

Load Test Report

Moscow Exchange Trading & Clearing Systems

24 September 2016

Contents

Testing objectives	2
Main results.....	2
The Equity & Bond Market trading and clearing system.....	2
The FX Market trading and clearing system	3
The Derivatives Market trading and clearing system	5
Subsystem for real time monitoring of the trading system parameters and market activity	6
ASTS Gateways	6
Communication channels and remote gateways.....	6
Bridges to the trading system.....	7
Latency for transactions in the Equity & Bond Market and FX Market trading systems	7
Latency for transactions in the Derivatives Market trading system	8
Transactional FIX gateways of ASTS	8
FIX/FAST UDP multicast marketdata of the Equity & Bond Market and FX Market	9
FAST UDP multicast marketdata servers of the Derivatives Market	9
Conclusions	10
The Equity & Bond Market, the FX Market.....	10
Derivatives Market.....	11

Testing objectives

1. To verify the trading and clearing systems operation under conditions of peak loading and an increased number of orders and trades. The trading systems of the following Moscow Exchange's markets were tested:
 - a. The Equity & Bond Market;
 - b. The FX Market;
 - c. The Derivatives Market.
2. Operational testing of the trading systems in the DataSpace data center, with a technical configuration as close as possible to the configuration to be used after DataSpace becomes the primary facility.
3. To estimate the time of order filling and data delivery from the trading and clearing systems at different load levels and software and hardware configurations.
4. To carry out a public testing of a new ASTS+ trading system of the Equity & Bond Market.
5. To allow developers of customer software and brokers to test their systems and estimate the throughput capacity of communication channels to the exchange venues.

Main results

The Equity & Bond Market trading and clearing system

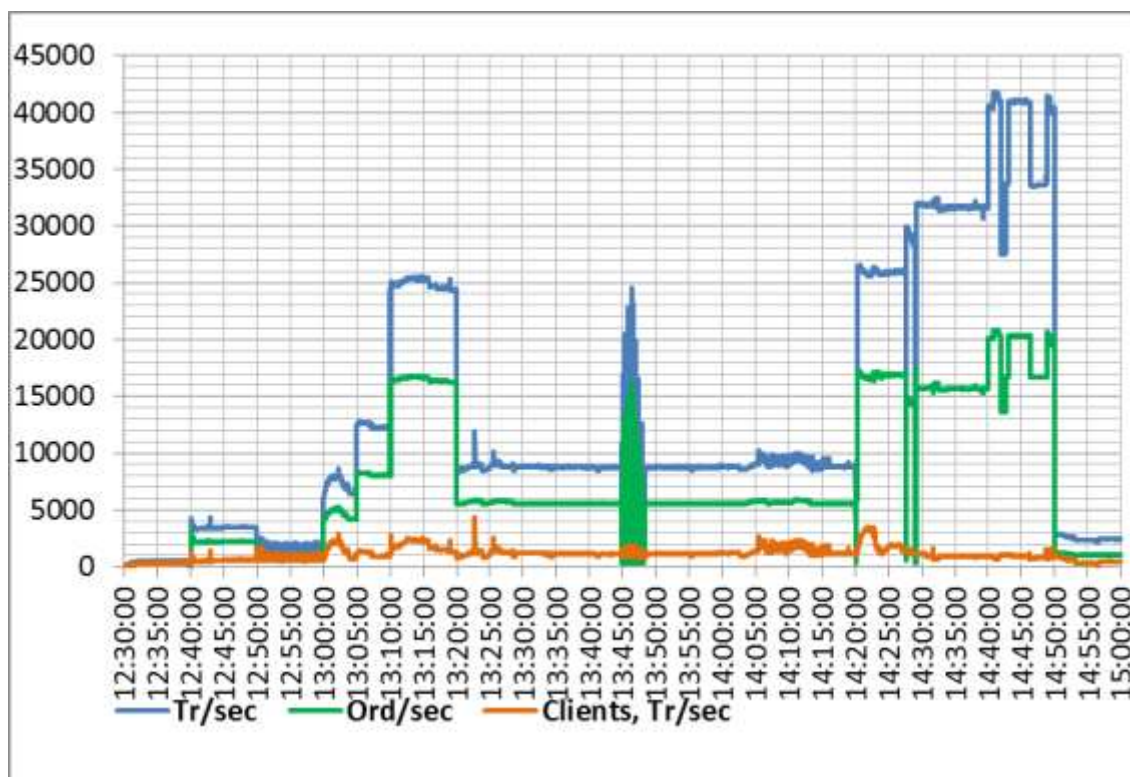
The testing was carried out on the hardware in DataSpace that is expected to become a primary data center in November 2016. We used the ASTS+ trading platform software to be deployed into production in February 2017. The ASTS+ trading platform is backward compatible with the version currently in production in terms of user interfaces and functionality which fact was confirmed during the testing.

