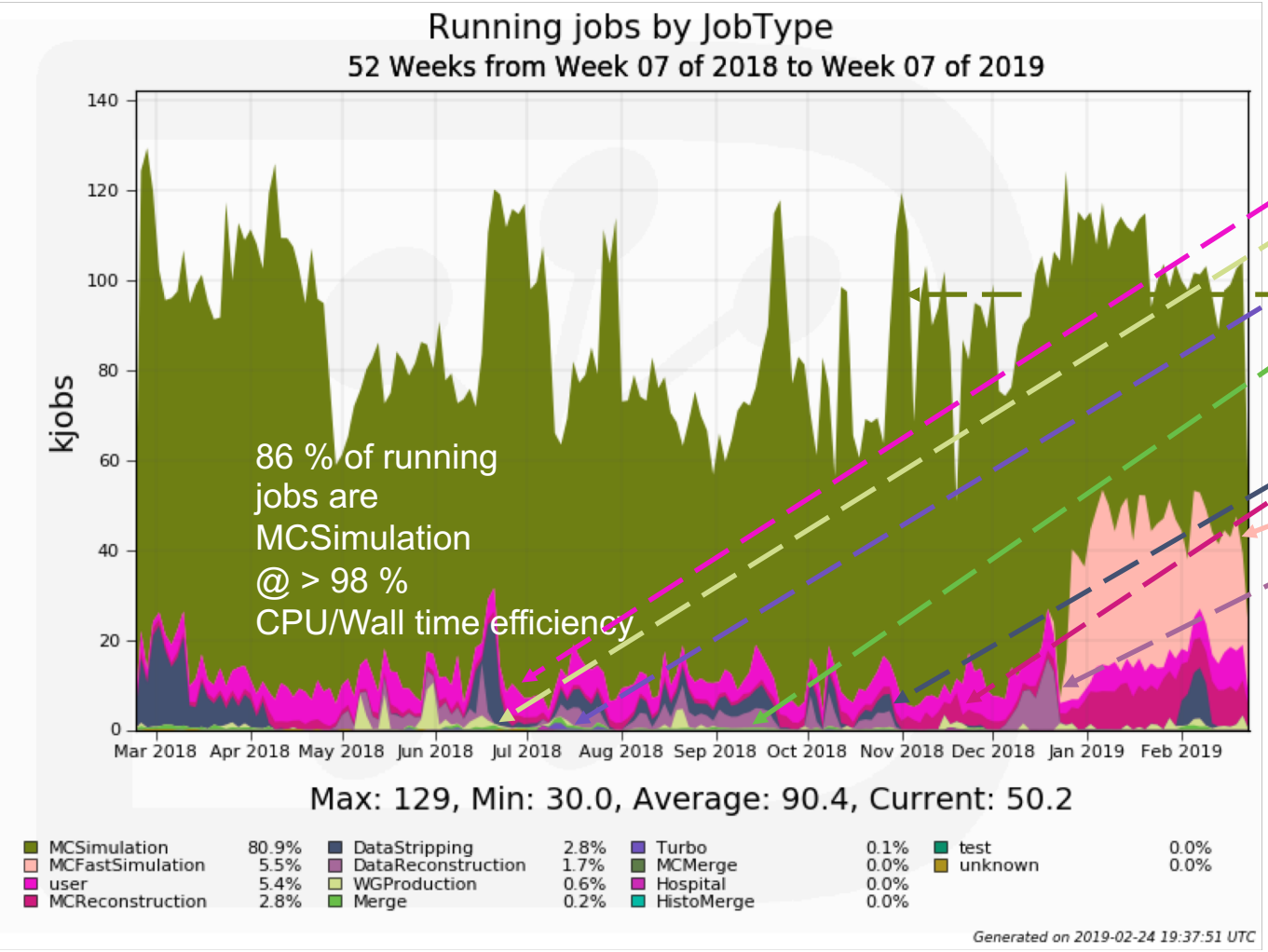
LHCb Software & Computing

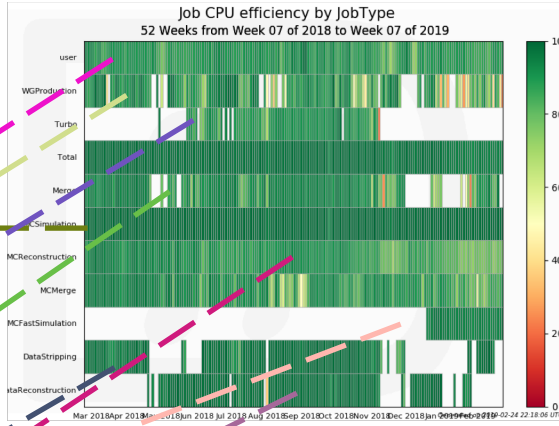
Concezio Bozzi

WLCG/LHCC Referees

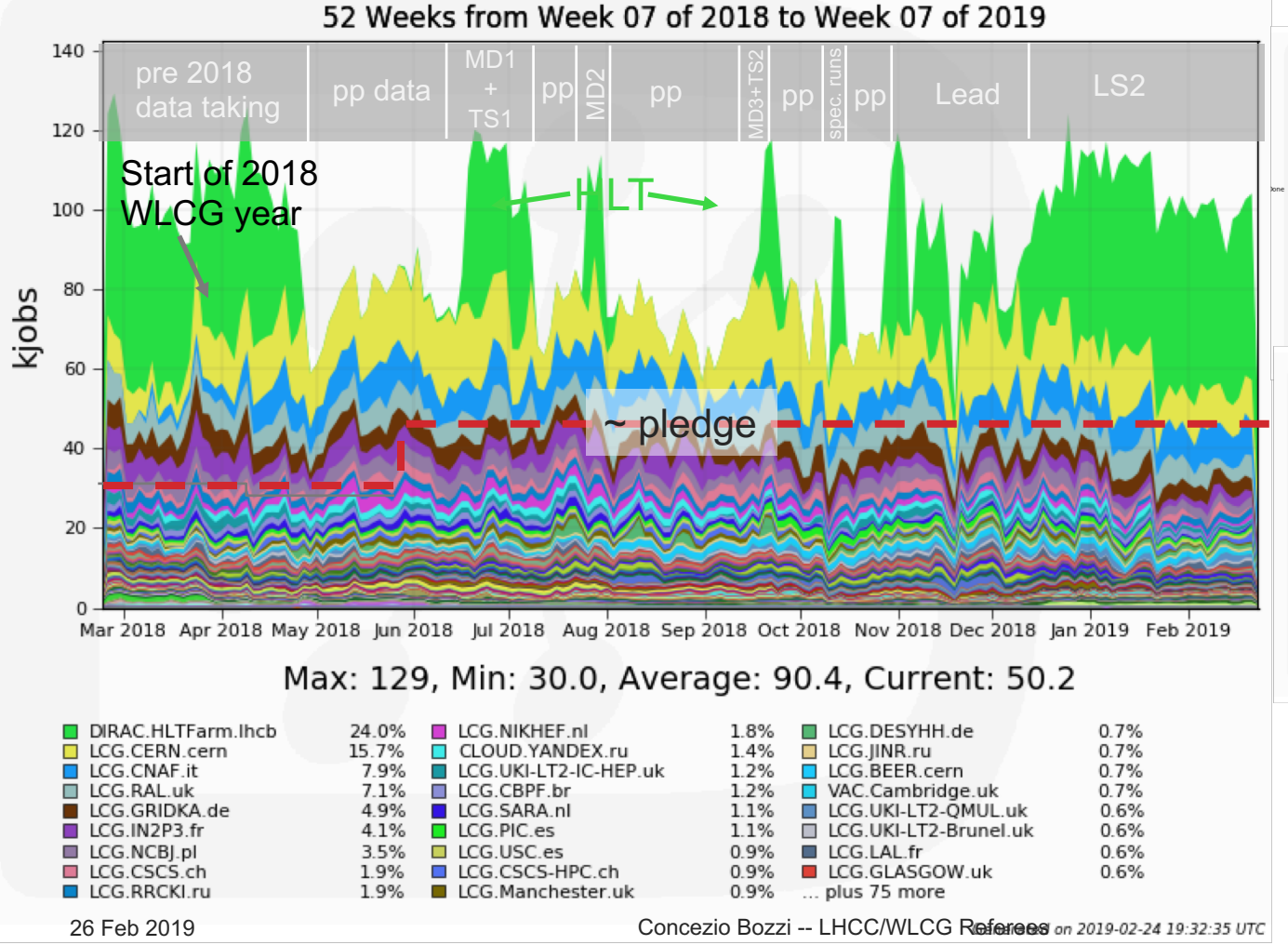
26 February 2019

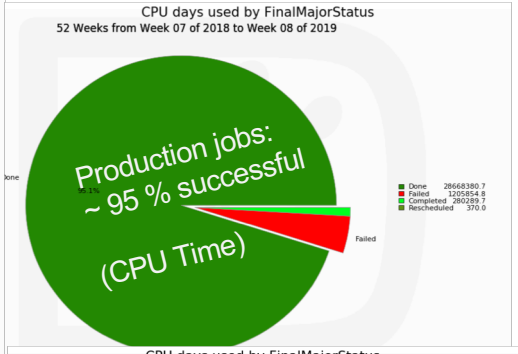
Operations

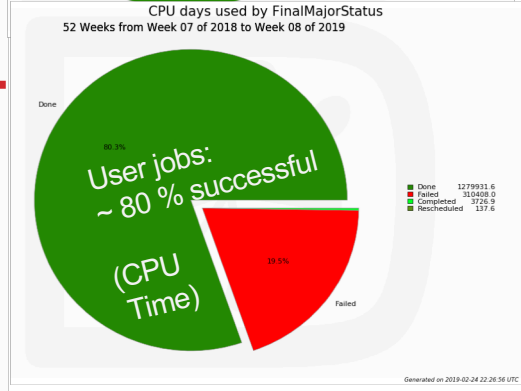

Running jobs by JobType
52 Weeks from Week 07 of 2018 to Week 07 of 2019
kjobs
140
120
100
80
60
40
20
0
Mar 2018
Apr 2018
May 2018
Jun 2018
Jul 2018
Aug 2018
Sep 2018
Oct 2018
Nov 2018
Dec 2018
Jan 2019
Feb 2019
86 % of running jobs are MCSimulation @ > 98 % CPU/Wall time efficiency
Max: 129, Min: 30.0, Average: 90.4, Current: 50.2
