

Jetstream NSF Annual Report

December 1, 2019 – November 30, 2020

1. Accomplishments

1.1. What are the major goals of the project?

Jetstream is the first NSF-funded, open, interactive production system targeted at the national science and engineering research communities to provide on-demand computing and storage anytime, anywhere. By design and in practice, Jetstream is programmable cyberinfrastructure in that users can reconfigure the system to suit their needs; this takes many forms, from multi-user platforms that use Jetstream to configure their services to single users who utilize Matlab interactively. Although the SUs allocated via XRAC process resulting in “Research Allocations” are reported quarterly to the NSF and are available in XDMoD, Jetstream is a very different system from other XSEDE resources: Jetstream is a highly usable system and serves those in the long tail of science.

As a system designed for high-throughput, interactive, on-demand workloads, it is useful to consider the capabilities and value of Jetstream in aggregate, just as one would for commercial cloud systems, rather than focusing exclusively on HPL performance and core-hour delivery. Jetstream was built so that scientists with no batch programming experience could use it, and so that those with more technical acumen could reconfigure it to suit their needs. Jetstream is programmable cyberinfrastructure and, thus, must be assessed in a more holistic fashion. In the *Disseminations of Results, Products, and Impact* sections of this report, we provide a view into the value proposition through the lens of researchers, the XSEDE community, and educators demonstrating impact beyond what is discussed in terms of return on investment¹, and beyond what is presented in XDMoD charts for quarterly research allocations. As the open cloud for science, Jetstream provides:

- “Self-serve” academic cloud services, enabling researchers and students to select a VM image from a published library or, alternatively, to create or customize their own virtual environment for discipline- or task-specific personalized research computing.
- Hosting of persistent VMs to provide services beyond the command line interface for science gateways and other science services.
- New modes of sharing computations, data, and reproducibility.

¹ Stewart CA, Hancock DY, Wernert J, Link MR, Wilkins-Diehr N, Miller T, Gaither K, Snapp-Childs W (2018) Return on Investment for Three Cyberinfrastructure Facilities: A Local Campus Supercomputer, the NSF-Funded Jetstream Cloud System, and XSEDE (the eXtreme Science and Engineering Discovery Environment). *2018 IEEE/ACM 11th International Conference on Utility and Cloud Computing (UCC)*, :223–236.

- Expanded access to the NSF XSEDE (the eXtreme Science and Engineering Discovery Environment) ecosystem by making virtual desktop services accessible from institutions with limited resources.

While Jetstream is a hardware acquisition, the system is both a pilot and production system that is made novel by its software components and interfaces. The Jetstream team's overarching goal is to continue improving upon existing functionality throughout the duration of the award by increasing stability and adding new features.

1.2. What was accomplished under these goals?

1.2.1. Major Activities

- The project team provided services to 2,601 researchers and educators, including 1,046 students, during this period. Jetstream is serving 393 projects in 75 fields of science for users at 402 institutions.
 - This user data is what is tracked for XSEDE-supported resources and represents the rolling 1-year allocation window for direct access to Jetstream through the user portal and APIs; it does not include indirect access such as users of science gateways.
 - The 48 science gateways on Jetstream have supported over 42,000 users to date, most without knowing they're directly using the resource. The science gateway community continues to improve tracking of unique users but some gateways (e.g., for geoscience data distribution) are open to the public and do not track individual users. These figures represent an increase in over 20 gateways and 10,000 users from the prior annual reporting period.
- The Jetstream team, researchers, and peers at other institutions continue to produce high-quality products in the form of journal articles and juried conference papers. Additionally, book chapters are now listed for this annual period. As of November 18, 2020, Jetstream has been cited 152 times and users have acknowledged Jetstream (and their XSEDE allocations) in additional publications. For reference, Bridges (a larger XSEDE resource which was funded by the NSF at the same time as Jetstream) has been cited 121 times. Additional details are available in Section 4.2.
- The Jetstream team met at the University of Arizona on January 28-29, 2020 to discuss project priorities and plan future work. OpenStack service enhancements and updates were planned and work began to bring the TACC cloud to the same OpenStack service release levels as IU. Alternative user interfaces were discussed and demonstrated. Usage reporting was discussed at length and a plan to unify reporting methods between the clouds was laid out for pursuing after the meeting. Metrics reporting, including ongoing efforts to integrate with XDMoD were discussed. Devising a plan to accommodate a growing need and desire for container orchestration was also put forth.
- The Jetstream Stakeholder Advisory Board meeting occurred on January 30, 2020 in Tucson, AZ. The Jetstream team presented the results of Operational Year 4 to date.

The SAB feedback was encouraging to the Jetstream team and provided guidance on future opportunities. Four of the eleven SAB members were in attendance as well as representatives from IU, TACC, and UA.

- Jetstream leadership from IU conducted a joint review of Operational Year 4 to date through a virtual panel with attendees from the NSF Office of Advanced Cyberinfrastructure on February 4, 2020. A recommendation to extend spending authority for Operational Year 5 was submitted with approval received on July 13, 2020 for the period through May 2021.
- The COVID-19 pandemic has resulted in primary operations of Jetstream being conducted remotely since March 2020. There has been no decrease in level of support or operational stability with this change.
- Jetstream joined the COVID-19 HPC Consortium in March 2020 to contribute resources to the pandemic response. The team has participated in joint reporting efforts to the NSF throughout the pandemic with numerous projects taking advantage of the resource. These COVID-19 related projects include gateways such as Galaxy (used for some of the earliest genomic analysis) and ChemCompute (used by students for remote learning), as well as for international medical records projects like OpenMRS, and regional information such as projections by UTSA for Texas hospital bed and ICU bed usage.
- An extension of Jetstream for an additional operation year was requested to continue operations through November 2021. The supplement was approved on August 21, 2020.
- Jetstream received a supplement for an REU program at IU. Five students began the 8-week program on June 1, 2020. The students worked remotely with IU staff and non-IU faculty on discipline-specific projects. The program concluded with student research poster submissions to SC20. Three posters were accepted and the students will be presenting in November 2020. Tenecious Underwood, from the Jetstream 2019 REU program returned this summer, and acted as a junior mentor to the NSF-funded students as another pilot of a hybrid funding possibility. The projects were: “HPC Ranking Based on Real Applications”, “Automatic Capture and Classification of Frog Calls”, “Visualizing metagenomic data in R using Jetstream”, “Best practices for building a HIPAA compliant Docker container for Jetstream”.
- AT PEARC20, numerous papers, tutorials, and posters involved Jetstream as they were offered by IU-affiliated authors or by researchers leveraging Jetstream for their work.
 - One half-day tutorial used the platform:
 - Deep Dive into Constructing Containers for Scientific Computing and Gateways presented by IU and Cornell (Jeremy Fischer, Eric Coulter, Steve Bird, Sanjana Sudarshan, Suresh Marru, and Peter Vaillancourt)
 - Multiple papers and talks that utilized Jetstream were given, including these:
 - A talk on a science gateway using Jetstream was presented. The title: *FutureWater Indiana: A science gateway for spatio-temporal modeling of water in Wabash basin with climate change in focus*

- *Custos: Security Middleware for Science Gateways* was a paper/talk submission from XSEDE and SGCI staff that utilized Jetstream in testing and for development of a future tutorial (Gateways 2020)
 - Invited panel presentation:
 - A plenary panel entitled *Introduction of the new NSF Innovative HPC Systems* included a presentation on Jetstream2 with initial details on the transition plan from Jetstream.
- The Jetstream team has made a decision to hold on the Rocky version of OpenStack. Proceeding beyond Rocky would require an upgrade to the underlying operating system from CentOS-7 to CentOS-8. This upgrade would require a lengthy downtime and could potentially represent a significant risk to the project. An end date for extended maintenance support for OpenStack Rocky has yet to be determined; but, a reasonable extrapolation from earlier versions would be beyond January 2022. Maintenance update support for CentOS-7 will continue until 30 June 2024.
- The Jetstream team has back ported features available in post-Rocky versions that were requested by users; e.g. features desired from the Train version of Magnum have been implemented.
- The Jetstream team continues to stay current with operating system security updates for core services as well as featured images available to the community.
- Reporting of Atmosphere VM utilization by the Texas Accounting System (TAS) to the XCDB has been constrained starting 13-Mar-2020. As of early November-2020, TACC has worked with XSEDE to bulk upload historical data, reestablish real-time statistics, and are working to clear any duplicate entries.
- Findings from the 2019 Jetstream Annual User Assessment (2020 survey is still in the field at the time of writing) include the following:
 - Overall, over nearly 88 percent of all respondents report they are either “satisfied” or “extremely satisfied” in their experience using the Jetstream system. Mean satisfaction with Jetstream is 4.30 (on a 5.0 scale).
 - Users note the highest levels of satisfaction with the quality of responses to questions, whether via help@xsede.org (4.53/5.0) or via direct email to Jetstream staff (4.51/5.0), followed by the speed of response to questions via direct email to Jetstream staff and/or help@xsede.org (4.48/5.0) and the quality of in-person training and workshops (4.35/5.0).
 - 74% of respondents report that Jetstream is “very important” or “essential” to their research activities with a mean importance rating of 4.14/5.0.
 - 73% of users indicate the resource is “very important” or “essential” to their education activities, with a mean importance rating of 4.00/5.0.

1.2.2. Specific Objectives:

Accomplishments relevant to the achievement of goals for this project are described below:

Indiana University

- The team performed general hardware maintenance, updated firmware, and security software updates. Specific details on operating system and OpenStack version updates are highlighted above.
- The annual Jetstream user survey was released in Q3 2020 and the results will be available in Q1 of 2021.

TACC

- The team performed general hardware maintenance, firmware updates, and security software updates.
- The Jetstream-TACC cloud was upgraded to include the Openstack:Magnum container orchestration package. Full integration and optimization will be completed by the end of the reporting period.
- Accounting systems Gnocchi and Ceilometer had revised to compensate for in warranty disk failures to prevent performance loss.
- TACC experienced some hard disk failures (without data loss). In-warranty replacement is in progress and will be completed during the reporting period.
- The TACC TAS-XSEDE AMIE accounting system has successfully processed the remaining backlog of accumulated accounting packets. Staff at TACC are working with XSEDE staff to identify and remove a small number of artifactual packet duplicates that resulted from the rapid clearing of the backlog.
- TACC's TAPIS (<https://tapis-project.org/>; https://www.nsf.gov/awardsearch/showAward?AWD_ID=1931439&HistoricalAwards=false) and ABACO (<http://useabaco.cloud/>; https://www.nsf.gov/awardsearch/showAward?AWD_ID=1740288&HistoricalAwards=false) technologies, supported directly by Jetstream, are the underlying technology for the SD2E (Synergistic Discovery and Design Environment, SD2E ; <https://sd2e.org>) initiative; DARPA, HR001117S0003) as well as TACC's own UTRC user portal (<https://utrc.tacc.utexas.edu>) that allows integrated graphical access-to, use-of, and job-deployment on the majority of TACC's cyberinfrastructure resources.

University of Arizona

The University of Arizona team monitors, maintains, and improves upon Atmosphere on an ongoing basis, generally in response to user support requests or partner feedback. Significant improvements include:

- Ensured compatibility of Atmosphere deployment and configuration process (via Ansible) with CentOS 8 and began work for compatibility to Ubuntu 20.04;
- Improvements to ansible processes to deploy and configure instances faster and more reliably;
- Added the ability to launch multiple instances within the Atmosphere UI;
- Integrated Argo workflows as a replacement to Python celery;

- Added scripts to provision multiple instances for a class or workshop for one or more user accounts in Atmosphere;
- Added a script to reset a user account to a default state, which is also useful for workshop demo accounts;
- Added enhancements, scripts, and other processes to better monitor Atmosphere or recover from certain failures.

Johns Hopkins University (JHU)

- JHU has continued to offload a significant portion of the workload from the Galaxy Main (usegalaxy.org) gateway onto Jetstream. Slurm-based clusters have been deployed at IU and TACC and integrated with other infrastructure available to Galaxy Main. In the course of this year, we have made these clusters dynamically scalable to balance peak and off-peak utilization. The resource capabilities available on Jetstream, such as the ratio of CPUs and memory, are not available from other resources allowing Galaxy Main to provide mass access to tools not otherwise possible. During the past year, 177,314 jobs were executed on Jetstream on behalf of 12,391 individual users.
- The Galaxy project has an active outreach program and has held many interactive workshops. During these workshops, it is critical to have responsive compute infrastructure to ensure predictable workshop flow. Using a shared resource, such as Galaxy Main, does not yield such experience. To streamline the user experience, we have established a Workshop instance of Galaxy that provides a dedicated resource for interactive workshops and is linked to Galaxy Main at the database level allowing users to continue where they left off with their analyses even after the workshop ends. This Workshop Galaxy instance is scaled manually prior to the workshop. Thus far six workshops were held using this infrastructure.
- JHU has completed the initial implementation of the next-generation deployment model for Galaxy that is based entirely on modern containerization technologies. This Kubernetes-based implementation provides a more robust deployment paradigm that is gradually being rolled out across the Galaxy community. Jetstream was one of the pilot deployment targets. Anyone is able to launch their own instance of Galaxy based on this paradigm using the self-serve CloudLaunch portal, available at <https://launch.usegalaxy.org/>. Blog posts detailing the advancements of the new model are also available: <https://galaxyproject.org/blog/2020-02-gvl5-beta/>, <https://galaxyproject.org/blog/2020-04-gvl5-beta2/>, <https://galaxyproject.org/blog/2020-06-gvl5-beta3/>, <https://galaxyproject.org/blog/2020-07-gvl5-beta4/>.

