

Table of Contents

Summary of GDB meeting, September 10, 2014 (CERN).....	1
Agenda.....	2
Introduction - M. Jouvin.....	3
EGI Role in the Evolving DCI Landscape - Y. Legre.....	4
Discussion.....	4
Identity Federation - R. Wartel.....	6
Actions in Progress.....	8
Ops Coordination Report - M. Allandes.....	8
Information System Update - M. Allandes.....	9
SL/CentOS Discussions.....	10
Cloud pre-GDB Summary - M. Jouvin.....	11
Site Availability Change Proposal - P. Saiz.....	13
Data Management Discussion - O. Keeble, F. Furano and W. Bhimji.....	14
Multicore: Dynamic Partitioning with Condor at UVic - F. Berghaus.....	15

Summary of GDB meeting, September 10, 2014 (CERN)

Agenda

<https://indico.cern.ch/event/272777/>

Introduction - M. Jouvin

As usual, looking for volunteers to take notes

Next GDBs

- Next ones at CERN on the second Wednesday of the month
- Next one outside CERN in March 2015. Please, let Michel know if you are interested in organising it

On the list for next pre-GDBs: to be confirmed

- October: no pre-GDB but a F2F meeting of the IPv6 WG (Monday-Tuesday)
- November: volunteer computing
- December: data management (storage protocols)
- February/March: cloud issues
- Send suggestions if you have something you'd like to see discussed

Summary of WLCG Workshop

- Around 85 people attended, similar to Copenhagen Workshop
- Notes available in <https://twiki.cern.ch/twiki/bin/view/LCG/GDBMeetingNotes20140707>

Actions in Progress:

- Migration to GFAL2/FTS3 clients in October 1st
- Volunteering sites to provide dual stack IPv6 endpoints
- Batch accounting assesment in progress
- SL/CentOS: no news

Upcoming meetings: see slides

O. Smirnova asks about the status of Batch accounting assessment regarding handling of failed jobs

- J. Gordon explains that there are still on going discussions on what should be considered as a failed job: currently not treated the same way by all batch systems with respect to accounting

O. Smirnova asks whether virtual WNs should offer IPv6 endpoints

- M. Jouvin suggests to contact the IPv6 WG for details but is not aware of such a requirement
 - ◆ First goal is to assess that enabling IPv6 on a service endpoint doesn't impact IPv4

EGI Role in the Evolving DCI Landscape - Y. Legre

Evolution is marked by fragmentation

- Fragmentation of infrastructures: technologies, provisioning, policies, access mechanisms
- Community/RI fragmentation
- Lack of collaboration and trusts between RIs and e-Infra

EU Vision for 2020 as drafted by the e-IRG in 2012

- Single point of contact to provision all ICT services required (e-Infra commons)
- Ability for users to connect to the best expert consultancy (services)
- Ability to reuse science output (open data)

Strategy for e-Infra commons

- federated services
- joint capacity planning
- Open and integrated resource provisioning and access, e.g. OpenAIRE

EGI mission in this context: be a key providers of e-Infra commons and knowledge/expertise

- Challenges due fragmentation
- Address needs both of individual users and communities
- Joint development of solutions with communities
- Primary focus remains data

EGI Strategic actions

- New governance model including stakeholders, not only EGI
- New sustainability and business model: who funds the effort
- User involvement
- e4ERA proposal involving all major DCI actors, ESFRI, in Europe

Discussion

M. Jouvin expresses his concerns on the fact that EGI relies on the control of resources they do not actually own and that funding agencies are in a process of having a greater control on them with EU-T0

- Yannick: there has to be a common strategy and we need to work together to be able to consolidate e-infrastructures in Europe, or otherwise there is a risk that funding agencies won't invest on different projects.
- Yannick: there has to be a national programs for e-infrastructures coordinated at the European level, a model similar to GEANT.
- Michel: there is a risk that EGI has less influence on the way the resources are accessed, the type of service deployed, if this is more directly controlled by communities

On a comment by O. Smirnova, Y. Legre explains that other communities apart from HEP are growing and have different needs and it has to make sure these needs are also covered.

Question about more details on the e-infrastructure commons proposal.

- Yannick explains some proposals have been already made and for some others there will be consultation with people like WLCG, etc.

