



Outcome of February Run & Plans for May

Jamie.Shiers@cern.ch

~ ~ ~

LCG CB, 24th April 2008

Agenda

- **Outcome of February Challenge and plans for May**
- Discussion:
 - ¿ *Were the goals met in terms of measured service reliability, targets for response to and resolution of experiment-related problems?*
 - ¿ *Did the regular meetings, conference calls and Tier-2 coordinators (put in place for CCRC'08) improve performance?*
- I will present a few slides on each of the above topics, primarily to stimulate discussion (main messages are not new...)
- **These points were also discussed during the on-going WLCG Collaboration Workshop**

CCRC'08 – Objectives

- Primary objective (**next**) was to demonstrate that we (sites, experiments) could run together at 2008 production scale
 - **This includes testing all “functional blocks”:**
 - Experiment to CERN MSS; CERN to Tier1; Tier1 to Tier2s etc.
- Two challenge phases were foreseen:
 1. **February :** not all 2008 resources in place – still adapting to new versions of some services (e.g. SRM) & experiment s/w
 2. **May:** all 2008 resources in place – full 2008 workload, all aspects of experiments' production chains

💣 **N.B. failure to meet target(s) would necessarily result in discussions on de-scoping!**

- Fortunately, the results suggest that this is not needed, although **much** still needs to be done before, and during, May!

Main Lessons Learned

- ☺ **Generally, things worked reasonably well...**
- **Still improvements in communication are needed!**
 - Tools still need to be streamlined (e.g. elog-books / GGUS), and reporting automated
 - ¿ **Service dashboards – should be in place before May...**
 - F2Fs and other meetings working well in this direction!
- **Pre-established metrics extremely valuable!**
 - As well as careful preparation and extensive communication!
- ↳ **Now in continuous production mode – this will continue – as will today's infrastructure & meetings**

Recommendations – mid-February!

- ✓ To improve communications with Tier2s and the DB community, 2 new mailing lists have been setup, as well as regular con-calls with Asia-Pacific sites (time zones...)
 - Follow-up on the lists of “**Critical Services**” must continue, implementing not only the appropriate monitoring, but also ensuring that the WLCG “standards” are followed for **D**esign, **I**mplementation, **D**eployment and **O**peration
 - We are having some (small) success in convincing others in the Grid....
 - Clarify reporting and problem escalation lines (e.g. operator call-out triggered by named experts, ...) and introduce (light-weight) post-mortems when MoU targets not met
 - We must continue to improve on open & transparent reporting, as well as further automations in monitoring, logging & accounting
- ⚡ We should foresee “data taking readiness” challenges in future years – probably with a similar schedule to this year – to ensure that full chain (new resources, new versions of experiment + AA s/w, middleware, storage-ware) is ready **[maybe also July 2008!]**

Experiment View

In Order of Appearance (March F2F...)

- CMS
 - [Very] Detailed presentation of up-front metrics per functional block
 - 100% success not reported, but well understood status
- ATLAS:
 - CCRC was a very useful exercise for ATLAS
 - Achieved most milestones in spite of external dependencies
 - It's difficult to serve the Detector, Physics and IT community!
- ALICE:
 - For ALICE, the CCRC exercise has fulfilled its purpose
 - Focus on data management
 - Brings all experiments together
 - Controlled tests, organization
- LHCb:
 - Initial phase of CCRC'08 was dedicated to development and testing of DIRAC3
 - CCRC'08 now running smoothly
 - Online->T0 and T0-T1 transfers on the whole a success
 - Some issues with reconstruction activity and data upload from the WNs
 - Investigating with Tier-1s recent problem of determining file sizes using gfal
 - Quick turnaround for reported problems

What Did We Achieve? (High Level)

