

# Research Computing Advisory Committee

Minutes Dec 2, 2013 (taken by Laurie Taylor)

**Present:** Sophia Acord, Paul Avery, Reed Beaman, Erik Deumens, Rob Ferl, Rolando Milian, David Nessler, Sanjay Ranka, Alberto Riva, Almuth Winterstein

## Discussion

### 1. Collaborative Research Data Storage project (Erik):

- a. Implementation started with DDN having been selected by statewide ITN on Monday Nov 18, 2013. Goal is to enable collaboration with large research data sets across SSERCA and with the rest of the world. The architecture supports replication and high performance connectivity, CIFS access to workstations, NFS access to HPC systems, and excitingly has a Dropbox-like interface for ease of sharing data with anyone (does not require a Gatorlink or other long process, so it is more usable). The goal is to have the configuration for UF defined in December, system ordered early January, with target date of end of March or April for production. Research Computing will provide training sessions for using the system. The system is WOS, web object storage, from the vendor, DDN, DataDirect Networks, a high-performance storage company.
- b. This will be the long term storage, with the existing pricing with Research Computing; researchers will be able to buy and use for their research data. Given the size and costs, planning is ongoing to also support general data sharing needs at no cost for small data sizes, like 20 GB per researcher.
- c. A system to support PHI (Patient Health Information) data is also being deployed based on the same hardware platform.
- d. The committee expressed strong support for this long-awaited solution for research data. It is very supportive of trying to provide a system that is open for everyone, not just researchers who have big needs that require buying storage from Research Computing. It also indicated that addressing the PHI problem should be a high priority and should be done in conjunction with this solution, lest there would be yet another partial solution with limited scope.
- e. A possible name for the system is GatorBox.

### 2. Summer school planning for 2014: developers wanted (Erik)

- a. Summer school in microbial computational genomics. Prof. Marco Salemi of the EPI organized a successful summer school at UF in August 2013. He proposes to develop a summer school at UF on a similar topic, but wants to make it broader so that it can attract more students and aligns with the broader scope of the Informatics Institute being created. The UF Conference Department took care of the logistics. The 93 students from 47 countries were very pleased. They paid their way. The conference had a budget of \$100K and the expenses were about \$50K, thus the process generated \$50K.

- b. Thus there is an opportunity for other faculty to teach modules. The structure of the summer school is to teach a module in a week, with lectures in the morning and hands-on labs in the afternoon. We could invite students to stay for another week or two and work on their data with resources at UF and with access to the experts who just taught them the tools and the theory. It is possible that faculty who teach a module get summer salary.
- c. The long term vision is that the modules can be expanded into a full semester course, physical or online, if appropriate.
- d. Sophia Acord and Laurie Taylor are working on a module based on the primer for big data, now called HPDA (high-performance data analytics), digital epigraphy, etc.
- e. Rob recommended discussing this with Eric Triplett based on innovative teaching experience that relates to this proposal. Instructors only need to do content, and the rest supported at the institutional level.
- f. Could do similar CH&PS funding with scholarships like with the digital storytelling scholarships.
- g. In the future, CPET (Center for Precollegiate Education and Training) can be engaged for doing outreach to high school teachers, for developing training and experience for bright students at the high school level.
- h. All RCAC members are asked to share this opportunity with anyone and everyone who may be interested.

**1. Supercomputing 2013 (Erik, Paul, others)**

- a. Booth for all institutions of SSERCA (Sunshine State Education and Research Computing Alliance), offers place to meet to connect people as a human infrastructure, network effect.
- b. Opportunity to show slide decks to promote research at UF.

**2. Continue development of outreach plan (all)**

- a. Paul will write a draft plan from the discussions the committee had in the past three months for review, comment, and approval.
- b. The plan with action items and metrics of impact and success is due soon to the CIO.

**Organization**

The web site has been updated to show the meeting schedule through the Summer 2014 at <http://www.it.ufl.edu/governance/advisorycommittees/researchcomputing.html>.

Next meeting will be on Monday January 6, 2014 at the usual location in NPB 2205 from 1:30 – 2: 30 pm.