Project Management Plan for the CDF, DØ CD Joint Run II Offline Computing Project

September 25, 2014

The task of developing the software and acquiring and commissioning the hardware for the Run II offline analysis is a massive undertaking. In scope, complexity, personnel, and budget it is equivalent to roughly half of one of the detector upgrade construction projects. A significant part of the effort is the sole responsibility of each experimental collaboration. Other parts, including the specification, procurement, and implementation of hardware and commercial software and some of the software development will be carried out by a collaboration of CDF, DØ and the Fermilab Computing Division (CD). This part of the project goes under the name "CDF, DØ CD Joint Run II Offline Computing Project" or, for short, the "Joint Offline Project". The purpose of this document is to describe the project management and oversight of the Joint Offline Project. Section 1 discusses the Joint Offline Project in the context of the whole Run II software effort and presents its motivation, goals, and history. It also discusses some special considerations which influence and, in some cases, constrain the approach that is taken to management; Section 2 presents the top few levels of the Work Breakdown Structure (WBS); Section 3 describes the overall management plan including the key management roles and organizations and the formal instruments that will be used in project management; Section 4 describes the mechanisms for project change control and Section 5 those for project oversight; and, finally, Section 6 explains how the current plan can evolve to meet changes in the needs or progress of the Joint Offline Project.

1 Introduction to the Joint Project

It is essential for CDF, DØ and FNAL that the upcoming collider run realize its full physics potential. To do this, new software must be developed and new computing hardware must be acquired and commissioned to run it on. The data volumes for Run II are very large – of order of 1 Terabyte per calendar day per experiment. Reconstructing, storing, accessing, tracking, and analyzing this quantity of information in a timely manner represents the offline analysis challenge for Run II.

1.1 Special Considerations

The offline problem, in addition to its sheer magnitude, is complicated by several factors, some of which are related to the nature of computing and some of which are related to the collaborative manner in which large HEP experiments are carried out:

- Requirements are difficult to establish in a 'research environment'. The system must be able to adapt to new experimental circumstances and theoretical or analysis insights. Since computer programs are, at least in principle, easier to modify and extend than hardware, sometimes it is harder to converge to firm decisions.

- Software practices – programming techniques and languages, software engineering methodologies, and supporting standards – are all in a state of rapid change. In particular, the move from procedural to object-oriented programming has caused significant portions of our successful legacy programming base to be abandoned and it now needs to be replaced with programs written in the new paradigm.

- A multiplicity of hardware platforms and operating systems makes the problem of developing and validating programs more complicated and time-consuming and introduces portability and maintenance problems. The fact that standards may not exist at all or still be evolving exacerbates this problem. However, support of multiple platforms is necessary to reap the economic benefits of an evolving hardware market.

- New hardware, while offering new power, may have problems of reliability and robustness and may have compatibility and interoperability problems.

- The offline system must interact with university supplied equipment on site and at remote institutions. While this brings additional resources to bear on the computing problem, in many cases the choice of equipment is dictated not by Fermilab but by the institutions' local circumstances, including their non-Fermilab commitments, rather than global considerations and results in a proliferation of approaches that complicates the support and maintenance issues for the collaborations and for Fermilab.

- It has long been understood that when software development projects get large they are dominated by complicated communications problems since any part of a system may interact with any other part. This can result in complex debugging and maintenance problems. Management of complexity at this scale is a challenge which requires some formality and structure.

- As discussed below, in order to reduce the resources required by the project to a reasonable level, significant collaboration and commonality of approach is being encouraged between CDF and DØẆhile this is highly desirable, it does complicate the process of arriving at decisions. In particular, to even have a reasonable chance to achieve commonality, the two experiments must be able to arrive at decisions at roughly the same time and this is not always possible.

These factors have been taken into account in developing the project management plan described below. This project differs from the Run II detector upgrade projects in its explicit joint character; the central computing hardware used for offline production and analysis computing in Run II will be purchased and maintained, and its control software and support personnel will be provided, by the CD for both experiments. Joint procurements, and mutually consistent if not identical solutions, are thus both more obtainable and more necessary than in the detector upgrade projects. The project management plan reflects this necessity in making the top of the reporting chain a steering committee staffed by the three collaborating organizations CDF, DØ and CD. This committee is an essential forum for achieving joint decision making, and full participation by the two experimental collaborations in the relevant CD project activities. To simplify directorate oversight, the PMP also provides a Project Office whose task coordinator(s) can handle the day-to-day reporting requirements with more efficiency than a committee could. The plan also must be viewed as a 'baseline' that can be modified and developed as the project itself takes shape and as various problems are encountered and overcome.

