

SciDB for Financial Markets

from Paradigm4

Case Study: Transaction Cost Transparency Boosts PDQ's Innovative Trading Platform Business

PDQ Enterprises (PDQ), the owner and operator of PDQ ATS, Inc., an independent alternative trading system, provides customized execution and unique liquidity generation for equity market participants. They came to [Paradigm4](#) to accelerate and streamline the creation of advanced transaction cost analysis (TCA). PDQ ATS puts TCA up front and center to help brokers and dealers optimize trading performance.

Transaction cost analysis (TCA) assesses the cost of acquiring liquidity by comparing realized trade prices to benchmark values. Assembling an aggregated order book of all trades and quotes for stocks traded on all major U.S. equity exchanges enables the calculation of the most accurate benchmark values by allowing all relevant microstructure data to be incorporated into the analysis, a challenging task which SciDB is well suited to solve.

After a thorough review of several data management platforms, they chose SciDB, a radically new computational database for mining insights from financial trade and quote data.

SciDB was developed by renowned database researcher and Turing Laureate, MIT Professor [Michael Stonebraker](#).

ABOUT SciDB

SciDB is a computational database management system designed for managing and analyzing highly dimensional and diverse data. Key features include:

- Rapid n-dimensional selects, joins, windowing, and aggregation accelerated by our unique MAC™ storage technology
- Multi-user support while preserving data integrity
- Transparent, scale-out, massively parallel processing, on premise or in a cloud
- Scalable, in-database advanced analytics like covariance and PCA
- Accessible/programmable from R, Python; plus GUIs like R Shiny

SciDB slices through massive amounts of data fast so teams can easily share data, analyze more, and program less.

SciDB for Financial Markets

from Paradigm4

BACKGROUND & BUSINESS REQUIREMENTS

PDQ ATS gives market participants time to achieve enhanced trade execution through an on-demand auction process that creates unique liquidity in response to each order. An alternative trading system (ATS) is an off-exchange equity market that matches buy and sell orders from subscribers to the system.

PDQ ATS offers:

- A modernized, inclusive market structure
- Patented auction models utilizing pause technology
- On-demand solutions that aggregate unique liquidity
- Control and choice for equity traders

In January 2016, PDQ ATS averaged 128.5 million shares executed per day in both the ATS and via PDQ's extensive custom routing options, with an average of 13.4 million shares per day traded in the ATS.

Traders and brokers are able to answer complex hypothetical questions like:

- What would be the average market impact of a segment of order flow in a variety of scenarios that look at: trade size, time horizon (measured in terms of elapsed volume after the trade time), and market impact?
- How does the cost of immediate liquidity, represented by the universe of exchange display books, relate to the actual realized cost of liquidity, as indicated by the stream of transactions?
- How does cost vary for each security? What explains that cost variance?

PDQ ATS needed to analyze market microstructure so that when subscribers wanted to acquire or liquidate a position, they could also forecast how much the particular transaction should cost. In order to maintain an optimal flow of information, PDQ ATS needed to ingest large amounts of market data quickly, which their previous system was unable to do.

SciDB's time series data management, parallel distributed processing, dynamic provisioning, and tight integration with R and Python, the most common languages used in financial markets, eliminate the hassles PDQ ATS experienced with data management and simplified dealing with computational complexity. In addition, SciDB handily managed ingesting data at the rate required.

SciDB for Financial Markets

from Paradigm4

THE TECHNICAL CHALLENGES

The technical data management challenges for PDQ Enterprises' Advanced Liquidity Studies and Analytics platform (ALISA) are significant. There is a lot of data preparation (i.e. extract, transform, load) that has to happen before PDQ can build the master order book that underpins their proprietary TCA analysis. PDQ is able to load, index, and cluster post-trade data simultaneously into SciDB at a high rate, which they were unable to do with their previous system.

