BMLL Data Feed

NBBO - OPRA

Date: 2025-11-18
Version: 1.3

1.1. Introduction

The OPRA-NBBO dataset gives a 1s, trade safe conflated view of the NBBO for US equity options.

Due to the very large size of the raw data, BMLL provides a conflated view of the Best Bid Offer updates allowing all market participants to get a sense of the prices in the NBBO markets. If multiple events occur during the same 1 second interval, BMLL will output the first event immediately and then provide a single quote update for all the subsequent events in the 1 second interval.

1.2. Dataset schema

Basic Types

Type	Content	Text representation
BOOLEAN	A True or False boolean value.	1,0
INT32	A 32 bit signed integer.	12345
INT64	A 64 bit signed integer.	12345
DOUBLE	IEEE 64-bit floating point value.	123.4567
BYTE_ARRAY	An arbitrarily long byte array.	“VODAFONE GROUP”
NUMBER	A number with defined precision (total digits allowed) and scale (number of digits allowed to the right of the decimal point). Unless otherwise specified, the default is precision 38 and scale 0.	12345
VARCHAR	Variable length string up to 500 characters. May contain quotes and special characters. All text will be utf8-encoded unless indicated differently. Ticker and ISO codes should be ASCII compatible.	“VODAFONE GROUP”
TIMESTAMP_NTZ	Nanoseconds since 1970 in UTC. Timestamp will be represented in ISO format in text format.	20220302-11:23.57.12345789
FLOAT	IEEE 64-bit floating point value.	123.4567
DATE	A calendar date consisting of a year, month, and day, without including any time or time zone components.	2025-01-01

1.3. Schema

One row for every quote flush in the conflation - trades are in the trade file.

Display Name	Type	Description
Date	INT32	The date containing the data for which this file was produced.
UnderlyingTicker	BYTE_ARRAY	Ticker of underlying equity or index.
Ticker	BYTE_ARRAY	Generated OCC / OSI option symbol name for options. The OSI symbol identifies the option contract that is traded.
ConflatSequence	INT32	The sequence number in the conflation for that specific `Ticker`. See Conflation algorithm for detail. Conflation sequences are used for trades and quotes and can be used to enumerate all the quotes on a contract.
EventTimestamp	INT64	UTC Time of the trade from the conflation algorithm based on timestamps of OPRA aggregation. Timestamps are expressed as the number of nanoseconds since 1970-01-01.
LocalTimestamp	INT64	Local time corresponding to the EventTimestamp.. Timestamps are expressed as the number of nanoseconds since 1970-01-01.
BMLLSequenceSource	INT32	The OPRA line number or a numerical partition identifier (2019 and earlier data).
BMLLSequenceNo	INT64	This sequence number will respect the order of messages in upstream protocol and will be populated even if the upstream venue does not provide a sequence number - where the source of the feed is indicated by BMLLSequenceSource.
ExchangeSequenceNo	INT64	The sequence number as provided by OPRA.

NBBOQuoteCondition	BYTE_ARRAY	Character indicating the quote condition associated with the BBO. Same encoding as OPRA quote condition for single venues but restricted to customer interest. ' ' : No customer interest. 'C' customer interest existing on both sides. 'B' customer interest on the bid side. 'O' customer interest on the ask side.
BestBidExchange	BYTE_ARRAY	Exchange that has provided that set the BestBidPrice, either by being the first one at that price, or by providing a size exceeding by at least 10 the previous quantity at the best bid price.
BestBidPrice	DOUBLE	Best bid price at the time of the event. Empty book represented via 0 is as provided by OPRA.
BestBidQuantity	INT32	Size (quantity) at the best bid price at the time of the event on the `BestBidExchange`.
BestBidTime	INT64	UTC time at which best ask price was set expressed as nanosecond since 1970-01-01.
BestBidExchangeQuantity	INT32	Size (quantity) at the best bid price at the time of the event on the `BestBidExchange`
BestAskExchange	BYTE_ARRAY	Exchange that has provided that set the BestAskPrice, either by being the first one at that price, or by providing a size exceeding by at least 10 the previous quantity at the best ask price.
BestAskPrice	DOUBLE	Best ask price at the time of the event. NaN if no ask orders.
BestAskQuantity	INT32	Size (quantity) at the best ask price at the time of the event across all venues.
BestAskTime	INT64	UTC time at which best ask price was set expressed as nanosecond since 1970-01-01.
BestAskExchangeQuantity	INT32	Size (quantity) at the best ask price at the time of the event on the `BestAskExchange`.
VenuesAtNBB	INT32	Bitfield indicating which venues are at NBB based on BMLL calculations (see notes).

