A national science & engineering cloud

Jetstream - A self-provisioned, scalable science and engineering cloud environment

Craig Stewart
ORCID ID 0000-0003-2423-9019
Jetstream Principal Investigator
Executive Director, Indiana University Pervasive Technology Institute

funded by the National Science Foundation
Award #ACI-1445604

Craig A. Stewart[a]
stewart@iu.edu

Timothy M. Cockerill[b]
cockerill@tacc.utexas.edu

Ian Foster[c]
foster@mcs.anl.gov

David Hancock[d]
dyhancoc@iu.edu

Nirav Merchant[e]
nirav@email.arizona.edu

Edwin Skidmore[e]
edwin@iplantcollaborative.org

Daniel Stanzione[b]
dan@tacc.utexas.edu

James Taylor[f]
james@taylorlab.org

Steven Tuecke[c]
tuecke@ci.uchicago.edu

George Turner[d]
turnerg@iu.edu

Matthew Vaughn[b]
vaughn@tacc.utexas.edu

Niall I. Gaffney[b]
ngaffney@tacc.utexas.edu

[a]IU Pervasive Technology Institute and IU School of Informatics and Computing

[b]Texas Advanced Computing Center
University of Texas at Austin
Road A, Austin, TX 78758

[c]Computation Institute
University of Chicago
5735 S. Ellis Ave Chicago IL 60637

[d]IU Pervasive Technology Institute
2709 E. Tenth Street
Bloomington, IN 47408-2671

[e]University of Arizona
1401 East University Boulevard
Tucson, AZ 85721

[f]Johns Hopkins University
Department of Biology
3400 N Charles St.
Baltimore MD, 21218

Award #1445604

Jetstream

pti.iu.edu/jetstream

Jetstream Cloud Services

Logout User

- 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: Community...

View as:

Popular Images from All Communities

Math Kernel Library
blas fft fortran lapack
Community: Mathematics
52 0 7

RNASeq Analysis Tools
bowtie2 blast blat edgeR R rnaseq tophat2
Community: Biology
30 2 4

Atmospheric Dispersion Modeling
aermod aermet aermap
Community: Atmospheric Sciences
20 0 0

MrBayes with TreeMix
bayesian inference mrbayes treemix
Community: Phylogenetics
25 1 10

NSF Award #1445604

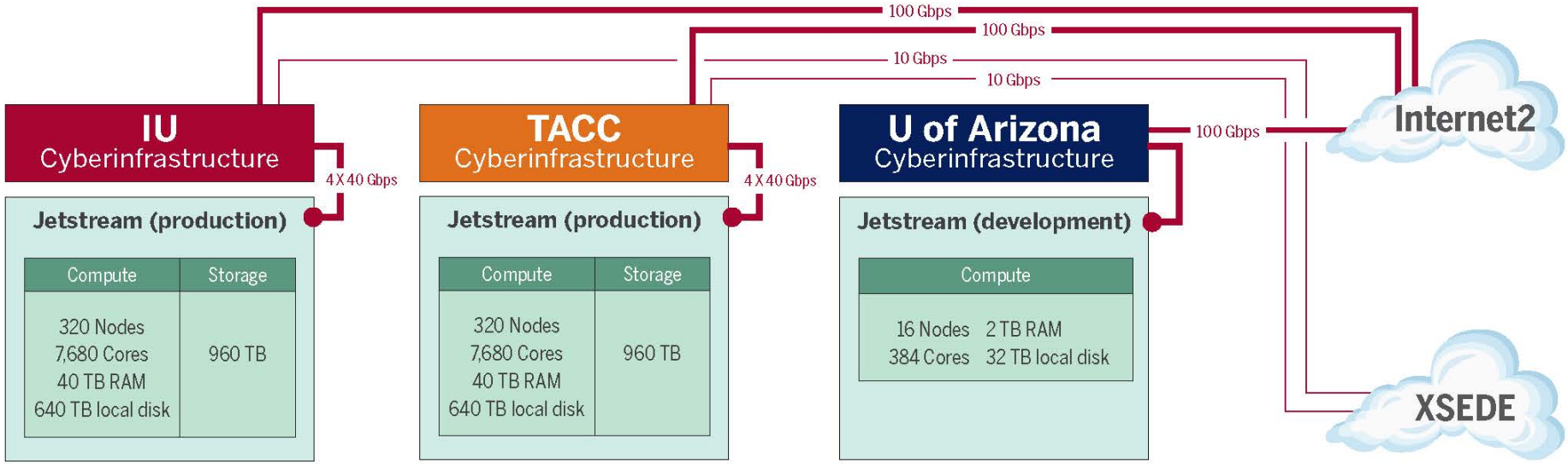
Jetstream

pti.iu.edu/jetstream

Jetstream characteristics

- First production cloud for science and engineering research across all areas of NSF-supported activity
- Interactive computing and data analysis resources “on demand”
- Focus on ease-of-use, broad accessibility
- VM library, custom VMs, or “private computing system”
  - Not every cloud computing service one might be able to buy from AWS
- Reproducibility: Store, publish via IUScholarWorks (DOI)
- Will support persistent gateways (iPlant, Galaxy, generic “SciGAP” build-a-gateway image)

