2007

US CMS Tier-2 Computing

Kenneth A. Bloom
University of Nebraska-Lincoln, kbloom2@unl.edu

Journal of Physics: Conference Series **119** (2008) 052004 doi:10.1088/1742-6596/119/5/052004

US CMS Tier-2 Computing

Kenneth Bloom

Department of Physics and Astronomy, University of Nebraska-Lincoln, Lincon, NE 68588-0111, USA

E-mail: kenbloom@unl.edu

Abstract. The CMS computing model relies heavily on the use of "Tier-2" computing centers. At LHC startup, the typical Tier-2 center in the United States will have 1 MSpecInt2K of CPU resources, 200 TB of disk for data storage, and a WAN connection of 10 Gbit/s. These centers will be the primary sites for the production of large-scale simulation samples and for the hosting of experiment data for user analysis – an interesting mix of experiment-controlled and user-controlled tasks. As a result, there are a wide range of services that must be deployed and commissioned at these centers, which are responsible for tasks such as dataset transfer, management of datasets, hosting of jobs submitted through Grid interfaces, and several varieties of monitoring. We discuss the development of the seven CMS Tier-2 computing centers in the United States, with a focus on recent operational performance and preparations for the start of data-taking in 2008.

1. Introduction

CMS will use a highly distributed computing system for data processing and analysis [1]. The Tier-0 center at CERN will do first-pass data processing, and immediately dispatch copies of the data to seven Tier-1 centers spread around the world, including one at Fermilab in the US. Tier-1 centers keep custodial copies of the data, re-process data when improved calibrations and algorithms are available, and create skimmed sub-samples of the data. There is no capacity for user analysis at Tier 1; all resources are committed to the above tasks. Analysis is deferred to the approximately 35 CMS Tier-2 sites, which will host the skims created at Tier 1 and analysis jobs running on those skims, and also generate simulated samples archived back to Tier-1 sites for use by the entire collaboration. The analysis resources are controlled by the user, and the simulation resources are centrally controlled by the experiment.

US CMS supports seven Tier-2 centers in the US, hosted by the universities indicated in Figure 1. The sites were designed to support the analysis work of about 40 physicists each, while also providing the required simulation capacity. When the LHC starts up, each site will have at least one million SpecInt2K of processing resources, 200 TB of disk, and 10 Gbit/s WAN bandwidth. (There are no tape resources.) About half of the CPU is dedicated to simulation, with the other half available for user analysis. Most of the disk will be occupied by data transferred from Tier-1 sites, or by user-generated data, with a small amount of disk needed for temporary caching of samples being simulated. The network bandwidth is sufficient to totally refresh the disk cache with new and/or improved dataets in under two weeks.

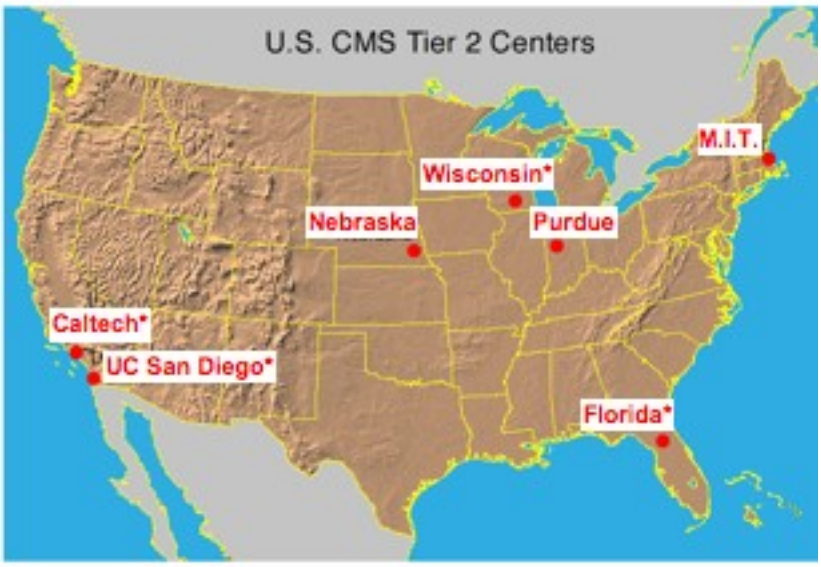

Figure 1. Figure caption

Table 1. Current capacity of US CMS Tier-2 sites.