The table below shows comparative performance in testing in 2015 and 2016.

	Transactions	Orders	Trades
Values reached (units), 2015	77,100,000	41,500,000	3,000,000
Values reached (units), 2016	113,454,567	65,924,897	1,300,000
Max processing rate (units per sec), 2015	19,314	7,826	700
Max processing rate (units per sec), 2016	41,822	20,862	1,775
Performance improvement, 2016/2015, %	+116%	+166%	

The maximum trade frequency was not reached in testing. The tested peak frequencies exceeded peaks of the real FX market trading sessions. The values in the table are not the maximum possible.

The graphs below show transaction and order frequency:



Clients generated 8.5% of the transactions.

We expect that the performance of the Securities market ASTS trading platform in production would be 24,000 transactions/sec when deployed in the DataSpace facility.

Maximum processing rate of the production matching engine in ASTS+ version will be artificially limited to approximately 35,000 transactions/sec to preclude data publishing lags at gateways. Higher values on the graph are given according to the testing with limiting component turned off.

Update and unification of the gateway hardware in accordance with the plan of moving all gateways into the exchange datacenter allowed increasing substantially the maximum number of orders and trades during the trading session. Information capacity of the Equity & Bond Market trading system exceed 10 million trades and 200 million orders.

The FX Market trading and clearing system

The testing was carried out on the production version of the ASTS+ trading platform deployed in production in 30 May 2016. The trading system was updated to replace production software and hardware operating in 2011 to 2016. Comparative analysis

makes sense with regard to 2014 as the FX market load tests in 2015 were run using a release candidate of ASTS+.

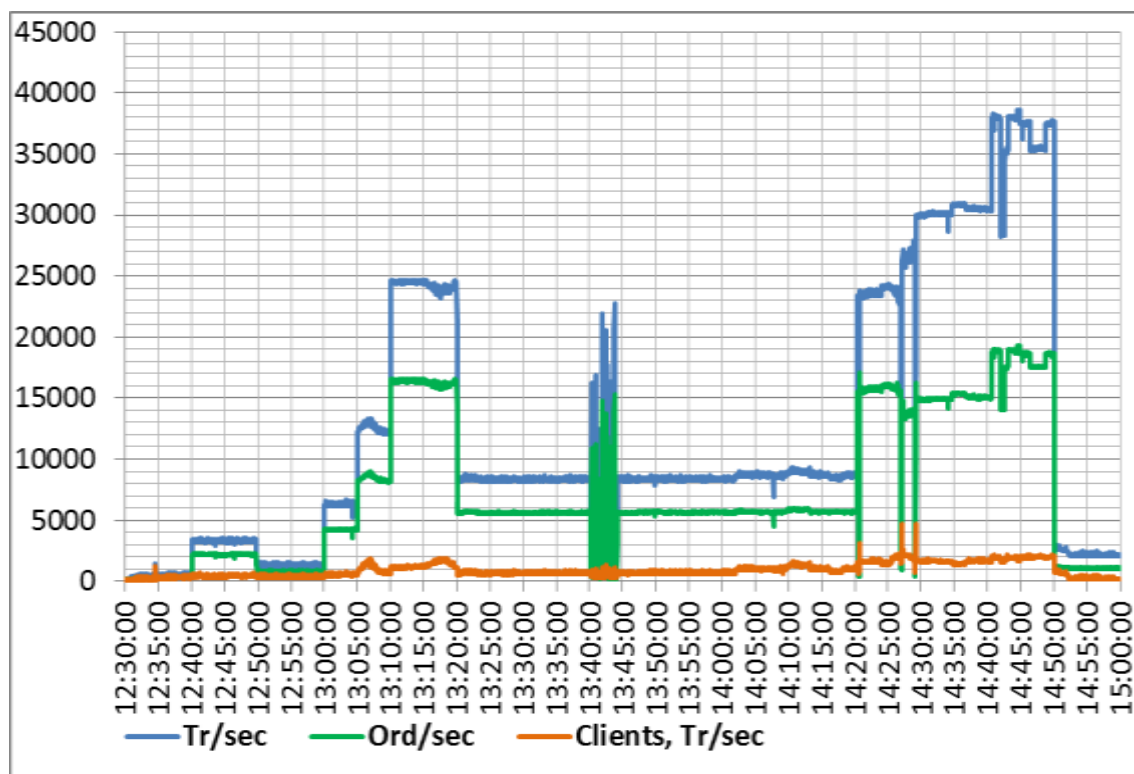
DataSpace was the primary site in the testing that allowed the participants to check whether they are ready to the migration of the primary data center in November 2016. The results obtained during the testing and internal testing data in M1 were identical as the DCs have identical server hardware.

