Dress Rehearsal Preparations for CMS (CSA07)

Ian Fisk

September 2, 2007

What is CSA07

The preparation tests in CMS are call Computing Software and Analysis Challenges (CSA07)

- The goal is to exercise aspects of the computing model and the software development program with analysis activities and users
  - Dedicated tests of components do not show interference problems
- CSA07 is intended to exercise the computing model at greater than 50% of the target for 2008
  - The CSA06 challenge was an exercise at 25% of scale

We have a number of elements that have not been exercised previously

- Integration of the computing components up to storage manager
- Some data transfer channels: Tier-1 to Tier-1 Transfers, Tier-2 to Tier-1
- Balancing of simulation and analysis

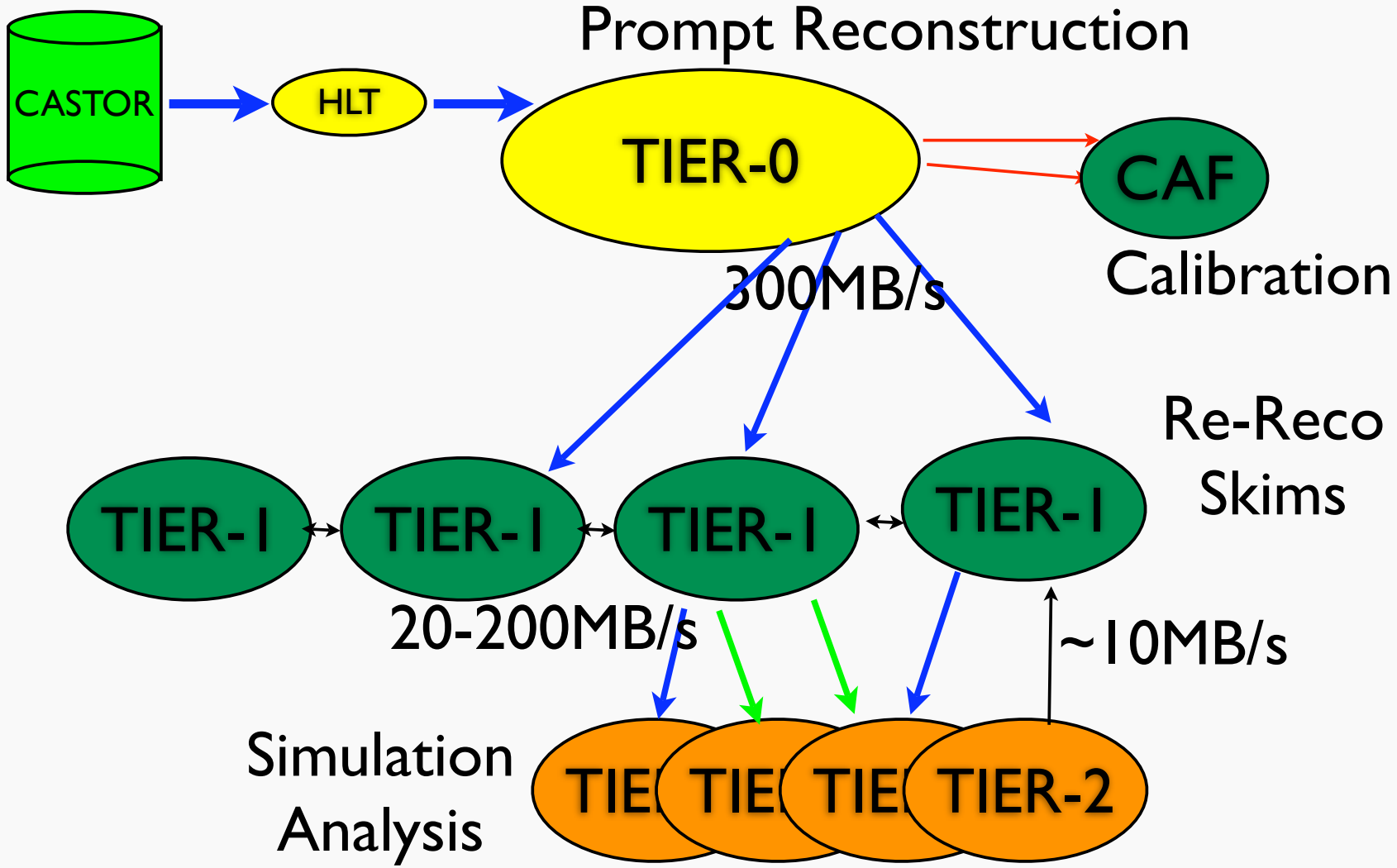
Desire to demonstrate computing and offline tools with a diverse and active user community

