Fermi National Accelerator Laboratory

FERMILAB-TM-2109

The Fermilab Computing Farms in 1999

Marina Albert et al.

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

May 2000

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CH03000 with the United States Department of Energy

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Distribution

Approved for public release; further dissemination unlimited.

Copyright Notification

This manuscript has been authored by Universities Research Association, Inc. under contract No. DE-AC02-76CH03000 with the U.S. Department of Energy. The United States Government and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government Purposes.

The Fermilab Computing Farms in 1999

Marina Albert, Troy Dawson, Jim Fromm, Lisa Giacchetti, Terry Jones
Tanya Levshina, Igor Mandrichenko, Ramon Pasetes, Marilyn Schweitzer
Karen Shepelak, Dane Skow, Stephen Wolbers

April 10, 2000

Introduction

The farms in 1999 changed in two major ways. First, PC's running Linux continued to expand and this allowed for the reduction of the SGI and IBM components of the farms. Second, the first large farms for CDF and D0 run II were purchased and installed in 1999. Simultaneously, a large increment for non-Run II computing was made.

The farms continue to provide large CPU resources for those experiments and calculations which benefit from this type of computing (large CPU, low I/O, dedicated resources). Farms usage will continue to increase given the demands of the user community (reflecting the scientific program) and the preparation for and beginning of Run II.

The year in review

The original PC farm (fnpc1-37) continued to provide large computing for the user community during 1999. These systems ran well and required little maintenance during the year. The I/O system for these machines (fnsfh) was augmented to provide adequate disk storage and tapedrives but otherwise was stable. The success in using the PC farm allowed us to decommission all of the IBM farms, including its I/O node and all of the single-processor SGI farms. The 4-processor O200 farm nodes were kept as part of the farm.

A major acquisition of farm computing was made in the summer/fall of 1999. 150 dual-Pentium III/500 MHz PC's were purchased after extensive pre-qualifying and an RFP. The resulting purchase consisted of 100 PC's from HiTec and 50 from Eternal Graphics. The PC's came with 1-6GB and 2-18GB diskdrives, 512 MB of memory, and 10/100 Mbit ethernet interfaces. The default allocation is for 50 nodes

to be dedicated to CDF, 50 to D0 and 50 to the rest of the laboratory's scientific program. In addition I/O nodes for the workers were acquired or upgraded. An SGI Origin 2000 was upgraded to serve as the I/O node for the 50 "general" nodes, an SGI Origin 2000 was acquired for the D0 I/O node and an SGI Origin 2200 was acquired as the CDF I/O node. Having a similar architecture for the I/O nodes is important to help reduce the support load and increase the ability to debug and solve problems in a common way across the farms. In addition, major switch acquisitions and installations were made during 1999 to give the worker and I/O nodes connections to the network with the required bandwidth.

The utilization of the farms during 1999 was lower than usual. The demand for the computing was not high enough during 1999 to use the full resource. Nevertheless, E871 finished their reconstruction of the 1996-97 data during 1999. This was a successful reconstruction and allows them to analyze the large sample of data that was acquired and get ready to reconstruct the data that was taken in 1999. E831 finished up some reprocessing and second-pass processing on the farms in 1999. No other major data reconstruction occurred in 1999 on the farms.

NuTeV ran simulation jobs during the last part of 1999 and made heavy use of the PC farms. A large sample of events was generated and this generation continues. Auger used some of the farms nodes to simulate high energy air showers. This computing is also run on FNALU. Most of the Auger work on the farms finished in mid-1999 but it may return if necessary. NUMI made small use of the farms for various calculations.

Two other large calculations were performed on the farms in 1999. The first was a calculation of beam dynamics in the Tevatron. This calculation took some time to optimize for the farms environment, due to some synchronization issues, but by the end of the year it was running smoothly on the 5 O200's left in the farm. There was a short but significant calculation performed by Walter Giele of the Theory Department during the middle of 1999.

CPU utilization

Table 1 provides the summary of CPU time (in VUP-equivalent units and in SpecInt95) for the whole farm in 1999. A plot of the CPU utilization (including all previous years of the UNIX farms) is shown in Figure 1.

Table 2 and Figure 2 show the utilization for each of the many experiments that have used the farms during 1999. E871 was the largest user of the farms. NuTeV and Auger simulations were also significant users of CPU on the farms. Theory

calculations, E831 cleanup, and Tevatron calculations were the remaining large users of the farm in 1999.