The table below shows comparative performance in testing in 2014 and 2016.

	Transactions	Orders	Trades
Values reached (units), 2014	65,568,622	32,994,049	747,539
Values reached (units), 2016	107,447,400	64,061,937	2,583,400
Max processing rate (units per sec), 2014	11,324	5,700	700
Max processing rate (units per sec), 2016	38,609	19,322	1,720
Performance improvement, 2014 / 2015, %	+240%	+238%	

During load testing, the number of trades exceeded 25x the peak values of the production market in 2015. The peak trade frequency was not reached. Peak frequencies of test trades exceeded 5x-6x peaks of the production FX market. The values in the table are not the maximum possible.

The graphs below show the frequency of transactions, orders and transactions by clients – testing participants.



Clients participating in testing generated 7% of the transactions.

Maximum processing rate of the production matching engine at FX market is artificially limited to approximately 30,000 transactions/sec to preclude data publishing lags at gateways. Higher values on the graph are given according to the testing with limiting component turned off.

Update and unification of the gateways hardware in accordance with the migration plan allowed increasing substantially the maximum number of orders and trades during the trading session. Information capacity of the FX Market trading system exceed 10 million trades and 200 million orders.

The Derivatives Market trading and clearing system

Due to the introduction of TWIME, a new transactional binary protocol, profile and distribution of transactions submitted via different access protocols during the testing changed compared to the previous load testing. Transactions distributed via TWIME, FIX and CGate in the following proportion that is as close as possible to the production environment:

CGate: 50% of transactions;

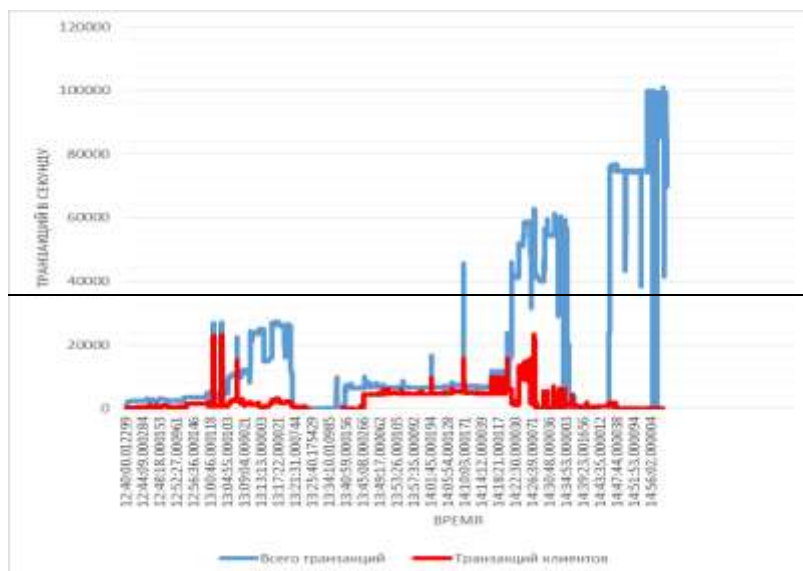
FIX: 10% and

TWIME: 40%.

The order-to-trade ratio in the testing was also near the production value. 209 million transactions were send and 3.5 million trades were executed during the testing. The peak performance was 101,000 messages/sec.

During the testing, we carried out the scheduled intraday clearing session. Despite large volumes of orders and trades, clearing was performed as usual within the established time frames.

The graph below shows a transaction load on the Derivatives Market trading system:



Subsystem for real time monitoring of the trading system parameters and market activity

The monitoring facilities operated well and provided data visualization in graphic form. Message signals were produced in accordance with the established criteria; data was collected to the monitoring database without fail. Operation of the monitoring system did not influence the facility performance.

ASTS Gateways

Outdated gateways and remote client gateways were disconnected during the testing session as they will be turned off after the DataSpace becomes the primary DC.