MCSimulation 80.9%
MCFastSimulation 5.5%
user 5.4%
MCReconstruction 2.8%
DataStripping 2.8%
DataReconstruction 1.7%
WGProduction 0.6%
Merge 0.2%
Turbo 0.1%
MCMerge 0.0%
Hospital 0.0%
HistoMerge 0.0%
test 0.0%
unknown 0.0%
Generated on 2019-02-24 19:37:51 UTC


Job CPU efficiency by JobType
52 Weeks from Week 07 of 2018 to Week 07 of 2019
user
WGProduction
Turbo
Total
Merge
MCSimulation
MCReconstruction
MCMerge
MCFastSimulation
DataStripping
Mar 2018
Apr 2018
Jun 2018
Jul 2018
Aug 2018
Sep 2018
Oct 2018
Nov 2018
Dec 2018

Running jobs by Site







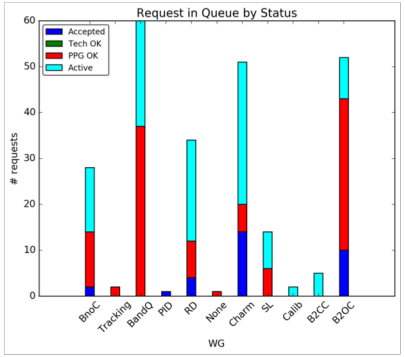
