

# LHCb Computing

## 2015 Q3 Report

---

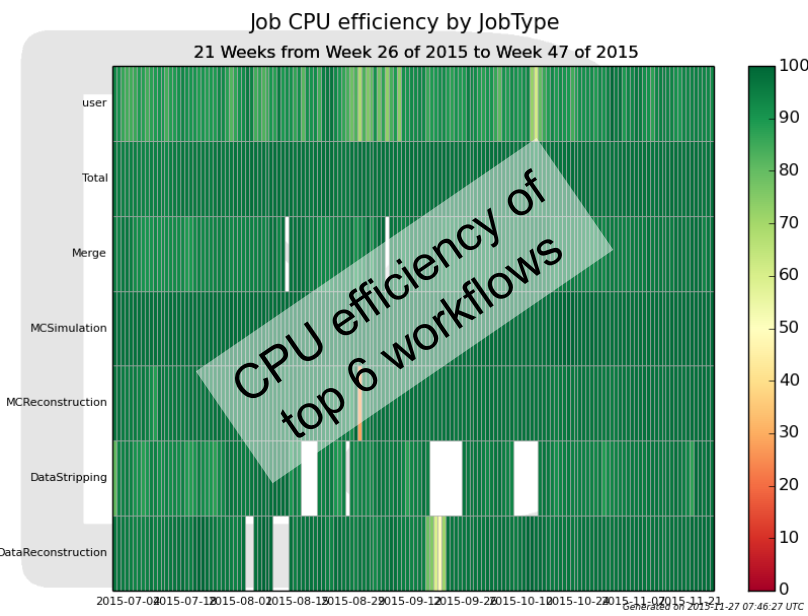
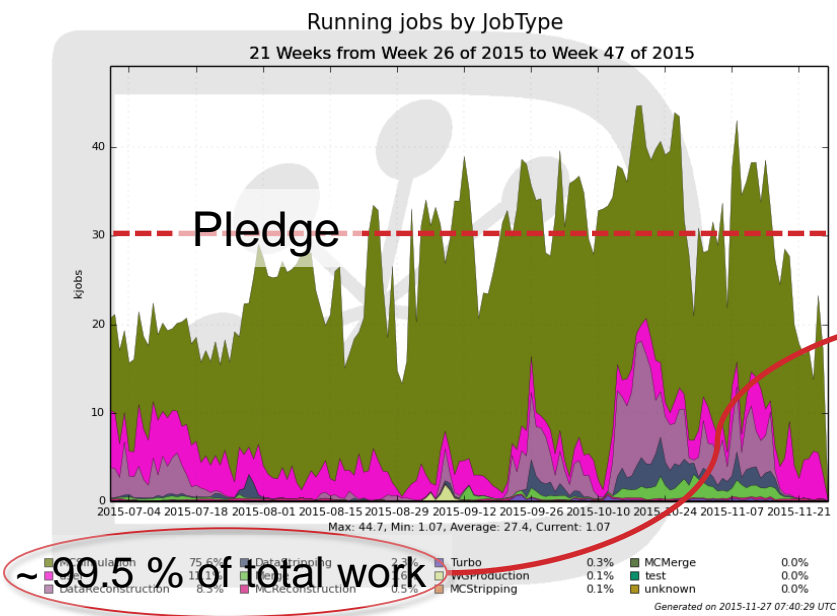
Stefan Roiser

