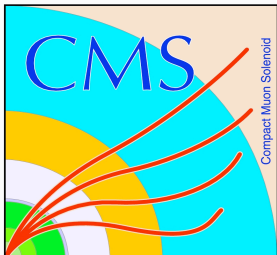


# CMS Computing Readiness

11/13/09

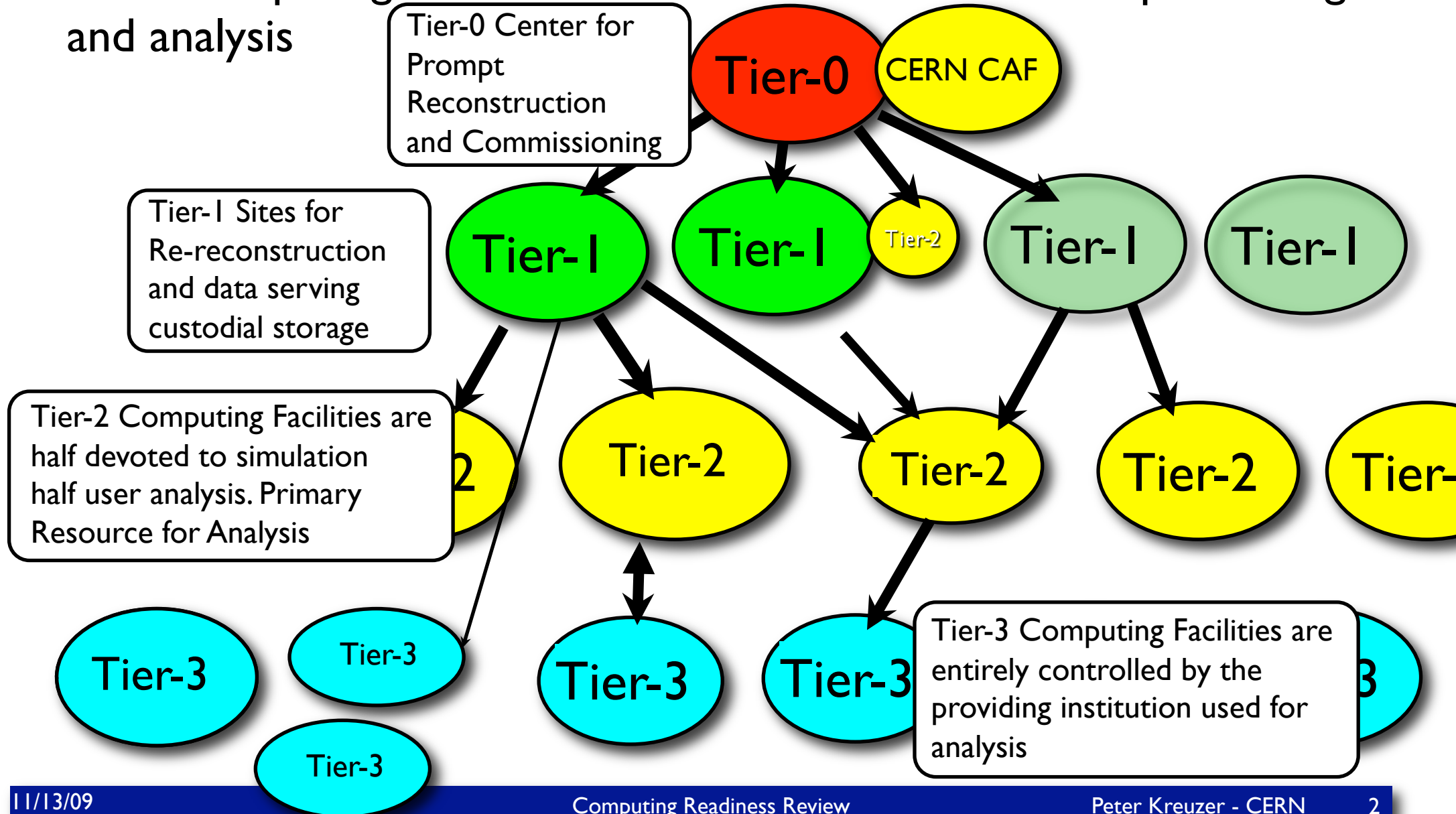
Peter Kreuzer (RWTH Aachen / CERN)  
For the CMS Computing Project





