

SRCPAC MEETING – SPRING 2019

Date: Thursday, April 25, 2019

Time: 10:00 a.m. – 11:30 a.m.

Location: 523 Butler Library

Meeting Called By: Chris Marianetti, Chair of SRCPAC

WELCOME & INTRODUCTIONS

Chris Marianetti, Chair of SRCPAC

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

HIGH PERFORMANCE COMPUTING UPDATE

Kyle Mandli, Chair of the HPC Operating Committee

George Garrett, Manager of Research Computing

George presented an update on HPC. Regarding HPC governance, previously there was an operating committee for each cluster. This was redundant due to similarities between clusters and membership and has been consolidated to a single HPC Operating Committee to handle all clusters.

Cloud computing consulting was highlighted as a complimentary CUIT service to help onboard researchers to cloud computing and find the best for researcher requirements. This includes help with budget planning, imaging machines, and installing software with cloud computing services such as Amazon Web Service, Google Cloud Platform, and Azure. Many of these cloud computing services give away research credits (\$10k, \$40k, etc.) and CUIT can help guide you in applying for those as well.

Habanero cluster utilization continues to be heavy. There was a memory enforcement change in late 2018 which had the effect of reducing core hours utilized, so there was a bit of a drop off. Additionally, some users began migrating to Terremoto. The Habanero average daily utilization is between 75-85%.

Terremoto usage ramped up over the last couple months and now averages around 80% utilization. Purchasing for Terremoto Round 2 is in mid-May 2019.

Singularity software, made specifically for HPC, allows Docker containers to run on HPC clusters, and is now available on Habanero and Terremoto. Singularity is easy to use, and more secure than Docker, and allows seamless transfer from laptop to HPC to cloud to other national HPC resources. Singularity enables rapid deployment of new versions of software and complex software stacks, as well as support for other flavors on Linux operating systems.

CONSUMER GPU CLUSTER EXPERIENCE

Sander Antoniades, Senior Research Systems Admin, Zuckerman Institute

Jochen Weber, Scientific Computing Specialist, Zuckerman Institute

Chris first provides some background to this topic, sharing how faculty asked why we don't provide a consumer GPU cluster. The current enterprise GPU servers are expensive and have expensive CPUs and other infrastructure that may not be needed in all cases. There are a lot of complicated factors, and we invited the Zuckerman Institute Research Computing group to speak about these.

Sander shares an overview of consumer versus enterprise GPUs, and shares Zuckerman Institute's experience running a consumer GPU pilot for Zuckerman researchers. Rajendra Bose states that the whole goal is to not replicate the resources already available across the University but find and fill gaps identified by researchers. Sander explains that enterprise cards are the typical ones provided by Columbia's shared HPC clusters.

The Zuckerman Institute wanted to complete a pilot and used an interested research group to test out to determine if this is a service they want to provide across the board. These GPU have much greater power demand and produce a lot of heat.

Bob Mawhinney inquires as to how quickly the consumer GPUs die and how much are they are being run. Sander and Jochen respond that some run for days on end but not weeks or months – users log in run model and then it stops. Bob shares his experience with consumer GPUs. The cost performance was good, but after 3-4 years they began failing. Bob also asks as to the research reproducibility since the consumer GPUs do not have error correcting code memory, as do the enterprise GPUs. There was also a discussion of using GPUs in the cloud which can be a good, cost effective option in many circumstances, especially when needs are bursty.

Chris has received feedback in the past encouraging us to explore the benefits and disadvantages of supporting consumer GPUs. He states that for SRCPAC to dive into consumer GPUs right now would be premature, and suggests as another option to try GPUs in the cloud.

FOUNDATIONS FOR RESEARCH COMPUTING UPDATE

Marc Spiegelman, Chair of the Advisory Committee

Patrick Smyth, Program Coordinator

Marc and Patrick present slides which report on recent activities of the Foundations for Research Computing and highlight future directions.

Foundations is working on expanding into CUIMC, and is available to all graduate students. The presentation moved on to a discussion of how to scale the Foundations initiative. Patrick mentions reaching untapped groups and said we're trying to maximize impact and reach of the Foundations programming.

Foundations is working on building a community of instructors and we will be training an additional 6 or 7 new instructors this June. Instructors have taken the lead in teaching many courses, including some courses uptown, Python workshops at the Business school, and a text mining workshop for around 50 people at the Columbia Population Research Center, to name a few.

Next year, we want to serve twice as many people, run a more "intermediate" level bootcamp. We are using feedback to figure out exactly what students want, and plan to have more half day intensives. Chris asks what is limiting factor in scaling up? Patrick responds that we now should have a good pool of instructors, but that perhaps Patrick's time will become a limiting factor in scaling up.

Patrick explains that they will vary the menu of offerings to see what works as well as experiment with different kinds of programming.

A general question was asked about the satisfaction level of bootcamp attendees. Patrick says that pre and post survey from students are taken - and overall feedback is good, but they have pointed to a few things – some attendees don't understand why they are learning things like Git, for example. Foundations can tune the curriculum to the audience, and more offerings at different levels would be helpful as well.

CUIT UPDATES - Globus

George Garrett, Manager of Research Computing

George provides an update about Globus and CUIT's Globus license. Globus transfers data and manages reconnecting even if you shut off system, and makes it easy to share data with researchers both within Columbia and externally with other institutions. It allows researchers to easily connect to cloud endpoints, as well on-premise or other external endpoints. CUIT and Zuckerman Institute co-hosted a Globus workshop on April 24th with over 25 attendees from Columbia and nearby institutions. At the workshop, Globus experts from the University of Chicago presented Globus features and technical topics in depth. Columbia has an enterprise Globus license which provides users with premium features. Nicky Agate asks whether Globus provides data publishing. George responded that Globus seems to be deemphasizing data publishing capabilities and may perhaps discontinue data publishing support going forward.

RCEC PLANNING

Chris Marianetti, Chair of SRCPAC

Chris asks researchers to send updates of publications that used Columbia's HPC resources to srcpac@columbia.edu.

Chris, Raj, and others provide an overview of the Research Computing Executive Committee (RCEC). The RCEC meets once a year, and attendees of this meeting are deans, the EVPR, EVP A&S, Chief Librarian, and the heads of CUIT and the DSI. At the next RCEC meeting in June, Chris will present about HPC, Marc will talk about Foundations, and CUIT will also provide an update. Presenters will report on what has happened in the past year and there will be budget asks to support future initiatives.

CLOSING REMARKS

Chris opens up the floor for feedback and discussion of any additional topics.

To date, the HPC Free tier requires sponsorship by a faculty member. Moacir Pereira notes that he had a grad student that had to wait for 2 months until faculty approved someone for the free tier. RCS acknowledges that this is a long time and there may be ways to alleviate this sponsorship hurdle to becoming a free tier user. One option that will help in this case is allowing additional people from the Libraries approve HPC Free tier memberships.

Rob Lane asks about how to fund and support an HPC Free tier for Terremoto. RCEC is coming up and may be worth bringing up at that to discuss how to secure the funding. Barbara Rockenbach adds that perhaps the HPC Free tier could be supported as an outgrowth of the Foundations initiative.

Chris Marianetti thanks everyone for their input and closes the meeting.