CERN–RRB–2024–012
8th April 2024

Computing Resources Scrutiny Group

C Allton (UK), P Christakoglou (The Netherlands), A Connolly (USA), D Elia (Italy), E Fede (France), J Hernandez (Spain), J Kleist (Nordic countries), T Mkrtchyan (Germany), M Schulz (CERN), P Sinervo (Canada, chair person), A Valassi (CERN, scientific secretary)

1 Introduction

The Computing Resources Scrutiny Group (C-RSG) is responsible for evaluating the computing needs for the Large Hadron Collider (LHC) experiments, ALICE, ATLAS, CMS and LHCb, and for making recommendations to the Computing Resources Review Board (C-RRB) regarding the World-Wide LHC Computing Grid (WLCG) resources required by the Collaborations to address their approved physics programme. The C-RSG has completed its spring 2024 review, which considered the computing usage in 2023, anticipated usage in 2024 and requests for 2025 computing resources for the four LHC experiments. This report summarises the results of this scrutiny. Unless otherwise noted, the start of each reporting period is 1st April.

The C-RSG solicited reports from each Collaboration on the utilization of computing resources in 2023, the anticipated utilization in 2024 and proposed resources in 2025. The C-RSG reviewed these reports, which also addressed the C-RSG recommendations from the most recent scrutinies in spring 2023 and autumn 2023. The C-RSG sent questions to each Collaboration before dedicated meetings were held with Collaboration representatives for clarifications. The face-to-face meetings with the Collaboration representatives provided the C-RSG with an informed understanding of the computing needs.

The C-RSG also had an informal meeting with the LHCC (LHC Experiments Committee) WLCG referees, who focus on the role of the WLCG in supporting the approved physics programmes of the experiments. This allowed the LHCC and the C-RSG to confer on areas of common interest related to computer resources and utilization. Although the mandates of the two groups are different, the overlap in the evaluation and oversight of computing resources made for a helpful and informative discussion. The C-RSG thanks the LHCC referees for their participation in this discussion.

The C-RSG recommendations for 2025 resource procurement by the WLCG are provided in this report. They are based on the current LHC Run 3 schedule for 2024 and 2025 [1]. The C-RSG recommendations assume that the 2025 operation of the LHC will increase the Run 3 datasets for all four Collaborations, with a particularly large increase for the LHCb experiment.

2 C-RSG Membership

There have been no changes to the C-RSG membership since the autumn 2023 report. The chair thanks the C-RSG members for their commitment and expert advice.

The C-RSG extends its appreciation to the C-RSG scientific secretary, Andrea Valassi, for ensuring the smooth running of the group, and thanks CERN management for its support of the group's work.

The committee is particularly grateful for the work of the Collaboration representatives to support the work of the scrutiny group and to address the previous recommendations by the C-RSG.

3 Interactions with the Experiment Collaborations

The Collaborations submitted their reports by 13th February 2024. The C-RSG thanks the Collaborations for the timely submission of their detailed documents [2–6], which also contained responses to the findings and recommendations from the autumn 2023 scrutiny round [7]. The group thanks the Collaboration computing representatives for their availability, their constructive responses to the questions raised by the C-RSG and responses to subsequent requests for further information. The dedicated meetings with experiment representatives were particularly helpful and greatly appreciated by the C-RSG.

Specific teams of C-RSG referees were assigned to review the ALICE and LHCb requests. As usual, by agreement with the ATLAS and CMS managements, a single team of C-RSG referees scrutinised 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 observations and recommendations in this report.

In anticipation of the autumn 2024 scrutiny, the C-RSG asks the experiments to submit their documents by Tuesday, 03rd September 2024. This timeframe is a challenge given summer schedules, but is unfortunately constrained by the need for the C-RSG to submit its report to the C-RRB by 14th October 2024. The C-RSG requests that as part of their submission, the Collaborations respond to its general recommendations as well as those specific to their experiments.

4 Resource Usage in Calendar Year 2023

4.1 Overall Usage

The C-RSG requested the Collaborations to summarise their use of both pledged and non-pledged computing resources for the period January 2023 to December 2023. Although this usage is taken from WLCG accounting records, the Collaborations find that various adjustments have to be made to properly account for the utilization of resources. The C-RSG at various stages in the scrutiny delve into the accounting records to better understand some trends and pledges, but it finds the information provided by the Collaborations to be accurate.

The Collaborations have been focussing their analysis efforts on the initial physics results from the first year of Run 3 data and on completing Run 2 studies. The WLCG resources are used for a mixture of data reconstruction, physics and detector simulation, and physics analysis. The production of large simulated event samples dominates the CPU use while disk and tape storage scale with the recorded data volume. This usage report will follow the usual classification of CERN (T0), Tier-1 (T1) sites and Tier-2 (T2) sites.

The 2023 resources pledged by the funding agencies were approved in March 2022 at a time when the Run 3 programme faced a number of uncertainties. As noted in the autumn 2023 scrutiny round, the 2023 data-taking was affected by a shortened proton-proton (pp) run for ATLAS, CMS and LHCb. The LHCb Collaboration used this data-taking period for detector commissioning while the Collaboration's physics analysis was focused on Run 2 studies. The ALICE Collaboration, on the other hand, had an extended Pb-Pb run that delivered more data than the total dataset recorded in Run 2. Despite these changes, the Collaborations were able to use the pledged resources effectively. The volume of papers arising from these studies is testament to their success.

Total CPU usage in 2023 was 13.5 MHS23-years, representing a 22% increase in CPU utilization over 2022. Over one-third of this increase came from increased utilization of opportunistic resources. Disk and tape storage was limited to pledged resources. Disk storage used by the Collaborations grew from 662 PB in 2022 to 764 PB in 2023, while tape storage increased from 880 PB to 1,150 PB during the

