DEUTSCHE BÖRSE GROUP

Xetra®

Insights into Trading System Dynamics

April 2016

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

Xetra technology roadmap

Xetra is very committed to following its technology roadmap to deliver innovative and superior technology and is constantly improving its technical offering.

In the course of end of Q2 to Q3 2015, Xetra® servers were upgraded to next generation Poulson Itanium processors.

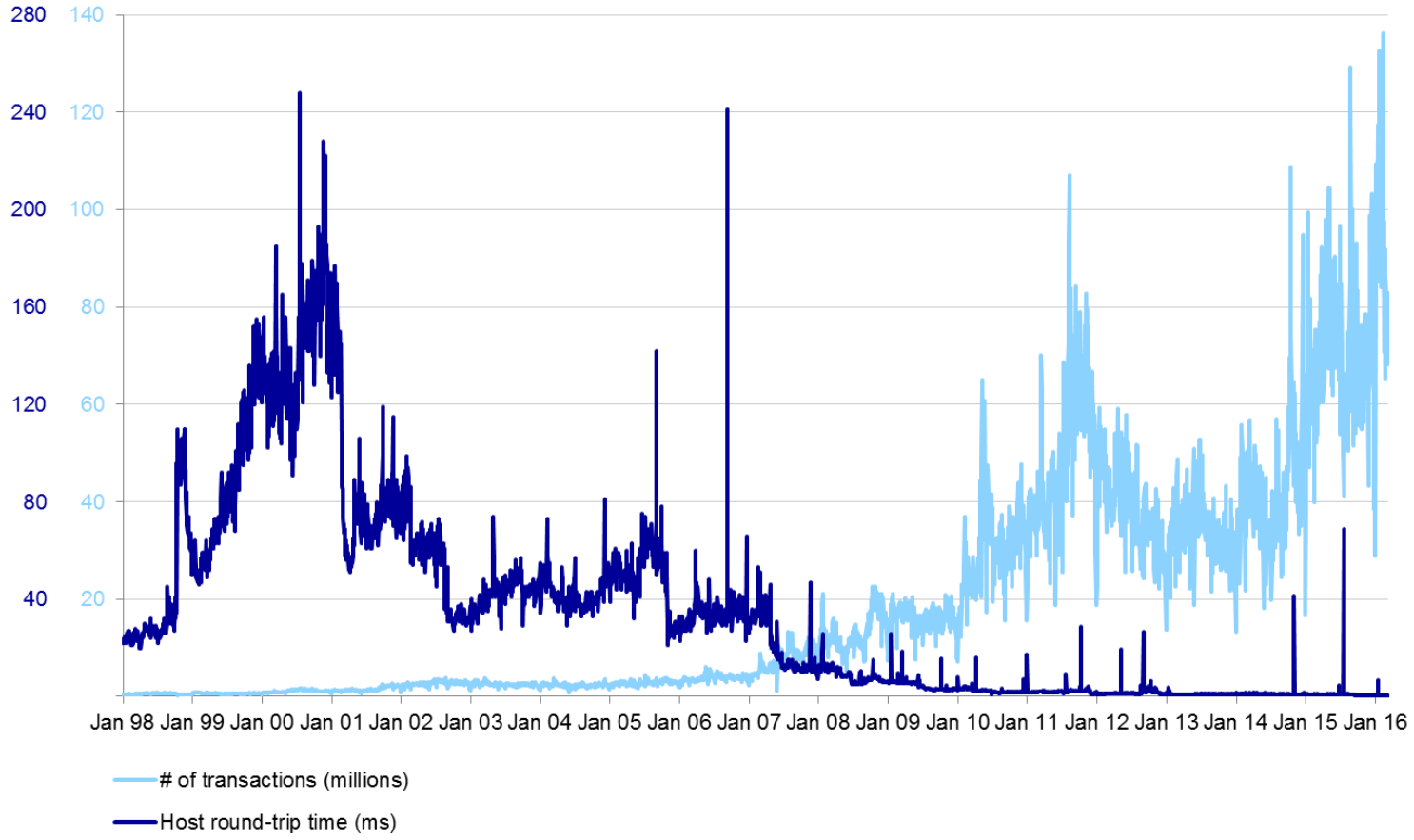
In addition, a configuration change on all the Xetra ETS Gateways has improved the performance.

As a result, median round trip times and in particular outliers (99 percentile) were significantly reduced (slide 5).

Deutsche Börse plans to migrate cash market asset classes onto its T7 system as part of its aligned technology/ business roadmaps.

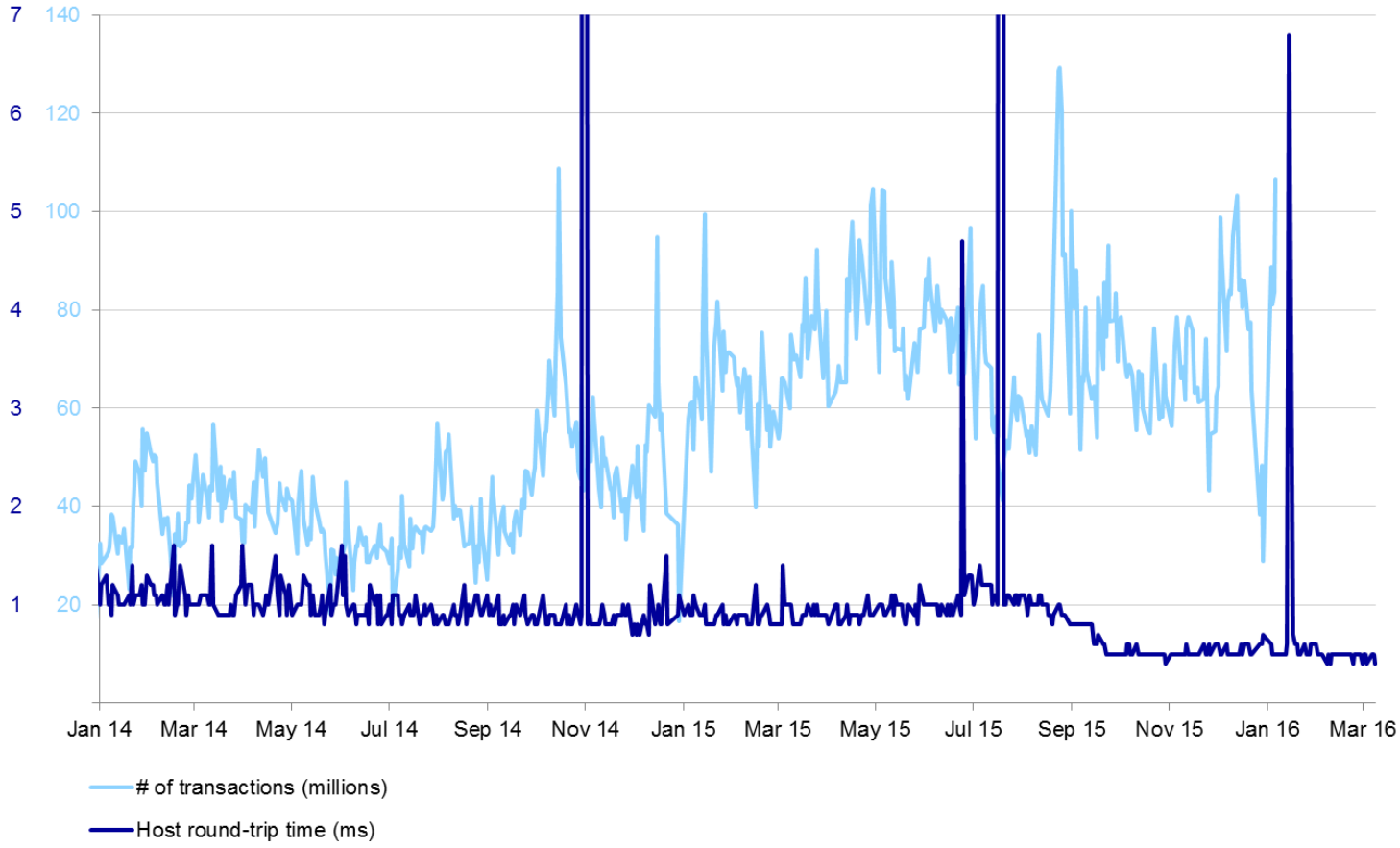
The cash market migration onto the T7 system will be guided by the existing T7 interface landscape and functionality.

