CMS

09Q3 Report

WLCG-MB 6.10.2009

Matthias Kasemann

Compact Muon Solenoid

Outline

- **Computing in 2010**
- **CMS Software status**
- **Site readiness**
- **T0, T1, T2 performance**
  - CRAFT
  - MC production
- **Computing Resources**
- **Analysis Operations**
- **Summary and Outlook**

Computing Organization: 2010

Computing
Ian Fisk
Deputies: Patricia McBride, Daniele Bonacorsi

Computing Resource Board
Chair - NN

Offline

Common Coordination
Integration CERN Fac+Services
Stephen Gowdy, Dave Mason
Integration Distr. Fac+Services:
Claudio Grandi, Jose Hernandez
Resource Coordination:
Marie-Christine Sawley

User Analysis Operation
Stefano Belforte
Frank Wuerthwein

Facilities Operations
Joseph Flix
Peter Kreuzer

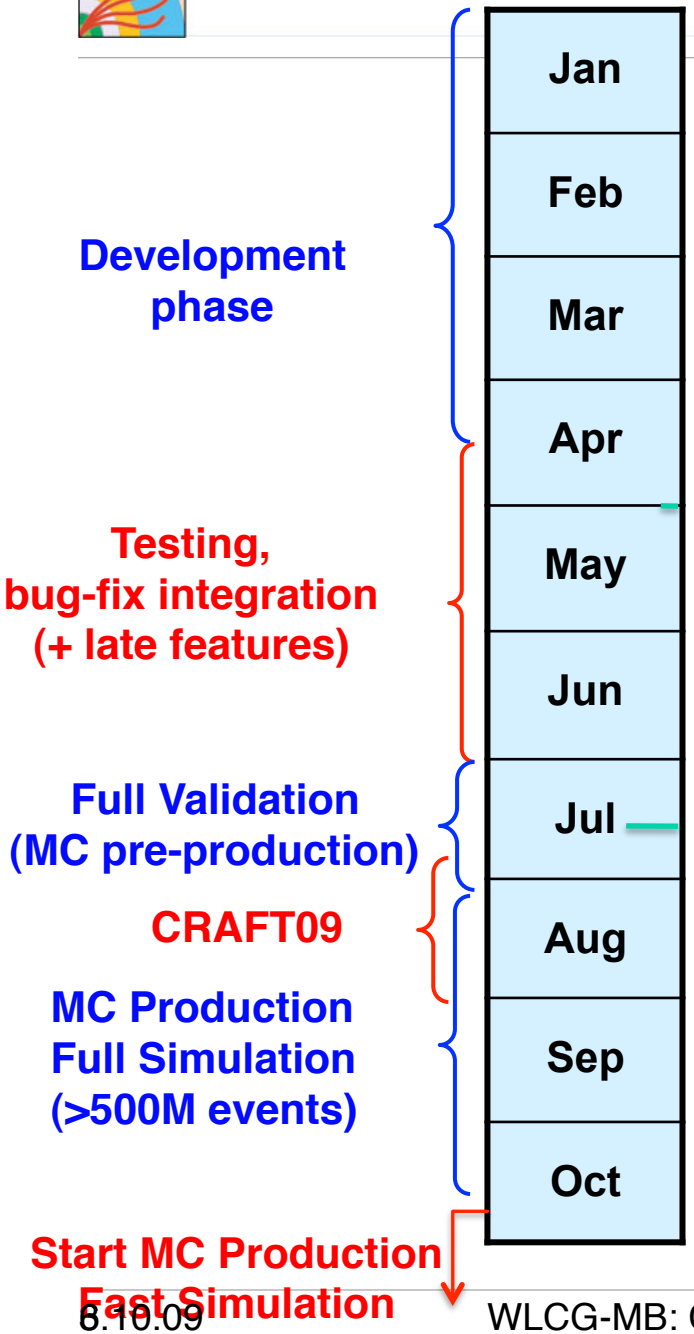
User Support
Kati Lassila-Perini
Sudhir Malik

Data Operations
Oliver Gutsche
Markus Klute

WBS:

Σ (over 49 Tasks)
1420.60 Mo 118.38

Status of CMSSW Release



- LHCC Feb 2009 : "Will continue to work on adding/ improving functionality in all CMSSW areas, preparing for tuning with real data"
- Long development period lasting >3 months
  - >500 package changes in last dev pre-release
- CMSSW_3_1_0 released on July 1st
  - Release targeted at MC production goals
  - Validation using 52 different pre-production samples
  - 3_1_1 (July 7th): first pre-production round (~15M events)
  - 3_1_2 (July 24th): second pre-production round
  - MC production started using 3_1_2 on July 29th
- CMSSW_3_2_0 released on July 19th
  - Release targeted at CRAFT09 data-taking, processing and re-processing (to be done with 3_2_7)

Software Performance

TTBAR events

Processing step	Metric		2_2_13	3_1_0	3_2_4
Full Simulation (GEN+SIM+HLT) (100 events)	CPU Time (s)		128	87.2	86.9
	Peak Vsize (GB)		1.25	1.14	1.14
		PU	1.97	1.40	1.40
	File size on disk (MB)		1.98	1.48	1.39
Reconstruction (RAWtoDIGI+RECO) (8000 events)	CPU Time (s)		4.16	3.96	4.06
	Peak Vsize (GB)		1.14	0.95	1.25[1]
	File size on disk (kB)		330	416	420
Fast Simulation (8000 events)	CPU Time (s)		2.68	2.33	2.47
	Peak Vsize (GB)		1.7[2]	1.88[2]	1.65[2]
	File size on disk (kB)		116	116	112

- **numbers obtained running Performance Suite on Intel(R) Xeon(R) CPU 5160 @ 3.00GHz**
  - **running on 1 core while loading other 3 cores with a cache-contained cpu intensive benchmark**
- **[1] memory leak – fixed in 327, [2] several memory leaks – fixed in 330_pre4**
- **expect gcc4.3.2 to give further small gains 'out-of-the-box'**

