On the employment of LCG GRID middleware

Luben Boyanov, Plamena Nenkova

Abstract: *This paper describes the functionalities and operation of the LCG GRID middleware. An overview of the development of GRID technologies and important GRID projects is made. Special attention is paid to the architecture and interaction of the components of LCG. The paper describes the experience and practice during the use of LCG for GRID tasks.*

Key words: *GRID technologies, LCG middleware.*

1. INTRODUCTION

GRID technology has become a popular and wide spread term during the last decade. The term is used for a unified computing resource, able to perform huge computing tasks. Its attractiveness is based on the popularity and wide-spread of Internet, high-speed network technologies, and low-cost hardware. Our paper reviews GRID architecture and examines in more details the functioning and abilities of an important middleware software component – the LCG.

2. DEVELOPMENT OF GRID TECHNOLOGY

Interconnecting centers with supercomputers for big problem solving is not a novel idea but in the mid 90-ies two such projects - FAFNER [1] and I-WAY [2] laid the foundations of the technology, known nowadays as GRID. FAFNER (Factoring via Network-Enabled Recursion) developed techniques for splitting up and distributing a big computational problem nowadays on the basis of seti@home (a scientific experiment that uses Internet-connected computers in the Search for ExtraTerrestrial Intelligence – [3]). I-WAY (Information Wide Area Year) influenced the development of the Globus Project, which is at the core of most GRID activities. During a workshop at Argonne National Laboratory, the term "computational GRID" emerged [4]. Soon followed the book regarded nowadays as the "GRID bible"- "The GRID: Blueprint for a New Computing Infrastructure" [5]. During and after the late 90-ies, both in the USA and Europe, many scientific teams started work on GRID projects. GRID is not only one of the very interesting and modern trends in contemporary computer science. The expectations of its development and importance are demonstrated in the fact that GRID is amongst the four technological priorities of the Sixth Framework Programme (FP6) of the Information Society Technology Programme of the European Union.

3. MAIN COMPONENTS OF GRID TECHNOLOGY

A GRID system coordinates resources that are not subject to centralized control, uses standard, open, general-purpose protocols and interfaces, and delivers nontrivial qualities of service [6].

As pointed in [7], the GRID components can be viewed as follows:

• *GRID fabric* - all the globally distributed resources that are accessible from anywhere on the Internet. Such can be computers running various operating systems, storage devices, databases, etc.

• *Core GRID middleware* – a layer with services such as remote process management, of resource reallocation, storage access, information registration and discovery, security, resource reservation and trading.

• *User-level GRID middleware* - application development environments, programming tools, and resource brokers for managing resources and scheduling application tasks for execution on global resources.

• *GRID applications and portals* - GRID applications are typically developed using GRID-enabled languages and utilities for parallel programming such as HPC++ (High-Performance C++) or MPI (Message Passing Interface). GRID portals offer Web-enabled application services, where users can submit and collect results for their jobs on remote resources through the Web.

Figure 1 presents the architecture of GRID in regard to its components.

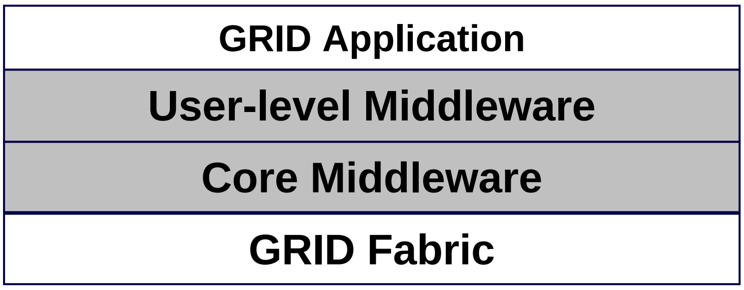

Figure 1. GRID components.

The intensive work on GRID is due to the fact that many applications can benefit from such distributed supercomputing paradigm – for example collaborative engineering, data exploration, high-throughput computing, etc. So far, one of the shortcomings is the lack of common standards due to the numerous requirements towards GRID services and applications. The development of Grid technology is closely linked with the results of a several successful GRID projects.

