Fermi National Accelerator Laboratory

FERMILAB-Conf-98/361

PC Farms

G.P. Yeh

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

December 1998

Presented at the *International Conference on Computing in High Energy Physics (CHEP '98)*, Chicago, Illinois, August 31 - September 4, 1998

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

FERMILAB-Conf-98/361
September, 1998

PC Farms

G.P. Yeh

Fermi National Accelerator Laboratory, Batavia, Illinois

Academia Sinica, Taipei, Taiwan

Abstract

In recent years, PCs have achieved performance comparable to the high-end UNIX workstations, at a small fraction of the price. Commodity PCs have become the solution to CPU needs, even for large scale computing in high energy physics, nuclear physics, space sciences, and many other fields. We report on their development and applications, and look forward to the important and exciting future of large scale PC computing.

1 Introduction

Since the introduction of the Intel Pentium Pro processors in 1995, the high performance and low cost of PCs (relative to RISC processors and UNIX workstations) have attracted many groups to test clusters of PCs for cost-effective parallel and distributed computing. By April '97, a dozen groups from NASA, DESY, Fermilab, Sandia, NIH, other labs and universities reported [1] their performance benchmarks for various applications, some with plans to build clusters with thousands of processors.

In high energy physics, by its nature of independent collisions, the event reconstruction from data and the Monte Carlo simulation are excellent tasks for loosely coupled parallel computing. Fermilab pioneered 10 years ago such parallel computing "Farms" as an alternative to the costly mainframe computers. The Fermilab UNIX Farms at their peak had a total of about 500 UNIX workstations.

With the severe CPU needs in the future, it is natural to evaluate using PCs, with their high performance and low price, as the most cost-effective solution. The advances in CPU chips are accompanied by similar great advances in commodity network and memory. The PC Farms and general PC computing also importantly provide an upgrade path.

1.1 Large experiments and CPU needs in next 10 years

Many major experiments in high energy physics and nuclear physics will have large, challenging computing needs. The three large experiments at three B Factories, HERA-B(DESY), Belle(KEK), BaBar(SLAC) are on schedule to start in early '99 to collect large amounts of data for CP violation studies. The BRAHMS, PHENIX, PHOBOS, and STAR experiments at RHIC/BNL plan to start heavy ion collisions in June, '99, with one of the main physics goals being to study quark-gluon plasma.

At Fermi Lab, CDF & D0 experiments will have data taking period Run II in 2000–2002, to study the newly discovered Top Quark, and search for B CP violation, Higgs, SUSY. The amount of data collected in Run II is expected to be 20 times that of Run I. The CPU requirement will be $N \times 10^5$ MIPS or 10^4 SPECint95 (Fermilab defines CPU equivalent to a VAX-11/780 to be 1 MIPS, and SPECint95 = 40 SPECint92 = 13 MIPS = 10 CERN_Units). The amount of data collected in Run III in 2003 –2005, is expected to be 15 times that of Run II.

For the Large Hadron Collider starting in 2005, the Atlas & CMS experiments have estimated CPU needs of $N \times 10^7$ MIPS or 10^6 SPECint95. There are strong, active Linear Collider study groups in the 3 regions of Asia, Europe, and North America. Their hope is to complete by 2008 a Linear Collider with (Phase I) center-of-mass energy of 500 GeV, which can later be upgraded up-to 1,500 GeV. There are also study groups for the longer term Muon Collider and the Very Large Hadron Collider.

1.2 Loosely-coupled parallel computing

At Fermilab, to fulfill the CPU needs at an affordable price, the ACP systems were used in '87-'89 for CDF Level 3 and Offline, and also for other experiments. These systems were based on commercial (Motorola 68020) CPUs, and hundreds of custom boards. The ACPMAPS system was developed for more tightly coupled lattice QCD calculations (http://hppc.fnal.gov/acpmaps/acpmaps.html). The system had 5 Gflops with 256 Weitek XL-8032 processors in '88, and 50 Gflops since '92 with 612 i860 chips based on dual CPU boards.

