Insights into Trading System Dynamics

Deutsche Börse’s T7®

March 2025

DEUTSCHE BÖRSE GROUP

Content

3
Introduction

T7® Technology Roadmap

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

Recent developments

- 17/31 August 2024: Introduction of Colo 2.0 Mid-Layer switch for Eurex
- 7 September 2024: Tech refresh of Colo 2.0 order entry switches for Xetra A-side
- Q3/Q4 2024: Technical upgrade of T7 core trading system and T7 backend capture infrastructure.
- 28 September 2024: Introduction of Colo 2.0 Mid-Layer switch and tech refresh of Colo 2.0 market data switches for Xetra B-side
- 12/26 October 2024: Tech refresh of Colo 2.0 order entry switches for Eurex
- 9/23 November 2024: Tech refresh of Colo 2.0 market data switches for Eurex
- 18 November 2024: T7 Release 13.0
- 25 January 2025: Optimization of latency for cross-side connections
- 15 February 2025: Tech refresh of Colo 2.0 order entry switches for Xetra B-side
- 22 February 2025: Introduction of additional Colo 2.0 Access-Layer switches for Eurex order entry

Outlook

- Ongoing: Replacement of customer cassettes for Colo 2.0 connections
- 24 March 2025: Go-live of Enhanced Drop Copy Interface in Production
- 19 May 2025: T7 Release 13.1
- 10 November 2025: T7 Release 14.0
- Timeline to be fixed: Improve normalization of Colo 2.0 connections (specific dates will be announced separately)

T7® Technology Roadmap

Timeline of updates

2024

Aug
Sep
Oct
Nov
Dec

2025

Jan
Feb
Mar
Apr
May
Jun
Q3
Q4

Introduction of Colo 2.0 Mid-Layer switch for Eurex
Start of technical upgrade of T7 core trading system and T7 backend capture infrastructure
Tech refresh of Colo 2.0 order entry switches for Xetra A-side
Tech refresh of Colo 2.0 market data switches for Xetra B-side
Start of replacement customer cassettes for Colo 2.0 connections
Tech refresh of Colo 2.0 order entry switches for Eurex
Tech refresh of Colo 2.0 market data switches for Eurex
T7® Release 13.0

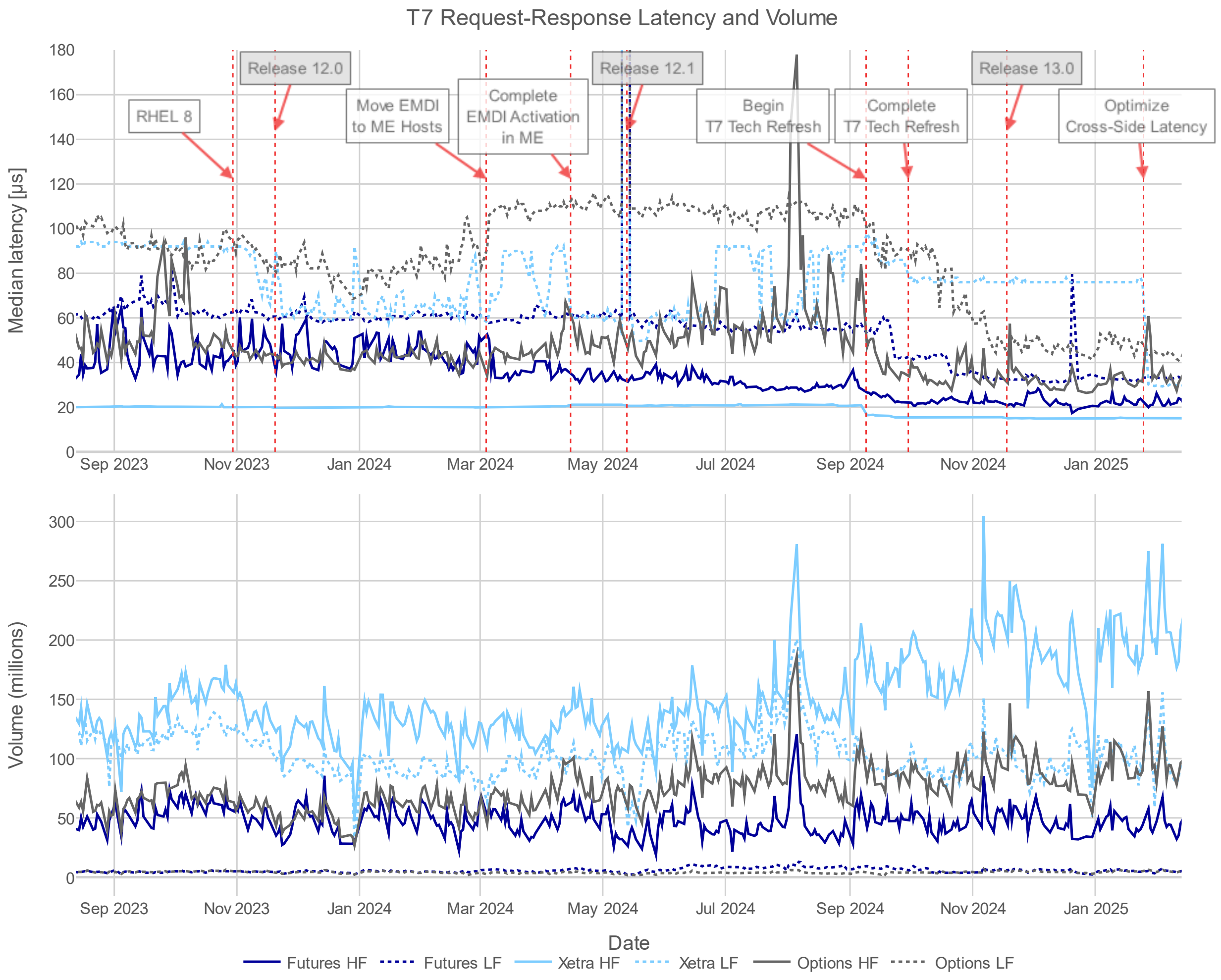
Ongoing replacement customer cassettes for Colo 2.0 connections
Optimization of latency for cross-side connections
Tech refresh of Colo 2.0 order entry switches for Xetra B-side
Introduction of additional Colo 2.0 Access-Layer switches for Eurex order entry
Go-live of Enhanced Drop Copy Interface in Production

T7® Release 13.1

T7® Release 14.0
Improve normalization of Colo 2.0 connections

Processed Transactions and Response Times

T7 request – response round-trip times



- Deutsche Börse continuously invests in its trading system and is holding up transparency while providing a low latency trading venue.
- We have continued to add functionality while at the same time tuning our system further.
- The technical upgrade of the T7 core trading system in Q3/Q4 2024 significantly improved throughput and latency. The positive effects can be observed in the median latency values as well as in higher percentiles.
- The optimization of latency for cross-side connections had a positive impact on Xetra LF RTT statistics.

7
Recent Developments

EOBI DSCP Field and Discard IP Range

Look back

What is speculative triggering?

Initiation of sending a request by a participant at a time when not all required information are available.

There are certain triggers in a market data packet that reveal information:

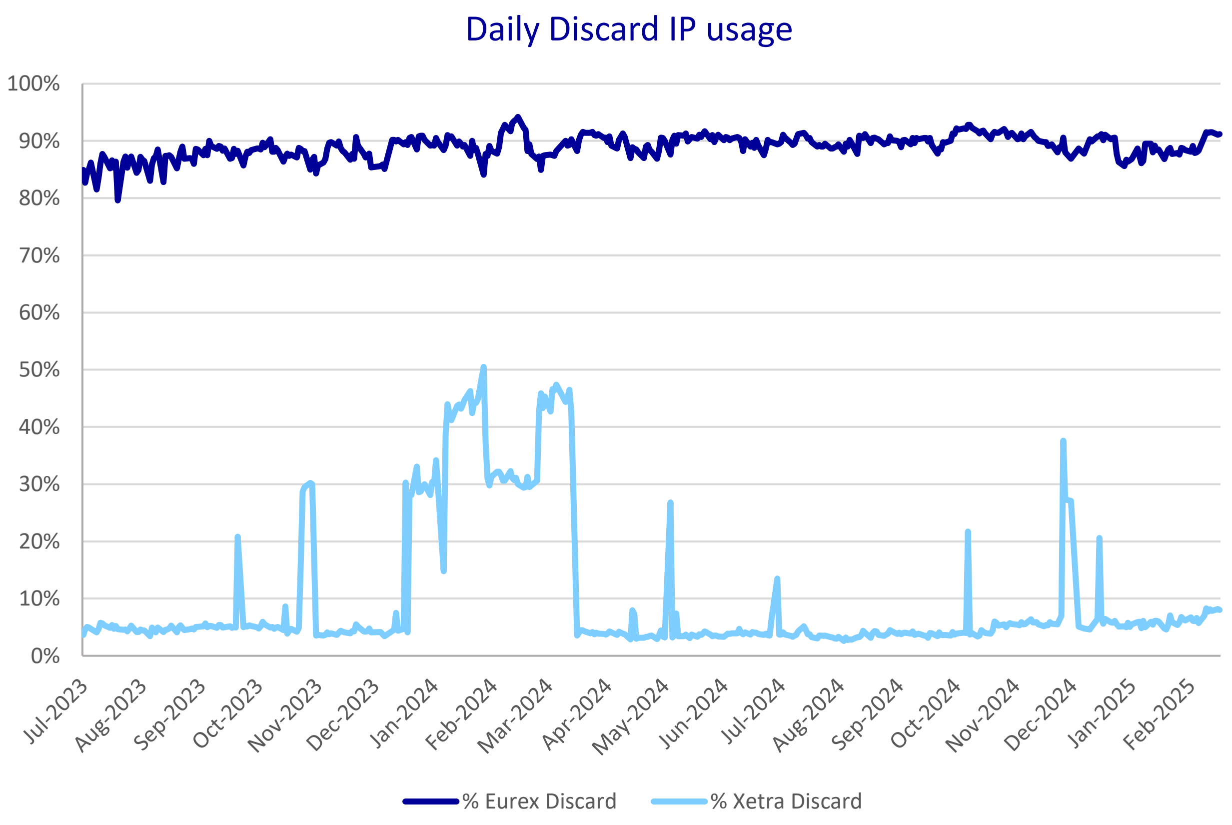
1. Ethernet preamble of market data (earliest time = 0 ns)
2. Destination MAC (ethernet header) = identification of the multicast stream
3. IP total length: Indicates the type of message contained, e.g. ExecutionSummary or single order add
4. Payload (up to 100 ns later): Quantity and Price.

Why is it used?

Speculative triggering is incentivized by the deterministic technical network architecture of the T7 trading system. The 10 Gbit/s Access Layer switches operate in a 'cut-through' mode and the first bytes of an ethernet frame reserve priority on the competitive network path for the uplink to the next switch. This incentivizes latency sensitive trading participants to send technical transactions purely to reserve switch priority creating a high load on the T7 system. To avoid unnecessary technical transactions on the T7 a technical solution has been implemented.

EOBI DSCP Field and Discard IP Range

Confining the effects of speculative triggering



In July 2020, Eurex introduced a Discard IP address range 172.16.0.0/16 on the 10 Gbit/s order entry networks. Xetra followed in May 2021. Trading participants may send falsely 'speculative' triggered packets to the discard IP range, instead of sending it to the exchange. These packets will be discarded at the Access Layer switch port and no other participant is influenced. Packets sent to the discard IP address are not considered to be orders and are not forwarded to the exchange.

To enable market participants to effectively use this discard IP address, the DSCP field of the IPv4 headers in EOBI market data packets is used. Four different bits indicate the most common ‘interesting’ market situations.

The number of packets reaching the Matching Engines decreased significantly after introduction of this measure. Since then, the number of Discard IP packets has constantly increased.

The graph on the left shows the percentage of Discard IP packets compared to total number of packets reaching the Access Layer switches as an average per day.

EOBI DSCP Field and Discard IP Range

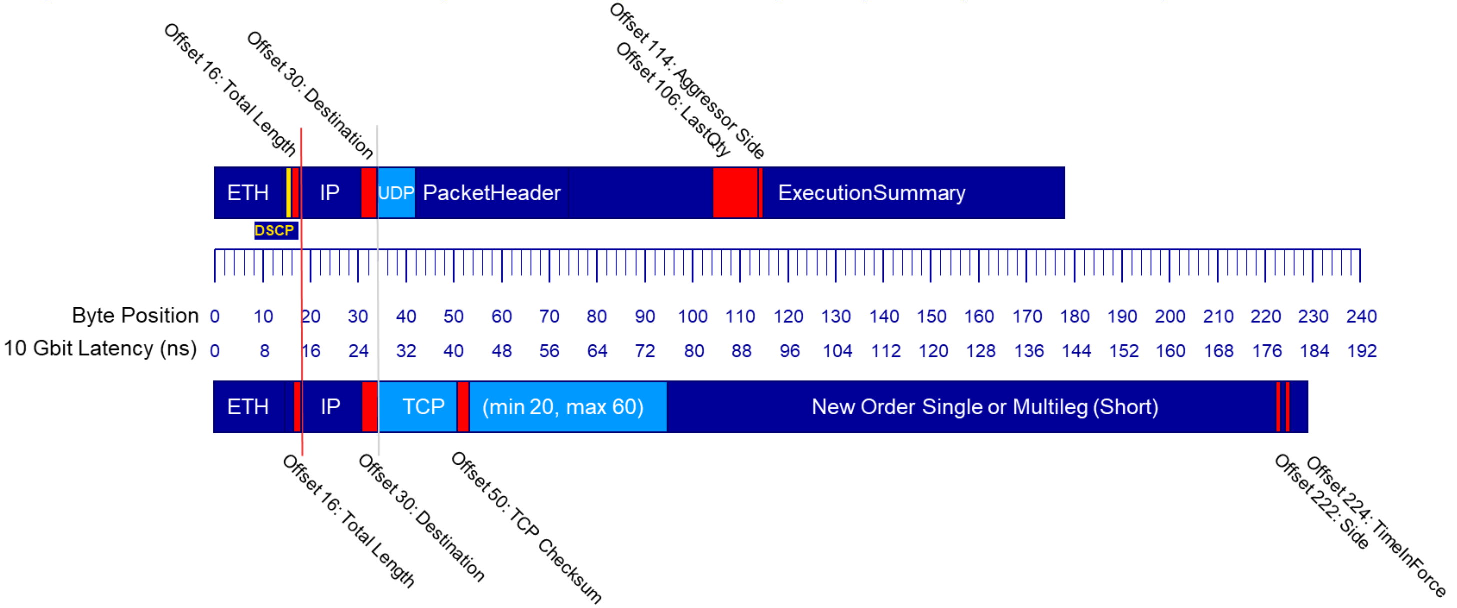
Confining the effects of speculative triggering (continued)

A technical solution to the speculative triggering problem:

- Mark potential triggers early in the IP header of market data packet with the help of DSCP (Differentiated Services Code Point) flags.
- Offer a non-competitive Discard IP destination address to enable packets to be discarded right on the Access Layer switch.

DSCP flags indicate Execution summaries and/or widening or narrowing of the bid/ask spread from orders (not quotes).

Examining these flags allows participants to still change the destination IP address of an in-flight outgoing message to the Discard IP address for uninteresting packets. These packets will be sent on the exchange network but will not reach the trading system as they will be rejected on the Access Layer switch port facing the participant sending the transaction.

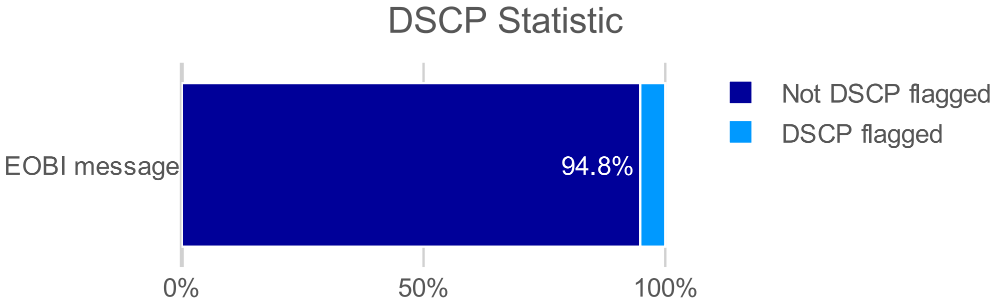


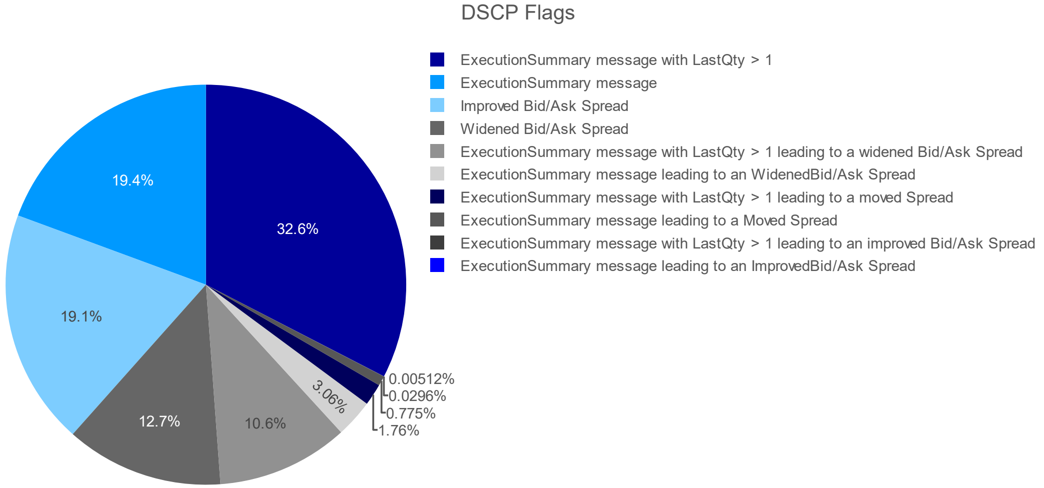
- The left figure shows at which byte position different information is available in the message.
- A market data packet is received at time 0.
- The fastest participants may react and send a response as early as the first bit of market data has been received, dynamically reading the market data packet while already streaming out the response.
- The response packet may be modified in-flight, after reading e.g. the DSCP flags, total length etc.
- Even with a reaction time of 0 the outgoing destination IP address can still be modified after evaluating the DSCP flags of the incoming market data packet.

