

Research Computing Advisory Committee

Minutes July 6, 2020 (taken by Erik Deumens)

Present: Paul Avery, S. Balachandar, Nikolay Bliznyuk, Erik Deumens, Jorge Frias-Lopez, Damon Lamb, Lauren McIntyre, Rafa Munoz-Carpena, Melissa Rethlefsen, Alberto Riva, Plato Smith, Laurie Taylor, Bruce Vogel, Chris Vulpe

Updates

- NVIDIA announced new GPU technology on May 14: The Ampere A100 GPU and the DGX A100 node with 8 A100 GPUs tightly coupled. UF is the first university in the world to get this technology. They arrived June 25th.
- More details of the UF-NVIDIA partnership hinted at during the Provost's AI Retreat on April 29 will be announced on July 21st.
- The new /blue file system is in production.
 - o All data has been copied.
 - o We now coordinate with each group (120 done, 269 to go) for a final sync and then the group is moved over.
 - o All data on the old /ufrc will remain there; we just make it not visible to users to avoid the same file being changed in two places.
 - o It is likely that /ufrc will keep that data until early 2021 when it will be turned off. Hence if any data inconsistency is discovered in the new /blue, the old data will be available until at least that time.
- The HiPerGator 3.0 ITN is in progress.
 - o We cannot provide specific details at this time.
 - o We hope to place a P.O. before the end of July.
- UFIT RC has completed one search and hired Brennan Jones, who will start at UF on August 7. The second search will be reposted. A staffing proposal to support the AI initiative is being reviewed for funding.
- The metric page is now in production at <https://metrics.rc.ufl.edu/>
 - o To get to it you need to authenticate with GatorLink (and DUO) and then you need to authorize with your HiPerGator credentials, which are the same. But not everyone with a GL can view these pages. This simplifies the support RC staff needs to perform on this page.

Discussion

Members ask for an update on the SAS license. The details were announced in January. Faculty have to renew their licenses at the new prices before August 30, so that we can see whether the cost model is sustainable.

- The importance and urgency of this is not clear to faculty.
- Erik will bring this to the attention of UFIT licensing and SAS licensees will be contacted as well as a reminder communication effort will be carried out.

With the expansion of equipment (new HPG 3.0, new GPUs, new AI training), the fact that UFIT Research Computing does not have enough FTE to support the systems

remains. New staff will be hired for the new AI initiative, but there was a backlog before the AI initiative. How to best make the case for new staff to UF administration?

- There is a tendency by administrators to view support for RC as a cost. It should be viewed as an investment in the university infrastructure for the faculty and students to be more successful. To give them the tools they need to outcompete all universities in the nation and make UF “top 5” and stay there.
- Part of the FTE can be charged to grants for consulting at the rate of \$65/hour. However, this does not recover the full cost because only about 1,000 hours out of the 2,000 hours per year of an FTE are allowable cost to grants. So to create a team of consultants that can deliver in a timely manner when grant-funded projects need their services requires an investment in 2 FTE to sustain a team of 4 FTE.
- The numbers to support this view do not leave any doubt about the financial soundness of that investment:
 - o A \$3M per year budget (mostly from UFIT, with some from the Office of the Provost and some from the Office of Research) supports the staff and operation of HiPerGator.
 - o Over 70% of the cost of HiPerGator hardware (CPUs, GPUs, nodes, and network equipment) and 100% of the cost of storage hardware is recovered from research grants.
 - o This investment supports over 300 active faculty who bring in \$280M in research grants and contracts every year. That means \$3 million annual investment supports \$280 million in grant expenditures per year. While the university invests in many other things, including building and IT infrastructure, to support this grant project activity, it remains that the investment in support for research computing staff and infrastructure is a small investment enabling a significant amount of activity in a highly competitive arena.

Next meeting will be on August 3, 2020 from 1:30 – 2:30 pm. The meeting will be Zoom only. An invitation will be sent.