Strategic Directions of Computing at Fermilab

Stephen Wolbers

Fermi National Accelerator Laboratory, Batavia, IL 60510

Fermilab computing has changed a great deal over the years, driven by the demands of the Fermilab experimental community to record and analyze larger and larger datasets, by the desire to take advantage of advances in computing hardware and software, and by the advances coming from the R&D efforts of the Fermilab Computing Division.

The strategic directions of Fermilab Computing continue to be driven by the needs of the experimental program. The current fixed-target run will produce over 100 TBytes of raw data and systems must be in place to allow the timely analysis of the data. The collider run II, beginning in 1999, is projected to produce of order 1 PByte of data per year. There will be a major change in methodology and software language as the experiments move away from FORTRAN and into object-oriented languages. Increased use of automation and the reduction of operator-assisted tape mounts will be required to meet the needs of the large experiments and large data sets. Work will continue on higher-rate data acquisition systems for future experiments and projects. R&D projects will be pursued as necessary to provide software, tools, or systems which cannot be purchased or acquired elsewhere. A closer working relation with other high energy laboratories will be pursued to reduce duplication of effort and to allow effective collaboration on many aspects of HEP computing.

Key words:

1 Current Computing Issues at Fermilab

Computing at Fermilab is driven primarily by the experimental program at the laboratory. The experiments produce data which must be processed through the computing systems. The results are then analyzed and lead to the physics results of the laboratory. Though the experimental program has the largest single influence on computing and computing directions there are certainly other influences and projects that also have an effect. There is a large amount of infrastructure which must be built and maintained. In addition there is an influence which the Computing Division has on future developments of Computing at Fermilab.

1.1 Collider run I

The major reconstruction of the data taken by CDF and D0 in run I was completed in 1996. The current emphasis is on the final physics analysis of the data. Both experiments are using a combination of Silicon Graphics Challenge systems, UNIX workstations, and VMS clusters to finish the analysis of the run I data. Adequate computing must be provided if the physics results of run I are to be fully achieved.

The computing farms at Fermilab allowed us to reconstruct CDF and D0 data during run I in "quasi-real time", i.e., as quickly as the data was taken but with a short delay (days to weeks) to allow for final calibration constants to be created.

1.2 Fixed Target

The accelerator is currently running in fixed target mode, with 10 experiments collecting data. The experiments cover a broad range of computing needs, with raw data samples ranging from the modest through many multi-Terabyte and up to over 100 TBytes for one experiment. The data will be analyzed on systems which have been purchased primarily for these experiments' offline needs.

The largest new addition is the UNIX farms expansion, which is currently being completed and which will add over 15,000 MIPS to the current 7,000 MIP farms. The new farms consist of SGI and IBM workstation and SGI SMP (O200) systems. More details can be found in the contribution on Fermilab UNIX farms, also given at this conference.[1] The farms will have sufficient compute power to allow timely reconstruction of the data.

For further analysis and data handling (splitting, DST-creation, etc.) a set of UNIX systems (FNALU) will be used. These systems contain both batch and interactive computing components, connections to robotics and 8mm tapedrives, as well as software required to handle most fixed target analysis tasks. Recent system acquisitions have increased batch capacity and disk space. Work has also been done to increase the stability and reliability of AFS, which is the primary filesystem of FNALU.

For the KTeV experiment a special set of requirements led to the purchase of a dedicated computing system for KTeV's offline needs, including event reconstruction, splitting, DST-creating, tape copy, Monte Carlo event simulation, and final physics analysis. The system consists of two DEC 8400's with 10 440 MHz processors each, 500 GBytes of disk, 16 DLT 4000 tapedrives, and a connection to mass storage. This system has just been put into production and is expected to provide rapid analysis for physics results.

1.3 General Laboratory

The rest of the laboratory and user community is also in need of computing. There are other scientific efforts at the laboratory which require computing. These include the Sloan Digital Sky Survey, theory calculations (perturbative and non-perturbative), future experiments, CMS, magnet design, accelerator design, and others. Some of these require significant computing power, and this is being provided on many of the central systems.

In addition, there are needs for Web servers, news servers, database servers, name servers, tape copy facilities, and other services which are provided by the Computing Division.

2 Run II

The next few years will be dominated by the preparation for the data and analysis which will be coming from the next run of CDF and D0 (run II), which will begin in 1999.

Recently a careful evaluation of the realistic data rates of the two detectors has led to a much clearer understanding of the huge task ahead. The volumes of data which are anticipated can be found in Table 1. The amount of raw data is truly huge compared to previous CDF and D0 runs and is a reasonably large fraction of the volumes expected for CMS and ATLAS.[2]

2.1 Strategies for run II

Various strategies for handling the large CPU, data storage, data access and analysis needs are being developed. Final solutions are not yet in-hand but the intent is to design and build the required computing systems during the next 3 years with the major requirement that sufficient computing systems must be in place when data is first being taken in 1999.

