One Year HPC Certification Forum in Retrospective

Julian Kunkel, K. Himstedt, W. Filinger, JT. Acquaviva, A. Gerbes, L. Lafayette

https://hpc-certification.org

HETET Workshop

2019-06-20

Copyright University of Reading
Outline

1. The Program
2. The HPC Certification Forum
3. Skills
4. Certification Process
5. Summary & Conclusions
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

Important!

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

Scope of the talk/paper

■ Achievements in one year of the program
■ Status of the effort
Status of the Activity

- Organizing regular meetings (see our webpage)
- Released a first skill tree
- 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 Skills

4 Certification Process

5 Summary & 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. Skills
4. Certification Process
5. Summary & Conclusions
**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 units**

![Skill Tree Diagram](image)

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

**Aim:**
- comprehend and describe the basic architecture of SLURM and its tools
- use relevant tools to run and monitor (parallel) applications

**Learning 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/](https://www.hpc-certification.org/wiki/)
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
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)
The Program

The HPC Certification Forum

Skills

Certification Process

Summary & Conclusions

The HPC Certification Forum Skills Certification Process Summary & Conclusions

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-----
Outline

1. The Program
2. The HPC Certification Forum
3. Skills
4. Certification Process
5. Summary & Conclusions
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)