LHCC Referees Meeting

1 December 2015

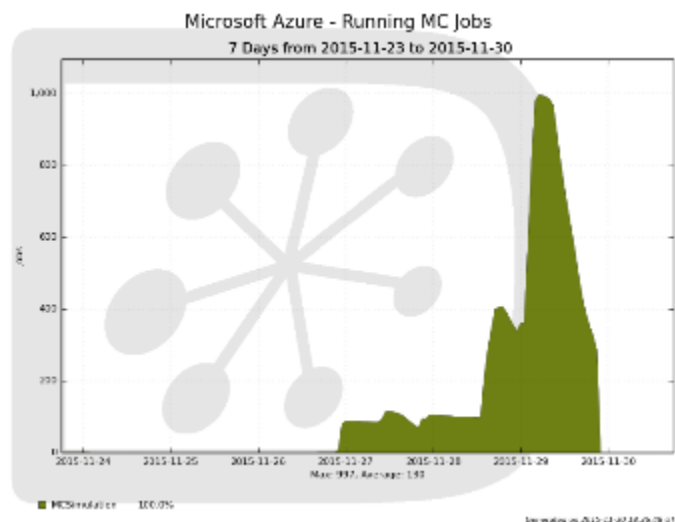
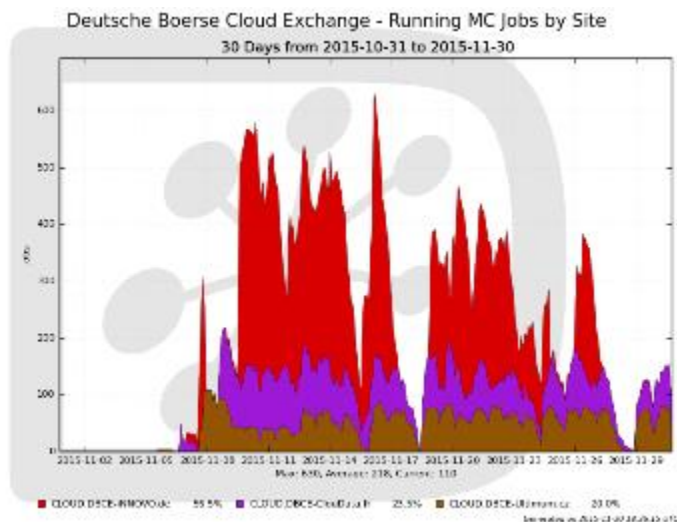


# CPU Resource Usage



- Usage consistent with pledges
  - Job execution continues at very high CPU efficiency
  - As of start of data taking no usage of the HLT farm
    - Deferred triggering occupies the resource for online processing
  - Continued usage of other opportunistic resources Yandex, Ohio, ...

# Cloud Resources



- DBCE: second cloud procurement from CERN IT
  - LHCb quota of ~1000 cores Nov to Jan, split across 3 to 5 providers
  - MC simulation: third procurement will target data processing too
  - Four-core VMs, one MC job (one Dirac pilot) per core
- In parallel, also tested some capacity available on Azure

# Disk Storage Occupancy

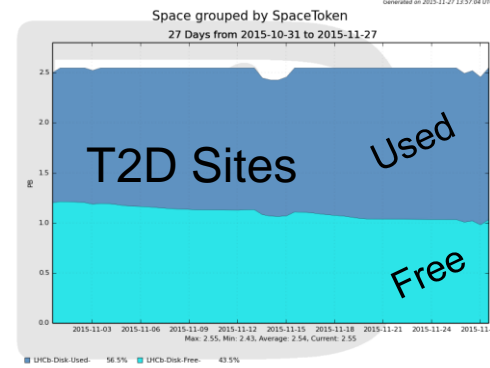
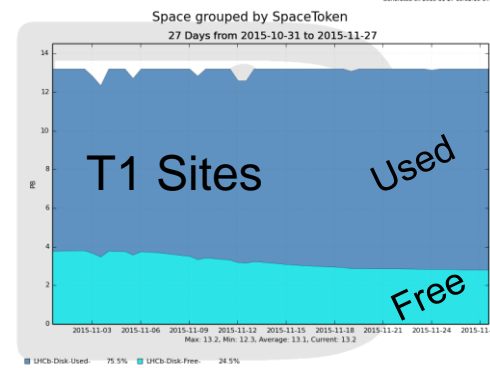
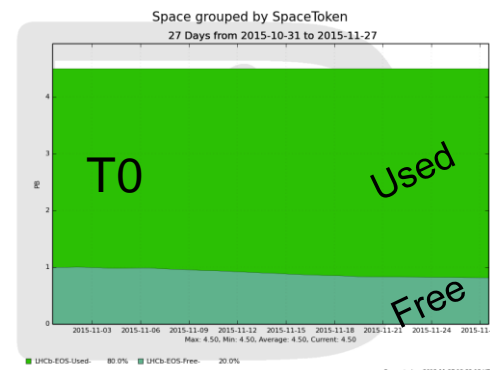


