

CRCF Annual Public Report

Academic Year 2023

Last updated 5/23/23

Contents

- Overview 2
- CRCF HPC Cluster Specifications 4
 - Augie 4
 - Alipi & Clusty 5
 - Key Personnel (May 2023) 5
- Strategic Plan 6
- Accomplishments 7
 - Administrative 7
 - Technical 7
- Metrics 9
- Work in Progress 13
- User Survey Responses 13
- Final Points 15
 - Good Citizenship 15
 - Interested in Helping CRCF? 15

Overview

The Centralized Research Computing Facility (CRCF) is Villanova's first core facility (established November 2021) and provides support for computational research. Most research support is in the form of high performance computing (HPC), but other forms of support are available as well. We encourage all Villanova researchers to contact us if they are interested in research computing but do not yet have a strategy towards achieving their goals.

CRCF supports three HPC clusters and their users:

- Augie: this is CRCF's largest and newest cluster. It was procured through an NSF Campus Cyberinfrastructure (CC*) grant in Summer 2020. The cluster was installed in March 2021. Augie is open to all Villanova researchers, external researchers as part of the Southeastern PA Consortium, and Villanova students taking Augie as part of a formal course collaboration between the instructor and CRCF.
- Alipi: this cluster is limited to College of Engineering researchers. Alipi is a couple of years older than Augie.
- Clusty: this is the Astrophysics cluster that has been opened up to all users. The cluster is managed by Andrej Prsa in the Astrophysics and Planetary Science department.

CRCF emphasizes that *Centralized computing is more than just HPC equipment*. The facility envisions the following at its maturity:

- State-of-the-art HPC resources for use by the entire Villanova community
- Central knowledge base/pool of expertise for research computing resources
- A partner to help faculty integrate HPC into their coursework and proposals
- A financial sustainability model that follows federal guidelines, is equitable, and controls resource growth to match demand.
- A mechanism to share the wealth of existing research computing knowledge
- A resource for developing computer programming and software applications skills for students, staff, and faculty. This resource follows Augustinian values of helping others (akin to service learning)

In summary, *CRCF is a partnership of faculty, technical staff and students that empowers Villanova to maximize its computational research potential.*

CRCF has three thrusts as part of this mission/vision:

Thrust #1: HPC Support through a Collaboration of Faculty and IT Staff. This thrust involves the management of Augie, Alipi, and Clusty.

Thrust #2: Research Computing Center of Expertise. CRCF provides training opportunities including monthly new user training, skills development courses, and a point of contact for interaction with experts at other institutions. All researchers interested in research computing or HPC should approach CRCF to help develop a strategy to meet their goals.

Thrust #3: HPC Peer User Support (a.k.a. The Coding Center). A wealth of research computing expertise exists on campus, and CRCF strives to facilitate communication between users seeking help and the experts that can help them. The Coding Center consists of volunteer experts helping those that need it

per Villanova's Augustinian tradition, providing support development and debugging for existing coding projects and support for the core software stack. Finally, CRCF can search for experts at other institutions for support if on-campus Coding Center personnel are unable to help address issues at hand.

CRCF HPC Cluster Specifications

Augie

Item	Values
Compute Nodes/Cores	29 compute nodes/1,856 AMD EPYC Series
RAM	13.3 TB Main Memory
Disk Space	292 TB
GPU	2x Tesla A100: 13.8k CUDA cores for throughput of 19.5 TFLOPS
Interconnect	Mellanox 200 Gbps
Software	LAMMPS Quantum Espresso COMSOL Python CP2K MATLAB VASP R PyMT TensorFlow Compilers: GCC, AOCC

The status of Augie's applications is as follows:

1. The following software applications has been installed and performance tested. Sample scripts are available to users: C/OpenMPI, Fortran/OpenMPI, MATLAB, Python, R, COMSOL
2. The following software applications has been installed, and performance testing and sample script development is in progress: QuantumEspresso, LAMMPS, CP2K, VASP, IQTREE, PyMT, CUDA, TensorFlow, PyTorch, Intel compilers, AMD compilers, PAML, RAXML
3. The following software applications are slated for installation: OpenFOAM, Singularity
4. The following software applications need further exploration or discussion prior to initiating an installation effort: Mathematica, Abaqus, ANSYS, Anjuta, Gaussian

Only those software packages listed in Items 1-3 above are currently supported by CRCF. Users are encouraged to install their own software packages in their home directories for their own use.

