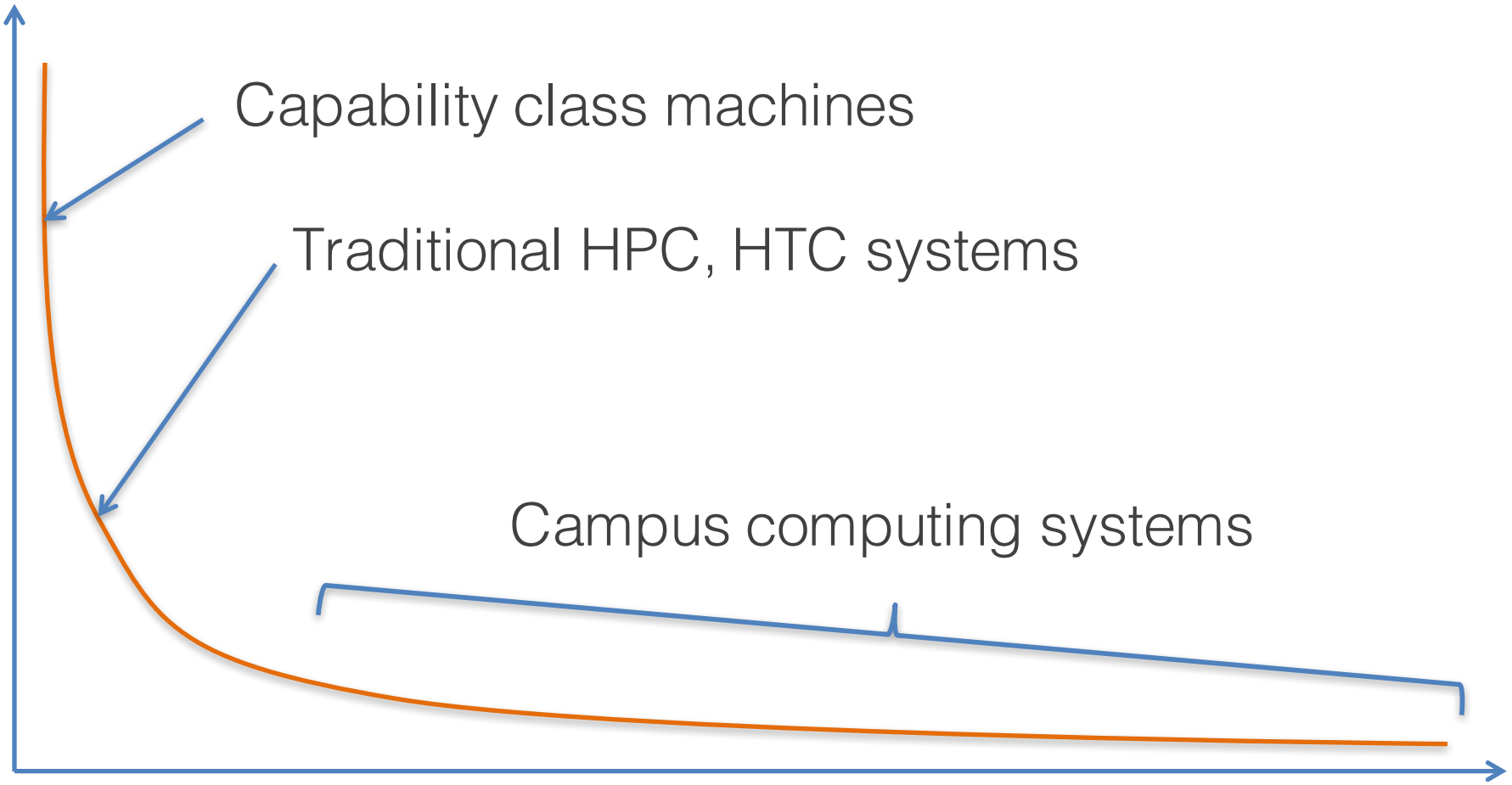
