Report of the Computing Resources Scrutiny Group

CRSG current composition

C.Bozzi (Italy), T.Cass (CERN), G. Lamanna (France), D.Espriu (Spain, *Chairman*), J.Flynn (UK), M.Gasthuber (Germany), D.Groep (The Netherlands), T. Schalk (USA), W.Trischuk (Canada), B.Vinter (Nordic Grid), H.Meinhard (CERN/IT, *Scientific Secretary*)

Report of the Computing Resources Scrutiny Group

Contents of this report:

- Overall usage of the WLCG resources during 2011 (January to July/August). *[Introductory part of the report]*
- The use the experiments made of the committed resources *[Part A of the written report]*
- Final scrutiny of the experiments' requests for 2012 and 2013 *[Part B of the report]*
- Usage of the Tier 2, by country *[Part C of the report]*

In the written report CERN-RRB-2011-0187 the scrutiny of CPU@Tier 2 is incorrectly stated, please refer to these slides for the CRSG estimates.

*In preparation of the April 2012 C-RRB we ask the experimental collaborations to provide their documents by **1st MARCH 2012.***

Report of the Computing Resources Scrutiny Group

Live time: 30 days/month = **720** hours
Folding in efficiencies 720 x 0.7 x 0.4 = 201.6 effective hours/month = **725760** s/month

RRB year	RRB year start	RRB year end	Months (max) Data taking	Total live time (in Ms)	pp	AA
2011	**April '11**	**March '12**	**8**	**5.9**	**5.2**	**0.7**
2012	**April '12**	**March '13**	**8**	**5.9**	**5.2**	**0.7**

- 2.8 Ms of pp beam have been delivered to the experiments from March 1st to July 31st
  LHC average live time closer to 20% than to the nominal 28%
  Average luminosity close to 10^33
- > 70% of the maximum number of events expected for this period.
  This large number of recorded events has been possible thanks to experiments using all the available bandwidth and effectively recording events at rates larger than the nominal ones
- Pile up started lower than expected but it reaches now ~ 15 events/crossing at fill

Report of the Computing Resources Scrutiny Group

- While it was already clear in the April 2011 scrutiny that the experiments had been able to cope with rapidly changing running conditions, but the total volume of data in 2010 was limited and there was a moderate sense of urgency for physics results.

- The situation has now evolved to steady running conditions with a massive use of the available resources. Some aspects of the computing models such as large individual non-organized computing usage, format and distribution of the data sets, the flexibility to cope with increasingly challenging running conditions (pile up, bunch separation), and the urgency to reprocess and analyze large amounts of data in a short time have represented a real challenge for the computing models and for the WLCG as a whole.

- During the early months of 2011 the collaborations reconstructed and analyzed the events recorded during the first AA run.

Report of the Computing Resources Scrutiny Group

Scrutiny of the WLCG resources utilization in 2011

Report of the Computing Resources Scrutiny Group

WLCG resources and accounting for 2011:

http://lcg.web.cern.ch/

EGEE accounting portal at CESGA.ES:

http://www3.egee.cesga.es/

Reports provided by the four experiments to the CRSG. T2 usage compiled by Ian Fisk (with thanks)

Overall usage 2011 (Jan-July)

Resource	Site(s)	Used/Available [mean occupancy] *(Apr il 2011)*
CPU	**CERN**	**52 %** *(32 %)*
	T1	**83 %** *(62 %)*
	T2	**117 %** *(122 %)*
Disk	**CERN**	**99 [99] %** *(110 %)*
	T1	**116 [112] %** *(110 %)*
	T2	**Not available**
Tape	**CERN**	**75 [64] %** *(69 %)*
	T1	**43 [47] %** *(60 %)*

Report of the Computing Resources Scrutiny Group

Efficiency of the utilization of the CPU at Tier 2s in 2011 (left column) compared to 2010(right column)

ALICE	**50 %**	**73 %**
ATLAS	**89 %**	**85 %**
CMS	**80 %**	**66 %**
LHCb	**98 %**	**88 %**

*In view of these figures we recommended a revision of the official figure of **60%** assumed for Tier2 up to **67%** for 2012. In view of the excellent performance we recommend adopting a **70%** nominal efficiency for CPU@T2 in 2013.*

ALICE: The efficiency for user/chaotic analysis is extremely low (35%). The collaboration should take vigorous steps to increase this low figure.

