Table of Contents

Summary of June GDB, June 12th, 2019

Agenda

For all slides etc. see:

Agenda

Introduction - Ian Collier

Speaker: Ian Collier (Science and Technology Facilities Council)

- July GDB has Dynafed and DOMA Access preGDBs
- No August GDB
- September GDB planned to be in FNAL, with Authz preGDB also at FNAL. Open

meeting on the Thursday on authz too. Final call to be put out this week. (This is now confirmed - Ian 2019-06-20)

- October preGDB on Benchmarking
- January 2020 (1 week later than normal) 2-day preGDB on LHCOPN-LHCONE

Policy update

Speaker: David Kelsey (Science and Technology Facilities Council)

Aim for overarching policies for EOSC. Use WISE templates. WISE since 2015 as joint policy group.

New/revised documents:

Draft privacy notice: general document.

Draft template privacy notice, for use by services whose processing of personal data is not covered by general one.

Overarching policy framework, based on WLCG requirement to abide by (security) policies, with GDPR changes.

WLCG DP Policy Framework, updated from 2017

Acceptable Use Policy, uisng WISE Baseline AUP v1. Accepted by all infrastructures (or most!)

Discussion about some details of the Privacy Notice, including right to be forgotten and transfer of details to others. To raised on GDB mailing list for further comments. There will be more time for discussion at a future GDB.

AUP template attached to agenda. Doesn't cover SLA-like aspects. Discussion about AUP, and what to include (eg giving acknowledgements.)

Please comment on Google doc.

LHCb Computing outlook

Speaker: Concezio Bozzi (CERN and INFN Ferrara)

30x increase in throughput from upgraded detector and fully software trigger (event rate and pile up)

Full reconstruction with final alignment within online farm.

Run3 computing model. Split HLT (RT alignment and callibration), TURBO stream; offline computing dominated by simulation; offline storage driven by trigger output bandwidth (MC saved in microDST so less impact on storage)

Run2->Run3 FULL/Turbo goes from 65%/26% to 29%/68%.

Data will be recalled from tape during year end technical stops. Not reco, just filtering and slimming pass. Expect ~4x increase of throughput required for staging wrt Run2.

CPU dominated by MC (~90% CPU). Expect the same after Upgrade.

First Run3 year assumed to be a commissioning year with half luminosity delivered.

Slide 17 has storage requirements projections as graph and table. Pledge evolution assumes constant budget model with +20% more every year. Some deviations during LS3 but in line with model by end of LS3. Similar for tape.

CPU also based on the same pledge evolution, but with the addition of opportunistic resources and the online farm.

Summary: major re-engineering of software framework; order of magnitude more physics events from the detector; offline CPU dominated by simulation; DIRAC will remain LHCb workhorse for CPU and data management; Run3 requires modest increases in resources wrt Run2, not far from constant budget scenario; mitigation strategies include fast simulation and WLCG/DOMA iniative to reduce storage costs; making efficient use of HPC nad co-processor resources is challenging.

Questions/answers/comments:

Q. Could you run all simulation on one big HPC? A. Yes, if such a machine was willing to provide it all! If it would consist of x86 processors only, and it would comply to our requirements such as network connectivity for the worker nodes and software installlation via cvmfs.

C. This kind of suggestion is very unhelpful as such resources aren't available to us in sufficient quantity, but it comes up in funding contexts ("Why don't you just run on HPCs?")

C. By Run4 experiments will need to be able to run on GPUs.

C. On Run3, 20% growth is seen as not feasible now. Also big steps up and down in budget difficult for some countries.

CMS Computing Outlook

Speaker: Stephan Lammel (Fermi National Accelerator Lab)

Data processing plans for LS2: re-reco of Run2 data; MC matching those calibration etc improvements; reprocess heavy-ion data; reco parked LS2 B-physics; MC for Run3 detector; MC for phase-2 upgrades; continuing user analysis and MC

PhEDEx to RUCIO during LS2. CMS grid info from SiteDB to CRIC. webDAV storage access "of interest". Small computing resource increase at sites: network becoming more important with cross-site data access (not pledged though.)

