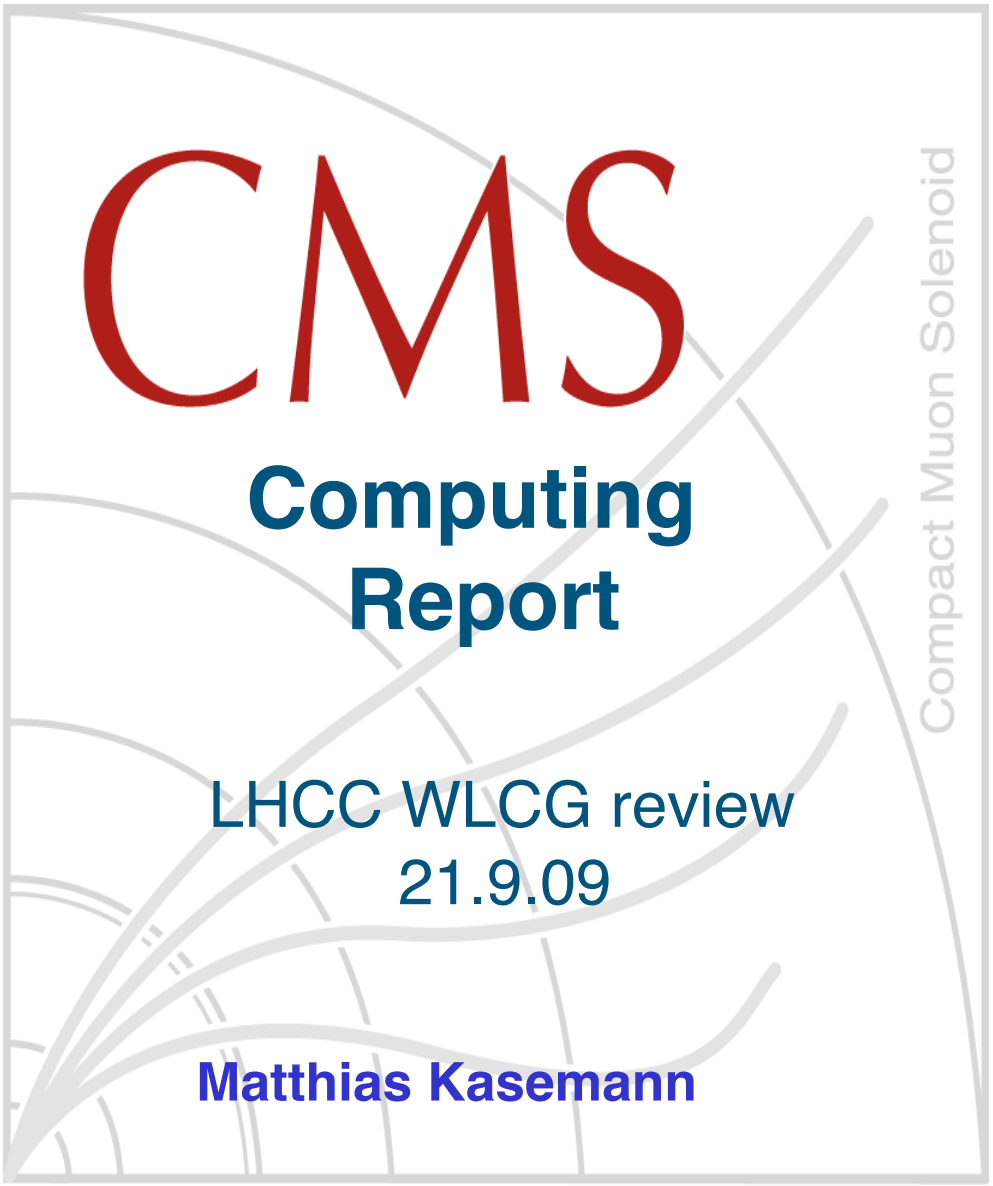



# Computing Report

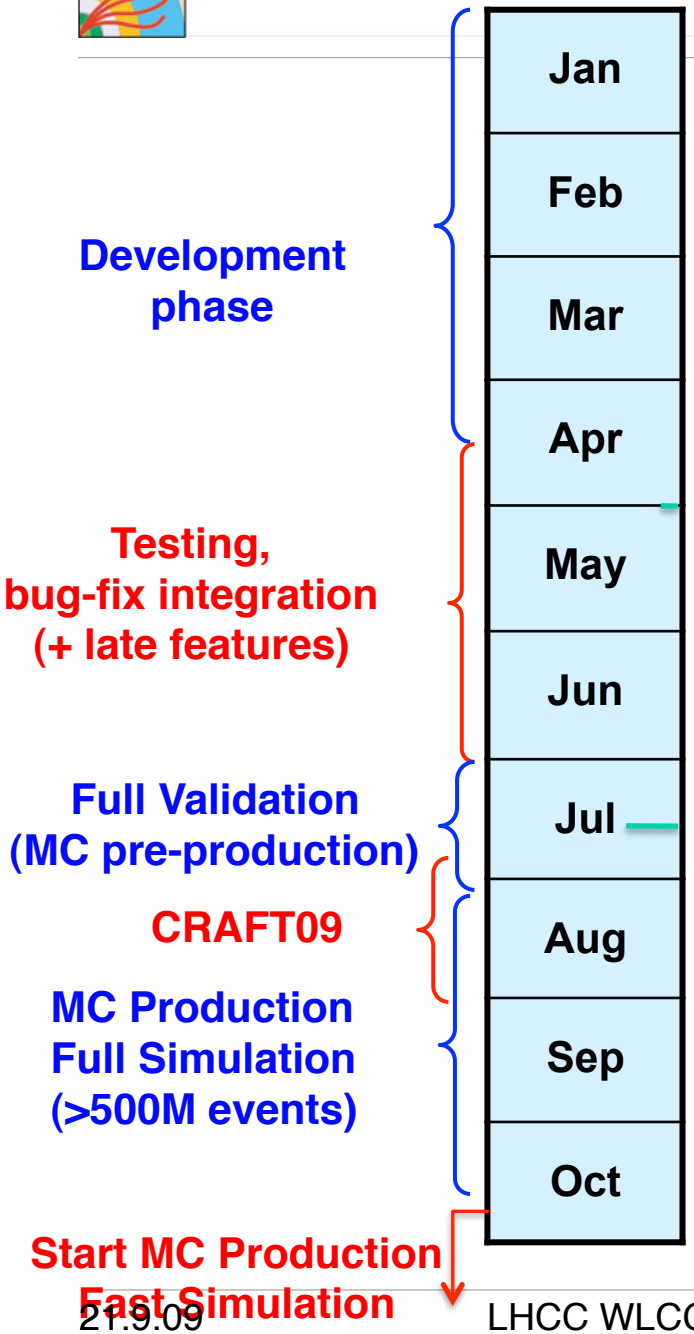
LHCC WLCG review
21.9.09

**Matthias Kasemann**

## Outline

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

# 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

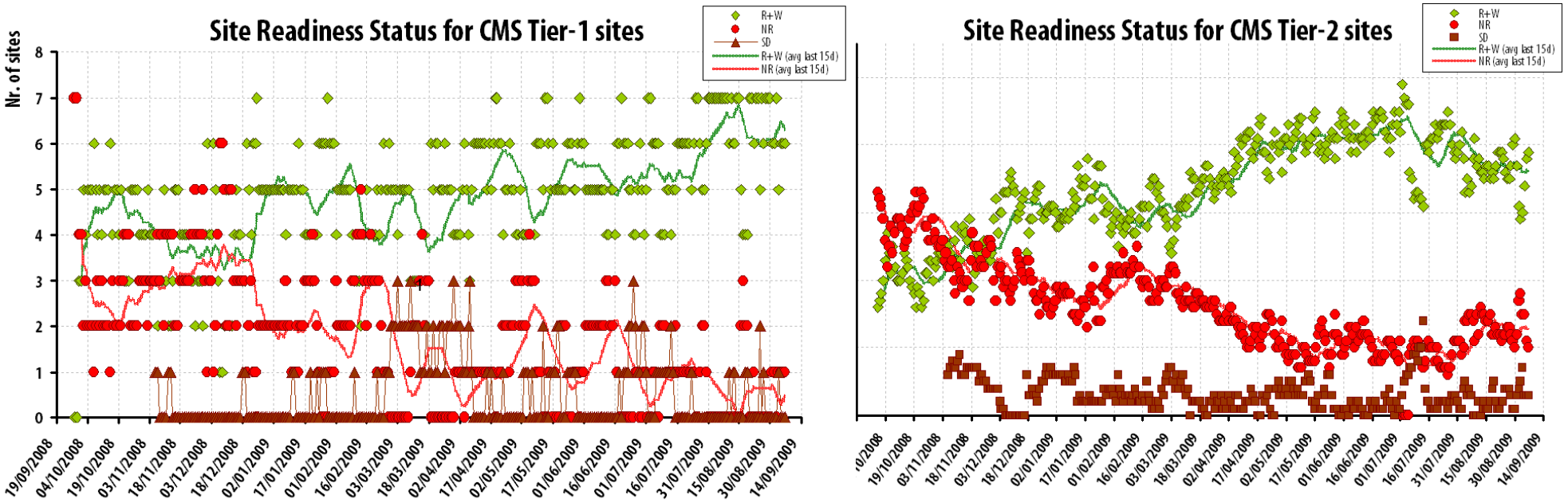
| Processing step | Metric | | 2_2_13 | 3_1_0 | 3_2_4 |
|---|---|---|---|---|---|
| Full Simulation (GEN+SIM+HLT) (100 events) | CPU Time (s) | | 128 | 87.2 | 86.9 |
| | Peak Vsize (GB) | | 1.25 | 1.14 | 1.14 |
| | | PU | 1.97 | 1.40 | 1.40 |
| | File size on disk (MB) | | 1.98 | 1.48 | 1.39 |
| Reconstruction (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'**

