Software packages can become difficult to manage as a network of computers grows in size. This poster covers the Arctic Region Supercomputing Center’s (ARSC) initial evaluation of the Spacewalk and Pulp software suites with the intent of potentially improving software repository (repo) management.

There are two sides to the RPM (Red Hat Package Manager) management system under consideration: the mirrored RPM repos and control of RPMs on the machines on the network. Mirroring a RPM repo involves several steps (organizing the repos, creating the repos, populating the repos, updating the repos) which makes management tedious. RPM control can be done by manually pushing updates to computers on the network, however such an approach is prone to human error and limitations (e.g. installing RPMs to the wrong computer, installing the wrong packages, etc.). RPM management systems regarding package installation help keep track of what machines need updates, as well as simplify the process of applying those updates.

ARSC currently uses comparatively simple methods for repository management. The process for updating the mirrored repositories has only been scripted, and system updates are conducted using the pdsh tool. An updated repo management system will vastly reduce the time required to keep machines up to date and the RPM repositories operating nominally.

To fulfill the requirements for the repository management system, we decided to use two software suites: Spacewalk and Pulp.

Spacewalk provides a large suite of tools focused on monitoring the software installed on computers under its management. Its ability to keep track of the state of RPMs on each machine enables it to report when a computer has available updates. In addition, it offers organizational features, enabling the administrator to group computers, and monitor the packages on that group as a whole.

Pulp handles the other side of the repository management system by providing a simple way of managing the mirrored RPM repositories. It automatically updates the mirrored RPM repo and can schedule times to synchronize it’s repos with their sources. Additionally, and most importantly for ARSC’s purposes, it provides an option to add individual RPMs to custom RPM repositories. ARSC produces several programs in house, which are hosted in our own RPM repositories. Pulp’s ability to automate the process of adding an RPM to a repository will save time interacting with the repos.

To begin testing the functionality and usability of these suites, we set up three test workstations to act as a production environment. Pulp was set up on a host machine named Cygnus and was set to mirror a few of the RPM repositories ARSC pulls from. Pulp then provided the mirrored repos on the internal network for another host called Slug. The Spacewalk suite (including all necessary databases and dependency packages) was installed on Slug. Spacewalk was configured to retrieve all it’s packages from the Pulp repositories on Cygnus. Lastly, the Spacewalk client was installed on a host called Lambda, which simply acted as an ordinary workstation. Lambda listened for updates and requests from the Spacewalk server, allowing updates to be pushed out from the Spacewalk server to Lambda rather than the Spacewalk server having to wait for Lambda to check in.

Testing the Spacewalk and Pulp suites has thus far proven promising. While each group of software has its quirks, they together fulfill ARSC’s repository management needs. Pulp efficiently mirrors external RPM repositories and provides easy to use commands to create and modify new repositories for ARSC’s software. Spacewalk has proven itself capable of serving the RPMs from the Pulp repositories to the Spacewalk clients in a timely manner, thereby simplifying software updates and installations.

There are a few quirks (some Spacewalk web features don’t work, problems with repo permissions) in the software however, which is cause for more testing. For software to be put into production environments, it must be stable and it’s flaws as well as their effects on the surrounding environment, must be well understood. Spacewalk is still under heavy development, so it is more prone to errors than Pulp. Therefore its limits will have to be tested much more thoroughly. It is likely ARSC will continue testing for several months to come.

References
Pulp: http://pulpproject.org
Spacewalk: http://spacewalk.redhat.com