

WHITE PAPER

Support the Changing Business Needs for Business Intelligence and Performance Management Through Data Warehouse Lifecycle Management

Sponsored by: WhereScape

Dan Vesset

Brian McDonough

January 2007

IDC OPINION

As organizations of all sizes and in all industries place increasing emphasis on business intelligence (BI) and performance management initiatives, demands on data warehousing (DW) professionals continue to grow. Increasing data volumes, number and diversity of users and ongoing user requests for decision support and reporting needs are increasing the complexity of DW projects. At the same time, the success rates of DW initiatives based on low satisfaction levels with poorly designed schema, long project lifecycles and unfulfilled promises remain high. The time, cost and risk associated with managing these broader and deeper demands can be reduced through the use of robust lifecycle management methodologies and supporting software.

DW lifecycle management tools, such as those provided by software vendor WhereScape, can play a pivotal role in supporting the needs of DW architects, data modelers and BI developers in their effort to address the demands of their internal user constituents. The same is true of IT professionals in consulting or system integration companies who need to deliver DW solutions to their clients faster, and increasingly, within pre-established, fixed-budget parameters.

IN THIS WHITE PAPER

In this white paper IDC describes the latest trends in the business analytics market which includes both data warehousing and business intelligence. It highlights the challenges faced by IT or information management departments in addressing the ever changing decision support needs of business users. Furthermore, it introduces WhereScape Software as a provider of DW lifecycle management software, and summarizes the experiences of two organizations with WhereScape RED software.

SITUATION OVERVIEW

Business Analytics Software Market

The business analytic (BA) tool, application and data warehousing platform market includes software for tracking, storing, analyzing, modeling, and delivering data in support of decision-making and reporting processes. The worldwide BA software market reached \$16.6 billion in 2005 and is expected to grow at the compound annual growth rate of 10% over the next 5 years.

The foundation of any business analytic solution is a robust DW platform that supports data integration, data management and calculation functions. Organizations employ different strategies for DW deployments. Some seek to implement an enterprise DW (EDW) strategy while others prefer distributed (but connected) data marts. In a 2005 survey of over 300 DW market participants, 67% indicated that they have a EDW strategy with 33% having a distributed data marts strategy.

Business Analytics Market Trends

Regardless of the DW approach, organizations in all industries and of all sizes must face an environment of increased demand on the IT department by business users and information management processes. The following three variables characterize current market trends:

- ☒ **Growth in data volumes.** Over half (56%) of organizations of all sizes indicate that they have between 2 and 9 source systems feeding their primary DW and 37% indicate that they have more than 10 source systems. According to IDC's recent survey of 750 respondents, 40% of DWs are over 1 TB in total size, with 18% of organizations expecting their largest DW to at least double in size during the next year. The largest percent (37%) expect a 25% to 50% growth rate for their DW.
- ☒ **Shorter decision cycles.** Compounding the complexity of providing end users the most relevant information is the increased demand among employees for more rapid access to the data. As decision cycles shrink, organizations are looking for near-real-time information to be delivered to their decision makers. Already, 13% of organizations indicate that they refresh their DW in real-time. While the use of real-time refreshes is still fairly low, increasingly, daily loads from transactional to analytics systems are becoming the norm.
- ☒ **More end users.** With the massive amounts of data collected, managed and refreshed at a more rapid rate, organizations are seeking to distribute business information to a wider audience of decision makers at all levels. IDC surveys indicate that 70% of organizations are planning to increase the number of internal users of BI tools over the next 12 months. Additionally, 33% of organizations plan to provide direct BI tool access to suppliers, 40% to customers, and 20% to government agencies.

In addition to the above-mentioned trends, it is often the DW-related business issues that act as a bottleneck to potential wider adoption of BI and analytic application. For example,

- ☒ 23% of organizations indicate that faster implementation of requested changes by end users, such as addition of new data sources or new views, will lead to wider adoption of BI/DW solutions.
- ☒ 20% of organizations indicate that more ad-hoc query and analysis functionality rather than static or pre-built views will lead to wider adoption of BI/DW solutions. This implies increasingly complex, dimensional schema development.
- ☒ 43% of organizations indicate that their DW can't be down for more than a few hours without having a severe impact on business operations.

