The Nucleus, **46** (1-2) 2009 : 1-8

**The Nucleus**
A Quarterly Scientific Journal of Pakistan Atomic Energy Commission
NCLEAM, ISSN 0029-5698

# GRID COMPUTING IN PAKISTAN: OPENING TO LARGE HADRON COLLIDER EXPERIMENTS

N. BATOOL, A. OSMAN, A. MAHMOOD and *M. A. RANA[1]

Grid Computing Laboratory, CISD, PIEAS, P.O. Nilore, Islamabad, Pakistan

[1]Physics Division, PINSTECH, P.O. Nilore, Islamabad, Pakistan

(Received February 11, 2009 and accepted in revised form May 15, 2009)

A grid computing facility was developed at sister institutes Pakistan Institute of Nuclear Science and Technology (PINSTECH) and Pakistan Institute of Engineering and Applied Sciences (PIEAS) in collaboration with Large Hadron Collider (LHC) Computing Grid during early years of the present decade. The Grid facility “PAKGRID-LCG2” as one of the grid node in Pakistan was developed employing mainly local means and is capable of supporting local and international research and computational tasks in the domain of LHC Computing Grid. Functional status of the facility is presented in terms of number of jobs performed. The facility developed provides a forum to local researchers in the field of high energy physics to participate in the LHC experiments and related activities at European particle physics research laboratory (CERN), which is one of the best physics laboratories in the world. It also provides a platform of an emerging computing technology (CT).

**Keywords:** Computing grid, Large hadron collider, CERN, High energy physics, South Asia, LCG, Virtual organization

## 1. Introduction

Grid computing is a method for the coherent use of piled up power of a large number of computers present in distributed networks around the globe. A schematic of the grid computing concept is shown in Fig. 1. Owing to the joint use of computers in a grid for executing computational tasks or jobs, it can perform a large number of complicated computations without a single huge computational facility. Grid computing offers the possibility of the data communication within the grid and external to the grid. Communication within the grid is important for sending jobs and their required inputs to distributed network points within the grid as some jobs require a large amount of data to be processed which may not always reside on the machine running the job. The bandwidth available for such communications can often be a critical resource that can limit utilization of the grid. External communication is also valuable as it provides the connectivity among the geographically distributed grid users. Grid technology is a part of global efforts to build service oriented science that would democratize science by decreasing the gap between “haves” and “have-nots”. This technology has the potential to ample individual and collective scientific productivity and to build a culture of shared understanding [1]. Data storage, computational capacity and access are three pillars of the expanding computing grids [2-6].

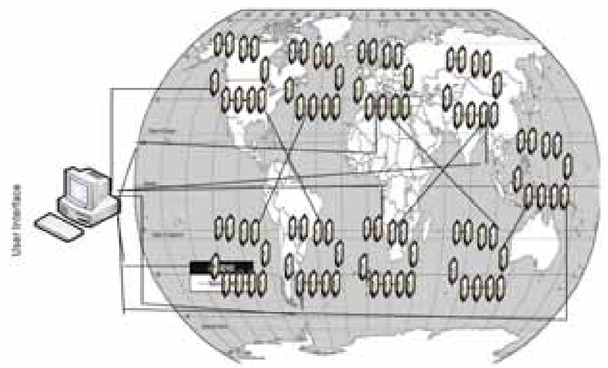


Figure 1. Grid Concept

This paper describes the implementation and function of the grid concept at PINSTECH/PIEAS achieved in collaboration with LHC, CERN, Switzerland. The functional status of the developed grid facility “PAKGRID-LCG2” is compared with similar facilities in South Asia. The aim of the paper is to introduce this facility to its potential users in Pakistan and offer them participate in the scientific challenges of this emerging CT.

