CMS
Service Challenge 4 and beyond

CMS SC4 targets/milestones

- CMS has set targets for SC4 since beginning of 2006
  - Communicated them to all our sites
  - Posted on Twiki, Web, Agenda's, mails…
- Aimed to smoothly roll in a WLCG service that CMS can use in October 2006 to test computing flow at 25% of 2008 scale (CSA06)
- Establish early in the year baseline of non-demaning continous file transfers (20MB/sec per site disk-only). Avoid SC3 syndrome.
- Targets for June
  - Have all CMS Tier1 and most Tier2 on board
  - Demonstrate data transfers at realistic rates (tape at T1)
    - 150MB/sec out of CERN total (600MB/sec in 2008)
  - Ramp up job submission rate over WLCG to 25K jobs/day
    - 90% submission efficiency
    - double to 50K jobs/day by October (200K target in 2008)
  - 1 MB/sec/execution-slot from disk to CPU

Spring summary

- Could not establish working baseline in the spring
  - in spite of start of activiyt in early April
  - srmCopy, srm push/pull, Castor2, WLCG tests, FTS, gLite 3.0 …
- Used the time to get ready for SC4:
  - FTS integrated in PhEDEx
  - Develop job submission/monitoring tools
  - Integrate in new CMS Software framework
- Started SC4 at beginning of June

June summary

- Steady progress, but not there yet

- Finally have all Tier1's and 20 Tier2's (including perspective ones) on board: run one CMS simulation job, transfer 100GB of CMS data, run one CMS analysis job on local data
- Transfers: a few channels work, mostly below target rate
  - much better on OSG side.
  - Ad hoc effort started in coordination with WLCG, daily meetings, several CMS persons involved
- Job submission rate: reached what seems to be LCG RB scale limit (5K jobs/day/RB)
- 90% job success
  - OK for WMS (OSG and EGEE)
  - OK for single sites at certain times
  - Not achieved yet on all sites for the same day (not even all T1's)
- Not started measuring disk/CPU throughput

Transfers

Scaling Tape Rates by pledge aiming for 150MB/s

- ASGC: 10MB/s to tape (Currently averaging 4MB/s)
- CNAF: 25MB/s to tape (Currently averaging 20MB/s)
- FNAL: 50MB/s to tape (As high as 50MB/s but currently averaging 5)
- GridKa: 20MB/s to tape (Currently averaging 1MB/s)
- IN2P3: 25MB/s to tape (No current transfers)
- PIC:20MB/s to tape (Sustaining 10MB/s)
- RAL: 10MB/s to tape (Averaging 10MB/s)

A number of items progressing in storage

- WLCG is working to commission and debug channels
  - ATLAS is having general better performance trying to understand
- Daniele and Lassi will begin systematically working with T1 on July 5th
- A much larger load test sample will shortly be available.

Job submission rates

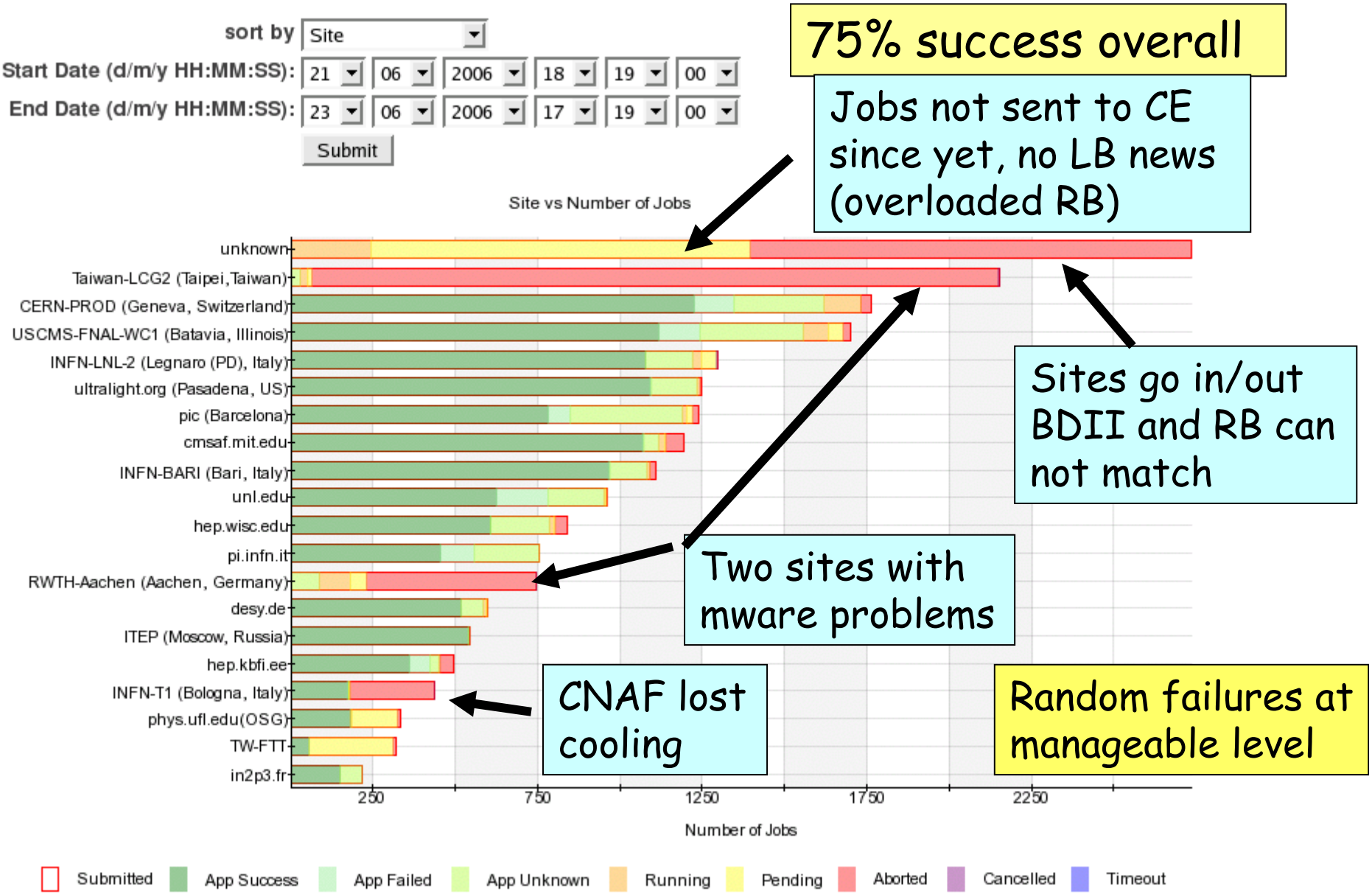
- Imperial College WMS monitoring list jobs by RB each day for all VO
  - http://gridportal.hep.ph.ic.ac.uk/rtm/reports.html

