HEP Software Foundation Status

Pere Mato (CERN)
for the HSF Startup Team

HEP Software Foundation Workshop
LAL, Orsay, 2-3 May 2016

Outline

- Reminder of Objectives
- HSF Timeline
- Working Groups
  - Status & Achievements
- Fostering Collaboration
- Food for Thought
  - Proposals for discussion
- Conclusions

HSF 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 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 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

HSF Timeline

- Early 2014: HEP software collaboration proposed first collaboration meeting
- Spring/Summer 2014: gathering White Papers from the community.
- Fall 2014: Startup plan and team, White Paper Analysis, community discussions
- 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
- May 2016: LAL HSF Workshop

HSF Activities and Working Groups

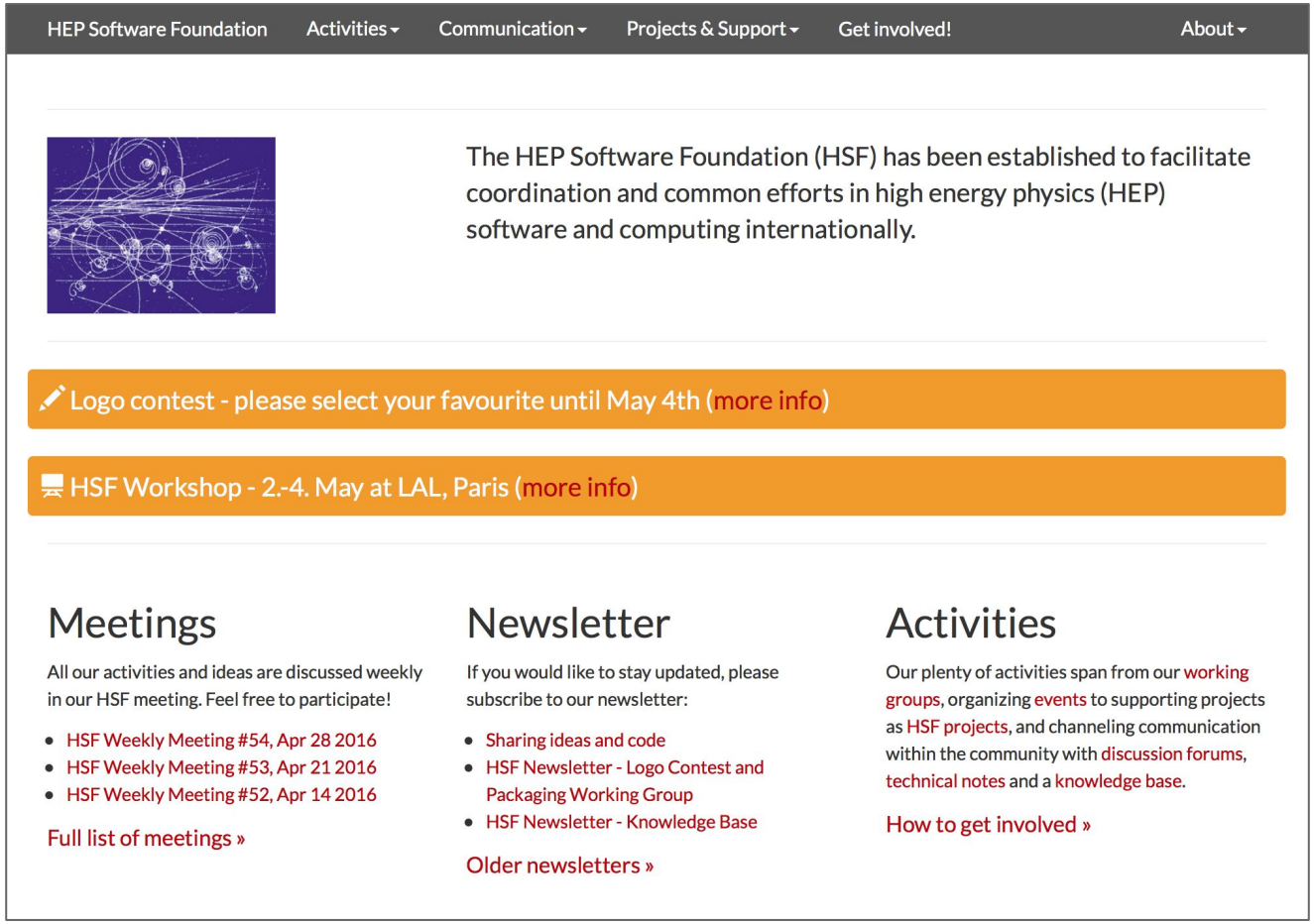
Working Group	Objectives	Forum - Mailing list
Communication and information exchange	Address communication issues and building the knowledge base Technical notes	hep-sf-tech-forum
Training	Organization of training and education, learning from similar initiatives	hep-sf-training-wg
Software Packaging	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

