PROCEEDINGS OF SCIENCE

The commissioning of CMS computing centres in the WLCG Grid

S. Belforte[a], A. Fanfani[b], I. Fisk[c], J. Flix Molina[d], J.M. Hernández[d], J. Klem[e], J. Letts[f], N. Magini[gh], V. Miccio[gh], S. Padhi[f], P. Saiz[g], A. Sciabà*[g] and F. Würthwein[f]

[a] *INFN Sezione di Trieste, Trieste, Italy*
[b] *Università di Bologna, Bologna, Italy*
[c] *Fermi National Accelerator Laboratory, Batavia, IL, USA*
[d] *CIEMAT, Madrid, Spain*
[e] *Helsinki Institute of Physics, Helsinki, Finland*
[f] *University of California at San Diego, La Jolla, CA, USA*
[g] *CERN, Geneva, Switzerland*
[h] *INFN – CNAF, Bologna, Italy*
E-mail: Stefano.Belforte@ts.infn.it, Alessandra.Fanfani@bo.infn.it, ifisk@fnal.gov, jflix@pic.es, Jose.Hernandez@ciemat.es, Jukka.Klem@cern.ch, Nicolo.Magini@cern.ch, Vincenzo.Miccio@cern.ch, Sanjay.Padhi@cern.ch, Pablo.Saiz@cern.ch, Andrea.Sciaba@cern.ch, fkw@ucsd.edu

The computing system of the CMS experiment works using distributed resources from more than 60 computing centres worldwide. These centres, located in Europe, America and Asia are interconnected by the Worldwide LHC Computing Grid. The operation of the system requires a stable and reliable behaviour of the underlying infrastructure. CMS has established a procedure to extensively test all relevant aspects of a Grid site, such as the ability to efficiently use their network to transfer data, the functionality of all the site services relevant for CMS and the capability to sustain the various CMS computing workflows (Monte Carlo simulation, event reprocessing and skimming, data analysis) at the required scale. This contribution describes in detail the procedure to rate CMS sites depending on their performance, including the complete automation of the program, the description of monitoring tools, and its impact in improving the overall reliability of the Grid from the point of view of the CMS computing system.

XII Advanced Computing and Analysis Techniques in Physics Research
November 3-7, 2008
Erice, Italy

*Speaker.

PoS(ACAT08)043

1. Introduction

The Large Hadron Collider (LHC), located at CERN, will be operational in 2009 and will produce p-p collisions at centre-of-mass energy of 14 TeV, with a luminosity eventually two orders of magnitude larger than current hadron colliders. The Compact Muon Solenoid (CMS) is one of the four detectors that will observe the collisions and it is foreseen to collect about 5 PB of data each year. To process these data requires to exploit computing and storage resources from several centres outside CERN. The resources are in fact provided by the Worldwide LHC Computing Grid (WLCG), which at most sites exploits the computing infrastructure of other Grid projects, like EGEE, Open Science Grid and NorduGrid.

In the case of the CMS collaboration, the sites involved are around 60 from about 20 countries all around the world. They are organized with a tiered structure, where different tier levels correspond to different functions. The Tier-0 site is CERN, and takes care of the prompt event reconstruction and detector calibration, the distribution of raw and processed data to external sites and the backup storage of the raw data. The seven Tier-1 sites run the subsequent reprocessing, including data skimming, keep an active copy of the raw data and store the Monte Carlo generated at Tier-2 sites. Finally, Tier-2 sites get samples of the skimmed data for analysis and are used to run the Monte Carlo simulation. A complete description of the CMS computing model and its services can be found elsewhere [1].

Given the complexity of the infrastructure, it is important to be able to measure its performance in a continuous way, in order to inform the sites of any problem CMS is encountering, or will encounter, when running at that site. The Grid projects operating the infrastructure have their own procedures to identify and correct problems, but these do not necessarily cover problems more specific to CMS. For this reason, CMS has established a set of techniques and tools intended to provide a better picture of the site performance and reliability.

The following sections describe in detail how this is performed: firstly, the procedure to test sites in an automatic and continuous way is described; secondly, the results of the test activities performed during the 2008 Common Computing Readiness Challenge (CCRC08) [2] at the Tier-2 sites for the data analysis are reported.

