SUMMARY FROM WLCG TEG WORKSHOP

Ian Bird

CERN

8th Feb 2012

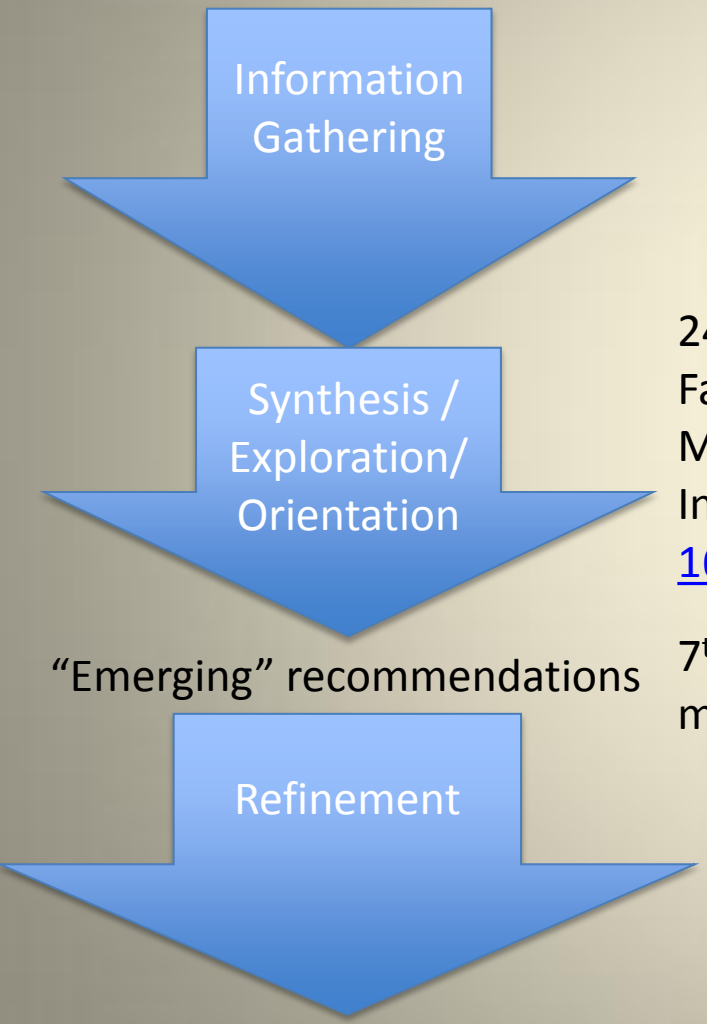
SOME CONTEXT
WHAT DOES THE FUTURE LOOK LIKE?

- WLCG is a collaboration that has made a lot of use of projects such as EDG, EGEE, EGI, PPDG, OSG, etc.
- These are now ending – EMI and EGI (at CERN) end Q2 2013, OSG enters a new round of funding (not yet clear at what level)
- Unlikely to attract generic funding for grid developments in the near future
- So what does WLCG become?
  - Much complexity has migrated from grid layer to experiment-specific layer (sometimes in common between experiments)
  - Sites need to focus on providing a extremely robust physical infrastructure and key access services
    - Simplifying the grid complexity can help do this
  - Sources of development effort are limited

SOURCES OF EFFORT
PROVOCATIVE … BUT…

- There will continue to be a certain limited(!) level at CERN (IT and PH-SFT)
- Hope that other large national grid projects can help (INFN? GridPP?)
- Can we build collaborative/community efforts to support key software?
  - Have not done this successfully very much
- The only way to really have a supportable software is to really collaborate across all the entities in WLCG – experiments and sites
  - And try to get key focussed developments from external sources where possible
- This really implies common solutions and collaborative efforts
- I doubt that supporting software and services for a single experiment will be easy to justify or fund in future

SM/ DM TEG Process so far



- Initial questionnaire
- Defined topics [TopicsDataStorageTEG]
- Questions to experiments:
  Experiment Presentations and Twikis
  [ALICE; ATLAS;CMS;LHCb]

24/25 Jan: Face-to-Face Meeting Indico : 165687

- Storage Middleware presentations

- Face-to-face session for each topic plus broader discussions.

