Towards an HPC Certification Program

Julian Kunkel
University of Reading
Reading, United Kingdom
j.m.kunkel@reading.ac.uk

Kai Himstedt
Nathanael Hübbe
Hinnerk Stüben
Sandra Schröder
Michael Kuhn
Matthias Riebisch
Stephan Olbrich
Thomas Ludwig
Universität Hamburg
Hamburg, Germany

Weronika Filinger
EPCC, The University of Edinburgh
Edinburgh, United Kingdom

Jean-Thomas Acquaviva
DDN
Paris, France

Anja Gerbes
Goethe-Universität
Frankfurt am Main, Germany

Lev Lafayette
University of Melbourne
Melburne, Australia

ABSTRACT

The HPC community has always considered the training of new and existing HPC practitioners to be of high importance to its growth. This diversification of HPC practitioners challenges the traditional training approaches, which are not able to satisfy the specific needs of users, often coming from non-traditionally HPC disciplines, and only interested in learning a particular set of competences. Challenges for HPC centres are to identify and overcome the gaps in users' knowledge, while users struggle to identify relevant skills.

We have developed a first version of an HPC certification program that would clearly categorize, define, and examine competences. Making clear what skills are required of or recommended for a competent HPC user would benefit both the HPC service providers and practitioners. Moreover, it would allow centres to bundle together skills that are most beneficial for specific user roles and scientific domains. From the perspective of content providers, existing training material can be mapped to competences allowing users to quickly identify and learn the skills they require. Finally, the certificates recognized by the whole HPC community simplify inter-comparison of independently offered courses and provide additional incentive for participation.

1 INTRODUCTION

There is a generally accepted set of skills and competencies necessary to efficiently use HPC resources. This skill set depends on the role and domain of the practitioner but also on the available infrastructure of the center providing the computing resources. For example, a scientist needing to run an application on a specific machine may need basic skills in Linux, MPI, environment modules, and knowledge about the batch scheduler, e.g., SLURM. Understanding SLURM is a good example of a very fine-grained skill, indeed we can identify "resource management" as a generic skill that illustrates concepts across the rich variety of available resource managers. Institutions which operate HPC systems typically offer regularly teaching events about general aspects of their supercomputer's hard- and software architecture and about the software environment around parallel programming and optimization. The learning material provided by an HPC center, however, is geared to the special demands of the institutions they support and its specific HPC environment. This content typically covers a small part of basic HPC skills which are necessary to use other HPC systems. Moreover, they do not attest the users to have a certain competence. Certificates are widely used in industry to attest certain knowledge but, so far, there is no similar approach for HPC training.

This lightning talk describes the current status of the certification program and the effort to create an independent body that curate the competences and issue certificates for the users.

The foundation for this work was laid with the Performance Conscious HPC (PeCoH) project, which among other goals, aims to establish an HPC certification program. A white paper about the approach containing technical details is provided in [1]. Within the project it become quickly apparent that the local effort is useful for the wider HPC community and could be extended to a global effort.

2 RELATED WORK

Relevant work can be classified into approaches to establish a curriculum or the creation of teaching material. In academia, individual universities offer their own curriculum around scientific computing and HPC, covering theoretical aspects like the software development of numerical applications. They are not tailored to the needs of a practitioner to actually use HPC systems effectively. Data centers offer their own material and courses to support their own users. Several projects address the generation and sharing of teaching material for HPC. The EuroLab-4-HPC project establishes training
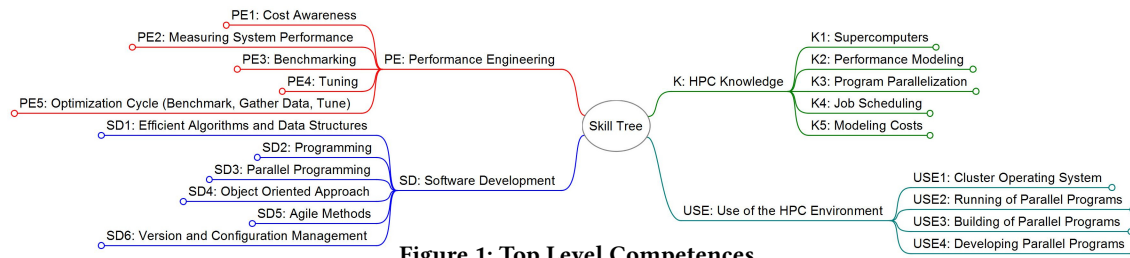
Figure 1: Top Level Competences