- Even **before** the official start date of the February challenge, it had proven an extremely **useful** focusing exercise, in helping understand missing and / or weak aspects of the service and in identifying **pragmatic** solutions
- Although later than desirable, the **main bugs** in the middleware were fixed (just) in time and many sites upgraded to these versions
- The deployment, configuration and usage of SRM v2.2 went **better** than had predicted, with a noticeable **improvement** during the month
- Despite the high workload, we also **demonstrated** (most importantly) that **we can support** this work with the available manpower, although essentially no remaining effort for longer-term work (**more later...**)
- If we can do the same in May – when the bar is placed much higher – we will be in a **good position** for this year's data taking
- However, there are certainly **significant concerns** around the available manpower at all sites – not only today, but also in the longer term, when funding is unclear (**e.g. post EGEE III**)

May Run – Introduction

- 2nd phase of CCRC'08 is from Monday 5th May to Friday 30th.
 - This is immediately preceded by a long w/e (Thu-Sun) at CERN
- May is also a month with quite a few public holidays
 - ~1 per week...
- **Everyone is keen that we improve on February in terms of being ready well in time**
 - And in terms of defining what “ready” actually means...
- Detailed presentations on middleware, storage-ware, databases, experiments plans & requirements at April's F2F
 - This can be summarized in a couple of tables, on the CCRC'08 wiki, as was done for February – *very small updates wrt February!*
- **I will also mention communication – but at the end!**

- Hardware upgrade – important to get this in place prior to CCRC'08 to be consistent with setup for this summer's LHC data taking

Upgrade to new h/w performed last week using DataGuard technique - minimum downtime; old h/w retained as "Standby" to new. Oracle (s/w) upgrade postponed.

Possibility to upgrade in June? TBD...

need joint decision on whether to postpone until after 2008 data taking or not...

LFC DB backend deployment: Current Understanding (TBC)

- Site deployment is a **site** decision – subject to meeting requirements of experiments (**negotiation**)
- LHCb requires Oracle backend for LFC in order for current **replication** mechanism from central global catalog at CERN to local R/O replicas to work
- ATLAS has stated requirement of **deleting** $O(100K)$ entries (catalogs + storage) per day – see separate talk on performance goals and results. Primary requirement, however, is **operational**.
- Survey of sites is given in following **table**:

WLCG Tier1 Sites

Site (ISO standard)	Experiments (ATLAS, LHCb)	Comments
CA-TRIUMF	ATLAS	MySQL, plan to migrate to Oracle in 2009.
DE-KIT	ATLAS, LHCb	Oracle for LHCb, MySQL for others
ES-PIC	ATLAS, LHCb	
FR-CCIN2P3	ATLAS, LHCb	Oracle
IT-INFN-CNAF	ATLAS, LHCb	Oracle
NDGF	ATLAS	<i>r/s?</i> (LFC on PostgreSQL(?) or MySQL)
NL-T1	ATLAS, LHCb	Oracle (+MySQL for non-HEP VOs)
TW-ASGC	ATLAS	
UK-RAL	ATLAS, LHCb	MySQL → Oracle in 2-3 weeks (ATLAS?)
US-BNL	ATLAS	LRC. Evaluating LFC. b/e decision later.
US-FNAL	N/A	N/A

