



# CMS Cloud Activity on the HLT (and AI)

Andrew Lahiff

33<sup>rd</sup> GridPP Collaboration Meeting, Ambleside



# Introduction

- Areas of cloud activity in CMS
  - **HLT cloud**
  - **CERN Agile Infrastructure** (Meyrin, Wiger)
  - Tier-0
  - User analysis
- Significant contribution from the UK
  - David Colling, Adam Huffman, Andrew Lahiff, Daniela Bauer,...

# HLT

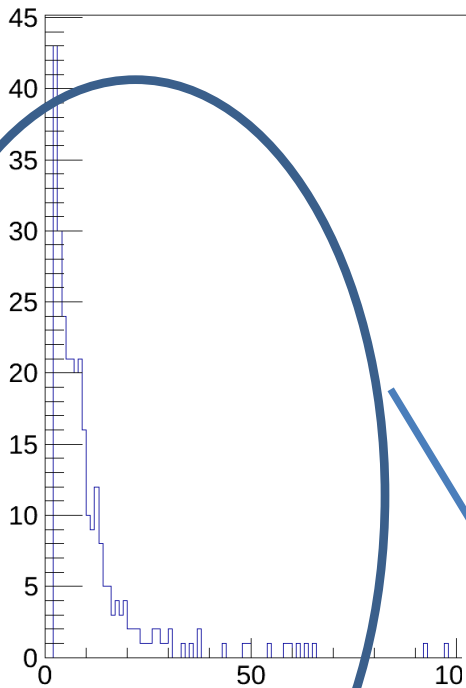
- What is the CMS High Level Trigger?
  - Compute farm with over 13000 cores located at Point 5
  - No large local storage system, some nodes have small disks
- Computing power
  - 195 kHS06 (150 kHS06 easily useable)
  - HLT is comparable in capacity to entire CMS T1 request
- We should make the most possible use of the resources
  - Use the HLT as production resource for processing during LS1
  - Use the HLT whenever possible during Run 2, e.g.
    - Maintenance & machine development periods
    - In the gaps between fills
    - Possibly even during running if conditions mean that the HLT isn't fully required

# HLT

- Between stable beams in 2012

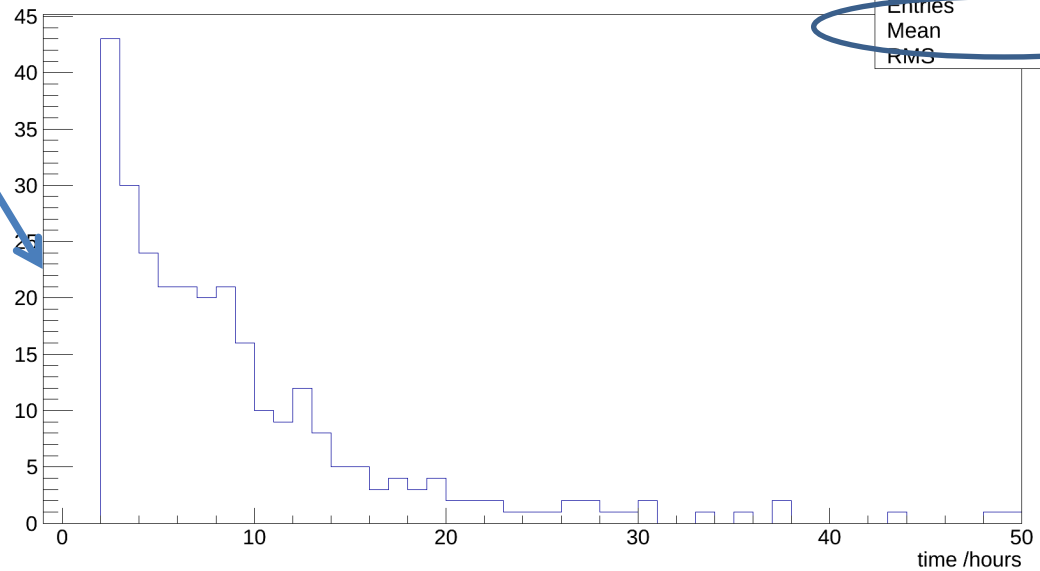
Time when beams not stable (end of SB in one fill to SB in next)

Entries	296
Mean	13.66
RMS	26.92



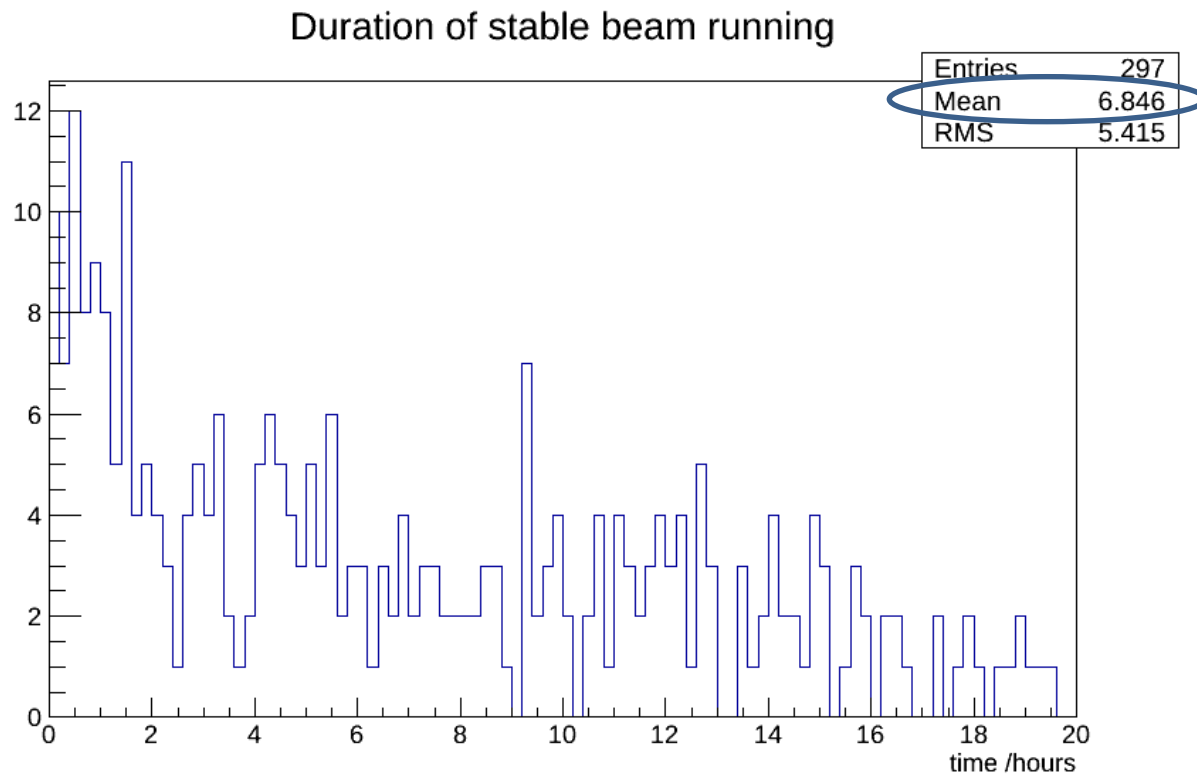
Time when beams not stable (end of SB in one fill to SB in next)