# T1 and T2 Site Readiness

## The Site readiness is closely monitored:

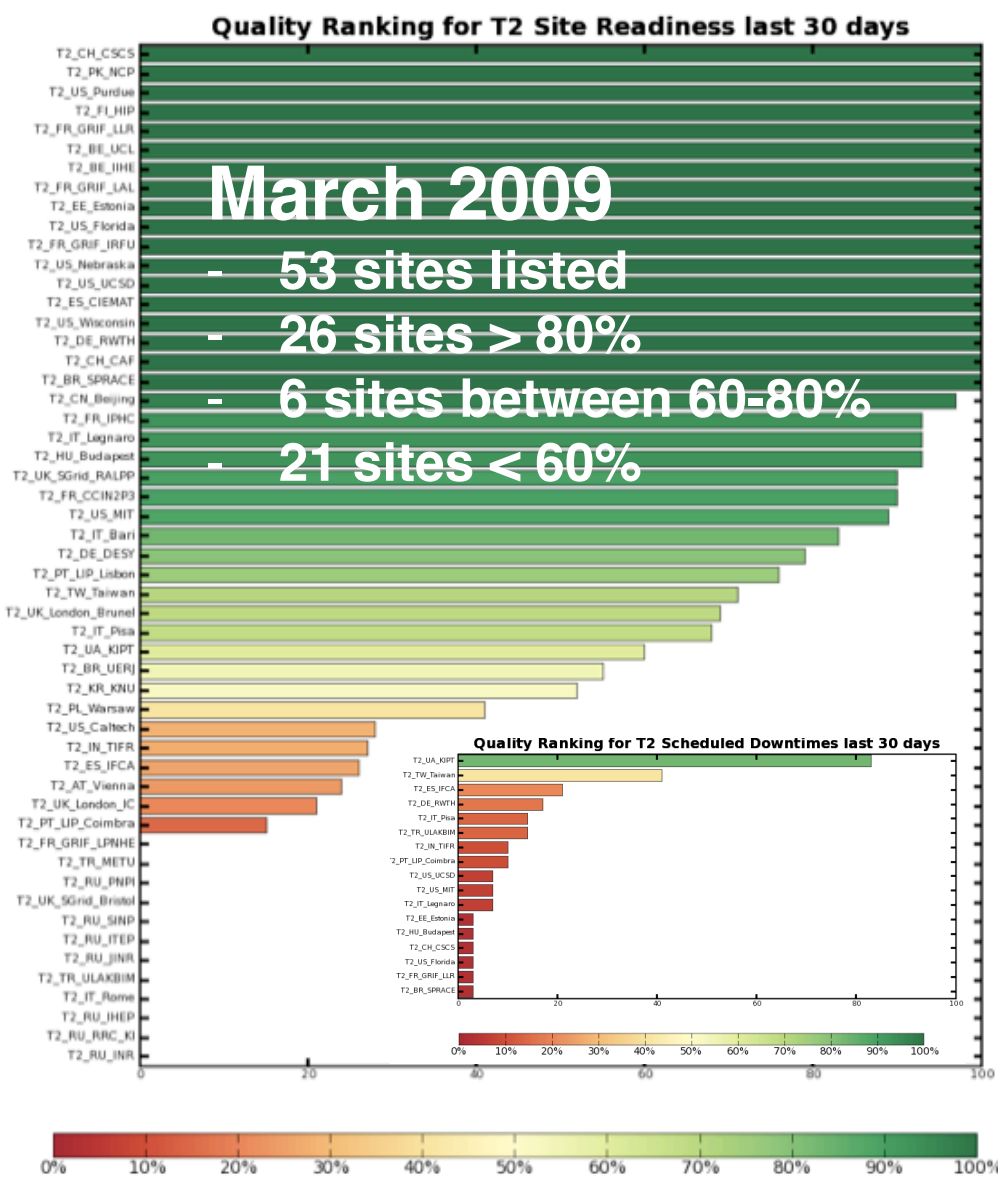
- Reports and follow-up during weekly Facility Operations meetings
- Additional meetings to focus on Asian and Russian&Turkish sites

