Cinnober®

A Cinnober white paper

Using Adaptive Micro Auctions to provide efficient price discovery when access in terms of latency is differentiated among market participants

Lars-Ivar Sellberg, 20 October 2010

Introduction

After the "flash crash" of May 6 2010 – during which the Dow Jones Industrial Average (DJIA) dropped briefly by almost 10% – high frequency trading (HFT) is being scrutinized by regulators. Voices have increasingly been raised regarding its potential negative effects and the fragility of today's markets.

Suggestions for how to avoid future "flash crashes" and recreate confidence in financial markets have mainly aimed at dealing with the symptoms: increasing market surveillance, imposing fines, introducing circuit breakers, message throttling and regulating a minimum "time in force" for orders. At Cinnober, we believe that there is an opportunity for some marketplaces to differentiate themselves by taking a whole new approach.

Exchanges globally are investing tremendous resources in data centers, offering hardware co-location for HFTs to cut transaction latency to an absolute minimum. In more recent developments however, some exchanges are seeking ways to attract the more long-term investors, such as asset managers and pension funds, by giving incentives for volume orders. This to lure back liquidity that has moved to dark pools as the "lit" markets have been increasingly fragmented with smaller trade sizes.

In April we published a white paper about the impacts of algorithmic trading – including HFT – on marketplaces and ways to accommodate it in an efficient manner.[1)]

In this paper we present an alternative method of handling non-constructive high frequency trading. Instead of treating the symptoms, we address the root cause directly, by changing the underlying market model to become insensitive to an environment where latency is differentiated among members.

At Cinnober, we call this new approach **Adaptive Micro Auctions (AMA)**. The AMA market model may appeal to marketplaces who want to offer their customers equal ability to act on new information, regardless of what latency conditions they are working with. Using this model, many of the symptoms that so many are trying to cure in today's financial markets won't even arise.

A market venue deploying such a market model is likely to attract liquidity that would otherwise be directed to dark pools.

The latency advantage

Market participants accessing an electronic marketplace experience two main types of latency:

1. Market data latency: the time it takes for market data disseminated by the marketplace system to reach the market participant.
2. Order entry latency: the time it takes for a market participant to enter an order on the marketplace and receive an acknowledgement of its initial matching status.

The source of both of these types of latency is multi-faceted and is affected by a number of factors, such as geographical location of the market participant, HW/SW deployed by the market participant, HW/SW deployed by the marketplace operator, etc.

In many cases both types of latency tend to be differentiated among market participants, implying that some market participants systematically have faster access to a marketplace than others. Having faster access to a marketplace system than competitors is generally considered to be an important advantage.

For example, consider an order entered onto a marketplace that offers some sort of arbitrage opportunity. The market participant who first receives this information (latency type 1 above) is able to act on it (latency type 2 above) and exploit the arbitrage opportunity. There are no prizes for the runners up in this type of game.

[1)] *"Algorithmic trading and its implications for marketplaces" can be downloaded at: http://www.cinnober.com/white-papers.*

Some of these factors affecting latency can be directly influenced by the market participants themselves, for example by:

- Locating their trading systems as close to the marketplace as possible, with proximity hosting or co-location in the same data center.
- Using the fastest available application protocol: most marketplaces offer a standard FIX-based protocol and a faster native protocol.
- Using superior SW/HW. Examples include in-house-developed network stacks that are faster than those commercially available.

The items listed above all have one thing in common; they demand significant investments, hence creating barriers to entry or equal competition conditions. In other words, it is possible for market participants to effectively buy an important advantage over competitors with less money to spend.

There are two ways of looking at this:

- *It is a positive thing* that encourages competition; thereby it increases liquidity and removes arbitrage. A marketplace should encourage it by offering services such as co-location of market participants' trading systems with the marketplace system, etc.
- *It is a negative thing*; by not offering equal access to all market participants, a marketplace fails to provide a fair market. Market participants not able to spend enough money will systematically be penalized by the ones that can. Followers of this idea usually point out that the phenomenon of flash trading is just a structured way for marketplaces to charge selected market participants in return for trading advantages, without taking a detour with complicated technical solutions like providing co-location services, etc.

The purpose of this paper is to show how very short-lived auctions – called Adaptive Micro Auctions (AMA) – measured in milliseconds (ms), can be used to provide a fair market with an efficient price discovery process, even in situations where latency is highly differentiated.