- Previous exercises have relied heavily on load generators

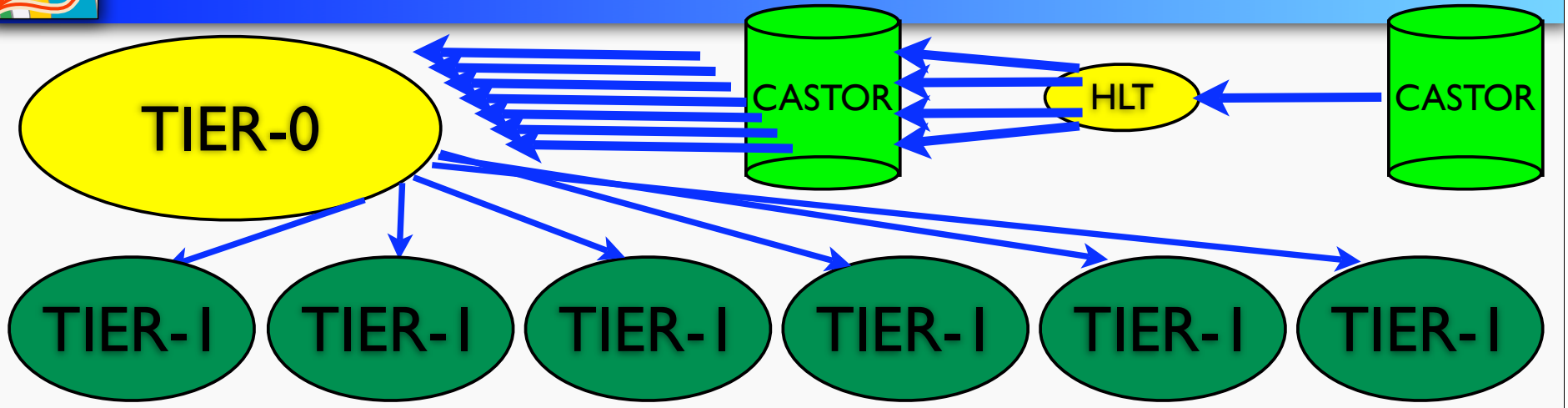
Basic Scaling Items Checked in CSA07

Service	2008 Goal	CSA07 Goal	CSA06 Goal	Status 2006
Tier-0 Reco Rate	150Hz - 300Hz	100Hz	50Hz	Achieved
Network Transfers between T0-T1	600MB/s	300MB/s	150MB/s	Achieved (6/7 cont.)
Network Transfers between T1-T2	50-500 MB/s	20-200 MB/s	10-100 MB/s	Achieved (15 sites)
Network Transfers T1-T1	100MB/s	50MB/s	NA	Not Attempted
Job Submission to Tier-1s	50k jobs/d	25k jobs/d	12k jobs/d	3k jobs/d
Job Submissions to Tier-2s	150k jobs/d	75k jobs/d	48k jobs/d	Achieved
MC Simulation	1.5 10^9 events/year	50M per month	NA	Not Attempted

CSA07 Workflows



Tier-0 Workflow



The Tier-0 workflow is similar to what was done for CSA06 with one large addition

- We have a pre-processing step to add the HLT information
  - Done before for technical issues (application limitations and memory footprint)
  - Keeps the Tier-0 workflow more realistic
- In CSA06 we had a minbias stream, a TTbar stream, a Jets stream based on simulation information
  - Now data from the Tier-0 looks more like experiment data based on the trigger selection

