

## Experience with CMS Offline and Computing from Commissioning to Collisions

---

### Markus Klute for the CMS Collaboration\*

*Massachusetts Institute of Technology*

*E-mail: markus.klute@cern.ch*

In this presentation we will discuss the early experience with the CMS computing model from the last large scale challenge activities through the early collisions runs. Between the initial definition of the CMS Computing Model in 2004 and the start of high energy collisions in 2010, CMS exercised numerous scaling tests. We will discuss how those tests have helped prepare the experiment for operation. We will outline how the experiment operation has evolved during the first few months of operation. The current state of the Offline and Computing projects will be presented and we will describe the initial experience with active analysis users and real data. We will include Tier0 processing, reprocessing steps on data and fast turn-around calibrations. We will address the issues that worked well in addition to identifying areas where future development and refinement are needed.

*35th International Conference of High Energy Physics*

*July 22-28, 2010*

*Paris, France*

---

\*Speaker.

## 1. Introduction

The Large Hadron Collider (LHC) [1] started operation with first proton-proton collisions in November 2009. This marked the transition of the CMS [2] computing system from preparation to operation. This article will describe the experiences of the CMS offline and computing projects with the processing of collision data and the corresponding production of Monte-Carlo events.

The CMS computing model [3] has been designed as a world wide distributed, hierarchical structure of computing resources with well defined roles. CERN is at the inner core of this structure hosting the Tier-0 computing center. The Tier-0 receives collision data directly from the experiment. The incoming data is used to extract calibration and alignment information that can be utilized in a first reconstruction pass. Beside the role in promptly processing the incoming data, the Tier-0 hosts a safe copy of the RAW data and distributes the RAW and reconstructed data to the next layer in the computing system, the Tier-1 centers. CMS has seven Tier-1 centers located in France, Germany, Italy, Spain, Taiwan, the United Kingdom and the United States. Tier-1 host a custodial copy of the raw data, perform skimming workflows promptly on the incoming data. With improved software and detector understanding, the data and Monte-Carlo is re-reconstructed at Tier-1 centers. Remaining resources are used for the production of Monte-Carlo events. The Tier-1 centers serve data and Monte-Carlo alike to the Tier-2 computing centers. The Tier-2 computing resources are located at about 50 sites on four continents. Tier-2 centers host data for user analysis. A main activity on Tier-2 centers is the production of Monte-Carlo events for which about half the resources are allocated. The Tier-3 level resources are mainly used for user analysis.

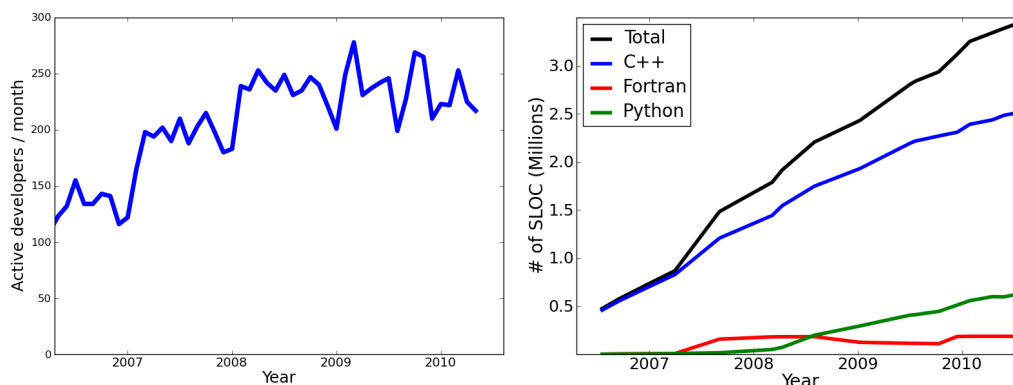
The CMS computing model is governed by a number of principles summarized below:

- CMS has two safe copies of collision RAW data on tape, with one copy at CERN and a second copy at a Tier-1 computing center.
- Resource planning foresees two or three re-reconstruction passes per year at Tier-1 sites of the complete set of recorded collision data and Monte Carlo.
- The number of generated Monte Carlo events produced by CMS matches the number of collision events recorded.
- Jobs are sent to the data.
- All CMS Tier-1, Tier-2 sites have good network connectivity to each other.

In preparation for LHC collision, the CMS computing model was tested in a number of computing challenges. Those challenges increased in scale and complexity in the years leading up to the first collision.

## 2. Offline Project

The offline project comprises a large number of sub-projects ranging from simulation and reconstruction to production tool development to name a few. Hundreds of software developers are working together to provide a stable software environment to the experiment while incorporating the latest developments, see Figure 1. An extensive program to optimize CMS code in terms of



**Figure 1:** Number of active CMS software developer per month (left) and number of lines of code (right).

CPU demand, memory consumption, data storage needs and maintainability has let to excellent performance of CMS software. Table 1 lists the processing time, event sizes and memory footprint as key performance parameters.

Workflow	Time	RECO size	AOD size	Memory usage
Data reconstruction	0.6 s/evt	400 kB	150 kB	900 MB
Top event simulation	90 s/evt	1400 kB	150 kB	980 MB

**Table 1:** Key performance numbers for CMS processing workflows for the first few months of data taking.

### 3. Computing Resources

Key ingredients for successful computing operation and the backbone of CMS computing are the world wide distributed computing facilities. The performance and availability of these resources are monitored centrally. A close relation between central operation and the facility representatives allows for prompt feedback. Leading up to the first collisions and throughout the first data taking campaign the readiness of the CMS sites was excellent. The resources deployed for CMS are listed in Table 2.