Entries	296
Mean	9.32
RMS	7.874



# HLT

- Stable beams in 2012



On average, stable beam is ~50% of time between stable beam.

Therefore, HLT cloud should be running for >60% of 2015.

That means a resource of ~2/3 of all our T1 CPU capacity.

# HLT

- Why a cloud?
  - Need to guarantee there is no interference with the HLT's main role
- Essential to be able to migrate on & off quickly
  - Much work has gone into ensuring we can start up jobs on the HLT cloud as fast as possible
    - Need to be able to start up as quickly as possible in order to make best usage of resources after LS1
  - Required lots of tuning/fixes/etc
    - OpenStack
    - HTCondor
    - glideinWMS
    - WMAgent

# HLT

- Using OpenStack Grizzly
- 60 Gbit link to CERN
  - upgraded from 2 x 10 Gbit in February
- Currently have just over 4000 cores available
  - 144 x C6100 (12 cores, 20 GB RAM)
  - 150 x C6220 (16 cores, 28 GB RAM)
- Moving nodes between DAQ and OpenStack
  - “Shifter” under development
  - GUI tool for allocating part or all of the cluster to either DAQ or OpenStack, and migrating between them

# CERN AI

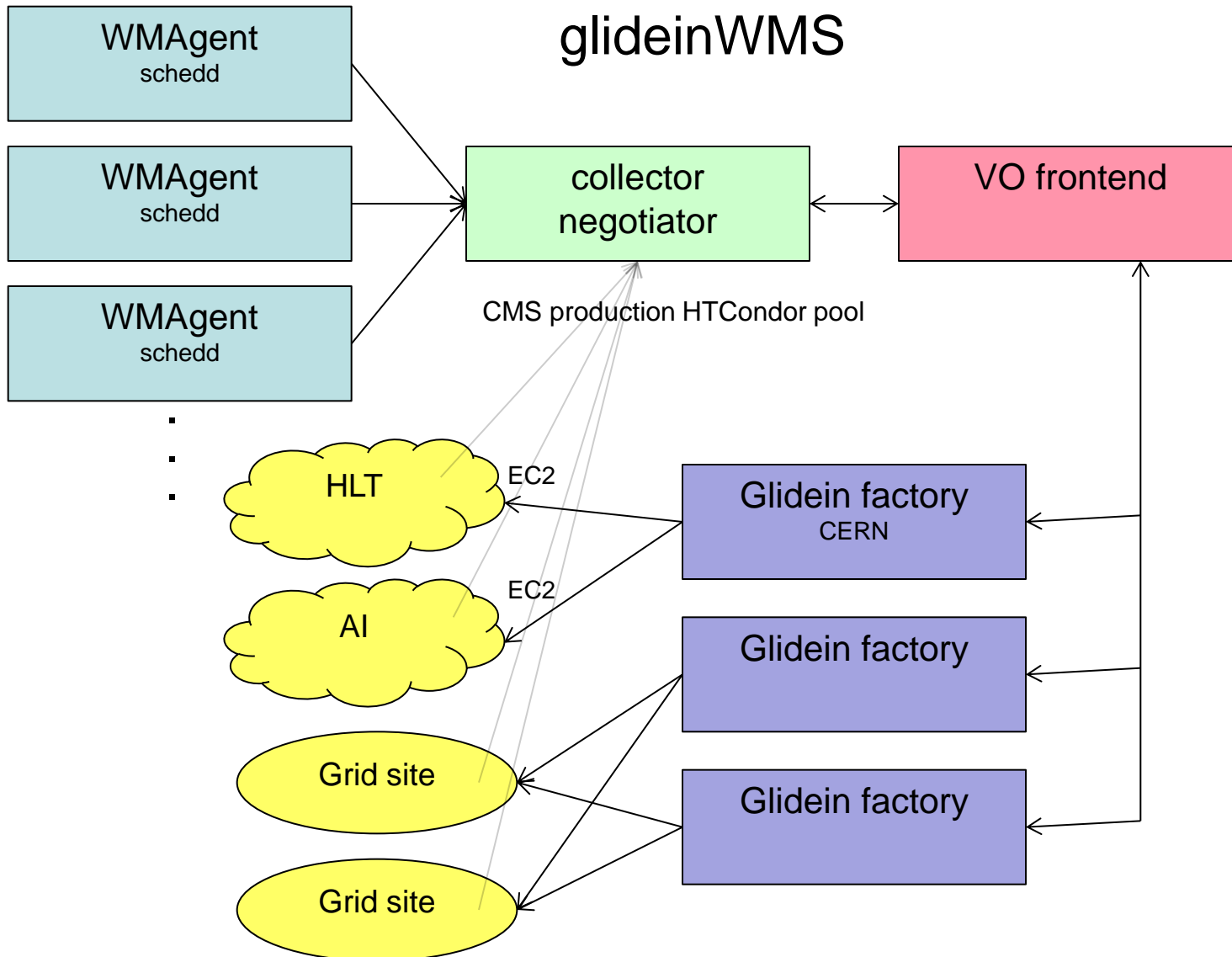
- CERN's OpenStack cloud
- Doesn't contain tuning that was done on the HLT OpenStack instance...



