

Computación financiera: High-Frequency Trading and Graphics processing unit

Jonathan Antognini C., Luis Salinas C.
Universidad Técnica Federico Santa María. Valparaíso, Chile

Abstract—En la vida cotidiana de la mayoría de los inversores individuales, se toman estrategias de compra y retención a largo plazo. Esto porque en su mayoría no disponen tiempo o experiencia para este tipo de comercio. Sin embargo es conocido que esta área es muy lucrativa, y además existen profesionales en el área conocidos como *dealers*. En el presente las tecnologías permiten a los individuos que no están trabajando para algún tipo de firma, tener a mano este tipo de mercado. Pero esto no es novedad, este tipo de mercados ha encontrado una fuerte componente digital, con sistemas computacionales robustos, que permiten a los usuarios realizar trading remotamente e incluso programando robots. Los nuevos medios informáticos permiten trabajar con datos de alta frecuencia, como transacciones en intervalos de tiempo muy pequeño, lo cual se ha incluido como una nueva e importante área de este tipo de mercado. El objetivo de este documento es interiorizar al lector al respecto del *electronic trading*, algunos enfoques del market making, al concepto de High-Frequency Trading, y a la computación de alto desempeño.

Index Terms—Financial Market, Electronic Market Making, High-Frequency Trading, GPU, HPC.

I. INTRODUCCIÓN

A través de los años el hombre ha presentado un cambio radical en su nivel de vida; los conocimientos que él ha logrado acumular y aplicar han sido para su beneficio, y han cambiado radicalmente su modo de vivir. Existe una notable diferencia entre el hombre de hace unas cuantas décadas y el hombre moderno, tal diferencia se ha dado por el desarrollo de la ciencia, que está estrechamente relacionada con la innovación tecnológica. Por esta razón se amplía el contenido de cómo ha evolucionado la ciencia y la tecnología en el mundo, su origen remoto, los países que más han aportado en esta área y su respectiva utilización, bien sea para el desarrollo o la destrucción.

Como se sabe, la tecnología se está haciendo presente en todas y cada una de las áreas de investigación, como física, química, biología, computación, y lo que es de interés para este documento es el área de los mercados financieros.

En tiempos pasados, las bolsas de valores estaban compuestas por agentes que compraban y vendían bonos o activos, esto de forma presencial, y dejando una impresión para una persona externa, de un cierto nivel de euforia y descontrol. Sin embargo la tecnología ha hecho una irrupción en esta área, permitiéndoles inicialmente realizar las mismas acciones, pero desde plataformas computacionales online, lo cual, de un comienzo se pensó como: ahorrar espacio, recursos, facilidad de acceso, etc. Pero por otro lado, se dejó abierta la posibilidad de incluir ciertas facciones robóticas dentro del actuar del ser

humano, como caso sencillo, dejar un robot esperando cierto precio de compra, y que el robot venda o compre. Cuando se habla de robot, se refiere a un programa que permita tomar decisiones y ejecutarlas también.

El primer gran salto desde el punto de vista de trading, fue el término asociado a High-Frequency Trading HFT, que es la negociación de acciones y productos derivados por medio de herramientas automatizadas, son programas informáticos desarrollados por las entidades financieras en estrecha colaboración entre el departamento de IT -técnico- y gestión de activos -funcional-. La parte técnica entre otras muchas tareas debe procesar, parametrizar y mapear grandes cantidades de datos relacionados con la cotización de los activos. En función de su configuración pueden enviar miles de órdenes -compra y venta- sobre un activo a lo largo de la sesión bursátil. Su esencia es puramente cuantitativa ya que los algoritmos, matrices de datos y parámetros son tratados mediante herramientas de estadística avanzada. Las órdenes en el HFT's se mantienen por un breve lapso de tiempo, la entrada y salida del mercado es muy rápida, normalmente por debajo del segundo. El principal motivo de la inmediatez se debe a que no son posiciones estratégicas en el capital de la sociedad cotizada.

En el siguiente documento se formularán conceptos básicos para que el lector pueda comprender el área en donde se está desarrollando el tema. Además se realizará un nexo entre la computación de alto desempeño y este tipo de mercado.