* Corresponding author : marana@alumni.nus.edu.sg

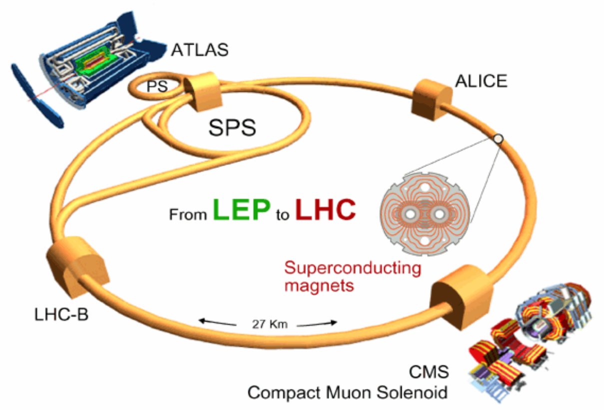


Figure 2 Layout of LHC. LEP is Large Electron Positron collider, SPS Super Proton Synchrotron and PS Proton Synchrotron. Other abbreviations are explained in the text.

## 2. Overview of LHC Related Experiments

The LHC (shown in Fig. 2) is the world's largest and most powerful particle accelerator and is expected to answer some of the fundamental questions about universe. The LHC will provide the highest energy proton-proton and ion-ion collisions ever happened in the laboratory. The smashing of protons in 27 km ring of accelerator produces enormous events which are expected to unfold the various mysteries of matter, time and space. These mysteries or questions include origin of mass, dark matter and energy, nature's favoritism for matter over antimatter, secrets of the big bang and matter of hidden dimensions. The LHC is a focal point for physics community around the world as it is expected to reply or resolve some of above mentioned big questions.

There are six distinct experiments in which LHC will be employed to investigate above mentioned issues. These experiments are run by international collaborations which bring together scientists from various institutes from all over the world. Subject of the two large experiments Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) involves a wide range of physics. These experiments will measure and analyze the myriad of particles produced by the collisions in the accelerator. These experiments have two independently detection systems for cross-confirmation of any new discovery made. Two medium-size experiments are A Large Ion Collider Experiment (ALICE) and Large Hadron Collider beauty experiment (LHCb). ALICE will study lead-lead ion collisions (quark-gluon plasma) whereas the LHCb experiment will investigate why universe is composed almost entirely of matter, but no antimatter. TOTal Elastic and diffractive cross section Measurement (TOTEM) and Large Hadron Collider forward (LHCf) are two other small size experiments. They focus on 'forward particles' (protons or heavy ions) which just brush past each other as the beams collide, rather than meeting head-on [7, 8].

The whole community of scientists and supporting staff, distributed around the globe, from LHC experiments are grouped into Virtual Organizations (VOs). These VOs operate under the umbrella of Worldwide LHC Computing Grid (WLCG). The service oriented architecture provides the standard procedures for users to access VOs to use these distributed resources. So the CERN operates in a massively complex, worldwide collaboration, comprising more than fifty four countries. Pakistan-CERN collaboration is mostly with CMS experiment and recently started with ALICE experiment as well.

## 3. LHC Grid Computing

The computational requirements of the LHC experiments are approximately 15 Peta Bytes of data (equivalent to more than 20 million CDs), which will be generated each year and more than 70,000 of today's fastest PCs processors would be required to analyze this data. The LHC experiments ALICE, ATLAS, CMS and LHCb are preparing for data acquisition planned to start in 2009. The LHC experiments are relying on WLCG [9]. The LHC grid infrastructure is based upon a Multi-Tier distributed model, addressing the needs of a computing system capable to handle LHC data, analysis and archival [10]. Fig. 3 illustrates the 4-Tier Model, where CERN as Tier-0 is the central production centre and will be responsible for distributing the raw data, and the Tier-1 centers spans several organizational units at sites distributed over a large geographical area, will be responsible for all the production/processing phases associated with the real data. The Tier-2 centers will be primarily Monte Carlo production centers, with both CERN and the Tier-1 centers acting as the central repositories for the simulated data [11]. The Tier-2 grid sites [9] provide services for CERN experiments and to local researcher's communities for grid-based analysis. The resources available with Tier-2 centers are accessible for accelerator experiments through the grid infrastructure. Tier-3 is bottom of the grid chain and it is the first level at which physicists have access to data under their own control. There are fewer constraints on bringing up a Tier 3 which is really the growth area. At Tier-3, scientists have