## Storage

- Currently using less disk than anticipated
  - LHC live time ~half of that initially foreseen
  - LHCb implemented changes in the computing model parameters (see later) which moderate the tape requirements
  - Also ~1.1PB disk space cleaned up following data popularity analysis
- Expect ~15% less disk usage and ~35% less tape usage with respect to pledge by the end of 2015 WLCG year

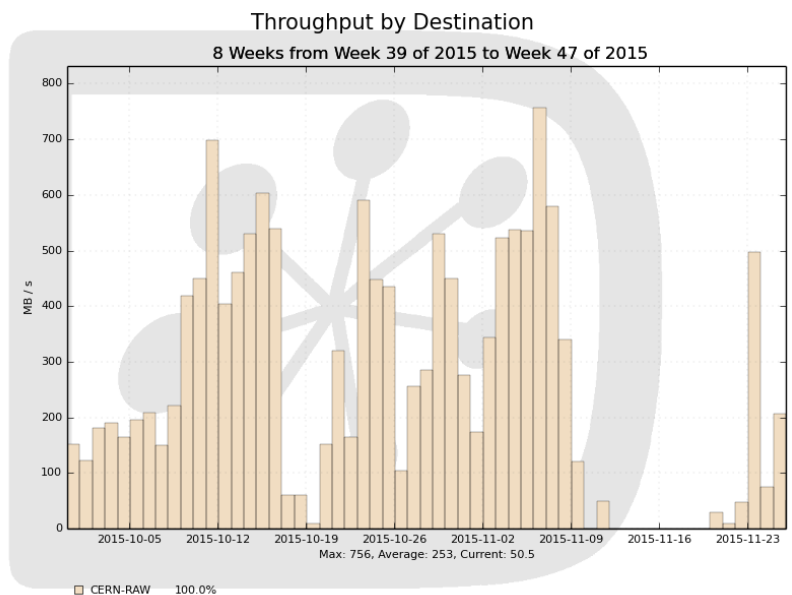
Disk (PB)	CERN	Tier1s	Tier2Ds
<b>LHCb accounting</b>	<b>3.32</b>	<b>8.36</b>	<b>1.20</b>
SLS T0D1 used	3.51	8.43	1.20
SLS T0D1 free	1.57	5.31	1.36
SLS T1D0 (used+free)	0.53	1.36	--
<b>SLS T0D1+T1D0 total1</b>	<b>5.61</b>	<b>15.10</b>	<b>2.56</b>
<i>Pledge '15</i>	<i>5.50</i>	<i>14.04</i>	<i>1.95</i>

Tape (PB)	CERN + Tier1s
RAW	6.5
FULL.DST	4.9
ARCHIVE	5.5
<b>TOTAL</b>	<b>16.9</b>
<i>Pledge '15</i>	<i>39.3</i>



