High Frequency Trading with Complex Event Processing

Conference Paper · December 2016

DOI: 10.1109/HiPCW.2016.014

Nandini Sidnal
Visvesvaraya Technological University

High Frequency Trading with Complex Event Processing

Ajay Acharya
Assistant Professor, Dept. of CS&E
KLE Dr. MSSCET, Belagavi, India
ajay.acharya@klescet.ac.in

Dr. Nandini. S. Sidnal
Professor, Dept. of CS&E
KLE Dr. MSSCET, Belagavi, India
sidnal.nandini@gmail.com

Abstract—**High frequency trading is steadily taking over the equity trading world. High frequency trading involves very high speed systems placing trades at sub millisecond speeds across multiple stock exchanges. HFT is a good example for Big Data analytics - especially the velocity aspect of big data. For HFT strategies to be profitable, real time processing of big data is essential. In this paper we discuss the challenges faced by HFT systems and the opportunity for big data processing with low latency in the field. Most HFT systems are designed using real time stream processing, which have certain drawbacks. We present a theoretical framework for building high frequency trading systems using the complex event processing paradigm which could overcome the drawbacks of stream processing. Complex event processing enables detecting patterns of events from disparate events streams and responds to the detected pattern. The applicability of the framework for HFT applications is discussed.**

Keywords- High Frequency Trading, Complex Event Processing, Big Dataanalytics

I. Introduction

Traditionally Big Data has been characterized by the 3Vs, Volume, Variety and Velocity[1]. While a lot of research has gone into the first two Vs, i.e. Volume and Variety, focus on the third V (Velocity) is now increasing. As the analytical capabilities of large amounts of diverse data improve, there is increasing need for real time analytics. In this paper we look at the financial domain and high frequency trading in particular to bring about aspects of real time analytics.

High Frequency trading (HFT) is one of the recent financial innovations to have taken over the world of equity trading. It is estimated that the majority of all trading activity in mature stock exchanges is a result of HFT activity[2]. HFT involves computers placing a large amount of trades at sub-second speeds with a view of profiting at the expense of slower traders. We believe that HFT offers a great opportunity for exploring the Velocity aspect of Big Data analytics. Based on the available literature, most of the work on HFT appears to be done within large financial entities or specialized hedge funds; little work is seen from the academic circles to navigate the opportunity for analytics in the HFT domain.

A majority of the literature on real time analytics focuses on stream processing. While it has its advantages, there are some disadvantages of stream processing as the complexity increases. In this paper we discuss about complex event processing (CEP) and reason that CEP is better suited for HFT analytics. Complex event processing enables detecting patterns of events from disparate events streams and responds to the detected pattern.

II. HFT

In the recent past, especially after the financial crisis of 2008, high-frequency trading has progressively gained a firm foothold in financial markets. This is a result of increased competition among hedge funds, enabling legislative measures and competition among stock exchanges. The ever increasing computing speeds and processing capabilities have also played a major part in the rise of HFT. In popular literature, the terms “electronic trading”, “high-frequency trading and “algorithmic trading” are used interchangeably which is incorrect. So it is necessary to differentiate between these terms. Aldridge[3] provides a good discussion on the differences between these terms. The differences are clearly brought about in figure 1.

Electronic trading refers to the ability to place and execute orders electronically as opposed to non-electronic modes like telephone, mail, or in person.

Algorithmic trading (AT) is more involved than electronic trading, it is not limited to mere execution of trades; algorithms are used to manage orders and to reduce market price impact by breaking large orders into smaller ones, and closely tracking benchmarks over the execution interval .

High-frequency trading (HFT) is a subset of algorithmic trading where a large number of orders (which are usually fairly small in size) are sent into the market at high speed, with round-trip execution times measured in microseconds [4]. Programs running on high-speed computers analyze massive amounts of market data, using sophisticated algorithms to exploit trading opportunities that may open up for milliseconds or seconds. Participants are constantly taking advantage of very small price imbalances; by doing that at a high rate of recurrence, they are able to generate sizeable profits.

