Unity: Federated HPC Cluster Operations

University of Massachusetts Amherst

Session Abstract

The University of Massachusetts Amherst launched a new High Performance Computing (HPC) cluster and operations model in 2018 known as the ***Unity*** cluster. Built as a pilot cluster, it has since moved to ***production***.

Unity uses common open-source tools, such as

- Slurm
- Shibboleth
- Open OnDemand (started with Jupyter)

Importantly, the cluster was designed to ***support federated identities*** for users and is also operated as a ***mutually operated collaborative*** serving the needs of several institutions.

Presenters

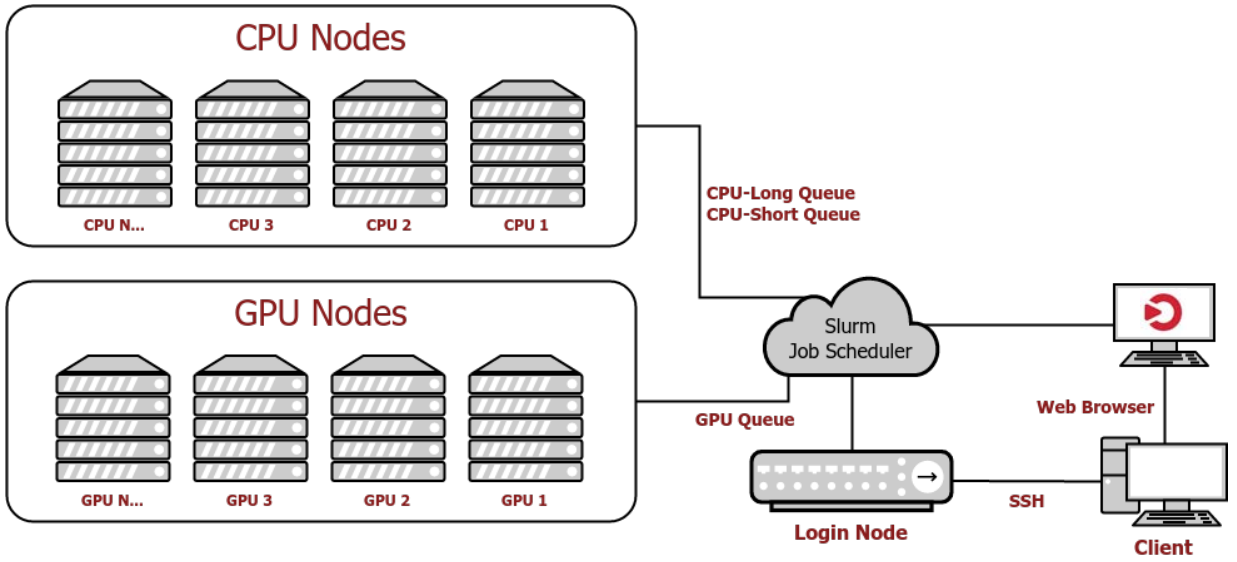
Chris Misra, Vice Chancellor and CIO
University of Massachusetts Amherst

Dave Marble, CEO
OSHEAN

The Unity cluster is a collaborative, multi-institutional high-performance computing cluster

The cluster is under active development and supports primarily research activities.

Started by UMass Amherst in 2018



Welcome to Unity!

Unity Goals

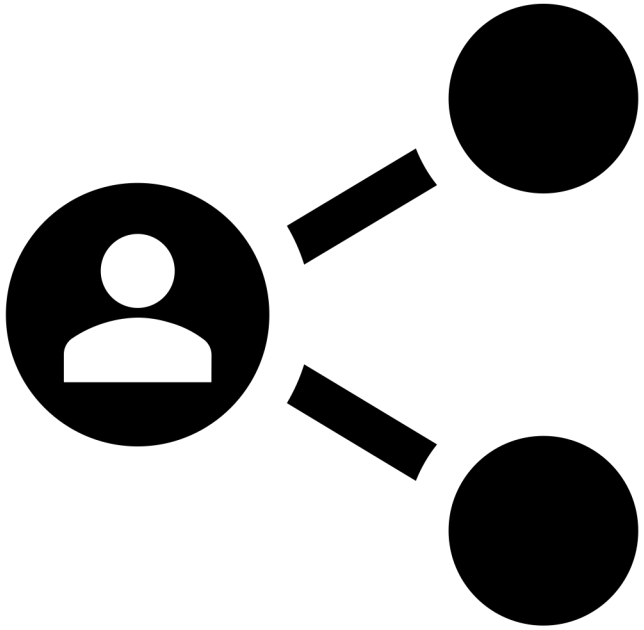
Designed for federated access from day 1

OpenOnDemand

(Stated with Jupyter)

Offer a true priority buy-in model for participating researchers

Democratize access to HPC / data science resources



- Built on Trust
- Shared Staffing Model
- Strong Governance

Principles of Operation

First UMass MGHPCC cluster was built on an LSF scheduler, focused on data-intensive life-science applications

Faculty surveyed could not effectively use cluster for their research

Capital funds targeted at an experimental cluster (Unity) with the goals of federated identity, heterogenous hardware, and feedback-based scheduler configurations

Built on trust and engagement with faculty

Unity Origin Story

Funded by initial capital contribution from UMass IT in 2018

Additional compute nodes donated by Center for Data Science in 2019

NSF MRI nodes added in 2020 (Jianhan Chen)

Buy-in nodes from Astronomy, Civil & Environmental Engineering (Weinberg, Brown, Gleason, Andreadis)

Unity Early Growth

Partnering institutions today include UMass Amherst, UMass Dartmouth, and University of Rhode Island. Over the past year alone these institutions collaborated and received 3 different federal grants of total value $2M to enhance the cluster. Federal agencies are strongly supportive of such regional collaborations on infrastructure.

