

Shared Research Computing Policy Advisory Committee (SRCPAC) - Spring 2016 Minutes

April 12, 2016, 2:30 p.m. - 4:00 p.m.

523 Butler Library

Attendees: Ryan Abernathy, Marley Bauce, Alex Bergier, Paul Blaer, Jennifer Brown, Michael Burke, Alan Crosswell, George Garrett, Victoria Hamilton, Andreas Hielscher, Roslyn Hui, Kathryn Johnston, Wojciech Kopczuk, Rob Lane, Gaspare LoDuca, Chris Marianetti, Mahdad Parsi, Dali Plavsic, Amy Nurnberger, Ingrid Richter, Matthew Rhodes, Bruno Scap, Brent Stockwell, Julien Teitler, John Villa, Eric Vlach, Michael Weisner, Delia Wu, Tian Zheng

Welcome & Introduction

Kathryn Johnston, Chair of SRCPAC
Chair, Department of Astronomy

Kathryn Johnston opens the meeting by welcoming all attendees, and asks everyone to introduce themselves. Each meeting attendee has been provided with a printed copy of the meeting [agenda](#).

Faculty Recruitment Policy

Kathryn Johnston, Chair of SRCPAC
Chair, Department of Astronomy

Kathryn Johnston mentions that there is finite capacity on campus for servers. If you are recruiting someone or applying for a grant to obtain compute resources which would be shared, please get in touch with Victoria Hamilton who will coordinate with CUIT to determine whether there is sufficient capacity. Ryan Abernathy notes that the current HPC Expansion should be used as a faculty recruiting tool since it provides faculty access to a resource larger than they can afford themselves, without having to worry about the purchase and installation process, finding space and energy, or system administrative team.

Electronic Lab Notebook

Halayn Hescock, Director of Research Computing, CUIT

The University has procured an enterprise-wide license and BAA for LabArchives' Electronic Lab Notebook software. This will be a free service which is paid for by CUIT and the Libraries and includes a classroom edition. The final stages of CUMC certification for HIPAA data are underway. Early adopters are using the system now. Brent Stockwell has been working with LabArchives to have certain features added and the vendor has been very responsive. Please contact Halayn Hescock for more information or if you're interested in becoming an early adopter. Amy Nurnberger, who is on the implementation team, notes that the Open Science Framework is another tool for collecting and tracking research data.

Update on Amazon Web Services

Gaspare LoDuca, Chief Information Officer
Vice President, CUIT

CUIT has signed an enterprise agreement with Amazon Web Services and developed a process to link existing AWS accounts to the enterprise agreement. Linked AWS accounts will get the benefit of protection in terms of IP, security, and compliance, and the ability to pay for AWS services with a Chart String rather than a P-Card, in compliance with the University's procurement policy. You can sign up to [link your AWS account](#) now and an automated tool for sign up will be available in about 10 days.

Research Computing Services will provide some advisory, system, and software support for researchers using AWS services. Rob Lane says that the intent is that in the future RCS will provide standard images with standard tools, as well as other software that is requested by researchers.

Alan Crosswell notes that there is no free tier available under the AWS enterprise agreement and that CUIT is waiting for AWS to come out with an educational offering that doesn't require a credit card.

Education Tier: Teaching on Yeti

Matthew Rhodes, Term Assistant Professor, Department of Biology (Barnard)

Matthew Rhodes is teaching an Intro to Genomics and Bioinformatics course at Barnard and has around 14 students working on Yeti on a regular basis both during class and outside of class. Yeti has been helpful for his class in a few different ways. Students log on to Yeti and have been introduced to Linux and the HPC platform. They've downloaded large databases and are using complex software. The students upload their assignments to Yeti and are starting to use the cluster for independent projects to analyze large datasets.

Please contact RCS at rcs@columbia.edu if you are interested in using Yeti for educational purposes.

The Yeti education tier is supported by the use of grant funded compute nodes. The HPC expansion will not have grant funded shared nodes. Chris Marianetti suggests that Deans should be asked to each contribute a few nodes for classroom education.

Ryan Abernathy then asked about the scope of support for education activities on Yeti. Rob replies that RCS will provide software installation and system support to the instructor and/or TAs, who will act as intermediaries to the students.

