Critical services in the LHC computing

A Sciabà
CERN, CH-1211 Genève 23, Switzerland
E-mail: Andrea.Sciaba@cern.ch

Abstract. The LHC experiments (ALICE, ATLAS, CMS and LHCb) rely for the data acquisition, processing, distribution, analysis and simulation on complex computing systems, running using a variety of services, provided by the experiments, the Worldwide LHC Computing Grid and the different computing centres. These services range from the most basic (network, batch systems, file systems) to the mass storage services or the Grid information system, up to the different workload management systems, data catalogues and data transfer tools, often internally developed in the collaborations. In this contribution we review the status of the services most critical to the experiments by quantitatively measuring their readiness with respect to the start of the LHC operations. Shortcomings are identified and common recommendations are offered.

1. Introduction
The Large Hadron Collider at CERN is expected to start its scientific programme at the end of 2009. All the experiments will then start collecting data, eventually reaching a volume of 15 PB per year. This data will be processed to reconstruct events from collisions, to calibrate the detectors and to perform physics analysis. Event simulation is ongoing since several years, while in the last two years data from cosmic ray interactions has been collected for the commissioning of the detectors.

The LHC computing activities are performed on a distributed infrastructure, built and maintained by the Worldwide LHC Computing Grid project since 2002 [1]. The WLCG not only provides the hardware infrastructure, in most cases via the EGEE, OSG and NorduGrid projects [2], but also distributes the Grid middleware, mainly developed by EGEE and VDT [3]. In addition, the WLCG encourages the interoperability among different middleware stacks, provides support to the LHC experiments in using the Grid, monitors the availability and the reliability of the resources at the participating sites and ensures an adequate level of performance.

Each experiment, during the past years, has developed its own computing system in order to run all their computing activities. Different requirements, event data models and architectural choices made the resulting systems very different from each other, although all of them are based on the tiered WLCG architecture. According to a simplified description, sites have these roles:

• Tier-0 at CERN: first-pass event reconstruction, custodial storage of all raw data, distribution of raw and reconstructed data to Tier-1 sites;
• Tier-1 sites: event reconstruction, data reprocessing, custodial storage of part of the raw data, of reconstructed data and simulated data, distribution of reconstructed data to Tier-2 sites, analysis;
2. Services in the WLCG computing

2.1. Overview

All the computing systems of the LHC experiments are rather complex and depend on a large number of services installed at the participating sites. They can be classified as follows:

- experiment-specific services;
- WLCG services;
- infrastructure services.

Experiment-specific services take care of a wide range of high-level functionalities; they are developed inside the collaboration and are normally very tied to the computing model of the experiment. Three main areas are covered by them:

- data management, including data bookkeeping and data replication and movement;
- workload management, that is the creation, submission and tracking of data processing jobs, both for organized processing (data reconstruction and skimming, Monte Carlo simulation) and end-user analysis;
- calibration and alignment database, including the distribution of calibration data to external sites.

The WLCG services constitute the Grid layer of the system, connecting the experiment layer to the underlying site resources. They are intended to implement the common functionality needed by all the LHC experiments, like a file transfer service, a generic access interface to storage and computing resources, a job management system, a repository for authentication and authorization data.

Finally, the infrastructure layer includes the basic services provided by sites, such as a storage system, a computing farm, a database cluster and everything else the experiments may be relying upon (a shared file system, a web server, a bug tracking system, etc.).

In figure 1 a generic service hierarchy is depicted for the Tier-0 and for a Tier-1 site: details can vary for actual experiments.

Figure 1. Service hierarchy at the Tier-0 (left) and at a Tier-1 (right).

- Tier-2 sites: generation of simulated data, data analysis.

In practice, roles depend also on the computing model of the experiment.
Table 1. Questions defining the service readiness.

| Software readiness                  |
|-------------------------------------|
| Is a high-level description of the service available? |
| Are the middleware dependencies and versions defined? |
| Is the code correctly released and packaged? |
| Does a certification process exist? |
| Are administration guides available? |

| Service readiness                   |
|-------------------------------------|
| Are the disk, CPU, database and network requirements defined? |
| Are the monitoring criteria described? |
| Is the problem determination procedure documented? |
| Is the support chain defined? |
| Is a backup/restore procedure defined? |

| Site readiness                      |
|-------------------------------------|
| Is a suitable hardware used? |
| Is the monitoring implemented? |
| Does a test environment exist? |
| Is the problem determination procedure implemented? |
| Is an automatic configuration system implemented? |
| Is the backup/restore procedure implemented and tested? |

2.2. Service readiness definition
The WLCG has defined a set of criteria to quantify how close a service is to be ready for LHC production. These consist in a series of questions (table 1), addressing several aspects (software distribution, documentation, monitoring, etc.): the fraction of affirmative answers can be taken as a measurement of the service readiness. It should be noted that these criteria do not include requirements on the performance and robustness of the services as it would be too difficult to define generic metrics for all the services: it is assumed that the required levels of performance and stability will be obtained by distributing the service on enough nodes using load balancing etc. It is worth to mention that WLCG also measures the site availability and reliability as a function of functional tests running on several types of Grid services and enforces certain targets to be met, in particular for CERN and the Tier-1 centres [4].

