ISTITUTO NAZIONALE DI FISICA NUCLEARE

CNAF–Bologna

INFN/TC-04/15
6 Settembre 2004

PILOT PRODUCTION GRID INFRASTRUCTURE FOR HIGH ENERGY PHYSICS APPLICATIONS

Sergio Andreozzi[1], Daniele Bonaccorsi[2], Vincenzo Ciaschini[1], Alessandro Cavalli[1], Flavia Donno[3], Sergio Fantinel[4], Federica Fanzago[5], Antonia Ghiselli[1], Alessandro Italiano[1], Mirco Mazzucato[5], Andrea Sciaba[3], Heinz Stockinger[5], Gennaro Tortone[6], Marco Verlato[5], Cristina Vistoli[1]

1) INFN-CNAF, Viale Berti Pichat 6/2, 40127 Bologna, Italy
2) INFN-BO, Viale Berti Pichat 6/2, 40127 Bologna, Italy
3) INFN-PI, Via F. Buonarroti 2, 56100 Pisa, Italy
4) INFN-Lab. Naz. di Legnaro, Via Romea 4, 35020 Legnaro, Italy
5) INFN-PD, Via Marzolo 8, 35131 Padova, Italy
6) INFN-NA, Via Cintia ed. G, 80126 Napoli, Italy

Abstract

Current Data Grid tools are enhancing and getting more mature. On the other hand, several application users express their needs and requirements to deploy Grid tools in their daily work. Thus, there is a strong need that Grid tools make the transition from prototype tools to production tools. We have made several improvements and additions to an early release of the European DataGrid software tools in order to meet the requirements of data intensive sciences. We show that our software release is stable and performs well in order to be used in an international production infrastructure for physics data simulation with one High Energy Physics experiment. Furthermore, we outline the experience with interoperability solutions from the Grid Laboratory Uniform Environment (GLUE) activity.

Published by ***SIS–Pubblicazioni***
Laboratori Nazionali di Frascati

1 Introduction

Several data intensive science domains like High Energy Physics (HEP) have started to rely on Grid technologies for solving parts of their computational challenges since Grid technologies have become more mature and robust. Whereas first prototype Grid solutions meet a majority of important user requirements, often interoperability or stability concerns are not addressed sufficiently due to the fact that provided solutions are still rather new. Only if these outstanding issues are dealt with in an efficient way, Grids can be used in a daily production system.

The Grid software development process usually involves several partners, starts with the requirement specification, includes prototyping, testing, improvements, and finally the deployment of the software in a production system. In our specific case, the main user community is the CMS (Compact Muon Solenoid) experiment [9] at CERN LHC (Large Hadron Collider, a new particle accelerator currently under construction). Since the computing model used by the High Energy Physics community requires the processing of Tera or even Petabytes of data, a distributed resource allocation and Grid software tools from different Grid projects are necessary and used as explained below.

The middleware developed by the European DataGrid project (EDG) [15] provided the basic software for the physics community and it is also the main system we started with. It includes workload management and data replication services built on top of Globus core services [18]. The European DataTAG project (EDT) [16] provided several additions and improvements to the EDG main baseline such as schema standardization, security enhancements and monitoring services. Moreover, DataTAG realized the integration of CMS applications with Grid services. The LHC Computing Grid project (LCG) [26], which purpose is to deploy a worldwide grid infrastructure for the LHC experiments, has adapted the above mentioned grid middleware encapsulating new features to make it flexible and extensible. LCG has also ensured support for software deployment on a CMS dedicated testbed and for application integration through the LCG/EIS (Experiment Support and Integration) group.

Within the Grid Laboratory Uniform Environment (GLUE) [2] initiative, DataTAG Work Package 4 and iVDGL [23] collaborated to the definition of a common Grid resources schema allowing a transparent access to computing and storage resources belonging to US and European Grids. During the international conference for High Performance Networking and Computing (SC2002) and the Information Society Technologies (IST2002) conference a joined EU-US intercontinental WorldGrid testbed demonstrated the validity of the GLUE activity [11]. The result of the integration of DataGrid, DataTAG middleware, and the WorldGrid experience is the LCG-0 software release [24].

In this article we report on the development and deployment activities (also referred to as ‘CMS-EIS-DataTAG initiative’) that resulted in the LCG-0 software release and the operation of a dedicated Grid infrastructure for the CMS experiment. Several new services and functionalities have been introduced at an early stage, for instance two Replica Location Service (RLS) [8] flavours, VOMS (Virtual Organization Membership Service) [4], and GridICE [3].

The main goal of the CMS-EIS-DataTAG initiative was to implement a pilot production infrastructure dedicated to a specific Virtual Organization (CMS experiment) in order to understand how well the Grid services developed so far by DataGrid and DataTAG fulfil the requirements, which improvements need to be considered and what kind of operations and support are necessary to set up a satisfactory service for users. In particular, the new enhancements made by DataTAG (VOMS and GridICE) were targeted at a production environment. Therefore, with this work a useful experience has been collected and feedback for further improvements to LCG and EDG middlewares have been provided.