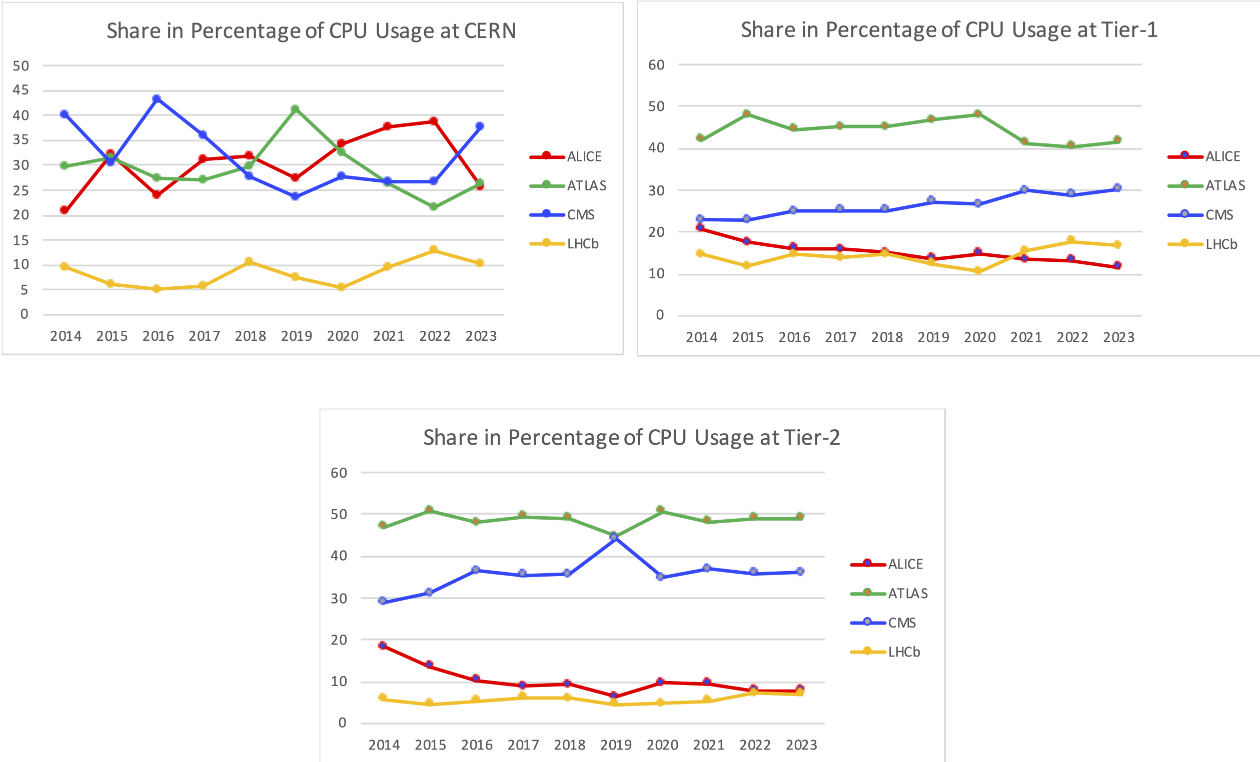

Figure 1 The figure on the top-left shows the percentage of CERN CPU that is used by each experiment. The figure on the top-right shows how the CPU resources at T1 sites are shared by experiment. The bottom figure shows the same sharing of T2 CPU resources. The shares for each year sum to 100%.

same period. Both T0 and T1 tape storage increased by 30% between 2022 and 2023, reflecting the increase in Run-3 recorded raw data. Together, the four Collaborations utilised ~1.9 XB of storage.

4.2 CERN, Tier-1 and Tier-2 Usage

The LHC Collaborations have fully utilized the pledged CPU resources, and the ATLAS and CMS Collaborations were able to obtain an additional 50% of CPU resources, which have traditionally been considered opportunistic.

Most of these opportunistic resources were used for production of simulated events. Although these resources are not committed beyond the current year, they have been consistently available. Hence the Collaborations do not believe there is significant risk to their computing plans. At worst, loss of these resources will delay the production of large Monte Carlo event samples, requiring more aggressive prioritization of these by the Collaborations.

The pledged disk space is completely utilised when one takes into account the “headroom” needed to manage large data sets and the distributed analysis efforts underway by each Collaboration. Although tape usage is below pledges, the expenditures for tape media by WLCG sites are made as experiments require increased storage within the limits of the pledges. The LHCb Collaboration has used less than 40% of the pledged tape storage due to the change to their data-taking plans for 2023.

Figure 1 shows the yearly evolution of the share of CPU usage by experiment at CERN (top left), T1 sites (top right) and T2 sites (middle bottom). In each plot, the percentage used by each experiment normalised to the total CPU cycles used is plotted so that they sum up to 100% year-by-year. The primary users at CERN are the ALICE, ATLAS and CMS Collaborations, with the LHCb Collaboration being a minor user at around the 10% level. T1 CPU usage is dominated by the ATLAS experiment at

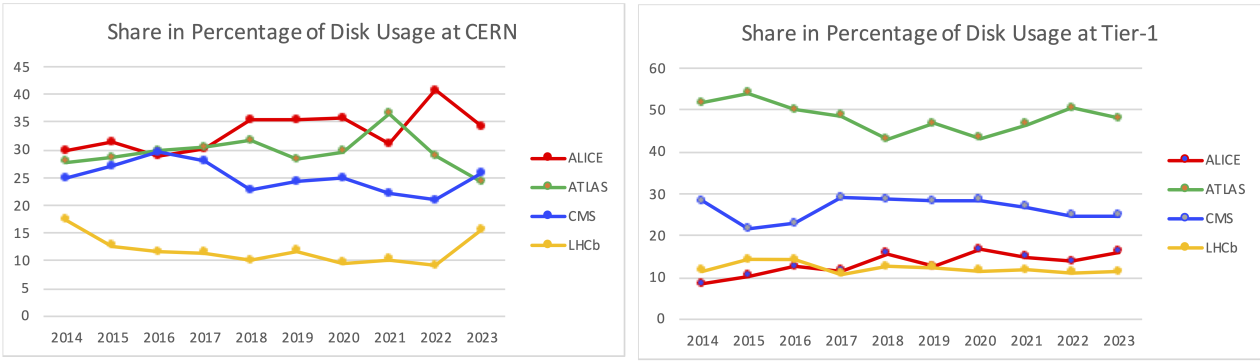

Figure 2 Year-by-year share of disk space by experiment used at CERN (left) and T1 sites (right). Data were obtained from CRIC [8] using the WLCG year.

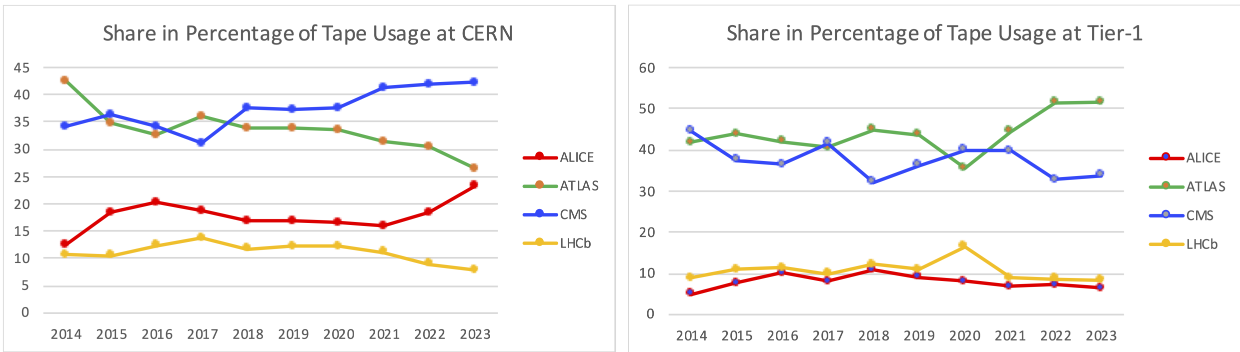

Figure 3 Year-by-year share of tape storage space used by experiments provided by CERN (left) and the T1 facilities (right). Data were obtained from the experiments using the calendar year.

40%, with the CMS experiment at around 30% and the ALICE and LHCb experiments around 15% each. The ATLAS and CMS experiments dominate the use of T2 computing resources, maintaining a consistent pattern even as T2 resources have increased.

