WLCG RRB

Minutes of the 33rd WLCG Resources Review Board Meeting
(CERN, Geneva, 24th April 2018)

C. De Clercq (FWO), Belgium
E. Cortina Gil (FNRS, Belgium)
R. Cesar (FAPESP, Brazil)
M. Nasser-Edine (Canadian Foundation for Innovation, Canada)
P. Krieger (University of Toronto, Canada)
O. Novak (Ministry of Education, Youth and Sports, Czech Republic)
M. Lokajíček (Institute of Physics AS CR, Czech Republic)
P. Hansen (Niels Bohr Institute, Denmark)
K. Huitu, K. Osterberg (Helsinki Institute of Physics, Finland)
A.-I. Etiennevre (CEA/IRFU, Centre d'etude de Saclay Gif-sur-Yvette, France)
P. Verdier, V. Beckmann, (IN2P3, France)
M. Gast (BMBF, Germany)
M. Fleischer, V. Guellzow (DESY, Germany)
W. Ehrenfeld (BMBF/PT-DESY, Germany)
S. Bethke (MPI, Germany)
A. Heiss (KIT, Germany)
T. Kolleger (GSI, Germany)
K. Borras (DESY/RWTH Aachen University, Germany)
F. Siklér (Wigner RCP, Hungary)
K. Mazumdar (BARC, Mumbai, India)
E. Kajomovitz (Technion, Israel)
L. Levinson (Weizmann Institute of Science, Israel)
A. Zoccoli, N. Pastrone (INFN, Italy)
M. Cobal (INFN and University Trieste, Italy)
R. Sawada (University of Tokyo, Japan)
Seonghoon Yoon (Ministry of Science, ICT and Future Planning, Korea)
Youngdae Ryu (National Research Foundation of Korea)
S. Noh, K. Cho (Korea Institute of Science & Technology Information (KISTI), Korea)
A. Bernotas (Lithuanian Academy of Sciences, Lithuania)
A. van Rijn (NIKHEF, The Netherlands)
F. Ould-Saada (University of Oslo, Norway)
B. Wosiek (HNIN, Polish Academy of Sciences, Poland)
J. Kitowski (AGH University of Science and Technology, Poland)
G. Stoicica (IFIN-HH, Romania)
V. Savrin (JINR, Dubna, Russia)
P. Strizenc (Slovak Academy of Sciences, Slovak Republic)
J. Cleymans (University of Cape Town, South Africa)
V. Spannenberg (iThemba Labs, South Africa)
M. Martinez Perez (Ministry of Economy and Competitiveness, Spain)
J. Flix (PIC, Spain)
N. Ottosson (Swedish Research Council, Sweden)
M. Rännar (Umeå University, Sweden)
1. Introduction. E. Elsen, Director for Research and Scientific Computing

E. Elsen opens the meeting and welcomes delegates.

2. Approval of the minutes of the last meeting. E. Elsen, Director for Research and Scientific Computing

The minutes of the last Computing Resources Review Board meeting CERN-RRB-2017-132 are approved without comments.

I. Bird presents the status of the WLCG collaboration, where one change is foreseen. The CMS Tier2 in Korea is moving to KISTI, which is already running a Tier1 for ALICE, with a consequent update of the MoU in preparation.

New record rates in data written to tape have been observed towards the end of 2017. Global transfer rates remain significant, with very little fluctuation during the end of year stops. New peaks in CPU delivered have been reached during the winter shutdown. Usage of CPUs continues to be above pledge, as the CRSG report during this meeting will discuss in more details.

A major operational issue in November 2017 was the flood of the CNAF Tier1 in Bologna. This caused damage to equipment including power supplies, CPUs, the tape library and a few tapes. The main impact was that CNAF was down until February 2018, luckily not during data taking. LHCb was the worst affected LHC experiment, as it had some data that were only located at CNAF. CERN and other Tier1 sites provided some contingency while the Tier1 was unavailable. CNAF is now back in production with full resources for 2018.

IPv6 deployment has progressed significantly. All Tier1 sites are now in dual stack mode and a campaign has started for Tier2 sites. I. Bird encourages all Tier2 sites to deploy IPv6 if they have not yet done so.

Another very visible operational issue at the beginning of 2018 was caused by the Spectre and Meltdown vulnerabilities. These triggered a massive patching and rebooting campaign that affected all sites, but luckily this happened during the end of year stop. Initial fears that the patches would cause significant loss of performance turned out to be unfounded.

