CERN–RRB–2013–008
April 2013

Computing Resources Scrutiny Group

T Cass (CERN), J Flynn (UK, chairman), M Gasthuber (Germany), D Groep (Netherlands),
G Lamanna (France), D Lucchesi (Italy), H Meinhard (CERN, scientific secretary), T Schalk (USA),
M Vetterli (Canada), B Vinter (Nordic countries)

1 Introduction

This report covers the use of computing resources by the four main LHC experiments, ALICE, ATLAS, CMS and LHCb, during the 2012 LHC run and those experiments' resources requests for 2014, with a preliminary discussion of requests for 2015 and beyond.

The long LHC run from April 2012 to February 2013 has been a great success with the announcement of the discovery of a Higgs-like boson on 4 July [1, 2] and many other exciting results from all four experiments. The LHC's instantaneous and integrated luminosity far exceeded expectations. The CRSG adds its praise for the outstanding performance of the LHC, of the four experiments and, particularly in this context, of the WLCG.

The LHC is now in its first long shutdown, LS1, extending through 2013 and 2014. The accelerator is scheduled to restart in April 2015 at centre-of-mass energy 13 TeV. The long shutdown and following higher energy run mean that the resource requirements for 2014 should be considered in conjunction with those for 2015, when significant increases are anticipated. Thus, although this report contains a detailed scrutiny of the 2014 requests, the 2015 requirements are kept in mind. The Scrutiny Group is prepared to conduct a more detailed review of the 2015 requests in time for the October RRBs, rather than waiting for next April, in order to give the funding agencies as much advance information as possible on the 2015 requirements.

The next sections of the report contain our narrative discussion. They are followed by a part on the overall and experiment-by-experiment usage of WLCG resources in 2012, and then a part on the experiment-by-experiment resources requests for 2014 (with some attention to anticipated requests for 2015).

2 LHC running conditions

The LHC centre of mass energy was increased to 8 TeV for 2012 pp running in order to access slightly larger cross sections for interesting processes, especially the Higgs search. Following the July announcements, the pp run was extended into December with the pPb run shifted to January and February 2013. The pp luminosity reached $7.7 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, more than twice the maximum reached in 2011. At the very end of the pp run the bunch spacing was successfully reduced to the design value of 25 ns from the 50 ns used hitherto. Table 1 shows the LHC stable beam time for 2012 and the expectation for 2015.

Table 2 shows the effective event rates for the experiments in 2012, which are above those foreseen in the original computing models. The LHCC has in the past accepted increases in trigger rates for extending physics reach, but has warned against sustained substantial increases on the grounds that computing resources could likely not be increased to match. The CRSG endorses this view.

RRB year	start	end	pp	HI
2012	April 2012	March 2013	6.6	0.7
2013	April 2013	March 2014	—	—
2014	April 2014	March 2015	—	—
2015	April 2015	March 2016	5.2	0.7

Table 1 Live time in Ms for LHC running in the 2012 resource year and expectations for years 2013 to 2015. For 2012 the values are the delivered live time, using data from the LHC Programme Coordination web pages [3].

ALICE	pp 180 Hz, HI 280 Hz	From recorded event numbers and LHC live time
ATLAS	prompt 450 Hz, delayed 140 Hz	From recorded event numbers and LHC live time
CMS	prompt 450 Hz, delayed 360 Hz	From monthly average trigger rates (prompt includes 25% dataset overlap)
LHCb	4 kHz plus 1 kHz deferred	From October 2012 submission to CRSG [4]

Table 2 Average event rates for LHC running in 2012. ATLAS and CMS delayed data is recorded for later reconstruction. LHCb defers (buffers) events in the HLT for processing during inter-fill periods, raising the effective rate.

The LHC is expected to run at centre of mass energy 13 TeV when it restarts in 2015, where cross sections for interesting physics processes are factors of two or more above those at 8 TeV. The running conditions will be more challenging and will depend on the bunch-spacing adopted; pileup is bigger at 50 ns spacing, while out-of-time events are more common at 25 ns spacing. This coupled with the desire of the experiments to increase trigger rates to extract the most physics, means that substantial increases in resource requests are expected. We comment more on this below.

3 Overall assessment

The WLCG resources are being intensively used and it is clear that the experiments have made substantial efforts to optimise their use of these resources. With experience gained in the first run of the LHC, the number of reprocessings has decreased, the numbers of copies of data are being reduced and the types of data kept on tape have been cut. Software improvements have reduced cpu usage per event.