Figure 2 shows the year-by-year usage of disk storage by the experiments at CERN (left) and T1 sites (right). The percentage is obtained from the space used by each experiment divided by the total disk space used at CERN and T1 sites, so that by definition they sum up to 100% year-by-year. At CERN, the ALICE, ATLAS and CMS Collaborations continue to use comparable amounts of disk space, though the CMS Collaborations has significantly reduced its usage in the last two years, reflecting changes that the Collaboration has made to its analysis model. The LHCb Collaboration's share of disk increased significantly due the need to have a large portion of commissioning data available for study, and to have a buffer for staging data to tape.

The largest disk storage increases have taken place at T1, with the total volume having increased over the last five years from 174 PB in 2018 to 311 PB in 2023. As shown in Figure 2, this increase has been distributed across all experiments.

Usage of tape storage at T0 and T1 sites is shown in Figure 3, illustrating the fraction of space used by each Collaboration. Tape usage has increased by 270 PB reflecting the increase in Run-3 raw data. The share of tape used by each Collaboration has not changed significantly, with ATLAS and CMS using approximately 40% each 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 CERN CPU resources and disk storage used by a given Collaboration compared with its total utilisation. CPU usage and disk space

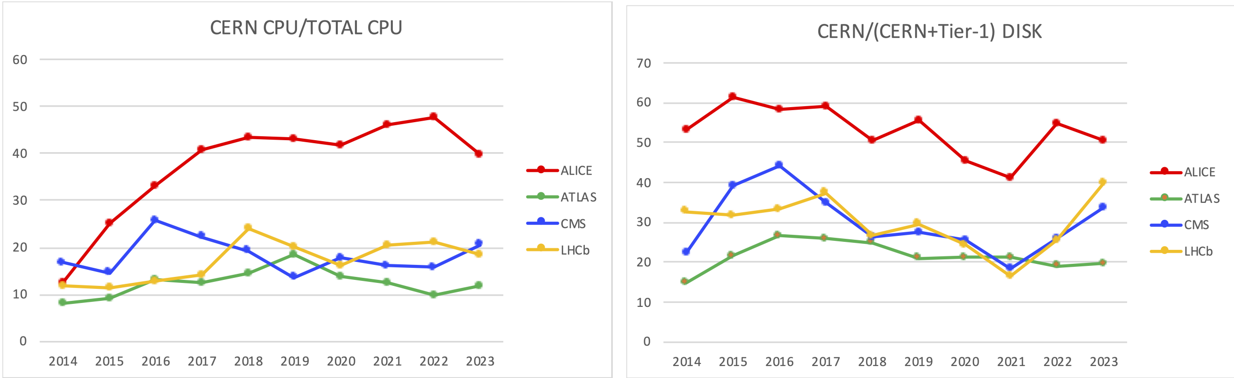

Figure 4 CPU usage at CERN divided by the total usage (left), defined as the sum of CERN, T1 and T2 CPU. Disk space usage at CERN divided by the sum of CERN and T1 usage (right). T2 disk space utilization is not available for the entire period.

utilisation are shown on the left and right, respectively. The CERN CPU contributions to the ATLAS, CMS and LHCb experiments' needs are around 10-20% , while ALICE relies on CERN for 40% of its CPU resources. These patterns have remained largely unchanged over the last 5 years.

The proportion of CERN disk space used by the Collaborations has shown a secular decline for all experiments until 2022, when the need for more CERN disk by the ALICE and CMS experiments increased. The overall trend reflects an increasing amount of T1 disk space available to all Collaborations. The ALICE Collaboration significantly relies on CERN disk resources, with CERN providing 50% of the combined T0 and T1 disk space. The CMS, ATLAS and LHCb Collaborations have between 20 and 30% of their disk space provided by CERN. These data illustrate the important role played by CERN in the WLCG computing model, with the T1 and T2 sites still being critical computing resources. Note that these plots do not take into account T2 disk resources, as these were not consistently tracked until 2018.

5 Resource Usage: ALICE

The report on the ALICE Collaboration's use of computing resources is based on the usage data and resource request provided by the ALICE experiment [2], a written set of responses to the C-RSG scrutiny questions, and a face-to-face meeting with the ALICE computing and physics coordinators. The ALICE report details the computing resources used from January 2023 to December 2023 and is summarised in Table 1.

After the IR8 Inner Triplet incident on July 17^{th}, the ALICE Collaboration had 7 weeks without beam to commission the apparatus in preparation for the heavy ion (HI) run. After operations resumed, the machine ran with limits on the injection intensity of protons. This led to the decision to postpone the planned pp reference run and proceed to the Pb-Pb run after only one pp reference fill at low intensity. The 2023 Pb-Pb run provided 32 days of physics. Due to various machine-related technical issues, the integrated luminosity fell below expectations but the delivered integrated luminosity of 1.96 nb^{-1} exceeded that collected from Runs 1 and 2. There was 43.8 PB of compressed-time format (CTF) files recorded during the HI period.

The complexity of the time-projection chamber (TPC) calibration has driven much of the ALICE Collaboration's computing needs. Difficulties with the TPC space charge distortion (SCD) correction required development of a new calibration method leading to a fourth asynchronous reconstruction pass for the 2022 pp data and, after the Pb-Pb run, a fifth pass. The first Pb-Pb data processing pass

ALICE		2023				
		C-RSG recomm.	Pledged	Pledged /C-RSG	Used	Used /C-RSG
CPU	Tier-0	541	541	100%	660	122%
	Tier-1	572	506	88%	446	78%
	Tier-2	592	567	96%	555	94%
	HLT	n/a	n/a	n/a	n/a	n/a
	Total	**1705**	**1614**	**95%**	**1661**	**97%**
	Others				*222*	*13%**
Disk	Tier-0	58.5	58.5	100%	51.6	88%
	Tier-1	63.5	57.6	91%	50.2	79%
	Tier-2	57.5	60.4	105%	47.3	82%
	Total	**179.5**	**176.5**	**98%**	**149.1**	**83%**
Tape	Tier-0	131.0	131.0	100%	115.0	88%
	Tier-1	82.0	87.7	107%	42.5	52%
	Total	**213.0**	**218.7**	**103%**	**157.5**	**74%**

*: Percentage taken with respect to the total C-RSG CPU recommendation

Table 1 Summary of ALICE resource usage in 2023 and the comparison with recommended and pledged resources. The CPU and storage resources are in units of kHS23-years and PB, respectively.

was carried out on only 20% of the collected statistics in 2023, but was also affected by calibration issues and required repeating.

Event selection for the 2022 pp data retained 0.1% of the collisions. For 2023 data, this was reduced to 0.04%. Together with the CTF skimming, the compression reduction for 2023 pp data was improved from 6% in 2022 to 3% in 2023. MC production of pp events has been undertaken to cover the different running conditions and calibration passes in 2022. Production of MC data for the 2023 HI period proceeded throughout the year together with dedicated MC production for specific physics working groups.

The CPU resources used at the T0 exceeded the pledge by 22% mainly due to the processing of the pp data. An additional 156 kHS23-years was provided by the event-processing nodes (EPN) with reconstruction at the EPN farm benefitting from the use of GPUs. In total the equivalent of 318 kHS23-years of CPU have been utilised. Due to the small amount of Pb-Pb data recorded in 2022, less CPU was required at the T1 sites. This resulted in a use of 78% of the pledged T1 resources.

