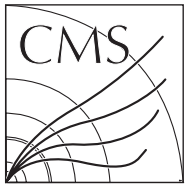
Available on CMS information server

CMS CR -2013/359



The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

28 October 2013 (v2)

CMS Computing Operations During Run1

Oliver Gutsche for the CMS Collaboration

Abstract

During the first run, CMS collected and processed more than 10B data events and simulated more than 15B events. Up to 100k processor cores were used simultaneously and 100PB of storage was managed. Each month petabytes of data were moved and hundreds of users accessed data samples. In this presentation we will discuss the operational experience from the first run. We will present the workflows and data flows that were executed, we will discuss the tools and services developed, and the operations and shift models used to sustain the system. Many techniques were followed from the original computing planning, but some were reactions to difficulties and opportunities. In this presentation we will also address the lessons learned from an operational perspective, and how this is shaping our thoughts for 2015.

Presented at *CHEP2013 Computing in High Energy Physics 2013*

CMS computing operations during run 1

J Adelman[6],S Alderweireidt[19],J Artieda[6],G Bagliesi[15],D Ballesteros[6],S Bansal[19],L Bauerdick[6],W Behrenhof[7],S Belforte[10],K Bloom[13],B Blumenfeld[24], S Blyweert[22],D Bonacorsi[3],C Brew[16],L Contreras[6],A Cristofori[3],S Cury[5],D da Silva Gomes[6],M Dolores Saiz Santos[5],J Dost[18],D Dykstra[6],E Fajardo Hernandez[5],F Fazango[23],I Fisk[6],J Flix[14],A Georges[18],M Giffels[5],G Gomez-Ceballos[12],S Gowdy[5],O Gutsche[6],B Holzman[6],X Janssen[19],R Kaselis[5],D Kcira[4],B Kim[8],D Klein[18],M Klute[12],T Kress[1],P Kreuzer[1],A Lahiff[16],K Larson[6],J Letts[18], A Levin[12],J Linacre[6],L Linares[5],S Liu[6],S Luychx[19],M Maes[22],N Magini[5],A Malta[5],J Marra Da Silva[17],J Mccartin[21],A McCrea[18],A Mohapatra[20],J Molina[5],T Mortensen[18],S Padhi[18],C Paus[12],S Piperov[5],D Ralph[12],A Sartirana[9],A Sciaba[5],I Sfiligoi[18],V Spinoso[2],M Tadel[18],S Traldi[3],C Wissing[7],F Wuerthwein[18],M Yang[12],M Zielinski[6],M Zvada[11]

[1]RWTH Aachen Univ., Aachen, Germany
[2]INFN Univ. di Bari, Italy
[3]INFN Univ. di Bologna, Italy
[4]California Institute of Technology, USA
[5]European Organization for Nuclear Research (CERN), Geneva, Switzerland
[6]Fermi National Accelerator Laboratory, Batavia, IL, USA
[7]Deutsches Elektronen Synchrotron (DESY), Hambug, Germany
[8]Univ. of Florida, USA
[9]LLR, Ecole Polytechnique, France
[10]INFN Univ. di Trieste, Italy
[11]Inst. fr Exp. Kernphysik Karlsruhe, Germany
[12]Massachusetts Institute of Technology, USA
[13]Univ. of Nebraska-Lincoln, USA
[14]Port d'Informació Científica (PIC), Universitat Autònoma de Barcelona, Bellaterra (Barcelona) and Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, CIEMAT, Madrid, Spain
[15]INFN Univ. di Pisa, Italy
[16]Rutherford Appleton Laboratory, UK
[17]Univ. Estadual Paulista (UNESP),So Paulo, Brazil
[18]UC San Diego, California, USA
[19]Univ. Antwerpen, Belgium
[20]Univ. of Wisconsin, USA
[21]Ghent Univ., Belgium
[22]Vrije Univ. Brussel, Belgium
[23]INFN Padova, Italy
[24]Johns Hopkins University, USA

E-mail: gutsche@fnal.gov

Abstract. During the first run, CMS collected and processed more than 10B data events and simulated more than 15B events. Up to 100k processor cores were used simultaneously and

100PB of storage was managed. Each month petabytes of data were moved and hundreds of users accessed data samples. In this document we discuss the operational experience from this first run. We present the workflows and data flows that were executed, and we discuss the tools and services developed, and the operations and shift models used to sustain the system. Many techniques were followed from the original computing planning, but some were reactions to difficulties and opportunities. We also address the lessons learned from an operational perspective, and how this is shaping our thoughts for 2015.

