PERVASIVE TECHNOLOGY INSTITUTE

UNIVERSITY INFORMATION TECHNOLOGY SERVICES
RESEARCH TECHNOLOGIES

Jetstream2

Research Cloud Infrastructure:

A Cloudy Introspective

Jeremy Fischer – Indiana University

Research Cloud Infrastructure Manager

RT Lunch Series

Bloomington, IN – June 8, 2023

Research Cloud Infrastructure: An Origin Story

- Who we are:
  - Jeremy Fischer – Manager, former JS/JS2 EOT/support lead
  - Mike Lowe – Senior Cloud Architect, JS/JS2 architect and lead
  - Steve Bird – Cloud Administrator, JS/JS2 admin and ResCloud admin
  - Aaron Wells – Cloud Administrator, ResCloud admin
  - Sarah Williams – Infrastructure Specialist, Monitoring wizard

Jetstream2

VOLTRON

What RCI does

We subvert the dominant paradigm by engaging services that utilize technologies hosted by a different organization than the end user.

Translation: We run cloud services and some services related to RT operations.

Jetstream2

No, really, what does RCI do?

- Jetstream2 – National Science Foundation funded research cloud
- IU Research Cloud (ResCloud) – a new cloud system built on the former Jetstream1 hardware
- Jump hosts
- Zabbix
- Vault
- Loghost
- Rundeck
- REDCap (transitioning to RCS)
- Other services to be named later

Jetstream2

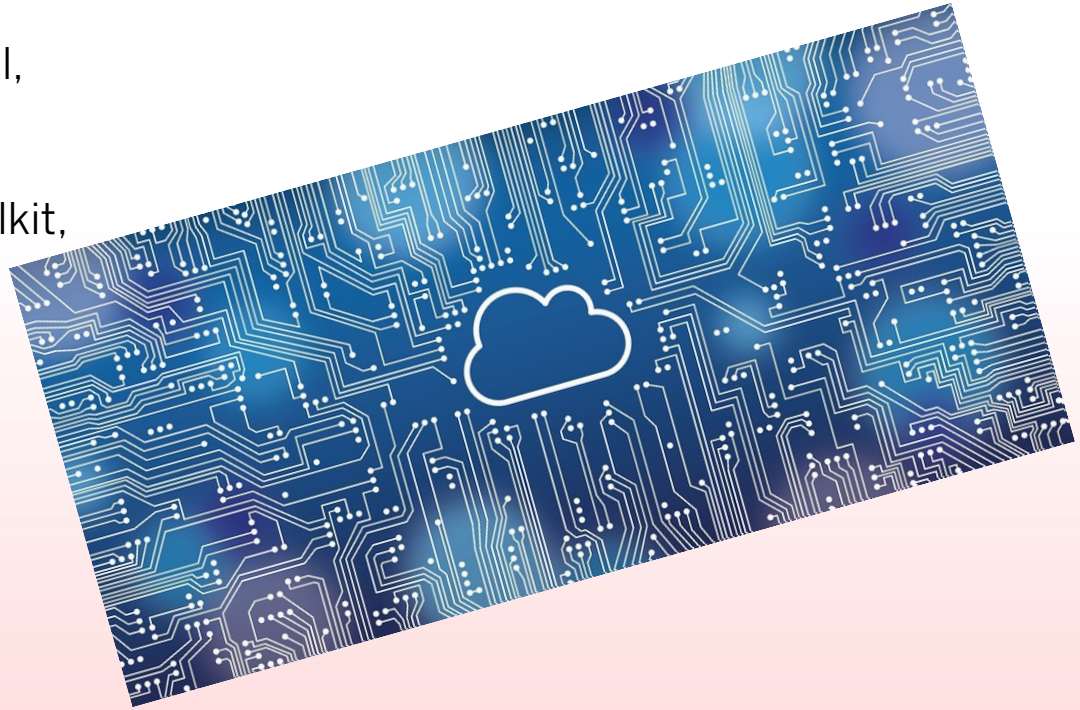
About Jetstream2

- Mostly the same as my two talks to RT in August 2022
- NSF-funded production cloud environment
- Ease-of-use focus, rapid on-ramp to ACCESS (allocated ONLY via ACCESS)
- On-demand interactive computing and persistent services for science gateways
- Enables configurable environments; *programmable cyberinfrastructure*
- Building on the success of Jetstream1
  - The 63 science gateways that utilized Jetstream indirectly supported over 183,197 people.
  - Six year of operations an overall availability of 98.54%, incl. planned and unplanned outages
  - An uptime of 99.9967% where the system was operating but at a reduced capacity

