AN INTERNATIONAL OVERVIEW OF HPC TRAINING PROGRAMS

I. Introduction

The objective of the present data collection and summary was to map training data from some European institutions that have considerable HPC training experience, and to identify the key features of the training programs organised in the past two years and scheduled for 2021. In the light of these data, a picture of the practices used by HPC centres similar to us and of the most important characteristics of such trainings emerges.

This knowledge could provide help and a starting point for establishing the training concept of the HPC-CC Coordination Office.

The table below shows the following data from the relevant institutions:
Name, title of the course, website/link, description of the course, form of training, date, duration, target groups, prerequisites, number of participants, prices, availability of the table of contents/schedule, downloadable training materials, course levels, label/category, others.

II. Characteristics of the institutions

1. PRACE

PRACE (Partnership for Advanced Computing in Europe) is a joint European supercomputer infrastructure and cooperation program established with support from the European Union. PRACE is involved in a wide range of education and training activities in the form of seasonal schools, workshops, and scientific and industrial seminars to promote the efficient use of the HPC infrastructure. While seasonal schools target a wide HPC audience, workshops focus on individual technologies, tools, scientific disciplines, or fields of research. PRACE trainings are free of charge. All trainings are publicised through the PRACE Training Portal, and training materials are also made available here.

In the case of PRACE, data from the seasonal schools and summer practice opportunities organised in 2020 and 2019 were assessed. PRACE trainings held by the 7 designated institutions were assigned to the institutions themselves both in the database and in the present summary.

The training activities consist of the following parts:

1. PTC (PRACE Training Center): PRACE operates 14 PRACE Training Centers offering high-level education and training opportunities for European academic users and researchers, and European industrial users. In the table, these PTC trainings are mentioned under the relevant institutions.

2. PRACE seasonal schools: PRACE seasonal schools complement the PTC training program with three such events usually held throughout the year. Seasonal schools are held at different geographical locations and usually in a non-PTC country.

3. Summer of HPC: This PRACE program offers summer practice opportunities in HPC centres across Europe for university students in their final year and/or MSc students. A maximum of twenty applicants are selected from Europe, and applications can usually be submitted in January every year. Participants spend two months on projects related to various technical or industrial fields.

4. International HPC Summer School on Challenges in Computational Sciences: The objective of this school is to collect and bring together the best computer science students actively interested in HPC and seeking to acquire a deeper understanding thereof from the countries that have joined the program. Students receive help for this from the trainers of the HPC centres and other supercomputer experts.

2. IT4I

IT4Innovations National Supercomputing Center – the leading research, development and innovation centre of the Technical University of Ostrava – is involved in a wide range of training activities. Every year, approximately 10 courses, workshops, and conferences focussed on HPC, HPDA, and AI are offered. In the field of deep learning for computer vision, IT4Innovations was awarded the NDVIA Deep Learning certificate, as well as the PRACE Training Center (PTC) status as a member of PRACE in 2017. The courses are led not only by in-house experts from IT4Innovations but also by lecturers from leading foreign institutions.

- **Number and form of the training programs:** During the past 2 years, 17 trainings were completed; these included 3 online courses, and the rest were conventional ones. As regards the trainings scheduled for 2021, no information is currently available.

  Duration and language: Typically, the courses last 1 or 2 days, and were held in English except for one program.

  Content levels: The indicated level of difficulty was basic for 7 courses, basic to intermediate for 5 courses, advanced for 1 course, and no such data was available for 4 courses.

- **Number of participants:** For 12 courses, it was maximised to 20 to 30.

- **Prices:** All listed training programs are free of charge.

- As regards the **target groups**, typically students, employees, and researchers from scientific institutions are invited to the trainings, while industrial players may inquire about opportunities to participate via email.

- **Downloadable training materials:** Training materials are uploaded for 58% of the courses.

- **Topics, labels:** Machine Learning / Deep Learning: 6; Scientific Visualization: 5; Programming Languages (C++): 2

3. HLRS

During the past 20 years, the High Performance Computing Center Stuttgart (HLRS) in Germany has gained leading expertise in the support and education of end users while focussing on engineering activities. Among others, HLRS offers training services in the following fields: Parallel Programming, numerical Methods for HPC, Visualization, Grid and Cloud Concepts, Big Data, Machine Learning, and Artificial Intelligence.

- **Number and form of the training programs:** During the past 2 years, 51 trainings were held; as regards 2021, the data for 17 scheduled courses are currently available. Due to the coronavirus-related restrictions, the courses went online from April 2020; before that, only conventional trainings were held.