1.2 Objectives and Benefits

The motivations for having a joint offline project are the following:

- There are strong similarities between the offline analysis problems of the two experiments in the quantity of data and the requirements for reconstruction CPU cycles and data access.
- The two collaborations already use many tools provided and supported by the CD.
- At the upper end of the analysis chain, both experiments are manipulating the same physics objects. They already use common analysis tools obtained from Fermilab, CERN, and elsewhere.
- The collaborations use or plan to use many commercial programs aimed at solving similar scientific and data management problems.
- Some important problems, such as code management and distribution, are almost identical between the two collaborations and can be readily

translated into terms that are familiar to computing professionals and do not require detailed 'physics knowledge' to work on.

- Both collaborations need similar training in new methods and similar advice and consulting support.

- Neither collaboration feels that they can supply all the people needed to provide a smooth and user-friendly analysis system and would like to get the help of computing professionals in CD to help them with the parts of the offline analysis problem that do not require experience in detailed physics and detector issues.

Fermilab management needs to provide assistance in constructing the offline analysis and to do so in a way that uses its resources as efficiently as possible. It is also charged with providing a vision for all computing at the laboratory and a strategy for achieving that vision. The Run II software must be a fully integrated component of that vision.

One key way to control the resources required to create the kind of offline system that will facilitate the physics analysis is to exploit various levels of commonality so that each collaboration can benefit from the efforts of the other one and both can benefit from the efforts of the Computing Division.

There are many different levels and types of commonality.

- The most obvious form of commonality is in the hardware. By adopting similar hardware solutions and combining them into single acquisitions, Fermilab will get the lowest prices, reduce the number of people needed to manage the acquisitions, get the most leverage out of its support and maintenance personnel, and simplify and streamline the interactions with its vendors.

- Common choices of commercial software are also desirable, particularly for the more expensive products like high-end database software. Again, we will get the best quantity discounts and will leverage consulting, maintenance, development, and support staff.

- Significant leverage can be obtained by simply having the same people in CD working on similar problems for both experiments. For example, there could emerge two somewhat different approaches to data access but by having both projects being worked by one development group in CD we are likely to get some commonality that would not be present otherwise and at least expertise and staff can be shared between the projects.

- In major development projects, it would be ideal if both collaborations could agree on a uniform set of requirements that could be addressed by a single set of programs (perhaps with some experiment specific customizations). This kind of common approach obviously has the greatest benefits.

> While every attempt will be made to bring this about, it should be understood from the beginning that it will be the hardest to achieve. The two collaborations have different histories, experiences, internal organizations, tastes, sociologies, decision making processes, and different issues associated with their university and foreign collaborators.

While commonality is a worthy goal, we have to acknowledge that selective diversity also has its virtues. There are areas where having two choices available at the Laboratory minimizes vulnerability in a very rapidly evolving computing environment. Potential sources of high risk are vendor business stability (for both hardware and software vendors) or cutting edge technology choices (again, in either hardware or software) for our large data handling needs. The steering committee provides the forum for evaluating the tradeoffs in each area where we have to balance the necessity for conserving resources versus the advantages of spreading out the risks.

1.3 History of the Joint Offline Project

This Joint Offline Project emerged from a series of discussions that started in late 1995 and early 1996. By early 1996, a series of 'common interest' areas were established which were a subset of the WBS items which now constitute the project.

As a first attempt at working collaboratively, we defined a 'configuration management' project which quickly converged to a similar approach to version management, software releases, and code distribution for CDF and DØ.

Another area of activity was related to the provision of basic C++ classes that were needed by both collaborations for the reconstruction and data analysis. A workshop was held to discuss how Fermilab should approach this in December of 1996. It was decided to assemble a task force to generate some basic C++ classes. A steering committee was appointed to guide and oversee the project, which began in early 1997. The task force evolved into the Fermilab Physics Class Library Project. The Steering Committee gradually acquired additional project management and oversight responsibilities and became the 'Run II CDF-DØ-CD Joint Offline Project Steering Committee', whose role and responsibilities are described below.