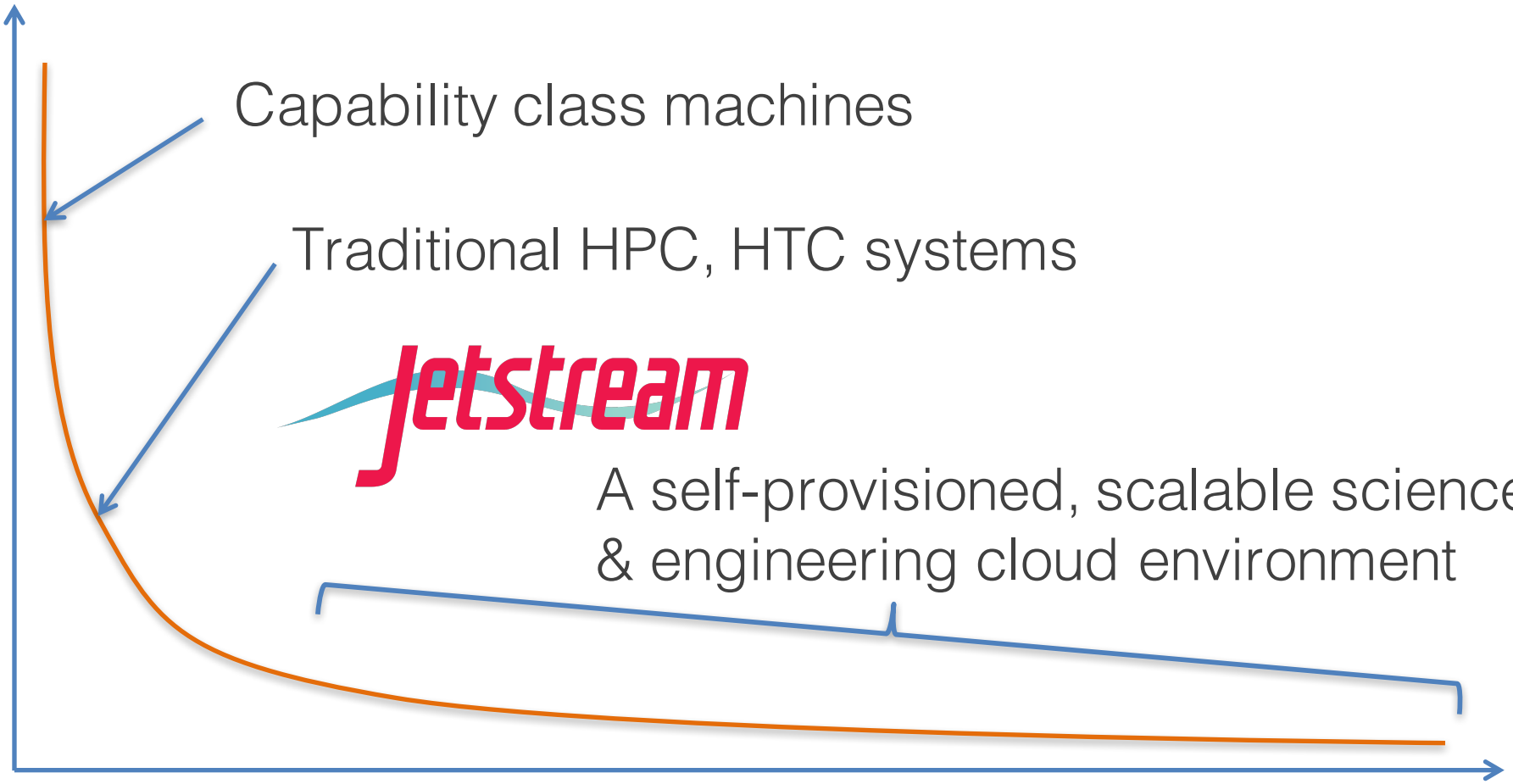
A national science & engineering cloud

