**Fall 2019 HPC Operating Committee Minutes**

November 12, 2019

**Attendees**: Kyle Mandli (Chair), Sander Antoniades, Cesar Arias, Jacqueline Austermann, Tom Chow, Gus Correa, Zachary Fuller, George Garrett, Lokke Highstein, John Pellman, Dali Plavsic, Axinia Radeva, Michael Shelter, Marc Spiegelman, Michael Weisner

**HPC OC Committee and Executive Committee Overview**

Kyle Mandli, Chair of the HPC Operating Committee, opens the meeting by welcoming members to the Fall 2019 meeting, and after introductions by attendees, turns the floor over to George Garrett, Manager of Research Computing Services within CUIT.

George gives an overview of the HPC Operating Committee, noting that this an open meeting, chaired by Kyle Mandli, that typically meets twice a year to share HPC operating updates and provide a forum to provide guidance, share feedback, and discuss policy.

The newly resurrected HPC OC Executive Committee exists to make decisions swiftly between OC meetings when critical to do so. The Executive Committee membership consists of Ryan Abernathey, Jacqueline Austermann, Andrei Beloborodov, Stefano Fusi, Kyle Mandli, Lorenzo Sironi and Alexander Urban.

**HPC Clusters - Overview, Usage, and Updates**

There are four ways to participate in the University’s high-performance computing initiative: Purchasing nodes, Renting nodes, the Free Tier (with limited access), and the Education Tier (supported by the Engineering School and Arts and Sciences).

The Terremoto cluster was launched in December 2018 and is undergoing expansion in December 2019. Including the expansion, Terremoto participation includes 29 research groups, 325 users who have utilized 13.4 million core hours. Terremoto has a 5 year equipment lifetime. It consists of 137 compute nodes, 510 TB storage, Dell hardware, Dual Skylake Gold 6126 cpus and uses the Slurm job scheduler. Please refer to the slides for additional details.

Terremoto usage has ramped up, frequently spiking to over 90% utilization, and has averaged 70% utilization in the past 3 months, and 74% utilization over the month of October. Users primarily run small jobs, with the average job size of 17 cores. There are a handful of users who run jobs which utilize between 50 and 1000 cores. Other job statistics include average job walltime of 5.8 hours, average queue wait time of 1.8 hours, and median queue wait time of 24 minutes.

Gus Correa inquires as to why the clusters are not maintained longer than 5 years, since he has maintained a cluster at Lamont for around 10 years. He notes it is difficult for researchers to purchase equipment at numerous internvals because there is no guarantee that funding to do so will be available in the future. Dali Plavsic mentions that the newer clusters have speed improvements and extending maintenance contracts can be expensive. George adds thas there also power efficiencies with newer equipment and that we extended our cluster maintenance period from 4 years on Habanero to 5 years on Terremoto to move in this direction.

Habanero participation continues to grow and has 44 research groups or departments, a total of 1897 users since launch, and over 215 active users. A total of 18 courses have used Habanero as part of the Education tier.

Habanero 1st round equipment is set to retire in December 2020. George notes that there are some discussions that some of those retired nodes may be used in the future to support the Educational or Free HPC tier. Tom Chow asks about the status of this planning and George responds that these discussions are still in the early planning phases and has not been finalized.

OpenOndemand HPC web portal software, is available for testing on Terremoto. The portal enables interactive HPC such as submitting and monitoring jobs, connecting to the cluster, and transfering files all via a web interface. If interested in piloting this software with us, please contact rcs@columbia.edu for details.

The Data Center Cooling Expansion project was successfully completed in July 2019 and assures HPC capacity for the next several generations.

With regard to upcoming cluster planning, each year we typically sent out an announcement in April about upcoming cluster buy-in opportunities. In 2020, if there is sufficicient demand, we may hold an RFP and work with researcher guidance and vendors to design a new cluster. The purchase round would likely commence in late Spring 2020, with Go-live of the cluster in late Fall 2020. If you are aware of potential demand, including new faculty recruits who may be interested, please contact us at rcs@columbia.edu.

**Overview of Business Rules**

Any special usage requests should be sent to hpc-support@columbia.edu and RCS will contact Kyle Mandli, and if needed, the HPC OC Executive Committee, for approval.

Nodes owned by individual users have the fewest restrictions, with priority access for the dedicated node owners. Vice versa, nodes owned by other accounts carry the most restrictions.

For Public nodes, there are few restrictions, but nobody has priority.

The maximum wall time is 5 days on nodes owned by your group or on public nodes. Jobs that ask for 12 hours or less can run on any node.

Every job is assigned a priority: “Fair Share” is determined by target share (determined by the number of nodes owned by each account) and recent use (number of core hours used recently; calculated at the individual and group level). If recent use is less than the target share, the job priority increases; users using more of their target share see their priority decrease.

**Support**

The RCS team offers a number of support methods, including open office hours and group information sessions, as well as training. The [Foundations for Research Computing](https://rcfoundations.research.columbia.edu/) initiative was highlighted as one useful resource for upcoming workshops and other events.

**Training Workshops**

Upcoming Introduction to HPC workshops to be held on:

* **Tuesday, January 28th , 2020**: [Introduction to High Performance Computing](https://events.columbia.edu/cal/event/eventView.do?b=de&calPath=%2Fpublic%2Fcals%2FMainCal&guid=CAL-00bb9e24-6dd8bf9f-016d-db52faff-00000d8eevents@columbia.edu&recurrenceId=), 1:00 p.m. – 2:30 p.m.
* **Tuesday, February 25th, 2020:** [Introduction to Linux](https://events.columbia.edu/cal/event/eventView.do?b=de&calPath=%2Fpublic%2Fcals%2FMainCal&guid=CAL-00bb9e24-6dd8bf9f-016d-db4a8956-00000b59events@columbia.edu&recurrenceId=), 1:00 p.m. – 2:30 p.m.
* **Tuesday, March 3rd, 2020:** [Introduction to Scripting](https://events.columbia.edu/cal/event/eventView.do?b=de&calPath=%2Fpublic%2Fcals%2FMainCal&guid=CAL-00bb9e24-6dd8bf9f-016d-db4d4ea0-00000bbfevents@columbia.edu&recurrenceId=), 1:00 p.m. – 2:30 p.m.
* **Tuesday, March 10th, 2020**: [Introduction to High Performance Computing](https://events.columbia.edu/cal/event/eventView.do?b=de&calPath=%2Fpublic%2Fcals%2FMainCal&guid=CAL-00bb9e24-6dd8bf9f-016d-db4f79b9-00000c16events@columbia.edu&recurrenceId=), 1:00 p.m. – 2:30 p.m.

**HPC Publications Reporting**

Reporting research publications which used local HPC resources is critical for demonstrating to University leadership the utility of supporting research computing. Please report new publications utilizing one or more of these machines to srcpac@columbia.edu.

**Feedback and Discussion**

In light of the Habanero 1st round retirement in December 2020, Sander Antoniades suggests that RCS reach out to researchers who have nodes that will be retiring so that they can make the appropriate decisions about whether to purchase new equipment and can consider buying into the next cluster. RCS agrees and will reach out to all researchers with retiring equipment.