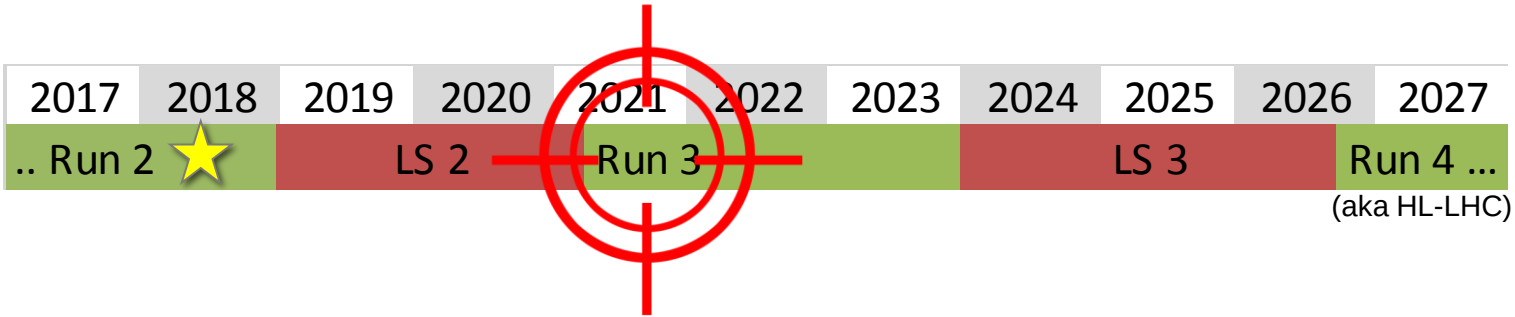
LHCb DOMA Input

Stefan Roiser

WLCG DOMA Kickoff

4 June 2018

Quick Reminder on the LHCb Upgrade Timeline



- Main Upgrade of the experiment happening in Run 3 (2021)
  - Full and final real-time data reconstruction at online resources
  - Trigger output bandwidth increases by several factors

- Run 4 upgrade (HL-LHC) in computing will see minor changes compared to Run3

LHCb Upgrade Distributed Computing in a Nutshell

- LHCbDIRAC will continue to operate distributed computing resources in upgrade era
  - Architecture allows re-engineering, scaling of individual components as needed

- Full exploit of “Turbo” paradigm
  - Final detector alignment and calibration during trigger processing
    - Already successfully exercised for 1/3 or the trigger rate in Run 2
    - → final physics objects out of the trigger → no offline data processing
    - → 90 % of distributed computing resources needed for MC simulation

- Factor 5 increase of luminosity (Run3) & removal of L0 hw trigger & reduction of event size via Turbo
  - → increase of trigger rate by one order of magnitude
  - → LHCb needs for simulated events scales with the above
    - → aiming for 60 to 80 % of fast and full parametrized simulated events
    - Re-decay, tracker-only, RICHless in production, more planned (shower libs, delphes, …)

- → Despite optimizations, offline data storage needs increase by factors

Distributed Workload Management – User Analysis

- User analysis with input data
  - Input data split and grouped according to storage sites
  - Jobs sent to data storage site location (T0, T1, T2D)
  - Fallback mechanism (Gaudi federation)
    - In case the local replica is not accessible → WAN fallback onto possible 2nd replica
  - Push from individual user analysis towards “working group productions”

- User analysis without input data
  - Runs on all WLCG resources

Distributed Workload Management – MC Simulation

- MC Simulation – Step 1, generator and detector response
  - 95 % of CPU used in this step.
  - Output upload of (intermediate) file to “buffer storage” (usually T0 & T1)
  - Runs on all computing resources (WLCG, Boinc, HPC, Cloud, …)
  - No input data needed for this step (start from random seed)

- MC Simulation – Step 2
  - Processed at disk storage site
  - Process remaining workflow (digitization, trigger, (reco), stripping (aka slimming, streaming, skimming)
  - Output upload to disk storage site (T0, T1, T2D)

LHCb Distributed Data Management

- Small number of disk storage sites
  - T0 + (8) T1 sites, 13 T2 sites with storage (>> 300 TB each)
  - User and buffer space on T0/1 sites
  - Data & MC replicated twice on all storage sites
- Mostly ”democratic” data replication policy based on free space
  - Both data and MC
- Launch of data management workflows done manually (takes O(Min))
  - Replication managed by Dirac, offloaded to FTS3 (including staging)
  - Remove replica, delete file done by Dirac
  - Transfer P8 (pit) to CERN done by Dirac

Data Management Catalogs & Data Access

- Dirac File Catalog (~ 20 Mio LFNs)
  - Information on centrally produced data replicas
  - Provides access to logical SEs (can be on same hardware SE)

- Bookkeeping
  - Contains data provenance information (~ 1 Billion files)
    - Origin (LHCb or MC), Conditions (Beam energy), Processing pass (reco, stripping, …), Event Type (mostly for simulation), File type (RAW, RDST, BHADRON.MDST, EW.DST, …)
  - Stores information on all files ever created and data quality information
    - Including location, worker node, cpu time, wall time, …)
  - Defines “data sets” used by physicists for selection of their input files

Space Tokens & Data Access

- Space tokens in use
  - LHCb-Tape (T1D0)
  - LHCb-Disk (T0D1)
  - LHCb-User (T0D1)

- Moving from srm to direct xroot access
  - Reading shall be possible soon for all sites
  - Writing possible on sites where tape and disk are not in the same namespace

Outlook

- Run 3 upgrade step is major for LHCb
  - Offline distributed computing needs increase by several factors
- WLCG developments aiming to reduce resource needs are welcome
  - NB: Different timeline of upgrade program to GPDs
- LHCb volunteers to try especially storage need optimizations early on