

Education in HPC: A Lifelong Effort

Introduction

Goals for this session

- Create an understanding for HPC education and training as a lifelong effort
- Show-case education and training efforts
- Stimulate feedback on specific needs for training and education
- Stimulate exchange on best-practices

Speakers

Education and training initiatives

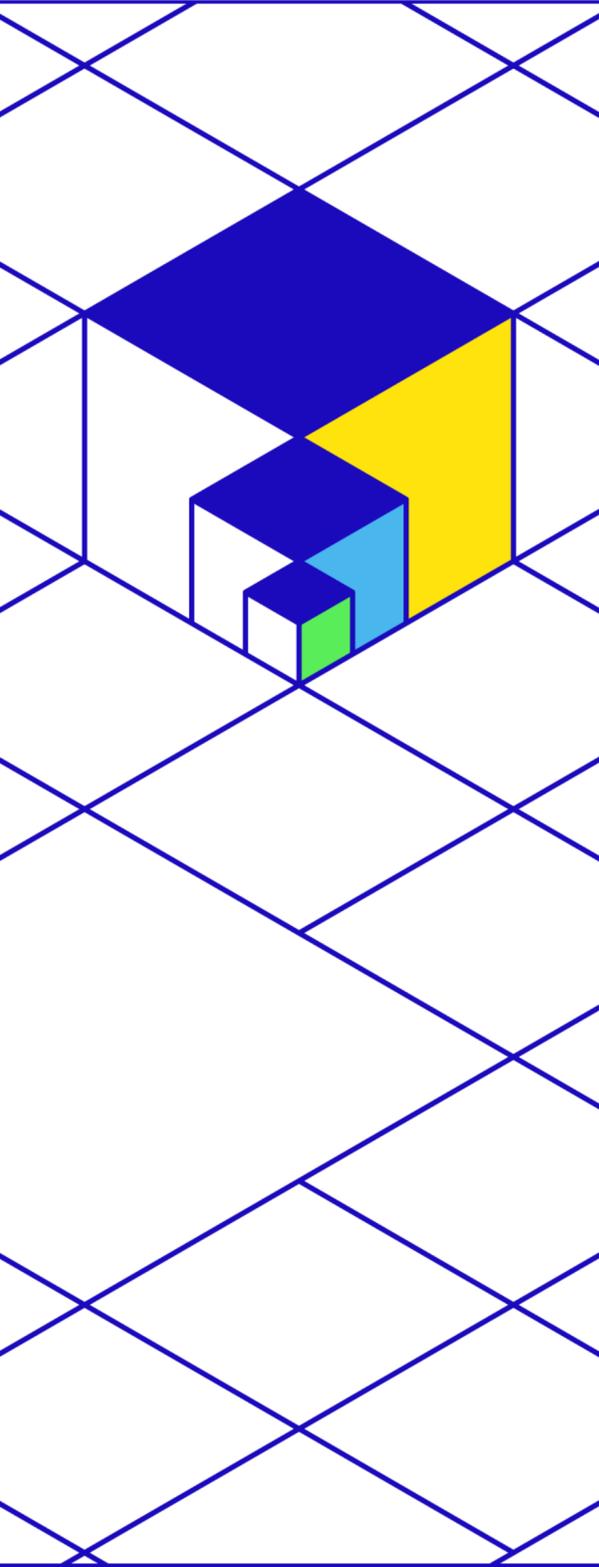
- Tiina Leiponen (CSC)
- Thor Wikfeldt (RISE)
- Pascal Bouvry (Unilu)

Needs for education

- Maria Ribera Sancho (UPC)
- Eric Monchalín (Atos)

Best practices in education and training

- Andreas Hertén (FZJ)
- Johan Hellsvik (KTH)
- Pratik Nayak (KIT)



Education in HPC: A Lifelong Effort

PRACE Training Centres

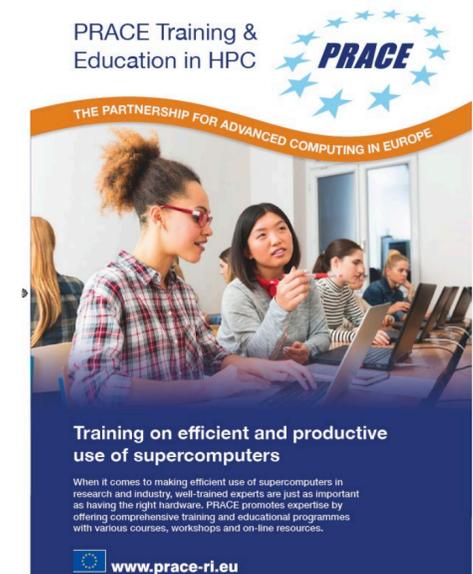
PRACE PTCs

- **14 PRACE Training Centres (PTCs) in Europe**
- **Established a state-of-the-art curriculum for training in HPC and scientific computing**
- **Free of charge training for European research and industry**
- Nurtured training quality
- Backbone of PRACE training
- Extensive and developing offering - steady marketing practises
- Knowing the users - reaction and adaptation to novel technologies
- Adaptation to environmental change and pandemic – rise of online training, MOOCS



People – The Power of the PRACE PTCs

- PRACE is a well-known, trusted and valuable brand in HPC training – based on customer experience and shared partnership promise of delivering top-class HPC training, already since 2008.
- Common to the PTCs: Competence, tradition and commitment to training, maintaining and developing the HPC training legacy
- People willing to collaborate, share experiences, discuss and agree with the common plans, guidelines, practises and development ideas are the glue that keeps the network alive
- Professional course delivery, integrated plans, continuous participant statistics, feedback (ICHEC, Simon Wong)
- Sustainable and experienced pool of trainers, established processes, sufficient funding, suitable venue
- Trust between partners – It is the people we know making the training events possible



Clear Training Strategy

- PTCs place in the HPC training ecosystem, place in the PRACE training portfolio
- PTC's goal to maintain the core course offering: courses participants expect to find and what they need in their profession
- Flexible, personal ways of reaching the common goals as partners
- Basic guidelines to annual plans, regular active discussion in the meetings
- Collaboration and multilateral partnerships between partners and the HPC community

Offering PRACE training

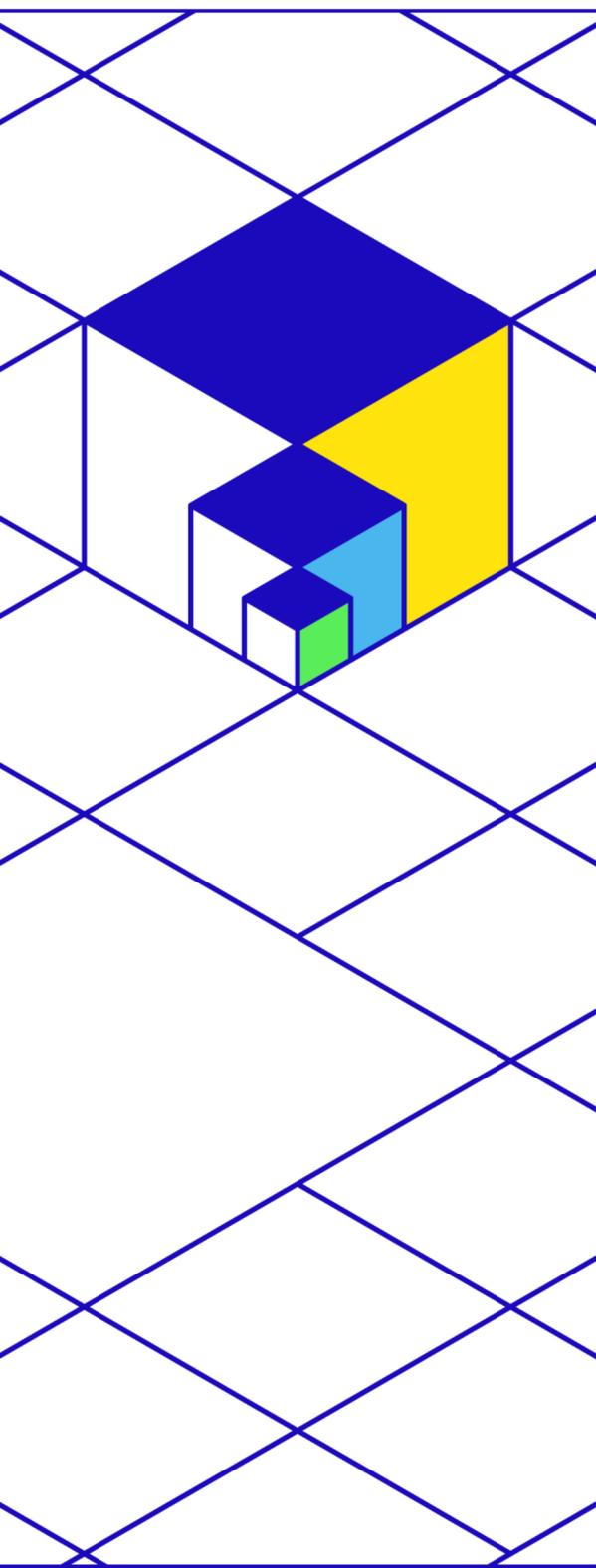
- PRACE training and events portal – the marketplace for PTCs
- Enable viewing and inputting PRACE PTC trainings as well as training offered by other partners and projects
- Materials, videos and more from the network
- One stop shop for training offering
- Developed and revamped over times
- Requires continuous maintenance and updates
- PTCs adapted Indico system as the primary source of all training meta data

The screenshot displays the PRACE Training Centre Events portal. The header features the PRACE logo and the text "PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE". Below the header, there is a navigation bar with "Home", "Create event", and "My profile" options. The main content area is titled "PRACE Training Centre Events" and includes a search bar and a "Create event" button. A message indicates that some events have been hidden. A list of events for December 2022 is shown, including:

- 26 Dec: [ONLINE] EPCC Template PATC event (DO NOT REMOVE) @ EPCC (past)
- 14 Dec - 16 Dec: [ONLINE] R in an HPC environment @ SNIC
- 13 Dec - 14 Dec: [ONLINE] Performance Optimisation on AMD EPYC @ EPCC 13-14 December 2022
- 13 Dec - 14 Dec: [Online] Introduction to Modern Fortran @ EPCC 13-14 December 2022
- 13 Dec - 14 Dec: [ONLINE] Hybrid Programming in HPC - MPI-X @ VSC Vienna
- 08 Dec - 18 Dec: [ONLINE] Porting of Codes for Next Generation GPU Architectures @SURF
- 26 Dec: [ONLINE] Introduction to machine learning in the application area of fluid mechanics and combustion using HPC @ JSC
- 27 Dec: [ONLINE] Data Management with IRODS and Compute @SURF
- 27 Dec - 08 Dec: [ONLINE] Parallel Programming with OpenMP @ IT4Innovations
- 05 Dec - 08 Dec: Software Carpentry @ EPCC at Queen's University Belfast 5-8 December 2022
- 06 Dec - 07 Dec: [ONLINE] Node-Level Performance Engineering @ LRZ

All time PRACE Training Achievements

Activity/Course	#events	#training days	#participants
PRACE Training Centres	955	2 608	26 081
Seasonal Schools and workshops	45	175	2 161
International HPC Summer School	11	54	808
On-demand events	14	43	388
TOTAL (short courses)	1 025	2 880	29 438
Feedback from the participants	Positive feedback		Average overall rating 8,4/10 remained stable through the years
Summer of HPC	Organised annually for 10 years since 2013; mentored 289 students in Europe.		
MOOCs	8 MOOCs developed since 2017; a total of 37 929 people joined these courses.		