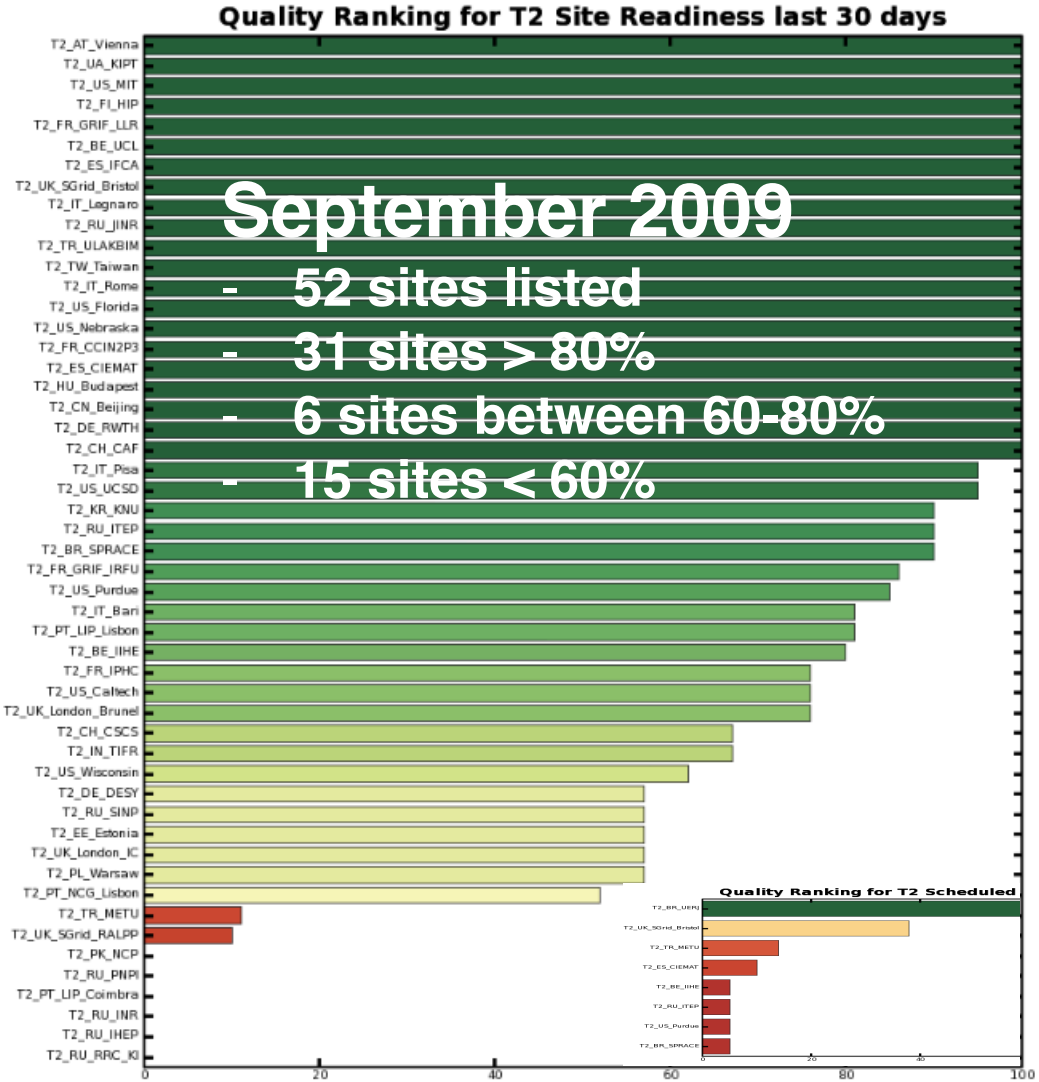



## Substantial improvement is observed for large number sites.

- **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.**

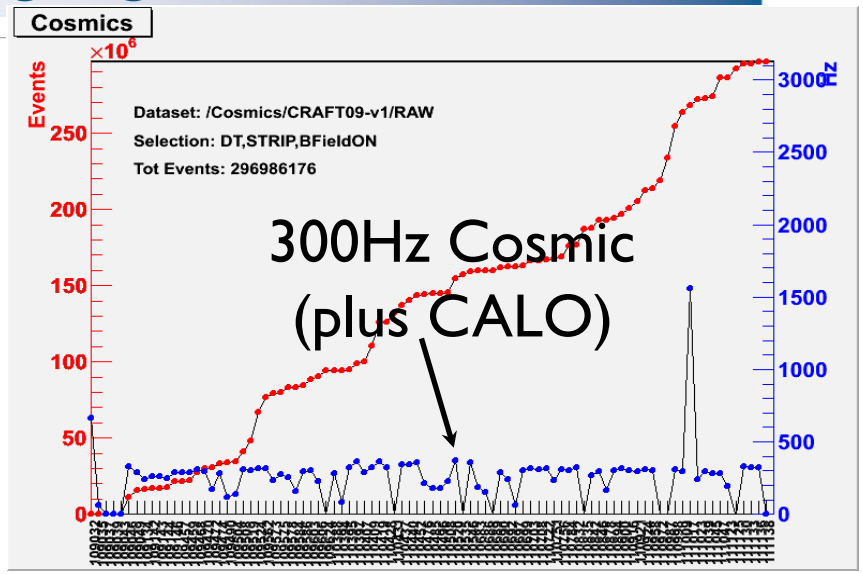
# Site readiness T2: substantial improvement

