DEUTSCHE BÖRSE GROUP

Open Day 2019

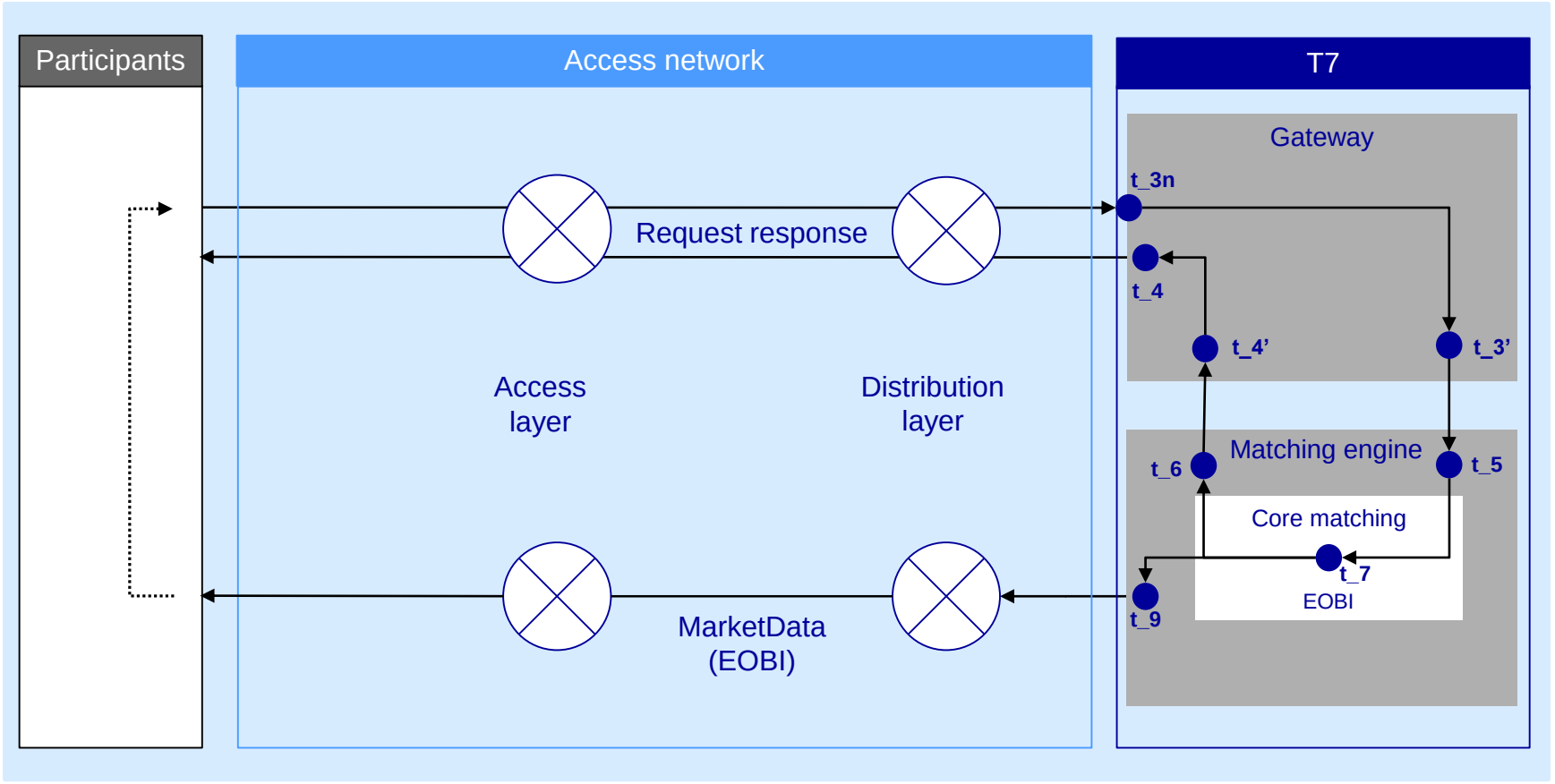
T7® infrastructure and latency

Andreas Lohr

26 September 2019

Contents

T7® Topology



● Timestamps provided in T7 API (in real time) in dark blue (t_3n: taken by network card, other: application level)

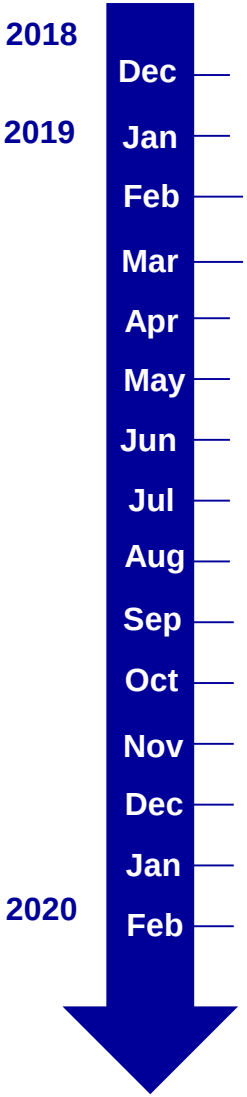
⊗ Cisco 3548X switches operating in cut-through mode.

