Data Management

No much since Lisbon…

Data Management

- Data placement policy
  - From 2 to 4 replicas depending on popularity
    - Mitigate extremely good LHC efficiency with 2 replicas for real data (for now)
- Disk replica reduction / removal
  - Based on Data popularity
  - For active datasets, 2 replicas is the bare minimum, 4 the maximum
    - Criterion is data accessibility and availability
- Data federation (only for user analysis jobs)
  - Ensured by *Gaudi data federation*
    - Works very well, transparent to users
    - Remote replicas used if local replicas not available
  - User jobs brokering still based on best knowledge data location: jobs to data
    - From file / replica catalog
- Much SRM-based still
  - …but trying to avoid SRM as much as possible
    - In the works, for data movement input and job tURL de-referencing

Baseline for LHCb

Data Management

- Storage reliability
  - Disk storage used as temporary storage
    - Before merging and transfer to more permanent storage
    - As pre-staging for tape datasets processing (could be lower level of service)
  - Important to not loose data (before merging) ⇒ requires reliability
    - Heavy burden to reconstruct it
  - Tape data loss may be dramatic ⇒ requires reliability
    - Used for reconstruction output (single replica)
- No dynamic removal / replication
  - Reaction time scale too slow
  - Not easy to optimize number of replicas (≥2, but…)
- Resource reporting
  - Need reliable storage usage and availability information
    - Currently SRM based
- Data formats
  - Reduce as much as possible event sizes for each use case
    - Using μDST more and more for simulation
    - Turbo processing: directly use online reconstructed data

Popularity and tape parking

Data Management

- Popularity used for removing datasets from disk
  - See talk-295 @ CHEP
  - Removed > 1.5 PB of disk replicas in 2016
  - Minimum number of replicas difficult to optimize
    - Intermediate step with 1 replica before full removal
    - Always keep one tape archive replica
- Users do use local (private) caching
  - Moving datasets to local storage
  - Non-pledged resources, no central operation
- Data parking
  - For LHCb this means
    - Wait before processing data (e.g. until next LS2)
    - Process when enough disk and CPU available
    - "Smart parking": select physics channels to be parked
      - Already done for some 2016 Turbo data (Charm physics)
  - Better than not record data
    - Arbitration not easy (management)