

Research Cyberinfrastructure Immediate and Long Term Staffing Needs and Vision

Allen Schiano 12-3-2015

Summary: Compared to similar research universities, including in the UC system, UCI's staffing service for 'RCI' are substantially below the optimum level. Comparisons of staffing levels at representative universities at the forefront of campus RCI often show more than 10 staff assigned in various roles to RCI. A survey of UCI staffing indicates that our staffing levels are a factor of 2-3 lower than universities with comparable research funding, faculty head counts, or stature in university ratings. Short term needs are focused in both areas where current services do not have enough staffing to sustain the current level of service or growth and in areas of critically to the campus, like research data storage, that have no staffing assigned. Longer term needs general focus on staff services to RCI areas that are growing rapidly, e.g., genomics and bio/medical research, where the need is on training and supporting new users of RCI.

Current State of Staff Services to RCI: At present UCI has approximately 3 FTEs assigned to RCI infrastructure support (Green Planet and HPC clusters and LightPath network) and approximately 3 FTEs assigned to supporting researchers in topic areas, tools, and general computing support. As an exemplar. Purdue or Indiana Universities' have 20-30 or more staff assigned to these tasks. The largest UC RCI environment has over a dozen staff assigned to RCI support.

UCI's RCI staff support areas include:

- Computing cluster architecture, design, operation, and development
- Research (high performance and size) data storage and transfer
- Parallel computing support both 'fine grained, i.e., OpenMP and MPI' and 'course grained'
- Bioinformatics, especially genomics and similar data processing and reduction
- Computational chemistry
- Graphical information systems and associated scientific visualization

Immediate Staffing Needs: The most critical short term needs are to bolster the support of computational clusters at UCI and expand the areas of research support in topical areas. In the first case, these needs are similar enough in character that a joint OIT and Physical Science focus is amenable. Research support in topical areas could be addressed either centrally if the need was campus wide or in Schools

where the need was focused in a School only. The most critical needs in order of importance are:

1. *Increase staffing to maintain and develop the existing computational clusters on campus.* That need is driven by the large increase in user support (software installation, training of new users to HPC, Linux, and related topics, code optimization and development) in 'new areas' of cluster computing such as Biological and Medical Sciences. **The highest current need is for cluster knowledgeable system administrators and engineers** at a level high enough to allow current high level staff to concentrate on the development of the systems including the very complex networking, queuing and job control, checkpointing, security and other tasks that are currently being not done or significantly delayed.
2. *Provide system administration/engineering staff capable of developing, maintaining, and training end users on the use of a new research storage system and service.* The main new service requested by the largest number of RCI using faculty (see 2013 FASRC results) cannot currently be built and operated due to a lack of high level technical talent on campus. Currently that level of technical talent is used to support the computational clusters. While costs to purchase and develop a research data storage system are likely to be available via a 'condo style' funding model already successful for Green Planet and HPC, **the high entry cost of a new, long term funding commitment for staff requires direct campus support.**
3. *Increase the number of research computing topical staffing whether centrally or in Schools.* While there are 'pockets' of staff services in a few research areas, the wide range of campus research areas and associated hardware and software support services increasing the staffing in the most 'pressing' or 'growth' areas would follow the path of other leading RCI campuses. **In the short term, another one or two staff topical experts is greatly needed. Which support areas to develop first would be an item for discussion among a campus governance group coming out of this visioning process.** Whichever topical area is chosen would require staffing that would work one-on-one with research faculty on specific projects and support as well as working through training and instructional efforts to the graduate and undergraduate communities. A list of possible topical areas is noted below.

Long Term Staffing Needs: As seen from comparisons with other leading universities, the growth of RCI services includes a commensurate increase in staffing in support of research users. Research support is NOT an 'operation which scales less than linearly with increasing number of users, e.g., an email service. RCI support is a 'hands on' operation where staff are often more colleagues than supporters to the research process. As such, adding staff is a daunting long term task for making improvements to UCI's RCI environment.

While it is possible to add more hardware and services to a computational cluster, for example, and therefore we could double the size of Green Planet or HPC without a doubling of system administration staff, the doubling (or more) increase in the number of users, software 'titles', and overall support costs would be likely at least doubling. Adding in new types of users and their research areas would also add to the need. Thus a long term vision for RCI staffing services implies a large increase in staffing, at least comparable to the addition of research users. This has not happened at UCI in many years and as such we are behind in the short term staffing of RCI.

In which topical areas these staff should provide expertise is a matter of determination by the overall research mission of the campus that requires RCI services. OIT's Research Computing Support and Physical Sciences Support staff have noted the following areas where they are receiving the most requests (the 2103 FASRC report also corroborates these assessments). In no priority order, they are:

- High performance programming skills and techniques (OpenMP, MPI, GPGPUs, etc.)
- High performance networking and data transfer
- High performance i/o and data storage, access, and sharing
- Bioinformatics, especially related to genomics sequencing
- Statistical support, especially training and usage of modern tools
- Mathematical software
- Scientific visualization and analysis
- Data management and curation
- Linux expertise, especially training of new users
- Scientific use of databases
- Web site development in support of research data distribution and publishing
- And many others that could be listed

A final addition to the list of staff services for RCI has often been called an 'RCI concierge'. This need comes from the often heard statement (at UCI and elsewhere) that faculty don't know about campus resources that are available to them in many cases. Information on such services travels best via word of mouth such that the typical IT information spreading tools like mailing lists and websites fails with researchers. As such, the need for one or more staff to spend their time evangelizing RCI services is needed. 'RCI concierges' would make themselves visible to research faculty and students by attending meetings and events where researchers circulate. However without a 'RCI specialty' of their own to support their value might be lessened. **As such, the workgroup believes that the model should instead focus on all RCI support staff to have a significant portion devoted to such 'outreach' and concierge service. What currently precludes this from happening is the understaffing of RCI at present.** With a significant increase of RCI staffing as

described above, RCI outreach to researchers that COULD use RCI services but don't know about them, would thereby strengthening the overall campus research output.