funded by the National Science Foundation
Award #ACI-1445604

“Long tail” of the NSF XD Ecosystem



“Long tail” of the NSF XD Ecosystem



Award #1445604

Jetstream

pti.iu.edu/jetstream

Jetstream characteristics

- First cloud for science and engineering research across all areas of activity supported by the NSF
- Interactive computing and data analysis resources “on demand”
- Focus on ease-of-use, broad accessibility
- VM library, custom VMs, or “private computing system”
- Reproducibility: Store, publish via IU Scholarworks (DOI)

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

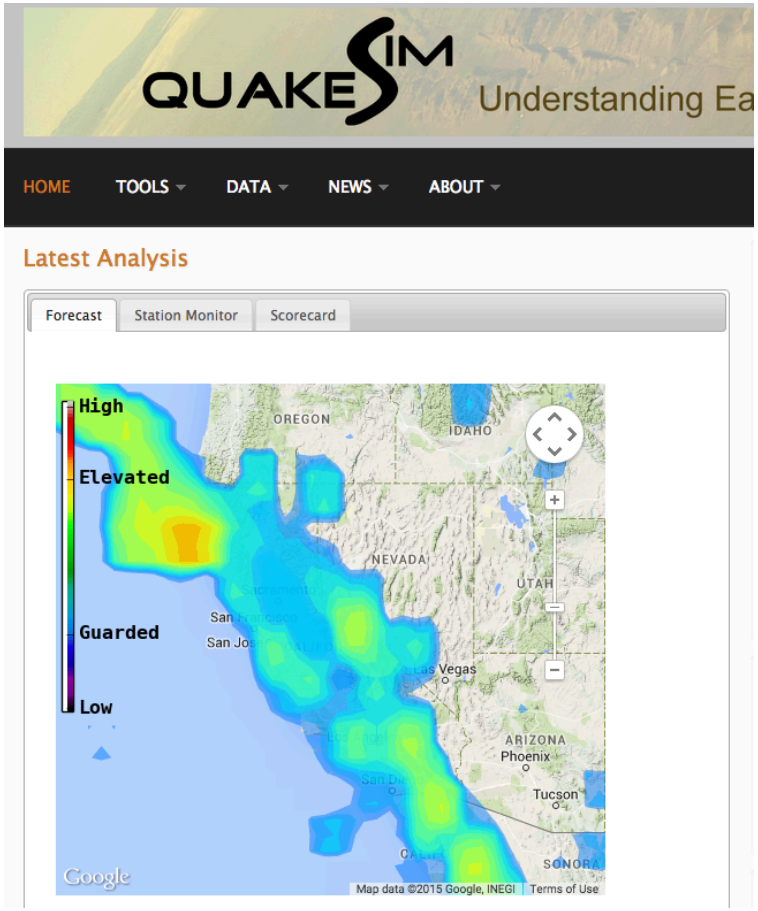
Science domains

- **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!

Types of applications supported

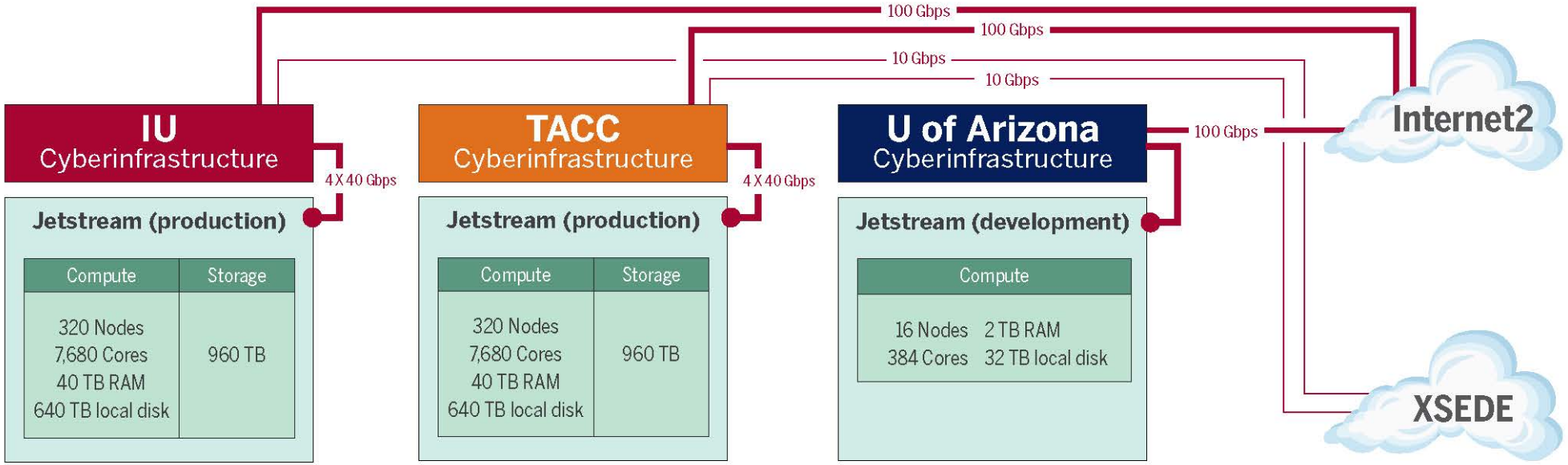
- Interactive, VM-based work
- Persistent science gateways
- Hadoop at modest scale