# HLT/AI cloud status

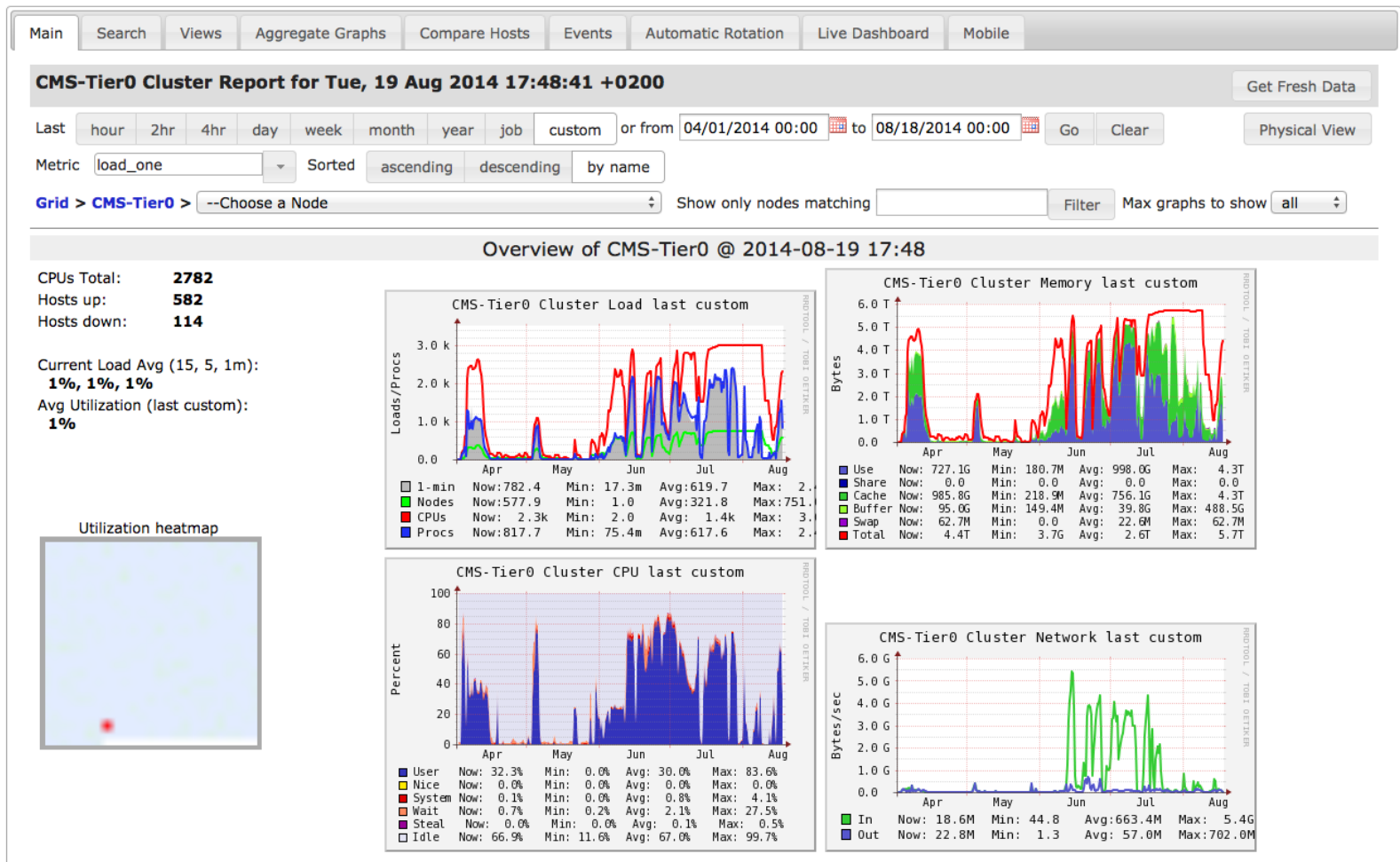
- HLT & AI clouds added to production glideinWMS infrastructure in May
- Being used routinely for production
  - Reprocessing: workflows assigned to CERN LSF, HLT, AI, reading input data from EOS
  - MC production: workflows assigned to HLT, AI + many T1s, T2s

