# Report of the First Review by the

# U.S. CMS Software and Computing Oversight Panel

**C. Boeheim, E. Blucher (chair), J. Branson, J. Butler, P. Mato, D. Petravick, W. von Rüden, T. Wenaus**

## Fermilab, October 25, 2000

## Table of Contents

# 1. Executive Summary

This report contains the results of the first Review of the U.S. CMS Plans for CMS Computing by the U.S. CMS Software and Computing Oversight Panel (SCOP). It is based on initial Project Management Plan received before the meeting, and on presentations and discussions during the meeting at Fermilab.

The committee was impressed by the presentations and with the progress of the project. The goals and scope of the project are clearly defined, and the participants seem focused on allowing U.S. physicists to participate as effectively as possible in CMS.

The Software and Computing project is divided into two subprojects: Core Application Software and User Facility. The U.S. CMS Core Application Software effort appears to be extremely well integrated into International CMS, with the U.S. playing a leading role in many areas. There are some staffing concerns, but in general, the work seems to be progressing as planned. The User Facility subproject is also making good progress. The plan for a Tier 1 computing facility at Fermilab is well developed. The Tier 2 computing model is less well developed and will require attention to be sure it is a successful component of the computing effort. The committee believes that the U.S. CMS S&C project management should take a leading role in defining the mission of the Tier 2 centers; the project management must provide oversight for the R&D required for effective implementation of the Tier 2 centers. In addition, coordination between the Tier 0 center at CERN and Tier 1 centers around the world is needed to insure that the distributed resources can be used effectively.

Several recommendations appear in the text of the report, and are listed below.

- Fermilab should try to create a nucleus for software and physics activity in the U.S. by supporting CMS software and physics activity at the lab. It should also create a physical center for CMS software and physics activity. Proximity of this site to the CMS physicist's offices and to the CMS software and computing staff offices would be of great benefit.

- Management should take into account the likelihood of less than full staffing of the S&C project, and have contingencies if key positions are not filled. Turnover should be anticipated by having good written records and documented procedures to allow new people to step into roles with as little ramp up time as possible.

- The committee feels that the S&C Level 1 manager should be given as much flexibility as possible to optimize the funding profile among the different subprojects. The committee recommends that the S&C project management work with Fermilab and the funding agencies to develop a plan to deal with short-term budget shortfalls.

- We recommend that Fermilab, perhaps in conjunction with the U.S. ATLAS Tier1 center and other Tier 1 Centers, open discussions with CERN on mechanisms for addressing the relevant issues of hardware, software, and networking for LHC computing

- We encourage the head of User Facilities and the U.S. CMS S&C Project manager to take a leadership role in defining the mission of the Tier 2 centers, supporting their selection and construction, and developing the appropriate coordination mechanisms and support agreements.

- The committee encourages Fermilab to join and actively participate in the Geant4 project in ways that will advance the CMS simulation effort.

- We encourage the group to investigate possible common projects between the U.S. CMS and U.S. ATLAS.

## 2. Introduction

The U.S. CMS Project Management and Fermilab have formed an oversight panel (U.S. CMS Software and Computing Oversight Panel - SCOP) to provide advice to the U.S. CMS Project Manager for Software and Computing. The main topics to be considered are

- project management, including integration into International CMS and responsiveness to needs of U.S. CMS;
- Core Application Software;
- User Facilities.

The complete charge to the committee is given in Appendix A.

The committee members are:

E. Blucher, University of Chicago, chair

C. Boeheim, SLAC

J. Branson, UCSD

J. Butler, FNAL

P. Mato, CERN

D. Petravick, FNAL

W. von Rüden, CERN, ex-officio as Chair of the corresponding U.S. ATLAS committee

T. Wenaus, BNL

The committee, which expects to meet every 6 months, held its first meeting at FNAL on Oct. 23, 24, and 25 (the agenda of this meeting is given in Appendix B). The committee's comments from this first meeting are summarized in the following sections.

# 3. Software and Computing Project

## 3.1 Project Management

The Project Management plan is now relatively stable, and clearly defines the project scope and organization. The roles of the principal leaders of the project, the Level 1 Project Manager, and the two Level 2 Project Managers are well defined. The boundaries between the two major sub-projects are also well defined and should allow the two projects to develop with reasonable independence but with enough coordination to insure success.

