

Jetstream

A national science & engineering cloud

Prepared for GlobusWorld 2015

14 April 2015

Kurt A. Seiffert
Enterprise Architect, Indiana University – Research Technologies

Please cite as Seiffert, K.A. 2015. Jetstream: A national science & engineering cloud. Globus World Lightning Talk. 2015. Chicago, IL. 14 April 2015. <http://hdl.handle.net/2022/20137>



License specifics in last slide

funded by the National Science Foundation
Award #ACI-1445604

What is Jetstream?

- NSF's first cloud for science and engineering research across all areas of activity supported by the NSF
- Jetstream will be a user-friendly cloud environment designed to give researchers and research students access to self-provisioned computing and data analysis resources.
- Jetstream will leverage Globus tools for data movement and authentication
- It will provide a user-selectable library of domain-oriented virtual machines that users can select from to do their research.



Award #1445604



pti.iu.edu/jetstream

Who uses Jetstream

- For the researcher needing a handful of cores TODAY rather than thousands next week.
- Software creators and researchers needing to create their own customized virtual machines -or- their own “private computing system” within Jetstream.
- It will enable countless discoveries across disciplines such as biology, atmospheric science, economics, network science, observational astronomy, and social sciences.



Award #1445604



pti.iu.edu/jetstream

What does the name mean? And is it really a cloud?

- Name
 - In the atmosphere the Jetstream lies at the border of two different air masses
 - The Jetstream system stands at the border of the existing NSF-funded XD program and advanced cyberinfrastructure resources and users who have not previously used such NSF funded infrastructure before.
- Yep, it's really a cloud, or at least a cloud environment (one could quibble over the definition of cloud vis-à-vis expansability). Software layers:
 - Atmosphere interface
 - KVM
 - OpenStack
 - CentOS Linux



Award #1445604



pti.iu.edu/jetstream

Dashboard

Images

Favorites

My Images

Projects

Cloud Providers

Quotas

Settings

Search Images

Search by App Images, Tag, OS, and more

Popular Searches: [R](#) [Bisque](#) [NGS](#) [Community: Astrophysics](#)


Quick Sort: Popularity Recency Rating

[Advanced Search Options](#)

Quick Filter:

View as:

Popular Images from All Communities




Math Kernel Library

[blas](#) [fft](#) [fortran](#) [lapack](#)

Community: Mathematics

52 likes, 0 dislikes, 7 comments



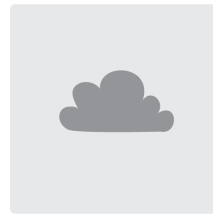
RNASeq Analysis Tools

[bowtie2](#) [blast](#) [blat](#) [edgeR](#)

[R](#) [rna-seq](#) [tophat2](#)

Community: Biology

30 likes, 2 dislikes, 4 comments




Atmospheric Dispersion Modeling

[aermod](#) [aermet](#) [aermap](#)

Community: Atmospheric Sciences

20 likes, 0 dislikes, 0 comments







MrBayes with TreeMix

[bayesian inference](#) [mrbayes](#)

[treemix](#)

Community: Phylogenetics

25 likes, 1 dislike, 10 comments


Award #1445604



pti.iu.edu/jetstream

VM Instance Sizes

Instance Type	vCPUs (48 total)	RAM (128 GB total)	Storage (2048 GB total)	Instances/Node
Tiny	1	2	20	46
Small	2	4	40	23
Medium	6	16	130	7
Large	10	30	230	4
X-Large	22	60	460	2
XX-Large	44	120	920	1

Node config: 2 Intel 2680 v3 “Haswell.” 2.5 GHz base frequency. Floating point intensive operations utilizing the AVX instruction set run at 2.1 GHz.



