DEUTSCHE BÖRSE GROUP

Insights into trading system dynamics

Deutsche Börse's T7®

July 2018

Agenda

2

Overview

T7® Technology Roadmap

Deutsche Börse is pursuing its Technology Roadmap to deliver innovative and superior trading technology.

Recent developments

- Introduction of partition-specific gateways in HY1 2018
- Rollout of T7 release 6.1 on 18 June 2018, including
  - support of standard orders via high-frequency sessions
  - support of short message layout for complex instruments as well as for standard orders
  - introduction of Book-or-Cancel (BOC) attribute for quotes
  - enhancements to Eurex EnLight
  - reduction of median request-response latency by 8 µs to 52 µs for partition specific gateways

Outlook

Launch of HPT file service in July 2018

Launch of White Rabbit time synch service in Q3 2018

Reduction of the number of Eurex low frequency gateways (4 instead of 6)

T7 Release 7.0 on 3 December 2018:

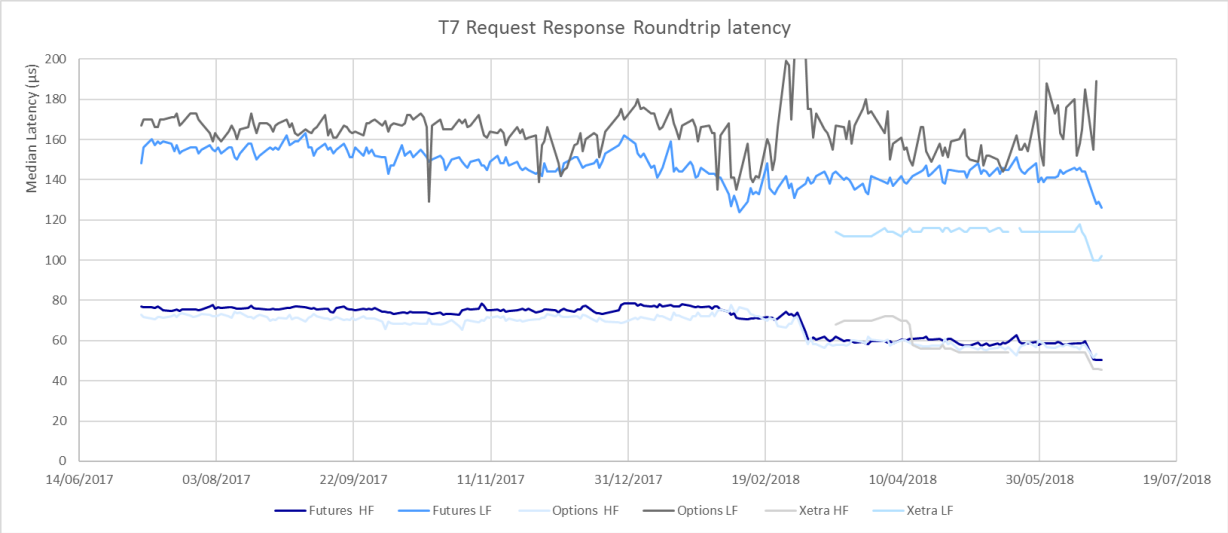
- Further optimization of partition-specific gateways outbound processing
- Support for passive liquidity protection for options

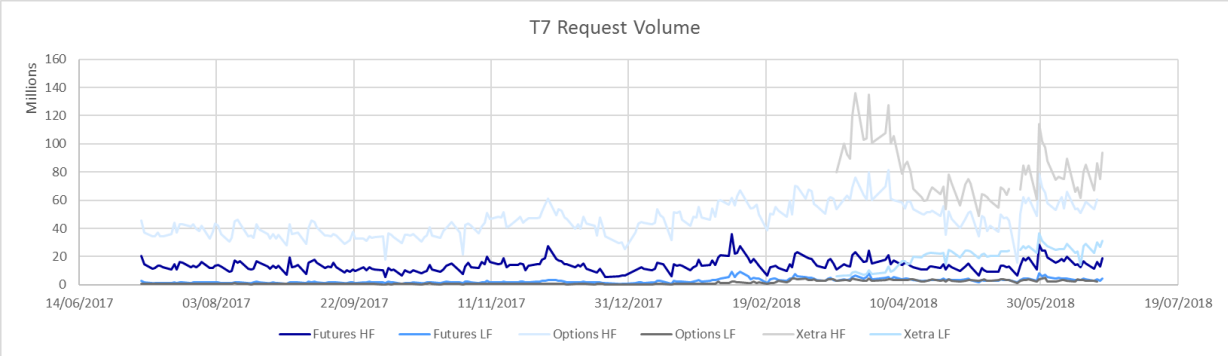
Launch of EOBI for all Eurex options in February 2019

For further details about T7 please visit our websites:
www.eurexchange.com/t7 and www.xetra.com/xetra-en/technology/t7

Processed transactions and response times

T7 request – response round-trip times



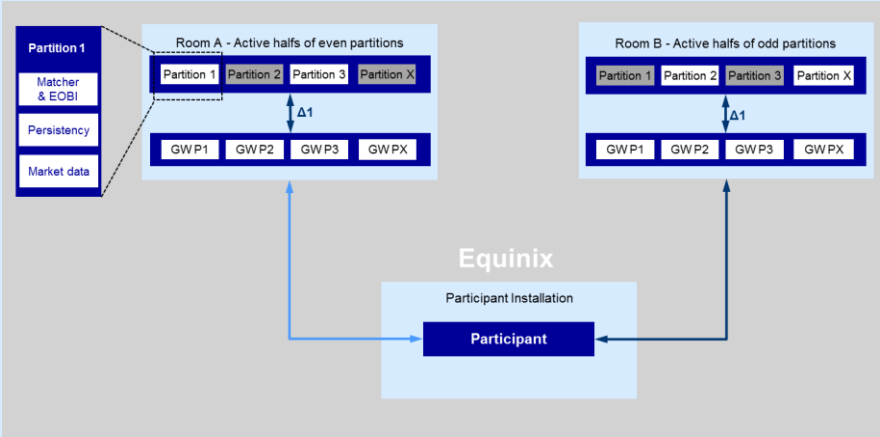


- Deutsche Börse has continuously invested in its trading system and has been able to reduce the processing time of technical transactions significantly.
- Introduction of PS gateway for Eurex in Q1/2018 and Xetra in April 2018 reduced the median round-trip by 19 µs to below 60 µs, while having an all time high of processed transactions.
- Release 6.1 further lowered the median latency for requests sent via PS gateways to below 52 µs.

Overview of T7

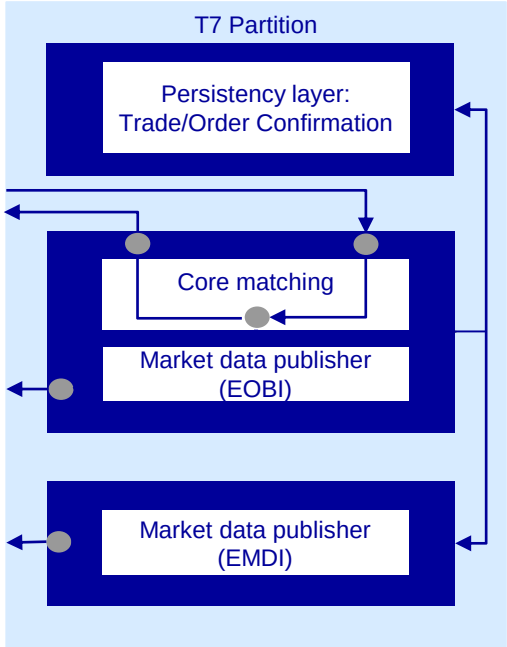
- T7 consists of partitions. Here, a partition is a failure domain in charge of matching, persisting and producing market data for a subset of products. Each T7 partition is distributed over two rooms in the Equinix data centre.
- There are 10 Eurex T7 and 10 Xetra T7 partitions.
- Separate partitions are used for Vienna (XVIE), Dublin (XDUB) and EEX (XEEE) markets.
- The reference data contains the mapping of products to partition IDs.
- With the introduction of partition-specific (PS) gateways there is a one-to-one mapping of active PS gateways to partitions. The default active PS gateways are located on the same side as the active matching engines.
- 6* low-frequency (LF) gateways allow access to all Eurex partitions and the separate EEX partition.
- 4 low-frequency gateways are shared between all cash markets (XETR, XDUB, XVIE).

- Note that the active half of a partition and its partition-specific gateway is either on side A (for even partitions) or on side B (for odd partitions).
- Only in case of the failure of a matching engine or a market data publisher, the active half of the service will shift to the other room.



*Note: the number of Eurex LF gateways will be reduced to 4 in Q3 2018

Processing inside a partition



Orders/quotes entered for a specific product are sent by the gateway server to the respective matching engine (residing in a partition).

The matching priority is assigned when the orders/quotes are read into the matching engine.

The core matching component works as follows:

- when an order/quote arrives, it is functionally processed (e.g. put in the book or matched).
- handover of data to the EOBI Market data publisher
- handover of all data resulting from the (atomic) processing of the incoming order/quote to the market data and persistency components in the partition.

Resulting responses and private broadcasts are sent out in the following order:

- direct response to the order/quote entered (for persistent as well as for non-persistent orders and quotes)
- fast execution information for booked orders/quotes (in case of a match)

In case that during this phase several new orders/quotes transactions arrive at the core matching component the processing remains unchanged, i.e. no batching takes place.

The generation of market data other than EOBI (by the market data publisher), listener broadcasts and trade confirmations (by the persistency server) are done on separate servers. Hence the order of the resulting messages from these servers is not deterministic.

Middleware, network, hardware and OS overview

T7 uses state-of-the-art infrastructure components

Intel Xeon E5-2667 v3 CPUs (Haswell) on all servers hosting core services (Matching engines, un-netted market data publishers)

Intel Xeon Gold 6144 CPU (Skylake) for partition-specific gateways.

Intel Xeon E5-2690 CPUs (Sandy Bridge) or E5-2683 v4 (Broadwell) on non-performance critical servers.

The operating system used is Red Hat Linux 6.9 with real-time kernel on all core components.

T7 internal communication between its core components is based on Confinity Low Latency Messaging using an Infiniband network in order to deliver the required speed, capacity and stability requirements.

T7 network access

Deutsche Börse offers Trading Participants to connect via 10 GbE cross connects to its T7 platform in the Equinix data centre.

The co-location 2.0 offering uses Cisco 3548x switches. All cables are normalized to give an overall maximum deviation between any two cross connects of less then +/- 1 m (+/- 5 ns).

Participant facing interface cards on the gateways and market data publishers use Solarflare EnterpriseOnload technology to bypass the kernel TCP stack.

9

Partition-specific gateway

Partition-specific gateway

Motivation

Latency jitter on parallel inbound paths has incentivized multiplicity to reduce latency. This led to higher system load at busy times and thus created higher, less predictable latencies. The introduction of a single (low-latency) point of entry addresses these issues.

Increased predictability

Each partition has only one partition-specific (PS) gateway assigned to it. This gateway operates in first-in first-out (FIFO) mode. Thus the reception sequence of the PS gateway determines the sequence of matching* (based on the time-stamp of the first bit of the frame that completes a ETI message).
Co-location 2.0 offers a highly deterministic, predictable and equal network access.

Reduced complexity

There is only one low latency entry point per partition. There is no need to probe multiple gateways to achieve best matching priority. All partitions are accessible via the low-frequency gateways.

Reduced latency

The PS gateway is tuned for highest throughput and offers a lower base latency than the previous HF gateways. In fact, the median inbound latency from gateway receive time (t_3n) to matching engine in (t_5) has been reduced by 8 µs compared to HF gateways. Improvements introduced with release 6.1 on 18 June 2018 reduced the median inbound latency by another 2 µs and the outbound latency (t6 to t4) by 5 µs.

