

# HEP Software Foundation White Paper Synthesis

Andrea Valassi (CERN IT-SDC)  
for the HSF Startup Team

HSF Workshop – SLAC, 20<sup>th</sup> January 2015



# The road to this workshop

- Jan 2014 – Proposal for a HEP Software Collaboration
- Apr 2014 – HEP Software Collaboration Workshop
  - Lively discussion, “Foundation” proposal, call for White Papers
  - *10 White Papers were received during the next two months*
- Oct 2014 – HSF Startup Team formed
  - Mandate: synthesize WPs, propose and implement a startup plan
- Jan 2015 – WP Analysis and Proposed Startup Plan v1.1
  - With minor updates over the earlier draft 1.0 circulated in Dec 2014
- This talk will highlight some inputs received, as summarised in this document – to stimulate further discussion in the WS
  - And avoid repeating past discussions: *we are not starting from scratch*

# The 10 White Papers

- Six papers from different geographical areas:
  - G. Quast et al [*German community input*]
  - J. Templon et al [*Nikhef input*]
  - M. Jouvin et al [*IN2P3 input*]
  - P. Spentzouris et al [*US input*]
  - D. Britton et al [*UK GridPP input*]
  - D. Menasce et al [*INFN input*]
- Four papers from different software domains:
  - O. Smirnova et al [*the Grid view*]
  - R. Mount [*personal input of ATLAS computing coordinator*]
  - M. Asai and M. Verderi [*Geant4 Collaboration input*]
  - A. Nowak [*the Openlab view*]
- A mixture of personal and institutional views

# White Paper synthesis document

HEP Software Foundation (HSF)

White Paper Analysis and Proposed Startup Plan

*The HSF Startup Team  
Version 1.1, January 7 2015*

For more information see [hepsoftwarefoundation.org](http://hepsoftwarefoundation.org)

## Table of Contents

[Revision history](#)

[Executive Summary](#)

[The White Papers](#)

[General motivations, goals and scope](#)

[Software focus areas](#)

[Technology challenges](#)

[Software process – Policies, guidelines](#)

[Software process - Common infrastructure and support teams](#)

[High-level coordination and support tasks](#)

[Communication and visibility within and outside HEP](#)

[Software developer skills and careers](#)

[Stakeholders and membership](#)

[Governance](#)

[Potential activities and deliverables](#)

[Proposed startup plan](#)

[References](#)

- 10 WPs from diverse range of teams
  - Different emphasis on different areas
  - Prepared first some notes on each paper, then a spreadsheet (matrix) of comment categories vs papers
  - Finally **summarized by categories**
- Synthesis document also includes the startup team subjective views and a proposed startup plan
- WPs and synthesis document are all available on the HSF web site

- The outline of the remainder of this talk will follow the document ToC
  - Report summaries/highlights of ideas and suggestions by categories

# Preliminary comments on the WPs

- *Good agreement on the general motivation and goals of HSF*
  - Different people focus on different issues – inputs are complementary
  - Agreement also on many specific suggestions, but not on every issue
  - We should aim to **prioritize** HSF goals/activities, understand/reconcile disagreements and identify the issues we have missed so far
- Organization (membership, governance) is the area where the widest range of opinions and suggestions was expressed
  - In the WPs as well as in the discussions during the first workshop
  - Within the startup team we find it premature to discuss this in detail
    - It is more useful to first agree/prioritize **what the HSF should be and do**
    - i.e. **what you need from HSF and what you are willing to bring to it**
  - This talk will not focus on the possible options for governance but will attempt to pinpoint in which areas it is more (or less) needed

# General motivations and goals

- *Evolve HEP software to optimize its performance*
  - Exploit new technologies, get ready for new experimental programmes
- *Promote common developments in HEP and with non HEP*
  - Avoid duplication of efforts, promote standard solutions
- *Facilitate compatibility, interoperability, integration testing*
  - Enable software sharing and reuse between different communities
- *Improve communication and expertise sharing on software*
  - Facilitate discussions within HEP and with non-HEP communities
- *Promote software development skills, careers, training*
  - Attract and retain software developers – and acknowledge their role
- *Incubate innovation & systematically foster new developments*
  - Provide a supportive environment without “management hostility”

# Software domains

- Wide agreement that HSF should initially focus on MC *simulation, reconstruction frameworks and data analysis*
  - Need to optimise them was one of initial motivations for launching HSF
  - Grid middleware and distributed software are welcome to participate proactively, but probably they would not be core effort at the beginning
- Focus on *software used by more than one experiments* or at least with a potential to evolve in that direction
  - General purpose toolkits like GEANT and ROOT are clearly in
  - Software specific to a single experiment is relevant as it has a potential to be reused by others and/or to be replaced by common software

# Technology challenges

- Exploiting *parallelism (multicore CPUs and vector processors)*
  - Widely recognized as the main current issue with HEP software
- Emergence of GPU accelerators, low power cores (e.g. ARM) and heterogeneous architectures: the end of x86 computing?
- Efficient access to large volumes of distributed data
- Exploiting diverse resources including (commercial) Cloud computing, HPC facilities and volunteer computing resources

