Computing Resources Scrutiny Group

C.Bozzi (Italy), C.Diaconu (France), D.Espriu (Spain, *Chairman*), J.Flynn (UK), D.Groep (The Netherlands), J.Knobloch (CERN), A.Lazzarini (USA), H.Marten (Germany), W.Trischuk (Canada), B.Vinter (Nordic Grid), H.Renshall (CERN/IT, *Scientific Secretary*)

Introduction

This report summarizes the deliberations of the Computing Resources Scrutiny Group (CRSG) established by the WLCG Memorandum of Understanding regarding the computing requests by the four LHC experiments for 2009 and 2010.

The scrutiny for 2009 and 2010 is now final. This scrutiny replaces the preliminary one contained in document CERN-RBB-2009-067.

The CRSG is an independent committee whose members are selected by the Funding Agencies represented in the C-RRB. The purpose of the CRSG is to inform the decisions of the Computing Resources Review Board (C-RRB) for the LHC experiments.

According to the WLCG MoU, every year the CRSG shall scrutinize

- The resource accounting figures for the preceding year
- The use the experiments made of these resources
- The overall request for resources for every experiment for the following year and forecasts for the subsequent two years
- The CRSG shall also examine the match between the refereed requests and the pledges from the institutions
- The CRSG shall make recommendations concerning apparent under-funding.

The starting point of the scrutiny is the resource request presented by the four experiments. The CRSG then enters into a sustained dialogue with each experiment seeking to understand to what extent the computing resource requests are well motivated, the usage made in the past of these resources and the accounting figures regarding usage and availability of the pledged resources.

The LHC schedule

The expected schedule of the LHC is an essential ingredient of the scrutiny even though it should be recognized that a fraction of the computing needs do not scale with running time.

According to the new schedule of the LHC restart following the 19 September 2008 incident, 2009 and 2010 (from October to October) will formally constitute a single run, with a short break at the end of the 2009 period. For the purposes of WLCG procurement and availability, it counts as two separate periods. One planned to extend from October 2009 to the end of March of 2010. We shall refer to this period as “2009”. The second period was planned to extend from April 1st 2010 to October 2010. We shall refer to this period as “2010”. The resources for 2009 were planned to be in place by September 1st 2009 and the resources for 2010 by April 1st 2010.

The most recent news following the circular by the CERN Director-General of July 15th indicates a further delay, with the LHC expected to be ready for beam injection in mid-

November. This leads us to recommend that the 2010 resources should be, exceptionally for this year, in place by June 1st 2010. This would give more time to the participating institutions to fulfil their pledges and represent some relevant savings. In practice this amounts to redefining the computing years '2009' and '2010', with the combined total running time unchanged.

The effective running time assumed is

Year	Pp	AA
	Beam time (seconds/year)	Beam time (seconds/year)
2009	1.7×10^6	0
2010	4.3×10^6	5×10^5 (*)

The AA estimate (*) corresponds to, at most, one month of data taking with an assumed efficiency x availability of around 20% (our estimate).

Note that the above numbers correspond to an effective beam time that differs considerably from the ones used in the 2008 report. Throughout the 2008 report to the RRB an efficiency of 50% was assumed, although already there it was mentioned there that it was probably more realistic to assume at most a 40% efficiency. In fact in the revised LHC schedule the efficiency factor is conservatively taken to begin at a modest 10% in 2009 and ramps up to 32% only at the end of 2010. This represents an important reduction in effective beam time and probably a more realistic estimate.

Usage of the resources

The WLCG project keeps an accounting of this at http://lcg.web.cern.ch/LCG/accounts.htm.

In 2008 for Tier1's and CERN the CPU usage was roughly at the 40% level of the pledges, whereas disk and tape are at the 60% level approximately. In the first two months of 2009 CPU usage was approximately 60% of the resources, disk at a 75% and massive storage at the 60% level. The figures showed a clear improvement in the usage of the GRID resources. Increasing numbers of users are submitting jobs as the middleware is becoming progressively more mature and the commissioning of the LHC for physics runs approaches.

The situation appears substantially unchanged after the STEP09 exercise in June. The statistics show that during the first six months of 2009 usage of the installed resources has been: 56% in CPU, 75% in disk and 62% in tape for CERN and Tier 1's. As of June 09 the installed CPU and Disk resources actually amounted to 67% and 83% of the pledges, respectively. Tape appears to have actually surpassed the pledge.

ATLAS and CMS have achieved an average 80% success rate for job submission during the recent STEP09 challenge. A few underperforming Tier 1's (some of them quite important) have been identified during the STEP09 exercise, while the performing ones well exceeded 90% efficiency. The CRSG considers it a worrisome signal that some centres appear to have difficulties in passing this type of test and encourages the respective Funding Agencies to follow the evolution of their performance closely.

For Tier 2's the level of usage is generally better but exhibits extremely large variations among the centers.

The rhythm of fulfilment of pledges has naturally slowed down as the date for 2009 was moved from April 1st to September 1st along with a general perception that those that are installed or planned to be installed by September 1st are already enough to deal with the 2009 data after the reduction in beam time. This perception is roughly correct as for instance the resources in place at the Tier 2's already surpass for some experiments the ones requested after the last revision.

In conclusion, the level of usage of the resources by the experiments of the CERN and Tier1 resources made available to them does not seem totally optimal and this is a cause of concern to some extent. It is also clear that the current pledge for year 2008 represents an overcommitment of resources, mostly due to the delay in LHC commissioning. While the long awaited first year of LHC operations is certainly not the right time to cut back substantially on computing resources, the CRSG deems it appropriate to recommend an adjustment based on the running conditions foreseen at the moment.

A more agile mechanism of scrutiny and better timing of the C-RRB meetings would surely help to better match between pledges and needs, thus optimizing expenditure. We recommend the C-RRB to take appropriate actions in this direction.

In preparation for the April 2010 C-RRB meeting the CRSG in fulfilment of its mandate is asking the experimental collaborations to provide a summary of the use of the computing resources made available to them. This report should be available by March 1st 2010.

Interactions with the experiments

Following the Chamonix review, the initial requests from the experiments for 2009 and 2010 were made public on April 7th. These requests were, roughly speaking, at the level of the foreseen pledges last year, in spite of the ensuing LHC delay.

The CRSG submitted a preliminary assessment to the C-RRB on April 26th.

In this preliminary scrutiny it was stated that experiments compared their requests for 2009 and 2010 to the current pledges, while in view of the expected beam time it would perhaps be more meaningful to compare 2009 to 2008, and 2010 to 2009. In order to substantiate requests at the level of the pre-existing 2009 and 2010 pledges even if the actual number of events to process will be roughly half of those considered for 2008 and 2009, respectively, the experiments needed to make sizeable modifications to the computing models, some of which were modified in ways that the CRSG did not expect.

We concluded that these modifications represented changes that, if agreed, could make some of the computing models unsustainable. The CRSG believes that the existing computing models have largely proven their validity and has no doubt that they will survive their first contact with real data in 2009. The CRSG stated that there was no reason for radical changes at this point without hard experimental evidence.

The CRSG also expressed its concern about the short time it was allowed to develop its scrutiny and the paucity of information provided to the scrutiny group by some experiments, which in fact prevented a complete scrutiny of the ATLAS request.

ATLAS and CMS challenged the preliminary report on the basis that more resources were needed in view of the special circumstances of the 2009+2010 run and expressed concern about all experiments receiving an equal treatment. They also emphasized that the pledged resources are never 100% available, implying the need for a safety margin.

The C-RRB instructed the CRSG and the experiments to finalize our 2009 and 2010 scrutinies in July and September, respectively, seeking convergence.

The CRSG recognizes the special characteristics of this combined run, extending over two years, and the uncertainties that remain in the current computing models and the need for redundancy/contingency. The situation for 2009 and 2010 should indeed be regarded as somewhat exceptional, and to a large extent driven by the transient character of the start-up months. On the other hand, the CRSG believes that

- The delay in the LHC commissioning should be visible to the funding agencies.

- Resources should be driven by needs and not vice-versa.

A new common scrutiny procedure was set up for ATLAS and CMS following a May 11th 2009 letter by the CRSG chair to the ATLAS and CMS management. This procedure would ensure that a coherent set of principles is applied. In a letter dated May 19th the management of ATLAS and CMS agreed to the new scrutiny procedure.

A joint meeting with ATLAS and CMS took place on June 19th, where the experiments submitted a revised request. This revised request is the basis of the present scrutiny.

A number of interactions with ATLAS and CMS followed. Some questions were referred to the LHCC for advice.

Meanwhile, the scrutiny for LHCb and ALICE continued separately. Our referees had a number of meetings with their representatives seeking full understanding of their computing requests and, in the case of LHCb, of the merit of some modifications to their computing model. Some issues were also referred to the LHCC.

As for the non-availability of the full committed resources the CRSG wishes to emphasize again that their recommendations to the C-RRB are, within reasonable uncertainties, well adjusted to the real needs and, in particular, that there is no contingency for late delivery or failure to meet the pledges included in our estimates.

Interactions with the LHCC

In carrying out the scrutiny of the experiments requests the scope of this group is largely limited to the implementation of the respective computing models which are periodically reviewed by the LHCC.

The evolution of the commissioning of the experiments as well as the implementation of the computing models in successive tests along with a better understanding of their needs have motivated a number of changes, sometimes representing limitations in the original model or assumptions.