* Note that in high load situations there is a chance that messages sent via LF gateway may reach the matching engine earlier, details in this presentation

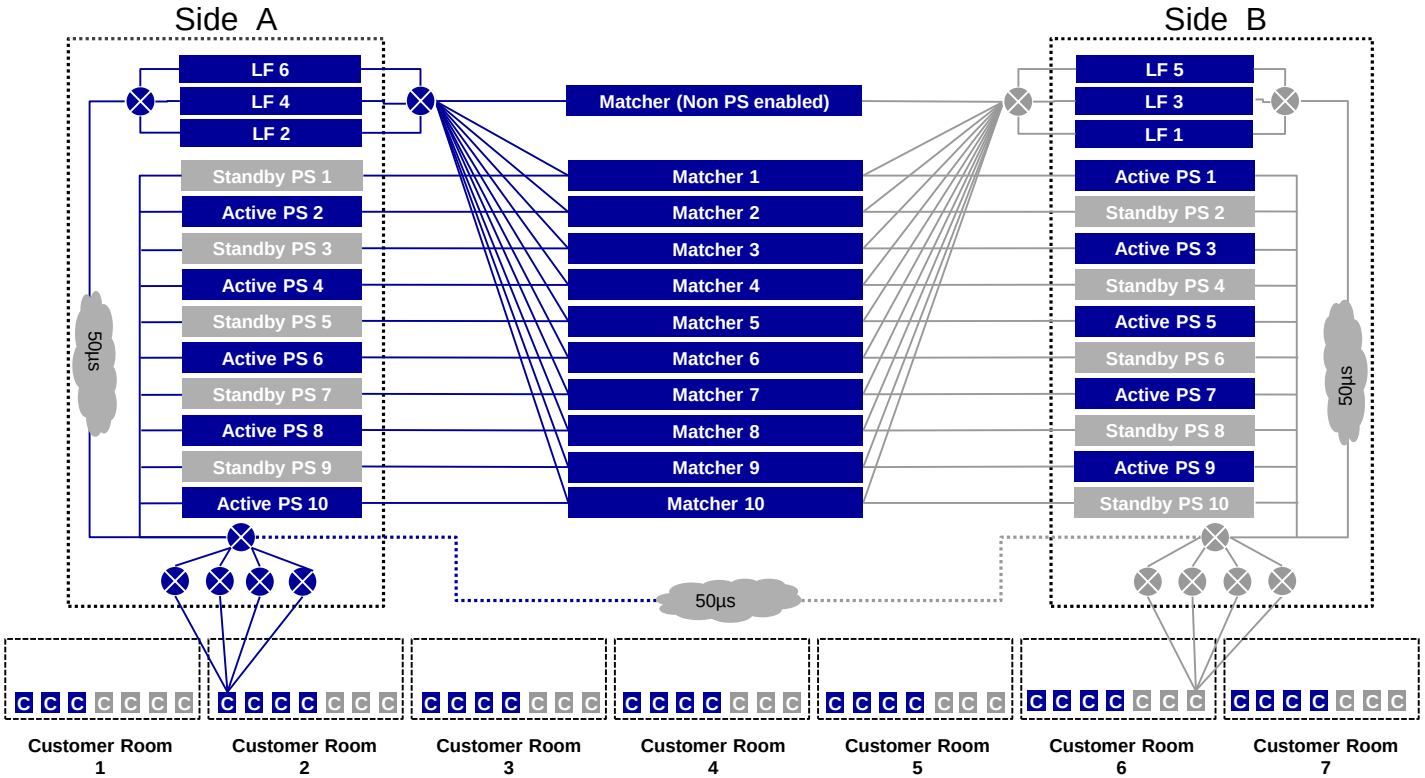
Partition-specific gateway

Migration

- Eurex migrated to partition-specific gateways between 22 January 2018 and 6 March 2018.
- Xetra migrated to the PS gateway architecture in two steps between 9 April 2018 and 16 April 2018
- Partition-specific gateways are the only low latency access to the Eurex and Xetra market on T7.
- EEX, Xetra Vienna and Xetra Dublin will not adopt the PS gateway concept.

Partition-specific gateway

Topology



The active partition-specific gateway resides on the same side as the active matching engine per default.
There is a network link between side A and B via the distribution layer switches with a one way latency of 50 µs.
This guarantees that all partition-specific gateways are reachable via a single line in case of a failure.

Note that PS gateways are available only for Xetra and Eurex markets, whereas EEX, Xetra Vienna and Xetra Dublin offer access via low-frequency gateways only.

The network link to LF gateways is around 50 µs slower than the access to PS gateways via 10 Gbit lines.

Partition-specific gateway

Implementation

The partition-specific gateway uses the same ETI protocol as the all ETI gateways (current low-frequency and previous high-frequency).

It offers the same functionality as the high-frequency gateways, but only for a single partition.

Session setup

All high-frequency sessions are eligible to connect to a partition-specific gateway.

A session may only connect to a single gateway at any given point in time.

There is a maximum number of sessions per Participant allowed to login to a single PS gateway at any given point in time. This limit is currently set to 80.

Connection

The connection process follows the three-step logon procedure, with a ConnectionGatewayRequest message to retrieve the assigned active and standby PS gateway from the connection gateway, followed by a Session Logon at the PS gateway. The initial Connection Gateway Request message has to contain the target partition ID.

You may send a Session Logon to the standby PS gateway to test network connectivity. Those logons will be rejected with the appropriate error code (refer to the respective ETI manual for details).

Partition-specific gateway

Failover

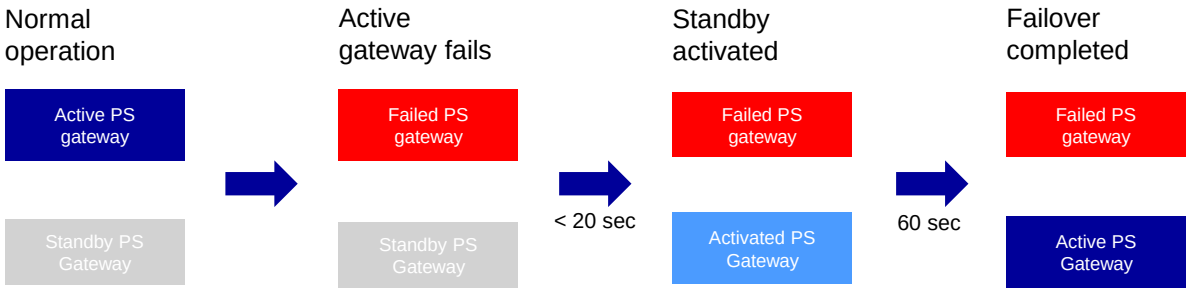
When a failure of a partition-specific gateway is detected, all sessions logged in via that gateway will be logged out and their orders and quotes are deleted.

Subsequently the standby PS gateway will be activated and allow session logins. A connection request should be sent to the connection gateway. The response will indicate the active PS gateway and the session can then login to this PS gateway.

There will be an activation phase during which no order management via the activated PS gateway will be possible. This is set to 60 seconds to allow Participants some time to evaluate the situation and re-login.

After the activation phase an ETI ServiceAvailability broadcast will be sent to the connected sessions and order management service will be available.

Schematic partition-specific gateway failover



Throttle and session limits

In order to protect the trading system, T7 has several measures in place to ensure that its most vital components are not harmed by a malfunctioning client application. Therefore transaction limits are imposed on T7 sessions.

All ETI sessions (HF and LF) are available with throttle values of 150 messages/sec or 50 messages/sec.

Furthermore LF sessions that cannot enter orders/quotes but can only receive trade and listener broadcasts are available (at a reduced price).

All ETI session types have an assigned disconnect limit of

- 450 for sessions with a throttle value of 150 messages/sec, i.e. a session will be disconnected in case of more than 450 consecutive rejects due to exceeding the transaction limit (throttle).
- 150 for sessions with a throttle value of 50 messages/sec, i.e. a session will be disconnected in case of more than 150 consecutive rejects due to exceeding the transaction limit (throttle).

Please note that in case the disaster recover facility is used, all ETI sessions will have a throttle limit of 30 messages per second.

For both limits, all technical transactions are counted using a sliding window.

The number of ETI sessions which can be ordered is limited. Currently, up to 80 sessions can be ordered. If more than 80 sessions are required please get in touch with your Technical Key Account Manager.

There will be a limit on maximum number of sessions per Participant and partition that can connect to a partition-specific gateway concurrently. This limit is currently configured to 80 sessions but subject to review.

15

Network access

Network access
Overview Co-location 2.0

Co-location 2.0 is an improved 10GbE connectivity introduced in parallel to the existing 10GbE network (Co-location 1.0).

Reduced complexity

- Fewer customer facing switches
- Customers may reach any switch from any data-center room
- Passive components in customer rooms, i.e. cables and patch panels only

Increased predictability

- Reduced latency variance
- Hardware refresh of switches (Cisco Nexus 3548x)
- Switches operate in cut-through mode, configured to use “warp mode” to minimize latency
- Switches exhibit very low switch jitter within the precision of measurement devices (+/- 4 ns), both for ETI and market data
- Equidistant cables with a tolerance of +/- 1m verified using an optical time-domain reflectometer (OTDR) and packet round-trip measurements
- One way latency of 2 µs

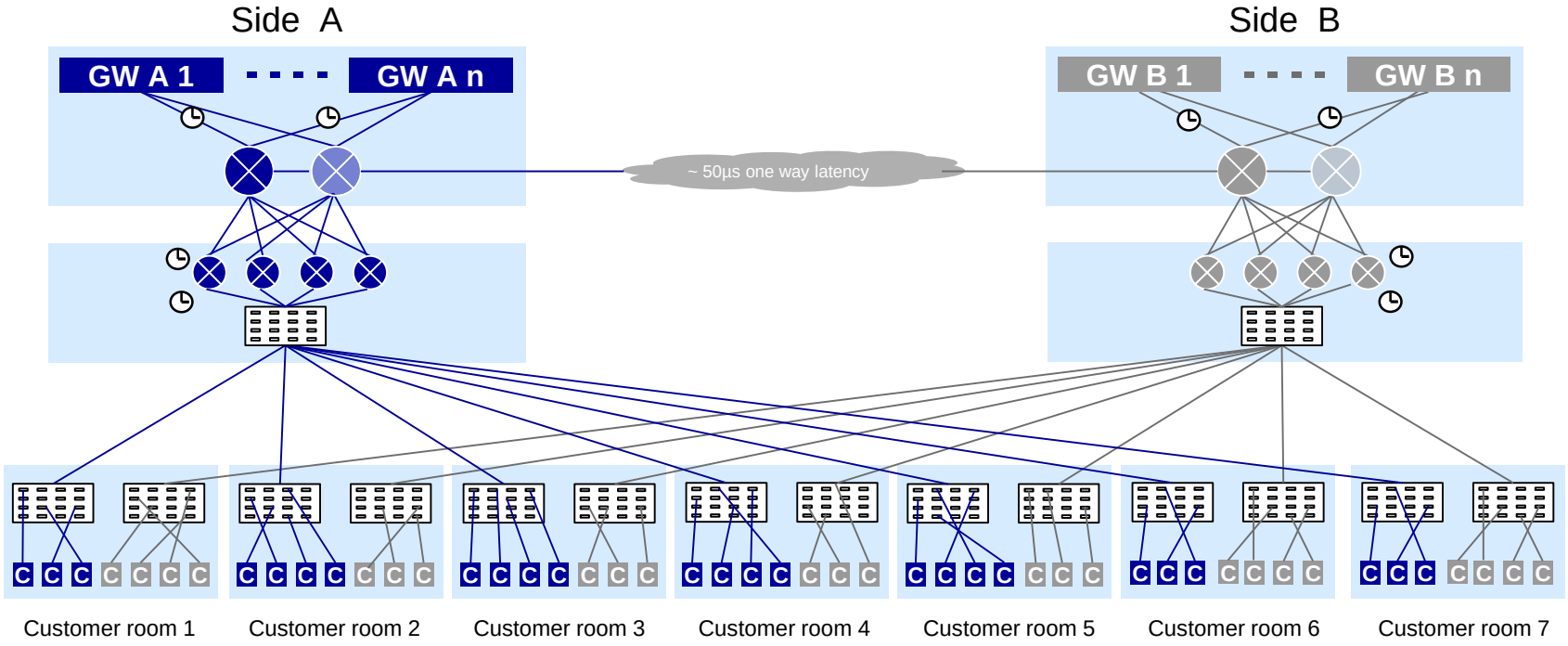
Improved monitoring

- Tapping and timestamping at network boundary and internally

Note: Co-location 1.0 will be decommissioned by 30 September 2018 (see Eurex Circular 013/18 and Xetra Circular 014/18)

Network access
Topology

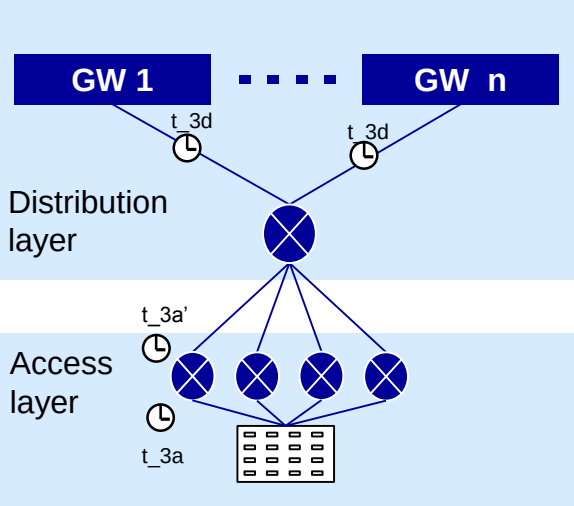
- 2 switches per gateway room per market (‘distribution layer’, only one market shown)
- Centrally located ‘access layer’ switches (Eurex: 4 per side*, Xetra: 2 per side)
- Customers can connect to any access layer switch from any of the 7 co-located rooms
- There is a separate Market Data network with same layout
- The network link between ETI side A and B distribution layer switches has a one way latency of ~50µs.



