

# Linux and MAC OS X Deployment at STScI

## Scientist User Feedback

*-compiled by C. A. Christian*

Introduction .....	1
Interview Results.....	1
User environment .....	2
Usage.....	2
Linux.....	2
MAC OSX.....	3
Configurations .....	3
Centralized Storage.....	4
Local Root/Administrator on machines?.....	4
Dual Boot?.....	5
Peripherals.....	5
Citrix .....	5
Other Comments .....	6
Recommendations .....	6
Short Term Prioritized Solutions.....	6
Longer Term Priorities.....	7
Hardware Purchases.....	7

## Introduction

Scientific staff at Space Telescope Science Institute (STScI), similar to their colleagues in the science community, are enthusiastic to use technologies, software, and systems that are effective tools for conducting their research. In addition, STScI scientists need facilities to support their functional activities. Like any user community, the demographics of the STScI science staff are varied, and their individual environments necessitate different deployment and support approaches.

In order to provide cogent information regarding the deployment of both Linux and MAC OS X systems, scientists were interviewed (by C.A.C) and their commentaries analyzed by a Task Group drawn from the STScI scientific staff. This document is a report of those activities, and the resulting recommendations. The purpose of the report is to inform the Center for Process and Technology (CPT) regarding user requirements, recommendations, specifications and a strategy for collaboration between CPT and the scientific staff to accomplish a rational plan.

## Interview Results

In the interviews, scientists were asked to describe their work environment: at STScI, at home, and while traveling. The interviews were generally free form, although CPT had identified some key topics for probing. Note that in most cases "Unix" refers to Sun Solaris Unix in the following discussions.

## ***User environment***

Scientists were asked to describe the various operating systems used. Most responses focused on the desktop environment.

- Desktop environments, if replaced by Linux (or MAC OS X), should be equivalent to what scientists have now. This means in function, in available software, in upgrades, patches and support.
- Some scientists are happy with their Sun/Unix workstations and there may be other Unix users who would object to mandatory replacement of their Unix box by a Linux machine. As scientists purchase new machines, they wish to choose the type of system that fits their needs (or their postdoc and student needs). Cost, performance, support, experience and other factors influence those preferences.
- Users experience that Red Hat Linux is full Unix and comes as such. Today, they have few problems with that flavor of Linux and find it easy to upgrade.
- Scientists have built up a significant amount of expertise with the Linux and MAC OS X systems they use.
- Some scientists are using the MAC OS X (usually a laptop) environment and prefer it to their desktop Sun UNIX workstations because it provides a seamless environment between the Office suite and an X environment.

## ***Usage***

Scientists were queried about their usage of their various systems. What software and utilities do they use?

### **Linux**

- Scientist users who have Linux systems, use them for both functional and research tasks.
- Science applications that are used regularly are IRAF, IDL, AIPS, AIPS++ and various other utilities.
- Many scientists use Linux on a dual-booted laptop. They would prefer a seamless environment between Linux and Windows applications, and so for now, many use dual-booting as a solution. This admittedly is a work-around. They wish to use the Office suite and other Windows-based utilities such as math routines, display routines and etc. for functional and research use. Certainly scientists need Word and Power Point on a mobile platform. Many also use LaTeX for word processing scientific material.
- Some scientist Linux users have been able to create identical environments on their laptops and their Unix desktops. This is great! Users desire to have this type of seamless environment. The level of cloning of the desktop and laptop (and home) environments depends upon the experience, need and available time a user has. Unfortunately there is no elegant synchronization tool to maintain the two environments this way. But it is quite useful to have the laptop duplicate the desktop environment so it can be mobile with the user.
- Some scientists found dealing with the synchronization problem too time consuming. They work around the situation by performing some tasks on their laptops that cannot be easily performed on their desktop. For example, while scientists prefer to do data reduction on desktops and laptops, some scientists may view attachments, prepare documents and etc. solely on their laptop to avoid using Citrix, Office emulators or other software on their desktop.
- Scientists very often use Windows applications for presentations (as noted above) and also opening attachments sent through email. All Linux users require some Windows applications

- for presentations. StarOffice, the only available native application that attempts to duplicate Office, is not really equivalent, but useful for reading attachments
- Scientists use other applications they need installed in the Windows environment. (e.g., Palm or Handheld PC SW), Meeting Maker, etc. One user reported using Jpilot as a calendar that I can easily sync with the Palm. The Meeting Maker page claim support for Linux and Java, but the experience in the Institute is limited in this regard.
  - Scientists require autonomous systems with dual boot when traveling. Typically scientists are or will be conducting both research, analysis, reading email and accessing Office and other Windows utilities while on the road. Sometimes this involves extended stays at another institution.
  - Most scientists add their own applications to Linux and to Windows. They strongly feel that the original disks are required to adequately maintain their own environment, especially their mobile environment.
  - Some individuals are using applications such as Xmacs, image processing and Video creation packages

