Impact of electronic trading on the evolving market and exchange structure

FIX market data optimization

By Matt Simpson (left), Associate Director, Electronic Trading Architecture, Chicago Mercantile Exchange Inc. (CME) and FIX Protocol Committee member & Richard Shriver, Industry Consultant, Jordan and Jordan

The relationship of market data to electronic trading

As electronic trading continues to expand and customers choose this venue over the trading floor, market data has become more important than ever. The trend toward electronic trading is creating an explosion in market data volumes and peak message rates which challenges even the most technologically savvy distributors and other direct consumers. The aggressive migration toward all-electronic markets and increased trading efficiencies is primarily responsible for this rapid upward spiral in market data volume and automation. While the pervasiveness of electronic trading is a tremendous stride forward it also presents equally tremendous challenges in terms of market data dissemination.

At one time, market data was conveyed by the alacrity of pit traders communicating bids and offers via hand signals and shouting, which were subsequently posted on the edifice-like wallboards of the trading floors

At one time, market data was conveyed by the alacrity of pit traders communicating bids and offers via hand signals and shouting, which were subsequently posted on the edifice-like wallboards of the trading floors. In this model, market data volume and subsequent dissemination was not a concern. The use of automation and communications networks has expanded and today this market information must be transformed into bits and bytes and transmitted electronically in the form of market data feeds. Market data has evolved into a lifeline for electronic trading platforms as market participants around the world make critical decisions on real-time information.

With the explosion of market information volumes and peak message rates, there is a renewed need to seek improved efficiencies in the dissemination of market data from the trading venues all the way to the desktop. While a broadcast-based model may alleviate some of the redistribution requirements that tend to create bottlenecks, the core infrastructure is under increasing pressure to respond. Recognizing the significant opportunity to address a serious industry issue in advance of potential problematic consequences, FPL formed the FIX Market Data Optimization Work Group.

The objective of the market data optimization working group

In January 2005, the FIX Market Data Optimization Work Group was formed to research and develop market data technology standards that keep pace with and support the needs of electronic trading. The group's charter is to formulate comprehensive and optimal solutions covering all aspects of the market data paradigm necessary to support efficient delivery and management of market data including:

- Business Model Support
- Book Management
- Data Representation
- Multicast Dissemination

The high-priority objective of the working group is to lay the groundwork for a flexible approach that can be implemented in a standard way across the industry. An industry-standard interface for market data has the potential to create many synergies and reduce costs to all participants. At this time, a market data consumer may need to support a myriad of different interfaces. The emergence of a standard would offer the potential to drastically improve efficiency in responding to market information needs.

A notable strength of the working group is its diversity which draws participants and representation from all types of firms participating in all major markets (including Equities, Equity Options, Listed Derivatives, Foreign Exchange and Fixed Income) across the US and European communities. The convergence of this diverse group is producing a solid market data solution, as well as a common model that can be adopted across a wide array of markets.

Through a concerted focus on the four key areas listed above, the market data optimization working group is forging ahead with development of new standards that will be critical for the evolution of electronic trading.

Business model support

Markets will evolve and mature at different paces and

standards must be able to span the varying requirements while providing a flexible framework for extension. In addition, terminology will vary across markets to the point where the same terms may have different labels and the same label may convey different meaning depending on the market context. A fundamental principle of the working group is to develop solutions that serve all business models of the FPL member community and to drive the standardization to common terminology as much as is feasible.

The ultimate goal of the working group from a business process, message, and data model standpoint is to ensure that the market data model covers the mainstream aspects of all market segments: equities, options, futures, foreign exchange, and fixed income

To gain an appropriate view of the requirements of various markets, the working group commissioned member representatives to present their respective markets' requirements for market data. The working group spent a considerable amount of time studying a diverse set of business models and requirements. The results of the study were intriguing and revealed that while markets will have some unique requirements, there is significant commonality across all classes at a basic level that allows a core set of processes and messages to meet the needs of the majority of business models.

Presenters took care to elucidate the high-level requirements of their specific market segment, describe how these requirements were manifested in their specific business, and finally address the implementation of those requirements with respect to technology, messaging, and data elements. Presentation materials from these briefings can be found in the working group web pages http://www.fixprotocol.org/working_groups/mdoptwg. The ultimate goal of the working group from a business process, message, and data model standpoint is to ensure that the market data model covers the mainstream aspects of all market segments: equities, options, futures, foreign exchange, and fixed income.

