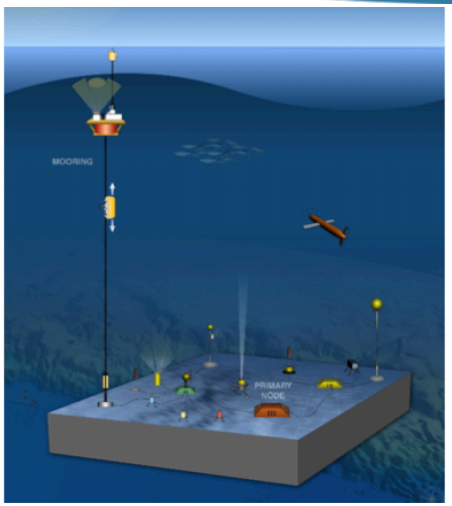
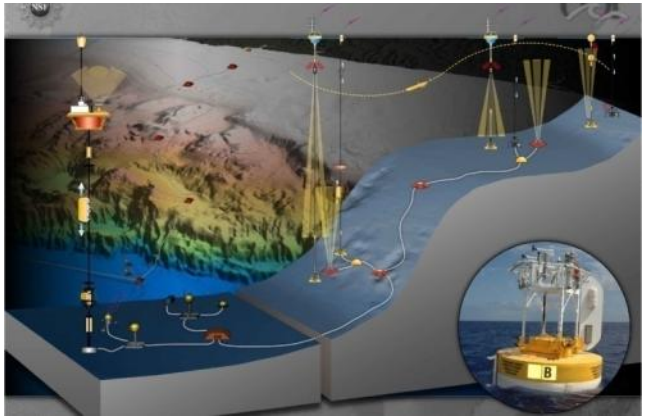
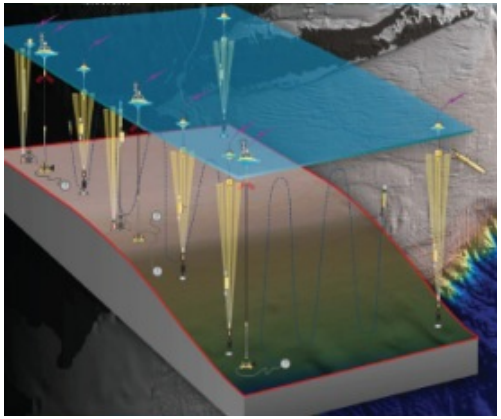
Ocean Observatories Initiative

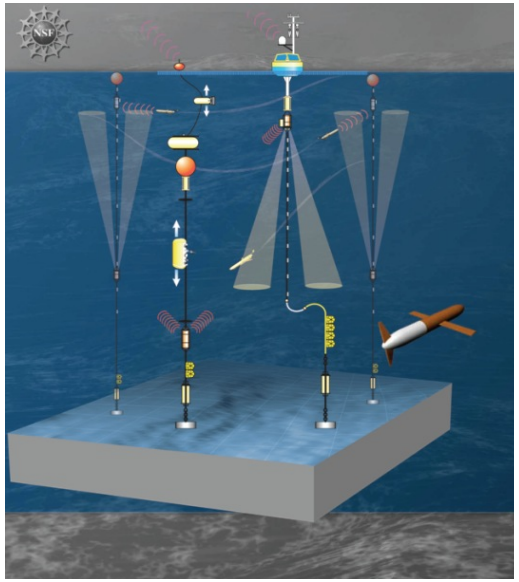


# OOI Data Management

January 15, 2015



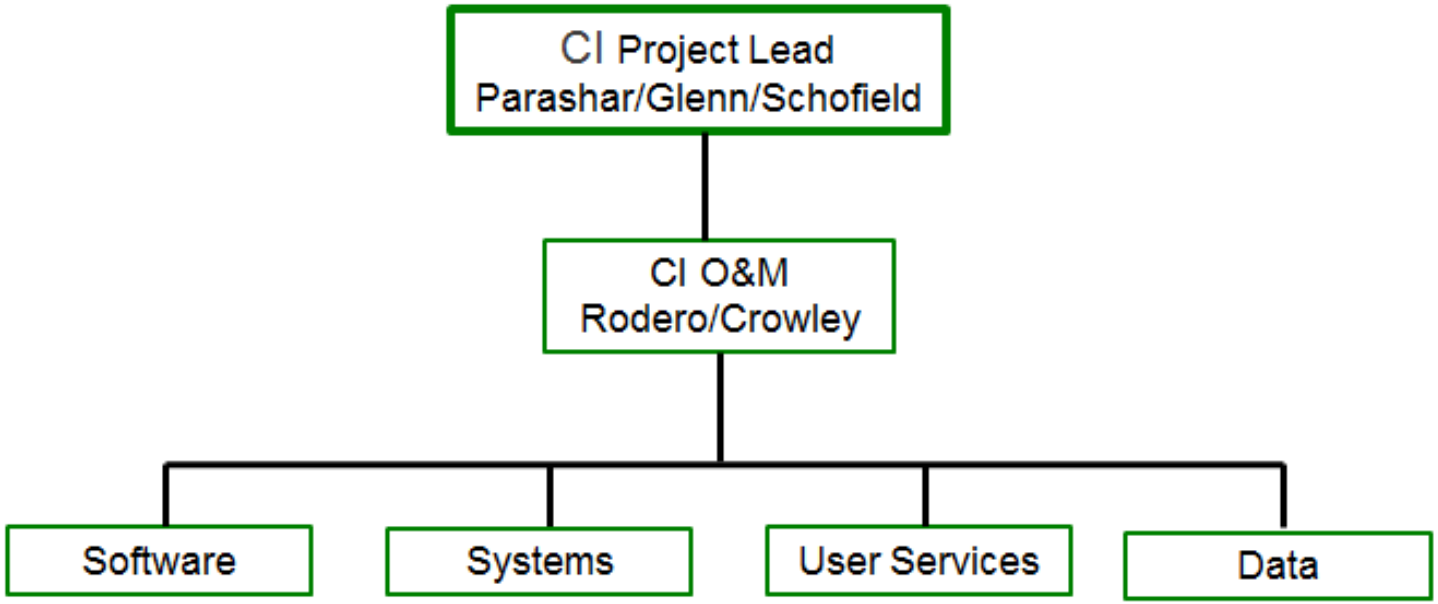




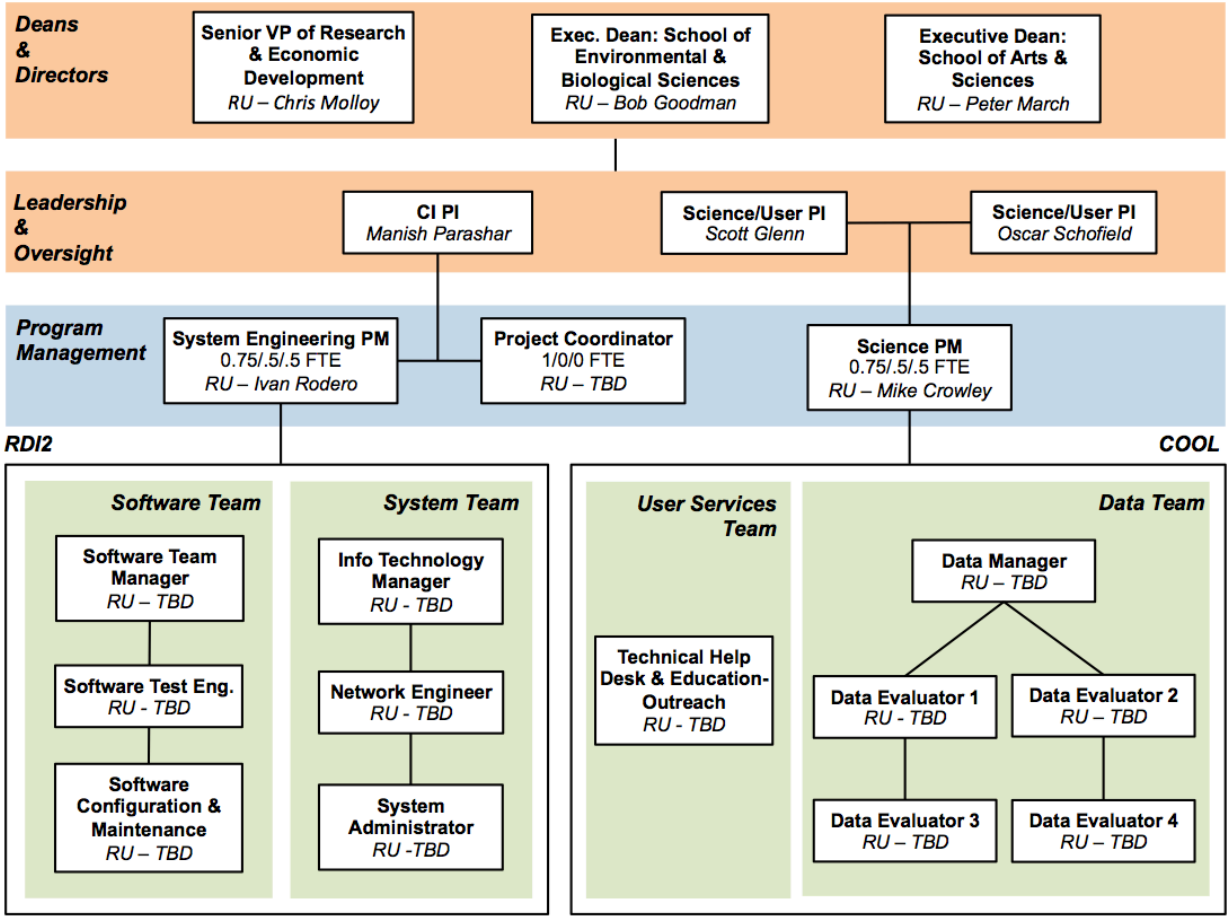
# Data User Groups

- OOI CI CONOPS has the following categories of users:
  - Science
  - Education and Outreach
  - OOI Operations Users
  - Data Manager/Evaluators
  - Configuration & Observatory Asset Data Management
  - Cyberinfrastructure Management
  - OOI Management & Observatory Director

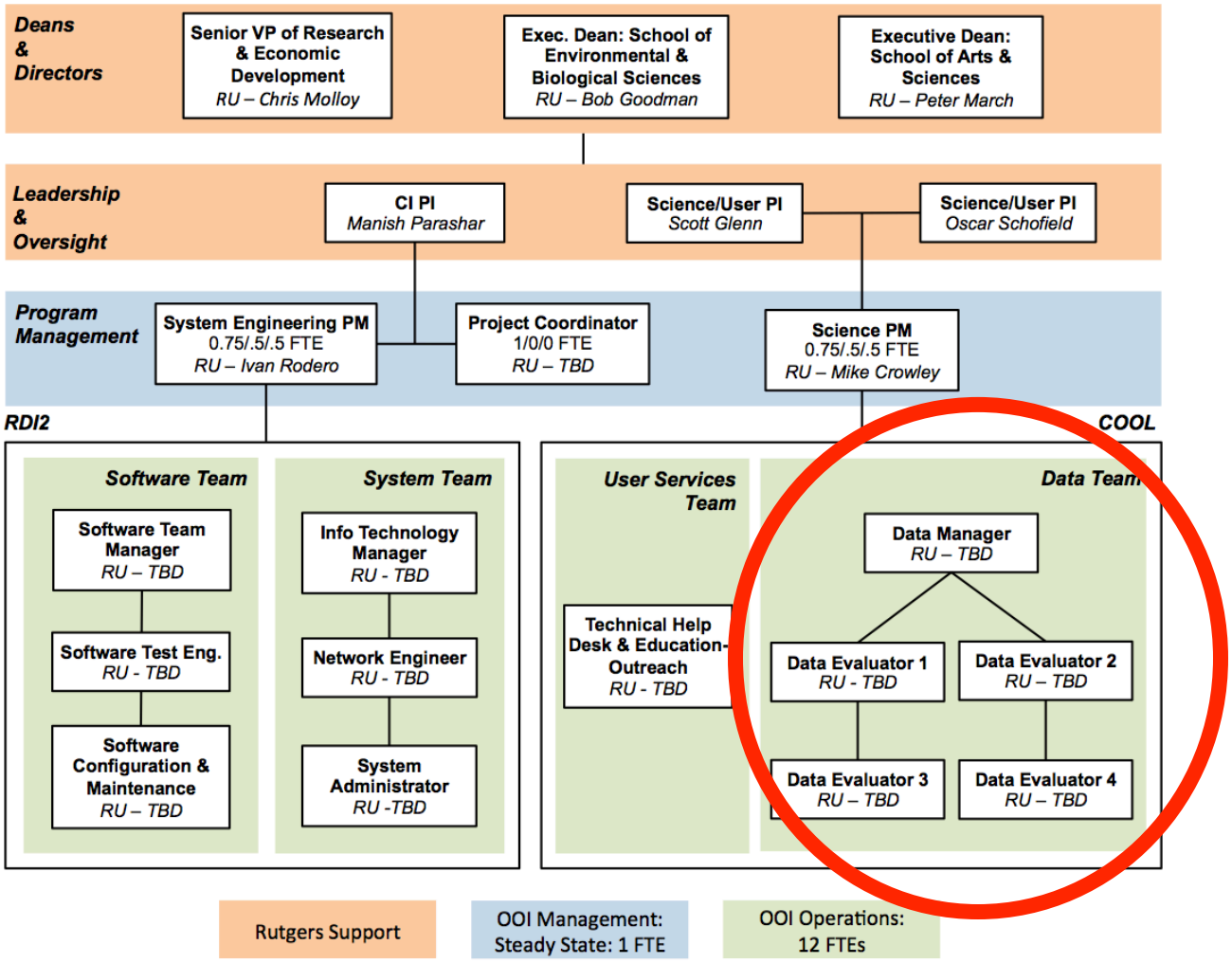
# CI O&M Organization



# Rutgers CI Structure in O&M



# Rutgers CI Data Team



# Data Management Team

- Data Manager (1)
  - Ph.D. + Observatory Experience
  - Manage the data lifecycle
  - Manage the data evaluators
  - Primary human interface for data access & feedback

- Data Evaluators (4)
  - Senior level - Ph.D. or MS + Experience
  - Regular level – MS or BS + Experience
  - Monitor automated Quality Control
  - Regular and directed human-in-the-loop QC
  - Designated for Global, Cabled, Endurance & Pioneer

# Data Manager - Hired

OSU Oregon State University

Find people and pages | Search CEOAS

College of Earth, Ocean, and Atmospheric Sciences

Home | Directory | Research | Academics | Ships | Facilities | About

Future Students | Faculty & Staff | Alumni | Friends & Donors | Employment | Outreach | News

## Directory of People

Search

Everyone (with photos)

Advisors (undergraduate)

Business Office

Job Category

Deans

Teaching & Research Faculty

----- Faculty Specialties

Research Assoc., Postdoc, Faculty Research Assist. (RAFRA)

Graduate Students

----- Geography

----- Geology

----- MRM

----- OEAS

## Michael Vardaro

**Assistant Professor (Senior Research) - OOI Assoc. Project Scientist**

*Discipline:* Ocean Ecology and Biogeochemistry
*Office:* Burt 295
*Phone:* 541-737-9350
*Fax:* 541-737-2064
*Email:* mvardaro@coas.oregonstate.edu
Vita or Résumé (PDF)

### Research Interests

Deep-sea biogeochemistry, with specializations in remote sensing, invertebrate biology, time-lapse underwater photography, climate variability and change, and surface-to-seafloor connectivity.

### Current Research

I am an associate project scientist on the Ocean Observatories Initiative, an NSF-funded project to record long-term environmental changes and short-term events in the marine ecosystem on a Coastal, Regional, and Global scale. My portion of the project is to help design, test, and deploy the Endurance Array off the coast of Oregon and Washington. Endurance will consist of a series of moorings and seafloor observatories, some self-contained and others connected to electro-optical cables to provide real-time data and power, and will be instrumented with various types of sensors, cameras, and autonomous vehicles.

# Deployed Scope of OOI (over 800 instruments distributed over all moorings, benthic packages, seafloor nodes, gliders and AUVs)

## Global Arrays

| Subsystems | Components | Instruments | Service Frequency |
|---|---|---|---|
| ***Global Arrays*** | | | |
| Station Papa | 1 Subsurface Hybrid Profiler Mooring | 12 | Yearly |
| | 2 Flanking Moorings | 32 | |
| | 3 Gliders | 9 | |
| Irminger Sea | 1 Surface Mooring | 23 | Yearly |
| | 1 Subsurface Hybrid Profiler Mooring | 12 | |
| | 2 Flanking Moorings | 32 | |
| | 3 Gliders | 9 | |
| Southern Ocean | 1 Surface Mooring | 23 | Yearly |
| | 1 Subsurface Hybrid Profiler | 12 | |
| | 2 Flanking Moorings | 32 | |
| | 3 Gliders | 9 | |
| Argentine Basin | 1 Surface Mooring | 23 | Yearly |
| | 1 Subsurface Hybrid Profiler | 12 | |
| | 2 Flanking Moorings | 32 | |
| | 3 Gliders | 9 | |

## Coastal Arrays

| Subsystems | Components | Instruments | Service Frequency |
|---|---|---|---|
| ***Coastal Arrays*** | | | |
| Pioneer | 3 Surface Moorings | 60 | Twice a year |
| | 2 Surface-Piercing Profilers Moorings | 18 | |
| | 5 Profiler Moorings | 29 | |
| | 3 AUVs | 18 | |
| | 6 Gliders | 30 | |
| Endurance (Oregon Line) | 3 Surface Moorings | 50 | Twice a year |
| | 2 Surface-Piercing Profilers Moorings | 18 | |
| | 1 Hybrid Profiler Mooring | 16 | |
| | 1 Benthic Experiment Package | 10 | |
| | 1 Multi-Function Nodes | 8 | |
| Endurance (Washington Line) | 3 Surface Moorings | 68 | Twice a year |
| | 2 Surface-Piercing Profilers Moorings | 18 | |
| | 1 Profiler Mooring | 5 | |
| | 6 Gliders | 30 | |

## Cabled Arrays

| Subsystems | Components | Instruments | Service Frequency |
|---|---|---|---|
| ***Regional Scale Nodes*** | | | |
| Hydrate Ridge | Seafloor: Primary and Secondary | 16 | Yearly |
| | Profiler – Winched | 10 | |
| | Profiler – Wire crawler | 5 | |
| | Midwater Platform@ 200m | 8 | |
| | Bottom Instrument Package | 6 | |
| Axial Seamount | Seafloor: Primary and Secondary | 26 | Yearly |
| | Profiler – Winched | 10 | |
| | Profiler – Wire crawler | 5 | |
| | Midwater Platform @ 200m | 8 | |
| | Bottom Instrument Package | 6 | |

**Connected by 880km of seafloor cable, with 10KW power, internet connectivity between 7 primary nodes, multiple secondary nodes, and all distributed instrumentation**

## Cyberinfrastructure

**Computing platforms, software applications, storage, and high speed network equipment**

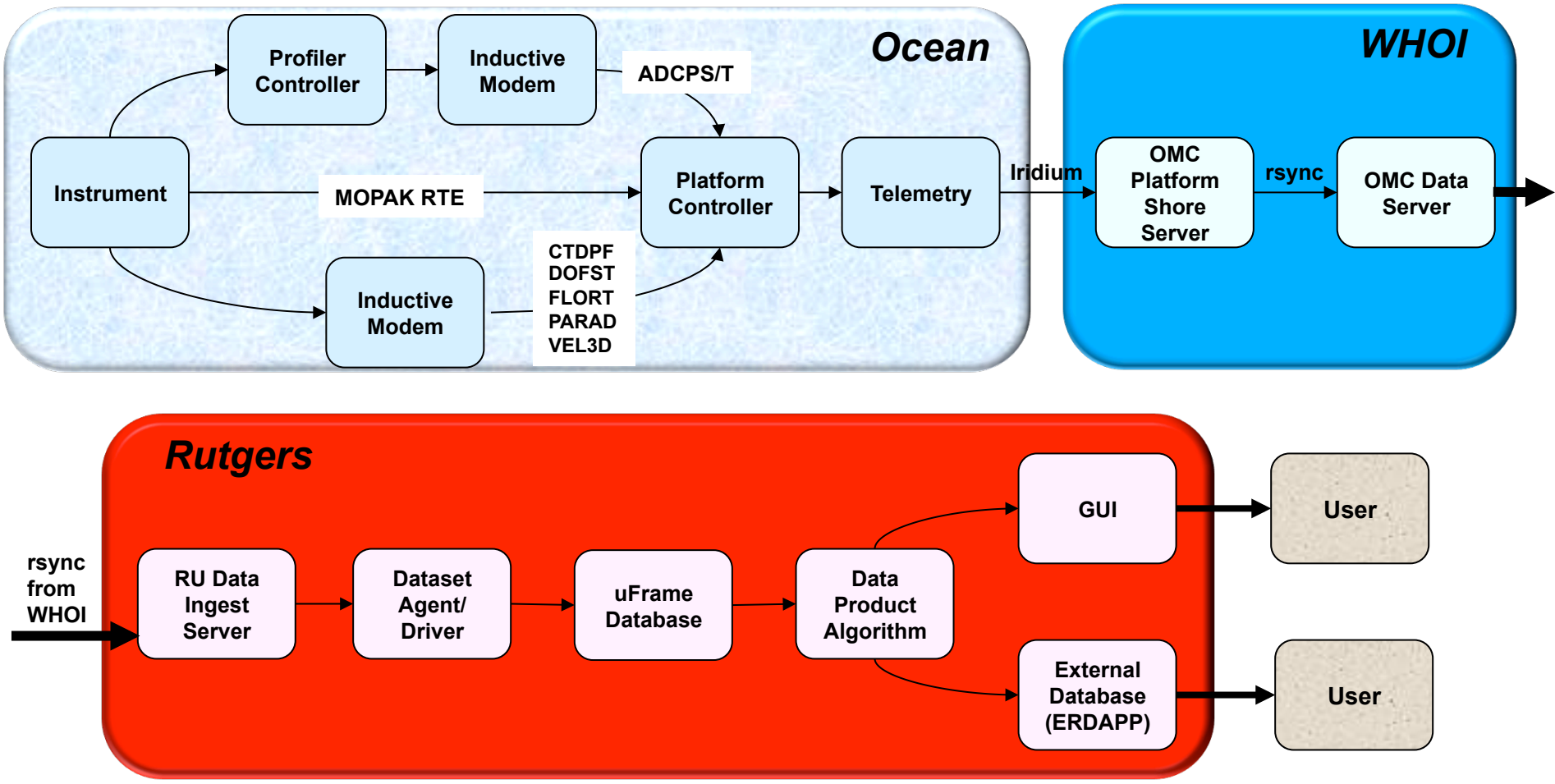
**Cyber Points of Presence (CyberPoPs)**
- **Acquisition Points**
- **Distribution Points**

**Integrated Observatory Network – OOI Net**
- **Hardware / Software**

**Redundant computing environment**

**Extensive details about each component can be found on the OOI website (http://oceanobservatories.org)**

# Data Flow Example: Pioneer Profiler



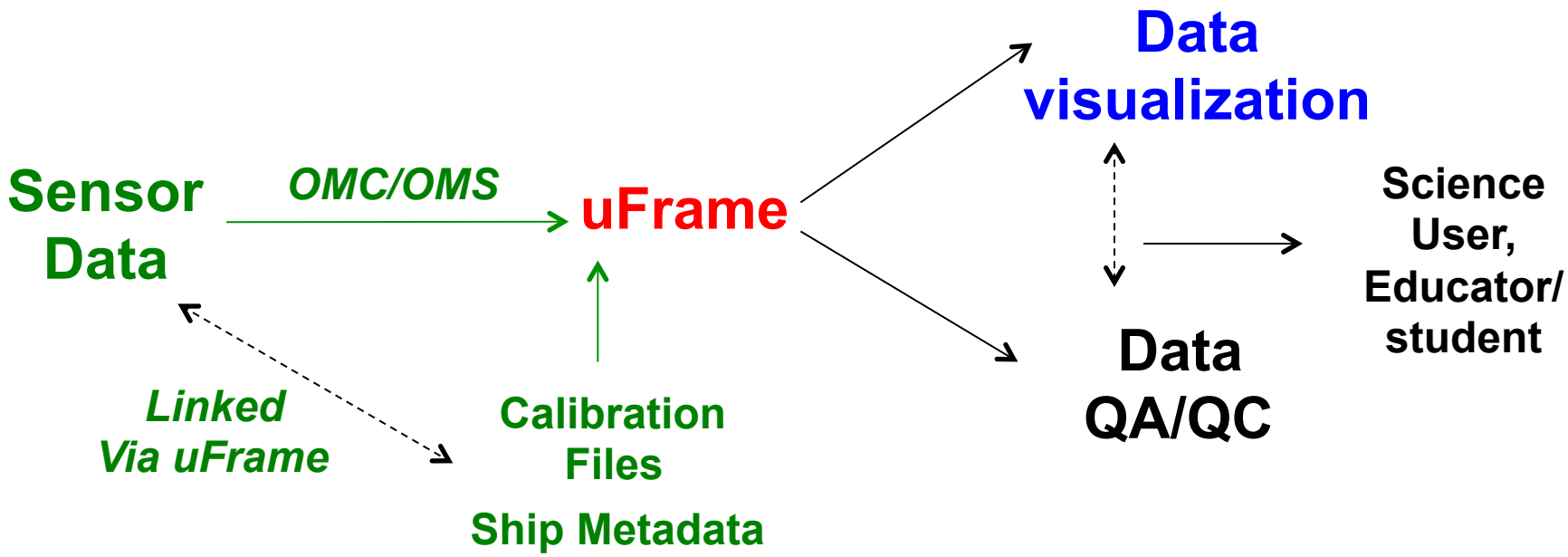
# Types of Data Acquired

- Telemetered *(Global, Pioneer, Endurance Arrays)*
  - Near real time dataset transmitted to OMC via satellite.
  - A smaller subset of the full recovered dataset.
- Streaming *(Cabled, Endurance Arrays)*
  - Real time science and engineering data sent to OMS via cable.
  - Nearly complete subset of the full recovered dataset.
- Recovered
  - Full dataset (science & engineering) brought back to shore via the recovery vessel.
- Calibration
  - Dedicated data collected on shore or at sea by manufacturers or Marine Operators to calibrate at-sea sensors.
- Shipboard
  - Environmental data collected by the ships in normal operations.
- Metadata

# Data Products Produced

- Level 0 (L0)
  - Raw sensor data acquired by the Marine Operators
- Level 1 (L1)
  - Calibrated data in science units:
    (a) Derived from L0 data (e.g. Conductivity, Temperature & Depth from a CTD), or
    (b) Delivered through vendor software (e.g. Glider data)
- Level 2 (L2)
  - Calibrated data in science units derived from multiple L1 products (e.g. Salinity & Density from a CTD)

# OOI Data Quality ConOps




| Marine IOs | OOI-net | Ux/Ui | Data Qa/Qc |
|---|---|---|---|
| *WHOI, Scripps OSU, UW* | *Raytheon, ASA, Rutgers* | *ASA, Rutgers* | *Rutgers, SMEs* |

# Data Quality Assurance & Control

- Quality Assurance Activities by Marine Operators
  - Factory Calibrations, Burn in, Deployment, Recovery
  - Feedback from Data Management Team

- Automated Quality Control Algorithms
  - Global Range, Local Range, Stuck Sensor, Spike, Gradient, Trend
  - Generate & accumulate automated Quality Control flags
- Quick Look Data Plots
  - Human in the loop flagging (red, yellow, green & events)
  - Routine workflow with standardized GUIs for Data Evaluators
  - Time Series, Profiles, T-S, Multiple Sensors on same platform
- Deep Dive Data Investigations
  - Human in the loop
  - Directed workflow by Data Manager with input from community
  - Recovered vs Telemetered, Comparisons w/ Calibration Data, Vicarious calibration comparisons, Climatologies/Models, etc.

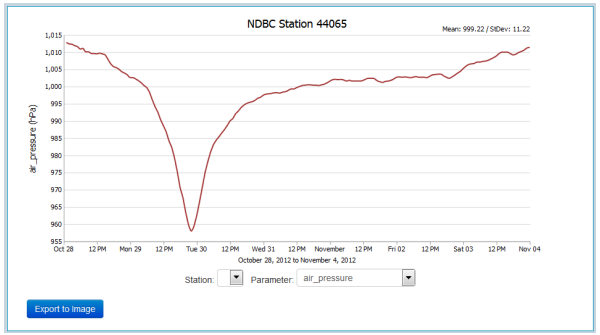
# User Interface Categories

- Status User Interface
  - The Status UI will rely-upon engineering, instrument, status log and asset information data that will be made available by the uFrame system.
- Science User Interface
  - Will contain tools to perform basic analysis of the data. Some of the tools are: GIS mapping of marine assets, visualization of time-series data, trend-analysis tools.
- Asset Management User Interface
  - Front-end for the asset management database. Allows users to view and modify the data within the asset management system. Also provides APIs to allow scripted updates.
- Command and Control User Interface
  - Front-end for the monitoring of marine assets and infrastructure. Provides capability of sending commands to the OOI cabled infrastructure specific to the Cabled Array needs

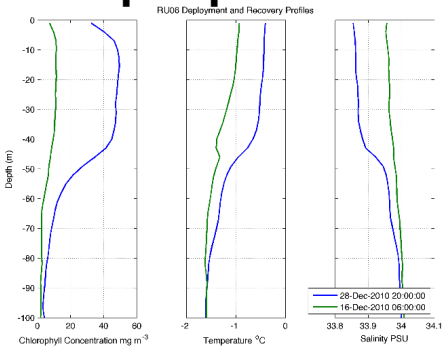
# Visualization Tools for Data QC

- Close iterative feedback between ASA visualization team and Rutgers data team

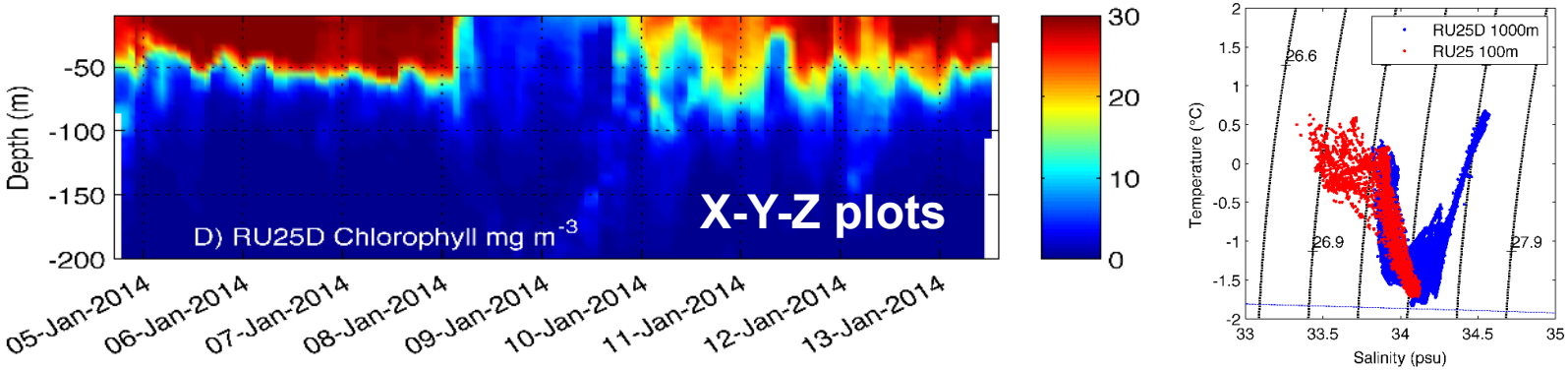
**Time series**




**Depth profiles**




**Parameter-parameter plots**

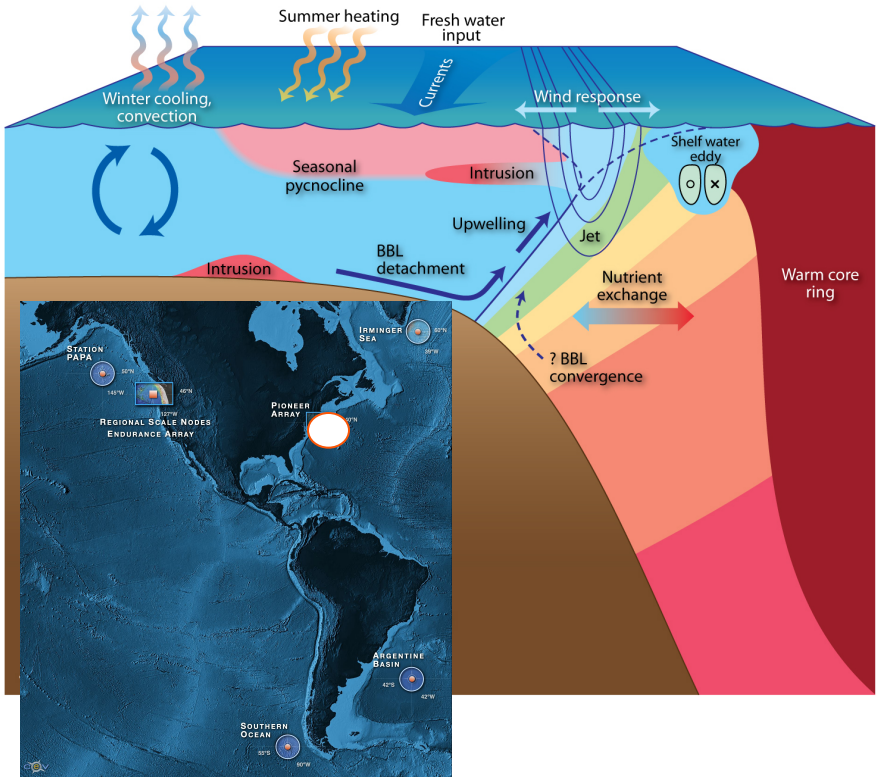


# QC Test Comparisons - ARGO

| ARGO QC Test | OOI QC Test |
|---|---|
| 1. Platform ID* | 1. Data is sorted by reference designator so this needs to be correct for the data to go into the file |
| 2. Impossible date* | 2. Time series check in quick look plots. An automated algorithm has been suggested |
| 3. Impossible location* | 3. Quick look maps generated for glider/mooring locations to determine close approach times for mobile and fixed assets are already being used for this |
| 4. Position on land* | 4. Same as 3 |
| 5. Impossible speed* | 5. Same as 3. Could be automated |
| 6. Global range test* | 6. Already part of automated QC algorithms |
| 7. Regional parameter range* | 7. Already part of automated QC algorithms |
| 8. Pressure increasing | 8. Less relevant. Profilers move both directions and can be impacted by turbulence or shallow water waves |
| 9. Spike test | 9. Already part of automated QC algorithms |
| 10. Top – bottom spike - obsolete | 10. Obsolete |
| 11. Gradient test | 11. Already part of automated QC algorithms |
| 12. Digit rollover | 12. Digital rollover? Not sure what this is |
| 13. Stuck value | 13. Already part of automated QC algorithms |
| 14. Density inversion | 14. Plan to implement this as part of automated QC as level 2 products are produced |
| 15. Grey list | 15. Not applicable |
| 16. Gross salinity or temperature drift | 16. Part of the multi-time scale quick look plot examination |
| 17. Visual QC – not mandatory in real time | 17. Weekly visual QC is mandatory |
| 18. Frozen profile | 18. Part of out of range test or stuck value test |
| 19. Pressure not greater than Deepest Pressure = 10% | 19. Pressure outliers identified in quick looks |

# Pioneer Array Science

**How exchanges between a broad shelf with the a deep ocean that is bounded by an energetic western boundary system structure physics, chemistry, and biology of continental shelves**



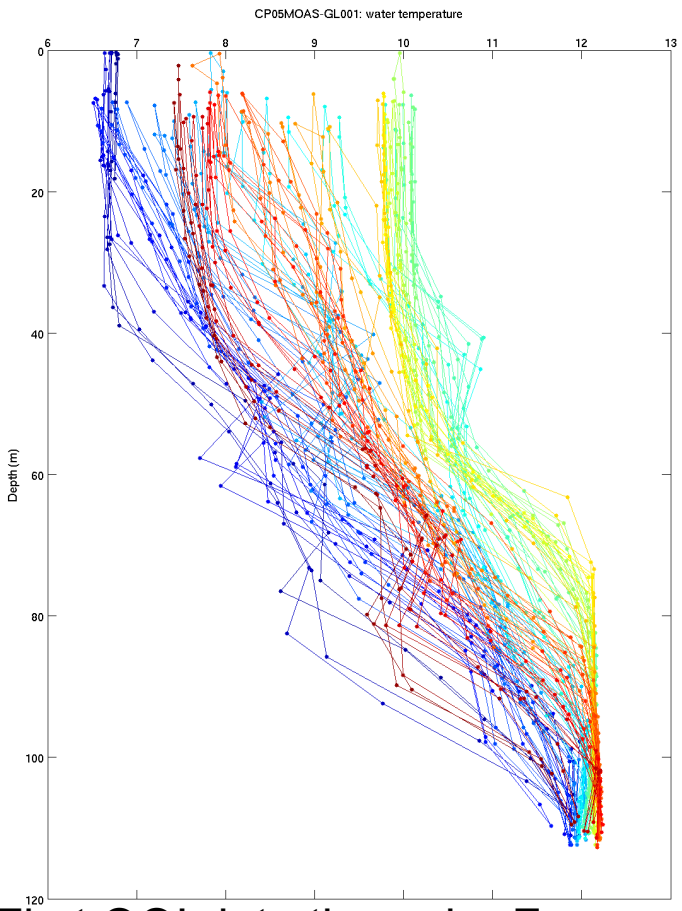

- **Observing Requirements:**

  **Nested simultaneous observations resolving short time scales and multiple spatial scales, data from air-sea interface to sea floor, multidisciplinary sensor suites, real-time data, high resolution adaptive sampling**

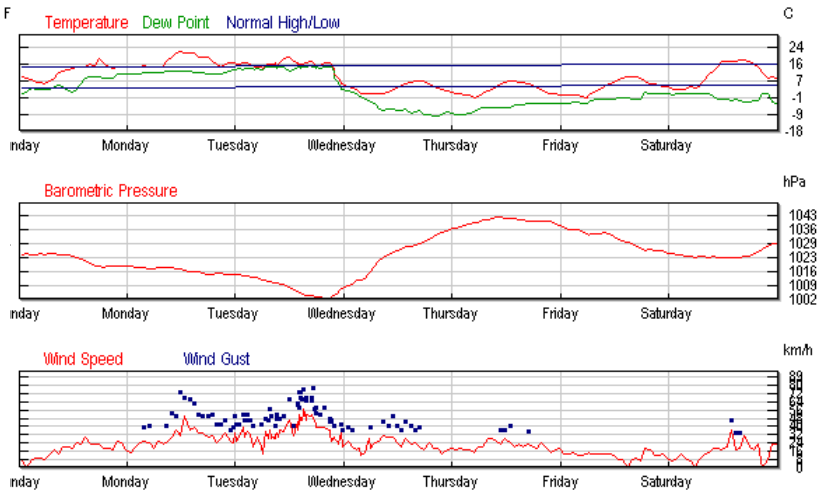
- **Engineering Drivers:**

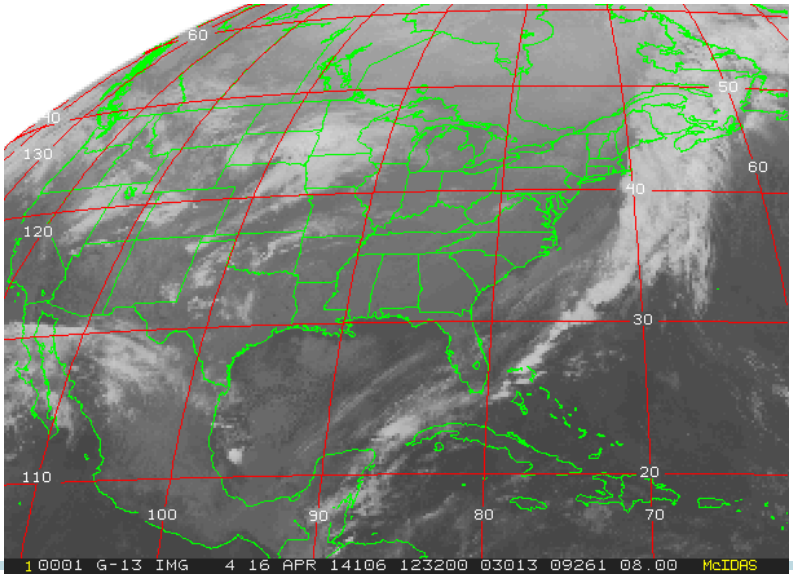
  **High turbulence resulting in high frequency heterogeneity in space/ time, high rates of bio-fouling, human presence, rapid response cabailities**

# Comparisons to Other Data Streams



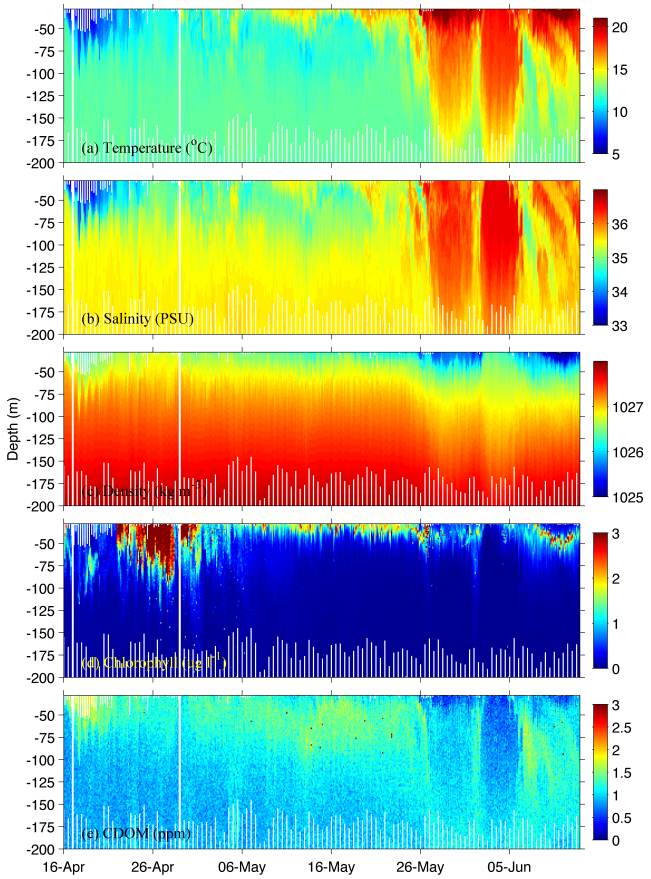

First OOI data through uFrame - Glider temperature profiles



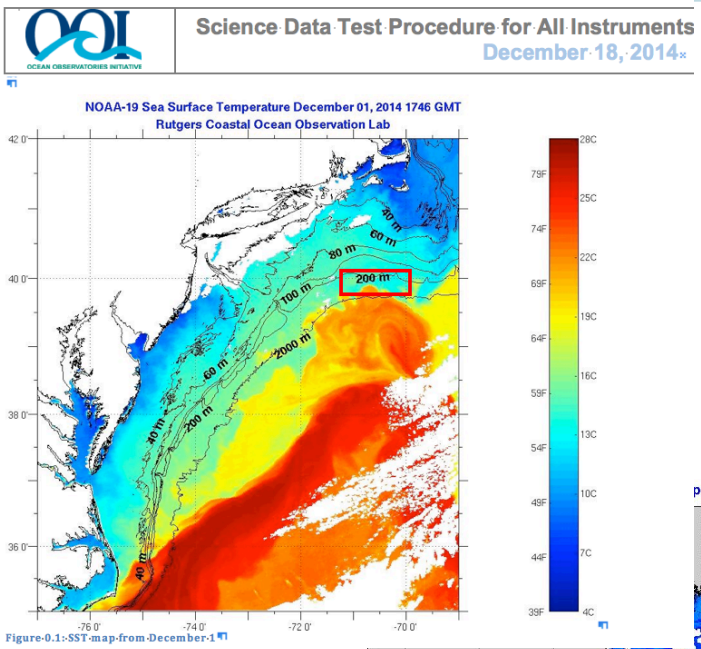



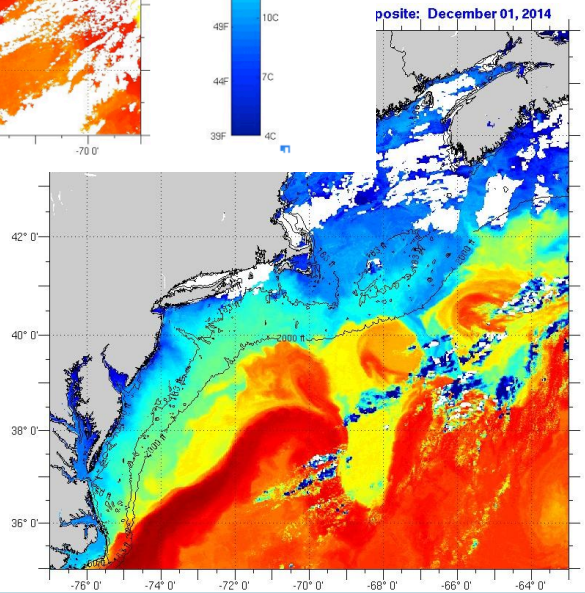

Met data showing storm responsible for surface cooling

# Comparisons to Other Data Streams



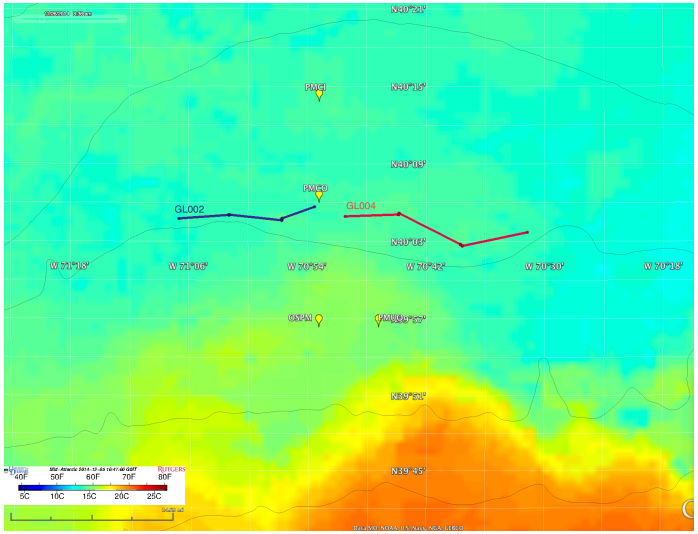

Warm/salty anomalies observed in glider data



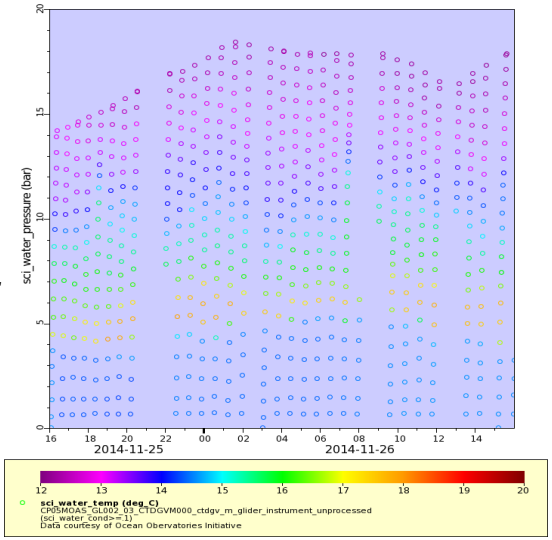




Satellite SST showing warm core rings responsible for warming

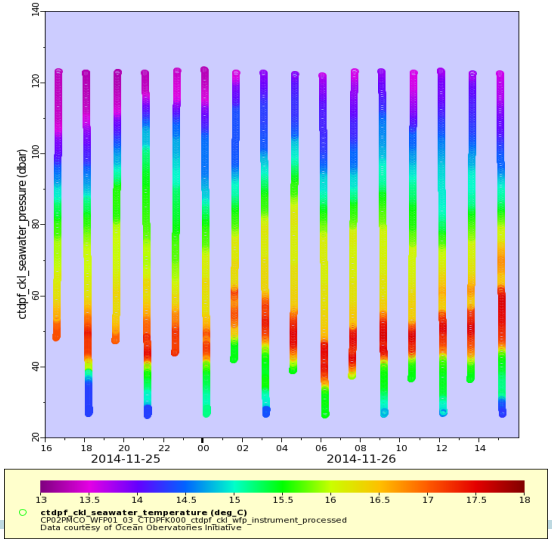
# Vicarious Calibration Opportunities



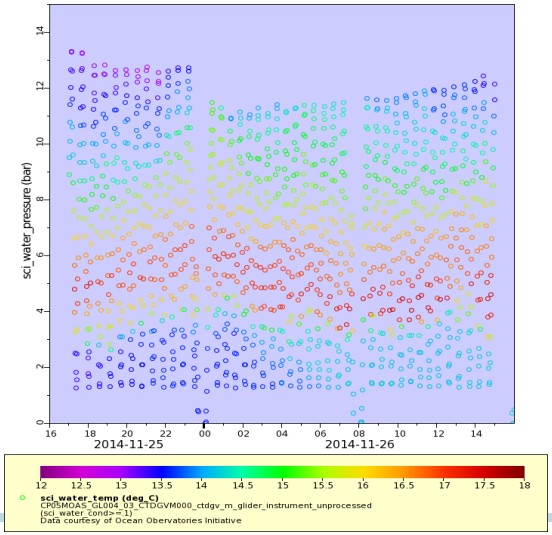

Two OOI Gliders fly by the OOI Wire Following Profiler inshore of the warm core ring observed in the SST




Glider 002



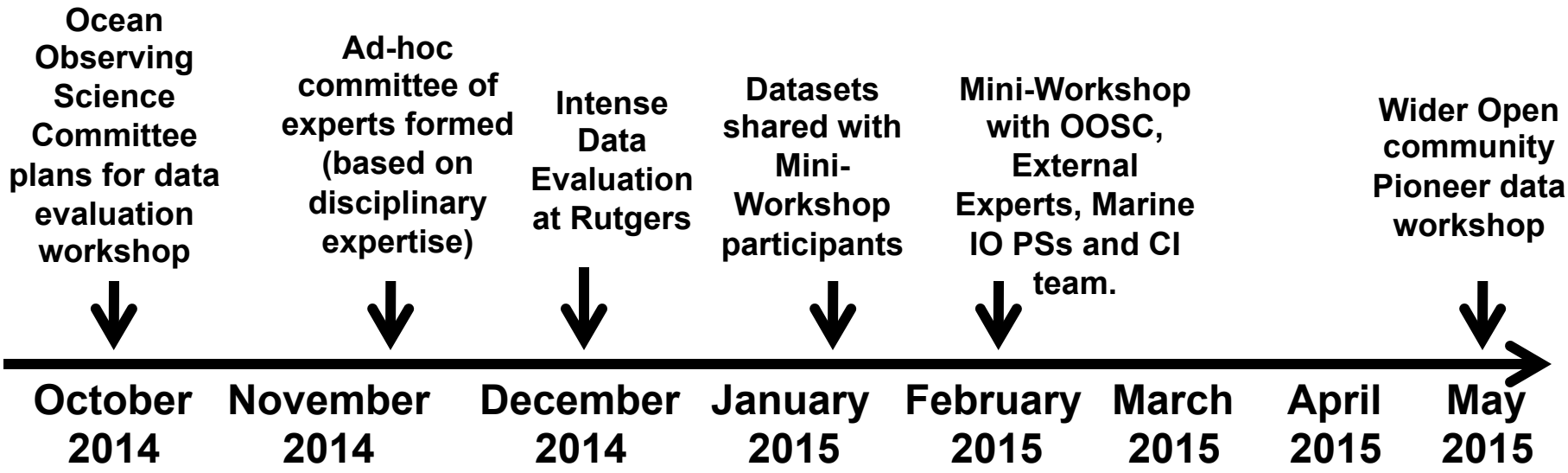

Profiling Mooring Central Offshore




Glider 004

# Data Quality Confidence Building

- Efforts to engage the external community
- Engender excitement and ownership of OOI data.
- Initial focus on Pioneer Array Mini-workshop
- Expandable to core OOI sensors at other arrays



# Pioneer Array Mini-Workshop

- 1.5 Day Mini-Workshop

- At Rutgers on Feb 12-13

- Overall Goals:
  (1) Sensors – Review QC procedures for first 4 sensor systems
  (2) User Information – Define mechanisms for outreach to user community
  (3) Community Pioneer Data Workshop – Decisions on timing, goals, agenda & constraints

# Mini-Workshop Feedback Matrix

- Provide feedback on the matrix of 3 data QC activities being applied to 4 types of sensors deployed on Gliders and Wire Following Profilers at Pioneer

| | CTD | ADCP | Ecopuck & PAR | Oxygen |
|---|---|---|---|---|
| Automated Algorithms | | | | |
| Quick Look Plots | | | | |
| Deep Dive Investigations | | | | |

- Based on the Pioneer data viewed during the feedback matrix exercise, provide a summary assessment of the data at this stage and recommendations for the future

# Mini-Workshop Participants

- **OOSC Members:**
  Larry Atkinson, Emmanuel Boss, Mary Jo Richardson, Steve DiMarco, Rouying He, Suzanne Carbotte, Raphael Kudela & Annette DeSilva
- **External Experts:**
