



# Data Management Services

## Experience from 2008 - Sites & Services

WLCG 2009 Data-Taking Readiness Planning Workshop  
13 November 2008

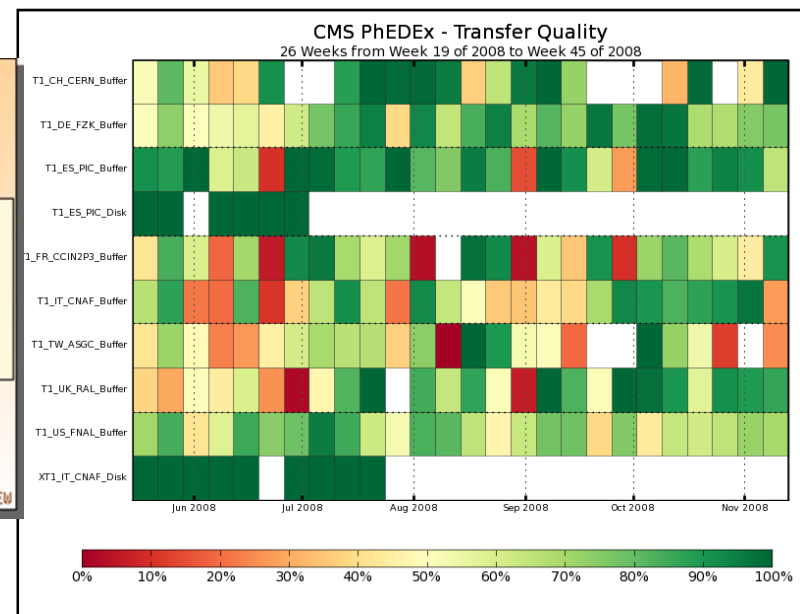
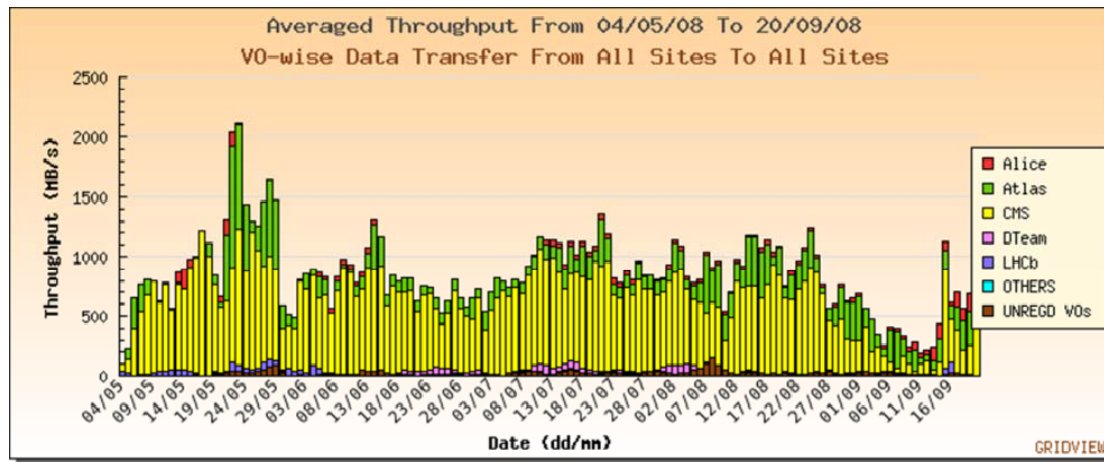
Gavin McCance  
CERN IT/FIO

# Outline

- Scope:
  - Review of last 5-6 months, post CCRC'08
  - Looking at the core WLCG (non-experiment) data management services (LFC, FTS, Storage)
  - Focus is mostly on T0 / T1, some comments on T2
- Based on comments from sites and experiments
  - IN2P3, GRIF, NDGF, PIC, SARA/NIKHEF, RAL, CERN, US-CMS
  - ATLAS, LHCb, CMS, ALICE
- What was achieved
- Current status of services
  - Review of incidents
  - Summary of issues
- General comments about data management
- Summary

# What was achieved

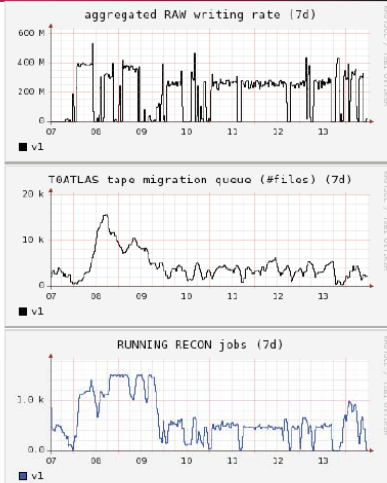
- Routine productions working! 😊
  - Continuous cosmics for several months! Some real data...
- Data distribution is working! 😊
  - Machine data (RAW, recon) and Monte Carlo data
- SRM infrastructure is mostly set up 😊
  - Storage spaces, space tokens
- Data generally available once it gets there 😊