1. Introduction

The first data taking run of the Large Hadron Collider (LHC) [1] at CERN in Geneva, Switzerland, started in Fall 2010 and ended in Spring 2013. The preparation for *LHC run 1* for the machine, the detectors of the major four experiments, the trigger and data taking chain, as well as the computing, was intense and culminated in a very successful physics harvest with hundreds of publications crowned by the discovery of the Higgs boson and by the Nobel prize in physics in 2013 [2] awarded for the theoretical prediction of the Higgs boson. The Compact Muon Solenoid (CMS) [3] experiment records collisions of the LHC. From the beginning, CMS was preparing to use a distributed computing approach departing from the previously favored host-laboratory centric setup. With this approach came many challenges to operate the computing infrastructure. The goal in the end was to enable all CMS physicists to perform their analyses and publish papers, which the collaboration succeeded in marvelously. This paper presents the operational experience of the CMS computing infrastructure during LHC run 1. After an introduction to the infrastructure and used systems, we go in detail through important aspects of the operational approach and point out lessons learned and plans for *LHC run 2* which will start in 2015.

2. CMS computing infrastructure

The CMS computing infrastructure is based on a tiered set of distributed computing sites connected through networks based on the MONARC model [4]. The computing infrastructure is organized in a strict hierarchy starting with the host laboratory CERN at its origin called *Tier-0*. On the next level, 7 regional computing centers called *Tier 1* sites (Tier-1) form the backbone of the system, followed by *Tier 2* sites (Tier-2) at universities and/or research institutes. The next level (*Tier-3*) is not discussed further, but also represents an important part of the infrastructure volunteered by CMS collaborators. All sites are interconnected via dedicated or general purpose scientific networks. The original design separated the Tier-2 level into groups of sites associated and connected to only one Tier-1. During LHC run 1, the strength and reliability of the networks allowed for a more flexible setup, realizing a full-mesh network topology where every Tier-2 is connected to every Tier-1 and every other Tier-2 (see figure 1).

Tier-1s provide access to large CPU farms and a mass storage system (MSS) with disk caches and tape libraries. They are different from Tier-2s in two main aspects. Tier-1s provide 24/7 support operation, while Tier-2s are operated during business hours and unattended in between. In addition, Tier-2s do not have tape libraries at the sites and operate purely with disk caches. The network between the sites is the backbone of the CMS computing infrastructure as it is based on transferring files between the sites for access. The network between the Tier-0 and the Tier-1s is dedicated and called the *LHC Optical Private Network* (LHCOPN) [5]. The Tier-2s are connected through general purpose scientific networks (GPNs) in their host countries. The main data flows can be separated into *archiving* and *serving*. Archiving is tape based where related files are grouped by physics content on separate sets of tape cartridges (known as tape families) to optimize writing, recall, and recycling. Serving is disk based. The main data flows are as follows:

