LHCb Computing Report

Stefan Roiser

LHCC WLCG Referees Meeting
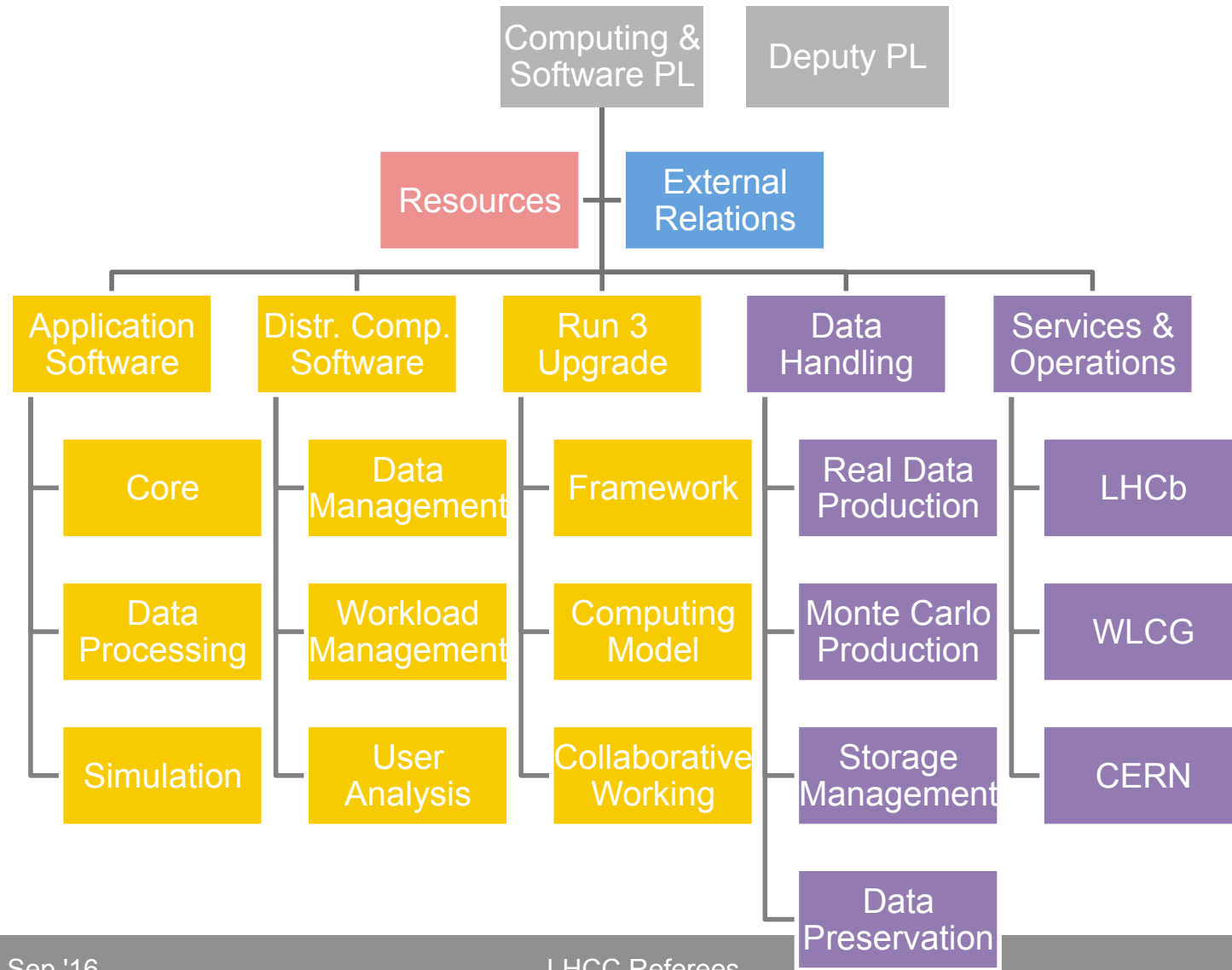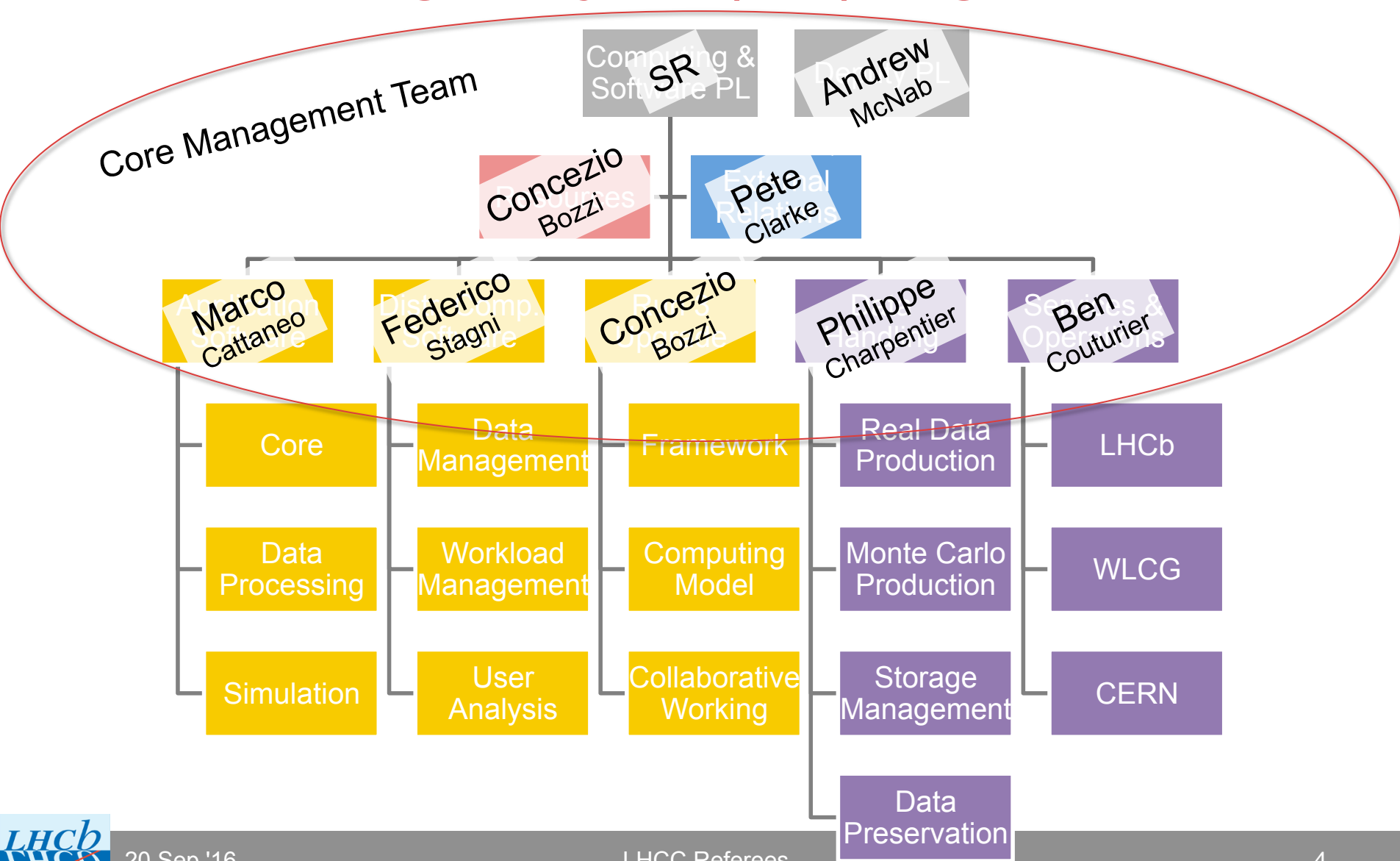
20 September 2016

