Journal of the Korean Physical Society, Vol. 55, No. 5, November 2009, pp. 2067∼2071

The Data Processing of e-Science for High-Energy Physics

Minho Jeung, Hyunwoo Kim, Kihyeon Cho* and Ok-Hwan Byeon
Supercomputing center, Korea Institute of Science and Technology Information, Daejeon 305-806

(Received 29 September 2008)

The goal of high-energy physics is to understand the basic properties of elementary particles and their interactions. High-energy physics is usually conducted at major accelerator sites, in which detector design, construction, signal processing, data acquisition, and data analysis are performed on a large scale. However, in order to study high-energy physics anytime and anywhere even if we are not on-site at accelerator laboratories, we have created a new research paradigm, e-Science. The e-Science for high-energy physics has three components: data production, data processing, and data analysis. In this paper, we focus on the data processing of e-Science for high-energy physics. We show current implementations and experiments of data processing for the ALICE (A Large Ion Collider Experiment) Tier 2 center and for CDF (Collider Detector at Fermilab) grid farms.

PACS numbers: 29.85+c, 07.05-t, 07.05.Bx
Keywords: e-Science, High-energy physics data grid, Grid computing, Data processing
DOI: 10.3938/jkps.55.2067

I. INTRODUCTION

The goal of HEP (high-energy physics) is to understand the basic properties of elementary particles and their interactions. Since the invention of the cyclotron by Ernest Orlando Lawrence at the University of California [1], HEP has usually been investigated at major accelerator sites, in which detector design, construction, signal processing, data acquisition, and data analysis are performed on a large scale. In order to cope with more data and more collaboration, we have created the concept of e-Science. The goal of e-Science for high-energy physics is to study high-energy physics anytime and anywhere even if we are not on-site at an accelerator laboratory. If computing processing is to be performed on the required HEP scale, data grid technology is a strong requirement [2]. The amazing advance in IT (information technology), such as Moore's law, and the wide spread use of IT help computing processing [3].

The objective of a HEP data grid is to construct a system to manage and process HEP data and to support a high-energy physics community. In this paper, we introduce data processing as a component of e-Science for ALICE (A Large Ion Collider Experiment) and CDF (Collider Detector at Fermilab) experiments. We have built an ALICE Tier 2 center and a CDF Analysis Farm based on LCG (large hadron collider (LHC) computing grid) farm.

*E-mail: cho@kisti.re.kr; Fax: +82-42-869-0789

II. DATA PROCESSING OF E-SCIENCE

1. e-Science

e-Science is a new research paradigm for science, which is computationally intensive science [4]. Consequently, e-Science uses immense data sets that require grid computing and is carried out in highly-distributed network environments [4]. HEP requires a particularly well-developed e-Science infrastructure due to its need for adequate computing facilities for the analysis of results and the storage of data originating from accelerator laboratories [5]. Therefore, HEP is one of the best applications for e-Science [5]. The components of e-Science are 1) data production, 2) data processing, and 3) data analysis. In this paper, we focus on data processing. First, data production refers to both an on-line shift and an off-line shift anywhere. Second, data processing processes data by using high-energy physics data and supports the high-energy physics community [6]. Third, data analysis is used by collaborations around the world to analyze and publish the results in collaborative environments.

2. Data Processing as a Component of e-Science

The main infrastructure of data processing for e-Science is network and computing resources. The first infrastructure is both a domestic network (KREONET) and an international network (GLORIAD). The KREONET (Korea Research Environment Open NETwork) is a 1 $\sim$ 40 Gbps national research and development network operated by the Korea Institute of Science and Technology Information (KISTI). The KREONET is also connected to GLORIAD (Global Ring Network for Advanced Applications Development). The GLORIAD is built on 10 Gbps of a fiber-optic ring of networks around the northern hemisphere of Earth and linked with Russia, China, Korea, Canada, USA, the Netherlands, and the five Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden). It provides scientists, educators, and students with advanced networking tools that improve communications and data exchange, enabling active, daily collaboration on common problems [7]. With GLORIAD, the high-energy physics community can move unprecedented volumes of valuable data effortlessly, stream video, and communicate through quality audio- and video-conferencing [7]. The KISTI is directly connected to CERN (European Organization for Nuclear Research) via a 10-Gbps network with GLORIAD.