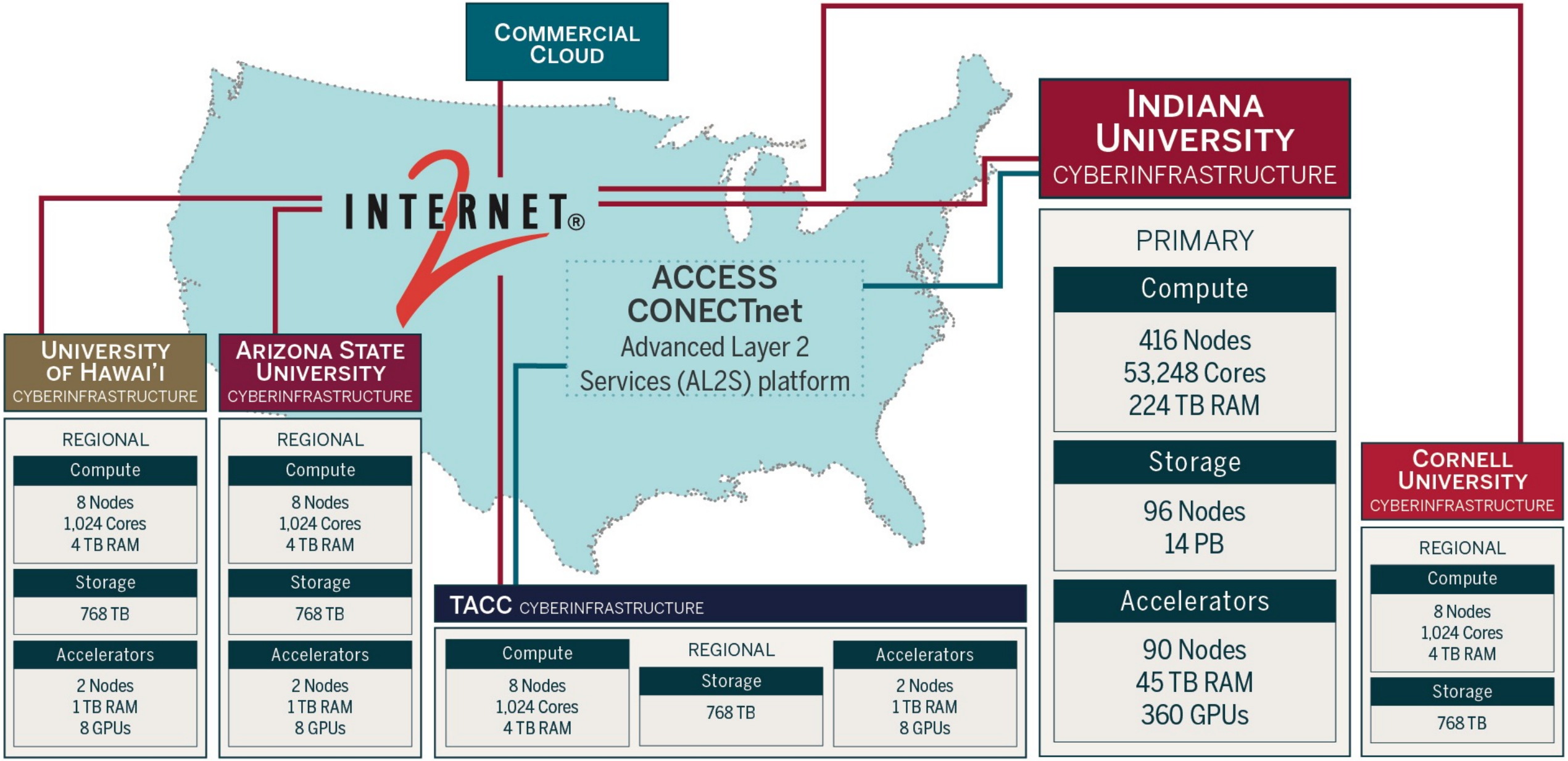
Jetstream2

Jetstream2 Features

- Primary Cloud (IU)
  - 400 compute nodes -- AMD EPYC 3rd Generation Milan CPUs – 128 cores per node + 512gb RAM
  - 90 GPU nodes – 4 x NVIDIA A100 40gb per node
  - 32 Large Memory nodes with 1TB of RAM
- Regional Clouds available by invitation/request (Arizona State, Cornell, Hawaii, TACC)
- Shared application store with common applications (NVIDIA HPC Toolkit, multiple compilers, R/Rstudio, Matlab, Anaconda, etc)
- Federated JupyterHubs, Virtual Clusters, and orchestration are all available with features being added and refined
- Commitment to >99% uptime
  - 99.87% *availability for 9-7-22 to 3-31-23* (*last* NSF reporting period)