## Tier-0 and data-taking activities

- We are taking continuously cosmic ray data since several months and until 3rd November
  - With only short breaks for detector work (and LHC data)
- The Tier-0 is coping well with nominal data rates and processing tasks
  - A few Castor glitches are usually sorted out with the Castor team within a very reasonable time
- In November hardware detector commissioning work will restart
  - But cosmic data-taking will carry on with partial read-out



Dario Barberis: ATLAS Computing

4



## CMS Computing Summary

- **CSA08 and CCRC08** Demonstrated all key performances of the T0, CAF, T1 T2 infrastructure
- **During the summer we**
  - Improved infrastructure reliability, production tools (Tier-0,...), monitoring and operations
  - started Computing and Offline Run Coordination and Computing shifts
- **Routine Cosmics and Commissioning Data taking performed over the summer**
  - Processed at T0, Calibration&Alignment performed, distributed on demand to several T1/T2 centers
- **Production of requested Monte Carlo samples performed routinely**
  - Huge production of MC Startup Sample (>200M) started when final software and configuration became available

### Computing tasks for the coming months:

- consolidate operations,
- commission Tier-2 sites
- roll-out improved production tools,
- work on the monitoring and fault detection.
- Configure and get experience with the CAF-T2 resources
- Global running and Cosmics data dating will allow for more systematic checks of the whole production and transfer chain.
- The Analysis of Data and Monte Carlo will move to Tier-2 centers.

WLCG MB: September 30, 2008

CMS Status Report

M.Kasemann

15



## LHCb recent activities

- **Commissioning of DIRAC3**
  - **Fully reengineered system**
  - **Main features:**
    - Single framework for services, clients and agents
    - Fully integrated Workload and Data Management Systems
    - Supports production and user analysis activities
      - ⊕ Allow to apply VO policy: priorities, quotas...
    - Uses pilot jobs as DIRAC2
      - ⊕ Ready for using generic pilot jobs (not switched on yet)
      - ⊕ Full scale test with generic pilots will take place in the coming weeks
    - New bookkeeping system (also integrated)
- **Production activities**
  - **Complete simulation and stripping of MC data (so-called DC06 as was launched in 2006)**
  - **CCRC-like activity at low rate (10%)**
  - **Start 2008 simulation**
    - Mainly for alignment and calibration studies
    - Wait for first data for tuning generators and detector response



LHCb QR, October 2008

PhC

15

## ALICE

- **Monte Carlo**
  - Continuous production in prevision of 2009 data taking: production in T1/T2, data and end user analysis in T2
- **Analysis**
  - CAF (fast), analysis train (organized) and end user analysis (chaotic) operational
- **Raw data**
  - Online production of condition parameters, first pass processing @ T0, replication in T1s, N pass processing @ T1s operational
- **Software**
  - Stable release ready for data taking; code evaluation and some refactoring to be done before LHC start...
- **Services**
  - New AliEn version deployed routinely with effectively no downtime
  - Job management in all its form: RB (phased out but still widely used), WMS, CREAM (very promising initial stability and scaling tests)
  - xrootd-enabled SE continuous deployment (T2s are a concern)
- **Accounting**
  - Used 40% of allocated CPU and 53% of required.
  - 27% of pledged storage is operational and 64% of that is used.
- **Resources**
  - Requirements for 2008/2009 had been reevaluated (before LHC incident)
  - New requirements for 2009, depending on LHC schedule; expect larger requirements with respect to C-TDR (CPU, disk)

26 August 2007

ALICE 2Q2008 report

17

# LFC – LHC file catalogue

- Deployed for 3 LHC VOs + 1 shared for other VOs
  - CMS: central (not used), Other VOs: central, shared
  - LHCb: central and replicated read-only, using Oracle Streams
  - ATLAS: 1 distinct LFC per regional cloud (hosted at T1s)
- Releases during period
  - 1.6.9 / 10 unstable (server crash in security layer)
    - IPv6 support, bugfixes
  - 1.6.11 Bugfix of above
  - 1.7.0 in certification
    - Client for Python 2.5 supported, SL5 support
- Current deployment
  - Most T1 sites @ 1.6.11
  - CERN conservatively running 1.6.8
- Work in progress
  - Web-service front-end (requested by US Atlas)
  - A couple of pending requests from LHCb

# LFC

- Service level generally stable
  - Extremely critical for experiment operations (ATLAS, LHCb)
- Very few incidents / service degradations:
  - 1.6.9/10 instabilities: July/August: Various sites
  - Streams aborted causing delayed updates (LHCb):
    - FZK 31/7 – 6/8, SARA 30/9, RAL 7/10
    - Central LHCb catalogue was not affected
- Bulk interfaces appear to be meeting performance needs
  - ATLAS reported an issue, but it was found to be elsewhere in the stack

