

Letter from Chair Marianetti to SRCPAC in lieu of meeting
Monday, December 7, 2020

Dear SRCPAC Members,

SRCPAC traditionally meets every Fall and Spring. As we have no pressing issues this Fall, I am writing to update you on items of interest, in lieu of yet another Zoom meeting. I will detail the following items:

- Our **next HPC** Cluster
- **Habanero** move
- Google Cloud Platform (GCP) for Research
- RCS Consulting
- **Foundations for Research Computing**

Next HPC Cluster

As voted on by you, the next shared HPC cluster will be named **Ginsburg!** We target go-live for **Ginsburg** in January 2021, which is only slightly behind schedule. This mild delay was due to a combination of the pandemic and a last minute add-on of \$300k worth of equipment from a new center grant within Columbia. **Ginsburg** will consist of 139 compute nodes (87 standard nodes, 30 high memory nodes, 18 RTX 8000 GPU nodes, and 4 V100S GPU nodes) and 570TB of storage. This compares to 110 nodes (92 Standard nodes, 10 High Memory nodes, 8 GPU Nodes) and 430 TB of storage of the initial Terremoto installation.

Habanero move

The **Habanero** cluster currently sits in the Zuckerman Institute data center. It will reach its four year end-of-life in December 2020. In the Spring, our solicitation of interest in the opportunity to extend the life of Habanero by one year garnered overwhelming participation. As a reminder, the highlights are:

- RCS will need to relocate Habanero from the ZIDC to CUIT's Morningside data center.
- A charge of \$250 per node will cover the move, installation, and warranty.
- If you chose not to pay to have your nodes moved, they will become a part of the edu/free tier.
- Participants will still have the same rights and priorities to their nodes.
- The move will occur in February 2021 which will start the year extension.
- The charge for the extended year will be considered 'services provided by CUIT', not 'equipment'.

We would like to thank Raj Bose and the ZIDC staff for their alliance and collaboration in hosting Habanero for the past four years.

Google Cloud Platform (GCP) for Research

CUIT can now provision you for Google Cloud Platform (GCP) for teaching and learning and research. GCP offers compute, storage, databases, servers and other services. Benefits include:

- CUIT's enterprise agreement and BAA
- Automated project provisioning
- Access with your Columbia UNI and password
- Multifactor authentication
- Chartstring payment
- Built-in security and privacy controls
- GCP training and consulting is available

For more information, see <https://cuit.columbia.edu/gcp>.

RCS Consulting

As a reminder, you can consult with CUIT's **Research Computing Services** team (**RCS**) on your research computing needs. Whether you have questions about cloud computing, on-prem HPC, or external resources such as XSEDE, please reach out to rcs@columbia.edu.

Foundations for Research Computing

The Foundations for Research Computing program has had to adapt to the pandemic by taking all in-person training online and updating offerings for a changing environment. This update on the program will share how the program has shifted in response to the events of 2020, and also how it has grown in significant ways despite these challenges. As always, we welcome any follow-up questions or discussion about Foundations for Research Computing.

As this is a longer update, we have broken it into three sections to focus on specific areas of interest:

- *Remote Transition*
- *Partnership Pilots*
- *New Programming*

Remote Transition

In mid-March, despite adverse circumstances and little time to prepare, Foundations for Research Computing instructors rose to the challenge and offered a full two-day Foundations for Research Computing Bootcamp over Zoom. Helpers who would normally circulate in the room assisted through chat, and CUIT was instrumental in setting up accounts and technology on short notice to accommodate the 56 researchers who were able to participate during this difficult time. While the number of participants was smaller than originally planned, this number still represents an increase from the 45 researchers served at the March 2019 bootcamp.

Following the March 16-17 bootcamp, all Foundations for Research Computing training has taken place online. Since this transition, the program has offered 24 events for 583 researchers, including the scheduled two-day bootcamp in August. This is a decrease from the 759 researchers trained over this period in 2019. This difference is mostly attributable to a reduction in the "intensives" category of full-day events for intermediate researchers, and we are optimistic that a reconfigured version of our intermediate intensives can resume in the fall.