In April of 1997, at a meeting called by the Fermilab Directorate, and attended by the offline project leaders and spokespeople of both collaborations, a project management plan was first outlined. The plan presented here is an elaboration of the plan presented at that meeting.

During this period, several working groups were formed to define tasks and develop general requirements for the project. A task force called the 'Data Access Management Needs Assessment Working Group', nicknamed DAMNAG, reviewed the requirements for production CPU, data storage and access, and CPU and storage for data analysis. They delivered a report that was reviewed

and accepted, with some comments and reservations, by the two collaborations in the spring of 1997 and which was submitted to the Fermilab Directorate.

Shortly thereafter, the Associate Director for Information and Technology at Fermilab set up a review process and appointed a review panel to advise the Directorate regarding the Joint Offline Project. The first review was held in June of 1997.

In the summer of 1997, a task force called the 'Software Needs Assessment Working Group', nicknamed SNAG, identified the programmer and physicist resources needed for the project and also began the discussion of the costs associated with the acquisition of commercial software. They delivered a report which was reviewed and accepted, with various comments and reservations, by the two collaborations and submitted to the Fermilab Directorate. Based on the needs for Computing Professionals identified in the SNAG report, the Fermilab Computing Division was reorganized to make sure that adequate human resources were available to work on the project and were organized to do so effectively.

Through the spring and summer of 1997, three models of data access were discussed and evaluated at the conceptual level. Reports on each model were written by teams including CDF, DØ and CD participants and presented and discussed by the two experiments. Several new projects were initiated including the 'reconstruction pipeline project' or RIP. A high level Work Breakdown Structure was also developed and a plan for the first major hardware and software procurements began to take shape.

A second follow up review was held in October of 1997. A written report was produced.

At the end of 1997 and the beginning of 1998, with the number of projects increasing and the first major hardware and software procurements approaching, the group formulated the project structure and management plan which is the subject of this document.

2 Current Project Definition: The Upper Levels of the Work Breakdown Structure

The process of defining the Joint Offline Project was described above. In general terms, the tasks in the Joint Project consist of 1) hardware acquisitions which need to be conducted by CD personnel, with the participation of members of CDF and DØ; 2) software license acquisitions which share many of the same issues as the hardware; 3) core infrastructure software which can be specified in an experiment independent way or involves the manipulation of objects that are dictated by general computing or physics considerations rather than experiment-specific ones (this, for example, includes the Fermilab Physics Class Library); and 4) other programs and projects which can be 'abstracted' into concepts that

are not tightly tied to the individual experiments but are shared by experiments of this type or scale. (These are the tasks which can best be done in common and can most easily be worked on by computing professionals and physicists who are not deeply knowledgeable about the details of either of the experiments.)

The detailed project definitions and staffing levels are (or will be) contained in 'charges' to project teams, project plans, MOUs, and other written agreements or contracts. The discussions of the last year have resulted in an upper level Work Breakdown Structure (WBS) which defines the general areas covered by the project. Some of these projects are underway and additional levels of the WBS are available for them. In some cases, various levels of plans or MOUs exist. Here, we present only the 'top-level' WBS because it is closely connected with the way the overall project will be managed.

WBS Level 1 is the project. Because of the collaborative nature of the project it is managed by a steering committee, assisted by task coordinators and a project office. These roles are all described below.

Below this, the Work Breakdown Structure consists of the following **four** major project areas (WBS Level 2) and **fourteen** major projects WBS Level 3).

- Basic Analysis Infrastructure
  - Fermilab Physics Class Library (a.k.a. ZOOM)
  - Simulation
  - Configuration Management
  - Support Databases
- Mass Storage Systems and Data Access
  - Storage Management and Access
  - Serial Media Working Group
  - Mass Storage System Hardware
- Reconstruction Systems
  - Reconstruction Farm and Processing Hardware
  - Networking Hardware
  - Production Management
  - Reconstruction Input Pipeline
- Final Physics Analysis Support
  - Physics Analysis Hardware
  - Physics Analysis and Collaborative Support Tools

– Visualization

Each WBS item at Level 2 will have one (possibly two) project coordinators, who report to the Run II Steering Committee and whose roles are described below. Each WBS item at Level 3 will have a project manager (possibly co-managers) who report to the Level 2 area coordinator and whose roles are described below.

