

HEP Software Foundation

HEPiX

March 23, 2015

Michel Jouvin

<http://hepsoftwarefoundation.org>

Why HSF?

Facilitate coordination and common efforts in HEP software and computing

- 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

Objectives

- Share expertise
- Raise awareness of existing software and solutions
- Catalyze new common projects
- Promote commonality and collaboration in new developments to make the most of limited resources
- Aid in creating, discovering, using and sustaining common SW
- Support training career development for S&C specialists
- Framework for attracting effort/support to S&C common projects
- Provide a structure to set priorities and goals for the work
- Facilitate wider connections: form the basis for collaboration with other sciences

HSF Prehistory

- April 2014: kick-off meeting for a HEP SW Collaboration
 - <https://indico.cern.ch/event/326823/>
 - Very large participation: ~150 people
 - Broad spectrum of views but preference for a lightweight structure
 - Call for White Papers (WP) to express what it could be or should not be
- Spring 2014: 10 WPs received from different geographical and “scientific” horizons
 - Differences but agreement that building a lightweight collaboration would be beneficial and that it should be bottom-up, based on motivated individuals and projects rather than organisations

HSF First Steps

- July 2014: Interim Foundation Board (iFB) created
 - Made of WP authors and other people interested
 - Misnamed: in fact a “general assembly”
- Sept. 2014: Startup Team created
 - Initially 6 people, now 14, reporting monthly to iFB, public summaries
 - People agreeing to spend part of their time to bootstrap HSF
 - Preparation of a HSF “kick-off” workshop at SLAC, January 2015
- Fall 2014: synthesis of WPs + proposal of a HSF startup plan
 - <http://hepsoftwarefoundation.org/sites/default/files/HSFwhitepaperanalysisandstartupplanV1.1.pdf>
 - HEP web site created: <http://hepsoftwarefoundation.org>
 - Contacts with several communities: IF, astrophysics, MC generators...

HSF Startup Team Members

- Amber Boehnlein (SLAC)
- Peter Elmer (Princeton)
- Daniel Elvira (FNAL)
- Frank Gaede (DESY)
- Benedikt Hegner (CERN)
- Michel Jouvin (LAL, IN2P3)
- Pere Mato (CERN)
- Dario Menasce (INFN)
- Elizabeth Sexton-Kennedy (FNAL)
- Graeme Stewart (Glasgow)
- Craig Tull (LBNL)
- Andrea Valassi (CERN)
- Brett Viren (BNL)
- Torre Wenaus (BNL)

SLAC Workshop

- January 20-22, 2015
- Important milestone: 1st workshop of HSF
 - Validate ideas elaborated by the startup team after the kickoff workshop (April) and the White Papers
 - Assess if enough people/projects were interested to implement them
- Good attendance: ~100 people (80 local + 20 remote)
 - Lower attendance than April workshop
 - Suffered the late announcement but a high quality participation
 - Good non-European participation: mainly US but also Asia
 - Many non “pure HEP” experiments: Dayabay, LSST, Photon science...

Workshop Topics

- Agenda: <http://indico.cern.ch/event/357737/other-view?view=standard>
 - 2 + 1 days
 - Designed to allow a lot of interactions: worked well
 - 41 Short presentations (6')
- Workshop main sessions
 - “Learning from others: 3 “long” presentations from “similar” projects
 - Views on HSF by experiments, projects, individuals
 - New projects that could benefit from HSF (7 presented)
 - Discussion on concrete next steps: ~1 day
 - ½ day with the full attendance
 - ½ day after the workshop (iFB) to digest/refine workshop discussion
 - Meeting with LSST, GEANT4 and LCLS/Photon science

Learning from Others...

- Very interesting and useful session
- Apache Software Foundation
 - Goal similarity with us: umbrella for related projects, no central management of projects, they remain autonomous
 - Difference: ASF started before projects, invented the model when developing
 - Do-ocracy: no long-term planning, active people have their say
 - Darwinian approach: ASF provides an infrastructure for projects, users decide the projects that will survive by their adoption
 - ASF focuses on providing an incubator for new project and on ensuring the project sustainability
 - Avoid projects bound to 1 individual (hit-by-the-bus problem!)
 - Transparency is essential: a pillar of ASF culture

... Learning from Others

- D. Katz on Building Scientific Software communities: a very nice summary on lessons learnt from successful and failed projects
 - Avoid too much planning, try-and-fail is the most productive approach
 - Governance: flat layer of peers generally better than benevolent dictator to create a community: forces to work together
 - Make easy for people to contribute, with little time and effort
 - Give credit for every work done, motivate people
 - Get people involved rather than having them reinventing the wheel
- Software Sustainability Institute (UK) - Neil Chue Hong
 - Helps SW projects to address sustainability, great focus on training
 - Same message as D. Katz, insistence on not designing the perfect HSF
 - Lobbying/communication about career path for Research SW Engineers