Award #1445604

pti.iu.edu/jetstream

Jetstream System Overview

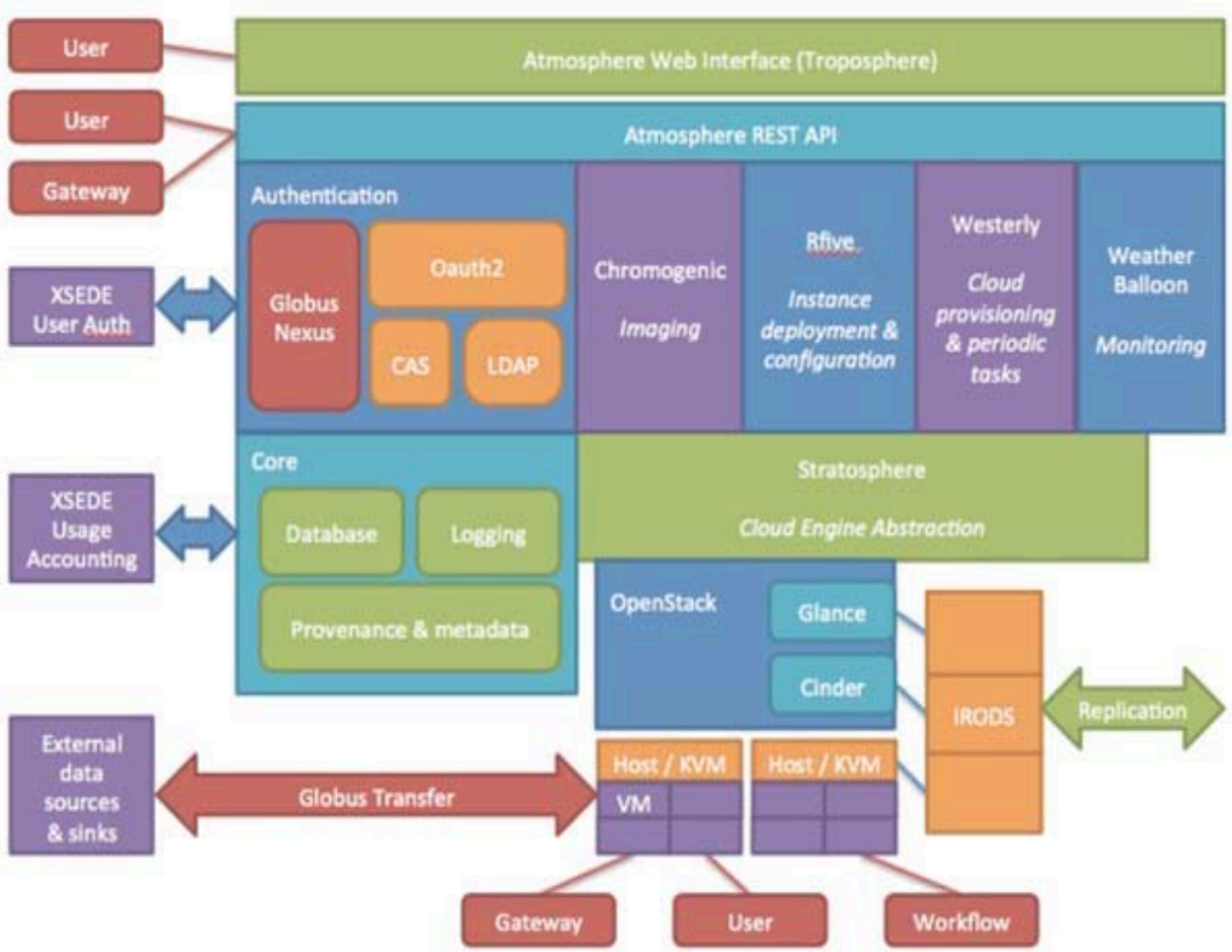


- Geographically distributed cloud; 0.5 PetaFLOPS
- Globus for large scale file transfer, authentication