The relationship between the project and the Fermilab Computing Division is also spelled out and seems workable. The scope of the software and computing project is limited; it is clear where it ends and where the physics software activities begin. In order to have a successful implementation of the analysis software, it is important that the scientific code be written in manner that uses resources efficiently and is easy to maintain over a long period of time. The project provides for expert programming support to help the writers of the scientific code achieve this objective.

The Committee is pleased to note that the Level 1 Project Manager has been appointed and will soon arrive and assume responsibility for the project. This will free the head of the Fermilab Computing Division to carry out his crucial oversight role without the added burden of functioning as the Acting Level 1 Project Manager. We look forward to the appointment of the permanent Level 2 project managers very soon.

Staff turnover and competition for staff with industry are realities of the current climate. Management should take into account the likelihood of less than full staffing of the project, and have contingencies if key positions are not filled. Turnover should be anticipated by having good written records and documented procedures to allow new people to step into roles with as little ramp up time as possible.

The planned staffing of the Project Office appears appropriate. To fully justify the staffing plan, we expect the Level 1 Project Manager to clearly define the roles of this staff and the division of responsibility between the Project Office and the lower level managers.

## 3.2 Needs of U.S. CMS

The Software and Computing Project should meet the needs of U.S. physicists to analyze data and produce physics results. It should also enable them to fulfil their detector and calibration responsibilities. The current project is generally designed to perform these tasks. From the presentations by the project managers, it was clear that they are committed to serving the needs of U.S. CMS. The addition of a fully committed Level 1 project manager is also a positive development.

The project primarily receives advice on how to satisfy the collaboration's needs through the Advisory Software and Computing Board. The broad representation of the collaboration on this board should provide a good way to keep the project on track. In addition, the current board has the expertise necessary to properly advise the project. We think that the project is well organized to serve the needs of U.S. CMS and that the current plan will allow it to succeed in this goal.

One important aspect of the User Facility at Fermilab is that it creates a nucleus for software and physics activity in the U.S**.** Fermilab should do what it can to help create this nucleus by fostering CMS software and physics activity at the lab. It should also create a physical center around which CMS software and physics activity may nucleate. Proximity of this site to the CMS physicist's offices and to the CMS software and computing staff offices would be of great benefit.

### 3.3 Connection to CMS

Another important aim of the Software and Computing Project is to cooperate with CMS as a whole to create software and computing tools and to do production computing. It is important to insure that the project is well integrated into CMS so that duplication and waste can be avoided, and so that CMS software and computing can be properly managed.

From many presentations of the current work and plans, we are convinced that this project is currently well integrated into CMS and expect that it will continue to be, at least over the next few years. Matthias Kasemann explained the connections to CMS. First, U.S. CMS has a structure parallel to CMS as a whole and this project fits into those structures in the same way as the successful U.S. CMS construction project does. Second, the U.S. is well represented in the CMS management. Third, the Level 1 project manager will be a member of all the relevant CMS software and computing boards. Fourth, the current members of the U.S. ASCB are deeply involved in CMS software and computing and hold important positions there.

We are convinced that the project management plan will assure good synchronization with CMS as a whole and that the current project plan is well designed to serve the needs of CMS. In particular, the Core Software Project is clearly well integrated into the CMS Core Software Project, with the U.S. taking the lead in many areas. There are also many technical channels open between U.S. CMS and CMS as a whole and a promising start between Fermilab and CERN.

The Tier1 Centers, one of which is at Fermilab, collectively represent a key resource for the CMS collaboration. It is very important for International CMS to coordinate the activities of these centers so that they form part of an integrated and reasonably transparent system for CMS data analysis. This coordination role should begin as soon as possible so that work now starting at Fermilab and elsewhere does not have to be redone.

Similarly, since the Tier 1 Centers must interoperate with the Tier 0 Center at CERN, we encourage CERN to define as quickly as possible the resources that it will bring to bear on the CMS reconstruction and data analysis since that is an important consideration for planning the Regional Centers. We also encourage CERN to begin immediately to provide coordination and guidance for the Regional Centers, which includes those for all four LHC experiments, so that that Centers can begin to understand how what they need to do to properly interoperate with the Tier 0 center at CERN.

