LHCb Computing Resource requirements 2009-10

Update 2.6, June 30th 2009

(N.Brook and Ph.Charpentier)

The new Computing Resource requirements of LHCb have been computed using the assumptions that are in the C-RRB letter. We would like hereby to clarify some points on request from the C-RSG as well as report some changes in the earlier assumptions and provide new estimates of requests.

History:

- v1.0: April 24th 2009. Additional information requested by C-RSG.
- v2.0 to 2.5: June 24th to 29th 2009. Add information on initial parameters, including some changes w.r.t. the original document.

Reminder: time is split into four phases:

- Phase 1 (early 2009): from 01.05.2009 until 30.09.2009, no data taking
- Phase 2 (late 2009): from 01.10.2009 to 31.12.2009, early runs of data taking (0.5 10^6 seconds)
- Phase 3.1 (early 2010): from 01.01.2010 to 31.03.2009, first physics run (1.2 10^6 seconds)
- Phase 3.2 (late 2010): from 01.04.2010, physics run (4.3 10^6 seconds)

Increase of CERN CPU resources

To ensure LHCb can react quickly to the early data in understanding the detector and producing early physics results it is felt necessary to increase the fraction of analysis at CERN. This will allow users to respond quickly to problems and avoid external debugging issues that could be associated with early Grid problems with first heavy use of the Grid with real data. In addition, using the Grid is often an overkill when dealing with small data samples (e.g. for detector studies or for testing and tuning physics selection algorithms). LHCb re-iterate that Grid computing is central to the analysis of data from the experiment. It is expected, and all necessary steps are taken in order to enforce that, that the fraction of analysis done at CERN will decrease down to the anticipated level of 25% (as stated in the Computing TDR). However we have used for the requirements for 2009-10 a decreasing percentage starting of 50% for 2009 and 40% for 2010.

Number of (re)-processing passes

LHCb will have no experience with real data for alignment and calibration prior to first beam, due to the low utility of cosmics events with vertical detectors. In addition the currently foreseen running scenarios indicate that despite LHC will run at low luminosity, the multiplicity of interactions per crossing in LHCb may

be higher than expected, and therefore imply additional tuning in the reconstruction algorithms. It is therefore anticipated that changes and tuning in the reconstruction code as well as in alignment and calibration constants will have to occur several times during the data taking, implying a re-processing of previous data. Despite this it will imply only reprocessing of a fraction of the whole period an unknown number of times, therefore we have assumed that the number of events that will be reconstructed is 3 times the number of events recorded (and not 4 as wrongly stated in the C-RRB document). Note that (re)-

Q: The old strategy (2008) contained 4 stripping passes per year which I roughly interpret as "2 per reconstruction pass" (or 4 strippings per initial data set). What is the new model? Does it imply 3 reconstrutions x 2 strippings = 6 strippings per initial data set (on average, of course)?

A: Here are the number of reconstruction passes and stripping passes (to be understood that in total *n* times the total number of events is processed / stripped):

	early 2009	late 2009	2010
Number of recons passes	10	3	3
Number of stripping passes	1	3	4

processing represents 10 to 15% of the overall resource needs at CERN and Tier1s, therefore a change on the number of reprocessing passes as a small effect on the overall needs.

Parameters used in the estimation

Now the first MC samples have been simulated for 2009 datasets (MC09), we have a better estimate of the events sizes for MC data, which is smaller than that used in the resource estimates earlier this year. This is the result of an extensive programme of work from the Core Software team to compact the event data using a persistent event model different from the transient one (untouched). However in the first estimates we had anticipated that a reduction in number of events would be feasible during the simulation job, before storing the data. As the main aim of MC09 is for understanding the data reduction (trigger and stripping) for various LHC running scenarios and various possibilities of physics, this reduction was not possible. There is therefore some compensation between a larger number of events stored and their smaller size.

The event sizes used in the estimation are the following[1]:

Real DATA RAW event size (kB)	35
Real DATA Reduced DST event size (kB)	35
Real DATA Full DST event size (kB)	50
Real DATA TAG event size (kB)	1
Monte Carlo RAW event size (kB) (b-phys)	250
Monte Carlo DST event size (kB) (b-phys)	50
Monte Carlo TAG event size (kB) (b-phys)	1
Monte Carlo RAW event size (kB) (non b-phys)	5
Monte Carlo DST event size (kB) (non b-phys)	25
Monte Carlo TAG event size (kB) (non b-phys)	1

[1] RAW here includes MC truth, this is why it is larger than for real data. The 5 kb size for RAW (non b-phys) comes from the fact that the RAW (100 kb) is kept only for 5% of the events.

Real DATA RAW was 35 kB in 2008 scrutiny, and 50 kB (2009) and 40 (2010) in our Apr'09 models – source of the latter two = presentation of N.Brook to MB on 31-March-2009 [1]

After discussion with Online experts they are confident with the 35 kB in 2009 going down to 30 kB in 2010. We had allowed for contingency earlier this year that seemed slightly too high, triggered by VeLo requesting art of the data not zero-suppressed. Of course only real data will tell us… Apologies for not putting the parameter for 2010.