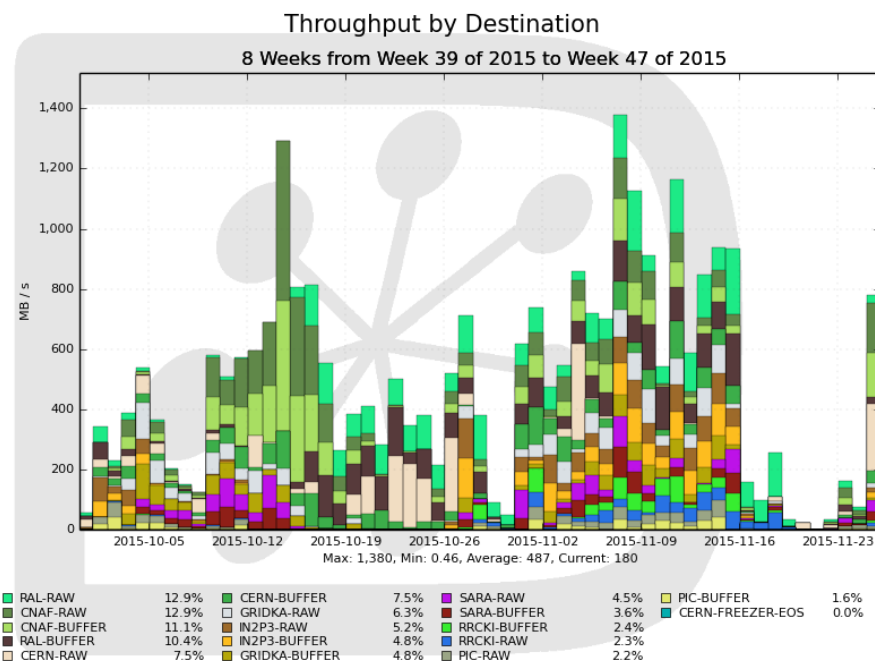
# Exercising the data flows

- Data transport at designed Run2 throughput rates
  - Further optimization of overall workflow by merging small RAW files in the pit -> less files -> higher throughput with FTS transfer service

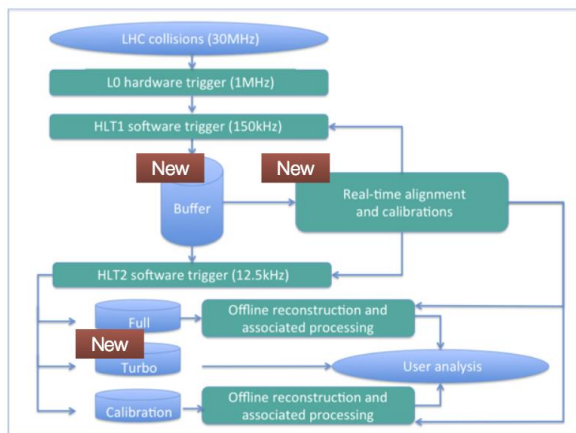


Pit export of RAW files at design throughput of up to 750 MB/s

## CERN RAW export to T1s at > 1.3 GB/s



# Exercising the workflows



- Turbo processing

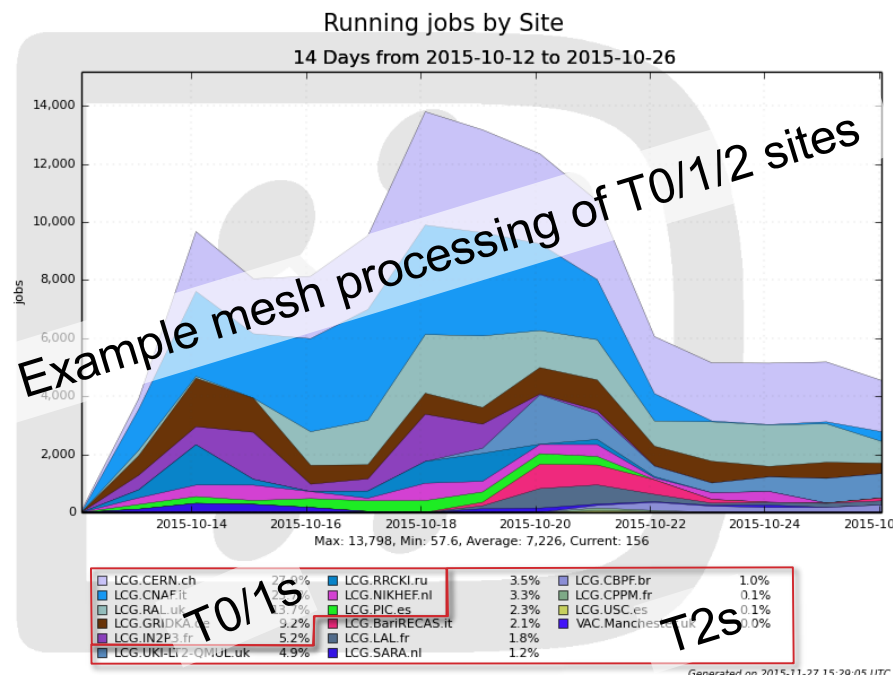
- Write out reconstructed physics objects directly from the pit after final alignment / calibration

