

Jetstream

Jetstream: A new national research
and education cloud

Jeremy Fischer (jeremy@iu.edu)

ORCID 0000-0001-7078-6609

Senior Technical Advisor, Collaboration and Engagement Support
UITS Research Technologies



funded by the National Science Foundation
Award #ACI-1445604

What is Jetstream?

- First production cloud for science and engineering research across all areas of activity supported by the NSF
- Interactive computing and data analysis resources “on demand”
- Part of the NSF eXtreme Digital (XD) program - an advanced, nationally distributed, open cyberinfrastructure comprised of various computational and scientific resources connected by high-bandwidth networks, integrated by coordinated policies and operations

What is Jetstream?

- Focus on ease-of-use, broad accessibility
- Reproducibility: Store, publish via IU Scholarworks (DOI)
- Will support persistent gateways (iPlant, Galaxy, generic “SciGAP” build-a-gateway image)
- The primary goal set by IU and its partners in implementing Jetstream is to create a resource that expands the users of XD program resources beyond the current community of users.
- VM library, custom VMs, or “private computing system”

Not just another XD resource (Why Jetstream?)

- National Science Foundation (NSF) estimates that 299,000 researchers, educators, and learners received direct support during the year ending September 2013, yet merely 1.5% completed a computation, data analysis, or visualization task on XD program resources and less than 3% had an account on the XSEDE portal.
- Jetstream, through its scale and flexibility, will dramatically enhance the diversity and size of the US researcher community benefiting from XD resources by focusing on lowering the barriers to scientific computing.

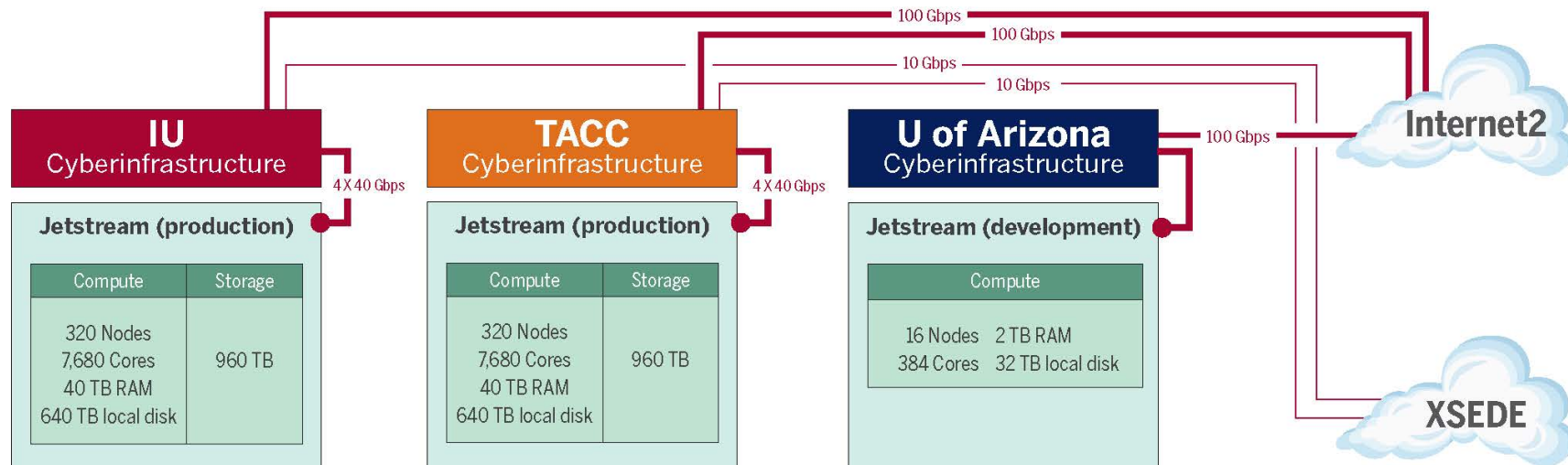
Who will use Jetstream?

- 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
- As a backend supporting science gateways
- Doing Hadoop at a modest scale

Science Domains

- Biology: iPlant and Galaxy VMs, enabling access to and use of new analytical codes in various modalities
- Earth Science: VMs capable of requesting NSIDC data and running common routines to enable more effective research and better analyses of data
- Field Station Research: VM-based data collection and analysis tools to support data sharing and collaboration
- GIS: Deliver the CyberGIS toolkit and provide access to ArcGIS in a VM using IU's existing site license
- Network Science: Build VMs with CShell tool builders to deliver network analysis tools interactively
- Social Sciences: Create VMs that allow selection of data from the Odum Institute in a way that retains provenance and version information
- Whatever you do, probably ...unless you run large scale MPI codes or HTC workloads!