The second infrastructure is computing resources. For current and future HEP activities for large-scale data, the HEP data grid is indispensable [5]. We need to maintain a mass storage system of hard disks and tapes in a stable state [5]. If the HEP data are to be made transparent, CPU power should be extendable and accessible [5]. Transparent HEP data means that the data should be analyzed even if high-energy physicists as users do not know the actual source of the data [6].

III. ACHIEVEMENTS

1. Data Processing with the ALICE Experiment

The objective of the ALICE experiment at CERN, as with the RHIC (Relativistic Heavy Ion Collider) experiments at BNL (Brookhaven National Laboratory) [8], is to study the physics of strongly interacting matter at extreme energy densities, where the formation of a new phase of matter, the quark-gluon plasma, is expected. ALICE has been conducted at CERN, where detector design, construction, signal processing, data acquisition, and data analysis are performed. The data size will be a few PB of data per year. In order to handle this amount of data, we use grid technology. For this work, we assembled a LCG farm at KISTI for the ALICE experiment. The LCG is a worldwide infrastructure where all the computations relevant to the analysis of the data coming out of the four LHC experiments are taking place. The LCG organization involves a hierarchy of computing centers from CERN, labeled Tier 1, Tier 2 and Tier 3.

In 2007, the MOST (Ministry of Science and Technology) in Korea and CERN had a MOU (Memorandum of Understanding) to build and operate the ALICE Tier 2 center at the KISTI. Table 1 shows the ALICE Tier 2 computing capacities at the KISTI on the MOU [9]. The ALICE Tier 2 center is using LCG farms (KR-KISTI-GCRT-01), which consist of 120 kSI2K CPU and 30 TB of storage [9]. The farm currently maintains a 96 percent operating capacity (8,000 jobs per month). Now, the ALICE Tier 2 center at the KISTI has become a federation of global ALICE farms, and consists of a 13,804 kSI2K CPU and a 99.52 PB disk around the world. Currently, around 1,000 physicists from 109 institutes in 31 countries use the ALICE farms.

2. Data Processing with the CDF Experiment

A. Introduction

The CDF is an experiment on the Tevatron in Fermilab. The CDF group began its Run II phase in 2001. CDF computing needs include raw data reconstruction, data reduction, event simulation, and user analysis [10]. Although very different in the amount of resources needed, they are all naturally parallel activities [10]. The CDF computing model is based on the concept of the CAF (Central Analysis Farm) [11]. To date, Run II has gathered more than 2 PB of data of 4 fb^{-1}. The increasing luminosity of the Tevatron collider has caused the computing requirement for data analysis and Monte Carlo production to grow larger than the dedicated CPU resources that are available [12]. In order to meet future demand, CDF has examined the possibility of using shared computing resources. CDF is using several computing processing systems, such as CAF, DCAF (Decentralized CDF Analysis Farm), and grid systems. A Korea group has built a DCAF for the first time [6]. Finally, we have constructed the CDF Grid farm at KISTI and the LCG farm for the ALICE experiments. Only 60 % of the CPU resources of the CDF experiments run at a Central Analysis Farm inside Fermilab. Now, global DCAF and grid farms constitute 40 % of the total CPU resources for the CDF experiment [6].

B. CAF (Central Analysis Farm)

In 2001, we built CAF, which is a cluster farm inside Fermilab in the USA. The CAF was developed as a portal. A set of daemons accept requests from the users via kerberized socket connections and a legacy protocol [10]. Those requests are then converted into commands to the underlying batch system, which does the real work [10]. The CAF is a large farm of computers running Linux with access to the CDF data handling system and to databases to allow the CDF collaborators to run batch analysis jobs [11]. The submission uses a CAF portal, which has two special features. One is that users can submit jobs from anywhere. The other feature is that

Table 1. The ALICE Tier 2 computing capacities at the KISTI.

