COASTMAP EDS - Environmental Data Server

Operational Data Management and Dissemination

SUMMARY

RPS ASA's EDS: Environmental Data Server™ collects a wide variety of oceanographic and meteorological data that is used for marine response and crisis management (i.e. oil and chemical spill modeling, emergency response, and search & rescue planning). The EDS provides superior data sources to environmental modeling applications. EDS connects US NASA, NOAA, NAVY, IOOS and other regional datasets to operational users worldwide.

EDS provides real-time as well as historical environmental data products. It is in charge of the management, analysis, visualization, and internet-based distribution through Web services. The system collects scientific data in disparate formats and makes it available to operational users via standard web services following OGC standards.

This document presents a technical description of ASA's EDS.

ASA | History

Applied Science Associates, Inc. (ASA) is a global science and technology solutions company and has been involved in developing and utilizing marine response systems, environmental data management, and Geographic Information Systems (GIS) for more than 20 years. ASA has been contracted to work for a diverse and demanding client base developing mission critical cutting edge marine response and GIS applications.

Since 1979, ASA has been building Search & Rescue (SAR) and oil spill response software systems for oil companies and international governments. ASA's has consolidated its presence with search & rescue systems and scientific expertise throughout America, Europe, and Asia. ASA's systems combine the latest technology in search & rescue drift prediction, search planning tools, as well as integrate real-time weather and vessel track data. Many coast guards and response groups rely on ASA's technical and scientific solutions.

Contact:
RPS ASA
55 Village Square Drive | South Kingston, RI 02879 USA
Email: ASAcontact@asascience.com
Phone: +1 401 789-6224 (GMT-5)

BACKGROUND

ASA's Environmental Data Server (EDS) has been implemented to operationally disseminate near-real-time environmental data. Available met-ocean sources include both observed as well as forecasted datasets. The system is regularly used by Coast Guards worldwide in support of their Search & Rescue Missions.

The Data Server regularly collects and disseminates many different data products, both public and private. It allows remote users to simply specify the geographic region and time range of interest and the historic or forecasted data is automatically retrieved.

This on-line data service, which is available 24h/7, allows users to obtain environmental data in three simple steps:

1. The user specifies the area of interest and a time range. The Server shows a list of data providers and data products for the specified area and time.

2. When the user selects a data product, a request is sent to the EDS. The Server processes the request and generates a customized file which contains a subset of the original dataset based on the specified area and time.

3. The user receives the requested environmental data file in typically less than a minute. The file, usually a wind and/or current file, is saved locally to the user's computer. All received files are a common standard format (NetCDF) and are ready to be used.

This on-line system can be crucial in support of Emergency Response operations by gathering valuable information about the present and future met-ocean conditions in an integrated and timely manner. Multiple met-ocean data sources are regularly collected: some are publicly available while others are private and thus require a specific contract subscription. ASA can provide specific data products to address client's needs: emergency response, mission support, planning, etc.

The Environmental Data Server was initially developed for the United States Coast Guard (USCG) and is currently maintained by ASA. EDS is a crucial part of the USCG dedicated system for search & rescue planning, called the Search and Rescue Optimal Planning System (USCG-SAROPS).

SAROPS has been running operationally since 2004, implemented in all USCG centers in charge of SAR missions. SAROPS includes a drifting model prediction and has the ability to handle multiple scenarios and search object types, model pre-distress motion and hazards, and account for the affects of previous searches. Acknowledging the importance of the input data acquisition component in near-real-time, SAROPS relies on the Environmental Data Sever (EDS) to request to and receive information from real-time gridded environmental data products.

RPS ASA has been supporting the U.S Coast Guard SAROPS program since 2003.

The SAROPS system has a number of components:

- ArcGIS Desktop (COTS from ESRI)
- SAROPS/SARTools extensions (developed by Northrop Grumman and ASA)
- SIM (developed by Metron)
- Environmental Data Server (EDS) (developed by ASA)

The Environmental Data Server (EDS) provides homogenous access to meteorological and hydrodynamic data via web services. It is designed to be scalable and modular and utilizes a Service Oriented Architecture (SOA). Users (clients) access the available data by using a web service, and the web service returns data based on the request. The web service may return data to the client from data products stored in distributed servers, or a more common approach is that the catalog server retrieves data to a central server. The EDS consists of three major sub-systems governing data acquisition (the Catalog Server), data storage (the Data Store), and data request (the EDS web services).