# Job submission



# Monitoring

- Ganglia monitoring for AI VMs

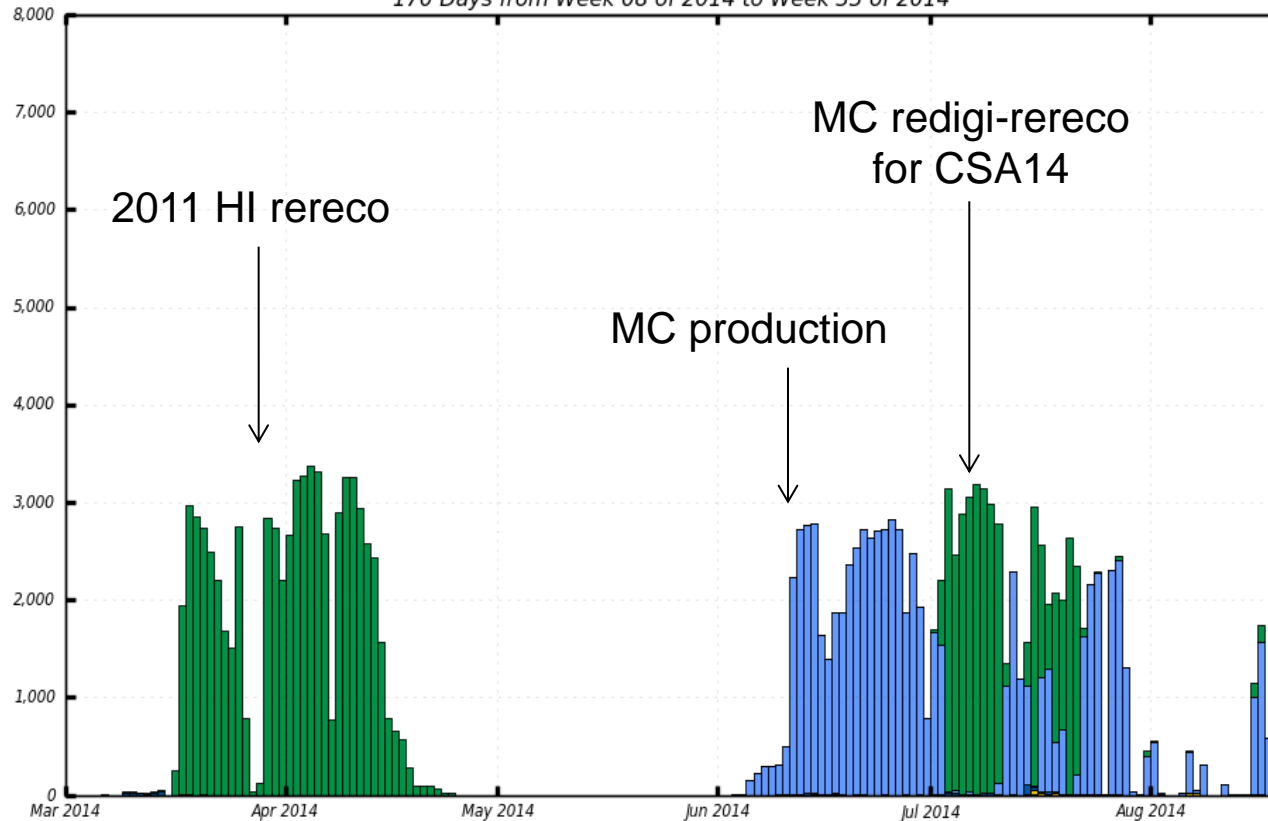


# HLT usage since March



## Running jobs

170 Days from Week 08 of 2014 to Week 33 of 2014

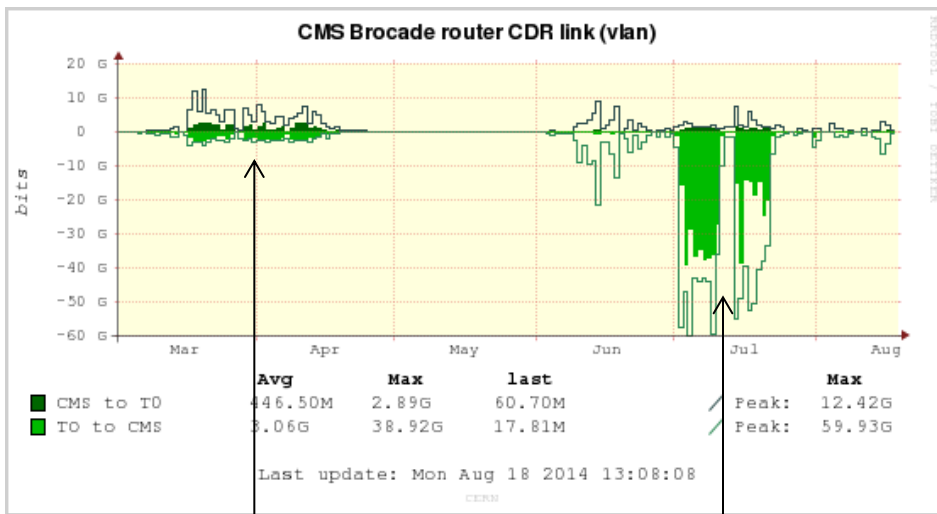


Maximum: 3,370 , Minimum: 0.00 , Average: 1,094 , Current: 2.00

# Network usage since March

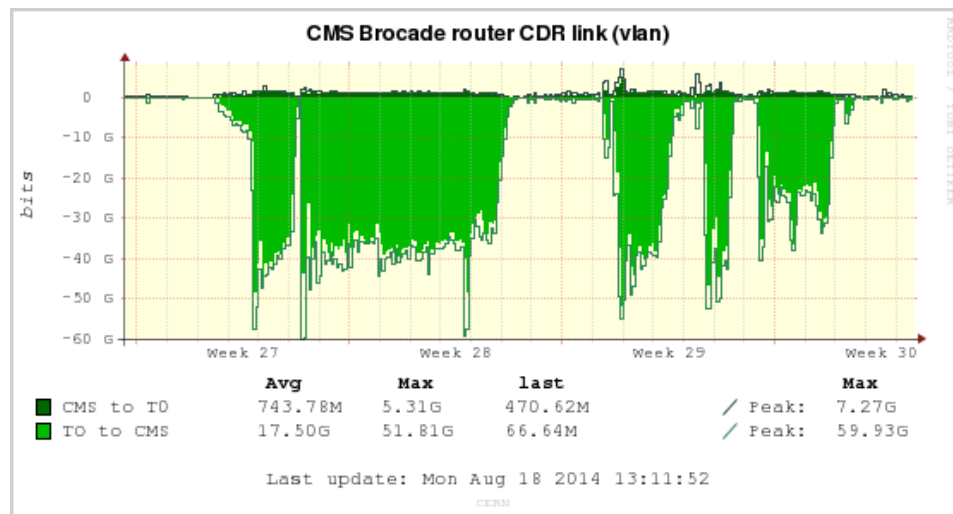
- 60 Gbit link from P5 to CERN
- Usage since March:

