



PERVASIVE TECHNOLOGY INSTITUTE



UNIVERSITY INFORMATION TECHNOLOGY SERVICES
RESEARCH TECHNOLOGIES

Jetstream2

Jetstream2: Accelerating cloud computing via Jetstream

Jeremy Fischer – Indiana University
Research Cloud Infrastructure Manager

STEPS Resource Provider Spotlight
Miami, FL – May 22, 2023

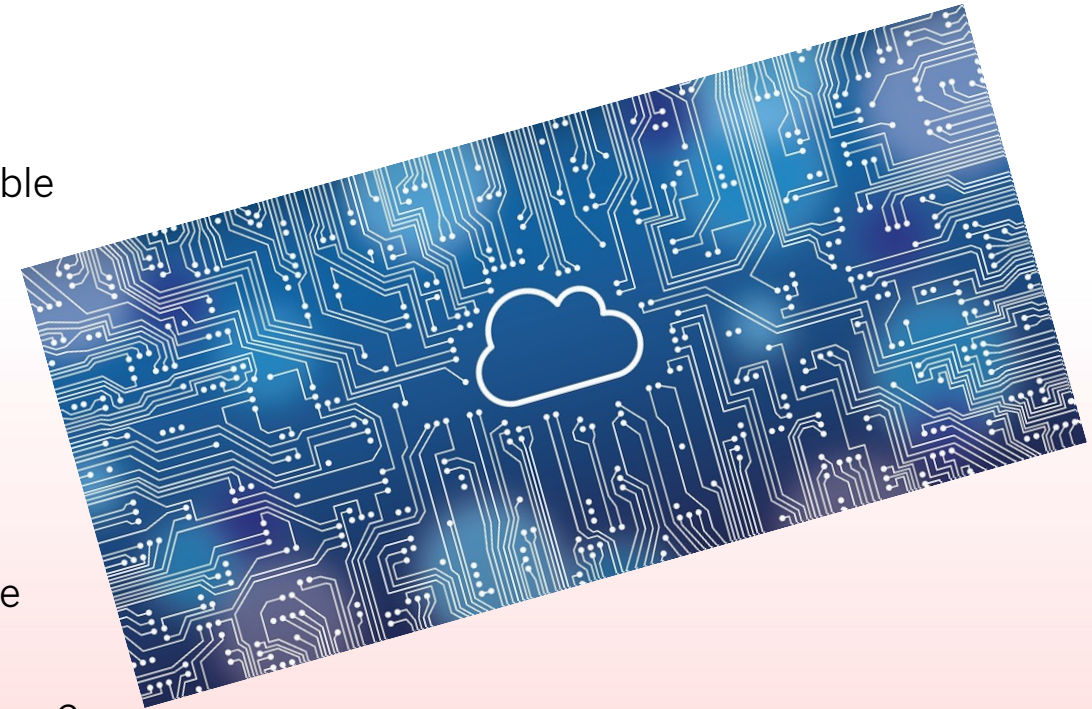
About Jetstream2

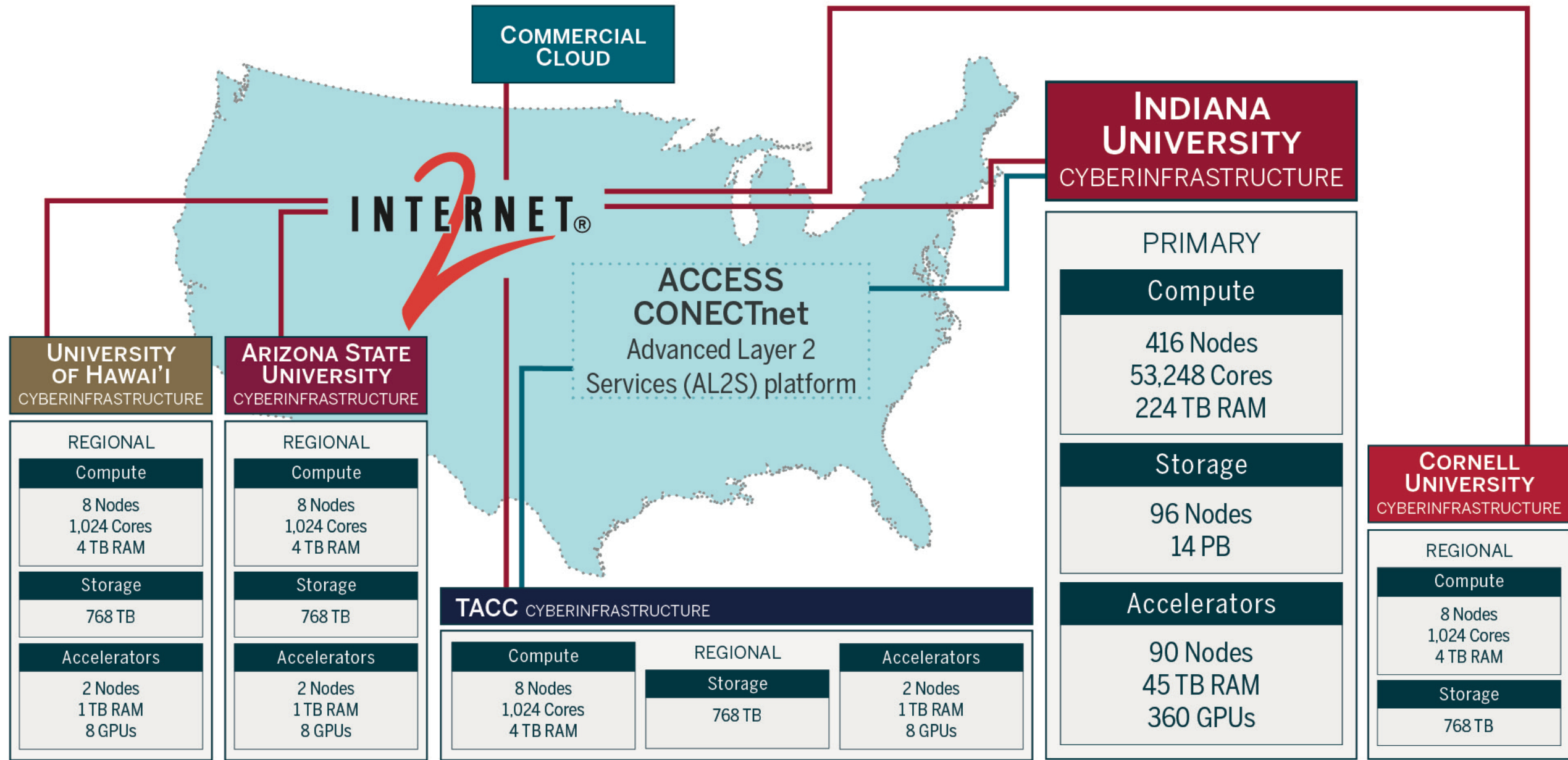
- NSF-funded production cloud environment
- Ease-of-use focus, rapid on-ramp to ACCESS
- **On-demand** interactive computing and persistent services for science gateways
- Enables configurable environments; *programmable cyberinfrastructure*
- Building on the success of Jetstream1
 - The 63 science gateways 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 Features

