Table of Contents

December 2016 GDB notes

Agenda

http://indico.cern.ch/event/394789/ ⧉

Introduction (Ian Collier)

https://indico.cern.ch/event/394789/contributions/2392195/attachments/1388433/2114503/GDB-Introduction-2016121

Note: Feb Pre-GDB on benchmarking

Next WLCG Workshop

Discussion of WLCG Workshop location between Manchester and Naples. Nordic countries go on vacation 23rd June. Manchester looking better for accessibility and timing?

Hold Naples in reserve as first choice for the next non-CHEP meeting.

Ratify at MB next week.

Downtime policies: A proposal (Maria Alandes Pradillo)

https://indico.cern.ch/event/394789/contributions/2392230/attachments/1388375/2113909/LongShutdowns.pdf ⧉

Q: Data migration, experiments decide on their own?

A: I guess yes, would like to have the time to do it

Ian Bird: take more than a month

Mattias: Small T2 could migrate in a couple of weeks

Maria: We can discuss this, CMS proposal, maybe data to be used in the next month?

See backup slides for experiment input

Q: Has anyone looked into the past years to see how many downtimes have been declared more than a month? I get the impression this happens rarely.

A: A few long downtimes, at least one was not announced in advance. Operational problems for some of the experiments. Could we do better, is there a policy? None except for A/R targets, sort the issue once and for all, grew into a bigger thing. Indeed, nominally on the same page, sites try to do a honest job. Even that downtime was external reasons, Downtime had no choice. To make this into a concrete policy.

Ian Collier: Just to be clear, examples that prompted this were not scheduled downtimes, no policy will change that.

Q: Cases of data migrations? Did it happen before, what happens when the site comes back?

Maarten: Migration of SARA, announced half year in advance, pointed out that experiments should effect data should be on disk by time of downtime because it was on tape. Done very nicely. Not out of site, but experiments had to take nontrivial action. T1 can't be vacated just like that. Rare occurance. T2, migration statement we should say that this is primarily for T2s, they have a chance to be vacated if necessary.

Ian Collier: T1 moving into new machine rooms isn't so unusual but is usually planned and wouldn't involved DTs of more than a month.

Q: Suprised that 1 day warning of DTs is OK. Remember from KIT had 2 day downtime at the same time as another T1, major complaint.

A: Avoiding T1s going into Downtime at the same time. True that we can extent policy to conver this. LHCb insisted to please try to avoid clashes.

Q: Prompted be VO to notify in advance, looking to history big sites are announcing in advance.

Maarten: Complex matter, not an easy guilty party, far from it. Any site can say won't let other site dictate when I can do downtime. Site has the last word. LHCb always affected the most, most reliant on T1s, others can tolerate this more easily. Can't say 1st to book DT get it, external pressures. Try to avoid this when we can, then sites shouldn't be punished

Q: May be good to state in the policy - we try to coord with expts, have some flexibility. Did ask experiments, ask experiments in advance. In policy, put that it is welcome for sites coordinate with experiments?

Dave Kelsey: Along that line, doesn't say just data migrated but agreement with experiment is what to do, may be case by case basis. Interesting to EGI as well. Would be really nice to have WLCG/EGI policies agree. VO specific calculation is getting more complex.

Ian Collier: Spoken to GocDB devs, could in principle have more complex policy engine per VO with different constraints per VO.

Maria: Talk to EGI but look at WLCG first. Would be good to have feedback and talk more with CMS.

Ian Collier: Whether this liaising should be part of the policy, may be too difficult to capture in a dependable way. It's the practice for T1s that we try to coordinate. Capturing that formally, this might be the way to do this.

Dave Kelsey: Short term might be security, want that done quickly

Ian Collier: Not entirely scheduled, that's effectively incident response.

Maria: Monday meetings is the place to talk about this kind of thing.

Ian Collier: As mentioned, a bit of analysis in the last couple of years might be useful. Don't want to put too much work into policy that only deals with one edge case.

Maria: Experiments are asking for this. Agreement that generally sites are doing very well.

Ian Collier: Comments by 5th January - people should be thinking about this before they go away.

HEPiX Report (Helge Meinhard)

https://indico.cern.ch/event/394789/contributions/2392202/attachments/1388333/2113814/go
https://indico.cern.ch/event/394789/contributions/2392202/attachments/1388334/2113824/2016-12-14-GDB-HEPiXRe

No questions

Downtime policies: A proposal (Maria Alandes Pradillo)   3

Alice use of HPC Facilities (Pavlo Svirin)

https://indico.cern.ch/event/394789/contributions/2392205/attachments/1388279/2113711/GDB_meeting_14.12.2016.p

Q: What is the advantage that PanDA is bringing to you? Don't have to implement a lot of stuff/reinvent the wheel?