The infrastructure and updated gateways at the Exchange's technical centers operated with minimal latency in updating data tables with transaction frequencies reaching 30,000 and 35,000 transactions per second on the FX Market and Equity & Bond Market, respectively.

Therefore, unification of gateway characteristics as part of the program aiming to update the hardware and move gateways to the Exchange's DC allowed the maximum frequency of normal operation of the trading system coupled with gateways to be 2.5 times higher than in 2015 (12,000 transactions/sec for the majority of servers).

The testing has proved that main hardware currently installed at the primary and DR data centers of Moscow Exchange is balanced in terms of its performance.

Communication channels and remote gateways

ASTS

The ASTS trading platform employs gateways operating in regional technical centers.

To ensure proper operation of gateway servers on each market, at least 8 Mbit/s of network channel bandwidth is required for every gateway software instance of every market.

When constant transaction frequency exceeded 15,000 transactions per second, some remote gateways started delaying data tables if hardware specifications and/or data channel throughput were not sufficient to sustain increased transaction load. Implementation of measures aimed at updating technical facilities in regional data centers would allow fixing this.

SPECTRA

Due to results of this year's load testing, minimum bandwidth requirements for network are set as follows:

To ensure proper operation of client gateways, at least 30 Mbit/s bandwidth with the ping time of 1,000 microseconds maximum is required. To ensure proper operation of client gateways/trading terminals at least 4 Mbit/s bandwidth is required per every Exchange software copy. If a feed includes full order/trade log (FORTS_ORDLOG_REPL/FORTS_DEALS_REPL), a bandwidth of up to 20 Mbit/s and up to 10 Mbit/s is required for access servers and bridges respectively. To get accelerated replication feeds, clients should have at least 50 Mbit/s bandwidth to the trading system.

We strongly recommend that participants check their bandwidth and the quality of their local networks as the network with the high packet losses deteriorates substantially latency and may cause considerable delays in data transmission.

Bridges to the trading system

To estimate the time needed for the trading system to respond to orders, we used a generator entering and deleting orders via a Linux version of ASTS Bridge (libmtesrl.so library). The generator was launched on a server typical for colocation clients of the Exchange that was installed in a trading network segment (the gateway server).

To estimate the time needed for the Derivatives market's trading system to respond, we used a monitoring system, logging the transaction feeds from participants on the side of the Exchange and the Cgate API installed in the M1 colocation zone.

Latency for transactions in the Equity & Bond Market and FX Market trading systems

On the Equity & Bond Market and the FX Market, the average time of receiving a response to a transaction by ASTS Bridge when frequency was up to 20,000 transactions per sec was as follows:

- Mean: 230 microseconds (vs. 350 microsecond in the range of up to 15,000 in 2015)
- 90% of responses: no more than 270 microseconds
- 99% of responses:
 - no more than 400 microseconds when the load was at the production standard level of less than 50 trades/sec;
 - 1,500 microseconds when the load was at 500 trades/sec.
- 99.9% of responses: no more than 600 microseconds upon the regular load for the production environment.

Consistency of response time on the Equity & Bond Market has improved manifold compared to that shown with the trading platform in production in the M1 data centre. The enhancement has occurred due to the hardware update and optimisation of the trading system code. We did not see any difference between the M1 and DataSpace with regard to response time consistency on the FX Market as the production FX system already runs on the new hardware.

Latency for transactions in the Derivatives Market trading system

During the testing, door-to-door latency was:

- less than 250 microseconds at the load of 50,000 transactions/sec at TWIME gateways and
- less than 300 microseconds at CGate.

At peak loads, i.e. 80,000 transactions/sec, response time might have increased at order entry.

The most likely response time (median) was 250 microseconds with 99% of responses coming not later than in 1,000 microseconds. In this case, we assumed that the peak frequency would be 20,000-30,000 transactions/ sec in the next year. At the same time, the response time may reach 2,000-5,000 microsecond in periods of high simultaneous activity of users due to the specially defined restrictions in the trading system.

Transactional FIX gateways of ASTS

Hardware of FIX gateways was renewed in early September 2016 to become identical for all addresses.

FIX gateways operated well throughout the whole transaction frequency range. It should be noted that participants demonstrated low activity. The Exchange's load generator did not use FIX protocol.

FIX/FAST UDP multicast marketdata of the Equity & Bond Market and FX Market

Hardware of FAST gateways was renewed in early September 2016 to become identical in each of the DCs for the Equity & Bond and FX Markets. In the testing session, multicast sources were moved to DataSpace.