While the demands of end users continue to grow as shown by the above research findings, many DW support groups continue to struggle either due to legacy practices which introduce inefficiencies into the DW development and maintenance process or due to a lack of appropriate software tools.

Data Warehouse Lifecycle Management

Years of DW implementations have made data warehouse lifecycle management (DWLC) a familiar concept for many organizations and BI practitioners. As shown in Table 1, the results from an October 2006 survey of 750 respondents indicate that a sizable percentage of organizations are in either an upgrade or maintenance mode with their DWs. However, many organizations do not have robust DW lifecycle methodologies or supporting software tools.

TABLE 1

Stage in DW Project

Starting a new/additional project in the next 12 months	38%
Upgrading an existing solution to the latest version within 12 months	30%
Maintaining a production solution without any significant change	33%
Not applicable	11%

Source: IDC, 2007

A typical DW process includes the following steps: define, design, develop, test, deploy. In a traditional approach, IT staff gathers end user requirements, designs the DW schema, develops ETL routines and all the DW objects, tests the new system and then deploys it for use by business users. The time frame from requirements definition to deployment varies, but can often stretch into months. Upon receiving access to the DW, business users begin to query the DW for answers to their daily questions by conducting analysis using BI tools. As some of their questions are

addressed, new questions arise. At the same time, various business events occur (reorganizations, product hierarchy changes, mergers, etc.) that change the types of views and information end-users require.

Several shortcomings are apparent in this process or lack thereof:

Shortcomings of Typical Data Warehousing Process

The first shortcoming is that traditional 'waterfall' development methodologies gather hypothetical requirements. Users with little understanding of the available data make requests or demands; IT personnel transcribe those requests, and requirements are produced. When those requirements are taken into design, and later implementation, so-called compromises are made, often unilaterally, by BI practitioners. Frequently the resulting data warehouses have substantial anomalies to the outcomes imagined by end-users months before during the hypothetical requirements gathering activities.

The second shortcoming of such a project-based approach to DW is its lack of flexibility in responding to ongoing business users' change requests. The lifecycle of a DW should not end with the initial deployment, but instead continue as the business requirements evolve.

The third resulting shortcoming is the use of inappropriate software tools to support the DW lifecycle management process.

Data Warehousing Tools

There are three primary tool categories used by organizations in developing data warehouses:

- Generic application development tools used for internal development or so-called homegrown solutions. These may include SQL and its database-specific variants, Cobol, or other development languages.
- Traditional extract, transform, and load (ETL) tools
- Specialty DW lifecycle management tools

Each tool set has its pros and cons. For example, in many cases home-grown solutions have been developed over time with low software costs and function as expected, but their total cost of ownership is high due to the high level of manual effort that goes into development, maintenance and upgrading. Homegrown systems also have a tendency to have poor documentation.

Many traditional ETL solutions overcome some of the challenges of internally built systems. Traditional ETL tools are powerful and appropriate for various types of data integration projects including DW generation. However, over the years the leading ETL tool vendors have expanded their offerings to include suites of data integration software. While these tools are in wide use, they carry, in general, a high price tag due to the breadth of functionality. If an organization is looking just for the DW lifecycle management software, it may not need to purchase other related tools or modules.

Specialty DW lifecycle management tools are focused exclusively on the creation and management of DWs and data marts. Their specialization results generally in a lower price point than that of traditional ETL tools. Focused functionality also makes such tools more flexible in supporting the change requests of business users.

Although all three types of tools can and are being used in DW environments, this white paper focuses on the specialty DW lifecycle management category.

Data Warehouse Development and Maintenance Challenges

Unlike many other IT projects, DW initiatives are deployed incrementally rather than in a single deployment instance that goes unchanged for long periods. Yet, as survey data in Table 2 suggests, there is significant room for improvement in the success rates of DWs.