The hierarchical distinction of which tasks run where, evident in the original computing models, is becoming less rigid. Simulation tasks can be floated up to higher-level (lower tier number) while analysis and even reconstruction tasks can be moved to lower levels. At an individual site, however, there are optimum CPU to disk ratios for different tasks, and tasks needing tape access may best be run at the T1 holding the data, or a T2 closely associated with it. As network connections continue to improve we can expect this trend to continue. This does not imply that the distinction between Tiers in terms of hardware (eg provision of mass storage or not) and its management is dissolving.

We recognise the work by ATLAS, CMS and LHCb to use their online farms (HLTs) for offline tasks during LS1, mitigating the growth in resource requirements. This use should continue during shutdown periods once the LHC restarts. CMS anticipates being able to extend this even to periods between LHC fills.

The experiments have benefitted from opportunistic use of resources from outside the WLCG, in some cases very substantial. We caution against assuming the continuing availability of these.

The experiments will be reviewing their computing models and strategies over Summer 2013 and are undertaking software development programmes to reduce processing times with faster algorithms, code improvements (eg faster linear algebra libraries, smaller memory consumption) and adaptation to changing computer architectures. Improvements here have lasting benefits for future resource usage and the CRSG hopes that these efforts can be effectively supported.

ALICE is keeping disk and tape requirements flat for 2014 compared to 2013, with CPU power increases below 10% at T0 and T1. Larger jumps are anticipated between 2014 and 2015, notably in tape storage. Starting from 2015 ALICE aims to collect $1\,\mathrm{nb}^{-1}$ worth of PbPb events before the second long shutdown.

To further study the standard-model-like Higgs boson and intensify searches at the energy frontier, ATLAS and CMS are proposing trigger rates around 1 kHZ in 2015. This, combined with the more challenging running conditions expected (more pileup, out-of-time events), leads to some substantial jumps in resource requirements, up to factors of 2 in CPU power, by 2015. These increases already assume that problems like out-of-time events can be solved and take into account anticipated improvements in software.

ATLAS' request for CPU power in 2014 is stable at T0, with around 20% increase for the sum of T1 and T2 (the availability of the HLT farm for simulation and reconstruction helps here). Total disk requests grow by 7% from 2013 to 2014, while total tape grows by 25%. Larger increases are anticipated for 2015, the largest of which is a doubling of CPU power at T0 by the end of 2015.

The CMS computing requests do not change in 2013 and 2014 except for a 14% increase in the T2 CPU and an 11% increase in T1 tape. The use of the HLT for simulation provides substantial extra CPU resources, mitigating increased requests. However, substantial increases, notably doublings in CPU power at T0 and T1 are anticipated for 2015.

The LHCb resources request shows a step of 14% in CPU power between 2013 and 2014, 43% in disk storage and 15% in tape storage. LHCb has demonstrated successful running with a 5 kHz trigger rate and proposes to start with this rate in 2015, but does not request increased CPU power over 2014; the headline change for 2015 is a near-doubling in tape capacity.

All the experiments anticipate being more certain about requests for 2015 by October, when they should have more reliable estimates of computing model parameters like event sizes and CPU times, and when they will also have been subject to further review by the LHCC (including on the question of trigger rates). For now we note that substantial factors in resources requested for 2015 compared to 2013 may well survive scrutiny and that it may be possible to meet these if a flat spending profile can be maintained in 2013 and 2014.

4 Recommendations and comments

1. There is a strong expectation that reasonable exploitation of the physics potential of the LHC and the experiments from 2015 will require significantly increased computing resources. To reach these levels looks achievable if stable funding at the 2012 level can be maintained in 2013 and 2014.

2. ALICE and LHCb's scrutinised requests have not been fully met by the Tier 1 centres supporting them. It even looks impossible in principle for 100% of the scrutinised requirements recommended by the CRSG to be met. We think this problem should be addressed urgently.

3. The CRSG is willing to work with WLCG management to review the request/review/pledge process in light of experience over the first period of LHC running, with the hope of better matching to funding and procurement cycles (and streamlining where possible).

4. Efforts to improve the efficiency of software are essential in view of the level of 2015 requests. Indeed, the resulting gains are already being assumed in making those requests. The CRSG strongly supports these software engineering efforts and hopes that sufficient effort can be funded to support them.

5. An agreed efficiency factor of 0.7 has historically been allowed when calculating disk space requirements. In practice disks are used more efficiently than this and ATLAS now prefers to dispense with the efficiency factor and make its own explicit calculation of disk space. The CRSG welcomes a change in the way disk requirements are calculated, to reflect more efficient use, provided the ability to compare experiments' requests on an equal footing is maintained.

6. The effectiveness of disk usage is only partly captured by disk occupancy figures. A metric which also takes account of frequency of use would be highly desirable.

7. The CRSG welcomes the effort to produce a new documentary description of the experiments' computing models after their evolution in the face of real data. We look forward to the results of this exercise, which are expected before the next RRB.

5 Interactions with the experiments

The experiments submitted usage reports and resources requests in time for or very soon after our 1 March deadline [9–16]. The CRSG thanks the experiments for their detailed documents and acknowledges in particular the computing representatives for their availability. The Scrutiny Group exchanged questions with the experiments and held at least one face-to-face meeting with each.

As agreed some time ago with the ATLAS and CMS managements, a single team of CRSG referees scrutinises the ATLAS and CMS reports and requests to ensure a consistent approach.

For the October 2013 RRB we ask the experiments to submit their reports and requests by 1 September 2013.

6 CRSG membership

D Espriu (Spain) stood down as chairman of the CRSG following the October 2012 RRB and has been succeeded by J Flynn (UK). After a gap of around a year, M Vetterli has replaced W Trischuk as Canadian representative. A replacement Spanish representative has been nominated.

Two long-standing members of the group, B Vinter (Nordic countries) and D Groep (Netherlands) will be standing down after the April 2013 scrutiny round and will need replacements.

The chairman thanks D Espriu for his long period of service; Espriu had chaired the group since its inception late in 2007. The chairman thanks all CRSG members for their dedication and the experiments' representatives for their collaboration with us. Thanks are also due to the CERN management for support and to H Meinhard (CERN) for ensuring smooth running of the group.

WLCG resource usage in 2012

7 Overall usage

Table 3 shows the use of different resources in the calendar year 2012, with the 2011 table reproduced for comparison. The figures show a growth in the use of pledged resources, especially tape, and the availability of resources beyond the pledges (which have been of great benefit to the experiments). The CRSG would welcome WLCG/EGI reporting of Tier 2 disk use.

Table 4 compares the efficiency of CPU use at Tier 2 by experiment and between 2012 and 2011. ALICE continues to lag the other experiments. However, the ALICE efficiency improved at the end of 2012 and it has averaged above 80% at all tiers between December 2012 and February 2013.

Disk usage is reported here by fractional occupation at a fixed point in time, or by time-integrated occupancy divided by time-integrated capacity over a time interval. A figure of merit which also takes account of frequency of access would be valuable. The experiments already record such 'popularity' information and use it to guide their data replication and cleanup policies. The CRSG would welcome progress towards a common synthetic metric which could be agreed upon for use in future scrutinies.

In table 5 we show the division of WLCG resources at CERN plus Tier 1 among the experiments (the first three columns sum to 100% up to rounding effects). The final column indicates what fraction of total CPU use by an experiment has been at CERN (hence the column values need not sum to 100%). Table 6 shows the division between experiments of CPU consumption at Tier 2. The similarity

		2012 used/pledged				2011 used/pledged	
		average	end of year			average	end of year
CPU	CERN	61%	—	CPU	CERN	55%	—
	T1	116%	—		T1	93%	—
	T2	171%	—		T2	166%	—
Disk	CERN	104%	111%	Disk	CERN	105%	119%
	T1	135%	141%		T1	121%	137%
	T2	—	—		T2	—	—
Tape	CERN	88%	101%	Tape	CERN	75%	97%
	T1	71%	86%		T1	47%	51%

Table 3 Usage summary for calendar year 2012 (left). Averages are calculated using time-integrated CPU power or storage capacity over the year. End of year disk and tape percentages are from capacities. We do not have T2 disk usage information. Data is from T1 and T2 accounting summaries in the CERN WLCG document repository [5]. On the right we reproduce the analogous table for 2011 from the April 2012 CRSG report [8].

	2012	2011
ALICE	64%	60%
ATLAS	88%	88%
CMS	83%	82%
LHCb	95%	98%

Table 4 T2 CPU efficiency by experiment in 2012 and 2011 calendar years. Efficiency is defined as sum of CPU time divided by sum of wall time. Data from EGI Accounting Portal [6] with udpate time 2013-03-20 12:33 UTC.

End of 2012

	CPU	Disk	Tape	% CPU at CERN
ALICE	12%	13%	10%	50%
ATLAS	55%	44%	39%	14%
CMS	24%	33%	42%	30%
LHCb	10%	10%	10%	22%

End of 2011

	CPU	Disk	Tape	% CPU at CERN
ALICE	15%	14%	12%	52%
ATLAS	51%	46%	39%	17%
CMS	23%	33%	41%	21%
LHCb	11%	7%	8%	26%

Table 5 Use of resources at CERN plus T1 by the experiments at the end of 2012 (left) and end of 2011 (right). The first three columns are the fractions of the total CPU (time-integrated), disk (at end of year) and mass storage (at end of year) used by each experiment at CERN and all the T1 sites. The final column is the fraction of the experiment's total CPU consumption which has been at CERN (hence the column need not sum to 100%). Data is from the master accounting summary in the WLCG document repository [5].