Thus for marketplace operators subscribing to the idea that the ability of market participants to buy faster access is a negative thing, such short-lived auctions might constitute an alternative to traditional continuous trading.

AMAs can also be useful in markets with a geographically outspread client base where transaction latency differs substantially between market participants depending on how close to the marketplace they operate.

Furthermore, AMAs are also an alternative to circuit breakers.

We start by looking at traditional auctions since many of the basic mechanisms of the AMA are taken from there.

Traditional auctions

Auctions in their traditional long-lived form – usually measured in minutes – are widely used as means of achieving efficient price discovery. They work by collecting buy and sell orders over a period of time, typically 5-15 minutes, and calculating an equilibrium price based on these orders. Once the equilibrium price has been established, all orders eligible for execution at the equilibrium price are executed at that price.

There are countless variations on the exact formula used to calculate the equilibrium price and how to distribute volume on the short side of the auction. The point is that an auction offers an efficient price discovery mechanism, which is not sensitive to an environment where latency is differentiated among market participants. Most high-frequency algorithms are effectively rendered useless in an auction-based environment.

In the arbitrage example given in the previous section, the market participant with the fastest access has no advantage over other participants, assuming the following:

- The auction allocation algorithm does not take into account in which sequence orders were entered during the auction phase.[2)]
- The proper anti-gaming procedures have been implemented, making sure that market participants cannot predict the exact execution time of the auction. Otherwise a market participant could wait until just before the execution time to enter his order, not giving other participants enough time to react to the new information by modifying their own orders.[3)] Note that anti-gaming procedures are only needed for open auctions, that is, auctions with some sort of market transparency.

[2)] *There are several standard algorithms which do not; an example being pro-rata.*
[3)] *There are a number of known and practiced anti-gaming procedures. Among the more well known are to simply randomize the exact execution time of the auctions.*

The above is well known within the financial industry and auctions are widely recognized as efficient price discovery mechanisms. So why haven't auctions replaced continuous trading on a large scale? One major reason is most likely that market participants put a high value on execution immediacy; they do not want to wait several minutes for the order to be executed. Auctions are therefore primarily used to establish an equilibrium price in the opening and closing of a market.

The next section presents the AMA concept and its use in providing a fair and efficient price discovery process in an environment with highly differentiated market access in terms of latency, while still offering near-immediate execution.

Adaptive Micro Auctions

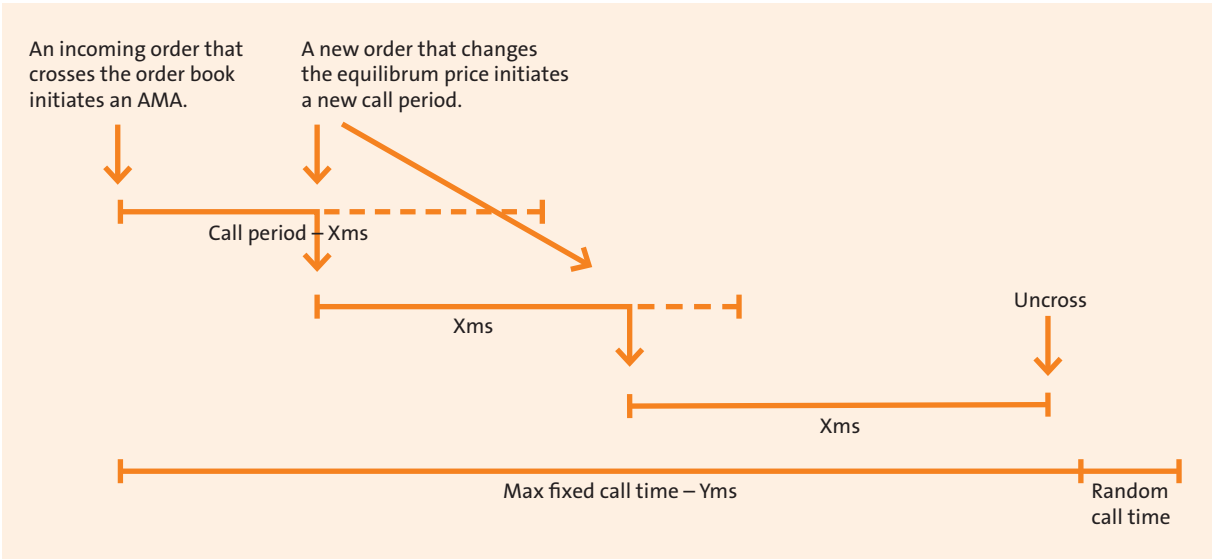
The idea behind Adaptive Micro Auction (AMA) is to preserve most of the positive aspects of auctions while offering near-immediate execution of continuous trading. The obvious solution is to decrease the running time of the auction to a level where executions are deemed immediate from all practical perspectives by market participants. This does, however, present several practical problems.

It is assumed that AMAs are transparent and at least disseminate the current equilibrium price, that is, the price at which the auction would execute if triggered at that point. But an AMA might very well be more transparent and disseminate market-by-level or even all individual orders.[4)]

Assuming that only the equilibrium price is being disseminated, the technical cost for disseminating market data in terms of bandwidth and hardware is drastically reduced. In high order flow environments, these cost savings can be significant for the marketplace as well as for the market participants.

In order to minimize the running time of the auction while still making sure that market participants have received all information relevant to their decision making, the AMA consists of the following basic steps:

1. When an order eligible for execution (that is, it crosses the book) is received, the current equilibrium price is calculated and disseminated.
2. A call period of X ms is initiated. If an order that changes the equilibrium price is received during this phase, the new equilibrium is disseminated and a new X ms call period is initiated. The total time spent in step 2 can however never exceed a randomized outer boundary of length Y ms.
3. The auction is executed.



[4)] Note that that it is generally considered bad practice to disseminate price levels better than the equilibrium price since this would discourage market participants from entering orders reflecting their true intention to trade. Instead these price levels are normally aggregated to the equilibrium price if, for example, the marketplace is disseminating market-by-level.

The idea is to configure X to be the sum of latencies of types 1 and 2, as defined in the introduction. This implies that all market participants have time to receive information about a newly created equilibrium price and respond to it by modifying, canceling or inserting new orders. If these new order amendments generate a new equilibrium price, then the process re-initiates. The randomized outer boundary is needed in order to prevent auctions from running too long if the equilibrium price keeps changing.

The outcome is an auction that runs just long enough to make sure that all market participants have received relevant information and have had the opportunity to act upon it, hence the name "Adaptive". Once this has been achieved the auction is executed.

A typical value for X might be in the range of 5-10 ms while Y might be 40-50 ms, that is, near-immediate execution of orders is achieved while still providing efficient and fair price discovery.

The basic mechanisms of an AMA are well known from traditional auctions. What is new and makes it adaptive is the timescale on which they are applied.

Note that there are several alternatives to using the changed equilibrium price as evaluation criteria in step 2 of the AMA. Some conceivable ones are:

- A changed imbalance in the auction.
- Whenever a new order is received.

Both of these alternative evaluation criteria are more sensitive than the equilibrium price, so using these, auctions are more likely to be repeatedly delayed in step 2. The benefit of this is that market participants get the opportunity to act on information they might find valuable; even though an order does not change the equilibrium price, it might constitute information to another market participant. The disadvantage is of course that auctions will run for longer and are more likely to hit the randomized outer boundary. Using the equilibrium price as evaluation criteria is a compromise between sensitivity and auction running time.

Larger trades and increased liquidity

In addition to creating a fair and efficient price discovery process, other potential positive effects of Adaptive Micro Auctions (AMA) are increased trade sizes and liquidity. Arguably, both deregulation such as MiFID and high-frequency trading have driven down trade sizes and liquidity. The reasoning behind this is usually as follows:

- Deregulation has an effect since competition between an increased number of marketplaces have provided incentives to decrease tick-sizes to levels where the volume on each price level is very low. This forces a large order to "eat" through a number of price levels, generating a large number of trades, in order to fully execute.
- High-frequency trading has an effect since institutional market participants are reluctant to have orders residing in the order books due to the risk of information leakage and sniping activity.

AMAs can be used to counter these negative effects to a certain extent. First of all, an AMA concentrates liquidity in time: all orders entered during the auction phase are executed at once, at a single equilibrium price. This will help decrease the number of trades and increase liquidity.

Regarding information leakage feared by institutional investors, the AMA can be configured to only disseminate a limited amount of market data, such as the current equilibrium price. With regard to sniping, the problem never occurs when using an AMA, since all market participants have time to react to new incoming orders. Indeed, this is the core contribution of AMAs.

A dynamic alternative to circuit breakers

Circuit breakers are temporary trading curbs that are imposed when a financial instrument, such as a stock, has moved rapidly in value. Trading is restarted, typically using an auction, after a "calm down" period. Usually the interesting case is when an instrument has dropped in value, as we discuss below.

