

# LHCb Networking Requirements for Tier-2 Sites

---

Stefan Roiser

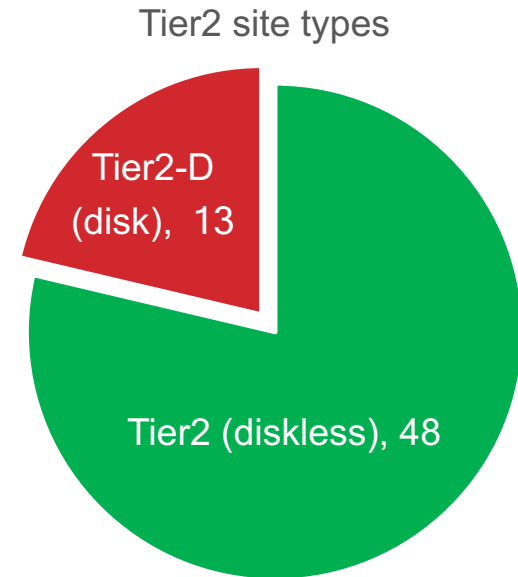
WLCG Workshop, Manchester

20 June '17



# LHCb Tier2 Sites

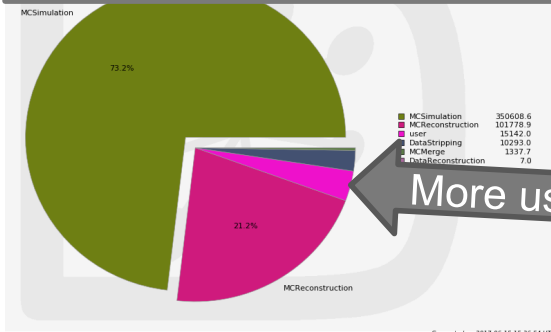
- LHCb uses 61 WLCG funded Tier2 sites
  - The majority of LHCb Tier2 sites do NOT provide disk storage to the experiment
  - 13 "Tier2-D" sites provide a minimum of 300 TB of storage for real and simulated data ready for user analysis (final physics format)
    - (Tier1 storages in addition also include "user", "buffer", "failover" and "tape cache" disk areas)



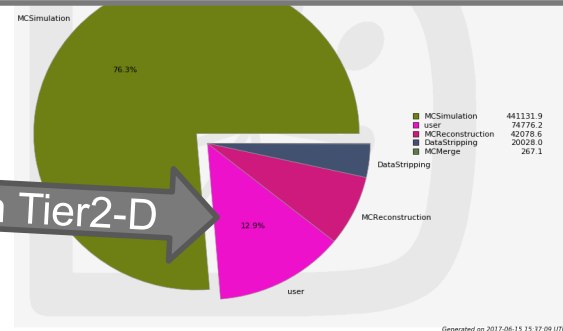
# Job Types at Tier2(-D) sites

	Tier2	Tier2-D
Monte Carlo Simulation	OK	OK
"Helper" for Data Processing	OK	OK
User Analysis	OK, if no input data ("toy MC")	OK
Working Group Productions	NO	OK

## Job Types at Tier2 Sites



## Job Types at Tier2-D Sites



More user analysis on Tier2-D

# Categories from the “Charge”

Category	Workflows being executed
“Full”	No, (user analysis on Tier2-D fit more into “disk rich” category)
“CPU rich”	No, intermediate caching on Tier2 sites does not exist for LHCb workflows
“Disk rich”	Yes, a small sub-set of Tier2 sites with disk storage run user <u>analysis jobs</u>
“Disk poor”	Yes, <u>working group productions</u> downloading local input data by protocol
“Disk less”	Yes, “default” site setup for LHCb Tier2 sites running <u>simulation</u> (no input data) and workflows downloading data from remote ( <u>helper data processing</u> , <u>mc reconstruction</u> )

- Majority of workflows are simulation jobs without input data
- Certain fraction are user jobs with local protocol access to input data
- Currently small usage of working group productions. Will increase

# Network Requirements

	Job Length	Input Data	Output Data
Monte Carlo Simulation	6 hours	None (start from random seed)	O(500MB), close T1
Monte Carlo Reconstruction	1 – 2 hours	Download O(5 GB), close T1	O(5GB), same T1
”Helper” for Data Processing	12 – 24 hrs	Download O(5-10 GB), close T1	O(5GB), same T1
User Analysis @ T2 (“toy MC”)	< 2 hours	None	Close T1 user area
User Analysis @ T2-D	< 2 hours	Protocol access, local storage	Close T1 user area
Working Group Production	< 3 hours	Protocol access, local storage	Close T1 data area

- In case of input data, always read by default from local site
- Output data always goes to different storage areas on T1 sites

# Special cases for network usage

- “Fast Stop”
  - Send signal to simulation job to stop processing after this event, e.g. useful for draining a rack/row/site in short time. Graceful stop happens usually within 1 – 5 minutes → Consequence all simulation jobs will upload their output data around the same time (more details in “optimization” session)
- “Gaudi federation”
  - Jobs are always sent to the location of one data replica. In case the local replica is not available the job will try to access the “next” replica from a remote site. According to workflow access can be protocol or download
- “Failover”
  - In case the destination storage site is not available for output data upload, the upload will be done to one of 7 “Tier1 failover storage areas”. Moving the file to the destination storage happens asynchronously from within Dirac