

Fall 2017 SRCPAC Meeting Minutes
Monday, December 4th, 2017

Attendees: Chris Marianetti (Chair), Kyle Mandli, Marc Spiegelman, Tian Zheng, Andreas Hielscher, Har-
men Bussemaker, Michael Burke, Lorenzo Sironi, Kathryn Johnston, Paul Blaer, Gaspare LoDuca, Alan
Crosswell, Victoria Hamilton, Tom Chow, Rob Cartolano, Maneesha Aggarwal, Halayn Hescock, Barbara
Rockenbach, George Garrett, Sander Antoniadis, Mark Newton, Rob Lane, Fred Palm, Marley Bauce, Dali
Plavsic, Eric Vlach, Jennifer Brown, Raj Bose, Jochen Weber, Josiah Trinidad-Christensen, Michael Weis-
ner, Michelle Benson, Dan Mechanic

Chair Chris Marianetti opens the meeting by introducing himself as the Chair of SRCPAC. The meeting attendees then introduce themselves.

Habanero Update

Chris introduces Kyle Mandli, Chair of the Habanero Operating Committee. The Habanero Operating Com-
mittee meets semiannually to review the business rules, and determine whether to make any adjustments.
Habanero launched with rules nearly identical to the Yeti business rules, save for necessary changes im-
posed by the new software. These rules were re-reviewed in the Fall 2017 Habanero OC meeting. The
meeting is open to all Habanero users; prospective attendees are encouraged to contact rcs@columbia.edu
with questions.

Kyle then asks George Garrett, Manager of Research Computing Services to review Habanero status. Kyle
also acknowledges George's new role as Manager, taking over for Rob Lane. The Committee warmly wel-
comes George to the new role.

George begins by showing the Committee a photo of the new Habanero cluster pre-expansion, which in-
cluded four high-density racks. The Habanero Part 2 expansion adds 18 new research groups (some repeat
customers) that have altogether bought 80 new nodes (or, 1,920 cores); in total, Habanero will include 302
nodes and 7,248 cores. The new cluster will include the same CPUs, specifically: 58 standard servers, 9
high memory servers, and 13 GPU nodes. The machine includes two types of GPU nodes – K80 and P100
(P100s are new for the expansion, offering additional features like higher memory and more advanced half-
precision performance, which is optimal for machine learning research).

Habanero also has a new visualization server to enable users to visualize data very quickly over their inter-
net browser. George encourages researchers to test this feature out, as it will reduce the amount of data
necessary to download onto the local desktop.

Shared HPC – since its inception via Hotfoot in 2009 – has grown steadily. In the current Habanero cluster,
44 research groups have purchased nodes. A Rental Tier includes 7 users, and a Free Tier includes 45 users.
Altogether, Habanero has registered 840 user accounts. There is an Education Tier funded by A&S and
SEAS; this year, four courses spanning computer science, biology, and environmental science have used
the Educational Tier that A&S and SEAS leadership funded. Instructors interested in using this Tier for
their courses are encouraged to contact rcs@columbia.edu to discuss options.

Much of Research Computing Services support is communicated via email, in addition to monthly office
hours in the Science & Engineering Library (3-5pm on the first Monday of each month).

RCS also holds two types of Information Sessions. Support staff will come to departments, individual labs
or research groups, which can be scheduled via hpc-support@columbia.edu to discuss specific issues. Al-
ternately, already-scheduled information sessions cover topics including a cluster overview, using the

scheduler, and identifying business rules. All new and interested users are welcome and encouraged to attend.

CUIT and the Libraries offer a number of workshops on three topics: Intro to Linux, Intro to Scripting, and Intro to HPC. These workshops are typically held once in the Fall and once in the Spring semesters; Spring workshops will soon be scheduled.

Jochen asks how many cores participation in the Rental Tier provides. George will follow-up, although believes the standard guarantees four cores, and more resources if and when they become available. Chris notes that current users of the Rental Tier are happy and have ended-up using more than their fair share.

Chris thanks George for his presentation, and notes that RCS has been doing a great job and that everyone thinks RCS is wonderful.

Raj seeks clarification about the Yeti retirement. George responds that 66 Yeti compute nodes remain until March 2019. Support is still available for Yeti.

Announcement of Training Subcommittee

Chris states that in addition to the CUIT and Libraries workshops that George mentioned, many other short courses and online resources exist. The traditional curriculum does not always cover research and HPC computing skills that students must know. Moreover, different disciplines may need different skills prerequisites, and these topics change frequently enough so that standard courses may not be sufficient. SRCPAC has created a Training Subcommittee – chaired by Marc Spiegelman – to organize and structure the compendium of offerings across the University and make recommendations for next steps.

Marc explains that the Subcommittee will focus on helping incoming graduate students get up to speed so that formal courses need to spend less time on this remedial training. The Subcommittee aims to complement the existing curriculum.

Harmen asks whether future plans are to provide course credit for students. Marc explains the goal now is to allow students to do better in for-credit courses. No plan exists to develop for-credit courses under the auspices of the Subcommittee.