Jetstream2


COMMERCIAL CLOUD
INTERNET2®
ACCESS CONECTnet
Advanced Layer 2 Services (AL2S) platform
INDIANA UNIVERSITY
CYBERINFRASTRUCTURE
PRIMARY
Compute
416 Nodes
53,248 Cores
224 TB RAM
Storage
96 Nodes
14 PB
Accelerators
90 Nodes
45 TB RAM
360 GPUs
UNIVERSITY OF HAWAI'I
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
ARIZONA STATE UNIVERSITY
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
TACC CYBERINFRASTRUCTURE
Compute
8 Nodes
1,024 Cores
4 TB RAM
REGIONAL
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
CORNELL UNIVERSITY
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB

Jetstream2

Jetstream2

Platform Overview

Horizon

OpenStack CLI

CACAO

Exosphere

Science Gateways

Authentication Service

ACCESS Accounting

OpenStack | Ceph | OpenStack

Jetstream2 Core System

Some sample use cases

- Science gateways
- Research-supporting infrastructure / Infrastructure as a service
- Education support – compute and desktops for courses, workshops, tutorials
- Domain science interactive compute
- Domain science long running compute
  - Small core counts, ”pleasingly parallel”, etc
- Jupyter notebooks and Hubs
- Research software development
- Machine learning – training and workflow development and data analysis
- [Your use case here]

Jetstream2

A few key stats...

1,900

Users using Jetstream2 directly

170,000

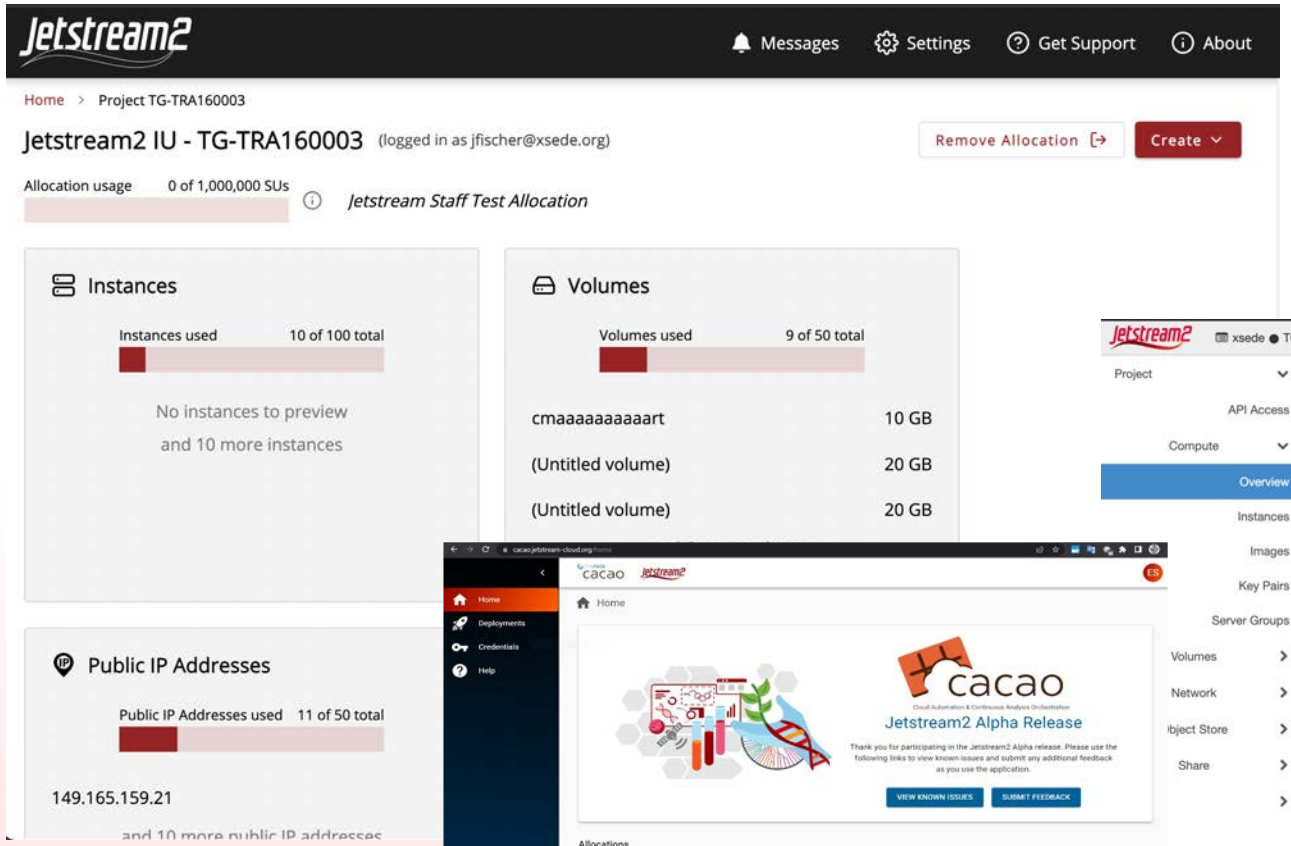
Users using Jetstream2 via Science Gateway