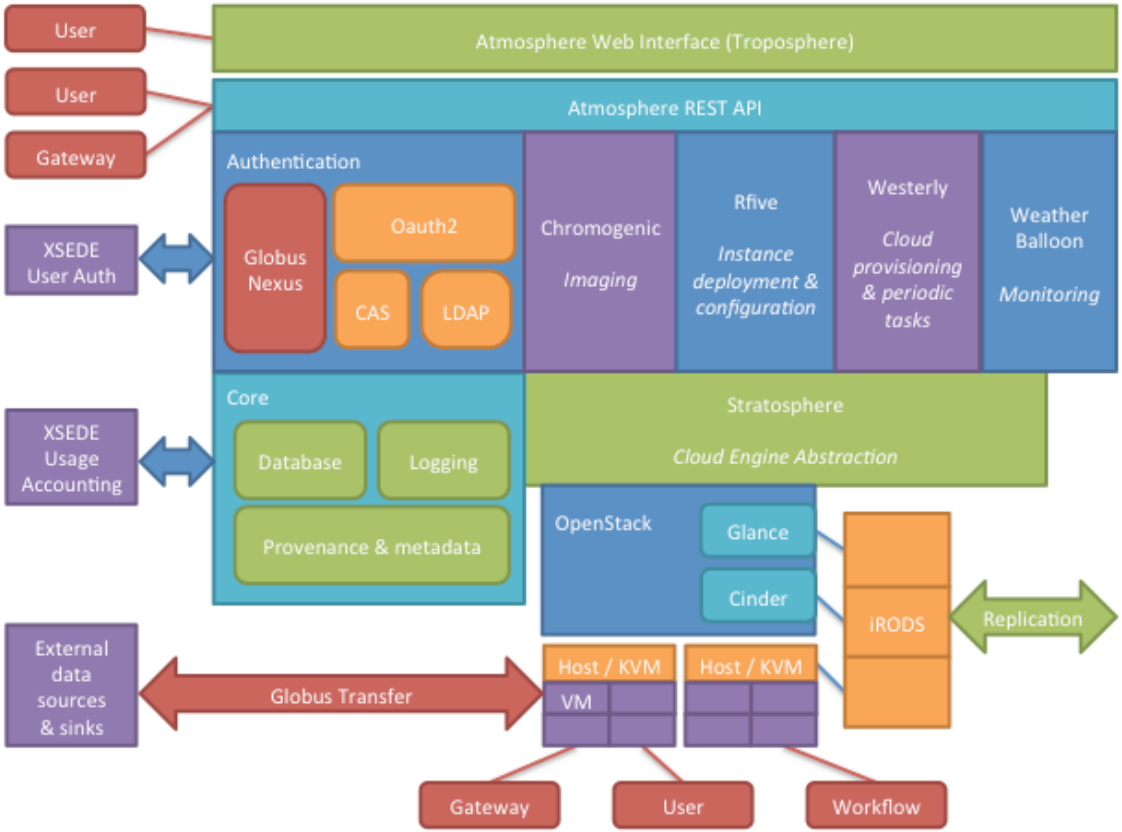
Jetstream System Overview



- Geographically distributed cloud; 0.5 PetaFLOPS
- Globus for large-scale file transfer, authentication

Award #1445604

Software Stack: Metal to Atmosphere



Award #1445604

pti.iu.edu/jetstream

Use cases drawn from 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 CIShell 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!

