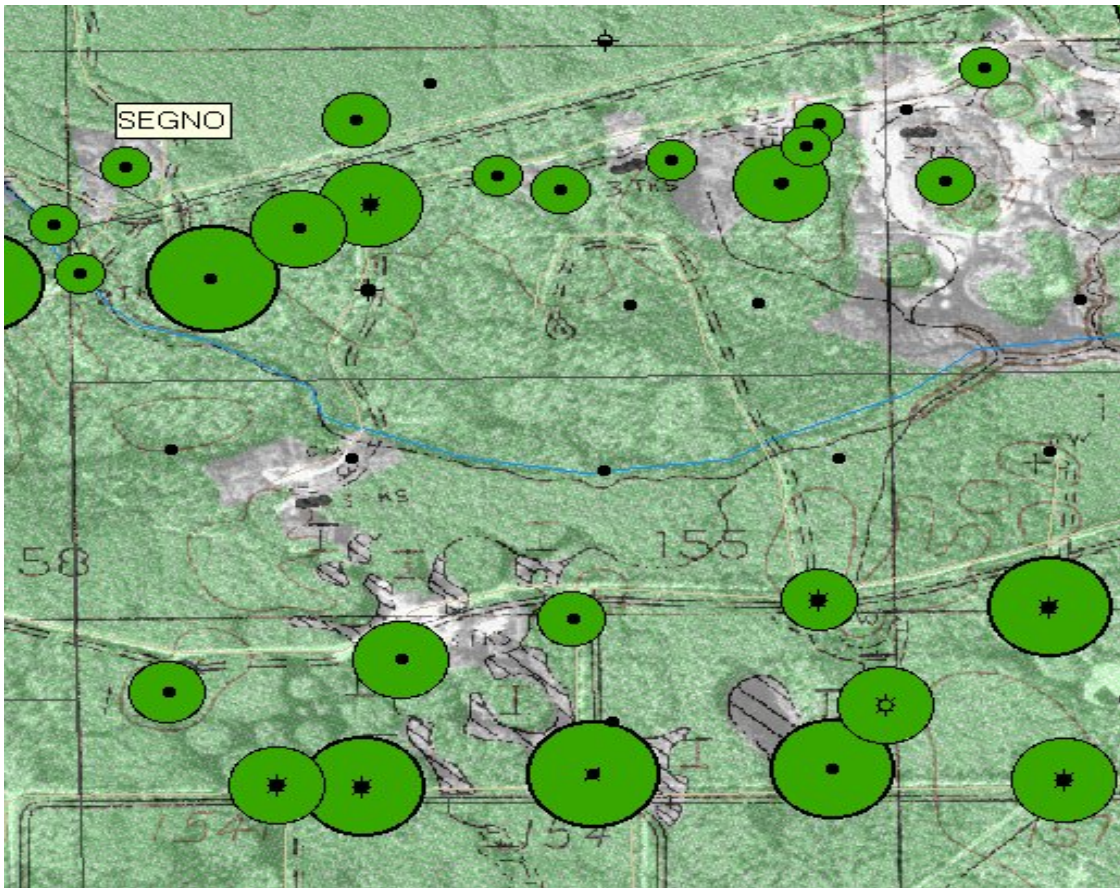


# Data Access and Integration Through GIS



**IHS Energy®**  
**Technical Solutions Group**  
**James R. Blaha, Director**

**August 2004**



---

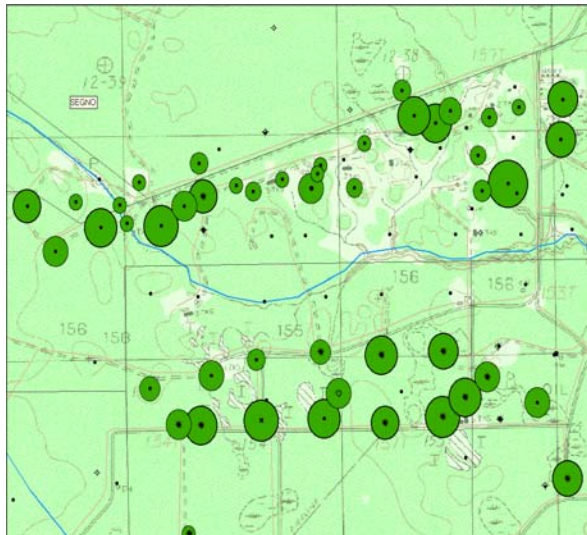
## Contents

|   |    |
|---|----|
| Executive Summary.....                    | 3  |
| Overview .....                            | 5  |
| Data Access .....                         | 6  |
| Spatial Layers .....                      | 6  |
| Spatial Layer Management .....            | 7  |
| Online Access .....                       | 8  |
| Data Integration .....                    | 9  |
| Simple Data Integration .....             | 9  |
| Enhanced Data Integration .....           | 10 |
| Advanced Data Integration .....           | 10 |
| Data Security and Portal Technology ..... | 12 |
| Solution Implementation .....             | 13 |
| Summary.....                              | 14 |



## Executive Summary

Oil and gas companies around the world are looking for ways to improve efficiency in all aspects of their business. To realize efficiencies, they need to be able to make better decisions faster and cheaper by analyzing all of the relevant data available to them in a systematic fashion. Good decision making starts with quick and easy access to clear and accurate data; no analytical solution, no matter how sophisticated, can compensate for incomplete or inaccurate data. Unfortunately, the data that is needed is very rarely stored in a common format or location and is rarely accessible through integrated, workflow-based, analytical solutions. Instead, users spend a large amount of time tracking down the data they need and loading it into the applications that they use before combining the results to make a final decision.



