

The HPC Certification Forum: Toward a Globally Acknowledged HPC Certification

Julian Kunkel¹, Weronika Filinger², Christian Meesters³, and Anja Gerbes⁴

¹University of Reading

²University Of Edinburgh

³Johannes Gutenberg-Universität Mainz

⁴Goethe-Universität Frankfurt am Main

May 20, 2020

Abstract

The goal of the HPC Certification Forum is to categorize, define, and examine competencies expected from proficient HPC practitioners. The community-led forum is working toward establishing a globally acknowledged HPC certification process, a process that engages with HPC centres to identify gaps in users' knowledge, and with users to identify the skills required to perform their tasks. In this article, we introduce the forum and summarize the progress made over the last two years. The release of the first officially supported certificate is planned for the second half of 2020.

Introduction

The ever-changing nature of High-Performance Computing (HPC) Computing has always compelled the HPC community to invest in continual efforts to train new and existing practitioners. Historically, these efforts were tailored around a typical group of users possessing, due to their background, a certain set of programming skills. However, as HPC has become more diverse in terms of hardware, software, and user backgrounds, the traditional training approaches have become insufficient to address the training needs of our community. An increasingly complicated HPC landscape makes the development and delivery of new training materials challenging. During training delivery, educators need to address the knowledge gaps resulting from the diverse backgrounds of their learners. It is not uncommon for an attendee to come with a specific learning objective related to their work tasks and not be interested in the core HPC knowledge. In that sense, the term HPC practitioner describes anyone involved in providing or using HPC systems, e.g. a user that runs an application on an HPC-resource, a developer for HPC-systems, or an administrator. At the same time, we define the term “HPC” inclusively, capturing parallel computing and cluster computing, e.g. High Throughput Computing (HTC) or Multi-Task Computing (MTC), as these too suffer from a lack of knowledge with regard to performance issues.

How should we develop training for users, often coming from disciplines that have not traditionally used HPC resources, and are only interested in learning a particular set of skills? How can we satisfy their training needs if we don't really understand what these are? HPC centres struggle to identify and overcome the gaps in users' knowledge, while users struggle to identify the skills required to perform their tasks.

The goal of the HPC Certification Forum is to clearly categorize, define, and examine competencies expected from proficient HPC practitioners. The HPC Certification Forum (HPCCF) is the central authority, and

curates and maintains the certification program. The program consists of three parts: competencies defined in a modular and easily expandable skill tree, an examination process to verify that practitioners possess those skills, and the certification demonstrating their knowledge. Although the forum is not involved in the development of any training materials or tools, it supports the ecosystem around the competencies.

The ultimate goal of the forum is to offer a free, globally acknowledged certification program that will make HPC education and training more transparent and quantifiable for training providers, and easier to navigate for practitioners. This article highlights relevant aspects of our activities, more details can be found in (Kunkel et al., 2020).

A Community-Led Forum

The forum is organized around several key roles, which include: the general chair, a publicity chair, and curators for the skill-tree, topics, and examinations. While the board leads the effort, members of the community are expected to contribute to the effort, and anyone is free to benefit from it. Active members can gain nomination and voting rights via an annual steering board election. Decision making is lightweight at the moment: while we have defined roles for steering board members that include final authority in the event it is needed, thus far we have made decisions democratically without the need to rely on this formal mechanism. Basically, any contribution is either accepted or discussed and modified until it is accepted.

The forum uses Slack for its monthly meetings and organizes two face-to-face meetings per year (one at ISC-HPC and the other at the annual SC). GitHub and the Forum’s [webpage](#) are used to coordinate the effort and publish information.

All software used by the forum is Open Source and freely available to allow everyone to participate. The forum aims to provide an ecosystem revolving around the certification specification (including the skill tree and the examination framework) which consists of tools that cover, for example, branding of teaching materials, referring and cross-linking to the competency definitions, and compiling curricula. In particular, we hope to catalogue and reference the existing content of third-parties to allow practitioners to browse the skills and navigate to relevant open and commercial teaching material.

Note that there is currently no direct funding for the effort, but we support all proposals and efforts that members bring forward and associate their work with the forum. For example, in the [ESiWACE](#) project, some contributions regarding HPC IO are expected. Ultimately, we believe that the sustainability of the effort depends upon the recognition of its importance and the voluntary contribution of institutions and individuals.

Categorization of Competencies

The forum groups a well-defined set of competencies into a skill, and a skill is identified by a set of learning outcomes and relevant metadata that clearly specifies what a practitioner should be able to do to be said to possess that skill. The skills are organized in a tree structure from a coarse-grained representation (corresponding to the tree branches) to a fine-grained representation mapped onto the tree leaves. On the leaf level, a skill is orthogonal to other skills – their narrowed scope means they intentionally can be taught in sessions ranging from a 1.5-hour lecture up to a 4-hour workshop. Skills may cover technology-specific knowledge such as the skill “USE1.1-B Command Line Interface” for Linux basics or the skill “K4.2-B SLURM Workload manager” that describes how a cluster manages user jobs. We believe this granularity allows practitioners to select skills relevant to their circumstances, and allows lecturers to prepare modular training sessions with well-defined content while still achieving comparable training outcomes for a varied range of practitioners’ backgrounds. Cross-linking between skills belonging to different branches is allowed and provides for the reuse of the skill definitions and eases the navigation of the tree.

Examination

As the Forum aims to keep the certification free for practitioners, an online examination process has been chosen. The summative assessment will be conducted primarily using multiple-choice questions. For each skill a pool of questions and answers will be created, drawing from both internal and external contributions, and an examination will consist of randomly selected questions. Future developments will include questions beyond the multiple-choice type.