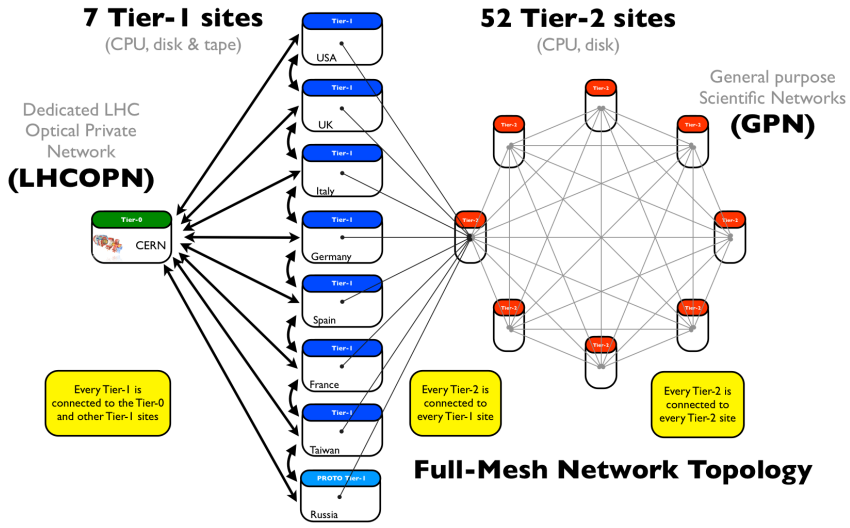


Figure 1. The tiered CMS computing infrastructure showing the Tier-0, Tier-1 and Tier-2 levels interconnected with dedicated or general purpose scientific networks in a full-mesh network topology.

T0→T1: *RAW* data recorded by the detector is recorded at the Tier-0 and stored on tape as a "cold" backup replica. The RAW is distributed across the Tier-1s via the LHCOPN network links and archived on tape. The tape replica at the Tier-1s can be recalled at any time and is called the *custodial* replica. The mass storage system at the Tier-1s also provides caching capabilities and the files from the current year of data taking are kept cached.

T2→T1: The process of Monte Carlo (MC) event generation and simulation is CPU intensive and performed mainly at the Tier-2s. The output of the simulation of a specific physics process is consolidated at a Tier-1 site where it is stored on tape.

T1→T1: Both RAW and simulated data are reconstructed for analysis on the Tier-1 level. This requires especially strong I/O capabilities. The resulting *Analysis Object Data* (AOD) holds the minimum amount of information needed for analysis and is archived at the Tier-1 that stored the custodial tape copy of the input data.

The main workflows can be separated into *production* workflows and *analysis* workflows and are characterized by the level on which they are executed:

Tier-0: Data recording Collisions are recorded by the detector and selected by the trigger system [6] and stored in binary format in files. They are transferred from the detector to the CERN computing center for further processing:

- 10% of the selected collisions are *express* processed, where the repacking step to transform the binary format into the CMS ROOT-based event format [7] is run within one hour of the recording of the collisions. It is followed by a reconstruction step to provide input to workflows to calculate alignment and calibration constants and for data quality monitoring.
- 100% of the selected collisions are repacked into the CMS event format, and split into primary datasets according to physics content by trigger path groups (introducing an overlap of 10-25% between primary datasets). The reconstruction of all of the collisions

is started 48 hours after their recording in the detector, in order to use the alignment and calibration constants calculated using the express processed collisions.

Tier-1/2: Simulation Collisions are generated using theory software packages and their detector response is simulated using GEANT4 [8]. This workflow is CPU dominated and needs little or no input.

Tier1: Reconstruction Collisions both from the detector and simulations are reconstructed on the Tier-1 level, in the case of detector data using updated alignment and calibration constants and/or software versions compared to the Tier-0 reconstruction. The workflows are executed at the sites where the input is stored custodially on tape and requires pre-staging of the input to disk for efficient access. For simulated collisions, additional proton-proton interactions in the same or nearby bunch crossings (pileup) are taken from a simulated pileup dataset, and superimposed on the hard collision. The average number of pileup interactions can reach 140 in later stages of LHC data taking and was at average 30 in 2012, making this the most I/O intensive workflow of CMS.

Tier-2: Analysis Analysis is performed by accessing both recorded and simulated collisions on the Tier-2 level. Physicists use a mixture of official and self-written code to extract physics measurements. In contrast to the production workflows, analysis involves multi-user access to files which have been pre-distributed to Tier-2s. Every analysis job is preferentially processed at the location of its input, and the user-generated output is stored outside the CMS file catalogues.