*** Note that for capacity reasons we may add a 5th switch on each side for Eurex ETI**

Network access

Latency profile



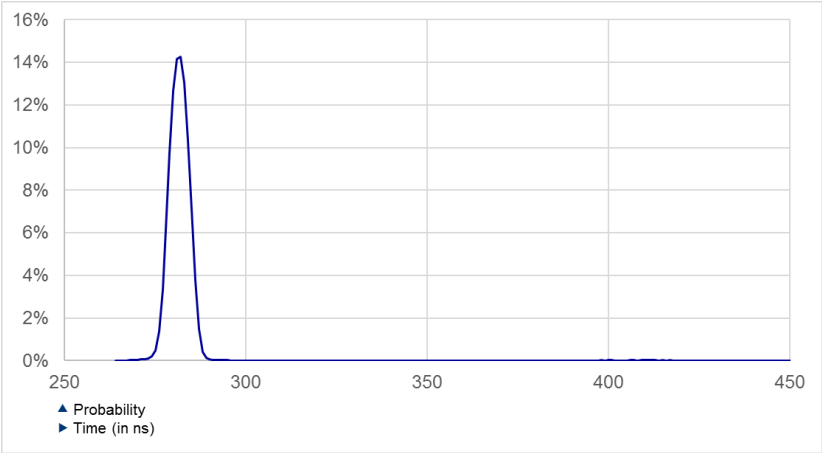
We are constantly measuring the latency profile of our co-location infrastructure using taps and aggregation switches that use hardware assisted timestamping on ingress port.

Extra care has been taken to ensure the best time synchronization between these timestamping devices. We use a combination of white rabbit and pps achieving a time synch better than 5 ns.

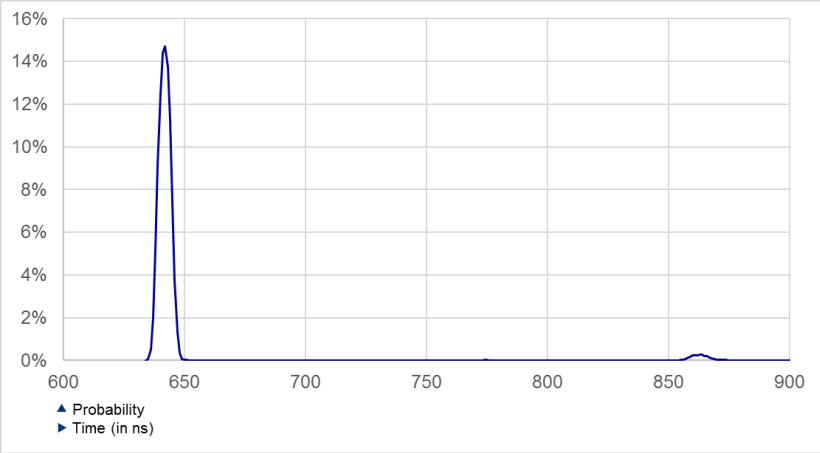
The latency profiles between the three measurement points is shown below. Latency between access layer and distribution layer is identical for all access layer switches within the measurement precision. The spectrum is very tight with a standard deviation below 5 ns.

Note that queues might occur in sharp bursts as all member cross connects are connected to access layer switches, which are connected to a single active distribution layer switch via a 10 Gbit line each. There are four access layer switches per side for Eurex and two for all cash markets XETR, XVIE, and XDUB. There is a single active 10 Gbit line connecting each PS gateway to a distribution layer switch. There are no PS gateways for the XVIE and XDUB markets.

Access layer switch latency (t_3a' - t_3a)

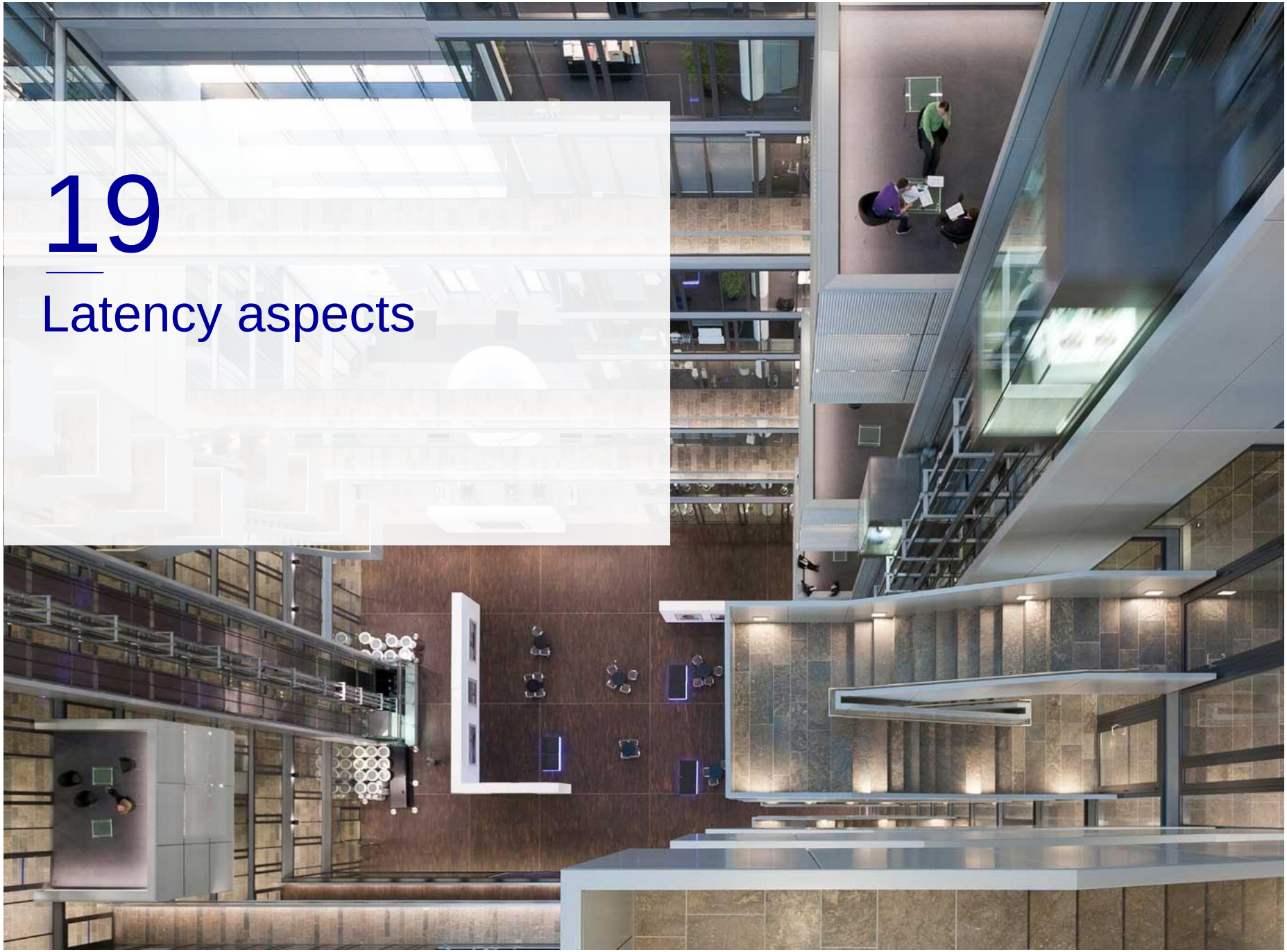


Access to distribution layer latency (t_3d - t_3a')



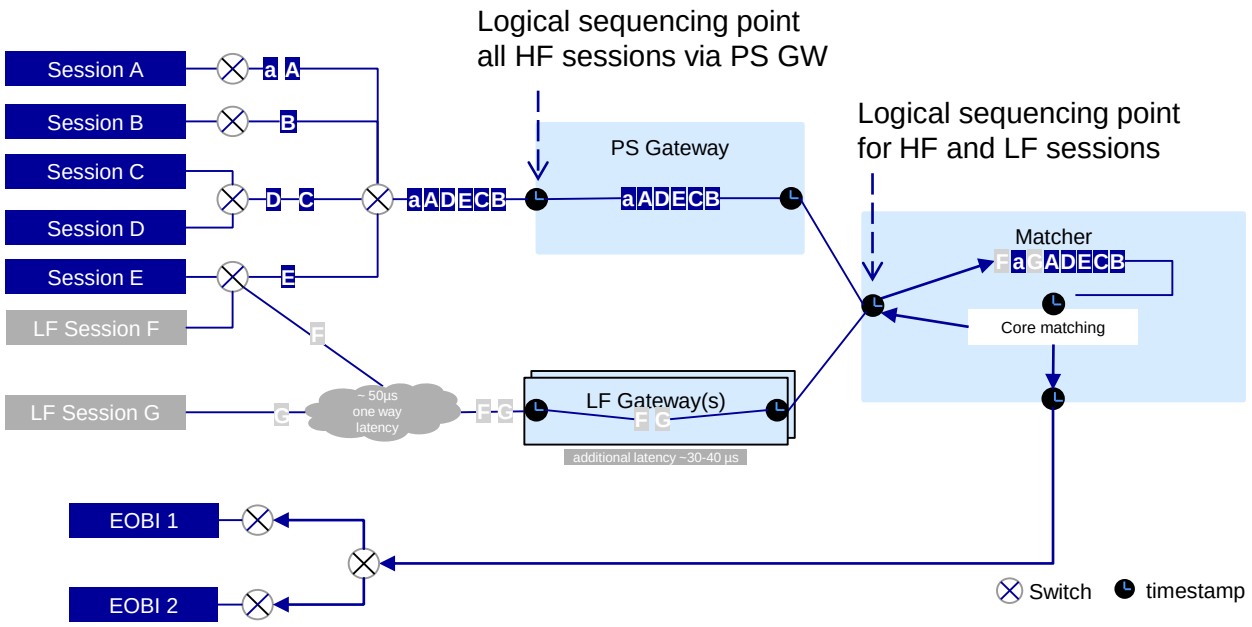
19

Latency aspects



Inbound message sequencing

PS gateway architecture



Inbound sequencing inside the T7 system takes place

- on the network in front of the PS trading gateway,
- in the matcher for messages from PS and LF gateways.
- Note that LF gateways have a 80-90 µs higher median latency (t_5 – t_3a).

Inbound ordering is preserved

- within all messages from one PS gateway
- between the messages sent from each LF/PS gateway to one matching engine (=partition).

T7 Latency

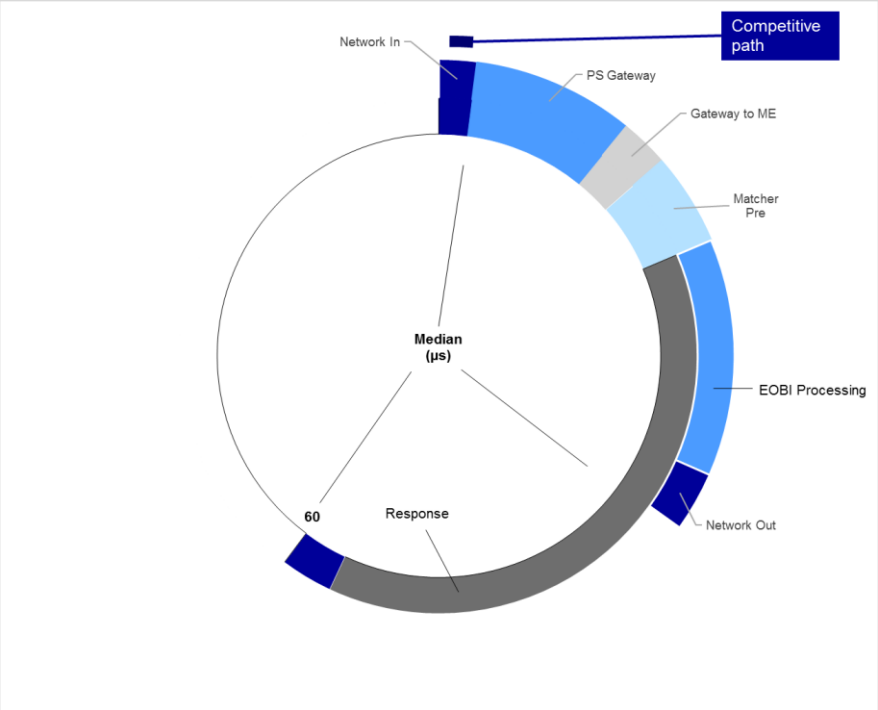
Composition

The diagrams below compare the composition of latency for Eurex Futures transactions for T7 release 6.0 and 6.1. The full circle represents 100 µs.

