Sequoia Standard Ultrix

Jon Forrest

Project Sequoia 2000

1. GOALS

Project Sequoia will make use of a large number of computers, so many that if they aren't managed carefully and consistently the goals of the project could be threatened. With this in mind, this document describes the computing environment in which Sequoia researchers at UC Berkeley will operate. Most of the features of this environment are based on comments from members of the Postgres research group, which has been using a similar environment for quite a while *.

For the sake of brevity, the new environment will be called SSU ("Sequoia Standard Ultrix") but what is really being described is the hardware and software environment for Project Sequoia.

Although other Project Sequoia sites are not required to run SSU, they are encouraged to replicate this environment as closely as local conventions allow.

2. HARDWARE ENVIRONMENT

SSU obviously assumes that the machines in its computing environment will be DEC machines running Ultrix. This is not to slight other manufacturers but, given DEC's generous financial contribution, other options aren't financially viable. On the other hand, existing non-DEC hardware can interact with SSU machines but the non-DEC hardware should not assume that all the features of SSU will be present.

Project Sequoia will have 2 fast server machines, called **HEEL** and **TOE**. These servers will be DECSystem 5900's and they will each have 128Mb of RAM, 10 1.3Gb RZ58 disk drives, Prestoserve, 1 8mm backup device, and a CD-ROM reader. Other backup media, such as 4mm and 9-track tape, will be added as needed. These servers will be configured with very large disk caches. Interactive use of these servers will be discouraged.

Each workstation will have a minimum of 32Mb of RAM and a local disk of at least 600Mb. Since most Sequoia researchers will use visualization tools, most workstations will include at least 8-bit color graphics using a 16" or 19" 72Hz color monitor. Researchers needing more graphics capabilities can order the various accelerators.

All systems at Berkeley will be connected by FDDI. There will be adequate control of the routing tables on all workstations to insure that routing loops do not occur. Routing to ethernets will be accomplished by a designated router on the FDDI net. More information on Sequoia network configuration will be sent out to system managers as the information becomes available.

Whenever possible, exotic hardware (e.g. hardware that requires special drivers) will not be placed on the servers, at least not initially. Driver development and device shakeout will be done on a workstation or some other non-critical machine.

There will be at least one 8-mm tape drive available on machines outside the machine room that can be used by people who don't have machine room access. Machine room access will be minimized.

---

* Machines that are currently part of Postgres will also be migrated into the environment at a time convenient to both Sequoia management and to their owner.

There will be at least 6 dialup modems connected to an Ethernet-based terminal server. As many of the dialup lines as possible will be connected to 9600-baud V.32 modems but, under no circumstances, will any speed less than 2400-baud be supported. The telephone lines used by these modems will be configured in a hunt group so that one generic incoming modem telephone number can be used.

3. OPERATING SYSTEM

All Sequoia machines will run the latest general release of Ultrix, currently Ultrix 4.2A. The only time exceptions will be made is when a special OS release is required to run special hardware, or when OSF/1 comes out. In the later case, we may wait a while to hear the word on the street since we don't want to suffer from potentially unstable software. We will also make use of patches from DEC's Software Support Center so that software problems can be fixed without having to do a complete OS upgrade.

4. USER ENVIRONMENT

The general model for SSU is that any user can log in to any machine and work in their environment of choice. Home directories will be available on all workstations. All printers can be accessed by all workstations. All users can send and receive mail on any workstation.

5. LOCAL AND REMOTE FILE SYSTEMS

Each workstation will have local root, /tmp, and /var file systems. Swapping will also be done to a local swap partition whose size will be three times the amount of RAM. For the sake of maintainability, users will not be permitted to modify the root file system. Workstations will be configured so that at least 20% of their memory is used for disk buffers.

Again, for the sake of maintainability, the /usr file system will be located on one of the server machines and served read-only to all the workstations. It is recognized that this may not be a popular decision but the amount of effort that could go into maintaining and upgrading multiple /usr file systems would be impractical. To help satisfy reliability concerns, the /usr file system will also be mirrored on another disk on the other server.

Each workstation will also have a local file system on which the owner of the workstation can keep any files they want. None of these files will be maintained by Sequoia management in any way. Users who copy files to a local disk do so entirely at their own risk. Workstations will not be configured as NFS servers so their local disks will not be served to other machines. Files that need to be shared should be located on one of the servers.

User login directories will be located on one of the servers and will be served using NFS. This makes it easier to backup and to share user files without requiring workstations to be NFS servers. In order to prevent down time when one of the servers goes down, a disk on the other server will act as a mirror for the disk containing user directories. An automounter will be used on workstations so that the switchover itself will be invisable. Although every effort will be made to carry out this goal, it may turn out that the mirroring is implemented by a nightly disk-to-disk copy so that the result of a failover won't be completely transparent. Other approaches using disk shadowing or RAID techniques might be used but they don't solve the problem of what to do when a server itself goes down.