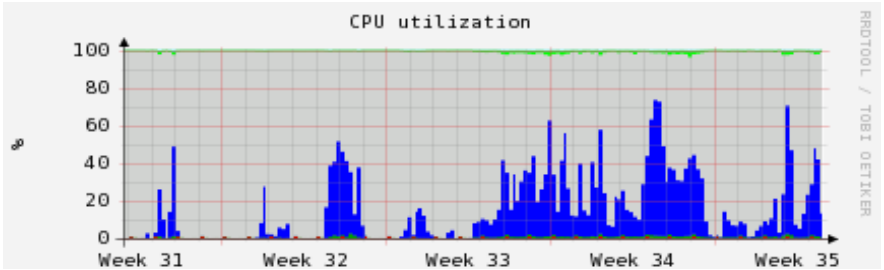


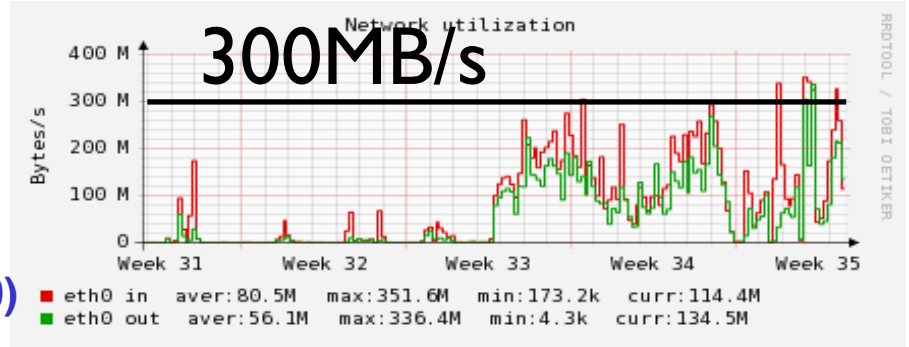



# CRAFT Tier-0

- **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)**



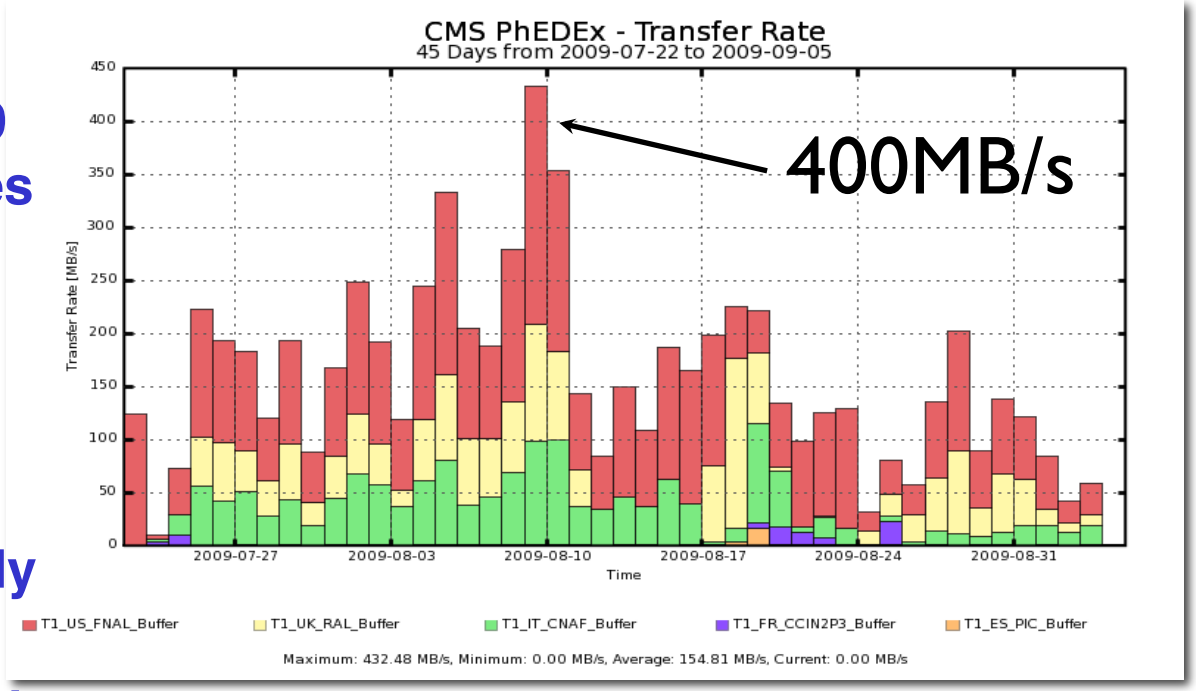



# Tier-1s

- **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**



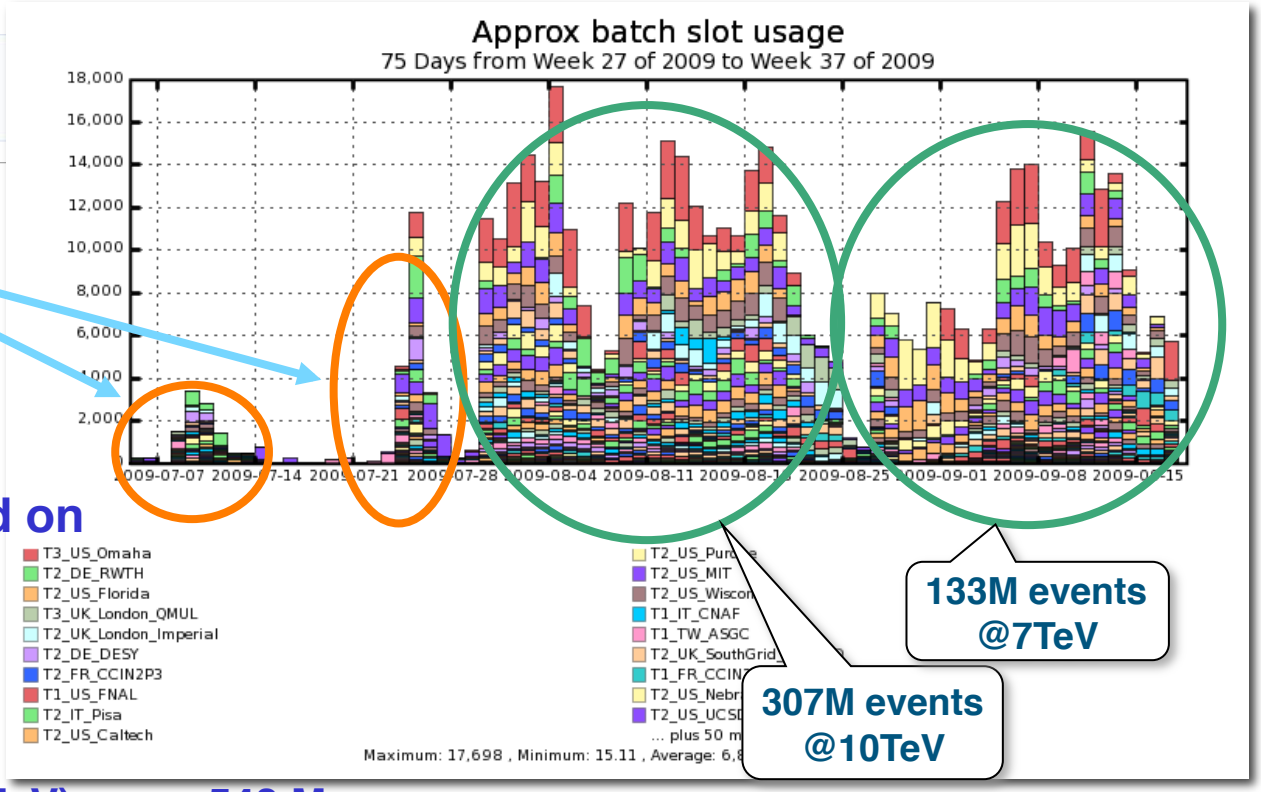

- **Not all Tier-1s are equally ready to accept collision data**
  - **Computing is preparing visits and exercises during the fall (repeat STEP'09 tests at some sites)**