In contrast to traditional algorithmic trading, high frequency traders do not hold long positions. They instead enter into short-lived positions and do not carry over risk into the next trading session. Algorithmic traders or “algo” traders generally try to deduce the market impact of large orders. In the past few years, there have been a number of studies of high frequency trading. A survey of 56 academic

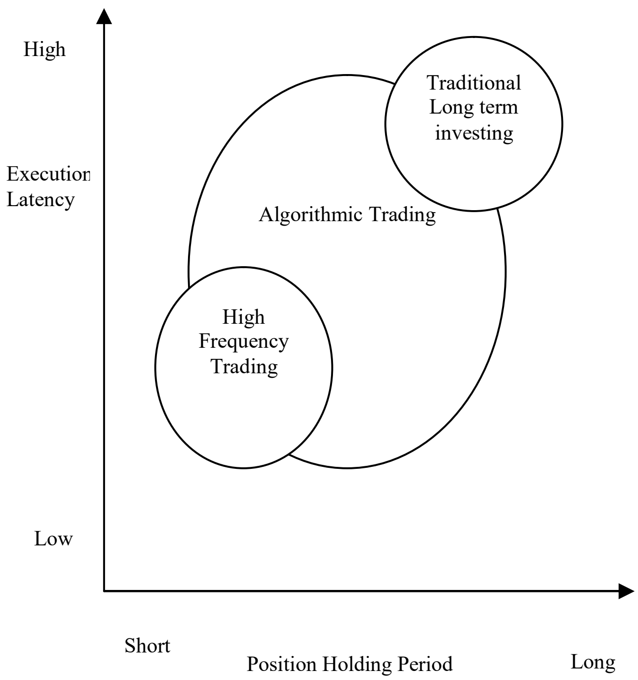

Figure 1. Differences between HFT and algorithmic trading.

research papers was conducted by [5]. As per the survey, the literature addresses the following issues: price discovery impact, economic impact, limit order book dynamic modeling, theoretical modeling, and behavior studies of HFT trading practices. There is very sparse literature on the technology use in high frequency trading and none deal with high frequency trading from a big data analytics point of view. What exactly constitutes high frequency trading is still matter of debate, but according to the US Securities and Exchange Commission (SEC) [6], high frequency trading can be described using the following characteristics:

1. Very fast and complex systems are used for placing and executing orders on the stock exchanges.
2. The HFT firms make use of sophisticated tools like individual feeds and co-location services of the exchanges to minimize latency.
3. The positions are maintained for a very short duration.
4. A lot of orders are submitted and cancelled subsequently.
5. At the end of the trading day, none or minimal positions are carried over to the next trading session.

III. HFT Strategies

Over time, HFT algorithms have continuously evolved: the first generation algorithms were fairly simple in their goals and logic; they were pure trade execution algorithms. The second-generation algorithms were for strategy implementation; these were much more sophisticated and were typically used to produce their own trading signals which were then executed by trade execution algorithms. The third-generation algorithms learn from the market activity and adjusts the trading strategy of the order based on what the algorithm perceives is happening in the market. HFT firms implement several strategies to maximize their profits. These strategies typically attempt to discover underlying temporary but recurring pricing phenomena that can generate trading opportunities. Generally, these strategies target the relationships between the stock prices, like those between pairs or even groups of stocks, or of an instantaneous stock price quote to that stock's historic price, or relationships between the prices and volatility. In general, they do not involve any fundamental analysis of the company.

Strategies targeted at providing liquidity try to replicate the traditional role of market makers – but unlike the market makers, electronic market makers (liquidity providers) have no formal obligation for market making[7]. These strategies involve making a two-sided market with an aim to profit by earning the bid-ask spread.

In statistical arbitrage strategies[8], the traders seek to correlate prices between stocks and to profit from imbalances in those correlations. The correlation could be between 2 stocks or even a group of stocks or any other parameter.

In liquidity detection strategies, traders aim to discover whether there are large open orders by sending out small orders (also known as pinging) to seek for large orders. If a small order is filled quickly, there is a good chance that there is a large order behind it.

IV. Challenges with HFT

We will now discuss the challenges of High Frequency Trading from a big data processing point of view. For most HFT applications, market data should be processed immediately as it is ingested for more critical latency computations. This feature is also required by many financial analytics other than HFTs, such as risk management and illegal trading detection. Also, the analytics of "new" real time data may need lookup of historical data, so the approaches should be built based on a high performance programming model. To achieve real time data analytics, optimizing the memory usage to keep temporal data in memory should also be considered.

The main challenge for HFTs related to big data analytics is to effectively extract valuable information in time. This will entail the designing of a highly efficient system to process the distributed historical and real time data. The traditional computation-centric model is not suitable for this scenario due to the frequent data movement. Programming models, such as MapReduce[9] can significantly reduce the I/O and overheads and realize a high aggregate bandwidth. However, the main goal of MapReduce is not to support low latency processing. Later approaches based on stream processing such as Spark try to utilize local memory to speed up computation; however the improvement is still not enough for real time analytics. Complex event processing (CEP) is better suited to dealing with the challenges when it comes to big data analytics with respect to HFTs.

