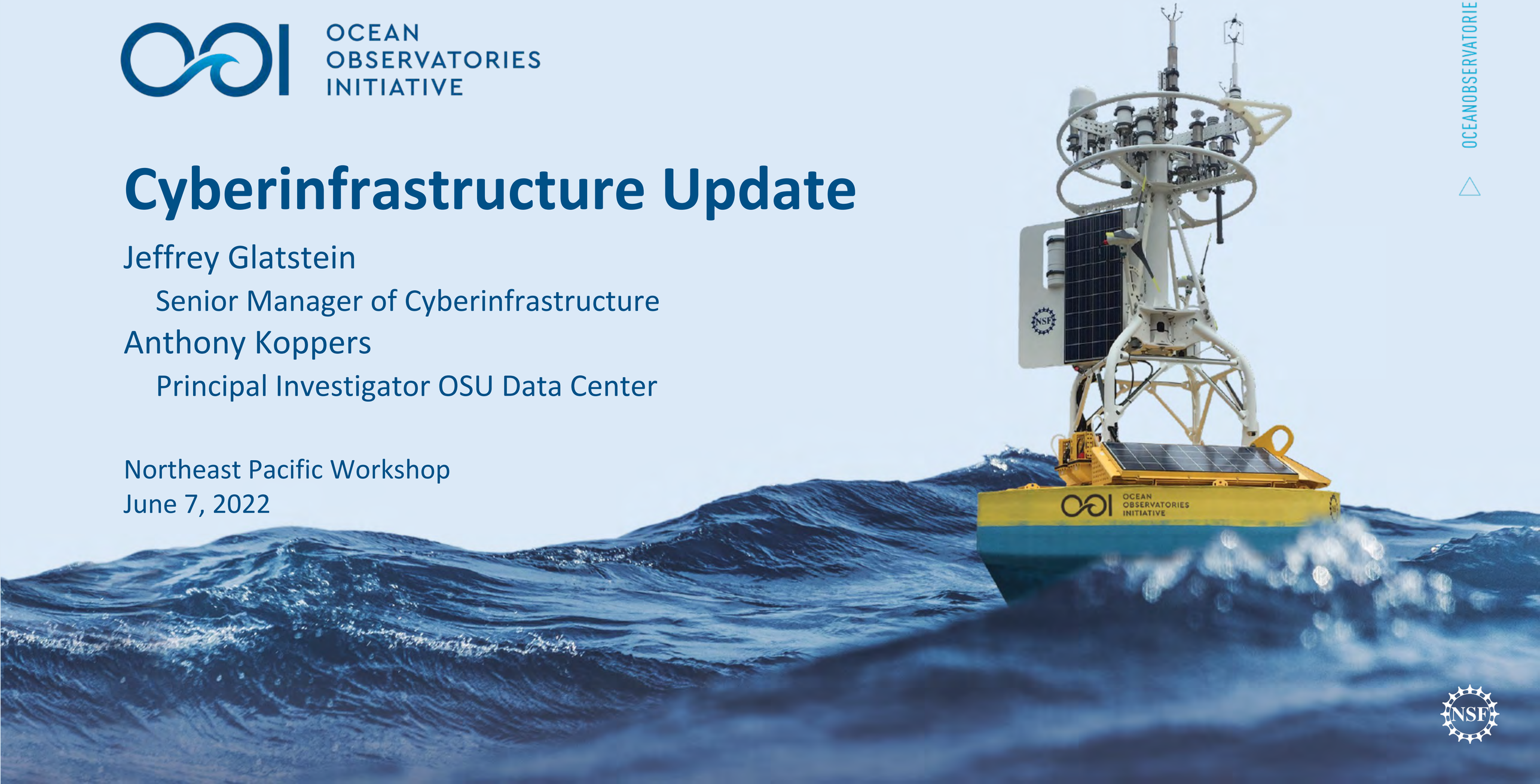
OCEAN OBSERVATORIES INITIATIVE

Cyberinfrastructure Update

Jeffrey Glatstein
Senior Manager of Cyberinfrastructure
Anthony Koppers
Principal Investigator OSU Data Center