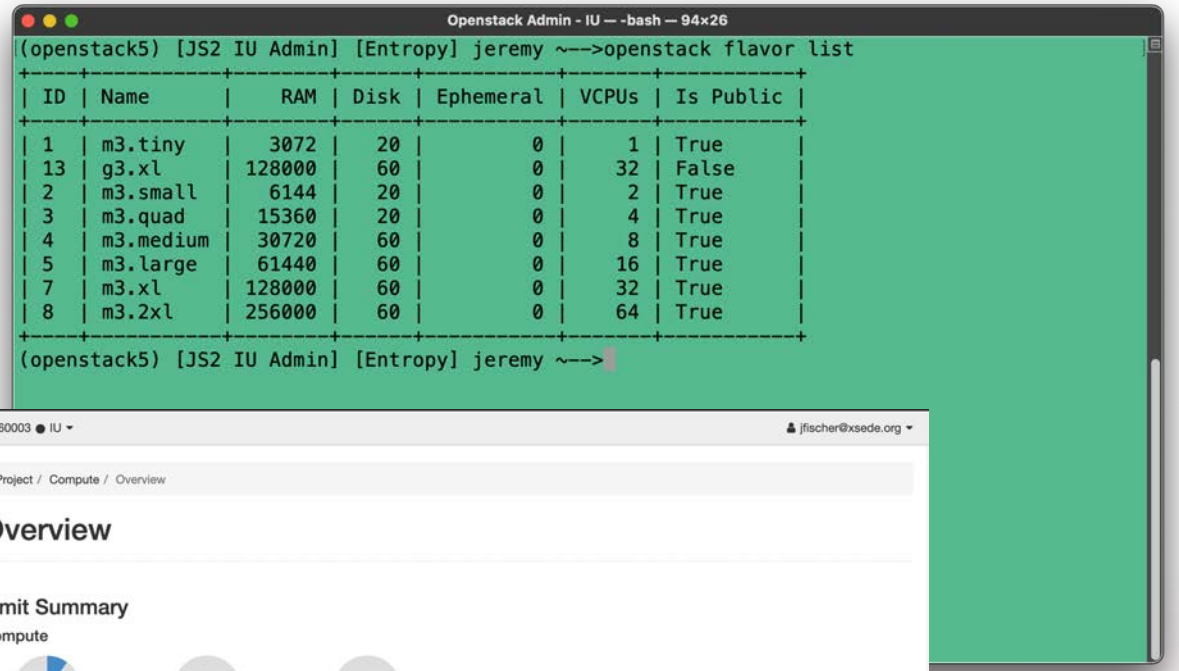
67

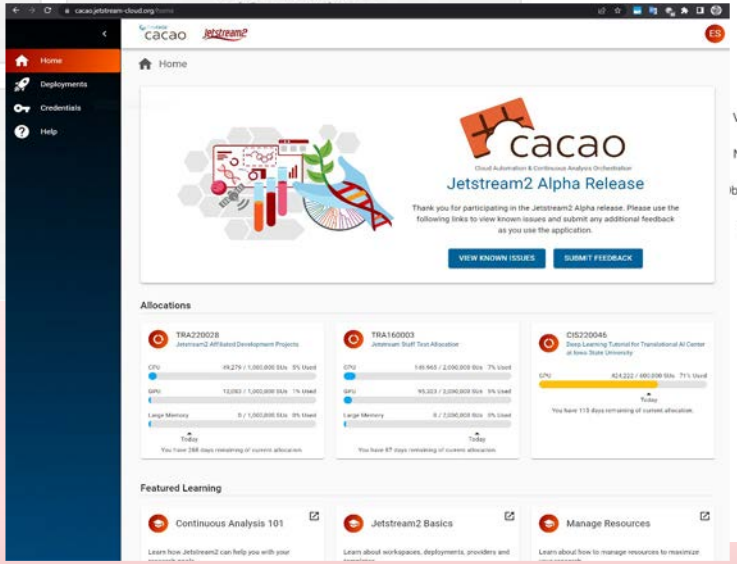
Fields of science represented on Jetstream2

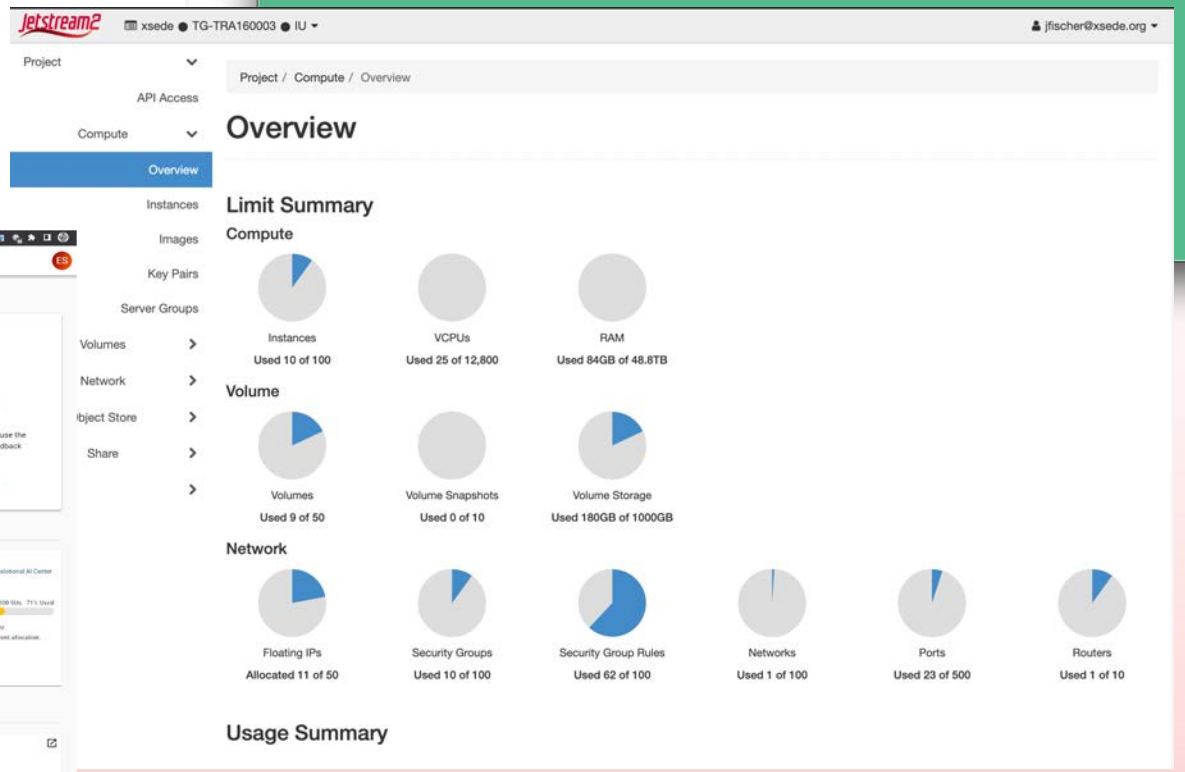
Jetstream2

How do I access Jetstream2?