Obviously there is a gray area where the respective competences of the LHCC and the CRSG overlap and it is not so clear what would represent a change of the computing model or what would just be a natural adaptation of it to the changing circumstances.

During the 2009 scrutiny, the CRSG felt that a number of issues required consultation with the LHCC. A first set of questions was discussed in an open meeting with the LHCC on February 16th 2009. A summary of the items submitted to the LHCC is:

- The event size has a very direct impact on the computing requirements. Some experiments made an effort to reduce the raw event size (and the size of all subsequent derived formats) by establishing a reduction profile after startup. We believe that this effort should be followed by all without unduly jeopardizing the physics.
- The potential proliferation of different data formats serving the same purposes should be watched closely as it seems a matter or concern. Strong overlap between different data sets as claimed by CMS now is a new matter of concern.
- More than one collaboration seems now to place heavier demands on CERN resources or suggest that a larger than normal part of their analysis should be done at CERN. This is hardly compatible with the worldwide LCG effort.
- Cosmic data taking is now much more emphasized by experiments; while it is clear that cosmics are extremely useful in commissioning for calibration, this data is by nature transient and it seems somewhat questionable to us to support substantial requests based on cosmic runs.

- The data analysis strategies should be reviewed by the LHCC very soon, in order to ensure both, a reliable start-up and a coherent long term strategy.
- A clearly defined sharing among the experiments of the resources installed at CERN should be considered, in particular for the analysis facilities, such that an optimal usage of these resources is achieved.

A second set of specific questions was referred to the LHCC in June 2009. These questions were reviewed in an open meeting with the experiments to which the CRSG chair was invited on July 6th.

- In the revised requests, some collaborations demand much more MC simulation. Which is the minimum amount of simulated data (and of which type) that does not jeopardize the physics program of the different experiments?
- Which is the scientific value of the initial high trigger rate for ALICE?
- CMS (trigger rate 300Hz) plans to reprocess at 380 Hz peak performance in 2009 and sustained rate 400Hz in 2010. ATLAS (trigger rate 200Hz) plans to reprocess at 400-500 Hz sustained rate throughout the run. These figures were not apparently foreseen in the TDRs. Are they necessary? Note that this may place constraints on the disk/tape rate and has a large impact on cost.
- Recent performance data from CMS suggest that by pre-staging tape one can get access for reprocessing at a rate that is sufficient. Should this procedure (which potentially represents large savings) be recommended to other collaborations?
- Is the 40% overlap between primary datasets in 2009 in CMS a realistic possibility? Why is it so different between ATLAS and CMS?
- Different collaborations estimate the impact of pile-up due to the foreseen running conditions in 2009/10 differently.

The CRSG was provided on July 22nd by the LHCC with a set of draft recommendations, addressing the previous issues, which have been taken into account in the present scrutiny. The CRSG is satisfied with the recommendations and has no further pending questions at the moment. We recommend that the present scrutiny is read in conjunction with the relevant LHCC documents.

Over the last months a fluid dialogue has been established between the CRSG and the LHCC, but both committees feel that more complete and earlier interactions are necessary. These should also include joint meetings with the experiments to avoid duplicate reviews.

It should however be borne in mind that the nature of the scrutiny carried out by the CRSG is different to the one implemented by the LHCC reviews. For the scrutiny group, acting on behalf of the Funding Agencies, the budgetary cost of the resources made available to the experiments and their efficient use are relevant issues.

On the scrutiny process

The CRSG is now satisfied with the quality and quantity of the information provided by the experimental collaborations during the second iteration of the 2009 and 2010 scrutiny.

The CRSG did not have on this occasion full access to the experiments internal spreadsheets, but was able to produce their own that are able to reproduce the experiments requests given the current computing model and with the assumed values of their parameters. In the process a number of discrepancies were found, some were due to simple errors, while on several occasions they revealed assumptions or parameters not sufficiently documented.

Adapting to the changing running conditions and modifications of the models/parameters has also proved to be quite challenging.

For future reviews we insist that

- Requirements and respective models be frozen during the review
- All changes to the models compared to the previous review be well documented
- All documents be provided sufficiently early to allow time for the review, a deadline for the revised requirements should be agreed upon well ahead the final report deadline. For the upcoming April C-RRB meeting this deadline is March 1st 2010.

The CRSG also plans to set up in advance a meeting calendar and will propose to hold joint review sessions with the experiments and the LHCC twice a year. We expect these sessions to take place in February and July.

It is clear that the current implementation of the WLCG MoU is not agile enough to respond to the changing environment we have been facing during the LHC commissioning. This has had a non-negligible cost for the funding agencies. The CRSG is committed to make recommendations to the C-RRB with the aim of optimizing the resources already invested and adjusting the future ones to the required computing needs.

General comments

Even if some experiments produced estimates for 2011 and beyond, we have not considered at all any projections beyond 2010. For a proper scrutiny we have yet to see real collisions and real data with the computing models going through a reality check and the CRSG prefers not to commit itself to any specific forecast beyond 2010. Experience gained once real data-taking is underway should reduce remaining uncertainties considerably allowing better estimates for 2011 and beyond.

It seems prudent to scrutinise the experiments' use of resources after the first months of data taking in 2009. The CRSG commits itself to provide such a report in the shortest delay which is feasible, hopefully for the April 2010 C-RRB. In order for this to happen, the CRSG will require timely resource usage reports from all the experiments no later than 1 March 2010.

The WLCG represents a computing effort of an unprecedented scale. We recommend that the different collaborations undertake a proper risk analysis and take stock of their results in future requests in order to cope with the most likely failures or shortfalls.

- The experiments should incorporate the running conditions into their models in a uniform way. At present there is still a discrepancy in the way that ATLAS and CMS treat the consequences of pile-up due to the foreseen long bunch crossings.
- The experiments are asked to actively pursue the policy of reducing the size of their raw events, and other derived formats, in future years as much as possible as detectors become better understood. No additional progress has been found along these lines in the present requests.
- We welcome efforts to remove `dark' or `orphaned' data and encourage the experiments to pursue this vigorously. Likewise we commend them to pursue a vigorous programme of purging data that is no longer required.
- We recommend the experiments make maximal use of the distributed resources in the GRID, thereby avoiding as much as possible the use of CERN facilities.
- In the case of CERN resources, we advocate for a very clear separation between the contributions used for calibration and first pass reconstruction and central analysis

(‘express stream’ or similar), and those used to perform physics analysis by the CERN based physicists.

- As far as data distribution is concerned, we notice different strategies which try to optimize the total CPU power required to analyse data, ranging from the maximization of replicas among the computing centers to the usage of streamlined primary datasets. Each approach has its own advantage and disadvantages. We recommend that the experiments exploit the upcoming data taking period in order to determine which strategy optimizes physics output while keeping resource requirements at a reasonable level in the understanding that some of the tenets of the original computing models do not appear to be sustainable anymore.

- The emphasis on non-beam data and the corresponding resources should become less and less important during the upcoming data taking period, as detector performance are better understood

- Any shortage of CERN resources implies potential disruption of data taking; therefore, we advocate for a full support of CERN resources. Resources at Tier1s and Tier2s are crucial for physics output; any shortage of non-CERN (non-custodial) resources does not impact data samples, but slows down the physics productivity of the experiments.

The CRSG wishes to state that the recommendations contained in this scrutiny are to the best of our knowledge rigorous. They correspond to the real needs of the experiments for a given LHC live time in the present stage of the commissioning and of their computing model implementation. There is no contingency for late delivery or failure to meet the pledges included in our estimates or for less than 100% availability of these resources.

Scrutiny of the ALICE request for 2009 and 2010

Overview

ALICE presented its updated 2009 and 2010 requests to the LCG management board in March and April 2009 and made a presentation to LHCC in May 2009 on responding to shortfalls in resources. We received the final 2009 and 2010 requests in September 2009. The experiment has not changed the parameters (event sizes and event processing requirements) in its computing model. Changes in resources requested stem largely from changes to the LHC operation schedule, but new features in the model are the addition of a large disk buffer at CERN and some immediate processing using the CAF during pp and AA running.

Modifications

In the ALICE computing model pp events are recorded at 100Hz in the LHC steady state, leading to 10^9 pp events recorded annually. The collaboration stresses that the initial running is “prime time” for them. To maximise the physics impact of the early data they plan to use a higher event rate in the 2009 and 2010 pp runs, aiming to record 1.5×10^9 events in total, 50% more than in a standard data-taking year. For our scrutiny we have split this fixed number using the ratio of running times to give 0.43×10^9 events from 1.7×10^6 s in 2009 and 1.07×10^9 events from 4.3×10^6 s in 2010. ALICE says that the 50% increase in pp data compared to a standard year leads to a 5% or less increase in resources requested. Our calculations agree with this. We note that the increased data collection rate has been endorsed by the LHCC.

ALICE assumes a 1-month heavy-ion run in October 2010 leading to a standard year’s worth of events (10^8). We follow this for our scrutiny.

Note that ALICE treats the 2009 and 2010 running periods as a single long run and hence the transient (disk) storage requirements for 2009 and 2010 are cumulative, and for 2010 correspond to 1.5×10^9 pp events and 10^8 AA events.

We confirmed that ALICE has only one data stream so there is no issue of increased storage needs arising from stream overlap.

