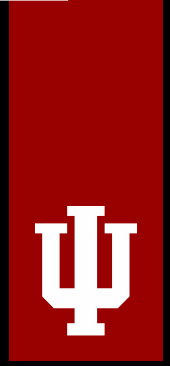


Open Science Forum, April 26, 2017

IVMOOC on Jetstream & Open XD Metrics on Demand Value Analytics

CNS, IUNI, UITS

INDIANA UNIVERSITY BLOOMINGTON

OPEN SCIENCE FORUM

Open Science Forum

Wednesday, Apr 26th 2017 at 4:00 PM
SSRC Grand Hall (Woodburn Hall 200)

Jetstream, funded by NSF and led by the Indiana University Pervasive Technology Institute (PTI), adds cloud-based computation to the national cyberinfrastructure (https://jetstream-cloud.org). IUNI collaborated with the Cyberinfrastructure for Network Science Center to deploy the Network Workbench tool on Jetstream for use by IVMOOC students (http://ivmooc.cns.iu.edu).

The "**Open XD Metrics on Demand Value Analytics**" NSF project, directed by Matthew Link (UITS), aims to measure and visualize the impact of campus-financed cyberinfrastructure (CI) on progress in science. Resulting visual analytics modules will be added to the existing CI metrics tool eXtreme Data Metrics on Demand (XDMoD) to present a view of financial, collaboration, and publication data, showing "return-on-investment" metrics in relation to CI usage.

New Ventures in Research, Engineering, and Educational Computing.

George Turner, Chief Systems Architect
Research Technologies, UITS, Indiana University

Open Science Forum, SSRC Grand Hall
Indiana University Bloomington, IN
26 April 2017

funded by the National Science Foundation
Award #ACI-1445604

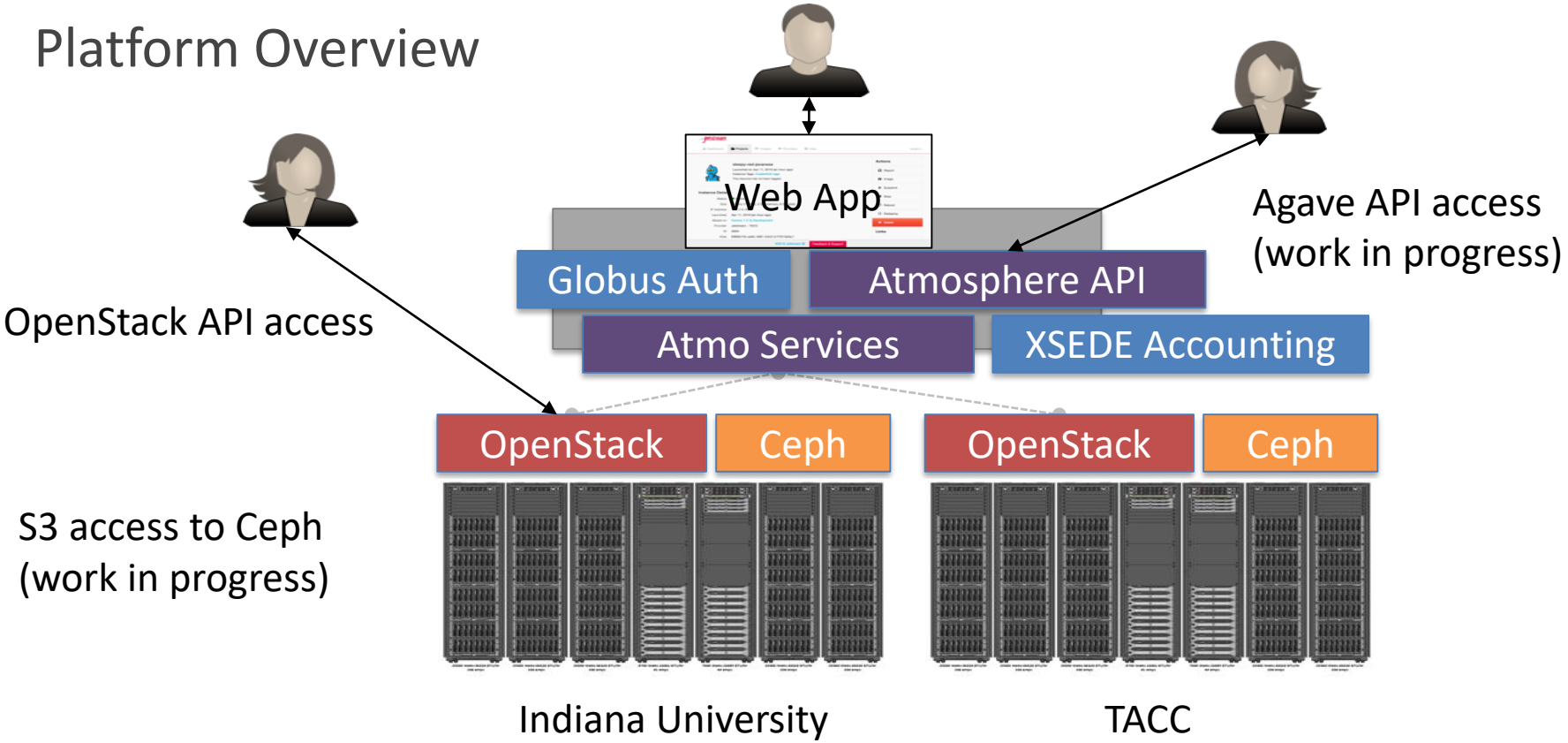
What is Jetstream?

- **User-friendly**, widely accessible cloud environment
    - **User-selectable library** of preconfigured virtual machines; no need for system administration skills.
    - **Programmatic API access** to implement modern cloud computing techniques

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Platform Overview



Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

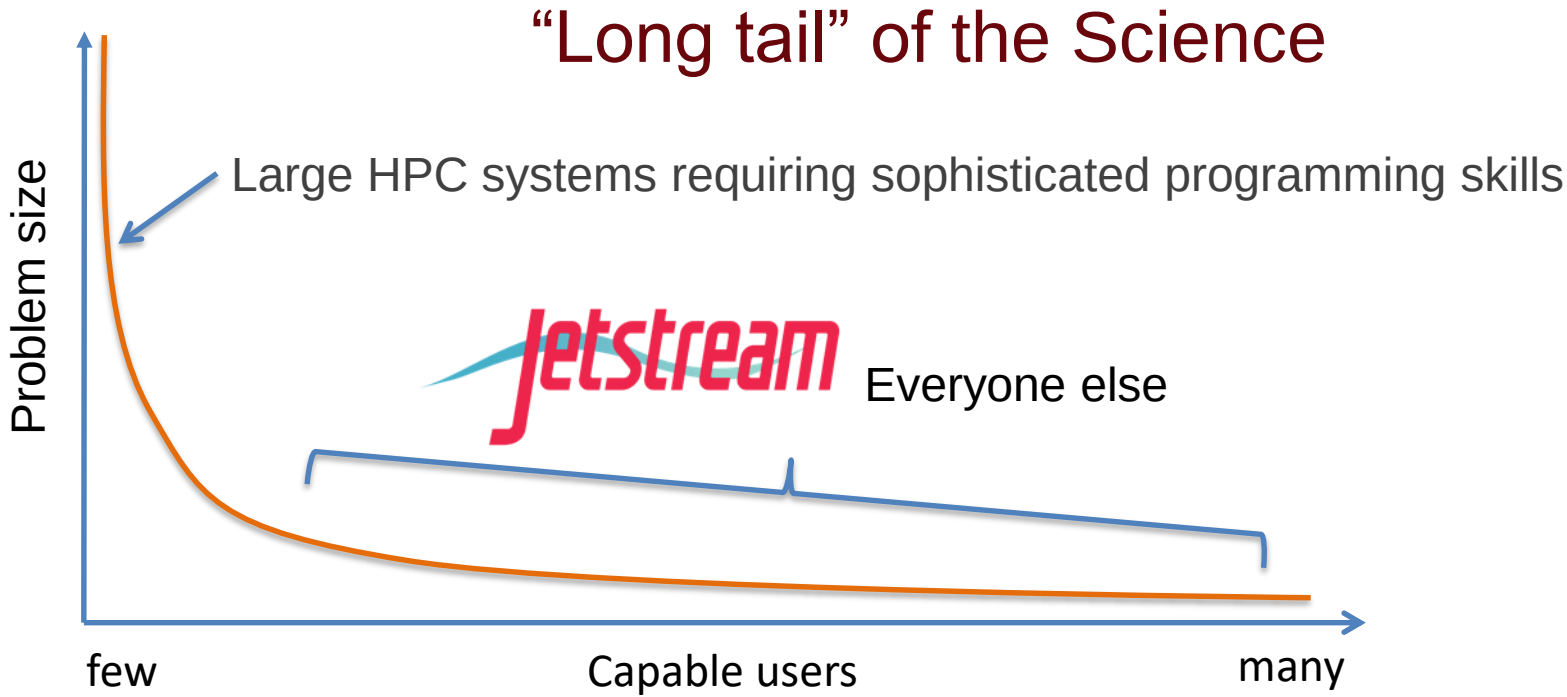
What is Jetstream?

- **Reproducibility**: store, publish via IU Scholarworks (DOI)
- **Cloudy**: clouds are more the just virtual machines (VM)
  - Old way: robust (expensive) infrastructure, weak (cheap) software
  - Cloudy way: commodity infrastructure, robust software
  - Cows, not pets : pets take great amount of care, feeding, and you name them; cows you intend to have high turnover and you give them numbers.
- **Primary goal** is to **expand the user base** of NSF's eXtreme Digital (XD) program resources beyond the current community of users.

Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

What is Jetstream? (cont)



Jetstream
http://jetstream-cloud.org/

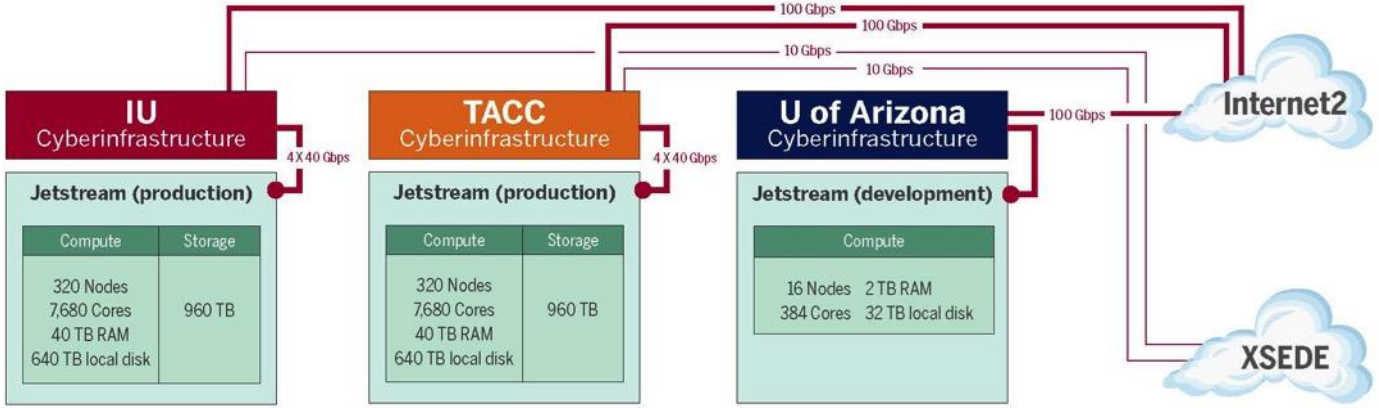
funded by the National Science Foundation
Award #ACI-1445604

What is Jetstream? (cont)

- **Software layers**
  - **Atmosphere** web interface
    - library of images, genertic, domain specific
    - simplify VM administration
  - **Openstack:** software tools for building and managing cloud computing platforms for public and private clouds.
  - **KVM** hypervisor: what the VMs run on
  - **Ceph**: storage platform that stores data on a single distributed computer cluster, and provides interfaces for **object**-, **block**- and *file-level* storage.
  - **Operating systems**: CentOS, Ubuntu, Windows?
  - **Applications**; e.g. software developed by the domain specialist, gateways, etc.

funded by the National Science Foundation
Award #ACI-1445604

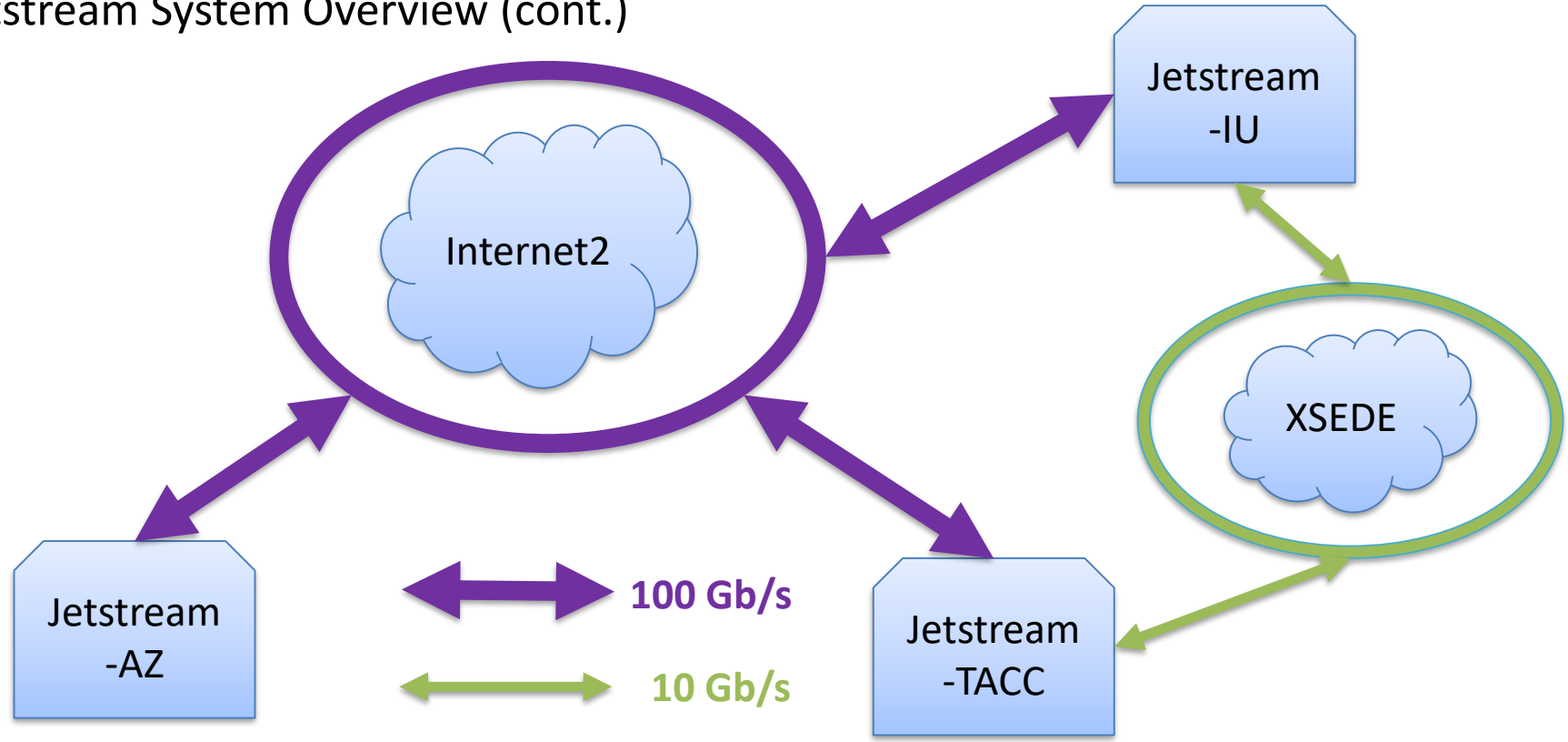
Jetstream System Overview



Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Jetstream System Overview (cont.)



Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Production Cloud Hardware (per site)

Hardware	Number	Specifications	Function (IU)
Dell PowerEdge M630 blades	320	2X Intel E5-2680v3 “Haswell” 24 cores @ 2.5 GHz 128 GB RAM 2 TB local disk	Compute hosts OpenStack services
Dell PowerEdge R630 1U server	7	2X Intel E5-2680v3 “Haswell” 24 cores @ 2.5 GHz 128 GB RAM 2 TB local disk	Cluster management High Availability Databases RabbitMQ
Dell PowerEdge R730xd 2U servers	20	2X Intel E5-2680v3 “Haswell” 24 cores @ 2.5 GHz 64 GB RAM 48 TB storage for Ceph pool	~1 PB Ceph storage
Dell S6000-ON network switches	9	32+2 40 Gb/s ports	Top of Rack Spine

Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

VM Instance Sizes (Flavors)

Instance Type	vCPUs	RAM(GB)	Storage(GB)	Instances/Node
Tiny	1	2	8	46
Small	2	4	20	23
Medium	6	16	60	7
Large	10	30	120/60*	4
X-Large	22	60	240/60*	2
XX-Large	44	120	480/60*	1

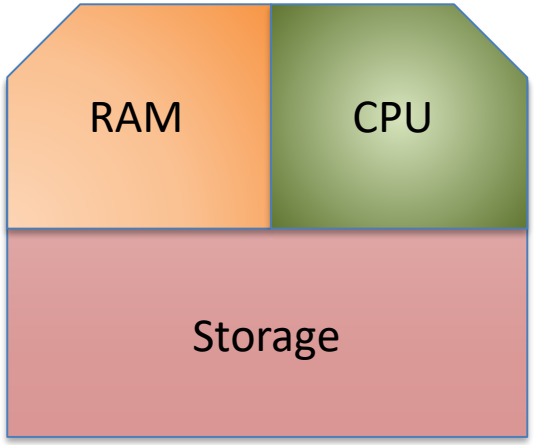
Node config: dual Intel E-2680v3 “Haswell”, 24 physical cores/node @ 2.5 GHz, 128 GB RAM, dual 1 TB local disks.

* Effective 29-Mar-2017

Jetstream
http://jetstream-cloud.org/

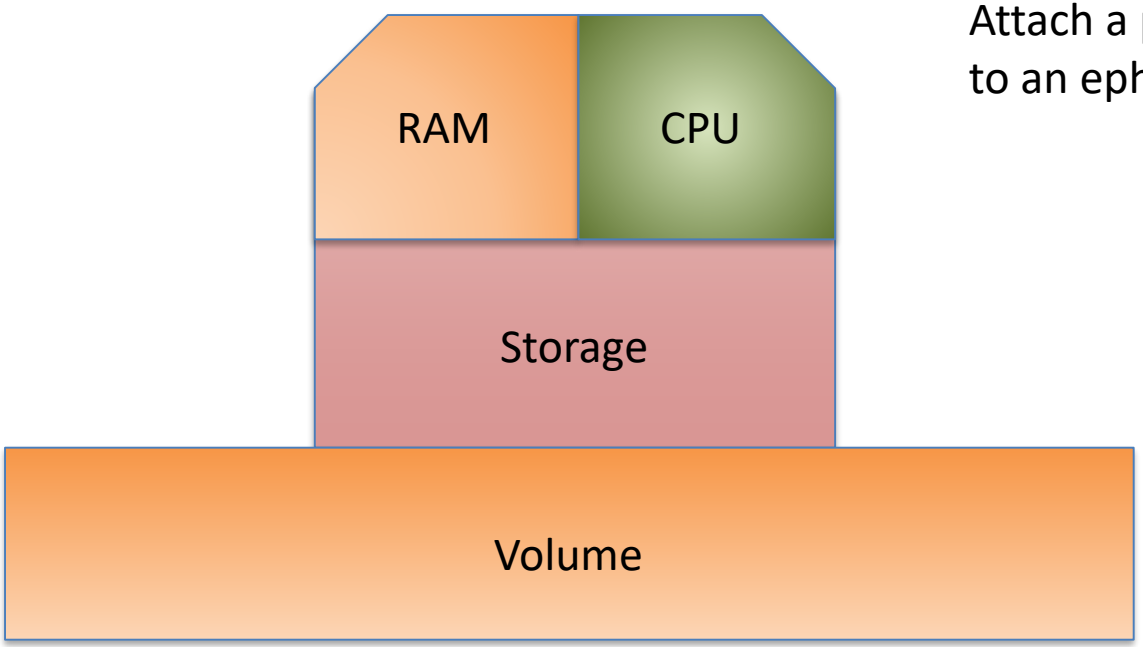
funded by the National Science Foundation
Award #ACI-1445604

What is an Instance



NSF
funded by the National Science Foundation
Award #ACI-1445604

Instance & Volumes

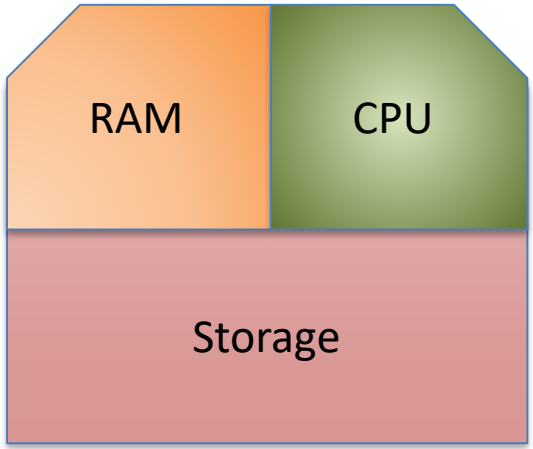


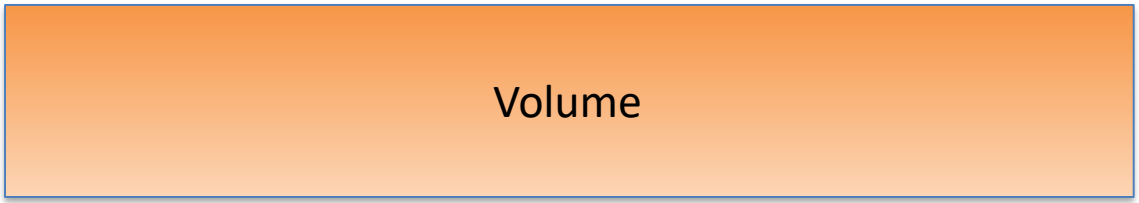
Attach a persistent volume to an ephemeral instance

Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

Instance & Volumes





Dettach the persistent volume to an ephemeral instance

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Instance & Volumes

The instance is gone
but the volume persists

Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

Terms

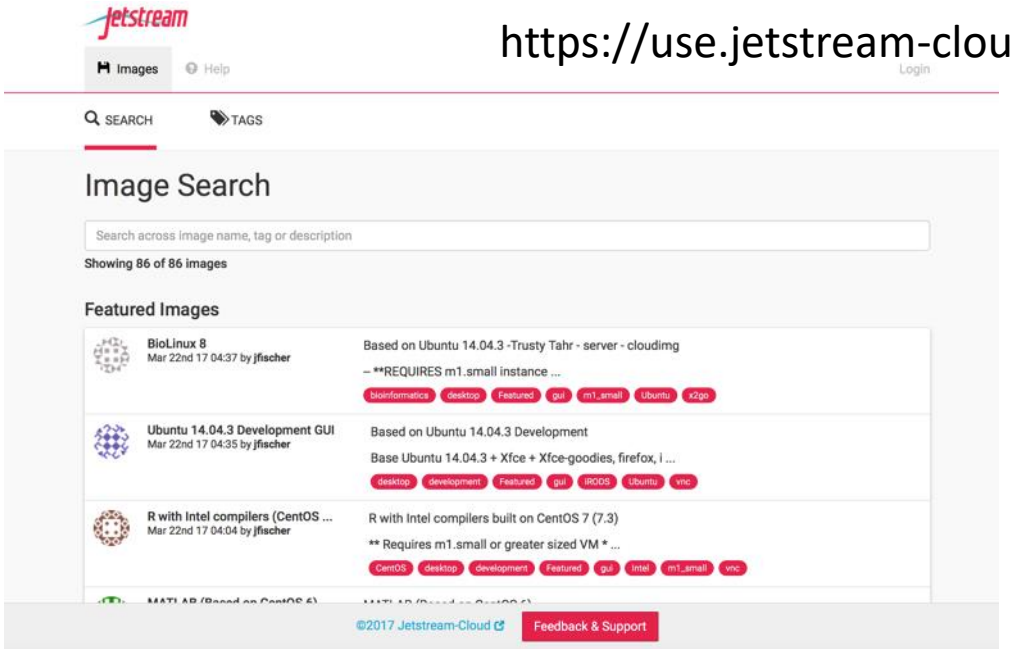
- Image: a file on a storage device
- Instance: a running
- Volumes: persistent storage device
- Flavors: the size of instance, #cpu, amt of RAM, amt of storage
- Host: hardware
- Guest:

NSF
funded by the National Science Foundation
Award #ACI-1445604

Jetstream's Atmosphere Interface

(no login required at this point)

https://use.jetstream-cloud.org/

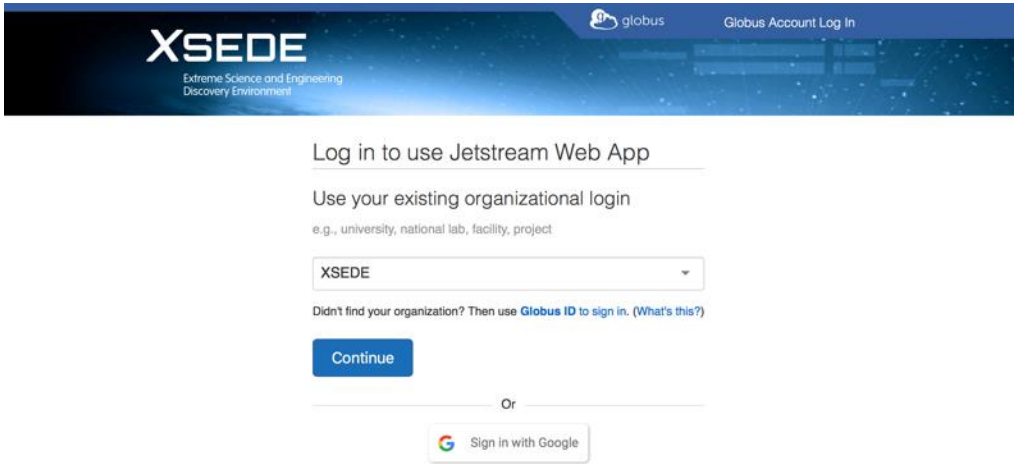


Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Jetstream's Atmosphere Interface

(Pick identity provider)

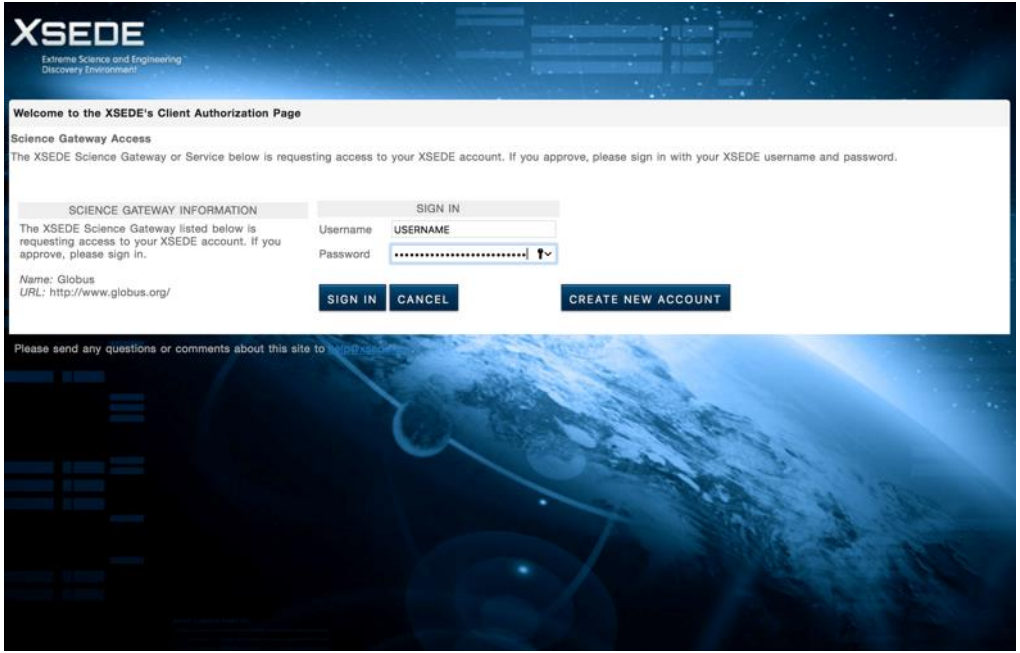


Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Jetstream's Atmosphere Interface

(Authenticate)

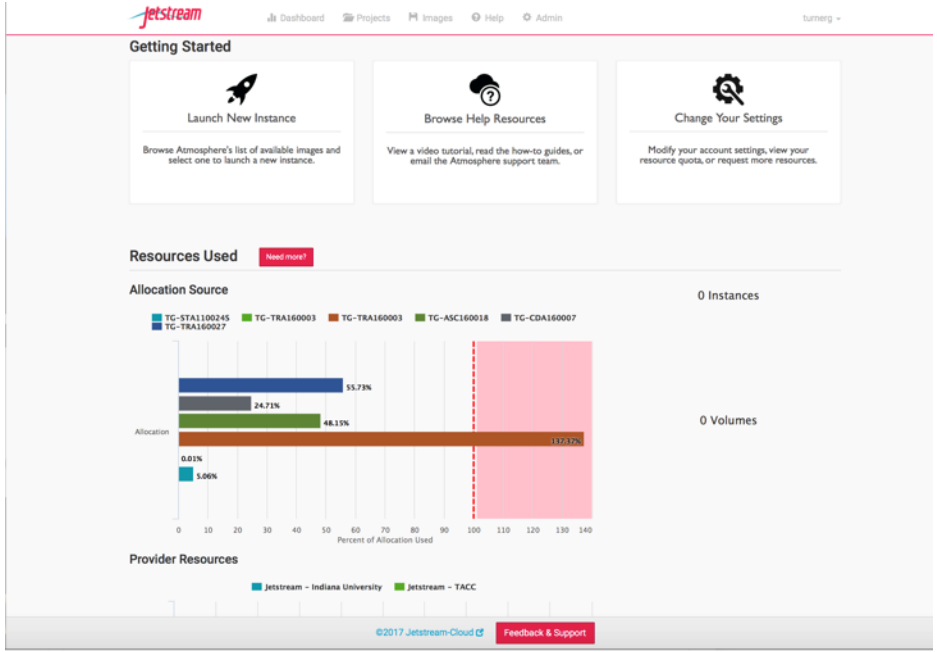


Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

Jetstream's Atmosphere Interface

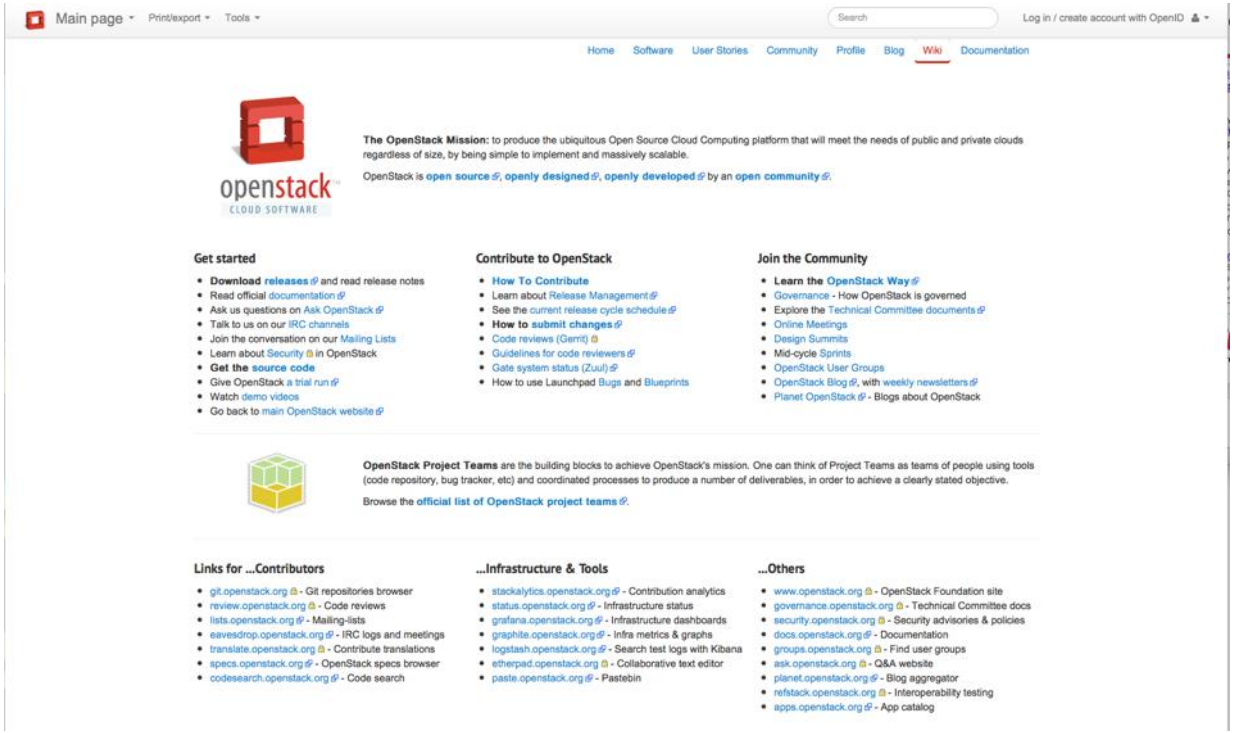
(user's home space)



Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

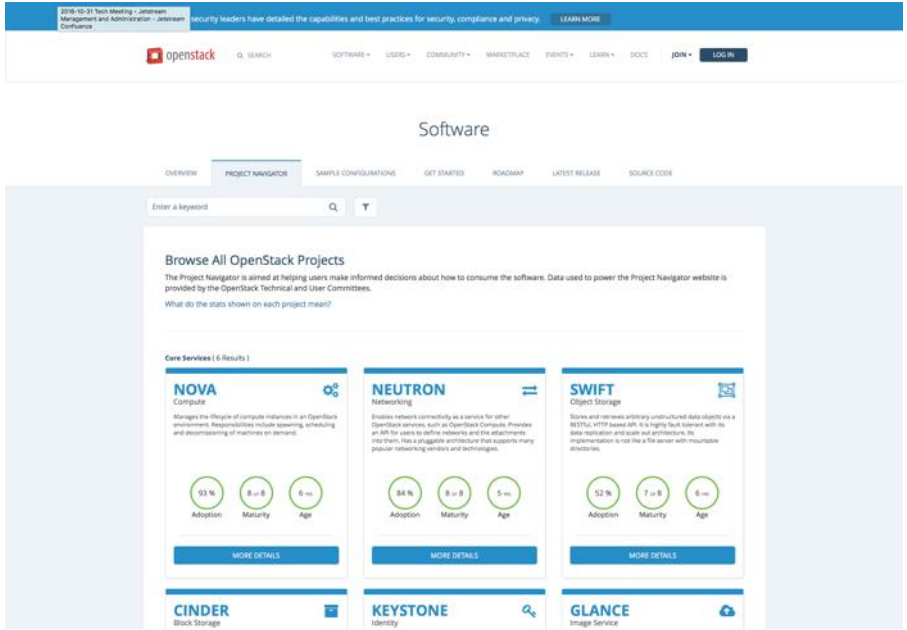
OpenStack Organization



Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

OpenStack Projects



http://www.openstack.org/software/project-navigator/

Jetstream
http://jetstream-cloud.org/

NSF

funded by the National Science Foundation
Award #ACI-1445604

Openstack Projects …the core services

Service	Name	Adoption	Maturity	Age
Identity	Keystone	96%	7/8	5 yrs
Images	Glance	95%	6/8	7 yrs
Block device	Cinder	88%	7/8	5 yrs
Networking	Neutron	93%	7/8	5 yrs
Compute	Nova	95%	8/8	7 yrs
Object device	Swift	52%	7/8	7 yrs

https://www.openstack.org/software/project-navigator/

Jetstream
http://jetstream-cloud.org/

NSF funded by the National Science Foundation
Award #ACI-1445604

Openstack Projects …some other services

Service	Name	Adoption	Maturity	Age
Dashboard	Horizon	87%	6/8	5 yrs
Telemetry	Ceilometer	55%	1/8	4 yrs
Database	Trove	13%	3/8	3 yrs
Orchestration	Heat	67%	6/8	4 yrs
Provisioning	Ironic	17%	2/8	3 yrs

https://www.openstack.org/software/project-navigator/

Jetstream
http://jetstream-cloud.org/

NSF funded by the National Science Foundation
Award #ACI-1445604

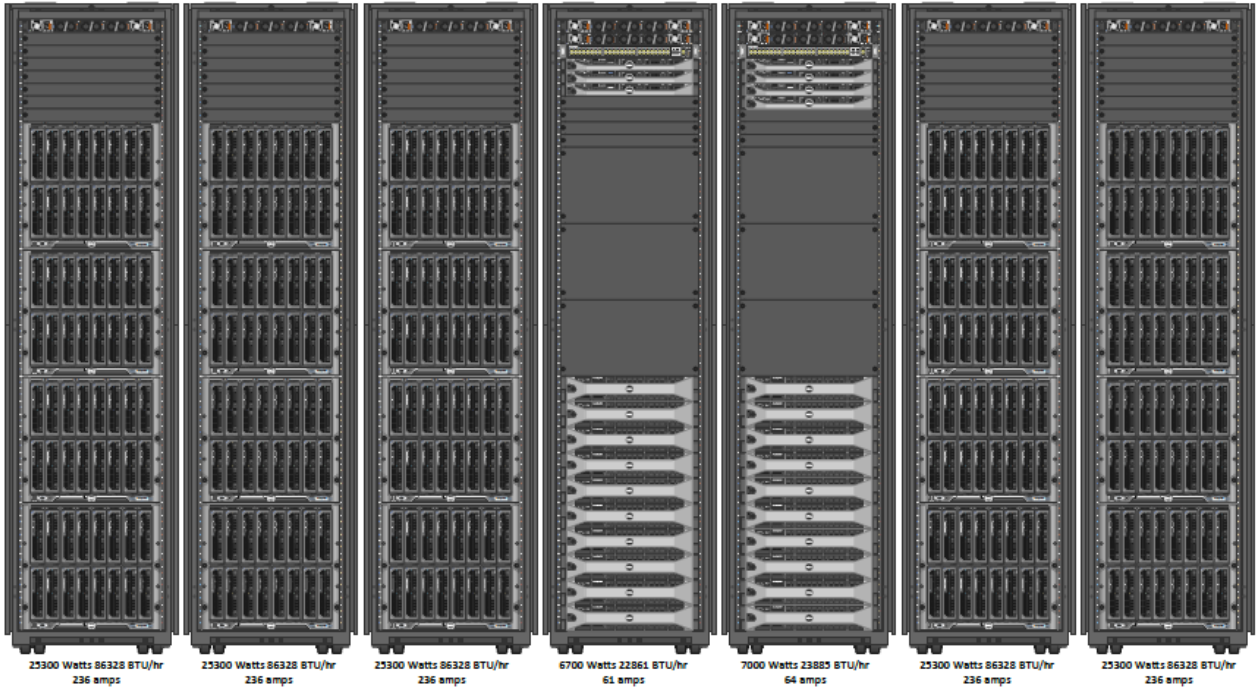
Openstack Projects …some other services

Service	Name	Adoption	Maturity	Age
Map/Reduce	Sahara	10%	3/8	3 yrs
Shared Filesystems	Manila	14%	5/8	3 yrs
DNS Service	Designate	16%	3/8	3 yrs
Containers	Magnum	11%	2/8	2 yrs
Application Catalog	Murano	11%	1/8	2 yrs

https://www.openstack.org/software/project-navigator/

NSF
funded by the National Science Foundation
Award #ACI-1445604

Jetstream Production Hardware



Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

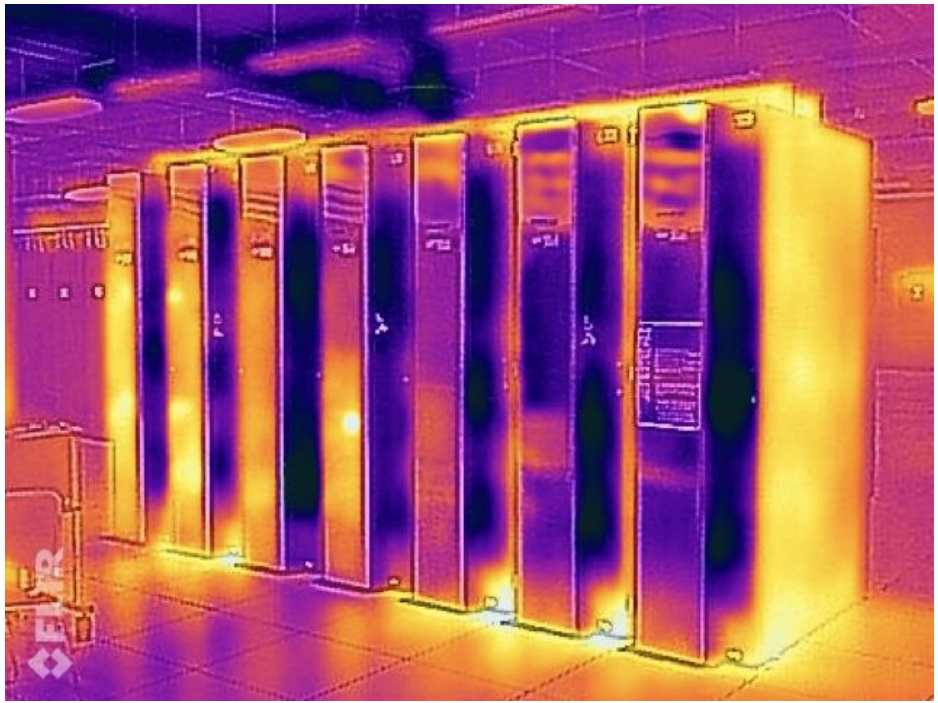
Just for fun: Happy Cluster – Mad Cluster

