LHCb: Plans for CCRC'08

Nick Brook

Planned tasks

- Raw data distribution from pit → T0 centre
  - Use of rfcp into CASTOR from pit - T1D0
- Raw data distribution from T0 → T1 centres
  - Use of FTS - T1D0
- Recons of raw data at CERN & T1 centres
  - Production of rDST data - T1D0
  - Use of SRM 2.2
- Stripping of data at CERN & T1 centres
  - Input data: RAW & rDST - T1D0
  - Output data: DST - T1D1
  - Use SRM 2.2
- Distribution of DST data to all other centres
  - Use of FTS - T0D1 (except CERN T1D1)

All tasks envisaged during data taking in 2008

Useful Links

- LHCb planning meeting
  - http://indico.cern.ch/conferenceDisplay.py?confId=23406
- LHCB CCRC08 Twiki page
  - https://twiki.cern.ch/twiki/bin/view/LHCb/CCRC08

Planned tasks

- Preparation of quasi RAW data (hopefully) will occur before Christmas
  - Will use current generated MC data but need to merge datasets into ~1.8GB files
  - Available data means we have to re-use files 4 times
  - Each file will take ~24 hours to reconstruct
- February activities
  - Maintain equivalent of 2 weeks data taking
  - Assuming a 50% machine cycle efficiency

Activities across the sites

- Breakdown of processing activities (CPU needs)

Site	Fraction (%)
CERN	14
FZK	7
IN2P3	12
CNAF	8
NIKHEF/SARA	25
PIC	4
RAL	30

Will want to test conditions DB access & LFC service at sites

NB: No other production activities envisaged but user analysis will continue

February's Activities

- 42 TB of data from pit to CERN T0
  - Corresponding 23k files
- Same 23k RAW files from CERN to be distributed over T1 centres
- 14% of rDST production at CERN, remaining 86% at T1 centres (see table on earlier slide)
  - LHCb responsibility to ensure unique files are recons across CERN & T1 centres
  - Additional 23k (rDST) files produced (integrated across all sites) in proportion to figures in previous table
  - Corresponds to an additional 21 TB of data

February's Activities

- Stripping on rDST files
  - 8k DST files produced during the process (and stored on T1D1) - corresponds to 8TB of data
  - All files are distributed to other sites
    - 7x8k files
    - 7x8 TB
- Total number of jobs accessing the data
  - Recons: 23k
  - Stripping: ~8k