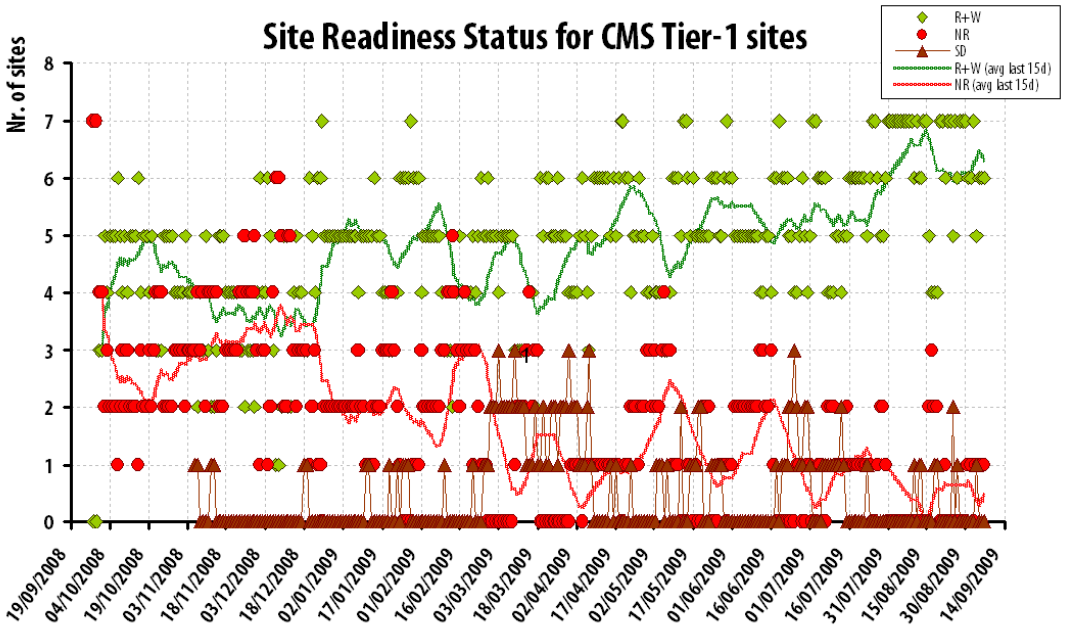
Software Status:

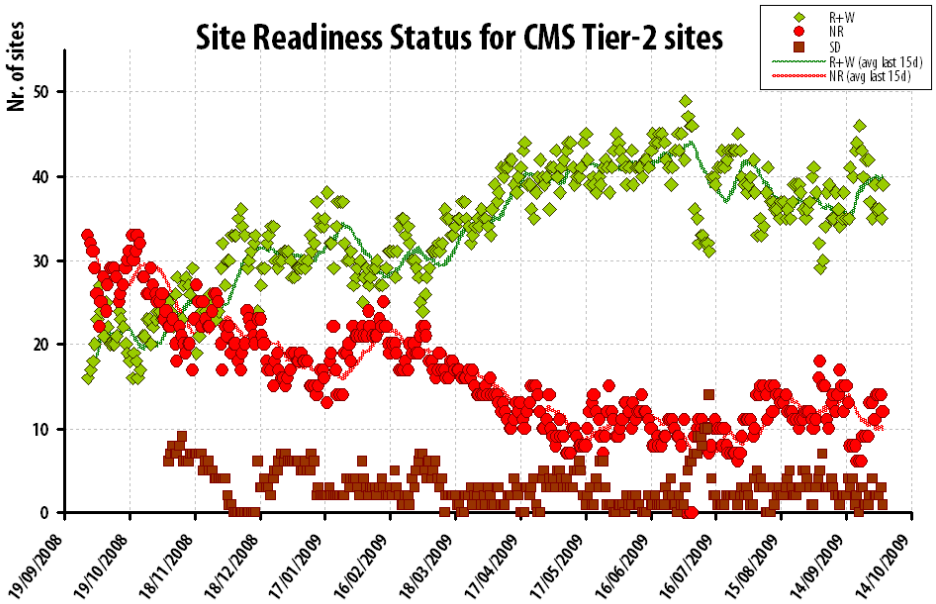
- Since June focus has been on integration & validation of CMSSW Releases (3_1_x/3_2_x) for use in MC production and CRAFT09
- Steps have been taken to improve validation procedures (PVT)
- CRAFT exercise has been very valuable for adapting software to run in data-taking conditions
  - Tier0 workflows ran very reliably - very few failures in prompt reconstruction
  - Prompt calibration loop exercised at CAF for the first time
  - Deployment of online/offline patches in operation
- Further tests of T0/T1 production systems using Monte Carlo samples are in progress
- Automation of prompt calibration workflows using CRAB is high priority
- Trying out a new model for managing releases better adapted to data-taking, including shorter release cycle (~ 6 weeks)
- Validation of SLC5/gcc432 planned for CMSSW_3_3_0

T1 and T2 Site Readiness

The Site readiness is closely monitored:

- Reports and follow-up during weekly Facility Operations meetings
- Additional meetings to focus on Asian and Russian&Turkish sites