Two versions of the EDS support USCG SAROPS operations, a version that is hosted and maintained at the USCG Operational Systems Center (OSC) in West Virginia, and a back-up server hosted by ASA in a commercial server farm. ASA also hosts an internal EDS system for development and redundancy. As the EDS has evolved, in terms of its technology and the vast amount of disparate data that it stores and distributes, it has become clear that the level of expertise, both technological and scientific, to meet the demands of the system and liaise with the federal and regional data providers is significant. This proposal suggests that it may be more cost effective for the U.S government to centralize EDS operations at ASA where ocean scientists and programmers work closely together and high reliability can be provided through production, backup and development servers. As ASA also works on multiple ocean-related missions, we believe the EDS can support other USCG missions beyond search & rescue by providing instantaneous access to global and regional metocean conditions.

Unique in the community

There are many data providers that may comply with a given standard. However, the ways in which an individual or client application accesses that data vary greatly from provider to provider. The EDS collects ocean and meteorological data from disparate sources and offers that data to consumers using consistent international standards and documented delivery formats. This allows geospatial applications beyond SAROPS to access the EDS for metocean data.

The ability to provide consistent, homogenous data output for source data provided by a wide range of sources is unique within the community. The EDS aggregates data spatially and temporally; it allows consumers of the data to use one interface to access any of the data sets within their catalog of data. This level of abstraction means that data consumers need not know anything about the many different methods required to obtain data from the source, they just need to know how to use the EDS Data Request services. Issues such as data gaps, quality control and data redundancy are automatically handled by the EDS infrastructure.

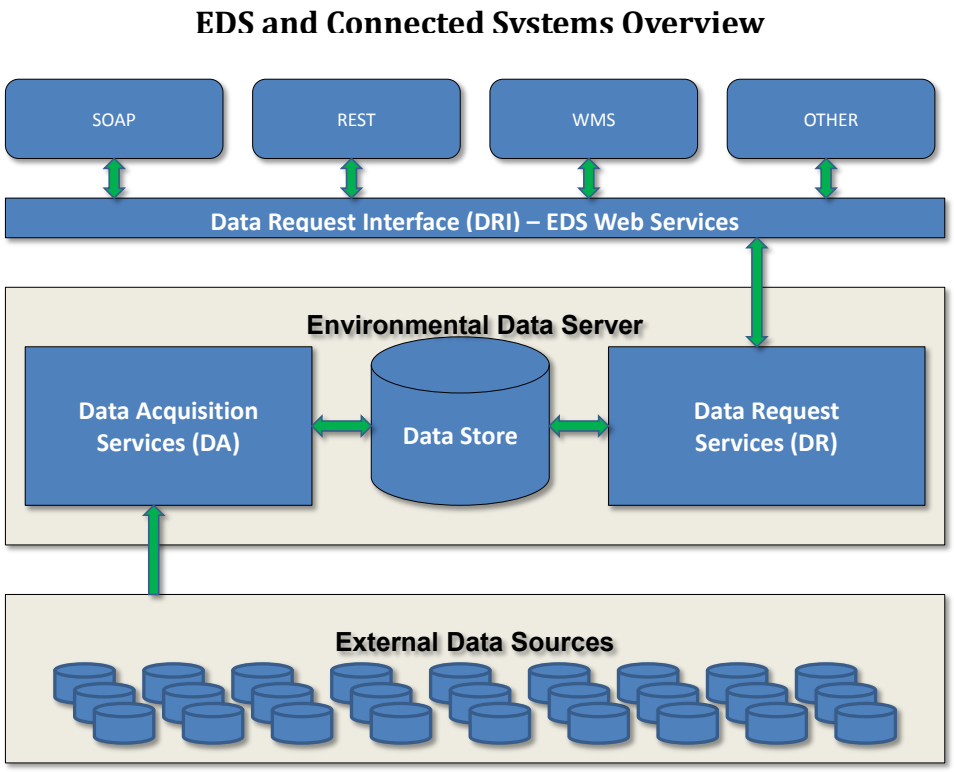

Figure 1: Systems overview

SAROPS EDS

The U.S. Coast Guard currently uses the EDS as a service supporting its SAR program. SAR Controllers may select wind and current data to feed into the SAROPS Simulator in order to generate fields of search object particles and probability grids to support SAR planning. The ASA Production EDS is currently used by the USCG in the following manner:

- The source for SAROPS wind and current data for users that are not connected to the CGDN.
- A staging platform for USCG subject matter experts to vet new sources and determine prototype guidance documentation to accompany their migration from Coastmap to the EDS instance located on the CGDN at OSC.
- ASA also maintains deep archives of the data. ASA archives most source data from the point that the new data source was added to the Coastmap EDS. The archived data is available for restoration to Coastmap Production servers with very little lead-time.
- The backup EDS service when there are issues with the OSC EDS – the USCG SAR controllers (and SAR School) can elect to use the ASA EDS Server.
- Flexibility – ASA hosts 3 instances of the EDS, Development, Staging, and Production with the ability to quickly add instances. In 2011, during Hurricane season, ASA deployed an instance of the EDS on the Amazon Cloud in anticipation of outages caused by Irene.

Cost effective

An application written to work with the EDS is guaranteed able to consume any data added to the server. This greatly reduces IT costs associated with adapting applications to the multitude of different interfaces available online.

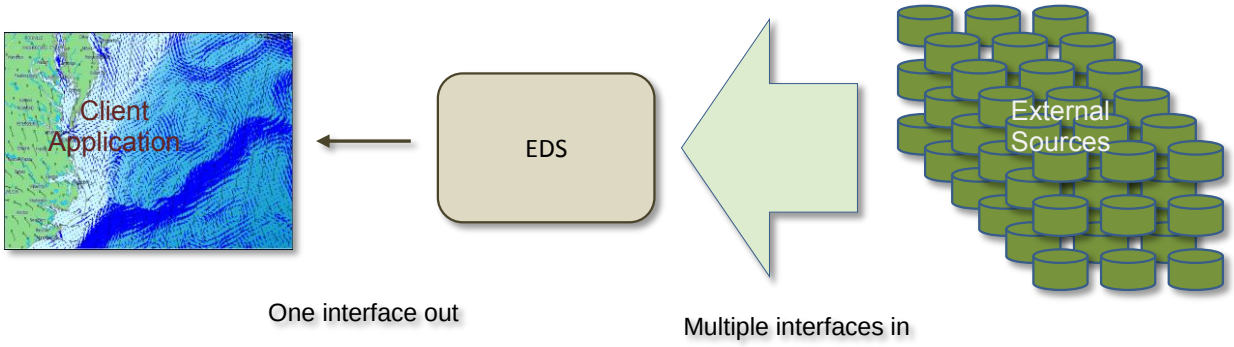

Figure 2: Interface consolidation

Since 2006, the EDS has been updated to acquire and manage over 150 different data sources, with new sources added frequently throughout the year. Over the same time period, the interface to the EDS has changed once, and the old interface has been preserved for the few clients that still require it. ASA archives the data it collects to network-accessible media that allows fast restoration of data when required and allows clients to avoid costly data backup considerations.

By using open standards, virtual machines, having extensive domain experience in managing metocean data and supporting many international marine agencies, ASA can provide a cost-effective solution that meets the needs of the U.S Coast Guard.

Availability

The EDS is hosted within a secure, monitored facility with backup power on servers configured to DHS and DoD hardening guidelines. ASA has additional servers available in Spain and Australia for our commercial clients, at our home office in RI, as well as on the Amazon EC2 cloud.

The EDS has been available 24/7 since its introduction, and the Production servers have always had a redundant backup available. Average yearly uptime percentage has been in excess of 99%.