in form of (online) courses[1]. The Barcelona Supercomputing Centre (BSC) aims to develop a professional training curriculum [2]. The virtual organization XSEDE[2] provides an online system to train the usage of an HPC system, structuring the corresponding information on their website into major topics like "Getting Started". The user can navigate the topics and receive further information.

3 CERTIFICATION PROGRAM

The certification program serves two purposes: 1) the definition and organization of fine-grained competences (skills); 2) the establishment of certificates and (online) exams that confirm that users possess a certain skill. Note that the certification program does not regulate the content – the definition of skills and certificates is separated from content creation. This allows the re-use of existing content but also allows to create a new ecosystem in which HPC centers or commercial companies could offer the best teaching material. Teaching material should be marked to indicated which skills it covers. In the future, the program may provide means to register and reference existing content of third-parties allowing users to browse the skills and navigate to teaching material. We assume that the collaboration of scientific institutions will complement each other in producing a rich variety of content for the different learning styles.

3.1 Skills

The skills represent competences in a fine-grained fashion. A skill is defined by a unique key, short name, background, level, and a description of what it encompasses. This model can be compared to the classification of school knowledge, for example, the skill with the short name "addition" could describe the math skill of being able to add numbers successfully. The level serves the purpose of distinguishing the expertise further, a *basic level*, for example, may mean to be able to add two numbers between 1-20 while the *expert level* of that skill could indicate to be able to add any numbers.

The individual skills are organized into a tree that shows generic competences close to the root and refined skills on the leafs. The two top-levels of the skill tree are shown in Figure 1; we have identified more than 35 skills. In respect to the granularity, we expect the basic level of a leaf like "Bash programming" can be acquired by novel users in a workshop day. The tree serves the purpose of navigating across the skills, and at the same time a parent node defines the scope of its children. For example, the *USE* competence provides means of using the HPC environment to perform various tasks, such as running parallel applications, using core services of the

operating system. While the tree structure has been chosen as a graphical representation, some competences are cross-referenced, particularly in the skills of the USE branch. The tree is managed in an XML file and we offer tools to visualize the skills as mindmap or embed them into a webpage – more information are found in [1].

3.2 Certificates

Certifying users with a certain competency is a core element of the program. Since we identified so many skills, it is not useful to perform an exam on a single competence like "SLURM". Therefore, initially, we aim to group sets of skills into certificates and establish an online examination that attests users that they mastered a certain skill. To be meaningful these tests must prevent cheating to some extent. However, as any examination can be cheated with sufficient effort, we focus on practical aspects like a huge corpus of questions and some instantiated questions like *how would you start ProgramX with 4 MPI processes?*

4 CONCLUSIONS

The HPC Certification program offers a strategy to classify and organize HPC competences. While the idea started with the PeCoH project, we are supporting the HPC Certification Forum[3] which is an independent international body that aims to sustain the work. The certification forum has the role of a (virtual) central authority to curate and maintain the skill tree and certificates. Moreover, the forum supports tools and an ecosystem around the competences.

REFERENCES

[1] Kai Himstedt, Nathanael Hubbe, Julian Kunkel, and Hinnerk Stuben. 2018. An HPC Certification Program Proposal Meeting HPC Users' Varied Backgrounds. https://www.hpc-certification.org/downloads/hpccp-concept-paper-180201.pdf

[2] Maria-Ribera Sancho. 2016. BSC best practices in professional training and teaching for the HPC ecosystem. *Journal of Computational Science* 14 (2016), 74–77.

---

[1]https://www.eurolab4hpc.eu/

[2]https://portal.xsede.org/web/xup/training/overview

[3]https://hpc-certification.org