For parallel computing on loosely-coupled processors, Fermilab developed CPS (Co-operative Processes Software). Other inter-processor communication software such as PVM and MPI have been developed, supported, and widely used by other labs.

UNIX workstation farms have provided the main source of CPU at Fermilab and other labs since '91. The use of commercial UNIX workstations and Operating Systems instead of the custom board based systems such as the ACP provided the ability to upgrade the systems easily. To satisfy the large increase in CPU requirements in the future, however, a change in hardware platform is necessary.

1.3 PC related talks in CHEP '95

CHEP '95 was in Rio de Janeiro, in Sept. '95 (www.hep.net/conferences/chep95/). In its Systems and Facilities Session, there were the presentations "PC as Physics Computer for LHC ?" [S. Jarp], and NICE, desk top cluster [D. Foster]. The Session Chairman M. Miranda noted: 1) the importance of good network and its progress, 2) the increasingly important role of parallel processing in high performance computing, 3) PC's commodity market with good price/performance, 4) Commercial products becoming more suited for HEP, 5) proprietary Microsoft vs. Open Systems in UNIX world.

1.4 PC related talks in CHEP '97

CHEP '97 was in Berlin, in April '97 (www.ifh.de/CHEP97/). There were several talks on PC support and applications: WNT at D0 [S. Fuess], Compiler [S. Jarp], Dynamic Libs [V. Fine], FNAL PC Support [M. Kaletka], DESY WNT [T. Poser], CERN WNT [A. Pace], Linux Lab [G. Organtini], WNT/Linux PC [M. Athanas], and two PC farms talks: Hermes Linux PC Farm [W. Wander], FNAL Linux PC Farm [S. Wolbers].

1.5 Pentium-Pro Cluster Workshop, April '97

The First Pentium-Pro Cluster Workshop, April '97 (www.scl.ameslab.gov/workshops/PPCworkshop.html) was held in Des Moines (hosted by SCL at DOE Ames Lab) at the same time as CHEP '97. Representatives from NASA, Caltech, Fermilab, NIH, and other institutions shared experiences and successes with Pentium-Pro cluster computing. The Linux / NT ratio was about 3 / 1.

1.6 Super Computing '97

The Super Computing '97 Conference (www.supercomp.org/sc97/) was in San Jose, in November '97. Several groups from several labs and universities exhibited "Beowulf projects" or PC Farms. Almost all of these groups used Linux as Operating System. Fermilab and CDF demonstrated using PC Farms for Online Level 3 trigger and for Offline Reconstruction.

2 Non-HEP systems

2.1 Beowulf Projects

Among HEP & DOE labs, a loosely coupled computer system is known as a "Farm" Following the UNIX farms, the PC based systems are "PC Farms". The NASA Beowulf Project started in Summer '94 at CESDIS to produce software for off-the-shelf clustered commodity PC-class hardware, using a high-bandwidth low-latency network and open source Linux operating systems & message-passing. The system should provide super computer performance at a fraction of the cost ("GigaFlops at $ 50K"). An important contribution to the PC based parallel computing is the Linux device drivers for Ethernet by Donald Becker at CESDIS. Examples of the dozens of existing Beowulf projects/systems are shown in Appendix A. There are Pentium-Pro Cluster "Cookbooks". The recipes and products from NASA are called "Do-it-Yourself Rocket Science", and we find that "We are not alone" in PC based parallel computing and building PC farms.

2.2 Linux Alpha

In the movie Titanic, Digital Domain did the visual effects simulation on 100 Alpha nodes running Linux. The Avalon Linux Alpha cluster at Los Alamos has 70 nodes. The University of Wisconsin-Milwaukee used 48 Linux Alpha nodes to process data for the LIGO Experiment.

2.3 ASCI Teraflops Supercomputers

For large scale parallel computing, Teraflops supercomputers ASCI Red at Sandia has 9000 Intel P-Pro-200 processors, while ASCI Blue Mountain at LANL is based on SGI Origin 2000 clusters, and ASCI Blue Pacific at LLNL is based on IBM SP-2. Development and progress are being made from Teraflops supercomputers towards Petaflops super-supercomputers.

3 Linux Operating System

Linux is the Operating System of choice for both the non-HEP Beowulf projects presented in the last section and the high energy physics PC Farms described in the next section. In recent years, there has been a very large UNIX base in high energy physics and other scientific fields. Linux is "generic UNIX", and enables easy, quick migration from UNIX workstations to PC computing. Http://hepwww.ph.qmw.ac.uk/HEPpc/ is a useful reference with many links.

Linux[2] is an open source, well supported, widely used Operating System with multitasking capability. Most of the Linux system infrastructure (libraries, compilers, utilities) outside of the kernel comes from the Free Software Foundation (GNU). Codes written in the C language can be compiled using GNU CC (gcc), and Fortran codes are often compiled using g77. Other compilers are also available (e.g. Microway, Absoft, f2c converter, g++, egcs, and KAI).

Many of the software packages that physicists use as part of everyday work are available for Linux. These include TEX/LATEX, editors (Nedit, EMACS), TCP/IP, X11R6, Netscape, HTTP, FTP, Telnet, SLIP, CERNLIB, etc. Linux is a complete, copylefted UNIX clone. It runs on many platforms, including Intel 386/486/Pentium/P-Pro/P-II, DEC Alpha, MIPS R4x00 and SPARC/Fujitsu(AP+).

Various pieces of the Fermilab UNIX environment have been ported as well, including UPS (a configuration and product management system), UPD (a product distribution system), and CPS (Cooperative Processes Software). Porting of software to Linux has in general been straightforward. Useful porting advice is readily available [3],[4]. Linux is now used in many experiments. In most cases, the porting takes only a "few weeks".

3.1 Physics Code

Versions of CERNLIB with different compilers (v95b(f2c), v96a(g77), v96a_absoft, v97a, v98, v99) are available from CERN FTP (ftp://asisftp.cern.ch/).

Data analysis in high energy physics may require a lot of Monte Carlo studies, which need large samples of generated events with detector simulations. Currently PYTHIA, ISAJET, GEANT, VECBOS, NJETS, JETRAD, DYRAD and other programs can all be run on PCs under Linux. PC Farms can be an excellent and cost-effective platform for Monte Carlo studies.

In CDF, for example, most of the offline codes, enough for Run I Production (UNIX in Run I), event display, and Run II codes have been ported to Linux. Most modifications are on OPEN() statement and system calls. The byte-swapping from UNIX is similar to reading/writing VAX format. We use the # derivatives of C Preprocessor (cpp). The g77 Fortran compiler automatically calls cpp to preprocess *.F before compilation. Linux and PCs are now widely used for Run II CDF code development. With the successful porting of the reconstruction and analysis software packages to Linux, the collaborators will be able to do analysis from their PCs at home or in offices.

We also use PC Farms for online level 3 triggers. The open source nature of Linux has been particularly useful in testing, understanding and improving the Level 3 PC Farm.

4 Price/Performance

The main reason for PC computing is its high performance at low price. By Summer '98, the price is less than \$7 per MIPS (CPU equivalent to a VAX-11/780). The Price/Performance is decreasing at the rate of about 1% per week.

The 200 MHz Pentium-Pro chip has a processing power of 113 MIPS as indicated by timing studies using our benchmark programs Pythia and Geant, "Tiny", and the CDF offline reconstruction code. For comparison, the corresponding power of a R4400 node on SGI Challenge was 83 MIPS. Our tests, as well as tests done by others, show that the performance of the newer P-II chips scale with the clock speed. The 400 MHz P-II is twice as fast as the 200 MHz Pentium-Pro. The doubling in speed every 12-18 months is expected to continue, with GHz chips planned by Intel for the year 2000.

4.1 Dual Processor PC

A dual processor PC running e.g. Linux, which has Symmetric Multi-Processing (SMP) capability, is twice as fast as single processor PC, for HEP Monte Carlo or CDF Data. The dual Pentium-Pro 200 MHz PC had the same combined speed as SGI R10000 (which has about 2.5 times the speed of R4400). By February '97, the Price/Performance of dual Pentium-Pro PC was more than a factor of 3 better than a R10000 workstation.