3

Developments since last Open Day

Developments since last Open Day

Timeline



- Dec: T7® Release 7.0
- Feb: Launch of EOBI for Options
- Mar: Launch of High Precision Timestamps “All” for Futures
- Apr: Hardware refresh LF gateways
- May: T7® Release 7.1
- Jun: Launch of Passive Liquidity Protection (PLP) Pilot Phase
- Aug: Launch of High Precision Timestamps “All” for Options
- Sep: Deutsche Börse Open Day 2019
- Nov: T7® Release 8.0
- Jan: Activation of LF tunnelling

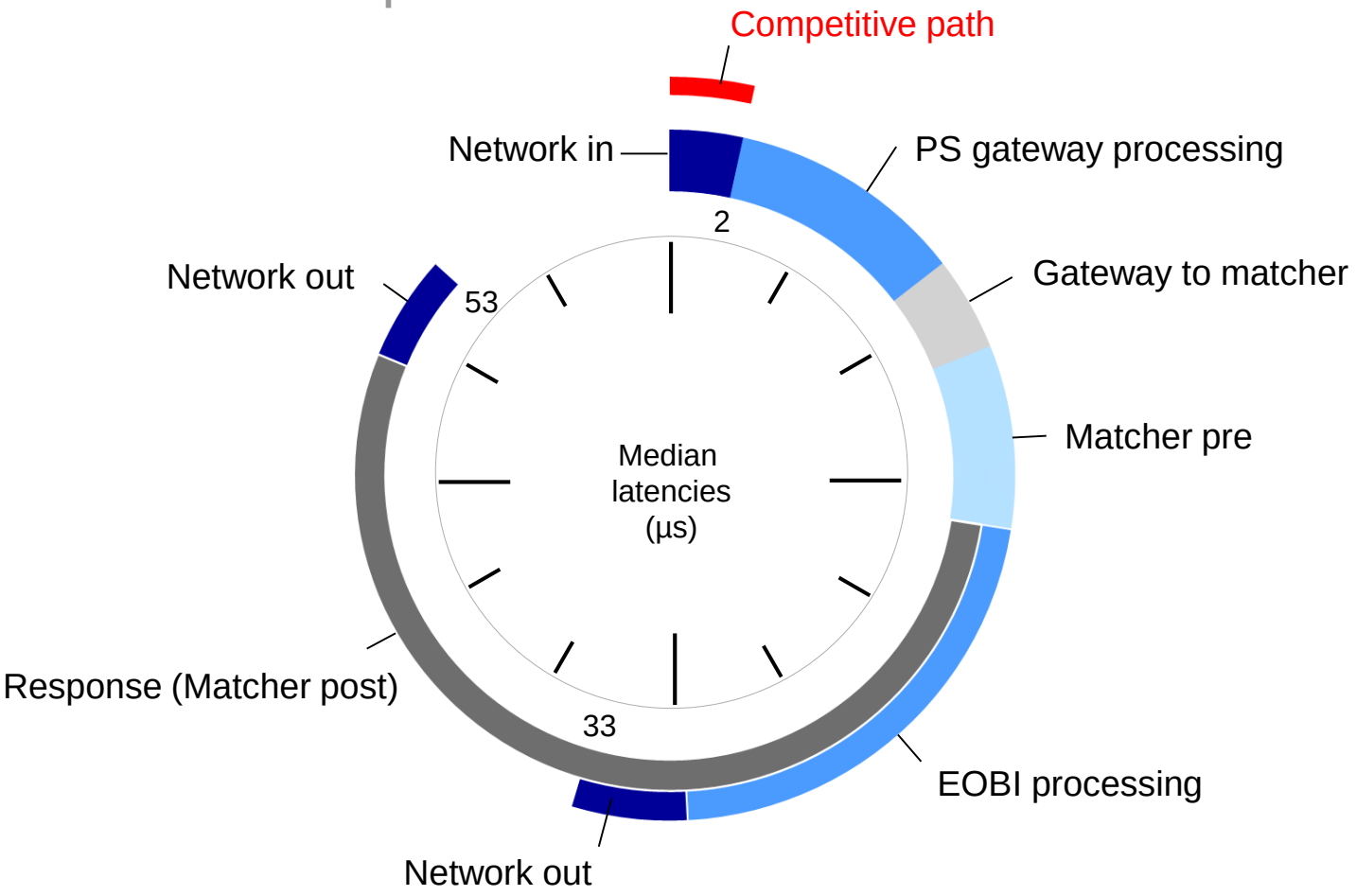
Developments since last Open Day

All time highs in processed messages – stable latencies

Type	Date	Value
Daily Transactions (T7®)	15-Aug-2019	322 Million
Daily Transactions (Eurex)	15-Aug-2019	133 Million
Latency (t_3n to t_4 in µs)	Aug-2019	All / Opt HF / Fut HF Average: 337 / 427 / 119 Median: 56 / 54 / 53 Minimum: 20 / 22 / 22
Market Data (EOBI) Latency (t_3n to t_9, µs)	29-Aug-2019	All / Options / Futures Average: 229 / 595 / 117 Median: 42 / 43 / 41 Minimum: 20 / 20 / 20
Percentage of orderbook updates published first via public market data (EOBI)	29-Aug-2019	All / Agency / Trades Futures 99.7 / >99.9 / 99.9 Options 79.0 / 98.6 / 92.3
Max Matcher Input Rate per second (T7®)	29-Aug-2019	97 kHz
Max Matcher Input Rate per second (Eurex)	03-Sep-2019	50 kHz
Max Matcher Input Rate per second (Eurex partition)	12-Dec-2018	11 kHz (p 6)
Max Matcher Input Rate per second (Xetra partition)	22-May-2019	12 kHz (p 59)
Max Matcher Input Rate per second (single product)	17-Jun-2015	9 kHz (FESX)