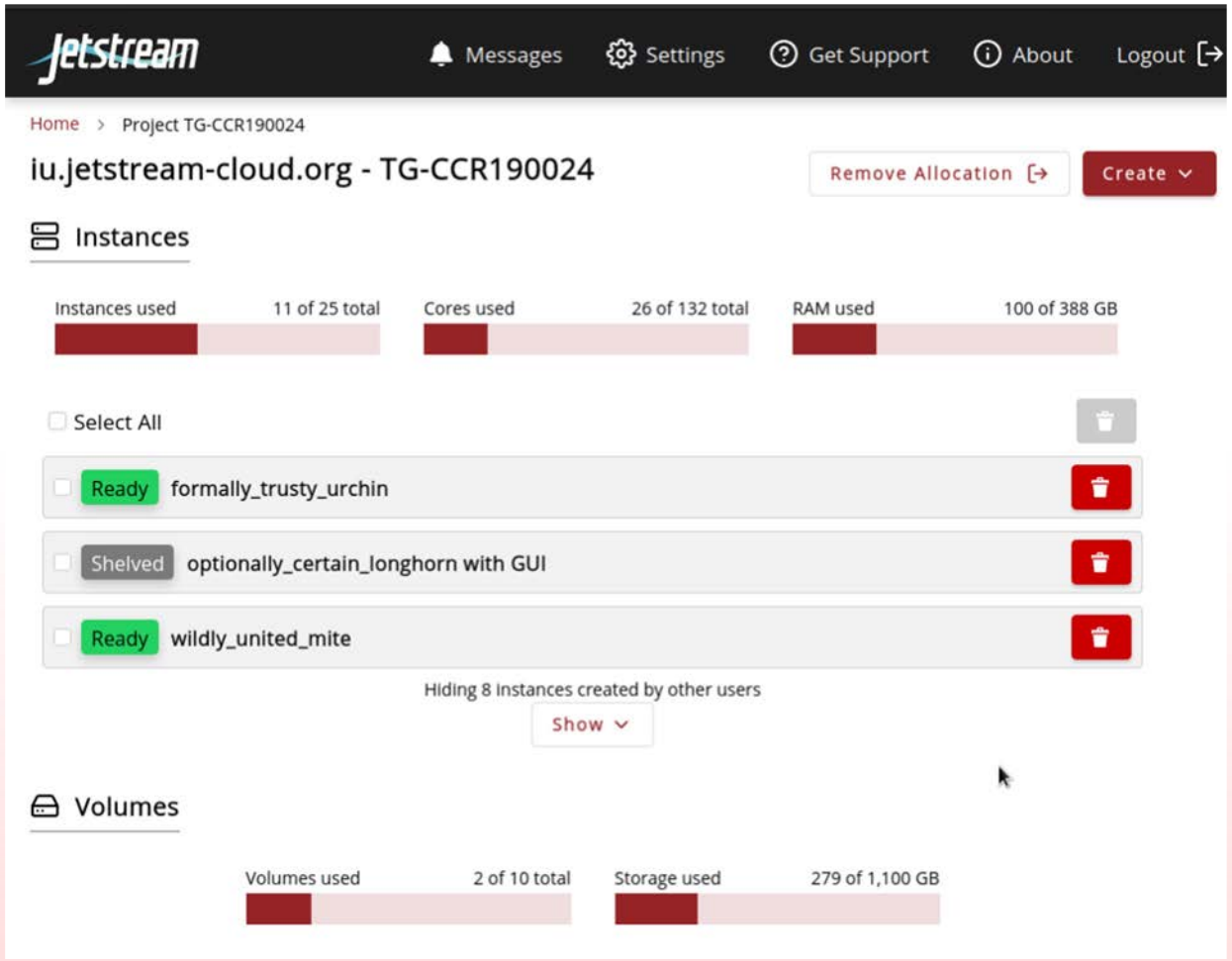


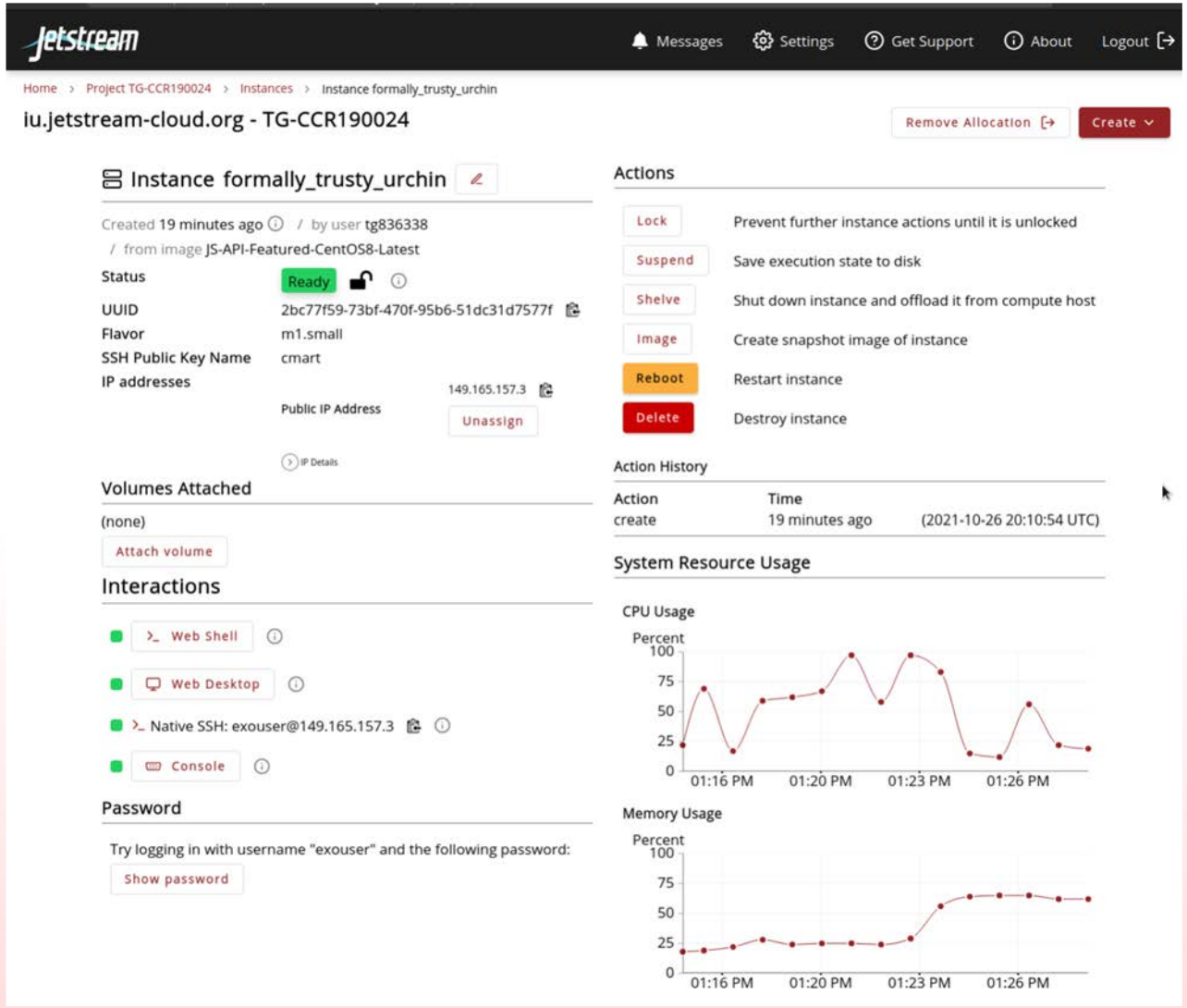


Jetstream2

https://docs.jetstream-cloud.org/overview/overview-doc/

Exosphere





https://exosphere.Jetstream-cloud.org or try.exosphere.app

Using and preserving VMs

