LHCb Computing Resources: 2020 requests and preview of the subsequent years

LHCb Public Note

Issue: 0
Revision: 0

Reference: LHCb-PUB-2019-003
Created: 4th February 2019
Last modified: 18th February 2019

Prepared By: LHCb Computing Project
C. Bozzi/Editor

Abstract

This document presents the computing resources needed by LHCb in 2020, as resulting from the current experience and the foreseen computing activities. It also gives a preview of the computing resources for Run3 and LS3, based on the recently submitted Computing Model Technical Design Report for the LHCb Upgrade.

Table of Contents

1. INTRODUCTION 5

2. RESOURCES NEEDED IN 2020 5

3. SUMMARY OF 2020 REQUESTS 6

4. COMPUTING MODEL FOR THE LHCB RUN 3 UPGRADE 9

4.1. BASIC FEATURES OF THE LHCB UPGRADE COMPUTING MODEL 9
4.2. MAJOR DRIVERS IN OFFLINE RESOURCE REQUIREMENTS 10
4.3. OFFLINE RESOURCE NEEDS 11
4.4. CONCLUSION 13

List of Tables

Table 2-1: Estimated CPU work needed for the different activities 5

Table 2-2: Disk Storage needed for the different categories of LHCb data. 6

Table 2-3: Tape Storage needed for the different categories of LHCb data....................... 6

Table 3-1: CPU power requested at the different Tier levels. 8

Table 3-2: LHCb Disk request for each Tier level. For countries hosting a Tier1, the Tier2 contribution could also be provided at the Tier1. 8

Table 3-3: LHCb Tape request for each Tier level. 8

Table 4-1: [taken from LHCb-TDR-018] Extrapolated throughput to TAPE and to disk (after offline processing), for the FULL, TURBO and CALIBRATION streams during the Upgrade.. 10

Table 4-2: [taken from LHCb-TDR-018]: Summary of the LHCb upgrade computing model requirements. Top section: main assumptions of the model. Bottom section: resource requirements, with yearly growth factors.. 11

Table A-1: Summary of the scrutinized and pledged resources in 2019, evolution of requests in 2020, 2021.. 14

1. Introduction

The computing resources needed by LHCb in the 2019 and 2020 WLCG years[1] were shown in a document (LHCb-PUB-2018-001), that was subsequently reviewed by the C-RSG in Spring 2018. The C-RRB made a preliminary endorsement of the 2019 requests in April 2018.

The LHCb computing model, its implementation, recent changes and processing plans were described in LHCb-PUB-2018-001. In the following, the resources that will be needed and a summary of the requests are given in Sections 2 and 3, respectively. The LHCb computing model and its parameters did not change in the last six months. Therefore, the 2019 and 2020 requests contained in this document are unchanged with respect to those of LHCb-PUB-2018-001. Finally, Section 4 reports a preview of the evolution of the computing model for the Run 3 upgrade of LHCb, with emphasis on the main changes that drive the determination of the offline computing resources.

2. Resources needed in 2020

There are no changes in the resource requests with respect to what has already shown in the previous report. Table 2-1 presents, for the different activities, the CPU work estimates when applying the model defined in LHCb-PUB-2018-001.

CPU Work in WLCG year (kHS06.years)	**2020**
Prompt Reconstruction	0
First pass Stripping	0
Full restripping	0
Incremental (re-)stripping	13
Processing of heavy ion collisions	0
Simulation	571
VoBoxes and other services	4
User Analysis	42
Total Work (kHS06.years)	**631**

Table 2-1: Estimated CPU work needed for the different activities

Table 2-2 presents, for the different data classes, the forecast total disk space usage at the end of 2020 when applying the baseline model described in the previous section. Table 2-3 shows, for the different data classes, the forecast total tape usage at the end of 2020.

---

[1] For the purpose of this document a given year always refers to the period between April 1st of that year and March 31st of the following year.

Disk storage usage forecast (PB)	**2020**
Stripped real data	24.2
TURBO Data	4.2
Simulated Data	20.7
User Data	2.2
Heavy Ion Data	4.2
RAW and other buffers	1.6
Other	0.6
Total	**57.6**

Table 2-2: Disk Storage needed for the different categories of LHCb data.