In the early 1990s, it was believed that monolithic database models and centralized repositories would solve application interoperability concerns. In spite of huge investments and the dedication of leading industry experts, this approach failed. The resulting data models were too complex, and the data conversion costs were too high. Consequently, the petroleum industry data management vision underwent a major transformation during the mid 1990s. A data-centric focus was adopted with a resulting shift from the data model to the data itself and the use of metadata catalogs and middleware to integrate data and applications. This approach has yielded some success, but the proprietary nature of the middleware formats and the cost of development of the associated conversion tools make the resulting solutions extremely expensive and complex to implement and manage.

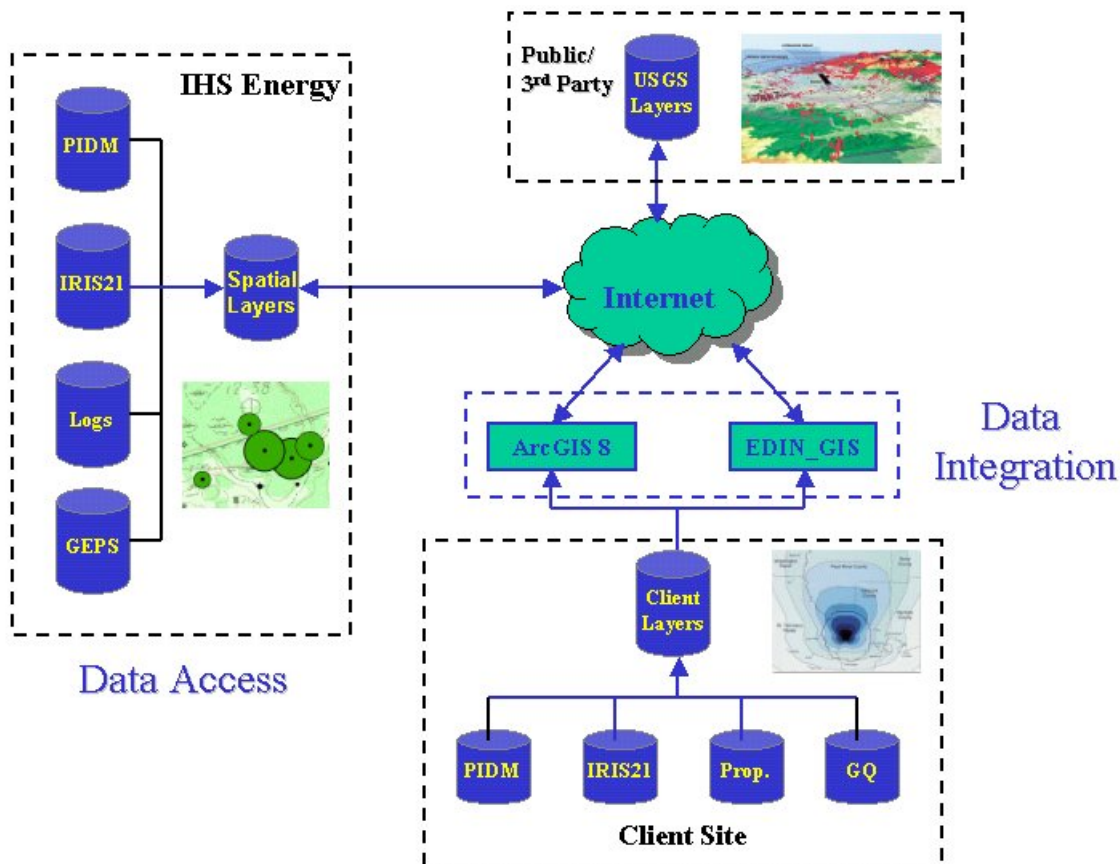
Over recent years, evolutionary technological advances have presented a more powerful and flexible option for multi-source data access and integration. Data can be managed in its native format and location but is referenced and represented through a de facto industry standard spatial file structure using the SDE® technology from ESRI®. This establishes a common data access format that then provides a platform for the integration of information from multiple sources that can be displayed, analyzed, and exported through commercial and customized Web-based or desktop map interfaces.

This White Paper discusses the common Data Access format and how it provides a foundation for different levels of Data Integration from straightforward to more sophisticated solutions. To support the concept, details on the underlying technology are provided as well as the challenges associated

with the design, development, and implementation of an integrated solution that meets essential performance and security requirements.

## Overview

The following diagram illustrates the concepts that will be described in this White Paper. Many commercial and public data vendors are making spatial representations of their meta-data available online in the SDE format using ArcIMS™ technology. These layers can then be accessed through turnkey Web browser solutions such as EDIN-GIS from IHS Energy or through desktop GIS solutions, including ArcGIS 8™. The ability to access data from multiple sources and locations within the same application immediately provides a basic level of data integration. This level of integration can be expanded upon to enable cross data-source queries by standardizing attribute sets, then further by providing extensions to ArcGIS, for example, that facilitate deeper levels of data mining.



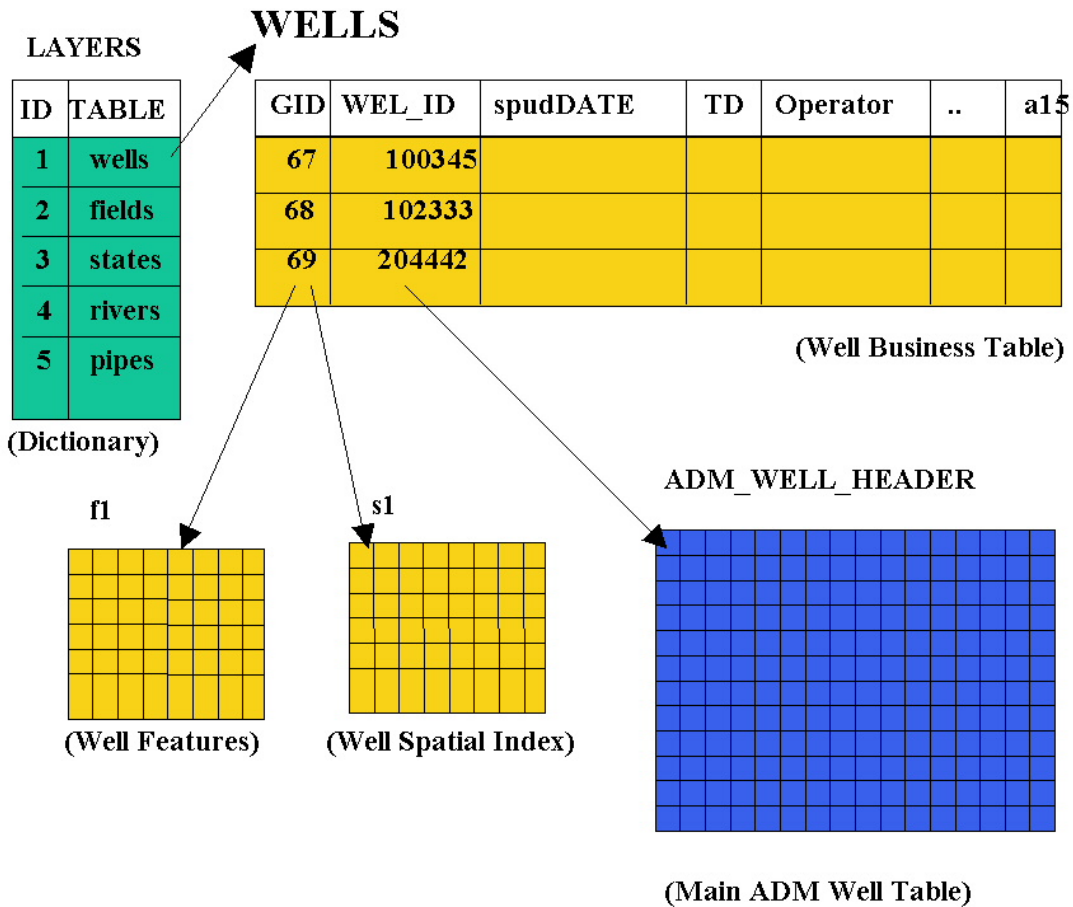
This Enterprise-level solution provides several ways to access the same information depending upon the needs of the user. Casual users wanting to perform straightforward browse and query functions might be satisfied with a lightweight browser interface such as that offered by EDIN-GIS. Users needing access to more sophisticated mapping and geographical query capability will want to interface with the information through a heavier desktop application such as ArcGIS 8. The important thing to note is that all sets of clients access the same information and that it is managed remotely by the vendor. The vendor ensures that the information is the most accurate and up-to-date available and reduces the complexity of the client's data management challenge which will, in most cases, lead to an overall cost reduction.

## Data Access

Establishing a common Data Access format is the foundation for developing a platform for integration. The SDE technology from ESRI combined with the appropriate management tools provides the necessary structure to create this common format.

### Spatial Layers

Using the SDE, spatial information is stored and managed using a robust RDBMS that is linked to associated attribute information. This spatial information can then be accessed, displayed, and queried either from desktop applications such as ArcGIS, or from a Web browser using ArcIMS technology.



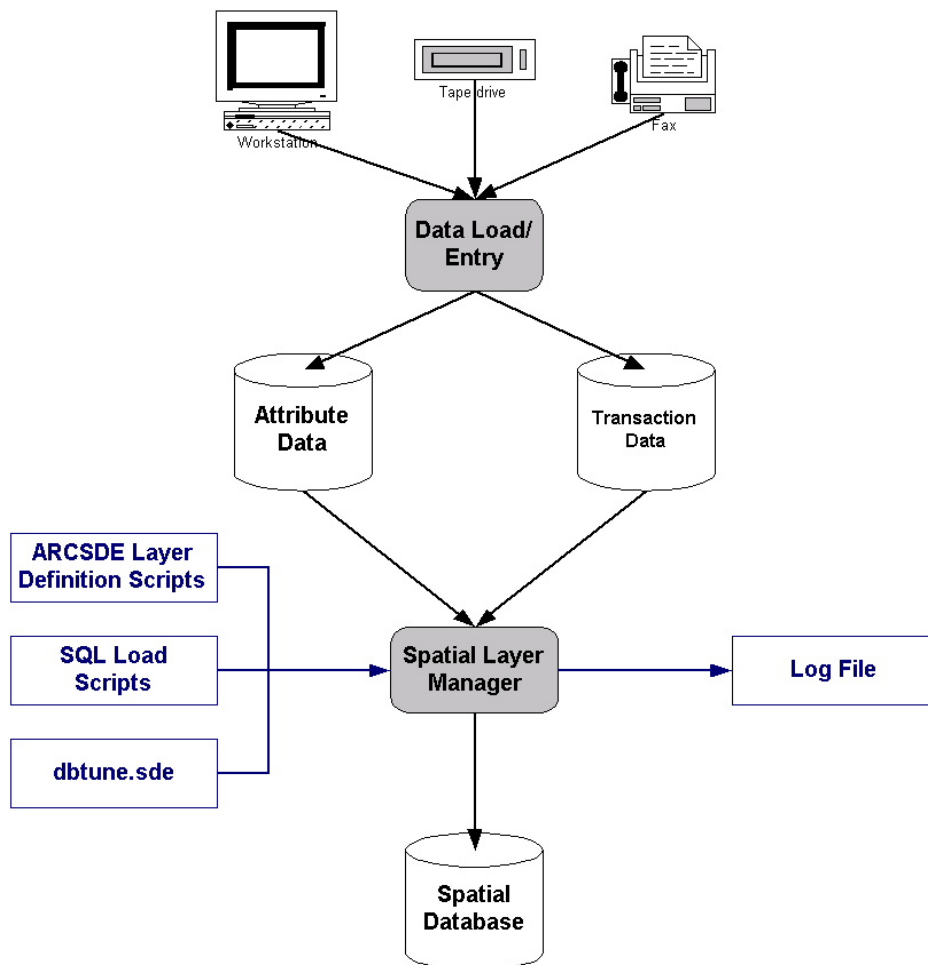
There are two primary methods for associating database attributes with spatial layers. The first is to create a view of the desired attributes using a standard SQL statement. The advantage of this approach is that the attribute data remains in the original tables so that there is no storage overhead and the attribute information is always guaranteed to be up-to-date. The disadvantage of this approach is that the performance of queries that access the attributes is much slower, particularly if the view has to perform table joins. The alternative is to directly store the attributes in the Spatial Layer Business Tables, as illustrated in the above diagram. This obviously increases the storage

requirement and requires strong procedures to ensure that the attributes in the Business Tables are continuously synchronized with the core database, but it delivers the best performance.

In an extensive series of tests, IHS Energy determined that using views was acceptable from a display perspective but that the performance of attribute queries was unacceptable for an Internet-based solution. Indeed, queries on Spatial Layers with attributes stored in the Business Tables performed five to ten times faster than the view approach, so this solution was adopted.

### Spatial Layer Management

The level of effort involved in updating and managing the Spatial Layers once they have been defined should not be underestimated with this approach. This effort is particularly cumbersome when large volumes of data are involved in creating the initial layer. In some cases it can take hours if not days to create the layers and associated Business Tables from scratch, which often must be done each time the data changes. To overcome this problem, IHS Energy developed the Spatial Layer Manager. This tool simplifies the process of defining, creating and populating the initial layer, but its true value is in ongoing synchronization of data. Instead of recreating the content from scratch, the tool is able to analyze the core database and only update those records that have been modified. This process can reduce update times from hours down to minutes and is used to manage IHS Energy's online spatial environment.

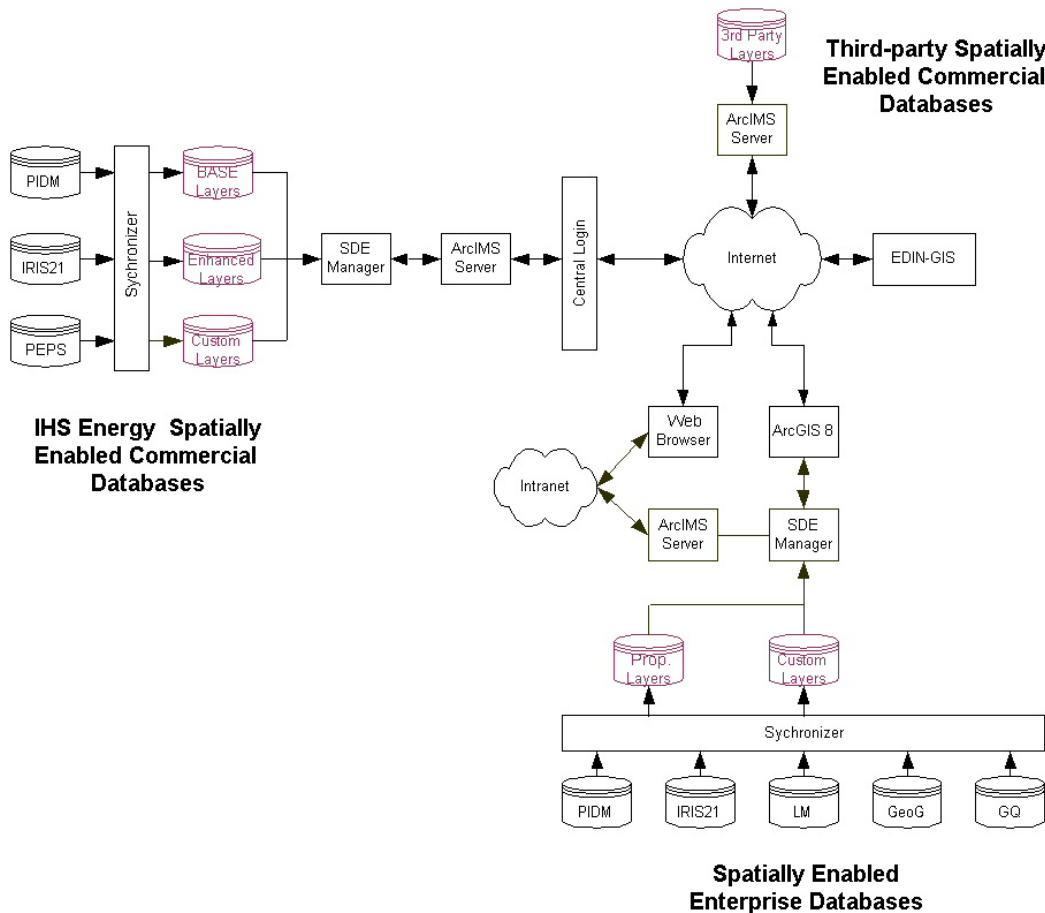


### Online Access

Once the Spatial Layers have been created in an SDE format, they can be accessed in a variety of ways:

- ❖ Across a LAN or WAN, using one of the ArcGIS 8 series of products
- ❖ Across an Intranet, through ArcGIS 8 or a Web browser, using ArcIMS technology
- ❖ Across the Internet, also using ArcIMS technology
- ❖ Through a commercial application such as EDIN-GIS from IHS Energy
- ❖ Any combination of the above

The next diagram illustrates the different ways in which the data can be accessed along with the components of a typical Enterprise system.



In this scenario, a client can use a Web browser or desktop GIS client to access Spatial Layers either from internal sources or from external public and commercial sources. One of the major advantages of this flexibility is that large, complex commercial databases can be efficiently managed at their source by the vendor, thereby guaranteeing access to the most up-to-date information. This eliminates a major burden from the client and should lead to significant cost savings and performance improvements.

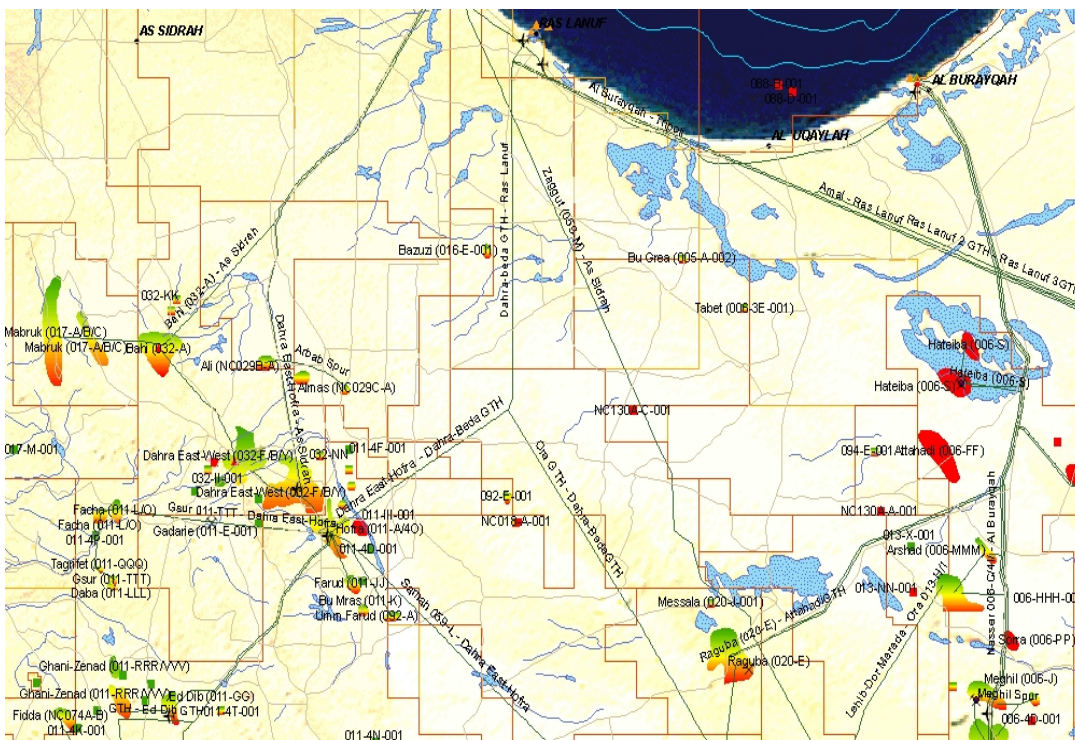
IHS Energy provides access to a whole suite of U.S., Canadian, and International spatial layers through its secure and reliable Denver-based Hosting Center. Clients may choose to subscribe to any set of layers and then access them from anywhere in the World on a 24x7 basis.

## Data Integration

In this section we will present and discuss three levels of Data Integration that build from the Data Access foundation. The level of complexity increases with the increasing level of integration, so information is also included on the underlying technological architecture, as well as services available to help implement Enterprise class solutions.

### Simple Data Integration

A major problem for users is the inability to view all available data at the same time through a single interface. Establishing a common Data Access format provides a foundation for solving this problem. Using standard GIS tools, it is possible to make graphical queries, perform simple attribute queries, browse data lists, and create simple downloads in either a graphical or text-based format. This approach is useful for finding what data is available, either internally or out on the Internet, and then using it to pull together project data for analytical work. Furthermore, enhanced spatial layers can be developed that present a specific view of the underlying data to particular categories of users. In the following example, a series of views of global political risk have been developed from a comprehensive database and application to present users with an easy to comprehend and up-to-date representation.



IHS Energy has developed a comprehensive catalog of enhanced layers that are available online and can be accessed from any application that supports the SDE format. For example, this can be a sophisticated desktop GIS, a thin-client Web browser, or a proprietary application built around map objects. The advantage is that all users have access to the same information through the same back-end technology.

The limitation of this approach, however, is that the user is only able to access attributes directly associated with the layer, and there is no standardization between similar layer types from different sources to permit cross-layer queries.

### **Enhanced Data Integration**

If further integration is desired, it is possible to create the same spatial layer with identical attributes from different data sources for comparison purposes. For example, a client can create a Well spatial layer from a proprietary data source with the same attributes as the base Well layer from IHS Energy. The two layers can then be plotted on top of each other to provide a visual comparison, and then a single query could be run against both layers and the data exported in a common format. This is particularly useful for performing Data Matching between overlapping data sets and for identifying missing or incorrect data.

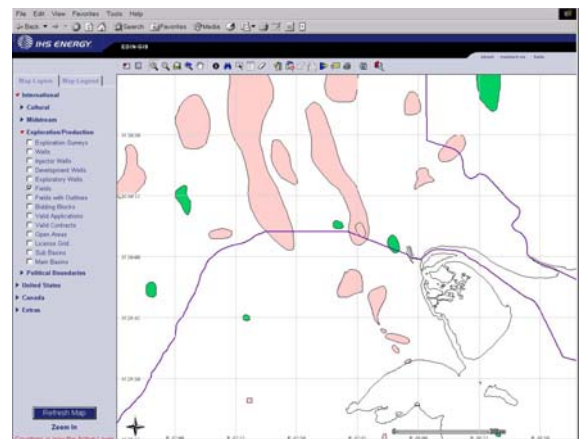
By matching attributes between layers of a similar type, we will have achieved not only common Data Access, but also a straightforward yet effective level of Data Integration.

### **Advanced Data Integration**

The limitation of the previous two levels of data integration is that it is only possible to run queries and perform analysis and exports on the attributes stored or linked to the Spatial Layers through the Business Table. In most cases, there is a wealth of additional information stored in the underlying database that needs to be accessed in order to support effective decision making.

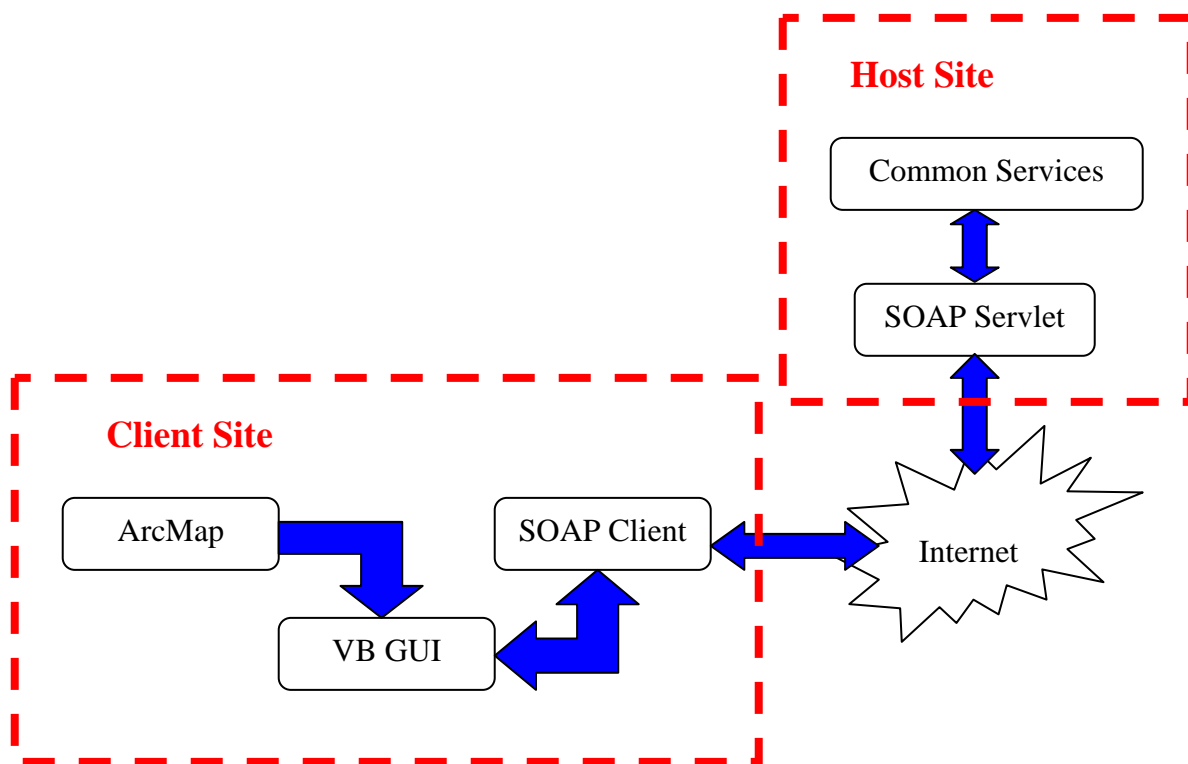
To access this information, many vendors supply turnkey solutions that have proprietary functionality that interacts with the map layers to perform specific functions against the underlying databases. As an example, EDIN-GIS from IHS Energy is a thin-client, Web browser solution that builds upon ArcIMS and SDE technology by accessing the core data models to enable users to:

- ❖ Perform a wide range of standard or ad hoc queries
- ❖ Create custom browse lists of information
- ❖ Generate a wide list of attribute-based reports and exports
- ❖ Generate graphs and charts
- ❖ Create effective maps through additional labeling and symbolizing features
- ❖ Export maps in a variety of different formats through an interface to the FME tool from SafeSoft



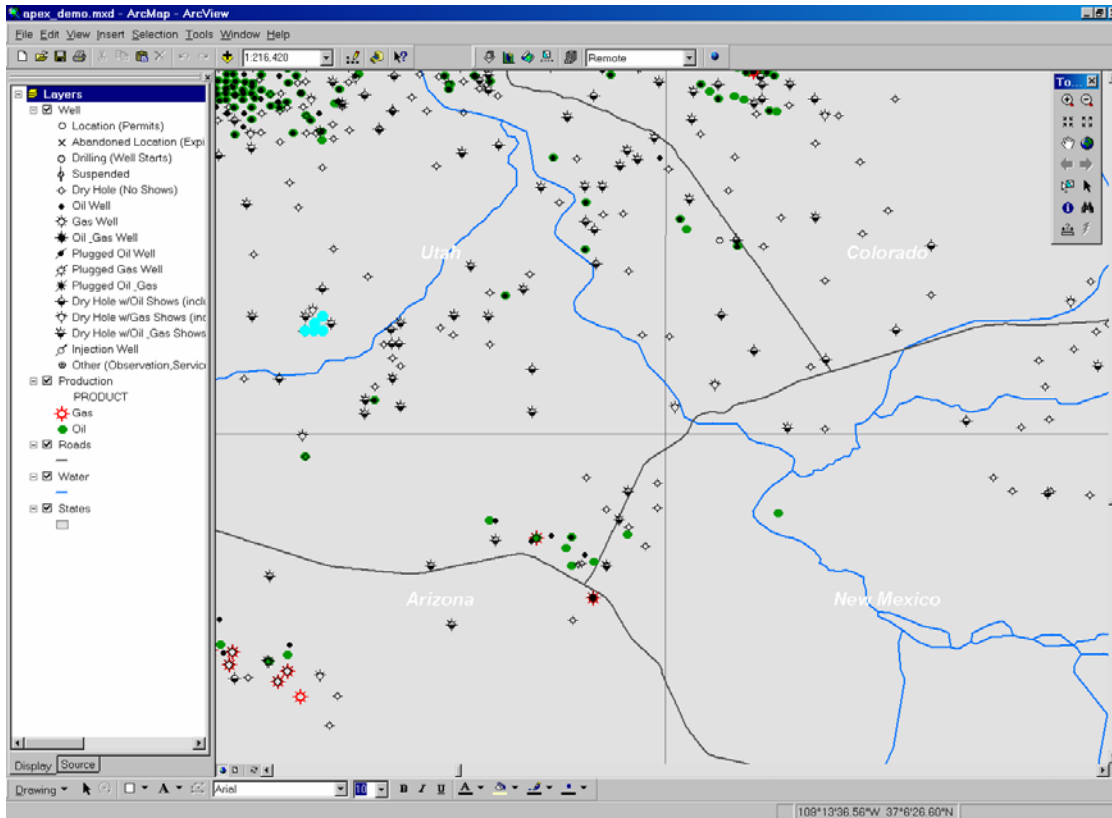
The main issue with turnkey solutions is that while they are generally excellent at accessing and processing in-depth information from the databases that they know about, they are less flexible in pulling information in from proprietary and other vendor data sources. In addition, it is often difficult or impossible to modify the applications themselves without going back to the vendor for custom development work.

One solution to this is to adopt a toolkit approach to development and to standardize the development so that tools from different vendors can be plugged into the same environment where they can interoperate. In the GIS world, ArcGIS 8 provides an environment for common data access as already discussed, but also an environment in which tools (extensions) can be plugged in to enable users to drill deeper into the underlying information. At IHS Energy, we have been investigating the development of a lightweight VB interface to online services so that users can make selections using standard GIS spatial query tools and then execute services such as Reports, Graphs, and Exports, returning the subsequent information as an XML, pdf, or xls file. The basic architecture is illustrated in the next diagram.



Using this approach, we have been able to develop a simple prototype that enables the user to access Spatial Layers and Report, Graph and Export Common Services across the Internet (see next image).

At the client site, the information can be integrated with information pulled from other sources to create consolidated project files that can be used to generate sophisticated maps and reports or be pushed to other applications for further analysis.



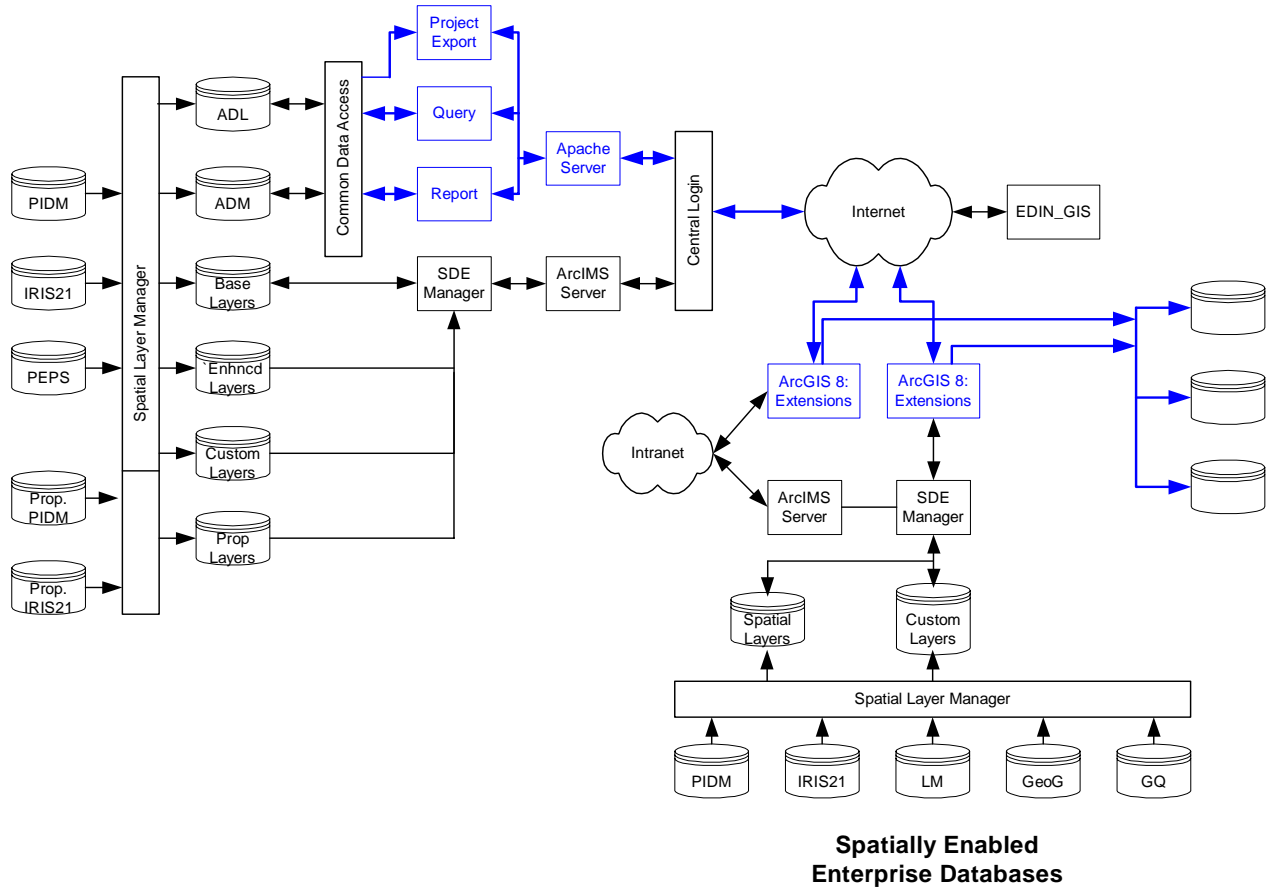
Providing this toolkit completes the Enterprise Data Access and Integration solution. In the configuration illustrated in the next diagram, several options are provided for accessing the same information through different platforms depending upon the level of integration necessary. It should be noted, however, that the technical implementation of this Enterprise class solution requires careful planning to minimize problems with security, performance, and reliability.

## Data Security and Portal Technology

One of the biggest challenges with providing access to online information is to make sure that the interface with the information is seamless and that users see only the data that they are entitled to see. It would be a major impediment to require users to log in to each data source that they wanted to see. Fortunately, there are a number of technologies available that enable security systems to exchange authentication information in the background to eliminate this step. In some cases, vendors are coming together to establish standards that will facilitate this type of security information exchange, the Liberty alliance being but one example of this.

Enabling secure communications opens the door to the implementation of portal technologies. Rather than simply providing a link to vendors' Web sites, it is possible for a vendor to become an active part of the customer's Web site or portal. The advantage is that information of value can be pushed to the client to eliminate them having to sort through volumes of data and numerous

applications. As an example, information on a new bidding round or on new wells added to the database could be pushed down to the client portal as a headline to trigger further investigation.



## Solution Implementation

Clearly analysis, design, development, and deployment of a fully integrated Enterprise GIS solution is no trivial matter. To realize all of the benefits of an advanced Data Access and Integration solution requires a partnership between clients and vendors that extends beyond simple data sharing. It is important to ensure that both vendors and clients set expectations early, and that each partner has the technical expertise and infrastructure necessary to be successful.

To meet this need, IHS Energy has created a TSG that has the technical expertise and industry experience to assist clients with the design, development, implementation, and support of systems that will meet their Data Access and Integration goals.

## Summary

This White Paper has discussed the adoption of a widely accepted spatial data format as a meta-layer for the access of information from disparate and dispersed data sources. Providing a common data access layer establishes a foundation for the integration of information into applications that are used for browse, query and analytical purposes. Several stages of integration can be achieved that involve increasing degrees of design sophistication and implementation. A full solution, as illustrated in the previous diagram, provides access to the same sources of information from a variety of applications that meet the needs of different types of users. Casual users who are interested in the straightforward browse and query of commercial data might be satisfied with a turnkey Web browser solution such as that offered by EDIN-GIS. More GIS-oriented users can view all of the information available through a tool such as ArcGIS 8 and perform data matching or sophisticated map creation, for example. Finally, a customized desktop solution can be created to match user workflows and decision-making processes.

As noted in this paper, a full solution can only be developed as a partnership between clients and vendors. There are numerous issues to be resolved, including Spatial Layer formats, integration levels, common systems architecture, and security. The benefits to companies that have implemented this type of solution, however, are enormous in terms of improved efficiency and improved decision making through access to the most up-to-date and comprehensive sources of data available integrated within a single application.