

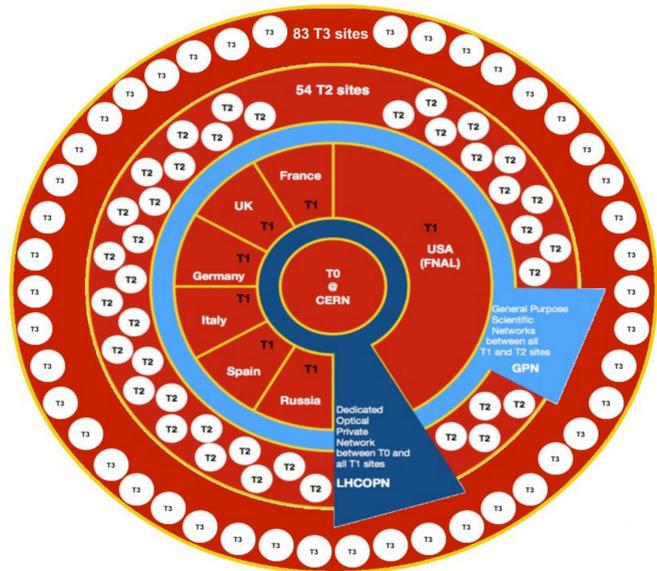


Grid Site Availability Evaluation and Monitoring at CMS



Introduction

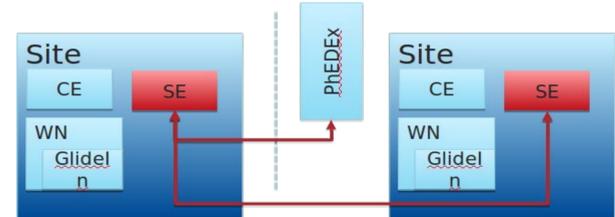
The Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) uses distributed grid computing to store, process, and analyze the vast quantity of scientific data recorded every year. The computing resources are grouped into sites and organized in a tiered structure. A tier consists of sites in various countries around the world. Each site provides computing and storage, in total about 125 sites contribute. In such a large computing setup scheduled and unscheduled outages occur continually and are not allowed to significantly impact data handling, processing, and analysis. Unscheduled capacity and performance reductions need to be detected promptly and corrected. CMS developed a sophisticated site evaluation and monitoring system for Run 1 of the LHC based on tools of the Worldwide LHC Computing Grid (WLCG). For Run 2 of the LHC the site evaluation and monitoring is being overhauled: A new Computing Resource Information Catalog (CRIC) is under development by WLCG. This will enable more flexible and dynamic resource management. Site/service tests are refined and new performance metrics are being setup. The frequency of the core tests, SAM and Hammer Cloud, was increased to 15 minutes so that site readiness and usage flags (statuses) can be evaluated for periods of hours instead of a one day value.



SAM

SAM is a complete monitoring framework for grid services and grid operational tools. It currently tests:

- Operation of grid services at each site.
- Functionality needed for a specific experiment like:
 - Site configuration
 - Specific software installation
 - Data access via xrootd protocol
 - Stageout to a remote site
- Execution of production-like jobs to test that the software stack for the execution of workflows is complete and functioning.



Transfer quality

CMS uses the Physics Experiment Data Export (PhEDEx) system as data transfer management system.

- The system measures the transfer quality between sites using both requested data movements and small test transfers scheduled at regular intervals.

HammerCloud

Hammer Cloud is a system developed by CERN IT to run custom tests on Grid sites to test functionality and detect site issues.

- It submits realistic analysis jobs via the experiments batch system to each site.
- It runs tests continuously (functional tests) or on demand (stress tests). Continuous testing is accomplished by releasing a short job for execution at a site every 5 minutes.

