An interoperable spatiotemporal weather radar data dissemination system

Y. Cao, C. Yang*, D. W. Wong

Joint Center for Intelligent Spatial Computing, and
Earth Systems and GeoInformation Sciences
College of Science, George Mason University, Fairfax, VA, 22030, USA

* Corresponding author. Phone: 703-993-4742, Email: cyang3@gmu.edu
Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)

Abstract

This paper investigates an interoperable framework to disseminate Earth Science data to different application domains. The proposed framework can manage different Earth science data products and raster snapshots over time through the use of relevant metadata information. The framework generates images to be accessed by GIS software for various Earth science and web-based applications. The access is enabled through the compliance with OpenGeospatial Consortium’s Web Map Service (WMS) for interoperability such that any WMS viewer can access the service. The framework can provide GIS users the capability to incorporate geospatial information from other WMS servers. Using the United States NEXt generation weather RADar (NEXRAD) data, we demonstrate how the proposed framework can facilitate the dissemination of Earth Science data to a broad community in a near real-time fashion. The proposed framework can be used to manage and disseminate various types of spatiotemporal Earth science data.

Keywords: Interoperability, Earth Science, NEXRAD, Near Real-time, Web Map Service (WMS)
1. Introduction

The volume of Earth observation data has been increasing exponentially with remote and in-situ sensors constantly monitoring the Earth. Numerous users depend on these data to develop various day-to-day applications. For example, the National Aeronautics and Space Administration (NASA) Distributed Active Archive Centers (DAAC) and the Earth Science Information Partners (ESIP) distributed over 2 terabytes of data to over 2 million users during 2002 (Ramapriyan, 2003). Along with the flourishing of Earth science data has been the development of a large number of complex, but to some extent, similar algorithms and/or datasets tailored to specific applications (Votava et al. 2002). While the objectives of those applications may not be the same, researchers and scientists have duplicated their efforts in developing these similar or overlapping algorithms and/or datasets without leveraging each other’s accomplishments. Therefore, a current challenge is to reduce these possibly duplicated efforts and to utilize what are already developed and available in other applications efficiently.

To achieve an integrated environment with reduced algorithmic and data redundancies, the use of remote sensing and Geographic Information System (GIS) technologies can be beneficial in studies involving geospatial information in various domain applications: land use and land cover changes (Ö’Hara et al. 2003; Yang and Liu, 2005; Saraf and Choudhury, 1998), ground water analysis (Tcherepanov et al. 2005; Pankaj et al. 2006; Jaiswal et al. 2003), and coastal management (Host et al. 2005). These environmental applications successfully reduced redundancy by utilizing a case-by-case interface. However, redundancy is still a problem when Earth science data are utilized in multi-disciplinary research. For example, it took tremendous efforts to integrate Global Position System (GPS) data with remote sensing data to provide information through viewers like Google Earth™ and Virtual Earth™ to evacuated residents in the aftermath of Hurricane Katrina (Womble et al. 2006). Besides ordinary citizens who use common web-based viewers to browse data and obtain information, experts such as planners, meteorologists, and environmental engineers would prefer a
This paper introduces a framework to bridge the gap in data sharing between Earth sciences and various disciplines. Earth science data, such as satellite, radar, lightning, wind profiler and model data, as is the case here, may need to be integrated for specific domain applications. The proposed framework is based on the following three requirements:

1) Must be generic enough to be applied to different types of Earth science data.
2) The need for near real-time data access by users.
3) High level of interoperability on data operations across disciplines to reduce algorithm and data redundancy.

The next section reviews relevant frameworks and geospatial interoperable solutions, followed by an overview of the proposed framework. Section 3 describes in detail the design and implementation of the proposed framework. Section 4 presents a case study using the framework to disseminate the NEXt generation weather RADar (NEXRAD) weather data. Finally, Section 5 provides a conclusion and discusses remaining research issues.

2. General approach

2.1 Related work

Considerable progress has been made in data sharing and dissemination since NASA initiated the Earth Observing System (EOS, http://eospso.gsfc.nasa.gov/) and the Federation of Earth Science Information Partners (ESIP, http://archive.esipfed.org/index.html). However, near real-time data handling has not been sufficiently addressed. To improve the performance of Earth science data processing for near real-time applications, Votava et al. (2003) parallelized several Earth science algorithms on a cluster of workstations and adopted the Message Passing Interface (MPI, Gropp et al. 1999) wrapper as the interface between the new
Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)

framework and the legacy system. The performance issue was also addressed by Tehranian et al. (2006) for distributed processing of satellite data where reliability and real-time response are critical.

Lessons learned from NASA’s Earth Science Enterprise (ESE) experiments indicate that rather than using a single system to deal with all kinds of Earth science data, it would be more efficient to design and develop a group of systems with dataset-specific settings, implementations, and operational decision mechanisms (Ullman et al. 2003). In most cases, components or facilities are implemented according to user needs in a specific domain. For example, the Global Land Cover Facility (GLCF) provides land cover datasets through a high performance computing architecture which has three levels of storage and access hierarchy (Lindsay, et al. 1999). Therefore, the desired framework should leverage specific capabilities provided by different data systems. While each implemented component has its focus, all components should communicate or share data with each other.

Some domain specific frameworks were developed to test the feasibility of such an interoperable system environment. For example, Hofmann (1999) proposed the GIStermFramework as a three-tier framework to access distributed, heterogeneous spatial data in federated environmental information systems. GIStermFramework integrates GIS components into a service-oriented architecture and each component complies with the Open Geospatial Consortium’s (OGC) Simple Feature Specification for geometric attributes storage. The Integrated Data Viewer (IDV), a Java-based tool developed by Unidata, is widely used to analyze and visualize Geoscience data. Because IDV is intended to be a powerful visualization and analysis viewer, it can access other Web Map Servers. However, it does not provide a Web Map Service (WMS) to other compliant viewers (Beaujardiere, 2004). Thus, its interoperable communication is unidirectional. Similarly, the National Climate Data Center (NCDC) developed an interoperable viewer for the highly demanded real-time radar data, NEXRAD (NCDC, 2007). This viewer provides an OGC-compliant service, but cannot be used by other WMS viewers or systems for NEXRAD data integration.
These individual applications have limited levels of interoperability and therefore, they may not be efficient in supporting interdisciplinary studies when multiple geospatial technologies or Earth science datasets are required. To facilitate interoperability within a system, both the service and access to services should be implemented according to standard interoperable interfaces, such as the specifications proposed by OGC, and the standards proposed by the United States Federal Geographic Data Committee (FGDC) for building the National Spatial Data Infrastructure (NSDI, Clinton 1994). One of the popular interface specifications is WMS, which produces spatially referenced information, such as maps in image formats (Portable Network Graphics-PNG, Graphics Interchange Format-GIF, or Joint Photographic Experts Group - JPEG) and metadata in eXtensible Markup Language (XML). WMS supports three operations: getCapabilities, getMap, and getFeatureInfo. These operations are explained in Table 1. Besides supporting the spatial dimension, WMS also supports temporal dimensions by providing an animation service extension to deal with temporal data (LaMar, 2006). In our framework, we adopt the properties of WMS in sharing Earth science data.

2.2 A general framework

We propose a framework (Figure 1) that utilizes WMS, Web GIS, and near real-time data dissemination technologies to develop interoperable and bi-directional data sharing capabilities. The remote or in-situ GeoSensors collect data in real time and transfer the data to the Data Ingest Manager to be stored on a data server, where the Data Broker is responsible for data preprocessing and generating metadata. According to different data formats and characteristics, processing components in the Data Broker will be triggered by data receiving events (Muhl et al. 2006). Thereafter, the GIS server is responsible for providing geospatial mapping capabilities and communicating with viewers in an interoperable manner by adopting the WMS interface. The temporal WMS and animation service extension can be included in the WMS to support
various spatiotemporal operations, which are submitted or requested by users on the client side. Although this framework is relatively simple, it is highly flexible as each component can be modified or enhanced. It is also versatile because the components support general functions, but can be tailored to meet the requirements for specific applications. The framework structure is not application dependent. Below, we will elaborate on each of these components.

3. Design and implementation

The major components of the framework include Geosensors, Data Ingest Managers, Data Broker, and GIS Server (Figure 2), which are the basic or minimum set of components for the proposed framework. Other application-specific components can be added to the framework to meet specific system requirements and/or user specifications.

*Geosensors:* The term “geosensors” in this paper refers to sets of sensors distributed over different locations or covering different parts of the Earth’s surface to collect geospatial data or monitor geographical phenomena simultaneously. These sensors are connected through networks such that data assembled from these sensors can provide a broad coverage of the concerned geospatial phenomenon. There are numerous such networks in different forms. At a local scale, we may have security cameras connected by a network monitoring various parts of a commercial building or a university campus. The set of satellites in the Earth Observing System (EOS) is an extensive global geosensor network monitoring various aspects of the Earth’s systems, including the atmosphere, hydrosphere and lithosphere. In between these two extreme scales, we have numerous geosensor networks continuously gathering data that need to be disseminated to users as quickly or as close to real time as possible. The NEXRAD system, focused on in this paper, and the Advanced Very High Resolution Radiometer (AVHRR) system are additional examples.

*Data Ingest Manager:* To handle the data gathered by Geosensors efficiently and effectively, a data management subsystem or the Data Ingest Manager is needed to receive the near real-time data from the
geosensor network. Various functions can be performed by this subsystem. These functions can include, but are not limited to, handling specific data formats provided by the geosensors and managing the data stream over time or providing temporal management of the data.

**Data Broker:** When the Data Broker receives new data products from the Data Ingest Manager, it will save the data under a preconfigured file path. The data broker processes new data to be ‘GIS ready’ in real time, and archives the data on the data server for temporal queries. The introduction of ‘GIS ready’ data can benefit many users for different applications. To facilitate map overlay operations, geo-referencing and re-projection tools are used to prepare the data. When new data/files are received, an automated computer script will update the most recent file with this new one under the ‘now’ path and will also archive the previous data product with its timestamp. Most data mentioned here are raster/image data. They are snapshots of a time series (Cheng et al. 1988). For Earth science data, there is a tradeoff in saving data in the original data formats or in one of the commonly accepted image formats, such as PNG and GIF. For mapping purposes, PNG format is recommended since very little information is lost during format transformation (raster to image) and transparency can be applied on the resultant images for map overlay inside the client viewers.

**WMS Support:** To support WMS operations, metadata are stored in XML files, and a metadata file is created for every data layer (Figure 3). Since metadata remain unchanged when time series data are updated, the WMS server only needs to read the metadata once. File reading and parsing time can be saved for the server to generate responses for the getCapabilities operation when the operation is called on multiple times by different viewers. An important metadata entry is geographical projection. The Geospatial Data Abstraction Library (GDAL, 2007) can be used to transform one projection to other GIS-compliant projections.

**Client-server communication:** As illustrated in the sequence diagram (Figure 4) and detailed in the case study in Section 4, client-server communication is captured by three scenarios: real-time WMS, map
animation with temporal WMS, and spatiotemporal analysis. The viewer can request georeferenced data from a WMS server through the HyperText Transfer Protocol (HTTP). Any WMS-compliant viewer can also communicate with the server and request Earth science data products. For the remaining steps, special modules supporting specific functions send requests through HTTP, then parse and render the spatiotemporal results in the viewer.

4. Case study: NEXRAD data dissemination and analysis

4.1 Background of NEXRAD and the IDD systems

The NEXRAD system is a joint effort of the National Weather Service (NWS), the US Departments of Commerce (DOC), Defense (DOD), and Transportation (DOT) to measure weather parameters in the U.S. over approximately 160 sites using Doppler radar (OFCM, 2006). The system can operate in clear-air mode, precipitation mode, or severe weather mode depending on different weather conditions (Crum et al. 1993). The NWS has a long-term objective to utilize emerging technologies to perform the most comprehensive weather impact diagnosis. All possible means of observations, such as radar, satellite, surface, and other data can be fitted into a four-dimensional (the two spatial dimensions, plus elevation and time) grid coverage (NRC, 2002). This case study presents not only data integration and dissemination through WMS, but also a four-dimensional analysis, which can contribute to such a grid.

The NEXRAD data are categorized into four levels according to different characteristics. In this study, Level III data (NEXRAD Information Dissemination Service, or NIDS) are adopted. The data are a series of 24 products (selected products are detailed in Table 2) created by the Radar Product Generator (RPG). Specifically, base reflectivity (N0R) data obtained in clear-air mode with the highest sensitivity and one hour total precipitation (N1P) data are used. Therefore, the N0R data with a large range (15DBz to 55 DBz) are better for visualization and animation purposes. The precipitation data are available only when weather
becomes severe, and therefore the data are appropriate for quantitative analysis of certain severe weather phenomena, such as hurricanes and flash floods.

The Unidata Internet Data Distribution (IDD) system has been used to manage and disseminate Earth science related data, including the NEXRAD outputs. Through the IDD system, users can select data from available products for various applications. Registered users can use the Local Data Manager (LDM) as a subsystem to obtain real time data products. The LDM provides client and server programs as well as event driven protocols within the IDD system to support near real-time data distribution (Davis and Rew, 1994). Because LDM supports data distribution in many user-specified data types (such as GOES data) and data formats (such as NetCDF), the entire data system can be extended to support various Earth science data products with highly flexible protocols.

In this case study for the proposed framework, the implementation is based on the following four major strategies: (1) the LDM is adopted as the Data Ingest Manager to receive near real-time data products from the Geosensors of the Unidata system; (2) the Data Broker converts the newly received data products from the LDM into geo-referenced images, and manages the versioning of products over time; (3) metadata management ingests proper meta information; and (4) the Data Broker pushes the GIS-ready data to the distributed GIS components, which implement near real-time WMS, map animations and spatiotemporal analyses. Major components include the LDM, Data Broker, GIS server modules, and GIS client modules (Figure 2). The LDM ingests data products from a configured data provider and transforms the formats of various earth science data products through a set of decoders. Due to the use of various legacy data formats in Earth science research, these decoders are designed according to the “add as needed” principle.

4.2 System design

Level III data are delivered through the IDD in near real time. Two data providers are configured for our system. If the connection to the upstream site is lost, a backup upstream site can be used to ensure that
data can be delivered continuously. Complying with the framework design in Section 3, the system includes four major components as illustrated in Figure 1. The LDM is responsible for data ingest. The data broker has four functions: 1) converting the NIDS level III radar data to composite, geo-referenced images, 2) populating data archives with dated image files, 3) maintaining one PNG image as the most recently updated data for each data layer, and 4) populating metadata for new data products or new data layers. The map servlet component will acquire service metadata and communicate with map viewers. The different scenarios of communication include real-time WMS, map animation and spatiotemporal analysis. We leverage on the Unidata IDD physical network in this study by installing the LDM, the GIS server module, and the Data Broker on a 2-CPU 3.06Ghz hyperthread server with 4 Gigabytes of memory and 100 Mbps connection to the Unidata IDD system. The client can be any web browser that supports Java Applets with the minimum configuration of 1.0 Ghz CPU and 512 Mbytes memory, and a 128 kbps connection to the Internet.

4.3 Scenarios

4.3.1 Real-time WMS. The system developed here can be accessed in the following sequence (Figure 4):

1) A user submits a request and the viewer issues the getCapabilities request to the server.

2) The server returns metadata for service and data to the viewer in XML format.

3) The user selects the Region of Interest (ROI) and map layers.

4) The server generates the ROI map and returns the map to the client.

These operations on the server side are implemented by different components of the request parser and response generator. The server always uses the current N0R map for visualization (Figure 5).
4.3.2 Map animation with temporal WMS. In addition to the near real-time WMS, the system also supports animation of temporal images in the following sequence (Figure 4):

1) The client side issues a request to obtain a list of all available radar images within a given time frame using the following selection parameters: desired start time, end time, and the bounding box representing the ROI.

2) The server parses the request to obtain the relevant parameters and searches for data files from the spatial data repository according to the submitted parameters. The list of available images will be prepared by the server and sent back to the client in XML format.

3) The user selects images according to this returned list for animation. The NEXRAD servlet will receive multiple getMap requests according to the time parameters.

4) The applet viewer overlays both the static background images and the images for animation.

When the animation reaches a new frame, its corresponding time will be highlighted on the time list (Figure 6). This communication method has at least two advantages: 1) the web communication load is leveraged by a series of NEXRAD images instead of a single movie clip or animated image. It is possible that some frames may be lost occasionally during the transmission and rendering, but most of the frames should be rendered. 2) The animated map is generated in an applet viewer using asynchronous image loading to avoid unnecessary delay. On the client side, it is also easier for the user to track each frame.
4.3.3 Support for spatial-temporal analysis. The server can also obtain precipitation values from the NEXRAD data repository, and select data according to the user’s defined time frame and ROI specifications. As a subprogram of the IDD system, Nex2img can generate the N0P precipitation data in gray scaled PNG images according to the parameters in the Lookup Table File (LUTFILE) (Unidata, 2007). N0R reflectivity data are not used here because information is partially lost after categorization according to the color table.

Higher resolution results may be obtained from the N1P product, but that product does not have a linear value setting throughout its full range, making it inappropriate to use. Table 3 depicts the categorization of precipitation based on pixel values, and thus precipitation maps can be generated.

After precipitation values are obtained, statistical summaries and analyses of the data can be performed on either the server side or the client side, depending on the level of operation complexity and the existing workload on both sides. We have developed server side analysis capabilities. For example, given a time frame and ROI, the server will compute and return the daily average precipitation in XML format. A histogram can be generated to show the daily precipitation for the time span (as in Figure 7). A trend analysis can be extended by incorporating certain aspects of hydrological models.

The interoperable interfaces and spatiotemporal analysis modules can be customized to accommodate different application domains according to the specific research needs. For example, the daily average precipitation information can be used to support a road weather information network (Koonar et al. 2005).

5. Conclusion and discussion

Recognizing the current redundancy problems between Earth science data dissemination and application domains, we propose an interoperable framework to perform near real-time geospatial processing through the WMS. Using event-driven protocols, the data server can efficiently obtain near real-time data products from data providers, and these data products can be further processed and archived by the image-processing module. When the resultant map is ready for GIS applications, the web-based GIS server can
communicate with any WMS viewer, through which domain experts or application researchers can utilize the data in near real-time and integrate them with their own data. The temporal dimension supported by WMS is also leveraged to provide capabilities in spatiotemporal data operations and analyses.

To fully exploit the potential of such an interoperable framework, we have to address some remaining issues. For example, precipitation data derived from radar need extensive validation and verification to ensure their quality (Ahnert, 1984). Rain gauge data and Digital Elevation Model (DEM) data may be integrated into the system to enhance the precipitation estimation (Vieux and Vieux, 2002). These integration and validation procedures should precede the generation of radar data images, and they can be added as additional components within the data broker. Because the number of snapshots keeps increasing over time, Earth science data products are often very large. The data will take up too much disk space if the system keeps both the data in original formats and the images for GIS representation. Near real-time data compression is needed to handle this data volume issue.

It is our plan to implement the proposed framework using a series of interoperable GIServices in a grid computing-based environment to support natural hazards analyses (Foster and Kesselman, 2006), such as flood forecasting (Lien et al. 2004), near real-time routing (Cao 2007), and other environmental decision support applications (Manca et al. 2006). This paper has chosen the NEXRAD data provided by the IDD system to demonstrate the feasibility and potential of the proposed framework. As mentioned before, the IDD system also supports the GOES weather data, which can be processed and disseminated using the same framework. In fact, the Unidata’s IDD system disseminates many Earth science data products (for more information about the four major types of data supported by IDD, please visit http://www.unidata.ucar.edu/data/) and therefore, the proposed framework can be used to process and disseminate many other data products, such as model data, and point data for lightning and wind. We incorporated several functions supported by IDD in our system components (Data Ingest Manager and Data
Conceptually, the proposed framework will also work with other data systems if similar functions exist in those systems. Otherwise, they can be developed. Although the specific implementation of the framework described in this paper is about NEXRAD and relies on IDD, the applicability of the proposed framework is not limited to NEXRAD data and is not dependent upon the Unidata’s IDD system. The adoption of interoperable interfaces in the proposed framework makes it easy to integrate other datasets. For example, it took about one year to develop the entire framework, but only several days to incorporate the AVHRR data with coverage of the entire United States.

Acknowledgement:
This research was supported by a Chinese 973 project (2006CB701306), US FGDC 2005 CAP (05HQAG0115), and the NSF/Unidata 2003 Equipment Award. We acknowledge the assistance of the following individuals: Wenwen Li supported XML parser development; Tom Yoksas and Steve Chiswell helped extensively on Unidata’s IDD system; and Daryl Herzman provided useful suggestions on georeferencing NEXRAD images. Comments from the Editor and two anonymous reviewers are greatly appreciated.

Reference


Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)


GEOSPATIAL DATA ABSTRACTION LIBRARY (GDAL), Available online at: http://www.gdal.org/ (accessed 15


Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)


Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)


Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)

Figure 1. General framework for disseminating earth science data.
Figure 2. Framework design: *data preprocessing* modules provide geospatial data and metadata so that interoperable communications can be performed between *map servlet* and *map viewer*. 
Figure 3. Service and layer metadata.
Cao, Yang, and Wong, 2008, An interoperable spatiotemporal weather radar data dissemination system, IJRS (in press)

Figure 4. Client server communication for the scenarios of real-time WMS, map animation, and spatiotemporal analysis.
Figure 5. Real time NEXRAD Level III base reflectivity map (with DEM data from WMS integrated as background).
Figure 6. Animation of NEXRAD level III base reflectivity data (with AVHRR data integrated from WMS as background).
Figure 7. Trend analysis of daily average precipitation.
Table 1. WMS operations

<table>
<thead>
<tr>
<th>Operations</th>
<th>Request format</th>
<th>Return format</th>
<th>Return content</th>
</tr>
</thead>
<tbody>
<tr>
<td>getCapabilities</td>
<td>(Service URL)?request=</td>
<td>XML</td>
<td>List of data layers on the server, associated legends, projections, and other service metadata</td>
</tr>
<tr>
<td></td>
<td>getCapabilities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>getMap</td>
<td>(Service URL)?request=</td>
<td>Image</td>
<td>A mapped image for a specific geographic location, with a given set of layers and rendering options</td>
</tr>
<tr>
<td></td>
<td>getMap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>getFeatureInfo</td>
<td>(Service URL)?request=</td>
<td>XML</td>
<td>Attributes associated with a given point on the map</td>
</tr>
<tr>
<td></td>
<td>getFeatureInfo</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. NEXRAD data products

<table>
<thead>
<tr>
<th>Macro name</th>
<th>Product content</th>
</tr>
</thead>
<tbody>
<tr>
<td>N0R</td>
<td>Reflectivity</td>
</tr>
<tr>
<td>N1P</td>
<td>1 Hour Precipitation</td>
</tr>
<tr>
<td>NTP</td>
<td>Storm Total Precipitation</td>
</tr>
<tr>
<td>N0V</td>
<td>Velocity</td>
</tr>
<tr>
<td>N0S</td>
<td>Storm Relative Motion</td>
</tr>
<tr>
<td>N0Z</td>
<td>Long Range Reflectivity</td>
</tr>
</tbody>
</table>
Table 3. N1P value setting.

<table>
<thead>
<tr>
<th>Category</th>
<th>Precipitation value (in inch)</th>
<th>Gray pixel value (0-255)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$0 &lt; p &lt; 2$</td>
<td>$10 \leq v &lt; 110$</td>
</tr>
<tr>
<td>II</td>
<td>$2 \leq p &lt; 6$</td>
<td>$110 \leq v &lt; 210$</td>
</tr>
<tr>
<td>III</td>
<td>$6 \leq p &lt; 12$</td>
<td>$210 \leq v &lt; 255$</td>
</tr>
</tbody>
</table>