FIX Market Data Optimization Work Group
January 2006

Table of Contents

Authors:

David Cohen, BT Radianz
Michael Cooper, BT Radianz
Avner Gelb, SIAC
David Sawyer, BT Radianz
Matt Simpson, CME
Jeff Welch, SIAC

Recommended Practices for FIX Market Data over a Multicast Transport

Multicast transport is generally recognized as a very efficient means of distributing large quantities of identical data to a wide audience. It is widely used in the financial industry by organizations that provide highly critical market data services.

IP multicast networking provides mechanisms that support the forwarding of IP data packets to a set of receivers at a number of locations. By way of contrast, IP unicast networking supports the forwarding of IP data packets to a single receiver at a particular location and IP broadcast supports the forwarding of IP data packets to all receivers at a particular location.

The principle advantage of IP multicast networking is that a single IP packet may be generated by an application Server and then forwarded to a group of interested receivers. This permits the optimization of both Application and Network resources.

Because the Application Server is only required to generate a single packet for all receivers, then Server capacity and performance is no longer affected by an increase in the number of receivers (which would otherwise impact CPU, Memory, Network connectivity etc.) and the Application Server does not need to be concerned with the maintenance of session overheads. Similarly, the amount of bandwidth consumed across the network to support these application flows is reduced because only a single packet needs to be forwarded over any given path.

Architecturally, the de facto standard network protocol model is a Sparse Mode model whereby data is multicast across the network to network devices to which interested receivers are attached. Dense Mode models that correspond to a ‘push model’ have been depreciated.

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 following recommendations are being proposed.

Multicast Configuration

IP multicast utilizes UDP (User Datagram Protocol) at the transport layer and the application data is encapsulated in an IP|UDP frame.

Burst control is critical as a server could easily send a burst which will overwhelm network bandwidth (sub second level not just n seconds). All multicast sources should have some sort of pacing mechanism to allow their bandwidth use to be regulated.

It is normal industry practice to logically segment the data being published by instrument or some similar metric. This data is then transmitted on unique prescribed multicast group addresses and with a unique prescribed UDP Port numbers. This has the effect of creating discrete data ‘channels’ or ‘lines’.

The creation of data channels and lines introduces a level of service granularity that can be used to create product sets and permits the sharing of processing overhead.

However, the introduction of too much granularity can introduce network and other processing overheads (the result could be the creation of a very large number of multicast groups for instance). So, while the segmentation of data is in general of benefit, care should to keep the number of groups being transmitted to a reasonable minimum.

One of the key issues in multicast scaling is group churn. Any model that allows customers to subscribe/unsubscribe to groups at too fine a granularity does not work (for example if every stock code was a group). There is a trade off between group size (how many instruments) and bandwidth waste (delivering the data when no-one is interested). Larger groups waste more bandwidth, smaller groups do not scale.

Consequently, it is normal practice to logically group content in a manner that introduces a reasonable granularity and to group the data in a manner that permits the data to be shared relatively evenly across the delivery channels. Typically metrics include information content (e.g. Equities, Options), and Data volumes (the set of instruments "A - D" for example).

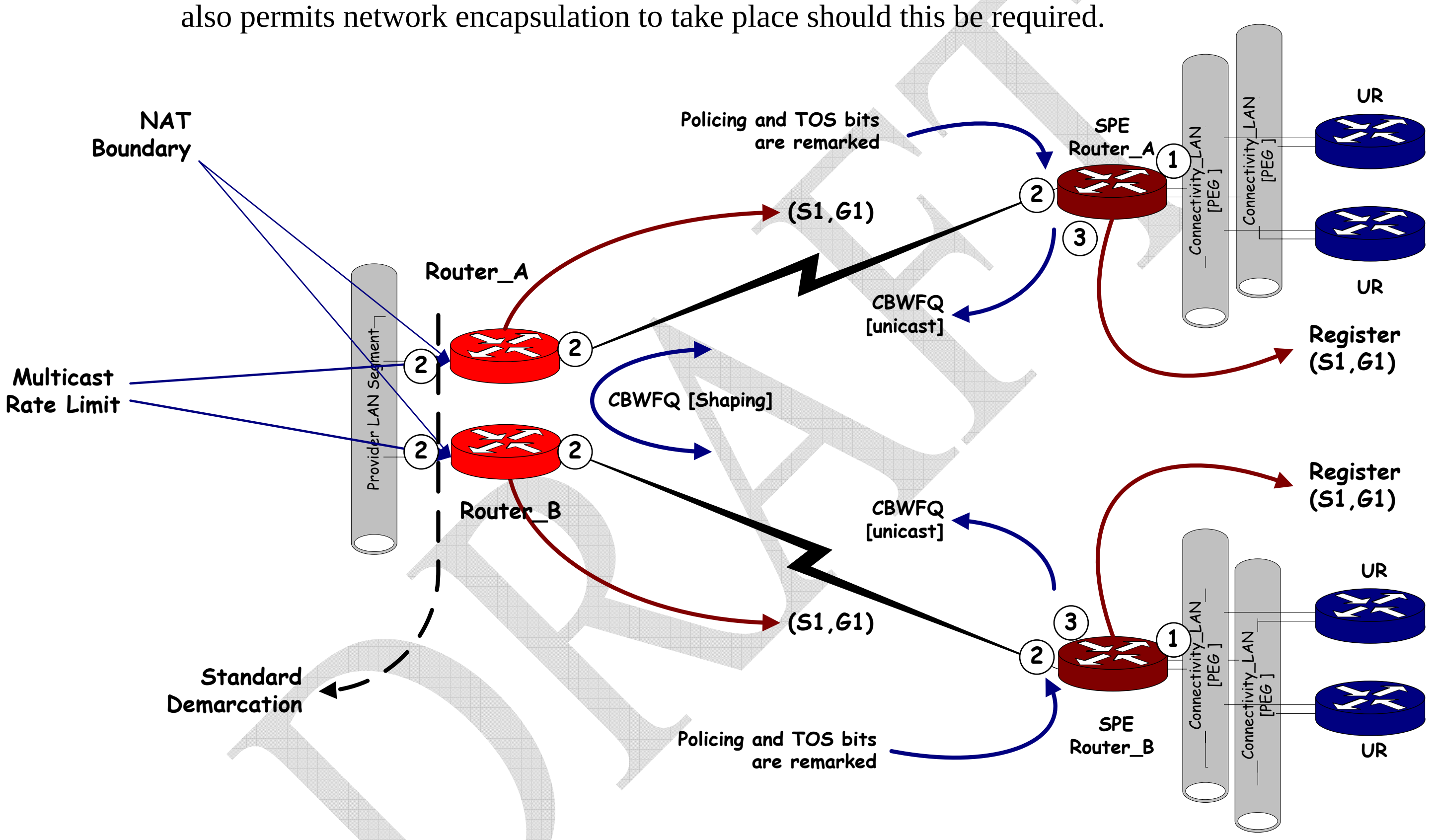
In order to facilitate application service provider, network service provider and subscriber integration it is best practice to use IP Source and Destination addresses that are globally unique.

For the unicast IP source address this would be an IP address that is registered by the IANA organization or the appropriate Registry to either the application or network service provider. For the multicast destination multicast group addresses this means that application service providers should only use addresses that have been assigned by IANA or that conform to industry standards e.g. RFC3180.

It is considered best current practice to optimize network and server resources, and to minimize overheads by packing multiple messages into a single data packet. This typically requires the application to transmit a packet on the expiry of some (minimal) clock cycle with packets being transmitted before the clock cycle end where the packet is fully packed.

It is also considered best practice that transmitted packets should be uniquely identifiable with an absolute reference and that packets should not be referenced by their relationship to each other. This enables retransmission requests to reference a specific packet being transmitted to a specific group and does not place any requirement on the receiving host that it must process a specific packet before processing any others.

Consideration should be given to the size of packets that are generated by the application. As mentioned above, it is recommended that multiple messages should be carried in each packet in order to optimize processing and forwarding. However, in general, it is recommended that packet size should not exceed 1400 Bytes to stay under MTU. This also permits network encapsulation to take place should this be required.



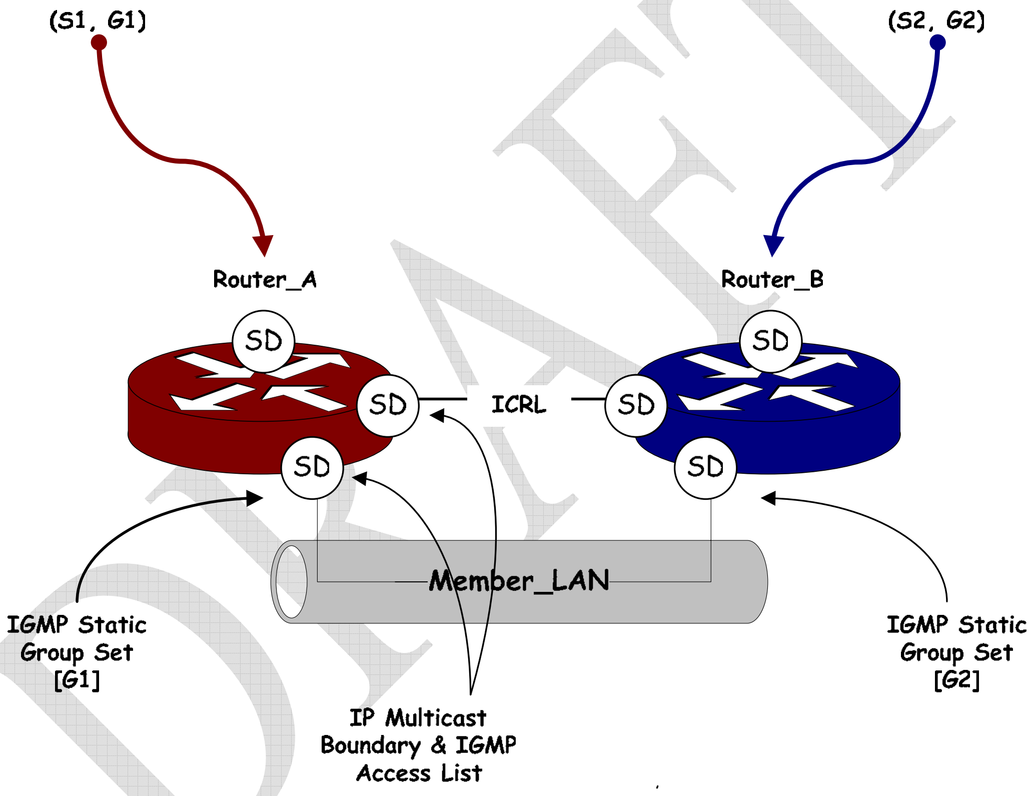
Live/Live Configuration

IP packets may (under certain circumstances) be delivered out of order, duplicated or may be lost. Like IP, UDP is a connectionless protocol and does not provide the same mechanisms that TCP does with respect of lost or out of order packets. As a consequence; it is the responsibility of the application to manage out of order, duplicated or lost packets.

Applications that join a multicast group and receive multicast packets should be able to tolerate duplication and out of sequence packets.

Application service providers utilize a range of mechanisms to provide a retransmission service for multicast applications. In addition to these many application providers use a common approach | mechanism to mitigate the risk of a subscriber not receiving a data packet.

This approach | mechanism is to duplicate the data and then to deliver each copy over separate channels (multicast group | UDP port). These channels can then be delivered over separate physical network infrastructure further extending the diversity and further mitigating the risk of lost packets. The expectation is that the subscriber | receiver of this data should listen to both channels and is responsible for arbitrating between them to create a single set.



Channel Grouping and Definition

Care should be taken to keep channels to a reasonable minimum. As a practical example, , this could approximately 10 primary channels, for a total of 20 channels in a Live/Live configuration. One reason for restricting the proliferation of channels is the increasing scarcity of Global Unique Multicast ID's, which have a limited IP Address range. Many factors need to be considered to specify the number of channels. For instance; volume of data is a determining factor - we do not want 10 channels at 1 KBit/s (or even 64 KBit/s

Communicating Channel ID's

Best current practice is to separate the definition of what groups, ports etc. constitute a channel from the actual forwarding mechanism.

Therefore, incorporating IP address information (including multicast group addresses) in the data portion of the message is discouraged. Separation of the address from the data portion of the message also provides more separation and permits more flexibility in the use of IP and UDP addresses (this is particularly important when changes are required).