Table 1
Run II data rates

Category	Parameter	D0	CDF
DAQ Rates	Peak Rate	53 Hz	75 Hz
	Average Event Size	250 kB	250 kB
	Average Data Rate	5 MB/s	7 MB/s
Data Storage	Number of Events	600 M/year	900 M/year
	Raw Data	150 TB/yr	250 TB/yr
	Reconstructed Data	75 TB/yr	135 TB/yr
	Physics Analysis Data	53 TB/yr	79 TB/yr
	Total Data Volume	280 TB/yr	464 TB/yr
CPU estimate	reconstruction/event	1000-2500 MIP-s/ev	1200 MIP-s/ev
	Reconstruction	34000-83000 MIPS	56000 MIPS
	Analysis	60000-80000 MIPS	90000 MIPS

2.2 CDF/D0/Computing Division Cooperative Efforts

It has been recognized that the preparation for run II will not be successful without a high degree of cooperation among the Computing Division, CDF and D0. This cooperation is being seen in the following efforts. First, physicists from CDF and D0 are members of the Computing Division staff as part of the CDF and D0 Computing and Analysis Groups in the Division. These physicists provide very close cooperation between the CDF and D0 collaborations on the one hand and the Computing Division on the other.

The largest coordinating activity which is specifically aimed at run II planning is the Run II CD/CDF/D0 joint committee. This committee consists of physicists and computer programmers from the two collaborations and from the Computing Division. Its role is purely to prepare for run II, and to that end many working groups have been formed to formally study issues and make recommendations, coordinate activities, write formal reports, and to provide actual effort towards solving specific problems. Working groups have been established in languages (C++ and class libraries), configuration management (code management and distribution), software development tools and methodologies, and needs assessment (understanding the parameters of the hardware and software required). In addition there have been reviews of various Computing Division projects (e.g. HPSS) to understand how these projects can be focussed on run II. Additional working groups will be established as required to get the necessary computing in place by the beginning of run II.

It is expected that the physicists on CDF and D0 will write all of the code required for the reconstruction of the raw data. The Computing Division will help wherever possible with training and with some of the common infrastructure required.

2.3 Strategies for Handling run II Needs

There are some general strategies which are being used to design systems which will meet the needs of run II computing. It is hoped that some combination of robotics, hierarchical mass storage (in particular HPSS) and possibly Object-Oriented Data Bases will be used to handle the access to data. This is clearly not yet a solved problem and much more work needs to be done to model the access patterns and understand the robustness and capabilities of the systems. Event reconstruction is most likely handled by PC farms, possibly running the LINUX operating system. There are certainly other possibilities but in general this looks like a part of the offline problem which is easily solved. Networks are an important part of the computing strategy. This includes networks internal to the laboratory which will be required to move data around among the large central computing systems, the connections to workstations and other desktops at Fermilab and the network going out to University and other remote collaborators. Video-conferencing will continue to be an important part of remote communication to allow the widespread collaborations to communicate effectively. For some of the needs there are not yet clear solutions. Some R&D work will have to be performed in advance to understand what can be done to solve the problems. Finally, there will have to be close coordination amongst all the participants at the laboratory and in HEP as a whole in order that the most effective use possible is made of the limited and in many case diminishing resources available.

3 Beyond run II

In the years during and beyond run II Fermilab has to shape its computing to meet the requirements of the physics which will be pursued at the laboratory. There are many program options being discussed as well as many already approved experiments during the period of run II and beyond. These include the Main Injector experiments MINOS and COSMOS, a Main-Injector based fixed-target program, astrophysics experiments (SDSS and Auger), very high rate collider experiments specializing in charm and b physics, higher luminosity $p\overline{p}$ experiments, and efforts going beyond the current accelerator complex. It also includes the participation in the CMS experiment at CERN, both as a collaborator and as a possible regional computing center.

4 Conclusion

Fermilab is facing great challenges in computing and physics. A large effort will be required to provide adequate computing for the CDF and D0 run II. This is expected to keep the Computing Division busy for the next 2-3 years and beyond. Other efforts at the lab will receive support as well during this period. At the same time, new initiatives will be proposed and studied, leading to more demands in the future. Fermilab must continue to innovate and respond to these needs in order that the physics results of the laboratory are not compromised by a lack of computing but instead are enabled by excellent computing.

References

[1] Ron Cudzewicz, et al., Next Generation Farms at Fermilab, submitted to this conference.

[2] CMS and ATLAS Computing Technical Proposals, CERN/LHCC 96-45 and CERN/LHCC 96-43, December, 1996