- ***Globus***

Many GRID projects are being built on protocols and services provided by the Globus Toolkit [8] - an open source reference implementation of key protocols for GRID computing. The Globus Toolkit uses an object oriented approach and provides a uniform programming model for GRID applications. It offers the tools to employ services, needed to implement security, resource location, resource management, and communications.

- ***European DataGRID (EDG)***

DataGRID was a European Union funded project [9]. It was a successful project, building a test computing infrastructure which provides shared data and computing resources (EDG software). The project used the Globus toolkit and created a functioning GRID testbed. After its end in the spring of 2004, a followup of the project - EGEE started.

- ***EGEE (Enabling GRIDs for E-science in Europe)***

EGEE project [10] will concentrate on building a consistent, robust and secure GRID network and to improve and maintain the existing GRID middleware. EGEE aims to build and operate GRID infrastructure, providing GRID service to a variety of scientific applications. It will attract additional computing resources and deliver reliable service to users.

One should also mention the **Condor** [11] software – a specialized workload management system for compute-intensive jobs and the **Virtual Data Toolkit** (VDT) [12] - an ensemble of grid middleware, which eases the deployment, maintainance and use of GRID middleware.

- ***LCG (Large Hadron Collider Computing GRID)*** :

The European Laboratory for Particle Physics - CERN, one of the world's most prestigious centers for fundamental research has started an ambitious GRID project - the LCG [13] in 2003. It aims to solve huge data storage and analysis challenges for tasks, necessary for scientists, who will use the Large Hadron Collider -the world's largest and most powerful particle accelerator intended be in 2007. LCG will provide common infrastructure of libraries, tools and frameworks required to support the physics application software [14]. It is a collection of geographically distributed resources and services allowing people working in LHC experiments, organized in Virtual Organizations (VO) to run their programs for simulation, processing and data analysis. In a functional way it can be considered as the middleware GRID component. LCG has deployed Globus, EDG, VDT and Condor software. One should also mention that LCG is used as the production platform for EGEE. In March 2005 the LCG project announced that the computing GRID it is operating includes more than 100 sites in 31 countries [15].

4. LCG ARCHITECTURE

The LCG software incorporates services at the User-level- and Core Middleware architectural level. These services perform resource management (Workload Management System - WMS), data management (Data Management System - DMS), information management (Information System - IS) and security (the Authorization and Authentication Systems). Other services - for accounting, monitoring, installation and configuration are also implemented.

The current version of the LCG-2 middleware has several major components:

UI - User Interface, RB – Resource Broker, RLS – Replica Location Service, BDII – Berkeley Database Information Index, CE – Computing Element, gateway to computing resource, WN – Worker Node provides the computing cycle for the CE, SE – Storage Element.

The software packets for the above mentioned services can be grouped and installed on one or more computers. Each organization running LCG-2, plans its GRID site with nodes depending on available resources and application requirements. The interaction of different components on the submission of a job by the user is described bellow and illustrated in Figure 2.