  Duration and language: As regards the duration of the trainings, most (65%) of them are longer ones lasting 4 to 5 days, which are followed in frequency by 1- to 2-day courses (22%).
  28% of the trainings are in German but the slides are (also) provided in English in these cases.

  Content levels: As regards the level of difficulty, 37% and 20% of the courses are classified as advanced and intermediate, respectively, and the rest is of mixed difficulty (beginner/intermediate level in 10% of the trainings, all three levels in 14%, and no such data is indicated for the rest). 5-day courses typically begin with contents for beginners, which is then followed by the presentation of intermediate and advanced materials. Therefore, the main topics are often presented at a beginner level first, and then at an intermediate level during the second half of the week.

👤 As regards the **number of participants**, no such data is available is most cases.

💲 **Prices:** While PRACE trainings are free of charge, the following prices were typically indicated for the rest of the courses:
Students (undergradute): EUR 35 to 40.
PhD students of universities in Germany: EUR 75 to 90.
Staff members of universities and public research institutes in Germany: EUR 75 to 90.
Staff members of universities and research institutes in EU or PRACE member states: EUR 150 to 180.

◎ **Target groups:** In 66% of the courses, no information has been found on the target groups. The available data suggests that in most (16%) of the cases, university lecturers, staff members of research institutes, and students are indicated, while HPC users and researchers are targeted in 8% and 7% of the courses, respectively.

Repeated trainings: It is typical of the HLRS training activities that most courses are repeated each year or even more frequently; this suggests that these courses are demanded at that frequency, and that these are the key topics.

🔖 **Topics, labels:** Based on the labels, the following topics are covered:
Parallel Programming (PAR), MPI, OpenMP, Code Optimization: 33 courses, 47 %; Programming Languages (LNG): 17 courses, 25%; Computational Fluid Dynamics (CFD): 8 courses, 12%

⬇ **Downloadable training materials:** Training materials are only uploaded to the website for 4 courses.

4. JSC:

Jülich Supercomputing Centre operates one of the most powerful supercomputers of Europe, JUWELS. About 200 experts and contacts for all aspects of supercomputing and simulation sciences work at JSC. JSC's research and development concentrates on mathematical modelling and numerical, especially parallel algorithms for quantum chemistry, molecular dynamics and Monte-Carlo simulations. JSC offers a wide range of courses to members of the Forschungszentrums Jülich and to supercomputer users. Furthermore, students and staff members of universities or research institutes may also participate.

- **Number and form of the training programs:** During the past 2 years, 31 trainings were held: 12 in 2020 and 20 in 2019. As regards 2021, the data for 14 courses are currently available on the website. Due to the coronavirus-related restrictions, the courses went fully online from March 2020; before that, trainings were only held under conventional frameworks.

  Duration and language: As regards the duration of the trainings, the picture is mixed: 42% of them last 3 to 4 days, 31% of the courses are of 5 days or longer, and 27% are 1- to 2-day events. Most of the courses are in English, and 14% of them are held in German.

  Content levels: While levels are not indicated for half of the training programs, 18% and 20% of them are classified as beginner and intermediate, respectively.

- **Prices:** 23% of the courses are free of charge but no information is available for the rest.

- **As regards the target groups**, the data shows the following distribution:
  Researchers: 27%; HPC users: 22%; Programmers: 16%; Staff members: 13%; Others: 22%

  Course levels: Course levels are indicated for 44% of the training programs, and they show that 18%, 10%, and 16% of them are classified as beginner, intermediate, and intermediate/advanced, respectively.

- **Topics, labels:** Programming Languages: 18 courses, 32%, Parallel Programming (PAR): 11 courses, 20%, General HPC: 10 courses, 18%, Machine Learning / Deep Learning: 4 courses, 7%

- **Downloadable training materials:** Training materials are downloadable from the website for 16% of the courses but they are not available for the rest of them.

5. LRZ:

Leibniz Supercomputing Centre (Leibniz-Rechenzentrum, LRZ) has operated world-class supercomputers for decades, and is the IT service provider for all Munich universities as well as for a growing number of research organisations.
LRZ also plays an important role as one of the members of the Gauss Centre for Supercomputing (GCS), delivering top-tier HPC services on the national and European level.
Skills acquired in the field of parallel programming and code optimisation are indispensable for the efficient use of HPC computing architectures. LRZ offers regular courses on various topics in the field of HPC. Since 2012, as part of the Gauss Centre for Supercomputing, LRZ is one of 14 PRACE Training Centers offering HPC trainings for scientific and industrial researchers from across Europe.