Community and Project Views

- Every community and project mentioned that HSF could help in some ways
- No real conflicting view but different focus
 - Experiments: SW knowledge base to increase SW reuse, consultancy for new projects, SWAT teams, consistent build/packaging tools across projects, build/test infrastructures, teaching, licensing
 - Projects: technical forums, help in organizing technical discussions with other projects, help in organizing meetings with users, build/test infrastructure for smaller projects, licensing
 - Common SW or expertise: avoid to reinvent the wheel (example with HPC), help with convergence and sustainability (pyroot/rootpy)

New Project Initiatives

- Examples of innovative projects that could benefit from HSF
 - fads : Go-based detector simulation toolkit (1 individual)
 - Condition DB for Belle2: discussions started with CMS and ATLAS
 - Find grained event processing with an event service, based on ATLAS experience
 - Acceleration simulation/modelling framework (BNL)
 - HEP SW Knowledge Base based on existing prototype
 - HepSim: repository of theoretical predictions for HEP

Non Topic: Governance

- Big difference with kickoff workshop in April
 - Probably everybody convinced it was the thing to avoid...
- Large consensus established in the last 6 months that HSF should be a light structure without a too formal management
 - Apache model seen as a good reference
 - Continue with the existing Interim Foundation Board + Startup Team
 - iFB: misnamed, in fact a “general assembly” of all people interested by HSF, meeting once a month with Startup Team
 - Startup Team: ~12 volunteers to propose ideas and help with their implementations
 - Encourage volunteers to take responsibilities in the different activities promoted by HSF: already several raised their hands at SLAC!
- Be transparent and open to other communities

HSF Website

- <http://hepsoftwarefoundation.org>
- Website update reflecting workshop outcomes and maintained up-to-date
 - **Main contact point for HSF**
- Main pages:
 - **Foundation**
 - **Events**
 - **Activities**
 - **Get involved**
- Left-side website overview box with direct links to activities/WGs (also linked from Activities page and referenced from Get involved page)

Activities

[View](#) [Edit](#) [What links here](#) [Outline](#) [Revisions](#) [Track](#) [Access control](#) [Nodequeue](#)

Working groups

The HSF initiates, plans and coordinates activities via Working Groups. Those so far established or being established are as follows. All are welcome and encouraged to participate.

- [Training](#)
- [Software Packaging](#)
- [Software Licensing](#)
- [Software Projects](#)
- [Development tools and services](#)
- [Communication and information exchange](#)

Discussion Fora

- [Concurrency Forum](#)
- [Reconstruction Algorithms Forum](#) (under discussion)

Projects

HSF seeks to serve new and emerging common projects through a project incubator activity. Templates to guide and aid new projects are being established. The Projects WG is working to establish project membership/participation/life cycle levels and for the principal focus of incubator projects, put together useful templates. Projects with a declared interest in involvement with the HSF include the following. If you'd like your project to be involved just let us know. Talk to any member of the startup team or email hep-sf-startup-team@googlegroups.com.

- [ROOT](#)
- [Geant4](#)
- [xrootd](#)
- [Gaudi](#)
- [pyroot](#)
- [rootpy](#)
- [GenFit2](#)

Services

The HSF is establishing services identified at the SLAC workshop and in consultations as useful to the community, including

- [HEP Software and Computing Knowledge Base](#) hosted at this website
- [Software development and build services](#) leveraging the resources at our major Laboratories

Mailing lists

- Google-based self-signup lists (no need for a Google account)
 - Send a mail with 'subscribe' as the *subject* (no content) to <listname>+subscribe@googlegroups.com, e.g. for the list above, hep-sw-comp+subscribe@googlegroups.com
- HEP S&C community website (190 members)
 - <http://groups.google.com/d/forum/hep-sw-comp>
 - General S&C discussions, not related to HSF
 - Please encourage your communities to sign up!
- HSF Forum (currently 101 members)
 - <http://groups.google.com/d/forum/hep-sf-forum>
- HSF Startup team: the only list without self-signup
 - hep-sf-startup-team AT SPAM NOT googlegroups.com
 - Submission open to everybody
- WG-specific lists: open to anybody interested

Working Groups

Working Group	Objectives	Forum - Mailing list
Training	Organization of training and education, learning from similar initiatives	hep-sf-training-wg
Software Packaging	Package building and deployment, runtime and virtual environments	hep-sf-packaging-wg
Software Licensing	Recommendation for HSF licence(s)	hep-sf-tech-forum
Software Projects	Define incubator and other project membership or association levels. Developing templates	hep-sf-tech-forum
Development tools and services	Access to build, test, integration services and development tools	hep-sf-tech-forum
Communication and information exchange	Address communication issues and building the knowledge base Technical notes	hep-sf-tech-forum

Next Steps: Projects WG

- Experiment the incubator idea with a few guinea pig projects
 - What services to support projects (what they would benefit from HSF)
 - What they can bring to the community
 - Which criteria to move out of the incubator: avoid too high thresholds, remain pragmatic
- Experiment with the inter-project relationships under HSF umbrella
- What kind of membership or HSF affiliation for projects
 - One or different types?
 - Membership doesn't mean a recommended project
- Provide some templates for new projects
- A few projects declared their interest: rootpy, Gaudi, Vac/Vcycle, FADS, GenFit2 (from Belle2), xrootd

