Training Experiences by Skills for HPC Ecosystems

Carlos J. BARRIOS H.
Universidad Industrial de Santander (UIS)
Bucaramanga, Colombia
0000-0002-3227-8651

Robinson RIVAS
Universidad Central de Venezuela (UCV)
Caracas, Venezuela
0000-0001-7900-0764

Xavier BESSERON
University of Luxembourg
Esch-sur-Alzette, Luxembourg
0000-0002-7783-2185

Gilberto J. DIAZ
Universidad Industrial de Santander (UIS)
Bucaramanga, Colombia
0000-0001-8188-5784

Yiannis GEORGIOU
Ryax Technologies
La Coursive d'Entreprise
69190, Saint-Fons, France
0000-0003-1264-7234

Pedro VELHO
Ryax Technologies
La Coursive d'Entreprise
69190, Saint-Fons, France
0000-0002-9767-468X

***Abstract*—High-Performance Computing (HPC) is one of the pillars of developing modern science and disruptive technologies, uniting computer architectures and parallel programming into multidisciplinary interactions to face domain-specific problems. That is why different areas of knowledge require their future professionals (scientists or not) to acquire skills in using HPC. The Super Computing and Distributed Systems Camping School, SC-Camp, is a non-profit activity that proposes a series of courses about HPC with an important focus on practical sessions (more than half of the time) addressed to undergraduate and graduate students that could benefit from HPC by demand. It is an itinerant school, bringing the HPC knowledge to a different place every year, with an important focus on diversity, sustainability, and humanity. This paper shows our history, primary goals, organization, and results over a decade of work in different countries.**

***Index Terms*—Education, HPC, Training, Outreach**

I. Introduction

Supercomputing is one of the pillars of developing modern science and supporting technology development and knowledge. Skills in HPC are important for various professional fields and much sometimes crucial in scientific and technology areas, building complete ecosystems. Several initiatives as the HPC Certification Forum [2] and the Special Interest Group in High-Performance Computing (SIG HPC) [1] Education of the ACM present an organization of the skills necessary in HPC by goals and roles. The two visions, built from discussions between the specialists and users, propose three main actors: HPC systems engineers, HPC software engineers, and computer scientists.

These actors are easily related to the general computer science players: HPC system engineers integrate, monitor, tune, optimize, and evaluate the performance of each element and process of the HPC hardware and software environment. HPC software engineers develop parallel programs, debug and analyze the applications' performance, and optimize the HPC software cycle. For the latter, computer scientists can be associated with users. But in the case of HPC, they need specific skills: coding, running, and deploying parallel programs efficiently using HPC Environments. These three actors are in continuous interactions and must share common knowledge.

Besides the formal courses in different universities around the world, the HPC community has been proposing for years summer schools, seminars, tutorials, workshops, and certifications. In the scientific community's interest, the specialists (among them the engineers) and the industry were involved in developing activities on computer architectures, applications, methodologies, and programming paradigms. The deployment of different specific schools around the world allows for democratizing knowledge and the creation of previously non-existent collaborations in multicultural, multidisciplinary, and inclusive ways.

II. The Advanced Computing Context

Complex scientific computational projects from both theoretical and practical angles coincide with using the term *Advanced Computing* [3]. This involves hardware and software technologies and methodologies for using supercomputers to support large-scale data processing. Indeed, advanced computing projects need large computing capacities in storage and processing. Additionally, to guarantee the achievement of good results in predictable times, they require specialized execution and deployment and efficient exploitation of computing resources. These elements constitute the stakeholders' competencies in these projects: the HPC infrastructure operators, software developers, and HPC scientists.

We introduce three actors involved in advanced computing projects: HPC technology and operations engineers (HPCTOE), HPC software developers (HPCRSE), and HPC scientists or users (HPCUSE). All of them are in continuous interactions to achieve the goals and ensure the best deployment and performance of the applications of the advanced computing projects. Figure 1 shows the different elements impacted by each one of the roles in the HPC environment in levels, from infrastructure to applications. From those roles, the responsibilities and actions of each actor are visible, allowing the identification of the involved knowledge. Elements related to the support of application execution and development require skills to deploy environments or development services.

III. Integrate HPC Ecosystem Actors Skills

We showed how all actors of the HPC ecosystem have common skills, and there is a continuous and necessary interaction between them to improve their own capabilities and the performance and impact of their work.

