CERN-RRB-2011-066

Computing Resources Review Board

Minutes of the 19th Meeting held at CERN on 12 April 2011

DRAFT 2 – 18.07.11

Present:

Europe:

J. Sacton (FNRS, Belgium), J. Lemonne (FWO, Belgium).

M. Lokajicek (Institute of Physics AS CR, Czech Republic), O. Novak (Ministry of Education, Youth and Sports, Czech Republic).

J. D. Hansen (NBI, Denmark).

D.-O. Riska (HIP, Finland).

E. Augé (CNRS/IN2P3, France), U. Bassler (CEA/IRFU, France), G. Lamanna (CNRS/IN2P3, France), P. Rebourgeard (CEA/IRFU, France).

S. Bethke (MPI, Germany), K. Ehret (BMBF, Germany), M. Fleischer (DESY, Germany), V. Guelzow (DESY, Germany), P. Malzacher (GSI, Germany), P. Mattig (BMBF, Germany), M. Nagel (BMBF, Germany), B. Neumair (BMBF, KIT, Germany).

T. Csorgo (KFKI, Hungary), G. Vesztergombi (KFKI-RMKI, Hungary).

C. Bozzi (INFN, Italy), L. Silvestris (INFN, Italy), F. Ferroni (INFN, Italy).

A. Bernotas (Lithuanian Academy of Sciences).

F. Linde (NIKHEF, Netherlands), A. van Rijn (NIKHEF, Netherlands).

B. Jacobsen (Norwegian Research Council).

G. Polok (HNIN, Academy of Sciences, Poland).

G. Barreira (LIP, Portugal).

F.-D. Buzatu (Institute of Atomic Physics, Romania), M. Dulea (Romania).

D. Filatov (Ministry of Education and Science, Russia), I. Golutvin (JINR, Russia), Y. Kozlov (Ministry of Education and Science, Russia), A. Petrov (Mission of Russia in Geneva), V. Savrin (Institute of Nuclear Physics, Russia).

N. Colino (CIEMAT, Spain), G. de Cordoba (MICINN, Spain), F. del Aguila (Universidad Granada, Spain).

T. Ekelöf (Uppsala University, Sweden), E. Olsson (Swedish Research Council).

T. Nakada (CHIPP, Switzerland).

G. Zinovjev (National Academy of Sciences of Ukraine).

A. Medland (STFC, United Kingdom).

Americas:

R. Mcpherson (University of Victoria/IPP, Canada).

J. Butler (FNAL, USA), S. Gonzalez (DOE, USA), H. Gordon (BNL, USA), K. Shank (Boston University, USA), M. Tuts (Columbia University, USA).

Asia Pacific:

C. Jiang (NNSF, China), Q. Chang (NNSF, China).

R. Iyer (Department of Atomic Energy, India), K. Mazumdar (TIFR, India).

L. Levinson (Weizmann Institute of Science, Israel), E. Rabinovici (Hebrew University, Israel).

T. Kawamoto (University of Tokyo, ICEPP, Japan).

D.-J. Chung (MEST, Republic of Korea), J.-H. Lee (National Research Foundation of Korea), S. Riu (KISTI, Republic of Korea).

S.-C. Lee (Academia Sinica, Taiwan).

Africa:

D. Adams (Department of Science and Technology, South Africa).

CERN:

S. Bertolucci (Chair), J. de Groot (Scientific Secretary), F. Hemmer (IT Department Head), T. Lagrange (FP Department Head), J. Salicio Diez (PH Department), E. Tsesmelis (LHCC Scientific Secretary), E. van Hove (FP Department).

WLCG:	I. Bird, S. Foffano.
C-RSG:	D. Espriu.
ALICE:	P. Giubellino, Y. Schutz.
ATLAS:	D. Charlton. F. Gianotti, A. Lankford.
CMS:	D. Bonacorsi, T. Rodrigo, J. Incandela, G. Tonelli.
LHCb:	P. Campana, A. Golutvin.

