The HackHPC Model: Fostering Workforce Development in High-Performance Computing Through Hackathons

Jeaime Powell
Texas Advanced Computing Center
University of Texas at Austin
Austin, TX, USA
jpowell@tacc.utexas.edu

Linda Bailey Hayden
Science Gateway Community Institute
San Diego, CA, USA
haydenl@mindspring.com

Amy Cannon
Omnibond Systems, LLC
Liberty, South Carolina, USA
amycannon@omnibond.com

John Holly
Southeastern Universities Research Association
Washington, D.C., District of Columbia, USA
jholly@sura.org

Charlie Dey
Texas Advanced Computing Center
University of Texas at Austin
Austin, TX, USA
charlie@tacc.utexas.edu

Alexander Nolte
University of Tartu
Tartu, Estonia
Carnegie Mellon University
Pittsburgh, PA, USA
alexander.nolte@ut.ee

***Abstract*—Developing a workforce with applied skills in the high-performance computing (HPC) field has highlighted a gap in personnel experiences needed for capabilities towards scientific purposes. The exposure provided in current academic programs, especially on an entry level is insufficient. To increase student participation HackHPC, a collaboration between the Science Gateways Community Institute, Omnibond Systems, the Texas Advanced Computing Center, and the University of Tartu was formed to address that gap through the application of events known as hackthons. Hackathons are time-bounded events during which participants form teams to work on a project that is of interest to the them. Hosting a hackathon that has the desired long term outcomes involves a number of crucial decisions related to preparing, running and following up on an event. In this paper we report on the development and refinement of the "HackHPC Model" which includes methodologies, participants, procedures, and refined implementation of practices used to plan and host hackathon events to foster workforce development in HPC.**

***Index Terms*—science gateways, hackathon, learning, student participation**

I. Introduction

High-performance computing (HPC) has become an essential tool to advance science [1], [2]. At the same time there is a gap between the skills needed to ensure that HPC capabilities can be utilized for scientific purposes and the exposure provided in current academic programs especially on an entry level [3]. To address this challenge the Science Gateways Community Institute (SGCI) together with Omnibond Systems, the Texas Advanced Computing Center, and the University of Tartu started an initiative to increase student participation in HPC. A cornerstone of this initiative was to establish a series of hackathons for students to pave their way towards becoming future HPC professionals [4]. Hackathons are time-bounded events during which participants form teams to work on a project that is of interest to them [5]. Such events promised to be a great fit for this endeavor since prior work has shown that hackathons can foster learning [6] and aid community development and growth [7].

Organizing a student hackathon sounds simple enough, but creating a suitable event design for the outlined purposes is indeed anything but. Running a hackathon that has the desired long term outcome involves a number of crucial decisions related to preparing, running and following up on an event [8]. Starting with an initial design in 2018 we have since conducted nine events that were connected to various conferences within the HPC community. Reflecting after each event, we iteratively refined our model resulting in the design we present in this paper. Part of the refinement was moving from an "in-person" to an "online" event design due to the COVID-19 pandemic [9].

In the following we will discuss iterations developed (section II) prior to the current revision of our model (section III). Additionally we will provide information regarding our participants (section II-A) and outline our future plans (section IV).

II. History

The initial concepts that became the HackHPC Model were derived from a hackathon training workshop hosted at the Computer-Human Interactions Conference in 2018 [10]. Hackathons were considered a hands on method to introduce HPC concepts into workforce development. The workshop resulted in a collaboration between the Science Gateways Community Institute and the Institute for Software Research at Carnegie Mellon University with targeted to hold an event at the 2018 Practice and Experience in Advanced Research

Presented at Gateways 2022, San Diego, USA, October 18–20, 2022. https://zenodo.org/communities/gateways2022/

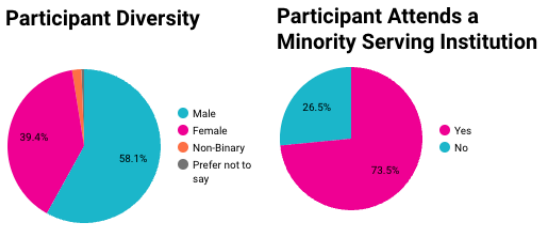

Fig. 1. Participant diversity

Computing (PEARC) Conference. The then named "Science Gateways Hackathon Event" took place over the course of two days and led to the addition of Omnibond Systems as key partner and collaborator. The success of the initial hackathon collaboration was soon followed by the first Supercomputing Conference (SC) co-located event sponsored by Omnibond in 2018 then named the "CloudHPC Hack".

