

# 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015)



Contribution ID: 221

Type: poster presentation

## QuerySpaces on Hadoop for the ATLAS EventIndex

The new ATLAS EventIndex catalogue uses a Hadoop cluster to store information on each event processed by ATLAS. Several tools belonging to the Hadoop eco-system are used to organise the data in HDFS, catalogue it internally, and provide the search functionality. This presentation will describe the Hadoop-based implementation of the adaptive query engine serving as the back-end for the ATLAS EventIndex. The QuerySpaces implementation handles both original data and search results providing fast and efficient mechanisms for new user queries using already accumulated knowledge for optimisation. Detailed description and statistics about user requests are collected in HBase tables and HDFS files. Requests are associated to their results and a graph of relations between them is created to be used to find the most efficient way of providing answers to new requests. The environment is completely transparent to users and is accessible over several command-line interfaces, a Web Service and a programming API.

**Primary author:** Dr HRIVNAC, Julius (Laboratoire de l'Accelérateur Lineaire (FR))

**Co-authors:** FAVARETO, Andrea (Università degli Studi e INFN Genova); GLASMAN, Claudia (Universidad Autonoma de Madrid); Dr CRANSHAW, Jack (Argonne National Laboratory (US)); TOEBBICKE, Rainer (CERN); YUAN, Ruijun (Laboratoire de l'Accelérateur Lineaire (FR))

**Presenter:** Dr HRIVNAC, Julius (Laboratoire de l'Accelérateur Lineaire (FR))

**Track Classification:** Track3: Data store and access