We believe the incentive to deliberately cheat during the assessment, e.g. by having the exam filled by someone else, is low. Therefore, we address this issue in a lightweight and cost-effective fashion. Our process deploys several strategies to minimize the risk of cheating, such as raising the examinees' awareness, using a large pool of questions, setting time limits for each question, and a delay between registering for and taking the actual examination. Since knowledge can quickly become obsolete, each certificate needs to indicate when (month and year) the qualifying examination took place. Also, because the examination of a single fine-grained leaf-level skill would be too easy to pass with short-term memorization and more prone to cheating, the certificates bundle multiple skills together. To ultimately provide trust, the Forum hopes to provide the automatic generation of short tests for prospective employers that would allow validating the knowledge of applicants under their own supervision or in assessment centres.

We manage exam questions internally within the HPCCF consortium. To facilitate external contributions, we provide on our wiki-webpage a light-weight interface for suggesting a question for each skill-set which can be used by anyone – those skill-related questions are reviewed and ultimately managed by the exam curator who is part of the steering board. A well-defined process is being created for prospective contributors to verify the suitability of the contributed questions. For a question to be approved, it will need to undergo a review process where comprehensibility, logic (does it have at least one clear answer?) and rigor (does the question lead to the expected answer?) have to be met. Particularly, each exam will be designed considering a variety of aspects such as competency assessment, learning goals, exam objectivity, and taxonomy of educational objectives (L. W. Anderson, 2000).

As HPCCF exams are potentially interspersed between courses for a particular examinee, examinees may profit from the forward learning effect, i.e., the outcome of such a lightweight test may influence how a practitioner will study the material further. Lecturers, who will be provided with “their” examinees summarized outcomes, may profit from insights to conceptual issues course participants may exhibit.

The questions are the only proprietary component of the HPCCF – using restrictive license terms for authors while giving them credit. By providing potential employers with our tests, too, will have the means to test how an exam is seen and complete it from a student's point of view; this is a measure with established relevance (Roelien Brink, 2011).

Related Efforts

The Certification Program is a new community-wide effort and to the best of our knowledge nothing similar was attempted at such a scale, both with respect to the comprehensibility of the covered content and the international reach of the effort. The other related efforts mainly focus on either providing a comprehensive catalogue of existing training materials and opportunities, providing simple badges confirming participation in a specific training event, or establishing a branded well-defined content and teaching practices recognised by the community. A number of institutions and organisations attempted to catalogue the existing training materials, keep a list of training events and bring the HPC training community together, the ongoing efforts include: [PRACE Training Portal](#), the [HPC University](#) and the [ACM SIGHPC Education Chapter](#). An example of a badging effort is [XSEDE Training Badges Program](#). Finally, the last example refers both to the [Carpentries initiative](#), and [HPC Carpentry](#) which was developed in recent years. The HPC Certification Forum recognises the importance of these efforts and is actively engaging with their contributors.

Conclusions

The program of the HPC Certification Forum allows the existing content to be re-used but also makes it possible to create a new ecosystem in which HPC centres, research labs, academic institutions, and commercial companies could offer the best of their teaching material. The HPCCF aims to support existing activities and complements them by providing a unified and clear way of mapping out the relevant HPC competencies. It should be emphasised that the HPCCF does not regulate the content of training material; we purposely separate the definition of skills, the examination, and the certification from the content delivery. The program does not prescribe a curriculum or any fixed order in which skills should be obtained, thus providing flexibility. It eases the navigation between different competencies without being overly restrictive. We are hoping that a majority of the existing and newly created teaching resources can be branded indicating the skills they cover.

We believe the program will bring multiple benefits to everyone involved in HPC teaching and training. Making clear what skills are required or recommended for a competent HPC user would be helpful to both the HPC service providers and practitioners. Training providers could bundle together skills that are most beneficial for specific user roles and scientific domains, which would allow practitioners to browse through skills to quickly identify and learn the skills required to perform their tasks. The variety of training offered within the HPC community makes finding the right resources more complicated than it should be. We hope that the certification program will eventually provide useful information on where the desired skills are taught. The examination confirming that a certain set of competencies has been acquired makes the learning process more complete and meaningful.

By participating in the program, HPC training providers can increase the visibility of their teaching opportunities and share their resources more effectively. The mapping of the skills defined by the program onto the existing training materials should also help to identify any potential gaps and improve the integrity of the offered training. Finally, the certificates recognized by the whole HPC community will simplify the inter-comparison of independently offered courses and provide additional incentives for participation. Overall, the flexibility of the program allows for the construction of more personalized and just-in-time pathways to learning about HPC.

To achieve these goals, the forum welcomes contributions from volunteers. For this initiative to truly fulfill its role the involvement of the members of the HPC training community with diverse backgrounds and experiences is required. As the HPCCF community is managed collaboratively according to self-managed evolving rules (Ostrom, 2015), it is welcoming and the expected contribution is not demanding, due to the informal nature of the forum.

References

- One Year HPC Certification Forum in Retrospective. (2020). *The Journal of Computational Science Education*, 11(1), 29–35. <https://doi.org/10.22369/issn.2153-4136/11/1/6>
- A Taxonomy for Learning, Teaching, Assessing. A Revision of Bloom's Taxonomy of Educational Objectives.* (2000).
- Electronic assessment in higher education. (2011). *Educational Studies*, 37(5), 503–512. <https://doi.org/10.1080/03055698.2010.539733>
- Governing the Commons: The Evolution of Institutions for Collective Action.* (2015). Cambridge University Press. <https://doi.org/10.1017/cbo9781316423936>
- Webpage of the HPC Certification Forum.* <https://www.hpc-certification.org>