

Report of the Computing Resources Scrutiny Group 2009

The purpose of the CRSG is to inform the decisions of the Computing Resources Review Board (C-RRB) for the LHC experiments.

Every year the CRSG should 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 composition of the CRSG is established in the WLCG MoU. Members are independent and have no link whatsoever to the experiments they review.

According to the new schedule of the LHC 2009 and 2010 (from October to October) will formally constitute as a single run, with a short break over the end of 2009 period but for the purposes of WLCG procurement and availability it counts as two separate periods:

- From October 2009 to the end of March of 2010. We shall refer to this period as “2009”.
- The second period extends from April 1st 2010 to October 2010. We shall refer to this period as “2010”.

The resources for 2009 have to be in place by September 1st 2009 and the resources for 2010 by April 1st 2010.

Report of the Computing Resources Scrutiny Group 2009

The WLCG project keeps an accounting of usage 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 pledged resources, disk at a 75% and massive storage at the 60% level.

The figures show a clear improvement of 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. Yet, 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 to us and this is a cause of concern to some extent

The usage of Tier2 resources is higher in general, although extremely irregular in distribution.

Report of the Computing Resources Scrutiny Group 2009

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). The above numbers correspond to an effective beam time that differs considerably from the ones used in the 2008 report.

In the 2008 report 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 the revised LHC schedule the efficiency factor is conservatively taken to begin at a modest 10% in 2009 and ramp up to 32% only at the end of 2010.

Report of the Computing Resources Scrutiny Group 2008

The CRSG proposed a standard set of assumptions on beam time. These assumptions have been used for scrutinizing all experiments. They differ considerably from previous scenarios in the case of 2008, and only slightly in 2009. An 'efficiency' of 50% has been assumed in order to extract useful beam time from the total amount that the accelerator will be running. This is an optimistic assumption (recent public presentations suggest that 40% is closer to reality and this is perhaps still too optimistic for the first months of running).

These beam times would correspond to 3 months of data-taking in 2008 and 7 months of data-taking in 2009 for proton-proton (pp) operations, and 0 months in 2008 and 1 month in 2009 for heavy ion (AA) operations. These were rather optimistic, but attainable, expectations.

However the 19 September events forced a last-minute change of the scrutiny. The new ones assume that 2009 will be a nearly normal year as far as running conditions.

Year	pp	AA
	Beam time (seconds/year)	Beam time (seconds/year)
2008	0.3×10^7	0
2009	0.9×10^7	10^6
2010	10^7	10^6

Year	pp	AA
	Beam time (seconds/year)	Beam time (seconds/year)
2008	0	0
2009	0.9×10^7	10^6
2010	10^7	10^6

Report of the Computing Resources Scrutiny Group 2009

The present experimental request was made public on April 7th and it was not previously reviewed or endorsed by the LHCC or the CRSG.

The CRSG apologizes for making available the scrutiny with such short notice.

The CRSG regrets that it had only two weeks to carry out a review process of such complexity.

The CRSG stresses that this scrutiny has to be considered preliminary.

In carrying out the present scrutiny the scope of the CRSG is largely limited to the implementation of the respective computing models. The review of the computing models themselves has been in the hands of the LHCC.

While the methodology used has varied, uniformity in the scrutiny has been a priority. The methodology has been identical to the one used in the 2008 report.

For a proper scrutiny we have yet to see real collisions and real data with the computing models going through a reality check and we have to be flexible and prepared to respond to unforeseen circumstances.

Report of the Computing Resources Scrutiny Group 2009

Experiments compare their requests for 2009 and 2010 to the current pledges, while it would perhaps be more meaningful to compare 2009 to 2008, and 2010 to 2009 in view of the delay in the LHC commissioning for physics and the reduced beam time.

The expected beam time in 2009 (1.7×10^6 s) is approximately half the beam time assumed in our 2008 scrutiny for 2008 (3×10^6 s). Likewise for 2010 (4.3×10^6 s versus 9×10^6 s considered in the 2008 assessment for 2009) but direct comparison may be too simple:

- Experiments have been collecting cosmics over the last months and will continue to do so and this data needs to be stored (partially).
- There may also be updates of the computing models or modifications.
- More Monte Carlo simulated data may be required.

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, one needs to make considerable modifications to the computing models.

Some of the computing models have changed in ways that the CRSG did not expect.

We fear that these some of these modifications represent changes that, if agreed, make some of the computing models unsustainable.

The CRSG has also been left with the impression that some of the modifications proposed do not have an understandable bottom-up justification

Report of the Computing Resources Scrutiny Group 2009

Taking as starting point our baseline report submitted to the C-RRB in November 2008, we have made those modifications that we understand justified, rejecting others that appear to us unjustified, unrealistic or unsustainable.

Taking into account

- the new efficiency and availability conditions,
- the convenience of some modifications for the first year or running
- the necessity of additional Monte Carlo simulated data in some cases

we conclude that the experimental needs for 2009 and 2010 should be roughly satisfied by granting the resources originally planned for 2008 and 2009, respectively.