Table 3 is a sum of all the CPU time used by all the experiments that have used significant CPU time on the farms during the last 9 years, along with the totals used by each. E831 is the largest farm user, with E871 second. Given the huge new capacity that has been added to the farms in 1999 (84,600 MIPS or 6180 SpecInt95) the list of integrated CPU will change dramatically once the new systems are fully integrated and utilized.

SpecInt95 and MIPS

The unit that we have been using to report utilization (the MIP) was and continues to be defined by us by running a simple program (called TINY). The program is an HEP track generator which subsequently performs pattern recognition and fits the tracks. This program has served us well, especially when the vendors' claims were exaggerated, highly tuned and otherwise not always trustworthy. The SPEC suite (see www.spec.org) improved and has been able to provide reasonably reliable and complete ratings for machines. We have decided to report 1999 in SpecInt95 units, as well as MIPS. SPEC has a new unit, Spec2000, which is just now being used. It is not clear yet how soon Spec2000 will become useful for the purposes of reporting CPU time. For now a switch to SpecInt95 will be made using published numbers. Table 4 gives the 1999 utilization and the integrated 1991-1999 utilization in SpecInt95 units. The factor used to convert from MIPS to SpecInt95 is 0.062653. Please note that this factor is good only for the farms and its mix of CPU types. Every other system at Fermilab will have to use its own conversion factor to connect old (MIPS or VUPS) data with new (SpecInt95) data.

Table 1 – Total CPU use on the Farms – 1999

Month	CPU delivered (Vax-Months/month)	(SpecInt95)
January	11315	709
February	9247	579
March	5518	346
April	7987	500
May	9084	569
June	6928	434
July	8453	530
August	6349	398
September	3750	235
October	4293	269
November	6088	381
December	4916	308

Table 2 – CPU use by experiment – 1999

Experiment	CPU time (Vax-years)	(SpecInt95)
E871	3333	209
NuTeV	1280	80
Auger	834	52
Theory	568	36
E831	459	29
Beams	259	16
E872	32	2
NUMI	26	2
TOTAL	6791	425

Table 3

Integrated Farm Use
(In units of MIP-years)
Through December, 1999

Experiment	1991	1992	1993	1994	1995	1996	1997	1998	1999	Total
E831							234	5988	459	6681
E871							191	1349	3333	4873
D0 offline		59	570	1072	1184	1743	96			4724
E706	28	82	732	992	795	686	19			3335
E791		100	1232	1249	214	11				2806
Auger							880	1023	834	2737
CDF		110	320	438	752	956				2576
E781							302	1566		1868
NuTeV									1280	1280
D0 MC	101	162	396	197	108	23				987
E665	14	105	733	91	2	10				955
Theory							309		568	877
E771		94	211	339	219					863
Beams							105	112	259	476
E866						55	409	8		472
NUMI								412	26	438
E789		156	247							403
E687	99	235	29	30	2					395
Minos							221			221
Recycler						61	130	19		210
E872							21	60	32	113
E760		54								54
E835								52		52
E731	38									38
Magnet							28			28
Total	267	1156	4541	4408	3276	3545	3217	10707	6791	38908

Table 4

Integrated Farm Use
(In units of SpecInt95-years)
Through December, 1999

Experiment	1999	Total
E831	29	419
E871	209	305
D0 offline		296
E706		209
E791		176
Auger	52	171
CDF		161
E781		117
NuTeV	80	80
D0 MC		62
E665		60
Theory	36	55
E771		54
Beams	16	30
E866		30
NUMI	2	27
E789		25
E687		25
Minos		14
Recycler		13
E872	2	7
E760		3
E835		3
E731		2
Magnet		2
Total	425	2438

