Oregon State University

**OOI-CI DATA CENTER**

# Toward an Extensible Cyberinfrastructure for the Ocean Observatories Initiative

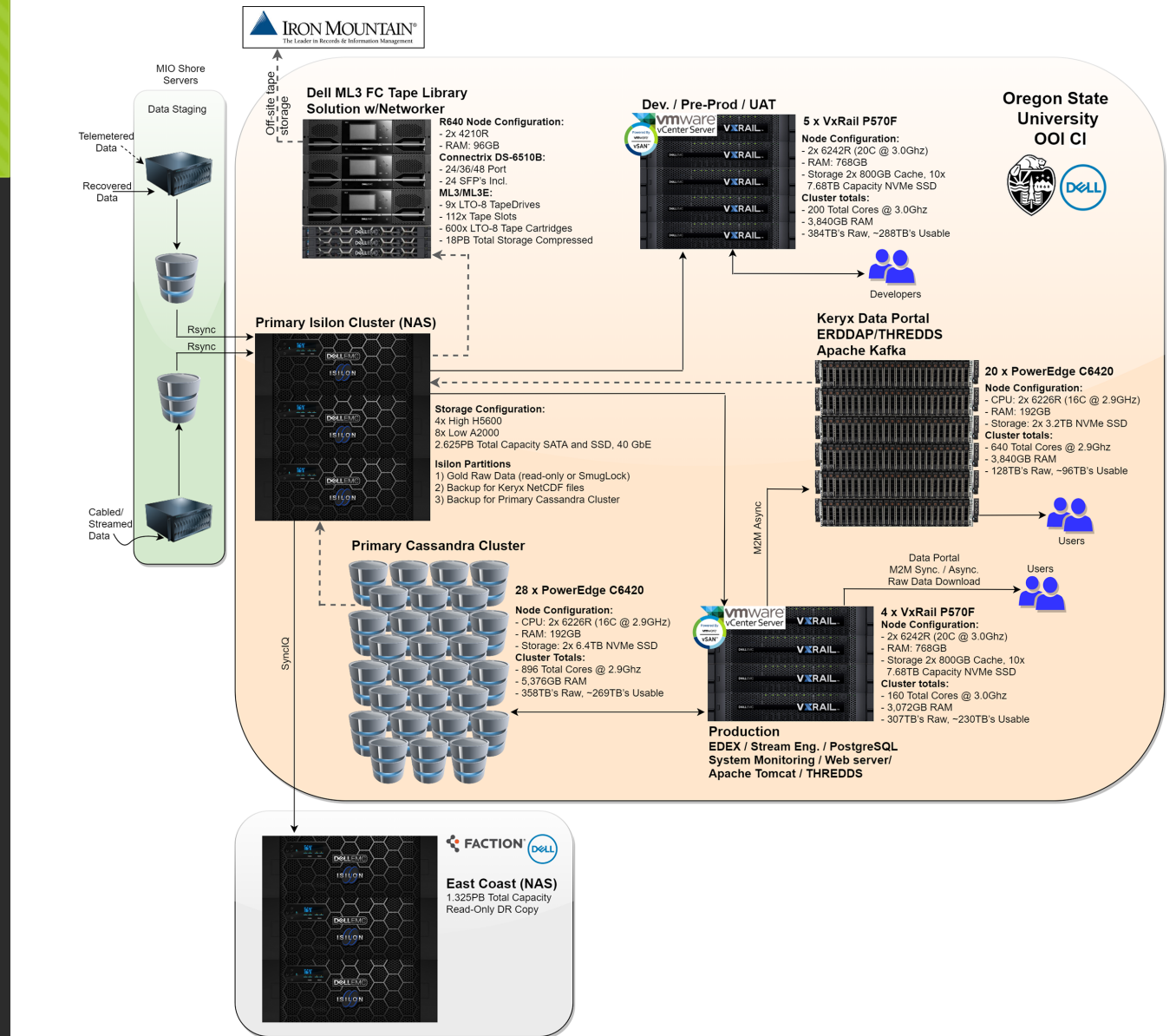**Anthony Koppers** (Lead-PI, Professor, Associate Dean for Research)

**Mark Keever** (Co-PI, Director of Research Computing)

**Craig Risien** (Operations Manager, OOI Endurance Array Engineer, Data Portal POC)

# Modernized, Simpler More Capacity, Faster

# New OSU Architecture

- Five Bare-metal Servers
  - Primary Isilon NAS (2.6 PB)
  - Prod VxRAIL VM Server (5 nodes)
  - Test/Dev VxRAIL VM Server (6 nodes)
  - Cassandra Cluster (28 nodes)
  - Keryx Data Portal (20 nodes)
- Secondary Isilon Server on Faction
  - East Coast Isilon NAS (1.3 PB)
  - Active:Active replica
  - Can serve as AWS read-only node
- Tape Library (19 PB)
  - Incremental and Snapshot Backup
  - Offsite storage

IRON MOUNTAIN®
The Leader in Records & Information Management

MIO Shore Servers

Data Staging

Telemetered Data

Recovered Data

Off-site tape storage

Cabled/ Streamed Data

Rsync
Rsync

SyncIQ

**Dell ML3 FC Tape Library Solution w/Networker**
R640 Node Configuration:
- 2x 4210R
- RAM: 96GB
Connectrix DS-6510B:
- 24/36/48 Port
- 24 SFP's Incl.
ML3/ML3E:
- 9x LTO-8 TapeDrives
- 112x Tape Slots
- 600x LTO-8 Tape Cartridges
- 18PB Total Storage Compressed

**Dev. / Pre-Prod / UAT**
vmware vCenter Server
VxRAIL.
**5 x VxRail P570F**
Node Configuration:
- 2x 6242R (20C @ 3.0Ghz)
- RAM: 768GB
- Storage 2x 800GB Cache, 10x 7.68TB Capacity NVMe SSD
Cluster totals:
- 200 Total Cores @ 3.0Ghz
- 3,840GB RAM
- 384TB's Raw, ~288TB's Usable

**Oregon State University OOI CI**
DELL

Developers

**Primary Isilon Cluster (NAS)**
DELL EMC ISILON

Storage Configuration:
4x High H5600
8x Low A2000
2.625PB Total Capacity SATA and SSD, 40 GbE

Isilon Partitions
1) Gold Raw Data (read-only or SmugLock)
2) Backup for Keryx NetCDF files
3) Backup for Primary Cassandra Cluster

**Keryx Data Portal ERDDAP/THREDDS Apache Kafka**
**20 x PowerEdge C6420**
Node Configuration:
- CPU: 2x 6226R (16C @ 2.9GHz)
- RAM: 192GB
- Storage: 2x 3.2TB NVMe SSD
Cluster totals:
- 640 Total Cores @ 2.9Ghz
- 3,840GB RAM
- 128TB's Raw, ~96TB's Usable

Users

M2M Async

**Primary Cassandra Cluster**
**28 x PowerEdge C6420**
Node Configuration:
- CPU: 2x 6226R (16C @ 2.9GHz)
- RAM: 192GB
- Storage: 2x 6.4TB NVMe SSD
Cluster Totals:
- 896 Total Cores @ 2.9Ghz
- 5,376GB RAM
- 358TB's Raw, ~269TB's Usable

Data Portal
M2M Sync. / Async.
Raw Data Download
Users

vmware vCenter Server
VxRAIL.
**4 x VxRail P570F**
Node Configuration:
- 2x 6242R (20C @ 3.0Ghz)
- RAM: 768GB
- Storage 2x 800GB Cache, 10x 7.68TB Capacity NVMe SSD
Cluster totals:
- 160 Total Cores @ 3.0Ghz
- 3,072GB RAM
- 307TB's Raw, ~230TB's Usable

**Production**
EDEX / Stream Eng. / PostgreSQL
System Monitoring / Web server/
Apache Tomcat / THREDDS

FACTION DELL
DELL EMC ISILON
**East Coast (NAS)**
1.325PB Total Capacity
Read-Only DR Copy

# Modernized, Simpler
# More Capacity, Faster

- Five Bare-metal Servers
  - Primary Isilon NAS (2.6 PB)
  - Prod VxRAIL VM Server (5 nodes)
  - Test/Dev VxRAIL VM Server (6 nodes)
  - Cassandra Cluster (28 nodes)
  - Keryx Data Portal (20 nodes)
- Secondary Isilon Server on Faction
  - East Coast Isilon NAS (1.3 PB)
  - Active:Active replica
  - Can serve as AWS read-only node
- Tape Library (19 PB)
  - Incremental and Snapshot Backup
  - Offsite storage

# New OSU Architecture

Challenges addressed will be:

- **Performance improvements**
  - Cassandra cluster on SSDs and larger to help keep data volume per node in optimal range
  - New hardware will add compute power
  - Increased network bandwidth

- **Understanding of the system in how it was constructed**

- **Offsite back-ups will provide better disaster recovery**

- **Similar dev, test, uat and prod instances**

- **Improved monitoring for better insight into system and server performances**

- **Ability to add more Ingestion or Stream Engine compute power as needed**

- **Ability to add new management tools (e.g. Toad for Postgres)**

# Transition from Rutgers to OSU Data Center

- First two months:
  - Fast procurement by Dell and OSU.
  - Expedited hiring by OSU.
  - Install by Dell EMC to be finished by the beginning of December.
  - Zoom meetings with Rutgers+WHOI are critical to collaboratively prove the transition timeline and procedures.
- Procedures carried out will include:
  - Hardware testing by Dell during burn-in period, testing that will continue in normal daily monitoring routines.
  - Disaster Recovery testing of various backup and recovery scenarios.
  - Testing with WHOI and Rutgers.

# Overall Timeline

| | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **YEAR 1** | | | | | | | | | | | | |
| Hiring IT Staff | | | | | | | | | | | | |
| Procurement | | | | | | | | | | | | |
| ECC Installation | | | | | | | | | | | | |
| Cassandra | | | | | | | | | | | | |
| Data Transfer | | | | | | | | | | | | |
| VM Setup | | | | | | | | | | | | |
| Data Ingestion | | | | | | | | | | | | |
| Testing + Transition | | | | | | | | | | | | |
| Final Switchover | | | | | | | | | | | | |
| O&M | | | | | | | | | | | | |
| **YEARS 2 & 3** | | | | | | | | | | | | |
| O&M | | | | | | | | | | | | |

**8 Months**

# Detailed Gantt Timeline
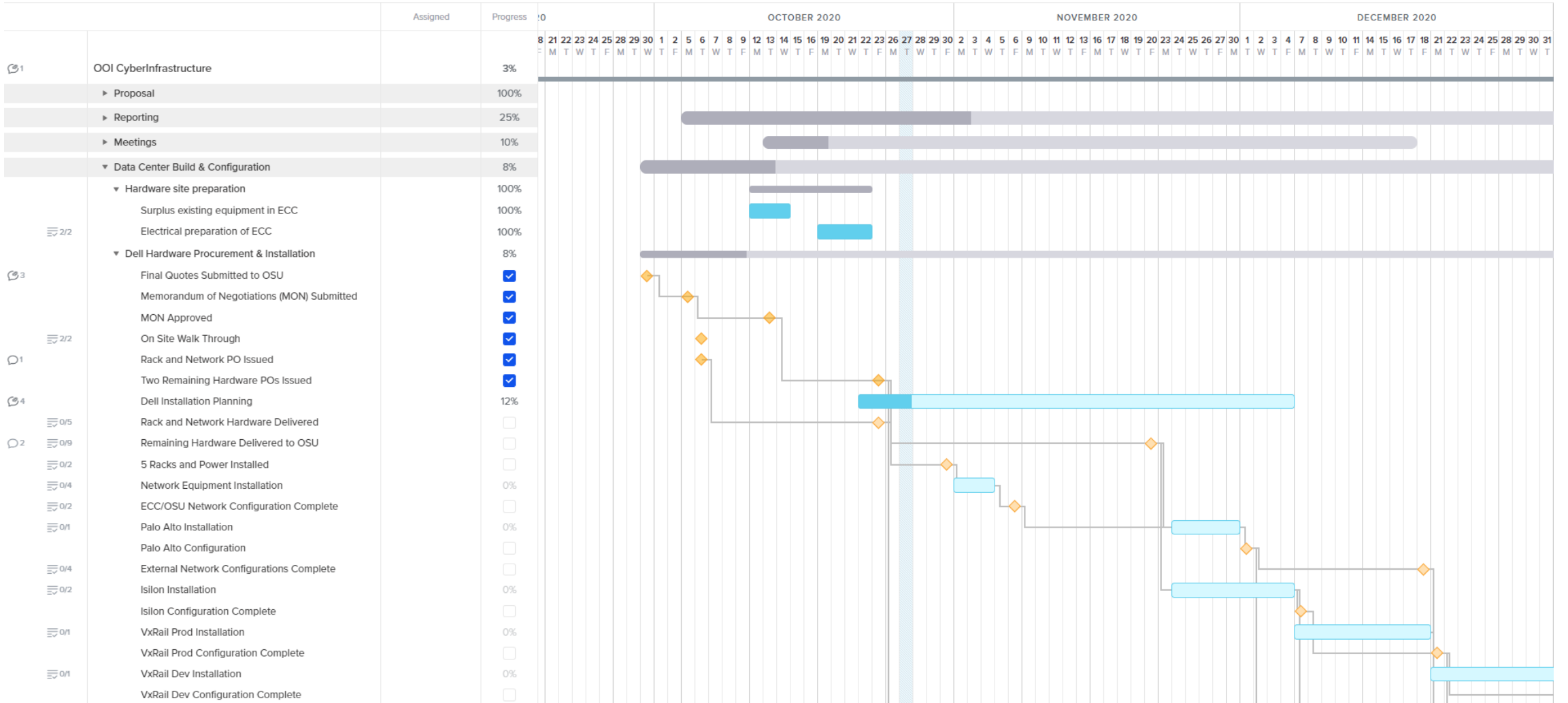
# Key Components in RU➡OSU Transition

1. Gold Raw Data Server
   - 600-700 TB of data files
   - DR copy on Faction on East Coast
2. Prod Software Stack
   - uFrame, Edex, M2M, etc.
   - MIO data ingress
   - Establish identical setup on Test/Dev
3. Cassandra Cluster
   - Upgrade software to 3.1.11
4. Kafka Cluster
   - Kyrex Data Explorer using Docker
   - 50-70 TB of data file storage
5. Tape Library for Offsite Backup

# Gold Raw Data Copying

CHALLENGE: RU has less than 1 Gbit of throughput

- Over internet copying will take more than 2-3 months

- RU will copy 500+ TB onto NAS and ship to OSU

- Remainder will be rsynced, likely from the more proximal and faster West Coast copy of the Gold Raw Data server

- The Faction Secondary Isilon will first be shipped to OSU, then data will be copied from pre-filled primary Isilon, and when the copy is complete, the secondary Isilon will be shipped to Faction in North Virginia

# Key Components in RU➡OSU Transition

1. Gold Raw Data Server
   - 600-700 TB of data files
   - DR copy on Faction on East Coast
2. Prod Software Stack
   - uFrame, Edex, M2M, etc.
   - MIO data ingress
   - Establish identical setup on Test/Dev
3. Cassandra Cluster
   - Upgrade software to 3.1.11
4. Kafka Cluster
   - Kyrex Data Explorer using Docker
   - 50-70 TB of data file storage
5. Tape Library for Offsite Backup

# Prod Software Stack Install

**CHALLENGE:** Complex Software Stack by Various Groups

- Detailed coordination is already starting up

- Providing root/sudo access over VPN using SSH

- OSU will install and maintain OS and major software installs, such as Ansible, Docker and Debian Linux

- Work towards common Grafana monitoring system across all five server components

# Key Components in RU➜OSU Transition

1. Gold Raw Data Server
   - 600-700 TB of data files
   - DR copy on Faction on East Coast
2. Prod Software Stack
   - uFrame, Edex, M2M, etc.
   - MIO data ingress
   - Establish identical setup on Test/Dev
3. Cassandra Cluster
   - Upgrade software to 3.1.11
4. Kafka Cluster
   - Kyrex Data Explorer using Docker
   - 50-70 TB of data file storage
5. Tape Library for Offsite Backup

# MIO Ingress Switch

**CHALLENGE:** In Year 1 ingress will take place to both RU and OSU from MIOs but then needs to switchover

- **Detailed coordination is already starting up**
- **Testing ingress from MIO➜OSU in parallel while RU remains primary ingress point and copy to OSU via rsync**
- **Switching over from RU➜OSU for ingress and copy to RU via rsync for the remainder of Year 1**

# Key Components in RU➡OSU Transition

1. Gold Raw Data Server
   - 600-700 TB of data files
   - DR copy on Faction on East Coast
2. Prod Software Stack
   - uFrame, Edex, M2M, etc.
   - MIO data ingress
   - Establish identical setup on Test/Dev
3. Cassandra Cluster
   - Upgrade software to 3.1.11
4. Kafka Cluster
   - Kyrex Data Explorer using Docker
   - 50-70 TB of data file storage
5. Tape Library for Offsite Backup

# MIO Ingress Switch

**CHALLENGE:** Timing of Cassandra Upgrade to 3.1.11

- Detailed coordination is already starting up

- Cassandra can be established via backup copy or start out as a fresh install following complete replay of all data

- This choice determines when upgrade to 3.1.11 is most opportune

- Replay is significantly faster in new CI data center because of improved hardware and the ability to run multiple Stream Engine server clones as VMs and queuing up of multiple scripts in parallel

- Replay will allow for clean up of thombstones and can deal with refilling of gaps

OOI DDCI Meeting
27 October 2020

Oregon State University

OOI-CI DATA CENTER

# Toward an Extensible Cyberinfrastructure for the Ocean Observatories Initiative

THANK YOU !!!

WE ARE HAPPY TO ANSWER QUESTIONS !!!