Updates are presented from each of the experiments with respect to resource utilization, data taking and processing. In all of the experiments, there is an ongoing effort to speed up processing and to reduce the amount of data volumes to be stored. The use of opportunistic non-Grid resources, such as HPCs, clouds and HLT farms, continues to be significant. The average pileup in 2017 was higher than expected in both ATLAS and CMS, causing backlogs and delays in data processing; for instance in CMS the average pileup was 45, instead of 35 that was used in the resource estimates for 2017.

For 2018, the pledge situation is good and the requests are mostly filled. The expected running conditions for 2018 imply an average pileup of 55, compared to the current value of 45; this implies a 20-25% increase in CPU needs. The expected total data volume for 2018 is 60 fb⁻¹, compared to the 45 fb⁻¹ collected in 2017. ALICE are looking forward in 2018 to the most significant data-taking run of heavy ions in Run2.

For Run3, there are still some unknowns on the running conditions and on the experiment plans for trigger rates and for luminosity levelling. Beam energy will be 7 TeV, compared to the current 6.5 TeV. The expected average pileup is around 60, compared to the value of 55 expected for 2018. It is possible that levelling is done at higher luminosities, which would imply an even higher average pileup, with a non-linear increase in CPU needs. There might also be less time between fills, i.e. more live time. The current best guess is that a 30% increase in resources will be needed with respect to 2018, but a more conservative estimate of a 50% increase will be used for resource estimates. In 2021, the first year after LS2, there might be only half-year live time, but the ramp-up to optimal conditions is expected to be rapid.
If the above estimate of a 50% increase in resources is assumed for 2021, CPU and disk needs still stay approximately within a 20% yearly increase from the beginning of Run1 in 2010, i.e. within flat growth. Conversely, the increase in tape needs is expected to be significantly above flat budgets, but still approximately within a 25% yearly increase from the beginning of Run2 in 2015.

ALICE and LHCb however are upgrading during LS2, so their expectations for Run3 do not follow the above assumptions. LHCb will see a factor 5 increase in both luminosity and pileup, with much higher trigger rate and HLT output bandwidth. LHCb have agreed with the LHCC to deliver their computing model and an engineering TDR in Q3 2018, therefore no firm estimates exist yet for their resource needs in 2021. ALICE will see a factor 100 increase in readout rates, but their data volume increase is mitigated by online reconstruction and raw data compression in the new O2 facility. The O2 TDR has been approved and the expected increases in their 2021 needs with respect to 2018 will be 48% for CPU, 74% for disk and 90% for tape.

The funding profile at CERN for WLCG between 2018 and 2021 shows a balanced situation for personnel, while the materials planning includes minimal purchases in 2019-2020 and purchases in 2021 at the level of a factor 1.5 higher than in 2018. The overall balance shows a slight deficit for 2021, but there are still large uncertainties, in particular in cost extrapolations.

Looking further into the future at HL-LHC, an important milestone has been achieved with the publication of the Community White Paper (CWP), under the aegis of the HEP Software Foundation. This includes ideas for change and improvement, gathered from the HEP community at large, across the whole spectrum of activities that are part of HEP Software and Computing. Based on the input from the CWP, WLCG has prepared a strategy document, which is prioritizing a programme of work, from the WLCG point of view, for R&D to be undertaken in view of HL-LHC on the 2025 timescale. The goal of this document is essentially to show that we are in control of the costs, while maximizing the physics output. The clear areas where work needs to be done include software performance for modern architectures, algorithmic improvements, data reduction, improved operational efficiency and hardware cost optimization. The draft strategy document is now being reviewed by the LHCC: it will be discussed at the next LHCC meeting and possibly published shortly afterwards. In the meantime, work has already started in WLCG on some of the R&D projects proposed in the document. Work in these areas is also ongoing inside the individual experiments, as discussed earlier in this presentation, and also in quite a few inter-experiment working groups that are active to address these issues in common.

The expected evolution of market costs is presented. At fixed cost, technological progress is expected to give performance improvements around 15% per year for CPU and 20% per year for disks. However, there are large unknowns, of the order of a factor 2, in cost estimates. Markets are dominated by very few companies and prices depend on demand and on the availability of fabrication plants. Our software also needs to be ready to exploit architectures that are not necessarily ideal for us: for instance, new processors are focussed on machine learning, which may be useful for some of our processing needs but not all of them. Hard disks are still important for us, as SSDs are not yet cost effective at scale. The tape market is also still a concern.

I. Bird finally presents the status of data preservation and open access. A data portal exists at CERN, backed by EOS storage, but not at other sites. This is used in different ways by the different experiments and is currently funded from the WLCG budget. A better medium term plan is needed about how the experiments will use this and the amount of resources required. It should also be clarified how this effort should be funded, and whether this is the responsibility of CERN alone or of the entire WLCG collaboration.
In summary, WLCG and the experiments are working well, with efficient and heavy usage of resources during the winter stop and new peaks reached. The major incident at the CNAF Tier1 was accommodated by the ability of some of the other centers to help out. The infrastructure and resources needed for 2018 are in place and adequate. R&D activities aimed at HL-LHC have started, around the concrete strategy laid out in the WLCG strategy document.

