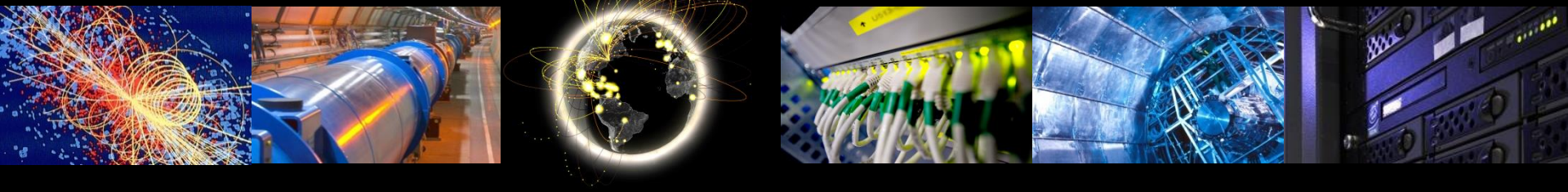
WLCG Archival Group update

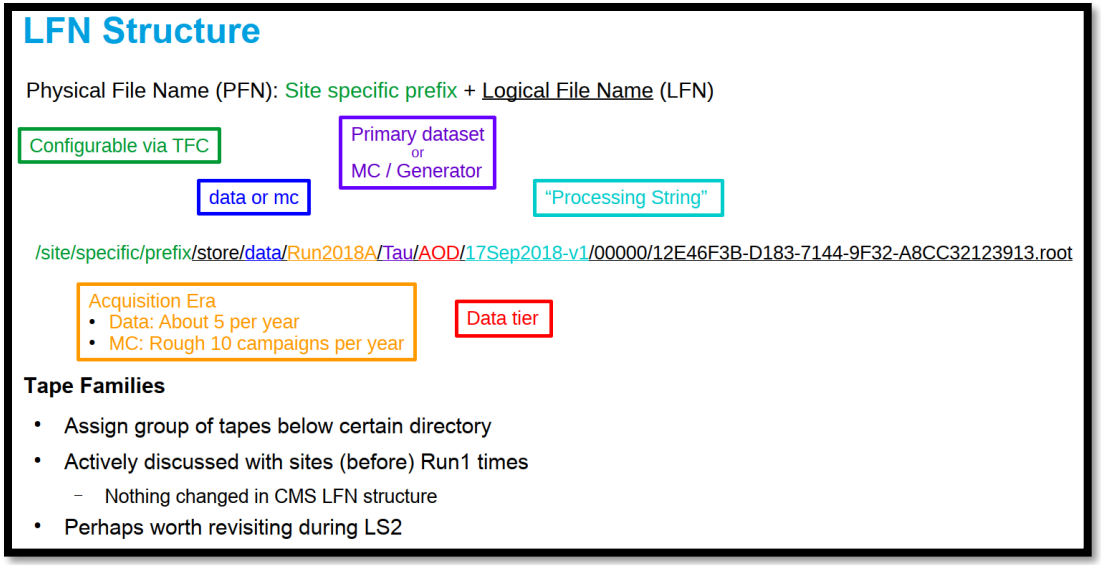
March 2019



Vladimír Bahyl, Oliver Keeble – CERN

Background

- Experiment (recall) requirements as presented at WLCG DOMA general meeting 28 November 2018:
  - https://indico.cern.ch/event/767209/
- Discussion about optimizing resources
- ATLAS & CMS experiments have logical data set centric structure:



- Experiments expect storage systems to benefit from that structure

- Multidimensional problem:
  - Optimize: READ, WRITE, buffer space, REPACK

Data Locality Issue

- Overall READ performance can be more higher if the related files are recalled together
- Issues:
  - Data sets have different sizes (= many files) which are unknown at creation time
  - To guarantee WRITE performance, input streams (files) are spread on multiple tapes
  - To be able to “re-assemble” the data sets, large buffer is required
  - Storage systems data set awareness is limited = do not take into account complex logical name space structures when migrating
    - tapeguy @ TRIUMF highly ATLAS specific
  - Repack reshuffles data further
  - Experiments have a limited knowledge of what they will recall
    - Datasets will be recalled in their entirety ... in unknown combinations

Complex problem

- Measures discussed:
  - Data tags (for explicit logical grouping of files) passed from higher level frameworks (Rucio)?
  - Could larger files help?
  - Hold files on large disk buffer (to complete a data set) before sending it to tape?
  - What about just preserving temporal collocation?
- Possible directions:
  - Increase dataset awareness in tape systems
  - Try to keep a dataset on the smallest number of tapes
  - Multiple datasets will always be recalled simultaneously
    - Optimising recall speed of single datasets is unnecessary
  - Reconsider file sizes
  - Renegotiate “contract” with experiments – shift some investment from capacity to I/O
  - Use repack as an opportunity to optimise locality
- Identify common ground covering many underlying storage systems
- Discussion continues, no conclusions (yet) …