2. Site evaluation techniques

The Site Commissioning activity is part of the CMS computing integration program and its mandate is to evaluate the readiness of every CMS site to execute the computing tasks assigned to it. This information can be used by the sites to become aware of problems, and by CMS to plan accordingly the distribution of the workload such to temporarily avoid sites experiencing problems, by automatically producing a list of "good" sites to which production managers can restrict job submission. In order to accomplish that, custom tests are regularly run at each site, and these tests are conceived to test every possible functionality exploited by CMS, aiming at having the highest possible correlation between failures of these tests and of real CMS jobs. Sites must satisfy certain lower limits on the success rate of these tests to be considered reliable.

The following information is used to evaluate the site readiness:

- the fraction of time during a day when all relevant functional tests were successful;

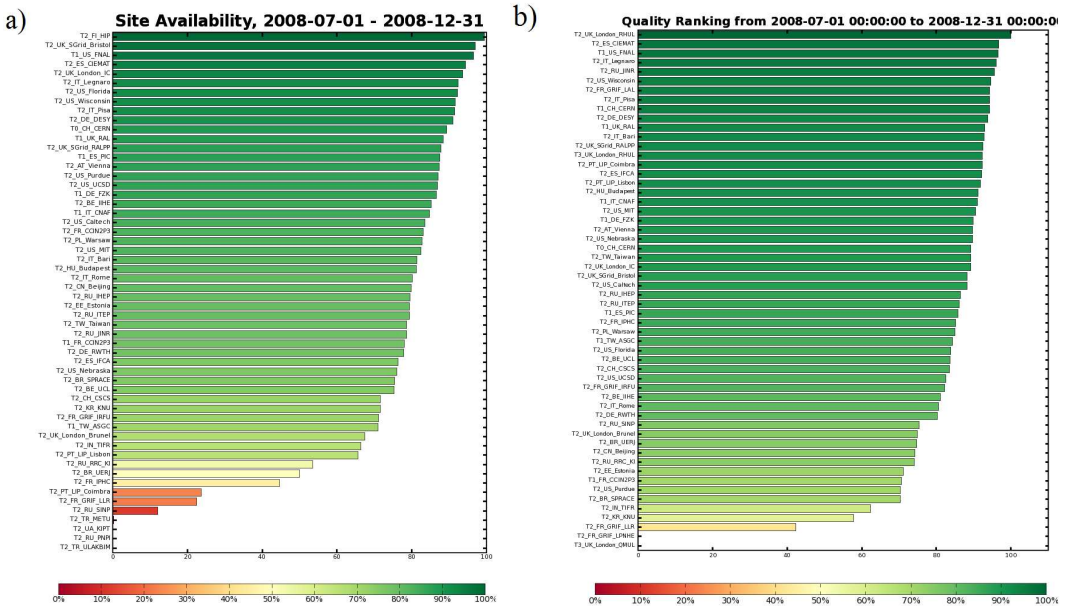


Figure 1: a) Average site availability; b) average Job Robot success rate for all Tier-1 and Tier-2 sites.

- the daily success rate of automatically submitted analysis-like jobs;
- the number of commissioned data transfer links with other sites.

These tests and monitor tools are detailed below.

2.1 Site availability

In the WLCG, all Grid services are periodically tested using a framework called SAM (Service Availability Monitor) [3], which executes periodic tests on all the Grid services within the infrastructure. SAM provides one of the main sources of information for the Grid operations and is used to measure the availability of Grid services.

CMS has adopted SAM to run custom tests on the Computing Elements (CE) and Storage Resource Manager (SRM) instances at the sites. Each service type has defined one or more critical tests. Given a time interval in the past, the *service availability* is defined as the fraction of time the service was passing all the critical tests, while the *site availability* is the fraction of time at least one instance of each relevant Grid service type at the site is available. The critical tests for the CE and the SRM are listed in table 1.

In the second half of 2008, the average availability of Tier-1 sites was 85%, while for Tier-2 sites it was 69%, including downtimes due to scheduled interventions. In Fig. 1a, all sites are shown by decreasing average availability.

2.2 Job Robot

Another complementary testing method consists of regularly submitting jobs similar to real analysis jobs. The difference with respect to the SAM tests is the fact that the statistics are much

Test name	Test definition
Computing Element	
js/jsprod	Checks the job submission via Grid
basic	Checks the local CMS configuration
swinst	Checks the locally installed CMS software
mc	Checks the file transfer mechanism to the local SE
frontier	Checks access to the calibration data via the local cache
squid	Queries the local calibration server
analysis	Checks the accessibility of a local dataset
Storage Resource Manager	
get-pfn-from-tfc	Performs logical-to-physical file name translation
lcp-cp	Transfers files to and from the local SRM

Table 1: CMS critical tests for the CE and the SRM services.

higher ($O(100)$ jobs/(site$\times$day)), the fact that the accessed data can be spread on several disks and a higher load on the site storage system. A tool called *Job Robot* was developed to implement such automatic job submission system using CRAB, the CMS analysis job submission tool [4].