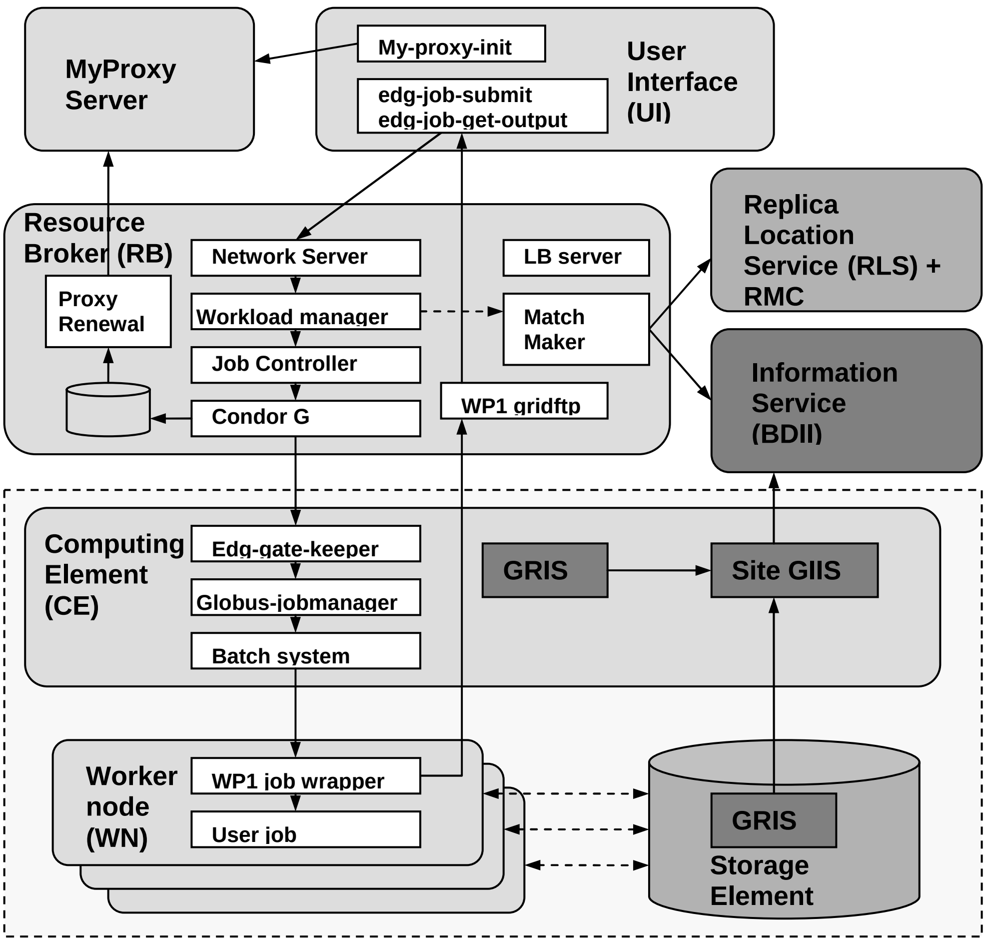

Figure 2. LCG components and the interaction amongst them.

Workload Management System (WMS)

The WMS manages the jobs submitted by users. It matches the job requirements to the available resources and schedules it for execution on an appropriate computing cluster. It then tracks the job status and allows the user to retrieve the job output when ready.

Users authenticate themselves using temporary credentials called proxy certificates. User logins to UI and submits job to RB.

RB runs several services. The Network Server (NS) authenticates the user, accepts the job, and supports job control functionality. When a job is submitted, the Workload Manager (WM) calls the Matchmaker to find the resource which best matches the job requirements, interacting with the Information System and the Replica Location Service. The Job Controller (JC) is then called to submit the job to the Condor-G component for control remote job submission. Afterwards the job is sent to the CE. The Logging and Bookkeeping service (LB server) logs all job management GRID events.

The CE is the front-end GRID interface to a computing node. The Gatekeeper accepts job requests and creates a job manager to submit or to cancel jobs. CE manages a computing cluster of WN, which executes jobs. The Local Resource Management System (LRMS) dispatches jobs to WN.

Data Management System (DMS)

Data Management System provides access to data in the GRID and allows for replication between different sites. The DMS relies on two kinds of services: the Replica Location Service and the Storage Element. Data is transferred via a number of protocols (GridFTP is the most commonly used). Two catalogs within the RLS provide the mapping between the unique file identifiers and physical file names. Physically the files are stored on a SE. The SE can range from simple single disk server to complex Mass Storage Systems.

Information System (IS)

The IS provides information about the LCG-2 resources and their statuses. The Grid Resource Information Service (GRIS) provides both static and dynamic information of a particular resource. The information is stored into an OpenLDAP (Lightweight Directory Access Protocol) database. Each GRID site also provides a Grid Information Index Service (GIIS) that aggregates all registered local GRISes. Finally, at the top level is the Berkeley Database Information Index. It regularly queries a list of site GIISes and/or GRISes to provide a global view of the information system for a Virtual Organization.

