Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia, 9-13 July 2012
5B Sensor networks and their applications

CREON – Integrating Disparate Sources of Remote Coral Reef Sensor Data

Krisanadej Jaroensutasinee[1], Mullica Jaroensutasinee[1], Scott Bainbridge[2], Tony Fountain[3], Sally J. Holbrook[4], Michael Nekrasov[3]

[1]CoE for Ecoinformatics, School of Science, Walailak University, Nakhon Si Thammarat 80161 Thailand
[2]Australian Institute of Marine Science, PMB 3, Townsville MC, QLD, 4810 Australia
[3]California Institute of Telecommunications and Information Technology, UCSD, CA, 92093 USA
[4]Marine Science Institute, UCSB, Santa Barbara, CA, 93106 USA
Corresponding author: krisanadej@gmail.com

Abstract. The Coral Reef Ecological Observatory Network (CREON) is a grass-roots, self-assembled group of coral reef scientists and engineers working to advance an understanding of coral reef dynamics by facilitating the development of practical environmental sensor networks. One of the goals of the group is to develop integrated data products and outcomes from the various CREON sites. This paper presents an overview of the use of cloud computing based data architectures, such as Pachube, to deliver integrated data outcomes using a range of disparate real time data from each site. The outputs include simple event detection systems with alerts delivered to social network sites such as 'Twitter'. This allows the multitude of collected data to be reduced to a small number of interesting or meaningful events. The system is being used to identify local events that may have a global context and as a tool for monitoring these sites. The group also has developed a suite of technologies that looks to lower the cost of entry for deploying sensor networks on coral reefs in areas that are poorly or under-sampled. A second goal of CREON is to include a greater range of real time coral reef sensor and other data through the use of an open, common cloud based computing infrastructure. This may include providing technology transfer to new sites and groups that are not currently engaged, especially those in areas that do not currently have access to sensor network technologies.

Key words: Sensor network, Coral, CREON, Environmental observing.

Introduction

Mass coral bleaching has been reported to be increasing in recent years and is directly correlated with increases in sea temperatures driven by climate change (Hoegh-Guldberg 1999, Hughes et al. 2003). The bleaching pattern in reef corals is a generalised response to stress and occurs as a reaction to many environmental factors such as high and low temperature (Hoegh-Guldberg and Smith 1989, Glynn 1991, Hoegh-Guldberg and Fine 2004), high light (Lesser et al. 1990), aerial exposure (Leggat et al. 2006), cyanide (Jones and Hoegh-Guldberg 1999), low salinity (Kerswell and Jones 2003) and herbicides (Jones 2005).

The Coral Reef Environmental Observatory Network (CREON) is a group of international institutions made up of scientists and engineers whose goal is to develop tools for coral reef research. It is a collaboration between a diverse team of ecologists, computer scientists, and engineers from the Marine Science Institute at the University of California Santa Barbara (www.msi.ucsb.edu/), the California Institute of Telecommunications and Information Technology at the University of California San Diego (CalIT2 UCSD, www.calit2.net), the Australian Institute of Marine Science (AIMS, www.aims.gov.au) and the Centre of Excellence of Ecoinformatics, NECTEC-Walailak University.

The deployment builds on the experiences of CREON members in establishing coral reef observatories that share and interchange data from multiple sites around the Pacific Rim (Fountain et al. 2009a,b). It is envisioned to be a living laboratory for long-term studies of marine ecology and as a test-bed for evolving technologies for environmental and biological sensing, communications and analysis. Using a variety of platforms and instruments, the CREON group hopes to solve some of the more technical aspects in a collaborative framework. Sensors used in coral reef observatories can generate surprisingly large amounts of data. The intersection of sensor technologies with optical networking represents a major research thrust.

Networks have been deployed on the Great Barrier Reef as part of the Australian Integrated Marine Observing System (IMOS) (http://imos.org.au), in Moorea in French Polynesia as part of the US Long Term Ecological Research (LTER) network Moorea

Coral Reef site (www.mcr.lternet.edu), at Racha Island in Thailand through Walailak University and the National Electronics and Computer Technology Centre (NECTEC) and in Kenting National Park in southern Taiwan through the National Centre for High-Performance Computing (NCHC) and Academia Sinica, Taiwan.