II. MERCADO FINANCIERO

El mercado financiero es un espacio con marco institucional que permite poner en contacto a oferentes y demandantes para que efectúen transacciones financieras. La idea de mercado como foro organizado a la que acuden agentes económicos para efectuar transacciones queda reducida en el mundo financiero como las bolsas de valores.

El concepto de mercado financiero se utiliza en general para referirse a cualquier mercado organizado en el que se negocien instrumentos financieros de todo tipo, como por ejemplo, acciones. Además el espacio para generar estas interacciones no necesariamente debe ser físico. Por otro lado el negociar instrumentos financieros implica a grandes rasgos: intercambiar instrumentos financieros, y definir su precio. Por ende, estos mercados están basados en las fuerzas de oferta y demanda, ubicando a todos los oferentes en el mismo lugar para facilitarle la búsqueda a los demandantes.

Una de las razones que hace importante este tipo de mercado, es su funcionalidad, ya que permiten:

- Aumentar el capital, siendo esto uno de los casos favorables, ya que también hay probabilidades considerables de disminuir el capital.
- Comercio internacional, como en los mercados de divisas, por ejemplo Forex.
- Reunir a quienes necesitan recursos financieros, con los que tienen recursos financieros.

Se puede encontrar una amplia variedad de tipos de mercados financieros, clasificados por: activos transmitidos, en función a su estructura, según la fase de negociación. Para efectos de este trabajo, este estudio se centrará en los mercados de tipo activos transmitidos, más en específico mercados bursátiles.

A. Mercados bursátiles

Los mercados bursátiles están clasificados en los mercados de capitales, en donde se negocian activos financieros. Este tipo de mercado provee financiamiento por medio de la emisión de acciones y permiten luego el intercambio de estas. La aplicación más directa de este tipo de mercados, son las bolsas de valores, cuyo origen se remonta a finales del siglo XV en las ferias medievales de Europa.

Las bolsas de valores se pueden definir como mercados organizados y especializados, en los que se pueden realizar transacciones de títulos de valores por medio de intermediarios autorizados. Estas bolsas ofrecen al público y a sus miembros facilidades, mecanismos e instrumentos técnicos que facilitan la negociación de títulos de valores susceptibles de ofertas públicas, a precios determinados mediante subasta.

La principal función de las Bolsas de Valores implican proporcionar a los participantes información veraz, objetiva, completa y permanente de los valores y las empresas inscritas en la Bolsa, sus emisiones y las operaciones que en ella se realicen, supervisión de actividades.

Las componentes de este sistema son los activos, instituciones financieras cuya misión es contactar demandantes y oferentes en los mercados donde se negocian los diferentes instrumentos o activos financieros.

Dentro de los estudios de la economía, se habla de que este tipo de mercado es de competencia perfecta, es decir posee características como: [9]:

- Existe un elevado número de compradores y vendedores. La decisión individual de cada uno de ellos ejercerá escasa influencia sobre el mercado global.
- Homogeneidad de los productos, es decir, no existen diferencias entre productos que venden los oferentes.
- Transparencia del mercado. Todos los participantes tienen pleno conocimiento de las condiciones generales en que opera el mercado.
- Libertad de entrada y salida de empresas. Todos los participantes, cuando lo deseen, podrán entrar o salir del mercado a costos nulos o casi nulos.

B. Electronic trading

Al observar el ferviente crecimiento en las tecnologías de información y acceso a ello, se impulsó la creación de los mercados de trading electrónico. Los primeros mercados de este tipo vieron la luz entre los años 1989 y 1992, pero el principal salto fue que en el año 1997 se dispusieron API's para los usuarios. Si se observa como aumenta la tecnología y la accesibilidad a ella, era de imaginar los sucesos importantes que seguían en un futuro. Es así como el 2006 nacen las principales Electronic Communications Networks (ECN): New York Mercantile Exchange (NYMEX), New York Stock Exchange (NYSE), Bolsa de Valores de Santiago, etc. Un ECN es el término que se utiliza en los círculos financieros para un tipo de sistema informático que facilita el comercio de productos financieros fuera de las bolsas de valores. Ya en el 2007 con la web 2.0 y los smartphones del mercado, se crean sistemas de trading vía celulares y se perfilan como una de las mejores alternativas para los inversores individuales o minoristas.

