LHCb data namespace

This page describes the namespace for LHCb data. By LHCb convention, the SURL or the file is formed by appending this LFN to the SAPath as provided by the sites. if the SAPath terminates with "/lhcb", it is not repeated twice, otherwise it is. We also indicate the SRM v2.2 token descriptions to which these files are written.

There is a need for 4 tape-sets, based on the namespace hierarchy's first 4 "directories" (indicated in bold). Note that for clear reasons of segmentation of data, one of the directories increments with the corresponding year. These are the 4 tape-sets to which we have given arbitrary names:

- lhcb_raw: RAW data
- lhcb_sdst: SDST. **Note** that the directory should be created in advance such that sites can set the tape family...
- lhcb_archive: Archive (replaces all the former T1D1 SRM spaces)
- default: general purpose, including Castor home directories

Changes:

- 21.01.09: added <stream> to RAW path, as discussed in November 08
- 25.01.11: added SDST, change /DST/<stream> to <stream>.DST
- 04.05.11: adapt to the 2011 namespace for real data productions, remove obsolete paths
- 04.04.12: make clear between "full year" (2012) and "short year" (12)
- 23.08.13: for lhcb_raw, no need to split by year as files come anyway chronologically. Reminder that the FULL.DST directory must be created early enough...
- 01.06.15: add RDST since it may replace FULL.DST sometime during summer 2015.
- 02.06.15: add early measurements directory for lhcb_sdst
- 08.02.17: add MDST.DST and specify other types of collisions

Path	SRM Space token	Comment	Tape set
Real Data:			
/lhcb/data/<full year>/RAW/ <stream>/<config>/<runtype>/<run no>/	LHCb-Tape (T1D0)	Raw (grouped chronologically by construction)	lhcb_raw
/lhcb/LHCb/Collision<short year>/FULL.DST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	FULL DST grouped by year	lhcb_sdst
/lhcb/LHCb/Collision<short year>/RDST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	RDST grouped by year	lhcb_sdst
/lhcb/LHCb/Collision<short year>/MDST.DST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	DST for rebuilding MDST grouped per year	lhcb_sdst
/lhcb/LHCb/Collision15em/FULL.DST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	FULL DST for 2015 early measurements	lhcb_sdst
/lhcb/LHCb/<beams><short year>/FULL.DST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	FULL DST for special collisions grouped by year	lhcb_sdst
/lhcb/LHCb/<beams><short year>/RDST/ <prodID>/<subID>/	LHCb-Tape (T1D0)	RDST for special collisions grouped by year	lhcb_sdst
/lhcb/LHCb/Collision<short year>/<stream>.DST/<prodID>/<subID>/	LHCb-Disk (T0D1)	DST grouped by year	none

/lhcb/LHCb/Collision<short year>/<stream>.MDST/<prodID>/<subID>/	LHCb-Disk (T0D1)	µDST grouped by year	none
Simulated Data:			
/lhcb/MC/<MCSet>/SIM/<prodID>/<subID>/	LHCb-Tape (T1D0)	MC hits (very few)	default
/lhcb/MC/<MCSet>/DIGI/<prodID>/<subID>/	LHCb-Tape (T1D0)	MC digits (very few)	default
/lhcb/MC/<MCSet>/GEN/<prodID>/<subID>/	LHCb-Disk (T0D1)	Generator level simulation (HEPMC)	none
/lhcb/MC/<MCSet>/XGEN/<prodID>/<subID>/	LHCb-Disk (T1D0)	Generator level simulation (with MCParticles)	none
/lhcb/MC/<MCSet>/[<stream>.]DST/<prodID>/<subID>/	LHCb-Disk (T0D1)	MC DST grouped by year	none
/lhcb/MC/<MCSet>/XDST/<prodID>/<subID>/	LHCb-Disk (T0D1)	MC eXtended DST grouped by year	none
User Data:			
/lhcb/user/<initial>/<user>/ - with <user> defined in VOMS generic attribute field	LHCb_USER (T0D1)	User files	none
Archives:			
/lhcb/archive/ <LFN>	LHCb-Tape	Tape archive of real and MC data	lhcb_archive
Test Data:			
/lhcb/MC/test/<whatever>	any depending on the test	For testing new version of the simulation	none
/lhcb/test/<whatever>	any depending on the test	test files	none
/lhcb/validation/<whatever>	any depending on the test	workflow validation	none
/lhcb/certification/<whatever>	any depending on the test	files from certification productions	none
Failover and debug areas:			
/lhcb/failover/<LFN>	LHCb-Disk (T0D1)	temporary hop for files	none
/lhcb/debug/<LFN>	LHCb-Disk (T0D1)	Copy at CERN for debugging	none

- Key:
  - <year> - corresponding to year that RAW data was collected. For RAW data, it is the full year (yyyy) while for non-RAW data, the convention is <activity>yy (e.g. Collision11)
  - <MCSet> - For MC, the convention is to use "20xy" for the samples used to analyse the year 20xy. "Dev" or "/DEV" is used for exploratory work, and "Upgrade" for studies on the upgrade of LHCb.
  - <run no> - from DAQ
  - <prodID> - corresponds to the production pass that uniquely defines the configuration
  - <subID> - 4-digit number (0000, 0001...) used in order to limit the number of
  - <config> - indicates the DAQ partition ("LHCb" for full readout)
  - <runtype> - calibration, physics, ...
  - <stream> - name of the stripping output stream

    - ♦ <beams> - types of collisions other than p-p (labelled Collision), for example *Ionproton, Protonion, Protonhelium...*

- Space tokens deployment status at Tier1 sites
- T1 Storage System setup updated at September the 4th

-- PhilippeCharpentier

This topic: LHCb > NameSpace
Topic revision: r24 - 2017-02-08 - PhilippeCharpentier