Dashboards to MonIT at CERN; end of CREAM-CE support; AFS to EOS at CERN; SL6 to CentOS 7: asked CMS sites to deploy Singularity last spring - maybe CentOS 8 + Python 3 for some CMS services; expect golden releases of site storage systems for Run3 and sites to upgrade.

Questions/Answers/Comments:

C: AFS community came alive again, so continue to end of Run3. No longer a rush to migrate away for AFS as a high performance shared filesystem. C: May be able to use CentOS 8 more generally for some services, eg WWW C: dCache golden release should be ok

DOMA TPC report

Speaker: Alessandra Forti (University of Manchester)

Third party copy phase 2: all sites providing more than 3PB of storage have non-GridFTP endpoint in production (not necessarily for all experiments) as required. Not used for TPC yet.

Explicitly stated storage software baselines, needed for TPC to work

For authorization, need delegation if xrootd or tokens for http; and checksums.

Run http smoke tests: once a day very detailed tests; can be run manually by site admins to check changes; xrootd version being worked on. Rucio functional tests on any site that asks to participate via http/xrootd.

DOMA stress tests, for produciton sites with baseline, for http since January, and adapting for xrootd too now.

Extending to experiments: CMS can already add non-gridftp protocols for TPC to phedex; ATLAS needs development in Rucio - nneds to handle sites with different protocol preferences in AGIS; setting up stress tests in ATLAS

To go into production also need FTS. Specific verson of FTS/gfal2/xrootd and davix have been installed for this to work. Need to switch off httpd streaming (to avoid ambiguity when monitoring TCP.)

More sites: need more production sites, particularly DPM (TF you can refer to) but also dCache, xrootd, EOS and SRM.

Conclusions: getting there; sites upgrading to baseline will help

WLCG CRIC update and demo

Speakers: Julia Andreeva (CERN), Mr Panos Paparrigopoulos (CERN)

CRIC combines information from multiple sources; allows cross checking; allows authenticated data modification and retrieval; logging of modifications; a central entry point for WLCG generic topology and configuration information, required for central WLCG operations, testing, accounting and monitoring.

CORE CRIC, then plugins: ATLAS, CMS, and WLCG (for ALICE/LHCb use and central ops tasks), and DOMA CRIC for TPC tests.

WLCG CRIC clients: SAM (ETF), WLCG central monitoring; WLCG accounting; all clients requiring pledge info; WLCG dissemination (Google Map etc); interactive clients for pledge injection or navigation.

Working prototype of WLCG CRIC ready for testing, validation and feedback (wlcg-cric-dev-1.cern.ch). Stable version already used for WLCG Storage Space Accounting (wlcg-cric.cern.ch). Need feedback before going to production.

Demo of WLCG CRIC web dashboard. See screenshots on the Indico agenda.

Questions/Answers/Comments

Q. What about pushing fixes upstream to GOCDB for instance? A. Need to be able to do something immediately while wait for upstream fixes

Q. Is there a way to view data by country? A. No, only by federation at

Q. Could we require a ticket to be associated with discrepancies between say GOCDB info and what is stored in WLCG CRIC A. Yes C. Who should do all this? C. Worried about long lived tickets because site and experiment don't agree or configuration can't be fixed.

Q. When can REBUS be closed? A. An MB decision. Two to three step process. Another round here to check everyone is ok with it. Not before summer break. Soon after, able to conclude that CRIC is now better, and then a smooth transition? Ask MB next week to consider it.

