



Contribution ID: 32

Type: **Oral Presentation**

Evolution of the ATLAS data placement model

Tuesday, 12 April 2011 12:00 (30 minutes)

Overview

The ATLAS experiment at the LHC is fully relying on the usage of grid computing for its offline data placement, processing and analysis. For data placement the ATLAS Computing Model defines a set of policies, which establish a hierarchical tier organization according to the network topology which was laid out for data distribution. However, since the original date of the Computing Model, network capabilities have significantly increased and it is convenient to gradually start relaxing some of the imposed boundaries. This talk will focus on the work that is being carried out in the ATLAS Distributed Data Management project in order to evaluate more dynamic constraints and provide the necessary framework for network link commissioning.

Impact

New data brokering models such as PanDA Dynamic Data Placement have been recently introduced by ATLAS. The idea behind this model is to replicate datasets from any Tier1 to any Tier2 only after they have exceeded a popularity threshold, thereby eliminating the replication of unpopular datasets. The presented work is a first attempt in the ATLAS Distributed Data Management project to optimize these cross-cloud transfers, improve the network usage and provide the necessary statistics needed for link commissioning activity with the final goal of reducing cloud boundaries.

A similar initiative exists in the CMS experiment, another of the LHC experiments. Their approach is somewhat different, as instead of measuring the transfer events, the statistics are collected by parsing html files retrieved from each one of the gLite File Transfer Service servers. Also the statistics are not fed back into the system in order to optimize transfers, but are only displayed for link commissioning.

Description of the work

The ATLAS Distributed Data Management is the project built on top of the WLCG middleware and is responsible for the replication, access and bookkeeping of the multi-Petabyte ATLAS data across the Grid while enforcing the policies defined in the ATLAS Computing Model. Following this model ATLAS sites are grouped into ten clouds by geographical and organizational reasons. Each cloud is formed by one Tier1 that must provide a high level of service and is responsible for data storage and reprocessing and several Tier2s and Tier3s, which are used for analysis and Monte Carlo production and depend directly on the Tier1.

Network connections between the Tier0 (CERN) and the Tier1s are guaranteed by the Optical Private Network and inside a cloud are generally performant, while inter-cloud links are less guaranteed. Thus, Tier1s usually act as the access point to get data in and out of the cloud and direct cross-cloud communication between Tier2s is generally avoided. However, since networking capabilities have significantly evolved, the Computing Model is moving towards more dynamic data management policies and the cloud boundaries are gradually being reduced in a controlled fashion.

Consequently, the data distribution framework in the ATLAS Distributed Data Management project has been instrumented to measure the durations of gLite File Transfer Service (FTS) transfers between sites and store them in an Oracle database. The transfer statistics will be used as feedback to optimize the source selection and choose between multi-hop transfers through the Tier1s or direct cross-cloud transfers. The statistics are also visualized in a dynamic web page in order to monitor the throughput performance of the network links. In parallel, an ad-hoc load generator will trigger transfers on the complete mesh of ATLAS sites and will provide the information for a first attempt of link commissioning.

URL

<http://bourricot.cern.ch/dq2/ftsmon/>

Conclusions

The collected transfer statistics are known directly by the FTS servers, but are not made available through an programmatic interface. Having access to them is not an ATLAS particular interest, but can be of useful to any grid based community. This contribution wishes to open the discussion between the different user communities with the ultimate goal of moving in the future towards a common, central solution.

Primary authors: Mr TYKHONOV, Andrii (Jozef Stefan Institut); Mr BARREIRO MEGINO, Fernando Harald (CERN); Dr CAMPANA, Simone (CERN)

Presenter: Mr BARREIRO MEGINO, Fernando Harald (CERN)

Session Classification: Data Management

Track Classification: Data Management - Results