In addition to basic business model analysis, the working group strives to develop a common understanding or view of requirements stemming from the regulatory structure that shapes some of the different business models. It is important to grasp this dynamic, as well as the role played by each organization, in order to determine the impact on the new standards. The working group reviewed current industry regulatory activities that have the potential to significantly change the landscape including a brief overview of Reg-NMS or specifically the "trade-through rule" and a summary of the requirements of The Markets in Financial Instruments Directive (MiFID) that could be satisfied by an FPL market data standard. These requirements and the potential for changed market behavior are being considered and incorporated into the working group activities.

The FIX Market Data message suite was first introduced in FIX 4.2 and revisited again in FIX 4.4. The Market Data Model has recently been re-evaluated as part of the working group activities with the goal of ensuring coverage and inclusiveness with respect to the full spectrum of business models and requirements. For the most part, the framework proved to be in excellent shape.

Order book management

Central to any market venue (exchange, ECN, ATS, etc.) is its book of orders or interest in tradable securities or contracts. As part of the market data service, a market venue will provide views of its order book to participants and interested market data consumers. These views may not include the full scope or content of the book and the level or granularity of information may vary between different levels of service.

Order book management covers a broad range of topics from start of day procedures, to basic book management techniques, to contingency behavior when attempting to recover from missed messages. The working group formed a subgroup of industry experts tasked with developing the recommendations for book management considering the needs of the various business models. The subgroup first outlined the problem in general, studied the proper behavior of an order book and then addressed the

Exhibit 1

An Order is added to the book, Buy 3 @ 9750

Book Before Order is Applied

Top Of Book -BBO		Aggregate-Price Depth		Detail -Order Depth	
BID	**ASK**	**BID**	**ASK**	**BID**	
2 / 9740	9760 / 5	2 / 9740	9760 / 5	9740	ID=213, Qty=2
		5 / 9730	9770 / 3	9730	ID=223, Qty=4; ID=227, Qty=1
		1 / 9720	9780 / 5	9720	ID=973, Qty=1
			9790 / 7	**ASK**	
				9760	ID=230, Qty=5
				9770	ID=231, Qty=3
				9780	ID=232, Qty=5
				9790	ID=233, Qty=4; ID=234, Qty=3

a) Update the Top Of Book

Tag number	Tag name	Value	Description
TBD	MDBookType	1	Top Of Book
279	MDUpdateAction	1	Change
269	MDEntryType	0	Bid
271	MDEntrySize	3	Quantity
270	MDEntryPx	9750	Price, it's the key to insert the entry
107	SecurityDesc	GEZ5	Instrument, it's the key to the book

b) Update the Aggregate (Price Depth)

Tag number	Tag name	Value	Description
TBD	MDBookType	2	Aggregate
279	MDUpdateAction	0	New
269	MDEntryType	0	Bid
107	SecurityDesc	GEZ5	Instrument, it's the key to the book
271	MDEntrySize	3	Quantity
270	MDEntryPx	9750	Price, it's the key to insert the entry
TBD	MDPriceLevel	1	PriceLevel of Bid/Ask

c) Update the Orderbook Detail (Order Depth)

Tag number	Tag name	Value	Description
TBD	MDBookType	3	Order Detail
279	MDUpdateAction	0	New
269	MDEntryType	0	Bid
107	SecurityDesc	GEZ5	Instrument, it's the key to the book
271	MDEntrySize	3	Quantity
270	MDEntryPx	9750	Price, it's the key to insert the entry
TBD	MDPriceLevel	1	PriceLevel
290	MDEntryPositionNo	1	Position at this price level
278	MDEntryID	274	Order Reference

Book After Order is Applied

Top Of Book -BBO		Aggregate-Price Depth		Detail -Order Depth	
BID	**ASK**	**BID**	**ASK**	**BID**	
3 / 9750	9760 / 5	3 / 9750	9760 / 5	9750	ID=274, Qty=3
		2 / 9740	9770 / 3	9740	ID=213, Qty=2
		5 / 9730	9780 / 5	9730	ID=223, Qty=4; ID=227, Qty=1
		1 / 9720	9790 / 7	9720	ID=973, Qty=1
				ASK	
				9760	ID=230, Qty=5
				9770	ID=231, Qty=3
				9780	ID=232, Qty=5
				9790	ID=233, Qty=4; ID=234, Qty=3

requirements in the context of efficiency and reliability, all with the intent of providing a robust and well-defined market data standard. The findings of the subgroup have been published as a Recommended Practices document which is available on the FIX website for FPL members.

A full spectrum of market data book service types were evaluated inclusive of Top of Book, Order Depth, and Price Depth. These book types reflect two basic properties of book management; Aggregation and Depth. "Aggregation" refers to the convention of expressing a book in terms of order-level detail or conversely, summarizing orders at a given price level. "Depth" refers to the number of price levels maintained as either a detailed or summarized book.

