01-WE

World Multiconference on Systemics, Cybernetics and Informatics

July 22-25, 2001
Orlando, Florida, USA

PROCEEDINGS

Volume I

Information Systems Development

Organized by IIIS
International
Institute of
Informatics
and Systemics

Member of the International
Federation of Systems Research

IFSR

Co-organized by IEEE Computer Society
(Chapter: Venezuela)

EDITORS
Nagib Callaos
Ivan Nunes da Silva
Jorge Molero

A Web-Enabled System for Storage and Retrieval of GOES-8 Meteorological data from a Semantic Database*

Martha E. GUTIERREZ, Naphtali D. RISHE, Scott GRAHAM, Gianina G. ROCHA, Xiangyu YE
High Performance Database Research Center
NASA Regional Applications Center
Florida International University
Miami, FL 33199, USA

Ouri WOLFSON
Department of Electrical Engineering and Computer Science
University of Illinois at Chicago
Chicago, IL 60612, U.S.A.

ABSTRACT

Meteorological data is a major component in the study and understanding of environment behaviors and their impact on society. Examples of the effects of hazardous weather conditions range from mild structural damage and agricultural and commercial losses to massive death and destruction. Meteorological data enables the prediction and analysis of weather phenomena and its possible life-threatening and property-damaging effects.

The Geostationary Operational Environmental Satellite (GOES) system provides timely global weather information, including advanced warning of developing storms. Using GOES imagery, meteorologists are able to better study and understand atmospheric circulation patterns and behaviors. Capable of generating over 18 GB of meteorological data per day over the eastern United States, the GOES-8 is the first of this new generation of Earth-observing satellites.

The High Performance Database Research Center (HPDRC), in conjunction with the NASA Regional Applications Center (RAC), has developed a web-enabled system which houses several gigabytes of GOES-8 meteorological data during each hurricane season. This system uses semantic database (Sem-ODB) technology to efficiently store and retrieve GOES-8 atmospheric data. Combining newly developed technologies and algorithms with the broadcast capabilities of the World Wide Web, HPDRC and NASA RAC make this vast source of information available to the general public and fellow scientists around the world.

Keywords: Semantic Database, GOES-8, Internet Application, Meteorological Data, and Image Data.

1. INTRODUCTION

Understanding atmospheric behaviors aids meteorologists and weather forecasters in providing improved advanced warnings of thunderstorms, flash floods, hurricanes, and other severe weather conditions. Accurate forecasts save lives, preserve property, and benefit agricultural and commercial interests.

Aided by the rapid evolution of remote sensing technologies, the usage of meteorological satellites has become a major method for the acquisition of weather imagery and data. Data acquired by these satellites or satellite systems provides highly accurate and detailed information facilitating the study, analysis and understanding of weather phenomena. One such system, the GOES, is a basic element of US weather monitoring and forecast operations and a key component of the National Weather Service (NWS) modernization program. The GOES system is an essential cornerstone of weather observations and forecasting used for short-range warning and "now-casting."

Seeking to encourage further research advancements and development, HPDRC has designed a web-enabled system housing large amounts of GOES-8 weather and atmospheric data. This system uses semantic database technology to achieve efficient storage and retrieval of large amounts of data. Provided with a Web-interface, this database system makes queries of weather measurements, imagery and data available to a vast area of the scientific community.

2. HPDRC

The HPDRC is a division of Florida International University (FIU), School of Computer Science. It conducts research on database management systems and various applications, leading to the development of new

* This research was supported in part by NASA (under grants NAG5-9478, NAGW-4080, NAG5-5095, NAS5-97222, and NAG5-6830) and NSF (CDA-9711582, IRI-9409661, HRD-9707076, and ANI-9876409).

types of Database Management Systems (DBMS), new database techniques, and the refinement of existing ones.

The development of Sem-ODB, a prototype for a massively parallel Semantic/Object Oriented DBMS, is the largest project at the HPDRC. Our system is useful for most typical database applications, as well as for specialized domains such as Earth Sciences [2].

