1 Introduction

This report from the Computing Resources Scrutiny Group (C-RSG) summarizes for the four Large Hadron Collider experiments, ALICE, ATLAS, CMS and LHCb, their computing usage in 2018, anticipated usage in 2019 and requests for 2020. Unless otherwise noted, the start of each reporting period is 1 April.

The information from previous years together with the resource requests for 2020 from each experimental collaboration, subsequent questions from the C-RSG to each collaboration and face-to-face discussions with collaboration representatives for clarifications constitutes the inputs examined to arrive at the C-RSG recommendations for 2020 resource procurement. In addition, the C-RSG was provided with earlier estimates of resource requirements for the 2021 year (and in some cases for later years) to evaluate the longer term resource implications, and will comment on those as appropriate.

2 C-RSG membership

Since the autumn 2018 scrutiny, the Funding Agency of The Netherlands has nominated Panos Christakoglou (Nikhef) as a new representative. Panos has already actively contributed to the spring 2019 scrutiny. The C-RSG asks the Resources Review Board (RRB) to officially endorse his membership in the C-RSG.

Donatella Lucchesi (Italy) ended her mandate as chair; according to the WLCG MoU, the CERN Director for Research and Computing, Eckhard Elsen, has nominated Pekka Sinervo (Canada) as chair of C-RSG starting in January 2019.

The members of the C-RSG warmly thank Donatella for her tireless work and her very important contributions during more than six years on the group as a member and then as the chair.

The C-RSG awaits the nomination of a new representative for Italy; it kindly reminds the Funding Agency of Italy to nominate one as soon as possible.

The chair thanks the C-RSG members for their commitment and expert advice, and the experiment representatives for their collaboration with the C-RSG and this review process. Thanks are also due to the CERN management for its support and to the C-RSG scientific secretary, H Meinhard (CERN), for ensuring the smooth running of the group.

3 Interactions with the experiments

The experiments were asked to submit their reports by February 18th, 2019. The C-RSG thanks the experiments for the timely submission of their detailed documents [1–5]. The group would like
to thank the computing representatives of the experiments for their availability, their constructive responses to the questions raised by the C-RSG and subsequent requests for further information. The face-to-face discussions were particularly helpful and greatly appreciated by the C-RSG.

Specific sets of C-RSG referees were assigned to review the ALICE and LHCb requests. As usual, by agreement with ATLAS and CMS management, a single team of C-RSG referees scrutinized the ATLAS and CMS reports and requests to ensure a consistent approach. The referees subsequently reported to the full C-RSG, which then developed the recommendations contained in this report.

In anticipation of the Fall 2019 scrutiny, the C-RSG asks the experiments to submit their documents for the October 2019 RRB by August 28th 2019. The C-RSG requests that as part of their submission the experiments respond to the general and experiment-specific recommendations. According to decisions taken in 2017 about the scrutiny process, this submission is the first opportunity to discuss resource requests for 2021, allowing for in-depth discussion with the experiments and LHCC during winter 2019/2020. The C-RSG plans to make recommendations to the RRB on the 2021 resource requests at the Spring 2020 scrutiny.

4 Overall Resource Usage in 2018

4.1 Overall Usage

The C-RSG assembled the overall patterns of usage from data obtained from the experiment submissions as well as the EGI portal [6]. The time period goes from April 1st 2018 up to March 31st 2019. Disk and tape space occupancy are only available up to end of February 2019. The specific usage by each experiment will be summarized in the sections below.

The experiments have effectively utilized the computational resources pledged to them, and in several cases have exceeded the amounts pledged for pure computation – CPU cycles – by the WLCG by significant fractions, taking advantage of on-line systems associated with their High-Level Triggers (HLT) when these are not needed for data-taking, as well as opportunistic resources they have identified. Data storage in disk and tape have been essential resources, with experiments using 431 PB of disk (79 PB at CERN, 174 PB at T1 and 178 PB at T2) and 575 PB of tape storage (245 PB at CERN and 330 PB at T1). The C-RSG notes that the average disk and tape storage at CERN and T1 used by the experiments has grown by 16% and 23% annually, respectively.

HPC centers outside the WLCG are becoming an important contributor to LHC computing and therefore C-RSG asked the experiments to report on the utilization of those as well. C-RSG reports on the usage following the usual classification of CERN (T0), Tier-1 (T1) and Tier-2 (T2).

4.2 CERN, Tier-1 and Tier-2 Usage

The usage relative to the pledged resources for CERN, Tier-1 and Tier-2 is shown in Table 1 for the last five years. Values are averaged over the four experiments. CPU usage is calculated as the average of time-integrated CPU power over the RRB year for 2018, 2017 and 2016, while calendar year is used before. Disk and tape numbers give the occupancy of available resources at the end of the RRB or calendar years.

CPU usage continues to exceed the amount formally pledged by the WLCG, using both HLT farms and other centres that the experiments have made arrangements with. In all cases, the creation of simulated Monte Carlo (MC) data sets are the single largest CPU requirement and also dominate the storage utilization. Disk space is completely utilized when one takes into account the “headroom” needed to manage large data sets and the distributed analysis efforts underway by each collaboration.
Table 1  Usage summary for different Tiers for 2018, 2017 and 2016 RRB years and for calendar years 2015 and 2014. Data is from Tier-1 and Tier-2 accounting summaries for WLCG obtained from EGI [6]. Tier-2 disk usage is taken from experiment submissions to this report. The CPU data are the usage reported from April 2018 to February 2019 and part of March 2019. Disk and tape utilization is average monthly usage of available pledged resources for the first ten months of the year. Pledged resources are actual available resources reported by REBUS [8]. The CERN percentage is not taken into account properly before 2016 and this explains the difference with the latest years.

