# Computing Division Self Assessment FY04

## *Hardware Management*

The Computing Division (hereafter referred to as CD) manages 24,000 square feet of computer rooms in three separate buildings which house over 3,000 computers and related equipment including networking and 6 automated tape libraries. The laboratory's scientific mission requires 7 x 24 operations. Physical infrastructure includes electrical power, air conditioning and fire protection. The current power consumption is 1 Megawatt and requires 300 tons of air conditioning. In the Feynman Center we utilize UPS and generator power backup systems. These utilities are maintained by service contracts and are monitored and controlled via an automated system, Johnson Controls Metasys, with automated paging implemented. The result has been our ability to manage a coherent computing operation across three separate sites with minimal personnel. Staff members are on-call at all times to respond to problems.

Computer room planning, including computer layouts and electrical distribution, are managed and documented in the computer-based Aperture CAD/database system. All change management is performed using this tool.

All computer rooms and utility areas are protected using security ID card readers for entry. We ensure that all computer rooms conform to applicable building codes and NFPA 75 for Information Technology Centers. This includes automated unlocking of all doors in a fire emergency. CD personnel and others are granted access to these rooms based upon need. Required reading of computer room work rules and hazard analyses are used to insure personnel safety. Our safety record exceeds 1,000 days without a restricted or lost time injury.

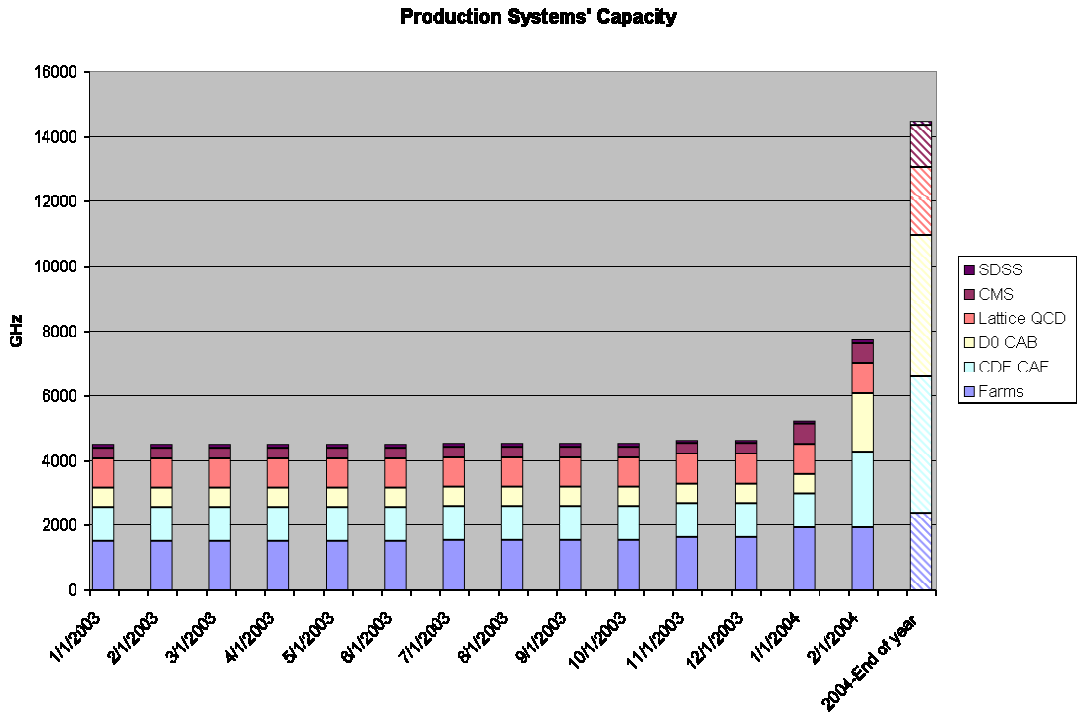
Trending and projection of computer growth is used to plan future facility needs. Increased effectiveness of the Fermilab Accelerator complex and a successful continued experiment program has led to an increased need for data analysis which translates into more and faster computers which require more space and higher density power and cooling. New computer rooms have been built in existing, unused buildings in order to meet this demand at lower cost. This year we are adding 1 ½ Megawatts and 330 tons of air conditioning just to get through the FY05 period. Engineering plans for additional infrastructure for FY06/07 is already in progress.

The principal mechanism by which computing needs have been met is through the use of "farms" of computers—sets of identical PC's all arranged to compute independently. The hardware procurement and management of these computers is tuned to take advantage of the savings that can be achieved through a carefully structured bid and acquisition process. Specifically, Fermilab holds a Linux Server Vendor Qualification once every two years and in this process, the vendors are tested for familiarity with the Linux operating system, quality of hardware construction, on-site service capability, and price-performance. 21 vendors submitted evaluation units in the most recent qualification and 18 were found to be technically qualified. The top 5 vendors in price/performance were selected to be our eligible Linux vendors.
. As the need arises for purchase of computing analysis farms (e.g., in the circumstance of increased numbers of collisions as described above), the five vendors are asked to bid. The procurement is for an entire racked system, including computers, rack, cables, console servers, and power controllers, as well as 3 years on-site parts and labor warranty.