Tape storage usage forecast (PB)	**2020**
Raw Data	40.8
RDST	14.8
MDST.DST	0.7
Heavy Ion Data	3.7
Archive	31.6
Total	**91.6**

Table 2-3: Tape Storage needed for the different categories of LHCb data.

3. Summary of 2020 requests

Table 3-1 shows the CPU requests at the various tiers, as well as for the HLT farm and Yandex. We assume that the HLT and Yandex farms will provide the same level of computing power as in the past, therefore we subtract the contributions from these two sites from our requests to WLCG. The required resources are apportioned between the different Tiers taking into account the capacities that are already installed.

The disk and tape estimates shown in previous section have to be broken down into fractions to be provided by the different Tiers using the distribution policies described in LHCb-PUB-2013-002. The results of this sharing are shown in Table 3-2 and Table 3-3. It should be noted that, although the total storage capacity is given globally for the 8 Tier1 sites pledging resources to LHCb, it is mandatory that the sharing of this storage between Tier1 sites remains very similar from one year to another: the annual increments are small, existing data is expected to remain there, runs assigned to each Tier1 are also expected to be reprocessed there and the analysis data is stored at these same Tier1s. There is a level of flexibility offered when replicating the data to a second Tier1 that takes into account the available space. However, a baseline increase is mandatory at all sites, otherwise they can no longer be used for the placement of new data.

LHCb will be upgraded during the LHC shutdown of 2019-2020, and will resume data taking in 2021. There will be many changes in the computing activities in the upgrade era, that will need to be prepared and properly tested before 2021. These activities are reported in the Technical Design Report of software and computing for the LHCb upgrade[2]. The part more related to the computing model and the required computing resources in Run3 and LS3 is reported in the LHCb Upgrade Computing Model TDR[3], with excerpts given in the next section.

---

[2] CERN-LHCC-2018-007 ; LHCB-TDR-017; [http://cdsweb.cern.ch/record/2310827]

[3] CERN-LHCC-2018-014; LHCB-TDR-018; [https://cds.cern.ch/record/2319756]

CPU Power (kHS06)	**2020**
Tier 0	98
Tier 1	328
Tier 2	185
Total WLCG	**611**
HLT farm	10
Yandex	10
Total non-WLCG	**20**
Grand total	**631**

Table 3-1: CPU power requested at the different Tier levels.

Disk (PB)	**2020**
Tier0	17.2
Tier1	33.2
Tier2	7.2
Total	**57.6**

Table 3-2: LHCb Disk request for each Tier level. For countries hosting a Tier1, the Tier2 contribution could also be provided at the Tier1.

Tape (PB)	**2020**
Tier0	36.1
Tier1	55.5
Total	**91.6**

Table 3-3: LHCb Tape request for each Tier level.

4. Computing model for the LHCb Run 3 Upgrade

As mentioned in the previous report, the LHCb detector will undergo a major upgrade for the data taking during the LHCb Run 3 and beyond, with major changes in sub-detectors and trigger that will impact the computing model of the experiment.

The Computing Model for the LHCb Upgrade and its physics foundations are discussed at length in the recently released Technical Design Report (LHCb-TDR-018). The rest of this section presents a brief recap of the basic feature of the LHCb Upgrade Computing model, and the required offline computing resources.

4.1. Basic features of the LHCb Upgrade Computing Model

The novel concepts that were developed and implemented during the Run 2 data taking will become predominant.

- The splitting of the High Level Trigger in two parts, synchronous (HLT1) and asynchronous (HLT2) with data taking, enables the final detector alignment and calibration to be performed online in real time, thereby allowing for an offline-quality event reconstruction in the HLT and avoiding almost completely the necessity of costly offline reconstruction.
- The trigger system will be entirely based on software. This will increase the trigger efficiency for most of the physics channels by at least a factor 2. This, coupled with a five-fold increase of the instantaneous luminosity and the fact that the triggered events are highly pure in signal, will increase the event throughput to offline storage by at least an order of magnitude.
- From the processing flow point of view, the majority (70%) of triggered events will be sent to the TURBO stream, where only high-level information is saved to offline and the raw events are discarded. This is the case of events selected by exclusive trigger lines, as in the case of e.g. charm decays.
- More inclusive trigger lines as well as calibration lines (about 30% of the total) will be saved in the “classic” FULL and CALIBRATION stream, where the entire event is persisted. The FULL stream is then further processed offline, where slimming and filtering criteria are applied, aimed respectively at saving only the interesting parts of the event and to increasing signal purities.
- The “classical” offline event reconstruction will be run only part of the data corresponding to use cases such as detector commissioning, reconstruction studies, and to reconstruct streams that cannot be reconstructed online (e.g. due to timing constraints). These involve mainly events in the CALIBRATION stream