Distributed Computing Productions in 2019

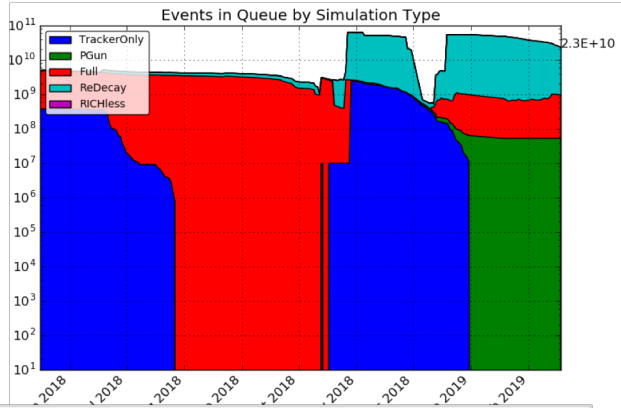
- Monte Carlo productions
  - 2017 MC simulation in full steam, starting with 2018 MC simulation
  - Estimate to complete current “queue” of requested events account for several months
  - New Sim10 simulation framework under validation

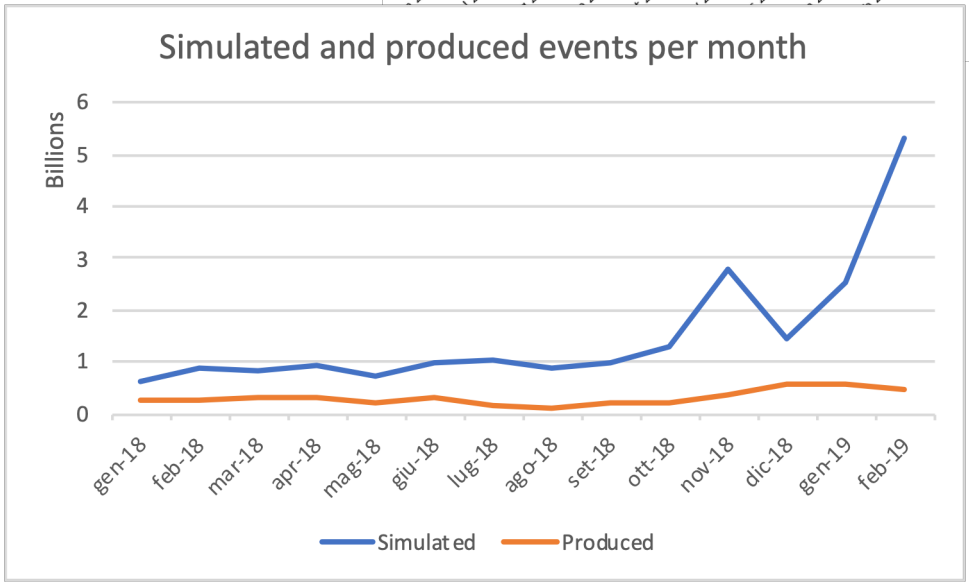
- Incremental stripping campaigns for all Run1 and Run2 data under way
  - Plan to finish processing of data by fall 2019
    - 6 - 8 weeks for each Run 2 data taking year and Run1 combined
  - Tight schedule, mostly limited by tape staging throughput

