well as other Grid projects, are as follows:

1. The existing D0 data handling system, SAM, is a mature and robust system. When properly abstracted, it plays a principal role in the overall architecture and in particular, has direct effects on the Job Management services.
2. We assign a critical role to the Monitoring and Information Services. We widen the boundaries of this principal component to include every service that provides, or receives, information relevant for job and data management. For example, we include the job logging and bookkeeping service into this area, as well as the popular concept of replica catalog.
3. We aggressively base the preliminary job handling architecture on the Condor technology, relying on the collaboration with the University of Wisconsin Computer Scientists within PPDG and beyond.

The purpose of this draft is to stimulate discussions and especially to focus on the *information flow* within the system, because any progress in the job handling area and in the project in genera will hinge sooner or later on the progress in the area of information services.

Principal Components

We identify three principal components (see the Figure), roughly aligned with the three principal areas of work in the D0 Grid project. These are (1) Job Handling, or Job Management, or Workload Management, (2) Monitoring and Information Services, (3) The Data Handling System.

Job Handling

The sub-architecture of this component follows that of the EDG and we will not repeat its description (we might outline it in the final draft however). There are two major differences, however.

First, we would like not to assign too much responsibility to the Request Broker (Resource Broker in EDG terminology) which has a danger of incorporating all the undecided, and therefore obscure, intelligence into a single piece (which is what D0 SAM did with the “global” resource optimizer), as well as the danger of deviating from the standard grid technologies toward D0-specific framework. At least initially, we’ll base the RB almost entirely on Condor’s Match-Making Service (our acronym - MMS). Such a usage of the MMS is contingent on the projected extensions of the Condor software.

Second, the Resource Broker will not take low-level input from computing elements and storage elements. Rather, it will rely on higher-level, pre-processed information that is a combination of the computing and data-handling resource information. Hence the above stated principal role of the Data Handling system which provides intelligent, high-level services to the Job Handling principal component.

Monitoring and Information Services

These services are also well covered in the Grid literature. They encompass more than e.g. MDS-based resource information service. We point out that, since we focus on Condor initially, we explicitly treat Condor's Class Ad Collector as an Information Service. We anticipate to make additional requirements to this service; for example, support for being globally distributed or dissemination of information.

The celebrated Replica Catalog concept, whether grid-standard or D0-specific, is another explicit part of this principal component.

We envision a services that processes and converts the information. For example, there may be conversion between Condor Class Ads format and Globus MDS format, or between d0-specific information services (the relation database at present) and the grid-standard building blocks (see the bricks in the picture).

The Data Handling

This architecture shows how the data handling system, SAM, should be augmented with the capabilities to use the standard Grid protocols to generate (and use) appropriate information. For example, we are likely to teach SAM to produce Condor Class Ads to be used as part of the Ads that the RB will use. As another example, we may need to have SAM stations receive updates from MDS about (abnormal) job status changes, in a standard way that will replace the existing, proprietary means of the "sam submit" mechanism.