When the bid is awarded, follow-up quality assurance procedures are employed as part of the procurement. The goal is to correct any problems that may exist before the units and the racks

are shipped. These include visits to the vendor site by Fermilab personnel and visits to Fermilab by the vendor before. The combination of choosing hardware that is more reliable through the evaluation process and of purchasing the hardware warranty up front has led to significant cost savings during the post-installation period as well as much-improved uptime. In the case of problems relating to service, quality or timely delivery from a vendor, that vendor is removed from the list and the next one in line takes its place.

The following graph shows the growth of Farm computing in the past year.



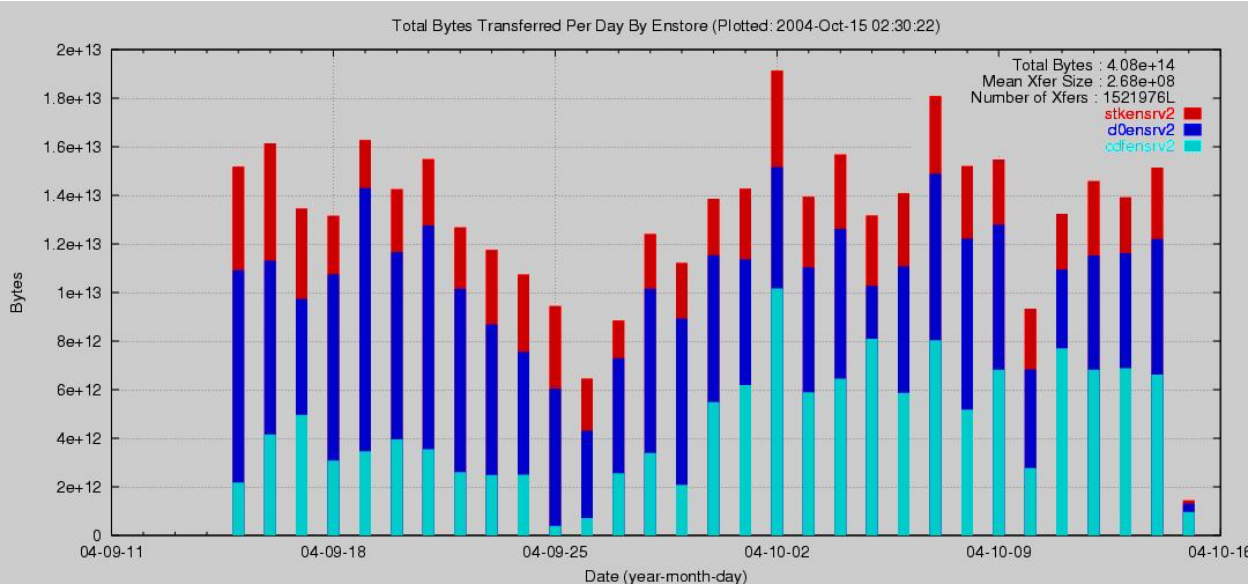

The laboratory's critical scientific data sets are maintained in the Feynman Computer Center. There are now over 2.2 Petabytes ($10^{15}$) of active data sets stored in automated tape libraries (ATLs) in six tape silos. Additionally, there is a tape vault containing legacy (relatively inactive) data collected by the earlier Run I and Fixed Target experiments .

The management goals of these critical data sets are to: Steward the data securely, provide good access for experimenters, and to operate efficiently.

Security: The ATL systems routinely monitor the retention qualities of the archive. Newly written files are sampled when written; moreover a random tape is sampled each hour. This year the automated checking of newly written data discovered a drive with defective hardware, writing files in a defective manner. This monitoring prevented a severe case of data loss. Accidental over-writing of data is prevented with a systematic program of physically activating "write-protect" tabs, even in the ATLs.

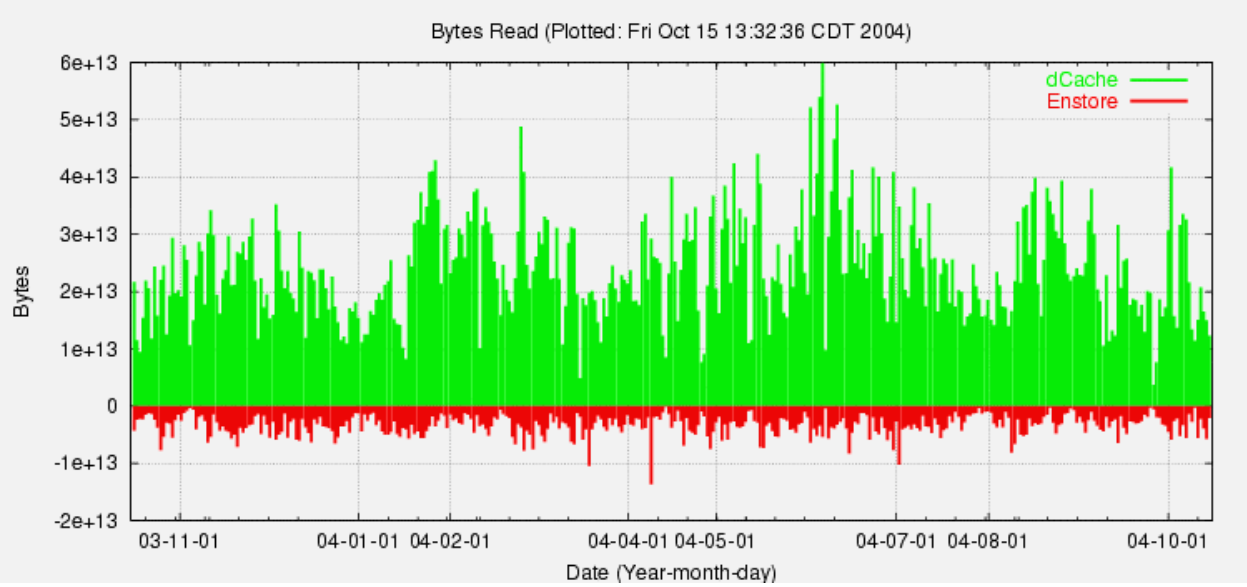
Performance and efficiency: The automated tape systems were able to move up to 25 TB of data per day in FY2004. The two most active legacy data sets (SDSS and KTeV) were migrated into the automated libraries thereby reducing the frequency of hand-