Simulation Queue

- Submitted requests are dominated by 3 working groups
- Fast simulation for majority of requested events
- Number of simulated events per month increases due to the usage of fast simulation
- Most productions are filtered before being written to disk

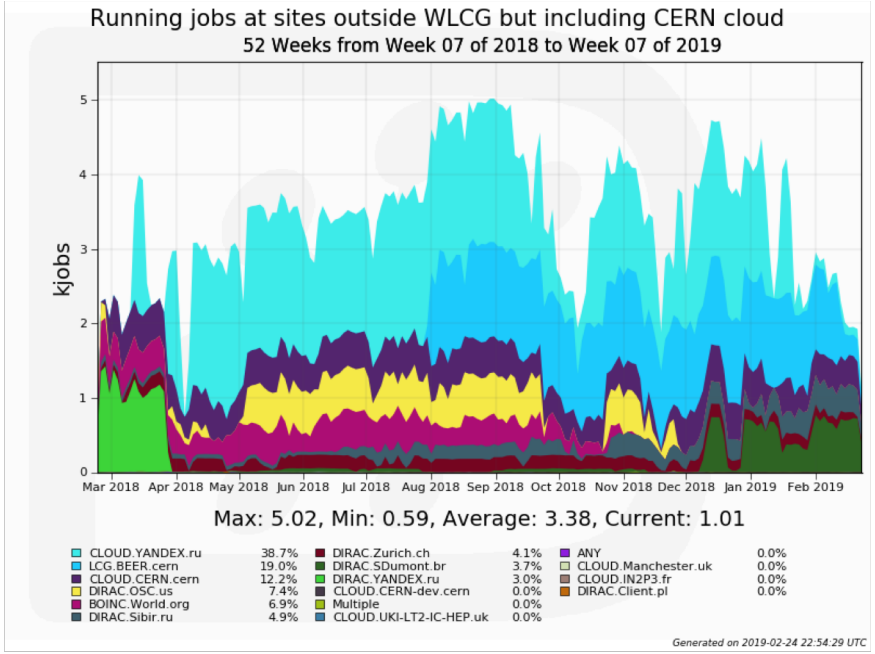






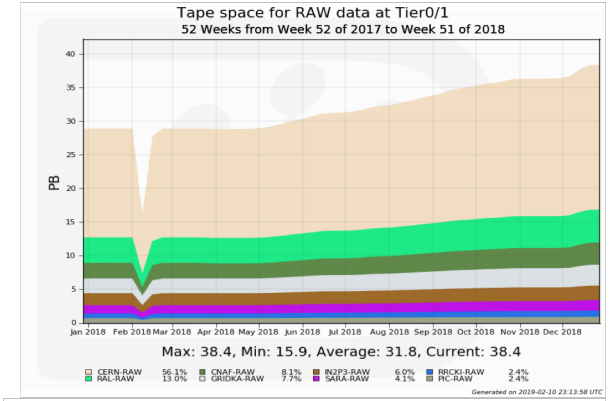
Opportunistic Computing

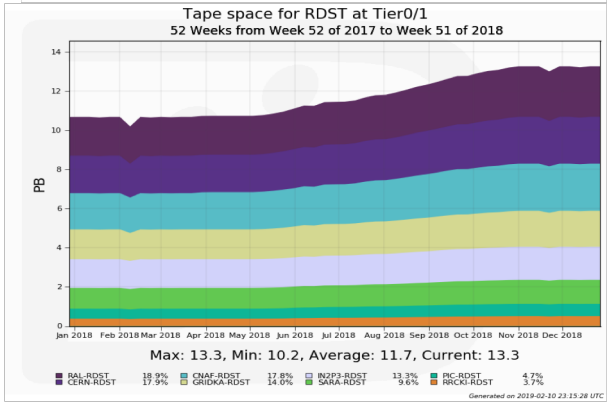
- Opportunistic resources continue to be successfully used
- Significant usage of BEER at CERN and CERN cloud

