



# **U.S. CMS Software and Computing**

## **Progress Report for the 4th Quarter FY2004**

Technical, financial and management status is reported for the period of July 1<sup>st</sup> to September 30<sup>th</sup>, 2004.

### **1 Technical Status**

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

- Preparing for and 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
- Defining the work plan towards the fully functional computing and software systems

#### **1.1 User Facilities subproject**

##### **1.1.1 Tier-1 Facilities**

This quarter the User Facility subproject continued 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 continues to influence all areas of the work breakdown structure.

###### **1.1.1.1 WBS 1.1 Facilities**

WBS item 1.1 is the User Facility activity for 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 has been participating in a series of service challenges in collaboration with US-ATLAS and the LCG community. The first challenge was to demonstrate the reliability and performance of the Storage Resource Manager (SRM). SRM is the basis for the interface to all Storage Element (SE) development work. The SE is a fundamental grid component. The service challenge started this quarter. The transfer rates are within expectations and the challenge has identified some reliability issues to be investigated and repaired. Currently the total throughput is limited by the availability of disk servers. We expect higher scale tests during the next quarter.

Also in the area of research and development, US-CMS has procured a small group of 64 bit processors for evaluation purposes. To stay on the technology curve CMS will have to master the transition to 64-bit architectures – including addressing the software and systems compatibility issues. The procured systems represent the next generation of processing elements and are expected to be more similar to the systems we will procure for the start of the experiment.

In the area of procurement and ramping of facility capacity, US-CMS commissioned and validated the FY2004 procurements during July. There are 100 new dual systems that represent a doubling of the total Tier-1 capacity, which is what was called for in the long-term procurement plan. The Tier-1 facility has been very heavily used both by US-CMS simulated event production and jobs submitted by Grid2003 users.

The 100 new nodes are also serving as the hardware for our grid interoperability work with the LCG. The new systems have been configured to appear as resources in the LCG. The computing resources and information providers report correctly. These systems will also be configured to appear to the US grid infrastructure. Once the systems appear to both grids we will have successfully completed the interoperability milestone.

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. The silo was delivered during the final days of Q4 and is expected to be installed in the first few weeks of FY05.

The final element of the Facilities WBS is operations. The facility continues to efficiently produce events for CMS and for US-ATLAS. During DC04 CMS made heavy use of resources operated by US-ATLAS. ATLAS is in the middle of DC2, their second data challenge. Significant use of US-CMS Tier1 and Tier-2 resources was achieved with grid-submitted ATLAS jobs.

### **1.1.1.2 WBS 1.2.1 Preparations for Analysis**

After completion of DC04 at the 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.

In last quarter's report we reported on our improvements in data serving, user services, and enabling the facility for grid submitted analysis jobs.

The data serving capabilities continue to improve with more data sets available. We see an increasing number of users accessing the service and running analysis jobs. This is both driven by the needs of the new LHC Physics Center, but also from people doing physics studies on data sets hosted at the Fermilab Tier-1 center from Universities from throughout the U.S. and also other countries. The data serving infrastructure requires operational support, but is responding well to the current load.

In the area of user services, the User Analysis Facility grew by 50% in response to heavy user load on the system. The facility's team is currently investigating how to enable users to submit to the larger simulation systems for some types of analysis applications.

### **1.1.2 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. The DPE was updated this quarter to the current version of the underlying middleware components. The changes allow us to stay current with the US and LCG grid efforts as well as allow us to take advantage of improvements and fixes in the new software releases.

### **1.1.3 Tier-2 Facilities**

The two primary activities of the US-CMS Tier-2 facilities were producing simulated events and the preparations for the first round of Tier-2 center selections.

The three current prototype centers, at Caltech, UCSD, and UFL, continued to provide significant computing resources to US-CMS and to other virtual organizations submitting through the Grid3 interfaces. US-ATLAS, as part of their Data Challenge DC2, is making significant opportunistic use of CMS computing resources through Grid3.

US-CMS is also preparing for the first selection of permanent Tier-2 centers. There was a workshop hosted by the Advisory Software and Computing Board (ASCB) last quarter at Caltech to define the scope and capacity of a US Tier-2 center. Early this quarter a request for proposals was circulated by the US CMS S&C project manager and responses are expected by the end of October.

## **1.2 Core Application Software**

### **1.2.1 Architecture Framework Toolkits and Reconstruction Software (WBS 2.1)**

During this quarter, much of the effort of Tanebaum, Paterno and Brown was concentrated on the COBRA/ORCA/ROOT interface to evaluate the requested functionality of access to the ORCA data from the ROOT prompt, or access to ROOT from a COBRA prompt. A series of notes were written to communicate the use cases and approaches, and a review of the framework in view of RunII experience was started, on initiative of the LPC group. It appears that with the current versions of software these physics drive use cases are not straight-forward to be implemented. A survey of what best practices from the current D0/CDF collaborations frameworks has been made and a set of specifications has been drawn up.

Tanebaum collaborated with LPC physicists to develop use cases for the event data model for the LPC.

Zhen Xie continued work on the implementation of Relational database (RDBMS) interfaces in POOL called Relational Access Layer (RAL). She is worked on the internal CMS interfaces in COBRA. She provided the CMS requirements on distributed databases to the LCG -3D (Distributed Deployment of Databases) group. She has continued work on the POOL Filecatalog. She has also worked on optimizing CMS scripts that access database information.

Michael Case continued work on DDD, mainly bug fixes for parts of the code that had not been used until now. DDD also required minor modifications to support the most recent COBRA release. The bulk of his time was spent on Conditions DB work. He has started working with the POOL RAL, developed by Zhen Xie. A large fraction of his time this quarter was spent discussing with DB users and developers to better understand the requirements of the Conditions DB work.

