U.S. CMS Software and Computing

Progress Report for the 3d Quarter FY2004

Technical, financial and management status is reported for the period of March 31st to June 30, 2004.

2 Technical Status

U.S. CMS Software and Computing efforts were driven by the major milestones and activities for FY04:

- Supporting data analysis for the Physics TDR
- Preparation and demonstrations for the US-LHC Data Streaming Project.
- Rollout of the LCG Grid Service and federating it with the US facilities.
- LCG Interoperability

2.1 User Facilities subproject

2.1.1 Tier-1 Facilities

The bulk of the effort in this quarter was to ramp up the analysis and data serving capabilities of the Tier-1 center and increase the production capability to meet the targets set by CMS. This has required activity in all areas of the work breakdown structure.

In addition to increasing the facility scale and capacity, the User Facility group is currently in an "introspective" period. We are working to document and assess the lessons learned during the first CMS data challenge (DC04). The next major CMS data challenge is not scheduled until after the completion of the analysis for the Physics TDR, and much technical progress needs to be made in preparation.

2.1.1.1 WBS 1.1 Facilities

WBS item 1.1 is the User Facility activity assigned with developing facility services, ramping up computing capacity, and operating the CMS computing facilities. Progress was made this quarter in all task areas.

In the area of research and development, US-CMS developed and deployed a prototype storage element for use in DC04. The Storage Element is based on the Storage Resource Manager (SRM) for transfer management and a disk caching system, dCache, for managing local storage and serving data. The technology performed well during the data challenge. As a next step, US-CMS is working in collaboration with US-ATLAS and the LCG to perform a series of service challenges to demonstrate the performance and reliability of the SRM implementation. We will also demonstrate interoperability between the US SRM implementations with a dCache backend and the CERN SRM implementation on top of CASTOR.

In addition to the storage element development, US-CMS continued the collaboration with LCG to allow jobs submitted through LCG interfaces to run on resources available to interfaces installed for Grid2003. The goal is to allow one set of physical resources to be shared by multiple grids and the usage to be controlled by site policy.

In the area of procurement and ramping of facility capacity, US-CMS is expecting delivery of the FY2004 facility upgrades in July. The Tier-1 facility has been very heavily used both by US-CMS simulated event production and jobs submitted by Grid2003 users. US-CMS made heavy use of US-ATLAS facilities during the preparations for DC04 and US-ATLAS is utilization resources during the production phase of their data challenge (DC2).

In addition, US-CMS has moved forward the procurement of a tape silo, originally planned for FY05, to late FY04. The large increase in data volume from the DC04 simulation combined with increasing scarcity of tape slot space in the existing common silo prompted us to move the silo purchase forward. A change request to apply management reserve to fund this in FY04 was made and approved.

In the final WBS area of operations, the bulk of our effort has been spent improving the stability of the facilities and quality of service. US-CMS has seen a steady increase user account requests, dataset serving requests and facility services requests. The increase in users requires an increase of storage areas, CPU resources, and complexity of environments. The increase in the number of datasets to host at FNAL has increases the scale of the mass storage system requirements. All of these increases in demand and operational issues are valuable for the facility because they help us prepare for the beginning of the experiment, but it taxes our available operations staff.

2.1.1.2 WBS 1.2.1 Preparations for Analysis

After completion of DC04 end of April 2004, CMS has entered a year of intense analysis activity preparing for the Physics TDR. Therefore US-CMS needs to increase the analysis capability of the Tier-1 center and the services provided. US-CMS has seen a doubling of the number of active analysis users at the Tier-1 since last year. We expect to roughly double the number of active users every year until the start of the experiment.

There are three primary areas that US-CMS has been concentrating on to improve the capacity for analysis at the Tier-1 center. The first is dataset serving. The datasets created both for DC04 and in subsequent productions need to be transferred to FNAL, if they were not created there initially; prepared for analysis, which involves creating the meta data files required by the analysis framework; and finally served the data from mass storage to analysis applications. The second area is user services. The Tier-1 center is aiming to provide users with sufficient updateable space to work. The User Facility project has deployed a high available, high performance distributed file system to allow sufficient workspace for users. Commissioning of this system was started during this period of performance. The final area of development is the interface to computing resources. Initially US-CMS gave interactive login accounts to users who wished to work at the Tier-1. While we are still willing to do this, there has been a desire from some CMS members to enable grid submissions of analysis jobs. The User Facility is working to enable access to the data and metadata through the LCG interfaces.

2.1.2 WBS 1.3 Distributed Production Environment

The Distributed Production Environment effort is the US-CMS task to develop distributed computing infrastructure. The primary product of this task is the Distributed Processing Environment (DPE). The DPE has been used as the basis for CMS distributed event production on top of both in the dedicated US-CMS grid installations and Grid2003, a multi-virtual organization grid developed by US-CMS, US-ATLAS, and the US grid projects. This infrastructure was used to produce the bulk of the US contribution to the events produced for DC04. This quarter the DPE upgraded the version of the underlying middleware used and released a DPE version compatible with the latest Grid2003 software environment.

