



HEP Software Foundation Workshop Summary

Benedikt Hegner

PH-SFT Group Meeting
2.01.2015

The HEP Software Foundation

- New effort to coordinate/communicate on common efforts in HEP software and computing
- Motivation:
 - The current HEP software generation is > 20 years old
 - Technology (hardware and compilers) evolved
 - Paradigm shift due to many-core CPUs
 - New resources such as HPC, commercial clouds, etc waiting to be exploited
 - New software needs to interoperate with other software - inside and outside our community
 - We need to attract new people with the right skills
- The [Concurrency forum was a nice example](#) how this might work

What happened so far...

- On 3-4 April 2014 a (brain) storming workshop at CERN
- 10 White Papers with ideas for scope, goals, formation process, governing models, etc.
- **Interim Foundation Board (iFB)**
 - Formed by WP authors and other interested people
 - Started meeting from mid July
 - Favored a **bottom-up approach**: invite projects to join ('endorsed' and 'hosted'), produce specific proposals on services and eventually agree on a governance strategy
 - Decided to **assemble a startup team** of volunteers with a broad representation and expertise interested in getting the HSF started
 - Participation from 10 different labs/universities

Goals identified

- 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 developers and users in creating, discovering, using and sustaining common software
- Support training career development for software and computing specialists
- Provide a framework for attracting effort and support to S&C common projects
- Provide a structure for the community to set priorities and goals for the work
- Facilitate wider connections; while the HSF is a HEP community effort, it should be open enough to form the basis for collaboration with other sciences
- Apache foundation as an example for how the collaboration could look like

Communication channels

- Website:
 - hepsoftwarefoundation.org
- HSF forum (open-list, iFB):
 - <https://groups.google.com/forum/#!forum/hep-sf-forum>
- HEP S&C community website:
 - <https://groups.google.com/forum/#!forum/hep-sw-comp>
 - 180 people registered so far

The HSF workshop

- ...and of course the HSF workshop (20-22 Jan) at SLAC:
- <http://indico.cern.ch/event/357737/>
- 10 sessions
- 44 presentations - 7 from our group!
- 80 local and ~20 external participants

The Workshop's goals

- Hear from small and large projects what HSF could bring to them and what they can bring to it
- Hear from other similar initiatives in different contexts like Apache SW Foundation
- Hear from experiments, science communities and individuals
- Hear views from funding agencies
- Discuss related initiatives that could benefit from HSF or could be launched under the HSF umbrella
- Set up the relative priorities between the possible activities/services for HSF and ensure we have the required resources committed by participants and funding agencies

Workshop Agenda

- Organized in 10 sessions:
 - Current Status
 - Software Collaboration - Learning from others
 - Community views: Sciences & experiments
 - Project views: Learning from experience
 - Community views: Common software needs & opportunities
 - New project initiatives
 - National, agency and Institution Views
 - Next steps
 - iFB Meeting
 - HSF-community discussions

Workshop Agenda

- Organized in 10 sessions:
 - Current Status
 - Software Collaboration - Learning from others
 - Community views: Sciences & experiments
 - Project views: Learning from experience
 - Community views: Common software needs & opportunities
 - New project initiatives
 - National, agency and Institution Views
 - Next steps
 - iFB Meeting
 - HSF-community discussions

Learning from others

- **The Apache Foundation** - Brian Behlendorf
 - Umbrella for related technologies. Addressing the concern of standardization
 - Currently 300 projects, 600 members and 2000 contributors
- **Darwinian approach**
 - Provide an environment for projects, but let users decide who survives
 - Right to fork is essential!
 - No benevolent dictator!
- However, foundation guarantees that every project is survivable on its own
 - Projects go through an **incubator step** to become members
 - checking participation, aligning to community, usability, ...
 - Projects are reviewed from time to time and sometimes the board “pulls the plug”
- Transparency is essential
- A **do-ocracy** - no overall planning and 3 year-plans and active people have the say

