US Equities Multicast Latency Feed Specification

Version 1.1.0

December 22, 2011

Contents

1 Introduction

1.1 Overview

BATS members may use the Multicast Latency Feed to monitor various Matching Engine unit performance statistics. Updated Matching Engine level statistics will be distributed via the Latency Feed every 15 seconds during market hours.

Measurements and metrics for each measurement to be distributed are as follows.

1.1.1 Measurements

- Order to Quote - The elapsed time from when an order enters the BATS network to the time that order appears on the BATS Multicast PITCH Gig shaped A feed. Only visible orders will have Order to Quote metrics.

1.1.2 Metrics

- Begin Time - Number of milliseconds from midnight Eastern Time.
- End Time - Number of milliseconds from midnight Eastern Time.
- Count - Count of orders used in calculation of all other metrics from *Begin Time* to *End Time*.
- Minimum - Minimum order processing latency from *Begin Time* to *End Time* in seconds.
- Maximum - Maximum order processing latency from *Begin Time* to *End Time* in seconds.
- Average - Average order processing latency from *Begin Time* to *End Time* in seconds.
- Mode – Mode of the order processing latency from *Begin Time* to *End Time* in seconds.
- Percentile – Percentile based metrics of the order processing latency from *Begin Time* to *End Time* in seconds. Initial values to include 99.9%, 99%, 95%, 90%, 75%, 50% (Median), and 25%.
- Standard Deviation - Standard deviation of sample.

1.2 Feed Connectivity Requirements

The Multicast Latency Feed is available only from the BATS primary data centers of the BYX Exchange and the BZX Exchange. Given the limited amount of messaging required, there is no minimum bandwidth requirement for the Latency Feed. All statistics, at the time of this writing are available from a single multicast group. There will be no subdivision by symbol range. Dropped messages are not recoverable for the Latency Feed.

2 Protocol

The Latency Feed is to be used for monitoring Matching Engine unit performance. The Latency Feed cannot be used to enter orders. For order entry, refer to the BATS FIX Specification.

All information presented on the Latency Feed is anonymous and does not contain any member identifying information.

2.1 Message Format

The messages that make up the Latency Feed protocol are delivered using the BATS Sequenced Unit Header which handles sequencing and delivery integrity. This unit header mirrors the unit header used in Multicast PITCH. Members familiar with the Multicast PITCH protocol should find it very easy to reuse that code to process the latency feed. The `Hdr Unit` field from the sequence unit header will always be zero filled in the latency feed. Within the context of the latency feed a single message can carry multiple statistics from different units. The applicable unit will be identified within each statistic carried in the message.

All UDP delivered events will be self-contained. Developers can assume that UDP delivered data will not cross frame boundaries and a single Ethernet frame will contain only one Sequenced Unit Header with associated data.

The Latency Feed is comprised of a series of dynamic length sequenced messages. Each message begins with a Length field. BATS reserves the right to grow the length of messages without notice. Members should develop their decoders to deal with unknown messages and messages that grow beyond the expected length. Messages will only be grown to add additional data to the end of a message.

Message sequence numbers are incremented by one for every sequenced message.

2.2 Data Type

The following field type is used within the `Sequenced Unit Header` and the `Latency Stat` messages.

- **Binary** fields are unsigned and sized to “Length” bytes and ordered using Little Endian convention (least significant byte first).
- **Double** fields are 64-bit floating point ordered using Little Endian convention (least significant byte first).

2.3 Message Framing

Messages will be combined into single UDP frame where possible to decrease message overhead and total bandwidth. The count of messages in a UDP frame will be communicated using the BATS `Sequenced Unit Header`. Framing will be determined by the server for each site.

2.4 BATS Sequenced Unit Header

The BATS `Sequence Unit Header` is used for all BATS Multicast Latency Feed messages.

Sequenced and un-sequenced data may be delivered using the `Sequenced Unit Header`. Un-sequenced headers will have a 0 value for the sequence field.

Sequenced messages have implied sequences with the first message having the sequence number contained in the header. Each subsequent message will have an implied sequence one greater than the previous message up to a maximum of count messages. Multiple messages can follow a `Sequenced Unit Header`, but a combination of sequenced and un-sequenced messages cannot be sent with one header.

The sequence number for the first message in the next frame can be calculated by adding the *Hdr Count* field to the *Hdr Sequence*. This technique will work for sequenced messages and heartbeats.