Content

- Computing Project Organisation
- Situation with 2016 Data Processing & Simulation
- Outlook to 2017, 2018 and 2019 resource requests
- Run 3 Upgrade
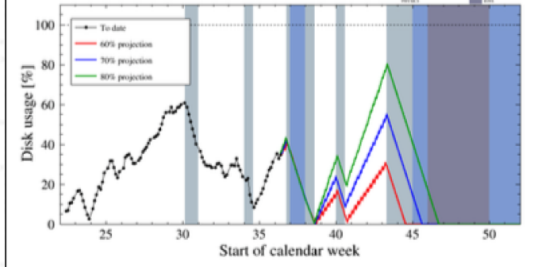
Computing Project (Re)Organisation

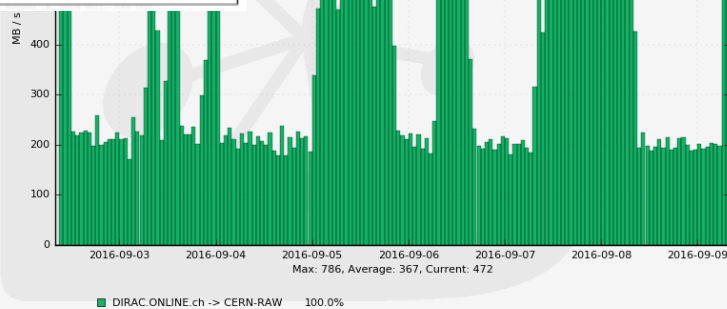# Computing Project (Re)Organisation

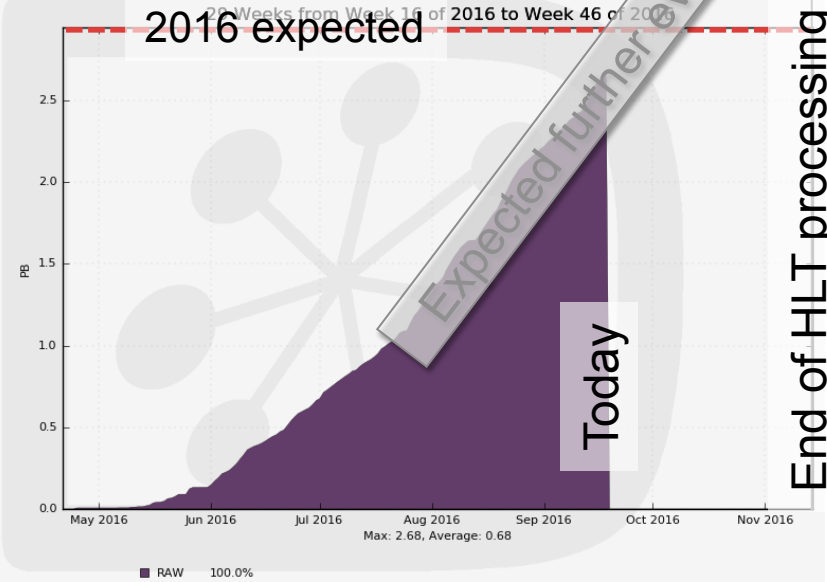# Data Export from LHCb Pit

HLT Disk Buffer Occupancy

RAW data volume

2016 expected

Throughput by Channel
Days from 2016-09-02 to 2016-09-09

Pit Export Throughput

Expected further evolution

End of HLT processing

Today

Max: 786, Average: 367, Current: 472

DIRAC.ONLINE.ch -> CERN-RAW   100.0%

Generated on 2016-09-09 08:55:43 UTC

Max: 2.68, Average: 0.68

RAW   100.0%

Generated on 2016-09-19 13:06:55 UTC

- **Continuous data export** from the pit to CERN/Castor
  - Increased LHC efficiency with repercussions on data export volume
    - Currently reaching exported RAW volume that was expected for all of 2016
    - Mitigation via reducing the size of the reconstruction output format
  - Output rate In-fill ~ 200 MB/s, Inter-fill ~ 600 MB/s with split HLT 1 & 2
  - Will continue until mid November until HLT pit disk buffer is emptied