# LFC

- Operational pain points reported:
  - LFC cleanup operations (consistency)
    - In case where files are lost in storage – bulk deletion by VO has to be done – manual procedure for
      - Experiment: cleanup LFC, cleanup internal catalogues
      - Site: cleanup storage namespace
    - Understood that we don't do auto-synchronisation between the various catalogues (not just an LFC issue!)
  - LFC manual DB operations
    - Bulk rename of storage elements is a significant man-power intensive operation (requiring downtime and direct DB access)
      - particularly for the LHCb streams-replicated central catalogue (coordinated distributed intervention)

# LFC: where are we?

## LFC

- LFC is the mo clouds
- Resolves LF
- Without the L
- Data transfe
- Production
- User analysi
- LFC failure is

## Conclusions

- LFC is a vital service for ATLAS in each cloud
- Replication of read-only LFCs using Oracle Streams is possible
  - But read-only replicas are not useful enough and using them will require changes to many pieces of DDM and client code
    - *Complexity is the enemy of reliability*
- Data Guard is a promising technology and hides almost all the complexity of maintaining replicas
  - However, it is money and time to deploy
  - Need to balance benefits against the considerable costs
- For the moment Tier-1s should concentrate on making the LFC a highly available service for ATLAS with plans to cope quickly with known issues which would affect availability

## The Problem?

ATLAS Errors	Other Errors	Batch System Catalog Errors	Site Errors (SE/Batch)	Total Errors
3.8	2.8	7.5	0.1	10.9
7.3	0.9	19.5	0.0	25.4
7.2	6.8	6.1	0.2	22.8
3.5	0.8	7.7	0.0	19.1
10.6	2.3	10.4	0.0	15.9
5.3	2.3	9.1	0.0	20.7
3.6	3.3	5.1	0.0	8.0
7.7	4.0	4.8	0.1	7.1
26.9	3.6	9.2	0.0	20.9
15.2	5.7	1.4	0.5	4.3
14.4	3.7	3.7	0.1	11.1
0.4	1.1	0.8	0.0	59.0
3.5	2.9	22.5	0.0	23.2
6.0	4.2	1.6	0.1	1.0
4.5	3.4	8.7	0.1	16.2
48.7	1.3	2.6	0.0	2.0
14.7	2.4	5.6	0.0	11.9
2.4	1.0	3.5	0.2	0.2
4.5	4.3	9.5	0.0	15.3
4.2	3.0	7.1	0.1	11.7
8.0	3.2	6.3	0.1	15.2

on errors - not so bad in

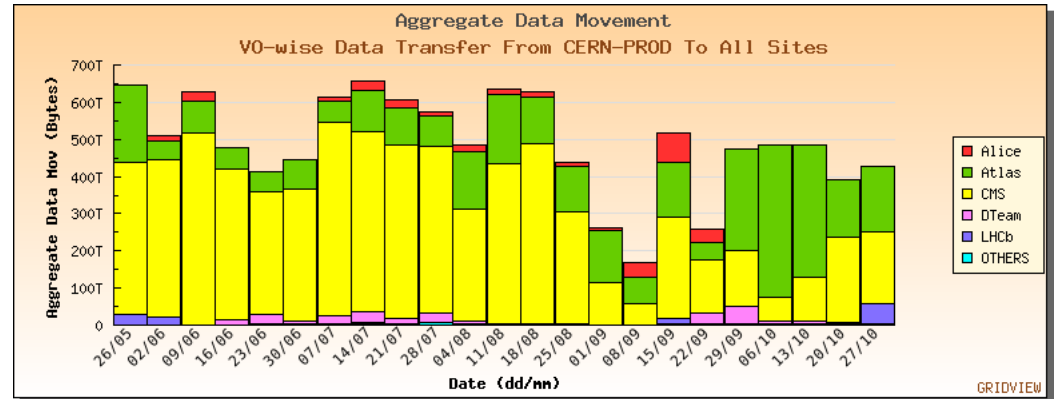
- Now at the level of ~0.1% of production errors for ATLAS but high impact when it happens
- Want to protect against cloud LFC / DB failure
  - DB replication / failover
  - ~Similar to LHCb central catalog replication using Streams
- Given the MoU, is it worth the cost for 0.1%?
  - Maybe better to focus on ensuring T1 HA



## FTS - File Transfer Service

- FTS 2.0 (SLC3) deployed at T1s and CERN
- Work on gLite 3.1 / SLC4 release: FTS 2.1
  - Much awaited...
  - Tested with over 750 TB of transfers on PPS to many T1 sites – multiple issues found and resolved in PPS
    - Error reporting problem (error message format change in API) noted on 1<sup>st</sup> day in production
      - Review PPS strategy?
  - Now deployed in production at CERN (in parallel to SLC3)
    - Requesting experiments migrate asap
    - Roll out to T1 sites
- Fixes a variety of niggling issues
  - Source space tokens can now be specified
  - Better robustness, more monitoring information