One of the goals of CREON is to integrate data coming from the various sites in order to ask broad scale questions using the collective data. This requires true data integration, the ability to deal with potentially large data sets and the ability to extract, from this large amount of data, events and patterns of interest. This paper presents an overview of the use of cloud computing based data architectures, such as Pachube™ (www.pachube.com), to deliver integrated data outcomes using a range of disparate real time data from each site.

Materials and Methods

Study sites

CREON sites such as Davies Reef (Lat: 18.83S, Long: 146.63E) on the Great Barrier Reef in Australia and Racha Island (Lat: 7.60488N Long: 98.3766E) in southern Thailand (Fig. 1a,b), were used as source sites for the integration work.

Figure 1: CREON sites. (A) Racha Island, Thailand and (B) Davies Reef on the Great Barrier Reef, Australia

Davies Reef is a lagoonal mid-shelf reef located some 70 kilometres off the Queensland coast in north-east Australia. The sensor network deployed consists of an above water weather station (Vaisala WXT-520 (Vaisala 2011) and a number of buoys with temperature [MEA Thermistors (MEA 2011) and SeaBird SBE39 (SeaBird 2011a) temperature / pressure sensors] and salinity sensors [SeaBird SBE37 (SeaBird 2011b) CTD units – Conductivity (salinity), Temperature and Depth (pressure)]. Recently a set of underwater light metres [LiCor LI-192 (LI-COR 2011)] measuring Photosynthetically Active Radiation (PAR) (Wikipedia 2011) and an underwater camera have been added to the site.

Racha Island is located south of Phuket Island in the south of Thailand on the Andaman Sea. The island has a number of fringing reefs and is a popular holiday destination being only a twenty minute boat ride from Phuket. The study site is located in a bay managed by a local resort (Banraya Resort and Spa) and has been set up in conjunction with the resort owners. The instrumentation includes a Davis Vantage Pro II Plus (Davis 2011) weather station located on the hill behind the resort, a number of logging (non-real time) HOBO Pendant (Onset 2011) temperature and light loggers, a real-time SeaBird SBE37 (SeaBird 2011) CTD that records conductivity (salinity), pressure (depth) and temperature (the same instrument used on the Davies Reef deployment) and a number of EcoCams (underwater video cameras).

Data integration

Initially, each country was to host and manage their own data and then make the data available via Web Map Services (WMS) and Web Feature Services (WFS). While this would provide universal access the complexities of this approach along with issues such as firewalls / security, maintaining externally facing servers and delays in making the data available made this approach less than satisfactory.

Using a different approach, it was decided that the data from each sensor network would be stored in a cloud-based computing infrastructure using a service based architecture allowing for data access and storage. Each field unit would 'push' its data into the cloud based store where it would be available for other services and processes such as event detection. This removes the idea of distributed data centres and replaces it with a single uniform cloud based system. With all of the data in one place and accessible via simple and transparent service calls, it becomes possible to do truly integrated queries and to build processing chains that deliver real time multi-parameter event detection and analysis and subsequently to generate information products.

Data standardisation

Most data conformed to a simple pattern of time, location-depth and value represented as X / Y / Z / T / Value and typically by latitude (X), longitude (Y), depth (Z), Date/Time (T) and the value being measured. This data quintet is then directly linked to parameter definitions, units of measure and a range of ancillary data such as descriptions of the sensor systems taking the measurement, metadata describing how and why the measurement is being made, the project or data owner, links to previous / next data and so on.

This analysis matches those done by others to develop schemas for implementations of the Open Geospatial Consortium (OGC) Sensor Observation Service (SOS) (OGC 2007) such as that implemented by 52 Degrees North (52 Degrees North 2011). These typically define Features of Interest which are spatial locations where observations occur (the X / Y / Z in our model schema), phenomena which are defined characteristics that can be measured (such as temperature, pressure, radiation, etc.), offerings which are implemented instances of phenomena and procedures which are mechanisms whereby measurements are made. SensorML records are used to describe the observing hardware and the process where the real world measurement is converted into an observation. Additional metadata can be delivered via International Standards Organisation (ISO) 19115 (ISO 2003) or equivalent spatial metadata systems, again as a service, linked into the main data.

Cloud computing

The CREON group investigated the use of cloud-based observational data stores to make the data freely available to the processes and services that will, in the future, be the backbone of how new data and information are delivered. The initial plan was to use the 52 Degrees North Sensor Observation Service (SOS) (52 Degrees North 2011). The problem with this approach was that none of the agencies were set up to be service providers. As an alternative, a commercial or quasi-commercial solution was investigated.

