

HPC CF

<https://hpc-certification.org>

Julian Kunkel (+ HPC Certification Forum)

The HPC Certification Forum in 2021



2021-06-27

Annual HPC CF meeting during ISC HPC

Outline

1 The Forum

2 Skills

3 Certification Process

4 Conclusions

Challenges for HPC (and Open Source) Training

- Not all users possess the right level of training
 - ▶ Inefficient usage of systems, frustration, lost potential
 - ▶ Good training saves compute time and costs!
- Diverse user background and goals
 - ▶ Science is the goal, HPC is the vehicle
 - ▶ Need to run an application to complete the PhD
- Learning is not easy
 - ▶ Users need to understand beneficial knowledge for tasks
 - ▶ There exist various different training material
 - ▶ Teaching of different data centers is hard to compare
- Data center have difficulties to verify the skills of users

The HPC Certification Forum

Goals

- Fine-grained standardizing HPC knowledge representation
 - ▶ What competences exist, how are they defined?
 - ▶ Puzzle of competences for everyone (practitioners, students, admins)
 - ▶ Supporting navigation and role-specific knowledge maps
- Establishing international certificates attesting knowledge
- Supporting an ecosystem around the HPC competences

Scope of the forum

- Central authority for competence representation, certification, and support
- Purposeful limitations of the forum:
 - ▶ We do not compete with content providers
 - ▶ We do not create a curriculum (university/centers responsibility)

The HPC Certification Forum

Organization Details

- An independent international body
- Organized into
 - ▶ Steering board (elected)
 - ▶ Full members (with voting rights)
 - Contributors to the project (e.g., 1-2 hours per month)
 - ▶ Associate members (anyone and any institution)
 - ▶ Collaboration with e.g., SIGHPC Education Chapter

Responsibilities

- Curating and maintaining the **Competence Standard**
- Providing tools and ecosystem around the competences

Governance

Various processes are documented [here](#).

Steering Board

- General chair: Julian Kunkel (University of Göttingen / GWDG)
- Skill-tree curator: Kai Himstedt (University of Hamburg)
- Topic curators:
 - ▶ HPC Knowledge: Lev Lafayette (University of Melbourne)
 - ▶ Performance Engineering: Anja Gerbes (University of Dresden)
 - ▶ Software Development: Roberto Villegas-Diaz (South Dakota State University)
 - ▶ Administration: Sudeep Narayan Banerjee (Indian Institute of Technology Gandhinagar)
- Examination curator: Christian Meesters (University of Mainz)
- Publicity chair: Weronika Flinger
- Other topics are jointly managed by the board

Organization

Organization of the members

- Webpage is the central hub (<https://www.hpc-certification.org>)
- Mailinglists (news, members, board)
- Monthly public meetings on our Slack channel
- Annual general assembly (form of a BoF at ISC or workshop)

Data handling

- Everything* is developed/available in the open
GitHub (<https://github.com/HPC-certification-forum>)
- Exception are examination questions

Outline

1 The Forum

2 Skills

3 Certification Process

4 Conclusions

Classification of Competences == Skills

- A **skill** defines background, objectives, learning outcomes
- The **skill tree** organizes the competences as hierarchical skills
- Certificates bundle several skills into attestable unit

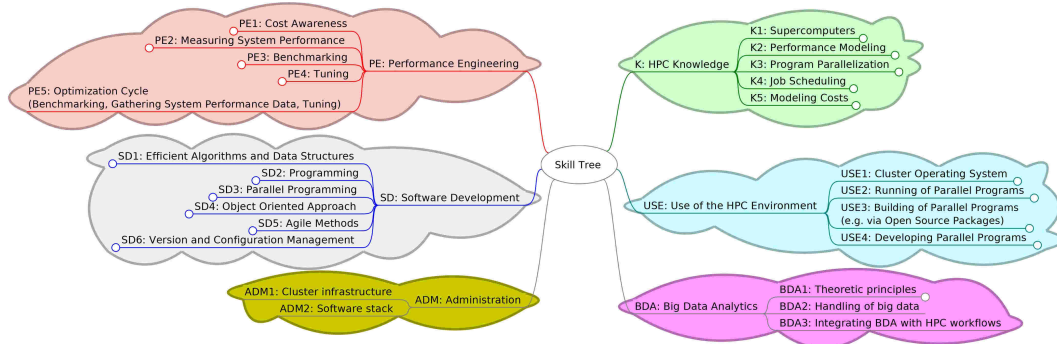


Figure: Top-levels of the skill tree (Initial ADM and BDA branches)

Example High-Level Skill (Excerpt)

- Name: Command Line Interface
- Id: USE1.1-B
- Background: HPC systems are usually accessed via a Linux-based Command Line Interface (CLI) that is provided by a shell. At its core, a shell is ...
- Aim:
 - ▶ describe the key principles of a shell
 - ▶ execute basic programs to query system information and manipulate...