ChilledDoor Rack Cooling System by motivair

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

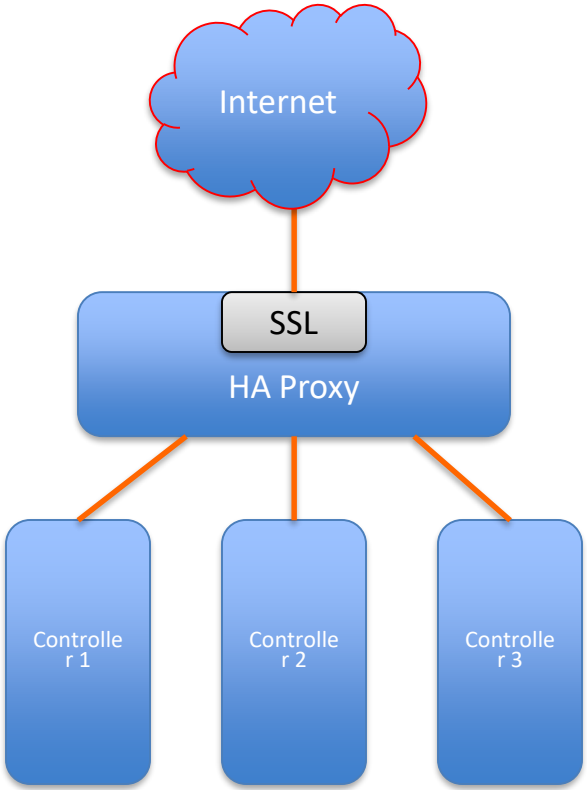
Infrared image of Jetstream



ChilledDoor Rack Cooling System by motivair

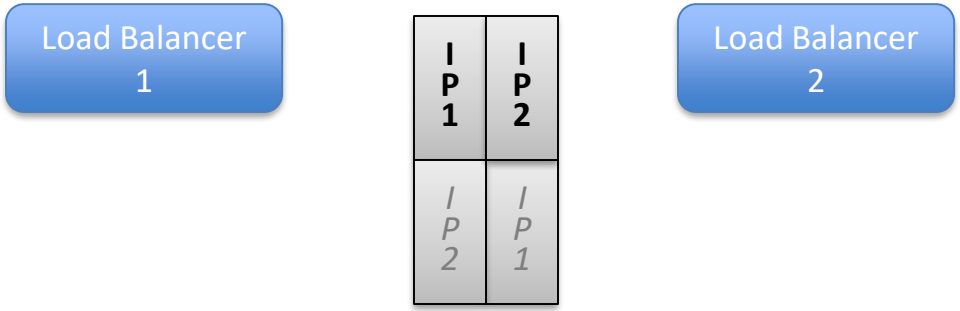
Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604



Jetstream
http://jetstream-cloud.org/

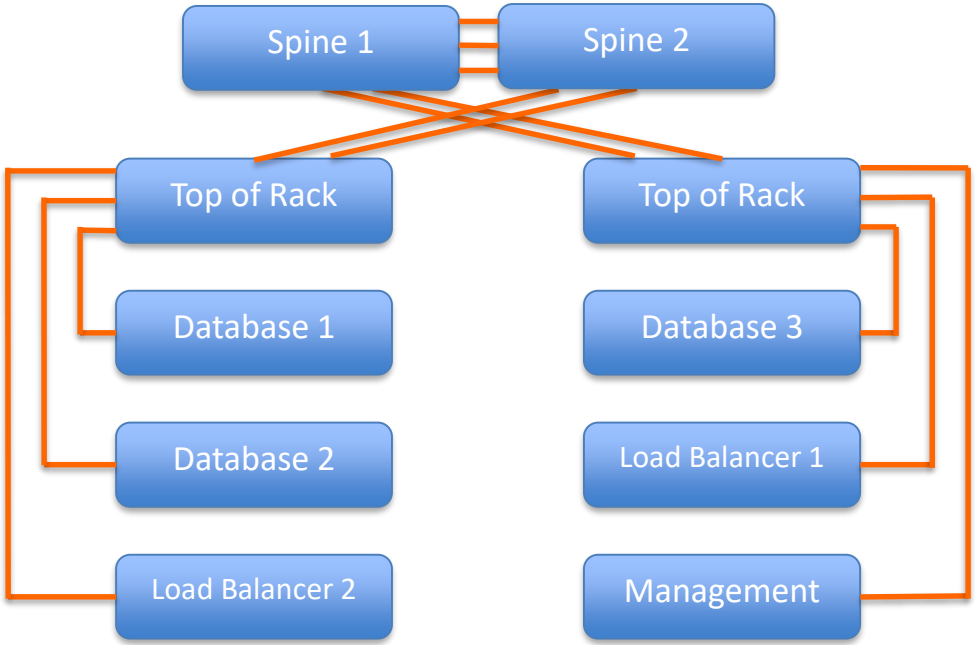
funded by the National Science Foundation
Award #ACI-1445604



http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

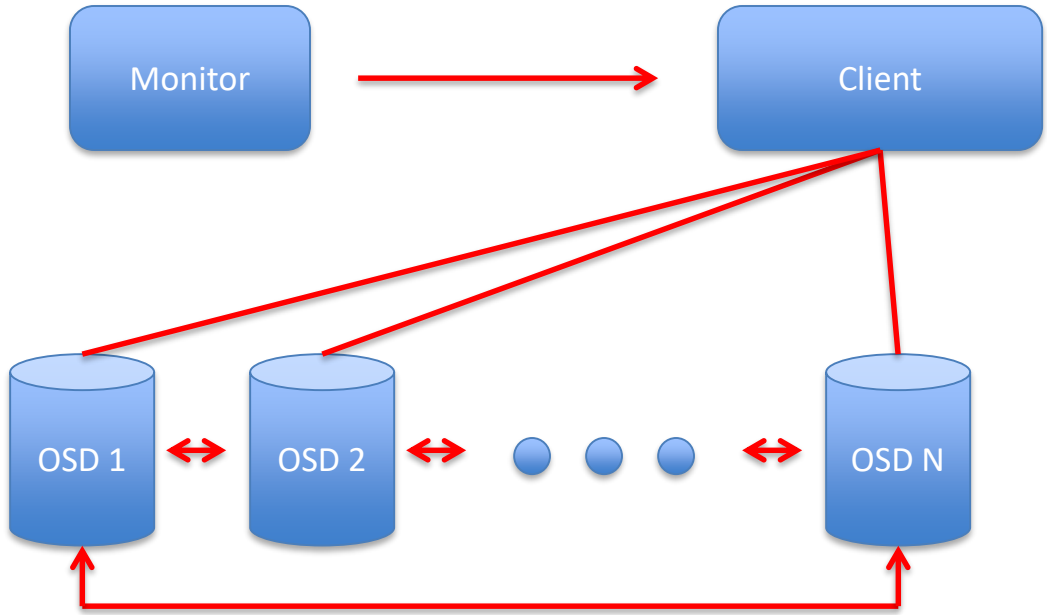
High Availability layout for the databases



Jetstream
http://jetstream-cloud.org/

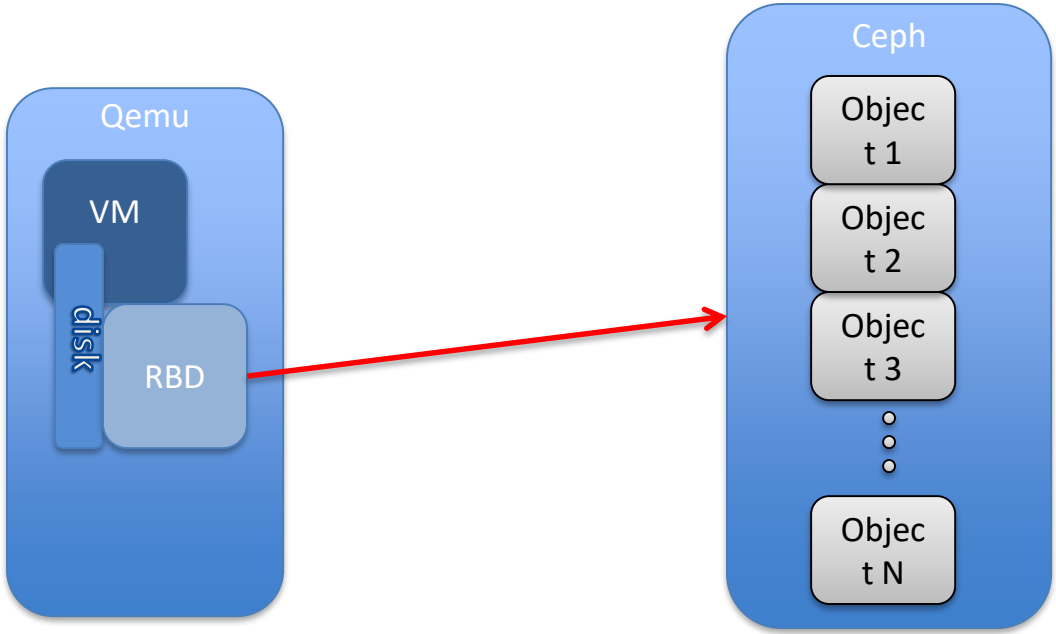
NSF
funded by the National Science Foundation
Award #ACI-1445604

Glance - Cinder - Ceph



Jetstream
http://jetstream-cloud.org/

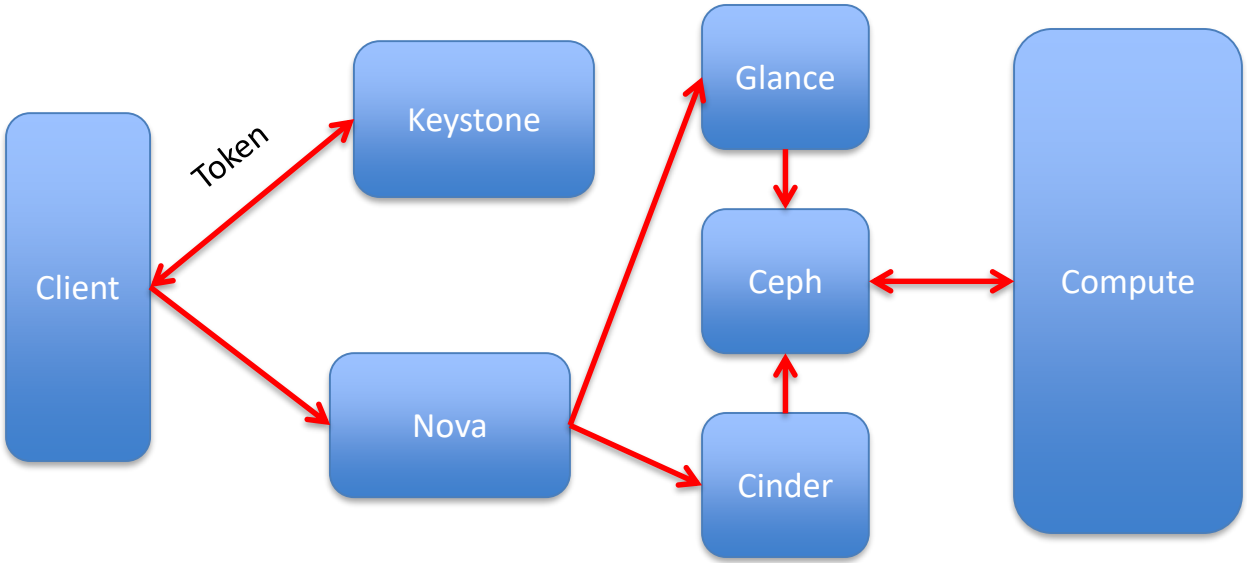
NSF
funded by the National Science Foundation
Award #ACI-1445604



Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

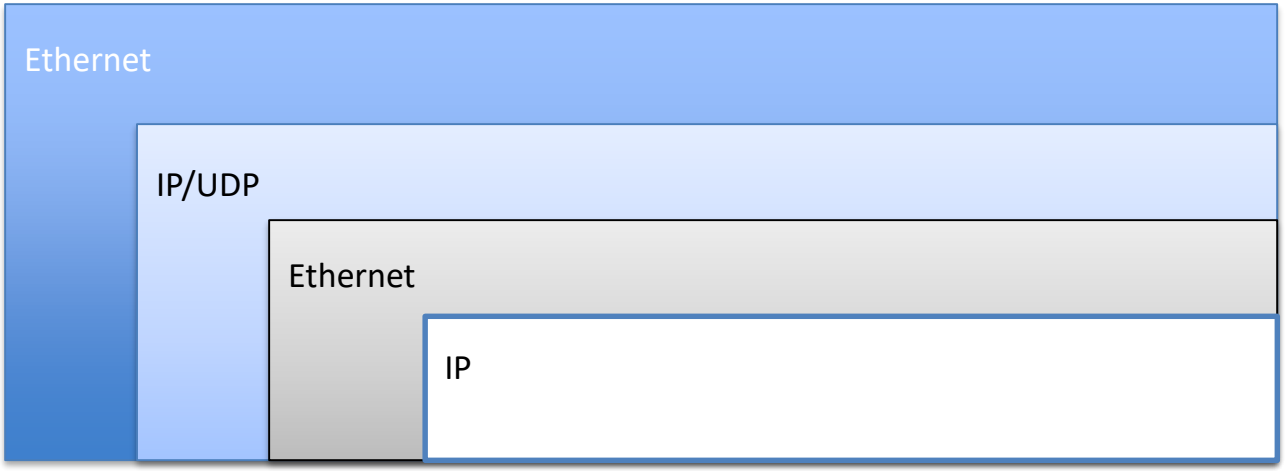
OpenStack Overview



Jetstream
http://jetstream-cloud.org/

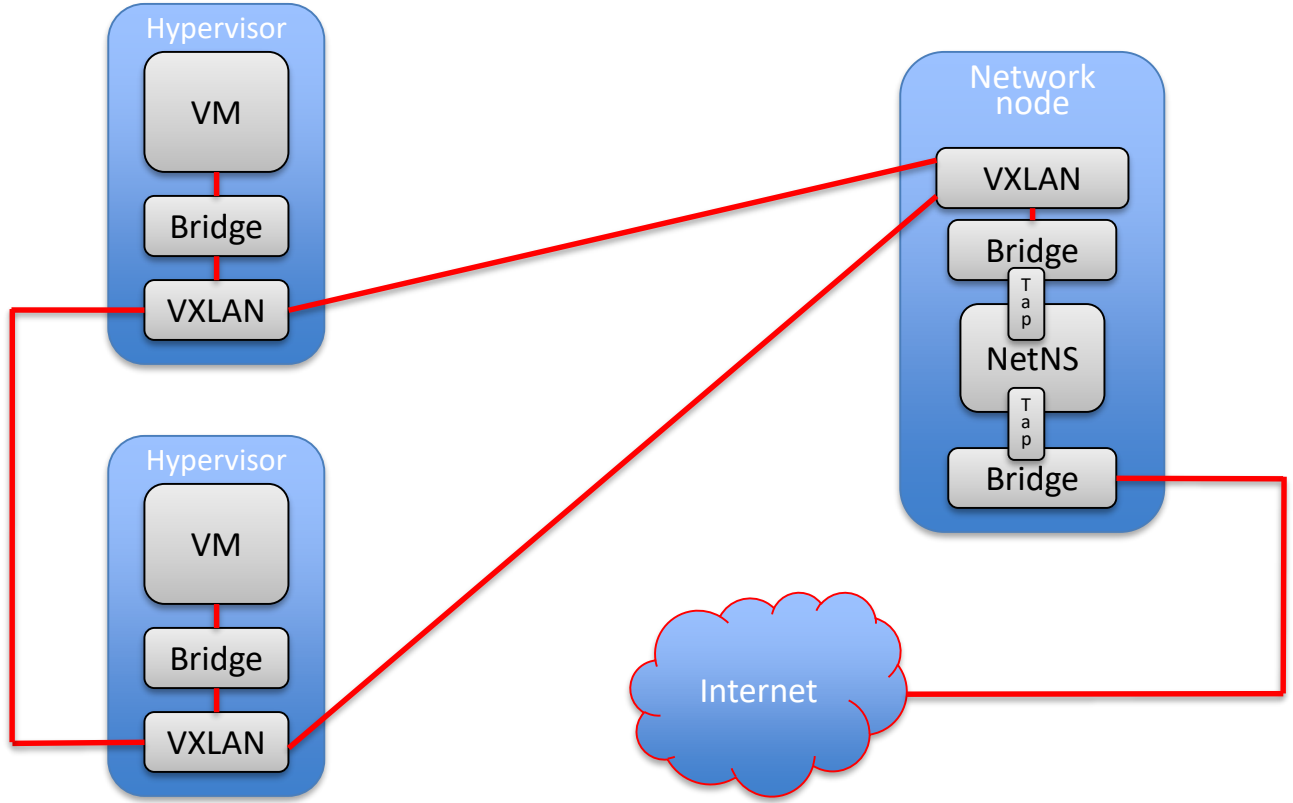
NSF
funded by the National Science Foundation
Award #ACI-1445604

VXLAN Packet



funded by the National Science Foundation
Award #ACI-1445604

Neutron Networking

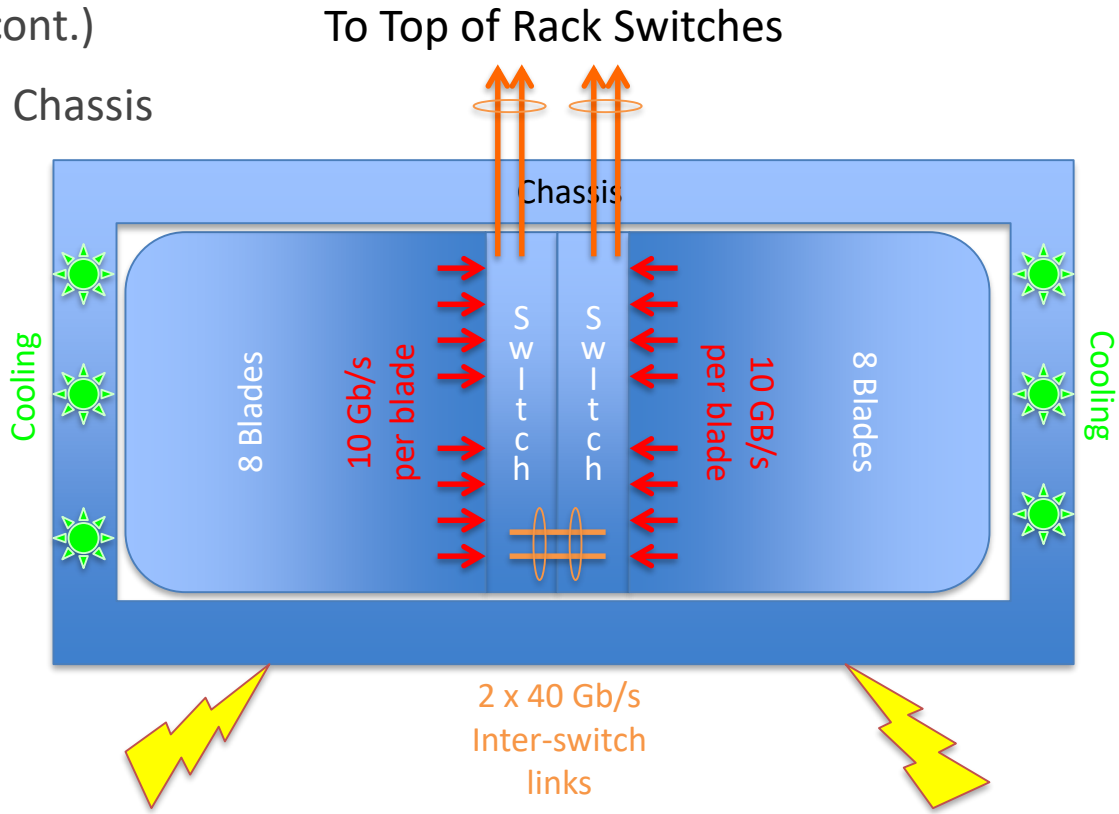


Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Network Topology (cont.)

Sixteen blades per chasses
Two switches per chassis



Jetstream
http://jetstream-cloud.org/

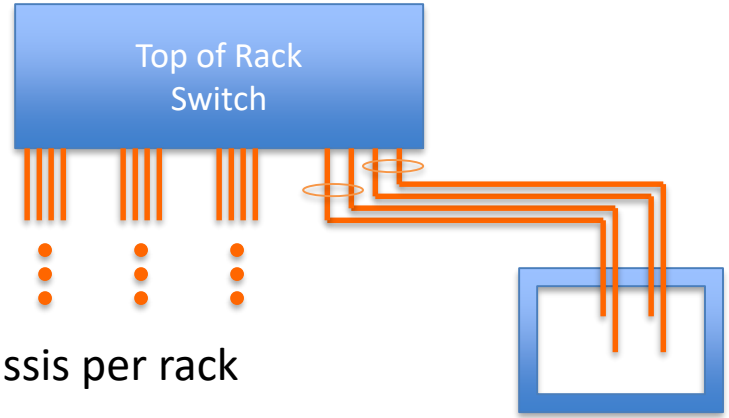
NSF
funded by the National Science Foundation
Award #ACI-1445604

Network Topology (cont.)

Chassis to Top of Rack

40 Gb/s

LAG



Four chassis per rack

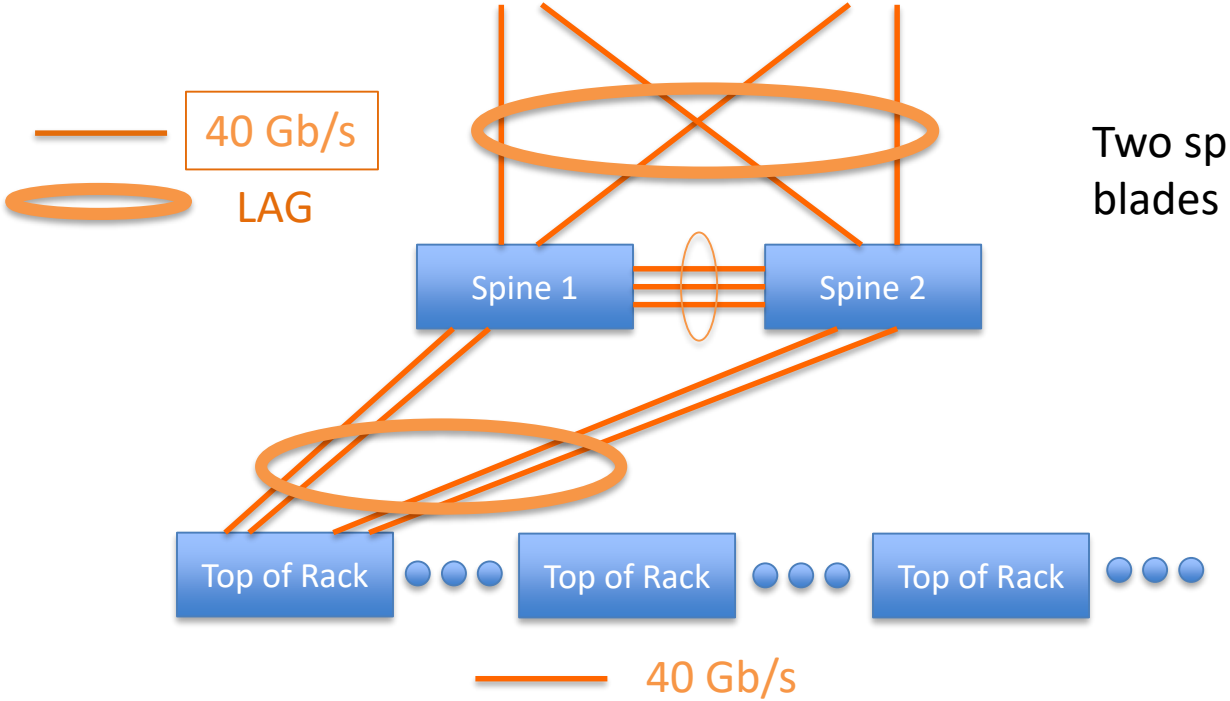
Two switches per chassis

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Network Topology (cont.)

100 Gb/s uplink to Internet2



Two spines tie into two network blades in the datacenter switch

Seven racks tie into the two spine switches

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

How do we onboard users onto Jetstream?

- An XSEDE User Portal (XUP) account is required. They are free! Get one at https://portal.xsede.org
- Read the Allocations Overview - https://portal.xsede.org/allocations-overview
- Write a successful allocation request – start with a Startup or Education request - https://portal.xsede.org/successful-requests

funded by the National Science Foundation
Award #ACI-1445604

Jetstream Information Sources

- Jetstream: https://use.jetstream-cloud.org/
- XSEDE User Portal is required to actually login: https://portal.xsede.org
- User guide: http://jetstream-cloud.org/training.php
- Paper describing Jetstream Jetstream: A self-provisioned, scalable scince and engineering cloud environment
- Configuration management: https://github.com/jetstream-cloud/Jetstream-Salt-States

Jetstream
http://jetstream-cloud.org/

funded by the National Science Foundation
Award #ACI-1445604

Jetstream Partners

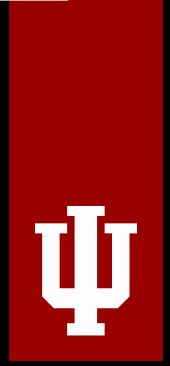
Jetstream
http://jetstream-cloud.org/

NSF
funded by the National Science Foundation
Award #ACI-1445604

Questions?

Project website: http://jetstream-cloud.org/
Project email: jethelp@iu.edu
Direct email: jomlowe@iu.edu, turnerg@iu.edu

NSF
funded by the National Science Foundation
Award #ACI-1445604



Open Science Forum, April 26, 2017

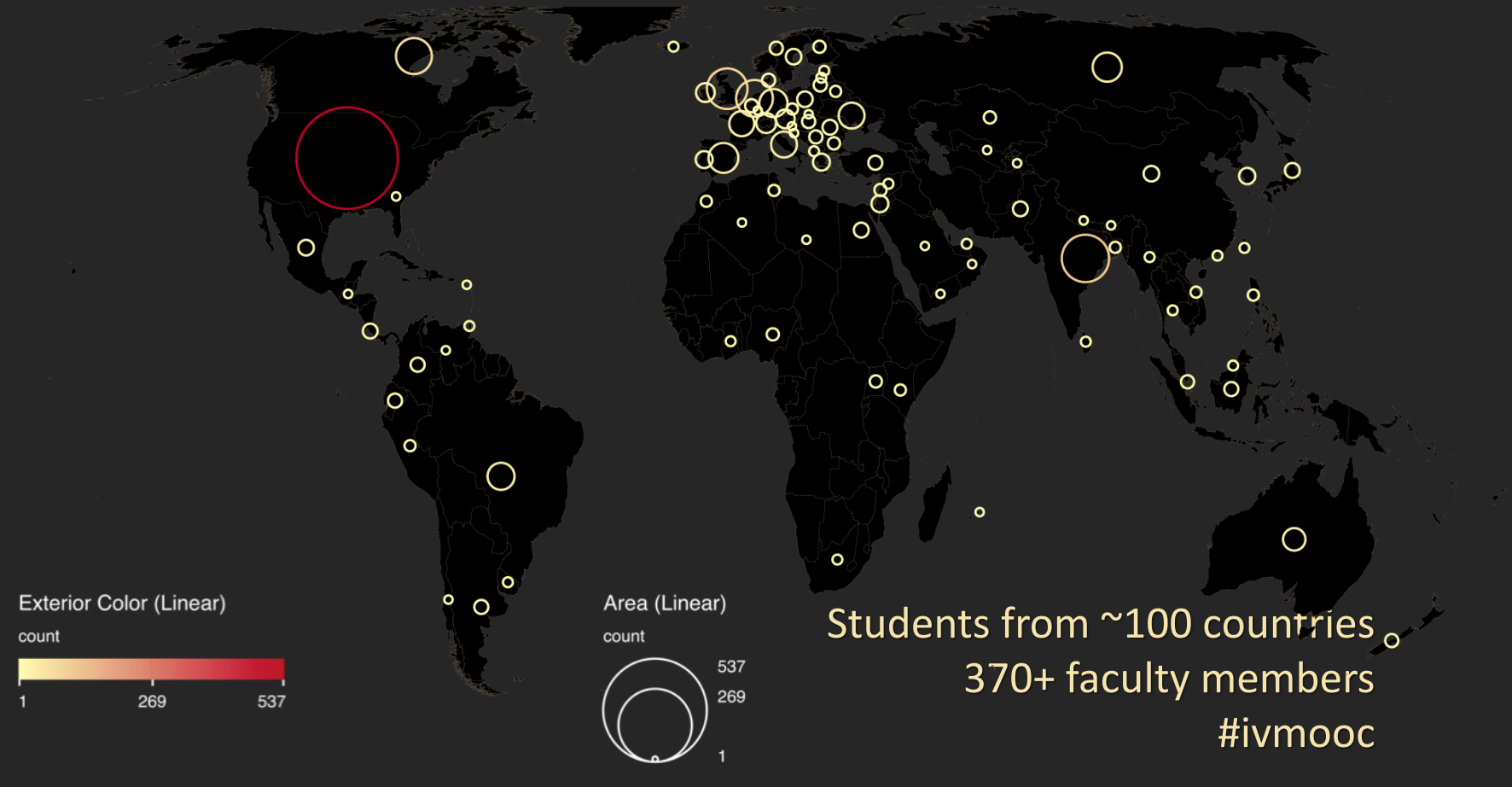
IVMOOC on Jetstream

CNS, IUNI, UITS

INDIANA UNIVERSITY BLOOMINGTON

Register for free: http://ivmooc.cns.iu.edu. Class restarts Jan 09, 2018.

The Information Visualization MOOC
ivmooc.cns.iu.edu



Students from ~100 countries
370+ faculty members
#ivmooc

Part 1: Theory and Hands-On

Session 1 – Workflow Design and Visualization Framework

Session 2 – “When:” Temporal Data

Session 3 – “Where:” Geospatial Data

Session 4 – “What:” Topical Data

Mid-Term

Session 5 – “With Whom:” Trees

Session 6 – “With Whom:” Networks

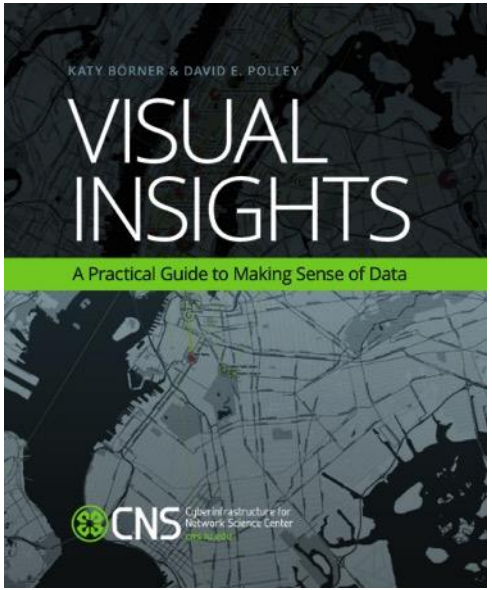
Session 7 – Dynamic Visualizations and Deployment

Final Exam

Part 2: Students work in teams on client projects.

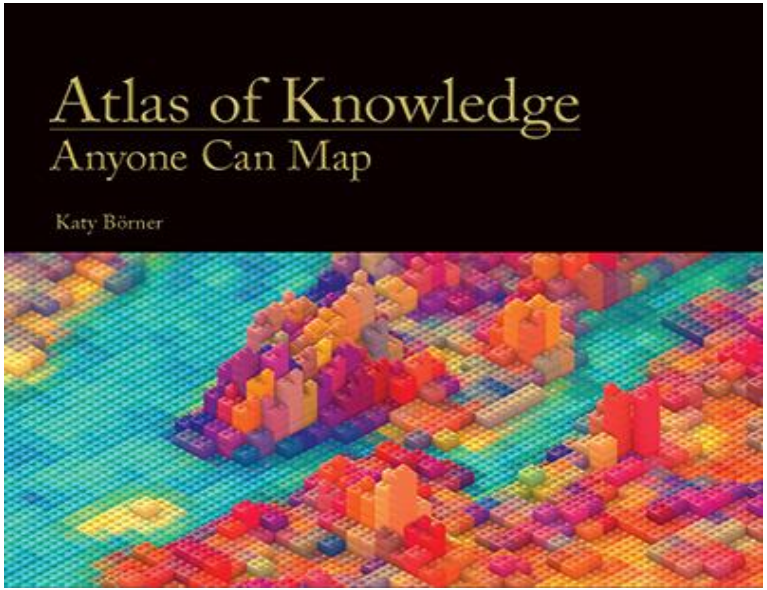
Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%)

Books Used in the IVMOOC



Teaches timely knowledge:

Advanced algorithms, tools, and hands-on workflows.

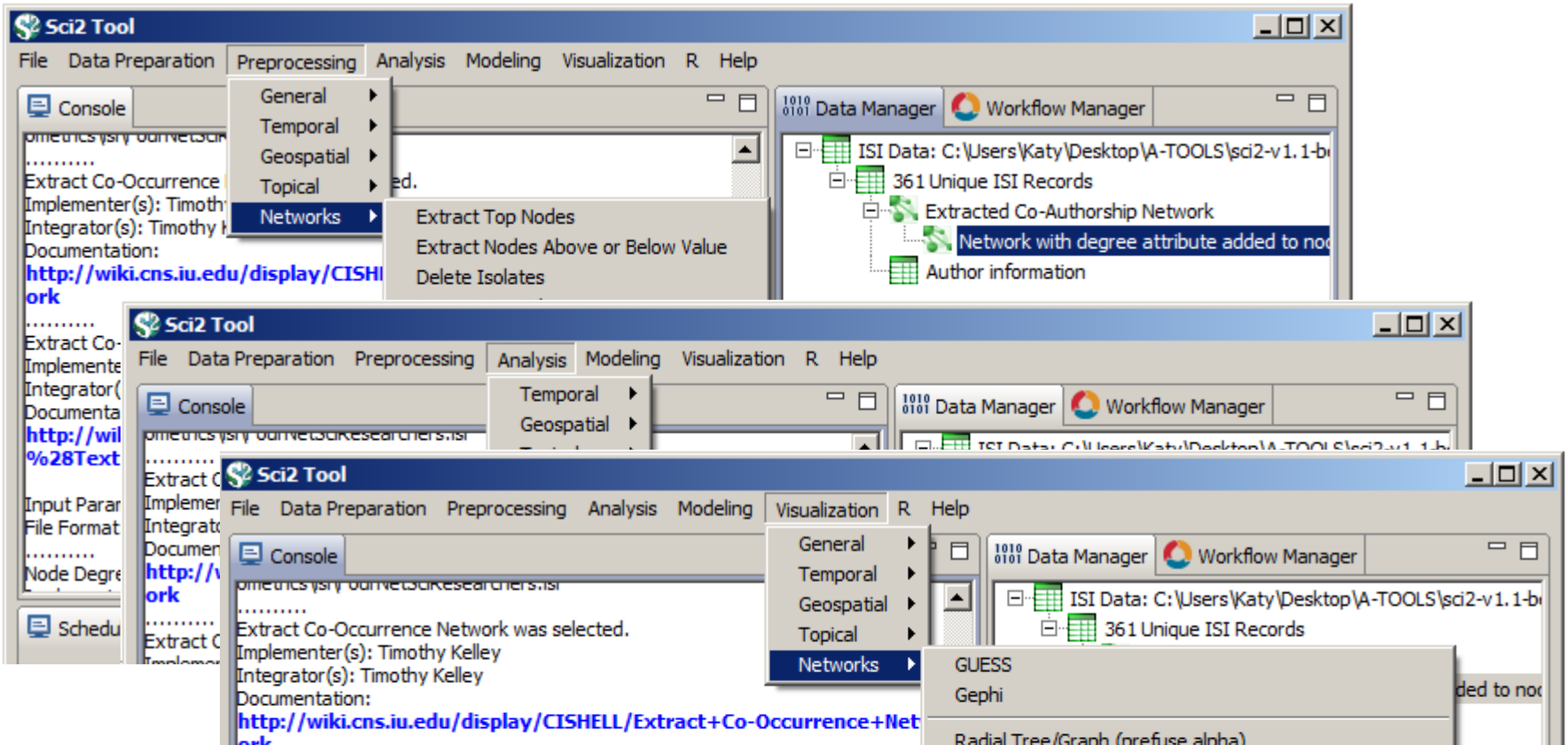


Teaches timeless knowledge:

Visualization framework—exemplified using generic visualization examples and pioneering visualizations.

Sci2 Tool Interface Components

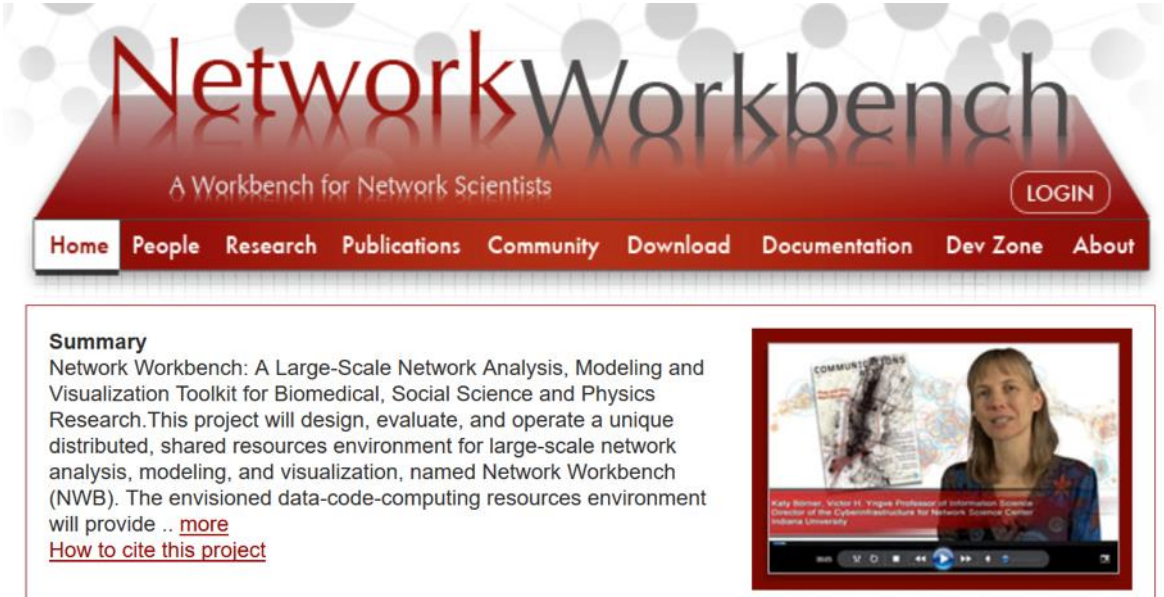
Download tool for free at http://sci2.cns.iu.edu



IVMOOC 2017: Using NSF XSEDE and Jetstream to Run Data Analysis and Visualization Workflows in the Cloud

Katy Borner, 2017.04.20

In Spring 2017, IVMOOC students are invited to beta-test a novel cloud computing setup that supports low- to high-bandwidth users in running more compute intensive data analysis and visualization workflows in the cloud. Specifically, students will create a user account for the Extreme Science and Engineering Discovery Environment (XSEDE) and then use the Jetstream cloud-computing environment to run specific workflows using the Network Workbench Tool.



Hi students,

We wanted to let you know that there is a new assignment posted today that students can complete to earn an extra 2% to their final grade. (https://iu.instructure.com/courses/1595331/quizzes/2271360)

To earn the extra credit, students will have to create an account with XSEDE cloud computing program, and then run a compute instance on the Jet Stream platform to test running a workflow in the cloud environment.

To earn the extra credit, you will need to create your XSEDE account and submit your username to us by Monday April 24 at noon, then we'll link your account to Jetstream. Then from April 25-April 27th at noon, you will need to create an instance on the Jetstream, run a workflow, and submit your result to us.

Best,

Michael, Katy, and Andreas

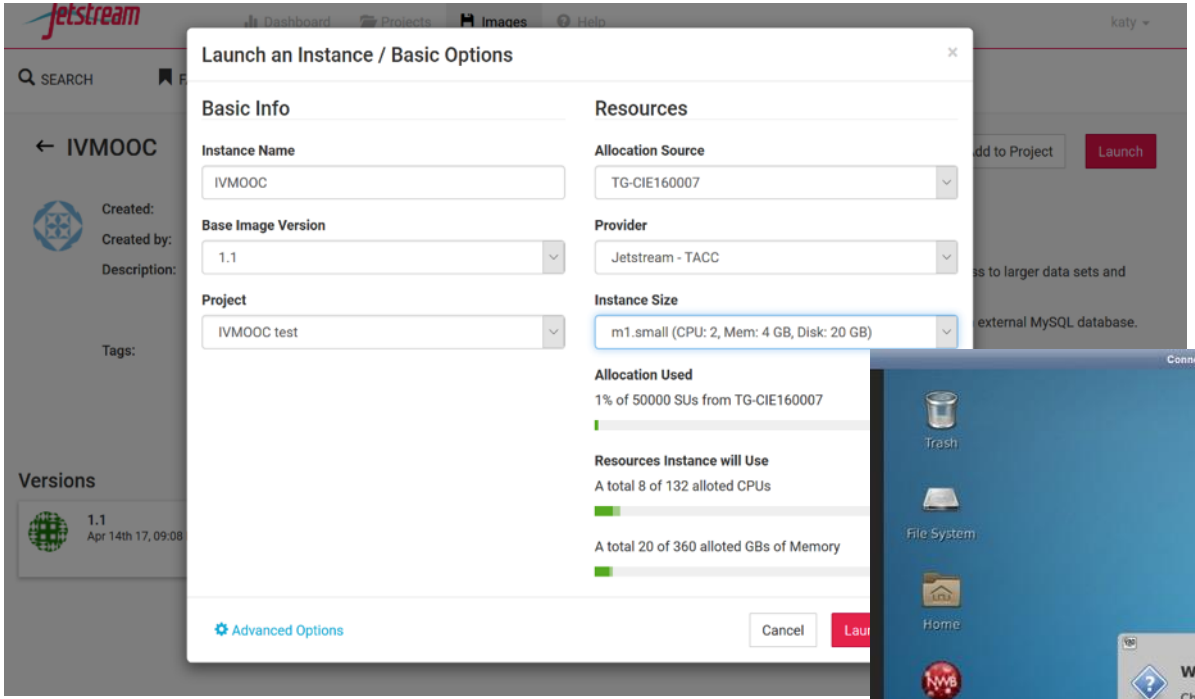
Account Setup

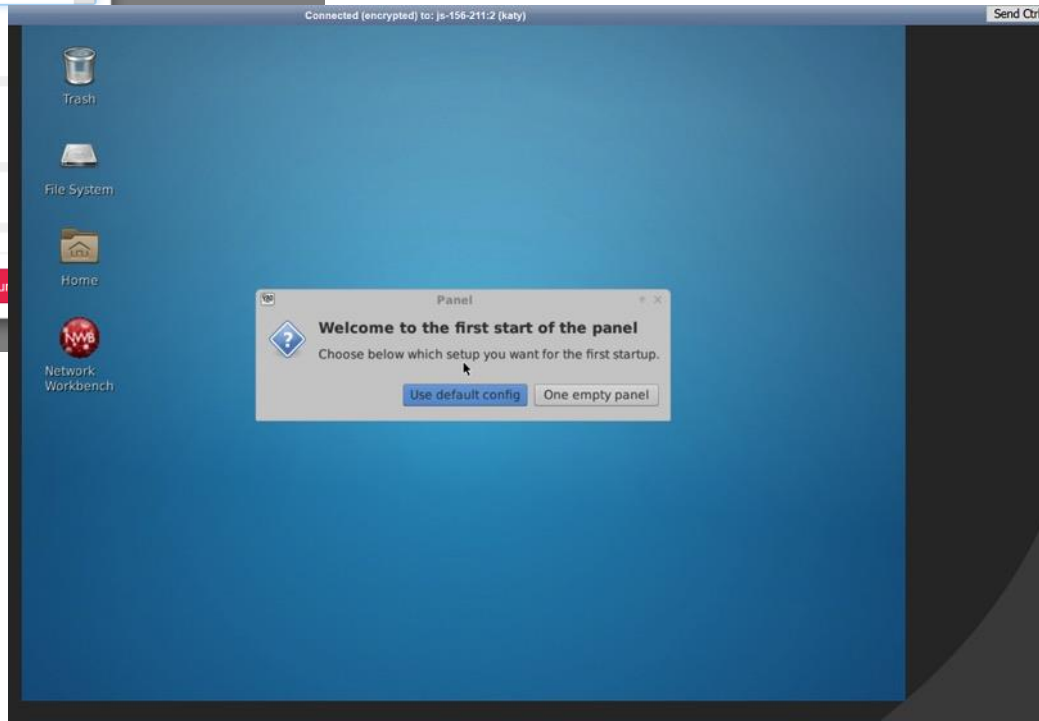
In order to use the IVMOOC Virtual Desktop via the Jetstream cloud environment, you need to get an XSEDE user account and this account needs to be linked to the IVMOOC cloud instance so that you can use this unique resource.

- Create an XSEDE user account via XSEDE User Portal.
- Submit your XSEDE user name via Canvas by 4/24, noon.
  (XSEDE user names were batch-added to IVMOOC VM)
- All IVMOOC students who submitted their XSEDE user name on time gained access to IVMOOC instance on Jetstream by 4/25.

Create to IVMOOC Instance on Jetstream

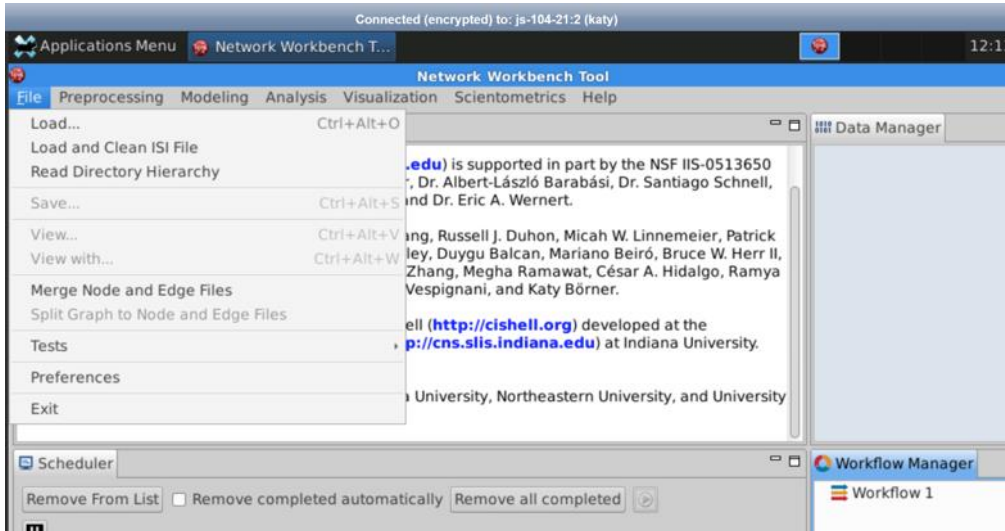
- After 4/25, noon, go to https://use.jetstream-cloud.org/application/images
- Login to use Jetstream using your existing organizational login, select “Indiana University” and use DUO.
- Select “Launch New Instance”
- Search for “IVMOOC” or directly go to https://use.jetstream-cloud.org/application/images/366
- Click on “Launch” in top right.
- Keep default values for the instance but increase “Instance Size” to **m1.small** (2 CPUs, 4096 GB memory, 20 GB disk). Click “Launch Instance” in lower right, see next slide.
- Wait until Status is “Active” then click on Instance with Name “IVMOOC.”
- Wait until “Activity” is N/A (about 1-2 mins). The VW is now ready for usage.
- Reload page. “Open Web Desktop” should now be visible in lower right. Click on it to open virtual desktop in web browser.
- Confirm “Use default config.”


Jetstream
Dashboard
Projects
Images
Help
katy
SEARCH
IVMOOC
Launch
Launch an Instance / Basic Options
Basic Info
Instance Name
IVMOOC
Base Image Version
1.1
Project
IVMOOC test
Resources
Allocation Source
TG-CIE160007
Provider
Jetstream - TACC
Instance Size
m1.small (CPU: 2, Mem: 4 GB, Disk: 20 GB)
Allocation Used
1% of 50000 SUs from TG-CIE160007
Resources Instance will Use
A total 8 of 132 alloted CPUs
A total 20 of 360 alloted GBs of Memory
Advanced Options
Cancel
Created:
Created by:
Description:
Tags:
Versions
1.1
Apr 14th 17, 09:08


Connected (encrypted) to: js-156-211:2 (katy)
Trash
File System
Home
Network Workbench
Panel
Welcome to the first start of the panel
Choose below which setup you want for the first startup.
Use default config
One empty panel

Run NWB Tool

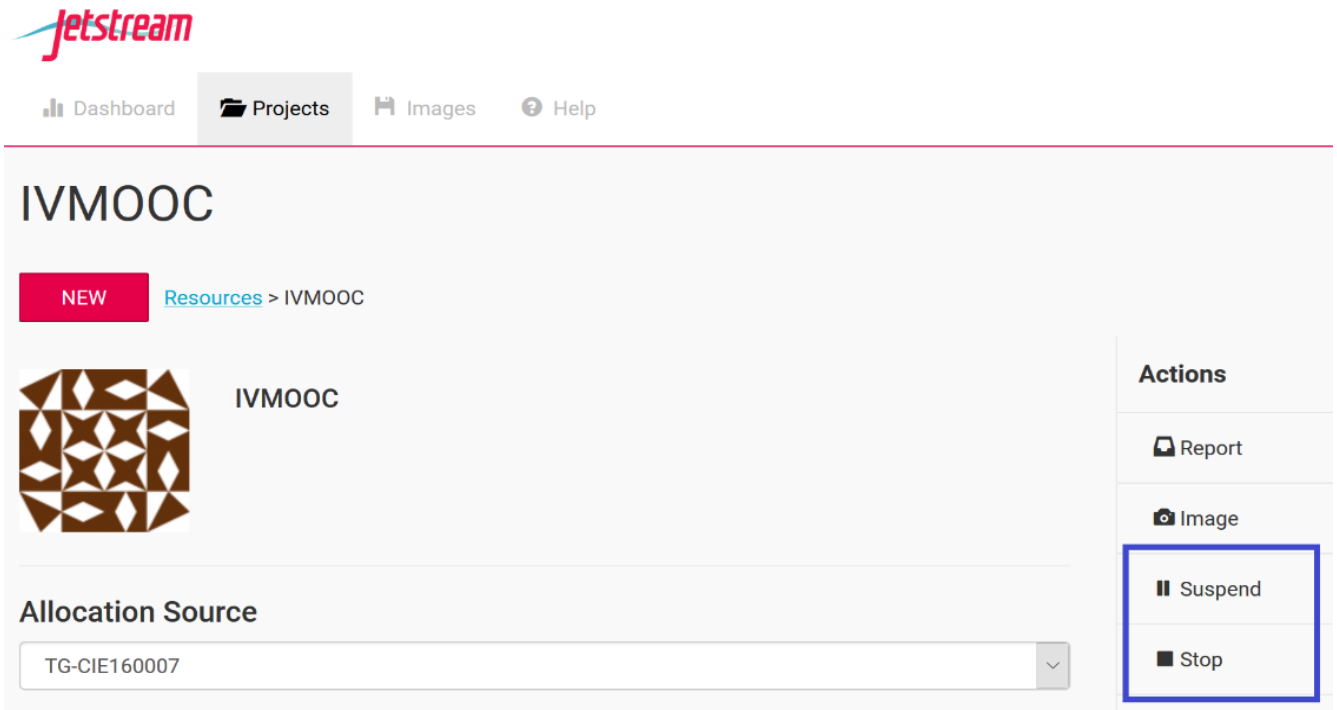
- On Web Desktop, double click Network Workbench (NWB) to run the tool.
- Run “Model > Hypergrid” and generate a network with 10,000 nodes and a Maximum degree of each node: 4.
- With the Hypergrid Network Model being selected in the Data Manager, run “Analysis > Network Analysis Toolkit (NAT).”
- Record the “Average degree” for nodes and submit value via Canvas.
- Feel free to try out other workflows.

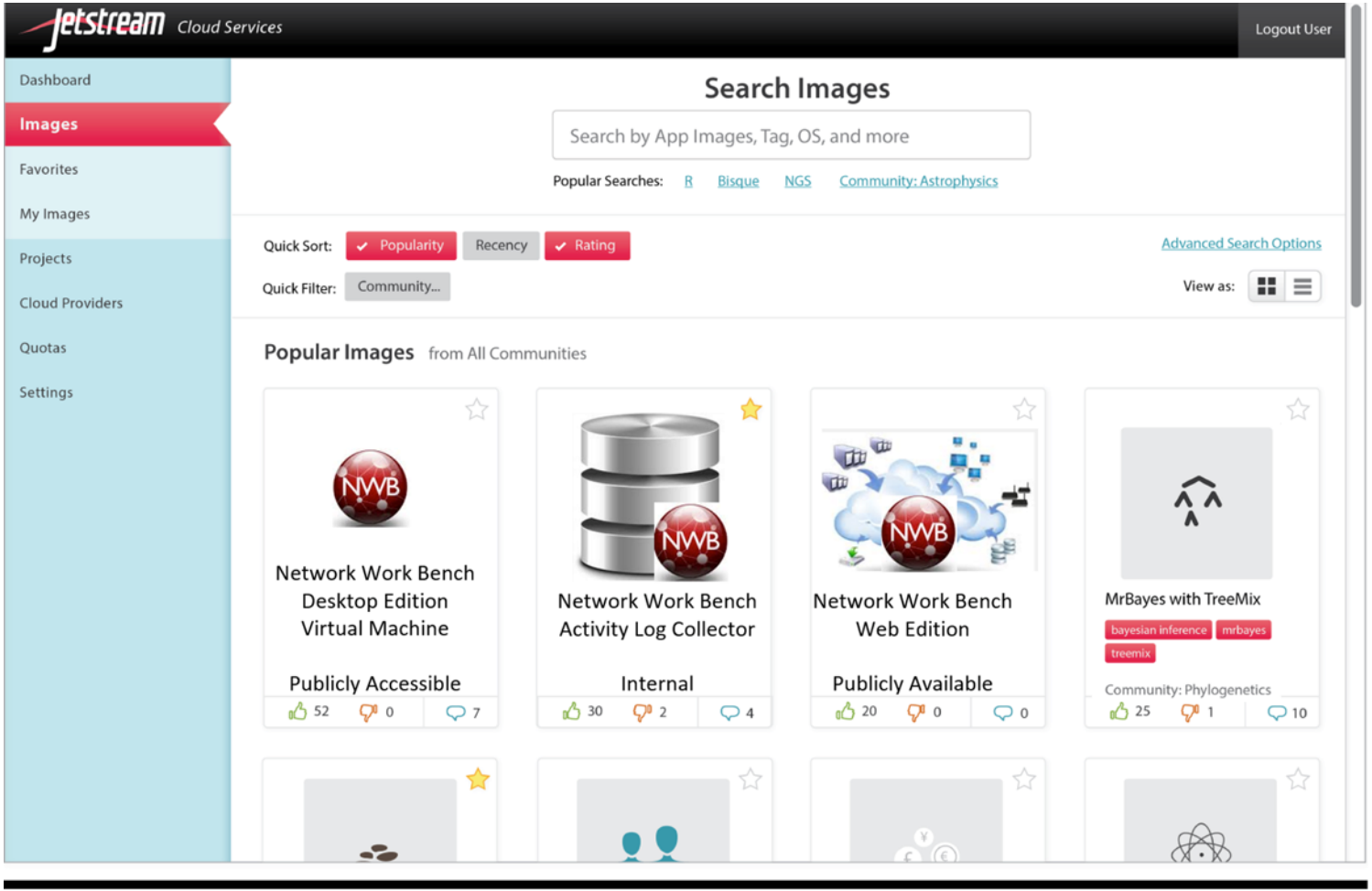


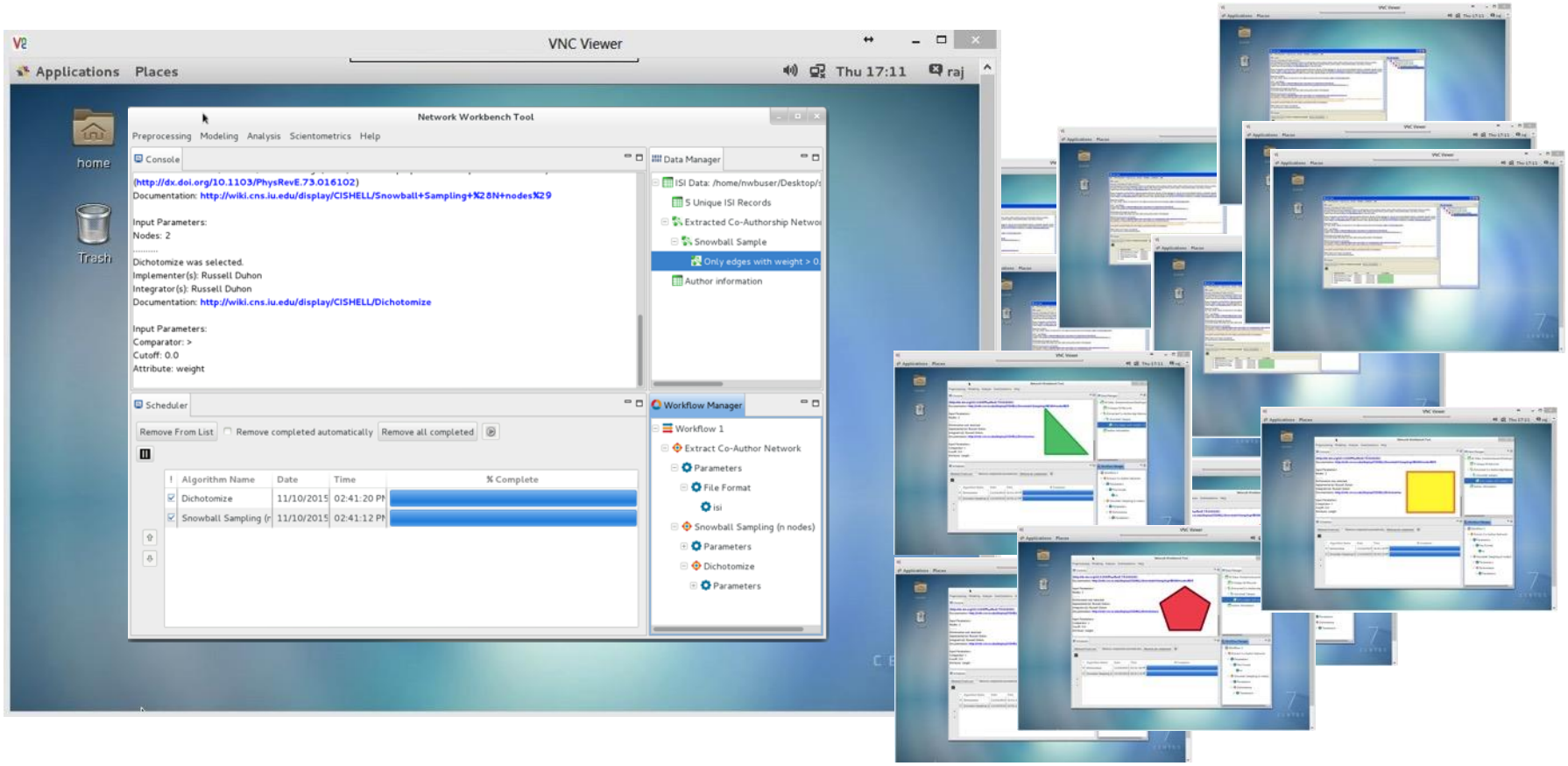
Fletcher, George, Hardik Sheth, and Katy Börner. 2005. "Unstructured Peer-to-Peer Networks: Topological Properties and Search Performance". *Agents and Peer-to-Peer Computing: Third International Workshop, AP2PC 2004. New York, NY* 14-27

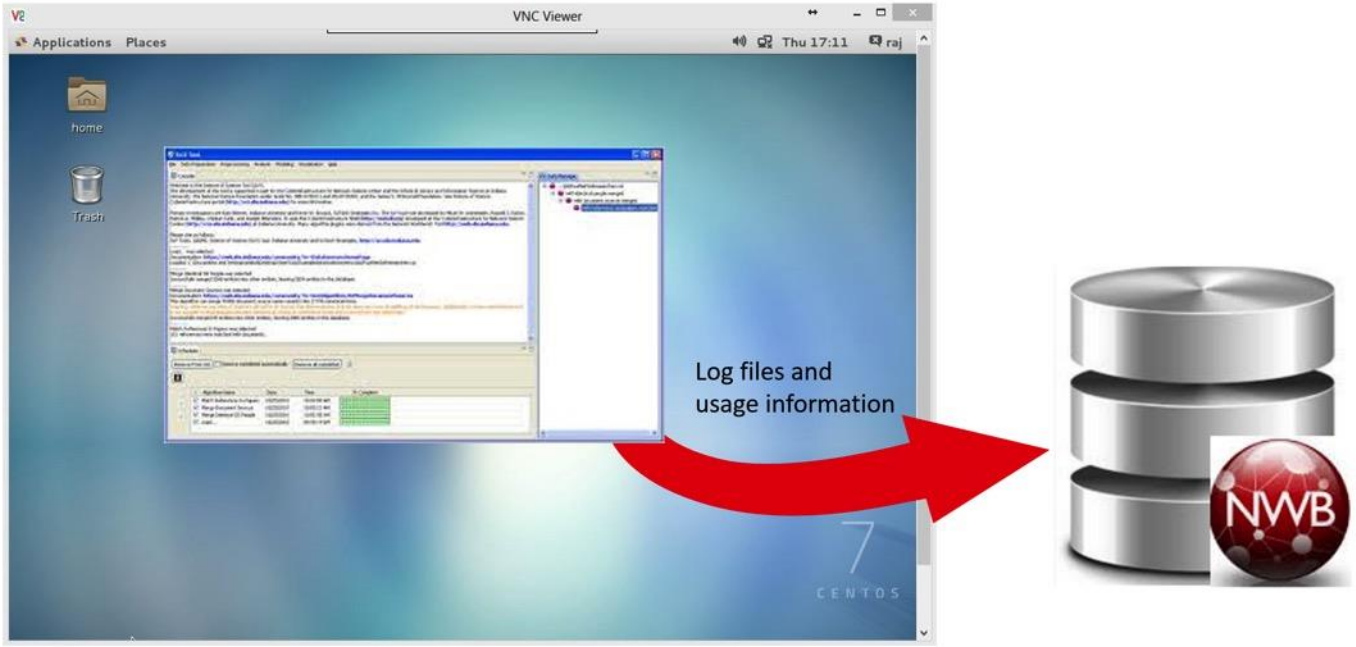
Suspend/Stop IVMOOC Instance on Jetstream

When done, go back to Jetstream web interface and click on “Suspend” or “Stop” in Actions list on right:

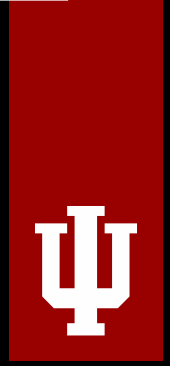



Jetstream Cloud Services
Logout User
Dashboard
Images
Favorites
My Images
Projects
Cloud Providers
Quotas
Settings
Search Images
Search by App Images, Tag, OS, and more
Popular Searches: R Bisque NGS Community: Astrophysics
Quick Sort: Popularity Recency Rating
Advanced Search Options
Quick Filter: Community...
View as:
Popular Images from All Communities
Network Work Bench Desktop Edition Virtual Machine
Publicly Accessible
52 0 7
Network Work Bench Activity Log Collector
Internal
30 2 4
Network Work Bench Web Edition
Publicly Available
20 0 0
MrBayes with TreeMix
bayesian inference
mrbayes
treemix
Community: Phylogenetics
25 1 10


VNC Viewer
Applications
Places
Thu 17:11
raj
home
Trash
Network Workbench Tool
Preprocessing
Modeling
Analysis
Scientometrics
Help
Console
(http://dx.doi.org/10.1103/PhysRevE.73.016102)
Documentation: http://wiki.cns.iu.edu/display/CISHELL/Snowball+Sampling+%28N+nodes%29
Input Parameters:
Nodes: 2
Dichotomize was selected.
Implementer(s): Russell Duhon
Integrator(s): Russell Duhon
Documentation: http://wiki.cns.iu.edu/display/CISHELL/Dichotomize
Input Parameters:
Comparator: >
Cutoff: 0.0
Attribute: weight
Data Manager
ISI Data: /home/nwbuser/Desktop/
5 Unique ISI Records
Extracted Co-Authorship Networ
Snowball Sample
Only edges with weight > 0
Author information
Scheduler
Remove From List
Remove completed automatically
Remove all completed
Algorithm Name
Date
Time
% Complete
Dichotomize
11/10/2015
02:41:20 PM
Snowball Sampling (r
11/10/2015
02:41:12 PM
Workflow Manager
Workflow 1
Extract Co-Author Network
Parameters
File Format
isi
Snowball Sampling (n nodes)
Parameters
Dichotomize
Parameters



- Logging allows for capturing all user actions.
- Supports student evaluation and guidance.
- Helps identify frequency of dataset/algorithm usage per branch of science in support of future tool development.



Open Science Forum, April 26, 2017

Open XD Metrics on Demand Value Analytics

CNS, IUNI, UITS

INDIANA UNIVERSITY BLOOMINGTON

XDMoD

METRICS ON DEMAND

VALUE ANALYTICS MODULE

Supported by the National Science Foundation

Matt Link

Associate Vice President (Acting)

Director, Systems

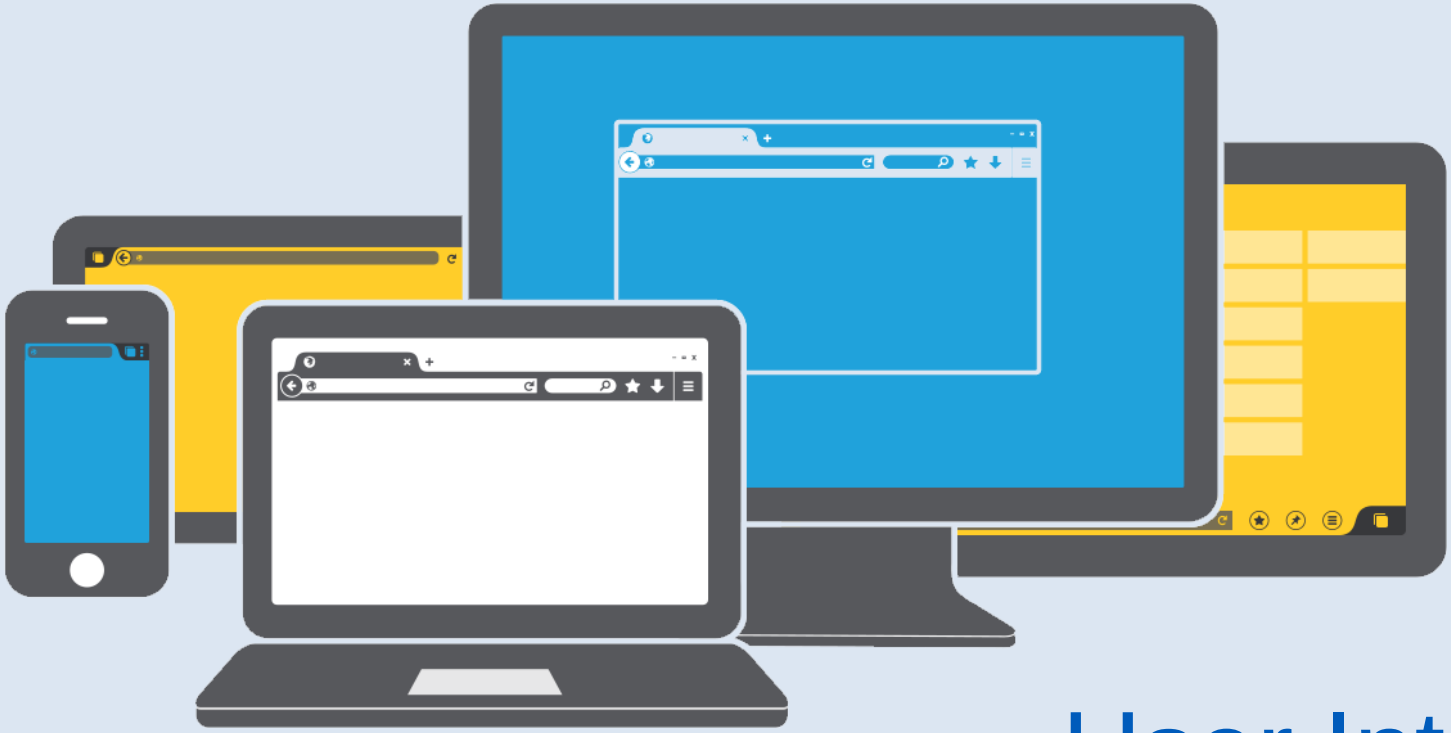
Research Technologies, Pervasive Technology Institute

Office of the Vice President for IT, Indiana University

- Developed by the University at Buffalo Center for Computational Research
- Comprehensive resource management for HPC systems
- Provide detailed operational and usage data
- Support optimization of HPC resource utilization
- Facilitate planning and analysis
- Used for XSEDE metrics

XDMoD – Value Analytics

- NSF eager award to Indiana University and University at Buffalo
- Collaboration between the **Center for Computational Research** at the University at Buffalo, and the **Pervasive Technology Institute** and **Center for Network Science** at Indiana University
- Enables academic institutions to better understand Return On Investment (ROI) on advanced Cyberinfrastructure (CI)
- Shows the value of:
  - Fostering collaboration
  - Supporting scientific publications
  - Show relationship between campus CI and external grant funding
- Local XDMoD installation required – all data kept locally[2]



User Interface

VA module will be integrated into local[1] version

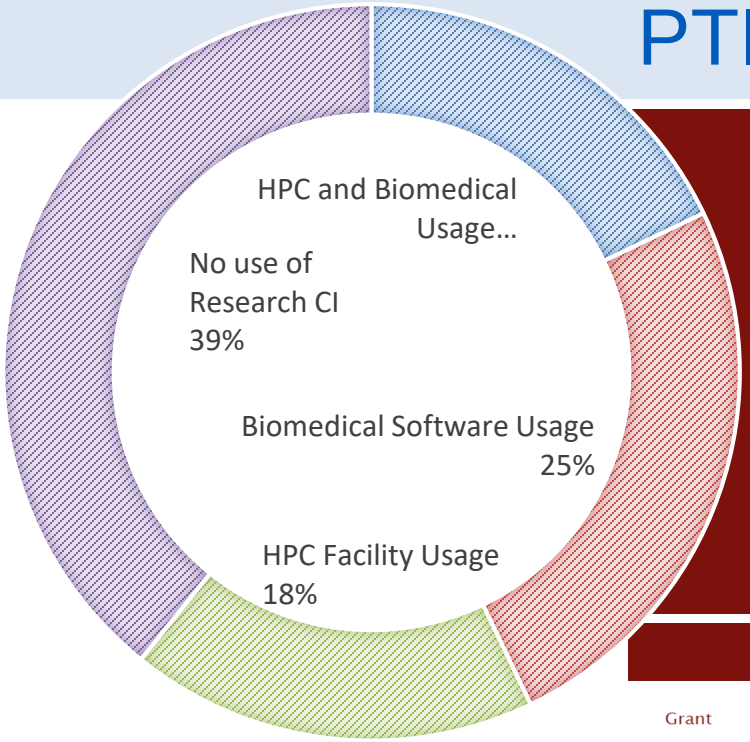
Grant data

Why it’s important

"Since Big Red’s installation in 2006, users of that system were PIs or project directors on a total of $253 million in external funding, which includes $65.4M in facilities and administration funds. In addition to the research dollars flowing into the University, the ability to leverage IU’s leadership in research cyberinfrastructure has aided many other grants awarded to IU.”

–2010

PTI: Funding analytics



Screen shots of PTI developed statistics tool.

RT Stats

Biochemistry/Molecular Biology (Indianapolis)

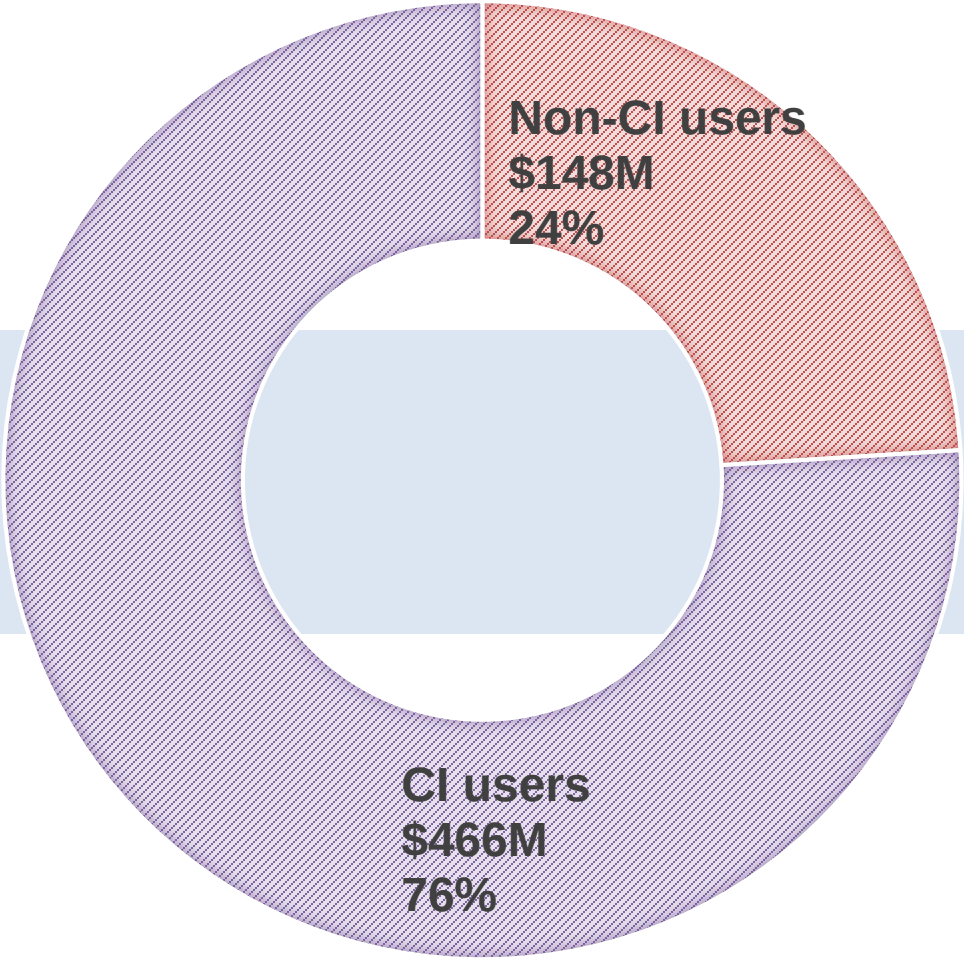
School Of Medicine

Biochem Auxiliary Services	Bioinformatics	Biom Chemical Genomics Program	Diabetes Basic Science Res Ctr
Abdul Sater, Zahi A	Arthur, Jack W	Cerabona, Donna	Conteh, Abass
Craven, Kelly	Edenberg, Howard J	Folck, Anthony F	Fox, Melanie J
Fusakio, Michael	Gendron, Jaimie M	Georgiadis, Millie M	Heyen, Joshua W
Heyerdahl, Darcy	Hoang, Quyen Q	Huang, Fei	Hunter, Gerald Q

Grants
24 totaling $6,056,352.00

Grant	Project	Status	Dates	Award Total
043956-00005B	Docking uPAR for Selective Targeting of Cancer Metastasis	New	January 2015 – June 2016	$180,000.00
044265-00004B	The Bipolar Genome Study	New	May 2014 – April 2016	$87,728.00
053198-00014B	Collaborative Study on the Genetics of Alcoholism (COGA)	New	September 2015 – August 2016	$1,480,906.00
054541-00005B	Regulation of RNA Polymerase II Transcription by the Phosphatase Rtr1	New	August 2015 – July 2016	$291,461.00
054541-00007B	Regulation of RNA Polymerase II transcription by the phosphatase Rtr1	New	August 2015 – July 2016	$34,710.00
056553-00006B	Non-Homologous End Joining Repair in Human	New	April 2015 – March 2016	$319,550.00
056564-00005B	Metabolic stress resones and EIF2 kinase GCN2	New	May 2014 – April 2016	$343,929.00
056564-00008B	Metabolic stress responses and eIF2 kinase GCN2	New	May 2014 – May 2016	$116,099.00
057512-00006B	Early binge drinking and gene regulation	New	September 2015 – August 2016	$225,424.00

2016 IU grant income
$614M



Grant data capabilities

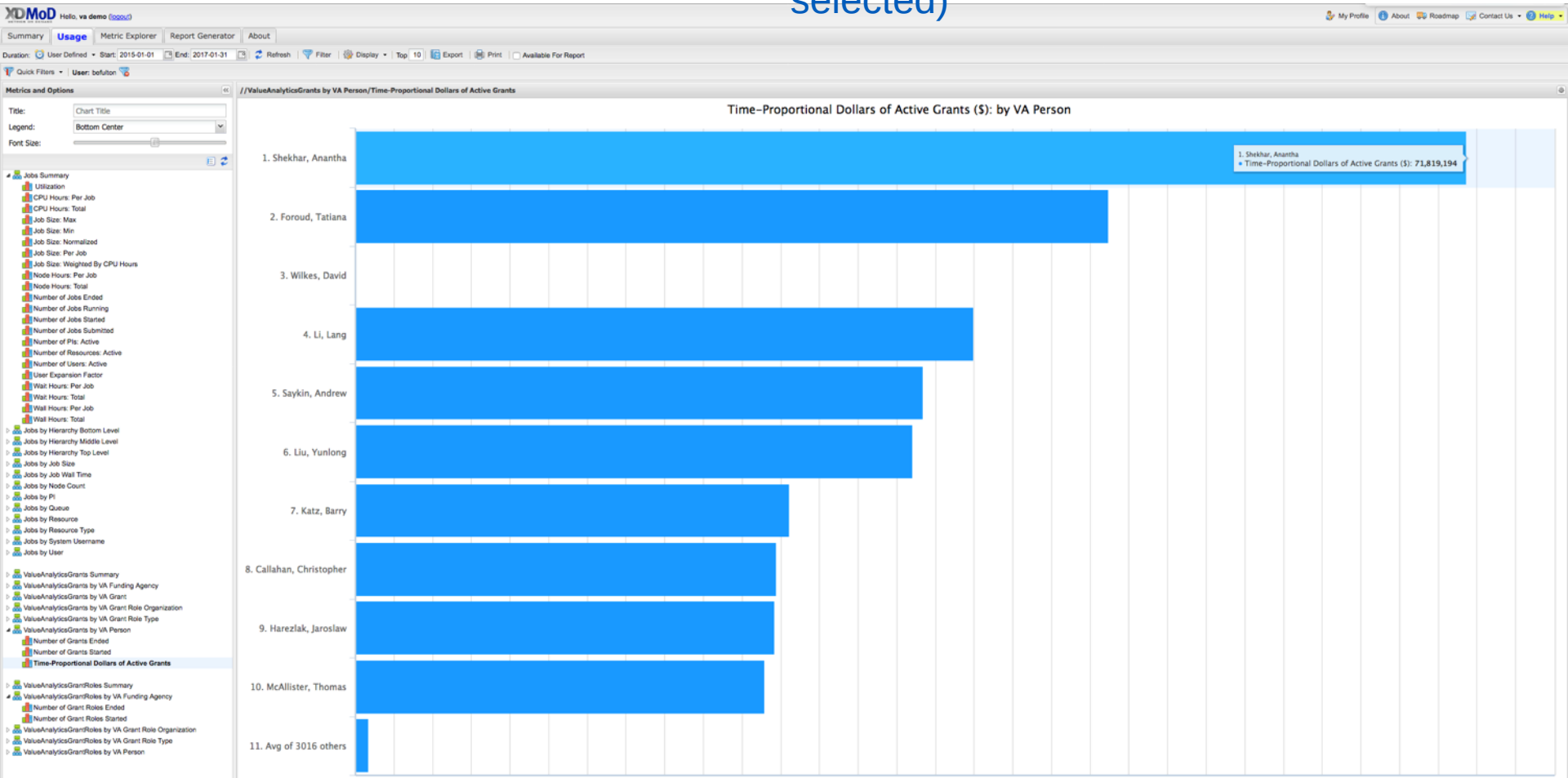
Current

- IU KFS export -> JSON convert -> XDMoD_VA ingest
- NIH and NSF from public datasets
  - IU to provide scripts for NIH and NSF data imports to local[3] XDMoD_VA instance

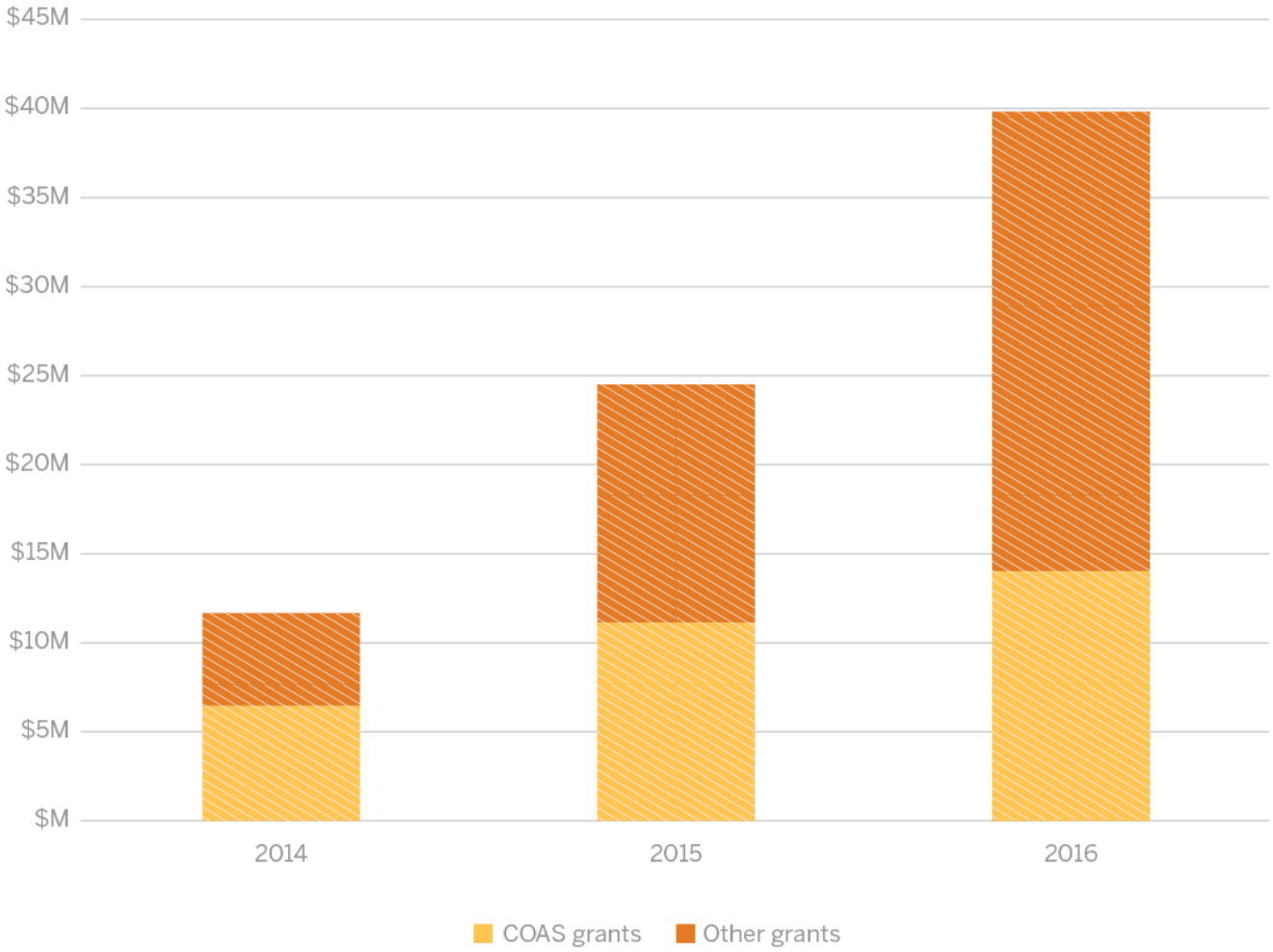
Future

- Ability to group by organizational structure(as seen in PTI slide) (IMS export)

