Contribution ID: **3**   Type: **not specified**

ATLAS DDM/DQ2 & NoSQL databases: Use cases and experiences

Tuesday 7 June 2011 12:00 (30 minutes)

The Distributed Data Management System DQ2 is responsible for the global management of petabytes of AT-LAS physics data. DQ2 has a critical dependency on Relational Database Management Systems (RDBMS), like Oracle, as RDBMS are well-suited to enforce data integrity in online transaction processing application. Despite these advantages, concerns have been raised recently on the scalability of data warehouse-like work-load against the transactional schema, in particular for the analysis of archived data or the aggregation of data for summary purposes. Therefore, we have considered new approaches of handling vast amount of data. More specifically, we investigated a new class of database technologies commonly referred to as NoSQL databases. This includes distributed filesystem like HDFS that support parallel execution of computational tasks on distributed data, as well as schema-less approaches via key-value stores, like HBase, Cassandra or MongoDB. These databases provide solutions to particular types of problems: for example, NoSQL databases have demonstrated that they can scale horizontally, deliver high throughput, have automatic fail-over mech-anisms, and provide easy replication support over LAN and WAN.

In this talk, we will describe our use cases in ATLAS, and share our experiences with NoSQL databases in a comparative study with Oracle.

Proposed speaker

V.Garonne

Authors: Dr MOLFETAS, Angelos (CERN-PH-ADP-CO);   ZANG, Donal (IHEP);   Mr DIMITROV, Gancho (BNL);   Dr CANALI, Luca (CERN-IT-DB);   Mr LASSNIG, Mario (CERN-PH-ADP-CO);   Dr GARONNE, Vincent (Conseil Europeen Recherche Nucl. (CERN)-Unknown-Unknown)

Presenter: Dr GARONNE, Vincent (Conseil Europeen Recherche Nucl. (CERN)-Unknown-Unknown)

Session Classification: Implementations II