Learning from others

- **Building scientific software communities** - Dan Katz (doi.org)
 - Not software, but people is the problem to solve
 - Motivate people, give credit - as well in their career
 - Help people to participate (low entry level, documentation)
 - Be careful about licensing (GPL too 'restrictive')
 - Re-invention of the wheel a continuous problem
 - Favours Apache governance model
 - Do rather than plan

Learning from others

- **The Software Sustainability Institute** - Neil Chue Hong (<http://www.software.ac.uk/>)
 - His personal experience in BaBar - no gratification for his work
 - Created “Software Carpentry” for training and exchange; raising awareness to others
 - Communicating need for career path for “research software engineers”
 - Impact of bad or missing software on projects’ progress and success
 - His advice to us: don’t try to come up with the perfect solution from the start
 - Benevolent dictator a too common pattern in SW software
- The website is really worth a read !

Community Views

- Input from astro particle community, LHC experiments, linear collider, FCC, Belle 2, ...
- Topics raised multiple times:
 - Need for a software knowledge base to increase re-use (“steal before you write”) and to know status of projects (e.g. are they still alive?)
 - Expertise and consultancy for newly starting projects
 - Value of discussion fora and exchange
 - Confusion on which license to use
 - Build/testing infrastructure and continuous integration
 - Consistent software distributions and build tools
 - Urgent need for better training
 - Lack of expertise in web based technologies
 - Motivation is the key to success

Community Views

- Concerns about
 - Kind of open-source license - L(GPL) vs. Apache style
 - Re-usability of components if too experiments specific
 - Level of support
 - “only” common build tools vs. providing actual binaries
 - Freedom within the HEP Software Foundation
- Only a few presentations listed existing or previous cooperations. :-)

Summary Discussion

- **HSF organization**

- The Apache Foundation seems a good model
 - Adopt incubator approach
- Stay open and transparent
- Don't convey that 'member' projects are 'recommended' projects
- Try with focussed project teams to address the issues discussed
- Common (?) copyright issue needs to be resolved
- Stay open to other communities...
 - ...but start with ours first

Proposed HSF efforts

- **SW knowledge base to know what projects are around**
 - sign up and put your favorite software in
 - Don't hesitate updating what is already there
- **Project hosting infrastructure**
 - Broad consensus that this is well covered by open-source tools and portals (e.g. github) already
- **Build, packaging and testing infrastructure**
 - Assemble people to look into possibilities here
 - Discussion among LIM participants started already
 - Welcome side effect on porting to new compilers, etc

Proposed HSF efforts II

- **Access to computing resources on many platforms and architectures**
 - CERN volunteered to look into this via the TechLab
- **Access to development tools**
 - can community benefit from lab-owned licenses?
 - Need people to look into this
- **Consultancy for new experiments and projects; peer review**
 - **High priority**, in concert with knowledge base
 - Will set up a list of experts for the various fields
 - Provide a 'project template' for an easy start
 - May think about "SWAT" teams if urgent need somewhere

Proposed HSF efforts III

- **Addressing the licensing issue**

- bring experts together and see if we can quickly agree on a zeroth version of recommendations.
- OpenLab at CERN has a lot of expertise. G4 put hundreds of FTE-hours into selecting a license.
- Open source licensing addressed at CERN.
- The startup team should volunteer or find someone to coordinate this.

- **Access to scientific software journals**

- HSF technical notes as first steps
- Maybe a new journal with Springer? They are open to this idea

Proposed HSF efforts IV

- **Training and technologies and tools**
 - Get inspired by software-carpentry.org
 - Prepare repository of existing teaching material and webinars
 - Need to identify volunteers and “training officer”
 - Newly created mailing list:
 - hsf-sf-training-wg@google.com
 - Fear that this task may become too huge

Next steps

- The startup team will call for volunteers to work on the items mentioned
- The next face to face meeting to review the progress is planned at CHEP
- **HSF will only be a success if people join !**
 - **How and where can our group contribute?**