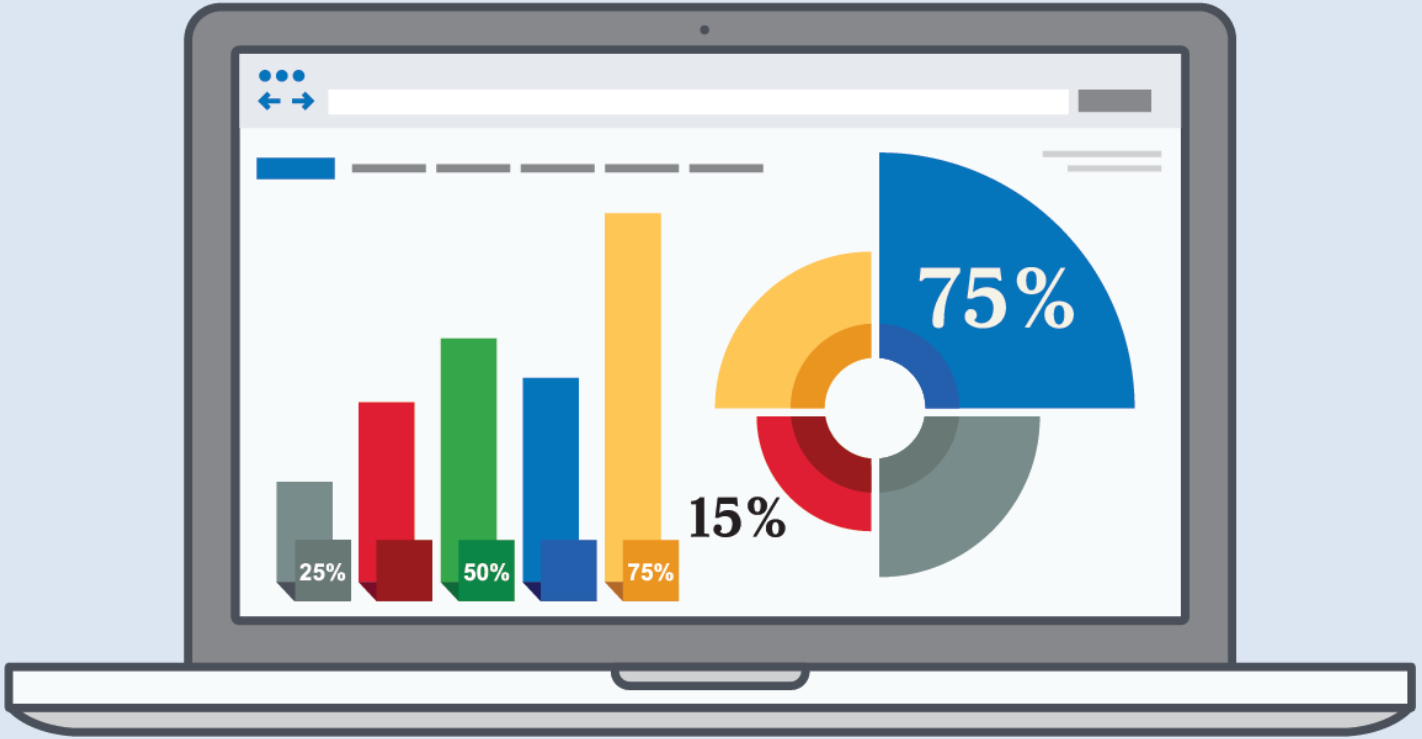
Grant dollars per PI (based on timeline selected)



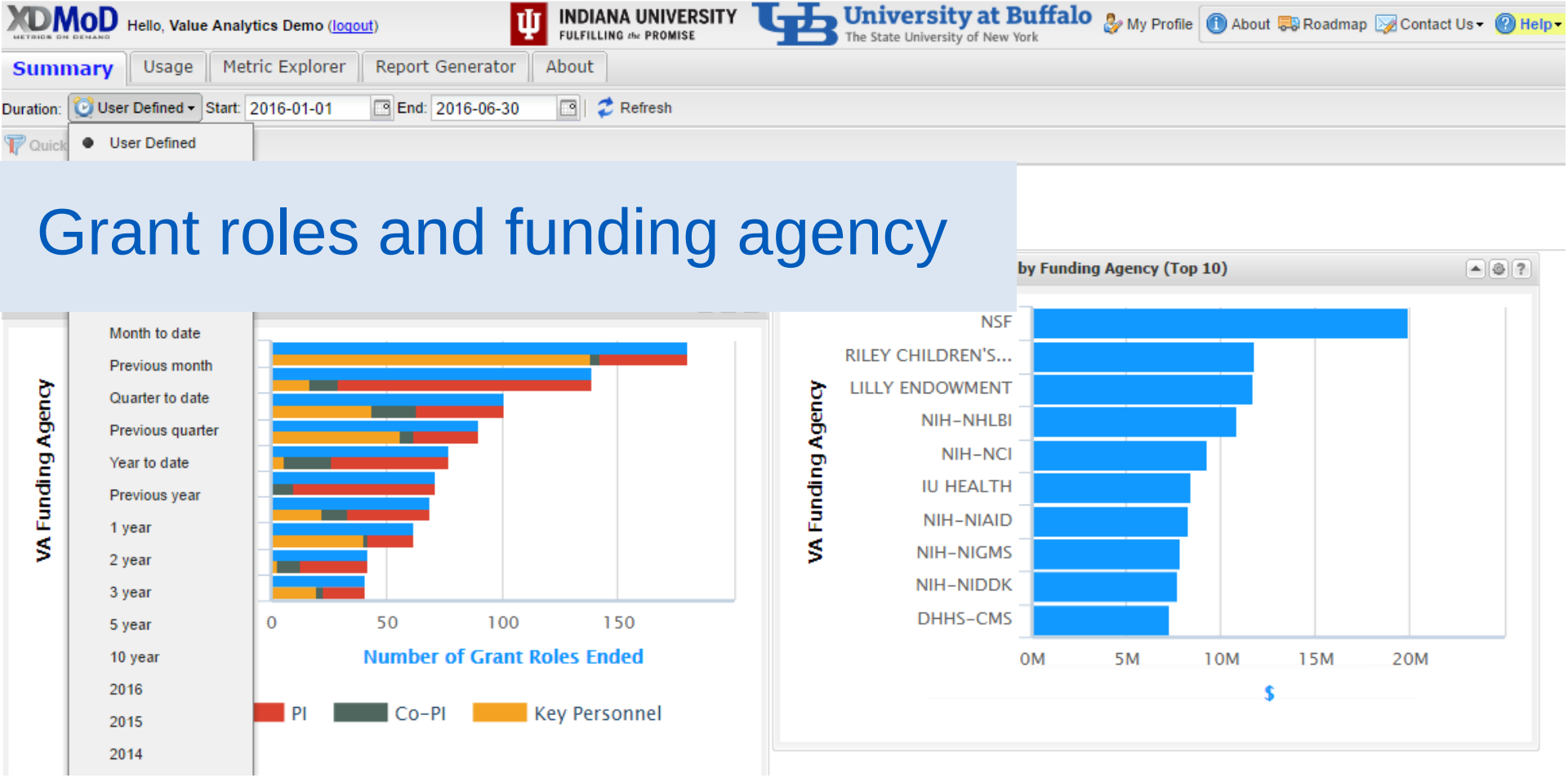
Future capability (grouped by organizational structure)