Full / Turbo Stream Partitioning

Full Stream vs Turbo RAW Recorded Data
20 Weeks from Week 17 of 2016 to Week 37 of 2016
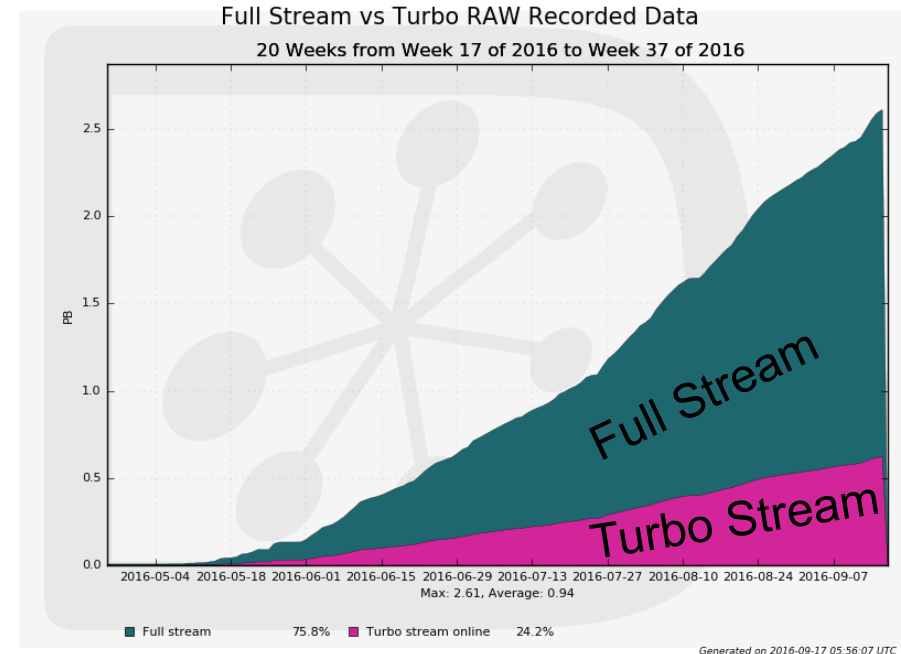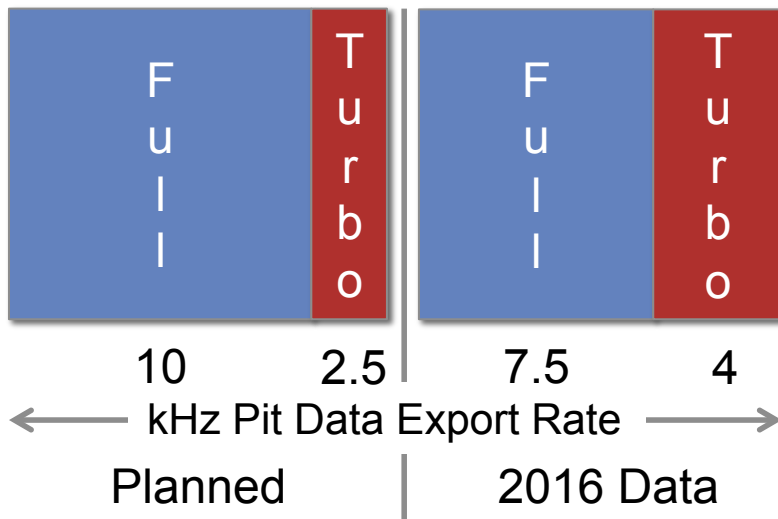
- Decided beginning of the year to also allow reconstruction information in Turbo stream data in addition to signal info (Turbo++)

  - Moved data from Full to Turbo/Turbo++ → reduce offline CPU needs
  - Event size increased from initially 10 kB to 50 kB, needed tape resources ~ the same but → strain on disk

TURBORAW Stream

- Split HLT Turbo RAW data exported to CERN/CASTOR

  - Safety net for eventual need of re-running the trigger on Turbo data

  - No additional strain on final tape resources, planned to be removed by the end of the year

