T0 post mortem.

Service runs generally smooth. (few issues with SRM interference, garbage collection in CASTOR for CMS). Alert mailing list used the first time (ATLAS). CMS has been the main VO putting load on CASTOR disk cache. However a throughput of about 14 GB/s sustained for some (short) period by CMS T1D0.

Some incidents experienced with CASTOR. The incidents had not big impact however (just degradation of the service).

Experience of the alert system…some issues with the procedure it self (server side). It should be revisited. The operator should call the service expert on stand-by and also a contact on the experiment should be provided for asking more information. Reliability is still issue ~10 incidents.

Shared DB serving all VOs…during CCRC migrated Atlas (going to be the same for the other VOs in the next days)

How to improve things:
- Support for multiple stagers with new CASTOR2.1.7 (making it more reliable)
- redundant backends (for transparent interventions)
- Improve the monitoring.

Tape: writing efficiency continues to improve….bigger files however tiny files is still issue. High read activity from tape (data volume per mount is still too low)
Tape service passed CCRC but more load is expected for the future…
Operational issue on the tape system ran w/o the service being disrupted.

Production users higher priority than normal users…(optimization: avoid multiple mounts for reading files on the same tape)

Monitoring improvements..
Message form that monitoring: T0Atlas is working well.

Monitoring regarding tape mounts. Max number of times a tape is mounted per day (average is actually high about 60) just for reading…this is heating us (CASTOR) in terms of performances….we know who's mounting…however user does not usually know that his running program opens files and mounts tapes.

Power cut: what happened? Communication with control center has bee good and also the recovery fairly good. We have experienced some problems with DNS and with the time server. This happened because incorrect power connection and problem with computing center operations and physics database.
We envisage every 3 months live tests…

Batch farm:
No major problems
2.4 M jobs (490K from the grid) up to 80K jobs per day.
Anticipating the expected load few measures has been early adopted…
Dedicated T0 activity, new h/w resources for CE and WN and a new version of LSF (fixing the double logging problem before CCRC), publishing CPUs instead of cores. (Discussion on the reason of that…. Steve suggested that)
The number of pending jobs dropped. We are draining the queues faster than before because more powerful machines. Fluctuation in the number of jobs observed. Is this production running level (same question for the DMS part of the talk)….if yes we're ready to sustain it. 10K jobs were affected by the power cut. LFC service: smooth operations: two bugs found. WMS smooth operations with few problems. VOBOX kernel upgrade. This has triggered a discussion mainly with ATLAS people about a problem with version of the Kernel. PPS is there for that. ATLAS confirms that this is not possible because you should have a dedicated machine on the PPS just for running Atlas (DDM) service. CMS: is the LFS issue experienced in the past fixed for sure? No we are not sure the problem is definitely fixed.Not easy to reproduce (test on the same conditions) being the problem from Grid jobs and not easy to reproduce exactly.Try to simulate,.

ATLAS Presentation:

Phase 1 (Feb) mainly a SRMv2 test (installation and configuration) …FDR was running concurrently…
Phase 2 (May) was instead carried along all the month, not overlap with FDR though detector was taking cosmic (higher precedence).

Focus on the test of the overall data distribution chain:
Metric have been set (and were very demanding). We required more than what we need for data taking in 2008.

Load generator agent used to emulate the PIT/ONLINE (fake) data generation.
Double registration problem is the main problem. File with double entry in LFC (same GUID, DDM bug).

Week 1: warm up period, data subscribed to T1 tape and disk endpoints accordingly the share for Computing Model. Not T2 activity for this week. Breakdown per all sites: few

problem at CNAF/NDGF and SARA preventing to have 100% dataset complete. Metrics were relaxed however.

Week 2: full test T1-T1 matrix. Every single channel has been extensively tested. It was agreed on testing concurrently with CMS.
18TB were to be replicated to each T1. Subscribe in one go: aim to check whether the system is able to throttle or it collapses. In 2 days and half 90% should be transferred to the T1. → 90MB/s import sustained rate to each T1.Error rate is quite high. Very aggressive target.

INFN being slow importing data
SARA being slow exporting data.

 Very good: ASGC, PIC (peak of 500MB/s) and NDGF. FTS global tuning of parameters is needed at least (RAL too conservative ASCG very aggressive)
ATLAS does not want internal FTS retries.
Channel <everysite>-NIKHEF must be set everywhere.
FTM at al T1 and FTS logfiles must be exposed.

Week3 throughput test. push the system to its limit.
Peak: 24/hday nominal rate is14hours/day.
No oversubscription share accordingly computing model. Question on the expected rate per site. Overall efficiency fairly low (successful transfers over total transfers)..site by site main issues described.

Week 4: full exercise test of all data movement chain.T0-T1 throughput is quite promising. No problem are envisage (largely beyond the most demanding metric)
Even more: double registration (DDM problem) is proving the system can give even more job than required.

Week 4 and beyond: production. Despite production competition, power cut and double registration (three unexpected external variables) they haven't been prevented to meet the metric.

1PB deleted in less than a day (nice plots from SLS showing the space occupancy as function of time).

DPM main issue (Michel) : balance of the file systems.

Network issue mainly BNLtoCERN weren't managing to run upto 100MB/s.
 Iperf server at CERN should be permanently installed (for catching problems)?

Space tokens disappearing: sometimes mistake (NDGF) some time bug (SARA) sometime reconfiguration problem. Not operational issue.

FTS: SRM negotiation makes extremely inefficient small files transfer.