2.3. Service criticality definition
In order to measure the degree of criticality of a service in the context of the computing system of an experiment, each experiment has provided a list of services by tier and rated their criticality as a number in a scale from 0 to 10. An explicit meaning has been assigned to different criticality levels: for completeness, these meanings are listed in table 2. It can be seen that the meaning of a criticality level differs in general from one experiment to another: nevertheless, they all provide a reasonable quantitative indication of the impact of a service downtime for the experiment.

The list of the critical services has been provided by each experiment as part of the preparation for the CCRC’08 data challenge [5]. Tables 3-6 show the services defined as critical together with the criticality rank according to the definitions listed above.
Table 2. Criticality levels by experiment.

| Level | ALICE | ATLAS | CMS | LHCb |
|-------|-------|-------|-----|------|
| 0     | 0     | Not needed | Not critical | Not critical |
| 1     | 1     | Not critical | Not critical | Not critical |
| 2     | 2     | Reduced effectiveness | CERN stops | Reduced effectiveness |
| 3     | 3     | Service interruption perturbs software development or part of computing operations | CMS development stops | Major reduction in effectiveness |
| 4     | 4     | Reduced efficiency, maximum downtime 12 hours | CMS development stops | Major reduction in effectiveness |
| 5     | 5     | Serious disruption, maximum downtime 8 hours | Service critical but not needed all the time | Serious disruption |
| 6     | 6     | Service interruption seriously perturbs offline computing operations | Tier-1/2 production or analysis stops | Tier-0 production stops |
| 7     | 7     | Service interruption affects data taking or stops all offline operations | Tier-0 production stops | Critical |
| 8     | 8     | Service interruption affects data taking or stops all offline operations | CMS stops transferring data from the pit | Critical |
| 9     | 9     | Service interruption affects data taking or stops all offline operations | CMS stops transferring data from the pit | Critical |
| 10    | 10    | Critical, maximum downtime 2 hours | CMS stops transferring data from the pit | Critical |

2.3.1. **ALICE.** The ALICE collaboration has defined the service criticality as in table 3 [6]. The ALICE computing is entirely based on AliEN [7], which encompasses the data management and the job submission and tracking; it includes central services and site services, running on “VO boxes” (i.e. nodes running experiment services). Together with the CERN storage system, based on CASTOR and xrootd, it gets the maximum criticality and a maximum downtime of two hours is allowed. Mass storage at Tier-1 sites acts only as a backup for data which exists also at CERN, and which must be accessible from disk anyway: therefore a downtime can last up to 12 hours without consequences. The File Transfer Service [8] at the Tier-0 is needed to move data to Tier-1 sites, which is necessary for running event re-reconstruction; a downtime of up to 8 hours is acceptable. The gLite Workload Management System (WMS) [9] has several instances installed at different sites, so the criticality of any given WMS instance is low; moreover it is expected that the WMS will become irrelevant when all sites will have a CREAM [10] computing element available. Finally, the PROOF [11] farm at the Tier-0 analysis facility has a similarly low criticality, its unavailability affecting only end-user physics analysis.

2.3.2. **ATLAS.** The criticality of the services used by ATLAS is described in table 4 [12]. The most critical components are the online database, without which acquisition of raw data is impossible, and the catalogues of the Distributed Data Management (DDM) [13], used by all offline activities.

An interruption of data transfer from the pit to CASTOR is acceptable as long as it does not last more than one day. Similarly, downtimes limited in time are acceptable for the exporting of
Table 3. Critical services in ALICE.

| Rank | Services                                                                                     |
|------|-----------------------------------------------------------------------------------------------|
| 10   | AliEN central services, Site VO boxes, CASTOR and xrootd at Tier-0                           |
| 7    | FTS at Tier-0                                                                                |
| 5    | Mass storage and xrootd at Tier-1, gLite WMS, PROOF at Tier-0 CERN analysis facility         |

Table 4. Critical services in ATLAS.