High Level Trigger Streams

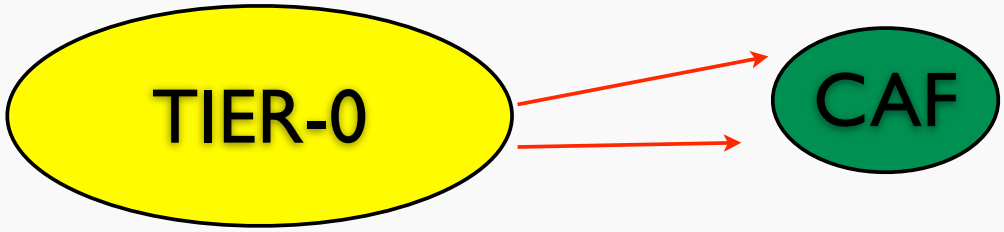
We are producing ~10 primary datasets

- constituents from about 80 simulated data samples mixed up
- Should be a more realistic analysis exercise

From a computing standpoint the creation of these samples has been interesting to watch

- Big variations on the constituents and the relative sizes
- Data hosting proposals are made and modified as these samples change
- Still work to do to come up with a final set

Calibration and Online



Mostly we will be using the CERN CAF resources for prompt calibrations

- Small data subsets are created and stored in a dedicated Castor pool
- Queue of batch resources through LSF are available for the calibration teams
- Currently no grid interface

This is ~100 CPUS, but between this and the Tier-0 tests we have consumed nearly all of the CMS share

Data From CERN

Once events are reconstructed they head for Tier-1 centers

- Total rate of reconstruction is expected to be 100Hz
  - Application is more complete than CSA06.
  - Large variations in the time of reconstruction. Memory requirements are higher than expected
- Expected aggregate rate is 300MB/s

Once data gets to Tier-1 centers it is the hot copy

- There are very few resources at CERN except for the Tier-0 and the small queue at the CAF
- Skimming and Re-Reconstruction will happen at Tier-1 centers
  - Reprocessing will involve creating the samples from the individual simulation constituents
    - TTbar, Jets, DrellYan, etc.

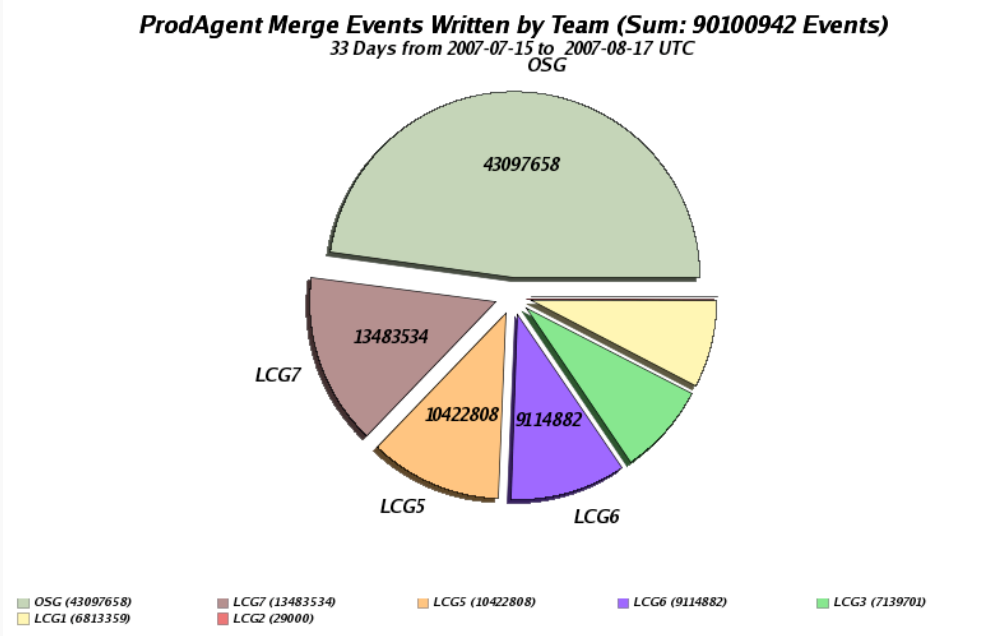
Status

Nearly 140M events have been simulated for the challenge

- A bit of transfers to CERN left to complete, but most of the data is successfully staged