The agenda and presentations to the meeting are available from the INDICO web site:

http://indico.cern.ch/conferenceOtherViews.py?confId=128046&view=lhcrrb&showDate=all&showSession=3&detailLevel=contribution

---

Introduction

S. Bertolucci welcomes the participants to the meeting.

The minutes of the previous meeting are approved without correction.

WLCG Status (I. Bird)

I. Bird presents a short status report of the WLCG project, concentrating on progress since the last C-RRB meeting.

Records were broken towards the end of 2010 during the heavy ion run. Rates of up to 5 GB/s of data to tape were achieved, much higher than anything anticipated. During proton running in 2010 about 2 PB per month were written to tape. Averaged data input over the year into CASTOR was 2 GB/s. with peaks of 11 or 12 GB/s. These are very large numbers indeed and are far in excess to what was planned originally. Furthermore, this was realized without major disturbances.

The volume of data written into CASTOR in 2010 was 19 PB with 10 PB read. The continuous evolution of the software meant that CASTOR was much more efficient. Only about 50 drives were used at peak rates, much better than expected.

The activities in the experiments during the winter break resulted in large data transfers out of CERN. I. Bird summarizes these activities. Proton-proton data for ATLAS, CMS and LHCb have been reprocessed. ALICE has copied the raw data to the Tier-1s and has reconstructed the lead-lead data, with second pass reconstruction in progress. CMS is in the process of distributing the lead-lead data following offline zero suppression.

I. Bird shows some plots of the number of jobs running as a function of time in 2010, showing the reprocessing strategies for the different experiments. Towards the end of 2010 the experiments used at times all of the available job slots.

Plots of Tier-1 and Tier-2 reliabilities are shown. Reliabilities have much improved compared to the situation of early 2010.

The number of service incidents is not decreasing and their characteristics do not change. Debugging of networking incidents has proven difficult because responsibility for the incident is not clear, indicating that better procedures are needed to solve such problems.

Network monitoring has improved with more tools developed to identify network problems. This was an area of concern. A prototype LHCOPN (LHC Optical Private Network) dashboard has been created as well as an interactive diagnostic tool to monitor LHCOPN traffic.

The decision to run the LHC both in 2011 and 2012 and the experience gathered in 2010, in particular the much higher pile-up than expected, has an impact on the evolution of the requirements. I. Bird shows graphs of the evolution of the requirements in 2011 – 2013 for CERN, the Tier-1 and Tier-2 centres. This will be examined in the report of the C-RSG.

During 2010 a meeting took place to examine how to improve data access and availability. Some prototyping was started that is presently being investigated by the experiments.

It also became clear that the experiments' computing models would have to evolve to make better use of the network. The evolution points to a model where there is data transfer not only from the Tier-0 to Tier-1s and from there to the Tier-2s, but also between the Tier-1s and Tier-2s. This will of course have implications for the network. Until now data has been placed 'by

design'. This is now evolving towards a system where the data are moved on request. The consequences for network usage are not yet clear.

A proposed solution for these questions is LHCone (LHC Open Network Environment), supplementing LHCOPN. Comments on this proposal are welcome. I. Bird adds that the network requirements should be re-introduced in future resources requests and budgeted in the pledges as was done in the early days of the project.

A large number of proposals have been received following the invitation for interested countries to submit proposals for a remote Tier-0. Discussions are on going and more formal steps are not expected until later this year. The additional capacity will be needed by 2014.

Summarizing, I. Bird says that 2010 was a very successful year for LHC computing with data rates exceeding by far those expected. The full system of Tier-0, Tier-1 and Tier-2 sites managed these without problems. Resources usage reached peak values in the latter part of the year with Tier-1 and Tier-2 sites reaching full capacity. Activities continued during the technical stop. Computing models are evolving based on the experience gained and a better understanding of the LHC conditions.

Discussion:

T. Nakada enquires how the Tier-0 storage requirements evolve following the change of the LHC schedule. I. Bird replies that the required capacity is well within what is already available.

