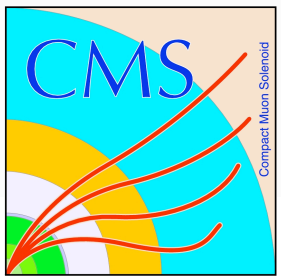
CMS Operations Report

January 12, 2010

Ian Fisk

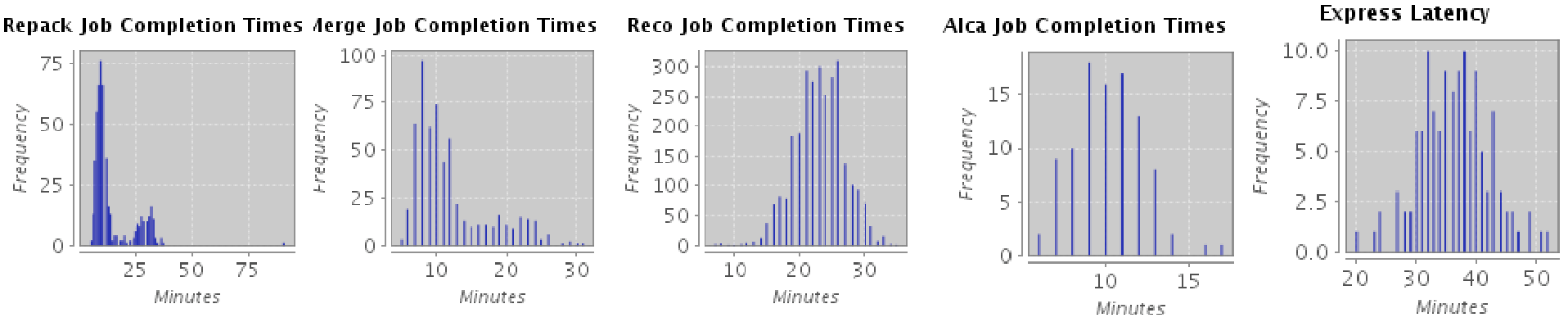


Introduction

- The CMS Distributed Computing System generally performed well with the addition of collision data
  - The data rates and sample sizes are still quite low
  - The system was not resource constrained during this early period
  - The workflows and activities were generally what was expected from the computing model
    - Workflows executed much more frequently
      - Data Multiply Subscribed (More T1 and T2 subscriptions)
      - Re-processing occurred every 2-3 days
- Data Reconstruction, Skimming, Re-reconstruction at Tier-1s went nicely in parallel with distributed user analysis and MC production at Tier-2s

Data Collection Infrastructure

- Tier-0, Tier-1 Re-reco and Data Distribution Systems functioned with early collisions
  - Events were reconstructed and exported to Tier-1 sites
  - Express stream latency at target levels
  - Re-reconstructed using Tier-1 centers
  - Prompt Skimming system moved into production



Readiness of T0 reconstruction

- Tier-0 Facility had been routinely exercised with cosmic data taking and simulated event samples
  - Performing Stably with Cosmics

Job Type	Total Jobs	Failures	Success Rate
Express	342186	31	99.99%
Repack	134730	2	100.00%
PromptReco	38911	18	99.95%
AlcaSkim	41659	3	99.99%

