HIGH-PERFORMANCE COMPUTING SYSTEM FOR HIGH ENERGY PHYSICS

M. ORLANDEA, C. COCA, L. DUMITRU, E. TEODORESCU

Horia Hulubei National Institute of Physics and Nuclear Engineering, P.O.Box MG-6, RO-077125
Bucharest-Magurele, Romania, E-mails: orlandea@ifin.nipne.ro; cornelia@lhcbcc.nipne.ro;
dlaur@ifin.nipne.ro; eli_teo@ifin.nipne.ro

Received September 4, 2009

In this work we present technical details and recent developments for a computing cluster working in a GRID environment, configured for high energy physics experiments at the National Institute of Physics and Nuclear Engineering. Main ideas and concepts behind the GRID technology are described. Two *Virtual Organizations* (VO) LHCb and ILC using GRID resources for Monte Carlo production, data analysis and data storage are presented together with the recently initiated development of their specific tools.

Key words: GRID computing, high energy physics.

1. INTRODUCTION

Nowadays, research without high-tech computers is unconceivable. Their cost and the maintenance cost is most often a real challenge in research. Therefore, in the last years great efforts have been made in order to find alternative and cost efficient solutions in computing development. Usage of low cost personal computer clusters and the internet to connect them offered the solution for the problems described. The GRID technology consists of computing and storage resources distributed worldwide. Distributed computation is well suited for massive scientific calculations like detector simulations and data analysis in high energy physics. There are various materials describing GRID technology [1], [2] therefore we will focus further only on describing the specific aspects which appear in our case.

GRID technology is used in international scientific collaborations with *Large Hadron Collider* (LHC) [3] and *International Linear Collider* (ILC) [4]. Our group is involved in LHCb collaboration at LHC and FCAL collaboration at ILC. On our computing cluster (RO-15-NIPNE [5]) are installed and are running the specific software tools for two VO's LHCb and ILC. These two VO's are part of the *EU project Enabling Grids for E-SciencE* (EGEE) [6].

1.1. THE GRID INFRASTRUCTURE AND MIDDLEWARE.

The EGEE infrastructure consists of GRID services and GRID resources. The computing and data storage resources are shared by the participating GRID sites and distributed globally. The site operate *Computing Elements* (CE), which act as logical entry points to clusters of *Worker Nodes* (WN) and *Storage Elements* (SE) which provide interface to data storage systems. The core GRID services are *VO Membership Services* (VOMS), VO *Membership Registration Service* (VOMRS), *LCG File Catalogues* (LFC), *Workload Management Systems* (WMS) and *Resource Brokers* (RB). These services must exist at least once per VO.

The gLite middleware mediates between the user application and the operating system. It is deployed on most of the systems supporting the GRID infrastructure. It also provides an interface for the end-user, called *User Interface* (UI).

Access to the resources is provided by means of a service infrastructure. People, institutions and resources are collected by *Virtual Organizations* (VO) which are based on common sharing rules. Users authenticate against the GRID services of personal certificates and are authorized to use resources within a VO.

1.2. VIRTUAL ORGANIZATION LHCB AND ILC

VO LHCb is a global virtual organization dedicate to LHCb (*Large Hadron Collider Beauty Experiment*). LHCb is a High Energy Physics Experiment set on finding the solution to the mystery of the matter-antimatter imbalance studying CP violation in the universe at CERN's Large Hadron Collide.

VO ILC is a global virtual organization for the next important accelerator *International Linear Collider* (ILC). Stretching approximately 35-kilometers in length, this electron-positron collider will allow researchers to discover the Terascale, an energy region that may answer some of the most fundamental questions of all time.

1.3. DIRAC INFRASTRUCTURE DISTRIBUTOR AND GANGA JOB WIZARD

The DIRAC (*Distributed Infrastructure with Remote Agent Control*) software architecture is based on a set of distributed, collaborating services. Designed to have a light implementation, DIRAC is easy to deploy, configure and maintain on a variety of platforms. The relationship between resources, services, agents and clients which form the main components of DIRAC will be briefly discussed in turn below Figure 1.

Clients in DIRAC can simply be considered as submitters of jobs or requests. Clients include the Bookkeeping Query Webpage, which requests information about datasets and their replicas on the Grid. For distributed analysis and user production jobs, clients interact with the central services via the DIRAC *Application Programming Interface* (API). For LHCb production tasks, the Production Manager Console is used. This provides a general framework for the construction and management of production tasks and provides a GUI for users [7]. There is also a File Catalogue Browser which makes use of the Data Management components of DIRAC [8].

Services The DIRAC Job Management Services perform vital operations for production and distributed analysis jobs, such as uploading any necessary files for application steering and checking any requested input data is available. The Services highlighted accept requests from Clients and Agents. The configuration Service provides necessary site dependent information for Agents. The Job Monitoring Service keeps track of changes in job status. Similarly, the Bookkeeping Service will log selected results to provide a history about jobs in case of failure. The role of the Job Accounting Service is to provide statistics on job efficiency and resource usage at the site and user level.

