The commissioning of CMS sites
improving the site reliability

Stefano Belforte, INFN Trieste, Italy; **Ian Fisk**, Fermilab, US; **Josep Flix**, CIEMAT, Spain; **José Hernandez**, CIEMAT, Spain;
Jukka Klem, Helsinki Institute of Physics, Finland; **James Letts**, University of California San Diego, US;
Nicolò Magini, CERN, Switzerland; **Pablo Saiz**, CERN, Switzerland; **Andrea Sciabà**, CERN, Switzerland;
On behalf of CMS Offline and Computing

ABSTRACT

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.

SITE COMMISSIONING

https://twiki.cern.ch/twiki/bin/view/CMS/PADASiteCommissioning

The CMS site commissioning is one of the activities of the PADA (Processing and Data Access) Task Force, whose objective is to guarantee that the data processing workflows at Tier-1 and Tier-2 sites can be performed efficiently and reliably.

CMS distributed computing requires stable and reliable behavior of the underlying infrastructure at all times. Heterogeneity and different amount of computing resources, plus support.

The site commissioning makes use of several sources of information to assess the readiness of a site to run CMS workflows:

- ❏ **CMS SAM tests**: Jobs sent to sites to test specific services;
- ❏ **Job Robot load generator**: Simple jobs reading data;
- ❏ **Data transfers**: Transfer quality and commissioned links;
- ❏ **Scheduled Downtimes**: from SAM/GridView DB on services used by CMS.

Site readiness metrics were established to guarantee data processing can be performed efficiently and reliably.

SITE AVAILABILITY

Like other LHC experiments, CMS uses the Service Availability Monitoring framework to run specific tests on computing (CE) and storage (SRM) resources at the CMS sites. These tests allow to determine, among other things, whether:

- ❏ It is possible to send and run jobs;
- ❏ The CMS software is correctly installed and configured;
- ❏ It is possible to access in a job local CMS data;
- ❏ It is possible to copy data in and out of the local storage.
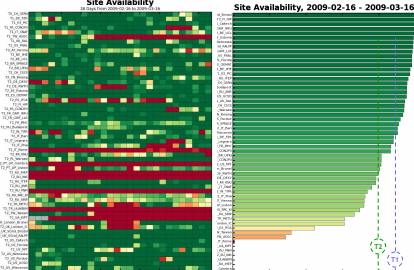
A failure of a critical test means that the site is considered unavailable, until all critical tests are passed. The site availability is used since a very long time to estimate the quality of the site.

"Latest View" of all CMS critical SAM tests run at T1 sites

Site Availability last month history (left) and ranking (right) for all CMS sites tested

JOB ROBOT

The Job Robot is a tool to automatically create fake analysis jobs:

- ❏ A set of agents (automatic preparation, submission, collection);
- ❏ Uses CRAB, the CMS analysis job submission tool;
- ❏ Reads a dataset of ~500 GB available at the site.

Two operation modes:

- ❏ **Monitoring Mode**: constant low rate job submission;
- ❏ **Stressing Mode**: filling sites with jobs:
  - To test sites under stress up to pledged slots. [not tested yet]

Running in monitoring mode: ~25,000 jobs/day [~50 sites]. Fraction of successful jobs over the total number of finished jobs is calculated.

Job Robot success rate last month history (left) and ranking (right) for all CMS sites tested

DATA TRANSFERS

For sites to be usable, they need to have sufficient data transfer connections among them. A link is a [source site → destination site] pair for data transfer. These links need to be operational:

- ❏ Tier-0 → Tier-1: to export raw and reconstructed data;
- ❏ Tier-1 → Tier-1: for synchronisation and large-scale data reprocessing;
- ❏ Tier-1 → Tier-2: for distribution of data to be analyzed at Tier-2 sites;
- ❏ Tier-2 → Tier-1: to upload Monte Carlo events generated at Tier-2 sites.

The Debugging Data Transfers (DDT) task force, since July 2007, defined metrics, provided a procedure and the tools to test the links and assisted sites solving problems. Tests regularly run to commission new links on **Debug PhEDEx (Data Management) instance**. → reference [1]

The minimum requirements to commission a link are:

- ❏ 20 MB/s sustained for 24h for Tier-0 → Tier-1 and Tier-1 (↔)→ Tier-X links;
- ❏ 5 MB/s sustained for 24h for Tier-2 → Tier-1 links.

Each commissioned link is enabled for Production use. At March 2009:

- ❏ All 56 T0 to T1 and T1-T1 cross-links commissioned;
- ❏ 300/352 (85%) T1 → T2 downlinks commissioned;
- ❏ 140/352 (40%) T2 → T1 uplinks;
- ❏ 27 T2-T2 cross-links (not in Computing Model).