At regular time intervals, a new analysis task is created for each site, to be run on a specific dataset. The task is then split into several jobs, which are submitted as a collection to the gLite WMS [5]. Each job performs a trivial data analysis on a fraction of the dataset. All submitted jobs are classified as successful, as failed at the application level or as aborted at the Grid level.

At Tier-1 sites, the Job Robot can also be used to emulate data skimming, which is input-bound and hence more demanding on the storage infrastructure.

The Job Robot daily statistics are used to measure the success rate for each site. Currently, around 25,000 jobs are submitted daily to approximately 60 sites. However, this load is not capable to fill all available slots at the sites. The Job Robot can be tuned to saturate all available CMS slots at the sites, and then compared to the resource pledges to CMS; this is useful also to uncover possible bottlenecks or scaling problems in the site services.

During the second half on 2008, the Tier-0 and Tier-1 sites had an average success rate of 89%, and Tier-2 sites of 85%. In Fig. 1b, all sites are shown by decreasing success rate.

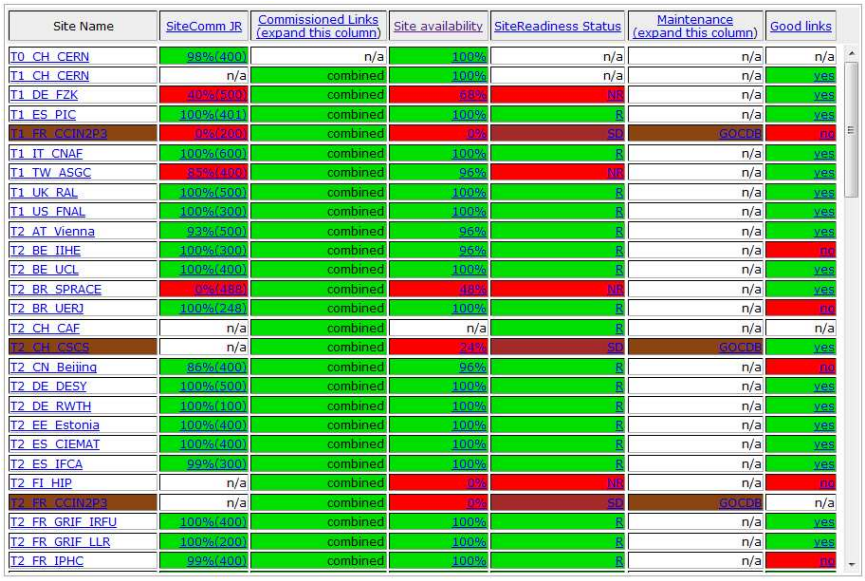
2.3 Data transfer links

A site needs to have sufficient data transfer connections to other sites in order to perform CMS workflows. In 2007, a Debugging Data Transfers (DDT) task force was created to design and enforce a procedure to debug problematic links [6]. This procedure uses a traffic generator to test the quality of a link and considers a link to be commissioned when it demonstrates:

- for links with source at Tier-0 or Tier-1 sites, 20 MB/s averaged over 24 hours;
- for links with source at Tier-2 sites, 5 MB/s averaged over 24 hours.

The numbers of commissioned links at the end of 2008 were:

- 56/56 Tier-(0,1)$\leftrightarrow$Tier-1 links (100%);



