Installing and supporting scientific software for users of a HPC system is a time consuming and complex task. EasyBuild and Lmod are community-driven tools commonly used in combination to simplify the process of installing and managing software on a HPC system.

With the introduction of CVMFS, the workflow of building software is unchanged, and ACCRE management software ensures that each compute node is linked to the software built against its particular CPU family. However, demands for external cloud usage and interactive tool access have increased. As a result, ACCRE management software providers have had to develop strategies for managing software access and providing a seamless experience for users.

Practically, this means providing a client that can access software from CVMFS or GPFS, and how external cloud compute nodes or user workstations can access the software stack. This is particularly important for users who require access to proprietary software or who need to access software specifically compiled for their CPU architecture family. This framework has been used to provide access to the software stack hosted on public OSG OASIS servers, while a private ACCRE CVMFS instance allows only internal access to proprietary code. Additional redirection allows for individual compute nodes to access software specifically compiled for their CPU architecture family. This has allowed us to meet the above goals at our facility.

To provide fault tolerance and handle client load, a set of Stratum-1 CVMFS servers may be used to serve as a backup to the Stratum-0 servers. In the event of a failure on the Stratum-0 server, the links will revert to the identical files hosted on GPFS. This has allowed us to meet the above goals at our facility.

The operation of ACCRE with hosting software modules over CVMFS has so far shown this to provide better consistency, reliability, and improved cloud capability over our traditional deployment to read-only network filesystems. While Lmod and ACCRE have commonly been used in combination in HPC environments, here we suggest that the further combination of Lmod, EasyBuild, and CVMFS forms a complete solution for managing and distributing scientific software to cluster users and allows users to also access the same software configuration from external workstations or other resources.

REFERENCES


Cloud Deployment

Proof-of-concept cloud deployment of the open source portion of the software stack was tested on an Amazon EC2 medium node with 2 virtual CPU cores and 4 GB of memory, using a CentOS 7 base image provided by Amazon. Installation of CVMFS was performed using the DIS package manager and the CVMFS configured in just a few lines to connect to OSG OASIS directly. For large cloud deployments, additional Squid proxy servers should be configured.

Lmod 7.1 was installed manually and the appropriate symlinks created for the CPU architecture of the node. This process may be performed in a few minutes. Once configured, the software stack is accessible via the command line exactly as in the ACCRE cluster, except that the proprietary software modules are neither available nor displayed.

With a simple installation script, a new virtual machine could be provisioned and ready on AWS in under 10 minutes!