*Astronomical Data Analysis Software and Systems XVII* **P10.9**
*ASP Conference Series, Vol. 394, © 2008*
*R. W. Argyle, P. S. Bunclark, and J. R. Lewis, eds.*

# System Administration Best Practices: An Astronomical Perspective

Patrick L. Shopbell

*California Institute of Technology, 770 S. Wilson Avenue, Pasadena, CA, 91125, USA*

**Abstract.** The literature contains numerous discussions of best practices concerning the administration of computing systems and networks. In this paper, we review some of these strategies and techniques, with a special emphasis on those which are particularly applicable in the field of Astronomy. We remind the reader of many common practices of importance, such as communication (both upward and downward) and documentation, but moreover emphasize a number of issues critical for the successful management of computer systems used in astronomy. As an example, the large, and rapidly increasing, data flows in astronomy today, as well as the increasing use of pipeline software on the desktop, are creating a much more dynamic environment with respect to data backups. We draw upon the experience of ourselves and others in this field over the past twenty years to suggest a few best practices for system and network administration which should be especially useful for those in the field of Astronomy.

## 1. Basic Principles of System Administration

The management of computing systems and networks presents challenges arising from both the variety and rapidity of change in the field of information technology (IT). As such, every computing environment is as individual as the variety of requirements and people being served. In this paper, we briefly present an overview of a selection of "best practices" for system administration, following which we present some observations of those areas of generic IT in which the field of Astronomy brings special considerations.

The underlying objectives of system and network administration can be expressed in a way that is representative of essentially all computing environments. These are the "goals" of a well-managed computing system:

**reliability** - User systems can now tolerate less downtime than in the past, due to increasing remote access opportunities. This includes hardware, software, and networking. Real-time systems, such as observatory systems and archives can tolerate even less.

**stability** - Users expect consistency across the board, between users, machines, environments, and interfaces (e.g., command-line, web, and mobile web).

**accessibility** - Data needs to be available anywhere, anytime.

**manageability** - System administrators need to do more with less. Efficiency is key.

**scalability** - Computing and Astronomy are both in eras of rapid change and increasing interdependency. Hardware and software must remain flexible.

## 2. Selected System Administration Best Practices

The practice of system administration has been reviewed and discussed by numerous on-line and print documents (see References). Most of these discussions are relatively generic, with occasional emphases on architecture-centric environments (e.g., Windows-based networks). We have collected here a selection of general system administration practices:

1. State knowledge is critical: an understanding of the current state of your hardware, operating systems, and users gives you a baseline for problem diagnosis, security management, and planning for growth. Sample tools: log analysis, system monitors and alarms, Tripwire.
2. Communication is paramount: between admins and their users (down), between admins and management (up), with vendors (external), and amongst fellow admins (lateral). The more widely information is dispersed, the happier the user environment tends to be. Sample tools: helpdesk, e-mail lists, user education (web pages, manuals, FAQs), policy statements.
3. Standardize and automate as much as possible. Lower hardware costs are leading to a higher system/admin ratio, requiring automated installs, patching, and monitoring. Develop habits and mantras to handle smaller tasks in a more automated fashion as well. Sample tools: scripting languages, JumpStart/ KickStart installs, cfengine, OS consolidation agreements.
4. Document, document, document. As in software development, good documentation in system administration is critical. User documentation leads to more informed and happier users; maintenance documentation leads to more consistently managed systems and more quickly debugged system problems. Sample tools: wikis, blogs, asset databases.
5. Software application maintenance is a balance between functionality and stability. Major upgrades (e.g., Microsoft Office 2007) may require extensive testing and rollout periods. There is no substitute for planning.
6. Security affects all aspects: hardware, software, network. It often requires dedicated personnel to guard against external and internal threats. The development of security policies is critical. Sample tools: Kerberos, VPN, forensic tools, TCP wrappers, nmap.
7. Profitability (monetary and otherwise) often maps directly to reliability. A primary goal of many of the above practices is to ensure system availability for internal and external users, despite an every increasing complexity, growing numbers of systems, and security threats. Sample tools: redundant hardware elements, load-balanced and failover systems, industrial-strength operating systems.
8. This is the information age. Processing power is cheap, making data reliability and security a cornerstone of system management. Data backup equipment, procedures, and policies are more important than ever, as are disaster recovery plans. All such plans must be tested and revised regularly.

