

# ThingSpan™: The Distributed Graph Platform for Financial Institutions

## The Business Opportunity

The financial services industry is rapidly evolving, both in terms of its business requirements and the enabling technologies needed to improve decision-making and gain competitive advantage. It's crucial to adopt a massively scalable distributed graph platform that can handle:

- Graph analytics and real-time relationship discovery
- Integration with the open source stack – Spark, Kafka, HDFS, YARN
- High-speed ingest with parallel querying
- Petabyte-scale to trillions of nodes and edges



Forrester Research estimates that graph databases will be the fastest-growing area in database management systems, with more than 25% of enterprises using graph by 2017<sup>2</sup>.



An enterprise-grade graph analytics platform can be used to augment these three broad categories:

### · **Smart trading**

Big data analytics is crucial for evaluating long-term investment strategies and generating alpha. A graph analytics platform lets financial companies optimize their portfolios by enabling analysis of all various asset classes' past performance and their predicted growth trajectory while highlighting how they impact the valuation of related assets.

### · **Regulation and compliance**

Compliance can be difficult to track for institutions dealing with hundreds of millions or billions of trades across many geographic regions and industries. Using a graph analytics solution to Know Your Customer (KYC) is the key to accurately understanding risk, as well as catching rogue activities, such as insider trading and securities fraud.

### · **Cybercrime prevention and detection**

Cybercrime is the top market threat to financial organizations<sup>1</sup>, and many security breaches are not identified until the damage has been done. A graph analytics solution can instantly spotlight any unusual patterns that may be indicative of a security breach by correlating data from security and network solutions with internal and semantic web applications. Unauthorized activity can then be shut down before theft can occur.

While the value of big data is undeniable, many financial institutions struggle with adopting a solution that can help them rapidly process data to enable immediate decision-making. When they are not able to realize the value of their data quickly, billions of dollars may be at stake.



## Use Case: Smart Trading

When evaluating a trade, traders or asset managers should be able to see how the factors at play (accounts, sectors, exchanges, market data, reference data, sentiment, etc.) compare to historical trends and be able to determine the relationships between various transactions to accurately forecast the stock's value and identify undervalued equities for investment.

In addition, smart trading practices must involve analyzing all data streams from a plethora of markets, news and social media in order to create new data models to generate alpha. Application of sophisticated machine-learning techniques on graphically stored data in real time is the right solution for the problem.

<sup>1</sup> CFTC Commissioner J. Christopher Giancarlo, Dec. 1, 2015  
<http://www.fiercefinanceit.com/story/cybercrime-number-one-market-threat-cftc-commissioner-says/2015-12-08>

<sup>2</sup> Big Data Analytics News:  
<http://bigdataanalyticsnews.com/4-predictions-for-nosql-technologies-in-2016/>

## The Technical Challenge

Many organizations rely on big data analytics solutions that involve a data ingestion layer that processes all data points about financial transactions and customer data—including market, news, and social media—before breaking them down into stored memory where they can be queried. However, while large volumes of data may be processed in this manner, institutions often rely on micro-batches and may need to wait days until the relevant data points can be surfaced. In many cases—particularly when dealing with profit opportunities and fraud—such delays are not an option.

In order to generate the high-performance, high-speed processing power and the sophisticated contextual analysis needed within the financial industry, an enterprise-class graph analytics solution is essential.

The challenge, however, is finding a distributed graph platform that can scale to petabytes of data and perform queries in parallel to data ingestion. While there are many platforms available, few of them offer real-time data analysis at scale, enabling organizations to capture streaming data and analyze it in relation to historical and contextual data to identify opportunities and risks across a broad array of use cases.

Institutions need a highly scalable, real-time graph analytics platform to analyze massive volumes of financial and customer data for risk, fraud, compliance, business performance, predictive analytics, and other benchmarks that are high priority to organizations.

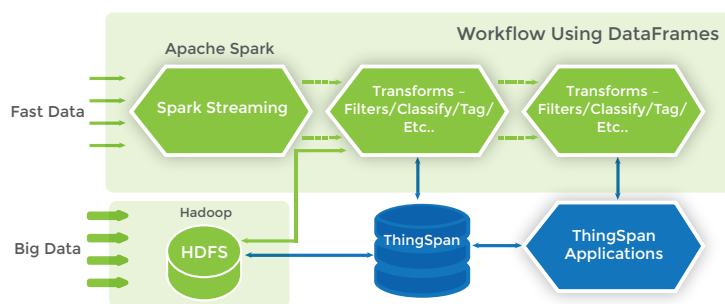
## The ThingSpan™ Solution

ThingSpan™, Objectivity's massively scalable graph analytics platform, is compatible with the distributed open-source data management framework, **Hadoop**, and utilizes **Spark** for high-speed streaming ingest and enriched data processing. It has the power to transform and analyze real-time financial and IT data in context, offering a single logical view of all data—wherever located—to accelerate parallel processing.

ThingSpan leverages open-source tools by supporting the Hadoop and Spark ecosystem atop a high-performance, distributed graph database purpose-built for relationship and pattern discovery. It runs natively on top of POSIX or HDFS as a YARN application while using Spark for workflow and data transformation. It is also designed to support streaming systems based on Kafka, Flume and other distributed messaging tools for streaming data. Integration with Spark via DataFrames allows ThingSpan to ingest streaming data while maintaining and persisting relationships as first-class logical models.

This model allows for enriched and transformed data to simplify the support of complex, multi-dimensional queries associated with financial services applications and analytics. ThingSpan enables financial organizations to capture powerful insights around data relationships to make better-informed decisions in financial trading and asset management, avoid regulatory and compliance problems, and remain vigilant about cybersecurity attacks and financial fraud.

### ARCHITECTURE DIAGRAM



Objectivity, Inc. is a pioneer in high-performance distributed database platforms that power mission-critical applications for the most demanding and complex data sources in the enterprise. Objectivity enables organizations to accelerate time-to-value of their data assets at scale by enriching Big Data with Fast Data. With a rich history of serving Global 1000 customers and partners, Objectivity holds deep domain expertise in fusing vital information from massive data volumes to capture new revenue opportunities, drive competitive advantages, and deliver better business value. Objectivity is privately held with headquarters in San Jose, California. Visit <http://www.objectivity.com> to learn more.



ThingSpan is a massively scalable distributed graph platform designed specifically for the complex issue of extracting actionable insights from Fast Data and Big Data sources to enable real-time relationship discovery. It is architected to integrate with major open source Big Data technologies, such as HDFS and Apache Spark. ThingSpan leverages Objectivity's core object and graph data-modeling technologies and the company's rich experience in building fusion solutions.