mount tape requests to the point where tape operators need not be present at FCC continuously to service tape mount requests.




.

Bytes read from CDF dcache system



### *Software Management*

The Fermilab CD works with Divisions, Sections and experiments to provide common tools, methods and standards for management and development of software used for physics, engineering and business applications. In addition, the CD provides direct support for Windows and Linux operating systems, and administers Fermilab-wide contracts for software licenses and maintenance.

**Support of Physics & Engineering Software :** Development and management of the software used for physics data acquisition and analysis is primarily the responsibility of the individual experiments. The CD works with the experiments to recommend and provide commons set of tools such as compilers, debuggers, code managers, performance tuners and libraries. It also works with the experiments to develop standards and best practices for software development methods, code testing and production release, and documentation. During the previous year, both of the two Run II experiments released new versions of production analysis software while maintaining stable acquisition and analysis of production data.

**Support of Linux:** The primary operating system for physics applications is Linux. The CD provides and maintains a locally-supported version of Linux which is built from open source and based on a popular commercially-supported distribution. Fermi Linux includes specific customizations for physics software development and analysis, and for Fermilab security requirements. Fermi Linux is supported by approximately two FTE's in the CD, with additional effort from the user community.

During the previous year, support for Fermi Linux 7.3 continued. This is the version of Fermi Linux most commonly used on production systems, especially for the Run II experiments. Support for Fermi Linux 9.0 was dropped, and these systems were migrated to Fermi Linux LTS 3.0. Fermi Linux has been accepted in the HEP and broader scientific community as Scientific Linux, and the support for Scientific Linux has broadened into an international community effort, shared by several laboratories and institutions. A significant milestone was the acceptance of Scientific Linux by the LHC Grid as an alternative to commercially-supported Red Hat Linux.

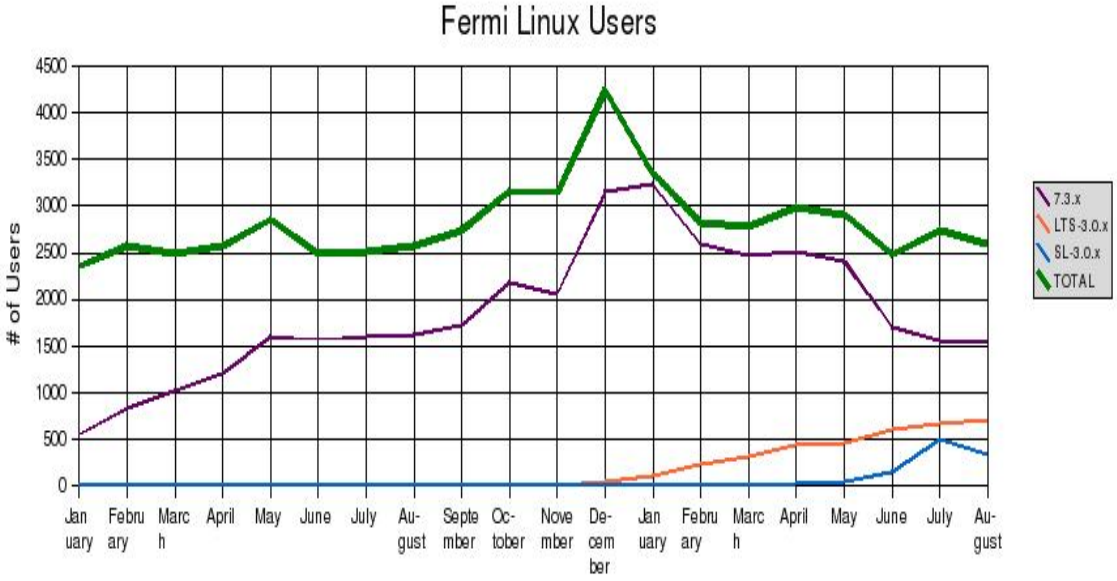
Fermi Linux includes an automatic update and configuration management system, the Yellowdog Update Manager (YUM). YUM automatically distributes Linux packages and updates from a central repository, including security updates. Important security updates are installed on most systems within two days of becoming available on the repository.

**Support of Windows:** Microsoft Windows is the primary operating system for administrative work at the Laboratory. There are over 3000 desktop and servers systems currently installed in the Fermilab Windows domain.