Site Name	SiteComm JR	Commissioned Links (expand this column)	Site availability	SiteReadiness Status	Maintenance (expand this column)	Good links
T0_CH_CERN	98%(400)	n/a	100%	n/a	n/a	n/a
T1_CH_CERN	n/a	combined	100%	n/a	n/a	yes
T1_DE_FZK	40%(500)	combined	68%	NR	n/a	yes
T1_ES_PIC	100%(401)	combined	100%	R	n/a	yes
T1_FR_CCIN2P3	0%(200)	combined	0%	SR	GOCDB	no
T1_IT_CNAF	100%(600)	combined	100%	R	n/a	yes
T1_TW_ASGC	85%(400)	combined	96%	NR	n/a	yes
T1_UK_RAL	100%(500)	combined	100%	R	n/a	yes
T1_US_FNAL	100%(300)	combined	100%	R	n/a	yes
T2_AT_Vienna	93%(500)	combined	96%	R	n/a	yes
T2_BE_IIHE	100%(300)	combined	96%	R	n/a	no
T2_BE_UCL	100%(400)	combined	100%	R	n/a	yes
T2_BR_SPRACE	0%(488)	combined	48%	NR	n/a	yes
T2_BR_UERJ	100%(248)	combined	100%	R	n/a	no
T2_CH_CAF	n/a	combined	n/a	R	n/a	n/a
T2_CH_CSCS	n/a	combined	24%	SR	GOCDB	yes
T2_CN_Beijing	86%(400)	combined	96%	R	n/a	no
T2_DE_DESY	100%(500)	combined	100%	R	n/a	yes
T2_DE_RWTH	100%(100)	combined	100%	R	n/a	yes
T2_EE_Estonia	100%(400)	combined	100%	R	n/a	yes
T2_ES_CIEMAT	100%(400)	combined	100%	R	n/a	yes
T2_ES_IFCA	99%(300)	combined	100%	R	n/a	yes
T2_FI_HIP	n/a	combined	0%	NR	n/a	no
T2_FR_CCIN2P3	n/a	combined	0%	SR	GOCDB	n/a
T2_FR_GRIF_IRFU	100%(400)	combined	100%	R	n/a	yes
T2_FR_GRIF_LLR	100%(200)	combined	100%	R	n/a	yes
T2_FR_IPHC	99%(400)	combined	100%	R	n/a	no

Figure 2: A view of the Site Status Board: each row corresponds to a site and each column to a site readiness variable.

- 280/352 Tier-1→Tier-2 links (80%);
- 42/44 Tier-2→Tier-1 regional links (95%);
- 92/308 Tier-2→Tier-1 non-regional links;
- 17 Tier-2→Tier-2 links.

In total, 487 links had achieved commissioning.

Recently, a procedure has been set up to continuously monitor the data transfer quality on all active links by probing them with low rate data transfers. A bad transfer quality can be due not only to problems at the network level, but also in the data transfer services and the storage infrastructure.

2.4 The Site Status Board

The *SiteStatus Board* (SSB) is a synoptic view of the status of all CMS computing sites [7]. It is designed to allow users to correlate the output of their workflows with known problems at sites, and to provide experts with a single entry point to the full suite of CMS monitoring tools. The provided information is often changed as the understanding of what is most relevant for making a good diagnosis of problems improves.

The SSB is a flexible presentation layer above a dynamic framework where information is stored in the columns of a database table, having the site name as key. These columns are filled by processes collecting data from the internal CMS Dashboard database [8], the WLCG information system and ASCII files on the web. Columns can be defined and added via a web form without any additional development work. The time history of any column can be graphically displayed or retrieved in XML format. Finally, several columns can be displayed together in "views" (Fig. 2).

Tier-1 sites	**Tier-2 sites**
daily SAM availability $\geq 90\%$	daily SAM availability $\geq 80\%$
daily Job Robot efficiency $\geq 90\%$	daily Job Robot efficiency $\geq 80\%$
downlink from T0 commissioned	commissioned uplinks to T1s ≥ 1
commissioned downlinks with T2s ≥ 10	commissioned downlinks from T1s ≥ 2
commissioned down/uplinks with other T1s ≥ 4	

Table 2: Site Commissioning daily metrics for Tier-1 and Tier-2 sites.

2.5 Site commissioning criteria

The quantities defined above must satisfy some constraints, to consider the site as "ready". Ideally, these constraints should be defined in such a way as to *a)* allow temporary glitches, *b)* enforce a reasonable level of reliability over a period of time and *c)* allow sites to quickly recover their "ready" status when problems are solved. In addition to that, downtimes due to scheduled maintenance and failures during weekends (for Tier-2 sites) should not be negatively considered in the site evaluation.

The evaluation of a site is expressed by a flag with three possible values: *Commissioned*, *Warning* and *Uncommissioned*, which mean respectively that the site is fully usable, that it is usable but suffering from temporary problems and that the site is unusable. Each day, a site is evaluated as *good* if the conditions in table 2 are satisfied, and as *bad* otherwise.

As a function of the site status in the previous seven days, the readiness flag is defined as follows:

- *Commissioned*: the site was good in the last 2 days, or in the last day and for ≥ 5 days in the last 7;
- *Warning*: the site is bad in the last day and was good for ≥ 5 days in the last 7;
- *Uncommissioned*: otherwise.