University of Chicago

- The University of Chicago team provided user authentication services for Jetstream (Globus Auth) that allowed users to login to Jetstream's web user interface using XSEDE, campus, and/or national identity provider services.
 - This year, the team refreshed the user interface for logins.

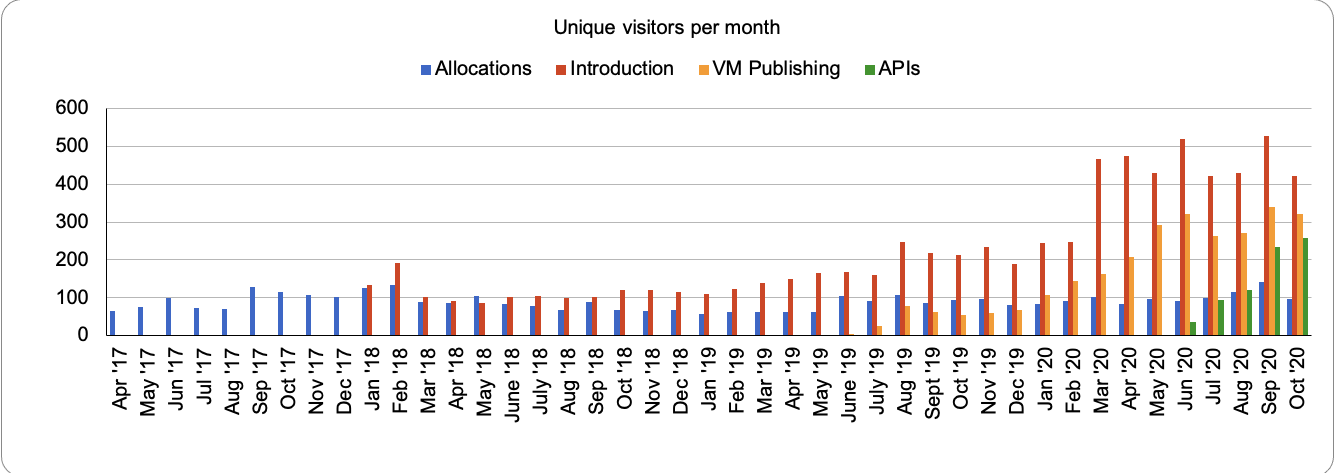
- The team provided routine weekly updates with bug fixes, security fixes, and new features for other systems. It also verified Jetstream logins after each update to ensure there were no unexpected issues.
- The University of Chicago team also reviewed and approved routine requests for Globus Plus access by researchers using the Jetstream system. Globus Plus allows researchers to transfer between two Globus Connect Personal endpoints. This enables Jetstream users to transfer data between their Jetstream VMs and their personal systems at home or on campus, with Globus Connect Personal installed in both places.

The University of Texas at San Antonio

- The UTSA team develops and maintains AI workshop training (<https://github.com/utsabigdata/AI-Workshop>) and Big Data Technology hands-on training (<https://github.com/utsabigdata/bigdata>) through Jetstream and the JupyterLab environment (PyTorch library) for classroom training as well as local community outreach and professional education training.
- The UTSA team integrates and maintains cyber security online labs, called SEED Lab (<https://github.com/seed-labs>), for cyber security education on Jetstream. SEED Lab, funded by NSF and now used by over a thousand educational institutes worldwide, is developed with the objective of supplying students with experimental cybersecurity learning.
- The online training topic “Introduction to Jetstream for Data Science using Elastic” developed and conducted by UTSA Open Cloud Institute as a virtual workshop in Sept. 2020 for minority undergraduate and graduate students at UTSA campus.
- The UTSA team developed a single pane of glass for the United States COVID-19 Dashboard on Jetstream. An Interactive dashboard that gathers data from various locations to explore COVID-19 related statistics, community mobility, and Hospitalization Rates (for a certain period of time). Data is gathered from the CDC, the New York Times, Google's Community Mobility Database, and UsaFacts.org on the National, State, and County levels.

Cornell University

- The Cornell team develops and maintains Jetstream online training in the Cornell Virtual Workshop (CVW) format. CVW topics are freely available at all times to the entire scientific community – researchers, cyberinfrastructure practitioners, students, and educators. We are currently developing the final CVW topic on the Jetstream API and other advanced topics.
- The online training topic “Introduction to Jetstream” developed in the Cornell Virtual Workshop was completed in January 2018.
- The online training topic “Introduction to Jetstream” developed in the Cornell Virtual Workshop was updated in May 2018 and again in November 2018.
- A third workshop on Publishing Jetstream Virtual Machines was completed in July 2019.
- The fourth workshop “Using the Jetstream API” was completed in June 2020 by the Cornell Virtual Workshop team.



Unique visitors to Jetstream virtual workshops by Cornell

1.2.3. Significant results

We have the following significant technical results so far:

Metric	Goal per program year	Q1PY4 Dec19 - Feb20	Q2PY4 Mar20 - May20	Q3PY4 Jun20 - Aug20	Q4PY4 Sep20 - Nov20 **	Achieved Dec20 - Nov20 **
System availability (uptime of an element of the production hardware, as % of wall clock time)	95%	100%	100%	100%	100%	100%
Capacity availability (% of the total capacity available for NSF use over time)	95%	99.96%	99.52%	99.39%	99.93%	99.70%
Job completion success - % of jobs submitted should complete without having to be resubmitted as a result of a failure in the hardware or system software.	96%	99.03%	96.52%	96.13%	99.46%	97.79%
Average number of active VMs	320	1121	1299	1081	1215	1179
CPU % utilization*	6%	7.62%	5.18%	5.42%	5.38%	5.90%
Core cloud environment software will be upgraded to match current versions of components such as operations systems and cloud software environments.	Updated prudently, generally within < 12 months of major releases	Rocky	Rocky	Rocky	Rocky	Rocky

***Note:** Dropping CPU utilization is a result of improved efficiencies in the storage system.

****Note:** November 2020 stats are up to and including 12-Nov-2020

Additional statistics of note:

Metric	Goal per program year	Achieved December 2019 - November 2020	Notes
Capacity of system allocated via NSF-specified allocation process	90%	>100%	Annual goal exceeded
Total number of distinct users (annual total)	1,000	2,601	Annual goal exceeded
Total number of students having used Jetstream in an educational or training setting (aggregate total)	1,750	1,046 - Dec 2019 to Oct 2020	3,715 - aggregate total through October 2020 There is considerable churn over the course of the year. Our aggregate total far exceeds the goal.
Total number of science gateways using Jetstream (running total)	15	48	Annual goal exceeded
SUs available to user community per month	1.47M (per month)	>4.0M	Annual goal exceeded
% of SUs available to user community that were used per month	Goal not established	>100%	
Total number of VM images and/or datasets published with a DOI via IUScholarWorks	150	12 (running total)	Despite advertisement of this service, there has been low uptake.

Allocation statistics for December 2019 through October 2020:

	Q1 PY5 Dec 2019 to Feb 2020	Q2 PY5 Mar-May 2020	Q3 PY5 Jun-Aug 2020	Sep 2020	Oct 2020	Total to Date
<i>Startup</i>						
Total requests	20	29	25	6	8	491
SUs requested	8,322,560	1,939,000	1,450,000	750,000	630,000	57,574,742
SUs awarded	8,322,560	1,939,000	1,450,000	750,000	630,000	57,574,742
<i>Educational</i>						
Total requests	11	8	16	4	4	177
SUs requested	5,919,120	1,209,260	4,154,340	2,362,200	411,000	48,551,300
SUs awarded	5,919,120	1,209,260	4,154,340	2,362,200	411,000	43,551,300
<i>Campus Champion/Staff</i>						
Total requests	35	39	35	7	15	619
SUs requested	1,710,000	1,910,000	1,750,000	350,000	750,000	33,596,500
SUs awarded	1,710,000	1,910,000	1,750,000	350,000	750,000	33,596,500
<i>Supplemental/Discretionary/Net Transfers</i>						
Total requests	21	29	10	5	5	284
SUs requested	2,243,480	15,034,193	1,306,365	2,200,039	3,670,000	75,462,061
SUs awarded	2,243,480	15,034,193	1,306,365	2,200,039	3,670,000	75,462,061
<i>Research allocations</i>						
Total requests	11	14	14	N/A	12	158
SUs requested	6,666,806	5,626,880	22,288,021	N/A	9,170,300	140,242,503
SUs awarded	9,125,046	6,891,880	19,267,941	N/A	9,170,300	152,188,242
Total request and allocations						
Total requests	98	119	100	22	44	1729
Total SUs requested	24,861,966	25,719,333	30,948,726	5,662,239	14,631,300	355,427,106
Total SUs awarded	27,320,206	26,984,333	27,928,646	5,662,239	14,631,300	362,372,845

1.2.4. Key outcomes or Other achievements

New (major) features:

- The Magnum project was made available on Jetstream-TACC (Rocky release).
- Jetstream continues to support the Manila project in beta; this allows users to create network file shares easily from block storage.
- Jetstream has ported future OpenStack project features into the Rocky install when requested by users (e.g. installing the Train version of Magnum to provide better container orchestration integration with Kubernetes).

Jetstream REU program:

Students focused on four projects. This summer's program resulted in three poster presentations given at the SC20 conference. The three projects were:

HPC Ranking Based on Real Applications. REU students: Nolan Baker, Aaron Jarsmuch. Performance benchmarks are used to stress test hardware and software of large scale computing systems. A corporation known as SPEC has developed a benchmark suite, SPEC ACCEL, consisting of test codes representative of kernels in large applications. This project ranks the published results from ACCEL based on different criteria. The goal is to prepare a ranking website for the work-in-progress real-world SPEC HPC benchmark suite, HPC2021 that is soon to be released.

Automatic Capture and Classification of Frog Calls. REU student: Eliza Foran. Junior mentor: Tenacious Underwood (from REU 2019). Global frog populations are threatened by an increasing number of environmental threats, such as habitat loss, disease, and pollution. Traditionally, in-person acoustic surveys of frogs have measured population loss and conservation outcomes among these visually cryptic species. However, these methods rely heavily on trained individuals and time-consuming field work. We propose an end-to-end workflow for the automatic recording, presence-absence identification, and web page visualization of frog calls by their species. The workflow encompasses recording of frog calls via custom Raspberry Pi's, data pushing to the Jetstream's cloud, and species classification by three different machine learning models: Random Forest, Convolutional Neural Network, and Recursive Neural Network. The goals of this project were to: 1) Create an end-to-end analysis system for frog surveys. 2) Handle multiple frogs calling in the same recording. 3) Identify a variety of species, including closely related species. 4) Handle unknown calls and sounds in recording (i.e. birds, talking, etc.). 5) Provide user-transparent information on probability of call identification.

Visualizing metagenomic data in R using Jetstream. REU student: Haley Leffler.

- After the 2010 BP oil spill, it was estimated that half of the oil was removed from the ocean by physical and chemical methods and the other half was unaccounted for at the time.
- Metagenomic techniques allow researchers to study microbial communities, and revealed their ability to degrade oil or hydrocarbons.
- Data for this project was downloaded from a research study focusing on the microbial response to hydrocarbon seepages (the natural release of oil or gas bubbles from the ocean floor shown in Fig 1). This study found that hydrocarbon seepages can significantly alter the microbial community.
- Researchers applying metagenomics techniques generate large datasets to capture the complexity of the microbial community and face challenges during visualization. Using several visualization methods to display the data from multiple samples allows for exploratory analysis and represents the data from different points of view.
- This project performs an exploratory analysis of metagenomes using Jetstream, a cloud-based infrastructure that aids analysis of large datasets.

An additional project, not accepted at SC20:

Best practices for building a HIPAA compliant Docker container for Jetstream. REU student: Jonathan Holzmann. Many organizations including hospitals, financial institutions, and universities are adopting cloud-based infrastructure. To facilitate replicability and ease of deployment, specifically containers are being implemented for the flexibility and scalability they offer. Some industries, for example, the healthcare industry have some complex processes to maintain HIPAA compliance. Containers, however, are not inherently compatible with HIPAA. Here we explore building a HIPAA compliant Docker container (for cloud computing systems) adopting best practices for encryption of data at-rest and data in-transit.

Major activities supported by discretionary usage

- Jetstream staff, with assistance from XSEDE staff, continue to support *Jetstream Trial Accounts* via an XSEDE-User-Portal-integrated mechanism that facilitates adoption of Jetstream by researchers. Since its inception, this mechanism has enrolled nearly 1 user per day on average. During the past year we have moved these users to leverage discretionary time. Discretionary usage is tracked in the XSEDE central database for Jetstream.
- The Jetstream teams work Arizona State University on expansion of Jetstream led to development of an Major Research Instrumentation response, a dedicated GPU system purchased by ASU installed in the test and development resource hosted at the University of Arizona, and an ASU partnership in the Jetstream2 award.
- IU has added additional large memory and GPU nodes to Jetstream-IU for discretionary testing and usage. In early performance testing there is no degradation when using GPUs from within a virtual machine versus bare metal. IU also tested virtual GPU functionality (ability to slice a single GPU and share resources to multiple VMs/users) and found little to no reduction in performance. A pre-print of a paper in development is available at <https://scholarworks.iu.edu/dspace/handle/2022/24567>.

- Testing this functionality within Jetstream was an important part of an “*AI for Everyone*” initiative at IU with a goal of spreading AI-enabled tools to more domains and users. These efforts were also beneficial for developing the Jetstream2 architecture. This specific paper was presented at the SC20 SuperCompCloud workshop in November 2020.
- A number of project have explored GPU usage on Jetstream:
 - The SASSIE gateway (<https://sassie-web.chem.utk.edu/>) experimented with using the GPUs for their workflows. This gateway presently uses Jetstream resources via the GenApp gateway and may move to Jetstream once GPUs are more readily available. They saw significant speedup in their work using Jetstream GPUs. They did a full comparison of all flavors at our request to benchmark performance with different sized vGPUs.
 - Namdrunner (<https://namd.jetstream-cloud.org/namdrunner/>) also was an early tester of GPU resources. They did a full comparison of all flavors at our request to benchmark performance with different sized vGPUs. Their work using NAMD sees significant improvement using GPUs.
 - Dr. Paul Rad’s Open Cloud Institute at University of Texas San Antonio is presently using the GPUs both as part of educational efforts, teaching students about machine learning as well as his research in real-time recognition analysis.
 - Dr. Ali Maga from University of Washington is prototyping a workflow for 3D Slicer, the open-source biomedical visualization platform (another NSF-funded project) on Jetstream. He anticipates scaling it up on Jetstream2 when it comes online with additional GPU facilities.
 - Dr. Mark Frank of the IU School of Medicine is prototyping a machine learning and analysis workflow for analyzing chest CT scans for identifying potential health issues.
 - Dr. Xukai Zou of the IU School of Informatics is doing research on computer security systems, specifically targeting protected health information. His work is using the GPUs presently for machine learning and analysis.
 - Dr. Vitaly Ford of Arcadia University used the GPUs for his computer science Masters’ students capstone projects in machine learning.
 - The Galaxy Project has recently begun developing GPU workflows for the Galaxy gateway using Jetstream GPUs.
 - NCGAS developed a workflow for BEAST and BEAGLE bioinformatics software for one of their researchers.
 - Several other researchers have expressed interest in using the Jetstream GPUs, though have not begun using them at this time. We anticipate several more projects being onboarded by early Q1 2021.

1.3. What opportunities for training and professional development has the project provided?

The Jetstream team has enhanced the training and development of many CI professionals, domain scientists, and students through outreach efforts. Outreach events have been a key driver for the Jetstream team and while the planning was very promising for this period outcomes radically deviated from the plans due to COVID-19. As we settled into the realities of 2020, a number of these events were canceled but many became virtual, allowing the team to continue with education and outreach efforts in a different way. Many of these events are focused on training and professional development for researchers, educators, and students.

Numerous members of the Jetstream team virtually attended the PEARC20 conference in July online to deliver and attend workshops and tutorials as well as the Gateways 2020 conference in October.

Jetstream has provided the following training opportunities:

- Jeremy Fischer conducted hands-on tutorials at the virtual RMACC HybridCloud event in May and SGCI Coding Institute virtual event in July 2020.
- Sanjana Sudarshan conducted a Hands-on with Jetstream Atmosphere tutorial in February at Indiana University Bloomington with remote participants from remote IU campuses as well as virtually for the Consortium for Computing Sciences in Colleges Northwestern event in October.
- Jeremy Fischer, Sanjana Sudarshan, Steve Bird, Eric Coulter, and Suresh Marru along with Peter Vallaincourt from Cornell conducted 2 half day tutorials and 1 full day tutorials titled “Deep Dive into Constructing Containers for Scientific Computing and Gateways” for the SGCI Coding Institute, PEARC20, and Gateways 2020.
- Brian Beck taught at the TACC Summer Institute, “Reproducible Science” on July 1-22, 2020: <https://www.tacc.utexas.edu/education/institutes/reproducible-science>
- Jetstream was also an integral platform for two other TACC Summer Institutes:
 - “Advanced Computing Essentials (ACE)”, June 9-30, 2020 : <https://www.tacc.utexas.edu/education/institutes/ace>
 - “Machine Learning”, August 10-14, 2020: <https://www.tacc.utexas.edu/education/institutes/machine-learning>

In addition, the five students who were recruited for the REU program attended IU-based trainings that covered a variety of topics including Intro to Unix, Intro to R, Intro to Python, and the OmniSOC cybersecurity learning experience with Elastic Corporation.

1.4. How have the results been disseminated to communities of interest?

Results have been communicated to communities of interest through published papers in peer-reviewed journals and conference proceedings (listed in the Products section). In addition, Jetstream team members held three in-person events in January and February as well as 11 webinars since May 2020. This is lower than in years past as many events were cancelled outright due to the COVID-19 pandemic.

Date	Event Name	Location	Presenter(s)/Attendees	Notes, Link to slides/materials, etc.
Jan 11-15	Plant and Animal Genome Conference (PAG XXVIII)	San Diego, CA	Bhavya Papudeshi, Tom Doak, Sheri Sanders, Carrie Ganote	Genomics and other Science in the NSF-Funded Jetstream Cloud - https://plan.core-apps.com/pag_2020/abstract/0150dda9-3bb4-469b-b84b-4d9faa5750e7 Mining Microbial Genomes from Datasets on the Sequence Read Archive - https://plan.core-apps.com/pag_2020/abstract/60658e794c2a6d06f3aa20dcfb5cb51a Multiple posters
Feb 24	Jetstream Introduction - Indiana University - CSCI-B649	Bloomington, IN	Jeremy Fischer	https://www.jetstream-cloud.org/files/Jetstream-Outreach-IUB-CSCI-B649-Feb2020.pdf
Feb 27	Jetstream for IU Researchers	Bloomington, IN	Sanjana Sudarshan	http://new.emerging-researchers.org
May 20-22	RMACC HybridCloud Event	Webinar	Jeremy Fischer	https://www.jetstream-cloud.org/files/Jetstream-Outreach-RMACC-HPC-Symposium-May2020.pdf
July 9	SGCI Coding Institute	Webinar	Jeremy Fischer	https://www.jetstream-cloud.org/files/Jetstream-API-Tutorial-SGCI-Workshop-July2020.pdf

July 10	SGCI Coding Institute	Webinar	Jeremy Fischer, Sanjana Sudarshan, Mike Lowe, Steve Bird, Eric Coulter	Slides here: https://www.jetstream-cloud.org/research/publications.php (Multiple slide decks)
July 26-30	PEARC20	Webinar	Jeremy Fischer, Sanjana Sudarshan, Mike Lowe, Steve Bird, Eric Coulter	https://github.com/XSEDE/Container_Tutorial
Aug 27	RT Infoshare	Webinar	Jeremy Fischer, David Hancock	Slides at https://www.jetstream-cloud.org/research/publications.php
Sep 3	SC20 Student talk	Webinar	Sanjana Sudarshan	https://www.jetstream-cloud.org/files/SC_student_talk_Sep_3.pdf
Sep 4	Big Data Neuroscience Workshop - Lightning Talk	Webinar	Jeremy Fischer	https://www.jetstream-cloud.org/files/Jetstream-ACNN-Lightning-Talk-Sep2020.pdf
Oct 3	CCSC - NW	Webinar	Sanjana Sudarshan	https://www.jetstream-cloud.org/files/CCSC-NW-Oct-3-2020.pdf
Oct 15	Gateways 2020	Webinar	Jeremy Fischer, Sanjana Sudarshan, Mike Lowe, Steve Bird, Eric Coulter	https://github.com/XSEDE/Container_Tutorial
Oct 16	CASC Hybrid Cloud Panel	Webinar	Jeremy Fischer	https://www.jetstream-cloud.org/files/Fischer_CASC_HybridCloudPanel_Oct_2020.pdf
Nov 13	Unidata User Committee	Webinar	Jeremy Fischer	https://www.jetstream-cloud.org/files/Fischer_UnidataUserCommittee_Nov_2020.pdf

Meetings with research groups interested in being future users of Jetstream:

- Jeremy Fischer met with Carlos Maltzahn and Kevin Tyle to discuss a follow-on project to Big Weather Web and appropriateness for Jetstream.
- Jeremy Fischer met with Dr. Hetland, Dr. Thyng, and Dr. Perez at Texas A&M concerning their research and education needs. Ultimately resulted in an allocation (TG-OCE200001).
- Multiple Jetstream team members met with Brian Wheeler, Sr Systems Engineer with IU Libraries in regards to using GPU resources on Jetstream for several projects.
- Utah Valley University - George Rudolph - Computer Science education

- Indiana University School of Medicine - Mark Frank - AI for radiology research. Ultimately resulted in an allocation.
- Texas A&M - Robert Hetland - education/remote learning for summer REU using container orchestration (Kubernetes)
- Protein Data Bank - John Westbrook - moving PDB101 into production on Jetstream
- Center for Applied Cybersecurity Research - Mark Krenz - virtual security camp for middle schoolers
- Los Alamos National Laboratory - Mark Flynn - using Jetstream for COVID-19 research
- Bethune-Cookman University - Raphael Isokpehi - remote education environment for summer and fall 2020
- Met with multiple researchers on the eBird/citizen science biodiversity project about using Jetstream resources
- Met with Dr. Mark Holland from IUPUI as well as multiple collaborators on HERA-MI Projects using Jetstream
- Worked with Chris Layton of Oak Ridge National Laboratory in regards to having ORNL users prototype workflows on Jetstream that exceed the capabilities of their on-premises cloud. Multiple projects are in development presently, including one using 1000 cores per run via elastic computing techniques.
- Consulted with XSEDE XCI and George Mason University on their on-premises cloud - policies and practices
- Met with Brian Stuck of University of Florida about hosting the Biocode Gateway
- Met with the Open Science Grid about with with Jetstream for on-demand resources

Courses and workshops utilizing Jetstream resources:

- E631: Quantitative Research in Music Education - Indiana University
- E632: Advanced Quantitative Research in Music Education - Indiana University
- XSEDE Workshop at Clark Atlanta University/Spelman College
- Deep Dive into Constructing Containers for Scientific Computing and Gateways at the SGCI Coding Institute
- Deep Dive into Constructing Containers for Scientific Computing and Gateways at PEARC20
- Deep Dive into Constructing Containers for Scientific Computing and Gateways at Gateways 2020
- CSCI 455: Principles of Database Systems - University of Puget Sound
- Advanced Computing for Social Change (ACSC) and Computing for Change (C4C) Workshops at PEARC2020
- Advanced Computing for Social Change (ACSC) and Computing for Change (C4C) Workshops at SC20
- XSEDE Faculty Development Workshop at Atlanta University Center Consortium - Data Science Initiative
- CS 440: Capstone in Computer Science - University of Puget Sound
- ACSC Faculty Curriculum Development Workshop - affiliated with PEARC2020

- BMI5007: Methods in Health Data Science - University of Texas Health Science Center at Houston
- BIO 316: Introduction to Computational Biology
- Neurepiomics Course at The Biggs Institute at UT Health San Antonio
- NCGAS R for Biologists virtual workshop (summer 2020)
- UC Davis GGG 201b: lab section
- Metagenomics Discovery Challenge – CTF 125 - CUNY Kingsborough Community College
- BIOL 360: Genomic Analysis - Albion College
- Engaging Students and Researchers in High Performance Computing R&D Projects - UT San Antonio
- High Performance Computing with Accelerators course - hosted by Argonne Laboratory
- Python Data Analytics Training Series - hosted by University of Iowa
- CSCI-B 649: Topics in Systems: Science Gateway Architectures - Indiana University
- B547/I533: Systems & Protocol Security & Information Assurance - Indiana University
- SDSC HPC Students: A program for educating and training the next generation of HPC professionals and researchers.
- CS504 Data Analytics Engineering (DAEN) - George Mason University
- Advanced Biostatistics Seminar: Big Data Analysis with Python - University of Iowa
- Software Engineering I: Fall 2019: CS 4320/7320 - University of Missouri
- S 490 / CS 491 Capstone - Arcadia University
- I535: Management, Access, and Use of Big and Complex Data - Indiana University (Spring 2020 and Fall 2019)
- JupyterHub for data science and neuroscience teaching platform - University of Missouri, Columbia
- NFO B513: The Design, Implementation, and Evaluation of Electronic Health Record Systems - IUPUI
- INFO B535 / NURS I631: Clinical Information Systems - IUPUI
- INFO B585: Biomedical Analytics - IUPUI
- Jetstream REU program - Summer 2020
- CPSC 474 Class - Advanced Architecture and Parallel Computing - Slippery Rock University
- CS4663 Course - Distributed and Cloud Security - UT San Antonio
- Mathematical modeling of chemical engineering processes using JupyterHub - Carnegie Mellon University
- Reproducible & Collaborative Data Science - University of California, Berkeley
- Applied Metagenomics - University of Detroit, Mercy
- Workshop in Molecular Evolution at the Marine Biological Laboratory at Woods Hole, MA (May and June 2020)
- JPL HBCU Summer Internship Project - NASA Ames Research Center
- NASA 2020 JPL-SAGAn Bioinformatics Workshop - Jet Propulsion Laboratory in Pasadena, CA, USA
- Math-592/492 Methods of Applied Mathematics - University of South Dakota

- XSEDE ICEBERG Users Workshop - Rutgers
- I520/B544 Security for Networked Systems - Indiana University
- BE429/BE529/ISTA429/INFO529/PLS429/PLS529 - Applied Cyberinfrastructure Concepts - University of Arizona
- Geospatial, Temporal, and Network Data Science - SUNY College at Oswego
- IS 6733 Big Data Technology - University of Texas San Antonio
- MSCI: 6040 Data Programming in Python - University of Iowa
- CHARMM-GUI Lecture Series on Molecular Modeling and Simulation - Lehigh University
- Biology 492: Genomics - SUNY College at Oswego
- Basic Bioinformatics for Genomics Research - Nova Southeastern University
- Data Intensive Biology Summer Institute for Sequence Analysis at UC Davis
- Gen 711/811 Genomics and Bioinformatics - University of New Hampshire
- MBL Physiology Boot Camp: Bioinformatics - Woods Hole Oceanographic Institute
- BTEC 395: Translational Bioinformatics - University of Maryland, Baltimore County
- Introduction to Bioinformatics Fall 2020: Full Genome, Transcriptome, Metabolome Analysis - Nova Southeastern University
- BIOL597/697GE - Evolutionary Genomics & Bioinformatics - University of Massachusetts Amherst
- PPPL-Simons Summer School: “Stellarator Optimization” - Princeton University
- Applied Engineering Design 2-5 (AED 2-5) - Marmion Academy
- Computational Methods for Imaging Science - Washington University, St. Louis
- Scientific Computing for Biologists (Summer course) - University of Arkansas
- Application of Python programming for University Physics (PHYS-191) - Montclair State University
- GIS 5577 – Spatial Databases and Administration - University of Minnesota
- Big Data Analytics CS594 (007) / CS690 (001) - University of Tennessee, Knoxville
- ITEC 4220: Advanced Data Analytics - Georgia Gwinnett College
- BE429/BE529/ISTA429/INFO529/PLS429/PLS529 - Applied Cyberinfrastructure Concepts - University of Arizona
- GEOG 5560: Application Development with Python Programming - University of North Texas
- CST 210: Fundamentals of Computational Science - Doane University
- Introducing Computing and Technology Through Problem-Solving and Discrete Math - UC San Diego Math Project (for high school students)
- Science Gateways Community Institute - Summer Coding Institute - SGCI/TACC/IU
- XCRI Workshop at University of Central Florida - XSEDE/IU

1.5. What do you plan to do during the next reporting period to accomplish the goals?

Technical highlights of this work will include:

- Deploying the latest maintenance and feature release to the Jetstream portal, Atmosphere.
- Continued application of security updates (in both featured images and Jetstream hosts) to maintain stability for PY6.
- Continued backporting of any critical feature requests from the user community for OpenStack projects that are needed beyond the Rocky release.
- Continued resolution of data consistency between the TACC Accounting System and the XSEDE central database.
- Investigation of the new accounting system in early testing within XSEDE. The RESTful API features should be a better match for future reporting due to the large number of records generated by Jetstream.

Administrative aspects of this work will include:

- Reviewing Jetstream annual survey results.
- Continuing collaborations with Jetstream researchers to provide letters for proposals to funding agencies.
- Scheduling and preparing for the Jetstream annual operations review.
- Auditing financial spending for PY6 to ensure partners adhere to plans.

User service and access elements of this work will include:

- Developing the annual Jetstream user survey.
- Continuing development of Cornell Virtual Workshops for Jetstream.
- Updating user documentation.
- Continuing to present at various conferences, workshops, etc.

Education, outreach, and training aspects of this work will include:

- Virtual outreach events at University of Cincinnati for users at that organization as well as potentially for other campuses in the area.
- Ongoing webinars for University of Texas at San Antonio for their computer science, engineering, and other interested areas.
- Further discussions with institutions that were forced to cancel scheduled events due to COVID-19 in regards to scheduling in-person events as virtual outreach events.

2. Products (resulting from this project during the specified reporting period)

Journal Articles and Juried Conference Papers

Jetstream project team generated (in bold). User-generated in plain text.

1. Skluzacek, T. J., Chard, R., Wong, R., Li, Z., Babuji, Y. N., Ward, L., ... & Foster, I. (2019, December). **Serverless Workflows for Indexing Large Scientific Data**. In *Proceedings of the 5th International Workshop on Serverless Computing*(pp. 43-48).
2. Skluzacek, T. J. (2019, December). Dredging a data lake: decentralized metadata extraction. In *Proceedings of the 20th International Middleware Conference Doctoral Symposium*(pp. 51-53).
3. Phillips, T., Xiaoyuan, Y., Haakeson, B., & Zou, X. (2019, December). Design and implementation of privacy-preserving, flexible and scalable role-based hierarchical access control. In *2019 First IEEE International Conference on Trust, Privacy, and Security in Intelligent Systems and Applications (TPS-ISA)* (pp. 46-55). IEEE.
4. **Wu, C., Summer, T., Li, Z., Woodard, A., Chard, R., Baughman, M., ... & Foster, I. (2019, December). ParaOpt: Automated Application Parameterization and Optimization for the Cloud**. In *2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)* (pp. 255-262). IEEE.
5. Redondo, C., Arora, R., & Ba, T.N. (2019, December). Theoretical and practical approaches for teaching parallel code correctness. In *2019 26th International Conference on High Performance Computing, Data and Analytics Workshop (HiPCW)* (pp. 48-55). IEEE.
6. Wright, A. M., Schwartz, R. S., Oaks, J. R., Newman, C. E., & Flanagan, S. P. (2020, March). The why, when, and how of computing in biology classrooms. *F1000Research*, 8(1854), 1854.
7. Raychel M. Benson, Edward Munsell, Nicholas Bertrand, Michael Baynton, Evan F. Bollig, and Jeffrey McDonald. 2019. A Multi-Environment HPC-Scale Puppet Infrastructure for Compliance and Systems Automation. In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning) (PEARC '19)*. Association for Computing Machinery, New York, NY, USA, Article 29, 1–8. DOI:<https://doi.org/10.1145/3332186.3332240>
8. Duplyakin, D., Ricci, R., Maricq, A., Wong, G., Duerig, J., Eide, E., ... & Akella, A. (2019). The design and operation of CloudLab. In *2019 {USENIX} Annual Technical Conference ({USENIX}{ATC} 19)* (pp. 1-14).
9. Edwards, R. A., Vega, A. A., Norman, H. M., Ohaeri, M., Levi, K., Dinsdale, E. A., ... & Bibby, K. (2019). Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nature microbiology*, 4(10), 1727-1736.
10. Cai, G., Bossé, Xiao, F., Kheradmand, F., & Amos, C.I. 2020. Tobacco Smoking Increases the Lung Gene Expression of ACE2, the Receptor of SARS-CoV-2. *Am J Respir Crit Care Med*. 2020 Apr 24. PMID: 32329629.
11. Cai, G., Cui, X., Zhu, X., Zhou, J. A Hint on the COVID-19 Risk: Population Disparities in Gene Expression of Three Receptors of SARS-CoV. Preprints. 2020 Feb 28. DOI 10.20944/preprints202002.0408.v1.

12. Perri, M. J. (2020). Online Data Generation in Quantitative Analysis: Excel Spreadsheets and an Online HPLC Simulator Using a Jupyter Notebook on the Chem Compute Web site. *Journal of Chemical Education*. [acs.jchemed.0c00565](https://doi.org/10.1021/acs.jchemed.0c00565)
13. Choi, I. K., Abeyasinghe, E., Coulter, E., Marru, S., Pierce, M., & Liu, X. (2020). TopPIC Gateway: A Web Gateway for Top-Down Mass Spectrometry Data Interpretation. In *Practice and Experience in Advanced Research Computing* (pp. 461-464). <https://doi.org/10.1145/3311790.3400853>
14. Brown, C. T., Moritz, D., O'Brien, M. P., Reidl, F., Reiter, T., & Sullivan, B. D. (2020). Exploring neighborhoods in large metagenome assembly graphs using spacegraphcats reveals hidden sequence diversity. *Genome biology*, 21(1), 1-16. <https://link.springer.com/article/10.1186/s13059-020-02066-4>
15. Keahey, K., Anderson, J., Zhen, Z., Riteau, P., Ruth, P., Stanzione, D., ... & Mambretti, J. (2020). Lessons Learned from the Chameleon Testbed. In *2020 {USENIX} Annual Technical Conference ({USENIX}{ATC} 20)* (pp. 219-233).
16. **Hong, Z., Tchoua, R., Chard, K., & Foster, I. (2020, June). SciNER: Extracting Named Entities from Scientific Literature. In *International Conference on Computational Science* (pp. 308-321). Springer, Cham.**
17. **Jalili, V., Afgan, E., Gu, Q., Clements, D., Blankenberg, D., Goecks, J., ... & Nekrutenko, A. (2020). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Research*.**
18. Chacón, H. D., Kesici, E., & Najafirad, P. (2020). Improving Financial Time Series Prediction Accuracy Using Ensemble Empirical Mode Decomposition and Recurrent Neural Networks. *IEEE Access*, 8, 117133-117145.
19. Klesse, S., DeRose, R. J., Babst, F., Black, B., Anderegg, L. D. L., Axelson, J., ... & Harvey, J. E. Continental-scale tree-ring based projection of Douglas-fir growth—Testing the limits of space-for-time substitution. *Global Change Biology*.
20. Zhu, R., Ding, C., & Huang, Y. (2020). Practical MPC+ FHE with Applications in Secure Multi-Party Neural Network Evaluation. *IACR Cryptol. ePrint Arch.*, 2020, 550.
21. Hubbard, A., Bomhoff, M., & Schmidt, C. J. (2020). fRNAkenseq: a fully powered-by-CyVerse cloud integrated RNA-sequencing analysis tool. *PeerJ*, 8, e8592.
22. Brown, D. W., Ford, V., & Ghafoor, S. K. (2020, May). A Framework for the Evaluation of Parallel and Distributed Computing Educational Resources. In *2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 261-268). IEEE.
23. Salam, S., Brandt, P., D'Orazio, V., Holmes, J., Osorio, J., & Khan, L. (2020). An Online Structured Political Event Dataset based on CAMEO Ontology. [Vrt4a.htm](https://doi.org/10.1111/2041-210X.13440)
24. Norman, K., Chamberlain, S., Boettiger, C. (2020). Taxadb: A high-performance local taxonomic database interface. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.13440 software
25. Boettiger, C. (2020). Ecological management of stochastic systems with long transients. *Theoretical Ecology*. doi:10.1007/s12080-020-00477-4, code
26. Martin, B.T., Douglas, M.R., Chafin, T.K., Placyk Jr., J.S., Birkhead, R.D., Phillips, C.A., & Douglas, M.E. (2020). Contrasting signatures of introgression in North American box

