

NOAA IDEA IT Study Tsunami Use-Case Scoping Guidelines

Vision

Collaborating with key stakeholders (identified herein), develop a system that will archive, QC and provide access to tsunami focused global sea level station data, especially near real time, *event* based data that originate from NOAA sensors (such as DART buoys), but is extensible to international agencies that support the same tsunami warning function. Revise and extend current schemas to conform to new standards (such as ISO 19115) and include new data elements and application requirements. This new schema will build on existing schemas for depicting sea level stations (such as NOAA CO-OP, PTWC). Further, a key feature of the system will be to archive the metadata itself.

Project background

In 2005, the National Science and Technology Council (NSTC) and the Sub-Committee for Disaster Reduction released a report outlining the President's strategy for reducing the tsunami risk (NSTC, December 2005). The NSTC is the principal means for the President to coordinate science and technology policy across the Federal government. The National Tsunami Hazard Mitigation Program (NTHMP) is identified as providing the organizational framework for implementing the President's strategy. The NTHMP is a partnership of U.S. coastal states, territories, and federal agencies, including the National Oceanic and Atmospheric Administration (NOAA) and the U.S. Geological Survey (USGS). To support the national strategy for minimizing the impact of tsunamis as outlined in *A Framework for Action*, NOAA relies on a network of global data, acquired and processed in real-time, as well as high-quality global databases supporting advanced scientific modeling. NOAA has upgraded the majority of its sea level stations for near shore monitoring and has expanded the network of Deep-ocean Assessment and Reporting of Tsunami (DART™) stations in the Atlantic, Caribbean, Gulf of Mexico, and Pacific regions as part of the Global Earth Observation System of Systems (GEOSS).

NOAA's National Geophysical Data Center (NGDC) operates the World Data Center for Solid Earth Geophysics (WDC-SEG, including tsunamis). The WDC/NGDC provides the long-term archive, data management, and public access to global tsunami data supporting research and forecast and mitigation of tsunami hazards. Archive responsibilities include the global historic tsunami event and run-up database, hazard photographic archive as well as other related hazards, coastal digital elevation models and underlying bathymetric data and information. In addition, the WDC-SEG/NGDC are now archiving and providing access to NOAA's high-resolution pressure and temperature data from deep-ocean tsunameters and from the coastal water-level tide gauge stations. A significant gap in data management for the coastal water level data is the lack of a long-term archive. It is also very difficult to obtain water level information from various sources because there is a lack of a central portal. In an effort to preserve and improve access to

these important data for tsunami research, an effort is underway to provide access to retrospective water level data through web services, using an XML-based schema. The WDC-SEG/NGDC are committed to supporting open, standards-based access to data by we distributing data in standard formats. The deep-ocean and coastal water level data would be no exception.

Another tributary effort is led by agencies united under the Pacific Region Data Integrated Data Enterprise (PRIDE) program by the National Oceanic and Atmospheric Administration (NOAA) to develop a plain language semantic description of a sea-level station as applicable to tsunami hazards. The Tsunami of 26 December 2004 exposed a grave need for sea-level station data to be at the fingertips of warning center personnel and disaster managers. Specifically, in the critical moments after the earthquake was reported, it was unclear to officials if there were any sea-level stations located nearby operating in real-time and, if so, whom to contact to obtain the water level records. In post-disaster hindsight, it became obvious that information about sea-level stations that could be used to support detection and warning (such as location, collection and transmission capabilities, operator identification) are insufficiently known or simply inaccessible. This semantic description was later formalized into an XML schema definition. Thus began, in the aftermath of the Indian Ocean event, a multiyear effort to develop a distributed metadata system describing sea-level stations starting with pilot activities in a regional framework and focusing on tsunami detection and warning systems being developed by various agencies. This activity has matured at the time of writing (May 2008) to whereby metadata for stations monitored by the Pacific Tsunami Warning Center's "Tide Tool" are now available for consumption by XML enabled client applications. Specifically, the service is available in two formats: KML, which can be opened with Google Earth, and XML. Through the service provided by the NOAA IDEA Center (<http://www.sealevelstations.org/>), users can visually locate sea level stations in an area of interest (such as in immediate proximity of an under sea earthquake) and simply click on station icon to access the Wave and Water Level (WWL) records at the station.

Above all, the tsunami "use-case" aspect of this project derived from ongoing projects at the NOAA IDEA Center, specifically the concept of an Enterprise Service Bus (ESB), based on Information Technology (IT) Design Study. It is also informed by the High Seas theme within the Pacific Region Integrated Climatology Information Portal (<http://www.pricip.org>).

Stakeholders

Data Providers

These are individuals and organizations that are stewards of real time and near time sea level data (at a minimum we want this set to cover all of NOAA).

- Stu Weinstein (PTWC, datasource – minimal dataset, limited to warnings)
- Paul Whitmore (ATWC, datasource)
- Rich Bouchard (NDBC, DART)
- Steve Gill (NOS CO-OP – datasource – domestic – NOAA)
- Mark Merrifield (UHSLC, datasource – research focus)
- Peter Pissierssens (IODE, datasource – international)
- Geoff Clitheroe – GNS Science, New Zealand
- May be other unique providers such as JMA, Australia, IOTWS, ODINAFRICA and others.

Research/User Community

Includes data providers who would like their data to be archived by NOAA. The group below are interested in conducting historical research, validating models etc:

- Vasily Titov, PMEL (Modeler)
- John Marra (Coastal Hazards Specialist)
- Sasha Rubinovich (Tsunami Scientist – Russia)
- Laurie Dinger (Humbolt University)
- May be other tsunami scientists, coastal hazards specialists, climate change

Risks

Besides the usual project management and technical risks we see the following risks unique to this project:

- Project is dependent on resource commitments from stakeholders identified above. We need to identify a project plan that has a buy-in from all stakeholders.
- Until recently, funding for development was a risk. And we expect this risk to persist. However, for the scope of work defined in the phased release will be defined to be within the program funding. This overall vision document is broken into several components that will be funded time to time. To ensure success, we strongly recommend an extensible architecture that implements integrated of one source with one client, first.

- Resistance to adoption of standards. We need to walk a fine line to have just the right group to champion this (within NOAA at least) as complementary capability to their core competencies (as opposed to replacing their work). This groups also needs be not too small that we are operating in a vacuum and not too big that we unable to effectively pursue a concrete implementation of the tsunami user-case.

Assumptions, Limitations

- The implementation will be limited to realizing the data archiving portion of the so called “tsunami use-case” developed for the IDEA Center IT study.
- For non-NOAA datasets we are limiting the storage to “events”. Stakeholders will determine what combination of parameters are “events”.

List of features

- Plotting
- Search
- Google Maps, Earth front end (for asynchronous retrieve capability)
- ArcIMS front end (hazards map overlay)
- Data download, access (in various formats TBD)
- provide access to long-term archive high resolution
- station metadata tracking history
- direct machine link to station (for status etc, i.e. no manual updates)

Development Phases

- Define station metadata for catalog, time-series, FGDC.
- Develop a relational database design for metadata.
- Detailed Requirements, Architecture and Implementation.
- Minimum: get data from NOAA NOS CO-OP and NDBC DART buoys.
- Ingest into NGDC (through existing services like “pipeline”).
- ETL (Extract Transform Load)/Data Access Service Layer: Search, Plots etc.
- Enable web-based clients: GIS, Google, ESRI etc; web-forms, event-based portals.
- Extend to include non-NOAA Data Sources.