

Table of Contents

| | |
|---|-----------|
| Summary of GDB meeting, February 12, 2014 (CERN)..... | 1 |
| Agenda..... | 2 |
| Welcome - M. Jouvin..... | 3 |
| HEP SW Collaboration - I. Bird..... | 4 |
| IPv6 Update - D. Kelsey..... | 5 |
| Future of Scientific Linux - T. Ouveley..... | 6 |
| Actions in Progress..... | 7 |
| Jobs with High Memory Profiles - J. Templon..... | 7 |
| Ops Coord Summary - S. Campana..... | 7 |
| SAM Test Scheduling - L. Magnoni..... | 8 |
| WLCG Transfer Dashboard - A. Beche..... | 9 |
| Data Preservation Update - J. Shiers..... | 10 |
| WLCG Monitoring Consolidation - P. Saiz..... | 11 |
| LHCOPN/ONE Evolution Workshop Summary - E. Martelli..... | 12 |
| perfSONAR - S..... | 13 |

Summary of GDB meeting, February 12, 2014 (CERN)

Agenda

<https://indico.cern.ch/event/272618/>

Welcome - M. Jouvin

HEP SW Collaboration - I. Bird

Kickoff meeting, April 3-4, is really about the organization of the collaboration

- The 3 days before are the technical discussions: Concurrency Workshop, March 30-April 2

Computing will not scale easily to post LS1 LHC conditions (higher pileup, reduced bunch spacing): need to radically evolve the SW

- Need parallelism everywhere, at each level: have to tackle the micro-parallelism (vectors, instruction pipelining, new architectures...)
- Huge effort, lots of training will be required, benefit will not come immediately: target libraries and toolkits with the highest impact on the community (GEANT, ROOT, ...)
 - ◆ Will require an incremental approach with a lots of checking on the physics output
- Already a 2 year effort on this topic that helped to build some consensus on what should be done: Concurrency Forum

GEANT and ROOT are 2 key tools for HEP: try to have more commonalities

- Need a more modular infrastructure: a lot of work to achieve it!

SW Collaboration

- Discussed with Director of Research, IT and PH
- Initially focus on HEP community but opened to others
- A formal collaboration production open scientific SW, with an agreed governance
 - ◆ Build on the work in concurrency forum
 - ◆ Framework for increasing collaboration and give credit and recognition for the work done: reuse successful WLCG experience
 - ◆ Publish roadmaps and priorities: will help to get additional funding
 - ◆ Establish a team of experts to provide consultancy and help in SW development (optimisation, debugging...)
- Idea of a follow-up meeting in the US (at least outside Europe) later in the year

Discussion

- Philippe: what is the intent regarding LCG Application Area? Will it be replaced by the new collaboration?
 - ◆ Ian: too early to say exactly... but this may be one possible outcome... if it makes sense!
- Concerned people are not in GDB but GDB members must be ambassadors for the initiative, convincing appropriate people that this is an important and open initiative!
 - ◆ Invitation has been widely circulated...

IPv6 Update - D. Kelsey

CERN news

- Progressing well with IPv6 deployment: firewall, DNS ready, campus-wide deployment foreseen in March
 - ◆ Still some issues with dhcpv6
- OpenStack Havana still not controlling well IPv6: CERN developing its own control

Other sites

- DESY ready for deployment
- KIT has a test cluster
- Imperial very active with dual-stack testing

Applications

- Application survey making progress, a few critical application without IPv6 support: dCache, EOS, Frontier
 - ◆ <http://hepex-ipv6.web.cern.ch/wlwg-applications>
 - ◆ Still a lot of unknown
 - ◆ For some clients, it requires the last version (e.g. GFAL2)
- Xrootd v4 (required for IPv6 support) still not available?
- dCache: 2.6 not IPv6-ready, 2.8 successfully tested at NDGF
 - ◆ Caveat: gridftp1 through FTP door doesn't work
 - ◆ gridftp is the difficult part of IPv6 compliance: other protocols (e.g. http) are easier

Transfer testbed

- Successful mesh of transfers before CHEP but several unrelated issues afterwards as people were not looking at it any longer
 - ◆ Not IPv6 issues per say

Experiment readiness

- ATLAS (A. Dewhurst)
 - ◆ AGIS: flags added for IPv6-ready storage endpoints
 - ◆ Frontier is a problem: 2.8 not supporting IPv6, 3.1 should be ok