DSCP Statistics for Selected Products

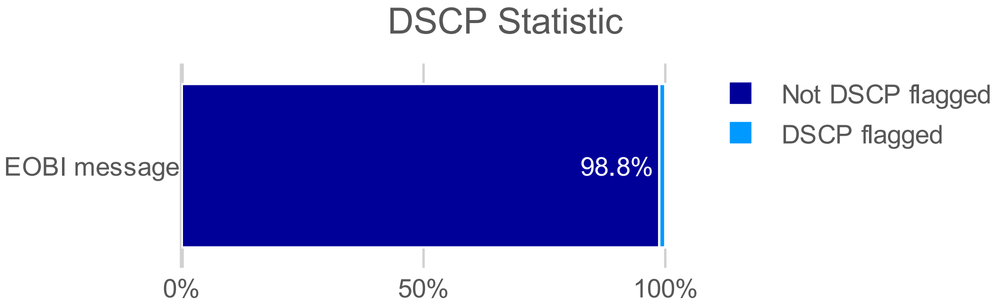
FESX and OESX

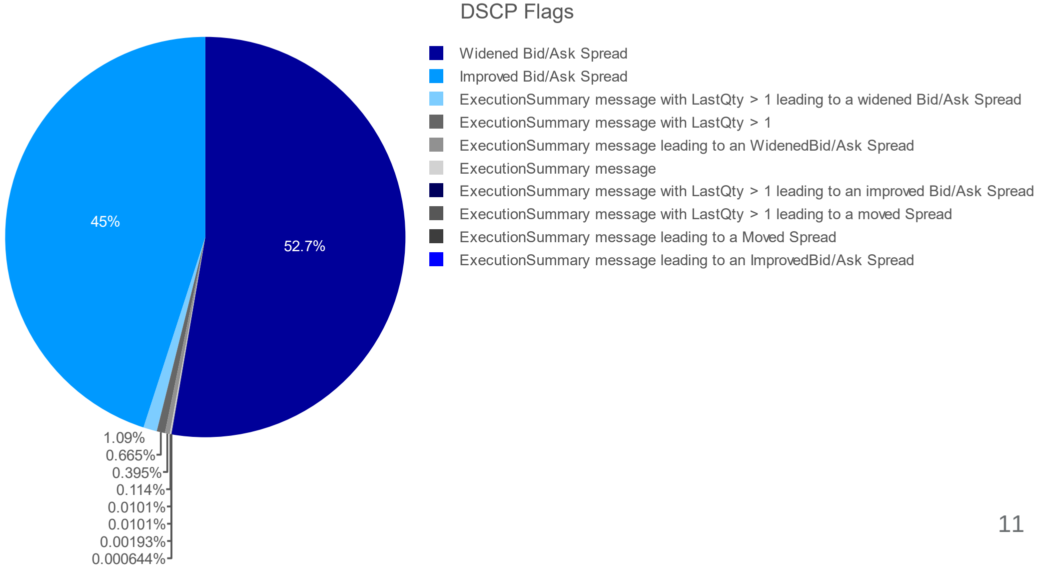
FESX DSCP Statistics 2025-02-13





OESX DSCP Statistics 2025-02-13



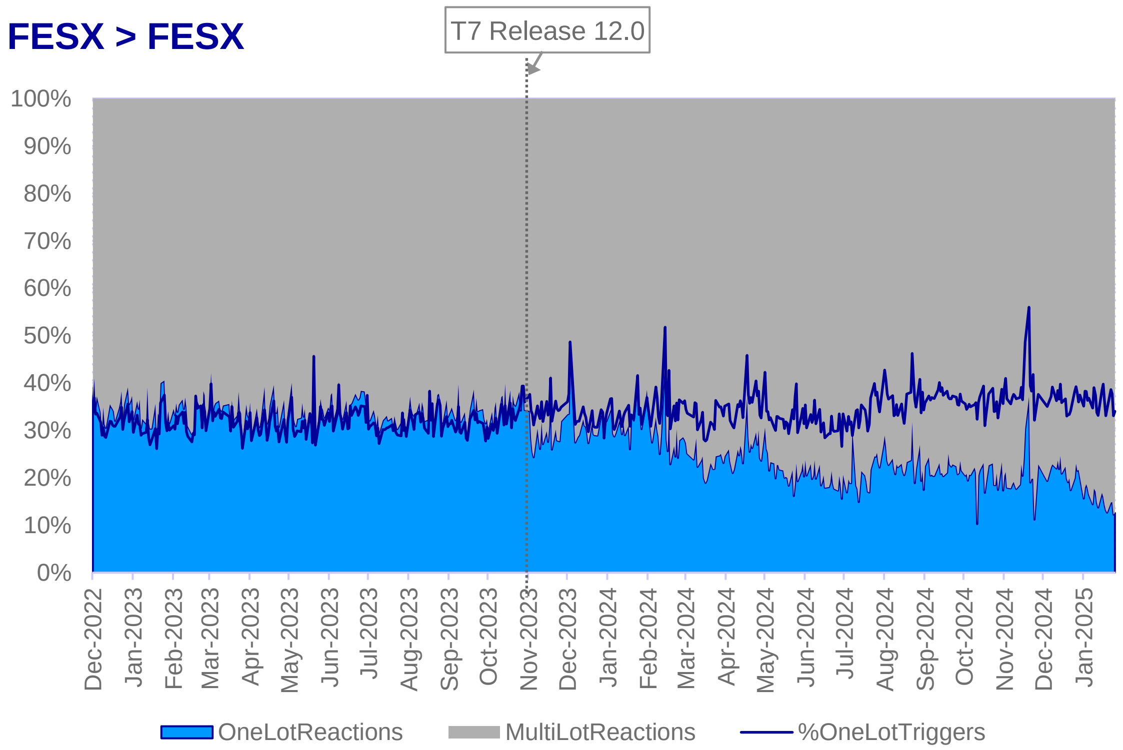


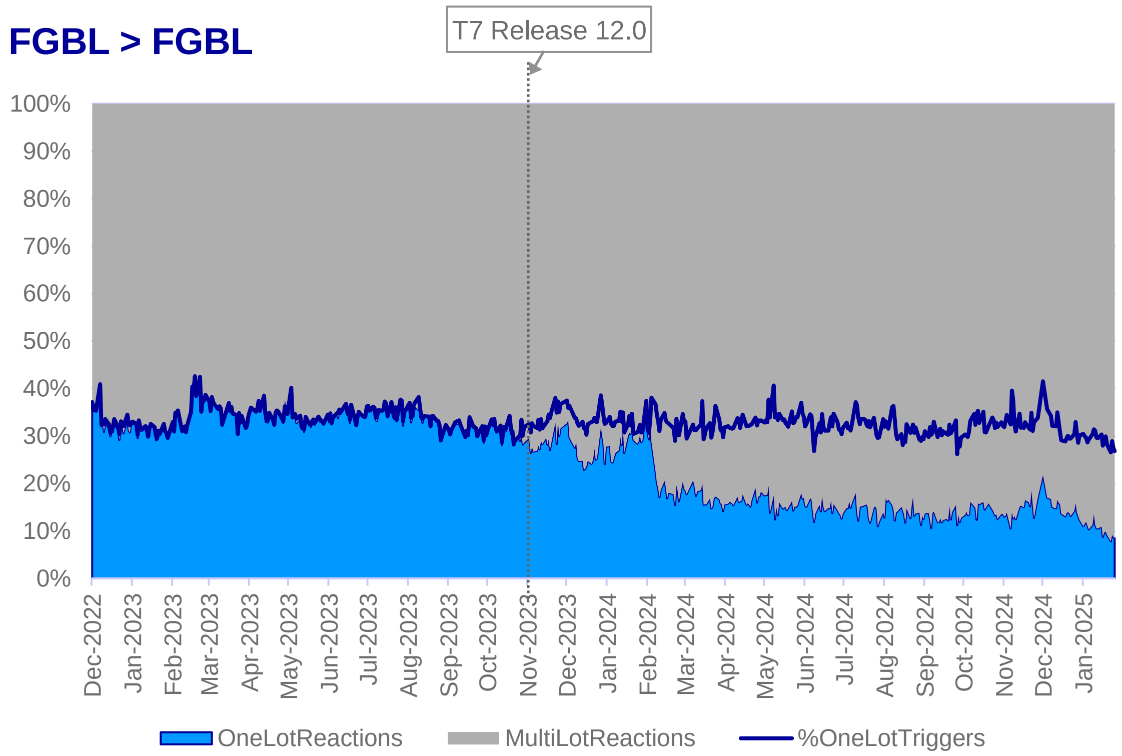
EOBI DSCP Field

Speculative Triggering on 1-lot Trades

With T7 Release 12.0, Eurex introduced a DSCP flag indicating whether the traded quantity was 1-lot or bigger.

This led to a significant reduction in number of speculative reactions* to 1-lot trades in several benchmark Futures products.





* Reactions with a difference between t_9d and t_3a of below 2890 ns

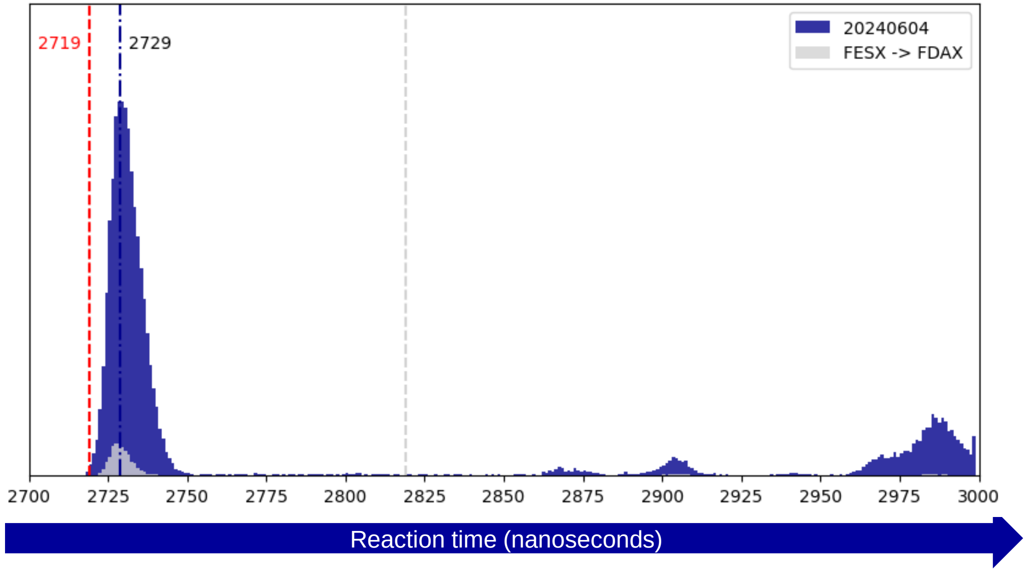
Participant Reaction Time

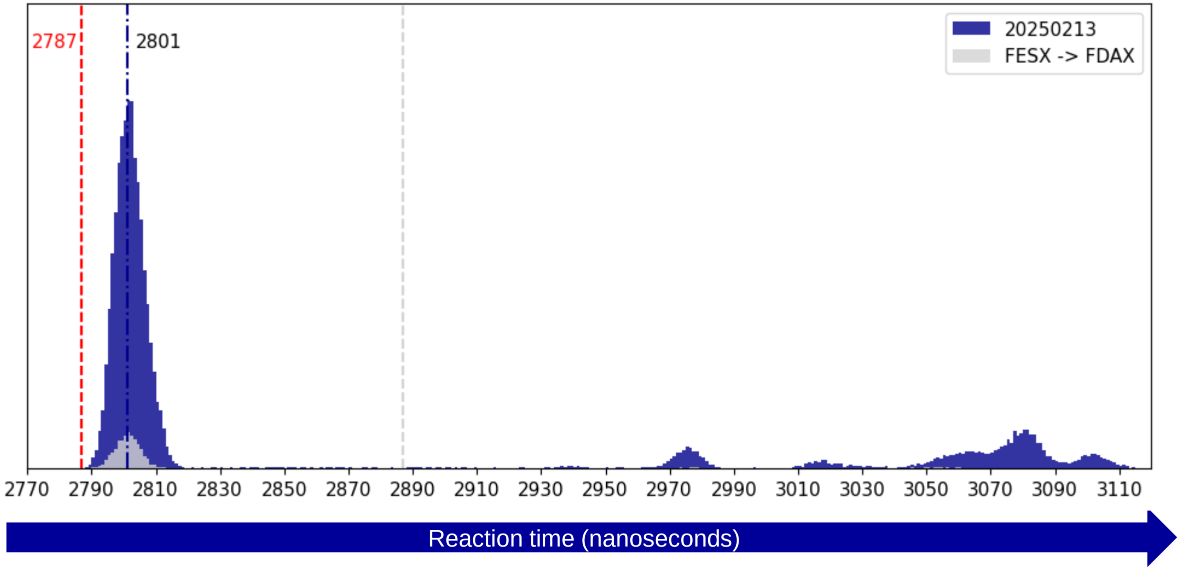
Insights from the High Precision Timestamp file - Eurex

We define reaction time as the latency between a ‘trigger’ market data packet and a request that leads to an execution.
We use measurement point t_9d for the market data packet and t_3a for the request. We measure 2787 ns as the minimum latency between t_9d and t_3a determined by the T7 infrastructure. The minimum reaction time changed due to hardware upgrades in Colo 2.0 and tech refresh of the T7 core system and backend capture devices.

The below charts show the distribution of the reaction times for all Eurex products (blue) and FESX to FDAX (grey) from 13 February 2025 (right) compared to the last figures from 4 June 2024 (left).

We observe a high level of competition (there are around 10 trading participants with reaction times < 2830 ns for most active products).





Participant Reaction Time

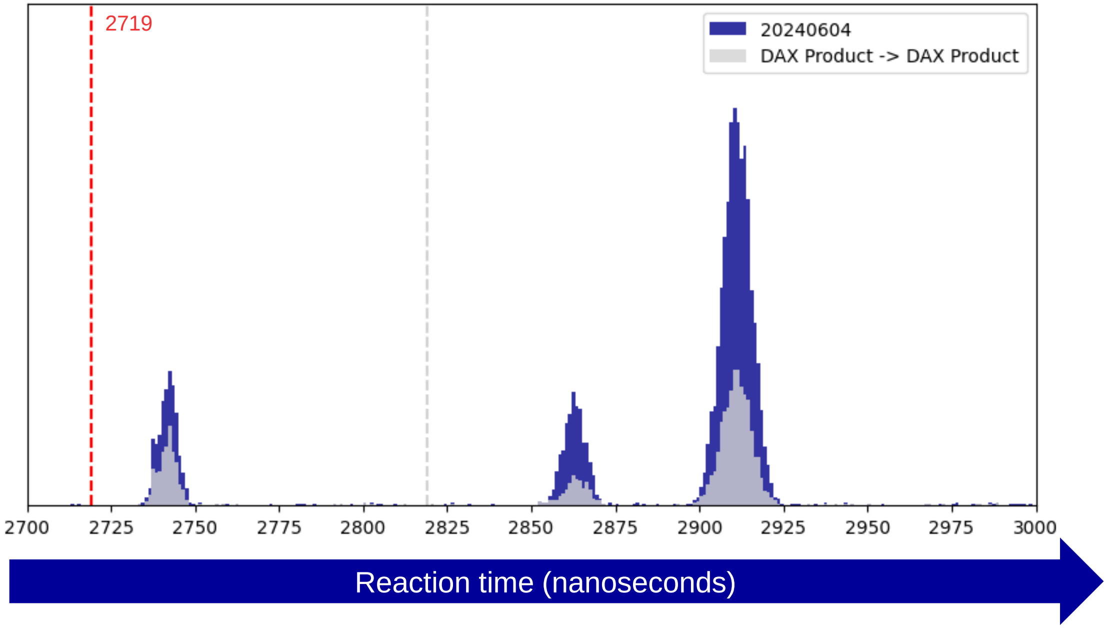
Insights from the High Precision Timestamp file - Xetra

The charts below show the distribution of the reaction times for all Xetra products (blue) and DAX to DAX products (grey) from 13 February 2025 (right graph) compared to the reaction times for Xetra B-side products from 4 June 2024.

