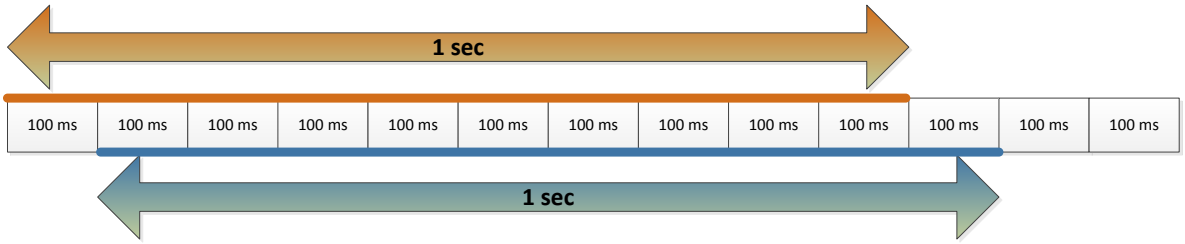
ORDER TRANSMISSION LIMIT (THROTTLING)

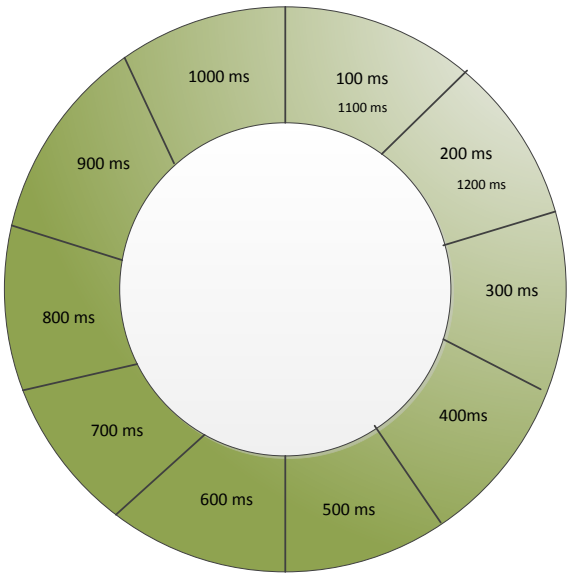
IMPORTANT NOTE:

This throttling mechanism valid from March 6th, 2017 for the BISTECH Equity and Derivatives Market Trading System.

BISTECH FIX System returns a Business Message Reject (35= j) in case members exceed the order limit per second (Throttling Quota) that have been predefined for their users while sending an order through FIX - BISTECH system. The capacity of the order limit per second is calculated and processed in the following one second period. There is a “sliding windows” structure consisted of ten 100 millisecond-units. Throttling limit is not checked according to the system clock but via 10 consecutive windows each lasts 100 milliseconds through the entire time. These windows follow each other in a circular manner on FIX gateways. In case the capacity of the order limit per second is exceeded, FIX session will not be terminated.

You can see the details of the structure below;





Because of Throttling Mechanism, the recommendations described below should be taken into consideration on client side during FIX connection.

1- Throttling limit can be managed in millisecond-level instead of second-level on client side.
2- SendingTime of sent FIX messages should not be used (generated via FIX Client tool). Instead timestamp on the Execution Report message should be used. A time interval should be planned by considering the time passed on network.
3- When the throttling limit is reached, incoming messages are not accepted by FIX servers and they are rejected with a Business Message Reject (BMR). It can be started again processing messages according to capacity utilization within the next 100 milliseconds period by looking backwards through a total of 10 windows.
4- Order transmission timing can be determined by controlling time lag between sent messages and reply execution messages dynamically (Tag52 : SendingTime in Execution Report – Order Ack (out)). For example, the time lag($t_{diff.}$) for Execution Report (ER) messages of first order and the last order within the capacity in a single second is calculated (t_n-t_1= $t_{difference}$). For the next message packet, client should wait 1 second after the execution message of last message is received (t_n +1000). Timing of second packet can be brought to an earlier time for a certain amount, as long it is not brought earlier more than the time difference ($t_{difference}$). The amount of time to bring earlier should be determined by user experience.

Instances resulting from Throttling mechanism can be seen below example.

Window sequence number	1	2	3	4	5	6	7	8	9	10	11	…	20
Total time (ms.)	100	200	300	400	500	600	700	800	900	1000	1100		2000
Sliding window size(ms)	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.	100 ms.
The amount of Order sent within the window	30 messages	56 messages	14 messages	0	0	0	0	0	0	0	100 messages (Pacing on), 70 messages rejected with BMR.		

✓ Orders coming from a user having throttling limit of 100 messages/second can be received in separate time windows as above.
✓ After sending 100th order, if the second group of 100 orders is sent in 1001th ms. without doing any control, only first 30 messages will be accepted and remaining 70 messages will be rejected with BMR. This is because at 11th time interval, 30 of the capacity are freed at 1st time interval.