ALICE now has a disk buffer of order 1.5PB in front of the mass storage system at CERN which can hold a standard year's worth of pp data. This allows data to be checked and possibly discarded before going to mass storage. A new feature of the computing model is to perform some immediate reconstruction and analysis on the CAF during pp running using data from the buffer: this will run a few hours ahead of and guide the main Tier0 reconstruction. Similarly, fast analysis of a fraction (1/4) of the AA events is foreseen using the CAF during AA data-taking[1]. The collaboration intends to run the detector up to its maximum rate for pp (up to around 500 Hz for the TPC and up to around 1 kHz for other detector elements) and to saturate the data bandwidth for AA running. An added benefit is that most analysis will be able to use faster disk-based access to recent data.

The ALICE software can automatically purge data that is no longer required. ALICE has not suffered from dark or orphaned data to date and expects that their bookkeeping systems should not let such data accumulate. PROOF-enabled facilities have been established at two Tier 2 sites in addition to the CAF, although the CAF is the only one accessible to all members of ALICE.

The detector has additional elements in place for 2009 startup compared to 2008 (increased coverage of the TRD and EMCAL). This leads to small fractional increases in event size (the TPC dominates the output). It could be offset by implementing ROOT compression, which is expected to give a saving of 10% or so (the compression would increase CPU use in the online system, but that is outside the scope of CRSG scrutiny). In any case, ALICE has stuck to its TDR event size for estimating resources and says it is very close to this from above.

Data sizes and processing requirements per event (real and simulated) have not been changed. ALICE does not assume additional reprocessing runs in their calculations for 2009 and 2010: their figure of three reconstructions encompasses prompt reconstruction, many short passes and a final complete reprocessing. CPU capacities have been converted to HEPSpec06 (HS06) units using 1 kSI2k = 4 HS06. Note that for ALICE 2009 and 2010 refer to calendar years, whereas we use these to denote the 2009 and 2010 "resource periods" Oct 2009 to Mar 2010 and Apr 2010 to Mar 2011 respectively.

We strongly encourage ALICE to adopt for their computing requests the common usage of "computing years" extending from 1st April to 31st March.

Real data (including cosmics)

The experiment's expectations for the amount of collision data recorded were noted above. Based on previous experience, ALICE expects to retain up to 400TB of cosmic events, with only one copy kept in permanent storage. Processing of this data will probably not continue once collision data-taking starts: if it does, it will be for calibration/alignment purposes.

Simulated data

ALICE aims to simulate both pp and AA data corresponding to a standard year's worth of events in each of 2009 and 2010 (calendar years). For their mass storage request, they count one standard year's worth of pre-existing MC data in custodial storage at the start of October 2009.

---

[1] Events will be available for analysis since some AA reconstruction will happen at T0 during data-taking even though the bulk is performed afterwards.

“Living with less”

ALICE continues to recruit new partners to provide additional computing resources. The experiment requests the resources it believes are required but in response to a request from the LHCC referees in May 2009, ALICE has proposed ways to accommodate reduced CPU and disk storage resources. The experiment says it assumes that sufficient mass storage will be provided: “It would not make sense to stop the experiment while beam is on.”

- ALICE notes that resources are typically made available in April, whilst their biggest spike in processing needs comes in October with the reconstruction of AA data at T0. Thus they suggest a re-profiling of provision and have broken down their resources requests to quarter years.
- “Plan B” encompasses a reduction in disk storage, decreasing the “replication factor” for ESDs on disk (from 3 to 2.5) and reducing the fraction of raw data on disk at each T1 from 10% to 5%. Scheduled analysis trains could run through data sequentially, staging from tape as needed, but chaotic analysis is less predictable and so will be more affected with consequent loss in efficiency. Calibration/alignment and reconstruction checks will also be affected. The impact of this change is much worse if the Grid fails to allow effective load balancing and data transfer between T1s and T2s.
- “Plan C” considers a reduction in AA MC simulation which will affect systematic errors. ALICE quotes an example of their charm measurement. Part of the error in acceptance calculations stems from MC statistics. Reducing the number of events by 1/2 (or 1/3) increases the systematic error from 16.3% to 17.7% (or 17.0%).

Summary tables

In the summary tables below we have compared our estimates (CRSG) of the resource requirements with the ALICE requests made in April 2009. Our calculations are independent of the experiment and based on our understanding of the ALICE computing model and the scheduling of data recording, reconstruction and analysis. The summary tables should be read in conjunction with the comments in the text.

For CPU capacities, we made estimates based on a “minimal” processing scenario with scheduled and chaotic analysis of reconstructed collision and MC data. In addition we made estimates assuming additional analysis (using an ‘analysis multiplier’ of 1.5) is needed in the early years while algorithms are being checked and debugged (this additional analysis is also a surrogate for additional reconstruction processing needed during early running). Accepting that some additional processing is needed, we adopt the larger estimates.

Our preliminary estimate for CPU capacity at CERN for 2009 was based solely on the T0 requirement to keep up with LHC running when reconstructing pp data together with on-the-spot calibration and alignment activity. However the request includes CAF capacity for prompt (‘fast’) reconstruction and analysis to shadow LHC runs of several hours’ duration and help decide which data from the disk buffer should be saved to mass storage. This is now included in our estimates.

In steady-state years the large AA initial reconstruction requirement essentially fixes the request for T0 processing capacity, and during pp running the extra capacity is reassigned as T1 resource. There is no AA running in the 2009 resource period, but the experiment still requires T1 capacity at CERN and assumes this for its requests. Without this, the 2009 CPU capacity at CERN would be small.

It should also be noted that we split the number of pp events for the 2009 and 2010 runs in the ratio 2:5 reflecting anticipated LHC duty cycles, whereas the experiment assumes a constant pp data-taking rate throughout the long pp run. Using an equal split and no ‘analysis multiplier’ produces a CRSG CPU estimate in close agreement (in total) with that using our split and the multiplier. The equal-split values are used for the 2009 CPU in the tables.

The difference in assumptions for the split of pp events between 2009 and 2010 has very little effect on storage requirements given the amount of pre-existing data and the generation of MC data during the run. For the 2010 run AA computing needs dominate and again the different split of pp events has little influence

Our disk estimates agree quite well with the requests, being slightly smaller overall. For mass storage our estimates are substantially larger for T1s external to CERN. An explanation may lie in the mismatch between an experimental request for resources sufficient to reach the end of the calendar years 2010 and 2010 compared to our estimates for the running periods ending in March 2010 and March 2011. We queried ALICE about this and they provided estimates for resources to reach the ends of the first quarters of 2010 and 2011. For mass storage the (CERN,T1ext) values in PB are (3.4,10.2) for 2010Q1 and (8.1,21.0) for 2011Q1, in much better agreement with our estimates (last row of the tables below). We encourage ALICE to adopt the standard practice in this regard.

Significant discrepancies between the ALICE request and our scrutiny, or numbers we have been unable to fully scrutinize will be denoted in the tables below using ***<u>bold underlined italic font.</u>***

2009

		CERN	T1ext	T2ext	Total
CPU/kHS06	Request	42.4	42.8	36.0	**121.2**
	CRSG	39	39	33	**111**
Disk/PB	Request	3.6	4.5	6.7	**14.8**
	CRSG	3.1	3.7	6.7	**13.5**
MS/PB	Request	2.4	***<u>7.2</u>***		***<u>9.6</u>***
	CRSG	3.0	***<u>12.7</u>***		***<u>15.8</u>***

2010

		CERN	T1ext	T2ext	Total
CPU/kHS06	Request	46.8	57.6	89.6	**194.0**
	CRSG	50	61	95	**206**
Disk/PB	Request	5.5	10.8	12.6	**28.9**
	CRSG	5.7	12.3	10.0	**27.9**
MS/PB	Request	6.3	***<u>16.3</u>***		***<u>22.6</u>***
	CRSG	7.2	***<u>24.1</u>***		***<u>31.1</u>***

Comments and recommendations

In the ALICE computing model, data and processing are nominally assigned to different tiers, but in practice data and tasks can be moved between tiers (more easily up, T2 to T1 or T1 to T0, than down). The experiment has used this freedom to do some balancing of the requests between CERN and external T1s and T2s. For our estimates we too made a nominal distribution of data and tasks between tiers. Without subsequent redistribution, we expected our total resource estimates to agree better with ALICE than our distribution between tiers, which was indeed the case.

The computing representatives commented that while the global total of guaranteed resources is most important, the experiment depends on the requested resources at CERN and a distribution between T1 and T2 reflecting the reality of provision. We note that other LHC experiments have heavy demands on CERN resources. Thus we are willing to accept a redistribution compared to our 'raw' numbers. This has most impact on the CPU capacities and we have therefore redistributed our CPU values in the same ratios as for the experiment's CPU requests, since those ratios are our best information on the distribution assumed by the experiment. For the 2009 CPU requirements we obtained values (12.3, 24.1, 74.3) kHS06 for (CERN, T1ext, T2ext) before redistribution, becoming (39, 39, 33) kHS06 afterwards. Likewise for 2010 CPU, values (38, 85, 83) became (50, 61, 95) after redistribution.

Our final estimates are in reasonable agreement with the ALICE requests (with the exception of T1 mass storage, noted above).

We queried the computing model's reliance on the dependability of the CERN T0 and CAF. ALICE agreed that failure of the CAF could lead to them storing data 'blind' and thus less effectively. However, carrying out prompt analysis at external sites would be less efficient because of network latencies for data transfer.