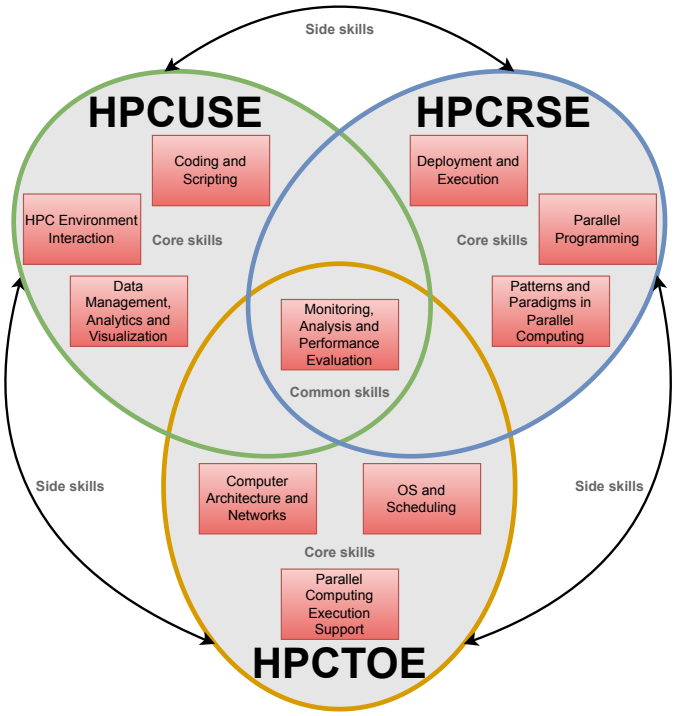

Fig. 1. HPC Technical Skills and Spaces of Development

Observing the span of technical skills, we identify common spaces of training and development. Figure 1 shows three main integrated spaces with common skills that can guide the development of courses or practical sessions of HPC schools with the participation of all involved actors of the HPC ecosystem. The acronym **HPCTOE** refers to the HPC Technical and Operation Engineers, **HPCRSE** is to the HPC Research Software Engineers (or Developers), and **HPCUSE** for the HPC Scientific Users or simply users.

Considering the skill set listed in the previous section, common knowledge, core skills, and side aptitudes are identified for each role; the Figure 1 highlights how the three roles of HPCUSE, HPCRSE and HPCTOE share some common skills, e.g., monitoring, analysis, and performance evaluation[1] but also a set of specific skills or core knowledge for the role. All the technical skills are shared between the different actors is the sense that the core skills of one role are actually side skills for the other roles. Side skills show competencies that are not required to perform the main tasks of the roles but for which basic knowledge is desirable to understand the system's complexity properly. It points out clearly how much the competencies of the three types of actors overlap and why it makes sense to train them in interaction with each other.

[1]This skill includes debugging, profiling, benchmarking, and testing of applications or the HPC environment, including platform and infrastructure.

IV. Discussion

In this contribution, we presented how we identified different roles in an advanced computing ecosystem and described how common and specific skills are distributed among the different actors of HPC. The interaction between them and common actions following good practices and communication are important to perform better HPC development, seek to build better and sustainable applications, optimize resource use, and achieve big impact in the community.

Following these aspects and strongly aimed at democratizing knowledge, we develop the Supercomputing and Distributed System Camping School, SC-Camp, to create common teaching spaces and leverage interaction, enhance technical and scientific multidisciplinary skills, and promote the appropriation of knowledge. The most notable scope of the SC-Camp has been the increase in collaboration between actors, even from different places and knowledge areas. Equally important is the impulse in the appropriation of knowledge by the participants. In practice, the follow-up of the alumni was not developed in this paper (and is difficult to organize). However, from their own initiative, the students have created groups in different social networks and continue to interact (i.e., it is possible to follow the hashtag #SCCAMP or #SCCAMPStudents).

The COVID-19 pandemic forced new ways to implement SC-Camp. We presented how the SC-Camp-V was performed during the 2020 and 2021 years, contributing to the digital acceleration and maintaining the initial spirit and objectives of the camping school. After consideration of issues related to the economic and environmental impact and the inclusion and participation of remote students using the Internet, the SC-Camp committee proposes to organize test versions of a Hybrid SC-Camp (called SC-Camp-H) for the next years. For 2022, we will experiment with in-situ training in a few selected sites for 1 day for 6 weeks, but with the possibility to participate online remotely. In 2023, we return to face-to-face activity for the most part, but with a previous remote component. Of course, the hybrid mode comes with additional challenges, for example, how to get the online and physical participants to interact together, which are issues which need to be addressed as we move forward.

More information about SC-Camp in www.sc-camp.org.

References

[1] ACM SIGHPC Education. URL: https://sighpceducation.hosting.acm.org/.
[2] HPC Certification Forum. URL: https://www.hpc-certification.org/.
[3] National Research Council. Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020. doi:10.17226/18972.