

Computing Facilities in the PDSL

(last updated: July 11th, 2007)

Important: Please be sure to report any equipment failures (no matter how small) to pgraham@cs.umanitoba.ca immediately. Broken equipment takes time to repair/replace and small problems (e.g. a noisy fan) often turn into big problems (e.g. a dead machine) if not reported and dealt with quickly.

Workstations

There are six (6) shared, general purpose workstations available for use in the PDSL. Each is a Dell Optiplex GX620 with a hyperthreaded Pentium 4 processor running at 3.0GHz, 2GB of RAM and 160GB SATA hard disk. Displays are 17" LCD panels at 1280x1024. All workstations are dual boot Linux FC4 and Windows XP Pro. Three of the machines boot to Linux by default while the other three boot to Windows. Feel free to reboot to another OS if necessary but please leave the machine in its default state when you are finished. The Linux side of each machine is intended for software development work and access to the main computing facilities in the PDSL. Accordingly, the Linux side of each machine gets its userids, home file system, etc. from the internal lab server (which is also the head node for the compute cluster – see below). For this reason, departmental Linux accounts, ACN accounts, etc. that you may have will not work in the PDSL. You must request a Linux account from pgraham@cs.umanitoba.ca. The windows side of each machine is intended to support development of presentations, etc. and is managed exactly as any other Windows machine in the department. Hence, your `CompSci` domain account will work in the lab. All machines are also dual-homed (i.e. they exist on two different networks). On the general Computer Science network (130.179.{28,29,30,31}.x) the machines are named `pigeon-01.cs.umanitoba.ca` through `pigeon-06.cs.umanitoba.ca`. On the PDSL-local network (192.168.1.x) they are known as `pdsl-01` through `pdsl-06`. On the Linux side, the disks are divided as follows:

<code>/</code>	The normal Linux/Unix filesystem. Do not store files here.
<code>/home/UID</code>	Your PDSL local home directory, where <code>UID</code> is your login id. Files stored here are available on other Linux machines in the PDSL (including cluster nodes – see below) but not on machines outside the lab. Automated backups are not provided on <code>/home</code> (see below).
<code>/share</code>	This is machine-local space shared with the Windows-side of the machine. (The <code>G:</code> disk on Windows.) It is formatted as FAT-32 so there is no protection. Also this filesystem is not backed up automatically (see below). Each user should create their own directory and put their files in that directory only. Please respect the privacy of the data of others as well.
<code>/psdl/backups</code>	A place where you can put critical data to get automated backups (see below). Your files should be placed in <code>/psdl/backups/users/UID</code> , where <code>UID</code> is your login id.

`/pdsl/stage` A directory where commonly used cluster software is maintained. This is available on cluster nodes and, for convenience, on workstations in the lab.

On the Windows side, the disks are divided as follows:

- C: The Windows system partition – do not put your files here.
- D: The DVD/CD combo drive.
- E: The applications partition – do not put your files here.
- G: A large partition shared with the Linux-side. Files may be stored here but they are local to this machine only and are not backed up automatically (see below). Further, since the file system is FAT-32, there is no protection. Each user should create their own directory and put their files in that directory only. Please respect the privacy of the data of others as well.
- S: Your `CompSci` domain mounted home directory (shared throughout the department). This is where you should probably be putting your Windows files.

Finally, since these are shared machines, please always abide by the machine locking policy posted in the lab. Repeated violations will result in loss of access privileges!

The CAB Cluster

The primary compute resource in the lab is the “Crash And Burn (CAB)” cluster. This is an 8 node research cluster (Beowulf style) with a variety of interconnects. Each node in the cluster is based on dual core, dual processor Opterons with 2GB of RAM and 250GB SATA hard drive. The cluster nodes are interconnected using Fast Ethernet, Gig Ethernet, Myrinet and CLAN switches as well as SCI point-to-point links. The name comes from the fact that the cluster is intended to support the development of low-level systems software and hence may be inherently unstable. If you have a large scale parallel compute application, there are alternative (stable) facilities available external to the PDSL (see below). The nodes in the cluster are named `b1` through `b8`. The CAB is hosted by a head node (see below) which provides userids and shared file system for both the cluster and the Linux side of the workstations in the lab. Only parallel compute jobs should be run on cluster nodes. Various versions of MPI targeting the various interconnects available are provided for this purpose. If you plan to run something else on the cluster or need dedicated access (e.g. for timing runs), please contact your advisor well in advance. In addition to the Linux file systems described earlier and available on the Linux side of each lab workstation, each cluster node also has a local file system where data may be stored. This may be important for performance reasons when doing parallel computing as NFS access can quickly become a bottleneck. On each cluster node (`b0` through `b8`) there is a file system `/data` where machine local data may be stored. You should store such data in `/data/users/UID`, where `UID` is your login id.