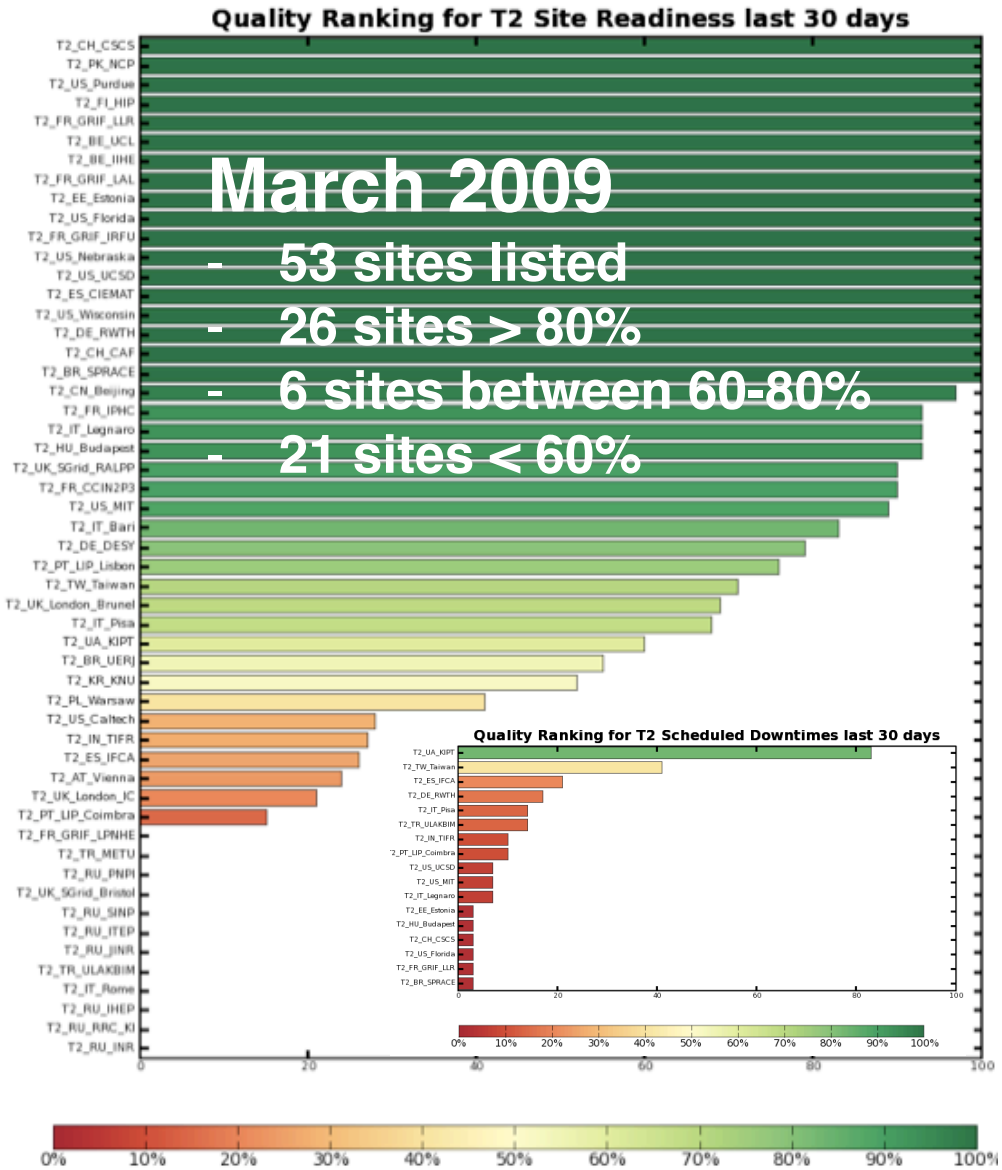


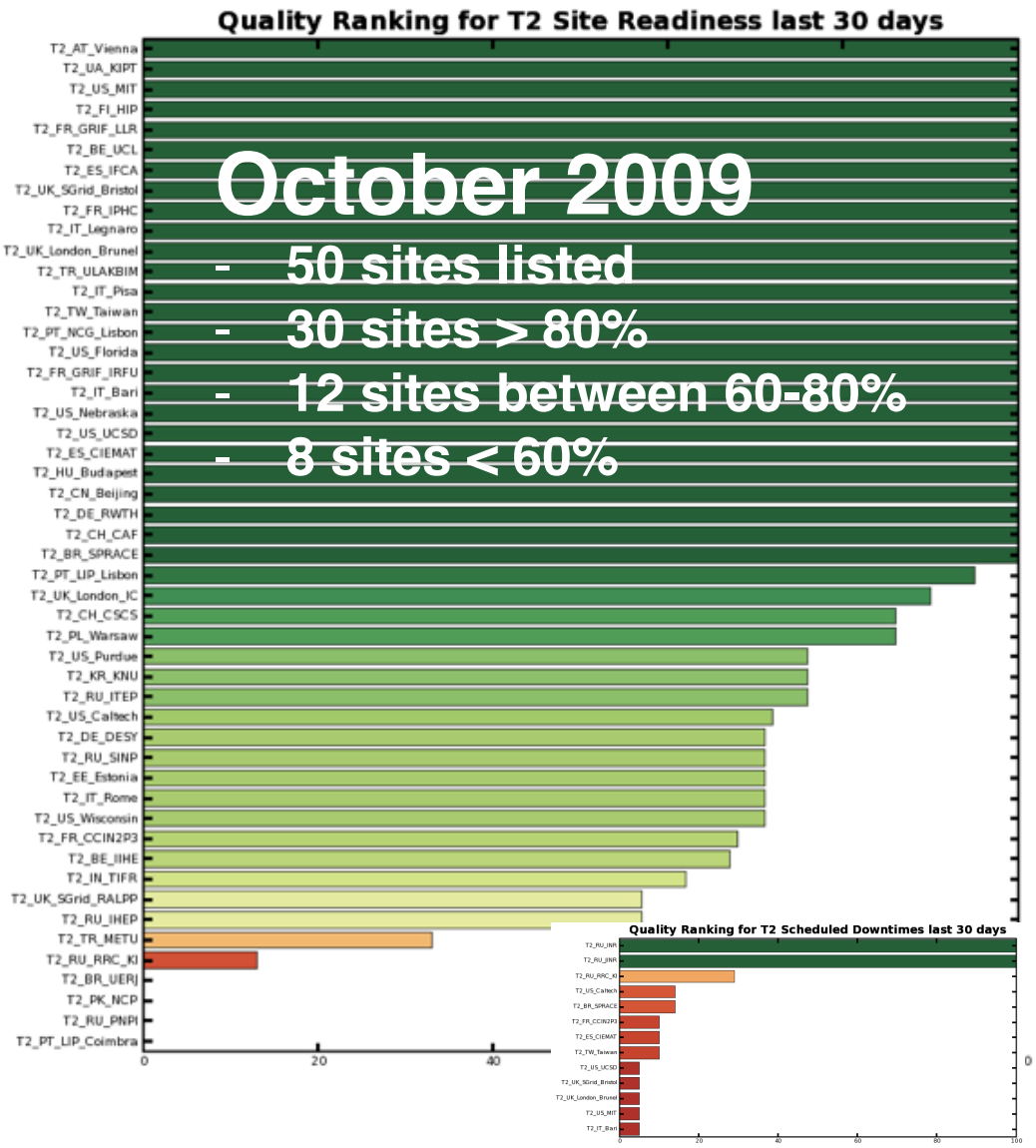


Substantial improvement is observed for large number sites.