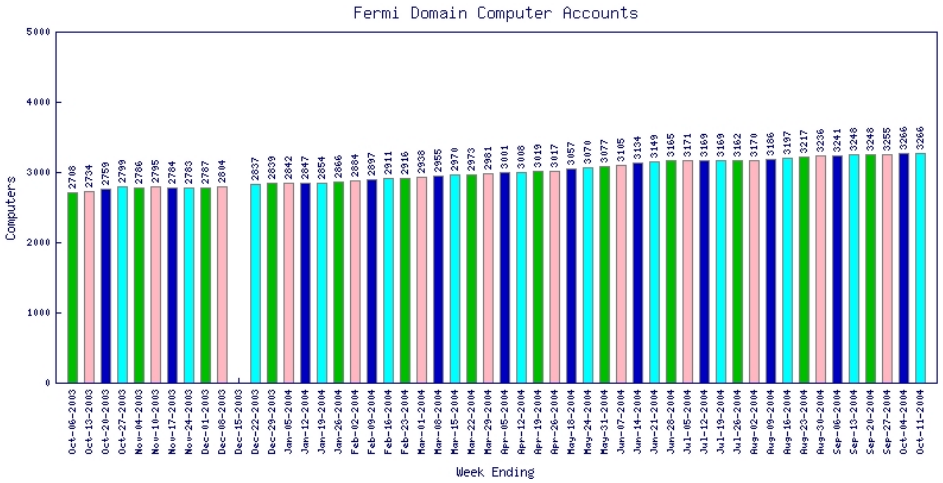
During the previous year, the Laboratory put in place an enterprise-wide software support agreement with Microsoft. This agreement was a significantly lower cost than other available alternatives, and covers new licenses and software updates and maintenance for a three-year period. It allows the Laboratory to better control its Microsoft software costs by centralizing the administration and tracking of licenses. The cost of this agreement in FY04 was $503,000, which was front-loaded for additional licenses. During FY05, the cost will be approximately half this, subject to the addition of new licenses acquired during FY04.

The Laboratory also instituted a two-tier automatic patching system for Windows systems for the over 3000 computers in the Windows domain. The first tier is managed locally by support organizations and is used to push both routine updates and critical security updates. The second tier exists as a “safety net” and pushes critical security patches to any systems missed by the first tier. The system dramatically reduced the time needed to roll out critical security updates, from as much a few weeks to only a few days, and nearly eliminates the need to visit desktop systems to install patches.

**Software Contracts:** CD centrally manages software support contracts for operating systems (Microsoft, Sun, Silicon Graphics) and common packages, such as compilers and debuggers (Kai), CAD engineering and design (IDEAS), and databases (Oracle). Costs of these centrally-administered contracts were reduced by approximately $100,000 (10%) between FY04 and FY05. The Laboratory extended the current Oracle agreement to the end of CY 2006, at a cost of $113,500 per year, while continuing to negotiate a new agreement with Oracle to avoid an estimated annual increase of $2M to $3M per year.



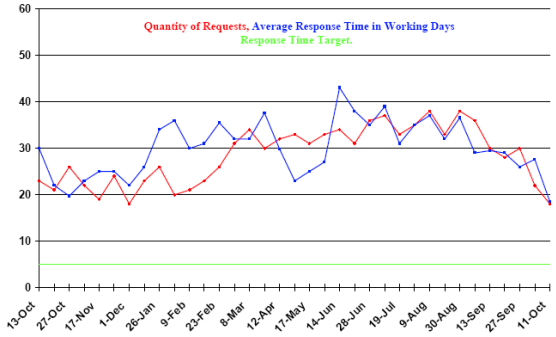

Number of Fermi Linux installations (based on YUM updates)



### *Network Management*

The mission of Network Section is to provide all aspects of data communications support, including design, acquisition, installation, operation, maintenance & documentation of cabling plant, device infrastructure, and network services necessary to support the Laboratory's computer network facilities. Self-assessment of the Section's network management efforts breaks down into analysis of five general areas of support that encompass the bulk of network support activities at the Laboratory:

Physical Infrastructure is the cabling plant and associated hardware that support data communications equipment. In the past year, a new, more cost effective cable infrastructure was developed, tested, and subsequently deployed in the new High Density Computing Facility. Cabling support for experiment farms expansions has consistently been provided ahead of the systems arrival and burn-in. The general cabling installation work queue has remained too high, but in August we contracted outside technician support which has reduced the installation queue time, and we expect that outside technician support to continue to be used until the queue length is acceptable.



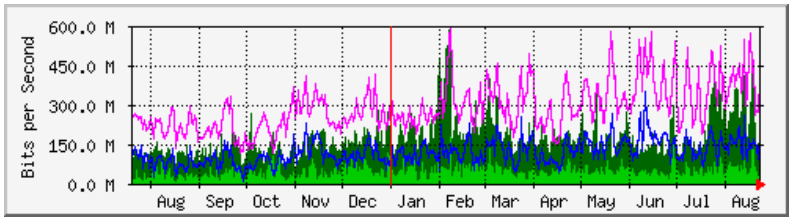

Operational support involves management of the network devices & services that comprise the functioning facility network. In the past year, there was no disruption to core network switches outside of scheduled maintenance periods. The absence of redundancy in core network devices did result in disruption during maintenance periods, but we are actively planning to improve core network redundancy in FY05. The 7x24 off-hours support coverage was provided within the published response time goals, with no coverage escalation beyond the primary & secondary on-call personnel required. Monitoring of essential services with enhanced with a monitoring tool that checks for response of network services themselves, not just the connectivity of the network server. Core network device monitoring was enhanced by deploying automated paging via HP OpenView.

