Architects Forum - Minutes

Date / Time: 15 November 2018 - 14h
Agenda: https://indico.cern.ch/event/766401/
Present: P Mato, G Ganis (chair, notes), D Piparo (for A Naumann), G Cosmo,
J Blomer, C Delort, M Schulz, W Lampl, A Sailer
Next meeting: 17 January 2019 - 14h

Introduction (G Ganis)

G Ganis reminded that the AF is the main point of contact between SFT and its stakeholders and that it is open to suggestions for topics of discussion and to new participants whose activities are connected to SFT.

In this respect, he welcomed M Schulz, CERN IT-DI-LCG, leader of the project 'Understanding Performance', which joined the forum as of this meeting. Understanding Performance is a long-term activity started in 2016 with the goal of analysing the software and workflows in HEP, providing tools to understand, measure and improve performance, in close collaboration with the community (HSF, relevant fora).

Several excused absences influenced the attendance of today's meeting. He reminded that the tentative dates for the next meetings are available on the Indico calendar; people are invited to inform SFT of conflicts so that a solution can be envisaged in case of need.

SFT line management (P Mato)

Personnel movement:

- Staff
    - P Mendez Lorenzo is moving to EP-DT as of the coming December
- Fellows
    - G Amadio was accorded an extension of 6 months
    - A junior fellow co-financed EP and SFT was selected to work on the Drupal 8 transition

A discussion is going on between SFT and IT concerning the support for the SWAN service. This is likely to require at some point input from the experiments; it will be further discussed in this meeting.

Report from the SFT Projects

ROOT

Busy period for the team, with the preparation of the end of year release 6.16/00 and of patch releases 6.14/06, tagged on Nov 7th, and 6.14/08 foreseen by end of November. These patch releases include important fixes for bugs affecting mostly ATLAS; they also include important performance improvements in multi-threaded mode; experiments are advised to test these versions and evaluate the performance improvement.

Simulation

Geant4

Team busy with the preparation of the end of year release v10.5, expected for early December. A preview is available for test on CernVM-FS. Given the time required to validate and adopt a release, this version is a good candidate of the version which will be used during Run 3.

CernVM

CernVM-FS

The preparation of the patch release v2.5.2 is almost completed and should be out before the Christmas break; as a reminder, this is a bug fix release fixing some issues with the S3 backend. Working with IT on the optimization of the `unpacked.cern.ch`, a repository dedicated to (unpacked) container images.
The preparation of version 2.6, expected for Q1/2018, proceeds well. The major new features expected include: the `shrinkwrap` utility a new tool to extract and pack information from cvmfs intended to replace the unofficial `uncvmfs` tool; a notification systems to broadcast repository changes and reduce propagation delays; a component to collect and monitor publication and garbage collection statistics; a better integration for containers distribution.

As a reminders, experiments/groups suffering from performance issues during publication, are encouraged to negotiate with IT the move of the backend to S3; experience shows that one can expect a factor 5 speed increase for the publication phase.

CernVM

The CernVM 3 and CernVM 4 appliances got regular security hotfixes, following the Scientific Linux and EPEL security updates.

SPI

The release LCG_94 - available since mid September - has been completed with additional platforms combinations, including python 3 and Ubuntu 16.04 and 18.04.
The team is preparing LCG_94a, which will include (probably) ROOT v6.14/08 and a few fixes in some packages.
The team is continuing to work at the consolidation of the build infrastructure and to extend the integration test coverage.
A quick discussion took place about when the experiments will move to Python 3. No answer was give, Run 3 remains a possible target. It was mentioned that the BE users of SWAN are using Python 3 LCG releases.

Report from related Projects

COOL/CORAL (C. Delort)

As part of the maintenance activities, CORAL has been ported to Python 3, fully tested in the LCG builds. Support for Python3 in COOL is the next step.

Understanding Performance (M Schulz)

Main activities of the team aim to understand the performance cost model, in particular with respect to network, IO, CPU utilization/exploitation, and to the provision of tools to improve performance profiling.
Currently the team is investigating the impact of storage redundancy (RAID), latency and bandwidth, including also the impact on performance of data the access method and of the lifecycle of data (how much and how long they are used).
Tools like Trident allow to deepen the understanding of the factors limiting performance. Investigation with LHCb and ALICE have shown that experiments codes are not very optimized, in particular wrt memory management, with estimated potential improvements of 10-20%. Also, code optimized from the start - e.g. ALICE HLT on GPUs - can be much more efficient.

Stakeholders feedback

ATLAS

Decided not to reprocess 2018 data. Open sourcing ATHENA getting serious; the problem is that 25% of the code in under GPL (the rest mostly Apache). Removing violations of license

agreement, using as much as possible the BFG tool. All this means rewriting (git) history, whihc is the price to pay.

Linear Collider

Team is busy with contribution to European Strategy contribution and related performance studies.

G Ganis, 11 December 2018