7th Feb GDB meeting

We have:
- **Agreed set of observations**
- **"Emerging" recommendations and direction**

We need:
- Overall architecture "layer" diagram
- Concrete proposals under each topic

Proposed Steps until mid-March

- Work on the straw-man of current architecture
  - component diagram, description and responsibility split
- In parallel in the topical working groups
  - agree concrete recommendations from information which is already available
- Focused iteration to document some key model elements
  - Strategy for use of storage with lower level of service
    - e.g. cache level storage for a) pushed or b) on-demand data
  - Table with main data types
    - associated service level, authorization + quotation needs, r/w vs r/o
  - Minimal user/group model to implement desired SE security model
  - Table of interface & protocol requirements
    - needed by whom but also *for what*

WM TEG report, summary

- Feedback from the audience
  - Novel input from experiments, that did not come up when discussing similar topics during TEG work. This is normal, and shows that TEG reports are useful also as catalyzers.
  - Discuss extension of environment variables to include:
    - HS06 from WN (this was actually in the HEPiX-virt proposal already)
    - Remaining job lifetime (discuss how)
    - Signaling for job termination (LRMS dependent)
  - Include analysis (a couple of years from now) about the need for a "gLite WMS" service. Currently:
    - SAM tests: all experiments
    - SW installation: ATLAS, will move that to PanDA
    - Production and user jobs: CMS (but moving more activity toward glideinWMS); LHCb, until direct submission works everywhere
  - Volunteer computing?

- We should consolidate the documents currently in the WM TEG wiki – timeframe 1 month.

WM TEG, next steps

Item	Conclusion	Timeframe (tentative)
glexec, accepted by all VO's for handling pilots	**ATLAS will adopt glideinWMS** for this part	Dep. on ATLAS tests
Streamed submissions	**Define and recommend** extension of the CE interface; run tests **(*)**	Definition, 1m; implementation, TBD
Common ways of running pilots	**Conclude evaluation** on problem scope and potential solutions if any	Conclusions, 1-2m?
Support of multi-core jobs	**Recommend** using JDL extensions; **define** environment variable, interactions with IS; run tests **(*)**	PoC, by CHEP?
CPU- vs. I/O-bound jobs	**Recommend** extension of the CE interface; run tests **(*)**	PoC, by CHEP?
CPU affinity, cgroups	**Evaluate** possible solutions; run tests **(*)**	Evaluation, 1m; implementation, TBD
Requirements for a CE	**Require** adoption of a "virtual CE" alias. **Require** support of common LRMS.	CE alias implementation, EMI-3?
Requirements for a WMS	Decommissioning not feasible in the short term. **Evaluate** if/when possible with experiments.	Evaluation, by CHEP?
Virtualization technologies	**Find a (more permanent) forum** for sharing experiences between sites and VO's	2012Q2
Cloud computing	**As above**; **adopt** HEPiX-virt recommendations for existing Cloud test sites; **evaluate** AuthN/AuthZ; **explore** dynamic provisioning of resources.	2012Q3
Information system	**Expectation**: in the future WLCG experiments will continue to need mostly a simple discovery service.	N/A
Extension of environment info	**Discuss** including environment info about HS06 (easy), remaining job lifetime, soft/hard signals for job termination	Evaluation/PoC, by CHEP?

(*) This TEG? Other groups?

Security TEG - Priorities

- Reminder: the most important risks for WLCG, according to the risk analysis (http://cern.ch/go/dt9S):
  - Misused identities (WLCG identities are not limited to x509)
  - Attack propagation between WLCG sites
  - Exploitation of serious OS vulnerabilities

- Provide recommendations to mitigate these risks
- Concentrate on clarifying the traceability requirements
- Continue discussions in the “AAI on the WN” subtask
- Concentrate on other subtasks, in particular:
  - AAI on the storage systems
    (work likely to be conducted directly within the Storage TEG)
  - Usability vs security

Security TEG - Timeline

- Finalise the risk analysis document and submit it to the MB
  - By end Feb 2012
- Prepare recommendations for the most important risks
  - Gradual delivery, from Feb 2012 until end 2012
- Propose a trust model for the WN
  - By May 2012 (CHEP)
- Provide a report on the storage systems security status/ requirements
  - By May 2012 (CHEP)? (to be agreed with the storage TEG)
- Prepare reports on other areas
  - Timeline depending on contributions to the TEG

Databases for Conditions Data

- CORAL (ATLAS, CMS, LHCb) and COOL (ATLAS, LHCb) are examples of very successful common projects
- The CMS conditions software relies on CORAL for accessing Oracle (both directly and via Frontier) and SQLite.
  - CMS expects that CERN IT will provide the users with an adequate support for the CORAL framework, which should be devoted mainly in bug fixing and in improving performance bottlenecks when identified.
- ATLAS relies on COOL for the conditions database infrastructure, on CORAL to access the database layer, and on CoralServer for database access in the online environment.
  - ATLAS expects that these products will be supported by CERN as long as they are used by ATLAS and other experiments.
- LHCb needs CORAL and COOL to continue to be supported, as they are essential components of the LHCb software.