## MAC OSX

- Scientist users utilize their MAC OS X systems for both functional and research tasks.
- Scientists are interested in running IRAF and IDL on their systems. Note some IRAF utilities on MAC OS X are not really mature and lag both Unix and Linux in implementation.
- As for PC Windows / Linux users, scientists are interested in having other utilities such as the Office suite, mathematical, computational, and management (calendar, project management, etc.) available on their machines.
- Manager scientists would like to have a project management program compatible with MS Project (since that appears to be the standard used in-house).
- Some scientists (and certainly technical staff) need FileMaker and Framemaker.
- Most scientists add their own applications to the MAC. They strongly feel that the original disks are required to adequately maintain their own environment, especially their mobile environment.
- Scientists would like to run the Starview interface, but may have to live with the web interface to the archive instead.
- MAC users would like to use the wireless environment now.

## Configurations

Users were asked about the configurations they used. How did their system get configured? Did they customize the configurations?

- All Unix applications currently in use are expected on Linux configurations to enable scientific research.
- Users will expect to have access to data regardless of its origin – Unix, Linux or even PC or MAC. Users expect data to be shared across multiple platforms. Further investigation by CPT in collaboration with the science staff would be beneficial to explore solutions and cost/benefit.
- Over time, a large number of macros, scripts and other utilities have been built for the SUN Unix systems to aid in scientific research and other tasks performed at the Institute. Scientists will expect those utilities to be available on their Linux systems. Detailed specifics can be explored by CPT in collaboration with the scientific staff.

- Some users heavily invested in their Unix environment as they have created it for its own use. Scientists are reluctant to relinquish such a familiar customized environment unless they receive support to re-create them on the replacement Linux platforms. Again, CPT can work in detail with the science staff to specify the situation if replacement of a Unix system by a Linux system becomes an imperative.
- Users desire Linux systems to have automatic patch updates, etc. so that they (the users) do not have to do so much system management on their Linux systems.
- Some users already using Beowulf cluster and would like to see such facilities expanded.
- Some commercial compilers do not match from brand to brand of Unix or Linux so some commonality and standardization would be needed. For example, most scientist seem to be happy with RedHat Linux. This situation mostly is a problem when running large application packages that may be from a source other than the user. Users, through a combination of a Linux user's group and action of CPT, need to be alerted about certain applications that may not migrate well between systems, before they purchase such systems.

### **Centralized Storage**

Users were asked about the usefulness of centralized storage and their experience with using centralized or remote storage (i.e., storage not located on their desktop system).

- Some scientists make use of centralized storage, but it is painful because the network is so slow. Typically users try to copy what data is needed to a local area for concentrated work.
- Most users have localized storage and find this beneficial.
- Scientists who work in groups need to share data. Some scientists use centralized storage, others put the data on their own machines and require their postdocs and students to access those files. This is becoming a problem because both large and small collaborative groups are being inhibited by the network speed. A postdoc or collaborator sees a slow speed if the data is in central storage *or* on their colleagues system. It makes little difference, the network reduces productivity. Users cannot afford to purchase large amounts of storage space for each machine in their workgroup and also synchronize all the data and ancillary information on every machine to get around the network speed problem.

### **Local Root/Administrator on machines?**

Scientists were asked about their experience with troubleshooting their own system, installing software and getting upgrades. Scientists were asked specifically if they needed local administrator access to their desktop and laptop machines.

- Scientists require administrator privileges on their mobile machines. This situation is not restricted to Linux. The situation applies to Windows, and the MAC. Scientists need control of the machines to add software and to troubleshoot.
- Most scientists who have Linux or MAC OS X have been administrating their own systems for a long time. Several scientists used Linux before coming to STScI and have experience in that regard.
- Most scientists do not insist on local root or administrator on desktop machines, *if and only if* help from CPT is swift. Scientific staff appreciate that CPT is limited in resources and feel confident in handling some of their own administration. Many scientists are quite technically competent in this area. Many stories regarding trivial problem fixes were told in which the scientist was "admonished" or "forbidden" from fixing the system or contacting a vendor directly. Many scientists related that they often could find fixes, upgrades, drivers and other

- items by browsing the web and conferring with colleagues so that they knew in detail what action needed to be taken.
- Some scientists who are power users wish to have administrator privilege on the desktop. This should be examined on a case-by-case basis.

