Getting Started on Jetstream2

Danny Havert – Indiana University
Systems Programmer/Analyst
Research Cloud Services
djhavert@iu.edu

Zach Graber – Indiana University
Systems Programmer/Analyst
Research Cloud Services
zegraber@iu.edu

Jenn Taylor – Indiana University
IT Project Manager
Research Cloud Services
jlrobiso@iu.edu

Julian Pistorius – Indiana University
Research Software Developer
Research Cloud Services
jupist@iu.edu

UITS IT Training, November 30, 2023
Agenda

Introduction to Jetstream2 (45 min)
- ACCESS account and training allocation
- What is Jetstream2? Whom is it for?
- How to apply for an allocation - ACCESS

Getting access to Jetstream2 & Break (15 min)
- Logging in to Exosphere

Using Jetstream2 - Basics (45 min)
- Creating and accessing your instance
- Using the software collection
- Volume/Storage management
- Transferring Data to and from your instance
- Saving and Sharing Images
- Shelving Instance

Using Jetstream2 - Advanced (45 min)
- Other Interfaces: Horizon and CACAO
- GPU instances and 3D accelerated Desktop
- Additional Requests by Attendees
Accessing Jetstream2

Before you can use Jetstream2 you must first have an ACCESS ID and be on an active Jetstream2 allocation! Fill out this form to participate in the tutorial.

https://forms.gle/JSbMkfXURPLMcaYA9
What is Jetstream2?

- A user-friendly cloud computing environment for researchers and educators.
- NSF funded. Available to any US based researcher or educator.
- CPU, GPU, and large memory nodes.
- On-demand interactive computing and data analysis resources.
- Infrastructure for gateways and other “always on” services.
- Available for Education - use it for courses!
- Virtual clustering, programmable cyberinfrastructure
What’s Special About Cloud Computing?

● Create your own virtual machine (VM or ‘instance’)
● Full sudo (admin) access. Use your instance how you want.
  ○ Install the software you need.
● On-demand resources. No sharing or queue times.
  ○ Keep your jobs running as long as you want.
● Interactive computing environment
  ○ With graphical desktop
● Full internet access with Persistent IPs
  ○ Great for web hosting
Jetstream2 Instance Sizes

- Three different resources:
  - CPU
    - Up to 64 cores, 250 GB RAM
  - Large Memory
    - Up to 128 cores, 1000 GB RAM
  - GPU
    - Up to 32 CPU cores, 125 GB RAM, and a full NVIDIA 40GB Ampere A100 GPU
  - Storage
    - Default: 1 TB
    - More on Request

View the full list: https://docs.jetstream-cloud.org/general/vmsizes/
Use Cases

- On-demand access to powerful computer for high performance analysis.
  - Useful for teaching a course in which students have this need.
- Teach a course where each student gets their own VM to work on.
  - Network security, reproducible development environment, etc.
- Write, debug, and execute code with interactive GUI applications
  - Use a GPU if needed.
- Host a file server or database. Receive, store, and serve data to colleagues.
- Host a website or web-app to share with students or colleagues.
Use Case: ML & 3D Biomedical Visualization
Use Case - Science Gateways and Portals

Galaxy:
“enables scientists to share, analyze and visualize their own data via a browser.”
Use Case - Science Gateways and Portals

Brainlife: “providing an online, community-based platform where users can publish code (Apps) and Data while integrating HPC and cloud-computing resources to run Apps.”
Use Case - Live Data Processing and Distribution

Unidata:
“sharing geoscience data and the tools to access and visualize that data.”
How to Use Jetstream2

● We have three Dashboards
  ○ Exosphere - The Primary dashboard. Most user friendly. Start here!
    ■ Stay for the interactive tutorial to see this in action.
  ○ Horizon - The default Openstack dashboard. For advanced users. More tools but not as user friendly.
  ○ CACAO - Still in active development but available to use. For advanced users. Useful for building clusters.

● Use Dashboard to…
  ○ Create, Shelve, and Delete Instances
  ○ Create and Attach/Remove Volumes
  ○ Access your instance via Web Shell or Desktop (Exosphere only)

● Command Line Interface - for programmatic interaction
Data Storage

● Jetstream2 is NOT primarily a Storage Device
  ○ Treat storage as a temporary location to stage data for analysis.
  ○ Default of 1 TB storage.
● Storage Options
  ○ Volumes - The default storage type
    ■ Attach to an instance to read and write your data from your programs.
  ○ Manilla - Shared file systems
    ■ accessible by multiple instances
  ○ Object Store - ‘bucket’ to store large amounts of data without a filesystem
    ■ S3 compatible

*Data that must be protected by Federal security or privacy laws (e.g., HIPAA, FERPA, ITAR, classified information, export control, etc.) is NOT permitted to be stored on Jetstream2 unless such storage and usage is specifically authorized by the responsible University administrator and complies with any processes for management of access to such information.
How to Access Jetstream2

Allocations are managed by ACCESS
What is Access?

Provides Access to National Advanced Computing Resources

https://access-ci.org/

The NSF’s ACCESS (Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support) program builds upon the successes of the 11-year XSEDE project, while also expanding the ecosystem with capabilities for new modes of research and further democratizing participation.
Allocation Process

CREATE ACCOUNT

SELECT OPPORTUNITY

REQUEST ALLOCATION

RECEIVE CREDITS

EXCHANGE CREDITS
Getting an ACCESS ID

In order to Apply for an allocation OR be able to login to any resources an Access ID is required.