In CHEP '97 and the Pentium-Pro Cluster Workshop, the PCs had P-Pro 200 MHz and DRAM (66 MHz) at \$4 / MB (vs. \$40 / MB for UNIX Workstations). In CHEP '98, we see clusters with 400 MHz P-II and SDRAM (100 MHz) at \$1.5 / MB. The price/performance came down by a factor of 2 between CHEP'97 and CHEP'98.

It would be great if PCs with higher number of processors per node could be competitive in price/performance, since multi-processor nodes could reduce floor-space and system maintenance, and even improve performance. For now, the dual processor PCs offer the best price/performance.

Combining many PCs provides the computing power of super computers at a small fraction of the cost. In the next section, we describe the wide spread usage of PC Farms in the high energy physics and nuclear physics community.

5 PC Farms in HEP

PC computing related talks in CHEP'98 are shown in Appendix B. The number of talks and the scale of the PC Farms have increased by a factor of several since CHEP '97.

5.1 DESY

The HERMES experiment has been successfully running a Linux Production Farm since Sept '96, with 10 Dual P-Pro PCs. This was one of the earliest PC Farms.

The ZEUS experiment has 20 Linux PCs, which processed 30% of ZEUS data (10 Million events) in '97. The system has reprocessed data during the '97-'98 HERA shutdown, and has also been used for data analysis.

The HERA-B experiment, starting in '98-'99, has 100 P-II Linux nodes installed for Second Level and Third Level Triggers. The Fourth Level Trigger currently has 10 Linux nodes, with a goal of 100-200 nodes.

The Linux-PC Farms at DESY Zeuthen have been running since February '98. The three farms from AMANDA, H1, and HERA-B have 40 PCs, which doubled the CPU for the experimental groups. The installation of Linux and Post-Installation for setting up a farm PC (e.g. batch system, file server access) and for setting up DESY-Zeuthen environment (e.g. AFS, user environment) are all via network.

5.2 RHIC at BNL

The 4 experiments BRAHMS, PHENIX, PHOBOS, STAR starting in June '99 will record 1.5 PB of raw data per year, requiring 10,000 SPECInt95 (500 P-II 400 MHz) of CPU. The Mock Data Challenge has started in Sept. '98. The Linux PC Farm with 40 dual Pentium II is a 10% prototype for the full system.

Linux PC farms will be based on either the commodity PCs, or custom 8-processor nodes being evaluated. The in-house control software is written in perl. Remote Management of individual PCs and a console multiplexing system, as well as a batch processing system are available. RHIC Computing Facility (RCF) is planning 3 additional farms: Central Reconstruction Server Farm, (part of) Data Mining Farm, and (part of) Central Analysis Server Farm.

5.3 CERN

Most experiments at CERN still rely on UNIX workstations, since LEP experiments have been running for many years, and plan to stop running in 2000. The PC based Simulation

Facility running Windows NT (PCSF) began in Summer '97, with 35 dual processors in spring '98. Some tools have been developed to deal with this OS. PCSF is using LSF, a commercial batch scheduler. Scalability issues are being studied. The system now allows for some reconstruction, which is more I/O demanding.

There are also Linux PC Farm and Linux PC clusters, which are used in Chorus, NA48, NA49 and other experiments (wwwinfo.cern.ch/dci/linux/).

It has been estimated (J. Knobloch in www.ifh.de/CHEP97/) that the CPU requirement for the Atlas experiment for Event Reconstruction, Monte Carlo Production, and Physics Analysis will be 7×10^4, 5×10^4, and 15×10^4 SPECint95, respectively. The Online triggers will have even higher CPU requirement. Similar numbers are expected for CMS. PC Farms and PC computing are clearly the leading solutions to such CPU challenges for LHC.

5.4 Jefferson Lab

The 3 experimental areas record 1 TB of raw data per day. To process this large amount of data, the Linux PC Farm has 10 dual P-II 300 MHz (in production), 20 dual P-II 400 MHz (adding now) , and 20 more Dual P-II (end of '98). Currently the PC Farm has $\geq 85\%$ usage. There is also a small Linux PC interactive cluster.

