

Fall 2018 SRCPAC MeetingFriday, December 7th, 2018**Welcome & Introductions**

Chris Marianetti called the meeting to order and asked that each person introduce themselves.

High-Performance Computing Update

Kyle Mandli (attending by phone) made welcoming remarks and asked George Garrett to give the update. Highlights included:

- *Terremoto*, SRCPAC's latest cluster went live this week. Most groups are on board and the rest will be shortly. With a five-year life, rather than the previous standard of four, Terremoto has 50% more memory than Habanero and faster CPUs.
- *Habanero*'s participation continues to grow with approximately 400 users added since the Spring SRCPAC meeting for a total of about 1,400. Four more courses have been added to the Education Tier since the Spring. 80M core hours to complete 2.7 million jobs. It is running at about 85% utilization, so a busy popular cluster. Full retirement in 2021.
- *Yeti* is half retired, with final retirement in March. Utilization has been low as most users have migrated to other clusters.
- The Data Center Cooling expansion now scheduled to be completed in February, 2019. It will create 500KW of cooling capacity.

Foundations for Research Computing Update

Marc Spiegelman brought the committee up-to-date on the progress of the Foundations program, highlighting in particular the pre-semester boot camps and monthly lectures.

Foundations will shortly announce the January 17-18 boot camp. To honor the support of the program, Ph.D. track applicants will receive priority.

Marc and Barbara noted that they anticipate announcing the hiring of a Coordinator shortly who will support further expansion. Topics of discussion included:

- *How to expand the capacity for instruction*. Already six additional instructors have undergone Carpentries Training, with more interest from CUIMC and possibly SPS.
- *How to address discipline-specific and intermediate needs*. Schools and departments have expressed a desire to expand offerings, and some are experimenting with doing so.
- *How to maximize the effectiveness of workshops and office hours*. Several experiments are underway and others planned.

Research Data Survey

Barbara Rockenbach updated the committee on the Libraries' intent to survey needs for research data (and storage). This survey results from questions raised by the Provost's Academic Committee for the Libraries (PACL). The Libraries tested the clarity of the survey by submitting it to SRCPAC members; Barbara thanked anyone who took it for their comments. The survey took as its base a similar survey that Yale used.

The Libraries have spoken with Yale about their experiences and results. One surprise to Yale was the need for much more education and publicity about the resources that existed.

Potential for New Cloud and GPU Subcommittees

Chris raised for discussion and input the question of whether or not SRCPAC should form subcommittees around (a) Cloud resources and/or (b) GPU clusters. Chris indicated while both topics have been of interest for some time, he was reluctant to form subcommittees until clear charges existed.

CLOUD Chris pointed out that SRCPAC is always struggling with what role(s) the Cloud will play for research computing.

- CUIT has made use of professional cloud providers easier. Columbia now has an Enterprise license with AWS (discussed further below) that is HIPAA approved and a direct connect has been operating for almost a year. Columbia also has arrangements with Google and Microsoft.
- Should Columbia's demand for HPC outstrip the Data Center, bursting Columbia's HPC system to the Cloud or establishing a separate "HPC in the Cloud" remain alternatives. Technical details on how to accomplish those alternatives, how to allocate jobs to make the best use of local and Cloud capacity, and the financial model(s) would be non-trivial to work out, but doable.
- Gaspare LoDuca suggested that perhaps it made sense to create a subcommittee for External Resources. Chris agreed that "Cloud" should also mean some supercomputing centers.
- Raj Bose indicated that some professional societies --- recognizing that many universities were grappling with variations of this conversation -- were publishing papers re the academic use of cloud. Raj offered to pull together a short list and provide to SRCPAC. Chris applauded this idea as these papers could educate SRCPAC and help forestall reinventing the wheel.

GPU

The cost of enterprise grade GPU's jumped for Terremoto due to a number of market factors. Should SRCPAC form a committee to explore the alternative of consumer grade GPUs?

Conversation ensued. Some Columbia academic users are experimenting with consumer grade clusters and several members said they would report back on capabilities, downtime/burnout and staff time.

CUIT Updates

Secure Data Enclave – Michael Weisner, RCS Systems Engineer

Michael discussed the Secure Data Enclave (SDE), a secure, remotely accessible, virtual Windows 10 desktop environment to store and collaboratively analyze PII and PHI data as an alternative to traditional cold room computing environments. <https://cuit.columbia.edu/sde>

After a period of piloting, SDE is now available university-wide. It has certain statistical software and already possess several data sets. For example, CUIT has worked with Libraries to make the SDE the primary source for EuroStat Data. Halayn reminded the audience that SDE is for working datasets; it is not meant for storage.

GLOBUS – George Garrett



Shared Research Computing
Policy Advisory Committee

George indicated that Columbia now has an Enterprise license for this web-based interface for HPC data transfer between users within and outside Columbia. GLOBUS already has 200 users at Columbia and 100 endpoints.

AMAZON WEB SERVICES (AWS) – Jimmy Chiong, Lead Infrastructure Engineer

Columbia has negotiated an Enterprise Agreement with AWS. It allows purchasing under the University's procurement system, IP protection, liability and other protections. Importantly, Columbia has also signed an agreement to enable Columbia users a way to host PHI, while maintain HIPAA compliance. Jimmy answered a number of questions and there is further description and an FAQ at this site. <https://cuit.columbia.edu/aws>

Closing Remarks

Chris Marianetti thanked everyone for their input and closed the meeting.