Communication and Information Exchange

- HSF Web
- Mailing Lists (Fora)
- Knowledge Base
- Technical Notes
- Newsletter

HSF Website

- http://hepsoftwarefoundation.org



Mailing Lists

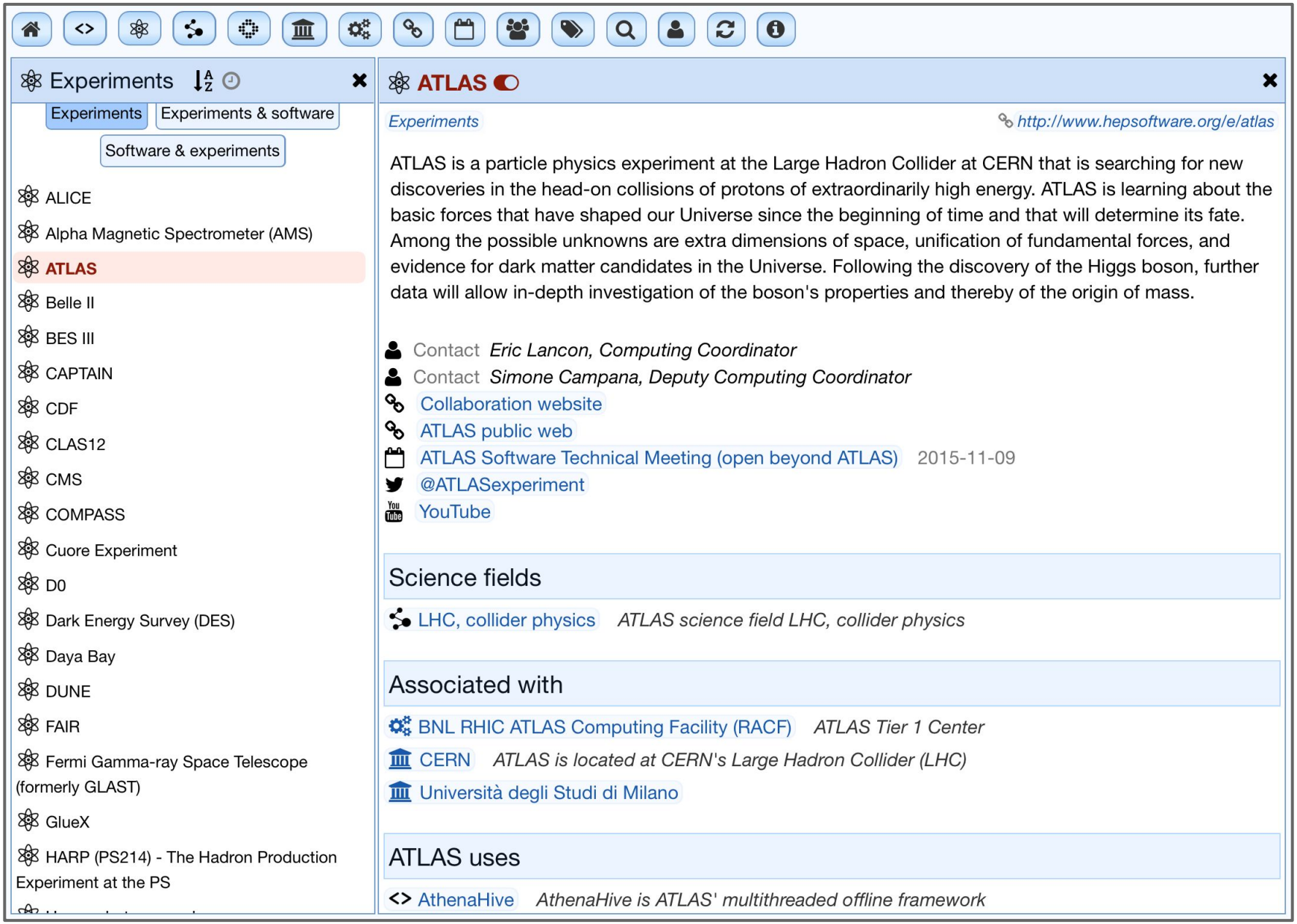
- HSF Forum
  - http://groups.google.com/d/forum/hep-sf-forum
  - 147 members
- HEP S&C community forum
  - http://groups.google.com/d/forum/hep-sw-comp
  - 337 people have signed up
  - General mailing list everybody in our field should subscribe to
- Other specialized lists
  - Training - hep-sf-training-wg
  - Packaging - hep-sf-packaging-wg
  - General HSF technical discussion forum - hep-sf-tech-forum
  - Ready to create more as needed
- Reminder: Google-free self-signup to lists
  - Simply send a mail (subject and content irrelevant, can be empty) to <listname>+subscribe@googlegroups.com, e.g. for the list above, hep-sw-comp+subscribe@googlegroups.com
- See the 'Get involved' page on the website for details

