OOI 2.0 CI and Data Delivery Update

November 12, 2019

Jeffrey Glatstein
OOI Data Delivery Manager
Agenda

- CI achievements since last OOIFB meeting
- System availability
- Redmine ticket metrics
- Analysis of alternatives
- CI PYII work plan
- Cyber security
- System architecture graphic
- Organization
- Change management process overview
- Data usage management
- Management dashboard
- Collaborations
OOI 2.0 CI achievements since last OOIFB meeting

• Quality
  • QC proof of concept for gross range test
  • Adoption/implementation of QARTOD code set
  • Migration of ADCP bin depth fix

• Stability
  • Email of ingestion start, completion and error
  • Implementation of Nagios monitoring software
  • Implementation of Grafana system metrics tracking
  • Build-out of UAT system as mirror of production typography

• Data
  • Add time range filter to ingest process (data replay)
OOI 2.0 CI achievements since last OOIFB

• Metrics
  • Number of ingestions or data requests in queue
  • Number of particles (i.e. lines of data) entering the system on a 30 second interval
  • Tracking of Data Portal requests, actual time to service and size
  • Tracking of synchronous data requests
  • Tableau Beta
    • Redmine ticket metrics
    • Data request trending (Alpha)

• User Experience
  • M2M curl and Python examples posted on plotting page
  • Changed default plot style, time range and preferred timestamp
PY1 CI Availability

99.999%

(99.94% considering scheduled downtime)

- System and network utilization requirements have been exceeded.
- High percentage of uptime due to design and management best-practices.
- System utilization ~20% - Right-sized for anticipated future growth.

- PY1 CPU usage for Apache Cassandra and uFrame
OOI 2.0 Open Ticket Trend by Month

28% fewer tickets at close of PYI
• Bulk of the work is Enhancement, Bug and Support

• Many tickets represent user feedback

• Support represents the operational burden of the team

• Closed tickets start to outpace Open tickets as PYI progresses
OOI 2.0 Analysis of Alternatives - Process

• Process: started 1/19 and completed 6/19
  • Evaluation team: 1 voting member from each IO, PMO; DDCI representation
  • Set scope, goals, roles, evaluation criteria and procedures – socialize with program Stakeholders
  • Built list of alternatives (including OOI CI) and research
  • Researched alternatives – produced short list
  • Determined list of alternatives to formally evaluate

• Statistics
  • 34 vendors, technologies and reference resources researched
    • 6 measured by evaluation matrix
    • 7 determined not appropriate for replacement but good technologies to watch for future integration
    • 17 determined to not be a fit
    • 3 have a neutral status
OOI 2.0 Analysis of Alternatives - Findings

Report 2100-10011_AoA_Recommendations_2019-06-06_ver_1-00 released as controlled document by the OOI Control Change Board.

1. The current CI applications and architecture as they exist today are fixable. A full replacement is not required to achieve the OOI program’s mission.

2. The areas most in need of attention in the current CI system are user experience (UX), Asset Management and Quality Assurance/Quality Control (QA/QC).

3. There are very few existing end-to-end oceanographic data collection systems available that can come close to delivering 100% of the OOI requirements. Concentrating on an incremental, component by component “best of breed” approach, emulating, not re-inventing solutions would be more time and cost effective.

4. In order to deliver a more modern interface and meet users' needs for versioned data, OOI will need to move towards providing a pre-processed data source. This can be achieved with a hybrid model where calculate-on-demand is still available for those users who need it, but is not the primary means of data dissemination.

5. UX is the area most impactful to the end user and perception of the program, and arguably the area needing most improvement.
OOI 2.0 Analysis of Alternatives - Recommendations

It is the recommendation of this AoA panel that the current CI architecture is fixable and that it should be replaced only in part.

✓ Engage Axiom Data Science to re-architect the Data Portal UX – POC in PYI (Sep 2019), Project start PYII (Oct 2019)

✓ Develop the Roundabout asset tracking solution into a full asset management application utilizing WHOI IS, internal and Axiom developers – Project start PYII (Oct 2019)

✓ Continue with current QA/QC plan to improve the understanding of OOI data quality and leverage QARTOD standards and community – Project start PYI (Jan 2019)
OOI 2.0 CI PY II Work Plan

• Apply 2,200 hours to prioritized mission critical enhancements and bugs from the PI/PS reviewed ticket list to include:
  • Continue from PYI priorities
  • Fill values (analysis from 1.0)
  • Annotation indicators
  • Data Management – be able to maintain database entries more flexibly
  • Digital Object Identifier (DOI) on raw data
  • Instrument centric approach on ticket resolution of bugs, enhancements and data (e.g. ADCP, PCO2, HYDBB, etc…)

• Analysis of Alternatives recommended projects
  • Engage Axiom Data Science to re-architect the Data Portal and implement user experience improvements
  • Develop the Roundabout asset tracking solution into a full asset management application
  • QA/QC – continue implementing QARTOD logic and communicating QC results (dashboard)

