Software installation and condition data distribution via CernVM FileSystem in ATLAS
On behalf of the ATLAS collaboration

The ATLAS Collaboration is managing one of the largest collections of software among the High Energy Physics Experiments. Traditionally this software has been distributed via rpm or pacman packages, and has been installed in every site and user’s machine, using more space than needed since the releases could not always share common binaries. As soon as the software has grown in size and number of releases this approach showed its limits, in terms of manageability, used disk space and performance. The adopted solution is based on the CernVM FileSystem, a fuse-based http, read-only filesystem which guarantees file de-duplication, on-demand file transfer with caching, scalability and performance.

Changes to CernVM-FS repository are staged in a "build box" using a readwrite file system interface. There is a dedicated installation box for each repository. CernVM-FS tools installed on these machines are used to obtain up-to-date packages, publish new and updated files, and to publish new repository snapshots. The machines act as a pure interface to the database and can be safely re-installed if required.

Installed items:
1. Stable Software Releaes;
2. Nightly builds;
3. Local settings;
4. Software Tools;
5. Conditions Data

Stable Software Releases
The deployment is performed via the same installation agents used in the standard Grid sites. The installation agent is currently manually run by a release manager whenever a new release is available, but can be already automated by running periodically the same procedure or kept in the installation OB, and dynamically queried by the deployment agent when needed. The installation OB is part of the ATLAS Installation System, and updated by the Grid installation service.

Once the releases are installed in CernVMFS, a validation job is sent to each site. For each validated release a tag is added to the site resources, meaning it can be used for with the given release. The installation System can handle both CernVMFS and non-CernVMFS sites, transparently.

Software Tools
The installation infrastructure tools from CernVM-FS is an ideal tool for the Physicist end-user to access Grid Resources and for analysis activities on locally managed resources, which can range from a grid-enabled Data Center to a user desktop or a CernVMVirtual Machine. The software, comprising multiple versions of Grid Middleware, front-end Grid job submission tools, analysis frameworks, compilers, data management software and User Interface with diagnostics tools, is first tested in an integrated environment and then installed on the Grid site by the same mechanism that allows local Grid installation on non-privileged Tier accounts. The User Interface integrates the Grid software with the Stable and Nightly ATLAS releases with the CernVMFS Data, the latter two residing on different CernVM-FS repositories.

Local Settings
The local parameters needed before running jobs, ATLAS commonly uses an hybrid system, where the generic values are kept in CernVM-FS, with a possible override at the site level, or an additional location point by an environment variable, defined by the site administrators to point to a local shared directory. Periodic installation jobs write the site-wide configuration in the shared local area of the sites. The fully enabled CernVM-FS option, where the local parameters are dynamically retrieved by the ATLAS Grid Information System (AGIS) is being tested.

Performance
Some measurements on CernVM-FS access times against different sites in the ATLAS have shown that the performance of CernVM-FS is either comparable or better performing when the system caches are filled with non-event data.

The results of a metadata-intensive test performed at the T2-martinsican show the comparison of performance between CernVM-FS and CHS over 6GB. The tests are equivalent to the operations that are performed by each ATLAS site in the Tier-2. The Tier-2 access times with CernVM-FS are ~3s when the cache is filled and ~5s when the cache is empty, compared with the CHS values of ~15s and ~25s respectively with page cache on and off. The test performed over the Grid is ~2× even faster than the standard distributed file systems.

According to other measurements, CVMFS also performs satisfactorily better than NFS in wide area networks with additional bonus that the system can work in disconnected mode.

---