Summary of the HSF Workshop in Paris

Michel Jouvin (inspired from Benedikt Hegner)
WLCG GDB, May 11, 2016

HSF

HEP Software Foundation

hepsoftwarefoundation.org

- LAL Orsay, May 2-4, 2016
    - 1st workshop at SLAC, January 2015
- Agenda: https://indico.cern.ch/event/496146/timetable/
    - Mix of general discussions and topical sessions
    - Several hands-on sessions
    - Geant4 Technical Forum
- Good participation: ~70 people
    - A good proportion of new faces, compared to previous workshop
    - ~30 affiliations
    - Good balance Europe vs. North America (US)
    - Various experiments represented, in particular Intensity Frontier and Belle II
- Co-located event: Geant4 Technical Forum (2h)

Quite a while ago the HSF called for a logo contest.

There is now a winner:

(submitted by J. Lingemann; 17 votes in 37 responses)

- 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 career development for SW and computing specialists
- A framework for attracting support to S&C common projects
- 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

- Jan 2015: SLAC HSF workshop established concrete activities and next steps

- Apr 2015: HSF meeting at CHEP 2015 on progress, opportunities, next steps

- Mid 2015: Survey and discussions of tools & approaches in Packaging WG

- Sep 2015: Technical Notes policies published and TN series begun

- Sep 2015: HSF on WikiToLearn

- Oct 2015: Evaluation of new Knowledge Base finished, hepsoftware.org deployed

- Nov 2015: Second generation HSF website deployed

- Early 2016: Startup team meetings replaced with weekly open HSF meeting

- Early 2016: Intensive packaging activity around Spack

- April 2016: Asked to organize a review of GeantV

- April 2016: Refurbishing of the HSF website

Working Group	Objectives	Forum - Mailing list
Communication and information exchange	Address communication issues Build the SW&C knowledge base Technical notes	hep-sf-tech-forum
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

A very dense programme

- **DIANA-HEP** (USA)
  - Centered around analysis tools with ROOT as its core
- **AIDA2020** (EU)
  - Many cross-experiment projects with the same spirit as HSF
- **Future Conditions Database**
  - Common project between ATLAS/CMS for next-generation conditions handling
  - Belle-II and LHCb had a closer look at it as well; no conclusion yet
- **HEP S&C Knowledge Base**
  - Central point to add and find information about SW, **now in production**
- **WikiToLearn**
  - Platform for "Collaborative Textbooks", not HEP-specific but committed to support HSF
  - Growing community (160 contributors)

HSF is built for and rely upon SW projects
- What does it mean being an HSF project? What's the benefit to projects and users?

Two concrete advancements in the last year
- Software Best Practices Document (draft)
- Project Starter Kit (*hsf_create_project.py* in https://github.com/HEP-SF/tools)
- Both the document and the starter kit considered good approach

Future
- Help with visibility of projects
- Interoperability of projects ⇒ should actively be worked on
  - Explore the possibility of automatic integration tests
- Project peer-review: will start with GeantV

During 2015 the Packaging WG made a review of the packaging tool landscape
- What tools are around - inside and outside the community?
- What are the requirements of the various stakeholders?
- Is there a chance to reduce the zoo in HEP SW?

One very promising candidate identified: Spack
- From HPC world: one requirement was to support multiple version of the same package
- Discussion (~30 people) in a dedicated session, introduced by Patrick Gartung (FNAL)
- Great interest: a very good tool for a broad range of use cases... but probably not all

Future
- Continue work for adopting it in HEP
  - Maintain add-on repo "hep-spack" for HEP specific SW
- Invest in development of features not yet supported (many PR exist already)

Bioconductor project (W. Huber; EMBL)

- Open source project portal for an entire SW ecosystem
  - Supported by a core team of people
- Peer review of software important
  - Development on github, open-peer review
  - Research papers get corresponding paper about the SW
  - Initially people feared their ideas/work are taken away, now their careers benefit from it
- Trade-off between standard settings / non-duplication and being inclusive/open

Netherlands eScience Center - D. Remenska

- An organisation bridging scientific communities and computing infrastructures
  - Core team of experts in various computing fields
- Funding by NeSC: allow to retain expertise and to do the "technology transfer"

