ID46- ARCHITECTING THE CYBERINFRASTRUCTURE FOR THE NATIONAL SCIENCE FOUNDATION OCEAN OBSERVATORIES INITIATIVE (OOI)

IVAN RODERO[172] AND MANISH PARASHAR[173]

Abstract – The NSF Ocean Observatories Initiative (OOI) is a networked ocean research observatory with arrays of instrumented water column moorings and buoys, profilers, gliders and autonomous underwater vehicles (AUV) within different open ocean and coastal regions. OOI infrastructure also includes a cabled array of instrumented seafloor platforms and water column moorings on the Juan de Fuca tectonic plate. This networked system of instruments, moored and mobile platforms, and arrays will provide ocean scientists, educators and the public the means to collect sustained, time-series data sets that will enable examination of complex, interlinked physical, chemical, biological, and geological processes operating throughout the coastal regions and open ocean. The seven arrays built and deployed during construction support the core set of OOI multidisciplinary scientific instruments that are integrated into a networked software system that will process, distribute, and store all acquired data. The OOI has been built with an expectation of operation for 25 years.

Keywords – Research Infrastructure, Cyber-Infrastructure.

I. INTRODUCTION

The National Science Foundation-funded Ocean Observatories Initiative (OOI) is an integrated infrastructure project composed of sciencedriven platforms and sensor systems that measure physical, chemical, geological and biological properties and processes from the seafloor to the air-sea interface.

The OOI network was designed to address critical sciencedriven questions that will lead to a better understanding and management of our oceans, enhancing our capabilities to address critical issues such as climate change, ecosystem variability, ocean acidification, and carbon cycling.

The OOI has transformed research of the oceans by integrating multiple scales of globally distributed marine observations into one observing system and allowing for that data to be freely downloaded over the internet in nearreal time. The OOI will continue to deliver data and data products for a 25-year-plus time period within an expandable architecture that can meet emerging technical advances in ocean science.

Building on last century's era of ship-based expeditions, recent technological leaps have brought us to the brink of a sweeping transformation in our approach to ocean research – the focus on expeditionary science is shifting to a permanent presence in the ocean. As technological advances continue over the lifetime of the OOI, developments in sensors, computational speed, communication bandwidth, Internet resources, miniaturization, genomic analyses, high-definition imaging, robotics, and data assimilation, modeling, and visualization techniques will continue to open new possibilities for remote scientific inquiry and discovery.

The OOI is funded by the National Science Foundation and is managed and coordinated by the OOI Program Office at the Consortium for Ocean Leadership (COL), in Washington, D.C. COL is leader, owner, and operator of the OOI and its infrastructure. Implementing Organizations (IOs), subcontractors to COL, are responsible for construction and development of the different components of the program. Woods Hole Oceanographic Institution is responsible for the Coastal Pioneer Array and the four Global Arrays, including all associated vehicles. Oregon State University is responsible for the Coastal Endurance Array. The University of Washington is responsible for cabled seafloor systems and moorings. Rutgers, The State University of New Jersey, is implementing the Cyberinfrastructure component. The OOI Data Management and education and public engagement team is co-located with the Cyberinfrastructure group at Rutgers University [1].

II. OBSERVATORY COMPONENTS

The design of the OOI enables multiple scales of marine observations integrated into one observing system via common design elements and overarching, interactive Cyberinfrastructure Technology. The coastal assets of the OOI expand existing observations off both U.S. coasts, creating focused, configurable observing regions. Cabled observing platforms 'wire' a single region in the Northeast Pacific Ocean with a high speed optical and high power grid. And the global component addresses planetary-scale changes via moored open-ocean buoys linked to shore via satellite.

This unprecedented and diverse data flow is coming from 89 platforms carrying over 830 instruments which provide over 100,000 scientific and engineering data products. Design, construction, and full deployment of these systems was completed late last year. This infrastructure includes 12 surface moorings, 8 subsurface flanking moorings, 22 profiler moorings, 20 cabled seafloor packages, 32 gliders, and 2 Autonomous Underwater Vehicles (AUVs). Some of this cutting edge instrumentation had never been fielded in an operational format before and are now in the water and actively collecting data, including the in situ massspectrometer, particulate DNA sampler, and other vent chemistry sensors [1]. The map with the location of the seven OOI arrays is shown in Figure 1.