Agents are deployed close to resources and form an Overlay Network. On LCG, Pilot Agents are deployed to Worker Nodes via the Resource Broker. These Pilot Agents in turn start a DIRAC Agent which on an individual PCs and site clusters this can be done manually.

Resources DIRAC can integrate resources such as Individual PC's, site clusters and Grids.

The Ganga (Gaudi/Athena and Grid Alliance) [9] [10] Grid front-end is used to submit distributed user analysis tasks for LHCb. Ganga is a joint project between ATLAS and LHCb that provides a common framework for cooperation in the configuration and management of tasks. Ganga provides a seamless way to submit jobs to several 'backends', these include: local; LSF; PBS; LCG; gLite; Condor and DIRAC. However, for LHCb Grid jobs, the default mode of submission is via DIRAC. Ganga makes use of the DIRAC API to configure, submit and monitor jobs. The Ganga client also offers a GUI, that will provide a seamless way for users to query the LHCb Bookkeeping Database for LFNs, as well as client-side splitting of jobs into smaller tasks. The functionality to support more complex workflows, is currently not available via Ganga. However, this is anticipated in the future. With Ganga submitting Grid jobs via the DIRAC API to the WMS, the LHCb VO has a seamless system that allows users to transparently submit jobs to batch systems, such as LSF, and the Grid. Users require little or no Grid knowledge to submit jobs via Ganga to DIRAC. The DIRAC job status machine is very refined in order to aid in the debugging of Grid jobs and increase redundancy. Ganga provides a simplified view of this for the user, in order to mask the underlying complexity.

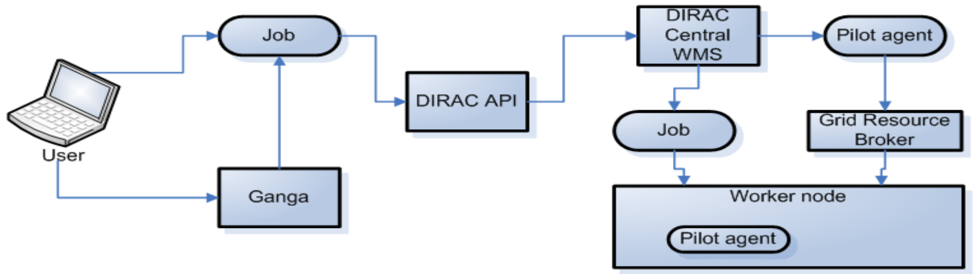

Fig. 1 – Jobs are submitted from Ganga via the DIRAC.

1.4. TECHNICAL DETAILS

Our middleware system was designed to take into account both EGEE requirements and local security implementation. As shown in figure 2, the system is divided in two parts. One has direct external access and the other is a local area

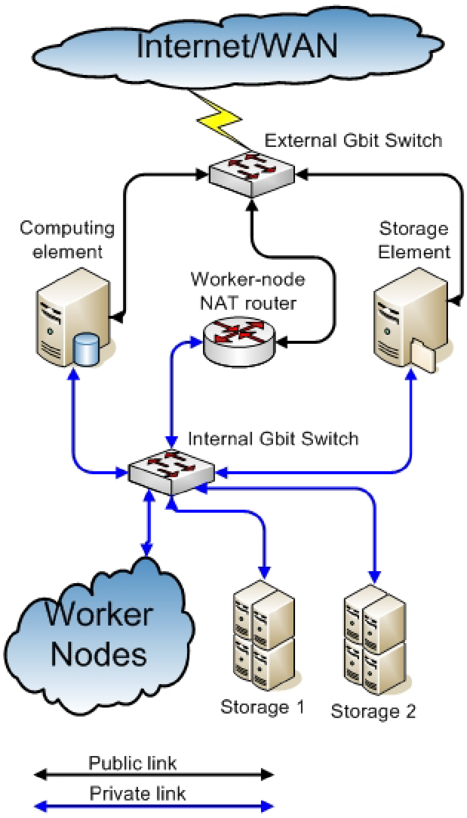

Fig. 2 – Hardware local configuration.

network that uses private addressing. According to LCG requirements, the CE and SE have externally visible IP addresses. The third system with public access is the router. It assures internet connectivity to Worker Nodes and also provides a general firewall for these. As for the hardware, one fully 100/1000G ethernet switch connects the worker nodes internally with the router, another 100/1000G switch connects the router, CE and SE to the external network. The Storage Element provides access to 2 x 6.4Tbytes of storage that is located on two dedicated, raid-based, storage systems. The actual installation started with the router installation and configuration, based on a Gentoo-Linux system. It provides network services such as NAT, DHCP, DNS, WEB and TFTP for network boot [11]. Other nodes are SLC4 based, now at subversion 8 and gLite 3.1. The storage machines that are connected to the SE, both run Gentoo Linux. Worker node installation was done in an automated fashion, using network boot and scripted anaconda installation.

1.5. PHYSICS ADAPTED SOFTWARE TOOLS

Events can be simulated independently and, in typical detector simulations, there is no state information carried from one event to the next. This implies that the order of the single events is unimportant and that multiple events can be simulated on different computers at the same time. Our site is involved in this process, the results being depicted in Fig. 3.