Education in HPC: A Lifelong Effort

EuroCC Training

Thor Wikfeldt (RISE/ENCCS), 2023-03-22

With input from:
Martina Blazkova
Carlos Teijeiro Barjas



EuroCC



- 33 National Competence Centers
- Bringing together expertise
- Increase national strengths in HPC, HPDA, AI
- One-stop shop for HPC

Target groups



Industry



Academia



Public administration

Services



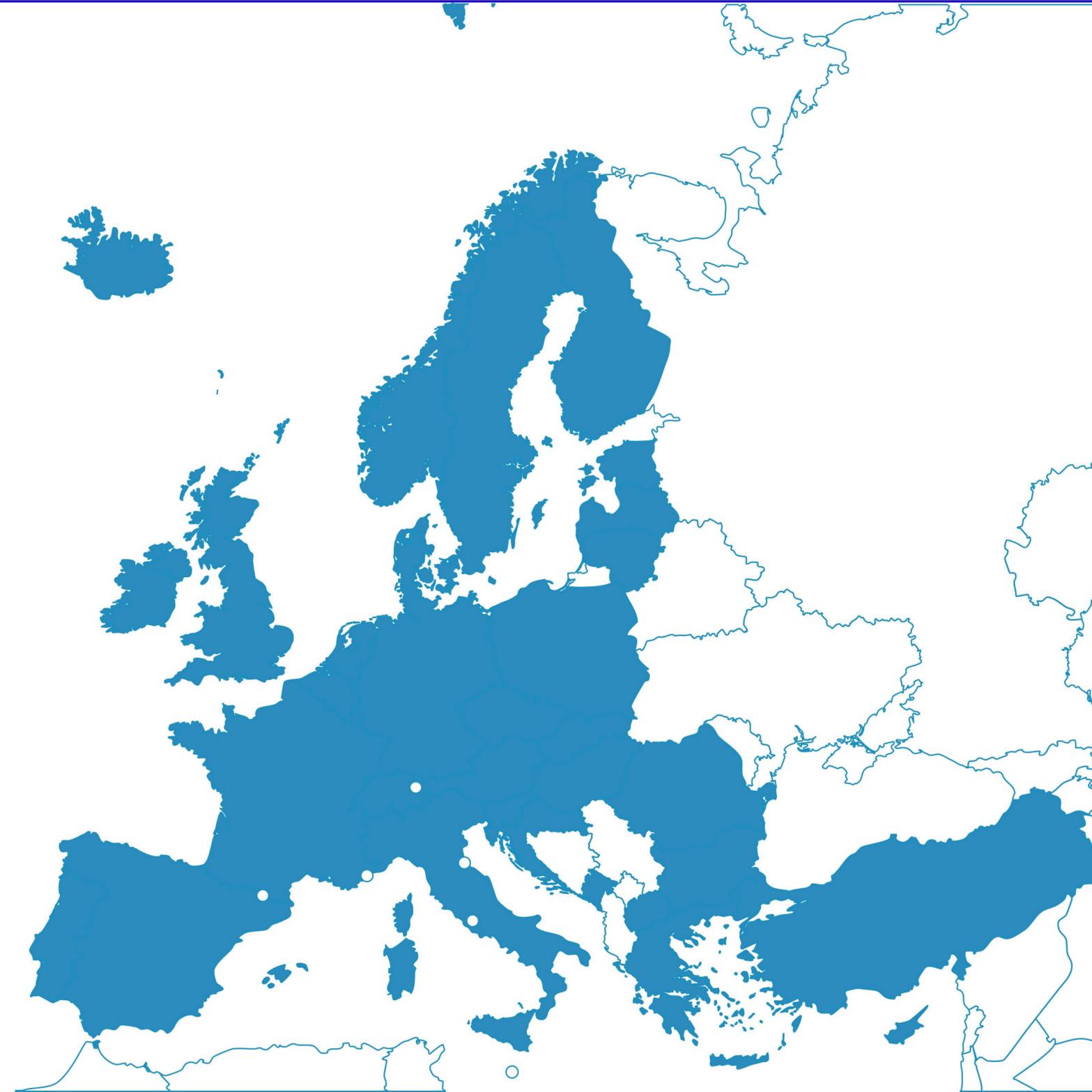
Training



Software support



HPC allocation help



EuroCC



- 33 National Competence Centers
- Bringing together expertise
- Increase national strengths in HPC, HPDA, AI
- One-stop shop for HPC

Target groups

Industry Academia Public administration

Services

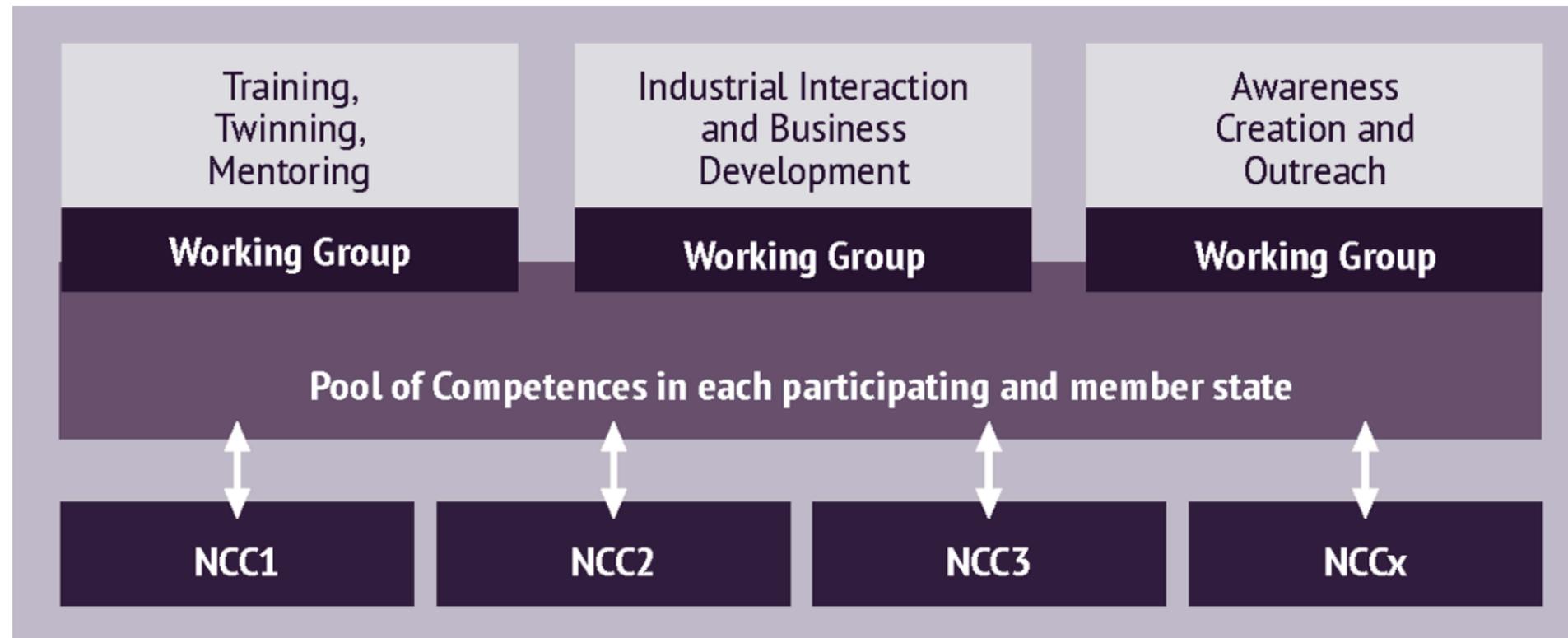
Training Software support HPC allocation help

A map of Europe is shown with several event cards overlaid on it. The cards provide details for various HPC-related events, including:

- Introduction to ANSYS CFX on LRZ HPC Systems** (Germany, Mar 16 - Apr 27 2023)
- Iterative Linear Solvers and Parallelization** (Germany, Mar 20 - Mar 24 2023)
- Modern C++ Software Design (Advanced)** (Austria, Mar 20 - Mar 23 2023)
- Artificial intelligence and machine learning** (Estonia, Fri Mar 31 2023)
- Fundamental Quantum Computing Algorithms and Their Implementation in Qiskit** (Czech Republic, Apr 03 - Apr 05 2023)
- Parallelization with MPI and OpenMP** (Germany, Apr 11 - Apr 14 2023)

Each card includes a description of the event, a link to the event website, and icons for Google and Outlook.

CASTIEL



- **Connecting NCCs** – recruiting teachers, co-organizing workshops
- **Internationalisation** – advertising NCC workshops
- **Workshops** with hardware providers
- **Mentoring workshops** – “HPC, HPDA and AI Applied to Industrial Use Cases”, and “Quantum Computing”
- **Best practice guides** – Train the Trainer, How to Organise an HPC Event, How to Find Attendees for Trainings, In-Person and Online Training Courses

Training portals

- Several portals exist
 - <https://hpc-portal.eu/>
 - <https://www.eurocc-access.eu/services/training/>
 - <https://events.prace-ri.eu/>
- NCCs encouraged to share training events
- Enables HPC users to find online training not available locally

Challenges

- Multiple portals
- No standardised curriculum, taxonomy, certification
- Not widely known in key user communities

Training

Find upcoming training offers from the National Competence Centres here! From beginner to expert courses, whether C++ or Open MP, there's something for everybody. If you want to see all training offers, including past courses, go to the [HPC in Europe Portal!](#)

Event feed: [Download](#) [Copy URL](#)

Search events:

All items per page

		
Germany Mar 13 - Mar 17 2023	Online Wed Mar 15 2023	Online Mar 15 - Mar 16 2023
CFD with OpenFOAM®	Working on the VSC Clusters	Einführung in ParaView zur Visualisierung von wissenschaftlichen Daten
Germany	Austria	Germany
OpenFOAM® is a widely-used open-source code and a powerful framework for solving a variety of problems mainly in the field of CFD...	In this course we will help you getting started on the VSC clusters, Austria's most powerful supercomputers. We'll focus on using the...	ParaView ist eine auf dem Visualization Toolkit (VTK) basierende Open-Source Software, mit der wissenschaftlich-technische Datensätze...
Click here for more details	Click here for more details	Click here for more details

Training materials

- CASTIEL shares material and recordings
 - <https://www.eurocc-access.eu/services/video-library/>
- ENCCS is building a repository of public and open-source training material
 - <https://enccs.se/lessons/>
- Training materials are increasingly being shared among NCCs, CoEs and others

Challenges

- Lack of coordination and collaboration in training material development
- No generally agreed-upon content level specification
- Much high-quality training material is not shared

 <p>High-performance Data Analytics with Python</p> <p>COURSE MATERIAL</p>	 <p>Julia for High Performance Scientific Computing</p> <p>COURSE MATERIAL</p> <p>VIDEO RECORDING</p>	 <p>SYCL Workshop</p> <p>COURSE MATERIAL</p> <p>VIDEO RECORDING</p>	 <p>A.I. as a Tool for Change</p> <p>VIDEO RECORDING</p>
 <p>Graph Neural Networks and Transformer Workshop</p> <p>COURSE MATERIAL</p> <p>VIDEO RECORDING</p>	 <p>Upscaling A.I. Workflows</p> <p>COURSE MATERIAL</p>	 <p>OpenFoam Workshop</p> <p>COURSE MATERIAL</p>	 <p>OpenACC Workshop</p> <p>COURSE MATERIAL</p>
 <p>Intermediate CUDA Workshop</p> <p>COURSE MATERIAL</p>	 <p>Intermediate MPI Workshop</p> <p>COURSE MATERIAL</p>	 <p>OpenMP for GPU Offloading Workshop</p> <p>COURSE MATERIAL</p>	 <p>NEK5000 Workshop</p> <p>COURSE MATERIAL</p>

Training collaborations

- Many NCCs collaborate on training
 - Example 1: NCC1 contributes instructors, NCC2 provides hardware access, NCC1+NCC2 advertise to local audiences
 - Example 2: Collaborative development of training material with multiple NCCs
- NCCs disseminate training events from other NCCs
- CASTIEL Twinning/Mentoring program often used for training

Challenges

- Not all training providers can offer HPC access for workshops
- We tend to collaborate mostly with NCCs we already know



GPU programming: why, when and how?