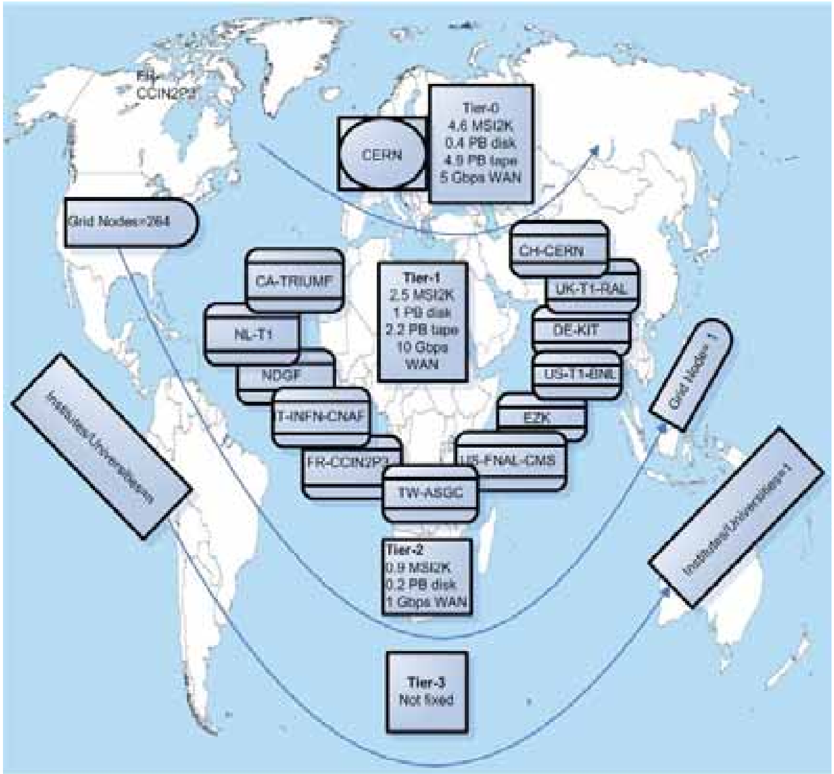


Figure 3. Tier Model of LHC computing grid.

more freedom to run the analyses they want and more flexibility in testing and debugging the code before it is used [12]. The participating sites have varying level of resources, organized hierarchically in Tiers. An important benefit of this approach is facilitating scientists all over the world to contribute intellectually without their presence at CERN [13].

The requirements for LHC data handling are very large, in terms of computational power, data storage capacity, data access mechanism, inter-tiers/inter-sites communication, performance and the associated human resources for development/operation/ support. It is not possible to get consolidated resources at CERN and even not considered feasible to find all of the resources at few sites; so that the WLCG computing service is implemented as a geographically distributed resource platform (Fig. 3). This means that computing resources, both computational and storage, installed at a large number of regional computing centers in many different countries, interconnected by fast networks. All the WLCG distribution is further handled under five virtual organizations, users are grouped and access to resources is provided under the umbrella of VOs such as CMS, ATLAS and ALICE etc. [14-16]. To use WLCG resources one must be affiliated at least with one of the VOs of CERN. The registration with VO is required so as the information will be forwarded to the VO administration and grid resource providers for validation. The VOs manage immediate allocation, provisioning of resources, based on their availability and the provision of secure access to the resources. A VO must be capable of providing facilities for dynamic collection of resources based upon users to solve specific problems associated with its task. The associated grid production infrastructure is comprised of more than 250 sites across 54 countries and approximately 140 Regional Centers which are participating as Tier-2 centers of CERN computing infrastructure. The WLCG architecture is developed for Particle Physics; it is also suited for other computer intensive applications.