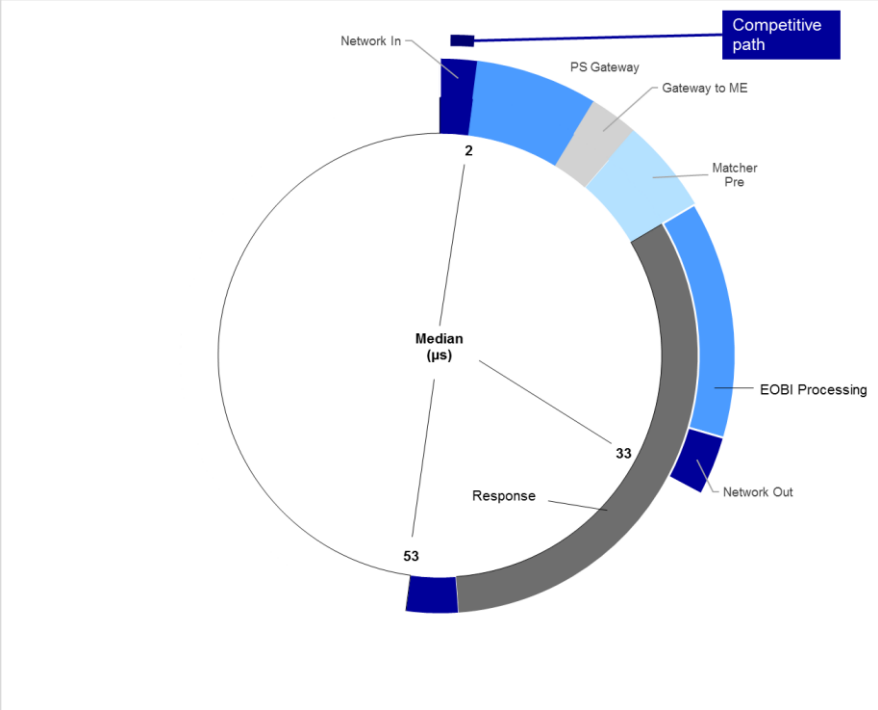
The latency in the request and response path has dropped significantly.

Note the latency difference between order book updates sent on the public path via EOBI and the private response.

Median latencies Release 6.0



Median latencies Release 6.1



T7 Latency
Composition

The below charts show a comparison of latencies since the last update (March 2018) for Eurex futures sent via PS gateways.

Tuning measures for the PS gateway lead to significantly reduced processing times for requests and responses.

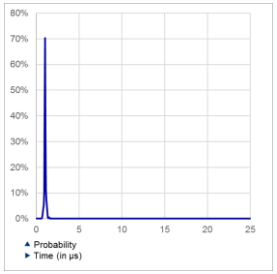
Other latencies are stable since the introduction of PS gateways.

Network response and market data include TCP/UDP stack on the server.

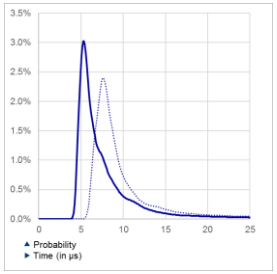
Dotted lines represent data from 19 March (T7 Release 6.0) and solid lines data from 19 June 2018 (T7 Release 6.1).

Request path and market data (EOBI)

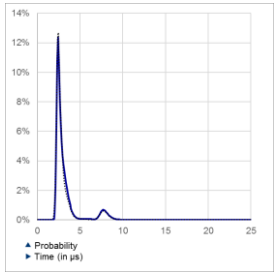
Network request (t_3n – t_3a)



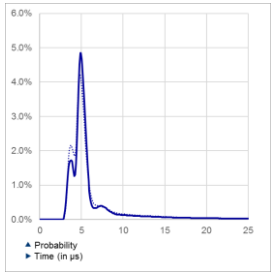
Gateway request (t_3' – t_3n)



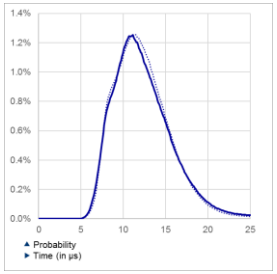
Gateway to matcher (t_5 – t_3')



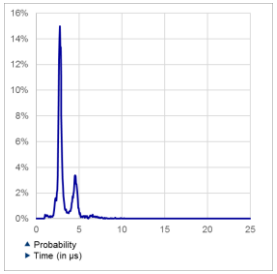
Matcher pre (t_7 – t_5)



EOBI processing (t_9 – t_7)

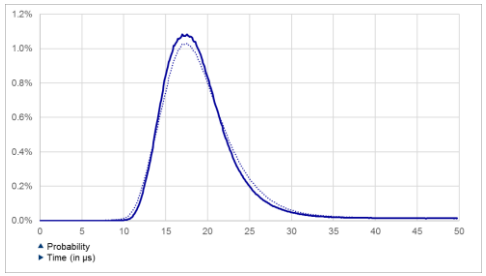


Network EOBI (t_9a – t_9)

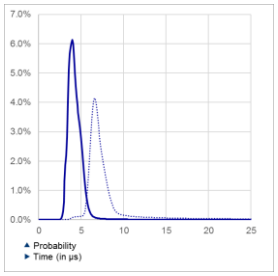


Response path

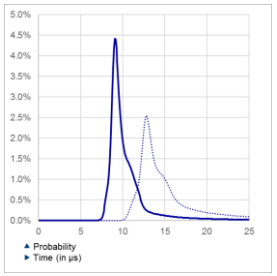
Matcher processing (response) (t_6 – t_7)



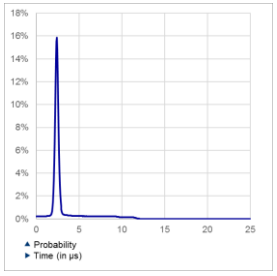
Matcher to gateway (t_4' – t_6)



Gateway response (t_4 – t_4')



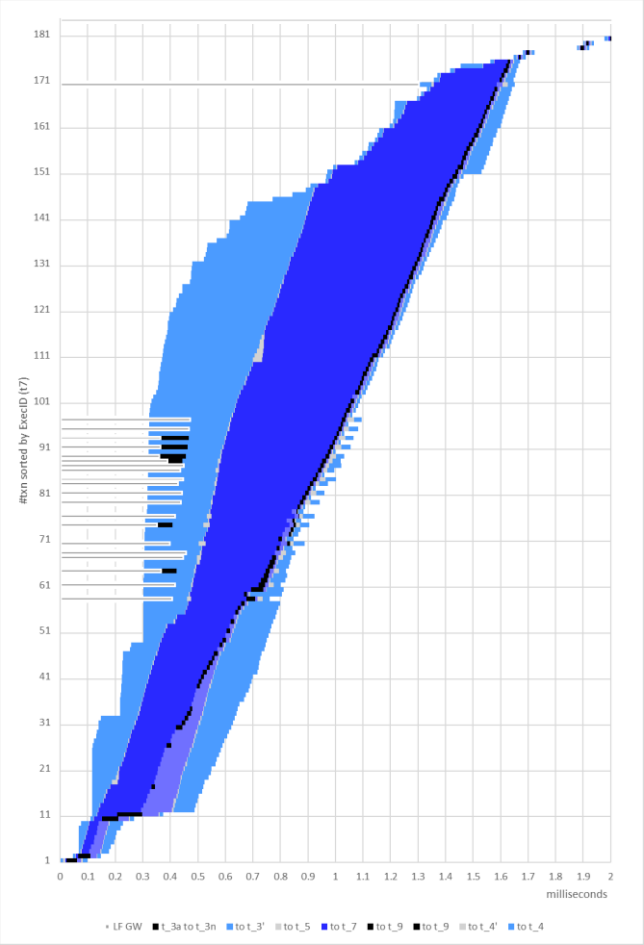
Network (t_4a – t_4)



PS gateway --- Motivation revisited

Microburst dynamics

Microburst dynamics FESX (18 Jun 2018)



During micro-bursts, the input into the trading system may be greater than the throughput capabilities. This in turn causes queuing which results in higher latencies.

Higher latencies cause risk (i.e. it takes longer to place/pull an order).

T7 provides real-time performance insights by providing up to seven timestamps with each response and key timestamps with every market data update.

The chart on the left shows the paths

- Co-location 2.0 access layer to t_3(n) to
- Gateway request Sending time (t_3') to
- Matching engine in (t_5) to
- Start matching (t_7) to
- EOBI SendingTime (t_9) [where available] to
- Matching engine out (t_6) to
- Gateway response received (t_4') to
- ETI SendingTime (t_4).

Typical throughput rates (1/ms) are 8000 at t_3n, ~170 at t_3', 140 at t_7 and 150 at t_4.

EOBI send times are usually well before the gateway send time of responses.

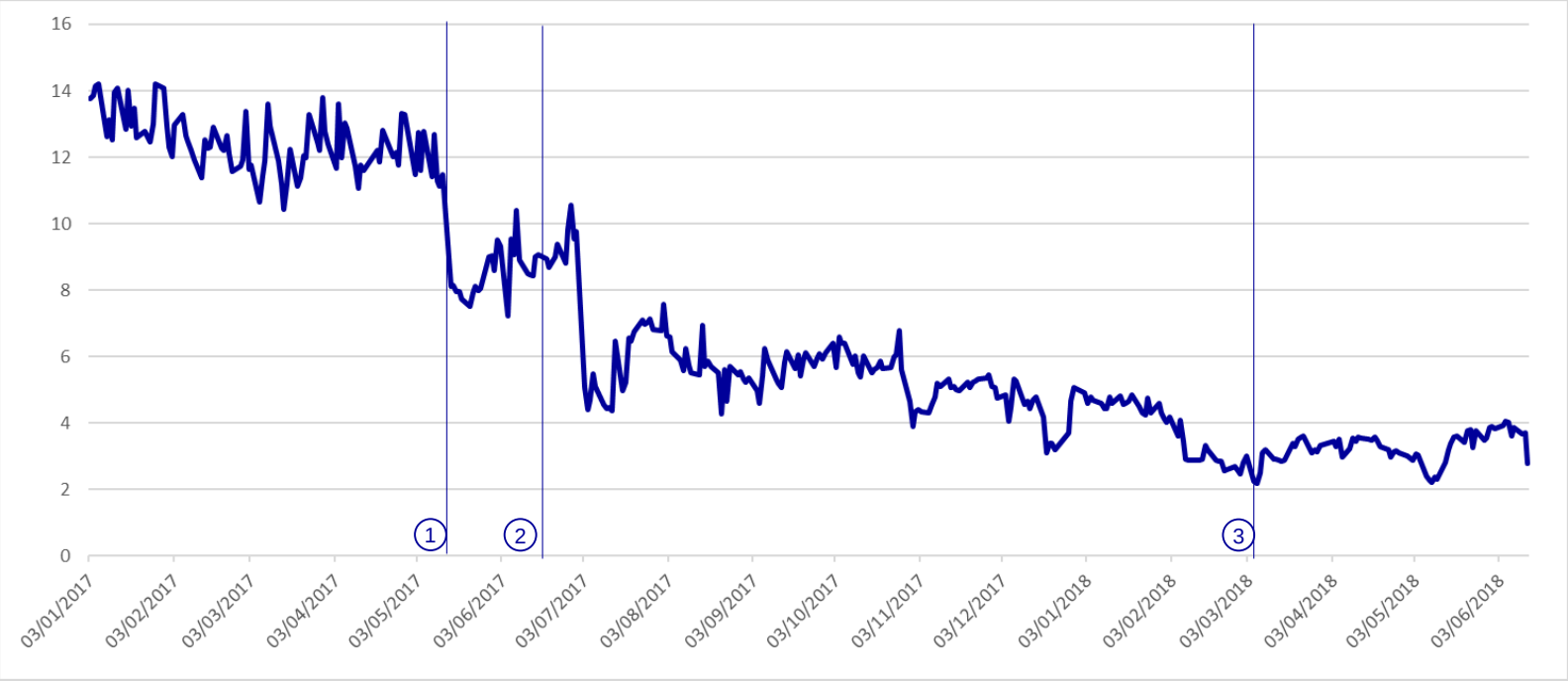
Note that base latency for requests entered via LF Gateways is 80-90 µs higher (~50µs network + 30-40 µs LF Processing and internal network). Nonetheless, in micro-bursts there is a possibility of requests sent via LF to overtake those that are queuing in the PS Gateway. This usually happens after the first 20-25 PS Gateway requests have been processed by the matching engine.