Search docs

THE LESSON

- Why GPUs?
- Introduction to GPU hardware
- What problems fit to GPU?
- GPU programming concepts
- GPU software environments
- Preparing code for GPU porting
- GPU programming types
- Recommendations
- Problem example: heat equation

REFERENCE

- Quick Reference
- Glossary
- Instructor's guide

Home / The hitchhiker's guide to GPU programming

Edit on GitHub

The hitchhiker's guide to GPU programming

Graphical processing units (GPUs) are the workhorse of many high performance computing (HPC) systems around the world. The number of GPU-enabled supercomputers on the [Top500](#) has been steadily increasing in recent years and this development is expected to continue. In the near future the majority of HPC computing power available to researchers and engineers is likely to be provided by GPUs or other types of accelerators. Programming GPUs and other accelerators is thus increasingly important to developers who write software which is executed on HPC systems.

However, the landscape of GPU hardware, software and programming environments is complicated. Multiple vendors compete in the high-end GPU market, each vendor provides their own software stack and development toolkits, and even beyond that there is a proliferation of tools, languages and frameworks that can be used to write code for GPUs. It can thus be difficult for individual developers and project owners to know how to navigate this landscape and select the most appropriate GPU programming framework for their projects based on the requirements of a given project and technical specifics of any existing code.

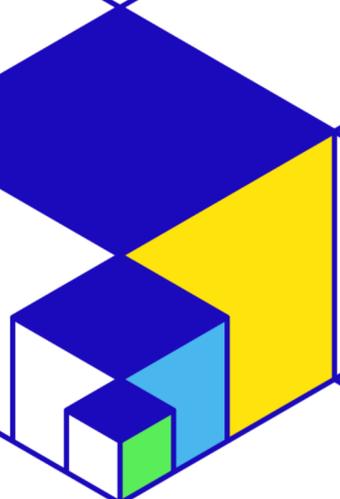
This material is meant to help both software developers and decision makers navigate the GPU programming landscape and make more informed decisions on which languages or frameworks to learn and use for their projects. Specifically, you will:

- Understand why and when to use GPUs.
- Become comfortable with key concepts in GPU programming.
- Acquire a comprehensive overview of different software frameworks, what levels they operate at, and which to use when.
- Learn the fundamentals in at least one framework to a level which will enable you to quickly become a productive GPU programmer.

In EuroCC2, we need to ...

- Consolidate all training events in single widely disseminated training platform
- Meet more often to discuss and initiate joint training efforts
- Agree on common format and collaboratively develop EuroHPC training material
- Share more instructors between NCCs
- Build a coherent and coordinated training catalogue with learning paths
- Introduce EuroCC-wide certification for HPC training
- Streamline HPC access for EuroCC training workshops





Education in HPC: A Lifelong Effort

EUMaster4HPC

The first pan-European Master Programme
for High Performance Computing

Pascal Bouvry (U Luxembourg), 2023-03-22

Objectives

Pan-European Master programme for HPC: 120 ECTS (2 year)

- ✓ Train the next generation of HPC experts in Europe
- ✓ Joint modular curriculum on core courses and specializations
- ✓ Establishment of strong collaboration between universities, HPC centres, industry and promoting students' mobility

The consortium and Partners

8 awarding Universities

1. Universitat Politècnica de Catalunya, Spain
2. Sorbonne Université, France
3. Politecnico di Milano, Italy
4. Università della Svizzera Italiana, Switzerland
5. Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
6. University of Luxembourg, Luxembourg
7. Sofia University "St. Kliment Ohridski", Bulgaria
8. KTH Royal Institute of Technology, Sweden

Contribution to:

- ✓ Body of knowledge
- ✓ New educational content, materials, teaching methods
- ✓ Joint/dual diploma
- ✓ Quality and standardization
- ✓ Implementation of pilot program

29 Non-awarding universities

- ✓ Input on the curriculum
- ✓ Joint effort on the creation of new teaching material
- ✓ Future adopters of the standard programme

16 Contributing parties (industry and research)

- ✓ Definition of industry/research needs
- ✓ Sponsorship of students/events
- ✓ Organization of student training
- ✓ Future employers

36 Additional supporting partners

- ✓ Promoting
- ✓ Lobbying

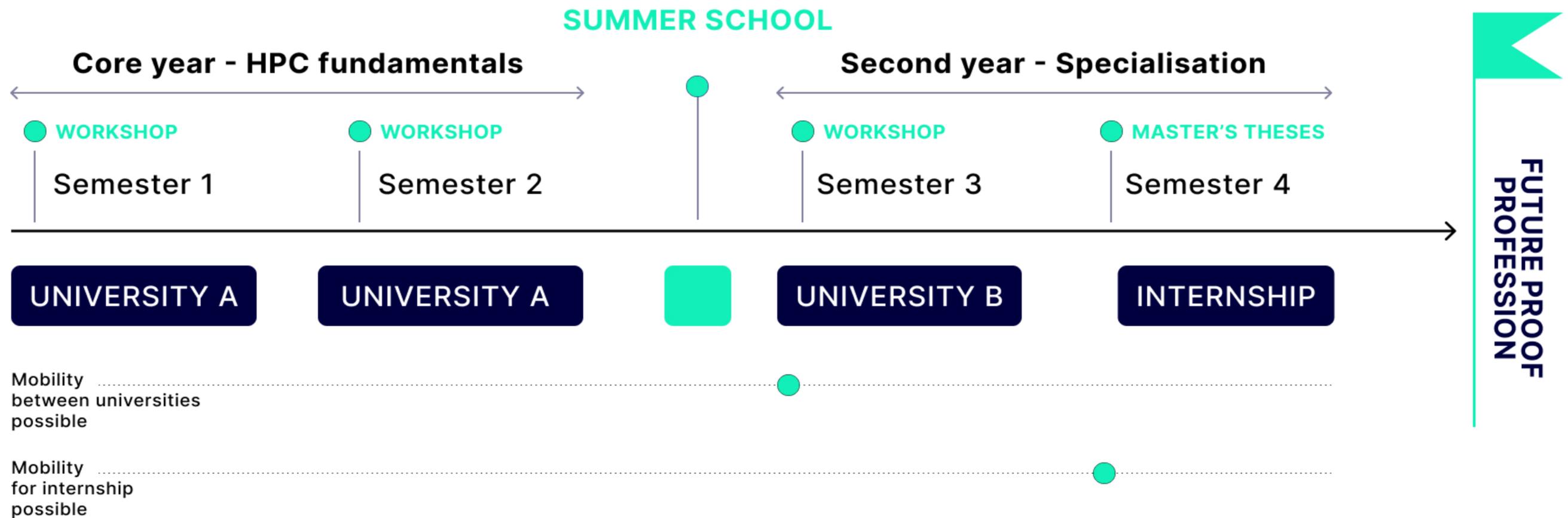
List of target Expertises

- ✓ Research Software and Domain Expertise
- ✓ Numerical and Data Analytics Expertise
- ✓ Performance Analysis and Engineering Expertise
- ✓ System Development, Operations and Support Expertise
- ✓ System Architecture Expertise

Academic Journey



ACADEMIC JOURNEY



WP2 Modules: 1st year (fundamentals)

4 topics covering first year of the EUMaster4HPC curriculum:

1. Mathematics and Statistics
2. Software Engineering
3. Parallel Programming
4. Computer Architecture

Ongoing: setting up list of suitable courses for each topic.

WP2 Modules: 2nd year (specialisation)

5 specialisations based on industrial and academic needs

1. Application Domain Expert
2. Numerical and Data Specialist for Science Domains
3. Performance Analyst and Advisor
4. System Development and Support
5. System Architect

Ongoing: setting up list of suitable courses for each topic.

WP3 Modules - MOOC

TARGET USE CASES

- ✓ Pre-course (entire course topic)
- ✓ Pre-class (a specific topic)
- ✓ In-class (replacing a specific topic lecture)
- ✓ Post-class (additional specific topic lecture)
- ✓ Independent learning (pre-requisite for some courses)

Planned credits:

- ✓ 1 ECTS
- ✓ 2 ECTS

Expected hours of video:

- ✓ 1 ECTS = 3-5 hours of video on theory
- ✓ 2 ECTS = 6-10 hours of vide

ADDITIONAL MATERIAL PROVIDED WITH THE MODULES

- ✓ Slides
- ✓ Notes
- ✓ Open/Multiple-choice questions for self-assessment
- ✓ Evaluation material for peer review
- ✓ Other online evaluation material
- ✓ Lab exercises

Planned timeline:

- ✓ Call close: May 1
- ✓ Proposals selected: May 15
- ✓ MOOCs published: November 2023

WP6: pan-European education in HPC

Common roadmap and recommendations for the pan-European curriculum in High Performance Computing (HPC)

- ✓ Collect and analyse lessons learned from each WP and deliver a set of recommendations
- ✓ Promote and support the implementation of the updated curriculum for a European master in HPC within the interested participating universities
- ✓ Explore and design bridges between this EUMaster4HPC and HPC-courses in various HPC related Master programs
- ✓ Map the network created in this pilot
- ✓ List possible new candidates and partners for a further collaboration

Education in HPC: A Lifelong Effort

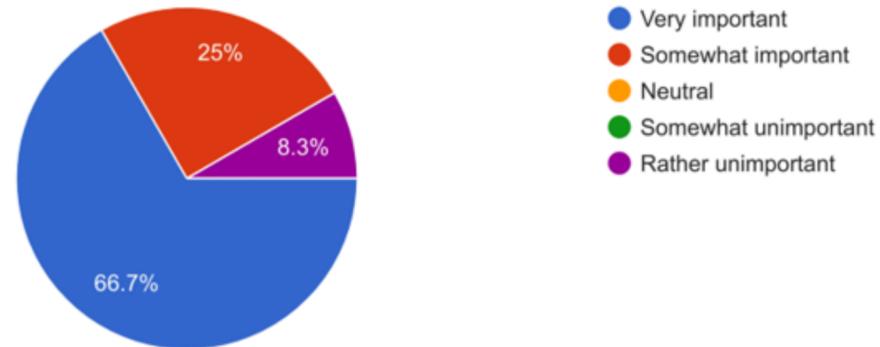
Needs for Education: Academia

Towards an European Curriculum for a Master in HPC

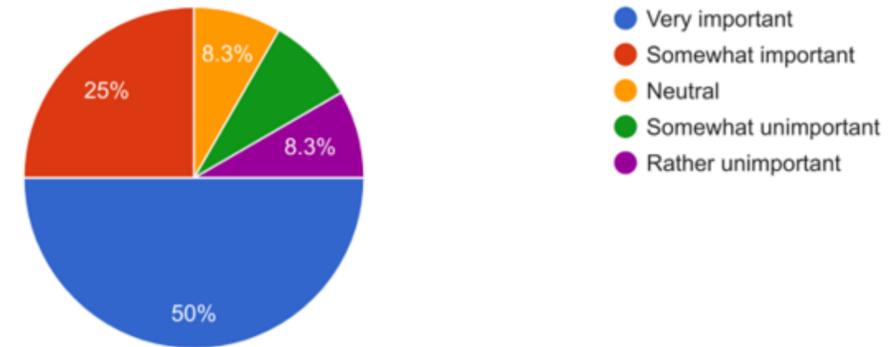
Maria Ribera Sancho (UPC), 2023-03-22

Requirements and Methodology

JOB PROFILE #1 / SKILL #1: HPC applications execution, debugging, profiling and performance tuning
12 responses



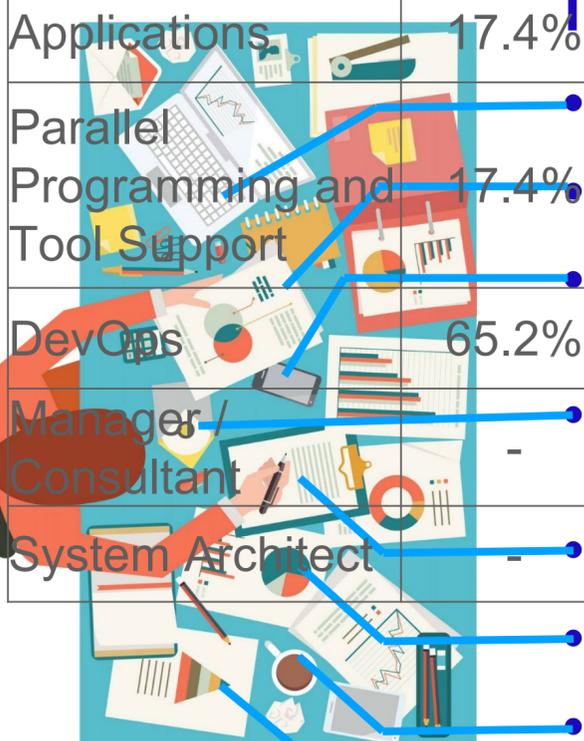
JOB PROFILE #1 / SKILL #2: Code development for massively parallel CPU-based computer systems using Fortran, C/C++, MPI, OpenMP
12 responses



ion
chec
cket

Profile	HPCWire	HiPEAC Jobs	BSCJobs	AVG
Applications	17.4%	13.4%	58.8%	29.8%
Parallel Programming and Tool Support	17.4%	17.0%	19.7%	18.0%
DevOps	65.2%	23.2%	5.4%	31.3%
Manager/Consultant	-	1.1%	1.1%	0.4%
System Architect	-	4.4%	15.2%	20.5%

Methodology



- Collect offers for HPC-related jobs in the academic and commercial sector
- Analyse offers to extract job profiles and related sets of skills
- Use analysis results to design a questionnaire for an online survey as well as structured interviews
- Compile results from stakeholders' feedback to establish a set of necessary skills
- Proposal for Fundamentals
- Proposal for Specializations
- Initial Gap Analysis
- Proposal for Transversal Skills

Profile (profile count)	Subarea	Job Names	Hard Skills
ARCHITECT (52/112)	Computing systems (Architecture, Microarchitecture, Compilation, firmware), memory subsystem	PostDoc, Senior Architect, RTL Engineer, FPGA Developer, Computer Architect, System Designer, IC Designer	<p>Fundamentals</p> <p>OS/Compiler interaction</p> <p>Hardware/Software co-design</p> <p>Microarchitecture, Multiprocessor/Multi-threading architecture</p> <p>Memory systems including storage</p> <p>Interconnection networks (on-chip, off-chip, cluster-level)</p> <p>Etc.</p> <p>Specific tools</p> <p>Modeling/Simulation/Tools (at the different levels: ISA, cycle-accurate,...)</p> <p>Benchmarking/Profiling, FPGA prototyping</p> <p>EDA tools (HLS, verification), RTL programming</p>
HiPEAC Jobs (32/112)	Accelerators (e.g. AI/ML, Task-Based parallelism, Financial applications)		
BSCJobs (67/442)	Hardware security		

Initial Results: Job Profiles

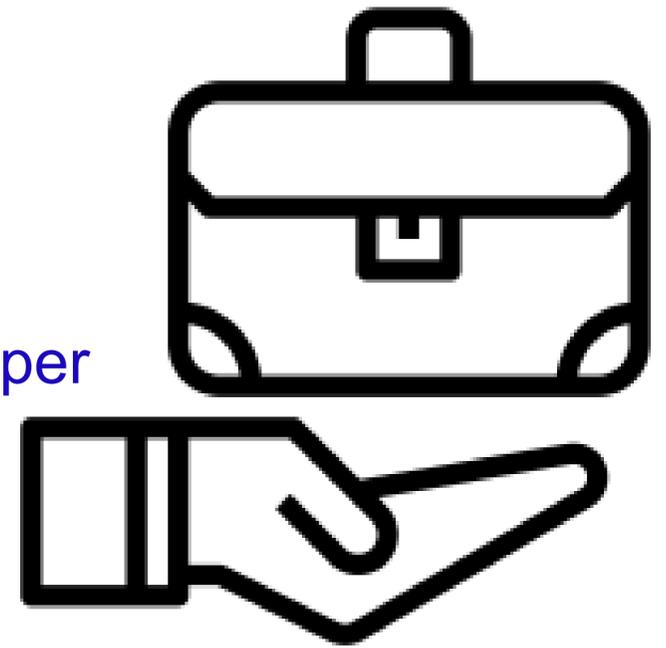
Academic sector

- Analysis of ~600 job offers
- Job profiles:
 - Application / Domain Expert
 - Parallel Programming and Tools Support/Solution Designer
 - DevOps (System Support and Development)
 - System Architect



Commercial sector

- Analysis of ~30 job offers
- Job profiles:
 - Computational scientist
 - Hardware developer
 - System software developer
 - HPC architect
 - System administrator
 - Application software developer



Identification of Specializations



Application Domain Expert

- Specialist on a science domain, getting the skills on parallel programming, and computer architecture, to participate in application development in interdisciplinary teams from the perspective of the Science Domain

Numerical and Data Analyst for Science Domains IT4I, VSB

- Focus on the software engineering techniques for the development of algorithms for supercomputers, parallel programming, computer architecture, knowledge on Application Domains, works tightly with Domain Scientists, and Application Domain Experts

Performance Analyst and Advisor

- Specialist on performance analysis and tuning for supercomputers. Focus on algorithm analysis, knowledge on performance analysis tools and models, get the best of heterogeneous systems. Use of containers, virtualization

System Development and Support

- Development of the software for supercomputers. Focus on operating system, programming models, compiler, performance analysis tools, middleware, virtualization. Knowledge on the low level characteristics of the hardware, synchronization techniques, communications, from chip to data center

System Architect

- Design and development of the supercomputers. Focus on processor, multiprocessor, supercomputer architectures. Memory and I/O systems, networking, circuit design, verification and test, low power techniques, fabrication

Current proposal for Fundamentals: Main topics

Mathematics and Statistics



Parallel Programming:

- Shared-Memory, Middleware Technologies (currently OpenMP), Parallel Programming and Constructs, Vectorization/SIMD, GPU Computing, Embodied
- Software Engineering
- Application Design, Methods and Tools
- Component Integration
- Memory Hierarchy
- Software Releases



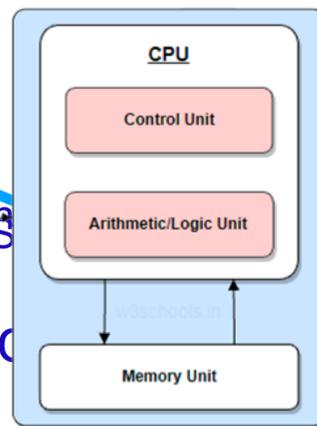
Performance Analysis and Models:

- Metrics, Scaling, Amdahl's and Gustafson's Laws, Roofline Model, Energy-Efficiency, Advanced Mathematical Models and Algorithms



Advanced Parallel Programming: Devices:

- HPC Applications Execution (Intra Node Work Distribution, I/O, I/O Buses (currently USB, PCIe), Storage, Disks, S, Node Work Distribution), Interoperability, HPC Infrastructure Resource Management



Computer Architecture

- Programming Paradigms, Parallelism (Data Centric), Computing, Parallelism (currently GPUs, FPGAs, CGRAs), Instruction Set Architecture (ISA), Debugging
- Processor Architecture
- Functional Analysis
- Multicore, Manycore, Multithreading, Memory Hierarchy, NUMA, I/O, Clusters, Cloud
- Data Centric Computing (Data Life Cycles)

Initial Gap Analysis

Disciplines	Uni.Lu	Sofia	POLIMI	FAU	Sorbonne	KTH	USI	UPC
Applied Math			√	√	√	√	√	
Applications			√	√	√	√	√	
Parallel/Het. Prog.	√	√	√	√	√	√		√
Distributed/Cloud	√	√	√	√		√	√	√
DevOps		√						√
Big Data/HPDA	√	√			√		√	√
AI	√	√	√			√	√	√
Comp. Architecture					√			√
Transversal skills	√							√

POLIMI

- Starting a new Master this current year
- Increased strength towards Computer Architecture
- Including now new topics: Quantum Computing

FAU

- Current master program
- Good coverage of the Fundamentals
- Additional strength on Applications

Sorbonne

- Current master program
- Good coverage of the Fundamentals
- Additional strength on Mathematics

UPC

- Current Master Program
- Increased strength towards Computer Architecture
- Including now new topics: CMOS Technology and Verification

Current State

Fundamentals:

