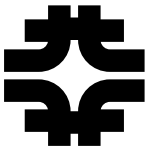

Fermi National Accelerator Laboratory

FERMILAB-Conf-98/365

Linux Support at Fermilab

D.R. Yocum et al.

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

December 1998

Published Proceedings of *Computers in High Energy Physics 1998 (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

LINUX SUPPORT AT FERMILAB

D.R. Yocum, C. Sieh, D. Skow, S. Kovich, D. Holmgren, R. Kennedy

Fermi National Accelerator Laboratory, Batavia, IL 60510

October 15, 1998

ABSTRACT

In January of 1998 Fermilab issued an official statement of support of the Linux operating system. This was the result of a ground swell of interest in the possibilities of a cheap, easily used platform for computation and analysis culminating with the successful demonstration of a small computation farm as reported at CHEP97. This paper will describe the current status of Linux support and deployment at Fermilab.

The collaborative development process for Linux creates some problems with traditional support models. A primary example of this is that there is no definite OS distribution ala a CD distribution from a traditional Unix vendor. Fermilab has had to make a more definite statement about what is meant by Linux for this reason. Linux support at Fermilab is restricted to the Intel processor platform. A central distribution system has been created to mitigate problems with multiple distribution and configuration options. This system is based on the Red Hat distribution with the Fermi Unix Environment (FUE) layered above it.

Deployment of Linux at the lab has been rapidly growing and by CHEP there are expected to be hundreds of machines running Linux. These include computational farms, trigger processing farms, and desktop workstations. The former groups are described in other talks and consist of clusters of many tens of very similar machines devoted to a few tasks. The latter group is more diverse and challenging. The user community has been very supportive and active in defining needs for Linux features and solving various compatibility issues. We will discuss the support arrangements currently in place.

1. WHAT IS SUPPORTED

It has been shown at Fermilab and elsewhere (DESY, CERN, etc.), that farms of commodity PC's running Linux are feasible for such tasks as processing large amounts of data. This has been facilitated by the creation of a standard Linux distribution based on commercial releases. At Fermi, the decision has been made to use the Red Hat Software version 5.0 and modify it to our needs by including the current software updates as well as several Fermi specific modifications such as a boot strap installation of the Fermi Unix Product Support and Unix Product Deployment (UPS/UPD) packages (Heavey 1998). Our distribution is accordingly called Fermi Red Hat Linux v5.0.2 (hereafter, v5.0.2)

Further decisions have been made to exclusively support desktop systems based on the Intel architecture in order to limit the number of support personnel so that the hardware cost savings are not overshadowed by the increase in manpower.

Installation on site is via the local area network from a central Fermi Red Hat distribution server. Currently, individuals are able to obtain an installation floppy from two sites at the Lab, and following step-by-step installation instructions (Yocum, et al. 1998), can install Linux on their computer in approximately 30 minutes. We allow users to administer their own systems but we anticipate moving to a more clustered environment in the near future for those users who do not wish to administer their own machines. A tiered level model of support is currently being researched, with a proposed model being similar to that of the Linux support system at DESY (Wetter 1998).

Off site collaborators can request and will receive v5.0.2 CD's. CD's will not be distributed on-site to maintain the integrity of future Fermi Red Hat Linux releases, and to prevent users from installing earlier, unpatched versions inadvertently.

Software that comes directly from the standard Red Hat v5.0 distribution is treated as an "accessory" and it is not clear that such packages such as Emacs, Netscape, Pine, etc. need to be made into Fermi UPS/UPD products. Products such as tcl, tk, Python, et al. will be available as UPS/UPD products due to the need for specific versions of these products.

We also support the capability to dual boot certain user machines into either v5.0.2 or Windows NT. This decision has been made due to user needs to use certain sets of office software which are not available on Linux.

2. LINUX GROWTH RATE AT FERMILAB

As of August 31, 1998 the number of Linux machines had increased significantly since the official Linux support announcement was made in January, 1998. Table 1 outlines the numbers of Linux boxes at the Lab since the announcement.

	Numbers	CPU's	type	Date
Farms				
	8	single/dual	pilot	1997
	58	dual	production	Oct, 98
Analysis				
	35	dual	production	May, 98
Desktops				
	25	single	production	April, 98
	~150	single	production	Aug., 98
Console				
	6	single	production	Aug., 98

Table 1

3. RED HAT LINUX DISTRIBUTION

As with any operating system release, numerous patches to software packages follow, thus creating a period of time when it becomes difficult to keep up with new releases. We have made the best effort possible to stay on top of this situation. Since the v5.0.2 distribution freeze in late August until the writing of this paper in mid October, only 2 minor software patches have been released by Red Hat.