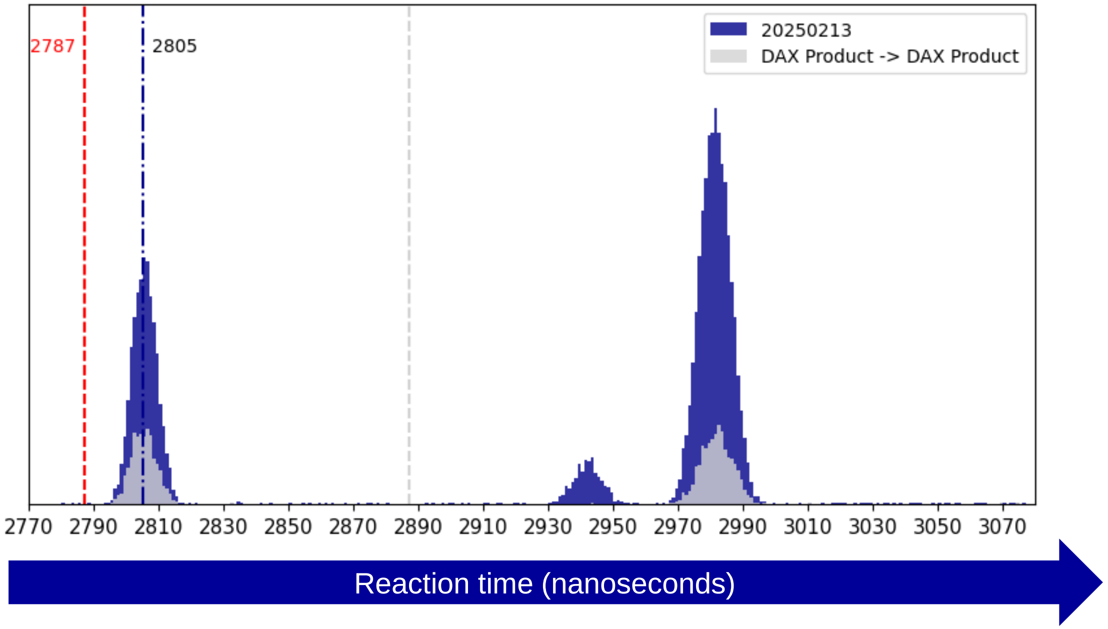
Trigger products are limited to Xetra.

The minimum reaction time changed due to hardware upgrades in Colo 2.0 and tech refresh of the T7 core system and backend capture devices.

Reaction times Xetra products B-side



Reaction times Xetra products both sides



15

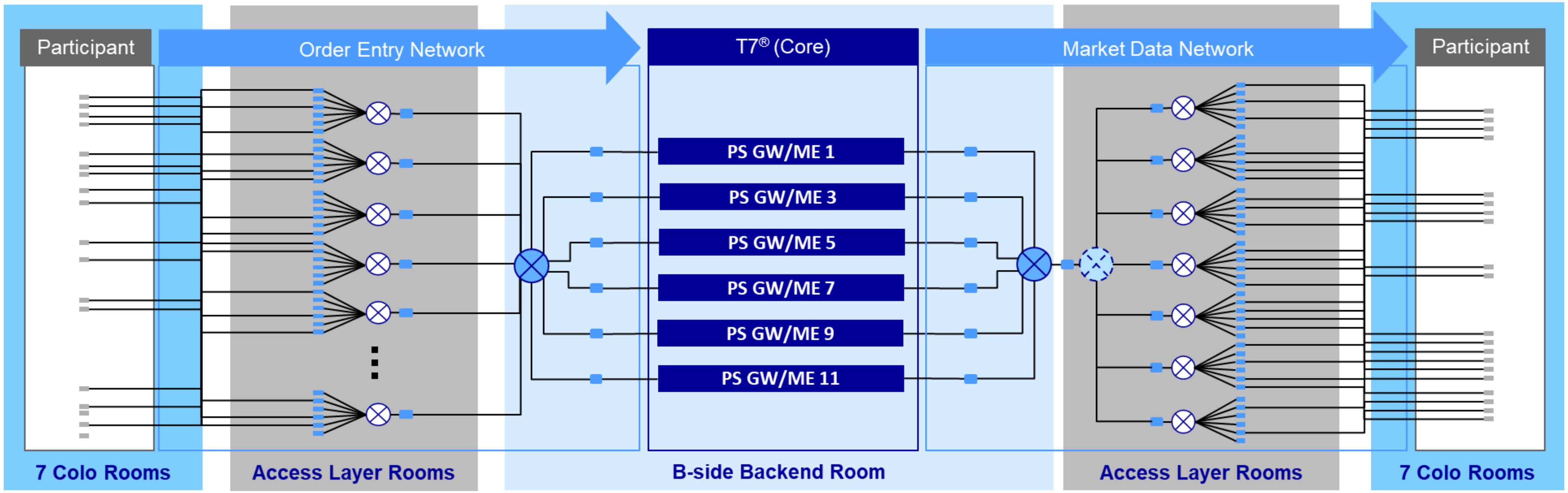
Latency Analysis

T7® Topology

At a glance

The below shows the Eurex B side (uneven partitions) as a schematic example of the topology of the T7 system.

Note that Xetra has only two Access Layer switches per side.



Eurex currently has 9 Order Entry switches on each side.

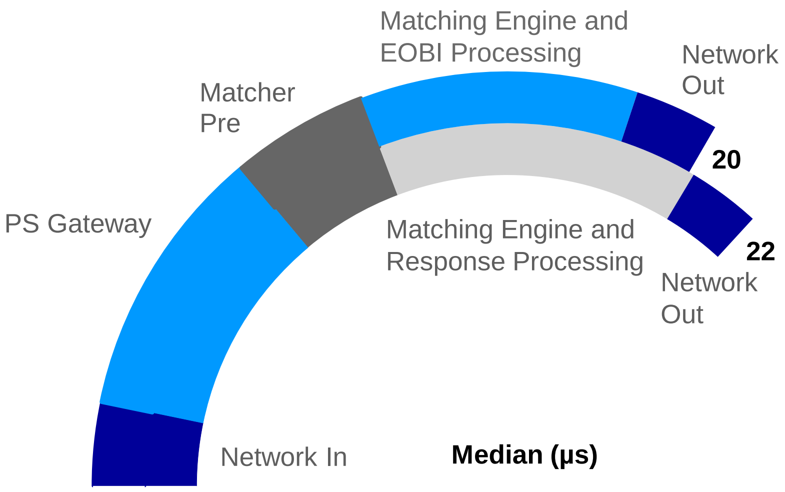
- Cisco 3548X Distribution Layer Switch
- Cisco 3550T Mid-Layer Switch
- Cisco 3548XL Access Layer Switch for Order Entry
  Cisco 3550T Access Layer Switch for Market Data
- Optical tap (capture and time measurement)

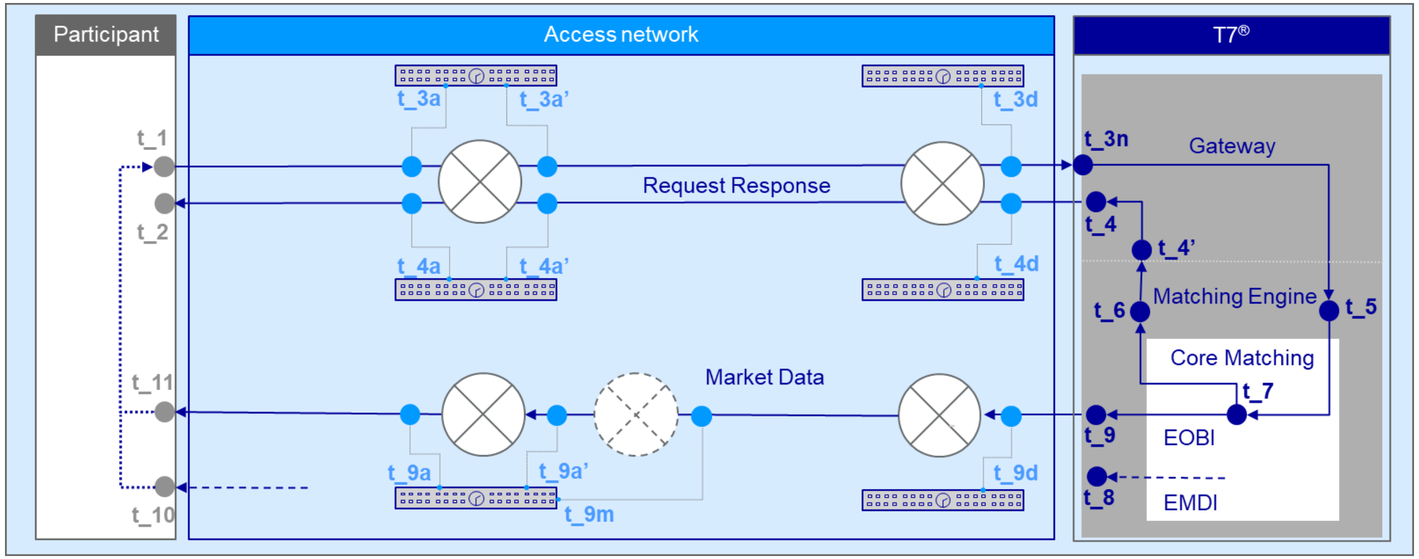
T7® Latency

Composition

The T7 trading system provides utmost transparency about its latency characteristics. Most of the timestamps taken at the various measurement points within T7 Core are included in each ETI response and EOBI market data. With the high precision timestamp file, we also make three network timestamps available for each EOBI market data packet (t_9d) and it's triggering transaction (t_3a, t_3d).

The latency circle shows the median latencies for the request-response/EOBI market data path for Eurex futures sent via high frequency sessions measured on 13 February 2025.

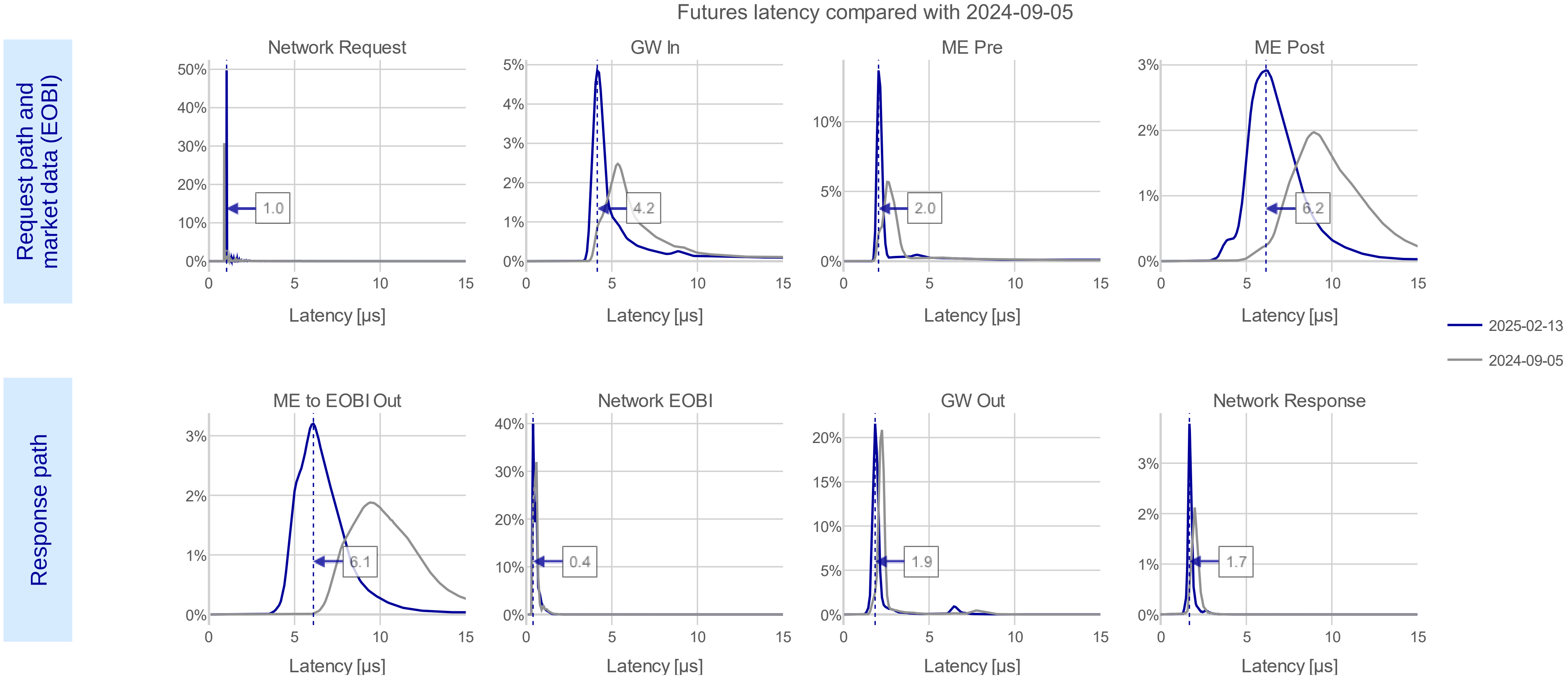




T7® Latency

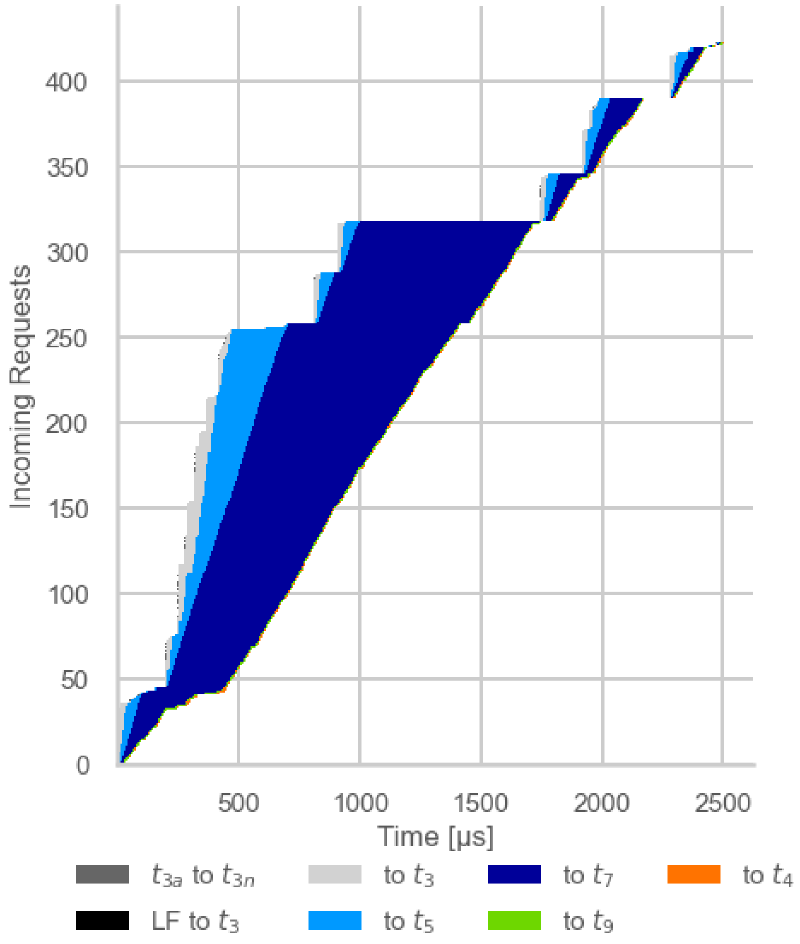
Composition (continued)

The below charts show a comparison of current latencies with the spectrum from before the technical upgrade of T7 core trading system. The data is for Eurex Futures sent from HF sessions. 'Network response/EOBI' include the TCP/UDP stack.



T7® Latency

Eurex Micro-burst dynamics



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 six timestamps with each response and key timestamps with every market data update.

The left chart shows for one example burst in FESX on 13 February 2025 the following paths:

- Network Access Layer (t_3a) to PS GWIn (t_3n) to PS GW SWIn (t_3) to
- Matching Engine in (t_5) to
- Start matching (t_7) to
- EOBI SendingTime (t_9) [where available] to
- ETI SendingTime (t_4).

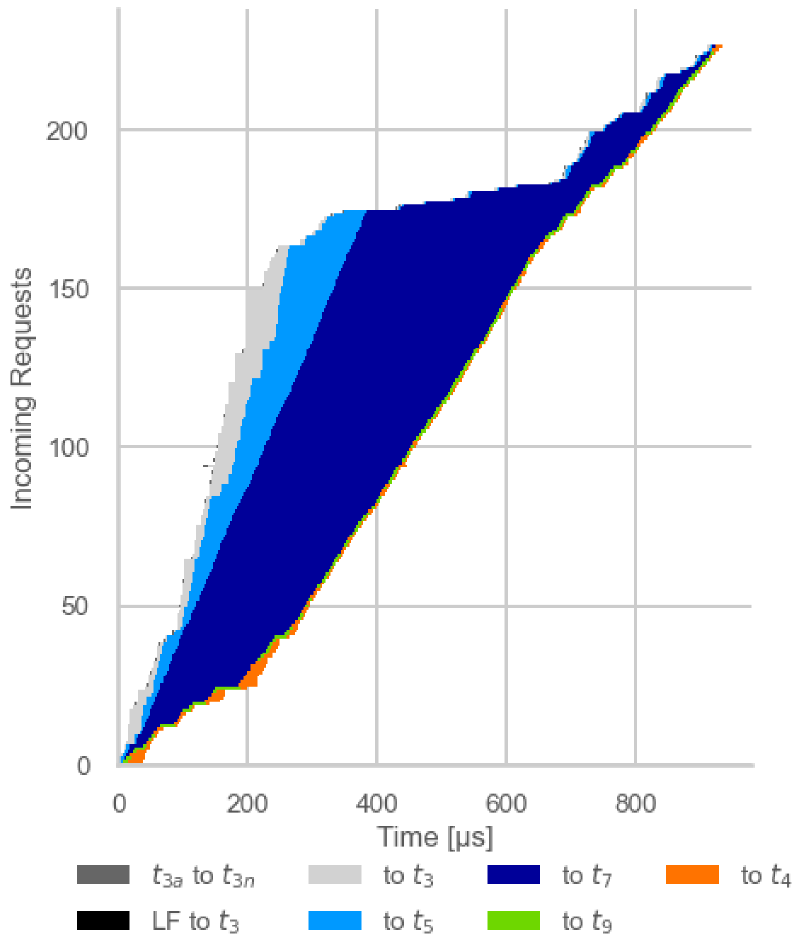
Typical throughput rates in kHz (1/ms) are 8000 at t_3n, ~500 at t_5, ~300 at t_7.