In support of this it is recommended that a notification or advisory approach is adopted for the communication of product definitions and attributes (which symbols are available on what multicast group address for example).

Multicast Session Layer

This section discusses the session layer control that should be applied when disseminating market data over a Multicast Transport. 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.

Entitlement – If channel customers connect directly to the provider's distribution network, entitlement may be controlled at the provider network edge by PIM Filtering or IGMP control mechanisms. This entitlement acts as a substitute for a Login. This entitlement method cannot be used on a network infrastructure that is not directly controlled by the provider (or a proxy), such as the Internet. A separate session level Login would only be necessary in order to affect template exchange and communicate from sender to receiver which message types that will be used. In short, an explicit Login by the Receiver to register on a channel or set of channels is not required.

Sequence Gap Handling – Conventional FIX Sequence Gap Handling behaves differently in a multicast scenario for several reasons. First, the Live/Live configuration requires that both feeds be interrogated for a potentially missing sequence number. Sequence Gap Handling can still be performed, but the algorithm must determine whether the message sequence number was present on either feed. Second, it is not recommended that resend requests be made over a multicast session due to reliability issues. In the case where a sequence number is not available on either feed, an *out-of-band request* for retransmission of the missing data can be made using a Resend Request (see Application-level Requests for Retransmissions below). It is recommended that a separate session connection be maintained via TCP-IP for this purpose as it provides reliable delivery of the resend request.

Requests for Retransmission via Resend Request must always specify the primary channel over which the requested messages were originally transmitted. The beginning and ending sequence numbers alone are not sufficient since they may not be guaranteed to be unique across channels. Note that since the channel ID's are subject to change they should be made a configurable parameter in the Resend Request.

Heartbeat **–** Used by the receiver to verify that connection is still active. Heartbeats must be provided per channel and will carry a Message Sequence Number (MsgSeqNum/Tag 34) per the FIX specification. Each channel will have its own series of sequence numbers that must be unique within channel but not across channels. The MsgSeqNum can be used by the receiver to validate the integrity of the feed. If the MsgSeqNum on the Heartbeat is equal to the Receivers next expected sequence number then the feed is intact. Otherwise, the feed has been disrupted and a retransmission or reconnection is necessary.

Time Beacon – serves roughly the same purpose as the Heartbeat, which is to provide notification that a channel is active. However, a Time Beacon does not carry a Next Expected Sequence Number.

Retransmission of Data

Because UDP is an unreliable, connectionless datagram delivery mechanism; it is the responsibility of the application layer to cater for packets that may be duplicated, may arrive out of order and may be lost. Note that in terms of packet forwarding, either TCP or UDP may be susceptible to this.

Consequently, and even in the presence of delivery mechanisms such as multicast LIVE | LIVE, most applications that are distributing critical data also support a retransmission capability.

There are a number of different retransmission mechanisms being used today with many different characteristics.

Practically all of these mechanism use an 'out of band' request mechanism. An 'out-of-band' request mechanism is considered best and the recommended mechanism is to use a TCP/IP unicast request model – 'in band' multicast based requests are discouraged.

In this model the receiver Host device would establish a TCP session with the appropriate application service provider device (which may or may not be the same device publishing the multicast data streams) and request that a packet or range of packets be retransmitted.

The requested packets may be delivered in a number of ways: (i) packet(s) is retransmitted using the primary | production multicast Groups, (ii) packet(s) are returned on the TCP/IP unicast session, (iii) packet(s) are transmitted on a discrete group identified as being the retransmission group.

It is recommended that 3rd option – where a separate channel or channels are used- be adopted as this provides separation from the primary production forwarding mechanism while still taking advantage of the efficiencies of multicast.

Frequently, where this model is used, there are also multiple multicast retransmission groups. Where there are multiple groups, the best current practice associates each retransmission group with the data content.

Retransmission Channels

It is recommended that the retransmission mechanism utilize IP multicast to disseminate packets being retransmitted. Further, it is recommended that retransmitted packets should not be published on the same channels on which the original packet(s) were transmitted.