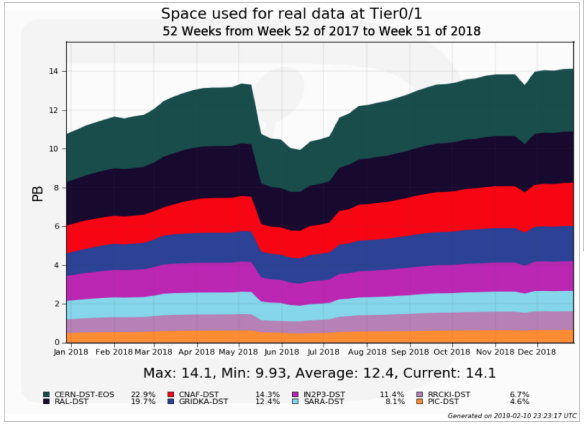


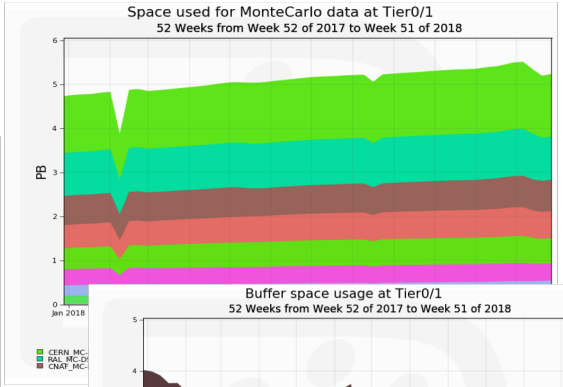
Storage usage

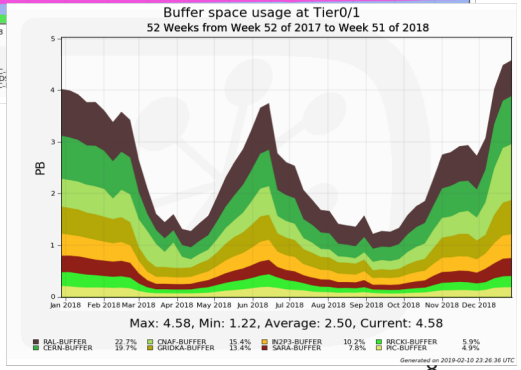
LHCb		2018						
		C-RSG	Pledged	Pledged / CRSG	Used	Used / CRSG	Deployed capacity	Deployed cap./CRSG
Disk	Tier-0	11.4	11.4	100%	7.72	68%	9.8	86%
	Tier-1	24.5	26.2	107%	22.13	90%	25.0	102%
	Tier-2	5.7	3.7	65%	3.1	54%	5.3	93%
	Total	**41.6**	**41.3**	**99%**	**32.95**	**79%**	**40.1**	**96%**
Tape	Tier-0	33.6	33.6	100%	29.2	87%		
	Tier-1	45.6	56.9	125%	39.7	87%		
	Total	**79.2**	**90.5**	**114%**	**68.9**	**87%**		
		disk usage Feb 12 2019 tape usage Dec 31 2018						











Resources in 2019-2021

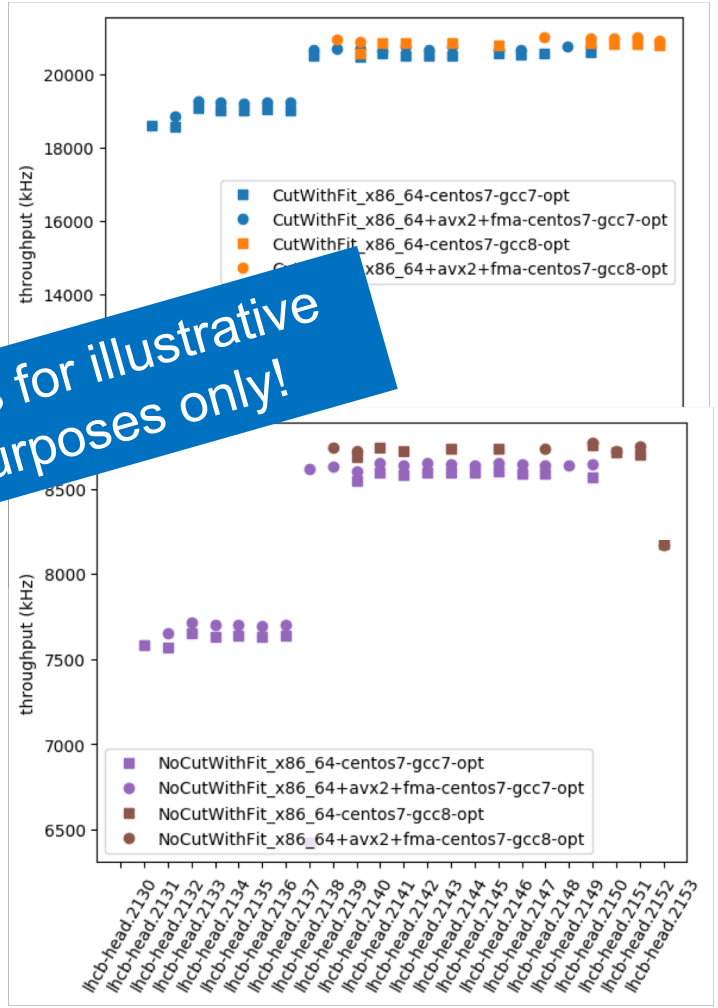
- Increase in 2020 due to generation of large simulated samples to tune the reconstruction and selection algorithms to be run in the HLT
- Resources in 2021 and beyond: see below