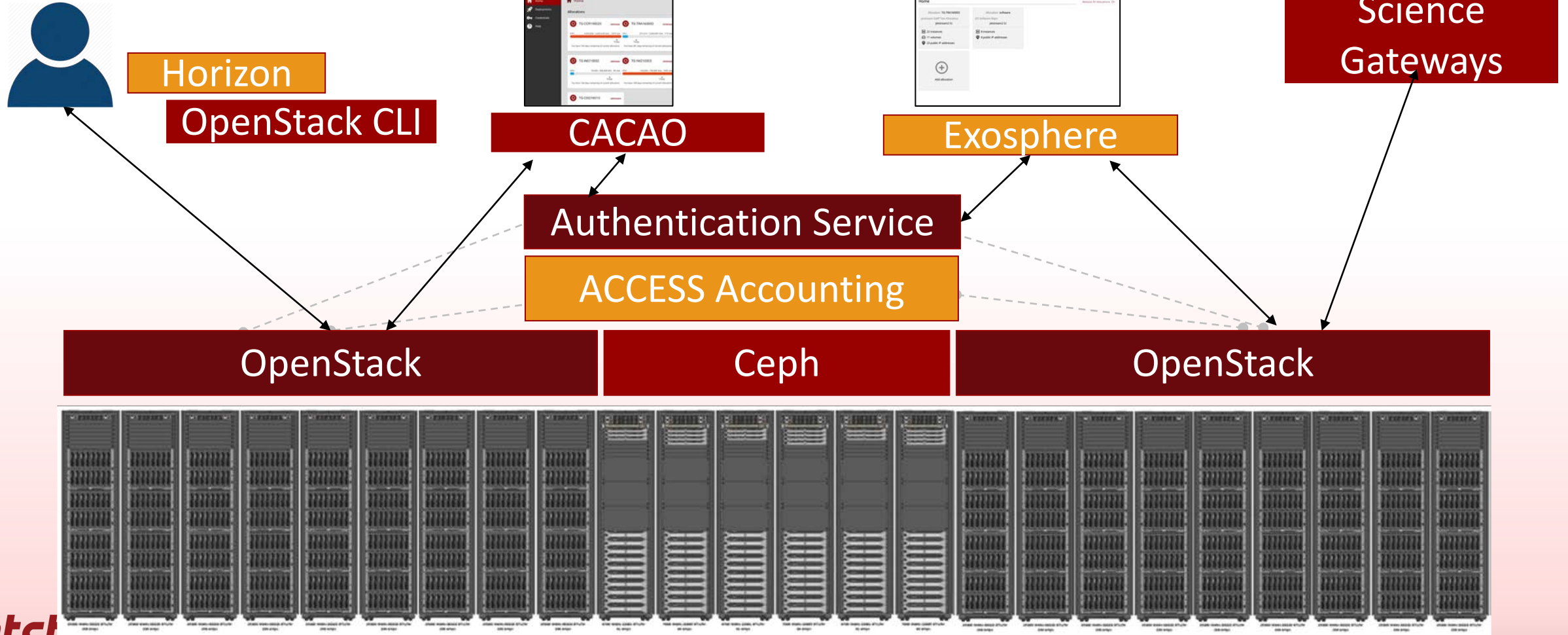
- Primary Cloud (IU)
 - 400 compute nodes -- AMD EPYC 3rd Generation Milan CPUs – 128 cores per node + 512gb RAM
 - 90 GPU nodes – NVIDIA A100 40gb
 - 32 Large Memory nodes – up to 1TB of RAM
- Regional Clouds available by invitation/request (Arizona State, Cornell, Hawaii, TACC)
- Default VM root disks and storage are NVMe. Large dataset storage available on HDDs
- Filesystems-as-a-service – natively shared filesystems between VMs
- Load-balancing-as-a-service recently deployed
- 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
- Support for commercial cloud integration and funding when using Jetstream2 coming this year