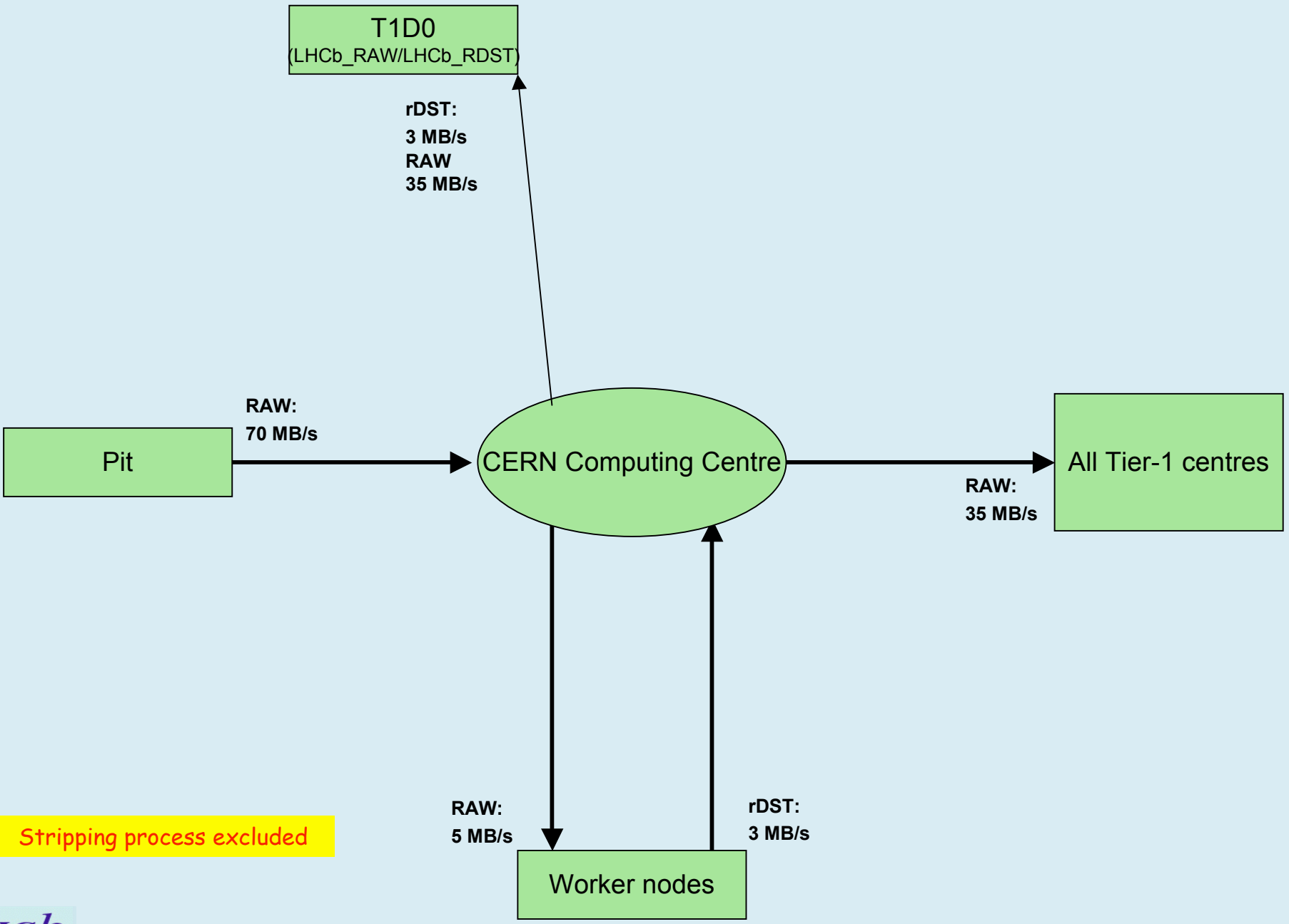
Nos of jobs/site

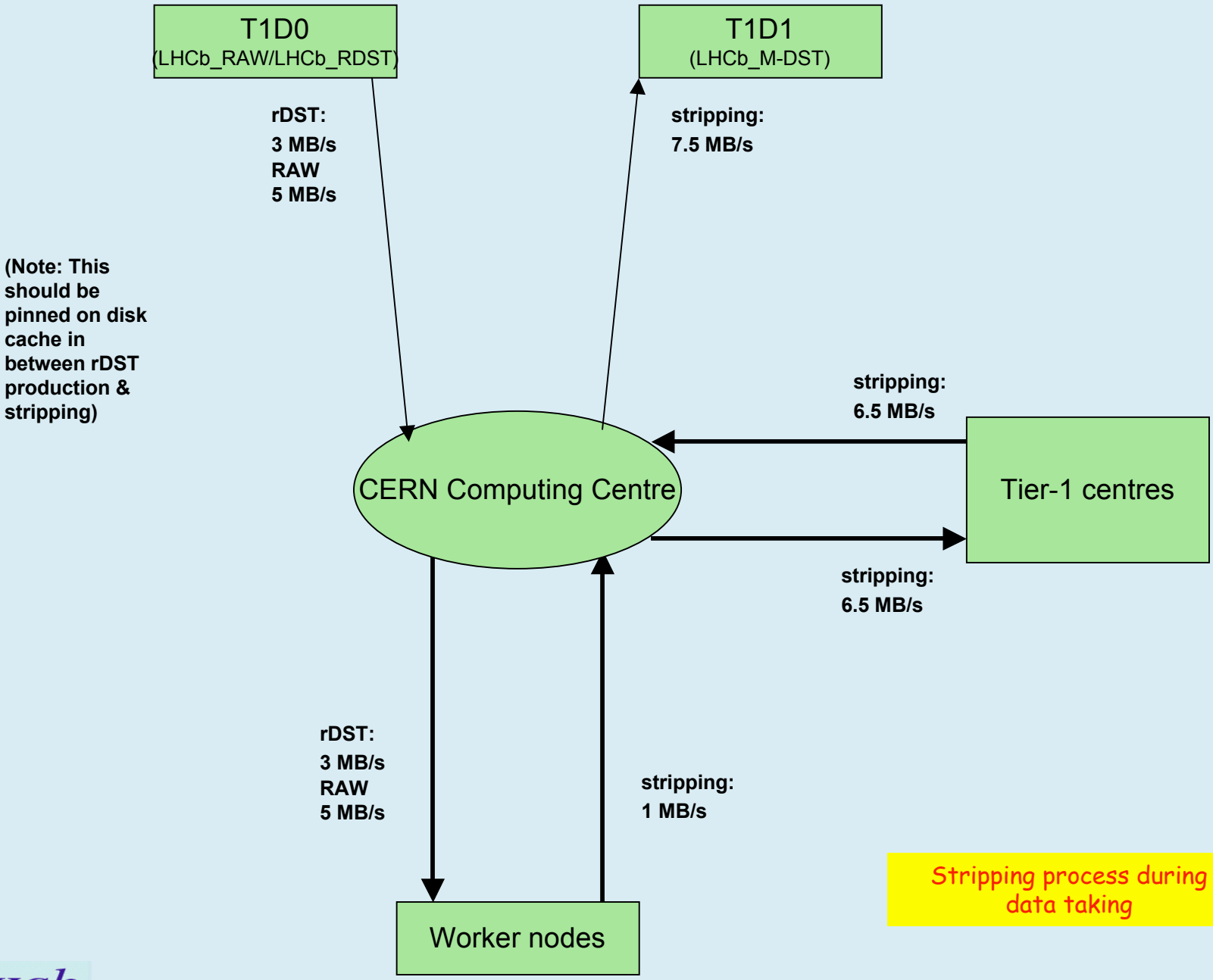
	Total Jobs			Simultaneous jobs		
	Recons	Strip	Total	Recons	Strip	Total
CERN	3300	400	3700	236	29	265
FZK	1700	200	1900	122	15	137
IN2P3	2700	300	3000	193	22	215
CNAF	1800	200	2000	129	15	144
NIKHEF	5700	700	6400	408	50	458
PIC	900	100	1000	65	8	73
RAL	6900	900	7800	493	65	558
Total	23000	3000	26000	1643	215	1858

Amount of data/site

	T0D1	T1D1	T1D0	Disk	Tape
CERN	0	8	45	8.0	53.0
FZK	7.4	0.6	5.2	8.0	5.8
IN2P3	7.1	0.9	8.1	8.0	9.1
CNAF	7.4	0.6	5.6	8.0	6.2
NIKHEF	6.0	2.0	17.2	8.0	19.2
PIC	7.7	0.3	2.8	8.0	3.2
RAL	5.6	2.4	21.0	8.0	23.4
Total	41.1	14.9	105.0	56.0	119.9

