Contribution ID: **134**   Type: **presentation**

Allocation Optimization for the ATLAS Rebalancing Data Service

Monday 9 July 2018 15:45 (15 minutes)

The distributed data management system Rucio manages all data of the ATLAS collaboration across the grid. Automation such as replication and rebalancing are an important part to ensure the minimum workflow execution times. In this paper, a new rebalancing algorithm based on machine learning is proposed. First, it can run independently of the existing rebalancing mechanism and can be modularised. It collects data from other services and learns optimality as it runs in the background. Periodically this learning agent takes a subset of the global datasets and proposes them for redistribution to reduce waiting times. The user can interact and choose to accept, decline, or override the dataset placement suggestions. The accepted items are shifted continuously between destination data centres as a background service while taking network and storage utilisation into account.

Authors: VAMOSI, Ralf (CERN); LASSNIG, Mario (CERN); SCHIKUTA, Erich

Presenter: VAMOSI, Ralf (CERN)

Session Classification: T6 - Machine learning and physics analysis

Track Classification: Track 6 –Machine learning and physics analysis