| Rank | Services                                                                                      |
|------|-----------------------------------------------------------------------------------------------|
| 10   | Oracle (online), DDM central catalogues, LFC                                                  |
| 7    | Data transfers from pit to CERN, online-offline database connection, CASTOR, Computing Element, Oracle (offline), FTS, VOMS, LFC, dashboard, Panda/Bamboo servers, DDM site services |
| 4    | 3D streaming, gLite WMS, Storage Element, CAF, CVS, Subversion, AFS, build system, Tag Collector |

| Tier-1 services |                                                                                               |
|-----------------|                                                                                                |
| 7               | LFC, FTS, Oracle                                                                             |
| 4               | 3D streaming, Storage Element, Computing Element                                               |

| Tier-2 services |                                                                                               |
|-----------------|                                                                                                |
| 4               | Storage Element, Computing Element                                                            |

| Other services  |                                                                                               |
|-----------------|                                                                                                |
| 7               | AMI database                                                                                  |

conditions data from the online to the offline database, the Tier-0 Grid resources and services, the production servers (Panda and Bamboo [14]) and the DDM site services: a service interruption determines a growing backlog of operations (data transfers and processing) giving a grace period of 1-2 days. Services with moderate criticality may impact software development or a subset of the resources.

2.3.3. CMS. Critical services for CMS are listed in table 5 [15]. As for ATLAS, a downtime of CASTOR will stop data transfer from the online system to CERN, reconstruction, calibration and data export to Tier-1 sites; the buffer at the online farm will fill in 1-2 days. A downtime of Oracle or DBS will stop the data acquisition and the event reconstruction; the latter will be stopped also in case of an unavailability of the processing farm.

Data export to Tier-1 sites will stop in case of an unavailability of FTS or PhEDEx. For the Frontier launchpad the effect will be stopping the export of calibration data. AFS is used as repository of several scripts and files used by the production system and the CERN analysis facility to run high priority analysis and calibration.
Table 5. Critical services in CMS.

| Rank | Services |
|------|----------|
| 10   | Oracle, CASTOR, SRM, DBS, processing farm, Kerberos, data transfer and processing from pit to Tier-0 |
| 9    | FTS, PhEDEx, Frontier Launchpad, AFS, CAF |
| 8    | gLite WMS, VOMS, MyProxy, BDII, WAN, other production tools |
| 7    | APT servers, build and testbed machines, tag collector, CMS web server, Twiki pages |
| 6    | SAM, dashboard, PhEDEx monitoring, Lemon monitoring |
| 5    | Web tools, e-mail, Hypernews, Savannah, CVS server |
| 4    | Linux repository, phone conferencing, Valgrind machines |
| 3    | Benchmarking machines, Indico |

Production and analysis activities on the Grid will stop in case a central Grid service like the WMS, VOMS [16] or the BDII is down (these services are normally highly redundant, however). The same applies to CMS services, like CRAB and the ProdAgent (used respectively to manage analysis and production jobs).

Other services, mainly used for monitoring, software development and information exchange, have lower levels of criticality.

2.3.4. LHCb. The LHCb collaboration has indicated as critical the services in table 6 [17].

At CERN, on the experiment side, the maximum criticality is attributed to the DIRAC3 central services and the central catalogue, without which all offline computing stops. CASTOR is needed for all data movement and AFS is used by the central services and to access the gLite user interface.

A VOMS downtime of a few hours is tolerated as it just implies the impossibility to create fresh user proxies to submit jobs; the proxy validity is normally of the order of a few days. FTS is needed to export data to Tier-1 and the Conditions database to run data reprocessing. The CE is needed to run reconstruction and analysis at CERN.

At Tier-1 sites the SE is essential for all data movement and processing, while the CE is essential only for the data processing. The FTS is used to redistribute DST data among Tier-1 sites. Finally, downtimes of the gLite WMS, of the replica LFC [18] and the Conditions database will just cause a reduction in computing capacity or in the fault tolerance.

2.4. Survey of the service readiness status

During November 2008, and again in March 2009, a survey was done by WLCG to assess the status of the LHC computing in terms of service readiness in view of the imminent start of data taking. The motivation was to understand if there were shortcomings which could seriously affect offline operations if not addressed in time.

The survey was conducted by asking the service managers to answer the questions in table 1. The survey was mostly focused on CERN for the WLCG and the infrastructure services, whose status is summarized in table 7. In general, it appears that most services are already in a very good condition, the most significant problem being the lack of a “piquet” service (a rotation of
Table 6. Critical services in LHCb.

| Rank | Services at Tier-0 |
|------|-------------------|
| 10   | CERN network, CASTOR, DIRAC3 central services (on VO boxes), AFS, master LFC, Oracle |
| 7    | VOMS, FTS, CE, processing farm, Conditions database, LHCb bookkeeping service, Oracle Streams, SAM |
| 5    | gLite WMS |
| 3    | Replica LFC, dashboard |

Services at Tier-1

| Rank | Services |
|------|----------|
| 7    | SE       |
| 5    | CE, processing farm, Conditions database |
| 3    | VO boxes, Replica LFC, FTS |
| 1    | gLite WMS |

Table 7. CERN WLCG services readiness.

