



Contribution ID: 352

Type: Poster

Monitoring of services with non-relational databases and map-reduce framework

Thursday, 24 May 2012 13:30 (4h 45m)

Service Availability Monitoring (SAM) is a well-established monitoring framework that performs regular measurements of the core services and reports the corresponding availability and reliability of the Worldwide LHC Computing Grid (WLCG) infrastructure. One of the existing extensions of SAM is a Site Wide Area Testing (SWAT), which gathers monitoring information from the worker nodes via instrumented jobs. This generates quite a lot of monitoring data to digest, as there are several data points for every job and several million jobs are executed every day. The recent uptake of non-relational databases opens a new paradigm in the large-scale storage and distributed processing of systems with heavy read-write workloads. For SAM this brings new possibilities to improve its model from performing aggregation of measurements to storing raw data and subsequent re-processing. Both SAM and SWAT are currently tuned to run at top performance reaching some of the limits in storage and processing power of their existing Oracle relational database. We investigated the usability and performance of non-relational storage together with its distributed data processing capabilities. For this, several popular systems have been compared.

In this contribution we describe our investigation of the existing non-relational databases suited for monitoring systems covering Cassandra, HBase, OpenTSDB and MongoDB. Further, we present our experiences in data modeling and prototyping map-reduce algorithms focusing on the extension of the already existing availability and reliability computations. Finally, possible future directions in this area are discussed, analyzing the current deficiencies of the existing Grid monitoring systems and proposing solutions how to leverage the benefits of the non-relational databases to get more scalable and flexible frameworks.

Primary author: BABIK, Marian (CERN)

Co-authors: Mr COLLADOS POLIDURA, David (CERN); SOUTO MOURE, Fabio (University of Vigo (ES))

Presenter: BABIK, Marian (CERN)

Session Classification: Poster Session

Track Classification: Software Engineering, Data Stores and Databases (track 5)