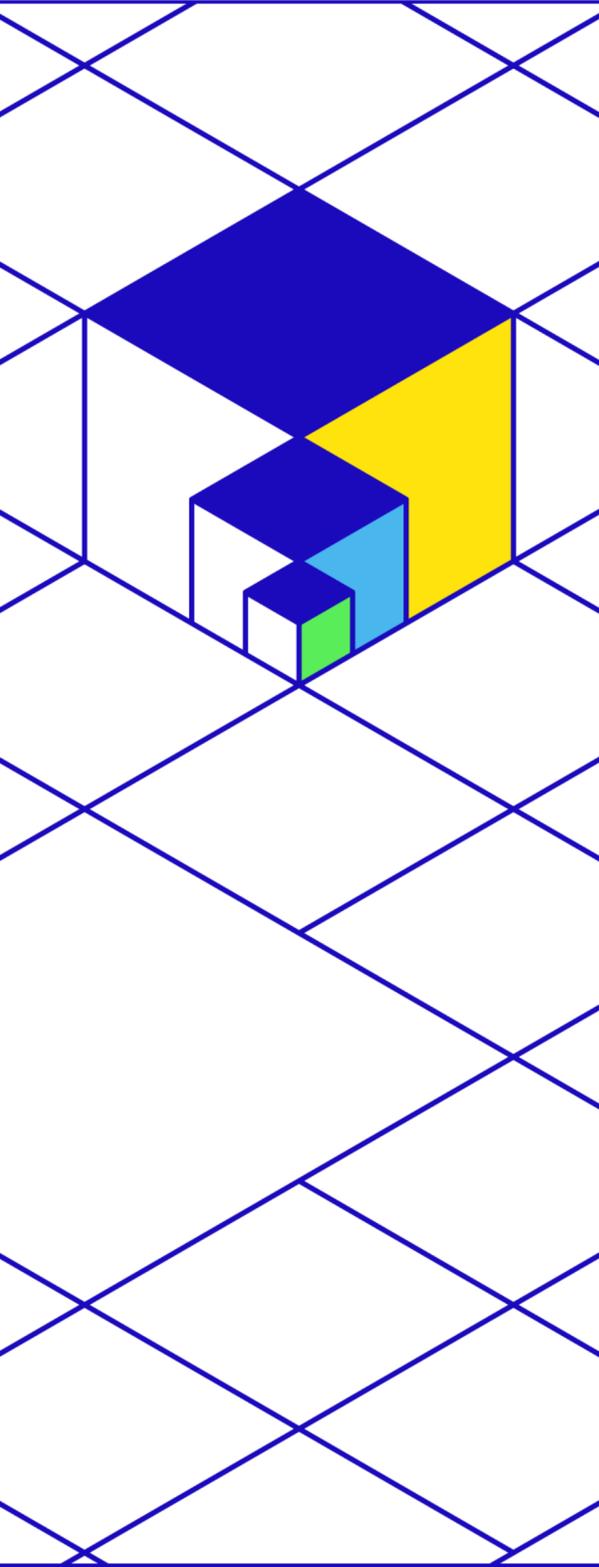
- Detailed definition of skills
- Learning outcomes, modules and content definition in progress

Specializations:

- Initial definition of skills
- Working groups for skills detailed definition established and progressing

Transversal Skills

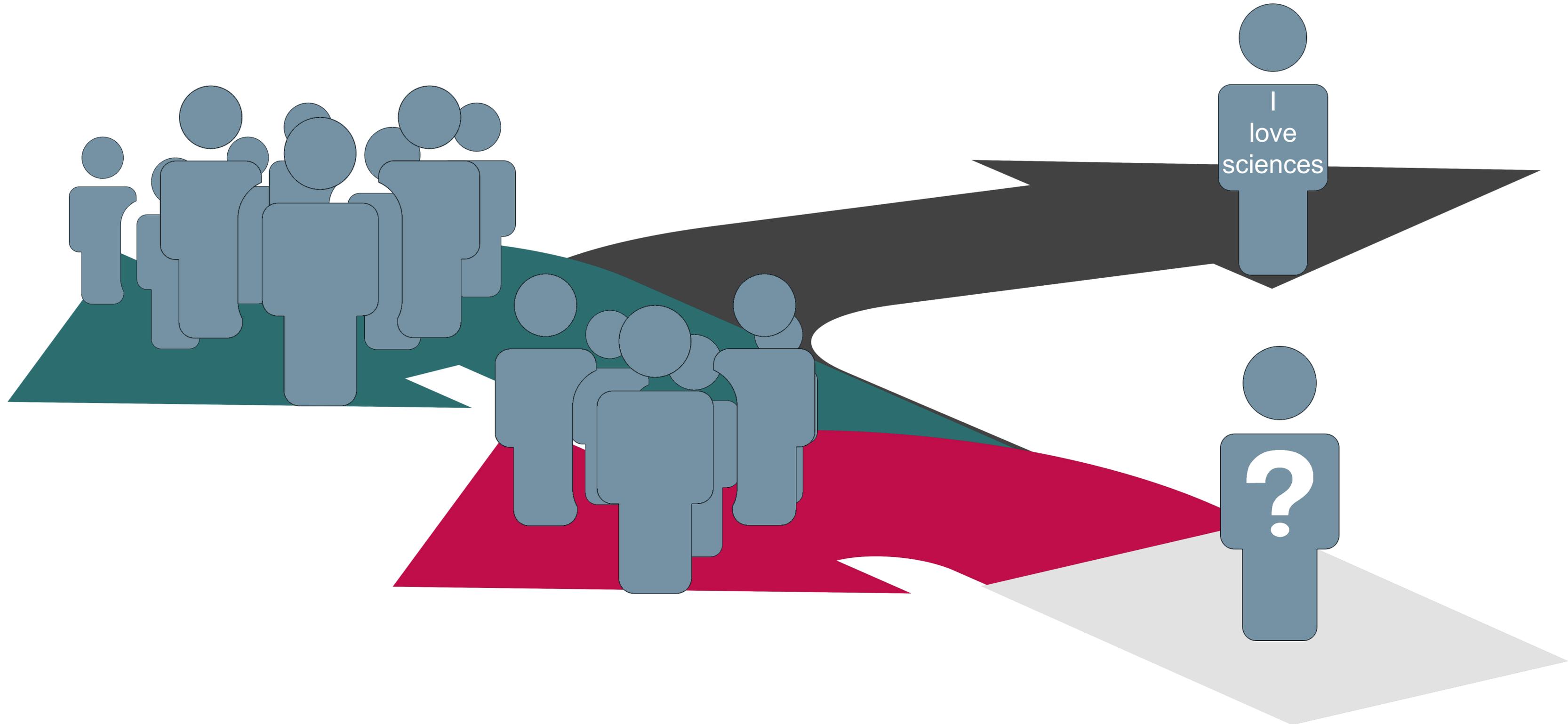
- Detailed definition of main skills to address
- Initial work on how to address them in the CV



