Experiments and Large External R&D Collaborators
Improving Integration and Planning

Paolo Calafiura

Background

Last year, the ATLAS International Computing Board (ICB) created a working group with the following mandate

1. Identify ongoing and possibly future R&D projects and activities relevant to ATLAS software and computing
2. Assist with rapid integration of useful outcomes, and avoid duplication of efforts
3. Make preliminary assessments of cooperation opportunities (with R&D projects)
4. Provide rapid feedback to ICB […] on sustainable pool of computing and s/w experts

In June, I presented to ATLAS the WG's findings.

Following Gordon's brief, this talk attempts to reframe these findings also from the POV of an external R&D collaborator

Large External Collaborators (LEC)

How do we define “large”?

- 15+ FTE*years?
- Multi-institutional
- Cross-cutting
- Run by an “EB” with advice from a “SB”

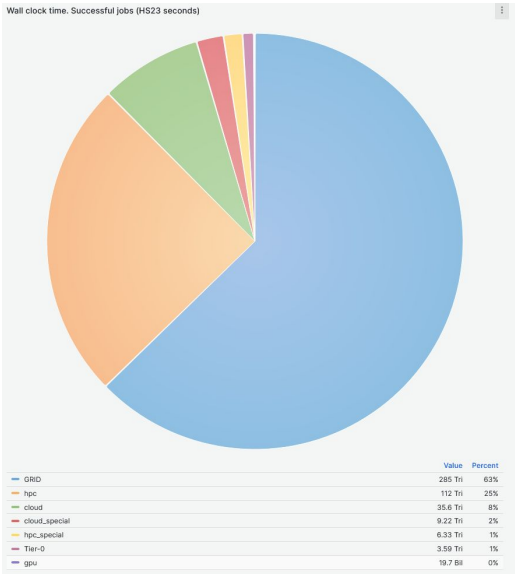
Examples: IRIS-HEP, HEP-CCE, CERN OpenLab, SWIFT-HEP, Excalibur-HEP

Not so large: AdePt, Celeritas, Optix, Exa.TrkX, MadFlow,...

Not so R&D: ROOT, Geant4

Not so HEP: Simons Foundation, SCIDAC Institutes, HPC early science programs

What's in it for the Experiments?

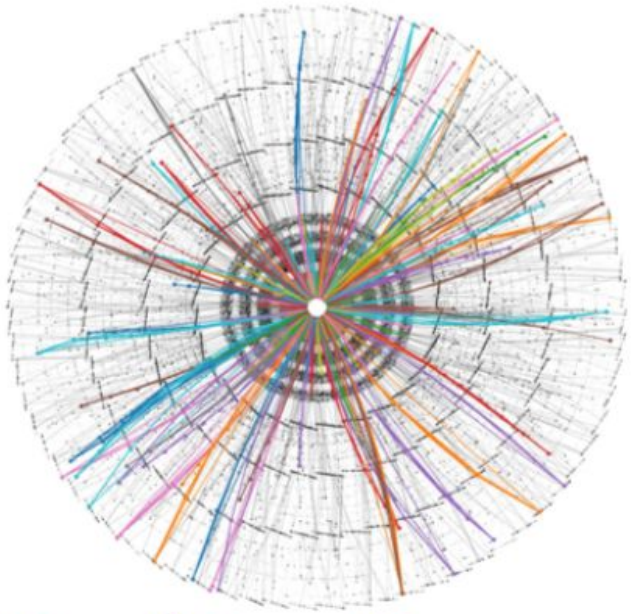


- **Gain access to new resources:**
  funding, facilities, know-how
  - HPCs provided 26% of ATLAS CPU over the last two years
    - After five years in production still not using HPCs as efficiently as the GRID.
- External R&D helps establish a sustainable **pool of computing and software experts**
  - Attract new people and train them on "hot skills"
  - Reward and motivate experienced developers
    - Encourage them to go outside their technical "comfort zone"
      - While making sure they remain focused on concrete needs.
- Outsource early-stage R&D projects
  - Share risk with other experiments
- Integrate successful R&D
  - Involve experiment developers since early stages

A *potentially* ✔ successful example of convergence and integration

ATLAS-specific, sorry!

New Tracking · acts · TrkX AND L2IT · detray · VecMem · EF Tracking · to the benefit of ATLAS EXPERIMENT



HL-LHC Tracking Demonstrators

A. Salzburger (CERN)
on behalf of and talking about work from **many others from many universities and institutes**

https://indico.cern.ch/event/1268248/contributions/5326275

Building a “framework” to integrate algorithms across projects

(As of May 2023, not updated since)

Demonstrator Matrix - Projected timeline

Name	*Environment*	*Implemented*	*Validated*	*Comment*
FastTracking *Reference*	21.9	Done.	Done.	Decommissioned when master validated.
FastTracking *Trk*	master	Done.	Q3/2023	Decommissioned when/if ACTS validated.
ACTS Tracking	master	Q2/2023*	End of 2023 (first version)	Ongoing improvements until feature freeze.
traccc Prototype	standalone	Q3/2023*	CKF chain	prototype on dummy detector
GNN4ITk	master	Done.	Q3/2023	Ongoing improvements, Q3/2024 EF demonstrator
GNNACTS	standalone/ master	Q3/2023	End of 2023 (proof of principle)	Should become a flavour of GNN4ITk
GNNtraccc		?	?	Not yet worked on

https://indico.cern.ch/event/1268248/contributions/5326275

* these are not final implementations.

Improving Integration and Planning

What should a successful "external collaborator" do?

- Gain stakeholders' trust
  - Navigate experiments technical, sociological, and bureaucratic challenges
    - Timeline, legacy code, jargon, access to documentation, data, publication rules, and many others
  - Establish close connections with (and across) experiments, but also facilities, CS/Math/AI research communities
  - **Push through the 80/20 barrier:** deliver solutions, not papers.
- Let experiments set the strategy. Respond to their evolving priorities
  - Embedding collaborators into experiments
- Deliver on a coherent, cross-cutting program that is more than the sum of its parts.
  - Promote cross-experiment R&D, avoiding duplication

Improving Integration and Planning

- What should experiments do:
  - Follow progress of proposed and active R&D projects, provide feedback on priorities, collaborations, and duplications
    - Best way to do that is to embed experiment folks in projects
  - Gather and disseminate information on funding opportunities and funding agencies priorities
  - Help CS researchers understand the experiment's computing needs
    - Present needs at CS venues (particularly funding agency "research needs" workshops).
    - Document needs in a set of white papers targeting CS researchers.
  - Help form collaborations with CS researchers and submit research proposals
    - Connect external researchers and experiment experts
    - Help navigate experiment requirements (data access, approvals, ...)
- Much of this work could be shared across experiments, facilitated by projects, through p2p groups like HSF

Coming Up: Data Intensive Computing

Funding agencies around the world are updating their supercomputing strategy to make it "data-intensive"

- CERN and CERN experiments seem to be at the center of European "data-intensive" planning
- In the US, initiatives like IRI and HPDF could have a big impact on HEP computing, and benefit from decades of HEP experience in this area.
  - Can HEP-CCE and IRIS-HEP help the US HEP community speak with one voice?