In addition to the pledged resources, the ALICE Collaboration benefited from opportunistic resources at US HPCs (Lawrencium and Perlmutter). The C-RSG congratulates the ALICE Collaboration for their ability to undertake remote reconstruction of the Pb-Pb data at the Lawrencium HPC - reading data over the network from CERN. A new US Analysis Facility (AF) at LBNL has been added to those at GSI and Wigner. In total the US AF facilities provided 6.4 kHS23-years of unpledged CPU resources.

The CPU efficiency (i.e. the CPU usage/wall-clock time) at T0 for all tasks was 72%, while it was 65% and 62% at the T1 and T2 sites, respectively. This has resulted in an ongoing assessment of the sources of the inefficiency.

The MC simulation workloads ran at an efficiency of 87%, comparable to other experiments. The efficiency of the organised Run-2 analysis is 40%, but it is a legacy activity and so little effort will be invested in improving its performance. The organised Run-3 analysis had an efficiency of 60% and the asynchronous processing had an efficiency of 55%. The ALICE Collaboration reported to the scrutiny group that focused efforts are now underway to understand the root causes of these inefficiencies and to identify mitigation strategies.

The total available disk capacity, including the disk buffers for the tape systems, was 187.6 PB at the end of 2023. The installed T0 capacity was 100% of pledges. The installed disk capacity at the T1s and T2s was 104% and 115% of the pledges, respectively. The reported disk capacity at the T2s included the unpledged storage of 2 PB at Wigner and LBNL.

At the T0, 88% of the pledged storage has been used. At the T1s, 79% of the resources recommended by the C-RSG have been used while 82% of the T1 resources have been used. It was noted that at the end of 2023 the event size of the Pb-Pb analysis object data format (AO2D) and the corresponding MC data had increased by 70 and 50%, respectively. Attempts to reduce the AO2D size for data events have been effected by larger-than-expected data volume, and for MC events by the retention of the full event record.

At the T0, out of the 131 PB of pledged tape storage 115 PB have been used. The volume increased since the last report by 53.6 PB. At the end of 2023, the 2023 *pp* data (corresponding to 40.8 PB) remained on T0 tape. These will be removed in 2024 to free space for 2023 Pb-Pb data. The archiving of the HI data from 2023 is currently in progress. At the T1s, 3 PB have been stored since the last report bringing the total to 42.5 PB of the 87.7 PB pledged. The skimmed CTF data from 2022 and 2023 will be written to tape after the completion of the processing.

6 Resource Usage: ATLAS

The report of the ATLAS Collaboration's computing usage is based on the information provided by the ATLAS experiment [3], written responses to questions by the C-RSG, and a face-to-face meeting with the ATLAS computing coordinators.

The ATLAS experiment has provided summaries of resource usage for the calendar year 2023. The average total CPU usage has been 7879 kHS23-years, well above the pledged 4266 kHS23-years, mostly from unpledged resources provided by T2 sites and HPC facilities. The HPC category was divided into two categories: "HPC" are the sites that accept all type of workloads, dominated by the Vega EuroHPC system that provided more than 2/3 of HPC CPU resources. "HPC-special" are sites that accepted only pre-defined jobs. A small fraction of unpledged resources, 185 kHS23-years, was also provided by the T3 sites, Amazon Web Services and the Google Cloud Platform. The working experience and the "total cost of ownership" evaluation of the Google Cloud Platform is expected to be published during the first half of 2024.

The overall ATLAS CPU efficiency on Grid sites has been consistently above 85%. The HPC systems have shown less efficiency: 67% for HPC and 61% for HPC-special. Though MC simulation scales linearly with the number of cores, the reconstruction jobs show a sub-linear scaling when the number of cores exceeds 8. There is an ongoing R&D activity within the ATLAS software team to address this deficiency.

In 2023 ATLAS has fully utilized 341 PB of disk space out of 342 PB pledged. The tape utilization has reached 473 PB out of 530 PB pledged. The data on disk is divided into three categories: i) persistent data, ii) temporary data that is stored on disk while it is being processed but scheduled for deletion in the future, and iii) cached data that is eligible for deletion when space is needed. Such separation aims to optimise job scheduling and minimise data recall from tape.

The new data formats, DAOD_PHYS and DAOD_PHYSLITE, are being increasingly employed by the ATLAS user community. One of the fist papers published in 2024 is based entirely on analysis of data in the DAOD_PHYSLITE format. Moreover, the analysis tutorials and the first release of Open Data sets for general public are also using the DAOD_PHYSLITE format.

ATLAS		2023				
		C-RSG recomm.	Pledged	Pledged /C-RSG	Used	Used /C-RSG
CPU	Tier-0	740	693	94%	675	91%
	Tier-1	1430	1462	102%	1613	113%
	Tier-2	1747	1841	105%	3457	198%
	HLT	n/a	n/a	n/a	467	n/a
	Total	**3917**	**3996**	**102%**	**6212**	**159%**
	Others				*1667*	*43%**
Disk	Tier-0	40.0	40.0	100%	36.4	91%
	Tier-1	136.0	140.9	104%	149.0	110%
	Tier-2	168.0	160.6	96%	155.0	92%
	Total	**344.0**	**341.5**	**99%**	**340.4**	**99%**
Tape	Tier-0	174.0	174.0	100%	127.0	73%
	Tier-1	353.0	360.3	102%	346.0	98%
	Total	**527.0**	**534.3**	**101%**	**473.0**	**90%**

*: Percentage taken with respect to the total C-RSG CPU recommendation

Table 2 Summary of planned and used resources for ATLAS in the calendar year 2023. The CPU and storage resources are in units of kHS23-years and PB, respectively.

The ATLAS software porting to ARM CPUs has continued. In 2023 the validation of the reconstruction software was performed. The validation was done on the AWS, as no ARM-based grid resources were available for ATLAS. After automatic software builds, the pipeline is finalized and verified by a group of physicists.

7 Resource Usage: CMS

The report on the CMS Collaboration's usage is based on the information provided by the CMS experiment [4], written responses to questions by the C-RSG, and a face-to-face meeting with the CMS computing coordinators.

During 2023, the CMS experiment was able to record 92% of the integrated luminosity delivered by LHC and made effective use of the pledged computing resources. Table 3 shows an overview of the CMS resource usage in 2023 along with the comparison with the amount of resources recommended by the C-RSG and the pledged resources by the sites.

The aggregated CPU usage at T0, T1 and T2 sites and by the high-level trigger facility (HLT) amounts to 5045 kHS23-years, which is about 71% higher than pledged. Significant amounts of beyond-pledge and opportunistic processing capacity at HPCs and WLCG sites have reliably been available during the past few years, allowing the CMS Collaboration to extend its baseline physics program. Nearly three-quarters of beyond-pledge or opportunistic CPU was provided directly by CMS sites and the Run-2 HLT, and most of the rest through allocations at HPCs. The CPU efficiency of CMS jobs at T0 has dropped from by 6% with respect to 2022 and is now at 67%. The CMS experiment is working to identify and address the causes of this drop in efficiency.

