

Table of Contents

Summary of GDB meeting, December 12, 2012.....	1
Agenda.....	2
Welcome - M. Jouvin.....	3
Operations Coordination.....	4
Introduction - G. Merina.....	4
Discussion.....	4
CVMFS - S. Roiser.....	4
Discussion.....	5
perfSONAR - S. Campana.....	5
Discussion.....	6
Actions Update.....	7
EMI Migration Update - P. Solagna.....	7
Requirements for GLUE2 -M. Alandes Pradillo.....	7
Discussion.....	8
Oracle at T1s.....	9
ATLAS - D. Barberis.....	9
CMS - I. Fisk.....	9
Discussion.....	9
DPM Community and Workshop - O. Keeble.....	10
Discussion.....	10
Virtualized WNs: where are we? what next?.....	11
Final Report for HEPiX WG - T. Cass.....	11
Discussion.....	11
Software Defined Networks for Big-Data Sciences - I. Monga.....	12
Discussion.....	12
MW Provisioning Lifecycle - M. Schulz.....	14
Discussion.....	15
NDGF Update - O. Smirnova.....	16
Discussion.....	16

Summary of GDB meeting, December 12, 2012

Agenda

<https://indico.cern.ch/conferenceDisplay.py?confId=155075>

Welcome - M. Jouvin

2013 GDBs

- January: moved to 3rd Wednesday (January 16) as CERN is reopening only the 2nd Monday
 - ◆ A pre-GDB confirmed on January 15: F2F meeting of the Operations Coordination
- March: 2nd Wednesday (March 13) but at KIT (Germany)
 - ◆ Details at January GDB
- April: moved to 1st Wednesday (April 3) because of the clash with EGI Community Forum

Planned pre-GDB

- January: F2F Operations Coordination meeting (confirmed)
- February: AAI on storage
 - ◆ To be confirmed in January by a Doodle poll to check all required parties are available
- Not yet scheduled: virtualized WNs and cloud resources

GDB summaries seem to be appreciated but lack of volunteers to take notes

- Should not be a specialized role handled by 1 or 2 persons
- Michel agrees to finalize them and takes responsibility for the contents...
 - ◆ No more post-processing of notes required from volunteers
- Michel will try to build a team of potential volunteers and establish a rota among them

Main forthcoming meetings

- IPv6 WG F2F meeting, CERN, January 24-25
 - ◆ <http://indico.cern.ch/categoryDisplay.py?categId=3538>
- ISGC 2013, Taipei, March 17-23: registration open, participation encouraged
 - ◆ Call for papers is officially closed but late submissions are still welcome
- EGI Community Forum, Manchester, April 8-12
 - ◆ The call for participation and registration is now open
 - ◆ http://cf2013.egi.eu/call_for_participation/index.html

Operations Coordination

Introduction - G. Merino

All meeting minutes available on WLCGOpsCoordination Twiki

- 6 coordination meetings since September
- F2F meeting at pre-GDB in January

Several dedicated task forces

- SHA-2: need to test RFC-proxy readiness already "now", while SHA-2 CERN CA should become available in January
 - ◆ MW validation will be done with EGI
 - ◆ OSG ahead of EGI: preparing a compliant release very soon but still to be deployed. Bestman status still unclear (JGlobus-2 enhancement to be confirmed). Not really coordinated by TF.
 - ◆ dCache still not ready: waiting for an update (not clear they can go for the Bestman approach)
- MW deployment: moving away from unsupported MW, migration to SL6
 - ◆ EMI-2 SL5 WN: ok for all VOs, tarball expected in January
 - ◆ EMI-2 SL6 WN: ATLAS not yet ready, other VOs ok. Sites should not upgrade if they support ATLAS
- glxexec: SAM test added to EGI ROC_OPERATIONS, failures raise alarms followed up by EGI/NGI
 - ◆ Goal: reach 75% early next year
 - ◆ VO: CMS ready, ATLAS/ALICE work to be done, LHCb needs to check if current support is ok
- Tracking tools: email list created, people interested should subscribe
 - ◆ Hot topic: Savannah-JIRA migration + GGUS bridge
- FTS3: functional tests ongoing for CMS and ATLAS, several FTS3 features demonstrated already, asked to install FTS3 clients in a public area
 - ◆ MySQL backend already available
- xrootd
- Squid monitoring: move monitoring of Squid as a WLCG responsibility rather than a VO responsibility
 - ◆ Currently using a tool developed at FNAL based on MRTG and awstats
 - ◆ Evaluating the possibility to use Aggregated Topology Provider (ATP) as the source of configuration information (feeds available from OIM and GOCDB, an extension to GOCDB planned that may be enough)