### 3.4 Budget

Although the committee did not attempt a detailed review of the proposed budget, the presentations showed that reasonable care has gone into setting the requested funding levels. For the CAS project, the support level is based on a 25% contribution to the overall CMS CAS project. For the User Facility, personnel requirements were estimated with two independent methods that gave similar results. Hardware requirements for the User Facility are based on estimates from the MONARC study and the Hoffman review at CERN; hardware costs assume

(perhaps optimistically) that "Moore's Law" extrapolations will be valid for all CPU and data storage.

The budget includes at 10% management reserve; the committee felt that this level of contingency was justified. Budget planning is complicated by uncertainty in the NSF budget and profile. Based on the current expected DOE and NSF funding, there is a shortfall of close to 20% in the overall S&C project. In making the case for the funding agencies to close this gap, it would be useful for the S&C group to understand how the project would have to be descoped or delayed if additional funds do not become available. The networking cost associated with Tier 2 centers (~6 million dollars between 2001-2006) is particularly uncertain, and might be reduced significantly by choosing Tier 2 sites that already have good network connections. The current budget profile also indicates a particularly large deficit in 2004. Delaying some Tier-2 hardware purchases until closer to LHC turn on may mitigate this profile problem. The committee feels that the Level 1 manager should be given as much flexibility as possible to optimize the funding profile among the different subprojects.

There are also significant budget shortfalls in the first two years of the plan**.** The committee recommends that the S&C project management work with Fermilab and the funding agencies to develop a plan to deal with these shortfalls.

## 4. Core Application Software

The presentation of the CAS sub-project included a great deal of detail and demonstrated a clear understanding of the project's objectives. The project planning appears to be quite advanced. The scope of CAS is well balanced and covers most of the major areas in core software development for International CMS. In addition, the fact that U.S. CMS has leadership roles in some of the projects in CMS Core Software is a very good sign of the health and fulfillment of its goals. The participation in the software architecture design and documentation is very positive and will certainly have long term benefits. The group's strong presence at CERN, which it plans to sustain, is an important asset.

The rolling planning process, with fine-grained planning for the near future and coarse-grained planning for activities in the more distant future is appropriate. The development process should be iterative and incremental, so the rolling approach to planning fits very well.

Care has to be taken in avoiding duplication of software development efforts between the UF and CAS, in particular in areas close to user support for software development, which appear in both WBSs. The structure of the U.S. CMS Software and Computing with the Level 1 project manager overseeing the two sub-projects should be sufficient to insure that duplication of effort does not occur in the future.

One area of concern was the status of Geant4. The Geant4 based simulation has been progressing with U.S. CMS involvement but is behind schedule, and could soon be on the critical path. U.S. CMS is increasing the effort on Geant4 to help rectify this, but the committee feels that an important additional contributor could be Fermilab itself. The committee encourages Fermilab to join and actively participate in the Geant4 project in ways that will advance the CMS simulation effort.

The committee noted that CAS plans include no mention of the possibility of collaboration between the core software development being done in U.S. ATLAS and U.S. CMS. Although the "grid" participation will induce some kind of collaboration, investigation of possible common projects between the two collaborations should be encouraged. Perhaps this investigation should be driven at the level of the international collaboration, but the U.S. projects have a direct interest in exploring this possibility.

The required manpower resources for CAS (increasing from 9 FTE in 2000 to 13 FTE in 2005), although not deduced from a resource loaded WBS, is based on a 25% contribution to the overall CMS CAS project. The approach taken by CMS, preferring the commitment of a certain "level-of-effort" for software development rather than commitments to deliver final products, maintains flexibility and seems reasonable.

## 5. User Facility

The committee believes that the plans for the Tier 1 computing facility at Fermilab are well developed. It is well understood by the Fermilab staff how to plan for, procure, and deploy a computing system of the size planned for CMS. The development timeline with respect to the CMS schedule at CERN is realistic, and they have already been adjusting the Fermilab schedule to mesh with changes in the CMS schedule.

While the scale of the data is larger than that handled to date by Fermilab, it is not so much larger that their methods would not extend to cover it. Being a remote site for data generated elsewhere is a new role, but planning seems well along for that. The exact modes of transporting the data from CERN to Fermilab are still vague, but that is appropriate given the developments in networking that are likely to happen between now and first data. We note that Fermilab's preparations are based on data volumes and processing speeds predicted by the CMS collaboration, and inaccuracies in those estimates would necessitate changes in the Fermilab plans. Fermilab staff should follow developments in the CMS computing model and size estimates closely.