Site Readiness

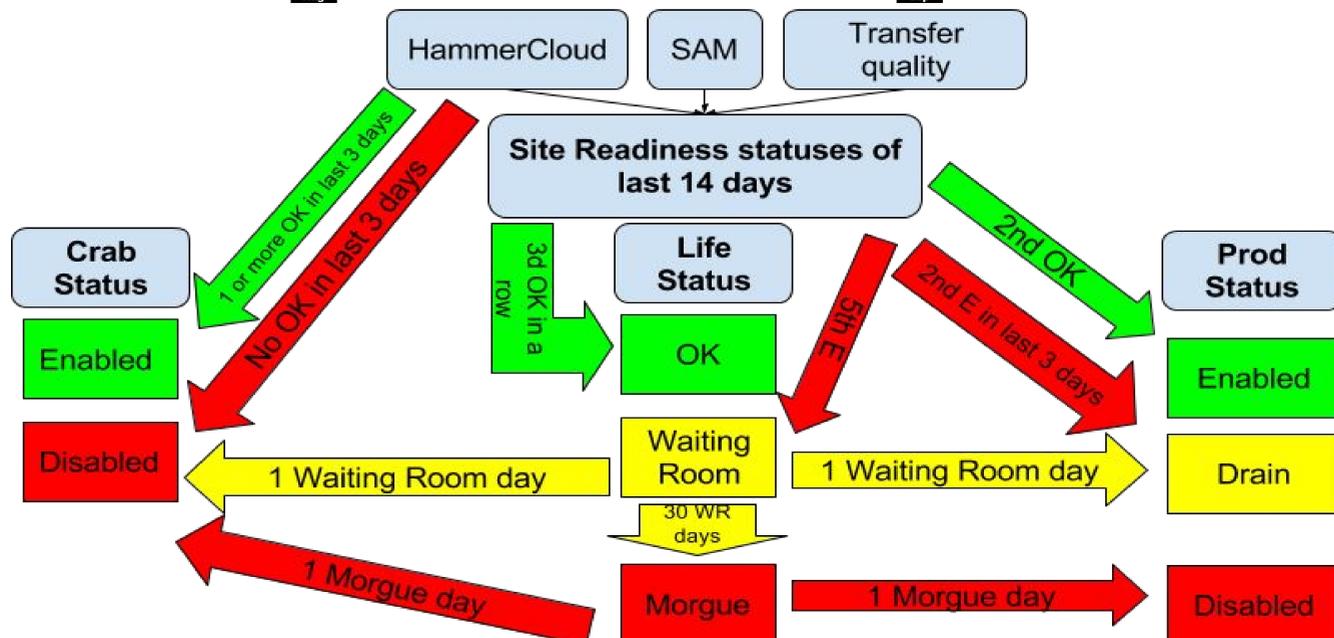
Data from tests (SAM, HammerCloud, Transfer quality) are aggregated in regular intervals into a site metric named 'Site Readiness'. The evaluation of the previous day happens the next morning based on the test results collected the previous day. The site readiness metric reflects the performance of the site. If 90% of SAM, Hammer Cloud, and transfer test were successful the site is considered good/ok.

Site Statuses

The site statuses for the day are set each morning after the Site Readiness is evaluated for the previous day.

- Life Status is the lowest of the three statuses and determines if a site is part of the CMS computing grid.
- Prod Status signals if a site is enabled for production (Tier-0, data processing, Monte Carlo workflows). Possible states: enabled, drain, disabled, test.
- Crab Status signals if a site is likely to run analysis programs successfully. Possible states: enabled, disabled.

	SITE_NAME																					
LifeStatus:	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Site Readiness:	0	E	0	0	0	E	E	E	0	0	0	0	0	0	0	0	0	0	0	0		
Maintenance:	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up		
HammerCloud:	90%	96%	90%	100%	99%	66%	50%	50%	100%	99%	88%	100%	100%	93%	97%	97%	100%	100%	93%	83%		
SAM Availability:	100%	77%	97%	96%	99%	93%	89%	100%	100%	100%	92%	100%	97%	100%	100%	100%	100%	100%	92%	89%		
Good T2 links from T1s:	15/16	15/16	14/16	16/16	16/16	15/16	15/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16		
Good T2 links to T1s:	13/16	14/16	10/16	13/16	14/16	14/16	14/16	13/16	13/16	13/16	13/16	14/16	14/16	12/16	14/16	14/16	13/16	13/16	14/16	15/16		
Active T2 links from T1s:	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14		
Active T2 links to T1s:	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14		
	19	20	21	22	23	24	25	26	27	28	29	30	31	01	02	03	04	05	06	07	08	09
	Aug									Sep												



Future improvements

- A new Pilot Startup Site Test (PSST) checks the functionality needed for successful CMS production and analysis job execution on a worker node before a job starts. It is a unified version of the worker node test from the GlideInWMS pilot, SAM, and HC tests.
- CMS plans to use the Computing Resource Information Catalog (CRIC) under development by WLCG to store experiment specific site/service information. The tool will replace the CMS SiteDB that did not keep up with grid development and changes.
- Success and failure information of user analysis and production jobs provides information about sites. Evaluation of this information is more complex than test jobs but the site/service use is also more complete and realistic than test jobs.