

WLCG Collaboration Workshop
21 - 25 April 2008, CERN

Introduction / Roadmap

Jamie.Shiers@cern.ch



Worldwide LHC Computing Grid
Distributed Production Environment for Physics data Processing



General

- Most sessions are [here](#) in the CERN main auditorium
- live Webcast & recording has been requested
- A few parallel sessions - **Tuesday**, Thursday and Friday - are in the IT amphitheatre, B160, B40, B28 & B513
 - ATLAS sessions Tue B40, Thu IT amphi; CMS Thu B40
- Coffee breaks organised morning and afternoon (except Thursday pm) outside main auditorium
- **Not for other sessions!**
- 🍸 “Collaboration drink” Wednesday evening in main restaurant
 - Watch out for gate-crashers!



And the modify key?



Registration

- 223 people registered
 - List with guard at Gate B
 - Normally a surge of registrations at (or after) the deadline
 - Typically 20% no-shows and 20% "VIPs"
- 💣 Closed definitively Wednesday 16th April**

CCRC'08 – Agenda for Collaboration Board presentation

➤ 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 co-ordinators (put in place for CCRC'08) improve performance?*

- In other words, many of the issues to be discussed – and hopefully resolved (where appropriate) – at this workshop!



Main Workshop Themes

- Mon am: Reliable Services; pm: storage
- Tue: review of **February** CCRC'08
- Wed: planning for **May** CCRC'08
- Thu / Fri: Operations Track
- Thu / Fri: Database Track
- Thu pm: Collaboration Board (by invitation)
- **👉 Tue / Thu: ATLAS jamboree (B40, IT amphi)**
- **👉 Thu: CMS sessions (B40)**



Service Reliability = D.I.D.O.

- The Goals - WLCG MoU Targets & Experiment Critical Services Lists
- The Techniques: middleware developers' golden rules
- The Techniques: deployment strategies
 - application DNS load balancing, Oracle Maximum Availability Architecture', ...
- **Measuring What We Deliver & Problem Resolution Targets**
 - Middleware & experiment-ware surveys - how do we rack up?
 - Testing Experiment Services - SAM tests et al
 - ATLAS Grid Services

Design, Implementation, Deployment, Operation

(Re-)implementing Services

- Cannot take existing non-HA service, add SLA and make it HA
- Several techniques available - need to be further studied
 - openlab et al
- Need to discuss options and impact early on with users
 - e.g. DataGuard, Replication, RAC all offer solutions in area of HA
 - Most appropriate depends on needs of service
 - How much downtime can be tolerated?
 - Pros & cons of multiple cheap boxes versus more expensive (and more complex) hardware
- Learnt a lot from implementing ~24 x 7 services for LCG
- Extend to other services as part of 'group vision' (standards!)

Possible H/A DB Solutions

<u>Technology</u>	<u>RAC</u>	<u>DataGuard</u>	<u>Replication</u>
	Requires special H/W	Logical or physical - complete DB	Finer level of granularity
	More complex than single box to manage	Good for reports on standby	(A)synchronous
Notes:	Best paper solution? (But most expensive)	Widely used in conjunction with RAC	Performance / management overhead?
Real Application Clusters (RAC) = DB clusters	Automatic application failover (R&W)	Logical standby has some limitations on schema	Multi-master (and more complex setups) possible
DataGuard = Standby DB	We are using them! (On Sun with SAN...)	Physical standby is 'closed'	Stop-gap for distributed RLS? (CMS...)



ATLAS Critical Services (PDF)

Tier	Service	Criticality	Consequences of service interruption
0	Oracle database RAC (online, ATONR)	Very high	Possible loss of DCS, Run Control, and Luminosity Block data while running. Run start needs configuration data from the online database. Buffering possibilities being investigated.
0	DDM central services	Very high	No access to data catalogues for production or analysis. All activities stops.
0	Data transfer from Point1 to Castor	High	Short (<1 day): events buffered in SFO disks, backlog transferred as connection is resumed. Long (>1 day): loss of data.
...			
0-1	3D streaming	Moderate	No export of database data. Backlog can be transferred as [soon as] connections are resumed.

... more ...