Site	CPU (kSI2K)	Disk (TB)	WAN (Gbit/s)
Caltech	586	60	10
Florida	519	104	10
MIT	474	157	1
Nebraska	650	105	10
Purdue	743	184	10
UCSD	932	188	10
Wisconsin	547	110	10

2. Resources and Services

Table 1 shows the resources available at each US Tier-2 site. The sites are currently at about half the needed capacity for LHC startup, putting us on schedule for our planned facility ramp-up. Sites use the Scientific Linux 4 operating system. Most use Condor as a queuing system; some use PBS. Disk pools are managed using dCache [2], with SRM as an interface for remote data transfers. This provides a name space across storage devices, excellent I/O and load balancing. The Open Science Grid [3] software stack provides interfaces to the grid for remote submission of jobs; users generally do not have login privilege. Gratia [4] provides accounting services. The sites also run many CMS-specific services, such as PhEDEx [5] for wide-area data transfers, interfaces to the Dataset Bookkeping Service [6], Frontier squids [7] for caching database values, and the CMSSW software system [8] for data analysis and simulation.

3. Operations and Performance

The performance of the US Tier-2 sites can be measured relative to that of other CMS Tier-2 sites, and also relative to the benchmarks which have been established for the 2007

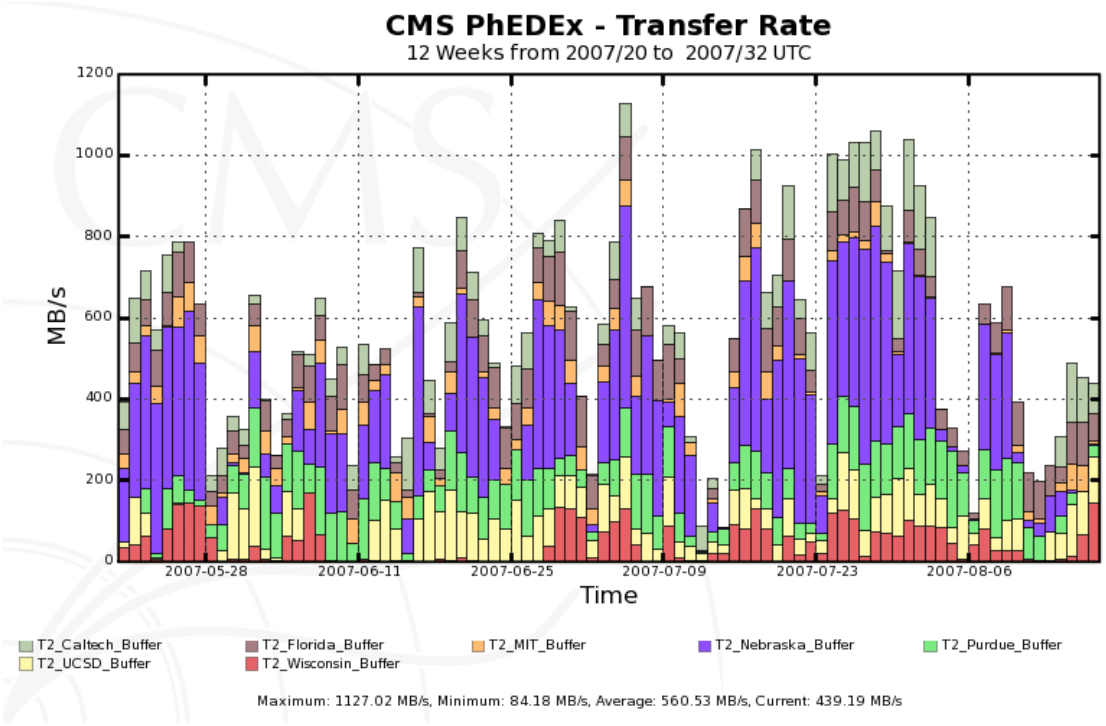

Figure 2. Daily average data transfer rates from CMS Tier-1 sites to US CMS Tier-2 sites for May 15-August 15, 2007.

CMS Computing, Software and Analysis Challenge (CSA07) that will start shortly after this conference, in which we will operate the full computing system at 50% of LHC-startup scale.