# FTS – service incidents



- Service level generally OK
  - Works without too much operator effort
  - A few 'lowlights' and a number of issues remain
- Channel agent resilience against node failure
  - 24 Oct CERN-PROD FTS outage (3-4 hours on some channels): <https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortemFts24Oct08>
  - Impact: outage of any channels hosted on that node, Risk: low
    - The recovery procedure is manual
    - N.B. Load-balanced web-services are resilient to node failure
- 29 Jul: Oracle DB issue seen at CERN (a 'standard' schema copy of the production schema corrupted a sequence):
  - <https://twiki.cern.ch/twiki/bin/view/FIOgroup/FtsPostMortemJul29>
  - Has not been reproduced

# FTS – service incidents/issues

- [logical] DB corruption (TRIUMF 26 Oct) – in case of agent type swap (srmcopy <-> urlcopy)
  - Impact: Channel outage until DB intervention, Risk: ~low
  - The 'correct' swap procedure is fairly complex
    - Fixed in FTS 2.1
- Issue: Delegation race-condition in web-service still not fixed
  - Potential impact: outage for given VO for ~hours, Risk: low
  - 'Worked-around' by experiments mostly
    - ...but still the cause of the odd outage / degradation (noted on CERN-PROD / FZK )
- Issue: DB 'history' cleanup 'tool' still needs to be integrated into core FTS
  - affected FZK, SARA
  - Potential impact: without the cleanup tool, everything gracefully slows down to a halt over a period of months, Risk: medium
- Issue: better handling of timeouts on large files [10GB+](CMS)
  - Study use of transfer marker timeouts to make this work better
  - Big files are handled 'manually'

# FTS – operations

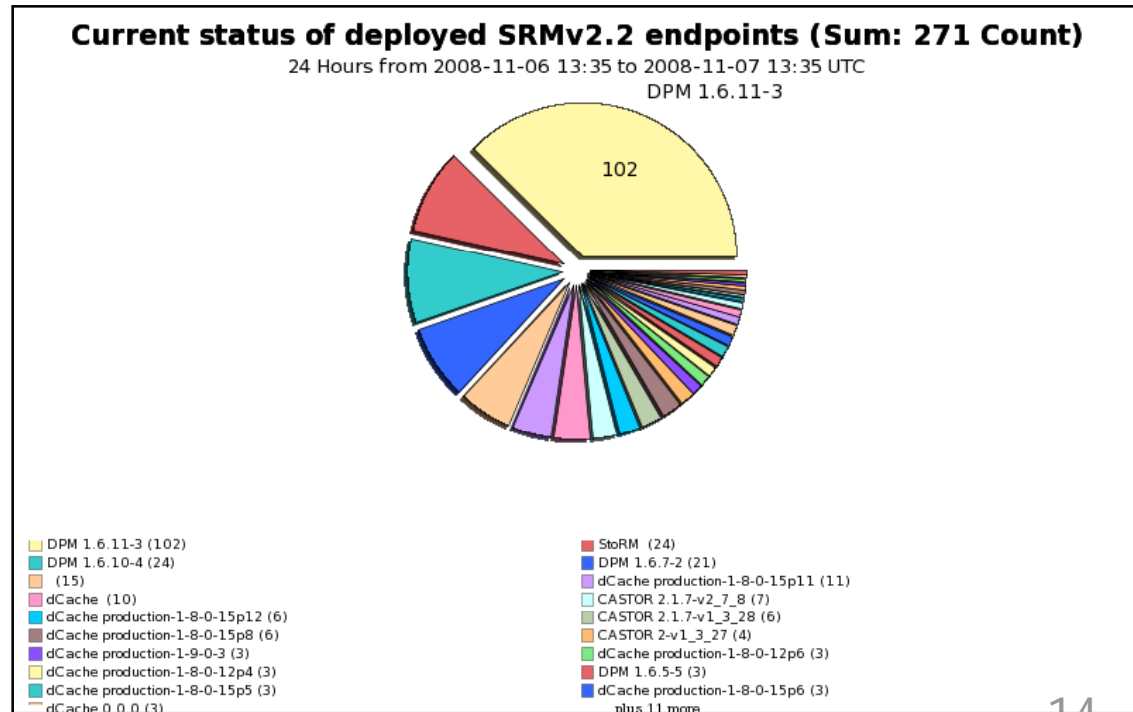
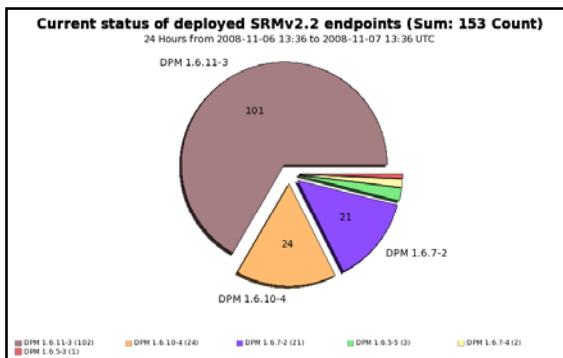
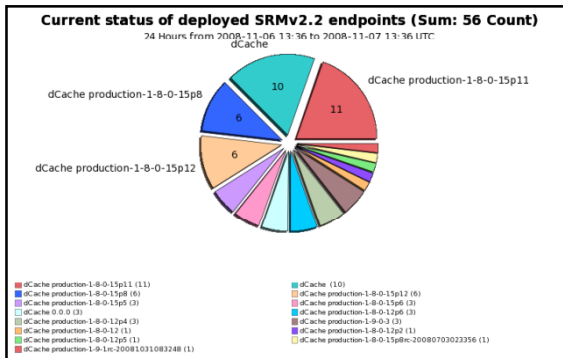
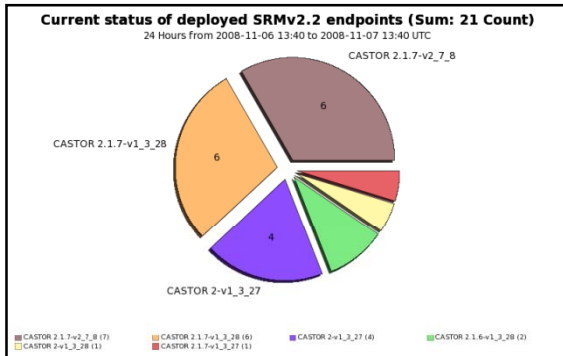
- Operations pain points:
  - Channel setup: the best way to partition the work by channels is difficult to decide (ATLAS, LHCb, CMS)
    - Everything on \* channels: good for capacity planning, bad when one site blackholes the single channel
    - Split channels: better control per site, bad: management fragmentation
    - Sometimes a mixture is needed
    - Investigate usage of “group” channels
  - Multiple FTSs overloading a single site’s SRM (ATLAS/LHCb/CMS)
    - Balancing / tuning often done ‘by hand’ across sites
    - Investigating ways of allowing SRM to signal ‘BUSY’ to its clients
  - Disambiguation of stage-in timeout errors (file was on tape) from something's-got-stuck timeout errors (ATLAS)
    - This would help experiments and WLCG operations a lot
    - That SRM.get does issues an implicit stage-in also complicates matters a little
  - More later... (operational cost of whole stack)