  (a) Wendell Brown (U Mass) - CTDs
  (b) Libe Washburn (UCSB) – ADCPs
  (c) Heidi Sosik (WHOI) – Puck & PAR (Optical)
  (d) Mark Moline (U Del) - Oxygen
- **Marine Operator Project Scientists (PS):**
  (a) Al Plueddemann (Pioneer), in person
  (b) Jack Barth (Endurance), Deb Kelley (Cabled) & Bob Weller (Global), via webex
- **Rutgers CI Data Management Team, OL & NSF**

# Future Mini/Community Workshops

- Targeting 2-3 Community Workshops in 2015.
- February Pioneer Mini-Workshop leads to the Pioneer Community Workshop
- After Pioneer Mini-Workshop, begin work on potential Mini-Workshops with Deb Kelley on Cabled Array, Bob Weller on Global, Jack Barth on Endurance.
- Endurance is a natural extension of the Pioneer & Cabled Array workshops, and will benefit from that experience.
- Global Array has already held a successful Irminger Sea workshop to engender international community involvement.

# Pioneer Data Evaluation Philosophy

- Gather a small group of experts to evaluate the scientific quality of a Pioneer data subset that passes through uFrame.
- Generate recommendations in time to help prioritize remaining construction & transition.
- Pioneer data subset through uFrame release 1A of 5 planned releases is the first opportunity.
- Data streams may not be 100% bug-free, just sufficiently bug-free to enable an initial scientific evaluation.

# Pioneer Data Evaluation Process

- Quick reminder of Pioneer platform and sensor configurations.
- Introduce existing transition processing and visualization tools.
- Review existing test data that has passed through uFrame release 1A as documented in existing reports.
- Review automated QC Algorithms and coefficient look-up table.
- Generate recommendations for OOI Net Transition Team

# Pioneer Data Recommendations

Thursday – develop recommendations for:

- uFrame core data processing algorithms
- GUI display screens to enable internal data evaluation and external scientific use
- Automated QC coefficients in look-up table
- Additional data investigations led by the Data Manager
- Feedback to Marine Operators

Friday – gap fill and prioritize (1,2,3). Submit.