Gonzalo moving to a new job, thanks for the work done!

- Probably need to find a replacement for Ops Coord

Discussion

- ATLAS jamboree: move toward Xrootd and HTTP/WebDAV
 - ◆ followup in task force after Wahid's WG has concluded

CVMFS - S. Roiser

Version 2.1.5 now in beta: shared cache and NFS export

- NFS export is intended as a "last resort" configuration when there is no other option

- Stable version foreseen in January

ALICE recently joined the TF: will move to CVMFS for SW deployment

- Now all 4 VOs involved in TF
- ALICE first targets the HLT farm, deployment at sites later next year after validation
- ATLAS and LHCb set a target for full CVMFS deployment by end of April 2013
 - ◆ A GGUS ticket will be sent starting in Feb. for missing sites
 - ◆ No new SW deployment outside CVMFS after the deadline

When the TF started, 97 sites targeted

- 26 already deployed it
- 24 have planned to deploy it early 2013
- 10 have plans but no date
- 32 sites have not responded: GGUS tickets opened

Discussion

- NFS export is supported long term and has been tested at a site with 2k cores
- Unresponsive sites can be discussed at the F2F meeting
 - ◆ for LHCb they provide little resources
 - ◆ we may raise the bar without losing a lot

perfSONAR - S. Campana

perfSONAR "requires" 2 different hosts for bandwidth and latency tests

All sites should run a perfSONAR instance but started by list of sites provided by experiments

- Sites grouped in region or clouds

Until recently, the perfSONAR configuration was local to each instance: does not scale with a meshed configuration

- Will move to a central configuration: feature available but not yet natively built into the current perfSONAR release: need to run a shell script to activate it
 - ◆ Will be in next release planned next quarter
 - ◆ Sites should not wait for that release to install perfSONAR
 - ◆ Existing sites are requested to move to the central configuration

Sites must register their perfSONAR instances (both) in GOCDB/OIM

Scope of tests

- One mesh per region + OPN mesh
- T2s to T1s: Each region against OPN
- Inter-region: considered important by CMS and ATLAS
- Bandwidth test frequency depends on the scope of the tests
 - ◆ Every 6h for intra-region/cloud tests

TF manpower mainly from IT, US, UK

- Would like to have one person per region

- ◆ 1 common to all experiments
- ◆ Small FTE fraction: more a contact

Documentation being streamlined

A modular dashboard [↗](#) is available to display the results.

- Code being moved from BNL to GitHub
 - ◆ Modified BSD license, contributions are welcome
- Regions welcome to start their instance

From January, larger deployment campaign will start.

Discussion

- EGI: a GGUS support unit will start in April, handled by DANTE
 - ◆ Will be for perfSONAR-MDM
 - ◆ Test interoperability with perfSONAR-PS not yet demonstrated, but backend data is compatible
 - ◆ Tests being investigated, to be followed up next month
- The BNL dashboard has an API that is used for injecting the results into the ATLAS SSB
- The WG will look into the desired evolution of the API, alarms, GUI, ...