# Incident summary

- Over the 5-6 month period
- LFC: less than 1/month in WLCG (T0 and T1)
  - Incidents related to bad 1.6.9 release
  - 1-2 related to streams
- FTS: ~1.5/month in WLCG (T0 and T1)
  - 1 hardware related (resilience of service)
  - 1 Oracle related weirdness
  - 3 configuration related (missing cleanup, complex procedures)
  - 2 related to delegation bug
- DB downtimes excluded
  - Mostly scheduled, the odd unscheduled downtime too
  - 0.22 % annual server unavailability (patches deployment, hardware) – see Maria's talk...

# Storage deployment

- What sites are using (Greig's plots)
  - Snapshot of WLCG info system endpoints
  - Instance in info system:  
DPM: 150, dCache: 51, Storm 24, Castor 17
  - Sites: T0/T1s: 4 Castor, 1 Storm, 8 dCache



# Storage-related incidents

- There have been a large number of incidents
  - “Good week” ~1 major/week in WLCG
    - Impact varies from a some hours to many days
  - Several smaller ones per week tracked in daily meeting (< MoU degradations)
    - Wide variety of random problems: timeouts, things stuck, files not accessible
    - **Investigation of these is very time consuming**
- I won't focus on individual incidents – try to draw some common threads...
- See <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGServiceIncidents> for all incidents on all services

# Incidents - Storage

- **Not at all exhaustive** – excludes a lot of small degradations reported in daily meeting – and weighted heavily towards the last 2 months when we got the “automatic post-mortem” procedure going
  - **Thanks to site for reporting these issues**
  - The last couple months have seen a large improvement in the reporting and systematic ‘post-mortem’ follow-up by sites
- 
- BNL (early July) - ATLAS
    - inaccessible files
  - BNL (14/8) - ATLAS
    - pnfs issues
  - CERN (25/8) – Weekend CMS, LHCb
    - Networking related
  - CERN (28/8) – 6 h CMS
    - DB related SRM
  - CNAF (7/9) 12+h – All
    - DB related
  - CERN (10/9) – ATLAS 1st data
    - Diskserver hotspots
  - RAL (17/9) 17h LHCb, 12h ATLAS
    - DB related, 14k files lost
  - CERN (8/10) - CMS
    - data corruption
  - CERN (12/10)- LHCb
    - Hotspots / pre-staging
  - RAL (18/10) 55 hours - ATLAS
    - DB related
  - SARA (22/10) - ATLAS
    - High load
  - ASGC (25/10) week+ – ATLAS, CMS
    - DB related
  - IN2P3-CC (25/10) - ATLAS
    - Pnfs?
  - SARA (28/10) 10 hours – All?
    - Tape system issues
  - PIC (31/10) 10 hours - All
    - Pnfs