A: Yeah. Different approach, in next versions, yes.

A: Also consideration that ATLAS has strong presence working with Titan, want to leverage that, they have a team we have one part time person from CERN. Very good response from ATLAS team. Long list of things that need to be adopted. Again reinventing a lot of things needed to run on supercomputer, coming not only from few corners but people trying to find resources elsewhere. Everyone is doing this, probably not most economical way to have everyone do that. Especially with this level of application code, submission system, all HPC are different, unique requirements. Perhaps have to make cohesive effort for all experiments together, not have each experiment discover on their own. Tremendous amount of work. if we are going to start using these resources seriously, should have central effort.

AARC Report (Hannah Short)

http://indico.cern.ch/event/394789/contributions/2392215/attachments/1388481/2114469/20161214_AARC_Summary

AARC is a European Commission funded project that brings together 20 different partners from among National Research and Education Networks (NRENs) organisations, e-Infrastructures service providers and libraries. AARC aims to develop and pilot an integrated cross-discipline authentication and authorisation framework, built on existing AAIs and on production federated infrastructures.

For more details of the project, see: https://aarc-project.eu

Summary of the talk:

The 4th annual meeting was held at CERN at the end of Nov. 2016 (https://indico.cern.ch/event/569445/). A review of the progress/achievements is available at https://aarc-project.eu/achievements/. All of the deliverables are on track, and production ready pilots will soon be available. Strong progresses on the last year for the EGI and WLCG pilots (see slides for technical details). Sirtfi (Security Incident Response Trust Framework for Federated Identity) was widely accepted by identity federation community during 2016 (116 IdPs in EduGAIN already). Gèant (GN4) will fund/provide Operational Support, which includes Security Incident Report Support. An Incident Response Procedure has been proposed by AARC, a hierarchal approach to incident response with roles assigned for FedOPs & EduGain. In line with GN4 proposal. An agreement on the baseline requirements for levels of assurance has been set, targeted to concrete use cases from currently running infrastructures. A review on Data Protection is pending, and it takes approaches from both WLCG and EGI. Next Steps include moving FIM4R from version 1.0 to 2.0 (1st workshop planned for 20th Feb. 2016, cohosted with TIIME); developing scalable policy models in light of the Scalable Negotiator for Community Trust Framework in Federated Infrastructures (Snctfi) blueprint (ISGC2017 will be focusing in this); AARC2 phase, 2-years duration - starting 1st May 2017.

Questions:

(Q) I. Collier: avoid diff. projects doing the same type of things. How to merge and bring all other projects together?

(A) H. Short: INDIGO, EUDAT, are doing similar things. Communities should avoid doing similar things, and work in adapting. AARC is trying to do that, but it will take some additional efforts.

(A) D. Kelsey: this is not a single solution to impose. There are other products, using EduGAIN, but AARC brings everything together. The idea is not to impose, but let the people know on the product, easily to be adopted by any community. Indeed, AARC is very international; it shouldn't be seen as a EU-centric project.

SOC Working Group Report (David Crooks)

http://indico.cern.ch/event/394789/contributions/2392206/attachments/1388428/2114476/SOCGDBSummary-Dec2016

Summary of the talk:

Last meeting occurred in 8/12. Check https://indico.cern.ch/category/8128/☑ for all of the meetings made so far, since its creation in July 2016.

Security Operations Centers (SOCs) are complex, with many components in different areas, all integrated to deal with security log data and incidents. Several tools have been identified and are being exploited, for a minimum viable product: threat intelligence (MISP), IDS (Bro), reference framework (Metron). This falls under the working group's mandate to examine current and prospective SOC projects & tools.

There is strong use of Bro in the US. EU would benefit from more investigation in this area. Bro status: In the WG, some UK sites working with it in addition to CERN. One option being considered in new deployments is to monitor WNs through NAT as first step. MISP: threat intelligence sharing (see previous GDB talks). Status: RAL and Glasgow, including sync. of data. Also, tested sync. between WLCG(CERN) to Glasgow instance. MISP + SIRTFI: Glasgow added to the UK Access Federation SIRTFI pilot. Two MISP training events since last report (Brussels and Zurich). Training materials for MISP (including those used at the Brussels and Zurich events) can be found at the following URL: https://www.circl.lu/services/misp-training-materials/☑

CERN runs a MISP instance for WLCG/HEP. This enable sites to share or simply pull data for their own use, and/or enable direct sharing with other MISP instances. Alls IoCs from MISP being fed into Bro: 5-20 notifications/day total in general. Careful work to resolve concerning False Positives.

