21st International Conference on Computing in High Energy and Nuclear Physics **CHEP2015** Okinawa Japan: April 13 - 17, 2015

Contribution ID: **76** Type: **oral presentation**

NoSQL technologies for the CMS Conditions Database

Monday, 13 April 2015 14:30 (15 minutes)

With the restart of the LHC in 2015, the growth of the CMS Conditions dataset will continue, therefore the need of consistent and highly available access to the Conditions makes a great cause to revisit different aspects of the current data storage solutions.

We present a study of alternative data storage backends for the Conditions Databases, by evaluating some of the most popular NoSQL databases to support a key-value representation of the CMS Conditions. An important detail about the Conditions that the payloads are stored as BLOBs, and they can reach sizes that may require special treatment (splitting) in these NoSQL databases. As big binary objects may be a bottleneck in several database systems, and also to give an accurate baseline, a testing framework extension was implemented to measure the characteristics of the handling of arbitrary binary data in these databases. Based on the evaluation, prototypes of a document store, using a column-oriented and plain key-value store, are deployed. An adaption layer to access the backends in the CMS Offline software was developed to provide transparent support for these NoSQL databases in the CMS context. Additional data modelling approaches and considerations in the software layer, deployment and automatization of the databases are also covered in the research. In this paper we present the results of the evaluation as well as a performance comparison of the prototypes studied.

Primary author: SIPOS, Roland (Eotvos Lorand University (HU))

Presenter: SIPOS, Roland (Eotvos Lorand University (HU))

Session Classification: Track 3 Session

Track Classification: Track3: Data store and access