CHEP 2007



CMS Critical Services (wiki)

Rank	Definition	Max. Downtime	Comments
11	CMS Stops Operating	0.5 hours	Not covered yet
10	CMS stops transferring data from Cessy		Cessy output buffer time
9	T0 Production stops		min(T0 input buffer/Cessy output buffer) or defined time to catch up
8	T1/T2 Production/analysis stops		
7	Services critical when needed but not needed all the time (currently includes documentation)	0.5	
6	A service monitoring or documenting a critical service	8	
5	CMS development stops if service unavailable	24	
4	CMS development at CERN stops if service unavailable		

... more ...

CHEP 2007



ALICE critical services list

- WLCG WMS (hybrid mode OK)
 - LCGRB
 - gLite WMS (gLite VO-box suite a must)
- FTS for T0->T1 data replications
 - SRM v.2.2 @ T0+T1s
- CASTOR2 + xrootd @ T0
- MSS with xrootd (dCache, CASTOR2) @ T1
- PROOF@CAF @ T0

CHEP 2007



LHCb Critical Services (CCRC08 wiki)

Service	Criticality
CERN VO boxes	10=critical=0.5h max downtime
CERN LFC service	10
VOMS proxy service	10
T0 SE	7=serious=8h max downtime
T1 VO boxes	7
SE access from WN	7
FTS channel	7
WN misconfig	7
CE access	7
Conditions DB access	7
LHCb Bookkeeping service	7
Oracle streaming from CERN	7

... more ...

CHEP 2007



Definitions of "Critical"

Experiment	Down	Seriously Degraded	Perturbed
ALICE	2 hours	8 hours	12 hours
ATLAS 	As text	As text	As text
CMS	30'	8 hours	24 hours (72)
LHCb	30'	8 hours	24 hours (72)

Quite significant differences in list of services under each heading:

- ATLAS: *only* 2 services are in top category (ATONR, DDM central services)
- CMS: (long) contains also numerous IT services (incl. phones, kerberos, ...)
- LHCb: CERN LFC, VO boxes, VOMS proxy service
- ALICE: CERN VO box, CASTOR + xrootd@T0 (?)

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 e-mail	99%
1 hour	Operator response to alarm / call to x5011 / alarm e-mail	100%
4 hours	Expert intervention in response to above	95%
8 hours	Problem resolved	90%
24 hours	Problem resolved	99%



February Post-Mortem

- CCRC08 - View from WLCG Service
 - CCRC08 - View from Site Preparation
 - ATLAS summary - lessons learnt and things to be improved
 - CMS summary - lessons learnt and things to be improved
 - ALICE summary - lessons learnt and things to be improved
 - LHCb summary - lessons learnt and things to be improved
-
- **Panel on Support Issues**
 - **The Metric - as a Common Challenge, involving all experiments and sites, how did we do overall?**



May Planning

- Resource / Service Review - What can we count on for 2008 production?
 - ATLAS Status, Plans and Requirements for May run and beyond
 - LHCb Status, Plans and Requirements for May run and beyond
 - CMS Status, Plans and Requirements for May run and beyond
 - ALICE Status, Plans and Requirements for May run and beyond
 - Baseline Middleware Versions - what is in production!
 - Baseline storage-ware versions - what is in production
-
- **Running, monitoring and measuring the challenge**
 - **CCRC'08 Post-mortem workshop in June**



Database Track

- DB service and streams intervention responsibilities
- Experiments' updates: new applications / deployment plans
- Transportable table spaces for scalable re-instantiation
- FTS / LFC DB resource needs and configuration
- Experiment connectivity requests and firewall update
- DB server limits (processes/sessions), session sniping
- New storage back-ends, storage performance
- Experience with 32 to 64 bit migrations

- DBA Q&A Session - A collection of "hot" DBA topics



Key Discussions...

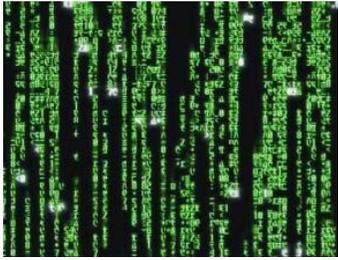
- Storage issues - Monday afternoon
- "The Metric" - Tuesday afternoon
- Wednesday - understanding and agreeing not only experiment requirements but also middle / storage-ware versions for May
- Thursday / Friday - WLCG Operations in post-EGEE (/EGI?) era
- Useful to have "rapporteurs" to take notes from these sessions...