- turtle (*Terrapene* spp.) contact zones. *Molecular Ecology*. (pp. 1-17). DOI: 10.1111/mec.15622
27. Zhou, J., Smith, K., Wilsbacher, G., Sagona, P., Reddy, D., & Torkian, B. (2020). Building Science Gateways for Humanities. PEARC20. July 2020. (pp. 327-332). <https://doi.org/10.1145/3311790.3396628>
 28. Nine, MD, Zulkar, SQ, & Kosar, T. "A Two-Phase Dynamic Throughput Optimization Model for Big Data Transfers." *IEEE Transactions on Parallel and Distributed Systems* 32.2 (2020): 269-280.
 29. Spoor, Shawna, Connor Wytko, Brian Soto, Ming Chen, Abdullah Almsaeed, Bradford Condon, Nic Herndon et al. "TriPal and Galaxy: supporting reproducible scientific workflows for community biological databases." Database 2020 (2020). DOI: <https://doi.org/10.1093/database/baaa032>
 30. Arms, S., J. Chastang, M. Grover, J. Thielen, M. Wilson, and D. Dirks, 2020: Introducing Students to Scientific Python for Atmospheric Science. Bull. Amer. Meteor. Soc., 101, E1492–E1496, <https://doi.org/10.1175/BAMS-D-20-0069.1>.
 31. **Hong Z., Tchoua R., Chard K., Foster I. (2020) SciNER: Extracting Named Entities from Scientific Literature. In: Krzhizhanovskaya V. et al. (eds) Computational Science – ICCS 2020. ICCS 2020. Lecture Notes in Computer Science, vol 12138. Springer, Cham. https://doi.org/10.1007/978-3-030-50417-5_23**
 32. X. Yu, B. Haakenson, T. Phillips and X. Zou, "User-Friendly Design of Cryptographically-Enforced Hierarchical Role-based Access Control Models," 2020 29th International Conference on Computer Communications and Networks (ICCCN), Honolulu, HI, USA, 2020, pp. 1-9, doi: 10.1109/ICCCN49398.2020.9209656.
 33. Clements, Timothy, and Marine A. Denolle. "SeisNoise. jl: Ambient Seismic Noise Cross Correlation on the CPU and GPU in Julia." *Seismological Research Letters* (2020). DOI: <https://doi.org/10.1785/0220200192>
 34. Lwowski, Brandon, and Peyman Najafirad. "COVID-19 Surveillance through Twitter using Self-Supervised and Few Shot Learning." In Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020.
 35. Tabataba-Vakili, F., J. H. Rogers, G. Eichstädt, G. S. Orton, C. J. Hansen, T. W. Momary, J. A. Sinclair et al. "Long-term tracking of circumpolar cyclones on Jupiter from polar observations with JunoCam." *Icarus* 335 (2020): 113405. DOI: <https://doi.org/10.1016/j.icarus.2019.113405>
 36. **Tchoua, R.B., Ajith, A., Hong, Z., Ward, L.T., Chard, K., Belikov, A., Audus, D.J., Patel, S., de Pablo, J.J., Foster, I.T. Creating training data for scientific named entity recognition with minimal human effort. *Lecture Notes in Computer Science*. Vol. 11536, pages 398-411. DOI: 10.1007/978-3-030-22734-0_29**
 37. Lyons, E., Zink, M., Mandal, A., Wang, C., Ruth, P., Radhakrishnan, C., ... & Rodero, I. (2020, January). DyName: Scalable Weather Workflow Processing in the Academic MultiCloud. In *100th American Meteorological Society Annual Meeting*. AMS.
 38. Thareja, K., Wang, C., Ruth, P., Mandal, A., Baldin, I., & Stealey, M. (2019, October). COMET: Distributed Metadata Service for Multi-cloud Experiments. In *2019 IEEE 27th International Conference on Network Protocols (ICNP)* (pp. 1-2). IEEE.

39. Bullock D, Takemura H, Caiafa CF, Kitchell L, McPherson B, Caron B, Pestilli F (2019) Associative white matter connecting the dorsal and ventral posterior human cortex. *Brain structure & function*, 224(8): 2631–2660.
40. Rossini PM, Di Iorio R, Bentivoglio M, Bertini G, Ferreri F, Gerloff C, Ilmoniemi RJ, Miraglia F, Nitsche MA, Pestilli F, Rosanova M, Shirota Y, Tesoriero C, Ugawa Y, Vecchio F, Ziemann U, Hallett M (2019) Methods for analysis of brain connectivity: An IFCN-sponsored review. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, 130(10):1833–1858.
41. Kumar S, Sreenivasan V, Talukdar P, Pestilli F, Sridharan D (2019) ReAI-LiFE: Accelerating the Discovery of Individualized Brain Connectomes on GPUs. *Proceedings of the... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence*, 33(1):630–638.
42. Kaneko T, Takemura H, Pestilli F, Silva AC, Ye FQ, Leopold DA (2020) Spatial organization of occipital white matter tracts in the common marmoset. *Brain structure & function*, 225(4):1313–1326.
43. Rheault F, De Benedictis A, Daducci A, Maffei C, Tax CMW, Romascano D, Caverzasi E, Morency FC, Corrivetti F, Pestilli F, Girard G, Theaud G, Zemmoura I, Hau J, Glavin K, Jordan KM, Pomiecko K, Chamberland M, Barakovic M, Goyette N, Poulin P, Chenot Q, Panesar SS, Sarubbo S, Petit L, Descoteaux M (2020) Tractostorm: The what, why, and how of tractography dissection reproducibility. *Human brain mapping*, 41(7):1859–1874.
44. Caron B, Bullock D, Kitchell L, McPherson B, Kellar DA, Cheng H, Newman SD, Port NL, Pestilli F (2020) Advanced mapping of the human white matter microstructure better separates elite sports participation. <https://doi.org/10.31234/osf.io/dxaqp>
45. Mejia J, Mejia A, Pestilli F (2019) Open data on industry payments to healthcare providers reveal potential hidden costs to the public. *Nature communications*, 10(1):4314.
46. Puzniak RJ, Ahmadi K, Kaufmann J, Gouws A, Morland AB, Pestilli F, Hoffmann MB (2019) Quantifying nerve decussation abnormalities in the optic chiasm. *NeuroImage. Clinical*, 24:102055.
47. Takemura H, Pestilli F, Weiner KS (2019) Comparative neuroanatomy: Integrating classic and modern methods to understand association fibers connecting dorsal and ventral visual cortex. *Neuroscience research*, 146:1–12.
48. Bertò G, Bullock D, Astolfi P, Hayashi S, Zigiotta L, Annicchiarico L, Corsini F, De Benedictis A, Sarubbo S, Pestilli F, Avesani P, Olivetti E (2020) Classifyber, a robust streamline-based linear classifier for white matter bundle segmentation. *NeuroImage*, 224:117402.
49. Kurzawski JW, Mikellidou K, Morrone MC, Pestilli F (2020) The visual white matter connecting human area prostriata and the thalamus is retinotopically organized. *Brain structure & function*, 225(6):1839–1853.
50. S. Sinkovits R, D. Soto O (2020) Introducing Computing and Technology through Problem-Solving in Discrete Mathematics. *Practice and Experience in Advanced Research Computing*: 429–435.

51. Whitnell RM, Reeves MS (2019) Process Oriented Guided Inquiry Learning Computational Chemistry Experiments: Revisions and Extensions Based on Lessons Learned from Implementation. *Using Computational Methods To Teach Chemical Principles*, 1312:65–77.
52. Zonca A (2020) Deploy Kubernetes and JupyterHub on XSEDE Jetstream. <https://doi.org/10.17605/OSF.IO/ZYHWT>
53. Skorupa Parolin E, Salam S, Khan L, Brandt P, Holmes J (2019) Automated Verbal-Pattern Extraction from Political News Articles using CAMEO Event Coding Ontology. *2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, :258–266.
54. Salam S, Khan L, El-Ghamry A, Brandt P, Holmes J, D’Orazio V, Osorio J (2020) Automatic Event Coding Framework for Spanish Political News Articles. *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, :246–253.
55. Farquharson JI, Amelung F (2020) Author Correction: Extreme rainfall triggered the 2018 rift eruption at Kīlauea Volcano. *Nature*, 582(7811):E3.
56. Shih M-FM, Davis FP, Henry GL, Dubnau J (2019) Nuclear Transcriptomes of the Seven Neuronal Cell Types That Constitute the Drosophila Mushroom Bodies. *G3*, 9(1):81–94.
57. Sumner JT, Wax S, Andrasz C, Anderson P, Keeling E, Davidson J (2020) Hybrid Genome Assembly of a Regenerative Chordate. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 34(S1):1–1.
58. Pryor SC, Shepherd TJ, Volker PJH, Hahmann AN, Barthelmie RJ (2020) “Wind Theft” from Onshore Wind Turbine Arrays: Sensitivity to Wind Farm Parameterization and Resolution. *Journal of Applied Meteorology and Climatology*, 59(1):153–174.
59. Wang Z, Chu T, Choate LA, Danko CG (2019) Identification of regulatory elements from nascent transcription using dREG. *Genome research*, 29(2):293–303.
60. Perri, M.J., Akinmurele, M., Haynie, M. (2019). Chem Compute Science Gateway: An Online Computational Chemistry Tool. *Using Computational Methods To Teach Chemical Principles. ACS Symposium Series*, 1312, p. 79-92. DOI: 10.1021/bk-2019-1312.ch007
61. Barthelmie RJ, Shepherd TJ, Pryor SC (2020) Increasing turbine dimensions: impact on shear and power. *Journal of physics. Conference series*, 1618(6):062024.
62. Barthelmie RJ, Shepherd TJ, Aird JA, Pryor SC (2020) Power and Wind Shear Implications of Large Wind Turbine Scenarios in the US Central Plains. *Energies*, 13(16):4269.
63. Mondejar ME, Haglind F (2020) The potential of halogenated olefins as working fluids for organic Rankine cycle technology. *Journal of molecular liquids*, 310:112971.
64. Pryor SC, Schoof JT (2020) Differential Credibility Assessment for Statistical Downscaling. *Journal of Applied Meteorology and Climatology*, 59(8):1333–1349.
65. Pryor SC, Barthelmie RJ, Bukovsky MS, Leung LR, Sakaguchi K (2020) Climate change impacts on wind power generation. *Nature Reviews Earth & Environment*, <https://doi.org/10.1038/s43017-020-0101-7>

66. Lu Y, Tzeferacos P, Liang E, Follett RK, Gao L, Birkel A, Froula DH, Fu W, Ji H, Lamb D, Li CK, Sio H, Petrasso R, Wei MS (2019) Numerical simulation of magnetized jet creation using a hollow ring of laser beams. *Physics of plasmas*, 26(2):022902.
67. Gao L, Liang E, Lu Y, Follet RK, Sio H, Tzeferacos P, Froula DH, Birkel A, Li CK, Lamb D, Petrasso R, Fu W, Wei M, Ji H (2019) Mega-Gauss Plasma Jet Creation Using a Ring of Laser Beams. *The Astrophysical Journal Letters*, 873(2):L11.
68. Lu Y, Li S, Li H, Flippo KA, Barnak D, Birkel A, Lahmann B, Li C, Rasmus AM, Kelso K, Zylstra A, Liang E, Tzeferacos P, Lamb D (2020) Modeling hydrodynamics, magnetic fields, and synthetic radiographs for high-energy-density plasma flows in shock-shear targets. *Physics of plasmas*, 27(1):012303.
69. Song MJ, Potter BI, Doyle JJ, Coate JE (2020) Gene Balance Predicts Transcriptional Responses Immediately Following Ploidy Change in *Arabidopsis thaliana*. *The Plant Cell*, 32(5):1434–1448. <https://doi.org/10.1105/tpc.19.00832>
70. Coate JE, Schreyer WM, Kum D, Doyle JJ (2020) Robust Cytonuclear Coordination of Transcription in Nascent *Arabidopsis thaliana* Autopolyploids. *Genes*, 11(2)<https://doi.org/10.3390/genes11020134>
71. Arora R, Arnaout R (2020) Private Antibody Repertoires Are Public. *Cold Spring Harbor Laboratory*, :2020.06.18.159699. <https://doi.org/10.1101/2020.06.18.159699>
72. Dierauer JR, Zhu C (2020) Drought in the Twenty-First Century in a Water-Rich Region: Modeling Study of the Wabash River Watershed, USA. *WATER*, 12(1):181.
73. Dumett Torres D, Pamidighantam S, Jain PK (2020) Crystal Symmetry, Strain, and Facet-Dependent Nature of Topological Surface States in Mercury Selenide. *Journal of Physical Chemistry C*, 124(19):10344–10352.
74. Dierauer JR, Zhu C, Gong L, Walsh A, Pamidighantam S, Wang J, Christie M, Abeyasinghe E (2020) FutureWater Indiana: A science gateway for spatio-temporal modeling of water in Wabash basin with a focus on climate change. *Practice and Experience in Advanced Research Computing*, :252–261.
75. Pamidighantam S, Wang J, Christie M, Abeyasinghe E, Ellett K, Kammer R, Yaw S, Middleton R, Hoover B (2020) A Science Gateway for Simulating the Economics of Carbon Sequestration Technologies: SimCCS2.0. *Practice and Experience in Advanced Research Computing*, <https://doi.org/10.1145/3311790.3396654>
76. Schneider BI, Bartschat K, Zatsarinny O, Hamilton KR, Bray I, Scrinzi A, Martin F, Gonzalez Vasquez J, Tennyson J, Gorfinkiel JD, Lucchesse R, Pamidighantam S (2020) Atomic and Molecular Scattering Applications in an Apache Airavata Science Gateway. *Practice and Experience in Advanced Research Computing*, :270–277.
77. Canchi S, Swinton MK, Rissman RA, Fields JA (2020) Transcriptomic analysis of brain tissues identifies a role for CCAAT enhancer binding protein β in HIV-associated neurocognitive disorder. *Journal of Neuroinflammation*, 17(1)<https://doi.org/10.1186/s12974-020-01781-w>
78. Xia L, Collins AGE (2020) Temporal and state abstractions for efficient learning, transfer and composition in humans. *Cold Spring Harbor Laboratory*, :2020.02.20.958587. <https://doi.org/10.1101/2020.02.20.958587>