Learning outcomes (these must be examinable)

- Utilize the bash shell to execute individual programs with arguments
- Describe the meaning of the exit code of a program
- Run multiple programs after another depending on the exit code `;`, `&&`, `||`
- List the set of basic programs and their tasks:
 - ▶ `pwd`

Classification of HPC Competences

■ Granularity of skill descriptions

- ▶ Too fine \Rightarrow content of a skill is predefined at leaf level
- ▶ Too coarse \Rightarrow no help for structuring the material
- ▶ Guiding principle: leaf node should be coverable in 1-4 hour lecture/workshop

■ Organization of HPC skills

- ▶ Skills are typically depending on sub-skills \Rightarrow tree structure
- ▶ References to skills are possible; still skills are building blocks for various tasks
- ▶ One skill can have multiple instances for different skill levels (basic, ..., expert)

■ Verification of skill tree and certification approach

- ▶ Feedback by the HPC community/practitioners justify the approaches

Further Considerations

■ Certificate definition

- ▶ Bundles a set of useful skills together
- ▶ A users' HPC qualification is certified by successful exams
- ▶ Testing a single (fine-grained) skill may be too easy with a cheat sheet

■ Separation of **skill**, **certificates** and **content provider**

- ▶ Similar to the concept of a high school graduation exam
- ▶ Learning material can be provided by different institutions
- ▶ Teachers can put badges on material: this "trains skills X, Y, Z"

■ External information can be linked to the skills providing different **views**

- ▶ Suitability for a user role (Tester, Builder, Developer)
- ▶ Suitability for a scientific domain (Chemistry, Physics, ...)
- ▶ View: purpose-specific representation / coloring / content
 - Groups/institutions can derive a new skill tree with their own emphasis
 - What should people know to effectively work in your environment?

Status / Previous Activities

- Development version of the Competence Standard is online
 - ▶ Git managed Markdown files
 - ▶ Files are also available in a Wiki (for interaction)
- Developed various processes

2021 News

- Experts Adopting Skills
 - ▶ Enable experts to curate skills that are in their field of expertise
 - ▶ Similar to code maintainer
- Working on sponsoring
- Developed first (but limited exam!)



This training covers (partially)
- K1.1 System architectures
- K1.2 Hardware architectures
See <https://hpc-certification.org/c/1.0>

All our developments are under open licenses (except the exam questions)

Wiki for Skills

The screenshot shows a web browser window with the address bar displaying `hpc-certification.org/wiki/skill-tree/k/b`. The page title is "K-B HPC Knowledge [HPC x]". On the left, a "SkillTree" sidebar lists various categories, with "K-B HPC Knowledge" expanded to show sub-items like "K1-B Supercomputers", "K2-B Performance Modeling", "K3-B Program Parallelization", "K4-B Job Scheduling", "K5-B Modeling Costs", "PE-B Performance Engineering", "SD-B Software Development", and "USE-B Use of the HPC Environment". The main content area is titled "K-B HPC Knowledge Background" and includes an "Edit" button. Below the title, a paragraph states: "The theoretical knowledge of HPC provides the background to understand how supercomputers and HPC environments operate. This enables practitioners to effectively use such environments." To the right of the main text is a "Table of Contents" with links to "K-B HPC Knowledge", "Background", "Aims", "Outcomes", and "Subskills". Below the paragraph, there are sections for "Aims" and "Outcomes", each with a bulleted list of points and an "Edit" button. The "Aims" section lists three points: providing background knowledge, providing theoretical background to judge system behavior, and providing technical understanding. The "Outcomes" section lists eight points: explaining hardware/software/operation, constructing/judging performance models, understanding performance frontiers, explaining the challenge of speedups/efficiencies, comparing parallelization paradigms, constructing/executing HPC workflows, comprehending job scheduling principles, and applying a cost model. The bottom of the page shows the start of a "Subskills" section.