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

Note that base latency for requests entered via LF gateways is ~ 12 µs higher. As all requests are routed via PS gateways no overtaking between LF gateway and PS gateway requests is observed.

T7® Latency

Xetra Micro-burst dynamics



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 six timestamps with each response and key timestamps with every market data update.

The left chart shows for one example burst on Xetra partition 57 (4 June 2024) the paths:

- Network Access Layer (t_3a) to PS GWIn (t_3n) to PS GW SWIn (t_3) to
- Matching Engine in (t_5) to
- Start matching (t_7) to
- EOBI SendingTime (t_9) [where available] to
- ETI SendingTime (t_4).

Typical throughput rates in kHz (1/ms) are 8000 at t_3n, ~500 at t_5, ~300 at t_7.

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

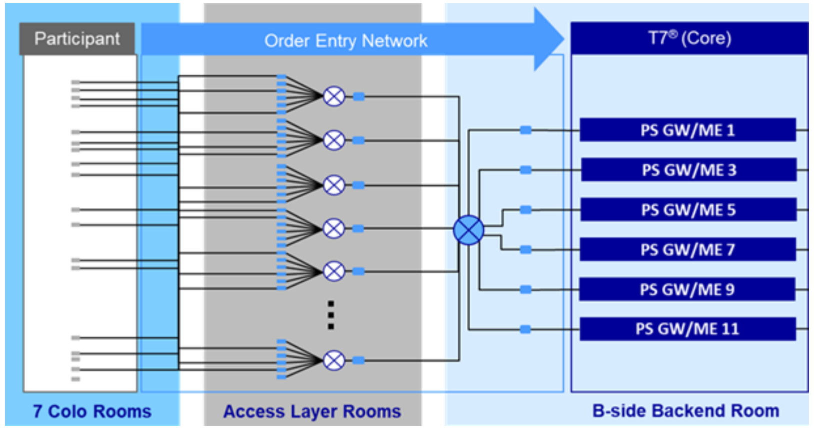
Note that base latency for requests entered via LF gateways is ~ 12 µs higher. As all requests are routed via PS gateways no overtaking between LF gateway and PS gateway requests is observed.

T7® Latency

Order entry network

The order entry network in Colocation 2.0 uses a two staged hierarchical funnel in approach.

All cables are normalized to guarantee a maximum deviation of ±0.35 m (±1.75 ns) between any two cross connects to the exchange. The reduction from ±0.5 m was achieved as an intermediate step and Deutsche Börse plans to further improve the cable length normalization.



Every inbound and outbound packet on this path is captured via passive network TAPs at 3 different locations.

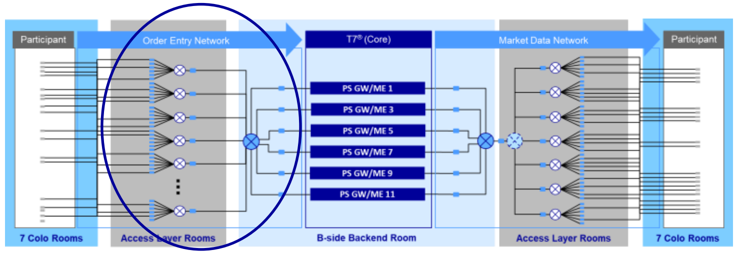
The packets are timestamped with sub-nanosecond resolution and nanosecond accuracy.

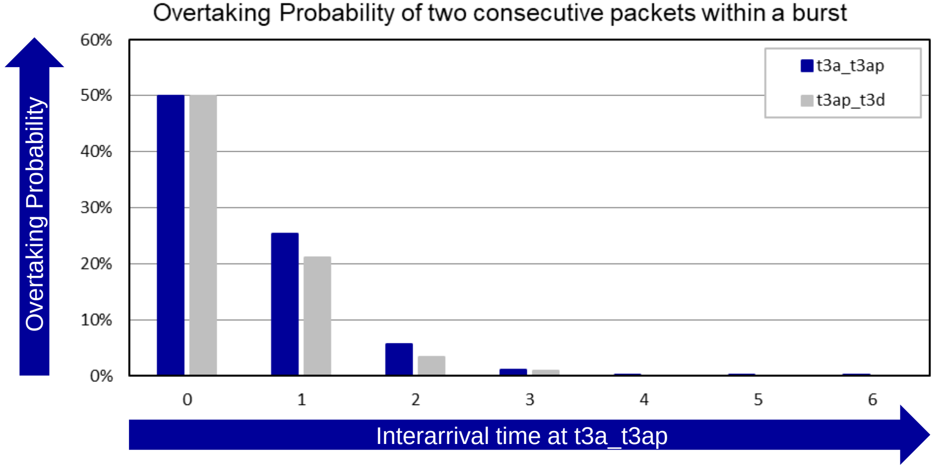
This capture infrastructure allows early detection and analysis of many kinds of technical network problems and an in-depth latency analysis on network level, like overtaking probabilities between measurement points.

The high precision timestamp file service enables participants access to timestamps t_3a, t_3d and t_9d for each request leading to an EOBI market data update.

Order Entry Network

Latency aspects



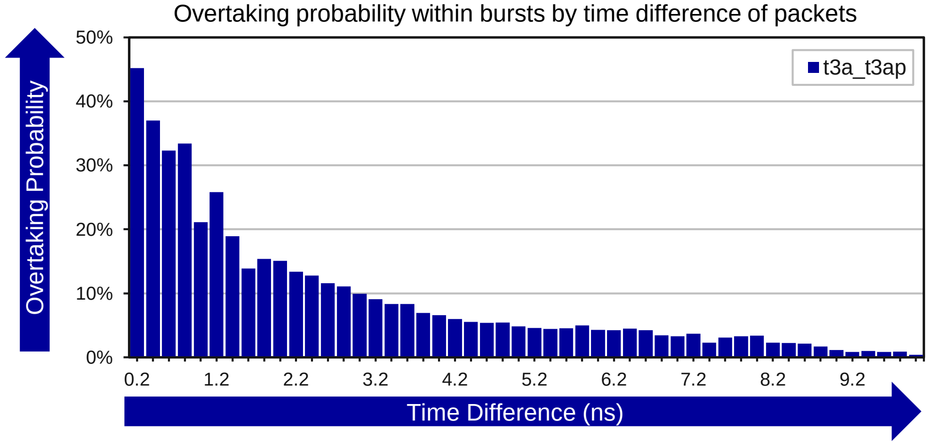


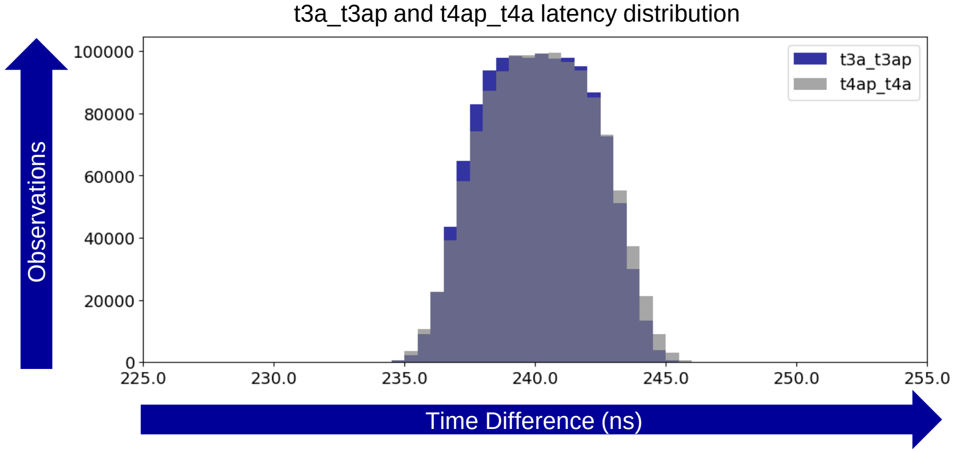
The chart on the left shows the overtaking probability between a first and a second message within a burst on one switch. The overtaking drops sharply and there is no observed overtaking beyond 3 nanoseconds delta between messages.

With improved timestamp accuracy and introduction of sub-nanoseconds on capture devices, more granular analyses could be performed. The bottom left graph shows the overtaking probability of all messages within a burst based on their interarrival time.

The base latency and latency jitter is identical for all Access Layer switches within the measurement precision (bottom right graph).

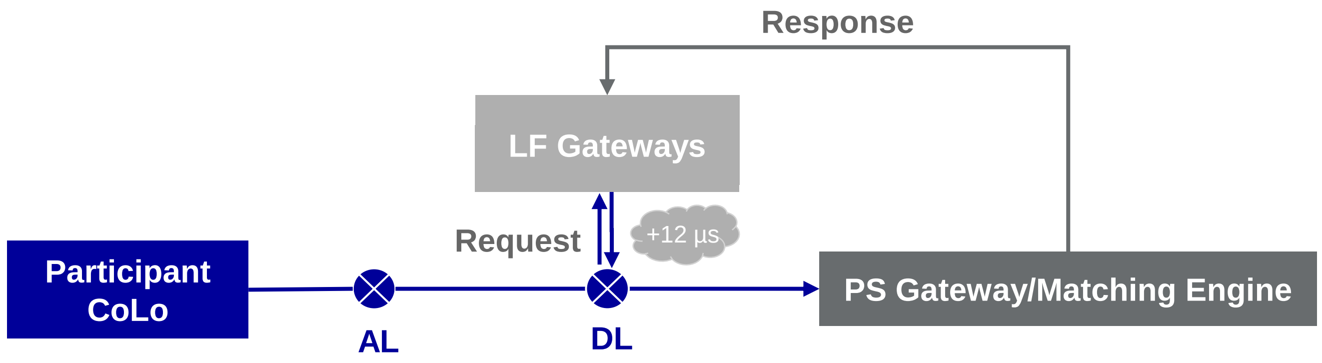
Data is taken from the Eurex A side.





PS Gateway versus LF Gateways

Latency comparison



- Using 10 Gbit/s cross connects and access via PS gateways provides the fastest way for order and quote management in T7.
- LF gateways on the other hand allow access to all partitions of a market via a single session.
- Some markets (e.g. XMAL) are using LF gateways only.
- Most markets have combined PS gateway/ Matching Engines in place for which all Matching Engine bound requests sent to LF gateways are routed via PS gateways.
- The base inbound latency for LF gateway transactions is around 12 µs higher compared to directly accessing the PS gateways. Note also that requests that have to cross sides between LF and PS gateways take another ~1 µs to reach the Matching Engine. This was significantly reduced from ~45 µs to ~1 µs end of January 2025.

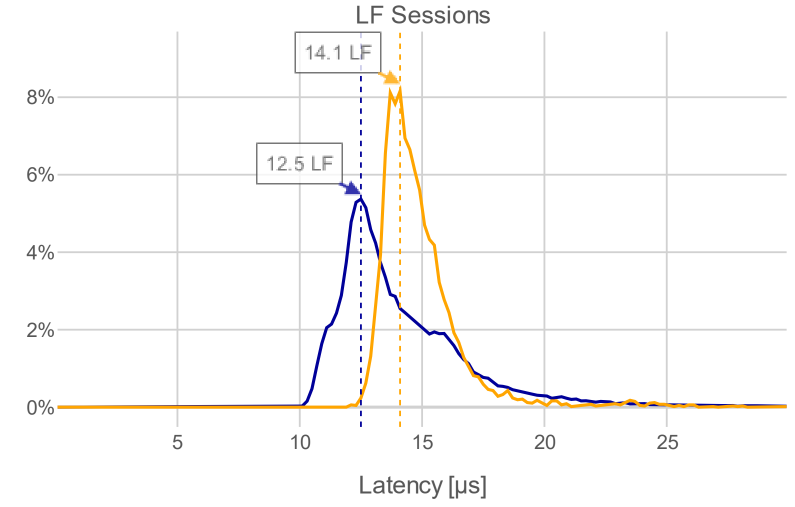
Order Entry Latency

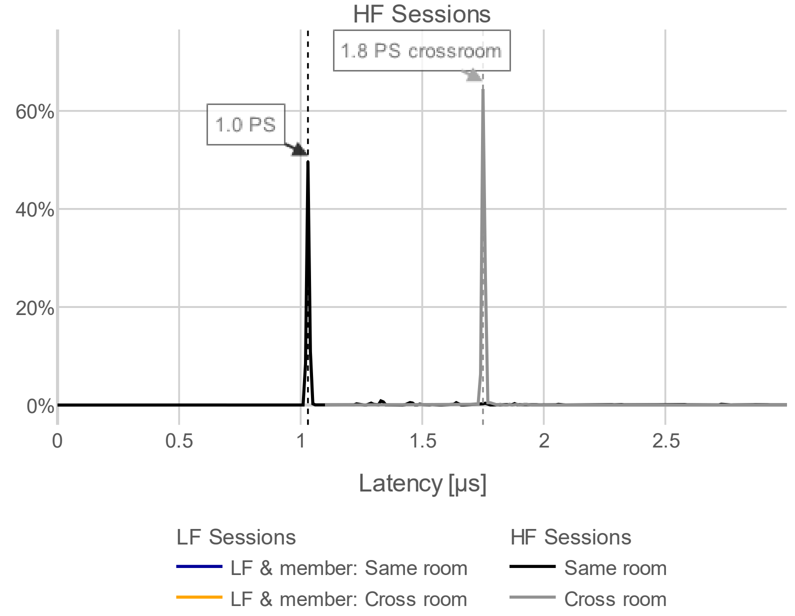
Comparison of access types

The table below shows a comparison of different access options to the T7 system. All times given are in microseconds.

Network timestamps (t_3a) are synchronized using white rabbit. The time synch quality between these timestamps is thus ~1 ns. Other T7 timestamps are subject to jitter of up to ±50 ns.

Gateway type	Same side line	Same side partition	t_3a to t_3n uncongested latency
PS	☑	☑	1
	☐	☑	2
LF	☑	☑	12
	☑	☐	13
	☐	☑	13
	☐	☐	14

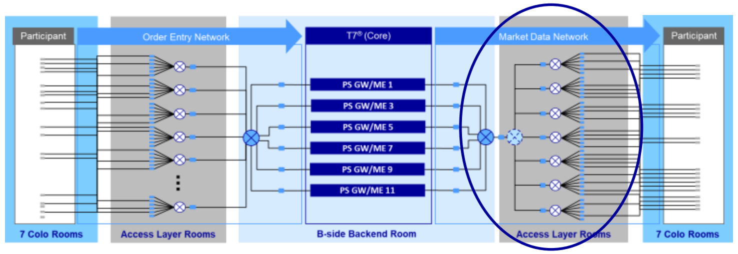




25
Market Data

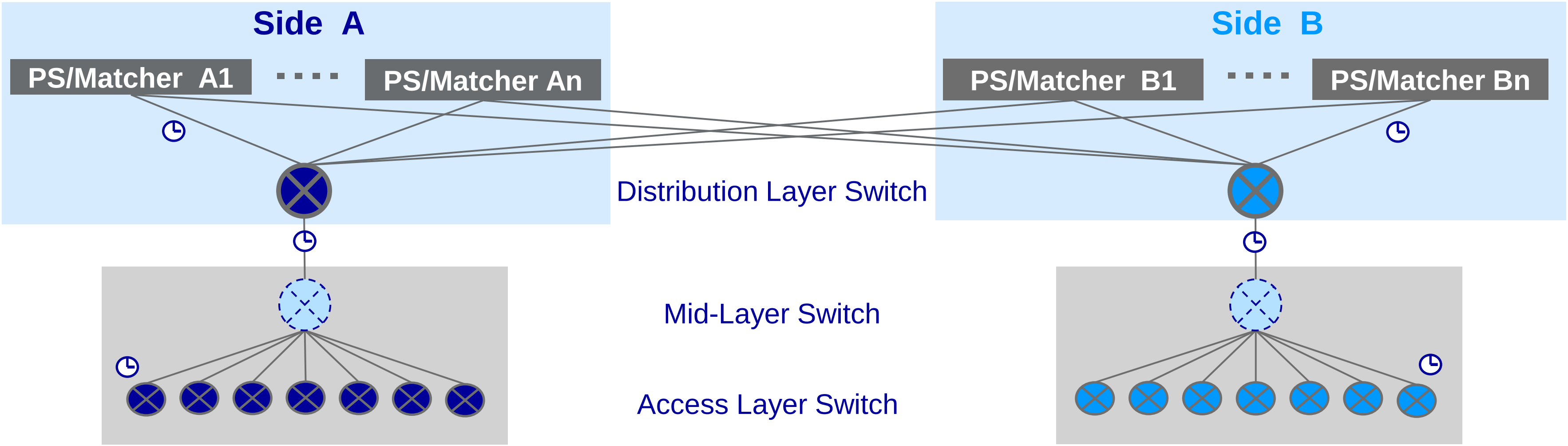
Market Data Network

Mid-Layer Switch



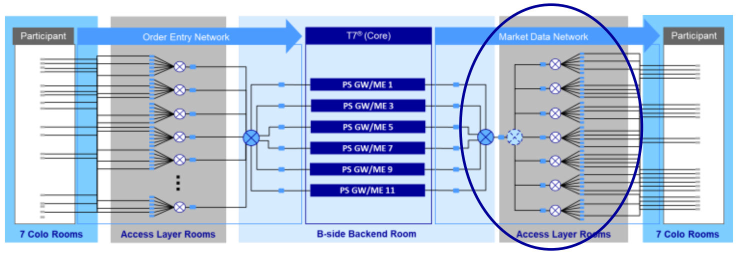
Opposite to the Order Entry network, the Market Data network has an additional Mid-Layer Switch between the Distribution Layer and Access Layer switches for Eurex and Xetra. The purpose of the Mid-Layer is to:

- Enable equal cable length to Access Layer switches, this is achieved by physical proximity of Mid-Layer and Access Layer switches.
- Serialize the connection from T7 backend to Access Layer by having just one connection between Distribution Layer and Mid-Layer.
- Make use of the positive characteristics of the Cisco 3550T to achieve better distribution of market data to the Access Layer switches.