Next steps

- PhEDEx/SRM/dCache tests on testbed
- Move to use some production endpoints at volunteer sites
 - ◆ Need clearly defined use cases from experiments about the IPv6-only clients
 - ◆ Some sites ready to participate but only if accepted by experiments and funding agencies that availability may suffer (directly or indirectly): similar to MW readiness testing, work in progress by Simone, to be discussed at MB
- IPv6 pre-GDB on IPv6 management on June 10: aimed primarily at T1s but others welcome
 - ◆ Encourage IPv6 support and move to dual-stack WLCG services
 - ◆ Try to do as much as possible before restart of data taking
- Also monthly meetings by Vidyo (13/2, 13/3) and a 2-day F2F meeting in April (post-GDB)

Future of Scientific Linux - T. Ouveley

Beginning of January, CentOS announced that it is joining its forces with RH as part of the open source team

- Governance published
- Remains a separate group and product at RH: remains an upstream vendor

Key changes

- RH is offering to sponsor the build system and content delivery network
- git.centos.org remains the central place for CentOS developments and is open to everybody
- Possibility of SIGs

Impact on SL for SL5/6

- May be no longer SRPMS: direct use of Git sources, not really a problem
- Hope to make the changes transparent

SL(C)7: 3 options

- Continue to build from source as previously
- Create a Scientific CentOS SIG
 - ◆ SIG board defines the rules and can diverge from core
 - ◆ Right to override packages, use specific logos...
- Use CentOS7 as SL7: in fact now SL has just very few diffs with CentOS (in particular openAFS)
 - ◆ Main issue was the process openness but will improve greatly with recent evolutions
- No decision yet
- RHEL7 final expected this summer, CentOS7 beta soon

Actions in Progress

Jobs with High Memory Profiles - J. Templon

Many jobs in all LHC experiments requires up to 4 GB of vmem

- Mostly ALICE has a significant number of jobs requiring more than 4 GB of vmem
- NIKHEF uses a pvmem limit of 4 GB

Why to use PVMEM

- Translate into a ulimit in the process
- Gets an out-of-mem signal rather than a kill: gives a chance to handle it
- doesn't take into account overhead outside app space (wrappers...)
- Limit set at the queue level

If a user needs more than 4 GB, several ways to request it

- In particular through JDL... but in fact result is not necessarily consistent (see slides)
- User asks for mem rather than vmem: not necessarily a major problem has WNs have not a good amount of memory

Big Damned Jobs vs. many process

- MAUI doesn't properly understand a request for 32 GB and 8 cores: understood as 32 GB per core (256 GB!)
- Not clear how to work around it and if it is worth the effort

Ops Coord Summary - S. Campana

~40 people, good representation of both sites and experiments

- Lots of discussions

glexec

- CMS SAM tests not yet critical, due to the SAM test scheduling issue
- Still 20 sites haven't deployed it yet but mostly not CMS sites

perfSONAR: the core WLCG tool for network monitoring

- Still 20 sites without it
- 20 sites with obsolete/old versions

Tracking tools

- Savannah-to-JIRA migration in GGUS still not done because of missing features in JIRA: recent agreement they were not critical and we'll leave without them
 - ◆ Fixing JIRA is out of the scope of the TF

SHA-2

- Progress in VOMS-admin: experiments need to adapt to it

- Will start soon the campaign to update VOMS server host cert
- When SHA-2 migration is over, will tackle RFC proxy compliance

Machine/job features

- Prototype available for testing, testing started at some sites
- Good collaboration with OSG to integrate bi-directional communication

MW Readiness

- Use experiment testing framework, e.g. HammerCloud
- Sites will deploy "testing production instances": monitoring will distinguish them from normal production instances
- A demonstrator being setup in Atlas
- Discussing how to reward sites: discussion at next MB

Baseline versions

- Enforce defined baseline versions: need a tools for reporting sites failing baseline versions
- Proposal to look at Pakiti as a way to report baseline compliance

WMS decommissioning

- No change in plans: CMS/LHCb WMS will be shut down end of April

Multicore deployment

- ATLAS and CMS foresee different usage of multicore job slots: TF is not discussing which model is the best one
 - ◆ ATLAS: both single core and multi-core jobs, sites in charge of doing the backfilling
 - ◆ CMS: always using multicore slots, doing the filling optimization
 - ◆ LHCb working on an approach similar to CMS
- TF goal: find a way to handle both approaches
- MAUI is a great concern regarding multicore job support

FTS3 : consensus on development model

- 3 to 4 servers at maximum: shared configuration but no state sharing/synchronization