# General storage issues

- Negative feedback ...?
  - More often than not, things work OK 😊 😊
  - When things work fine they work fine 😊
  - When it starts to go bad, failures build up and quickly kill the storage ☹️
- DB-related issues
  - Oracle-related problems (RAL, ASGC) – tickets open
  - Cleanup: pnfs hoovering, Castor SRM cleanup, DPM cleanup, FTS table cleanup – fiddly ops procedures
- Data loss resynchronisation
  - It's a bit of a pain (for site, for VO)
- Data corruption – checksumming is desirable *everywhere*
  - Hardware disk corruption (issues with CMS at CERN)
  - It brings particular operational pain if the corrupted files are detected after distribution
  - More generally e.g. FTS / gridFTP check-summing requested
- Infrastructure / manpower effects
  - Procurement, installation, power, cooling, etc
  - Operations cost of running the service (later...)

## dCache pnfs performance / interactions with FTS

- Reports from various sites (PIC, NDGF, IN2P3, SARA) about poor pnfs performance
  - Affects directory listings and 'deep' SrmLs (stat with details)
  - Improved version available (Postgres, factor 2)
  - Scales with hardware – this helps...
- Bad interactions between FTS and dCache?
  - Doing a srmLs -l stat is expensive in dCache (e.g. hits pnfs hard) – FTS does it once per attempt, and more to do the mkdir -p
    - When things go bad, FTS starts timing out and (experiment) retries, so the rate of stats increases, the system slows down, the retries fail
    - CMS 'gentle' retry policy with FTS handles this better
  - Similar issues noted in ATLAS pre-stage exercise (also doing deep stats to determine the online status of files)
- Contributes to the “when things go bad, they go really bad” slippery-slope effect
  - Agreed possible solutions with FTS team (option to not check directory)
  - Use (SRM\_INTERNAL\_ERROR) to signal to clients to back-off
  - This would help CASTOR too

# Interactions...

- 3<sup>rd</sup> party copy with dCache and FTS / lcg-util
- Patch to Globus gridFTP2 client from NorduGrid to allow 3<sup>rd</sup> party transfers to work much more efficiently (hand-off of transfer to the backend dCache diskserver)
  - Now incorporated into VDT
- SRM Copy protocol probably not necessary?

# Storage DB issues (RAL, ASGC)

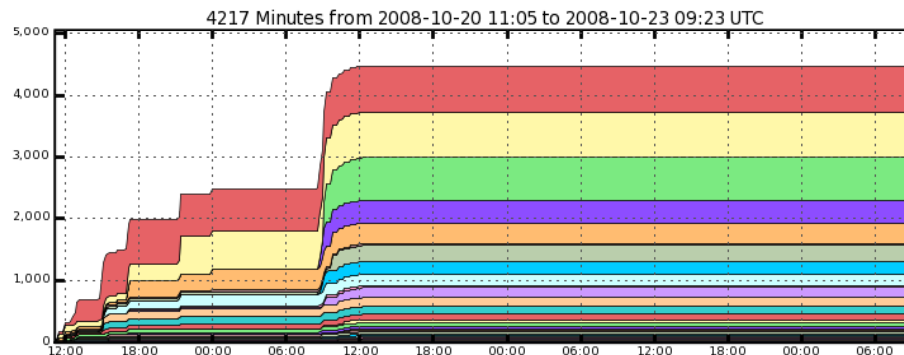
- Reported in DB workshop Tuesday (Gordon)
  - Maria's talk for summary...
- A number of Oracle related incidents affecting Castor:
  - Schema cross-talk – seen before in previous Oracle versions and “fixed”, but probably recurring?
    - Difficult to reproduce. Oracle SR open.
    - 14 K files lost
  - Other causes of service degradations being followed up on the DB side
    - Cursor invalidation
      - SR open
    - PK violation
      - Corrupted file entries in nameserver

# Handling of D1T0 files

- ATLAS, LHCb, CMS have reported issues with files being inaccessible (Castor / dCache )
  - Diskservers fail from time-to-time, and require intervention during which files will not be accessible
    - Intervention times can be long...
  - LHCb asking if this could somehow be signalled – so that files could be marked temporarily inaccessible in the catalogue
    - reduce job failures, distinguish from terminal failures (file lost)
    - SRM file status should UNAVAILABLE ! (fixed Castor/dCache)
- DPM@T2 (GRIF) Automated handling of disk-server outage is desirable
  - Currently, putting a server back online is a manual operation

# Stage-in / Bulk prestaging

- Stage-in / Bulk prestaging
  - Backlog recovery when files are on tape – issues noted at CERN for ATLAS
    - Efficient use of the mass storage system can be tricky
      - Alice noted they explicitly want to avoid this
    - Some problems caused by tape defragmentation also contributed
  - Bulk pre-stage for reprocessing at T1s - ATLAS
    - Generally works OK on Castor sites
    - dCache interfaces to multiple tape systems
      - Serious problems noted with massive pre-staging at some sites
      - IN2P3 noted explicitly an issue with the HPSS backend handling (dCache drips-in the stage-in request in rather than sending all at once, so tape mount efficiency is impacted – config?)