Education in HPC: A Lifelong Effort

Needs for Education: Industry

Make scientific education great again



Gen Z motto: You Only Live Once

Leave an impact on the world

Align job and interests

Entrepreneur

Tech-savvy

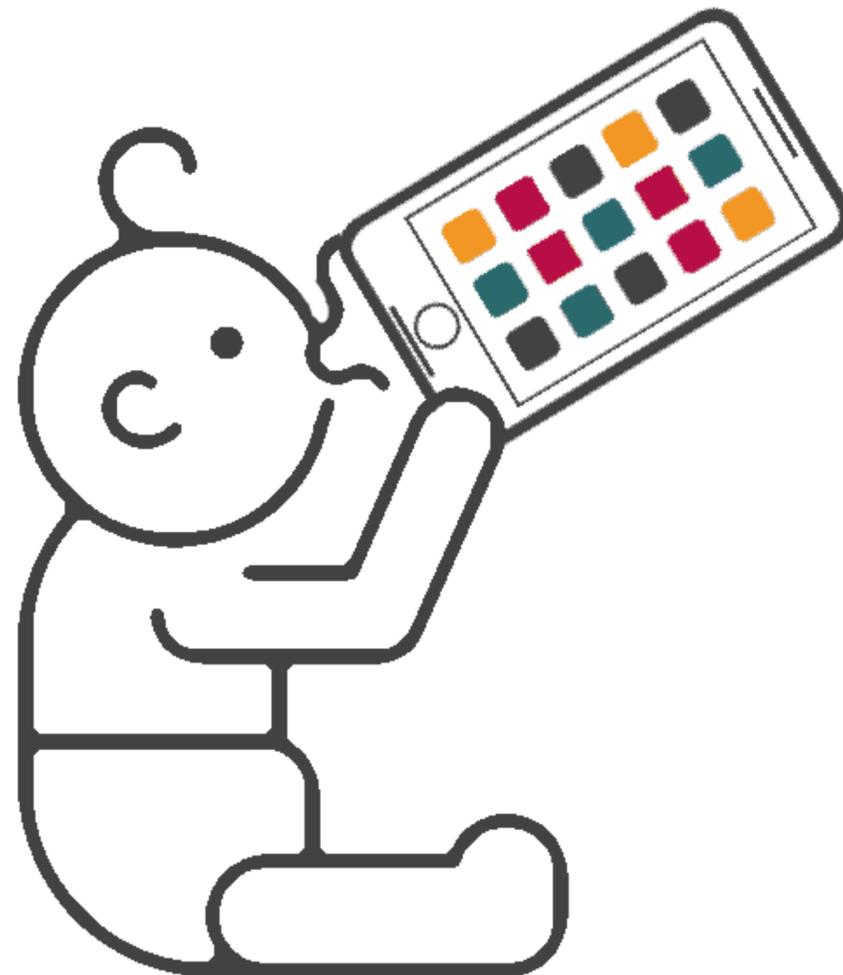
Multi-taskers

Few focus

work-life balance

Freedom

Digital interactive



HPC, a key asset for the future of Humanity

AI

Digital
Twins

Data
Analytics

Energy

Climate

E-Health

HPC at
The Edge

Innovation



Be ready to get hired for your 2nd job

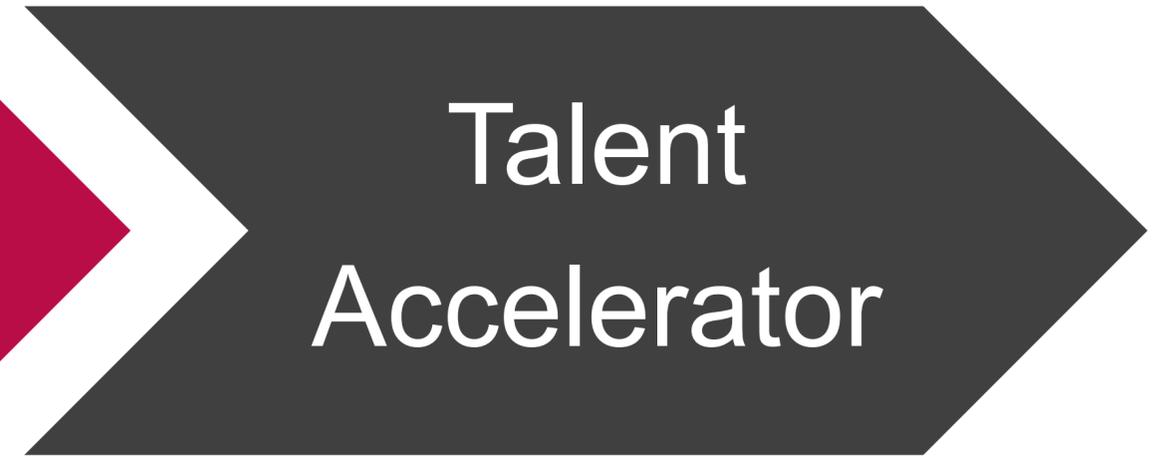
High School



Bachelor - Master



Professional life



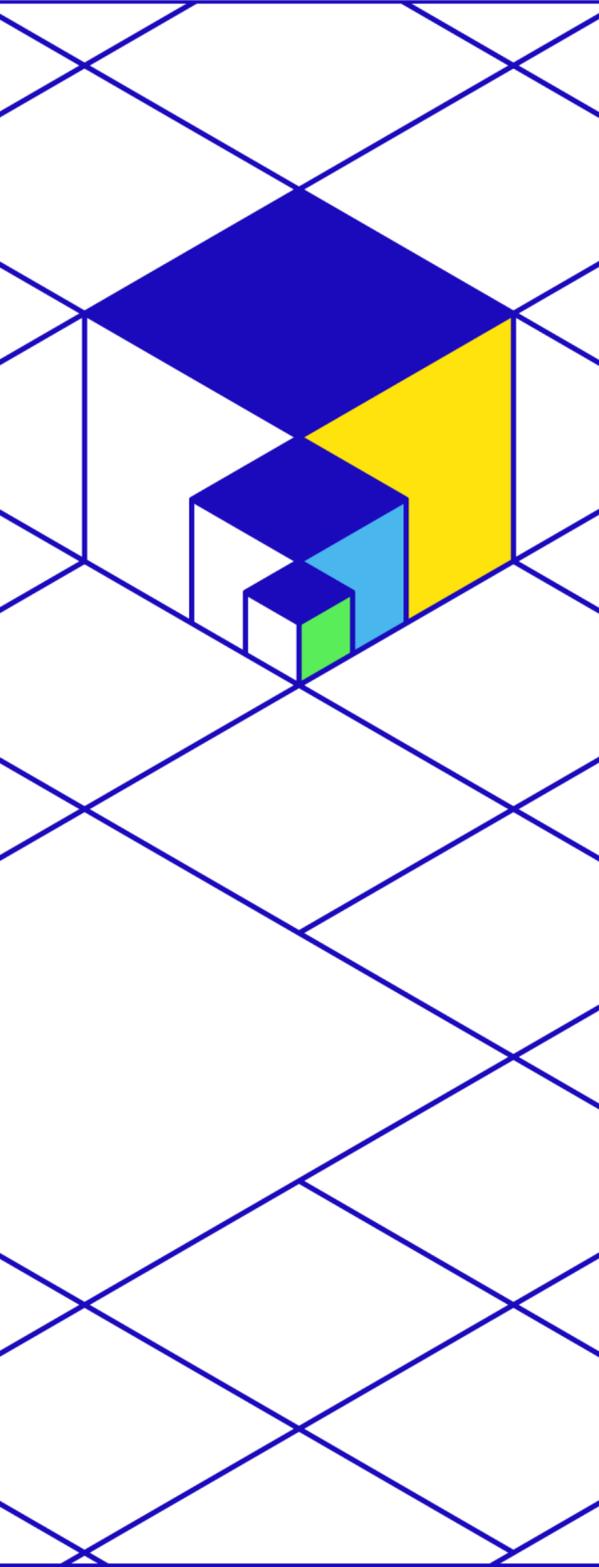
Empower **yourself** with the right basic skills **for** sustainable **innovation**

Take a step ahead
Be above competition

- Ecosystem & economy
- Math background
- Fluent in HW & SW concepts
- Master solvers and AI principles

- Customer value
- Production quality
- Curiosity - Self-learning
- Cooperation – Autonomy
- Leadership
- Problem solving

- Extend your expertise
 - Sell your ideas
- Develop your strengths



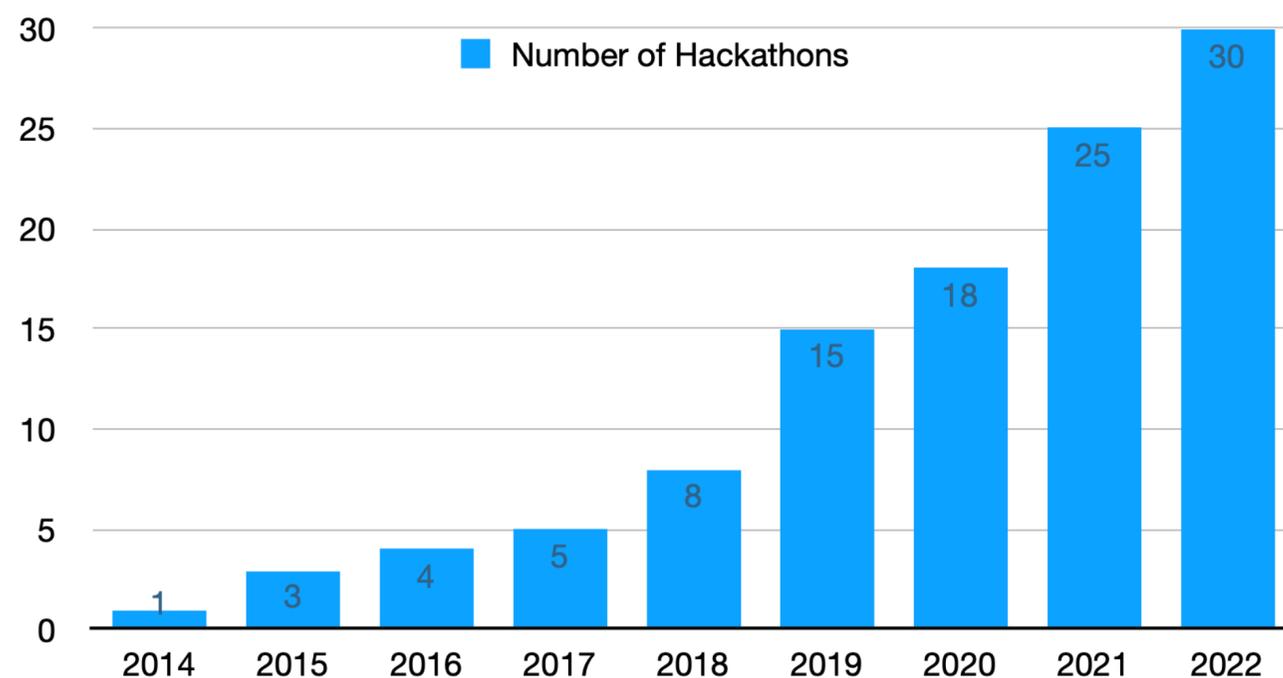
Education in HPC: A Lifelong Effort

Practices: GPU Hackathons



History

- Since 2014: Start at ORNL
- 2016: German Hackathon hosted by Dresden/Jülich
- 2022: >30 Hackathons world-wide
- Organization: ORNL → OpenACC → Open Hackathons, with local support





EuroHPC Summit

2023 Göteborg

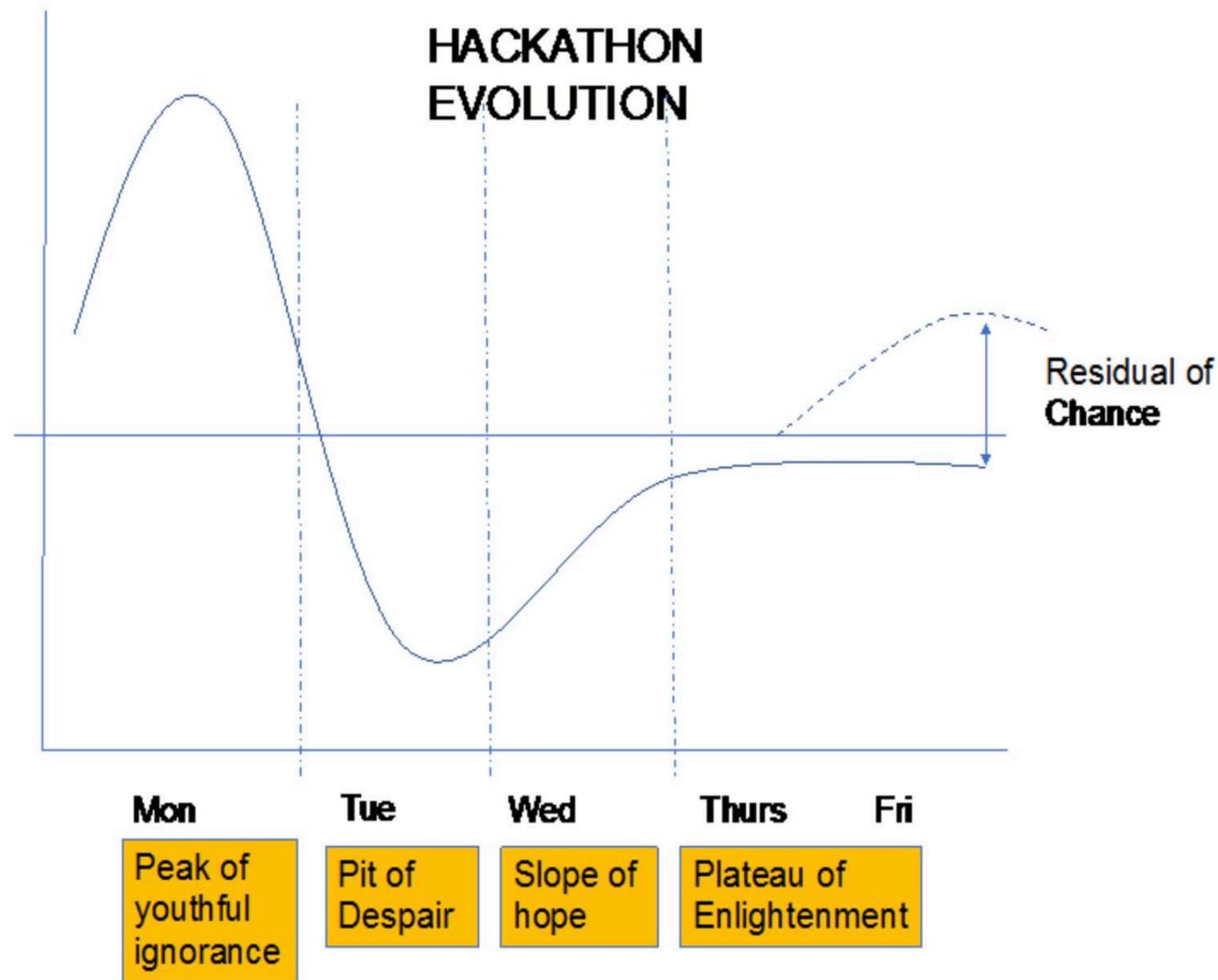
Key Aspects

- 3+1 **intensive days** of nonstop *hacking*
- Teams (min. 3 members) with **own application**, goals; **reviewed**
- 2 **expert mentors** per team (centres, vendors, universities)
- Regular **status updates** (short/long)
- On **HPC infrastructure**
- Possibly how-tos, lectures
- One common **room** + breakout