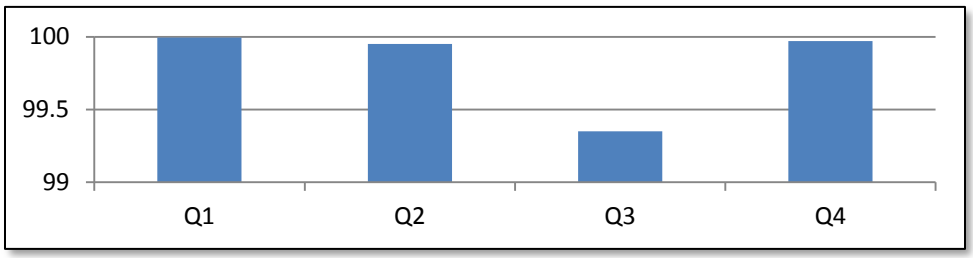

Figure3: 2011 EDS uptime percentage

Production environment downtime is normally less than 24 hours per year for maintenance purposes, and coverage is always available through the use of the alternative servers. Planned additions to infrastructure this year will allow for expanded EDS support, a disaster recovery site that will be collocated in the central or western US and improvements in our Production virtual environment in RI.

Support

ASA has a team dedicated to the support and improvement of the EDS. Developers and architects have an extensive knowledge of the EDS workflow and the data that it hosts. EDS support personnel include oceanographers and other subject matter experts who are available to assist in the vetting of new data sources, quality assurance of integrated sources and on-going checks of all data acquired by the system.

EDS support personnel leverage COTS tools, custom tools and dashboards to aid our systems administrators in catching issues as they occur. Support personnel are most-active during normal business hours. Staff are available for any emergency after-hours support on a call-in basis. Most data source – related issues are flagged and usually resolved without clients seeing interruption in availability for that source.

ASA provides standard working hours support in the U.S. as well as emergency call details. ASA also has a fully operational EDS in Australia with support staff that allows for support to SAR users outside the continental U.S and provides on-line support during non-U.S. working hours (the EDS is used operationally by SAR agencies in New Zealand and Australia).

Security

EDS development has been conducted using DHS security guidelines, with all servers complying with DHS server hardening guidelines. The Production servers are located within a locked rack within a secure collocation facility that is monitored 24/7. The Staging servers are located within the secure ASA server room – ASA holds facility security clearance at its offices in South Kingstown.

Architecture

The EDS design is based on a Service-Oriented Architecture (SOA) that provides Simple Object Access Protocol (SOAP), Representational State Transfer (REST) and Web Map Service (WMS) interfaces for client applications. The core of the EDS is made-up of multiple services which provide for a highly-scalable system. Processing is shared among several services that communicate with one another through the use of message queues and may be spread across many hardware or virtual servers. Workflow and configuration are data-driven, and the array-of-services design allows EDS architects to plug-in new functionality with little effort and little or no down time.

The EDS is made-up of three primary systems: Data Acquisition services, Data Request services and Data Store services. Clients of the ASA Environmental Data Server would be concerned only with the Data Request services, though they may be setup to receive data reports from the Data Store services.