The purpose of a circuit breaker is to prevent any excessive, usually downward, movement in value of the traded instrument. The assumption is that such excessive movements are caused by a temporary lack of liquidity in the order book. Once market participants have been given a chance to examine and evaluate the current prices in the order book, additional liquidity will be made available from them in the form of new buy orders. The idea is that the temporary trading curb introduced by the circuit breaker gives market participants the time they need to examine and evaluate the situation.

But there is a drawback to circuit breakers that is often neglected. If market participants know that trading might be curbed for an extended period due to a circuit breaker, they might increase their risk premiums; they may lower the price they are willing to pay for an instrument. This by itself will cause the instrument to fall further. With the rise of high-frequency trading, it is likely that market participants have already implemented automated algorithms such "we are 1% away from a circuit breaker imposed trading curb, sell the instrument", which will further accelerate the fall of the instrument.

This implies that the choice of length of the trading curb used in a circuit breaker mechanism is important since the benefits from giving market participants time to re-evaluate a situation needs to be weighed against the increased risk premium. Current implementations of circuit breakers typically use time periods of fixed length, chosen because they are presumed to be a good compromise. A typical value is 5-30 minutes.

There are two problems with fixed time periods measured in minutes.

- In general, a period measured in minutes is extremely long in a trading environment where the time scales have been shrinking rapidly with the rise of high-frequency trading and a generally increased trading tempo.
- Using a fixed period implies that it is possible to find an optimal time for all situations. A potential drop of 15% during high turnover must be treated in the same manner as a drop of 5% during light turnover. It is unlikely that a trading curb period will be found that is optimal for all cases.

The concept of Adaptive Micro Auctions (AMA) offers an attractive alternative to traditional fixed length circuit breakers, given that the AMA parameters X and Y are made a function of the circuit breaker triggering condition.

Instead of curbing trading altogether when the circuit breaker limit is reached, the length of the AMA parameters are simply dynamically increased. That is, a drop of 5% yields AMA parameters X and Y while 6% yields X+n and Y+n and so on.

Note that the function can be made more complex and take other parameters, such as turnover, into account. While the example above uses discrete steps, this is not a prerequisite; it might just as well be a continuous function. This means that a function can be found that dynamically adjusts the trading "curb" time to the optimal length for each situation.

For most marketplaces the timescale for the X and Y parameters used in conjunction with circuit breakers should be measured in seconds and milliseconds, not minutes.

Using dynamic AMAs instead of circuit breakers to limit excessive downward market moves gives a smooth slow-down of trading instead of an abrupt stop, with all the consequent disruption it implies.

Conclusion

An Adaptive Micro Auction (AMA) is a concept that utilizes the good qualities of traditional auctions: fair and efficient price discovery with near-immediate execution.

Using the AMA model establishes an environment in which market participants are given equal ability to act on new information, regardless of which latency conditions they are working with, relieving investors of the fear of their orders being sniped by market participants who have faster access to new information. Using this model, many of the symptoms we're trying to cure in today's financial markets won't even arise.

Adaptive Micro Auctions is one way for marketplaces to differentiate themselves in today's competitive and fragmented markets and attract customers who otherwise would direct much of their liquidity to dark pools.

Passion for change | Cinnober is the world's leading independent provider of innovative marketplace and clearing technology. Our solutions are tailored to handle high transaction volumes with assured functionality and low latency.

We are passionate about one thing: applying advanced financial technology to help trading and clearing venues seize new opportunities in times of change.

Among our customers are leading exchanges such as the Chicago Board Options Exchange, the London Metal Exchange and NYSE Liffe. We also power new initiatives and alternative trading systems such as Alpha Trading Systems, Burgundy and Markit BOAT. Our clients rely on our platform-independent, Java-based technology for leveraging change quickly and cost-effectively.

Cinnober is headquartered in Stockholm and employs 170 people together having more than 1500 man years of experience from developing exchange and post-trade systems.

We are an independent technology provider and do not operate a market of our own, avoiding any conflicts of interest. We are not owned by — nor have any ownership interests in — any market operator.

Our track record says it all. We help our customers turn change into a competitive advantage.

Cinnober®

Cinnober Financial Technology AB
Kungsgatan 36
SE-111 35 Stockholm
Sweden
+46 8 503 047 00
cinnober.com