Actions Update

EMI Migration Update - P. Solagna

Sites still running unsupported MW today: 44 sites, 82 services, not including DPM/LFC/WN

- BDII, CREAM, lcg CE, VOMS, WMS, LB, dCache
- 26 sites had plan to migrate by end of Nov. but failed to do it
- New deadline: Dec. 17. After this date, sites must have at-risk downtime declared for the impacted services
- No site suspended yet: a downtime for unsupported services is considered enough if there is a reasonable plan to upgrade

DPM, LFC and WN now unsupported: alarms are opened against sites since Dec. 11 (already 44)

- WN status not clear after CE migration
 - ◆ gLite 3.2 WN still being used at many sites
- Handled by ROD rather than COD
- Ticket asks for plan to be communicated in the next 10 days
- Jan. 31 is the deadline for upgrading

Next steps

- Add probes for ARC unsupported MW (prior to EMI-1)
 - ◆ 5 sites affected
- Deployment of probes for EMI-1 products that will be obsolete end of April
 - ◆ Warning from start of January
 - ◆ Alarms from March 1

Requirements for GLUE2 -M. Alandes Pradillo

Monthly meeting between IS experts, experiments, EGI and OSG

- First meeting last week
- Twiki page to track discussions

GLUE2 deployment status: 10 more sites in EGI

GLUE2 validation: validator tool being upgraded to comply with EGI GLUE2 profile

- Will be added as a Nagios probe
- Longer term plan: have the GLUE validator part of the resource BDII: will prevent publication of invalid information
 - ◆ BDII will include it in next version but it will take time to get it working and deployed for all products

ginfo: extended use cases, not only service discovery

- In contact with experiments to get feedback

Discussion ongoing about the information needed in the BDII, between static info provided by GOCDB and dynamic info provided by messaging.

OSG confirmed it has no deployment request from its stakeholders: wants to see GLUE2 proven in EGI first

- Will revise their plans accordingly in the future

Discussion

- `ginfo` will *not* be extended with support for GLUE 1.3
- The idea is to try moving to GLUE 2 in a gradual manner such that 1.3 can be phased out one day
 - ◆ Try to avoid investing further in obsolete technology
 - ◆ The monthly meeting has all stakeholders involved

Oracle at T1s

ATLAS - D. Barberis

3 databases

- geometry, calibration data... : Oracle + FronTIER
 - ◆ Confirmation from 4 T1s that they will continue to support it in the future. Will continue with these sites until end of RunII
 - ◆ DB releases will be phased out: direct use of FronTier when required
 - ◆ Condition data files now moved from HOTDISK to CVMFS
- TagDB: reducing from 5 to 3 sites
 - ◆ Commitment from these 3 sites for 12-18 months: enough to develop a new backend
 - ◆ Longer term plan: use Hadoop instead of Oracle
- Muon calibration centers
 - ◆ Calibration centers don't profit from CERN license
 - ◆ Revising plans to attempt moving to standard license (rather than the one including Streams)
 - ◆ Also working on alternative plans

CMS - I. Fisk

Conditions DB accessed by FronTier

- Oracle at T1 only as an insurance against catastrophic events: will reduce to 1 backup

Other usage at T1 is for FTS: will be happy to move to a new backend (MySQL) when available

- FNAL is very interested in participating in the testing: the only Oracle use at FNAL is for the FTS

Some Oracle DB at T0: may consider reducing the number if asked

- Positive experience with MongoDB and CouchDB
- PhEDEx is probably hard to move out of Oracle but only required at CERN

Discussion

- Tony: discussion with Oracle ongoing, no decision yet whether license will continue covering T1 usage or not
 - ◆ CERN is trying to ensure that T1 will have conditions at least as good as today
- ATLAS: each T1 can decide which backend they prefer for their FTS-3 instance

DPM Community and Workshop - O. Keeble

DPM workshop [↗](#) last week at LAL: 25 people registered

- First day was for product presentations and site/experiment feedback
- Second day was more hands-on and demo of new features: http/DAV interface, DMLite installation, memcache, new backends
- An important milestone in the process of creating a DPM community

Site feedback about main issues

- I/O concurrency limits
- Draining performance
- Rebalancing storage
- Hot file replication
- Metadata consistency
- Inter-VO quotas
- Redundant dpmdaemon
 - ◆ Nameserver can already be done
- Some of these issues already (partly) addressed by GridPP-maintained admin toolkit

Gridftp improvements planned to support redirection

- Will allow SRM-less operations required by services like Globus Online

DMLite

Configuration: Puppet use demonstrated and will be used in standalone mode as the future standard configuration tool in replacement of YAIM

