

USFWS National Wetlands Inventory Project Synopsis

Title

Wetlands Mapper

Description

The U.S. Fish and Wildlife Service (Service) is the principal Federal agency providing information to the public and other agencies on the extent and status of the Nation's wetlands. The Service provides stewardship for the wetlands data that comprise the Wetlands Layer of the **National Spatial Data Infrastructure** and makes these data available via the Internet

The U.S. Fish and Wildlife Service's strategic plan for digital wetland data is focused on the development, revision, and dissemination of wetlands data and information to Service resource managers and the public. The present goal of the Service is to provide the citizens of the United States and its Trust Territories with current geospatially referenced information on the status, extent, characteristics and functions of wetlands, riparian, deepwater and related aquatic habitats in priority areas to promote the understanding and conservation of these resources.

Updates to the Wetlands Mapper is in direct response to the need to integrate digital map data with other resource information to produce timely and relevant management and decision support tools.

This new version of the Wetlands Mapper is designed to promote greater awareness of wetlands geospatial data applications and to deliver easy-to-use, map like views of America's wetland resources in a digital format. It has been developed in collaboration with the U.S. Geological Survey (USGS). This Federal partnership has yielded tremendous benefits in ongoing efforts to configure, improve and distribute the wetlands map information using newer technologies in digital mapping and web-serving capabilities.

Currently, the web site is serving wetlands data that cover over 57,000 quadrangles in a seamless ESRI File Geodatabase. This represents wetland map data for approximately 68 percent of the conterminous U.S.

The updated Wetlands Mapper is designed to provide reliable geographical information and direct users to other potential sources of information. It is not intended to provide legal or regulatory products. Unintended use of the information or products is discouraged. Persons intending to engage in activities involving modifications within or adjacent to wetland areas should seek the advice of appropriate federal, state or local agencies concerning specified agency regulatory programs and proprietary jurisdictions that may affect such activities.

Contact person

Andrew Cruz
U.S. Fish and Wildlife Service (NSST)
505 Science Dr
Madison, WI 53711
608-238-9333 extension 114
Andrew_Cruz@fws.gov

Operating System

Microsoft Windows Server 2008

Application Platform: (i.e. java, ASP, C#)

Adobe Flex viewer; ActionScript programming language.

Software to be included

The Wetlands Mapper runs on ESRI ArcServer and IIS.

Database SW requirements

None. This is a file based geodatabase.

Bandwidth requirements

The current mapper is served from a facility that has a potential 18MB per second. This connection is shared with other USGS applications. Last quarter use statistics:

Mapper user requests 9,047,723

WMS user requests 2,713,735

KMZ user requests 252, 228

Data size

The data sets combined take approximately 12.6GB of data space. This number will increase with each new update but it will probably never exceed 20GB. There are approximately 16 million polygons in the geodatabase.

US Census Bureau TIGER/Line Project Synopsis

Title

Census Bureau TIGER/Line® Shapefiles Data Server

Description

TIGER/Line® Shapefiles are spatial extracts from the Census Bureau's MAF/TIGER database, containing features such as roads, railroads, rivers, as well as legal and statistical geographic areas. Currently the Census Bureau offers the files to the public for download from a Census Bureau data server. The files are accessible from the download web site three ways:

via anonymous ftp: <ftp://ftp2.census.gov/geo/tiger/TIGER2009/>

via http (direct access): <http://www2.census.gov/geo/tiger/TIGER2009/>

via http (CGI application): <http://www2.census.gov/cgi-bin/shapefiles2009/national-files>

We would like to utilize the GeoCloud Sandbox Initiative to evaluate the performance and cost of hosting TIGER/Line data in a cloud computing environment per the Geospatial solutions #1 - Data Case. We would provide the 2009 TIGER/Line® Shapefiles in a compressed format.

Contact person

Lynda Liptrap
U.S. Census Bureau
4600 Silver Hill Road, Geo 7400
Washington, D.C. 20233
301-763-1058
Lynda.a.Liptrap@census.gov

Operating System

N/A

Application Platform: (i.e. java, ASP, C#)

N/A

Software to be included

N/A

Database SW requirements

N/A

Bandwidth requirements

500 Mb/s

Data size

56 gigabytes (compressed)

NOAA IOOS Project Synopsis**Title**

Integrated Ocean Observing System (IOOS) Registry, Catalog and Viewer

Description

We are considering GeoCloud hosting of the following components:

- 1) A registry of IOOS-related data servers, a Catalog of metadata (not data) from those servers describing data holdings, geographic and temporal coverage, and other essentially service metadata
- 2) A harvester that periodically polls the servers to refresh the metadata catalog
- 3) A viewer interface that allows map-based and forms-based queries against the catalog and displays results on a map or in textual format.

The Catalog is intended to be fine-grained (i.e., to record data availability at the level of individual observing platforms, map layers or grid coverages, not at the more coarse service level). Software will be based on existing implementations of similar components; evaluation of components is now in progress.

Contact person

Jeff de La Beaujardiere
NOAA Integrated Ocean Observing System Program
1100 Wayne Ave., Suite 1225
Silver Spring, MD 20910
Phone: (301) 427-2420
Fax: (301) 427-2073
Jeff.deLaBeaujardiere@noaa.gov

Operating System

Probably Linux

Application Platform: (i.e. java, ASP, C#)

TBD (probably Java, Javascript, Perl)

Software to be included

TBD (probably Apache, PostGIS/PostgreSQL, and govt or open-source sw)

Database SW requirements

Probably PostGIS/PostgreSQL

Bandwidth requirements

TBD

Data size

Modest (metadata catalog, not full data holdings)

NOAA ERDDAP Project Synopsis**Title**

Environmental Research Division's Data Access Program (ERDDAP)

Description

ERDDAP (<http://coastwatch.pfel.noaa.gov/erddap/index.html>) is a RESTful web service that can serve both local data as well as read from remote web services, and return the data in a variety of formats. ERDDAP also provides stand-alone OPeNDAP, WMS and IOOS SOS services, as well as various pub/sub capabilities. The project would be to have a cloud ERDDAP service, as well as move selected of our most used datasets (see either <http://coastwatch.pfeg.noaa.gov/erddap/info/index.html> or <http://oceanwatch.pfeg.noaa.gov/thredds/catalog.html> for a list of possible datasets to be hosted in the cloud).

Contact person

Roy Mendelssohn
3333 North Torrey Pines Court
La Jolla, CA 92037-1022
(831) 648-9029
Roy.Mendelssohn@noaa.gov

Operating System

Linux

Application Platform: (i.e. java, ASP, C#)

Java

Software to be included

Tomcat, Thredds Database

Database SW requirements

N/A

Bandwidth requirements

TBD

Data size

TBD

USGS NED Project Synopsis

Title

National Elevation Data Set Delivery

Description

This is a set of applications and services to deliver map services, and data downloads of the complete National Elevation Data Set in support of The National Map. This entails a complete copy the the data, applications, and services.

Contact person

Sara Jean Paulson
U.S. Geological Survey
Earth Resources Observation and Science (EROS)
47914 252nd Street
Sioux Falls, SD 57198-0001
(605) 549-6560
paulson@usgs.gov

Operating System

Windows (possibly multiple - depending upon configuration - Windows and Solaris is the current configuration)

Application Platform: (i.e. java, ASP, C#)

.NET, Java, PHP

Software to be included

Microsoft IIS, Apache Tomcat, Apache Tomcat ISAPI Connector, Helicon ISAPI_Rewrite, Sun MySQL, Zend PHP, Sun Java Runtime Environment, Microsoft .NET, ESRI ArcIMS, ESRI ArcGIS Server, ESRI ArcGIS Image Server, ESRI ArcSDE

Database SW requirements

SDE capable environment - We use Oracle 10g. SQL Server might be a possibility.

Bandwidth requirements

~ 172,000,000 requests per day, ~ 26,220 KB of data transfer per day

Data size

~ 2.75 TB of storage

USGS NATIONAL MAP Project Synopsis

Title

National Map Viewer

Description

This configuration will cover the infrastructure of the base map services, tile services and the implementation of the TNM Viewer configuration.

Contact person

David R. Hughes
US Geological Survey
DFC NGTOC DENVER
Lakewood, CO 80225

(303) 202-4112
drhuges@usgs.gov

Operating System

Windows, Server 2003, Solaris is the current configuration)

Application Platform: (i.e. java, ASP, C#)

ArcGIS Server, Apache, Tomcat, Java, IIS

Software to be included

Microsoft IIS, Apache Web Server, Apache Tomcat, Apache Tomcat ISAPI Connector, Oracle 11G, Sun Java Runtime Environment, Microsoft .NET, ESRI ArcGIS Server, ESRI ArcSDE, Perl

Database SW requirements

SDE capable environment, Oracle 11G. (While we list the Oracle/SDE environment requirements, we do have an option to support the viewer application without this dependency on the database hosting within the cloud. In other words, the database hosting is optional since we could direct the database requests to in-house systems.)

Bandwidth requirements

TBD

Data size

Tile Storage 1 TB, DB Storage 1 TB, Total: 2 TB

EPA LAKES AND PONDS Project Synopsis

Title

US EPA Region 1 EMAP Lakes and Ponds

Description

The Region 1 EMAP Lakes and Ponds project created a database of lake and pond sampling data collected from 2006 – 2009 as part of the National Lakes Survey. The sample and resulting analytical data have been QA/QC'd by the EPA's Narragansett Lab and various reports, maps, and models have been generated.

Currently, the published information is in a static format and the project manager wants to see a more dynamic solution to promote reuse of the data and methods. The concept for the application on the GeoCloud is to publish dynamic maps, reports, models, and data extracts from a relational database through a web page interface. The application will run on a Windows 2003 server with an Oracle or Postgre-SQL database instance. Spatial data will be published through ArcGIS Server as standards based services.

A mapping interface will be developed to graphically display sampling locations linked to results, documents, etc. The map will leverage the EPA's license agreement for Microsoft Bing Maps to provide background data and a simple SDK.

Reports and data extracts will be user defined and produced on the fly. The user will enter query parameters for the desired information and the application will produce the specified products based on the query results.

The models produced for this project utilize R-Script for statistical analysis. A potential solution for publishing the models involves a combination of Python scripts, R Scripts, and ArcGIS. The combination of these software packages will allow for models to be executed on the web by a user with no programming skills by entering parameters for the specific model.

This configuration will cover the infrastructure of the base map services, tile services and the implementation of the TNM Viewer configuration.

Contact person

Tom Giffen
US EPA Region 1
5 Post Office Square Suite 100
Mail Code OARM02-1
Boston, MA 02109-3912
617-918-1444
giffen.tom@epa.gov

Operating System

Windows 2003 Server standard edition on VMware or Open Virtual Format compatible infrastructure

Application Platform: (i.e. java, ASP, C#)

ASP.NET, Jscript, R-script

Software to be included

ArcSDE, ArcGIS Server, Model Builder

Database SW requirements

Oracle or Postgre-SQL

Bandwidth requirements

10 Mbps

Data size

10 GB

NOAA LarvaMap Project Synopsis

Title

Particles in the Cloud

Description

Particle tracking models are a good application to test both the computing aspects of the cloud and the storage and transport of large datasets needed to run the models. LarvaMap is a web-based larval transport model that can utilize Common Data Format for Climate and Forecast (NetCDF CF) available through a Thematic Realtime Environmental Distributed Data Services (THREDDS) data server (TDS). LarvaMap is a web-based service that allows users to specify starting location of particles and run models on top of circulation model outputs. Challenges include transport and access to large data sets, the need to subset data spatially and temporally, time and space aggregation of model data, and the rapid return of results for visualization. The LarvaMap service is an ideal case study for applying cloud computing resources to problems which are very difficult using traditional IT resources. This project will explore how services and data can be deployed on the cloud and will investigate and report on issues such as implementation challenges, performance, and security.

The project would install the existing LarvaMap service and its associated data on the GeoCloud. It would take advantage of a number of concurrent developments in NSF/NOAA funded infrastructure for data management in the cloud. We would install the LarvaMap service, 10 – 20TB of ROMS data from PMEL, and data management tools being developed by the NOAA/NSF funded OOI/IOOS collaborative data exchange project. LarvaMap is a prototype service, representative of those required by NOAA to make management decisions based on ocean science data. It is an example of a category of analysis which is ideally suited to operating in the cloud. The particle tracking model uses Eulerian model data which has complete coverage of a region in space and time as input, returning much smaller Lagrangian data for the location of a parcel over time. This information is both more valuable to the user while also represented by a much smaller number of bytes. It is therefore efficient to move the analysis service to input data rather than move the data to the service or the user. The computation process is computer intensive. While particle tracking is a fundamentally iterative process which requires serial execution, stochastic modeling of many particles is ideally suited to parallel, concurrent execution. These properties of the LarvaMap service make it an ideal case study for applying cloud computing resources to problems which are very difficult using traditional IT resources.

Cloud computing provides an opportunity to do more than host a single webservice using abstracted resources. It is a trivial matter to start an image of the existing service on an Amazon EC2 instance. The value of the proposed prototype comes from integration with other infrastructure efforts lead by OOI CI and OPeNDAP. These groups hope to provide the critical infrastructure which will deliver on the promise of scalable cloud computing and data caching. Early integration with these new technologies will provide critical insight to NOAA's future development, helping align with other major efforts in ocean science data management.

As part of a collaborative effort between the OOI CI (NSF-funded) and IOOS (NOAA funded) a prototype system for distributed data management using message based services in the cloud is being developed. The Data Exchange prototype will be one of the first products of the OOI CI effort and provides an opportunity for early integration with the next generation of ocean science data management. The prototype will provide tools to collect datasets in a cache on the cloud. Analysis services such as LarvaMap can then be integrated with the system to deliver valuable new information to the user. While the Data Exchange prototype is the most advanced system to target for integration, other more straightforward options exist which reduce the risk of project failure. The Data Exchange prototype is using the Hyrax data server from OPeNDAP as the cache request agent. The Hyrax server itself has a distributed architecture, providing a platform in which the LarvaMap service could be integrated as an alternative. There are many options to investigate in this exploratory work.

Contact person

Tiffany Vance
NMFS Alaska Fisheries Science Center
Seattle, WA
tiffany.c.vance@noaa.gov

Scott Cross
NESDIS National Coastal Data Development Center
Stennis, MS / Charleston, SC
Scott.Cross@noaa.gov

Operating System

TBD

Application Platform: (i.e. java, ASP, C#)

TBD

Software to be included

Hyrax data server, NOAA/NSF data management tools

Database SW requirements

TBD

Bandwidth requirements

TBD

Data size

The data sets combined take approximately 10 – 20TB of data space.