PFN space usage by EventType

20 Weeks from Week 17 of 2016 to Week 37 of 2016

Max: 3.99, Average: 1.46

Full stream	48.7%	TURCAL	1.9%	stream?	0.0%
TURBORAW	27.1%	Beam Gas	1.6%	Nobias stream	0.0%
Turbo stream online	15.5%	Calib stream	1.5%	Luminosity stream online	0.0%
Express stream	3.5%	na	0.0%		
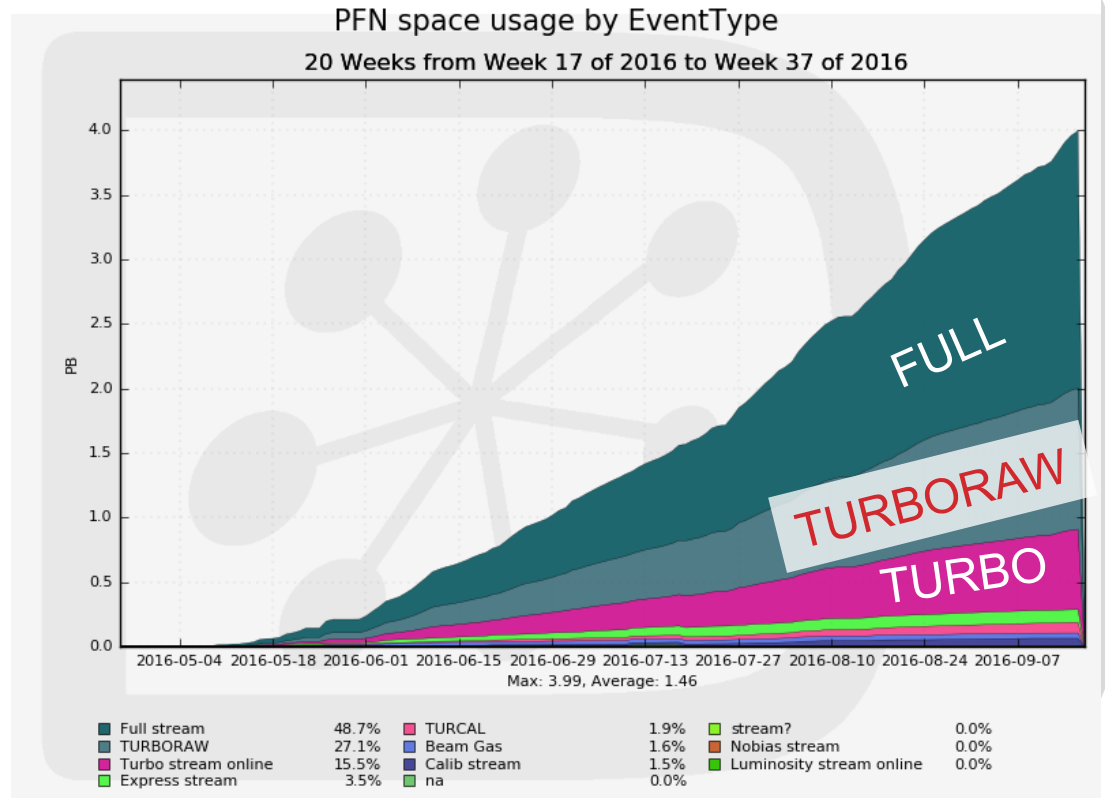
Generated on 2016-09-17 06:45:46 UTC

Offline Data Processing

- Overall usage well above pledged resources

- Usage dominated by Simulation Workflows (~ 70 %)

- All Data Processing Workflows setup and running, i.e.

  - Full Stream Reconstruction & Stripping

  - Turbo Stream Processing

Running jobs by JobType
15 Weeks from Week 22 of 2016 to Week 38 of 2016

Pledge

Max: 78.4, Min: 6.34, Average: 42.9, Current: 47.7

MCSimulation	72.5%	DataStripping	5.4%	Turbo	0.4%
user	10.4%	MCReconstruction	0.9%	WGProduction	0.0%
DataReconstruction	9.4%	Merge	0.8%	MCMerge	0.0%

MCStripping	0.0%
unknown	0.0%
test	0.0%

Generated on 2016-09-19 14:32:56 UTC
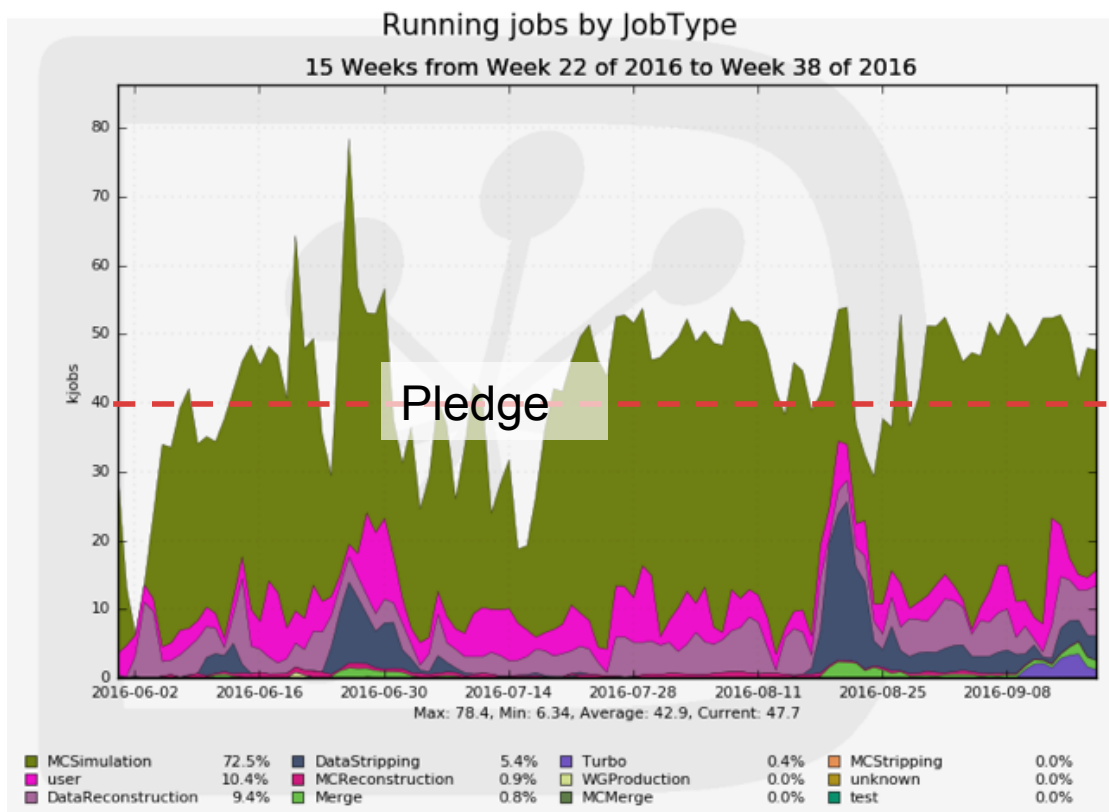
Data Processing Workflows

- DataReconstruction running continuously since beginning of data taking

- DataStripping re-started mid-August with lower bandwidth to cope with disk resources

Running Jobs by Type
15 Weeks from Week 22 of 2016 to Week 37 of 2016

Max: 28,473, Min: 67.4, Average: 6,718, Current: 404

		Merge	5.0%
DataReconstruction	58.9%	Turbo	2.2%
DataStripping	34.0%		

Generated on 2016-09-17 06:06:36 UTC

  - Processing in June/July is part of the same data sample

- Development of streaming of Turbo data needed b/c of increased LHC efficiency & data format