University of Massachusetts Amherst

THE UNIVERSITY OF RHODE ISLAND

UMass Dartmouth

MGHPCC

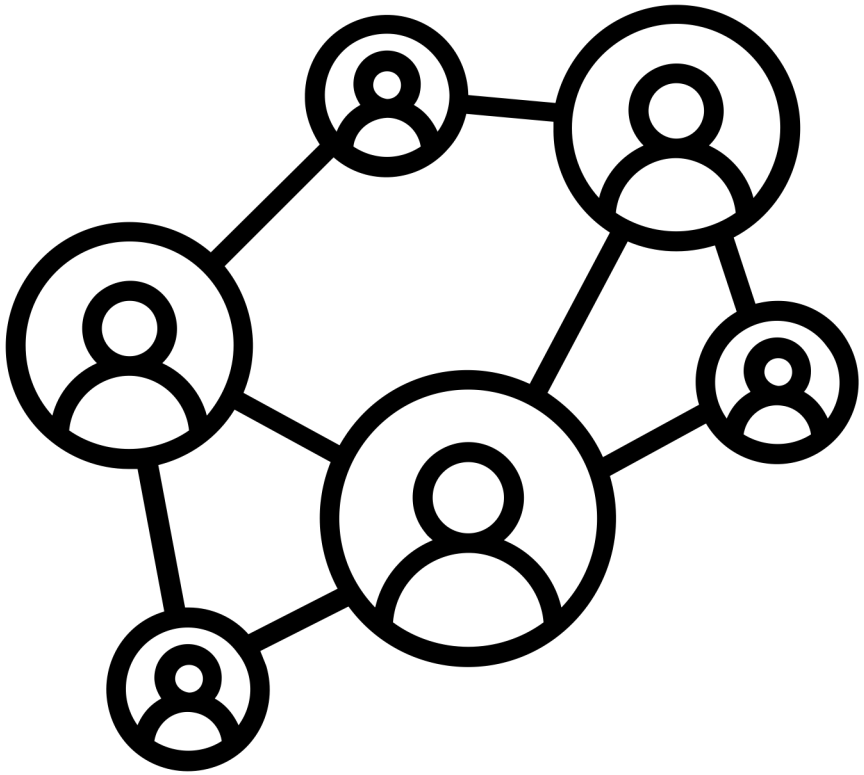
UMass Boston and Lowell are currently on-boarding new users to Unity in Spring 2023

Partner Institutions

Unity Support

Unity operations and support are managed using a novel approach with contributions of ***personnel from 3 universities***.

Unity is also serving as a platform for ***research collaboration*** in computational work across the UMass system and URI.



Building Partnerships

with faculty to succeed in all aspects of the mission.

Implementing Governance

and processes to coordinate and manage departmental and institution investments and priorities.

Assessment

on how the organization and cluster is performing, structured, and able to respond to changing needs.

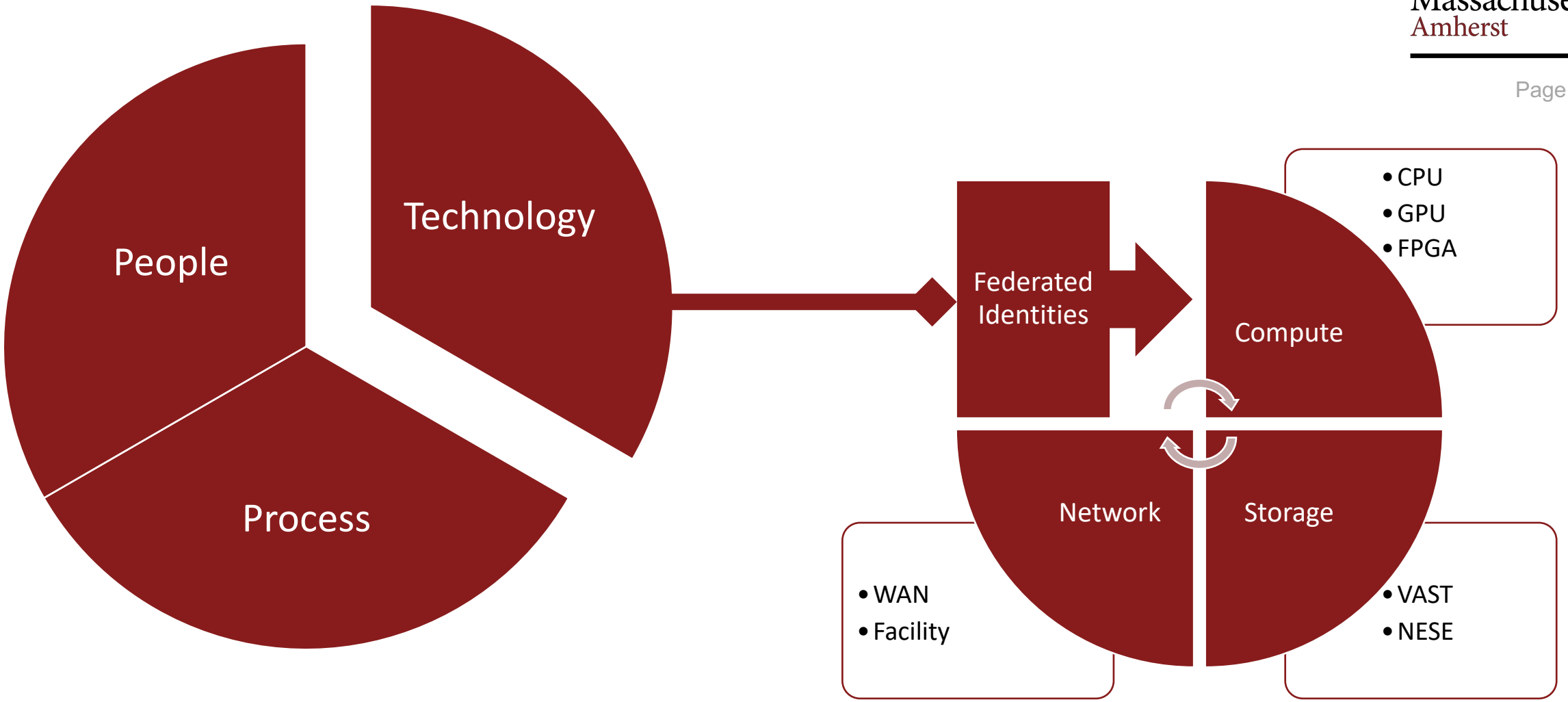
Operational Stability

through documentation, project management, change management, and communications and training plans.

Well-informed Decision Making

supported by data and analytics with input from many voices.

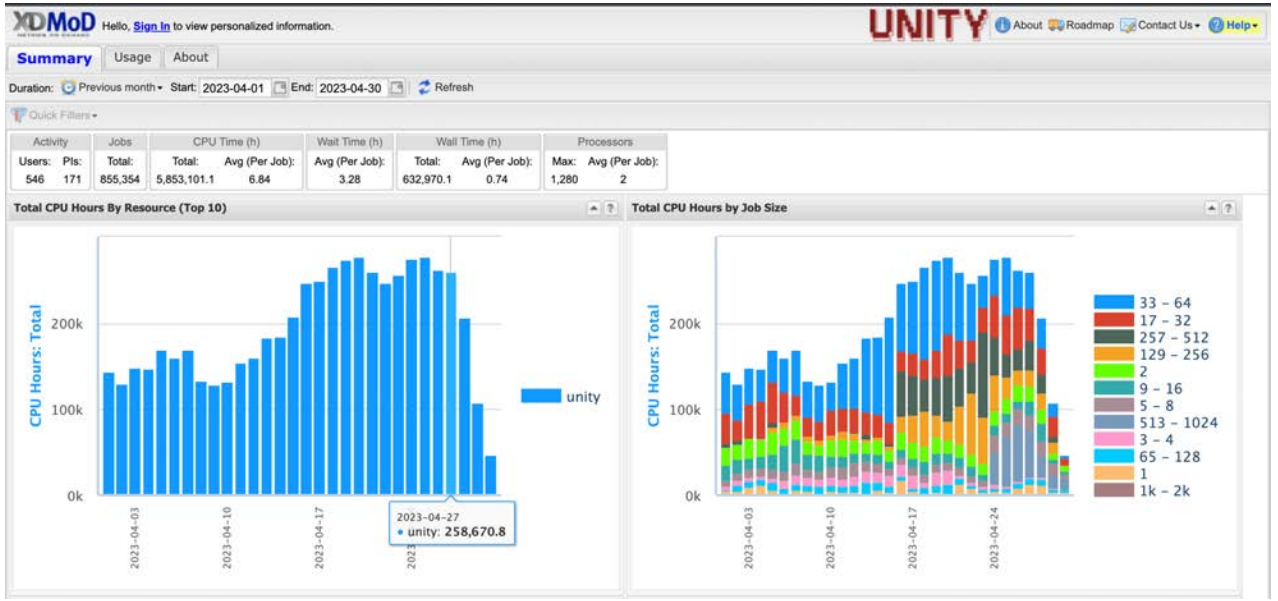
Decision Making Guidelines



Unity Project Topology

Expanded over the last 5 years

- **Currently has ~15,000 cores and ~1,000 GPUs for HPC/AI computations.**
- **Heterogeneous architecture x86, Power, ARM, GPU, (FPGA soon).**
- **Special features include access via InCommon (Shibboleth)**
- **Web-based Jupyter and Open OnDemand**



Unity Compute

Unity Storge: NESE

NESE is the Northeast Storage Exchange, a ***shared regional storage collaboration*** funded by the National Science Foundation

- Operated as a long-term partnership between the Universities and institutions.

Our main goals are to meet the storage needs of the data revolution for science, engineering, education and technology, particularly for researchers in the northeastern part of the U.S.

Also located at MGHPCC

- The Unity Cluster is located at the ***Massachusetts Green High Performance Computing Center***
- The Massachusetts Green High Performance Computing Center (MGHPCC) is an intercollegiate high-performance computing facility, located in Holyoke MA.

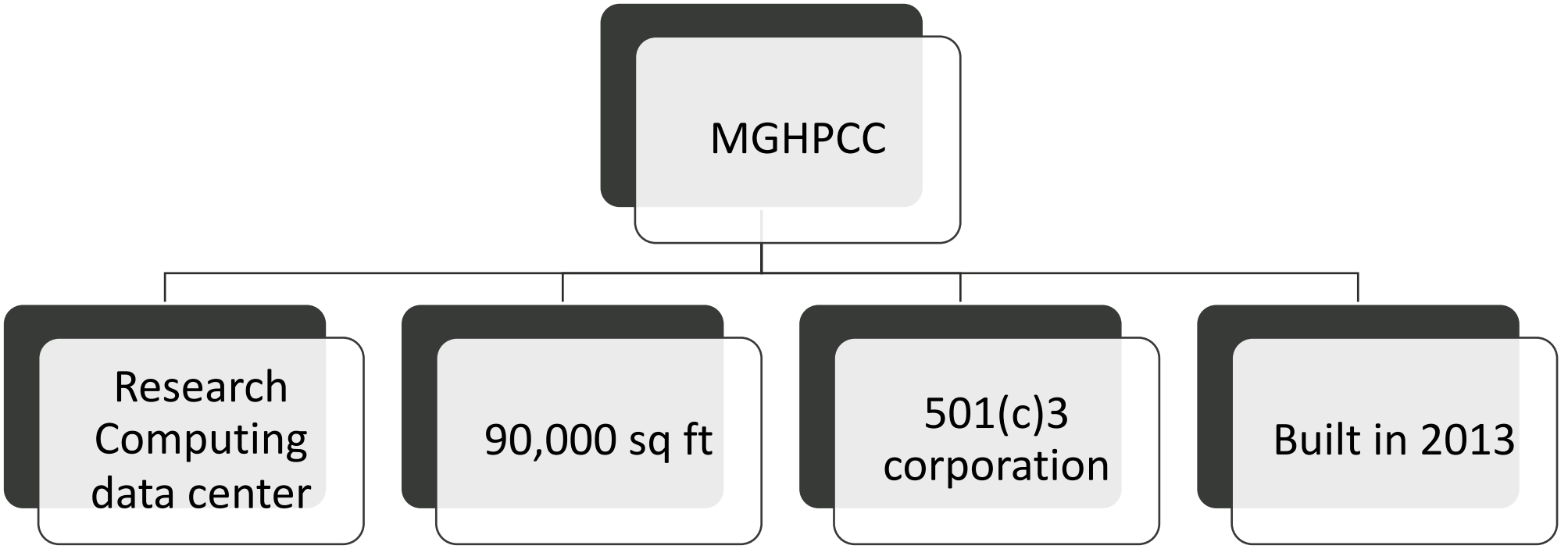
Massachusetts Green High Performance Computing Center (MGHPCC)

https://en.wikipedia.org/wiki/Holyoke_Dam#/media/File:Holyoke_Dam_during_thaw,_2018.jpg

The MGHPCC provides space, power, and cooling capacity to support over 680 racks (80 racks for UMass) of computing and storage equipment, drawing up to 15MW of power.

The MGHPCC facility was designed and built to be a leader in green computing and has been awarded LEED Platinum status.

MGHPCC Power Source

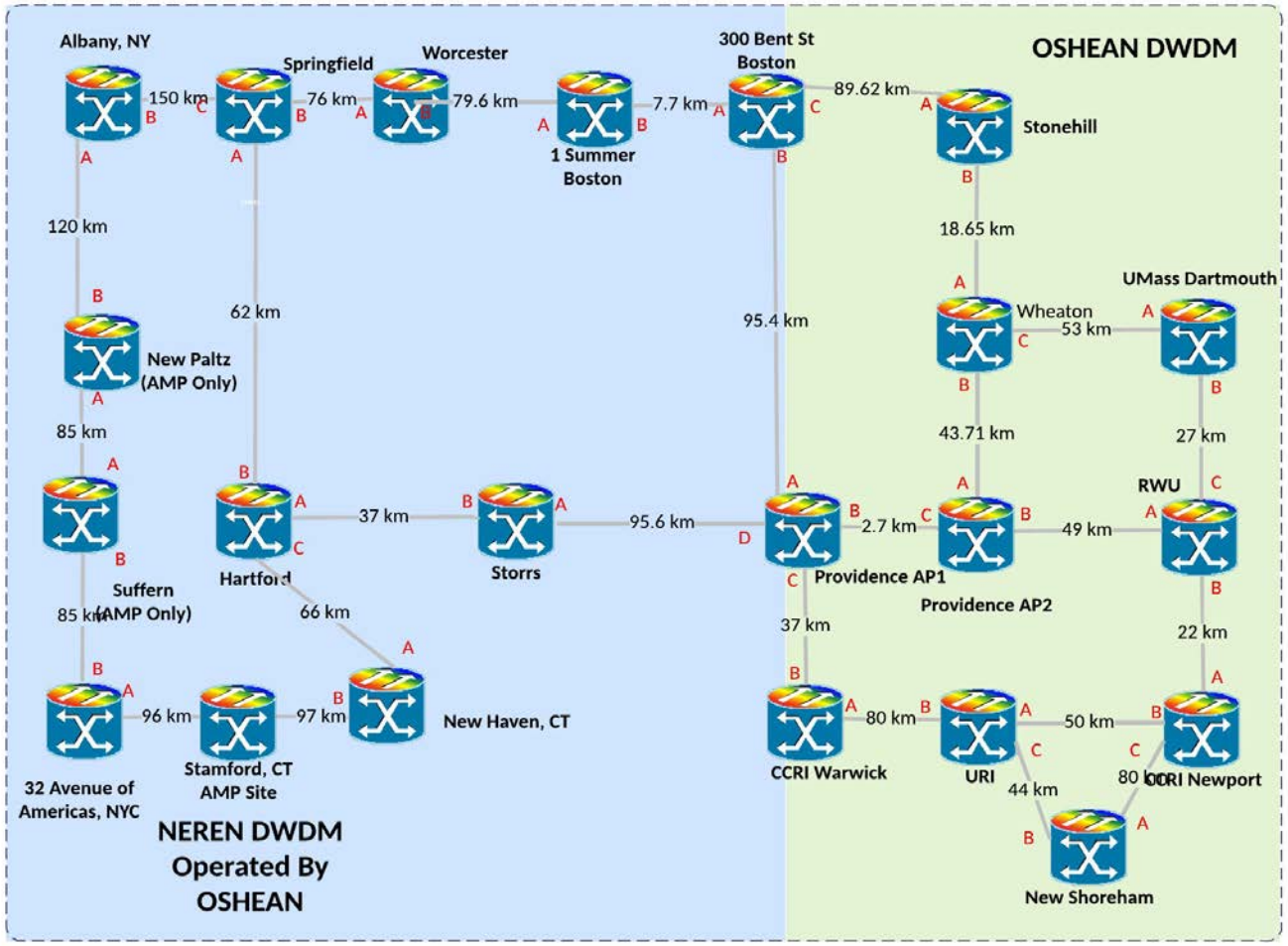


Massachusetts Green High Performance Computing Center

- **160 Member REN in RI**
- **HE, Healthcare, K12, Libraries, State and Local Gov.**
- **Segment Routed 10-100Gbps Packet/Optical backbone (250+nodes)**
  - Heavy use of I2 Cloud Exchange, Private Peering and HPC Resource Access
- **NEREN Fiber Mesh Backbone provides private access across the northeast**
  - NE and NY collaborative
  - I2 access in Albany and NYC
  - Colo Suite in Boston – Building an OSN PoD there this week
    - NSF's OSN PoD will compliment the NESE storage used today
  - Looking to add MGHPCC node; 400G Flexgrid

Who is OSHEAN?

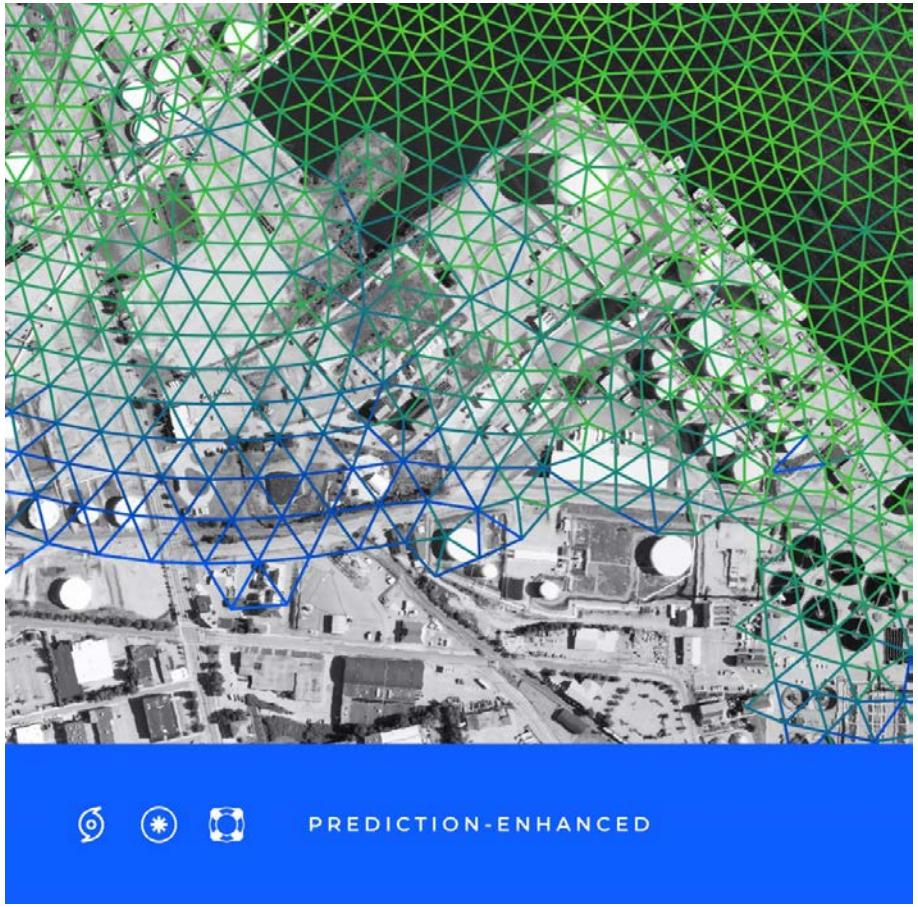
- **REN's and MGHPCC in the NE are highly collaborative for Research solutions**
  - CEN, OSHEAN, NYSERNET, UMaine, UMass
  - UMaine CC* 400Gbps project to connect to MGH
  - Fabric node at UMass
  - NSF Open Storage Node on the OSHEAN backbone
- **Collaborative with Ecosystem for Research Networks (ERN)**
  - Advanced Research techniques
    - Remote Access for high end research endpoints (i.e. Cryo-EM)



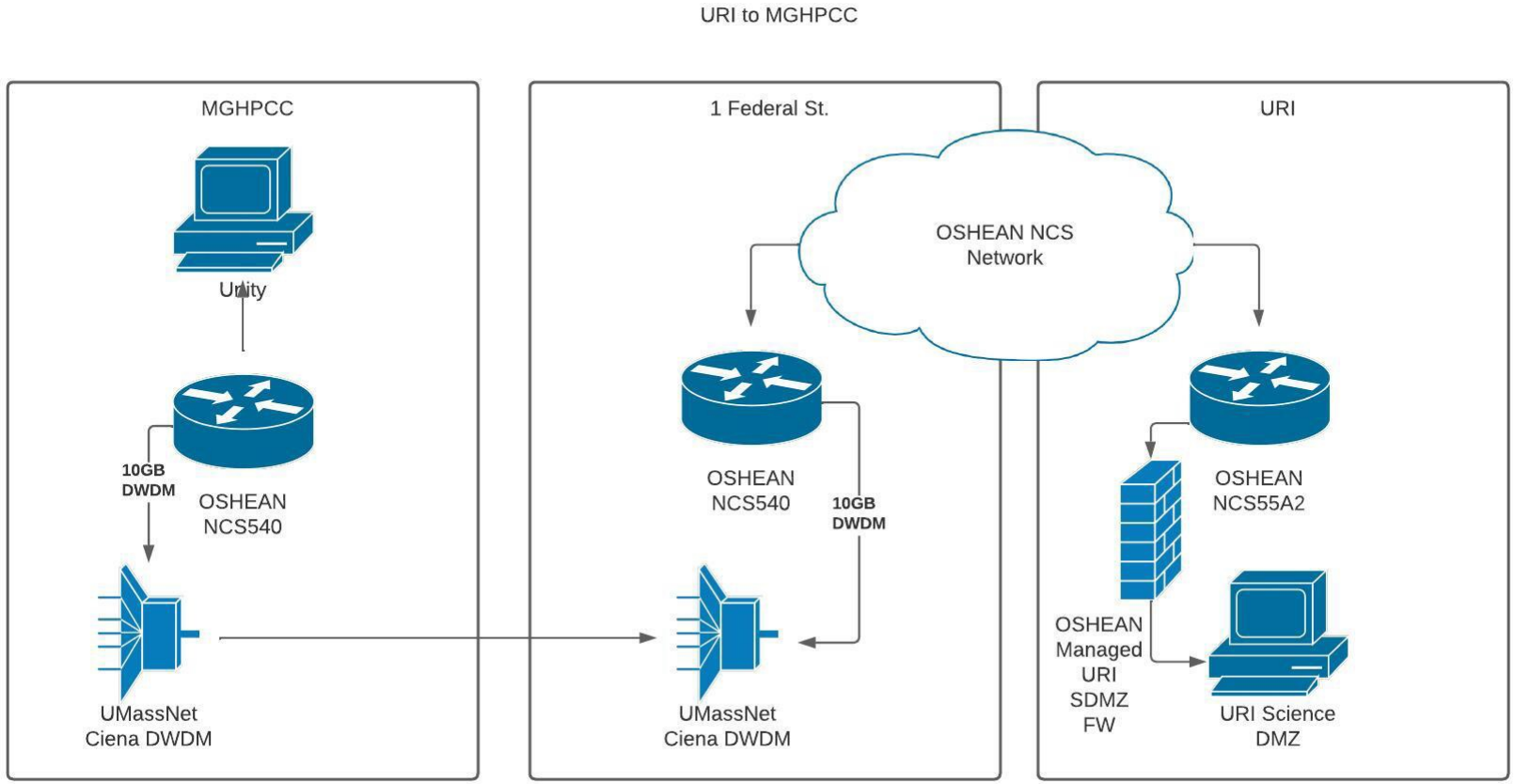
Working together

RICHAMP is a production service led by URI that provides near-real-time storm surge / flooding predictions for the entire New England south-coast covering CT, RI and MA in the event of a major storm.

- This data is used by Emergency Management in CT, RI and MA to plan evaluations and responses.
- Unity is the primary computational resource for RICHAMP



Use Case: Rhode Island Coastal Hazards, Analysis, Modeling & Prediction



URI to MGHPPC Topology

Research Facilitation Matters

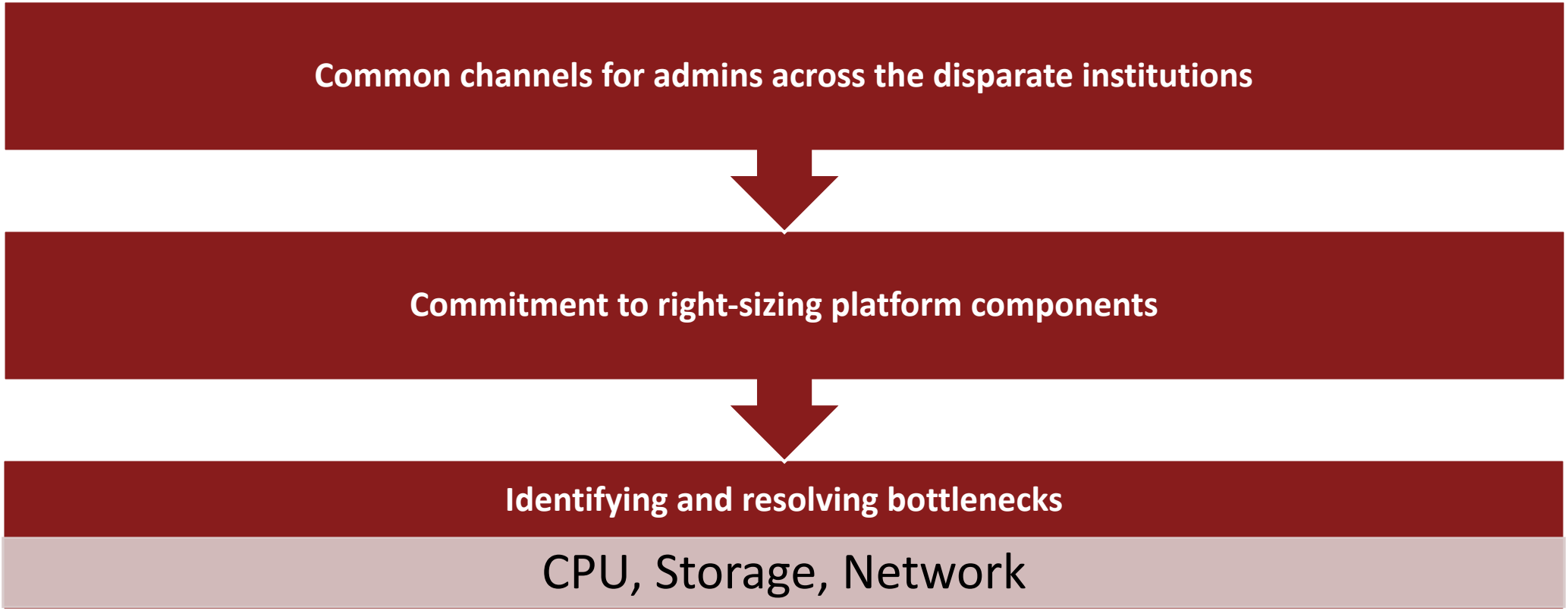
Commitment to open-source, whenever possible

New GitLab (using CILogon)

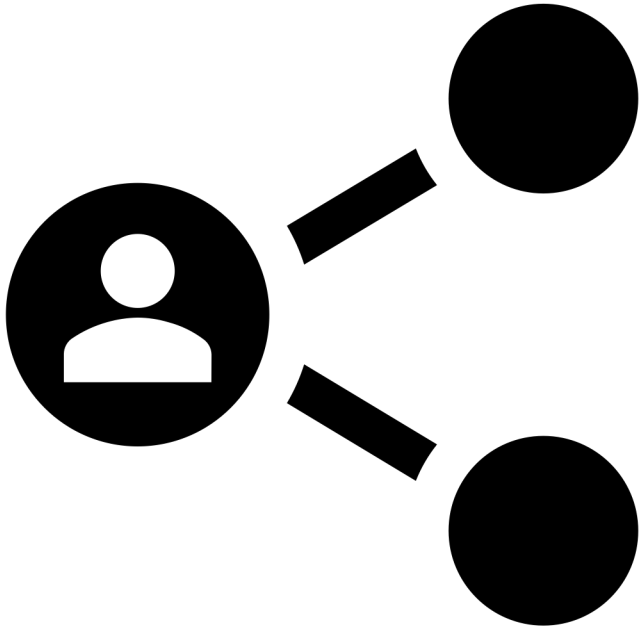
Building a blueprint for future successes

Contributing back to the community

What next?


Common channels for admins across the disparate institutions
Commitment to right-sizing platform components
Identifying and resolving bottlenecks
CPU, Storage, Network

What helped?



- Built on Trust
- Shared Staffing Model
- Strong Governance

Takeaways

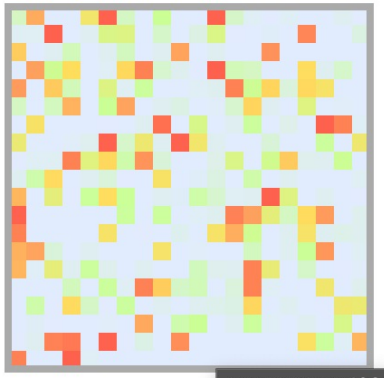
Appendices

Overview of Unity @ 2023-05-01 17:11

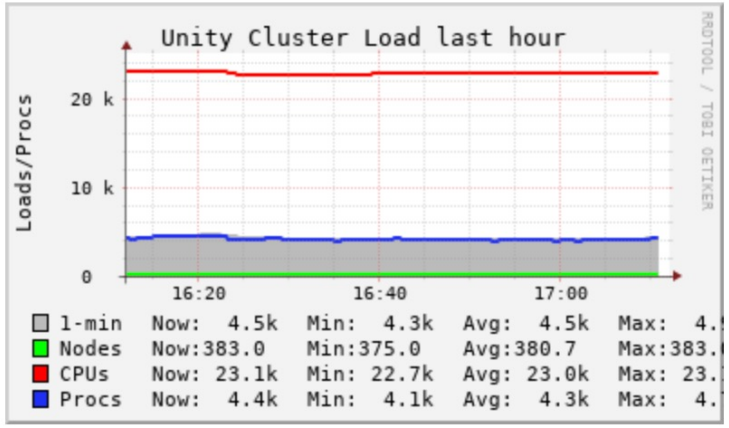
CPUs Total: **23064**
Hosts up: **385**
Hosts down: **0**

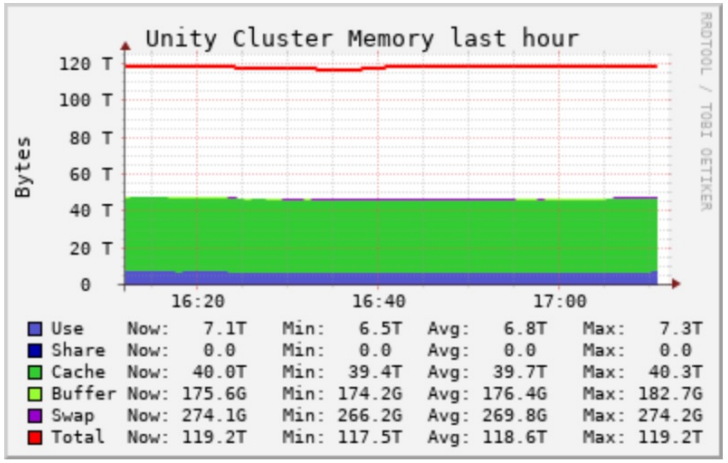
Current Load Avg (15, 5, 1m):
19%, 19%, 19%
Avg Utilization (last hour):
19%

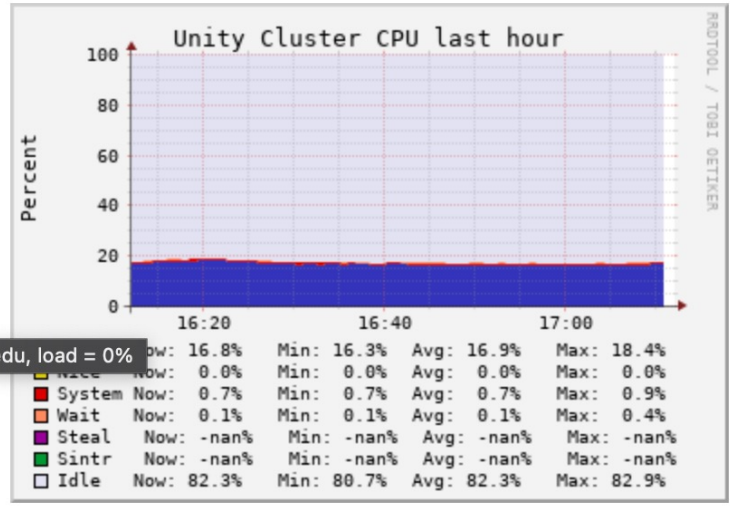
Server Load Distribution

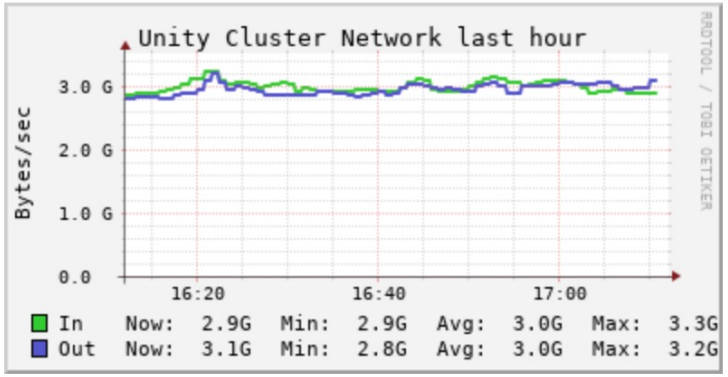


gypsum-gpu122.unity.rc.umass.edu, load = 0%







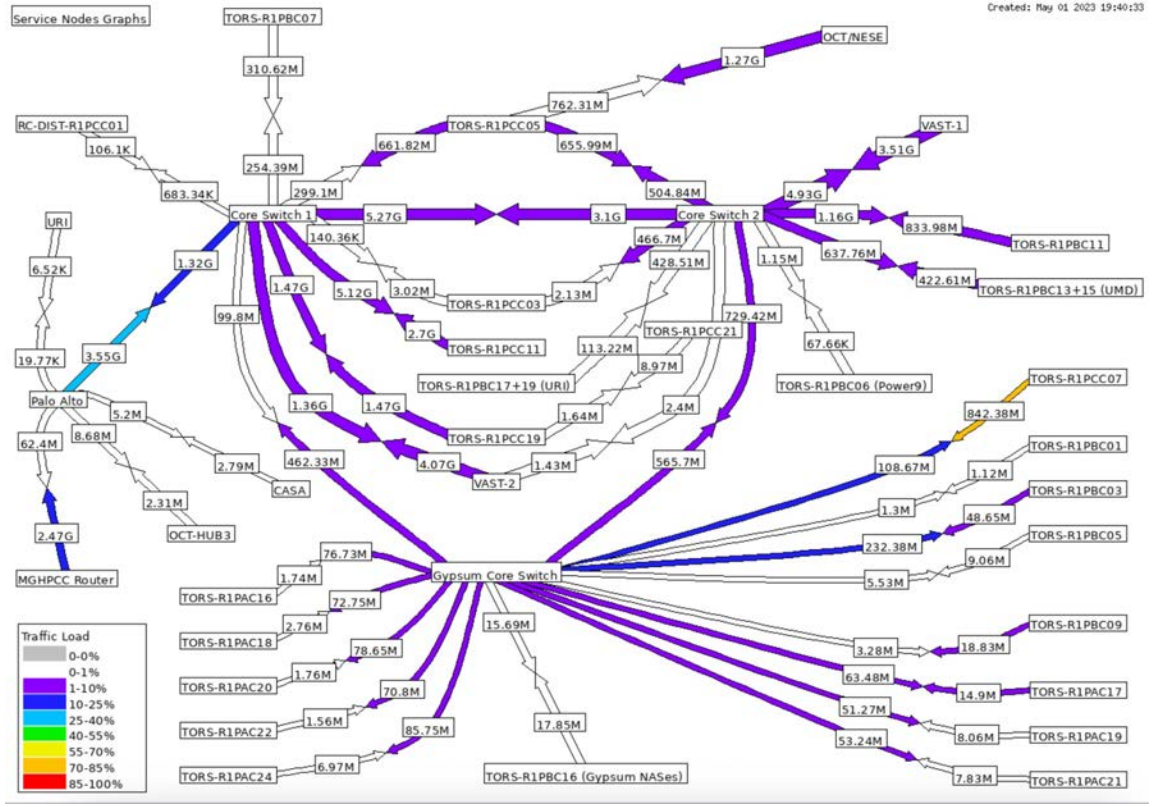


Unity Metrics and Measurement

The facility network needs some work

Active project to review and assess options

Aggregation of two clusters (Unity and Gypsum)

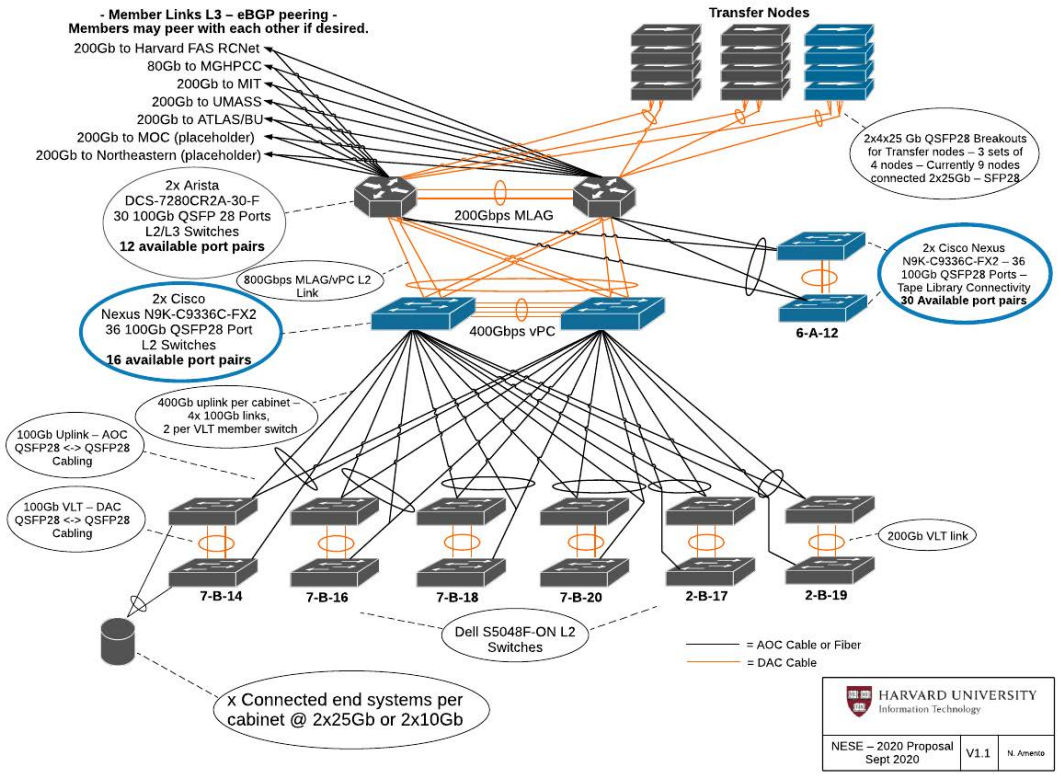


Unity Facility Network

NESE is the Northeast Storage Exchange, a ***shared regional storage collaboration*** funded by the National Science Foundation

- Operated as a long term partnership between the Universities and institutions below.

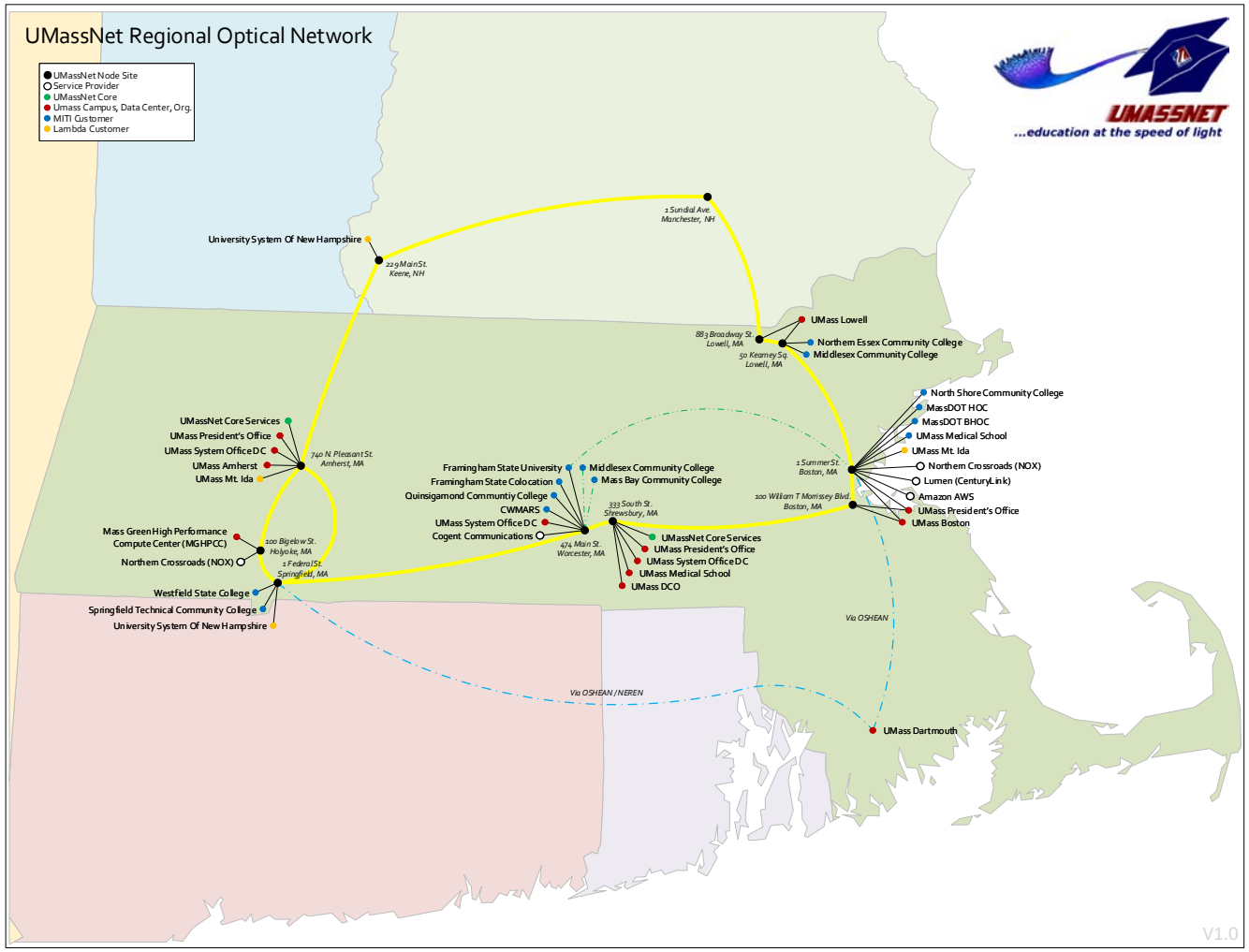
Our main goals are to meet the storage needs of the data revolution for science, engineering, education and technology, particularly for researchers in the northeastern part of the U.S.



NESE Topology

- Active job viewer
- Job builder/template
- File Explorer and text editor
- Desktop environment
- JupyterLab
- Rstudio
- Matlab
- Mathematica

Open OnDemand



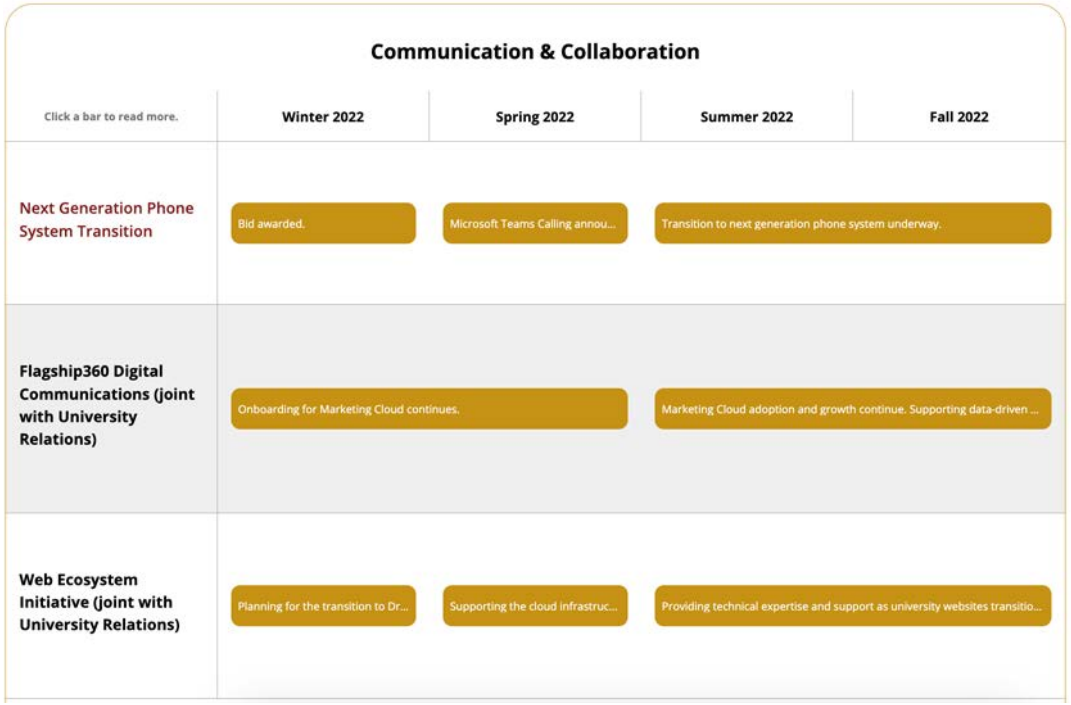
Internet/Internet2 connections

- UMass campuses
- 5 Colleges
- Several State and Community Colleges
- MGHPCC
- Participant in regional efforts (NEREN, etc)

UMassNet

- **An agile technology portfolio** to support on-site and remote delivery of courses, research, and university operations.
- Technology to **build a diverse, accessible, and inclusive campus**.
- **Data-driven IT service management**.
- **Operational excellence**, continuous improvement & digital transformation.

https://www.umass.edu/it/strategic-plan-2021/initiatives



IT Strategic Plan Principles

University of Massachusetts Amherst