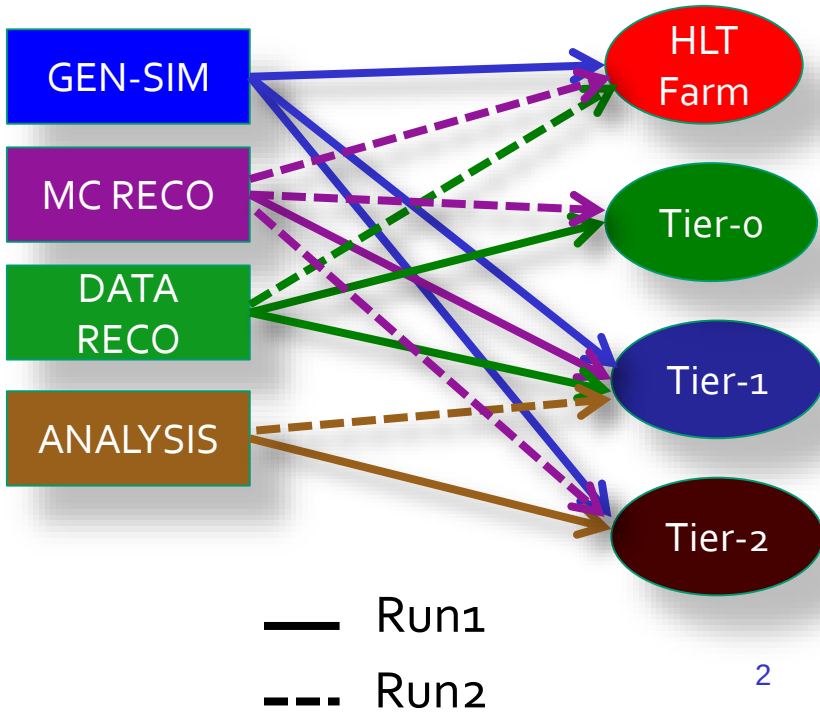
CMS Software and Computing Update

Maria Girone and David Lange

Computing: Ready for Run2

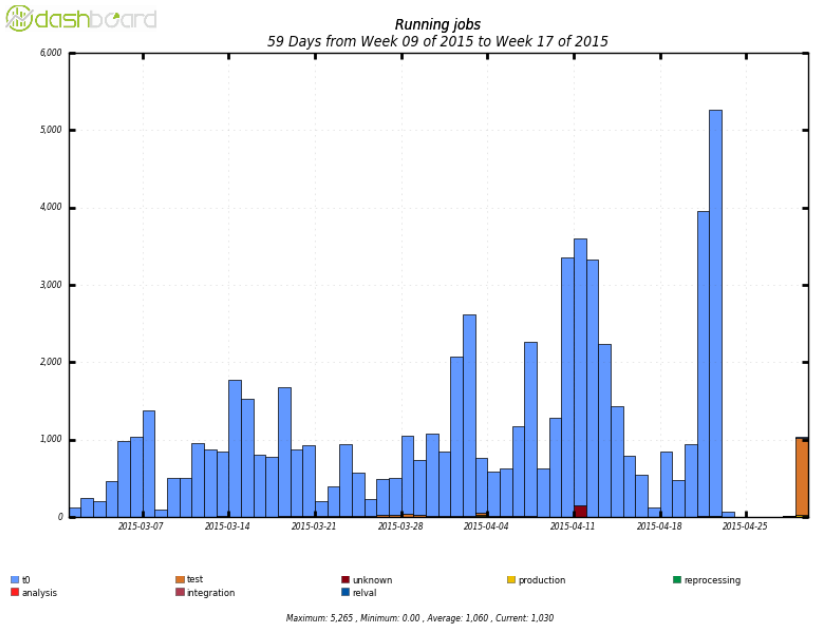
- The software and computing systems, underwent a significant overhaul during the shutdown
  - Improved simulation and reconstruction as a multi-threaded application
  - HLT fully integrated as a production resource, outside data taking
  - Rework of the grid computing facilities to increase their flexibility in handling workflows and reduce the time needed to produce analysis datasets
    - **Data Federation** on Tier-1 and Tier-2 sites
    - **One Central Condor Pool** for all types of resources and applications
    - **Dynamic Data Placement** and automatic clean-up
    - New **Mini-AOD** format in production
    - Deployed a new Distributed Analysis Tool (**CRAB3**)