# Computing Model in CMS

- CMS Computing makes use of more than 60 sites for processing and analysis





# Readiness of T0 reconstruction

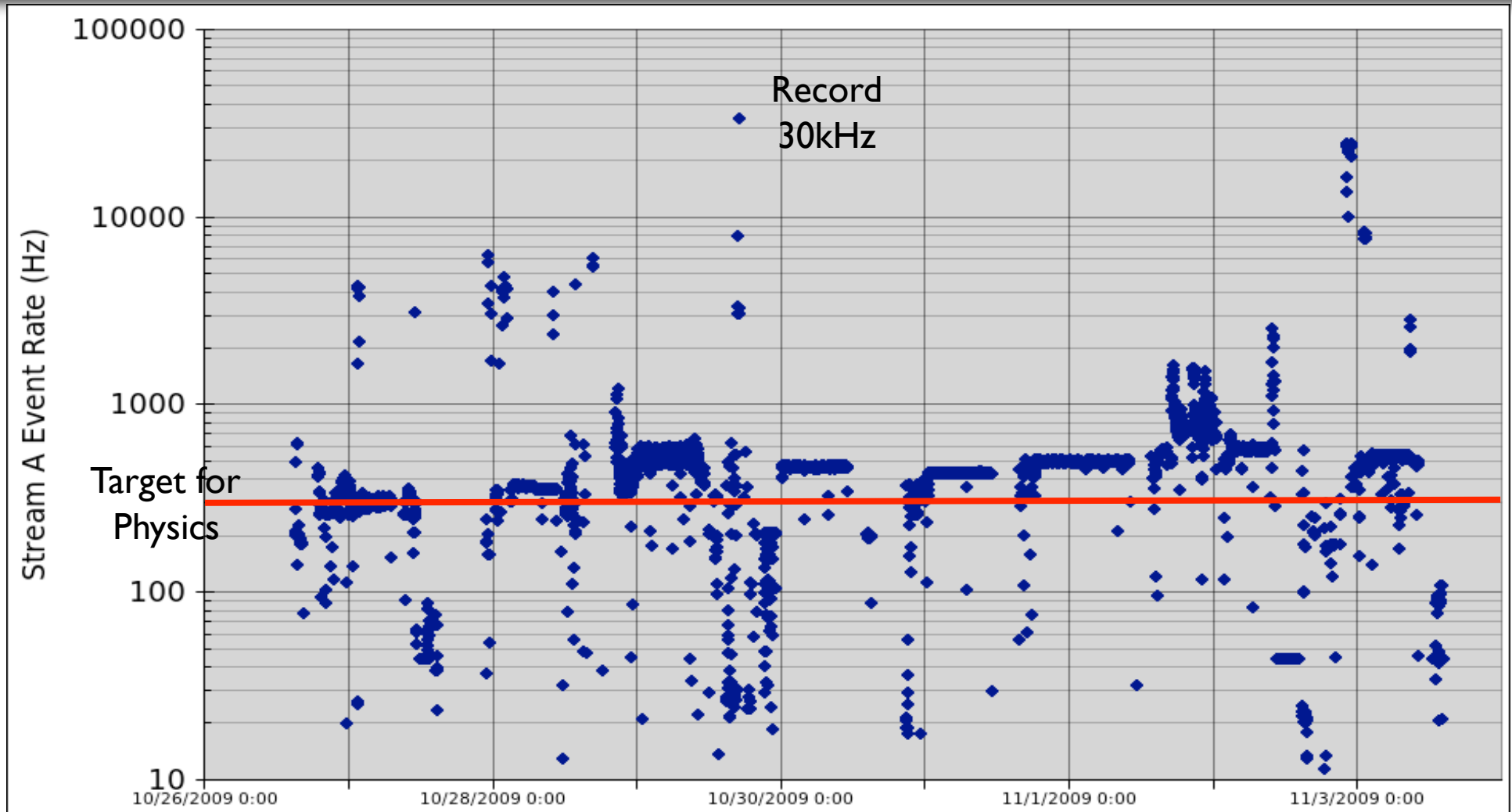
- ▶ Tier-0 Facility has been routinely exercised with cosmic data taking and simulated event samples

- ▶ Performing well

Job Type	Total Jobs	Failures	Success Rate
Express	342186	31	99.9909%
Repack	134730	2	99.9985%
PromptReco	38911	18	99.9537%
AlcaSkim	41659	3	99.9928%

- ▶ ~3000 cores
- ▶ Local submission to farm with multiple workflows
- ▶ Good stability and performance of CMS software

# Readiness of T0 reconstruction



- ▶ Large dynamic range of data into Tier-0
- ▶ Working to keep sustained operations