Community Topical Fora

- Software Technology Forum
  - This forum replaces the ***Concurrency Forum*** and widens its scope
  - 1 meeting/month
- Reconstruction Algorithms Forum
  - All matters of event reconstruction and pattern recognition software
  - 2 meetings so far, “Connecting the Dots” workshop
- Machine Learning Forum
  - The Inter-Experimental LHC Machine Learning Working Group IML launched recently
  - ML discussions and code development in the context of HEP, development of relevant tools, methodology and applications

Knowledge Base

- **Software catalog, software categories, science fields, community, and events**
  - implementation is a browser-based app (javascript client, node.js server, json in between, MySQL)
  - authentication is via github, google, facebook etc.
  - **emphasizes easy adding/editing of content, and extensibility. Adding content should be fun.**
- Available at http://hepsoftware.org
  - See Torre's presentation
  - Comments/feedback are welcome!
  - Just start filling it!
- *Implementation based on ATLAS DKB (data knowledge base) prototype*


Experiments
Experiments
Experiments & software
Software & experiments
ALICE
Alpha Magnetic Spectrometer (AMS)
ATLAS
Belle II
BES III
CAPTAIN
CDF
CLAS12
CMS
COMPASS
Cuore Experiment
D0
Dark Energy Survey (DES)
Daya Bay
DUNE
FAIR
Fermi Gamma-ray Space Telescope (formerly GLAST)
GlueX
HARP (PS214) - The Hadron Production Experiment at the PS
ATLAS
Experiments
http://www.hepsoftware.org/e/atlas
ATLAS is a particle physics experiment at the Large Hadron Collider at CERN that is searching for new discoveries in the head-on collisions of protons of extraordinarily high energy. ATLAS is learning about the basic forces that have shaped our Universe since the beginning of time and that will determine its fate. Among the possible unknowns are extra dimensions of space, unification of fundamental forces, and evidence for dark matter candidates in the Universe. Following the discovery of the Higgs boson, further data will allow in-depth investigation of the boson's properties and thereby of the origin of mass.
Contact Eric Lancon, Computing Coordinator
Contact Simone Campana, Deputy Computing Coordinator
Collaboration website
ATLAS public web
ATLAS Software Technical Meeting (open beyond ATLAS) 2015-11-09
@ATLASexperiment
YouTube
Science fields
LHC, collider physics ATLAS science field LHC, collider physics
Associated with
BNL RHIC ATLAS Computing Facility (RACF) ATLAS Tier 1 Center
CERN ATLAS is located at CERN's Large Hadron Collider (LHC)
Università degli Studi di Milano
ATLAS uses
AthenaHive AthenaHive is ATLAS' multithreaded offline framework

