


# Table of Contents

<b>Summary of GDB meeting, March 13, 2013 (KIT)</b> .....	<b>1</b>
<b>Agenda</b> .....	<b>2</b>
<b>Welcome - M. Jouvin</b> .....	<b>3</b>
<b>EMI3 Highlights - C. Aiftimiei</b> .....	<b>4</b>
<b>WLCG Information System - L. Field</b> .....	<b>5</b>
<b>Pre-GDB summary - M. Jouvin</b> .....	<b>6</b>
<b>Storage Accounting - J. Gordon</b> .....	<b>7</b>
<b>Actions in Progress</b> .....	<b>8</b>
Ops Coord Report - J. Flix.....	8
FT3 - M. Simon.....	9
<b>Storage WGs Report</b> .....	<b>10</b>
Storage Interfaces - W. Bhimji.....	10
IO Classification and Benchmarking - D. Duellmann.....	10
<b>Storage at GridKa - X. Mol</b> .....	<b>12</b>
<b>Federated Identity Pilot - R. Wartel</b> .....	<b>13</b>

# Summary of GDB meeting, March 13, 2013 (KIT)

# Agenda

<https://indico.cern.ch/conferenceOtherViews.py?view=standard&confId=197801> 

# Welcome - M. Jouvin

See slides.

- April meeting **canceled**
- May be a pre-GDB in May: to be confirmed mid-April
  - ◆ Follow-up for cloud work/discussion? Other proposals/needs?
- Topic on the list for future meetings
- Forthcoming meetings

## **Still need more people to participate to the note taking rota**

- Hardly sustainable with the current number of people...

Michel would like to get some feedback about meetings off-CERN: are they to be continued? New sites volunteering?

- CERN participation reduced for meetings off-CERN: always many meetings conflicting locally
- Keep in mind that March meeting conflicts with Geneva motor show: prices for travelling and accomodation much higher
- Contact Michel directly or through the GDB list

# EMI3 Highlights - C. Aiftimiei

## History

- EMI1: 148 products, 23 updates
  - ◆ 4 critical updates
- EMI2: 107 products, 9 updates
  - ◆ Only 1 critical update
  - ◆ In // with EMI1

## EMI3 released March 11

- 50 products
- Full support of SL5 and SL6, partial support of Debian6
- Not yet released: ARGUS-EES, CREAM, FTS4, HYDRA
- Security: EMI Security Token Service (STS) easing use of X509 infrastructures, Common Authentication Library (CAnL)
- Compute services: implementation of EMI ES across all implementations
- Main known backward incompatibilities:
  - ◆ APEL Publisher + schema change
- Major VOMS Java API redesign + clients
  - ◆ Upgrade possible
  - ◆ VOMS no longer configured by YAIM

## Discussion

- YAIM future : probably none... some products (eg. VOMS) already moved off YAIM. Compensate with better documentation...
  - ◆ Risk of more load put on sites
- EPEL packaging issue due to non backward compatible changes: will lead to renaming many RPMs (adding a major version number as part of the RPM name, like it was done for EPEL)
  - ◆ Risk of fetch-crl syndrom: EMI missed that fetch-crl is now fetch-crl3 and continued to distribute an obsolete version...

# WLCG Information System - L. Field

EMI2 now published by all EGI sites: only OSG sites are not publishing

- Need a risk assessment about consequences of OSG not publishing: when GLUE2 requests may come? What are the mitigation possibilities and cost?

GLUE2 validator: final version almost ready

- Not yet in EMI3 but used to find bugs in EMI3 info providers
- Based on EGI GLUE2 profile current draft

GLUE2 validation status

- 273 errors: all related to wrong total published for storage
- 30K warnings! missing mandatory attributes or wrong types
- 31K info!

ginfo is the GLUE2 replacement for lcg-info(sites)

- Early feedback received that experiments wanted more than static information on service endpoints
- New 1.0.0 release ready and soon in EPEL

EMI3 BDII: mostly changes for being able to release it in EPEL

WLCG Global Information Registry

- Current situation requires 2 complementary sources aggregated by every experiment
  - ◆ Pledged resources: EGI GOCDB4 and OSG OIM
  - ◆ Actual service states: BDII
- Idea is to build a global registry merging these informations that can be queried by experiments
  - ◆ Built on top of REBUS taht already contains service pledges
  - ◆ Using GLUE2 as the internal integration model but independent of the GLUE version
  - ◆ OSG and EGI
  - ◆ Avoids duplication of effort, improves information quality
  - ◆ Can produce both plain text and JSON feeds
- Next step: validate prototype with AGIS
  - ◆ LHCb also interested for CE discovery
- Timescale for production cannot be defined yet

OSG GLUE2 support and risk assesment: MB should ask OSG to provide information on what is the effort required for them, what is the possible time frame to implement it to serve as an input to the risk assessment

# Pre-GDB summary - M. Jouvin

See slides and pre-GDB summary.