The article is organised as follows: in Section 2, requirements of the HEP users community and description of the architecture is discussed; in Section 3, deployment details of our pilot production infrastructure are given; in Section 5, a brief overview on similar activities is recalled; finally, in Section 4, the results of this experience are summarized.

2 Requirements, Architecture and Implementation

2.1 Requirements

LCG is the project which will deploy a Grid software release for LHC experiment collaborations distributed around the world. In order to guarantee interoperability of the LCG facility with other Grid infrastructures, the most important requirements to satisfy are the following:

- common information modelling of Grid resources
- standard protocols to access computing and storage resources
- global security infrastructure, authentication and authorisation tools

Common information models of Grid resources are required due to the heterogeneity of involved resources and the need for interoperability among different Grid middlewares. Abstractions of different flavours of resources and services and conceptual schema of domain specific entities require a collaboration effort in order to enable a coherent information services cooperation. The LCG-0 software release has been the first experience in a production environment based on the GLUE Schema, a common information model that is the outcome of a collaboration between iVDGL, DataTAG, DataGrid, and Globus (see Figure 1 for the Computing resources model).

The GLUE Schema has been designed in respect of the following principles: (1) clear separation between system (a set of connected items or devices which operate together as a functional whole) and service (actions that form a coherent whole from the point of view of service providers and service requesters) entities; (2) generalisation, that is capturing common aspects for different entities providing the same functionality (e.g. uniform view over different batch services); (3) deal with both monitoring needs and discovery needs, in particular the former concerns those attributes that are meaningful to describe the status of resources (e.g., useful to detect fault situation), while the latter concerns those attributes that are meaningful for locate resources on the base of a set of preferences/constraints (e.g., useful during matchmaking process).

The GLUE Schema currently captures two main system categories: cluster systems providing computing services, and storage systems providing storage spaces. Within core Grid services, the computing service (called Computing Element) and the storage manager service (called Storage Element) have been defined. Each modelled service has a unique identifier, a human-readable name, a set of policies, a set of access rights, and a state. Both system and service information models have been implemented in LDAP for the Globus Monitoring and Discovery Service (MDS) [7]. Information providers collecting information from computing and storage system services have been implemented in order to provide upon request updated information to be considered for resource brokering.

As regards the need for a standard way for accessing computing and storage resources, on the computing side, this is currently achieved through the Globus GRAM [18] protocol for job submission. On the storage side protocols like GridFTP [1] and the emerging SRM [32] standard are important to facilitate data access. As for the deployment of the services, also the use of a common configuration (such as directories sharing between computing nodes) had to be adopted.

The last key requirement refers to the need for a global security infrastructure, authentication and authorisation system. As regards the authentication, this is gained via the Grid Security Infrastructure (GSI) [17]. In addition, every single user needs to be authorised and authenticated within a particular VO, which is done via the Virtual Organisation Membership Service (VOMS) [4]. VOMS ensures that the membership in a VO is treated as sensitive information.

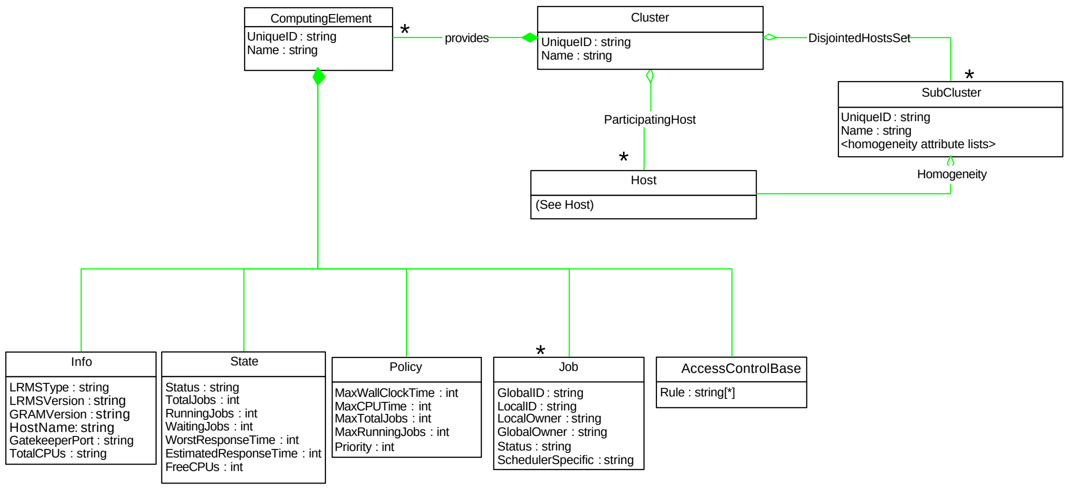

Figure 1: GLUE Schema for Computing Element in UML notation

The VOMS represents a significant improvement against the EDG LDAP-based VO server in terms of hierarchical user classification instead of the prior flat list. During this experience, we have measured that scalability and security of VOMS are superior to the previous LDAP-based VO server solution.