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%**

Run 3 Upgrade

Upgrade Software Status

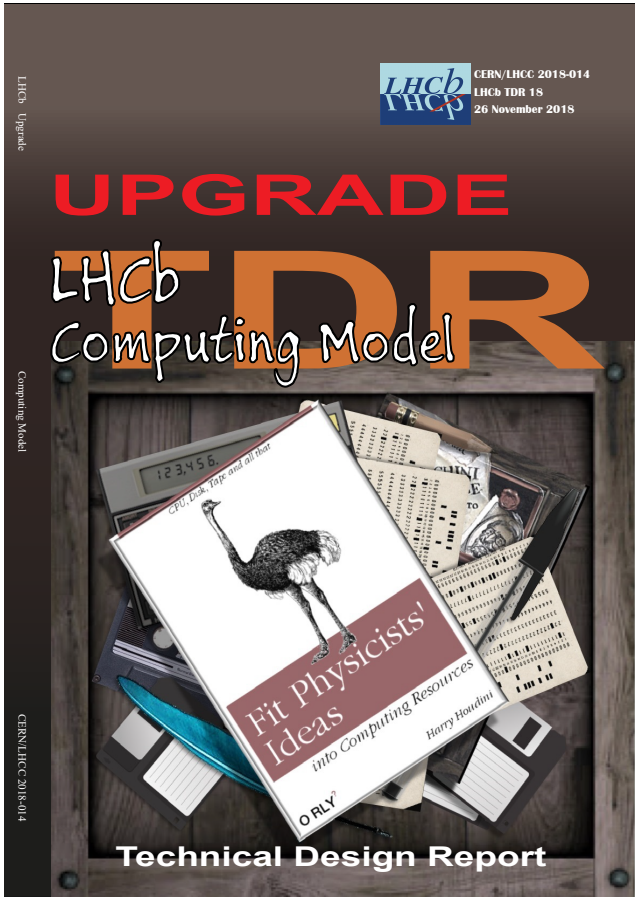
- New real-time analysis (RTA) project in charge of the HLT and reconstruction/selection
- Strict collaboration with computing project, which takes care of software engineering aspects
  - New scheduler, framework improvements, better memory management, new interfaces to conditions data and detector description (DD4HEP)
- Steady improvements in the past months for HLT1
- More improvements in the pipeline
  - Refactoring of algorithms
  - Working on data structures to enable vectorization
- Moving focus on HLT2 application