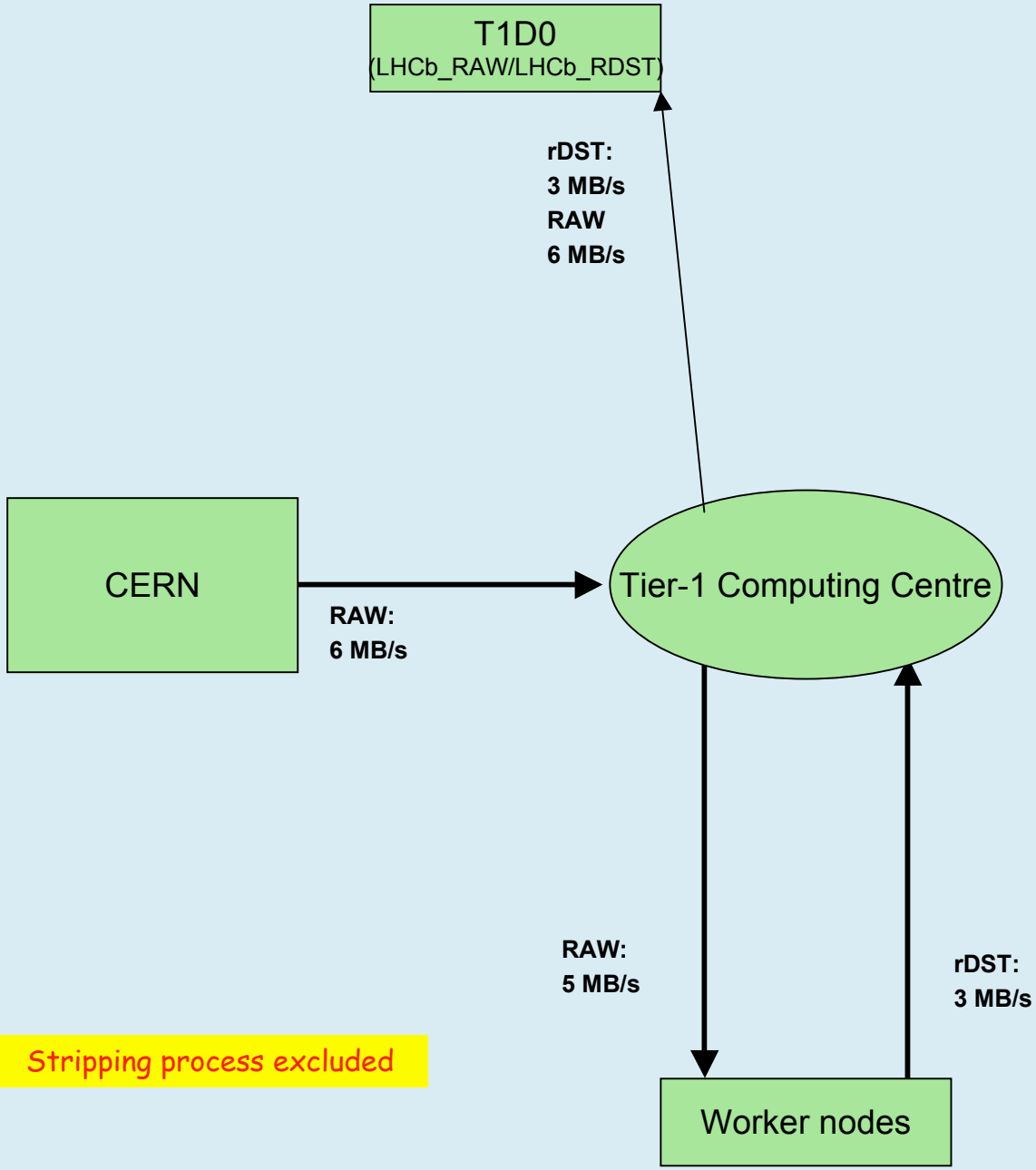
CERN CCRC08


T1D0
(LHCb_RAW/LHCb_RDST)
rDST:
3 MB/s
RAW
35 MB/s
RAW:
70 MB/s
Pit
CERN Computing Centre
All Tier-1 centres
RAW:
35 MB/s
RAW:
5 MB/s
rDST:
3 MB/s
Worker nodes
Stripping process excluded