Award #1445604



pti.iu.edu/jetstream

Science Domains and Users

- Biology
- Earth Science/Polar Science
- Field Station Research
- Geographical Information Systems
- Network Science
- Observational Astronomy
- Social Sciences
- Jetstream will be particularly focused on researchers working in the “long tail” of science with born-digital data
- Enabling analysis of field-collected empirical data on the impact and effects of global climate change will be one of the specific foci of Jetstream
- Whatever *you* do



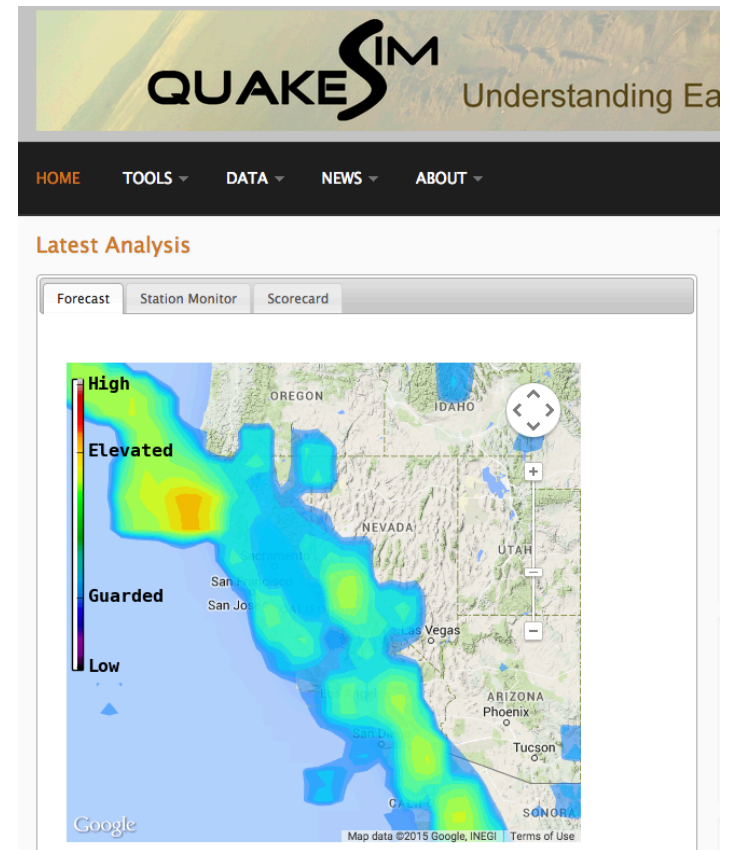
Award #1445604



pti.iu.edu/jetstream

Types of applications supported

- Interactive, VM-based work
- Persistent science gateways
- Hadoop at modest scale



Award #1445604



pti.iu.edu/jetstream

21st century workforce development

- Jetstream will include virtual Linux desktops and applications specifically aimed at enabling research and research education at small colleges and universities including HBCUs (Historically Black Colleges and Universities), MSIs (Minority Serving Institutions), Tribal colleges, and higher-ed institutions in EPSCoR (Experimental Program to Stimulate Competitive Research) States
- Jetstream will also support deployment of user-friendly Science Gateways