Technical Notes

- Technical Notes can be proposals, ideas, anything useful to the S&C community
- The first TN with the TN policy was published Sept 2015
- Other have followed: Licensing guidelines, Packaging tools, Machine/job features, Project recommendations, ...
- Repository and version control in GitHub

TN Number	Title	Authors	Download
HSF-TN-2016-03	HSF Packaging Working Group Report	B. Hegner, L. Sexton-Kennedy	PDF GitHub
HSF-TN-2016-02	Machine/Job Features	M. Alef et al.	PDF GitHub
HSF-TN-2016-01	Software Licence Agreements HSF Policy Guidelines	J. Harvey et al.	PDF GitHub
HSF-TN-2015-01	HSF Technical Notes policy	A. McNab	PDF GitHub

Training

- People having knowledge are rare
- People having time to actively share their knowledge even more rare
  - not a problem of motivation!
- So how to make best use of what is there?
  - First of all - make it visible!
  - Lower the bar for collaborative editing and re-use
- Visibility of training material
  - The KB is the place for advertising things
- Easier collaborative editing
  - Put material under a Creative Commons license
  - HSF invested into **WikiToLearn**

join the hep-sf-training-wg

WikiToLearn

- **WikiToLearn** is a wiki-based platform tailored at training and teaching
- Initiated in the context of italian universities
  - Basic idea was that students can improve and extend the material of their professors, while still being quality-controlled
- See whether HSF can take advantage of it
  - Started adding material to this site
- Investing in providing interactive tutorials
  - Combination of jupyter style notebooks and a privately owned sandbox - start tutorial now, resume later (collaboration w/ the ROOT team)
- WikiToLearn is only the shell, content has to come by the community (i.e. you!)

Software Packaging

- Topics:
  - package building, deployment, runtime environment, new technologies like Dockers, cmake best practices
- Organized a series of discussions/presentations on packaging and build tools (6 meetings)
  - Current practices inside and outside HEP
  - Document to summarize findings as TN
- Trying a hands-on approach to increase share of actual code even if existing experiments and projects locked-in to a certain packaging solution
  - Common “build recipes” protocol
- Tomorrow’s Spack Hackathon