Market Data Network

Cisco 3550T



The graphs below show the processing times measured between t_9m and t_9a based on production data.

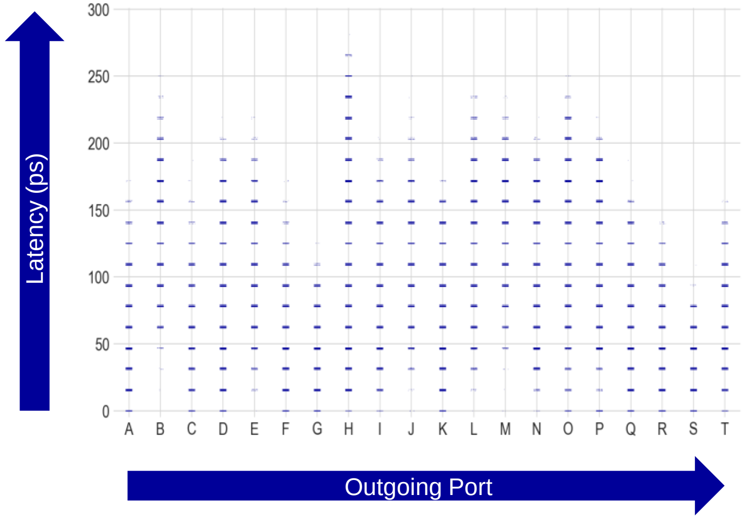
The first graph represents the per port processing time differences on one Eurex Access Layer switch normalized to the fastest port for each packet. The 15 ps steps between the clusters are due to the measurement device resolution.

The second and third graphs show the difference in per switch median processing time between two Access Layer switches, in this case for Xetra as an example. Eurex and Xetra market data switches have a similar latency profile.

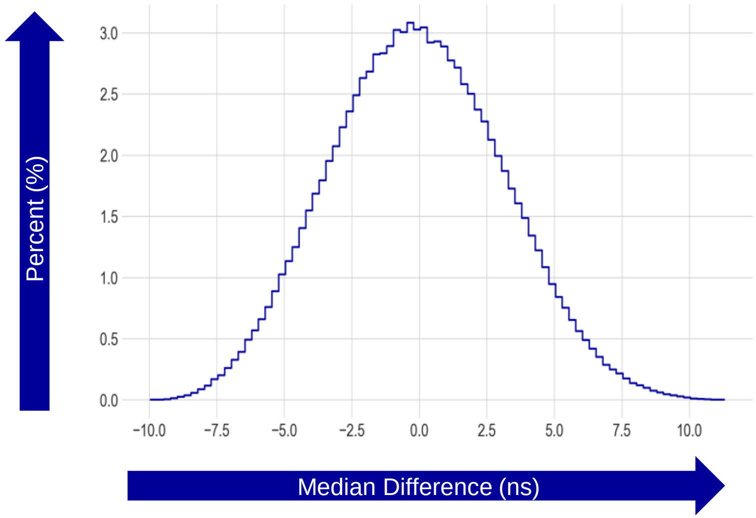
We observe the following behavior of the Cisco 3550T:

- Processing times within a single switch are consistently distributed within a range of ~300 ps across all ports after the latest firmware update.
- The median processing time differences between two switches follow a normal distribution with a higher variance than the processing time on a single switch.

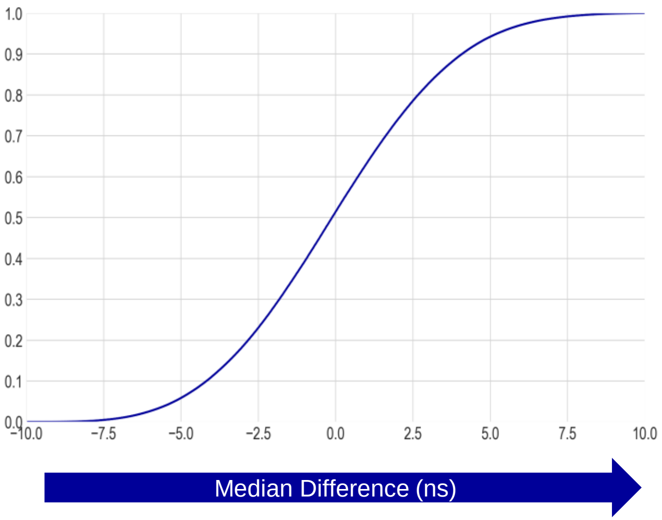
t_9m to t_9a (one 3550T)



t_9m to t_9a median difference (two 3550T)

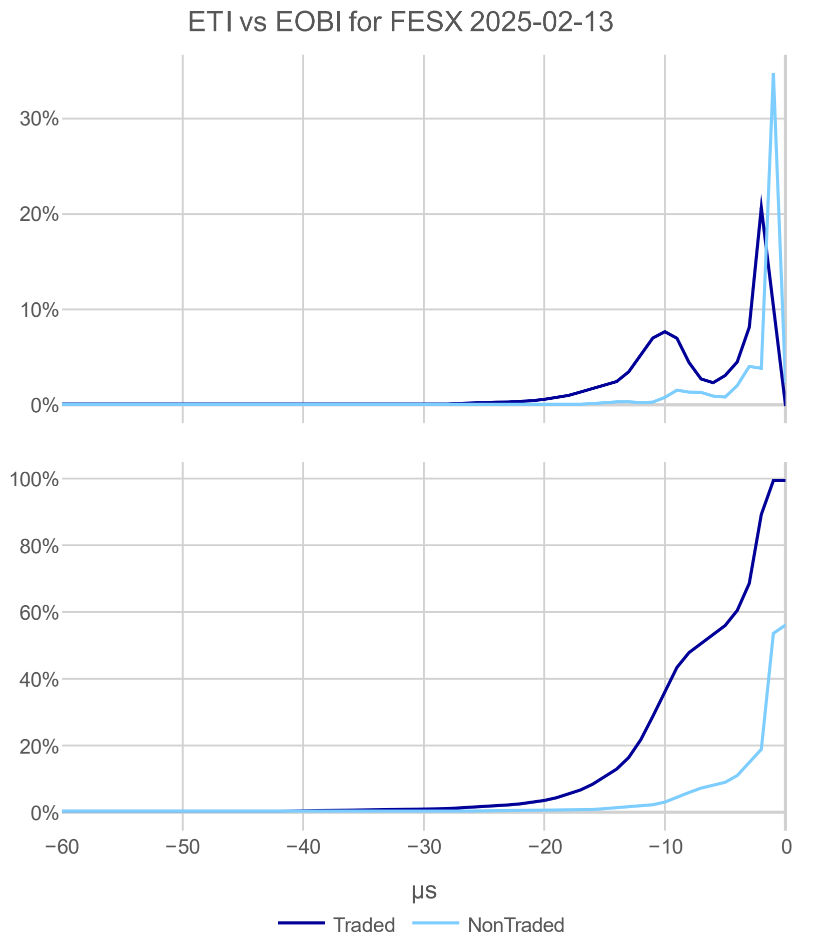


t_9m to t_9a CDF median difference (two 3550T)



Trading System Dynamics

Latency characteristics of EOBI versus ETI for Futures



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

The diagram shows the time difference distribution between public and private data in microseconds (EOBI first datagram vs ETI responses, t_9 - t_4).

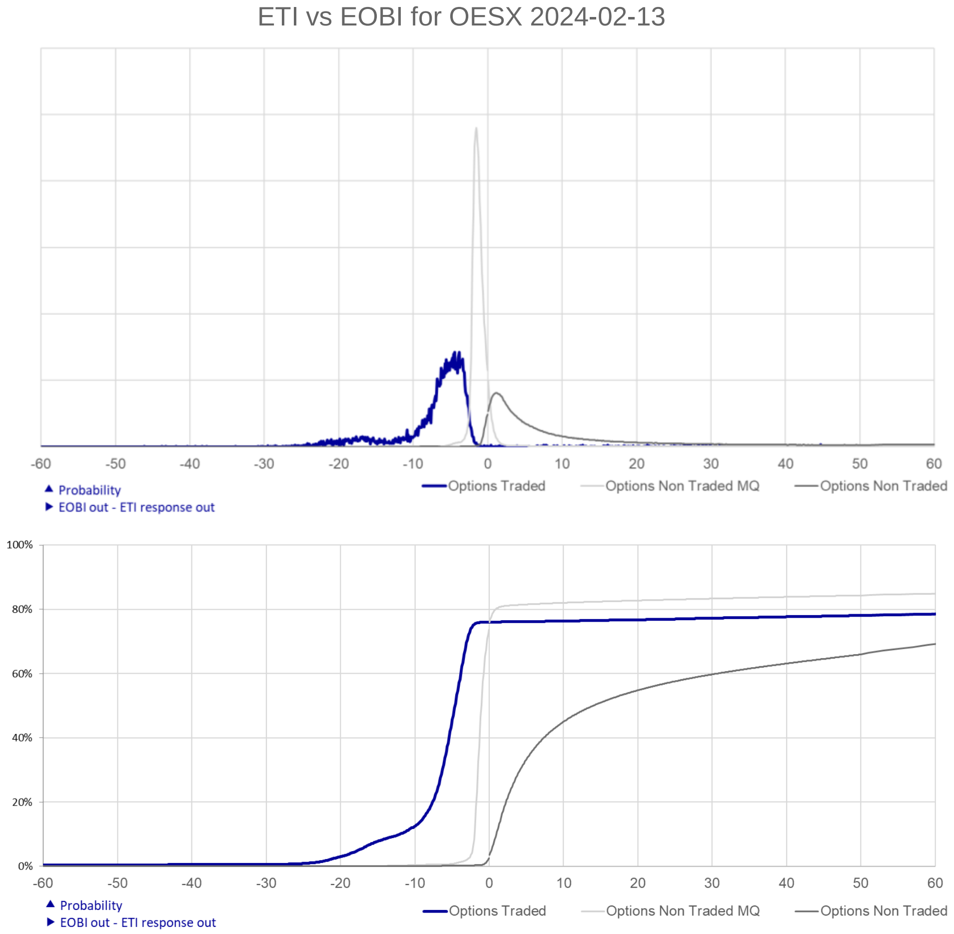
The data is a production sample from 13 February 2025.

EOBI market data is in median 1.5 µs faster than the ETI response for order book updates and 7.5 µs faster for executions.

The first EOBI datagram is faster in more than 99.4 percent of the cases compared to the ETI response, and also the first passive ETI book order notification for simple transactions.

Trading System Dynamics

Latency characteristics of EOBI versus ETI for Options



The data is a production sample from 13 February 2025 for OESX Options.

We distinguish between orders leading to a trade (quotes cannot match aggressively in OESX due to PLP), orders and single quote updates and mass quotes updates.

Trades are received first on EOBI in around 76 % of the cases with a median latency advantage of 3.5 µs.

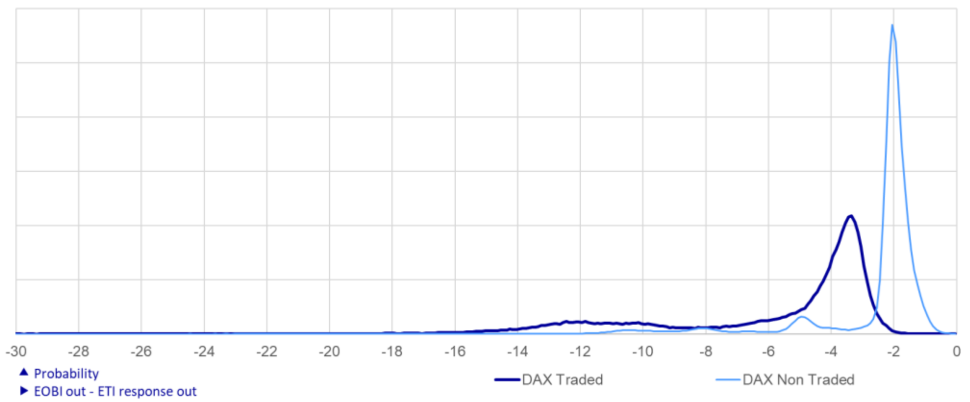
There are two main reason for EOBI delays:
The transaction is delayed by preceding messages (queues).
The transaction causes a market maker protection with many quote deletions.
Additionally, the integration of EMDI into the Matching Engine process had an impact on the EOBI processing time.

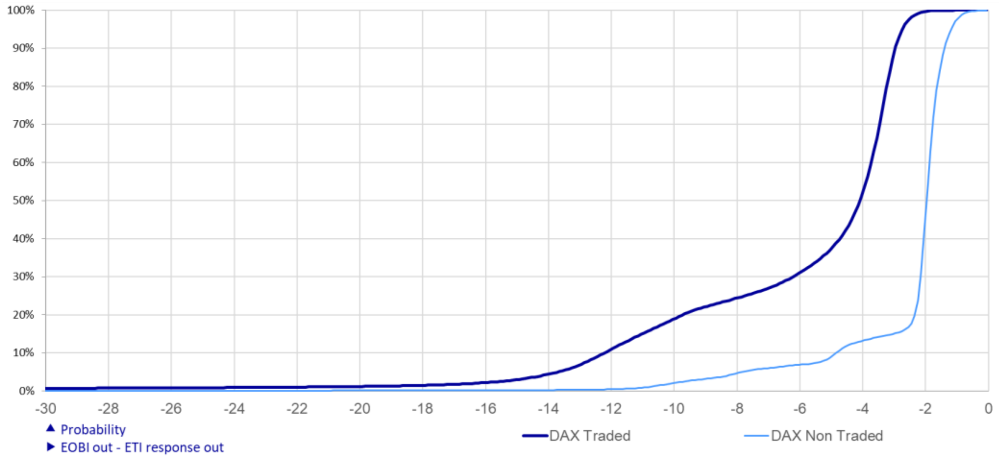
The latency profile for mass quotes is dominated by larger mass quotes, where the EOBI publisher has to broadcast each quote, leading to longer delays and queues in the EOBI path, while the ETI path only deals with a simple mass quote ack.

Trading System Dynamics

Latency characteristics of EOBI versus ETI for Xetra

ETI vs EOBI for Xetra DAX 2025-02-13





The diagram shows the time difference distribution between public and private data in microseconds for XETRA DAX products (EOBI first datagram vs ETI responses, t_9 - t_4).

The data is a production sample from 13 February 2025.

The latency distribution is similar to the Eurex futures, trades are received on EOBI in median 4 µs faster, whereas single orderbook updates are usually 2 µs faster on EOBI.