	2012	2011
ALICE	7%	9%
ATLAS	53%	53%
CMS	35%	30%
LHCb	5%	7%

Table 6 Distribution of total CPU consumption (time-integrated CPU power) by the experiments at Tier 2 in 2012 and 2011. Data from EGI Accounting Portal [6] with update time 2013-03-20 12:33 UTC.

		installed/pledged
CPU	CERN	100%
	T1	111%
	T2	169%†

		installed/pledged
Disk	CERN	85%
	T1	113%
	T2	—

		installed/pledged
Tape	CERN	100%
	T1	100%

Table 7 Fulfilment of pledges. The table reports the situation at the end of 2012. Data from the master accounting summary in the WLCG document repository [5]. †: the T2 CPU percentage is delivered/pledged for December 2012 from the Tier 2 reports in [5].

between 2012 and 2011 indicates some stability in the implementation of the computing models. ALICE continues to use the largest fraction of CPU at CERN of all the experiments.

8 Usage by the individual experimental collaborations

8.1 ALICE

Our report uses the documents provided by ALICE, [9] and [10], and takes into consideration exchanges of information and meetings with the ALICE computing representatives, Yves Schutz, Predrag Buncic and Latchezar Betev.

Table 8 summarises ALICE's resource usage. The ALICE requirements in 2012 for CPU, disk and tape (to a lesser extent) have been covered by the pledged resources (although after accepting the request by the CRSG to adjust the requirements to the anticipated pledges).

The 2013 pledges by Tier 1s do not cover the requirements and are lower than the pledges in 2012 (see table 8). This situation results from a substantial reduction of pledged resources by a few major sites: the ALICE Collaboration has experienced discrepancies among funding agencies in respecting the collaboration's constitution on 'fair shares' (but acknowledges that some have contributed much more than expected).

The collaboration is especially troubled by the reduction in disk space from 2012 to 2013 at the Tier 1s and has asked the CRSG to endorse a request to converting some of the tape pledges to disk in Tier 1s, thus compensating for the shortfall. Analysis using derived data (ESD and AOD) on tape suffers high latency in data retrieval and this motivates the desire to have all ESDs and AODs resident on disk (analysis using the smaller AOD format is encouraged). However, when one accounts for the additional Tier 1 disk used for tape buffering, the disk shortfall is less obvious (the disk buffers are sized to store the largest raw data set to be collected in the resource year, usually from the heavy-ion run; the total T1 buffer is divided among the sites in proportion to their share of raw data storage). What is more evident is that only about 50% (30%) of pledged tape resources have been used in Tier 0 (Tier 1), since simulation data is no longer saved on tape and only raw data meeting a set of 'good physics' criteria is shipped to Tier 1 sites.

The CRSG discussed with ALICE possible scenarios to optimize the use of resources:

- The possibility to exploit disks available in Tier 2 would imply a modification of their computing model, moving out some raw data reconstruction tasks from Tier 1s. This is problematic because of latency in data transfer, but the situation may improve once the LHCONE network is in place.

Resource	Site(s)	2012 pledged	2013 pledged	2013 needed	2012 used
CPU (kHS06)	T0	90	90	126	78
	T1	113	101	120	83
	T2	178	188	145	189
Disk (PB)	T0	8.1	8.1	11.0	8.1
	T1	8.1	7.7	10.8	9.0
	T2	10.9	12.8	15.8	10.1
Tape (PB)	T0	20.0	22.8	22.8	10.2
	T1	12.0	14.1	21.0	4.6

Table 8 Summary of resource usage by ALICE in 2012 (April 2012 to March 2013), with 2013 pledges and 2013 needs (as in October 2012). The Tier 0 disk takes into account a 2.9 PB buffer for the tape system and the Tier 1 disk takes into account similar buffer space of 2.35 PB.

- To rely more on using Tier 1 tape systems by optimizing the logical data model for data retrieval (minimizing the frequency of staging). However the staging is already optimized and the speed and efficiency of tape robot systems is not homogeneous among Tier 1s.
- Reducing the CPU needs estimated for 2013 once the most recent improvement in CPU efficiency (increase of about 20%) is taken into account. Actually this gain in efficiency (due mainly to data caching by ROOT `TTreeCache`) has more effect on analysis since the main reconstruction tasks at Tier 1s were already more efficient.

8.2 ATLAS

As can be seen in table 9 ATLAS has continued to be very successful in using resources beyond pledges at ATLAS sites and with opportunistic use of resources otherwise idle to support the large simulation requests from their physics groups. They have generated 80% more full simulation and 45% more fast simulation events than predicted in the fall of last year. They have implemented improvements to the generation and digitization code for a 15–40% reduction in needed CPU. They have also demonstrated the ability to run simulation on cloud resources and their HLT farm. This code improvement work is an ongoing activity complete with milestones.