Real DATA Reduced DST was 20 kB in 2008 scrutiny, and 35 (2009) and 25 (2010) in our Apr'09 models – source of the latter two = [1]

This is unchanged since April, again forgot to mention 25 kB for 2010. The full table is here:

	early 2009	late 2009	2010
Real DATA RAW event size (kB)	**35**	**35**	**30**
Real DATA Reduced DST event size (kB)	**35**	**35**	**25**

Real DATA Full DST was 75 kB in 2008 scrutiny (not 110 as cited in our Apr'09 report) and 100 in our Apr'09 models. However, implications on storage and tape are small as this is neither fully kept on disk nor on tape. Only influences final disk/tape requirement for stripped(RAW + DST) and disk for analysis.

Apologies again, there was a confusion here… To be clear: what is stored on disk for stripped data is DST+RAW, i.e. 85 kB for 2009 and 80 kB for 2010. In April there was a misunderstanding that these two would be additive, therefore the 100 kB in April were *in addition* to the RAW size. Also since then there was a considerable effort in reducing the size by packing the content on a limited number of bytes depending on the error on quantities.

A difference in the size of Monte Carlo RAW and DST b and non-b is a new feature of the model. Sizes always were 150 kB for RAW and 250 for DST for both cases, in the 2008 scrutiny as well as in our Apr'09 models.

Our mistake, it was indeed different but not explicitly said. Clearly b events are larger than non-b events. In addition as explained for non-b (minimum bias events) used for trigger optimisation, the MC truth is kept only for a small fraction of events (5%). Therefore the average "RAW" event size that contains this MC truth is much smaller.

Also the CPU required by the LHCb 2009 applications have been measured. This performance has been determined for an average number of interactions per bunch crossing of 1 (close to the LHCb nominal setting of 0.7). However a LHC running scenario that sees a lower number of bunches, with smaller "squeezing", may lead to higher values that will reflect on the processing time.

Brunel Real DATA Processing time (HEP06*s/evt)	**12**
Stripping Real DATA Processing time (HEP06*s/evt)	**0.8**
Gauss Monte Carlo Processing time (b phys)(HEP06*s/evt)	**360**
Gauss Monte Carlo Processing time (non-b phys)(HEP06*s/evt)	**100**
Boole Monte Carlo Processing time (b-phys) (HEP06*s/evt)	**4**
Boole Monte Carlo Processing time (non-b phys) (HEP06*s/evt)	**4**
Brunel Monte Carlo Processing time (HEP06*s/evt)	**12**

A difference in MC simulation time of b and non-b is a new feature of the model. This was 75 kSi2k*s/ev (300 HEP06*s/evt) for both cases, in the 2008 scrutiny as well as in our Apr'09 models.

No in April it was already like now, but maybe not explicitly repeated in all presentations.

The number of MC events to be generated in early 2009, late 2009 and 2010 respectively are:

	early 2009	late 2009	2010
MC b-phys Sample (evt)	**4.00E+08**	**1.00E+08**	**1.00E+09**
MC non-b phys Sample (evt)	**2.00E+09**	**1.00E+09**	**2.00E+09**

We used the information from [1] in our April'09 model: 600M b and 2000M non-b in 2009, and 1000M b and 2000M non-b in 2010.

Yes, these are the new numbers form the physics community… and it is due to change again probably. Note that for 2010, we have assumed that only ½ (1/5) of simulated b (non-b) events would be stored by applying loose trigger cuts just after digitisation (Boole), hoping that the first data allows that.

Until real data comes, it is foreseen to also use Tier1s for MC simulation and the following sharing has been assumed:

	early 2009	late 2009	2010
Monte Carlo Fraction at CERN	**10%**	**0%**	**0%**
Monte Carlo Fraction at Tier1s	**40%**	**0%**	**0%**
Monte Carlo Fraction at Tier2s	**50%**	**100%**	**100%**

As the RAW and RDST files are in T1D0 storage classes (i.e. not permanent on disk), we have foreseen that 10% of these data sets would be kept on disk for detector and performance studies; this is compatible with the proposal in the computing TDR.

Concerning analysis, some assumptions were made on the number of jobs that need to run on the whole available datasets (group analysis) and user analysis (small number of events or microDST analysis). The result is the number of jobs and average number of events per jobs reported in the following table:

	early 2009	late 2009	2010
Total number of jobs	3.7E+04	1.2E+04	4.9E+04
Average nos of events per job	3.8E+05	2.3E+06	2.3E+06

The CPU needed for analysis of real and simulated data is set to 1.2 HS06.s/event. It is also assumed (as observed) that a similar amount of CPU as for data analysis is needed for toy-MC and heavy fitting analysis that can be performed at Tier2s (at the 50% level) since they do not require data access.