CMS operates the following services based on GRID technologies from the *EGI* [9], *ARC* [10] and *OSG* [11] GRID middle-wares under the umbrella of the *Worldwide LHC GRID* (WLCG) infrastructure [12] to support the execution of the CMS workflows:

Transfer system: Files are organized in datasets with similar physics content. The CMS transfer system is called *PhEDEx* [13] and replicates datasets to destination CMS sites on request using GRID transfer protocols. It is based on well established and constantly monitored site-to-site PhEDEx links, using the ability of PhEDEx to mark enabled and tested network paths.

Bookkeeping system: Metadata of files and datasets is catalogued in the CMS *Dataset Bookkeeping System* (DBS) [13].

Constants system: Alignment and calibration constants can be accessed by every executable on every site through the *Frontier* [14] system based on a hierarchy of *SQUID* caches [15].

Software distribution: CMS software is available as pre-compiled releases and installed at the sites through GRID jobs [16] or accessed through a shared GRID based file system called *CVMFS* [17], based on SQUID caches.

Submission infrastructure: Different submission infrastructures are used by CMS to submit jobs to all sites to execute production workflows and analysis. The infrastructures used are the *gLite*[18], *HTCondor_G*[19], and *glidein* Workload Management Systems (WMS)[20].

Site availability monitoring: The availability of the GRID services at the sites is probed periodically using the WLCG's *Site Availability Monitoring* (SAM) [21] system.

Site stress tests: All sites are tested periodically with complete analysis workflows submitted by the automated *HammerCloud* [22] infrastructure.

Monitoring: All jobs are instrumented to report status information to the WLCG *DashBoard* [23] which is combined with site downtime information from the WLCG infrastructure as well as site readiness information compiled from the SAM and HammerCloud tests [24], taking into account valid PhEDEx links.

Production system: All production workflows are executed through a state based production system called *WMAgent* [25].

Analysis system: The CMS analysis system helps users to find the location of the input to their analyses and splits the workflows into individual jobs which are then submitted through the submission infrastructure. Unlike the production system, the analysis system, called the *CMS Remote Analysis Builder* (CRAB) [26], allows the user to move user-specific code to the GRID worker nodes of the jobs.

3. CMS computing operations

CMS supports the infrastructure and the sites, both through central teams of operators and experts and through site contacts who bridge CMS to site administrators. Team members are both physicists and engineers. Both physicists and institutes get awarded service credit by CMS to fulfill author list requirements. In 2012, CMS awarded 40 FTEs of service credit to members of the central operation teams and 60 FTEs of service credit to sites and their contacts. The central teams are divided into *computing operations* and *physics support*. Central services, site support, and the execution of production workflows are handled by computing operations with experts managing parts of the infrastructure and operation teams for the different components: Tier-0 operation, central workflows execution and production infrastructure operation, site support, submission infrastructure operation, and monitoring. Physics support handles the support of analysis users and their interaction with the analysis system. Because of the large user base of CMS with up to 500 concurrently active users at a given time, a significant amount of effort is spent in supporting individual users and small groups in succeeding in their analysis efforts.

The CMS collaboration requires its collaborators for authorship to contribute to central shift operation of the detector and data taking chain in addition to the general service requirements. Computing is a vital part of the data taking chain and established computing shifts to monitor all running systems. Throughout the year, non-expert collaborators man three shifts covering 24 hours each day, following the three main time zones of the collaboration: Europe, Asia and the Americas. A shifter is called a *Computing Shift Person* (CSP) and follows shift instructions to monitor critical aspects of the computing infrastructure. In case of deviation from normal operation parameters, the shifter alarms experts or sites via tickets or direct communication through electronic logs, email, phone, or instant messaging. Most of the tickets opened to sites originate from CSP shifter observations. A computing expert oversees the CSP shifts, triaging problems found by the shifters and acting as a contact between computing and the detector operation and data taking coordination teams. The expert shifter is called the *Computing Run Coordinator* (CRC). A CRC shift lasts one week, and the CRC must be located physically at CERN during data taking. The CRC is in permanent contact with the different CSP shifters and is able to handle interventions on critical systems and services outside business hours when the specialists are off duty. Both CSP and CRC shifts proved vital to the operation of the computing infrastructure during LHC run 1. One lesson learned is that the system has to be monitored constantly and its operation has to be compared to nominal operations parameters. Most CSP checks could be automated. This has not happened yet due to time and manpower constraints. CMS plans to pursue the automation of the CSP monitoring checks during LHC run 2. This process is expected to require a significant amount of time and effort. The tasks of the CRC role will not be automated, because it is important that trained experts are always available to handle critical situations.