III. OOI CYBER-INFRASTRUCTURE SERVICES

The primary functions of the OOI Cyber-Infrastructure are data acquisition/collection, storage, processing and delivery.

(a) Data Collection and Transmission to the OOI Cyberinfrastructure:

Data is gathered by both cabled and un-cabled (wireless) instruments located across multiple research stations in the Pacific and Atlantic oceans. Once acquired, the raw data (consisting mostly of tables of raw instrument values – counts, volts, etc.) are transmitted to one of three operations centers: Pacific City, directly connected via fiber optic cable to all cabled instruments in the Cabled Array; Oregon State University (OSU), an OperationalManagement Center (OMC) responsible for all un-cabled instrument data on the Pacific coast; and Woods Hole Oceanographic Institute (WHOI), the OMC for Atlantic coast-based un-cabled instrument data. The data from the operations centers is transferred to the OOI Cyberinfrastructure for processing, storage and dissemination.

(b) Data Management, Storage, and Processing:

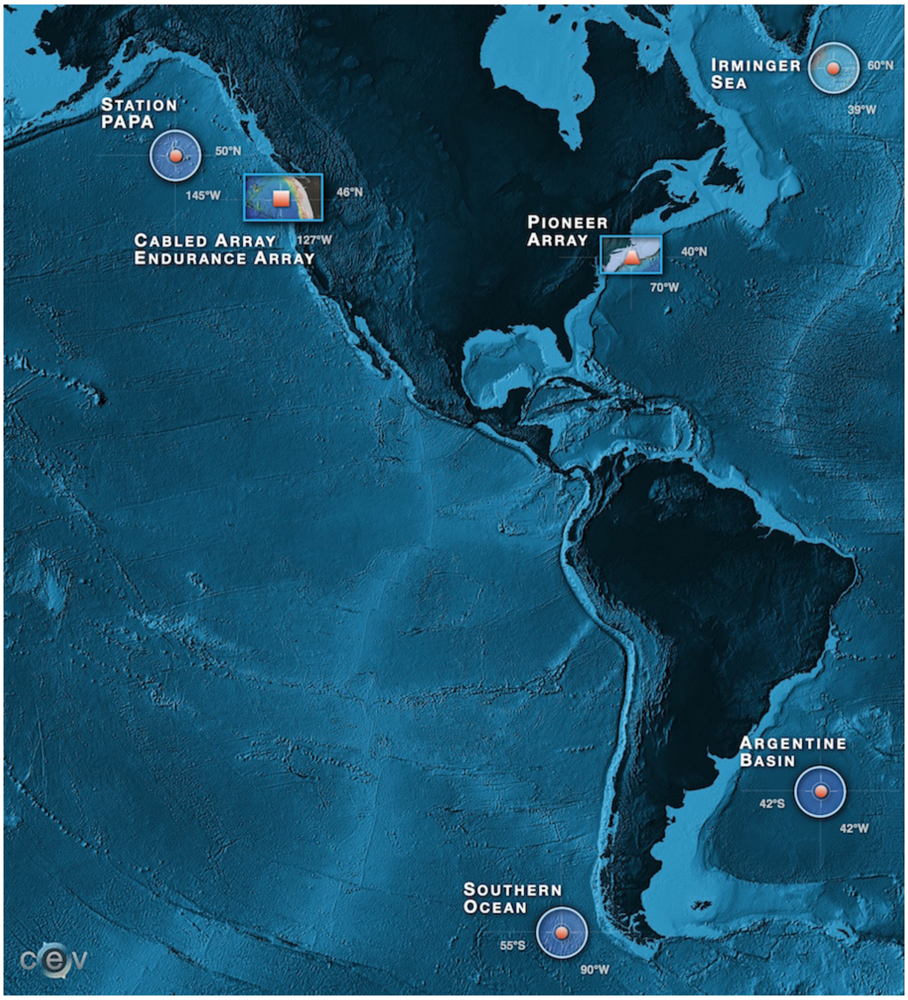
Two primary cyber-infrastructure (CI) centers operated by the Rutgers Discovery Informatics Institute (RDI2) are dedicated to OOI data management: the West Coast CI in Portland, Oregon, and the East Coast CI, at Rutgers University. While data from the Cabled Array components are initially received at the Shore Station in Washington, it is the East Coast CI that houses the primary computing servers, data storage and backup, and front-facing CI portal access point, all of which are then mirrored to the West Coast CI over a high-bandwidth Internet2 network link provisioned by MAGPI (Mid-Atlantic GigaPOP in Philadelphia) on the east coast and PNWGP (Pacific- Northwest GigaPOP) on the west coast. The data stores at the OMCs at OSU and WHOI are continuously synchronized with the data repositories located at the East and West Coast CI sites.