The CPU needs in the LHCb Upgrade era will be therefore dominated by Monte Carlo simulation. Given that the number of events to be simulated scales with the integrated luminosity, and that the simulation time scales with pile-up, the CPU requirements will scale accordingly. Therefore, a full Geant4-based simulation of the detector will not be possible as it would require an increase of at least one order of magnitude in the resources that might be made available to LHCb in Run 3. Faster simulation options are being introduced to mitigate the CPU requirements (see below).

The main data workflows will then be

- the processing of this TURBO stream data to convert the LHCb-specific online format to the ROOT I/O-based offline format, and the subsequently streaming of these data. This workflow currently accounts for 0.01% of the CPU work on the Grid and will stay negligible in Run 3 and beyond.
- The slimming and filtering of data in the FULL stream, and their subsequent streaming. Also in this case, the expected CPU work on the Grid is no more than a few percent of the total
- The processing of the CALIBRATION stream

4.2. Major drivers in offline resource requirements

As already mentioned, the production of simulated events dominates the offline CPU computing needs. LHCb is mitigating this by exploiting faster simulation options. As already mentioned in previous reports, several of them are already in production. Speed-up factors between 5 and 10 are achievable with these fast simulation options.

More fast simulation workflows are under preparation and will be used in Run 3, or before if they become available, such as the utilization of shower libraries and/or machine learning techniques to parametrize the response of the calorimeters. A full parametric simulation, based on the DELPHES package, is also in development.

In all the above cases, the simulation workflow will stay as in Run 2. In a first step, events are generated and the detector response is simulated. This accounts for the vast majority of computing work. The subsequent steps are the digitization of the detector signals, the trigger emulation and the event reconstruction.

The simulation is currently being moved to the new Gaudi software framework that has been developed to allow running the LHCb application in a multi-threaded environment. This allows a significant reduction of the memory footprint, thereby opening the possibility to use resources, such as HPC farms and many-core architectures, where the memory per logical core is smaller than that of the usual grid computing nodes.

The storage needs are dominated by data and crucially depend on the HLT output bandwidth. A bandwidth of 10GB per live second of LHC is deemed sufficient to carry on the physics program of the LHCb Upgrade. While associated tape needs are incompressible, some mitigations are possible for disk. As already mentioned, about 70% of triggered events will be saved in the "light" TURBO format. However, the majority (6.5GB/s out of 10GB/s) of the bandwidth is taken by the remainder 30% of events in the FULL and CALIBRATION streams, where the entire event is saved. The events in these two streams will therefore be slimmed and/or filtered offline, as in the stripping process that has been already in place, so that the total (logical) bandwidth to be saved of disk is only 1/3 of the original. Table 4-1, taken from LHCB-TDR-018, shows the extrapolated throughputs to tape and disk for the three data streams that will be used in the LHCb Upgrade.

stream	rate fraction	TAPE throughput (GB/s)	TAPE bandwidth fraction	DISK trhoughput (GB/s)	DISK bandwidth fraction
FULL	26%	5.9	59%	0.8	22%
Turbo	68%	2.5	25%	2.5	72%
Calibration	6%	1.6	16%	0.2	6%
Total	100%	10.0	100%	3.5	100%

Table 4-1: [taken from LHCb-TDR-018] Extrapolated throughput to TAPE and to disk (after offline processing), for the FULL, TURBO and CALIBRATION streams during the Upgrade.

The impact of simulated events on storage requests is small, since all data produced during the intermediate production steps are not saved and the simulation output will be migrated to the microDST format, thus achieving a size reduction per event of a factor up to 20. Also, analysis-dependent filtering criteria that are already applied in a fraction of the current productions, will be further exploited to reduce the number of events written on storage.