Long-term processed transactions and response times



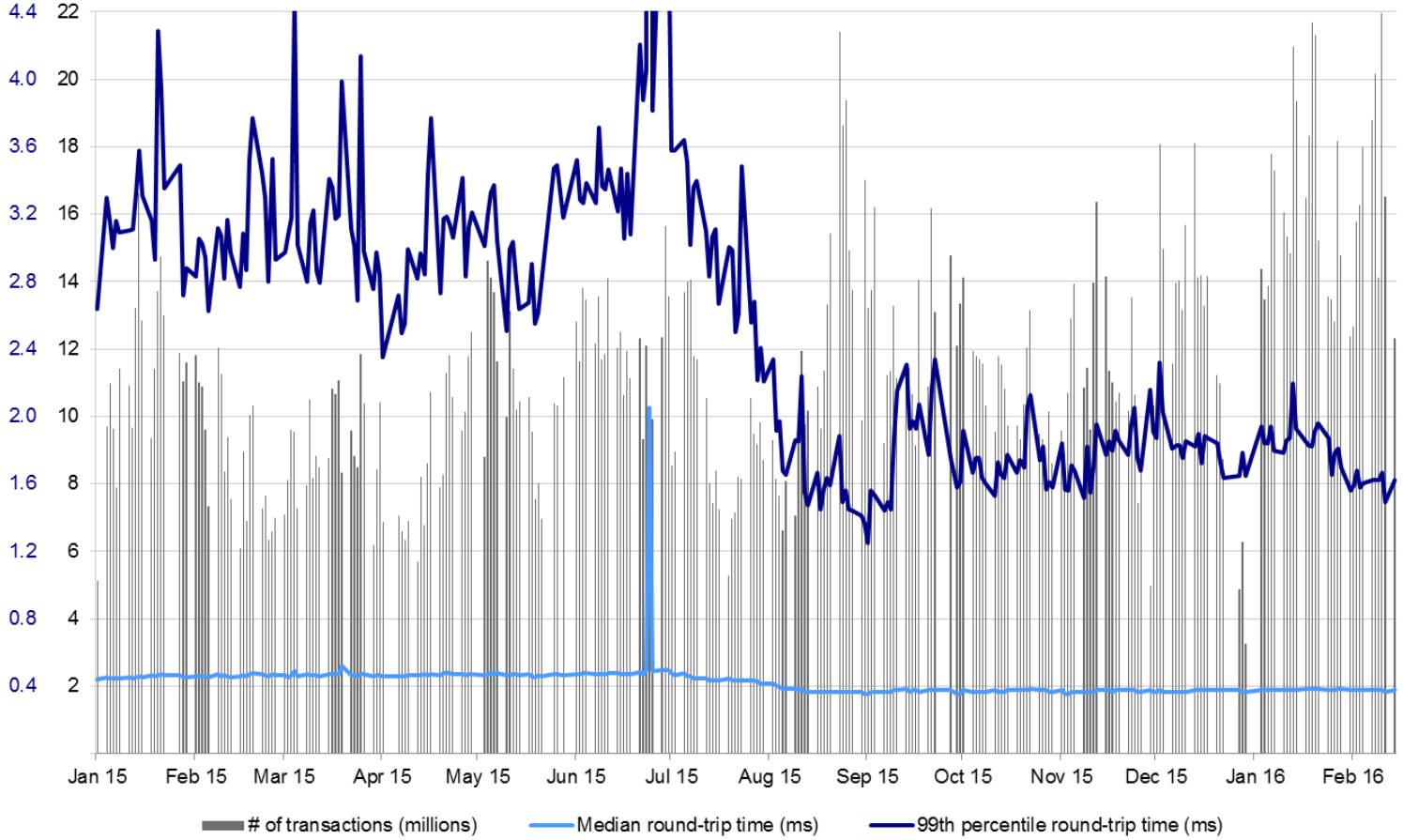
Xetra® has continuously invested in its trading system and has been able to reduce the processing time of technical transactions significantly although the daily load on the system has grown significantly.

Recent processed transactions and response times



Over 99 percent of order transactions are non-persistent and entered via the Enhanced Transaction Solution interface. The host latency for all transactions has been significantly reduced to a level of a daily average of about 500µs.

Round-trip times DAX non-persistent orders in Xetra®



During the past months, we have optimized the Xetra system by replacing all fourteeen Xetra hosts with improved hardware. At the same time, we changed the configuration on all Xetra ETS Gateways to lower the processing times. The positive effects can be observed for example in the significantly reduced mean and especially the 99th percentile of the (gateway-to-gateway) round-trip times for DAX non-persistent orders.

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

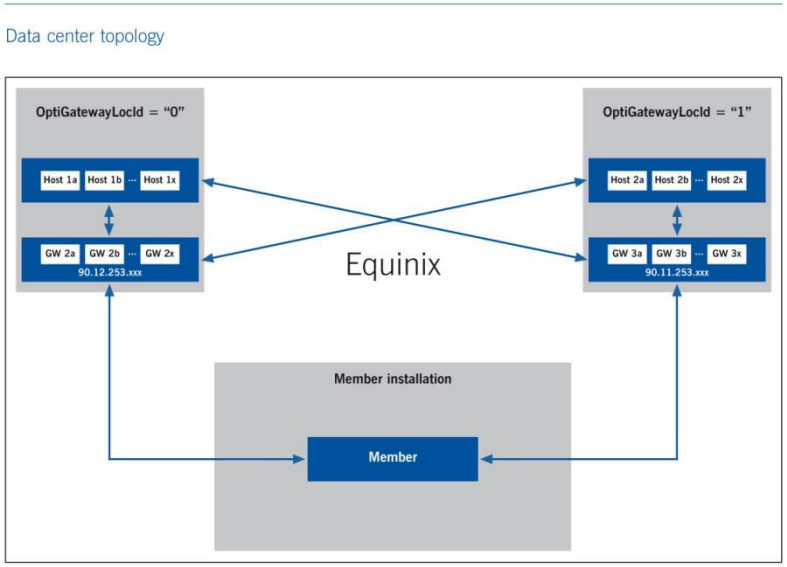
8 Appendix

Enhanced Transaction Solution – Overview

The **fourteen** Xetra® matching engines (hosts) are equally distributed over two rooms in the Equinix data centre. Six of them are dedicated to DAX® instruments, one to the top 10 MDAX instruments, the other eight are used to match all remaining Xetra instruments.

The Enhanced Transaction Solution gateways in the Equinix facility provide the fastest access to the Xetra trading system. There are **sixteen** such gateways in the Equinix data centre shared by all Xetra trading participants.

Xetra's Enhanced Transaction Solution reference data contains the OptiGatewayLocID, providing an exceptional level of transparency and customer service (the OptiGatewayLocID is relevant for choosing the right gateways and for listening to the correct market data stream in order to minimize latency, also see pages 8 and 11).



Currently, we have the following mapping for DAX® 30 instruments (this is not static):

OptiGatewayLocID = "0":

ADS, BAYN, CON, DAI, DBK, FME, IFX, LHA, MRK, MUV2, RWE, SAP, PSM*, SIE, VNA

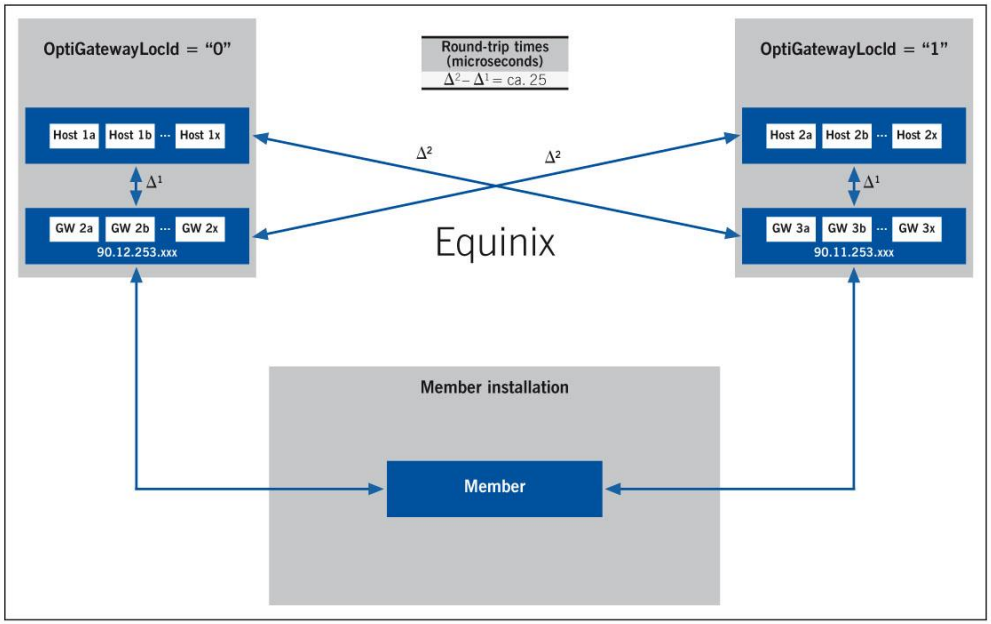
OptiGatewayLocID = "1":

ALV, BAS, BEI, BMW, CBK, DB1, DPW, DTE, EOAN, FRE, HEI, HEN3, LIN, TKA, VOW3

*) PSM has replaced SDF since March 21st.