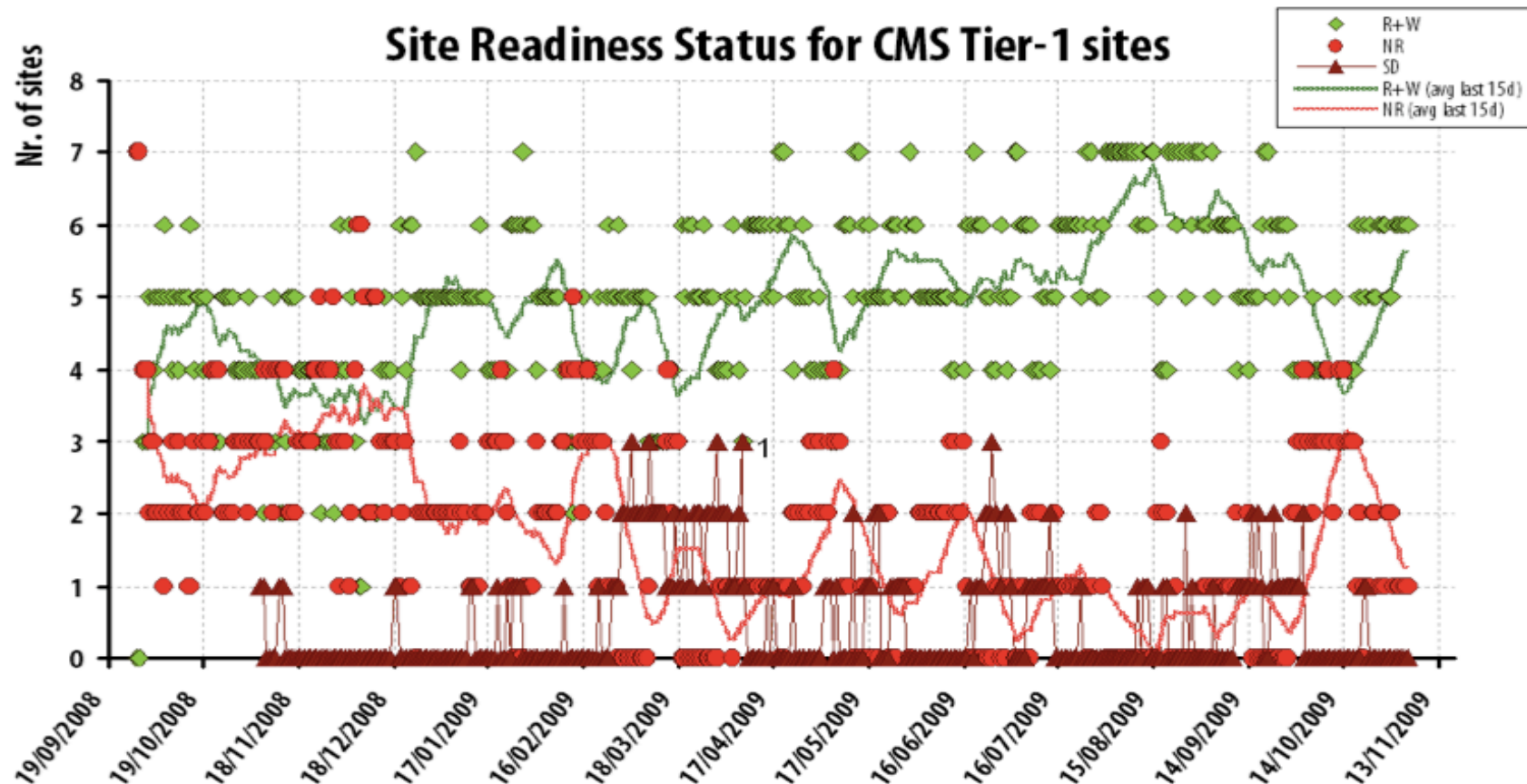
# CAF Readiness

- ▶ CERN Analysis Facility (CAF) is used for commissioning and low latency work
- ▶ CPUs on the CAF are heavily used
  - ▶ CERN pledge for T0+CAF will soon increase from 29.6 to 44 kHSPEC06  $\approx \times 1.5$ . Some will go into much needed increase in commissioning resources
  - ▶ Currently using multiple batch queues for prioritization. OK for startup
  - ▶ Need to provide more monitoring by group for a better overview of CPU fair share and ensure equitable usage
- ▶ Data management
  - ▶ CAF Production and User disks well separated
  - ▶ Disk space monitoring will be done by groups
  - ▶ Need automated clean-out of older samples



# Tier-1 Availability

- ▶ CMS Stores Data Custodially at Tier-1 Centers
- ▶ Reliability has improved, but we still see issues
  - ▶ Fall site visits were productive and action items being followed
  - ▶ During First Running non-custodial samples will ensure availability





# Readiness of data distribution to T I for custodial copy

- ▶ In order to maximize availability CMS will be storing data on both sides of the Atlantic
  - ▶ We can do this in the opening of the experiment
  - ▶ Good use of distributed Tier-I centers