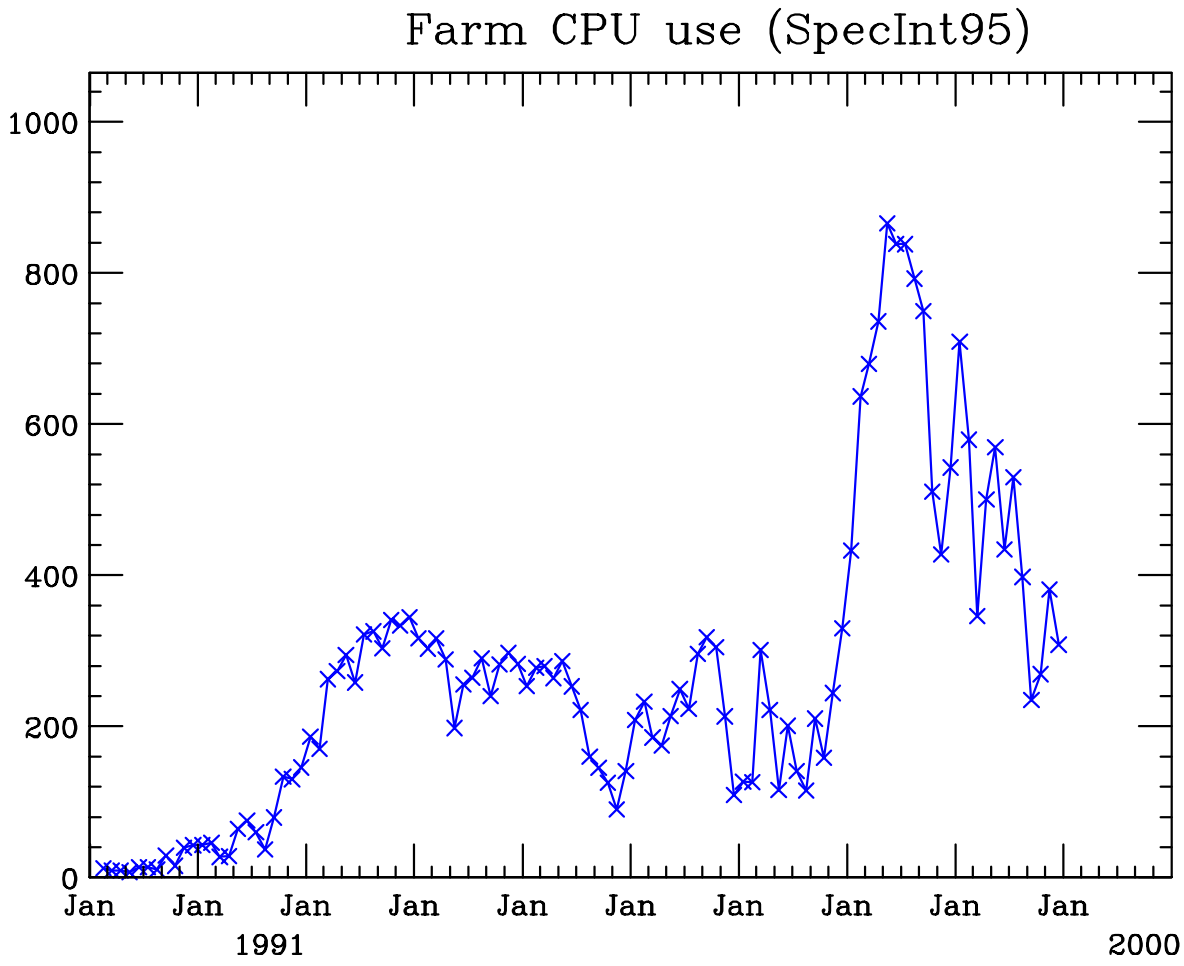
Farm CPU use (SpecInt95)
1000
800
600
400
200
0
Jan
Jan
Jan
Jan
Jan
Jan
Jan
Jan
Jan
Jan
1991
2000

Figure 1.