  - New development allows to park 1/3 of data on tape → reduce disk strain
  - Development finished, production started 9 Sep and already caught up
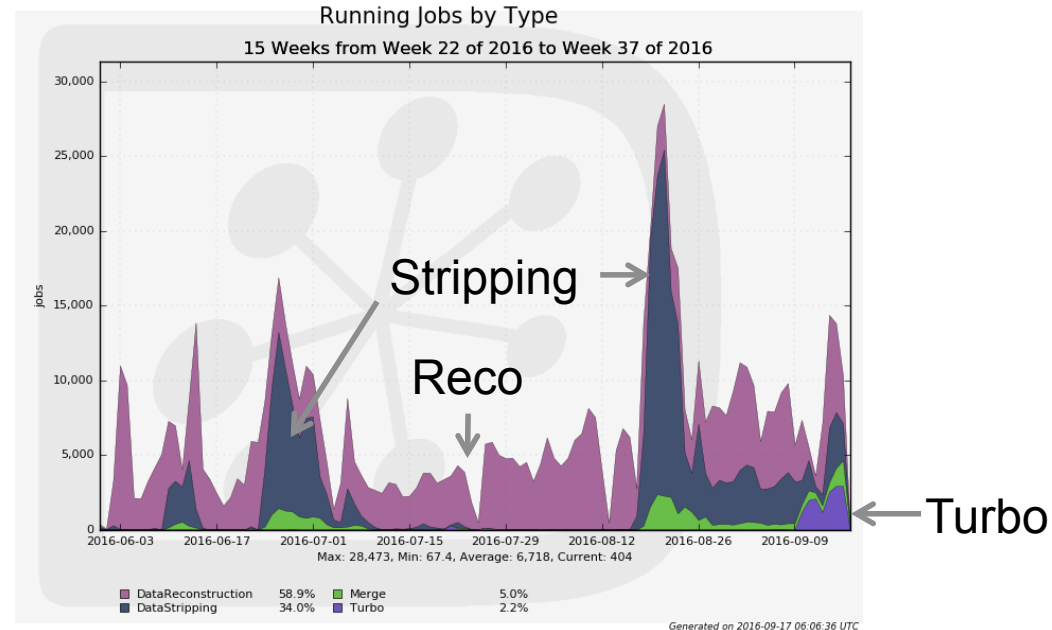
Disk Resident Data

- 2016 Mitigation Policies
  - Reduce number of disk replicas
  - Turbo streaming and parking 1/3 of data volume on tape
  - Currently ongoing discussion about physics need for full end of year restripping of Run2 data
    - NB: Trying to avoid any possible hit on the experiments physics performance and output
  - Use data popularity to remove ~ 1 PB of currently not needed Simulation datasets

Aug '16 CRSG Resource Usage Data Popularity Plot

10113.4

4128.8 ← was 5183 in Jan '16

189.9

Amount of physical data (TB)

Number of Disk dataset usages (last 26 weeks)

Unused older | Unused from period | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 | 10-11 | 11-12 | 12-13 | 13-14 | >14 | Used
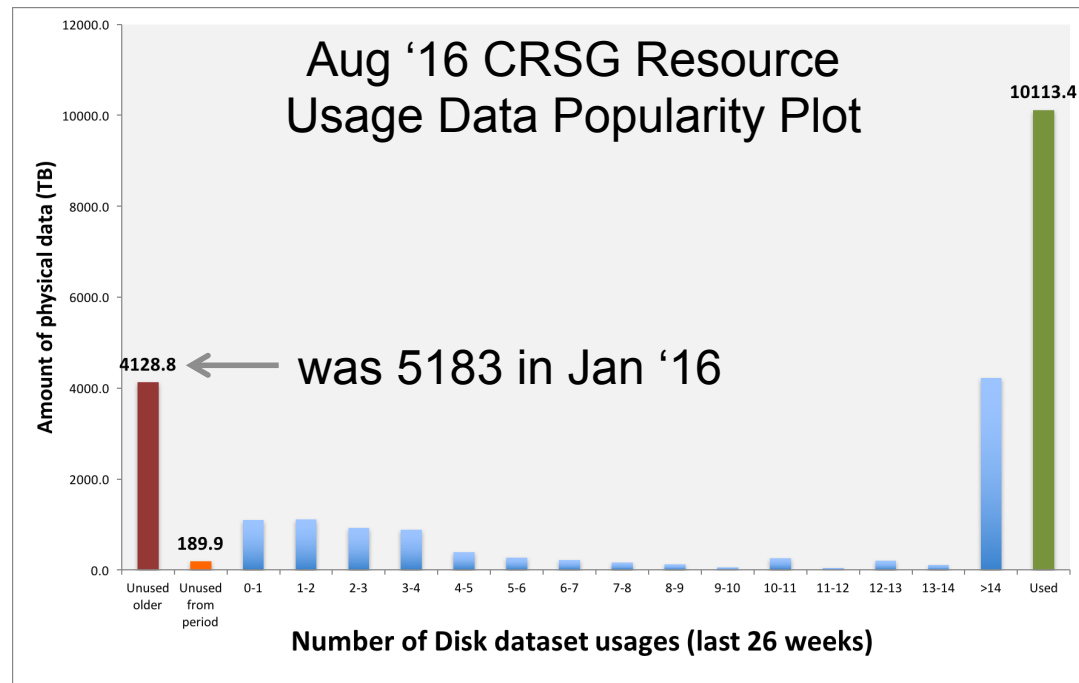
Simulation