	CPU	Disk	Tape
Tier-0	55kHS06	3PB	9 PB
Tier-1	100kHS06	11PB	20 PB
Tier-2	192kHS06	12PB	

**Table 2:** CMS computing resources deployed in 2010.

### 4. Central Processing

CMS data is first written to a disk buffer located close to the experiment by the data acquisition

system. The typical rate of events is about 300 Hz. With the LHC duty cycle of 20-50% and a raw event size of less than 300 kB, the total raw data volume of CMS can be estimated. The raw data is automatically transferred via two 10 Gb links to disk pools at CERN, where it is repacked into primary datasets and the final CMS raw data format. Those datasets are archived on tape. The arrival of the data triggers a sequence of workflows. A small fraction of the data is extracted to measure alignment and calibration constants and to monitor the quality of the data within one hour of the data been taken. Prompt reconstruction jobs, which can already use the newly estimated detector and beam conditions, are launched with some time delay. The raw data and results of the prompt reconstruction jobs are transferred to the Tier-1 facilities. Subsequent workflows at the Tier-1's generate specialized datasets used for detector commissioning and analysis activities. With improved reconstruction software, calibration and alignment constants, the raw data can be re-reconstructed at the Tier-1's later. In the first four month of 7 TeV data taking, 10 such campaigns have successfully produced improved datasets for CMS. The corresponding Monte-Carlo has been re-reconstructed three times.

CMS matches the number of recorded collision events with simulated events. The workflows have been exercised for many years. A total of 64 computing sites participated in the production of Monte-Carlo events. The current capacity of CMS is a sustained production of 100 million Monte-Carlo event per week. In the first half of 2010 the Monte-Carlo production focused on so called "data-like" events, which where mostly minimum bias or low transverse momentum QCD samples with reflect the detector and beam conditions as closely as possible.

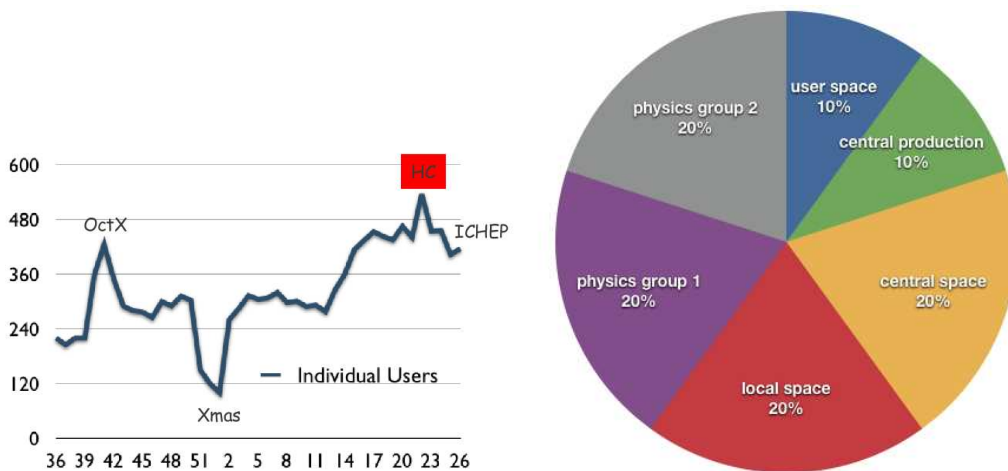
## **5. Data Distribution and Tier-2 Activities**

The distribution of data to the Tier-1 centers is organized centrally to balance the resource utilization while keeping in mind that CMS production jobs are executed at the site hosting the data. The storage space available at the Tier-2 facilities serves as a temporary buffer to the experiment. The data distribution to the Tier-2's is organized in a three layers and the storage is maintained by different user groups, see Figure 2. Data significant to the entire experiment is centrally distributed. This includes the collision data and main background Monte-Carlo samples that are needed by a large part of the community. Samples that are of interest to a smaller community are transferred by those groups, e.g. Higgs signal samples are transferred by the Higgs physics group to dedicated Tier-2 centers. And lastly, individual user can request the transfer of datasets to their local Tier-2 center.

A mandatory requirement for the success of the distributed CMS computing model is the reliable transfer of data. CMS established connectivity between all Tier-1 and all Tier-2 sites. This so called full meshed approach allows not only the transfer from all sites to each other without temporary storage at a third site. The transfer tool also re-balances itself because data sinks become sources as soon as a transfer is completed. The average total Tier-1 to Tier-2 transfer rate during the first three month of 7 TeV collision data taking was 560 MB/s with a maximum of 1407 MB/s. Transfers from Tier-2's averaged to 254 MB/s with a maximum of 853 MB/s.

The quick and reliable transfer of datasets to the Tier-2 centers is a prerequisite for distributed user analysis. The activities at Tier-2 centers are split between the central production of Monte

Carlo events and analysis activities. Up to five hundred individual user submit jobs using grid resources in CMS. Those user dominate the resource usages on Tier-2's, see Figure 2.



**Figure 2:** The typical storage breakdown of Tier-2 space in CMS (right). The number of active CMS grid user per week in 2009 and 2010 leading up ICHEP (left).

## 6. Conclusion

The CMS computing model was put to a test during the first four months of LHC 7 TeV operation. Detailed preparations of the system and the experience gained in numerous computing challenges paid dividends. Computing and offline projects have proven to be flexible enough to cope with the rapidly changing environment and demands. The 2010 data taking campaign has been completely successful and data was delivered with short latencies for first physics analysis. One of the big challenges for the coming years is operation of the computing model in a resource constraint environment.

## References

- [1] O. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, "LHC design report. Vol. I: The LHC main ring", *CERN-2004-003*
- [2] The CMS Collaboration, "CMS, the Compact Muon Solenoid: Technical proposal", *CERN-LHCC-94-38*
- [3] The CMS Collaboration "CMS Computing Technical Design Report", *CERN-LHCC-2005-023*