5.5 Other labs

At KEK, the Online groups have provided software for Linux (VMElib). The Belle experiment officially supports and has ported their software to Linux. K2K and other experiments use Linux for DAQ.

At SLAC, for BaBar, the officially supported Operating Systems have been AIX, Solaris, OSF, and HPUX, but the BaBar software has been ported to Linux. At RAL, an NT Farm is operating on 11 Pentium-Pro nodes . There are also small Linux PC farms/clusters at LBL and Argonne, and many universities and other labs.

6 Fermi Lab

At CHEP & the Pentium-Pro Cluster Workshop in April '97, Fermilab and CDF presented PC Farm benchmarks based on studies using Pythia, Geant, CDF Run I Production, and the CDF Event Display. At the Super Computing Conference in November '97, Fermilab and CDF exhibited a CDF Offline Production PC Farm and a CDF Online Level 3 PC Farm. For the Online Level 3 PC Farm, the initial discussion was in January, the hardware was ordered in June and arrived in August. The system with simplified Run I Level 3 code ran successfully by October '97. The Demo systems ran at San Jose Convention Center (trucked to San Jose and back) and ran in Fermilab HighRise lobby during Dec '97 - Feb '98. The systems are robust. By '97, the Fermilab Theory Group, and other experiments such as E815, E871, and SDSS also took advantage of PC computing using Linux.

In January '98, Fermi Lab's Computing Divison announced official support for Linux (www.fnal.gov/cd/unix/linux). The Computing Divison selected/incorporated the RedHat distribution, and makes recommendations and guide lines for Fermilab users on all aspects of the PC computing.

6.1 CDF Level 3 Trigger

With the 20 fold increase in luminosity and data for Run II, the online Level 3 trigger will need to input, select, and output data at rates about 20 times Run I rates. The Level 3 trigger for Run I required 3,000 MIPS and was based on 4 SGI Challenges. PC Farms provide a cost-effective solution to online Level 3 challenges for Run II and beyond.

I/O performance may be a major concern for Level 3 PC Farm configurations. CDF has demonstrated and adopted the solution of using ATM switch distributing data to Linux PC Farms. The ATM switch has 16 inputs and 16 outputs. The event size is expected to be 200 KB. It has been measured and demonstrated (with 5 PCs) that each output can perform at 15 MB/s (while the peak requirement is 3.8 MB/s), for a system total of 240 MB/s. The Level 3 Trigger using PC Farms can process 4 times the Run II TDR rate requirement. The output rate is limited mainly by the offline storage, accessing and analysis of data. The system can easily be upgraded for Run III.

The open source nature of Linux has been valuable, especially for Level 3, to test, understand, and improve the performance of the system. The system has also been demonstrated to be stable. Additional 20 dual PCs are being installed to further test the I/O, stability, and scalability of the system, and to exercise Level 3 trigger software.

6.2 D0 Level 3 Trigger PC Farm

For Run II, D0 Level 3 group plans for 50 dual or quad processor nodes, each node with 128-256 MB memory and 2-4 GB local disk, running Windows NT. The system should be able to process at an input rate of 1500-2000 Hz (designed for 10,000 Hz for Run III). The output rate is expected to be 50-70 Hz, with event size of 256KB.

Data will be distributed to each node via custom high speed hardware. The aggregate bandwidth is 192*3=576 MB/sec for the 3 initial segments of PC farms. This is based on VME technology, and there is a PCI-VME bridge to get the data into the node (runs around 20-25 MB/sec for these VME boards). Custom distribution hardware is run by both Windows NT and Windows CE. Output will be gigabit ethernet or other commodity network. Individual nodes of a prototype farm are running, with supervisor processes running on the same machine. The custom distribution hardware is in the TDR stage.

6.3 Fermi Lab Plans for Offline Farms

Offline Production for experiments at Fermilab, including CDF and D0 Run II, will be accomplished by Linux PC Farms. By the year 2000, CDF & D0 Offline Production will require CPU equivalent to 285 - 500 dual P-II 400 MHz PCs.

In the last few months, Fermilab has been De-Commissioning UNIX workstation farms, and Commissioning Production PC Farms. Presently, the E871 Production PC Farm has 64 P-II 333 MHz processors.