Finally, the efficiencies used for CPU, disk and tape usage are those agreed within WLCG:

Scheduled CPU Efficiency	**75%**
Chaotic CPU Efficiency	**60%**
Tape Efficiency	**100%**
Disk Efficiency	**70%**

Updated LHCb requirements summary

The tables below compare the newly estimated LHCb requirements to those presented to the April C-RRB.

Date	Site	kHS06	Disk (TB)	Tape (TB)
Oct'09	CERN	17	780	1200
	Tier-1	31	2800	1300
	Tier-2	30	20	0
Apr'10	CERN	25	1470	1800
	Tier-1	45	4400	2100
	Tier-2	38	20	0
Oct'10	CERN	28	1470	2300
	Tier-1	49	4400	2900
	Tier-2	40	20	0

Table 1: CPU Power, disk and tape storage needed in place to meet LHCb requirements for the 6 month period commencing (a) October 2009, (b) April 2010 and (c) October 2010, as submitted to the April 2009 C-RRB

As suggested by the C-RSG, we have not included in the CPU requirements the 2 kHS06 at CERN and 1 kHS06 at Tier1s corresponding to the VOBOXes. However the 70 TB of disk storage corresponding to the cache for T1D0 service classes was included at CERN and each Tier1.

Date	Site	kHS06	Disk (TB)	Tape (TB)
Oct'09	CERN	15	720	1000
	Tier-1	31	1740	1000
	Tier-2	31	20	0
Apr'10	CERN	21	1290	1500
	Tier-1	41	3290	1800
	Tier-2	36	20	0
Oct'10	CERN	23	1290	1800
	Tier-1	44	3290	2400
	Tier-2	38	20	0

Table 2: CPU Power, disk and tape storage needed in place to meet LHCb requirements for the 6 month period commencing (a) October 2009, (b) April 2010 and (c) October 2010, as best estimate end of June 2009.

Q: Can you please provide an additional table in kSH06*year, analogous to Table 4 in the April-6 version of the document?

A: Here it is, but note 2 kHS06 at CERN and 1 kHS06 at Tier1s is removed (VOBOXes) compared to April. Note that for 2010, these numbers have essentially not changed if one takes this into account. Small changes in the power w.r.t. April come from a different (better?) estimate of profile.

kHEP06*year	2009	2010
CERN T0 + T1	8.54	17.19
Tier1s	11.70	32.99
Tier2s	17.12	31.74
Total	37.36	81.91

As can be seen, this re-evaluation of the LHCb Computing Resource needs comes to a slight to medium reduction compared to the estimates of April 2009. However we would like to outline the large uncertainty due to unknowns in the schedule of the machine as well as to the understanding of the detector.

Scaling of resource requirements

We have estimated by how much the resource requirements depend on the amount of running time. In this estimate, we consider that the amount of Monte-Carlo simulation is independent on this running time, as it is solely used for evaluating the sources of background and analysis biases. The actual determination of the corrections or systematic errors will be done using the real data (i.e. scaling with signal by definition). We estimate that about 50% of the analysis is on real data. A small fraction of analysis can also happen at non-Tier1 sites (toy-MC studies) as they do not require input data, and these are assumed to scale with running time.

The fraction of resources that scales with the amount of data taken is given in the following tables for CPU, disk and tape at CERN, Tier1s and Tier2s. Here "2009" means 1.7 10^6 seconds until April 2010 and "2010" means 4.3 10^6 seconds as of April 2010.

Table 3: Fraction of resources that scale with data taking time in 2009

Site	CERN	Tier1s	Tier2s
CPU	50%	30%	7%
Disk	35%	55%	
Tape	50%	45%	

Table 4: Fraction of resources that scale with data taking time in 2010

Site	CERN	Tier1s	Tier2s
CPU	65%	75%	10%
Disk	45%	70%	
Tape	75%	80%	

The main difference between 2009 and 2010 is for CPU at CERN and Tier1, which is explained by the fact that due to the low statistics of data collected in 2009, we have foreseen to use these sites also for MC simulation, which is no longer true in 2010.

The decrease in resource requirement is f(1-1/R) where f is the fraction quoted in the table above and R the reduction factor in real data (2 in the example below).

Example: twice less real data in 2009 (R=2) would reduce the CERN CPU requirements by 25%, the disk requirements by 17.5% and the tape requirements by 25%.

Previous data

Currently LHCb has collected very few real data (beam and cosmics) that are all kept at CERN (about 1TB). The requests presented to the C-RRB concern data being simulated or recorded in 2009 and 2010. It is understood that the existing MC data (so-called DC06) will be removed from storage when having been replaced by newer simulation. A cleaning campaign is already taking place for

reducing the amount of data kept on tape (mainly DC06 MC RAW and rDST that are no longer needed).