DISTRIBUTED COMPUTING GRID EXPERIENCES IN CMS DC04

A. Fanfani, P.Capiluppi, C. Grandi *INFN-Bologna, Italy*
I. Legrand, S. Suresh, *Caltech, USA*
S. Campana, F. Donno, W. Jank, N. Sinanis , A Sciabà,*CERN, Switzerland*
P. Garcia-Abia, J. Hernandez, *CIEMAT, Madrid, Spain*
M. Ernst, *DESY, Germany*
A. Anzar, I. Fisk , L. Giacchetti, G. Graham, A. Heavey , J. Kaiser, N. Kuropatine, T. Perelmutov, R. Pordes, N. Ratnikova, J. Weigand, Y. Wu, *FNAL, Batavia, USA*
D.Colling, B. Macevoy, H. Tallini, L.Wakefield*, Imperial College London, United Kingdom*
N. De Filippis, G. Donvito, G. Maggi, *INFN-Bari, Italy*
D. Bonacorsi, L. Dell'Agnello, B. Martelli, *INFN-CNAF, Italy*
M. Biasotto, S. Fantinel, *INFN-Legnaro, Italy*
M. Corvo, F. Fanzago, M. Mazzucato, *INFN-Padova, Italy*
L. Tuura , *Northeastern University, Boston, USA*
J. Letts, T. Martin, *University of California at San Diego, UCSD, USA*
K. Bockjoo, C. Prescott, J. Rodriguez, A. Zahn, *University of Florida, UFL, USA*
D. Bradley, *University of Wisconsin, USA*

Abstract

In March-April 2004 the CMS experiment undertook a Data Challenge (DC04). During the previous 8 months CMS undertook a large simulated event production. The goal of the challenge was to run CMS reconstruction for sustained period at 25Hz input rate, distribute the data to the CMS Tier-1 centers and analyze them at remote sites. Grid environments developed in Europe by the LHC Computing Grid (LCG) and in the US with Grid2003 were utilized to complete the aspects of the challenge. A description of the experiences, successes and lessons learned from both experiences with grid infrastructure is presented.

INTRODUCTION

The Compact Muon Solenoid experiment (CMS) is one of the four particle physics experiments that will collect data at the Large Hadron Collider (LHC). The CMS collaboration will have to process large amounts of events that will be available when the detector will start collecting data. The size of the resources required, the complexity of the software and the physical distribution of the CMS collaboration naturally imply a distributed computing and data access solution. The Grid paradigm is one of the most promising solutions to be investigated, and CMS is collaborating with many Grid projects in order to explore the maturity and availability of middleware implementations and architectures.

The preparation and building of the Computing System, able to treat the data being collected, pass through sequentially planned challenges of increasing complexities [1]. The Data Challenge for CMS during 2004 (DC04) was planned to reach a complexity scale equal to about 25% of that foreseen for LHC initial running. Its goal was to run CMS reconstruction at CERN for a sustained period at 25Hz input rate, distribute the data to the CMS regional centres and analyse them at remote sites. About 50 millions simulated events were required to match the 25 Hz input rate for a month. Actually more than 70 millions events were requested by the CMS physicists. The Pre-Challenge Production (PCP) was the preliminary phase comprising the simulation and the digitization of about 70 millions of events, at the different CMS Regional Centers. It started in July 2003 and is currenlty running the last step of the simulation chain (digitization).

GRID PRE-CHALLENGE PRODUCTION

CMS software for Monte Carlo Production

CMS Monte Carlo production consists of several steps: generation of physical processes (CMKIN), simulation of tracking in the CMS detector based on GEANT3/GEANT4 (CMSIM/OSCAR), reconstruction of CMS detector response and physical information for final analysis (ORCA).

The collection of tools for managing the CMS production system, OCTOPUS, is illustrated in Figure 1. RefDB [2] is a central database located at CERN where all information needed to produce and analyse data are kept. McRunjob [3] is a tool for workflow configuration to automate job creation, following the directive stored in RefDB, and submission to a variety of environments. CMSProd is another tool that provides the same functionality, supporting reading from RefDB and submitting to a local scheduler or LCG scheduler. BOSS [4] is a CMS-developed system that provides information about the job execution status.