Sem-ODB

Sem-ODB, the Semantic DBMS developed by HPDRC, is based on the Semantic Binary Model. In the Semantic Binary Model, the information is represented by logical associations (relations) between pairs of objects and by the classification of objects into categories. The Semantic Binary Model is a natural and convenient way of specifying the logical structure of information and for defining the concepts of an application's world. [1]

The Semantic database models are potentially more efficient than the conventional models for two main reasons. The first reason is that all the physical aspects of the representation of information by data are invisible to the user. This creates a potential for optimization by allowing more changes without affecting the user programs. The second reason is that the system knows more about the meaning of the user's data and about the meaningful connections between such data. This additional knowledge can be utilized to organize the data so that meaningful operations can be performed faster at the expense of less meaningful operations. [1]

The mathematical abstraction of the relational model has allowed the introduction of powerful and easy-to-use languages for retrieval and updates of databases. The semantic model however, offers a higher degree of abstraction, which results in more concise user programs, speedier processing (due to optimization), and a wealth of other features. Relational databases are good for general conventional database applications. However, in situations where the structure of information is complex, where greater flexibility is required (objects with unknown identifiers or where objects move from one category to another, etc.), or where non-conventional data is involved (spatial data, long text, images, etc.), semantic databases need to be considered.

The efficient retrieval and updates are a requirement of the semantic database. Performance is maximized by decomposing queries into atomic retrieval operations; each atomic retrieval request is optimal. Each query result is composed of a set of facts. These facts can state that the objects belong to a category, they can state that there is a relationship between objects or they can be facts relating objects to data, such as numbers, texts, dates, images, etc. HPDRC's Semantic DBMS contains semantic facts and inverted semantic facts. This fact inversion scheme assures efficiency of queries including range queries and content access and also exhibits low entropy of data blocks, which facilitates compression. [1]

3. RAC

The NASA Regional Applications Center (RAC) at FIU is a constituent part of HPDRC. The RAC Program was initiated by the Applied Information Sciences Branch of NASA Goddard Space Flight Center (GSFC) to extend the benefits of its information technology research and cost-effective system development to a broader user community. The RAC objectives are based on the goal of fostering the use of environmental and Earth resource data by regional institutions. The ultimate goal of the RAC is to establish a fundamental set of remote sensing technologies that can be assembled by a specific user community, to meet the information needs of that community.

4. A WEB-ENABLED WEATHER SYSTEM

Aided by semantic database technology, our innovative weather system efficiently acquires, stores, organizes and provides access to GOES-8 data enabling its use for the investigation of oceanographic and meteorological phenomena. Four major components comprise this web-enabled system. These components are: a vast archive of GOES-8 meteorological data, a database system capable of efficiently storing and manipulating large amounts of data, an ingest system constantly acquiring up-to-the-minute imagery and measurements, and a World Wide Web interface.

GOES-8 Meteorological Data

Spacecraft and ground based systems work together to accomplish the GOES mission of providing weather imagery and quantitative sounding data for weather forecasting and related services. GOES satellites provide the continuous monitoring necessary for intensive data analysis. They circle the Earth in a geosynchronous orbit, providing a constant vigil for atmospheric "triggers" of severe weather conditions such as tornadoes, flash floods, hailstorms and hurricanes. When such conditions develop, these geostationary meteorological satellites monitor storms and track their movements.

Provided with a five-channel imager and a nineteen-channel sounder, the GOES meteorological satellites acquire high-resolution visible and infrared data as well as temperature and moisture measurements of the atmosphere. Furthermore, they introduce two new features. The first, flexible scanning, allows small-area imaging; a feature that allows the focus of localized weather trouble spots to improve short-term weather forecasts and now-casts. The second feature is simultaneous and independent imaging and sounding. This feature allows the usage of data from both, imager and sounder instruments, to increase the accuracy of weather analyses and forecasts.

In this manner, data provided by the GOES-8 satellite, also called GOES-EAST, includes detailed weather measurements, frequent imagery, and new types of atmospheric soundings. Located at 75°W, the GOES-8 satellite generates data covering most of the eastern US and South America. Furthermore, this meteorological data, combined with information from new Doppler radars, aids water resource managers as they make critical decisions about allocating precious water resources.

The Database System

Based on Sem-ODB technology, our weather database is primarily used for monitoring of severe tropical weather phenomena during hurricane season. Specifically, our Hurricane Season 2000 database archives over 26GB of GOES-8 Continental United States (CONUS) imagery and data. A high growth potential is observed due to the constant acquisition of imagery. Implementation of the semantic database model and an efficient schema design enable the fast querying and efficient manipulation of such considerable amounts of data.

The database is categorized by hurricane season, which in turn is categorized by date. Figure 1 presents the semantic database schema.

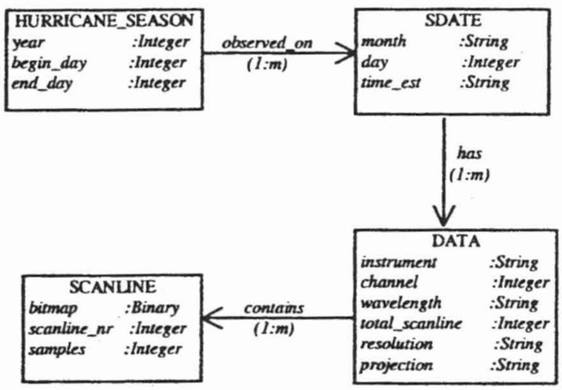

Figure 1 – Hurricane Database Semantic Schema

The schema contains four Categories: HURRICANE_SEASON, SDATE, SCANLINE, and DATA. The HURRICANE_SEASON category contains the year of the hurricane season, and the starting and ending Julian day for the season, this distinguishes one hurricane season from another and allows storage of multiple seasons in one database. The HURRICANE_SEASON category contains many SDATE, which describes the month, day, and time of the data set. Each DATA object may contain many channels and many instruments. Each Channel can have many SCANLINEs, which contain data of each pixel in a scanline.

For GOES-8 GVAR, each CONUS data file is approximately 65 MB in size and is stored by scanlines. Storing the data by scanlines permits efficient scaling down of the image during retrieval. To scale an image that is stored by scanline with the scale factor *N*, only every N^{th} scanline needs to be retrieved, and then every N^{th} pixel within each scanline is picked up. It efficiently reduces retrieval time for other applications that need information about only specific area or scanlines in a similar way. This scheme improves the efficiency and flexibility of database retrieval for related applications.

Ingest System

The automatic ingest system uses UNIX scripts to communicate with the ingest machine, provide WEB access, and automatically load data into the hurricane database. This system is composed of four main modules: (1) File Downloading, (2) Data Processing, (3) Image Generating, and (4) Database Updating.

Figure 2 shows the control and data flow of the whole system. The ovals in the figure contain mainly processes the rectangles contain data and the arrows represent data/control flows.

The functionality of each module can be described as follows:

(1) File Downloading: A list of the current files in the system is downloaded via FTP from the ingest machine to the processing server. From this list, the latest file is selected. This file is then downloaded to the server for further processing. This is a timed process and varies depending on the requirements.

(2) Data Processing: The GOES-8 ingest system broadcasts the data in GVAR (GOES Variable Retransmission) format. This defined format consists of twelve distinct blocks numbered 0 through 11 including 10-bit pixel data and the detector information arranged in correct order. [3] When the GVAR image is processed, the data is organized into five bands and each of the five

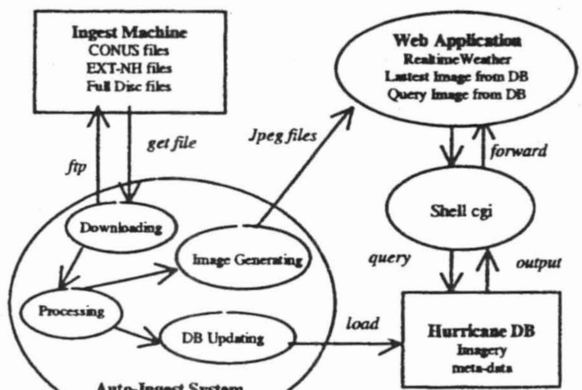

Figure 2 – Auto-Ingest System Structure

bands are output to a file in BSQ format for easier manipulation.

(3) Image Generating: The BSQ file is processed and an image is generated. This image is scaled down, converted into JPEG format and then put on the web page (in the latest data section). The web site is constantly updated with the latest weather image of the Eastern United States.

(4) Database Updating: The image is loaded into the hurricane database at set intervals. The method used for storing remote sensed digital imagery is directly related to the performance obtained at retrieval time. The decision of which method to use depends on the size of the image, the characteristics of the data and the related applications. There are several ways to store the data. For low resolution images, or images were algorithms will be performed, it is efficient to store the data by scanline. For high resolution images, or images where only a certain portion will be viewed, it is efficient to store the data by tiles. A combination of the two techniques can be used to store the data by swaths. For this system, we chose to store the data by scanline since this is a low-resolution image and the data attributes are relative to each scanline.

Web Interface

The web interface represents a gateway to our vast archive of meteorological imagery. It displays the latest weather image and facilitates the querying of archived GOES-8 data. The Web interface is divided into three frames, as shown in Figure 3.

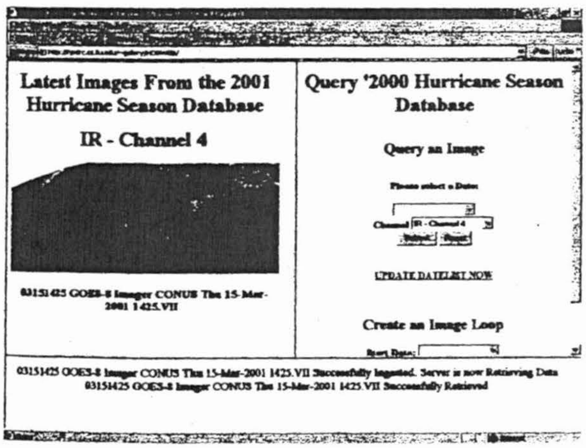

Figure 3 – Web Interface

The first is the "Latest image" frame. This frame displays the latest CONUS image acquired from our GOES ground station and processed by our ingest system. The "Latest image" is updated from the main automatic ingest system once every 30 minutes.

The second frame shows the Hurricane Database list. This frame provides links to available Hurricane Databases for the current as well as previous years. Following a database link provides access to the database query form. This form allows the user to query CONUS images of 12-hour intervals corresponding to specific dates and channels. In addition, image loops may also be created by specifying a loop start and end date. Image loops resemble the animations used during evening newscasts. They allow the observation of past atmospheric patterns and behaviors over a given time period.

The third frame is the console. It displays the server's current status as well as the date and time associated with the last ingested image. Server error messages are also displayed in the console frame.

5. CONCLUSION

The usage of remote sensed satellites, such as the GOES-8, has made major contributions and advancements to the study and monitoring of atmospheric and oceanographic phenomena as well as to the detection of forest fires and extreme changes in the environment. As the demand for meteorological data continues to grow, a specialized system that efficiently handles large amounts of detailed data produced by remote sensing devices is required.

The merger of Sem-ODB technology with Web resources has enabled the development of a reliable and efficient source of meteorological imagery and data. This system not only fulfills the need for fast data manipulation, but also facilitates access, regardless of possible limitations on connections speeds. Real time and historical data is then made available and can further be used in conjunction with other types of data or meteorological tools.

6. REFERENCES

[1] N. Rishe, Database Design: The Semantic Modeling Approach, McGraw Hill, 1992.

[2] N. Rishe, W. Sun, D. Barton, Y.Deng, C. Orji, M. Alexopoulus, L., Loureiro, C. Ordonez, M. Sanchez, A. Shaposhnikov, "Florida International University High Performance Database Research Center", SIGMOD Record, 1995, vol.24, no. 3, pp.71-76.

[3] Space Systems/Loral Operations Ground Equipment (OGE) Interface Specification, DRL 504-02-1 Part 1, Palo Alto, CA, 1995.

980-07-7541-2
57000>
9 789800 775417
ISBN: 980-07-7541-2