# 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 (20.9.09):**

- **Events requested (10TeV + 7TeV): 543 M**
- **Events produced: 439 M**
- **Events to produce: 105 M**
- **Events available in DBS@Tier-1: 398 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**

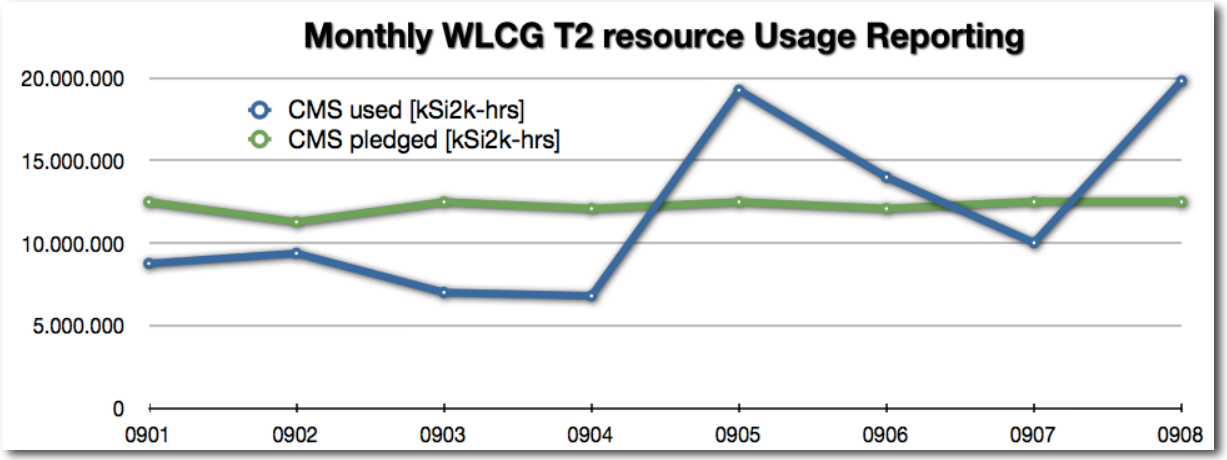
# Sites running MC production (22.7.-20.9.)

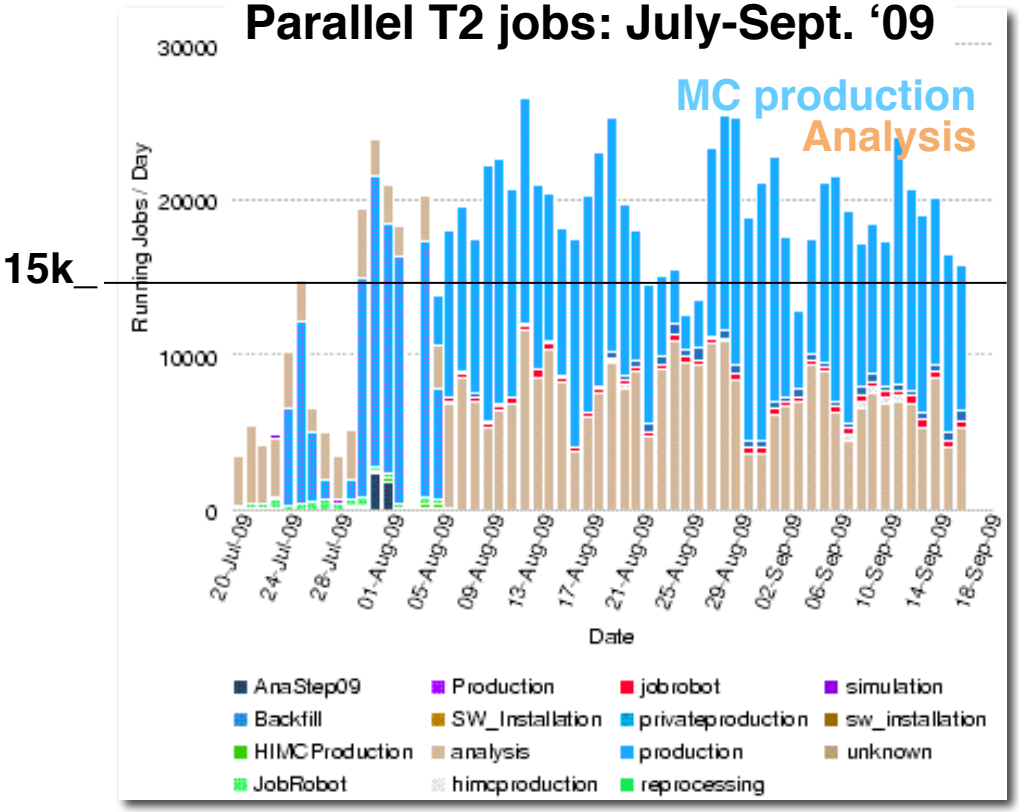
**Parallel Running Jobs (Average/Maximum)**