E. Elsen thanks I. Bird for his comprehensive report and invites questions to this presentation.

After congratulating WLCG on their success in running the infrastructure, V. Guelzow expresses his surprise that computing was not mentioned during Monday’s RRB discussion about Phase II upgrade projects. He expresses his concern that it may be impossible to harvest our physics goals with a flat budget for computing, despite technological progress and software efforts in the experiments, and states that this is a real risk that should be mentioned. I. Bird agrees and acknowledges this risk, but notes that currently we cannot have a good prediction of costs. The strategies started by WLCG to reduce data volumes and improve performance by factors will help, but understanding real costs will only be possible much closer to the time of running. E. Elsen notes that this concern has been mentioned in every RRB and in discussions with the Finance Committee. He points out the positive development with respect to two years ago, when the projected shortfall in computing resources was estimated to be more than an order of magnitude. He adds that maybe the overall budget will not remain constant, but it will not be higher by a large factor. He also notes that our software must remain fully flexible with respect to the hardware, for instance because the recent trend towards machine learning architectures may not help us. Now we have a better understanding of what is feasible, but there are still many uncertainties and we would fail if we decided to fix costs today based on some specific architectures.

M. Procario expresses his concern about tapes and asks for clarifications about what is happening in the tape market. I. Bird replies that the major recent change is that Oracle stopped producing enterprise drives, which CERN and many Tier1 sites were using. Sites have refocused on LTO drives, but the components of LTO drives are almost entirely produced by IBM, with all risks implied by a single-vendor market. A second notable evolution is that the tape market has slowed down: there are no technological limits and it would be possible to produce much denser tapes and higher performance drives, but there is no demand for this. The market is driven by companies and institutions that are using tapes much more than we do, but future evolutions are difficult to predict and tapes may be replaced by other types of cold storage. The growth of tape pledges on slide 17 (where the units are TBs of capacity) reflects the fact that people have used tapes more heavily to mitigate disk needs. Today we can buy tape media very cost-effectively compared to disks, without needing large increases in the tape drive and library infrastructure.

E. Elsen notes that new regulations on energy-efficient computing will come into effect in the US next year and asks if this will have an impact on WLCG. I. Bird replies that this issue should be addressed specifically by each national lab and it is not something that WLCG can plan around. He adds that there is however a growing trend of being asked to be using HPCs, which are not necessarily ideally suited to our applications, and we have to adapt to that.

E. Elsen comments that the CERN Open Data portal is a welcome development. Its usage is growing, but this is a spontaneous activity of CERN. The question of duplicate storage requirements for the same data needs to be addressed. E. Elsen invites ideas and proposals on how to access the data.

There are no further comments or questions.

On behalf of the LHCC, T. Wengler reports that, following considerations at the December 2017 and February 2018 sessions, the LHCC congratulates the WLCG and the experiments for the efficient data processing.

Concerning future computing requirements, the LHCC encourages all experiments to continue working on mitigation methods to cope with tight resources. Additional requests for special run data taking should be formulated well in advance by the experiments. The LHCC supports efforts of analysis preservation and open data access and encourages the experiments to investigate common solutions. Both tasks require resources that should be planned for. The LHCC notes that progress needs to be made soon on the HL-LHC R&D projects, to reduce uncertainties on the scale of the problem and to define the direction of further developments for the WLCG TDR due in 2020.

There are no comments or questions.


G. Cavallo presents the WLCG CERN Personnel and Materials expenses per year from 2005 to end-December 2017 inclusive. He then shows the 2 kCHF balance of the LCG Common Fund in December 2017 and the additional contributions received up to April 2018 amounting to 190 kCHF, declaring it in a healthy state. These contributions from the experiments are used to support staff in the IT Database Group. Finally, on the Indian WLCG Team Account the 10 kCHF of expenses, mainly COAS subsistence, led to a positive balance of 28 kCHF in December 2017.

There are no comments or questions.


D. Lucchesi begins by presenting the composition of the group, welcoming P. Sinervo as the new Canadian member, thanking the out-going Dutch representative J. Templon and reminding the Dutch Funding Agency to nominate a replacement. The other members of the group, the CERN Management and the experiments are also thanked for their support.