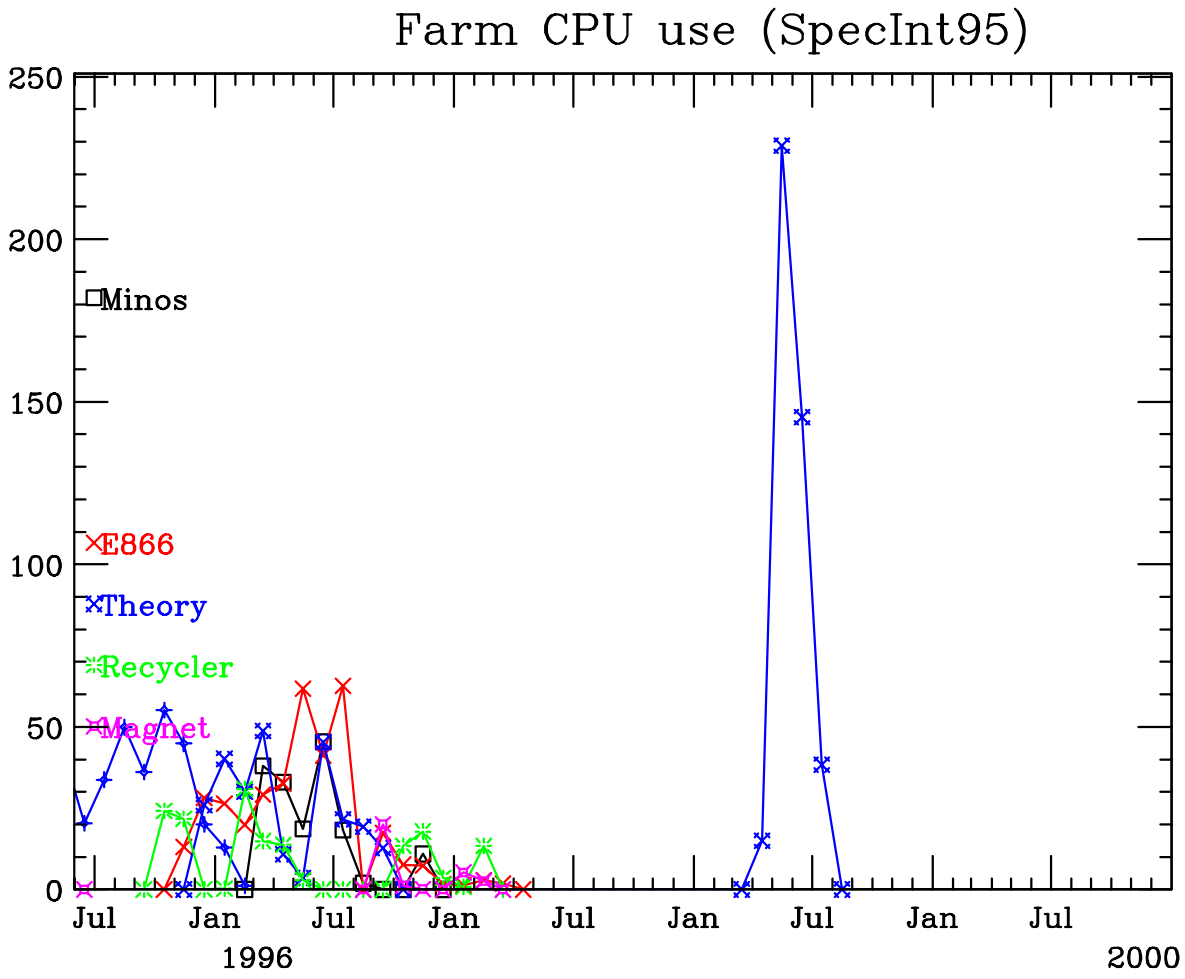
Farm CPU use (SpecInt95)
250
200
150
100
50
0
Minos
E866
Theory
Recycler
Magnet
Jul
Jan
Jul
Jan
Jul
Jan
Jul
Jan
Jul
1996
2000

Figure 2(a).

