Computing Resources Scrutiny Group Report

Pekka K. Sinervo, C.M., FRSC

University of Toronto

For the Computing Resources Scrutiny Group

October 29, 2019

C-RSG membership

C Allton (UK)	J Hernandez (Spain)
V Breton (France)	J Kleist (Nordic countries)
G Cancio Melia (CERN)	H Meinhard (CERN, scient. secr.)
P Christakoglou (Netherlands)	P Sinervo(Canada)
A Connolly (USA)	V Vagnoni (Italy)
F Gaede (Germany)	

- V Vagnoni is the new representative for Italy. He had observed the spring scrutiny and was an active member this fall
- C-RSG thanks the experiment representatives and to CERN management for their support

Fall 2019 Scrutiny Process

- The four LHC experiments gave updates on their computing and data processing activities and plans
  - Described computing activities for 2019 year (April 2019 – March2020)
  - Updated plans for the 2020 year – pledges approved at Spring 2019 RRB meeting
  - Updated estimates for 2021 year (April 2021 – March 2022)

- No surprises for 2019 and 2020 years
  - Continue data processing and scientific analysis of Run 2 data
  - Preparations continuing for Run 3 and some work on HL-LHC

- 2021 presents greater uncertainty
  - Biggest uncertainty is volume of LHC data

Resource Requirements for 2020 and Estimates for 2021

- 2020 is part of Long Shutdown 2
- Total increases below “flat budget model”
- Computing models being changed for Run-3, so some uncertainty for 2021 and beyond

Alice Requests for 2020 and Estimates for 2021

ALICE		2019		2020			2021	
		CRSG recomm.	Pledged	Request	2020 req. /2019 CRSG	C-RSG recomm.	Request	2021 req. /2020 CRSG
CPU	Tier-0	430	350	350	81%	350	471	135%
	Tier-1	365	331	365	100%	365	498	136%
	Tier-2	376	370	376	100%	376	515	137%
	HLT	n/a	n/a	n/a	n/a	0		
	Total	**1171**	**1051**	**1091**	**93%**	**1091**	**1484**	**136%**
	Others							
Disk	Tier-0	34.3	31.2	31.2	91%	31.2	45.5	146%
	Tier-1	37.9	35.1	44.0	116%	44.0	53.3	121%
	Tier-2	33.9	33.5	39.0	115%	39.0	44.8	115%
	Total	**106.1**	**99.8**	**114.2**	**108%**	**114.2**	**143.6**	**126%**
Tape	Tier-0	44.2	44.2	44.2	100%	44.2	80.0	181%
	Tier-1	37.7	41.1	37.7	100%	37.7	55.0	146%
	Total	**81.9**	**85.3**	**81.9**	**100%**	**81.9**	**135.0**	**165%**

- Flat utilization in 2020, allows for 3 passes through p-p and Pb-Pb data
- ”Flat budget” increases for disk at T1 & T2
  - T0 increase in disk in 2019 sufficient for 2020 operations
  - No additional tape space for 2020
- Increases in CPU and storage in 2021
  - Run 3 data-taking
  - Increased volume due to upgrades

ATLAS Requests for 2020 and Estimate for 2021

ATLAS		2019		2020			2021	
		CRSG recomm.	Pledged	Request	2020 req. /2019 CRSG	C-RSG recomm.	Request	2021 req. /2020 CRSG
CPU	Tier-0	411	411	411	100%	411	550	134%
	Tier-1	1057	1083	1057	100%	1057	1230	116%
	Tier-2	1292	1293	1292	100%	1292	1500	116%
	HLT	n/a	0	0	n/a	0	0	n/a
	Total	**2760**	**2787**	**2760**	**100%**	**2760**	**3280**	**119%**
	Others			0		0%		
Disk	Tier-0	27.0	26.0	27.0	100%	27.0	30.0	111%
	Tier-1	88.0	94.4	88.0	100%	88.0	107.0	122%
	Tier-2	108.0	101.2	108.0	100%	108.0	132.0	122%
	Total	**223.0**	**221.6**	**223.0**	**100%**	**223.0**	**269.0**	**121%**
Tape	Tier-0	94.0	94.0	94.0	100%	94.0	97.0	103%
	Tier-1	221.0	216.8	221.0	100%	221.0	249.0	113%
	Total	**315.0**	**310.8**	**315.0**	**100%**	**315.0**	**346.0**	**110%**

- 2020 “flat-flat” growth in CPU, given LS2 and Run-3 preparations
  - Driven by 20B MC event simulation
  - Large T2 utilization
- 2021 disk and tape storage modest increase
  - Tape “gap” of 230 PB used vs 311 pledged
  - Expects to need that over next 2 years

