

# Automatic rebalancing of data in ATLAS distributed data management

*Tuesday, 11 October 2016 16:30 (15 minutes)*

The ATLAS Distributed Data Management system stores more than 180PB of physics data across more than 130 sites globally. Rucio, the new data management system of the ATLAS collaboration, has now been successfully operated for over a year. However, with the forthcoming resumption of data taking for Run 2 and its expected workload and utilization, more automated and advanced methods of managing the data are needed. In this article we present an extension to the data management system, which is in charge of detecting and forecasting data imbalances as well as storage elements reaching and surpassing their capacity limit. The system automatically and dynamically rebalances the data to other storage elements, while respecting and guaranteeing data distribution policies and ensuring the availability of the data. This concept not only lowers the operational burden, as these cumbersome procedures had previously to be done manually, but it also enables the system to use its distributed resources more efficiently, which not only affects the data management system itself, but in consequence also the workload management and production systems. This contribution describes the concept and architecture behind those components and shows the benefits made by the system.

**Tertiary Keyword (Optional)**

**Secondary Keyword (Optional)**

**Primary Keyword (Mandatory)**

Distributed data handling

**Primary author:** BARISITS, Martin (CERN)

**Session Classification:** Posters A / Break

**Track Classification:** Track 4: Data Handling