Report of the Computing Resources Scrutiny Group

Percentage of use of the resources by experiment in 2011 (CERN+Tier 1s)

Collaboration	% of tape in T1+CERN used at end of period	% of disk in T1+CERN used at end of period	% of CPU in T1+CERN used	% of which at CERN (Oct 2010)
ALICE	11 %	13 %	13 %	59 % (33%)
ATLAS	41 %	47 %	52 %	18 % (14%)
CMS	42 %	32 %	23 %	18 % (20%)
LHCb	7 %	8 %	12 %	28 % (46%)

- The large collaborations seem to converge in the relative fraction of CERN resources they use.
- LHCb does now a reasonable fraction of their total computing at CERN. ALICE has increased enormously their dependence on CERN resources.

Report of the Computing Resources Scrutiny Group

Percentage of use of the resources by experiment in 2011 (Tier 2s)

Collaboration	% of total disk in T2 used at end of period	% of total CPU in T2 used in 2010 *(October 2010)*
ALICE	N/A	**6 %** ***(7 %)***
ATLAS	N/A	**54 %** ***(59 %)***
CMS	N/A	**31 %** ***(30 %)***
LHCb	N/A	**9 %** ***(4 %)***

Statistics show a marked stability and quite definite patterns.

Disk @ Tier 2 not centrally accounted yet.

Report of the Computing Resources Scrutiny Group

Delivered versus pledged

Resource	Site(s)	Available / pledged
CPU	CERN	100 %
	T1	99 %
	T2	117 % (*)
Disk	CERN	100 %
	T1	101 %
	T2	Not available
Tape	CERN	100 %
	T1	80 %

- (*) Pledged CPU. The large turnout in CPU at the Tier 2 indicates that the percentage installed is actually above 100%, and that the efficiency is large.

- PIC, RAL and TRIUMF have installed disk capacities well above their nominal pledges.

- Tier 1 have adapted to the low usage of tape.

Report of the Computing Resources Scrutiny Group

PART A

Usage by the experimental collaborations

Report of the Computing Resources Scrutiny Group

Generally speaking, the experiments' computing models have demonstrated their capability to record, distribute and analyze the substantial amounts of data delivered to them by the LHC during the current year.

The performance throughout the year has been regular and without any noticeable difficulties, with periods where usage has been quite intensive corresponding to simulation and reprocessing campaigns and intensive physics analysis before the seasonal conferences. Noticeable peaks of individual user analysis have been detected in the weeks preceding these events but the efficiency of the system remained high throughout (with one exception).

The data placement policy and the detailed computing activities have been different from what was envisaged in the computing models, in the direction of making the some of the computing models more hierarchical and organized.

On the contrary, the reprocessing policy is quickly converging to the one indicated in the computing models.

CPU resources are generally exceeding the experiments' needs at this point and the experimental collaborations have had substantial headroom that they have employed to increase simulation production. There is also some headroom for disk, but this is not automatic now because it requires good data management policies.

Some collaborations and sites have experienced inefficiencies associated with large memory footprints.

Report of the Computing Resources Scrutiny Group

ALICE

Resource	Site(s)	Pledge	Use	Used/Pledge	Average CPU efficiency
CPU/kHS06	T0+CAF	62	74 (82)	119 % (55 %)	47 %
	T1	71	54 (57)	76 % (82 %)	45 %
	T2	81	68 (78)	84 % (61 %)	50 %
Disk/PB	T0+CAF	6.1	3.6	59 % (47 %)	-
	T1	5.5	3.7	67 % (32 %)	-
	T2	7.3	N/A	--	-
Tape/PB	T0+CAF	6.8	5.9	87 % (48 %)	-
	T1	8	2	25 % (14 %)	-

Report of the Computing Resources Scrutiny Group

ATLAS

Resource	Site(s)	Pledged	Used	Used/ Pledged	Average CPU efficiency
CPU (kHS06)	T0+CAF	75	64	85 %	66 %
	T1	250	224	90 %	88 %
	T2	281	324	115 %	89 %
Disk (PB)	T0+CAF	7.0	4.9	70 %	-
	T1	27	22	82 %	-
	T2	34	15	44 %	-
Tape (PB)	T0+CAF	12	13	108 %	-
	T1	32	15.5	48 %	-