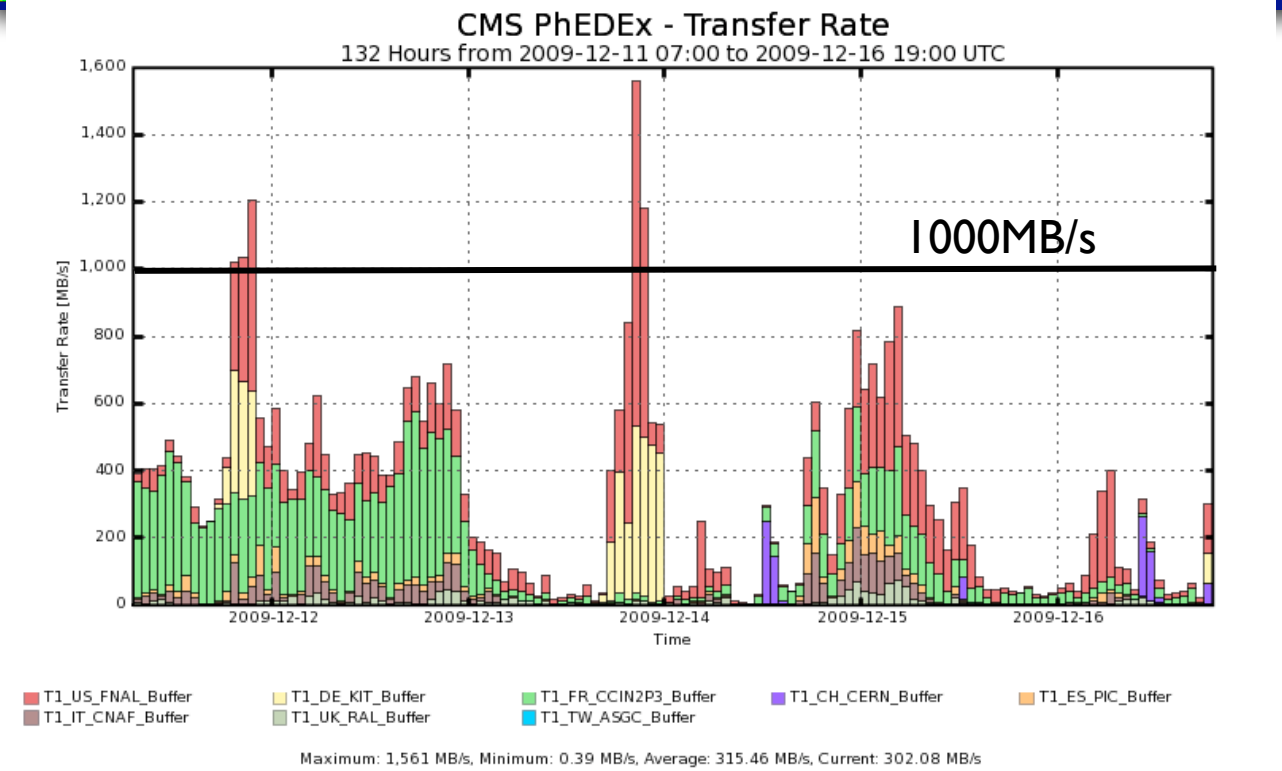
- With Collisions (Failures concentrated in setup)

Job Type	Total Jobs	Failures	Success Rate
Express	404546	9442	97.72%
Repack	86982	69	99.92%
PromptReco	209773	2875	98.64%
AlcaSkim	17631	431	97.61%