Tape storage is currently underutilized by several experiments relative to the amounts pledged by the WLCG; the C-RSG will return to this issue in its recommendations.

Figure 1 shows the yearly evolution of the share of CPU usage by experiment at CERN (top left), Tier-1 (top right) and Tier-2 (middle bottom). In each plot, the percentage used by each experiment normalized to the total CPU cycles used is plotted, therefore they sum up to 100% year by year. The major users at CERN are ALICE, ATLAS and CMS, with LHCb being the minor user at the 5% level. Tier-1 CPU usage is dominated by ATLAS, with CMS at 25% and ALICE and LHCb around 15%. Tier-2 sees ATLAS and CMS as dominant users with ALICE and LHCb consuming 10% of the resources.

Figure 2 shows the year-by-year usage of disk storage by experiment at CERN (left) and Tier-1 (right). The percentage is obtained from the space used by each experiment divided by the total disk space used at CERN and Tier-1, so that by definition they sum up to 100% year-by-year. At CERN, ALICE, ATLAS and CMS continue to use comparable amounts of disk space, though CMS has significantly reduced its usage in the last year, reflecting changes that the collaboration has made in its analysis model. LHCb’s share continues to decline, both in relative terms (now below 10%) as well as in absolute terms, which is also a reflection of changes it its analysis strategies.

The largest disk storage increases have taken place in Tier 1 usage, with total disk storage utilization increasing from 128 PB in 2017 to 175 PB in 2018. Although dominated by the requirements of the ATLAS collaboration, both CMS and ATLAS had similar absolute increases in disk utilization (increases of approximately 10 and 16 PB, respectively) while both ALICE and LHCb significantly increased their use of T1 disk storage in both absolute terms (12 and 9 PB, respectively) and in the total share (4% and 2%, respectively).

Usage of tape storage is shown in Figure 3, illustrating the fraction of space used by each collaboration. Tape usage has increased by 171 PB, with usage of tape storage at CERN increasing by 84 PB and at Tier-1 by 87 PB. The share of tape used by each collaboration has not changed significantly, with ATLAS and CMS using approximately 40% and ALICE and LHCb using between 10 and 15% of the total tape storage.

Finally, Figure 4 explores the year-by-year evolution of the fraction of CPU resources and disk storage
that a given collaboration uses at CERN compared with its total utilization at CERN, Tier-1 and Tier-2 sites. CPU usages and disk space utilizations are shown on the left and right, respectively.

For CPU, an increase in 2016 of CERN’s share is common to all experiments due to the fact that CERN provided the requested resources while the same did not happen at the facilities supported by all the funding agencies. The CERN contribution for LHCb and ATLAS is around 15%, while for CMS it is now about 22%. The CERN CPU contribution to ALICE is slightly more than 40%.

The proportion of disk space used by the collaborations and provided by CERN decreased significantly for ALICE, CMS and LHCb, while the proportion for ATLAS stayed at around 25%. ALICE is the experiment with the largest reliance on CERN CPU disk resources, with 40% and 50% of its CPU and disk storage resources, respectively being provided by CERN. Note that these plots do not take into account Tier-2 disk resources.

4.3 CPU efficiency

CPU efficiency, defined as the CPU time provided divided by wallclock CPU time, has continued to incrementally improve. This is shown in Figure 5, where on the left, the Tier-1 efficiency is shown for the last RRB year and on the right, the Tier-1 efficiency is shown for the last four years for each experiment. The largest changes have come from an improvement by CMS in its overall efficiency, arising from efforts to reduce inefficiencies from payload scheduling effects.

Figure 1 Year-by-year share of CPU resources by experiment at CERN (top left), Tier-1 (top right) and Tier-2 (bottom). Data were obtained from EGI [6] using the RRB year April 2018 to March 2019.
5 Resource Usage: ALICE

The C-RSG report is based on the usage and resource requests provided by the ALICE experiment [1], a written set of responses to C-RSG scrutiny questions, and an in-person meeting with the ALICE computing coordinators. We thank the ALICE team for providing a detailed response to the C-RSG recommendations from the Autumn 2018 scrutiny report. The focus of this scrutiny round is the resource usage for 2018, the allocations for 2020, and initial projections of resources required for 2021.

ALICE data taking in 2018 comprised of two runs. From mid-April 2018 p-p data was collected with a luminosity of \(2.6 \times 10^{30} \text{ cm}^{-2} \text{s}^{-1}\) and an effective run time of \(5.6 \times 10^6\) s (compared to an assumed \(7.39 \times 10^6\) s). A total of 5.3 PB of RAW data were collected (5.4 PB was the expected data volume) resulting in 10.6 PB of RAW data (two instances) stored on tape at the T0 and T1 sites.

Pb-Pb data taking occurred over a 24 day period in November and December 2018. Due to a delay in the beam start and an increase in beam size the time in Stable Beams was 78% of that expected. To compensate for the reduction in the availability of the beam ALICE adjusted its triggers to increase the number of minimum bias events (159M from the predicted 100M), reduce the central events (133M from 250M), and add mid-central events (118M). ALICE applied a standard gzip compression to the non-TPC RAW data to complement the HLT compression resulting in a 17% reduction in data written to tape. A total of 11 PB of data (two instances) were written to T0 and T1 tape (compared to the expected 18 PB).