PS gateway --- Motivation revisited

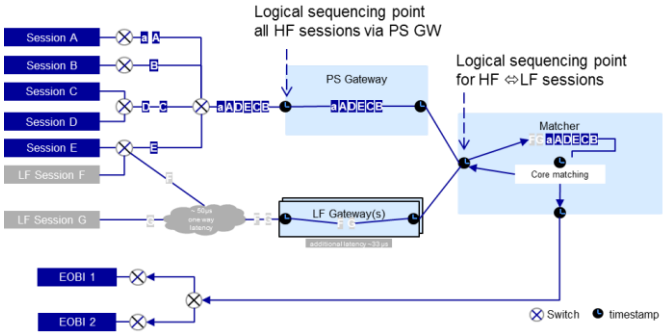
Multiplicity

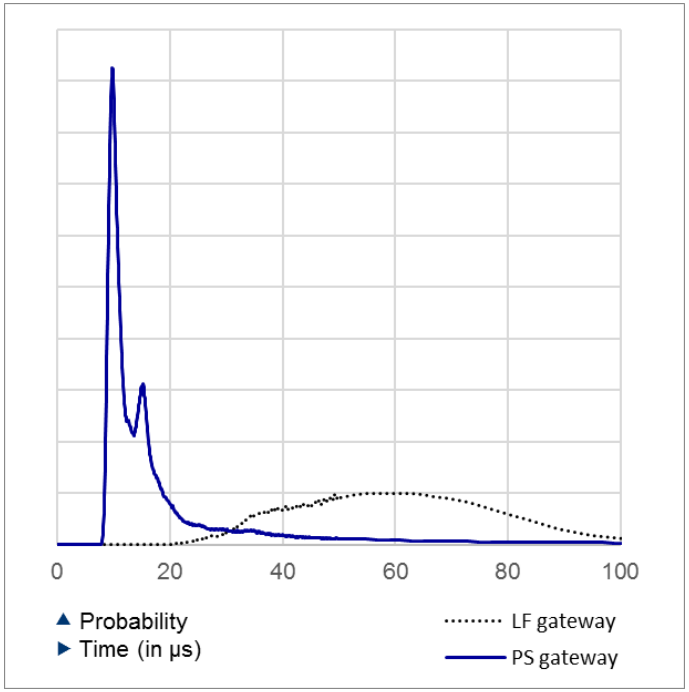
Latency jitter on parallel inbound paths has incentivized multiplicity to reduce latency.
This led to higher system load at busy times and thus created higher, less predictable latencies.
The introduction of a more deterministic network infrastructure (1), first-in-first-out (FIFO) processing of high-frequency gateways (2) and the recent migration to PS gateways as a single (low-latency) point of entry (3) led to a sizable reduction of multiplicity.

Below graph shows the number of sent vs. executed IOCs for the Eurex market.



PS Gateway versus LF Gateways





Using 10 GbE cross connects in co-location 2.0 offering access to the partition specific ETI gateways in the Equinix facility provides the fastest way for order and quote management in T7.

Low frequency gateways on the other hand allow access to all partitions of a market via a single session.

The base latency of the path to the matcher is more than 85 µs higher for LF gateways,

- 50 µs because of a slower network access from customer to the LF gateway (see top left and page 20)
- 35 µs because of slower processing of requests by the gateway and a slightly longer path to the matching engine (see bottom left)

In case of microbursts however the number of requests being sent to the PS gateway may be significantly higher than the throughput capabilities.

In those cases requests sent via a lightly loaded LF gateways may overtake those that queue in the PS gateway. This currently happens at around queue position 20-25, i.e. when requests queue behind at least 20-25 messages in the PS gateway.

The graph on the left hand side shows the latency distribution from gateway to matching engine, from t3_n for PS gateway, t3 for LF gateway to t5 (for partition 3).

Latency

Comparison of access types

The table below gives an overview of current latency figures of the T7 trading system. All times given are in microseconds.

Network timestamps (t[3489][aa'd]) are synchronized using pps and white rabbit. The time synch quality between these timestamps is thus <<5ns. Other T7 timestamps are subject to jitter of up to +/- 50 ns, so that t3d_t3n and t4_t4d have to be read with a grain of salt.

The latency difference of the inbound path (t_3a to t_5) of LF vs PS Gateways is 82/91 µs respectively.

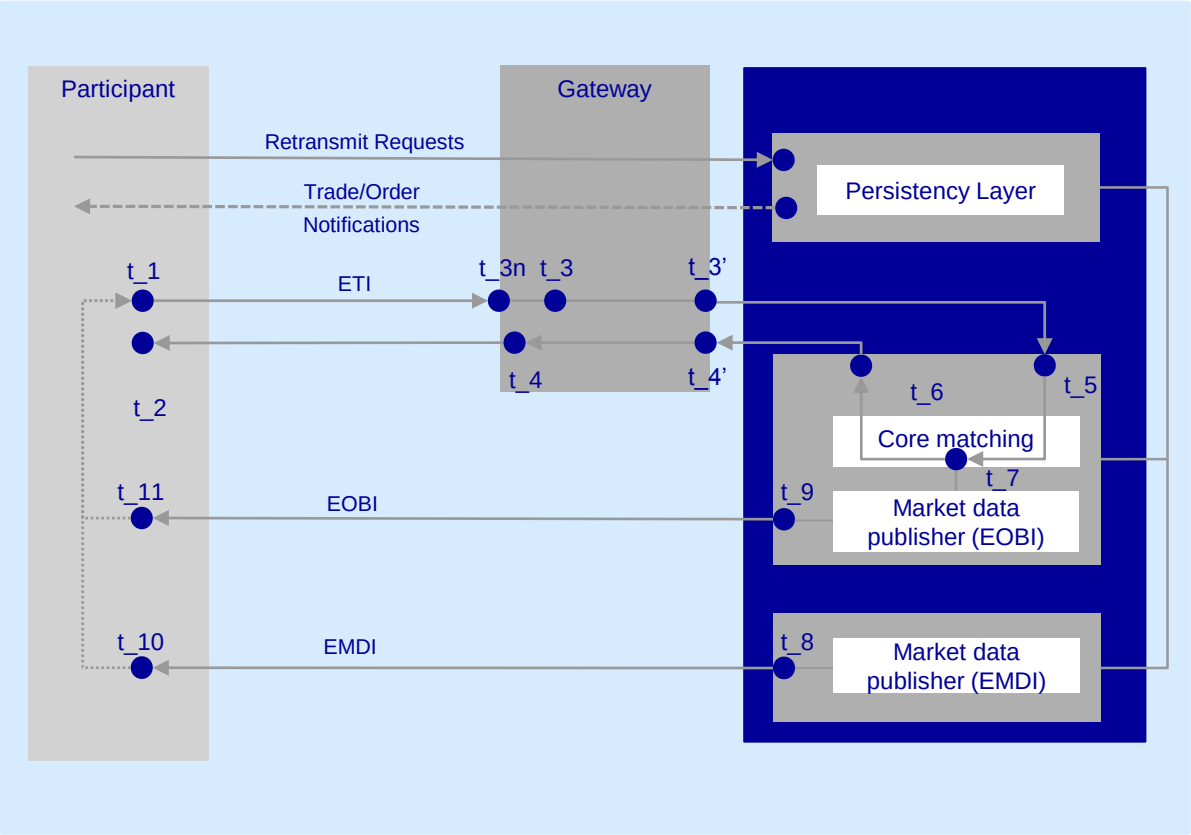
Please refer to the timestamps glossary at the end of this presentation.

Market	Gateway Type	percentile	t3a_t3a'	t3a'_t3d	t3d_t3n	t3a_t3n	t3n_t3'	t3'_t5	t4_t4d	t4d_t4a'	t4a'_t4a	t4_t4a	t3a_t4a
Eurex	PS	50%	0.272	0.652	0.133	1.059	7.865	2.636	1.419	0.678	0.270	2.366	52.671
	LF	50%	0.271			53.267	40.119	9.465				46.138	229.213
	Delta LF - PS	50%	-0.001			52.209	32.255	6.830				43.773	176.542
	Delta LF - PS	50%		|----------	91.3		----------|						
Xetra	PS	50%	0.272	0.653	0.131	1.056	7.457	2.635	1.426	0.800	0.270	2.495	52.301
	LF	50%	0.273			49.449	23.910	19.836				41.202	208.892
	Delta LF - PS	50%	0.001			48.393	16.454	17.202				38.707	156.592
	Delta LF - PS	50%		|----------	82.0		----------|						
Eurex	PS	25-75%	0.004	0.003	0.066	0.070	5.749	0.886	0.335	0.003	0.004	0.335	24.653
	LF	25-75%	0.003			4.677	19.283	3.402				2.377	53.450
Xetra	PS	25-75%	0.004	0.004	0.067	0.067	3.344	1.090	0.344	0.003	0.003	0.344	12.983
	LF	25-75%	0.004			3.111	6.493	10.682				4.340	29.039

27

Inside the matching engine

T7® topology



Matching engine:

- order book maintenance & execution
- creation of direct responses as well as execution messages for passive orders/quotes
- creation of EOBI order book messages
- creation of EOBI order book snapshot messages

Persistency:

- persistent order storage
- trade/execution history
- transaction history for standard orders
- creation of listener broadcast for standard orders

Market Data (EMDI):

- creation of order book delta messages
- creation of order book snapshot messages

PTP based synchronization of clocks using hardware support is used for PS gateways, matching engines and market data servers in production (and also in simulation). Hence timestamps on these servers can be used to analyze one way transport times down to sub microsecond level. For accuracy on nanosecond level we recommend to use the High Precision Timestamp service that makes use of network timestamps synchronized to the 1 ns level (see datashop.deutsche-boerse.com/High-precision-timestamp)

Orders/quotes – detailed performance data

Our transparency

For the top 15 futures products, daily statistics about the matching engine processing times as well as Eurex Enhanced Transaction Interface gateway processing times are provided via the 'Member Section' on Eurex Exchange's website. The ETI round-trip times are calculated based on t_4 – t_3 (gateway SendingTime – gateway application start).

Since introduction of PS gateways reduced multiplicity has resulted in much lower average matching engine processing times. Optimizations introduced with T7 release 6.1 lowered latencies further.
The table below additionally contains latency figures for DAX equities. All data displayed below refer to 19 Jun 2018.