We note ALICE's recent purge of data and encourage the experiment to maintain this good practice.

Scrutiny of ATLAS requests for 2009 and 2010

Overview

As mentioned previously, the ATLAS and CMS requests were scrutinized in parallel by the combined team of ATLAS and CMS referees. After a joint meeting on June 19th, ATLAS was asked more questions which were answered in a phone meeting on June 29th. The CRSG chairman and a CRSG referee had the opportunity to attend the LHCC mini-review on July 6th, where further questions were asked and subsequent clarifications were provided. The level of detail provided by ATLAS has been excellent and has allowed the referees to reproduce independently the ATLAS computing model to a good level of accuracy. The majority of the requests are firm and justified, given the special data taking scenario for 2009 and 2010.A final iteration with ATLAS in early October 2009 resolved the major disagreement between our scrutiny and their original 2010 requests. This report reflects that modification, namely that the ATLAS 2010 request for Tier 1 CPU capacity for reprocessing was revised downward, from 59.0 kHS06 to 35.0 kHS06.

Generally speaking, the referees feel that the ATLAS computing model is adequate to survive the first impact with data taking in 2009 and 2010; nevertheless a substantial amount of work in the areas of data size, processing times, data format and distribution is required in order to match the challenges of steady-state running from 2011 onwards.

The ATLAS Computing Model

At the highest level, the model they use today is essentially unchanged from the one appearing in their 2005 TDR. ATLAS points out that many of the parameters driving the resource requirements predicted by their model have changed as the model has been confronted with (and calibrated by) the reality of data taking and data analysis on cosmic rays in 2008 and 2009Q1.

As noted in the 2008 scrutiny, their event sizes are larger now than they were in the original model assumptions. These have not changed since 2008, and there is no immediately obvious way to reduce them without compromising the experiment quality according to ATLAS. According to the referees, there is some space for reduction after a few months of data taking, by reducing the amount of information from the highest consumers; an example which comes to mind is the usage of zero-suppressed data from the calorimeters when the detector performance has been understood sufficiently well.

Experience with calibration and reconstruction of cosmic ray data has shown that recall speed of raw data from tape limits ATLAS ability to keep up with the data. As a consequence, the experiment has begun to rely much more heavily on disk storage for their reprocessing. This attitude is reasonable at the start of data taking, but not sustainable in the long term. Recent tests within the STEP09 challenge have shown data recall rates from tape well above targets. We therefore encourage ATLAS to further explore tape usage for scheduled activities, by using in particular pre-staging.

The time needed to generate and reconstruct a simulated event is a point of concern. ATLAS explained in great detail the single contributions, and the plans which have been implemented in order to reduce the required CPU power to a more manageable level. The referees encourage such plans, and recommend to use background samples from data whenever possible.

The analysis strategy foreseen by ATLAS is organized by trigger streams. This implies that analysis groups perform data reduction and filtering by processing big fractions of the data sample and creating in several cases derived formats. In order to make this process sustainable, data are replicated several times in the ATLAS computing clouds so that the associated CPU load can be shared. This strategy is different from CMS's, where data are centrally organized and managed as primary datasets, and each analysis group reduces them by processing relatively smaller chunks of data, with reduced requirements on data replication on disk. Each approach has its advantage and disadvantages. We encourage the experiments to use the upcoming data taking period in order to review critically if their data distribution model can be modified in order to maximize their physics output while keeping the overall resources at a manageable and scalable level.

Comments and recommendations

The ATLAS liaisons provided the CRSG team with a revision of their April 2009 resource request for 2009+2010. The T1 resources have been revised downward; in particular, the disk requests have gone down by about 40%. This is due to the reduction of data replicas at the Tier1s. The same data are nevertheless available on each the Tier2 clouds, as well as on tape. Other resource requests remain mostly the same or decreased by 10-20%.

The tables below show the resource estimations for 2009 and 2010. For each year, the results of the present scrutiny are shown in the column "Scrutiny", whereas the ATLAS requests are shown in the column "*ATLAS*".

Generally speaking, we feel that the resources are well justified, with the exceptions noted below. Discrepancies between the ATLAS request and our scrutiny or numbers we have been unable to fully scrutinize will be denoted in the tables below using ***<u>bold underlined italic font.</u>***

The CERN resources are reasonable and should be supported; as noted above, any shortage of resources at CERN would reflect in potential disruption of data, whereas shortage at Tier1s and Tier2s slows down the physics productivity of the experiments.

The recommended resources are shown in the following table. A detailed breakdown is given in the following paragraphs.

CERN resources	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Tier0 CPU (kHS06)	28.9	29.0	33.6	30.0
CAF CPU (kHS06)	28.0	28.0	36.5	37.0
Tier0 disk (PB)	0.7	0.7	0.7	0.7
CAF disk (PB)	2.6	2.6	2.9	3.2
CERN tape (PB)	5.1	5.1	8.9	9.0
Non-CERN resources	**2009**		**2010**	
	Scrutiny	**ATLAS**	**Scrutiny**	**ATLAS**
Tier1 CPU (kHS06)	78.8	81.0	190.8	216.0
Tier1 disk (PB)	11.8	12.2	21.4	21.9
Tier1 tape (PB)	8.0	8.3	13.7	14.2
Tier2 CPU (kHS06)	109.2	111.0	239.4	240.0
Tier2 disk (PB)	10.6	10.8	20.1	20.9

Tape resources

The ATLAS requests are consistent with our estimates.

CERN Tape (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Pre-existing data	2.80	2.80	5.10	5.10
Cosmic ray data	1.38	2.30	1.49	3.90
LHC data	0.92		2.19	
Total	**5.10**	**5.10**	**8.78**	**9.00**

Tier1 Tape (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
LHC RAW data	0.54	0.54	1.92	1.92
Real ESD+AOD+DPD data	0.67	0.59	2.38	2.07
Simulated data	2.29	2.65	4.44	5.13
Group+user	0.00	0.02	0.00	0.07
Pre-LHC data	4.00	4.50	4.00	5.00
Cosmics	0.50		1.00	
Total	**8.01**	**8.30**	**13.74**	**14.19**

Disk resources

The Tier0 disk resources are on the critical path. ATLAS foresees plenty of buffers which serve as safety factor to prevent data to “fall on the floor”.

T0 Disk (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Buffer for RAW & processed data	0.39	0.40	0.39	0.40
Buffers for I/O	0.34	0.34	0.34	0.34
Total	**0.73**	**0.74**	**0.73**	**0.74**

The CAF provides also critical services. The scrutinized disk resources should therefore be warranted to ATLAS.

- The “***group data***” requirement is assumed to be 20% of AODs + 200TB of already existing data.
- ***User data*** were not scrutinized

CAF Disk (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Fraction of current RAW data	0.08	0.08	0.59	0.67
Real AOD+DPD data	0.13	0.20	0.46	0.70
Simulated RAW+AOD+DPD data	0.50	0.47	0.98	0.94
Calibration and alignment outputs	0.05	0.05	0.16	0.16
Group data	***0.34***	***0.25***	***0.33***	***0.33***
User data (scratch)	***0.40***	***0.40***	***0.40***	***0.40***
Cosmics	***1.14***	***1.14***	***0.00***	***0.00***
Total	**2.64**	**2.59**	**2.92**	**3.20**

The disk space at Tier1 is used to store copies of the different data tiers, as well as non-beam data and buffers.

- There are two overall copies for each current data tier (ESD/AOD/DPD data/MC), one overall copy for the previous version.
- As expected, the disk space for real and MC data is dominated by ESDs (1PB in 2009, 3.7PB in 2010 for real data, 1.2PB in 2009 and 2.3PB in 2010 for MC data).
- The space required by RAW data (0.78 and 2.74 PB in 2009 and 2010) and ***cosmic ray*** data (3.7PB and 1.6PB in 2009/2010) can be reduced by using pre-staging from tape. In particular, the disk space required for ***cosmics*** does not seem justified.

- The requirement due to ***processing and I/O buffers*** for 2010 (2.9 PB) should be clarified.
- ***User data*** was not scrutinized.

Tier1 Disk (PB)	**2009**		**2010**	
	Scrutiny	**ATLAS**	**Scrutiny**	**ATLAS**
Current RAW data	0.78	0.78	2.74	2.74
Real ESD+AOD+DPD data	1.44	1.19	5.09	4.64
Simulated RAW+ESD+APD+DPD	2.35	3.19	4.56	5.33
Calibration and alignment outputs	0.50	0.45	1.60	1.50
Group data	1.87	1.79	2.38	2.59
User data (scratch)	***0.00***	***0.00***	***0.60***	***0.60***
Cosmics	***3.67***	***3.67***	***1.57***	***1.57***
Processing and I/O buffers	***1.14***	***1.14***	***2.90***	***2.90***
Total	**11.76**	**12.21**	**21.45**	**21.87**

In summary, the areas marked in ***bold underlined italics*** in the above table correspond to transient storage and cosmics, and account for about 5 PB disk space in 2009 and 7 PB in 2010. Given that ATLAS has 10 Tier1s, this means that in each Tier1 0.5 PB and 0.7 PB are dedicated to transient storage or cosmics. We feel these "buffers" somewhat large and that ATLAS should justify them better, especially for 2010, and extrapolate them to subsequent years.

We note incidentally that if ATLAS were to use the same data placement strategy at Tier1s as CMS, the total disk requirements at Tier1 would be about 9 PB and 15 PB in 2009 and 2010, to be compared with CMS's 6.5PB and 13.4PB. The difference between the two experiments is therefore largely due to the different data distribution strategies.