Report of the Computing Resources Scrutiny Group

CMS

Resource	Site(s)	Pledged	Used	Used/ Pledged	Average CPU efficiency
CPU (kHS06)	T0+CAF	106	34	32 % (22 %)	55 %
	T1	132	86	65%	88 %
	T2	315	259	82%	80 %
Disk (PB)	T0+CAF	4.5	3.5	79%	-
	T1	16	14	87%	-
	T2	20	14	70 %	-
Tape (PB)	T0+CAF	22	10	46%	-
	T1	44	21	48%	-

Report of the Computing Resources Scrutiny Group

LHCb

Resource	Site(s)	Pledged	Used	Used/ Pledged	Average CPU efficiency
CPU (kHS06)	T0+CAF	21	20	95 %	79 %
	T1	70	57	81 %	93 %
	T2	48	65	135 %	98 %
Disk (PB)	T0+CAF	1.5	0.9	60 %	-
	T1	3.8	2.0	53 %	-
	T2	--	--	--	-
Tape (PB)	T0+CAF	2.5	1.1	40 %	-
	T1	3.9	1.7	44 %	-

Report of the Computing Resources Scrutiny Group

Mitigation of the growth in resources

• Experiments have made an effort to reduce the raw event size (and the size of all subsequent derived formats) and event processing times. These efforts have mitigated the serious challenge that pile-up represents. This has allowed experiments to record events at a higher rate, indicating some margin of safety and redundancy in the resources available.

• Collaborations have made substantial changes in their data distribution policies, reducing the number of copies stored in Tier 1 or Tier 2 and moved to more compact datasets such as AOD's.

• Reconstruction times have generally improved.

• Substantial progress in the implementation of fast Monte Carlo simulations has been made

• Experiments have been very active in redistributing tasks among CERN, Tier 1 and Tier 2 to optimize the usage of resources. The CRSG welcomes an improved, more equilibrated distribution of the usage.

• Experimental collaborations have implemented aggressive data cleaning policies.

•The user efficiency is better than planned (with the exception of ALICE).

Report of the Computing Resources Scrutiny Group

PART B

Scrutiny of the requests for 2012 and 2013 (final)

Report of the Computing Resources Scrutiny Group

ATLAS

CPU [kHS06]	2012	2013
CERN	*73 → 111*	*111*
Tier-1	*259*	*273*
Tier-2	*266*	*289 → 281*
Disk [PB]		
CERN	*9*	*10*
Tier-1	*27*	*30*
Tier-2	*47*	*53*
Tape [PB]		
CERN	*18*	*18*
Tier-1	*36 → 29*	*40 → 33*

Report of the Computing Resources Scrutiny Group

CMS

CPU [kHS06]	2012	2013
CERN	*120*	*120*
Tier-1	*145*	*145*
Tier-2	*315*	*315 → 306*
Disk [PB]		
CERN	*5 → 7*	*4 → 7*
Tier-1	*22*	*27*
Tier-2	*26*	*26*
Tape [PB]		
CERN (including HI)	*23*	*23*
Tier-1	*51*	*59*

Report of the Computing Resources Scrutiny Group

LHCb

Date	Site	CPU (kHS06)	Disk (PB)	Tape (PB)
2012 period	CERN	*34*	*3.5*	*6.4*
	Tier-1	*113*	*9.5*	*6.2*
	Tier-2	*43*	*0*	*0*
2013 period	CERN	*33*	*4.0*	*7.7*
	Tier-1	*110*	*11.1*	*8.0*
	Tier-2	*42*	*0*	*0*

Report of the Computing Resources Scrutiny Group

ALICE

In April 2011 we asked ALICE to revise their request to make it commensurate with the expected resources. After the reconstruction and analysis of the AA events it is clear that enough computing resources are not available and are unlikely to be offered in the future.

After a number of conversations with the collaboration it became clear that ALICE prefers to submit a request based on their physics needs, even if unlikely to be fulfilled.