Disk usage peaked at 212 PB amounting to 84% of pledged capacity. Tape storage usage peaked at 429 PB amounting to 81% of pledged capacity. The disk space is managed by Rucio and divides the data into three categories: local control, data “locked” by a Rucio rule (effectively “persistent”), data not locked in place by such a rule, called “dynamic” (or “cached”), and unutilized or free space. The tape buffers are not managed by Rucio but rather set manually by the site administrators.

CMS		2023				
		C-RSG recomm.	Pledged	Pledged /C-RSG	Used	Used /C-RSG
CPU	Tier-0	720	720	100%	884	123%
	Tier-1	800	916	115%	1159	145%
	Tier-2	1350	1313	97%	2505	186%
	HLT	n/a	n/a	n/a	359	n/a
	Total	**2870**	**2949**	**103%**	**4907**	**171%**
	Others				*524*	*18%**
Disk	Tier-0	45.0	45.0	100%	39.2	87%
	Tier-1	98.0	96.8	99%	76.7	78%
	Tier-2	117.0	109.7	94%	96.1	82%
	Total	**260**	**251.5**	**97%**	**212.0**	**82%**
Tape	Tier-0	228.0	228.0	100%	202.8	89%
	Tier-1	316.0	303.7	96%	226.4	72%
	Total	**544.0**	**531.7**	**98%**	**429.2**	**79%**

*: Percentage taken with respect to the total C-RSG CPU recommendation

Table 3 Summary of planned and used resources for CMS in the calendar year 2023. The CPU and storage resources are in units of kHS23-years and PB, respectively.

An aggressive on-tape data deletion campaign was initiated in 2023, which resulted in a freeing up of 20 PB, or 7% the total pledged tape capacity.

Though the contribution from JINR is reported as pledged resources following the guidance of the WLCG and CERN management, CMS has considered the T1 site at JINR as an opportunistic site for CPU resources (e.g., Monte Carlo production), but no longer for the storage of unique detector or simulated events. The remaining T1 sites made extra beyond-pledge tape resources available in 2023 in sufficient amounts to close most of the resulting deficit in tape.

8 Resource Usage: LHCb

The level of utilization of the offline computing resources throughout 2023 by the LHCb Collaboration was conditioned by the incident in the LHC vacuum system that prevented the VELO from operating fully closed, rendering the originally scheduled physics data-taking during that period of limited utility. The available CPU resources were efficiently used mainly for the production of simulated events (close to 90% of the total CPU usage), followed by data analysis activities, the processing of the pp and HI commissioning data taken in 2023, and the reprocessing of the data collected in 2022.

Table 4 shows an overview of the LHCb resource usage in 2023 along with a comparison to the amount of resources approved by C-RRB and the resources pledged by the sites. The usage information is based on the WLCG accounting [9] records from the LHCb 2023 resource usage report [5]. The delivered CPU work for sites not included in the WLCG accounting (e.g., the HLT farm) is calculated from the LHCb Dirac accounting [10]. Pledged resources are extracted from CRIC [8].

WLCG sites provided 1402 kHS23-years of CPU work, 12% higher than the pledged CPU for LHCb at those sites. The contribution from non-WLCG sites, including HPC centers, was modest, 26 kHS23-years. The LHCb online HLT farm was exclusively used by the software trigger system to process the commissioning datasets collected by the LHCb detector, without contributions to offline data processing or simulation.

LHCb		2023				
		C-RSG recomm.	Pledged	Pledged /C-RSG	Used	Used /C-RSG
CPU	Tier-0	215	215	100%	258	120%
	Tier-1	707	598	85%	652	92%
	Tier-2	391	434	111%	492	126%
	HLT	50	50	n/a	0	n/a
	Total	**1363**	**1297**	**95%**	**1402**	**103%**
	Others				*26*	*2%**
Disk	Tier-0	30.3	30.3	100%	23.4	77%
	Tier-1	60.5	54.7	90%	35.2	58%
	Tier-2	11.6	7.9	68%	3.6	31%
	Total	**102.4**	**92.9**	**91%**	**62.2**	**61%**
Tape	Tier-0	91.0	91.0	100%	37.8	42%
	Tier-1	157.0	133.7	85%	55.8	36%
	Total	**248.0**	**224.7**	**91%**	**93.6**	**38%**

*: Percentage taken with respect to the total C-RSG CPU recommendation

Table 4 Summary of LHCb resource usage in 2023 and the comparison with recommended and pledged resources. CPU numbers are taken from the EGI and Dirac accounting portals. Disk and tape occupancy are taken from the WLCG storage space accounting and refer to 31st December 2023. The CPU and storage resources are in units of kHS23-years and PB, respectively.

Storage resources were heavily underutilized due to the absence of substantial physics data collected in 2023. The total tape utilization by the end of the year was 93.6 PB, 65% lower than the original request of 248 PB, which was made under very different assumptions for data taking in 2022 and 2023. At the end of the 2023 calendar year, LHCb was using 62.2 PB of disk space, 20% lower than the total available disk space of 78.2 PB, and 33% lower than the pledge of 92.9 PB.

Disk storage at T0 was under stress and saturated for two reasons: i) commissioning data were kept on disk to enable a detailed understanding of the performance of the LHCb detector, its calibration and its alignment; and ii) the migration to the CERN Tape Archive (CTA) system required an unanticipated increase in the disk buffer in front of the tape drives. As a result, the pledged disk space at T0 (30.3 PB) was full in the summer and autumn of 2023. CERN alleviated the peak demand by lending the LHCb Collaboration additional disk space.

9 Anticipated 2024 Resource Utilization

The Collaborations are planning for a full-year of data-collection according to the LHC running schedule for 2024. This can be seen in an expected increase in tape storage from ~1.1 XB to ~1.9 XB.

Both ATLAS and CMS will continue data-taking, reconstruction, simulation and analysis activities without significant change in computing operations, while both ALICE and LHCb will be collecting and processing data at much higher rates than in Run 2. This change is perhaps most dramatic for LHCb, which will be taking its first full year of Run-3 data.

Overall, the C-RSG believes the resources available to the Collaborations in 2024 are well-matched to their physics programmes.

ALICE		2023 C-RSG recomm.	2023 Pledged	2023 Used	2024 C-RSG recomm.	2024 Pledged	2025 Request	2025 req. / 2024 C-RSG	2025 C-RSG recomm.
CPU	Tier-0	541	541	660	600	600	680	113%	680
	Tier-1	572	506	446	630	540	690	110%	690
	Tier-2	592	567	555	650	641	730	112%	730
	HLT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	Total	**1705**	**1614**	**1661**	**1880**	**1781**	**2100**	**112%**	**2100**
	Others			222					
Disk	Tier-0	58.5	58.5	51.6	67.5	67.5	78.0	116%	78.0
	Tier-1	63.5	57.6	50.2	71.5	61.9	79.0	110%	79.0
	Tier-2	57.5	60.4	47.3	66.5	69.8	77.0	116%	77.0
	Total	**179.5**	**176.5**	**149.1**	**205.5**	**199.2**	**234.0**	**114%**	**234.0**
Tape	Tier-0	131.0	131.0	115.0	181.0	181.0	220.0	122%	220.0
	Tier-1	82.0	87.7	42.5	107.0	102.4	123.0	115%	123.0
	Total	**213.0**	**218.7**	**157.5**	**288.0**	**283.4**	**343.0**	**119%**	**343.0**

Table 5 ALICE resource requests for 2025 and C-RSG recommendations. For reference, the 2023 utilization and 2024 pledged resources are also given.

