Linking ArcIMS With an Internet File System

Steven H. Wong, Hongli Luo, and Sandra L. Aguilar

Abstract - Our approach offers an alternative way for dynamic web-based mapping. An Internet file system can be accessible via web browsers in which users can upload, download, view, and execute files. Our system will allow users select shape files from the web interface of the Internet file system for mapping using the ArcIMS services. The integration of the ArcIMS with the Internet file system provides substantial benefits to GIS users because the application takes advantage of the powerful search, versioning, security, backup, and other features offered by the Internet file system and its underlined enterprise database management system (DBMS). On the other hand, the value of the Internet file system and the associated DBMS is enhanced by the ArcIMS mapping capabilities.

Introduction

The advance of Internet mapping has been impressive. ArcIMS is one of the most popular platforms for publishing enterprise geospatial data over the World Wide Web. We believe that strengthening the integration of ArcIMS with the mainstream Database Management System (DBMS) will empower users, especially those who do not perform geospatial work on a regular basis.

One of the mainstream web-based DBMS capabilities is the content management technology. This technology provides many benefits to users and organizations, such as desktop-like user interfaces, sophisticated security features, easy upload/download of files, and powerful content search. In this paper, we describe Oracle Internet File System (iFS), and our preliminary effort on integration iFS with ArcIMS. We expect that this type of integration will make it more likely for mainstream users to accept web mapping concept and tools.

Oracle iFS has evolved to the current Content Management Software Development Kit (CMSDK). However, most iFS features described in this paper are applicable to the current Oracle CMSDK without much modification. In this paper, we only get into a subset of the iFS features. We will first go over some innovative iFS techniques of searching and rendering information. We then will take a look at how we link these capabilities with ArcIMS. We believe that the coupling of efficient text search in iFS with spatial search/visualization in ArcIMS will allow better and increased use of geospatial data and functionality.

Enabling Metadata Integration with Data Contents in iFS

The iFS expands the database platform to present documents and media as files and folders that users may access through familiar interfaces such as Windows, the web (HTTP), FTP, and AFP (Apple File Protocol) [1]. In this section, the authors attempt to introduce an important iFS feature that allows the innovative integration of metadata with data contents. Then, in the next section, the authors show the iFS text search capability that enables efficient online searches on data content, metadata, and the combination of the two (data plus metadata). We will describe our enhancements to the iFS. Readers are referred to [2] for detail description of the iFS and its

rich features.

Because of their critical role in data quality control, data discovery, and data longevity, metadata are recognized as a vital part of any geospatial data set. Without the metadata and the associated information on data quality and other issues, the usefulness of the geospatial data set in question is significantly compromised. However, the enabling technology to seamlessly integrate metadata with data contents in a geospatial data set has been elusive. Metadata and data contents are typically separated into different files for storage, search, and retrieval. In common practice, data contents in the form of files or DBMS tables are accompanied by one or more metadata files. Most metadata files contain linkages in the form of Uniform Resource Locators (URL) to the data set. While it is advantageous to serve metadata as files in clearinghouses for search and data discovery, for systems that already contain both the metadata and data contents, the systems are not fully and efficiently utilized in terms of information search if the systems serve metadata as separate files.

The following two examples illustrate the inefficiency of searching information from separated metadata files and data files. Both examples use the same text search criterion: Find all data sets that contain the text "fish" in the Abstract field of the FGDC metadata AND the text "seagrass" in their data content. Notice the above criterion implies a combination search on metadata and data contents.

In the first example, a user initially went through the metadata search from an online metadata clearinghouse. The search resulted in a list of metadata files that contain the text "fish" in the Abstract field of the FGDC metadata. Unfortunately, that was the easy part. In order to find data sets that contain the text "seagrass" in their data content out of the resulted list of metadata files, the user needs to download all the data files through linkages within the metadata files before searching the data content. The search on the data content from downloaded files may have to be done on the user's desktop computer. One can imagine the monumental task of performing this type of search for 200 returned metadata files!

The second example illustrates performing the same search as above but from the opposite direction, *i. e.,* searching from the data content to the metadata content. The user initially searched for all data files that contained the text "seagrass" in their data content. Now the difficult part comes when the user has to compile all metadata files associated with the resulted data files from the initial search before searching for the compiled FGDC metadata files that contain the text "fish" in the Abstract field. At the end, the user needs to match the data files with the resulted metadata files to obtain the search result, *i. e.,* data sets that contain the text "fish" in the Abstract field of the FGDC metadata AND the text "seagrass" in their data content.