## **Dual Boot?**

The CPT made it clear they prefer not to support dual boot systems or more than one system on a given desktop. A similar sentiment was expressed regarding laptops. Scientists were asked how much they depend upon dual boot systems. The Task Group felt that dual boot is an interim *albeit* poor substitute for seamless integration of the tools they ultimately need.

- Scientists require dual boot systems on their PC based mobile environments. A seamless access to Linux and Windows applications is desired and the current interim workaround is to install Red Hat Linux and Windows. A minority of users wish to have Linux (or Unix) and MAC OS X available for some applications to have access to applications released on Linux (Unix) but not the MAC OS X system. Perhaps a solution is to provide robust access to central Sun Unix or Linux services to avoid maintaining two systems in one office.
- Users noted that many science applications are released on Linux and Unix nearly simultaneously, but on MAC OS X more slowly. Therefore MAC users may want access to Unix systems as well. This sentiment means that MAC users are inclined to have two systems in their offices. Perhaps a solution is to provide robust access to central Sun Unix to avoid maintaining two systems in one office.
- Most users did not require dual boot systems on the desktop, however some did because the Windows applications they used were more extensive than Office and did not run well on Citrix.
- Several users suggested investigating VMWare or Cross Over instead of dual boot, or some other clever system for access to Windows applications in a seamless fashion.

## **Peripherals**

Scientists were asked if they had any unusual needs for peripherals or had any experience in this regard.

- Scientists in general expect the usual peripherals to be available – hard disks, printers, tape drives (Exabyte, DLT, DAT, DVD CDROM burners). Users remarked that a few years ago peripherals and drivers were problematic but that is not the case today with Red Hat and certified machines.
- A small number of specific problems such as “the latest” type of storage device or driver were related, but in general, users felt that commercial support of peripherals was available rapidly for new devices.

## **Citrix**

The CPT was specifically interested in user experience and needs regarding Citrix.

- Scientist users obviously cannot ignore Windows applications. They need the Office suite and typically other utilities and programs installed on Windows. They need a seamless method for accessing Windows based software from Linux. For MAC OS X this is not such an issue.
- Desktop Citrix satisfies most scientists who only occasionally use Windows applications. In those cases the desire for dual boot systems or multiple systems on the desktop is small.

Other users who regularly use Windows applications and power users who use Windows software more extensively have had more problems with the Citrix system. Many of those users have resorted to confining their extended Windows application use to their laptops for this reason.

- However laptop users find Citrix is too big and does not satisfy their needs. Also Citrix is unusable when traveling. Users find that a dual boot system performs better especially since some stand-alone PC utilities are needed.

### **Other Comments**

- User recommend that the Center create a transition plan that enables scientists and other staff to switch from Sun Unix desktop systems to Linux and/or MAC OS X as needed.
  - Can this be quick if it is relatively seamless?
- Propose making available well equipped central Suns with disk space that will enable scientists to give up their SUNs without forcing them to convert every last bit of code to be Linux or Mac O/S compatible.
- Scientist users do not agree that the schedule for deploying MAC OS X and Linux on laptops should be so long.

### **Recommendations**

This section contains the recommendations and solutions from the scientific staff at large, the interviewees and the Task Group. The scientific staff at large have discussed these issues and wish for CPT to work directly with a subgroup appointed to represent them in developing the detailed plan for Linux and MAC OS X deployment. This subgroup of the scientific staff has been designated and will also work with CPT on networking issues, firewall issues and the IT strategic planning.

The scientific staff, being diverse in needs, work habits, and technical expertise, will not be well served by only one set of operating procedures. Specifically, some users who are more technically capable can and are willing to take on more responsibility for their systems. This can be done with an agreed upon mutual understanding of risks and consequences. Other users prefer to have more standard systems with standardized support. This dichotomy in the user base is reflected in these proposed solutions.

### **Short Term Prioritized Solutions**

- STSCI should create an environment that is Linux-friendly. What does that mean?
  - Acknowledge that Linux and MAC OS X systems are already in use.
  - Define standard systems for individuals who wish to purchase now.
  - Allow systems to be purchased with Linux installed or MAC OS X installed.
  - Purchase a site-wide license for using Red Hat's update facility.
  - Create two user groups (mostly email based), one each for Linux and MAC OS X – encourage scientists and others to share experiences with their systems, fostering questions and answers. Encourage a member of CPT to contribute to the discussion, tap into the experience and share it within CPT.
- Enable as much autonomy as possible with these systems for now.
  - Allow local administrator privilege on laptops.
  - Allow user to get RedHat patches (if they want to). Red Hat users can keep their systems up to date with their online *up2date* tool, and do not necessarily have to rely on CPT to do this. Users need to have the license to do that.