- **Tier-1: sites readiness a concern. Improving lately.**
  - **Plan expert visits to improve the situation .**
- **Tier-2: readiness state improved significantly over the last year. Need to sustain efforts.**

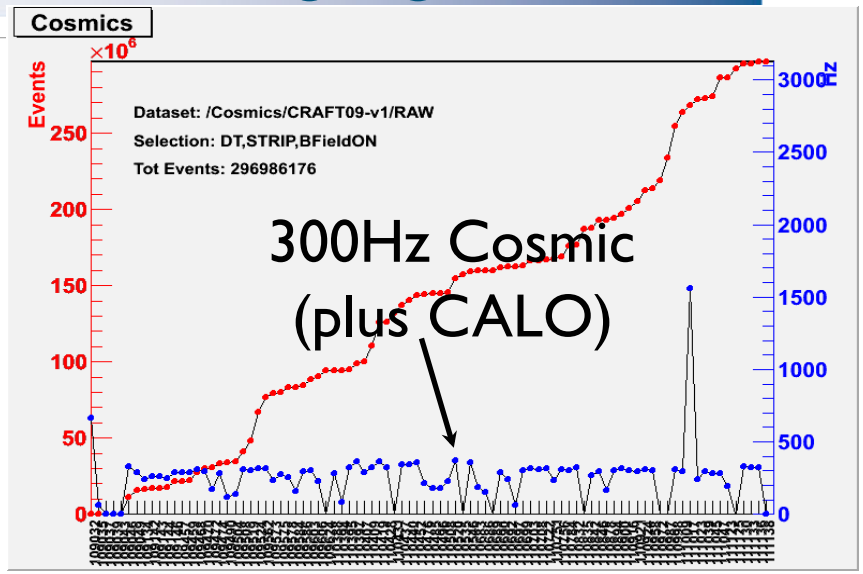
Site readiness T2: substantial improvement

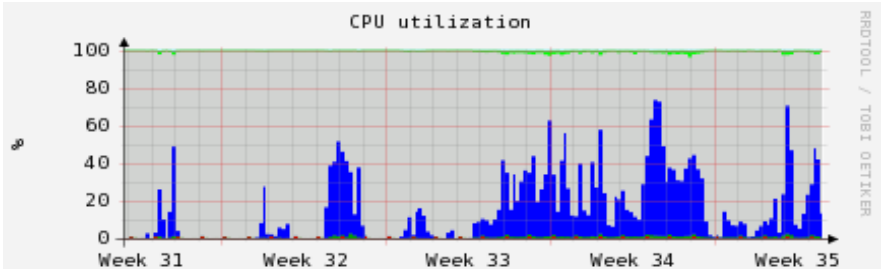


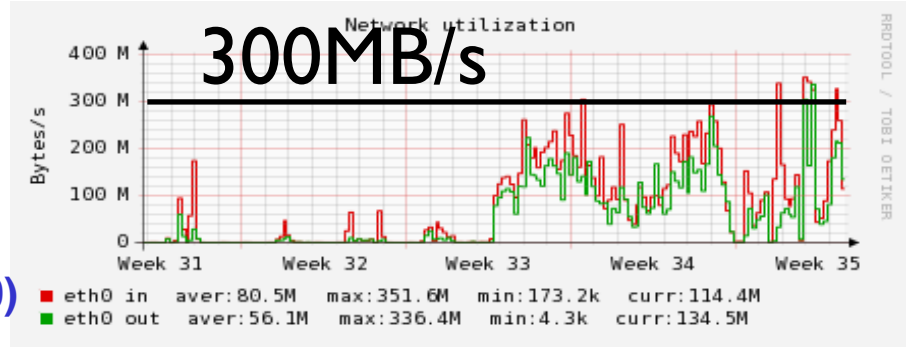


Cosmics run: CRAFT Tier-0

- **CRAFT rate of prompt reconstruction is higher than expected collision data**
  - **Processing and data movement is smaller because the reduced event size and complexity**
    - Setting up a MC based T0-test to exercise other elements
  - **CRAFT data does not fully utilize the Tier-0 reconstruction farm**
    - IO Rates also lower
    - Collision data will utilize the full resources
- **Express Stream was commissioned**
- **DQM Harvesting of Express was integrated**
- **Prompt_Calibration_Hold for Prompt Reco was introduced (>run 110500)**