Si bien es cierto, cuando se le habla a cualquier persona respecto a las Bolsas de Valores, lo primero que se le viene a la cabeza son una gran cantidad de personas comprando y vendiendo acciones, y llamando por teléfono de forma desenfrenada. Y en realidad los sistemas en la actualidad han ido dejando de funcionar así por algunas razones como: Falta de escalabilidad, dificultad para ampliar la base de clientes, insatisfacción del cliente, costos operativos, complejidad de integración, costo de desarrollar tecnología internamente, etc. Existen bastantes razones por las cuales este tipo de sistemas se vio destinado a desaparecer en el tiempo, dejando la puerta abierta a los sistemas electrónicos.

Las características de la digitalización de los mercados mundiales fue impulsada por los siguientes factores:

- La globalización, expansión de redes de telecomunicaciones y con acceso fácil a la información.
- Nuevas tecnologías disponibles, más confiables, más seguras y más fáciles de usar.
- Los mercados y las empresas de corretaje necesitan crecer y captar nuevos clientes para no perder mercado frente a la competencia.
- Los clientes demandan mejor calidad de los servicios a costos más bajos.
- Hay una necesidad de contar con mercados más transparentes, más confiables, escalables, sin cuellos de botella y que permitan ser auditados fácilmente.

C. Definición

En este documento se abordará el concepto de Electronic Market Maker. Para entender el concepto es necesario definir algunos términos: [6]

- Dealers: compran y venden "activos" o "valores". Lo que buscan es comprar bajo y vender alto.
- Se fijan dos precios: **Bid** (dealer compra, cliente vende). **Ask** (dealer vende, cliente compra). Lo esperado es que: $ask > bid$. El término **spread** corresponde a $ask - bid$.

lo que corresponde a la diferencia entre lo que el dealer vende y compra (concepto de ganancia). Esto se entiende como la fuente de ingresos de los dealers. [8]

- Market order: es una instrucción del cliente al dealer (comprar o vender al mejor precio posible). Asegura la realización de la transacción, pero no el precio.
- Limit order: es una instrucción del cliente al dealer de pedir una transacción a un precio específico (o más ventajoso). No garantiza la realización de la transacción pero si se conoce una cota del precio.
- Existe un order book, el cual es un registro de dos colas: una de venta y una de compra. Si una nueva orden entra y no es satisfecha por el limit order, esta ingresa al order book (cola de venta o compra).

La definición de estos conceptos son recurrentes en la literatura, y son la base con la que los investigadores se expresan.

D. Áreas de investigación

El Electronic Market Maker se descompone en dos áreas grandes de investigación:

- Establecer el bid-ask spread.
- Actualizar el bid-ask spread, de manera predictiva o no predictiva.

La primera decisión para el market maker es donde establecer el spread inicial. Es decir encontrar la valoración del activo al ser negociado:

- Para un stock de cierto activo, determinar su valor para la compañía. Si el activo corresponde a un bono, buscar el valor presente del pago prometido.
- Si no hay mercado establecido, o el mercado tiene poca liquidez, entonces la valorización es la única aproximación. [10]

La segunda, actualizar el bid-ask, es considerado como el corazón del market maker. Existen formas predictivas al corto plazo teniendo en cuenta el comportamiento de la curva (bid-ask), esto es posible cuando se tienen valores históricos, y en base a ellos se pueden calcular aproximaciones a los nuevos valores posibles. El segundo caso posible, es que no se posea información histórica, y que se tenga la información actual del mercado. Estas últimas son inherentemente más simples.

En el estudio de la actualización del bid-ask mediante formas predictivas es donde se utilizan la mayor parte de procedimientos matemáticos. Además en el presente se trabaja con datos de alta frecuencia, es decir, data proveniente de los servidores con diferencia temporal muy pequeña, lo que deja abierta la posibilidad de realizar investigación tanto matemática como computacional en la optimización de estos cálculos.

