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

Contribution ID: **222** Type: **oral presentation**

Distributed Data Collection for the ATLAS EventIndex.

Monday, 13 April 2015 15:45 (15 minutes)

The ATLAS EventIndex contains records of all events processed by ATLAS, in all processing stages. These records include the references to the files containing each event (the GUID of the file) and the internal "pointer" to each event in the file. This information is collected by all jobs that run at Tier-0 or on the Grid and process ATLAS events. Each job produces a snippet of information for each permanent output file. This information is packed and transfered to a central broker at CERN using an ActiveMQ messaging system, and then is unpacked, sorted and reformatted in order to be stored and catalogued into a central Hadoop server. This talk describes in detail the Producer/Consumer architecture to convey this information from the running jobs through the messaging system to the Hadoop server.

Primary author: SANCHEZ, Javier (Instituto de Fisica Corpuscular (ES))

Co-authors: FERNANDEZ CASANI, Alvaro (Instituto de Fisica Corpuscular (ES)); SANCHEZ, Javier (Universidad de Valencia (ES)); Dr GONZALEZ DE LA HOZ, Santiago (IFIC-Valencia)

Presenter: SANCHEZ, Javier (Instituto de Fisica Corpuscular (ES))

Session Classification: Track 3 Session

Track Classification: Track3: Data store and access