- “Mesh processing”

- ie. allow T2 sites to dynamically contribute to workflows
- Very flexible way of distributing work across sites / tier levels

- + more than pp collisions

- SMOG – ie. p-He/Ne/Ar collisions
- PbPb program ongoing ...



# Next steps & winter shutdown activities

- Data processing
  - Several more planned processings upcoming
    - Run 2
      - Proton-Argon 5TeV & 13TeV, led-led first pass processing
      - Full pp data re-stripping
    - Run 1 incremental stripping
  - + proton-Neon, proton-Helium and Turbo reprocessing
    - Small datasets
  - No pp reprocessing as laid out in the planning
- Simulation
  - Finalize Sim09 validation
    - I.e. next major step in the simulation framework using new versions of Geant4, generators + tuning, Root 6, ...
      - Supports both Run 1 and Run 2 data

# 6<sup>th</sup> LHCb Computing Workshop

- This edition dedicated to “Run 3 brainstorming”
  - 16 – 20 Nov in LPNHE @Paris (first time outside CERN)
    - Plenary + 5 parallel tracks (event model, hardware, framework, data access, collab. tools)
  - ~ 70 participants
    - Core computing team, physicists, “externals” (CERN/SFT, ATLAS, CMS, FCC, Google, Intel, skykitlearn, HPX, ... )
- Shall lead to a “Run 3 roadmap”
  - Further evolve into the Computing TDR due by Q4 2017

## In conclusion

- A very productive week
- Lots of enthusiasm
- Critical assessment of current model, weak points, scalability issues
- Several ideas on how to improve / redesign
- Beware of timescales / effort required
- Keep up momentum!

C Bozzi, Workshop Summary

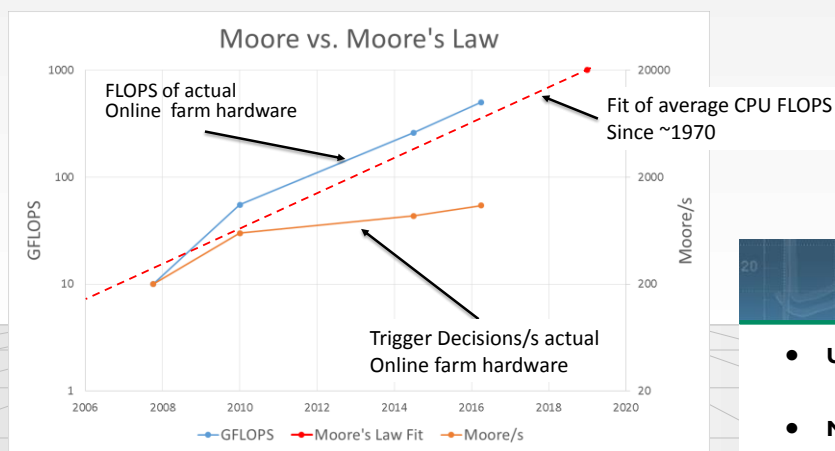


# Some impressions from the workshop ...

# Vectorization

## Moore vs. Moore's Law

- Decreasing number of trigger decisions / FLOP
  - Need better usage of CPU vectorization capabilities



- Tests done with 2014 Farm tender benchmark
- Dates are release dates of CPU

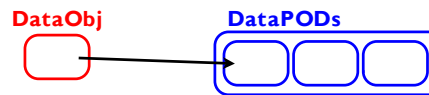
R Schwemmer, Online

## Separation of Concerns

- **Using PODs is a good idea, but...**  
... they are a little bit too dump to support all what is needed.
- **Need smart layers on top of the PODs**
  - Dealing with ownership
  - Allow referencing between objects
  - Deal with non-trivial I/O operations