TABLE 2

Success Rates of Data Warehouse Projects

Successful	39%
Needs improvement	49%
Not applicable	12%

Source: IDC, 2007

Some project failures result from unsatisfied business users that realized their true requirements after interacting with the initial deployment. In a 2005 study conducted by IDC, IT and business users indicated the following top challenges associated with their DW initiatives (see Table 3).

TABLE 3

Top Data Warehousing Challenges

IT related challenges	Business related challenges
Data Integration	Finding budget
Keeping up with end user change requests	Getting business users to share data
Keeping up with business change	Deciding on key performance indicators
Managing metadata	Gathering user requirement

Source: IDC, 2007

These results show that defining user needs and designing the solution are the key challenges facing IT departments. These and other challenges are relevant not only during the first development and deployment project, but are ongoing challenges throughout the lifecycle of a DW, which can last for decades. For other application development efforts user change requests are often collected and kept in a queue for implementation during pre-set time intervals, which are based on IT release schedules rather than business users' needs. Given the user sentiments shown through IDC research and described in the previous sections of this paper, this approach is unacceptable in DW projects.

Whenever a change is requested, typical IT tasks include building or changing tables, rewriting ETL load routines, enhancing or changing calculations, editing schemas, or rebuilding cubes, testing the changes and deploying them to business users. IT professionals, including DW architects, data modelers, DBAs, cube and report developers need to expedite this process as much as possible to satisfy business users. They can do so by adopting best practices in application development that are relevant to DW. Such best practices include presenting business analysts with incremental prototypes to show them their own data prototypes prior to a production version of the solution. Instead, many DW initiatives are approached as "one-off" projects rather than ongoing and constantly changing processes.

Scoping the Need for DW Life Cycle Management

In evaluating the need for DW lifecycle management software, organizations should consider the efficiency and quality of their existing DW development and maintenance process.

If the efficiency of internal operations can be improved and if internal use satisfaction metrics are anywhere near the survey results presented in this paper, a DW lifecycle management solution should be considered as part of the overall DW software evaluation process.

Some of the features and functionality that should be considered include:

- A single tool for supporting all steps in the DW lifecycle
- Rapid prototyping to enable iterative development and deployment
- Automated documentation to mitigate risk of staff turnover
- Support for multiple operating and database platforms

WHEREscape SOFTWARE, LTD.

Founded in 2002, WhereScape is a vendor of software for data warehouse (DW) lifecycle management. The company has offices in U.S., Europe and Asia Pacific and as of the writing of this white paper has over 150 customers for its flagship product WhereScape RED, including companies such as Gillette, Wells Fargo, Bridgestone Firestone, Vodafone and Air New Zealand.

WhereScape RED provides integrated DW lifecycle management functionality for Microsoft SQL Server, Oracle and Teradata deployments. The company's go-to-market approach includes direct sales to clients, OEM relationships with transactional application vendors, and a network of resellers and consulting partners in the geographic regions in which it operates. WhereScape also provides a freely downloadable working version of its software, which has been an effective method for allowing prospective clients to test-drive the software prior to production deployment, by building fully-functional small-scale warehouses and marts with WhereScape RED prior to purchase.

Product Offering

As shown in Figure 1, WhereScape RED is a metadata driven DW lifecycle management environment.