Jetstream System Overview



The Jetstream Interface

Jetstream DEMO IMAGES Help Login

SEARCH TAGS

Search across image name, tag or description

Showing 7 of 7 images

Featured Images

MAKER: P 2.28 with C...
Nov 18th 2015 02:31 am
EST by edwinintest3
Featured
MAKER: P 2.28 installed as well as wq_maker used to combine Work Queue and MAKER for flexibility.

All Images

MAKER: P 2.28 with C...
Nov 18th 2015 02:31 am
EST by edwinintest3
Featured
MAKER: P 2.28 installed as well as wq_maker used to combine Work Queue and MAKER for flexibility.

TSW Workshop Willa...
Nov 18th 2015 02:31 am
EST by edwinintest3
Featured
MAKER: P 2.28 installed as well as wq_maker used to combine Work Queue and MAKER for flexibility.

Trusty Tahr (v64)
Oct 23rd 2015 12:08 am
EST by atmodadele
This is the base image for Trusty Tahr

circos-0.3.4-x86_64
Oct 23rd 2015 12:08 am
EST by atmodadele
Imported Application - circos-0.3.4-x86_64

©2016 Jetstream CF Feedback & Support

Jetstream DEMO IMAGES Help Login

SEARCH TAGS

Filter by tag name or description

Showing all tags

ascp	ascp and the utility
bcftools	Tool for BCF calling
bowtie2	bowtie2
bwa	bwa
ccools	ccools
cutdiff	cutdiff
cummeRbund	cummeRbund
cytoscape	creates interconnected networks for genomic data, e.g. annotated proteins
edgeR	edgeR
fastqc	A quality control tool for high throughput sequence data
fastqscreen	FastQ Screen (Screen NGS reads for contamination v0.4.2
featured	featured images
Featured	featured images
goseq	goseq is an R package for evaluating enrichment of GO terms within categories
igb	Integrated Genome Browser
igv	igv
komododit	Komodo Edit is a popular development environment
knsnp	find SNPs between bacterial genomes
maker	maker
R	R statistical software
rstudio	rstudio

©2016 Jetstream CF Feedback & Support

Jetstream Timeline...what comes next?

- Test system has completed acceptance testing at IU
- Program Execution Plan has passed peer review and been conveyed by DACI (Division of Advanced Cyberinfrastructure) to DGA (Division of Grants and Awards) for modification of award instrument
- SOW with vendor (Dell) has been executed
- Production system has arrived and is under construction!
- Friendly user mode by SC15
- ***Early operations mode Jan – Mar 2016, targeted start date of January 20, 2016***

How can I use Jetstream?

- An XSEDE User Portal (XUP) account is required. They are free! Get one at <https://portal.xsede.org>
- Read the Allocations Overview - <https://portal.xsede.org/allocations-overview>
- Write a successful allocation request – start with a Startup or Education request - <https://portal.xsede.org/successful-requests>

Where can I get help or learn more?

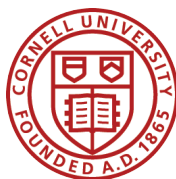
- Production:
 - User guides: <https://portal.xsede.org/user-guides>
 - XSEDE KB: <https://portal.xsede.org/knowledge-base>
 - Email: help@xsede.org
 - Campus Champions: <https://www.xsede.org/campus-champions>
 - Training Videos / Virtual Workshops (TBD)
- Early use:
 - <http://jetstream-cloud.org/>
 - Early use: jethelp@iu.edu

Jetstream Partner Organizations

Initial construction (funded partners)



Management & Operations partners



Application / community lead partners



funded by the National Science Foundation
Award #ACI-1445604



Questions?

Project website: <http://jetstream-cloud.org/>

Project email: jethelp@iu.edu Direct email: jeremy@iu.edu

License Terms

- Fischer, Jeremy. January 2016. Jetstream Overview – Prepared for 2016 ESIP Winter Meeting. Washington, D.C. Also available at: <http://hdl.handle.net/2022/20577>
- Jetstream is supported by NSF award 1445604 (Craig Stewart, 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.