Storage Baseline Versions for CCRC08 in May



Implementation	Addresses
<p><u>CASTOR:</u></p> <ul style="list-style-type: none"> • SRM: v 1.3-20, • b/e: 2.1.6-12 	<p>Possible DB deadlock scenarios; srmLs return structure now conforms; Various minor DB fixes; Fix for leaking sockets when srmCopy attempted; correct user mappings in PutDone; log improvements; better bulk deletion(?)</p>
<p><u>dCache:</u></p> <ul style="list-style-type: none"> • 1.8.0-15 	<p>http://trac.dcache.org/trac.cgi/wiki/manuals/CodeChangeLogs T1D1 can be set to T1D0 if the file is NOT in a token space; There is a new PinManager available.(improved stability); Space Tokens can be specified in a directory, to become the token used for writing if no other information is provided; dCache provides to the HSM script : Directory/file, Space Token, Space Token Description (new)</p>
<p><u>DPM:</u></p> <ul style="list-style-type: none"> • 1.6.7-4 	<p>Issues fixed for 1.6.7-4: Pool free space correctly updated after filesystem drain and removal ; SRM2.2 srmMkdir will now create directories that are group writable (if the default ACL of the parent gives that permission)</p> <p>Known issues or points: No srmCopy is available; Only round robin selection of filesystems within a pool; No transfer stream limit per node</p>
<p><u>StoRM</u></p> <ul style="list-style-type: none"> • 1.3.20 	<p>http://storm.forge.cnaf.infn.it/documentation/storm_release_plan Upgraded information providers, both static and dynamic; Fix on the file size returned by srmPrepareToGet/srmStatusOfPtG for file > 2GB; Support for ROOT protocol; "default ACLs" on Storage Areas that are automatically set for newly created files; srmGetSpaceMetaData bound with quota information; Improved support for Tape1Disk1 Storage Class with GPFS 3.2 and TSM</p>

SRM v1.1 Services – R.I.P. ??

- **When can sites de-commission SRM v1.1 services?**
 - Pre-May is presumably not realistic... (i.e. end-April)
- Experience from May will presumably (re-)confirm that SRM v2.2 is (extremely fully) ready for business...
- Decide at June post-mortem workshop?

‡ **Set tentative target of end June?**

¿ *Is this too aggressive? Too mild?*

➤ **Looks like by end-2008 WLCG can be declared an "SRM v1.1-free zone!"**

Middleware Baseline

CCRC May recommended versions

Component	Patch #	Status
LCG CE	Patch #1752	Released gLite 3.1 Update 20
FTS (T0)	Patch #1740	Released gLite 3.0 Update 42
FTS (T1)	Patch #1671	Released gLite 3.0 Update 41
gFAL/lcg_utils	Patch #1738	Released gLite 3.1 Update 20
DPM 1.6.7-4	Patch #1706	Released gLite 3.1 Update 18

Icg-CE

- Patch #1752
 - <https://savannah.cern.ch/patch/index.php?1752>
- Fixes
 - Significant performance improvements
 - Limit on jobmanager processes
 - Persistent daemons, not re-invoked perl
- rpms
 - 2 new rpms on CE
- Released in glite 3.1 update 20
 - <http://glite.web.cern.ch/glite/packages/R3.1/updates.asp>

FTS (FTA)

- Patch #1671
 - <https://savannah.cern.ch/patch/index.php?1671>
- Fixes
 - [bug #33148: FTS: error in srmGetSpaceTokens in SRM copy push transfers](#)
- Rpms
 - [glite-data-transfer-url-copy-2.0.1-8.i386.rpm](#)
- Released in gLite 3.0 Update 41
 - <http://glite.web.cern.ch/glite/packages/R3.0/updates.asp>

gfal/lcg_util

- Patch #1738
 - <https://savannah.cern.ch/patch/index.php?1738>
- Fixes numerous bugs
- rpms
 - [GFAL-client-1.10.9-1.slc4.i386.rpm](#)
 - [lcg_util-1.6.9-2.slc4.i386.rpm](#)
- Released in glite 3.1 update 20
 - <http://glite.web.cern.ch/glite/packages/R3.1/updates.asp>

DPM 1.6.7-4

- Patch #1706
 - <https://savannah.cern.ch/patch/index.php?1706>
- Fixes
 - Bug #34799 (umask/ACL issue)
 - Numerous other bug fixes
- Released in glite 3.1 update 18
 - <http://glite.web.cern.ch/glite/packages/R3.1/updates.asp>