3 The Management Plan

This is a joint project among three organizations, CDF, DØ and the Computing Division. The reporting, advisory, and oversight roles of the different entities are shown in Fig. 1.

3.1 The Run II Joint CDF-DØ-CD Project Steering Committee

The overall project management will be carried out by a steering committee which consists of representatives of DØ, CDF, and the Computing Division. The steering committee will consist of two members from DØ, two members from CDF, and two members from CD. In order to insure a smoothly functioning committee, it is permitted, but not required, that each regular member have an alternate to substitute in his or her absence. Alternates may attend all meetings as observers.

This committee is responsible for creating and maintaining the WBS and for ensuring that it is being properly executed. It appoints the top-level WBS managers and must concur with the organization of each each major project and any major assignments of responsibility by the top-level managers. It is responsible for making sure that each project is defined properly through specification of requirements accepted by all three collaborating organizations and that the appropriate manpower levels are assigned to it. It is responsible for proposing the budget and for seeing that the money is spent in accordance with the approved project plan. It has the responsibility for monitoring milestones and adjusting resources where appropriate. It has the responsibility for change control and must approve major new projects and major changes to the scope, deliverables, schedule, and resources for all major projects. It has the task of setting up reviews and status reports to monitor all aspects of the project. It has the responsibility to keep FNAL and the CDF and DØ collaborations fully informed about the status of the project. All activities are subject to the review and oversight of the Fermilab Directorate.

The CDF and DØ collaborations choose their members of this committee who are then appointed by the Fermilab Director. The collaboration members should be able to authoritatively represent their organization within this

committee, specifically to be able to allocate resources, conduct discussions, propose solutions, request collaboration input, and most especially resolve issues – including review and acceptance of plans – related to the Joint Project expeditiously within their collaboration, according to its own mechanisms.

The Computing Division members will include the Division Head and one of the Task Coordinators whose role is described below. The Deputy Division Head serves as an alternate to the Division Head. The other Task Coordinator will serve as the second alternate. The head of the Computing Division has a special role in this process in that he/she is responsible for the overall health of the lab's computing and must make sure that this project is consistent with lab goals and vision for the future. He/she is also the line manager who approves the major purchases of hardware and software for the project and is accountable for the effective use of CD personnel on the project and for the operation and maintenance of all hardware and software acquired for the run. Most detailed management functions for the CD portion of the project are delegated by the CD Head to the Task Coordinators.

It is hoped that the committee can reach decisions by consensus but if votes need to be taken in the Steering Committee, each of the three organizations has one vote.

3.1.1 The Run II Joint Offline CDF-DØ-CD Committee

This group is an advisory body to the Steering Committee. Its function is to provide a forum where issues may be discussed and debated in an attempt to reach consensus. The status of the ongoing project is also reviewed. Membership is determined by the Steering Committee and consists of members from each collaborating organization in roughly equal numbers. While not officially a decision making body, this committee has proven to be important for forming consensus and occasionally 'straw polls' are taken on the 'sense' of the group which then constitute recommendations to the Steering Committee. The Steering Committee appoints the chairperson or persons and a recording secretary for this meeting. The chairperson(s), in consultation with the Steering Committee, call the meetings and set the agenda.

3.2 The Run II Task Coordinators and the Joint Offline Project Office

There will be two Run II Task Coordinators. Their role is to carry out the project management and coordination activities of the Run II project for the Run II Steering Committee and will include

- Tracking the effort expended on all aspects of the project and supervising the preparation of reports for the Steering Committee and for lab management.

- Assembling the budget for submission to the Steering Committee and the Directorate
- Tracking the budget and supervising the production of reports for the Steering Committee and lab management.
- Organizing budget and project reviews and tracking milestones required by the Run II Steering Committee or the Directorate.
- Identifying significant problems to the Steering Committee.
- Monitoring the preparation of various hardware and software acquisitions from the requirements phase through the specification, source selection process, final procurement, and implementation and providing status information to the Steering Committee.
- Supervising the preparation of MOU's or other formal agreements where required.
- Helping the Steering Committee respond to requests for status and budget information on a steady and also on an ad hoc basis for the directorate.