V. COMPLEX EVENT PROCESSING

In equity markets, businesses need to be able to spot and take advantage of opportunities with speed and precision. Real-time data streams, such as market feeds, are becoming quite common. Complex Event Processing could be used to support the rapid analysis and identification of opportunities and threats in real time.

Essentially Complex Event Processing is used to refer to the process which combines data from multiple sources to create inferences or recognize patterns that suggest more complicated circumstances[10]. The ultimate goal is to identify meaningful patterns, such as threats or opportunities, and respond to them as quickly as possible. Complex event processing deals with the evaluation of a multitude of events and then taking action as required. The events may be of different types and may occur over a different duration. The correlation among the events may be financial, temporal, or spatial. CEP entails the usage of sophisticated event interpreters, matching, and correlation techniques. According to [11], complex event processing, "provides a software infrastructure that can detect patterns of events by filtering, correlating, contextualizing and analyzing data captured from disparate live data sources to respond as defined using the platform's development tools." Figure 2 shows a typical CEP structure.

CEP is used in conjunction with other similar technologies like business activity monitoring, operational intelligence and message-oriented middleware. CEP is capable of operating on live data arising from diverse sources producing homogeneous or heterogeneous events.

The major disadvantage of stream processing is that the process of pattern detection happens only within a stream. Events occurring in a different stream will not be considered even though they may be relevant. Effectively, the stream processing method is a silo based technique which could miss the bigger picture.

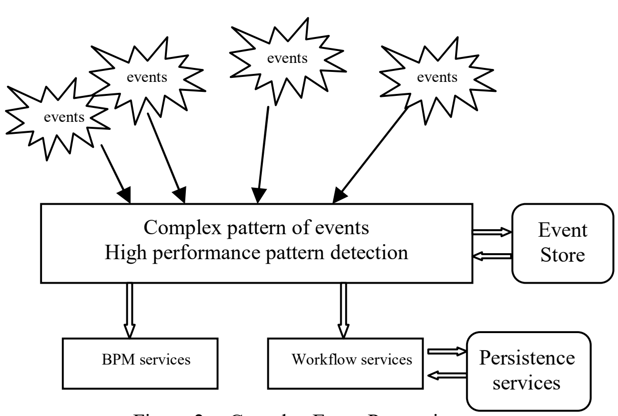

Figure 2. Complex Event Processing.

VI. HFT PERFORMACE ASPECTS

To improve the performance and robustness of HFTs, the following issues need to be addressed:

A. *Message latency*

To achieve message latency, the underlying network should be made of high performance computing elements. Low latency networks are critical for timely message passing. Proximity of the CEP compute grids to the capital markets would reduce the network latency to a great extent. A lot of stock exchanges are providing co-location facilities to minimize network latency. Improving the message latency will also reduce the time taken between the decision making and execution.

B. *High performance computing*

A high performance computing setup is vital for HFTs. This will enable the ability to price and calculate risks and positions at portfolio scale in near real time[12]. It will also support the ability to analyze data and performance after the execution to identify and evaluate market strategies thus creating a platform for informed decision making. The CEP based architecture will enable the execution of strategies automatically, in real time; this is a primary requirement of HFT.

VII. ARCHITECTURE

An architecture based on complex event processing for HFT is presented in figure 3. It consists of the following components:

A. *Input Event Adapter*

The Input Event Adapter receives events that are coming to the CEP. This component is a combination of different adapter implementations. Most common implementations are available by default in the CEP. Because of the CEP's pluggable architecture, custom adapters can also be implemented and plug into the CEP. From a HFT perspective, the input events could come from various streams such as a real time market feed, news feeds etc.

B. *Event Pre-Processor*

The Event Pre-processor converts the events received by the Input Event Adapter to Event Stream and handles the input mapping related tasks in the CEP. This will include tasks like filtering, loading the cache etc. The aim of this step would be to format the input events from different source streams into one which is understandable by the event processor.

