Remote Operation of the CMS Data and Workflows

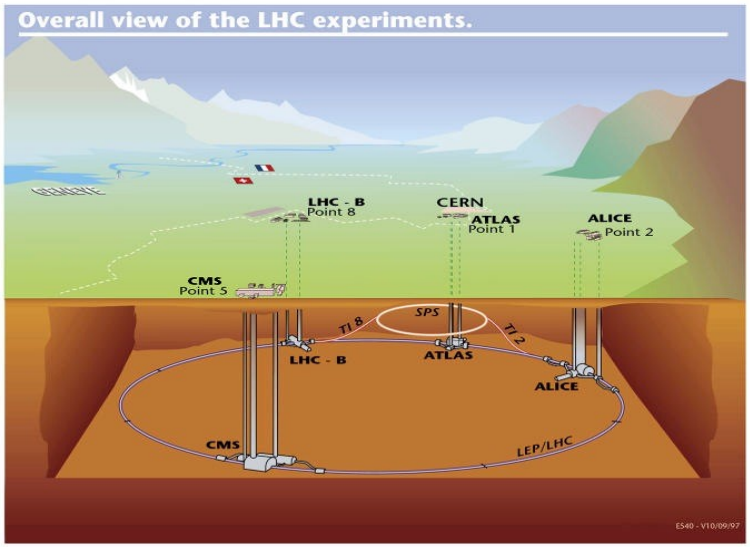
David A. Mason, Fermilab

On behalf of the CMS Offline and Computing Projects

Outline

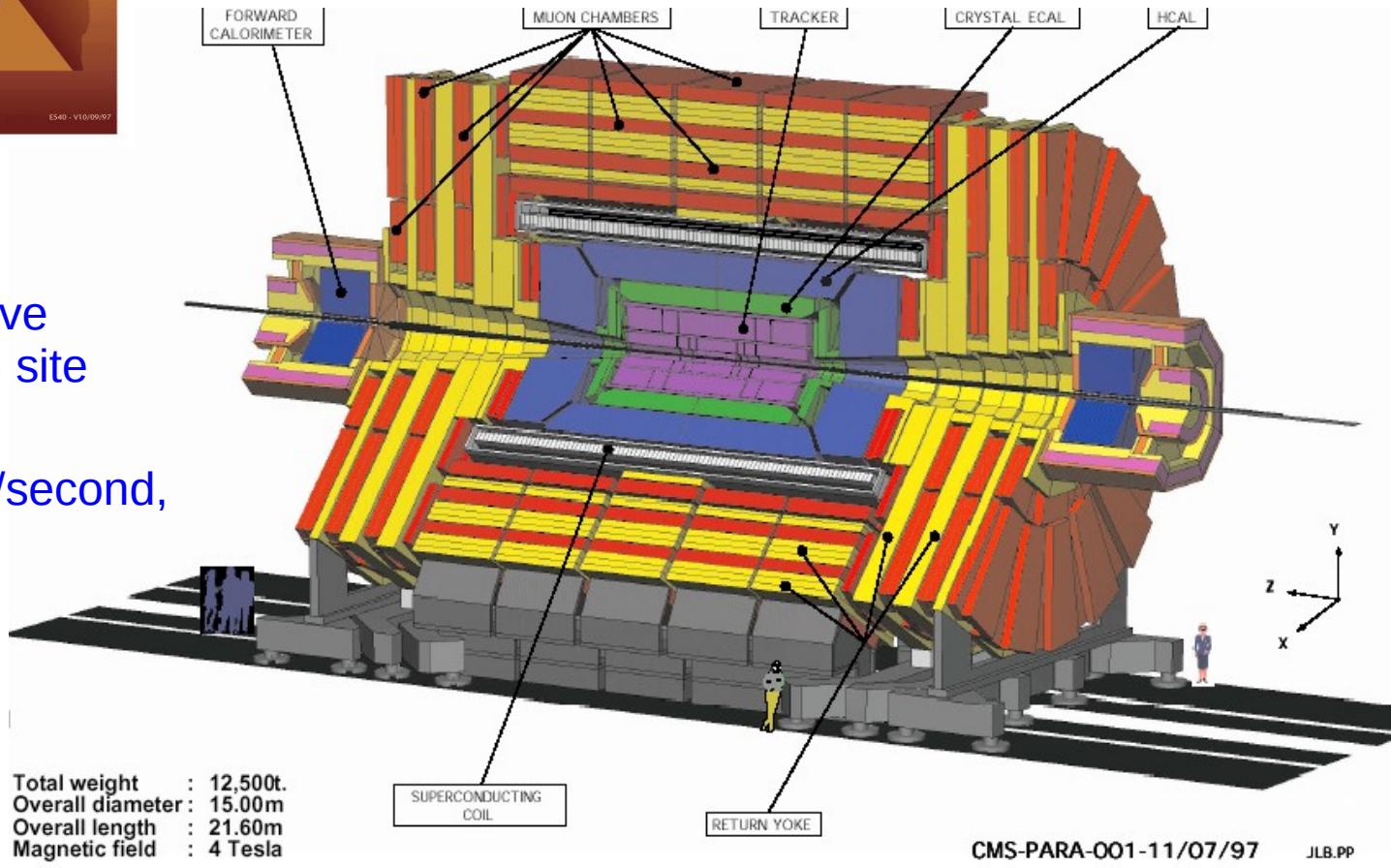
- Introduction to CMS and its Computing Model
- CMS Data Operations
- Fermilab Remote Operations Center
- The CMS Tier 0 system
- CRAFT and global run datataking experience