Common Commissioning Excercise (STEP14): consensus there was no need for it

- Production activity is at a sufficient level
- Specific commissioning activities coordinated by TFs, Ops Coord: concurrent testing for some specific cases still remains an option if possible and necessary

SAM Test Scheduling - L. Magnoni

SAM test framework based on Nagios to do the check scheduling

3 categories of tests

- Public grid services: FTS, SRM, LFC...
- Job submission: ability to execute a job in a given time window
- WN: ability to use various services from a WN

Currently WMS is used to submit job submission tests and WN tests

- Move in progress to Condor-G submission
- WN tests is the difficult part as it has to bundle in a job payload Nagios aspects, MTA....

Test timeouts: WN test needs specific groupe/role, inherited from job submission test

- Job submission may thus time-out too: will result in MISSING status rather than critical
- One timeout for each job state: in fact 4 different timeouts
- Timeouts represent 1/3 to 1/2 of the job submission test failures for CMS and ATLAS
- Most of the timeouts happening in waiting state at the WMS (~80% for ATLAS and CMS): should probably increase the timeout at WMS (45 mn currently)
 - ◆ Very low level of timeouts at site level
 - ◆ Waiting state in WMS is purely related to match-making operations
 - ◆ An overloaded CE sets its status to Draining where the requirement is for production state
- Concentrate effort on issues independent from WMS itself as the main goal is to move asap to CondorG submission

In the future we may want to decoupe job submission from WN tests

- glxec for WN tests
- Submit with experiment frameworks
 - ◆ HC is a good candidate to ship WN tests but not Job Submission tests
- Need to keep a generic framework with not too many differences from what we have today...
 - ◆ Looking at Zabbix as an alternative, more scalable scheduler
 - ◆ Having a different framework for WN tests and others may add dependencies between different infrastructures: not necessarily desirable

WLCG Transfer Dashboard - A. Beche

Original dashboard designed for FTS: difficulties to incorporate data for other transfer methods like Xrootd

- Some specific data not relevant to FTS
- Data duplication: tool specific dashboards + WLCG one

New architecture will be an aggregator for tool specific dashboard

- No data duplication
- Access to tool specific data through the tool specific dashboard
- Data retention policy depending on tools: FAX/AAA kept a long time to feed data popularity calculation
- New GUI views: map view

Data Preservation Update - J. Shiers

Things are going well!

- Full Costs of Curation workshop: HEP context, focus on medium term plans to ensure continuity with next experience
 - ◆ 100 PB today, 5 EB in 15 years
 - ◆ Start with a 10 PB archive now, increase every year with a flat budget
 - ◆ Estimated ~2M\$/year: affordable compared to experiment costs and WLCG costs (100 ME/year)
 - ◆ A team of 4 people could "make significant progress"
 - ◆ Need to get sites other than CERN involved in the budget effort/forecast, need from support from RRB

DPHEP Portal built over

- Digital library tools and services (Invenio, INSPIRE, CDS...): already well organized collaboration taking care of backward compatibility when there is a major technical evolution
 - ◆ Well funded in FP7 and will continue to be funded in H2020
- Sustainable SW coupled with virtualization techniques: mostly exists (e.g. CERNVM and CVMMFS) with sustainability plans
 - ◆ Sustainable bit preservation as a service is also fundable: based on RDA WG RFCs for interfaces
- Open data/access: more standard description of events....
 - ◆ Well identified goal for H2020
- DPHEP portal built in collaboration with other disciplines and a proposal is being prepared

Next actions

- Launch (and drive) a RDA WG on bit preservation at next RDA meeting in March
- Prepare an H2020 project on bit preservation for the sept. call: funding to start in 2015
- Strengthen multi-directional collaboration (APA, SCIDIP-ES, EUDAT)
- Agree on functionality for DPHEP portal
- Do a demonstrator of using CERNVM to preserve old environment for current experiments (CMS)
- (Self-)Audit, reusing OAIS, ISO16363
- Establish core team of ~4 people this year

SWOT analysis (see slides)

- Strengths and opportunities clear
- Weakness: manpower needed but strong support statements from various bodies + H2020 funding opportunities
- Threats: commitment from funding agencies, non-CERN experiments/sites agreement

WLCG Monitoring Consolidation - P. Saiz

A lot of work done, good feedback from experiments but site feedback weaker

- Prototype phase completed with a detailed report
- Now entering deployment phase