DPM 1.6.10

- Patch #1605
 - <https://savannah.cern.ch/patch/index.php?1605>
- Fixes
 - Numerous bug fixes
 - Bug #34799 (umask/ACL issue) already fixed in 1.6.7-4
- Released in glite 3.1 update 20
 - <http://glite.web.cern.ch/glite/packages/R3.1/updates.asp>

Summary – (Part 1)

- The February run of CCRC'08 was largely **successful** and introduced the important element of **up-front** and **measurable** metrics for many aspects of the service
- We still have a lot to do to **demonstrate** full 2008 readiness – **May and beyond will be a busy time with no let up in production services prior to then**
- We have developed a way of planning and operating the service that works **well** – re-enforce this and build on it incrementally

➤ **This is it – the WLCG Production Service!**

Agenda

- Outcome of February Challenge and plans for May
- Discussion:
 - ¿ *Were the goals met in terms of measured service reliability, targets for response to and resolution of experiment-related problems?*
 - ¿ *Did the regular meetings, conference calls and Tier-2 coordinators (put in place for CCRC'08) improve performance?*
- I will present a few slides on each of the above topics, primarily to stimulate discussion (main messages are not new...)

Discussion

- The 2nd of these points is easier to answer than the 1st:
 - Yes – the meetings, conference calls, wiki pages, e-logbooks etc made a concrete difference.
 - Daily (except Mondays – weekly operations, and during F2Fs) meetings certainly helped CERN service providers & experiments communicate and resolve problems reasonably rapidly
 - More work is needed to optimize these for outside sites
 - More awareness of existing minutes & summaries to MB
 - Some of the ideas need to be taken back into existing systems (e.g. GGUS)
 - This is not the first time that this has happened – rapid prototyping of a down-time scheduling / logging tool was eventually superseded by corresponding enhancements to GOCDB / CIC portal
 - The number of problems reported per day was relatively low throughout the challenge and the follow-up good (not perfect)
 - Derek Ross gets prize for most consistent attendance (>>me)

Discussion (cont.)

- **Tier2** (and **DB**) coordinators – this was proposed mid-January, and additional names added to the corresponding lists – but it is not clear if the people involved were aware of their new responsibilities...
- Certainly, Tier2s still say that they are not fully in the loop / unsure of what is required of them
- This is also true of Tier1s, who continue to request more detailed information from the experiments
- Specific efforts were made to help Asia-Pacific sites
 - Regular con-calls plus a workshop prior to ISGC
- Collaboration workshops – such as this week – together with experiment weeks / jamborees etc – are very valuable communication tools – but there are limits!

Communication

- Is an on-going issue...
 - We have made – and continue to make – a lot of progress
 - It is unlikely that everyone will 'agree' about every detail
 - And we must – as has been repeatedly pointed out – be both pragmatic and efficient
- **Let's focus on the things that seem to be broken, whilst continuing to make progress in other areas...**
- For example, it is a perfectly fair comment that the daily WLCG operations meeting (@15:00) is 'like joining a CERN meeting remotely'.
- It *is* useful for service – experiment (- site?) interaction – we should try and make it more so
- ¿ *Make more people aware of the 'minutes' and weekly summaries*
- ¿ *More pro-active use of wlcg-xxx-contacts roles?*

Week-Daily Operations Calls

- ***Not*** on Mondays (joint operations at 16:00)
- ***Not*** during F2F meetings (1st(?) week / month)
- ***Short; focussed; minutes available***
- ***Agenda & call details:***
<https://twiki.cern.ch/twiki/bin/view/LCG/AgendaandCallDetails>
 1. Dial +41227676000 (Main) and enter access code 0119168, or
 2. To have the system call you, click [here](#)
- Minutes (one wiki page per week) are summarized to the MB of the following week (see following)