Of these data, the p-p data have been calibrated and processed through the first of three passes including...
the generation of Monte-Carlo (MC) simulations. The Pb-Pb data have been calibrated with a first pass processing of the data expected to be completed before the end of March. Simulations for the first pass are in production.

As of February 15 2019, tape usage for 2018 amounts to 10.8 PB at T0 giving an accumulated total of 41.4 PB. At the T1 sites the accumulated tape storage is 35.8 PB. Together this corresponds to 84% of the C-RSG approved and pledged capacity. The underutilization of tape is due in part to the performance of the LHC, the change in the mix of Pb-Pb events, the ability to maintain the HLT compression factor of 8.5, and the use of the gzip compression on non-TPC data. The unused tape and improved compression has been accounted for in the resources requests for 2019 and 2020.

Disk deployed for 2018 is 31.2 PB, 30.4 PB, and 25.6 PB at the T0, T1 and T2 sites, respectively. This includes two disk buffers, one of 3.2 PB at T0 and a second of 3.0 PB at T1. The 5 PB at T0 provided by CERN beyond the pledged amount for 2018 will be integrated into the 2019 resource allocation. Usage at the T0, T1, and T2 amounts to 27.5 PB, 26.3 PB, and 23.2 PB, respectively (or 90%, 87%, and 90% utilization). From August 2018 ALICE removed approximately 5 PB of space from T0 to enable the creation of the disk buffers used in the Pb-Pb data taking. ALICE expects to generate an additional 3 PB of data before the end of the 2018 allocation with the completion of the first processing pass for the Pb-Pb data.

ALICE used 541 kHS06 at T0 (55% more than the pledged resources), 340 kHS06 at T1 (22% more than the pledged resources), and 311 (94% of pledged resources) at T2 sites. The HLT contributed 2.2 kHS06 and other opportunistic resources a further 3.6 kHS06. The C-RSG congratulates the ALICE team on their ability to make effective use of opportunistic resources. ALICE continues to demonstrate good compute efficiency at the T1 sites with CPU-to-wall clock ratios of ~86%. Lower
efficiency is achieved at the T0 and T2 sites (82% efficiency).
Simulation remains the dominant factor in the CPU usage with ALICE intending to continue their use of GEANT3 for the Run 2 processing.

6 Resource Usage: ATLAS

Figure 7 shows an overview of ATLAS' resource usage for 2018. This information is based on the report from ATLAS [2], with pledged resources extracted from REBUS; used resources are obtained from the EGI accounting portal [6].
As a consequence of the exceptional LHC performance, ATLAS computing resource usage has been very high during 2018.

CPU utilisation has been dominated by Run 2 physics study activity (MC simulation, MC/raw reconstruction and event generation), only a small part has been dedicated to HL-LHC simulation production. Over 55B events were processed during the year. The fraction of resources used for analysis was 21%, and around 4% were used for raw data processing. CPU efficiency ranged between 80% (T0) and 84% (T1). T0 resources were in constant usage, with Grid jobs backfilling queues during non-data taking periods. Deployment of the ATLAS Event Service allowing for fine-grained event allocations has been completed for MC simulation on Grid and volatile opportunistic resources. Around 17 PB (158 B events) of raw data were collected then promptly reconstructed. Out of these 4 PB (9.1 B events) originated from Pb-Pb collisions during November, where rates to T0 exceeded 3 GB/s. DAOD generation does now fully integrate the usage of the SharedWriter merging component that avoids a separate I/O-intensive merging step.

Overall, CPU utilisation across CERN and T1/T2 sites amounted to 3348 kHS06 (35% over pledged resources, 2479 kHS06). As in previous years, the largest fraction of beyond-pledge resources originate from T2 sites (562 kHS06), with another 213 kHS06 from HLT usage during non-data taking periods. In addition, around 13% of the simulation samples were provided by HPC resources. As in previous years, and despite their continued availability, C-RSG points out the substantial utilisation of beyond-pledge and opportunistic resources as a potential risk.

C-RSG congratulates ATLAS on the fact that fast simulation is rapidly being adopted by the physics groups (7.6 B or 43% of all simulated events, compared to 1.6 B events in 2017).

ATLAS disk utilisation remained very high across all tiers, with AOD/DAOD samples dominating (over 60% of the total space). DAOD processing in 2018 was focused on 2016/2017 data but analysis of 2018 samples increased towards the end of the year; ATLAS expects 2017 and 2018 DAOD data usage to become comparable during 2019. Data distribution is balanced across sites and subject to automated cross-site re-balancing and lifetime policies, resulting in 15-20% of the global disk capacity holding “secondary” data (redundant replicas for access optimisation or expired data that can be subject to deletion).

ATLAS has started several activities for optimising its comparatively high disk occupancy. C-RSG encourages these activities and looks forward to updates in these areas:

- ATLAS has set up a dedicated study group for reducing the AOD and DAOD (derived AOD) footprint by optimising the amount of different derivation formats and their physics data payload. The target is to reduce AOD/DAOD occupancy by 30% in time for Run 3. The study group will issue a final report during summer.

- A new workflow component, the Event Streaming Service, is being prototyped in conjunction with the DOMA WG. Its goal is to enable finer input data access granularity and to the usage of WAN data access, potentially reducing the need for multiple replicas and thus helping to reduce the ATLAS disk footprint.