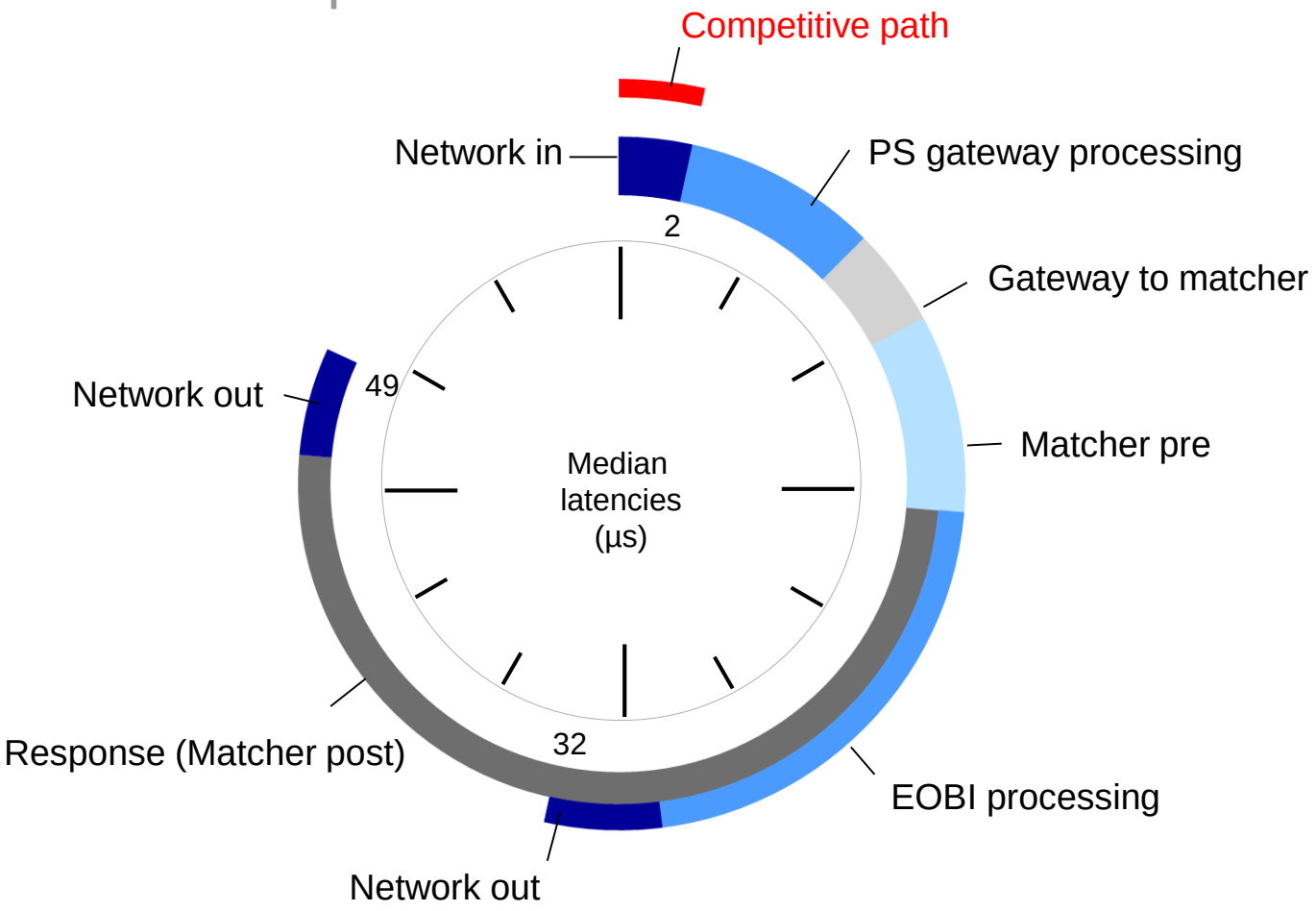
Developments since last Open Day

The micro clock – September 2018



Developments since last Open Day

The micro clock – September 2019



Gateway processing reduced by 2 µs inbound and 2 µs outbound

Public first principle untouched

Developments since last Open Day

T7® latency composition

The charts below show a comparison of latencies for Eurex futures sent via PS gateways.

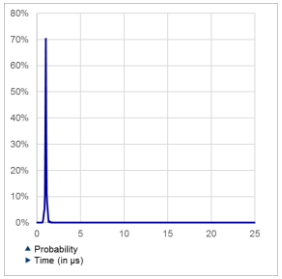
Dark blue are recent figures, light blue are from August 2018.

Most noticeable is an increase in matcher processing times due to addition of pre trade risk functionality.

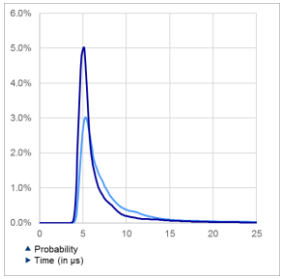
