

Collecting conditions usage metadata to optimize current and future ATLAS software and processing

General Goal of this project: Improve understanding of Conditions data usage by event-wise processing



- This helps us to develop improved tools and resources for future processing!

First step: Build upon an existing collection of metadata about Conditions database folder and global tags. New Information added:

- Conditions DB Metrics: database table row counts, average row volumes and counts of distinct references to external data by folder and folder tag (red columns)
- Conditions Usage patterns of existing representative event data processing jobs (yellow tables)

Using this enhanced information, we successfully built "Customized DB Releases":

- SQLite files containing a subset of Conditions data which provide sufficient data for future specific types of processing to be executed in network isolated environments (i.e. HPC)
 - volume is ~10% of the previous files produced for this purpose.
 - So future jobs can be executed on nodes with more limited resources.

We have many further use cases for this collected metadata:

- These are being explored for better Run 2 ATLAS conditions monitoring and will be useful for the design of corresponding systems in LHC Run 3.