A basic premise of efficient book management is that updates to the book be concise and incremental. Messages used to convey updates to a view of the book should include only the data necessary and should eliminate or reduce data redundancy. This can result in requirements for implicit book management behavior in consuming applications. Book management techniques include the ability to add, change and delete entries based on a predefined or pre-established key and thus there is a transactional nature to the message handling. Incomplete updates to the book will result in a loss of integrity and data quality.

Book management techniques include the ability to add, change and delete entries based on a predefined or pre-established key and thus there is a transactional nature to the message handling

Working from these premises, one can then proceed to elaborate the techniques by which these different books can be maintained. A Price Depth (Aggregate) book is best maintained by price, bid/ask and price-level where the book is sorted on price (descending for bids and ascending for asks). Bid/Ask and Price act as keys to referencing entries in the book with Price-level functioning in the role of an integrity check. Additionally, an action code of add, update or delete specifies the operation to be carried out on that price level. Top of Book, also qualifying as a Price Depth Book, would be managed as a series of changes to the top price level. An Order Depth book will be maintained at an additional level of granularity due to the presence of orders. In order to sequence orders within a price-level a position number within that price level should be specified.

In the following example (Exhibit 1) the Top-Of-Book, Price Depth Book and Order Depth Book behavior is illustrated showing the books before and after update messages are applied. In the Top-Of-Book and Price Depth Book the cells show total quantity and price. On the buy-side it is shown as qty/price and on the offer side it is price/qty. In the Order Depth Book you see the orders at each price level. For each order you see the reference ID and the quantity. The vendor may make additional information available but this is the minimum required.

The mechanics of the process do not necessarily assume the use of FIX messages and tags but work well in the context of the FIX vocabulary. The Market Data Incremental Refresh message provides a repeating group at the entry level. This entry can be used to specify the bids, asks, trades in addition to a number of price events such as high and low price. Within a repeating entry one would also specify instrument, action, price, size, etc. In concept, all the market data artifacts associated with an Order event can be encapsulated in a single message.

Data representation

Data Representation is concerned with the format of data as it is placed on the wire and seeks to reduce the size of market data messages without removing any of the content. The working group is taking advantage of a unique opportunity to develop new techniques that will conserve bandwidth and reduce the size of data as it traverses the network. This work is extremely relevant given today's struggle with market data volumes and has been commissioned by a coalition of industry participants. FPL has engaged a team of expert engineers who will execute a proof of concept program to empirically demonstrate the effectiveness of a number of different techniques using live industry data as its subject. The ultimate goal of this exercise is to demonstrate that market data can be delivered in a radically reduced form, obviating the need for greater bandwidth capacity, while keeping the complexity of data representation itself to a minimum.

Data Representation can be viewed as a series of layers where each layer provides further optimization without reducing content. The first layer, referred to as Implicit Tagging, is a form of FIX tag=value in which the tags are implied based on the existence of a template. Implicit Tagging has two beneficial results with respect to reducing message size; tags are removed and replaced with delimiters, and fields become variable length. Variable length fields are not a new concept in FIX. However, many market data formats in use today use fixed length fields and would reap significant savings by adopting variable length fields. Below is a FIX book update message (Exhibit 3) represented as tag=value immediately followed by the same message formatted with implicit tagging (Exhibit 4). The template for building and parsing the implicitly tagged message is provided in Exhibit 5.

The second layer, referred to as Field Encoding, is applied while the data is still represented as ASCII and involves the removal of redundant data. For example, a message that has a repeating group of book updates is likely to carry redundant price information. By representing the price in the first group in its entirety and prices in subsequent groups as the arithmetic offsets, the message size can be reduced. The particular technique is referred to as "Delta Value Coding". Another technique used at this level is called "Copy Value Coding" in which the current tag is not provided if it is equal to the previous instance of that tag. Exhibit 6 shows the message as it appears when Field Encoding has been applied. Exhibit 7 shows the template which has been applied to produce the message. Tag 270 (EntryPrice) is denoted as "270-" indicating Delta Value Coding. The working group has produced a proposal that recommends more than 10 such optimization techniques.