| Site | all | production | Production |
|---|---|---|---|
| BEIJING-LCG2 | 285/709 | 187/441 | 151/284 |
| BEgrid-ULB-VUB | 402/1267 | 88/172 | 79/155 |
| BUDAPEST | 177/475 | 89/322 | 125/232 |
| BelGrid-UCL | 263/514 | 236/413 | 197/407 |
| CIEMAT-LCG2 | 516/1039 | 356/1005 | 273/483 |
| CIT_CMS_T2 | 787/1611 | 331/766 | 273/661 |
| CSC | 1/2 | 0/0 | 0/0 |
| CSCS-LCG2 | 310/1915 | 156/741 | 157/503 |
| DESY-HH | 450/1235 | 237/967 | 248/899 |
| GLOW | 1123/2449 | 470/1342 | 572/1218 |
| GRIF | 313/1500 | 138/435 | 113/305 |
| HEPGRID_UERJ | 3/58 | 0/0 | 0/0 |
| HEPGRID_UERJ_OSG64 | 37/221 | 0/0 | 0/0 |
| Hephy-Vienna | 347/1337 | 145/404 | 123/305 |
| IFCA-LCG2 | 626/5417 | 43/432 | 1623/5242 |
| IN2P3-CC-T2 | 435/1264 | 323/1243 | 207/369 |
| IN2P3-IRES | 105/396 | 64/385 | 147/394 |
| INDIACMS-TIFR | 191/354 | 144/304 | 148/233 |
| INFN-BARI | 148/481 | 59/226 | 120/212 |
| INFN-LNL-2 | 316/1515 | 46/199 | 192/409 |
| INFN-PISA | 814/1544 | 496/1164 | 385/782 |
| INFN-ROMA1-CMS | 181/677 | 136/472 | 27/39 |
| ITEP | 103/279 | 49/119 | 55/85 |
| JINR-LCG2 | 209/628 | 156/554 | 157/280 |
| Kharkov-KIPT-LCG2 | 34/104 | 23/53 | 42/96 |
| LCG_KNU | 150/333 | 105/245 | 87/154 |
| LIP-Coimbra | 10/77 | 0/0 | 0/0 |
| LIP-Lisbon | 149/532 | 85/269 | 131/323 |
| MIT_CMS | 1037/1781 | 384/1196 | 724/1286 |
| NCG-INGRID-PT | 153/1833 | 0/0 | 0/0 |
| NCP-LCG2 | 4/9 | 0/0 | 0/0 |
| NDGF-T1 | 195/901 | 160/892 | 135/556 |
| Nebraska | 456/1410 | 224/637 | 436/692 |
| Purdue-RCAC | 1498/5117 | 1224/5064 | 696/2715 |
| Purdue-Steele | 227/539 | 241/539 | 94/215 |
| RRC-KI | 24/287 | 4/14 | 11/17 |
| RU-Protvino-IHEP | 42/144 | 29/79 | 43/81 |
| RWTH-Aachen | 1172/2813 | 744/2711 | 620/2213 |
| Ru-Troitsk-INR-LCG2 | 29/125 | 37/93 | 0/0 |
| SPRACE | 25/102 | 0/0 | 0/0 |
| T2_Estonia | 247/1012 | 161/1008 | 258/534 |
| TR-03-METU | 26/122 | 24/69 | 44/116 |
| TW-FTT | 282/1240 | 129/472 | 112/248 |
| UCSDT2 | 371/1696 | 121/363 | 152/323 |
| UFlorida-HPC | 601/2603 | 286/929 | 260/831 |
| UKI-LT2-Brunel | 166/474 | 111/469 | 131/313 |
| UKI-LT2-IC-HEP | 711/2312 | 427/1082 | 419/1432 |
| UKI-SOUTHGRID-BRIS-HEP | 26/98 | 6/15 | 12/26 |
| UKI-SOUTHGRID-RALPP | 495/1714 | 338/1648 | 446/962 |
| WARSAW-EGEE | 138/399 | 88/334 | 81/203 |
| ru-Moscow-SINP-LCG2 | 68/165 | 44/123 | 65/99 |
| ru-PNPI | 11/125 | 0/0 | 54/125 |
| ucsdt2-b | 424/1693 | 158/352 | 184/390 |
| Summary | 17151/26641 | 0/0 | 8729/18750 |

**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)**
  - **40 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 ~ 1/3 of slots**

