

Megatable Inputs

Dataflow

RAW data is reconstructed:

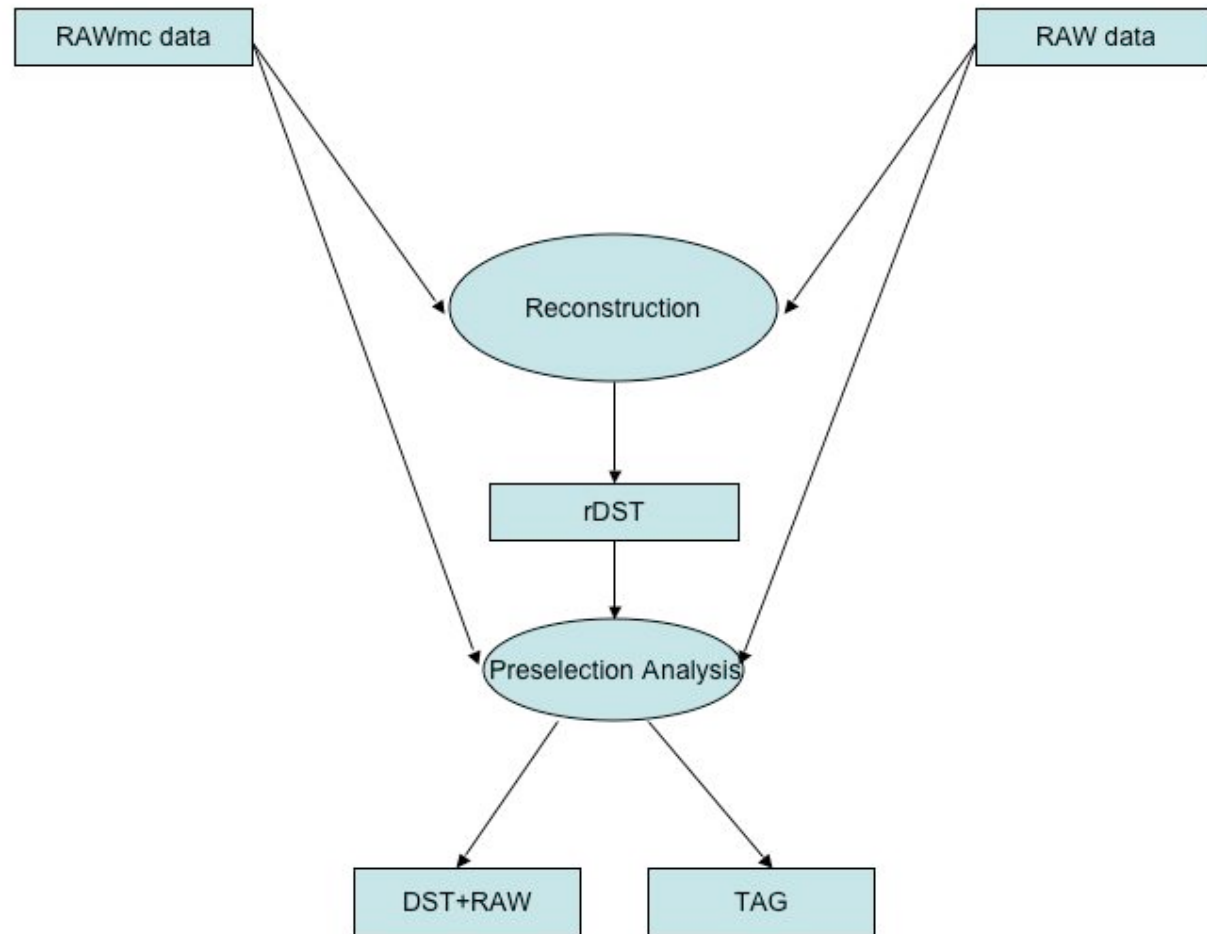
e.g.

Calo. Energy clusters

Particle ID

Track momentum

...