Spring 2016 Yeti Operating Committee Meeting

Rob Lane, Manager of Research Computing, CUIT on behalf of
Greg Bryan, Chair of the Yeti Operating Committee
Professor, Department of Astronomy

Rob Lane stood in for Greg Bryan who was unable to attend. At the Spring OC Meeting, Yeti usage and knowledge transfer were discussed, but no rules or policies were changed. Minutes and slides are available on the [SRCPAC website](#). All Yeti users are invited to join these meetings.

Brent Stockwell, who is deciding how many nodes and what type of nodes to purchase in the HPC expansion, asks what would be the appropriate formula to calculate fairshare given that GPU servers are two and half times the price of regular nodes. Should fairshare be calculated by number of nodes or by cost? Chris Marianetti notes that you get priority access to the hardware you pay for. Kathryn Johnston suggests to take the discussion offline. The operating committee for the HPC expansion will be formed after all the purchases have been made.

HPC Expansion

Chris Marianetti, RFP Committee Member
Associate Professor, Department of Applied Physics & Applied Mathematics

Chris Marianetti summarizes the HPC Expansion Request for Proposal (RFP) Committee process and presents these [slides](#). Chris mentions the faculty members involved, the meetings to review technical requirements, finalize the RFP and distribute it to vendors, and to choose vendor finalists and then winners. A list of key technical decisions, compute and storage vendors and winners, and a menu of server pricing was provided. Chris pointed out that the new Nodes had 12 cores each versus the 8 core nodes of Yeti.

Chris mentions that it could be a good goal to see if we can get on the [Top500](#) list of supercomputers. This would look good and could help rally the troops to join the HPC Expansion.

Victoria Hamilton emphasizes the fact that the storage price has dropped significantly and is now \$250 per TB versus \$2000 the last round. She then thanks Gaspare LoDuca who got RCS supported so that CUIT could eliminate the \$1000/node support fee that existed for Yeti. Chris Marianetti adds that SEAS will provide 25% matching for SEAS faculty' orders.

Ryan Abernathey points out that this storage system is very fast because it's connected over Infiniband and uses an optimized parallel file system which is scalable. He notes that it is much faster than Yeti's storage system which uses NFS.

Kathryn Johnston suggests that researchers put orders in early as there is a snowball effect that could spur other researchers to buy-in. Chris Marianetti notes there are stragglers who should be corralled to join the expansion. Gaspare LoDuca offers to help put together a brochure to get the word out. Ryan emphasizes that free support provided for this system is a big selling point.

Rob Lane mentions that RCS is available to come out and talk to people about the expansion and answer any questions. HPC Expansion Information Sessions will be held for three straight Thursdays at noon on 4/14, 4/21, and 4/28 in 105 Low Library.

Orders for the HPC Expansion are due Monday, May 2nd.

Please contact Rob Lane at rcs@columbia.edu if you have any questions about the HPC Expansion and see the [HPC Expansion webpage](#) for technical, pricing, and ordering details.

Kathryn Johnston announces that the HPC Expansion cluster will be named Habanero.

Yeti Publications Reporting

Kathryn Johnston, Chair of SRCPAC
Chair, Department of Astronomy

Kathryn asks Yeti users to submit a list of publications that have made use of Yeti as this helps communicate and demonstrate the success of the cluster.

Closing Remarks

Kathryn Johnston, Chair of SRCPAC
Chair, Department of Astronomy

Kathryn Johnston opens the floor for topics that anyone would like to discuss.

Ryan Abernathey brings up [Software Carpentry](#), which is a group that does instruction on basic skills for research computing, such as introductions to Unix, Python, and R. His department thinks this might be worthwhile and it's not very expensive. Amy Nurnberger notes that her group has also been interested in this for some time and Rob Lane mentions that colleagues in peer institutions have said very positive things about Software Carpentry. Sander Antoniadis mentions that Continuum Analytics has also offered to come teach about Anaconda Python. Paul Blaer adds that the CS department offers 1 credit courses in programming languages and a course about using Unix which will be offered this Fall. Kathryn suggests that this useful discussion be pursued in email.