The Distributed Production Environment has also been developing more advanced submission and job tracking infrastructure. This is designed to reduce the operational load on the people handling CMS production and expand the number of people capable to managing production operations. This tracking and submission infrastructure builds on the production configuration infrastructure developed by US-CMS.

2.1.3 Distributed production and Tier-2 Facilities

The primary activities of the U.S. CMS prototype Tier-2 facilities during this reporting period were producing CMS and supporting facilities and grid interfaces to enable production operations for US-ATLAS and SDSS (Sloan Digital Sky Survey). The CMS physics groups have to prepare the analysis required to write the Physics TDR, which is due to be submitted in late 2005. The current estimate is that 10 million simulated events are required every month starting this fall and running for 12 months consecutively. The CMS production apparatus globally is working to ramp to this production level.

The other significant activity has been the support for Grid2003 interfaces and virtual organizations. US-ATLAS recently started the simulation for the DC2 data challenge. US-CMS had significant benefit from US-ATLAS resources during DC04 pre-production and we are happy to reciprocate. SDSS is working to analyze a larger dataset using distributed facilities. US-CMS is learning about service requirements for data intensive analysis applications from the experience supporting SDSS.

2.2 Core Application Software

2.2.1 Architecture Framework Toolkits and Reconstruction Software (WBS 2.1)

Bill Tanenbaum worked on a number of features in COBRA to aid in the reconstruction of missing or incomplete metadata. This work was especially useful in the light of DC04, where often the data was available before it was attached to the metadata. This adds flexibility to COBRA. He also continued work on COBRA to incorporate POOL collections. Collections can now be written. Work is ongoing to allow the collections to be read.

During this quarter, attention was turned to the CMS event data model (EDM), as one of the outputs of DC04 was the first attempt at a DST. Bill, along with Marc Paterno, worked on this. In addition, Marc Paterno started investigating the application of "run control parameters" developed by D0 to the CMS environment. This would be a kind of "meta-data" for the numerous parameters that control the execution of CMS programs such as ORCA and OSCAR.

Zhen Xie, in addition to continuing work on POOL catalogs, started the implementation of Relational database (RDBMS) interfaces in POOL. This is driven by CMS requirements to access relational data as C++ objects. An example is the access to the Conditions DB from within ORCA. The requirements collection started last year; implementation started in April. Several implementations exist by now, ranging from local disk, through middle range DB (e.g., MySQL), to enterprise scale (e.g., ORACLE). She has shown that POOL catalogs have good scalability and good performance.

Michael Case continued some development of DDD, based on user requests. The implementation of divisions within the XML geometry was completed. This should be the final piece of development code for DDD. The HCAL group has successfully used his examples to develop the geometry for the HCAL barrel sub-detector. Case has implemented geometries for the HCAL and EMU test beam setups for Summer 2004. All milestones have been met for the DDD.

He has also started work on the ORCA framework to read objects from the ConditionsDB. Initially, this will enable the test beam groups to access the ConditionsDB. This experience will be used in developing the CMS ConditionsDB.

Lassi Tuura has continued the development of the tools produced for DC04. It is clear, in the light of DC04, that data management will be a key issue for CMS. Lassi will continue to work in the context of the newly formed CMS CCS Data Management task (DM) as well as the cross-project Distributed Analysis (DAPROM) effort.

Lassi also attended an Apple developers workshop, as a first step in porting the CMS software to the Apple platform. Currently it is not possible, due to, among other things, shortcomings in the Apple dynamic loader. This should be rectified with the next release of the Apple MAC operating system, expected early next year.

Giulio Eulisse started work on a "virtual file system" to allow physicists an easy access to data and to analysis programs. A proof of concept was demonstrated at the May CPT meeting. This has caught the interest of CMS physicists, and he is now working with other developers in the DM task as well as the other newly formed CCS task, Workload Management (WM)

2.2.2 Visualization and Analysis (WBS 2.2)

Ianna Osborne implemented the visualization of DST objects in IGUANACMS, enabling physicists to see exactly what was contained in the DST. The muon visualization continued to be developed. A nice feature is that now the results of all of the muon reconstruction algorithms can be viewed, each with a different color. Work has started on the visualization of the output of the fast simulation program, FAMOS.

2.2.3 Distributed Computing Software (WBS 2.3)

Although McRunjob was used for DC04, it was determined that a more broadly useable product could be made available by using the re-factored kernel called ShahKar. Consequently, CMS McRunjob code has been moved into the new ShahKar code base. Anzar Afaq is now coding the replacement of the EVD services.

In the process of analyzing the data from DC04, work has been done on generalizing how the analysis will be done, and creating a process which will be easier for users to follow.

Documentation for these areas of change has been updated and improved.

2.2.4 Support

The CAS subproject continued with various support activities related to production and data challenge support. Also work continued on various software support tools. DAR and REFDBDAR continued to be developed.

CAS developers are working together with the LPC project to ensure that CMS software will be readily useable by physicists working at the LPC. Work is ongoing to develop procedures to release CMS software at FNAL within 24 hours of its release at CERN.

