



Jetstream Overview: Programmable Cyberinfrastructure: Learning the Jetstream API

Research Bazaar Tucson
May 20, 2019, Tucson, AZ

Jeremy Fischer – Jeremy@iu.edu - Indiana University

Manager, Jetstream Cloud,
UITS Research Technologies

Fischer, J. (2018). Jetstream Overview - Programmable
Cyberinfrastructure: Learning the Jetstream API. Tucson, AZ. Retrieved
from <https://jetstream-cloud.org/research/publications.php>

What is Jetstream and why does it exist?

- NSF's first production cloud facility
- Focus on ease-of-use, broad accessibility
- User-selectable library of preconfigured virtual machines
- Provides **on-demand** *interactive* computing and analysis or persistent services such as gateways
- Enables *configurable* environments; ***programmable cyberinfrastructure***



What Jetstream isn't...

- It's not traditional HPC
- There's no shared filesystem (think cloudy!)
- There's no high-end interconnect fabric (keep thinking cloudy!)
- There aren't GPUs (yet...working on integrating them now!)
- It isn't Amazon, Azure, or GCE (similar, but...)

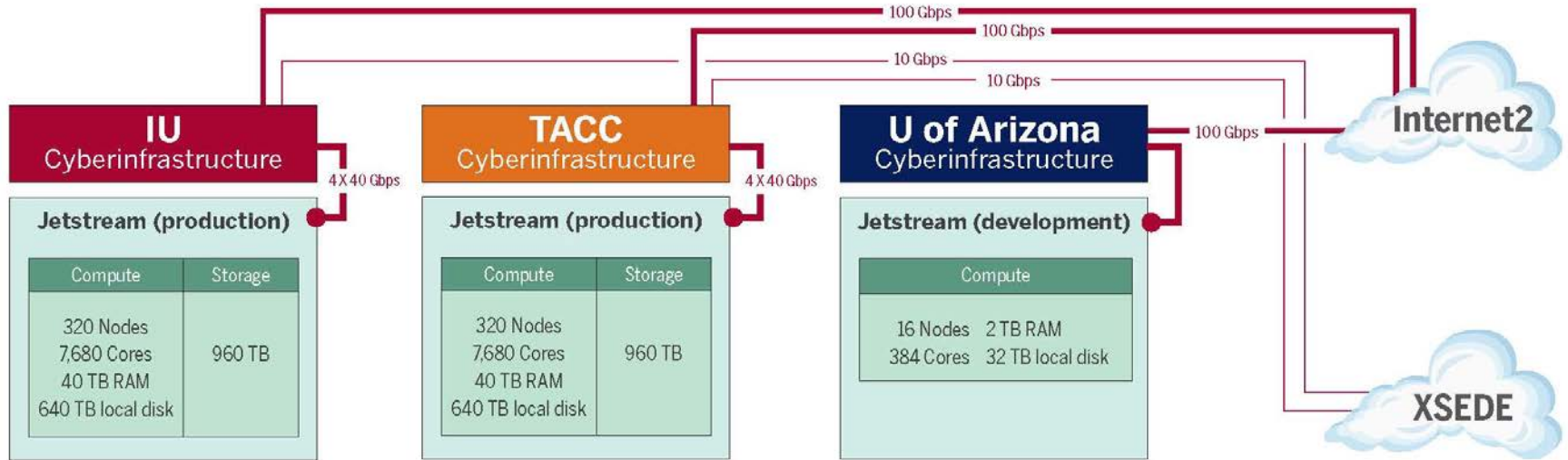


Who uses Jetstream?

