

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



Contribution ID: 359

Type: poster presentation

Filesize distribution of WLCG data at the Rutherford Appleton Laboratory Tier1

The Rutherford Appleton Laboratory (RAL) operates the UK WLCG Tier1 facility on behalf of GridPP. Tier 1's provide persistent archival storage (on tape at RAL) and online storage for fast access data analysis. RAL is one of the few Tier-1s which supports data management for all the major LHC experiments, as well as a number of smaller Virtual Organisations. This allows us to compare usage of the infrastructure across many different experiments and use cases.

In this post, we look at the distribution of file sizes between VOs and across different storage work flows within the same VO. We also report on how this distribution impacts standard operational procedures such as tape usage and decommissioning of storage nodes.

Further, we also show how a hypothetical change on data storage policy of log files would change the profile of filesize distribution of a significant subset of data currently being stored at RAL and how this change would impact operations to the advantage of the user communities.

Primary author: DAVIES, Brian (STFC (RAL) GB)

Co-author: DE WITT, Shaun (STFC)

Presenter: DAVIES, Brian (STFC (RAL) GB)

Track Classification: Track3: Data store and access