		Matching engine Round-trip times (in µs)			Enhanced Trading Interface Round-trip times (all GWs, t_4 - t_3 in µs)			Enhanced Trading Interface Round-trip times (PS GWs, t_4 - t_3n in µs)		
Product	Product	Average	Median	99th percent	Average	Median	99th percent	Average	Median	99th percent
EURO STOXX 50® Index Futures	FESX	64	26	580	128	56	940	127	56	1000
STOXX® Europe 50 Index Futures	FSTX	36	24	289	112	59	900	121	54	1250
DAX® Futures	FDAX	31	23	175	79	52	358	70	52	398
Mini-DAX® Futures	FDXM	34	25	199	72	52	356	71	54	382
MDAX® Futures	F2MX	33	24	280	103	64	760	100	54	1000
SMI® Futures	FSMI	30	24	152	75	52	275	66	52	288
Euro-Bund Futures	FGBL	66	26	540	135	64	840	139	58	1000
Euro-Bobl Futures	FGBM	50	25	386	114	60	760	110	56	1000
Euro-Schatz Futures	FGBS	36	24	290	105	91	500	90	54	750
Euro-Buxl® Futures	FGBX	40	26	225	93	58	467	91	58	750
Long-Term Euro-BTP Futures	FBTP	33	25	161	92	62	356	78	56	438
Euro-OAT Futures	FOAT	37	24	237	94	56	520	91	54	750
EURO STOXX® Banks Futures	FESB	38	25	292	109	55	900	117	54	1250
VSTOXX® Futures	FVS	34	27	227	91	56	620	81	52	1000
STOXX® Europe 600 Index Futures	FXXP	35	24	291	92	51	860	96	52	1000
DAX® Equities		36	28	180	68	53	271	66	54	288

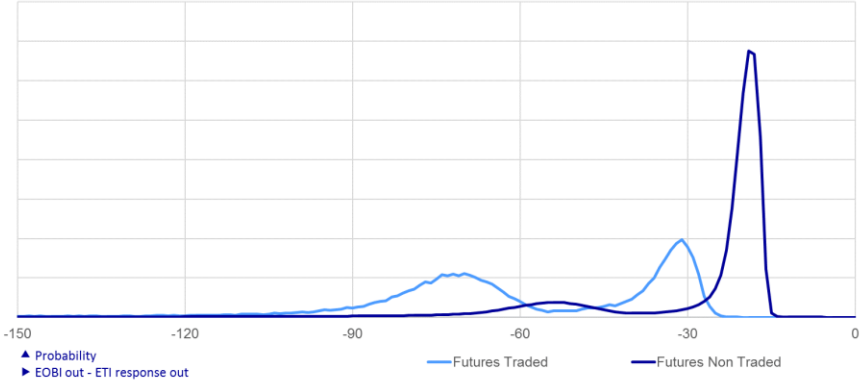
30

Market data

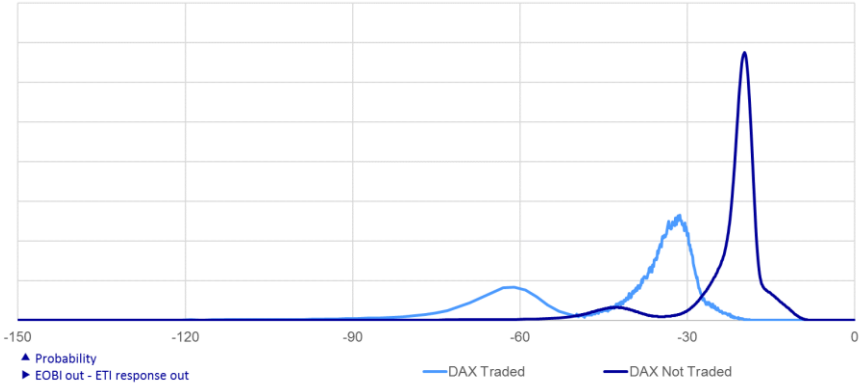
Trading system dynamics

Latency characteristics of EOBI versus ETI

EOBI vs ETI (Eurex FESX Future)



EOBI vs ETI (Xetra DAX equities)



T7 is designed to publish order book updates first on its public data feed.

The top diagram shows the time difference distribution between public and private data for the Eurex FESX future (EOBI first datagram vs ETI responses, t_9-t_4), the graph below shows the same for Xetra DAX equities.

The data is a production sample from 19 June 2018.

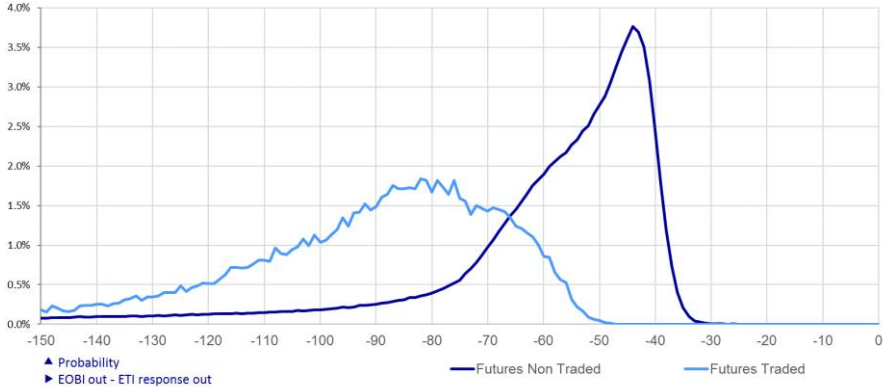
T7 release 6.1 has reduced the median latency of responses via PS gateways by 5 µs, but EOBI market data is still 20 µs faster for order book updates and 35 µs faster for executions. Note that the slower mode of the distribution is caused by responses via low frequency gateways typically being more than 30 µs slower.

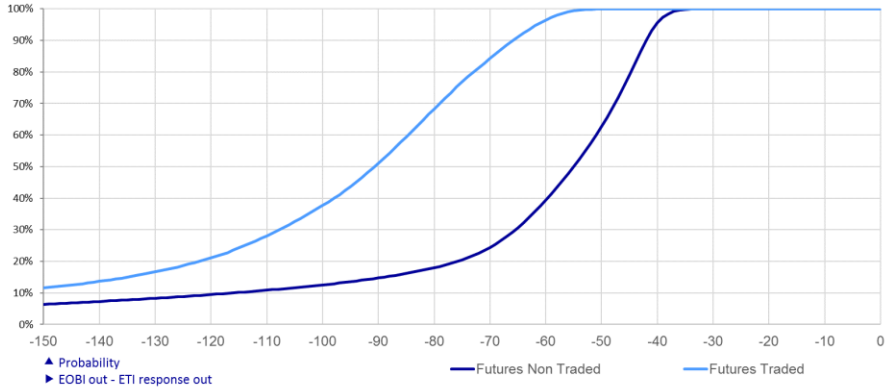
The first EOBI datagram was faster in approximately 99.9 percent of the cases compared to the ETI response and also the first passive ETI book order notification (not shown).

Trading system dynamics

Latency characteristics of EOBI versus EMDI

Latency characteristics of EMDI versus EOBI





Market data updates provided via EOBI are almost always faster than EMDI updates.

The top diagram shows the distribution of t_9 minus t_8, i.e. EOBI first datagram versus EMDI sending time, the bottom diagram shows the cumulative distribution.

The graphs show data of EURO STOXX 50® Index Futures (FESX) for 19 June 2018 (after introduction of T7 release 6.1).

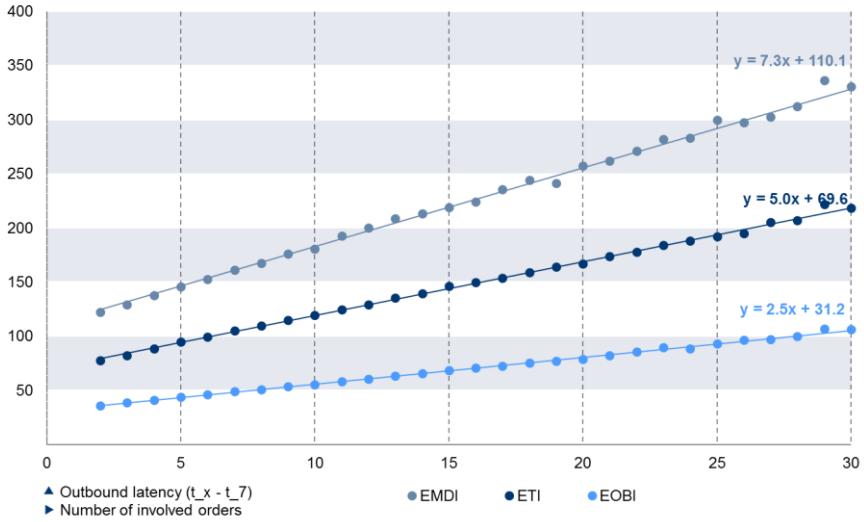
EOBI was faster for more than 99.99 percent of all order book updates.

A very similar latency characteristic applies to Xetra.

Trading system dynamics

Latency characteristics of ETI versus EOBI versus EMDI

Outbound latency for FESX trades (19 Jun 2018)



This diagram displays the dependency of the median latency on the complexity of a trade for ETI (t_4-t_7), EMDI (t_8-t_7) and EOBI (t_9-t_7). Note that for ETI we display the gateway sending time of the first passive notification and for EOBI the sending time of the UDP datagram containing the Execution Summary message.

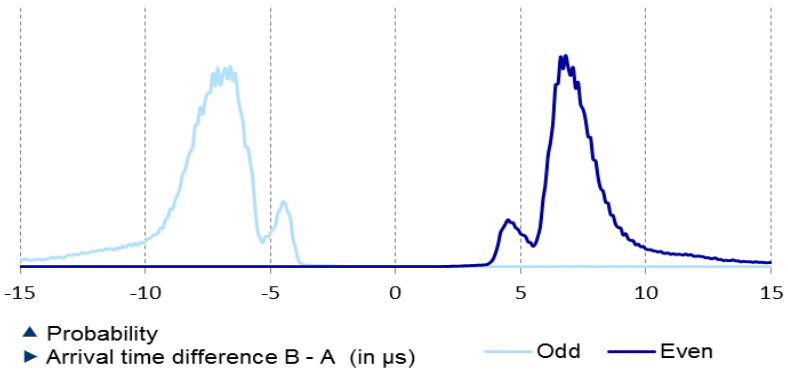
In about 99.95% of all trades, we disseminate order book data on EOBI first (even true for larger trades).

ETI response latency has slightly increased since the migration to PS gateways.

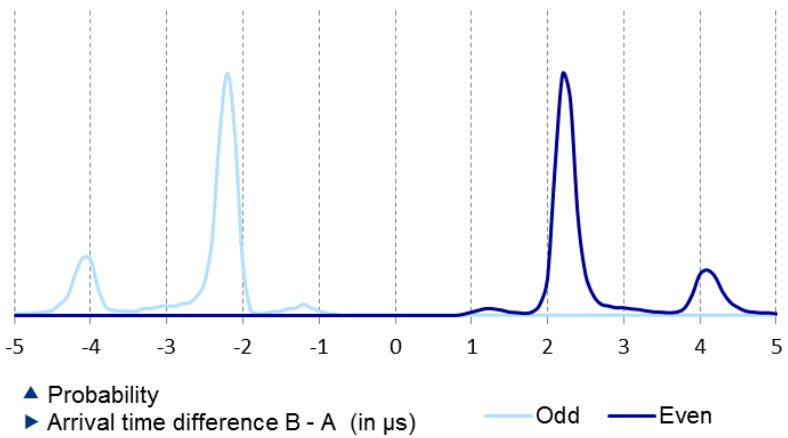
Market data distribution

Latency difference of primary and secondary feed

EMDI



EOBI

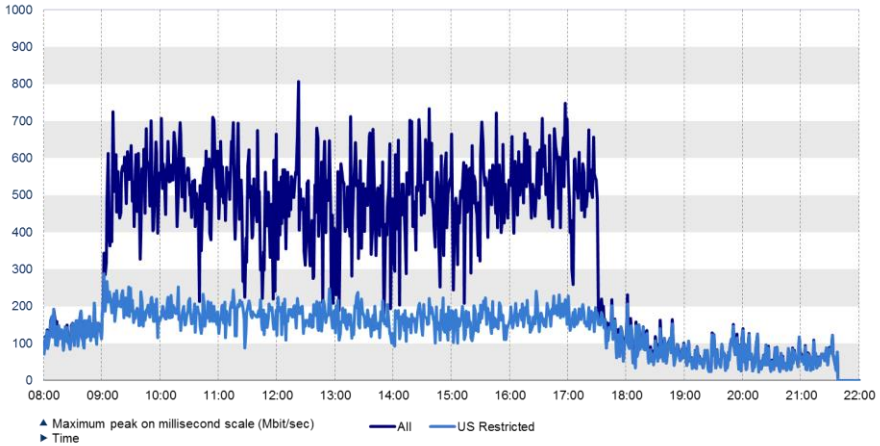


Please note that for products assigned to an even partition, market data is published first on the A stream and then on the B stream whereas, for products assigned to an odd partition market data is published first on the B stream and then on the A stream.
The partition ID / product ID is contained in the UDP datagram header of the order book incremental messages and can be used for filtering on UDP datagram level for EMDI / EOBI.
Furthermore, a UDP datagram on the T7 EMDI / EOBI order book delta or snapshot channel will only contain data of exactly one product (e.g. EURO STOXX 50® Index Futures).