Authentication and Authorization System

The Authentication and Authorisation system contains the list of all the people authorised to use LCG-2, according to a VO. This list is normally downloaded to all machines running GRID services in order to map the LCG-2 users to the local users of the machine. Centralized LDAP servers provide user certificate subjects for each VO. The VO server is queried by the RB, CE, and SE to build a list of authorized users.

Monitoring Systems

LCG-2 has monitoring services. One of them - GridICE is a complete GRID monitoring service to collect, store and visualise the status of the LCG-2 resources (the number of jobs running, the storage space available). Another one is R-GMA (Relational Grid Monitoring Architecture). This monitoring and information management service uses SQL as a query language and gives an RDBMS-like (Relational Database Management System) view of information for distributed resources.

5. OPERATING LCG

The Institute for Parallel Processing at the Bulgarian Academy of Sciences (IPP-BAS) participates in a number of GRID related projects (EGEE, SEEGRID, CoreGRID). Through them IPP-BAS actively participates in the development of a pan-European GRID infrastructure, integrating networking and middleware environment. The projects are also very important in terms of the creation of regional GRID infrastructure.

As a first step, IPP-BAS introduces the pilot LCG-2 GRID middleware thus forming a medium size GRID site (an organization providing all GRID Services is called "GRID site"). The GRID Resources at a site are the hardware, software, data and supporting infrastructure required to provide the GRID Services. BG04-ACAD is one of the Testbed sites at IPP-BAS.

5.1. Network topology, services and resources

Eight computers at the BG04-ACAD site (called nodes) are inter connected via Gigabit Ethernet LAN as well as to GEANT, which provides capacities for GRID experiments. All nodes are protected by firewall.

The LCG-2 middleware has modest requirements on the hardware. The minimal configuration for all nodes includes a Pentium II processor with clock rate of 500MHz and 256MB memory. The middleware uses about 1GB of disc space in addition to at least 5GB for each job run at the same time. The total storage for the SE depends on the particular GRID application. A small shared file system is currently required for the storage of the experiments software. The requirements of the High Energy Physics (HEP) experiments for their productions are much more demanding.

The BG04-ACAD site (Figure 3) provides computing resources (for the WNs) based on a PC cluster with 6 processors Pentium IV, 2,8 GHz. The capacity of the storage resources (SE) is 80 GB. The CE of BG04-ACAD, which is visible for external sites uses Dual Intel IV processor running at 2,8 GHz with 512 MB memory.

