Bridging the Gap between Education and Research: A Retrospective on Simulating an HPC Conference

Sarah Neuwirth
Institute of Computer Science
Goethe University Frankfurt, Germany
s.neuwirth@em.uni-frankfurt.de

Abstract—High Performance Computing (HPC) is playing an increasingly important role in industry, research and everyday life. Moreover, a central core of the European HPC strategy is the Modular Supercomputing Architecture (MSA), which breaks with traditional HPC architectures by integrating heterogeneous computing resources in system-level modules. Nevertheless, HPC and especially MSA content only rarely find their way into the curriculum of computer science courses at German universities. In addition, the necessary competencies for independent scientific research are hardly addressed, although these skills are essential for students for writing their final theses.

This paper presents a blended learning based module concept that promotes the understanding and application of modular supercomputing while connecting it with the techniques of scientific project work. The module was first implemented at Goethe University in Summer 2022. The initial feedback and evaluation results are quite encouraging both in terms of learning outcomes and student engagement and interest.

Index Terms—HPC, Modular Supercomputing, Blended Learning, Computer Science Education

I. INTRODUCTION

Reaching Exascale compute performance and beyond at an affordable budget requires increasingly heterogeneous high performance computing (HPC) systems. Hence, in recent years, highly specialized processors suited for different tasks and use cases have been developed, including graphics processing units (GPUs), tensor processing units (TPUs) and quantum processing units (QPUs). The Modular Supercomputing Architecture (MSA) [1], [2] breaks with traditional HPC systems and exploits these heterogeneous resources by integrating them at the system level. An MSA system consists of a collection of modules with different architecture and/or performance characteristics and can supply any combination or ratio of resources across modules. It is not bound to fixed associations between, for instance, CPUs and accelerators as will be found in clusters of heterogeneous nodes and is therefore ideal for HPC centers running a heterogeneous work-load mix. The goal is to provide cost-effective computing at extreme performance scales fitting the needs of a wide range of computational sciences. Each application run can dynamically decide which kinds and how many nodes to use, mapping its intrinsic requirements and concurrency patterns onto the hardware (as depicted in Figure 1), therefore improving the parallel efficiency, time to solution and energy use.

While an introduction to high performance computing [3] and parallel programming is now offered as an elective module at several German universities, the programming and architecture paradigm of the modular supercomputing architecture is not yet part of the standard curriculum. To the best of the author’s knowledge, only the Modular Supercomputing and Quantum Computing research group at Goethe University Frankfurt has offered a course on this topic, although modular supercomputing is the core concept of the European Exascale strategy and the first European Exascale-class computer JUPIO (short for “Joint Undertaking Pioneer for Innovative and Transformative Exascale Research”) [4]. In addition, students at German universities often lack the necessary research skills to work on their final theses, which is a mandatory module in both bachelor’s and master’s degree programs.

Previous work [5], [6] has already highlighted the necessity of implementing education on supercomputing in leading research and technology universities. Multiple different formats have been proposed including teaching formats for inexperienced programmers [7] and graduate students [8], but also on-demand formats [9] have been explored. However, these teaching formats mainly focus on imparting fundamental and application-related knowledge, but neglect the integration of research aspects and methods.

In this paper, the author presents a seminar module that, based on blended learning [10], combines understanding of basic knowledge of modular supercomputing with the methods of scientific work. Not all people learn effectively in the same way. This is the ulterior motive of blended learning: existing (learning) concepts are mixed with new concepts in order to make learning opportunities as diverse as possible and to achieve the greatest learning effect. The goal of this seminar format is for students to acquire the necessary skills to conduct

Fig. 1: Modular Supercomputing Architecture.
Fig. 2: Research Development Framework [11].

The aim of seminars is to help students in advanced semesters to acquire the competencies in scientific work and research methods necessary for their later theses, while also deepening their knowledge about the chosen seminar topic. In computer science, seminars usually consist of an introductory session. Afterwards, each student works individually on their topic and presents it afterwards. Often, the work is limited to merely summarizing a given paper. Seminars take place either once a week in 1.5 hour sessions or at the end of the semester as block courses. By the time students begin their thesis, they should have the competencies of Level 4 (Student-Initiated).

In addition, extracting and presenting central core statements and findings are also important qualifications for their further professional life (within and outside academia). For bachelor students, it can be useful to start with a Level 2 method, since their bachelor theses are classified between Level 3 and 4.

Nevertheless, seminars at German universities often focus only on the theoretical aspect of the seminar topic and neglect the teaching of scientific methods and their application to research projects. Still, seminars are often the only course modules designed to bridge the gap between education and research. Therefore, in this paper, we present a seminar course design that combines theoretical and applied aspects of high performance computing with scientific project work.

