Benchmarking GPU-Accelerated Databases Against Traditional Databases

(Technical Paper)

A Sociotechnical Analysis of High Frequency Trading

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Paul Lee
Spring, 2020

Technical Project Team Members
Paul Lee

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Introduction

On May 6, 2010, a trillion dollars worth of economic value was wiped out from the economy, all within a span of 36 minutes (Kirilenko et al., 2017). Stock prices plunged at a rapidly in what seemed to be a near catastrophic series of events. Just as reminders of the Crash of 1929 filled the minds of traders, prices rapidly recovered. After this "flash crash," regulators and traders immediately began looking for the root cause of this event. The ability to influence market prices at this scale was unprecedented, and eventually, high frequency traders were found to be the perpetrators.

In a capitalism, every purchasable asset has some sort of financial representation, such as cash. Across centuries, financial instruments, from securities to bonds, have been formulated to model the complex dynamics of the financial world (Hu, 2011). Trading flourished around these financial assets, leading to the need for an exchange. An exchange is a platform through which buyers and sellers could buy and sell, confident that each party will meet a counterparty to trade with. The ability to fulfill these trades is called liquidity, a necessary property of an exchange (Britannica, 2018). today, billions of trades are made on a daily exchange (NYSE, 2019). However, the ability to match billions of trades coming in at multiple exchanges around the world is a technological challenge, a feat too difficult for humans to manually achieve. This need lead to the birth of high frequency trading.

These high frequency trading (HFT) firms utilize powerful software and computers to conduct massive volumes of trades at scale and performance. With the surge in computer technology, these firms have been able to gain tremendous influence over financial markets. Indeed, HFT firms were responsible behind the Flash Crash of 2010. Various algorithms detected queues in markets that triggered massive selloffs followed by massive buyouts. A

testament to their influence, the crash showed the ability for HFT firms to control and potentially manipulate prices. As such, for the STS topic, this paper will address how HFT firms gained infamy as predatory market participants. The dissonance between their image as market manipulators and their fundamental market-making role will be explored. Understanding the underlying societal and technological dynamics behind HFT firms will be critical, as their increasing presence in the financial world will be critical in capitalistic functions.

For the technical topic, this paper will explore a promising new technology that can be used to enforce regulations over HFT firms in order to prevent manipulative practices. HFT activity regularly produce terabytes of data on a daily basis (Aldrige, 2013) and thus, make it extremely hard to monitor and regulate. The sheer volume of activity HFT firms produce make it a challenge to regulate HFT activity. Databases, a critical component used in most modern computer systems, are used to enforce monitoring over HFT practices. Databases are used to store and retrieve data, which are critical in logging and tracking through trading activity. There are several types of databases, with each type being specialized for their use cases. GPU (Graphics Processing Unit) databases will be explored, a cutting-edge database that utilizes specialized hardware called GPUs, for their potential to be used in monitoring trading activity. GPUs are frequently used in computationally intensive processes such as graphics rendering and drug simulations (Peddie). It has become increasingly known that GPUs can be used to power databases and allow for extremely fast data ingestion and queries (Patrizio, 2018).

To reiterate, for the STS topic, this paper will explore on how HFT firms have taken on a "manipulator" reputation, despite their fundamental market making role. For the technical topic, this paper will explore a certain technological component that regulators can potentially use, GPU databases, and see how they can be leveraged to increase transparency among HFT and

enforce regulation. Addressing and understanding the fundamental role of HFT firms is important in justifying their value as market makers or price manipulators. In addition, having a functional technology to properly regulate HFT firms will be critical in deterring HFT firms from market manipulation and allow markets to benefit from HFT's market making activities.

Technical Topic

At a high level, databases perform two fundamental operations: read data and write data. Data is written to a database, where it is stored in computer memory. Data is read from a database, where data is written to a different source, such as a report. Databases are nearly universal in software systems. The storage and processing of data is a fundamental process, and databases enable this essential process. However, as the volume and complexity of data have grown, databases have constantly evolved to perform for increasingly intensive data operations. There are two general breeds of databases: OLTP (Online Transaction Processing) and OLAP (Online Analytical Processing). OLTP databases are involved in rudimentary, repetitive operations such as displaying a post on a forum. OLAP databases are used for internal analytical purposes, such as monitoring customer purchasing behavior over the past year (Alvi, 2019). Naturally, OLAP database operations are more complex and large.