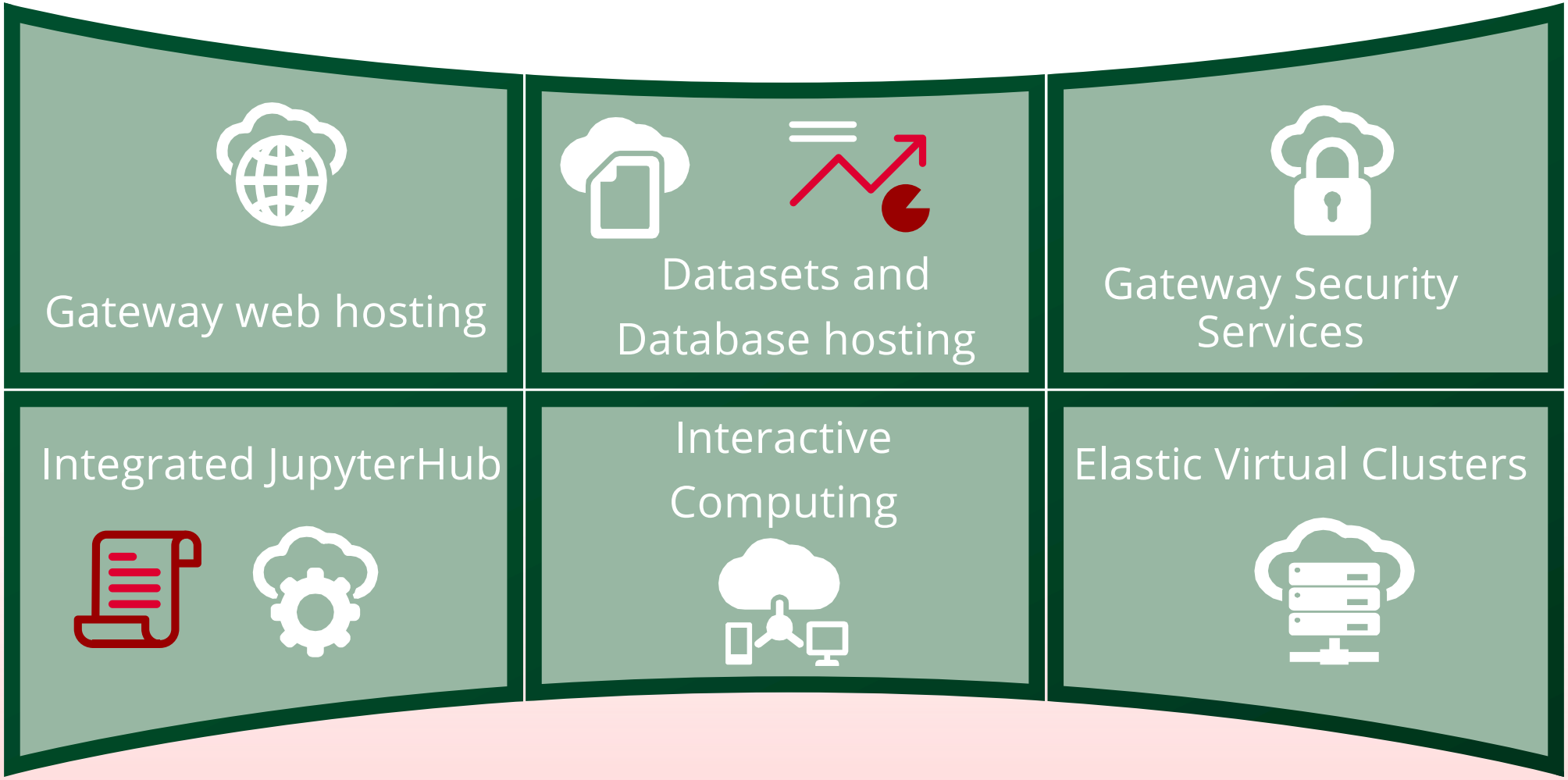
- You can install just about anything*
  - But generally limited to Linux**
- Snapshots are fairly simple and easily shared with your allocation
- One general practice is often to pull from Git(hub/lab) or pull a container