The Tier2 disk requirements are dominated by the real and MC replicas (1 replica per cloud; 10 clouds overall). The group and user data are generally used for the final user activities and are somewhat transitory in nature, the persistent data being generally represented by DPDs.

The data placement strategy at Tier2 is the following:

- 10 overall copies of the current version of AOD and DPD, both data and MC
- 10 overall copies of the previous version of AOD data (no DPDs are retained)
- 4 overall copies of the previous version of AOD MC (no DPDs are retained)

We are concerned that this distribution model may not be sustainable in the long term, and that ATLAS should consider revising it by e.g. streamlining data at an earlier stage. A thorough investigation of the ATLAS data distribution model must be carried on during the first year of data taking.

Tier2 Disk (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Current RAW data	0.78	0.78	0.27	0.27
Real AOD+DPD data	1.97	1.97	6.94	6.96
Simulated RAW+AOD+DPD data	4.38	5.12	8.49	9.23
Calibration and alignment outputs	0.05	0.05	0.16	0.16
Group data	0.62	1.08	0.79	1.87
Existing group data	1.00		1.00	
User data (scratch)	1.20	1.20	1.80	1.80
Processing buffers	0.60	0.60	0.60	0.60
Total	**10.60**	**10.80**	**20.06**	**20.89**

For illustrative purposes, the table below shows the disk requirements in a scenario similar to CMS, i.e. by halving the number of copies requested by ATLAS. The overall disk requirements are 7 PB and 12 PB for 2009 and 2010. For comparison, CMS requires 3.7PB and 9.2PB for 2009 and 2010, respectively.

Tier2 Disk (PB)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Current RAW data	0.78	0.78	0.27	0.27
Real AOD+DPD data	0.98	1.97	3.47	6.96
Simulated RAW+AOD+DPD data	2.28	5.12	4.42	9.23
Calibration and alignment outputs	0.05	0.05	0.16	0.16
Group data	0.31	1.08	0.40	1.87
Existing group data	1.00		1.00	
User data (scratch)	1.20	1.20	1.80	1.80
Processing buffers	0.60	0.60	0.60	0.60
Total	**7.20**	**10.80**	**12.12**	**20.89**

CPU resources

The Tier0 CPU resources needed for full reconstruction are straightforward to compute starting from the trigger rate end event processing time; partial processings are assumed to require 20% of full processing. Given that the duty cycle of the LHC will be small, the ATLAS requirement to perform ***first-pass processing at 200Hz***, i.e. their full trigger rate, is somewhat unnecessary.

Tier0 CPU (kHS06)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Full reconstruction	14.1	14.0	18.8	16.0
Partial processing and valdation	2.8	3.0	3.8	3.0
Merging and monitoring	***4.0***	***4.0***	***4.0***	***4.0***
Automatic calibration	***8.0***	***8.0***	***7.0***	***7.0***
Total	**28.9**	**29.0**	**33.6**	**30.0**

The tasks to be performed at CAF are high priority and should be granted.

CAF CPU (kHS06)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Partial reco, debug, monitoring	5.0	5.0	5.0	5.0
Non-automatic calibrations	4.0	4.0	3.5	4.0
Group activities	10.0	10.0	19.0	19.0
User activities	3.0	3.0	3.0	3.0
Servers	6.0	6.0	6.0	6.0
Total	**28.0**	**28.0**	**36.5**	**37.0**

ATLAS performs RAW data reprocessing at Tier1s. They require a reprocessing rate of 415 events per second. This figure has been scrutinized by LHCC and found adequate. About half of the Tier1 CPU resources are required for MC production, which accounts for most of the difference between ATLAS and CMS (which requires 46/100.4 kHS06 for 2009/2010). The number of MC events to be generated by the experiment has been also reviewed by the LHCC and found adequate.

Tier1 CPU (kHS06)	2009		2010	
	Scrutiny	ATLAS	Scrutiny	ATLAS
Re-processing	24.9	26.0	30.8	35.0[a]
Simulation production	35.9	34.0	99.5	98.0
Simulation reconstruction	3.2	7.0	11.8	12.0
Group (+user) activities	14.8	14.0	48.6	47.0
Total	**78.8**	**81.0**	**190.8**	**192**

[a] *Originally ATLAS requested 59.0 kHS06 for reprocessing in 2010. In a subsequent discussion with ATLAS, they noted an error in their calcualtion and agreed that the number should be revised downward. Their corrected request is 35.0 kHS06.*

The Tier2 CPU requirements are well balanced among simulation production and physics groups/user activities. The latter have been obtained by assuming reasonable processing times per event.

Tier2 CPU (kHS06)	2009		2010	
	Scrutiny	**ATLAS**	**Scrutiny**	**ATLAS**
Simulation production	68.1	68.0	116.7	114.0
Group activities	0.0	0.0	31.6	32.0
User activities	41.1	43.0	91.1	94.0
Total	**109.2**	**111.0**	**239.4**	**240.0**

Scrutiny of the CMS requests for 2009 and 2010

Overview

As mentioned previously, the ATLAS and CMS requests were scrutinized in parallel by a single team of referees. After a joint meeting on June 19th, CMS was asked more questions that were answered in a phone meeting on June 29th. The CRSG chairman and a CRSG referee had the opportunity to attend the LHCC mini-review on July 6th, where further questions were asked and subsequent clarifications were given. The level of information provided by CMS has been very good and allowed the referees to reproduce independently the CMS computing model to a good level of accuracy. The majority of the requests are firm and justified, given the special data taking scenario for 2009 and 2010; some requests for 2010 in the area of non-beam data seem less justified at this point and should be further clarified.
Generally speaking, the referees feel that the CMS computing model is adequate to survive the first impact with data taking in 2009 and 2010; minor work might be required in order to match the challenges of steady state running from 2011 onwards. The CMS computing model

As with the ATLAS experiment, the CMS computing model has not undergone significant changes since it was first published in 2005. In particular important parameters such as the data-taking rate, event sizes (RAW, RECO and SIM) and the CPU required to process a single proton collision have not changed. Still when the CRSG referees received the updates to the CMS computing model in April 2009 a number of secondary parameters had changed, some of which have important consequences for the total computing resources CMS anticipates using in the 2009/2010 running period. We summarise some of these updates here.

As a result of the published LHC machine commissioning schedule, CMS now anticipates that the collision overlap in the 2010 running will average 7, as the LHC expects to deliver the same current in fewer bunches. Simulation studies carried out by CMS show that this will require twice as much CPU per event for reconstruction. While this was always part of the CMS computing model as LHC design luminosity was reached, it was not foreseen for the low luminosities expected in 2010, but results from the machines stated commissioning plan.

Further, with the appearance of new LHC operating points (10 TeV and more recently 7 TeV) CMS anticipates generating 50% more MC than was foreseen in the original computing model (where all data was taken at 14 TeV).

As a result of the compressed time-scale for the 2009 and 2010 runs (one follows the other with only weeks of down-time in between) CMS plans a much more aggressive re-processing schedule than foreseen in their original computing model. Instead of three re-processings in the first year of data-taking and two in subsequent years, CMS now plans for two re-processings of the 2009 data-set and three more of the 2010 dataset. While this does not result in five 'full re-processings (they present a detailed, staggered re-processing schedule that seems well adapted to the LHC machine running schedule) between late 2009 and late 2010, it does mean that these re-processings must follow each other after only one or two months. While it seems feasible to validate new re-construction packages on this time-scale it does result in an average demand of 400 Hz of processing power 24/7/365, a significant increase relative to their original computing model.

Finally, CMS physics analysis plans have matured as a result of their computing exercises and analysis of cosmic ray data. They currently plan to stream events based on trigger information (electrons, muons, jets, missing Et, etc.) If an event fires both the electron and jet trigger (something their simulation says should happen 20% of the time in 2009) then it will be streamed to both datasets and thus will be reconstructed (and re-reconstructed) twice. A careful analysis of their trigger algorithms leads them to conclude that there will be a 40% increase in the number of events on analysis streams, relative to the number of raw triggers. While this vastly simplifies the analysis of secondary and tertiary data-sets – greatly simplifying the Tier2 analysis stage – it does result in a 40% increase in the Tier0 and Tier1 event processing CPU and a similar increase in the storage needs for RAW and RECO datasets. Based on their simulation studies and the validation of these studies using the 2009 datasets CMS anticipates reducing the trigger overlap in 2010 to 20% (or less). However neither of these factors appeared in the computing model prior to April 2009.

Requests and recommendations

The CMS liaisons provided the CRSG team with a revision of their April 2009 resource request for 2009+2010 which showed an overall reduction of 10-20%.

The tables below show the resource estimations for 2009 and 2010. For each year, the results of the present scrutiny are shown in the column "Scrutiny", whereas the CMS requests are shown in the column "CMS".

Generally speaking, we feel that the resources are well justified, with the exceptions noted below. The CERN resources are reasonable and should be supported; as noted above, any shortage of resources at CERN would reflect in potential disruption of data, whereas shortage at Tier1s and Tier2s slows down the physics productivity of the experiments.