III. GPU COMPUTING

Las siglas GPU provienen de Graphics Processing Unit, o en español Unidad de procesamiento gráfico. Para efectos de hardware, la GPU funciona como coprocesador, pudiéndose utilizar de forma simultánea a la CPU y así aprovechar el

potencial que puedan ofrecer ambas al mismo tiempo. Una GPU es un procesador diseñado para llevar a cabo cálculos necesarios implicados en la generación de gráficos, ya sea, para un video juego, o como para una aplicación que utilice gráficos en 2D, o 3D. Hoy en día las GPU son muy potentes y pueden incluso superar frecuencias de reloj de una CPU antigua.

Una GPU está formada por cientos de pequeños núcleos que trabajan juntos para procesar los datos de alguna aplicación. Esta arquitectura de procesamiento paralelo masivo es la que proporciona al GPU su alta capacidad de cálculo. Existen numerosas aplicaciones aceleradas en la GPU que brindan una forma rápida de acceder a la computación de alto desempeño (High Performance Computing).

El concepto implícito en todo este tema es el paralelismo, que es una forma de computación en la cual varios cálculos pueden realizarse simultáneamente, siempre y cuando no existan dependencias secuenciales entre cálculos. Basándose en "divide y vencerás", principio que busca dividir los problemas grandes, para obtener varios problemas pequeños, que son posteriormente solucionados en paralelo.

La evolución de las tarjetas gráficas ha venido acompañado de un gran crecimiento en el mundo de los videojuegos y las aplicaciones 3D, realizándose grandes producciones de chips gráficos por parte de grandes fabricantes, como NVIDIA, AMD (ex ATI).

En los últimos años también han aparecido conjuntos de herramientas y compiladores que facilitan la programación de las GPUs, como por ejemplo, NVIDIA CUDA, que cuenta con la comunidad más activa hasta la fecha en programación de GPUs.

A. NVIDIA CUDA

Las primeras GPU fueron diseñadas como aceleradoras de gráficos y admitían apenas procesos específicos de funcionamiento fijo. En las últimas dos décadas, el hardware cada vez se volvió más programable, lo que culminó con la primera GPU de NVIDIA en los años 1999. En poco tiempo que se desarrollara el concepto de GPU, investigadores empezaron a utilizar el rendimiento de estas tarjetas en cálculos con punto flotante.

Los dilemas que vivieron en futuro los programadores, fue que la programación para GPU estaba lejos de ser fácil, hasta que investigadores de la universidad de Stanford se propusieron reimaginar la GPU como un coprocesador de flujos.

Puesto que NVIDIA sabía que su hardware era bueno e iba creciendo muy rápido, debían combinarse con herramientas de hardware y software intuitivas, invitaron a un equipo de investigación y desarrollo, para empezar a evolucionar una solución que eyecturara el lenguaje de programación C a la perfección en el GPU. Al reunión software y hardware, NVIDIA lanzó al mercado CUDA en el año 2006. La competencia AMD tuvieron intentos forzosos en generar algo similar, pero su comunidad de desarrollo no tuvo la misma motivación que si tuvo NVIDIA CUDA.

CUDA es una plataforma de computación paralela y un modelo de programación creado por NVIDIA. La tecnología implementada por NVIDIA, es un entorno basado en el lenguaje C, que permite a los programadores escribir software para resolver problemas computacionales complejos en menos tiempo aprovechando la gran capacidad de procesamiento paralelo de las GPU multinúcleo. Miles de programadores están utilizando las herramientas gratuitas de desarrollo de CUDA (válidas para millones de GPU que circulan en el mercado), a fin de acelerar todo tipo de aplicaciones, desde herramientas de codificación de audio y video, diseño de productos, investigación científica, etc.