4.3. Offline resource needs

The basic assumptions that go into the calculation of the offline resource needs are

- Increased output bandwidth as a consequence of the increased luminosity and trigger rate, mitigated by processing online as much data as possible in the TURBO stream, and by an aggressive offline data reduction of the FULL and CALIBRATION streams
- (following past trends) simulation of Run 2 data continuing during LS2 and ramp down between 2021 and 2023, and the bulk of the simulation of a year of data taking starting in the following year
- A mixture of full/fast/parametric simulations, with the fraction of the fast and parametric ones gradually increasing during Run 3
- Most of the simulation output written in the microDST format
- 2021 will be a commissioning year for both LHC and LHCb; data taking at nominal conditions is expected in 2022 and 2023

The basic parameters and the resource requirements, including yearly growth factors, are reported in Table 4-2

Model assumptions						
L ($cm^{-2}s^{-1}$)	2×10^{33}					
Pileup	6					
Running time (s)	5×10^{6} (2.5×10^{6} in 2021)					
Output bandwidth (GB/s)	10					
Fraction of Turbo events	73%					
Ratio Turbo/FULL event size	16.7%					
Ratio full/fast/param. simulations	40:40:20					
Data replicas on tape	2					
Data replicas on disk	2 (Turbo); 3 (FULL, TurCal)					
Resource requirements						
WLCG Year	Disk (PB)		Tape (PB)		CPU (kHS06)	
2021	66	1.1	142	1.5	863	1.4
2022	111	1.7	243	1.7	1.579	1.8
2023	159	1.4	345	1.4	2.753	1.7
2024	165	1.0	348	1.0	3.467	1.3
2025	171	1.0	351	1.0	3.267	0.9

Table 4-2: [taken from LHCb-TDR-018]: Summary of the LHCb upgrade computing model requirements. Top section: main assumptions of the model. Bottom section: resource requirements, with yearly growth factors.

Despite the increase in the data volume from the detector of more than one order of magnitude, the offline resource increase is much more modest and does not exceed the funding model used for the original LHCb experiment by a large amount. The yearly evolution of the storage needs is much more pronounced during data taking years, with negligible increase during LS3, due to the fact that these needs are dominated by the HLT output bandwidth and no data reprocessing is foreseen. The growth profile of CPU requirements follows the requirements for Monte Carlo production, where the simulation of data taking conditions for a given year start typically the following year. This suggests that the strategy for resource procurement might be optimised by considering the entire period from the start of Run 3 (in 2021) until the end of LS3 (in 2025) and by prioritizing the type of resource needed.