2.2 Architecture

The architecture of LCG-0 is based on the general Data Grid architecture of the EDG software distribution (version 1.4.3) as described in [12]. Here, we only point out the main architectural components and stress the ones that were modified or enhanced. The general layered architecture is depicted in Figure 2. The figure also shows which services are provided by the Globus, the Condor, EDG and DataTAG. In addition, services enhanced or added by DataTAG are highlighted and further discussed in the remainder of the article.

In brief, the basic services used to implement the layered architecture are given below. In addition, Figure 3 shows the service interaction from a user's point of view assuming the user submits a computing and/or data intensive task (job) or interacts with the certain services directly.

- **User Interface (UI)**: This refers to all client software components that are actually used by the end users. Typically, the software is installed on the 'User Interface machine' which then provides the main access to the Grid infrastructure. Client software comprises workload and data management as well as interfaces to the information service.

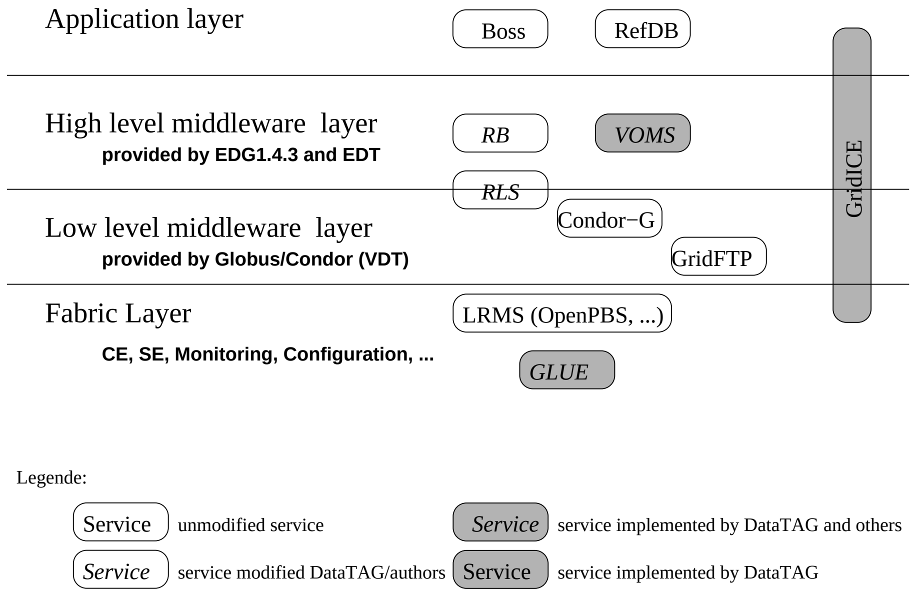

Figure 2: General layered architecture used in LCG-0 with some example services as they are provided by several partners

- **Information Service (IS)**: The Information Service provides basic information about features and status of all Grid nodes and services.
- **Replica Location Service (RLS)**: The RLS can be regarded as a special purpose information service that provides replica locations and logical identifiers of files stored in Storage Elements.
- **Virtual Organisation Membership Service**: This service provides the user with a specific proxy certificate that is then interpreted/used by Grid services for authorisation.
- **Resource Broker (RB)**: The Resource Broker is the core of the workload management system and responsible for deciding where the actual user job will be executed. All workload management services are supposed to run on a single machine.
- **Computing Element (CE)**: The Computing Element provides the gateway to several **Worker Nodes** where user jobs are executed.
- **Storage Element (SE)**: The Storage Element provides a Grid storage system.
- **GridICE**: Tool used to monitor activities on Computing Element, Worker Node, Grid services etc.

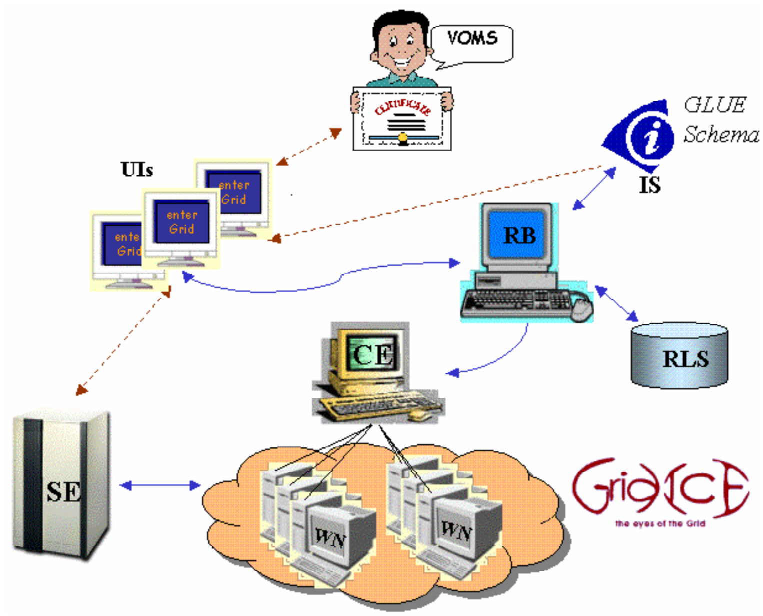

Figure 3: Main LCG-0 services and their interactions. Full lines correspond to the main workflow for a job submission task. Dashed lines represent direct user interaction with services.