- ATLAS has initiated a “data carousel” R&D activity to study the performance of I/O intensive workflows running from different storage resources with varying quality of service parameters such as cost, latency and throughput. This includes evaluating the possibility of running more tape-intensive workloads on T1 sites such as tape-based DAOD production.

ATLAS tape occupancy has significantly grown over 2018, but still remains below requests approved by RRB (231 PB occupied, corresponding to 77% of 2018 pledge). During LS2, ATLAS expects tape volume to increase by 50 PB (281 PB or 89% of 2019/2020 pledges). According to ATLAS, this
leaves some margin space for Run 3 preparations; also, ATLAS assumes that they are required to keep some headroom for site-specific operational needs such as media repacking and migration. C-RSG requests ATLAS to provide greater details on LS2 tape utilisation for the next scrutiny round.

As requested by C-RSG, ATLAS 2018 requests included a contingency for 20% more luminosity— which did eventually not fully materialise. As discussed in more detail below, ATLAS therefore has reduced 2020 requests to 0% increase with regard to 2019 and does not require significant additional resources during LS2.

### 7 Resource Usage: CMS

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>Disk</th>
<th>Tape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier-0</td>
<td>423</td>
<td>26.0</td>
<td>99.0</td>
</tr>
<tr>
<td>Tier-1</td>
<td>600</td>
<td>60.0</td>
<td>188.0</td>
</tr>
<tr>
<td>Tier-2</td>
<td>900</td>
<td>70.0</td>
<td>188.0</td>
</tr>
<tr>
<td>HLT</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Total</td>
<td>1923</td>
<td>156.0</td>
<td>287.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Pledged</th>
<th>Pledged /C-RSG</th>
<th>Used</th>
<th>Used /C-RSG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier-0</td>
<td>423</td>
<td>100%</td>
<td>365</td>
<td>86%</td>
</tr>
<tr>
<td>Tier-1</td>
<td>600</td>
<td>94%</td>
<td>540</td>
<td>90%</td>
</tr>
<tr>
<td>Tier-2</td>
<td>900</td>
<td>103%</td>
<td>1432</td>
<td>159%</td>
</tr>
<tr>
<td>HLT</td>
<td>n/a</td>
<td>n/a</td>
<td>148</td>
<td>n/a</td>
</tr>
<tr>
<td>Total</td>
<td>1915</td>
<td>100%</td>
<td>2485</td>
<td>129%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Others</th>
<th>Others /C-RSG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>30</td>
<td>2%</td>
</tr>
</tbody>
</table>

Figure 8 Summary of planned and used resources for CMS in WLCG year 2018. Note that CMS manually keeps T2 disk usage at 90% utilization. * Percentage taken with respect to total C-RSG CPU recommendations.

As a consequence of the exceptional LHC performance in 2018, CMS has recorded 64 fb$^{-1}$ of pp collisions resulting in overall computing resource usage being very high.

The CPU utilization has been dominated by Run 2 physics study activity (MC simulation, MC/raw reconstruction and event generation). A total of 47 B physics events were recorded and reconstructed and 16 B Monte Carlo events were produced.

The CPU efficiency ranged between 66% at T2 and 75% at T1, considerably improved with respect to last year due to the minimization of payload scheduling effects, yet still lower than for the other experiments. CMS continues to work on improving the CPU efficiency but also pointed out that the CMS efficiency comes from a global optimization of computing resources, taking into account also storage costs.

Overall, CPU utilization across CERN and T1/T2 sites amounted to 2485 kHS06, 34% (29%) over pledged (C-RSG recommended) resources.

The beyond-pledge resources originate largely from T2 sites (536 kHS06) with an additional 148 kHS06 from the usage of the HLT in non-data taking periods. An additional 30 kHS06 (2% of the total pledged CPU resources) originate from other opportunistic resources, mostly exploiting the
new Dynamic On Demand Analysis Cluster (DODAS) mechanism. CMS plans to extend the use of DODAS in the future, potentially also allowing its use for smaller T2 centers.

As in previous years, and despite their continued availability, C-RSG points out the substantial utilization of beyond-pledge and opportunistic resources as a risk.

CMS had a total of 131 PB of data stored on disk, were the occupancy is kept at 90% automatically by the CMS Dynamic Data Management system. A rather low disk occupancy of 69% at CERN at the end of the year is a special situation and should be cured by recalling 7 PB of Parking B data from tape for processing.

CMS continues its effort to carry out more physics analyses in the new nanoAOD format with the goal to further reduce requirements for disk space in the future.

CMS is using a total of 219 PB of tape resources at CERN and T1 sites. The T1 usage corresponds to 68% (77%) of pledged (C-RSG approved) resources. CMS assumes that they are required to keep 15% headroom for site-specific operational needs such as media repacking and migration. As this constitutes significant resources C-RSG requests CMS to revisit this assumption and check with resource providers on the actual required/recommended headroom for tape storage.

CMS reported in the handling of the CNAF incident reported earlier, where their second biggest T1 suffered a major incident due to a water leak. CNAF was back in full operation by April 2018 and CMS has managed to carry out its planned physics programme by finding individual solutions to specific issues encountered. Also no irreproducible (ie. RAW) data was lost due to the CMS policy of having two copies of such data.

The pledges for CMS have not fully met the C-RSG recommendations for 2018 but the shortfall is reduced with respect to previous years (96% CPU, 94% disk, 91% tape). We encourage the funding agencies to provide the C-RSG approved resources in the future.