Network upgrade and planning is required to ensure that network capacity meets customer needs. The architecture of the facility network is based on work group LANs, customizable to meet the specific needs of affinity groups, such as experiments or Laboratory Divisions. 10-gigabit ethernet backbone links were implemented in the CDF & CMS work group LANs to meet capacity needs for those two experiments. A 10-gigabit ethernet upgrade to the D0 work group LAN is under procurement, and should be in place in time to meet D0's increasing network capacity needs. Overall, in the past year the Network section deployed 26 new network switches (19.5% increase), supporting 1944 new switch ports, of which 1116 were gigabit ethernet. As a result, the capacity of the facility internal network has been kept well ahead of demand for bandwidth in general, and gigabit network connections in particular. Wireless

network coverage was expanded by the deployment of 40 additional wireless access points, a 28% increase. Progress in migrating existing shared-network connections to dedicated switch ports, however, was less successful. The equivalent of one floor of Wilson Hall was upgraded to dedicated network connections; the year's target was two floors. Our plan for the coming year remains to upgrade at least the equivalent of two Wilson Hall floors. Only half of the remaining shared network connections at D0 were upgraded to dedicated connections, although the remaining shared network connections are expected to be upgraded within the next several months.

Network support of the Laboratory Computer Security Protection Plan (CSPP), and computer security in general, represents a major area of network management effort. The automated scan blocker for the facility's off-site network traffic has been rewritten to generate fewer false positive automated blocks. It was modified to be more robust against intensive external scanning, after a virus outbreak earlier in the year caused performance problems for off-site traffic. The current implementation of the automated scan blocker does not perform true intrusion detection, which is a recognized weakness. It has been designed to support existing intrusion detection products in parallel, but that function has yet to be implemented. The Network Section also has developed and deployed a DHCP Registration utility to require registration information from visitors utilizing the facility network.

Off-site network connectivity has been the network management area of greatest change in the past year. The laboratory depends on the ESnet for its production network connectivity. The capacity of the existing production network off-site link, a 622Mb/s service provided by the Energy Sciences Network (ESnet), is barely sufficient to meet current off-site bandwidth needs. The Laboratory's off-site network traffic doubles in volume every year, and clearly requires additional bandwidth:




Upgrade of the existing ESnet link entails funding and approvals that are beyond the control of the Laboratory. However, we are pursuing that ESnet upgrade path. To support advanced systems development, the Laboratory has leased optical fiber cable to an network exchange point known as StarLight. Optical network equipment has been deployed on the StarLight fiber that provides 12,000 Mb/s of off-site bandwidth for the Laboratory, with an upgrade path for greater capacity, if needed. Non-production network traffic that supports development of the CMS Tier One center is now being carried on the StarLight link. We are prepared to route overflow production traffic onto this link to mitigate production network performance problems should, as seems likely, the existing ESnet link become congested. It has not been a design goal for the StarLight link to have the same level of reliability as the ESnet link. Production network off-site capacity and reliability will remain a concern until the ESnet Chicago MAN is funded and put into place.

## *Planning for Future Systems*

### Support of Physics & Engineering Software

Another mission of the CD is to design, develop and support advanced electronic systems for parallel computing, data acquisition, event triggering and front-end electronic readout for Fermilab's experimental program, providing the necessary technical and/or cost/schedule management information required for a proposed or ongoing project. An additional component of the mission is to investigate future technology and architectures emphasizing advanced circuitry and high-speed data movement for parallel computing, data acquisition, trigger and front-end readout systems.

In support of the Run II experiments, the engineering section carried out many projects for the CDF and D0 collaborations and the test beam efforts. Projects that were successfully completed and helped the experiments keep up with the increasing luminosity of the accelerator complex were:- CDF trigger throttle to mitigate resonances in the trigger at high rates, D0 AFE II upgrade module board layout, D0 mixer system memory upgrades, CDF SRC module FPGA upgrades to increase available processing capability, D0 VRB Controller firmware upgrades to increase the trigger rate and cure legacy logic errors, and the firmware for the D0 luminosity monitor input to the Level 1 trigger.

In support of future experiments, the section contributed to the BTeV collaboration and SNAP/JDEM collaboration efforts by performing groundwork R&D projects and providing contributions to WBS and technical support documentation. These are ongoing projects with specific work areas in the past year of :-

- Major involvement in the Pixel R&D efforts since 1994. In 2004 the section built the systems used to characterize the pixel sensors and readout ASICS. Radiation dose and beam test experiments were performed and discoveries reported. The baseline design for the collision hall readout electronics from the readout chips outputs to the optical fiber to the counting room was completed and pre-pilot units will begin construction in FY05.

- Work on the BTeV trigger has been ongoing since 1998 and the baseline hardware design has been stable for several years. The Section produced a working version of the FPGA segment tracker and operated it with GEANT4 simulated physics data. This was a significant accomplishment and milestone in the R&D.

- The BTeV DAQ project status and stage of development is almost identical to the trigger. A prototype Level 1 buffer was designed in FY04 and will be combined with a pilot Data Combiner Board in FY05 to support sub-detector design efforts.

- SNAP/JDEM R&D tasks this year were (a) to understand the effects of typical spacecraft radiation levels on non-radiation qualified FPGAs and (b) to understand and quantify the data movement issues within the spacecraft data acquisition system.