In fact a small reduction with respect to the resources approved for one year earlier (at the 10% level) may even be justified (10% is the assumed error of our estimates anyway).

In the case of massive storage a substantially larger reduction may be advisable as the original plan envisaged data taking already in 2007.

Report of the Computing Resources Scrutiny Group 2009

ALICE

	Old report (October 2008)		This report (April 2009)	
	2008	2009	2009	2010
CPU (kHS06)	98	172	94.6	206
Disk (PB)	14.4	22.1	12.1	21.9
Tapes (PB)	6.7	17.1	4.5	15.1

Report of the Computing Resources Scrutiny Group 2009

CMS

2009

Resource	Tier0	CAF	Tier1	Tier2	Total	Fraction of 2008
CPU(kHS06)	22.0	8.0	36.4	45.6	111.9	0.83
Disk (TB)	286	1670	7010	7010	15975	1.00
Tape (TB)	2200		5000	-	7200	0.72

Report of the Computing Resources Scrutiny Group 2009

CMS

2010

Resource	Tier0	CAF	Tier1	Tier2	Total	Fraction of 2009
CPU(kHS06)	29.2	6.8	46.8	65.6	148.8	0.85
Disk (TB)	286	2580	7200	5300	15365	1.00
Tape (TB)	4300		12000	-	16300	0.66

Report of the Computing Resources Scrutiny Group 2009

LHCb

Resource	Old report (October 2008)		This report (April 2009)	
	2008	2009	2009	2010
CPU (kSH06*yr)	26.24	71.24	35,68	52,12
Disk (TB)	1162	3501	1738	3018
Tapes (TB)	1150	4626	1072	3564

Report of the Computing Resources Scrutiny Group 2009

ATLAS

In the 2008 scrutiny, the CRSG reviewers were able to drill down in detail to confirm that the basic elements of the ATLAS 2008 request were understandable and that the requested (modest) growth at that time reflected identifiable developments derived from improved experience-based performance data.

However, we warned the C-RRB that it would not be appropriate to harm experiments with a more consolidated CM in favour of those whose CM is less defined or consolidated at this stage.

This year, our understanding, after the proposed changes in the ATLAS CM is more qualitative at this point. Based on the information provided and given the uncertainties in the model and the input parameters used by ATLAS, the latest requests are not understood at the same level of detail and depth as was possible last year. In particular, exploring the sensitivity of the model to changes in parameters, needed for a comprehensive requirements analysis, was not possible and it is not feasible at this point to provide quantitative estimates.

However, there is no reason to believe that the trend common to all experiments, in particular the ability of the resources shifted by one year to cover contingencies, should be different for ATLAS in view of the substantial reduction in beam time.

Report of the Computing Resources Scrutiny Group 2009

- 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 possible delay.
- It is important to inform the Tier1 and Tier2 computing centres of the resources required for calendar year 2011 as soon as possible. We would suggest the C-RRB to modify the timing of their meetings to facilitate the procurement process.
- 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 information provided to us by ATLAS and CMS about their AA program remains all too sketchy.
- As bunching conditions may vary in the future the collaborations should be aware that this has to be accommodated within the existing resources by decreasing the event rate or similar measures.
- 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 no longer required.
- We recommend the experiments make maximal use of the distributed resources in the GRID 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 and those used to perform physics analysis by the CERN based physicists.

There is no contingency for late delivery or failure to meet the pledges included in our estimates.

Report of the Computing Resources Scrutiny Group 2009

Our scrutiny has identified some new issues that need to be brought to the attention of the LHCC. Below we provide a list of them along with some older ones that are still pending

- The LHCC recommend that ALICE undertakes a full assessment of how their physics reach might be affected by requested computing resources not materializing. We are waiting for this statement.
- 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 the experiments with the largest computing needs 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 of concern. Strong overlap between different data sets as claimed by some 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 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.
- We recommend that the issue of the data analysis strategies is reviewed by the LHCC very soon, in order to ensure both, a reliable start-up and a coherent long term strategy. In particular it is recommended that the strategy for official reprocessing as a part of the computing model are better specified and then reviewed by the LHCC
- We recommend that a clearly defined sharing among the experiments of the resources installed at CERN is considered, in particular for the analysis facilities, such that an optimal usage of these resources is achieved.

Report of the Computing Resources Scrutiny Group 2009

In the process of scrutinizing the 2009 and 2010 requests of the four LHC experiments we have critically examined all possible aspects of the different computing models and their implementation. This scrutiny is a preliminary one.

While, apart from some potentially troublesome issues, the 2008 report substantially agreed with the requests, we have seen unexpected changes in this year's request. Our scrutiny recommends taking into account only part of them.

Care has to be taken not to harm experiments with a more consolidated CM in favour of those whose CM is less defined or consolidated at this stage.

The scrutiny after the first round of real data will be of great relevance.

The CRSG believes that the existing computing models have largely proven their validity and we have no doubt that they will survive their first contact with real data in 2009. There is no reason for radical changes at this point without hard experimental evidence.

The C-RRB should have the flexibility to react to unexpected changes quickly and eventually commit more resources if they are really needed.