Actualmente NVIDIA ofrece un kit de herramientas de CUDA, las cuales incluyen un compilador, bibliotecas de matemáticas, herramientas para corregir y optimizar rendimiento de aplicaciones. Encontrándose también con muestras de código, guías de programación, manuales de usuario, referencias de la interfaz de programación de aplicaciones (API) y otra documentación para ayudar al usuario dar los primeros pasos en el área. Cabe destacar que ofrece todo esto de forma gratuita, incluyendo NVIDIA Parallel Nsight for Visual Studio, el primer entorno de desarrollo del sector para aplicaciones masivas paralela que usan tanto GPU como CPU, esto en sistemas operativos Windows.

IV. OBJETIVOS

Uno de los conceptos que ha incurrido actualmente en el área, es el High-Frequency Trading (HFT), en español corresponde a negociación de alta frecuencia. Esto es un tipo de negociación que se lleva a cabo en los mercados financieros utilizando intensamente herramientas tecnológicas sofisticadas para obtener información del mercado y en función de la misma, intercambiar valores financieros como activos (vender o comprar). Las características principales:

- Su componente cuantitativa involucra algoritmos informáticos para analizar datos del mercado e implementar estrategias de negociación
- Cada posición de inversión se mantiene solo durante breves períodos de tiempo, para rápidamente ejecturar la posición y comprar o vender, según convenga, y el activo de que se trate. En algunas ocasiones esto se realiza miles de veces en un solo día. Los períodos de tiempo en los que se mantienen las posiciones pueden ser de solo fracciones de segundos.
- Suelen llevarse a cabo por importantes salas de mercado, como fondos de inversión o bancos de inversión con carteras de activos de gran volumen y diversificadas.
- La característica clave y de interés, es que este tipo de neogicación es muy sensible a la velocidad de procesamiento del mercado y al propio acceso al mercado.

Las oportunidades que se presentan en esta área son importantes, debido que los defensores del HFT destacan que la aplicación de la tecnología da liquidez al mercado, abarata costos y contribuye a eliminar posibles ineficiencias en la formación de los precios.

El HFT es una especialización de los Algorithmic Trading, que es la práctica de operar en los mercados financieros siguiendo unos algoritmos predefinidos. De este modo las operaciones se realizan por parte de máquinas conectadas a los mercados financieros en vez de por traders tecleando. Actualmente existe una competición entre los High Frequency Traders para tener las máquinas más avanzadas y las conexiones más rápidas. Por supuesto esto se tiene que aplicar utilizando el software más rápido y avanzado del que sean capaces de diseñar. Es este motivo el cual hace nexo entre los mercados financieros y la computación de alto desempeño.

Por lo tanto, lo que se pretende buscar es alguna aplicación de alguna estrategia de HFT bajo algoritmos hechos en HPC.

V. APLICACIONES PREVIAS

A. High-Frequency Trading

En las estrategias predictivas, se han realizado varios modelamiento mediante herramientas que proveen las matemáticas: [4], como Series de tiempo. Una serie temporal es una secuencia de datos medidos en determinados momentos de tiempo y ordenados cronológicamente, por lo general, los datos están espaciados entre si de manera uniforme, pero para efectos de aplicación en datos financieros, esto no es una realidad, debido a que la actualización del bid/ask no es estándar en tiempo. El análisis comprende métodos que ayudan a interpretar este tipo de datos, extrayendo información representativa, tanto referente a los orígenes o relaciones como la posibilidad de extrapolar y predecir comportamientos futuros. En [5] el autor desarrolla e implementa una técnica para estimar el modelo del bid/ask, y el valor del spread. El spread lo descompone en dos grandes componentes, uno debido a la asimetría de la información y el otro debido a los costos de inventario. La data utilizada la obtuvieron de la Bolsa de Nueva York, y contaron con los precios de transacción de acciones comunes en el período 1981-1983.

Si se busca en internet respecto a High-Frequency Trading, se pueden ver muchos artículos relacionados a una fecha en específica del mercado de EEUU, 6 de mayo del 2010. Fecha en la cual se acuña el término de *Flash-Crash*, en [2] definen tanto el comercio algorítmico y el HFT, y su intención es informar al público los responsables políticos y los reguladores, para proporcionar una base para las discusiones y las necesidades de regulación.

Respecto a lo anterior se puede escribir bastante, debido a que existen detractores (como el paper citado anteriormente), mientras que otros autores no implicitan su postura debido a que no poseen tanta certeza respecto a lo ocurrido [7].