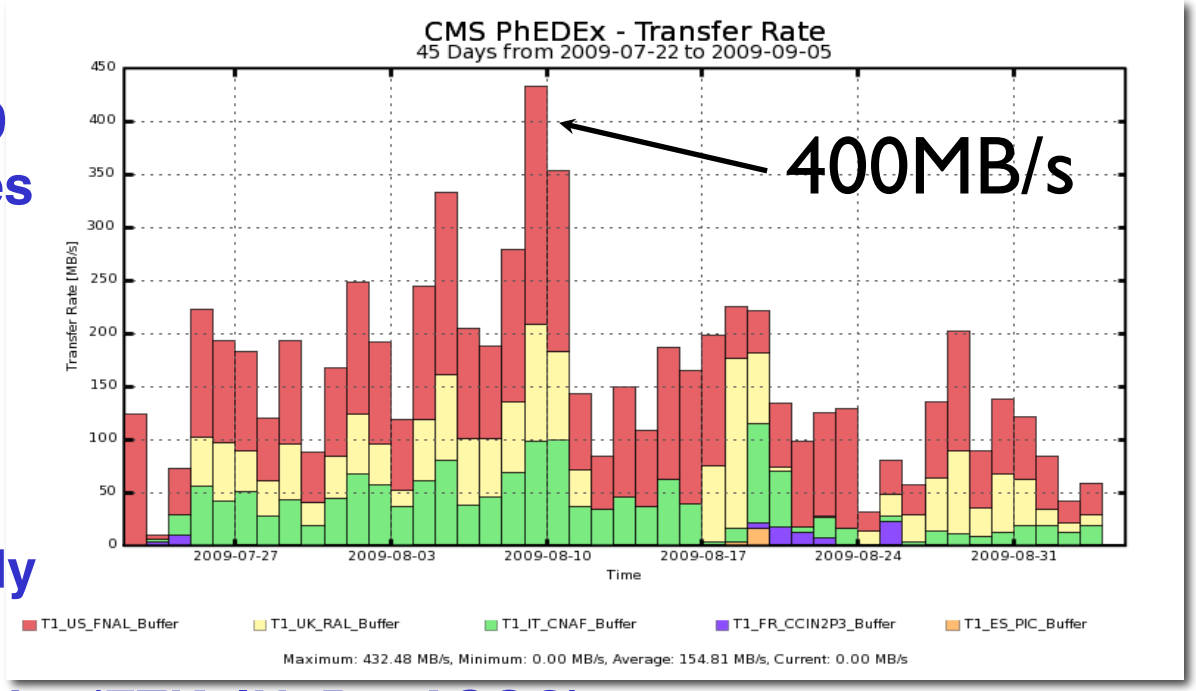




Tier-1s

- **Tier-1 traffic during CRAFT was custodial transfers to RAL and CNAF**
  - **Additional copy of the data to FNAL**
    - Generally successful
      - Interesting exercise in recovery after RAL cooling failure

- **Preparing for CRAFT08 and 09 re-reconstruction at Tier-1 sites**



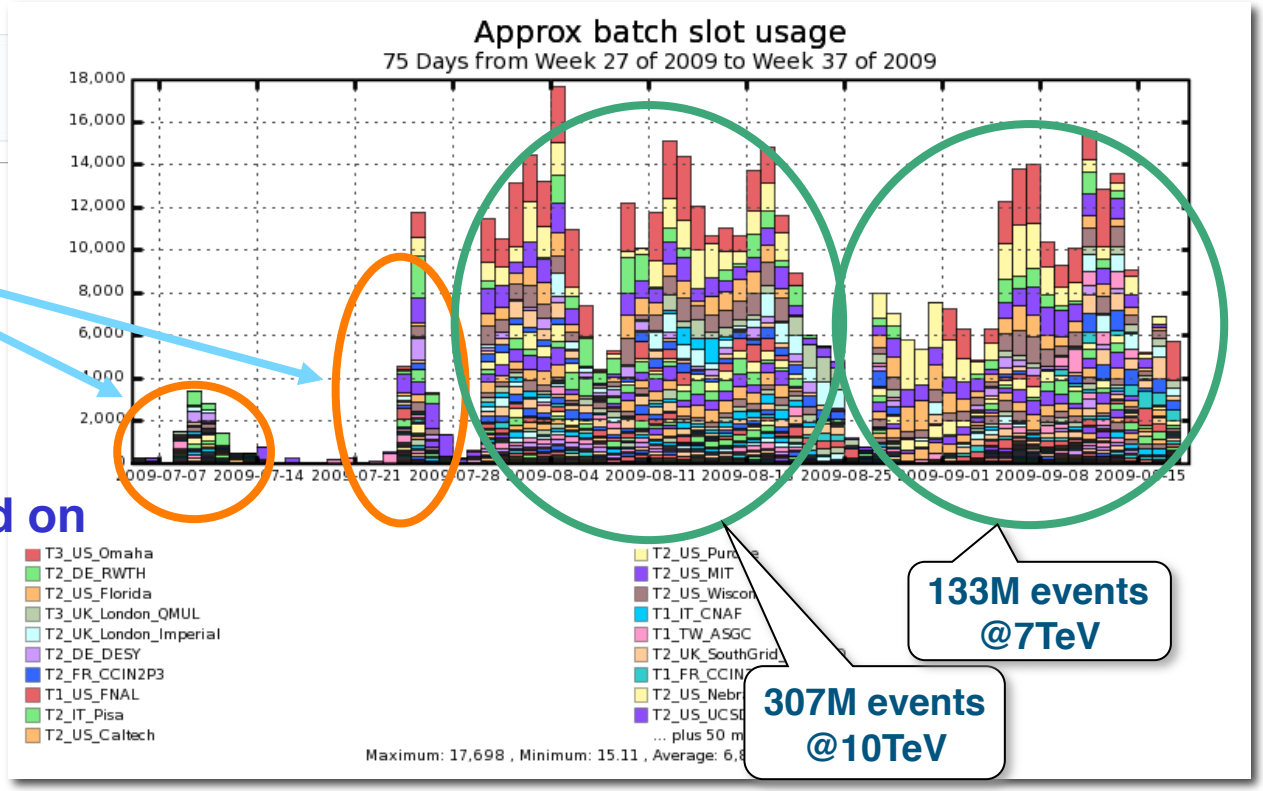
- **Not all Tier-1s are equally ready to accept collision data**
  - **Computing is preparing visits (FZK, IN2P3, ASGC) and exercises during the fall**
    (repeat STEP'09 tests at some sites)

MC production

Pre-production of MC generation in July.

- **~15M events produced within few days**

Summer'09 MC production started on July 29, when pre-production validation finished.



Production status:

- **Events requested (10TeV + 7TeV): > 550 M**
- **Events produced: > 500 M**

Large number of resources could be grabbed and used efficiently

- **About 15k slots pledged during STEP'09**
- **Quite some opportunistic non-CMS-T2-pledged resources used (T3, beyond pledge, non-CMS)**
- **Analysis ongoing in parallel**

Sites running MC production (1.8.-30.9.)