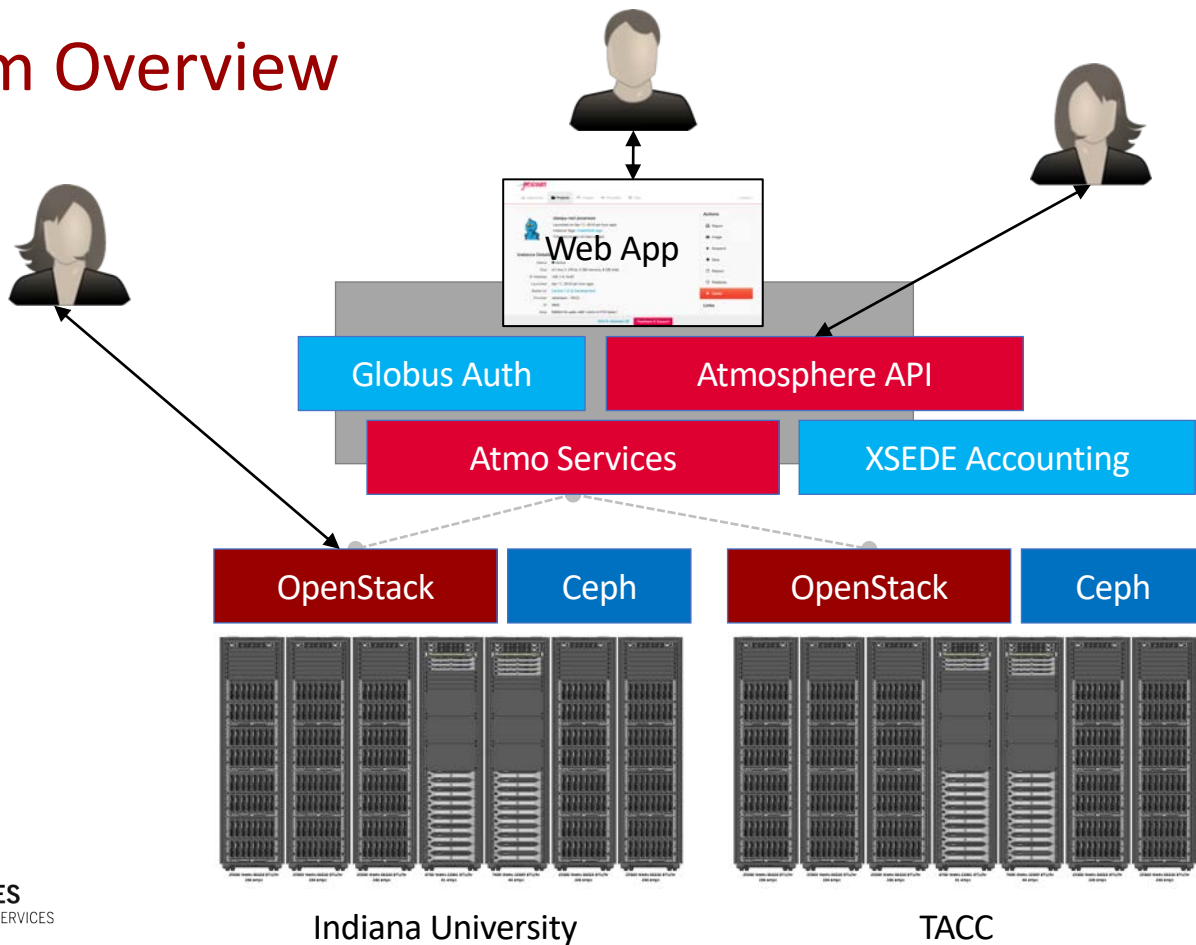
- The researcher needing a handful of cores (1 to 44/vCPU)
- Software creators and researchers needing to create their own VMs and workflows
- Science gateway creators using Jetstream as either the frontend or processor for scientific jobs
- STEM Educators teaching on a variety of subjects