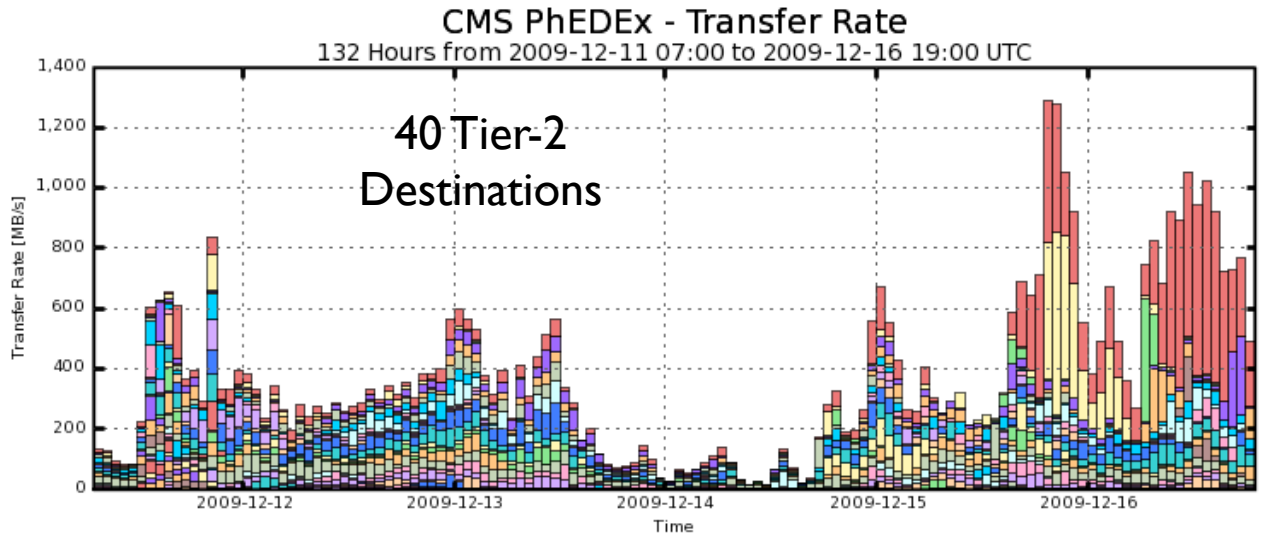
4 9's

- ~3000 cores
  - Local submission to farm with multiple workflows
  - Good stability and performance of CMS software
- Received confirmation from CERN on T0+CAF pledge in 2010

Distribution, Processing, Access



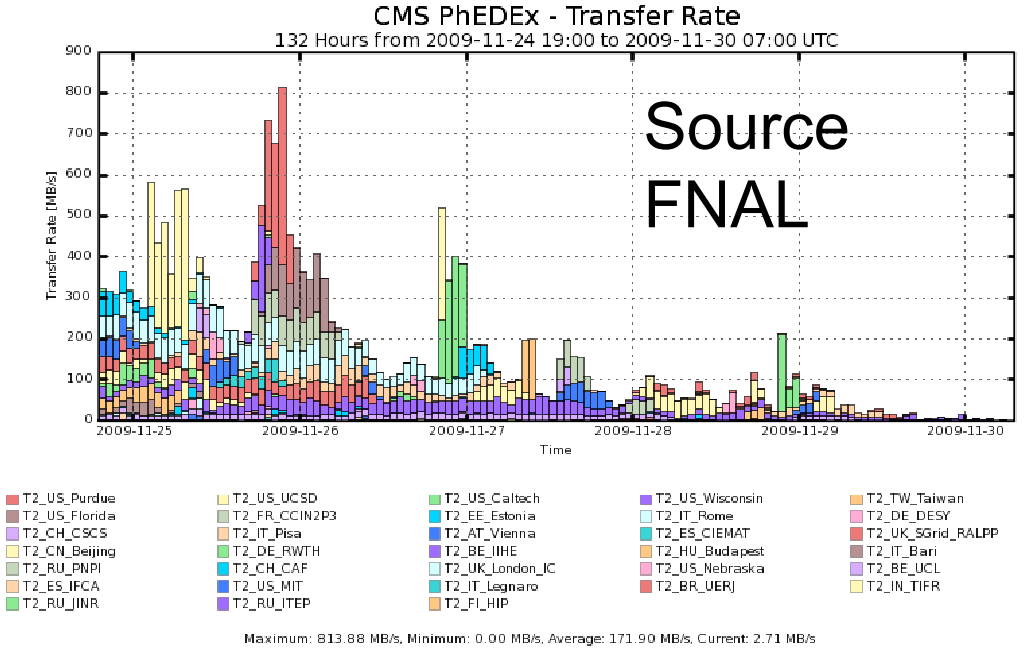
Source CERN or Tier-1 going to destination Tier-1