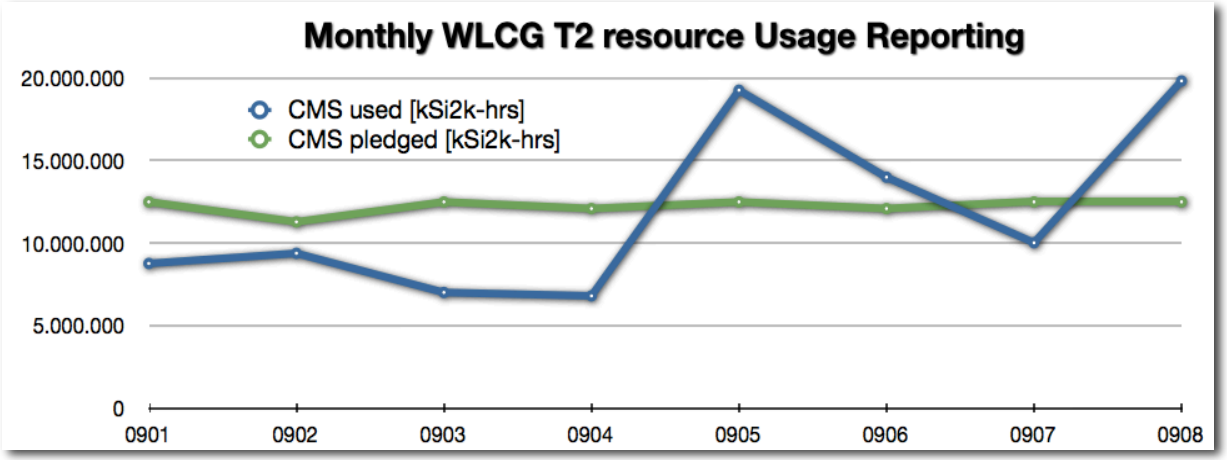
Parallel Running Jobs (Average/Maximum)

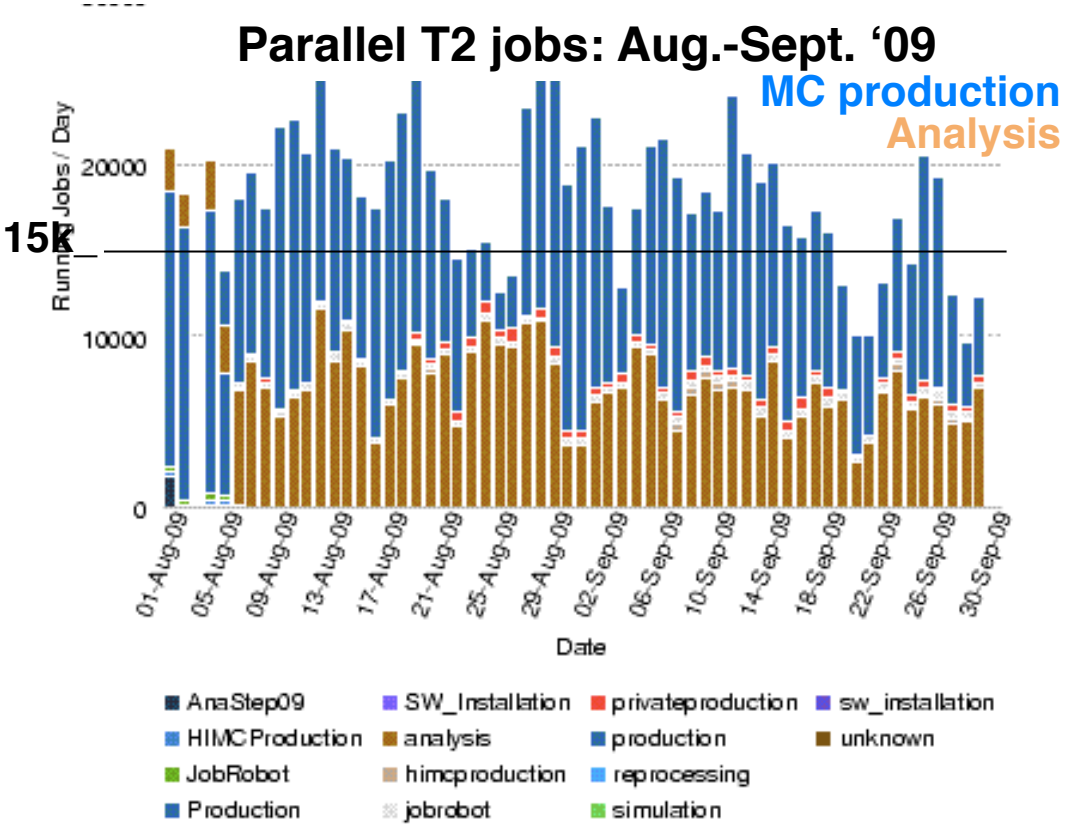
Site	all	production	Production
BEIJING-LCG2	299/709	199/441	176/279
BEgrid-ULB-VUB	386/1267	89/172	85/155
BUDAPEST	194/475	112/322	167/232
BelGrid-UCL	252/514	227/413	250/401
CIEMAT-LCG2	550/1039	372/1005	310/483
CIT_CMS_T2	741/1611	360/766	362/642
CSC	13/76	0/0	0/0
CSCS-LCG2	369/1915	212/741	206/503
DESY-HH	531/1235	299/967	198/466
GLOW	1115/2492	516/1631	624/121
GRIF	339/1500	139/435	130/305
HEPGRID_UERJ	1/8	0/0	0/0
HEPGRID_UERJ_OSG64	50/221	0/0	0/0
Hephy-Vienna	374/1337	171/404	162/305
IFCA-LCG2	519/4089	50/432	1056/40
IN2P3-CC-T2	449/1264	342/1243	212/294
IN2P3-IRES	132/392	88/385	164/316
INDIACMS-TIFR	196/354	144/304	156/233
INFN-BARI	169/481	71/231	137/212
INFN-LNL-2	342/1515	74/305	260/409
INFN-PISA	879/1544	545/1164	569/782
INFN-ROMA1-CMS	210/677	143/472	37/37
ITEP	113/279	45/119	58/84
JINR-LCG2	204/628	165/554	191/260
Kharkov-KIPT-LCG2	34/104	25/65	41/96
LCG_KNU	163/336	120/245	111/154
LIP-Coimbra	11/77	0/0	0/0
LIP-Lisbon	150/532	94/269	113/241
MIT_CMS	1010/1781	391/1196	589/1239
NCG-INGRID-PT	147/1833	0/0	0/0
NCP-LCG2	4/9	0/0	0/0
NDGF-T1	204/901	145/892	186/556
Nebraska	461/1410	226/637	470/692
Purdue-RCAC	1473/5117	1251/5064	1138/2715
Purdue-Steele	240/588	244/588	103/197
RRC-KI	14/115	5/14	11/16
RU-Protvino-IHEP	46/144	30/79	51/81
RWTH-Aachen	1185/2813	826/2711	525/1707
Ru-Troitsk-INR-LCG2	41/132	37/93	0/0
SPRACE	25/102	0/0	0/0
T2_Estonia	262/1012	183/1008	218/307
TR-03-METU	28/120	24/69	6/17
TW-FTT	279/1240	133/472	114/240
UCSDT2	327/1696	130/363	184/323
UFlorida-HPC	601/2603	297/929	341/831
UKI-LT2-Brunel	161/474	119/469	201/313
UKI-LT2-IC-HEP	770/2312	540/1291	321/1072
UKI-SOUTHGRID-BRIS-HEP	35/98	8/15	8/11
UKI-SOUTHGRID-RALPP	614/2017	416/1648	619/962
WARSAW-EGEE	150/399	108/334	82/169
ru-Moscow-SINP-LCG2	84/166	46/123	66/95
ru-PNPI	11/125	0/0	43/125
ucsdt2-b	369/1693	153/352	221/390
Summary	17616/26641	0/0	11134/16584