Enhanced Transaction Solution – Latency

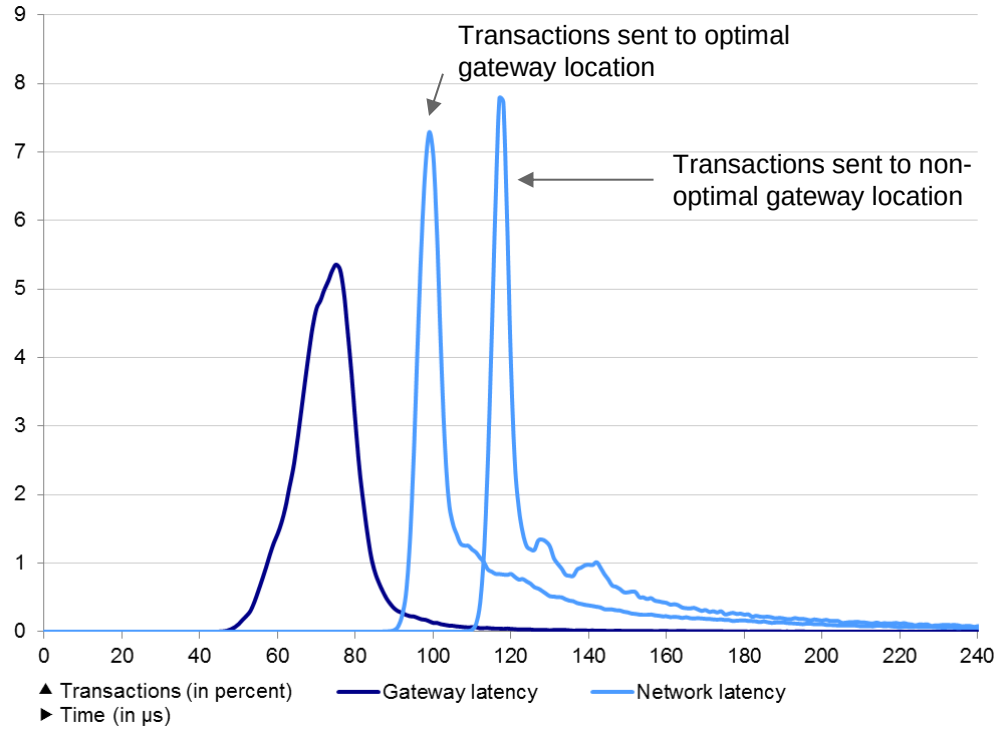
Enhanced Transaction Solution – latency



- Fastest access to Xetra® is provided via the Enhanced Transaction Solution gateways in Equinix (see IP subnets in the diagram).
- Participants can use the reference data to find out in which room an instrument trades. Routing a transaction to a gateway located in a room where the instrument does not trade increases the in-bound latency by 12.5 µs on average (see also next slide).

- Daily statistics about private "last mile" performance between the gateways and participant servers as well as best in class numbers (per location and system wide) are provided.
- Please note that the overall latency advantage of the 10 GbE connections compared to 1 GbE connections is approximately 50 µs round-trip time for order transactions.
- Xetra expects that a good daily average TCP/IP round-trip will be less than 8 µs for 10 GbE connections.

Enhanced Transaction Solution – Processing times



- Both graphs are histograms of February 16, 2016 for all order / quote transactions entered for instrument DBK via gateways located in Equinix.
- The light blue lines represent the transport time between the gateways and matching engines. According to the timestamps that are provided by Xetra®, this is (t_4' – t_3') – (t_6 – t_5).[1)]
- The dark blue line reflects the processing time on the gateways which is defined by (t_3' – t_3) + (t_4 – t_4').[1)]

The diagram shows that the bulk of the round-trip time that is not spent on the matching engine is actually spent on the transport of messages from the gateways to the matching engine and back. Furthermore, one can clearly see the impact of transactions routed to non-optimal gateways on the round-trip times between the Xetra gateways in Equinix and our matching engines.

1) The definitions of all Xetra timestamps can be found in the special section on timestamps on page 23 and in the appendix (p. 30).

Throttle limits

In order to protect its trading system, Xetra has several measures in place to ensure that its most vital components are not harmed by a malfunctioning client application. In particular Xetra uses the following transaction limits (throttle):

Xetra ETS sessions are available with throttle values of 150 txns/sec and 50 txns/sec (ETS Light sessions), respectively.

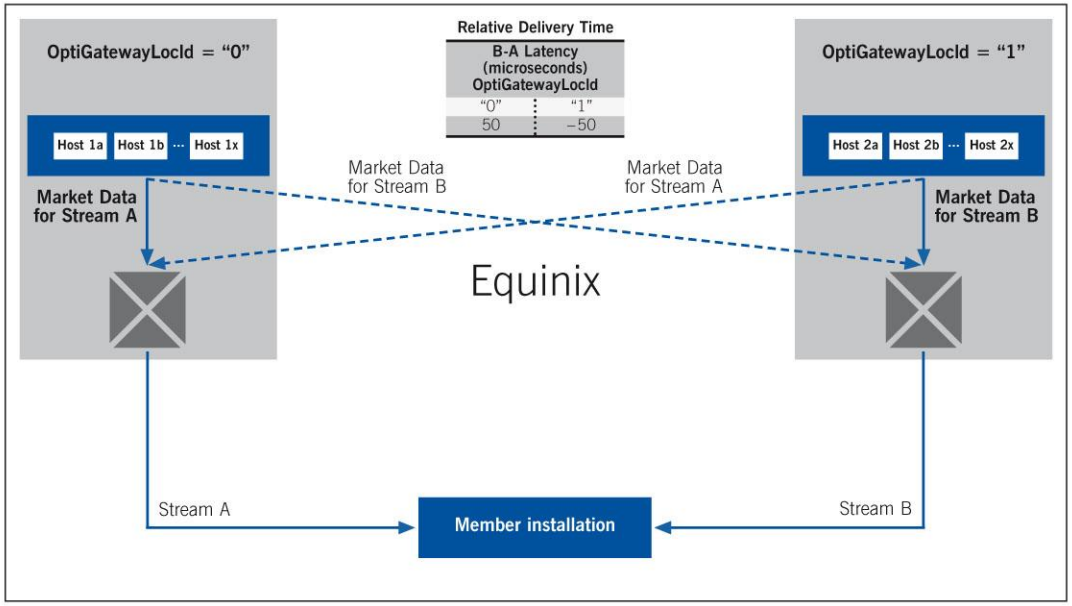
All Xetra ETS sessions have an assigned disconnect limit of

- 450 for sessions with a throttle value of 150 txn/sec, i.e. a session will be disconnected if it sends more than 450 transactions within a second (resulting in more than 300 consecutive rejects due to exceeding the throttle limit).
- 150 for sessions with a throttle value of 50 txn/sec, i.e. a session will be disconnected if it sends more than 150 transactions within a second (resulting in more than 100 consecutive rejects due to exceeding the throttle limit).

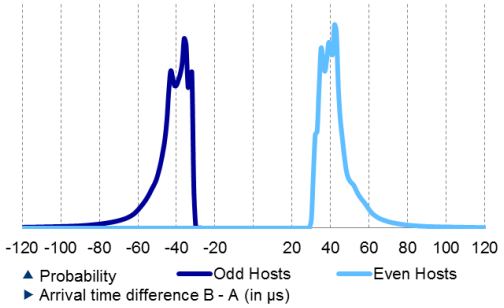
For both limits all technical transactions are counted using a sliding window.