79. Yang B, Kalyanam R, Willis C, Lambert M, Kirkpatrick C (2019) CHEESE: Cyber Human Ecosystem of Engaged Security Education. *Proceedings of the 20th Annual SIG Conference on Information Technology Education*, :189–190.
80. Bendre N, Ebadi N, Prevost JJ, Najafirad P (2020) Human Action Performance Using Deep Neuro-Fuzzy Recurrent Attention Model. *IEEE Access*, 8:57749–57761.
81. Silva SH, Alaeddini A, Najafirad P (2020) Temporal Graph Traversals Using Reinforcement Learning With Proximal Policy Optimization. *IEEE Access*, 8:63910–63922.
82. Pryor SC, Barthelmie RJ, Shepherd TJ (2020) 20% of US electricity from wind will have limited impacts on system efficiency and regional climate. *Scientific reports*, 10(1):541.
83. Letson F, Shepherd TJ, Barthelmie RJ, Pryor SC (2020) Modelling Hail and Convective storms with WRF for Wind Energy Applications. *Journal of physics. Conference series*, 1452(1):012051.
84. Pryor SC, Shepherd TJ, Bukovsky M, Barthelmie RJ (2020) Assessing the stability of wind resource and operating conditions. *Journal of physics. Conference series*, 1452(1):012084.
85. Pryor SC, Shepherd TJ, Volker PJH, Hahmann AN, Barthelmie RJ (2020) Diagnosing systematic differences in predicted wind turbine array-array interactions. *Journal of physics. Conference series*, 1618(6):062023.
86. Aird JA, Barthelmie RJ, Shepherd TJ, Pryor SC (2020) WRF-Simulated Springtime Low-Level Jets over Iowa: Implications for Wind Energy. *Journal of physics. Conference series*, 1618(6):062020.
87. Letson F, Shepherd TJ, Barthelmie RJ, Pryor SC (2020) WRF Modeling of Deep Convection and Hail for Wind Power Applications. *Journal of Applied Meteorology and Climatology*, 59(10):1717–1733.
88. Morgan, E. L., Abeysinghe, E., Pamidighantam, S., Coulter, E., Marru, S., & Pierce, M. (2019). The Distant Reader: Tool for Reading. In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)* (pp. 1-4).
89. **Doak, T.G. Sanders, S.A. Ganote, C. Papudeshi, B. Fischer, J. Hancock, D.Y. (2020). National Center for Genome Analysis Support (NCGAS): Genomics and other Science in the NSF-Funded Jetstream Cloud. *Plant and Animal Genome*.**
90. Cotta, L., Teixeira, H.C., Swami, A., and Ribeiro, B. (2020). Unsupervised Joint k-node graph representations with compositional energy-based models. *Adv. Neural Inf. Process. Syst.* 33.
91. Doane, MP, Morris, MM, Papudeshi, B, Allen, L, Pande, D, Haggerty, JM, Johri, S, Turnlund, AC, Peterson, M, Kacev, D, Nosal, A, Ramirez, D, Hovel, K, Ledbetter, J, Alker, A, Avalos, J, Baker, K, Bhide, S, Billings, E, Byrum, S, Clemens, M, Demery, AJ, Lima, LFO, Gomez, O, Gutierrez, O, Hinton, S, Kieu, D, Kim, A, Loaiza, R, Martinez, A, McGhee, J, Nguyen, K, Parlan, S, Pham, A, Price-Waldman, R, Edwards, RA, and Dinsdale, EA. (2020). The skin microbiome of elasmobranchs follows phylosymbiosis, but in teleost fishes, the microbiomes converge. *Microbiome*. (8)1: 93.

92. Lima, LFO, Weissman, M, Reed, M, Papudeshi, B, Alker, AT, Morris, MM, Edwards, RA, de Putron, SJ, Waidya, NK. (2020). Modeling of the Coral Microbiome: the influence of temperature and microbial network. *Microbiome*. (11)2.
93. Burnett, KM, Elshall, A, Wada, CA, Arik, A, El-Kadi, Aly, Voss, C, Delevaux, JM, Bremer, LL. (2020). Incorporating historical spring discharge protection into sustainable groundwater management: A case study from Pearl Harbor Aquifer, Hawai`i. *Front. Water*. (2)14.
94. Fornace, ME, Porubsky, NJ, Pierce, NA. (2020). A unified dynamic programming framework for the analysis of interacting nucleic acid strands: Enhanced models, scalability, and speed. *ACS Synth. Biol.* (9)10: 2665-2678.

Pre-prints

95. Nüst D, Eddelbuettel D, Bennett D, Cannoodt R, Clark D, Daroczi G, Edmondson M, Fay C, Hughes E, Lopp S, Others (2020) The Rockerverse: Packages and Applications for Containerization with R. *arXiv preprint arXiv:2001.10641*, <https://arxiv.org/abs/2001.10641>
96. Das A, Rad P (2020) Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey. *arXiv [cs.CV]*, <http://arxiv.org/abs/2006.11371>
97. Silva SH, Najafirad P (2020) Opportunities and Challenges in Deep Learning Adversarial Robustness: A Survey. *arXiv [cs.LG]*, <http://arxiv.org/abs/2007.00753>
98. Bendre N, Marín HT, Najafirad P (2020) Learning from Few Samples: A Survey. *arXiv [cs.CV]*, <http://arxiv.org/abs/2007.15484>
99. Caron, Bradley, Ricardo Stuck, Brent McPherson, Daniel Bullock, Lindsey Kitchell, Joshua Faskowitz, Derek Kellar et al. "Collegiate athlete brain data for white matter mapping and network neuroscience." *bioRxiv* (2020). <https://www.biorxiv.org/content/biorxiv/early/2020/08/26/2020.08.25.267005.full.pdf>
100. Hryckowian, Andrew J., Bryan D. Merrill, Nathan T. Porter, William Van Treuren, Eric J. Nelson, Rebecca A. Garlena, Daniel A. Russell, Eric C. Martens, and Justin L. Sonnenburg. "Bacteroides thetaiotaomicron-infecting bacteriophage isolates inform sequence-based host range prediction." *bioRxiv* (2020). Retrieved from <https://www.biorxiv.org/content/biorxiv/early/2020/03/05/2020.03.04.977157.full.pdf>
101. Cazares, Daniel, Adrian Cazares, Wendy Figueroa, Gabriel Guarneros, Rob Edwards, and Pablo Vinuesa. "Dynamics of infection in a novel group of promiscuous phages and hosts of multiple bacterial genera retrieved from river communities." *bioRxiv* (2020). Retrieved from <https://www.biorxiv.org/content/biorxiv/early/2020/08/10/2020.08.07.242396.full.pdf>
102. **Pistorius, Julian L., Chris Martin, and Sanjana Sudarshan. "Exosphere--Bringing The Cloud Closer." arXiv preprint arXiv:2008.10640 (2020). DOI: <https://arxiv.org/abs/2008.10640>**
103. Das, Arun, Jeffrey Mock, Henry Chacon, Farzan Irani, Edward Golob, and Peyman Najafirad. "Stuttering Speech Disfluency Prediction using Explainable Attribution Vectors of Facial Muscle Movements." *arXiv preprint arXiv:2010.01231* (2020). DOI: <https://arxiv.org/abs/2010.01231>

104. Li, Feng, Dali Wang, Feng Yan, and Fengguang Song. "ElasticBroker: Combining HPC with Cloud to Provide Real-time Insights into Simulations." arXiv preprint arXiv:2010.04828 (2020). DOI: <https://arxiv.org/abs/2010.04828>
105. Venkatesh I, Mehra V, Wang Z, Simpson MT, Eastwood E, Chakraborty A, Beine Z, Gross D, Cabahug M, Olson G, Others (2020) Computational approaches identify novel transcription factor combinations that promote corticospinal axon growth following injury. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/2020.06.12.146159v1.abstract>
106. Caron, B., Bullock, D., Kitchell, L., McPherson, B., Kellar, D., Cheng, H., Newman, S., Port, N., & Pestilli, F. (2020, January). Human white matter microstructure predicts elite sports participation. Preprint. PsyArXiv.
107. Cai, G., and Xiao, F. SCANNER: A Web Resource for Annotation, Visualization and Sharing of Single Cell RNA-seq Data. Preprint. bioRxiv 2020.01.25.919712; doi: <https://doi.org/10.1101/2020.01.25.919712>
108. **Vaillancourt, P. Z., Coulter, J. E., Knepper, R., & Barker, B. (2020). Self-Scaling Clusters and Reproducible Containers to Enable Scientific Computing. *arXiv preprint arXiv:2006.14784*.**
109. Chard, K., Gaffney, N., Hategan, M., Kowalik, K., Ludaescher, B., McPhillips, T., ... & Turk, M. J. (2020). Toward enabling reproducibility for data-intensive research using the Whole Tale platform. *arXiv preprint arXiv:2005.06087*.

Other Conference Presentations / Papers

1. **Sudarshan, S. (2020). Jetstream Introduction. Indiana University, Bloomington, Feb. 2020.**
2. Suderman, K., Ide, N., Marc, V., Cochran, B., & Pustejovsky, J. 2020. AskMe: A LAPPS Grid-based NLP Query and REtrieval System for Covid-19 Literature. 30 June 2020. ACL 2020 Workshop NLP-COVID submission. <https://openreview.net/forum?id=WQwNh1RShX>
3. **Leffler, H., Sanders, S., Papudeshi, B. "Visualizing metagenomic data in R using Jetstream." (2020) <http://hdl.handle.net/2022/25761>**
4. **Foran, E., Underwood, T.A., Snapp-Childs, W., Sanders, S. "Automatic Capture and Classification of Frog Calls." (2020) <http://hdl.handle.net/2022/25760>**
5. **Baker, N., Jarmusch, A., Chandrasekaran, S. "HPC Rankings Based on Real Applications." (2020). <http://hdl.handle.net/2022/25759>**

Books

None to report.

Book Chapters

None to report.

Other Products

Survey Instruments

N/A

Software or netware

1. Autamas <https://github.com/alecbcs/autamus>
2. <https://materialhub.org>
3. <http://osome.iu.edu>
4. <https://geo-gateway.org>
5. <https://cord.distantreader.org/carrels/INDEX.HTM>
6. <https://carrels.distantreader.org>

Dataset

7. Costa C (2020) Dataset: XSEDE Return on Investment Data and Analysis (July 2014 to August 2019). <https://doi.org/10.5967/6tc4-sb46>

Theses / Dissertations

1. Willis, Craig Alan. "Trust, but verify: An investigation of methods of verification and dissemination of computational research artifacts for transparency and reproducibility." PhD diss., University of Illinois at Urbana-Champaign, 2020. DOI: <http://hdl.handle.net/2142/108468>
2. Zhu, R. (2020). New Techniques and Programming Tools for Efficient MPC with Real-world Applications (Doctoral dissertation, Indiana University).

Other publications

1. Wernert, J. Jankowski, H., Hancock, D.Y., and Snapp-Childs, W. Jetstream Annual User Survey - 2020 Summary Report. Available from <http://hdl.handle.net/2022/25954>.

Other (not included in research.gov report)

Other presentations

None to report.