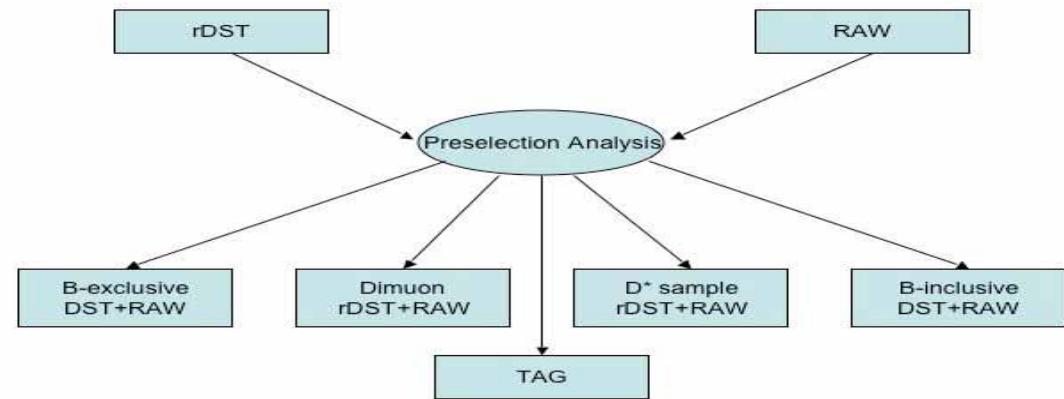


At recons time only enough info is stored to allow physics pre-selection to run at a later stage - reduced DST (rDST) - stored separately from RAW

RAW is distribute from CERN so an extra copy is distributed over all T1

During re-processing rDST produced at "pit" is partially shipped to T1's

Dataflow - Stripping



rDST is analysed in production-mode → event streams for further analysis; 4 high level categories (≡ min. 4 streams)

Algorithm developed by physics working groups - use as i/p rDST & RAW

Event to be output will have additional reconstructed info added: (full) DST+ RAW data

Event Tag Collection (ETC)- created to allow "quick" access to data; contain "metadata"

The DST & ETC are distributed to all T1 centres & T2 centres supporting analysis

Megatable - Nomenclature

There are 3 storage type envisaged within LCG

- Tape-1Disk0: pure custodial reqt with seldom access
- Tape1-Disk1: custodial reqt plus fast access needs
- Tape0-Disk1: custodial dataset elsewhere fast access needed

Network traffic is broken down into:

- Network traffic into the site from T0, T1 & T2
- Network traffic out of site to CERN (CAF), T1 & T2
- ... big discussion whether to use peak or average rates, for T2 both are quoted in following slide, otherwise peak

Megatable - Input Assumptions

Network traffic is "raw" rate - no efficiency factors included in data transfer rates

Disk storage usage has a 70% efficiency overhead (as used in the TDR) which is meant to cover associated overheads with storing data on disk NOT fluctuation in evts sizes, overlap of physics channels etc

T2 resources used in the calculations were based on the commitment given to the LCG management

Megatable - Input Assumptions

Country	CPU(kSI 2k)
France	252
Italy	550
Poland	111
Russia	1036
Spain	300
Switzerland	265
UK	796
Brazil	0

Assume a fraction of the resource in Russia (~10% of LHCb needs) & Manno (5% of needs) for analysis

Await Brazil to give commitment to LCG

Tier1-Tier2 Mapping

Not critical (in most part) for LHCb computing model:

CNAF: it

RAL: uk

PI C: es

Karlsruhe: (de), ch (excluding CERN), pl

CERN: ru

Lyon: fr, br + west of meridian line (undefined resource reqt)

NIKHEF: (nl)+ east of meridian line (undefined resource reqt)

Megatable - based on latest 2008 reqts

Including efficiencies:

Data on	RAW	rDST	Stripped	Simulation	Analysis
tape 2008 (TB)	560	320	483	128	-
Data on	RAW	rDST	Stripped	Simulation	Analysis
disk 2008 (TB)	76	43	775	375	114
CPU needs	Recons.	Stripping	Simulation	Analysis	
in 2008 (MSI2k.yr)	1.4	0.5	4.6	0.5	

LCG commitment still only ~70% simulation needs currently

Megatable - based on latest 2008 reqts

Estimated 4×10^6 secs physics (including machine efficiency)

Assume 8×10^9 events to Tier-0

	CPU (MSI 2k.yrs)	Disk (TB)	Tape (TB)
CERN	0.36 (40%)	350 (42%)	631 (46%)
Tier-1's	1.77 (40%)	1025 (42%)	860 (41%)
Tier-2's	4.55 (59%)	-	-

Rate out of CERN is averaged over 8×10^6 seconds i.e. 50% machine+expt efficiency

Megatable - Input Assumptions

For peak rates the following assumptions were made:

- Any network traffic during data taking is averaged over $8 \cdot 10^6$ seconds for peak rates
- Any network traffic during re-processing is averaged over 2 months for peak rates
- Any network traffic during re-stripping is averaged over 1 months for peak rates
- No catch-up assumed in peaks - purely amount of data that needs to be shipped over a fixed period
 - An aside - same arguments apply for calculating peak CPU as for peak networking

For average rates, assume the amount of data is moved over 6 months

Storage includes analysis (estimate analysis need on network are negligible)