CMS computing operates many different systems as described in Sec. 2, many of which rely on web interfaces like the transfer system or the metadata service. Stable operation of these systems is one of the cornerstones of successful operation of the infrastructure. But stable oper-

ation is in contradiction to rapid development cycles and roll-out of new features caused by the evolution of data taking, especially crucial during the early phases of data taking. To enable the services to evolve while providing stable services, CMS established a specialized deployment and testing procedure for its web cluster providing access to the services called *cmsweb*. All services under the cmsweb platform are installed on a common load-balanced web platform. Not only web interfaces to services are provided, but also specialized functionalities like NoSQL databases. All services go through a structured deployment procedure. Every month, a list of changes to be deployed is compiled and packages in RPM format of the services are cut and installed on a testbed instance of the cmsweb cluster. After a two week period of testing and fixing any bugs, the changes are rolled-out in production. This regulated and regular deployment procedure, including the structured deployment platform, kept the system stable while rolling out new developments and bug fixes successfully. Although at times painful, it was necessary to formalize the procedure to introduce predictability for upgrades and updates and increase the chance for problem-free production roll-outs. This was achieved with great success. In LHC run 1, not all services were integrated into the cmsweb platform and using the cmsweb deployment procedures. CMS's goal for LHC run 2 is to transition all services (if possible) to this common testing and deployment scheme.

The Tier-0 operation has a special role in the overall operations effort. In case of problems with the Tier-0 operation, data taking can be impacted up to a complete stop of detector operation if the buffer space at CERN fills up. If the express processing was to stop, the data quality monitoring could not be guaranteed to the fullest extent, which in turn could mean reduced data quality because of the lack of feedback to detector operation. To guarantee the flawless operation of the Tier-0, CMS designated two full-time operators who are located on either side of the Atlantic to increase the business day coverage. This turned out to be crucial especially during the startup phase of data taking. The Tier-0 operators were available to handle crisis situations reliably during most of the day, augmented by the CRC. Apart from covering both sides of the Atlantic, one additional important lesson learned was the involvement of the Tier-0 operators in the development of the Tier-0 processing infrastructure. It quickly became apparent that only operating the infrastructure did not sufficiently prepare the operators to uphold optimal operation efficiencies. They needed to become proficient in the infrastructure itself, with expertise close to that of a full-time developer. Without this intimate knowledge of the system, the operators would not have been able to react to the rapidly changing data taking conditions. It was for example crucial to alarm detector operation of trigger rates that would endanger the subsequent processing of the data, possibly rendering the recorded data useless. Only experts are able to spot such issues and reliably report back.

An overview of the different workflow types that the central processing infrastructure needs to support are given in Sec. 2. They come in many versions which makes the operation of the production infrastructure a manpower intensive endeavor. Because of the reliance on many different central systems and on the readiness of the sites, it needs a team of trained expert operators to understand the many failure codes returned by the jobs and to solve all issues quickly to keep the latency low and tails short. Especially the last 5% of the completion of a workflow is very time and manpower intensive. Ever-changing requirements and modifications in the workflows themselves need to be incorporated and executed and sometimes are cause of latencies as well. This learned lesson is not expected to be solved completely for LHC run 2. But CMS expects to reduce the manpower needs for the operation of the production infrastructure further because workflows are expected to be more stable and experience gained in LHC run 1 will help.