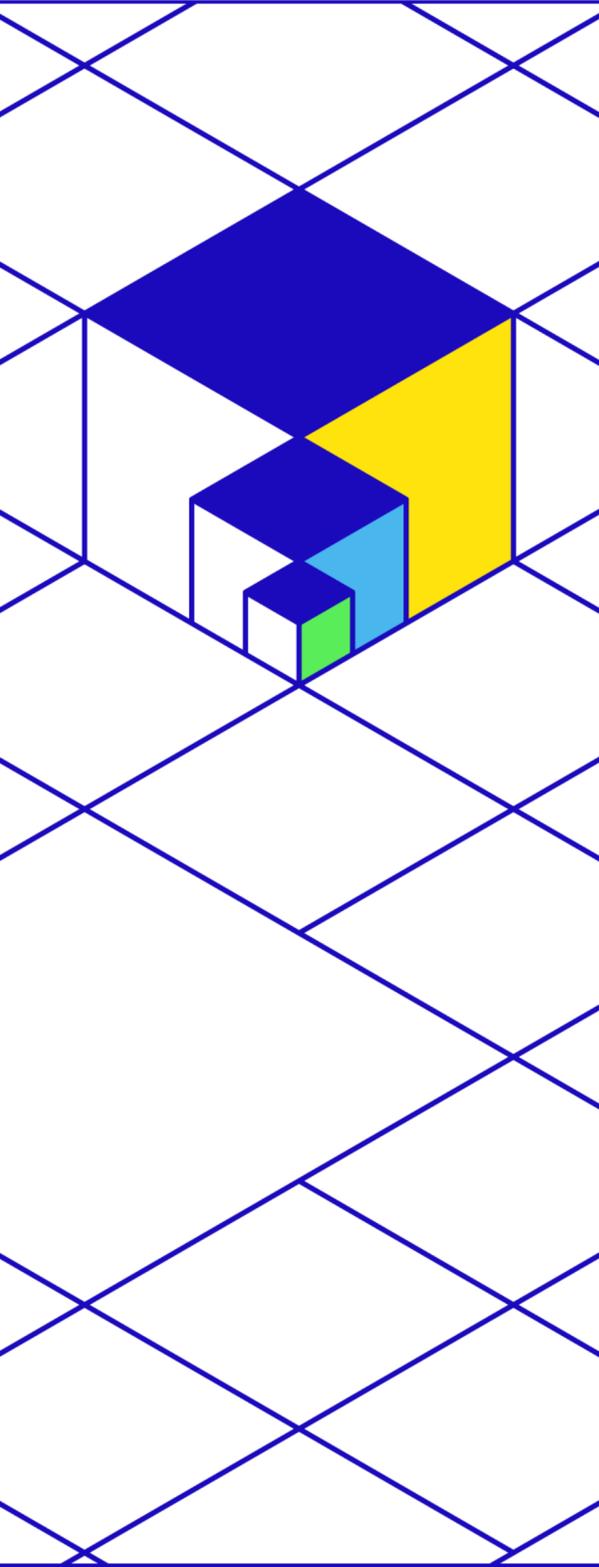
Goals/Results

- **Jump-start** for fresh GPU projects
- Apply best GPU **coding practices** to own code
- Obtain skills for **performance analysis**, apply to own code
- Improve **soft skills** (team work, ...)
- **Sustainable** impact on own application

More: go.fzj.de/hack1, go.fzj.de/hack2



Hackathon Stages



Education in HPC: A Lifelong Effort

Practices: CodeRefinery

CodeRefinery

coderefinery.org

- We teach tools and practices for the development of reproducible and reusable research software
- Team and project: Started in 2016 as a Nordic collaboration and NeIC project. Is now in phase 3 until 2025.
- Partners: NeIC (1 FTE), Aalto University, CSC; Sigma2/NRIS, USIT, UiO; ENCCS, NAISS, PDC, UPPMAX; DeIC, T1C for interactive HPC
- Workshops, since 2016
 - 31 full workshops
 - 25 other workshops/events
 - 3 instructor training events
 - Over 2000 persons trained
 - Last four workshops: 8 volunteer instructors, 80 volunteer helpers



CodeRefinery: bridging the gap

Specialist training

Traditionally run by computing centers
CodeRefinery provides collaboration network

CodeRefinery

Expert training for reusable software
Not broadly taught in all degree programs that need it

The Carpentries

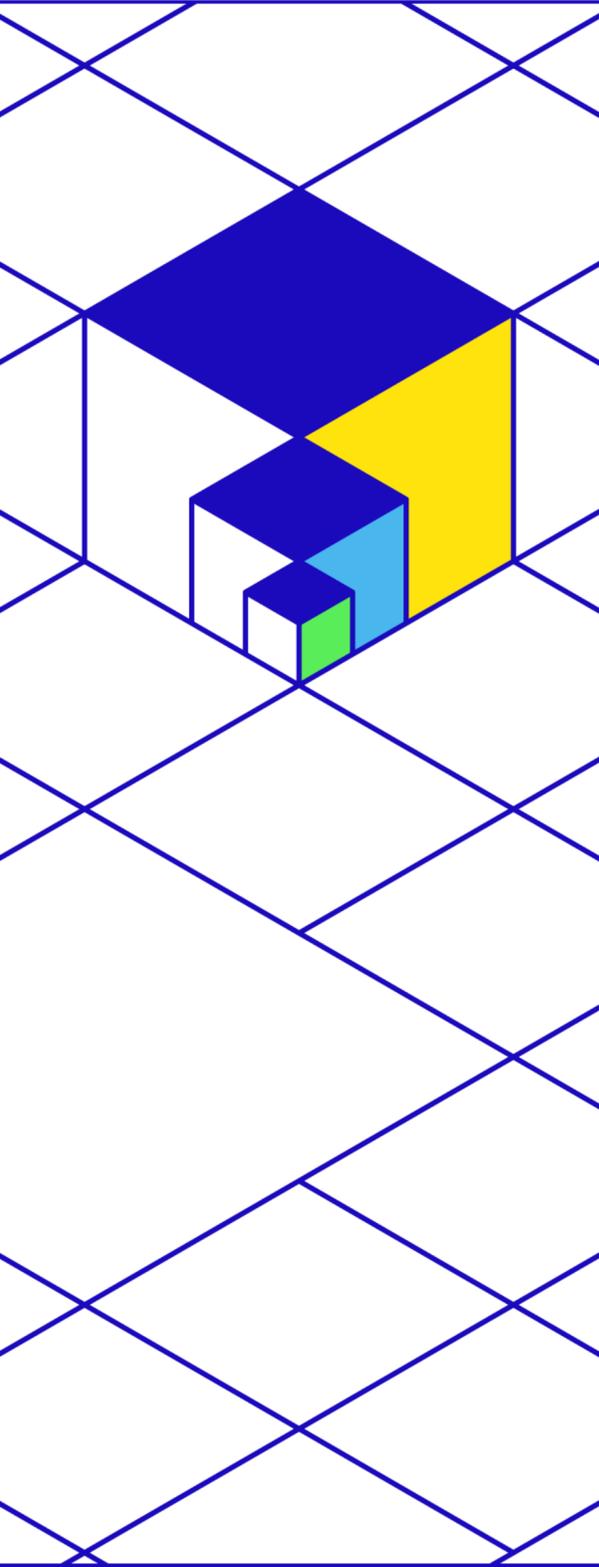
Basics training in programming
and data science for novices

Typical curriculum

- Basic and collaborative Git 3 days on-site (next: Uppsala May 2023)
- Code review 6 half-days online (ongoing: March 2023)
- Social coding and open software
- Reproducible research
- Jupyter notebooks
- Code documentation All material is open and CC-BY
- Automated testing <https://coderefinery.org/lessons>
- Modular code development <https://github.com/coderefinery>

Future

- Establish a CodeRefinery non-profit organization
- Continue the fruitful collaboration with partners/members inside and outside the Nordic countries
- Challenge: Coordination effort starts to outweigh lesson preparation and teaching
- The importance of communicating value to organizations: for learners and for staff participating in hosting and teaching workshops



Education in HPC: A Lifelong Effort

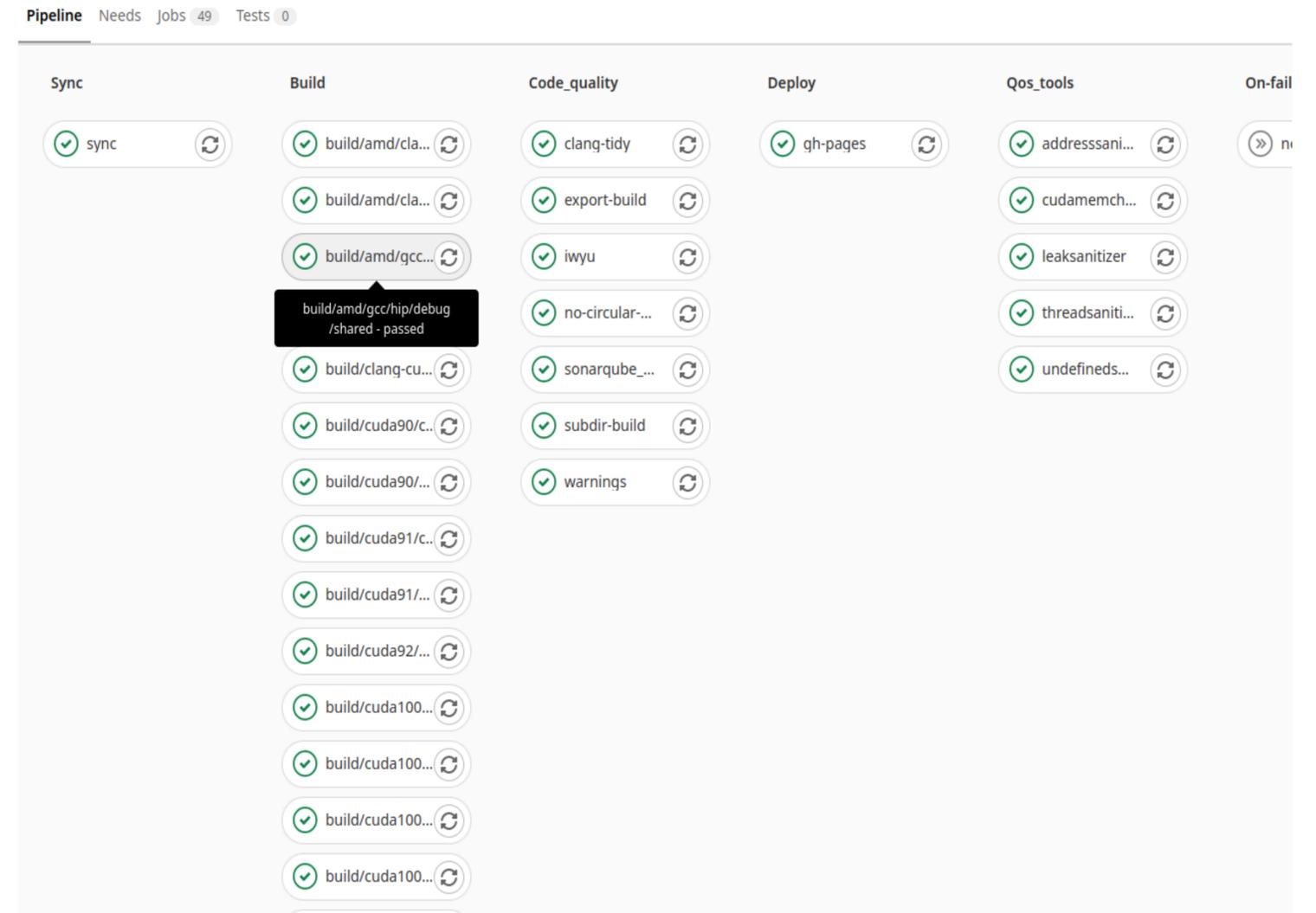
Practices: Teaching sustainable software development practices

Why ?

- Almost all research employs some form of software.
- Software lifecycle often exceeds hardware lifecycles.
- Good sustainable software is **THE** key component of computational science.
- Ingraining good software practices in students is important to their careers in industry and in academia.

What ?