The FX Market servers operated as usual.

On the Equity & Bond Market, FAST servers delayed versus the market with regard to quote publication when the load exceeded 20,000 transactions/sec. Any delays at the production site are not possible.

When the ASTS+ is rolled out in February 2017, the Equity & Bond Market FAST servers settings will be optimised to preclude any delays throughout the whole transaction frequency range, similar to the FX Market.

UDP multicast traffic reached the following values in each copies A and B:

- 18 Mbit/sec in every feed with anonymous active orders;
- 18 Mbit/sec in every feed with aggregated orderbooks;
- 15 Mbit/sec for market statistics updates;
- 1 Mbit/sec as aggregate for snapshot and instrument descriptions.

Participants connecting to the service via data distribution channels are recommended to schedule carefully their subscriptions to data feeds and take into account network bandwidth as the total traffic of two FAST lines of the FX Market and Equity& Bond Market in copies A and B in total may reach 400 Mbit/sec.

In the production environment, peak FAST traffic for the FX market would most probably correspond with network requirements stated above. Requirements for the Equity & Bond Market may be decreased twofold considering its statistics.

Recommendations given at <http://www.moex.com/a1160> are applicable to each FAST line.

FAST UDP multicast marketdata servers of the Derivatives Market

Measurements of trade velocity at co-location facilities showed that data flows via multicast are identical to those via native API.

We used servers running the "full ORDERS-LOG" topic:

- Grouped into 750 microsecond quanta similar to Plaza II/CGate;
- Fully on a real-time basis.

The bandwidth per different transaction loads is given in tables below.

The aggregated OrderBook and trade topics; Incremental and Snapshot:

Load, tr/sec	3,000	6,000	12,000	25,000	40,000	50,000
FUT-BOOK-1	0.114	0.183	0.259	0.368	0.546	0.672
FUT-BOOK-1-snap	0.016	0.016	0.016	0.016	0.016	0.016
FUT-BOOK-5	0.188	0.306	0.465	0.690	1.028	1.275
FUT-BOOK-5-snap	0.024	0.024	0.024	0.024	0.024	0.024
FUT-BOOK-20	0.302	0.440	0.778	1.072	1.628	2.002
FUT-BOOK-20-snap	0.024	0.024	0.024	0.024	0.024	0.024
FUT-BOOK-50	0.416	0.579	1.153	1.643	2.531	3.214
FUT-BOOK-50-snap	0.040	0.040	0.040	0.040	0.040	0.040
FUT-TRADES	0.650	1.297	1.702	0.988	1.164	1.454
FUT-TRADES-SNAP	0.016	0.016	0.016	0.016	0.016	0.016

The full ORDERS_LOG topics, data groups (ORDERS_LOG_Grp) and pure real-time (ORDERS_LOG_Onl); Incremental and Snapshot:

Load, tr/sec	3,000	6,000	12,000	25,000	40,000	50,000
ORDERS_LOG_Grp	1.960	3.586	8.011	19.073	46.380	56.120
ORDERS_LOG_snap	1.000	1.000	1.000	1.000	1.000	1.000
ORDERS_LOG_Onl	2.417	4.559	15.527	42.847	67.383	82.881
ORDERS_LOG_snap	1.000	1.000	1.000	1.000	1.000	1.000

Traffic is in megabits per second per one feed (FEED A)

Conclusions

The Equity & Bond Market, the FX Market

1. Performance of the FX market trading and clearing system has increased by 3.4x in the release of 30 May 2016 vs. 2011-H1 2016.
2. It is expected that performance of the Equity & Bond Market's trading system will increase twofold in 2017 vs. 2011-2016.
3. The gateway hardware has been renewed and unified. It demonstrated sufficient performance within the established transaction frequency range.

4. The migration to using only the gateways located inside the Exchange's data centers may be completed in November 2016.

Derivatives Market

1. SPECTRA's performance is sufficient to meet demands of participants even at peak loads with respect of order procession and market data distribution.
2. Participants are strongly advised to consider abandoning leased gateways that are an out-of-date part of the "exchange-participant" infrastructure.
3. The required network bandwidth for clients who wish to use the FAST service to receive ORDERS-LOG on a "pure real-time basis" is minimum 100 Mbit/sec per feed. To receive two feeds, FEED A and FEED B, or data from more than one market, the 1 GBit/sec bandwidth is recommended.

New network requirements will be published on the Moscow Exchange's website at <http://www.moex.com/a1160>.