CMS

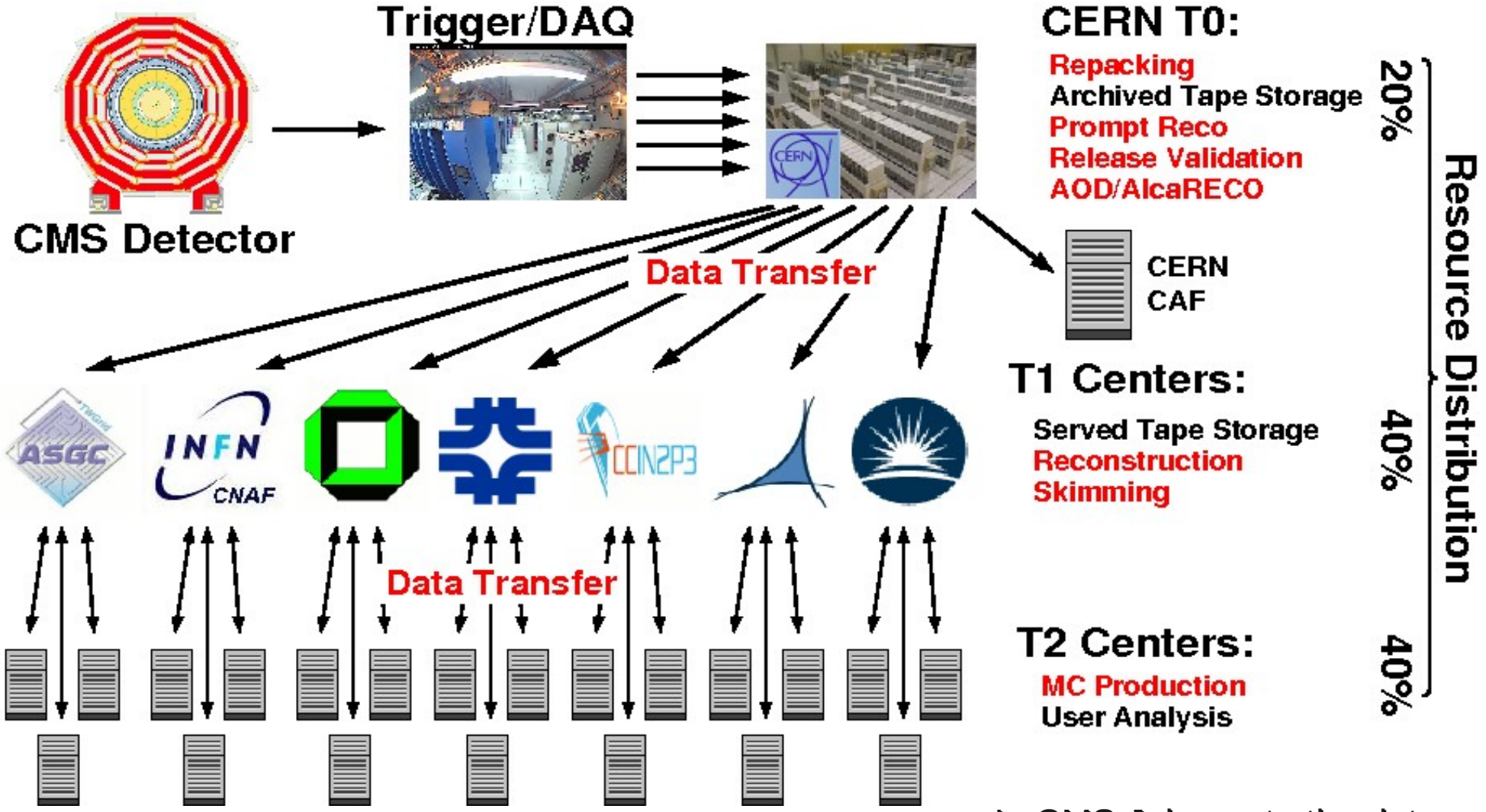


LHC p-p collider experiment located about a half hour drive from the main CERN Meyrin site

Nominal datarate of 450 MB/second, with bursts of 1 GB/second



CMS Computing Model



Tasks in red are performed by central data operations teams

CMS Data Operations

- Everything in red in the previous slide is in the hands of CMS Data Operations.
- CMS employs two teams, one at CERN, the other at FNAL
  - Both teams take full advantage of grid based tools to manage CMS workflows independent of physical location