LUNCH

################### The EISCAD_3D Data Solutions project (John White)

see slides

NeIC : Nordic e-Infrastructure Collaboration EISCAT: study of the auroral ionosphere EISCAT radars: total dataset since 1981 is less tan 100 TB EISCAT_3D project: nex type of radars, several PB of data data management operations similar to the LHC experiments embargo for n years of the data (make AAI difficult), FAIR principles EGI Check-in implemented Data management services: FTS, rucio User management through VOMS, X509 is too complicate for users -> looking at EGI Check-in

discussion

Q: what about the network to the datacenters ? John White: Nordunet will bring a link to them. Nordunet will make sure they have dark fibers. Some of them are shared by the universities, a part is reserved, like 1-3 Gbps. Currently four datacenters are candidates. 1st data: fall 2021

Maarten Litmaath: with respect to AAI, in WLCG we figured there is two types of users. We want to ease life of 99% of them. There might be X509 around for the power users (like production jobs). You might consider that instead of going 100% to EGI check-in. John: in EISCAT_3D we have only one or two power users, users are basically using mathematica Maarten: it is the same for SKA, it does not fit on your laptop anymore. We try to simplify and you may pick from our catalog as much as you can to make your life easier

CMS Computing Outlook   7

#################### LHCONE/LHCOPN Workshop report - Umsea (SE) (presenter: Erik Mattias Wadenstein)

see slides

---

discussion Stephane Jezequel: CMS is claiming the FTS transfers will increase and ATLAS plan to inscrease xrootd transfers. How do we manage both ? Is there a monitoring ? Mattias: it was discussed in the workshop. The quick answer is that the FTS transfers is know about, the rest is not. Maybe you can get it from the job monitoring. Maarten: I was about to say something similar, you can do better than 20 years ago. Data locality is still very important. You have better to make sure that your job is close to the data. Most sites are not ready to make their network increase by 25% let's say. This morning was said: there is no network in the pledges. For the many years to come we have to be cautious and see where they break. I would not recommend anybody to change how the jobs go to the data. Mattias: they is also a large variety in the way sites have network. In some part of the world it is also more expensive ect. Maarten: we may go to the lakes. We have to make sure it benefits and we have to see over time if the sites go to the data lakes model or if it does not happen. We should not run away with an idea that will not happened. Mattias: some sites could also profit from HPC centres and do not do analyses anymore, just simulation. Alessandro Di Girolamo: My points: - your slide 26: big topic, it is a pity you say we can't do anything about it TPC is moving 2 PB/day for ATLAS if you move 50% of the jobs reading remote, you put the load on the WAN my main concern is that in this we have no handle and no monitoring an example in France recently, we did not know what to debug we would like to know who did what - 2nd point: about the cost of WAN, experiment should know that, if not we move things around. We have to know.

Mattias: and the cost of network should be on the same form of the "20%"

Alessandro: 1st we have to interface the experiments with the network

Mattias: discussed by Shawn, there is ESNET and GEANT involved and it is not soon to happened

Alessandro: if it is not possible to have real time info, like on the last hours (to adapt our workflows), it is good to start with something like one week

Mattias: another part was to create a map of the traffic in a country and make a nice topolgy map, like to have a collection of these weathermaps (slide 13) Alessandro: it might be difficult for an experiment to use this we don't know the paths Mattias: it is part of the discussion: what is it the information you need ? there is a learning curve Maarten: we will be slowly moving forward in the coming years to come the current projects, even small scale, will spark further ideas in the last years, there was a push towards SDN... the other extreme and as Alessandro pointed out, he does not know the route between 2 sites ideally, we would avoid to be network experts to optimise transfers between A and B Mattias: you should have some knowledge on what kind of demand is feasible or not

Alessandro: we might select 25 sites and would be alerted if there are a large utilisation for more than 12 hours for example Mattias: a map with this amount of links is readable Maarten: we should see Eduardo and see if it is feasible Alessandro: I discussed with Eduardo and the info is available for cern links Mattias: a good overview should be doable if the providers share their info Alessandro: I am cautious because some unmonitored transfers might distrurb the monitored transfers Xavi Espinal: this kind of monitoring could help also in understanding the impact of processing data. Before doing some statements we need to measure the real impact

//end-15h40//

-- IanCollier - 2019-06-19

This topic: LCG > GDBMeetingNotes20190612
Topic revision: r3 - 2019-08-01 - unknown