One of the major contributors to the completion latency of central workflows was identified

in LHC run 1 to be the Tier-1 storage model. The model foresees that workflows will run at the site which holds the custodial tape replica of the input. This is not a big issue if many different workflows for inputs stored at different Tier-1s need to be processed, because the distribution of inputs across the Tier-1s would even out the CPU usage. During LHC run 1, on many occasions this was not the case. For example the reconstruction of only a few primary datasets was necessary to bring high priority analyses to a conclusion quickly. These patterns were not foreseeable when the datasets were placed at the sites. This situation could lead to some of the Tier-1s being busy while others were idle, therefore introducing an inefficiency in CPU usage on the Tier-1 level. To remove this source of inefficiency, CMS is asking all Tier-1s to reconfigure their MSS setup to separate disk and tape into a large disk-only pool used for serving data to the CPU farm and storing the output coming from the workflows, and a small disk pool connected to the tape system to archive files. The transfer system will be used to archive files on tape and to recall them to the disk-only pool if processing is needed. Due to the disk/tape separation, input files can be recalled to multiple sites and output can be consolidated to single tape archive instances, therefore removing the inefficiency observed in LHC run 1. This setup also enables the publication of all files on disk at the Tier-1 sites to the CMS remote access federation (AAA, for *Any data, Any time, Anywhere*) [27], something which was not possible before in most cases due to the strict coupling of disk and tape. This is expected to increase the flexibility of the CMS processing infrastructure further. In the end, this setup will also allow analysis workflows on the Tier-1s which has not yet been possible due to the possibility that random access to all samples stored on tape could overwhelm the tape systems.

Physics support is concerned with the other side of CMS computing operations: the support for the analysis of the data and MC samples by the CMS physicists. Unlike the central production workflows, the analysis workflows come in a multitude of different forms and flavors with diverse requirements on resources on the worker nodes as well as non-standard code that analysts wrote themselves. As users are in control of the definition and validation of their analysis workflows, this is a much more challenging multi-user problem than the production case, where workflows are run using the same production role. CMS has a dedicated support team of experts to help users in analyzing the data and in supporting the analysis GRID tools. Inefficiencies observed during LHC run 1 are connected to the way CMS distributes samples to the Tier-2s and how jobs are only sent to Tier-2s that provide local access to them.

In the current model, the available Tier-2 disk space is logically separated into a portion managed by the transfer system holding datasets and a second portion for users to store their analysis outputs outside the metadata booking system and the transfer system. The portion managed by the transfer system is further subdivided into quotas for centrally placed samples like commonly used MC samples (typically simulations of common standard model physics processes) and data AOD samples, and quotas for the individual physics groups. This setup requires that the different quotas are managed manually and decisions are made which samples to keep in the disk space allocations. This leads to inefficiencies manifested in samples stored on Tier-2 disk that are accessed only rarely or not at all. CMS is planning to simplify the setup of the managed disk space by combining the central and physics group quotas and having all the managed disk handled by an automated system. The popularity service [28] tracking the access patterns of the samples by the different analysis and production jobs is planned to be used as the basis to make automated decisions about which samples need to be replicated on the Tier-2s and which of these replica caches need to be released. In addition, the queue depth of the submission infrastructure can be used to make additional replicas of popular samples.

The submission infrastructure itself underwent an evolution during LHC run 1, from a mixed setup using a direct submission mode to a setup based on the pilot-based submission infrastructure glideinWMS [29]. The mixed setup used prioritization on the level of the *Compute*

Element (CE, grid door into worker node farm) at the sites to prioritize between production and analysis workflows (CMS initially required the Tier-1s to run 100% production jobs and the Tier-2s 50% production and 50% analysis jobs) which introduced inefficiencies in prioritizing between the different activities and users, and also presented a very rigid setup to change the prioritization for individual users or groups as happened frequently during LHC run 1. Now, CMS has moved completely to a pilot-based submission infrastructure that is able to do all of the prioritization in the WMS itself without the need for prioritization on site level, a much more flexible and efficient model. Especially attractive is the possibility to redirect jobs according to queue depth and input requirements to other sites with free CPU capacity and read the input through the CMS remote access federation. If the demand for a sample is sufficiently high, an additional replica can be made automatically which increases the efficiency even further.