In 2012 ATLAS averaged a prompt trigger rate of 400 Hz along with a 150 Hz delayed trigger. They observed an average pileup of around 21 and recorded 3B events (+ 0.9B events from the delayed trigger). The pileup was lower than expected which resulted in reduced CPU time for event reconstruction.

Disk usage numbers do not include the 70% efficiency factor which is used for the requests. This generates some confusion when comparing used versus pledged. Since the disk usage efficiency is close to 1, ATLAS and WLCG agreed to set the factor to 1 for 2013. For the 2012 disk used reported in table 9, the numbers should be multiplied by the factor 0.7 in order to compare quantities consistently.

The Tier 1 usage was 100% showing up in real ESD+AOD+DPD data space, group data and some increase in simulation data. The disk-resident copy of the full raw data was essentially removed after the end of pp running to help with space issues. The disk space at the T1's and T2's is actively managed with 10–20% used as buffer space for the dynamic replication software.

Resource	Site	Pledged	Used	Used/Pledged	Average CPU efficiency
CPU (kHS06)	T0+CAF	111	111	100%	89.8%
	T1	285	420	147%	92.1%
	T2	332	634	191%	88.2%
Disk (PB)	T0+CAF	9	10	111%	
	T1	30	47	157%	
	T2	45	52	116%	
Tape (PB)	T0+CAF	18	29	161%	
	T1	38	31	82 %	

Table 9 Fulfilment of pledges. The table reports the ATLAS situation at the end of the 2012 RRB year. CPU usage is time-averaged capacity March 2012 to February 2013; disk and tape usage are volumes occupied at the end of February 2013. 'Used' data from ATLAS.

8.3 CMS

CMS resource use for 2012 is shown in table 10. Estimates used for planning proved to be generally accurate with the exception of the raw and reco record sizes being smaller than expected (35% and 85%). The increased resource need generated by the longer running time was compensated for by reducing the 'parked' data triggers and by a smaller than expected average event pileup. In 2012 CMS ran with a 350–435 Hz prompt trigger rate and a 300–600 Hz parked event rate with a spike in December of 734 Hz for prompt triggers (540 Hz parked).

The full 2012 data sample is planned to have a reprocessing pass completed by spring of 2013. The CMS reconstruction time for 2012 averaged 20 s/event. The Tier 1 centers were used for substantial simulation production during 2012 thus freeing up Tier 2 resources for analysis. A total of 7B simulation events were generated during 2012 mostly at a centre-of-mass energy of 8 TeV. The T2 sites averaged pledged resource use for the year of 134% with an average CPU use of 83%. CMS analysis jobs continue to concentrate on AOD usage. The average CPU efficiency at the T1's was measured to be 88% averaged over the year.

For several years CMS has successfully used a full mesh scheme for data transfer among all sites. They currently have around 3300 commissioned links supporting this operation with about 2/3 of the total traffic T2 ↔ T2 and T1 ↔ T2.

Resource	Site	Pledged	Used	Used/Pledged	Average CPU efficiency
CPU (kHS06)	T0+CAF	121	75	62%	90%
	T1	137	142	104%	88%
	T2	320	429	134%	83%
Disk (PB)	T0+CAF	7	6	84%	
	T1	21	21.5	102%	
	T2	27	25	93%	
Tape (PB)	T0+CAF	23	22	96%	
	T1	47	42	89%	

Table 10 Fulfilment of pledges. The table reports the CMS situation at the end of 2012. Data from WLCG accounting.

8.4 LHCb

The LHCb usage report is based on the document supplied by LHCb [15], email exchanges and an informative face to face meeting.

The resource usage in 2012, table 11, shows the expected profile and scale compared to the previous period. The events/luminosity taken in 2012 are twice that of 2011, having in general the same resource usage per event as in 2011, resulting in roughly twice the resource usage. The resource numbers from the tables of 2011 are in good agreement with that scaling.

Compared to the initial planning for 2012, the LHC runtime extension and new physics studies resulted in about 40% more raw data being generated, processed and stored (with corresponding increases in the quantity of derived data). LHCb took several measures to adapt the computing model to this situation, especially by further reducing copies on disk and tape. The extension increased the resource pressure, especially for tapes (with about 3 PB shortfall at T1s) which was mitigated at the end of the period by some generous T1s providing their 2013 pledges ahead of schedule.