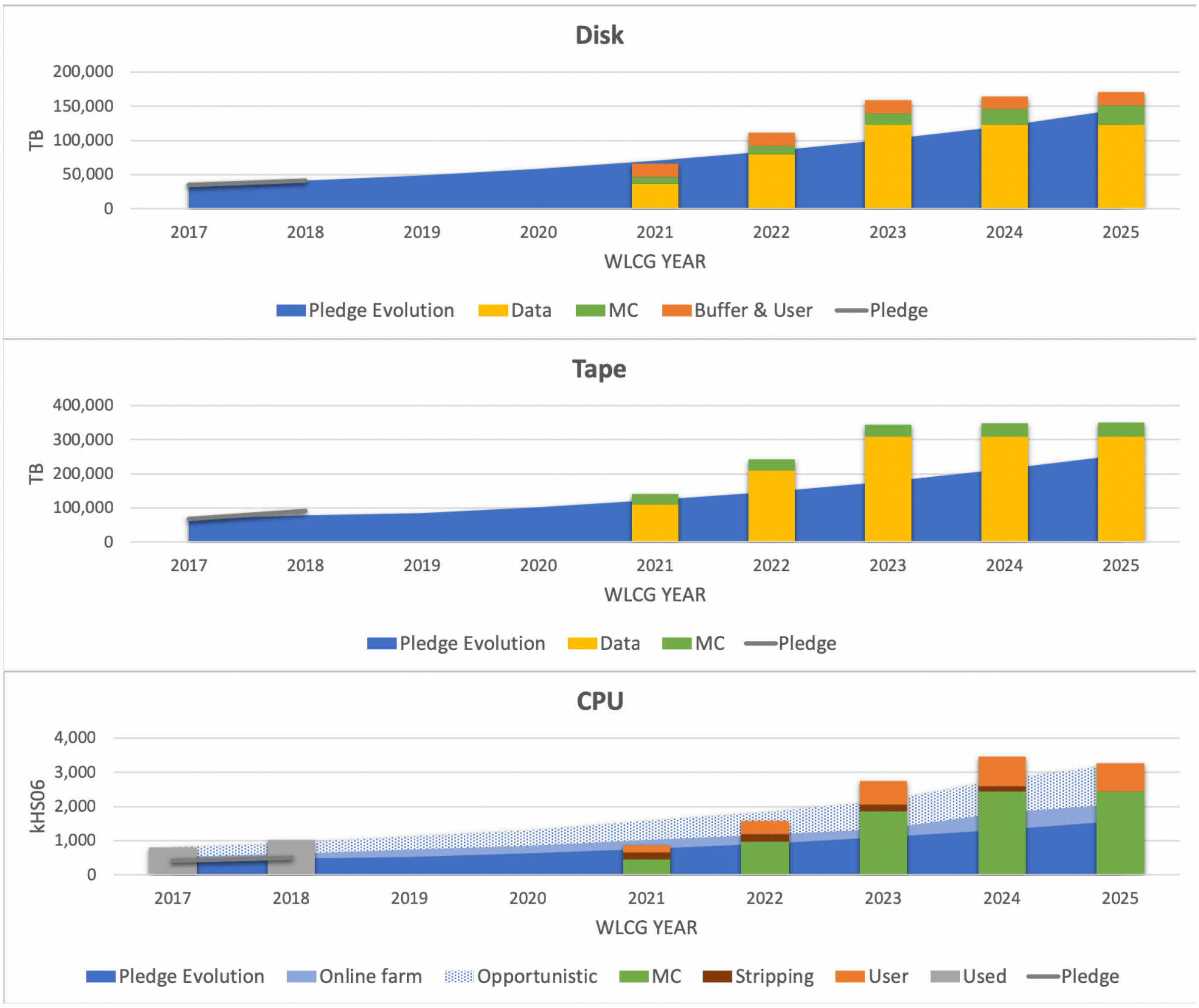


Figure 1 [taken from LHCb-TDR-018]: Baseline requirements for disk (top), tape (center), CPU (bottom) as a function of year. The breakdown is given by usage type. The current WLCG pledges for LHCb and their evolution to the Upgrade assuming the constant budget model are also shown. For CPU, the dark blue solid band and the light blue shaded band indicate the expected evolution of HLT farm and other opportunistic resources respectively.

It is assumed that the storage resources are to be entirely procured within WLCG. For CPU resources, it is expected that the HLT farm will contribute to offline activities during end-of-year shutdowns of the LHC, corresponding to an integrated availability of 30%, which increases to 50% during LS3, when major infrastructure interventions will take place. It is not possible to reliably estimate the additional opportunistic CPU resources that might be available. Figure 1 (bottom plot) assumes that they will scale according to the constant budget model, however there is no guarantee that this will be the case.

4.4. Conclusion

The LHCb Upgrade experiment will collect data volumes that exceed by more than an order of magnitude those of the original LHCb experiment. Consequently, the offline resources that would be required, in absence of mitigation measures, would be significantly larger than those achievable with the resource evolution scenario that has been employed so far. Innovative practices are being put in place by the collaboration to significantly reduce the storage and computing requirements of the LHCb Upgrade experiment compared with a naive scaling of those utilised by the LHCb experiment.