Next Steps: Technical Forum

- General agreement that some sort of technical forum is needed
- Started as a mailing list: hep-sf-tech-forum@googlegroups.com
 - Evaluating other forum technologies
 - Favor wide discussion, split to specific discussions only when really needed
 - Goal: have different expertises meet together
 - Publish technical notes on various topics relevant to the community: not blessed recommendations, more sharing of expertise
 - Later, could discuss a RFC-like process for some recommendations
- An already existing, successful “prototype”: the concurrency forum
 - Propose to continue as part of the HSF Technical Forum, with the specific focus on parallelism and concurrency

Next Steps: Training

- Consensus that it should be one of the initial HSF focus
 - Several volunteers... and several existing initiatives being leveraged in the HSF context
 - Should cover different kind of training from “traditional schools” to virtual seminars or webinars
- Learn from Software Carpentry very successful experience
 - Launched by Sustainable Software Institute
 - A very valuable input from the workshop
- A working group created and active, led by G. Stewart
 - **Volunteers welcome to join**
 - hep-sf-training-wg@google.com

Training WG

- Existing schools of HEP computing (CERN, Bertinoro, GridKa,...) are notable excellences but:
 - Format is restricted to, at most, one/two weeks
 - Intended audience knowledgeable on programming
 - Expensive: tutors and students must travel and be hosted
- Significant gap between programming expertise acquired at university and what is required by physicists/scientist to perform good quality research
- Students would benefit from lessons aimed at bridging this gap
 - Format must be light-weight: best one is webinars
 - Very cost effective, requires only volunteers, no travel
 - Lessons covering broad spectrum of specializations
 - Already some initiatives like [Software Carpentry](#) or [WikiFM](#)
 - Collaboration with them is a possibility to explore
- Other possible approach: establish and maintain our own platform

Next Steps: SW Knowledge Base

- Started at <http://hepsoftwarefoundation.org>
 - May migrate to a new platform in the future but data will be migrated
- About all the software developed and used in the community
 - not a blessing of registered SW
 - no formal review
 - ability to comment and cross-reference the usage by experiments
- **Everybody encouraged to register its favorite software**
 - Request an account on the web site

Next Steps: Development Tools

- Build/Testing/Continuous Integration
 - A clear need mentioned by several projects
 - Already some experience at several (big) sites willing to contribute
 - Try to come up with more concrete proposal in the next months
- **Will not do project hosting:** use GitHub if needing one
- Several labs with hardware and software resources that could benefit to other projects
 - E.g. CERN TechLab, FNAL...
 - Currently discussing how to make it possible for HSF projects to get access to them
- WG being formed

Next Steps: SW Packaging

- Would be useful if the HSF provided a validated, coherent software stack of (near) current versions of popular packages
 - e.g. the externals used by Geant4
- WG created: already active discussions
 - Package Building and Deployment
 - Runtime Environment / Virtual environments
 - Role of new technologies like Docker
 - Define a build 'protocol' to combine various projects into combined builds
 - Allow adoption by well established projects and new users
- Created an HSF organization at GitHub

Next Steps: Licensing

- An open-source license is mandatory to participate to HSF
- A priority, much discussed at the workshop
 - Many SW projects in the community without an explicit license
- Should HSF recommend one type of open-source license?
 - Need to start with what exists...
- A WG being formed
 - Building on CERN's investigations: CERN study available on the web site
 - First conclusions/recommendations by the summer

Other Possible Activities

- Consultancy, SWAT teams, peer reviews...
 - Ready to start some activities on these but waiting for a concrete request to refine how to do it
- Access to scientific journals
 - Not a priority for the short term, concentrate on technical notes
 - Springer expressed some interest for HSF and providing some “space” for it

Conclusions

- HSF is a major initiative to organize collaboration around SW development
 - Bottom-up: do-ocracy, based on motivated individuals/projects
 - Will be what people make out of it
 - Open to other communities sharing interest and SW with HEP
- Concrete actions are starting... several “doable” areas identified
 - Already some volunteers, more welcome!
- Next milestone: CHEP face-to-face meeting (Friday afternoon)
 - Report on first concrete actions
- Engage with more projects and more communities
 - First target for communities: nuclear physics and astroparticles
- **Encourage people to join!** (see web site)

To Learn More about HSF...

- HSF web site: register to mailing lists, ask an account and contribute!
 - <http://hepsoftwarefoundation.org>
- Meetings and summaries
 - <http://indico.cern.ch/category/5816/>
- White Papers
 - Attached to <http://hepsoftwarefoundation.org/documents>
- Synthesis of White Papers
 - <http://hepsoftwarefoundation.org/sites/default/files/HSFwhitepaperanalysisandstartupplanV1.1.pdf>
- SLAC workshop agenda and summary
 - <http://indico.cern.ch/event/357737/other-view?view=standard>