# Storage Accounting - J. Gordon

## Status of StAR

- Dcache and DPM publishing from test systems
- StoRM missed the EMI3 cutoff
  - ◆ Italy NGI working on publishing StAR records from BDII
- SSM transport

## Issues from testing

- No GOCDB sitename: may be an artifact of test systems
- No ResourceCapacityAllocated

## Open issues: frequency of publishing

- Sites could collect frequently but publish daily
- Repository could to monthly average like for CPUs

## Accounting portal: proposal made for storage accounting data based on fake data

- per VO, per country, per month...
- Allocated vs. used
- Ian: Need to get some real data in to produce a first set of reports to get feedback

## Reports: when storage accounting is deployed at most sites, report for T1 and T2 could contain storage data

- For T1 replace manually filled data by those collected

## Slow progress, not much data yet

## Action on John/Cristina: check possibility to deploy StAR publisher on existing versions without deploying EMI3

- Would help to get real data quickly

## EGI taking requests for features in the accounting portal

- Contact John directly or send emails to GDB list

## Possibility to deploy StAR providers on EMI-2 ?

- If only EMI-3, risk of delay before we get any real data
- John thinks there is not much dependency on EMI-3 but in some products, there may be DB schema changes. To be checked for each storage products



# Actions in Progress

## Ops Coord Report - J. Flix

Improving communication with T2 sites is a **priority**

- Get more people from sites participating/coordinating TF

Reduced frequency of daily meetings: twice a week

Agreed monthly report by EGI at the Ops Coord fortnightly meetings

MW deployment TF

- WMS fix for EMI2 UI crash available
- Site **required** to install the latest Frontier/squid RPM by end of April
- dCache 1.9.12 support supposed to end end of April. May be extended as T1s will not be able to meet the deadline
  - ◆ Non trivial migration

CVMFS: still missing sites. Deadline: end of April.

glexec

- 193 CEs
- CMS wants to have 90% of its CEs with glexec enabled by July 1st

SHA-2: new CA foreseen to be available soon

Xrootd deployment going on for CMS and ATLAS

- Emphasis now put on monitoring infrastructure: proposal by TF expected soon
- work going on on consumers

Tracking tools: deadline for JIRA migration set to end of 2013

perfSONAR

- Growing deployment: must now be treated as a prod service
  - ◆ Nagios probes now available
- TF testing new 3.3 version
- Testing "disjoint meshes"

SL6 migration TF just created

- First meeting expected next week
- Good input provided by experiments
- Deadline for bulk migration set to October 2013

UMD release plans

- Proposal to create a UMD 3.0 release with separate repositories, end of April
  - ◆ Avoid automatic updates of sites

- ◆ Feedback expected before March 15: is there really any other option?

We again see the benefit of the work done by the Ops Coord: need to find new site experts participating

- \*An action for all the T1/country representatives to help identify/motivate new people
- It is a best effort contribution to the community: sites also benefit from this involvement

## FT3 - M. Simon

FTS3 status

- Currently at the same level of functionality as FTS2
- Tested by ATLAS with Russian T1 and by CMS debug transfers
- Scaling tests will be reported at next FTS3 TF, next week
- Deployment plans not yet clear
- Deployed at several sites (CERN, RAL, PIC, ASGC...)

Main new features

- Channel less, SE centric configuration: no quadratic effect with the number of SEs
  - ◆ One asked if it would still be possible to put some limits on some specific pairs: currently not supported but can be discussed in the future if needed
- Multiple database backends: Oracle and MySQL currently
- Session reuse for gridftp/gSOAP
- Support for xrootd and httpd through plugins
- VOMS based authentication
- Improved monitoring: currently publishing the same info as FT2, in the future will send message for every transition allowing to federate FTS instances

Several new requests on the list, mostly from ATLAS

- REST interface for submission/status
- Multiple source/destination submission

Development process based on an iterative process with frequent interactions with sites and experiments

- One meeting of the FTS3 TF every 3 weeks

Release in EMI3 expected end of March

- In EPEL in April

Deployment still to be worked out but it is envisioned running the whole FTS service with only one instance at CERN

- With may be a failover instance in the US

# Storage WGs Report

## Storage Interfaces - W. Bhimji

Experiment's Data management plans

- CMS has no issue with non SRM sites
- ATLAS has some issues but most of them to be solved in Ruccio
- LHCb has similar requirements as ATLAS

Area requiring some development identified

- Main issues related to space token handling (space report, upload)

Recent developments

- FTS3 now in pre-production providing support for non SRM systems (xrootd and httpd)
- CMS requires xrootd at all sites
- ATLAS will soon need xrootd and WebDav
  - ◆ WebDav initially for Ruccio mass renaming
- In fact the addition/prevalence of xrootd/httpd after EMI upgrade is adding complexity in the short term as legacy interfaces are still used
- gFal2 providing client side abstraction that will help to replace some the server-side abstraction provided by SRM

General feeling is that move off SRM is doable but still to be done...