  - 7 hour time difference allows FNAL team to “take over” after CERN day ends.
- All Data Operations are Remote Operations
  - Generally are not sitting at the worker nodes themselves
- CERN CMS Centre and FNAL Remote Operations Center (ROC) serve as central “home bases” for the two data operations teams.
- Here we'll concentrate on the operational experience of running CMS data processing workflows away from CERN – via the FNAL ROC.

The FNAL Remote Operations Center

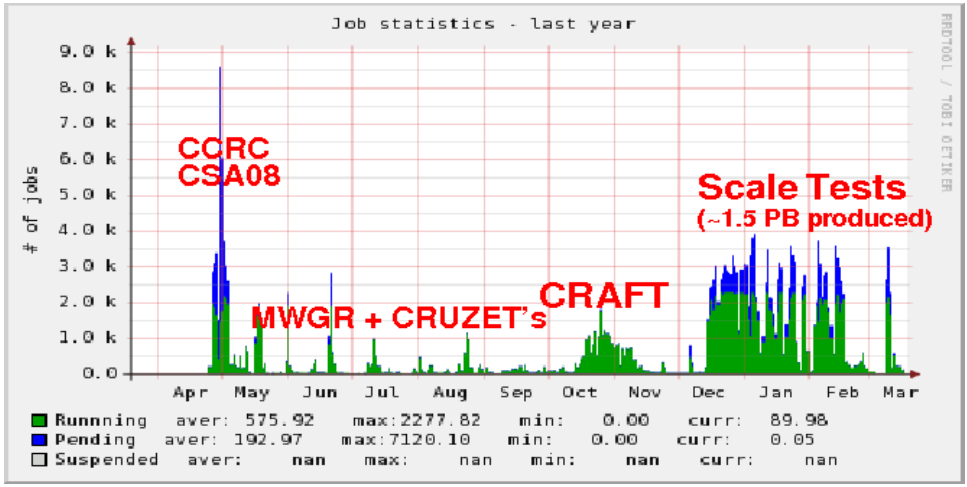
Preparations for first LHC beam, September 10 2008