The LHCb tools [12] *for simulations*:
Gauss is the simulation program of b-physics, **Brunel** is the LHCb event reconstruction application using Geant4 and **DaVinci** is the physics analysis software, for the LHCb experiment, based on the **Gaudi** framework.
The ILC tools [13] *for simulations*:
Guinea-Pig is a tool for the simulation of beam-beam interactions at electron-positron colliders, **Mokka** is a full simulation using Geant4 and a realistic description of a detector for the future linear collider and **Marlin** is an analysis and reconstruction framework designed for simulated and measured data in ILC-related studies, based on the **LCIO** framework. **LCIO** defines a data model for ILC detector studies. It is a toolkit that helps to organize detector- and analysis-related data in running applications.

Our system is designed to meet the EGEE requirements and the local security implementation. It is already being used in the international scientific collaborations with Large Hadron Collider (LHC) and it successfully took part on the preliminary simulation tests in the collaboration with International Linear Collider (ILC). Software specific tools for both experiments were installed and tested.

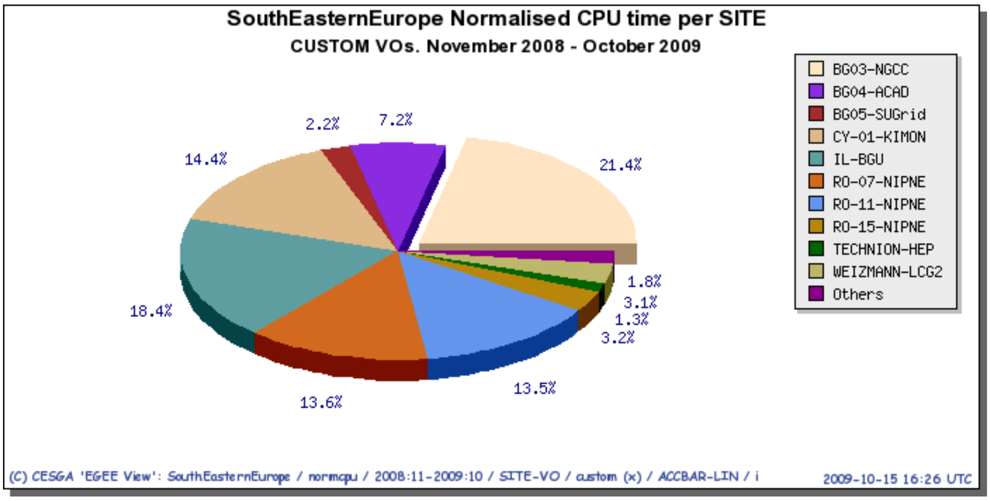

Fig. 3 – Normalised CPU time per SITE.

1.6. CONCLUSION

Simulations, for BeamCal detector studies, to test the reconstruction algorithms as well as to perform Monte-Carlo-based analyses have been carried out. Some results are presented in ref. [14].

Acknowledgements. This work has been supported by the Ministry of Education, Research and Innovation of Romania through the Authority CNCSIS under the IDEI Programme contract number 253/2007-2010 and Authority IFA under CEEX-M1 Programme contract 05-D11-81/2005 mainly in infrastructure.

REFERENCES

1. I. Foster and C. Kesselman, "The Grid 2: Blueprint for a New Computing Infrastructure", Morgan Kaufmann, 2003
2. F. Berman, G. Fox and A.J.G. Hey, "Grid Computing: Making the Global Infrastructure a Reality", John Wiley & Sons, 2003.
3. European Organization for Nuclear Research, http://www.cern.ch/.
4. International Lilear Collider project, http://www.ilc.org/.
5. http://goc.grid.sinica.edu.tw/gstat/RO-15-NIPNE/
6. http://www.eu-egee.org/
7. J. Closier, G. Kuznetsov, G. Patrick, and A. Tsaregorodtsev. DIRAC Production Manager Tools. In CHEP06, 2006.
8. A. C. Smith and A. Tsaregorodtsev. DIRAC: Reliable Data Management for LHCb. CHEP07, 2007.
9. Gaudi/Athena Grid Alliance. http://ganga.web.cern.ch/ganga/.

10. U. Egede, K. Harrison, D. Liko, A. Maier, J. Moscicki, A. Soroko, and C. Tan. GANGA - a Grid user interface. In CHEP06, 2006.
11. C. Coca, M. Orlandea, A. Rosca and M. Stanescu-Bellu "A Distributed Computing System Dedicated to High Energy Physics Experiments" Proceedings of the IASK International Conference E-Activity and Leading Technologies 3 - 6 December 2007, Porto, Portugalia (book) , ISBN: 978-972-99397-5-4 , pag 323- 328
12. http://lhcb-comp.web.cern.ch/lhcb-comp/
13. http://ilcsoft.desy.de/portal/software_packages/
14. E. Teodorescu, C. Coca, W. Lohmann, M. Orlandea, A. Sapronov "Expected Electromagnetic and Neutron Doses for the BeamCal at ILD" Rom. J. Phys. (in press)