Tier-0 testing in progress

- Up to 2000 job slots utilized
- Prod_Agent adopted for the Tier-0 workflow earlier in the year



- Running into interesting scaling limitations in Castor at CERN
  - Some of this was the CMS configuration

Status

We believe all the hardware at the sites is available to complete the challenge

- Need roughly 2500 batch slots at CERN
- Need functional tape resources at a nominal Tier-1s (600slots and 150TB storage)
  - Large variation in Tier-1 capacity
- Need credible disk mass storage at Tier-2s (150slots and 20TB)
  - Even larger variation here than in the Tier-2s
- Storage scales as time not as percentage of the final system

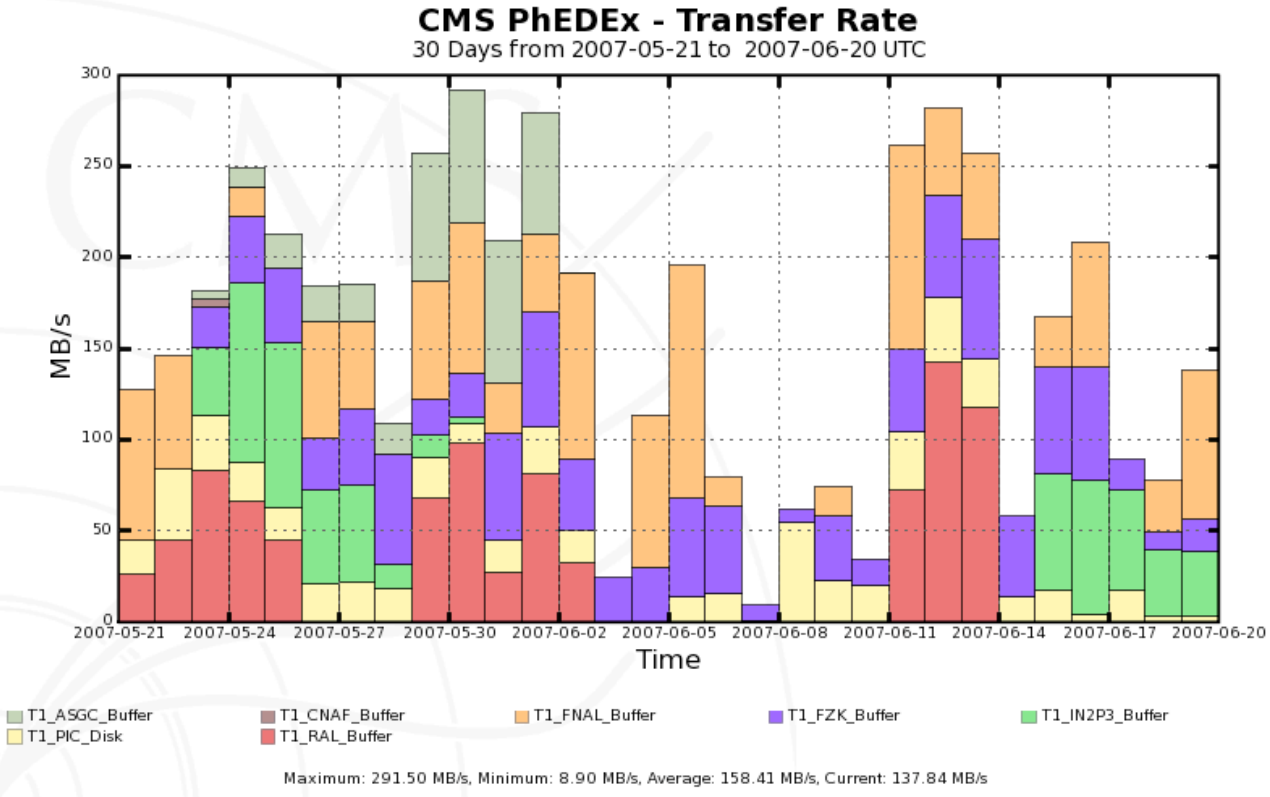
We expect a software release to start the pre-challenge steps on Monday