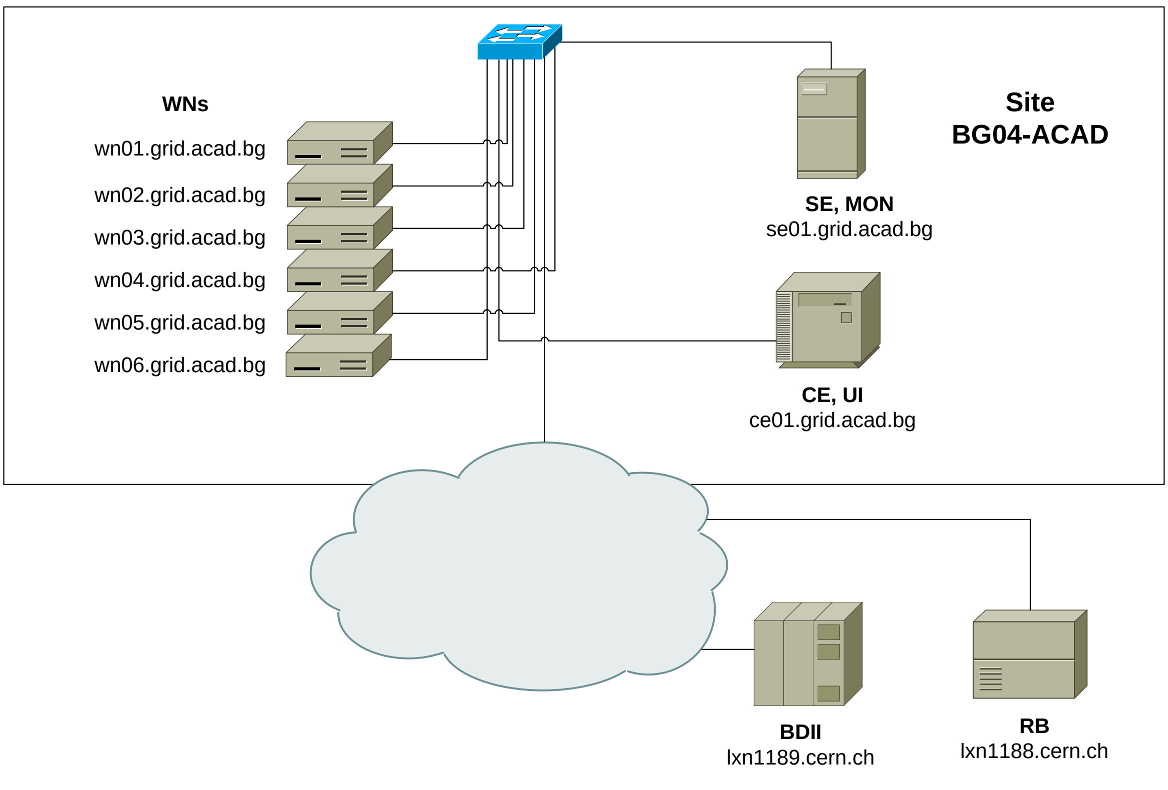

Figure 3. The BG04-ACAD GRID Testbed Site.

5.2. Making use of LCG-2

GRID middleware installation and configuration

We have installed and configured LCG-2 on top of Scientific Linux 3.0. This is the recommended OS but LCG can be also installed on the Linux Red Hat 7.3. The installation and configuration method (after LCG-2_3_0) is based on the Debian *apt-get* installation tool and on a set of shell scripts for the various LCG node types (WN, UI, CE, SE, etc) built within the *yaim* (Yet Another Installation Method) framework [16]. Yaim uses some example configuration files which we build. They contain specific data for the local site - domain name, the name of local service nodes (CE, SE, WN), parameters of the machines like memory and speed, name of external services (RB, BDII, Proxy Server), and VO names. The additional tuning of the middleware (in respect to the various services) and local customization has been made. Scripts have been run for the configurations, but due to errors, some were tuned manually. At present all nodes have been upgraded with the last available version - LCG-2_4_0 (giving fewer errors during installation and configuration).

Site components

Many of the emerging sites are with limited resources, but aiming get involved and gain experience with GRID deployment and usage. For a minimal site it is typical to have a user interface node which allows jobs to be submitted to the GRID. This node uses the information system and resource broker from other sites. A site providing resources adds a computing element (a gateway to the computing resources) and/or a storage element (a gateway to the local storage). Worker nodes can be added to provide computing power. Smaller sites usually add RGMA monitoring (MON) node functionality to their SE, while medium to large sites add a separate node as a MON node.

During the creation of our site, we dedicated nodes for CE, SE, and WNs. To avoid the necessity to add another node, MON service was installed on the SE node and UI - on the CE node. If the computing resources and number of job submisson grows, we intend to install a local RB and a BDII node.

Site registering and testing

The organizations that set up a GRID site using LCG have to register their services in the General Operation Center (GOC) at CERN. GOC is responsible for the monitoring of the deployed GRID services.

The support for sites is organized in a hierarchical way. Each site has a Resource Administrator of a GRID Service and Site manager responsible for the operation of a site. The ROC (Regional Operation Center) manager assists sites through the deployment of middleware releases and operational procedures. LCG Service-related software is distributed from CERN under the control of the GRID Deployment Team.