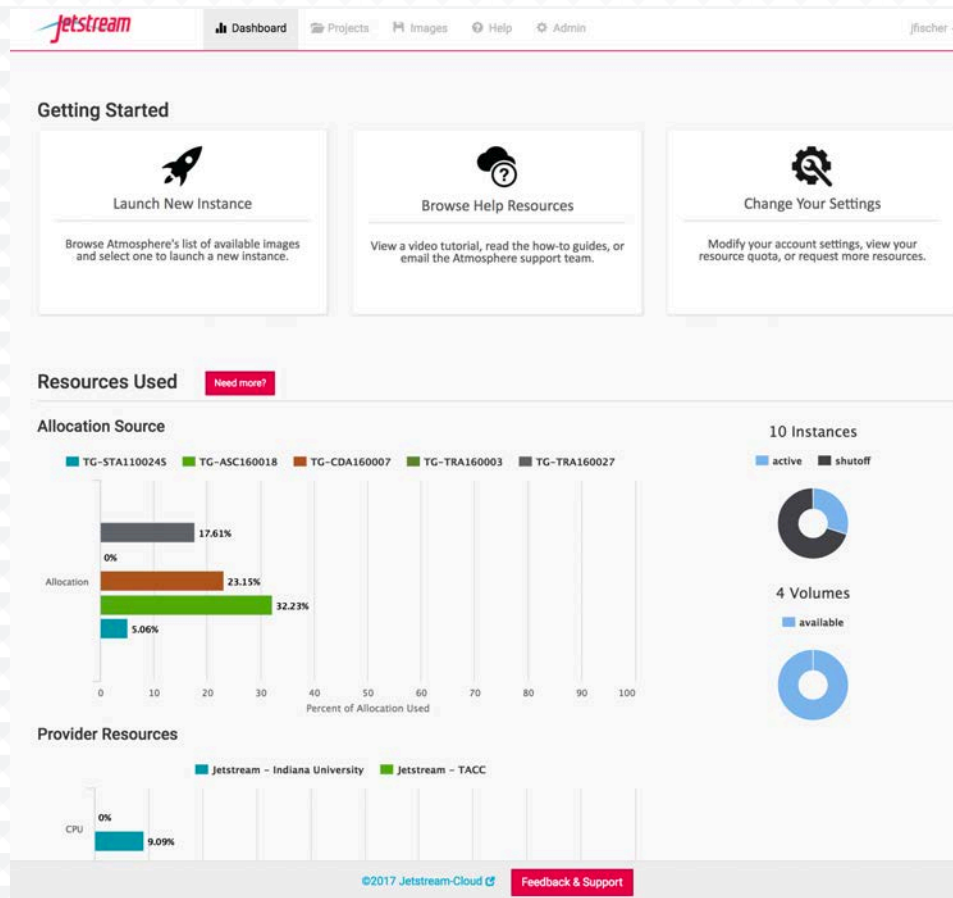
Jetstream System Overview



Platform Overview



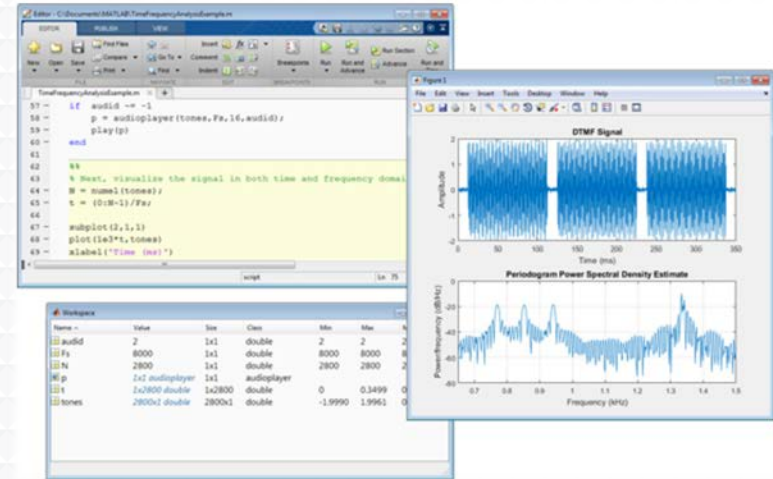
The Jetstream Atmosphere web interface



Discipline or area of interest	#of Jetstream allocations	SUs allocated on Jetstream	SU Increase/Decrease on Jetstream over previous year	% of SUs allocated on Jetstream	% of all SUs allocated on other XSEDE-supported systems
Behavioral Sciences	6	3,465,516	100% Increase	4.24%	0.61%
Biological Sciences	89	15,041,928	72.18% Increase	18.40%	3.59%
Biophysics	86	3,627,026	44.15% Decrease	4.44%	13.56%
Computer Science	72	6,883,269	33.28% Increase	8.42%	2.98%
Earth Sciences	37	5,476,250	37.06% Increase	6.70%	4.60%
Education and Training	128	16,599,512	2.62% Increase	20.31%	4.66%
Engineering	13	520,690	71.21% Increase	0.64%	1.75%
Materials Science	6	1,035,508	100% Increase	1.27%	13.89%
Mathematics	13	688,505	150.37% Increase	0.84%	0.90%
Molecular Science/Biochemistry	21	4,254,643	10.15% Decrease	5.20%	5.83%
Neuroscience	19	4,708,180	327.87% Increase	5.76%	1.98%
Physics	10	2,440,581	15.58% Decrease	2.99%	8.65%
Social Sciences and Humanities	28	2,409,633	192.27% Increase	2.95%	0.81%