Plots are presented showing the evolution of the resource shares used by the four experiments between 2014 and 2017, for CPU at CERN, Tier1’s and Tier2’s, and for disk and tape at CERN and Tier1’s. The CERN contribution to global CPU resources in WLCG remains stable around 10-15% for ATLAS and LHCb and around 20-25% for CMS, but has increased for ALICE from 15% in 2014 to 40% in 2017. The ratio of used to pledged resources for CPU, disk and tape at all sites is also presented. D. Lucchesi notes that disk usage accounting reports are not available for Tier2’s and invites WLCG and the experiments to make these reports available if possible.

The evolution of CPU efficiency at all Tier1’s is presented. Progress is also being made on agreeing on a better definition of CPU efficiency.

Details of resource usage experiment by experiment are presented. All four experiments are making good use of unpledged opportunistic CPU resources, with sizeable contributions from HPCs for ATLAS and from HLT for LHCb. All experiments are also doing efforts to control their
storage needs. In particular, ALICE have made good progress in recovering disk space occupied by unused data and in reducing tape space needs thanks to improved HLT compression. The CNAF incident had a high impact for LHCb, with 18% of the data unavailable for many weeks.

The experiment requests for 2019, a shutdown year, are presented. The Tier0 management expects pledges identical to 2018, to re-profile a maximum of purchases at the beginning of Run3. No resource increase at the Tier0 for ATLAS and CMS and a modest increase for LHCb are requested, while the increase is more substantial for ALICE because of the Pb-Pb data taking in 2018. Overall the CRSG is satisfied with the many measures taken by the experiments. The result of the scrutiny is presented, including the following recommendations for each experiment:

- The CRSG encourages ALICE to keep working to reduce the event size of Pb-Pb data and MC. It welcomes the ALICE work on Geant4 multi-threading and encourages further developments to exploit opportunistic resources, in particular HPCs.
- The CRSG is concerned about the ATLAS growth in disk usage and reliance on unpledged CPU resources. It notes that ATLAS have significantly more disk resources than CMS at Tier1 and Tier2 sites and that the difference continues to grow in 2019. It encourages ATLAS to investigate smaller data formats, similar to what CMS have achieved.
- The CRSG notes that the 2018 CMS pledges for disk and tape are now less than 5% and 8% below the requests, which lowers previous concerns about the lack of pledged resources. It applauds the CMS work on reduced data formats such as miniAOD and nanoAOD. It recommends that CMS continue to seek improvements to the relatively low CPU efficiency and appreciates the work done to understand this.
- The CRSG congratulates LHCb for the successful management of their computing model and the links to data processing operations. It considers that LHCb have an appropriate contingency plan if more data than planned are accumulated in 2018. It takes note of the LHCb needs reported for data preservation and participation in the Open Data project.

The CRSG asked the experiments, as part of future resources assessments, to provide a mitigation strategy to address changes in the assumed running conditions at the level of a 20% increase. The CRSG also asked the experiments for an outlook for 2020 resource requests. The answers received from the four experiments are presented.

D. Lucchesi finally summarises the CRSG comments and recommendations.

- The CRSG notes that disk usage accounting reports are not available for Tier2’s and requires that these reports be available from the next scrutiny.
- The purchasing profile can be different in shutdown years with respect to data-taking years and it can vary in the different agencies. The CRSG asks the experiment to provide requests for the next scrutiny under two scenarios: that purchasing is delayed until the beginning of Run3, or that resources are provided year by year during the shutdown.
- The CRSG requests that the CERN expectations for Run3 resources as a function of time be documented and communicated to the experiments prior to the start of Run3.
- The CRSG commends the experiments for their activity on data preservation and their involvement in the CERN Open Data portal. It recommends that this be discussed within WLCG and the experiments to have structured projects with appropriate funding.
- The CRSG appreciates the work done on increasing computational efficiency and reducing CPU and disk storage requirements. It hopes that all experiments reach the same level of involvement in pursuing improvements in these areas.
- The CRSG asks the experiments, as part of future resources assessments, to provide contingency plans to address changes in the assumed running conditions at the level of a 20% increase.
The experiments are encouraged to gather accounting information on the use of non-WLCG resources, in order to be able to report them to the CRSG. The CRSG applauds the efforts of the experiments to use their HLT farms to augment CPU resources.

E. Elsen thanks D. Lucchesi for her clear presentation and recommendations.

There are no further comments or questions.

7. Summary. E. Elsen, Director for Research and Scientific Computing

E. Elsen summarises the meeting. He notes that 2018 is under control, even if the data volumes may be larger than expected. For 2019 and beyond, the strategy is to hold back investments as much as possible to profit from further technological advances, with slower spending during the shutdown years and purchases delayed until the beginning of Run3. The TDRs will help to clarify the needs of the experiments.

E. Elsen thanks the delegates for their support and closes the meeting.