The Grid has witnessed rapid development in the last few years. It currently provides reliable software components for resource integration, task scheduling and computer safety, adapted to distributed, heterogeneous, and dynamic resources. Tier-0 center located at CERN while the data will be distributed to a series of large computer centers which are part of Tier-1 in the model. These centers provide services for data archiving, reconstruction, calibration and other data-intensive analysis operations. Tier-1 centers will make data available to Tier-2 centers which provide adequate computing power for end-user analysis tasks and Monte Carlo simulations [17]. The job of the WLCG is to prepare the computing infrastructure for the simulation, processing and analysis of all detectors before and after the LHC experiments. This includes both the common infrastructure of libraries, tools and frameworks required to support the physics application software and the development and deployment of the computing services needed to store and process the data, providing batch and interactive facilities for the worldwide community of scientists.

## 4. PAKGRID Computing Facility

### *4.1. Implementation*

The cooperation between Pakistan and CERN has been started since 1994 to participate in innovative and challenging initiatives by CERN. Pakistani engineers and scientists are taking active part in CMS and ALICE experiments, such as, Fabrication of Magnet Support Structure for the CMS Magnet Barrel Yoke, Fabrication and Production of RPCs, Precision Position Monitoring/Alignment System specific activities of WLCG. Pakistan participated in Grid developments at CERN since 2002, although at the early stages more than five establishments started developing the grid site but later on two institutes PINSTECH, Pakistan Atomic Energy Commission (PAEC), Islamabad and National Center of Physics (NCP), Islamabad succeeded in this effort. The PINSTECH-CMS Production Center project began in 2002, initially a test-bed was installed with the dial-up link to CERN. The lab is renovated for CERN CMS production activities and established the Giga Bit LAN to setup a PCs Farm. High computing resources and storage devices were deployed to replicate the CERN computing environment.

A PCs Cluster using Linux environment was established at PINSTECH to offer computing resources since 2003 and computed millions of events (collision of two particles in a bunch of particles is termed as an event). It produced simulated data for CMS detector in collaboration with other countries. In mid-2003, PAKGRID emerged as one of the first production-level Grid site in Pakistan and continued the grid activities up-till now. Events were assigned to computing collaborators of CERN according to their capacity. The site is registered with the name of PAKGRID-LCG2 and participating in WLCG. The first big goal of all such activities is to establish as Tier-2 production center and establish the grid node. The main focus of the project is to collaborate with large-scale scientific research carried out through distributed global collaborations by CERN and to promote a grid infrastructure in Pakistan which is an emerging technology. Following the vision of Grid about access of computing resources to all, PAKGRID-LCG2 has a collaborated work with a number of institutes at national level to provide high end computing solutions to local professionals. This group helps local researchers to share the experience of deployment of PCs Cluster infra-structure. PAKGRID-LCG2 production grid exploits a total of 11 CPU’s including core services and Worker Nodes (WNs). Scientific Linux Operating System 4.6 and grid deployment software Glite 3.10 are being used at this computing facility. The facility is available to users 24 hours a day, 7 days a week.

### *4.2. Grid infrastructure in South Asia*

In South Asia Pakistan and India are linked with WLCG-CERN as Tier-2 Centers, there are two sites from Pakistan and four sites from India. Though some sites are newly introduced and three sites such as PAKGRID-LCG2, NCP-LCG2 and INDIACMS-TIFR are operational since 2003 with collaboration of CMS experiment at CERN. There are some basic parameters which can be analyzed for generic comparison of grid sites of South Asia, such as computing resources, storage resources and the connectivity of these sites with rest of the grid frame-work. The Grid infrastructure at all the sites in this region is more or less the same, based on the most recent LCG-2 middleware and contains most of the important elements so as to be a Grid Node, such as; Computing Element (CE) with WN, Information index, User Interface (UI) and a Storage Element (SE) with an attached storage system. The Grid Information System provides the information about the structure of the grid using a set of rules which is same for all. While analyzing and reviewing these sites the other important factor is the connectivity and bandwidth allocation which plays the main role in performance. The overall resource structure of all these sites is given in Table 1 [18].