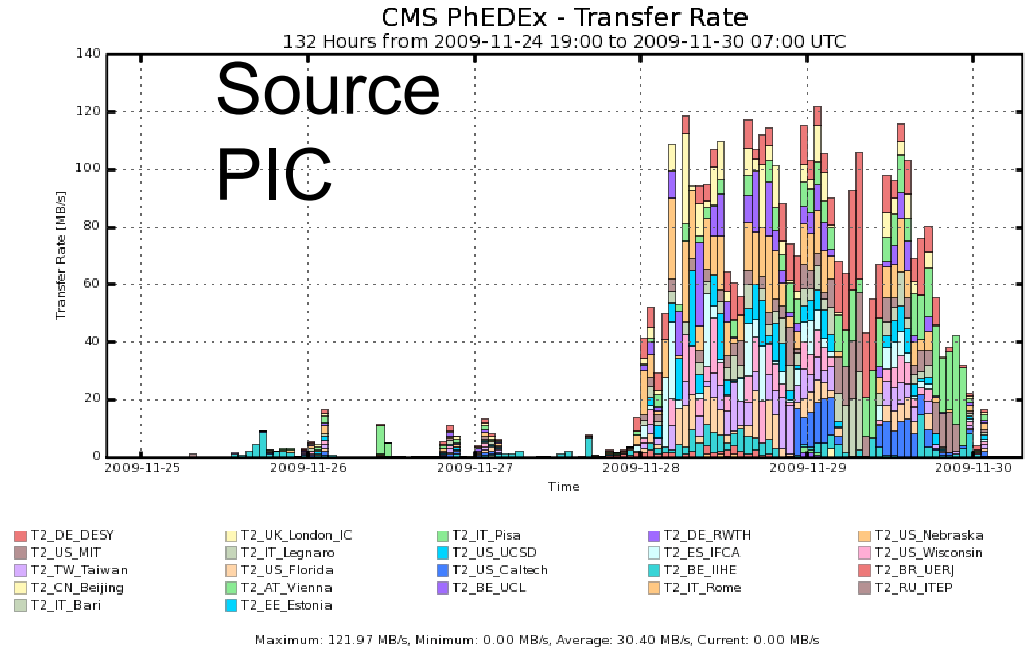


Source Tier-1 going to destination Tier-2

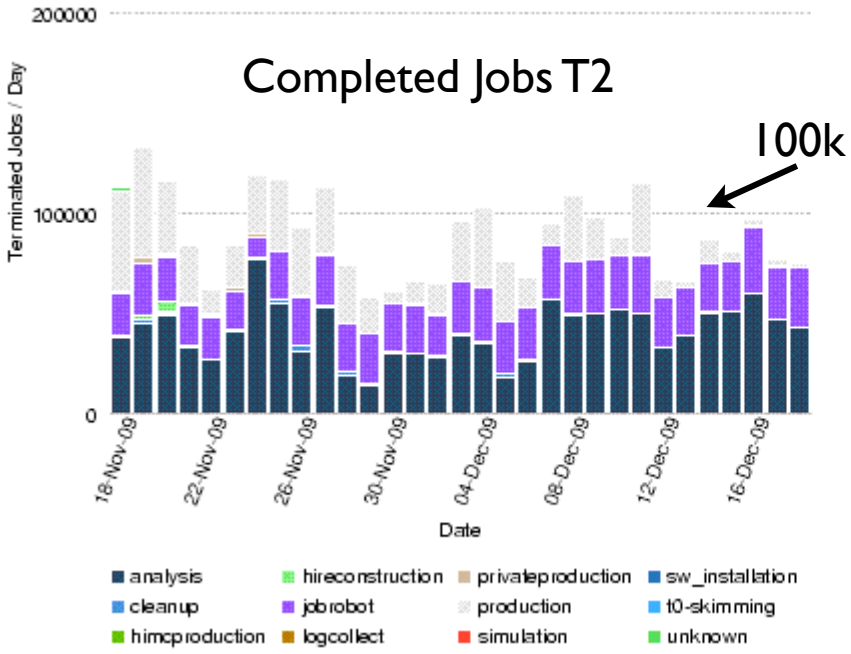
Load Balancing

- We subscribed the MinBias primary dataset to PIC between the 27 to 28th of November
  - Transfer system balanced the load to destination Tier-2s
  - Good performance from both sites