As a consequence it is considered best current practice to establish a set of retransmission groups that participants may join either if they anticipate needing to receive retransmitted packets or as they need to receive retransmitted packets. However, it is expected that generally participants who anticipate utilizing a retransmission mechanism will always join the dedicated retransmission groups.

Retransmission should always take place over a separate channel dedicated to this purpose. Generally speaking, there should be a single channel over which all the data is retransmitted. As is the case for any channel connection, an edge router will be statically joined to the channel meaning that the channel will always be in active mode and subscribing to the feed. This configuration has the advantage of simplicity in that all retransmitted data will always be received over a single channel.

The one comparatively small disadvantage it presents is that it does not make optimal use of bandwidth. The retransmission channel will carry data that has been requested by all market participants, not just the Receiver listening for the data. A possible alternative is to consider manually connecting to the channel when a retransmission is expected and disconnecting when one is not expected.

A possible alternative is to have multiple retransmission channels, perhaps as many as one per primary channel. The advantage of this approach is that the Receiver knows which data product group is being received over a given Retransmission channel. The disadvantage is that it violates the directive to keep channels to a reasonable minimum and increases the overall complexity of a Multicast infrastructure. For this reason it is not recommended.

Application-Level Requests for Retransmission

Application-level retransmission requests will be supported via the Market Data Request message. It is advised that Retransmission Requests be made over a TCP-IP connection which provides reliable delivery of the request. It is incumbent on the receiver of the request to respond with a response acknowledging receipt of the message. To the extent that validation is performed on the request, a Market Data Request Reject can be issued by the responder with an “unable to process” response and a qualifying reject reason.

Retransmission Requests are generally made by:

- Time range where the beginning and ending timestamps are specified
- Channel ID
- Instrument
- Current State of the Book

Note that the Market Data Request should not be used for requesting retransmissions by message sequence number as this is performed automatically at the session level.

Multicast Entitlement

The architectural model most frequently employed to support publication of data using IP multicast is a unidirectional model in which the application service provider publishes the data and subscribers signal their interest at the network layer.

As a consequence there is currently generally no end-to-end, bi-directional application layer communications that would support a logon sequence that could be used to validate that the subscriber is entitled to receive the data being published. Multicast is not normally the best technology for “any by any” interest reception.

Therefore it is incumbent on a trusted network provider (which may be the application service provider or a separate network service provider) to provide an entitlements model at the network layer.

In the absence of a trusted network, data encryption may be used, however data encryption is usually not used for multicast market data. Encryption is generally not considered necessary unless data is being transmitted over an untrusted network such as the internet. The security provided by a private line or a private extranet makes the additional latency introduced by encryption unattractive.

The primary goals of the entitlements model should be to provide (i) secure transmission of the IP multicast published data that also guarantees that the sender is authoritative; (ii) dissemination of the data to approved recipients only.

Multicast Entitlement provides a layer of security for the sender of data. Only approved receivers of the data will be capable of joining a channel. It doesn’t provide security in the sense of protecting the sender’s internal resources, but it does help ensure that only approved recipients get the channel data. Static configuration determines who is able to join a channel. A receiver must be set up in the router of the sender in order to join a multicast channel.

PIM Filters are used to provide entitlement at the router. No LDAP lookups or other validation is necessary.

The nature of IP multicast precludes the use of those mechanisms used in unicast communications to authenticate a requestor and there is no ubiquitous industry wide alternative. As a consequence, the most frequently adopted approach is to utilize network layer controls implemented on boundary edge devices to ensure that the dissemination of data is to approved recipients only.

Message Header

This section discusses the FIX Header and its application in a Multicast environment. The FIX Header, which is carried as part of the application message, is not to be confused with the Multicast Header, which specifies among other things, the IP Address of the Sender. Because Multicast is a broadcast rather than a unicast transport, messages can be built using just a subset of the FIX Header fields.

Minimally, the BeginString, MsgType, MsgSequNum and BodyLength should be used in a multicast environment. The FIX specification currently requires that the SenderCompID and TargetCompID fields be provided as well. However, these fields lose their utility in a multicast session and will be required only in point-to-point sessions.

Retransmitted messages can optionally carry the PosDupFlag, which is used in standard FIX sessions, when the message is being sent in response to a Resend Request.