Table 1. Comparison of technical specifications between PAKGRID and other similar grid nodes.

| Site Name | CPU | SE Available TB | Bandwidth |
|---|---|---|---|
| PAKGRID-LCG2 | 11 | 4.07 | 4 Mbps |
| NCP-LCG2 | 38 | 1.68 | 10 Mbps |
| INDIACMS-TIFR | 193 | 72.30 | 34 Mbps* |
| IN-DAE-VECC-01 | 26 | 0.11 | 34 Mbps* |
| IN-DAE-VECC-02 | 156 | 67.63 | 34 Mbps* |
| IN-DAE-VECC-EUINDIAGRID | 18 | 1.24 | |
| Data taken at PAKGRID-LCG2 on 09-03-2009<br>*TIFR Mumbai, VECC Kolkata at 34 Mbps [18] | | | |

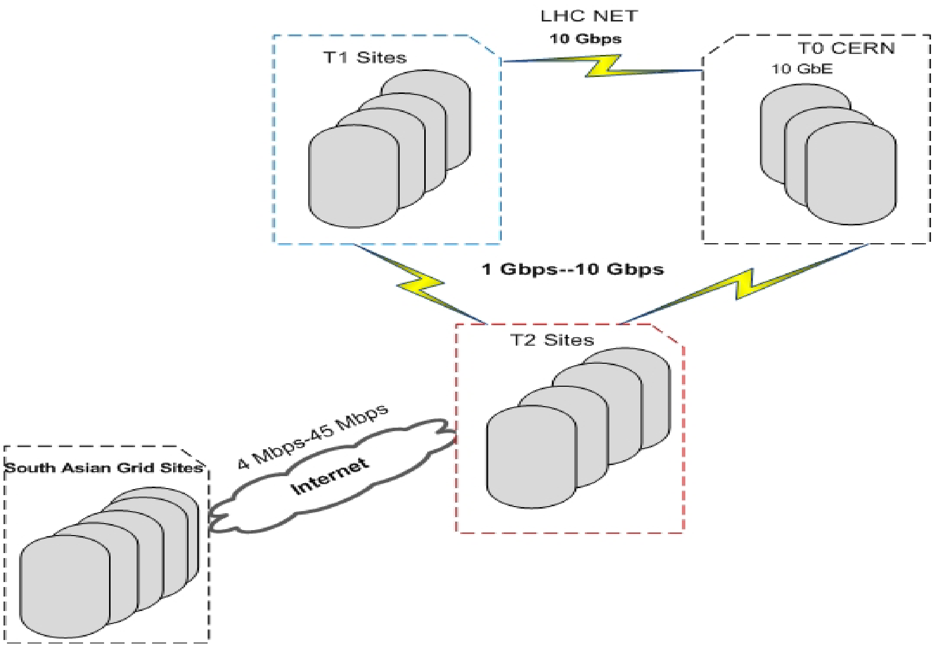


Figure 4. Limitation of internet connectivity of South Asian grid sites.

The Tier-2 centers are intended to provide computing resources for processing and analysis capabilities to local and remote user communities. Tier-1 centers are expected to store a portion of the raw data, and Tier-2 sites needs to provide minimal set of services and expected to provide the network capacity so as quality of data transfer can be maintained. The user's jobs requirements are commonly based on the bandwidth capacity of the end-to-end network path. The operational model for Tier-2 sites is developed with an understanding that set of network capacity requirements for Tier-2 should be flexible enough so as to manage the service, the minimum 2.5 Gbps [19] is recommended.

The sites operational in the South Asia region are generally using South-East Asia, Middle East, and Western Europe (SEAMEWE-3,4) submarine optical fiber link [20]. Pakistan has so far been connected to the outside world using various circuits having the capacity of 2.4-10 Giga band-width. The most important link is the 155 Mbps SEAMEWE-3 submarine optical fiber link while the bandwidth in India at the moment, is slightly higher [20-21]. However, the limited bandwidth is allocated to all grid infrastructures (see Table 1) in South Asia as compared to rest of the grid sites associated with LCG. The smoke ping results [22] show the packet loss and delays which the South Asian Grid sites experience because of the unavailability of dedicated WAN links to LCG (Fig. 4).