Pachube (www.pachube.com) was selected by the CREON group. Pachube offers a services based architecture, a range of graphing and visualisation tools, event detection via triggers and cost-effective data storage. The main issue is that the data format is not OGC SWE compliant (i.e., not using Observation and Measurement ML) but rather it uses EEML (EEML 2011) as its data format. While this may have future implications for direct integration of the Pachube services with other OCG compliant services, the EEML format is very simple and conforms to the X / Y / Z / T / Value schema discussed earlier. The service holds a wide range of observational data, enabling data products that span a range of observation environments and hooking more directly into the Internet of Things paradigm. For example, using the Pachube infrastructure, it becomes possible to link the data to social networking systems, to internet connected devices and into a range of graphing and visualisation systems.

Data infrastructure

At both sites the underwater instruments are interfaced to a commercial logger and then locally integrated using the DataTurbine (Fountain et al. 2009a) middleware which pulls data from each of the instruments, gives them a common date/time stamp and presents them in a uniform fashion for further use. Currently the data are extracted from DataTurbine into databases and also directly to MatLab (MathWorks 2011) for further processing. The systems running the DataTurbine middleware are Internet connected and so it is possible to push the data directly from DataTurbine to the cloud based storage to integrate the data from the two sites. Initially a small Java program was written to push the data from data files created by the logger to the cloud based data systems, but this was replaced by a new program that integrates DataTurbine directly with the cloud based systems. The program was written by the group and uses a small XML configuration file to configure the relationship between the data on the local DataTurbine system and that on the cloud storage.

The use of a ‘push’ model, where the local system pushes the data up to the cloud based storage, removes most firewall or security issues as most firewalls are tolerant of data going out of the system as distinct from external agents coming into the network to retrieve data. This model therefore proved to be easy to implement and one that worked with, not against, institutional IT systems.

Results

Small programs were written to initially interface the data files produced by the logger systems and then the DataTurbine middleware into the Pachube infrastructure. The programs use the DataTurbine API [23] to request data and then format the data as an EEML XML packet and then do an HTTP ‘put’ request to the Pachube API. The software can get the date and time of the last data in Pachube and use this to know what data to extract from the DataTurbine system.

Two Pachube data feeds were created, one for each site, and within these a number of sub-feeds were established for each particular data stream. Initially these were set to water temperature, salinity and depth and air temperature. Data from early 2011 were pushed up to the system from each site allowing for comparisons between the sites.

In order to visualise the integrated data a simple web site was developed (www.coralreefeon.org). This allows for the data from each site to be graphed and compared delivering, in a simple way, data integration from the two CREON sites. A simple event detection system, written using the Pachube services, was developed with the results being delivered automatically to Twitter®. In this case two threshold events were defined, one where the local water temperature in Thailand went above 30 °C and one where the difference in water temperature between the two sites was more than 10 °C.

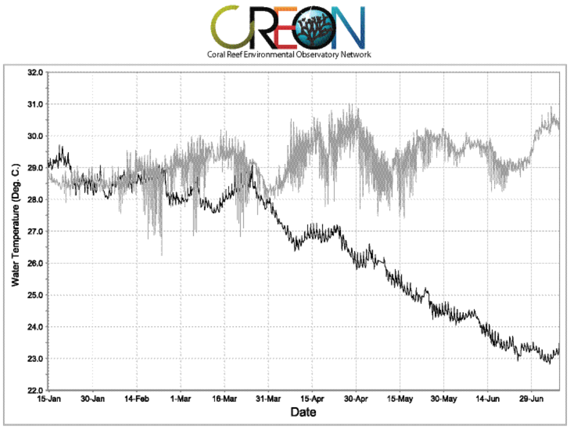

Figure 2: Temperature from the two CREON sites. Grey and black lines represent Racha Island, Thailand and Davies Reef, Australia

Temperature data from the two sites for early 2011 are shown in Fig. 2. The sites have differing seasonality as shown by the temperature decline during this period as the southern hemisphere Davies Reef site moves into winter. The data for Racha Island shows a peak, reflecting the local summer. Note that in 2009 wide spread bleaching was observed at Racha Island with water temperatures reaching over 33 °C (pers. observation, authors M.J. and K.J.).