Lassi Tuura has continued the development of the data transfer tools, originally produced for DC04. This is becoming part of the new Data Management are. The Transfer Management DataBase has been cleaned up and is being developed into a new product called Phedex. This service is being used

for data transfers and data set placement at CMS regional centers, including Fermilab. This is now in its second release.

Giulio Eulisse started work on a “virtual file system” as an interface for physicists to easily access data and analysis programs. The system is called Physh. He is working together with other people in the CCS Data Management task to develop user tools to easily publish data. This interfaces into the PubDB work led by Wildish (see below).

### **1.2.2 Visualization and Analysis (WBS 2.2)**

Ianna Osborne continued development of IGUANA. Parts of the code dealing with DST and Digis have been separated to allow easier maintenance. A new version of the Lego displays has been incorporated. With the new CCS management structure, it was decided to break off some portions of IGUANACMS into the new APROM (Analysis PROject Management; the CCS-PRS cross-project effort on analysis).

### **1.2.3 Distributed Computing Software (WBS 2.3)**

LCG-2 support in MCrunjob was re-implemented, and Grid production jobs have been successfully run on the LCG-2 infrastructure.

With the CCS reorganization, the Production task has been split into two: development, managed by Tony Wildish, and Operations, managed by an off-project group in Bari, Italy. This has allowed Wildish to devote more time to development. In the past quarter, he has developed a new database, PubDB, that allows the easy publishing of data catalogues. PubDB is also be used by the APROM developers as part of the data discovery effort.

The infrastructure is being put in place to be able to process 10 million events per month, to be able to deliver the needed data for the PRS groups in advance of the Physics TDR. The P-TDR is scheduled to be published by the end of 2005.

### **1.2.4 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.

## **2 Financial Status**

### **2.1 FY04 Budget and Allocation**

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.

A change request for applying \$400k of management reserve to Tier-1 equipment procurements was approved by project oversight. Those funds are for overruns in the costs for networking equipment, for adding additional storage to host Higgs and SUSY data sets, as requested by the PRS groups, and to procure a tape robot a few months earlier than planned, already in FY04, to address the shortage of storage slots in the current shared tape library.

The following table shows the total FY04 funding allocation, and the budgeted funding and effort. It should be noted that for the Tier-1 equipment budget, items from the FY03 procurement plan, which had been obligated in FY03 actually were paid for only in FY04. These items totaled to \$360k, while the carry-over in equipment funds was only \$245k. So FY03 had an overrun in costs of about \$115k or about 10% of the Tier-1 equipment budget, which carried over into the FY04 numbers.

<b>FY04 Funds and Costs</b>	<b>FY03 total</b>	<b>FY04 total</b>	<b>FNAL</b>	<b>Universities</b>
<b>Funding Allocated</b>				
CAS Personnel	\$1,370k	<b>\$1,498k</b>	\$425k	\$1,074k
UF Personnel	\$2,016k	<b>\$3,306k</b>	\$2,123k	\$1,183k
Tier-1 Equipment	\$550k	<b>\$1,345k</b>	\$1,345k	\$0k
Tier-2 Equipment	\$192k	<b>\$187k</b>		\$187k
Edge Computing		<b>\$305k</b>	\$305k	\$0k
Project Office, Reserve	\$265k	<b>\$667k</b>	\$667k	\$0k
<b>Total Allocated</b>	<b>\$4,393k</b>	<b>\$7,308k</b>	<b>\$4,864k</b>	<b>\$2,444k</b>
<b>Effort Allocated</b>				
CAS FTE years	10	<b>11</b>	3	8
UF FTE years	19.7	<b>25</b>	15	10
<b>Effort Spent (first 9 months)</b>				
CAS FTE years		<b>14.4</b>	2.4	6.0
UF FTE years		<b>21.8</b>	7.0	7.4

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

<b>Total Allocated</b>	\$4,393k	<b>\$7,308k</b>	<b>\$4,864k</b>	<b>\$2,444k</b>
USCMSSC DOE Funds	\$3,115k	<b>\$6,013k</b>	\$4,814k	\$1,199k
USCMSSC NSF Funds	\$750k	<b>\$750k</b>	\$50k	\$700k
iVDGL Funds	\$528k	<b>\$545k</b>		\$545k
<b>Effort Allocated</b>				
CAS FTE years	10	<b>11</b>	3	8
UF FTE years	19.7	<b>25</b>	15	10
<b>Effort Spent (12 months)</b>				
CAS FTE years		<b>11.7</b>	3.7	8.0
UF FTE years		<b>19.6</b>	9.7	9.9
<b>ACWP in k\$, as invoiced</b>				
CAS Labor		<b>\$1,273k</b>	\$622k	\$651k
UF Labor		<b>\$1,730k</b>	\$1,548k	\$182k
T1 Equipment		<b>\$1,508k</b>	\$1,508k	\$0k
T2 Equipment		<b>\$0k</b>	\$0k	\$0k
Project Office		<b>\$180k</b>	\$180k	\$0k
<b>Total</b>		<b>\$4,691k</b>	<b>\$3,858k</b>	<b>\$833k</b>

### **3 Management Status**

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

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 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.

A CMS Requirements and Technical Assessment process was successfully conducted with the help of US CMS, to define the requirements and approaches to data management and data organization. A report was submitted to and accepted by CMS CPT Joint Technical Board. A Data Management task has been formed as part of CCS, and Lothar Bauerdick has been asked to coordinate and to help defining the scope of this activity.

#### ***3.1 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/>

### ***3.2 Documents***

MOU/SOW for 9 months NFS money with UCSD, UFL, Caltech and NEU completed