A Perspective on the SPPEXA Collaboration from France

Nahid Emad

As the French member of the Steering Committee of SPPEXA, it is my great pleasure to give a short address to this volume from the perspective of the French partners in this German-French-Japanese cooperation. To highlight the types of software supported by SPPEXA, we first present a classification of high-performance software types. We then take a look at the recent activities of HPC software in France under the SPPEXA umbrella. Next, some local impacts of the SPPEXA collaboration on the French HPC community is provided, and lastly, an outlook to future collaborations.

1 HPC Software in Three Phases

High-performance numerical software targets at obtaining relevant scalability in space and time for large-size applications by using a large number of cores/processors/nodes of powerful computers. They can be classified into three phases: pre-treatment, treatment, and post-treatment. Obviously, such a software often belongs to more than one of the categories mentioned.

Pre-processing Software The precise definition of these phases depends on the context, but the main role of pre-treatment software is the preparation of the input data for the treatment phase. This preparation sometimes consists of a rather complex parallel algorithmic and programming processing. Big data compression, uniform data formatting, and conditioning improvement of data matrices are some examples of the pre-treatment phase.

N. Emad (✉)
University of Paris Saclay, UVSQ, LI PARAD, Maison de la Simulation,
Versailles, France
e-mail: nahid.emad@uvsq.fr

Treatment Software The HPC software in the processing phase concerns mainly high-performance numerical simulation of physical phenomena, social networks, etc. These softwares could be classified as (1) "standard" libraries and (2) ad-hoc libraries done by application field programmers, which implement their application by making use of building blocks (partially or totally) of the libraries of class (1). The latter implement numerical methods with the main parallel/distributed programming methodologies, such as ScaLAPACK, PETSc, SLEPc, ATLAS, etc. In the fields of application , HPC ad-hoc software targets epidemiology, electromagnetism, gamma astronomy, safety, or health and nutrition. Some ad-hoc software examples are CEDRE which targets simulation for energy and propulsion, CELESTIA and STELLARIUM which is a space simulator and a planetarium for observing the solar system and the rest of the universe in real time and in 3D, GAUSS, which is a flexible platform for data analysis, and AREPO, which is a cosmological hydrodynamical simulation code on a dynamic unstructured mesh.

Post-processing Software Post-processing software essentially involves the analysis, visualization, and performance evaluation of the treatment phase results. Some examples of such software are ParaView (a multi-platform data analysis and visualization application), VisIt (an interactive platform for visualization, animation and data analysis), MAQAO (sets of software tools for code optimization in the core or node level of a parallel architecture), and Maya (a software for modeling, simulation, and 3D animation).

All these software packages generally translate a physical phenomenon, social behaviour, etc. into mathematical equations. Their high-performance implementation on parallel and/or distributed systems is a delicate task and requires a huge ecosystem with people having interdisciplinary skills. This makes the existence and use of accompanying software necessary, which provides the logistics of high-performance computing. These software frameworks provide the environment for high-performance programming and often conceal the complexity of underlying parallel and/or distributed architectures. As a consequence this allows the users to focus on main objectives. MPI, OpenMP, Globus, Condor, XMP, YML, MUST, etc. are very few examples of this kind of software.

2   Trilateral Projects in SPPEXA and Their Impact

SPPEXA targeted fundamental research on different aspects of HPC software and covered software categories cited before with a co-design approach. Thanks to ANR, DFG, and JST, the trilateral French/German/Japanese projects have been funded within SPPEXA. Some of these projects, such as MYX, have benefited from pre-existing bilateral collaborations. This allowed a dynamical and productive work from the beginning and for a rapid progress towards the objectives set. In addition to project meetings, the cross-project SPPEXA workshops have given a new dynamic to the trilateral collaborations paving the way for the organization of other conferences, workshops or seminars.

The EXAMAG (Exascale Simulations of the Magnetic Universe) project is an example of an SPPEXA trilateral French/German/Japanese project with the aim of improving the astrophysical moving-mesh code AREPO and extending its range of applicability for high scale computing platforms. EXAMAG is an ad-hoc HPC software of the processing category of the classification given in the previous section.

MYX (MUST Correctness Checking for YML and XMP Programs) also is a trilateral French/German/Japanese project which aims to offer a guideline how to limit the risk to introduce errors and how to best express the parallelism to catch errors at runtime. From a practical viewpoint, MYX aims at the design and the application of a scalable correctness checking tool MUST to YML and XMP. In the MYX project, the main developed software packages (YML, XMP and MUST) belong to the last category of HPC software; the ones providing the logistics of high-performance application programs. However, in order to validate the design and development of these softwares, many other benchmarks and/or real applications are developed. Among them are the multiple restarted Krylov methods/HPCS ad-hoc, matrix generator/pretreatment , epidemic HPCS ad-hoc, etc.

The SPPEXA funding of workshops with several projects involved added an extra dimension of interdisciplinarity. In collaboration with the DASH and ESSEX-II projects, MYX members organized four trilateral (German/Japanese/French) workshops. Two of them have been hold at university of Paris Saclay/Versailles in France. With the prominent invited speakers and the talks of SPPEXA-involved project members, these workshops have been very attractive (40 and 60 attendees, respectively). An important number of indirect outcomes of SPPEXA activities (workshops, "open" trilateral meetings, doctoral retreats, etc.) generated new connections between German and French colleagues and students. A few examples are ▷ the review of the PhD dissertation of a non-SPPEXA funded French student by Sabine Roller, professor at Siegen University, Germany, ▷ Xinzhe Wu, who finished recently his PhD, funded by ANR part of ANR/DFG/JST MYX project and, who is currently in a post-doctoral position at *Jülich Research Centre* in Germany, ▷ M.A. Diop, currently PhD student, funded by a French CIFRE followship (with ATOS/EVIDIAN company), who participated in SPPEXA doctoral retreats as well as several SPPEXA workshops, ▷ a workhop organised by Sabine Roller and Nahid Emad at HPC Asia, or a large number of BSc and MSc students benefiting from the collaboration.

3   What Will Be Next?

The on-going convergence between machine learning, data analysis, and high-performance computing is creating new algorithmic and co-design approaches that need to be taken into account for the future. With the three tri-national workshops in Tokyo, SPPEXA has contributed to this on-going development, and we are all looking forward to a continued collaboration in the future.