## 5. Discussion on Functional Status

Present activities concerning distributed analysis are in at R&D stage with the focus on providing an end-to-end analysis system [24]. In general end-user analysis is a chaotic, non-organized task, carried on concurrently by many independent users that do not have a deep knowledge of distributed environment. Some of the tools that act as an interface to the physicist have been deployed at PAKGRID for global grid file systems ROOT and ALIEN to provide analysis environment for ALICE and CMS experiments. Monitoring services are needed to monitor and archive information about site resources (collecting e.g. host-related metrics, network activity, disk space, etc.). Job monitoring allows the user to extract job specific information and thus to monitor the job status in real time. The Grid as a whole is monitored after every hour, jobs processing/replication status, checking network stability and reliability during the daily grid sites activities. The grid network infrastructure worldwide is mostly based on dedicated wide area network (WAN) links with 10Gbps bandwidth support. Our limitation is to access the same Grid infrastructure by using the limited bandwidth of about 4Mbps.

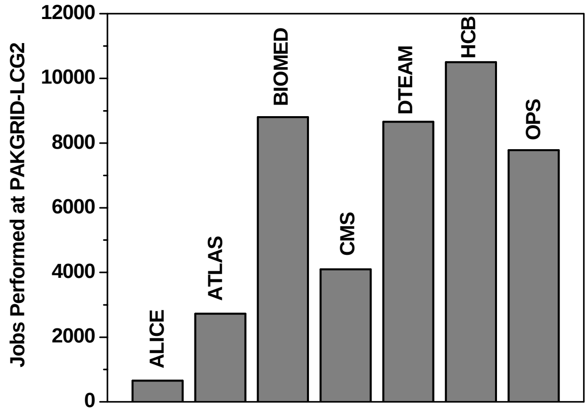


Figure 5. Total number of jobs (SITE and VOs) at PAKGRID-LCG2 during November 2007 – October 2008.

Considering the IT infrastructure in Pakistan, it has been a difficult task to establish and then maintain its functional status. A workload of 6000-10000 jobs/month was achieved by PAKGRID-LCG2. Fig. 5 shows total number of computational jobs completed at PAKGRID-LCG2 during the period of 1 year (November 2007 – October 2008). Computational jobs related to various LHC experiments were completed as clear from Fig. 5. Since the end nodes and network paths are globally distributed, usually across organizational and management boundaries, it can be difficult to locate and identify the factors that are effecting the grid sites performance. In order to analyze these types of performance problems the role of ISPs and the configuration at their end, matters. The network path problems are obviously beyond the scope and control of grid sites' local setups. If the bandwidth analysis does not uncover significant concerns, then "number of jobs" analysis of sites [21] provides the comparison of expected versus actual performance pattern of the grid sites. The accounting scenarios, which can be viewed as potential of all the sites in this region (Fig. 6), exhibit its status briefly. The statistics shown corresponds to the period of 2005-2008 [16]. The number of jobs on PAKGRID-LCG2 decreased in 2008 due to down time of 2 months in which physical shifting of grid facility from PINSTECH to PIEAS was carried out. Internet service provider (ISP) PAKNET was merged with another ISP Pakistan Telecommunication Corporation Limited (PTCL) with future responsibility of the service on PTCL. Problem related to resolution of grid machines IPs in the new merged environment added an additional down time of 5 weeks.