On a further question by T Nakada about networking I. Bird replies that there are of course other research communities with large data sets and consequent networking needs, but very few, if any, come close to the LHC needs. It is also not clear that Grid Computing is the solution for all research communities.

T. Ekelof notes that there appears to be a backlash from other sciences against the dominance of LHC computing. Diverse needs must be addressed. Is CERN addressing this question? I. Bird replies that the primary mission of WLCG is to make LHC computing work, but CERN is trying to be helpful and is involved in several EU projects and transatlantic collaborations. He cites data management as very LHC specific. The evolution of data management should take into account the needs of other communities. In this context F. Hemmer cites the unique CERN collaborative model

S. Bethke enquires about the implications of the inclusion of network requirements in the resources pledges. I Bird replies that the cost of network evolution needs to be understood and included in the resources pledges.

G. Barreira remarks that the power bill is a concern for all sites and asks if there any good practices in power management and if this will be discussed with suppliers of hardware. I. Bird replies that he is not aware of any plans in this area, but that if there is a need a workshop could perhaps be organized to look into the question. There must be a lot of accumulated experience in this area.

L. Levinson says that the LHCone should not result in a 'digital divide' where some Tier-2 sites are well connected and others are not. I. Bird replies hat this issue is taken into account in all discussions.

Principal LHCC Deliberations (E. Tsesmelis)

E. Tsesmelis, Scientific Secretary of the LHCC, notes that a paper has been submitted to the meeting (CERN-RRB-2011-017). The overall conclusions of the LHCC are consistent with the status report presented by I. Bird. The LHCC congratulates the WLCG Project and the experiments for the successful processing and analysis of the LHC data. The paper covers the WLCG operations and goes into some more detail concerning the experiments.

Status of Common Projects Accounts (T. Lagrange)

T. Lagrange says that there are no movements to be reported since the cut-off date for the report that was sent to the Board.

Report from the C-RSG (D. Espriu)

D. Espriu, Chair of the Computing Resources Scrutiny Group (C-RSG) starts his presentation by congratulating the experiments and the WLCG project for the very successful operation in 2010.

He recalls the membership of the C-RSG. A number of changes have occurred since the last C-RRB meeting, in particular the secretary of the Group, H. Renshall, who is about to retire from CERN, and has been replaced by H. Meinhard. He thanks the former members of the C-RSG for their contributions.

The task of the C-RSG is to scrutinize the resources accounting for the previous year and the use the experiments made of these resources. The C-RSG furthermore examines the request for resources for the following year, a forecast for the subsequent two years and the match between these requests and the pledges from the institutions. In case there is a mismatch between requests and pledges the C-RSG will make recommendations to address the problem. He mentions that the report from the C-RSG is preliminary and will be finalized for the October C-RRB.

D. Espriu then gives an outline of his talk, reflecting the main task of the C-RSG and emphasizing that the report is preliminary at this stage.

The running schedule for 2011 and 2012 will be similar to 2010 with a total of $5.9\ 10^6$ second of live time of which $5.2\ 10^6$ for proton-proton and $0.7\ 10^6$ for heavy ion running. He notes that with the increasing maturity of the LHC more of the scheduled time should be available to the experiments in 2011 and 2012. He adds that, given the current run parameters, the experiments have to cope with a higher pile-up than expected. In 2010 most of the data have been accumulated towards the end of the run, showing the flexibility of the experiments' computing models.

D. Espriu summarizes the overall usage in 2010 of CERN, Tier-1 and Tier-2 resources, comparing the usage of resources with those at the time of the

October C-RRB. There is substantial increase in CPU, disk and tape usage at CERN and the Tier-1 and Tier-2 centres.

The Tier-2 CPU usage efficiency is shown for the four experiments, comparing the 2010 data with the data at the time of the last C-RRB. The efficiencies are stable. Tier-2 CPU usage includes scheduled processing as well as 'chaotic' analysis tasks. The C-RSG recommends a revision of the target efficiency of 60% to 66%. This issue will be taken up with the experiments and eventually approved in the October 2011 C-RRB.