* Standard warnings about licensed software here.

** Here there be dragons.

Jetstream2

Gateways use JS2 in several ways



Jetstream2

Elastic Virtual Clusters

One Click OnDemand Cluster Augmenting the cloud capabilities

- Bundled lightweight HPC Stack, including SLURM.
- Users deploy scientific software with complete OS control.
- Dedicated and Responsive scheduler for rapid testing and development like workloads.
- Mounted persistent storage.

Jetstream2

Jetstream2 Allocation/Usage Considerations

- No scheduled downtime for upgrades
  - Upgrades are done while the system stays live overall
  - 99.87% availability for 9-7-22 to 3-31-23
- Persistent IP addresses (for the life of an allocation if desired)
- No runtime limits – VMs can exist as long as there is an active allocation with SUs available
- No allocation limits for SUs – if you can justify it and we can provide it, we do
- Instance, core, and ram limits are flexible and extendable – if you can justify it and we can provide it, we do
- Storage allocations are reasonably generous – 1TB default up to 50TB in volume, shared, or object storage

Jetstream2

The forecast calls for clouds...

- IU Research Cloud is in the process of being built
  - Repurposed Jetstream1 hardware with new storage and controllers
  - Condensed RAM from 320 computes with 128gb to 96 computes with 368gb
  - Additional RAM purchased for 1 chassis (16 nodes) of larger memory (768gb) nodes
  - 16 Intel Skylake nodes and 8 Intel Skylake nodes + NVIDIA V100s will also be integrated into ResCloud from Jetstream1 hardware

Jetstream2

ResCloud Use Cases

- Primary short-term goal 1: provide system for Research Desktop services with room to expand
- Primary short-term goal 2: condense the RT services scattered over multiple virtualization systems onto one cloud
- Longer term goals: Provide a platform for new RT-run research services and for IU researcher-run VMs
- Other goals: Exploring accelerated research desktops with the V100s

Jetstream2

Present state of ResCloud

- Several compute chassis online with more coming online
- Completing the network fault tolerance
- RED is up and running on ResCloud after a few cloud-induced hiccups
- Research Cloud Services has access to ResCloud now to start their testing and planning for service migrations to ResCloud

Jetstream2

National Science Foundation
Award #ACI-2005506

Acknowledgements

NSF Awards 1053575 & 1548562 (XSEDE), 1445604 (Jetstream), and 2005506 (Jetstream2)

This document was developed with support from the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Special thanks to the Research Cloud Infrastructure team – Mike, Steve, Aaron, and Sarah as well as the Jetstream2 PI David Y. Hancock, Malinda Husk, Winona Snapp-Childs, and George Turner (ret.)

PERVASIVE TECHNOLOGY INSTITUTE

UNIVERSITY INFORMATION TECHNOLOGY SERVICES
RESEARCH TECHNOLOGIES

Jetstream2

National Science Foundation
Award #ACI-2005506

Partners

Arizona State University

The University of Arizona

Cornell University

University of Hawai'i

Indiana University

Johns Hopkins University

TACC

UCAR

Jetstream2
AI for EVERYONE
NSF