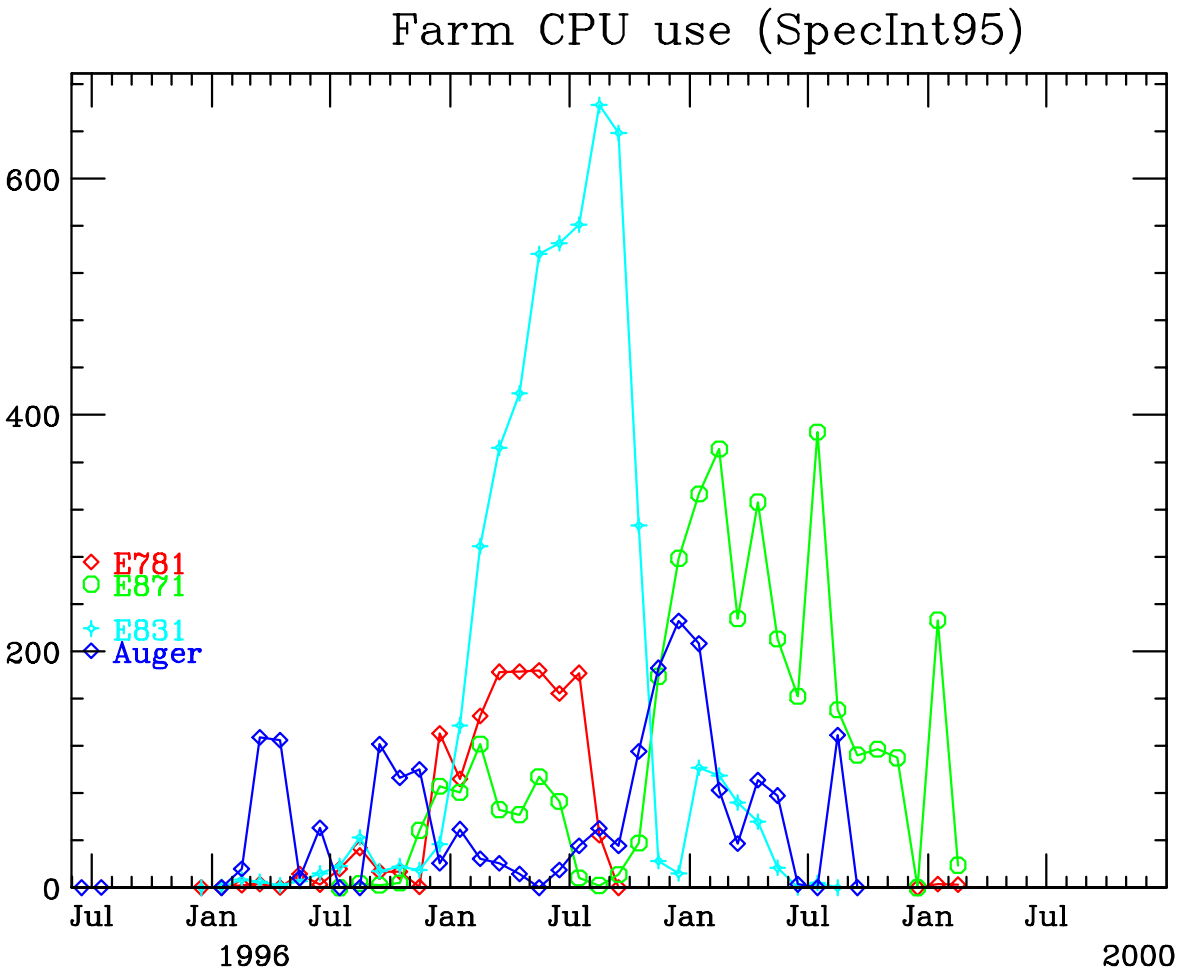
Farm CPU use (SpecInt95)
600
400
200
0
E781
E871
E831
Auger
Jul
Jan
Jul
Jan
Jul
Jan
Jul
Jan
Jul
1996
2000

Figure 2(b).