FNAL Remote Operations Center

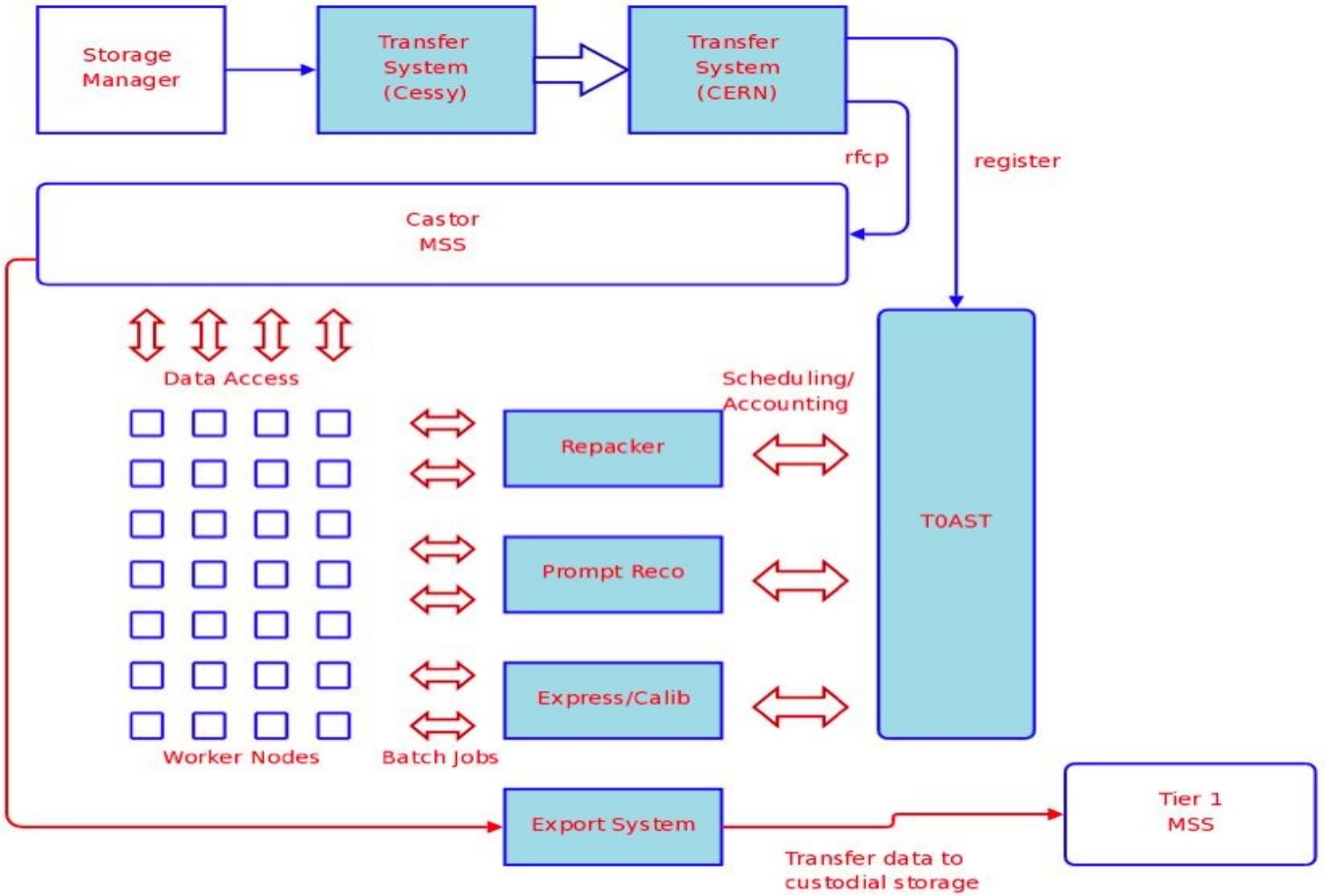
- 8 Linux workstations, each with 3 screens, one manned by the FNAL data operations team, another by the FNAL Tier 1 facilities team. (other terminals manned by various detector subgroups, data quality monitoring shifters)
  - See Erik Gottschalk & Lassi Tuura's talks in this conference
- State of the art communications technology
  - Steerable webcams at multiple points around the room
  - Polycom videoconferencing terminals for face to face communication
  - 4 Macintosh powered projector screens for overhead display of relevant plots, monitoring, status webpages, etc.
  - Live high definition “Window to CERN” audio/video link
    - Very much like having the two sides of the Atlantic joined by TV screen
    - Profound psychological effect of seeing + hearing people on the other side – with only phone connections, learning what was happening at CERN was like “pulling teeth”. With the video connection, CERN shifters see the remote shifters, and voluntarily inform us of run changes and plans.

CMS Data Ops Commissioning Activities

- The CMS data operations teams have been in existence since spring of 2007, working to commission CMS computing systems in concert with detector commissioning activities
- Throughout this past spring and summer CMS took data in “global runs” of cosmic ray datataking, ramping up to first beam in September.
  - Because these global runs were actual datataking events, both the FNAL ROC and CMS Centre were hubs of debugging information exchange between detector and computing/software experts.
  - Global runs at regular intervals have provided “tools down” checkpoints to push on and evaluate the current state of the computing systems.
    - Time between global runs allowed new components to be developed, debugged and commissioned
  - Through first beam in September, in terms of computing, the global runs were primarily dedicated to commissioning the Tier 0 systems.



The CMS Tier 0 System

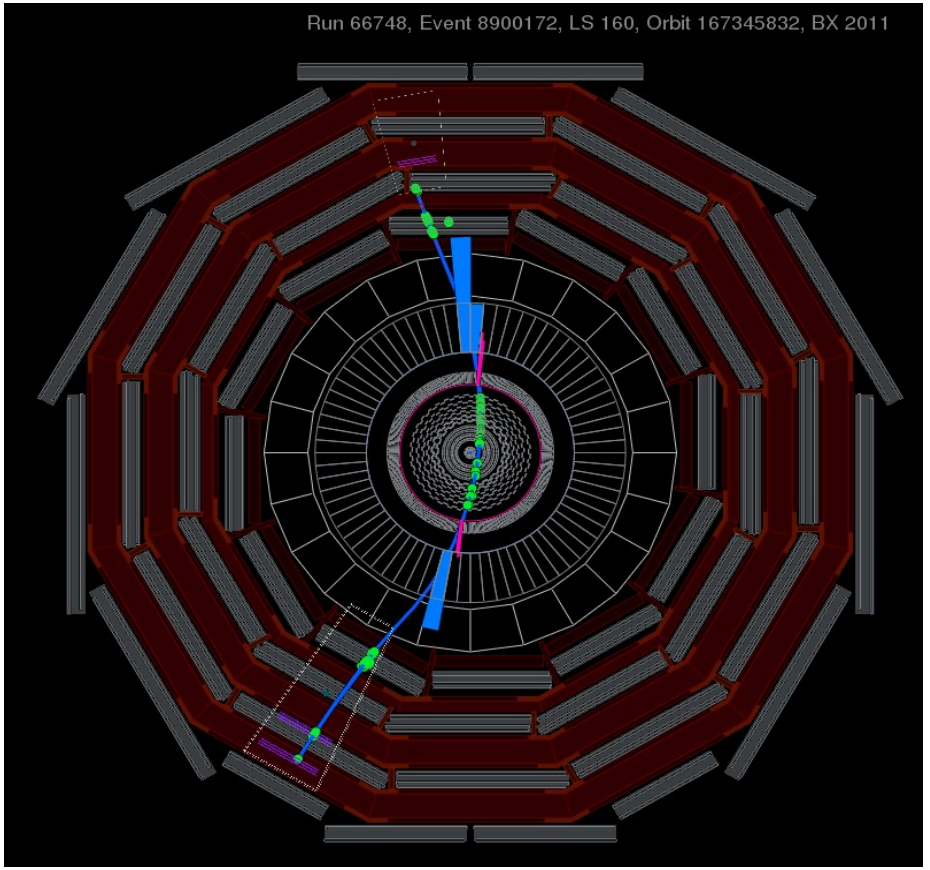


The CMS Tier 0 System

- The CMS Tier 0 system handles:
  - Repacking – reformatting of raw data and splitting into primary datasets
  - Prompt Reconstruction – first pass with a 24 hour turnaround
  - Automated subscriptions for data transfer & disk buffer management
  - Alignment and Calibration data skimming
- Processing jobs submitted at CERN to the CMS T0 LSF queue.
- The T0 system is based around an oracle state tracking database, “Tier 0 Activity State Tracker”, or T0AST.
  - The entire lifecycle of data through the T0 systems and out to the T1 are to be tracked by this state system
- After much hard work from developers, enough components of this system were ready to be able to process collision data last September.
- The September 19 LHC magnet failure however left us to exercise these systems with cosmic ray datataking.

CRAFT:

- 1 month of continuous datataking from mid October to mid November last year
- Detector ran with magnetic field on – in “beam on” condition, but without beam.
- Data rates held at nominal collision rates – with an average event rate over the month around 300Hz.
- Data Operations maintained 24 hour shift coverage throughout the CRAFT run, with shifts at the CMS Centre and FNAL ROC.
- Operators from either site monitored and controlled the T0 (repacking, prompt reco), T1 systems (skimming) and data transfers to T1's throughout the month long datataking period.