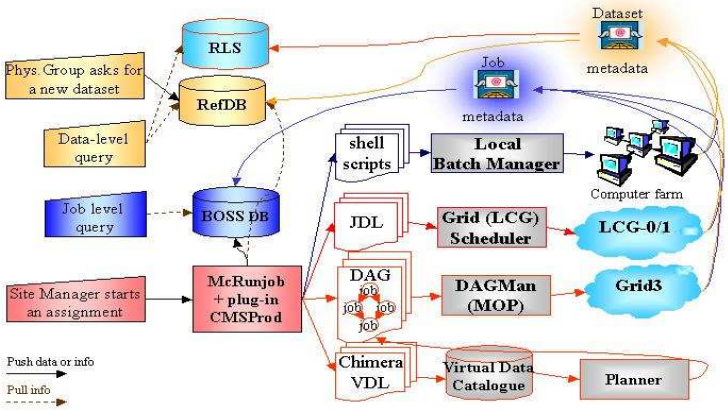

Figure 1 Description of the OCTOPUS system fro CMS Production.

Two grid-based prototypes have been used in Pre-Challenge Production: the *USMOP Regional Center* in the US and the *CMS-LCG Regional Center* mainly in Europe.

USMOP Regional Center

The USMOP Regional Center is using US CMS dedicated resources, about 500-600 CPUs (Caltech, UCSD, Florida, FNAL), and resources of the Grid3 System [5], consisting of about 2000 shared CPU's. Grid3 is a functional Grid deployed by a joint project of USCMS, USAtlas, LIGO, SDSS, BTeV, iVDGL[5] and PPDG[5]. A Pacman [5] package encoded the middleware installation based on the Virtual Data Toolkit (VDT), which provides services from the Globus Toolkit [5], Condor [5], GLUE schema for the MDS Information service, Virtual Organization management tools developed by the European Data Grid (VOMS) and monitoring tools such as Ganglia for cluster monitoring and Server and client software for the MonALISA [6] monitoring framework.

The CMS software includes McRunjob and MOP [7] for production control. The MOP system is a CMS DAG (Directed Acyclic Graph) writer that uses DAGMan and Condor-G for job specification and submission. A Condor-based match-making process selects the resources. All data produced were returned using GridFTP[9] to the Tier1 Storage Element facility at FNAL.

The system produced 8 million CMKIN events, 2.3 millions of CMSIM events and about 17 million OSCAR events, corresponding to about 1000 KSI2000-month and 13TB of data. The OSCAR production jobs are long (average is 30 hours) and not all sites have been able to accommodate running them. Simultaneous usage of CPU resources peaked at 1200 CPUs, controlled by a single FTE. The effort required to run the application has been about 2 FTEs, split between the application administrator and site operations support.

Approximately 70% of CMSIM and OSCAR jobs completed successfully. Jobs often failed due to site configuration problems, or in groups from site service failures. There were few random failures: more frequently a disk would fill up or a service would fail and all jobs submitted to a site would die. Service level monitoring needs to be improved and some services probably need to be replaced. For example, storage reservation (e.g., as provided by SRM) would have prevented various storage-related service failures.

CMS-LCG Regional Center

The CMS-LCG Regional Center used the CMS/LCG-0 system for about 3 months and then part of the official LCG1-testbed for about another month.

The CMS/LCG-0 was a CMS-wide testbed owned by CMS. It has been a joint CMS, DataTAG-WP4 and LCG-EIS effort started in June 2003. It is based on the LCG pilot distribution (VDT1.1.6 and EDG 1.4.X) but includes also components from DataTAG (VOMS for VO management, GLUE schemas and info providers, GridICE for monitoring). It consisted of about 170 CPU's and 4 TB disk distributed over several sites.

McRunJob (or CMSProd) is installed on the User Interface, where the job preparation and job submission took place. The available resources were matched against the Resource Broker, according to the job requirements defined in the Job Description Language (JDL) file. For example a requirement was to run on sites with CMS software installed and validated. The LCG data management tools were used to discover the physical location of an input file and to transfer the produced output data to an LCG Storage Element, registering it in the RLS. BOSS was used for real-time job tracking and job statistics purposes.

The system produced 0.5 million heavy* Pythia events, 1.5 million CMSIM events and 0.6 million OSCAR events (on LCG-1) corresponding to about 10000 jobs, 100 KSI2000-month and 2TB of data.

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 for few

* The duration of these jobs is about 8 hours, compared to the 1.5 minutes of the usual Pythia jobs.

periods. Other sources of instability were site mis-configuration, network problems and hardware failures. The success rate on LCG-1 was lower with respect to CMS/LCG-0 because a consistent site configuration was not always guaranteed and there was less support for services and sites (running over Christmas). Good efficiencies and stable conditions of the system were obtained in comparison with that obtained in previous challenges [8], showing the maturity of the middleware and of the services, provided that a continuous and rapid maintenance is guaranteed by the middleware providers and by the involved site administrators.