Good production performance:

- **Large number of sites participating**

- **Daily averages:**
  - **53 active sites (more than 0 jobs/day of any kind)**
  - **46 sites ran MC production (more than 0 MC jobs/day)**
  - **41 sites with more than 50 MC jobs/day**





T2 resource usage reported to WLCG on monthly basis
(Installed resources: queried by CMS in June'09)

Since May'09:
CMS uses ~all T2 resources
(plus opportunistic resources)

- **May-June: STEP'09**
- **August - September: MC production**
- **Analysis uses ~ 40% of slots**

T2 resources can be used effectively for MC and Analysis
(…if they pass the site readiness tests)

Resources: LHCC & C-RSG reviews

LHCC Mini Review of Computing Resources (July)

"…important investment in the construction of the LHC and the detectors… **physics outcome using very first LHC data should be maximized and not limited by computing resources**. … current estimates suffer from large uncertainties … not an appropriate time to cut back substantially on computing resources."

RRB Computing Resources Scrutiny Group

"Generally speaking the resources are well justified."

C-RSG estimates vs. CMS requests:

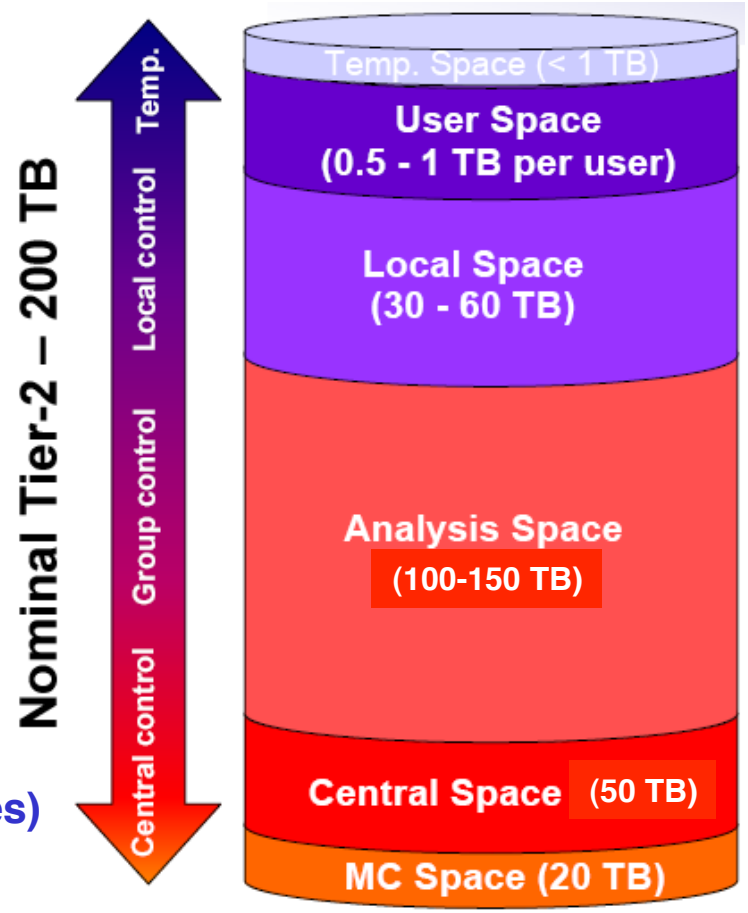
CERN resources	2009		2010	
	Scrutiny	CMS	Scrutiny	CMS
Tier0 CPU (kHS06)	37.1	37.1	61.9	61.9
CAF CPU (kHS06)	7.6	7.6	34.7	34.7
Tier0 disk (PB)	0.4	0.4	1.0	1.0
CAF disk (PB)	1.3	1.3	3.1	3.1
CERN tape (PB)	7.4	7.3	12.6	14.6
Non-CERN resources	**2009**		**2010**	
	Scrutiny	CMS	Scrutiny	CMS
Tier1 CPU (kHS06)	45.8	46.0	101.3	100.5
Tier1 disk (PB)	6.5	6.5	13.7	13.4
Tier1 tape (PB)	10.3	11.9	23.3	23.3
Tier2 CPU (kHS06)	73.2	75.2	199.0	195.0
Tier2 disk (PB)	3.7	3.7	9.0	9.2

Computing support of Analysis

Analysis Operations Task started in June
(S.Belforte, F.Würthwein and J.d'Hondt)

- **Data Placement Operations**
  - **Update T2 Associations.**
  - **Management of Central Space at T2's.**
  - **Group Skim Transfer Service.**
  - **Increase group space allocation to 50 TB.**

- **Support analysis based on CRAB.**
  - **To achieve fewer failed jobs and more streamlined data analysis operations for CMS physicists.**
  - **Operate several CRAB-servers (allows better user support, assist access to log-files)**



Improvement of analysis success rate is the biggest challenge.