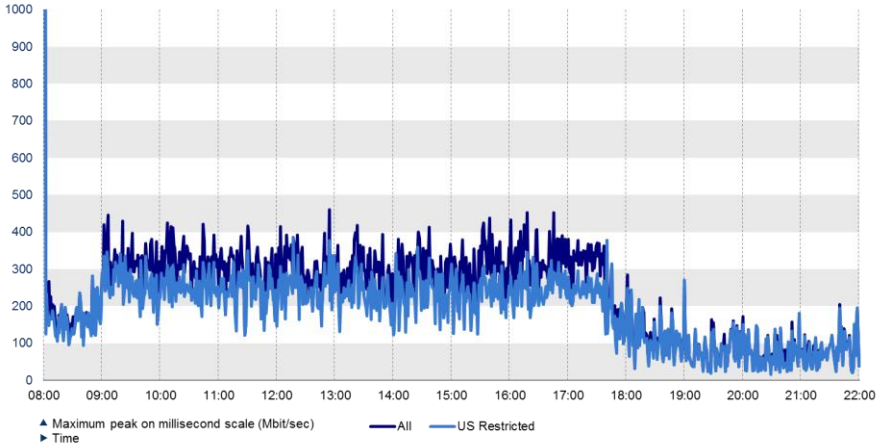
The median latency difference between the A and the B incremental feed is about 8.4 µs for EMDI (see top diagram to the left) and 2.3 µs for EOBI (see bottom diagram to the left).
T7 provides a csv file on a daily basis with the minute-by-minute network latency (minimum, average, maximum 99 per cent) for the A and B streams of EMDI for non co-location access points. This information can help you determine whether your environment or T7 had an issue causing a market data delay. The file is provided in the member portal (member.deutsche-boerse.com).

Eurex: Market data volume

Eurex EMDI



Eurex EOBI



Each data point equals the maximum bandwidth produced on a 1 millisecond scale by the incremental B stream in Mbps.

The provided data shows one data point per minute for 19 June 2018.

Eurex market data peak data volume can be significantly higher on high volume trading days. Hence Participants that want to receive data for all Eurex Exchange's products or U.S. only products with less than 1 ms queuing delays need to use a connection with a bandwidth of more than 1,000 Mbps (all products) or 300 Mbps (for U.S. only products), respectively.

Enhanced Order Book Interface market data is currently only available to Trading Participants using 10 GbE connections.

Trading Participants are advised to use two 10 GbE connections (one for each market data stream) in co-location to receive market data.

Xetra: Market data volume

Each data point equals the maximum bandwidth produced on a 1 millisecond scale by the incremental B stream in Mbps.

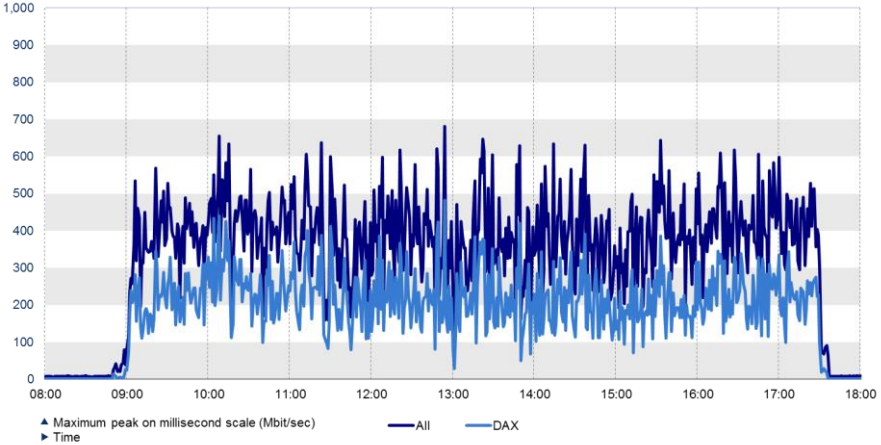
The provided data shows one data point per minute for 19 Jun 2018.

Xetra Market Data peak data volume can be a significantly higher on high volume trading days. Hence Participants that want to receive data with less than 1 ms queuing delays need to use a connection with a bandwidth of at least 700 Mbps (EMDI, All products) or 500 Mbps (EMDI, DAX® equities only).

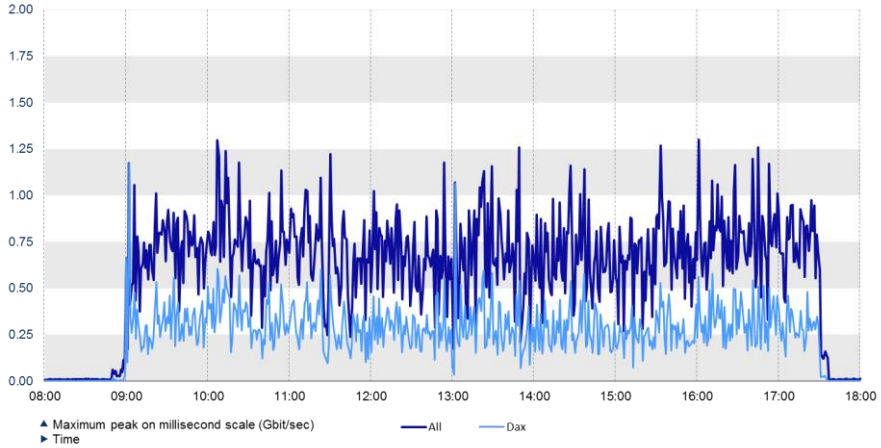
Enhanced Order Book Interface market data is currently only available to Trading Participants using 10 GbE connections.

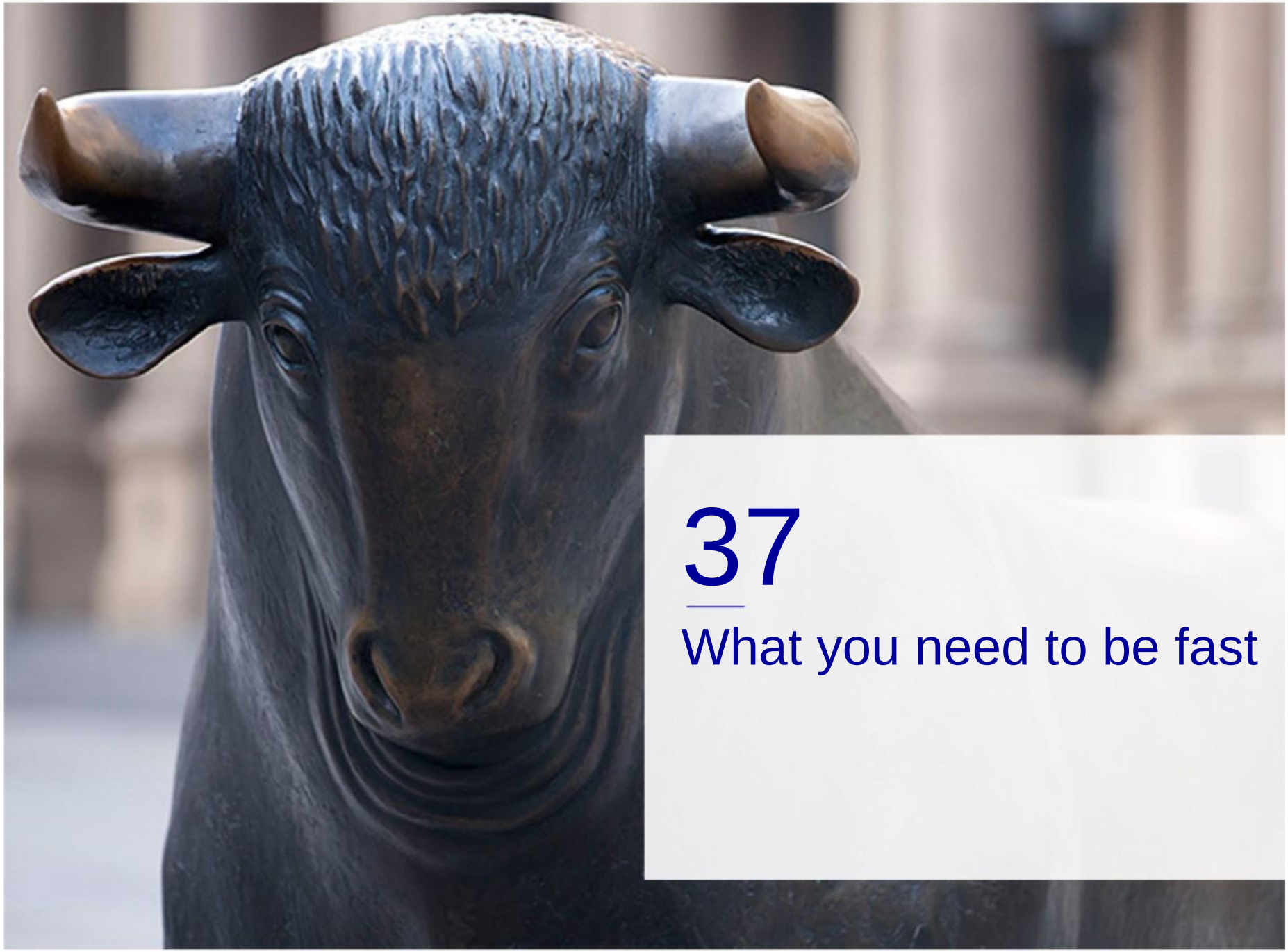
Trading Participant are advised to take two cross connects (one for each market data stream) in co-location to receive market data.

Xetra EMDI



Xetra EOBI





37

What you need to be fast

What you need to be fast…

A few recommendations to achieve the lowest possible latency:

Use the Equinix co-location facility to be close to Deutsche Börse T7.

Use state-of-the-art switches (if any) and only have at most one hop between the exchange network and your server.

Use good network interface cards and TCP/IP acceleration, e.g. a Linux kernel-by-pass library.

Use at least two 10 GbE (cross-) connections in co-location 2.0 for EMDI or EOBI market data and two 10 GbE (cross-) connections for T7 ETI.

Use PS gateways and make sure you use the cross connect on the same side as the gateway you are connecting to (compare time-to-live values in the IP header in the responses from both sides).

Measure and analyze your own timestamps (e.g. the reaction time as recommended on the next slide).

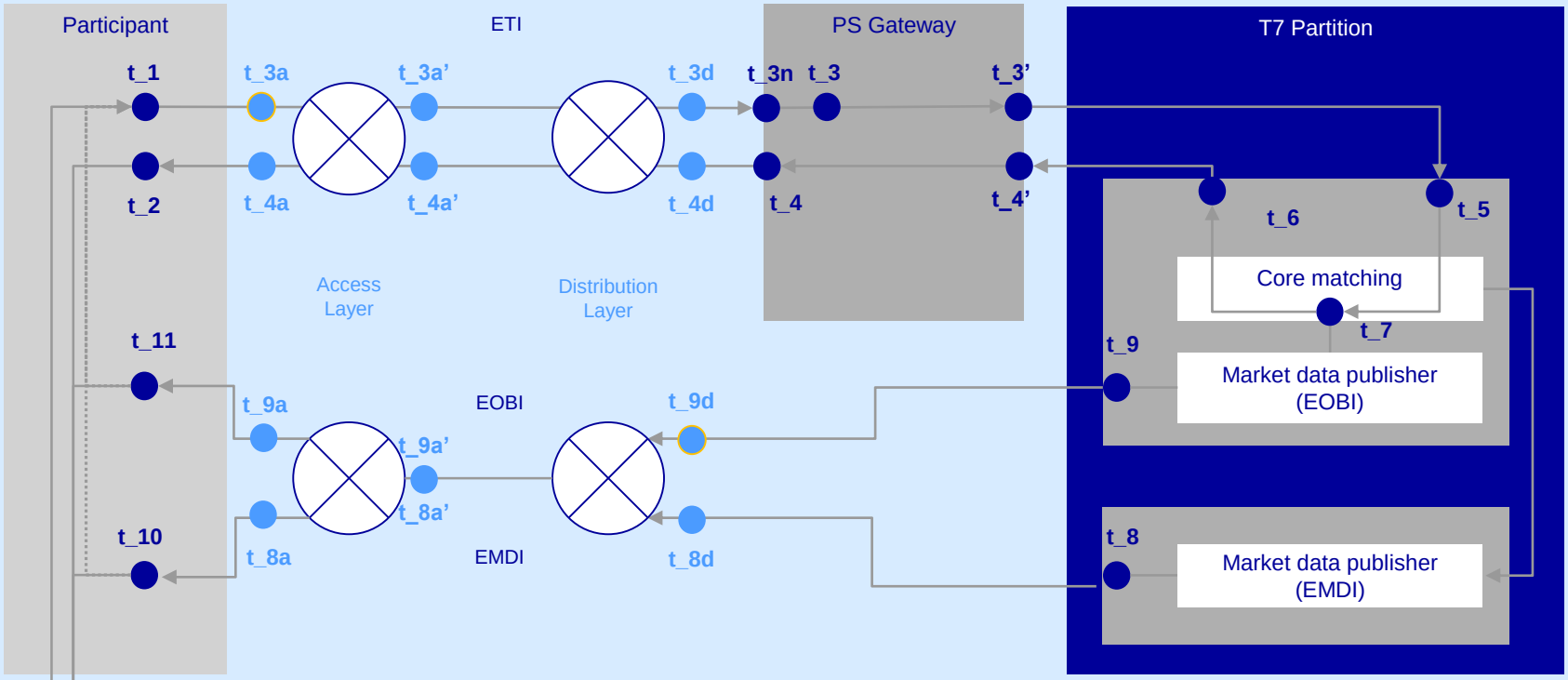
Use state of the art time synchronization, i.e. GPS clocks and a high quality time distribution. The PTP signal you can get from us has a quality of +/- 50 ns. For our network timestamps we use the White Rabbit protocol and PPS breakouts. Note that we aim to launch a White Rabbit service in Q3 2018. We provide highly accurate network timestamps for all executions via the high precision timestamp file service, see https://datashop.deutsche-boerse.com/High-precision-timestamp.

