The HPC Certification Program

J. Kunkel on behalf of the HPC Certification Forum

https://hpc-certification.org  ISC HPC

2019-06-18
The HPC Certification BoF: Agenda

Short presentations

- Introduction to the HPC Certification Program (10 min)
- HPC tools and usage: a view from the community (5 min)
- Skill tree introduction (7 min)
- Mapping content to skills: a case study from PeCoH (3 min)
- Certification process introduction (3 min)

Community discussion (20 min)
We prepared several stations in the room allowing you to give feedback

Wrapup (10 min)

- Roadmap, Summary & Conclusion
- Questions to the audience & joint discussion
Outline

1. The Program
2. The HPC Certification Forum
3. Certification Process
4. Roadmap & Conclusions
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

Important!

- We do not compete with content providers
- We do not intent to create a curriculum
Content of the Certification Program

- 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 (We are working on ADM and BDA branches)
Classification of HPC Competences

- **Organization of HPC skills**
  - Skills are typically depending on sub-skills ⇒ 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

- **Granularity of skill descriptions**
  - Too fine ⇒ content of a skill is predefined at leaf level
  - Too coarse ⇒ no help for structuring the material
  - Guiding principle: leaf node should be coverable in 2-4 hour lecture/workshop

- **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?
Further Considerations

- **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"

- **Verification of skill tree and certification approach**
  - We utilize the HPC community/practitioners to justify approaches
Status

- Organizing regular meetings (see our webpage)
- Released a first skill tree (we will discuss this after the talks)
- Released technical representations of the HPC skills
- Released JavaScript for visualization of skill tree (demo)
  ▶ Enables views: adjustable/embedable in your webpage
- Developed prototype for exam process: legal framework
- Designed seal of endorsement
- Engaged with various stakeholders (e.g., SIGHPC Edu)
- Conducted survey to verify the skill tree (more to come!)

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

1. The Program
2. The HPC Certification Forum
3. Certification Process
4. Roadmap & Conclusions
The HPC Certification Forum

The HPC-CF is 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 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: was not seated in 2018
- Publicity chair: Weronika Filinger
Outline

1. The Program
2. The HPC Certification Forum
3. Certification Process
4. Roadmap & Conclusions
Certification: Assessment Prototype

1. User takes multiple-choice test online (any time!)
   - A combination of JavaScript and a web service
   - System selects number of questions randomly from a pool
     • The questions are managed with rigorous license agreement
   - System draws 4-5 responses from 10 possible responses (some randomized)

2. Choices are submitted to the web server

3. Manual approval of the result

4. 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)
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-----
Outline

1. The Program
2. The HPC Certification Forum
3. Certification Process
4. Roadmap & Conclusions
Roadmap: Supercomputing 2019

- Finalizing documentation how to create views with the JavaScript
  - This will allow to outsource roles (e.g., tester) but also links to material
- Completing the online exam for the certification (of some base courses)
  - Explore the use of interactive input, e.g., simulate shell for job submission
- 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
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
Questions and Open Issues for Discussion

- Naming, e.g., Software engineering vs. software development
- How easy should editing of a skill be for members?
- How relevant are different branches of the tree (dev. vs. testers)?
- Who would be willing to create a data center/user specific view?
- How willing are people to share teaching material and what formats?
- How shall we link to it?
- Would a site like to get feedback about the examination of their users?
- What would be expected from HPCCF to support teaching?
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](https://hpc-certification.org)