Sequenced Unit Header				
Field	**Offset**	**Length**	**Value/Type**	**Description**
Hdr Length	0	2	Binary	Length of entire block of messages. Includes this header and *Hdr Count* messages to follow.
Hdr Count	2	1	Binary	Number of messages to follow this header.
Hdr Unit	3	1	Binary	This byte will be zero filled.
Hdr Sequence	4	4	Binary	Sequence of first message to follow this header.
Total Length = 8 bytes				

2.5 Heartbeat Messages

The BATS `Sequenced Unit Header` with a count field set to “0” will be used for heartbeat messages. During trading hours heartbeat messages will be sent from all multicast addresses if no data has been delivered within 1 second. Heartbeat messages never increment the sequence number for a unit, but can be used to detect gaps on the real-time multicast channels during low update rate periods.

Heartbeats on the real-time multicast addresses during trading hours will have a *Hdr Sequence* value equal to the sequence of the next sequenced message to be sent for the unit.

Outside of trading hours BATS sends heartbeat messages on all real-time channels with a sequence of “0” to help users validate multicast connectivity. Heartbeat messages may not be sent from 12:00 am – 1:00 am ET or during maintenance windows.

3 Latency Feed Message

The Latency Feed supports a single message type that is used to present a variety of metrics for noted key measurements.

3.1 Latency Stat Message

A `Latency Stat` message demonstrates performance for a specific matching engine unit for a specific timeframe. The unit and metric are defined in the body of the message as described in the table below.

Latency Statistic				
Field Name	**Offset**	**Length**	**Type**	**Description**
Length	0	1	Binary	Length of this message including this field.
MessageType	1	1	Binary	0x90
Measurement type	2	1	Binary	0 = Order to Quote
Matching Unit	3	1	Binary	Refer to Multicast PITCH specification for Matching Unit mappings.
Begin Time	4	4	Binary	Number of milliseconds from midnight (ET).
End Time	8	4	Binary	Number of milliseconds from midnight (ET).
Count	12	4	Binary	Count of orders used in calculation of all other metrics from *Begin Time* to *End Time*.
Minimum	16	8	Double	Minimum order processing latency (seconds) from *Begin Time* to *End Time*.
Maximum	24	8	Double	Maximum order processing latency (seconds) from *Begin Time* to *End Time*.
Average	32	8	Double	Average order processing latency (seconds) from *Begin Time* to *End Time*.
Standard Deviation	40	8	Double	Standard deviation.

Mode	48	8	Double	Mode of order processing latency (seconds) from *Begin Time* to *End Time.**
99.9 Percentile	56	8	Double	*99.9 Percentile* of order processing latency (seconds) from *Begin Time* to *End Time*.*
99 Percentile	64	8	Double	*99 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
95 Percentile	72	8	Double	*95 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
90 Percentile	80	8	Double	*90 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
75 Percentile	88	8	Double	*75 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
50 Percentile (Median)	96	8	Double	*50 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
25 Percentile	104	8	Double	*25 Percentile* order processing latency (seconds) from *Begin Time* to *End Time*.*
Total Length = 112 bytes				

** Adding effective 01/20/12.*

4 Multicast Configuration

4.1 US Equities Production Environment Configuration

4.1.1 Limitations/Configurations

The following table defines BATS current configuration for network limitations.

Period/Type	Limit/Setting	Notes
MTU	1500	BATS will send UDP messages up to 1500 bytes. Members should ensure that their infrastructure is configured accordingly.

4.1.2 Rendezvous Points

Data Center	Rendezvous Point
BZX Exchange Primary Data Center	208.90.208.253
BYX Exchange Primary Data Center	174.136.167.1
BZX Certification Data Center	TBD

4.1.3 Latency Feed Distribution Addressing

Data Center	IP Port	Multicast & (Source) IP Address
BZX Exchange Primary Data Center	30001	224.0.62.184 (208.90.209.217)
BYX Exchange Primary Data Center	30001	224.0.62.185 (174.136.167.113)
BZX Certification Data Center	TBD	TBD

5 Support

Please e-mail questions or comments regarding this specification to tradedesk@batstrading.com.

Revision History

Document Version	Date	Description
1.0.0	05/16/11	Initial version 1.0.0.
1.0.1	06/17/11	Corrected typos in Latency Statistic table.
1.0.2	10/21/11	Corrected Offsets in Latency Statistic table.
1.1.0	12/22/11	Extending `Latency Stat` message to include *Mode* statistics as well as *99.9, 99, 95, 90, 75, 50, and 25 Percentile* statistics. *Effective 01/20/12.*