There are several regulations across exchanges that require HFT firms to report trading activity. However, with the sheer number of trades they execute, reporting every activity they engage in is an engineering challenge. The sheer volume of data regulators must process and report can be impractical to sift through and analyze. In addition, regulators need to monitor trading activity in real time, an engineering challenge testing the limits of computational power (Easley, 2013). Querying and fetching terabytes of data is a timely and expensive process, as it is a computationally intensive workload. As such, HFT firms have been challenging to regulate.

OLAP databases are the database of choice for regulators as monitoring trading activity is an analytical process requiring complex insights and large volumes of data. With the computational limits of databases being tested, a new breed of OLAP database is needed to work with the sheer amounts of data HFT firms generate. Currently, databases utilized in HFT regulation have fallen behind HFT technology. Regulators currently use traditional databases used for lower cardinality datasets. As trading volumes increased, regulatory technology has not scaled up. Databases need to be able to make computations in parallel and utilize multiple CPU cores to ingest and analyze the incoming flow of data. Software architects argue for centralized and monolithic databases, which can handle heavy loads (Easley, 2013). However, with the distributed nature of stock market exchange servers, simply having a monolithic database will not help. Databases will need to be fundamentally changed, from the bottom up, to tackle the challenge of monitoring HFT firms for regulatory purposes. In addition, databases must be distributed in nature in order to work with collocated servers.

GPU databases show promise as the next generation of databases that will enable accurate regulation of the HFT sector. With the potential computational power they promise, GPU databases can make it possible to monitor a sector that has been infamous for lacking transparency. In this project, it will be determined whether or not GPU databases outperform traditional databases. GPU databases will be benchmarked against traditional databases to see if they can perform faster reads and writes for a wide array of operations. These operations will be run on large sample datasets, and the speed at which GPU databases and traditional databases perform will be measured. A comprehensive benchmark testing framework will be designed to holistically determine whether or not GPU databases outperform traditional databases.

STS Topic

The role of HFT firms in society can be navigated through the Social Construction of Technology (SCOT) framework. The SCOT framework is the framing of a technological phenomena in terms of its involved stakeholders, societal context, and technology. Using this framework, the interaction between market participants, regulators, traders, and these exchanges can be analyzed in a societal context, where the use of computational technology to engage in HFT is a phenomenon where humans leverage technology to achieve feats never previously conceivable. However, the SCOT framework is limited in that it superficially addresses the consequences of HFT in a broader context and merely emphasizes how HFT gained prominence. In addition, it avoids analyzing the deeper cultural values behind HFT as a technology, potentially disregarding the cultural elements behind HFT's infamy.

The SCOT framework is critical in seeing whether or not the practices of HFT firms are beneficial or harmful. To begin with, the SCOT framework will allow us to see how actors leverage technology to achieve a goal or make improvements on previous practices (Pinch, 1984). SCOT will help us understand the current state of high frequency trading and see all the involved participants and how they are affected. To elaborate, the utilization of HFT technology sits in the middle of two kinds of stakeholders: HFT traders and the counterparties who trade with HFT traders. In one way, both parties leverage and use the technological capabilities of these HFT servers. When a participant wants to sell or buy an asset, they are leveraging the liquidity provided by HFT firms so that they can have their orders executed and fulfilled. On the other end, HFT traders are taking on the role of fulfilling these orders through HFT technology. Both parties benefit. The traders make money, and the counterparty has their liquidity needs fulfilled.

Another STS framework that will be used to analyze HFT firms will be co-production. Co-production states how people leverage technology to generate knowledge and better understand the world around them (Jasanoff, 2004). Co-production draws upon the concepts of casuality loop, in which how each variable in a system essentially create and sustain each other. In terms of co-production, the game of trading is to best price an asset (CME, n.d.). In other words, those who make the most accurate pricing of any asset based on its fundamental value will make the most money. As such, trading is very much related to co-production because people are trying to discover the fundamental value of assets, a complex process. HFT firms, by providing liquidity, expedite the price discovery process because of liquidity. For HFT firms to trade in the first place, counterparty traders must exist, so by trading going on between these firms, both parties have their needs more fulfilled. However, with manipulative practices such as spoofing, where traders artificially inflate or deflate the price of an asset, the price discovery process is inhibited because prices are swung in arbitrary ways (Hendershott, 2012). In essence, the process of coproduction is inhibited. However, co-production can potentially be risky in the case of noetic flatness, where two parties, the HFT traders and their counterparties, influence each other equally. In such a scenario, the need for having a distinguishing boundary between these two groups diminishes, and the framework becomes ineffective in analyzing the individual involved parties behind HFT trading.