| Service                  | Readiness                                                                 |
|--------------------------|---------------------------------------------------------------------------|
| **Data services**        |                                                                           |
| CASTOR, SRM, FTS, LFC    | 100% ready                                                               |
| Oracle                   | No piquet service available                                               |
| **Computing services**   |                                                                           |
| CE, batch services       | No expert piquet service available                                        |
| gLite WMS                | Insufficient monitoring and problem detection, no expert piquet service, no backup |
| **Other Grid services**  |                                                                           |
| MyProxy                  | Procedures not fully documented, no expert piquet service                 |
| VOMS                     | Problem determination procedure does not cover everything, no expert piquet service |
| BDII                     | Some documentation slightly outdated, no expert piquet service            |
| VO box                   | No expert piquet service                                                  |
| Dashboard                | No certification process, no automatic configuration                      |
| SAM                      | 100% ready                                                                |
| **Other non-Grid services** |                                                                       |
| AFS, Kerberos            | No certification process at CERN, problem determination procedure not documented, no test environment |
| Twiki                    | Relies heavily on AFS for backend data and backups                         |

available staff) to provide expert on-call support outside working hours. Nevertheless, also in this case, a “best effort” support is provided. The second most significant problem is related to documentation of procedures being not always complete. In this case it is expected to improve with the operational experience of the services.
2.5. Readiness status of the experiment services

According to the survey ALICE services fulfill all the readiness criteria. The ATLAS central services partially rely on people on shift to determine problems but the solution very often requires an expert. A certification process exists for the DDM site services but the pre-production instance is actual a subset of the production system. Hardware requirements are well known, apart from the case of data analysis, where there are still uncertainties in the amount of resources needed by the ATLAS services to sustain the load in the coming months. All backups are managed via Oracle at CERN.

The CMS computing system is also very close to be fully ready: what should be improved is the monitoring of the production system and some problem determination procedures. Finally, the Tier-0 system is lacking in documentation and does not have a proper certification process, which is explained by the need to make updates in a very short time scale. Nevertheless, the system is able to run efficiently. A strong point of CMS is a thorough documentation of procedures for installation, configuration, startup and testing. All relevant backups are done by the CERN IT.

In LHCb, the results of the survey are also very good the only shortcomings being related to incomplete documentation both for the administration of the services and for the problem solving.

3. Conclusions

Six months before the foreseen start of the LHC data taking, a survey indicates that the vast majority of the services used in the LHC computing are fully ready to be run in a production system. The only concerns are related to monitoring, documentation and the lack of a piquet service at CERN for the support of several central services, replaced by a “best effort” support. Monitoring and documentation are expected to improve naturally as operational experience is collected which will have a beneficial impact on the manpower requirements for operations. The effectiveness of best effort support will be fully appreciated only on a longer time scale.

Acknowledgments

The author would like to thank Patricia Bittencourt, Simone Campana, Birger Koblitz, Patricia Méndez Lorenzo, Roberto Santinelli and the CERN IT department for their input, without which this review would not have been possible.

References

[1] http://leg.web.cern.ch/LCG/
[2] http://eu-egee.org/
  http://www.opensciencegrid.org/
  http://www.nordugrid.org/
[3] http://www.eu-egee.org/
  http://vdt.cs.wisc.edu/
[4] Duarte A, Nycz P, Retico A and Vicinanza D 2008 J. Phys.: Conf. Ser. 119 052014
[5] https://twiki.cern.ch/twiki/bin/view/LCG/WLCGCommonComputingReadinessChallenges
[6] https://twiki.cern.ch/twiki/bin/view/ALICE/CCRC08, private communication
[7] Saiz P et al 2003 Nucl. Instr. Meth. A 502 437
[8] Abadie L et al 2007 24th IEEE Conference on Mass Storage Systems and Technologies (Los Alamitos, CA: IEEE Computer Society) 60-71
[9] Andreetto P et al. 2008 J. Phys.: Conf. Ser. 119 062007
[10] Aiffimieci C et al 2008 Using CREAM and CEMON for job submission and management in the gLite middleware these proceedings
[11] Ballintijn M et al 2006 Nucl. Instr. Meth. A 559 13-16
[12] https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasCriticalServices, private communication
[13] Campana S 2009 Experience Commissioning the ATLAS Distributed Data Management system on top of the WLCG Service these proceedings
[14] Maeno T 2008 J. Phys.: Conf. Ser. 119 062036
[15] https://twiki.cern.ch/twiki/bin/view/CMS/SWIntCMSServices, private communication
[16] Alferi R et al 2008 VOMS, an Authorization System for Virtual Organizations European Across Grids Conference Proceedings 33-40
[17] https://twiki.cern.ch/twiki/bin/view/LHCb/CCRC08
[18] Frohner Á et al 2008 Data Management in EGEE these proceedings