The Task Coordinators will be members of the CD. The Steering Committee must concur with their appointment and, since they are a key point of contact between the project and the Directorate, the Associate Director for Information and Technology and the Fermilab Director must also concur with their appointment. One of the two Task Coordinators will serve as a member of the Steering Committee. The other one will serve as alternate.

The function of the Run II Joint Offline Project Office is to support the project management activities of the Run II Project. It is an instrument of the Run II Steering Committee. However, this office resides in the Computing Division. The staff are members of the Computing Division and they report to the Computing Division management.

The broad responsibility of this office is to provide, under the direction of the Run II Steering Committee, all necessary coordination, monitoring, and accounting, effort tracking, tracking of milestones, and other management functions for the Run II project. This includes maintaining an up-to-date resource loaded cost and schedule complete with project milestones and deliverables. This office will be the means for fulfilling the Steering Committee's obligation to provide information on a regular and, where required, on an ad hoc basis to the Director's Office concerning all aspects of the status of the project.

To assist the two Task Coordinators, the Computing Division budget officer will serve part time in this group and will generate the agreed upon budget, effort, and other supporting information under the direction of the Task Coordinators. There will also be a Project Technical and Administrative Aide who will help the Task Coordinators in carrying out their job and who will have specific responsibility to maintain the resource loaded cost and schedule.

The Task Coordinators can draw on Computing Division technical and administrative support, in consultation with the Division Head, to help them in various aspects of their tasks.

3.3 The WBS Level 2 Project Area Coordinators

The Run II project currently consists of four 'project areas'. Each project area includes several of the projects from among the fourteen projects described above. Definition of the project areas is to facilitate more close interaction between related projects, to provide a focus for integration and interface issues, and to allow for the reallocation of funds and people, up to some threshold, between such projects as needed. These project areas are defined at Level 2 of the WBS.

Each Level 2 Project Area has a Project Area Coordinator drawn from the collaborating organizations. The Project Area Coordinators will in general have significant involvement in one or more of the projects within their area. They will be appointed by the Task Coordinators with the concurrence of the Steering Committee for a fixed term - initially for one year.

The appointments of Project Area Coordinators are renewable. The Task Coordinators and the Steering Committee will consult with key personnel in the three collaborating organizations to make sure that the people holding the positions can establish good working relations with the people participating in the projects they will be involved with, and will consider carefully the ability of these people to interact with the WBS Level 3 Project Leaders in their area.

The primary functions of the Level 2 Project Area Coordinators are to focus on integration issues, identify interfaces between projects, and promote uniformity and consistency across closely related and interacting projects - for example between a procurement project and the construction project that will be developed for the hardware or software thus procured. The Level 2 Project Area Coordinators will help the WBS Level 3 Project Managers and the Task Coordinators in understanding the progress of a project, the resource allocations, and the status of the budget, and presenting them to the Project Management.

Specific responsibilities include, but are not limited to the following:

- Together with the Level 3 Project Managers and the Task Coordinators, setting and reviewing the schedules and milestones for the projects and making sure that all integration, interface and cross-ties within the project area are understood.

- With the Task Coordinators and Level 3 Project Managers, organizing reviews for individual projects and the group of projects within their area.

- Identifying issues that interact with and interface to other Level 2 Project Areas and bringing them to the attention of the Task Coordinators.

- Working with the Level 3 Project Managers, the Task Coordinators and the organizational Line Managers on the reallocating of resources to deal with problems within their project areas, up to some threshold.
- Representing their project area in discussions and interactions with the other Project Area Coordinators and the Task Coordinators.