III. Course Structure

This module introduces fundamental concepts of high performance and heterogeneous computing, specifically modular supercomputing. The course is an elective module offered in the form of a block seminar and can be taken by students in the advanced semesters of the Bachelor’s program in Computer Science and the Master’s programs in Computer Science and Information Systems. The module is divided into different phases and is designed as a 14-week course, which also includes elements such as the submission of a short paper and a conference presentation. Each phase comes with its own format, teaching material, and student interaction. Typically, 10 to 15 students enrol in a seminar each semester. There are no prerequisites to participate in this course.

A. Introductory Session and Seminar Organization

In the course of this seminar, students practice the individual steps of a small HPC research project, from finding a topic to writing and presenting a scientific paper. At the beginning of the lecture period, a 90-minute introductory lecture provides an introduction to "Modular Supercomputing", the learning objectives, and the structure and timeline of the seminar.

After the introductory lecture, the students pursue their research project in individual phases, as outlined in Figure 3. The students first engage in finding their own research topics. Under guidance, they then research and read current scientific publications in digital libraries, such as IEEE Xplore and ACM Digital Library, and extract the key statements relevant to their topic. This allows students to become familiar with the current state of the literature and its proper citation. This is followed by a phase of individual project work under the guidance of the lecturer and the mentoring team. Each student writes a short paper about their topic. Then, the paper is submitted and undergoes a peer review stage. Finally, the seminar concludes with a mini conference at the end of the lecture period.

B. Learning Objectives

The learning and competence goals of this module combine two different areas: (1) apply scientific methods and (2) understand the fundamentals of modular supercomputing. Following Bloom’s Taxonomy [12], the learning objectives for area (1) fall into Application, Analysis, and Synthesis, and for area...
(2) in Comprehension and Application. Having completed this module, students should be able to:
1) understand the motivation behind the design of the modular supercomputing architecture
2) describe and distinguish between the main differences of traditional HPC and MSA systems
3) recognize current research challenges with scaling and deploying MSA systems at exascale
4) apply literature search and analysis techniques
5) prepare and communicate complex issues
6) formulate constructive peer feedback
7) organize, prepare, and present small research projects.

The author believes that the combination of theory and practice will consolidate the knowledge among students and enhance their curiosity and interest in the subject area of HPC. Also, intrinsic motivation of students is key to the learning success and an increased interest in the topic itself. By working on individual projects, a deeper understanding is created and at the same time an insight into the world of research is provided.

C. Topics

The following HPC and MSA related topics have been selected to be briefly discussed in the introductory session:
- Evolution of Modern Supercomputing [13]–[15]
- Modular Supercomputer Architecture [1], [2], [16], [17]
- Overview of HPC Benchmarks [18]–[20]
- Introduction to Hybrid Programming (MPI+X) [3], [21]
- Potential Programming Models for MSA systems (e.g., OmpSs [22], [23] and oneAPI [24])
- Parallel I/O and Storage Models [25]–[28]
- Characterization of Emerging HPC Workloads [29]–[32]
- Energy Efficiency in HPC Data Centers [33]–[35]
- Containerization and Cloud Technologies [36], [37]
- Posits: A Potential Replacement for IEEE 754 [38], [39]

These topics serve as a basis for the formulation of the individual research topics and the initial literature search.

D. Seminar Phases

1) Phase I – Topic Identification and Initial Research:
During the first week of the lecture period and after the introductory session, each student independently researches seminar topics of interest to him or her and chooses a topic for an individual project. To ensure an equitable distribution of the various topics, each student submits three topic proposals from which one topic is selected per student.

From this stage on, a mentoring team consisting of doctoral students and the instructor can be consulted regarding open questions and problems. Consultation meetings are arranged individually upon request. All students are expected to participate and report regularly on their progress.

Selected student projects from summer 2022:
- MSA – From Idea to Production: In this topic, the historical development of MSA was highlighted. In addition, a critical analysis was conducted on the weaknesses of the current MSA systems.
- Porting Applications to a Modular Supercomputer: Experiences from Mapping a Radio Astronomy Workflow: This topic used the example of a workflow from the field of radio astronomy to show how individual code parts can be mapped to the most suitable MSA modules and what advantages this offers over traditional HPC systems.

A Beginner’s Guide to Ranking Supercomputing Systems: In this project, various supercomputing ranking lists such as TOP500, Green500, Graph500, HPCG500 and IO500 were critically reviewed and compared. In addition, the applicability to MSA systems was discussed.
- HPL, HPCG, HPL-AI and Co: Why are different HPC Benchmarks needed? This topic addressed the weaknesses of various HPC benchmarks and, in particular, their validity for current workloads. This included, among other things, the analysis of the various algorithms.
- Comparison of Storage Models in HPC: One student analyzed the differences between block, file, and object storage, and compared their applicability to HPC environments. This included implementing code examples to demonstrate POSIX I/O weaknesses and hands-on testing on the FUCHS-CSC cluster.

2) Phase II – Literature Research and Individual Projects:
In the second phase of the seminar, students work on their individual projects. Approximately four weeks are allotted for this phase. Students create their bibliography, formulate their research questions, and conduct their research. Depending
on whether it is a seminar paper of a bachelor or master student, the topics are rather theory- or application-oriented. For Bachelor students, the focus is on understanding and presenting complex topics, while for Master students, independent project work is expected. This includes, for example, hands-on experience with writing and running of code on the FUCHS-CSC cluster at Goethe University. In addition to the practical work, this phase also serves to identify open comprehension questions about modular supercomputing, to learn the methods of scientific work and for the initial project organization.

In order to learn the methods of scientific work and research projects, short instructional videos (5 to 10 minutes each) are provided on topics such as literature research, reading scientific texts, and project time management. Students again have the opportunity to schedule individual consultation meetings to discuss problems and the progress of their work.

3) Phase III – Writing a Scientific Paper: The third phase of the seminar is devoted to writing the scientific paper using the conference template from the Institute of Electrical and Electronics Engineers (IEEE) [40]. Approximately four weeks are planned for this phase. The seminar paper should not exceed four single-spaced, double-column pages using 10-point size font including everything excluding references.

Again, short instructional videos on proper citation and academic writing are provided. In addition, a tutorial on creating the outline of scientific papers with examples is made available via Moodle. As in the previous phases, the mentoring team is available to provide advice and support.

4) Phase IV – Paper Submission and Peer Reviews: The papers are submitted electronically in PDF format via the conference portal easyChair [41], which offers a simple and free possibility to organize conferences and workshops. Among other things, it provides the possibility of peer reviewing.

After the submission, all students are assigned two papers for peer review. The assignments are done in a double-blind fashion to preserve anonymity. The students receive an introduction to constructive feedback methods and criteria of good reviewing prior to the peer review phase through a short 10-minute video. In addition, the students are presented with a standard template for peer reviews, as seen in Figure 4, and sample reviews. To ensure consistency and quality of the reviews, mentors serve as the third reviewer for each submission. Through the peer review process, students receive early feedback on their work, but can also compare how their peers approach research projects and academic writing.

Up to two weeks are allocated for the peer review phase. After that, each student receives three reviews of his or her work. As in a real-world conference, this feedback is used to prepare the camera-ready paper submission. In preparation for the HPC conference and the peer feedback session on scientific presentations, a sheet with observation dimensions for good technical presentations is provided.

5) Phase V – Simulation of an HPC Conference: Towards the end of the lecture period, everyone meets again for the conference. Prior to the conference, the students receive instructional material about how to prepare effective science presentations. The conference is organized in technical paper sessions and each student has 30 minutes to present their work. After each talk, there is a 15-minute Q&A and peer feedback session to discuss the presented work and provide students with additional feedback on their presentation style.

At the end of the conference, two awards are presented: (1) Best Paper Award and (2) Best Reviewer Award. Figure 5 displays examples of the award certificates. This is intended to increase motivation in the preparation of the paper and the writing of the reviews, and indirectly to increase the quality of each step of the project work. Similar to a real-world conference and based on the submitted reviews, ratings and a program committee discussion, the recipients of the best paper award and best paper runner-up are selected. The best reviewer awards are based on the overall quality of the reviews and focus on whether the major review criteria has been met. In addition to the certificates, book prizes are being awarded.

6) Closing – Submission of the Conference Material: Upon successful completion of the conference, students submit their final presentation slides and camera-ready papers for evaluation at the end of the semester. The overall grade for the seminar is composed of 45% from the submitted paper, 45% from the presentation, and 10% from the peer reviews.

IV. Evaluation

This module was first introduced in the summer of 2022 in a combined undergraduate and graduate seminar at Goethe
University Frankfurt and attended by 15 students. Seminars are normally limited to 10 to 15 participants. This class size is therefore representative. The evaluation results are from an early pilot study to share first experiences and outcomes.

A. Design of the Evaluation

The design of the teaching research project is guided by two (quasi-)experimental test designs to evaluate the personal learning and the individual learning experiences.

1) One-group pretest-posttest design to capture the personal learning success: The students’ expectations of the seminar are recorded at the beginning of the introductory lecture by means of self-disclosure via an initial survey. After the conference, students are asked to assess their personal learning success and the suitability of the chosen seminar format, also by means of a final survey. The questions for the surveys are measured on a 5-point scale, with 1 for strong disagreement and 5 for strong agreement. Various subject characteristics such as previous experience with methods of scientific work, degree program, and semester of study are also queried in order to be able to take into account confounding variables in the subsequent evaluation.

2) One-shot-design to capture individual learning experience: After the conference, students conduct a retrospective and reflection on their learning experiences using Gorman and Gottesdiener’s 4L method [42]. The 4Ls stand for liked – learned – lacked – longed for. The survey is conducted anonymously via padlet.com, a digital whiteboard service. The results of this survey are also used as constructive criticism for future installments of this module.

B. Learner Demographics

The students were asked a series of questions in the initial survey including:
- What is your gender identification?
- What is your field of study?
- What is your academic classification (including version of your examination regulations / PO)?

The students were able to freely describe how they identify their gender and field of study in open text boxes. Figure 6 shows the self-disclosed demographics of the students. 20% of students identified as female, which corresponds to the percentage of enrolled female students in the Institute of Computer Science at Goethe University, and 80% identified as male. None of the students identified as transgender or non-binary. Out of the 15 students, 50% are enrolled in the bachelor’s degree program in computer science according to examination regulations as of 2019 (abbreviated as PO 2019 in Figure 6a), 36% in the bachelor’s degree program in computer science according to examination regulations as of 2011 (abbreviated as PO 2011), and only 14% are enrolled in the master’s degree program in computer science. The version of the examination regulations plays a decisive factor in prior knowledge, as only students according to PO 2019 can take a module on scientific writing at the end of their studies. Still, this module only focuses on the theoretical aspects of scientific writing and does not teach them any applied research skills.

C. Survey Evaluation

The students were asked about their motivation for choosing this seminar and their interest in the seminar topic. The results are displayed in Figure 7. Overall, students were very interested in the future potential and application of the seminar topic. In addition, there was a great interest in the research area and the preparation for a possible thesis in the area of HPC was a great motivating factor as well. Furthermore, the students had the opportunity to provide information on further expectations and motivating factors in a text box. Here it became clear that many have a great interest in learning the methods of scientific work through practical application. But modular supercomputing was also a decisive factor for many. One student stated that he or she knew the lecturer from another course and therefore decided to take the seminar.

In the second part of the survey, the students were asked to perform a self-assessment about their previous experience with methods in academic work. The average results from both the initial and the final survey are shown in Figure 8. Of particular interest was their previous experience with and knowledge of

- Constructive feedback methods
- Scientific discussion
- Presentation and communication methods
- Formal design of scientific work (e.g., layout, figures, citations, etc.)
- Time and project management
- Scientific language and style
- Organization of outline and structure
- Scientific writing
- Literature research and analysis
In the initial survey, the students’ experience level in all categories scored between 2.4 (little experience) and 3.4 (so-so). Most of them had little to no experience with research projects and scientific writing. Only the master students had previous experience from their bachelor’s degree programs, specifically from writing their bachelor theses. Even the bachelor students with previous experience from other seminars rated their experience levels rather low, which is in line with the author’s experience and observations from other seminars.

In the final survey, the students were asked to perform the self-assessment again. After the successful completion of the seminar, the students rated their experience and confidence level now between 3.4 (so-so) to 4.2 (experienced) on average. Figure 8 also reports the experience improvement from the self-assessments, which can be calculated as follows:

Experience Improvement (in %) = 100 × \left( \frac{\text{Score}_{\text{initial}}}{\text{Score}_{\text{final}}} - 1 \right)

The results show that the individual student experience and confidence increased by up to 42%. Also, the margin of error is statistically close to zero since all the students participated in both surveys. The students showed the most significant improvements in scientific writing, use of scientific language, and the ability to lead scientific discussions.

D. Retrospective

In addition to the surveys, students were asked to participate in retrospective feedback. As discussed, the 4L technique according to Gorman and Gottesdiener was used for this purpose. Each student was asked to anonymously write down what they Liked, Learned, Lacked, and Longed For – one per digital sticky note. Figure 9 displays the students’ responses. For reasons of space, each answer is included only once.

Overall, the seminar format was very well received. Most students indicated they were very satisfied with the topic selection. According to the students, the scope of topics was chosen so that HPC novices and students with prior knowledge could find engaging projects. The organization of the seminar in different phases and learning formats were also very well liked. In particular, the individual support provided by the mentoring team was highly appreciated. In addition, the early feedback through the peer review was very well received. Here, the comparison between the perceptions of fellow students and the academic staff was found to be quite helpful. Another very popular aspect was the insight into all phases of scientific project work and the structured time frame.

In the Learned category, most students indicated that writing a research paper was a new and helpful experience for them. Here, points such as learning how to cite correctly and how to structure scientific texts were highlighted in particular. Furthermore, almost all students emphasized that the seminar gave them a lot of insights into the subject area of HPC and its importance for everyday life. It was also praised that the seminar provided good insights into the European Exascale strategy of modular supercomputing.

Overall, there were three comments in the Lacked category that were meant as constructive feedback for future iterations of the seminar. Most notably, one student emphasized that it would be helpful to move the submission deadline of the preliminary paper and reviews forward. This would allow feedback to be considered not only for the camera-ready version of the paper but also for the conference presentation. Another suggestion was to hold an additional (online or offline) meeting with all students halfway through the semester to better prepare for the paper writing phase.

The last feedback category was Longed for. Here, students expressed a desire to hear more presentations on HPC. Another suggestion was to also accept higher-level topic suggestions from students in the future in order to be able to respond even better to the interests. Finally, it was suggested to invite additional members of the department to the conference in order to have more lively discussions.

E. Lessons Learned and Outcomes

Overall, the evaluation results are very satisfactory. These results are also confirmed by the overall very good grades.
The opportunity to work on individual topics and to present the results in scientific papers under guidance was very well received. From the feedback, it also became clear that the chosen format can ideally prepare for the requirements of bachelor and master theses. Most students regularly took advantage of the possibility to schedule individual meetings and found them very helpful. The choice of topics was also very well received by the students; following the seminar, the lecturer has already received five inquiries from participants for theses in the field of HPC.

Since there are very few HPC-oriented modules in the computer science programs in Germany and also the teaching of techniques of scientific work are often neglected, this module represents a good first attempt to combine research and teaching. At Goethe University, for example, there is only one elective module on supercomputer architectures and one elective project lab on parallel programming. In addition, there is an introductory lecture on modular supercomputing, but this is purely theory-based. For future iterations of the seminar format, the following findings and ideas can be summarized.

**Peer Review Phase:** The idea of peer review was very well received by all students. In the first implementation, the submission of the preliminary paper was in the 10th week of the semester. Since there were two weeks to prepare the reviews, there was only one week between the conference days and the sending of the reviews. In future iterations, the review phase will be moved forward by one week. This will allow feedback from the reviews to be taken into account when preparing the conference presentations.

**Topic Selection:** Although the range of topics was well liked by the students, there was a desire for more freedom in the choice of topics. In the future, own topic ideas should also be considered in consultation with the lecturer.

**Meetings and Communication:** Even though most students took advantage of the opportunity to meet individually, they were some students who never contacted the mentoring team. Interestingly, the same students said they wished for better communication of deadlines, although the instructor and mentors always made an effort to communicate all information multiple times and regularly. The learning platform Moodle, e-mails, and a Discord server were actively used for communication. In the future, the mentoring team will proactively contact those students to ask them about their progress.

**Poster Presentations:** An alternative to paper presentations would be lightning talks followed by a poster session. The effective design of posters and delivery of elevator pitches is another valuable skill in both the academic world and industry. Future installments of the seminar will explore posters and extended abstracts as an alternative to scientific papers.

**Scaling the Class Size:** In the initial implementation of the seminar, a total of two doctoral students and the instructor were available for individual supervision and mentoring sessions. With a course size of 15 students, the concept could be implemented very well, as the workload could be divided equally. Students attended advising meetings only once or twice during the semester, so the amount of supervision remained manageable. However, due to the variety of project topics, it can be assumed that the course format cannot be implemented with a larger number of students.

**V. Conclusion and Future Work**

This paper presented a module for exposing both undergraduate and graduate level students to high performance computing and scientific project work. The seminar module covers fundamental HPC and modular supercomputing topics while exploring the methods of scientific project work through hands-on experience. The preliminary evaluation from the first implementation is very promising both in terms of student learning outcome and engagement. The design of the module...
seems to stimulate the intrinsic motivation of the students, which ultimately leads to better results.

In the future, the author plans to adapt the timeline of the seminar to incorporate the feedback, but also to explore poster presentations as an alternative to scientific papers.

REFERENCES