

Table of Contents

Summary of May GDB, May 8th, 2019 (@ EGI Conference, Amsterdam)	1
Agenda	2
Introduction - Ian Collier	3
Benchmarking Update - Domenico Giordano (DG)	4
Middleware Evolution - Erik Mattias Wadenstein, Martin Litmaath	6
Report on Cream CE Migration Workshop - Jose Flix Molina	7
Bringing services in to EOSC - Matthew Viljoen	8
Privacy update - David Kelsey	9
HEPiX San Diego report - Helge Meinhard	10
Security Challenge debrief - Sven Gabriel	11

Summary of May GDB, May 8th, 2019 (@ EGI Conference, Amsterdam)

Agenda

Agenda [↗](#)

Introduction - Ian Collier

slides [↗](#)

When the room was asked about FNAL Dune hosted GDB in September there was general nodding and no objections. Dave Kelsey highlighted that there may also be a FIM4R (Federated Identity Management for Research) meeting in Chicago being planned around the same time.

Dave Kelsey (DK) - TNC in Tallinn should be added to list

Catalin C - CVMFS in June at Cern can be added to list

IC requested that the details of any missing meetings are emailed to him

Benchmarking Update - Domenico Giordano (DG)

Presented remotely

slides [↗](#)

Volker Guelzow (VG) do you expect companies to run the benchmark, who will run it? If companies, what will you allow in terms of optimised libraries and code modification

DG - this goes beyond preparation of the tool and needs wider discussion in WLCG. We are trying to make the barrier to running tool as low as possible. To this end also making clear to experiments that code must be publicly licenced and sample data fine to be publicly accessible. Another aspect to the answer to this question is: if they are able to identify areas of optimisation then why not get this feedback? It will be useful for the general optimisation of the code for the experiments.

Brian Bockelman (BB) - HEPSPC06 is useful as it has been the same for 10 years - talking about continuous integration and a benchmark seems a bit strange. Do you see this benchmark as something for HEP, WLCG, or CERN.

DG - We can guarantee the same stability - you can choose to use the same container version for the next 10 years. However, this is not an advantage as the experiment code evolves. There is an interest in including the latest applications in the latest version of the benchmark.

BB- differences in versions have an impact on pledges, there is a need for a fixed point of reference.

Maarten Litmaath (ML) - HEPSPC is diverging more and more from our code base and is less representative of this. There will need to be a process for moving between versions of the benchmark.

BB - Need to use benchmark for at least 5 years for use cases like pledges.

DG - Slide 3. Correlation for many years with HEPSPC but we are now diverging and so there is no good justification for remaining on HEPSPC

Helge M - there are good arguments for having one and only one benchmark and keeping it stable for as long as possible. Last time we coupled ourselves to the idea of SPEC CPU, and we have been stable on that for over a decade. Whilst there are alarm bells being rung, we are only seeing a 20% divergence, at worst, from this for our workloads. This is a success story, but it is unclear if we can repeat this success. We should be clear that changing the benchmark every year would be a mistake, but maybe 10 years is not possible. We must seek to have one benchmark and not end up with multiple benchmarks to reflect different workloads, which ultimately vary by only a few percent.

BB - we have created a tool that creates benchmarks - perhaps we have a standard reference for WLCG, but we can provide ability to generate new benchmarks for others to use.

HM - need to be careful as not only used for purposes such as pledges, but also used for procurement. There is a risk of causing confusion with companies that we work with if there are multiple benchmarks.

VG - Agreed that community specific benchmarks may be ok, but benchmarking costs money, so we need to make sure we don't ask companies to run multiple benchmarks.

DG - we are working on ensuring report from benchmark provides detail on e.g. code version. The discussion here is beyond the work that is being done.

IC - on the point of who runs benchmark, surely both sides run it - we procure to meet a set benchmark and provide benchmark. We then run the benchmark ourselves to validate the procurement.

Brief discussions in the room suggested this wasn't the case for all sites.

Mattias Wadenstein - we need to think about what the trigger point for using a new benchmark is. E.g. major changes in hardware

IC - This new framework allows us to change the benchmark when we feel the need, unlike the current situation where we are reliant on outside developers. It might also be the framework can be transferred to other communities

ML - We are at EGI. EGI adopted HEPSPEC06. If we do this right we can provide this benchmark to EGI and they also benefit.

BB - raised a concern that if the benchmark is hosted at cern.ch other communities may ignore it.

IC - may make sense in the future to change that, it is in our control. Approach should be theoretically extensible.

IC - should be noted that there is an action from the last management board for the benchmarking and cost modelling groups to work together

Middleware Evolution - Erik Mattias Wadenstein, Martin Litmaath

slides [↗](#)

IC - It is often easy to forget that this kind of migration is a normal kind of thing for us. It is not that it is no work and does not require attention and people to learn things; but we do know how to do it.

Report on Cream CE Migration Workshop - Jose Flix Molina

slides [↗](#)

Slide 5 - DK - Security evaluation is complete (in response to "being worked out")

IC - it may be interesting to not look just at the number of deployments, but at the amount of resources behind them.