# Diskserver Hotspots

- There have been various degradations due to disk “hotspots”
  - Disk-server balancing is difficult to get right – since it’s not up to the storage system which files are requested
    - Better suited algorithms?
    - And recovery from a hotspot can be a tricky operation
  - Interaction with hardware downtime (D1T0 especially)
    - After one day downtime, any ‘missing’ files in datasets that are on the affected disk servers are suddenly *really* popular
- User pools / production pools
  - Castor: SRM may select a user pool for production exports...
    - Affected 1<sup>st</sup> ATLAS Real Data files. Also CMS.
    - Contributed to the backlog clearing problem
  - Source space tokens in FTS will help here (FTS 2.1)

# Sharing experiences?

- It is not uncommon for an experiment to run into problems seen, analysed and resolved by another
  - Same point for sites sharing information
- More sharing of “solutions” would mean less pain and better use of available effort
- Examples:
  - pnfs overload issue: CMS have seen this before and made the workaround (keep directories below 1000 files), ATLAS just fell over it
  - Cleanup of databases – FTS ‘history’, DPM ‘cleanup’, Castor SRM ‘cleanup’, pnfs ‘hoovering’
  - CMS transfer retry policy is interesting for other VOs
- Some good sessions in the pre-GDB on this!



# Tier-2 status

- Some input on this (GRIF, ATLAS, US-CMS)
  - DPM / dCache / Storm / BestMan deployed at T2s
    - Alice testing xrootd interfaces at T2
  - Generally fairly positive for 'production' – no serious operational issues reported
    - USCMS: Noted dCache can be hard to run for a T2 – lack of standard admin tools
  - T2 user analysis is a concern for ATLAS – activity started

Outlook

- Analysis challenges are necessary to stress the system under full load
- Already discovered systematic problems – sites have to be closely involved
- Will expand to other clouds
- Just started - much more automation needed

*Where do we go from here?*

- Already we have learned from these challenges.
  - How the 1 Gb network limitations manifest as errors.
  - Our analysis challenge toolkit is building:
    - Manual job preparation and submission is a solved problem.
    - Results merging, parsing, and publishing is always good, but perhaps more metrics would be useful:
      - Improved error identification
      - Correlation of error types with SE, CE, time, ...
- We plan to maintain continuous stressing of the Italian cloud.
  - We will repeat "Seconda" tomorrow at 10:00 CET.

21

- T2 user analysis tested in CCRC'08 for CMS and work has continued
- Simulation reasonably well exercised

# Features?

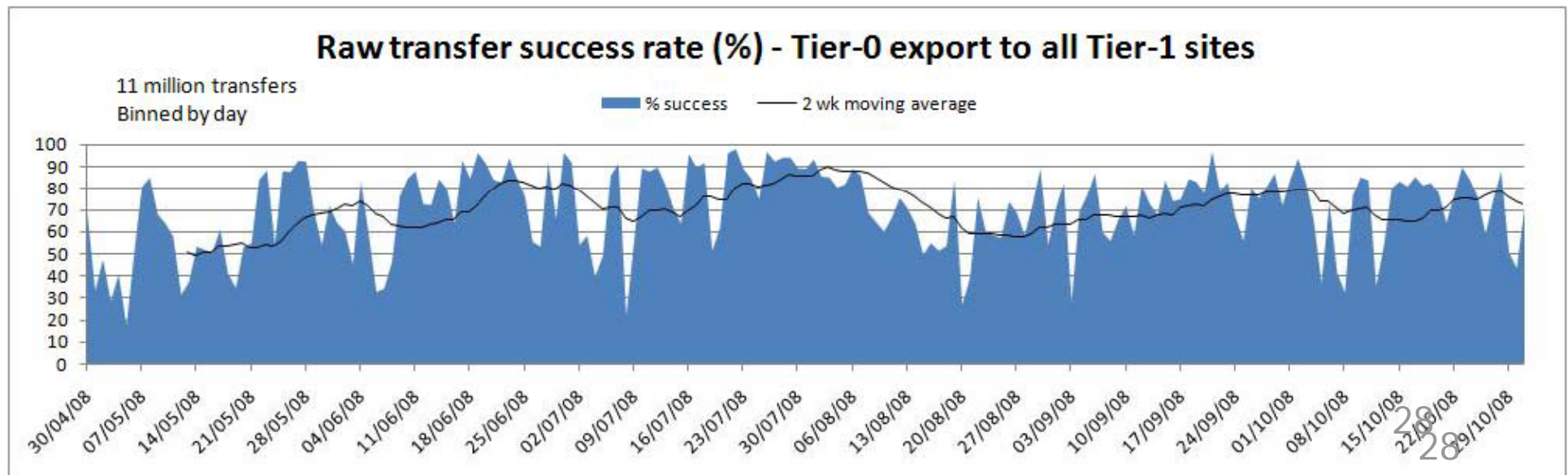
- Some feature requests remain
  - See Flavia's talk at GDB yesterday...
  - Point made by ATLAS that some operations are very expensive due to "missing tools" and can only be accomplished by doing it the "wrong way" – these operations stress the whole system
    - dCache SRM v1 -> SRM v2 migration "metadata" change to add "space" attributes required files to be 'copied' internally
    - FTS source space tokens (...)
- This is the yardstick I would use to prioritise feature requests

