DIAL Access to Distributed Digital Earth Remote Sensing Data

Ken McDonald[1] Liping Di[2] R. Suresh[2]

[1]*NASA Goddard Space Flight Center, Code 423.0*
Greenbelt, MD 20771
E-mail: ken.mcdonald@gsfc.nasa.gov

[2]*NASA/RITSS*
4400 Forbes Boulevard
Lanham, MD 20706
E-mail: Liping.Di@gsfc.nasa.gov

ABSTRACT Data and Information Access Link (DIAL) is a Web client-server based data and information system for ingesting, documenting, managing, and distributing digital Earth remote sensing data over the Internet.While powerful, it is compact, easy to set up and use, and has minimal computer power and maintenance requirements. It permits data producers to set up a server rapidly on their desktops, making data available via the Internet. The system provides both catalog and data services to data users through web browsers. The catalog services permit data users to search data archives at directory and inventory levels with spatial, temporal, and parameter criteria. The data services allow users to interactively browse, animate, plot, subset, subsample, reformat, and download the selected data. The system is also designed to interoperate at both catalog and data object levels, allowing one DIAL site to interoperate with other DIAL sites or other data systems by using interoperability protocols such as NASA V0 IMS and CEOS CIP. The interoperability of DIAL allows formation of a federation of data providers, enabling users to find and access distributed digital Earth archives in the federation by searching only one of the federated sites. The DIAL system has been developed and is available for major UNIX and Windows 95/NT platforms. Several national and international research programs have selected DIAL as their distributed data and information system. The DIAL software can be downloaded at no cost from http://dial.gsfc.nasa.gov

KEY WORDS Open Distributed System, DIAL, Catalog Interoperability, Data Interoperability, Data and Information System, Digital Earth

1. Introduction

DIAL is a web-based data and information system that enables scientists and small data producers to format, document, and distribute digital Earth remote sensing data through their sktop computers. The system allows data users, by using common Web browsers, to interactively search and select data in DIAL sites (*catalog services*), manipulate and visualize the selected data, and download selected data in their favorite spatial and temporal coverage and resolution, parameters, and data formats (*data services*). Multiple DIAL sites can form an interoperable DIAL network (federation) through DIAL's interoperability protocols. While maintaining the independence of individual DIAL sites, the network of DIAL sites can provide a unified information space, which would enable data users to search and access data in all sites in the network by querying just one of them. DIAL's demonstration site is at http://dial.gsfc.nasa.gov [DEG, 1999]

DIAL system was developed in response to recommendations of the National Research Council (NRC) for the future generation of data and information system for Earth Science. Instead of a large, more centralized system, NRC recommended a more distributed, principal Investigator (PI) controlled, individually small and flexible but collectively large and interoperable data and information system for Earth science (data federation) [NRC, 1996]. Recent development in information technologies, especially Internet and World Wide Web, makes such a distributed system possible.

DIAL was designed to support the federation concept and has many essential features of data interoperability required for the successful implementation of a federation, including a common user interface and the use of data, metadata, and interoperability protocols and standards. DIAL uses WWW standards such as http, html and Java applets in its implementation. It can work with any metadata standards, such as

FGDC [FGDC, 1998], since its metadata catalog system is totally configurable and a mapping mechanism between different metadata systems is provided. DIAL can work with many independent data providers at distributed

locations within a federation and provide interaction within the federation and outside users.

The DIAL software package is compact (less than 5 Megabytes) and does not need programming knowledge to use. Since it is completely developed from public domain software and protocols, it is not dependent on any COTS package. It is available on multiple platforms such as SGI, SUN, DEC, Linux, and PCs (Windows 95/NT). The source code is written using C and Java. We have distributed more than 1,000 copies of the DIAL software. Current DIAL installations include EOS PIs, NOAA's Pacific Marine Environmental Laboratory, JPL, international partners (Japan, UK, China, IGBP START Regional Centers) and others. It is also in operational use by NASA's Advanced Composition Explorer (ACE) satellite project.