“On one side, there is data provided by vendors and collected by us that exists in various formats: CSV, binary, network traffic capture files, etc. On the other side, there is a need for the analyst to test hypotheses and make conclusions. We needed a solution that let us focus on analysis. That is where our expertise lies, and we turned to SciDB to bridge the gap between the two sides and bring the data conveniently to the analyst.”

*Alex Nazaruk, Data Platform Architect
PDQ Enterprises*

PDQ queries SciDB to select a full day of market data for sets of instruments and exchanges. SciDB serves up this data as R data frames that can be fed into their custom R program to compute the order books for each exchange and aggregate that into a master book. Like PDQ, many trading companies have proprietary algorithms in R, Python, and C++ that they want to feed from a tick database.

“SciDB is tightly integrated with R and insulates us from programming the database. With a simple ‘getSeries’ function that is executable from an R environment, we just ask for the data and it is there for us in our preferred programming environment. This simple interface is why we chose SciDB.”

*Alex Nazaruk, Data Platform Architect
PDQ Enterprises*

SciDB for Financial Markets

from Paradigm4

PDQ's scalable R solution—developed by Paradigm4—lets them scale out their compute environment on-demand so that they can meet peak loads without over provisioning. Plus, shorter execution times accrue from massively parallel computation.

With an integrated SciDB and R solution running in a cloud, building and analyzing market data for all major U.S. equity exchanges happens overnight. Transaction costs are analyzed throughout the next day using this up-to-date master order book.

WHY SciDB?

PDQ Enterprises evaluated SciDB against a leading column-store SQL database and a legacy time series database. SciDB provided outstanding performance (on ingest, selects, and aggregates) and offered a way to control the system from, and return results to industry-standard and analyst-friendly R. With the other systems considered by PDQ, their analysts would have had to learn an esoteric database language or use SQL. SciDB eliminated this concern, enabling their analysts to do their job without getting bogged down in the management of a tick database.

SciDB benefits:

 A massively parallel processing architecture allows PDQ to distribute data, to scale storage and compute resources, and to retrieve data quickly using compute cloud.

 SciDB's array versioning and flexible data model support requirements for multiple named or time-stamped versions of data.

 Multidimensional Array Clustering technology MAC™ powers the fast range selects and joins. SciDB preserves the natural order of data and exploits that order to accelerate range selects and joins.

 Built-in support for time series data helps PDQ represent time in flexible and intuitive ways, enabling them to produce numerous aggregates on time ranges.

“Paradigm4 delivered much more than just a tick database... they provided a complete and scalable solution, freeing us from the need to become infrastructure technologists.”

*Keith Ross, CEO
PDQ Enterprises*

SciDB for Financial Markets

from Paradigm4

WHAT'S NEXT?

Since 2008, financial markets firms have dealt with a seismic shift in their day-to-day landscape arising from greater data variety and complexity to increased regulatory accountability. In the aftermath, many have re-capitalized, restructured, and reformed. Those actions were necessary but they alone will not provide a competitive edge go forward. The edge will come from innovation and differentiation.

Move forward.

Contact Paradigm4.

[http://www.paradigm4.com/about/contact/
financialmarkets@paradigm4.com](http://www.paradigm4.com/about/contact/financialmarkets@paradigm4.com)

ABOUT PARADIGM4

Paradigm4's SciDB—the latest innovation from renowned database researcher, Turing Award Laureate, and entrepreneur, MIT Professor Michael Stonebraker—is a radically new computational database for mining insights from genomic, clinical, image, financial markets, instrument, and sensor data. Paradigm4 is changing what's possible with Big Data by answering bigger and harder questions.

Industry leaders like MSCI, PDQ, Novartis, MIT Lincoln Laboratories, Foundation Medicine, the National Institutes of Health, NASA, and others use SciDB to drive the creation of new products and services.



281 Winter St #360
Waltham, MA 02451
www.paradigm4.com

FinancialMarkets@paradigm4.com

 [@paradigm4](https://twitter.com/paradigm4)

 [/company/paradigm4](https://www.linkedin.com/company/paradigm4)

MAC is a trademark of Paradigm4, Inc.