The Software Wiki was established this past year. The Software Wiki features pages devoted to the apps in Items 1 and 2 above, where each page is managed by a superuser and includes tips and notes regarding running the software. In addition, the Augie User guide features links to online tutorials and videos for running basic cases of each software tool.

Alipi & Clusty

Cluster	Alipi	Clusty
Cores	208	212
RAM	1,408 GB	208 GB
Disk Space	26 TB	15 TB
GPU	N/A	N/A
Interconnect	InfiniBand	10 Gbps
Software	Python, LAMMPS, QuantumEspresso , VASP, MATLAB Compilers: Akantu , Intel	Ubuntu server, Python, Perl, R, sqlite3, octave

Key Personnel (May 2023)

CC* Committee: Joseph DeMarco (student, CLAS), Jonathan Hardy (UNIT), Michael Robson (CLAS, Technical Director), Aaron Wemhoff (COE, Administrative Director)

Key UNIT & CLAS IT Support (Augie): Peter Palladino, Gavin Printz, Christopher Washburn, Jonathan Graziola

HPC Advisory Board: Andrej Prsa (CLAS), Daniel Smith (CON), Bill Wagner (VSB), David Cereceda (COE), Justin DiBenedetto (CLAS)

External Advisor: Jason Simms (Swarthmore College)

COE HPC Committee (Alipi): David Cereceda (ME), Zuyi Huang (CBE), Kyle Juretus (ECE), Chengyu Li (ME), Ondrei Miller (UNIT), Virginia Smith (CEE), Aaron Wemhoff (ME, chair)

Strategic Plan

Mission: Support and promote computational research for Villanova researchers and their collaborators

Vision: CRCF, as a center of expertise in computational research, is one of the top contributors to Villanova's research enterprise.

Goals & Metrics (5-year goal after colon):

1. CRCF is a significant enabler of research productivity
 - a. Number of grant proposals submitted that use CRCF resources: 10/yr.
 - b. Number of peer-reviewed publications produced using CRCF resources: 10/yr.
2. CRCF is the center of expertise for computational research
 - a. Number of workshops or tutorials offered: 4/yr.
 - b. Number of researchers using external services (OSG, PATH, XSEDE/ACCESS): 5/yr.
 - c. Number of CI grant proposals stemming from the center: 1/yr.
3. CRCF has a robust multidisciplinary community of computational researchers
 - a. Number of total users: 170
 - b. Number of active users: 40/mo.
 - c. Number of departments represented by active CRCF researchers: 12
 - d. Number of superusers per software application: 3
 - e. Number of software applications with superusers: 12
 - f. Number of classes that use Augie for instruction: 8/yr.
4. CRCF's resources grow to match demand
 - a. Low idle time: < 15%
 - b. Small average job wait time: < 12 hours (NOTE: we are still looking on ways to track this)
 - c. Number of new condoers per year: 2
 - d. Acting as a host site for REDCap
 - e. Dedicated HPC IT Admin
5. CRCF operates for the good of all humankind beyond research
 - a. Request that hosting data center must meet energy efficiency requirements (PUE < 1.5)
 - b. Request that data center have PPA in place for renewable energy

Accomplishments

CRCF's accomplishments for the AY have been split into administrative and technical categories.

Administrative

The team spent the AY continuing to grow CRCF to ensure long-term use and management of the facility. Highlights of the facility's administrative accomplishments include:

1. CRCF took over the financial management of COMSOL and ANSYS-Fluent research licenses, ensuring a fair, sustainable approach for use of these key software tools for researchers.
2. The Augie Management Plan and Augie User Guide documents were created. These documents provide much improved organization of the various policies surrounding using Augie for research and education. The Augie User Guide is especially useful for users as it is a central hub for all information they are seeking to maximize their benefit of using Augie, from technical specifications, a blurb for addition into proposals, tutorials on the benefits of HPC as well as specific software tools, and Acknowledgement statements in research publications.
3. A CRCF strategic plan was developed and implemented. The specifics of the plan are listed in the previous section.
4. The onboarding process for new Villanova users was overhauled and streamlined. The new process is much more efficient: users can be onboarded in a couple of days with personalized support to help them get off the ground with their research. This is accomplished in part by shifting away from once-a-month new user training sessions towards a Blackboard-based learning module. These users are added to the appropriate Teams and SharePoint pages, and mailing lists, within a few hours. In addition, the new user request form was transitioned to the current UNIT ticketing system, which has resulted in fewer technical issues.
5. CRCF office hours are now offered weekly for onboarding or to address user technical issues.
6. The Software Wiki was initiated for app/language-specific notes, superuser identification and Q&A. Pages were made for several software packages. The Wiki was transitioned to user-editable Word files on the AugieUsers SharePoint site since MS Teams announced it was phasing out the Wiki tool in June 2023.
7. CRCF met with CEE faculty candidates & provided info for CPE faculty candidates.
8. An HPC IT Administrator job search was conducted and completed.
9. Two nodes were procured for addition to Augie's general compute pool. These nodes are intended to eventually be converted into interactive nodes, one of which with the Open OnDemand app. These interactive nodes will make code debugging and execution more efficient and make integration into coursework for student hands-on exercises more feasible.
10. Michael Robson became Villanova's first XSEDE/ACCESS Champion.
11. The CRCF glossary of terms was expanded.
12. CRCF started emailing monthly updates and reminders to users this past AY.

Technical

Augie continues to evolve technically using the combined efforts of UNIT staff and faculty. This AY saw UNIT staff increase their role in evolving Augie's capabilities, whereas the faculty played more of a role in helping users achieve their research and teaching goals through working within Augie's existing capabilities. Some highlights:

1. Augie's SLURM limits (number of cores, memory allocation, and running time) started to be enforced. These limits, in particular the running time and number of cores, were also adjusted based on user request. Finally, the SLURM rules were adjusted to make access more equitable for all users.
2. Fortran and python scripts and performance tests were performed.
3. Augie's partitions and access rules were adjusted to enable the necessary 20% of Open Science Grid allotment while reducing idle time and Villanova user wait time.
4. CRCF helped users install the apps BLAST and sage in their Augie home directories.
5. Several users had their passwords reset at their request.
6. Several COMSOL toolboxes were added, and the overall COMSOL license was updated to the latest version.

Metrics

Total number of users (as of 3/6/23)

- Augie: 132
- Alipi: 52
- Clusty: 106 (note: all students in astrophysics & planetary science are on by default)

Journal papers using CRCF resources

- Crinnion, A.M., Toscano, J.C., & Toscano, C.M. (2022). Effects of experience on recognition of speech produced with a face mask. *Cognitive Research: Principles and Implications*, 7, 46. <https://doi.org/10.1186/s41235-022-00388-4>
- Bauer, A. M.; Sadlier, R. A.; Jackman, T. R. (2022) A Revision of the Genus *Bavayia* Roux, 1913 (Squamata: Gekkota: Diplodactylidae), a Non-adaptive Radiation of Microendemic Species. *Proceedings of the California Academy of Sciences*, 67, Series 4, Supplement I. 1-236.
- Qian, Y.; Gilbert, M. R.; Dezerald, L.; Nguyen-Manh, D.; Cereceda, D. (2023) Ab initio study of tungsten-based alloys under fusion power-plant conditions, *J. Nuclear Mat.*, 581, 154422.

Conference papers and presentations using CRCF resources

- Wemhoff, A. P., Trinh, A., 2022 ASME SHTC, Philadelphia, PA, "Extension of Cylindrical Inclusion Percolation Theory Towards Non-Uniform Distributions" (July 2022).
- Conference publication/presentation by Sergey Nersesov
- Gonzalez, A.S., Sarrett, M.E., & Toscano, J.C. (2022, October). Time-course of processing for syntactic properties in spoken word recognition. Poster presented at the 14th Annual Meeting of the Society for the Neurobiology of Language, Philadelphia, PA.
- Montañez, O., Sarrett, M.E., & Toscano, J.C. (2022, December). Time-course of processing for syntactic properties in spoken word recognition. Poster presented at the 183rd Meeting of the Acoustical Society of America, Nashville, TN.
- Sarrett, M.E., McMurray, B., & Toscano, J.C. (2021, November). Contributions of bottom-up and top-down processes in speech cue encoding: Evidence from EEG and machine learning techniques. Poster presented at the 62nd Annual Meeting of the Psychonomic Society. Virtual.
- Sarrett, M.E., McMurray, B., & Toscano, J.C. (2021, December). Neural representations of speech: Decoding bottom-up acoustics and examining top-down effects using electroencephalography. Poster presented at the 181st Meeting of the Acoustical Society of America. Seattle, WA.

Proposals submitted that would utilize CRCF resources (Items in bold are awarded grants)

- Caplan S (Principal), Toscano JC (Sponsor), "Identifying the computational mechanisms of perceptual learning", National Institutes of Health, F32 NRSA Postdoctoral Fellowship (sub: 12/8/2020, start: 8/1/2021, end: 7/31/2024).
- Caplan S. (Principal), Toscano J.C. (Faculty Sponsor), "Identifying the computational and neural mechanisms of perceptual learning", National Institutes of Health, F32 NRSA Postdoctoral Fellowship, \$197,994 (sub: 8/8/2021, start: 8/1/2022, end: 7/31/2025).

- Toscano, J.C. (Principal). "Immediacy of computation in cognitive processing and perceptual learning", National Science Foundation, \$413,509 (sub: 1/31/2022, start: 1/1/2023, end: 12/31/2025).
- **Freshmen Match proposal (\$1k), Michael Robson**
- **Lorente (Co-PI): ARPA-E COOLERCHPS proposal.**
- 3 proposals submitted by Sergey Nersesov
- **Wemhoff, Aaron P (Principal) "Improving Adsorption Bed Thermal Diffusion in Water-Based Air Conditioning," Sponsored by PA Manufacturing Fellows Initiative, State, \$69,267 (\$72,650 Villanova cost match). (September 2023 – August 2024).**
- Prsa (PI), NASA TESS GI, "TESS legacy in the making: an extended census of eclipsing binary stars", \$250,000
- Prsa (PI), NASA TESS GI, "TESSting the transfer of energy and mass in contact binary envelopes", \$70,000
- Prsa (PI), NASA TESS GI, "The study of Kepler's non-Keplerian orbits with TESS", \$70,000
- Prsa (PI), NASA ADAP, "Solving 150,000+ eclipsing binary light curves with an AI-powered model", \$499,841
- Singh, Pali. Magnetic field mapping-based Battery Management System for Energy Storage Systems, submitted to Consortium for Battery Innovation, \$300,432.

Courses where Augie, Alipi, or clusty have been integrated

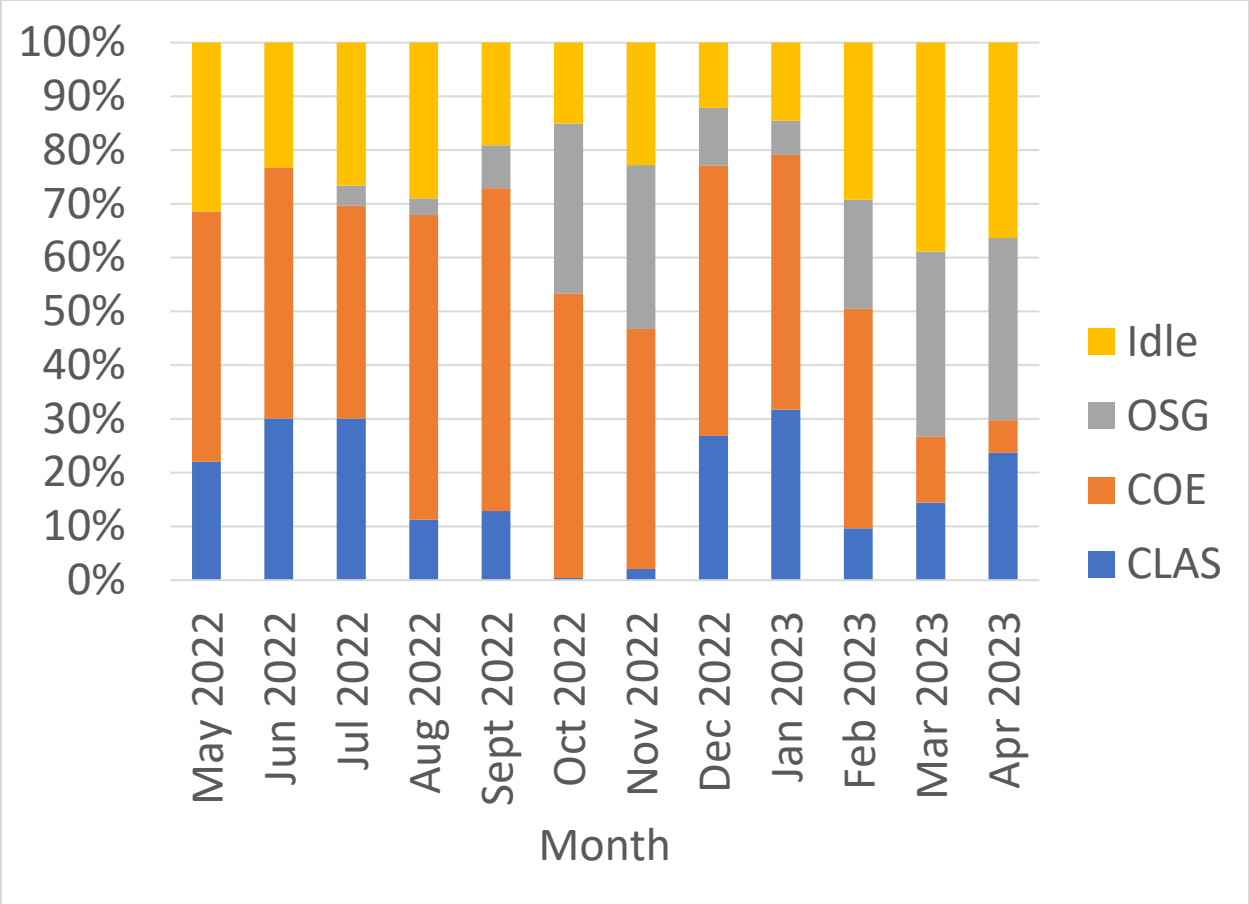
- AST 3148: Principles of Scientific Modeling (Clusty)
- BL 2149: TOP: Cyber Law (Augie)
- CSC 2405: Computer Systems II (Augie)
- CGS 5900/PSY 8900 Cognitive Science Seminar (Augie)
- CGS 5990: Fairness in AI (Augie)
- ME 3600: Fluid Mechanics (Augie)
- ME 7030: Numerical Methods in Engineering Simulation (Augie)

CRCF workshops offered

- Tutorial on PHOEBE (Andrej Prsa)
- Webinar on the Open Science Grid (Aaron Wemhoff)

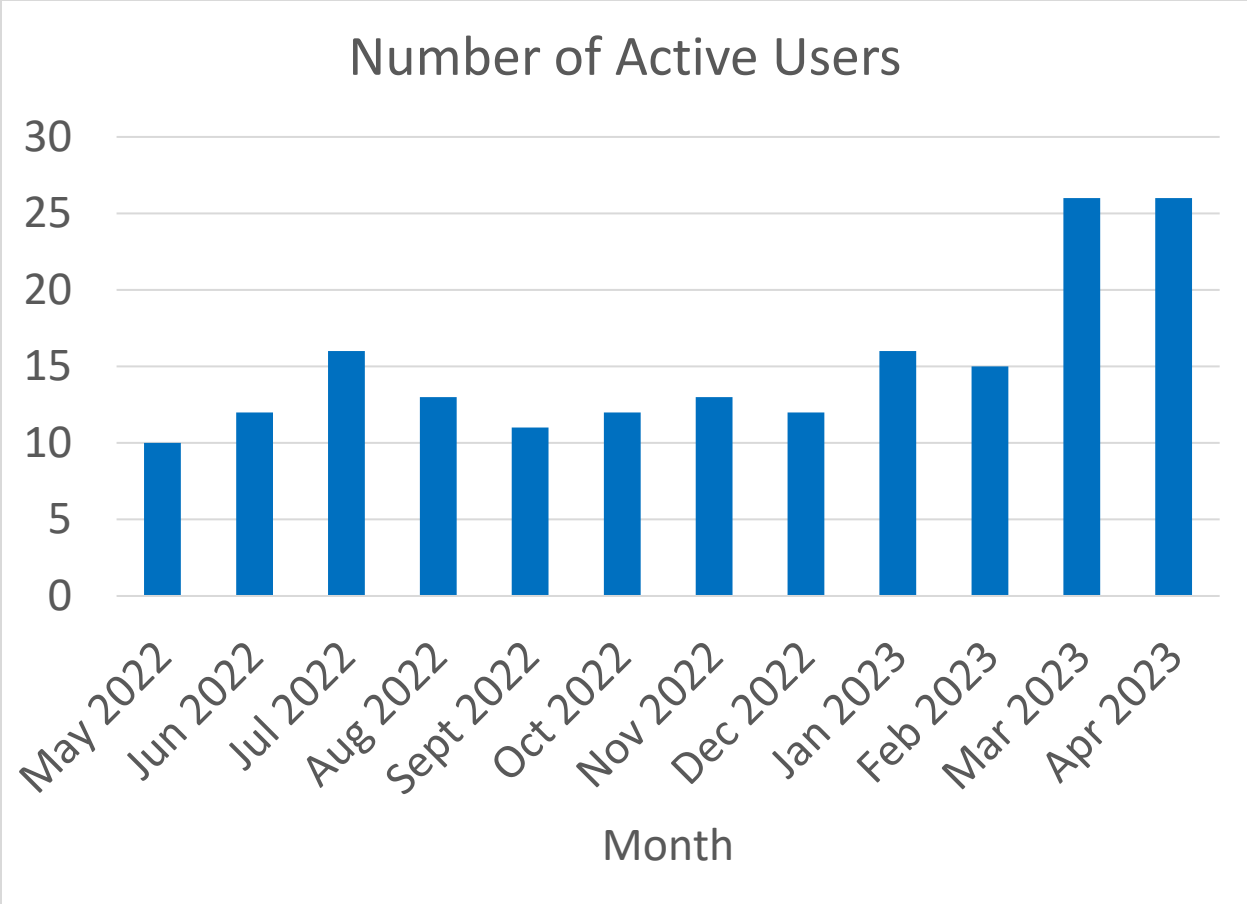
Augie Usage

The usage goals for Augie are (1) less than 20% idle time, and (2) an average of 20% usage by the Open Science Grid (OSG). In the last 12 months, the idle time is slightly high (25%), and the OSG usage is slightly low (15%). These are both acceptable values for purposes of ensuring that Augie's growth matches demand and that the NSF CC* grant requirements are met. It is seen that the CLAS demand varies widely, with little demand for October 2022 but exceeding COE demand in March and April 2023. The average Augie usage per college is 18% for CLAS and 42% for COE, with negligible usage by VSB and zero usage by FCON.



Number of Active Users

Active users are defined as those that submitted jobs to Augie within a specified month. The number of active users over the past 12 months has roughly held steady at 15 per month, with spikes of 26 active users in March and April 2023. These spikes in spring 2023 are possibly due to the integration of Augie into specific classes where students are required to submit jobs.



Work in Progress

The following efforts are currently ongoing:

- CRCF
 - Make a CRCF logo, which formalizes the facility and enhances its visibility among the Villanova community.
 - Finalize hire of HPC IT administrator, which is a major step forward in having a dedicated HPC IT expert who can continue to evolve the cluster and address issues related to users in a rapid manner, enhancing the user's productivity.
 - Formalize the external user onboarding process. This process involves streamlining a workflow that currently requires several university offices.
 - Set up and manage a CRCF social media account. This account will enhance CRCF's presence on campus and draw in more users.
- Augie
 - Patch the operating system (OS), including setting up a process for regular OS maintenance. Formalizing this process will avoid extended downtimes due to major software glitches.
 - Transition the current OS to an alternative one. The current OS is CentOS 7, whose support is ending in December 2023. Possible alternative OS options are Rocky Linux or Ubuntu.
 - Enable containers (singularity) and install OpenFOAM
 - Establish two interactive nodes with Open OnDemand on one of them
 - Address Quantum Espresso performance issue
 - Enable single sign on (SSO) for HPC access. This will prevent users from needing their unique Augie or Alipi passwords to be reset when they lose or forget them.
 - Establish ability to run MATLAB in both modes of parallelization. CRCF successfully characterized the performance of, and provided test scripts for, MATLAB parallelization using basic loop parallelization. Other, more flexible forms of parallelization with MATLAB are now being investigated.

User Survey Responses

A survey was provided to the Augie Users in March 2023. The following feedback was received, and CRCF's response to the comments is provided.

What is going well?

1. I plan to use the cluster as a demo in my course later this semester, so I have not tried to use it in the class yet.
2. Incredibly good support from the team. Any questions I have are immediately answered.
3. Running research experiments that cannot run on my VU laptop

What barriers are you facing?

1. For education: Because many students in the course do not have a background in using these resources, it was impractical to train them to use Augie. The course activities also rely on use of MATLAB with interactive sessions. For research: In some instances, we have experienced

bottlenecks due to the number of jobs in the queue and had to wait 12-24 hrs for our jobs to run.

Response: The College of Engineering and Mechanical Engineering department have chipped in to enable Augie to get two additional compute nodes. Both nodes are planned to act as interactive nodes, one of which containing Open OnDemand. Other universities have used Open OnDemand as an easy way for students to get introduced to HPC.

2. Time limit---I have simple calculations but they would take about a month or two to finish on my laptop. I could use the same simple program on Augie if I were allowed to run a job for a month or two. To get around this, I need to learn how to use the parallel processing efficiently. I have not had time to do this yet.

Response: You may want to discuss with with the Technical Director to see if it is feasible to use parallel processing. In some cases (e.g., MATLAB), it is easy to do.

3. Password reset, data storage limits

Response: Resetting the password should be easy – you just email the CC committee to reset it, and then attend CRCF office hours to get re-established. You can request additional data storage if you need it. Note that UNIT is working on integrating Augie access into single sign-on (SSO), which will alleviate password problems.*

4. Just me and time. I was using clusty much more regularly at one point. But last ~year have not been able to use either.

Response: N/A

Suggestions regarding the overall user experience

1. A database of SLURM commands with tips and tricks from other Villanova Users

Response: The Software Wiki (now the Augie App Notes on the SharePoint site) contains a number of pages dedicated to different applications software, and there is a single page for SLURM commands. Tips & tricks should be posted on these pages. Some (e.g., R, COMSOL) are more populated than others.

Final Points

Good Citizenship

Some reminders:

- No running jobs on the head node – submit batch jobs instead
- Don't use the debug queue for production runs
- Don't submit lots of jobs to occupy a large percentage of cores on the cluster
- Be sure to acknowledge use of the HPC clusters – see user terms and conditions documents:
 - Augie: Augie Users SharePoint, file [Administrative/Augie-HPC-TC 20210507.pdf](#)
 - Alipi: Villanova HPC Team → Alipi Users Channel, file Alipi Usage Policies.docx
 - Clusty: contact Andrej Prsa
- Need help on Augie/Alipi/Clusty?
 - Don't contact UNIT directly or put in ticket
 - Augie:
 - Find appropriate document in the [Augie App Notes](#) documents. See if the document specific to your app has the information you need.
 - Many documents have listed superusers; contact the superuser for help.
 - If the above two items don't work, then email ccstarcommittee@villanova.edu.
 - Alipi: post issue in Teams channel. If no response in 24 hours, then email engineering-hpc@villanova.onmicrosoft.com.
 - Clusty: contact Andrej Prsa

Interested in Helping CRCF?

Contact ccstarcommittee@villanova.edu if you are interested in...

- Supporting other Villanova researchers as a superuser
- Collaborating with a researcher at a small local college
- Providing additional feedback on CRCF operations
- Working as an undergrad support software technician this summer
- Providing your thoughts about how to make AugieFest a popular, fun event