CASTOR studies point out about 5 second SRM negotiation…need to merge these evaluations….

Need to setup 2 FTSes CERN-CNAF_disk CERN-CNAF-tape. Very painful managing 2 FTSes. (in the same path two different endpoints the lsower one might penalize the faster one that can support less SRM hits than the oteher).

Atlas is impressed with responsiveness of site admins.

Missing reprocessing activity no analysis

LHCb talk

 Computing Model of LHCb sketched. First pass recons at T1 (out of CERN) to reduce the amount of interesting data to be analyzed.
Planned task for CCRC are the one envisaged for data taking.

Maintain1 month data taking with 50% machine cycle efficiency.
Not fully stressed analysis though some small activity has carried out.
FAILOVER used at CERN. MC space tokens not used. Highly depend on sites for the success of the exercise (IN2p3 and SARA alone 51%)
Pit-T0 1.6 G every 30 seconds rate on castor 70MB/s. ONLINE storage issues.
Export to T1 follows closely the migration to tape. Issue of sending BK information.
Because no checksum failures were verified at the first phase we switched to do checksum after the transfer to T1 but this was problematic in the internal logic.

Transfer rate to T1s meets the target anticipated. Peaks visible because of backlog cumulated.
T0-T1 some problems. The empty vertical band is a problem generic affecting central services. Overall transfer efficiency at the first attempt is OK.

Reconstruction: preserved data share accordingly pledged resources.
Lesson from first phase: queue length issue → reduced the number of events per recons job. High luminosity (more large files were causing stalled jobs in Feb phase)
Recon's job automatically created after the file is transferred; a check on status of a file (is ONLINE?) then job submitted. Number of issues at SARA (file status reported incorrectly) and similarly a problem at IN2P3 with gfal_ls.

Nick showed a breakdown of the recons activity across sites
41.2 K recons created accordingly our metrics was also fine…the percentage of done jobs is not as good as expected however for the two most important sites (SARA and IN2P3 that got the largest share).

Not all done jobs processed their 25K events (77% of Done). Out of these 13% went through the failover mechanism. 25K events jobs (21.2K is the real number of successful jobs). Site by site analysis of the behavior…

CERN once data available 99% efficiency. Limited to 300. LFC issue? No clear it is LFC service or client. Under investigation. Proposed to increase the threads per LFC.
Also GridKA few issue but 99% efficiency…proved to be able to recovered huge back log very efficiently (up to 2K recons jobs running concurrently)
CNAF has been also very efficient (if we opportunely interpret jobs and inefficiencies).
IN2P3 initially efficient then problems…upgraded to latest (not yet in production version of gfal 1.10.11) because the old one was core-dumping .

NIKHEF 4% success rate.
Initially was OK only a problem uploading rdst data there. dCache upgrade and the nightmare started.
If the file is not immediately accessible it is NEARLINE. Fundamental problem there for running in some other mode (ex. copying to WN first and then open locally)→Abandon the stager approach too inefficient.

PIC 99%: just a problem with one WN.

RAL rfio problem not fully understood causing a efficiency of about 68% efficiency observed. Fall back solution of copying data locally on the WN. At RAL we can copy the file at IN2p3 often even form the local SE we could copy the file into the WN. \
Question about rootd at RAL raised.

We know about the issue in RFIO that once times out in rfio connection cannot be recovered. This has been fixed by rootd. Asking officially to WLCG to support rootd. This is supported by root people.

Wallclock/CPU site breakdown. General degradation because of downloading data to the WN sometimes takes longer w/o consuming CPU. Second timeout on lcg_utils to kill the command if stuck? This was in place in the custom distribution of lcg_utils for LHCb 2 years ago.

dCache sites breakdown:
Guess on the security of dcap: dcap is OK gsidcap introduces some issue? Not an issue for us to open through the site LAN in an unsecure way apart from NL-T1 where they are firewalled. At some sites however the unsecure dcap is forbidden.
 LHCB reminds that file access protocol are only for reading and not for writing and we do not care if someone else can read (and even more surprisingly understand our data). Others non LHC VO however do mind someone else can read data. We have a order according which we pick up the protocol to use: first dcap, then gsidcap then root rfio file:….
Is this security a site policy? We move data through lcg_utils.
Used Condition Databases successfully by recons jobs…

The LFC streams-distributed service has been only partially used.

Stripping has been mainly affected by a BK issue preventing to send ancestor information. The small fraction of jobs that had chance to run went OK at CNAF CERN PIC RAL GRIDKA with very high efficiency.

Despite jobs run OK they just produced very few data that has been replicated smoothly (reduction factor too important because a bug on the application). The mechanism in place has been however exercised at least and not major problems have been spotted. Dirac3 tools improved significantly since Feb.

Sites want to know what we are doing: where do they can find information about our activities and how they are going? This is a fundamental question that the Julia working group will solve with LHCB help. LFC people will increase the number of the threads on Monday. Number of threads must be configurable however. Monitoring of LFC should also be welcome.
Rootd xrootd…does LHCb ask all sites or only some sites…
We only asked for one service class at CERN and one at SARA (already in place) and we have started testing there; rootd protocol has been refused at RAL at some point on time. Philippe's remarked on xrootd implementation on CASTOR that requires 2 boxes per each service class. Very demanding!
Dcache has xrood integrated into the system and does not require 2 boxes per pool as in CASTOR. If WLCG support xrootd development must be envisaged. Patrick: we have to know whether xrootd is going to be used widely