July



Heavy ion rereco

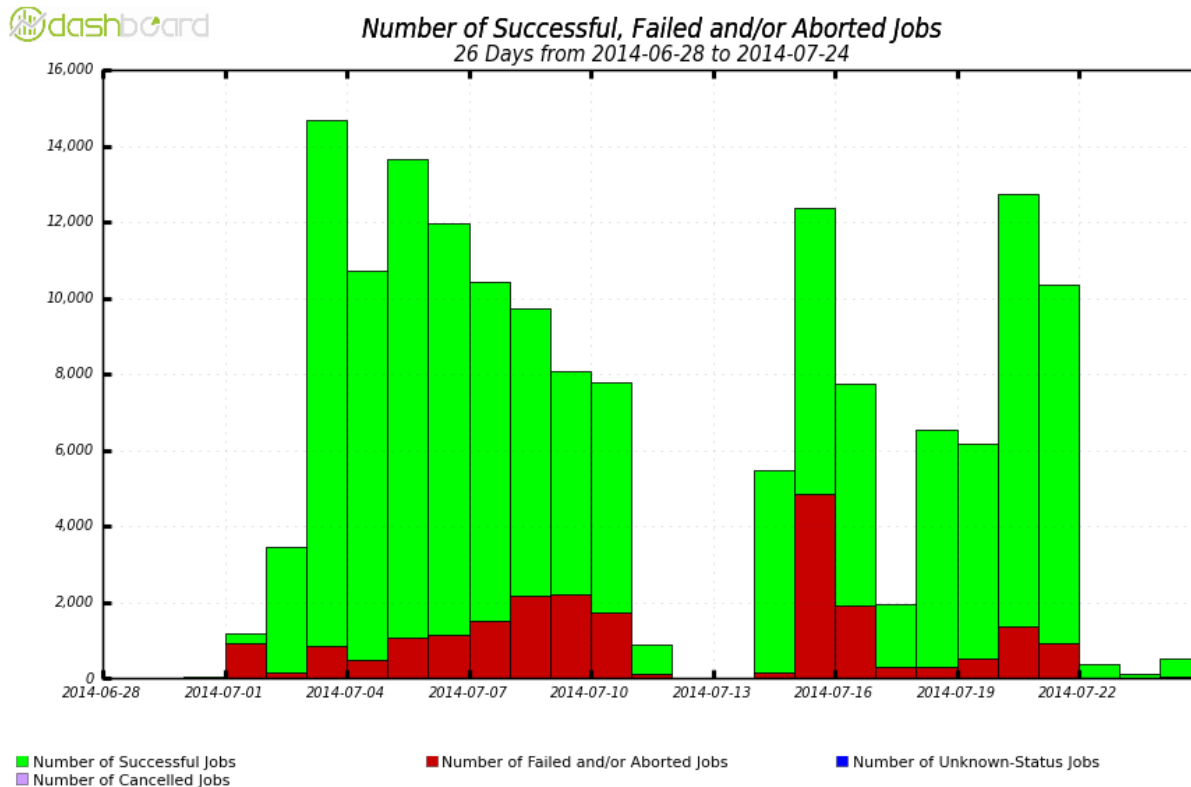
MC redigi-rereco  
for CSA14



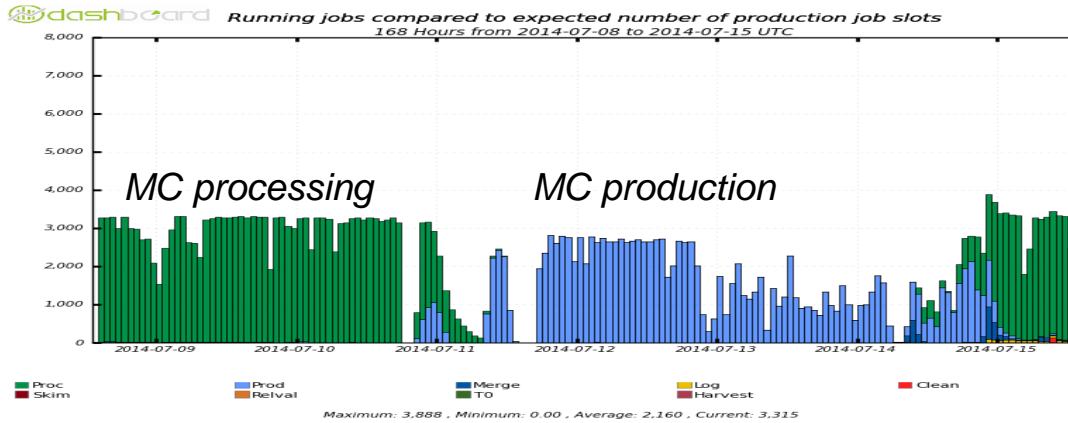
~3000 MC processing jobs running

# HLT efficiency

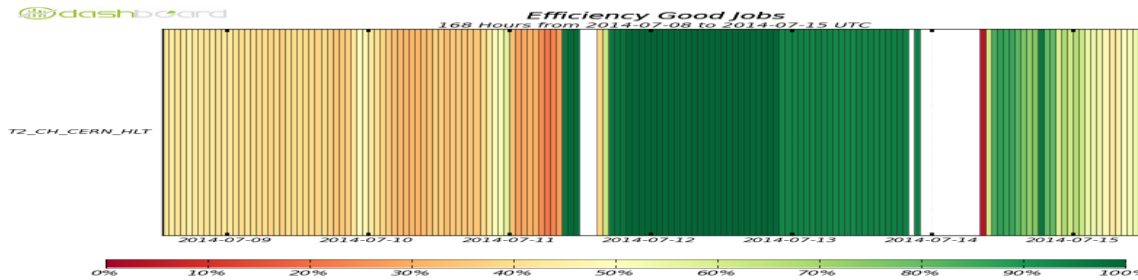
- Job successes & failures of CSA14 MC processing
  - Failures due to file access problems
    - Not cloud-related: “hot” files on EOS (replication needed)