T. Wildish asks why Y. Legre believes there is a lack of trust, as mentioned in his presentation

- Yannick thinks people are not talking enough together.
- Moreover EGI was originally set up for HEP and then started to look to other communities, which may have lead to some frustration because it didn't have the resources to take care of all of them.

M. Jouvin: has the feeling that WLCG and EGI had a good collaboration so far, focused mainly on operations. At risk that with reduced EGI funding for operations, this will make it weaker..

Yannick repots that it is difficult to identify who can be a WLCG official representative

- WLCG could decide to be member of the EGI.eu council as CERN is not a WLCG representative in this body
 - ◆ But the council is made of legal entities...

T. Bell about EGI plan developments: we are trying to avoid the development of in-house solutions (or project funded solutions)

- Once the project comes to an end, has to be maintained by us

Identity Federation - R. Wartel

Goals reminder

- Enable people to use their home organization credentials: grid jobs, web portals, other X509 services
- Allow the use of home credentials for other services: Vidyo, Indico...

Build on existing federations and infrastructures: eduGAIN

- CERN involved via SWITCHaai
- CERN eduGAIN integration into CERN SSO (web apps) now done
 - ◆ Included a lot of legal work
- Next steps: enable eduGAIN on core service like Indico, OpenStack..., discuss with experiments to enable federated identity access to several experiment core services

Grid service integration through a portal, using EMI STS to translate federated identity to X509 (WLCG pilot)

- No concrete progress yet

Policy and trust issues: there is not trust/guaranty about the identities issue

- Our current security models relies on Authorization: not identity vetting during Authentication
 - ◆ Identity vetting done during registration: registration by experiments is CRITICAL
- eduGAIN is currently an operational security wild west...!
 - ◆ No consistency between IdPs on operational rules
 - ◆ In fact inter-IdPs eduGAIN authentication has not been used widely in production so far... but a moving target
- How can we open access to our services without trusting identities and a trust framework? Still an open question...
- 2 main issues
 - ◆ Operational security: incident handling, coordinated incident response, exchange of confidential information between actors
 - ◆ Privacy and traceability: must comply with EU policy on privacy and personal data protection, "user consent" no longer sufficient, must find the right balance between logging all (current practice) or nothing (liability issues of collecting data), WLCG AUP and data protection policy must be reviewed to comply with new EU policies

Several ongoing efforts

- SCI: an EUGRIDMPMA initiative, many participants from different e-Infra
- Sir-T-Fi: largely based on SCI work, SCI members + Internet2 + InCommon + eduGAIN + REFEDS
- FIM4R
- AARC: H2020 proposal on AAI
- All these projects try to address the 2 main issues
 - ◆ Operational Security: propose an incident response procedure, a trust framework and a minimum set of (verified) requirements on IdPs and SPs
 - ◆ Privacy/traceability: propose a trust framework and guidelines, compliant with EU/US (conflicting!) policies and allowing safe operations, ensure convergence/compatibility between participants

GEANT's Data Protection Code of Conduct

- Encourage IdPs to release attributes
- Push eduGAIN SPs to enforce basic data protection practices

- Proposal that WLCG should endorse this CoC
 - ◆ Already endorsed by several communities

Conclusions

- Lots of trust/policy issues: not only technical
- Web use case almost sorted out: required a lot of work
- Need to adapt to the new trust and policy landscape

Maria: VOMS future? Probably difficult to adapt to federated identities...

- VOMS is the source of attributes today. This is the very critical piece of our AAI current infrastructure
- In the future, we have to see if a similar service more integrated with IdP infrastructure will emerge...

Maarten: EU/US legislation incompatibility is major threat for federated identity, WLCG should take care of this topic

Oliver: IOTA-profile compliant CA availability? How to convince site to trust it?

- One prototype set up at CERN
- This prototype IOTA CA is not yet IGTF accredited
 - ◆ IOTA profile also has to be endorsed by WLCG (and EGI): planned next year

M. Alandes asks about CERN Data Protection Policy and WLCG

- S. Lueders explains it affects all data stored at CERN.
- An assessment will be done: in principle it should be aligned with all the other EU policies and codes of conduct mentioned during the presentation.

O. Smirnova asks how eduGAIN will be integrated with computing services. She believes this will be more complicated for sys admins.

- R. Wartel explains this will be much easier for the users and the technical aspects will have to be of course understood.
- S. Lueders adds that this makes service management easier too since there will be a generic way to define who is authorised to use the service.

Transition issue: how to manage users with several accounts

- Probably no generic solution but per service

Actions in Progress

Ops Coordination Report - M. Allandes

Dates for next meetings defined until the end of the year: see slides or Indico

ARGUS support: critical situation

- No clear commitment from different partners to take over from SWITCH
 - ◆ SWITCH has now ZERO effort on ARGUS
- Mid august proposal to create an ARGUS collaboration sent to INFN, NIKHEF, CESNET and Helsinki Univ: positive reaction but no commitment
- Urgency to find a solution: problem in the last (unofficial) ARGUS rpm, used at CERN, intended to fix problems in the last official one...
- Follow-up at next MB

T0

- SL5 decommissioning on Sept. 30th
 - ◆ Deadline for feedback: Sept. 14
- Plan to decommission AFS UI: work in progress to understand use cases
- FTS2 decommissioned
- CVMFS stratum 0 upgraded to 2.1 for all VOs

Experiments

- Smooth operations during the summer for all VOs
- CMS planning AAA tests at a larger scale soon: sites may see increased activity
- Atlas rucio commissioning ongoing: may also create additional load

Concluded TFs

- Tracking Tools Evolution
- FTS3 integration: now in production for all VOs

SHA2: hard deadline end of November

- Failing sites ticketed (44) but some already OK
- SAM tests will start using the new VOMS servers on Sept. 15
- Experiment workflow assessment to be done

WMS decommissioning

- SAM will start to use Condor submission on Oct. 1st

Multicore deployment: extended to address passing of parameters to batch systems

Network and Transfer metrics: kickoff last Monday

- Still 9 sites with incorrect perfSONAR version
- New US funded project: PuNDIT
 - ◆ Coordinated by Shawn
- More detailed report planned at next GDB

Information System Update - M. Allandes

BDII: 3 releases since April (last update)

- Minor fixes, mostly glue-validator fixes

New beta version of ginfo, lcg-infosites replacement for GLUE2

- **Give it a try**, report problems to project-grid-info-support@cern.NOSPAMPLEASE.ch

GLUE2.1: GLUE2 extension being discussed

- New objects and attributes to better describe cloud resources and GPU resources
- Intended to be backward compatible with GLUE 2.0
- Will require of a new version of glue-schema package: not before end of the year
 - ◆ WLCG not really interested by these extensions

BDII overload incident in August: tuned LDAP configuration to better handle it

- To be released soon for deployment at all top BDIIs

BDII deployment status stable

- All EGI + OSG sites
- Around 400 BDII instances: 323 site BDII + 79 top BDIIs

SW tags clean up campaigns: from 195000 one year ago to 66000 now!

- Reduced by More than 20% the size of BDII data

glue2-validator runs as a Nagios probes: automatic ticketing in case of error

- Now steered by EGI
- More than 300 ticketed sites, 290 fixed
- Now: mostly info message, error/warning at a very low level
 - ◆ Mostly default value problems (e.g. 444444 waiting jobs): default changed from undefined value
- Some sites having problems during a very long period

SRM vs. BDII consistency

- ATLAS: similar check to what was done for LHCb, integrated in dashboard
- Numbers well aligned, main inconsistencies related to known issues
 - ◆ Means that BDII info can be used by VOs as an alternative to SRM

GLUE2 Storage information providers fixes

- CASTOR/EOS: since August
- DPM: issues identified, fix expected in October
- StoRM: issues fixed
- dCache: no known issue

OSG and GLUE2: no change, no plan until experiments require them

- Alessandro: experiments need an incentive for GLUE2 as GLUE1 has been mastered and limitations work around to provide what is needed. New resource types (cloud, storage http endpoints...) present in the BDII could be one as this is a lot of work to add them to experiment specific information system solutions.
- New CE may be an opportunity to make progress as there is currently no information provider for it: Alessandro in contact with OSG colleagues to push them to write a GLUE2 provider
 - ◆ Would help a lot as storage is mainly dCache and dCache already has a GLUE2 provider

Alessandro (ATLAS): BDII remains a critical source of information

- Importance of the validation work done: experiments can now be confident that they can rely on the BDII information
 - ◆ SRM-BDII consistency assessment is a good example: experiments can now be confident in using BDII rather than SRM
- Maria: validation work done only for GLUE2 but also benefits to GLUE1 as the providers are generally the same
 - ◆ Nevertheless GLUE2 validation may be one incentive to migrate

AGIS as a generic experiment IS

- Maria: no real activity since a prototype was delivered to CMS
- T. Wildish: evaluated positively by CMS (several CMS service development groups, including PHEDEX) but higher priorities to address in the short term to be ready for Run 2
 - ◆ Still a plan to use it

SL/CentOS Discussions

(coordination problem: see slides attached to the agenda but not presented at the GDB)

CERN started to deploy for evaluation some CERN CentOS7 machines

- Currently only the infrastructure team
- Using the CentOS rpms
- Good collaboration with CentOS people: CERN contributed experience with Koji to help setting up the build system

Cloud pre-GDB Summary - M. Jouvin

See agenda[↗] and summary:

WG explores possibility of using clouds as a replacement of CEs.

- Some progress done in different areas:
 - ◆ Machine/Job features TF
 - ◆ Accounting done by EGI Cloud in APEL
 - ◆ vcycle and Openstack fair share scheduler initiatives
 - ◆ Cloud technology seen as more pervasive technology, no MW development required by the community
- Idea is to foster on going work to study how clouds could replace CEs: the goal is to come up with realistic milestones to achieve this in shared clouds

Target shares

- Achieve something similar to fair share in batch systems: deploy vcycle (already in a few UK sites) and openstack fair share in some sites
- vcycle: Vac concept in a cloud, implemented via a service running as a cloud client and doing the VM provisioning according to VO target shares
 - ◆ Currently for OpenStack but EC2 and OCCI being added, allowing to work with any cloud MW
 - ◆ Several deployment in the UK
- FairShareScheduler: request queueing, fairshare-based scheduling of request
 - ◆ Based on SLURM scheduler
 - ◆ Being tested at Bari
- Both solutions potentially complementary
- Milestone: wider deployment and usage of both solutions

Accounting

- Work @CERN[↗] on benchmarking VMs presented
 - ◆ VMs classified according their "HW characteristics" (HW type) obtained from live information (CPU type, number of core, memspeed)
 - ◆ HS06 benchmark run on a large number of instances for each HW type and a table built, with one entry per HW type, with some sort of mean of the results
 - ◆ A VM can query the table to know its performance: this could be exposed by machine/feature mechanism.
 - ◆ Showed accuracy better than 10% at CERN
- Some features missing in APEL for cloud accounting, in particular publication of number of cores and HS06 perf of the VM
 - ◆ To be fixed in the next 6 months

Security

- Liability in case of problem with a VM: significant shift from site to VO, as the VO is responsible to build/manage/endorse the VM image
- Incident handling: not such a big change compare to today (pre-glexec) incident handling. Need cooperation between sites and VOs but proved to work. Must ensure that the necessary information is collected (same one as the grid basically but VM information is lost when it is stopped)
- Proposal: a TF with VO and site experts doing a traceability gap analysis and proposing solutions compatible with existing manpower.

- Avoid to work around problems after going to large production: try to anticipate operational/policy changes needed, always easier.

Stefan: machine/feature implementation for the cloud is progressing, first release was expected during the summer, slight delays to fix inconsistencies between the 2 implementations (OpenStack and OpenNebula)

A new meeting during the winter.

Site Availability Change Proposal - P. Saiz

Part of the monitoring consolidation effort

- Combine existing tools to reduce the operational costs
 - ◆ Common database for SAM and SSB
 - ◆ Common metrics and combination engine
- Add flexibility

Current reports

- T0-T1 report
- T1 6 month historical report
- T2 report
- T2 6 month historical report
- MB reports

Recent changes in the new monitoring tools

- Experiment defines topology
- Distinguish metrics by FQANs
- Combine all services in availability calculation

Change in the availability formula

- Calculation algorithm explicitly mentioned in the report
- All SRM endpoints at sites taken into account for several VOs

Proposed changes

- T1 6 month summary: replace federation name by site name to be consistent with the other report
- T2 report: replace global (all VO) Phys/Log CPU available by resources offered to the VO
 - ◆ Also available in REBUS: check numbers
- Monthly report: use the same kind of plots as for the other reports
 - ◆ Also change the title to "T0-T1 summary"

Proposed transition

- Continue to do SAM2 and SAM3 report until the end of the year, with SAM3 the primary since October

Discussion

- Remove phys./log. CPU from reports: no importance, quite often wrong numbers
 - ◆ Only important information is HS06 and it is cross checked with accounting
 - ◆ Could discuss in the future removing them from REBUS too
- Possibility to use HammerCloud tests rather than WMS ones
 - ◆ Submission: WMS is almost legacy, move to Condor direct submission almost completed
 - ◆ HC results could be easily inserted into availability calculation but there is no plan to do it presently

Data Management Discussion - O. Keeble, F. Furano and W. Bhimji

Xrootd v4: major improvements, many new features including IPv6

- Do experiments need it deployed soon?
 - ◆ Impact on SW providers relying on Xrootd client: storage, ROOT...
- Today a restriction with the client: cannot have both v3 and v4 on the same machine and there are not API/ABI compatible
 - ◆ Being worked on to solve this restriction
- CMS: would like to see IPv6 support (thus v4) deployed by the end of next year to use it in AAA
- Server deployment: needs to be done on every servers in the Xrootd cluster at the same time
 - ◆ No experience with rolling upgrade
- EPEL availability: required for DPM, FTS...
 - ◆ M. Ellert in charge of building it... expected soon
 - ◆ In the meantime, requires to maintain 2 versions of Xrootd-related components
 - ◆ No clear if the new version will be xrootd or xrootd4: Lukasz pushing for xrootd4, WLCG may push into this direction

SRM for metadata campaigns

- Recent SRM high load from Atlas deletion campaign: seems to have worked but a risk of creating problems
- Are they any other experiment plans for massive metadata campaigns through SRM?
 - ◆ CMS: massive deletion done locally at sites, not through SRM
 - ◆ LHCb: still relying on SRM for deletion, moving to xrootd for data access
 - ◆ ATLAS: not a new activity, need new data management to use SRM alternative
- Future metadata requirements by experiments
 - ◆ CMS: planning to expand its data moving interface at the timescale of end of Run2. Will look at other protocols during this work
 - ◆ ATLAS: very encouraging numbers with WebDAV, no change in the plan to move
 - ◇ 85% of ATLAS sites are WebDav ready
 - ◇ Endpoint used (and thus protocol) taken from GOCDB: is WebDav endpoint in GOCDB?
- FTS developers have been asked to add ability to do (massive) deletion through FTS: being added
 - ◆ Could use gridftp or other protocols

How to progress with interface rationalization?

- Plan a dedicated meeting (pre-GDB) to come up with milestones that could be followed by Ops Coord

Multicore: Dynamic Partitioning with Condor at UVic - F. Berghaus

IaaS clouds interfaced with HTCondor batch scheduler through CloudScheduler

- Users (or pilot framework) submit Condor jobs
- Take care of starting appropriate VMs based on jobs in queue
 - ◆ Typically CernVM3 images contextualized through CloudInit and Puppet, with dynamic Condor slot configuration (partitionable slot)
 - ◆ Rely on CVMFS for OS (microCERNVM) and SW
- Can support many different clouds
- Used for several different experiments including ATLAS and BELLE
 - ◆ ATLAS cloud (North America, Australia, CERN): 1.2 M jobs last year

Dynamic batch slots

- Jobs specify its CPU requirements
- defrag daemon used to recover fragmented slots

Condor group used to define and implement target shares

Dynamic Squid discovery for CVMFS through Shoal

- Shoal instance at UVic with connected squids at UVic, TRIUMPH, CERN and Oxford

EMI Dynamic Federation as the data access solution as no storage provided in the participating clouds

- Stateless, no persistency
- Very efficient and scalable
- http-based: allow support of non HEP communities, a requirement
- Data access either from the nearest endpoint or all the endpoints in //

-- MariaALANDESPRADILLO - 10 Sep 20

This topic: LCG > GDBMeetingNotes20140910

Topic revision: r5 - 2014-10-08 - MichelJouvin



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