join the hep-sf-packaging-wg

Build and Packaging Software Review

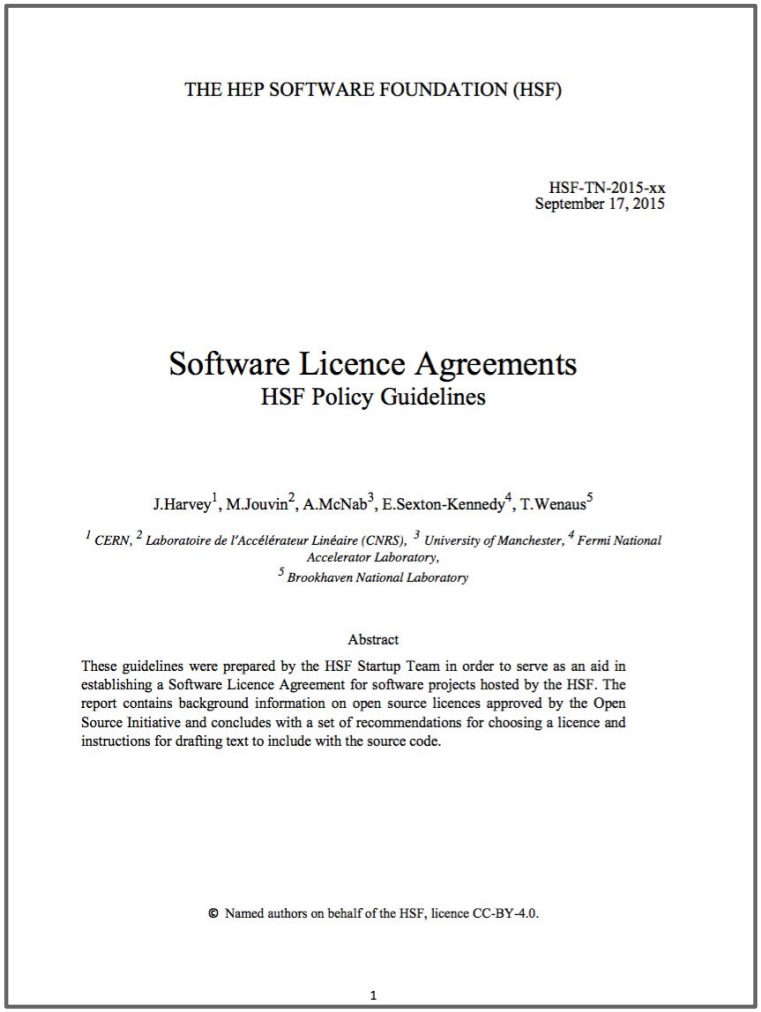
- Looked at many tools, in particular
  - worch, cmsBuild, aliBuild, LCGCMake, SciSoft, contractor (HEP), homebrew, Nix, conda, Spack
- Main problems in HEP software
  - reinvention of the wheel
  - non-share even within the community
- Main problems in non-HEP software
  - non HEP-tools prefer rolling releases, care less on reproducibility
  - little support for multi-environment setups
- Note HSF-TN/2016-03
- Tomorrow's Spack Hackathon

	Linux	MacOS X	Windows	Xcompiler
aliBuild	+	+	-	+
cmsBuild	+	+	-	+
Contractor	+	+	-	+
Homebrew	o	+	-	-
LCGCMake	+	+	o	o
Nix	+	+	o	o
SciSoft	+	+	-	-
Spack	+	+	-	+
Worch	+	+	-	o

join the hep-sf-packaging-wg

Software Licensing

- TN for the HSF Licensing Guidelines was finalized
- Background information on open source licences approved by the Open Source Initiative
- Set of recommendations for choosing a license and instructions for drafting text to include with the source code
- A second version is planned in the coming months based on **your feedback**



THE HEP SOFTWARE FOUNDATION (HSF)

HSF-TN-2015-xx
September 17, 2015

Software Licence Agreements
HSF Policy Guidelines

J.Harvey[1], M.Jouvin[2], A.McNab[3], E.Sexton-Kennedy[4], T.Wenaus[5]

[1] CERN, [2] Laboratoire de l'Accélérateur Linéaire (CNRS), [3] University of Manchester, [4] Fermi National Accelerator Laboratory,
[5] Brookhaven National Laboratory

Abstract

These guidelines were prepared by the HSF Startup Team in order to serve as an aid in establishing a Software Licence Agreement for software projects hosted by the HSF. The report contains background information on open source licences approved by the Open Source Initiative and concludes with a set of recommendations for choosing a licence and instructions for drafting text to include with the source code.