2.3 Implementation

We have carried out work in several of the layers presented in Figure 2. With respect to the services described in the previous section and depicted in Figure 3, changes, enhancements and addition to the software architecture have been done in the following areas:

- Integration of the GLUE schema in the Grid Information service
- Security enhancements (VOMS)
- Modifications to the Resource Broker
- Modifications to the Replica Location Service client libraries
- Monitoring (GridICE)

Architectural details on each of these services are given below with the respective implementations.

GLUE

Integration with the GLUE schema requires modifications to the Information Service (IS) as well as all services that use the new schema and data provided by the IS. Considerable effort has been put into the integration of the GLUE schema for Computing Element (CE) and Storage Element (SE) within the Globus and EDG services. Specific information providers have been built in order to fill the parameters of the CE and SE data models. The new GLUE based Information Service (the MDS of the Globus Toolkit) has been set up and made available to the upper layer services like the Resource Broker.

VOMS

In the EDG release membership information is stored in an LDAP server that is not configured with secure access. These LDAP VO servers are now substituted with VOMS since its maturity has already been demonstrated in [16,4,30]. This allowed us to test its functionalities on a widely distributed environment, and to finally meet one of the security requirements, namely that membership in a VO should be treated as a sensitive information. On the US CMS testbed we tested its integration with local security systems and proxy certificate generation, effectively achieving a full test of its functionality.

Another important step was to allow the Resource Broker to choose where to run a job based solely on VO membership, without requiring local sites to publish the list of allowed users (another security requirement). This would have normally been done by integrating VOMS with both Resource Broker and LCAS/LCMAPS components. LCAS and LCMAPS are components that are part of the EDG software stack: Local Centre Authorisation Service (LCAS) handles authorisation requests to the local computing fabric and the Local Credential Mapping Service (LCMAPS) provides all local credentials needed for jobs allowed into the fabric. Since those components were not ready when the testbed was setup, we achieved the same effect via the Resource Broker, which allowed to specify the VO in the JDL (Job Description Language).

Resource Broker

The EDG Resource Broker and thus the workload management system in EDG version 1.4.3 had some known scalability and stability issues that have been addressed in EDG version 2.0. However, at the time of the first LCG-0 pre-release in February 2003, several of these modifications were not publicly available. We still needed these improved features for a stable environment and made several modifications and additions to the Resource Broker. In more detail, DataTAG did the following modifications:

- Stability improvements: we tried to use multiple instances of the Resource Broker and balance the load to overcome the limitations outlined in [10].
- The GLUE schema was integrated with the Resource Broker as well as the Computing and Storage Elements.
- Integration with the VOMS. Although we did not use the full set of VOMS features like extended roles, the Virtual Organisation was specified by the user in the JDL simulating the functionality of the VOMS and then used by the Resource Broker.
- Integration with the RLS system to allow for several flavours as explained in the next section. In this way the user can tell the Resource Broker which replica cataloguing system to use.

The modified Resource Broker has proven to be stable and functional.

Replica Location Service

The LDAP based Globus Replica Catalogue as provided by the EDG 1.4 middleware showed some scalability limitations for running a production service for High Energy Physics experiments. During stress tests executed by CMS on the EDG 1.4 [10] testbed it has been outlined that the implementation did not satisfy all user requirements. Therefore, the EDG developers have worked together with the Globus team to come up with a new design for a distributed Replica Location Service (RLS) [8] that allows for the localisation of physical files on the Grid.

Two different implementations of the joint design have been produced by both the Globus and EDG teams. The EDG middleware enforces a specific schema in the catalogue and uniqueness of mapping between a logical file name (*lfn*) and a physical file on the Grid. In other words, the same logical file name cannot be used to point to two different physical files with different data content. For this reason, EDG has introduced the Replica Metadata Catalogue which maps a logical identifier to a Grid Unique file ID (*GUID*). The mapping between GUID and the physical storage URL (*SURL*) of the file is stored in the RLS as designed together with Globus. Another feature introduced with the RLS is that the service is no longer centralised but distributed.

The two implementations of RLS introduced however an incompatibility between US and European Grids since they currently use different protocols. This problem is still open. In order to allow CMS to try the new features and robustness of the RLS system, the EDG 1.4 software and in particular the edg-replica-manager client [31] has been modified in LCG-0 to be interfaced to the Globus implementation of the RLS and to the EDG implementation. In more detail, the globus_rls_client library had to be modified. Via a simple configuration file that can be specified on a per job base, the user can choose to use any one of the offered systems. Three Replica Location Services were configured: the LDAP based Globus Replica Catalogue at CNAF, the Globus implementation of RLS at CERN (using a MySQL back-end) and the EDG implementation (using Oracle) operated by the IT Department at CERN. Tests were executed with the three systems, and the EDG version RLS at CERN was used to execute real production jobs.

