



# Plans for CSA07

Ian Fisk  
CMS Meeting  
January 26, 2007



# What are CSAs

Computing Software and Analysis Challenges exercise the CMS Computing and Software systems at a defined scale and functionality

CSA06 was a 25% activity including many workflow elements

- ➔ Perform event reconstruction at the Tier-0 center on a mix of samples at  $\sim 40\text{Hz}$  for the period of a month
- ➔ Distributed the data to Tier-1 centers for archiving and data serving purposes. Average rate of  $150\text{MB/s}$ - $180\text{MB/s}$ 
  - Rate from Tier-0 to Tier-1 based on MOU pledges
    - Custodial archiving to tape where possible, or disk based archives for 30 days
- ➔ Serve the data to Tier-2 computing centers (Bursts of  $20\text{MB/s}$ - $100\text{MB/s}$ )
- ➔ Create selections of data storage at Tier-1 centers
  - Defined by physics groups
- ➔ Execute re-reconstruction jobs at Tier-1 centers
- ➔ Submit analysis jobs to Tier-2 centers
  - 30k-50k jobs per day integrated over centers



# Basic Scaling Items to Check in CSA07

Service	2008 Goal	2007 Goal	%
Tier-0 Reco Rate	150Hz - 300Hz	100Hz	50%
Network Transfers between T0-T1	600MB/s	300MB/s	50%
Network Transfers between T1-T2	50-500 MB/s	20-200 MB/s	40%
Network Transfers T1-T1	100MB/s	50MB/s	50%
Job Submission to Tier-1s	50k jobs/d	25k jobs/d	50%
Job Submissions to Tier-2s	150k jobs/d	75k jobs/d	50%
MC Simulation	1.5 $10^9$ events/year	50M per month	50%



# Work and Resources at CERN

## Tier-0 Reco Rate

- ➔ Negotiate with IT for ~2000 CPUs for the challenge
- Increase the completeness of the reconstruction algorithm
  - CSA06 Reconstruction rate was ~3s for minbias. Expected to be many times that when the experiment starts
- Majority of computing will simply re-reconstruct simulated events, but some fraction will hopefully reconstruct data from P5

## Network Infrastructure and data export capacity at CERN is probably OK

- ➔ Export capacity demonstrated at the end of CSA06 and service challenges indicates the rate can increase to CSA07 goals



# Work and Resources at Tier-1 and Tier-2

Tier-1 capacity should be ~600 CPUs at a nominal Tier-1

- ➔ Should not be an issue at most places
- ➔ Nominally 150TB of disk storage
  - Need to assess
  - This is much smaller than the computing model calls for even for a 50% challenge, but the disk storage is scaled down because it's only a 1-2 month challenge and now a full year
- ➔ Tape ingest rate at all Tier-1s needs to be verified
  - ~half the total capacity was actually demonstrated in CSA06
  - May be an issue

Tier-2 capacity would be ~150CPUs at a nominal Tier-2

- ➔ Disk storage 20-30TB
- ➔ This is for a nominal Tier-2 and there is a wide variation
  - Still we need sites that are credibly large



# Functionality Untested in CSA06 (1/5)

## Technical Items:

### Add Tier-1 to Tier-1 Transfers

- ➔ Tier-1 to Tier-1 transfers were tested opportunistically and only for short times
- Roughly half the permutations were tested at all
- ➔ CMS Computing Model calls for the ability for predictable bursts of data transfers between Tier-1 centers

### Incorporate all Tier-1 to Tier-2 transfers

- ➔ As daunting as it sounds all 200 permutations need to succeed
- Otherwise the Tier-2 will be limited for analysis

### Include Tier-2 to Tier-1 Transfers

- ➔ Rate is modest, but must demonstrate the Tier-1 sites can simultaneously receive data from CERN and Tier-2s, while sending data to Tier-2s

For this we require effort and time, but development is under control



# Tier-0 to Tier-1

Doubling expected import rates from CERN in CSA06

Site	Rate (MB/s)
ASGC	30
CNAF	50
FNAL	100
FZK	50
IN2P3	50
PIC	20
RAL	20

Goal is success transfers on 95% of the challenge days

➔ Target rate on 50% of the challenge days



# Tier-I to Tier-I Transfers

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

- ➔ The more data a Tier-I 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-I center