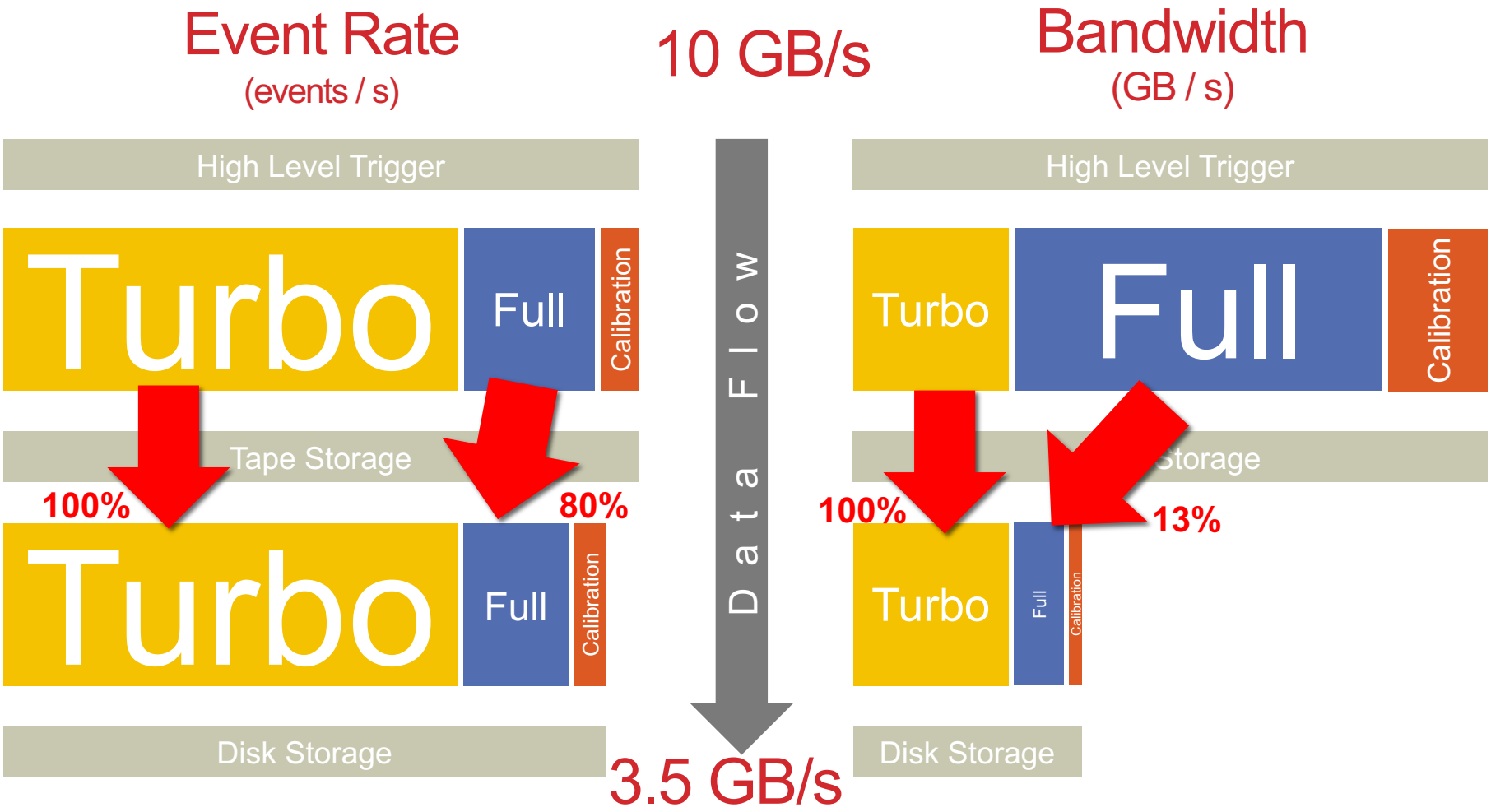
Run 3 Computing Model

- Concepts developed and implemented during Run 2 to become predominant
  - Split HLT → real-time alignment and calibration
  - TURBO stream for majority of physics program → RAW events discarded
  - FULL and CALIBRATION streams to insure flexibility → skim&slim offline
- Offline CPU computing needs dominated by simulation
  - Number of events to be simulated scales with luminosity
  - Simulation time per event scales with pileup

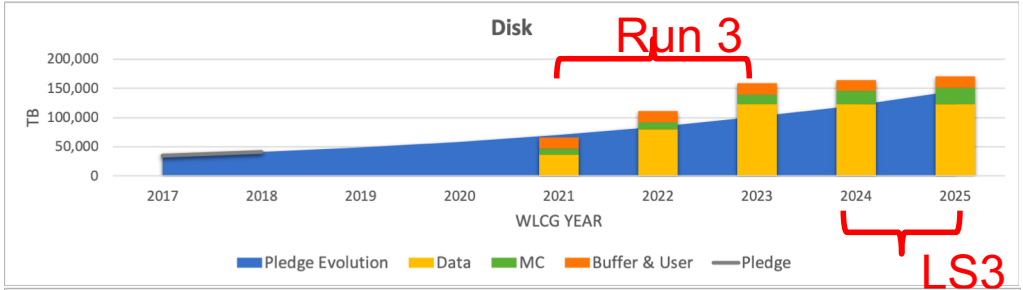
  →CPU simulation explodes → need for faster simulations
- Offline storage driven by trigger output bandwidth
  - MC saved in μDST, so little impact on storage