Tier-2 sites must transfer data from Tier-1 sites. Figure 2 shows the rate of data transfer achieved from seven Tier-1 sites to the US Tier-2; the rate is totally dominated by transfers from Fermilab. The average daily transfer rate per site is about 80 MB/s. The CMS-wide Tier-1 to Tier-2 transfer rate is 16 MB/s, putting the US sites well above average. The benchmark rate for CSA07 is 20-200 MB/s, which the US sites have reached for Fermilab; work continues on improving transfers from other Tier-1 sites.

Tier-2 sites must host user analysis jobs. Figure 3 shows the number of jobs hosted per three days by the US sites; the seven sites routinely host about 1K analysis jobs per day. Across all of CMS, the average number of analysis jobs hosted per day is 8K; the seven sites carry about their load among the $\simeq$40 world-wide sites. The benchmark for CSA07 is 75K jobs/day hosted across all Tier-2 sites; it is difficult to achieve this without a focused effort from users, but we have demonstrated that we can host jobs at that scale.

Finally, simulation is an important responsibility of Tier-2 sites. Figure 4 indicates the number of events generated by each US Tier-2 site in the past three months; the total is 66M events. During that same period, all CMS computing centers (including Tier-1 centers and CERN) combined produced 196M events. The CSA07 goal is 50M events/month at Tier-2 centers alone; the US Tier-2 sites are crucial for meeting the goal.

4. Conclusions

The CMS computing model relies on the success of computing sites around the world. The seven US CMS Tier-2 sites are fully integrated into CMS computing, and are performing as well as any CMS Tier-2 site. They have shown excellent performance in data transfers, job hosting and

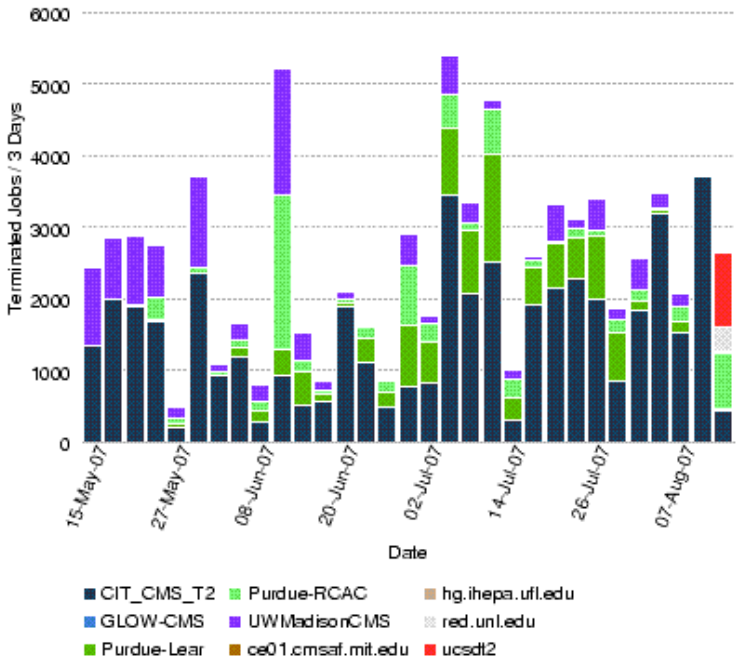

Figure 3. Number of analysis jobs hosted by US CMS Tier-2 sites in three-day intervals for May 15-August 15, 2007.

