Table of Contents

Summary of January 2020 GDB, January 15th, 2020

DRAFT

Agenda

For all slides etc. see:

Agenda↗

Introduction

Speaker: Josep Flix (CIEMAT (ES))

slides⤢

- March GDB at Taipei
- Mattias and Pepe as new chairs - Thanks Ian!
- Storage QoS white paper v1.0 released for feedback
- Volunteers needed for SVG Deployment Expert Group
- New K8s e-group created
- Let the new chairs know about topics and/or organizational aspects for GDBs and pre-GDBs

Operational intelligence

Speaker: Alessandro Di Girolamo (CERN)

slides

- Automation of the repetitive tasks for monitoring and reacting to errors.
- Machine learning on the monitoring and alerting flow might help to get a better flow of useful alerts.
- A first approach is being done with Rucio operations for ATLAS.
- ATLAS and CMS are also working on classifiers for job errors, to identify the most important underlying causes.

Discussion:

- Valassi: Could we just start with very concrete list of problems to get fixed, instead of being so big, maybe vague or theoretical? For ML techniques one needs to be very concrete. Alessandro: The idea is to start simple, and grow, but being ambitious. Valassi: one should be more concrete, there are lot of things happening everywhere, which share things or not. Demonstrating with specific things would help. Alessandro: there are concrete activities: Rucio, WF management, etc the forum allows for people to speak together.
- Graeme: Is there an Open source set, or open data, that people can use to get traine, put effort and cooperate with? Alessandro: It was asked to computing coordinators, no answer yet. That s the idea
- Duellman: try to have data in some place in order to classify, but we need to be aware on how to filter the data. The framework that collects the data is also very important. It could be good to have target metrics. Alessandro: Agreed.
- Valassi: We have more data that indeed we look at (!). Alessandro: the aim is to have all of the monitoring data in a single place.
- Pere: you need to have actuators, single recipes, and react in such a way an operator should react. Alessandro: this is the intention.

Geant4 Strategy for HL-LHC

Speaker: Pere Mató (CERN)

slides⧉

- Implementing lessons from GenatV, while GeantV's basic vectorization approach did not show to give the expected speedup.
- A rewrite of Geant4 with modern software engineering could give a 50-100% speedup.
- Fast simulation and accelerator support is also upcoming, which can give lots of speedup where applicable.

Discussion

- Valassi: FastSim is experiment dependent, is there a unified validation framework? How to validate this? Pere: fastSim shouldn t be much experiment dependent, fastsim should go on certain techniques, no on the experiment setup and config, the validation will be done with full sim, which is validated with testbeams, etc
- Mattias: G4 sensitive to CPU architectures? How can we get benefited? Pere: Vector instructions, quantitative evx2 or not, we need to go in this direction
- Pepe: Static/Dynamic compiling? Pere: better static. Compiler version is as well important.

Compute provisioning ideas

Speaker: Andrew McNab (University of Manchester (GB))

slides⧉

- Looking at medium or short term compute provisioning, reservation-focused.
- Medium term today is slides in meetings, maybe a json file could be better?
- Short term there is so many systems, maybe some more commonality could be better?

Discussion

- Maarten: is there a problem in mind you want to solve? HPCs work very different, and their system could be handy. Andrew: LSST ask in UK for resources in short campaigns, chunk of jobs to be executed. Manually accommodated. This is not going to scale. Maarten: how to structure this? CRIC/Rebus, 1 or 2 experiments fell paint, not much effort is going to be available in here. Andrew: do we need this in the future?

Maarten: More use cases, and experiments should be confronted. All this work needs to compete with other things. We need strong case.

- Pepe: pre-GDB would be fine, to know about experiments interests
- M: There is things going on, with priorities and resource provisioning, to integrate all of this might be a big thing
- A. Forti: LHC experiments define their own priorities, and the sites decide how to handle the priorities in the site
- A: Based on what?
- Forti: agreement for some specific workflows
- M: Why experiments should modify their priority if other experiments have priority jobs? How they can accept to be affected?
- A: The priorities can be adjusted, adhoc, manually.
- M: How to encourage the experiments not to run in priority mode and they accommodate in flat submissions
- Simone: we should avoid giving the idea you can submit in peaks, which opens the door on who has priorities. These communities need to be educated to use computing, in a distributed system, in a fair mode. What has priority for me, even within the experiments
- ???: In reality you are always late. You want everything in high priority. This is unsustainable.
- Simone: Everything is processed late, according to the planning.
- Julia: Requests can define them quarterly, but no one is using it. It is much complex.
- Mattias: Which consumer is, common APIs could benefit the stakeholders. It could be benefits in the commonalities.
- Discussion in the room said that the medium term planning was probably not very interesting for the LHC experiments.

Updates on OSG/WLCG perfSONAR Network Monitoring and Analytics

Speaker: Shawn Mc Kee (University of Michigan (US))

slides⧉

- Report on perfsonar.
- SAND project to look at more data and do richer extraction.
- Building network topology from traceroutes is challenging, but a very rich data set.
- Several student projects based on the data ongoing.

Discussion

- Dave K: Ipv6 slide 16. How to select one or the other. S: Effectively they are different measurements. Dual-stack, you get two measurements if you don t specify ipv4 or ipv6
- ???: Slide 14. LHCONE? S: All of the mesh
- Pepe to Shawn? MTU? S: No MTU is available for the moment - this come from trace path
- Mattias? Different routes for no reason? We need to understand why there are changes in the routes and paths. Maybe some paths are load balanced, we can check the performance.
- Pepe: We need to check with FTS, etc to see the effevts s: agreed

What CPUs to support, at what cost

Speaker: Helge Meinhard (CERN)

slides⬀

- Raising the question of regular updates on both baseline complier versions as well as CPU flags to support, or at least do the homework to show that the current situation is as good as it gets.

Discussion

- Markus: how to support multiple versions, as general. Static builds of g4 with ATLAS geometry is not possible. She is using the CMS geometry. There is lot of effort needed. The orchestration is done with python modules, the activity is fully supported but there is long distance to travel
- Maarten: But, we do need HEPOSlibs - it is linked in in the application
- Johannes: Nice initiative, ATLAS would like to support it. Speeds ups: simulation and rec. with avx2, and so on, there are studies being done. We don t see much gain in the full sim or full reconstruction. Word of warning.
- Shawn: HS06 in the sites
- Mattias: requirements would go down
- Rob, ATLAS: static build, efforts spent, not managed, complex
- Graeme: we need to look into this
- Andrew: which benchmark we need to use. LHCb DIRAC is compiled matching CPU architectures. They use AVX2, etc   they run mostly simulation
- Concezzio: This is very experiment specific, you enable the flags after optimizing the code. There is lot of homework done before enabling vectorization. And this takes years.
- Markus: physics validation is still very complex. It is very complex.
- Maarten: it makes sense to study this, and see the benefits of this. We need a report on how we asses the things in the way they are. We need to do this, definitely.
- Markus: 100 MCHF in WLCG - 10% improvement means 10 MCHF, which is some money.
- The conclusion was that this should proceed, and that we need to do this work as a community.

LHCOPN/LHCONE Workshop summary (pre-GDB)

Speaker: Shawn Mc Kee (University of Michigan (US))

slides ☑

- Meeting report.

Discussion

- No issues raised.

-- JosepFlix - 2020-02-10

This topic: LCG > GDBMeetingNotes20200115
Topic revision: r1 - 2020-02-10 - JosepFlix