Lorenzo notes this subject is very vast, and asks whether the Subcommittee is soliciting feedback on topics to focus upon. Marc answers the core foci will be Python and HPC and Linux. Rob Cartolano suggests to all SRCPAC members that they pass along advice based on experience of what is in most needed by the graduate student body. Tian recognizes that it is very hard to design an informal curriculum that works for everyone. She believes it is the Subcommittee's task to design this curriculum plan. She notes that posting some of the content posted online using videos and other media would be optimal. Marc responds that the goal is a hierarchy of content, mapped to specific courses. All interested individuals should contact Marc directly at mspieg@ldeo.columbia.edu.

Plans for New Cluster

In Spring 2018, SRCPAC will initiate the process of purchasing a new cluster to replace Habanero with orders due in June 2018 in order to accommodate new faculty recruits so they can potentially use a portion of their start-up package to buy into the cluster. Halayn notes that a major renovation of the existing Data Center should be completed in time to house the new cluster once it is live in late 2018.

Chris raises the topic of how best to exploit the Cloud, especially as prices gradually drop. While the Cloud currently will not work for everyone as none of the Cloud options have Infiniband and use of massive datasets may be cost-prohibitive, however, many researchers can do much or all of their work on the Cloud. It behooves us to continue exploring available options both in terms of the technology and the financial model. Currently, researchers can use a Columbia account to purchase Cloud computing time. Halayn notes that RCS has also hooked Habanero up to multiple Cloud-based resources to explore the technical and managerial feasibility of, for example, bursting the shared cluster to the Cloud during periods of great demand. Critical questions include how to determine which jobs would move to the Cloud and the funding model. Chris recognizes the import of continuously refining our thinking on the topic.

Raj mentions that Columbia is part of the Coalition for Academic Scientific Computation (CASC) where the Cloud is already a hot topic. Once this Consortium releases its report, Raj will share it with the larger SRCPAC community.

Halayn mentions the need to form a Faculty RFP Committee to help guide the new cluster. In January, SRCPAC will ask for faculty volunteers to participate on this Committee, which will require four meetings across the Spring 2018 semester. This Committee was absolutely critical for guiding CUIT in for the design and purchase of the Habanero cluster, and our goal is to replicate this success with the upcoming Committee. Any Faculty interested in participating on the RFP Committee should email srcpac@columbia.edu.

Kyle asks whether we have conducted a survey to understand researcher needs. Chris recognizes that surveys are typically conducted prior to the hardware purchase, but agrees this should happen soon. Sander notes that Habanero benefits from being quite homogenous, however GPU nodes become cost-prohibitive. Therefore the committee should consider whether we should have a separate GPU cluster.

Harmen asks whether it is feasible to create a forum for requests and comments: Currently researchers only interact with RCS but do not know of the other requests emerging out of HPC users. Users might find it useful to know what requests are emerging from others. RCS is exploring options for different communication methods such as Slack.

Lorenzo asks what fraction of Habanero jobs or CPU hours *could* be transferred to the Cloud. RCS cannot give a precise number, but George believes there is a large percentage that could transition, given that many users do not currently exploit the Infiniband resources.

Rob Lane notes that research computing is not technology- or funding model-independent. Researchers doing something on one platform may not do the same thing if moved to another platform. For example, under the purchase option, a user may run as many iterations as the user wishes since the price is fixed. Conversely, typically the Cloud is used on a pay-for-use model so researchers become much more sensitive to iterations. So activity may not translate solely based on prior experience.

Tools for Researchers

Maneesha describes a range of resources that CUIT and the Libraries host to aid researchers. Code Ocean is a repository for your code, which allows you to keep it running, share with others across institutions, collaborate with others, and keep it safe once it is finalized. CUIT will also be further evaluating the CUNIX environment. Storage will be free without any cost to researchers; the goal is an enterprise license in approximately three years. Chris asks whether basic version control like via GitHub should also be offered first before Code Ocean. Alan Crosswell has this on his to-do list, and will update SRCPAC soon.

CUIT is also collaborating with other universities to procure a group license for Globus for moving data to Habanero.

CUIT and the Libraries are also evaluating new technologies, and are keen to know of emerging technologies they should explore further. Send suggestions to her at maneasha@columbia.edu.

The Office of Research Compliance and Training has assembled a Research Data website, which provides a one-stop shop for data retention and security, as well as outlining storage options and ownership issues. This website outlines all policies and resources in one single page: <https://research.columbia.edu/research-data-columbia>.

Publications Reporting

A compendium of publications emerging out of work hosted by Habanero and/or Yeti is very important in order to persuade research leadership to further support these efforts, as well as to demonstrate to the NIH the value of the G20 award for a Research Data Center. Gaspare also notes that he gives these numbers to the Board of Trustees, and they recognize and appreciate this. All SRCPAC members will soon receive an email asking them to report this information. Publications can be reported at any time by emailing srcpac@columbia.edu. Rob Cartolano suggests creating an easy URL for researchers to insert in their papers to note the use of the Shared Research Computing Facility as this would both aid researchers and create a way to scan publications to identify whether they should be included in the list.