Los principales aportes científicos que realizan los investigadores del área, están ligados a sus definiciones de estrategias de HFT y sus algoritmos utilizados. En [1] definen pautas para la clasificación de estrategias:

- Provisión de liquidez automatizada: algoritmos cuantitativos para optimizar el precio y ejecución de la posición del market making. Tiempos alrededor a 1 minuto.
- Microestructura de trading de mercado: la identificación de participantes en la negociación, y el flujo de órdenes

a través de ingeniería reversa de las cotizaciones observadas. Tiempos alrededor a 10 minutos.

- Eventos de trading: macro eventos de trading. Tiempos alrededor de una hora.
- Desviaciones arbitrarias: desviaciones estadísticas del equilibrio, triangulos de trades, bases de trades, y cosas por el estilo. Tiempos alrededor de un día.

Este libro se considera como uno de los más importantes en el área, y contiene bastante información relevante de estudio.

El trabajo actual/futuro presenta desafíos para varios sectores de la ciencia, y algunos puntos están centrados en: [3]

- 1) Generalizar el problema del market maker para:
 - Stock multiples.
 - Opciones múltiples.
- 2) Problemas en donde el market maker pueda:
 - Ajustar cantidades en el bid y en el ask.
 - Ejecutar órdenes estratégicamente.
- 3) Modelamiento de selección adversa:
 - Salto entre precios de stock.
 - Correlación entre el retorno del stock y la posición del inventario.
 - Autocorrelación en la marca del market order (o limit order).

VI. CONCLUSIONES

Con la investigación realizada, se puede observar un amplio espectro de posibilidades de desarrollo. Sin embargo la memoria de pregrado se centrará en unir los mercados de High-Frequency Trading (con datos reales) con GPU. La razón de querer realizar esto, es la capacidad que ofrecen estos dispositivos, sus ventajas comparativas frente a cualquier programa que pueda ser procesado de forma secuencial en una CPU. A grandes rasgos la visión que se tiene para este problema es buscar una cartera de acciones, con su secuencia histórica de valores (bid-ask), y el objetivo es intentar realizar una buena extrapolación para los siguientes precios del activo en el mercado. Como se piensa trabajar con una cartera de acciones, lo más lógico sería procesar cada acción por separadas, y asociar un riesgo de compra venta por cada acción (dependiendo de factores de importancia). Como se trabajará en mercados de alta frecuencia de cambios de precios, una fracción de segundo es importantísima, por lo que se buscará implementar una metodología eficaz y eficiente, ya sea en el modelado del problema, como también en la implementación posible.

El marco de esta memoria de pregrado incluye como profesor guía al PhD. Luis Salinas, y el centro de investigación de la universidad CTI-HPC, que gracias a sus nexos y contactos en el área, conseguirán datos de prueba. Por otro lado se pretende utilizar el cluster que cuenta el centro de investigación y ahí realizar pruebas.

Actividad Realizado	Tiempo utilizado
Lectura documentos previos	20 horas
Preparación del documento (formato, diseño, etc)	1 hora
Desarrollo del documento*	14 horas
Revisión Ortográfica y redacción	2 hora

*Al momento de leer cada documentos e realizaba un resumen pertinente que facilitaba la labor de redacción del documento.

REFERENCIAS

- [1] I. Aldridge. *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*, volume 459. Wiley, 2009.
- [2] B. Arndt, M. Lutat, T. Uhle, and P. Gombert. High frequency trading. 2011.
- [3] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.
- [4] E.M. Azoff. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.
- [5] L.R. Glosten and L.E. Harris. Estimating the components of the bid/ask spread. *Journal of financial Economics*, 21(1):123–142, 1988.
- [6] L.R. Glosten and P.R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100, 1985.
- [7] X. Guo. High-frequency trading. 2012.
- [8] T. Ho and H.R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics*, 9(1):47–73, 1981.
- [9] N.G. Mankiw. *Principles of economics*. South-Western Pub, 2011.
- [10] D.J. Seppi. Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies*, 10(1):103–150, 1997.