The above examples elaborate that the conventional text search for geospatial data could become rather complicated and cumbersome when the search criterion involves both the metadata and data contents. The cause of the problem is the separation of metadata from the data contents even when both the metadata files and data contents are physically stored in the same server or information system. This problem can be resolved by innovatively applying a feature in the Oracle iFS, *i. e.,* Extensible File Properties to integrate metadata with the associated data contents.

With the Extensible File Properties [3], iFS makes it possible to add properties to specific file types, such as the abstract of the data set. One may use this feature to incorporate FGDC metadata fields into file properties of iFS files. Since file properties are part of a data file that also includes data contents, iFS provides a platform for the integration of metadata with data

contents.

Figure 1 shows three views of a data file in the Oracle iFS. The panel in the middle contains data files in the iFS displayed with a web browser via HTTP protocol. The panel to the right shows data content of one of the data files in the middle panel. The user accesses the data content by mouse clicking on the file name in the middle panel to open the data file. The panel to the left displays the file properties of the data file concerned. The user accesses the file properties by checking the selection box next to the data file in the middle panel, then selecting Properties from the Edit tool.

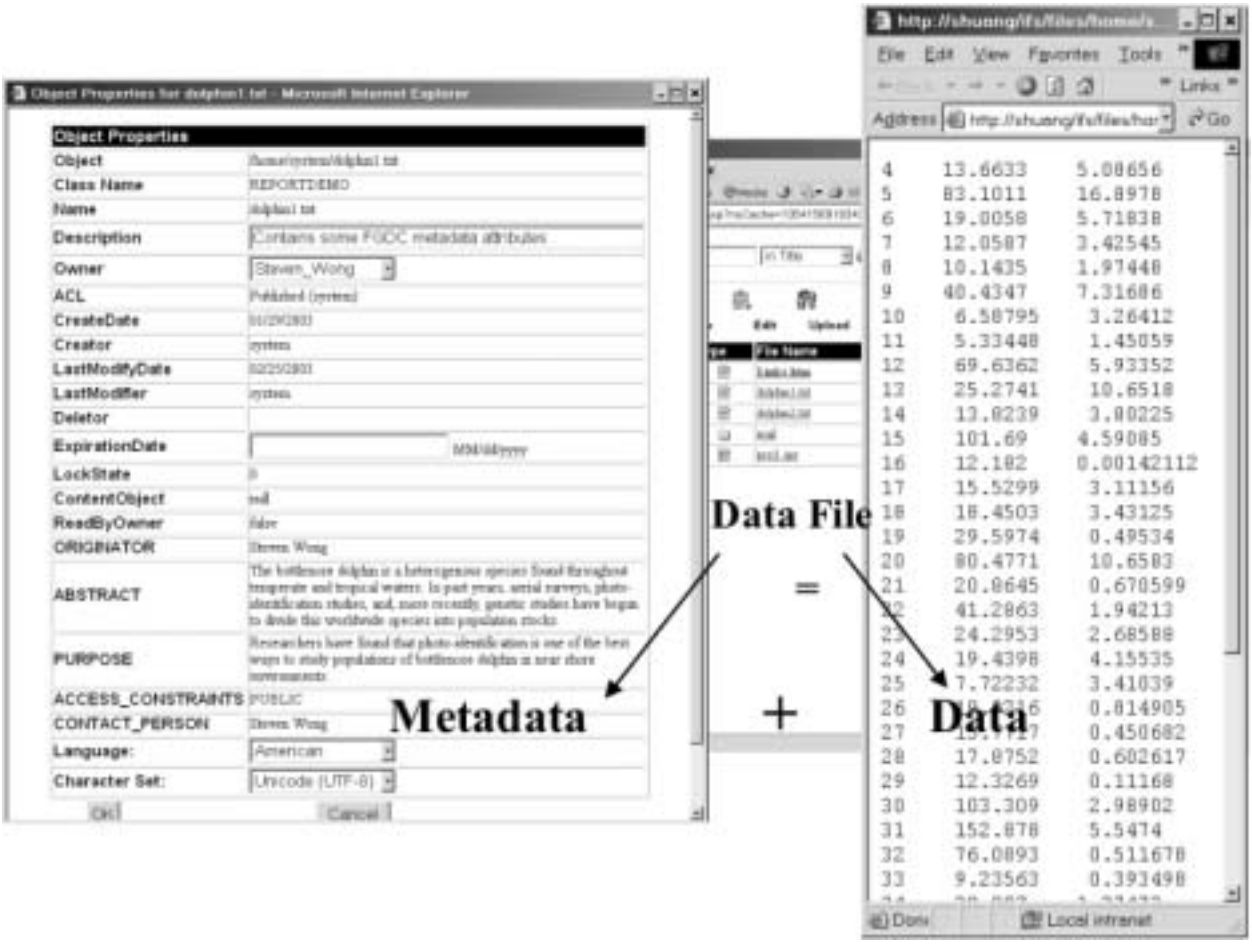

Figure 1. Three views of a data file in the Oracle iFS. The middle panel displays the files names, the right panel the data content of one of the files in the middle panel, and the left panel the metadata content of the data file in concern. The metadata and data contents are integral parts of a data file in the iFS.

Upon close examination of the left panel in Figure 1, one may notice that besides some generic fields for file properties such as CreateDate, and LastModifyDate, there are fields such as ABSTRACT, ACCESS_CONSTRAINTS that resemble certain fields in the Federal Geographic Data Committee (FGDC) metadata standard. In fact, the latter fields are added to the built-in iFS file properties by our research team using a technique called Subclassing in the iFS. With some more efforts using the Subclassing method, one can map all or parts of FGDC metadata fields into the iFS file properties. This innovative application of the iFS feature (Extensible File Properties) realizes the integration of metadata with the data contents. A data file in the Oracle iFS is composed of the data content file and the metadata in the form of file properties. One can access both the data contents and the metadata from a single user interface (Fig. 1). The metadata contents go along with the data contents whenever the data file is moved in the Oracle iFS. This tight association of metadata and data contents reduces the cost of managing metadata files separated from the data files caused by the conventional approach of handling metadata. With the incorporation of FGDC metadata into the iFS file properties, information on data quality becomes integral part of any data files.