Northeast Pacific Workshop
June 7, 2022



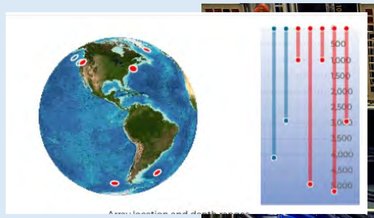
OCEANOBSERVATORIES.ORG

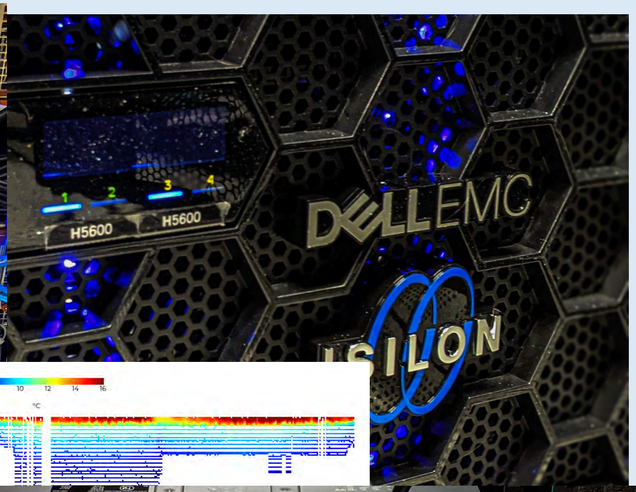
NSF

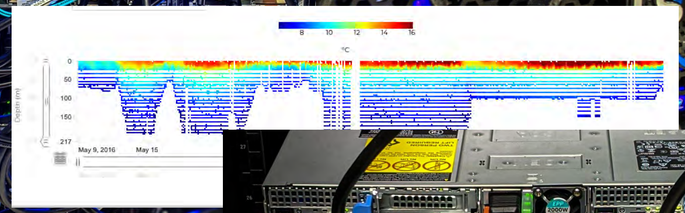
Agenda

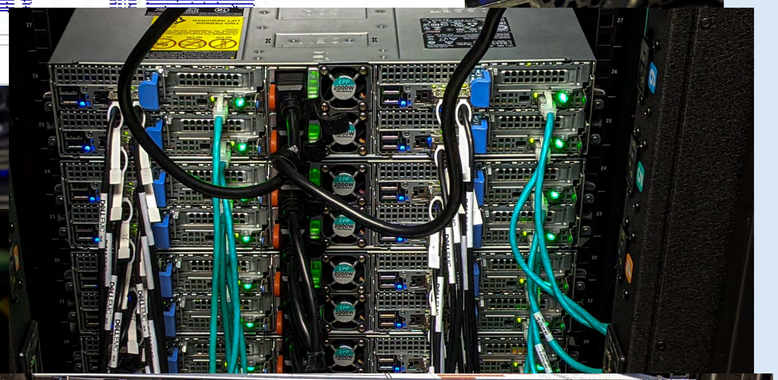
Discuss what is new in the cyberinfrastructure area:

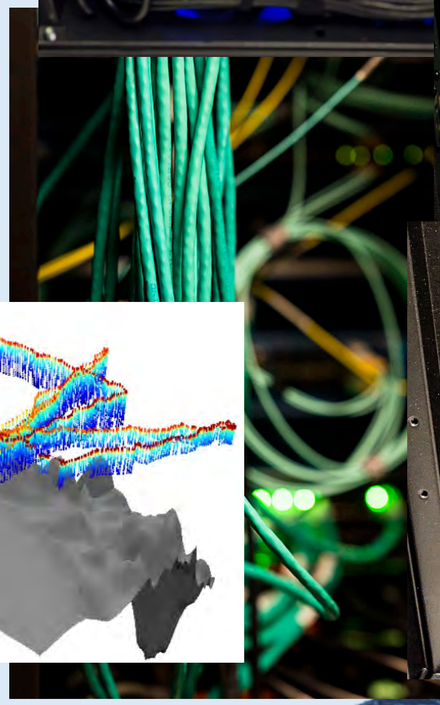
- OSU Data Center
- Cybersecurity & Data Protection
- System Monitoring
- New Data Explorer Features
- How to Get Help
- Data Interfaces
- JupyterHub Demo

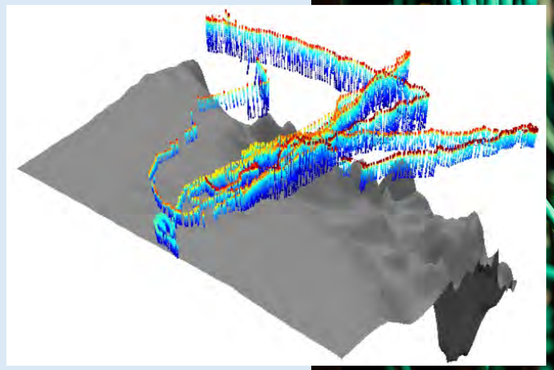








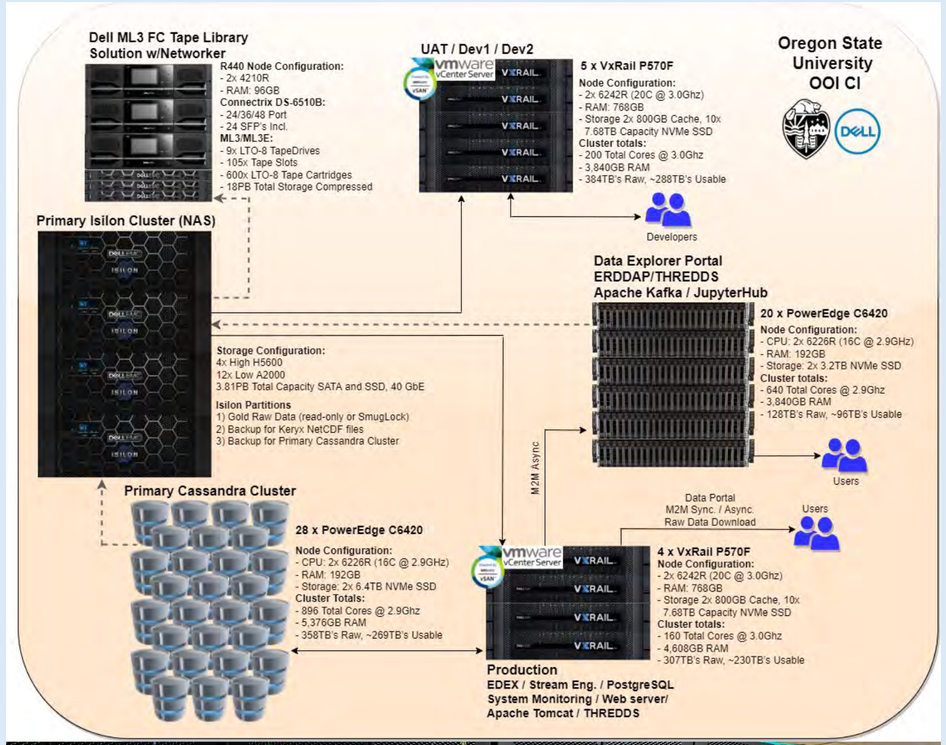


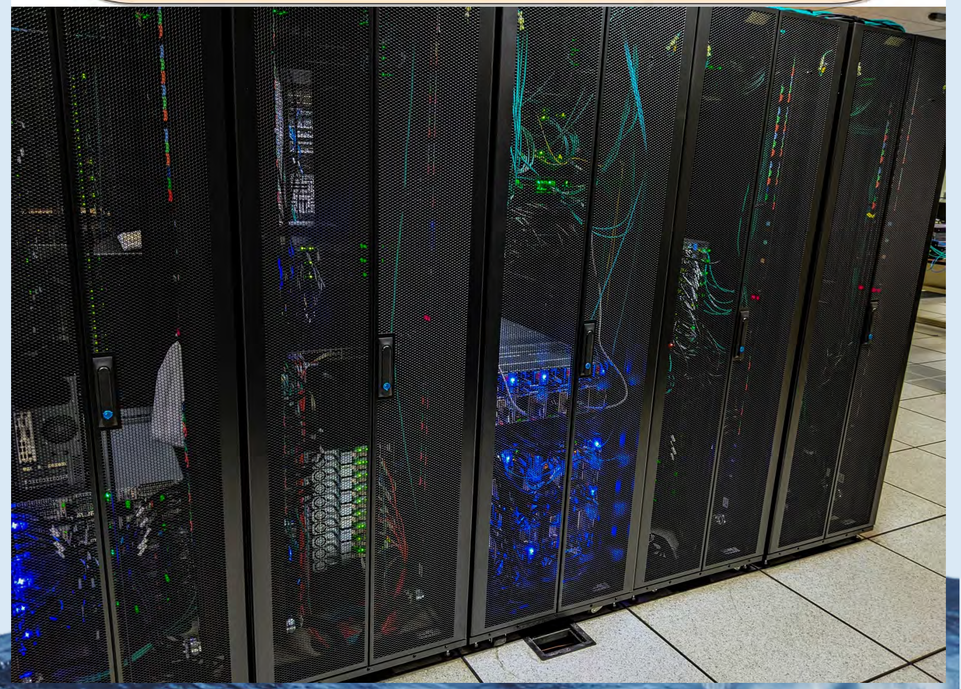


OSU Data Center

- OOI system of record since July 30, 2021
- Isilon Storage: 3.81PB
- VxRail Compute: 360 Cores; 8.4TB RAM; 684TB Storage

  194 Virtual Machines (Prod, UAT, Dev1, Dev2)
- DataExplorer / JupyterHub: 640 Cores; 3.8TB RAM; 128TB Storage
- Prod. Cassandra Cluster: 896 Cores; 5.4TB RAM; 360TB Storage
- 2 x Palo Alto Networks Next-Gen Firewalls
- 14 x Dell 100GbE PowerSwitches





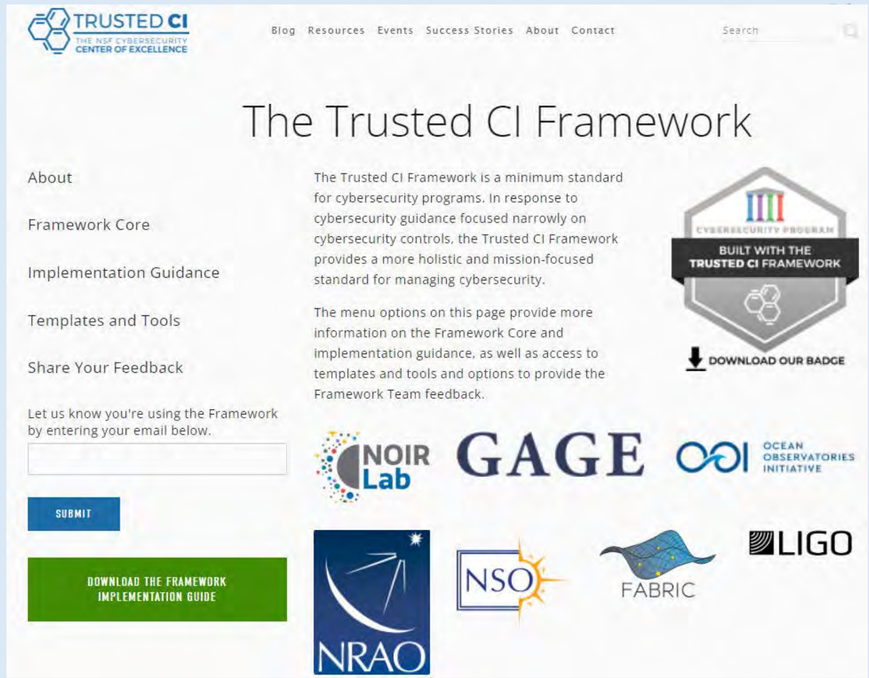
OCEAN OBSERVATORIES INITIATIVE

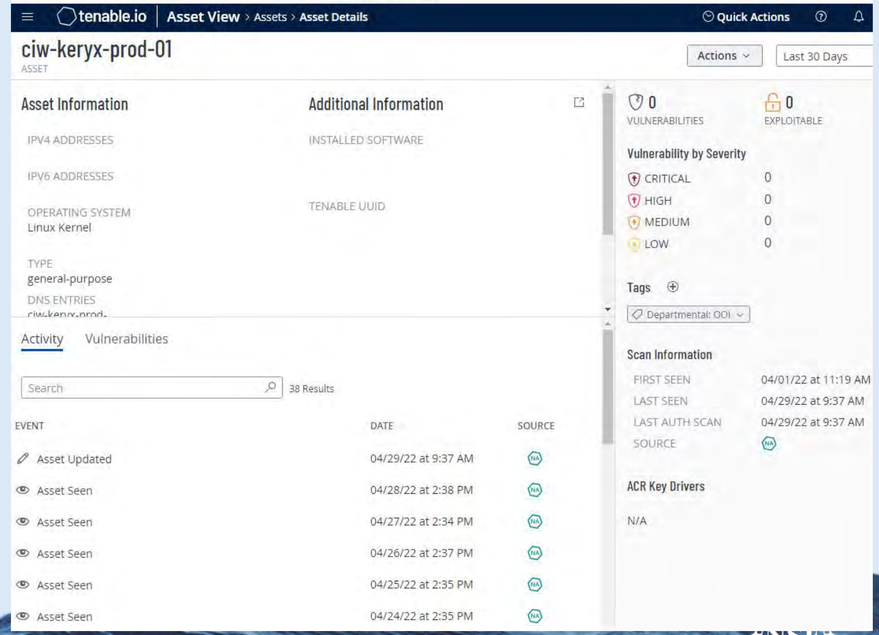
Cybersecurity

- Working with Trusted CI (TCI) to adopt and implement the TCI Cybersecurity Framework.
- Duo Multi-Factor Authentication (MFA) for Virtual Private Network (VPN) connections to OOI-CI Palo Alto firewalls.
- Implementing Center for Internet Security (CIS) Critical Security Controls.
- Internal and external vulnerability scanning using Tenable.io

Data Protection

- Isilon data protection level is set to +3d:1n1d, meaning that the cluster can survive three simultaneous disk failures or one entire node failure plus one disk.
- Daily data backups to a US East Coast Faction Inc. data center and to local LTO tape.

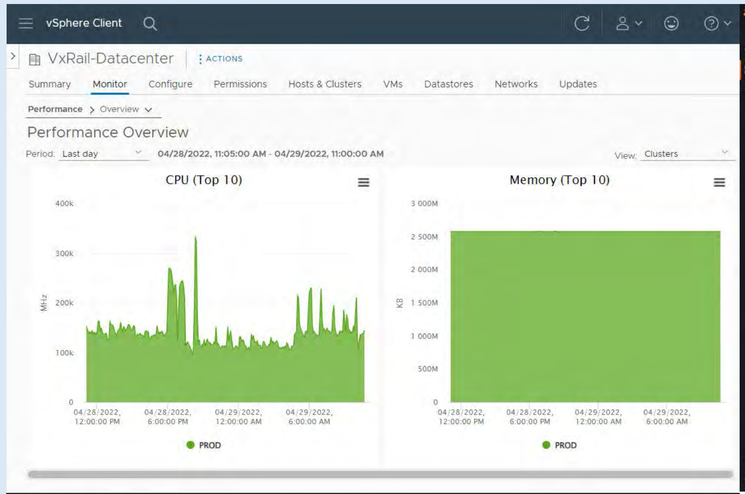


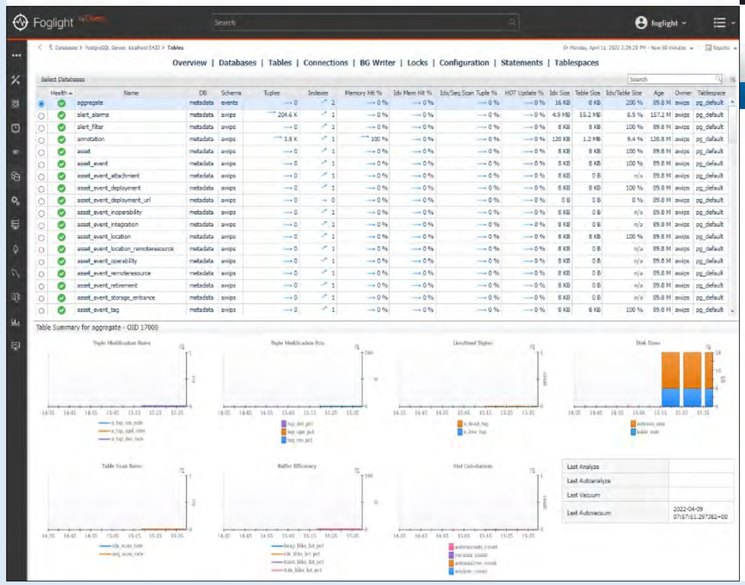


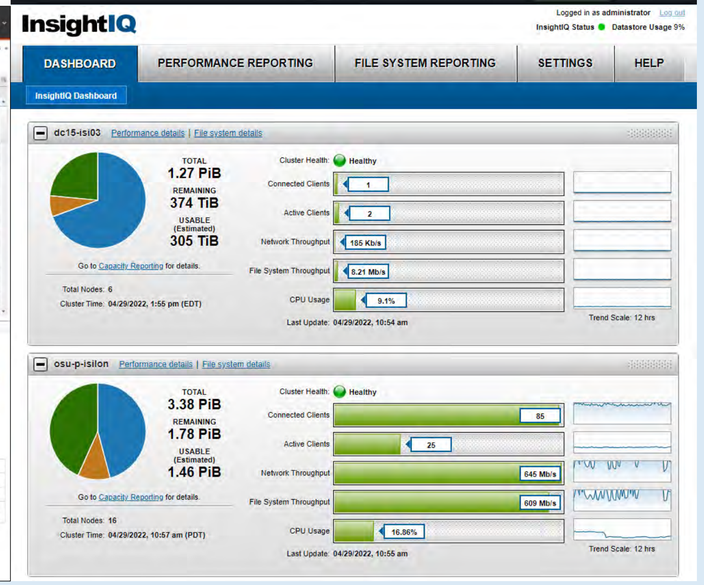
OCEAN OBSERVATORIES INITIATIVE

System Monitoring

- Zabbix, Grafana, and Nagios for overall data center monitoring.
- Quest Toad Edge and Foglight for database management and monitoring.
- InsightIQ for Isilon storage performance monitoring and reporting.
- VMware vSphere for VxRail management and monitoring.
- Panorama for firewall management and monitoring.





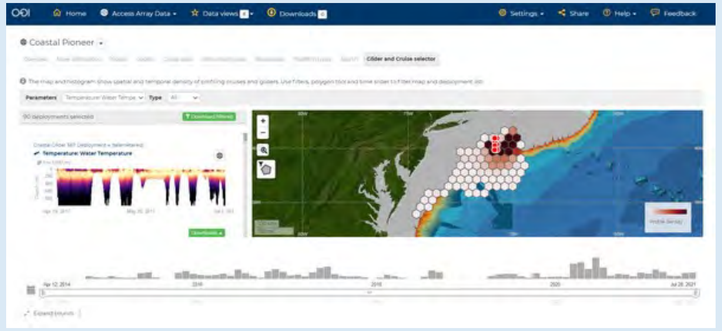


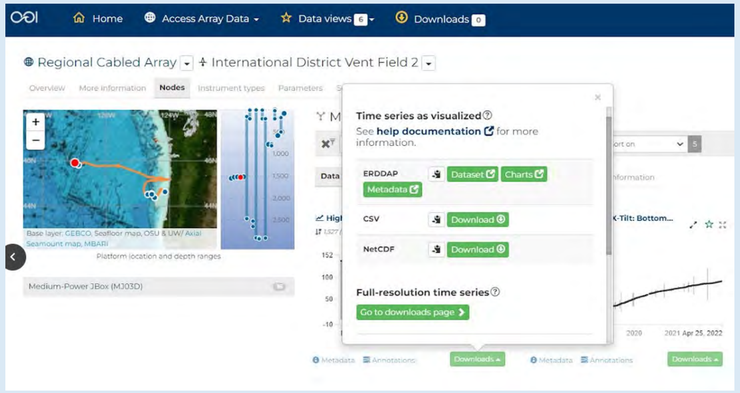
OCEAN OBSERVATORIES INITIATIVE

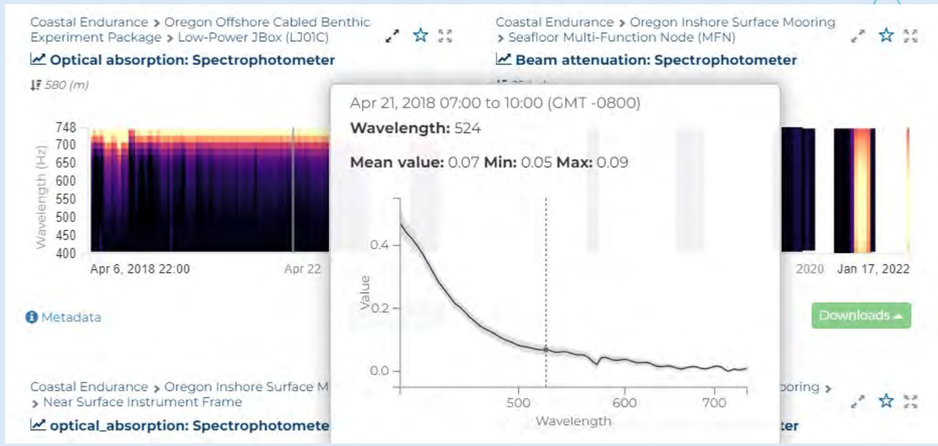
NSF

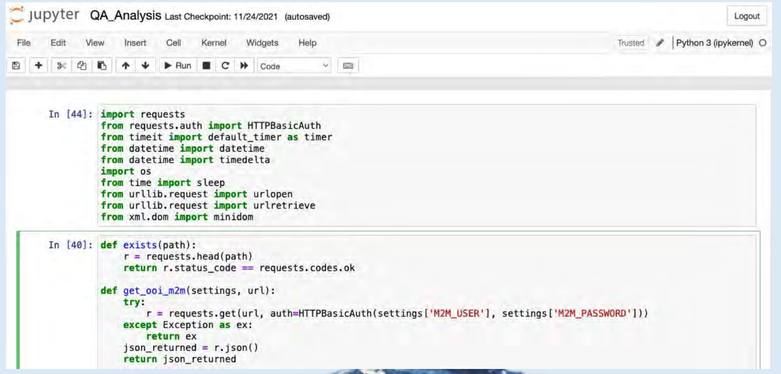
New Data Explorer Features

- Discrete Sample data available
- Glider and Discrete data search grid
- Custom polygon search capability helps visually find and compare instrument measurements to profiles collected by gliders and cruises near to instrument location
- Inset map displays data location within profile
- Clarification on data download source options
- Interactive plotting of OPTAA data
- Significant backend performance and stabilization updates









Coming Soon…

- Media server – HD pictures and videos
- Realtime data plotting
- Full resolution data visualization
- JupyterHub

OCEANOBSERVATORIES.ORG

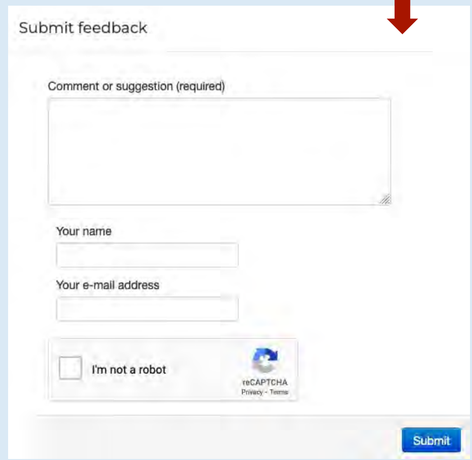
OCEAN OBSERVATORIES INITIATIVE

NSF

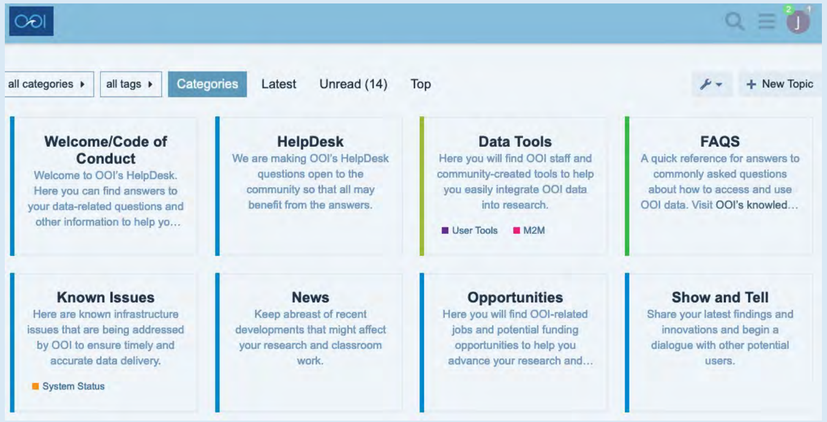
How to Get Help

- Data Explorer Help provides documentation and an interactive guide
- Data Explorer Feedback allows submission of text for feedback and/or issues
- Help and other topic are on the OOI Discourse site: https://discourse.oceanobservatories.org/ This the recommended method to search, comment and post on topics for Help and others of interest such as Data Tools and Known Issues
- OOINET and OOI queries, comments or issues use helpdesk@oceanobservatores.org. Links can be found on OOINET and the OOI main website

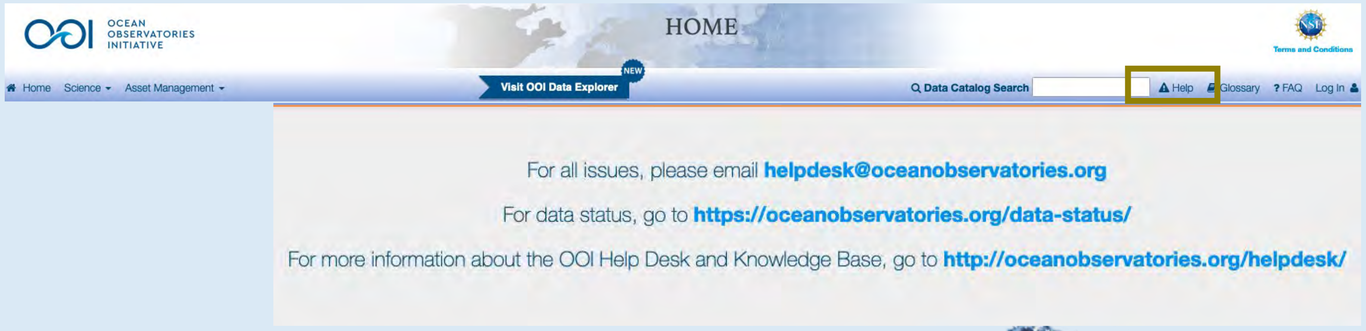
Data Explorer



Discourse site



OOINET



OOI Data Bus

More information and URLs can be found at: oceanobservatories.org/how-to-access-data/

Data Type / Source ● Current best method ○ Coming soon ◐ Phasing out	**Raw Data Archive** Raw instrument and engineering datasets presented in an Apache file system structure for download.	**Thredds** Pre-computed scientific numerical data products with calibrations applied alongside engineering data. Full resolution datasets are accessible by deployment and stream.	**Data Explorer** Primary gateway to visualize and access OOI data. Search across data points, download full datasets using ERDDAP, compare datasets across regions and disciplines, and generate shareable custom data views.	**ERDDAP** Underlying data server for Data Explorer providing access to ~600 datasets organized by OOI arrays. Download datasets in common file formats and make graphs and maps.	**M2M** Access to science and engineering data using both synchronous and asynchronous interfaces. NetCDF and JSON files are the standard outputs.	**OOI Website** Oceanobservatories.org provides access to datasets compiled by Principal Investigators who have added instruments onto OOI arrays.	**Alfresco** Document repository for instrument vendor information including calibrations	**Jupyter Hub** Hosted by OOI, this hub provides access to full resolution datasets and raw data server, allowing users to share notebooks and process data in a larger server environment.	**Data Explorer Media** Preview HD Photo, HD Video, visualized Hydrophone and ZPLS data along side science data in the Data Explorer tool	**OOINET** Legacy access to scientific and engineering data with the ability to search and plot data for review. Download requests are queued for system processing. User will be notified when download is ready. This interface will be slowly phased out.
Numerical Raw Data	●		○					○		
Processed Data Sets (NetCDF)		●	●	●	●			○		◐
Provenance data (JSON)		●	●		●			○		◐
Asset information						●	●			
Hydrophone	●							○	○	
Realtime data plots			○							◐
ZPLS data	●							○	○	
PI Data						●				
HD Video	●							○	○	
HD Photo	●							○	○	

JupyterHub Demo

Presented by Will Koeppen

Axiom DATA SCIENCE

Notebooks authored by:

Axiom Data Science
Jim Case (Case Ocean Services)
Chris Wingard (OSU)

jupyter Notebook

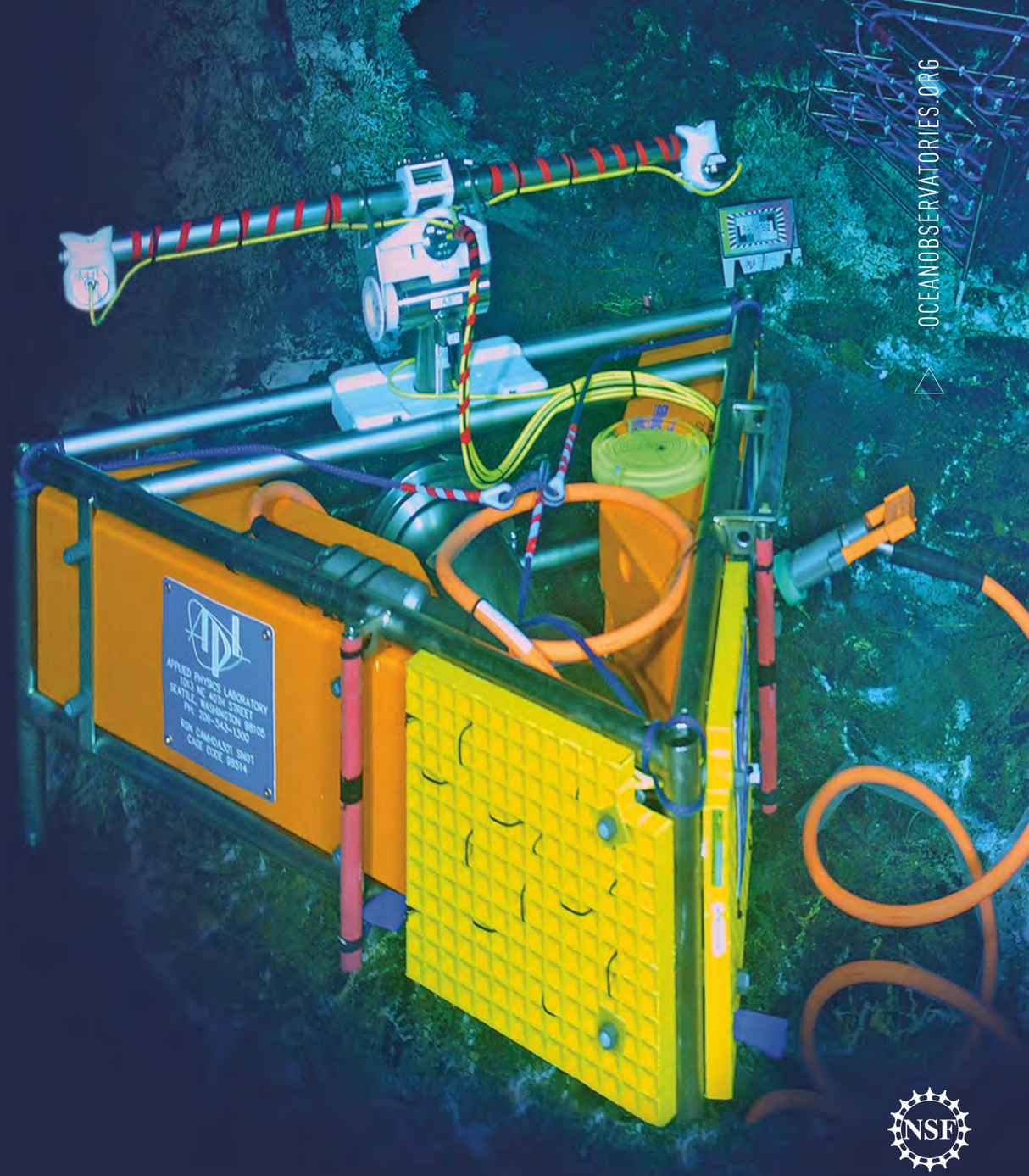
OCEANOBSERVATORIES.ORG

OCEAN OBSERVATORIES INITIATIVE

NSF

OCEAN
OBSERVATORIES
INITIATIVE

Questions?



OCEANOBSERVATORIES.ORG

NSF