- A wrapper is being developed to allow small sites to continue to use the YAIM interface they are familiar with

DPM collaboration: first meeting to establish the collaboration

- Representatives from CERN, France, Japan, Italy, Taiwan, UK
- Decided to proceed with establishment of the collaboration
- Statement: see slides

Discussion

- ATLAS: sites are asked to configure/enable xrootd and http access but sites expressed some concerns about the complexity
 - ◆ Probably more information must be released to site about what is involved exactly: additional RPMs required, config files to produce
 - ◇ Extra rpms are in EMI repository and/or (increasingly) EPEL
- The main issue was and still is the migration from gLite to EMI
 - ◆ Sites need to do that first
- RFIO use is deprecated

Virtualized WNs: where are we? what next?

Final Report for HEPiX WG - T. Cass

Initial concern about root access is no longer the main issue for many sites, eg. CERN

- Logs for traceability remain an important concern, in particular if the user has root access

Image endorsement model now endorsed by EGI Federated Clouds TF

Initial steps envisioned not really followed!

- Step 1: users choose between images prepared by sites. In fact, virtualised WNs are transparent to users
- Step 2: distribution of images between several sites. Has been demonstrated: technology is there.
 - ◆ CERNVM emerged as the main image
- Step 4: VM images connect directly to pilot framework
 - ◆ Clearly one of the issue that may be worked on together in the future
 - ◆ Need to work on how to transmit VO credentials to images
 - ◆ Role of the pilot factory, impact on batch system needs
 - ◇ How to avoid queuing requests for VMs?

Discussion

LHCb interested in working on direct use of cloud resources from DIRAC without going through batch systems

- A prototype exists
- Would like to use multi-core VMs
- Concerns about VM lifetime vs. fair share and shutdown grace period
 - ◆ The share needs to be measured in HEPSPEC, not how many VM or cores
 - ◆ Accounting should be on wall-clock time, as done at Amazon
 - ◆ Mechanism for communicating shutdown time has been agreed
 - ◆ Long lifetime is good for efficiency (less overhead), short lifetime may be better for traceability
 - ◆ VM per job would simplify usage of glxexec
 - ◇ Or let the VM run multiple tasks all coming from the same user

ATLAS / BNL also interested in participating in step 4 proposal

Start an email discussion between interested persons and come up with more concrete plans for possible work together

- Plan a pre-GDB on this topic early next year?

Software Defined Networks for Big-Data Sciences

- I. Monga

Why is there a problem with big-data sciences requiring bulk data transfers

- TCP is the underlying transfer protocol but it is a "fragile" workhorse
 - ◆ Sensitive to packet loss
 - ◆ Prevents using the underlying infrastructure efficiently
 - ◆ 2 possible solutions: replace/change TCP or layer 2 protocols like RoCE

Whatever the solution used will require

- Enough capacity provisioning to provide loss-free, high-bandwidth infrastructure
- Fast lanes (virtual circuit) end-to-end
- Big buffer
- Efficient monitoring for proactive resolution of problems

OpenFlow opens new possibilities

- Allows control of the data path (HW) from a central controller
- Based on flow tables where entries define actions taken based on rules based on 11-field tuples
 - ◆ Main actions are forward to a port (possibly after some transformations) or drop packet (firewall), ...
 - ◆ Can also forward to the openflow controller if no rule matches (or if a specific rule matches)

SDNs: provide a virtualized view of the global topology to applications (users)

- Allows an application to define rules applied to its view of the network resources and have them distributed throughout the network
- Example: ESnet can be presented to apps as a WAN virtual switch interconnecting a few sites
- Demonstrated at last SC with a 3-port VirtualSwitch
 - ◆ Specify edge OF ports
 - ◆ Specify backbone topology and bandwidth: was done using OSCAR mesh
 - ◆ Policy constraints like flowspace
 - ◆ Store the switch in a topology service

Future work

- Harden architecture and implementation: move from experiment to test service
- Verify scaling of the model, in particular size limit of a virtual switch
- Automation and intelligent provisioning

Discussion

- Network virtualization performance overhead will be manageable with newer HW generations
 - ◆ HW can/will even be designed for usage via OF
- User-defined networks might lead to greater disruption when something goes wrong
 - ◆ A user would be able to mess up his/her own SDN, but ought not be able to interfere with the SDN of others
- A big issue in WLCG has been (and occasionally still is) the end-to-end debugging of virtual circuits
 - ◆ How to do that with OF? Looks more complex, particularly when multiple domains are involved...