KISTI,	Pledged	Planned to be pledged				
Daejon, Korea	2007	2008	2009	2010	2011	2012
CPU (kSI2K)	50	100	150	150	150	150
Disk (TB)	30	30	50	50	50	50
Nominal Wan (Mbps/s)	10000	10000	10000	10000	10000	10000

job output can be sent directly to a desktop or stored on a CAF FTP server for later retrieval.

C. DCAF (Decentralized CDF Analysis Farm)

In 2003, we built DCAF, a cluster farm outside Fermilab for the CDF collaboration so that the CDF users around the world could use it like the CAF at Fermilab. The DCAF enables a user to submit a job to the cluster either at CAF or at the DCAF. In order to run the remote data stored at Fermilab in the USA, we used SAM (Sequential data Access via Meta-data) [6]. We used the same GUI (Graphic User Interface) used in the CAF. The difference is only in the selection of the analysis farm for the DCAF [6].

D. CDF Grid

In 2006, we built CDF grid farms in North America, Europe, and Pacific Asia Areas. HEP user activity patterns are CPU intensive and require large data file transport. These activity patterns require a change in the HEP computing model from clusters to a grid to meet the hardware requirements. Dedicated Linux Clusters on the FBSNG (Farm Batch System Next Generation) batch system were used when CAF was launched in 2002. However, the CAF portal has gone from interfacing to a FBSNG-managed pool to Condor as a grid-based implementation, without the need for users to learn new interfaces [10].

Now, we have adapted and converted the workflow to the grid. The goal of the movement to a grid at the CDF experiment is the world-wide trend for HEP experiment. We need to take advantage of global innovations and resources because the CDF still has many data to be analyzed. The CAF portal is allowed to change the underlying batch system without changing the user interface. The CDF used several batch systems: 1) FBSNG, 2) Condor, 3) Condor over Globus, and 4) gLite WMS (Workload Management System). The third and the fourth are grid-based production systems. The North America CDF Analysis Farm and the Pacific CDF Analysis Farm use a Condor over Globus model whereas the European CDF Analysis Farm uses a gLite WMS (Workload Management System) model. Table 2 summarizes a comparison of the Grid CDF Analysis Farm [5].

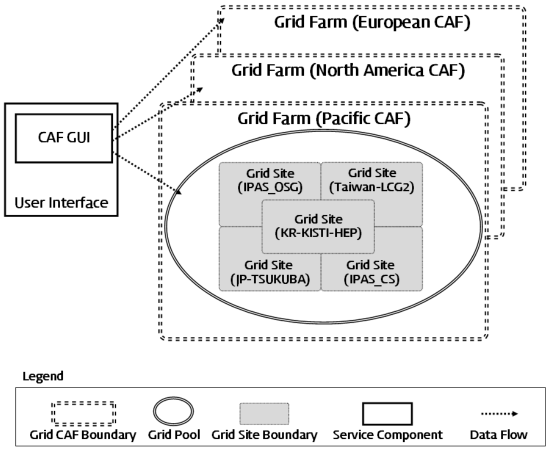

Fig. 1. The Scheme of CDF grid farms.

Figure 1 shows a schematic of the CDF grid farms. The users submit a job after they input the required information about the job (shell script to run, local path to executables, shared libraries, *etc.*) into a kerberized client interface [11]. The Condor over Globus model uses a created virtual private Condor pool out of grid resources [10]. Job containing Condor daemons are also known as a glide-in job [10]. The advantage of this approach is that all Grid infrastructures are hidden by the glide-ins [10]. The gLite WMS model talks directly to the gLite WMS, also known as the Resource Broker [10]. It allows us to use grid sites where the Condor over Globus model would not work at all and is adequate for grid job needs [10]. Because a Condor-based grid farm is more flexible, we applied this method to the Pacific CDF Analysis Farm.