The two events held in 2018 at PEARC18 and SC18 defined the initial approach that was later refined into the HackHPC Model. The approach included small (3-5 students) teams with subject matter experts as team mentors, varied levels of coding experiences within teams, mentor provided challenges that were used by teams to develop a project idea, provided cloud-based resources, pre-event training sessions, ice-breakers, team identity tasking, team progress focused challenges through "Checkin" titled mini-events, cross-team directed interactions for sharing of solutions, required deliverables including team GitHub repositories with source code, and judged final presentations with awards. Additionally, targeted adaptations based on prior event challenges have iteratively improved experiences for organizers, judges, mentors, vendors, and student participants alike. Table I lists targeted improvements by the associated HackHPC event and host conference. Each year since 2018 has seen a successive increase in participant registration with a total of 417 registered participants over nine hackathons.

A. Participants

For our events we invite undergraduate and graduate level students of varied skill levels as hackers/participants. We particularly aim to invite students from minority serving institutions[1]. So far we have successfully managed to attract a diverse hacker population related to gender (Figure 1) and study progress. Additionally, 75.3% of the participants at the time of a given event attended a Minority Serving Institution.

After each event we conduct a post-survey which includes a set of established 5-point scales [11] related to the perception of hackathon participants about their team and their project. In addition to scales we also ask participants to share their perception about the event in written form. Analyzing responses it becomes clear that the overall level of satisfaction related to their communication as a group, the team process and their project is very high (Figure 2). This finding is also reflected by open text responses such as: "*I loved the community and their willingness to teach and connect to resources*", "*I like that it was a good challenge and I learned so many different things*" and "*I liked everything! I learned a lot*".

[1]https://www.doi.gov/pmb/eeo/doi-minority-serving-institutions-program

TABLE I
HACKHPC MODEL IMPROVEMENTS

Event	Improvement
Science Gateways Hackathon (PEARC18)	Team identity, Mini-challenges, Deliverables, Viewer's Choice Award
CloudHPC Hackathon (SC18)	Scheduled downtime, Virtual mentoring, Cloud-based resources, Slack Team usage, Industry Sponsorship
HackHPC@ PEARC19	PEARC Student Program inclusion, Expanded mentor recruiting, Pre-event participant training, Social Media presence generated, YouTube Streaming/Recording
Cloud HackHPC@ SC19	Expanded student recruiting, Google Cloud Platform standardization, Staff specialists, HackHPC Page created
HackHPC@ PEARC20	Fully virtual event, Zoom usage, Team identity adjustments (Virtual Background, Theme Song), E-payments for prizes, GitPages event site
HPC in the City: Atlanta (SC20)	Mentor pre-training and support, Central theme based on host city, Paired student and industry mentors, SC Student Program Inclusion, SIGHPC Professional association sponsorship, Guest speaker lightning talks, Team formation kick-off night, Participation stipends
HackHPC@ PEARC21	Coding institute integration, Judge pre-training, Post event panels, Team intro videos by mentors, Partnered technical and subject matter export mentors for each team
HPC in the City: St. Louis (SC21)	Conference event review presentation, Registration Dashboard usage, Discord usage, SC Inclusivity Program inclusion, XSEDE sponsored faculty support, Giftcard prizes
HackHPC@ ADMI22	Faculty recruited school teams, Training presentation integration into the symposium, Vendor sponsored technology challenge

In addition to post-surveys we also try to stay in touch with participants after an event has ended. Through this we found that we had 66 multi-event participants and three successful student mentors. One student in particular has become a permanent member of our HackHPC family: Josselyn Salgado. She had no coding experience prior to a SGCI Coding Camp she had attended only two weeks before HackHPC@PEARC19. She and her team "*Cloud Busters*" won an award at that event. She returned twice more as a student participant in Cloud HPCHack@SC19 and HackHPC@PEARC20. In HPC in the City:Atlanta@SC20 Josselyn became the first student mentor of HackHPC. She has returned in that position and also has become known for her training session "*Beginning to End*" which was included as a part of the ADMI 2022 Symposium. Josselyn Salgado has become an example of how events like HackHPC can develop skill sets in students with little to no initial HPC experience.

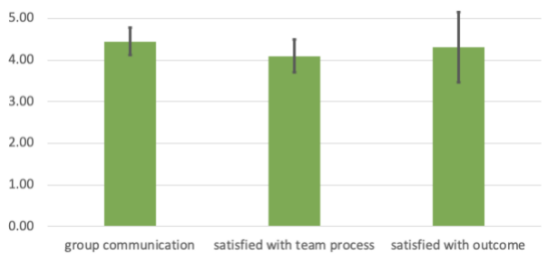

Fig. 2. Combined Participant Satisfaction Scores (N = 58). All responses were given on 5-point scales which were anchored between strongly disagree (1) and strongly agree (5). The bars indicate the mean (m) and standard deviation (SD) for each scale.

III. The HackHPC Model

In this section we will outline the HackHPC model. We elaborate on key roles that are necessary to run one of our events (section III-A) and explain the procedure we follow (section III-B). While our activities for in-person and online events are broadly similar, we would like to note that some of them need adaptation when running an online/virtual as opposed to an in-person event. These include the potential necessity to account for different time zones, parallel commitments, disengagement of participants, technical issues, and others. Online events generally need more deliberate planning and require closer monitoring. For more detailed information we would like to refer the reader towards two related publications [9], [12].

We would also like to note that the event design we present here is suitable for our specific purposes. Using the same design in a different context or for a different purpose might not yield the desired outcomes. For hints on how to adapt this design to fit specific needs we would like to refer the reader to the hackathon planning kit [8].

A. Key roles

Running a HackHPC event requires the combined effort of many people. In its current form, our model revolves around four (4) roles namely organizers/staff, mentors (faculty, technical, and student), judges, and sponsors. The **organizers and staff** consist of faculty, researchers, and vendors. They are responsible for planning, logistics, recruiting, hosting, training, and funding. Additionally members of the organizing team often served and continue to serve as committee members for conferences into which our hackathons are connected. We found this to foster better integration into the host conference and limit conflicts regarding hacker participation in conference sponsored student program/student volunteer activities. For **mentors** we differentiate between faculty mentors, student mentors and technical mentors. **Faculty mentors** can come from different communities including biology, geology, chemistry, computer science and others. They have prior teaching experience and are familiar with HPC technologies. Their role is to provide an idea or challenge that is related to their area of expertise and guide a team of hackers that wants to work on their proposed idea or challenge during the entire duration of a hackathon. **Student mentors** are commonly prior participants of our events. They are similar to faculty mentors in that they can come from different disciplines and that they can also provide an idea or challenge for hackers. We generally advise faculty and/or student mentors to work in teams of two because, supporting a team of hackers for the entire duration of a hackathon can prove difficult to manage for a single person. Finally **technical mentors** are individuals with extensive technical experience related to HPC. They are not tied to a specific team. Instead they provide technical support to teams upon request. Similar to mentors **judges** are often faculty members from different communities that are familiar with HPC technologies. Their role is to rate team projects based on a set of predefined criteria. These criteria include creativity, technical complexity, collaboration and others.

Sponsors are commonly providers of HPC technologies or educational institutions and groups. They can provide different types of support such as access to HPC resources, prizes, support towards student participation or others. For their commitment they are given the opportunity to present lightning talks during the hackathon event. Some of them may also serve as mentors or judges.

B. Procedure

Planning activities start six months or more prior to when a hackathon is scheduled to take place. Initially we focus on scheduling and logistics. Our hackathons are designed for a maximum of 50 participants, and they take place over a four day period. For an in-person event we look for locations with one large room that can fit all participants and multiple smaller spaces that can serve as breakout rooms.

After setting dates and a location we start **recruiting** mentors, judges, sponsors, and participants. For this we utilize an online form in tandem with affiliated relationships. This form includes questions related to participant demographics, affiliation, and prior hackathon experience. Mentors additionally are asked for a description of a challenge or idea they may want to offer for the teams (section III-A). We also **create an event page** that contains information about the event such as the event schedule, location, code of conduct, contact information of the organizers, and additional resources. In addition we invite everyone that is involved in the hackathon to a common Discord server to **facilitate asynchronous communication** between organizers, teams, mentors, judges and sponsors.

A few weeks prior to the event we run several online **training sessions**. For mentors we run a training session where we outline our expectations for interactions with students, and we discuss the ideas or challenges they provided in the registration form. For judges we run a training session where we explain the judging criteria and judging procedure. For hackers we run multiple tutorials where we explain technologies they may utilize during the event (e.g. Github, Discord, Python, Google Cloud Platform, CloudyCluster, and Jupyter Notebooks) as well as an example of a hackathon project from conception to presentation.

The hackathon itself begins with a **welcoming address** by the organizers which in the case of an online event takes place via Zoom. We outline the event agenda, present the code of conduct, explain the different prizes that teams can win (judges choice award, viewers choice award and an event specific award) including the associated criteria for each. This "Kick-off" session sets the tone for the event. We emphasize that while the event is competitive, the purpose of the event is for everyone to have a good time and learn. We also introduce the mentors, sponsors and judges. The introductions are followed **mentor pitches** for their respective idea or challenge. Afterwards the students are asked select an idea from the mentors in which they are interested, thereby **forming teams**. In an in-person setting this activity typically does not take more than one hour. In an online setting we allow additional time for breakout rooms visiting and synchronizing between mentors and interested hackers. After the hackers have decided which mentor they want to join, we adjust teams if necessary. They should be of similar size and between 3 and 5 hackers which in turn means that some mentors might not have sufficient students to form a team while others have enough to form multiple teams.