GridICE

Beside the Grid Information Service, allowing resource discovery and resource selection through specific parameters describing characteristics and state of the resource, a monitoring system collecting resource state information and reacting to critical resource conditions is fundamental for Grid management and control. Users require to monitor their jobs, check the Grid computing resource where their jobs run and control the disk/mass storage where their data are read or written. Without this 'eye' over the Grid, not only the Grid manager cannot control the Grid behaviour, but also the user has difficulty to 'trust' the Grid. For this reason, within the DataTAG project it has been decided to develop GridICE [3], a new monitoring infrastructure that is easy to integrate in the existing middleware. It is based on the Globus MDS information service and relies on the GLUE Schema information model. This system has been used in this context to validate its functionality, flexibility and capability to answer to grid-user and VO-manager needs. In particular, the CMS-VO-manager used it to check when to schedule new job submission storms.

Discussion

The slightly modified architecture and the changes in the implementation allowed for early tests of new services like VOMS, GridICE, RLS and its integration with higher level applications. This experience was then reported to later EDG and LCG releases. In addition, we were able to deploy the software system and give access to users in a production environment as discussed in the next sections.

3 Deployment

The LCG-0 software release presented in the previous section has been deployed to provide a **Grid infrastructure**. The software has been installed and configured on DataTAG-WP4 (summarised in Table 1) and CMS resources (summarised in Table 2), which allowed us to control the environment, give the necessary support to the applications as well as make necessary interventions. The Grid collective services topology has been designed in order to optimise CMS user access to the Grid as well as CMS applications.

EDG/DataTAG Service	**Used by**	**Site**
Resource Broker	CNAF, Bari UI	CNAF
Resource Broker	Padova, Lyon UI	CERN
RLS	Jobs on WN and UI	CERN IT
BDII	CERN, CNAF RBs	CNAF
VO server	all users	CNAF
GridICE server	Users	CNAF

Table 1: Service type and location used for the evaluation. This list contains the basic EDG services on the DataTAG-WP4 testbed.

Site	**CPUs per CE**	**SE disk space**
CERN	20	1.4TB + 700 GB (+ 700GB in SRB)
Bari	18	270 GB
Bologna	22	900 GB
Legnaro	50	370 GB
Padova	50	430 GB
Ecole Polytechnique	4	220 GB

Table 2: Computing and storage resources used during the CMS Pre-Challenge Production. These resources were only dedicated to the CMS experiment (i.e. a single VO) to increase the computing and storage power.

3.1 Deployment Details

The Resource Broker used (called ‘DataTAG/LCG-0’ broker) was based on the EDG broker version 1.2.21.2 from EDG 1.4.3, modified for a VO-based match-making. In addition, the support of the interaction with the RLS was achieved, allowing it to understand values of the JDL ReplicaCatalog attribute of the form rls://hostname:port and edgrls://hostname:port. The Replica Manager [31] of EDG 1.4 only supported the Globus version of RLS and had to be modified, together with some Globus libraries, to interface to both RLS versions.

A dedicated machine running the GridICE monitoring server was installed at CNAF (Bologna, Italy). The DataTAG cluster monitoring and information providers system was deployed on the testbed. A GLUE extended schema was installed on each resource. The installation of EDG information providers for the CE on top of the EDG software was done manually at each site according to instructions provided in collaboration with GLUE experts.

A Globus MDS-based Information Service was deployed. The central Index Services have been configured at CNAF. The Information Service was structured according to the EDG 1.4 style, with a MDS GRIS/GIIS hierarchy, a TOP GIIS and a BDII (Information Index), whose roles are outlined below. A GRIS resides on each CE and SE and produces information about the status of the resource and stores it in LDAP (ldif) format. The GRIS registers to a local GIIS. There is one GIIS per site, on the CE, to collect and cache information of the site nodes. The local GIIS registers itself on a TOP GIIS, which is configured on a CNAF CE, nevertheless served the entire testbed by collecting and caching information from all involved sites. The Berkeley DB Information Index (BDII), installed at CNAF, was used to provide the RB with a fast and stable caching of information about the Grid resources, helping to avoid the known instability of the Globus GIIS. It is a standard OpenLDAP server, which periodically (every 10 minutes) feeds itself with the content of the TOP GIIS.

3.2 Validation Phase

Before starting production, a long phase of site validation took place. After following precise installation and testing instructions published by the EIS group, a site was targeted with a specific set of CMS jobs to see if the results produced were as expected. Only then the site was admitted into the CMS/LCG-0 testbed and regularly registered in the information service. A CMS specific GlueHostApplicationSoftwareRunTimeEnvironment tag was published by the site, which allowed users to include the site among those able to process CMS requests.

3.3 CMS applications

The highest layer in the architecture of Figure 2 consists of the physics applications of the CMS experiment. The CMS applications executed on this Grid are part of the so-called Pre-Challenge Production (PCP) of the CMS Computing Data Challenge (DC04) foreseen to start in February 2004. It has to cope with the simulation of about 50 Million Monte Carlo (MC) events and is a world-wide distributed task.