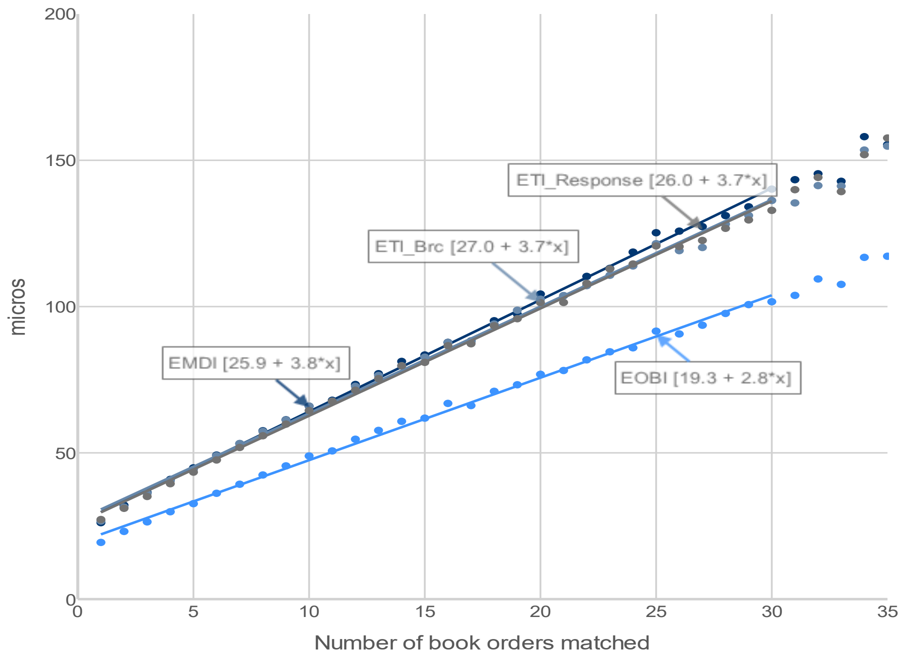
Trading System Dynamics

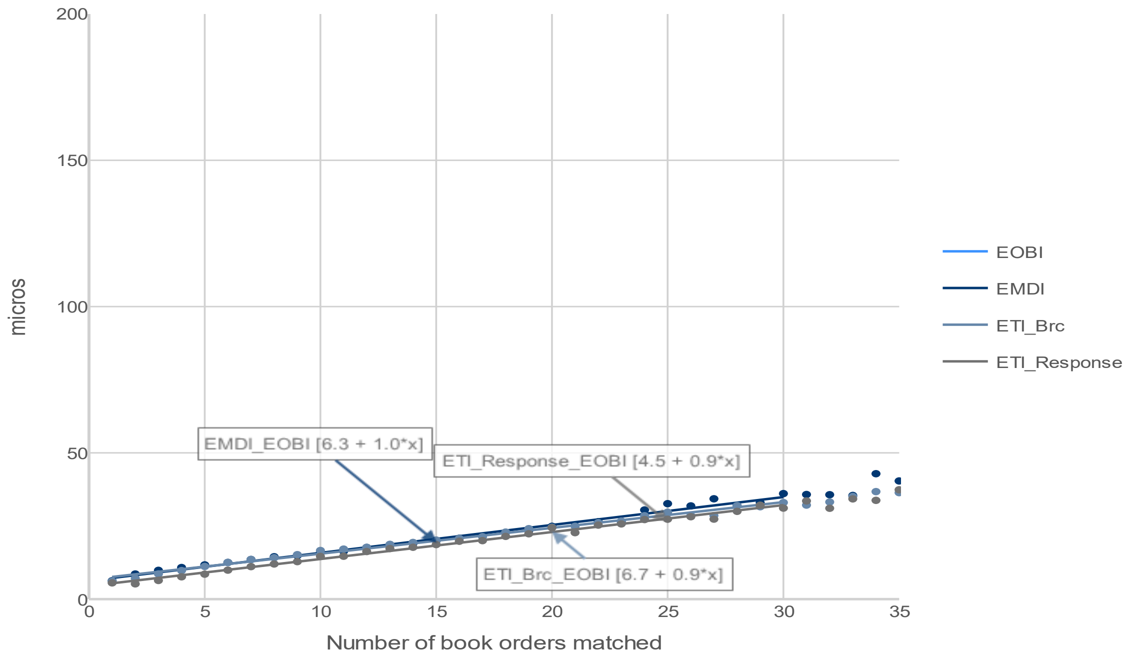
Latency characteristics of ETI versus EOBI versus EMDI

The diagrams below display the dependency of the median latency on the complexity of a trade for ETI (the right diagram is normalized to EOBI)

(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.

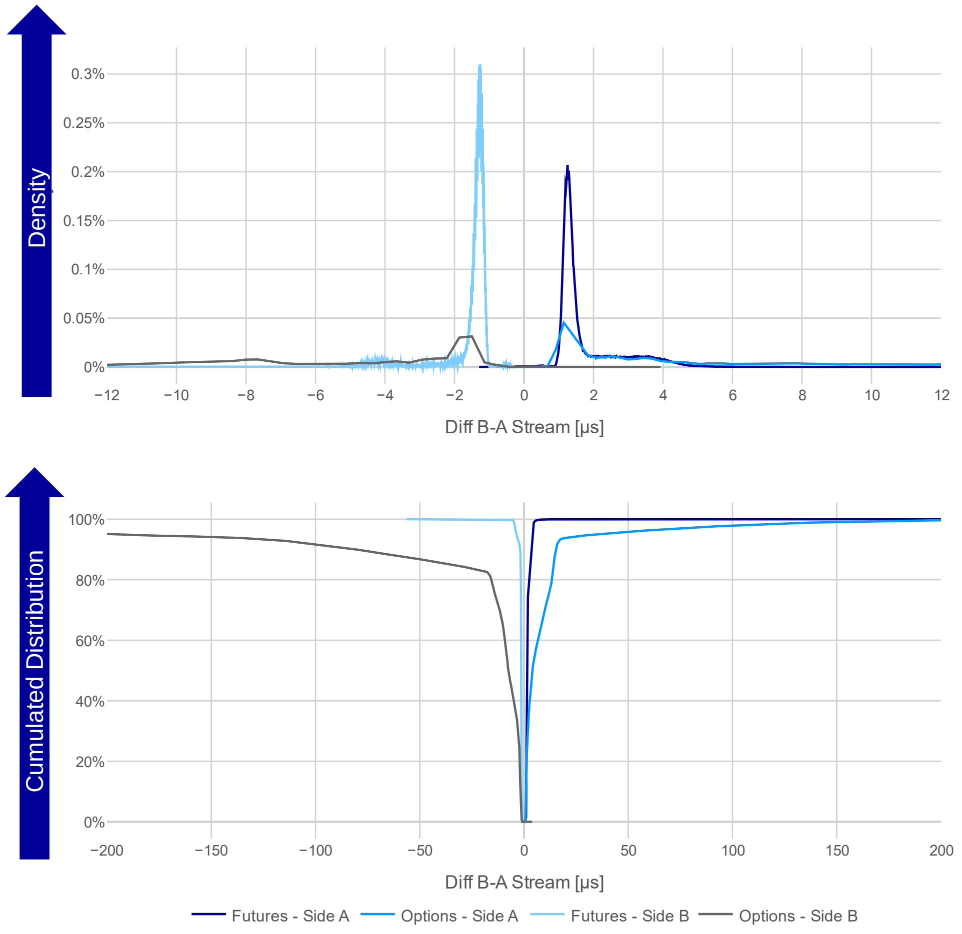
The difference between public and private data has slightly been narrowed with the integration of EMDI into Matching Engine process. However, the 'public data first' principle is still being ensured. In over 99% of all trades, we disseminate order book data on EOBI first (also for larger trades).





Trading System Dynamics

EOBI latency difference of primary and secondary feed



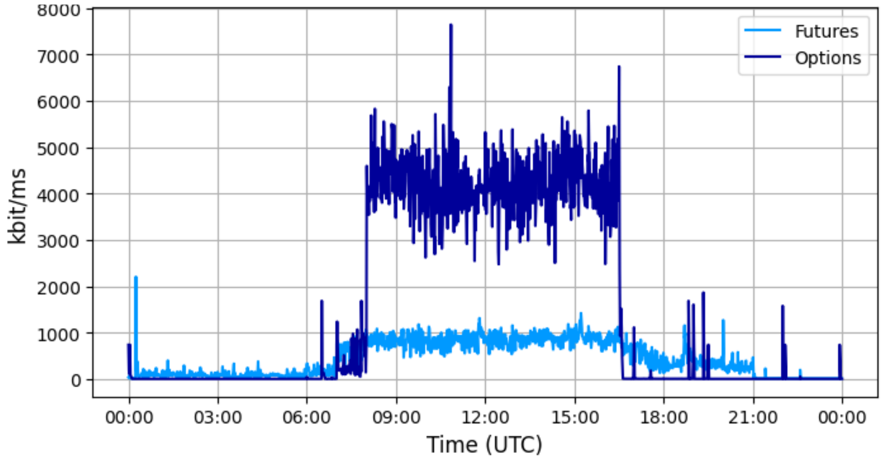
For products assigned to even partitions, market data is published first on the A and then on the B stream. For products assigned to odd partitions market data is published first on the B and then on the A stream. This also applies to a failover scenario.
The partition ID / product ID is contained in 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 contains data of exactly one product (e.g. EURO STOXX 50® Index Futures).

The data for both primary and secondary streams is disseminated by the same server using two ports connected to the two sides of the network. The process sends the data first on the primary interface. After all datagrams of a transaction are sent it starts sending on the secondary interface.

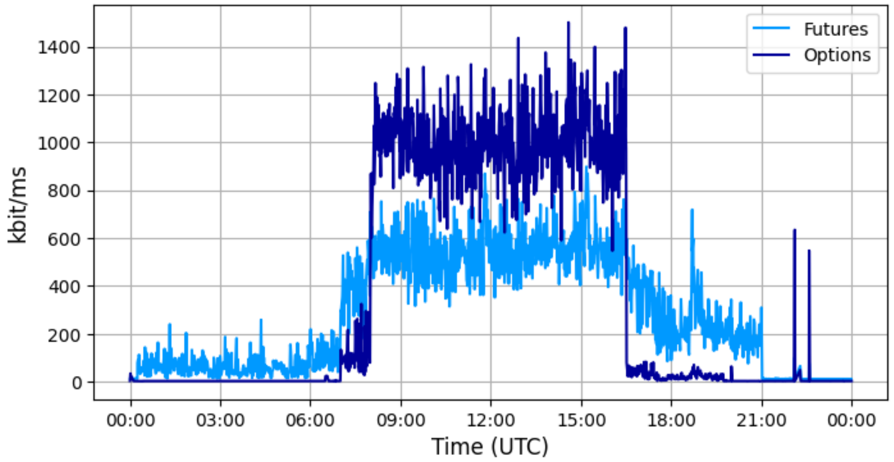
The median latency difference between the A and the B EOBI incremental feed is about 1.7 µs for futures and Cash products. For options the median is slightly higher and there are far more outliers (i.e. much slower secondary feed). The reason is that since data is published on the secondary feed only after all datagrams of a transaction are sent on the primary feed, the latency difference depends on the complexity of the transaction, i.e. a mass quote with 200 quote updates will lead to a higher delay than a single order entry.

Eurex: Market data volume

Eurex EOBI



Eurex EMDI



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

The provided data shows one data point per minute for 13 February 2025 – a busy trading day.

Enhanced Order Book Interface (EOBI) peak volume is significantly higher than price level aggregated data volume EMDI. EOBI market data is therefore currently only available to trading participants using 10 Gbit/s connections.

The EOBI for options incremental data stream peaks around 8 Gbit/s on millisecond level, while the futures stream peaks at 2.2 Gbit/s.

Participants that want to receive data for Eurex Exchange’s products on EMDI with less than 1 ms queuing delays need to use a connection with a bandwidth of more than 1.4 Gbit/s (options) or 900 Mbit/s (futures) respectively.

Trading participants are advised to use two 10 Gbit/s connections (one for each market data stream) in Co-Location to receive market data.

Xetra: Market data volume

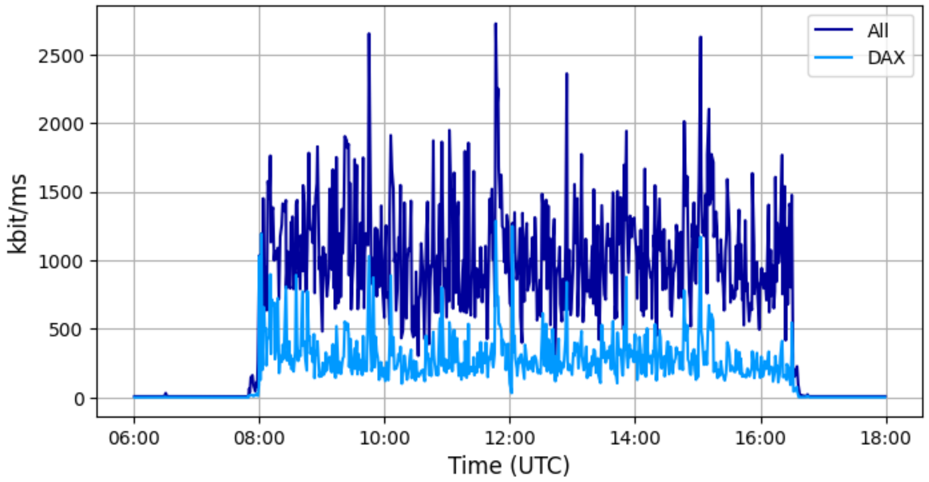
Each data point equals the maximum bandwidth produced on a 1 millisecond scale by the incremental B stream in Mbit per ms.

The provided data shows one data point per minute for 13 February 2025.

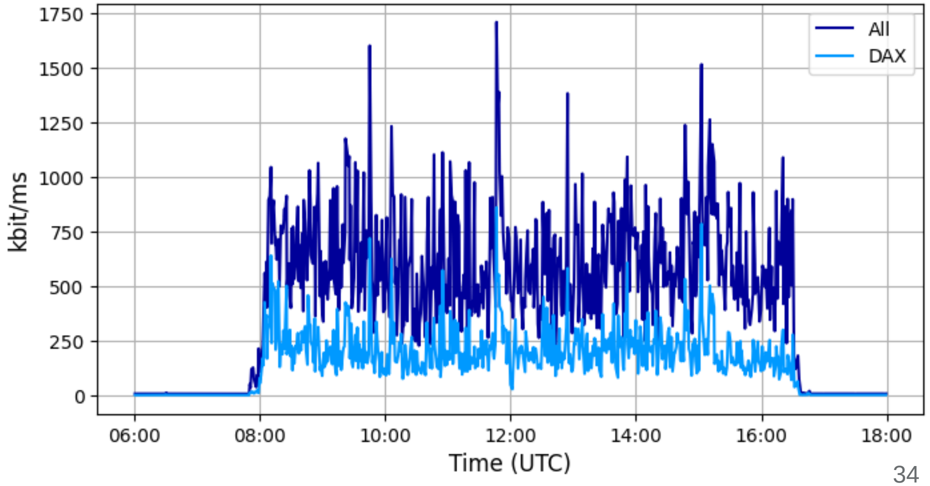
Enhanced Order Book Interface market data is currently only available to trading participants using 10 Gbit/s connections.
Trading participant are advised to take two cross connects (one for each market data stream) in Co-Location to receive market data.

Participants that want to receive EMDI data with less than 1 ms queuing delays need to use a connection with a bandwidth of more than 1.7 Gbit/s (All products) or 800 Mbit/s (DAX® equities only).

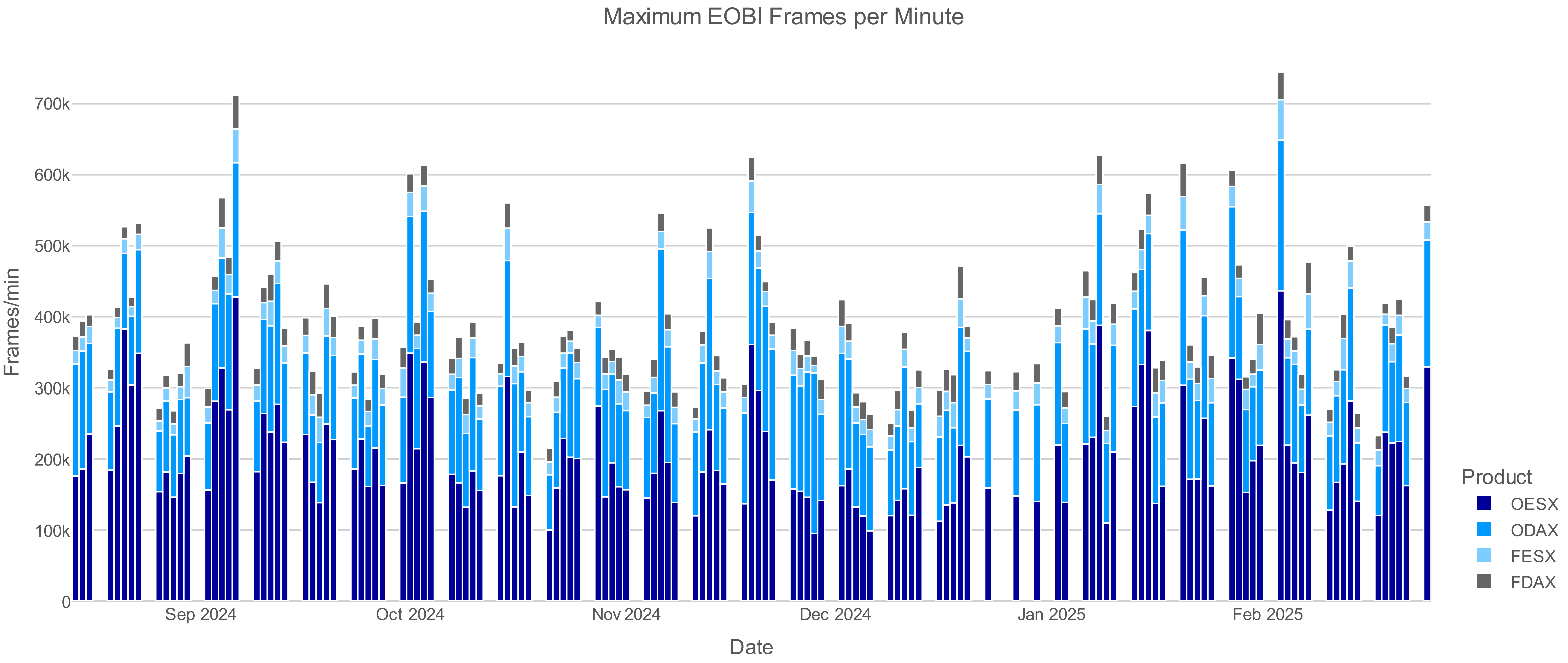
Xetra EOBI



Xetra EMDI



Maximum Frames per Minute - EOBI



36
What you need to be fast

What you need to be fast…

A few recommendations to achieve the low 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. Alternatively, use hardware solutions to connect directly without hops (e.g. FPGA).

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

Use at least two 10 Gbit/s (cross-) connections in Co-Location for EOBI market data and two 10 Gbit/s connections for T7 ETI.

Use HF sessions to connect to 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. You can connect to our white rabbit time service providing you a time synchronization quality of 1-2 ns max, see https://www.deutsche-boerse.com/dbg-en/products-services/ps-technology/ps-connectivity-services/ps-connectivity-services-time-services.