The percentage of CERN and Tier-1 tape, disk and CPU per experiment in 2010 is shown. The numbers vary widely, reflecting the different computing models. D. Espriu signals as anomalous the large fractions of CPU used at CERN by ALICE and LHCb (33 and 46% respectively). Similarly, the percentage of 2010 Tier-2 CPU resources usage by the experiments is shown. There is no change with respect to the October 2010 data.

D. Espriu then shows the delivered versus pledged CPU, disk and tape resources for the CERN, Tier-1 and Tier-2 sites. The situation is very satisfactory with all parameters showing 100% or more, indicating the availability of resources beyond what was pledged.

Turning to the usage by the experiments of the resources, D. Espriu shows summaries of the Tier-0 + CAF, Tier-1 and Tier-2 CPU, disk and tape resources used.

For ALICE this shows large variations over the year. This is not surprising because of the heavy ion run in November. Usage of central resources at CERN is very large, whereas usage of Tier-2 resources is modest. The need for tape resources at the Tier-1s is a small fraction of the pledged value and must be he result of a miscalculation of the capacity needed. D. Espriu notes that the scrutiny the ALICE computing model is necessarily tentative because of the fact that a large fraction of the resources was used only towards the end of the year.

For ATLAS the usage of resources is quite satisfactory. Most of it was fully used towards the end of the year. The C-RSG has had discussions with ATLAS in order to try and reduce the need for disk capacity at the Tier-2s, with some success. ATLAS has furthermore taken steps to reduce the need for tape storage.

CMS shows a rather equilibrated use of the pledged resources, but the usage of CPU at CERN and the Tier-1s is low. In view of this CMS is planning to move some of its activities there.

LHCb uses a large fraction of the CERN resources. On the other hand, the usage of Tier-1 and Tier-2 resources is less than optimal.

D. Espriu then turns to the scrutiny of he requests for 2012 and 2013, noting that this is preliminary and that the scrutiny will be completed during the year. The C-RSG is generally satisfied with the quality and the level of detail the experiments have provided to the referees. The C-RSG however, makes the following observations:

- ALICE is recommended to submit more detailed reports to the C-RSG.
- The C-RSG would like to see a better time granularity in the requests submitted by ATLAS
- The quantitative impact of pile-up on the need for resources in 2012 and 2013 is not yet fully documented.

D. Espriu notes that some experiments have proposed to the LHCC to take data at a higher rate than previously envisaged. The LHCC has endorsed as a valuable extension of its physics programme a request by LHCb to include charm physics. The LHCC encourages ATLAS and CMS to record data to maintain sensitivity to new physics at low thresholds. The C-RSG however does not see how a substantial increase in the data-taking rate can be accommodated with the existing requests for computing resources. He emphasizes that the C-RSG has carried out its scrutiny based on the existing running scenarios and requests for resources.

The ATLAS request for resources is presented. The request is endorsed by the C-RSG with some slight modification to the CPU resources at CERN and the Tier-1 and Tier-2 in 2013.

The requests by CMS are endorsed by the C-RSG. The request corresponds to the standard CMS 300 Hz. data-taking rate.

LHCb has substantially increased its requests because of the higher than expected pile-up and to the addition of charm physics. Obtaining sufficient computing resources may turn out to be problematic as a result. The referees express concern about the disproportionate use of resources at CERN. The C-RSG recommends reducing the fraction of computing done at CERN to about 25%. The C-RSG further recommends LHCb to have a 'Plan B' in case the available resources do not accommodate peak needs. The C-RSG urges LHCb to move forward with the long-promised online farm for physics.

The request for resources by ALICE is endorsed by the C-RSG. The C-RSG acknowledges that the computing resources requested by ALICE are not available and are unlikely to be available in the near future. The C-RSG therefore requests ALICE to revise its computing model to fit the available resources within a margin of 10%. The C-RSG requests for this revised model to be submitted to the October 2011 C-RRB for endorsement.