Jetstream2

Platform Overview



Horizon

OpenStack CLI

CACAO

Exosphere

Science Gateways

Authentication Service

ACCESS Accounting

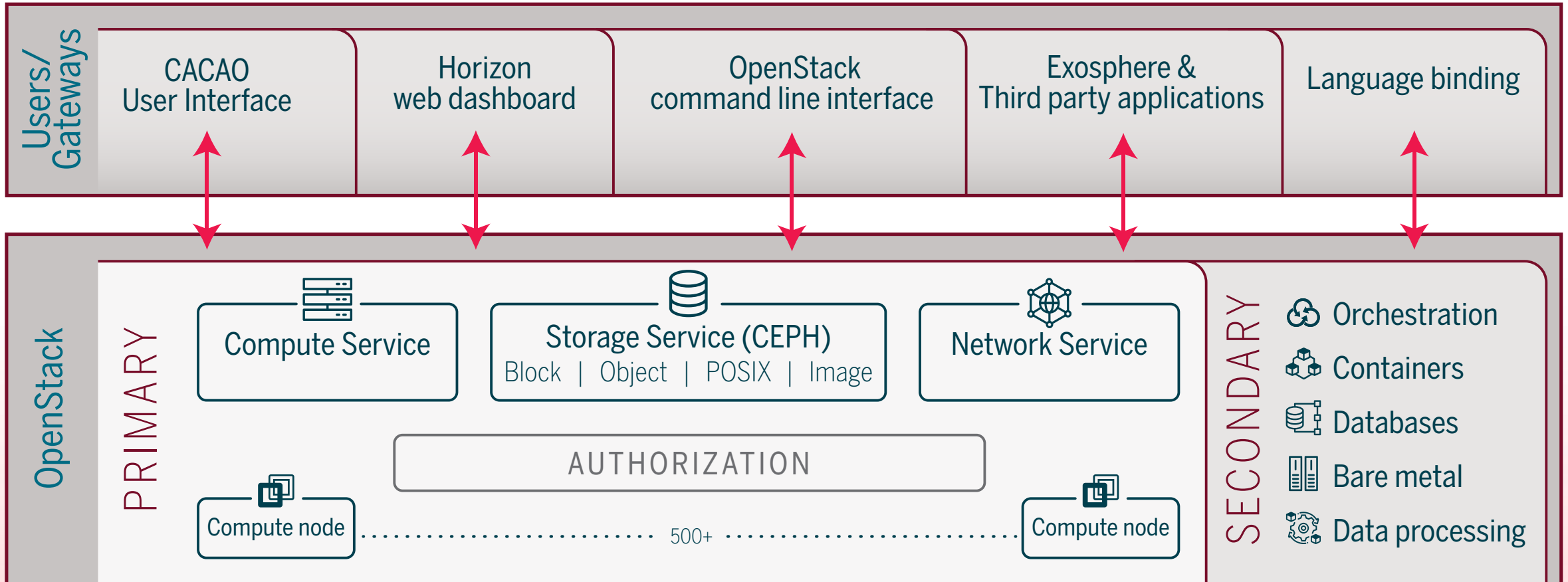
OpenStack

Ceph

OpenStack

Jetstream2 Core System

Conceptual Jetstream2 Architecture



Jetstream2 Capabilities

Enhancing IaaS model of Jetstream:

Improved orchestration support

Elastic virtual clusters

Federated JupyterHubs

Ease storage sharing (CephFS w/Manilla)

Commitment to **>99%** uptime

- Critical for science gateway hosting
- Hybrid-cloud support

Revamped User Interface

- Unified instance management
- Multi-instance launch



Feb 12, 2019 – Jet stream region called “Jet N6”
NASA/JPL-Caltech/SwRI/MSSS/Kevin M. Gill

- **>57K** cores of next-gen AMD EPYC processors
- **>360** NVIDIA A100 GPUs will provide vGPUs via NVIDIA’s MIG/vGPU feature
- **>17PB** of storage (NVMe and disk hybrid)
- 100GbE Mellanox network

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]

How do I access Jetstream2?

The screenshot shows the Jetstream2 web interface for a project named "Project TG-TRA160003". The user is logged in as "asjfischer@xsede.org". The interface includes a navigation bar with "Messages", "Settings", "Get Support", and "About". Below the navigation bar, there are buttons for "Remove Allocation" and "Create". A section titled "Allocation usage" shows "0 of 1,000,000 SUs" used. There are two main panels: "Instances" and "Volumes". The "Instances" panel shows "Instances used: 10 of 100 total" and "No instances to preview and 10 more instances". The "Volumes" panel shows "Volumes used: 9 of 50 total" and lists three volumes: "cmaaaaaaaaaart" (10 GB), "(Untitled volume)" (20 GB), and another "(Untitled volume)" (20 GB).

```
Openstack Admin - IU -- -bash -- 94x26
(openstack5) [JS2 IU Admin] [Entropy] jeremy ~-->openstack flavor list
+-----+-----+-----+-----+-----+-----+-----+
| ID | Name      | RAM   | Disk | Ephemeral | VCPUs | Is Public |
+-----+-----+-----+-----+-----+-----+-----+
| 1  | m3.tiny   | 3072  | 20   | 0          | 1     | True      |
| 13 | g3.xl     | 128000| 60   | 0          | 32    | False     |
| 2  | m3.small  | 6144  | 20   | 0          | 2     | True      |
| 3  | m3.quad   | 15360 | 20   | 0          | 4     | True      |
| 4  | m3.medium | 30720 | 60   | 0          | 8     | True      |
| 5  | m3.large  | 61440 | 60   | 0          | 16    | True      |
| 7  | m3.xl     | 128000| 60   | 0          | 32    | True      |
| 8  | m3.2xl    | 256000| 60   | 0          | 64    | True      |
+-----+-----+-----+-----+-----+-----+-----+
(openstack5) [JS2 IU Admin] [Entropy] jeremy ~-->
```

The screenshot shows the cacao website for the Jetstream2 Alpha Release. The website features a header with the cacao logo and a navigation menu. The main content area includes a "Jetstream2 Alpha Release" announcement with a "VIEW KNOWN ISSUES" and "SUBMIT FEEDBACK" button. Below the announcement, there are three "Allocations" cards showing resource usage for different projects. The "Allocations" section lists three allocations: "TRA220028", "TRA160003", and "DS220045". Each allocation card shows CPU, GPU, and Large Memory usage. The "Featured Learning" section includes links to "Continuous Analysis 101", "Jetstream2 Basics", and "Manage Resources".

