



Computing Commissioning Plan for 2007

*Stefano Belforte (East of the Atlantic)
TBN (West of the Atlantic)*



From Computing Project Plan



- Ensure CMS builds a fully functional computing systems
- Establish Computing- and Operations services at all participating Tier-0, Tier-1 and Tier-2 sites.
- Integrate CMS data and workload management systems with EGEE / OSG middleware
- Prepare for analysis, production, service/data challenges including regional centres



- Above list re-interpreted as action items
- Derive a plan for things to do in the short term and involved personnel
- A first iteration at the plan was done within the Computing Project
 - Responsibilities and scope better defined
 - Tier0 as a site: internals not in this project
 - take off from data injected into PhEDEx
- Commissioning means to hand operations a working system
 - Commissioning activities have large overlap with operations
- Operation personnel, esp. at sites, is naturally involved in commissioning
 - Beta test
 - Stress test
 - Continuous load
 - While of course supporting data operations

- So you will naturally participate in commissioning activities



Action items derived



- Make sure it all works: tie ends, watch for holes
- Exercise computing system (make sure is fully loaded at all times)
- Setup the tools and structure for facilities operations
 - tools to verify correct operations of the Tier's
 - ☞ CMS specific tests in SAM (aka Site Functional Tests)
 - Tools to find how what is working
 - ☞ e.g. job Dashboard, PhEDEx heartbeat
- Coordinate with OSG/EGEE operations, support and development
- Carry on tests of relevant middleware and work to address problems
- Organize work program of computing exercises targeted to make us ready for more integrated CSA-like challenges
 - Intersite data transfers
 - Tape usage at Tier1
 - Job submissions / executions
 - Etc.



- Contact with SAM developers/runners established via LCG EIS support
- Establish a "site OK for CMS" criteria, start from CSA06 checklist
- Preliminary list of what to add prepared
 - Submission of a CMS (not dteam) job
 - Verify SW installation, TFC
 - Verify local stageout
 - Verify Squid
 - Verify local SRM w/o depending on other sites
 - OK for production
 - OK for analysis
 - Test transfer to remote site ?
 - Verify published information to BDII
- Meant to be a service from "central commissioning team" to sites

- **Re-establish constant load on the system in SC4 spirit**
 - Higher rate targets
 - At the same time as production
 - ☞ deploy CPU resource sharing (aka job priorities)
 - ☞ separate PhEDEx instances, then prioritization by dataset (2.5)
 - With slight different focus
 - ☞ load disks and squids at sites
 - ☞ decouple site tests from job submission metrics
 - ☞ load site-to-site channels accordingly to target need
 - ☞ debug/commission data transfer channels following CMS data flow
- PhEDEx developers and Central Operation Team have developed tools to make load transfers easier, testing by voluntary sites to start shortly, deploy to all sites with PhEDEx 2.5. Work program being defined
- For job load will need one operator (from Central Team ?). Start with job robot with aim to transit to PA or CRAB Analysis.Server



Next transfer tests



- Over the course of February establish constant (high) load on T0 → T1 and T1 → T1
- Add T1→T2 as sensible
- New load generator makes it easy to test burstiness of traffic outgoing from Tier1
- Test will be driven by sites
 - Site operation people do the work this time
 - PhEDEx 2.5 lets sites (un)subscribe to dataset blocks
- With need some central help from PhEDEx developers and Operations

- Propose site operation people do the actual work, coordinated by Computing Facilities Operation
- Daniele will give details



Next job robot tests



- Validate and exercise analysis patterns at sites if users are not enough
- Currently job robot should be configured to have same priority as a normal user
 - In presence of real analysis load should backoff kindly and gently
 - More tuning may be possible
- Test system capability to accept and maintain full load
 - many short jobs will exercise WMS infrastructure
- Large scale test of Squid functionality
- Exercise data serving capability at sites



- New storage managers have functionalities
 - that we may want to use
 - that they want us to use
- SRM v2.2
 - Storage classes definitions for sites
 - Space token / Storage Classes usage in FTS/PhEDEx
 - Staging, pinning, unpinning, permissions
- gLite 3.x
 - New WMS
 - New gLite CE
 - Better error reporting in UI (important for dashboard)
- OSG
 - Condor status to dashboard
 - Stress test of current and future OSG CE's
 - Stress test of dCache



Running production and analysis at same time



- CPU share for production
- We asked every site to setup a 50% share for production
- Most have done it and filled the twiki accordingly
 - Thanks
 - Any site having technical difficulty ?
- We have to verify that this works in practice
 - Restart job robot (MC production already started)

- Please add priority for SW installation
 - need few hours latency
- Will not ask for more granularity this year

- New dynamic information provider
 - Publish information for each LRMS share
 - tell that CE is full of analysis jobs, not production
 - WMS can schedule production there (and viceversa)



- From WLCG Quarter 1
 - Tier0-Tier1 data export at **65%** of 2008 rates (to tape !)
 - Tier0-Tier1 data export at **50%** of 2008 rate in conjunction with T1-T1 / T1-T2 transfers
 - Tier0-Tier1 data export at **35%** of 2008 rate in conjunction with T1-T1 / T1-T2 transfers *and* reprocessing / analysis stripping at Tier1s
 - Provide SRM v2.2 endpoint(s)
- WLCG Quarter 2: do the same with SRM v2.2
- CMS milestones guide the way for the year
 - P5→T0→T1 starting in May
 - CSA07 in the summer
 - MTCC-3 in October
 - Pilot run in November



Timeline



- February
 - Deploy PhEDEx 2.5
 - T0-T1, T1-T1, T1-T2 independent transfers
 - Restart job robot
 - Start work on SAM
 - FTS full deployment
- March
 - SRM v2.2 tests start
 - T0-T1(tape)-T2 coupled transfers (same data)
 - Measure data serving at sites (esp. T1)
 - Production/analysis share at sites verified
- April
 - Repeat transfer tests with SRM v2.2, FTS v2
 - Scale up job load
 - gLite WMS test completed (synch. with Atlas)
- May
 - Start ramping up to CSA07
- June



- Details will be clarified and explained during Thursday's Integration/Commissioning/Operations meeting
- Usual time: 4pm GVA

- Today most sites are here "in flesh"
 - Any early questions from you ?



Questions for you



- Transfer tests/operations :
 - Do you want heartbeat ?
 - Do you want 20MB/load test matrix like monitor ?
 - Will only restart those if useful to sites

- Monitoring
 - Give feedback on dashboard
 - How can we get better handle on efficiency

- Data serving
 - A metric that "escaped attention" in 2006
 - 1MB/sec per execution slot disk to CPU
 - ☞ 200MB minimum for T2, 1GB for T1
 - How can we work to test this w/o affecting production
 - Piggy back on Merge jobs ?