The remainder of the event is structured around **checkpoints** which take place each day in the morning and in the evening. In case of an online event we again utilize Zoom sessions. Teams prepare and present their progress, discuss challenges they have faced, and outline their goals/tasks at each checkpoint. We also involve students and their mentors in social challenges where they e.g. have to come up with a team name, logo, song and handshake (in-person only) or where the mentor has to produce a hype video for their team. Each checkpoint includes an announcement for what teams are expected to present at the next session thereby allowing us to guide progress during the hackathon. In addition, we conduct checkpoints for mentors during which we discuss potential issues and direct resources to teams in need.

The hackathon ends with a **final presentation session** where all teams present their projects and demos to everyone involved in the event including organizers, mentors, judges, sponsors, and invited guests. The event is streamed live to accommodate other interested individuals such as HPC community members. The final session concludes with announcements of the winners for the various prizes.

After the event has ended, we send out the previously discussed **post-event survey** (section II-A) as well as we open invitations to our both alumni email list and LinkedIn group. These are then later used to share internships and job offerings and also to broadcast upcoming events thus growing our community.

IV. Outlook

We are currently in the process of planning our tenth hackathon. After organizing in-person and online only hackthons we are planning to expand towards hybrid events that welcome online as well as in-person participants. Returning to in-person events sounds appealing, but we do not want to loose the accessibility online events have provided many students. They may have otherwise not been able to participate due to various reasons related to finances, health, visa, or others.

While we will stick to a maximum of 50 participants per event we are at the same time planning to run three to four events per year instead of two. In addition we also aim to expand towards not running hackathons for faculty and research staff, but instead train them on how to host their own local events. We realize that our own resources are limited. To grow beyond running three to four events per year, and to have a more profound impact on HPC workforce development we aim to train others to run hackathons using the HackHPC model. The first of these events is currently in planning and will take place later this year. For this we are developing training materials and seminars. This paper can be perceived as an initial step.

References

[1] N. A. of Sciences, *Future directions for NSF advanced computing infrastructure to support US science and engineering in 2017-2020*. National Academies Press, 2016.

[2] N. R. Council *et al.*, *Getting up to speed: The future of supercomputing*. National Academies Press, 2005.

[3] C. Connor, A. Bonnie, G. Grider, and A. Jacobson, "Next generation hpc workforce development: the computer system, cluster, and networking summer institute," in *2016 Workshop on Education for High-Performance Computing (EduHPC)*. IEEE, 2016, pp. 32–39.

[4] J. Powell, L. B. Hayden, A. Nolte, J. Herbsleb, E. P. P. P. Than, M. Wong, R. Kalyanam, K. Ellet, S. Pamidighantam, K. Traxler *et al.*, "An analysis of the pearc 2018 science gateways community institute hackathon: Lessons learned," 2018.

[5] E. P. P. Pe-Than, A. Nolte, A. Filippova, C. Bird, S. Scallen, and J. D. Herbsleb, "Designing corporate hackathons with a purpose: The future of software development," *IEEE Software*, vol. 36, no. 1, pp. 15–22, 2019.

[6] J. Porras, A. Knutas, J. Ikonen, A. Happonen, J. Khakurel, and A. Herala, "Code camps and hackathons in education-literature review and lessons learned," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[7] A. Nolte, L. B. Hayden, and J. D. Herbsleb, "How to support newcomers in scientific hackathons-an action research study on expert mentoring," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–23, 2020.

[8] A. Nolte, E. P. P. Pe-Than, A.-a. O. Affia, C. Chaihirunkarn, A. Filippova, A. Kalyanasundaram, M. A. M. Angarita, E. Trainer, and J. D. Herbsleb, "How to organize a hackathon–a planning kit," *arXiv preprint arXiv:2008.08025*, 2020.

[9] J. Powell, L. Bailey Hayden, A. Cannon, B. Wilson, and A. Nolte, "Organizing online hackathons for newcomers to a scientific community–lessons learned from two events," in *Sixth Annual International Conference on Game Jams, Hackathons, and Game Creation Events*, 2021, pp. 78–82.

[10] E. P. P. Pe Than, J. Herbsleb, A. Nolte, E. Gerber, B. Fiore-Gartland, B. Chapman, A. Moser, and N. Wilkins-Diehr, "The 2nd workshop on hacking and making at time-bounded events: Current trends and next steps in research and event design," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–8.

[11] A. Nolte, E. P. P. Pe-Than, A. Filippova, C. Bird, S. Scallen, and J. D. Herbsleb, "You hacked and now what? -exploring outcomes of a corporate hackathon," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 1–23, 2018.

[12] W. Mendes, A. Richard, T.-K. Tillo, G. Pinto, K. Gama, and A. Nolte, "Socio-technical constraints and affordances of virtual collaboration - a study of four online hackathons," *arXiv preprint arXiv:2204.12274*, 2022.