

Effectively managing energy data

Sandy Fielden of Logical Information Machines, Inc explains why central databases are the key to success for the modern energy trader

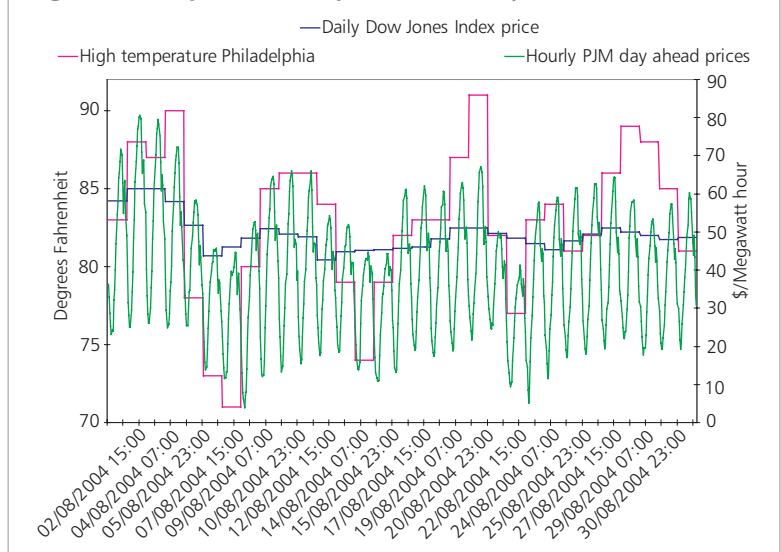
> In these turbulent times for energy trading it has become extremely important for companies to rely on a single database of information that can be accessed by all areas of an organisation. Evidence of accounting irregularities has highlighted the need for better audit controls. The FASB 133 accounting standards and Sarbanes-Oxley legislation both require higher standards to be applied to the valuation of market transactions and deals on a company's balance sheet. The issue of pricing derivative transactions is particularly complex. Inconsistency in pricing at varying levels in a firm can be a cause for millions of dollars of waste through incorrect billing/invoicing and the labour spent researching the correct price. A solution to these challenges is the implementation of a single repository for all prices, forecasts and transaction values.

Data quality and robust technology are the cornerstones for any data-warehousing project. Companies such as Logical Information Machines, Inc (LIM) provide just such a solution. The project can be segmented into three major areas: collection of market pricing, storage of internally generated information, and providing a technology solution around access to the data.

Market pricing

Our first area of focus is the collection, cleansing and maintenance of market-based pricing. In the energy arena, pricing is available from exchanges, websites, paper publications, electronic services, brokers and internally generated assessments. Gathering all that data in one place presents an immediate challenge. Adding to the difficulty are the varying characteristics of the information itself. Electricity pricing can be hourly, half-hourly, daily or real-time. Publication assessments can be daily, weekly, monthly or even quarterly. Transaction-based data requires information on a tick-by-tick basis. Forward prices use different bases depending on the commodity and the market structure. How can we ensure that the differing granularities are handled correctly?

Figure 1. Power prices and temperature in Pennsylvania



Time-series databases, such as the 'Market Information Machine' (MIM) from LIM, offer support for the various data granularities handled by energy market players. Support for reshaping the data is also important in that it will aid in the analysis of data quoted in different terms. For example, a trader may have access to data on an hourly basis from an independent system operator (ISO) and on a peak/off-peak basis from a publication provider. It is possible with the correct tools to adequately analyse the relationships in the data by asking the database for an hourly price from the ISO to compare with the peak average from the publication vendor, as well as to display the temperature in the area (see Figure 1).

Another crucial component of any storage solution is the ability to manage data quoted on the weekend. As traders are now structuring deals that span off-peak periods on the weekends and

vendors such as Platts provide data with flow dates across weekends, any solution must support Saturday and Sunday data, as well as offer the ability to analyse such data in relation to five-day data series, such as exchange-traded futures.

To make this type of analysis work, the user must be presented with the ability to fill information through the gaps on weekends (or other data gaps) by varying methods, according to need. For example, comparing a natural gas assessment that includes weekend data to the New York Mercantile Exchange (Nymex) natural gas futures contract, you must have the ability to fill the Nymex history through the weekends with a method of interpolation. Many times a simple fill forward or backward is adequate, but linear or logarithmic interpolation may be more appropriate when dealing with conflicting holiday calendars.