Since March, Foundations for Research Computing has sought out best practices for online teaching, and developed or experimented with new practices in cases where guidance was not available. The Carpentries has recently released additional guidance on conducting technical workshops using remote technology, and for the planned January bootcamp the program will pilot a four-day, half-day bootcamp format following their proposed model. Despite challenges associated with online learning, feedback for online events has remained positive, and our August two-day bootcamp received a Net Promoter score in the "excellent" range according to a survey of participants. Compared to this period in 2019, participation in two-day introductory bootcamps since March has been approximately the same: 169 in 2019 and 171 in 2020.

In addition to our standard programming, Foundations for Research Computing trialed a Curriculum Innovation Grant (CIG) program last spring. This program is targeted at creating specialized and mid-scale mini-offerings. CIG offered grants to seven recipients to create and teach a technical workshop or workshop module. Though some planned components of the program could not be completed, five of the seven CIG recipients have taught a technical workshop, and most have served as helpers or otherwise participated in Foundations for Research Computing. We are currently evaluating the effectiveness of the CIG program and will make recommendations in the Spring for proposed modifications.

Partnership Pilots

In its third year of operation, the Foundations for Research Computing program is experimenting with partnering with several groups around the university to enable the partners to independently run technical training based on the Foundations for Research Computing model. In FY21, the program is piloting three partnerships with specific groups at Columbia: the Division of Cardiology at CUIMC, the Department of Mechanical Engineering, and a collective of interested faculty in the humanities. Foundations will provide support, resources, and expertise. Specifically, Foundations has prepared a packet of materials, including communications templates, forms, and checklists, to share with partners. The program also provides initial planning sessions, specific strategic support for troubleshooting issues, and assistance in finding additional instructors. Partner groups receive up to two Carpentries training slots gratis, with an option for additional instructors to be trained at cost, and partner instructors are expected to participate in the larger instructor community.

These pilots are in an early stage, but have already seen some success. On August 19-20, Mechanical Engineering ran a bootcamp on the Foundations for Research Computing model overseen by Associate Professor Arvind Narayanaswamy. The bootcamp, "UNIX, Git, and Python For Mechanical Engineers," targeted an incoming class of 110 MS students in

mechanical engineering. A primary goal for these partnerships has been to leverage existing program resources to create a broader impact on campus. In addition to providing Carpentries training, resources such as checklists and communications templates, and light logistical support, Foundations for Research Computing connected Arvind with two experienced graduate student instructors who received a stipend for their work. As a partner, Arvind has significant experience in technical pedagogy, and it remains to be seen whether all such partnerships can be run as independently. A bootcamp for the Reilly Lab in the Division of Cardiology will be run in January in cooperation with Roger Lefort in Research Compliance and Training. We look forward to giving a full report on these partnerships at the Spring SRCPAC meeting.

New Programming

In coordination with CUIT's launch of the Google Cloud Platform (GCP) service, CUIT partnered with Foundations for Research Computing in offering two new workshops in a series on cloud computing this fall: Introduction to Cloud Computing for Research and High Performance Computing on Google Cloud Platform. We anticipate that these workshops will supplement the existing workshops offered by CUIT that prepare researchers to access cluster resources.

This semester, the program has begun piloting a Research Computing Reading Group targeting researchers with intermediate to advanced experience in research computing. The group, convened by a graduate student paid as a Libraries Digital Intern, discusses current or historical technical papers in areas such as natural language processing and machine learning that are relevant to research computing. We look forward to presenting this pilot for evaluation in the spring for a determination of its impact on researchers and whether it will offer continuing value to the program.

In 2020, despite challenges, the Foundations for Research Computing program continues to offer computational training that allows Columbia researchers to access shared resources at the university. Thank you to SRCPAC for the committee's continued support. Please reach out at any time with questions, either to rcfoundations@columbia.edu or to the Foundations for Research Computing program coordinator at patrick.smyth@columbia.edu

I look forward to seeing you early next semester. In the meantime, questions and comments welcome to SRCPAC at srcpac@columbia.edu.

Chris Marianetti, PhD
Chair, Shared Research Computing Policy Advisory Committee (SRCPAC)
Associate Professor, Department of Applied Physics and Applied Mathematics
www.columbia.edu/srcpac