The screenshot shows the Jetstream2 Overview page. The page includes a navigation menu with "Overview", "Instances", "Images", "Key Pairs", "Server Groups", "Volumes", "Network", "Object Store", and "Share". The "Overview" section displays a "Limit Summary" for Compute, Volume, and Network resources. The "Compute" section shows "Instances Used 10 of 100", "VCPUs Used 25 of 12,800", and "RAM Used 84GB of 48.8TB". The "Volume" section shows "Volumes Used 9 of 50", "Volume Snapshots Used 0 of 10", and "Volume Storage Used 180GB of 1000GB". The "Network" section shows "Floating IPs Allocated 11 of 50", "Security Groups Used 10 of 100", "Security Group Rules Used 62 of 100", "Networks Used 1 of 100", "Ports Used 23 of 500", and "Routers Used 1 of 10". The "Usage Summary" section is also visible at the bottom.



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

Exosphere

Jetstream Messages Settings Get Support About Logout

Home > Project TG-CCR190024

iu.jetstream-cloud.org - TG-CCR190024 Remove Allocation Create

Instances

Instances used: 11 of 25 total
Cores used: 26 of 132 total
RAM used: 100 of 388 GB

Select All

- Ready** formally_trusty_urchin 🗑️
- Shelved** optionally_certain_longhorn with GUI 🗑️
- Ready** wildly_united_mite 🗑️

Hiding 8 Instances created by other users Show

Volumes

Volumes used: 2 of 10 total
Storage used: 279 of 1,100 GB

Jetstream Messages Settings Get Support About Logout

Home > Project TG-CCR190024 > Instances > Instance formally_trusty_urchin

iu.jetstream-cloud.org - TG-CCR190024 Remove Allocation Create

Instance formally_trusty_urchin

Created 19 minutes ago / by user tg836338
/ from image JS-API-Featured-CentOS8-Latest

Status: **Ready** 🔒 ⓘ

UUID: 2bc77f59-73bf-470f-95b6-51dc31d7577f 🗑️

Flavor: m1.small

SSH Public Key Name: cmart

IP addresses: Public IP Address: 149.165.157.3 🗑️ Unassign

> IP Details

Volumes Attached

(none) Attach volume

Interactions

- 🟢 > Web Shell ⓘ
- 🟢 🖥️ Web Desktop ⓘ
- 🟢 > Native SSH: exouser@149.165.157.3 🗑️ ⓘ
- 🟢 🖥️ Console ⓘ

Password

Try logging in with username "exouser" and the following password: Show password

Actions

- 🔒 **Lock**: Prevent further instance actions until it is unlocked
- ⏸️ **Suspend**: Save execution state to disk
- 🛑 **Shelve**: Shut down instance and offload it from compute host
- 📷 **Image**: Create snapshot image of instance
- 🔄 **Reboot**: Restart instance
- 🗑️ **Delete**: Destroy instance

Action History

Action	Time
create	19 minutes ago (2021-10-26 20:10:54 UTC)

System Resource Usage

CPU Usage

Percent

Memory Usage

Percent



<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.