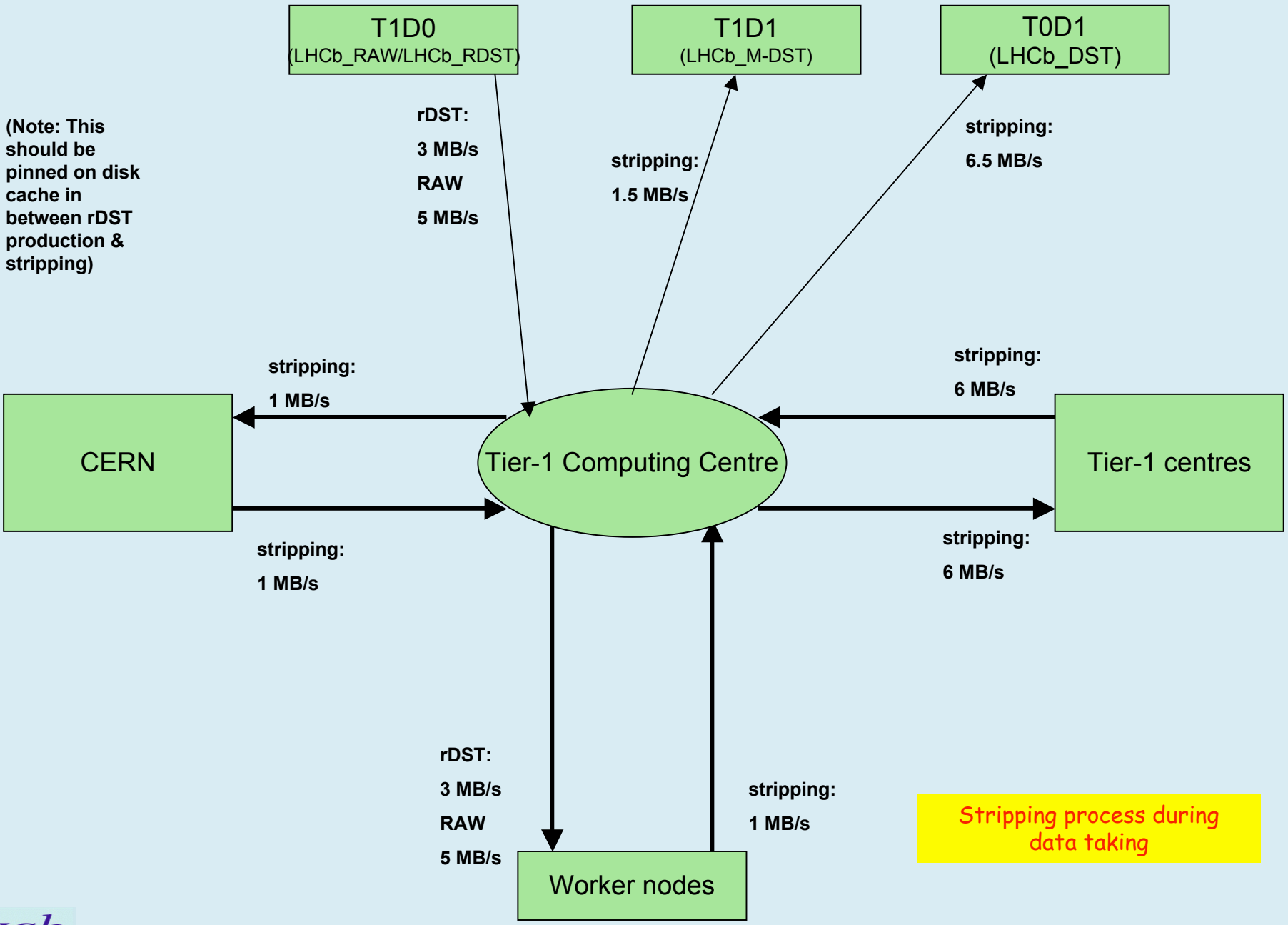


Stripping process during data taking

"Typical" Tier-1 CCRC08



Stripping process excluded


T1D0
(LHCb_RAW/LHCb_RDST)
T1D1
(LHCb_M-DST)
T0D1
(LHCb_DST)
(Note: This should be pinned on disk cache in between rDST production & stripping)
rDST:
3 MB/s
RAW
5 MB/s
stripping:
1.5 MB/s
stripping:
6.5 MB/s
stripping:
1 MB/s
stripping:
6 MB/s
CERN
Tier-1 Computing Centre
Tier-1 centres
stripping:
1 MB/s
stripping:
6 MB/s
rDST:
3 MB/s
RAW
5 MB/s
stripping:
1 MB/s
Worker nodes
Stripping process during data taking

Pit -> Tier 0

- Use of rfcp to copy data from pit to CASTOR
  - rfcp is the recommended approach from IT
  - Will have 25 TB capacity available online
  - A file sent every ~30 sec
  - Data remains on online disk until CASTOR migration
  - Rate to CASTOR - ~70MB/s
- Open questions:
  - How do we register a file in SRM SE with correct space token?
  - Can we have a guaranteed minimum migration to tape time?
  - How can we deal with error recovery with CASTOR 0 filesize problem?
  - Can CASTOR SRM publish the checksum value?