Access to Conditions Data

- CMS, ATLAS and soon LHCb use Frontier/Squid for Conditions DB access.
  - Central monitoring of worldwide Squids is very important to keep the Squids operating properly. The monitoring is now done by computers operated by CMS Frontier, but we recommend that a plan be made to transition the Squid central monitoring to WLCG
  - Locating the Squids is currently done separately per experiment and application, but we recommend that there be a WLCG standard way for jobs to locate Squids
  - We recommend that sites share Squids for all production services (currently Frontier and CVMFS)
  - Frontier/Squid should be recognized as a WLCG service and treated accordingly (GOCDB, GGUS, central rpm repository, monitoring)

NoSQL Recommendations

- It seems to us evident that there are valid use cases for providing and supporting at least one of the NoSQL technologies at CERN
- In order to be able to properly advise developers within the experiments groups, the CERN-IT-DB group should test the most common NoSQL products for the already known use cases and acquire expertise with them
  - Technology tracking and market survey should also be part of this task.
- ➡ Discussion on this point: CERN-IT-DB vs CERN-IT vs WLCG vs community support
- CERN IT should deploy a suitably sized Hadoop cluster
  - Focus on Hadoop rather than fragment effort over a variety of NoSQL tools
    - Other tools can, and will, be run ad-hoc by experiments as necessary
  - Hadoop clients, including pig/hive available on user interfaces (lxplus?)
  - Reasonably sized HBase installation
  - We make no operational requirements on the cluster, and appreciate that it will require training etc. for ops staff, so may run at low service level initially.
  - In the end it may need development, integration and production clusters
- Experiments would like to be involved in deployment discussions
- Build a community around the tools
  - Best practice doesn't really exist at CERN; have a forum to communicate what is learnt
  - Other groups may be interested in using these tools (Dashboard seems like a good candidate for example) but a central service is needed before expanding the user base

Global Recommendations

#	Title	Area	Timeline
R1	WLCG Service Coordination	Operations	From 2012
R2	WLCG Service Commissioning	Operations	From 2012
R3	WLCG Availability Monitoring	Monitoring	2012
R4	WLCG Site Monitoring	Monitoring	2012
R5	WLCG Network Monitoring	Monitoring	LS1
R6	Software deployment	S/W	2012/LS1
R7	Information System (WM TEG)	Underlying Services	2012/LS1
R8	Middleware Services	M/W	2012/LS1
R9	Middleware Deployment	M/W	2012/LS1

Order by Time

- Short term, specific time bounded and well defined targets
  - Availability, Site & Network monitoring
  - Software deployment
- Medium term, require a WG and need goals and metrics
  - Information system, Middleware Services and Deployment
- Long term, requires coordination and communication
  - Service Coordination and Commissioning

Ordering by Principle

- Reduce operations effort
  - Service Coordination and Commissioning, Site and Network Monitoring, Software deployment, Middleware Services
- Reduce complexity
  - Software deployment, Middleware Services and Deployment
- Minimize inter-dependencies (sites, experiments, services)
  - Software deployment, Information System
- Reduce effort to upgrade and reconfiguration
  - Middleware deployment
- Improve access to information
  - Information System, Availability, Site and Network monitoring
- Improve reaction to service/hardware failures
  - Site Monitoring
- Deploy scalable services (2-3 times above the average load)
  - Middleware Services

NEXT STEPS

- At next MB I will propose a small editing team to bring the work of the TEGs so far into an overall strategy
- Call out areas where it is clear further work needs to be done
  - Either as continuation of the TEG or
  - As a dedicated (short term) focused wg
- At future MB/GDBs
  - Review recommendations and set priorities
  - Plan for where effort needs to be invested
- Iterate until done

SUSTAINABILITY ANYONE?