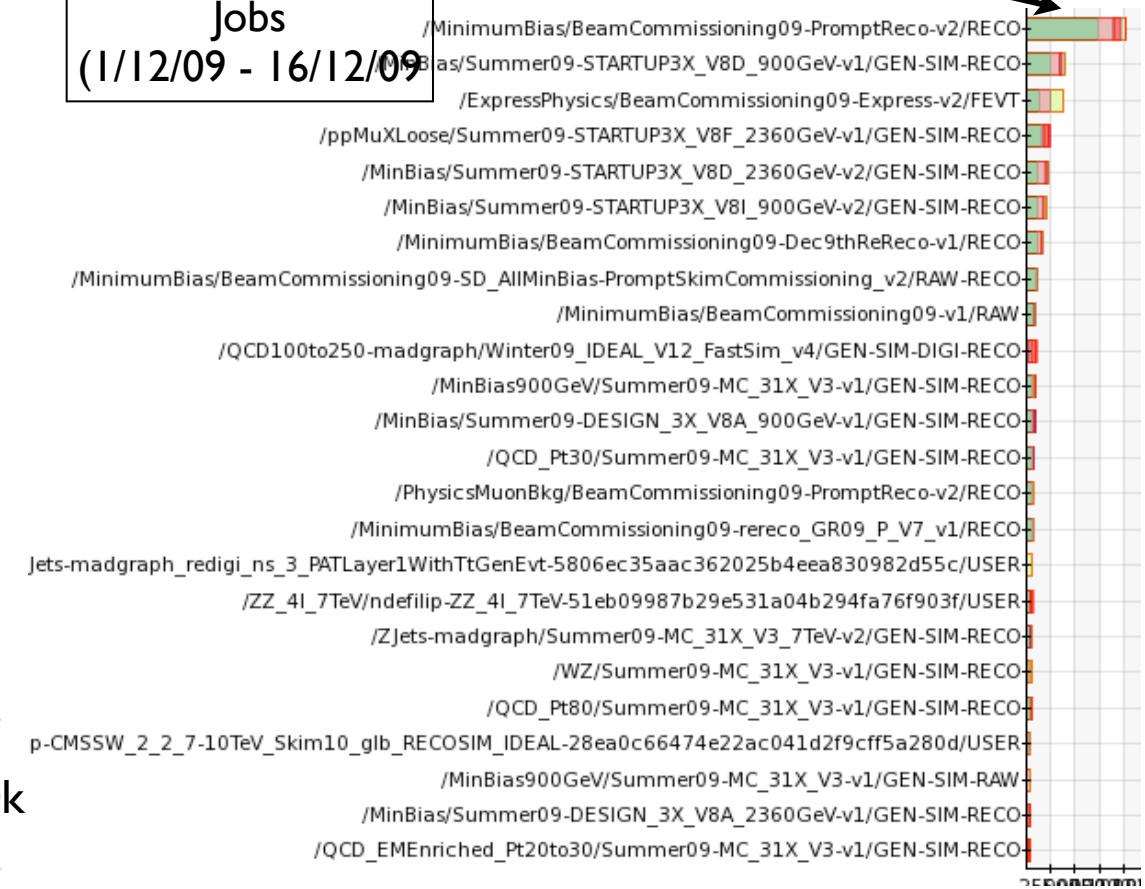


Access



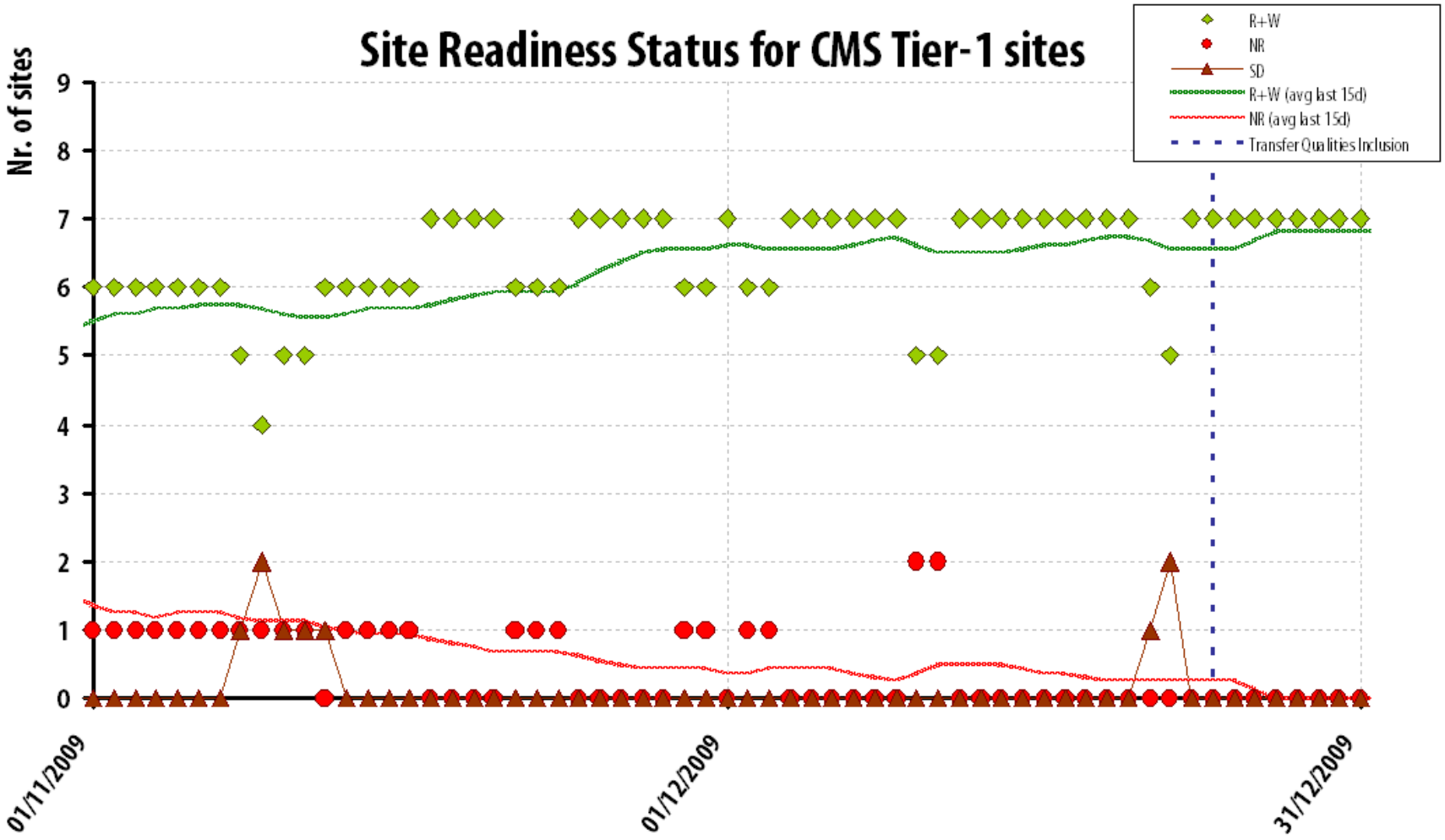
Dec Analysis Jobs (1/12/09 - 16/12/09

