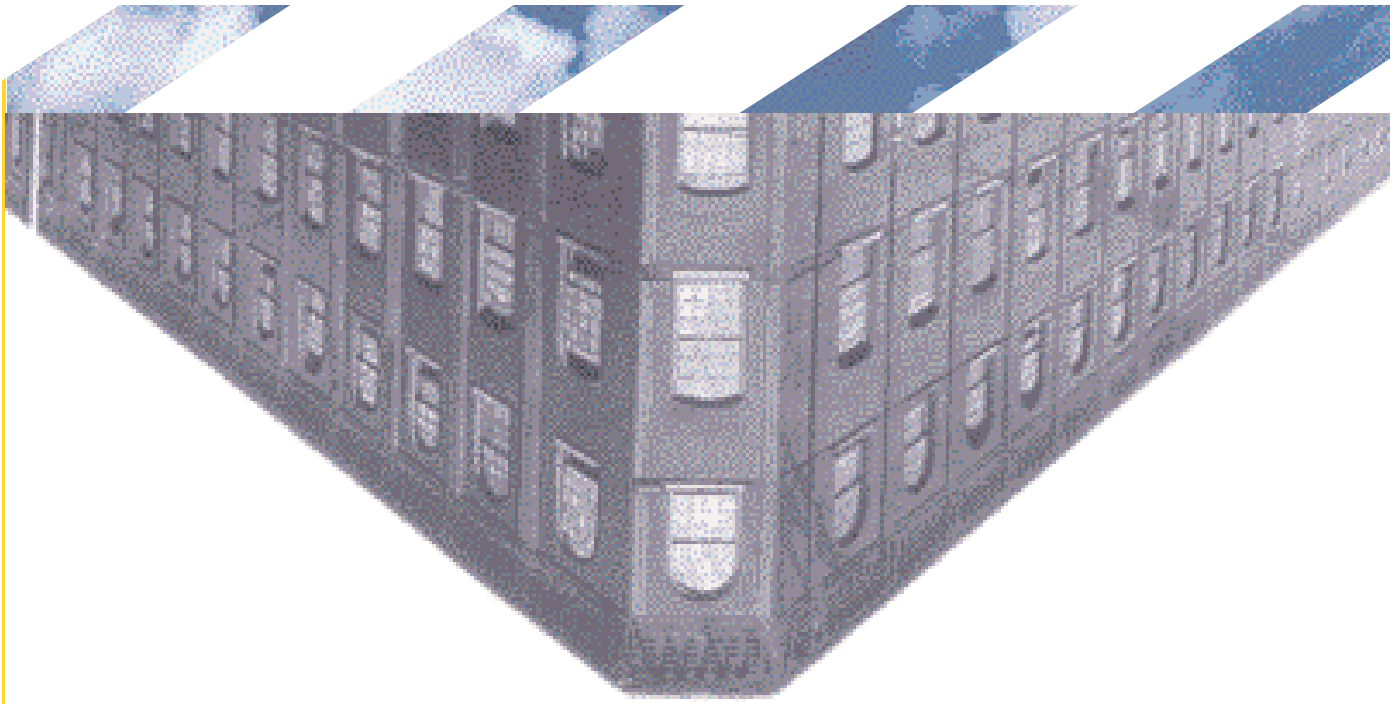
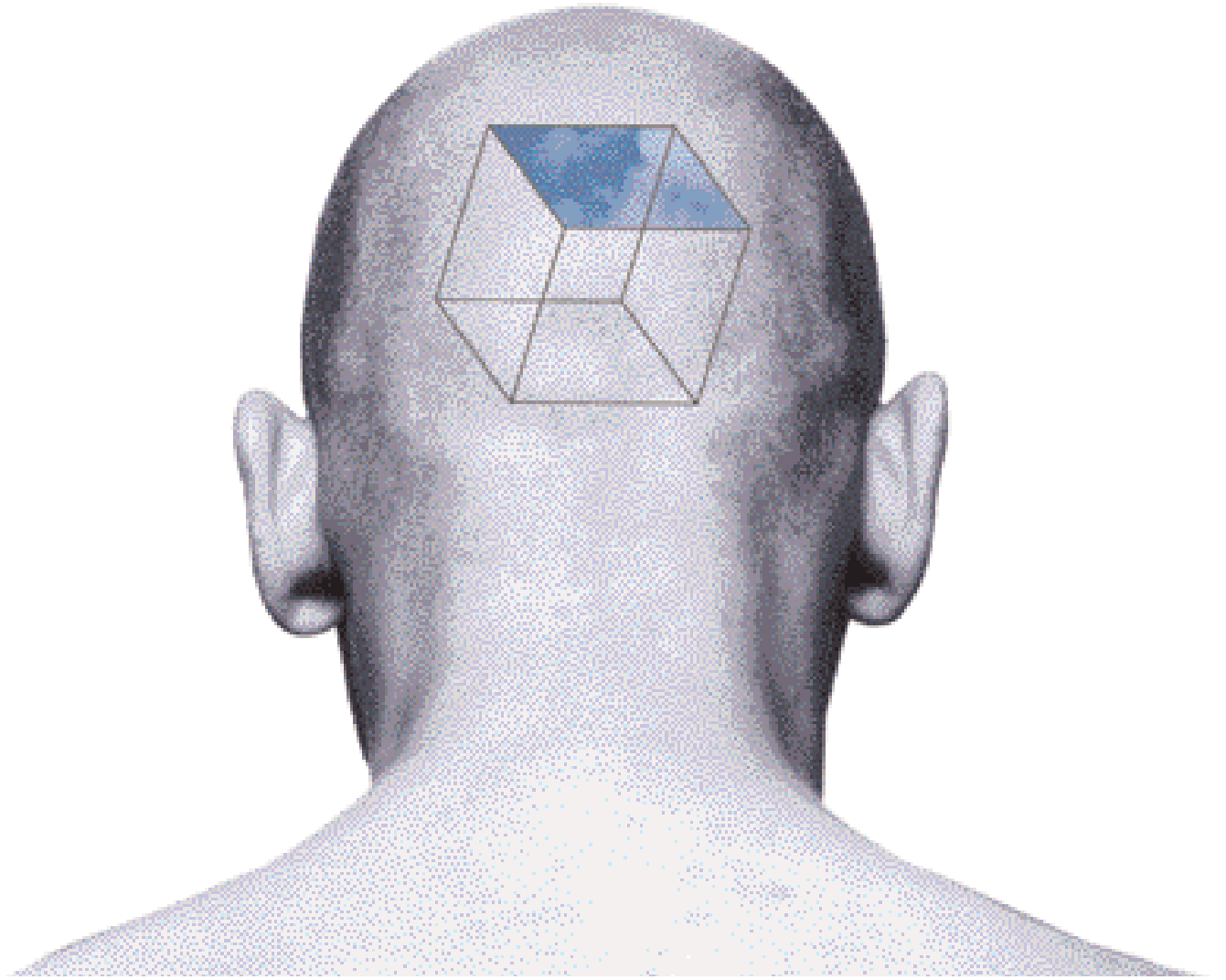
Integrated and smarter than ever, Oracle9i's new business intelligence technologies help your company make sense of customer data and use it to stay ahead. BY KELLI WISETH

<find>meaning

Not long ago, data warehousing and decision-support initiatives fell into the "nice to have" category of corporate IT strategies. That's no longer the case, with organizations large and small alike vying for the lead in the race for information. Business intelligence systems are competitive necessities in the arsenal of the successful e-business. According to Jagdish Mirani, senior product marketing manager for data warehousing at Oracle, "whoever is first in getting a handle on customer buying patterns, buying behaviors, discovering how internal processes are working or not working, and which channels of execution and delivery are most effective—and why—will be the leader in the marketplace."

Data warehousing and business intelligence initiatives serve two critical business needs: know your customer well, and know how well your





Oracle's End-to-End E-business Intelligence Platform

The latest releases of Oracle9i Database, Oracle9i Application Server, and Oracle9i Developer Suite provide an end-to-end software stack designed to let organizations build and deploy robust enterprise business intelligence systems. Some key new features that will benefit data warehousing and business intelligence implementers include improvements in reliability, scalability, and performance in the database server—for example, Oracle Real Application Clusters.

When it comes to business intelligence, however, it takes more than just muscle. Oracle9i Database now includes much of the underlying technology that enables business analysts and strategists to effectively analyze operations, learn about customers, and use that information to plan for the future. For instance, with the OLAP Option for Oracle9i, the Oracle database has been taught how to perform calculations previously available only in multidimensional databases. Features include an API for developers, scalable data store, integrated metadata, summary management, SQL analytic functions, and syntax extensions. The OLAP Option also includes functionality from Oracle Express Server. (Express Server is converging with the OLAP Option, and Express Server will no longer be a stand-alone product. See http://www.oracle.com/ip/analyze/warehouse/collateral/olap_sod.html for more information.)

Having OLAP functionality directly in the Oracle database makes it easier to manage and faster to implement OLAP applications—you needn't replicate data to an ancillary server, for example. Because OLAP Option is hosted on Oracle9i Database, it is also highly scalable and supports very large data sets, dimensions, and numbers of users. Developers can make use of the OLAP API by using Business Intelligence Beans (BI Beans), available in Oracle9i Developer Suite (JDeveloper). Oracle BI Beans include presentation Beans, data Beans, and persistent services Beans. A unique query builder Bean provides applications with a very simple, graphical user interface for defining queries.

The BI Beans components all have the standard JavaBeans architecture—they use the JavaBeans API from Sun—so developers can use them with any Java-compliant IDE. However, Oracle JDeveloper has additional functionality designed to improve developer productivity. For example, JDeveloper provides live access to the business data at design time, so that developers can change the layout and format of graphs and reports while they're designing or as they edit the queries.

THE POWER OF EMBEDDED DATA MINING >>> In addition to embedded OLAP, Oracle9i Database also now has embedded data mining algorithms. The first release provides two algorithms aimed at finding patterns and making predictions based on the patterns; additional data mining algorithms will be added in future releases. These data mining algorithms are exposed by means of an API that Java developers can use to write data mining or predictive applications.

TOOLS FOR HEAVY LIFTING >>> Arguably one of the more mundane aspects of data warehousing and business intelligence but nonetheless one of the most important is

organization is doing. According to "Harnessing Customer Information for Strategic Advantage: Technical Challenges and Business Solutions," an industry study conducted by the Data Warehousing Institute (www.dw-institute.com) in 2000, "Data warehousing and business intelligence technologies ... enable companies to collect, integrate, and analyze customer information so they can understand their customers' behaviors and intentions. Armed with this knowledge, companies can then optimize interactions with customers across all channels, and subsequently, improve revenues and profits."

Companies like Best Buy, NetZero, ON Semiconductor, and Wells Fargo are doing just that, and Oracle technology plays a key role.

KNOWING YOUR CUSTOMERS NetZero, a new model of ISP, has built a business out of using data mining technology to analyze massive volumes of consumer data from the internet and other sources. The company

the entire extraction, transformation, loading (ETL) process. ETL tools at the focal point of a typical warehouse staging area must bring together operational systems, business rules, transformations, enterprise models, and business intelligence classifications. Metadata is the linchpin that holds all the elements of decision support and business intelligence together. Until there was a common standard—the Object Management Group's (OMG) Common Warehouse Metadata (CWM)—each tool and each data source typically had its own proprietary mechanism for defining, storing, and managing the metadata.

Oracle Warehouse Builder (OWB), part of Oracle9i Developer Suite, is a unified ETL and metadata management tool. Metadata is stored in the OWB repository using the CWM standard and can be exported as XMI (Extensible Markup Language Metadata Interchange) formatted data.

The new release of OWB includes some unique features that simplify the process of building and maintaining a data warehouse. For instance, the OWB Mapping Editor lets you map source columns to target columns and define transformations easily, using a graphical user interface. Maintenance of a data warehouse is difficult at best, considering the iterative nature of data warehouse applications, so the life cycle management capabilities of OWB can be a boon to database administrators everywhere. For example, metadata reconciliation keeps all metadata definitions across source and targets in sync.

UNDERSTANDING MADE ACCESSIBLE >>>Of course, all the data in the world won't do you much good if you can't get to it quickly and easily. Oracle9iAS Discoverer is easier to use than previous versions and provides administrators, power users, and casual users a wide range of features designed to simplify ad hoc querying and speed up obtaining results. One key improvement is integration with Oracle9iAS Login Server, a single sign-on solution that provides one log-on for access to all Oracle9iAS systems, including Discoverer and Oracle Portal. Discoverer has also been more tightly integrated with Portal, and users can publish Discoverer results to a portlet. For performance improvement, Oracle9iAS Discoverer has also been integrated with Web Cache, for faster response to ad hoc queries.

Furthermore, Oracle9iAS Discoverer is integrated with OWB metadata so administrators can easily generate end-user layers (EULs) from metadata in OWB. This also provides the ability to drill down and see the source of the information behind a Discoverer query, so business analysts can be assured that the query is in fact providing the appropriate information.

Oracle Discoverer also leverages E-Business Suite technology, including security and flexfields. Oracle E-Business Suite in turn leverages the embedded data mining and embedded OLAP functionality in Oracle9i, as well as the metadata. In addition, Oracle Business Intelligence System 11i is an internet-ready performance management system that provides a framework for corporate performance management and continuous process improvement. Oracle Business Intelligence System can seamlessly analyze the back office and front office data provided by the Oracle E-Business Suite.

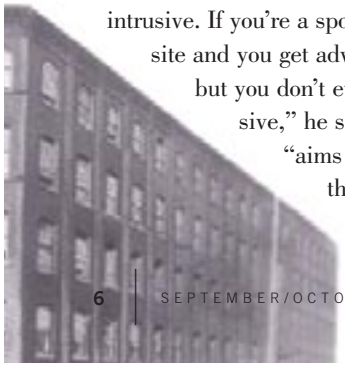
is redefining the internet access model by creating a service funded by advertising, marketing, and commerce arrangements, rather than by internet access fees. NetZero's CyberTarget division offers marketers and advertisers mass-scale, online market research and measurement services, using its own patented and patent-pending analytical tools to mine its Oracle data warehouse, which has amassed volumes of consumer internet data.

As NetZero subscribers browse the internet, ZeroPort, a patented Java application that runs on the browser's machine, collects the clickstream data and sends it to NetZero every 10 minutes. The information includes URLs visited, ads displayed, and ads the user has clicked on. The data is loaded into transactional data-

bases at various points around the internet and consolidated into an operational data store. From the operational data store, data from the user sessions and some other sources is all brought together in the data warehouse, which currently contains about 6 terabytes of data. Once in the warehouse, the data is summarized and analyzed.

Data mining finds hidden patterns among existing data that marketers, research analysts, risk analysts, and strategic planners can then use in the context of new information to do things such as predict outcome under certain conditions. Financial institutions use data mining to facilitate processes such as risk analysis, loan approval, and up-selling or cross-selling products based on patterns observed with existing customers. CyberTarget's raison d'etre is to mine and analyze the data to predict behavior so that the consumer browsing the internet at any given time will be presented only with content that will be of interest. According to Bob Hammer, director of CyberTarget operations for NetZero, the 6-terabyte Oracle data warehouse brings NetZero the ability to provide relevant content to any subscriber.

"If it's not relevant, content is going to be seen as intrusive. If you're a sports enthusiast visiting a sports site and you get advertising for mortgage banking—but you don't even own a home—that's intrusive," he says. Hammer says CyberTarget "aims at making sense of the data in the warehouse and providing value for advertisers."



SNAPSHOT

Best Buy

<<<In the early 1980s, a tornado wiped out what was then the Sound of Music chain of stereo component retail stores, founded in 1966. Founder Richard M. Schulze held a warehouse-like sale, and the concept was so successful that the state patrol had to manage traffic lining up to reach the store from as far as 20 miles away. Schulze went on to succeed with other innovative retailing concepts; the company became Best Buy in 1983 and went public in 1985. In 1989 Best Buy introduced a new retail format placing all inventory on the sales floor with noncommissioned product specialists. Today, Best Buy Co., Inc., (NYSE: BBY) is a large-volume specialty retailer of consumer electronics, personal computers, entertainment software, and appliances.

HARDWARE

<Sun E10000 running five separate domains

SOFTWARE & SERVICES

<Sun Solaris 2.8
<Oracle8 Release 8.1.5.0
<Oracle Recovery Manager
<Veritas Remote Backup
<Veritas Volume Manager
<Maestro Scheduling System
<MicroStrategy TKTK
<Informatica TKTK

Last year Best Buy's earnings rose 60 percent, to US$347 million, and revenues grew by more than 20 percent. Today, there are more than 350 Best Buy (www.bestbuy.com) stores in 39 states, and Eden Prairie, Minnesota-based Best Buy is on track to have 550 stores nationwide by 2004. The company operates retail stores and commercial Web sites under the names Best Buy (BestBuy.com), Magnolia Hi-Fi, Media Play (MediaPlay.com), On Cue (OnCue.com), Sam Goody (SamGoody.com), and Suncoast (Suncoast.com). The company reaches consumers through nearly 2,000 retail stores in the United States, in Puerto Rico, and in the U.S. Virgin Islands. Best Buy uses information about the business and its customers to tailor the product mix to its customer base, cycle its inventory for optimum results, and respond quickly to customer satisfaction metrics. At the center of this strategic information is the Oracle data warehouse, which consolidates information from about 350 different source feeds.

Best Buy, a Midwest-based specialty retailer, also uses an Oracle data warehouse to analyze customer behavior, but in this case the company uses the information to directly manage its own business operations. The ability to see precisely where their sales are moving and then react immediately is paramount to Best Buy's success. Sales from every store—more than 500 of them—are uploaded nightly to the 3.3-terabyte data warehouse (6.6-terabyte with mirroring), and by the next morning a complete picture of the entire company's sales is available to the key managers. According to Jeff Skochil, delivery project manager (Data Warehouse) for Best Buy, "The data warehouse gives visibility into the operation of the business and the trend of the business." Skochil credits the data warehouse with con-

tributing to Best Buys' high inventory turn rate in the last year, up to 7 from 4.6 just three years ago. "We're very efficient, not only with the volumes that we sell but also how much inventory comes in and how long it sits."

Best Buy uses the data warehouse to report sales information and for critical decision-making across the enterprise. Best Buy manages business performance using a number of financial metrics, including revenue and operating margin, in conjunction with trend analysis based on sales, inventory, pricing, and advertising. Actual performance of each store is taken into account, not only in terms of sales volume and type of goods sold but also in terms of store ranking and new sales opportunities at individual stores.

KNOWING YOUR OWN BUSINESS Customer relationship management and other marketing, sales, or customer-centric initiatives are a good reason for implementing business intelligence tools and a data warehouse, but there are other reasons as well. One result of the flattened organizational structure that's emerged in the last decade is that more people in every organization are decision-makers. As more people become decision-makers, easy-to-use yet powerful business intelligence tools and a consistent data store of enterprise-wide

SNAPSHOT

Wells Fargo and Company

<<<Wells Fargo and Company (NYSE: WFC) is a US$272 billion diversified financial services company providing banking, insurance, investments, mortgage, and consumer finance services through 6,000 stores, its internet site (www.wellsfargo.com), and other distribution channels across North America as well as internationally. With a staff of about 20, the Business Modeling Group, in the Consumer Deposits Group of Wells Fargo Bank, is responsible for developing and maintaining the 6TB Oracle-based corporate data warehouse.

HARDWARE

>Sun E10000 (for data warehous)

<EMC 8730 Symmetrix Array (for staging, ETL, and meta data management)

SOFTWARE & SERVICES

<Oracle8i Release 8.1.6

<Sun Solaris 2.6

<Veritas File System (VxFS 3.3.1)

< Volume Manager

<Computer Associates Erwin

<Brio Enterprise

<SAS Enterprise Miner

TECHNOLOGY BRIEF

Star Schema and Multidimensional Data

<<<Unlike an OLTP system, in which data is highly normalized, business intelligence systems—decision-support systems and data warehouses—are typically designed in a denormalized model called the "star schema." So-called because it is usually depicted as a fact table positioned in the center of many dimension tables that fan out from the center, the star schema enables the relational database to essentially mimic a multidimensional database. In a star schema, the dimensions are denormalized—they contain redundancies that eliminate the need for multiple joins on dimension tables. An example of a fact table might be Sales, containing columns for units and price; a dimension table might include Customer, Product, and Time.

The star schema enables a relational database to support a multidimensional business model and is commonly used as a format for storing historical data in the warehouse. In the Oracle database, the star schema has been optimized for execution of multidimensional queries, such as "What are the top ten products for time period T and customer X?" Oracle features such as materialized views (with query rewrite) take advantage of this design in the database. Materialized views have been enhanced further in Oracle9i, and other Oracle products—such as Oracle9iAS Discoverer and Oracle9iAS Discoverer Reports—are now more tightly integrated with this capability, offering better performance.

knowledge—accessible from anywhere, over the Web—become more important.

For example, according to Jim Hill, manager of data warehouse operations for ON Semiconductor, one of the most important ongoing initiatives for the data warehouse at ON Semiconductor is to "provide the business with a consistent way, across business unit or functional space, to tell how well the business is doing—to help upper management analyze the gaps against the goals that they've established, so that they can more quickly respond to the issues that come up."

Hill's team works with the sales organization to capture their sales forecast information; the data warehouse shows the sales organization that information against actual business performance, and they use that analysis to drive the decision-making process. Hill's team also works with the planning organization to ensure that the data warehouse provides information for capacity planning so that ON Semiconductor can respond effectively to changing business conditions. Similarly, for Wells Fargo, the data warehouse "is an information hub that collects all of the data from the various operational organizations and provides a single version of the truth," says Jesse Lund, chief information architect for the Business Modeling Group at Wells Fargo Bank. This is important in any company, but especially in a large company, says Lund, where many disparate groups may be performing similar types of analysis.

KNOWING HOW TO BUILD IT The business and technical challenges for a data warehouse that encompasses gigabytes or terabytes of information to be accessed by hundreds or thousands of users are not to be underestimated. What's needed for a successful implementation? According to NetZero's Hammer, "I think—and this isn't exactly a bolt of lightening from above—the most important thing is to make sure that you're close to the business. It's important to ensurethat you're not going off and building the world's technically best data warehouse that does not meet the needs of the business and add value to the business."

NetZero had the benefit of being a startup, says Hammer, and since the business model itself involves using intelligence gleaned from the data warehouse to provide a value-added product to customers, the issues are different than for organizations that have to build a data warehouse from a legacy of other systems. "We were lucky that we were able to build it from scratch the way we wanted it with Oracle, which...is one of the industry leaders in data warehousing," says Hammer.

Building a data warehouse will "test both your technical breadth and depth," says Best Buy's Skochil. "You need a good strategic partnership with Oracle, your hardware vendor, and your other layer vendors right off the bat—that's a key for success." Skochil warns, "You need to really be serious and very careful with large databases, because being a percent or two off can have an incredible cascade effect."

Wells Fargo's Lund puts it more bluntly: "The gloves come off when you start talking about big data. Big data is hard to work with." Lund says that some of the specific features of Oracle, such as parallel execution and

SNAPSHOT

NetZero

<<<With 6 million subscribers and 3.7 million active users, NetZero (www.netzero.net) knows a lot about a lot of people and what they're looking for on the Web. When it launched its service in 1998, NetZero quickly became one of the fastest growing ISPs, thanks in part to its pricing structure—it's free to users who agree to let banner ads display as they browse. (For users who don't want to have ads displayed, NetZero offers a dial-up service for $9.95 a month.) NetZero also offers consumers free e-mail and customizable navigation tools that provide "speed dial" to key sites on the internet.

HARDWARE
- <Sun E10000
- <EMC Storage System

SOFTWARE & SERVICES
- <Oracle8 Release 8.0.6
- <Oracle Portal
- <Oracle Consulting

Using its own patent-pending data mining and analysis algorithms in conjunction with its Oracle data warehouse, the company's CyberTarget division offers marketers and advertisers mass-scale, online market research and measurement services. By fiscal year end June 2000, NetZero had more than quadrupled its registered user

partitioning, that may not be used in smaller installations become very important. He cautions that before tuning the Oracle database, DBAs and system administrators should carefully tune the hardware platform, the operating system, and the volume manager to optimize IO, since disk is the slowest part of the equation. And, says Lund, it's very important for DBAs to understand the interrelationship of parallel execution and partitioning, particularly when it comes to partition pruning in the context of parallel execution.

Although ON Semiconductor's Hill agrees that the technical challenges of building a corporatewide data warehouse are formidable, he thinks the larger challenge at ON Semiconductor is getting people to change their ways. "I don't feel like the technical challenges are the most difficult, Hill muses. "I think the most difficult challenges are the people challenges, of helping people to migrate away from the toolset, where they did everything themselves, and utilize this new set of tools."

BRINGING INTELLIGENCE TO THE PEOPLE Despite the word "business" in their name, business intelligence tools have a role to play that extends beyond the corporate world into the scientific community and public sector. For example, the United States Geological Survey (USGS) National Water-Quality Assessment (NAWQA) program uses a full range of Oracle software to manage massive volumes of information that the public would otherwise have little opportunity to obtain. The data warehouse comprises 4GB of data and contains some 9 million result sets. "It's about US$70 million worth of water quality samples," according to Sandy Williamson, USGS NAWQA National Database team leader.

Access to the content in the data warehouse is through Oracle's ad hoc query and analysis tool, Discoverer 4i. Says Nate Booth, Oracle database developer for the USGS, "With [Discoverer] 4i and really Oracle8i Database, Release 8.1.7 in particular, quite a bit more analytic functions are embedded within the database and now integrated in Discoverer that really have been able to enhance the abilities of our analysts to write within the Discoverer tool. They no longer necessarily have to export the data and come into some more sophisticated scientific statistical tool."

However, the greatest benefit of Discoverer according to Williamson is the cost savings in terms of both reduced maintenance costs and time. "Development- and maintenance-labor costs and timelines are greatly reduced by avoiding the need for custom code and replacing with a fully-functional, off-the-shelf product."

SNAPSHOT

USGS NAWQA Program

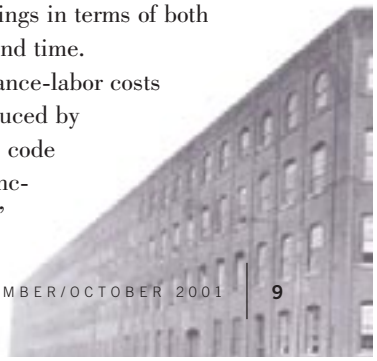
<<<Created by an act of the U.S. Congress in 1879, the U.S. Geological Survey (USGS) serves the United States as an independent fact-finding agency that collects, monitors, analyzes, and provides scientific understanding about natural-resource conditions, issues, and problems. It has no regulatory or management mandate but rather is focused on providing impartial science to serve a variety of needs. Begun in 1991, the National Water-Quality Assessment (NAWQA) program of the USGS was designed to provide information on water quality and a sound understanding of the natural and human factors that affect it. NAWQA's Oracle data warehouse was developed in 1999 to facilitate national and regional analysis of the data from the first 36 basins, with more to be added over time. The USGS NAWQA program uses Oracle Discoverer to perform ad hoc queries against the data warehouse. The NAWQA data warehouse is available for ad hoc queries from the public at http://water.usgs.gov/nawqa/data by means of an Oracle Discoverer 4i end-user layer.

HARDWARE

<Dell Poweredge 6300 (Oracle8.1.7)
<Dell Poweredge 1400 (Oracle9iAS)

SOFTWARE & SERVICES

<Windows NT
<Oracle8i Release 8.1.7
<Oracle Discoverer 4i
<Oracle Portal 3.0.6 on Linux
<Oracle Reports
<Oracle8i Lite
<Oracle Designer
<Oracle JDeveloper



The USGS hosts two different Discoverer end-user layers (EULs, the server-based metadata repository and query management engine for data warehouses), one for the internal USGS data analysts and one for the general public, available on its Web site (http://water.usgs.gov/nawqa/data/). According to Williamson, maintaining two EULs is not a problem, because Discoverer stores common metadata and routines only once. USGS plans to upgrade to Discoverer Plus, the Web-based version.

SNAPSHOT

ON Semiconductor Corp.

<<<Headquartered in Phoenix, Arizona, ON Semiconductor Corporation (www.onsemi.com) is a global supplier of high-performance broadband and power management integrated circuits and standard semiconductors. These products are used in numerous advanced devices ranging from high-speed fiber optic networking equipment to the precise power-management functions in today's advanced portable electronics. ON Semiconductor (NASDAQ: ONNN) was spun off of Motorola in 1999. With just over a year of independent operations under its belt, ON Semiconductor is making good on its strategic initiatives. For instance, last year ON Semiconductor introduced 422 new products, and its customers gave it 15 awards for service. Since becoming an independent company, ON Semiconductor has seen a 28-point improvement in delivery time and realized greater than US$200 million annualized cost savings over 1998. Part of meeting its initiatives includes being able to understand the ongoing operations in a consistent way with the help of the Oracle data warehouse.

HARDWARE
<Sun E10000
<Sun E450
<EMC? Veritas? TKTK
SOFTWARE & SERVICES
<Sun Solaris 2.X TKTK
<Oracle 8.1.X TKTK
<Oracle Warehouse Builder
<Oracle Portal 3.0.8
<Oracle Discoverer
<Oracle Reports
<Oracle E-Business Suite, including: Financials, Manufacturing, Purchasing, Oracle Shop Floor Management
<Oracle Consulting

SEEING INTO THE FUTURE How do you ensure that the data warehouse you build today will be of value tomorrow? Two gurus of the data warehouse, Bill Inmon (www.billinmon.com/) and Ralph Kimball (www.rkimball.com/), have always stressed the fact that building a data warehouse should be an iterative process, says Lund. "The most successful data warehouse installations have begun from some core focus and grown outward from there as they expand in the data and content."

From the beginning, however, the goal for the warehouse was that it would be developed in an iterative fashion, extended as needed to support future goals. To ensure the success of the system, the Wells Fargo Business Modeling Group's strategy included making sure they capture and store all source system data, regardless of whether it is needed immediately. The load process for the Wells Fargo data warehouse includes a series of staging tables where all raw source data, including attributes, is stored in base tables that can be referenced later to provide history if necessary. The tables comprise hundreds of columns and hundreds of millions of rows, and they enable the Business Modeling Group to pick up fields that aren't currently included for existing applications. This approach provides the assurance that a business need not know all of its requirements today but will be able to meet them in the future.

For Best Buy's Skochil, using information from the data warehouse to run the business is what it's all about. "Best Buy really uses the data in the data warehouse to run its business. Data is managed as an asset, it's recognized as a corporate asset, and that's the way it's managed here." And, he adds, as someone who's been a DBA for 18 years, "it's very exciting to work for a company that does that." [SLUG]

Kelli Wiseth *(kwiseth@alameda-tech-lab.com) is research director of the Alameda Tech Lab and Research Center, Inc.*