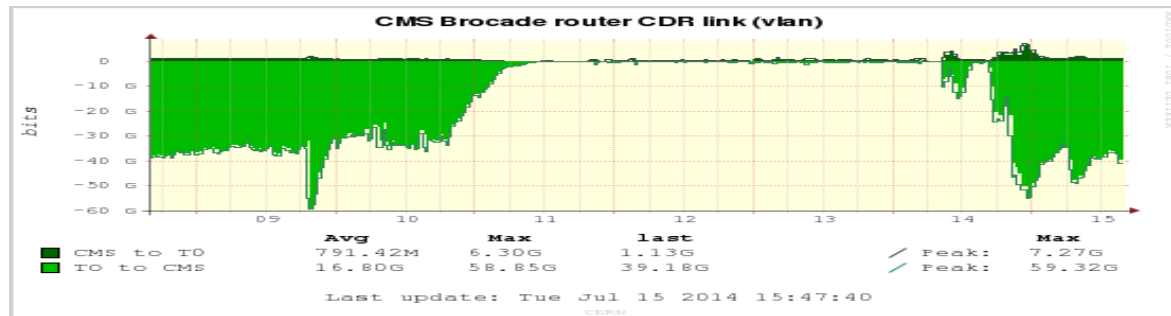
# HLT efficiency



activity



CPU efficiency



Network between P5 and CERN





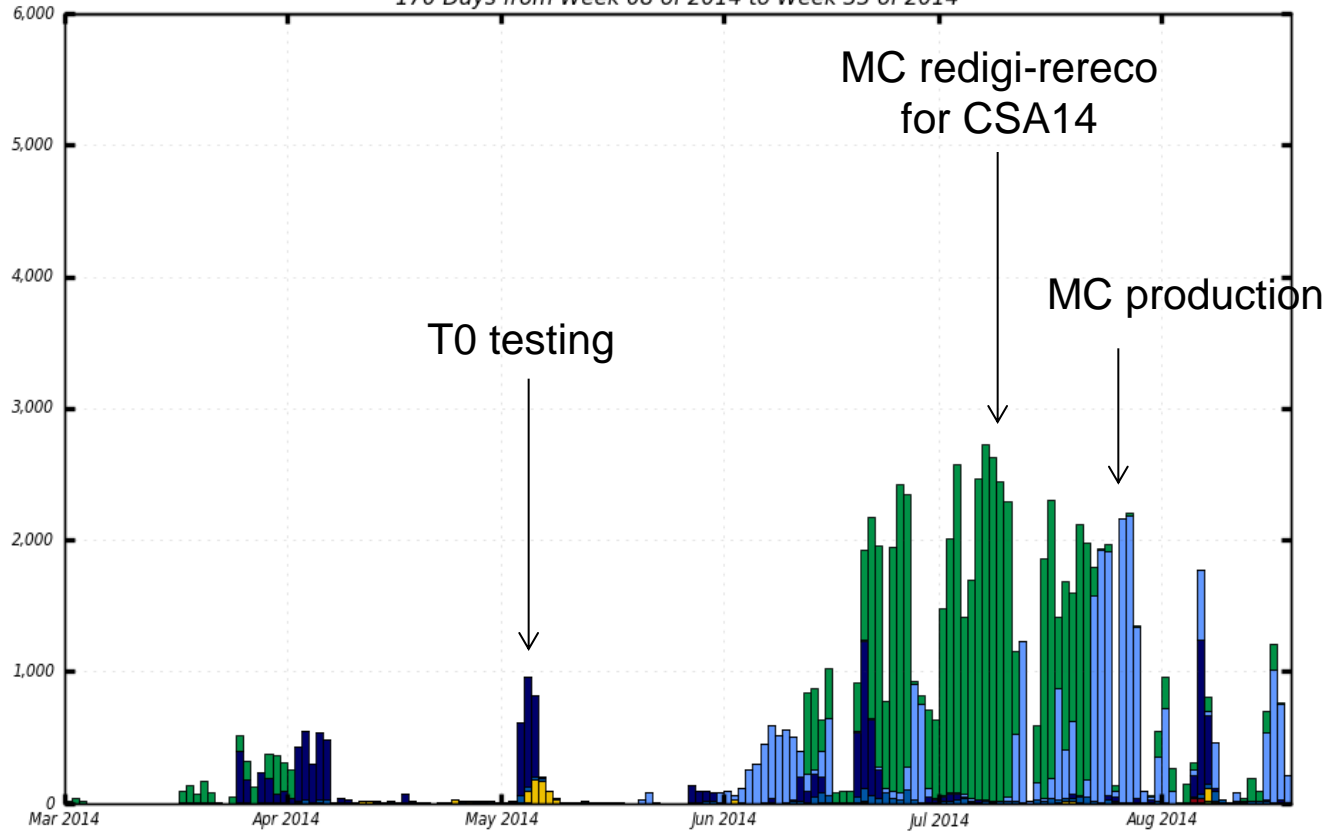
# AI usage since March

T2\_CH\_CERN\_T0 (OpenStack AI "CMS Tier0" project)