And *the upcoming and yet unknown challenges of the future...*



# SW process – policies, guidelines

- Agreement on a few points – the HSF should recommend:
  - *Open-source licenses* – and *peer reviews* (stimulate “social coding”)
  - *Open standards and existing solutions* (do not “reinvent the wheel”)
- More diverse view points on other issues, e.g. guidelines for:
  - Release management and software dependency management
  - Modularity, component interfaces and data format/representation
  - QA, component tests, stress tests – also performance benchmarking
  - Documentation, tutorials
- Should policies and guidelines be recommended or enforced?
  - These issues are related to membership and governance, see later...
  - General view that HSF should “support” projects, not “manage” them

# SW process – infrastructure, support

- Some WPs say the HSF could offer a *common infrastructure*
  - Access to common build and test infrastructure
  - Access to computing platforms including emerging ones (ARM,GPU)
  - Access to common tools (compilers, profilers, optimizers)
  - Access to common collaborative tools (web site, mailing list, twiki...)
  - Access to common issue trackers and source repositories
    - But others explicitly suggest using GitHub
- Some WPs say the HSF could offer *common support teams*
  - Support from common integration and certification teams
  - Support for licensing and IP issues
  - Support for organizing peer reviews
  - Support from dedicated task forces (SWAT teams) on specific issues
- Note that a few such coordinated common efforts already exist
  - e.g. for LHC at CERN via WLCG AA, PH-SFT, IT Openlab/Techlab...

# High-level coordination and support

- Some WPs say HSF should do some *high-level coordination*:
  - Actively promote collaborations between projects and existing solutions, and discourage fragmentation while encouraging diversity
  - Propose a general roadmap and risk analysis for HEP software
  - To these ends, maintain an up-to-date repository of projects
  - Incubate (and provide a “systematic process” for) innovation
  - Aim for sustainability, facilitate maintenance and lifecycle management
  - Provide support for and promote collaboration on funding proposals
- Some of these goals are “easy”, others are more ambitious
  - The HSF startup team has already started a repository of projects
  - But some other goals imply *choices* that require an agreed governance

# Communication within/outside HEP

- Widespread agreement that HSF should take an active role in several communication channels
- Communication within HEP
  - Set up forums to allow expertise sharing between developers
  - Set up forums/events for project/developer communication with users
  - Set up forums for communication between users on common needs
- Communication between HEP and non-HEP partners
  - Promote contacts and workshops with non-HEP partners
  - Promote the visibility of HEP software outside HEP (newsletter, outreach activities, HSF as entry point, public face of HSF web site...)

# SW developer skills and careers

- Widespread agreement that HSF should take an active role in promoting software developer skills and careers
- Some specific suggestions include in particular:
  - Provide training opportunities for developers (e.g. CSC, Bertinoro)
  - Provide career support for developers
  - Increase the visibility and recognize the value of developers for HEP
    - e.g. help define career paths comparable to research or detector work

# Organization – membership?

- “Membership” is generally seen as applying to projects
  - But HSF will be made first and foremost by motivated individuals
  - And there are many more stakeholders (see Richard Mount’s WP)
- Conditions for project membership?
  - Bottom-up initiative, membership is voluntary (buy-in)
  - Formal acceptance by HSF subject to policy acceptance/compliance?
- Benefits for project membership?
  - Visibility and recognition – and diversity is seen as a value
  - Better access to knowledge
  - Concrete support and access to common infrastructure
- Nice suggestion to separate *“hosted” and “endorsed” projects!*

# Organization – governance?

- Most WPs suggest a *non-prescriptive, lightweight organization*
  - With one or more Boards (one possibly with non-HEP members)
  - But at least one WP suggests a relatively prescriptive governance
  - GEANT4 reminds us of the importance of a *Collaboration agreement*
- A better definition of governance will be required by the time high-level coordination tasks and related choices are needed
  - Personal opinion: participating to HSF (i.e. to its common benefits) may imply not only contributing resources but also giving up some “freedom”
  - Still anyone contributing resources wants to see a return on investment
  - Not only elements of software “Darwinism” but also of “social contract”?
- Later in the year (CHEP?) may be the right time to discuss governance, after making progress on some concrete services
  - In the meantime, let’s hear what other Foundations have to say!

# A palette of possible activities

- In summary, many concrete activities/services were mentioned in the WPs:
  - Project hosting infrastructure
  - Building and testing infrastructure
  - Teams for certification and integration
  - Software repositories and package managers
  - Access to computing resources on many platforms and architectures
  - Access to software development tools
  - Training in software technologies and tools
  - Support for IP and licensing issues
  - Peer reviews
  - Access to scientific software journals
  - Task forces or “SWAT” teams to solve specific issues
  - Consultancy for new experiments or projects
- *“The initial portfolio of services will reflect the needs of the stakeholders that participate in the Foundation and the resources that it can attract”...*
- *“The HSF is what people bring to it”!*



# Concluding remarks: to start discussion!

- Many of the activities I described are already taking place
  - And in some cases already benefit from coordinated common efforts
- Appearance of HSF should be a smooth transition initially
  - What would the HSF allow you to change and do differently/better?
  - What can you afford to do differently to put things (more) in common?
- What HSF should be and do is a cost/benefit tradeoff
  - What would you like from it? What are you willing to bring to it?
    - “you” being all stakeholders (projects, but also users, funding agencies...)
  - And what are the appropriate and realistic timescales for all of this?