10

Depsy (J. Priem)

- Platform / infrastructure to promote credit for software in science
  - Depsy.org
  - Funded by NSF
  - http://www.nature.com/news/the-unsung-heroes-of-scientific-software-1.19100
- Automatic analysis of papers and software for citations/usage
  - Text-mining in papers
  - Code-mining in GitHub repositories
  - Focussing on Python and R
  - Including transitive credits / indirect contribution
  - Relies on project catalogues (like PyPI)
- Wasn't applied yet/tried for C++ code bases
  - May be harder: no well-established central repository/catalogue for packages

- **½ day :** variety of Machine Learning talks with a focus on Software and Tools
  - Summary of OpenLab Workshop (industry)
  - TMVA updates
  - Machine Learning Challenges (e.g. HiggsML)
  - Deep Learning
  - Common Tracking Software (ACTS)
- RAMP (Rapid Analytics and Prototyping) Hands-On Session on Anomaly Detection (3h)
  - Searching for anomalies (distortions) in a modified HiggsML challenge dataset
- Inter-experimental Machine Learning Working Group (IML)
  - A cross-experiment WG to share work and ideas about ML in HEP, wider than LHC experiments
    - Endorsed by all LHC experiments, participation from DUNE, NOvA…
  - Monthly meeting open to anybody interested
  - HSF forum for Machine Learning

- Organized by V. Innocente (CMS), P. Calafiura (ATLAS) and M. Schulz
- Contributions by ALICE, ATLAS, CMS, GeantV, ROOT, Art/LArSoft, and the Astroparticle community
- Followed by a discussion session
  - What hardware to focus on? Commodity vs. GPU vs. HPC
  - Still no common understanding and idea on what to expect
  - Can we decouple low-level optimization (experts) from high-level code (physicists) via libraries?
    - Interesting input from Astrophysics
- Lead to *more questions than answers*
  - All the details in the **meeting notes** and the upcoming workshop summary

Outcome

- Needs a more visible and continuous activity in the HSF
  - Concrete follow up in the SW Technology Forum (former Concurrency Forum)
- Important input to the Community Whitepaper (see later)

HSF demonstrated some initial collaborative activities but to address the challenges ahead (e.g. HL-LHC) need more and dedicated resources

- Community roadmap describing HEP S&C challenges and directions (similar to P5 for HEP experiments in the US)
  - Idea proposed/supported by several US colleagues and presented by P. Elmer
- Proposing a series of HSF-branded workshops with a conclusion mid-2017
- Need participation of funding agencies, coordination bodies...

Discussion outcome

- LHCC asked for similar document: WLCG in charge of preparing a computing TDR for HL-LHC
  - Timescale similar/compatible with the CWP proposal
- Overall consensus that this is high priority
- Possibly co-locate a kick-off workshop for the CWP with CHEP

- **Proposal by German colleagues and Springer**
  - Presented at HEPiX mid-April
- **A journal refereed, abstracted, indexed about HEP computing**
  - Authoritative and central reference archive
  - Help with career paths
  - Do not restrict to HEP strictly: open to Data-Intensive Physics. Fields organized as large collaborations around large-scale experiments
  - Scope should cover all aspects of computing: from infrastructure to data analysis…
  - Continuous publishing: no paper/volume
- **Several open questions**
  - Exact size/focus
  - Business model: hybrid (OA access per article, based on authors choice/constraints) vs. pure OA
    - Pure OA requires sponsorship.
  - Article types: regular articles, reviews, advanced tutorials, no letter, no proceedings
  - Editorial structure

- **HSF is well alive**
  - Significant progress and increasing motivation compared to one year ago
- **Community White Paper recognized as a good incentive to progress towards more common views and projects**
- **HSF communication: explore the use of StackExchange**
  - Well-identified, open forum for questions about HEP computing
  - An alternative for (some) mailing-list based forums in HSF?
- **Develop/increase project support of HSF**
- **Look for an "official blessing" of the HSF by bodies like ECFA/ICFA**
- **A legal entity to support HSF: still far from a consensus but agreed to explore the possibility with funding agencies and lawyers**
  - Main initial goal: IPR management, in a way similar to the Apache SW Foundation (IPR transfer)