K-B HPC Knowledge [Edit](#)

Background

The theoretical knowledge of HPC provides the background to understand how supercomputers and HPC environments operate. This enables practitioners to effectively use such environments. [Edit](#)

Table of Contents

- K-B HPC Knowledge
- Background
- Aims
- Outcomes
- Subskills

Aims

- To provide background knowledge that is relevant for all other branches.
- To provide theoretical background to judge the behavior and efficiency of systems.
- To provide technical understanding of HPC systems

[Edit](#)

Outcomes

- Explain the hardware, software, and operation of HPC systems
- Construct and judge simple performance models for systems and applications
- Understand that there are performance frontiers
- Explain why it is a special challenge to achieve good speedups and good efficiencies if the number of processing elements is steadily increased
- Compare different paradigms for the parallelization of applications
- Construct and execute an HPC workflow on an HPC system
- Comprehend job scheduling principles
- Apply a cost model to compute the costs for running a workflow on an HPC system

[Edit](#)

Subskills

Contribution to the Skill-Tree High-Level Editing

How can members contribute?

- Webpage with Markdown version controlled in Git
 - ▶ <https://www.hpc-certification.org/wiki/skill-tree/b>
 - ▶ GitHub: <https://github.com/HPC-certification-forum/skill-tree>
 - Pull requests, reviews, comments, ...
- Editing a MindMap, the structure of Skills
 - ▶ Synchronized with the skill tree in Git
 - ▶ Uses the OpenSource tool Freemind
- Discussion on our [Slack](#)
- Documented in our [processes section](#)
- See our videos on [YouTube](#)

Outline

- 1 The Forum
- 2 Skills
- 3 Certification Process**
- 4 Conclusions

Certification: Assessment

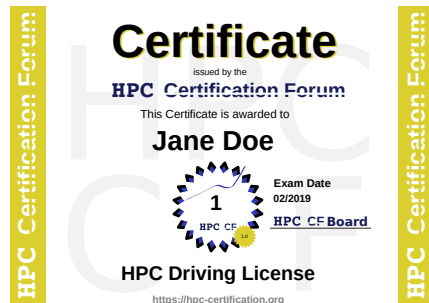
1. User registers to test, receives email
 2. User takes test online (any time!), consists of
 - ▶ Scenario
 - ▶ Multiple choice exam
 - System selects number of questions (and responses) randomly from a pool
 3. Results are submitted to the web server
 4. Automatic approval of response
 5. Automatic creation of certificate and returned by email
 - ▶ Permanent computer-verifiable proof that skill is created
 - Return a text version with GPG signature
 - Return a link that can be verified on hpc-certification.org
- Privacy: minimize information stored on servers, keep some for statistics
- Includes some measure to prevent cheating and brute forcing (e.g., delay)

Certification: Certificate

Text representation

```
-----BEGIN PGP SIGNED MESSAGE-----  
Hash: SHA512  
HPC Certification Forum Certificate  
This text confirms that "Jane Doe" has  
successfully obtained the certificate  
"HPC driving license" (id: 1) at 02/2019.  
Verification URL: https://hpc-certification.org/\[...\]  
-----BEGIN PGP SIGNATURE-----  
[...]  
-----END PGP SIGNATURE-----
```

Certificate



Outline

- 1 The Forum
- 2 Skills
- 3 Certification Process
- 4 Conclusions**

Outlook and Expected Benefits

HPC practitioners

- Increase motivation to participate as the certificates are recognized in a CV
- Validate knowledge via tests
- Browse relevant competences
- Identify recommended and required skills related to certain tasks
- Understand and compare teaching offers across sites

Data centers

- Increase sharing of teaching materials
- Simplifies documentation of taught skills
- Identify missing teaching activities
- Tailor skill-representation specifically to users
- Correlate lack of skills with efficient use

Summary

HPC Certification Program

- Effort to standardize representation/certification of relevant HPC skills
 - ▶ Hierarchical definition of skills for practitioners
 - ▶ Building blocks that can be cherry-picked for different tasks
 - ▶ It's goal is **NOT** to provide content or a linear curriculum
- Perspective for data centers
 - ▶ Use statistics and machine learning to direct users to right skills
 - ▶ Make certain skills a mandatory requirement?
- Customizable representation and navigation for data centers/domains
 - ▶ Interactive viewer to browse skills and related content
 - ▶ We will use the viewer to link good content to the skills, too!
- Visit us and join our Slack/ mailing lists: <https://hpc-certification.org>