10 Background for 2025 Resource Requests

The 2025 calendar year is planned to be the final year of Run-3 data-taking, and each experiment is expected to collect record numbers of events. The ALICE Collaboration will collect a somewhat smaller data sample than what was originally planned, resulting in a modest reduction in the required resources. It is also the year in which the LHCb experiment will require its largest increase in CPU and storage resources, with 60% increases in CPU and disk storage, and a 45% increase in tape storage.

The following sections summarize the justification for the requests and the recommendations of the scrutiny group.

11 Resource Requests: ALICE

The ALICE Collaboration's request for 2025 is shown in Table 5.

The ALICE request has been affected by the number of days of heavy-ion running in 2025 that is reduced compared to previous plans. In 2024 (2025) there will be 18 (15) days of Pb-Pb runs, making a Run-3 total of 5.5 weeks compared to the previous plan of 8 weeks. There will be no p-Pb run in 2025, contrary to the expectations in the Autumn 2023 report. Two short O-O and p-O runs are being considered that will have two orders of magnitude smaller resource requirements compared to the Pb-Pb runs and therefore will not be discussed further.

While the shorter Pb-Pb run time reduces the resource request, there are two data compression issues which act in the opposite direction. Firstly, the more aggressive "Strategy B" compression is not yet ready to be implemented due to remaining calibration issues with the TPC charge distortions. Secondly, "Strategy A", which is the compression currently used, results in compressed data 1.3 times larger than previously expected due to the presence of additional background clusters in its TPC data. ALICE hopes to implement Strategy B in late 2024 or early 2025.

The AO2D file size for physics events is 380 kB, and 2260 kB for simulated events. These are an increase from the 220 kB and 1500 kB previously estimated. Methods to thin these AO2D file sizes are under investigation.

During the pp runs in 2024 and 2025, ALICE will use their O2 computing facility to perform both the synchronous and first asynchronous passes. ALICE plans two asynchronous passes in total. Two instances of the skimmed pp CTF files will be stored, at the T0 and T1s.

The C-RSG notes that the ALICE Collaboration's ability to fully calibrate and optimize the reconstruction, simulation and analysis of Run 3 data has been effected by changes in the schedule and performance of the LHC (particularly for the Pb-Pb data). For example, for the Pb-Pb run in 2025, data collection is expected to be reduced by 47% from previous estimates. The corresponding reduction in the ALICE Collaboration's resource requests for 2025 is 14% in disk and 24% in tape. Tape and disk utilization for 2024 and 2025 will be driven by progress on finalizing the calibration of the TPC (which determines when the ALICE Collaboration will be able to adopt the more aggressive Strategy B compression), and the size of the AOD and MC AODs.

Conclusions

The C-RSG finds the ALICE Collaboration's 2025 computing requests to be appropriate to achieve its physics programme and recommends that they be approved.

The C-RSG makes the following observations and recommendations to the ALICE Collaboration:

ALICE-1 The C-RSG recommends that, given substantial potential savings in tape utilization under Strategy B compression, the ALICE Collaboration prioritizes evaluation of the physics impact of adopting this improved compression. Decisions about finalizing TPC calibration and adopting different compression strategies will be based on the requirements of the ALICE Collaboration's physics programme. To better understand this process and its impact on resource allocations, the C-RSG requests that the ALICE Collaboration report on the timeline and work undertaken for finalizing TPC calibration and adopting Strategy B compression, together with a description for how such decisions are made (e.g., how requirements are defined and how personnel are allocated for emergent problems).

ALICE-2 For future reports the C-RSG requests that the ALICE Collaboration tabulate their resource requests assuming both Strategy B and Strategy A (if Strategy A continues to be used).

ALICE-3 The C-RSG notes that the overall efficiency of ALICE CPU usage has declined in the last 12 months. Current efficiencies for simulations, Run 3 analysis, and asynchronous reconstruction are now 87%, 60% and 55% respectively. The C-RSG recommends that the ALICE Collaboration prioritize allocating the personnel resources necessary to improve the efficiency of their workflows. The C-RSG requests that the ALICE Collaboration describe the CPU efficiency (overall and separately) together with a high-level description of current sources of inefficiencies (e.g., size or complexity of the AOD data structures) and work that has been undertaken to reduce these efficiencies.

ALICE-4 The C-RSG recommends that the ALICE Collaboration evaluate the adoption of smaller derived AODs and MC AOD formats as a means to improve their analysis and IO efficiencies, and to reduce storage needs.

ATLAS		2023			2024		2025		
		C-RSG recomm.	Pledged	Used	C-RSG recomm.	Pledged	Request	2025 req. / 2024 C-RSG	C-RSG recomm.
CPU	Tier-0	740	693	675	936	936	1100	118%	1100
	Tier-1	1430	1462	1613	1516	1514	1635	108%	1635
	Tier-2	1747	1841	3457	1852	2074	1998	108%	1998
	HLT	n/a	n/a	467	n/a	n/a	n/a	n/a	n/a
	Total	**3917**	**3996**	**6212**	**4304**	**4524**	**4733**	**110%**	**4733**
	Others			1667					
Disk	Tier-0	40.0	40.0	36.4	49.0	49.0	56.0	114%	56.0
	Tier-1	136.0	140.9	149.0	163.0	163.1	186.0	114%	186.0
	Tier-2	168.0	160.6	155.0	200.0	194.0	227.0	114%	227.0
	Total	**344.0**	**341.5**	**340.4**	**412.0**	**406.1**	**469.0**	**114%**	**469.0**
Tape	Tier-0	174.0	174.0	127.0	207.0	207.0	258.0	125%	258.0
	Tier-1	353.0	360.3	346.0	452.0	460.0	561.0	124%	561.0
	Total	**527.0**	**534.3**	**473.0**	**659.0**	**667.0**	**819.0**	**124%**	**819.0**

Table 6 ATLAS Collaboration resource requests for 2025 and C-RSG recommendations. For reference, 2023 and 2024 pledged resources are also given.

12 Resource Requests: ATLAS

Table 6 shows the resource requests submitted by the ATLAS Collaboration for 2025, the resources used in 2024 and the pledged resources for 2023 and 2024 along with the C-RSG recommendations. The experiment is asking for increases of 10% for CPU, 13.7% for disk and 24.2% for tape, in line with the expected increases in recorded data.

The ATLAS Collaboration makes the following assumptions on the running scenario for 2025: one full year of operation with an integrated luminosity of $L = 120$ fb^{-1}, pile-up of 62 collisions/bunch crossing and a total running time of 6.3×10^6 s for pp runs. With these running conditions ATLAS expects to record 14.5×10^9 events in the main trigger stream and an additional 9.5×10^9 events in a delayed stream.

Some uncertainty around the 2025 running conditions remains. If the readout limitations of the inner detector can be overcome, ATLAS may push to even higher $< \mu >$, potentially up to $< \mu >= 70$. Given a sufficient physics case is made and the system is found to be operating well within the expected margins, it is possible that the trigger rates for the physics main and delayed stream will be increased by a further 5%. These possibilities have not been explicitly accounted for in the resource estimations.