Number and form of the training programs: During the past 2 years, 67 trainings were held (20 in 2020 and 47 in 2019). 15 of these were held online, and the rest were conventional ones. As regards 2021, the data for 6 scheduled courses are currently available, and they will be held online.

Duration and language: As regards percentages, most (62%) of the trainings last 1 to 2 days, followed by 3- to 4-day courses (25%) and courses that are of 5 days or longer (11%). Information on the language of the courses is not available.

Content levels: Based on the data available on the website, course levels show a mixed picture: beginner, intermediate, and advanced courses are held in almost identical numbers.

Prices: 41% of the courses are free of charge and no information is available on 50% of them.

Target groups: As regards target groups, the available data is insufficient: unfortunately, these are not indicated for 80% of the courses.

Downloadable training materials: Training materials are uploaded for 22% of the courses.

Topics, labels:
Programming Languages: 23 courses, 31%
Parallel Programming (PAR): 9 courses, 12%; Computational Fluid Dynamics (CFD): 8 courses, 11%; Deep Learning: 8 courses, 11%; General HPC: 5 courses, 7%, Machine Learning / Deep Learning: 4 courses, 5%

6. CINECA

Cineca is a nonprofit consortium of 67 Italian universities and 13 institutions.
SCAI (SuperComputing Applications and Innovation) is the high-performance computing division of Cineca, and is the largest information technology centre of Italy.
Using high-performance computing and the most innovative supercomputer systems, Cineca is based on the most advanced architectures and technologies, and provides support for public and industrial scientific research.
Education regarding novel technologies is an inherent part of Cineca's activities.
The institution is also a CUDA Research Center, and one of the PRACE Training Centers (PTC).

Number and form of the training programs: 18 and 30 courses were organised in 2020 and 2019, respectively, and data for 9 courses are available on the website. 38% of the trainings were held online and are planned to be online in the future.

Duration and language: As regards the duration of the courses, the following distribution emerges: Most frequently (in 68% of the cases), courses last 3 to 4 days followed by 5-day courses, which represent 19% of the training programs. 54% of the courses are held in Italian.

Content levels: No course levels are indicated by the institution.

Prices: All listed training programs are free of charge.

- **Target groups:** The available data suggests that in most (45%) of the cases, researchers and programmers are indicated, while students and scientist are targeted in 21% and 14% of the courses, respectively.

- **Downloadable training materials:** Training materials are uploaded for a great majority (79%) of the courses assessed.

- **Topics, labels:**
  Programming Languages (LNG): 21 courses, General HPC: 12 courses, Parallel Programming (PAR): 8 courses, Machine Learning / Deep Learning: 4 courses

7. EPCC

EPCC is one of the largest training services providers in high-performance computing (HPC) in Europe, and offers several courses for HPC users in the United Kingdom and Europe. EPPC is a PRACE Training Center, and is also involved in two accredited postgradual university master training programs with a duration of one year. Several courses are organised for undergraduate and postgraduate students, scientific researchers, and industrial players.

- **Number and form of the training programs:** 36 and 29 trainings were organised in 2020 and 2019, respectively. As regards 2021, the data for scheduled 6 courses are currently available on the website. 56% of the trainings were held online or are planned to be held online in the period to come.

  Duration and language: As regards duration, 82% of the trainings organised by EPCC last 1 to 2 days, followed by 3- to 4-day courses (18%). The trainings are held in English.

- **Prices:** 23 free courses are indicated but no information is available on the rest.

- **Target groups:** As regards target groups, few – only 20 – of the courses had an indicated target audience, and these showed the following distribution:
  HPC users: 9 courses; Students: 5 courses; Programmers: 3 courses; Researchers: 3 courses

- **Downloadable training materials:** Training materials are uploaded for 87% of the courses, and the lectures were recorded for another 29 courses (40%) with available videos/webinars.

- **Topics, labels:** General HPC: 31 courses, Parallel Programming (PAR): 10 courses, Programming Languages (LNG): 7 courses, Tooling: 6 courses, Big Data / Data Analysis: 4 courses

8. BSC

One central mission of BSC Barcelona Supercomputing Center is to educate and train researchers. As part of the PRACE cooperation and as one of the designated PRACE Training Centres (PTC), BSC has participated in the development of the European HPC training curricula during the past 8 years. BSC's training and education team builds its education program on the basis of the topics from cutting-edge scientific research, and from the development of the models, software tools, and simulation environment of the fields of high-performance computing and application.