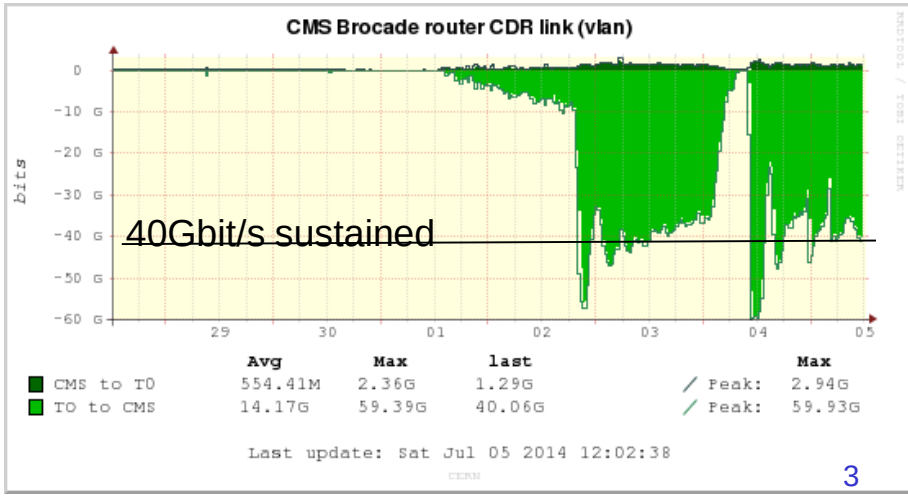
- In Run2 CMS computing resources will work like a coherent and more flexible single system
  - **HLT farm added as a new production resource**



The HLT and AI as Cloud Resources

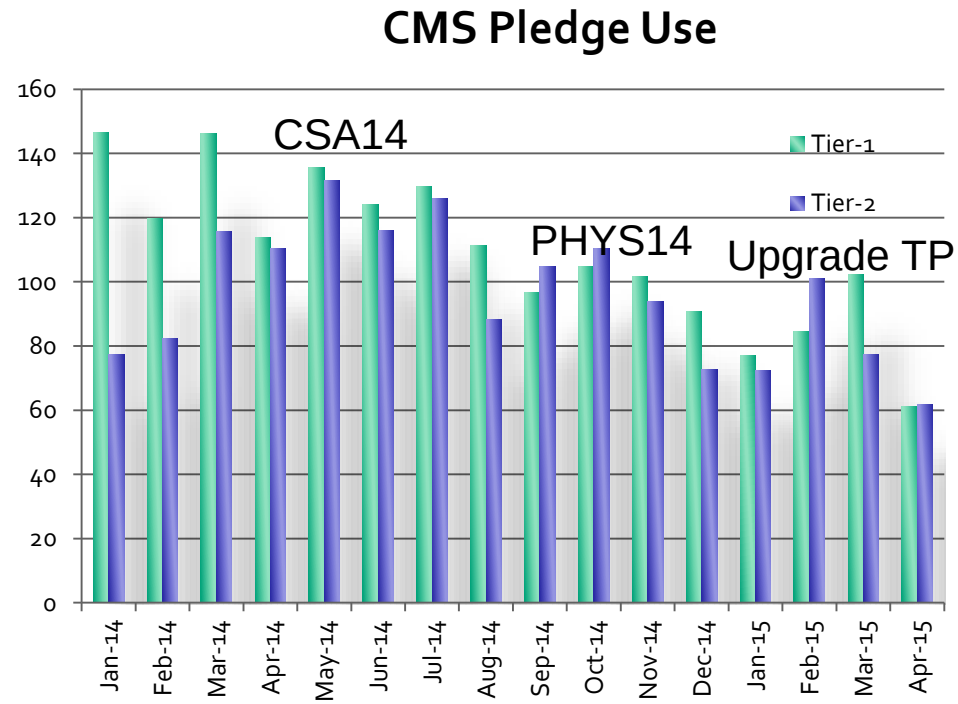
- An addition for Run 2 is the use of the High Level Trigger (HLT) farm for offline processing
  - It is a large computing resource (15k cores) that is similar in size to the Tier-0 in terms of number of cores
  - Successfully interfaced using cloud computing tools. It is similar to the Tier-0 AI
- In 2014 the network link P5 to the computing center was upgraded from 20 to 60Gb/s
  - Far larger than needed for data taking but necessary to access the storage in the computing center for simulation reconstruction
- In 2014, all production workflows have been commissioned including the Heavy Ion reprocessing, Gen-Sim, and Simulation reconstruction
  - All access to data is through the data federation and primarily served from CERN