Since April there have been changes to some of the ALICE computing model parameters:

- The data size of a raw pp event has more than doubled since the April 2011 request because of increasing pileup in pp running and a change in the trigger mix. For AA events, increasing event complexity has been balanced by compression, leaving the raw data size unchanged. Compression by a factor of 2 is assumed for the 2012 and 2013 requests, but the collaboration is confident that they can achieve a factor of 3.5
- CPU use per event for pp reconstruction and Monte Carlo is unchanged, but there have been significant reductions by factors 2.8 and 2.5 respectively for PbPb reconstruction and MC simulation, where the memory footprint has been decreased.

Report of the Computing Resources Scrutiny Group

ALICE intends now to take advantage of the heavy-ion event compression to allow an increase in trigger rate, maintaining saturation of the DAQ data-transfer rate and resulting in even more complex (biased by the trigger) events that will increase the aggregate reconstruction time.

The September 2011 ALICE resources request assumes recording 2 x 10^8 AA events in the 2011 and 2012 heavy-ion runs, whereas in the computing model a standard data-taking year assumed 10^8 such events. The T0 CPU request is essentially fixed by the need to reconstruct the heavy-ion events in 4 months before the subsequent pp run starts.

The collaboration's desire to record and process more events offsets improvements in CPU use per event and leads to more ESDs on disk. This would produce disk requirements at all tiers which would be unsupportable in practice.

In response to this the collaboration has re-evaluated its requirements, keeping multiple copies of ESDs for both real and MC events on disk only for the current reconstruction pass and one copy (or none) for earlier passes. This keeps the disk request closer to that made in April 2011, with some increase at T1s. However, the request still exceeds 2011 pledges by almost a factor of two at CERN and the T1s. We do not expect 2012 pledges to increase significantly (although ALICE should gain access to an additional T1 in Korea and possibly one in Mexico), so even the revised request looks unrealistic.

Report of the Computing Resources Scrutiny Group

Latest ALICE requests

Year	CPU (KHEP06)				DISK (PB)				Cumulated Tape (PB)	
	T0	CAF	T1	T2	T0	T1	T2	CAF	T0	T1
April RRB 2012	85	35	160	161	14	11	8	0.2	20	21
April RRB 2013	85	35	157	157	14	9	8	0.2	25	28
Pledged for 2011	62		71	81	6	6	6	0.1	7	8
New 2012	85	35	165	171	14	14	8	0.2	20	17
New 2013	85	35	165	171	14	14	8	0.3	24	21

Report of the Computing Resources Scrutiny Group

Recommendations and requests

- The CRSG is very concerned about the medium term sustainability of the ALICE computing model and about the low user efficiencies that have been reported to us. We recommend a substantial revision of this model. We suggest that it be a definite priority for action during the 2013 LHC stop if not already fixed by then.
- There is still room for improvement in the implementation of efficient staging strategies and/or dynamic data placement policies. The implications for best-use of resources of the interplay between improvements in network bandwidth and dynamical data placement policies should be evaluated.
- The CRSG encourages close collaboration of the different centres with the experiments to continue the implementation of intelligent storage management policies to allow efficient and cost-effective access to data
- The WLCG accounting of Tier 2 resources is improving steadily but is still insufficient: the Installed CPU compared to the pledged and installed disk capacity at the Tier 2 centres is not centrally accounted so far. It would be useful to disentangle the efficiency of organized/chaotic activities.

Report of the Computing Resources Scrutiny Group

- Care should be taken that the worldwide LCG resources are used as much as possible. A balanced use of the resources is essential for the long time coherence of the WLCG. A continuous increase in the request for additional resources at CERN is not sustainable.
- We encourage that a discussion on the issue of the memory footprint is undertaken as soon as possible. We note that in some cases lack of memory makes half the cores unusable.
- The CRSG recommends that CERN's policies of resource sharing when allocations are not fully used are clearly stated.
- In view of the current statistics we propose to make firm the April 2011 recommendation of assuming a Tier 2 efficiency of 67% for 2012 and increase this to 70% in 2013.
- While welcoming the experiments capability to record events at increased rates, the CRSG does not see how a substantial increase of the data taking rate could be accommodated with the existing computing resources and does not recommend a formal modification of the computing models in this direction.
- We encourage that the mechanisms of communication between experiments and the Tiers are continuously improved to optimize the usage.