View of the German Community on the proposed "HEP Software Foundation"

Editors of this document:

Christian Zeitnitz (Wuppertal, ATLAS), Thomas Kuhr (Karlsruhe, Belle II), Günter Quast (Karlsruhe, CMS), Matthias Kasemann (DESY Hamburg, CMS), Michael Schmelling (MPI Heidelberg, LHCb), Patrick Fuhrmann (DESY Hamburg, dCache)

Preamble

Scientific knowledge, in particular in High Energy Physics, but more and more in other research fields, is obtained from the analysis of huge data samples. The whole analysis process, from data taking to the extraction of results and from theoretical models to simulation of the detector response, relies on high-quality and performant software. The goal of this initiative is to organize ourselves, i. e. the people writing HEP software in different organisational structures ranging from large institutions to small university groups, to optimize the software development process and to provide better software products, aiming at maximized physics output.

Of course resources, in terms of people and funding, are limited. Problems to obtain sufficient funding should be addressed by enhancing the recognition of software activities. The optimization of the usage of available resources requires coordination among ourselves to exploit synergy and enhance the skills of each developer by knowledge transfer and training.

As the software environment - scientific projects, people, technology – is constantly changing, it requires continuous development and maintenance efforts. In particular software required for High Energy Physics experiments is usually written by physicists and not by professional programmers. Very prominent examples are ROOT, GEANT and event generators for different physics processes. Most of this software is currently running as monolithic, single thread programs, leading to severe performance losses on modern hardware, which is characterised by vector units, instruction pipelining supporting multiple threads, and multi- and many-core CPUs.

In order to fully utilize the capabilities of modern hardware, the current software has to be, at least partially, rewritten, and new software should be optimized for performance from the beginning. This requires, first of all, awareness and special knowledge and skills of the programmers.

A "HEP Software Foundation" has the potential to become a supporting structure for the developing teams of existing software projects, to stimulate new projects, and to attract and train new developers, thus enabling software development in our field to meet the technological challenges ahead.

Tasks of the Foundation

The foundation should provide support for HEP software projects on different levels:

- provide (or recommend) a technical hosting platform (similar to github or utilize hepforge for this purpose)
- provide clear guidelines for software projects of the foundation
    - clear licence rules (best to use one of the existing open source models)
    - propose and monitor quality standards
        - code management system and clear release management
        - performance measurements and benchmarking
        - organize peer reviews
    - compatibility among projects
    - documentation and tutorials
- provide a development and testing framework
    - access to professional development tools
      (like memory checkers, code optimizers, …)
    - run or provide access to different platforms
      (standard and new CPUs, GPUs, ARM, …)
    - provide an automatic environment to compile and run nightly builds and allow performance benchmarking on different platforms
- organize training and workshops or developer forums for (new) members of the foundation and the global community
    - training material and lecturers for the major schools on computing (such as CERN School of Computing, GridKA-school, national school on community computing, introductory C++ schools etc.)
    - introduction of coding recommendation and training on tools to realize such standards (debugging, profiling, memory leak search etc.)
    - eventually create a new, dedicated school on HEP software development
    - organization of topical workshops on new paradigms and trends in contemporary computing, stimulating contacts and exchange of expertise within and outside of the foundation, in particular involving computer scientists
- run an organizational structure to allow coordination of the supported projects
    - collect expertise to have an overview of technical developments and existing or planned HEP software projects
    - make recommendations and develop a general roadmap
    - propose possible collaborations to prevent duplication of work
    - support building collaborations on software projects and proposals to apply for funding
- establish communication with the communities of (potential) users
    - run an information platform providing an overview of supported projects
    - provide material and lecturers for schools on software organized by different communities (e.g. large and small HEP collaborations, astro particle physics, nuclear physics, …)
    - On request by users or projects: review projects and consult potential user communities
- improve the visibility of the HEP software community, of projects and individual developers
    - stimulate invitations of members of the foundation to major conferences
    - edit a (wide-spread) newsletter, eventually advocate and support an open-access journal

The structure of the Foundation

From the discussion at the initial meeting it is evident that a foundation or collaboration only works on a voluntary basis. There is no strong motivation - like the right to access the data - that makes people accept an organization like in HEP experiment collaborations.

As this initiative will only be successful if there is a sufficient number of supporters, it has to provide a clear benefit for its members. This is not easy, given, on one hand, the diverse structure of the HEP software community, ranging from large centres with strong IT divisions and full-time professionals to very small groups at universities with part-time developers, post-docs and students. On the other hand, there is a wide range of HEP-related software projects, including small and large-scale general (HEP) purpose and experiment-specific projects of small and large experiments. Ideally all groups with their different backgrounds and priorities should be able to identify themselves with the goals and structures of the foundation.

Currently the HEP software developments are usually organized in project teams (including the experiment software development teams). It seems natural that these teams will be the stakeholders of the foundation. Thus it is proposed to have a board of project representatives as the main organizational body of the foundation. Details of its composition and eventual weights of votes of different projects remain to be discussed. The foundation board would elect a chair person. As the HEP software foundation is built around a common set of guidelines regarding scope and quality of the associated software projects, the foundation board would decide about the admission of new projects and its members. The chairperson can appoint an executive team which will be approved by the board.

The tasks of the chair of the software foundation and of the executive team are

- assign reviewers on request by the projects or the project board
- ensure a solid and timely review process
- make recommendations of reviewers publicly available
- eventually appoint task forces for specific issues
- support organization of training and education
- support preparation of funding applications
- ensure that there is at least one healthy project for the technical services, for developing a roadmap for the HEP software development, and for reviewing technological developments
- establish and maintain communication channels with other communities and industry
- report to the project board, openly accessible to the whole community of the foundation
- ...

This proposal emphasizes the support that the executive team provides to the software projects in contrast to a direct management of the member projects. The influence on the HEP software activities is indirect via membership in the foundation, peer reviews, the roadmap, and recommendations.

Of course the foundation generates some overhead, in particular for the members of the executive team, but also for the other developers, e.g. by meetings or reviews. The benefits, like usage of common tools, better access to knowledge, and increased visibility, have to outweigh the drawbacks. New, small projects will probably benefit more than existing, big ones, but also for them the indirect support of smaller projects can add to their value so that they decide to join the foundation.