Jetstream tutorial presentations

1. Fischer, J. (February, 2020). Jetstream Introduction: Indiana University - CSCI-B649. Bloomington, IN. <https://www.jetstream-cloud.org/files/Jetstream-Outreach-IUB-CSCI-B649-Feb2020.pdf>
2. Fischer, J. (May, 2020). RMACC HybridCloud Event, Webinar [Online]. <https://www.jetstream-cloud.org/files/Jetstream-Outreach-RMACC-HPC-Symposium-May2020.pdf>
3. Fischer, J. (July, 2020). SGCI Coding Institute, Webinar [Online] <https://www.jetstream-cloud.org/files/Jetstream-API-Tutorial-SGCI-Workshop-July2020.pdf>
4. Fischer, J., Sudarshan, S., Lowe, J.M., Bird, S., Coulter, E. (July, 2020). SGCI Coding Institute, Webinar [Online]. <https://www.jetstream-cloud.org/research/publications.php>
5. Fischer, J., Sudarshan, S., Lowe, J.M., Bird, Sr., Coulter, J.E. (2020). Jetstream container tutorial at PEARC20, Webinar [Online]. https://github.com/XSEDE/Container_Tutorial

6. Fischer, J. & Hancock, D.Y. (2020). Jetstream for Research Technologies Infoshare, Webinar [Online]. <https://www.jetstream-cloud.org/research/publications.php>
7. Sudarshan, S. (2020). Jetstream for SC20 students, Webinar [Online]. https://www.jetstream-cloud.org/files/SC_student_talk_Sep_3.pdf
8. Fischer, J. (2020). Jetstream for Big Data Neuroscience Workshop, Webinar [Online]. <https://www.jetstream-cloud.org/files/Jetstream-ACNN-Lightning-Talk-Sep2020.pdf>
9. Sudarshan, S. (2020). Jetstream for CCSC - NW, Webinar [Online]. <https://www.jetstream-cloud.org/files/CCSC-NW-Oct-3-2020.pdf>
10. Fischer, J., Sudarshan, S., Lowe, J.M., Bird, S., & Coulter, J.E. (2020), Jetstream for Gateways 2020, Webinar [Online]. https://github.com/XSEDE/Container_Tutorial
11. Fischer, J. (2020). Jetstream for CASC Hybrid Cloud Panel, Webinar [Online]. https://www.jetstream-cloud.org/files/Fischer_CASC_HybridCloudPanel_Oct_2020.pdf
12. Fischer, J. (2020). Jetstream for Unidata User Committee, Webinar [Online]. https://www.jetstream-cloud.org/files/Fischer_UnidataUserCommittee_Nov_2020.pdf

3. Participants

3.1. Individuals

First Name	Last Name	Most Senior Project Role	Nearest Person Month Worked	Email (if new to project)	Affiliation	Contribution
David	Hancock	PD/PI	4		IU	PI
Ian	Foster	Co-PD/PI	1		UC	Co-PI
Matthew	Vaughn	Co-PD/PI	1		TACC	Co-PI
Nirav	Merchant	Co-PD/PI	1		UA	Co-PI
James	Taylor	Co-PD/PI	1		JHU	Co-PI
John	Lowe	Other Professional	6		IU	System Administration
Therese	Miller	Other Professional	1		IU	Project management
Enis	Afgan	Other	1		JHU	Developer
Brian	Beck	Other	9		TACC	User Services, Outreach
Stephen	Bird	Other	1		IU	System administration
Loren	Cain	Other	1	lcain@tacc.utexas.edu	TACC	Project management
Tharon	Carlson	Other	1		UA	Developer
Dave	Clements	Other	1		JHU	Outreach
Erik	Ferlanti	Other	1		TACC	Developer
Jeremy	Fischer	Other	11		IU	Outreach
Francois	Halbach	Other	2	fhalbach@tacc.utexas.edu	TACC	System administration
Cody	Hammock	Other	1	hammock@tacc.utexas.edu	TACC	System administration
Lee	Liming	Other	1		UC	Developer

Alex	Mahmoud	Other	1	amahmou4@jhu.edu	JHU	Developer
Calvin	McLean	Other	5		UA	Developer
Susan	Mehring	Other	1		Cornell	CVW Lead
Nathaniel	Mendoza	Other	1		TACC	Information Security Officer
Mike	Packard	Other	1		TACC	System administration
Akhil	Seth	Other	1		TACC	Developer
Edwin	Skidmore	Other	1		UA	Project lead for iPlant/Atmo coordination
Winona	Snapp-Childs	Other	2		IU	Administration
Craig	Stewart	Other	1		IU	Senior Investigator
Sanjana	Sudarshan	Other	12		IU	Outreach
Ben	Trumbore	Other	1		Cornell	CVW Developer
George	Turner	Other	12		IU	Chief Architect
Tenecious	Underwood	Other	2	tciousunderwood@gmail.com	IU	Mentor to REU participants
Peter	Vaillancourt	Other	1		Cornell	CVW Developer
Julie	Wernert	Other	1		IU	Assessment
Eliza	Foran	REU participant	2	egforan@iu.edu		REU participant
Haley	Leffler	REU participant	2	hleffler@iu.edu		REU participant
Aaron	Jarmusch	REU participant	2	jarmusch@udel.edu		REU participant
Nolan	Baker	REU participant	2	nolanb@udel.edu		REU participant
Jonathan	Holzmann	REU participant	2	jeholzmann42@tntech.edu		REU participant

3.2. Partner organizations

Type of Partner organization: Academic Institution

Name: University of Colorado/CIRES

Location: Boulder, CO

Partner's contribution to the project: Collaborative research

More detail on partner and contribution: Brian Johnson (University of Colorado-National Snow and Ice Data Center) has agreed to participate as a research collaborator using the Jetstream system for research in the polar science area.

Type of Partner organization: Academic Institution

Name: University of Texas at San Antonio

Location: San Antonio, TX

Partner's contribution to the project: Collaborative research

More detail on partner and contribution: Paul Rad (University of Texas at San Antonio and formerly of Rackspace) will contribute software expertise for system integration. Jetstream is also used as an integral part of the educational mission and as part of research projects involving the Open Cloud Institute at UTSA for minority undergraduate and graduate students.

Type of partner organization: Academic Institution

Name: University of Arizona

Location: Tucson, AZ

Partner's contribution to the project: Facilities; In-kind support

More detail on partner and contribution: The University of Arizona develops the software product called Atmosphere, which is the primary, initial user interface to Jetstream. It orchestrates cloud resources seamlessly across the two separate OpenStack sites, located at the Indiana University and TACC. Atmosphere was initially conceived and successfully used by CyVerse (formerly called iPlant Collaborative) for their users, and has been adapted to Jetstream. The Atmosphere development team primarily focuses on integrating with the TACC Accounting System (TAS), modifying Atmosphere to the XSEDE allocations model, improving the user experience of the Atmosphere user interface, and providing numerous enhancements and bug fixes related to Jetstream's specific use cases and requirements (another significant activity related to end-user and operational support tasks when issues arise within the Jetstream system).

Type of Partner organization: Academic Institution

Name: Johns Hopkins University

Location: Baltimore, MD

Partner's contribution to the project: Collaborative research; In-kind support

More detail on partner and contribution: Johns Hopkins University has continued to offload a significant portion of the workload from the Galaxy Main gateway onto Jetstream. Static Slurm-based clusters have been deployed at IU and TACC and integrated with other infrastructure available to Galaxy Main. The server then allocates suitable jobs (e.g., long-running tools) to Jetstream. The resource capabilities available on Jetstream are not available from other resources. In addition to the gateway usage of Jetstream, we are continuing to offer a standalone Galaxy virtual machine. We are currently working on the ability to launch such instances easily from an environment familiar to Galaxy users, namely Galaxy CloudLaunch (launch.usegalaxy.org).

Type of Partner organization: Academic Institution

Name: University of Chicago

Location: Chicago, IL

Partner's contribution to the project: In-kind support; Collaborative research

More detail on partner and contribution: UChicago made certain that Jetstream's authentication/login function--which relies on Globus-based XSEDE user authentication--worked continuously and satisfied user needs. The UChicago-based Globus user support team actively monitored XSEDE tickets related to Jetstream logins and was involved in debugging and resolving a handful of specific user requests related to this function.

Type of Partner organization: Academic Institution

Name: University of Texas at Austin, Texas Advanced Computing Center (TACC)

Location: Austin, TX

Partner's contribution to the project: In-kind support; Facilities; Collaborative research

More detail on partner and contribution: TACC supports the project by hosting a portion of the production hardware and by offering personnel to operate and function as tier-2 user support.

Type of Partner organization: Academic Institution

Name: Cornell University

Location: Ithaca, New York

Partner's contribution to the project: In-kind support

More detail on partner and contribution: Cornell has delivered three web-based training modules (Cornell Virtual Workshops). The completed modules cover "Jetstream Allocations" and an "Introduction to Using Jetstream," introducing users to the Jetstream system and to the Atmosphere cloud computing environment and a third module on publishing, archiving, and curating virtual machine images for maximizing reproducibility. One additional module is planned; the topic currently in the works will cover advance Jetstream usage via the application programming interface (API). The modules comprise short video clips, examples, exercises, and quizzes, with full-text discussions incorporating an HPC glossary.

Type of Partner organization: Academic Institution

Name: Penn State University

Location: University Station, Pennsylvania

Partner's contribution to the project: In-kind support

More detail on partner and contribution: Anton Nekrutenko, Penn State University, development of the Galaxy software suite.

Type of Partner organization: Academic Institution

Name: University of Hawaii

Location: Honolulu, Hawaii

Partner's contribution to the project: Collaborative research

More detail on partner and contribution: Gwen Jacobs (University of Hawaii) serves as an exemplar for colleges and universities across the United States in utilizing Jetstream's unique features for faculty and students as well as developing VM images supporting research needs in ocean science. Dr. Jacobs also leads the Jetstream Stakeholder Advisory Board.

Type of Partner organization: Academic Institution

Name: University of North Carolina at Chapel Hill

Location: Chapel Hill, North Carolina

Partner's contribution to the project: Collaborative research

More detail on partner and contribution: Thomas M. Carsey, University of North Carolina at Chapel Hill, The Odum Institute, advises regarding software tools of interest to the social science research community; works to increase awareness of Jetstream to our user community; and serves as a pilot site for the implementation of Jetstream's distributed cloud services.

Type of Partner organization: Academic Institution

Name: Jackson State University

Location: Jackson, Mississippi

Partner's contribution to the project: Collaborative research

More detail on partner and contribution: Jessie Walker (Jackson State University) is an unfunded partner, leveraging the Jetstream project in support of academic and research endeavors.

Type of Partner organization: Commercial vendor

Name: Dell, Inc.

Location: Round Rock, Texas

Partner's contribution to the project: In-kind support

More detail on partner and contribution: Dell manufactured and helped deploy the Jetstream hardware, and now provides ongoing hardware support for Jetstream equipment utilized in the test, development, and production systems.

Type of Partner organization: Commercial vendor

Name: The MathWorks, Inc

Location: Natick, Massachusetts

Partner's contribution to the project: In-kind support

More detail on partner and contribution: MathWorks enables end users of the Jetstream system to run MATLAB™ (and any other MathWorks products the end user is licensed to use). This is accomplished through the “bring your own license” arrangement which MathWorks has piloted with a small number of facilities worldwide. In addition, MathWorks, Inc. allows Indiana University to purchase a set of academic licenses of MathWorks products which can be made available to any academic user of the Jetstream system.

3.3. Have other collaborators or contacts been involved?

No.

4. Impact

4.1. What is the impact on the development of the principal discipline(s) of the project?

Jetstream has had a significant impact on the area of computational science as a first-of-a-kind production cloud funded by the NSF. As such, Jetstream has been a pilot, a pathbreaker, and a tremendous learning experience. The Jetstream project has published lessons learned in

peer-reviewed conference proceedings, which have a key impact on the computational science community. In the current operational year of this project, Jetstream has increased its impact on the community through extensive participation in working groups, workshops, Birds of a Feather (BOF) sessions, and small meetings and discussions that are shaping the community approach to implementation, provision, and support of publicly-funded cloud services for open science.