E. Pacific CDF Analysis Farm

The regional CDF Collaboration of Taiwanese, Korean, and Japanese groups have built the CDF Analysis Farm, which is based on grid farms. We called this federation of grid farms the Pacific CDF Analysis Farm. Figure 2 shows the components of the farm.

Table 2. Comparison of Grid CDF Analysis Farms.

Grid CDF Analysis Farm	Head node	Work node	Grid middleware	Method	VO (Virtual Organization)
North America CDF Analysis Farm	Fermilab (USA)	USSD (USA) *etc.*	OSG	Condor over Globus	CDF VO
European CDF Analysis Farm	CNAF (Italia)	IN2P3 (France) *etc.*	LCG	WMS (Workload Management System	CDF VO
Pacific CDF Analysis Farm	AS (Taiwan)	KISTI (Korea) *etc.*	LCG, OSG	Condor over Globus	CDF VO

Table 3. Comparison of the Pacific CDF Analysis Farm and the CGCC.

	Pacific CDF Analysis Farm at KISTI	CGCC at KISTI
Head node	paccaf.phys.sinica.edu.tw	New Head node
VOMS (Virtual Organization Management Service) server	voms.cnaf.infn.it	New VOMS server
Work nodes	13 nodes (25 core processes)	KISTI Supercomputer (Sun Blade 6048 Cluster system)
Storage	15 TB(NFS)	300 TB(dCache)

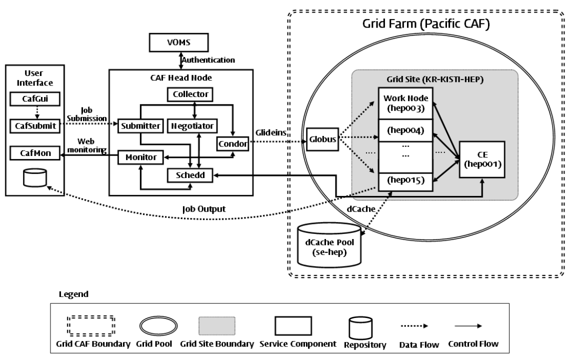

Fig. 2. The components of the Pacific CDF analysis farm.

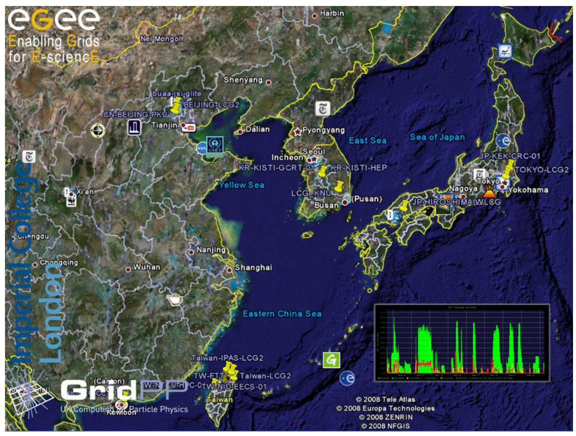

Fig. 3. The LCG monitoring system of the Pacific region. At the KISTI, we see both the ALICE Tier 2 Center (KR-KISTI-GCRT-01) and the Pacific CDF Analysis Farm (KR-KISTI-HEP).

The Pacific CDF Analysis Farm is a distributed computing model on the grid. It is based on the Condor glide-in concept, where Condor daemons are submitted to the grid, effectively creating a virtual private batch pool [10]. Thus, submitted jobs and results are integrated and are shared in grid sites. For work nodes, we use both LCG and OSG (Open Science Grid) farms. The head node of the Pacific CDF Analysis Farm is located at the Academia Sinica in Taiwan. Now, it has become a federation of one LCG farm at the KISTI in Korea (KR-KISTI-HEP), one LCG farm at the University of Tsukuba in Japan (JP-TSUKUBA-U-03), and one OSG and two LCG farms in Taiwan (IPAS_OSG, IPAS_CS, Taiwan-LCG2).