The Tier 2 Centers represent a new element of distributed computing in HEP. The success of the Tier 2 model depends in part on Grid tools that are not yet delivered. Flexibility needs to be maintained while these tools are developed. Without adequate Grid software, it may be difficult to achieve a reasonable level of support for the Tier 2 sites with the currently assumed support level of only 1.5 FTE per site; additional support from Fermilab might be required. The greater the uniformity of the Tier 2 sites in both setup and scale, the more this work may be reduced. The assumption of an ODBMS at all levels down through Tier 3 has been very difficult for current experiments to achieve, and will require significant work. While this work is more appropriately the responsibility of the CMS Collaboration, the outcome will affect the computing model at the Tier 1 and 2 sites.

We encourage the head of User Facilities and the U.S. CMS S&C Project manager to take a leadership role in defining the mission of the Tier 2 centers, supporting their selection and construction, and developing the appropriate coordination mechanisms and support agreements. This must start soon if the Tier 1 and Tier 2 centers are to form an effective and well-integrated component of CMS computing. R&D towards the implementation of Tier 2 centers should be conducted so as to achieve the intended final goal and should have strong input and oversight from the U.S. CMS S&C project through the manager of the User Facility Subproject.

## 6. Conclusions

The organization and progress to date of the Software and Computing project impressed the committee. TheWBS is established, first resource-loaded schedules exist, and there appears to be a lot of enthusiasm for the project.

Since the project is still at an early phase, many important decisions remain. Nevertheless, the committee believes that the basic strategy for the project is correctly established. Several recommendations appear in the text of the report, and are repeated below.

- Fermilab should try to create a nucleus for software and physics activity in the U.S. by supporting CMS software and physics activity at the lab. It should also create a physical center for CMS software and physics activity. Proximity of this site to the CMS physicist's offices and to the CMS software and computing staff offices would be of great benefit.

- Management should take into account the likelihood of less than full staffing of the S&C project, and have contingencies if key positions are not filled. Turnover should be anticipated by having good written records and documented procedures to allow new people to step into roles with as little ramp up time as possible.

- The committee feels that the S&C Level 1 manager should be given as much flexibility as possible to optimize the funding profile among the different subprojects. The committee recommends that the S&C project management work with Fermilab and the funding agencies to develop a plan to deal with short-term budget shortfalls.

- We recommend that Fermilab, perhaps in conjunction with the U.S. ATLAS Tier 1 center and other Tier 1 Centers, open discussions with CERN on mechanisms for addressing the relevant issues of hardware, software, and networking for LHC computing

- We encourage the head of User Facilities and the U.S. CMS S&C Project manager to take a leadership role in defining the mission of the Tier 2 centers, supporting their selection and construction, and developing the appropriate coordination mechanisms and support agreements.

- The committee encourages Fermilab to join and actively participate in the Geant4 project in ways that will advance the CMS simulation effort.

- We encourage the group to investigate possible common projects between the U.S. CMS and U.S. ATLAS.

Finally, the committee would like to thank U.S. CMS for its confidence in inviting it to review the Software and Computing Project. It also thanks Fermilab's management and support staff for the good organization of the review and the help given to the committee members.

# 7. Appendix A: Charge to the U.S. CMS Software and Computing Oversight Panel

This Panel is asked to review the joint effort by U.S. CMS, CMS and the other LHC experiments to provide a suitable set of tools for physics research. A large contribution to the computing software, hardware and physics analysis tools by the U.S. collaborators is essential for the success of the CMS experiment. A high level of participation by the U.S. collaborators is also necessary in order to meet our own goals for physics results when the experiment starts to operate. The efforts in the U.S. must be integrated seamlessly into CMS as a whole. It is essential to take into account the collaboration, not only with the whole of CMS, but also with the common elements of the LHC computing environment as coordinated at CERN and in the U.S. We are interested in evaluations of technical choices as well as considerations of cost and schedule. We begin this Oversight early enough that plans for the management of our project as well as the joint efforts can be evaluated in the light of your experience. CERN is just launching an intense study of these matters and will be interested in input.

The Oversight Panel should provide advice to the U.S.CMS Project Manager for Software and Computing. The reports will be transmitted to the Fermilab Associate Director in his oversight role, and the Chair of this Panel will be invited periodically to meetings of the Fermilab Physics Advisory Panel to present the status and plans for the project.