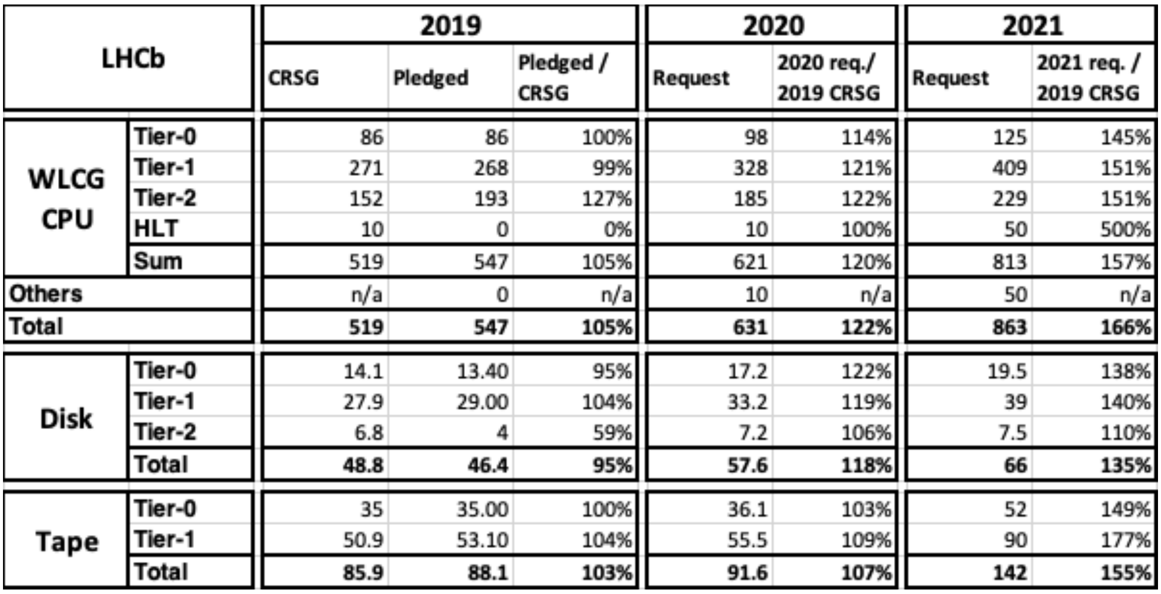
The computing model and its associated computing requests have been briefly summarized here; they are discussed in detail in LHCb-TDR-018.

The time profile of the computing requests from 2021 until 2025 show moderate increases during the Run3 data-taking years, followed by modest increases during LS3. The increase of CPU requests is slightly shifted with respect to storage requests. The “constant budget” model is not sufficient for resource provisioning during the Run3 data taking years. However, this model is violated by no more than a few tens of percent in any given year and any kind of resource. Moreover, the requests at the end of LS3 roughly correspond to what would be made available by a constant budget model. This suggests that a global 5-year planning of investment within a given envelope might be beneficial in optimizing the spending profiles and the procurement of the resources needed by the experiment.

LHCb is aware of the R&D work that has started in view of similar issues that the general-purpose detectors ATLAS and CMS will face in the HL-LHC era. While the timescale for such developments is necessarily that of Run 4 and beyond, LHCb welcomes any efforts that might possibly reduce costs and enable optimizations on the Run 3 timescale (such as the WLCG/DOMA initiative). LHCb encourages these initiatives, is very willing to contribute to them, and to continue the experiment’s tradition of being an innovator and “early adopter” of new computing resource practices.

Appendix

Following up a request made by the C-RSG in the previous scrutiny round, Table A-1 provides a summary of the scrutinized (C-RSG) and pledged resources in 2019, as well as the evolution of the requests in 2020 and 2021 and their corresponding increases with respect to the scrutinized 2019 requests. For 2021, a constant budget model has been used for the WLCG CPU requests.



LHCb		2019			2020		2021	
		CRSG	Pledged	Pledged / CRSG	Request	2020 req./ 2019 CRSG	Request	2021 req. / 2019 CRSG
WLCG CPU	Tier-0	86	86	100%	98	114%	125	145%
	Tier-1	271	268	99%	328	121%	409	151%
	Tier-2	152	193	127%	185	122%	229	151%
	HLT	10	0	0%	10	100%	50	500%
	Sum	519	547	105%	621	120%	813	157%
Others		n/a	0	n/a	10	n/a	50	n/a
Total		**519**	**547**	**105%**	**631**	**122%**	**863**	**166%**
Disk	Tier-0	14.1	13.40	95%	17.2	122%	19.5	138%
	Tier-1	27.9	29.00	104%	33.2	119%	39	140%
	Tier-2	6.8	4	59%	7.2	106%	7.5	110%
	Total	**48.8**	**46.4**	**95%**	**57.6**	**118%**	**66**	**135%**
Tape	Tier-0	35	35.00	100%	36.1	103%	52	149%
	Tier-1	50.9	53.10	104%	55.5	109%	90	177%
	Total	**85.9**	**88.1**	**103%**	**91.6**	**107%**	**142**	**155%**

Table A-1: Summary of the scrutinized and pledged resources in 2019, evolution of requests in 2020, 2021.