D. Espriu notes that the experiments have already made significant efforts to mitigate the growth in computing resources needed. The larger than expected pile-up in particular has meant a real need for optimization. He lists a number of measures implemented by the experiments.

D. Espriu concludes his presentation by a number of recommendations and requests:

- WLCG resources should be used as much as possible to counter the tendency to place more demands on CERN resources.
- CERN resources should be shared when allocations are not fully used. This should not increase the CERN-based share of the analysis.
- The interplay between improvements in network bandwidth and dynamical data placement policies should be evaluated.
- Experiments should quantify the impact of pile-up. The C-RSG expects an evaluation of this impact by the October C-RRB meeting.

- The C-RSG recommends revising the assumed Tier-2 efficiency to 2/3 from the present 60%, representing savings of 10% in Tier-2 CPU resources. A final decision should be taken in the October C-RRB.
- The WLCG accounting of Tier-2 resources, although improved, is still insufficient.

D. Espriu thanks again the members of the C-RSG for their work and congratulates the experiments for the wonderful year of physics, hoping for better things yet to come in 2011.

Discussion:

J. Shank, ATLAS Computer Coordinator, clarifies that the 50% Tier-2 disk usage shown in their usage report is a snapshot at the end of the year. In fact, ATLAS suffered for most of the year from the Tier-2 disks being completely full.

I. Bird comments that 80% CERN CPU resources are in fact shared. He further notes that on slide 10 the 129% of CERN CPU resources available represent a snapshot at the end of the year after some of the 2011 resources were installed. He asks for comments from the C-RSG on the LHCb request to double the tape copies kept at CERN. He also asks for comments on the large request from ALICE, particularly for 2011. D. Espriu replies that ALICE was asked to produce a revised computing model, taking into account the pledged resources. Concerning the second copy of raw LHCb data, D. Espriu says that the C-RSG referees considered the request to be reasonable.

A. Golutvin remarks that LHCb has been working with a pile-up rate six times higher than the experiment was designed for, producing nevertheless excellent physics without any noticeable degradation. But it does imply the need for corresponding resources. S. Bertolucci remarks that he would like to understand the need to keep two copies of the data at CERN.

A. Medland acknowledges the wish of the experiments to extract the physics from the data during the next two years. But this has to be balanced by the resources available. The requests for increased resources may become an issue for the United Kingdom and perhaps also other countries faced with downward pressure on budgets. He asks for any specific measures planned for the update of the experiments' computing models. D. Espriu replies that his understanding is that the increased requests can be accommodated in constant budgets. Some savings are expected from a more efficient use of the available resources.

S. Bethke asks if the long-term preservation of the early data (2009, 2010) is taken care of. D. Espriu replies that tape copies of the data are being kept.

A. Medland asks if the new LHCone network proposal is likely to improve efficiency. D. Espriu replies that moving to dynamic data placement led to some savings of disk resources at the cost of more network traffic. This may become problematic in the future. He notes that networking is not included in the mandate of the C-RSG. S. Bertolucci says that, according to his experience, not much in terms of savings is to be expected and that there has always been a tendency to saturate the available resources.

V. Guelzow explains that at present Tier-2s are relying on networking that is shared with other communities. LHC computing would suffer if other communities would increase their bandwidth. This is one of the main concerns behind the proposal for LHCone.

S. Bertolucci thanks D. Espriu for this presentation.

Status of Resources and Financial Plan (S. Foffano)

S. Foffano starts with an overview of her presentation.

The WLCG MoU has now been signed with 35 countries representing Tier-1 and Tier-2 centres. Two signatures have recently been obtained from LBNL as an ALICE Tier-2 centre and from the University of Ioannina, Greece, as a CMS Tier-2 centre. S. Foffano asks the delegates to communicate any changes, including the name of the representatives to the WLCG Collaboration Board,

with respect to the Tier-1 and Tier-2 centres, the latter being subject to occasional address changes or changes of configuration.