ATLAS full simulation remains a major CPU consumer. As in 2023, 1300 kHS23-years are assumed to come from beyond-pledge and HPC resources. These opportunistic resources continue to help greatly to broaden the ATLAS physics program. However, in a few months of 2023, the HPC resources fell below 1300 kHS23-years. If this trend continues, the projection will have to be reduced in 2026, which will require adjustments to computing priorities or a corresponding increase in pledged resources.

ATLAS expects a decrease in 2025 of RAW event size to 1.6 MB/event as a result of changes in the liquid-argon calorimeter (LAr) trigger data, but this can be enabled only after full commissioning of the new trigger system and sufficient operational experience.

Conclusions

The C-RSG considers that the ATLAS experiment's resource requests for 2025 shown in Table 6 are necessary to achieve the experiment's approved physics programme, and recommends that the C-RRB arranges for pledges correspondingly.

The C-RSG makes the following observations and recommendations:

ATLAS-1 The C-RSG recommends that the ATLAS Collaboration continues its effort on reduction of the raw event size.

ATLAS-2 The C-RSG recommends that the ATLAS Collaboration continues its effort on adoption of the DAOD_PHYS and DAOD_PHYSLITE data formats.

13 Resource Requests: CMS

The CMS Collaboration's resource requests for 2025 together with the C-RSG recommendations are shown in Table 7. The requested resources are comparable to the preliminary request provided to the C-RSG in autumn 2023, as reported in [7], with a slight increase in T0 disk by 2%.

CMS is asking for increases of 19% (+670 kHS23-years) for CPU, 19% (+62 PB) for disk and 27% (+187 PB) for tape, in line with the expected increase in data to be collected in 2025.

The CMS Collaboration makes the following assumptions on the running scenario for 2025: one full year of operation with an integrated luminosity of $L = 120\ \text{fb}^{-1}$, pile-up of 62 collisions/bunch crossing and a total running time of 6.3×10^6 s for pp runs and 1.4×10^6 s for Pb-Pb/O-O/Pb-O collisions.

The computing activities planned by CMS for 2025 are:

- Run 3 data taking: Prompt reconstruction at T0 of incoming events from the high-level trigger, followed by production of AOD, MiniAOD, and NanoAOD at T2 sites. The data collected during the beginning of the 2025 run will be reprocessed during the summer, while a full reprocessing will take place after final detector calibrations are derived.
- Run 3 MC production: A first round of sample production will take place to support detector studies and the development of reconstruction software, followed by a more substantial MC simulation for analysis.
- HL-LHC MC production: The simulated event samples to evaluate the performance of the Phase-2 upgrades will continue during 2025.
- Data analysis: Data analysis activities during 2025 will be focused on the data acquired during Run 3.

The T0 CPU need in 2025 is larger than that of their 2024 request, driven by the expected requirements for repacking of the RAW data and the prompt reconstruction, as well as the post-processing of the L1 trigger scouting data. One of the contributions to the iT0 disk size increase is an additional disk buffer to handle cases where the transfer of the RAW and derived data to T1 sites is interrupted.

The increases of CPU and disk resources at T1 and T2 sites are driven primarily by the larger Run 3 data that must be periodically reprocessed with the latest analysis algorithms and calibrations.

As in 2024, the CMS Collaborations in 2025 plans to use a portion (780 kHS23-years) of its expected beyond pledge compute capacity for MC production, or about a third of the total beyond-pledge CPU available to CMS in 2023.

CMS		2023			2024		2025		
		C-RSG recomm.	Pledged	Used	C-RSG recomm.	Pledged	Request	2025 req. / 2024 C-RSG	C-RSG recomm.
CPU	Tier-0	720	720	969	980	980	1180	120%	1180
	Tier-1	800	916	1173	930	1020	1100	118%	1100
	Tier-2	1350	1313	2544	1600	1484	1900	119%	1900
	HLT	n/a	n/a	0	n/a	n/a	n/a	n/a	n/a
	Total	**2870**	**2949**	4686	**3510**	**3484**	**4180**	**119%**	**4180**
	Others			0					
Disk	Tier-0	45.0	45.0	39.2	54.0	54.0	70.0	130%	70.0
	Tier-1	98.0	96.8	76.7	122.0	115.7	142.0	116%	142.0
	Tier-2	117.0	109.7	96.1	149.0	134.1	175.0	117%	175.0
	Total	**260.0**	**251.5**	**212.0**	**325.0**	**303.8**	**387.0**	**119%**	**387.0**
Tape	Tier-0	228.0	228.0	202.8	320.0	320.0	442.0	138%	442.0
	Tier-1	316.0	303.7	226.4	380.0	353.9	445.0	117%	445.0
	Total	**544.0**	**531.7**	**429.2**	**700.0**	**673.9**	**887.0**	**127%**	**887.0**

Table 7 CMS Collaboration resource requests for 2025 and C-RSG recommendations. For reference, 2024 is also given.

Conclusions

The C-RSG considers the CMS Collaboration's resource requests for 2025 shown in Table 7 necessary to achieve its approved physics programme, and recommends that the C-RRB arranges for pledges correspondingly.

The C-RSG makes the following request:

CMS-1 The C-RSG recommends that the CMS Collaboration continues its efforts in adoption of the NanoAOD data format.

CMS-2 The C-RSG recognises the effort to increase disk utilization on the T0 and T1 sites, and encourages CMS to raise this to the 90% limit considered optimal.

CMS-3 The C-RSG requests that the CMS Collaboration continue reporting on efforts to pinpoint and reduce the causes behind the relatively low 2023 CPU efficiency.

14 Resource Requests: LHCb

The computing resource requests submitted by the LHCb Collaboration for 2025 are shown in Table 8. The resources used in 2023, the pledged resources for 2023 and 2024 along with the C-RSG recommendations are also listed in the table.

LHCb requests for 2025 an increase of computing resources of 62% for CPU power, 57% for disk storage, and 46% for tape archival capacity, relative to values recommended by the C-RSG for 2024. This requested growth follows the LHCb computing model and is the result of a reassessed baseline for the needed resources in 2024, which takes into account the lack of substantial physics data collected during 2022 and 2023. The additional resources will be used to store and process the experimental data to be recorded in 2025 and to fulfill the growing data simulation needs of the experiment. The large growth in 2025 follows a modest increase in 2023 (about 10%) and no increase in 2024. Moreover, during the LHC Long Shutdown 3 (2026-2028) no resource increase is expected.

LHCb		2023 C-RSG recomm.	2023 Pledged	2023 Used	2024 C-RSG recomm.	2024 Pledged	2025 Request	2025 2025 req. / 2024 C-RSG	2025 C-RSG recomm.
CPU	Tier-0	215	215	258	174	174	283	163%	283
	Tier-1	707	598	652	572	542	928	162%	928
	Tier-2	391	434	492	319	394	518	162%	518
	HLT	50	50	0	50	0	50	n/a	50
	Total	**1363**	**1297**	**1402**	**1115**	**1110**	**1779**	**160%**	**1779**
	Others			26					
Disk	Tier-0	30.3	30.3	23.4	30.6	30.6	54.9	179%	54.9
	Tier-1	60.5	54.7	35.2	61.2	53.0	89.9	147%	89.9
	Tier-2	11.6	7.9	3.6	11.8	9.4	17.4	147%	17.4
	Total	**102.4**	**92.9**	**62.2**	**103.6**	**93.0**	**162.2**	**157%**	**162.2**
Tape	Tier-0	91.0	91.0	37.8	117.1	117.0	170.4	146%	170.4
	Tier-1	157.0	133.7	55.8	133.3	125.0	194.8	146%	194.8
	Total	**248.0**	**224.7**	**93.6**	**250.4**	**242.0**	**365.2**	**146%**	**365.2**

Table 8 LHCb Collaboration resource requests for 2025 and C-RSG recommendations. For reference, 2023 and 2024 resources are also given.