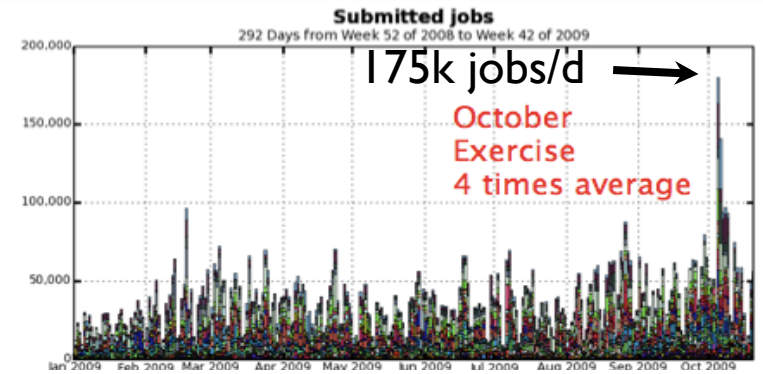
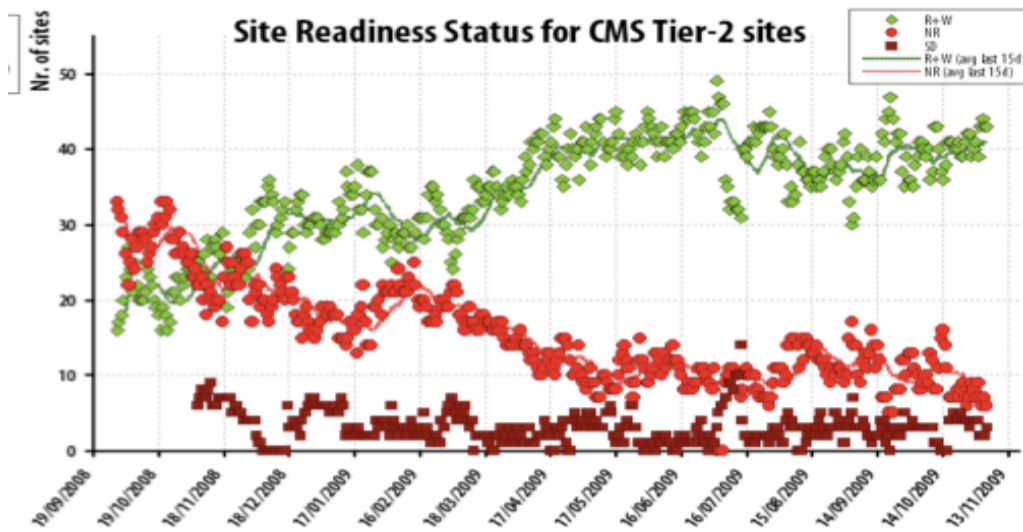
Site	CPU [slots]	Tot. Tape [TB]	Custodial [TB]	Non-Custodial [TB]	Currently Free [TB]	Currently Available (Tot.-Custodial) [TB]	Assigned PD	Available Space after tot. Storage [TB]
ASGC	1600	800	167	105	528	633	MinBias	427
CNAF	750	804	677	162	-35	427	ZeroBias	-136 *
IN2P3	1100	1309	608	50	651	701	BeamHalo,MinBias	530
FNAL	8000	7100	1370	1400	4330	5730	MinBias,ZeroBias, BeamHalo,Cosmics,Calo,FEDMonitor,HcalHPDNoise, MinimumBiasNoCalo, RandomTriggers, RPCMonitor, HcalNZS, TestEnables	3839
KIT	700	2000	472	1	1527	1528	Calo,HcalHPDNoise, HcalNZS,TestEnables	1396
PIC	560	974	297	217	460	677	FEDMonitor, RandomTriggers, RPCMonitor,MinimumBiasNoCalo	419
RAL	1000	1887	574	76	1237	1313	Cosmics	1136

\* Plan depends on additional media



# Readiness for Analysis

- ▶ During October CMS conducted a large scale analysis exercise
  - ▶ Have large experience and demonstrated success in users using CRAB (CMS Analysis submission tool) to process MC data,
  - ▶ Rate of jobs is similar to expectations in the Computing Model
  - ▶ Some scaling issues of workflow architecture specifically in staging out user files.



Introduced CRAB client that checks against common user errors before job submission.