Exhibit 3 - FIX tag=value message

```
262=123456|55=SP|48=CME000150112|22=4|460=7|167=FUT|200=200312|201=1|202=1055|
206=L|207=2|107=SPZ3C1850|268=14|290=1|269=0|270=1050|271=2|273=104444651|
272=20030930|290=2|269=0|270=1050|271=1|273=104458568|272=20030930|290=3|269=0|
270=1030|271=4|273=104 434395|272=20030930|290=4|269=0|270=1020|271=3|
273=104417468|272=20030930|290=5|269=0|270=1020|271=6|273=104419357|
272=20030930|290=6|269=0|270=1010|271=3|273=104416254|272=20030930|290=7|269=0|
270=980|271=5|273=104418643|272=20030930|290=8|269=0|270=93   0|271=7|
273=104415623|272=20030930|290=1|269=1|270=1460|271=3|273=104456539|
272=20030930|290=2|269=1|270=1520|271=6|273=104447792|272=20030930|290=3|269=1|
270=1520|271=2|273=104412561|272=20030930|290=4|269=1|270=1580|271=8|
273=104414953|272=20030930|2 90=5|269=1|270=1610|271=6|273=104411167|
272=20030930|290=6|269=1|270=1620|271=7|273=104411524|272=20030930|
```

Exhibit 4 - FIX implicitly tagged message

```
123456|SP|CME000150112|4|7|FUT|200312|1|1055|L|2|SPZ3C1850
<1|0|1050|2|104444651|20030930}2|0|1050|1|104458568|20030930}
3|0|1030|4|104434395|20030930}4|0|1020|3|104417468|20030930}
5|0|1020|6|104419357|20030930}6|0|1010|3|104416254|20030930}
7|0|980|5|104 418643|20030930}8|0|930|7|104415623|20030930}
1|1|1460|3|104456539|20030930}2|1|1520|6|104447792|20030930}
3|1|1520|2|104412561|20030930}4|1|1580|8|104414953|20030930}
5|1|1610|6|104411167|20030930}6|1|1620|7|104411524|20030930>
```

Exhibit 5 - implicit tagging template

```
262|55|48|22|460|167|200|201|202|206|207|107<290|269|270|271|273|272>
```

Exhibit 6 - FIX field encoded message

```
123456|SP|CME000150112|4|7|FUT|200312|1|1055|L|2|SPZ3C1850
<|0|1050|2|104444651|20030930}|||1|13917|}
||-20|4|-24173|}|||-10|3|-16927|}
|||6|1889|}||| -10|3|-3103|}
||-30|5|2389|}||| -50|7|-3020|}
|1|560|3|40916|}||60|6| -8747|}
|||2|-35231|}||60|8|2392|}
||30|6|-3786|}||10|7|357|>
```

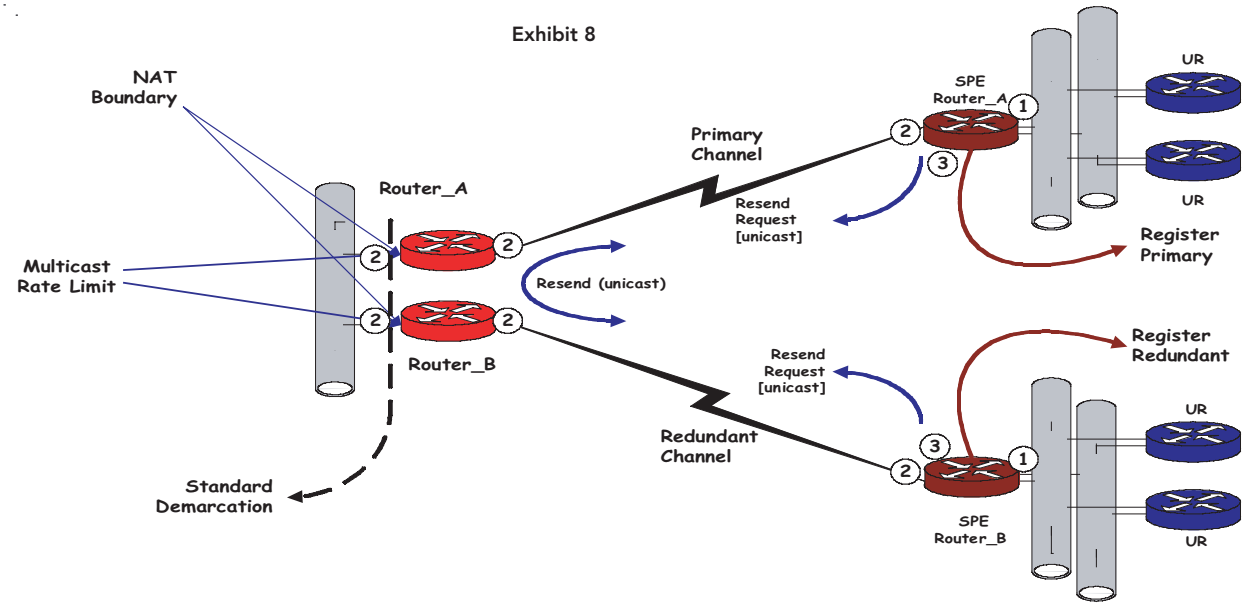
Exhibit 7 - field encoded template

```
262|55|48|22|460|167|200|201|202|206|207|107<290+1|269=|270 -|271!|273 -|272=>
```