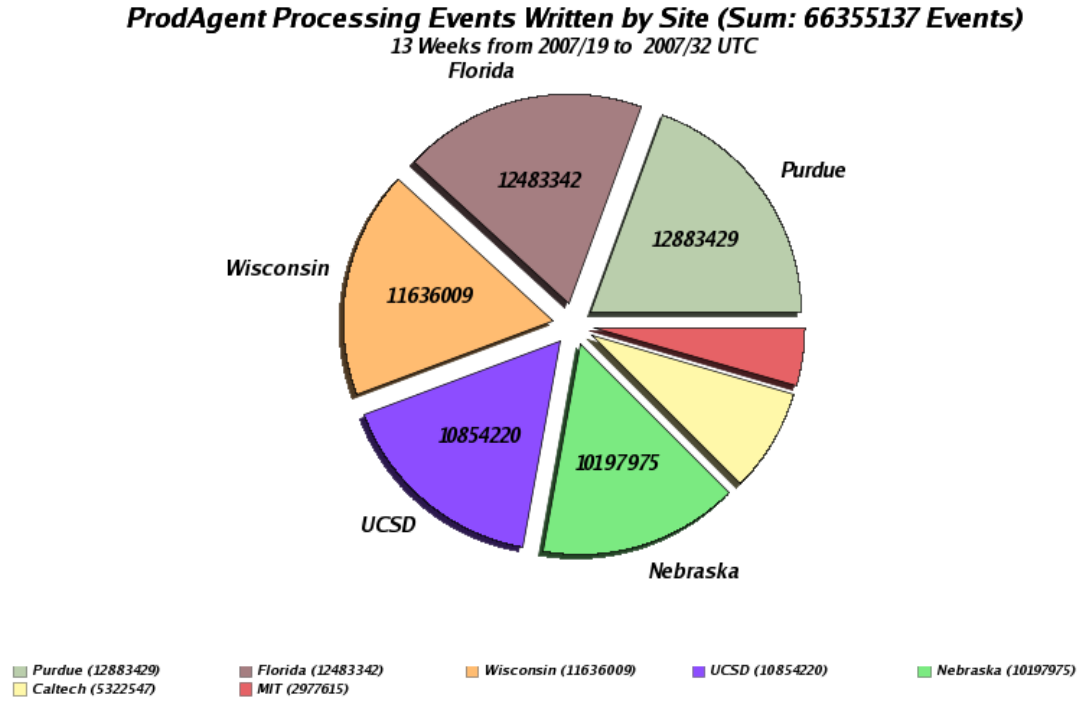

Figure 4. Number of simulated events produced by US CMS Tier-2 sites for May 15-August 15, 2007.

simulation. The upcoming CSA07 exercise will be an important test of site readiness for LHC startup, which is under a year away. But the successes that the US sites have shown so far bode well for our role in the discovery of new physics at the LHC.

[1] "CMS Computing Technical Design Report," CERN-LHCC-2005-023 and "CMS Computing Model," CERN-LHCC-2004-0035.
[2] `http://www.dcache.org`
[3] `http://www.opensciencegrid.org`
[4] "GRATIA – a resource accounting system for OSG," Philippe Canal, presentation #118, these proceedings.
[5] `http://cmsdoc.cern.ch/cms/aprom/phedex/`
[6] "The CMS Dataset Bookkeeping Service," Lee Lueking *et al.*, presentation #325, these proceedings.
[7] "CMS Conditions Data Access using FroNTier," Lee Lueking *et al.*, presentation #322, these proceedings.
[8] CMSSW reference