F. CGCC (CDF Grid Computing enter)

In 2008, we made a plan to build a CDF Grid Computing Center (CGCC), which is a big grid farm with a storage farm. Based on the CDF experiment in 2008, we knew that the expected CPU requirement would be about 65,000 kSPI2K. However, available CPUs at Fermilab are about 5,000 kSI2K [13]. In addition, guaranteed resources are needed because the LHC is starting. CGCC has been proposed and created in countries where there are big computing centers [13]. The candidate sites are IN2P3 in France, CNAF in Italy, and KISTI in Korea. By 2011, the KISTI will purchase 2,000 CPUs and 1,500 TB storages for high-energy physics experiments - CDF, ALICE, and Belle. Among the resources, we will make CGCC, which will be 300 TB of new systems for CDF data with Fermilab. Table 3 shows a comparison of the Pacific CDF Analysis Farm and the CDF Grid Computing Center at the KISTI.

IV. CONCLUSION

Figure 3 shows the LCG monitoring system of the Pacific region. At the KISTI, we see both the ALICE Tier 2 Center (KR-KISTI-GCRT-01) and the Pacific CDF Analysis Farm (KR-KISTI-HEP), both of which run successfully. We also see the work node of the Pacific CDF Analysis Farm in Taiwan (Taiwan-LCG2).

The e-Science for high-energy physics shows a new paradigm for science. As one component of e-Science for high-energy physics, we have introduced data processing. We succeeded in developing and installing high-energy physics data grid farms. Currently, we are successfully running the official ALICE Tier 2 center and the Pacific CDF Analysis Farm. Conclusively, high-energy physics is a good leading area for data processing in e-Science.

ACKNOWLEDGMENTS

We would like to thank Beobkyun Kim and Soonwook Hwang (Korea Institute of Science and Technology Information) for LHC computing grid farm, Yuchul Yang (Kyungpook National University), Mark Neubauer and Frank Wüerthwein (University of California at San Diego) for Central Analysis Farm, and Igor Sfillioi (Fermilab) and Hsieh Tsan Lung (Academia Sinica) for the Pacific CDF Analysis Farm.

REFERENCES

[1] W. T. Chu, J. Korean Phys. Soc. **50**, 1385 (2007).
[2] I. Foster, C. Kesselman and S. Tuecke, Intern. J. High-Performance Computing Appl. **15**, 200 (2001).
[3] K. Cho, Comp. Phys. Comm. **177**, 247 (2007).
[4] See http://en.wikipedia.org/wiki/E-Science.
[5] K. Cho, J. Korean Phys. Soc. **53**, 1187 (2008).
[6] K. Cho, Internat. J. Comp. Sci. Network Sec. **7**, 49 (2007).
[7] See http://www.gloriad.org/gloriad/index.html.
[8] E. Shuryak, I. Zahed and S.-J. Sin, J. Korean Phys. Soc. **50**, 384 (2007).
[9] B. Kim, in *Proceedings of 2008 LHC Physics Workshop at Korea* (Seoul, Korea, 2008).
[10] I. Sffligoi, Comp. Phys. Comm. **177**, 235 (2007).
[11] M. Neubauer, Nucl. Instr. Meth. A **502**, 386 (2003).
[12] A. Fella, S. C. Hsu, S. Sarkar, D. Jeans, F. Delli Paoli, D. Lucchesi, I. Sfiligoi, S. Belforte, E. Lipeles, M. Neubauer and F. Wüerthwein, in *Proceedings of Conference on Computing on High Energy Physics* (Mumbai, India, 2006).
[13] G. Compostella, D. Lucchesi, S. P. Griso and I. Sfiligoi, in *Proceedings of 3rd IEEE International Conference on e-Science and Grid Computing* (Bangalore, India, 2007).