The CPU requirements are dominated by the MC simulation for Run 3 data, consuming 73% of the total LHCb CPU budget. User analysis activities account for 17% of the requested CPU, while data processing consumes the remaining 10%. LHCb is expecting minimal additional requests to produce simulations for Run 1 and Run 2 conditions in 2025 and beyond. The fractions of events produced with full and fast simulations in 2025 are assumed to be the same as in previous years (i.e. 36% and 64%, respectively). A new multithreaded framework for the simulation is under commissioning that should bring gains in memory usage and CPU performance.

CPU estimates for the event simulation, processing, and analysis times are based on extrapolations from Run 2 performance and expected software improvements, rather than on measurements using collision data or simulations with Run 3 conditions. Although the events are busier, the CPU event times for processing and simulating Run 3 data are assumed to be the same as for Run 2, which comes from improvements in the reconstruction and simulation frameworks. The analysis CPU time per event is expected to be reduced by a factor of two, thanks to the centralization of the analysis productions.

The projected increase in data storage capacity is driven by the integrated data volume expected to be recorded in 2025. For a projected pp run time of 6.3×10^6 s, and a nominal throughput from the trigger farm to the offline system of 10 Gb/s, a volume of 126 PB of tape space will be required to archive two replicas of the collected data (one at CERN and the other replica distributed among the T1 sites). The Collaboration requires 5 PB of disk space to keep on disk two instances of the data in a reduced format for analysis.

The LHC schedule foresees a heavy ion run in 2025. The reconstruction of the heavy ion events will be done in the online farm, concurrently with data-taking. If the CPU resources are not sufficient to keep up with the collection rate, the data will be buffered on disk and reconstructed during the end-of-year shutdown. The CPU work required to simulate heavy ion collisions is assumed to be 10% of that needed for the reconstruction of the real data, or 93 kHS23-year.

The NCBJ site in Poland was endorsed as a new T1 for LHCb at the WLCG Overview Board meeting in December 2023. It is expected to provide about 5% of the total resources at the T1 level. The IHEP site in China made very good progress to become a T1 site. All hardware is in place and configured. A shared network link of 100 Gb/s is established to CERN through LHCONE. However, a connection through the LHC private optical network (LHCOPN), with a minimal bandwidth of 10Gb/s, will be established before the site can be endorsed as a T1.

Although the use of HPC centers during 2023 was relatively low, the Collaboration continues investigating efforts to exploit this kind of facility. The DIRAC workload and data management system is being extended to access HPC resources within locally isolated networks. The exploitation of HPC facilities relies on local LHCb groups that contribute efforts to get allocations and configure the resources.

The LHCb Computing Coordinators reported to the C-RSG a reorganization of the computing and software effort in LHCb, creating a single body (a "Software and Computing Board") that is responsible for coordinating all software and computing activities and reporting directly to the LHCb Management Board. The new structure is expected to raise the profile of this effort within the Collaboration and assist the team in allocating the necessary personnel resources to this effort.

Conclusions

The C-RSG considers that the LHCb Collaboration's resource requests for 2025 are necessary to achieve the experiment's physics program.

The following recommendations are made:

LHCb-1 Estimates for the required CPU for data processing, simulation and analysis are based on extrapolations rather than on measurements on real data or simulations using the Run 3 nominal running conditions. The C-RSG recommends reassessing those parameters using the data collected in 2024 and the corresponding simulations.

LHCb-2 The coordination of the software and computing areas has been reorganized within LHCb to optimize resources, increase efficiency, and improve communication. The coordinator will become a member of the experiment's management board, raising the project's visibility within the Collaboration. The C-RSG requests that the Collaboration provides a report in the next scrutiny round on the effectiveness of this new organization in addressing the shortage of personnel for computing and software support.

15 Comments and Recommendations

The C-RSG completed this scrutiny round impressed with the preparations of the experiments for data-taking in 2025. The Collaborations are actively addressing long-term issues such as increasing the efficiency and turn-around of physics analysis and making more effective use of storage resources through the use of smaller data formats and new forms of access. They have been doing this in a context where new technologies such as GPUs are being introduced into the WLCG environment.

The C-RSG makes the following overall recommendations that it hopes will assist the Collaborations in this evolution:

ALL-1 The use of compact data formats, first introduced by CMS and more recently adopted by ATLAS, has reduced some pressure on disk resources while increasing throughput of physics analyses to published results. The C-RSG encourages all the Collaborations to continue to focus development efforts in this area both for the analysis of the Run 3 data and for the HL-LHC computing frameworks.

ALL-2 Increasingly the LHC Collaborations are relying on the use of HPC systems. These systems progressively provide a large fraction of their capabilities in form of accelerators, such as GPUs, rather than CPUs. Their effective and efficient utilisation by the LHC community is non-trivial and requires significant changes to the software architecture. Furthermore the connectivity to external networks, the authentication and authorisation services, and batch systems differ. This mandates careful adaptation of workloads and workflow management systems. For this, additional sustained investment in expert developers is essential. The C-RSG recommends identifying adequate mechanisms to fund these activities.

ALL-3 To reduce confusion in terminology, the C-RSG requests that when experiments refer to a set of data it uses the term “instance” rather than “copy” or “replica”. This would mean that a data set stored twice (e.g., to ensure data integrity) would be described as two instances.

References

[1] CERN. LHC Schedule. http://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm.

[2] ALICE Collaboration. ALICE Computing Resources Usage January through December 2023, Resource Requirements for 2025. Submitted 13 February 2024.

[3] ATLAS Collaboration. ATLAS computing status and plans – Report to the C-RSG – April 2024. Submitted 13 February 2024.

[4] CMS Collaboration. Spring 2024 scrutiny – CMS utilization report 2023 and resource request for 2025. Submitted 13 February 2024.

[5] LHCb Collaboration. LHCb computing resource usage in 2023. *LHC-PUB-2024-003*, submitted 13 February 2024.

[6] LHCb Collaboration. LHCb computing resources: 2025 requests. *LHC-PUB-2024-002*, submitted 13 February 2024.

[7] Computing Resources Scrutiny Group. C-RSG autumn 2023 report. *CERN-RRB-2023-074*, 9 October 2023.

[8] WLCG CRIC. Computing Resource Information Catalogue. http://wlcg-cric.cern.ch.

[9] EGI Foundation. European Grid Initiative. https://www.egi.eu/federation.

[10] LHCb Collaboration. LHCb DIRAC Accounting Portal. https://lhcb-portal-dirac.cern.ch.