Each step of the production chain is managed in the CMS production environment

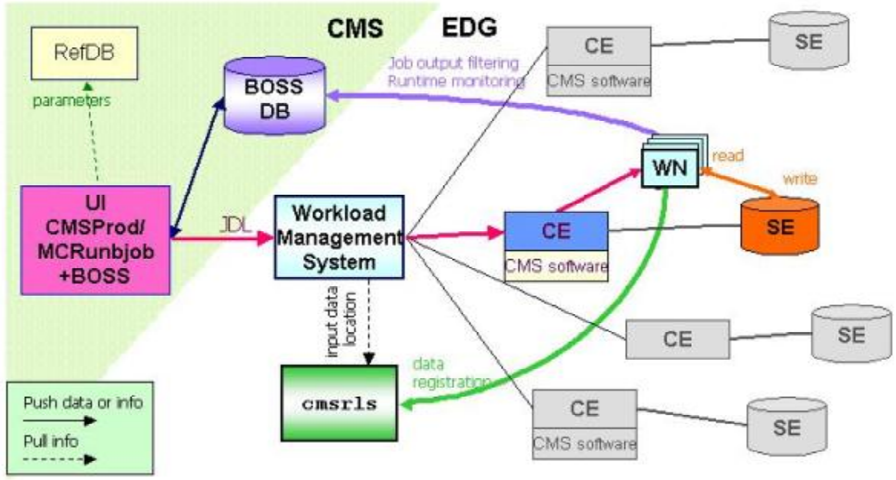

Figure 4: Integration of CMS production tools with LCG-0 middleware

by the following tools (see also Figure 4), developed in provide an automated subsystem taking care of input parameter management, robust and distributed request and production accounting, job preparation and submission on the batch system, job tracking and bookkeeping, data and replica management:

- **RefDB** [25] is a central SQL database, which allows for gathering of all relevant information to be given as input to the production component.
- **McRunJob** [19] (Monte Carlo Run Job)is a tool for job preparation that provides a metadata based approach for specifying work-flow patterns. It has plug-ins for submitting jobs to several resource management systems.
- **BOSS** [20] provides real-time monitoring and bookkeeping of jobs submitted to a computing resource, storing the information persistently in a relational database for further processing. Boss was installed on the UI.
- **CMSprod** is a program that allows for the creation and submission of production jobs into different schedulers including LCG-0 ones. It is a layer between middleware (LCG-0) and experiment specific software. The program is able to generate scripts for different job types, simulation jobs, analysis jobs, reconstruction jobs and submit them to Grid scheduler.

McRunJob or CMSProd was installed on the UI, where the job preparation and job submission took place. Both McRunJob and CMSProd were implemented to run produc-

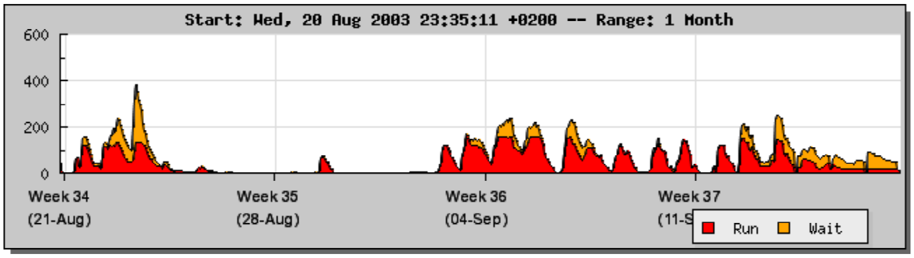

Figure 5: Number of jobs, both running and waiting, as a function of time, as shown from GridICE monthly graph

tion jobs in the specific LCG-0 Grid environment. The set-up of the CMS environment appropriate to run the job relied on the pre-installed CMS software in the Computing Element, as advertised in the Information service. The EDG data management tools were used to discover the physical location of an input file and to transfer the produced output data from the worker node to the Storage Element and its registration in the RLS. Along with the job scripts, JDL files were created with all the scripts and parameters needed.

4 Experience and Results

In this section we report on our experience on how application scientists have used the Grid infrastructure described in the article.

About 2 Million simulated events of the CMS production were produced, as reported in Table 3, corresponding to about 1.5 TB of data. The total number of successful 'long' (CPU intensive) jobs of 250 events each was about 8,000. According to GridICE monitoring, the number of jobs in the system over a period of one month is shown in Figure 5.

	Total no. of events	**data size**
CMKIN	500,000	20 GB
CMSIM	1,495,000	1,500 GB

Table 3: Total number of successfully produced CMKIN and CMSIM events and the corresponding size of stored data. CMKIN and CMSIM represent different steps in the physics simulation process.

The general results are very encouraging since the CMS experiment could efficiently use the infrastructure. The overall failure rate varied from 5-10% to 25-30%

depending on the incidence of some problems, in particular those related to the RLS unavailability, mainly for maintenance reasons. The main problem sources were identified as:

- The RLS unavailability, due both to planned interventions and to unexpected problems (e.g., disk failure, bug of the Oracle application server).
- Some instability due to a variety of reasons at various sites (e.g., wrong configuration, network problems, and hardware failures). These problems gave an overall inefficiency of about 5-10%.