Award #1445604

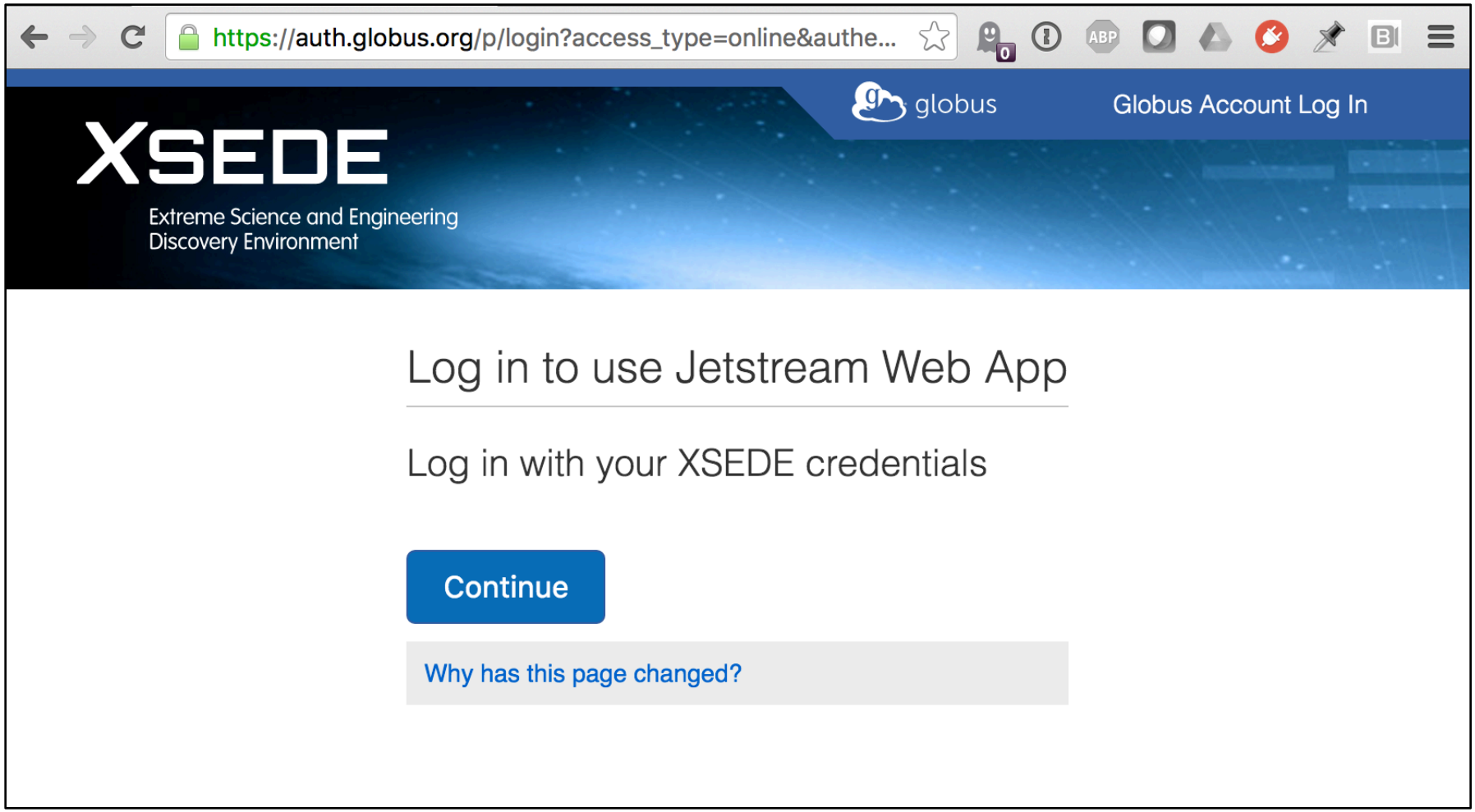
pti.iu.edu/jetstream

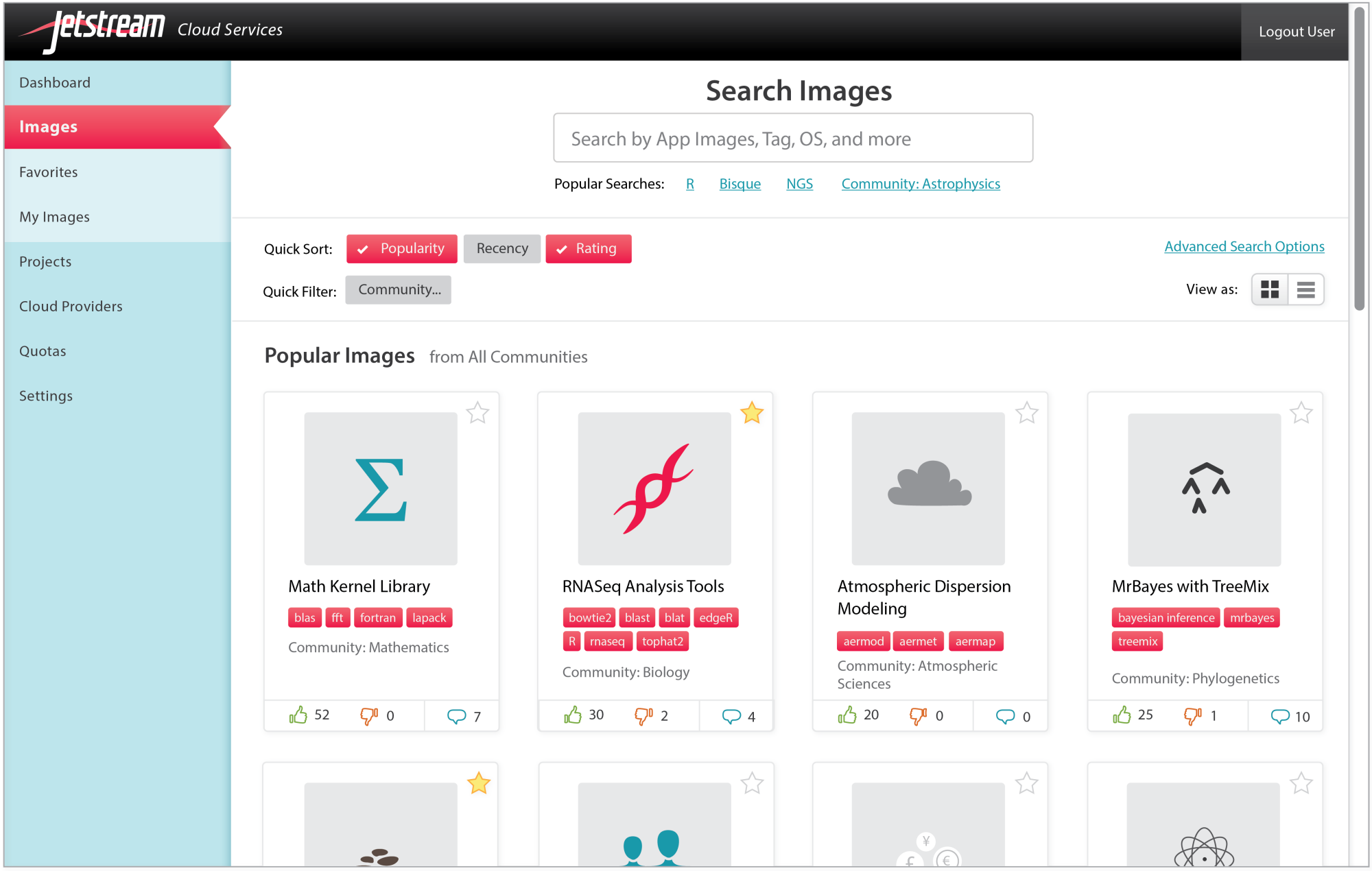
Software Stack: Metal to Atmosphere



Award #1445604

Login using Globus Auth





NSF
Award #1445604

Jetstream

pti.iu.edu/jetstream

VM Instance Sizes

Instance Type	vCPUs	RAM	Storage	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.

Planned Globus transfer use on Jetstream

- Globus Connect Personal in VM images
  - Being integrate with Atmosphere image management
- Globus Connect Server on object storage
  - Ceph RADOSGW

Jetstream Partner Organizations

Initial construction (funded partners)

Indiana University

The University of Texas at Austin

TACC

The University of Arizona

The University of Chicago

globus

Johns Hopkins University

Planned funded partners (O&M phase)

UTSA The University of Texas at San Antonio

Cornell University Founded A.D. 1865

Penn State

Unfunded partners

University of Hawai'i

University of Arkansas at Pine Bluff 1873

NSIDC National Snow and Ice Data Center

UNC The Odum Institute

NSF Award #1445604

Jetstream

pti.iu.edu/jetstream