The failure rates of analysis jobs are currently dominated by stage-out problems. Analysis jobs stage out their output directly from the worker node to remote storage elements at the Tier-2s. If the stage-out fails, the whole analysis job fails and needs to be repeated. CMS is planning to move to an asynchronous stage-out implementation. Analysis jobs first store their output temporarily at the site where the job is running. A subsequent step in the analysis workflow is then moving the output to the destination storage element using transfer system techniques. This will reduce the failure rates due to stage-out problems significantly and increase the efficiency of the CPU usage by analysis.

The planned improvements all rely on the sites working reliably over long periods of time. A problem at a site can have significant implications, even more when CMS moves to more dynamic and automated systems. As an example, the transfer system is functioning very well and only a few last-percent problems need to be solved actively by transfer support. The main problems are site and infrastructure related (slow network links, instabilities in storage elements, etc.). CMS plans to use PerfSonar [30] to monitor and solve network link inefficiencies and issues.

The site issues will be a major focus of computing operations in the next year till the start of LHC run 2 and beyond. CMS has a dedicated site support operations team working with site administrators to solve problems and improve long-term stability. SAM and HammerCloud tests are the basis to assess reliability and are crucial for the operation of all CMS sites. CMS has a large group of very reliable sites with only a few sites that are problematic. To be able to exploit all improvements, all sites need to reach sufficient readiness levels. CMS is intensifying the site support effort to work with the few less reliable sites to enable them to become fully reliable again. Site reliability will be an important factor for the success of CMS computing during LHC run 2.

4. Summary & Outlook

LHC run 1 provided a lot of operational challenges, sometimes more manpower-intensive than expected. Contributing to the intensity and complexity of the challenges were the rapidly changing conditions and requirements of data taking, central workflows, and analysis. These things were not unexpected and are expected to continue during LHC run 2 to some extent.

During LHC run 1, CMS gained valuable operational experiences and identified areas of improvement that are currently being worked on actively:

- Disk/tape separation at the Tier-1s
- Dynamic data placement and automatic cache release for samples on Tier-2s
- Full deployment of the global CMS remote access federation (AAA)
- Global glideinWMS pool and prioritization

With these improvements, CMS expects to be able to meet the challenges of LHC run 2.

5. Acknowledgements

We would like to thank the funding agencies supporting the CMS experiment and the LHC computing efforts.

