OCEAN OBSERVATORIES INITIATIVE

OCEANOBSERVATORIES.ORG

CI Systems PYV Accomplishments and PYVI Planned Activities

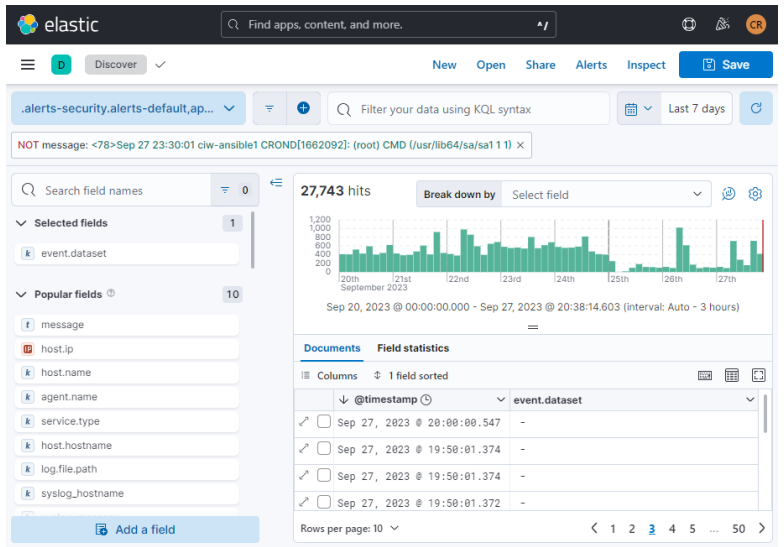
Anthony Koppers (PI)
Craig Risien (Co-PI, PM)
Jim Housell (IT Architect)
Casey Dinsmore (DevOps)
Pei Kupperman (Finance)

Fall OOIFB Meeting
October 12, 2023

PYV Work Plan Highlights – Presented 10/2022

Enhance Cybersecurity
- ✓ Continue working with Trusted CI CoP
- ✓ Apply appropriate CIS v8 controls
- ■ Endpoint management solution to help us to efficiently patch, manage, and secure our system
- ✓ Log monitoring and analysis (ELK Stack)
- ■ Develop Incident Response, Disaster Recovery & Acceptable Use policies

✓ Release JupyterHub
✓ Complete initial NOAA-NCEI and ■ TACC data archiving
✓ Upgrade Cassandra to version 4.0
✓ Add virtualized Data Explorer development env
✓ Ongoing system maintenance and support
✓ *Planning for OOI 2.5 Data Center refresh*



OCEANOBSERVATORIES.ORG

OCEAN OBSERVATORIES INITIATIVE

OOI 2.0 - OSU Data Center Goals

- Provide a **low-risk** and **cost-effective** OOI Data Center
- Introduce large **increases** in **compute power**, **modernize storage** solutions, and **improve backup** and disaster recovery
- Provide a **secure data store** with multiple layers of redundancy to significantly **reduce system downtime**
- Achieve a **seamless transition** from the OOI-CI operations at Rutgers University to OSU in Year 1
- Focus on **extensibility** in Years 2 & 3 by considering both bare-metal and/or cloud-like solutions based on OOI usages and needs

OOI 2.0 - Data Center Improvements

- More than **doubled** the **storage** capacity
  - 3.8 PB Isilon storage
- **Network** speeds **2.5 times faster**
  - 25 GbE switches
- **Improved Cassandra** database **speed and stability**
  - Upgraded to Apache Cassandra 4.0
  - Increased the number of nodes to 28
  - Added memory, CPUs, and solid-state storage
- **Increased fault tolerance, reduced system downtime**
  - Virtualized Uframe infrastructure
  - 250 VMs deployed on 12 VxRails running VMware
  - Dedicated VMs running specific services

Isilon storage nodes (16)

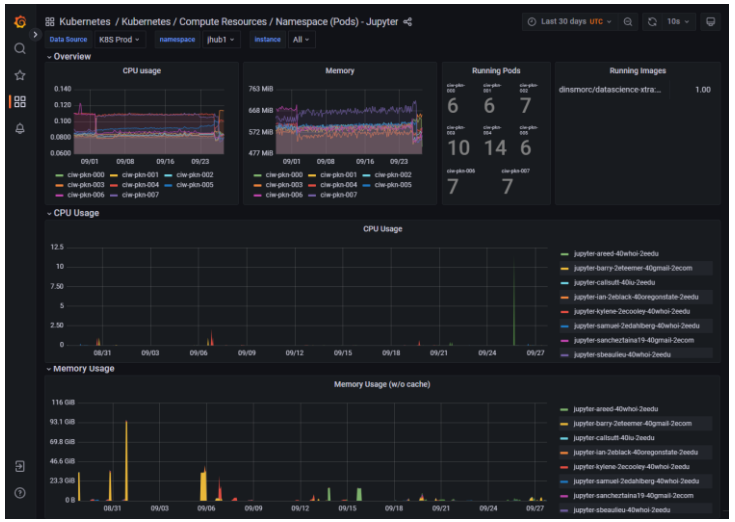
C6420 Cassandra nodes (28)

VxRail compute nodes (12)

OOI 2.0 - Data Center Improvements

- **Improved cyber security**
  - PA-5250s NGFWs (Firewalls)
  - Duo MFA for VPN connections
  - Trusted CI engagements
  - Network segmentation
  - Placed all development envs. behind the firewalls
  - Regular vulnerability scanning
- Developed extensive **monitoring system**
  - Panorama, Nagios, Grafana, InsightIQ, etc.
- Implemented **cloud-like data access**
  - 8-node Kubernetes cluster running JupyterHub
  - 256 cores; 3 TB memory; 100 TB storage
  - Standard (4GB; 4CPUs) >> X-Large (128GB; 48CPUs) servers
  - Support for Python, R, and MATLAB
  - Ultrafast, local, read-only access to raw data, NetCDF files
  - https://jupyter.oceanobservatories.org

Grafana Dashboard



JupyterHub Login Screen

OCEANOBSERVATORIES.ORG

OCEAN OBSERVATORIES INITIATIVE

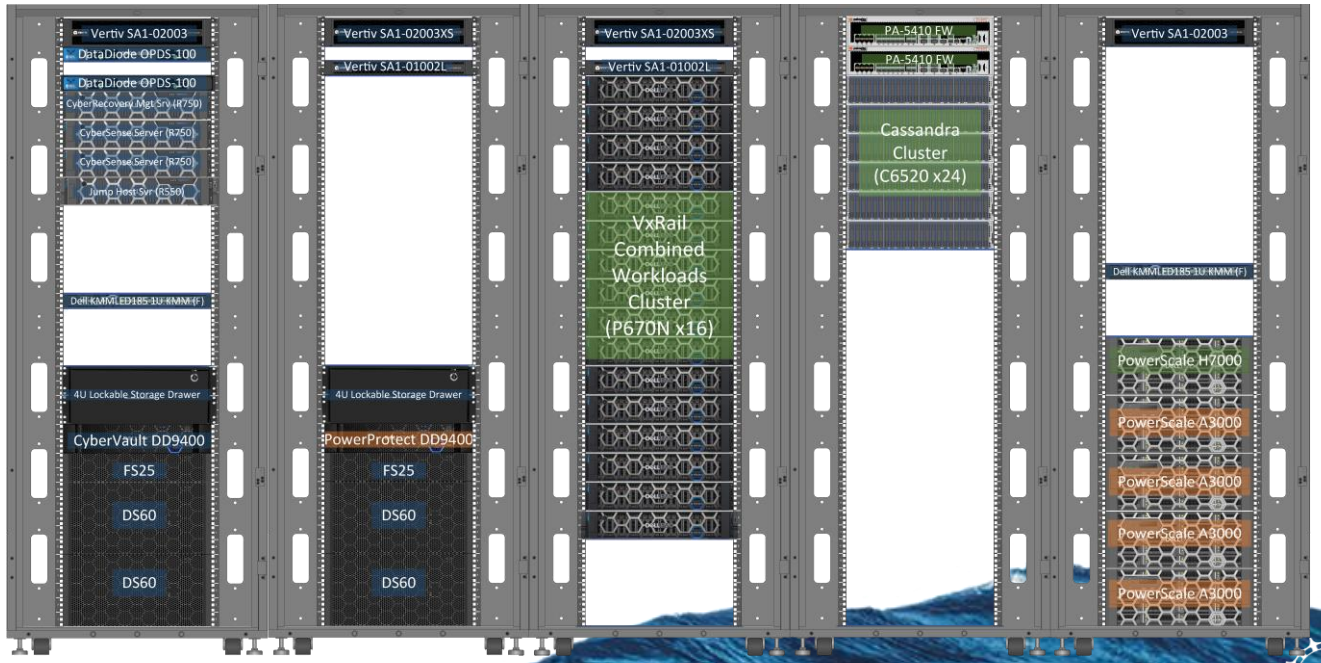
OOI 2.5 Planned Activities – Data Center Refresh (PYVI)