Research Question and Methods

This research will look into how market participants and regulators have grown distrustful of HFT firms. Addressing this will be important; with the role HFT firms play in today's capital markets, assessing the value they provide to society will have significant impacts to the dynamics of financial markets today. All market participants, from local store owners to

institutional hedge funds, will experience any ripples caused by HFT firms (Hendershott, 2013), as markets achieve increased levels of liquidity from their activity.

This topic will be analyzed using documentary research and policy analysis. Under the methodology of documentary research, data regarding how financial markets have been affected through HFT firms will be researched. Different sources of evidence will be organized and analyzed to ultimately support an interpretation of the value HFT firms provide. Through policy analysis, the current problems surrounding HFT firms will first be identified. Then, the context of the problem will be outlined before analyzing approaches that have been taken so far by regulators and exchanges toward HFT firms. The effectiveness of these policies in curbing market manipulation and increasing liquidity will be analyzed. Any downsides of these policies will also be considered in order to provide a holistic viewpoint of areas where HFT firms are questionable in their liquidity-providing functions.

Conclusion

For the STS deliverable, the following question will be addressed: how have HFT firms taken on a "manipulator" image and how have societal systems reacted? The immense power and secrecy of these firms have made regulators and market participants suspicious of HFT firms. As a result, much scrutiny has risen around these firms. Although HFT activity is a fundamental method of leveraging computational power to provide market liquidity, their immense power to influence markets have brought upon fear. As such efforts have been made to regulate their trading activity through computer systems that can track their activity.

However, due to the sheer scale of HFT activity, building software systems to properly regulate all activity is a challenge. Therefore, the technical portion of this project will attempt to determine if a promising new technology, GPU databases, could be utilized in collecting and

analyzing trading activity at scale and performance. In this project, a benchmark testing framework will be designed to compare the performance of GPU databases against traditional databases to ultimately determine if this novel database technology can provide a significant improvement in performance. Regulators could use this technology to monitor trading activity and detect manipulative practices, as current regulatory efforts lack the technological capacity to monitor HFT firms effectively (Aldrige, 2013).

References

Aldridge, I. (2013). *High-frequency trading: a practical guide to algorithmic strategies and trading systems*. Hoboken, NJ: John Wiley & Sons, Inc.

Emmer, P. C., & Gaastra, F. S. (1996). *The organization of interoceanic trade in European expansion, 1450-1800*. Aldershot, Great Britain: Variorum.

Alvi, A. (2019, October 10). OLTP vs. OLAP: Transactional Databases vs. Analytical Databases. Retrieved from https://datawarehouseinfo.com/how-does-oltp-differ-from-olap-database/.

Britannica. (2018, January 29). Liquidity preference. Retrieved from https://www.britannica.com/topic/liquidity-preference.

Easley, D., Mailoc López de Prado Marcos, & OHara, M. (2013). *High-frequency trading: new realities for traders, markets and regulators*. London: Incisive Media.

Hendershott, T., & Riordan, R. (2012). High Frequency Trading and Price Discovery. *SSRN Electronic Journal*. doi: 10.2139/ssrn.1928510.

Hu, J. C. (2011). *Asset securitization: theory and practice*. Singapore: John Wiley & Sons (Asia).

Jasanoff. (2004). Retrieved from https://sheilajasanoff.org/research/co-production/.

Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The Flash Crash: High-Frequency Trading in an Electronic Market. *The Journal of Finance*, *72*(3), 967–998. doi: 10.1111/jofi.12498

MMI. Improving Market Quality and Stability. (n.d.). Retrieved from https://www.modernmarketsinitiative.org/improving-market-quality-and-stability.

NYSE. (2019). Retrieved from https://www.nyse.com/data/transactions-statistics-data-library.

Patrizio, A. (2018, December 19). What a GPU-powered database can do for you. *InfoWorld*.

Peddie, J. (n.d.). Famous Graphics Chips: TI TMS34010 and VRAM. The first programmable graphics processor chip: IEEE Computer Society. Retrieved from https://www.computer.org/publications/tech-news/chasing-pixels/Famous-Graphics-Chips-IBMs-professional-graphics-the-PGC-and-8514A/Famous-Graphics-Chips-TI-TMS34010-and-VRAM.

Price Discovery - CME Group. (n.d.). Retrieved from https://www.cmegroup.com/education/courses/introduction-to-futures/price-discovery.html.

Rust, J., & Hall, G. (2002). Middlemen versus Market Makers: A Theory of Competitive Exchange. doi: 10.3386/w8883.