- Group should continue to monitor issues and progresses

Future interface directions

- Need to support both direct access and staging in the short/medium term
- xrootd and http are the more mature options
- rfio is the most obvious candidate for retirement
  - ◆ Forced (gradual) retirement of rfio is supported by the WG: who is in charge of following it
- Proposal of a concept of "core protocol": xrootd for the time being
  - ◆ required protocol at each site
  - ◆ Does not prevent the use of other more performant/appropriate protocol at a given site

Discussion: impact of xrootd requirement on StoRM

- Not integrated solution: a separate access to the same files
- Currently only readonly

## IO Classification and Benchmarking - D. Duellmann

Progress made on identification of what should be measured and what should be calculated and plotted

- Standardizing the metrics is critical to be able to share and compare measures on different systems

Discussions on how to analyse the data: similar work as standard physics analysis...

- Need to identify clusters of jobs with similar I/O patterns

- Pure file copy (xrdcp): sequential I/O, low sparseness, low randomness, not much influenced by client CPU
- LAN analysis: sensitive to sparseness and randomness, impact of client side CPU
- WAN analysis: same as LAN analysis + impact of the WAN throughput

File updates: need still to be understood

Read performances: analysis started based on ATLAS and CMS access at EOS (CERN)

- Different profiles
- ATLAS throughput showed a significant number of jobs throttled at 3-4 MB/s: is it something induced by ATLAS framework (decompression...???)
- Input from experiment and ROOT experts required for interpretation
- Patterns seen at CERN are likely not to be observed at other sites but methodology can be reused

# Storage at GridKa - X. Mol

GridKa supports all 4 experiments: until 2009 one single instance for all of them, then splitted into 1 instance per experiment

- ALICE: xrootd rather than dCache
- but similar dCache setup for every instance: separate read and write buffer, except for LHCb which can read from write buffer (due to limited disk size)
- Read mostly done with dcap (unauthenticated)

Xrootd: 2 instances for ALICE, 1 disk-only, 1 for tape buffering

- Backend is a GPFS cluster

Tape connection: TSM with eRMM

- eRMM is a library virtualizing access to the libraries
  - ◆ all tape libraries virtualized into one
- A restricted set of machines allowed read access to the tape system for better optimisation and reduced administration overhead
  - ◆ Will do the same in the future for writing tapes

HW for disk servers: medium number of cores (16+) but a lot of memory (32 to 64 GB)

- 10 GbE network connections
- DDN RAID6 as storage backend
- Deployment with ROCKS and configuration management with cfengine2
  - ◆ Looking at cfengine3 or Puppet for the future

Monitoring done with Icinga

- Moved from Nagios 1 year ago
- On-call service managed by Icinga
- Performance collected/displayed with Ganga

Challenges for the future

- dCache update
- https and xrootd deployment
  - ◆ Have to join the ATLAS FAX
- Improve tape throughput significantly

# Federated Identity Pilot - R. Wartel

Tried to refine foreseen use cases

- ALICE mainly interested in the web use case
- ATLAS equally interested by CLI and web
- LHCb interested mostly by CLI
- CMS: no feedback, no interest?

Test EMI STS instance installed at CERN

- Support a test IdP in Helsinki that supports the ECP profile
- Support the WS-Trust endpoint of CERN ADFS

CLI use case: the critical piece is ECP but only very few IdPs support it worldwide

- Externely difficult to convince hundred of IdPs to adopt it...
- ECP deployment will probably grow... but probably slowly
- Not real alternative without significant costs and compromises

Possible options

- Do not investigate CLI use case for now: focus on the web use case
- Continue with ECP pilot and wait for ECP adoption
- Investigate ECP alternatives based on CILogon
  - ◆ Download of a certificated from a web site
- Central credential repository: but a lot of issues and complexity

Conclusions

- Concentrate on the web use case that will allow to get operational experience and solve a few useful use cases
- Try to implement CILogon-based ECP alternative until there is enough deployment of ECP (that will take years)

---

This topic: LCG > GDBMeetingNotes20130313

Topic revision: r1 - 2013-03-17 - MichelJouvin



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback