

CRCF Annual Public Report

Academic Year 2022

Last updated 4/29/22

Contents

- Overview 2
- CRCF HPC Cluster Specifications 4
 - Augie 4
 - Alipi & Clusty 5
 - Key Personnel (April 2022)..... 5
- Accomplishments..... 6
 - Administrative..... 6
 - Technical 6
- Metrics 7
- Work in Progress..... 10
- User Survey Responses 11
- Final Points..... 13
 - Good Citizenship 13
 - Interested in Helping CRCF?..... 13

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 interested in research computing but do not 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, but communications between researchers needs improvement. CRCF aims to provide a peer-to-peer communication opportunity to help users with software and applications issues. 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. The Coding Center is slated to begin in the Fall 2022 semester.

CRCF HPC Cluster Specifications

Augie

Item	Values
Compute Nodes/Cores	20 compute nodes/1,440 AMD EPYC Series
RAM	~10 TB Main Memory
Disk Space	292 TB
GPU	2x Tesla A100: 6,912 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, MATLAB
2. The following software applications has been installed, and performance testing and sample script development is in progress: QuantumEspresso, LAMMPS, Python, Comsol, CP2K, VASP
3. The following software applications have been installed, but no performance testing has been done: R, PyMT, CUDA, TensorFlow, PyTorch, Intel compilers, AMD compilers
4. The following software applications are slated for installation: IQTREE, PAML, RAXML, OpenFOAM, Singularity
5. The following software applications need further exploration or discussion prior to initiating an installation effort: Mathematica, Abaqus, ANSYS, Anjuta

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.

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 (April 2022)

CC* Committee: Jules Blaustein (student, CLAS), David Cereceda (COE), Jonathan Graziola (CLAS IT), Jonathan Hardy (UNIT), Ryan Jorn (CLAS, Technical Director), Aaron Wemhoff (COE, Administrative Director)

Key UNIT Support (Augie): Peter Palladino, Gavin Printz, Christopher Washburn

HPC Advisory Board: Andrej Prsa (CLAS), Daniel Smith (CON), Bill Wagner (VSB)

External Advisor: Jason Simms (Lafayette College)

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

Accomplishments

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

Administrative

The CRCF team spent the AY focusing on growing visibility and Augie usage, establishing general administrative policies, and developing the core structure surrounding the facility to ensure long-term sustainability. Highlights of the facility's administrative accomplishments include:

1. CRCF established its governing structure to balance expertise, interest, and cross-campus representation. The core body of CRCF includes the CC* committee that consists of faculty, IT staff, and a student representative that meets most weeks. The CC* committee meets monthly with an internal HPC advisory board for feedback. Finally, an external advisor from Lafayette College will meet periodically with the team to provide guidance.
2. The first partnerships have been developed for the Southeastern Pennsylvania HPC Consortium. The partners are Joe Giammarco (Eastern U., partner with Andrej Prsa), Kamran Fouladi (Widener U., partner with Aaron Wemhoff), and David Schaffner (Bryn Mawr College, partner with David Cereceda). CRCF is currently developing a workflow for efficient onboarding for external researchers.
3. A web-based onboarding process, including user terms and conditions developed in collaboration with the Office of the General Counsel, has been created and implemented.
4. Websites were created for CRCF and Augie. The CRCF website also includes links to Alipi and Clusty websites, and the CRCF website includes recordings of facility webinars.
5. A Villanova HPC Team was created and includes channels for Augie and Alipi managers and users. Resources for Augie users has been provided via a separate SharePoint site.
6. NSF reporting and facility financial planning has been achieved.
7. The CRCF team met with several academic departments, staff offices, individual researchers, and two faculty candidates (Civil and Environmental Engineering)
8. User support documents and policies were created for Augie and Alipi users. These documents are planned to be organized into User Guides for convenience to users. Some documents include a user agreement, a condo computing protocol, forms related to integrating HPC into the classroom, proposal support sections (e.g., facilities blurb and budget justification template), and acknowledgements statements for inclusion in publications.

Technical

Technical evolution of the facility has centered around Augie's development due to its size and broad usage. CRCF procured and installed Augie, established SLURM queues, integrated Ryan Jorn's nodes into the cluster via CRCF's condo computing model, installed several applications (including remote job spawning via COMSOL), started benchmarking specific packages (MATLAB and C/OpenMPI), connected to the Open Science Grid, and established a secure sign-on procedure using Global Protect.

Metrics

Total number of users (as of 4/18/22)

- Augie: 85
- Alipi: 46
- Clusty: 106 (note: all students in astrophysics & planetary science are on by default)

Courses where Augie, Alipi, or clusty have been integrated

- AST 3142: The Principles of Scientific Modeling (clusty)
- CSC 5930-001 TOP:Quantum Computing (Alipi)
- ME 3950: Heat Transfer I (Augie)
- ME 7030: Numerical Methods in Engineering Simulation (Augie)
- SPA 3200: Intro to Spanish Translation (Augie)
- Note: HPC (Toscano lab machines) also used in CGS 5900 and PSY 8900

Journal papers using CRCF resources

- Prsa et al. (2021), ApJS 258, 16: <https://ui.adsabs.harvard.edu/abs/2022ApJS..258...16P>
- Crinnion, A.M, Toscano, J.C., & Toscano, C.M. (accepted). Effects of experience on recognition of speech produced with a face mask. To appear in Cognitive Research: Principles and Implications.
- Liu Yun, Angel Lozano, Tyson Hedrick, and Chengyu Li "Comparison of experimental and numerical studies on the flow structures of hovering hawkmoths," Journal of Fluids and Structures 107, 103405 (2021).

Conference papers using CRCF resources

- Menglong Lei, and Chengyu Li, "Effects of wing kinematics on modulating the odor plume structures in odor tracking flight," ASME FEDSM, Virtual Meeting, August, 2021.
- Menglong Lei, John Crimaldi, and Chengyu Li, "Navigation in odor plumes: How do the flapping kinematics modulate the odor landscape?" AIAA Aviation, Virtual Meeting, August, 2021.
- Zhipeng Lou, Adrian Herrera-Amaya, Margaret Byron, and Chengyu Li, "Hydrodynamics of metachronal motion: Effects of spatial asymmetry on the flow interaction between adjacent appendages," ASME FEDSM, Toronto, Canada, August, 2022 (accepted).
- Menglong Lei and Chengyu Li, "A Balance Between Odor Intensity and Odor Perception Range in Odor-Guided Flapping Flight," ASME FEDSM, Toronto, Canada, August, 2022 (accepted).

Proposals submitted that would utilize CRCF resources

- Robson VU SRF, Automatically Spreading Communication in Adaptive Applications, \$10k
- Nataraj: 6 proposals
- Chengyu Li, PI, "Collaborative Research: A benchmark investigation on the full wake flow of freely flying animals using a hybrid experimental and computational approach," NSF, \$153,424.
- Caplan S. (PI), Toscano J.C. (Faculty Sponsor), "Identifying the computational and neural mechanisms of perceptual learning", National Institutes of Health, F32 NRSA Postdoctoral Fellowship (sub: 8/5/2021, start: 8/1/2022, end: 7/31/2025).

- Toscano, J.C. (PI), "Immediacy of computation in cognitive processing and perceptual learning", National Science Foundation. (sub: 1/31/2022, start: 1/1/2023, end: 12/31/2025).

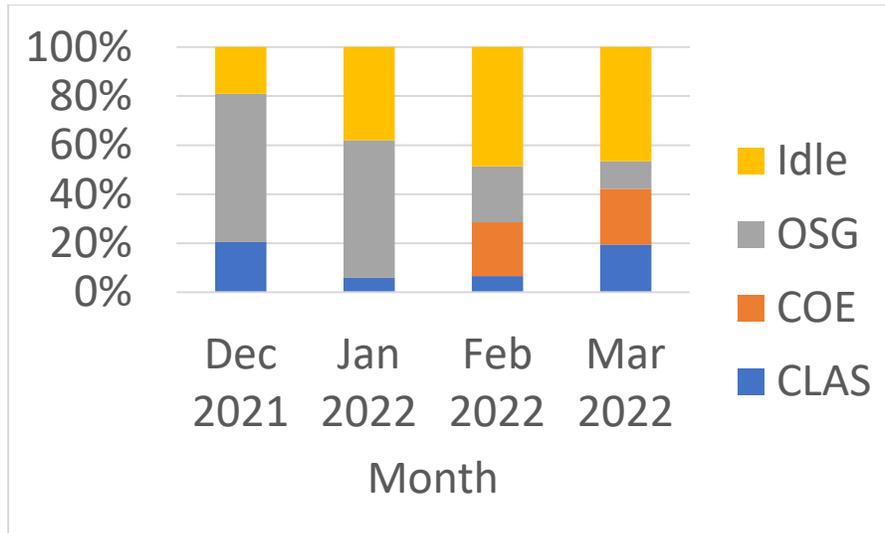
Grants won that utilize CRCF resources in the project scope

- Jorn, NSF, \$185k
- Nataraj, ONR, \$1.075M
- Robson VU SRF, Automatically Spreading Communication in Adaptive Applications, \$10k

CRCF workshops offered beyond monthly new user training sessions

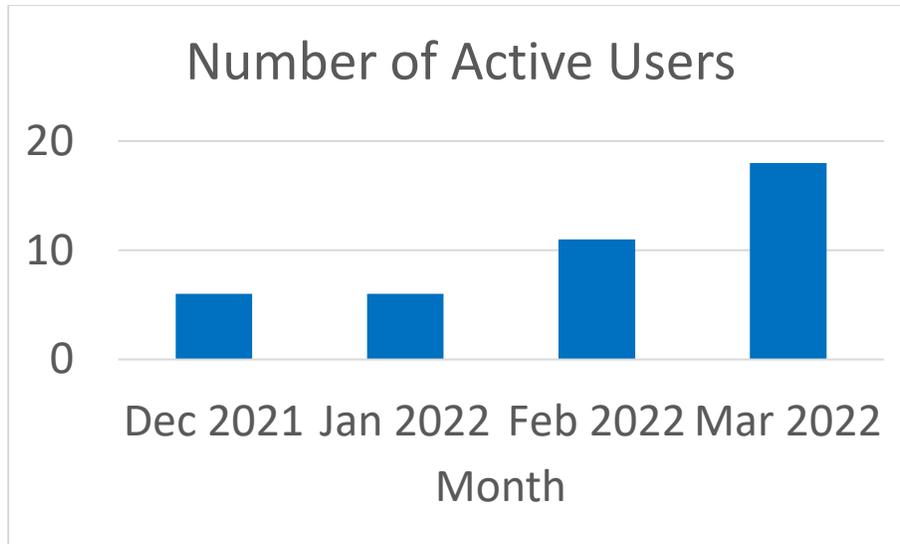
- Augie HPC Kickoff Meeting – Science Driver Users (May 2021)
- Augie HPC Kickoff Meeting – General VU Community (Nov. 2021)
- Integrating HPC into Your Research Plans (Dec. 2021)
- Tutorial on Message Passing in Distributed Systems (Mar. 2022)

Augie Usage



Note that the figure starts in December 2021 since that is the first month where accounting was available. Significant growth in on-campus use is seen between January and March. The large portion of idle time is believed to be related to users requesting more resources than are available at a given time.

Number of Active Users



One can see a threefold increase in active users (defined as running jobs during a given month) between January and March 2022. This trend corresponds to the growth in Augie usage by on-campus researchers.

Work in Progress

The following efforts are currently ongoing:

- Firm up SLURM queuing and enforcement. The current SLURM queues have not been enforced because of low usage in winter 2021-2022. Enforcement will be initiated soon because of increased usage.
- Install user-requested apps. Many users have requested specific applications to be installed. As a result, the waiting list for apps is long, and users are encouraged to proactively install the apps in their home directories.
- 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.
- Enable external (SEPAHPCC) access. CRCF is working on a workflow to more efficiently enable external user access to Augie.
- Complete first draft of Augie management plan. The management plan is an orderly way to enable the CC* Committee to access Augie policies and important documents. An Alipi management plan is already available for the COE HPC Committee.
- Complete first version of user guides. The user guides are similar to the management plans except have a focus on resources and policies pertinent specifically to the users.
- Establish simple onramp tutorials for common languages and apps. Several of these tutorials are envisioned to be enacted this summer.
- Implement more robust strategic planning process. A thorough strategic plan is necessary to establish a mission, vision, and goals with quantifiable benchmarks for the facility.
- Create plan for Coding Center. The third CRCF thrust will be established in the Fall semester. A draft description of the Coding Center will be created this summer and disseminated to the CRCF user community.
- Nominate a Villanova XSEDE Champion. The NSF XSEDE program enables researchers access to larger HPC clusters run by select institutions (e.g., U. Pittsburgh). An XSEDE Champion provides an important link between Villanova and the broader research computing community on best practices.
- Expand our HPC Glossary of Terms. An HPC glossary of terms document allows all users to know the “lingo” associated with this unique field. The glossary has been started but needs completion.
- Hold another tutorial session. Tutorials/workshops are planned this summer on using HPC in astrophysics (Prsa) and an overview of the Open Science Grid (Wemhoff)
- Update user training documents. The accuracy of the user training documents will be verified this summer.
- Establish organization on dissemination of different user issues. The Villanova HPC Teams site, with multiple channels, has a standard chat posting feature. However, it would be beneficial to the community to organize a mechanism for disseminating different issues rather than documenting all issues in the chat board.

User Survey Responses

A survey was provided to the Augie Users in March 2022. The following feedback was received, and CRCF's response to the comments is provided.

1. Augie time limits are causing a problem.

Response: please reach out to us if the queue time limits are causing an issue.

2. Using Augie doesn't fit my use case.

Response: please feel free to reach out to us to help you strategize the best approach towards your problem. I want all Villanova researchers to get into the habit of approaching CRCF when they have a research computing strategy question. HPC is not suitable for all problems, and we can help point users to other solutions. In fact, we are already doing this for some researchers.

3. We want to be able to use containers

Response: We are looking into this possibility.

4. How fast is Augie?

Response: We will add benchmarks that show the speed of Augie compared to serial code and the performance on other machines. We already have performance test files for MPI (C code) and MATLAB (parfor) on the Augie User SharePoint site. It must be stressed, though, that the benefits of Augie are only seen for specific applications (e.g., exceptionally large, parallelizable jobs). A small serial job, for example, may work best on a local workstation. Please contact the CRCF team if you have questions as to whether HPC is the best solution for your specific problem.

5. Logging into Augie is complicated.

Response: The requirement to be on the VPN (Pulse Secure) and Global Protect are because Augie needs to be available to an external network, the Open Science Grid, per NSF grant mandate. This availability requires extra security measures. We have a training video for logging into Augie from a variety of user OS platforms to help in the process.

6. How do I request that specific software apps be added?

Response: Please reach out to us directly to work on installing a particular package. Please note that we require a user to partner with us on software installation to ensure that the software is running correctly. Note that users are free to install their own apps in their home directory.

7. Hire TAs to help faculty when they hit a simple snag or if they want to make an app run faster.

Response: This is the idea behind the Coding Center, which is planned to be ramped up in the fall.

8. Allow for interactive sessions

Response: Purchasing a separate node for interactive sessions is part of a long-term plan.

9. The highest interest is for attending the AugieFest all-hands meeting and short (1-2hr) webinars on skills development. We'll hold off on day-long workshops for now.
10. Provide on-demand video training

Response: We have been meeting with UNIT staff regarding the creation of professional training videos.

11. Make some of the workshops in person.

Response: We plan to hold some future workshops in person.

12. Have people try running sample code.

Response: We already have provided some scripts that can be used after some minor modifications.

13. Onboarding training is long for users who are already familiar with HPC/Linux

Response: We just want to be sure that a baseline is reached for all who onboard.

14. Update new user training pdf. Some commands are old and not working anymore.

Response: We will update the training documents.

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 Augie-HPC-TC_20210507.pdf
 - Alipi: Alipi Users Teams 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: post issue in Teams channel. If no response in 24 hours, 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 through the Coding Center
- 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