The CDF/D0/Theory farm being installed starts with 18 Dual P-II 400 MHz processors, each with 256 MB SDRAM. Gigabit Ethernet is being evaluated for this system.

Several local "white box" companies have participated in the Fermilab purchasing/bidding process for the PC Farms. In the bidding and purchasing, Fermilab has specified hardware components such as the processor and motherboard. The PCs were delivered in 3 weeks, preinstalled with Fermi Linux, and other CERN, Fermilab, HEP software. The PCs were put together, including network, as a PC Farm within a few days of delivery.

6.4 Linux at Fermi Lab

The suggestions and instructions for distribution & installation of Fermi Linux are available (www.fnal.gov/cd/unix/linux). The software distributions are by free CD or download via network. CDF Run II software distribution is by similar methods. Fermilab users can also order PCs using the above web site, make hardware choices, and charge the order to a budget code. The PCs will arrive in about 3 weeks, with Fermi Linux pre-installed. Vendor support is available (one way is via contract with Red Hat).

To help improve the usability of PC computing, the Computing Division is offering Linux Administration (day long) classes. The monthly UNIX Users' Meeting has become the Linux Users' Meeting.

6.5 Recent port & support for Linux

Linux is officially supported by the leading companies providing database software, Oracle, Informix, Computer Associates, Sybase, IBM, Objectivity, and other companies such as Netscape, Corel, Interbase, Adaptec, Cygnus, Sun, SGI, and DEC. In September '98, Intel, Netscape, together with venture capital firms Greylock and Benchmark Partners, acquired equity positions in the Red Hat company. The strong support from such leading companies greatly strengthens future development of Linux.

6.6 PC Computing

In addition to the PC Farms, desktop computing is also migrating from UNIX workstations to PCs. Many related issues are being solved, e.g. clustering, batch system, I/O, even via wireless communication. Physicists will soon be able to easily analyze data using his/her Notebook PC.

6.7 Future of PC Farms

In '99, with HERA-B, Belle, BaBar, RHIC all starting to take data, there will soon be individual farms with hundreds of processors, and there will be many more PC Farms. E871 at Fermilab is running Production on their data, and CDF and D0 continue to develop experience and prepare for Run II. The Price/Performance of PC Farms will continue to improve at the rate of about 1% per week as technology improves.

The Beowulf projects will also continue to grow in size and applications as these projects develop further and reach for higher goals. The number of such projects will also dramatically increase. Software tools and manuals for clustered computing, both collected and developed by the NASA Beowulf Project, have been packaged as Extreme Linux (www.extremelinux.org, and www.redhat.com/extreme). Extreme Linux will allow a person/group to set up a computer cluster for parallel processing, with the OS and tools that are included in a $29.95 CD-ROM. A person can have PCs for a "Personal Super Computer".

PC Farms and PC Computing will be the main source of CPU. By CHEP Y2K, successes in large scale online and offline applications will be presented. Most high energy physics and nucler physics labs and experimental groups, many universities, and many other fields will all take advantage of the cost-effective commodity computing.

7 Conclusions

PC Farms are a cost-effective solution to future CPU needs. The commodity hardware has advanced to provide high performance at low price, and excellent upgrade path. Linux is an excellent Operating System for the PC hardware platform. PC Farms will be used for Monte Carlo studies, Offline data Production, and Online Trigger event selection. General PC computing will also benefit from, and contribute to the development of the PC Farms. These solutions are now well accepted by the high energy physics and nuclear physics community, and by many other scientific fields. The development of PC Farms and PC computing has been important and holds promise for a great future.

8 Acknowledgement

We thank the Fermilab PC Farms group, the CDF Taiwan group, and the CDF Level 3 group for their pioneering work on the PC farms. We thank Linus Torvalds and the Linux community for an excellent, open source Operating System. Thanks also to the Fermilab Computing Division, DESY, RHIC, CERN, and other labs for strong support for Linux and PC computing.