All ~500 production links are continuously exercised with fake transfers at 0.5 MB/s/link (in Debug instance). Adding Production transfers:
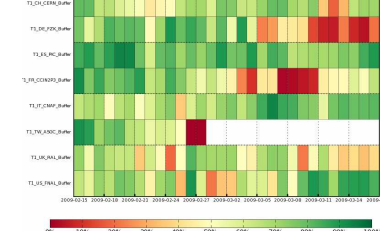
- ❏ Routine WAN transfers of ~50-100 TB/day (1-2 GB/s);
- ❏ Enough to detect systematic transfer problems;
- ❏ Use transfer quality information to decommission links.
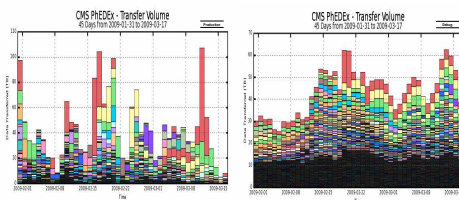
CMS Data transfers in Production (left) and Debug (right) PhEDEx instances, for all 'active' links

Daily average transfer quality on T1→T2 links for added Production/Debug instances

SCHEDULED DOWNTIMES

CMS service downtimes are extracted from SAM/Gridview DB:

- ❏ Consider services that have at least one critical test defined: CE and SRMv2;
- ❏ Only services in which CMS has run SAM tests last 30 days are included;
- ❏ Full-Maintenance when all instances of a service are in maintenance [CE-SD, SE-SD or SD (site's full maintenance)].

PUTTING ALL TOGETHER

→ reference [2]

Collect/display all site commissioning information in Site Status Board

Daily Metrics for Tier-1 sites	Daily Metrics for Tier-2 sites
daily SAM site availability ≥ 90%	daily SAM availability ≥ 80%
daily Job Robot efficiency ≥ 90%	daily JR-MM efficiency ≥ 80%
downlink from Tier-0 commissioned	having ≥ 2 commissioned links to Tier-1 sites
having ≥ 20 commissioned links to Tier-2 sites	having ≥ 4 commissioned links from Tier-1 sites
having ≥ 4 commissioned downlinks/uplinks with other Tier-1 sites	

CMS sites are subject to a set of **daily metrics**. Those combine into a single daily 'Site Readiness status':

**READY** - **WARNING** - **NOT-READY** - **SCHEDULED-DOWNTIME**

- ❏ It uses the history of last 7 overall Daily Metric;
- ❏ Allows some degree of unreliability in thresholds and stability;
- ❏ Intermediate warning state to give sites the time to recover;
- ❏ Weekend failures for Tier-2s do not negatively count in the evaluation [grey].
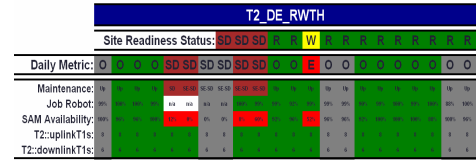
Example of the site readiness status and all daily metrics for a CMS Tier-2 site

Sites have an easy way to know if CMS is finding troubles at them:

- ❏ The program provides monitoring plots, XML feeds, nagios-plugin, and alerts;
- ❏ Results/problems are reported on weekly Facilities Operations meetings.
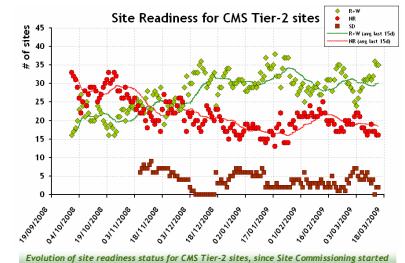
Evolution of site readiness status for CMS Tier-2 sites, since Site Commissioning started

Use site readiness status history to flag good/bad sites:

- ❏ Sites in **READY** or **WARNING** status are considered **stable & reliable**;
- ❏ The fraction of time a site has been stable & reliable (ignoring declared Downtimes) is shown for last 15 days;
- ❏ If a site has spent more than 75% days in Downtime, then is excluded in plot;
- ❏ So far used as indication for production/analysis but plan to enforce it. Jobs will be sent only to good sites.
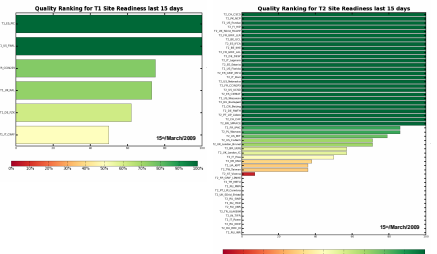
Site Readiness last 15 days ranking for Tier-1 (left) and Tier-2 (right) CMS sites

CONCLUSIONS

Site Commissioning activities **crucial** for bringing the CMS distributed computing system into stable & reliable operations → reference [3]

- ❏ Continuous monitoring of Grid & CMS services at sites;
- ❏ All the available information is now condensed in a single estimator, whose value takes in account also the stability of the site;
- ❏ Helps production and users to select reliable T2 sites;
- ❏ 6 months of data and positive trend for Tier-2 sites;
- ❏ Still room for improvements → inclusion of data transfers quality, production jobs statistics, etc...;
- ❏ Task Force created to determine usual failures, help sites to improve, feedback for robustness of CMS tools/services, increase reliability of sites…

@CHEP'09: [1] See poster #205 → J. Letts et al. "Debugging Data Transfers in CMS"
[2] See poster #275 → P. Saiz et al. "Generic monitoring solution for LHC site commissioning activity and LHC computing shifts"
[3] See talk #226 → J. Hernández et al. "Bringing the CMS Distributed Computing System into Scalable Operations"
→ The authors wish to thank especially the Dashboard team at CERN.