8 Resource Usage: LHCb

Offline computing resources were used in 2018 by LHCb for the continuous production of simulated events, analysis activities (execution of user jobs and centralized production of ntuples for working groups) and processing and stripping campaigns of proton-proton and heavy ion data. Simulation production largely dominates the CPU use (around 90%). Figure 9 shows an overview of the LHCb resource usage in 2018. This information is based on their report [4], which compiles usage information from the EGI [6] and LHCb DIRAC [7] accounting portals. Pledged resources are extracted from REBUS [8].

The usage of CPU in WLCG sites was around 30% higher than pledged (654 vs 502 kHS06). Opportunistic resources outside WLCG sites provided a sizeable amount of CPU power (103 kHS06). The largest contributors were the NCBJ (Poland) site, the YANDEX farm, idle CPU cycles from the CERN disk storage servers (BEER project), volunteer computing (BOINC) and a small contribution (few kHS06) from HPC centers. The BOINC queue was stopped in autumn 2018 due to lack of effort to manage the infrastructure. LHCb reported a seamless integration and use of the BEER resources.

The offline usage of the HLT farm provided a large source of unplugged CPU cycles, with periods where its computing power was comparable to that of all the other sites together. An average power of about 200 kHS06 (only 10 kHS06 were pledged) could be exploited in the HLT farm for offline computing activities throughout the year. It represented more than 20% of the total CPU used by LHCb in 2018. An unexpected higher performance of the HLT trigger and online reconstruction algorithms allowed the unplanned usage of this resource to fulfill the request of large simulated samples from the physics groups. By the end of 2018, LHCb was using 79% (87%) of the available disk (tape) resources. Most of the remaining space will be used by the legacy stripping campaigns of Run 1 and
Run 2 data that are under way. Data popularity is continuously monitored to identify datasets not used that can be purged. There was 2.6 PB of disk space (7% of the total volume) that was freed in 2018 in that way. The fraction of used data on disk slightly increased with respect to the previous utilization report (63% of the data have been accessed during the past 3 months and 78% during the past year). Data deletions are performed manually following close inspection of the popularity reports.

9 Background for 2020 resource requirements

The 2020 requests from all the collaborations reflect the fact that the LHC program is in the Long Shutdown 2 (LS2) period (2019-2020), during which the focus of the collaborations are on preparations for Run 3, scheduled to start in mid-2021, and in analysis of the large data sets collected during Run 2. This is reflected in all of the requests, where total resource requests by the collaborations for 2020 are essentially flat.

At the same time, the collaborations are working extensively on their computing models to prepare for the much higher luminosities anticipated during Run 3. Two collaborations in particular, ALICE and LHCb, are undergoing significant changes in their physics goals and attendant computing requirements. The impact of these changes are reflected in very preliminary estimates of computing resource needs for these two experiments. At the time of this scrutiny, the LHCb Technical Design Report for its computing model had just been approved by the Large Hadron Collider Committee (LHCC) while the ALICE Collaboration was in extensive dialogue with the LHCC referees to discuss technical details of computing upgrade for Run 3 (O2 project) along with its organization and milestones.

On the one hand, this makes the analysis of the 2020 requests straight-forward, and the C-RSG has taken the opportunity to better understand the improvements that the collaborations have been working on for their computing models for Run 2, and to the preparations of each of the collaborations for the Run 3 program. This review is summarized in the following sections, where the C-RSG reports on the computing resource requests for 2020 for each experiment.

### Table 9.1: Summary of used resources by LHCb in 2018

<table>
<thead>
<tr>
<th></th>
<th>CRG recomm.</th>
<th>Pledged</th>
<th>Pledged/CRG</th>
<th>Used</th>
<th>Used/CRG</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier-0</td>
<td>88</td>
<td>88</td>
<td>100%</td>
<td>141</td>
<td>160%</td>
</tr>
<tr>
<td>Tier-1</td>
<td>253</td>
<td>250</td>
<td>99%</td>
<td>307</td>
<td>121%</td>
</tr>
<tr>
<td>Tier-2</td>
<td>141</td>
<td>164</td>
<td>116%</td>
<td>206</td>
<td>146%</td>
</tr>
<tr>
<td>HLT</td>
<td>10</td>
<td>0</td>
<td>0%</td>
<td>206</td>
<td>2060%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>492</td>
<td>502</td>
<td>102%</td>
<td>860</td>
<td>175%</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
<td></td>
<td>103</td>
<td>21%</td>
</tr>
<tr>
<td><strong>Disk</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier-0</td>
<td>11.4</td>
<td>11.4</td>
<td>100%</td>
<td>8</td>
<td>68%</td>
</tr>
<tr>
<td>Tier-1</td>
<td>24.5</td>
<td>26.2</td>
<td>107%</td>
<td>22</td>
<td>90%</td>
</tr>
<tr>
<td>Tier-2</td>
<td>5.7</td>
<td>3.7</td>
<td>65%</td>
<td>3</td>
<td>54%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>41.6</td>
<td>41.3</td>
<td>99%</td>
<td>33</td>
<td>79%</td>
</tr>
<tr>
<td><strong>Tape</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tier-0</td>
<td>33.6</td>
<td>33.6</td>
<td>100%</td>
<td>29</td>
<td>87%</td>
</tr>
<tr>
<td>Tier-1</td>
<td>45.6</td>
<td>56.9</td>
<td>125%</td>
<td>40</td>
<td>87%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>79.2</td>
<td>90.5</td>
<td>114%</td>
<td>69</td>
<td>87%</td>
</tr>
</tbody>
</table>
The objectives for ALICE in 2019 and 2020 are to complete the three reconstruction passes for the pp and Pb-Pb data. This will occur at the T0 and T1 sites. Generation of the MCs (1.4x the number of RAW events) will utilize resources from all tiers. Analysis of the data will occur at all tiers and is expected to require about 20% of the requested resources. Two instances of the reconstructed and MC data are initially retained on disk. As each new reconstruction pass is completed earlier passes are reduced to a single instance (with a timescale of approximately one year).

