OOI Cyber-Infrastructure

Presented September 26, 2017
OOI CI Overall Architecture

Overall Capacity:

- ~25PB storage
- >100 servers (dual Xeon)
- 100G network backbone
- Palo Alto Firewalls

Robust, secure, and scalable production CI
Integrated Software Stack

OOI Data Portal + M2M

OOI Data

THREDDS

Raw data

uFrame Service-Oriented Infrastructure

OOINet

oceanobservatories.org

Alfresco (cruise and doc CM)

ERDDAP

cassandra

PostgreSQL

Qpid

RabbitMQ
UFrame Update

• UFrame/source code is available to OOI program without any limitation, based on all data provided
  – Final determination to be made by COL/NSF

• UFrame-based OOINet expertise/documentation exists
  – COL contracting for additional documentation

• Rutgers CI team is ready for UFrame-based OOINet operation and maintenance
  – Highly qualified and experienced personnel with the required skills
  – Knowledge transferred to Rutgers CI Team (e.g., has performed major modifications of the key components of the system)
UFrame Ingestion Functionality

• UFrame ingestion functionality developed, tested and delivered

• Demonstration delivered with easy-to-use capabilities (OOIFB members attended)

• Extensive training and support provided to data team

• RESTful API via M2M
  – Graphical clients available
  – Simple command-line scripting:
    
    ```
    ./ooi-m2m.sh -i GA02HYPM_D00001_ingest.csv (purge)
    ./ooi-m2m.sh -i GA02HYPM_D00001_ingest.csv (ingestion)
    ```
Ingestion Coordination Spreadsheet

- Provided access to ingestion functionality via M2M
  - Testing environments
  - Production environment

https://docs.google.com/spreadsheets/d/15CZiUCoRvITZGVhq4uqNwCOuZYKC0iTF1WrwmwAGuX8/
Overall CI Software Schedule

Legend:
- Green: Completed
- Orange: Pending (current target)
- Blue: To be scheduled
- Black Diamond: Deployment in production
- Black Arrow: Demo

Sprint 6 (completed)
Sprint 7 (completed)
Sprint 8 (completed)
Sprint 9 (completed)
Sprint 10 (current)
Sprint 11
Sprint 12

1/13 2/10 3/7 4/20 6/8 7/27 9/14 11/2
3/13 4/10 5/7 6/20 7/14 9/1 10/20 12/8

2017
Software Schedule – Updates

• Enhancement efforts **focused on data delivery**
  – Integration improvements are for data delivery work
• CI system issues have been moved to Rutgers CI scope of work
  – Examples: FOSS, Cassandra upgrades
• SW schedule until end of 2017 expected to be boarded
• Current status of Redmine tickets

<table>
<thead>
<tr>
<th>Redmine Category / Open Issues</th>
<th>Commissioning</th>
<th>Sep, 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drivers/Parsers/Data Products</td>
<td>169</td>
<td>118</td>
</tr>
<tr>
<td>OOINet Integration</td>
<td>55</td>
<td>113</td>
</tr>
<tr>
<td>UFrame</td>
<td>231</td>
<td>99</td>
</tr>
<tr>
<td>User Interface</td>
<td>394</td>
<td>126</td>
</tr>
<tr>
<td>Entered without a Project name</td>
<td>111</td>
<td>0</td>
</tr>
<tr>
<td>Other Projects</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>1,014</strong></td>
<td><strong>456</strong></td>
</tr>
</tbody>
</table>
Support for ERDDAP development

• Best practices followed in the ERRDAP environment deployment
  – DTAP model variant (i.e., quality of code, code review, testing unknown)
  – Deployment and source code under configuration management

• Architecture design by data team
  – Integration viability in OOI infrastructure discussed on April 2016
  – Resources/access provided in May 2016
  – Deployment requests to production begin on Nov 2016 (#11603, #10169, #12716)
  – ~300 Redmine notes involving Rutgers CI

• Testing environment
  – Full access to developers and no limitations to expedite ERDDAP development
  – Ingestion functionality fully available in the test system associated to ERDDAP