The recommended resources are shown in the following table. A detailed breakdown is given in the following paragraphs. Significant discrepancies between the CMS request and our scrutiny, or numbers we have been unable to fully scrutinize will be denoted in the tables below using ***bold underlined italic font.***

CERN resources	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Tier0 CPU (kHS06)	37.1	37.1	61.9	61.9
CAF CPU (kHS06)	7.6	7.6	34.7	34.7
Tier0 disk (PB)	0.4	0.4	1.0	1.0
CAF disk (PB)	1.3	1.3	3.1	3.1
CERN tape (PB)	7.4	7.3	12.6	14.6
Non-CERN resources	**2009**		**2010**	
	Scrutiny	**CMS**	**Scrutiny**	**CMS**
Tier1 CPU (kHS06)	45.8	46.0	101.3	100.5
Tier1 disk (PB)	6.5	6.5	13.7	13.4
Tier1 tape (PB)	10.3	11.9	23.3	23.3
Tier2 CPU (kHS06)	73.2	75.2	199.0	195.0
Tier2 disk (PB)	3.7	3.7	9.0	9.2

Tape requests

CMS provided the details for **CERN tape** in different categories than those used for ATLAS. They list “RAW data”, “Reconstructed data”, “AOD”, “Last years data”, but did not split it up in terms of commissioning, cosmic rays or beam data, nor could we infer it from other available documentation. For instance, they indicate they plan to store 11TB of cosmic data per day, but do not detail the relative proportion of RAW versus RECO/AOD. The same applies to commissioning data. We prefer the categories in the table below, similar to ATLAS, since it is immediately evident how the different numbers scale with LHC live time, cosmic data taking time.

In order to infer their requests we use the following information provided by CMS:

- Pre-LHC data: 3PB
- Cosmic ray data: 75 days in 2009, 25 days in 2010; 11TB to tape per each day
- CAF data: 0.9 PB in 2009, 2.6PB in 2010

CERN Tape (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Pre-existing data	***3.00***	***3.00***	7.40	***7.40***
Cosmic ray data	0.83	0.83	0.28	0.28
Group data (CAF)	***1.00***	***0.90***	***2.60***	***2.60***
LHC data	2.57	2.58	3.95	4.34
Total	**7.40**	**7.30**	**14.22**	**14.61**

CMS did not provide any details for **Tape at Tier1**. The breakdown per activity has been computed starting from the data volumes and the reprocessing factors, provided by CMS. The cosmic data of 2009 and pre-LHC data are disposed of in 2010.

Tier1 Tape (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
LHC RAW data	1.07	Data	2.48	Data
Real ESD+AOD+DPD data	1.29	not	5.15	not
Simulated data	3.15	available	7.92	available
Group+user	1.00	from	1.00	from
Pre-existing data	***3.00***	CMS	6.51	CMS
Cosmics	0.83		0.28	
Total	**10.33**	**11.90**	**23.33**	**23.30**

Disk requests

The CMS **disk request at Tier0** is reasonable; they are somewhat increased with respect to the past, but since the disk at Tier0 is on the critical path the increase adds an additional safety factor which prevents data to “fall on the floor”.

T0 Disk (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Buffer for RAW&processed data	0.31	0.31	0.77	0.77
Buffers for I/O	0.09	0.09	0.23	0.23
Total	**0.40**	**0.40**	**1.00**	**1.00**

The requests for ***disk at CAF*** have not been scrutinized in detail. Rather than requesting for given amounts of RAW, RECO, AOD and simulated data, CMS organizes CAF disk according to activities, but they not give any details on why each activity requires such disk space.

Given the critical service provided by the CAF, we warrant the resources requested by CMS (1.3PB in 2009 and 3.1PB in 2010)

CAF Disk (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Calib and alignment	***0.13***	***0.13***	***0.31***	***0.31***
Physics Monitoring	***0.52***	***0.52***	***1.24***	***1.24***
Commissioning	***0.65***	***0.65***	***1.55***	***1.55***
Total	**1.30**	**1.30**	**3.10**	**3.10**

The **Tier1 disk** requests are obtained as follows:

- For **real data**, in both 2009 and 2010, there are on Tier1 disk: a full copy of RAW data, 40% of the current RECO version, 40% of the previous RECO version over the whole Tier1 ensemble; 7 copies of the current AOD version (1 copy per Tier1), no copies of previous versions
- For **simulation**, in both 2009 and 2010, there are on Tier1 disk: 10% of sRAW data, 10% of the current sRECO version, 10% of the previous sRECO version over the whole Tier1 ensemble; 7 copies of the current sAOD version (1 copy per Tier1), no copies of previous versions
- In the scrutiny, only 50% of fast simulation AOD is retained on disk, i.e. thus we can assume that a re-generation supersedes the previous one.

The disk space is dominated by RAW real data (1.53PB in 2009, 3.54PB in 2010 for real data) and by the real + simulated AOD (2.46PB in 2009 and 5.74PB in 2010); the latter is due to the factor 7 replica for AODs.

Contrary to ATLAS, cosmics are not resident on disk and are expected to be (pre-)staged from tape when needed. The same technique could be used for RAW data, thus reducing the corresponding disk space considerably.

The analysis space and buffers are adequate.

Tier1 Disk (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
RAW data	1.53	1.54	3.54	3.93
RECO data	0.41	0.41	1.42	1.57
sRAW data	0.21	0.20	0.33	0.38
sRECO data	0.11	0.10	0.36	0.28
AOD (real+sim)	2.46	2.46	5.74	4.96
Analysis space	***1.42***	***1.42***	***1.69***	***1.69***
WAN buffers	***0.36***	***0.36***	***0.60***	***0.60***
Total	**6.50**	***6.49***	**13.68**	**13.40**

The **Tier2 disk** requests are obtained by CMS assuming that the Tier2 ensemble contains: a factor 4 replication of 30% of the current and 30% of the previous AOD (real and MC) versions; 35% of the current and 35% of the previous RECO (real and MC) versions.

The rest of the Tier2 disk space is used for group data and MC processing buffers
CMS did not give a detailed breakdown for the analysis activities, which are nevertheless found to be adequate. The group data for 2009 has been assumed to be 1PB; for 2010, it has been obtained by multiplying the number of working groups (16) by the AOD replication factor (4) and the space CMS assigns to a working group in a Tier2 (33TB), giving 2.11PB

Tier2 Disk (PB)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Real RECO	0.25		1.60	
Real AOD data	0.34		1.04	
Simulated RECO+AOD data	1.10	2.71	3.41	8.34
Calibration and alignment outputs	0.00		0.00	
Group data	1.00		2.11	
Processing buffers	***1.00***	***0.99***	***0.86***	***0.86***
Total	**3.70**	**3.70**	**9.03**	**9.20**

CPU requests

The **Tier0 CPU** resources needed for full reconstruction are straightforward to compute starting from the trigger rate end event processing time. Differently from ATLAS, CMS performs first-pass processing at 150Hz, i.e. half their trigger rate.

Since these are on the critical path, the resources required for the other tasks should be warranted.

Tier0 CPU (kHS06)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Reconstruction of RAW data	28.0	28.0	48.0	48.0
Reconstruction of express stream	***4.0***	***4.0***	***8.0***	***8.0***
VO servers	***4.0***	***4.0***	***4.0***	***4.0***
Reconstruction of calibration stream	***1.1***	***1.1***	***1.9***	***1.9***
Total	**37.1**	**37.1**	**61.9**	**61.9**

There is a substantial increase in the 2010 **CAF CPU** requirements, basically due to commissioning (which seems quite out of place in 2010) and physics monitoring. This requirement needs to be scrutinized more thoroughly in the immediate future.

CAF CPU (kHS06)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Calibration and alignment	***3.0***	***3.0***	***3.5***	***3.5***
Commissionino	***3.0***	***3.0***	***17.4***	***17.4***
Physics monitoring	***1.5***	***1.5***	***13.9***	***13.9***
Total	**7.6**	**7.6**	**34.7**	**34.7**

The **Tier1 CPU** requests are dominated by reprocessing of real and simulated data. There was quite some discussion in the review process about the reprocessing rate. As noted above, ATLAS plans for 415Hz steady rate in 2010, CMS is giving about the same figure (400Hz). The LHCC recommends keeping such a steady rate.

Contrary to ATLAS, there is no MC production at Tier1 in CMS. For this reason, the CMS resources at the Tier1 are significantly smaller than ATLAS's.

Tier1 CPU (kHS06)	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Re-processing of real data	18.4	18.7	44.1	44.2
Calibration tasks	***0.0***	***0.5***	***0.0***	***1.9***
Re-processing of simulated data	19.3	18.5	35.3	32.3
Group activities (skimming)	8.1	8.3	21.9	22.1
Total	**45.8**	**46.0**	**101.3**	**100.5**

The Tier2 CPU requests are reasonable. The breakdown of the user activities has not been detailed by CMS. Assumptions similar to the ones made for ATLAS lead to similar results.

Tier2 CPU (kHS06)	**2009**		**2010**	
	Scrutiny	**CMS**	**Scrutiny**	**CMS**
Simulation production	49.7	49.6	50.7	50.1
User activities	23.5	25.6	148.3	144.9
Total	**73.2**	**75.2**	**199.0**	**195.0**

Review of the LHCb experiment computing model and scrutiny of requests for 2009 and 2010

Introduction

In September 2008 the computing requirements of the LHCb experiment for 2008 and 2009 have been analysed in detail. In April 2009 a draft report has been delivered by the scrutiny group that compared the new LHCb requirements with estimates according to the new accelerator schedule.