For Tier-2 sites, errors occurring during weekends are ignored, as they are not supposed to provide support during off hours.

In Fig. 3, the number of sites in each readiness status is plotted as a function of time during 2008. A trend towards increasing numbers of good sites is evident for Tier-2 sites, from 14 at the beginning of October to an average of 26 at the end of December.

3. Analysis tests

During the last WLCG data challenge, CCRC08 [2], CMS performed extensive tests at Tier-2 sites to better understand the performance and the readiness of those sites when running data analysis.

The first phase consisted of three types of analysis workflows, centrally submitted:

- long-running CPU intensive jobs with moderate I/O and no output file stageout, with the goal of filling all batch slots available to analysis at a given site without stressing the site;

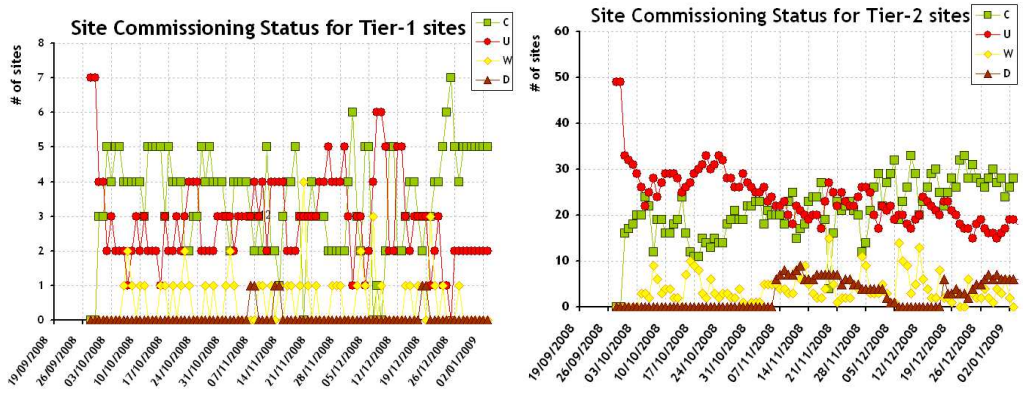


Figure 3: Number of sites in each readiness status as a function of time for Tier-1 *(left)* and Tier-2 *(right)*.

- long-running I/O intensive jobs to stress the storage infrastructure at the sites;
- short-running jobs of O(10 min) with local stage out of O(10 MB) file as output. These jobs ran for a short time, with many jobs finishing per hour, thus leading to a significant write request rate at large sites.

More than 10^5 jobs were submitted to 40 sites, with error rates varying from less than 1% at many sites to up to 50% at a few sites due to catastrophic storage failures. The failures were predominantly due to storage problems at sites. In most cases, those problems were detected and fixed by the site administrators within 24 hours. Ignoring the storage failures, the success rate in this exercise was found to be better than 99%. Overall success rate including the storage issues, ranged between 92-99% for these exercises.

A second exercise consisted in closely mimic physics group activities on a subset of good sites. In CMS, a physics group has at its disposal the resources of a given set of Tier-2 sites to run data analysis. The exercise consisted in:

- the definition of three physics groups;
- the association of a list of sites to each physics group;
- the use of the CRAB server for analysis tasks reading a dataset at all sites and running for about 4 hours with remote stageout of a O(20 MB) output file to a subset of Tier-2 sites (to simulate the model where each user has private space at a Tier-2).

More than 10^5 jobs were submitted in two weeks to 29 sites. The job distribution by the site is shown in Figure 4. Most failures were due to problems accessing the input data, with success rates normally between 0.1% and 10% but up to 50% in pathological cases, and remote stageout issues were due to old Grid clients affecting, at some sites, all submitted jobs. These stageout issues were promptly fixed by the site administrators. During the second week, the number of sites with efficiency above 90% significantly increased, as shown in Fig. 4.