References

[1] Evans L and Bryant P 2008 "LHC Machine" *JINST* **3** (2008), no. 08, S08001
[2] "The Nobel Prize in Physics 2013". Nobelprize.org. Nobel Media AB 2013. Web. 27 Oct 2013. http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/
[3] CMS Collaboration 2008 "The CMS experiment at the CERN LHC", *JINST* **3** (2008), no. 08, S08004
[4] Aderholz M et al. 2000 "Models of Networked Analysis at Regional Centres for LHC xperiments (MONARC) - Phase 2 Report" CERN/LCB 2000-001 (2000)
[5] http://lhcopn.web.cern.ch/lhcopn/
[6] CMS Collaboration 2000 "The TRIDAS Project Technical Design Report, Volume 1: The Trigger Systems" CERN/LHCC 2000-38, CMS TDR 6.1
[7] Brun R and Rademakers F 1996 "ROOT - An Object Oriented Data Analysis Framework", Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. And Meth. in Phys. Res. A 389 (1997) 81-86
[8] Agostinelli S et al. 2003 "Geant4a simulation toolkit" *NIM A* **3** (2003). no. 3, S0168, http://dx.doi.org/10.1016/S0168-9002(03)01368-8
[9] http://www.egi.eu/
[10] Ellert M et al. 2007 "Advanced Resource Connector middleware for lightweight computational Grids" Future Generation Computer Systems 23 (2007) 219-240
[11] Pordes R et al. 2007 "The Open Science Grid" *J. Phys. Conf. Ser.* **78**, 012057 http://dx.doi.org/10.1088/1742-6596/78/1/012057
[12] Bird I et al. 2005 "LHC computing Grid. Technical design report" CERN-LHCC-2005-024 (2005)
[13] Giffels M, Guo Y, Kuznetsoz V, Magini N and Wildish T 2013, *The CMS Data Management System*, submitted to CHEP 2013
[14] Dykstra D et al. 2007 "CMS conditions data access using FroNTier", Proceedings of the Conference on Computing in High Energy and Nuclear Physics (CHEP07), Victoria, Canada (2007)
[15] http://www.squid-cache.org/
[16] Behrenhoff W et al., 2011 "Deployment of the CMS software on the WLCG grid", J.Phys.Conf.Ser. 331 (2011) 072041
[17] Buncic P et al 2010 "CernVM a virtual software appliance for LHC applications" J. Phys.: Conf. Ser. 219 042003
[18] Cecchi M et al 2010 "The gLite Workload Management System" J. Phys.: Conf. Ser. 219 062039 http://dx.doi.org/10.1088/1742-6596/219/6/062039
[19] Thain D, Tannenbaum T and Livny M 2005 "Distributed Computing in Practice: The Condor Experience" Concurrency and Computation: Practice and Experience, Vol. 17, No. 2-4, pages 323-356, February-April, 2005
[20] Sfiligoi I, Bradley D C, Holzman B, Mhashilkar P, Padhi S and Wrthwein F 2009 "The pilot way to grid resources using glideinWMS" Comp. Sci. and Info. Eng., 2009 WRI World Cong. on 2 428-432 http://dx.doi.org/10.1109/CSIE.2009.950
[21] https://twiki.cern.ch/twiki/bin/view/LCG/SAMOverview
[22] van der Ster D et al. "HammerCloud: A Stress Testing System for Distributed Analysis" 2011 J. Phys.: Conf. Ser. 331 072036 http://dx.doi.org/10.1088/1742-6596/331/7/072036
[23] Andreeva J et al. 2008 "Dashboard for the LHC experiments" J. Phys.: Conf. Series 119 062008
[24] Flix J et al. 2012 "Towards higher reliability of CMS computing facilities" J. Phys. Conf. Ser. 396, 032041, 2012
[25] Fajardo E et al 2012 "A new era for central processing and production in CMS" J. Phys.: Conf. Ser. 396 042018 http://dx.doi.org/10.1088/1742-6596/396/4/042018
[26] Lacaprara S et al. 2006 "CRAB:a tool to enable CMS Distributed Analysis", Proceedings of the Conference on Computing in High Energy and Nuclear Physics (CHEP06), Mumbai, India, 13 Feb - 17 Feb 2006
[27] Bloom K et al. 2013 "CMS Use of a Data Federation", Proceedings of the Conference on Computing in High Energy and Nuclear Physics (CHEP13), Amsterdam, The Netherlands, 14 Oct - 18 Oct 2013
[28] Giodarno et al. 2012 "Implementing data placement strategies for the CMS experiment based on a popularity model" J. Phys.: Conf. Ser. 396 032047 http://dx.doi.org/10.1088/1742-6596/396/3/032047
[29] Gutsche O et al. 2013 "Evolution of the pilot infrastructure of CMS: towards a single glideinWMS pool", Proceedings of the Conference on Computing in High Energy and Nuclear Physics (CHEP13), Amsterdam, The Netherlands, 14 Oct - 18 Oct 2013

[30] Campana S et al. 2013 “Deployment of a WLCG network monitoring infrastructure based on the perfSONAR-PS technology”, Proceedings of the Conference on Computing in High Energy and Nuclear Physics (CHEP13), Amsterdam, The Netherlands, 14 Oct - 18 Oct 2013