75% success

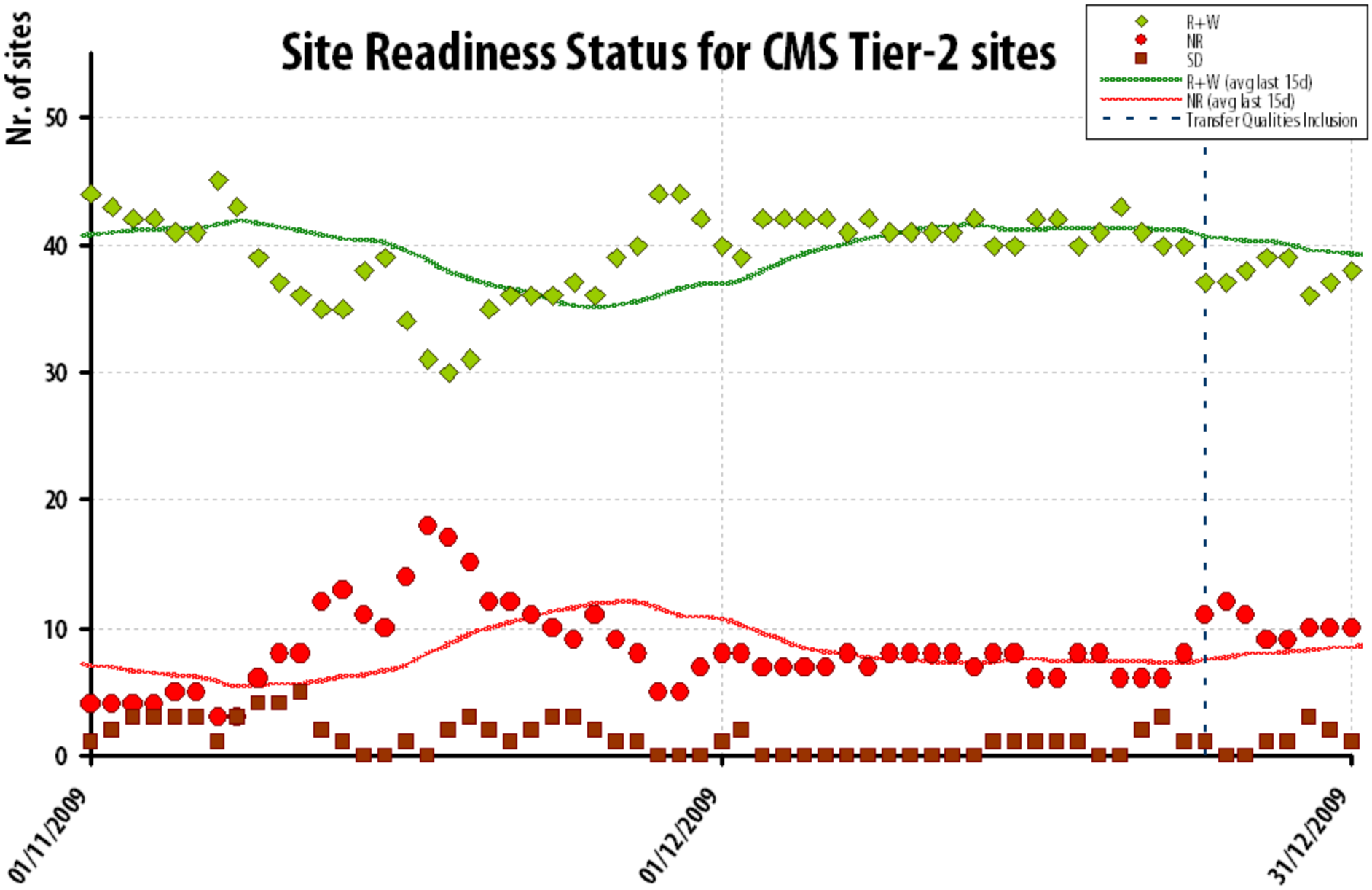


Site Stability

- Tier-1 Readiness November and December
  - Readiness defined as passing the CMS, SAM, Job Robot, and Transfer tests for a high percentage of a time window



