HSF and the Community
White Paper

Michel Jouvin
QCD Algorithms and Numerical Tools
Orsay, May 15, 2017
HSF Motivations: HL-LHC Software Challenges

• Pile-up x 10 = CPU x 100
  • Moore’s law over 10 years: only a x10
  • With a flat budget, Moore’s law is the real maximum that can be expected on the HW side

• HEP software hardly execute more than one instruction at a time (per thread)
  • Since ~10 years, CPU (core) power increase is due to the internal parallelism (Instruction-Level Parallelism: pipelines and vectorization)
  • x10 with the same HW only achievable if using the full power of processors: major SW reengineering required (but rewriting everything is not an option)
  • Accelerators like GPUs are of little use until the problem has been solved

• Increased amount of data requires to revise/evolve our computing and data mgt approaches
  • We must be able to feed our applications with the data efficiently

• « HL-LHC salvation » will come from software improvements, not from hardware

• HL-LHC is not unique: LSST, DUNE, FAIR...
HSF Goals

• Facilitate coordination and common efforts in software and computing across HEP and “friends”
  • HEP software must evolve to meet the challenges posed by new experiments
  • The computing landscape is evolving rapidly
  • No more free-lunch thanks to Moore’s Law: SW must use efficiently built-in HW parallelism, in particular Instruction-Level Parallelism (ILP)
  • Can’t just buy more hardware: budget and energy constraint

• Need to exploit all the expertise available in our community, and outside it, to meet the challenges and the affordable way to do it is collaboratively
  • All the HEP experiments are facing the same challenges, as well as next nuclear and astro physics ones
  • Some other communities have more experience/expertise with these parallelization issues
  • Not only a problem of computing techniques: many problems are intrinsically sequential, how to make them parallel: need to work on algorithms too.
  • Cannot afford anymore duplicated efforts/software: in the LHC experiments, each one has its own solution for almost everything (framework, reconstruction algorithms...)

31/01/2017
HSF Main Milestones

• April 2014: kick-off meeting for a HEP SW Collaboration (1st HSF workshop)
  • https://indico.cern.ch/event/326823/

• Second HSF workshop organized at SLAC in January 2015
  • Agenda: http://indico.cern.ch/event/357737/other-view?view=standard

• Third HSF workshop organized at LAL in April 2016
  • http://hepsoftwarefoundation.org/organization/2016/05/04/Workshop-summary.html

• Community White Paper Kickoff Meeting, SDSC, January 23-26, 2017
  • http://indico.cern.ch/event/570249/timetable/#all
A Bottom-Up, Transparent and Open Process

• A web site: http://hepsoftwarefoundation.org
  • Minutes of every (almost weekly) coordination meeting posted
  • Minutes of the main meetings from HSG WGs and activities
  • Information about activities going on, newsletter...
  • How to participate
  • Actively updated

• Several mailing lists: all open to everybody interested
  • Main one (low volume, announcements): HSF Forum, hep-sf-forum+subscribe@googlegroups.com
  • Topical ones can be found on the web site: http://hepsoftwarefoundation.org/forums.html
  • Also a general SW and Computing mailing list, not linked to HSF: hep-sw-comp+subscribe@googlegroups.com

• Need more volunteers: no dedicated manpower (yet)...
  • How to contribute: http://hepsoftwarefoundation.org/get-involved.html
HSF Activities...

- See web site for the details: http://hepsoftwarefoundation.org

- Training: identified as the critical short/medium term activity but not (yet) as active as we’d like
  - Online trainings: build/maintain a list of useful material
  - Link with WikiToLearn (https://en.wikitolearn.org/Main_Page) platform

- Software Project Visibility and Interactions
  - Software Knowledge Base: add the projects and events you know about!
  - Incubator: a few projects joined but not yet clear what can be done without a dedicated manpower

- Software Packaging: make easy to use an existing package, whatever build tools you are using
  - Active WG: progress made into identifying promising approaches, Spack as an appealing solution

