HPC CF

https://hpc-certification.org

Anja Gerbes

# Benefit for utilizing HPC CF - Partitiicipation of the TU Dresden

# 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
- Confusion of trainers and HPC practitioners regarding skills
  - What should one person know/train regarding "MPI"?

# The HPC Certification Forum

## Goals

- Fine-grained HPC knowledge representation ⇒ Competence Standard
  - 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

- Started in 2018 as a spin-off from a project
- An independent international body
- Organized into
  - Steering board (elected)
  - Full members (with voting rights) - Contributors
  - Associate members (anyone and any institution)
  - Collaboration with e.g., SIGHPC Education Chapter

## Activities

- 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 (Technische Universität Dresden)
Sofware Development: Marc-Andre Hermanns (RWTH Aachen)
Administration: Sudeep Narayan Banerjee (Indian Institute of Technology Gandhinagar)

Publicity chair: Weronika Filinger (University of Edinburgh)

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
- Monthly meeting of the board
- Annual general assembly (BoF at ISC or independent workshop)

## Data handling

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

# Motivation

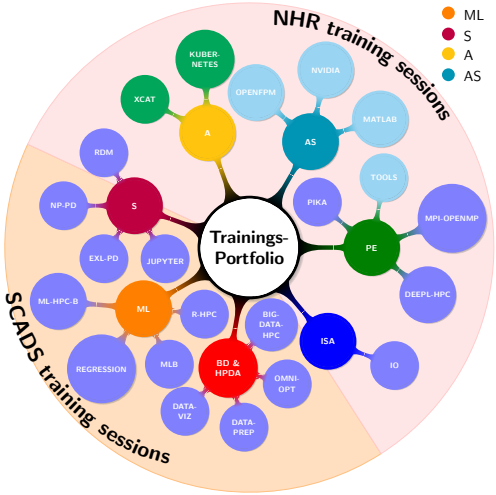
TU Dresden

- is a member of the National High Performance Computing (NHR) since January 2021.
  (https://tu-dresden.de/zih/hochleistungsrechnen/nhr-center)
- defined several competences in their NHR application.
- started their NHR Training Sessions in September 2021.
  (https://tu-dresden.de/zih/hochleistungsrechnen/nhr-training)

# NHR Training-Portfolio 2023 @ ZIH

**Target Group**
- HPC Beginner
- HPC User
- HPC Dev
- HPC Admin
- HPC Expert

| Keyword | NHR Competencies |
|---|---|
| PE | Performance Engineering |
| ISA | Innovative Storage Architecture |
| BD & HPDA | Big Data & High Performance Data Analytics |
| ML | Machine Learning |
| S | NHR Services |
| A | Administration |
| AS | Application Science |




**Course Type**
- NHR-Tutorial → Course with Hands-On
- NHR-Lecture → Course without Hands-On
- NHR-Workshop → Workshop

# Outline

# 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

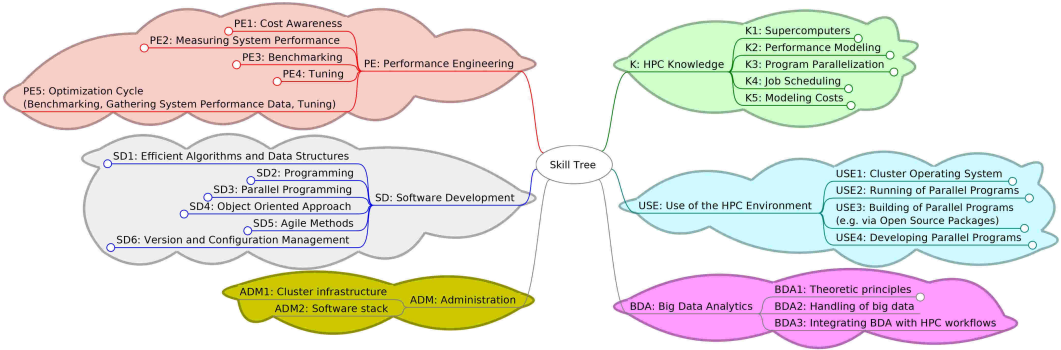



Figure: Top-levels of the skill tree (Initial ADM and BDA branches)

# 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 covered 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

# 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

# NHR Training-Portfolio 2023 @ ZIH

Speaker:

- Define Course Type
- Define Target Group
- Define Course Title
- Write Summary
- Define Agenda
- Create Reference Guide (optional)
- Define Questions for Survey
- Define Prerequisites
- Define Learning Objectives