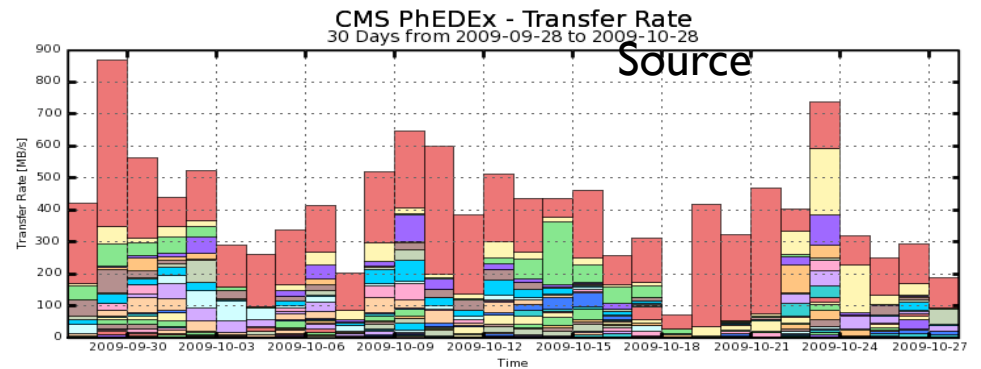
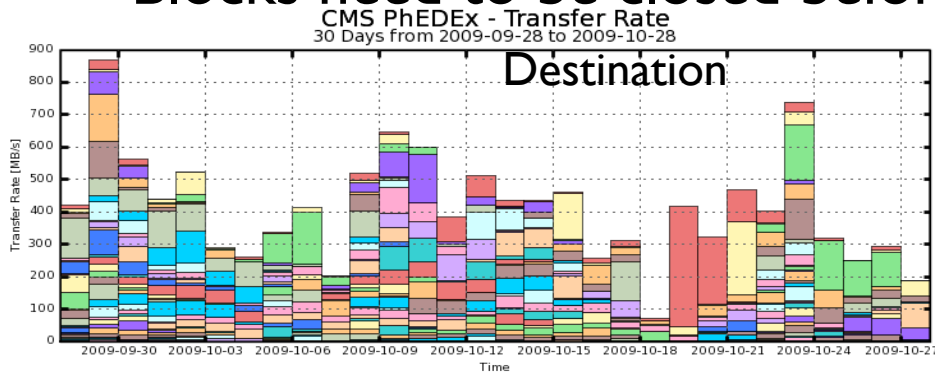




# Readiness for T2 data distribution

Samples successfully replicated in Oct Ex.

- ▶ 45 Tier-2s Received Data, 30 Tier-2s were a source of data
- ▶ Rates observed total out of Tier-1s are not as high as we expect from the planning.
- ▶ Working on dedicated load tests
- ▶ Latency for successfully closing a block of data varies widely by destination and is generally high. Need work with developers
- ▶ Blocks need to be closed before data is accessible



F2_US_Purdue	T2_CH_CAF	T2_FR_CCIN2P3	T2_US_Nebraska	T2_US_Wisconsin
F2_US_Caltech	T2_US_UCSD	T2_ES_IFCA	T2_DE_DESY	T2_US_MIT
F2_IT_Legnano	T2_US_Florida	T2_EE_Estonia	T2_DE_RWTH	T2_UK_London_IC
F2_FI_HIP	T2_FR_IPHC	T2_BR_URJ	T2_KR_KNU	T2_UK_SGrid_RALPP
F2_FR_GRIF_IRFU	T2_CH_CSCS	T2_ES_CIEMAT	T2_RU_JINR	T2_RU_SINP
F2_PT_LIP_Lisbon	T2_HU_Budapest	T2_AT_Vienna	T2_TR_METU	T2_IT_Rome
F2_IT_Bari	T2_UK_London_Brunel	T2_IT_Pisa	T2_BE_UCL	T2_UK_London_IC
F2_IN_TIFR	T2_TW_Taiwan	T2_CN_Beijing	T2_BE_IHE	T2_ES_IFCA
F2_RU_ITEP	T2_PT_RRC_KI	T2_PT_NCG_Lisbon	T2_UA_KIPT	T2_ES_CIEMAT

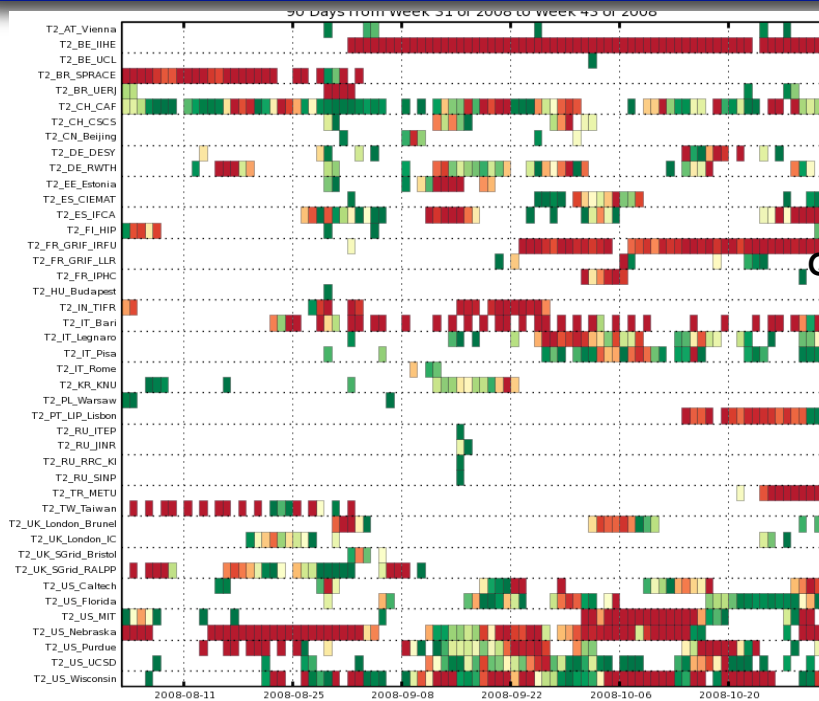
Maximum: 867.81 MB/s, Minimum: 71.25 MB/s, Average: 411.47 MB/s, Current: 187.30 MB/s