Object Land

POD Land

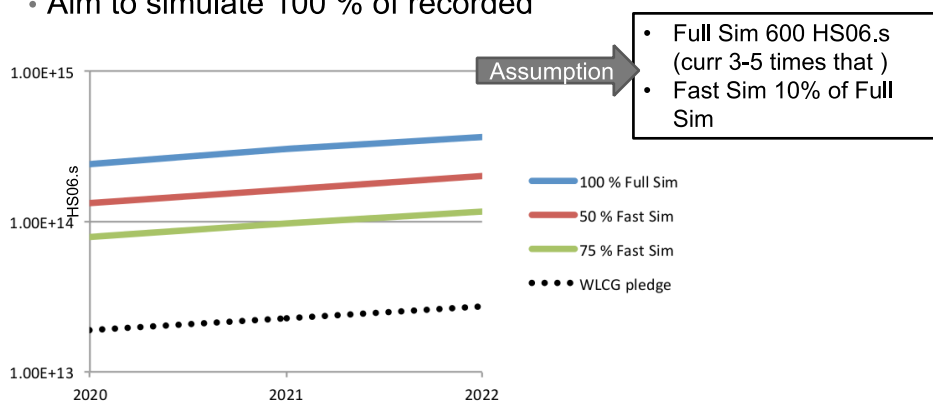


- **Whenever really performance critical - leave possibility of access to bare PODs**

# Simulation needs

## Monte Carlo Simulation

- Run 1 simulated events  $\sim 4.5 \cdot 10^9$  (spring '15)
  - $\sim 12\%$  of recorded
- Aim to simulate 100 % of recorded



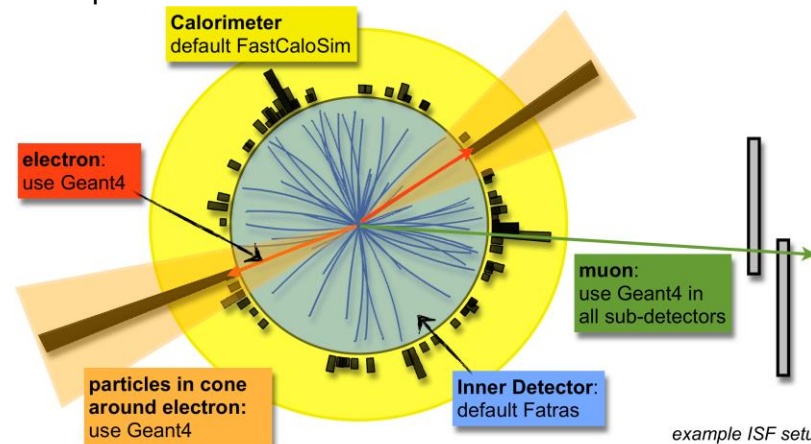
MC won't fit (by far) into the pledged resources :- ( **KO** )

19 Nov '15

F Stagni, Computing in 2020

## The Integrated Simulation Framework

- More recently we introduced the concept of the ISF in ATLAS
- Controls the stack; allows mixed simulation flavors in one event
- Hard questions about calibration and scale factors to be addressed



19/Nov/2015

Z Marshall Future ATLAS Fast Simu

Z Marshall, ATLAS Fast MC

- CPU work on distributed resources will be dominated by Simulation !!

# Detector Description & Conditions

Why did we choose XML at the time???