Trade notifications need to be processed to create safety (only the trade notifications contain legally binding information about a trade!). Therefore, we recommend to use either a low-frequency ETI session or a FIX trade capture drop copy to confirm the fast execution information provided by the execution reports via high-frequency sessions.

Try to use the EOBI Execution Summary for fast trading decisions and position keeping (passive executions).
For a consistent order book, all incremental updates following the Execution Summary should always be processed.

What you need to be fast…

Participant reaction time measurement



Measure the time between market data reception (t_10/t_11) and your reaction (t_1), take note of aggressor in timestamps (t_3n) of executions.

t_3a and t_9d are available via the high precision timestamp file service, see https://datashop.deutsche-boerse.com/High-precision-timestamp

41

Appendix

Enhanced Order Book Interface

Exceptions to fast trading decisions based on the Execution Summary

The Execution Summary at the beginning of an EOBI packet can be used for fast trading decisions (and passive position keeping) in the majority of cases. However there are certain scenarios where this may lead to a wrong perception of the order book on client side. This includes cases where:

The Request-In timestamp is not set:
This is for example in case of more than one market order being triggered by an incoming order/quote. In such a scenario, there will be one Execution Summary sent for each market order. The Execution Summary for the first market order will be at the beginning of the packet but the Execution Summaries for the other market orders will follow in the same or in the next packet(s) before CompletionIndicator is set to 1.

The Implied flag is set:
In case of synthetic matching, the LastQty shows the total matched quantity that result from all involved instruments‘ order books. At the same time only the instrument of the aggressing order is given. In order to have correct order books, Participants have to process all incremental updates following the Execution Summary. There is a potential shortcut in this case which is explained on the next two slides.

In equity markets the execution of hidden quantity of Iceberg orders is not reflected in the hiddenQty field of the EOBI Execution summary.

Enhanced Order Book Interface

Fast trading decision in case of synthetic matching (derivatives markets only)

In case of synthetic matching (TradeCondition in the Execution Summary set to ‚Implied Trade‘), Participants only interested in the front month instrument will also have to read all details that follow the Execution Summary.

However, we see that not each combination of instruments is traded with the same frequency (even during a roll when, on average, more synthetic matches take place). In our benchmark futures, we see that there are three highly liquid and three less liquid instruments. The front month, the second maturity and the spread between the two are defined as the liquid instruments, the third maturity and combinations thereof are considered as less liquid.

A possible shortcut which can speed up the decision-making process is available to Participants that base their trading decisions primarily on the Execution Summary. These Participants could potentially save processing time by considering the order books of the three most liquid instruments only.

Enhanced Order Book Interface

Fast trading decision in case of synthetic matching (cont.)

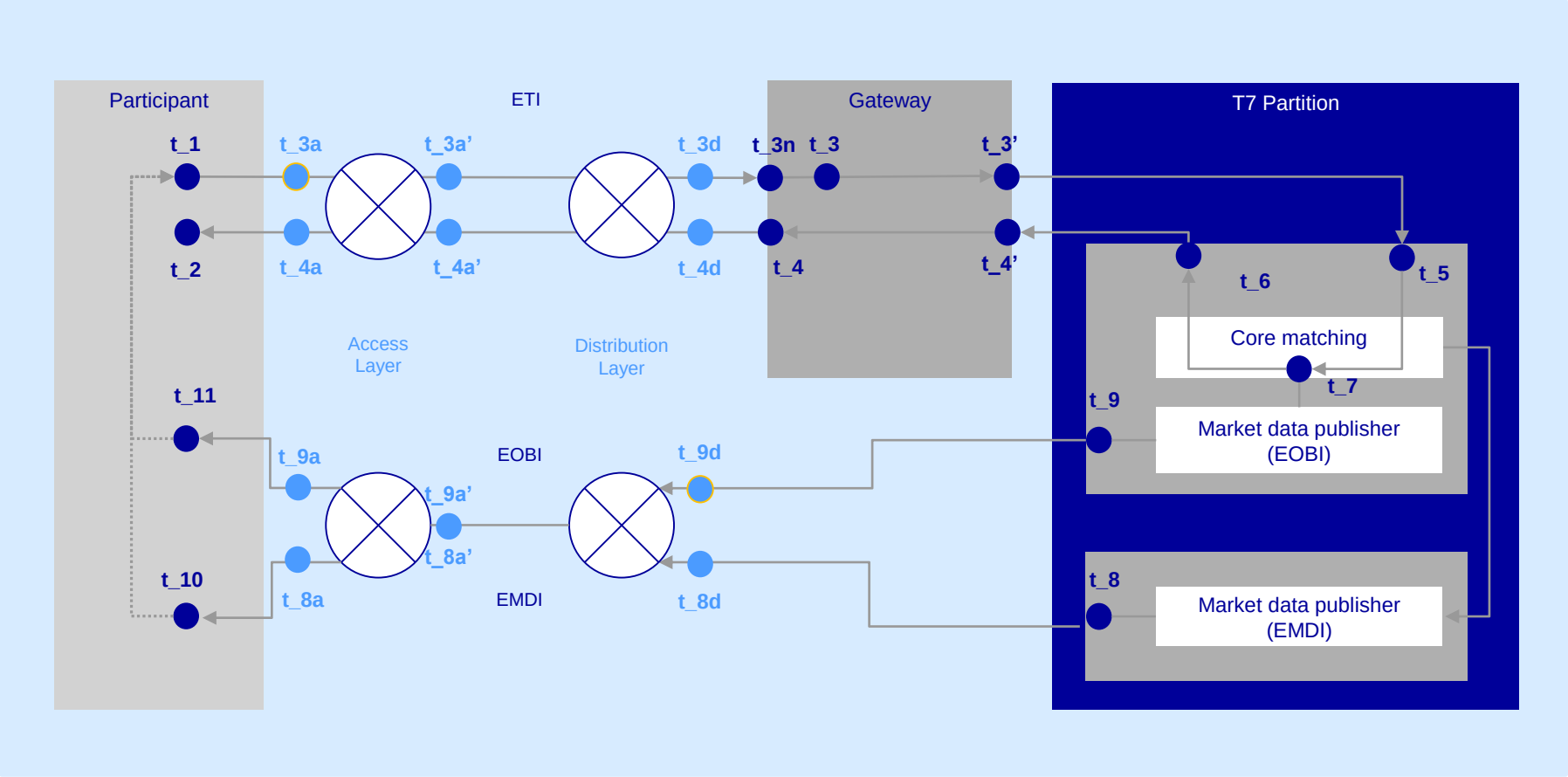
If the instrument in the Execution Summary belongs to one of the three liquid instruments and the TradeCondition field is set to ‚Implied Trade‘, Participants can apply the following logic to infer the order books for these liquid instruments:

- Check price (‚LastPx‘), total quantity (‚LastQty‘) and side (‚AggressorSide‘) in the Execution Summary.
- Delete orders in the incoming instrument‘s order book which have a better price than ‚LastPx‘.
- Check whether a combination of orders in the two other liquid books yields a better price than ‚LastPx‘. If true, delete the orders contributing to these combinations from the respective order books.
- Calculate the remaining quantity (‚LastQty‘ minus already deleted quantities). If possible, remove this quantity from the incoming instrument‘s order book at ‚LastPx‘. If there is not enough quantity available then try to delete the still remaining quantity from the synthetic order book combination at ‚LastPx‘.
- In case the total deleted quantities from the incoming instrument‘s order book and the combined synthetic order book do not add up to the ‚LastQty‘ it can be concluded that the order book is not correct afterwards.

Our data show that in over 99.7% of the trade cases, it is possible to build a correct order book by using this shortcut.

Please note that for a consistent order book, Participants should always process all incremental updates following the Execution Summary.

T7® timestamps



Network time-stamps shown in light blue, t_3a and t_9d are available via the high precision timestamp file service, see https://datashop.deutsche-boerse.com/High-precision-timestamp

Description of timestamps

Definition

t_1,t_2: can be taken by a Participant (e.g. via a network capture) when a request/ response is read from/written to the socket.

t_4: taken by the ETI gateway when a response/ notification is written to the socket on the Participant´s side of the gateway; contained in (private) ETI response/ notification.

t_3n: taken by the ETI gateway when the first bit of a request arrives on the PS gateway NIC; contained in (private) ETI response for transactions sent via PS gateways. Consecutive messages via the same session might be assigned to the same t_3n in rare cases.

t_3: taken by the ETI gateway application when a request is read from the socket on the Participant´s side of the gateway; contained in (private) ETI response for transactions sent via LF gateways.

t_3': taken by the ETI gateway right before a request is sent towards the matching engine; contained in (private) ETI response.

t_4': taken by the ETI gateway when a response/ notification is received by the ETI gateway from the matching engine; contained in (private) ETI response/ notification.

t_5, t_6: taken by the matching engine when a request/response is read/written; contained in (private) ETI response.

t_7: time at which the matching engine maintains the order book

t_8: time taken by EMDI publisher just before the first respective UDP datagram is written to the UDP socket.

t_9: time taken by EOBI publisher just before the first respective UDP datagram is written to the UDP socket.

t_10, t_11: can be taken by a Participant (e.g. via a network capture) when a UDP datagram is read from the UDP socket.

t_?a, t_?a', t_?d: taken by network capture devices in the access and distribution layers

T7 timestamp reference

The timestamps t_3,…,t_9 are available via the following fields:

t_3, t_3n:	Tag	5979	(“RequestTime”)	in the T7 ETI Response in the T7 EMDI Depth Incremental message, in case a trade is reported in the T7 EOBI Execution Summary message
t_3’:	Tag	7764	(“RequestOut”)	in the T7 ETI Response (from the matching engine)
t_4’:	Tag Tag	7765 25043	(“ResponseIn”) (“NotificationIn”)	in the T7 ETI Response (from the matching engine) in the T7 ETI Notification (from the matching engine)
t_4:	Tag	52	(“SendingTime”)	in the T7 ETI Response and Notification
t_5:	Tag Tag Tag	21002 21002 28820	(“TrdRegTSTimeIn“) (“TrdRegTSTimeIn”) (“AggressorTimestamp“)	in the T7 ETI Response (from the matching engine) in the T7 EOBI Order Add, Order Modify, Order Modify Same Priority and Order Delete messages in the T7 EMDI Depth Incremental message, in case a trade is reported in the T7 EOBI Execution Summary message
t_6:	Tag	21003	(“TrdRegTSTimeOut“)	in the T7 ETI Response and Notification (from the matching engine)
t_7:	Tag Tag Tag Tag	17 273 21008 60	(“ExecID“) (“MDEntryTime”) (“TrdRegTSTimePriority”) (“TransactTime”)	in the T7 ETI Response (from the matching engine) in the T7 EOBI Execution Summary message in the T7 EMDI Depth Incremental message in the T7 EOBI Order Add and Order Modify messages in the T7 EOBI Order Modify Same Priority and Order Delete messages
t_8:	no Tag		(“SendingTime”)	in the T7 EMDI UDP packet header
t_9:	Tag	60	(“TransactTime”)	in the T7 EOBI packet header
(t_8-t_5):	no Tag		(“PerformanceIndicator“)	in the T7 EMDI UDP packet header of the T7 EMDI Depth Incremental stream

Notes on timestamps:
All timestamps provided are 8 byte integers (in nanoseconds after Unix epoch).
The PerformanceIndicator is a 4 byte integer (in nanoseconds).
The Network timestamps (t_?a, t_?a’, and t_?d are not available in any field.

Thank you for your attention

Contact

Sebastian Neusüß, Andreas Lohr & Manfred Sand

E-Mail monitoring@deutsche-boerse.com

For updates refer to http://www.eurexchange.com/exchange-en/technology/high-frequency_trading
http://www.xetra.com/xetra-en/technology/t7/publications