This was more than compensated by a reduction of gateway response processing times.

Request and market data (EOBI)

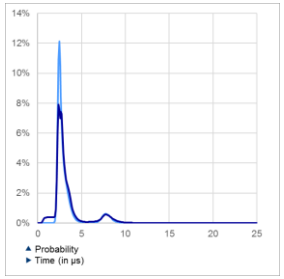
Network request (t_3a to t_3n)



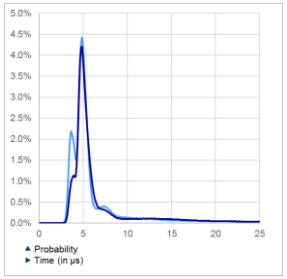
Gateway request (t_3n to t_3')



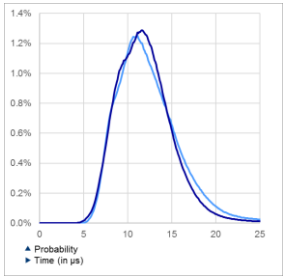
Gateway to matcher (t_3' to t_5)



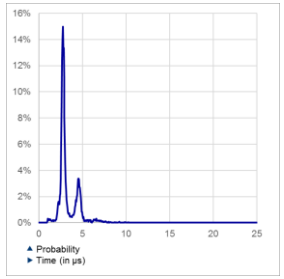
Matcher pre (t_5 to t_7)



EOBI processing (t_7 to t_9)

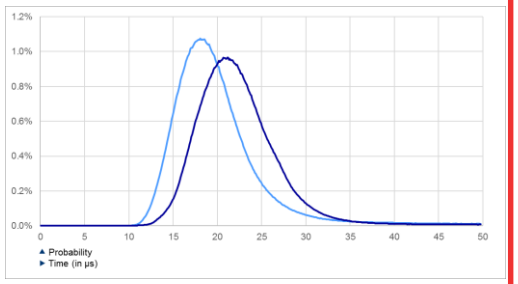


Network EOBI (t_9 to t_9a)

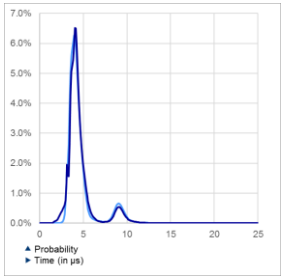


Response path

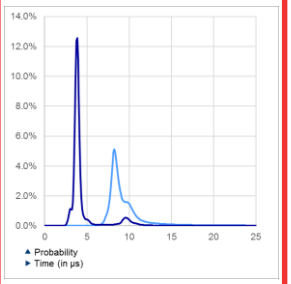
Matcher processing (response) (t_7 to t_6)