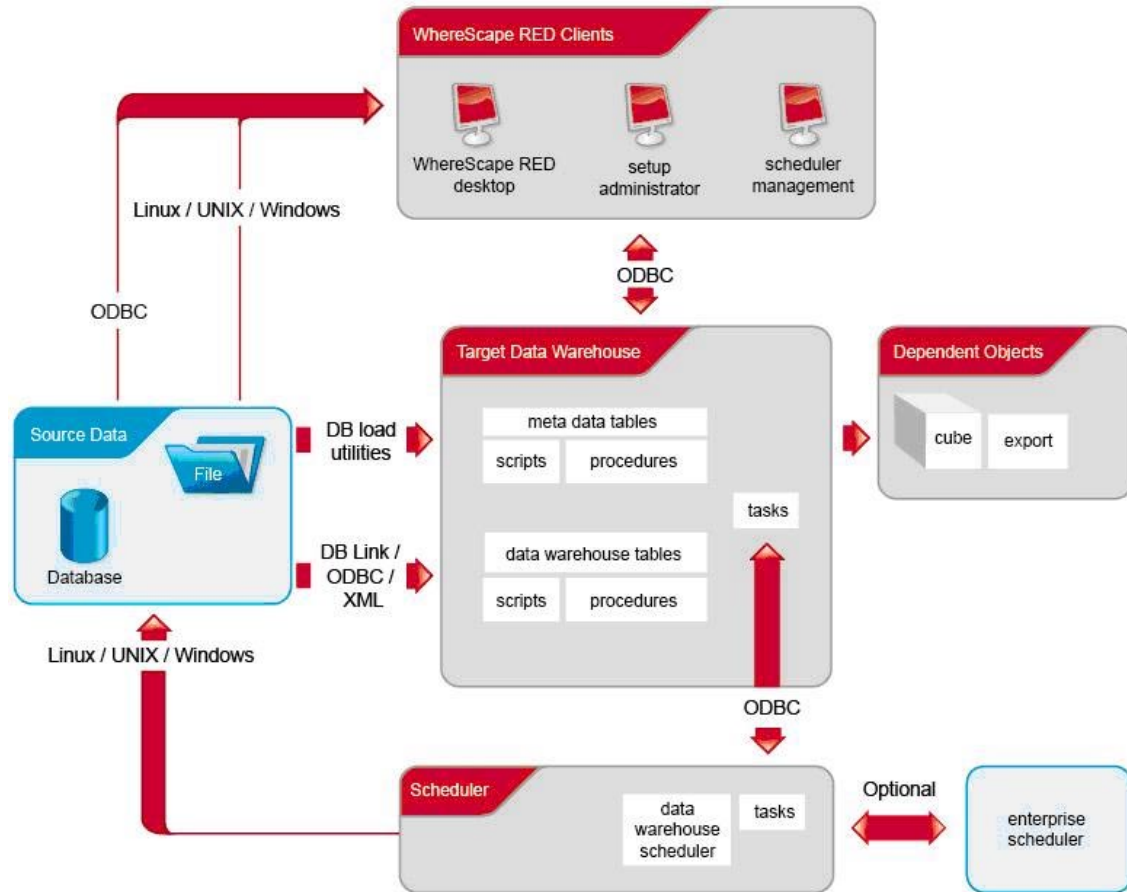
WhereScape RED includes 5 integrated components that support all stages of the DW lifecycle process:

- Source data extraction and loading facilities (native database connectivity, ODBC, XML, flat file)
- WhereScape RED Applications
 - Administrator – is a developer's interface for creating repositories within the data warehouse and maintaining schedulers.
 - Integrated Development Environment (IDE) – is a desktop interface for designing and managing the data warehouse
 - Scheduler – is an optional interface that is used primarily by administrators for job status information.
- Repository – housed within the target DW, the RED Repository is a set of database tables that are installed in each target DW environment. Metadata tables can be accessed by business intelligence tools for query, reporting and analysis purposes. The data warehouse objects created by WhereScape RED are standard database objects such as tables, views, scripts, procedures and indexes. They are always specific to the details of the target database's SQL or data manipulation dialect: scripts in Teradata can be MultiLoad or FastLoad, procedures are PL/SQL in Oracle and Transact SQL in SQL Server. RED uses the database as the data transformation engine and also stores all metadata about the objects created in the target DW and about the dependent objects within the repository.

- ☒ Scheduler – polls metadata to identify tasks that can be run. It also manages the workflow of dependencies that may exist among jobs or groups of tasks. Tasks can run in the target database or on a separate scheduler platform and initiate loads or procedures specific to databases being used. The scheduler supports multi-platform scheduling and can be called from an existing enterprise scheduler.
- ☒ Dependent Objects – located outside the RED repository, dependent objects include OLAP cubes or exports for other downstream systems whose metadata is stored within the RED repository. The scheduler also manages these objects.

FIGURE 1

WhereScape RED Product Portfolio and Logical Architecture Diagram



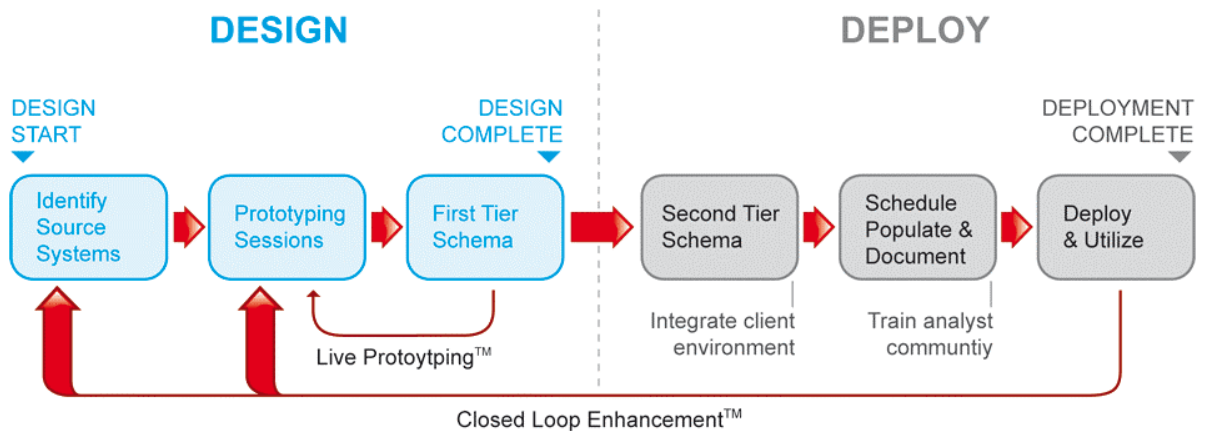
Source: WhereScape and IDC, 2007

Addressing the DW Lifecycle Management Process with WhereScape RED

WhereScape RED provides support for all steps of a DW lifecycle. Figure 2 represents a diagram of WhereScape's methodology that supports both the developers and business users in their collaborative efforts from the definition stage to deployment stage through a series of iterative steps and feedback loops.

FIGURE 2

WhereScape DW Lifecycle Management Methodology



Source: WhereScape and IDC, 2007

Some of the key strengths of the WhereScape RED software include:

Rapid Prototyping

WhereScape itself describes its software as enabling developers to "fail fast and fail early." With WhereScape RED, developers can create prototypes of candidate production schema, populate it immediately with source data (familiar to business users), and adjust the schema in response to business users' feedback as needed during joint design-prototype-iterate sessions – all within days and in some cases hours. This practice allows business users to see, in real time, the translation of their requests into working data warehouses and/or data marts, and helps BI/DW practitioners make appropriate accommodations while they themselves get comfortable, incrementally, with those accommodations.

Self-documentation

Self-documentation was highlighted as one of the key features of WhereScape RED during interviews with WhereScape customers. Within RED, metadata is created automatically while an architect interacts with the RED desktop interface during the various stages of the data warehousing process. Metadata and other documentation can also be edited or added manually on an as needed basis.

Single Source for Metadata

WhereScape RED stores metadata not only about source systems and the data warehouse objects, but also about the external dependent objects such as OLAP cubes as well as ETL routines. Having a single source for all metadata related to the various steps in DW lifecycle management enables administrators, architects and developers to manage and access this metadata more efficiently than if it was distributed in individual repositories of the component tools.

Retrofitting

Users of WhereScape RED can also reuse existing tables designed with third party DW modeling tools. In this case the software automatically recognizes that RED was not used as the design tool and retrofits existing DW tables into the RED Repository. This process has been identified as a considerable time saver at organizations, that for various reasons, may chose to work with another modeling tool. This is an example of WhereScape ensuring its product's compatibility with existing data warehousing tools and environments.

WhereScape Within an Existing DW Environment

Data warehousing is not a new market and, as is often the case, organizations have accumulated different tools for the different steps of DW development and deployment. Therefore, the flexibility of working within an existing DW environment which is seamless with other DW tools are important characteristics of WhereScape RED. Other DW tools may include generic application development software or purpose built DW software. As organizations evaluate their needs for a robust DW lifecycle management solution, they may find that wholesale replacement of all existing DW tools is not feasible for cost, risk management, or staffing reasons.

In some cases organizations will use WhereScape RED as a "light-weight" ETL tool for extracting, transforming and loading data into a data warehouse. In other cases, organizations will use RED in conjunction with a third party ETL tool, of which there are several options available in the market.

Retrofitting, highlighted in the previous section of this report is one example of the requirement to work with other existing DW tools. Another is the ability of WhereScape to support various target database environments such as Microsoft SQL Server, Teradata or Oracle.

As an independent software vendor, WhereScape has also ensured that its software can be deployed on both Windows and Unix operating environments – the two platforms that have the largest share of the DW market.

CHALLENGES/OPPORTUNITIES

The DW software market, as most other software markets, is composed of a few large vendors with broad portfolios of products and numerous smaller vendors who specialize in specific niches of the market. The specialization enables smaller vendors to provide functionality targeting a specific function. This market dynamic is common not only in the software market and is unlikely to change.

The opportunity and the challenge for WhereScape will be to stay firmly focused on the needs of its core user base. As the DW and BI markets expand in an environment of a shortage of skilled labor, tools that can improve the efficiency of DW lifecycle management are likely to be valued by client organizations. However, WhereScape will need to continue to enhance and effectively differentiate its product functionality and services while expanding its client base and partner network.

CONCLUSION

IT departments should change their DW practices from project- to process-based management to address the ever-changing BI and performance management needs of business users. DW lifecycle management methodology should take into account the need for rapid prototyping to mitigate the risk of costly DW schema definition, design and development work.

DW professionals must adapt their techniques to the reality that business users will never be able to provide precise, fixed and accurate requirements that can be developed into a successful DW on the first attempt. The reality of constant business change, such as corporate reorganizations, product hierarchy changes, mergers & acquisitions, and entry into to a new geographic region must be met with flexible IT systems of which the DW platform plays an important role as the basis for BI and performance management solutions. A robust DW lifecycle management tool can enable such flexibility.

When evaluating DW tools, organizations should consider not only what database, or BI software to use, but also the need to have DW lifecycle management software in the toolbox of its developers. One of the vendors to consider for this purpose is WhereScape Software with its RED DW lifecycle management software.

CASE STUDIES

Air New Zealand

Air New Zealand is an international and domestic airline group which provides air passenger and cargo transport services within New Zealand, as well as to and from Australia, the South West Pacific, Asia, North America and the United Kingdom.

In 1999, the company embarked on a data warehousing project. The company chose a distributed data mart architecture and has maintained this strategy to this day. Today, data marts support all aspects of Air New Zealand's business from sales and marketing to revenue and loyalty management to operations and HR.

Air New Zealand's Information Management team has gone through several stages in supporting the company's data warehousing platform. Initially, to develop its Oracle RDBMS-based data marts, the developers utilized internally built components using PL/SQL. The company later acquired a commercial ETL tool, which is now used in conjunction with WhereScape RED, which was purchased in mid 2005.

While the Information Management team appreciated the power and breadth of functionality of the ETL tool, it also felt that because of these characteristics the ETL tool was too complicated to use, resulting in developer errors, and not flexible enough for prototyping, developing and deploying data marts.

Instead, Air New Zealand was looking for a tool that would be better at building dimensional models. In WhereScape RED the company found a DW tool that was easy to use and met their expectations. "WhereScape RED cuts the time to build models significantly," said Colin Williams, Information Management team lead, adding, "While software will never eliminate the process of requirements gathering, we were able to produce the first data mart prototype with WhereScape RED in just 3 days."

Today, the Information Management team receives regular change requests from business users and is able to react to them quickly by using a DW lifecycle management approach as supported by the WhereScape RED software.

Air New Zealand has also taken full advantage of the software's self-documentation feature – to the point where such documentation has been used to populate PowerPoint slides for training purposes.

While the company did not perform a formal ROI analysis, it estimates that it is able to reduce DW development and management costs by up to 90% through a combination of lower labor costs (due to time saving in the development process) and lower software costs.

CIBER, Inc.

CIBER, Inc. is an international system integration consultancy. The company's global delivery services are focused on custom and enterprise resource planning (ERP) package environments. Founded in 1974 and headquartered in Greenwood Village, Colorado, the company has 8,000 employees and generates annual revenue of approximately \$1 billion delivered from 84 offices in 18 countries.

CIBER's first DW initiative was launched in 2005 with the primary, initial objective of changing the existing weekly reporting process by improving its accuracy and timeliness. The reporting solution had been used by CIBER operations, corporate management and the accounting department to forecast revenue and gross profit. The legacy reporting solution displayed data extracted directly from operational systems.

CIBER was looking for a DW that would consolidate data from disparate operational systems and make the information available to all decision makers. The DW initiative was sponsored by the CFO and COO, which provided the IS applications group with the necessary internal support.

CIBER adopted an EDW strategy. The company uses Microsoft SQL Server for its DW database and has also deployed SQL Server Analysis Services and Reporting Services as the BI components. Business users access reports through a web based interface and CIBER is evaluating the potential use of third party, end-user BI tools.

CIBER's evaluation of WhereScape RED DW lifecycle management software began with a proof of concept under the leadership of the company's enterprise architect, who had experience with the product at Wells Fargo. CIBER's IS team was pleased with the functionality of the software. Ahmad Malik, Director of IS Applications, elaborated "we were especially pleased with the auto-documentation and ETL routine building provided by WhereScape RED and the ease of building star schemas." These features differentiated WhereScape from an ETL tool that the team briefly evaluated.

CIBER was also impressed with WhereScape's customer service. The company provided CIBER with ample support during the initial stages of the DW process. When CIBER requested specific functionality that did not exist in the base product, WhereScape responded quickly by enhancing the product. Mr. Malik also noted that even though DW was a new initiative for CIBER, using WhereScape RED did not require any formal training for his staff of developers and database administrators. The ease of use of the product was a key variable in its uptake.

The DW process at CIBER included requirements gathering with business users to identify relevant subject areas and source system tables. Then the IS team addressed data profiling and data quality issues, built the DW model (using another software tool that is also used to manage the overall enterprise data model - not just for DW). CIBER's IS team then used the Retrofit feature in WhereScape RED to drag the staging tables into WhereScape RED by using a wizard provided within the software. As Mr. Malik pointed out "doing this process manually would have taken us significantly more time and resources."

In addition, CIBER uses WhereScape RED to build cubes based on Microsoft Analysis Services. Additionally, a big selling point for CIBER was the ability of WhereScape RED to call SQL Server Integration Services from their scheduler. In general, the smooth fit of WhereScape RED into the existing DW environment was well received by the IS team.

Today the DW-based reporting solution caters to an initial set of users within the several business groups including executive management, financial analysts, operations managers (upper management, area directors, and practice leaders), plus accounting and IS personnel. At the same time the DW is undergoing some fine tuning to be fully adopted by a broader audience of business users and its scope is expected to change in later phases to include additional data sources (e.g. general ledger) and user interfaces (dashboards).

The benefit of moving to a DW-based reporting solution as opposed to reporting directly from the source systems included consolidation and integration of data from multiple applications (e.g. PeopleSoft, Costpoint, SAP, Deltek) and availability of new functionality such as drill down capability, ability to produce the report for any period of time/date range, and ability to add new metrics and calculations (e.g. an estimated gross profit % for each P&L).

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2007 IDC. Reproduction without written permission is completely forbidden.