CMS Requests for 2020 and Estimates for 2021

CMS		2019		2020			2021	
		CRSG recomm.	Pledged	Request	2020 req. /2019 CRSG	C-RSG recomm.	Request	2021 req. /2020 CRSG
CPU	Tier-0	423	423	423	100%	423	517	122%
	Tier-1	650	620	650	100%	650	650	100%
	Tier-2	1000	960	1000	100%	1000	1200	120%
	HLT	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	Total	**2073**	**2003**	**2073**	**100%**	**2073**	2367	114%
	Others						50	
Disk	Tier-0	26.1	26.1	26.1	100%	26.1	31.0	119%
	Tier-1	68.0	63.4	68.0	100%	68.0	77.0	113%
	Tier-2	78.0	72.0	78.0	100%	78.0	93.0	119%
	Total	**172.1**	**161.5**	**172.1**	**100%**	**172.1**	201.0	117%
Tape	Tier-0	99.0	99.0	99.0	100%	99.0	144.0	145%
	Tier-1	220.0	188.8	220.0	100%	220.0	245.0	111%
	Total	**319.0**	**287.8**	**319.0**	**100%**	**319.0**	**389.0**	**122%**

- 2020 requests constant resources
  - Sufficient for Run 2 legacy work
  - Prepartions for Run 3
- 2021 increases are driven by Run 3 assumptions
  - Increase in T0 tape based on contingency assumptions of 42 fb-1 of data delivered by the LHC

LHCb Requests for 2020

LHCb		2019		2020			2021	
		C-RSG recomm.	Pledged	Request	2020 req. /2019 CRSG	C-RSG recomm.	Estimate	2021 est. /2020 CRSG
CPU	Tier-0	86	86	98	114%	98	112	114%
	Tier-1	271	268	328	121%	328	367	112%
	Tier-2	152	193	185	122%	185	205	111%
	HLT	10	10	10	100%	10	50	500%
	Total	**519**	**557**	**621**	**120%**	**621**	**734**	**118%**
	Others		10	10			50	
Disk	Tier-0	14.1	13.4	17.2	122%	17.2	20.7	120%
	Tier-1	27.9	29	33.2	119%	33.2	41.4	125%
	Tier-2	6.8	4	7.2	106%	7.2	8	111%
	Total	**48.8**	**46.4**	**57.6**	**118%**	**57.6**	**70.1**	**122%**
Tape	Tier-0	35.0	35	36.1	103%	36.1	56.0	155%
	Tier-1	50.9	53.1	55.5	109%	55.5	96.0	173%
	Total	**85.9**	**88.1**	**91.6**	**107%**	**91.6**	**152.0**	**166%**

- 2020 usages will be consistent with "flat budget " model
  - Expect increased use of HLT farm
- 2021 large increase in disk and tape resources
  - Reflects the change in computing model
  - Note that the tape increase of 55-73% is based on 3 fb^{-1} scenario
  - A further increase in tape (80 PB) would be needed to store data in 7 fb^{-1} scenario

C-RSG Summary

- Overall picture for 2019 and 2020 is consistent with plans
  - Legacy production of Run 2 data dominates
  - Preparation for Run 3 and HL-LHC moving forward
  - Plans for 2020 fit within the resources pledged to all experiments
  - C-RSG recommends that these be made available
- 2021 resource estimates still evolving, but increasing confidence
  - Work underway across all experiments to increase efficiency of utilization
- Tape utilization is still one area where further work is being done
- Overall, well positioned for Spring 2020 scrutiny

Comments and Recommendations

ALL-1 The C-RSG reaffirms that the computing resources requested by the four collaborations and pledged by the WLCG for the 2020 year are essential to address their approved physics programs.

ALL-2 The C-RSG recommends that the collaborations use a common approach to estimating tape resources at the Tier-0 and Tier-1 centers. The actual tape resources required for data storage should be the key driver for the future estimates, and the logistics of ``repacking'' data should be coordinated with the sites providing the tape storage. The scrutiny group understands that the WLCG Management Board is addressing this issue. The C-RSG notes that several collaborations have not used all of the Tier-1 tape resources pledged to them over the last several years.

ALL-3 The computing resources required for 2021 may exceed the available capabilities in the event that the amount of data collected by the experiments significantly exceeds baseline. The C-RSG recommends that a common strategy be developed to mitigate this, such as preparing to have Tier-0 provide the necessary tape resources to temporarily ``park'' the data until processing and disk resources become available.