Resource	Site(s)	2012 pledge	2012 used
CPU (kHS06)	T0	34	17.5 (26)
	T1	110	64.3(115)
	T2	48	45.1 (94)
	other		13.0 (30)
Disk (PB)	T0	3.5	2.6
	T1	7.5	5.6
Tape (PB)	T0	6.4	5.2
	T1	5.5	8.1

Table 11 2012 LHCb usage table. Disk is T0D1 class only, no buffer/scratch included. CPU is CPU power in kHS06 averaged over one year, whereas pledges take into account peak demands. Peak use of CPU power is shown in parentheses. CPU use includes a contribution from non-WLCG resources.

Thanks to the very flexible computing model (and the actual software framework), LHCb was able to use T2 and other CPU resources, like the online farm, with very little additional effort. The situation for the peak demands on CPU resources will have less impact in 2013 and 2014, because the main driver behind the peak usage (reprocessing after data taking and before conferences) is not present during LS1, resulting in a nearly flat CPU usage profile (assuming it is not limited by lack of free storage resources for the generated data).

The situation for disk space has continued to be fraught throughout the period, induced by increased demands from more raw data, late implementation of pledges by some T1s and share reduction by others. The demonstrated CPU and job efficiency was, similar to 2011, at a stable and high level (> 93%). The overall data processing (prompt reconstruction, reprocessing, skimming, MC production and user analysis) was completed in very good agreement with the time schedule anticipated at the beginning of the period and allows for continuation of 2013 processing on schedule.

Resource requirements for 2014

9 Resource requirements from the experiments

9.1 ALICE

A summary of 2013, 2014 and 2015 requirements is in table 12. We anticipate revised estimates for 2015 for the October RRB meeting and hence postpone detailed scrutiny of these.

The total ALICE archive since 2010 corresponds to: 10.2 PB raw data and calibration data stored on the Tier 0 tape system; 4.6 PB raw events passing physics selection replicated once to the Tier 1 tape systems; 18 PB data processed and archived plus the corresponding Monte Carlo data (obsolete data have been permanently deleted).

Storage expectations in 2013-2014 are:

- 9.2 PB of data will be generated by the 2013-2014 productions and will be available on disk, including a replication factor of 2 (distributed as 30% in Tier0, 30% in Tier1s and 39% in Tier2s).
- 11 PB in total of existing data on disk will be kept after reducing the current 22 PB by 50% by removing replicas.

This gives a total of 20.2 PB disk use in 2013 and 2014 to be distributed across Tier 1 and Tier 2 sites, with the additional total disk buffer space of 2.35 PB to be added at Tier 1. The Tier 0 disk is maintained at essentially its 2012 level.

During LS1 a major upgrade of the whole offline environment for reconstruction, analysis and simulation is foreseen to improve the software quality and performance. New developments from the R&D programme, including parallelization, vectorization, exploitation of GPUs and new algorithms, will be implemented. The collaboration does not expect the new environment to become available before the end of 2014.

In conclusion

- There is a request to trade tape for disk in 2013. If possible this strategy could help meet the disk storage increase required in 2015. We endorse this, but note a concern that the exchange

Resource	Site(s)	2013 needed	2014 needed	2015 needed
CPU (kHS06)	T0	126	135	190
	T1	101	110	110
	T2	188	190	200
Disk (PB)	T0	8.3	8.3	10.8
	T1	10.1	10.1	13.6
	T2	12.8	12.8	16.1
Tape (PB)	T0	12.0	12.0	27.0
	T1	6.0	6.0	21.0

Table 12 Summary of ALICE resource requests for 2013, 2014 and 2015. The 2013 requests are modified since October 2012 to take into account the current pledges. CERN CAF resources are included in T0.

rate for disk-to-tape conversion means a reduction in total storage capacity, with potential future implications.

- We are encouraged to see better CPU efficiency at the end of 2012. We think that software engineering and other efforts to improve the implementation of the computer model can have great benefit and strongly support the efforts to upgrade the offline environment during LS1.
- The CRSG has previously asked ALICE to reduce its resources request to match expected levels of support. The collaboration has responded to this. In return, we hope that funding agencies can, at least in principle, have the possibility to fully-fund the scrutinised requests.

9.2 ATLAS

Table 13 shows the latest ATLAS computing requests for 2013 and 2014 along with the current resource pledges and the CRSG recommendations. ATLAS projects the need for a substantial increase in the prompt trigger rate for 2015 (1 kHz). This would imply a significant increase in computing resources needed in 2015. Since there will be revised estimates at the October RRB meeting and information coming from the LHCC on the physics need for full analysis of the proposed 1 kHz trigger, we will leave most of the discussion of the 2015 request for this fall's scrutiny.

During the 2012 running ATLAS logged a 150 Hz 'parked' trigger. This data is now being processed and it would be instructive to learn how heavily it is being used.

