Cooperation in Software

Axel Naumann, 2020-05-06

Computingstrategie-Workshop

This is not just about ROOT

- ROOT as an example and personal motivation
- (Lack of) effective cooperation seems a general issue in the community
  - Few major software teams suffocate other smaller ones: not enough "meat" left to get funding. Lack of distribution of tasks.
  - Causes duplication of efforts, segregation, non-technical opposition.
- In an environment of scarce resources, difficult sustainability, competition for global relevance, and major technology challenges!

What brings me here?

- Community has problems to solve
- Would like to help increase relevance of German software contributions
  - Embedded in international projects, beyond experiments
- Germany has significant impact, expertise, resources for software
  - Is this underestimated? "Who would care?"
  - Or overestimated and thus "no need to look left or right"?

Collab + Coop Models

- Formal: Geant
- Cooperative: ROOT
- Others (US IRIS) mix this: formal for money, cooperative in coordination, seemingly ad-hoc in topics (depending on available contributors?)

- Not cast in stone: needs drive models, not the other way around

Module	Code Owner
I/O	Philippe Canal (**Fermilab**)
• Compression	Oksana Shadura (**Nebraska-Lincoln**)
Math, Statistics, Machine Learning	Lorenzo Moneta (CERN)
• RooFit	Wouter Verkerke (**NIKHEF**), Stephan Hageboeck (CERN)
• New histograms (RHist)	Hadrien Grasland (**LAL**), Axel Naumann (CERN)
Graphics	Olivier Couet (CERN), Sergey Linev (**GSI**)
RDataFrame	Enrico Guiraud (CERN)
Web-GUI, http Server	Sergey Linev (**GSI**), Bertrand Bellenot (CERN)
Event Display	Matevz Tadel (**UCSD**), Alja Mrak-Tadel (**UCSD**)
PyROOT	Enric Tejedor (CERN)
cling	Axel Naumann (CERN)
Build System	Oksana Shadura (**Nebraska-Lincoln**)
C++ Modules	Vassil Vassilev (**Princeton**), Oksana Shadura (**Nebraska-Lincoln**)
Documentation	Olivier Couet (CERN)

Why? Why now?

- Major roadblocks ahead for HL-LHC, see https://arxiv.org/abs/2004.07675 for analysis software:
  - High statistics, sampling multi-dimensional theory parameters, systematics must be under control to match stat precision (correlations correlations correlations). Data volume + compute volume.
- Need to counterweight today's nationalistic focus on funding agency with (at least) a European view, to **remain** globally **relevant**

Context: National Facilities

- Facilities and software should be designed for each other
  - Decisions on storage, networking, CPU/memory can benefit from input from software (see US exascale)
- Evolve hand in hand, ideally
  - Designing a facility should motivate software evolution; plans in software evolution change recommendations on facility design
  - Example: ROOT is redesigning TTree from the ground up, may we optimize it for your (future) facility?

Context: National Facilities

- How much, on which scale is this information exchange happening? See Exascale on the benefits
- How can we improve this, who are facilitators - WLCG?
  - Happy (happier?) to do this on the "kleiner Dienstweg"
  - Software projects are agile, funding is not, so initial setup might benefit from facilitators' involvement

Facilitators?

- WLCG as coordination body
- Experiments as stakeholders
- HSF as communication forum
- Projects such as those within ErUM-Data, or ROOT
- Personal connections

Conclusion

- Software needs massive effort for HL-LHC, we cannot just continue
- I would not rely exclusively on CERN for innovation (that's good news!)
- Please do involve existing projects in your innovation and R&D efforts, also for facility design
  - Beneficial for both sides, e.g. better chances for grant proposal
  - Can be game changer wrt deployment and sustainability
- Increase cooperation beyond existing personal connections?

Goal: Win-Win

- Better software for community
- Less duplication of effort
- More visibility and relevance for all contributors
  - Effectively creating software for production use, not for CHEP papers