Bringing services in to EOSC - Matthew Viljoen

slides [↗](#)

Virtual access: https://wiki.egi.eu/wiki/EOSC-hub:VA-TNA_FAQ [↗](#)

Q&A: Communications: already working quite well re operations of WLCG/EGI. Always areas for improvement, CREAM CE [session] fruits of this labour, shared issues. Presence at shared WGs can also help.

best way of gathering requirements for EGI services at OMB.

Ian: Is OMB open? Matt: For NOC managers. Ian: Is it clear that the relevant WLCG people are auto involved, or is there extra work needed? Maarten: Not been involved for many years Matt: Something to consider. Maarten: Already discussed over coffee, should have small committee? Group of people that see indeed where we could have better integration where it makes sense. Catalogue, don't use what we don't need, risk of discovering islands within this.

EOSC is still being made, can still influence [requirements]?

Matt: Yes - WLCG requirements feed into EGI, part of EOSC, need to make sure that reqs are being met. OMB this year understanding each tool in turn, getting feedback, good means to get requirements. Already APEL, GOCDB, will look at GGUS, Monitoring, Dashboards etc.

Ian. Double check, APEL, people from WLCG acc TF involved? That would be the obvious thing. Looking from view of APEL devs reporting to me, note times where the project reqs are so pressing that WLCG requirements are not so much of a priority? Hunch is historically there were individuals acting a bridges, but the thing is to know what the process is that's needed. Not fair to expect that devs to act in that role

DK: Replace John Gordon? Ian: The functionality Maarten: Requirements for accounting, benchmarks, etc. APEL is going to be funded

Brian: Comment about HTCondor CE - it was an example where coordination and planning would have been helpful to have accounting in place when needed.

Privacy update - David Kelsey

slides [↗](#)

DK: Historically have worked together on Security Policy, particularly in the WISE world. To extent privacy policies, hope useful across WLCG/EGI/EOSC...

Hoping that GEANT CoC v2 GDPR compliant would come to be. We could then follow that... hasn't happened. Need to do something in the interim.

CERN: All services need to go through "record of processing operations". However, WLCG decided to have single document which would be MBs view of provacy. Clear entry points etc. - need to address some draft comments, need to add those. Need to show to CERN.

Joint meeting in Utrecht week after next

QA: Going in the right direction. One of the key things is that you know who is accountable for the processing. Not sufficient to have privay policy that nobody owns...

DK: The privacy policy will have contact points per services. Multiple controllers.

And each contact must know they are responsible.

Ian: May not be enough to rely on individual institutions/countries, need to have work At WLCG level.

DK: Been round user consent, can't be freely given for professional work. Missing monitoring and legal elements. There we rely on the fact that it's only low risk stuff, things that would be professionally anyway. Fully legallly, bilateral agreements, 10s of thousands.

Maarten: Need to protect ourselves. "Prove to me that you're treating my data correctly". Price of this is that you will be temporarily suspended. Need to have something like that in my view.

DK: Would really like Refeds, acc community to have CoC...

Maaretn: Continue on this path. Be more careful with monitoring dasdhboards.

DK: If the Expt **needs** to expose identities in the dashboards, should better have a note of why in the privacy notice.

HEPiX San Diego report - Helge Meinhard

slides [↗](#)

Some people might say No proceedings is a bug, but...

... no proceedings, trust between people through discussions. People tell you what they tried that didn't work. Tells you at least as much as what worked.

- recordings not quite available yet but soon

- Also SOC Workshop intention for week following HEPiX

QA:

NTR

Security Challenge debrief - Sven Gabriel

slides [↗](#)

NOTES TO BE CLARIFIED

- Busy with other things and large amount of data to evaluate from different groups; this is a preliminary report.
- Modern landscape where traceability information is derived from a number of sources.
- Important to note that no single team (egi or VO csirt) can deal with such an incident, since no team has the full picture
- Planned in close collaboration with VO from the outset in 2018
- Better to find issues during exercise than real event.
- Final report, together with VO, in July
- Tracing jobs back through pilot framework?

QA:

Evgeny: No. of bots? Sven: 62 sites, around 250 bots? E: Spread between sites, how decided?

S: Just launched at all. Not deliberately fired at particular site. Vincent: Sites should receive same number of bots. [similar per site] Sven: Some commented "I can't see anything" - ask for more 😊 Sven: Not just one site where bots lived a long time

Sven: Not all sites react as they did before

Matthias: Examples of actions you would expect?

Sven: See activity, immediately stop it.

Ian: Do you think that the differences are due to the nature of the challenge, or something else?

Sven: Thinking about that this week. Turnover of admins, at least at small sites. Properly done handover. Small sites, grid admin, maybe just side project, more that ten years old, admins may leave... Maybe got a sense that not everyone understood what we were talking about? May have to do this more often.

-- IanCollier - 2019-05-08

This topic: LCG > GDBMeetingNotes20190508

Topic revision: r6 - 2019-05-20 - DavidCrooks



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