Contribution ID: **381**   Type: **Parallel**

Using Xrootd to Federate Regional Storage

Tuesday 22 May 2012 14:20 (25 minutes)

While the LHC data movement systems have demonstrated the ability to move data at the necessary through-put, we have identified two weaknesses: the latency for physicists to access data and the complexity of the tools involved. To address these, both ATLAS and CMS have begun to federate regional storage systems using Xrootd. Xrootd, referring to a protocol and implementation, allows us to provide data access to all disk-resident data from a single virtual endpoint. This "redirector" endpoint (which may actually be multiple physical hosts) discovers the actual location of the data and redirects the client to the appropriate site. The approach is particularly advantageous since typically the redirection requires much less than 500 milliseconds (bounded by network round trip time) and the Xrootd client is conveniently built into LHC physicist's analysis tools.

Currently, there are three regional storage federations - a US ATLAS region, a European CMS region, and a US CMS region. The US ATLAS and US CMS regions include their respective Tier 1 and Tier 2 facilities, meaning a large percentage of experimental data is available via the federation. There are plans for federating storage globally and so studies of the peering between the regional federations is of particular interest.

From the base idea of federating storage behind an endpoint, the implementations and use cases diverge. For example, the CMS software framework is capable of efficiently processing data over high-latency data, so using the remote site directly is comparable to accessing local data. ATLAS's processing model is currently less resilient to latency, and they are particularly focused on the physics n-tuple analysis use case; accordingly, the US ATLAS region relies more heavily on caching in the Xrootd server to provide data locality.

Both VOs use GSI security. ATLAS has developed a mapping of VOMS roles to specific filesystem authorizations, while CMS has developed callouts to the site's mapping service. Each federation presents a global namespace to users. For ATLAS, the global-to-local mapping is based on a heuristic-based lookup from the site's local file catalog, while CMS does the mapping based on translations given in a configuration file.

We will also cover the latest usage statistics and interesting use cases that have developed over the previous 18 months.

Authors: BOCKELMAN, Brian Paul (University of Nebraska (US)); GARDNER, Robert (UNIVERSITY OF CHICAGO)

Co-authors: HANUSHEVSKY, Andrew (STANFORD LINEAR ACCELERATOR CENTER); Prof. YAGIL, Avi (Univ. of California San Diego (US)); BRADLEY, Daniel Charles (High Energy Physics); Mr LESNY, David (Univ. Illinois at Urbana-Champaign (US)); BENJAMIN, Doug (Duke University (US)); WURTHWEIN, Frank (UCSD); DONVITO, Giacinto (Universita e INFN (IT)); ITO, Hironori (Brookhaven National Laboratory (US)); Dr SEVERINI, Horst (University of Oklahoma (US)); Mr SFILIGOI, Igor (University of California San Diego); BLOOM, Kenneth (University of Nebraska (US)); Dr BAUERDICK, Lothar (FERMILAB); TADEL, Matevz (Univ. of California San Diego (US)); ERNST, Michael (Unknown); Dr RIND, Ofer (BROOKHAVEN NATIONAL LABORATORY); MCGUIGAN, Patrick (University of Texas at Arlington (US)); WILLIAMS, Sarah (Indiana University (US)); Dr MC KEE, Shawn (University of Michigan (US)); Prof. DASU, Sridhara (University of Wisconsin (US)); YANG, Wei (SLAC National Accelerator Laboratory (US))

Presenter: BOCKELMAN, Brian Paul (University of Nebraska (US))