During 2013 and 2014 ATLAS anticipates improvements and adjustments to their resource model as well as substantial software architecture and speed improvements. These activities will be tracked, with milestones, by the software group. At present ATLAS is in the prototyping and testing stage for some of these changes. The reconstruction code has a two- to three-times speed-up goal using targeted algorithms, auto vectorization and modern efficient linear algebra libraries. Work on multi-process, event-level parallelism should also help with the memory footprint. The new multi-threaded framework is a joint activity with other LHC experiments. One critical reconstruction improvement is to make the code much less sensitive to out-of-time pileup to avoid a problem when the machine runs with 25 ns bunch spacing. Similar improvement is planned for the simulation code which is estimated to give a 10% speed up/year. Attention will also be given to the size of the heavily used AOD records, which are projected to remain at the 2012 size even with 2015 data, and to the optimization of the physics group and user analysis workflow.

This software activity is critical to constrain the overall computing resources needed to process any increase in the trigger rate. The CRSG strongly advocates support for the FTE needed for these

Resource	Site	2013 ATLAS	2013 CRSG	2014 ATLAS	2014 CRSG
CPU(kHS06)	T0+CAF	111(111)	111	111(111)	111
	T1	316(333)	319	373(327)	355
	T2	360(396)	350	408(399)	390
Disk(PB)	T0+CAF	10(10)	11	11(11)	11
	T1	35(36)	33	36(33)	33
	T2	51(49)	49	56(47)	49
Tape(PB)	T0+CAF	25(27)	23	31(31)	27
	T1	42(41)	40	53(43)	44

Table 13 ATLAS resources request. The numbers in parentheses () are pledged resources and the table contains our revised recommendations for 2014.

upgrades.

A new scalable data management system is being developed to handle data replication, deletion and file ownership. This should address the optimization of space usage. We encourage ATLAS to consider including the historically preplaced data under this management as well. This data distribution plan depends critically on adequate disk buffer space for the dynamically-placed transient copies of data at the T1s and T2s. These are explicit in the ATLAS request. A full data and simulation reprocessing will be done in 2014 for data consistency and long term archival purposes as well as a large sample of simulation at the new beam energy. Overall ATLAS projects the need for the same level of simulation/real data in the 2013–2015 requests but with a larger fraction using the new fast simulation code.

ATLAS, like CMS, plan to use their HLT farm for other work in 2013 and 2014. ATLAS need one or two days to set up the farm for non-HLT purposes, comprising up to six hours reconfiguring the farm itself for Grid operations and added time arising from trigger/DAQ operational procedures. CMS in contrast are much faster, planning even to use the time between fills when the LHC is running. We encourage ATLAS to review their activities in this area and to explore the option of using the HLT farm to augment Tier 1 processing capacity during LHC periods without beam in 2015 and beyond.

9.3 CMS

Table 14 shows the latest 2013 and 2014 CMS resource requests, the resource pledges and the CRSG recommendations. 2013 and 2014 are the 2 years of the LS1. Approximately 40% of the triggers CMS collected during 2012 were parked data. These are anticipated to be available to users by this summer. It would be useful to know how heavily these parked data are actually used.

CMS projects the need for a substantial increase in the prompt trigger rate for 2015 (1 kHz). This would imply a significant increase computing resources needed in 2015 so CMS has proposed that resources for 2014 and 2015 be considered together. Since there will be revised estimates for the October RRB meeting and information coming from the LHCC on the physics need for full analysis of the proposed 1 kHz trigger, we will leave most of the discussion of the 2015 request for this fall's scrutiny.

The 2014 resource request is anticipated to allow for a complete reprocessing pass on all data and simulation for consistent data treatment and legacy archive. A large simulation sample will also be produced for the increased energy running in 2015. The computing requests do not increase in 2013 and 2014 except for a 14% increase in the T2 CPU and an 11% increase in T1 tape. During 2014

Resource	Site	2013 CMS	2013 CRSG	2014 CMS	2014 CRSG
CPU(kHS06)	T0+CAF	121(121)	121	121(121)	121
	T1	165(150)	165	165(150)	175
	T2	350(400)	350	400(390)	390
Disk(PB)	T0+CAF	7(7)	7	7(7)	7
	T1	26(23)	26	26(22)	26
	T2	26(29)	26	27(29)	27
Tape(PB)	T0+CAF	26(26)	26	26(25)	26
	T1	50(48)	50	56(45)	55

Table 14 CMS resources request. The numbers in parentheses () are pledged resources and the table contains our revised recommendations for 2014.

CMS anticipate changing the Tier 0 setup to be more similar to that of the Tier 1 sites. This will allow for substantial prompt reconstruction at the T1s in 2015.