## Running jobs

170 Days from Week 08 of 2014 to Week 33 of 2014

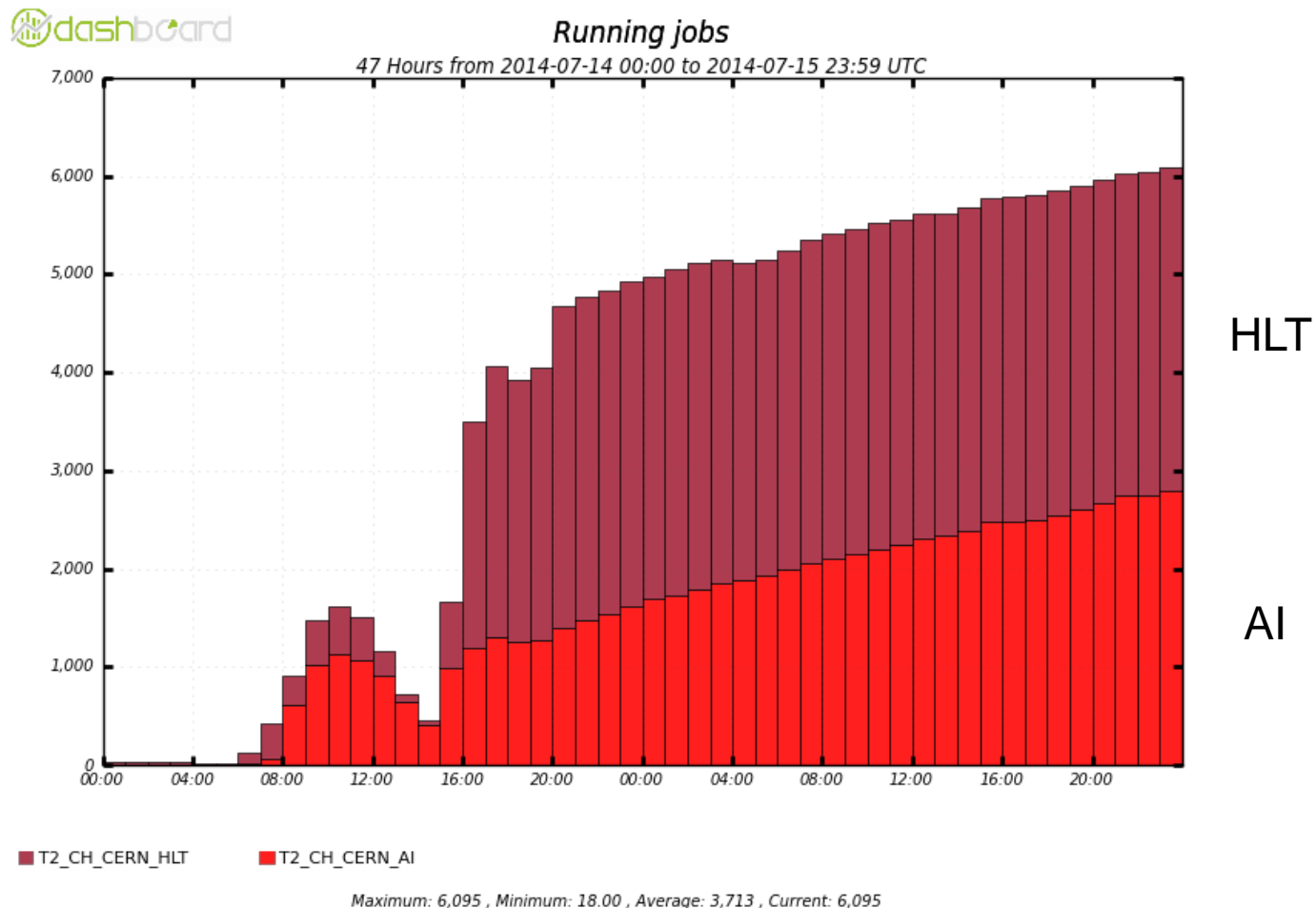


■ Proc ■ RelVal ■ Prod ■ Skim ■ T0 ■ Clean ■ Merge ■ RelVal ■ Log ■ Harvest

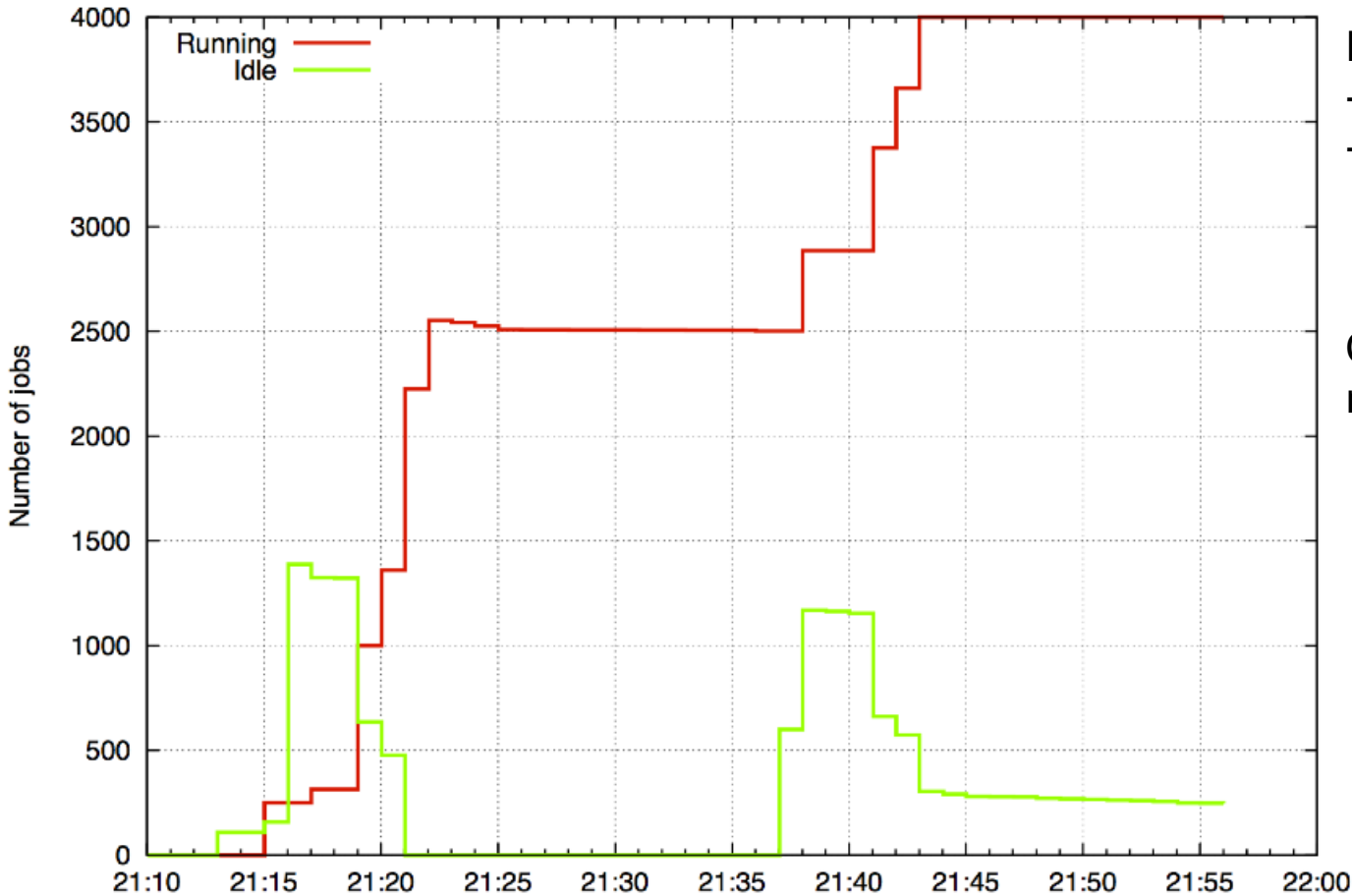
Maximum: 2,724 , Minimum: 0.00 , Average: 539.95 , Current: 217.00