Current activities – Services

Service	Issue
DBs	Two major activities. 1) Preparation of new RAC 5 and 6 hardware (both now on critical power) to migrate experiment services next week. 2) Upgrade of integration RACs to Oracle 10.2.0.4. This version will probably not be used for CCRC'08(?) More details in notes...
3D	CNAF down for the entire week due to maintenance. Streams replication to CNAF has been split and stopped, it will be resumed this week. ATLAS streams replication between offline and online database stopped on Good Friday due to an error from the streams apply process caused by an user application mis-configuration. The replication was resumed after a few hours.
CASTOR	The new CASTOR version (2.1.7) will be ready for pre-production certification on the 1st of April as originally planned. All tests and stress tests are successful.
SRM	A bug has been found in the CERN SRM v2 databases. This is in fact fixed in a rollout due on Monday so the rollout has been brought forwards.
FTS	Some Apache patches are in the pipeline for FTS servers.
CE	The fixes deployed at CERN to the LCG CE that reduce their load by an order of magnitude are being packaged for external distribution and should be ready next week.
DM	Patches to the 1.6.7 version of lcgutils should also be ready for distribution next week.

Current activities – ATLAS

Day	Issue
Wed	Throughput testing with junk data. Build up T0 to T1 transfers to 150% nominal, adding 50% each day (150% over the w/e) CNAF has a downtime so their share is going to BNL. NDGF have requested a reduced rate due to insufficient resources. SRMv2 is being used at CERN.
Thu	Some Castor fixes at the end of the day. Try to reach 100% of the rate today then 200% on Friday then throttle back to 100% for the weekend. A current problem is they no longer see BNL on the ATLAS dashboard.
Fri	1) Yesterday the export rate was supposed to be increased to 100% of nominal. Problem in the ATLAS T0 machinery and not enough LSF jobs were submitted. 2) Display problem with the dashboard. Some entries end up in the production dashboard but should be in the T0 dashboard and vice versa. 3) SARA is not getting data on disk since disk is full (being cleaned up now). NDGF does not get data on disk since disk is full (can not be cleaned up centrally since NDGF does not use LFC - the only supported catalog for the central deletion tools). NDGF people should do the cleaning (alerted).
LFC	ATLAS are planning a bulk CERN site change to about 60K LFC entries. The consensus was that this is a relative simple operation that should only take a few tens of minutes of real time. To be scheduled!

Critical Service Follow-up

- Targets (not commitments) proposed for Tier0 services
 - Similar targets requested for Tier1s/Tier2s
 - Experience from first week of CCRC'08 suggests targets for **problem resolution** should not be too high (if ~achievable)
 - The MoU lists targets for responding to problems (12 hours for T1s)
 - ¿ Tier1s: 95% of problems resolved < 1 working day ?
 - ¿ Tier2s: 90% of problems resolved < 1 working day ?
- **Post-mortem triggered when targets not met!**

Time Interval	Issue (Tier0 Services)	Target
End 2008	Consistent use of all WLCG Service Standards	100%
30'	Operator response to alarm / call to x5011 / alarm mailing list	99%
1 hour	Operator response to alarm / call to x5011 / alarm mailing list	100%
4 hours	Expert intervention in response to above	95%
8 hours	Problem resolved	90%
24 hours	Problem resolved	99%

Discussion (cont.)

- We were not very rigorous regarding these targets during February – concentrating mainly on major events (prolonged site down-times, prolonged service down-times)
- Post-mortems were carried out in a small number of cases and revealed areas to be improved
- **Unsurprisingly, these can be summarized in a single word – communication**
- Hopefully, the discussions earlier in the week will result in further improvements
- But we still need the systems that automatically measure where we are wrt these targets...
- (Not to mention follow-up on critical services...)

Overall Conclusions

- **Yes** – we are a working collaboration and **yes** – we have a service
- It is constantly improving – this can be clearly seen if we look back to Victoria or even the previous workshop at CERN
- We've shown that we can work through even complex and unpleasant situations / problems
- May will no doubt bring new challenges – most likely in areas that have not been fully tested or not at the right scale
- pp data taking will bring yet further demands – in particular, the excitement and pressures of a new machine