- Technical Forum: share expertise
  - Technical Notes: several issued in 2016
... HSF Activities

• Software Licensing: one technical note in 2016, another round of discussions last February
  • http://hepsoftwarefoundation.org/organization/2017/02/21/licensing.html
  • Several open-source licenses: compatibility may be a problem preventing code reuse
  • Distinction between contributor recognition and viral licensing
  • ATLAS and CMS wants to move to an Apache2 license but a problem with some GPLed code, in particular generators

• Development Tools and Services
  • Nothing operated by HSF: relying on existing tools (GitHub, GitLab@CERN....)
  • Access to CERN TechLab platform: various new fancy hardware!

• Google Summer of Code (GSoC) : HSF took over from CERN as a GSoC organization
  • Umbrella for High Energy Physics community
  • Great success: ~35 project proposed, 26 students “allocated” by Google
    • 12 students last year for CERN
    • High profile students
HSF Events

• Main events are the HSF workshops: ~1/year since HSF inception
• Several topical workshops to discuss one category of tools and the associated ecosystem
  • Visualization workshop (March 28-30, 2017): http://hepsoftwarefoundation.org/events/2017/03/28/VisualizationWorkshop.html
  • Analysis Ecosystem Workshop (May 22-24, 2017): http://hepsoftwarefoundation.org/events/2017/05/22/analysis.html
• GeantV peer-review (October 2016)
  • https://indico.cern.ch/event/570876/
  • First time for this peer-review activity, planned since the HSF inception
  • Panel formed from the simulation experts in the community
  • Public review: open to everybody interested (Geant4 main people participated)
  • 3 days, very intense: detailed review of every aspect of GeantV
  • Very constructive discussions: all participants very satisfied
  • Set of public recommendations at the end
Community White Paper (CWP)

• Goal: a Community White Paper (CWP) that describes a global vision for software and computing for the HL-LHC era and HEP in the 2020s
  • Elements common to the HEP community and those specific to an experiment
  • Relationship with other communities/sciences
• The CWP will identify and prioritize the software research and development investments required:
  • to achieve improvements in software efficiency, scalability and performance and to make use of the advances in CPU, storage and network technologies
  • to enable new approaches to computing and software that could radically extend the physics reach of the detectors
  • to ensure the long term sustainability of the software through the lifetime of the HL-LHC
• The HSF is engaging the HEP community to produce the CWP via a community process
  • Initiated as an HL-LHC planning process: official charge from WLCG
  • Aiming for a broader participation (LHC, neutrino program, Belle II, linear collider so far)
  • The resulting roadmap will be used for the HL-LHC computing TDR and other strategic plans
CWP Kickoff Workshop (San Diego, Jan. 23-26)

• ~110 participants, mainly US + CERN
  • Unfortunately very few Europeans outside CERN
  • Some non-LHC participation (FNAL-IF, China, Jlab, ILC...): we hope (and work!) to increase it
  • Not only the usual suspects: some physicists involved in trigger, reconstruction, machine learning...

• CWP-related material: http://hepsoftwarefoundation.org/activities/cwp.html
  • All the WG documents are public
  • Agenda: http://indico.cern.ch/event/570249/timetable/#all
  • Live notes (not polished) linked from Indico

• 2.5 days of parallel topical WG meetings
  • From infrastructure to reconstruction and analysis, through simulation, data management...
  • Notes from (almost) all WG discussions in the WG Google Docs, summary slides in the agenda
  • 1 day about Machine Learning, a hot topic in the community, with more and more coordinated activities covering an increasing number of areas

• Very productive kickoff with a general good will to collaborate around this process
  • Everybody is aware that this is the only chance to get some (limited) additional funding for our needs
CWP Current Activities

• Final CWP meeting in Annecy, June 26-30, 2017
  • Goal: each WP must have a final document by then
• Each WG currently preparing its CWP part
  • [http://hepsoftwarefoundation.org/cwp/cwp-working-groups.html](http://hepsoftwarefoundation.org/cwp/cwp-working-groups.html)
  • All document drafts publicly available
  • One mailing list associated with each WG
  • Entering the very active phase of the process!
• This is still time to join WGs: this is really an open process
  • Each WG has a mailing list: register to it
  • Be sure to register to HSF Forum and the CWP list: general information on the process sent there
How to Join/Follow HSF Activities?