Cosmic ray muon event passing through entire CMS detector

T0 Remote Monitoring in CRAFT

[T0Mon user documentation] [Contact]

T0 status @ Tue Nov 4 13:31:44 2008

"T0Mon" Web based T0 monitoring tool

Runs [count=647, query=201ms, analysis=2723ms]

All 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65

Run ID	Start	End	Process	Version	Reco version	Global tag	Status	Streamers	Repacked	Merged	Reconstructed
69276			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	CloseOutRepack	857	198 (64%)	28 (22%)	0 (0%)
69273			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	CloseOutRepack	1138	70 (30%)	26 (58%)	0 (100%)
69272			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	CloseOutRepack	523	36 (28%)	15 (52%)	0 (100%)
69269			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	10	7 (100%)	6 (100%)	2 (100%)
69207			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	157	57 (100%)	25 (100%)	8 (100%)
69165			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	352	127 (100%)	53 (100%)	18 (100%)
69109			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	832	363 (100%)	168 (100%)	41 (100%)
69094			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	534	428 (100%)	122 (100%)	2 (100%)
69080			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	464	235 (100%)	56 (100%)	20 (100%)
69071			HLT	CMSSW_2_1_9	None	CRAFT_V3P::All	Complete	196	102 (100%)	26 (100%)	9 (100%)

Repack stats [29ms]	
New jobs	0
Used jobs	99
Successful jobs	84724
Failed jobs	95
Abandoned jobs	0

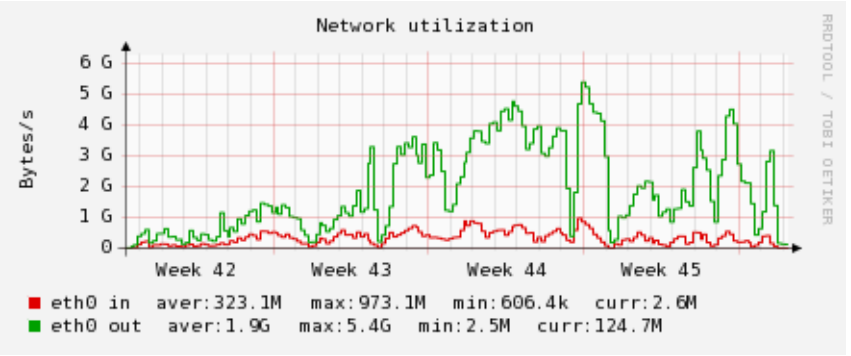
Merge stats [11ms]	
New jobs	1
Used jobs	0
Successful jobs	21549
Failed jobs	1
Abandoned jobs	0

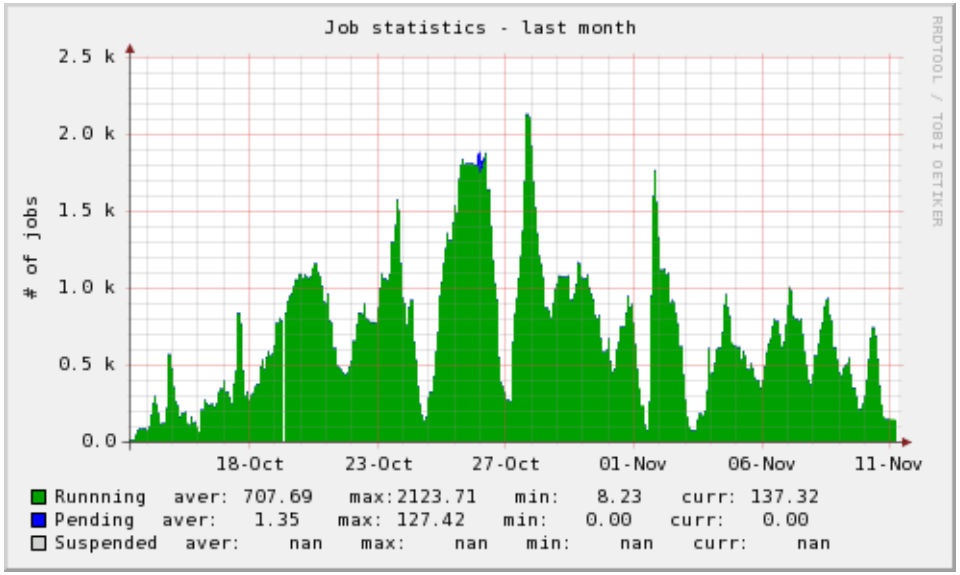
Reco stats [15ms]	
New jobs	67
Used jobs	0
Successful jobs	26478
Failed jobs	962
Abandoned jobs	516