Previous concerns about Ultrix NFS server performance should be at least partially overcome by using FDDI as the primary NFS medium.

To prevent the dreaded "nfs server not responding, still trying" message, read-only and read-mostly file systems will be mounted using soft mounts, although this policy might be changed based on experience with NFS on Ultrix.

6. BACKUP POLICY

It isn't possible to describe a definite backup policy for every disk that will be used on a Sequoia system due to variations in origin and usage of the data on these disks. For example, disks that contain frequently modified files obviously must be backed up more often than static satellite data. Some disks won't need to be backed up at all because they'll be used as a staging area for data that comes from CDROMS.

Every workstation's local disk will be backed up incrementally every night to a directory on one of the servers. These dump files will then themselves be dumped to tape everynight as part of the nightly incremental backups that will be done on server disks. Sequoia system management will be responsible for arranging for system disks (e.g. disks containing /usr, and login directories) and private file systems to be backed up, as well as for performing the actual backups themselves. Users can optionally choose to be totally responsible for backing up their local disk.

Sequoia management will be responsible for keeping a minimal set of backups of crucial files off site so that the Sequoia environment could be recreated in case of disaster.

7. OPERATING SYSTEM CONFIGURATION

The operating system on each workstation will be configured to access FDDI, Ethernet, 2 local disks, an 8mm and 4mm tape drive, a TK50 tape drive, a 3.5" floppy drive, a CD-ROM, and a QIC tape. Few, if any, workstation will actually have all these devices connected but having such a configuration reduces the need for later operating system modification. Also, each workstation will be configured with enough semaphores and shared memory to run Postgres.

8. ACCESSING POPULAR FREE SOFTWARE

Most users have become accustomed to using various "free" software packages and programs such as X-Windows, GNU software, TeX, and others.

There are currently several repositories for "free" software, e.g. the Software Warehouse, and the server maintained by the Workstation Group. The idea of using one or both of these resources is attractive but given that Sequoia can't control these resources, they will not be used as the primary location for free software that Sequoia depends on. Instead, space on disks connected to the Sequoia servers will be used for this purpose. On the other hand, free software not crucial to Sequoia will be accessed from the Software Warehouse or other locations.

It is inevitable that commonly used software will be copied to local disks. After SSU is stable a system may be set up in which a user can specify where they want the free software to reside and then automatic tools will handle the tasks of copying the software to local disks and making sure the software is up-to-date.

9. DISTRIBUTED SERVICES

Each workstation will be configured as a BIND client. At least one of the Sequoia servers will act as a secondary nameserver for the CS.Berkeley.EDU domain. All other services will be provided by NIS. Local password files and host files will not be provided by Sequoia management unless configuration problems require them.

10. PRINTING

Every machine in the Sequoia environment will be able to print to any printer in the environment. This will be accomplished by creating a symbolic link from /etc/printcap to /usr/etc/printcap, which will be the master printcap file. Any new printers purchased with Sequoia money will be Postscript printers.

11. EXTERNAL MAIL ADDRESSES

Every Sequoia user will have an e-mail address of user@CS.Berkeley.EDU. This will be their official e-mail address. Existing users that have been using other addresses can continue to use them but all new correspondance should go to the new address.

Every Sequoia machine will have a name of the form machine.CS.Berkeley.EDU. Again, existing names will not be disabled but all new machines will receive the new style names and all official correspondence will use the new names. At some point, exiting names will be "CNAME"ed to the new names using the CNAME facility in BIND. Outgoing mail will have the name of the Sequoia mail server appear instead of the name of the originating workstation. This will be accomplished by using an official SSU sendmail.cf file.

12. WINDOWING SYSTEMS

Each workstation user can select the windowing environment of their choice. It is recommended that the DEC-supplied X server be used, especially on workstations with exotic graphics hardware, but each user may choose their own windowing environment. The default windowing environment will use the xdm display manager but users will be free to choose their window manager.

13. SUPER-USER PASSWORD ACCESS

In Project Sequoia there must be a balance between the need to get work done and the maintainability of the environment. What this means is that we have to be very careful about who gets the superuser password but not so careful that users can't do what needs to be done. As a first attempt to solve this problem, there will be a separate superuser-like account for each workstation which the workstation owner can have the password for, if they want. This account will have the name "owner" and its password will be different on each workstation. This account should only be used in cases of emergency when the Sequoia system manager can't be found. It is not intended for everyday use. The Sequoia system manager will keep an eye on the use of this account to make sure things don't get out of hand. The password for the "owner" account will have periodic mandatory pasword changes. There will also be the normal "root" password which will only be known to a few people.

In rare cases, workstation owners will be allowed to setup accounts that are valid only on their workstation. These accounts might be for people who need temporary access to a workstation or who are temporary collaborators. In any case, these accounts are provisional and are subject to the scruntity of Sequoia management.