FGDC metadata within the iFS file properties can be edited easily by the user. We have developed utilities to facilitate FGDC metadata mapping between the iFS file properties and metadata files in XML, HTML, and other formats. Oftentimes, the user can generate FGDC metadata files with commonly used metadata tools. These FGDC metadata files can also be submitted to clearinghouses for metadata search. Using the metadata mapping utilities, one can easily move metadata contents between iFS file properties and FGDC metadata files. In the next section, the reader will gain more understanding on the power of incorporating FGDC metadata into the iFS file properties to perform advanced text search.

Advanced Text Search in iFS

In the previous section, we discussed two mechanisms for online text search: data content search and metadata search. In this section, we will demonstrate how one can perform these searches in the Oracle iFS. In addition, we will provide a new search mechanism that can perform combination searches on both FGDC metadata and data contents. This new search mechanism innovatively takes advantage of two iFS features: Extensible File Properties and Advanced Search.

An online search on data contents can be easily performed in the Oracle iFS. Figure 2 shows an example of online search on data contents in the iFS. The search criterion is: Find all data files that contain the text “seagrass”. Notice the search tool on the upper part of the web page (below the URL address). The user simply types “seagrass” in the Find text editor, and selects “in Document” as the search target before entering the Return key. The search result contains all documents in the iFS that meet the search criterion and accessible to the user. It is worthwhile to point out that the resulted documents are in various formats such as Microsoft Word, Acrobat PDF, and HTML. One of the key features of the Oracle iFS is the powerful text indexing for efficient text search on various file formats. It is also interesting to note that the user can easily access both the metadata and the data contents associated with a data file from the user interface in Figure 2 thanks to the integration of FGDC metadata with the iFS file properties. This easy access to metadata and data contents is less costly than the user’s having to

physically match the metadata files associated with the data files returned as the result from conventional search on data contents.

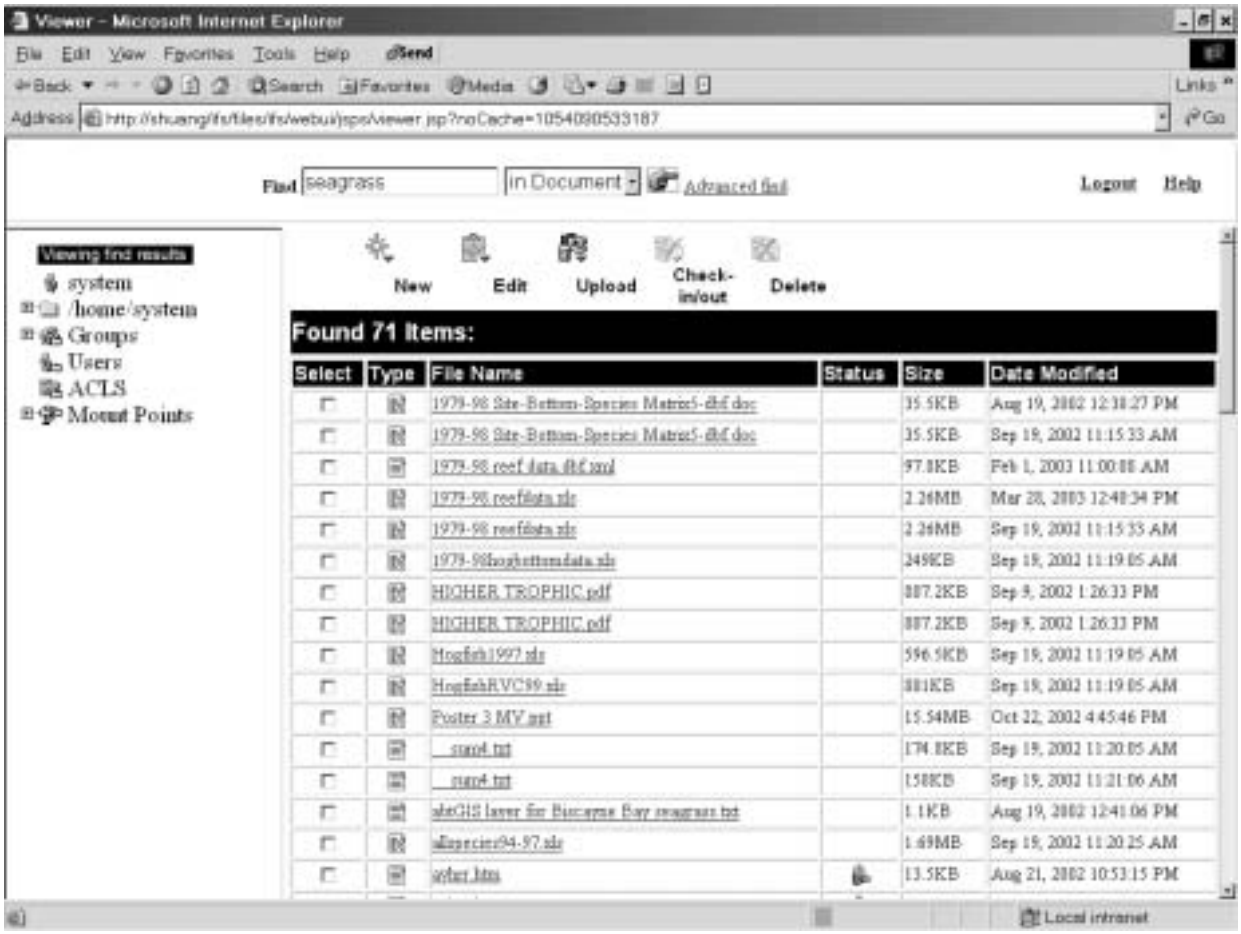

Figure 2. Example of an online search on data contents in the Oracle iFS. The search criterion is: Find all data files that contain the text "seagrass". The above search criterion is expressed in the iFS as: Find in iFS where Document contains "seagrass".

The incorporation of FGDC metadata into the iFS file properties makes it possible to conduct metadata searches using the iFS advanced search feature. Figure 3 shows the Oracle iFS metadata search interface with the similar search criterion: Find all data files that contain the text "fish" in their Abstract field of the FGDC metadata standard. The resulted data files are returned with an interface similar to that of Figure 2.

So far in this section, the authors have demonstrated the data content search and the metadata search in iFS style. There are clearly advantages of using iFS to perform these searches compared to conventional approaches. These advantages include: (1) availability on the Internet via HTTP protocol, (2) no restriction on file formats involved in the search, and (3) easy access to both metadata and data contents from the resulted interfaces. Advantage (3) is the

result of the integration of metadata with the data contents by incorporating the FGDC metadata standard into the iFS file properties.

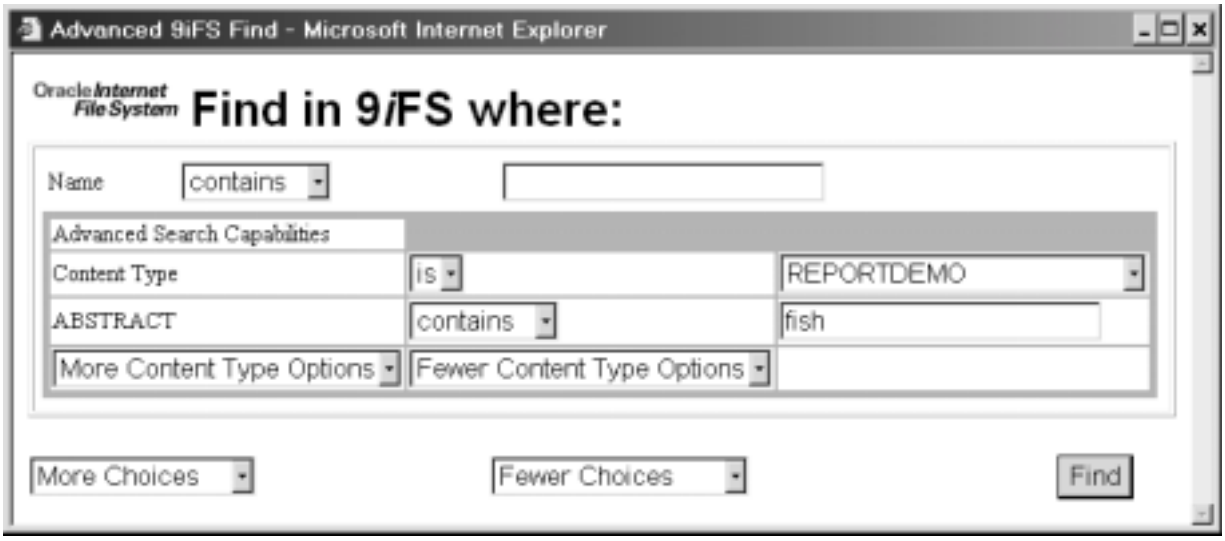

Figure 3. An example of the Oracle iFS advanced search. The search criterion is: Find all data files that have the text "fish" in their Abstract metadata field. The above search criterion is expressed in the iFS interface as: Find in iFS where ABSTRACT contains "fish".

Now we are ready to introduce an efficient online text search mechanism, *i. e.*, the combination search on FGDC metadata and data contents. As discussed in the previous section, the conventional way of information search does not handle the combination search efficiently and elegantly. Let us again use the abovementioned criterion for an example on the combination search: Find all data sets that contain the text "fish" in the Abstract field of the FGDC metadata AND the text "seagrass" in their data content. Once the FGDC metadata fields are mapped to the iFS file properties, the user can easily perform the combination search using the iFS interface for advanced search (Figure 4). With a simple user interface, the Oracle iFS handles the combination search elegantly. This type of combination text search on metadata and data contents should be facilitated by information systems that store both the metadata and data contents. This combination search mechanism saves the user from tediously looking for information from disparate and ever-increasing files containing metadata and data contents.

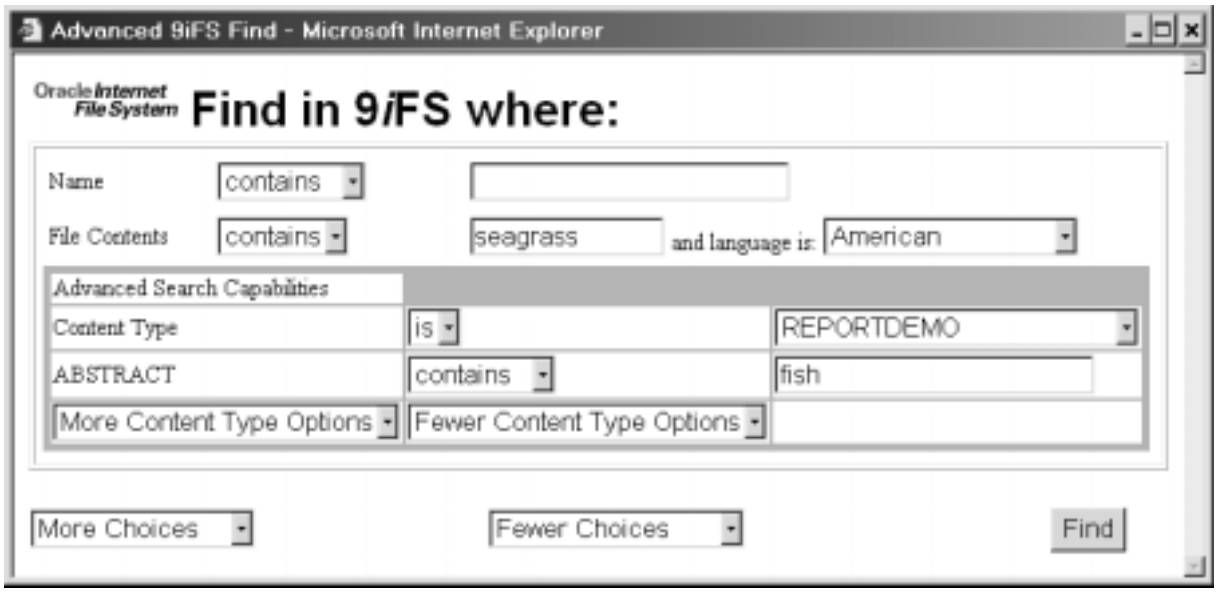

Figure 4. An example of the Oracle iFS interface for combination search on metadata and data contents. The search criterion is: Find all data sets that contain the text "fish" in the Abstract field of the FGDC metadata AND the text "seagrass" in their data content. The above search criterion is expressed in the iFS interface as: Find in iFS where File Contents contains "seagrass" AND ABSTRACT contains "fish".

Visualizing iFS Data with ArcIMS

The previous sections provide descriptions of some general and enhanced features of the iFS, especially the advanced text search functionality. Here we would like to show how we incorporate ArcIMS into Oracle iFS so that geospatial mapping can be performed seamlessly from the iFS interface.

Figures 5 and 6 show our design of data visualization initiated from the iFS web interface. In Figure 5, the user selects three shape files to be visualized, and clicks on the Map button. The selected files are then displayed in Figure 6 with a customized ArcIMS interface.

The integration of iFS with ArcIMS is made possible by the use of: 1) Oracle iFS Java APIs, 2) ArcIMS XML requests, and 3) ArcIMS map service.

It should be noted that the iFS user does not need to know anything about ArcIMS, or web mapping technology *per se*. The user (or innocent user) simply uses iFS to upload/download data, search data as described in the previous sections, browsing data and metadata contents, and visualize data. This is different from ArcIMS use paradigm by GIS practitioners or someone who purposely enters into the ArcIMS interface to map data. ArcIMS adds significant value to iFS by enabling innocent mainstream users easily visualize enterprise data in geospatial context without leaving the users' familiar interfaces. These familiar user interfaces are designed for common functionality in a web-accessible DBMS such as text search on metadata and data contents. Once the innocent user enters into the ArcIMS interface via iFS interface, the user may be exposed to additional datasets and geospatial functionality such as data search with drag-and-

drop spatial select, and spatial query.

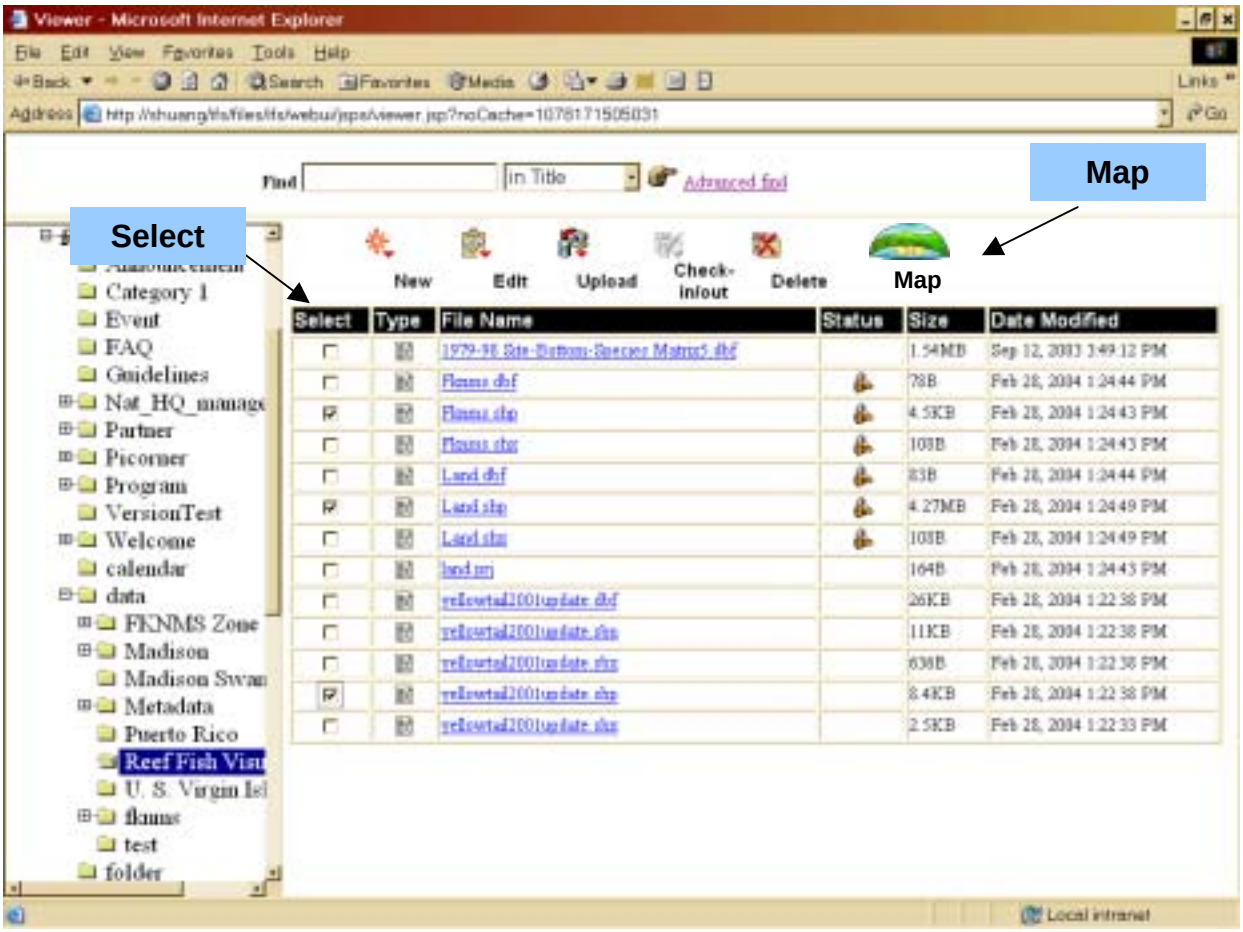

Figure 5. User selects 3 shape files in the above iFS web interface for data visualization in the ArcIMS.

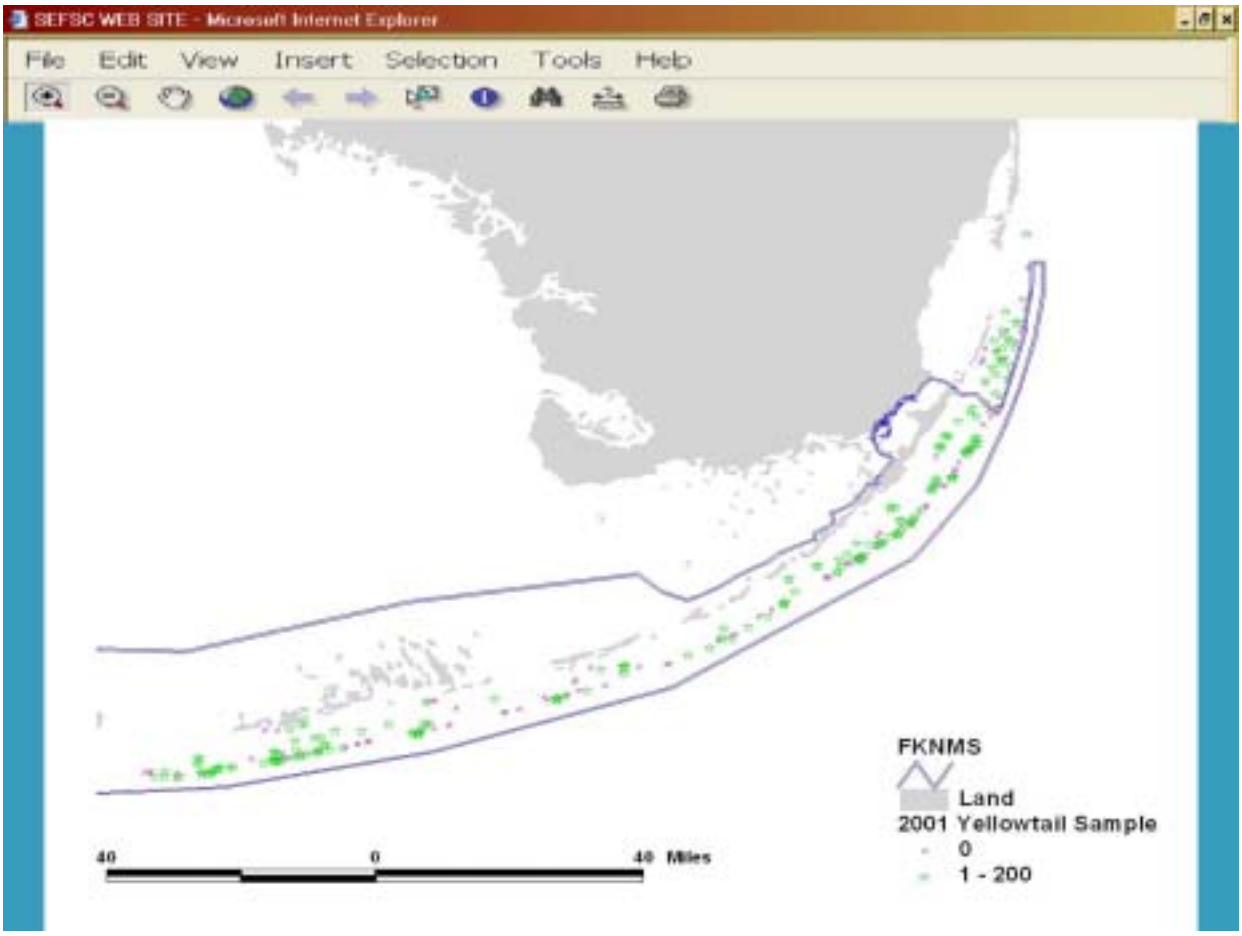

Figure 6. The 3 selected shape files in the iFS web interface (Fig. 5) are displayed in the customized ArcIMS interface.

Acknowledgment

The authors greatly appreciates Mr. Huaichen Yang for his technical assistance on Oracle products. Dr. Peter Thompson, Dr. Mei-Ling Shyu, and Mr. Doug Harper have contributed significantly to the research project in terms of their insights, reviews, and comments. The authors also expresses special thanks to the Office of Habitat Conservation of NOAA Fisheries for funding support for the project under the NOAA Coral Reef Initiative Program.

References

[1] Oracle Corporation. 2002. Oracle Internet File System Data Sheet. http://otn.oracle.com/products/ifs/htdocs/9ifsGeneralFAQ.html.

[2] Oracle Corporation. Oracle Internet File System Setup and Administration Guide. Release 9.0.1. http://download-west.oracle.com/docs/html/A81197_06/toc.htm.
[3] Oracle Corporation. 2002. Feature Overview of Oracle Internet File System (9iFS). http://otn.oracle.com/products/ifs/htdocs/9ifsFO.html.

Author Information

Steven H. Wong, Physical Scientist, National Oceanic and Atmospheric Administration (NOAA), 75 Virginia Beach Drive, Miami, FL 33149, 305-361-4496 (office), 305-361-4499 (fax), Steven.Wong@noaa.gov.
Hongli Luo, Graduate Student, Dept. of Electrical and Computer Engineering, University of Miami.
Sandra L. Aguilar, Jardon Howard Technology, Inc.