- ◆ ESNNet has a test/production service across multiple domains (countries) since a few years
 - ◇ Tested with PerfSONAR
 - ◇ More tools needed
- Topology and other meta data tools being worked on
 - ◆ Virtual circuits based on OGF NSI (Network Service Interface) standard

MW Provisioning Lifecycle - M. Schulz

Feedback shows the proposed approach is well accepted

No more central funding means all efforts become voluntary

- Fact that infrastructure depends on it doesn't change this
- Hopefully, most institutions involved are part of WLCG and committed to its success

Role of EGI, future open-ended EMI consortium, ScienceSoft

Scope: main focus is EMI components with gLite heritage

- OSG historically managed their MW independently
 - ◆ The two MW stacks have grown closer
- ARC: independent, no WN, not clear how much coupling is needed

Centrally managed rollback idea is rejected: no more discussion on how to do it

- Recipes will be published on how to do it for sites who may need it
- Site admins are generally not comfortable with the rollback idea
- In the Operations Coordination team there remain concerns about this matter
 - ◆ Effectively a fast rollback (or -forward) may be needed whenever an update has had a major negative impact

Versioned metapackages also rejected: will keep only unversioned metapackages

- Just describe the dependencies, not their versions
- Will evolve very slowly
- Does not help with the baseline versions required for various services

OS platform at discretion of PTs as long as RH and its derivatives are covered

- Coordination only required for OS major version changes

Configuration management: YAIM-core support ending soon and not clear what's the alternative for small sites

- DPM will try the idea of a "YAIM wrapper" to Puppet
- Transition period (~1 year) required to evaluate the appropriate solution

Pilot services: not necessary for most PTs

- In particular storage tends to be covered
 - ◆ Product communities and developers (e.g. FTS3)
- Main issue is probably the CE

Support: not all PTs are committed to provide level 3 support through GGUS

- EGI will do level 1 and 2
- GGUS must remain the place for tracking change requests: well known, no real alternative

Other questions

- MW discussions in WLCG as pre-GDB a few times each year
- Impact on other sciences: to be managed by PTs and their funding bodies
- Communication with PTs: they are asked to connect to pre-GDB or GDB when the plans for their products are discussed
- EPEL is strong on integration, weak on testing: material moving from epel-test to epel-stable can be time-driven, i.e. without sufficient testing
 - ◆ Need to put emphasis on automatic testing
 - ◆ Also need to do some real world testing involving a few production sites, as was done for the EMI2 WN

A new version of the proposal will be released soon, taking into account feedback received.

Discussion

- For the record: the idea of using rpm epochs was already removed from the latest document
 - ◆ There were good arguments against it, in particular that it makes difficult to identify what is the most recent RPM in a repository (no longer based on version in RPM name)

NDGF Update - O. Smirnova

NDGF T1 now part of NeIC which is an initiative of NordForsk

- NeIC director: Gudmund Høst
- HW and SW contributed by users

NDGF T1 main roles

- Overall coordinator: M. Wadenstein
- CERN liaison: Oxana
- Security officer: Leif Nixon
- Sites participating in T1 also host T2 resources
- Now extends outside nordic countries: Slovenia, Switzerland
- MW from EMI
- Huge variety of HW/SW configuration
- LHCOPN extending to all sites with T1 storage
- Pledges: 10% of ALICE T1 resources, 5% of ATLAS T1 resources

Discussion

- Funding period: no hard limit, at least the next 4 years should be OK
- Many proposals for new strategic areas are being evaluated

This topic: LCG > GDBMeetingNotes20121212

Topic revision: r4 - 2012-12-16 - MaartenLitmaath



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

or Ideas, requests, problems regarding TWiki? use Discourse or Send feedback