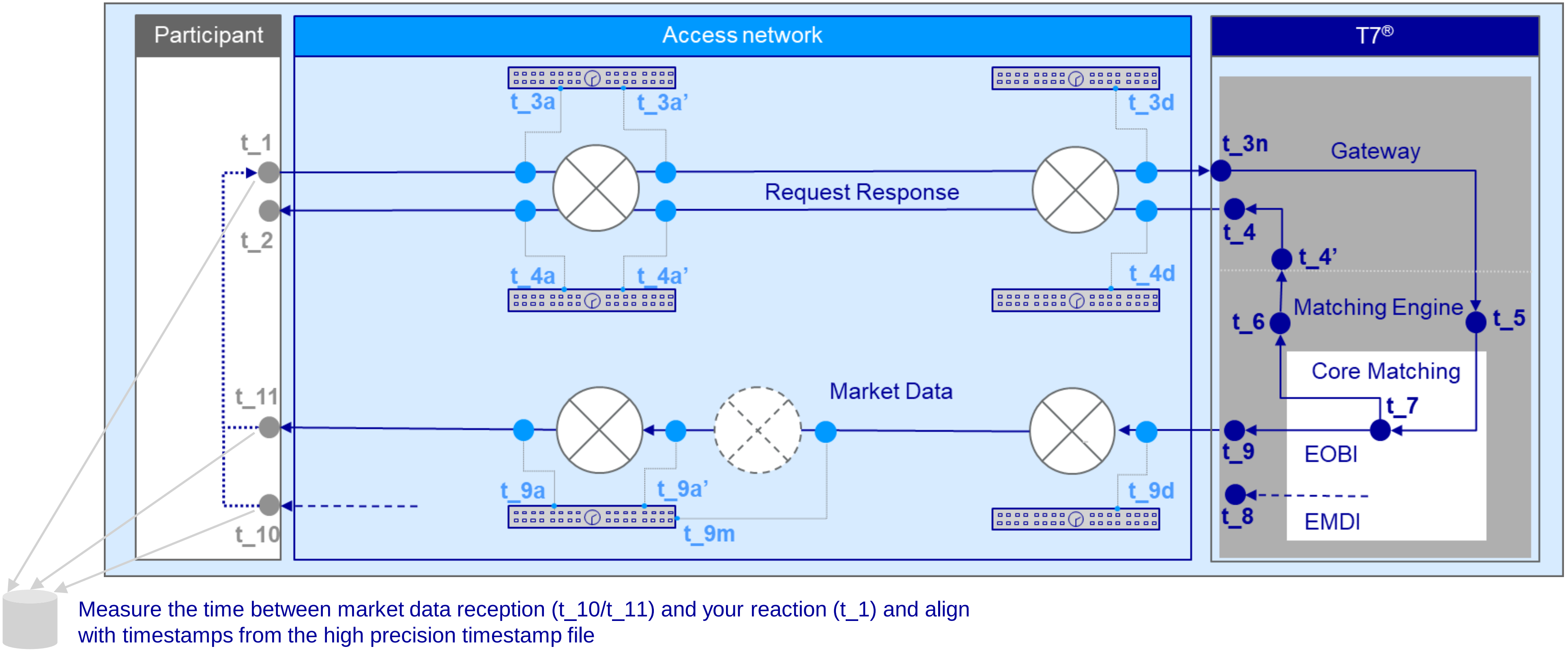
We provide highly accurate network timestamps of all orders leading to a market data update via the high precision timestamp file service, see https://www.mds.deutsche-boerse.com/mds-en/analytics/high-precision-timestamps.

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. For fastest decisions evaluate the market data classification based on the DSCP flags in the IPv4 header of EOBI market data packets.

Trade notifications need to be processed to create safety. 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.

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) and align with timestamps from the high precision timestamp file

t_3a, t_3d and t_9d are available via the high precision timestamp file service, see https://www.mds.deutsche-boerse.com/mds-en/analytics/high-precision-timestamps
Take our white rabbit signal to compare your timestamps with ours with ns accuracy
https://www.deutsche-boerse.com/dbg-en/markets-services/ps-technology/ps-7-market-technology/ps-n7/ps-connectivity-services-time-services

39

T7® Overview

T7® Architecture

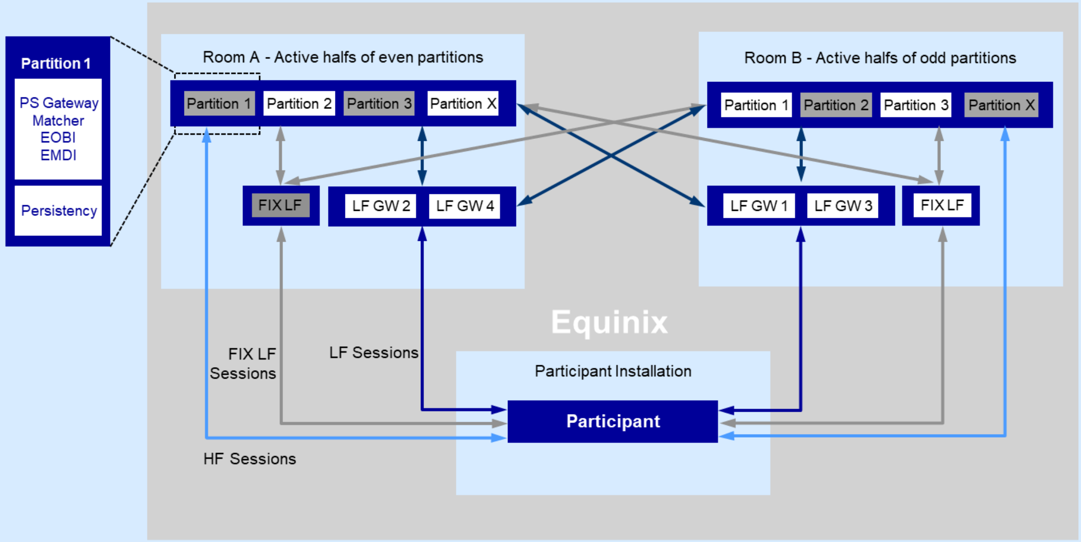
Overview

T7® architecture developed by Deutsche Börse:

- Uses state-of-the-art infrastructure and hardware for high performance
- Offers reliable connectivity and enriched trading functionality
- Is multi-market capable, operates Derivatives (Eurex, EEX) and Cash markets (Xetra, Börse Frankfurt, Vienna, etc.)

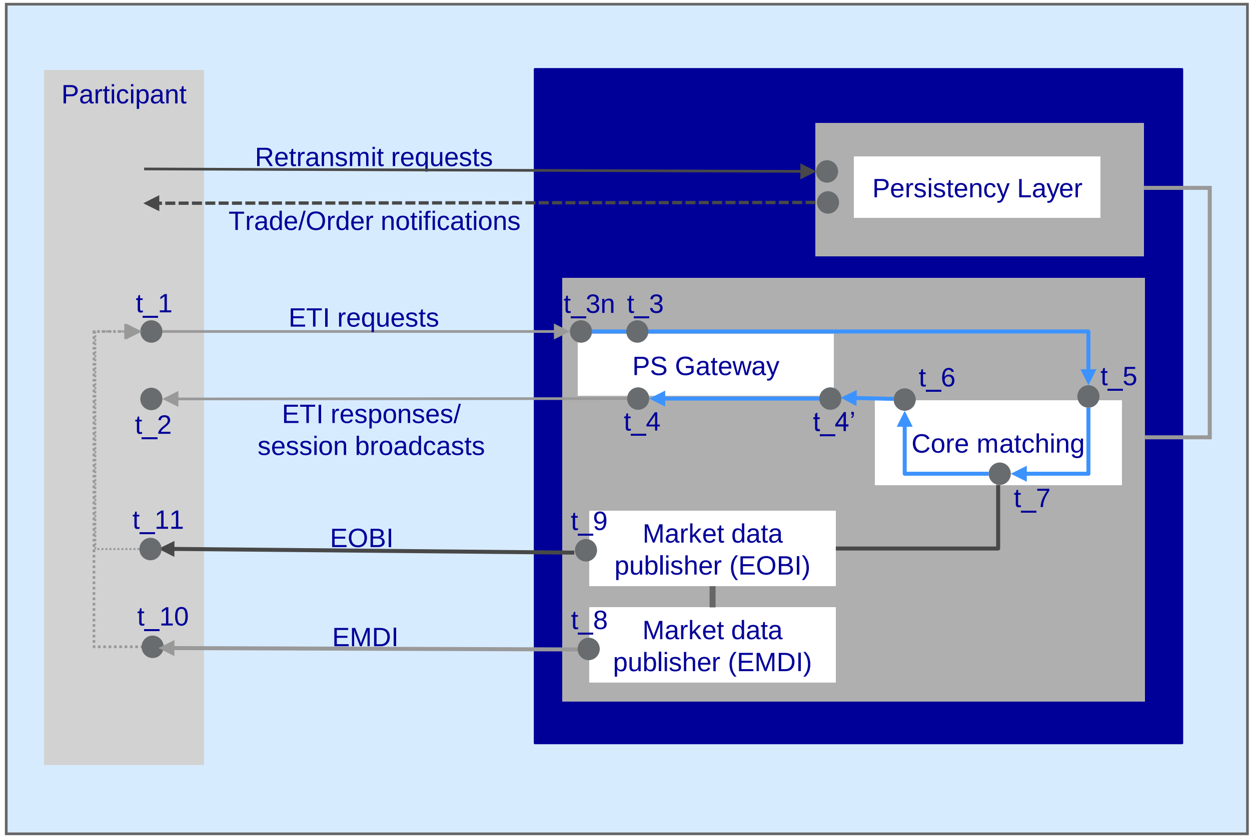
- T7® consists of partitions. A partition is a failure domain in charge of matching, persisting and producing market data for a subset of products. Each T7 partition is distributed over two rooms in the Equinix data centre.
- There are 12 Eurex T7 and 11 Xetra T7 partitions.
- Separate partitions are used for markets of other exchanges hosted on T7 (e.g. Vienna (XVIE), EEX (XEEE), Bulgaria (XBUL), …).
- The reference data contains the mapping of products to partition IDs.
- 4 LF gateways and one FIX LF gateway allow access to all Eurex partitions and the separate EEX partition.
- 4 LF gateways and one FIX LF gateway allow access to all Xetra partitions.
- 2 LF gateways and one FIX LF gateway are shared between Vienna and their partner exchanges.
- 2 LF gateways and one FIX LF gateway are shared between XBUL and XMAL

- Note that the active half of a partition is either on side A (for even partitions) or on side B (for odd partitions).
- In case of the failure of a PS gateway/Matching Engine or a market data publisher for EOBI or EMDI which is integrated into the Matching Engine, the active half of the service will shift to the other room.
- With consolidated PS gateway/Matcher process the active PS gateway and active Matching Engine act as a single failure domain within each partition, i.e. they will always fail as a single logical group.



T7® Topology

Overview



Matching Engine:

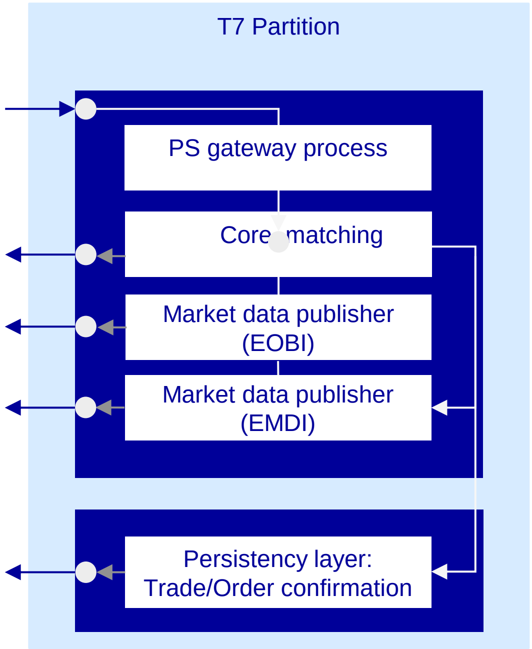
- order book maintenance & execution
- creation of direct responses as well as execution messages for passive orders/quotes
- creation of EOBI order book messages
- creation of EOBI order book snapshot messages
- creation of EMDI order book delta messages
- creation of EMDI order book snapshot messages

Persistency:

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

T7® Topology

Partitions



Orders/quotes entered for a specific product are sent by the PS gateway process to the core Matching Engine process (both residing on the same server in the same 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, followed by EMDI market data publisher and
- handover of all data resulting from the (atomic) processing of the incoming order/quote to persistency component in the partition.

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

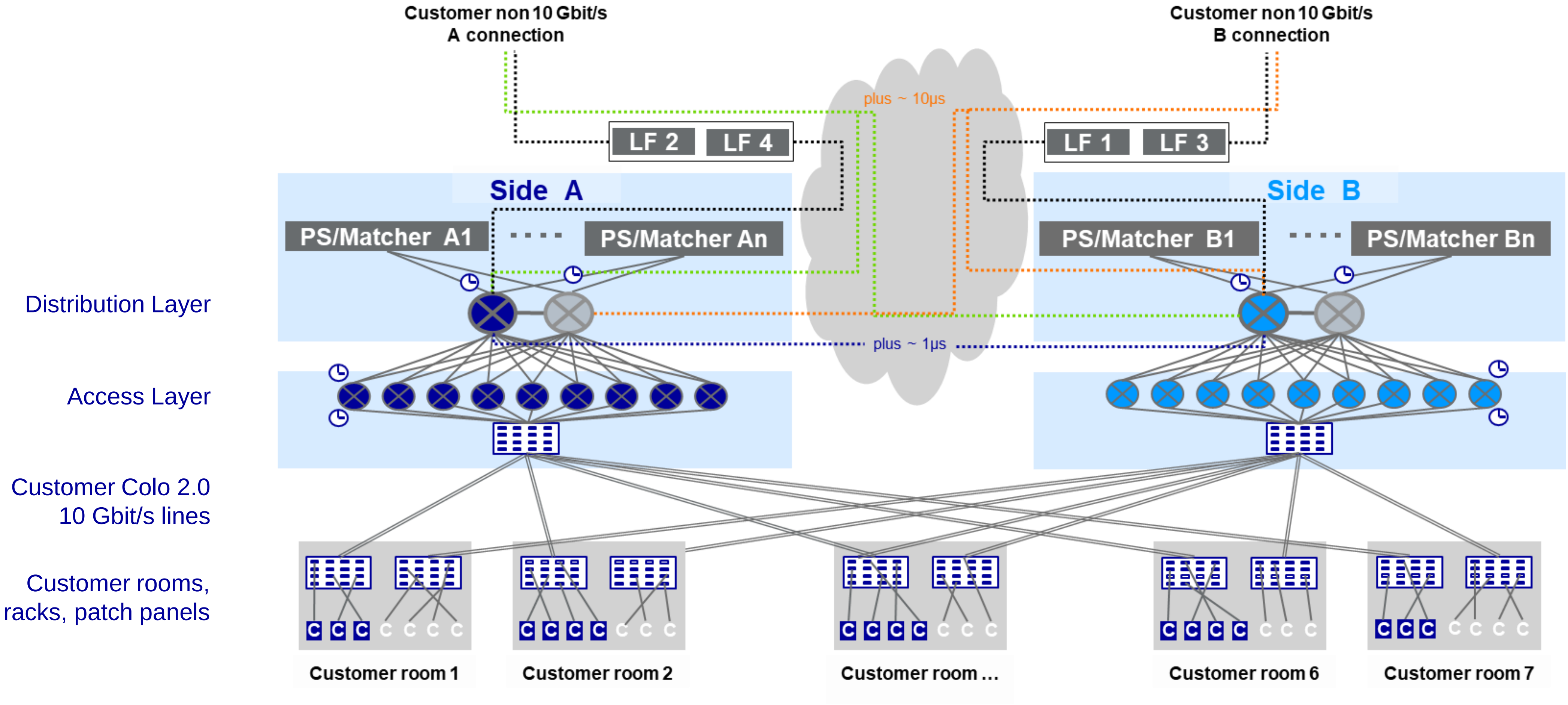
- direct response to the order/quote entered (for persistent as well as for non-persistent orders and quotes),
- fast execution information for booked orders/quotes (in case of a match).

The generation of listener broadcasts, trade confirmations (by the persistency layer) and of non-EOBI/non-EMDI market data is done by separate processes. Hence the order of the resulting messages from these processes is not strictly deterministic.