Resource requests for 2020 remain constant in CPU and tape across all tiers. There is a 16% increase in disk requested for T1 and a 15% increase for T2. Given the additional disk provided by CERN in 2018 there is no requested increase in T0 disk resources. All resource requests are consistent with flat growth in 2020.

2021 Outlook  ALICE provided a detailed review of resource requests for 2021. An increase of two orders of magnitude in event rate and an increase by a factor of five in data volume is expected in Run 3. The computing model for ALICE is undergoing a major refactoring to accommodate this increase in event rate both in terms of reconstruction, simulation, and analysis workflow. The revised TDR is in the process of final approval by the LHCC. ALICE anticipates a 2021 request that will see an 81% increase in T0 tape and a 46% increase in T1 tape. We note that assuming a flat budget across LS2 we would expect a 44% increase in tape by 2021.

Disk resource requests are 45.5 PB, 53.3 PB, and 44.8 PB for the T0, T1, and T2 sites, respectively, or a 46%, 21%, and 15% increase. CPU resource requests for 2021 are 455 kHS06, 482 kHS06 and 498 kHS06 for T0, T1, and T2 sites, respectively, or about a 32% increase in requested resources from the previous year. These increases are also consistent with a flat budget across LS2.

ALICE intends to transition from GEANT3 to GEANT 4 for Run 3. Given the length of simulation payload jobs (typically 5 hours with Geant3) will be reduced to < 1 hour through the use of multithreading and event level parallelization simulations will be better matched to the available time in opportunistic HPC resources. This will enable more effective use of HPC clusters in generating MCs. Current development work with GEANT4 has been mixed with simulations of lower energy events a factor of two slower in GEANT4 than GEANT3. ALICE expects to present to the LHCC a report
on the progress in developing GEANT4 simulations and predictions resource requirements for Run 3 before the next scrutiny. The C-RSG encourages the ALICE consortium to continue the work on the transition to GEANT4 given its importance for the CPU resource requests in 2022 and beyond.

For Run 3, ALICE will operate at a peak Pb-Pb collision rate of 50 kHz with all events read out, reconstructed, compressed and written to storage without selective triggering. This will result in ∼3.5 TB/s for Pb-Pb events. All current lossless compression techniques that have been previously adopted with the HLT in Run 2 will be applied in Run 3 together with the removal of clusters not associated with tracks. This is expected to reduce the data rate to ∼90 GB/s. The Pb-Pb event size is expected to be 1.8 MB (4× smaller than Run 2) with AODs also a factor of 2.5 smaller. These improved numbers were used in estimating the Run 3 resource requirements.

The resource projections for 2021 are consistent with a flat budget over the full LS2 with the exception of tape required at T0. Given the increase in data volume during Run 3 ALICE requests an 81% increase in tape for 2021 (compared to the 44% expected from a flat budget over the LS2). We note that there is a substantial underutilization of tape resources when considering all experiments and this could offset the required increase for ALICE.

**Recommendations**

**ALICE-1** C-RSG endorses the ALICE requests for 2020.

**ALICE-2** We request that the ALICE team when presenting the disk usage plots (e.g. Figure 4 and Figure 5 of the current report) include the disk buffer allocation in the plots (as a DC component). This will make the comparison with the pledged resources simpler to evaluate. We encourage ALICE to adopt a consistent way of discussing disk space (i.e. including the disk buffer in all plots and disk usage discussion) throughout their report.

**ALICE-3** We request that the ALICE team provide the C-RSG a copy of the approved TDR and the report on plans for simulations (including the status of the transition to GEANT4 and the decision on the fraction of simulations to RAW data events in Run 3) for the next scrutiny period.

**ALICE-4** Given the expected reliance of ALICE on GEANT4 for Run 3 and the current performance of the code relative to GEANT3 (which was used to benchmark the compute resources required for Run 3) we encourage the ALICE team to increase the amount of effort in transitioning the simulation framework to GEANT4.

### 11 Resource Requests: ATLAS

ATLAS resource requests for 2020 are shown in Figure 11. ATLAS is not asking for increases with regard to the resources approved by RRB for 2019. This represents a slight decrease compared to the requests presented at the October 2018 RRB, where safety margins for handling 10% of higher LHC luminosity were factored in, but the final luminosity delivered was in line with expectations.

During LS2, ATLAS plans to simulate over 20B events, which will dominate CPU utilisation. As in previous years, beyond-pledge and opportunistic resources will play a significant role, with an expected availability of around 700 kHS06 during 2019 and 2020, mostly from T2 sites that are expected to provide around 50% additional capacity with respect to pledges. Additional opportunistic, non-pledged resources will come from the HLT farm and from HPC centres.