- This is roughly three weeks later than we expected
- Application memory targets are higher than expected
  - After some effort from the developers should still work

Status

CERN to Tier-1 transfers are an area that is reasonable well tested

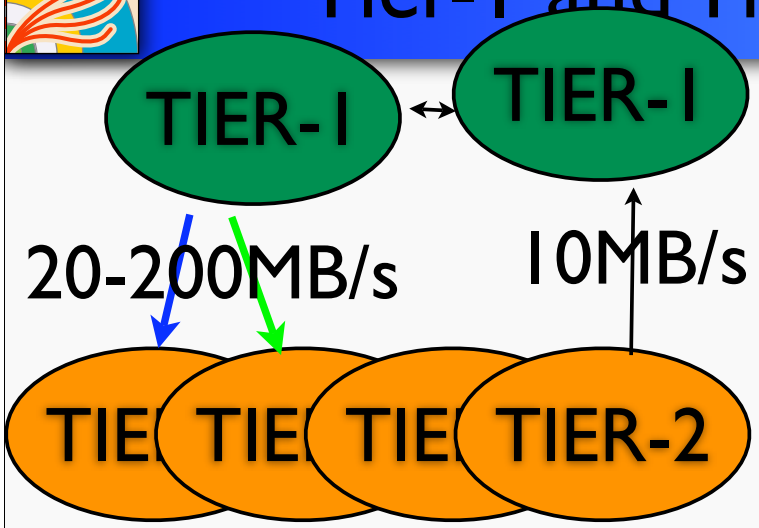
Site	Rate (MB/s)
ASGC	26
CNAF	37
FNAL	105
FZK	26
IN2P3	32
PIC	13
RAL	26



An open item is the ability to reliably write data to tape at all centers

- CSA07 also includes some elements specifically designed to exercise the ability to retrieve data from tape

Tier-1 and Tier-2 Transfers and Activities



Need the Tier-2 hardware to act as a dynamic cache

- Analysis activity at Tier-2s with data specified by user community
  - Low latency updates
  - Transfers need to work from arbitrary Tier-1 centers
    - Primary Datasets are hosted at particular Tier-1s.
    - Skims will be created at particular Tier-1 sites

Goal for Tier-2 analysis is 100k jobs per day

- Reached 50k jobs per day by the end of CSA06 but only with the heavy use of load generators

Tier-1 to Tier-1 Transfers

Tier-1 to Tier-1 transfers in the CMS Computing model are intended to synchronize a re-reconstructed copy of the AODs

- The more data a Tier-1 site hosts the less needs to be imported and the more needs to be exported. Metric is 3 consecutive days
- Goal is to hit 50MB/s on average in and out of a nominal Tier-1 center

Import

Export	ASGC	CNAF	FNAL	FZK	IN2P3	PIC	RAL
ASGC	0	4	4	4	4	4	4
CNAF	6	0	6	6	6	6	6
FNAL	18	18	0	18	18	18	18
FZK	4	4	4	0	4	4	4
IN2P3	6	6	6	6	0	6	6
PIC	2	2	2	2	2	0	2
RAL	4	4	4	4	4	4	0

Operations Change

During CSA06 Tier-2s were treated like mini-Tier-1 centers

- Subscriptions were handled centrally
- Data removal was manual and synchronization with the data management systems was not perfect