Greater improvements in terms of stability and efficiency for 'long' jobs were found with respect to the Stress Test production carried out documented in [10].

The first usage of VOMS on a real distributed system allowed us to obtain various results. First of all, its use and its ability to function as a scalable VO substitute for generation of grid-mapfiles. This is particularly evident if coupled with the results of a stress test done by US CMS community (e.g., generation of proxies, where roughly 95,000 successful requests were made to the server within one and a half days). Furthermore, this setup allowed us to test a mechanism for generation of grid-mapfiles as a compromise between LDAP based VO servers (completely open to the public) and the need for a prior registration of every grid service (e.g., CE, SE, and RB). This system configuration did not require such a pre-registration. In fact, the only way to access the list of users would have been to possess a valid certificate, verified by the VOMS server machine, though such access would be logged. Finally, its installation on a system like LCG-0, which was based on several different software versions, allowed us to discover and solve many portability problems, thus making VOMS consistently usable on many different kinds of Grid middleware.

Concerning the GridICE monitoring service, the CMS experience with the LCG-0 Grid middleware has been the first important opportunity for a wide deploy and test in a production environment. The integration with the middleware has been smooth as expected, and this proves the correctness of the design choices. The MDS Information Service has been loaded with extra data, but no scalability issues have been noticed. However, a recent version of GridICE avoids that the extra data needed for monitoring and analysis and not for resource brokering are injected in the MDS hierarchy. The GridICE web presentation of the Grid status showed to provide useful information to the production manager in order to plan the submission of jobs. The historical graph of waiting/running jobs aggregated for the whole Grid and the info related to the site view are concise representations of when new bulks of jobs can be submitted.

5 Related Work

Several Grid projects have the goal of design, deploy, and run Grid production infrastructure. As regards the physics community, the meaningful experiences we recall are related with the NorduGrid project [13], the Grid3 project [21], and the GridPP project [28].

The NorduGrid project set up a production Grid in the nordic countries of Europe. As regards information modeling, they use a proper solution, thus different than original Globus MDS schema and GLUE Schema. As regards standard protocols, they do not rely on the Globus GRAM, but they have developed a replacement called 'Grid Manager' based on the Globus Toolkit 2 libraries. As regards the security, they rely on GSI for the authentication, and they have included the VOMS system developed in DataTAG for the authorization.

The goal of the Grid3 project was to provide the infrastructure and services needed to demonstrate LHC production and analysis applications running at scale in a common grid environment [22]. As regard information modeling, they adopted the GLUE Schema and defined extensions, in particular for VO accounting info and site configuration [27]. As regards the resource access protocol, they rely on Globus GRAM. Concerning the Grid monitoring, they rely on the MonaLISA [29]. Finally, concerning the security, they adopt the GSI for the authentication and VOMS for the authorization.

The GridPP [28] project aims at creating and managing a production Grid infrastructure for the United Kingdom in order to support the LHC experiments. They rely on the EDG/LCG middleware and extend it with a monitoring service called R-GMA [6].

6 Conclusion

The enhancements and modifications that we did for the LCG-0 release provided the possibility to deploy an international production infrastructure for the CMS experiment in order to run successful physics simulations. This showed that current Grid technology is maturing as a computational and data distributed infrastructure in production environments.

As a second major contribution of our work we successfully tested and deployed new services and provided several useful feedbacks and software components that have been later included in subsequent software releases of EDG and LCG middlewares.

This experience will be used in current as well as future Grid projects in Europe. In particular, the EGEE project (Enabling Grids for E-science in Europe) [14] which main purpose is the integration of current national, regional and thematic Grid efforts to create a seamless European Grid infrastructure for the support of the European Research Area, will consolidate the results achieved and presented in this article as regards the production Grid infrastructure.

Acknowledgements

We are grateful to the DataTAG coordinators for the CMS experiment, Claudio Grandi and Paolo Capiluppi, for giving us the opportunity to test the most important DataTAG/WP4 solutions in a real application environment. We thank also Elisabetta Ronchieri, Salvatore Monforte and Massimo Sgaravatto from DataGrid/WP1 for their collaboration in the integration of the Resource Broker with the GLUE Schema and RLS. We thank David Smith (LCG) for his hints on the modification of the Globus libraries for RLS.

References

[1] W. Allcock. GridFTP: Protocol Extensions to FTP for the Grid. GGF Recommendation. GFD.20

[2] S. Andreozzi, M. Sgaravatto, C. Vistoli. Sharing a Conceptual Model of Grid Resources and Services. Computing in High Energy and Nuclear Physics (CHEP 2003), La Jolla, California, March 24 - 28, 2003.

[3] S. Andreozzi, N. De Bortoli, S. Fantinel, A. Ghiselli, G. Tortone, C. Vistoli. GridICE: a monitoring service for the Grid. In Proceedings of the 3rd Cracow Grid Workshop, Oct 27-29, 2003.