## Import

	ASGC	CNAF	FNAL	FZK	IN2P3	PIC	RAL
ASGC	0	5	5	5	5	5	5
CNAF	8	0	8	8	8	8	8
FNAL	17	17	0	17	17	17	17
FZK	8	8	8	0	8	8	8
IN2P3	8	8	8	8	0	8	8
PIC	3	3	3	3	3	0	3
RAL	3	3	3	3	3	3	0

Export





# Tier-1 to Tier-2 Transfers

With the current division of raw data

- ➔ 33% FNAL
- ➔ 16% CNAF, IN2P3, FZK
- ➔ 10% ASGC
- ➔ 7% RAL and PIC

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 10 Tier-2s at an aggregate of 200MB/s
- ➔ CNAF, IN2P3, and FZK to 5 Tier-2s at an aggregate of 100MB/s
- ➔ ASGC to 3 Tier-2s at an aggregate of 60MB/s
- ➔ RAL and PIC to 2Tier-2s at 40MB/s

Ideally would should be able to pick any random collection of Tier-2 permutations for the above

- ➔ Metric should be 8 hour burst on 50% of the challenge days



# Tier-2 to Tier-1 Transfers

Using the division of raw data from the previous slide

- ➔ Assuming a nominal Tier-2 produces MC needs to stage data to the Tier-1 center at 10MB/s
  - Modest but it adds up

Tier-1 import rates from Tier-2s

- ➔ FNAL 100MB/s
- ➔ CNAF, FZK, and IN2P3 50MB/s
- ➔ ASGC 30MB/s
- ➔ PIC and RAL 20MB/s

Goal should be 24 hour average rate on 25% of the challenge days



# Functionality Untested in CSA06 (2/5)

## Workflow Items:

Organized processing workflows (Re-reconstruction and skimming) were operated as proof of functionality and not a demonstration of operations

- ➔ Demonstration of Processing Request system
  - Demonstration of request system and prod\_manager components
  - Extension beyond Monte Carlo to skimming and re-processing production requests.
- ➔ Demonstrate that CMS can centrally trigger and track a re-processing step at the Tier-1s
  - Ensuring the processing of every event

Implies development and integration on Production suite of tools



# Functionality Untested in CSA06 (3/5)

## Workflow Items:

### Demonstrate Tier-2 management of local data

- ➔ CSA06 was a reasonable representation of the CERN to Tier-1 transfer workload
- ➔ Level of local control for Tier-2 data was not a good reproduction of the computing model concepts
  - Central operator support was needed
  - latency was high
  - A number of steps were manual

Requires the deployment of PhEDEx 2.5.1 and training of Tier-2 sites



# Functionality Untested in CSA06 (4/5)

## Work Flow:

### Increasing the people submitting jobs

- ➔ CSA06 was basically done with 20 people
  - Most of the submission was robots
- ➔ Even if we say only 10% of the collaboration works (hopefully wrong) we need 100 people for a 50% challenge exercise

Requires some training, but most importantly we need schedule the challenge activities to coincide with something the users are doing anyway



# Functionality Untested in CSA06 (5/5)

## Workflow:

Some Dataset Bookkeeping functionality was missing

- ➔ No connection with other DBs
  - Conditions, luminosity, data quality, etc.
- ➔ User data queries and discovery are limited

Relies on availability and integration of DBS2



# Challenge Schedule

## Items that drive the schedule

- ➔ The desire to overlap with the ATLAS “Dress Rehearsal”
  - Wide window, not clear how well fixed the start is
- ➔ When Tier-I sites can reasonably write to tape
  - The site capacity for processing and probably storage is OK, but it doesn't make sense to do the challenge without tape everywhere
- ➔ Availability of CMS components
  - Prod Components
  - DBS2
  - PhEDEx 2.5.x
- ➔ Desire not to run the challenge during entirely during August



# Schedule

CMS benefitted from Service Challenge 4 preparations to be ready for the challenge

- ➔ There is no service challenge, but a preparatory period would be helpful

Working on the schedule, but one possible scenario would be a rolling start during June

- ➔ Exercise the workflow without a formal measuring of metrics

Formal Challenge for 30 days in July

- ➔ Relies on the ability to generate 100M events with CMSSW 1.3 between March and July
- ➔ Availability of site resources and CMS services
- ➔ Progress on a lot of integration and commissioning activities

Subject to change with more careful assessment