The tasks in the U.S. CMS Software and Computing Project are structured under two categories, each with a leader reporting to the Project Manager for Software and Computing.

1. Core Application Software: Core software and detector simulation at the subsystem level and reconstruction software.
   Questions for the Panel include:
   - Are both software intended for common use and tools intended for individual experimenters given the right attention?
   - Are the technical choices being made correct?
   - Is the appropriate sharing of software and tools at all levels of collaboration?
   - Are experimenters being helped to enter the new computing environment sufficiently by training, help services and so on?
   - Is the right approach foreseen for Quality Assurance?
   - Are there provisions for assurance of the usability of the code in documentation, ease of data access, response time, consulting and so on?
2. User Facility: Regional Center (Tier 1); Remote Analysis Centers (Tier 2 and 3) and networking.
   Questions for the Panel:
   - Is there flexibility to respond to future technological developments?
   - Are technological differences between different parts of the LHC effort justified?
   - Are the capacity and robustness of all parts of the data processing and analysis chain adequate?
   - Are the capacities of the Tier 1 centers, with their infrastructure and support, adequate?
   - Are the plans for distributed computing and Tier 2 computing centers sensible?

The Panel should examine the following aspects of these the overall Project:

- Technical scope, capabilities and progress

- Costs
- Provisions for contingency
- Resource loaded schedules

It should examine the Management Plan and the ongoing management of the project as well.

We expect that the Panel will meet twice a year. We hope to receive the written report about six weeks after a review. We acknowledge the extensive nature of the questions we are asking and hope that they can all be addressed in the first few meetings.

# 8. Appendix B: Schedule of the Review Meeting

**US CMS FNAL SCOP 23-24/10/, 2000** **10/16/2000**

**10/23/2000 Conference room: FCC1 west**

| Time | Title | Speaker | Duration |
|---|---|---|---|
| 1:00 PM | Closed session | | 0:45 |
| 1:45 PM | Introduction | Kasemann/Gaines | 0:15 |
| 2:00 PM | Int'l CMS Software and Computing: Organization, Status and Plans | D. Stickland (by video) | 0:30 |
| 2:30 PM | The US-CMS Software and Computing Project | M. Kasemann | 0:45 |
| **3:15 PM** | **Break** | | **0:30** |
| 3:45 PM | The CMS Computing + Analysi Model, Monarc | I. Gaines | 0:45 |
| 4:30 PM | The US-CMS User Facility Subproject | V. O'Dell | 0:45 |
| 5:15 PM | FNAL CD: projects and tour through Feynman | S. Wolbers | 0:30 |
| 5:45 PM | Closed session | | 0:30 |
| **6:15 PM** | **Dinner** | **Reviewers and Presenters** | |

**10/24/00 Conference room: FCC1 west**

| Time | Title | Speaker | Duration |
|---|---|---|---|
| 8:30 AM | Closed session | | 0:30 |
| 9:00 AM | ORCA Software Process; HLT Simulated Event Production | D. Stickland (by video) | 0:30 |
| 9:30 AM | R&D for Distributed Analysis, Tier2 Centers | P. Avery (by phone) | 0:30 |
| **10:00 AM** | **Break** | | **0:30** |
| 10:30 AM | The US-CMS Core Appplication Subproject | I. Fisk + L. Taylor (by video) | 0:45 |
| 11:15 AM | Summary: Milestones, Funding profiles | M. Kasemann | 1:00 |
| **12:15 PM** | **Lunch** | | **1:55** |
| 2:10 PM | Closed session | | 1:00 |
| 3:10 PM | Discussion in parallel sessions | Reviewers and Presenters | 1:00 |
| 4:10 PM | Break | | 0:30 |
| 4:40 PM | Discussion in parallel sessions | Reviewers and Presenters | 1:00 |
| 5:40 PM | Closed session, formulation of report | | 1:00 |

**10/25/00 Conference room: FCC1 west**

| Time | Title | Speaker | Duration |
|---|---|---|---|
| 8:30 AM | Closed session, formulation of report | | 1:30 |
| **10:00 AM** | **Break** | | **0:30** |
| 10:30 AM | Closed session, formulation of report | | 1:30 |
| **12:00 PM** | **Lunch** | | **1:30** |
| 1:30 PM | Close out session | Reviewers and Presenters | 1:00 |