Since the release of Red Hat Linux v5.0 in late 1997, the Linux kernel has undergone 3 minor version changes from 2.0.32 to 2.0.35. Each version change has added a different aspect of enhancement of security or functionality. We also have made an effort to include the latest kernel version available in our distribution in order to gain the extra features. We treat this as any other patch release by including it in the distribution and testing via installation and simple user tests, i.e., does networking come up, can a user be created, does X work, etc.

Unfortunately, the kernel and modules on the install boot floppy lag behind our distribution kernel and modules. This leads to problems with users acquiring hardware which is supported in the latest kernel version, but which is not supported on the installation boot floppy. For example, large IDE hard drives (>8Gb) are not supported in kernel versions prior to 2.0.35, thus resulting in the entirety of these drive not being partitioned and formatted during installation. The remaining partition must then be partitioned and formatted after installation is complete. We have also encountered similar problems with ethernet and SCSI devices as well.

Another problem we have encountered by applying the patches to the base Red Hat v5.0 distribution is that the 5.0 installation boot floppy disk fails at a critical moment while installing LILO to the master boot record. While the reasons for this are not completely understood, we have been able to circumvent the problem by using the Red Hat v5.1 boot floppy with minor modifications to the install program. As a result we have also gained the 2.0.34 kernel and modules that come standard on the 5.1 installation boot floppy thus

gaining support for several more device drivers at install time. These modified installation boot floppies are availble for users to obtain at two central points on site.

A major factor in choosing Red Hat v5.0 distribution as the basis for the Fermi distribution is that it still uses the gcc-2.7.2 compilers and not the egcs 1.0.x compiler. This decision was made consciously due to the fact that the egcs compiler is under heavy development at this time and does not work completely with code developed at the Lab. It is anticipated that we will migrate to this compiler in order to gain it's enhancements, but until it becomes much more stable, we are only testing it. If users are in desperate need of this compiler it is available via Fermi UPS/UPD product which has been ported to our v5.0.2 distribution.

Where we are holding back with the compiler, we are forging ahead with the latest GNU C libraries, glibc-2.0.7, aka libc6. Again, this was a conscious decision. Red Hat v5.0 comes with these libraries by default, whereas Red Hat v4.2 does not. Unfortunately, this choice was not without some consequences, e.g, many applications built with the previous C libraries (libc5) needed to be re-compiled in order for them to function.

Red Hat does not expect to support SMP installs until the 2.2.x kernel is released some time in the near future, so we are left to recompile the SMP kernel on dual CPU machines. This has been acceptable so far since dual CPU machines have not been making their way to the desktops, and are currently only being used in the compute farms. In order to accomplish these large installations we have made a master hard drive which we have delivered to a vendor. The vendor has then duplicated this hard drive, installed them in the individual nodes and run a script that sets the IP address and node name for each machine. This has worked remarkably well, and the machines are ready to go after being installed at the Lab. The 37 node farm (see Table 1) is truly a plug-and-play supercomputer, and this will be replicated many times in the future as the Lab prepares for Run II (Wolbers 1998). We are considering creating our own SMP installation to facilitate the future arrival of dual CPU desktop machines.

Of particular interest is our success using Red Hat's Kick Start Installation. When used in the mode, one can insert a pre-configured installation floppy into a machine and literally walk away. Certain assumptions are made regarding the method of installation, specifically, this type of installation only can be done via NFS or CDROM, with NFS being the preferred method on site. NFS installation takes approximately 30 minutes and this time is roughly cut in half for the CDROM install.

4. LINUX AT FERMILAB

The generally accepted philosophy of supporting Linux on the desktop is that the machines should be completely disposable. To accomplish this end we are encouraging that no user has a local home area and that data that is on a machine can be restored from some external source. This requires that the users home area is on a master NFS server which is backed up regularly. The user can develop code locally, but the CVS code repository will also be on a CVS server. This poses some problems, particularly when writing large files to the NFS server. The latest supported version of NFS for Linux is version 2, which has relatively good read speeds (~700-800 KBps) but appallingly low

write speeds; as low as 25 KBps on an untuned server-client system. This number only approaches 100 KBps when the server is tuned on a Fast Ethernet. The hope is that these numbers will improve with the release of the 2.2.x kernel.

Currently we are exploring several avenues to facilitate upgrading software packages for security and feature enhancements. Several other high energy labs (DESY, CERN) are using cfengine and SUE (Cons 1998). These are being considered as well as autorpm (Bauer 1998), rsync, and the possibility of developing a difference engine locally. Currently, security and software enhancement announcements are made via email to those who have installed v5.0.2 themselves.