• OOI CI Software stack upgrade

• Infrastructure
  • Disaster recovery planning
  • Implement Two Factor authentication
  • Technology refresh plan
  • Storage upgrades as needed

• Support hours – helping users, issue resolution, non-scheduled work, tasks not classified as bugs or enhancements.
OOI 2.0 Cyber Security Plan

• Engaged Praetorian Group to assess current OOI Cybersecurity, perform a penetration test and an analysis of current security policies (September 2019)

• Plan going forward:
  ✓ Address analysis findings through UI code changes (PYII)
  ✓ Implement recommendation of software based security assessment on continuous integration (2nd QTR PYII)
  ✓ Update policy document incorporating Praetorian feedback and reflecting current technology environment (2nd QTR PYII)
  ✓ Research and implement developer training (PYII)
  ✓ Plan next third party analysis of environment (PYIII)
  ✓ Schedule regular testing of security policy (PYII)
OOI 2.0 System Architecture

Cabled

Telemetered

Recovered

Data Ingestion

MIO Shore Servers

Rsync

Gold Raw Data Server

Public Raw Data Server

Rsync

Cassandra Cluster

PostgreSQL

ERDDAP

Data Portal Stream Engine

M2M

Data Download

Raw Data Request

West Coast Mirror

Data Staging

Ocean Instruments

Rutgers Data Centers

Instruments

Data Portal Stream Engine

Stream Engine

ERDDAP
OOI 2.0 Software Administration Move to PMO

• ECR 1300-00633 *Transfer scope & budget for CI Software Administration from RU to WHOI* passed review and approved by NSF 5/3/19

• Transition
  • Software stack migration PYI included 18 packages supporting uFrame
  • Software stack migration PYII includes 3 packages supporting work tools
  • Activities associated with migration were:
    • Software/System monitoring
    • Issue resolution
    • Physical migration of code
    • Upgrades and patches
    • User notifications
  • Asset management updates

• Benefits of change
  • Closer developer relationship with production events for support
  • Appropriate privileges to support issue research
  • End to end oversight will reduce latency of diagnosis and solution implementation
  • Advantages that a matrixed organizations brings

Transition completed on schedule - September 30, 2019
OOI 2.0 Change Management Process Overview

- Add new OOI Instrument
- Add new stream
- Enhancement of existing software
- Update definition of parser requirements
- Fix existing software defect

- Project Tools:
  - Jira
  - Redmine
  - Ticket Management
  - Google Sheets

- Software Development Life Cycle
- Techopedia.com
- oceanobservatories.org/information-for-researchers
- ooi-science@nsf.gov

- Engineering design and build process
- Definition of CI requirements

- Redmine Ticket Created
- Prioritized by PI/PM Review
- CCB Approval
- Approved by PMO

- Assigned and Worked by Dev
- Unit Tested
- Code Tested
- Code Reviewed by Dev Team
- SME Tested
- Deployed on UAT
- Deployed In Production

- Engineering design and build process
- Definition of CI requirements
OOI 2.0 Data Usage Management

• Questions that can now be answered:
  • How many rows of data are stored in the system?
  • How many instruments do we have data for?
  • How many data requests does the system service?
  • What is the most downloaded data set?
  • Who has submitted the most requests?
  • Who is downloading data from the system at this moment?
  • What is the status of the data ingestion process by user or stream?
OOI 2.0 CI Management Dashboard - Beta

System health at a glance
Difference of data applied from one 30sec period to the next

Cumulative total of data received in 30sec increments

State of ingestion
OOI 2.0 Management Dashboard - Detail

Ingestion Heartbeat: Success
- Total rows of data in system: 67025511742
- Total streams applied to system to date: 8041084
- Total Downloads to date: 16811996
- No User actively querying the system
- Ingestion is ready for data

System Activity

Most active users:
- bill.candy
- ooi.copp@pioneering.com
- leigh.ranta
- jrand5@uw.edu
- jgray@uw.edu
- jhearn@uw.edu
- mark.pro
OOI 2.0 Collaborations

- **Attended Large Facility Work Shop**
  - Met with many peers at other institutions to discuss topics such as disaster recovery, collaboration opportunities among large facilities and common data models.

- **Attended CI Large Facility Work Shop**
  - Presented lightening talk on Analysis of Alternatives process
  - Met with peers to discuss common issues effecting CI departments such as staffing, training and budgeting.

- **Self Evaluation**
  - Interviewed OOI stake holders
  - External service providers
  - Sub-awardees
  - Users

- **Analysis of Alternatives**
  - Met with vendors inside and outside of oceanography and science
  - Discussions with other oceanographic service providers

- **DDCI Membership**

- **QA/QC**
  - Conversation with Sail Drone to share approach on QA/QC
  - Work with other developers and SMEs at Ann Arbor code week
Questions