Like any deployed grid node, PAKGRID-LCG2 has three functions which are user support, system support and grid operations. System support includes certification and testing of identified packages provided by the grid development projects or other middleware providers, system configuration and facing up the system failures or failures halting the functionality of the system. Number of grid jobs depends upon the access provided to the number of VOs. VO is a cyber environment or virtualization of a project which allows researchers of project (around the globe) to routinely access the computing resources including models and data. Our grid site PAKGRID-LCG2 provide access to seven VOs CMS, ALICE, ATLAS, OPS, DTEAM, LHCB and BIOMED. Another Pakistani grid site NCP-LCG2 does the same whereas Indian grid sites IN-DAE-VECC-01 and IN-DAE-VECC-02 are accessible to ALICE, OPS and DTEAM. Another Indian grid site INDIACMS-TIFR provides access to CMS, OPS and DTEAM.

To ensure smooth and trustworthy operation of a complex system (like WLCG) composed of interrelated data samples, stored at different centers around the globe and accessible online both locally and remotely, an exercise is needed to validate computing models, software, data models, and to ensure the correctness of all technical

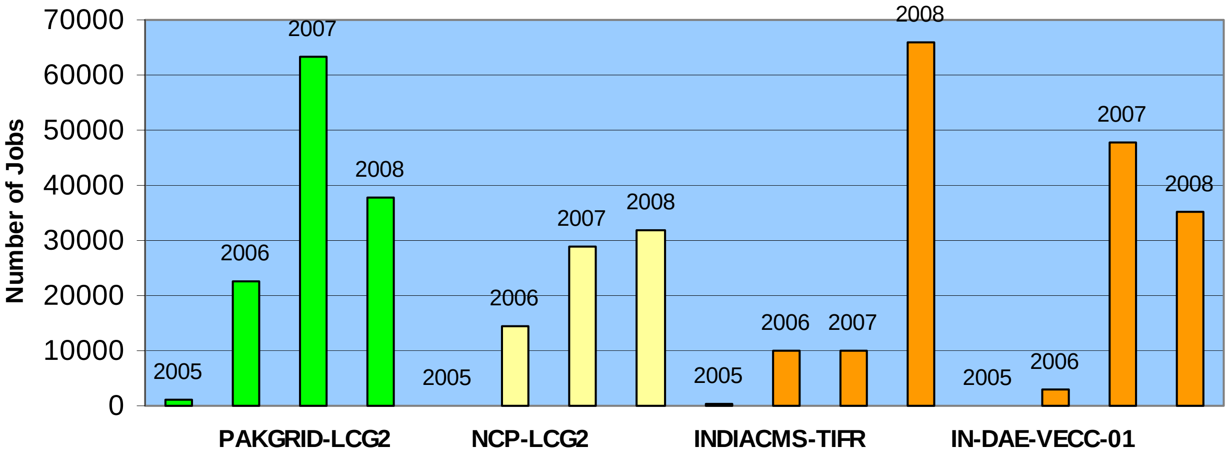


Figure 6. Jobs comparison of South Asia Pacific sites (2005-08) [16].

choices. This exercise is termed as Data Challenge which involves tens of institutes and hundreds of physicists. In a data challenge, large samples of data are simulated via advanced simulation programs and analyzed as if they were coming from the real experiment. In LHC Data Challenge 2004 and 2003 [23], the LCG environment was tested the functionalities for distributed computing. PAKGRID-LCG2 participated in addressing these challenges and development of infrastructure for data analysis.

## 6. Conclusions

This report is focused on the Grid Node essential bits and nuts as well as its uses and deployment in a research environment in Pakistan. A grid node provides an access to a layer of software that uses the Grid to gain access to data and resources related to the project involved and it also provide physicists with a user friendly interface for submitting the analysis jobs. The grid node PAKGRID-LCG2 provides access to worldwide dispersed users for the required grid services related to LHC computing grid. Major issues are related to the performances which are directly or indirectly linked to such components which are usually out of bound for grid node deployment team such as; power failure, Internet Protocol or IP resolving and internet bandwidth which is comparatively less than that required for LHC computing grid. Solution to these problems may be searched in a systematic study of functional grids in similar environments like India and some parts of China. On the other hand these grid nodes are the gate ways to participate in the scientific challenges and to launch this new CT in unexposed regions. Local adoption of this new service-oriented science and business culture will require Pakistani academic and research communities learn the use of grid technology and its evolution.