**HSF**

- A web site: [http://hepsoftwarefoundation.org](http://hepsoftwarefoundation.org)
  - Up-to-date information on HSF Activities and events
- Subscribe to mailing lists
  - Main one (mainly announcements): HSF Forum, [hep-sf-forum+subscribe@googlegroups.com](mailto:hep-sf-forum+subscribe@googlegroups.com)
  - One or more topical ones you are interested in: [http://hepsoftwarefoundation.org/forums.html](http://hepsoftwarefoundation.org/forums.html)
  - Also a good idea to subscribe the general SW and Computing mailing list, not linked to HSF: [hep-sw-comp+subscribe@googlegroups.com](mailto:hep-sw-comp+subscribe@googlegroups.com) (event announcements, moderated, very low volume)

**CWP**

- CWP main mailing list: [hsf-community-white-paper+subscribe@googlegroups.com](mailto:hsf-community-white-paper+subscribe@googlegroups.com)
- CWP-related material and WG documents: [http://hepsoftwarefoundation.org/cwp.html](http://hepsoftwarefoundation.org/cwp.html)
Algorithm-Related Activities in HSF

- Algorithm-related work is an important part of HSF goals
  - Performance challenge cannot be addressed only by engineering techniques
- An early attempts for track reconstruction
  - Currently mainly a set of activities related to this topic with some connection to HSF (and its people): Connecting the Dots conference series, ACTS experiment-agnostic framework, AIDA2020
- Machine learning: a strong cross-experiment activity in HEP
  - http://hepsoftwarefoundation.org/forum_ml.html
  - 1 full day devoted to ML during // sessions in CWP San Diego meeting
- Numerical tools for QCD are also fitting into HSF
  - HSF can give some visibility to existing developments
  - QCD is more than HEP and inline with our goal to be more than HEP
  - Many ways to benefit from HSF: project incubator, WG, events...
Journal on SW and Computing in Big Science

• Not really an HSF activity but closely related to HSF goal of sharing knowledge and helping with career recognition
  • Currently no place acting as a reference archive for SW and computing work done in our community

• Proposal of a journal filling this gap made by a few persons early 2016 and materialized into the Journal on Software and Computing for Big Science, launched early this year
  • Refereed journal in partnership with Springer editor
  • Pure open-access during the first 2 years: long-term business model will depend on its inclusion in SCOAP3 next phase
  • Covering all aspects of SW and computing, from online to analysis through tracking, reconstruction, algorithms, visualization, data management, SW performance, resource provisioning...

• Several key HSF members in the Editorial Board
  • See web site for the full list

• Article proposals are welcome!
  • More than HEP: all big sciences
Summary

• HSF has been as an attempt to build a world-wide collaboration around software used in HEP and related communities
  • HL-LHC as a driver but not as an exclusive stakeholder
  • Choice of a bottom-up approach to be inline with the project needs and to maximize the buy-in
• Despite the absence of dedicated effort, now well established and recognized as a positive initiative
  • Several cross-experiment projects going on: new generation of Conditions Database, packaging tools for easier interoperability, new approaches for tracking...
  • GeantV Review: a successful peer-review organized by the community
  • GSoC: HSF established as the umbrella organization with a large participation
• Community White Paper: a unique effort by the community to think about its future, identify challenges and to propose a roadmap to address them
  • Not a “one-size-fits-all” approach but an attempt to identify areas where collaboration is particularly important and/or possible, even if solutions may be different
  • Several topical WGs open to all people interested who are ready to spend some efforts to contribute
  • The CWP should be the base for upcoming TDRs and other experiment specific strategic plans