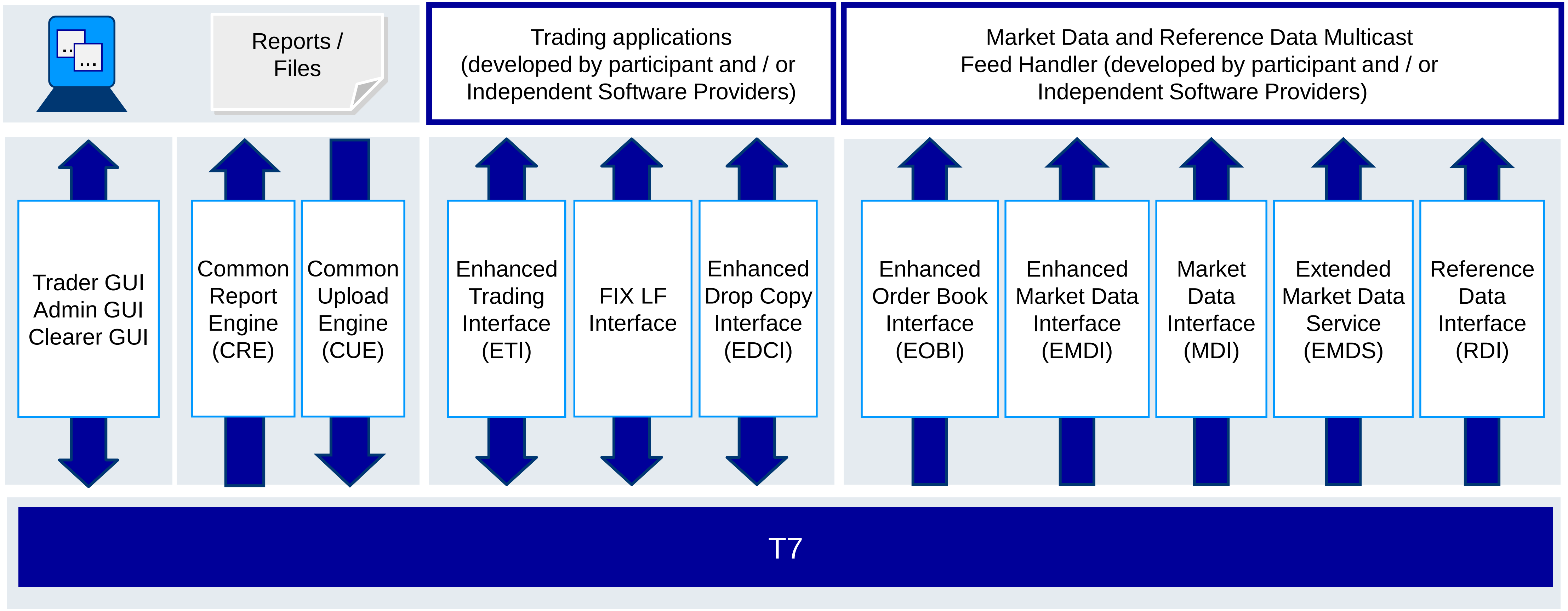
Note that the Matching Engine holds states of orders in memory. All responses, broadcasts, EOBI and EMDI market data thus are preliminary by nature.

Network Topology in Co-Location

Eurex Order Entry



T7® Trading System
Interfaces



Order Entry

Introduction

Requests sent to T7 will be routed via an access network and a gateway.

There are the following basic connection alternatives:

Choice of Network

There are two network classes connecting a participant’s installation with the T7 gateways:

- Co-Location with 10 Gbit bandwidth and a one-way base latency of around 2 µs.
- Other networks with less than 1Gbit bandwidth and a one-way base latency of minimum about 50 µs.

Choice of Session Type

T7 supports three session types:

- High frequency sessions connect to PS gateways for low latency access to a single partition (flat binary protocol: ETI). Please note that HF sessions can only be used within Equinix FR2 data center.
- Low frequency sessions connect to LF gateways for convenient access to all partitions, with a considerably higher base latency (ETI).
- FIX LF sessions connect to FIX LF gateway for convenient access to all partitions using the FIX protocol, with a considerably higher base latency than LF gateways.

Remarks

- LF gateway and FIX LF requests are routed via PS gateways.

Order Entry

Co-Location Network

Participants may use Co-Location to place their infrastructure in the datacentre that hosts the T7 system.

The Co-Location 10 Gbit network has the following properties:

Fair and equal access

Regardless of the Co-Location room we ensure all lines are created equal.

More precisely the latency between the handover point in the participant's rack and the first (Access Layer) switch is calibrated to below ±1.75 ns. The reduction from ±2.5 ns was achieved as an intermediate step and Deutsche Börse plans to further improve the cable length normalization.

Two redundant halves ('A' and 'B')

There are two independent order entry network halves.

As active gateways are placed in either half there is an optimal side for each gateway (even numbered gateways are on the A side, odd number on the B side). The only exception is FIX LF: There is only one active FIX LF gateway which is by default located on the B-side.

Crossing sides, e.g. connecting to a B side gateway via an A network, is possible but results in about 1 µs higher base latency.

Two hierarchical switch layers

Participants connect to Access Layer switches (currently 9 Eurex*, 2 Xetra per side).

The uplink of each Access Layer switch is connected to a Distribution Layer switch.

The Distribution Layer switches have a direct connection to the active gateways on the respective side.

* For capacity reasons two Access Layer switches were added to each side for Eurex on 22 February 2025.

Order Entry

Gateways

There are three gateway types to access the T7 system:

Partition-specific (PS) Gateway combined with Matching Engine

Protocol:	flat binary (ETI)
Allowed session types:	High Frequency Sessions only
Sequencing:	FIFO operation (Sequence guaranteed from network card to Matching Engine in)
Latency:	lowest, median latency ~ 12 µs network card to Matching Engine in
Versatility:	Allows routing to one partition only, only subset of broadcasts available

Low Frequency (LF) Gateway

Protocol:	flat binary (ETI)
Allowed session types:	Low Frequency Sessions only
Sequencing:	FIFO not guaranteed
Latency:	medium, (additional ~32 µs latency compared to PS gateway direct access)
Versatility:	Routes to all partitions (via PS gateway), all ETI broadcast types available

FIX LF Gateway

Protocol:	FIX
Allowed session types:	Fix Sessions only
Sequencing:	FIFO guaranteed
Latency:	high, requests to the Matching Engine are routed via PS gateways
Versatility:	Routes to all partitions (via PS gateway), all FIX broadcast types available

Market Data

Overview

Market Data can be consumed over two distinct types of networks and in various types

Choice of Network

There are two network classes available for market data:

- Co-Location with 10 Gbit bandwidth and a one-way base latency of around 2 µs.
  10 Gbit connections are equalized in length (cable latency difference of less than $\pm$1.75 ns) and provide the lowest jitter.
  The reduction from $\pm$ 2.5 ns was achieved as an intermediate step and Deutsche Börse plans to further improve the cable length normalization.
- Other networks with less than 1Gbit bandwidth with higher base latency.

Choice of Market Data Type

There are three market data types:

- Order by Order market data (EOBI) with highest granularity and lowest latency in flat binary format.
  EOBI is sent out directly from the Matching Engine and is only available via 10 Gbit network.
- Price level aggregated market data (EMDI) with slightly higher latency in FAST encoded format.
- Netted price level aggregated market data (MDI) in FAST encoded format.

Appendix

Middleware, Network, Hardware and OS Overview

T7 uses state-of-the-art infrastructure components

Intel(R) Xeon(R) Platinum 8462Y+ CPUs for all servers.

We currently use Red Hat Enterprise Linux 8.10 and plan to migrate to version 9.0 mid 2025.
T7 internal communication between its core components is based on Ethernet.

T7 network access
Deutsche Börse offers trading participants to connect via 10 Gbit/s cross connects to its T7 platform in the Equinix data centre.

The Co-Location offering uses Cisco Nexus 3548-XL switches for Order Entry and is operating in cut-through mode. For market data Cisco Nexus 3550-T are used.
All cables are normalized to give an overall maximum deviation between any two cross connects of less than ±0.35 m (±1.75 ns).
The reduction from ±0.5 m was achieved as an intermediate step and Deutsche Börse plans to further improve the cable length normalization.

Insight into network dynamics is offered by the High Precision Timestamp File service (see https://www.mds.deutsche-boerse.com/mds-en/analytics/high-precision-timestamps).

Participant facing interface cards on the gateways and market data publishers use Solarflare EnterpriseOnload wire order delivery API to bypass the kernel TCP stack and deliver messages in the same order received by the network card.

Cables connecting Line-of-Sight antenna cables have been equalized to ±1m by Equinix.

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.

ETI LF sessions are available with throttle values of 150 or 50 transactions/sec ETI HF sessions are available with throttle values of 250*, 150 or 50 transactions/sec. Furthermore, LF sessions that cannot enter orders/quotes but can only receive trade and listener broadcasts are available (at a reduced price).

The disconnect limit is set at:

- 750 for HF Ultra sessions with a throttle value of 250 transactions/sec. i.e. a session will be disconnected in case of more than 750 consecutive rejects due to exceeding the transaction limit (throttle).
- 450 for sessions with a throttle value of 150 transactions/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 transactions/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 recovery facility is used, all ETI sessions will have a throttle limit of 30 transactions 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 600 Eurex sessions and 200 Xetra sessions per participant can be ordered.

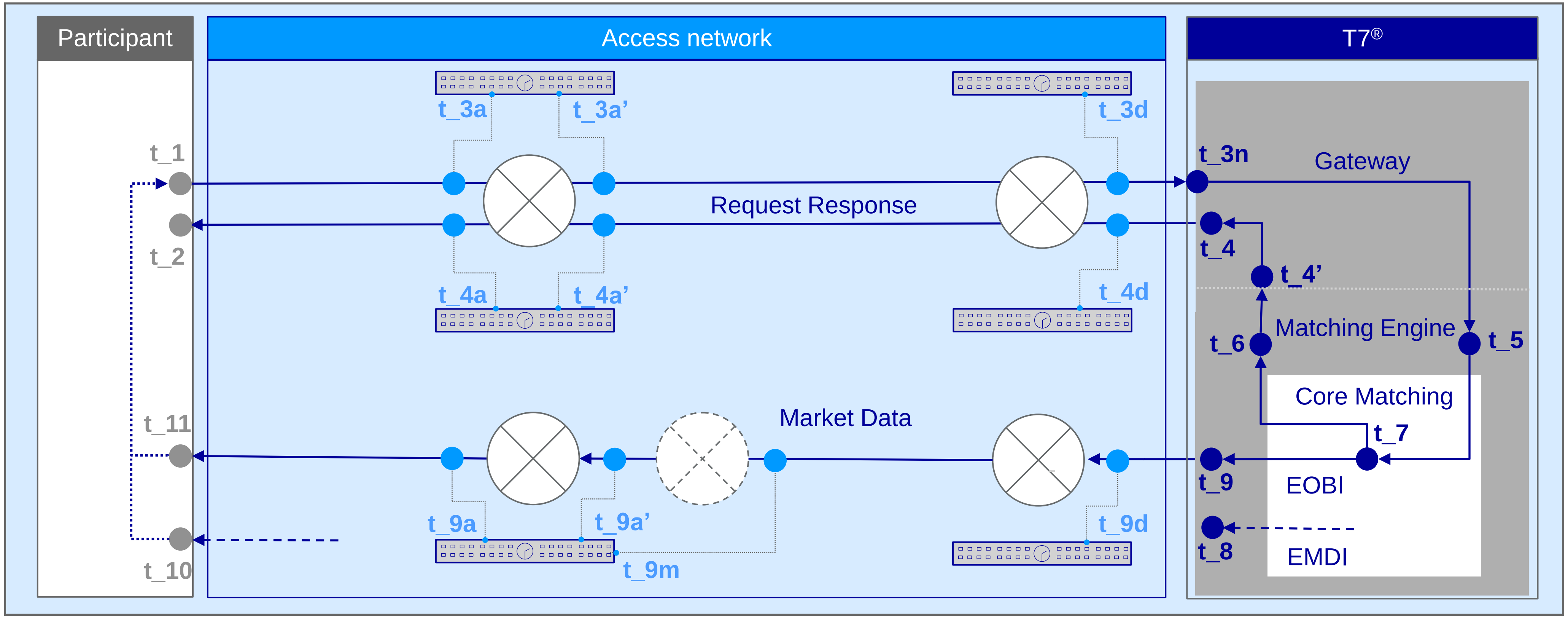
There is also a limit on the maximum number of sessions that can connect to a PS gateway concurrently per participant. This limit is currently configured to 80 sessions, see Eurex Circular 122/17.

On 1 July 2019, we introduced a limit on the maximum number of outstanding session and trader login requests possible per business unit and per session at any given point in time. This limit is set to 50 on business unit level, 10 on session level. We recommend a synchronous login procedure, where a login request is sent on a session only after the previous login has been responded to. Please refer to the Incident Handling Guide for details.

For Eurex the number of order entry cross connects in colocation that may be used concurrently on a single day is limited to 6 per Access Layer switch. In addition, the number of allowed ethernet frames per cross connect is currently limited to 30.000 per second and 600.000 per minute for Eurex and 25,000 per second and 600,000 per minute for Xetra.

T7® Topology

Timestamps



- Timestamps provided in T7 API (in real time) in dark blue (t_3n: taken by network card, other: application level)
- Network timestamps taken using TAPs and timestamping switches (Metamako)
- Timestamps possibly taken by participants shown in grey

T7® Timestamps

Description

t_[x]a, t_[x]a’	time taken by network capture devices in the Access Layers.
t_[x]d	time taken by network capture devices in the Distribution Layers.
t_9m	time taken by network capture device between Access Layer and Mid-layer.
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 network.
t_3n	time taken by the PS gateway when the first bit of a request arrives on the PS gateway NIC; contained in (private) ETI response for PS gateway enabled partitions. Consecutive messages via the same session may be assigned to the same t_3n.
t_3	time 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 for non-PS gateway enabled partitions (e.g. XVIE).
t_4’	time 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_4	time 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_5, t_6	time 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 starts maintaining 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.

T7® Timestamp Reference

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

Timestamp	Tag no.	Field name	Present in
t_3, t_3n	5979	RequestTime	ETI Response EMDI Depth Incremental message, in case a trade is reported EOBI Execution Summary, Order Add, Order Modify, Order Modify Same Priority and Order Delete messages
t_4'	7765 25043	ResponseIn NotificationIn	ETI Response (from the Matching Engine) ETI Notification (from the Matching Engine)
t_4	52	SendingTime	ETI Response and Notification
t_5	21002 2445	TrdRegTSTimeIn AggressorTime	ETI Response (from the Matching Engine) EMDI Depth Incremental message, in case a trade is reported EOBI Execution Summary message
t_6	21003	TrdRegTSTimeOut	ETI Response and Notification (from the Matching Engine)
t_7	17 273 60 21008	ExecID MDEntryTime TransactTime TrdRegTSTimePriority	ETI Response (from the Matching Engine) EOBI Execution Summary message EMDI Depth Incremental, Depth Snapshot and Top of Book Implied message EMDI messages for other events EOBI Order Modify Same Priority and Order Delete messages EOBI Order Add and Order Modify messages
t_8	No tag	SendingTime	T7 EMDI UDP packet header
t_9	60	TransactTime	EOBI packet header
t_8 - t_5	No tag	PerformanceIndicator	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_[x]a, t_[x]a', and t_[x]d, t_[x]d') are not available in any protocol field but some via the High Precision Timestamp File service.

Thank you for your attention

Sergej Teverovski, Manfred Sand, Phuong Hieke, Simon Braun, Raphael Colcombet
Trading System Analytics
Deutsche Börse AG
Mergenthalerallee 61
65760 Eschborn
Germany

E-mail: monitoring@deutsche-boerse.com

For updates refer to
https://www.eurex.com/ex-en/support/technology/t7 and http://www.xetra.com/xetra-en/technology/t7/publications

DEUTSCHE BÖRSE
GROUP

Disclaimer

This content is for informational purposes only. None of the information constitutes investment advice or an offer to sell or the solicitation of an offer to buy any contract, share or other financial instrument. This content is only for use as general information and all descriptions, examples and calculations contained are for illustrative purposes only.

Deutsche Börse AG, Frankfurter Wertpapierbörse (FWB®, the Frankfurt Stock Exchange), Eurex Frankfurt AG, Eurex Deutschland and Eurex Clearing AG do not represent that the information in this publication is comprehensive, complete or accurate and exclude liability for any consequence resulting from acting upon the contents of this or another publication, in so far as no willful violation of obligations took place or, as the case may be, no injury to life, health or body arises or claims resulting from the Product Liability Act are affected.

Securities traded on the Frankfurt Stock Exchange and Eurex derivatives (other than EURO STOXX 50® Index Futures contracts, EURO STOXX® Select Dividend 30 Index Futures contracts, STOXX® Europe 50 Index Futures contracts, STOXX® Europe 600 Index Futures contracts, STOXX® Europe Large/Mid/Small 200 Index Futures contracts, EURO STOXX® Banks Sector Futures contracts, STOXX® Europe 600 Banks/Industrial Goods & Services/Insurance/Media/Personal & Household Goods/Travel & Leisure/Utilities Futures contracts, Dow Jones Global Titans 50 IndexSM Futures contracts, DAX® Futures contracts, MDAX® Futures contracts, TecDAX® Futures contracts, SMIM® Futures contracts, SLI Swiss Leader Index® Futures contracts, RDXxt® USD - RDX Extended Index Futures contracts, Eurex inflation/commodity/weather/property and interest rate derivatives) are currently not available for offer, sale or trading in the United States nor may they be offered, sold or traded by persons to whom US tax laws apply.

The fund shares listed in XTF Exchange Traded Funds® are admitted for trading on the Frankfurt Stock Exchange. Users of this information service who legally reside outside Germany are herewith advised that sale of the fund shares listed in XTF Exchange Traded Funds may not be permitted in their country of residence. The user makes use of the information at their own risk.

Legal validity of this disclaimer

In the event that individual parts of or formulations contained in this text are not, or are no longer, legally valid (either in whole or in part), the content and validity of the remaining parts of the document are not affected.

Trademarks

The names and designations listed under www.deutsche-boerse.com/disclaimer are registered trademarks of Deutsche Börse AG or an affiliate of Deutsche Börse Group.

The names and trademarks listed there do not represent a complete list and, as well as all other trademarks and protected rights mentioned in this publication, are subject unreservedly to the applicable trademark law in each case and are not permitted to be used without the express permission of the registered owner. The simple fact that this publication mentions them does not imply that trademarks are not protected by the rights of third parties.

Information on trademarks of Stoxx Ltd. and ISS STOXX Index GmbH are available under Trademarks STOXX - STOXX

EEX® is a registered trademark of European Energy Exchange AG.

TRADEGATE® is a registered trademark of Tradegate AG Wertpapierhandelsbank.

Flexible is better.® is a registered trademark of Axioma, Inc.