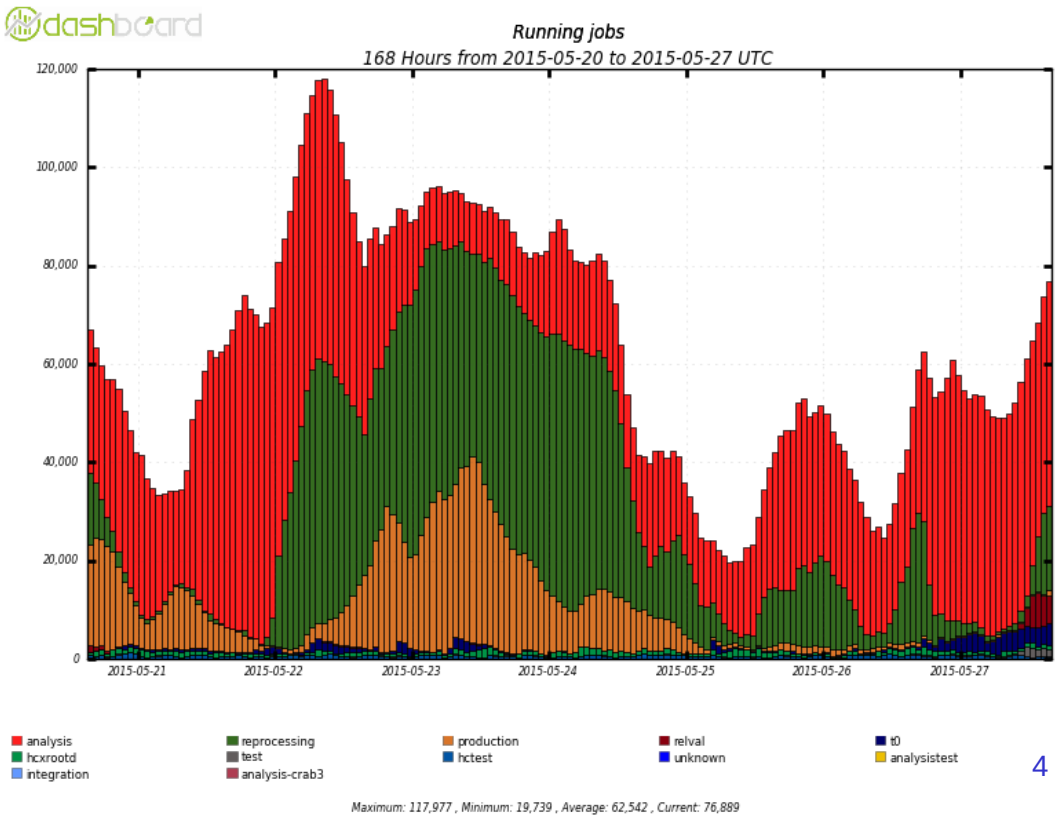




Resource Usage

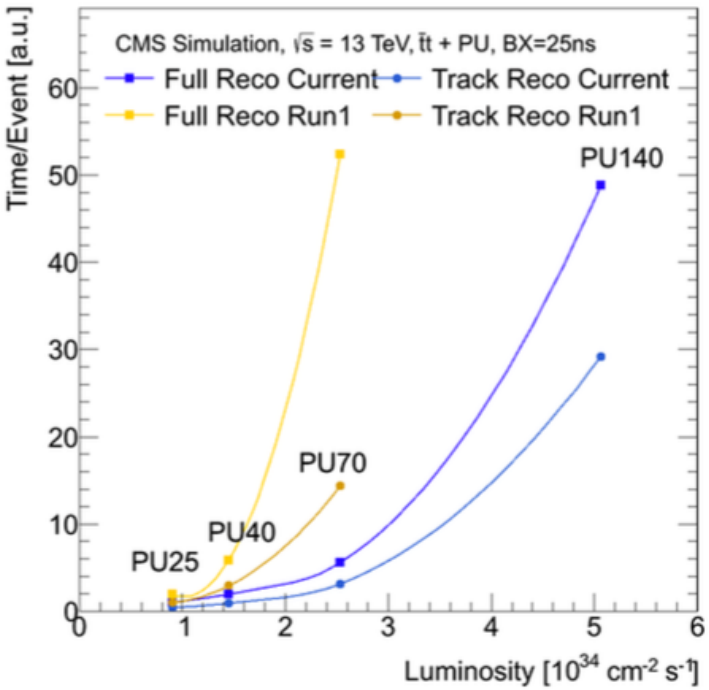
- Utilization was 109% for T1 and 96% Tier-2 from 2014 to now driven by large scale MC productions
  - Central queue peaks at more than 100k running jobs in the last week





Software: Ready for Run2

- Large technical performance gains achieved during LS1
  - Simulation: Factor of 2 gain in CPU utilization, primarily from Russian Roulette sampling algorithm to reduce time spent tracking low-energy particles in Geant4
    - Visible improvements already in the number events/month produced for CSA14(CMSSW6_2) and on-going 2015 production (CMSSW7_1)
  - Reconstruction: Large gains, particularly in tracking area and algorithms appropriate for 25 ns conditions)



These achievements were essential to meet Run2 challenges within resource constraints

Releases follow the LHC Schedule

- LHC Scheduled drives computing activities

7_4_0

- Multi-threaded Reconstruction
- Date: March 2015
- Target: DIGI-RECO and Prompt Reconstruction for start-up

7_5_0

- Date: **June 2015**
- Target: DIGI-RECO and Prompt Reconstruction for 25ns

7_6_0

- Date: September 2015
- Target: End of year Re-RECO

Current production release for Run 2 DIGI-RECO Monte Carlo simulation and Tiero operations

CMSSW_7_5_X is now closing for new features. Most significant changes are to HCAL 25 ns reconstruction and track identification

Collision at 13 TeV