(c) Data Safety & Integrity:

Data safety and protection is ensured in two ways: data security and data integrity. Data security is addressed through the use of a robust and resilient network architecture that employs redundant, highly available nextgeneration firewalls along with secure virtual private networks (VPN). Data integrity is managed through a robust and resilient Information Life-cycle Management (ILM) architecture.

(d) Public Data Access:

The OOI CI software ecosystem (OOINet) employs the uFrame software framework that processes the raw data and presents it in visually meaningful and comprehensible ways in response to user queries, which is accessible over the



Internet through the CI web-based portal access point [2]. The OOI CI portal also delivers command and control functionality. In addition to the portal, OOI CI provides the following data delivery methods:

• THREDDS Data Server: delivers any data and associated metadata that have passed the evaluation process are accessible as individual and aggregated datasets, via OPeNDAP and NetCDF Subset Services. It includes pre-calculated preliminary data sets for platforms identified by the OOI Science Oversight Committee and data products requested through the CI portal (i.e., generated asynchronously).
• Raw Data Archive: delivers data as they are received directly from the instrument, in instrument-specific format.
• Alfresco Server: provide cruise data, including shipboard observations.
• Other methods for data delivery will be available such as a machine-to-machine interface and an ERDDAP server.

OOI CI software ecosystem permits 24/7 connectivity to bring sustained ocean observing data to a user any time, any place. Anyone with an Internet connection can create an account or use CILogon (Federated Authentication) [3] and access OOI data.

IV. DESING AND IMPLEMENTATION ISSUES
The OOI CI design and implementation principles are based on industry best practises for the different aspects of the CI. The approach is based on a decentralized but coordinated architecture, which is driven by requirements, e.g., data storage capabilities, system and load stability, security, etc.

(a) Redundancy and resiliency:
The OOI CI is a mirrored infrastructure for high availability, disaster recovery and business continuity. It implements a resilient Information Life-cycle Management (ILM) architecture that integrates redundant enterprise storage area network (SAN) (disk-based) and a robotic library (tape-based). Redundancy is implemented at different layers, for example, an enterprise-level storage network of multiple hard drives managed by an intelligent device manager, reduces the data footprint by reducing data duplication while maintaining data integrity and access performance through storage redundancy, and tape storage, a "last tier" storage that is not dependent on power or cooling, supports longer-term backup and archiving, disaster recovery, and data transport.

(c) Service-oriented Architecture:
The core of the OOI CI software ecosystem (Uframebased OOINet) is based on a Service Oriented Architecture (SOA), a set of data dataset/instrument/ platform drivers and data product algorithms, which plug in to the uFrame framework. Uframe-based OOINet uses latest generation technologies for big management data such as Apache Cassandra, which is a state-of-the-art, scalable and highly available distributed database management system designed to handle large amounts of data. Uframe-based OOINet services are exposed through a web-services API and will be available as the machineto- machine interface for external access. The use of a well defined API based on standard protocols will enable other systems to interface and interact with OOI CI.

(c) Cyber-security:
The system is based on a multi-tier security approach with dedicated and redundant (highly available) appliances at the CI perimeter. The OOI CI implementation supports encryption of traffic, network traffic segregation, multilayer traffic filtering, multi-layer access control and comprehensive monitoring. Further, data delivery to external users is implemented through dedicated and distinct storage appliances (i.e., physical and logical isolation from core storage infrastructure) In addition to implementing industry best practices, the OOI CI cyber-security effort includes a comprehensive cyber-security program based on

engagement with the NSF Center for Trustworthy Scientific Cyber Infrastructure (CTSC) [4]. This program encompasses a set of policies and procedures as listed below:

• Master Information Security Policy & Procedures: represents the core information security policies and procedures, including information security-related roles and responsibilities; references to other, special purpose policies; and the core procedures for developing, implementing, and maintaining the information security program.
• Acceptable Use Policy: set of rules that a user must agree to follow in order to be provided with access to a network and/or resources. Used to reduce liability and act as a reference for enforcement of policy.
• Access Control Policy: defines the resources being protected and the rules that control access to them.
• Asset Management Policy: requirements for managing capital equipment including: inventory, licensing information, maintenance, and protection of hardware and software assets
• Information Classification Policy: used to ensure consistency in classification and protection of data.
• Disaster Recovery Policy: contains policies and procedures for dealing with various types of disasters that can affect the organization.
• Personnel Exit Checklist: form to be completed at the end of employment that addresses revoking access to resources, physical space and the return of organizational assets.
• Incident Response Procedures: a pre-defined organized approach to addressing and managing a security incident.
• Password Policy: a set of rules designed to establish security requirements for passwords and password management.
• Physical [and Environmental] Security Policy: details measures taken to protect systems, buildings, and related supporting infrastructure against threats associated with their physical environment.
• Training and Awareness Policy: outlines an organization's strategy for educating employees and communicating policies and procedures for working with information technology (IT).

Regular vulnerability scans/audits (both internally and externally) are also performed to the OOI CI.

V. CONCLUSION

The OOI network was designed to support of the needs of users to conduct research across a wide range of science themes, within an expandable observing infrastructure that spans widely differing ocean domains. The network is designed to provide observations of processes at multiple oceanographic scales, from ocean basin to continental shelf, over time scales from, short-term episodic events to decadal cycles.

OOI CI has initiated its operational phase and data (including science, engineering and data products) flowing from those instruments is freely available to users. The OOI CI portal provides all data, metadata and data processed via conventional algorithms or direct retrieval from OOI storage or data archives. Data quality and data management will utilize generally accepted protocols, factory calibrations and at sea calibration procedures.

During the first months of its operation, OOI community has been growing every day and is made up of a diverse set of users from 180 different organizations from around the world. At least 500 people has already registered on the OOI Data Portal, it has 3,000 unique visitors to the OOI website each month. In June alone, users downloaded over 6TB of data.

REFERENCES

[1] Ocean Observatories web site, "www.oceanobservatories.org", accessed: August 2016.
[2] OOINet web portal, "https://ooinet.oceanobservatories.org", accessed: August 2016.
[3] CI Logon web site, "www.cilogon.org", accessed: August 2016.
[4] NSF Center for Trustworthy Scientici Cyberinfrastructure web site, "http://trustedci.org", accessed: August 2016.

OOI is a NSF-funded effort and involves teams from Consortium for Ocean Leadership, Woods Hole Oceanographic Institution, Oregon State University, University of Washington, Rutgers University, Raytheon and RPS ASA. This document summarizes the contributions from these teams. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

ID47- OBSTACLE DETECTION ALGORITHM OF LOW COMPUTATIONAL COST FOR GUANAY II AUV

C. GALARZA[176], J. PRAT[177] AND S. GOMÀRIZ[175]

Abstract – Obstacle detection is one of the most important stages in the obstacle avoidance system. This work is focused to explain the operation of a designed and implemented for the overall detection of objects with low computational cost strategy. This strategy of low computational cost is based on performing a spatial segmentation of the information obtained by the SONAR and determine the minimum distance between the SONAR (AUV) and the obstacle.

Keywords – Obstacle Detection, Real Time, Low Computational cost.

I. INTRODUCTION

The obstacle avoidance system for AUVs are divided into two phases: the first obstacle detection and the second consisting of a strategy of evasion. At work [1], an obstacle avoidance system was designed for the vehicle Guanay II. In this work, the phase of obstacle detection is subdivided into three general stages. The first stage corresponds to the configuration of SONAR and communication with the control unit. The second stage is related to the configuration of the operating parameters of obstacle detection system [2]; these parameters affect the execution time of the algorithm. Finally, we have in the third stage the strategy of obstacle detection, from the data provided by the SONAR.

Initially in previous work [3], we have designed and implemented a detection strategy, based on the identification of objects automatically on an image, constructed from data provided by the SONAR.

The measured time of execution of this strategy is 22.7s. At that time, it includes the time for the acquisition of measurements (10.89s), the time used to process the image and perform the obstacle detection (11.89s).