- Make a list of “security risk” software for users to avoid.
- Allow users to install new software on their own: users will typically install IRAF, IDL and other reduction utilities as well as some windows utilities.
- Initially, CPT would not be expected to support the full configurations purchased.
- Allow dual boot on laptops as a reasonable interim solution. In the longer term, users desire some supported elegant solution to use Windows applications if it is found either by CPT or the science staff.
- Add Sun disk space on a Unix system, for Linux users to use for Unix-specific applications.
- Note: CPT should make it very clear that if they cannot accelerate the schedule, why this cannot occur. Scientists urge CPT to acknowledge that Linux and MAC OSX systems do exist, are on the network and are pretty much running without severely compromising the system. Duplicating those systems for others would accelerate the process of meeting the scientists’ needs.
- Note: Users would like to be warned of security issues and problems, but users do want to install their own applications. Users wish to retain installation disks. Disks distributed with systems and provided when software is purchased should NOT be confiscated. This point was mentioned in the strongest terms. Users who did not have disks have been stranded repeatedly away from the Institute with no recourse to fix problems. There is a very intense desire to eliminate the policy of confiscating disks.

### ***Longer Term Priorities***

- Create an environment where users have clear choices for future purchases based on their assessment and balance of cost, performance, support, etc.
- Users will expect standard SW and systems to be installed and patched by CPT. Users feel that for Linux, this is pretty easy since they are doing this themselves now.
- Work with the scientific staff subgroup to plan how major systems would be backed up. Work with individuals to backup large desktop storage systems in a rational fashion (once a week?)
- Continue to allow local administrator privilege on laptops. Allow administrator privilege on desktops on a case-by-case basis.
- Users would like to be warned of security issues and problems on a regular basis.
- Create a mechanism for backing up Linux systems.
- Create a mechanism for users to backup laptop systems on their own (storage space or local disk backup, e.g., Peerless system).
- Create seamless access between Unix and Linux systems for data/file access. (Note users are using ftp for now). Deploy NFS servers for Linux to the existing science cluster. Create a Linux science cluster (lower priority). Enable MAC systems and Linux to access cross mounted or shared disks – e.g., allow SAMBA clients and shares.
- Deploy new types of peripherals on a case-by-case basis.
- Create server to operate a web server to interoperate with Linux systems – this is primarily a “sharing data” problem.
- Burn an installation CD ROM configured for users to install with commonly used software – this is already done at ESO.

### ***Hardware Purchases***

- Cost will be a factor in some procurement. Some grantees may be very concerned about saving money or being able to provide resources for students and postdocs and so will prefer

- a Linux solution. Other grantees may need more powerful machines and configurations, depending upon their research. Allowing choices is a priority.
- Circulation of information regarding performance, especially for research packages would be very helpful. For example, scientists individually have accumulated information about various packages such as IRAF, IDL, AIPS, and AIPS++ that point to specific configurations that might work well for the scientists at large in considering all the relevant factors for new purchases.
  - PC laptops are built specifically for Windows. There are hardware and firmware modifications that are made by the manufacturer to optimize performance for Windows. It is wise to check hardware against Red Hat's Certification list before purchasing to insure that Linux will work properly and contain all the necessary drivers etc.
  - Procurement of hardware in general and specifically Linux hardware needs to be handled carefully. Scientists are aware of the pitfalls of executing procurement through buyers without deep technical background. A danger is that requirements might be stated loosely enough that low bids from hardware manufactures are solicited and accepted for computers that cannot address the user's actual needs. CPT could assist the science staff by being willing to collaborate on crafting language that would enable scientists to sole source or purchase systems robust enough to address their computing needs.
  - There are some very nice cheap disk backup solutions available. This could address the mutual scientist and CPT concern about backups of laptop systems and allow scientists to backup their systems at will. CPT could help by pro-actively circulating information they discover and users discover about such systems.

## **Interviewees**

Mark Dickinson  
Andy Fruchter  
Mauro Giavalisco  
Jesus Maíz-Apellániz  
Chris O'Dea  
Larry Petro  
Bernie Rauscher  
Mike Regan  
Dave Soderblom  
Roeland van der Marel  
Bob Williams

Additional input from Ralph Bohlin and Shardha Jongee

## **Task Group**

Carol Christian (lead)  
Stefano Casertano  
Mark Clampin  
Roelof De Jong  
Knox Long  
Antonella Nota  
Massimo Stiavelli  
Frank Summers