[4] R. Alfieri, R. Cecchini, V. Ciaschini, L. dell'Agnello, A. Frohner, A. Gianoli, K. Lőrentey F. Spataro. An Authorization System for Virtual Organizations. 1st Across-Grids Conference, Santiago de Compostela, Spain, February 2003.

[5] C. Baru, R. Moore, A. Rajasekar, M. Wan. The SDSC Storage Resource Broker, CASCON'98, Toronto, Canada, November 30 - Decemeber 3, 1998.

[6] A. Cooke, A. Gray, L. Ma, W. Nutt, J. Magowan, P. Taylor, R. Byrom, L. Field, S. Hicks, J. Leake, *et Al.* R-GMA: An Information Integration System for Grid Monitoring. In Proceedings of the Eleventh International Conference on Cooperative Information Systems (2003).

[7] K. Czajkowski, S. Fitzgerald, I. Foster, C. Kesselman. Grid Information Services for Distributed Resource Sharing. Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.

[8] A. Chervenak, E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunszt, M. Ripeanu, B. Schwartzkopf, H. Stockinger, K. Stockinger, B. Tierney. Giggle: A Framework for Constructing Scalable Replica Location Services. International IEEE Supercomputing Conference (SC 2002), Baltimore, USA, November 2002.

[9] The CMS Collaboration, Technical proposal, CERN/LHCC 94-38, LHCC/P1 (1994)

[10] CMS/EDG Stress Test Task Force. CMS Test of the European DataGrid Testbed. CMS Note 2003-14, May 1, 2003.
http://cmsdoc.cern.ch/documents/03/note03_014.pdf

[11] F. Donno, V. Ciaschini, D. Rebatto, L. Vaccarossa, M. Verlato. The WorldGRID transatlantic testbed: A Successful Example of Grid Interoperability Across EU and US Domains. Computing in High Energy Physics (CHEP03) La Jolla, California, March 24-28, 2003.

[12] EDG Architecture Task Force. The EDG Architecture. DataGrid-12-D12.4-333671-3-0, February 11, 2002.
https://edms.cern.ch/file/333671/3.0/DataGrid-12-D12.4-333671-3-0.pdf

[13] P. Eerola et al. The NorduGrid Production Grid Infrastructure, Status and Plans, 4th International Workshop on Grid Computing (Grid2003), Phoenix, Arizona, November 17, 2003.

[14] EGEE: Enabling Grids for E-science in Europe Project (EGEE). http://www.eu-egee.org

[15] European DataGrid Project (EDG). http://www.eu-datagrid.org

[16] European DataTAG Project (EDT). http://www.datatag.org

[17] I. Foster, C. Kesselman, G. Tsudik, S. Tuecke. A Security Architecture for Computational Grids. 5th ACM Conference on Computer and Communications Security Conference, San Francisco, California, November 2-5, 1998.

[18] Globus Alliance. The Globus Toolkit, http://www.globus.org

[19] G. Graham. Mc_Runjob - A High Energy Physics Workflow Planner for Grid Production Processing, Computing in High Energy Physics (CHEP03) La Jolla, California, March 24-28, 2003.

[20] Claudio Grandi. BOSS: A tool for Batch Job Monitoring and Book-keeping. Computing in High Energy Physics (CHEP03) La Jolla, California, March 24-28, 2003.

[21] Grid 3: http://www.ivdgl.org/grid2003

[22] The Grid2003 collaboration. Grid2003 Project Lessons. Technical Report.

[23] International Virtual Data Grid Laboratory (iVDGL). http://www.ivdgl.org/

[24] LCG-0 software distribution. http://cmsdoc.cern.ch/cms/LCG/LCG-0

[25] V. Lefebure, J. Andreeva. RefDB: The Reference Database for CMS Monte Carlo Production, Computing in High Energy Physics (CHEP03) La Jolla, California, March 24-28, 2003.

[26] LHC Computing Grid Project (LCG). http://cern.ch/lcg

[27] M. Mambelli on behalf of the Grid2003 Project. Grid3 monitoring metrics and architecture. 30 Dec 2003.

[28] A. McNab. A Technical Overview of the GridPP Testbed. UK e-Science All Hands Conference, Nottingham, September 2003

[29] H.B. Newman, I.C. Legrand, P. Galvez, R. Voicu, C. Cirstoiu. MonALISA: A Distributed Monitoring Service Architecture Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California

[30] Security Coordination Group - EU DataGrid Project, EU DataGrid Security Report, DataGrid-07-D7.7-0207-Security-2.2, January 2004. https://edms.cern.ch/file/414762/2.2/DataGrid-07-D7.7-0207-Security-2.2.pdf

[31] H. Stockinger, F. Donno, E. Laure, S. Muzaffar, G. Andronico, P. Kunszt, P. Millar. Grid Data Management in Action: Experience in Running and Supporting Data Management Services in the EU DataGrid Project. Computing in High Energy Physics (CHEP 2003), La Jolla, California, March 24 - 28, 2003.

[32] Storage Resource Management (SRM) working group. http://sdm.lbl.gov/srm-wg/

[33] Virtual Data Toolkit (VDT) http://www.lsc-group.phys.uwm.edu/vdt/