
| | |
|--------------------|---|
| Title | Linuxbrew and Homebrew-Science to Navigate the Software Dependency Labyrinth |
| Author | <i>Shaun Jackman, Inanc Birol</i> |
| Affiliation | Genome Sciences Centre, British Columbia Cancer Agency |
| Contact | sjackman@bcgsc.ca |
| URL | http://linuxbrew.sh/ , https://github.com/Homebrew/homebrew-science |
| License | BSD 2-Clause |
| Keywords | software installation, package manager, open source, reproducible research, high performance computing |

Installing software is at best a tedious experience, and is often a distressing experience. Each operating system (OS) provides its own package manager to install software and manage dependencies: for example, apt-get and yum for the various distributions of Linux, and Homebrew is popular for Mac OS. Using the package manager provided by the system is not however without its own set of challenges. The system's package manager may require administrator access to the machine, which is typically not available on high performance computing clusters. It installs versions of software from the era of the OS, and clusters are notorious for running ancient distributions of Linux. Current bioinformatics software is often not yet packaged and provided only as source. Compiling software from source can vary from difficult to impossible, for example when the compiler and libraries provided by the operating system are a decade old. Manually navigating the recursive dependency chain of the tool and its dependencies, and their dependencies, can feel like a labyrinth with no end, and can even result in conflicting dependencies that are mutually exclusive and impossible to satisfy.

Linuxbrew is a package manager for Linux derived from Homebrew, the Mac OS package manager. It is a cross-platform utility, compatible with any distribution of Linux and version of Mac OS released in the last decade, allowing you to use the same package manager on both your Linux server and your Mac laptop. It can be installed in your home directory, and does not require administrator access. Using Linuxbrew, challenging tasks are made easy; for example installing a modern compiler in your home directory takes a few minutes, even on an ancient distribution of Linux.

Homebrew-Science is a collection of scientific software packages installable by either Linuxbrew or Homebrew. A third of the 600 software packages available on Homebrew-science are bioinformatics tools. Software packages are maintained up-to-date by a fervent community of over 400 contributors, and the scripts to install historical versions of software are retained in version control.

Repeating an analysis of data starts with obtaining the original data and installing the software used for that analysis. Linuxbrew streamlines the installation of software in a repeatable fashion. Having the data and software is not itself sufficient, but is certainly a necessary component of reproducible science.