Jetstream for engineering researchers (and others)

- Matlab and SimuLink and additional toolkits are installed on Jetstream
- You do NOT need to have a local license to use MATLAB on Jetstream
- If you are a researcher, and MATLAB or SimuLink... you're ready to go!
- If you are an engineering researcher, and you need other tools... let us know!



Jetstream for Education

Jetstream has been used in multiple graduate and undergraduate courses

- Management, Access, and Use of Big and Complex Data
- Multiple informatics and general bioinformatics courses
- Business Intelligence (big data and analysis)
- Research Topics in Music
- Multiple genetics and sequencing courses
- Multiple information security and assurance courses

Research Data Alliance workshops, Galaxy workshops, data analysis in finance using R, security and intrusion detection, and principles in cloud computing...



Jetstream usage highlights – 1 April 2019

- 414 active XSEDE projects covering 73 fields of science and over 2300 **active users** representing **207 institutions**
- **80%** of Jetstream users have **not used any other XSEDE system**
- >190M CPU hours allocated to XSEDE projects since June 2016
- 24 active science gateways
- 54 education/teaching allocations serving over 1000 students
- 1458 mean active VMs in previous qtr, 1590 peak active VM count
- **Highest** user satisfaction in most recent XSEDE survey



Requesting access to Jetstream

- Trial allocations available **TODAY**
 - <http://wiki.jetstream-cloud.org/Jetstream+Trial+Access+Allocation>
- You can request **startup** allocations **anytime**. (Startups are simple!)
- <http://wiki.jetstream-cloud.org/Jetstream+Allocations>
- You can request allocations for **educational** use **anytime**.
- Next submission period for large allocations is 15 Jun 2019 – 15 Jul 2019
- Research allocation: Project desc (≤ 10 pages) and Scaling doc (≤ 5 pages)
 - **We can help!**

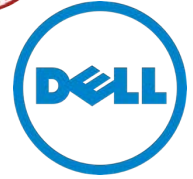
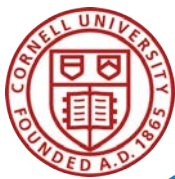


Where can I get help?

- **Wiki / Documentation:** <http://wiki.jetstream-cloud.org>
- User guides: <https://portal.xsede.org/user-guides>
- XSEDE KB: <https://portal.xsede.org/knowledge-base>
- Email: help@xsede.org
- Campus Champions: <https://www.xsede.org/campus-champions>
- Introduction to Jetstream Virtual Workshop: <https://cvw.cac.cornell.edu/jetstream/>
- Jetstream Allocations Virtual Workshop: <https://cvw.cac.cornell.edu/JetstreamReq/>



Jetstream Partners



funded by the National Science Foundation
Award #ACI-1445604



Questions?

- Project website: <http://jetstream-cloud.org/>
- Project email: help@jetstream-cloud.org Direct email: jeremy@iu.edu

License Terms

- Fischer, Jeremy. May 20, 2019. Jetstream Overview – Programmable Cyberinfrastructure: Learning the Jetstream API. Also available at: <http://Jetstream-cloud.org/research/publications.php>
- Jetstream is supported by NSF award 1445604 (David Y. Hancock, IU, PI)
- XSEDE is supported by NSF award 1053575 (John Towns, UIUC, PI)
- This research was supported in part by the Indiana University Pervasive Technology Institute, which was established with the assistance of a major award from the Lilly Endowment, Inc. Opinions presented here are those of the author(s) and do not necessarily represent the views of the NSF, IUPTI, IU, or the Lilly Endowment, Inc.
- Items indicated with a © are under copyright and used here with permission. Such items may not be reused without permission from the holder of copyright except where license terms noted on a slide permit reuse.
- Except where otherwise noted, contents of this presentation are copyright 2015 by the Trustees of Indiana University.
- This document is released under the Creative Commons Attribution 3.0 Unported license (<http://creativecommons.org/licenses/by/3.0/>). This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.