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

# Resource Needs 2009/10

## Resource requirements compared to installed resources:

| | Installed in Jun09†† | Installed in Oct09†† | 2009 CMS Request | 2010 CMS Request |
|---|---|---|---|---|
| T0 CPU | | | 37,1 | 61,9 |
| CAF CPU | | | 7,6 | 34,7 |
| **CERN total CPU** | **29,6** | **44,0** | **44,7** | **100,6** |
| T0 disk | | | 0,4 | 1,0 |
| CAF disk | | | 1,3 | 3,1 |
| **CERN total disk** | **2,2** | **2,2** | **1,8** | **4,0** |
| T0 tape | | | 6,4 | 12,0 |
| CAF tape | | | 0,9 | 2,6 |
| **CERN total tape** | **10,0** | **10,0** | **7,3** | **14,6** |
| **T1 CPU** | **49,3** | **58,9** | **46,0** | **100,4** |
| **T1 disk** | **6,1** | **9,7** | **6,5** | **13,4** |
| **T1 tape** | **9,0** | **16,5** | **11,9** | **23,3** |
| **T2 CPU** | **111,8** | **141,8** | **75,2** | **195,0** |
| **T2 disk** | **6,0** | **9,2** | **3,7** | **9,2** |
| **†† reporting of Tier2 resources & pledges incomplete** | | | | |
| **Units: kHEPSPEC06, PB** | Units are kHepSpec06, PB | | | |

### '09 Relative Pledges:

| | Raw % |
|---|---|
| **FZK** | 12,25% |
| **IN2P3** | 12,50% |
| **PIC** | 7,25% |
| **CNAF** | 8,00% |
| **ASGC** | 10,50% |
| **RAL** | 10,00% |
| **FNAL** | 39,50% |

**Resources needed for "2009-run" are either already installed or on course to be installed by October 2009.**

**Reviewed by the LHCC and by the C-RSG (reviews combined with ATLAS) in April-July 2009, draft reports available.**

# 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:

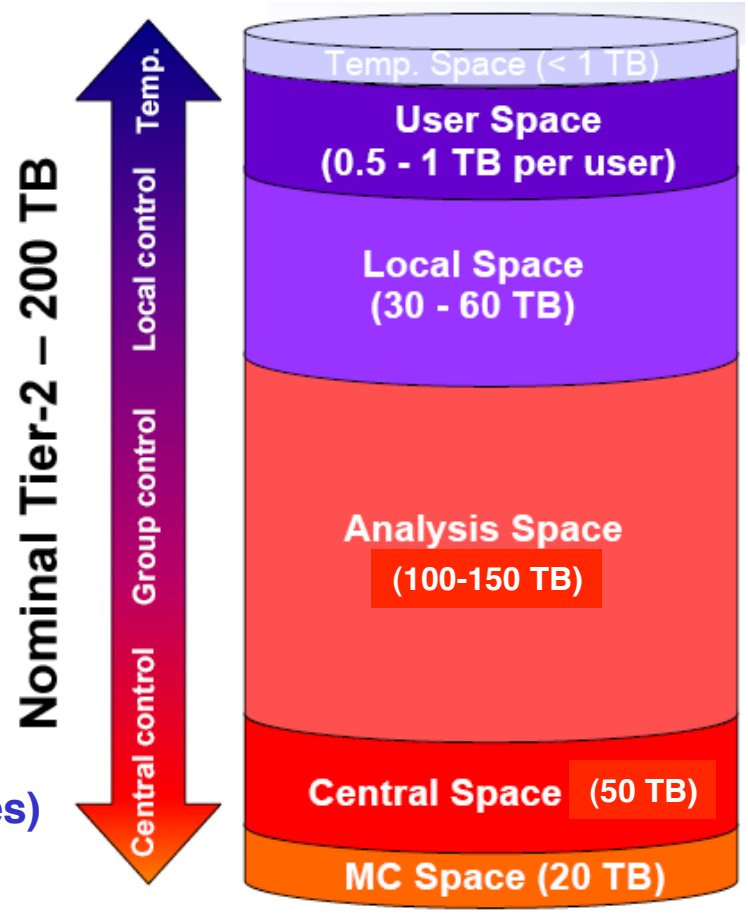
| CERN resources | 2009 | | 2010 | |
|---|---|---|---|---|
| | Scrutiny | CMS | Scrutiny | CMS |
| Tier0 CPU (kHS06) | 37.1 | 37.1 | 61.9 | 61.9 |
| CAF CPU (kHS06) | 7.6 | 7.6 | 34.7 | 34.7 |
| Tier0 disk (PB) | 0.4 | 0.4 | 1.0 | 1.0 |
| CAF disk (PB) | 1.3 | 1.3 | 3.1 | 3.1 |
| CERN tape (PB) | 7.4 | 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**

# Computing Support of October Physics Exercise

- **Data-Operations will produce "Secondary datasets" from MC**
  - **skimming on trigger quantities and pre-scales on the T1's**
  - **SD production schedule:**
    - 21.9. start testing, 25.9. start producing, 2.10. finish producing, 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 (regular computing shifts)**
  - **Active participation in October-Exercise-Meetings (weekly, daily)**

# ToDo

- **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**
  - **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**

# 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

# Conclusions

- **STEP'09 tests ran successfully in June.**
- **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…"*
- **CRAFT09 data successfully recorded and distributed in August.**
- **A large MC sample (500 M FullSim events) was produced at higher than normal rate in August/September.**
- **Computing is ready for LHC data.**