Tier 0 -> Tier 1

- FTS from CERN to Tier-1 centres
  - Transfer of RAW will only occur once data has migrated to tape & checksum is verified
    - Leads to 2 of the open questions already raised on previous slide
  - Rate out of CERN - ~35MB/s averaged over the period
  - Registration in LHCb bookkeeping & LFC occurs at same time as request to remove data from online disk is issued

Tier 1 <-> Tier 1

- FTS used to transfer data between from CERN/T1 to CERN/T1 of stripped files
  - LHCb mechanisms in place to do this but never been used in anger using FTS
    - Controlled using a central DIRAC services/site VO boxes
  - 1st Grid copy is stored locally
- Open question:
  - Not clear to us how we monitor actually usage corresponding to each space token
    - Currently we monitor this from entries in the LFC

Data Access

- Will need SRM 2.2 SE with correct space tokens
  - LHCb space tokens are:
    - LHCb_RAW (T1D0)
    - LHCb_RDST (T1D0)
    - LHCb_M-DST (T1D1)
    - LHCb_DST (T0D1)
    - LHCb_MC_M-DST (T1D1)
    - LHCb_MC_DST (T0D1)
    - LHCb_FAILOVER (T0D1) <- NEW!!!!
- Need access to lcg utils/GFAL on WN
  - Will be using lcg-gt & lcg-cp as a minimum as part of running applications

Databases

- Conditions DB at CERN & Tier-1 centres
  - For February will use static information replicated using "streaming" from CERN to Tier-1's
  - No plans to test replication of conditions DB Pit ⇔ Tier-0 (and beyond) during February
    - Standalone tests at a later date
- LFC
  - For February will use local T1 instance if available & tested (probably RAL, IN2P3 & CNAF)
  - Use "streaming" to populate the read-only instance at T1 from CERN
    - Programme of testing already under discussion with above sites

Service Reqts - Critical

- Max downtime 0.5 hour
- Critical services
  - VO boxes at CERN
    - Run central services - need availability
  - CERN LFC service
    - Central registration of all files
  - VOMS proxy service
    - Needed for any Grid activity, currently this includes running at CERN!

Service Reqts - Serious Disruption

- Max downtime 8 hours
- Serious disruption
  - T0 SE
    - No data gets transfer
    - Will have storage to cope with downtime of online system
  - (Individual) T1 VO boxes
    - Do have round robin system in place
  - SE access from WN/ CE access at site/Condtions DB access/WN misconfig
    - A large fraction of production will fail if data can't be processed at a particular site

Service Reqts - Serious Disruption

- Max downtime 8 hours
- Serious disruption
  - LHCb BK service
    - Separation of copy & registration process in place
  - No ORACLE streaming from CERN
    - Local LFC out of synch, round robin system in place
    - Real criticality here will be felt in May tests when condition DB will be updating
  - SAM service
    - Difficulties in assessing site availability for LHCb

Service Reqts

- Major reduction in effectiveness - max downtime 8 hours
  - LHCb RB
    - Using several RBs at CERN & T1 centres
- Reduced effectiveness - max downtime 24 hours
  - T1 LFC service
    - round robin system of other LFC'sin place
  - Dashboard
    - Difficulties in assessing job status from LCG point of view
    - LHCb have monitoring of jobs within DIRAC

Summary

- February dress rehearsal will test full chain
  - DAQ to T0 to T1
    - Data transfer & data access running concurrently
  - Current tests have tested individual components
  - Wish to test access to DB services at site, Conditions DB & LFC replicas