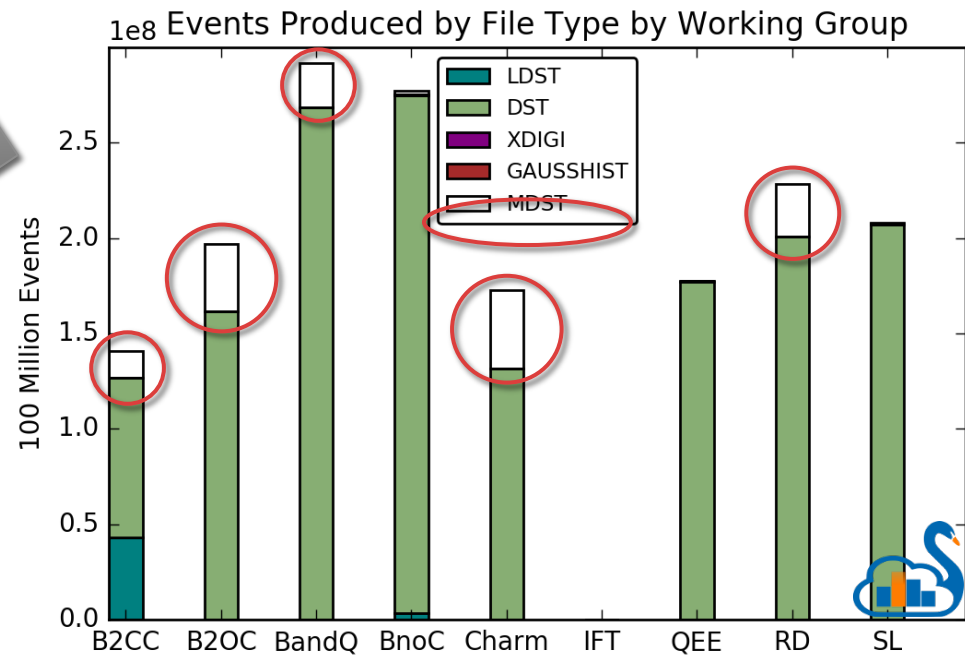- Move from "Sim08" to new simulation framework "Sim09" clearly visible
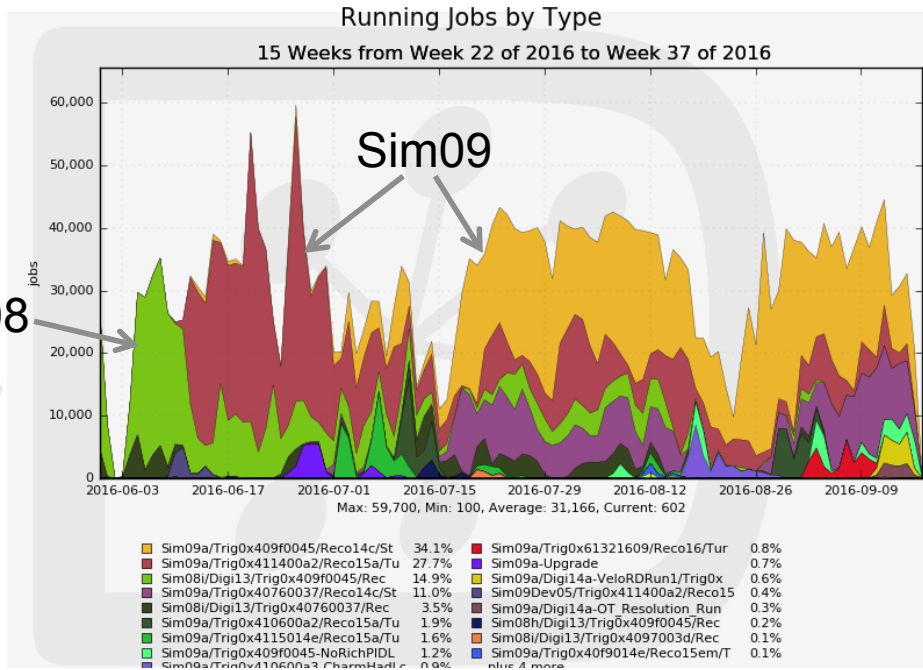
  - Speed improvements in RICH simulation

  - New microDST format starts being adopted by physics working groups

- "Sim10" developments targeting end of 2016:

  - Move to Geant4.10

  - Introduction of more generators (mostly for Pb)

  - Re-use of underlying event

Running Jobs by Type
15 Weeks from Week 22 of 2016 to Week 37 of 2016

Sim09

Sim08

Max: 59,700, Min: 100, Average: 31,166, Current: 602

Sim09a/Trig0x409f0045/Reco14c/St	34.1%	Sim09a/Trig0x61321609/Reco16/Tur	0.8%
Sim09a/Trig0x411400a2/Reco15a/Tu	27.7%	Sim09a-Upgrade	0.7%
Sim08i/Digi13/Trig0x409f0045/Rec	14.9%	Sim09a/Digi14a-VeloRDRun1/Trig0x	0.6%
Sim08i/Digi13/Trig0x40760037/Rec	11.0%	Sim09Dev05/Trig0x411400a2/Reco15	0.4%
Sim08i/Digi13/Trig0x40760037/Rec	3.5%	Sim09a/Digi14a-OT_Resolution_Run	0.3%
Sim09a/Trig0x410600a2/Reco15a/Tu	1.9%	Sim08i/Digi13/Trig0x409f0045/Rec	0.2%
Sim09a/Trig0x4115014e/Reco15a/Tu	1.6%	Sim08i/Digi13/Trig0x4097003d/Rec	0.1%
Sim09a/Trig0x409f0045-NoRichPIDL	1.2%	Sim09a/Trig0x40f9014e/Reco15em/T	0.1%
Sim09a/Trig0x410600a3_CharmHadLc	0.9%	plus 4 more	

Events Produced by File Type by Working Group

2016 Resource Usage Recap

- Increased LHC efficiency and change in Turbo data processing with several repercussions on resources
  - Increased usage of tape for RAW data, mitigated by change in the reconstruction output format
  - Increase of offline CPU needs for data processing mitigated by introducing moving more data from Full stream (offline CPU intensive) to Turbo stream (little offline CPU needs)
  - Increase of disk needs because of new Turbo++ format mitigated by usage of data popularity, reduction of replicas and parking of data
    - NB: From 2017 onwards also planned to reduce the Turbo++ format by further slimming of reconstruction information