Gateways use JS2 in several ways



Gateway web hosting



Datasets and
Database hosting



Gateway Security
Services

Integrated JupyterHub



Interactive
Computing



Elastic Virtual Clusters

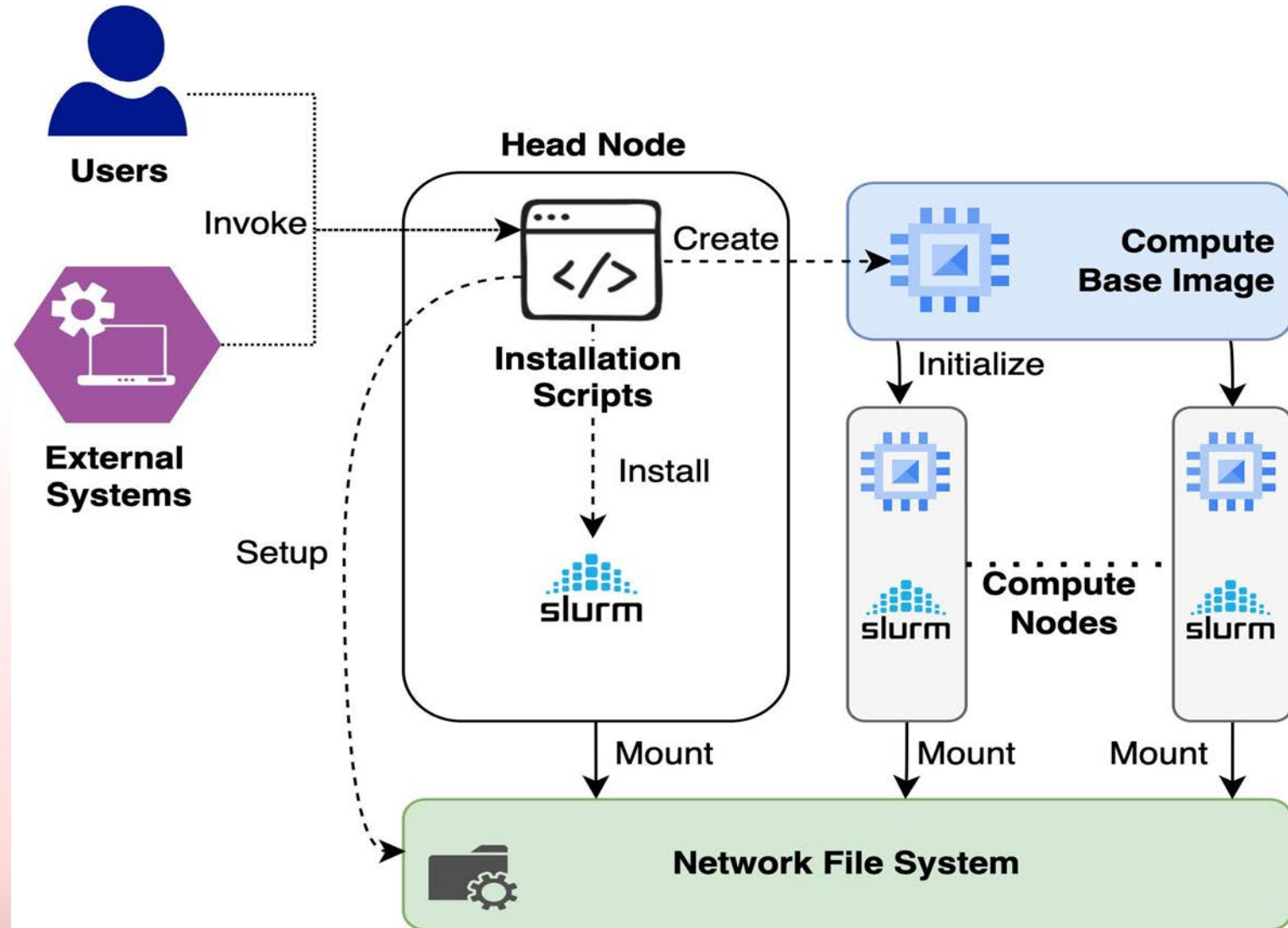


Elastic Virtual Clusters

- 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.

Virtual Clusters Architecture

- All these steps are bundled into the Ansible orchestrion.
- The entire system is a single click invoked from Exosphere.



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





PERVASIVE TECHNOLOGY INSTITUTE



UNIVERSITY INFORMATION TECHNOLOGY SERVICES
RESEARCH TECHNOLOGIES

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 contributors & Jetstream2 partners

*PI David Y. Hancock, J. Michael Lowe, Malinda Husk, Winona
Snapp-Childs, and George Turner*



PERVASIVE TECHNOLOGY INSTITUTE



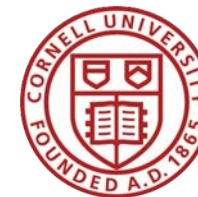
UNIVERSITY INFORMATION TECHNOLOGY SERVICES
RESEARCH TECHNOLOGIES

Jetstream2



National Science Foundation
Award #ACI-2005506

Partners





AI
for
EVERYONE

Jetstream2