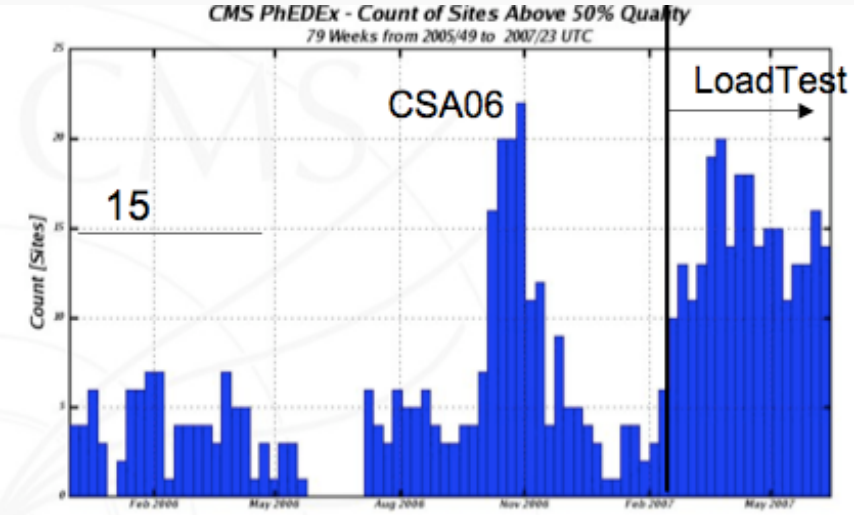
In the CMS computing model the Tier-2 data is intended to be dynamic and driven by the needs of the supported users and communities

- This year we have the concept of the site data manager
  - Approves transfer requests
  - Interacts with physics groups and users
  - Should be familiar with the size of the site and the needs of the users in order to make informed decisions about what data can be safely deleted
- Expect physics groups and users to interact with Tier-2 data manages

Tier-1 to Tier-2 Transfers

With the current division of raw data

- 40% FNAL
- 14% CNAF,
- 12% IN2P3
- 10% FZK, 10% ASGC, 10% RAL
- 5% PIC



[plots: courtesy of PhEDEx/B.Bockelman]

Assuming 30 Tier-2 sites participating in the challenge. The minimal rate from Tier-1 to Tier-2 should be

- FNAL should be export to 12 Tier-2s at an aggregate of 240MB/s, CNAF (4T2) 80 MB/s, IN2P3 (3-4T2) 60-70MB/s, and FZK, ASGC, and RAL (3T2) 60MB/s, PIC to 1 Tier-2s at 30MB/s

Maximum import rate should be much higher, so a Tier-1 site can satisfy the metric with a smaller number of Tier-2s

Processing in CSA

Aside from the Tier-0 prompt reconstruction, analysis and calibration object generation, there are four processing activities planned for CSA07

- Skimming
  - Selection of data at Tier-1 centers based on physics filters and transferred to Tier-2 centers for further analysis
- Analysis
  - Submission os analysis code to Tier-2 centers demonstrating local data access and the performance and the Tier-2 Storage Elements
  - Includes the contribution of the load generating job robot
- Re-Reconstruction
  - Processing of data at Tier-1 centers with access to new calibration constants through Frontier
- Simulation
  - Use the Tier-2 centers at 50% of the target rate of simulated event production