- Version control.
- Continuous Integration.
- Automated testing
- Collaborative peer review



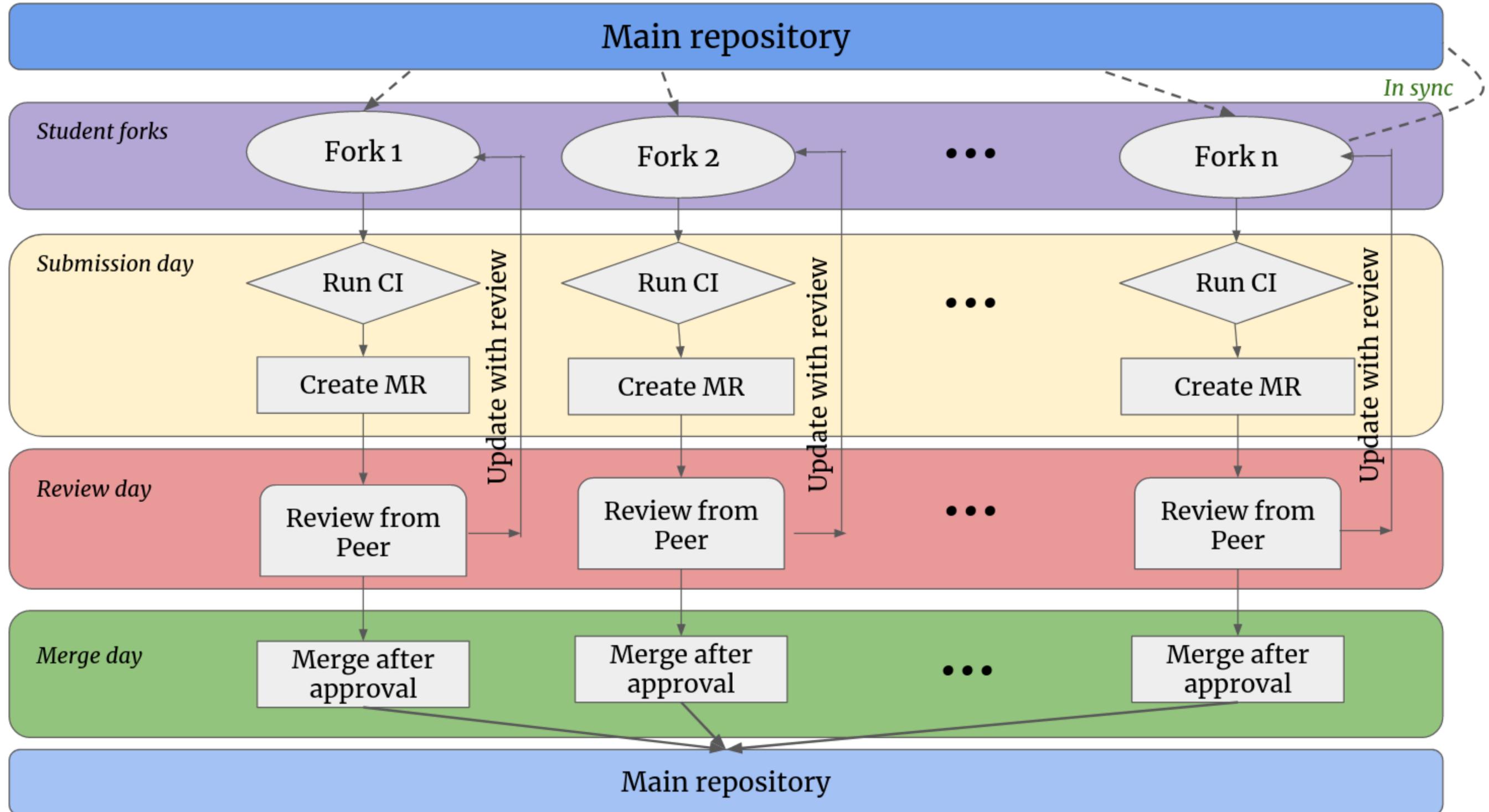
How?

- Incorporate these common practices into courses.
- Encourage students to experiment with algorithms and implementations and not be bogged down in build and platform issues.
- Review each other's codes, and encourage critical and constructive feedback.
- Improve your code from the code review.

How?

- Create a common Exercise framework for all to work on.
- Provide the building blocks: Compilation, testing and benchmarking frameworks and setup a Continuous Integration setup to automatically test the code on push.

<https://github.com/pratikvn/nla4hpc-exercises-framework>



An example

The screenshot shows a GitHub pull request interface. At the top, it indicates that a thread was started by a reviewer 2 weeks ago and last updated by an assignee 1 week ago. The code diff shows a CUDA kernel for a warp reduction operation. The code is as follows:

```
154 +   __shared__ ValueType tmp_res[default_block_size];
155 +   tmp_res[local_id] = tmp;
156 +
157 +   // do reduce operation on tmp_res until warp size is reached
158 +   for (int k = nt / 2; k > warp_size; k /= 2) {
159 +       __syncthreads();
160 +       if (local_id < k) {
161 +           tmp_res[local_id] += tmp_res[local_id + k];
162 +       }
163 +   }
164 +
165 +   // for last warp_sized entries use optimized __shfl() for reduce
166 +   tmp = tmp_res[local_id];
167 +   #pragma unroll
168 +   for (int k = warp_size / 2; k > 0; k /= 2) {
169 +       tmp += __shfl_xor_sync(0xffffffff, tmp, k);
```

Below the code, there is a comment thread. The first comment, from the reviewer 2 weeks ago, asks: "Why did you use xor and do you plan to try out also other options, or is there a reason why this option is the best?". The second comment, from the assignee 1 week ago, responds: "There are other warp shuffle operations. See the doc in the section about warp shuffles: (<https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html#warp-shuffle-functions>)". The assignee continues: "I chose the xor function because it was used in the lecture slides :) it is also the example used in the doc. But you could perform the warp reduction using the other options as well. As shfl-operation is performed on all threads in the warp simultaneously, the performance just depends on the number of operations you need. Here, that is 5. You could probably get the result using \"__shfl_down_sync()\" in the same number of operations with the only difference that with xor the result will be in each entry of tmp_res and with down only in tmp_res[0].". The assignee concludes with: "Hope that helps :) If not just ask again..".

On the right side of the interface, there are several settings and information:

- 0 Assignees (None - assign yourself)
- Reviewer: Reviewer
- Milestone: None
- Time tracking: No estimate or time spent
- Labels: None
- Lock merge request: Unlocked
- 3 participants
- Notifications: On
- Reference: nla4hpc/spring-2021...
- Source branch: pr_hw2_build

An example

- Assign HW reviews to peers.
- Merge once CI has passed and has been approved.
- Grading based on merged HW

The screenshot displays a GitHub pull request for the repository 'NLA4HPC'. The pull request title is 'Add HW1 of [redacted] (Review by [redacted])'. It is in the 'Merged' state, created 1 month ago by the user [redacted]. The interface shows the following workflow:

- Request to merge:** A request to merge the branch '[redacted]:pr_hw1_build' into the 'master' branch.
- CI Pipeline:** Pipeline #304098110 passed for commit 1e555e22 on the '[redacted]:pr_hw1_build' branch 4 weeks ago.
- Approval:** The merge request was approved by a reviewer.
- Merge:** The pull request was merged by Pratik Nayak 4 weeks ago. The changes were merged into the 'master' branch with commit 7686da6e. A 'Delete source branch' button is available.
- Final CI Pipeline:** Pipeline #304101780 passed for commit 7686da6e on the 'master' branch 4 weeks ago.

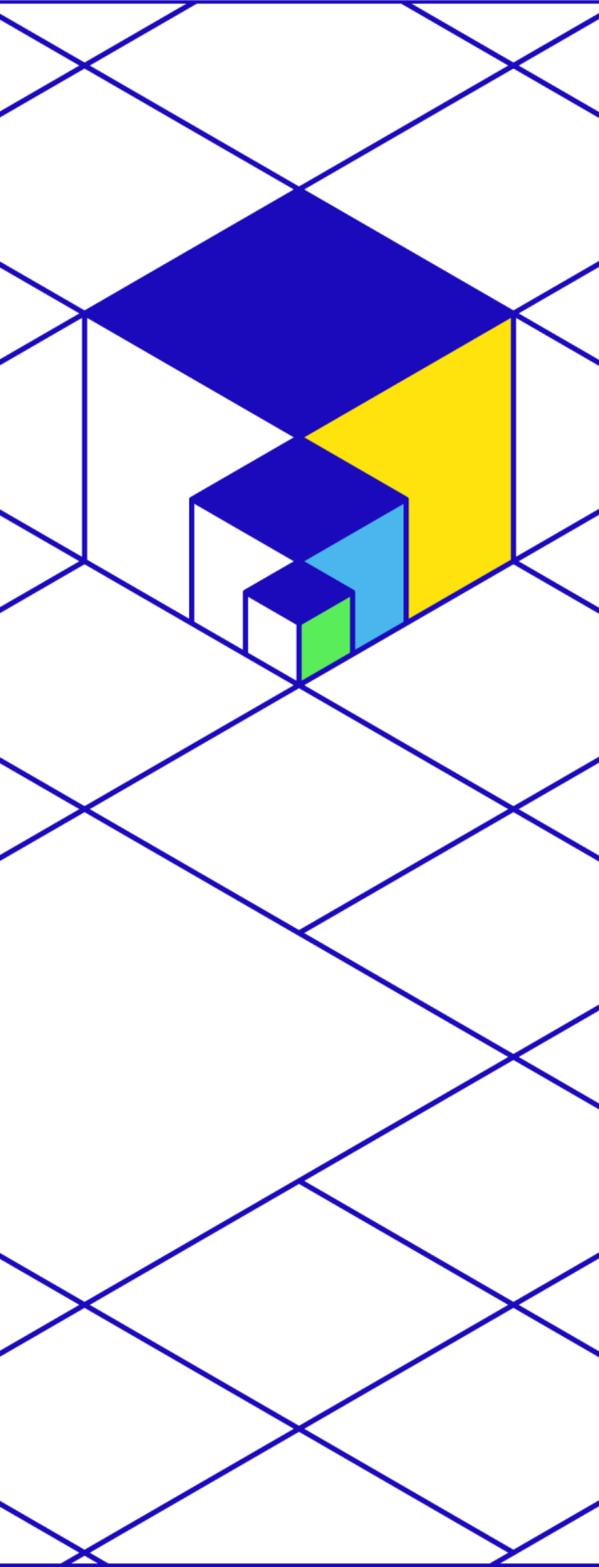
At the bottom, the commit history shows that Pratik Nayak changed the title and assigned the pull request to the reviewer [redacted] 1 month ago. A thread from the reviewer [redacted] is also visible, stating they started a thread on an old version of the diff 1 month ago, which was resolved by [redacted] 4 weeks ago.

Student Feedback

Question	Avg rating (1–5)
How easy was it to use the framework?	2
How useful did you find the exercises instructions?	2
How easy was it to compile and run the code as provided?	2.3
How useful was the code review from your peer?	1.6
How easy was the reviewing process?	3.6
Would you like to see this type of frameworks in other courses?	1

Conclusions

- We saw a marked improvement in code quality as the course progressed, which was not the case in our previous course offerings.
- This approach is scalable.
 - It can be almost completely automated.
- The students were able to focus on algorithms and optimizations rather than on build system and other orthogonal issues.
- It encourages students to showcase their code and makes them comfortable with contributing to open-source projects.



Education in HPC: A Lifelong Effort

Discussion and Q&A