Contribution ID: **31**   Type: **Oral Presentation**

An insight into the ATLAS Distributed Data Management

Tuesday, 12 April 2011 11:00 (30 minutes)

Overview

ATLAS, one of the four LHC experiments, fully relies on the usage of grid computing for offline data distribution, processing and analysis. 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 more than 100 distributed grid sites. It enforces data management policies decided on by the collaboration and defined in the ATLAS computing model. It is in production since 2004 and to date is considered one of the largest open data management environments and an example of a global multi-grid hybrid system. Tis contribution will give an overview of the architecture, operational and deployment strategies of this highly automated system, as well as details about different subsystems and monitoring solutions that could be of interest for other communities.

Impact

Given the scale of the project, ATLAS Distributed Data Management can be considered an important example of distributed data management on the grid. Although the implementation is experiment specific, it is based on common underlying middleware and is composed of independent subcomponents, auxiliary systems and monitoring solutions that can be of inspiration and reimplemented to fit the needs of any grid community. Relatively simple subcomponents such as the storage space accounting, which was initially implemented to account the used, free and total storage space of grid storage endpoints and was extended to allow the breakdown of the used space by different metadata, have been claimed to be needed by different communities in- and outside the WLCG community. The communication of the storage space accounting with a centralized site exclusion service or an automatic site cleaning agent are the next steps, in order to avoid replicating data to full sites or to trigger the cleaning of full sites respectively. In the case of ATLAS, these simple tools have helped to significantly reduce manual operations and improve day to day operations.

Description of the work

ATLAS Distributed Data Management is the system that manages the experiment's detector, simulated and user data while enforcing the policies defined in the Computing Model. It provides functionality for data placement, deletion, bookkeeping and access on a hierarchic grid model composed of around 100 sites with heterogeneous storage technologies, services and protocols.

The system is currently managing 50 PetaBytes of data corresponding to almost 200 million files and is being used to achieve aggregated throughput rates far beyond the initial requirement of 2GB/s, having reached a throughput peak of over 10GB/s. To ensure further scalability, the core of the system has been designed as a set of independent agents which work around a global database, the Central Catalogues. This architecture allows to run distributed, fault-tolerant services that interact with the grid resources and checkpoint centrally

the status of the requests.

Given the volume of files managed, it is critical to reduce the load on operations by providing compact monitoring information, curing automatically inconsistencies and automate systems as far as possible.

The talk will give an overview of the general architecture and furthermore explain in detail several independent subcomponents targeted towards resource optimization, such as the storage space accounting, the automatic site cleaning agent and the centralized site exclusion service and how these can be extended to interact between each other. The ideas and concepts presented will provide inspiration for any Virtual Organization that is currently planning to move their data to the grid or working on improvements to their usage of grid, network and storage resources.

URL

http://bourricot.cern.ch/dq2/accounting/global_view/30/

Conclusions

Data management is an area of constant expansion: Both experiment and user activity keep increasing and oblige ATLAS DDM to ensure the future scalability of the system by automating manual operations, optimizing the usage of available resources and adapt to evolving middleware, more performant infrastructure and new technologies.

To mention one example, DDM is currently replacing its messaging infrastructure, based on open protocols such as http, by more recent, asynchronous message queuing technologies, which should enhance the independence of subcomponents and therefore the overall fault-tolerance.

Another area of development, given the evolution of the network throughput, is the development of the hierarchic data distribution policies towards a more dynamic, relaxed Computing Model where gradually regional boundaries are relieved.

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

Presenter: Mr BARREIRO MEGINO, Fernando Harald (CERN)

Session Classification: Data Management

Track Classification: Data Management - Results