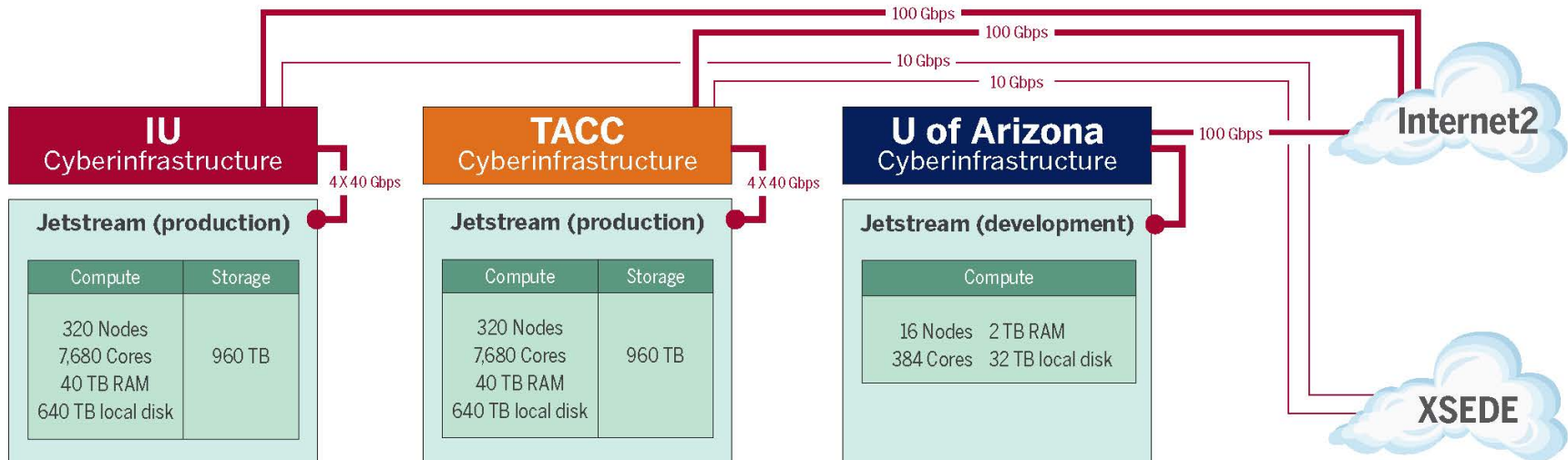


Award #1445604



pti.iu.edu/jetstream

Jetstream System Diagram



Award #1445604



pti.iu.edu/jetstream

Jetstream Deployment Partner Organizations

A seasoned team of organizations and experts:

- University of Texas Austin (TACC)
- University of Chicago (Argonne National Lab)
- University of Arizona
- University of Texas at San Antonio (Open Cloud Lab)
- Johns Hopkins University
- Pennsylvania State University



Award #1445604



pti.iu.edu/jetstream

Jetstream Deployment Partner Organizations

- Initial construction (funded partners):
 - University of Texas Austin (TACC)
 - University of Chicago (Argonne National Lab)
 - University of Arizona
 - Johns Hopkins University
- Planned funded partners for O&M phase:
 - University of Texas at San Antonio (Open Cloud Lab)
 - Penn State University
 - Cornell University
- Several unfunded collaborating partners:
 - University of Hawaii
 - University of Arkansas Pine Bluff
 - National Snow and Ice Data Center
 - University of North Carolina Odum Center



Award #1445604



pti.iu.edu/jetstream

Jetstream Application & Outreach Collaborators

- Cornell University –Ms. Susan Mehringer, Lead. Cornell® Virtual Workshops about Jetstream and applications running on jetstream.
- University of Arkansas at Pine Bluff – Dr. Jesse Walker, lead. cybersecurity education,, Minority Serving Education outreach
- University of Hawaii – Dr. Gwen Jacobs, lead. EPSCoR early adopter/user. Jacobs will chair Science Advisory Board
- National Snow and Ice Data Center (NSIDC) – Dr. Ron Weaver, lead. Data retrieval from NSIDC, application integration with ice sheet analysis applications
- University of North Carolina, Odum Center –Dr. Thomas Carsey , lead. Data retrieval from Dataverse Network
- National Center for Genome Analysis at Indiana University – providing genome analysis software. Includes TACC, PSC, and SDSC as partners



Award #1445604



pti.iu.edu/jetstream

Timeline

- Test gear arriving this quarter
- Production gear in mid-summer
- Friendly user mode prior to SC
- Production operation status by January 2016



Award #1445604



pti.iu.edu/jetstream

Questions

- Follow-up questions may be sent via email to stewart@iu.edu



Award #1445604



pti.iu.edu/jetstream

License Terms

- Seiffert, K.A. 2015. Jetstream: A national science & engineering cloud. Globus World Lightning Talk. 2015. Chicago, IL. 14 April, 2015. <http://hdl.handle.net/2022/20137>
- 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 2014 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.
- **FOR JETSTREAM RELATED WORK:**
- This research was supported in part by the National Science Foundation through Award ACI-1445604. 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.



Award #1445604



pti.iu.edu/jetstream