Tier-2 Stability

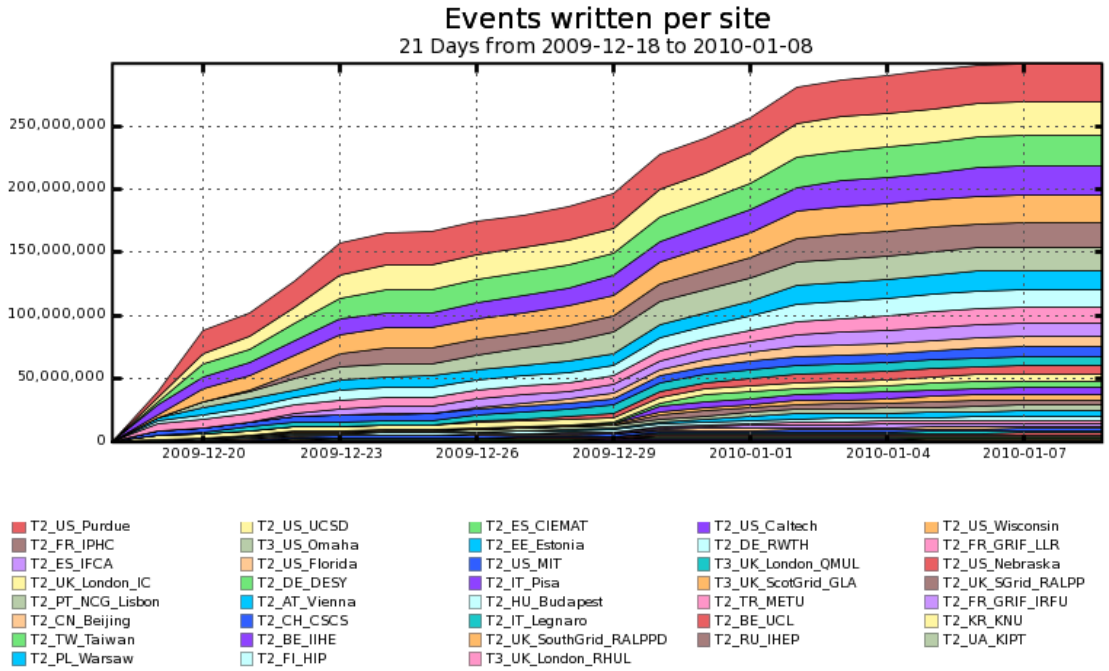


Activities over the break

- Data Processing Activities during the break
  - Re-processing and skimming of all good runs finished on 12/24 for the two large physics datasets
    - ZeroBias 22M RAW events, 1019 files processed
      - 11TB produced, 112M events in Secondary Datasets, AlcaReco etc
    - MinimumBias RAW 21.5M events, 1207 files processed
      - 10TB produced, 74M events in Secondary Datasets, AlcaReco etc distributed
      - Processed for two software releases (on SL5 and SL4)
        - Re-processing of MC datasets finished on 12/25
          - 20M MinimumBias
      - Re-processing of Cosmics MC finished on 12/25
        - 130M events
    - Almost problem-free processing of high-quality data
    - e.g. for the latest CMSSW version only one of >2000 job failed due to memory consumption all was done within 4-5 days

MC Production

- Smooth MC Production over break

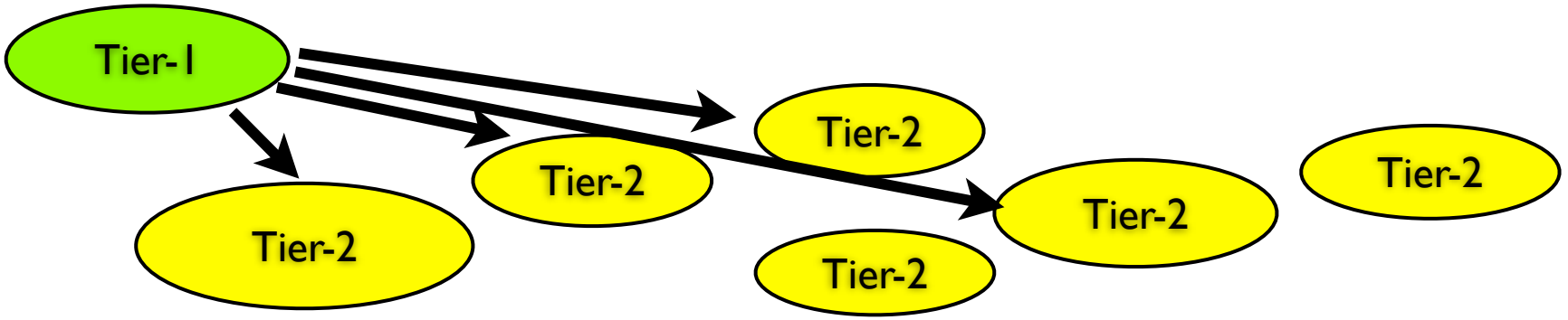


- ★ some 120M events produced (RAW, RECO, AOD)
  - ✦ including special MinBias samples for comparison with 900GeV and 2.36TeV data
  - ✦ most FullSim, some FastSim

Areas to Work On

- We were in an environment without resource constraints
  - Data rate and complexity is lower than expected in the final system
  - Allows many more passes and caused some complaints about lack of utilization
  - Number of users is also lower

- While we see the ability replicate data to Tier-2s. We are taking advantage of the oversubscription
  - Need to anticipate achieving good performance Tier-1 to Tier-2 when the data is accessible from fewer places



Improving Network

- The CMS Computing TDR defines the burst rate Tier-1 to Tier-2 as 50MB/s for slower links up to 500MB/s for the best connected sites
  - We have seen a full spectrum of achieved transfer rates
    - Average Observed Daily Max peaks at the lower end
- From the size of the facilities and the amount of data hosted, CMS has planning estimates for how much export bandwidth should be achievable at a particular Tier-1
  - No Tier-1 has been observed to hit the planning numbers (though a couple have approached it)
  - CMS would like to organize a concerted effort to exercise the export capability
    - Need to work with site reps, CMS experts, FTS and Network experts
      - Area for collaboration

Outlook

- Distributed Computing Worked well during the Opening collision data for CMS
  - Thanks to CERN and the Tier-1 sites for keeping things working
  - Some items to follow up on
- Not yet working at the rates anticipated in the planning
  - Interesting work for 2010