The third and final layer that is applied once Implicit Tagging and Field Encoding are complete is Binary Encoding. Byte-level Binary Encoding is simply the representation of numeric values in binary form. Because a single byte can represent a value up to 255 and two bytes can represent a value up to 65,535 binary is an extremely efficient form of representing data. The working group has capitalized on this fact by developing algorithms which cleverly combine simple integers with delimiters in order to conserve a byte and which "mask" information using a control byte at the beginning of each message. When all three data representation layers are applied to a set of data, the outcome is often a 70% - 80% reduction in message size on the wire.

The working group also plans to investigate bit-level approaches which are valid in their own right and provide an added degree of optimization. In a bit level approach data is packed very tightly. A byte may actually contain data from two or more fields since formal byte boundaries are not recognized. An important item to note is that both bit-level and byte-level binary encoding incur significantly less latency due to CPU utilization than do off-the-shelf compressor such as Zlib or Gzip.

Exhibit 8



The encoding algorithms are being developed and refined within the context of the pilot project mentioned above. The effectiveness of these methods at reducing market data will be empirically evaluated with results and metrics documented. At the conclusion of the study the finalized encoding/decoding code will be made available to FIX membership for integration into market data solutions across the industry.

Multicast dissemination of market data

Multicast transport is generally recognized as the most efficient means of distributing large quantities of data to a wide audience where all recipients must have the ability to receive the data at the same time. It is used broadly in the financial industry by organizations that provide highly critical market data services. Multicast is a broadcast distribution that supports a configuration in which many data consumers subscribe to the same data content. As part of the effort to optimize all aspects of Market Data, FIX is extending its specification to address the area of Multicast Transport and Dissemination of market data. With the help of several industry veterans that specialize in multicast data distribution the working group has defined a set of recommended practices for disseminating market data over a multicast protocol.

Multicast is based on UDP (User Datagram Protocol) and delivers packets of data in a broadcast paradigm. Multicast packets can be any size supported by the transmission medium but are generally configured at a size of 1KB. Multicast takes place over a Wide Area Network in which Consumers of the data are cognizant of the data source and subscribe to the Senders IP Address through static router configurations. By nature, UDP is not considered a reliable transport protocol because on the rare occasions that packets arrive lost or corrupted, UDP does not provide a retransmission mechanism. Retransmission and error-handling tasks are left to the application. Live/Live Configurations mitigate the lack of data reliability by providing a dual feed in which the same data stream is delivered over redundant connections. This configuration also protects from infrastructure failures in which a circuit, router or other hardware may fail.

From a logical standpoint, a channel is essentially a wide area network pipe used to deliver a related group of data. Primary groups should be defined by product or instrument. It is up to the Sender to determine the level of product granularity. This usually depends on the volume of data being generated across a set of products. Once the make up of the primary group has been determined, a channel can be created to the group to transmit the data. Exhibit 8 illustrates a standard multicast configuration as recommended by the FIX Optimization Group.

In general, a lighter FIX session layer is recommended for use in a multicast paradigm. However, there are still several session level controls that are useful in managing a Multicast session. For more details see the Recommended Practices for Multicast Transport which can be found on the market data optimization working group webpage.

Conclusion

To compete in the new era of electronic trading, exchanges, electronic trading venues and trading firms must consider the efficiencies of market data and its dissemination as a corner stone of their strategic technology architecture. While some industry pundits have predicted capacity chaos and mayhem if the increasing volumes and peak message rates are not mitigated in some way, it is clear that there are needs for more efficient and consistent standards in market data dissemination on a broader scale.

The FIX Market Data Optimization Working Group is committed to developing useful and relevant standards for market data dissemination. While there has been significant work already accomplished by the working group, there are still many challenges ahead and there remain several opportunities for firms and their representatives to get involved in defining these important industry standards. Individuals and firms interested in participating in these standards setting activities are encouraged to join FIX and engage in the activities of the working group.

Industry collaboration is the key to reaching these goals. FIX is on the verge of demonstrating once again that an industry-driven standard is the most effective way of solving industry problems. **FIX**

Any thoughts on this or other articles?
Please send any comments, refering to this article as Vol 1 Issue 6 GL1, direct to Edward at edward@fixglobal.com