Dealing with price corrections is another essential element to be addressed in the management of market pricing. If a contract is written based on an index price published by a vendor, the full history of the price must be available, as well as the history of changes to the individual prices. Improper handling of corrections can lead to disputes, or incorrect invoicing or trade settlement. The database must support rollback in order to determine what the publisher said the price was for a certain day, effective on a certain day.

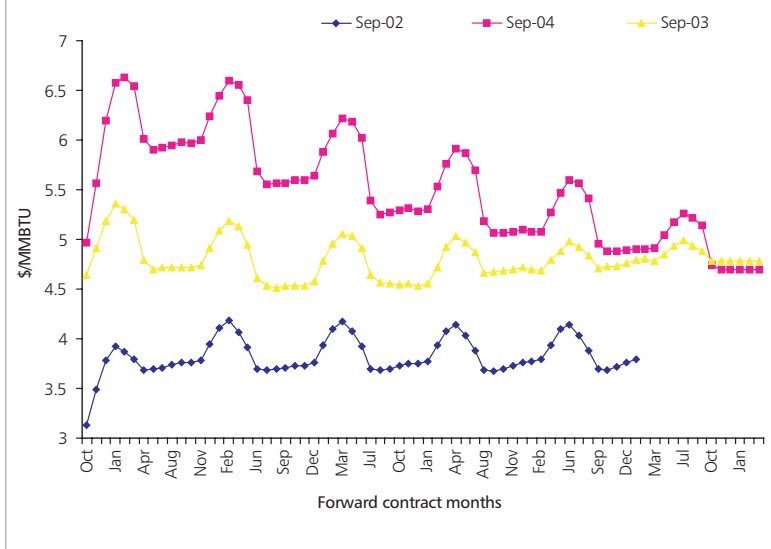
Forecasting and forward curves

In today's market, forward curves and forecasting are important methods for pricing structured transactions, trading over-the-counter derivatives, and marking portfolios to market. The ability to easily store and retrieve such curves is especially important in creating a structured deal quickly, while your counterparty is on the phone, as well as marking the deal to market at the end of the day. The need to capture forward curve

Data quality and robust technology are the cornerstones for any data-warehousing project

data generated by the trading desk illustrates the next crucial element of any data management solution: the loading and display of internal data. Companies generate internal curves that are really forecasts of prices on a daily or more frequent basis. Such forecasts should be stored as a record of that day's view of the forward market – even when the data is an hourly forecast going forward several years. It is important to store this data easily and then be able to retrieve a slice of data that tracks how your view of a particular forward period changed over time. In this way, forecasting approaches can be compared and the model fine-tuned (see Figure 2 for an example). Once deals are on the trading book, continued monitoring of internal and external forward curves is required by risk managers. Collection and organisation

Figure 2. Nymex natural gas forward curves



of audit-worthy views of the forward market are critical in the post-Enron trading world.

Ubiquitous data access

Storage of data is very important, but it has limited value unless all levels of the organisation can retrieve the data using the methods required by their group. Any data-warehouse solution must be implemented with a varied set of APIs and offer additional methods of data retrieval. Risk and portfolio applications often implement their own data store and require a physical file dump, back-office personnel require access to data in standard Microsoft 'Office' applications, and analysts need access to the information in their favourite statistical analysis application.

Loading data from the warehouse into other applications is critical. There are two main approaches – 'pull' and 'push'. Pull technology requires a scripting language or method to export files of information from the warehouse on a regular basis. Larger organisations implement system bus technologies to transport data between subscribing applications. The ability to push or publish new data from the warehouse onto the bus, as well as filter out corrections and identify new series, feeds these systems with a flow of data as soon as it is collected.

Lastly, for the enterprise to benefit from this centralisation of data, a robust technology must be employed to permit access from thousands of desktops at once. The speed and reliability of the server and its ability to service user queries at the same time as data is being loaded is essential. LIM's MIM server technology is being used throughout the energy and financial markets. It has also recently proved its strength when adapted to process thousands of transactions a second for an application used by GE Power to capture turbine data. ^{ERI}