Numerous tutorials and presentations have been given by several of the CAS engineers during this quarter. Also, much work has been performed on software documentation. Large parts of the CMS code has been examined, correcting grammar and spelling mistakes in error messages. Although this might seem trivial, it does help the end-user understand what might be going wrong.

3 Financial Status

3.1 FY04 Budget and Allocation

The funding guidance for FY04 was revised in March 2004, and then restored to the prior level in mid-June. This means the FY04 funding allocation from NSF is effectively arriving 9 months late.

The following table shows the total FY04 funding allocation, and the budgeted funding and effort.

FY04 Funds and Costs	FY03 total	**FY04 total**	FNAL	Universities
Funding Allocated				
CAS Personnel	$1,370k	**$1,498k**	$425k	$1,074k
UF Personnel	$2,016k	**$3,306k**	$2,123k	$1,183k
Tier-1 Equipment	$550k	**$963k**	$963k	$0k
Tier-2 Equipment	$192k	**$187k**		$187k
Edge Computing		**$305k**	$305k	$0k
Project Office, Reserve	$265k	**$667k**	$667k	$0k
Total Allocated	$4,393k	**$6,926k**	**$4,482k**	**$2,444k**
Effort Allocated				
CAS FTE years	10	**11**	3	8
UF FTE years	19.7	**25**	15	10

The following table shows the actual effort and costs incurred in this quarter, as far as data was available.

FY04 Funds and Costs	FY03 total	FY04 total	FNAL	Universities
Total Allocated	$4,393k	**$6,926k**	**$4,482k**	**$2,444k**
USCMSSC DOE Funds	$3,115k	**$5,631k**	$4,432k	$1,199k
USCMSSC NSF Funds	$750k	**$750k**	$50k	$700k
iVDGL Funds	$528k	**$545k**		$545k
Effort Allocated				
CAS FTE years	10	**11**	3	8
UF FTE years	19.7	**25**	15	10
Effort Spent (first 9 months)				
CAS FTE years		**14.4**	2.4	6.0
UF FTE years		**21.8**	7.0	7.4
ACWP in k$, as invoiced				
CAS Labor		**$1,593k**	$367k	$613k
UF Labor		**$1,440k**	$1,097k	$171k
T1 Equipment		**$1,065k**	$1,065k	$0k
T2 Equipment		**$0k**	$0k	$0k
Project Office		**$140k**	$140k	$0k
Total		**$4,238k**	**$2,669k**	**$785k**

4 Management Status

The project successfully passed the DOE/NSF mini review in July.

DOE funding guidance for Tier-1, Grid and CAS efforts stays about stable for FY05. It was not until June that NSF guidance was firmed up. Thus for FY04 the current effort funded through the NSF stays constant, providing CAS effort and one FTE system management effort at each of the US CMS

prototype Tier-2 centers, Caltech, UCSD and U.Florida. This was approved by project oversight in the PMG meeting.

Management recognizes that the Tier-2 program will need to be scoped out now that funding profiles are clear, due to the pressing time scales of data challenges and the need to be ready for physics in 2007. The project has started a process through the ASCB to get feedback from the collaboration on the requirements and expectations for US CMS Tier-2 centers. An ASCB retreat in June defined these issues, and a bidding process has started for Tier-2 centers at the end of FY04. The first US CMS pilot Tier-2 center can be started early in FY05.

US CMS is working with PPDG and the trillium projects to define technical solutions and an architecture for the Open Science Grid. In the Trillium project Joint Steering meeting the organizational structure for these activities was defined. Technical Groups were started for storage and security, and another one for operations support will start soon. The initial Open Science Grid deployment milestone was scoped out. OSG will be a production Grid as a follow-up for Grid3 that has additional functionality. For example, OSG will have storage services and a fine-grained authorization infrastructure. To come to a coherent architecture for the Open Science Grid, a Blueprint activity was started, in which US CMS is participating.

4.1 Meetings and Reviews

The following USCMSSC Program Management Group (PMG) meetings were held:

February 13, 2004, see http://www.uscms.org/scpages/projectoffice/pmg/2004-02-13/
April 9, 2004, see http://www.uscms.org/scpages/projectoffice/pmg/2004-04-09/
July 23, 2004, see http://www.uscms.org/scpages/projectoffice/pmg/2004-07-23/

The following DOE/NSF review was held:

July 8, 2004, see http://www.uscms.org/s&c/reviews/doe-nsf/2004-07/

4.2 Documents

FY04 MOU/SOW with Minnesota for DOE funded effort complete; PO complete
FY04 SOW with Wisconsin signed, PO complete
FY04 SOW with Caltech signed; PO complete
FY04 SOW with Princeton signed; PO complete
FY04 SOW for UCSD for iVDGL funds ready but not signed
FY04 SOW for Florida and Caltech for iVDGL funds waiting deliverables and dates

FY04 SOW for UCD (CAS) signed; PO complete.
MOU between iVDGL and USCMSSC is being signed
MOU/SOW for 9 months NFS money with UCSD, UFL, Caltech and NEU
Change Request: Transfer funds from Management Reserve to T1 Equipment