## 3. An Astronomy Perspective

The general system administration practices enumerated in the previous section each can be interpreted in the context of the field of Astronomy. The current rapid evolution of the practice of astronomy, together with its ever increasing volumes of data, makes system administration in this field a very dynamic endeavor. What follows are the same generic best practices of system administration from the previous section, but described with an emphasis on their implementation in the field of Astronomy. These points should be thought of as supplemental to the points in the previous section.

1. State knowledge is critical: distributed data and collaboration among widespread users make the concept of a 'system state' much more complex. Information often exists in multiple, perhaps redundant, locations (e.g., laptops). Significant effort must be placed into understanding and classifying data and its distribution.
2. Communication is paramount. Astronomy often involves disparate, sometimes remote, locations and personnel. Meanwhile, communication mechanisms are expanding: mobile phone, Skype, IM, etc. One must find a communication paradigm that works for everyone.
3. Standardize and automate as much as possible. An explosion in mobile computing has essentially doubled the number of systems to be installed and managed. Furthermore, research-related fields often use compute and disk farms of numerous inexpensive nodes, which demand automation in order to be managed by a staff of reasonable size. Sample tools: Rocks.
4. Document, document, document. Astronomy has a plethora of special-purpose hardware, instruments, software applications, etc. Good documentation is the only way to ensure successful use and maintenance of such unique systems.
5. Astronomy has a large number of field-specific applications, many of which require astronomical knowledge to adequately support. Moreover, maintenance to the latest versions is often not optional (e.g., APT, GMS). System stability can be a difficult goal.
6. Security affects all aspects: hardware, software, network. But in Astronomy the motivation for system security is primarily to ensure system reliability. Few astronomical sites are concerned with the theft of proprietary information, identity theft, credit cards, etc. We maintain security to a level that can ensure reliable and robust operation of our systems. Sample tools are defensive in nature: Antivirus, firewalls, Crack, ssh.
7. In Astronomy, the distribution of data and its requirement for wide availability have led to distributed computing and storage on new scales (e.g., VO). In addition, research-related fields such as Astronomy often utilize large-scale and lengthy computations, which require a high level of individual system reliability and stability. Scheduling of maintenance windows becomes paramount, as does the competing need for a rapid response to security and system bugs.
8. Astronomy is a data-centric field. The flow of data from observatories - both real and virtual - continues to increase, placing new demands on storage systems and backup procedures. The need for placing these data into the hands of the user, no matter where they are, is leading to a much more dynamic and distributed data environment. We will continue to see an evolution of techniques to ensure data integrity and security.

## 4. Conclusions

In this paper, we have reviewed a selection of best practices for system and network administration, and we have re-interpreted these practices in light of their implementation in the field of Astronomy. For those of us in the field of system administration within Astronomy, we must constantly confront the challenges of enabling our users in an extremely data-centric field, despite - or hopefully, by embracing - the rapidly changing computing systems and networks available to us today. This paper represents our initial thoughts in this field, and we plan to continue to develop the ideas presented herein into more complete and useful documents elsewhere.

## References

Avoloi, F. M. 2000, Best Practices in Network Security, http://www.networkcomputing.com/ 1105/1105f2.html?ls=NCJS_1105bt

Limoncelli, T. A. 2005, Time Management for System Administrators (Sebastopol, CA: O'Reilly)

Limoncelli, T. A. & Hogan, C. 2002, The Practice of System and Network Administration (Boston, MA: Addison Wesley)

Traugott, S., Huddleston, J., & Traugott J. C. 2007, Infrastructures.Org, www.infrastructures.org