# Start-up rates

- HLT seems much faster



# Ramp up time on the HLT



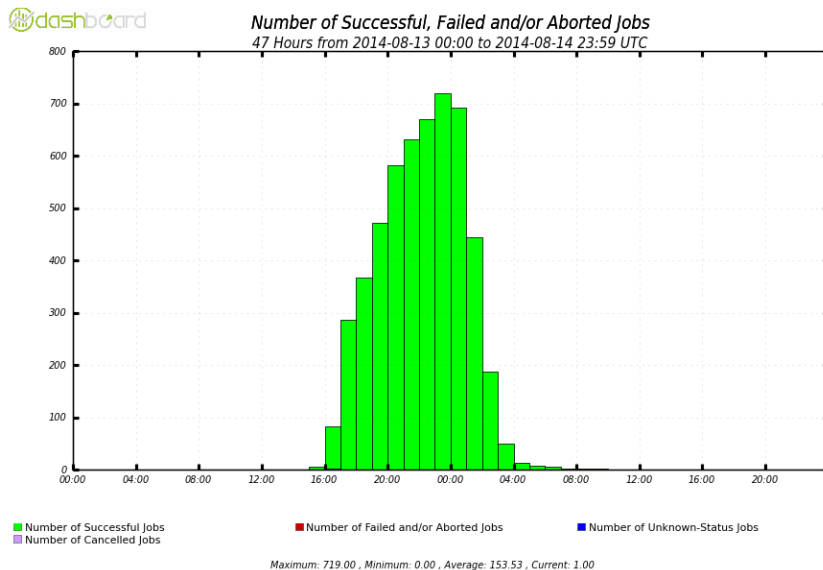
From scratch

- VMs booting
- HTCondor on VMs starting up & contacting the collector

Could get 4000 jobs running in ~25 mins

# Wigner

- Aim to treat nodes at Meyrin & Wigner transparently within a single site
- Currently have Wigner setup as a separate site for testing
- Example CSA14 MC processing workflow:



# 2011 HI rereco

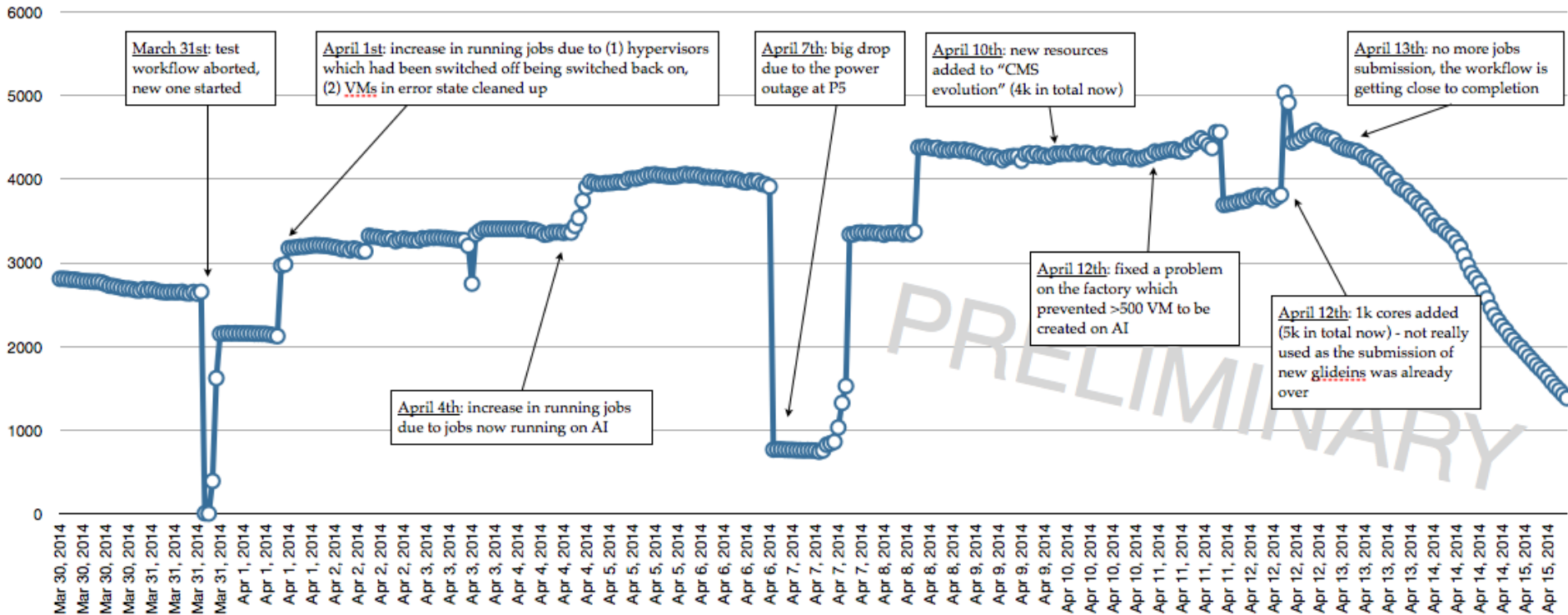
- Requests made in March 2014 for HI rereco
  - Essential for CMS contribution to Quark Matter conference in May
- Status by late March
  - A workflow running on HLT cloud making very good progress
  - A workflow running at Vanderbilt not expected to complete in time for the conference
    - Time per event much longer than expected
    - Storage problems at Vanderbilt

# 2011 HI rereco

- Decided to try 2 methods
  - Run a clone of the workflow on the HLT & AI clouds
  - Run another clone at multiple sites reading data using xrootd
    - Failure rate very high, didn't make good progress
    - Workflow aborted
- The workflow running on the HLT & AI clouds completed successfully in time for the conference
  - Major success of cloud resources for CMS

# 2011 HI ReReco

## Heavy Ion 2011 data re-reco in April 2014 - jobs running on HLT+AI



April 16th - Credits to D. Mason, A. Lahiff, A. McCreia, M. Sgaravatto. Data collected by A.Lahiff. Details on periods by A.Lahiff, M. Sgaravatto. Overview plot by D.Bonacorsi

# Summary

- HLT & CERN AI clouds used routinely for production
  - Both processing and MC production
  - Already an important resource
    - Without the HLT cloud, CMS wouldn't have been able to present work at the Quark Matter conference
- Future
  - Analysis jobs
  - Include Winger in routine production activities
  - Additional cloud sites