The present document provides an update of the April 2009 scrutiny after revision of the LHCb requirements for 2009 and 2010 on June 30, 2009, and a respective clarification of the underlying parameters of the LHCb computing model.

The LHC running model used here assumes the same pattern of physics data taking as in the April 2009 scrutiny, common to all experiments.

The scrutiny performed by the referees include a full re-simulation of the model applied and validated for the 2008 scrutiny, an analysis of the model parameter changes proposed for 2009 and 2010 and an attempt to take into account the parameter changes in the model simulation. A number of remarks and recommendations are also given.

In the scrutiny we used the updated LHCb requests and several documents containing some computing model parameters and detailed explanations of the changes. The interaction with the LHCb computing coordinators was very helpful and we would like to thank them for their collaborative attitude.

Resources model simulation (model A)

The model applied for 2008 scrutiny is repeated here for the new running scenario in 2009 and 2010. For better comparison with the old report it is assumed that the full resources for the second period (i.e. October 2009 – March 2010) will be installed by the sites in September 2009 at latest, and that they are counted for 2009.

The following two tables compare respective model assumptions in the old and the new report:

Year	pp operations **old**		pp operations **new**	
	Physics beam time (seconds/yr)	Luminosity (cm^{-2} s^{-1})	Physics beam time (seconds/yr)	Luminosity (cm^{-2} s^{-1})
2008	**$3•10^6$**	$2× 10^{32}$	**--**	--
2009	**$9•10^6$**	$2× 10^{32}$	**$1.7•10^6$**	$2× 10^{32}$
2010	$1•10^7$	$2× 10^{32}$	**$4.3•10^6$**	$2× 10^{32}$

Year	pp operations **old**		pp operations **new**	
	Expected DAQ events	Assumed MC events simulated	Expected DAQ events	Assumed MC events simulated
2008	**$6.0•10^9$**	**$1.6•10^9$**	--	--
2009	**$1.8•10^{10}$**	**$4.0•10^9$**	**$3.4•10^9$**	**$1.6•10^9$**
2010	$2.0•10^{10}$	$4.0•10^9$	**$8.6•10^9$**	**$4.0•10^9$**

The following table provides a comparison of the resources simulations for 2009 and 2010 according to the two scrutiny models with old and new beam parameters (a detailed breakdown of this exercise for different physics categories of the LHCb computing model are given in Annex 1). Please note that all CPU resources are now expressed in the new HEP-SPEC06 (HS06) unit. Numbers from the old report have been converted with 1 kSI2k = 4 HS06.

Summary of the SG calculations	**Old report (September 2008)**		**This report (model A)**	
	2008	**2009**	**2009**	**2010**
CPU	**26,24**	**71,24**	**23,64**	**61,56**
Disk (TB)	**1162**	**3501**	**870**	**2385**
Tapes (TB)	**1150**	**4626**	**707**	**2545**

It should be stressed again that, except for the new beam time and hence deduced expected data acquisition rates (i.e., number of events per year), the model applied here was exactly the one used in the 2008 scrutiny, which resulted in a good agreement with the LHCb requests, and which was in this way substantiated and supported by the Scrutiny Group.

Except for detector data that have already been taken in the absence of the accelerator beam (e.g. cosmics) and simulation results that have already been stored, the new running strategy for 2009 and 2010 should roughly correspond to a shift in the resource requirements estimated in 2008. Indeed, the results obtained by simply scaling down the running time in this model for 2009 compare well with the original 2008 estimates (are accordingly smaller), but they disagree by large factors with the new computing requirements announced by LHCb for 2009 and 2010. This raises the question of modifications of the computing model and its implementation. This issue is discussed in the next section.

Analysis of the computing model evolution

Since September 2008, and especially after the "Chamonix LHC Performance Workshop" at the end of February 2009, LHCb had extensive discussions with the physics analysis groups on their detailed physics programmes after the LHC re-start in October 2009. Consequently, the parameters of the LHCb computing model used as input for the new resource requests

2009/2010 underwent a significant number of modifications that are analysed in this section in more detail.

Data acquisition

While the overall trigger rate of the detector has been kept at 2 kHz the detailed physics programmes for Flavour physics, CP violation, rare decays and charm physics led to a significant change of rates of the different physics channels. Compared to the situation in 2008 (with 200 Hz for b-exclusive and 900 Hz for b-inclusive), analysis will not be primarily aimed at getting results on CP violation but mainly concentrate on the b-exclusive channel (1500 Hz) for measurements of cross-sections, branching fractions etc. as well as (together with signals from all channels) on getting an overall understanding of the detector behaviour and required calibration, alignment and reconstruction parameters under real collisions. Effort is ongoing to further reduce the RAW event sizes from 35 kB/evt in 2009 to 30 kB/evt in 2010.

Real data reconstruction

Reconstruction is performed by LHCb in two steps. In a first step, RAW data are reconstructed into an intermediate (reduced) rDST format. In a second step, rDST are selected and only a fraction of the data samples is converted into the final AOD format, which consists of respective fully reconstructed DST + RAW information. rDST data are used to select relevant decay modes during the so-called selection or stripping phase, thus significantly reducing the final number of events for physics analysis (so-called "stripping reduction factors"). However, compared to the initial strategy of 2008, a significant number of changes have been introduced.

- rDST sizes have been increased by 75% to keep more information at the beginning for better understanding the detector, while full DST sizes have been decreased by about 33%.
- The stripping reduction factors in the four physics channels (b-excl; b-incl; J/Ψ; D*) have been reduced from (10;100;5;5) to (1;1;1;1) in 2009, and changed to (10;10;10;10) in 2010, with large implications on the required disk storage. Note that the average number of stripping passes now is 3 in 2009 and 4 in 2010, compared to the originally planned 2 passes during the first one and half years of data taking.
- LHCb has introduced a new so-called "micro DST" format into the reconstruction, which reduces the DST event sizes. The micro DST format is especially applied to the b-exclusive data in 2009 which reduces the respective storage requirement by a factor of 10 and thus compensates for the reduced stripping factor by the same amount.
- Compared to the initial strategy in 2008, only 10% of RAW and rDST data are kept on disk during the stripping process.
- The original strategy to utilize the LHCb online farm for a second pass reconstruction during accelerator shutdown had to be given up since there is no winter shutdown at least at the end of 2009. Therefore, respective compute requirements had to be redistributed among CERN and the Tier1s.
- LHCb will have no experience with real data for alignment and calibration prior to first beam, due to the low utility of cosmic 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 data taking, implying respective re-processing of previous data. LHCb therefore assumes that the number of events that will be reconstructed is 3 times the number of events recorded (3 average reconstruction passes instead of the 2 according to planning in 2008
- Also, the distribution of processing fractions among CERN and the Tier1s changed, with respective implications on the distribution of CPU power and storage. While it was initially anticipated to do 1/7 of the reprocessing at each of the sites (CERN plus 6

Tier1s), these fractions were modified to 100% at CERN during early 2009, 40% at CERN during late 2009, and 25% at CERN during 2010, respectively.

- Calibration and alignment was not taken into account in the 2008 models. A CPU requirement of 0.4 kHS06*year at CERN is applied for this.

MC simulation and reconstruction

Several new parameters of the next LHC physics run have been made available (esp. beam energy near 5 TeV) and LHCb started the next MC simulation campaign at the beginning of 2009. LHCb plans for

- about $3•10^9$ minimum bias events (b-exclusive) in 2009 and $2•10^9$ in 2010 for the tuning of the L0 and HL trigger and stripping,
- about $5•10^8$ b-inclusive events in 2009 and $1•10^9$ in 2010 for evaluating background sources,
- a few 10^6 events of many channels to evaluate signal efficiency, acceptances and resolutions.

Compared to the strategy in 2008, this is an increase of the total number of events by about 80% for b-exclusive and a decrease of about 45% for b-inclusive, which is largely dictated by the physics programme. Moreover, new samples will be required if new (different) beam conditions (pileup, more accurate beam energy, luminosity) and detector conditions will become available.

While in 2008 MC event sizes as well as processing times for b-inclusive and b-exclusive events were identical, they are different in the new models now:

- MC RAW b-<u>in</u>clusive events have been increased from 150 to 250 kB/evt while respective DSTs decreased from 250 to 50, and MC RAW and DST for b-<u>ex</u>clusive events are much smaller now, with 10 and 20 kB/evt, respectively.
- Simulation times are now estimated with 360 instead of 300 HS06*s/evt for b-<u>in</u>clusive, and 100 instead of 300 HS06*s/evt for b-<u>ex</u>clusive events.
- Reconstruction times are slightly increased from 9.6 to 12 HS06*s/evt for both, b-inclusive and b-exclusive.

Finally, LHCb moved to the strategy of using all available resources for Monte Carlo production and reconstruction during early 2009 (10% at CERN, 40% at Tier1s and 50% at Tier2s) and plans to switch over to “MC production at Tier2s only” as soon as real data are available and compute resources at CERN and Tier1s are required for DATA processing and reconstruction tasks.

Analysis

Motivated by discussions with the physics groups, the LHCb analysis strategy has been significantly refined. Compared to the initial assumption in 2008, that 80% of the physicists will process about 10^6 events per job and 20% will run over larger samples of 10^7 events per job during a normal running year, these numbers are now 10^5, 10^6, and 10^9 events per job by 55, 40 and 5% of the physicists, respectively. Furthermore, the number of physicists active in analysis increased from 140 to 160, and the number of analysis jobs per week has been adjusted according to the physics programmes, such that the total number of analysis jobs per year roughly increased by a factor of three, and the average number of events processed per job increased by about 30%. Altogether, this leads to a significant increase of the analysis effort in terms of CPU power compared to the initial assumptions in 2008.