Discussion

It has been proven to be relatively easy to build a simple services based data management system around the Pachube infrastructure. The Pachube infrastructure provides out of the box functionality for data ingestion, date storage, data retrieval including graphical products, simple event detection and integration into other service based systems. What it does poorly is provide metadata including SensorML records, metadata records and associated ancillary data such as calibration data and so on. It is also a commercial system and so while the costs are reasonable this can be an impediment for uptake.

In terms of data integration the only main issue was with time zones. Pachube now has full time zone support so with some care it was possible to preserve the time zone information and to be able to compare the two sites using Universal Time Coordination (UTC)as a common time zone or using local time. This does require careful testing as many programs and databases will remove or alter time zone information and so support for sites with varying time zones needs to be tested in an end to end fashion.

The major challenge was in implementing the registry functionality using the machine tag functionality provided by Pachube along with the limited in-build registry functionality provided by the Pachube API. Machine tags are triple-tags in the form: "geo:long=50.123456" where 'geo' works like a namespace or domain name indicating, in this example, that the tag is a geographic coordinate. The 'long' second part indicates a parameter called 'long' within the 'geo' namespace while the '50.123456' is the value for that parameter. In this case the triple tag shows a value for longitude of 50.123456 in the geographic domain or namespace.

To be able to use the machine tags effectively there is a need to define these much as namespaces are defined in XML documents. This implies some level of standardisation or definition of that namespace and what elements are allowed and what content can be associated with each element. While any group can do this, to be meaningful, it has to be done at a level that reflects the potential users of the system. There is little point in the coral reef sensor network community developing standards for tags when the data are fundamentally the same as that generated by the larger marine and environmental observing communities. The lack of bodies to drive issues such as standard nomenclatures, keywords and so on is an issue and one that working within the OGC SWE specifications may have minimised given the size of the community engaged with these.

Conclusion

The work undertaken by the CREON group shows that advances in computing, in particular with cloud computing and services based architectures, provides new ways of integrating disparate information into systems that deliver via outlets such as social networking. The future vision is that individuals can create rules about events that are of interest to them, from a wide range of near real time data (including social data, weather data and crowd-sourced data), that deliver, from potentially millions of observations, a small number of knowledge outcomes.

The increasing need for evidence based management and planning requires a new range of data and information along with mechanisms to deliver this as a series of distilled, user defined, knowledge outcomes. The project described potentially represents one step on this path for coral reef systems.

Acknowledgements

We thank Jim Hendee from NOAA for comments on previous versions of this manuscript. The sensor network component of the Integrated Marine Observing System (IMOS) is funded by the Australian Government through the National Collaborative Research Infrastructure Strategy and the Queensland State

Government. The Racha Island work is funded through the Centre of Excellence for Ecoinformatics, Walailak University, Thailand, and the National Electronics and Computer Technology Centre (NECTEC), Thailand. All product/company names mentioned are registered trademarks of their respective owners or their associated companies. The support of Pachube is gratefully acknowledged.

References

52 Degrees North (2011) Sensor Web Community-Sensor Observation Service. http://52north.org/communities/sensorweb/sos/index.html, accessed May 2011

Bainbridge S (2010) The design of a novel time series client for ocean data. Proc Oceans'10 IEEE Conf, Sydney, May 2010

Botts M (2007) OpenGIS Sensor Model Language (SensorML) Implementation Specification. OGC 07-000, Open Geospatial Consortium, 2007

Botts M, Percivall G, Reed C, Davidson J (2007) OGC Sensor Web Enablement: Overview and High Level Architecture. OGC White Paper: OGC 07-165, Open Geospatial Consortium, 2007

CREON (2012) The Coral Reef Environmental Observatory Network, http://www.coralreefeon.org/

Davis (2011) Davis Vantage Pro II Weather Station. http://shop.davisnet.com.au/index.php?main_page=product_info&cPath=1&products_id=1, accessed July 2011

EEML (2011) Extended Environments Markup Language – EEML. http://www.eeml.org/, accessed May 2011

Fountain T, Tilak S, Hubbard P, Shin P, Freudinger L (2009a) The Open Source DataTurbine Initiative: Streaming data middleware for environmental observing systems. Proc 33rd Int Remote Sensing of Environment Symp, Stresa, Italy, 4-8 May 2009

Fountain T, Tilak S, Shin P, Holbrook SJ, Schmitt RJ, Brooks A, Washburn L, Salazar D (2009b) Digital Moorea cyberinfrastructure for coral reef monitoring. Proc ISSNIP 2009 243-248. DOI 10.1109/ISSNIP.2009.5416773