- **Doubling (again) the storage capacity**
  - 8 PB Isilon with +3d:1n1d protection
- **Increase Network speeds**
  - 100 GbE switches (4x faster than OOI 2.0)
- **Enhance security posture**
  - PA-5410 Series ML-Powered NGFWs
  - CyberSense and Ransomware Defender
- **Improve DR recovery posture**
  - PowerProtect DD9400
  - PowerProtect Cyber Recovery vault
  - Elastic Cloud Storage & GoldenCopy
- **Expand offsite DR data storage**
  - 8 PB Elastic Cloud Storage (2nd copy)
  - Texas Advanced Computing Center (3rd copy)

DELL Technologies

paloalto NETWORKS

superna



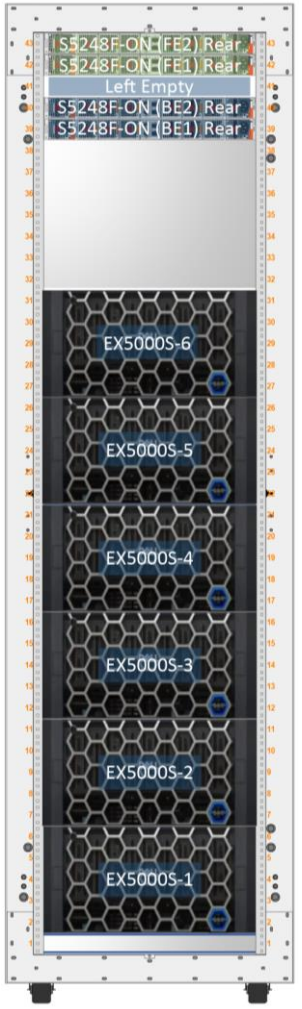
OCEAN OBSERVATORIES INITIATIVE

OCEANOBSERVATORIES.ORG

OOI 2.5 Planned Activities Continued

- Continue working with Trusted CI CoP
- Work with OOI CISO to develop Incident Response, Disaster Recovery and Acceptable Use policies
- Fully Implement log monitoring & analysis system
- Automate ~monthly data archiving to TACC
- Investigate upgrading to Cassandra version 5.x
- Create Enterprise Linux 9 software environments
- Develop a DR site in Bend, OR (more than 200 km away with Secure Internet2 Connection)
- Migrate Data Explorer to VxRail cluster
- Double Kubernetes/JupyterHub cluster.
  16-nodes; 512 cores; 6 TB memory; 200 TB storage
- Continue system maintenance and support

Elastic Cloud
Storage Cluster
CDDC – Bend, OR



jupyterhub

TACC

cassandra

OCEAN OBSERVATORIES INITIATIVE

Questions?

Anthony Koppers (PI)
anthony.koppers@oregonstate.edu

Craig Risien (Co-PI, PM)
craig.risien@oregonstate.edu

NSF