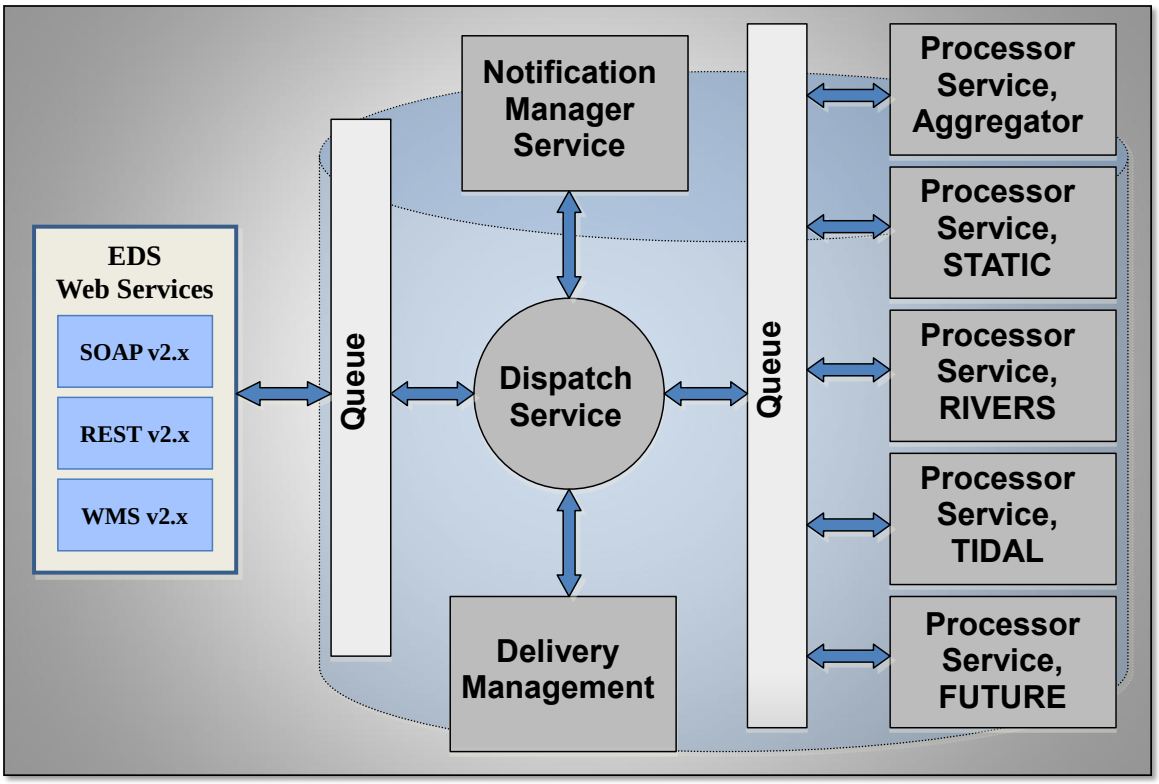

Figure 4: Data Request service array (high-level)

All requests for EDS data are stored within a database and tracked through their life cycles, allowing for reporting on order statistics and process metrics. Requests for data, regardless of source, are fulfilled through the creation of a CF-compliant NetCDF document that may be compressed for delivery-to or pickup-by the client application. Reports per client are available that would list all orders, their geographic coverage, data source and time range.

The Data Acquisition services also log all orders for external data in a database, and orders are tracked through their life cycle to a level of detail that allows ASA support staff to quickly determine when there may be a problem with a given external data source and, often, how to fix it. The Data Store services include more client-level reporting in the form of reports customized per client to show what data is in-house and when it was last downloaded.

In addition to the SOA services, there is an OGC-compliant Web Map Service (WMS) interface available for use by the U.S. Coast Guard for EDS data. This interface allows client applications to collect WMS capabilities for all available data sources, generate source overlay images for mapping applications and pull time series information for individual sources. This interface allows access to wind and current data as well as any other non-vector attribute data that may be specified as available through the EDS (see appendix A) for a given client. All EDS installations include a sample mapping application that takes advantage of this interface. There are non-OGC compliant layers available as well that are used to show SAROPS particle and probability grid outputs as well as oil spill model results.