1

Software Projects

- The essence of the HSF are the **Software Projects** under its umbrella
  - HSF does not enforce any particular software process, project management or methodology, however packages should conform to some standards to facilitate integration

- Defined preliminary Project Guidelines
  - Project name, public repository, web site, issue tracker, version naming, mandatory documentation, best practices,...

Software Project Templates

- Discussed this morning
- The idea is to develop a **project template** implementing these guidelines and best practices
  - For example using the PODIO project (Toolkit for Event Data Models) as guinea pig developed under the AIDA 2020
  - A few more in the pipeline
- Prototype template available at https://github.com/HEP-SF/tools
- To support small projects that do not have a collaboration environment available
- To serve as example for shared projects across collaborations
  - reducing impedance mismatch

Fostering Collaboration

HSF may have the role of fostering and publicising common software development initiatives. Some examples:

- Next-generation conditions data
  - Belle II / ATLAS / CMS / LHCb
  - 2 meetings: Dec 10, Jan 21
- Track reconstruction
  - Huge challenges ahead
  - Should try to not only share ideas, but concrete code
- Frameworks: Gaudi
  - ATLAS / FCC / LHCb

Satellite projects in the HSF constellation. Examples:

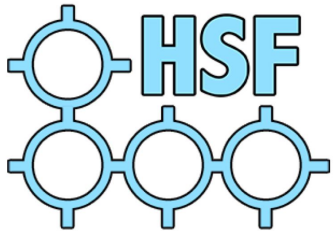
- DIANA (Data Intensive ANAlysis), 4-year NSF funded
  Focus on analysis software, including ROOT and its ecosystem
- AIDA-2020, EU funded, includes common software.

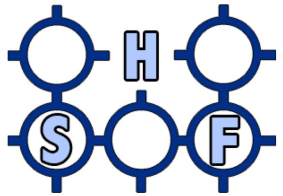
Food for Thought

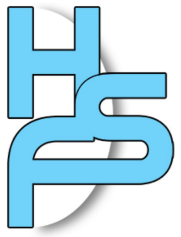
A number of proposals, ideas that we can discuss during the 3 days of the workshop and possible conclude at the last session

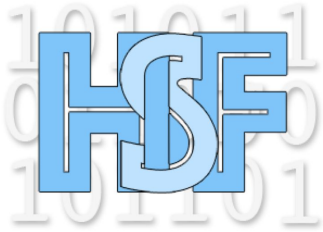
Logo Contest

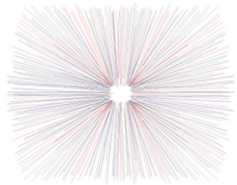
- 24 candidates (including variations on a theme)
- Voting ends May 3 (evening)
- Please vote your preferred choice! Will announce winner in the closing session (linked also on HSF home page)

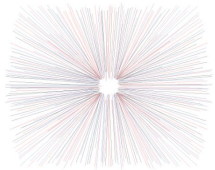




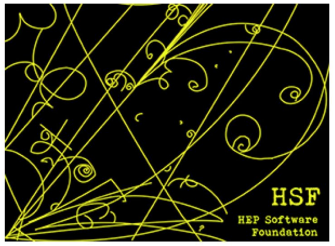








HSF
THE HEP
SOFTWARE
FOUNDATION



Project Reviews

- HSF was recently asked to organize and conduct a review of the GeantV project (to take place post-CHEP)
  - “Have been discussing who should review GeantV. Existence of HSF seems to make it the obvious choice, it is above the different parties and laboratories, can expect to get an objective view. In the GeantV project we want to start communicating in a more official way what we’re doing. Review will be an important step for the project.” - F. Carminati
  - State of the preparations so far: a doodle to set dates
- Project review is a foreseen HSF role from the beginning
  - Assemble experts & stakeholders to constitute an objective and representative review committee
  - Ensure the review is an open process
  - Review report and conclusions are produced by the committee
  - Communicate broadly the review itself and its outcomes
- If other projects could be interested, talk to us

HSF in StackExchange?