DATA CHALLENGE IN LCG

The main aspects of the Data Challenge in 2004 were:

- Reconstruction of data in the Tier-0 farm for sustained period at 25Hz
- Data distribution to Tier-1,Tier-2 sites
- Data analysis at remote sites as data arrive
- Monitor and archive resource and process information

with the aim of demonstrating the feasibility of the full chain.

The reconstruction jobs were submitted to a computer farm at CERN and the produced data were stored on a General Distribution Buffer (GDB) that was also a Castor [9] stage area, so files were automatically archived to tape. The data distribution to the Tier-1 centers was done supporting several data transfer tools: the LCG Replica Manager tools, native SRM (Storage Resource Manager) [9] and SRB (Storage Resource Broker) [9]. The analysis of the reconstructed data in real-time with their arrival was performed in some Tier-1 and Tier-2 sites [9].

The Spanish and Italian Tier-1 and Tier-2 were configured as LCG-2 sites. The full DC04 chain, but the Tier-0 reconstruction, was tested using LCG-2 components. The aspects of the Data Challenge involving LCG-2 components are described in the following.

Global Data Catalogue

The CERN Replica Location Service (RLS) provided the replica catalogue functionality for all the data distribution chains in DC04. The CMS framework uses POOL[10]. The RLS was used both as a file catalogue and as a metadata catalogue to store the POOL file attributes:

- The transfer tools relied on the Local Replica Catalog (LRC) component of the RLS as a global file catalog to store the physical file locations. The Resource Broker queried the LRC to submit analysis jobs close to the data. Inserting PFN was fast enough if the appropriate tools were used (0.1-0.2 sec/file with LRC C++ API programs).
- The Replica Metadata Catalogue (RMC) component of RLS was used as global metadata catalogue, registering the files attributes of the reconstructed data and querying it (by users or agents) to find logical collection of files. The Meta data schema was handled and pushed into RLS catalogue by POOL. Inserting files with their attributes was approximately usable with about 3sec/file in optimal conditions but slow otherwise. Querying information based on metadata was too slow (e.g. several hours to find all the files belonging to a given "dataset" collection).

The total number of files registered in the RLS during DC04 was ~570K Logical File Names each with typically from 5 to 10 Physical File Names and 9 metadata attributes per file. Some performance deficiencies inserting and querying information were identified. Several workarounds were provided to speed up the access to RLS during DC04, however serious performance issues and missing functionality, such as the overhead compared to direct RDBMS catalogues, fast queries and a robust transaction model, need to be addressed. During DC04 there was however no data loss or any extended service downtime.

Data distribution

A data distribution system was developed by CMS for DC04, built on top of available point to-point file transfer tools, to form a directed and scheduled large-scale replica management system [11] The distribution system was based on a structure of software agents collaborating through the Transfer Management Database (TMDB).

The data distribution Tier-0 →Tier-1 → Tier-2 was established using LCG Storage Elements and LCG transfer tools. The schema of the LCG distribution chain is shown in Figure 2. An Export Buffer agent running at CERN copies the files made available on the General Distribution Buffer to the LCG export buffer (classic disk SE), registers its physical location into RLS and updates the file state in TMDB. The Tier-1 transfer agent running at PIC and CNAF looks up in TMDB new files in the LCG export buffer and replicates them to the Castor SE at the Tier-1 using the LCG Replica Manager. The use of the Storage Element interfaced to Castor at Tier-1 was meant to safely store the data on MSS. File replication from the Castor SE at Tier-1 to the disk SE's at Tier-1 and Tier-2s was performed in order to serve data for analysis. The data transfer to both Tier-1s was able to keep up with the rate of data coming from the reconstruction at Tier-0 with good performances. Over 6 TB of data were distributed to PIC and CNAF Tier-1, reaching sustained transfer rates of 30 MB/s. The total network throughput was limited by the small size of the files being pushed through the system. Massive and parallel transfer of typically small files was also affected by the overhead introduced by the Replica Manager java command line, that was in some cases replaced with Globus globus-url-copy to make transfers using gsiFTP, and the Local Replica Catalogue C++ API to update the RLS. Dealing with too many small file increases the load in updating/querying catalogue and highlights both the scalability problem of the MSS and the CMS problem of producing small files. The Castor MSS at PIC was able to cope with it, while at CNAF transfer problems were experienced due to the performance of the underlying Castor MSS with too many small files.