Resources '16-'19

- Most strain on disk resources despite mitigation actions, ie.
  - replicas, parking, data popularity
- 2016 resources under control applying actions above
- 2017 resources re-submitted to CRSG with 20 % increase mainly for disk
- 2018/19 resources inline with constant budget increase of resource requests for cpu/ disk/tape on top of 2017
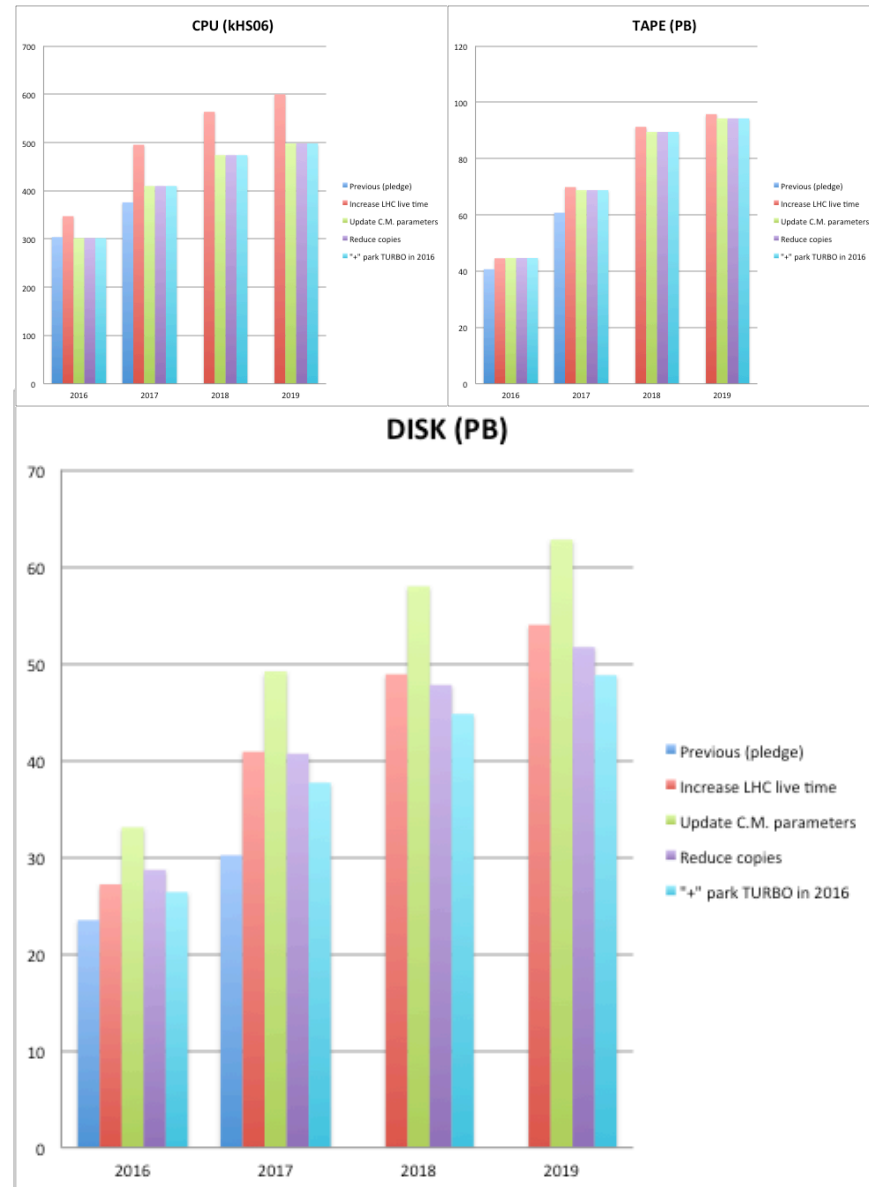
Computing Upgrade: Work Areas

- Revolution → task force established
  - Move Gaudi from single-threaded, sequential processing to task-concurrent framework by using Gaudi-Hive ideas
    - algorithms as re-entrant entities to be executed in parallel
    - Input/output data to become immutable
  - Re-develop Event Model to exploit SIMD / vectorization
    - Tightly coupled on usage of math libraries
  - Re-implement conditions database and detector description
  - Exploit alternative architectures (for trigger)
- Evolution → use Run2 as testbed
  - Simulation: fast / hybrid / parameterized / concurrent
  - Distributed computing
    - TURBO stream as default in Run3, centralized ntuple production, event indexing, incremental changes in DIRAC
  - Collaborative tools and analysis preservation