# Layers

- The current data management system has many layers
  - Experiment (more than 1 layer itself) <-> FTS <-> SRM <-> Stager <-> gridFTP <-> Tape
- Errors can and do happen in any layer
  - The frequency of errors is rather high
  - Interactions between components are sometimes not optimal (see earlier...)
  - The SRM 2.2 "standard" really isn't
    - Many comments that you still have to care about the SRM type
  - Biggest complaint! The error messages we get back remain rather obscure
    - ...and become even more obscured by their propagation through the various layers of software
  - Even identifying the site at fault is often a challenge for some transfers (gridFTP error messages being the classic here)
    - gridFTP's "the server dropped the connection"
    - More examples?

# Efficiency?

- Efficiency of the whole data management stack
  - (11 million transfer attempts binned by day, T0 to all T1 sites)
  - Expt framework $\leftrightarrow$ FTS $\leftrightarrow$ SRMs $\leftrightarrow$ Stagers $\leftrightarrow$ gridFTP $\leftrightarrow$ Tape
  - RAW transfer rate (#successes / total # attempts, per day)
    - Failure can and do happen in any layer at both ends of a transfer
  - About  $\frac{1}{4}$  of all transfer attempts fail
    - Most failures are ~quick (in the SRM) and don't consume network
  - All the files get there in the end! (multiple retries)



# Operational cost

- Operational cost of debugging all these is still too high
  - Impacts sites, experiment shifters (triage mode)
  - Blind retry policies don't help!
  - Inaccessible files (~hours), gridFTP mysteriously timing out (bouncing emails/phone calls back a couple of times to the other site) (~hours)
  - "We see reduced transfer rates, please could you check" (~hours)
    - Request from some sites for more specific tickets – but sometimes hard!
  - Performance variance is typically very large and not well understood
    - Some files go at 10MB/s, some go a 200KB/s, same site, same time
    - "It looks OK now" – ATLAS note problems sometimes appear 12am then disappear mysteriously at 5am
- It's not clear the situation is improving
  - Better monitoring does help. Starting initiatives in end-to-end monitoring to try and tie all the components together.
    - Aim: to be able to follow a WAN transfer through all components on both sites
  - Common and [semi-]automatic incident tracking, problem reporting/sharing and change management procedures are desirable (look at ITIL?)
    - Tomorrow afternoon's WLCG operations session...

# Summary

- Lots of achieved over last few months ☺
  - Overall, the system is working – the data is (usually) getting where it needs to get to, and is generally available once it's there
- LFC service stable (<1 incident/month in WLCG)
  - Errors at the level of 0.1% now (ATLAS)
    - Focus on ensuring (high-)availability of individual instances
- FTS service still work to do (1.5/month in WLCG)
  - gLite 3.1 / SLC4 version finally released after extensive testing
  - A few niggling issues remain
    - delegation bug, DB cleanup not integrated
    - resilience of channels to node-failure could be improved
    - Optimal choice of which channels to set up can be hard

# Summary

- The major challenge remains in the robustness and stability of the storage systems
  - We still see >> 1 degradations per week, often >1 major (MoU)
  - General stability, negative feedback?
  - Pnfs issue / improve interactions between components
  - Several DB-related issues / Schema cleanup issues – deja vu?
  - How the system handles diskserver downtime (D1T0) / Disk server hotspots
  - Issues with stage-in / bulk-prestaging
- Reports from T2 are positive
  - Concerns that analysis has not really been exercised, ongoing exercises
- The total operational cost of the DM system we have is still too high – particularly noticeable on the transfers
  - Complex multi-layered interactions with storage on both sites remain expensive to debug
  - Results in inefficiencies
- Some missing features exacerbate the situation

# Summary

- We're surviving with the current system
  - The work is getting done successfully! 😊
  - But ... higher than comfortable operational cost 😞
  - Still too many random incidents, particularly in the storage layer
  - Following up these incidents remains expensive and often takes (much) longer than the MoU limits
  - Had we been taking lots of Real Data this summer, it would have been 'exciting'
- The key challenge remains the operational cost and ensuring the sustainability of the core data management services
  - This is where we need to focus our efforts