**Fermilab Scientific Computing**

The Fermilab CD provides computing services for the current and future scientific programs. . These services include system administration for experiment specific needs, farm computing, robotic and disk storage, database administration and networking. In order to meet the needs of the experiments, the CD uses several planning processes.

For the Run II experiments (CDF and DØ), there is an annual Directors review of computing. Preparation for that review requires that the computing management for the experiments document how the current systems are performing and meet the needs of the stakeholders. In addition, both experiments have models for projecting computing needs into the next fiscal year (and into the out-years) based on the current computing use of the experiments and the history of computing costs projected as a function of time based on past purchases. A fundamental driver for computing resources is the online data collection rate, and both experiments plan to increase this rate. For both experiments, the computing needs exceeds the FNAL budget for Run II computing; thus the information for the review is also used as justification for international contributions to computing for the two experiments, as well as motivating and strengthening international collaboration on Grid computing projects in data handling, data movement, storage and job submission and management. In addition, the need to use hardware resources that are not located at FNAL at the same time as traffic for CMS is increasing has been used as input wide area networking planning. The most recent Director's review noted that for both RunII experiments, the computing models and computing hardware in place are performing quite well in meeting the current needs of the stakeholders.

In recognition that computing needs of the Run II experiments will extend far into the LHC era, the FNAL CD is undertaking to look for commonalities between all stakeholders in order to best leverage resources. In FY2004 a forum, "Grid and Data Management Planning" has been established to examine the needs of the stakeholders, the projects on which the FNAL CD is working and the resources available. The information developed in this forum is used to establish priorities for projects and within projects and to redirect effort when necessary. One goal is to virtualize experiment specific hardware resources in order to make most use of the available machines. The FNAL CD is working within framework of the Open Science Grid planning processes towards providing resources available via emerging or established interfaces on the grid. The Grid and Data Management planning forum is also used to inform the stakeholders of developments with respect to OSG.

Two otheractive experiments, MINOS and Sloan Digital Sky Survey (SDSS) also have planning processes in order to supply input to the computing division's yearly budget. These experiments rely to a large extent on general purpose resources. In case of SDSS, these resources include the general purpose farms for its most intense computing needs and the tape robot facilities for offline storage and backup. For specific needs of the SDSS, the FNAL CD also maintains two modest clusters, the Data Archive Server, which consists of a set of file servers, and the Catalog Archive Server machines, which stores a subset of the data in a database for collaboration and public distribution. The planning process is straightforward since the data collection rate is modest and relatively constant, and storage needs scale linearly with that rate. The MINOS experiment also makes use of the general purpose farm for data reconstruction and stores their data in the tape robot. The experiment will start taking beam data starting in January 2005. Planning for computing is carried out by a small number of people in the experiment. Plans for computing will evolve as they are confronted by real data.

**LHC Computing**

Fermilab is the host lab for U.S. CMS, and the CD is hosting the U.S. CMS Software and Computing efforts within the U.S. LHC Research Program. While the CMS department is leading the U.S. CMS efforts including hosting project management, several departments of CD are also involved with the deployment of basic hardware components and system services, the integration of those facilities into a grid-enabled distributed computing infrastructure, and the development of the applications and database services that run in that environment.

Fermilab provides computing resources to the CMS collaboration through the U.S. CMS Tier-1 computing center. We performed a large procurement of processors, storage devices, network switches and a robotic tape system this year to bring the Tier-1 facility in line with CMS requirements for this year. CPU resources doubled from 200kSI2k to 400kSI2k by adding 100 dual Xeon 3.0GHz systems. The storage capacities nearly doubled to 34TB. There is additional 4TB high-performance storage available to users of the User Analysis Facility (UAF). System support was strengthened and organized such that the CD-CMS Tier-1 manager leads a support group within the CSS department.

CMS achieved a major milestone in April 2004 with the DC04 Data Challenge; an end-to-end system test with simulated events being reconstructed at CERN and being distributed to the Tier-1 centers and then to Tier-2 centers for data analysis. In preparation for DC04, Fermilab developed important parts of the grid computing infrastructure necessary to build a truly distributed computing system. For more detail on Grid developments please refer to the later section of this document on Grids for Physics. Development work for the distributed systems concentrated on providing robust grid-enabled data storage and transport services through dCache, SRM and GridFTP. The new fiber link from Fermilab to Starlight enables high-throughput networking for these kind of data transfers. Other important work included providing information and configuration services, developing virtual organization services and interfaces to site-security, developing authorization mechanisms for grid-based storage, and providing interfaces between the different grid flavors in Europe and the U.S. to enable interoperability.

Through the UAF we serve large data samples to physics users and provide sufficient CPU resources for analysis jobs. Data sets are transferred into the Fermilab storage systems using the new grid systems, and are exported to the University based U.S. Tier-2 centers, as needed. The UAF is in heavy use by a growing user community working at the Fermilab LHC Physics Center on test beam analysis and on the preparations for the physics technical design report.

Reliability in providing large user storage space remains an issue, as the chosen software base solution (IBRIX) has proven to be not sufficiently robust. We continue to research a highly perform ant and reliable solution.

Fermilab is continuing to make significant contributions to the development of CMS core software. Fermilab software experts, many bringing valuable experience from Run II, are developing the plans and architectures for event data model support and the physics services. Work continued on the tools for binary distribution of CMS software, which are being used for all simulated event production, and now also increasingly for data analysis. A review of the requirements has been conducted working with the LPC, and work has started to implement these.