The Head Node/File Server

The cluster head node (`dovecot/pdsl-srv`) is a Dell PowerEdge 830 with dual core Pentium D processor running at 3.0GHz, 4GB of RAM and a 250GB RAID-1 array (i.e. mirroring) based on dual 250GB SATA hard drives. It provides DHCP, DNS, NIS and NFS

services to the cluster nodes and to the lab workstations via Gig Ethernet. When you store files in your Linux home directory, they are maintained on this machine and the use of RAID mirroring provides some measure of data access availability. Remote access to the cluster nodes can be achieved by logging into this machine though using one of the lab workstations (pigeon-01 to pigeon-06) for this purpose is preferable as it will not disrupt services provided by the head node.

The Simulation Server

Another compute resource available to users of the lab is a machine dedicated to performing simulations. Simulations are normally relatively compute bound and often require large amounts of memory to run efficiently. The machine `palladium.cs.umanitoba.ca` is a dual core, dual processor SunFire X4200 with 10GB of RAM and 140GB of high speed SAS hard drive storage. While this machine is generally available for use, there may be situations where dedicated bookings are required to meet publication deadlines, etc. Please discuss any long running simulations with your advisor well before initiating them.

Backups

There are no automated backup procedures for Linux-based machines in the PDSL. Further, there are no backups available for local disk storage on the Windows machines (though your S : drive is backed up regularly). Users of the lab are expected to do their own backups. This can be done in several ways. First, users of the lab are usually given a small (e.g. 128MB) USB memory stick by their advisor(s) when they join the lab. This is intended for day to day backups of your most recent work – use it! Second, under Linux, you can place *copies* of larger files in your `/pdsl/backups/users/UID` subdirectory for twice weekly automated backups. (Note that you should not put unnecessary material here as there is a limit as to how much can be backed up over the network overnight.) Finally, for archival backup purposes, there is a DVD writer in the lab on the machine `teak/t1` which is located near the machine room door. When you have finished work on a particular course/project, you can use one of the DVD-RW disks provided to permanently archive the related files. Such files should be subsequently removed from the system disks they were previously stored on. Given these resources, a suitable example backup regimen might be as follows:

1. Maintain two primary working directories (for both Linux and Windows, if you use both systems): CURRENT and ONGOING. At the end of each day, backup the contents of CURRENT to your memory stick. (Note that you will have to keep the size of CURRENT small enough to be stored on the memory stick you have.) To do this, move older material from CURRENT TO ONGOING as needed. Note that large volume, personal material such as MP3s, etc. should not be backed up. This is your personal responsibility and should be done with non-PDSL and non-university resources!
2. Twice a week, copy the contents of CURRENT and ONGOING to your `/pdsl/backups/users/UID` subdirectory.
3. At the end of each term, go through your ONGOING directory and decide what you want to keep, what you can discard and what you can archive to DVD and do so.

Please note that no special consideration will be given to those students who have lost work who did not take responsibility for performing regular backups of their work.

Printing

There is no printer in the lab. All machines in the lab (except for the cluster nodes), however, do have access to the printing resources provided by the department. Thus, you may print to departmental printers from both Windows and Linux machines in the PDSL. Further, there is a duplex laser printer two doors down from the lab (*Amber*) so access is convenient. Be aware that pages printed from the PDSL do count towards your monthly page limit. Be circumspect in your printing. Please read what you can online!

Casual/Visitor Network Access

There are facilities for visitors, casual users and/or students with their own laptops to gain network access within the lab. Wireless access is available using the `CompSci-PDSL` SSID. This is a WEP enabled network and the SSID is not broadcast. Hence, you will need to explicitly add the network to your list of wireless networks and provide a WEP key before getting access. The current WEP key is written on the upper left hand corner of the whiteboard in the lab. Wired access is available (via NAT) using the switch on the tables in the center of the lab. Just plug in and you will get a DHCP address. No printing or file access services are available via these networks unless you use them to log in to an established lab machine or to access a departmental network.

External Resources

A large number of additional resources are also available to PDSL users. These resources are provided by the Department of Computer Science generally. Most notable among these are the 80 core compute cluster run by the department for compute intensive applications, the primary login servers (offering e-mail access, etc.), and the print and shared storage facilities. Please ask your advisor for information about these resources if you believe you need access to them.