Jobs Submitted by RB

Resource Broker	Submitted	Success	% Success (% Resub)	Efficiency (%)	Registration Time	Match Time	DNs using	UIs using	CEs used
gdrb01.cern.ch	1431	414	28 (14)	59	19	19	23	24	102
gdrb02.cern.ch	2438	2088	85 (1)	32	20	85	9	7	201
gdrb03.cern.ch	3833	3065	79 (1)	44	10	13	5	7	120
gdrb04.cern.ch	2061	17	0 (0)	69	6	13	9	4	71
gdrb06.cern.ch	3372	2651	78 (0)	62	120	37	18	22	48
gdrb07.cern.ch	3955	1802	45 (0)	51	8	13	11	8	93
gdrb08.cern.ch	532	351	65 (0)	71	12	24	19	17	31
gdrb09.cern.ch	94	87	92 (16)	61	34	23	3	5	39
gdrb10.cern.ch	74	46	62 (8)	46	6	21	2	2	36
gdrb11.cern.ch	1948	1650	84 (0)	71	10	16	5	4	21
lcgrb01.gridpp.rl.ac.uk	3099	1982	63 (1)	50	23	11	17	13	135
gfe01.hep.ph.ic.ac.uk	189	23	12 (30)	24	101	29	4	7	17
gridit-rb-01.cnaf.infn.it	291	209	71 (3)	68	5	5	5	3	13
egee-rb-01.cnaf.infn.it	1125	376	33 (2)	77	6	12	15	10	55
egee-rb-02.cnaf.infn.it	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
egee-rb-03.cnaf.infn.it	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
egee-rb-07.cnaf.infn.it	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
prod-rb-01.pd.infn.it	55	50	90 (10)	41	5	14	5	4	24

- Can hardly find an RB with more then 5K jobs/day

Job success rate by site: two typical days



Sites issues

- We find sites in general responsive to CMS needs and problems
  - Thank you !!!

- gLite 3.0 not really there yet as a production service
  - FTS servers/channels not configured/exercised using CMS end points. Data transfer is an end-to-end activity.
  - Hints of configuration problems in gLite 3.0 LCG flavor CE
    - All CMS jobs fail with Maradona error on selected CE's.
      - other VO OK, SFT all green
      - a couple sites kept failing all jobs since > 2weeks
      - a 3rd site apparently went back to LCG 2_7_0 CE after a few days
- Not using the same infrastructure as WLCG throughput tests yet
  - E.g. Castor1 instead of Castor2 at Tier1's
    - Could not get a clean schedule for Castor2 avaibility for CMS
  - Want to be on final production infrastructure asap
    - Not be connected to something that in your opinion will not work

Summer plans

- Keep up basic SC4 activity:
  - File transfers up to target, then stay there
  - Job submission at ~12Kjob/day now, 10Kj/d demonstrated on OSG using Condor-G earlier on
- Add more complexity
  - Simulate 1M event/day in July/August to prepare for CSA06
  - Commission gLite WMS (RB, maybe also CE) to reach 50K jobs/day without 10 RB's
  - Add calibration/conditions data remote access
  - Test data serving throughput at sites (disk → CPU)
  - Allocate CPU resources separately for production and analysis jobs. Grid should not be a global FIFO queue
- It will be a busy summer

After the Summer

- CSA06 in October 2006
- From CERN disk to Tier0 to Tier1 and Tier2
- Demonstrating reconstration, analysis, calibration, reprocessing
- At 25% of 2008 scale
  - 35-40 Hz
  - Over a month
  - Would like to try higher througput

- Well captured in Harry Renshall's twiki
  - https://twiki.cern.ch/twiki/bin/view/LCG/SC4ExperimentPlans

- Ramp up to 2008 pledges
  - We caution sites against planning for a large increase in capacity in a short time. Every x2 is a challenge
  - True for Tier1 and Tier2
  - We look forward to stress test sites as soon as resources are available