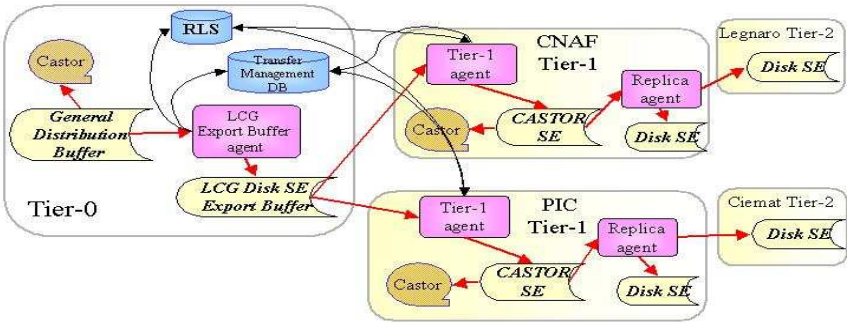

Figure 2 Data distribution system from CERN to LCG Tier-1 sites (CNAF and PIC) and Tier-2s (Legnaro and Ciemat).

Data Analysis

Data were distributed in quasi real time, being available for real time analysis in disk Storage Elements at the Tier-1 and Tier-2 centers. Several agents and automatic procedures were implemented to submit analysis jobs as new data came along [12]. The CMS software required for analysis was installed across LCG-2 sites by the CMS software manager via grid jobs. The LCG-2 Resource Broker was used to submit analysis jobs by selecting the LCG CMS resources hosting the data (Tier-1/2 centers in Italy and Spain). A Resource Broker and an information system (BDII) reserved for CMS were set-up at CERN. CMS could dynamically add or remove resources as needed. The PIC and CNAF RB were also used.

The analysis at INFN was run quasi continuously for two weeks submitting a total of more than 15000 jobs, with a grid efficiency of 90-95%. An average delay of only 20 minutes between the data becoming available from the reconstruction at CERN and the data being analyzed at the PIC Tier-1 was measured. The result obtained at PIC provided the best and most consistent result in this measurement. The LCG submission system coped with the rate of data coming from CERN.

Monitoring

A dedicated GridICE [13] monitoring server was setup in order to monitor the LCG-2 resources registered in the CMS BDII, collecting detailed information about nodes and information on the service machines (Resource Broker, Computing and Storage Element) with the possibility of notification in case of problems. MonaLISA [6] was also deployed at CNAF and PIC Tier-1s.

CONCLUSIONS

During the pre-Challenge phase CMS demonstrated that distributed productions based on grid middleware are possible. The prototypes were based on early deployed systems of LCG and on Grid2003 in the US.

Grid2003 was shown to be a reliable and scalable system for massive production, reaching a new magnitude in the number of autonomously cooperating computing sites for production, with peaks of 1200 CPUs simultaneous usage.

Large scale productions were performed in LCG exploiting high-level Grid components with good efficiency. The major concerns were the RLS being a single point of failure and the consistency of the distributed sites configuration and control.

In the 2004 Data Challenge the LCG environment provided the functionalities for distributed computing: global file and metadata catalogues, grid point-to-point file transfer tools and infrastructure for data analysis. The major issues were related to the performance of the data catalogues. The LCG data distribution and analysis chain successfully met the data challenge goals of large scale scheduled distribution to a set of Tier-1/2 and subsequent analysis.

REFERENCES

[1] C.Grandi, "CMS Distributed Data Analysis Challenges", IX Intl. Workshop ACAT, Japan 2003

[2] V. Lefebure, J. Andreeva "RefDB: The Reference Database for CMS Monte Carlo Production", CHEP03, La Jolla, California 2003

[3] G. Graham et al. "McRunjob: A High Energy Physics Workflow Planner for Grid Production Processing", CHEP03, La Jolla, California, 2003

[4] C. Grandi, A. Renzi "Object Based System for Batch Job Submission and Monitoring (BOSS)", CMS NOTE-2003/005

[5] I.Foster et al. "The Grid2003 Production Grid: Principles and Practice", 13th IEEE Intl. Symposium on High Performance Distributed Computing 2004 and references therein.

[6] I.C. Legrand et al, "MonALISA: a distribute monitoring service architecture", CHEP03, La Jolla, California 2003

[7] The MOP Project, http://www.uscms.org/s&c/MOP

[8] "CMS Test of the European DataGrid Testbed"; CMS NOTE-2003/014

[9] D. Bonacorsi et al. "Role of Tier-0, Tier-1 and Tier-2 Regional Centres during CMS DC04", CHEP04, Interlaken 2004, and references therein.

[10] POOL Project: http://lcgapp.cern.ch/project/persist

[11] T.Barrass et al. "Software agents in data and workflow management", CHEP04, Interlaken 2004

[12] N. De Filippis et al "Tier-1 and Tier-2 Real-Time analysis experience in CMS DC04", CHEP04, Interlaken 2004

[13] S.Andreozzi et al "GridICE: a monitoring service for the Grid", 3rd Cracow Grid Workshop, Oct 2003