Next steps: deploy Bro at more sites; integrating MISP with Bro at non-CERN sites; testing/integrating more SOC components, like ElasticSearch, etc   The aim of the WG is to create a reference design for larger sites/sites with experience with a security appliance for smaller sites/those that wish it. Sites are encouraged to be involved in these studies/tests.

Questions:

(Remark) I. Collier: good to see that the project is providing concrete outputs.

(Q) P. Flix: How to engage sites into this? (A) Romain: any site can connect to the WLCG MISP instance, so they can gather information from it. Others can be more proactive putting IoCs into MISP.

Security Policy Update (Dave Kelsey)

http://indico.cern.ch/event/394789/contributions/2392216/attachments/1387942/2113202/Kelsey14dec16.pdf☑
http://indico.cern.ch/event/394789/contributions/2392216/attachments/1387942/2113056/SPG_Drafts_Security_Policy

Summary of the talk:

Policies adopted by EGI during 2016, but not yet adopted by WLCG: AUP v2 to include other services, such as HTC, Clouds,   acknowledge support in publications. VM endorsement and Operation, important for the EGI FedCloud services (VM operator and VM consumer). LTOS AUP & Security policy. Grid becomes

e-Infrastructure; Site becomes Resource Centre. Policies revised by SPG yet still to be adopted by EGI (feedback from WLCG welcome): Personal Data Protection Policy; Acceptable Authentication Assurance

Top-Level Security Policy doc has been revised (see slides for the [small] changes introduced). Some work to adapt the well matured July 2010 policies, to include new infrastructure concepts. The new version has been circulated for comments. Example: User Community management, shown.

Plans for the future includes work on: Top-level security policy; revising the policies for the VOs. New versions will be proposed before the end of the EGI-engage; Security for Collaborating Infrastructures (SCI). V2 will be finalized in 2017, will include GEANT and NRENs and other infrastructures.

Asking WLCG MB to adopt the new top-level security policy. Provide feedback soon to the circulated document. Approval to adopt all of the other policy documents from EGI, as well.

Questions:

(Q) I. Bird: what about data privacy policy, in particular in the accounting data?

(A) D. Kelsey: this is much broader now, rather than accounting. No user information should go in public records. The frameworks should be compliant to this. The idea is to provide templates.

(Q) I. Bird: connection of this work and other initiatives/sciences, like NSF funded projects?

(A) D. Kelsey: not really too close today, in the past this was lead by individuals. OSG inputs to these modifications would be very welcome.

Community White Paper (Peter Elmer)

http://indico.cern.ch/event/394789/contributions/2392204/attachments/1388678/2114494/20161214-gdb-dec2016.pdf

Summary of the talk:

A CWP has been proposed to define a longer-term strategy for HL-LHC. 3 goals: improvements, scalability and performance, and to make use of the advances in CPU, storage and network; new approaches; long term sustainability of the SW during the lifetime of the HL-LHC.

All of the HEP community needs to be engaged. List of questions and topics to be discussed per area are presented. There is broad support for the idea, which has been presented to the LHCC. Some set of workshops will occur, a kick-off WS will happen in San Diego in January (http://hepsoftwarefoundation.org/events/2017/01/23/Workshop.html), and there will be a final workshop in summer 2017 (in Europe, probably CERN or near CERN). San Diego WS: many people from LHC and beyond already registered!

As of today, there is a request for contributed white papers on Computing Models, Facilities ,and Distributed Computing. Pre-existing docs can be linked and inserted later on.

Questions:

(Q) P. Flix: How this is organized? Are there editors chiefs for the different groups?

(A) P. Elmer: Groups are self-organized. People join and participate.

(Q) I. Collier: which are the groups which are active or not?

(A) P. Elmer: check the webpage, then you can see which ones are active, and which ones need more people.

(Remark) P. Elmer: during the San Diego WS, they will try to habilitate Vidyo in the rooms to allow remote people to join and participate, at least in the plenary sessions. Then, each WG can self-organize with their laptops, so this can be considered as well.

Wrap up (Ian Collier)

This topic: LCG > GDBMeetingNotes20161214
Topic revision: r3 - 2017-01-02 - JosepFlix