VenuesAtNBO	INT32	Bitfield indicating which venues are at NBO based on BMLL calculations (see notes).
CumulativeVolume	INT64	Cumulative volume across exchanges
RegularTradingHours	BOOLEAN	Boolean indicating if the line allocation is for Regular Hours or Extended Hours sessions - Extended / Global trading hours can be reported on different files

1.4. Format and file structure

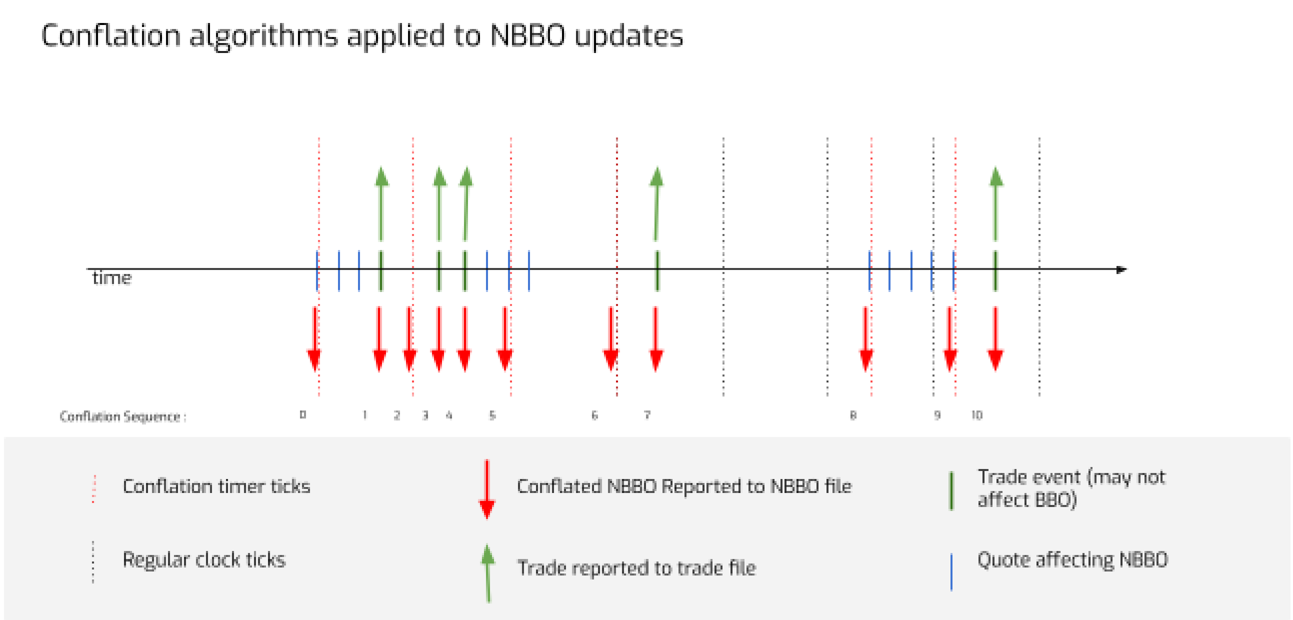
Data is delivered as a parquet file per date per sequence source.

OPRA-NBBO/{yyyy}/{mm}/{dd}/opra-nbbo-{yyyy}{mm}{dd}-{sequencesource}.parquet

1.5. Data arrival times

Full product coverage and arrival times are available at https://data.bmlltech.com.

1.6. Conflation algorithm



Updates to NBBO changes are conflated in the following manner:

- A timer at a specific frequency is started when there is a first quote that affects the NBBO - this timer may not align to a round time (e.g. 1s).
- The timer will remain active as long as there are quote updates in the window. As the timer starts the state of the NBBO will be output.

- All subsequent quote updates occurring during the duration of the conflation timer will be conflated together, and the state of the NBBO will be printed when the timer expires.
- Every time there is a trade, the NBBO will also be provided. This ensures that the NBBO is trade-safe
- Timestamps will be the timestamp of the NBBO flush - but the sequence number will be the sequence number of the last quote update event
- NBBO is flushed at the end of the conflation time only if NBBO is different from previously flushed NBBO record
- If the timer is not started, trades per se won't trigger the NBBO clock, but we expect if it is a lit trade that there will be an NBBO update at the same point in time and that the timer will be generally be started on this occasion.

A trade halt means that the venue stops contributing to the BBO and will be associated with empty books.

Note that there can be trades that don't affect books if they are executed off-exchange.

1.7. Representation of venues at BBO as a bitfield

OPRA venues have 32-bit ASCII characters - We would use the same bit as the OPRA encoding of the exchange A=bit 1. If a bit is set to 1 in the bitfield it means the last quote of the venue is at the NBBO price and the quote condition associated with that quote is eligible for BBO. It is technically possible for venues to display non-firm quotes better than NBBO but as those are not included in the NBBO this volume will make the venue disappear from the NBBO (even if it previously had a quote at the BBO, as there is no way for real-time feed to capture multiple prices) - The choice of the price identified as best-price for a specific venue is the choice of the participant. Also, we require the quantity available at the price to be greater than 0 for the venue to be considered to have a valid quantity.

1.8. 2019 and earlier data

BMLL had to use a different upstream data source for data prior to 2019. Some of the fields such as quote condition may not be available during that period.