**Support for the Accelerator Complex**

The CD was involved in a number of projects in support of the accelerator complex. The most notable has been the TeV BPM project. However, CD personnel have contributed in a number of other key areas. One CD scientist was a project manager for the Run II upgrade project and helped to organize a number of reviews for the Accelerator Division (AD). She did such a good job in this capacity that she recently transferred to AD.

We have also been part of a project to provide Ionization Profile Monitors, providing engineering and software support for this project. The electronics board required for this project is similar in design to a buffer board required by the BTeV experiment and the development well matched to the skills of the engineering team in CD. This project is proceeding on schedule for delivery in FY05.

The Blastman DB project was another new effort in CD this year. A database applications developer from CD worked with physicists and engineers from the Accelerator Division to develop modern database tools to handle the data that describes the magnets and their alignment in the Tevatron tunnel. This work has proceeded well and the first phase of the project has been completed. It has also led to a request for database development and support from the PPD alignment group. This request will be evaluated in FY05 and may lead to a new database project.

CD developers have been providing support for Shot Data Analysis, Booster Monitoring, Labview applications and Vxworks. We have been able to manage the requests for continued support of these products and will continue to do so in FY05.

The CD is leading an effort to upgrade the readout electronics and software of the Tevatron Beam Position Monitor (TeV BPM) system. The TeV BPM system measures the position of the proton and anti-proton beams in 240 locations around the Tevatron. The current system was designed and built in the early 1980s and is no longer capable of reliably measuring the beam position with sufficient precision to satisfy the requirements of the Tevatron. Therefore, a project to upgrade the electronics and the front-end, online and offline software was undertaken by the Computing and Accelerator Divisions (AD). The original goal for completion of the upgrade was October 1, 2004.

The project is staffed by a mix of CD and AD scientists, computer professionals, engineers, and technicians. Approximately 80% of the effort comes from the CD. The primary responsibility of CD is project management, electronics, front-end software and offline software. AD is responsible for project management, technical coordination, online software and application and accelerator requirements.

The project has completed all requirements and design work, has been reviewed at various stages, and is currently in the midst of fabrication and/or acquisition of all hardware and software components. First beam tests with upgraded hardware and software was made prior to the August, 2004 shutdown. All tunnel and service building preparations for the new system has been made. Plans are made to install and commission the 27 VME crates of hardware when the beam returns in November, 2004. The new system should be fully operational by April 1, 2005.

The project was delayed by delivery schedules for the main digital signal processing boards as well as the increased complexity of the timing system and analog signal processing as well as the need for antiproton position measurements. Total materials costs have not increased.

**Simulation and Theory Computing**

The mission of the Simulation Group is to support the modeling efforts of Fermilab experiment users in the Fermilab computing environment. This goal is achieved by ensuring that the necessary simulation tools are available and operational in this environment, and by providing the required expertise for their utilization. In addition, we develop simulation tools and for computationally demanding HEP theoretical and beam physics applications, by utilizing Fermilab's high performance, tightly coupled parallel clusters. During FY04, the six members of the Simulation Group were involved in the following projects/activities:

HEP simulation tools packaging and maintenance. We successfully made available HEP simulation tools developed here and elsewhere to the Fermilab user community. These tools were packaged and distributed using the Fermilab products database and were kept up to date with computing environment (operating systems/libraries) and physics developments. We also maintain user e-mail lists, to provide some minimal consulting service to the users.

PATRIOT. The PATRIOT (Physics Archives & Tools Required to Investigate Our Theories) project provides a comprehensive database of HEP simulated data samples under controlled and tested conditions. It focuses on Run II analysis issues, with the intension to expand to LHC and other physics. The project took off during FY04 and it was used by members of the D0 and CDF experiments. We helped organize and/or participated in many workshops regarding physics generator issues and their correct use. We also provided expertise to Run II HEP physics generator users.

Gean4. Geant4 (G4) is the detector modeling tool of choice for current HEP experiments. We provided up-to-date versions of G4 to the general Fermilab users, and provided application support and development for the CMS experiment. Our CMS support efforts included modeling and analysis of the CMS test beam data (essential to its successful use), and simulation infrastructure related to the CMS physics center. Due to lack of manpower we were not able to provide detailed help for the development of G4 applications for other experiments during the past year.

LQCD computing. We provided the necessary software infrastructure to utilize the Fermilab lattice QCD (lQCD) parallel computing facilities. Our responsibilities included physics input selection, running scientific applications, analysis of the results and distribution to the lQCD SciDAC collaboration. The cluster was fully utilized during FY04. We also participated in design work done by the metadata and middleware working groups of the International Lattice Data Grid. Our participation was partially funded by SciDAC.

Accelerator Modeling. We continued developing Synergia, an accelerator modeling framework with collective effect simulation capabilities. The framework focuses on solving beam physics problems requiring the use of tightly coupled parallel computers, complementing the Accelerator Division modeling efforts. Synergia incorporates existing beams physics libraries when possible and is portable to both super computers and parallel PC clusters. We used Synergia to develop a simulation of space-charge effects in the Fermilab Booster. Synergia was also used to model the A0 flat beam transformer. The work on this project was mainly funded by SciDAC.