GridKa

Average

Approximately factor of 2 between average & peak network traffic at all T1 centres

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	114.7	181.0	6.3	0.2	11.6	11.4	8.8	0.0	6.7	0.0	0	108	12	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	114.7	181.0	6.3	0.2	20.4	11.4	18.1	0.0	6.7	0.0	0	108	12	

IN2P3

Average

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	218.7	165.3	10.5	0.1	10.5	0.0	14.3	0.0	3.7	0.0	194	21	91	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	218.7	165.3	10.5	0.1	17.8	0.0	31.8	0.0	3.7	0.0	194	21	91	

CNAF

Average

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	119.8	138.3	6	0.3	11.9	0.0	8.5	0.0	7.8	0.0	101	11	78	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	119.8	138.3	6	0.3	20.7	0.0	17.1	0.0	7.8	0.0	101	11	78	

SARA-NIKEF

Average

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	184.0	157.1	9.1	0.0	10.0	0.0	11.6	0.0	0.0	0.0	166	18	92	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	184.0	157.1	9.1	0.0	18.2	0.0	26.8	0.0	0.0	0.0	166	18	92	

RAL

Average

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	131.2	140.3	6.3	0.4	12.5	0.0	9.7	0.0	11.2	0.0	108	12	75	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	131.2	140.3	6.3	0.4	20.8	0.0	18.6	0.0	11.2	0.0	108	12	75	

PIC

Average

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	61.3	122.1	3.5	0.2	11.6	0.0	4.3	0.0	4.3	0.0	51	6	75.2	

Peak

TOTALS			T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte			
	Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	61.3	122.1	3.5	0.2	21.7	0.0	8.6	0.0	4.3	0.0	51	6	75.2	

CERN (CAF)

Average

Note: factor of >2 between average & peak network traffic at CAF

TOTALS	T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte				
Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s aver.	MByte/s in	MByte/s aver.	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	287.9	321.3	-	0.5	8.5	11.4	19.8	0.0	13.9	0.0	166	108	103

Peak

TOTALS	T0=>T1	T2=>T1	T1=>T1	T1=>T2	T1<=T1	Storage for T2 TByte			Storage for T1 TByte				
Total Tape Tbyte	Eff. Disk Tbyte	MByte/s	MByte/s peak	MByte/s in	MByte/s peak	MByte/s out	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Tape1-Disk0	Tape1-Disk1	Tape0-Disk1	Cache-Disk
LHCb	287.9	321.3	-	0.5	21.1	11.4	51.1	0.0	13.9	0.0	166	108	103

LHCb view of CERN(CAF) facility:

- Same function as any other LHCb T1, with the exception of 1st pass reconstruction which occurs at T0
- Supports re-processing & stripping
- Supports user analysis & detector calibration/alignment
 - Anticipated to support 25% of the above activities as only centre with access to all RAW data
- Currently assume access to CERN (T0 & CAF) through Grid

T2 network traffic

		Pledge for 2008		TOTALS T2 Average		TOTALS T2 Peak	
Tier-2	Experiment	CPU kSI2k	Disk TByte	T2=>T1 MByte/s	T1=>T2 MByte/s	T2=>T1 MByte/s	T1=>T2 MByte/s
Tier-2s listed in the WLCG MoU							
FR/GRIF	LHCb	192	56	0.1	0.0	0.09	0
FR/LPC	LHCb	60	5	0.0	0.0	0.03	0
INFN	LHCb	550	1	0.3	0.0	0.25	0
Poland	LHCb	111	6	0.1	0.0	0.06	0
RU/RDIG	LHCb	1035	273	0.5	11.4	0.5	22.8
SP/LHCb	LHCb	300	1	0.1	0.0	0.14	0
CH/CSCS	LHCb	265	65	0.2	11.4	0.17	22.8
UK	LHCb	796	100	0.4	0.0	0.36	0
Tier-2 candidates presented to C-RRB							
Brazil	LHCb			0.0	0.0	0	0

In general T2-T1 network traffic for LHCb is negligible

Only data into Russian & Swiss T2 significant ~20 MB/s in bursts of ~month

Summary

A breakdown of T1-T2 mapping for LHCb given & consequences on T1 centres

- With the exception of GridKa CPU/storage ratio ~correct
- Latest request from LHCb to GridKa through their user board reflected these updated numbers
- T2 storage are ~10% of needs at T1

Still to do is disk cache needs

- LCG MB is in process of setting up a WG to look on how best to calculate this

Member of LHCb National Computing Board are aware of the planning behind megatable numbers

- local LHCb representative can be first point of contact