

The HPC Certification Program



J. Kunkel on behalf of the HPC Certification Forum

<https://hpc-certification.org>

HPC-SIG meeting

2019-02-13

Outline



1 The Program

2 The HPC Certification Forum

3 Conclusions & Roadmap

HPC Certification Program



Motivation

- Not all users possess the right level of training
 - ▶ Inefficient usage of systems, frustration, lost potential
 - ▶ Good training saves compute time and costs!
- 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 Program



Goals

- Standardizing HPC knowledge representation
 - ▶ What competences exist, how are they defined?
 - ▶ Supporting navigation and role-specific knowledge maps
- Establishing international certificates attesting knowledge

The work was bootstrapped and is supported by the PeCoH project

Important!

- We do not compete with content providers
- We do not intent to create a curriculum

Contributions by the PeCoH Project¹



Past contributions

- 1 Initial classification of competences
- 2 Initial development of a certification program
- 3 Prototype for the online examination

The program will be curated by the **HPC Certification Forum**

Pending independent contributions

- Creation of workshop material for (a) basic certificate(s)
- Providing an online tutorial for (a) basic certificate(s)

¹PeCoH was supported by the German Research Foundation (DFG) under grants LU 1353/12-1, OL 241/2-1, and RI 1068/7-1.

Content of the Certification Program

- Skill is characterized by ID, background, objectives, outcomes
- Skill tree defining the organization of the competences
- Certificates bundle several skills into attestable unit

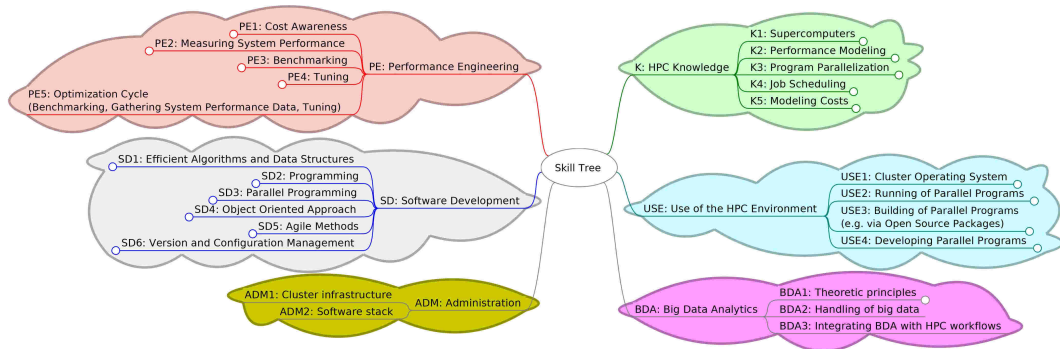


Figure: Top-levels of the skill tree (We are working on ADM and BDA)

Example High-Level Skill



- Name: SLURM Workload manager
- Id: USE4.2.2-B
- Background: SLURM is a widely used open-source workload manager providing various advanced features.
- Objectives:
 - ▶ comprehend and describe the basic architecture of SLURM and its tools
 - ▶ use relevant tools to run and monitor (parallel) applications

Outcomes

- run interactive jobs with salloc, a batch job with sbatch
- explain the architecture of SLURM, i.e., the role of slurmd, srun
- explain the function of the tools: sacct, sbatch, salloc, ...
- explain time limits and the benefit of a backfill scheduler
- (see <https://www.hpc-certification.org/wiki/>)

Classification of HPC Competences



- HPC skills are generally built upon one another
 - ▶ Skills are typically depending on sub-skills \Rightarrow tree structure
 - ▶ References to skills are possible; still skills are building blocks for various tasks
- Additional attributes are currently used to describe:
 - ▶ Level of a skill (Basic, Intermediate, Expert)
 - ▶ Suitability for a user role (Tester, Builder, Developer)
 - ▶ Suitability for a scientific domain (Chemistry, Physics, ...)
- We will extract these attributes to generate "views" on the content
 - ▶ View: purpose-specific representation / coloring / content

Considerations



- 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 2-4 hour lecture/workshop
- Certificate definition
 - ▶ Bundles a set of useful skills together
 - ▶ A users' HPC qualification is certified by successful exams
- 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 XYZ"
- Support flexible usage (views on skill tree) and skills is key
 - ▶ Institutions can derive new skill tree with own groups
e.g. users in weather/climate, single program, testers

Status



- A structure of the skill tree is released (35 basic skills)
- We are working on descriptions for (learning) objectives and outcomes
- Idea been discussed to DKRZ user group; they want it yesterday
- Technical representation of the HPC skills
 - ▶ Database for the HPC certification program
 - ▶ Implementation is based on XML with Schema
 - ▶ A markdown version and editable wiki exists but is yet not synchronized
- JavaScript for visualization of skill tree ([Live Demo](#))
 - ▶ Can be embedded in your webpage and adapted
 - Role/software-specific knowledge
 - What should people know to effectively work in your environment?

Certification



Approach

- User takes online multiple-choice test
 - ▶ A combination of JavaScript and a web service
 - ▶ System selects number of questions randomly from a pool
 - ▶ System draws 4-5 responses from 10 possible responses
 - Some responses may be parameterized (random)
- Choices are submitted to the web server
- Manual approval of the result
- Automatic creation of certificate and return by email
- Permanent computer-verifiable proof that skill was obtained
 - ▶ Return a text version with GPG signature
 - ▶ Return a link that can be verified on hpc-certification.org
- Includes some measure to prevent cheating and brute forcing (e.g., delay)

Outline



- 1 The Program
- 2 The HPC Certification Forum
- 3 Conclusions & Roadmap

The HPC Certification Forum



The HPC-CF the central authority for the development of the program

Organization Details

- An independent international body
- Organized into
 - ▶ Steering board
 - ▶ Full members with voting rights
 - ▶ Associate members

Responsibilities

- Curating and maintaining the skill tree and certificates
- Providing tools and ecosystem around the competences

Governance



- We have an initial set of governance rules
- Splitting responsibility across different roles

Current Chairs

- Program chair: Julian Kunkel (University of Reading)
- Curriculum chair: Kai Himstedt (University of Hamburg)
- Topic chairs:
 - ▶ HPC Knowledge: Lev Lafayette (University of Melbourne)
 - ▶ Performance Engineering: Anja Gerbes (University of Frankfurt)
 - ▶ Use of the HPC Environment: Jean-Thomas Acquaviva (DDN)
 - ▶ Software Development: Waseem Kamleh (University of Adelaide)
 - ▶ Administration: Sharan Kalwani (DataSwing)
 - ▶ Big Data Analytics: Julian Kunkel // Interim
- Examination chair: not seated this year
- Publicity chair: Weronika Filingner

Outline



- 1 The Program
- 2 The HPC Certification Forum
- 3 Conclusions & Roadmap

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 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 mailing lists: <https://hpc-certification.org>

Roadmap: ISC-2019



- Finalizing the first version of the skill-tree (V1)
- Finalizing documentation how to create views with the JavaScript
 - ▶ This will allow to outsource roles (e.g., tester) but also link to material
- Finalizing the seal that can be added to training material
 - ▶ We are looking forward to work with together with anyone
- Completing "Big Data Analytics for HPC" and Admin subtrees
- Creating converters for markdown version of the skill-tree
- Completing the online exam for the certification (of some base courses)
 - ▶ Determine legal disclaimer / terms of service (help welcome)
- Link third-party workshop material (of some base courses)
- Discuss affiliation program for institutions/companies
- Map out content by the *SIGHPC Education Chapter Content Committee*