Hyak Governance Board Meeting
Notes for August 16, 2012

I. Updates & Enhancements

We've been gradually adding more supported end-user codes and applications with broad utility:

- Mathematica: the Foster School of Business uses this extensively and it is now up and running on the wiki.
- Matlab DCS: a combined effort with the Matlab people to help support applied mathematics. One issue is that it doesn't work with Windows. Chance Reschke is working to get the solution from OSU.
- ARE: has been updated to the ARE Revolution Enterprise and is supported by companies, so we don't have to maintain it. The iSchool are heavy users of this.
- Tuned applications: working with IBM to target tuned applications that have caused problems in the past and/or have benefits in the future.
    - Gromacs: IBM has it up and running; Jim's people have been testing it. The tuned version has improvement of 7%.
    - GPU support
    - WRF and CESM

II. Hyak Capacity Projections

No real change from last meeting, except for two potential large deployments:

- Institute for Protein Design (David Baker's lab) has a budget for computational requirement and could grow to 300 nodes in Hyak if funded.
- pChem MRI: Jim's group is leading.

III. lolo Capacity Projections

- The Baker lab will move their archive (1/4 TB) from the old data center on Boat Street and evaluating lolo as a storage option.
- Genome Sciences (John Stamatoyannopulous' group) asked for 3/4 TB.

The good news is that the rates drop as capacity rises. The current rate is $190 per TB a year. It is possible to lower the rates mid-year. With the above two additions, we can potentially lower the rates to $80 per TB a year. Also, there will be a 20-30 TB addition to lolo collaboration that is available to campus to use as a buffer.

IV. Proposal successes, failures, and pending

- NSF CC-NIE: the proposal to support upgrades for network infrastructure was funded with about half a million dollars with Ed Lazowska as the PI. This was a recycling of an old proposal (NSF ARI) which was restructured to provide a cohesive narrative. We were asked to provide a science DMZ, which we already have (the diagram of hyak/lolo with blue oval vast connection coming out of it). This will upgrade the research network backbone to 20GB/sec for research traffic for on campus and off campus has an upgraded connection to gigapop. Martin Savage requested a map of the UW data network showing where the congestion points are. Brad Greer offered to track this down.
- NSF MRI: pending
- NSF CRI: this proposal was rejected, but the comments from the reviewers highlighted the apparent lack of institutional commitment to Hyak/lolo as a reason. Mary suggested we create standard letters from UW leadership which can be attached to all proposals depending on Hyak/lolo. Chance Reschke indicated he could draft these.

- NIH HEI: new round of funding has not been announced, but the genomic science groups have an interest.

Mary shared that future funding will most likely be very tight due to the upcoming sequestration. Christy Gullion our federal relations officer believes that this will happen.

V. IBM relationship

We met with IBM to discuss issues with general file system, College of the Environment, and GPU. In October, an IBM representative will come and discuss issues with the College of the Environment. With GPU, IBM will comp $100,000 that we invested with new gear when the engineering flaw is fixed.

VI. The Regional Scene

The universities across Oregon (Oregon State University, University of Oregon, and Oregon Health Science University) have established an Oregon universities research collaboration which is similar to the Hyak/lolo collaboration.

VII. Next Steps

Mary asked the group's preference for the meeting schedule and agenda. The group decided to continue meeting for 1 hour monthly.

Martin Savage asked about recommitting to the elements in support of scalable scientific computing (SSC) that were discussed in the initial planning meetings for the eScience Institute. These include:

- Outreach to identify researchers hoping to scale their scientific computing efforts
- Expert, hands-on assistance to help researchers scale their scientific computing efforts
- Expert, hands-on assistance to help researchers be most effective using Hyak and lolo

Mary offered to investigate the history of SSC plans and priorities .

Bruce Nelson and Mary both raised the issue of identifying sources of funds to support these efforts. Chance Reschke offered to provide the Board with information about the approach to this problem taken by Cornell. Jeff Gardner agreed that Cornell was a good model to follow.

Tom Ackerman indicated that lolo Archive rates were still too high for some users

A group discussion followed in which it was acknowledged that lolo is not an ideal solution for all groups. Use cases for which the lolo Archive seem to be a clearly good fit are groups:

- with more than a few TB of data
- who require a guarantee that their data will be preserved
- who need their data archived long term
- who require high speed data transfer with lolo
- who require high speed data transfer with off-site systems such as super computer centers
- who require high speed data transfer with a variety of on-campus systems