Jetstream Overview:
A national research and education cloud

A (Brief) Jetstream Overview – C2Exchange Resources for Teaching Panel
September 27, 2019 – Washington, D.C.

Jeremy Fischer – Jeremy@iu.edu - Indiana University
Manager, Jetstream Cloud, UITS Research Technologies

Expanding NSF XD’s reach and impact

Around 350,000 researchers, educators, & learners received NSF support in 2015

- Less than 2% completed a computation, data analysis, or visualization task on XD/XSEDE program resources
- Less than 4% had an XSEDE Portal account
- 70% of researchers surveyed* claimed to be resource constrained

Why are the people not using XD/XSEDE systems not using them?

- Perceived ease of access and use
- HPC resources – the traditional view of what XSEDE offers - are often not well-matched to their needs
- They just don’t need that much capability

*XSEDE Cloud Survey Report - http://hdl.handle.net/2142/45766
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
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
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...they’re coming!)
• It isn’t Amazon, Azure, or GCE (similar, but...)

RESEARCH TECHNOLOGIES
UNIVERSITY INFORMATION TECHNOLOGY SERVICES
HPC vs Cloud

Adapting to a different environment:
• No reservations, no queueing – more interactive usage
• Being your own admin – hey, we have root!**
• You really can have almost any (linux) software you want**
• Users seem to like few/no scheduled downtimes

** Here there be dragons...
Jetstream System Overview

**IU Cyberinfrastructure**
Jetstream (production)
- Compute: 320 Nodes, 7,680 Cores, 40 TB RAM, 640 TB local disk
- Storage: 960 TB

**TACC Cyberinfrastructure**
Jetstream (production)
- Compute: 320 Nodes, 7,680 Cores, 40 TB RAM, 640 TB local disk
- Storage: 960 TB

**U of Arizona Cyberinfrastructure**
Jetstream (development)
- Compute: 16 Nodes, 2 TB RAM, 384 Cores, 32 TB local disk

Network Connections:
- 100 Gbps from IU to Internet2
- 10 Gbps from TACC to Internet2
- 100 Gbps from U of Arizona to Internet2
- 10 Gbps from TACC to XSEDE
- 100 Gbps from U of Arizona to XSEDE

**Network Speeds:**
- 4x40 Gbps
- 10 Gbps
- 100 Gbps

**Research Technologies**
University Information Technology Services
Platform Overview

- Atmosphere API
- Globus Auth
- Atmo Services
- XSEDE Accounting
- OpenStack
- Ceph

Indiana University

TACC
The Jetstream Atmosphere web interface
The Jetstream Atmosphere web interface

## Featured Images

<table>
<thead>
<tr>
<th>Image</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Centos 7 (7.2) Development GUI</strong>&lt;br&gt;Jan 13th 17 03:31 by jfecher</td>
<td>Imported Application - Centos 7 (7.2) Development GUI</td>
</tr>
<tr>
<td><strong>BioLinux 8</strong>&lt;br&gt;Jan 2nd 17 03:34 by jfecher</td>
<td>Based on Ubuntu 14.04.3 -Trusty Tahr - server - clouding - <strong>requires m1.small instance</strong></td>
</tr>
<tr>
<td><strong>Ubuntu 14.04.3 Development GUI</strong>&lt;br&gt;Jan 2nd 17 03:34 by jfecher</td>
<td>Based on Ubuntu 14.04.3 Development Pached up to date as of 12/15/16 Base Ubuntu 14.04.3...</td>
</tr>
<tr>
<td><strong>Intel Development (CentOS 7)</strong>&lt;br&gt;Nov 30th 16 12:04 by jfecher</td>
<td>Intel compilers and development environment - **requires m1.small or larger VM to la...</td>
</tr>
<tr>
<td><strong>R with Intel compilers (CentOS 7)</strong>&lt;br&gt;Nov 30th 16 11:53 by jfecher</td>
<td>R with Intel compilers built on CentOS 7 (7.3) - <strong>requires m1.small or greater sized VM</strong></td>
</tr>
<tr>
<td><strong>Galaxy Standalone</strong>&lt;br&gt;Nov 15th 16 04:49 by admin</td>
<td>Galaxy 16.01 Standalone - based on Ubuntu 14.04.4 LTS - This is a standalone Galaxy server...</td>
</tr>
</tbody>
</table>
The Jetstream Web Desktop
Not just the usual suspects...

Physics, chemistry, and other “usual” HPC suspects are represented, but Jetstream also is home to projects on:

- Financial analysis / Economics
- Political science
- Humanities / Text analysis
- Network analysis
- Computer Science / Machine learning
- Satellite data analysis
Matlab for researchers and educators

• 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...
Expanding the reach: Jetstream REU Program

NSF Supplement for undergraduates
• 4 students participated in 2017
• 6 students participated in 2018
• 7 students participated in 2019

• REU student videos on YouTube
  https://www.youtube.com/user/IUPTI
Jetstream usage highlights – 1 August 2019

• 390 XSEDE projects covering 72 fields of science and over 2200 active users representing 213 institutions
• 80% of Jetstream users have not used any other XSEDE system
• >240M CPU hours allocated to XSEDE projects since June 2016

• 26 active science gateways
• 51 education/teaching allocations serving over 900 students
• 1438 mean active VMs in previous qtr, 1614 peak active VM count
• Highest user satisfaction in most recent XSEDE survey
Jetstream Timeline...what comes next?

- Completed our second year of operations with extension to November 2020
- Soliciting Research allocation requests plus Startup and Education allocations – including Science Gateways!
- Adding services as deemed useful/mature (Heat, Magnum, Trove, Manila, etc)
- Atmosphere enhancements on a regular cycle
- Working on partnerships with groups like HubZero and others to extend the value of Jetstream
Requesting access to Jetstream

- Trial allocations available TODAY
- You can request startup allocations anytime. (Startups are simple!)
- You can request allocations for educational use anytime.
- Next submission period for large allocations is 15 Sep 2019 – 15 Oct 2019
- Research allocation: Project desc (<10 pages) and Scaling doc (<5 pages)
  - We can help!
Where can I get help?

- User guides: [https://portal.xsede.org/user-guides](https://portal.xsede.org/user-guides)
- XSEDE KB: [https://portal.xsede.org/knowledge-base](https://portal.xsede.org/knowledge-base)
- Email: help@xsede.org
- Campus Champions: [https://www.xsede.org/campus-champions](https://www.xsede.org/campus-champions)
- Introduction to Jetstream Virtual Workshop: [https://cvw.cac.cornell.edu/jetstream/](https://cvw.cac.cornell.edu/jetstream/)
- Jetstream Allocations Virtual Workshop: [https://cvw.cac.cornell.edu/JetstreamReq/](https://cvw.cac.cornell.edu/JetstreamReq/)
Questions?

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

License Terms

- 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.