Enhanced Broadcast Solution – Topology

Enhanced Broadcast Solution – topology

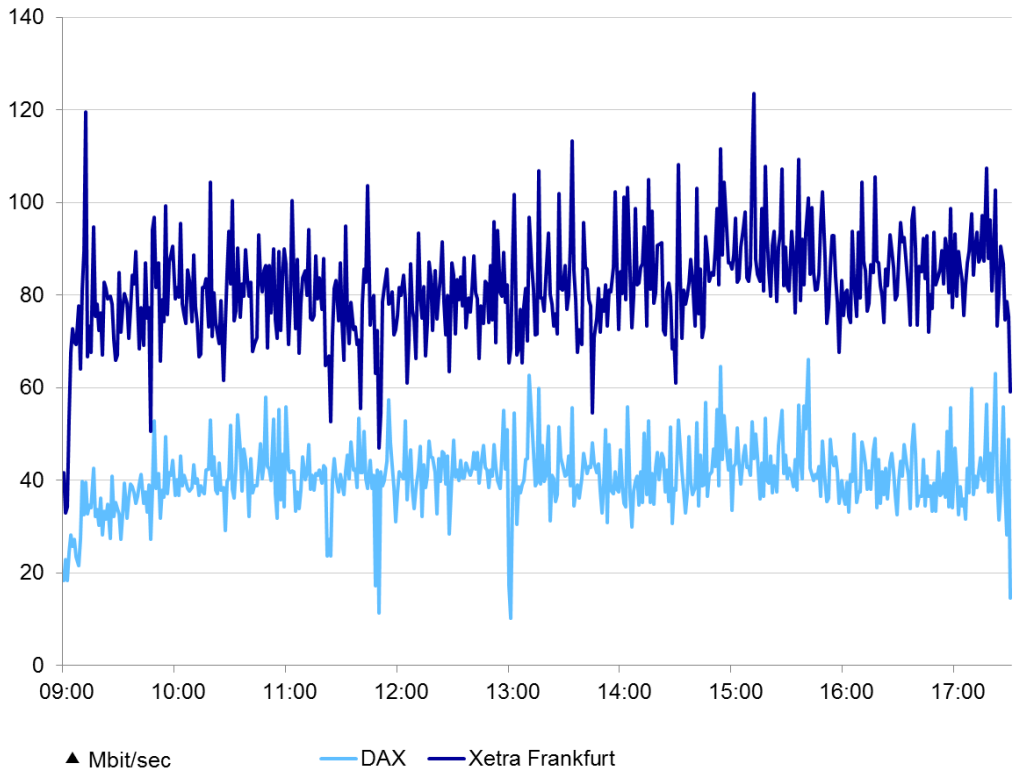


- Please note that market data for instruments with OptiGatewayLocID = "0" is published first on the A stream and then on the B stream whereas for instruments with OptiGatewayLocID = "1" market data is published first on the B stream and then on the A stream.
- Products with the same "SrcID" are traded on the same physical host.
- The latency advantage of a 10 GbE connection compared to a 1 GbE connection is approximately 60 µs for order book incremental messages.



- The average latency difference between the A and the B feed is about 45 µs (see diagram to the left).
- Xetra® provides a csv file on a daily basis with minute-by-minute network latency (minimum, average, maximum, 99 percent, 99.5 percent) for the A and B stream of the Enhanced Broadcast Solution. In case of issues with market data latency, this information can help determining whether you had an issue or Xetra did.

Enhanced Broadcast Solution – Data volume



- The provided Enhanced Broadcast Solution data of Xetra® shows one data point per minute interval for March 10, 2016 for all Xetra instruments and for DAX instruments only.
- Each data point equals the maximum bandwidth in use on a 1 ms scale by the incremental B stream in Mbit/sec.
- As the sample day did not show the highest market activity, microbursts on peak trading days may be higher. Hence, participants who want to receive data for DAX instruments with less than 1 ms queuing delays need to use a connection with a bandwidth of more than **50 Mbit/sec**. For all Xetra instruments, a connection with a higher bandwidth would be required.

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

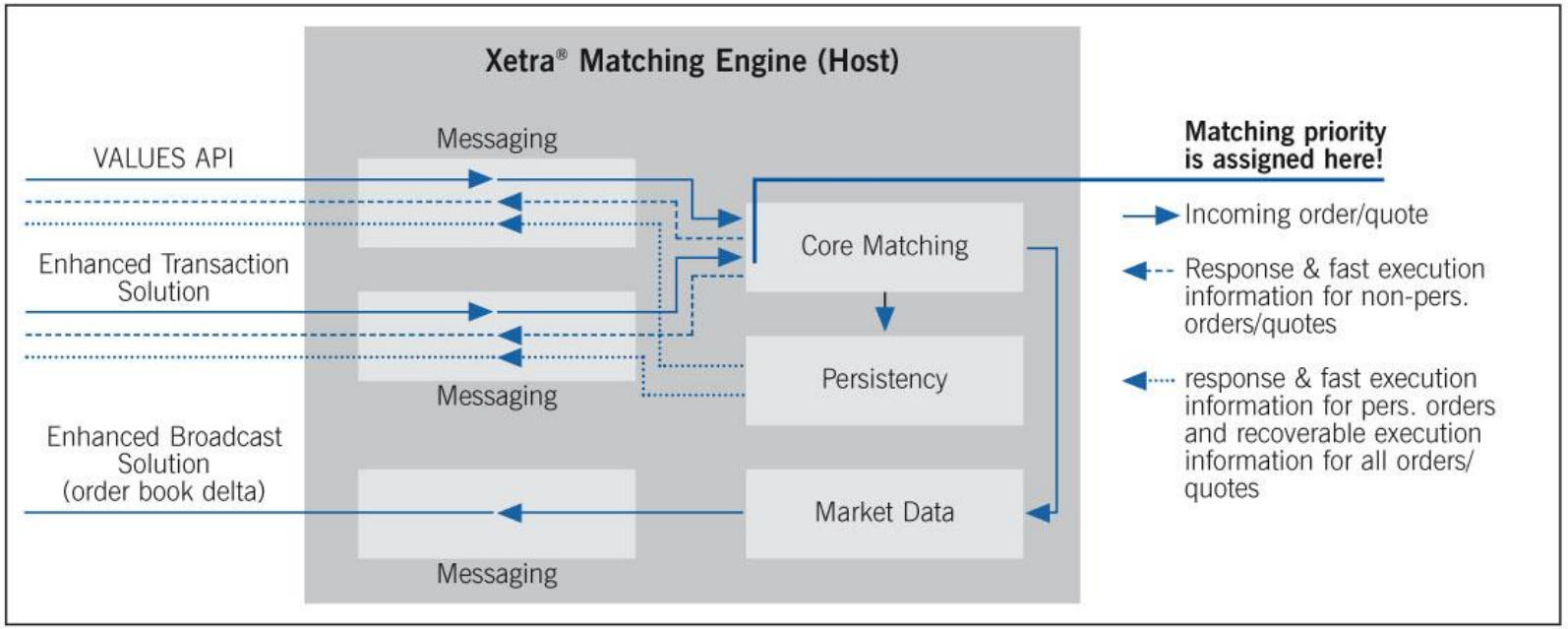
5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

Inside the matching engine – Overview



- Orders / quotes entered via VALUES API go through a MISS and a so-called communication server (not streamlined for low latency).
- Orders / quotes entered via the Enhanced Transaction Solution (ETS) go through a gateway (streamlined for low latency).
- Co-location in combination with a 10 GbE ETS connection offers the lowest possible latency.