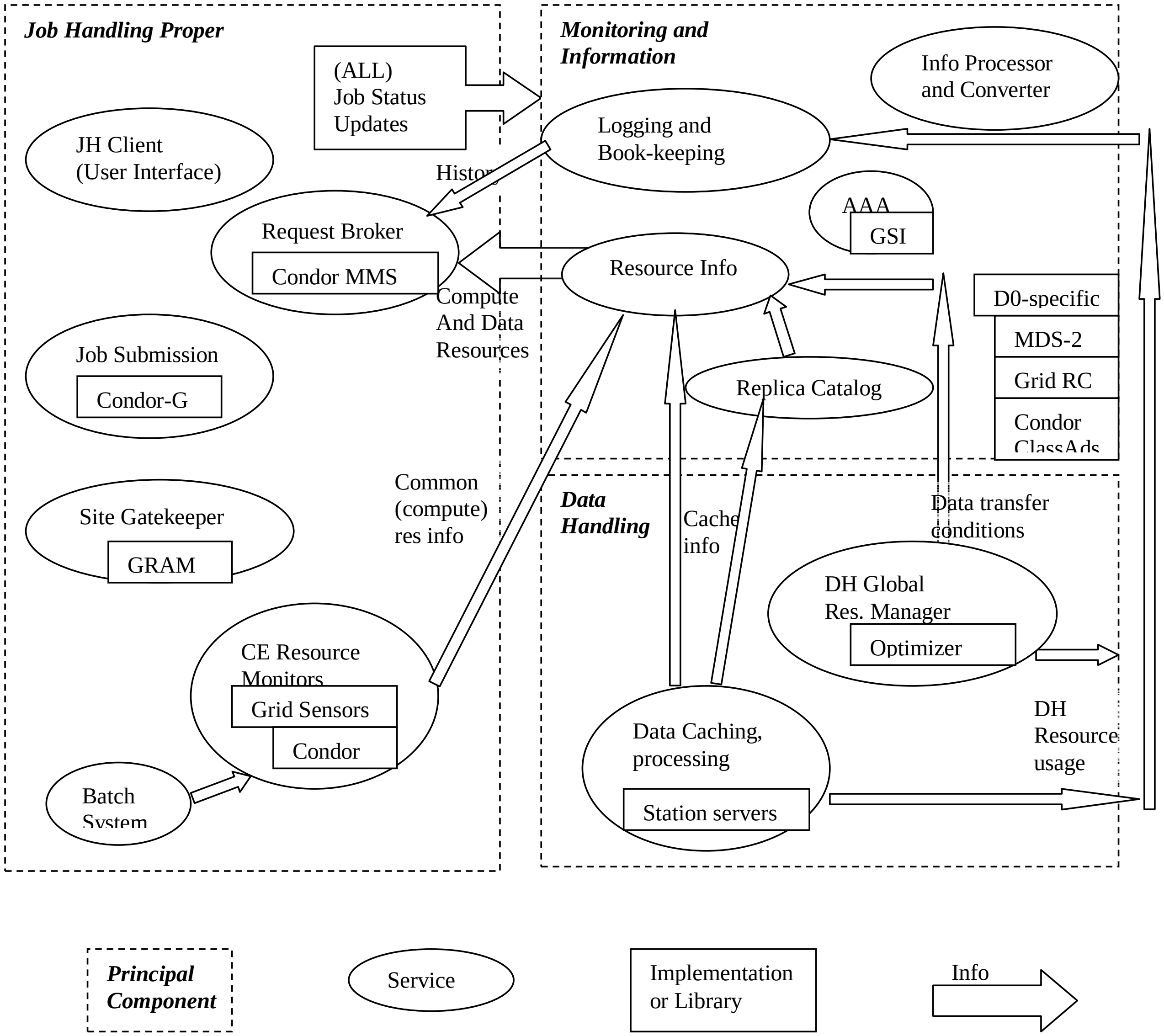
At a minimum, SAM will maintain a replica catalog (required in order to provide cached files information for analysis job placement service), as well as data transfer conditions among the stations and MSS' (an elaboration on the "Network Weather Service")

The Information Flow

In this Section we summarize what is already clear from the above descriptions. All Entities from the Job Handling and Data Handling principal components will send information to the Logging and Bookkeeping Service. Some of this information may be needed, in the form of history, for any "fair share" resource usage balancing, if we aim at such job scheduling (in the spirit of SAM's fair shares). Obviously, all this information is needed by the monitoring clients as well (not shown).

The information about the available resources will be combined somehow from that about the Computing Elements (common issue in the Grid world) with the high-level information about the data movement and availability (D0 architecture seems to be pioneering in this regard).

Authentication and Authorization information will permeate these services (hence its placement in the MIS principal component), it is therefore not explicitly connected to the other elements of the picture. The information access itself will probably (in the case of MDS – definitely) also be guarded by a security infrastructure, GSI being the obvious choice.


Job Handling Proper
(ALL)
Job Status
Updates
JH Client
(User Interface)
History
Request Broker
Condor MMS
Compute
And Data
Resources
Job Submission
Condor-G
Site Gatekeeper
GRAM
Common
(compute)
res info
CE Resource
Monitors
Grid Sensors
Condor
Batch
System
Monitoring and
Information
Info Processor
and Converter
Logging and
Book-keeping
AAA
GSI
Resource Info
D0-specific
MDS-2
Grid RC
Condor
ClassAds
Replica Catalog
Data
Handling
Cache
info
Data transfer
conditions
DH Global
Res. Manager
Optimizer
DH
Resource
usage
Data Caching,
processing
Station servers
Principal
Component
Service
Implementation
or Library
Info