As mentioned in the resource utilisation section, ATLAS is implementing several actions for reducing disk storage footprint, with a target for completion by beginning of Run 3. The evolution of
AOD/DAOD processing and tape access efficiency will play a significant role here. C-RSG also looks forward to understand in greater detail how ATLAS tape utilisation (which is still under pledged values) will evolve during LS2.

2021 Outlook

As the LHC running parameters for the first year of Run 3 are still not clear, ATLAS resource requirements for 2021 still have a large uncertainty.

ATLAS expects an average bunch-crossing interaction rate of $\mu = 60$. In terms of CPU resources, this represents an increase of 50%-70% for T0 processing. An increase of 50% will be required for simulation, already accounting for an extensive usage of fast simulation (over 50% of the generated events).

Disk resources are also expected to grow around 50%-70% during Run 3. This assumes on one hand that the ATLAS analysis study group will achieve its goals to reduce AOD/DAOD footprint by 30%, and on the other that there will be also progress in other areas such as increasing the usage of tape-based processing via “data carousel” based workflows.

Tape resources will grow with the amount of simulated, raw and analysis data to be archived; ATLAS expects a total capacity increase of around 50% during Run 3 across T0 and T1 sites.

Recommendations

**ATLAS-1** C-RSG endorses ATLAS requests for 2020. C-RSG encourages ATLAS to continue activities for optimising its comparatively high disk occupancy and looks forward to progress updates.

**ATLAS-2** As in previous years, C-RSG congratulates ATLAS for the successful utilization of opportunistic and beyond-pledge resources. The team is encouraged to keep using them in a truely opportunistic way and avoid dependence of their computing model on their availability.

**ATLAS-3** C-RSG requests ATLAS to revisit its tape under-utilisation at T1 and to provide greater details on foreseen usage during LS2 for the next scrutiny round. The need and extent of operational headroom figures should be clarified with tape resource providers.
ATLAS-4 C-RSG encourages ATLAS to produce specific 2021 request figures for the next scrutiny round. Having this information available as soon as possible will allow funding agencies to timely prepare budgets and purchases, therefore minimising impact of a steep resource increase.

12 Resource Requests: CMS

CMS resource requests for 2020 are shown in Figure 12. CMS is not asking for increases with regard to the resources approved by RRB for 2019. This is unchanged from the CMS requests presented at the October 2018 RRB.

The main computing activities of CMS planned for LS2 in 2019/2020 are:

- a complete legacy processing of all Run 2 data
- continuation of Run 2 analyses based on the legacy samples
- Run 3 Monte Carlo production aiming at 10 B events providing an adequate scale test of the CMS system
- HL-LHC Monte Carlo production for the preparation of various TDRs
- deployment and readiness demonstration for new computing infrastructure for Run 3, like the introduction of Rucio as data management system
- freeing of disk and tape resources for Run 3

2021 Outlook As the LHC running parameters for the first year of Run 3 are still not clear, CMS assumes 5.6 million seconds of pp running time in 2021 with luminosity leveled at $2.0 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ and an average pile-up rate of $\mu = 45$ at $E_{\text{cm}} = 14 \text{TeV}$.

With these assumptions, the CMS modeling results in an overall increase in resource needs relative to 2019 of 30% in CPU (summed across T0, T1 and T2), 20% for disk (summed across T0, T1 and T2) and 25% of tape (summed across T0 and T1). The storage requirements are somewhat reduced with respect to the Fall 2018 report and assume an aggressive cleanup of Run 2 and LS2 commissioning

![Figure 12 CMS resource requests for 2020 and C-RSG recommendations.](image-url)
samples as the 2021 data taking period starts up. CMS has also made some very first estimates of the 2022 resource needs, assuming that this will be a nominal Run 3 data taking year and resulting in similar growth rates as for 2021.

Due to the substantial uncertainties in the LHC running scenario and CMS operating scenario, these estimates contain significant uncertainties.

**Recommendations**

**CMS-1** C-RSG endorses CMS requests for 2020. C-RSG encourages CMS to continue the commissioning and eventual use of nanoAODs for physics analyses with the goal of reducing the disk resource needs and looks forward to progress updates.

**CMS-2** As in previous years, C-RSG congratulates CMS for the successful utilization of opportunistic and beyond-pledge resources. The team is encouraged to keep using them in a truely opportunistic way and avoid dependence of their computing model on their availability.

**CMS-3** C-RSG requests CMS to revisit its tape under-utilisation at T1 and to provide greater details on foreseen usage during LS2 for the next scrutiny round. The need and extent of operational headroom figures should be clarified with tape resource providers.

**CMS-4** C-RSG encourages CMS to produce specific 2021 request figures for the next scrutiny round. Having this information available as soon as possible will allow funding agencies to timely prepare budgets and purchases, therefore minimizing impact of a steep resource increase.

13 **Resource Requests: LHCb**

Figure 13 shows the resource requests submitted by LHCb for 2020, as well as the used resources in 2018 and the pledged resources for 2019. There have been no changes from LHCb in the resource requests for 2020 with respect to the Fall 2018 C-RSG scrutiny round. The C-RSG considers the requests adequate to the planned activities, which involve intense testing of the new computing model for Run 3.

A 20% increase in CPU and disk resources, and 7% increase of tape volume is requested for 2020. In addition to continue with Run 2 data processing and simulation production activities, a large sample of simulated events with the Run 3 conditions will need to be produced in order to fully test and validate the new software, data and computing workflows for the LHCb upgrade.