During the last two weeks of the challenge, real users were encouraged to submit jobs to produce a more chaotic submission pattern (Fig. 5). The total latency from dataset download

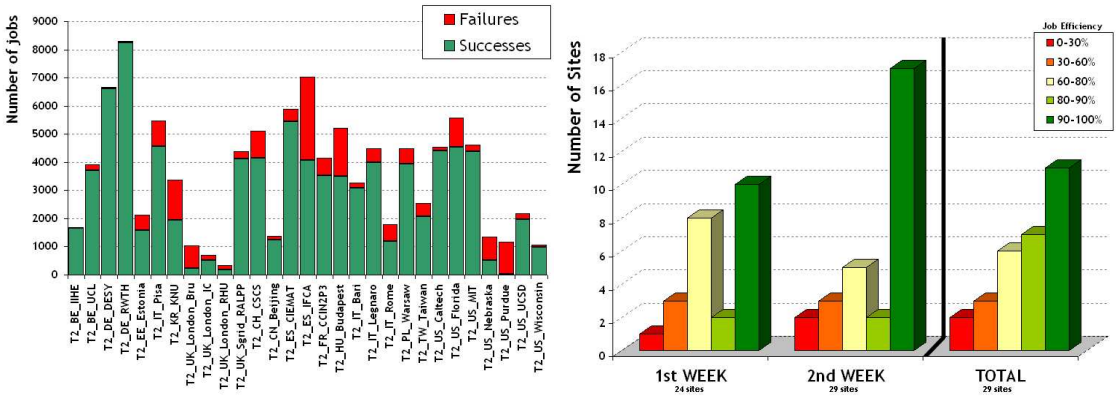


Figure 4: *(left)* Job distribution by site, and *(right)* distribution of the job efficiency by site, when simulating physics groups workflows during CCRC'08.

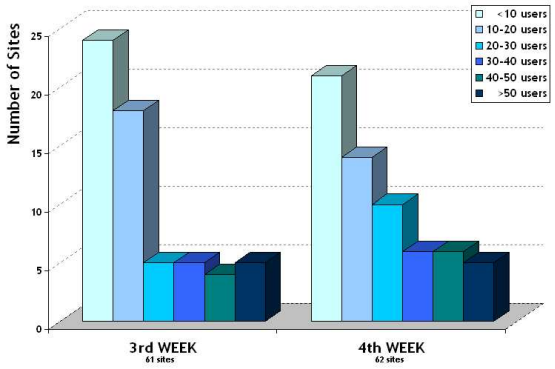


Figure 5: Distribution of the number of active users at sites, when simulating chaotic job submission on CCRC'08.

to analysis task completion was measured, giving results ranging from a few hours to days and dominated by the data transfer and the job completion times.

4. Conclusions

The CMS site commissioning activity has provided a series of tools to monitor the behaviour of CMS sites when running CMS workflows, and defined a single estimator that can be provided to the Monte Carlo production managers and the physicists to understand which sites can be reliably used. At the same time, sites have the tools they need to see and debug site problems. To this end, the analysis tests during the CCRC08 provided useful information about the functionality that needs to be tested.

The CMS computing shift team is responsible for identification and notification of a problem via tickets to problematic sites, which are also encouraged to provide a better service, if they want to be used to execute CMS workflows. Results obtained so far show a clear improvement in the number of "ready" Tier-2 sites, while for Tier-1 sites there is no significant change, compatibly

with the fact that they are constantly used and the level of reliability is not expected to change much over time.

References

[1] CMS Collaboration, *CMS Computing Project: Technical design report*. CERN-LHCC 2005-023, 2005.

[2] J.D. Shiers *et al.*, *The (WLCG) Common Computing Readiness Challenge(s) - CCRC'08*, contribution N29-2, session *Grid Computing*. Nuclear Science Symposium, IEEE (Dresden), October 2008.

[3] A. Duarte, P. Nyczyk, A. Retico, D. Vicinanza, *Monitoring the EGEE/WLCG Grid Services*, J. Phys.: Conf. Ser. **119** (2008) 052014.

[4] D. Spiga *et al.*, *The CMS Remote Analysis Builder (CRAB)*, LNCS vol. 4873, pp. 580-586 (2007).

[5] P. Andreetto *et al.*, *The gLite workload management system*, J. Phys.: Conf. Ser. **119** (2008) 062007.

[6] G. Bagliesi *et al.*, *The CMS Data Transfer Test Environment in Preparation for LHC Data Taking*, contribution N67-2, session *Applied Computing Techniques*. Nuclear Science Symposium, IEEE (Dresden), October 2008.

[7] R. Rocha *et al.*, *Experiment Dashboard for Monitoring of the Computing Activities of the LHC Experiments on the Grid*, contribution N29-4, session *Grid Computing*. Nuclear Science Symposium, IEEE (Dresden), October 2008.

[8] J. Andreeva *et al.*, *Dashoard for the LHC Experiments*, Proceedings of International Conference on Computing in High Energy and Nuclear Physics (CHEP 07) J.Phys.Conf.Ser.119:062008,2008.