References

[1] Super Computing '96, Pittsburgh
Computing in High Energy Physics CHEP '97, Berlin
Pentium Pro Cluster Workshop, April '97 Iowa

[2] www.linux.org and www.redhat.com

[3] M. Welsh, "Porting Applications to Linux", 1st International Dutch Linux Symposium, December 8, 1994 (http://www.cl.cam.ac.uk/~mdw24/porting-linux/paper.tex).

[4] S. Goldt *et al.*,"The Linux Programmer's Guide", Linux Documentation Project (ftp://sunsite.unc.edu/pub/Linux/docs/linux-doc-project/programmers-guide /llpg-0.4.tar.gz).

Appendix A Beowulf Projects

Naegling	CACR at Caltech	photo rendering; 114 P-Pro nodes
DragonWulf	Drexel University	atmospheric dynamics
Grendel	Clemson University	
The Hive	Goddard GSFC	plasma, imaging; 64 nodes
Whitney	NASA Ames	fluid dynamics; goal of 5000 PCs
Loki	Los Alamos	gravitational N-body simulation
	LANL	64 Dual P-II
Megalon	Sandia Lab	14 quads
Stone Souper	Oak Ridge (ORNL)	126 nodes at 0 dollars per node
LoBoS	NIH	lots of boxes on shelf; goal 5000
Alice	DOE Ames Lab	
PACET	George Washington U.	
AENEAS	UC Irvine	physics; 65 P-II 300
MAGI	CTU Prague's	speech recognition
Brahma	Duke University	
Topcat	Southern Queensland	
	Texas Tech	
SWARM	Oregon State	Shelves With Affordable ... 32 P-II
Smile	Kasetsart, Thailand	
Brahma	Duke	
Trumpf	Mannheim, Germany	
MadDog	Manchester UMIST	
JazzNet	NIST	
PAPERS	Purdue	video wall, 32 P-II
HPVM	UIUC	32 Dual P-II + 32 Dual P-Pro
	Paderborn Germany	
ZUSE	Koblenz, Germany	
	RWCP, Tsukuba, Japan	
	Academia Sinica, Taiwan	

Table 1: Examples of Beowulf projects and some of their applications

Appendix B PC related talks in CHEP'98

(www.hep.net/chep98/)

011 LINUX PC farm for Reconstruction and Analysis in ZEUS
023 Reconstruction Farm Software for Mock Data Challenge at RHIC
025 Microprocessors: Are they all equal ?
032 Choosing CPUs: System Performance Testing for BaBar Online Farm
043 Batch Computing Facility based on PC (BCFpc)
048 Evaluation of network performance for triggering using a large switch
068 Porting of the BaBar Offline Code to the Linux Operating System
072 NA48 Data Acquisition PC Farm
077 Processing Farms Plans for CDF and D0 for Run II
085 Emulation of Event Builder using ATM LAN of RCNP Institute
090 Porting of CERN-NICE for PC management to university environment
093 PCSF - A PC based Simulation Facility running Windows NT
110 A Real-time Data Network Using ATM
113 Experience with PCs for Reconstruction in ZEUS Experiment at HERA
115 SND Experiment data processing with Linux PCs
138 Fourth Level Trigger Online Reconstruction Farm of HERA-B
144 Linux PC-based Reconstruction Facility at Jefferson Lab
159 Experience with Nile Running Hera-B Code
161 Online farm of KLOE at Frascati: event building and monitoring
162 HEPNT group - recent work, recommendations and future plans
170 Porting and supporting the Fermilab CDF Run II Software under Linux
177 Second and Third Level Trigger Systems for the HERA-B experiment
180 The Linux-PC Farms at DESY Zeuthen
183 Application of PC's and Linux to the CDF Run II Level-3 Trigger
191 Linux Reconstruction Farm at RHIC
197 Apple's Rhapsody OS (Simultaneous UNIX and PC Application)
202 CODA Data Transport using Switched Fast Ethernet
220 Linux Support at Fermilab
240 Fast PCI Interfaces for On-Line Distributed Processing
248 Supported Linux Computing at DESY
519 Use of NT in Quasi-Realtime Environments for D0
520 Software Release Tools on NT for D0
522 RAL NT Farm