Inside the matching engine – Details (1/3)

- Orders / quotes entered for a specific instrument are forwarded by the gateway / communication server to the respective Xetra® matching engine.
- Two **messaging components** handle VALUES API and Enhanced Transaction Solution messages and hand these over to the core matching component where time priority is assigned.
- Statistically speaking, most “delays” during high load situations in the Xetra system occur here.

Inside the matching engine – Details (2/3)

During continuous trading, the core matching component works as follows:

- When an order / quote arrives, it is functionally processed (e.g. put in the book).
  Resulting messages for **non-persistent orders / quotes** are sent out in the following order:
  - direct response to the order / quote entered
  - fast execution information for booked orders / quotes according to matching priority
  - fast execution information for the aggressing order / quote
- A handover of data to the **market data component** which calculates the inside market and sends out an order book delta / incremental message including **up to 20 trade prices** in case a match takes place.
- Resulting messages for persistent orders and executions are sent to the **persistency layer**.
- After being persisted, the resulting messages for **persistent orders** and **executions** are sent out. Those are the response and fast execution information for persistent orders as well as the recoverable execution information for all orders.
- Finally, in case of a match, the **All Trade Price message** is published.

Inside the matching engine – Details (3/3)

If one or more orders / quotes arrive while another order / quote is being processed:

- The core matching component will batch these orders / quotes and treat them in one **unit of work**. Direct responses and fast execution information will be sent out one by one for non-persistent orders / quotes while this unit of work is processed.
- At the end of the unit of work the resulting final order book and all trade prices are handed over to the market data component which publishes the market data. Together with the final book, a **maximum of 20 trade prices** is then published in one delta / incremental message.
- Data will be handed over to the persistency layer, and the resulting messages for persistent orders and executions are sent out. Those are the response and fast execution information for persistent orders as well as the recoverable execution information for all orders.
  Finally, in case of a match, the **All Trade Price message** is published by the market data component.

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

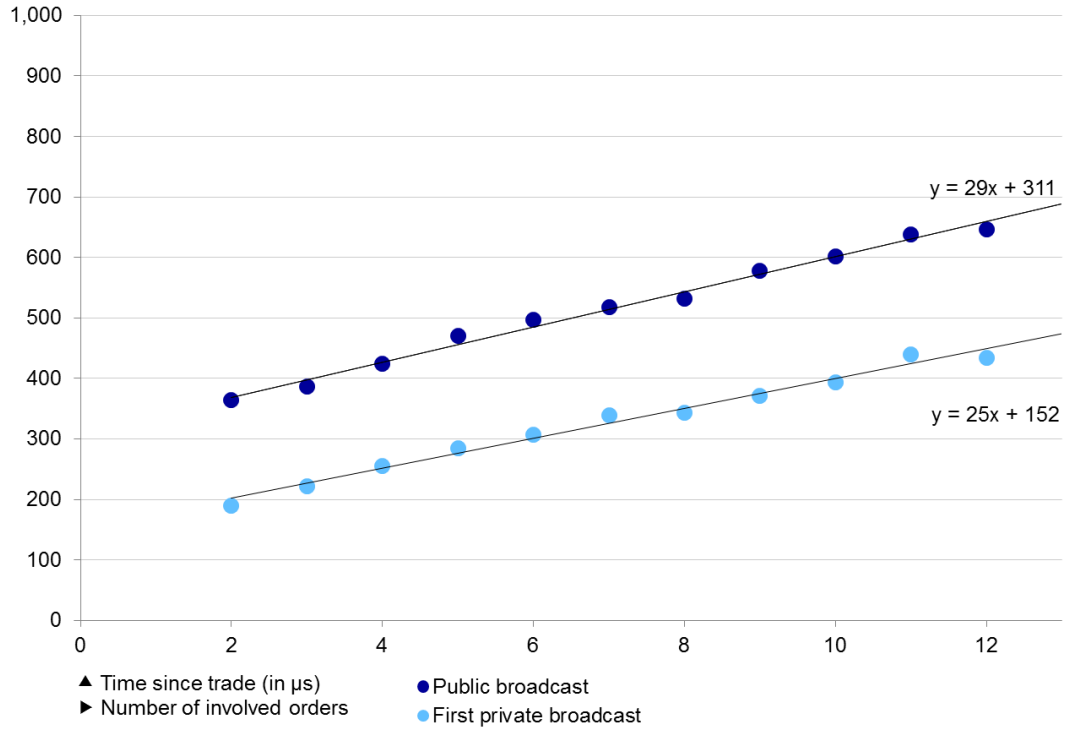
8 Appendix

Trading system dynamics

Latency of executions

The diagram on the right shows DBK production data from February, 8 2016 on the relationship between trade complexity and broadcast latency.

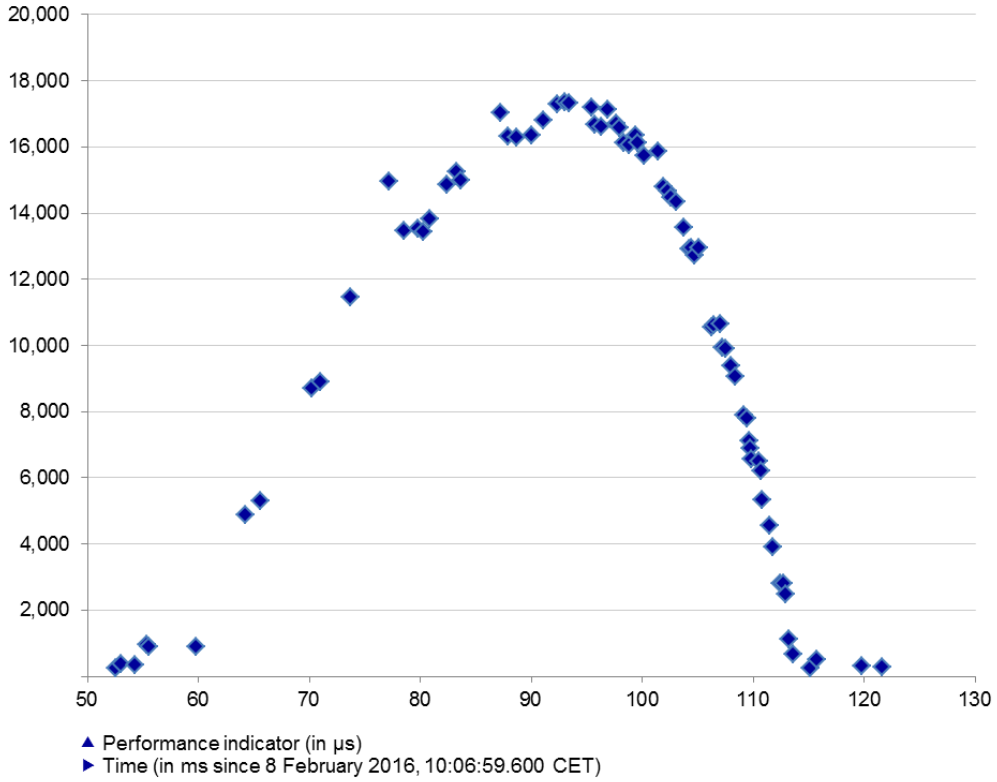
- The dark blue dots represent the median time it took from a match to the publication of the trade (t_8 – t_7).
- The light blue dots represent the median time it took from a match to the first private broadcast (t_6 – t_7) leaving the host.
- The median difference in time between two consecutive passive execution messages leaving the host is about 15 µs.