• Production environment
  – ERRDAP production log files fully available to developers (as other log files in production)
  – Deployment upon request (request under configuration management w/ rollback, etc.)
    – 11 deployment requests in 10 months (~1 request per month)
    – All deployments required follow up
## CI Operations Update

- **CI availability**
  - In 2017, OOI Cyberinfrastructure **availability** exclusive of scheduled downtime **has been 100%**.
  - It exceeds the requirement of “The CI network shall provide 97% availability exclusive of scheduled downtime”

- **Cyber-security**
  - 1,250 attacks mitigated January-June, 2017

- **Deployments to production**

<table>
<thead>
<tr>
<th>Date</th>
<th>Component</th>
<th>Notes</th>
<th>Date</th>
<th>Component</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/2/2017</td>
<td>GUI</td>
<td>Deployed 1.3.2 bug fix version</td>
<td>8/2/2017</td>
<td>GUI</td>
<td>1.4.3 and 1.4.4</td>
</tr>
<tr>
<td>1/26/2017</td>
<td>GUI</td>
<td>Deployed 1.3.0 feature version, and 1.3.1 bug fix version shortly after the same day</td>
<td>8/2/2017</td>
<td>uFrame</td>
<td>1.2.5</td>
</tr>
<tr>
<td>1/26/2017</td>
<td>uFrame</td>
<td>Deployed 1.2.1, including the previous 1.1.3 release which was not deployed before</td>
<td>8/15/2017</td>
<td>GUI</td>
<td>1.4.5 and 1.4.5.1</td>
</tr>
<tr>
<td>3/6/2017</td>
<td>GUI</td>
<td>Deployed 1.4.0</td>
<td>8/15/2017</td>
<td>uFrame</td>
<td>1.2.6</td>
</tr>
<tr>
<td>3/6/2017</td>
<td>uFrame</td>
<td>Deployed 1.2.2</td>
<td>9/5/2017</td>
<td>GUI</td>
<td>1.4.5.1</td>
</tr>
<tr>
<td>4/3/2017</td>
<td>uFrame</td>
<td>Deployed Sprint 6 features</td>
<td>9/5/2017</td>
<td>uFrame</td>
<td>1.2.6-sr1</td>
</tr>
<tr>
<td>5/22/2017</td>
<td>GUI</td>
<td>Deployed 1.2.3</td>
<td>9/5/2017</td>
<td>ERDDAP</td>
<td>Uncabled frontend - Version Number: 1, Uncabled backend - Version Number: 1</td>
</tr>
<tr>
<td>5/22/2017</td>
<td>uFrame</td>
<td>Deployed 1.4.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Download Statistics (Jun’16 – Jun’17)

<table>
<thead>
<tr>
<th></th>
<th>OOINet (UI Portal)</th>
<th>THREDDS Server</th>
<th>Raw Data Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visits</td>
<td>28,341</td>
<td>3,681</td>
<td>18,829</td>
</tr>
<tr>
<td>Distinct countries</td>
<td>104</td>
<td>36</td>
<td>57</td>
</tr>
<tr>
<td>Direct entries</td>
<td>22,446 (79%)</td>
<td>3,324 (90%)</td>
<td>17,021 (90%)</td>
</tr>
<tr>
<td>Search engines</td>
<td>227 (1%)</td>
<td>51 (1%)</td>
<td>(&lt;1%)</td>
</tr>
<tr>
<td>From websites</td>
<td>3,228 (26%)</td>
<td>306 (8%)</td>
<td>1,792 (10%)</td>
</tr>
<tr>
<td>Distinct websites</td>
<td>131 (540 distinct URLs)</td>
<td>17 (92 distinct URLs)</td>
<td>30 (158 distinct URLs)</td>
</tr>
</tbody>
</table>

**Data transferred**
- OOINet: 75.31 GB
- THREDDS Server: 923.3 GB
- Raw Data Server: 41.85 TB

**Top Countries by Visits**

**OOINet**
- United States: 22,029
- China: 1,942
- Brazil: 965
- France: 659
- United Kingdom: 426

**Raw Data**
- United States: 16,989
- China: 929
- Germany: 215
- Canada: 127
- Brazil: 116
- Israel: 84
Thank You!

facebook.com/rutgersRDI2
@RDI2_rutgers
rdi2.rutgers.edu