Examples of significant impact on the field of computational science include:

- Through the current project year hundreds of students took classes in computational sciences and did independent research in computational sciences based on the use of Jetstream. The Jetstream team has impacted 3,715 students (aggregate total through October 2020) through classes and research uses of the system.
- Stewart et. al [2015] – the main paper describing plans for the Jetstream system - has already been cited by 152 other published technical papers. Read statistics from Researchgate.net indicate that this paper and the follow-on 2016 publication have been read 448 times just through that venue.
- Stewart et. al [2016] – the paper describing the implementation of, acceptance of, and early experiences with Jetstream, which received the best technical paper award at the XSEDE'16 conference, was invited to submit to a special issue journal that includes the notable papers from the PEARC '17 conference. The follow up paper was published in the Wiley journal *Concurrency and Computation: Practice and Experience* in September 2018.
- The significant number of products from the Jetstream team and highlights from researchers that have used Jetstream are included to demonstrate the breadth of effort the team has put into outreach and the impact already being made within specific disciplines.
- The number of peer-reviewed conference papers and journal publications from Jetstream users and their affiliates has increased to 82 for this reporting period. Tracking such data is difficult for many projects and something XSEDE is interested in improving. Jetstream relies on self-reporting or self-citation by the users with aggregation provided by Google Scholar, Scopus, and ResearchGate.

4.2. What is the impact on other disciplines?

The impact on other disciplines includes:

- Jetstream has been used by over 1000 students (this year) through educational or startup allocations.
- Multiple students have completed/made considerable progress on their dissertations.
- Jetstream is being used as a case study for IU and XSEDE Return on Investment (ROI) analysis. A paper on this subject was presented at the IEEE Utility & Cloud Computing 2018 conference in December 2018, and was nominated for the conference best paper award. Further analysis will be performed as the system moves into the final year of operations.
- Jetstream has active allocations in 76 disciplines from 193 institutions.

- As of November 18, 2020, Jetstream has been cited 152 times and users have acknowledged Jetstream (and their XSEDE allocations) in additional publications. For reference, Bridges (a larger XSEDE resource which was funded by the NSF at the same time as Jetstream) has been cited 121 times. Examples of how Jetstream is being used in different disciplines are below:
 - Astronomy: As the recent imaging of the event horizon of the M87 black hole has demonstrated, astronomy has a complex set of computational requirements for processing and analyzing data in addition to its traditional role in collecting observational data. This significant scientific accomplishment was enabled by the Event Horizon Telescope (a telescope array consisting of a global network of radio telescopes), a large number of scientists, NASA spacecraft, and a variety of computing resources with an increasing portion being done on cloud resources. For example, for the M87 black hole image, two critical steps were done in the cloud and were piloted on Jetstream: 1) correcting for anomalies, so that further image processing could occur, and 2) large survey study of how image reconstruction algorithms affect the final images. In addition, the team is also developing new methods to correlate data from multiple telescopes (to reduce data from petabyte scale to terabyte scale) in the cloud, thereby reducing cost and time.
 - Biophysics: Nanopore sequencing is becoming an increasingly popular method for RNA/DNA sequencing. The basic idea is that an ionic current can be passed through nanopores (or nano-scale “holes”), and one can measure the changes in current as biological molecules pass through the nanopore or near it; changes in current are then used to identify molecules of interest. This method provides inexpensive sample preparation, but the techniques have reported error rates from 2-15%. Aleksei Aksimentiev (University of Illinois, Urbana-Champaign) and collaborators are using computer-based simulations to offer a fuller picture of nanopore technology. With the aid of Jetstream’s on-demand, cloud-based system, the team has developed a robust, inexpensive model called the [steric exclusion model \(SEM\) of nanopore conductance](#). This method is far more efficient than current methods, but is still sensitive enough to account for the atomic structure of both the nanopore and the analyte involved in the sequencing process.
 - Genomics: The Sequence Read Archive (SRA) provides a platform to make DNA and RNA sequences openly available. This resource is already more than 2PB in size, and is still growing. For the field of genomics, working with a repository of this size presents a novel and difficult problem. Rob Edwards (San Diego State University) has created [a new platform](#) that allows researchers to search the SRA with ease using Jetstream and Wrangler to host the search capabilities. The project effectively eliminates the learning curve associated with searching the SRA repository by allowing scientists to search for genes, bacteria, viruses, etc. Then, within a few short hours, the SRA provides a summary of similar sequences and other environments in which the searched entity has been found.

- Molecular Biology and Evolution: Users of Jetstream are also making important contributions to our understanding of population dynamics. Ariella L. Gladstein and Michael F. Hammer published a [recent paper](#) studying population structure and size changes for Ashkenazi Jews to examine the effects these processes have had on the distribution and frequency of disease alleles. Ashkenazi Jews have a high prevalence of known disease-associated mutations which are attributed to “founder effects” resulting from population bottlenecks. The major findings of this study support a view of different population growth rates in Western versus Eastern Ashkenazi Jews that likely have varying effects on the frequency of deleterious mutations in these subpopulations. The results of this important study give historians and medical scientists insight into disease history and potential insight into future population dynamics. This work was made possible by resources offered through the Open Science Grid (Jetstream, Comet, Bridges).
- Political Science: Journalists, through their writing and reporting, share information about global events to which many of us would not otherwise have access. They send missives directly from event sites, recording what’s happening during protests, summits, speeches, and violent actions. For political scientists, these articles offer a largely untapped mine of data about these events. The [Temporally Extended, Regular, Reproducible International Event Records](#) (TERRIER) project aims to be a catalyst for transformation in the methods available for political scientists and policy experts. TERRIER extracts event data from roughly 300 million news articles and puts it into a form usable by researchers. The TERRIER team (University of Oklahoma, MIT) built an open source tool to write scripts, and run Natural Language Processing (NLP) packages on the raw text in order to annotate the sentences with grammatical information and coupled this with the events’ geolocations. This project is built on Jetstream infrastructure, which gives the project the storage and structure to launch the pipeline and process the many available documents.

4.3. What is the impact on the development of human resources?

Jetstream has been used directly and indirectly by thousands of people to date, thus adding significantly to the development of these individuals as computationally-oriented scientists.

Jetstream is having significant impact as a resource in the development of human resources, and Jetstream dissemination efforts play an important role in providing opportunities for self-directed learning. Key points include the following:

- Jetstream has been used directly and indirectly by more than 28,000 people to date through specific allocations and jobs launched via science gateways.
- More than 2,000 people have attended a seminar, tutorial, or talk about Jetstream.

- Hundreds of people have now read peer-reviewed publications about or involving Jetstream, thus adding significantly to the development of these individuals as computationally-oriented scientists.
- Every XSEDE staff member was given a short term startup account on Jetstream, providing specific professional development activities for cyberinfrastructure professionals already working in this field with at least some part of their salary funded by the NSF award to operate the eXtreme Science and Engineering Discovery Environment (XSEDE).
- Jetstream provides trial allocations to allow any user to test-drive the system while considering their longer-term CI needs. This has significantly lowered the barrier to entry for users new to XSEDE.

4.4. What is the impact on teaching and educational experiences?

Of the resources available through the XSEDE program, Jetstream is far and away the most heavily used system for educational purposes. In CY2019, Jetstream provided six times more service units for educational allocations than any other XSEDE resource. In the current project year, Jetstream supported 43 Educational allocations requesting 14,055,920 SUs, and a total of 1,046 students used Jetstream. The courses and workshops utilizing Jetstream continue to grow as listed in Section 1.4. These efforts are inclusive of small colleges, universities, and well-established research centers. In addition, when surveyed, 73% of users indicate the resource is “very important” or “essential” to their education activities, with a mean importance rating of 4.00/5.0.

4.5. What is the impact on physical resources that form infrastructure?

Nothing to report.

4.6. What is the impact on institutional resources that form infrastructure?

Nothing to report.

4.7. What is the impact on information resources that form infrastructure?

Jetstream has proved to be a viable mechanism for large research projects to host and leverage as a service provider. Projects such as IRIS and UNAVCO currently leverage multiple physical

data centers in different regions for service availability. Jetstream provides the opportunity for those types of projects to attain external out-of-region services. The services Jetstream provides indicate that, through expansion, the system could provide widespread cloud service hosting that could replace multiple discrete data centers for domain specific projects within the NSF.

4.8. What is the impact on technology transfer?

There has been an important transfer of information from the Jetstream project to major open source technology projects, particularly OpenStack. The configuration files and methods used to deploy Jetstream are also publically available and have been shared for community adoption. There have been no formal invention disclosures, but Jetstream staff and research faculty play an important role in the development of OpenStack. Members of the Jetstream team are also leading collaboration efforts in the scientific community through the OpenStack Scientific Working Group. OpenStack itself is continually transferred from its open source repository to anyone in the US that wishes to download, install, and use it.

4.9. What is the impact on society beyond science and technology?

Through its impact on the OpenStack project, the team plays an important role in the ongoing improvement of the most widely used open source cloud environment for small business and industry. This aids the development of the cloud-based economy in the US generally. In addition, through REU participants, the team is able to provide high-impact workforce development opportunities to students across the nation, as students who participate will take the skills and knowledge back with them to their home institutions and will be able to share and positively impact others at their home institutions.

4.10. What percentage of the award's budget was spent in a foreign country?

0%

5. Changes/ Problems

5.1. Changes in approach and reasons for change

Nothing to report.

5.2. Actual or Anticipated problems or delays and actions or plans to resolve them

All planned and unplanned outages for the annual reporting period are documented below.

Entity	Start	End	Duration	Planned(P) or Unplanned (U)	Service, System, or Components	User Impact
Atmo	2020.10 .27 11:00E DT	2020.1 0.27 16:43E DT	4.00hr (5h72m)	Planned	Atmosphere upgrade	Atmosphere users will not be able to control virtual infrastructure; running VMs and API users will not be affected.
TACC	2020.09 .13 11:00E DT	2020.0 9.13 12:45E DT	1.75hr (1h45m)	Planned	TACC core network	Inability to control or access instances on TACC cloud; running instances were not affected
TACC	2020.09 .09 14:14E DT	2020.0 9.09 21:02 EDT	6.48hr (6h48m)	Unplanned	Power outage	Due to a power outage at TACC manual intervention was required to recover some number of the compute chassis/hosts.
TACC	2020.08 .11 19:00E DT	2020.0 8.12 15:00E DT	20.00hr (20h00 m)	Unplanned	Power outage	Due to a power sag at TACC manual intervention was required to recover some number of the compute chassis/hosts.
TACC	2020.07 .19 11:00E DT	2020.0 7.19 14:45E DT	3.75hr (3h45m)	Planned	TACC core network	Inability to control or access instances on TACC cloud; running instances were not affected

TACC	2020.06 .20 08:00E DT	2020.0 6.20 17:00 17:00E DT	9.00hr (9h00m)	Unplanned	Power outage	Due to a power sag at TACC manual intervention was required to recover some number of the compute chassis/hosts.
Atmo	2020.06 .16 14:00E DT	2020.0 6.16 14:28E DT	0.47hr (0h28m)	Unplanned	Redeploy of Atmosphere container	Users were unable to control VMs via Atmosphere. Running VMs were not affected.
Atmo	2020.06 .02 11:00E DT	2020.0 6.02 18:42E DT	7.70hr (7h42m)	Planned	Migrate Atmosphere to new server	Running VMs were not affected during the migration. Atmosphere services to change the state of VMs or image VMs will not be available, nor will the web shell or desktop be available, during the migration.
TACC	2020.05 .30 05:00E DT	2020.0 6.02 11:15E DT	44.25hr (44h15 m)	Unplanned	Expiration of AddTrust External CA Root certificate caused communications issues with TACC LDAP	Inability to control or access instances on TACC cloud; running instances were not affected
IU	2020.05 .30 05:00E DT	2020.0 5.30 12:20E DT	7.33hr (7h20m)	Unplanned	Expiration of AddTrust External CA Root certificate	Inability to control or access instances on IU cloud; running instances were not affected

					e caused OpenStack commands to fail	
TACC	2020.05 .10 11:00E DT	2020.0 5.10 15:00E DT	4.00hr (4h00m)	Planned	TACC core network	Inability to control or access instances on TACC cloud; running instances were not affected
Atmo	2020.05 .07 19:20E DT	2020.0 5.07 20:45E DT	1.42hr (1h25m)	Unplanned	Celery	Users would have experienced problems starting VMs via Atmo; API and running VMs were no effected
TACC	2020.03 .25 10:00E DT	2020.0 3.25 16:00 EDT	6.00hr (6h00m)	Planned	Provider network	Inability to control or access instances on TACC cloud; running instances were not affected
TACC	2019.12 .15 9:00 EST	2019.1 2.15 17:00 EST	8.00hr (8h00m)	Planned	Provider network	Inability to control or access instances on TACC cloud; running instances were not affected

5.3. Changes that have significant impact on expenditures

Nothing to report.

5.4. Significant changes in use or care of human subjects

Nothing to report.

5.5. Significant changes in the use or care of vertebrate animals

Nothing to report.

5.6. Significant changes in the use or care of biohazards

Nothing to report.

5.7. Has there been a change in your primary performance site location from the originally proposed? If so, please provide the location of your new primary performance site and the reason for the change in location.

Nothing to report.