T0AST DB tracks state of all data as it passes through the T0

T0 LSF job monitoring

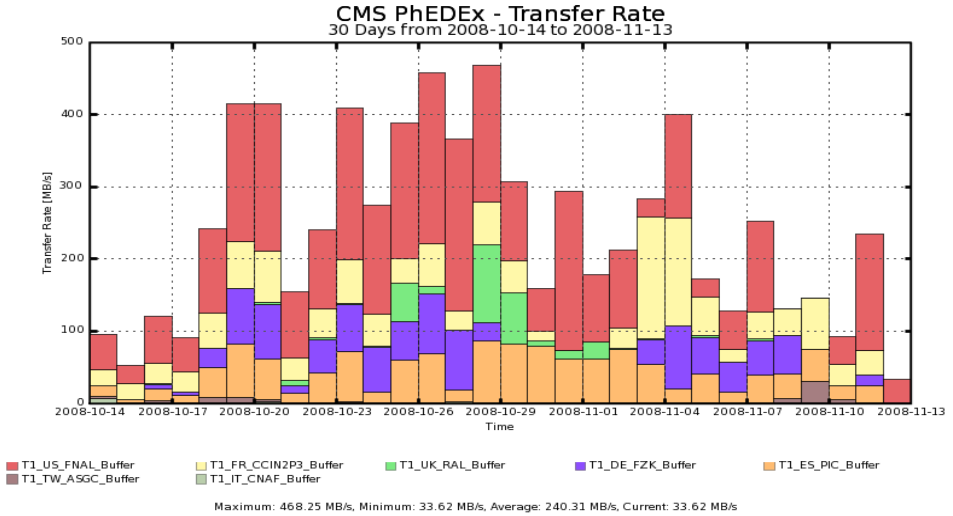
- CERN's “lemon” plots are an indispensable tool for monitoring the state of T0 jobs
- To the left is file I/O for the T0 farm (top) and total job counts (bottom) throughout CRAFT
- Peaks and valleys are changes in trigger rate as run conditions changed throughout the month.

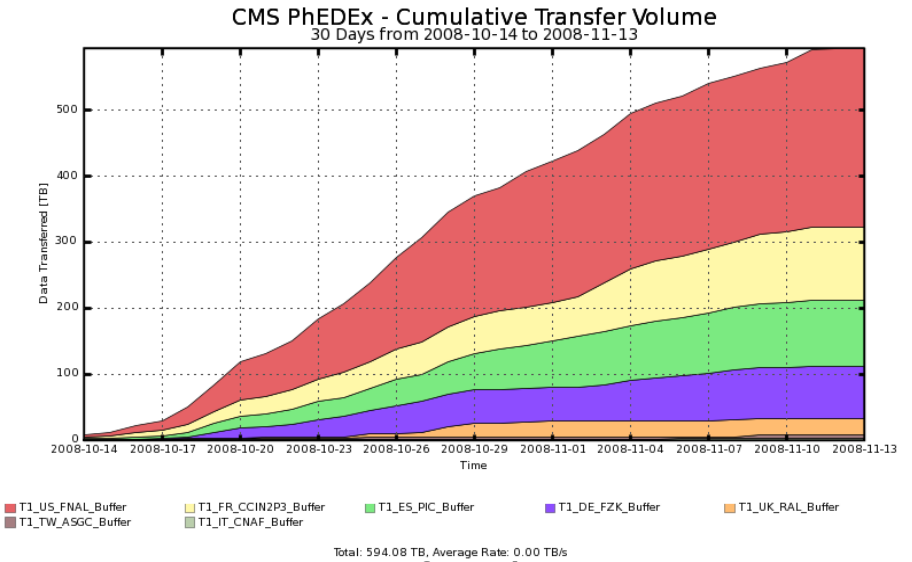




Data Transfer Monitoring

- Processed files must be transferred to their custodial T1 within 24 hours of being produced to free up T0 disk buffers.
- Problems must be detected quickly or we risk losing data.
- CMS' PhEDEx ("Physics Experiment Data Export") system provides monitoring tools to detect transfer issues
- CRAFT transfers shown at left