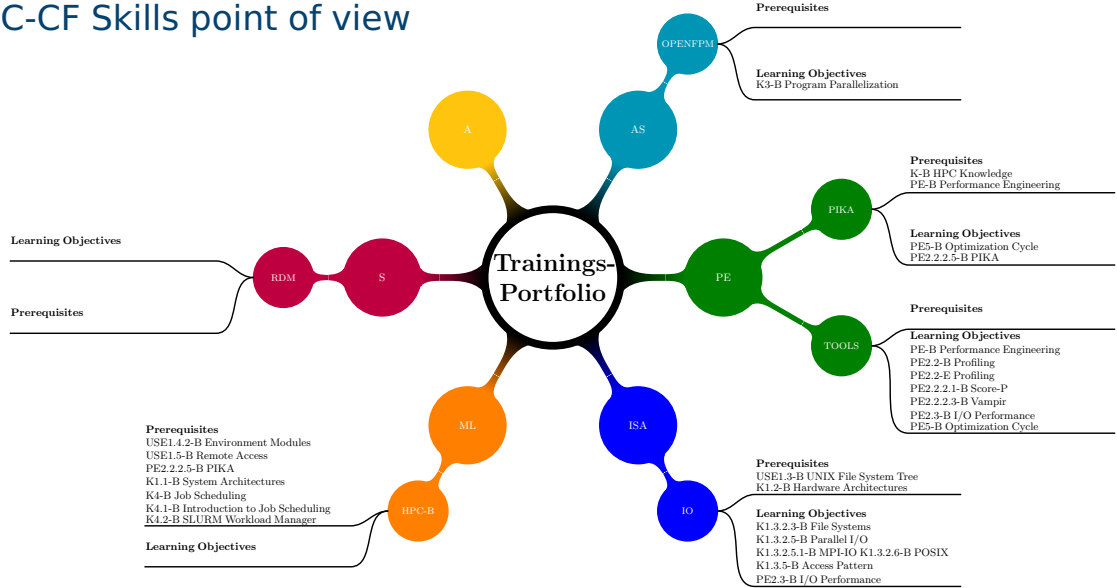
**mapping**

- Search/Define HPC-CF Skill Tree Entry
  - Background
  - Aim
  - Outcomes

NHR Coordinator:

- Define Questions for Survey
- Create Course Website Link
- Create Registration Link
- Create Survey Link
- Create Certificate of Participation

# HPC-CF Skills point of view

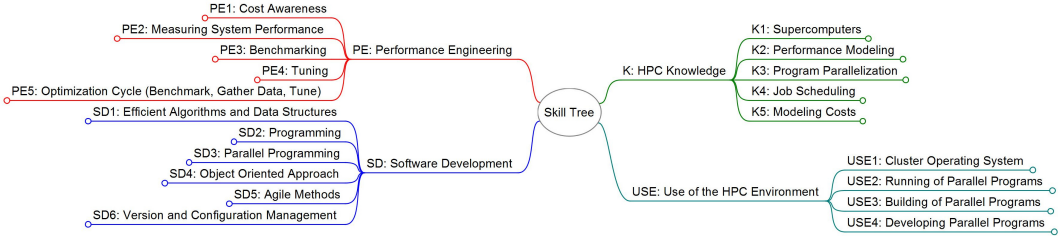


# HPC-CF Skill Tree

**HPC-CF ID**
PE2.2.2.5-B → position in the skill tree

**Target Group**
B → Beginner
I → Intermediate
E → Expert

| **Keyword** | **HPC-CF Competencies** |
|---|---|
| ADM | Administration |
| BDA | Data Analytics |
| PE | Performance Engineering |
| K | HPC Knowledge |
| SD | Software Development |
| USE | Use of the HPC Environment |



# HPCCF Skills of NHR Trainings

- **Description**
- Skill updated in HPC-CF
- Skill didn't exist in HPC-CF
- Skill already exist in HPC-CF

https://www.hpc-certification.org/wiki/skill-tree/b

- PE5-B Optimization Cycle
- PE2.2.2.5-B PIKA
- K1.3.2.3-B File Systems
- K1.3.2.5-B Parallel I/O
- K1.3.2.5.1-B MPI-IO
- K1.3.2.6-B POSIX
- K1.3.5-B Access Pattern
- PE-B Performance Engineering
- PE2.2-B Profiling
- PE2.3-B I/O Performance
- PE2.3-I I/O Performance
- PE2.2-E Profiling
- PE5-B Optimization Cycle
- PE2.2.2.1-B Score-P
- PE2.2.2.3-B Vampir