C. *Event Processor*

This is the core event processing unit of the CEP which handles actual event processing. It handles the different execution plans and processes events with the help of the

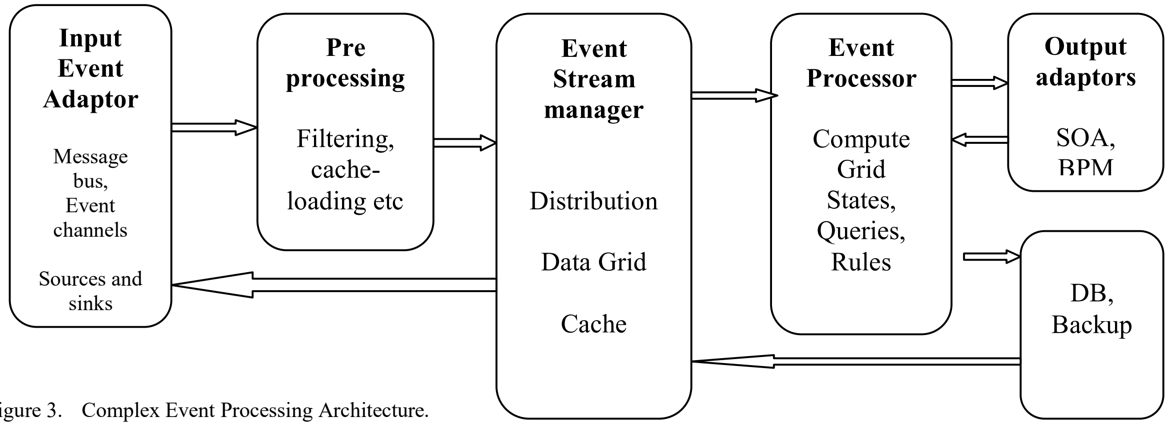

Figure 3. Complex Event Processing Architecture.

core computation engine. This would be comprised of a grid of high throughput computing engines having capability for in-memory processing. All processing on received events and triggering of new events happen in the runtime engine of each execution plan.

D. Output Event Adapter

The Output Event Adapter publishes events. Similar to the Input Event Adapter, this component is also a combination of different adapter implementations. It publishes events to the receiving server using various transport adapters. Database adapters are used to dump the data to relevant databases for future analysis.

E. Event Stream Manager

Event Stream Manager is an important component that manages the stream definitions. Users can add new stream definitions through this component. Event Stream Manager stores stream definitions in the registry of the server. Event Pre Processor and Event Processor components interact with Event Stream Manager to retrieve the information regarding the streams.

VIII. CONCLUSION

HFT has introduced a new paradigm in the capital markets and is now moving beyond equity markets to other financial markets like the currency market and derivatives markets. HFT presents a fantastic opportunity for big data analytics, more specifically for real time big data processing. Traditional stream based processing do not match up to the requirements of HFT as the domain has multiple event streams and a complete pattern may consist of events from multiple event streams.

We explore the possibility of using complex event processing to step into this area as it is suited on both fronts: understanding complex patterns across event streams and low latency processing. We also presented a theoretical framework based on CEP which could be used to build a HFT system. In future we would like to build a proof of concept for such a system.

REFERENCES

[1] Laney, Doug. "3D data management: Controlling data volume, velocity and variety." META Group Research Note 6 (2001): 70.

[2] Gomber, Peter, et al. "High-frequency trading." Available at SSRN 1858626(2011).

[3] Aldridge, Irene. High-frequency trading: a practical guide to algorithmic strategies and trading systems. Vol. 459. John Wiley and Sons, 2009.

[4] Brogaard, Jonathan. "High frequency trading and its impact on market quality." Northwestern University Kellogg School of Management Working Paper 66 (2010).

[5] Khashanah, K., Florescu, I., & Yang, S. (2014). High-Frequency Trading: A White Paper. from http://irrcinstitute.org/pdf/HFT_Practioner-Summary.pdf

[6] SEC, U. "Concept release on equity market structure." concept release(2010): 34-61358.

[7] Menkveld, Albert J. "High frequency trading and the new market makers."Journal of Financial Markets 16.4 (2013): 712-740.

[8] Avellaneda, Marco, and Jeong-Hyun Lee. "Statistical arbitrage in the US equities market." Quantitative Finance 10.7 (2010): 761-782.

[9] Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." Communications of the ACM 51.1 (2008): 107-113.

[10] Gualtieri, Mike, and John R. Rymer. "The Forrester wave™: complex event processing (CEP) platforms, Q3 2009." CEP (2009).

[11] Cugola, Gianpaolo, and Alessandro Margara. "Processing flows of information: From data stream to complex event processing." ACM Computing Surveys (CSUR) 44.3 (2012): 15.

[12] Cugola, Gianpaolo, and Alessandro Margara. "Low latency complex event processing on parallel hardware." Journal of Parallel and Distributed Computing 72.2 (2012): 205-218.