High-Performance and High-Throughput Computing at CSU

Pat Burns, VP for IT

September 30, 2015

Background

For decades, High-Performance Computing (HPC) has been essential to the conduct of research and higher education, complementing analysis and physical experimentation as another way of conducting science and engineering. Recently, however, the field has broadened significantly as the amount of data available for analysis has exploded into the realm of "big data," to evolve to a model of High-Throughput Computing (HTC). HPC tends to be compute intensive and HTC tends to be data intensive. Both aspects need to be supported to facilitate and enable researchers to be productive and competitive in our research environment.

Strategy

CSU's strategy for providing adequate HPC/HTC infrastructure (hardware and software) is threefold:

1) Continue to use its Cray obtained in 2010 under a previous NSF MRI grant, including funding needed upgrades,
2) Periodically apply to the National Science Foundation (NSF) for funding of new HPC/HTC resources through its restricted Major Research Instrumentation (MRI) program, and
3) Institute a "condominium" model for HPC resources, where users contribute each year to building a large, HPC/HTC resource that is operated by central IT and users share when they are not using the resources they purchased.

In the spring of 2015, with contributions from numerous departments, CSU upgraded its Cray to an XE6 model, including expanding the compute nodes by about 25% (item 1). However, the storage system was not upgraded, and it is likely that this upgrade will be needed within the next year or two. In addition, CSU and CU-Boulder collaborated to obtain a $3.9 million MRI grant from the NSF for a large, shared HPC system (item 2) that will be housed and operated at CU-Boulder to which we are fiber connected (making it perform as if it were on CSU's Local Area Network).

CSU Environment

In its research environment, CSU has numerous wet labs staffed by the world's best scientists, and also has some of the world's best analysts. Complementing this type of research with HPC and HTC offers the potential to increase innovation, and increase the amount of funding at CSU for research. However, an effective HPC/HTC environment requires: 1) access to contemporary hardware and software (addressed in the previous section), 2) support staff of high quality in needed numbers, 3) a rich educational environment to deliver training and education, and 4) a culture of collaboration and cooperation. Currently, CSU has adequate resources in each of these areas, except for staffing. Additional staffing is needed due to the expansion of our HPC/HTC activities from one activity (our Cray) to the three mentioned above in the previous section.

Staffing

CSU has staffed its HPC activities, with 1) a 0.5 FTE Manager of HPC, Richard Casey, who oversees CSU's entire central HPC/HTC environment, provides user support, installs and supports software packages, and responds to trouble tickets, and 2) a 0.25 system administrator, Dan Hamp, who manages and operates the Cray hardware, storage, and operating system. These staff have been barely adequate to support just the Cray. Now that we are embarking upon two additional major extensions to our HPC/HTC environment, the new environment will be vastly understaffed. Adding more staff is the principal recommendation from the Research Computing Committee (they recommended increasing to about four FTE staff), constituted in 2011 under the authority of the Vice Presidents for Research and IT to make strategic recommendations for our HPC/HTC environment. However, to get started and consistent with limited resources being available, it is proposed to elevate Richard Casey to full time, add two half-time hourly graduate students for user and application support, and elevate Dan Hamp from 0.25 to 0.5 FTE. A budget for this may be found as Appendix A.

Metric for Success

The proposed amount of staffing should be adequate initially to embark upon the new, more complex model consisting of three components, and should be regarded as an initial investment in elevating our HPC/HTC activities to higher levels. The Return on Investment (ROI) would be the increase in research volume enabled by this increased investment. The Office of the Vice President for Research has begun tracking grant activity in the HPC/HTC area, allowing the trend in ROI to be determined from that data. Should the volume of funded research grow in the areas of HPC/HTC, then adding additional staff in the future should be considered as an investment strategy to accelerate even more growth.

Appendix A

Cost Proposal

1.	Salary, Rick Casey, elevate to 6 months, equity increase	$46,166
2.	Benefits (25.4% of 1)	$11,726
3.	Student Hourly employees (2,000 hours at $15/hr.)	$30,000
4.	Benefits (0.6% of 3)	$180
5.	Total	$88,073
6.	Total fringe	$11,906
7.	Net base request (assumes fringe would be paid centrally)	$76,166

The budget consists simply of the additional salary for Rick Casey (item 1), his fringe (item 2), salary for two hourly half-time graduate students (item 3), their fringe (item 4), totaling $88,073 (item 5). The base budget request of $76,166 (item 7) assumes the total fringe (item 6) would be paid centrally and is therefore not requested).

Note: ACNS will absorb the cost of increasing Dan Hamp from 0.25 to 0.5 FTE support for HPC/HTC, so this budget element is not included in this request.