Turning to funding and expenditure, S. Foffano presents the final, 2010 book-closed situation for personnel and material expenditure. Such data were not presented to the last C-RRB meeting because the WLCG was in the midst of planning for 2011 and 2012. She acknowledges as very helpful the possibility to carry forward unused CERN project funds from one year to the next.

S. Foffano shows a summary of the expected funding and expenditure in 2011 – 2016 for personnel and material. For personnel the evolution is rather stable whereas for material expenses there are relatively large fluctuations from one year to the next. A small contribution of external 2011 manpower financed by India, Israel and Italy is handled via dedicated team accounts. Negotiations are under way for the extension of these arrangements.

S. Foffano recalls the process for the collection of monthly accounting data for the Tier-1 and Tier-2 centres. Of the cases of three Tier-2 federations mentioned as not reporting at the October C-RRB, two have been resolved whereas the third (VECC Kolkata) is still under investigation. She notes that if parts of Tier-2 federations do not report their data this results in a distortion of the accounting for that federation.

The results for CERN and the Tier-1 centres are shown for all of 2010 plus the first two months of 2011. The data show in graphical form what D. Espriu said during his presentation. Tape capacity was only partially used whereas towards the end of the year both CPU and disk resources were fully used.

A graph is shown comparing the 2010 and 2011 Tier-2 CPU pledges for the different Tier-2 federations with usage. She invites the delegates to examine the data for errors concerning their sites, in particular in cases where there is no activity reported for a given site.

Concerning the pledges the procedure is that funding agencies present their confirmed pledges to the October C-RRB. The corresponding resources are expected be installed and available in April of the following year. Estimates for one or two subsequent years are requested as well. At the October 2010 C-RRB meeting the list of pledges was nearly complete. Full details of the changes with respect to the pledges presented in October 2010 are provided in the written report. Pledges are now recorded using the REBUS (Resource Balance and Usage) tool. For the pledges presented to the October 2011 C-RRB delegates are requested to input the data directly into the REBUS tool.

A summary of the revised pledges is shown, comparing the data with the data presented to the last C-RRB meeting. Overall the changes with respect to the previous data appear to be minor. S. Foffano emphasizes that the balance for the experiments is based on the experiments' requirements as approved last year.

S. Foffano then proceeds to show an up to date summary of the Tier-1 installed resources. Globally the picture is very good. Almost all centres have fully installed the resources as pledged. She notes that in a number of cases more than the pledged resources have actually been made available.

A summary is shown of the Tier-2 installation status, grouping the different Tier-2 federations by country. In most cases the pledged resources have been fully installed. Where this is not the case there is an understood timetable for installation of the missing part. For the time being this information is collected manually, it is intended to collect the information automatically in the near future.

Resources requirements, approved by the C-RSG, will be published soon. On the basis of these requirements pledges for 2012, i.e. resources to be installed and available on 1 April 2012, will be invited as well as estimates for 2013. The data will have to be entered directly into REBUS by each federation. The deadline for submission is 30 September 2011 for the data to be presented to the C-RRB meeting in October.

Discussion:

V. Guelzow enquires about the consequences of the cut in CERN materials budget in 2010. I. Bird replies that, in place of a three-year replacement cycle, a four-year replacement cycle was adopted of which three years under warranty. Experience shows that this will not have a significant effect on reliability. He does not exclude a slight reduction of capacity.

Summary

S. Bertolucci summarizes the meeting. WLCG was one of the key factors for the success of the LHC program in 2010. The LHC and the experiments are entering a period of maturity. There is a clear need to adapt the computing models. Some rethinking about networking will be necessary, taking care that this should not lead to a 'digital divide'. S. Bertolucci is confident that the process put in place between the C-RSG and the experiments will allow arriving at the appropriate decisions. He urges the C-RSG and the experiments to prioritize requests in case there is a reduction. In his view the cost of LHC computing is not exploding and there is still some margin for optimization. He expresses the hope that future requests can be accommodated in essentially constant budget envelopes.