As stated earlier, the on-site installation is via NFS mounting a local v5.0.2 distribution server. This is an essential element to the success of deploying v5.0.2 to desktop machines. On site users who wish to install Linux on their system need only to obtain a modified installation boot floppy from one of the two distribution points and follow the online v5.0.2 installation guide. Off site collaborators can receive a CDROM upon request, or if they are local to the Fermilab area, they can order a pre-installed v5.0.2 system from a number of local hardware vendors with whom we have fostered contacts and have given a v5.0.2 CDROM.

Several vendors have agreed to adhere to the stringent hardware specifications that we require and we have had some success with purchases made from them. On farm nodes, the vendors have also performed the burn-in of the machines in their shops prior to their shipment to the Lab. To ensure that we receive systems that are usable we specify hardware one notch back on the technology curve. For example, at the time of the 1998 CHEP conference in late August, Linux was just beginning to support the Intel BX motherboard chipset. This is approximately 3 months after the release of the chip, but what we lose in cutting edge technology, we have gained in price cuts, therefore making it feasible to purchase more systems than we previously could on a limited budget.

5. SECURITY AND VULNERABILITIES

One of the by-products of treating these systems as completely disposable, is that if a system is compromised by an unauthorized individual, we can simply disconnect the machine from the network, determine the method of entry, and re-install from pristine sources, i.e., CDROM or the installation server.

This is an acceptable means of restoring a machine to the network, however, we have made many security enhancements from the base Red Hat v5.0 distribution to help reduce the need to do so. Besides including all the patches provided by Red Hat, we disable several services by default, including IMAP and POP2 and 3 mail servers, named domain name server, and Apache web server. In order for a system to be upgraded to the most recent patches, security or enhancement, the user only has to use the initial boot floppy and choose the ’Upgrade’ option when prompted. This will upgrade only the packages that have been patched, leaving the rest of the system intact. We also supply, to on-site users only, the Secure SHell (ssh) (SSH 1998) for installation via RPM (Bailey 1997).

6. SUPPORT AND APPLICATIONS

As of the time of this conference the majority of the installed desktop systems are user installed and administered. Request for support are beginning to come in, but they are via the same channels that any other Unix service request come through, and they have not been overwhelming. We expect that with the deployment of large clusters of machines that the support load will be similar to any other operating system on site. Many user questions can be answered by referring to the Linux at Fermilab web site at http://www.fnal.gov/cd/unix/linux.

In June, 1998, the Computing Division User Application Support group began porting Fermi UPS/UPD products to v5.0.2. These include, but are not limited to: upd, tcl, tk, expect, blt, ical, ups, gmake, gtar, bison, flex, ftt, netscape, lynx, perl, bash, tcsh, fmb, tkman, python, IRAF, egcs, and KAI.

We are committed to the ideal of upholding the open source environment of Linux. If purchases of a software license are necessary, we try to purchase site wide licenses so that we don't have to monitor the usage of the software.

For problems that extend beyond our range of knowledge, we have negotiated a 3Third Level support agreement with Red Hat Software. Support requests are sent via email to a closed mailing list of five Fermi computing division people and three Red Hat people.

We are also offering Linux System Administration Class to on site Unix System Administration personnel. These are one day classes focussing on the basics of PC hardware and Linux specific topics. Response to these classes has been very good and at the time of writing approximately 33 people have attended our first set of 3 one day classes at the beginning of October (Sieh, et al. 1998).

7. SUMMARY

Overall, the response to our support of Linux on the desktops has been favorable which can be seen by the increase in numbers of machines. Support requests have not been overwhelming yet, but, this is expected to increase in the future as more clusters of machines are installed with fewer people having root access to the machines on their desktops. However, we do not anticipate the level of support to be different than the other Unix operating systems supported at the Lab.

We are providing for easy installation and upgrades to people who administer their own machines as well as developing methods to carry out these tasks on farm and cluster nodes.

REFERENCES

Bailey, E.C. 1997, Maximum RPM, 3

Bauer, K. 1998, Linux Software by Kirk Bauer, www.kaybee.org/~kirk/html/linux.html

Cons, L. 1998, SUE: Standard UNIX Environment at CERN, wwwcn.cern.ch/dci/sue/

Heavey, A. 1998, UPS and UPD v4 Reference Manual, 1-1

Sieh, C., & Yocum, D.R. 1998, Linux Administration Class, www-oss.fnal.gov/fss/documentation/linux/training.html

SSH 1998, SSH Users Group, www.ssh.org

Wetter, D. 1998, International Conference on Computing in High Energy Physics

Wolbers, S. 1998, International Conference on Computing in High Energy Physics

Yocum, D.R., Kovich, S., & Sieh C., Installation guide for Fermi Red Hat V5.0.2, www-oss.fnal.gov/fss/documentation/linux/install-guide.html