Go To https://identity.access-ci.org/new-user
Getting Your Own Allocation

https://access-ci.org/

Applying for an allocation is like applying for a small grant that provides computational resources.

The first level of application “Explore ACCESS” is a single form and only takes a few minutes. Allocations are usually approved in a few days. ~$40,000 worth of cloud resources for a few minutes of effort.
## Allocation Tiers

<table>
<thead>
<tr>
<th>Explore</th>
<th>Discover</th>
<th>Accelerate</th>
<th>Maximize</th>
</tr>
</thead>
<tbody>
<tr>
<td>400,000 credits</td>
<td>1,500,000 credits</td>
<td>3,000,000 credits</td>
<td>N/A</td>
</tr>
<tr>
<td>CV and Brief Summary</td>
<td>CV &amp; 1-page proposal</td>
<td>3-page proposal &amp; panel review</td>
<td>10 page proposal &amp; panel review</td>
</tr>
<tr>
<td>Grad student projects</td>
<td>Modest resource needs</td>
<td>Mid-scale resource needs</td>
<td>Large-scale projects in need of massive amounts of computational resources</td>
</tr>
<tr>
<td>Small labs</td>
<td>Mid-sized projects</td>
<td>Multi-grant collaborative projects</td>
<td></td>
</tr>
<tr>
<td>Benchmarking</td>
<td>Campus Champions</td>
<td>Science Gateways with growing communities</td>
<td></td>
</tr>
<tr>
<td>Small classes</td>
<td>Large classes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Allocation Details

- Allocation must be submitted by the PI of the project
- Who can be a PI?
  - Any educator or research at a U.S. academic, non-profit research, or educational institution
- What about grad students?
  - Yes! Can receive Explore tier, just need a letter of collaboration from your PI.
- Anyone with an ACCESS ID can be added to a project
  - Getting an ACCESS ID is free and easy - useful for adding students
- Allocations must be renewed annually
- Projects can be “upgraded” by applying for the next-level allocation tier
Allocations Are Awarded Credits

- Credits are exchanged for computational time or storage on supercomputers
- For Jetstream2, 1 ACCESS credit = 1 Service Unit (SU) = 1 CPU hour
- What are some common usage scenarios?
  - Running the equivalent of a high-end laptop (8 cores, 32 GB RAM) for 40 hours a week for 2 years
    - 35,000 SUs
  - A server that can handle moderate traffic for 5 years uninterrupted
    - 175,000 SUs
  - Using an entire A100 GPU on average 40 hours a week for 2 years
    - 530,000 SUs
    - Note: We have smaller GPU instances that aren’t as costly
### What Resources are Available?

<table>
<thead>
<tr>
<th>Resource</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana University (IU)</td>
<td>Jetstream2</td>
</tr>
<tr>
<td>Johns Hopkins University (JHU)</td>
<td>Rockfish</td>
</tr>
<tr>
<td>National Center for Supercomputing</td>
<td>Applications (NCSA)</td>
</tr>
<tr>
<td>Open Science Grid</td>
<td>Open Storage Network (OSN)</td>
</tr>
<tr>
<td>Pittsburgh Supercomputing Center (PSC)</td>
<td>Bridges-2</td>
</tr>
<tr>
<td>Purdue University</td>
<td>Anvil</td>
</tr>
<tr>
<td>San Diego Supercomputer Center (SDSC)</td>
<td>Expanse</td>
</tr>
<tr>
<td>Stony Brook University</td>
<td>Ookami</td>
</tr>
<tr>
<td>Texas Advanced Computing Center (TACC)</td>
<td>Stampedede2, Ranch</td>
</tr>
<tr>
<td>Texas A&amp;M University (TAMU)</td>
<td>FASTER</td>
</tr>
<tr>
<td>University of Delaware</td>
<td>DARWIN</td>
</tr>
<tr>
<td>University of Kentucky</td>
<td>KYRIC</td>
</tr>
<tr>
<td>ACCESS</td>
<td>Advancing Innovation</td>
</tr>
</tbody>
</table>
Accessing Jetstream2

Go To https://jetstream2.exosphere.app/

Before you can use Jetstream2 you must first have an ACCESS ID and be on an active Jetstream2 allocation!
Documentation and Support

Jetstream2 Website
https://jetstream-cloud.org/index.html

Reach out to Jetstream2 support at
help@jetstream-cloud.org

Documentation:
https://docs.jetstream-cloud.org/

Or directly to Research Cloud Services:
rcsadm@iu.edu

Jetstream2 Login (Exosphere)
https://jetstream2.exosphere.app/

My email:
djhavert@iu.edu
Demo
Acknowledgements

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

This material is based upon work supported by 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 Jetstream2 contributors

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

And the rest of the Jetstream2 Support team

- Zach Graber, Julian Pistorius, Jenn Taylor and Le Mai Weakley.
Jetstream2 partners

ASU Arizona State University
THE UNIVERSITY OF ARIZONA

THE UNIVERSITY OF ARIZONA

UNIVERSITY OF HAWAII

THE UNIVERSITY OF ARIZONA

Johns Hopkins University
TACC

UCAR

http://jetstream-cloud.org/
National Science Foundation
Award #ACI-2005506