

Proposal for an HEP Software Collaboration

23rd January 2014

This proposal summarises the High Energy Physics community's motivations for creating a Software Collaboration for building and maintaining advanced scientific libraries and tools of general interest, as well as facilitating engagement with other scientific and industrial partners. It represents a collection of ideas that can serve as a useful starting point for informed discussion.

Motivations

- HEP adopted C++ 20 years ago. Much of our software is now old; it needs to be adapted to more modern standards (C++11) and to use up-to-date standard libraries.
- Moreover the field of software development is currently undergoing a paradigm-shift resulting from the evolution of CPU architectures towards the use of parallelism (vector units, many-core etc); our code has to be re-engineered to make use of the full capabilities in order to deliver the full performance offered.
- Our software needs to be adapted to make use of all resources available to our community, including HPC facilities, commercial cloud computing facilities, as well as volunteer computing resources.
- To address these issues we must attract people with the required advanced skills and experience
- A new approach is needed to make sure our software is interoperable with software developed by other scientific communities
- We must provide more opportunities for sharing software between different experimental programmes conducted in different laboratories and research institutes.

Proposal

To establish a collaboration with a view to developing open scientific software packages that are guaranteed to work together; this includes frameworks that assist in assembling full applications in the various software domains (data processing, validation, simulation, analysis etc). Important features of this proposal are:

- We can build on the work already done in the Concurrency Forum that has been very successful in bringing the HEP community together; to date work has focused on R&D studies (building demonstrators) to show various capabilities of emerging technologies and by having a forum within which to disseminate the knowledge acquired.
- A more formal HEP Software Collaboration would provide a framework for increasing the level of collaboration and involvement amongst

colleagues with experience in concurrent programming; this would provide opportunities to attract experts in our HEP institutes and in other scientific communities.

- Having a more formal collaboration would bring the means to provide recognition and give credit to contributors to the effort; experience shows that that recognition/credit is important in bringing potential collaborators into a common effort.
- This collaboration would work together to elaborate a coherent set of proposals for acquiring resources for the work programme from our funding agencies and external funding programmes, such as those offered by the EU (H2020) and NSF/DOE. It would also be a vehicle for engagement with industry, for example through mechanisms such as the CERN Openlab.
- It would provide roadmaps and priorities for investigation and development; having such roadmaps would allow potential collaborators to see in which areas they can contribute or where their expertise may be relevant.
- Another role would be to establish a team of experts who are able to provide consultancy and practical help for experiments to develop their code, in particular in the complex tasks of optimising performance, debugging concurrency issues etc.
- This initiative is predicated on the assumption that there is a high degree of commonality in the core software services of the experiments i.e. in the use of libraries, frameworks, tools.

Work Packages

The collaboration would organise its activities in Work Packages to be defined by the interested collaborators. Some ideas include:

- R&D studies conducted over short time-scales to aid in the choice of both hardware and software technologies;
- development projects to re-engineer existing libraries and toolkits;
- development of new common software components that are of general applicability ;
- establishment of an infrastructure that could provide essential resources for the development and benchmarking of 'parallel software' on a representative range of candidate computing architectures; it includes emerging hardware e.g. Xeon/Phi, AMD, NVidia, ARM, as well as giving access to development and analysis tools (compilers, optimisers, profilers, etc) ;
- provision of the essential development infrastructure in terms of tools and processes (repositories, build and test systems etc.)
- creation of task forces dedicated to addressing concrete needs as they arise e.g. to help migrate an experiment to use a new technology or tool.

Governance

The proposed collaboration should have a relatively light governance model encapsulated in a collaboration agreement. For example:

- A spokesperson or chair ;
- A steering committee – representing the major contributors;
- Technical coordination – responsible for producing roadmaps, workplans, and creating task forces;
- Development effort organised in project teams associated to each WP