And onward...



Cloud Computing Terms...simplified

Image: a file on a disk. It will be booted to create an...

Instance: a running virtual server; i.e. something you can log into.

Running: the *instance* is up & running

Suspended: the *instance* is memory resident but not running

Stopped: the *instance* is shutdown akin to powering down

Shelved: the *instance* is shutdown, backedup, and stored



Cloud Computing Terms...simplified

Flavor: the size of a running instance; i.e. #cores, RAM, disk

Hypervisor: the thing the instance runs on and is managed by; something akin to a software defined hardware compute server.

Snapshot: the process of taking an instance and turning it to an image.

State: something worth remembering; i.e. the state of the system



Cloud Computing Terms...simplified (Cont.)

Object store: a blob of bits; it has a starting address & a size. There may be metadata associated with the object. The data is consumed in a streaming manner.

Block store: a software defined entity akin to an unformatted hardware disk drive.

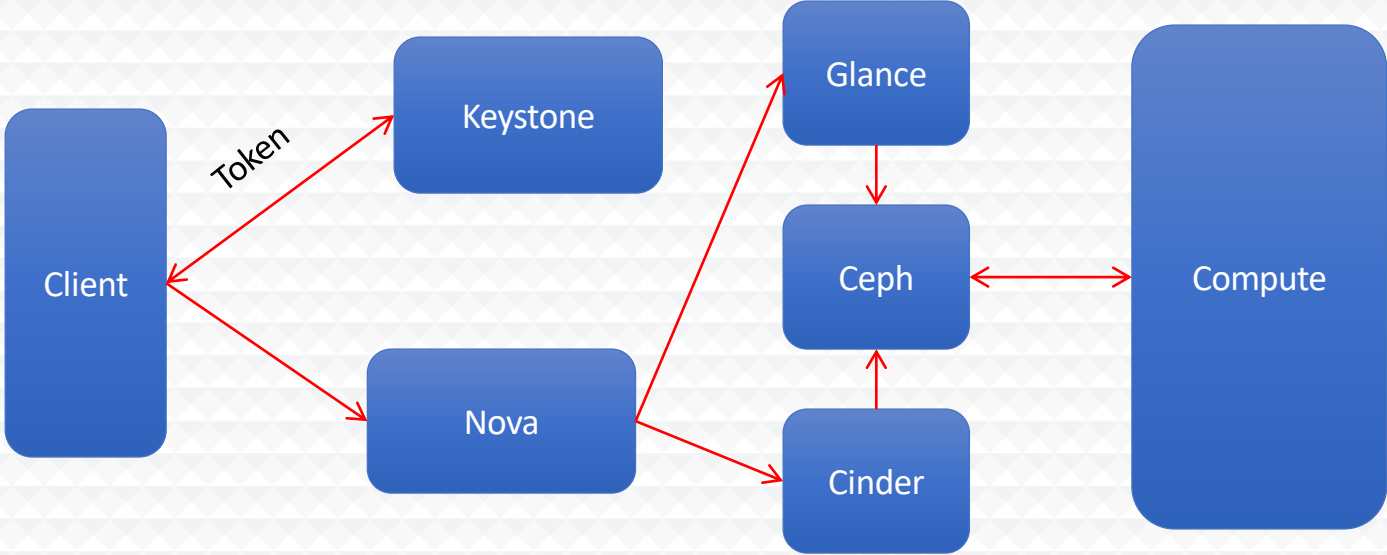
Filesystem: hierarchical in nature, directories & files, ability to open, seek, read, write.

Persistent storage: If you pull the plug, it will still exist when power is restored. Safe to store data or state here.

Ephemeral storage: If you pull the plug, it no longer exists. (Don't put your data here!!!)



OpenStack Overview



Getting into the hands on part -

Open <https://goo.gl/7X3HQz> in your browser

You'll be connecting to a prepped host...but you might want to wait until I tell you something key first. 😊

ssh trainXX@tutorial.jetstream-cloud.org