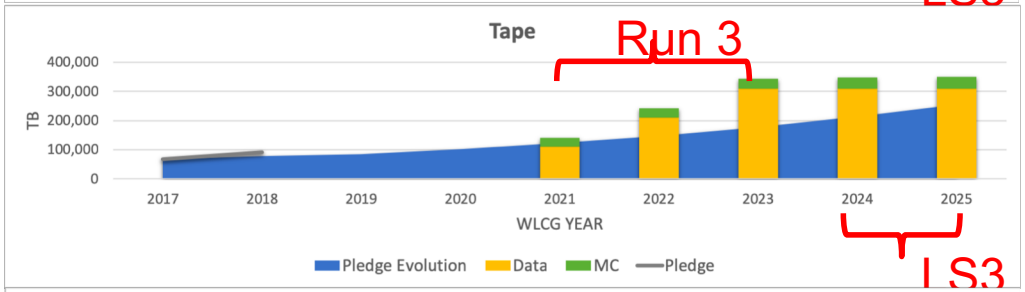


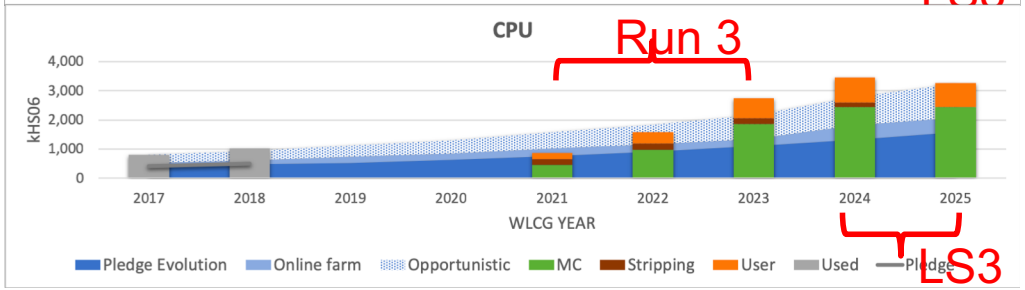
https://cds.cern.ch/record/2319756


Event Rate
(events / s)
10 GB/s
Bandwidth
(GB / s)
High Level Trigger
Turbo
Full
Calibration
Tape Storage
100%
80%
Turbo
Full
Calibration
Disk Storage
Data Flow
3.5 GB/s
High Level Trigger
Turbo
Full
Calibration
Storage
100%
13%
Turbo
Full
Calibration
Disk Storage

Resource needs – 2021 to 2025







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

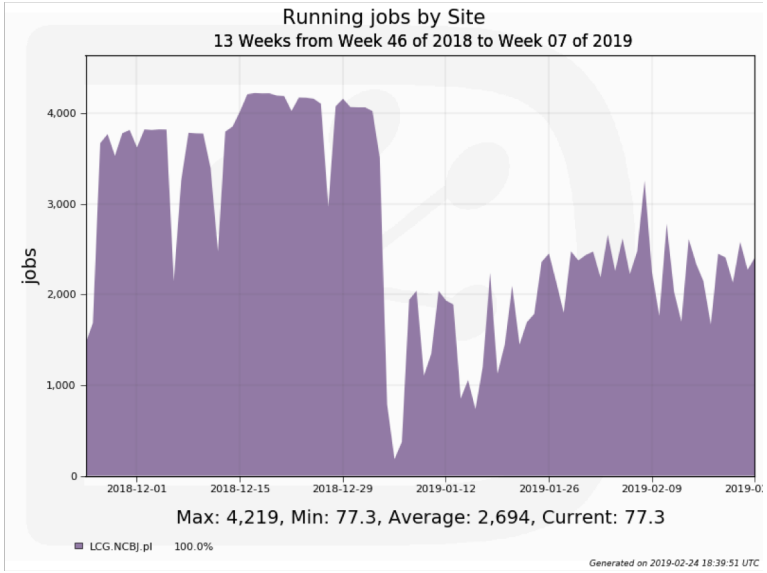
Support for Run 3 Resource Needs

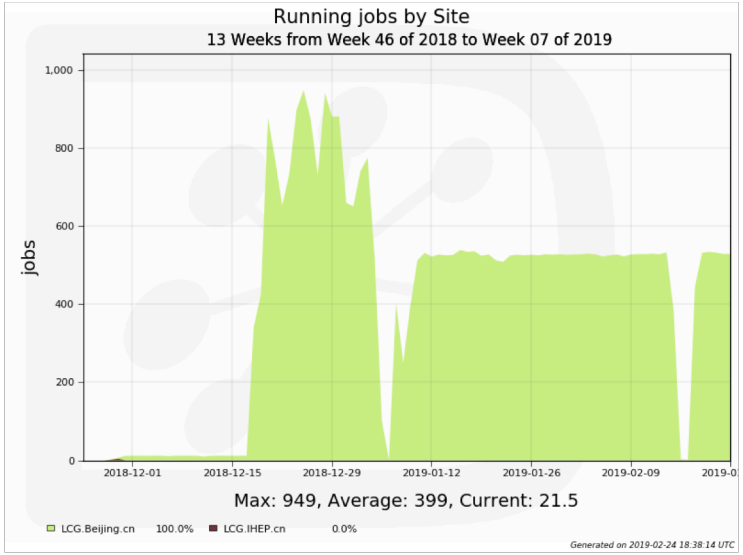
- Feedback received so far
  - Funding agencies understand and support our request for increasing resources beyond “constant budget”
  - They express their willingness to help but no concrete statements were made
  - They wish to better understand the physics justifications
  - And to discuss at the C-RRB

- LHCb is attempting to add more CPU and disk resources in the next years especially in China and US
  - Very good feedback and positive outlook!
- Poland planning to upgrade NCBJ to a Tier 1 site, including tape
  - Note that this site is currently not pledging to WLCG

Tier2 sites

- NCBJ reliably providing CPU comparable to a Tier1
- IHEP Beijing up and running
- MIT (US) being commissioned





Summary

- Computing operations continue at high usage and efficiency
  - This year usage on average 2x above WLCG pledged resources
  - Keeping up with MC production requests
  - Additional efforts to secure and utilize more opportunistic and unpledged resources, as well as more “structured” ones

- Run 3 upgrade work progressing well
  - Continued improvements for HLT1 CPU-based trigger
  - Computing Model TDR shows that the foreseen 30x larger data sample from the detector can be processed by a moderate resource increase.