Jetstream2: Accelerating Science and Engineering on Demand

David Y. Hancock – Indiana University
Director for Advanced Cyberinfrastructure
Jetstream & Jetstream2 Primary Investigator

PEARC20 – July 29, 2020
Panel: Introduction of the new NSF Innovative HPC Systems
What is Jetstream2 and why does it exist?

- Significant evolution of the Jetstream cloud resource
- Under 10% NSF investment → support for 24% of institutions, 23% of active PIs, and 32% of users*
- Jetstream has provided 6x more SUs than any other XSEDE resource for Education
- Emphasis on ease-of-use, broad accessibility, *AI for Everyone*
- Will provide on-demand interactive computing and persistent services for science gateways
- Enables *configurable* environments; *programmable cyberinfrastructure*

*Based on XDMoD data at Workload Analysis Report: http://arxiv.org/abs/1801.04306
Jetstream2 Capabilities

Enhancing IaaS model of Jetstream:
• Improved orchestration support
• Elastic virtual clusters
• Federated JupyterHubs

Commitment to >99% uptime
• Critical for science gateway hosting
• Hybrid-cloud support

Revamped User Interface
• Unified instance management
• Multi-instance launch

>57K cores of next-gen AMD EPYC processors
>360 NVIDIA A100 GPUs will provide vGPUs via NVIDIA’s MIG feature
>18PB of storage (NVMe and disk hybrid)
100GbE Mellanox network
XSEDEnet:
Advanced Layer 2 Services (AL2S) platform

**Primary**
- **Compute**
  - 416 Nodes
  - 53,248 Cores
  - 224 TB RAM

- **Storage**
  - 96 Nodes
  - 15 PB

**Regional**
- **Compute**
  - 8 Nodes
  - 1,024 Cores
  - 4 TB RAM

- **Storage**
  - 869 TB

**Accelarators**
- **Compute**
  - 2 Nodes
  - 1 TB RAM
  - 8 GPUs

- **Storage**
  - 869 TB

**Cornell University Cyberinfrastructure**
- **Regional Compute**
  - 8 Nodes
  - 1,024 Cores
  - 4 TB RAM

- **Storage**
  - 869 TB
Conceptual Jetstream2 Architecture

Users/Gateways:
- Atmosphere user interface
- Horizon web dashboard
- Atmosphere & OpenStack command line interfaces
- Third party applications
- Language binding

OpenStack:
- Compute Service
- Storage Service (CEPH)
- Network Service

Authorization:
- Compute node 500+ Compute node

Secondary:
- Orchestration
- Containers
- Databases
- Bare metal
- Data processing
Platform Overview

- Atmosphere
- XSEDE Accounting
- OpenStack
- Ceph
- Jetstream2 Core System

OpenStack

- Horizon
- OpenStack CLI

Authentication Service

XSEDE Accounting

3rd Party Clients

Science Gateways

Jetstream2 Core System
Timeline

• Jetstream now in 5th year of operations
• Jetstream extension requested through November 2021
• Jetstream2
  • Early operations planned for August 2021
  • Production operations by October 2021

Flickr user Oiluj Samall Zeid - Lejos de Yulín
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
• Jeremy Fischer, J. Michael Lowe, Therese Miller, Maria Morris, Winona Snapp-Childs, and George Turner