## Acknowledgement

NB and AO wish to acknowledge the hospitality of CERN for their visits to CERN which provided them a chance of getting training on Grid Computing. AM thanks to Pakistan-CERN Collaboration for sponsoring his visit to CERN.

## References


[1] I. Foster, Science **308** (2005) 814.

[2] LHC Computing Grid (http://public.web.cern.ch/public/en/LHC/Computing-en.html).

[3] M. Girone, J. Phys. Conf. Series **119** (2008) 052017.

[4] J. M. Hernández et al., J. Phys.: Conf. Series **119** (2008) 052019.

[5] J. Shiers, Comp. Phys. Comm. **177** (2007) 219.

[6] A. Opitz, H. Koning and S. Szamlewska, J. Grid. Comp. **6** (2008) 385.

[7] The LHC experiments (http://public.web.cern.ch/public/en/LHC/LHCExperiments-en.html).

[8] G. Cancio et al., Current Status of Fabric Management at CERN, Proc. of the Computing in High Energy Physics (CHEP) Conference (2004), p. 1145.

[9] Stefan Stonjek, Dirk Düllmann, Gordon Brown, Distributed Database Deployment for the LHC Experiments, a talk on 02 Nov. 2005 at Edinburgh (http://www.nesc.ac.uk/talks/608/Day2/3D talk_Edinburgh_Nov-2005.ppt).

[10] Lyndon Evans and Philip Bryant, The Journal of Instruments **3** (2008) S08001; also see other articles in the same issue.

[11] I. Foster, J. Gieraltowski, S. Gose, N. Maltsev, E. May, A. Rodriguez, D. Sulakhe, A. Vaniachine, J. Shank, and S. Youssef, The Grid 2003 Production: Principles and Practice in 13th International Symposium on High-Performance Distributed Computing (HPDC-13 2004), June 4-6, 2004, Honolulu, Hawaii, USA

[12] Kathryn Grim, Tier-3 computing centers expand options for physicists, A Feature by International Science Grid This Week, ISGTW, Jan. 14, 2009 (http://www.isgtw.org/?pid=1001578).

[13] J. Andreeva; B. Gaidioz; G. Maier; R. Rocha, P. Saiz; C. Catalin, I. Sidorova, S. Wakefield and J. Herrala, New monitoring applications in the experiment dashboard in 3rd EGEE User Forum, Feb. 11-14, 2008, Clermont-Ferrand, France.

[14] CMS ECAL Hardware (www.hep.ph.ic.ac.uk/~mryan/cms/CMS5.jpeg).

[15] The CMS detector ( ireswww.in2p3.fr/.../stephanie/images/CMS_3D.gif).

[16] EGEE Accounting Portal for grid http://www3.egee.cesga.es/gridsite/accounting/CESGA/tier2 view.html).

[17] D. Bonacorsi and Tiziana Ferrari, WLCG Service Challenges and Tiered Architecture in the LHC Era, Springer Milan, Italy (2007).

[18] D. Singh, EU-India Network Infrastructure & Indian Network Perspective, EU Funded Project No. RI-031834, Feb. 20, 2007.

[19] The LHC Computing Grid project, (http://lcg.web.cern.ch/LCG/).

[20] Optical Fiber Submarine Cable System Linking South East Asia to Europe via the Indian Sub-Continent and Middle East (http://www.seamewe3.com/inpages/cable_system.asp).

[21] K V Ratnakar, Indian Council of Medical Research (ICMR) Library Bulletin, Shri, Apri-June, 2008.

[22] ASGC Grid Monitoring using Smoke Ping (https://osadm.grid.sinica.edu.tw/smokeping/smokeping.cgi ).

[23] W. von Rueden and R. Mondardini, IEEE TCSC Newsletter **5** (2003) article No. 3, Issue 1.

[24] J. Andreeva et al., IEEE Nucl. Sci. Symp. Conf. Record **4** (2004) 2029.