StackExchange: network of 150+ Q&A *communities* to find, ask and answer questions, focused on each community's needs (take a tour)

- Users may also comment, edit, up/down vote, earn reputation
- Private company, contents contributed under Creative Commons Attribution Share Alike license (should check with legal for HSF)
- Initially StackOverflow for software programmers, then added SuperUser, ServerFault, now over 150 sites (including sites for physics, statistics/ML, data science, computational science…)

We could launch a HEP S&C **hepswcomp.stackexchange.com** site

- First go through "Area 51" definition (agree on 40 hypothetical questions), then commit, finally private/public beta ("*10 questions per day is a healthy beta*") - see stats proposal progress history
- Which questions are not in other stackexchange sites already?
- Overlap with other HEP tools (roottalk forum,…) or FCE?

HSF Journal?

- Publishing HEP Software and Computing papers in peer reviewed journals?
  - We tend only to publish as conference proceedings
  - HSF Technical Notes are good, but perhaps not sufficient
  - Larger visibility across communities
  - Need to worry about career progression of our people
  - Authoritative, central archive and reference for science and technology transfer, preservation and documentation in the field
- Springer ready to collaborate with us
  - Possibility of a dedicated HEP SW&C journal (OA?)
  - Possibility to publish a book or book series
- Discussion on Wednesday afternoon

HSF Resources: Need for Effort

- Need resources. Dedicated people.
- Effort so far is the spare time of a tiny number of overcommitted people
- This is not serious if the HSF is to have a serious role
- Broad agreement that much work in (especially) software performance, efficiency, new algorithms & approaches to effectively using new architectures is essential
  - e.g. LHC computing resources scrutiny group report last week: *"Improving software efficiency is essential to meet growth in requests. CRSG strongly supports funding to continue this work."*
- **If it's an important activity that should be funded**, and it should be pursued in common where possible...
- Endorsement from ICFA?

HSF as a Legal Entity?

- We may call HSF a foundation but it isn't really, it has no such standing
- Should HSF be a legal entity?
  - Legal basis for protecting intellectual property through copyright
  - Attract HEP common projects to use HSF as their copyright holder & protector of intellectual property rights
  - Facilitate collaboration with industry, e.g. on the OpenLab model
  - Able to accept direct material support, or generate it through e.g. training course fees, support from industry, ...
  - A means of achieving critical mass as a real HEP community enterprise (it is not there today), e.g. through a representative Board?
- Should HSF *not* be a legal entity?
  - CERN and other labs play these roles, e.g. for ATLAS, CERN holds copyrights "for the benefit of the ATLAS Collaboration"
  - Enters HSF into a morass of legality, complexity, governance that drowns the HSF as a grass roots software enterprise?
- Can we make it a realistic goal with help of labs and real support?
  - We need do-ers in our do-ocracy engaged to make this happen
- Recall that the Apache Software Foundation is a legal entity (and still manages to be grass-roots)
- Discuss during the workshop and we will give it (constrained) discussion time Wed afternoon, see if we have a consensus and if so make a plan

Conclusions

- We have made progress in some areas but at a slower pace than anticipated
  - Please continue contributing with your ideas
  - And with your direct participation
- We welcome new ideas and proposals for HSF activities and roles, such as GeantV's approach to conduct a review
- Areas for which we would like to have more help
  - Engagement of more software projects, help further develop the project template and tool set
  - Fill the knowledge base, contribute training material
- Please join and contribute to any of the working groups, the startup team, discussion fora
  - subscribe to the fora to follow progress and contribute
- Important questions are open as to HSF's direction & future
  - Continue as a tiny unresourced hobby of a few? (was not the plan)
  - Find real resources and real effort? (is it the time for this?)
  - Make the HS**F** real? (practical, beneficial, timely?)
- We have much to discuss here and can draw conclusions on questions and set next directions on Wednesday afternoon