¿ **Volunteers?**



WLCG Storage Issues

- Storage Performance Issues, Monitoring etc.
 - This session will discuss how to improve the performance of tape-backed MSS systems
 - or what the experiments could/should do to give hints to the system
 - experiences, suggestions for dCache and Castor, etc.
 - See also: { One and the same monitoring tool is needed for tapes for all T1's. Action on: Jamie Shiers }
 - We have another use case, but this is primarily for CERN. We still have users reading from tape and we didn't know.
 - Actually there are hundreds of tapemounts per day for reading by users and we didn't know.
 - If we would have had any monitoring we would have been able to spot this much earlier.
 - Now we learned for the first time in a presentation by Tim Bell.
- Data (access) security issues
 - This is a big topic that never gets properly addressed. 'Afterwards' would be too late...
- Summary of recommendations / proposals regarding SRM v2.2
 - for May and perhaps beyond...
- **Storage is something we should clearly revisit during the CCRC'08 "post-mortem" workshop in June...**



The Metrics

- *Based on the site / experiment presentations, how well did we do in:*
 - Defining the scope of the challenge;
 - Preparing for it;
 - Running it;
 - Measuring & reporting our success?
- ¿ **What needs to be fixed before “steady state production”?**



m/w & storage-ware

- m/w situation should be clear... [earlier]
- But what about storage-ware?
 - Likely needs some clarification and discussion...
- This is not the same discussion as Tuesday ("hints", optimisation, monitoring, ...) but more "Storage Solutions" - umm - solutions
 - N.B. see also tentative agenda for June post-mortem on (possible!) xxx-ware futures!



Post EGEE III operations



Comments on EGI design study

- In 2010, the LHC will be running
- In 2010, EGEE I will be running
- It is inconceivable to have:
 - a. Don't run the LHC
 - b. Run the LHC and EGEE I
 - c. Run the computer centre
- **This is required for the transition that are dependent on the LHC**
- The transition must be:
 - a. On time
 - b. Non-disruptive
- **This is a fundamental requirement for discussion**
- Goal is to have a fairly complete blueprint in June
- Main functions presented to NGIs in Rome workshop in March
- Essential for WLCG that EGI/NGI continue to provide support for the production infrastructure after EGEE-III
 - We need to see a clear transition and assurance of appropriate levels of support
 - Transition will be 2009-2010
 - **Exactly the time that LHC services should not be disrupted**
- Concerns:
 - NGIs did not all seem convinced that a large European production-quality infrastructure is a goal
 - **Not clear that there is agreement on the scope – hard to see how a clear agreed blueprint will be available in June**
 - Tier 1 sites (and existing EGEE expertise) not well represented by many NGIs
- WLCG representatives must approach their NGI reps and ensure that EGI/NGIs provide the support we need



Future Events(?)

- CCRC'08 post-mortem workshop 12-13 June
- HEP application cluster at EGEE '08
 - Around that time -overlapping with data taking - we will have to start thinking about 2009: middleware versions, storage-ware, resources etc and the plans for testing it!
 - ¿ Follow-on CCRC'09 F2F (2 day?) Nov / Dec 2008?
- WLCG Collaboration workshop in Prague, prior to CHEP '09 - *number of parallel sessions? (DM? DB? OPS?)*
- *Possible workshop prior to EGEE '09?*
- *WLCG Collaboration workshop in CERN, Spring 2010*
- *Pre-CHEP workshop in Taipei prior to CHEP '10*



Summary

- All pertinent information is (intended to be) on Workshop page
 - e.g. how to get free bus passes from airport, how to register MAC address of your PC / Smart Phone,
- If it's not, we'll add it...
- Includes a "Certificate of Attendance" which you are welcome to print, sign, eat, ...
- Hope to see you back in June for the CCRC'08 post-mortem workshop...

Questions?

If not, on with the workshop...



Worldwide LHC Computing Grid
Distributed Production Environment for Physics data Processing