4 different sets of tasks

- Underlying application support (job, transfer...): basically support of the current system
- Running the services: main action was the move of services to Agile Infrastructure, almost done
 - ◆ Also the move of SAM EGI to GRNET
- Merging applications with similar functionalities or significant overlap: reduce the number of apps to maintain
 - ◆ Merging testing frameworks (Nagios and HC): still in discussion
 - ◆ SAM & SSB: very good progress, in particular for storage and visualization
 - ◆ Keep the 2 UI look and feel over the same storage/virtualization framework: SSB and SAM
- Technology evaluation: a good part of the work done during the evaluation phase, close collaboration with AI monitoring

Next steps

- Downtime in avail/reliability reports
- Report generation

LHCOPN/ONE Evolution Workshop Summary - E. Martelli

Very good participation

- ~60 people
- ~11 NRENs
- Europe, N.A. and Asia

WLCG viewpoint expressed by I. Bird

- Network has been working well and LHC needs is expected to fit into technology evolution
- Network is central to the evolution of WLCG computing
- LHCOPN must remain a core component to connect T0 and T1s
- Main issue is connectivity to Asia and Africa

Experiments usage may differ but conclusions are the same

- WAN more and more important
- Reduced difference between T1s and T2s
- Need for more connectivity at T2s

Sites: diverse situations, in particular with respect to LHCONE

- T1s happy with LHCOPN
- T2s will increase their connectivity: several in the US are planning 100G for WAN
- Main demands
 - ◆ Better network monitoring
 - ◆ Better LHCONE operations, in particular for troubleshooting problems

LHCONE P2P service: NSPs should be ready soon to provide a production service

- CMS may exploit the service: early prototype being developed for PHEDEX
- Sites don't have clear plan/need for it: usual over-provisioning vs. complexity balance
- Billing is very unclear
- NSPs interested to get WLCG testing the service

Actions

- Keep LHCOPN as it is: increase bandwidth if really necessary and affordable
- Several sites would like to move T1-T1 traffic to LHCONE as topology may be more optimal than LHCOPN (routing through CERN)
- Several sites planning to use LHCONE as their backup for LHCOPN: will often provide better perfs than a slow backup link
- LHCONE L3VPN
 - ◆ Improve support in particular with global tracking system and cross-organization team
 - ◆ Improve monitoring: remains based on perfSONAR
 - ◆ Improve usage efficiency of transatlantic OPN/ONE resources
 - ◆ Sites asked to announce only LHC prefixes: a requirement for bypassing firewalls, one of the key feature for LHCONE (science data only)
- LHCONE P2P: still lacking interested sites and developers, lack of time to really discuss further plans
- To be followed-up at next meeting in April (Roma)

perfSONAR - S.

Current situation

- perfSONAR PS 3.3.2 released Feb. 3: bug fixes, improvements
- Modular dashboard prototype orphaned but still in GitHub

Evaluating new dashboard solution: MaDDash by ESNET

- A product of the perfSONAR-PS project
- Missing abilities to edit mesh and metric numbers are not displayed at the highest level
- Drill-down capabilities to get more details
- Exploring OMD to complement MaDDash to do the perfSONAR service testing
 - ◆ Based on Nagios + Check_MK
 - ◆ Very promising to monitor PS services, still missing the capability to monitor the mesh configuration
 - ◆ Graphs created automatically with checks

Mesh configuration status

- 3.3 has all the functionality for the mesh built-in
- Plan to automate the mesh generation from GOCDB or similar sources: currently manual, labor intensive

WLCG deployment details

- Sites organized in regions
- Every site expected to deploy perfSONAR
- Every site expected to check PS results every day

Debugging network problems: still triggered by a human!

- PS clearly helps but not necessarily enough: need to correlate with other facts/metrics at site

Known issues not yet addressed

- 10G vs. 1 G (vs. 100G?)
- Improve end-to-end path monitoring
- Alerting: difficult to converge on what, when, who
- IPv6 monitoring: preliminary work at Imperial

Network metrics collected by PS may be consumed for other uses

- Characterizing paths with cost to optimize placement decision (ANSE projet)
- Select data sources based on network connectivity

OSG proposing to host a network service for WLCG

- Need to revise current metrics: are they appropriate? Are they enough?

PS-PS and PS-MDM should converge eventually

- Not clear if PS-MDM is still developed
- PS-PS will incorporate some views provided by PS-MDM

This topic: LCG > GDBMeetingNotes20140212

Topic revision: r7 - 2014-03-12 - MichelJouvin



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback