

# DATABASE

## TRENDS AND APPLICATIONS

Solutions for the Information Project Team • [www.dbta.com](http://www.dbta.com)

Volume 19, Number 7 • July 2005

## SEC Uses Its IQ

*Improved performance and storage are primary benefits.*

### By Walt Jordan

The U.S. Securities and Exchange Commission's (SEC) mission is to protect investors and maintain the integrity of the securities markets in the U.S. Formed as part of the New Deal legislation in the wake of the stock market crash of 1929, the SEC's responsibilities and activities have once again become front-page news in light of the corporate scandals in recent years.

The fundamental concept underlying the securities laws in the United States is simple. Markets operate most efficiently when there is a free and open flow of information to all investors ranging from private individuals to large institutions. Consequently, the SEC mandates that all public companies disclose meaningful financial and other information to the public. It also oversees the key participants in the securities world, including stock exchanges, broker-dealers, investment advisors, mutual funds, and public utility holding companies.

With four divisions and 18 offices, including 11 regional offices throughout the country, enforcement of the law is one of the SEC's key responsibilities. Each year, the commission files between 400 and 500 civil enforcement actions against individuals and companies that break the securities laws. Violations include insider trading, in which individuals act on information that is not freely available to the public, accounting fraud, and providing false or misleading information about securities and the companies that issue them.

### Pre-Y2K Infrastructure

With the free flow of information one of the SEC's primary responsibilities, clearly the effectiveness of the commission's information infrastructure is a critical component in its ability to carry out its mission. Like many large government agencies prior to the year 2000, the commission had a mixed IT environment, which included a mainframe running the Adabas database and using applications developed in the Natural language; there were COBOL applications; some Java development; and applications created with Sybase PowerBuilder. "We needed all these skill sets just to keep us running," said Lewis Walker, who joined the commission just prior to Y2K and now serves as the assistant director for application development in the Office of Applications and Software Development.

With the impending Y2K deadline just one of the motivating factors, the decision was made to consolidate the infrastructure, moving to a common database, common application development environment, and a common hardware platform. "We had been using PowerBuilder, but Y2K gave us impetus," Walker said. "The first step was to decide what our architecture would be." The commission opted to migrate its mainframe applications and consolidate on Sun Solaris running the Sybase ASE enterprise database.

### Blue Sheets Application

When the migration off the mainframe began--it was not completed until after the Y2K deadline--there was not that

much data under management, less than a terabyte according to Walker. But with the new platform in place, the commission began developing new applications using Java. "We did exceptionally well," Walker said, "The data began to grow."

One of the key applications was internally known as Blue Sheets, named after the color of the form that was used when the process was manually executed. If an SEC analyst wants to investigate suspicious trade data, that data is requested via Blue Sheets.

The process works like this. The SEC receives a data feed called Equities Cleared from the Securities Industries Automation Corp (SIAC). A subsidiary of the New York Stock Exchange and the American Stock Exchange, SIAC runs the computer systems and communications networks of the stock exchanges and disseminates the information generated. Equities Cleared is a record of trades in which buyers and sellers have been matched. The SEC receives that information in an aggregated format. "There may be one million trades of a particular stock in a certain time period," Walker said. "But you don't know who made the trades."

If after being tipped off by a news report or other source, SEC analysts believe that if the activity at the aggregated level looks unusual or peculiar, using the Blue Sheets application, they can request the underlying detail-level data. That data is then loaded into the database, and SEC analysts can assess it using Business Objects tools. Initially, the Blue Sheets application was migrated to the Sybase ASE platform. The migration was so successful that the number of Blue Sheets requests soared. "That meant the amount of data increased dramatically," Walker observed.

### Move to IQ

As the number of Blue Sheets requests climbed and the amount of data to be analyzed concurrently increased, Walker's database administrator suggested that the commission look at implementing Sybase IQ, a relational database optimized for BI applications, as the repository for Blue Sheets data.

It was not IQ's data warehousing functionality that was of primary interest. "We didn't have a need for a traditional data warehouse. We didn't need the aggregation of data from disparate sources. We were going for high volume data, speed, and saving on storage. What was compelling to us was that IQ could both save storage space and increase the speed at which data can be retrieved," Walker said. Also, analysts were requesting ever-larger amounts of data to analyze from an ever-increasing database of information. IQ was added as an extension of the current infrastructure.

### Benefits

The move has paid off. According to Walker, as a result of Sybase IQ's compression technology, the commission has been able to save as much as 50 percent on storage requirements, more

than the 18 to 20 percent savings originally anticipated. "That has saved us from buying storage and the associated expenses," Walker said. Moreover, according to Samuel Foster, president of FosterSoft, Inc., a government contractor that worked on the project, query response time decreased by as much as 35 percent.

And there was a secondary benefit. The SEC uses EMC storage devices. Walker had long wanted to use EMC's replication technology to create the hot backup copies of the database needed for high availability. In traditional database applications, hot backup relies on the database transaction logs, making that approach challenging. But the EMC replication is Sybase IQ-aware. Consequently, Walker said, "I don't have to worry about a replication server."

Several terabytes of data are now stored in IQ. It is accessed by hundreds of SEC investigators and analysts working with Business Objects tools. The next step is to make the data available to the commission's economists. They use SAS tools to identify long-term economic trends. Currently, Walker noted, the 10 to 20 economists run their queries against a file system. "They use heavy duty algorithms and there have been

problems," he said.

Walker also plans to use IQ's large binary objects (LOB) functions to begin loading evidentiary documents into the database. "We are experimenting with that," he said. "We think we can get better control and parse information from the documents into the database to help with queries." He is also exploring search software solutions.

Over time, Walker hopes to incorporate the management of XML documents directly into the database platform. Internally, the commission has made a commitment to XML. "The idea is to store the XML documents in the database and query them *in situ* without having to parse the data into the traditional relational tables," he said. "Then people can create documents and put them in the database without having to create table designs where to parse the data."

From an IT perspective, the move to IQ has given the commission improved response times while saving on storage costs. More importantly, the new platform has enabled the commission to make more data available to more people for more purposes efficiently. "That opens the possibility for better analysis and improved enforcement," he said.