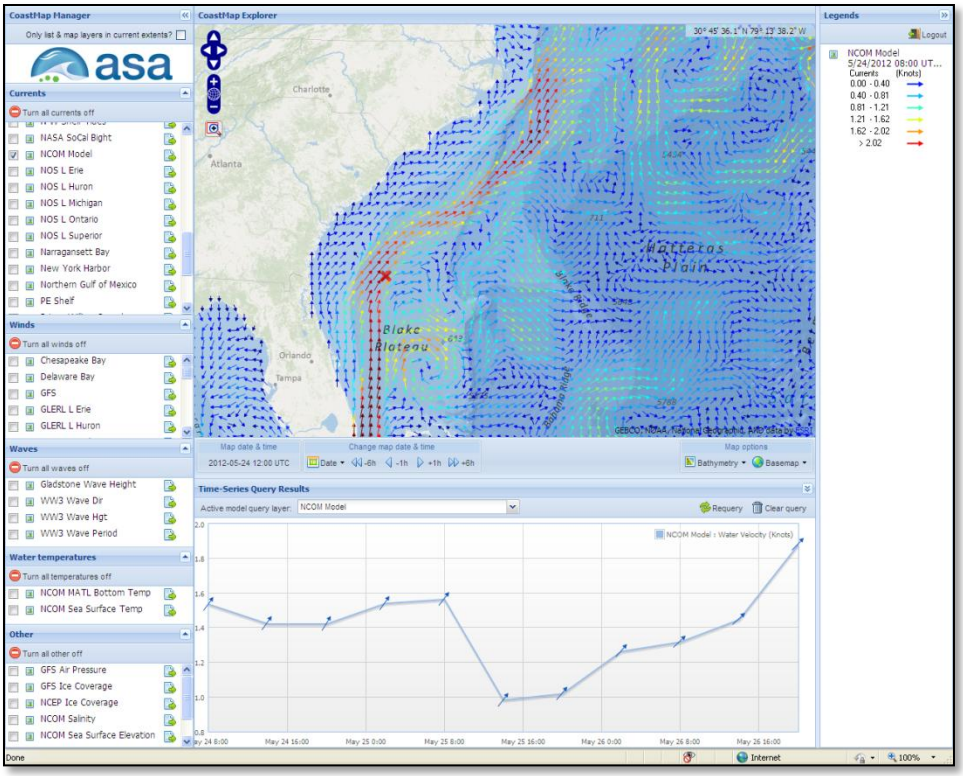

Figure 5: Sample mapping application provided with Coastmap EDS

The service array architecture of the EDS – the use of multiple, atomic services, fast messaging and data-driven configuration – provides for a highly-scalable system. Each service, in turn, handles multiple processes at once through the use of multithreading. This architecture distributes processing among many services, and these services may be started on any number of servers, effectively spreading the EDS installation's workload across hardware or virtual assets. The use of service arrays allows for the modification of existing services or the addition of new services without the need to bring the entire system down.

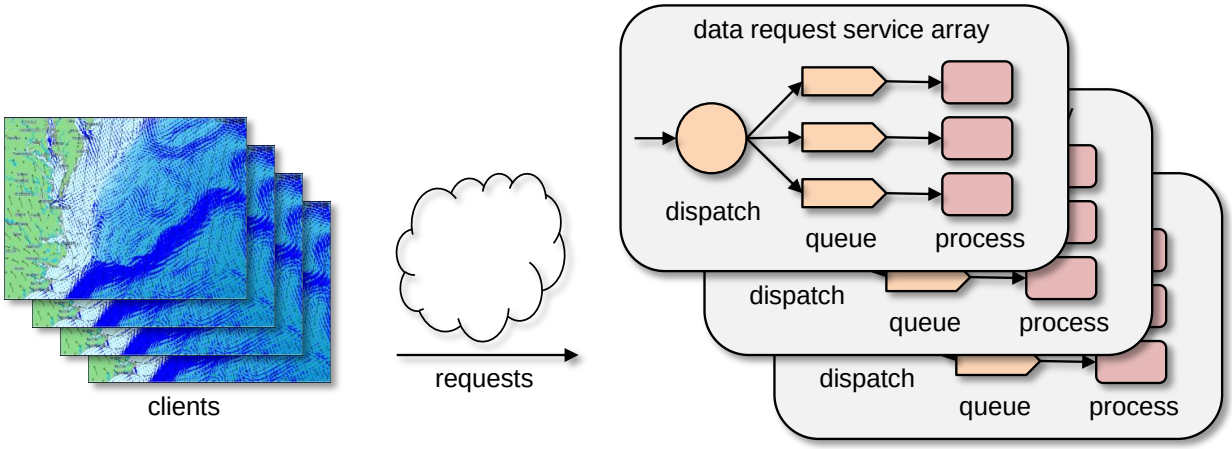

Figure 6: Multiple service arrays serving many client requests

Monitoring, quality, notification

Quality concerns are applied on many levels:

Data source integration. With the benefit of having integrated so many data sources, the integration process for new sources has evolved into a systematic approach. After technical-level negotiations with the source provider are complete and the source has been added as another to be acquired regularly by the EDS, a subject matter expert (SME) is brought on-board to evaluate the results. If a service that already handles one or more data sources is modified in any way to accommodate the new source, QA performs full regression testing in addition to SME reassessments of all potentially-affected sources.

Routine source acquisition. Once the source has been integrated, the process of data acquisition is monitored by EDS support personnel through the use of source audit trails, process notifications, an automated test harness, temporal checks and nightly reports.

Periodic source quality checks. Checks on all source output are performed periodically by ASA SMEs in order to ensure that all sources are still working properly and in order to identify points of improvement for subsequent EDS builds.

Automated source vetting. Certain sources, specified by records within the database as sources of interest, are periodically queried and compared to observation data available within source bounds. Presently at an early stage of development, the information gleaned from this quality analysis will be made available to foster greater awareness of each source's strengths or weaknesses within given areas.

Service-level client control, catalogs

The EDS is designed to provide catalogs customized to individual clients' needs. Through the use of 'client keys', access is limited to the data sources of interest to the client. Catalogs of data may be built to include global and regional data sets with winds, currents and a variety of scalar attributes. Please refer to the source list at the end of this document.

All EDS installations make use of client keys in order to create catalogs, provide capability extents and track Data Request orders. SAROPS users currently have one client key. It is recommended that there be a shift to a key-per-district (or other delineation) scheme – this would allow tracking of EDS usage per district.

Use of client keys allows the Coastmap EDS to treat all of the USCG Enterprise as one client group while still allowing treatment of USCG partners on a case-by-case subscription basis.

Change Governance

EDS service modification has been governed by a process of issue identification and prioritization, development life-cycle progression, change request review, application and promotion. Issues requiring immediate fixes and deployment typically shortcut only the first step of this process but proceed as normal through the rest of the workflow. With Operations relying 100% on the Coastmap EDS, it would be advisable to have a Change Control Board (CCB) available for prioritization at the front-end and

review/approval for promotion for any patches or new versions. This CCB should include at least one USCG representative.

The process of adding source data to be used by the U.S. Coast Guard is documented. The identification and ultimate vetting of new source data candidates are governed by a USCG representative. Retirement of any data source would also rely on USCG representative guidance.

With the addition of any new data source, RPS ASA collects data about that source, including contact data, to include in its own documentation for systems analysts' reference. This metadata is also forwarded to the USCG data representative for his use in the creation of training aids.

Concurrent access

Since inception in 2006, concurrent access (load testing) scenarios have been run against the Coastmap EDS. The EDS Test Harness application in use by Systems Administrators (SA) has a load test function that allows them to push multiple orders. Tests show success in scenarios where sets of 50 and 100 orders are generated simultaneously (for example: sources=GFS, NCOM; AOI=2degree square; Time range=72hrs; Total time for order completion=24secs for all 50 together).

The EDS is not likely to be used to capacity in the near future, although we recognize that the USCG SAR School may require many concurrent users and the EDS can manage that load.

If metrics reports show that the EDS is being taxed beyond an optimal level, ASA would bring additional application server virtual machines online to handle any excess workloads. The scalable nature of the EDS architecture, as mentioned above, allows this to be done with little effort (figure 6).

Sufficient access, coupled with high-availability will allow the EDS to be used by the U.S. Coast Guard for additional missions as well, including oil spill and emergency planning.

Community Cooperation

Through ASA's extensive relationships with related agencies, including NOAA and Navy and oceanographic institutions, the ASA scientists have access to the latest available data and model forecasts. A common understanding between these groups allows ASA to quickly remedy any issues with data providers and add data products as needed. As an example, during a subsea search (downed helicopter), ASA was able to integrate new sea-bottom currents into EDS to provide support to the search in a matter of hours.

Mission Support

The focus of EDS has been SAR operations support, but the system has also been used by USCG Ice Patrol to evaluate new data products. It is also used by other clients for oil and chemical spill response. The USCG can access the EDS to support other missions and new data is being added to the systems such as ice cover, atmospheric conditions, wave conditions, and in-water column conditions to support an even broader set of operational activities.