In 2011, NDVIA awarded BSC and the Universitat Politècnica de Catalunya (UPC) with the status of a GPU Center of Excellence (GCOE) recognising the role of these leading institutions in cutting-edge parallel computing research.

Number and form of the training programs: A total of 43 courses are listed on the website, 14 of which are scheduled for 2021. Because of the pandemic, these will be held online, or the decision whether trainings in the conventional form are feasible will only be made later. 29 trainings were held in the past two years, and these were in the conventional form.

Duration and language: As regards duration, 1- to 2-day courses are the most common (65%). 3- to 4-day and 5- to 8-day courses represent 21% and 14%, respectively. The language of the courses is not indicated but assumed to be English.

Content levels: The indicated level of difficulty was intermediate for 20 courses (48%), intermediate to advanced for 8 courses, and no such data was available for 32% of the courses.

Prices: The listed training programs are free of charge.

Target groups: For most (76%) of the courses, the indicated target audience included interns and students. No such data is available for the rest of the trainings.

Downloadable training materials: Training materials have not been uploaded for any of the courses.

Topics, labels:
Parallel Programming (PAR): 12 courses, General HPC / interdisciplinary: 9 courses, Programming Languages (LNG): 7 courses, Code Optimization: 5 courses, Big Data / Deep Learning (DATA): 4 courses.

III. Summary

The present assessment covered a total of 392 trainings of 8 European institutions.
As regards the form of the courses, online trainings were exceptional before the coronavirus pandemic. In view of the situation due to the coronavirus, all institutions plan online courses during the first semester of 2021.
The number of participants is usually unrestricted but is bound to registration in all cases.
In the case of HLRS and LRZ, course prices were also indicated for certain trainings. Such data is missing in many cases; these courses are presumably free of charge. PRACE trainings are always free of charge.
Every institution has certain trainings that are repeated every year or even more frequently. These may enable conclusions regarding the courses having the highest demand: these trainings constitute the main educational profile of the respective institutions.

As far as downloadable training materials are concerned, the assessed institutions show considerable differences. EPPC and Cineca are outstanding among the institutions in these terms: training materials have been uploaded for most courses. Further contents and recorded lectures are also available for many EPPC courses.
47%, 34%, and 19%, respectively, of the assessed trainings lasted 1 to 2 days, 3 to 4 days, and 5 days or longer.
In about half of the courses, target groups were not specifically indicated; the practice is rather to specify prerequisites, i.e. the required level of knowledge / expected experience, which are almost always indicated among the course data. Therefore, general conclusions on the target groups are difficult to make on the basis of the available data; however, one or more of university students, researchers, scientists, and interns are indicated in most cases. In 38 cases, HPC users are specifically invited.

Course levels: The assessed institutions indicated course levels in less than half of the cases (42%).
11% of the trainings are of beginner to intermediate level, 11% are of intermediate level, 10% are of advanced level, and another 10% are of intermediate to advanced level.
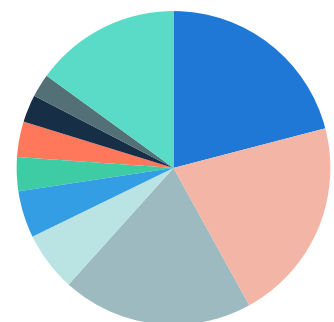
Fields of training: Following the practice used by HLRS, the topics (Column P) of all courses were identified to reveal the most frequent fields of training.

Most frequently, the assessed institutions organise trainings on Parallel Programming (21%), Programming Languages (21%), and HPC in General (20%), followed by Machine Learning (6%), Computational Fluid Dynamics (5%), Code Optimization (3,5%), Scientific Visualization (3,3%), Big Data / Deep Learning (DATA) / Data Analysis (2,8%), and Tooling (2,5%).

Others: Molecular Dynamics (1.7%), Deep Learning (2.2%), CUDA (1.2%), GPU Programming (1%), Sysadmin (1%), Cloud Computing, Unreal Engine, Compute Cluster and Administration (ADM), Performance Engineering, ExaHype.

AREAS OF TRAINING

- 21% *Parallel programming*
- 21% *Programming languages*
- 20% *General HPC*
- 6% *Machine Learning*
- 5% *Computational Fluid Dynamics (CFD)*
- 3% *Code optimization*
- 3% *Scientific Visualization*
- 3% *Big Data/Deep Learning (DATA)/Data analysis*
- 3% *Tooling*
- 15% *Other*

EURO

KIFÜ