Trading system dynamics

Matching engine and market data performance

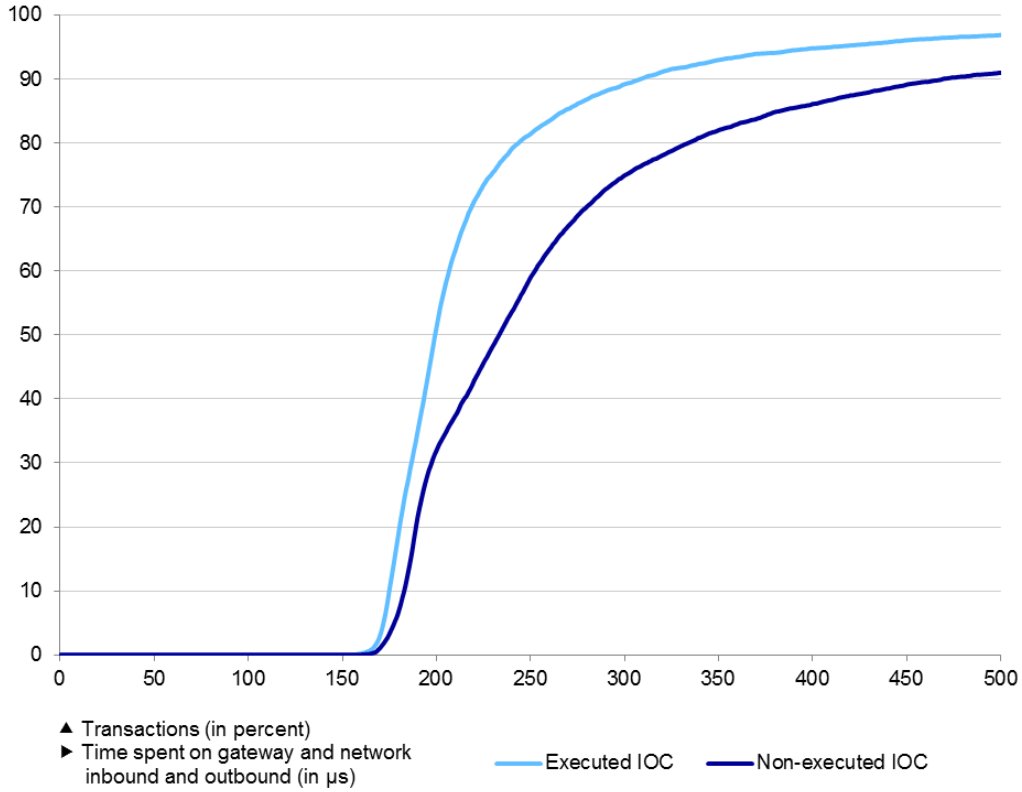
- During microbursts, our matching engines typically experience a temporary performance degradation.
- Unpredictable latencies cause risk (i.e. it might take longer to pull an order).
- Xetra® provides real-time insight into the matching engine and market data component performance by providing the performance indicator (t_8 – t_5) in the Enhanced Broadcast Solution packet header.
- This graph depicts an example of the performance indicator use case where a microburst caused a latency spike in DBK on February 8, 2016.



Trading system dynamics

Gateway and network latency

- Congestion on the gateways and network between the gateways and matching engines can cause significant latency.
- Participants experience this component to be relatively erratic.
- The extent of this behaviour can be illustrated by comparing the cumulative frequency distributions of the network latency for executed and non-executed IOCs.
- The diagram shows the time that executed versus non-executed IOC orders in DBK – on February 16, 2016 – spent on the gateways and between the gateways and matching engines, i.e. ((t_4 – t_3) – (t_6 – t_5)).
- The graph shows that for executed IOC orders the latency profile approaches a step function.



Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

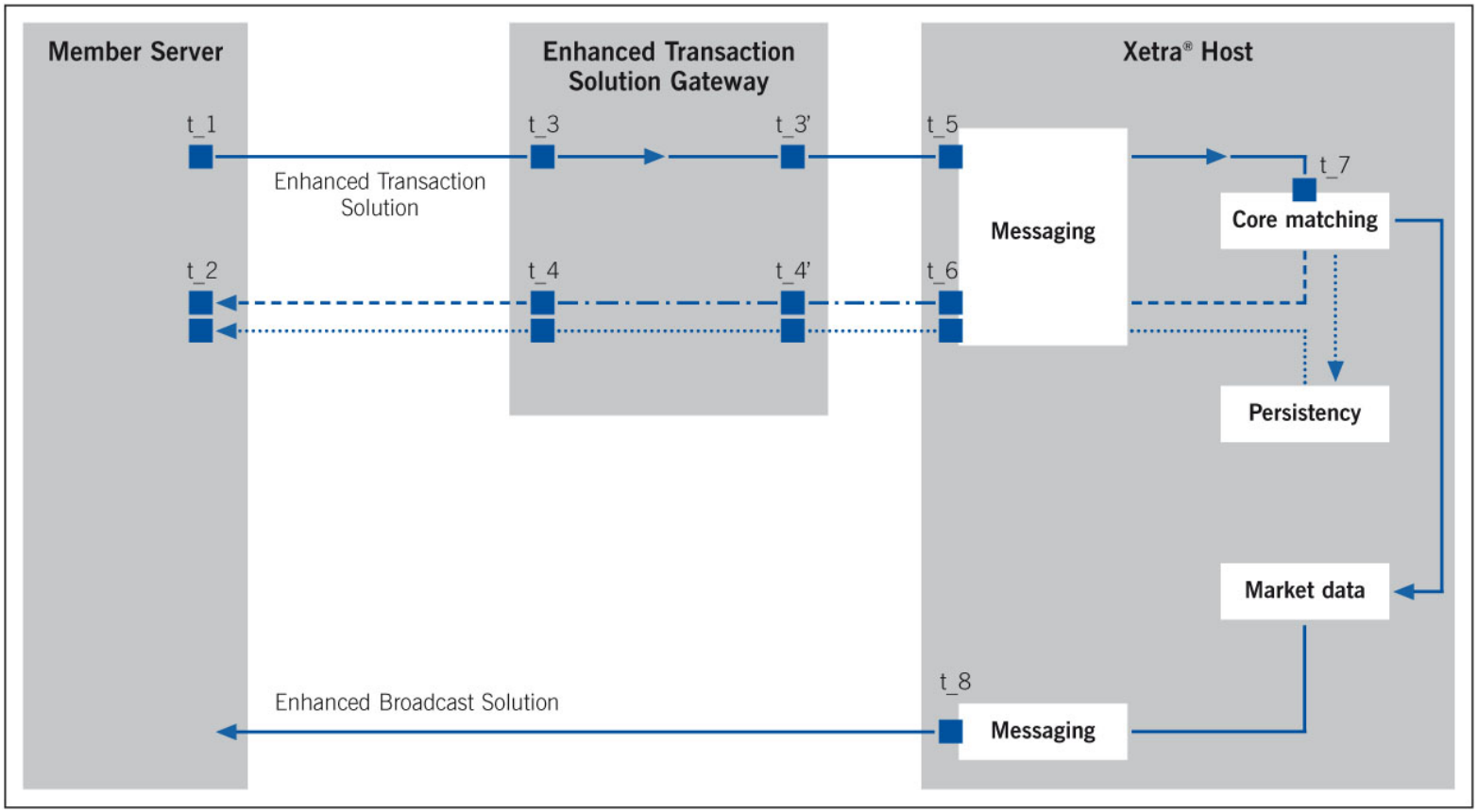
5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

Timestamps in Xetra® – Overview



Timestamps in Xetra® – Definitions

- **t_1 and t_2** can be taken by member applications when a request / response is read from / written to the socket.
- **t_3 and t_4** are taken by the Enhanced Transaction Solution gateway when request / response is read from / written to the socket on the participant's side of the gateway. This data is contained in the (private) Enhanced Transaction Solution response.
- **t_3' and t_4'** are taken by the Enhanced Transaction Solution gateway when request / response is read from / written to the socket to / from the matching engine. This data is contained in the (private) Enhanced Transaction Solution response.
- **t_5 and t_6** are taken by the matching engine when request / response is read from / written to the socket. This data is contained in the (private) Enhanced Transaction Solution response.
- **t_7** is taken when the order is functionally processed (e.g. put in book or matched). It is available in the (private) Enhanced Transaction Solution response and Enhanced Broadcast Solution order book delta and Enhanced Broadcast Solution trade message in case the order matches.
- **t_8** is taken just before a market data UDP datagram is written to the socket.

For every Enhanced Broadcast Solution datagram that contains an order book incremental message, Xetra publishes a **performance indicator** in the packet header. The value of the indicator is the time difference
t_8 – t_5. Here t_5 is the time when the "oldest" aggressing order which is contained in this market data update reached the Xetra matching engine.

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

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 the Xetra® system.
- Use state-of-the-art switches and only have at most one hop between the exchange network and your server.
- Use good NIC cards and TCP/IP acceleration, e.g. a Linux kernel bypass library.
- Use 10 GbE (cross-) connections for the Enhanced Broadcast Solution and Enhanced Transaction Solution.
- Try the two different Enhanced Transaction Solution gateways assigned to each session to see which delivers the better performance for your strategy (try it out and compare your timestamps as well as P&L for different days).
- If possible, measure and analyze your own timestamps.
- Use the Enhanced Transaction Solution fast info notifications for trading decisions although executions reported by such messages are only indicative (and not legally binding).
- Recoverable messages need to be processed to create safety. Therefore, we recommend to use either a FIX trade capture drop copy or the trade notification on member level in the Enhanced Transaction Solution interface to confirm the fast info notifications.

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

Outlook

In the light of the planned Xetra migration to T7 and the possible postponement of MiFID II, Xetra is currently reviewing its release planning for 2016. Member communication will be sent out in due time.

This presentation will be revised from time to time to provide current data.

Finally, Xetra will continue to investigate possibilities of extending the transparency with respect to latency figures – in case you have any suggestions, please get in touch with us!

Thank you for your attention!

For more information please contact
Fabian Rijlaarsdam
Phuong Hieke
Sebastian Neusüß
Market Microstructure and Latency
hft@xetra.com

Agenda

1 Achievements

2 Topology

3 Inside the matching engine

4 Trading system dynamics

5 Timestamps

6 What you need to be fast

7 Outlook

8 Appendix

Timestamps – Reference

The timestamps t_3 – t_8 are available via the following fields:

▪ **t_3**:	gatewayReqInTimeStamp	in the Enhanced Transaction Solution response notification
▪ **t_3'**:	gatewayReqOutTimeStamp	in the Enhanced Transaction Solution response notification
▪ **t_4**:	gatewayRespOutTimeStamp	in the Enhanced Transaction Solution response notification
▪ **t_4'**:	gatewayRespInTimeStamp	in the Enhanced Transaction Solution response notification
▪ **t_5**:	coreSystemInTimeStamp	in the Enhanced Transaction Solution response notification
	aggressorTime	in the Enhanced Broadcast Solution delta message
▪ **t_6**:	coreSystemOutTimeStamp	in the Enhanced Transaction Solution response notification
▪ **t_7**:	lastEventTrnId	in the Enhanced Transaction Solution Enter / Modify / Delete Order / Quote / Mass quote response and
	entryTime	in the Enhanced Broadcast Solution Delta, Snapshot and All Trade Price message
▪ **t_8**:	SendingTime	in the Enhanced Broadcast Solution UDP packet header
▪ **(t_8 – t_5)**:	PerformanceIndicator	in the Enhanced Broadcast Solution UDP packet header