Introducing the HPC Certification Program

J. Kunkel on behalf of the HPC Certification Forum

https://hpc-certification.org Webinar

2018-10-31
Outline

1. The Program
2. The HPC Certification Forum
3. Presentations from Representatives
4. Conclusions & Roadmap
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
  - Supporting navigation and role-specific knowledge maps
- Establishing international certificates attesting knowledge

*The work was bootstraped and is supported by the PeCoH project*
Contributions by the PeCoH Project\footnote{PeCoH was supported by the German Research Foundation (DFG) under grants LU 1353/12-1, OL 241/2-1, and RI 1068/7-1.}

Past contributions

1. Initial classification of competences
2. Initial development of a certification program

The program will be curated by the HPC Certification Forum

Pending contributions

- Creation of workshop material for basic certificates
- Providing an online tutorial for basic certificates
- Enabling an online examination
Content of the Certification Program

- Skill is characterized by unique key, background, knowledge covered
- Skill tree defining the organization of the competences
- Certificates bundle several skills into attestable unit

**Figure:** Top-levels of the skill tree

- Content is **NOT** covered and subject to content providers
  - We may link good content on our page
Example High-Level Skill (Version 0.5)

- Name: Hardware Architectures
- Id: K1.2-B
- Level: Basic
- Category: HPC Knowledge

Description of knowledge, i.e., what will the user learn

The skill delivers a high-level overview of:

- Elementary processing elements like CPUs, GPUs, many core architectures
- Vector systems, and FPGAs
- The NUMA architecture used for symmetric multiprocessing
- Network demands for HPC systems (e.g. high bandwidth and low latency)
- Typical network architectures like fast Ethernet (1 or 10 Gbit) or InfiniBand
Classification of HPC Competences

- HPC skills are generally built upon one another
  - Skills are typically depending on sub-skills → tree structure
  - References to skills are possible; still skills are building blocks for various tasks

- Additional attributes are 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, ...)

- Skill tree supports different "views" on the content
  - View: purpose-specific representation / coloring / content
Considerations

- **Granularity of skill descriptions**
  - Too fine ⇒ content of a skill is predefined at leaf level
  - Too coarse ⇒ no help for structuring the material
  - Actual skill tree contains 76 skills

- **Certificate definition**
  - Bundles a set of skills
  - 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
A first version of the skill tree is released (35 basic skills)

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
  - Corresponding XML Schema (XSD) assures consistency
  - Contribution is available on GitHub
    https://github.com/HPC-certification-forum/curriculum

JavaScript for visualization of skill tree
  - Can be embedded in your webpage and adapted
    • Role/software-specific knowledge
    • What should people know to effectively work in your environment?

Live Demo
Outline

1. The Program
2. The HPC Certification Forum
3. Presentations from Representatives
4. 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 – 2018

Governance Rules

- 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 (to be confirmed by the board): Sharan Kalwani (DataSwing)
- Examination chair: not seated this year
- Publicity chair: Weronika Filinger
Outline

1. The Program
2. The HPC Certification Forum
3. Presentations from Representatives
4. Conclusions & Roadmap
Outline

1. The Program
2. The HPC Certification Forum
3. Presentations from Representatives
4. 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
  - Mandatory requirements
- Customizable navigation for compute centers
  - Interactive viewer to browse skills and related content
- Visit us and join our mailing lists: https://hpc-certification.org
Roadmap: ISC-2019

- Finalizing the first version of the skill-tree (V1)
- Finalizing the seal that can be added to training material
  - We are looking forward to work with together with anyone to link material!
- Finalizing documentation how to create views with the JavaScript
  - This will allow to outsource roles (e.g., tester) but also link to material
- Adding a "Big Data for HPC" skill subtree
- Creating a markdown version of the skill-tree (with converters)
- Embed a version that can be edited in an online Wiki (by members)
- Creation of workshop material (of some base courses)
- Creating an online certification (of some base courses)
Certification

Approach

■ User takes online multiple-choice test
  ► Likely to use 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 scores

■ Automatic creation of certificate + permanent record

■ Permanent record proofs that a certain user has a skill
Open Questions

- Affiliation program for companies?

- Examination
  - Cheating has been a discussion, but we won’t focus on this right now
  - How to ensure that a person does not brute force the exam?
    - Delay between retry?

- Determine legal constraints for exams (help welcome)

- Create a unique proof that a user has a skill
  - Hash instead of person name?