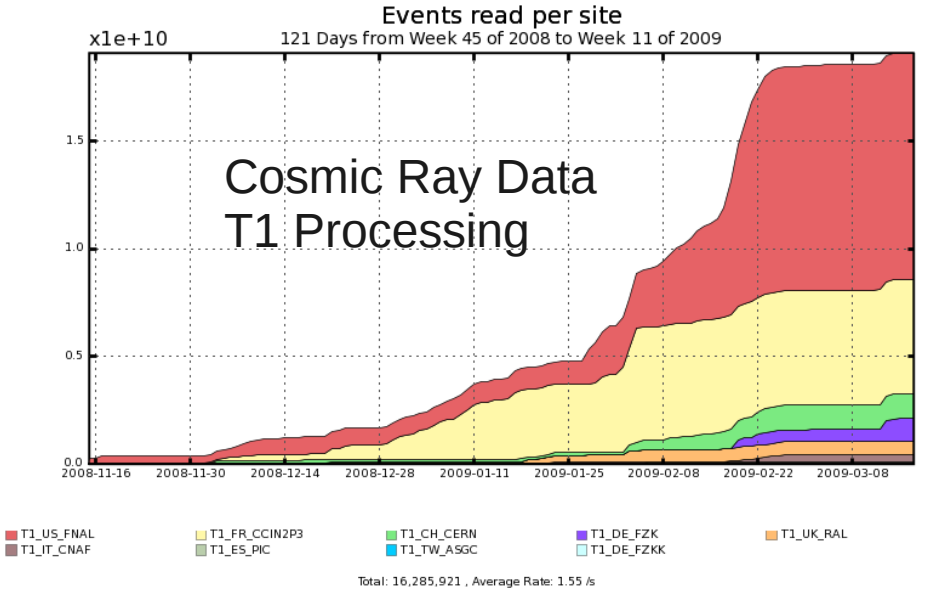


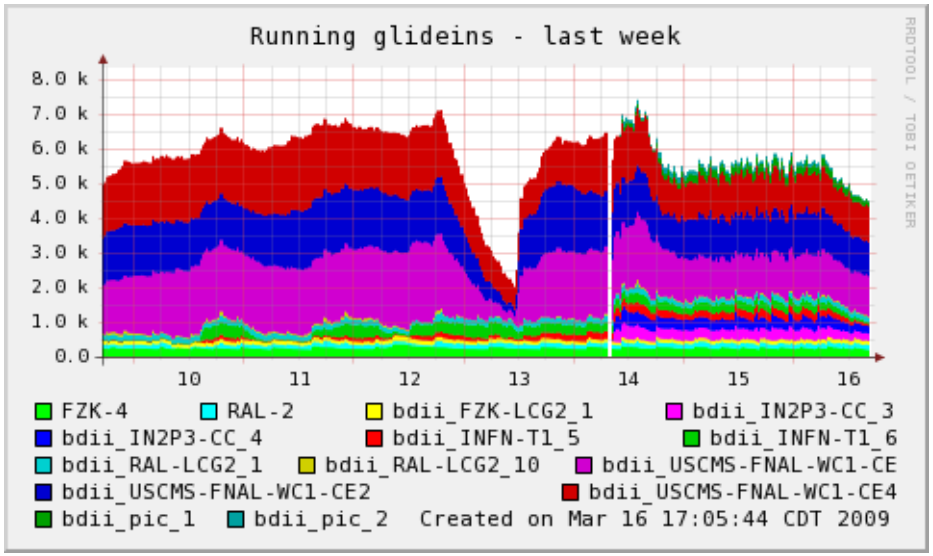


T0 – T1 Data Transfers in CRAFT

Subsequent Tier 1 Processing

- Data is shipped to T1's for immediate skimming, as well as archival storage and later rereconstruction.
- CMS cosmic ray data has been living at T1's for several months and has undergone several processing passes.
- As with the T0, T1 processing has been run successfully between the two central teams.
- Cosmic ray data reprocessing exercises now continuously used to maintain T1 site readiness





Monitoring CMS Data flow at the FNAL ROC

Summary

- CMS computing is based on a distributed, tiered model operated by centralized teams at CERN and FNAL.
- In such a system all data operations are remote operations
- The remote operation centers utilized by CMS help facilitate comunications between the remote teams and detector experts on opposite sides of the world.
- The FNAL ROC has been used successfully to manage CMS data processing workflows over extended periods from the Tier 0 through data transfer and processing at the Tier 1's.
- Remote operations centers will continue to play an important role in CMS data operations as we once again ramp up to acquire and process important collision data this fall!