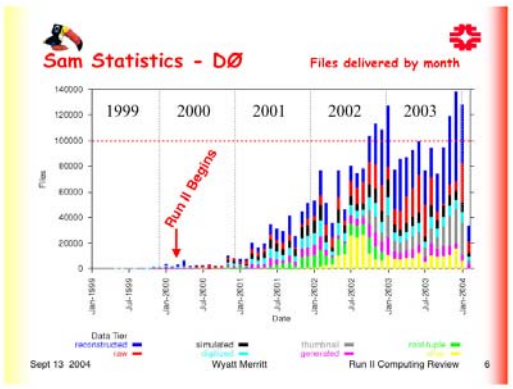
**Grids for Physics**

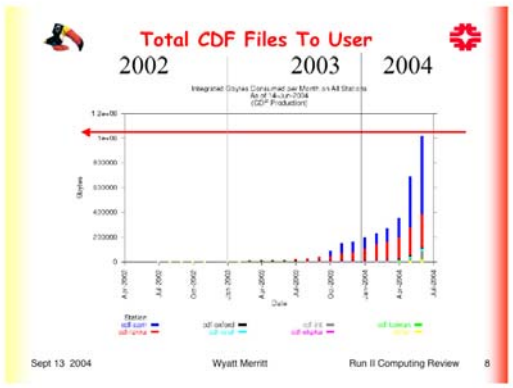
The Fermilab CD works with experiments and physics projects software and computing groups to build and operate their data storage, access and analysis systems. These distributed systems are now increasingly based on Grid technologies and principles. The CD and the physics software and computing projects are active contributors to the developments and collaborations that are bringing the Grid to fruition as a reliable, production quality infrastructure.

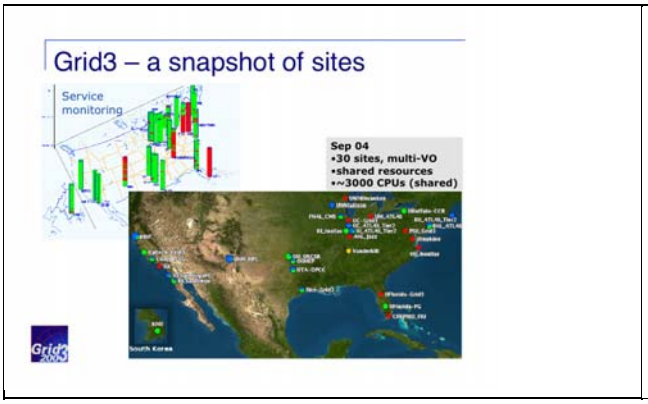
The goal for the SAMGrid data management project for 2004 was to have deployments for CDF, support reprocessing for D0, to continue support for normal data handling and simulation for D0 and to have the first production deployments for Grid job submission mechanisms. SAM is now installed at 40 DØ sites and 26 CDF sites, and over a peta-byte of data has been moved between the tape and disk based storage systems and caches [ Figure 1]. The reprocessing was successful. Grid based data movement is in use at all sites. Grid scheduling is in production at about 10 sites in D0, but not yet at CDF.

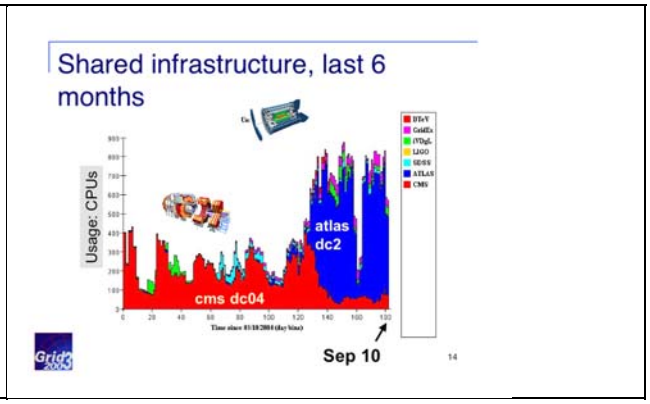
The Grid2003 project met the initial metrics in November 2003 for the number of CPUs and jobs. Grid3 has grown over the past year to 30 sites, over 3000 cpus and support for more than 10 different applications. By providing opportunistic use of their resources to each others' simulations US ATLAS and US CMS experiments have gained up to 50% in throughput. The ability to have central submission of jobs across a set of ~20 computing sites has enabled CMS, for example, to reduce the effort required to manage the jobs from 2 to 1 FTE.

CD has several projects to develop needed capabilities for Grid computing: the common Runjob job management software was integrated as planned into the CMS software infrastructure and is being integrated into the production systems at D0. The system for Virtual Organization (person) registration was completed, is in test for US CMS and under evaluation by the CERN LHC Computing Grid Project. Extensions to support group based roles and site authorization policies are under development as planned. The Particle Physics Data Grid project Year 3 milestones were met and an extension proposal funded for 2 more year by the DOE. The Quarknet-Grid education project is working to provide a "virtual data" based repository for sharing of cosmic ray data. Use in schools was not yet achieved, due to the immaturity of the middleware and need to provide a robust very easy to use interface for students.






We need to extend the capabilities and effectiveness of the common infrastructure by about an order of magnitude over the next 5 years. The Open Science Grid Consortium is proposed to build a sustainable, production quality Grid infrastructure for the distributed US scientific community (from the local to the remote and poorly connected) and to partner with other such infrastructures both in the US and internationally. Initial meetings and planning for the consortium have taken place.