T1_US_FNAL_Buffer	T1_IT_CNAF_Buffer	T1_UK_RAL_Buffer	T2_US_UCSD	T3_US_FNALLPC
T1_DE_FZK_Buffer	T0_CH_CERN_Export	T1_ES_PIC_Buffer	T1_FR_CCIN2P3_Buffer	T2_US_Caltech
T2_US_Wisconsin	T1_TW_ASGC_Buffer	T2_DE_RWTH	T2_IT_Legnano	T2_DE_DESY
T2_US_MIT	T2_US_Nebraska	T2_FR_CCIN2P3	T2_IT_Pisa	T2_US_Purdue
T2_CN_Beijing	T2_CN_KIT_Buffer	T2_AT_Vienna	T2_UK_London_Brunel	T2_UK_SGrid_RALPP
T1_DE_KIT_Buffer	T2_ES_IFCA	T2_AT_Vienna	T1_CH_CERN_Buffer	T2_FI_HIP
T2_FR_IPHC	T2_CH_PSI	T2_BE_UCL	T2_US_Florida	T3_US_Omaha
T2_CH_CSCS	T2_US_UMD	T2_FR_GRIF_LLJ	T2_FR_GRIF_IRFU	T2_FR_GRIF_IRFU
T2_RU_JINR		T2_US_Colorado	T2_PT_LIP_Lisbon	T2_PT_NCG_Lisbon

Maximum: 867.81 MB/s, Minimum: 71.25 MB/s, Average: 411.47 MB/s, Current: 187.30 MB/s



# Readiness for T2 Data Distribution



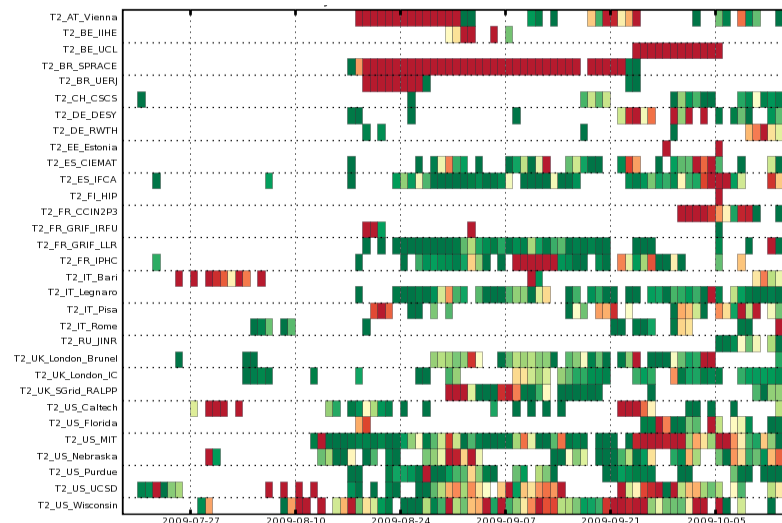
Considerable effort in commissioning transfers:  
Transfer Quality  
Tier-1 to Tier-2  
3rd quarter