As in previous years, a small amount of CPU power (10 kHS06) is pledged for the offline use of the HLT farm in 2020. In the past, LHCb has been able to exploit the HLT farm resources in a much larger extend than planned (20 times higher in 2017 and 2018). It is likely that the same trend will continue in 2019. However, in 2020 LHCb plans to upgrade and move the HLT farm to the surface, so the availability of this resource for offline processing seems to be uncertain.

**2021 Outlook** The LHCb detector is undergoing a major upgrade to face Run 3 data taking. A factor 5 higher instantaneous luminosity and a more efficient fully-software trigger (factor 2) will provide a 10-fold increase of the output rate. Due to the increased level of pile-up, the data volume would increase a factor 30 compared to Run 2. LHCb has upgraded its computing model to handle such a large data flow and drastically reduce the needed offline computing resources to store and process the data. About 70% of the output events will be processed by the TURBO stream where raw data will be discarded and only high-level information from the online reconstruction will be saved offline. This
will generate a 2.5 GB/s throughput to disk and tape. The remaining 30% of the events will be entirely archived to tape (7.5 GB/s), prompt-reconstructed and slimmed so that the effective bandwidth to disk is reduced to 1 GB/s.

While real event data will dominate the storage needs, the offline CPU needs will be driven by the production of simulated samples. A large reduction in the event full simulation time and a larger use of fast simulation will be required to limit the CPU needs.

LHCb has made an attempt to estimate the offline computing resource needs for Run 3 within the upgraded computing model. Significant yearly growth factors are anticipated, at the level of 1.4-1.8 during the period 2021-2023.

**Recommendations**

**LHCb-1** C-RSG congratulates LHCb on the successful management of their computing resources, in particular the successful exploitation of opportunistic CPU resources within WLCG and the large use of the HLT farm CPU to enhance the physics program with additional simulated samples. The usage of opportunistic resources in supercomputing centers (HPC) is small though. A larger use of HPC resources seems to be currently hampered by the unavailability of enough effort to integrate HPC centers in the LHCb distributed computing system. In particular, centers with no outbound network connectivity cannot currently be used. C-RSG recommends to invest more effort in this direction, specially given the large CPU needs foreseen for Run 3.

**LHCb-2** C-RSG recommends the implementation of an automatic procedure to delete cold data from disk using the information from the popularity monitoring and following pre-established lifetime policies for the various data formats. This dynamic data management should result in a better utilization of the available disk space.

**LHCb-3** C-RSG considers appropriate the resource requests for 2020 and recommends its granting. The requested resource increase will be essential for the thorough testing of the upgraded computing model for Run 3.

**LHCb-4** C-RSG congratulates LHCb for the enormous effort in reducing resource needs for Run 3 in view of the foreseen 10-fold increase in output event rate (and 30-fold data volume rate).
Nevertheless, still large yearly resource growth factors are anticipated, significantly above the growth rate experimented in Run 2. We note that aggressive assumptions had to be made in order to reduce CPU, disk and tape needs to manageable levels, resulting in a significant risk of underestimation of the required resources. Insufficient available resources would lead to data parking on tape and delayed physics analysis.

14 Comments and recommendations

The analysis of the requests presented above have resulted in the following findings and recommendations by the C-RSG.

**ALL-1** The C-RSG finds that the computing resources requested by the four collaborations for the 2020 year are essential to address their approved physics programs and recommends that these resources be made available by the WLCG.

**ALL-2** The C-RSG recognizes that the approved physics program for LHCb and the proposed program for ALICE will require significant increases in CPU, disk and tape storage by these experiments starting already in 2021. Given relatively large unutilized but pledged tape resources, the C-RSG recommends that the ATLAS and CMS collaborations review for the Fall 2019 scrutiny their tape storage requirements.

**ALL-3** The C-RSG recommends that a common set of planning assumptions related to instantaneous and integrated luminosity for Run 3 by adopted by the experiments in the computing resources scrutiny process. These should be informed by the best information available from the LHC Machine Committee, along with estimates of upper and lower performance to allow the experiments to make estimates of computing resources for the 2021 and future years.

**ALL-4** The C-RSG appreciates the efforts to estimate what increases can be assumed under a “flat budget” model for the acquisition of computing resources and data storage, using the experience of CERN and several Tier-1 centres in the WLCG. It recognizes that this will be information critical to the success of the Fall 2019 scrutiny, when the resources needed for the first year of Run 3 data-taking will be evaluated.

**ALL-5** The C-RSG recognizes that the 2020 year is unusual in that the required computing resources will be similar to the 2019 year, but that early indications are that the resources needed for the Run 3 program starting in 2021 will require increases in CPU and storage that exceed what a year-by-year flat budget model will provide. It notes that it may be possible to meet most of the additional computing requirements for Run 3 if a multi-year approach is taken to acquiring the financial resources needed to ramp up in the three years starting in 2021 by making use of funds available in the 2020 budget year. This may require close coordination between different funding agencies over the next three years.

**ALL-6** The C-RSG requests that as part of resource assessments for Run 3 the experiments provide a strategy to address the case where running conditions for the experiment (e.g., pile-up, luminosity or other effects) are 20% larger than anticipated. For this mitigation strategy the experiments should assume no increase in pledged disk or CPU.

**ALL-7** The C-RSG requests that in future submissions the experiments provide a section in their reports that responds to the findings and recommendations from the previous scrutiny. This response should address both the experiment specific recommendations and general recommendations relevant to all the experiments.
References