We note that CMS plans to use their HLT farm (40% of the T1 capacity) to augment their Tier 1 data processing capacity during 2013 and 2014 and note the advanced state of work to exploit virtualisation techniques to improve their ability to dynamically reconfigure this farm. We encourage CMS to further develop their work in this area and explore the option of using the HLT farm to augment Tier 1 processing capacity during LHC periods without beam in 2015 and beyond.

2013/14 will see substantial work on the production software including an effort to substantially reduce the impact of out-of-time pileup on the reconstruction time. CMS is in process of setting milestones for a new multicore framework and other code improvements in the 2013/14 time period.

9.4 LHCb

The LHCb computing resource requests for 2014 are based on the fact that there is no data taking in 2013 and thus few unknowns in the 2014 requests. There is a large amount of data to be processed and reprocessed. The resources used in processing the 2011 and 2012 data during 2012 gives a very solid base for the 2013 and 2014 requests.

This scrutiny is based on resource request document provided by LHCb [16]. The CRSG referees had fruitful email exchanges and a meeting and would like to thank the LHCb colleagues (namely R Graciani, M Cattaneo and Ph Charpentier) for their collaborative attitude, timely and detailed responses.

The request for 2014 computing resources, separated by tier level, is summarized in table 15. The total CPU request is 197 kHS06 · yr at T0/1/2, rising to 237 kHS06 · yr once resources from outside WLCG are included.

The overall situation for the 2014 request is affected by the following conditions:

- There is little margin for reduction since the data is all acquired and LHCb have already limited their resource use to the extent possible. In fact without exceeding the request they would fail in storing their existing data.

Resource	Site(s)	2014 request
CPU (kHS06)	T0	40
	T1	110
	T2	47
	unpledged + HLT	10+30
Disk (PB)	T0	6.4
	T1	14.0
Tape (PB)	T0	6.6
	T1	11.1

Table 15 LHCb resources request for 2014. The table shows the resources needed to process the backlog of data not yet processed and to reprocess the very large existing dataset. Storage needs are nearly doubled since last year, a logical consequence of the data taking in 2012. We remark that a nontrivial portion of the CPU power, 40 kHS06, is expected to come from outside the WLCG, although 75% of this is from LHCb's own HLT farm.

- LHCb are now in a state where they can do all processing through the Grid and thus are less sensitive in 2014 to the precise location of the resources, that is at T0 or T1 sites, provided that the T1 sites keep a reasonable amount of disk (a scenario with primary storage at CERN and processing at T1 sites is not feasible).

In conclusion, the resources request for 2014 is a highly realistic estimate. The data is acquired and the dataformats are stripped to the best possible size, except for the Monte Carlo simulated data where LHCb still expects to be able to find another 30% compression.

The referees support the ability of LHCb to efficiently use any CPU resource available, independent of the Tier level. This should help ensure that any resources that can be allocated for LHCb can be used efficiently. The use of unpledged resources is welcomed, and even expected, but dependence on these resources is now at a level where failing to obtain them will have significant consequences.

References

[1] ATLAS Collaboration, Physics Letters B716 (2012) 1–29

[2] CMS Collaboration, Physics Letters B716 (2012) 30–61

[3] LHC Programme Coordination web pages, http://lpc.web.cern.ch/lpc/

[4] LHCb Collaboration, R Graciani (Ed), LHCb Computing Resource Usage in 2012(I), LHCb–PUB–2012–013

[5] WLCG Document Repository, CERN, https://espace.cern.ch/WLCG-document-repository/

[6] EGI Accounting Portal, CESGA, Spain, http://accounting.egi.eu/

[7] REBUS Pledge Summary, CERN, http://wlcg-rebus.cern.ch/apps/pledges/summary/

[8] CRSG, T Cass et al, CERN–RRB–2012–045

[9] ALICE Collaboration, ALICE 2012 Computing Resources Usage, 28 February 2013

[10] ALICE Collaboration, Computing Resources Requirements 2013–2017, 6 March 2013

[11] ATLAS Collaboration, ATLAS Computing Resource Usage October 2012 – February 2013, 3 March 2013

[12] ATLAS Collaboration, ATLAS Computing Resource Requirements for Data-taking in 2013–2015, 3 March 2013

[13] CMS Collaboration, CMS Resources Utilization in 2012, 1 March 2013

[14] CMS Collaboration, CMS Computing Resource Request 2013/2014/2015, 1 March 2013

[15] LHCb Collaboration, Ph Charpentier (Ed), LHCb Computing Resource Usage in 2012(II), LHCb–PUB–2013–001, 11 February 2013

[16] LHCb Collaboration, M Cattaneo (Ed), LHCb Computing Resources: 2013 reassessment, 2014 request and 2015 forecast, LHC–PUB–2013–002, 26 February 2013