Yeti Operating & Executive Committees – Fall 2015 Meeting
November 5, 2015

Attendees: Greg Bryan (Chair), Chris Marianetti, Rajendra Bose, Rob Lane, Halayn Hescock, George Garrett, Marley Bauce, Dali Plasvic, John Villa, Marc Spiegelman, Michael Burke, Brent Stockwell, Roslyn Hui, Ingrid Richter, Tian Zhang, Wojciech Kopczuk.

Greg Bryan welcomes attendees and sets platform for discussion being an open forum discussion, with the Executive session being the last portion of the meeting.

Group makes personal introductions.

Rob Lane introduces meeting as primarily a review and reminders of policies, procedures, and usage activities.

Over 2015, there was a ramp-up when new systems went online in the new year. We have hit a steady plateau at 60,000 core days, although our theoretical maximum is 80,000 (so running at 75% of maximum); we will never hit this due to scheduler inefficiencies.

310 users in 2015 have run at least one job on the system, with 24 groups (does not include renters, TIP users, Free Tier, etc.). 4.5M jobs, with an average job size of 1.6 cores, and an average wall time of 0.1 days. Average job size is 1.1 nodes, wait time 2.3 hours, and average execution time 3.2 hours. These numbers are highly dependent upon high-throughput users.

Rob then shows all the queues split-up by system. Batch0 was added after the last Yeti Operating Committee meeting; batch2a was a temporary queue introduced prior to the introduction of Infiniband. Interlong and long are special queues that users can request if they must go beyond special parameters; these special requests are approved by Greg Bryan.

Since the Spring 2015 meeting:

- Full-node infiniband nodes have been added
- Access to infiniband is limited to the infiniband queue and the batch0 (2 hour) queue
- Added GPU queue
- Removed stack limits
- Job submission from execute nodes, which required a recompilation of the torque scheduler
- Hasn't happened yet, but migration to CUIT's Converged Infrastructure has opened up more space, so users will have more space for their home directory quotas (from 80MBs to 5GBs). This will allow for greater storage, and prevent sudden failures of MatLab.

Rob plans to ask Executive Committee to vote about usage of Free Tier, which is open to any researcher, or anyone with a CUID and faculty endorsement. Free users may use $1/64^{th}$ of a node (comparatively, renters get $1/16^{th}$ of a node). Brent Stockwell asks whether anyone at Columbia is able to use the Free Tier; Rob replies that this is open to anyone with a UNI, although Barnard users have not yet asked. The CUIT policy is to be as "open as possible," then asks for feedback about whether Free Tier should be open to affiliated institutions. Question can be re-raised at upcoming SRCPAC meeting.

Education Subcommittee explored options for letting Free users have access to more cores but for less time, to afford a simulated experience with an HPC-type system. Current system is to allow Free users more time on fewer cores.

Rob proposes to raise Core limit to 32-soft;48-hard, and restrict jobs to batch0, batch1, and Infiniband. This would allow Free users to use two Infiniband nodes for two hours. Rob asks whether a two-hour limit is acceptable.

No objections in the room, which indicates consent. Rob confirms there will now be a two-hour limit for Free Infiniband jobs, which will be confirmed at upcoming Executive session.

For the Education Tier, which has already been approved and set-up, but there has not been any use so far. Rob announces that any courses that utilize computational/HPC methods may have students gain access to this Tier. Marc asks whether this Tier is utilized for the CUIT/Libraries training courses. They have been using Hotfoot for these trainings, though could begin to use Yeti Free/Education.

CUIT is making another purchase round, and will begin 2016 with an RFP out to multiple vendors, with the purchase during the early-Summer (though there are restrictions on some funding sources). This will be a new cluster with new nodes and a new scheduler… from the ground-up. Thus, the Trial With Intent to Purchase (TIP) program is restarted, where prospective users are given a good-faith trial period to try the system before making a significant investment to a shared service.

For Publications (this will be a permanent slide), the most important way to justify the existence of Yeti service is by quantifying the number of papers emerging out of Yeti-covered research. **If you have any publications to report, please send to mb3952@columbia.edu at any time throughout the year.** Executive leadership is keen to count these publication accomplishments.

Rob asks whether job queue limits – job time, memory – are adequate. Group is silent, which indicates general satisfaction (or the lack of dissatisfaction).

Chris notes that it is possible to run into an incident where a group runs many nodes, and perhaps we can set maximum user run to the maximum nodes of the group, or whether this would be too complicated. Rob replies that this may be possible to achieve. Chris recommends making Infiniband time limit 48 hours, but that if it was turning down overall efficiency, the limit should be adjusted. This would afford greater flexibility and room so that they do not have to restart, reprogram, or complain to CUIT.

Greg asks who in room is using Infiniband; Chris Marianetti and Marc Spiegelman indicated they are. Chris explains that raising time limit to 48 hours would be useful for him, and there is no reason to raise beyond 48 hours. Rob replies that we can raise the limit now, and revisit issue for the Spring Yeti meeting.

Greg then opens the floor for general group discussion (“complaints, comments, congratulations”) on Cluster performance and policy. Greg requested comments prior to meeting, and received two replies. While there is still work to be done, for the most part there are no “burning issues.”

Brent Stockwell asks about batch jobs to help move jobs off of Hotfoot. George replies that he will begin work today with Schrodinger. Brent would like to understand whether the delay is a resource allocation issue (whereby CUIT cannot keep up with demand), or whether this is a staff time issue. George replies that Schrodinger is the most sophisticated software CUIT uses; the process is exceedingly complicated, and RCS has less resources, though has made a new hire (John Villa, started two weeks ago), who will help improve response time for more complicated issues. CUIT is committed to maintaining a two-week response time; if this is impossible, they will keep communication open in order to meet client expectations for performance. Rob notes that most software will take at least two weeks to install, and that CUIT will be “tightening-up” procedures, and are enthusiastic about new hire to ensure that install queue is properly managed.

Brent additionally mentions that he negotiated a deal with Schrodinger to make their software available – unlimited – to the entire university, which may result in more users who want to pilot/TIP. Chris comments that Schrodinger is an advanced piece of software, and commends CUIT to making it work on the cluster.

Raj asks whether there are trends in complications or requests of users for asking for different things. Does growth lead to diversity of atypical requests? Rob notes that, previous discussion notwithstanding, there have not been any great issues, and RCS has been able to successfully manage all new user requests.

Marc Spiegelman asks whether some queues are saturated with jobs that should not be run. Rob believes that the new-generation Infiniband is heavily used, while first generation Infiniband is not heavily used. Rob notes that there is more demand for specialized hardware now than there was even two years ago.

Greg closes floor and asks Executive Committee to stay for voting items. Remaining attendees leave.

Executive Committee Voting Items

1. **Change to Free Tier: Increasing core count, but decreasing usage time.** Brent wants to monitor utilization of Free Tier over time to see whether change impinges upon anything else. Brent asks whether Free Tier users are satisfied. Rob notes that users are very quiet – nobody has taken advantage of Tier. Rob believes that people do not know that this exists. Brent thinks that if we were to enhance this Tier, we could receive more resources from Central Administration, potentially buying more nodes to meet unmet demand. Greg notes that this option was not heavily publicized out of fear for a massive sign-up that would overwhelm machine, although this has not happened. Greg believes we should publicize opportunity more widely. Just as Rental Tier is publicized on front page of Yeti log-in, same could be done for Free Tier. Chris is teaching a 22-23 student course that may be prime for utilizing the Education Tier, which may strengthen case to Central Administration for more funding of resource.
2. **Increasing Infiniband time for 36 to 48 hours.** Committee agrees.

Halayn notes that CUIT did increase funding for RCS and HPC staff support, which culminated in their new hire. CUIT also received one-time funding that may be part of new cluster capacity, and is still figuring out how to spend this new funding. RCS also has some soft-money staff. On behalf of the Yeti Executive Committee, Greg thanks CUIT for this new investment. Chris calls Yeti a “huge success” and echoes that we have a lot to be proud of, and have ample evidence that the University should continue to support this.