3.4 The WBS Level 3 Project Managers

The Level 3 Project Managers are directly responsible accomplishing the joint projects we have defined for Run II Computing. The responsibilities of the Level 3 project managers are

- To develop, when at an appropriate stage, all the elements of a complete project plan, including requirements, specifications, resource estimates, schedules, budgets, deliverables, and cross ties to other projects in the form of extensions to the current WBS and resource loaded cost and schedule.
- To propose staffing for each project and subproject.
- To provide management and technical input to carry out the project.
- To nominate leaders of major sub-projects in consultation with the Level 2 project area coordinator, subject to concurrence by the Run II Steering Committee.
- To prepare reports – including regular bimonthly status reports – for the Steering Committee.
- To assist the Task Coordinaotors in responding to regular and ad hoc requests for information from the Directorate.
- To organize internal reviews of projects and subprojects.
- To assist in preparing for external reviews as required.
- To allocate resources within their part of the project.
- To bring any major problems in the project to the attention of the Steering Committee via the Level 2 project area coordinator and to prepare formal change control requests when required.

3.5 Interaction with the Experiments

The interaction between the project management and teams and the experiments is the most critical element of this project. The experiment consensus must be expressed in the requirements and specifications on which the projects

are based. Moreover. continuous involvement of key members of the experiments is essential during the implementation phase. The experiment members must make sure that any changes in the experiment's expectations or needs is immediately transmitted to the project team and that appropriate changes are made to the plan. The experimenters must also monitor the implementation of each project to make sure that the team delivers what was specified and to provide feedback as quickly as possible on whether the result is acceptable to them. We expect that each experiment will have a 'liaison' to each major project in which it has an interest. We expect experiment representatives to serve on the 'source selection' teams that will be involved in the acquisition of hardware and software. Finally, we expect each experiment to have a formal mechanism for reviewing and reacting to major plans, requirements documents, specifications, acquisition documents, etc.

4 Change Control

Change control will be handled on a graded basis. Changes in cost or schedule of an incremental nature will be handled by the Level 2 managers and reported to the Task Coordinators (for example, up to $10K or two weeks). More significant changes can be dealt with by the Task Coordinators (for example, $50K or one month). Changes above their threshold must be reported to the Steering Committee for approval. The Directorate must be informed of all changes at this level and must approve all changes above a final threshold (e.g. $100K or two months). The exact thresholds will be determined once the plan is in place and will be added as an amendment.

5 Project Oversight

Project oversight takes several forms. The Steering Committee, in addition to its very high-level management functions, must carry out oversight of the many components of the project. The Computing Division is responsible for the effective use of laboratory resources and for the successful implementation of the computer hardware and a portion of the software. It has responsibility for the hardware and software procurement associated with the project. The Fermilab Directorate has the ultimate oversight responsibility for all these activities. The readiness of the offline computing is also a 'scheduling' issue – that is, it affects when the experiments are ready to take data. The CDF and DØ experiments are the intended beneficiaries for this project and have concerns of their own.

The Steering Committee will design its own mechanisms for 'internal' oversight which will include frequent status reports, periodic internal and/or external reviews of major projects, and detailed tracking of milestones and accounting for failure to meet them. The Steering Committee is responsible for updating

the organization of the Joint Offline Project, as required, to address such problems as changing understanding of the needs of the two experiments, or inability to use the present structure to achieve the necessary milestones and final results. The Task Coordinators will assist them in their oversight function.

Collaboration oversight may be achieved in many ways. The collaborations will be well connected to the project by their representatives on the Steering Committee and by their participation in the Run II committee. Members of the project, especially the top-level project leaders, are available to participate in experiment reviews. Either experiment may request a major review specific to an aspect of the Joint Offline Project or obtain information on any aspect of the project by making a request through the Steering Committee.

The laboratory currently has an oversight structure for the project. The Associate Director for Information and Technology (ADIT) has appointed a a review committee which has met periodically at his request to review the project and make recommendations. The spokespersons of CDF and DØ have been asked to participate in these reviews.

With the establishment of the Run II Offline Project Office, the Directorate will be able to get regular status reports. The Run II Project Office, and especially the Run II Task Coordinators, are the main point of contact for regular and ad hoc requests for information. Regular bimonthly project status reports will be generated and will be submitted to the Directorate.

6 Evolution of the Project Plan

The project plan described here is only the upper level of a complete plan. Some of the projects – such as RIP, ZOOM, and configuration management – are in the implementation stage and those projects can be specified with a more complete Work Breakdown Structure, milestones, and deliverables. This is being done now. Other projects are still being defined and details will be added to the WBS as soon as they are understood. We expect to begin the process of assigning resources and budget to the parts of the project that are well understood.