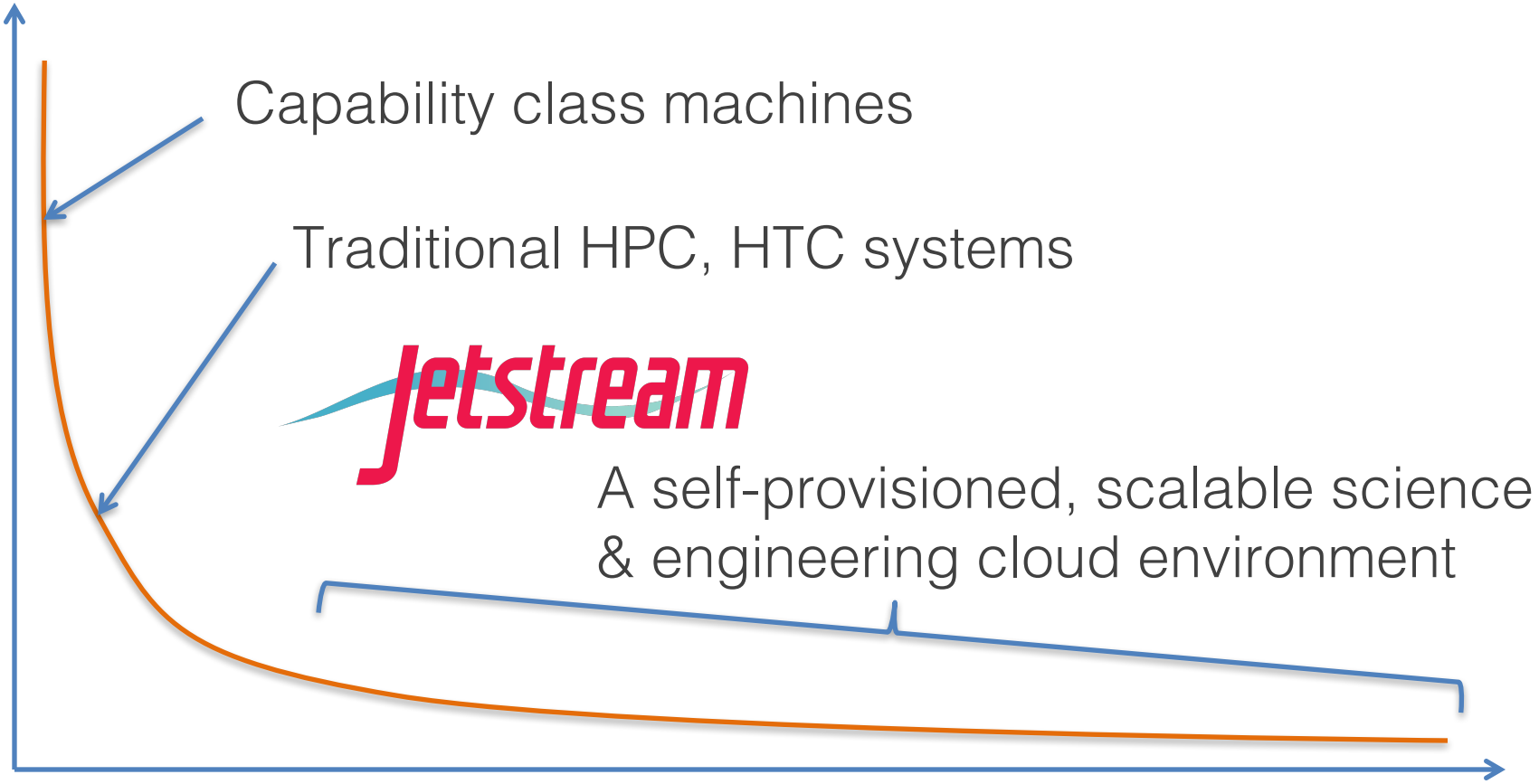
Needs based on mode of use

- Campus bridging. An old XSEDE use case: the ability to initiate an interactive computing session, detach with it running, and re-attach and continue working. Jetstream VM image featuring a user-friendly virtual Linux desktop running on Jetstream with screen images delivered to tablet devices on cellular connections or to older PCs on slow networks.
- Enable use of proprietary software.
- Facilitate reproducible data analyses.
- Enhance ease of science gateway deployment. Provide a gateway builder's toolkit, including VMs with commonly used workflow engines installed and ready to configure, XSEDE tools, and a platform for persistently hosting web services.
- Visualization and analysis. Many researchers would like to interactively use visualization.

21st-century workforce development

- Specialized virtual Linux desktops and applications to enable research and research education at small colleges and universities
- HBCUs (Historically Black Colleges and Universities)
- MSIs (Minority Serving Institutions)
- Tribal colleges
- Higher-ed institutions in EPSCoR States

“Long tail” of the NSF XD Ecosystem



Award #1445604

Jetstream

pti.iu.edu/jetstream

Challenges in implementation

- Well, it's an FOAK
- Allocations
- Appropriate use
- Scope of user community
  - Depending on existing community structures and XSEDE

Newest news

- Done
  - Test system in 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 been ordered
- Planned
  - Friendly user mode by SC15
  - Early operations mode Jan – Mar 2016

Jetstream Partner Organizations

Initial construction (funded partners)

Management & Operations partners

Application / community lead partners

NSF
Award #1445604

Jetstream

pti.iu.edu/jetstream

Citation, Acknowledgments, & Disclaimers

- Please cite as Stewart, C.A. 2015. Jetstream Overview – XSEDE '15 Panel - New and emerging US cyberinfrastructure resources Presentation. XSEDE'15, July 26 - 30, 2015, St. Louis, MO, USA. http://hdl.handle.net/2022/20338. Jetstream is supported by NSF award 1445604 (Craig Stewart, PI)
- XSEDE is supported by NSF award 1053575 (John Towns, UIUC, PI)
- This work was also supported by the Indiana University Pervasive Technology Institute, which was initiated with major funding from the Lilly Endowment, Inc.
- Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF) or other supporting organizations.
-