- Instead of inventing our own format use a standard one

a) Easy to read and to parse  
 b) Extensible  
 c) Easy to convert  
 d) Many tools  
 e) Extended using references

Consider

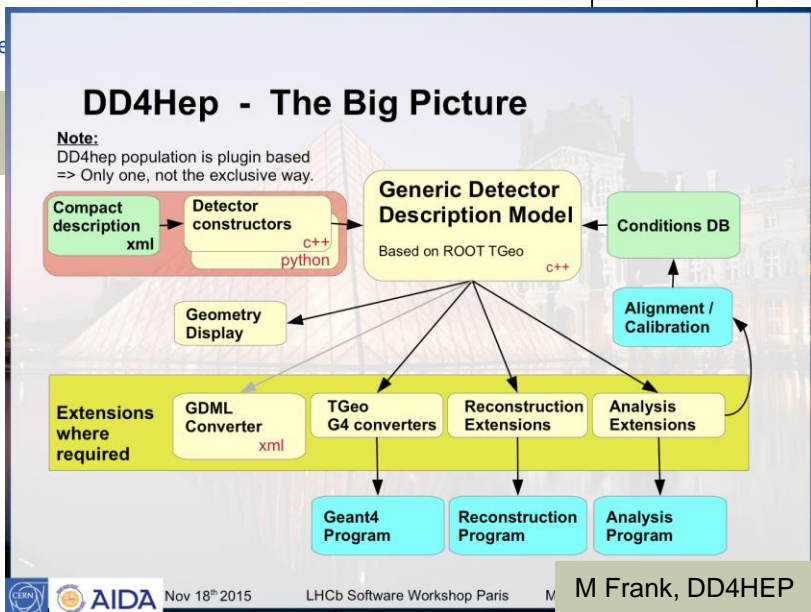
Introduction ○○○○  
 Current developments ○○○○  
 Future developments(?) ○○○○  
 Summary ○○○○

## Possible work plan for upgrade

- Joining forces with ATLAS and CMS
  - serverless implementation
  - we could easily test with current XML
- Review persistent format
  - DD4hep friendly?
  - pre-compiled?
- Versioning and tagging
  - re-think what tags should be/mean
- Migration path
  - conversion?
  - backward compatibility?
- Commissioning
  - Online/Offline population procedures
  - Deployment (SQLite or servers?)

M Clemencic, Detector + Conditions

G Corti,  
Current Simulation



- Current detector description in XML
  - New approach with DD4HEP
- New Conditions DB project upcoming

# More questions from the workshop

- Do we need stripping at a ~ 100 % retention rate?
  - Or is it enough to do “streaming”
- Which data formats do we need?
  - More “turbo like” processing to come
- Multi-threaded applications
- Improvements in user analysis workflows?
- Infrastructure needs
  - Improvements with code repositories, continuous integration, documentation, collaborative tools, development processes, ...



Growing up

- An incremental approach to deployment of new features
  - Adapting to changing requirements and environment
    - ↳ E.g. one order of magnitude increase in HLT output rate
  - Learn from the past, throw away what works badly, keep and improve what works well
- Development in parallel with running production system
  - Physics software in production since 1996
    - ↳ detector design, detector optimisation, HLT+physics preparation, physics exploitation
  - Production system continuously supporting major productions and analysis since 2004
- Strong constraint also for the future
  - Continue to support running experiment
  - Continue to support analysis of legacy data
    - ↳ Minimise pain of maintenance by supporting legacy data in new software
  - Do not underestimate training effort to keep users on board

# Answers?

- What / how much can be done
  - Constrained by maintaining a detector in operations
- Many collaboration opportunities with other experiments
  - Conditions data, detector description, event model description, fast simulation, software framework, ...
- Re-thinking of data processing workflows
  - Which will have implications on the distributed computing model
  - Evisage first discussion at WLCG Workshop @ Lisbon Feb'16
- Currently working towards a roadmap by Q1/16
  - Which will lead us to the Computing TDR by Q4/17

# Summary

- Run 2 data processing in full swing
  - pp processing finished but more programs to exercise
  - Working at design levels
  - All Run 2 workflows foreseen have been successfully exercised
- Brainstorming towards Run 3 has started
  - Very well attended workshop as official kick-off meeting
  - Several seeds have been planted, now see if they grow
  - Many opportunities to collaborate with other experiments
  - Detailed roadmap will be available early next year