GOC Database (GOC-DB) requires contact information for registering the site resources. The BG04-ACAD site was registered in the GOC-DB. The CE and SE services have been published in the testZone BDII (lxn1189.cern.ch) which is used by RB (lxn1188.cern.ch) at CERN. Access to the database is done through secure session via a web browser and the use of an X.509 certificate issued by a trusted LCG Certification Authority. All sites offering a CE Service provide accounting information to the Grid Operations Service to enable LCG-wide accounting statistics to be maintained.

Data and status about LCG sites can be retrieved from the GOC URL http://goc.grid-support.ac.uk/. Daily tested sites status can be extracted from the sites functional tests webpages: http://lcg-testzone-reports.web.cern.ch/lcg-testzone-reports/cgi-bin/lastreport.cgi .

The current workload of BG04-ACAD can be monitored via GridICE monitoring pages [17]. Our site is being tested by jobs from different VO, mostly from the Biomed test application (Biomed VO) [18].

6. CONCLUSSIONS

The LCG team follows for LCG software updates and new packages through the LCG_Rollout mailing list. A new LCG/EGEE GRID Middleware (gLite 1.0) has been released in April 2005. In addition to the other LCG functionalities, it also supports Web services. gLite GRID services follow a Service Oriented Architecture and allow compliance with upcoming standards, such as OGSA.

BG04-ACAD at IPP-BAS supports and runs HEP (ALICE, ATLAS, CMS and LHCb Virtual Organization) and Biomed (Biomed VO) test application. We look ahead to run regional GRID applications in the near future.

REFERENCES

[1] FAFNER. http://www.lehigh.edu/~bad0/fafner.html.
[2] I-WAY. http://www.sdsc.edu/SDSCwire/v1.15/7037.iway.html.
[3] seti@home. http://setiathome.ssl.berkeley.edu/.
[4] Argonne Workshop. http://www.crpc.rice.edu/newsletters/fal97/news_grid.html .
[5] Foster, I., C.Kesselman. The GRID: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, 1999.
[6] Foster, I. What is the Grid? A Three Point Checklist. http://www-fp.mcs.anl.gov/~foster/Articles/WhatIsTheGRID.pdf, 2002.
[7] Baker, M., R. Buyya, D. Laforenza. GRIDs and GRID technologies for wide-area distributed computing. http://www.gridbus.org/~raj/papers/gridtech.pdf, 2002.
[8] Foster, I., C.Kesselman. Globus: A Toolkit-Based GRID Architecture. *The GRID: Blueprint for a New Computing Infrastructure.* Morgan Kaufmann, 1999, 259-278.
[9] European DataGrid Project. http://eu-datagrid.web.cern.ch/eu-datagrid/.
[10] Enabling Grids for E-science in Europe. http://eu-egee.org/.
[11] Condor. http://www.cs.wisc.edu/condor/.
[12] VDT. http://www.cs.wisc.edu/vdt//index.html.
[13] LHC Computing Grid Project. http://lcg.web.cern.ch/LCG/.
[14] LCG-2-UserGuide. https://edms.cern.ch/file/454439//LCG-2-UserGuide.html
[15] LCG Press. http://info.web.cern.ch/Press/PressReleases/Releases2005/PR03.05ELCG100site.html.
[16] LCG-2 Manual Install. http://grid-deployment.web.cern.ch/grid-deployment/documentation/LCG2- Manual-Install/.
[17] GridICE. http://gridice2.cnaf.infn.it:50080/gridice/site/site.php.
[18] Biomed. http://egee-na4.ct.infn.it/biomed/.

ABOUT THE AUTHORS

Luben Boyanov, Research Assistant, PhD, Institute for Parallel Processing, Bulgarian Academy of Sciences, Phone: +359 2 9796616, Email: lb@acad.bg

Plamena Nenkova, PhD student, Institute for Parallel Processing, Bulgarian Academy of Sciences, Phone: +359 2 9796608, Email: plamena@acad.bg