- **Improve shift procedures ongoing**

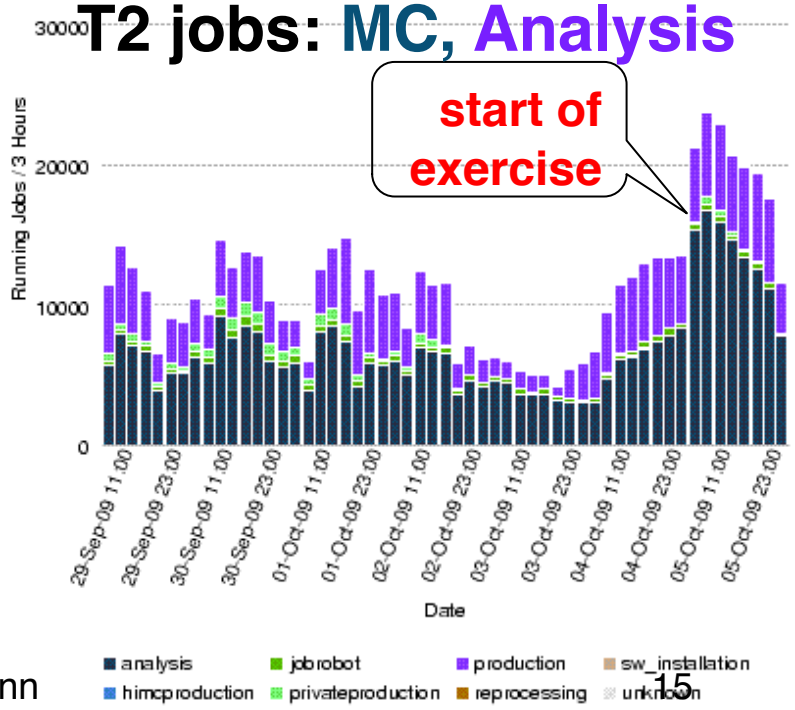
- **Distributed Analysis Metrics Evaluation**

October Physics Exercise 5.-18.10.2009

- "Data-Operations" produced "Secondary datasets" from MC
  - skimming on trigger quantities and pre-scales on the T1's, distribute to T2's

- "Analysis-Operations" supports data transfer and analysis
  - Operates CRAB-submission servers, to ease job-failure analysis
  - Performs data transfers and management of "central data samples"
  - Supports physics group production of "group data samples" (concept of 1 priority-user per group)
  - Supports registration and transfers of "group data samples"
  - Provides tailored documentation and training

- Computing has 'October-exercise-contacts'
  - Quick response to operational questions
  - Active participation in October-Exercise-Meetings (weekly, daily)



ToDo for Computing

- **Analysis operation support getting strong**
  - **October Exercise, data management, analysis jobs success-rate**

- **Program until the LHC start-up**
  - **Finalize data distribution of RAW, RECO and AOD to CAF, T1 and T2 centres**
  - **Tier-0: repeat scale tests using simulated collision-like events**
  - **Tier-1: STEP'09 tape and processing exercises where needed, T1 visits scheduled: GridKa, IN2P3, ASGC (at CERN)**
  - **Tier-2: Support and improve distributed analysis efficiency (Analysis Operations)**
  - **Review Critical Services coverage**
  - **Fine tune Computing Shifts procedures**
  - **Make sure (2010) resources pledges are available**

Conclusions

- **Resources requests for 2009/2010 were reviewed. From the C-RSG report:**
  "…Generally speaking the resources are well justified…"
  "…2010 resources should be … in place by June 1st 2010…"

- **CMSSW_3_x deployed for LHC data taking + MC production**

- **CRAFT09 data successfully recorded and distributed in August.**

- **Large MC samples (500 M FullSim events) were produced at higher than normal rate in August/September.**

- **Computing is ready for LHC data.**

Backup slides

HI Computing Workshop: Sept. 12. 2009

All aspects of HI computing reviewed on Sept. 12, '09

Goals:

- Review resource needs
- Review preparations for
  - Data recording
  - T0 use
  - CAF use
  - Data distributions
  - Analysis
- Review operational aspects

Saturday 12 September 2009

09:00	Opening Remarks: Workshop Goals (15')
09:15	CMS-HI Physics and Instititutions (25') (Slides)
09:40	Overview of CMS-HI Computing Plans (20') (Slides)
10:00	T0/CAF Operations for CMS-HEP (25') (Slides)
10:25	AlCa for CMS-HI at T0/CAF (30') (Slides)
10:55	Coffee Break
11:25	DQM for CMS-HI (30') (Slides)
11:55	Raw Data Reconstruction for CMS-HI (30') (Slides; docume
12:25	Hardware Configuration at ACCRE (20') (Slides)
12:45	CMS Software Operations at ACCRE (20') (Slides)
13:05	Lunch break
14:30	Proposed/Pledged Resources For CMS-HI (40') (Slides)
15:10	Analysis and Simulation Computing Plans for CMS-HI (45') (Sl
15:55	CMS-HI Software Integration (30') (Slides)
16:25	Coffee break
16:55	Closeout Discussion and Action Items (1h05')

Follow-up during Offline&Computing week in October.