GeoNetwork opensource (2011) GeoNetwork opensource. http://geonetwork -opensource.org/, accessed May 2011

Glynn PW (1991) Coral reef bleaching in the 1980s and possible connections with global warning. Trends Ecol Evol 6:175-179

Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. Mar Freshw Res 50:839-866

Hoegh-Guldberg O, Smith GJ (1989) The effect of sudden changes in temperature, light and salinity on the population density and export of zooxanthellae from the reef corals *Stylophora pistillata* Esper and *Seriatopora hystrix* Dana. JEMBE 129:279-303

Hoegh-Guldberg O, Fine M (2004) Low temperatures cause coral bleaching. Coral Reefs 23:444

Hughes TP, Baird AH, Bellwoood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J, Lough JM, Marshall P, Nystrom M, Palumbi SR, Pandolfi JM, Rosen B, Roughgarden J (2003) Climate change, human impacts and resilience of coral reefs. Science 301:929-933

International Standards Organisation (2003) ISO 19115:2003 Geographic information – Metadata. ISO 19115:2003(E), 2003

Jones R (2005) The ecotoxicological effects of photosynthesis II herbicides on corals. Mar Pollut Bull 51:495-506

Jones RJ, Hoegh-Guldberg O (1999) Effects of cyanide on coral photosynthesis: implications for identifying the cause of coral bleaching and for assessing the environmental effects of cyanide fishing. MEPS 177:83-91

Kepler Project, "Real-time Environment for Analytical Processing (REAP)," https://kepler-project.org/users/projects-using-kepler-1/reap-project, accessed May 2011

Kerswell AP, Jones RJ (2003) Effects of hypo-osmosis on the coral *Stylophora pistillata*: nature and cause of 'low salinity bleaching'. MEPS 253:145-154

Leggat W, Ainsworth TD, Dove S, Hoegh-Guldberg O (2006) Aerial exposure influences bleaching patterns. Coral Reefs 25:452

Lesser MP, Stochaj WR, Tapley DW, Shick JM (1990) Bleaching in coral reef anthozoans: effects of irradiance, ultraviolet radiation and temperature on the activities of protective enzymes against oxygen radicals. Coral Reefs 8:225-232

LI-COR (2011) Underwater PAR Measurements. http://www.licor.com/env/products/light/underwater.jsp, accessed July 2011

Long Term Ecological Research – MCR (2011) Moorea Coral Reef LTER – MCR, http://www.lternet.edu/sites/mcr/, accessed July 2011

MathWorks (2011) MatLab – The Language of Technical Computing. http://www.mathworks.com/products/matlab/?s_cid=global_nav, accessed July 2011

Measurement Engineering Australia (2011) MEA Thermistor String. http://mea.com.au/products/tempstrat-thermstring, accessed July 2011

Onset (2011) HOBO Pendant Temperature and Light Data Logger. http://www.onsetcomp.com/products/data-loggers/ua-002-64, accessed July 2011

Open Geospatial Consortium (2005) Web Feature Service Implementation Specification, OGC Implementation Specification. OGC 04-094 V1.1.0, Vretanos PA (ed), May 2005

Open Geospatial Consortium (2007) Sensor Observation Service. OGC Implementation Specification OGC 06-009r6 V1.0, Na A, Priest M (eds), October 2007

Pachube (2011) Pachube – data infrastructure for the Internet of Things. http://www.pachube.com/, accessed May 2011

SeaBird Electronics (2011a) SeaBird SBE-39-IM Inductive Modem. http://www.seabird.com/products/spec_sheets/39imdata.htm, accessed July 2011.

SeaBird Electronics (2011b) SeaBird SBE-37-IMP Inductive Modem and Pump. http://www.seabird.com/products/spec_sheets/37impdata.htm, accessed July 2011

Vaisala (2011) Vaisala Weather Transmitter WXT520. http://www.vaisala.com/en/meteorology/products/weatherinstruments/multiweathersensors/Pages/WXT520.aspx, accessed July 2011

Wikipedia (2011) Photosynthetically active radiation. http://en.wikipedia.org/wiki/Photosynthetically_active_radiation, accessed July 2011

Wikipedia (2011) Tag–metadata. http://en.wikipedia.org/wiki/Tag_(metadata), accessed July 2011.

Wikipedia (2011) XML Namespace. http://en.wikipedia.org/wiki/XML_namespace, accessed July 2011