Organization of task force

- "Revolutionary" program attacked with task-force approach
- 1st Taskforce hackathon – 26-27 May @ CERN
  - Infrastructure ready, start of real work
- 2nd hackathon – 7-8 July @ CERN.
  - Implementing further changes, start involving subsystems
- 3rd hackathon: September 19-20 @ CERN
  - https://indico.cern.ch/event/562730/
  - Progress towards implementing APIs
  - GAUDI workshop September 21-23
- 4th hackathon: Nov 14-18 during 8th Computing Workshop
  - APIs ready, subsystems more involved
  - Currently planning a two-day hackathon in parallel with training sessions to enable newcomers to quickly contribute
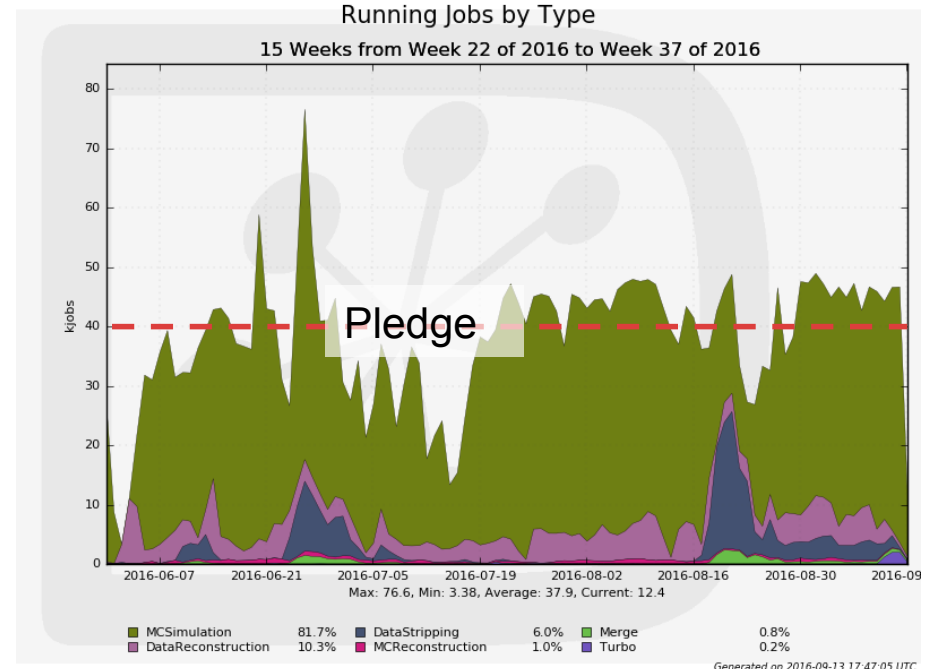

Currently Ongoing

Summary

- Overall high to very high usage of resources
  - New Simulation framework in full swing with 80 % of CPU usage
  - Increased LHC efficiency & new Turbo processing with strain on disk
- 2016 data processing situation
  - Repercussions of increased LHC efficiency but situation under control with several mitigation policies applied
- 2017 resource needs increased and resubmitted to CRSG
  - Together with evolution of resource requests for '18 + '19 assuming constant budget increase
- Computing Upgrade ongoing
  - Improved situation for the upgrade of Gaudi framework but stays on critical path
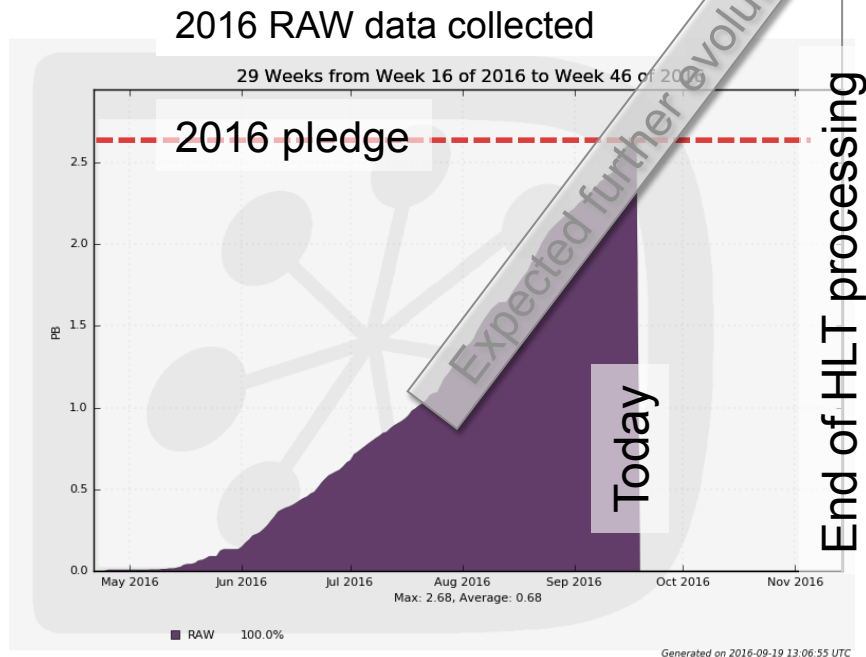
Executive Summary Slides

LHCb 2016 Distributed Data Processing

- Continuous pp RAW data export from pit expected until mid November
  - Result of HLT buffered data and its deferred processing
- Overall very high usage of distributed computing CPU resources
  - Adaptation of data processing workflows needed due to high LHC efficiency
  - All workflows now running in production
  - Release of new Simulation framework immediately adopted
    - 80 % of CPU on distributed compute resources needed for MC

Running Jobs by Type
15 Weeks from Week 22 of 2016 to Week 37 of 2016

Pledge

Max: 76.6, Min: 3.38, Average: 37.9, Current: 12.4

MCSimulation	81.7%	DataStripping	6.0%	Merge	0.8%
DataReconstruction	10.3%	MCReconstruction	1.0%	Turbo	0.2%

Generated on 2016-09-13 17:47:05 UTC

Data Storage

2016 RAW data collected

29 Weeks from Week 16 of 2016 to Week 46 of 2016

2016 pledge

Expected further evolution

Today

End of HLT processing

PB

2.5
2.0
1.5
1.0
0.5
0.0

May 2016   Jun 2016   Jul 2016   Aug 2016   Sep 2016   Oct 2016   Nov 2016

Max: 2.68, Average: 0.68

■ RAW   100.0%

Generated on 2016-09-19 13:06:55 UTC

- Increased disk resource needs in 2016 b/c of higher LHC efficiency and change in Turbo
- Mitigation actions taken:
  - Reduction of disk replicas
  - Parking of 1/3 of Turbo data on tape
  - Use of data popularity to reduce not needed Simulation samples

- Decided to allow Turbo data format to contain reconstruction information in addition to signal
  - Moved from Full to Turbo stream
  - → less CPU offline resources
  - Turbo event size increased
  - → additional strain on disk

- 2017 resource requests re-evaluated and re-submitted to CRSG with 20 % increase mainly for disk in respect to initial request
  - 2018/19 requests follow evolution of a constant budget increase