Ramping up Analysis Activities

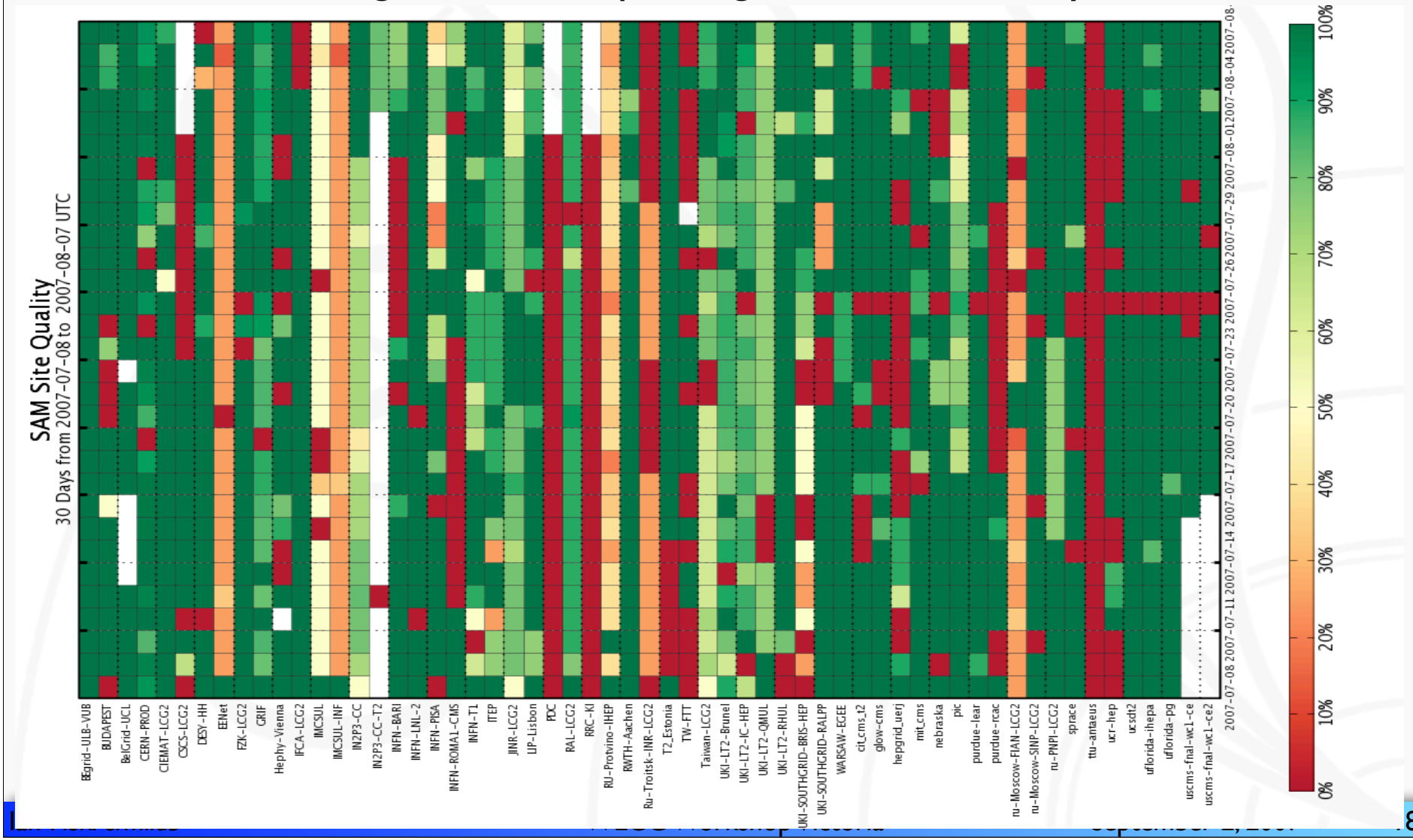
We would like to exercise the elements of the CMS computing model driven by user activities

- Skimming and data selection at Tier-1 centers
  - New sub-datasets are created based on physics selection code
- Active use of data caches at Tier-2 centers
- Submitting analysis jobs to Tier-2 centers
  - In the past much of the load has been generated by automatic load generators
- We would like to transition analysis users to Tier-2 centers
  - Currently a lot of activity is concentrated at a few sites
  - The CMS model calls for sharing the load across Tier-2 centers
    - We need the resources at CERN to complete the challenge
      - We estimate most of the resources available to CMS will be committed
    - We need many of the Tier-1 resources for reconstruction

Commissioning Activities

One goal this spring was to prepare the Tier-2 sites to host analysis datasets

- Commissioning has been improving the site availability with SAM



Schedule

We need to convert the simulated events to looking like events that came from the HLT farm

- This is divided into 3 steps and we expect this will take about three weeks
- Start hopefully Monday

Begin Tier-0 reconstruction activities on September 24

Simulation at the Tier-2s will continue from the beginning

About a week after the beginning we expect to start the skimming at the Tier-1 sites

- Data movement and analysis access

By two weeks we expect to begin reprocessing at Tier-1 sites

Outlook

Event production for the challenge is finished

- This went nicely. Largest number of events produced per month

Site commissioning and data channel commissioning is in progress

- The SAM tests are a good indicator for the eventual success of user jobs
- CMS computing model call for the mesh of Tier-1 to Tier-2 transfers to work efficiently. Testing in progress

Still testing a number of workflows

- Software releases and techniques are new

Previous challenges have validated elements of the CMS computing model

- Also gave directions for development and refinement
- It will be a busy fall.