2. DIAL Architecture

The DIAL architecture is modular, extendible, and is based on standards (Figure 1). It consists of five major parts:

1) A catalog interoperability layer to support major protocols for interoperation among DIAL servers and between DIAL and other data systems.
2) A scaleable scientific data server (SDS) to *serve complex scientific data in multiple formats* and to allow data users to interactively manipulate the selected data so that they can obtain the data in their

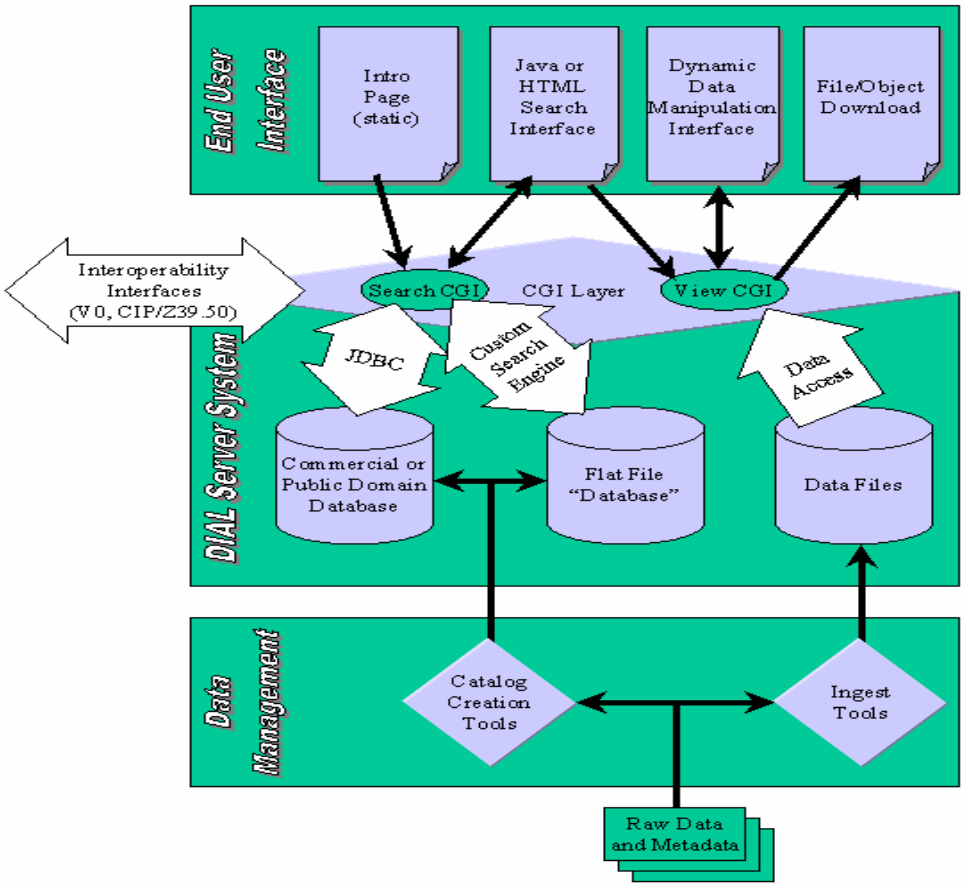


Figure 1. The architecture of DIAL system

favorite form in terms of spatial and temporal coverage and resolution, parameters, and format. Although data users can download data managed by DIAL in multiple formats, in the server side it currently works only with HDF and HDF-EOS formatted data sets. Data translation programs, which convert data between HDF-EOS and commonly used GIS formats, such as Arc/INFO Exchange format, GeoTIFF, ERDAS LAN, and ArcView Shape, are also available. In addition, DIAL's architecture allows extension for accessing data in other formats.

3) An open database interface layer to interface with Open Database Connectivity (ODBC)-capable metadata catalog databases for powerful data search and finding. In case the data producer does not have an expensive commercial database system, DIAL has a search engine with a file-based catalog database (binary table) to provide basic metadata database capability. The binary table option provides the same search capabilities as the ODBC-compatible databases, but the search speed will slow down when the table is very large.
4) A user interface layer to support user-friendly interaction between client and server. Because the user interface is built within Web data browsers through html protocols, DIAL end users can easily integrate other application software as Web helper applications.
5) A suite of software tools to help data producers to ingest the data into the system and to help data users to download and analyze the data.

The technical details of each component of the DIAL system can be found in Di et al., 1999.

3. Catalog Services

Catalog services provide users access to the catalog of a data archive so that users can find their requested data in the archive. They are one of the basic services that any data and information systems have to provide to data users. DIAL provides a full range of catalog services to data users, including

- Spatial, temporal, and parameter-based catalog search at both inventory and directory levels.
- User-friendly Java search interface with query preview
- Automatic creation of searchable catalog based on metadata
- Support of ODBC/JDBC compatible databases for storing metadata/catalog
- HTML search interface for slow network connection

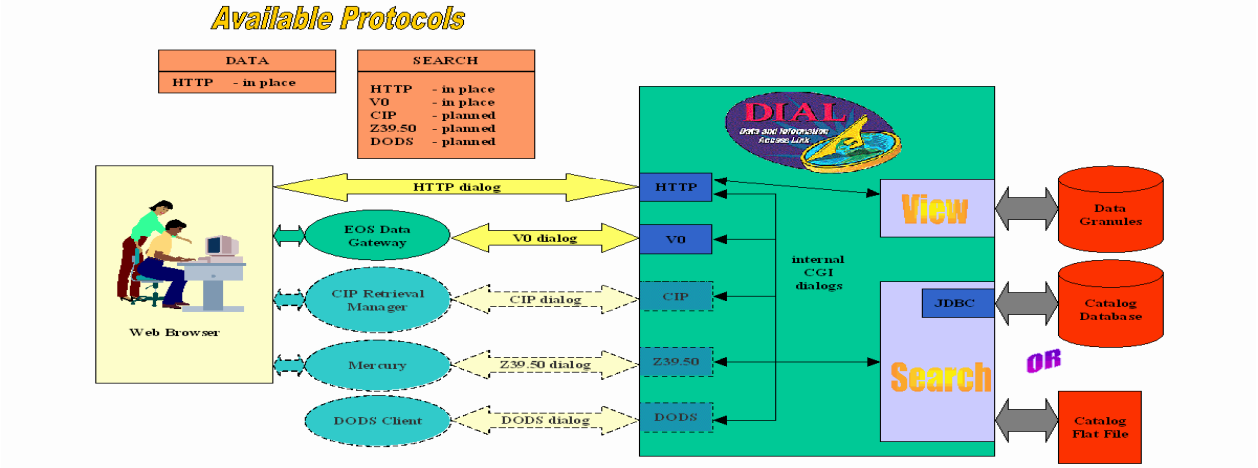
4. Data Services

Once the data users have find the needed data in a DIAL site, the next logical steps are to browse the data and associated metadata, visualize the data, manipulate the data if necessary, and download the selected data to users' machines. Ideally, a data server should be able to 1) present data to users in a unified form, hiding all physical storage characteristics of the diverse data in the server; 2) allow users to browse the data interactively to determine their usefulness before downloading; 3) permit users to subsample and subset the data in spatial, temporal, and parameter domains, and possibly reproject the data spatially; and 4) let users download the data in their favorite formats so that the data can be ingested into their analysis system easily. DIAL is able to meet all above requirements by providing the following on-line data services to users:

- On-line access to data and metadata;
- Multi-granule and single-granule data subsetting and subsampling based on array coordinates or record numbers;
- Multi-granule and single-granule data subsetting and subsampling based on geographic/map coordinates and physical parameters;
- Dynamic/interactive visualization of two dimensional data with on-the-fly browse image generation, both linear and histogram equalization image enhancement, and color palette selection;
- Client-controllable animation of time series or high dimensional data;
- Interactive color composites of multi-spectral data;
- Coastal and political boundaries overlay;
- On-line downloading of data in multiple formats, including ASCII, binary, HDF, and HTML;
- X-Y plotting for non-image data.

5. Access to Distributed Digital Earth Remote Sensing Data

With the above-mentioned catalog and data services, DIAL is a powerful data and



information system for disseminating digital Earth remote sensing data to users. Users can discover data and information contents of a particular DIAL site by querying its catalog service. However, with relevant scientific data located at tens or hundreds of DIALs and other scientific data servers around the world, it is impractical and time consumptive for data users to find each of these data servers and search the servers individually with a same set of search criteria. A better approach should be to build a network of data servers that links all relevant servers together so that users can search all relevant servers in the network by searching any one of servers at once. This requires that related catalog services interoperate to form a single, federated information space, which would enable users to search all sites in the federation by querying just one of them. The network of distributed interoperable scientific data servers is called a data federation. DIAL has been implemented with a catalog interoperability layer for this purpose.

The key to a successful federation is the ability to exchange metadata and data among participants and users of the system. This can be accomplished by using a set of interoperation protocols at both catalog and data levels. Currently DIAL supports the Earth Observing System Data and Information System (EOSDIS) Version 0 interoperability protocol [Suresh et al., 1999] and is implementing the Committee on Earth Observing System (CEOS) Catalogue Interoperability Protocol (CIP) [CEOS 1996], a geospatial profile of ANSI Z39.50 [ANSI, 1995]. These protocols enable a DIAL server to interoperate with not only other DIAL servers but also other data systems implemented with those protocols. Figure 2 shows the implemented and planned protocols in DIAL. With those protocols, DIAL is able to form a distributed data network for users to search and find the digital Earth remote sensing data in the network with just one search. Once the data are found, DIAL is able to provide data services to users through the HTTP protocols.

6. Examples of Distributed DIAL Networks

Currently several research programs have selected DIAL system to build distributed data networks for providing data accesses to scientists nd the general public. One example of such networks is the NASA Earth Science Information Partnership (ESIP) DIAL network. Working Prototype-Earth Science Information Partnership (WP-ESIP) Federation is a NASA prototype for future data product generation, distribution and services to the users. The objective of the WP-Federation is to experiment with evolving processes to make digital Earth science data easy to archive, locate, and distribute for multiple applications. One of the challenges for the Federation prototype is how to link various partners through data and information systems to provide interoperable catalog and possibly data services to users. DIAL is an ideal candidate for building a distributed data network for ESIP because of its interoperability and functionality.

Currently, a DIAL network has been formed in the NASA ESIP. The network has linked the University of New Mexico ESIP, the John Hopkins University/IBM ESIP, the JPL GENESIS ESIP, the NASA Goddard Space Flight Center (GSFC) Distributed Active Archival Center (DAAC), and

the NASA EDC DAAC together through the EOSDIS V0 protocol. More ESIP members are likely to join the DIAL network.

The ESIP DIAL network is very diverse internally. It consists of two distinguished data and information systems, the EOSDIS V0 and DIAL. While DIAL is mainly for small data producers and individual scientists, the EOSDIS V0 system was developed specifically for large data centers. The DIAL system can provide both catalog and data services while the EOSDIS V0 system provides interoperable catalog services. The NASA DAACs are large data centers who distribute large amounts of standard Earth science data products to users through the EOSDIS V0 system. Other ESIP DIAL network members are middle and small data producers who use DIAL for data distribution. The interoperability among the individual nodes in the DIAL network enables the network to hide those diversities from the data users.

The experimental international DIAL network is another example of DIAL networks. The network was set up as one of the Global Observation Information Network (GOIN) experiments. The goal of GOIN is to promote global environmental research by providing information and network support for disseminating and sharing the digital Earth science data. GOIN was started as a joint effort between U.S. and Japan and currently has been expended to Asia-Pacific countries. The international DIAL network has much wider geographic distribution than the ESIP DIAL network. The current DIAL network has nodes at NASA, National Space Development Agency (NASDA) of Japan, the START Southeast Asia regional center in Bangkok, the START East Asia regional center in Beijing, the National University of Singapore, and the National Taiwan University. The experimental DIAL network was successfully demonstrated at the joint CEOS/GOIN meeting in September 1999. The demonstration showed that the international DIAL network can provide both catalog and data services with satisfactory performance. It also showed that we could build an operational DIAL network across the globe with the current Internet capacity.

7. Conclusion and Future Plan

The DIAL system is a compact yet powerful data and information system that can provide both catalog and data services to users across the globe. The major advantage of the DIAL system is that it enables the individual scientists and small data producers to disseminate data to global users quickly and directly. The interoperability built in the DIAL system allows the formation of a global network of distributed data and information system.

Currently, we have developed and released DIAL version 2.4, which is available for SUN Solaris, SGI IRIX, Compaq Alpha UNIX, PC Linux, and Windows 95/98/NT. You can obtain the software without cost at http://dial.gsfc.nasa.gov. In the next years, we plan to implement CIP and other interoperability protocols, enhance the functionality for both catalog and data services, improve the user interface, and provide user support.

Reference

ANSI/NISO 1995 Information Retrieval: Application Service Definitions and Protocol specification: Z39.50. American National Standrds Institute, New York, New York.

CEOS 1996 Catalogue Interoperability Protocol (CIP) Specification - Release B. CEOS/WGISS/PPT /CIP-B, European Space Agency/ESRIN, Frascati, Italy.

Data Engineering Group (DEG) 1999, DIAL Release Version 2.4. Raytheon ITSS Corporation, Lanham, Maryland. (http://dial.gsfc.nasa.gov)

Di, L., R. Suresh, K. Doan, D. Ilg, and K. McDonald 1999. "DIAL-an Interoperable Web-based Scientific Data Server". In M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman edit, Interoperating Geographic Information Systems, Section 4. System Experiences. Kluwer Academic Publishers, Norwell, MA, pp 301-312.

Federal Geographic Data Committee (FGDC), 1998. FGDC content standard for digital geospatial metadata, version 2.0, Reston, Virginia. (http://www.fgdc.gov).National Research Council (NRC), 1996. ...

Suresh R., L. Di, L. Kumar, D. Ilg., and K. McDonald, 1999. A Web-based access to distributed heterogeneous geospatial data archives using NASA's EOSDIS V0 protocol. International Archives of Photogrammetry and Remote Sensing, V 32 Part 4W12: 243-247.