2008

2009



Tier-2 to Tier-2 Transfers:  
Many permutations improving  
quality





## Readiness for creation of secondary datasets and skimming

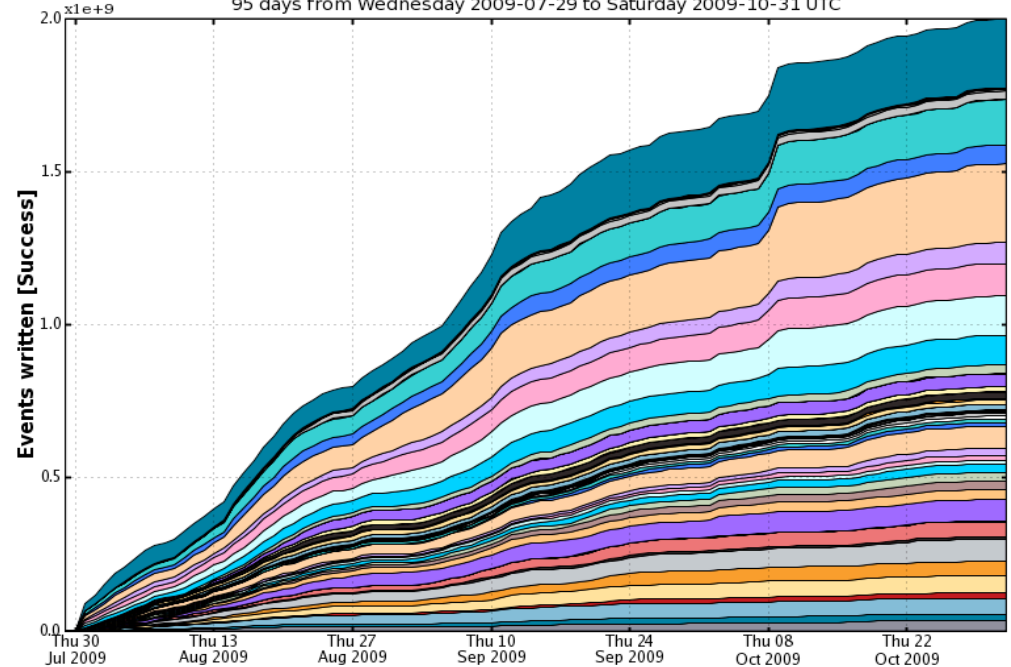
- ▶ During October Exercise we exercised the Tier-I samples to produced skimmed samples
- ▶ 41 workflows: 100% complete, 7 workflows: >99.3% complete
- ▶ Exercised Skimming
  - ▶ Samples Skimming From Stored Data
    - ▶ Based on tool used for producing simulated event
      - ▶ Operationally intensive
  - ▶ Samples Produced with the prompt skimming system
    - ▶ Based on system used in the Tier-0
- ▶ Preparing configurations to produce analysis samples for physics



# Readiness for MC production

- ▶ Simulated Event production using distributed computing infrastructure has been extremely successful in 2009
- ▶ 2B successfully written events
- ▶ Good contribution from more than 50 sites
- ▶ Moving to more advanced simulated events with multiple interactions

Events written [Success] (Processing jobs) sorted by Site matching T2|T3  
 95 days from Wednesday 2009-07-29 to Saturday 2009-10-31 UTC



T2_AT_Vienna [33.1M]	T2_IT_Pisa [70.9M]	T2_UK_SGrid_RALPP [2.47k]
T2_BE_IJHE [20.4M]	T2_IT_Rome [12.7M]	T2_UK_SouthGrid_Bristol [223k]
T2_BE_UCL [52.0M]	T2_KR_KNU [11.8M]	T2_UK_SouthGrid_RALPPD [28.5M]
T2_CH_CSCS [18.1M]	T2_PL_Warsaw [9.85M]	T2_PT_LIP_Lisbon [10.1M]
T2_CN_Beijing [55.9M]	T2_PT_NCG_Lisbon [484k]	T2_PT_NCG_Lisbon [484k]
T2_DE_DESY [47.1M]	T2_RU_JHEP [4.53M]	T2_RU_JHEP [4.53M]
T2_DE_RWTH [72.3M]	T2_RU_JINR [1.45M]	T2_US_Nebraska [69.6M]
T2_EE_Estonia [8.04M]	T2_RU_ITEP [3.57M]	T2_US_Purdue [257M]
T2_ES_CIEMAT [49.0M]	T2_RU_JINR [18.1M]	T2_US_UCSD [61.6M]
T2_ES_JFCA [1.61k]	T2_RU_PNPI [1.07k]	T2_US_Wisconsin [148M]
T2_FL_HIP [2.21M]	T2_RU_RRC_KI [13.8M]	T3_IT_Padova [948k]
T2_FR_CIN2P3 [69.6M]	T2_RU_SINP [4.87M]	T3_UK_London_QMUL [26.7M]
T2_FR_GRIF [32.5M]	T2_TR_METU [252k]	T3_UK_London_RHUL [5.20M]
T2_FR_GRIF_LLJ [28.1M]	T2_TW_Taiwan [20.8M]	T3_UK_ScotGrid_GLA [3.68M]
T2_FR_IPHC [29.0M]	T2_UA_KIPT [1.53M]	T3_US_FIT [2.04M]
T2_HU_Budapest [27.5M]	T2_UK_London_Brunel [15.6M]	T3_US_Omaha [225M]
T2_IN_TIFR [12.5M]	T2_UK_London_IC [65.1k]	
T2_IT_Bari [16.4M]	T2_UK_London_Imperial [42.7M]	
T2_IT_Legnano [22.3M]		