According to the initial strategy data analysis was planned to be performed in batch mode through Grid jobs with 25% at CERN and 75% at the Tier1s. However, to ensure that LHCb can react quickly to the early data in understanding the detector and producing early physics results LHCb plans to increase the analysis fraction at CERN to 70% during initial accelerator start-up, with a decreasing percentage starting of 50% in late 2009 and 40% in 2010, while the target level of 25% will be reached in 2011.

Also, LHCb has introduced the new concept of ToyMC analysis running at 50% at Tier2s and the other 50% distributed over CERN and Tier1s.

CPU power for servers

In April 2009 LHCb added several kHEP-SPEC06 to the requirements for dedicated servers at the CERN and Tier1s. Following the recommendation of the scrutiny group this has been removed again from the total compute requirements.

Other comments / disk cache in front of T1D0 storage

LHCb added requirements for a disk cache of 70 TB in front of tape systems per Tier0 and Tier1, which increases the total requirement by 490 TB. This is in fact a kind of temporary storage that is not available for effective storage of real physics or simulation data which might produce a significant cost that has not been accounted for in the past. Following the recommendation of the scrutiny group this fraction of resources is still included in the requirements.

The previous changes in the computing model were analyzed by the LHCC in their July mini-review.

Tests of the new model assumptions (model B)

Given the changes of old and introduction of new parameters discussed in the previous section we have updated our model and performed a new requirements simulation. The following three tables provide a comparison of the resources requested by LHCb with these new simulations for 2009 and 2010 in terms of CPU, disk and tape capacity per task or data category. Note that these tables provide a comparison of both models per calendar year.

Significant discrepancies between the LHCb request and our scrutiny, or numbers we have been unable to fully scrutinize will be denoted in the tables below using ***<u>bold underlined italic font.</u>***

CPU (kSH06*year)	**LHCb request June'09**		**New RSG model (model B)**	
	2009	**2010**	**2009**	**2010**
Stripping	1,77	4,31	1,94	4,36
Full reconstruction	2,37	17,57	2,43	17,05
MC	22,66	24,64	22,66	24,79
Analysis	***<u>10,16</u>***	***<u>34,99</u>***	***<u>5,28</u>***	***<u>19,48</u>***
Calib & alignment	0,40	0,40	0,40	0,40
Total	**37,36**	**81,91**	**32,72**	**66,07**

Disk (TB)	**LHCb request June'09**		**New RSG model (model B)**	
	2009	**2010**	**2009**	**2010**
Full RAW	6	48	6	48
Full Reduced DST	6	40	6	40
Stripped RAW+DST	1960	***<u>3355</u>***	1850	***<u>2226</u>***
TAG	47	241	42	217
Analysis	234	431	200	398
Total	**2253**	**4115**	**2104**	**2929**

Tape (TB)	LHCb request June'09		New RSG model (model B)	
	2009	2010	2009	2010
Full RAW	83	755	84	756
Full Reduced DST	163	1039	167	1007
Stripped RAW+DST	678	1776	720	1640
TAG	14	107	17	111
Total	**938**	**3677**	**988**	**3514**

In our model the following simplifications were assumed:

- LHCb assumes a split of 2009 into an early and a late phase (April – September 2009 and October – December 2009, respectively). Since our current model doesn't allow for a respective split, a few averaged parameters had to be used. Especially, LHCb assumes 10 average reconstruction passes and no stripping in early 2009 and 3 reconstruction and stripping passes during late 2009, while we use 4 reconstruction and 3 stripping passes during the whole 2009.
- The total number of analysis jobs in 2009 has been summed up while the average number of events per analysis job is a weighted average over the early and late phase in our model B.
- The data analysis fractions at CERN and Tier-1s were assumed with 50% both during the whole year 2009.
- The new LHCb model assumes a breakdown of CPU usage between "real" data analysis + ToyMC and "MC" data analysis + MCToy, with varying analysis fractions at CERN, T1s and T2s. MCToy in general and analysis at T2s in detail is a new feature of the LHCb model hasn't been taken into account in our model.

A comparison of the tables above shows that our simplified model reproduces the basic features of the total LHCb requirements in the different data categories per calendar year with two exceptions:

- Our model underestimates the CPU requirement for analysis in 2009 and 2010 by about a factor of two. Provided with the original LHCb model data sheets we verified that this is a result of the missing MCToy in our simulations. Switching off MCToy in the LHCb model reduces the total CPU analysis requirements in 2009 and 2010 to 5,08 and 20,8 kSH06*year, respectively, and doesn't have any implications on the storage requirements.
- Our model undershoots the disk requirement for AOD data in 2010 by about 30%. On the other hand, our model reproduces the overall tape requirements in the different data categories including AOD within 10%, indicating that the overall amount of data produced (and stored on tape) by this model is nearly correct. This discrepancy could not be clarified completely.

In summary, this comparison demonstrates that, taking into account most of the new and changed parameters our simplified model is able to well reproduce the basic features of the total LHCb requirements for the different data categories per calendar year. It should also be noted that due to the changed strategy the LHCb model is not dominated by MC any more. Compared to the 2008 scrutiny, 43% of the CPU power in 2010 is requested for analysis, 21% for reconstruction and only 30% for simulation. Also, storage requirements are largely dominated by AOD data.

An interesting question is the amount of proportionality between the requested resources and the LHC running time for physics. The amount of Monte-Carlo simulation is considered to be independent of 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). It is estimated 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 $1.7 \cdot 10^6$ seconds until April 2010 and “2010” means $4.3 \cdot 10^6$ seconds as of April 2010.

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%	

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, it is 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%.

Conclusions and recommendations

The resources requirements of the LHCb experiment are presented in the following table.

Date	Site	kHS06	Disk (TB)	Tape (TB)
2009	CERN	15	720	1000
	Tier-1	31	1740	1000
	Tier-2	31	20	0
2010	CERN	23	1290	1800
	Tier-1	44	3290	2400
	Tier-2	38	20	0

These requirement present significant reductions when compared to April 2009 requests. Nevertheless, they present significant increases compared to the 2008 estimations reflecting significant changes in the computing model, justified by the effort to prepare for the data taking and the first physics analyses. A number of features have been identified:

- The LHCb CPU computing requirements are expressed in kHS06 and reflect the peak computing requirements. This is unusual since it assumes periods of high activity interleaved with periods of idleness, and is also not in line with the LCG accounting practice. The unused resources and the associated extra costs induced in the computing centres (like cooling, electricity and storage space) must be balanced by strong scientific or policy arguments. Upon request, LHCb provided us with the following, additional table in units of kHS06*years, which has been analyzed and justified with model B in section 4 of the report.

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

- The changes of the computing model ingredients are largely dictated by concerns related to the LHC start-up. Security factors are applied (for instance in the event size) which lead sometimes to significant increases. These factors represent for the time being reasonable estimations and will have to be adjusted once real data is taken.
- The CPU resources do not directly scale with the LHC running time since the resources are dominated by the Monte Carlo simulation which proceeds largely independently of the data taking. The amount of scaling for CPU is 50% at CERN and 25% in Tier1. The situation is different for disk and tapes where the amount of proportionality is 75% (90%) for CERN (T1) resources. A reduction in the LHC running time will translate therefore in an almost proportional reduction of the amount of disk and tapes resources.
- The increased analysis power at CERN (50%) is related to the need of quick turn-around in the first analysis of the data. LHCb presented a clear plan to reduce in the next years the computing power at CERN to 25%, as expected from the TDR. This should be accompanied by clear measures to improve the distributed analysis, in collaboration with the LCG project.
- The uncertainty related to the calibration and reconstruction software reliability is taken into account by proposing an average number of three reconstruction passes for each period. The resources allocated to this activity amount for about 10 to 15% at CERN. A general strategy is not mentioned and an intuitive approach for the large scale reprocessings is envisaged at present.
- A new type of analysis “Toy Monte Carlo” has been introduced which almost doubles the total CPU requirements for analysis. The relevance of this type of analysis has to be established in the next evaluation.
- The number of Monte Carlo events is in net increased with respect to the previous estimations. The relevance of large samples of simulated events with non-definitive reconstruction is to be demonstrated.
- A new data format (mDST) is introduced such that access to the full data set is made possible without increase in the storage resources.
- We recommend considering an average disk cache size per site in front T1D0 storage in future requirements tables as this might be a significant cost factor (depending on the data handling strategies) at least at LHC start-up when remaining background storage

is still comparatively small. This additional disk space of 70 TB at CERN and each of the six T1s is already included in the requirements table 3 above, but not included in the comparison of model B with the respective LHCb model in section 4.

In conclusion, the CRSG referees evaluated the LHCb computing requirements and observed a significant effort to adapt the computing model to the conditions of the data taking and to take into account the unexpected running conditions. Since the computing must not be the limiting factor in the first year of the data taking, we recommend that the requested resources are allocated for 2009 and 2010. We identified however a number of points, listed above, where the assumptions will have to be verified and followed closely after the first real data processing period or after experience with large Monte Carlo samples. We recommend a full review of the computing model during the first long shutdown of LHC, involving also the LHCC.