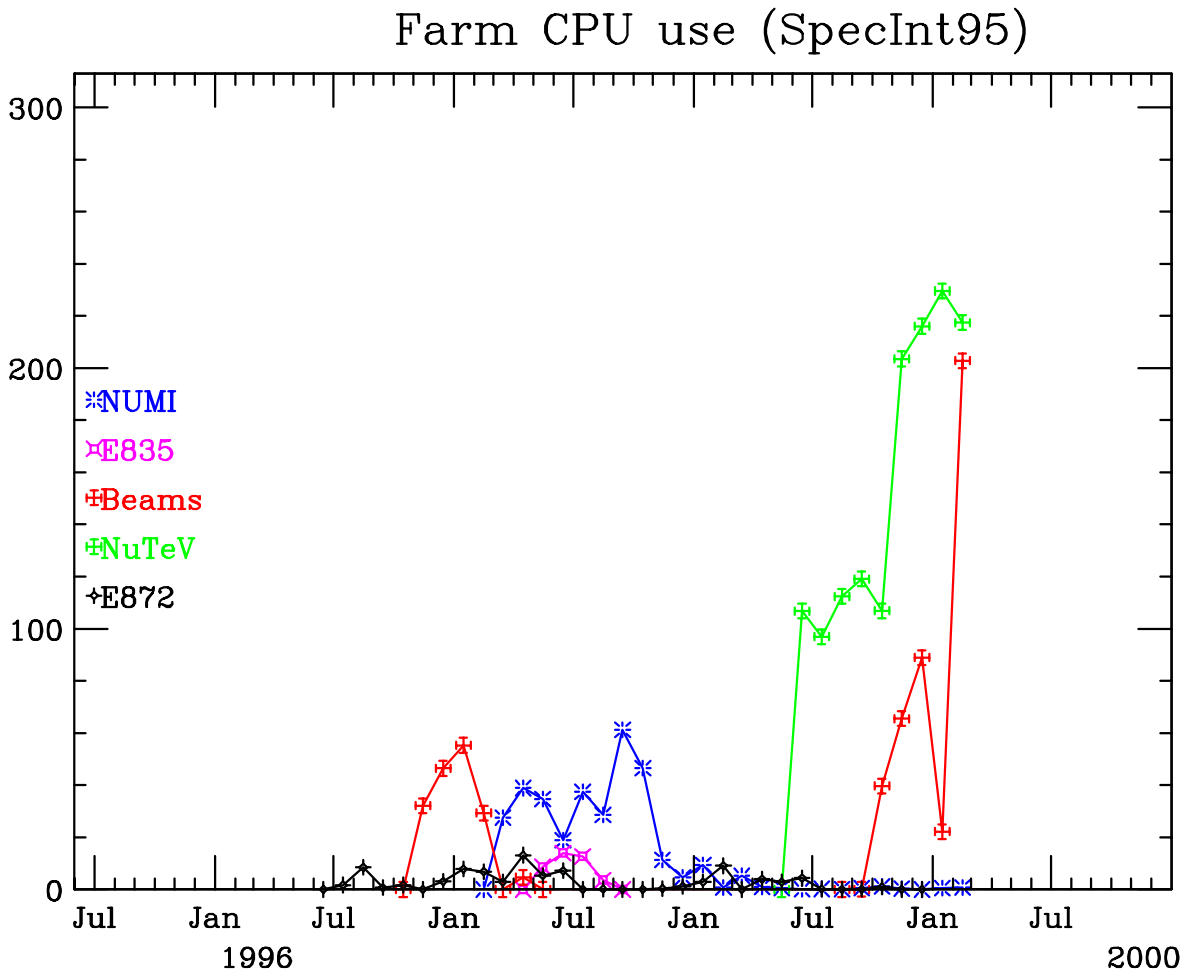
Farm CPU use (SpecInt95)
300
200
100
0
NUMI
E835
Beams
NuTeV
E872
Jul
Jan
Jul
Jan
Jul
Jan
Jul
Jan
Jul
1996
2000

Figure 2(c).

Allocations

For completeness the allocations of worker nodes as of March 24, 2000 is given in tables 5-6. These allocations are based on the laboratory's priorities, the ability of the group to use the resource and of course the need of the various groups for computing. These allocations are shifted fairly often to reflect the changes in the circumstances of the groups as they (hopefully) finish what they need to do and move on to other tasks.

Table 5
Current Farm Allocations

Experiment	**Allocation in MIPS**	**Allocation in SpecInt95**
CDF	28200	1767
D0	28200	1767
BTeV	15228	954
KTeV	8460	530
E871	6696	420
Beams	4160	261
NuTeV	4092	256
E781	3384	212
NUMI	2232	140
Linear Collider	1128	71
TOTAL allocated	101780	6378

Table 6
Current Farm Allocations
All Nodes

Expt	O200	PC(333)	PC(500)	MIPS	SpecInt95
CDF			100	28200	1767
D0			100	28200	1767
BTeV			54	15228	954
KTeV			30	8460	530
E871		36		6696	420
Beams	5			4160	261
NuTeV		22		4092	256
E781			12	3384	212
NUMI		12		2232	140
Linear Collider			4	1128	71
unused or test		2		372	23
TOTALS	5	72	300	102152	6400

Plans

2000 will see the ramp up of the use of the 150 new PIII/500 PC's which were purchased in 1999. 50 of them are connected to the general purpose or "fixed-target farms", 50 of them are dedicated to CDF and 50 to D0. The 50 new and 37 original general purpose farms will be rapidly allocated to a combination of E871, E781, NuTeV, BTeV, KTeV, Beams and Linear collider calculations. The 5 O200's may be moved to FNALU to provide CPU in that environment or they may be decommissioned.

The CDF and D0 farms will be dedicated most of the time to those experiments for testing of rates, functionality, mock data challenges, generation of simulation samples, etc. Additional nodes will be purchased in late 2000 or very early 2001 for the beginning of the run, scheduled for March 1, 2001.

Additional nodes for general purpose users may be acquired during 2000, though this depends on the utilization and demand from the various scientific users at the lab. There are potentially very large demands from BTeV, Linear Collider, Muon Collider and vlhc calculations.