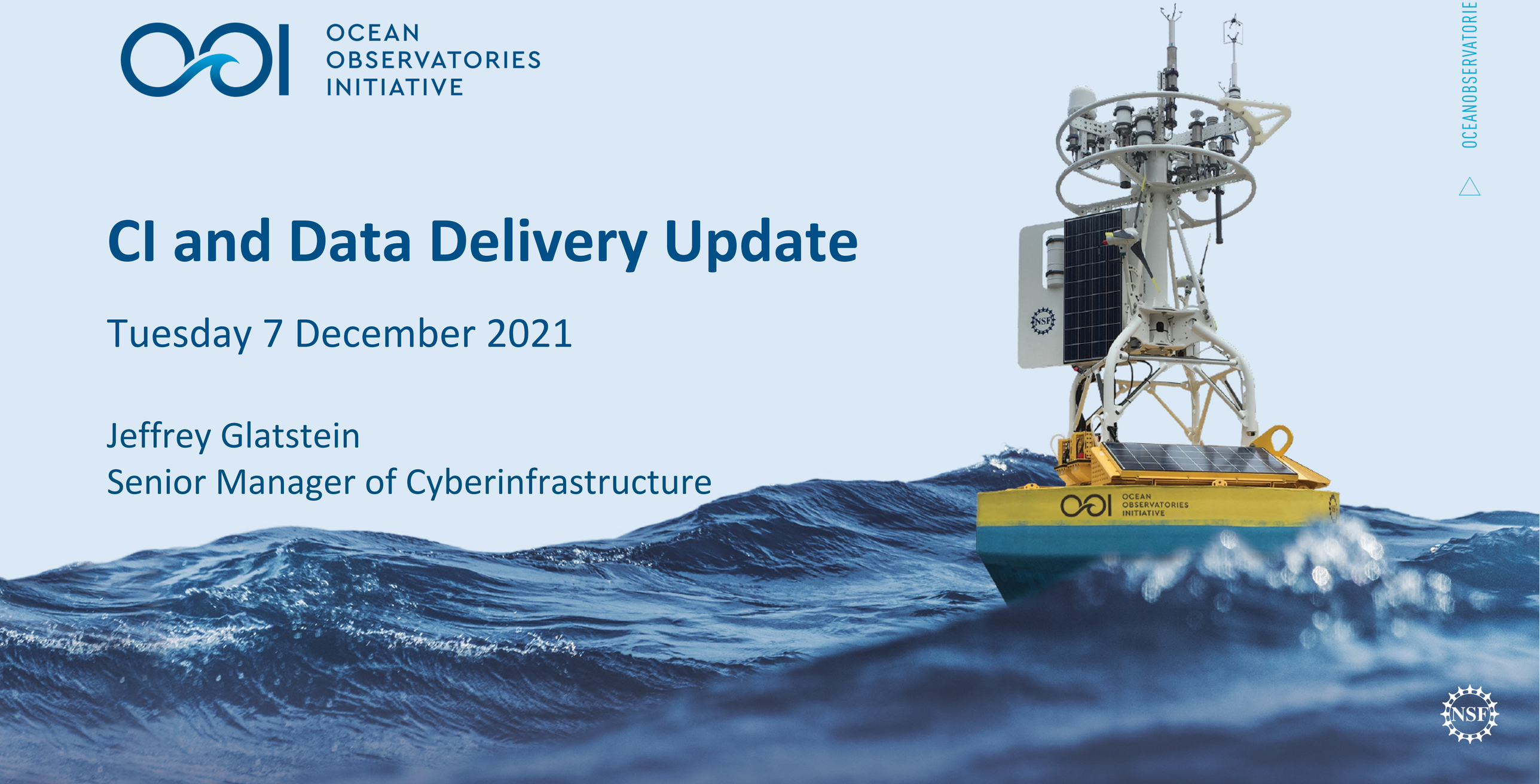
OCEAN OBSERVATORIES INITIATIVE

OCEANOBSERVATORIES.ORG

CI and Data Delivery Update

Tuesday 7 December 2021

Jeffrey Glatstein
Senior Manager of Cyberinfrastructure



NSF

Agenda

- PYIII Work Plan (13-Oct-20)
- PYIII Highlights
- PYIV Work Plan Highlights
- Questions
- PYIII Year In Review (for reference)

PYIII Work Plan Presented at OOIFB meeting October 13, 2020

- Data Center migration to OSU including Dev through Prod set-up
- OOI CI Software stack upgrade
- QA/QC – continue implementing QARTOD logic
- Complete Roundabout development for connectivity to uFrame
- Data Explorer v1.x
  - Continued increase in science data provided
  - Discrete sample data
  - Prototype for visualization of data <1min resolution
  - Login function
  - Support Jupyter notebook access
- Media server for HD video, Hydrophone, HD pictures and Echograms
- Migration of real-time streaming from OOINET to Django
- Instrument centric approach on ticket resolution
- Pre-load and NetCDF compliance
- “Redmine” redesign including public access

PYIII Highlights

- Data Center migration to OSU including Dev through Prod set-up
  - Data migration utility eliminated the need to delete instrument data for all time in order to address changes to a small amount of data.
  - Corrected Cassandra large bin performance and stability issue by applying new bin parameter values.
  - Migration of data (both telemetered and cabled) improved the data quality by eliminating duplicate and prior bad data ingestions.
  - OOI CI Software stack upgrade (300+ drivers and libraries)
    - Cassandra upgrade improved functionality by improving performance and reducing timeouts.
- Data Explorer v1.1 and 1.2
  - New user feature to print timeseries and data comparison charts.
  - View discrete sample and glider data in grid view.
  - Multiple viewing enhancements for gliders including polygon search feature.
  - Ingestion and viewing of discrete sample data.
  - Prototyped visualization of timeseries data <1min resolution.
  - Support Jupyter notebook access
    - POC utilizing the Thredds Gold copy data source processed OOI data interactively in place without the need to first download.

PYIV Work Plan

- Data Explorer v1.3 and v1.4
  - Visualization of multi-dimensional plot data (e.g. spectral)
  - Tuning of data ingest re-load
  - Ingest ZPLS* echopyp generated data streams and produce echograms
  - Media server assets – video and photos
  - Expansion of Jupyter notebook access
- Stream Engine modernization
- Instrument centric approach on ticket resolution
  - Pressure/Depth
  - PHSEN
  - Dissolved Oxygen
- QARTOD support
  - Climatology and gross range implementation
  - Gap and timing requirements gathering.
- Complete the Roundabout connectivity to uFrame
- Pre-load and NetCDF compliance improvements continues

PYIII Year in Review

- Data Center migration to OSU including Dev through Prod set-up [Completed]
  - Built utility to move data from one Cassandra database to another yielded the ability to read and update the Cassandra data by timestamp. This eliminates the need to delete instrument data for all time in order to address a small amount of data.
  - Corrected Cassandra large bin performance and stability issue by applying new bin parameter values.
  - Migration of data (both telemetered and cabled) improved the data quality by eliminating duplicate and prior bad data ingestions.
  - Improved data ingestion throughput.
  - OOI CI Software stack upgrade [Completed]
    - Upgraded 300+ drivers and libraries.
    - Utilizing latest uFrame release.
    - Improved supportability across all packages (some off which had been support was no longer available).
    - Cassandra upgrade improved functionality by improving performance and reducing timeouts.
    - Postgres upgrade provided stability and the ability measure performance.
  - New Dev system architecture mirrors production set-up (apart from size) allowing for more comprehensive unit testing. [Completed]

PYIII Year in Review – continued

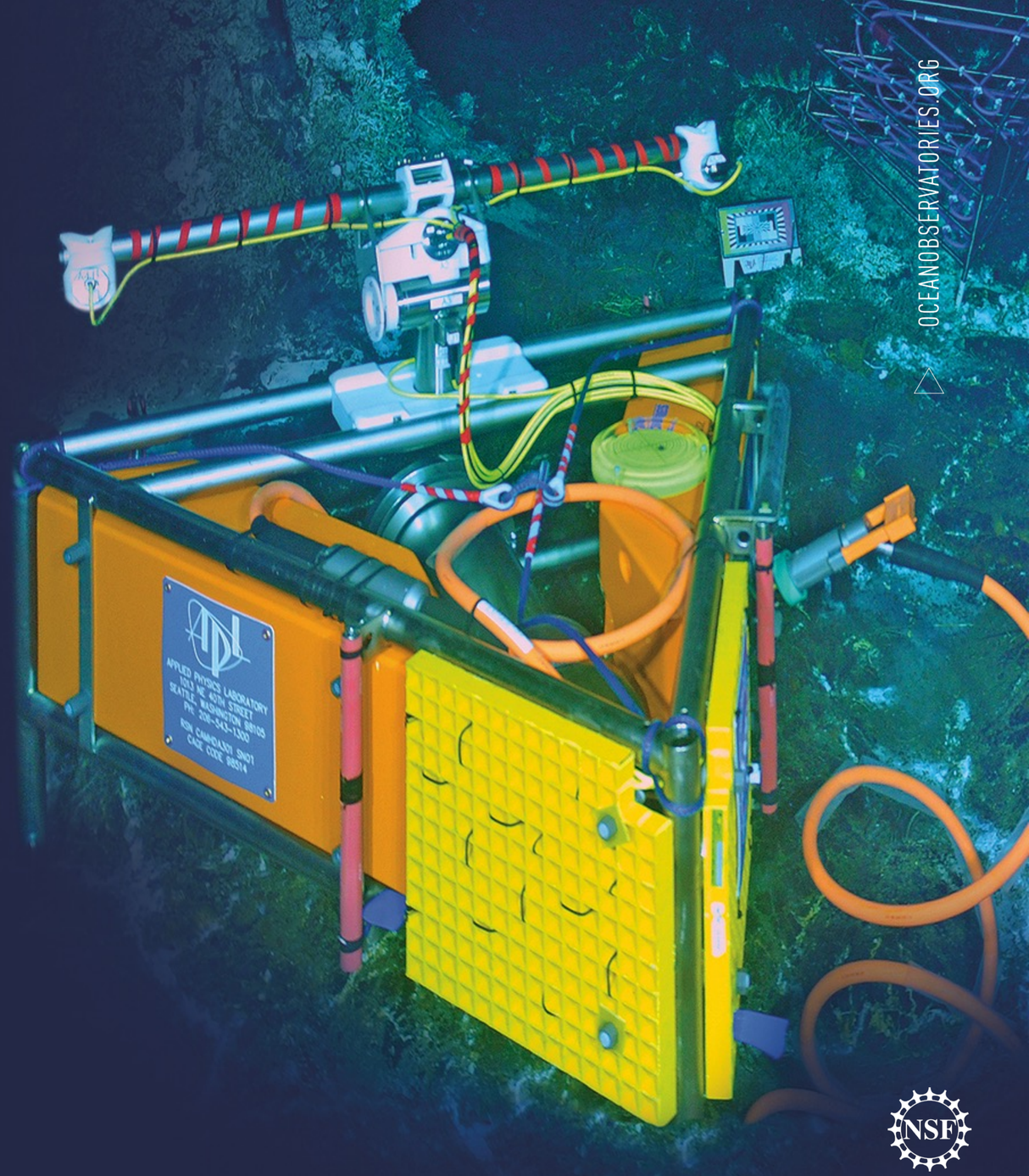
- Data Explorer v1.1 and 1.2 [Completed]
  - Continued increase in science data provided [Completed]
    - Added support for wire-following profilers, surface piercing profilers, cabled deep profilers, shallow profilers, and cabled VELPT.
    - Added subregions to ERDDAP for PHSEN data.
  - Support for PHSEN data through ERRDAP by implementing a method to flatten multi-dimensional data. [Completed]
  - New user feature to print custom configurations of data portal timeseries and data comparison charts. [Completed]
  - View discrete sample and glider data in grid view. [Completed]
  - Multiple viewing enhancements for gliders including new polygon search feature. [Completed]
  - Ingestion and viewing of discrete sample data. [Completed]
  - Prototype for visualization of timeseries data <1min resolution. [Completed]
  - Login function [Deferred]
    - Login function was primarily prioritized so that tracking of user requests could be reported on. A direct connection to Google Analytics for reporting has resolved this issue.
    - Login will still be added at a future date to better understand user behavior at an individual level and provide features such as saved downloads, history and data view sharing.
  - Multiple usability enhancements and defect fixes - view release notes for further details. [Completed]
  - Support Jupyter notebook access [Completed]
    - POC utilizing the Thredds Gold copy data source was completed and demoed. Application processed OOI data interactively in place without the need to first download.
- Media server for HD video, Hydrophone, HD pictures and Echograms [WIP]
  - Media files have been processed and back-end data delivery mechanism has been developed.
  - User interface has not been fully developed. Currently researching moving development into Data Explorer.
- Migration of real-time streaming from OOINET to Django [WIP]
  - Data delivery mechanism completed utilizing timeseries Postgres extension.
  - Realtime plotting visualization is being moved to Data Explorer instead of Django.

PYIII Year in Review - continued

- Instrument centric approach on ticket resolution [Ongoing]
  - Timestamp use across multiple instruments and deployments
  - HYDBB (dependency on vendor)
  - Pressure
- Pre-load and NetCDF compliance [Ongoing]
- “Redmine” redesign including public access [Completed]
  - Public access is through the Discourse implementation
- QA/QC – continue implementing QARTOD logic and supporting MIO testing [Ongoing]
- Complete Roundabout development for connectivity to uFrame [WIP]
  - The bulk of Roundabout development is complete. Work is till required to connect to uFrame. Little to no code changes will be required on the uFrame side. A web route needs to be built into Roundabout.
- Support [Ongoing]
  - Asset Management migrations
  - Software releases

OCEAN
OBSERVATORIES
INITIATIVE

Questions?



OCEANOBSERVATORIES.ORG

NSF