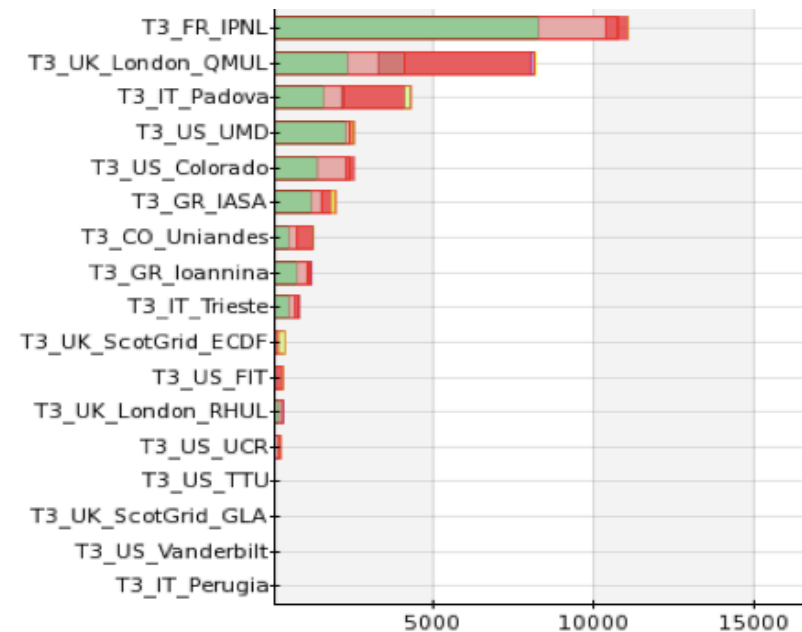
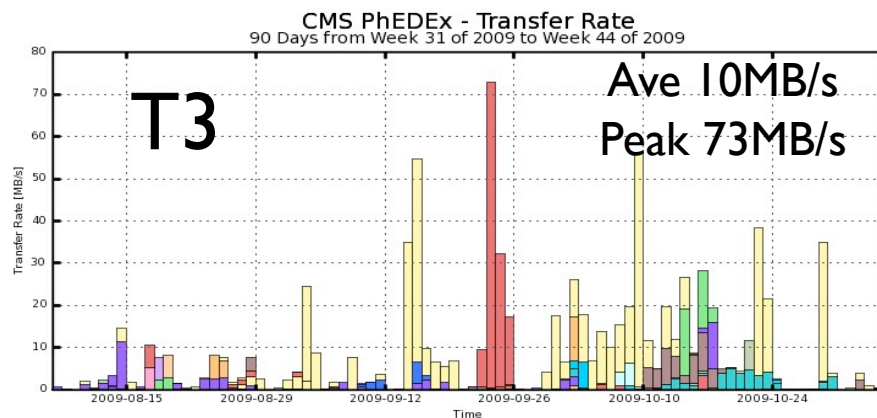
# Readiness for Operations

- ▶ Computing Operations still require intense effort
  - ▶ Good shift team in 3 time zones, weekly coordinators, and a team of analysis operators
  - ▶ Counting on continued improvements from Offline in capabilities of tools to improve operational efficiency and reliability and availability of sites
  
- ▶ Computing Operations relies heavily on monitoring information
  - ▶ Globally distributed system with transfer and processing requests coming from hundreds of people
  - ▶ Rely heavily on Dashboard, PhEDEx pages, Reporting from workflow tools
  - ▶ Still work to do on having better homogeneity in the monitoring system



# Readiness for Tier-3s

- ▶ The Tier-3s are analysis facilities for CMS
  - ▶ Though they are not very strictly defined and vary by size and capability.
- ▶ Current CMS has 45 Tier-3 centers registered in PhEDEx
  - ▶ 25 have successfully received data during the last quarter
  - ▶ About an equal mix of Tier-1 and Tier-2 sources
    - ▶ Compares to 48 Tier-2 destinations



- ▶ The Computing Project believes we are ready for the start of collision data
- ▶ There are open issues, some of which can be addressed in the final weeks and some will need to be worked around
  - ▶ Working on handling the occasional high data rates
  - ▶ We believe we can analyze data, but important issues were identified in the Oct Ex, not all of which can be fixed in the short term
  - ▶ The Tier-1s, Tier-2s, and a growing number of Tier-3s are ready for running
    - ▶ We still see stability issues, but redundancy in the early run will alleviate some issues
  - ▶ Good team of shifters though more effort will be needed for long running.