Roles and funding

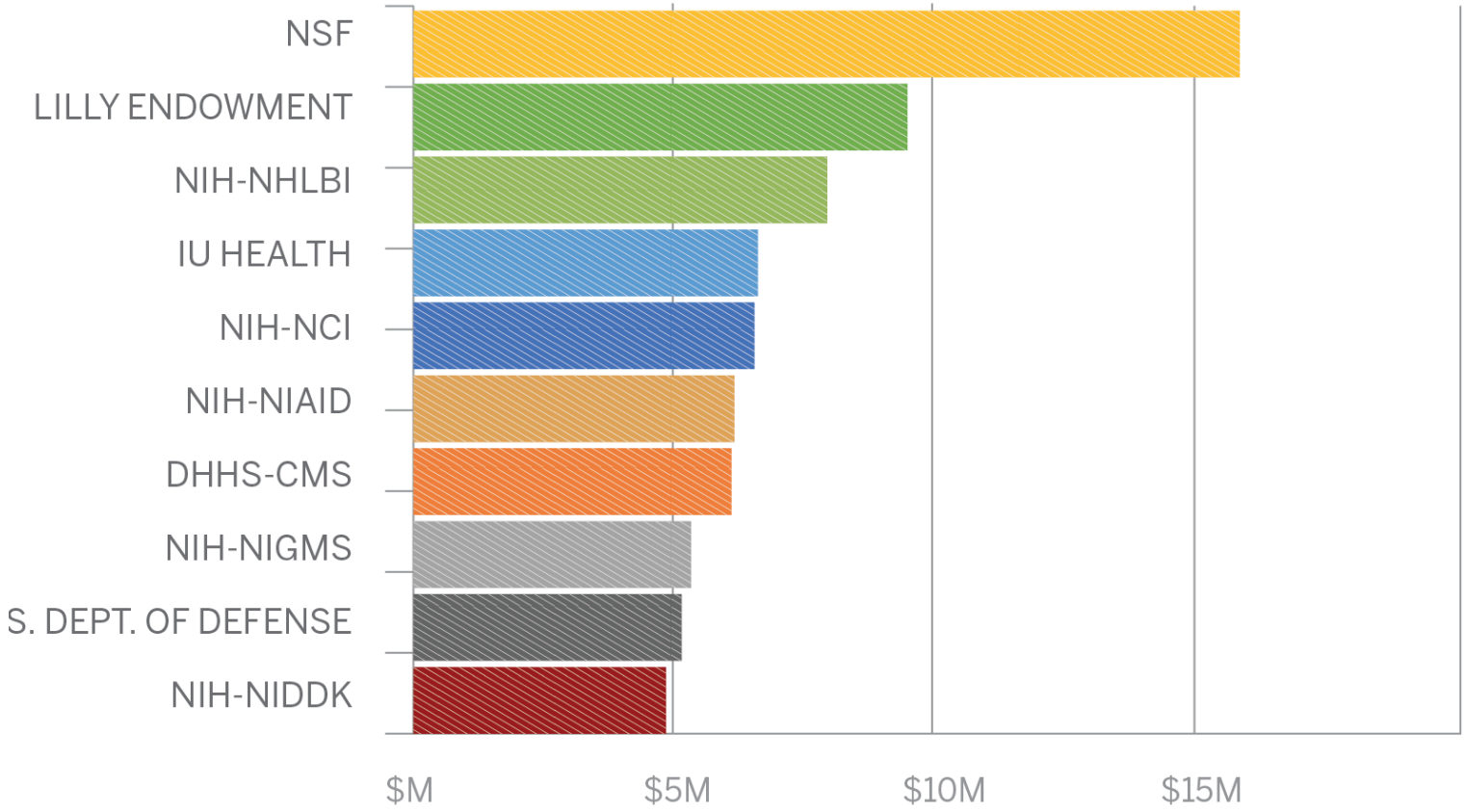


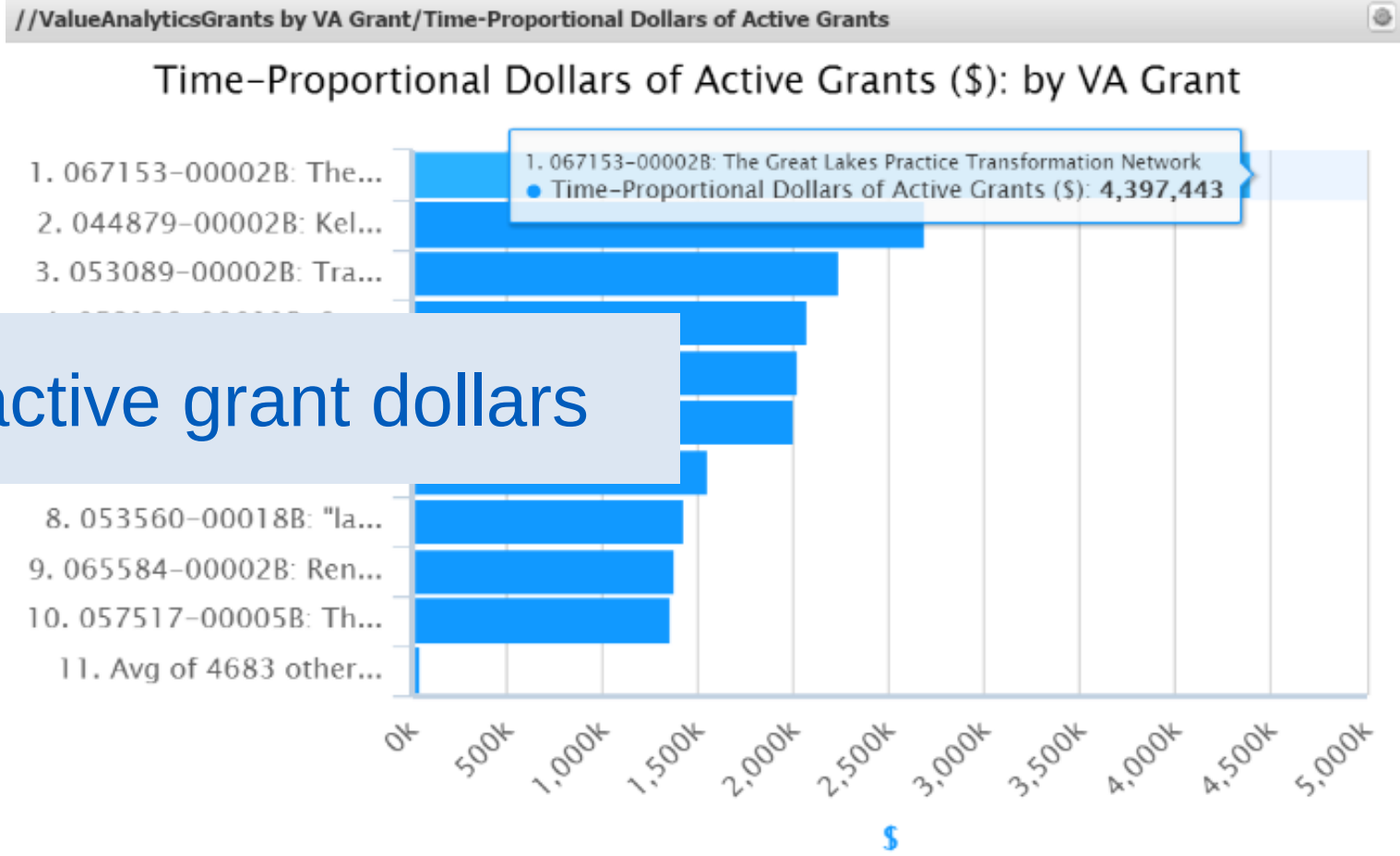
Grant roles and funding agency



Full XDMoD integration (job data and VA data) if available from

Grant dollars by funding agency

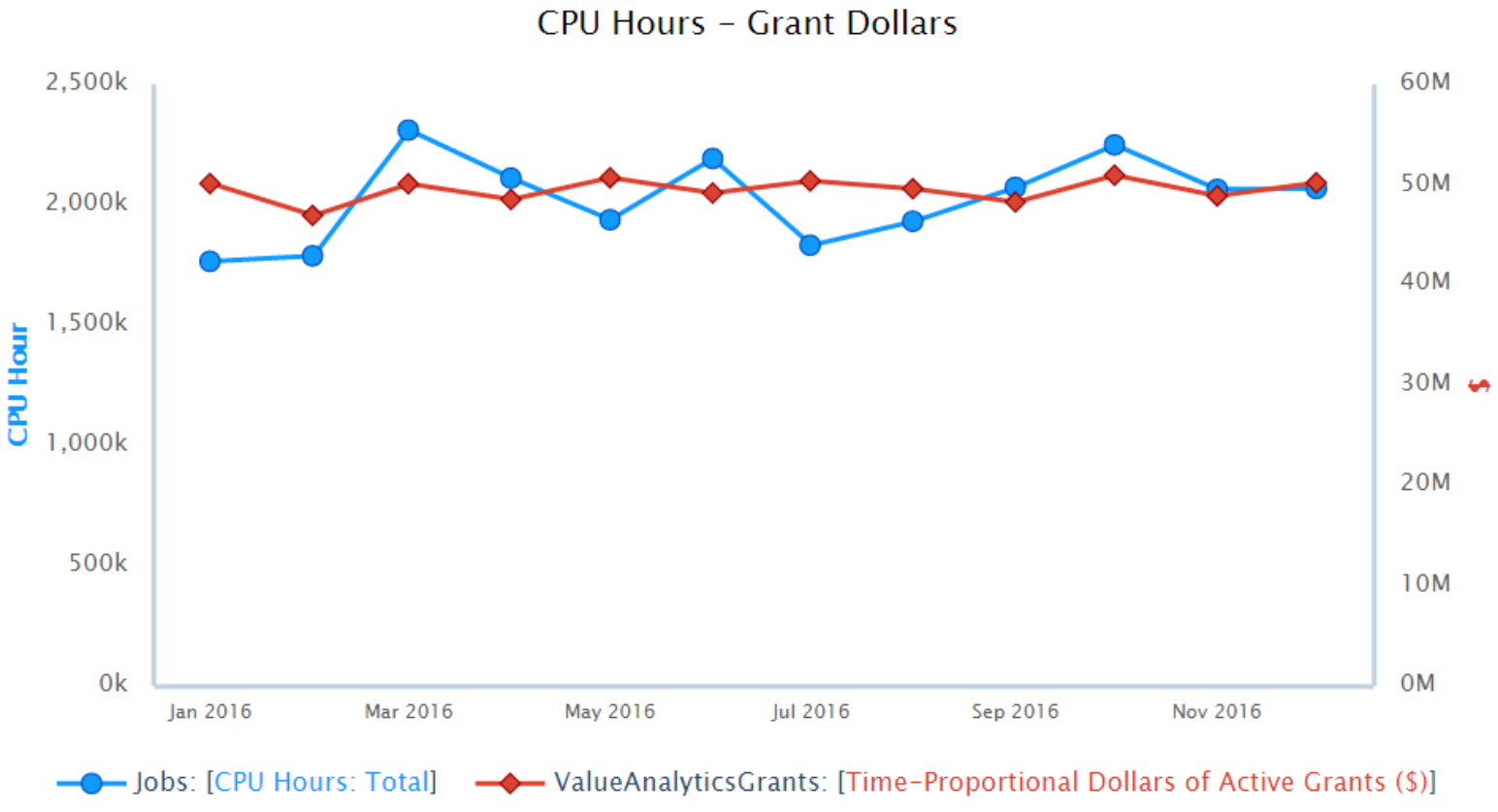


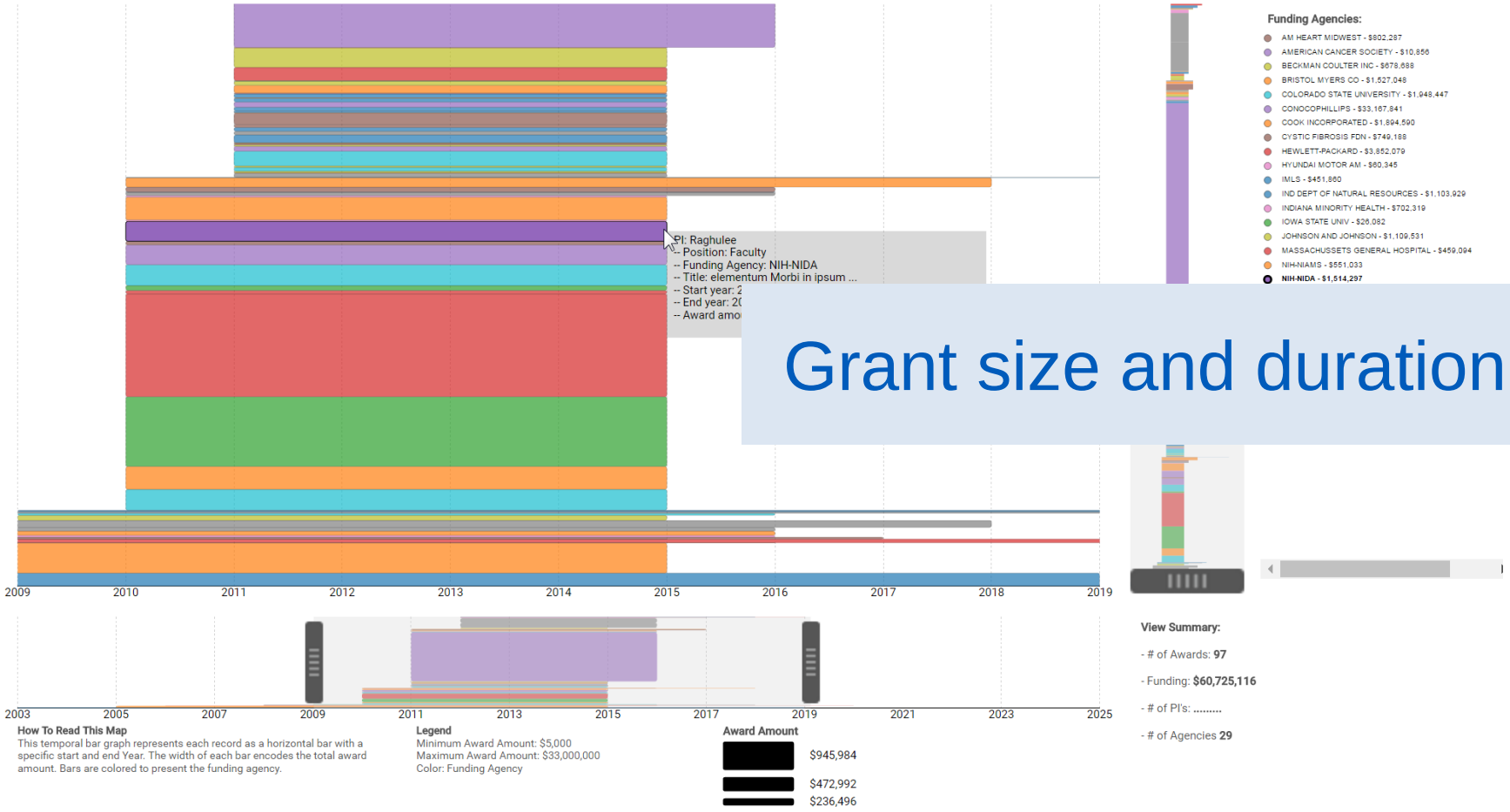


Prorated active grant dollars

Ability to drill down to individual grants

Combined HPC job statistics and value analytics metrics

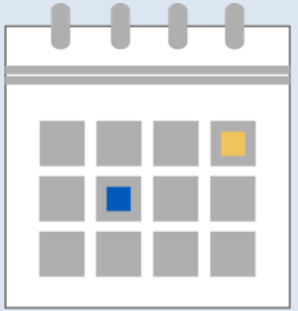


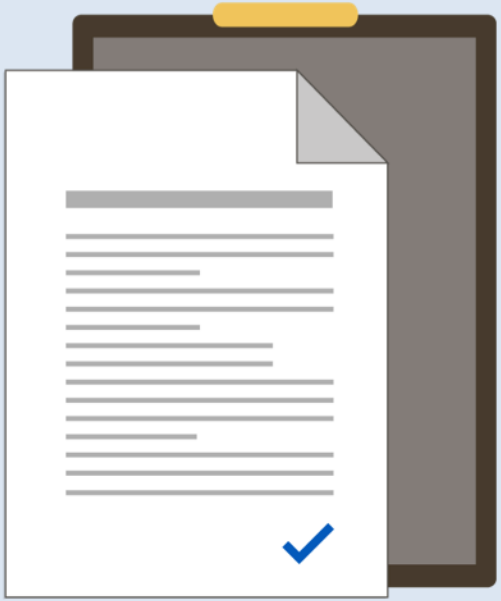


Grant size and duration

Grants over time associated with funding agencies

Publication data





Publication data capabilities

Current

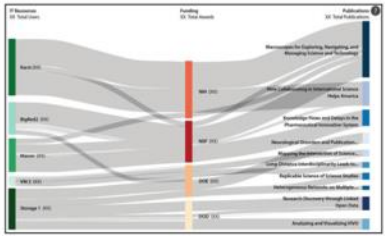
- Working to develop XDMoD schema for ingest
- Access to NIH grant and publication data available
  - NIH Exporter (https://exporter.nih.gov/)
  - Limited visualization capabilities

Future

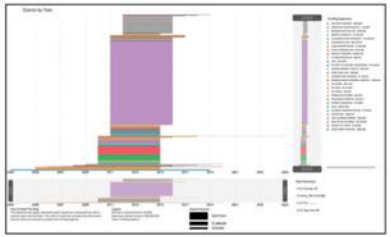
- Roadmap to link publication and grant data
- Ability to group by organizational structure
- Visualization capabilities integrated

XDMOD Value Analytics

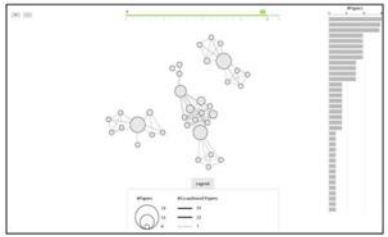
Value Analytics



Funding and Publication Impact
Impact of IT resources on external funding and publications.

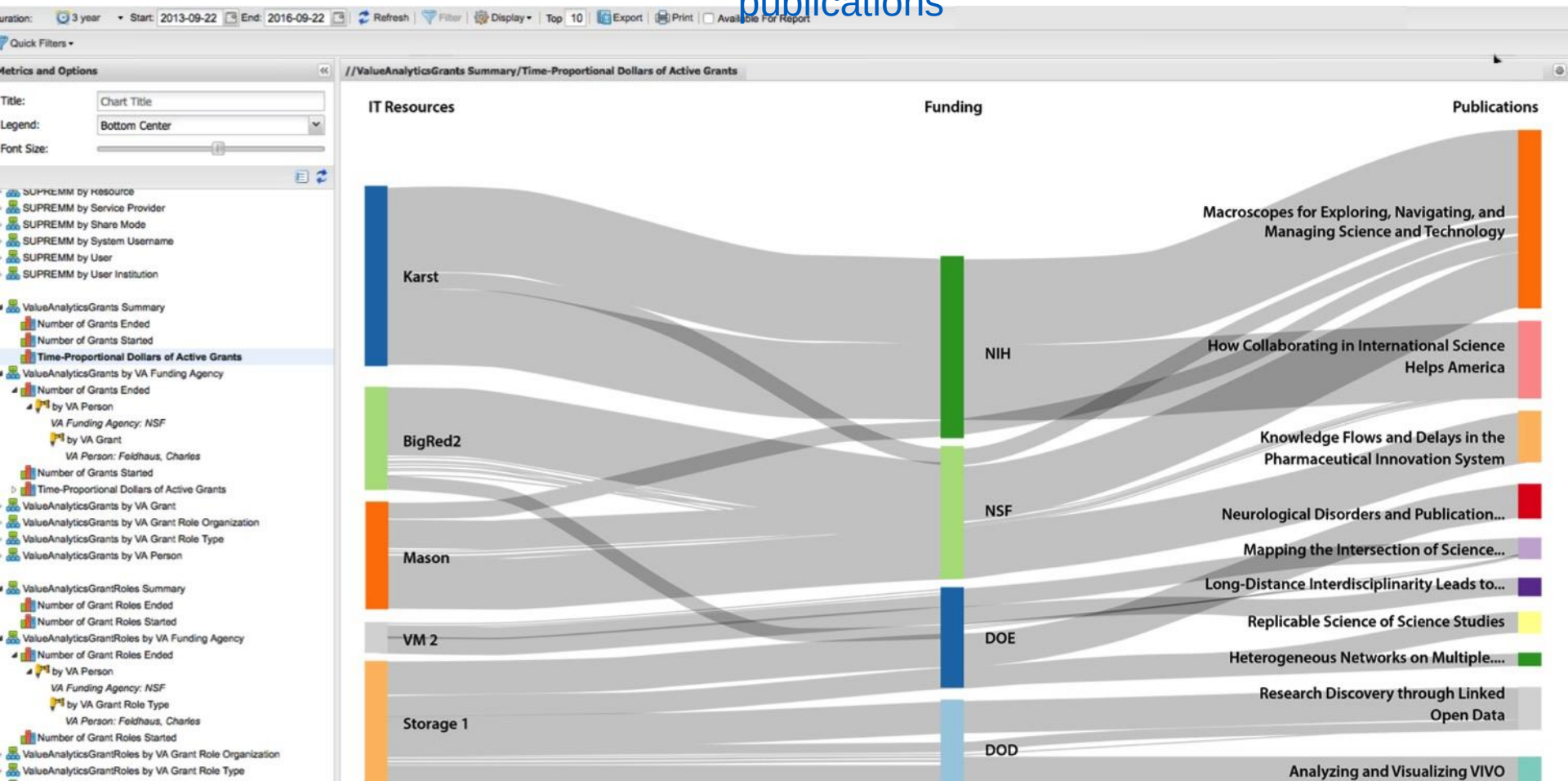


Funding Overview
Funding duration, amounts, and types over time.



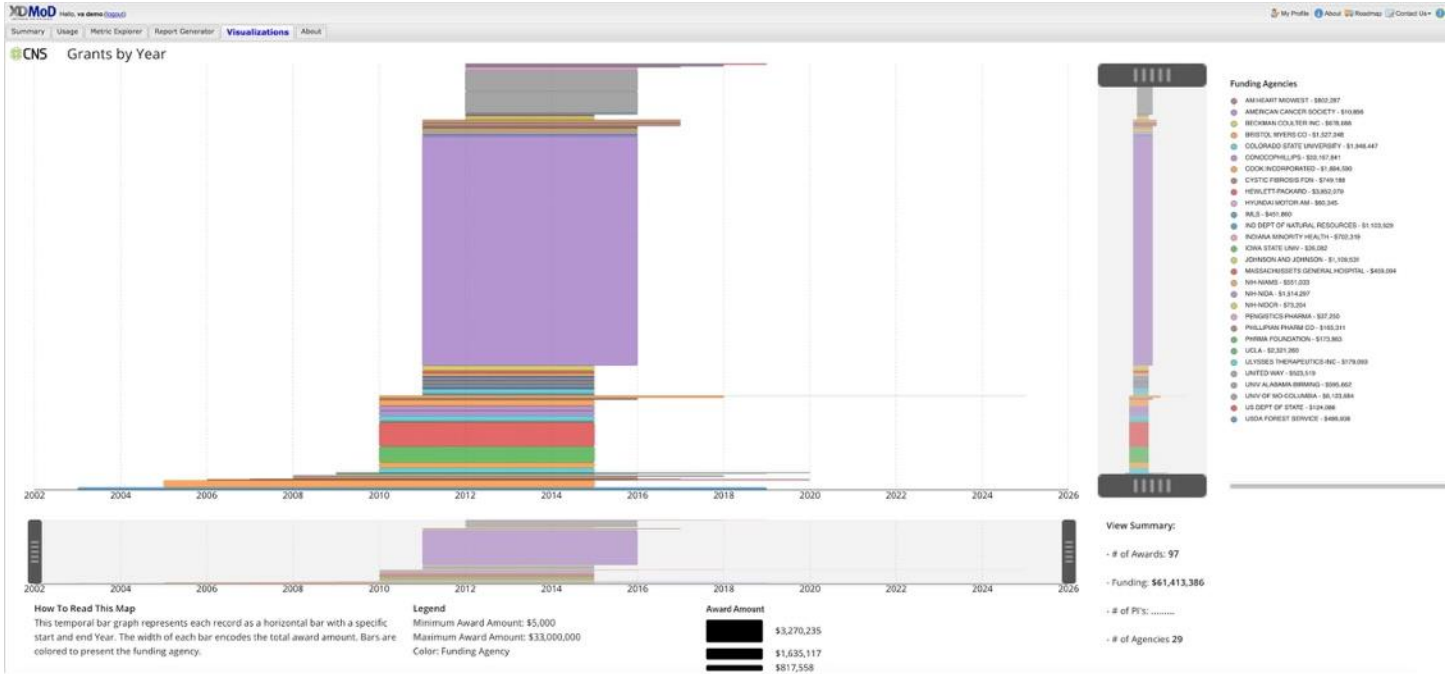
Co-PI Collaboration Network
Co-PI collaboration network based on NSF funding data.

Integrated view: CI resources, funding, and publications

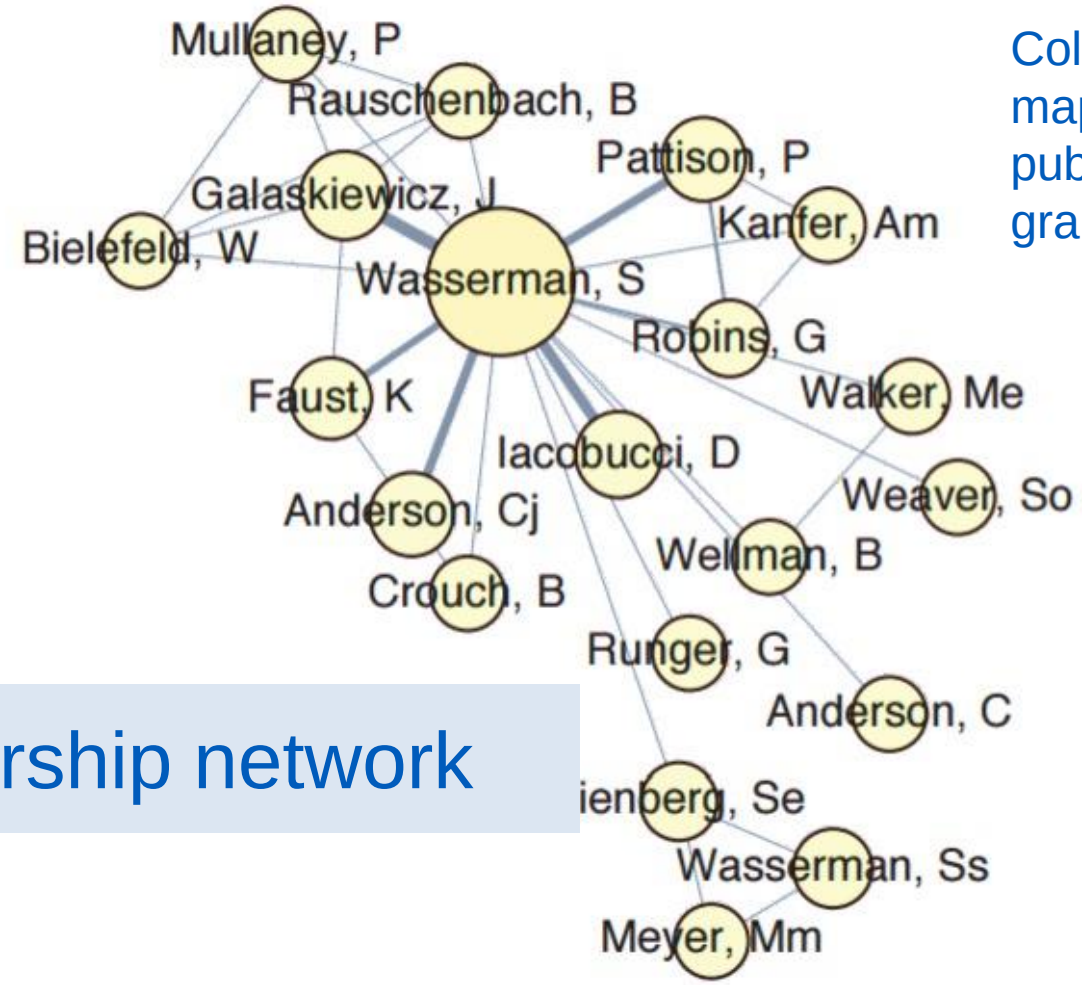


Visualization: Temporal Bar Graph

Project: XDMoD



This temporal bar graph represents each record as a horizontal bar with a specific start and end year. The width of each bar encodes the total award amount. Bars are colored to represent the funding agency (legend of funding agencies are listed on the right). This graph shows funding duration, amounts, and types over time.



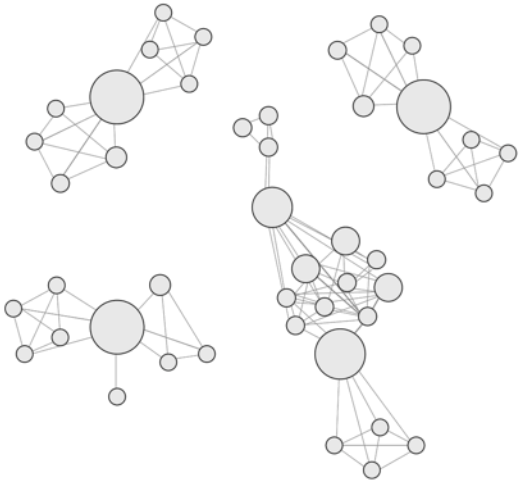
Collaboration mapping for publications and grants

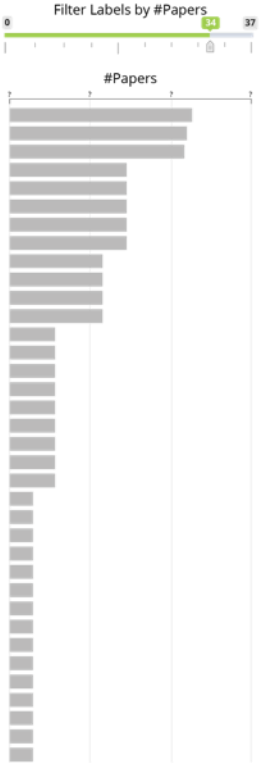
Co-authorship network

Visualization: Co-PI Network

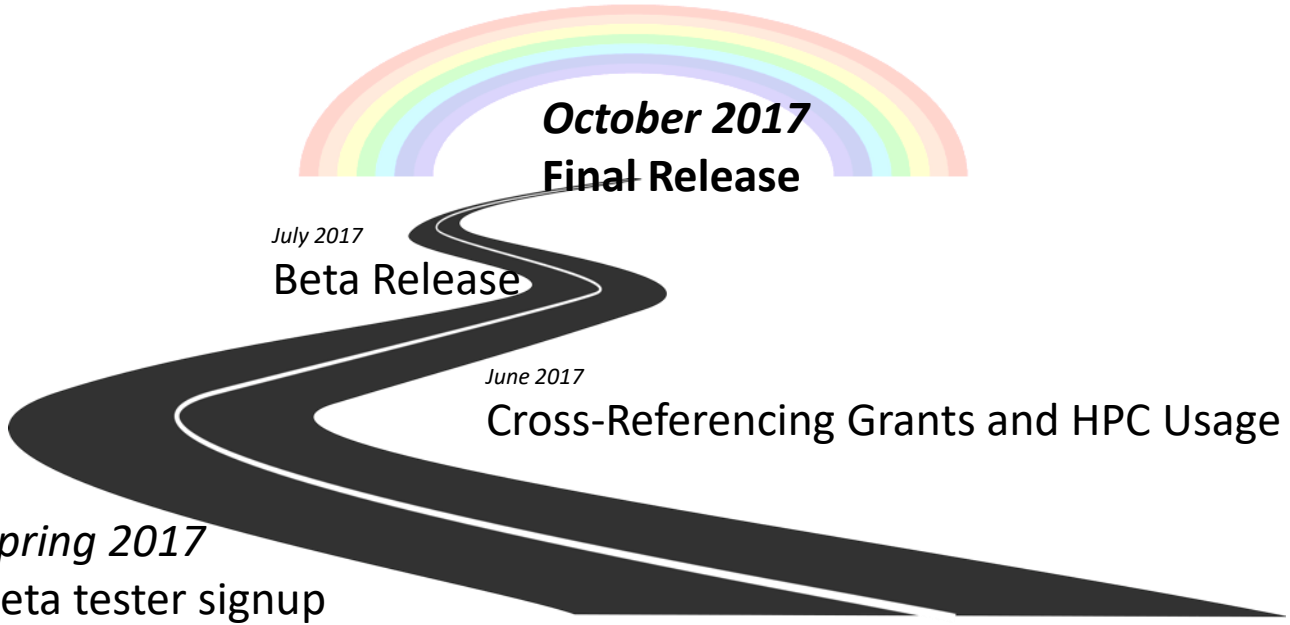
Project: XDMoD

This network represents collaboration patterns based on NSF funding data. Each node represents a principal investigator (PI or Co-PI). It is labelled by the PI's name, size coded by the total value of all awards for each PI, and color coded by the number of collaborators. An edge between two PIs denotes that they have a grant together with edge thickness denoting the number of times they collaborated and edge color reflecting their success in terms of total sum of all their joint awards. The listing on right rank orders PIs by total dollar amounts.





XDMoD Roadmap



October 2017
~~Final~~ Release

July 2017
Beta Release

June 2017
Cross-Referencing Grants and HPC Usage Data

Spring 2017
Beta tester signup
Publication data available
Aggregation charts available

February 2017
Pre-Beta Testing Under Way
Grants and PI information available in XDMoD

Roadmap

February 2017

Pre-beta testing under way

Grant and PI information available in XDMoD.

Spring 2017 | Beta tester signup

Publication data available.

Aggregation charts available.

June 2017 | Cross-referencing grants and HPC usage data

July 2017 | Beta release

October2017 | Final release

Email xdmodva@indiana.edu to participate.

Beta program requirements

- Install and maintain Open XDMoD local[4] instance
- Ability to ingest job logs from computational systems
- Ability to get your grant data from your Office of Research Administration
  - We will help you write a script to convert your data to JSON format for ingestion
- Engagement from your institution (you, ORA)
- Rice and SDSC are our two beta testers so far
- Several institutions expressed interest – we'll be in touch

Thank you.[5] Questions?

Contact:
xdmodva@indiana.edu

License terms

- Please cite as: “XDMoD Value Analytics Module”
- Items indicated with a © are under copyright and used here with permission. Such items may not be reused without permission from the holder of copyright except where license terms noted on a slide permit reuse.
- Except where otherwise noted, contents of this presentation are copyright 2013 by the Trustees of Indiana University.
- This document is released under the Creative Commons Attribution 3.0 Unported license (http://creativecommons.org/licenses/by/3.0/). This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.