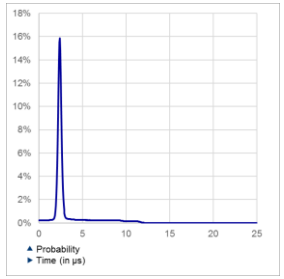
Matcher to gateway (t_6 to t_4')



Gateway response (t_4' to t_4)



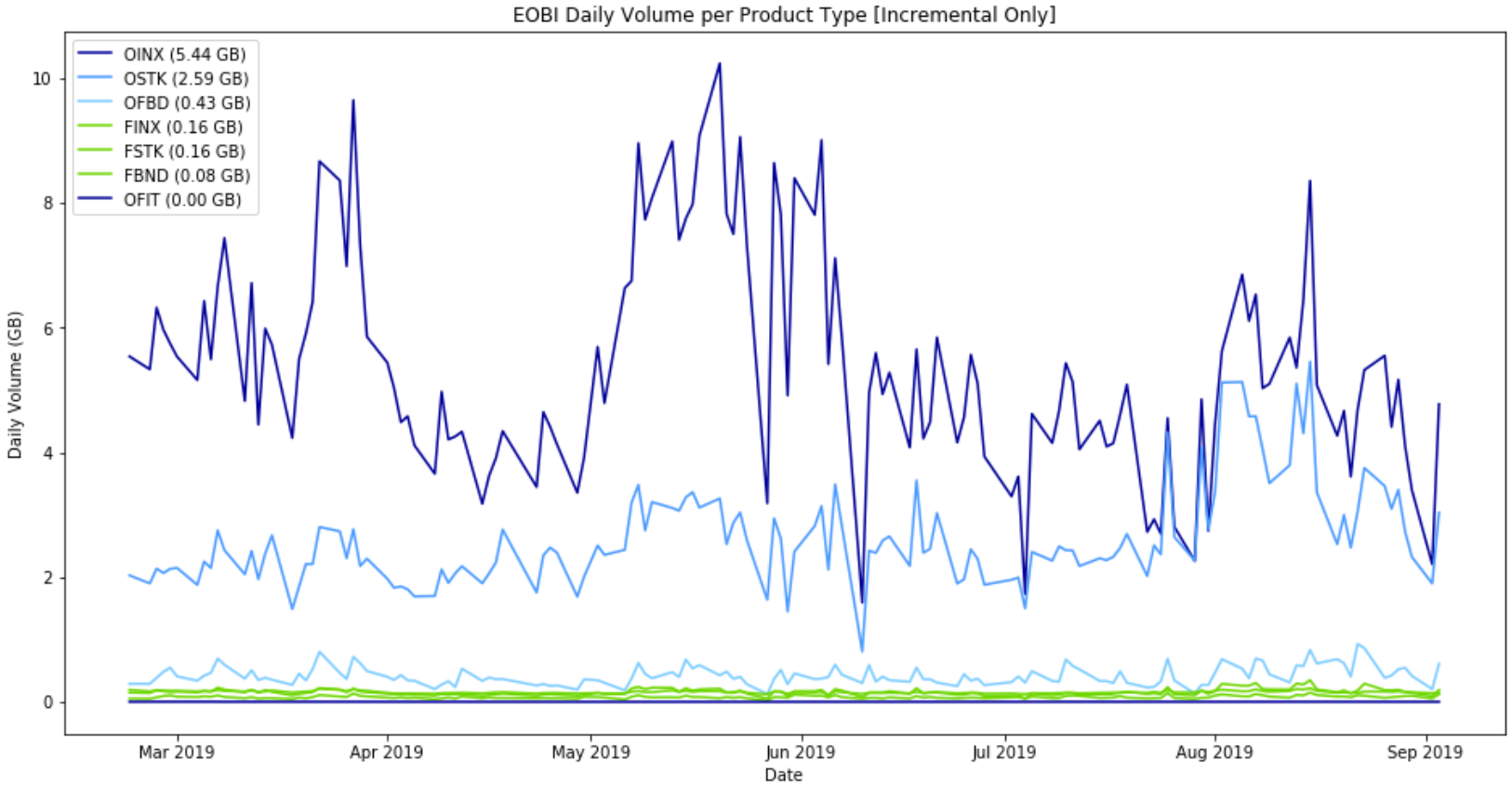
Network (t_4 to t_4a)



Developments since last Open Day

Launch of EOBI for Options in Feb 2019

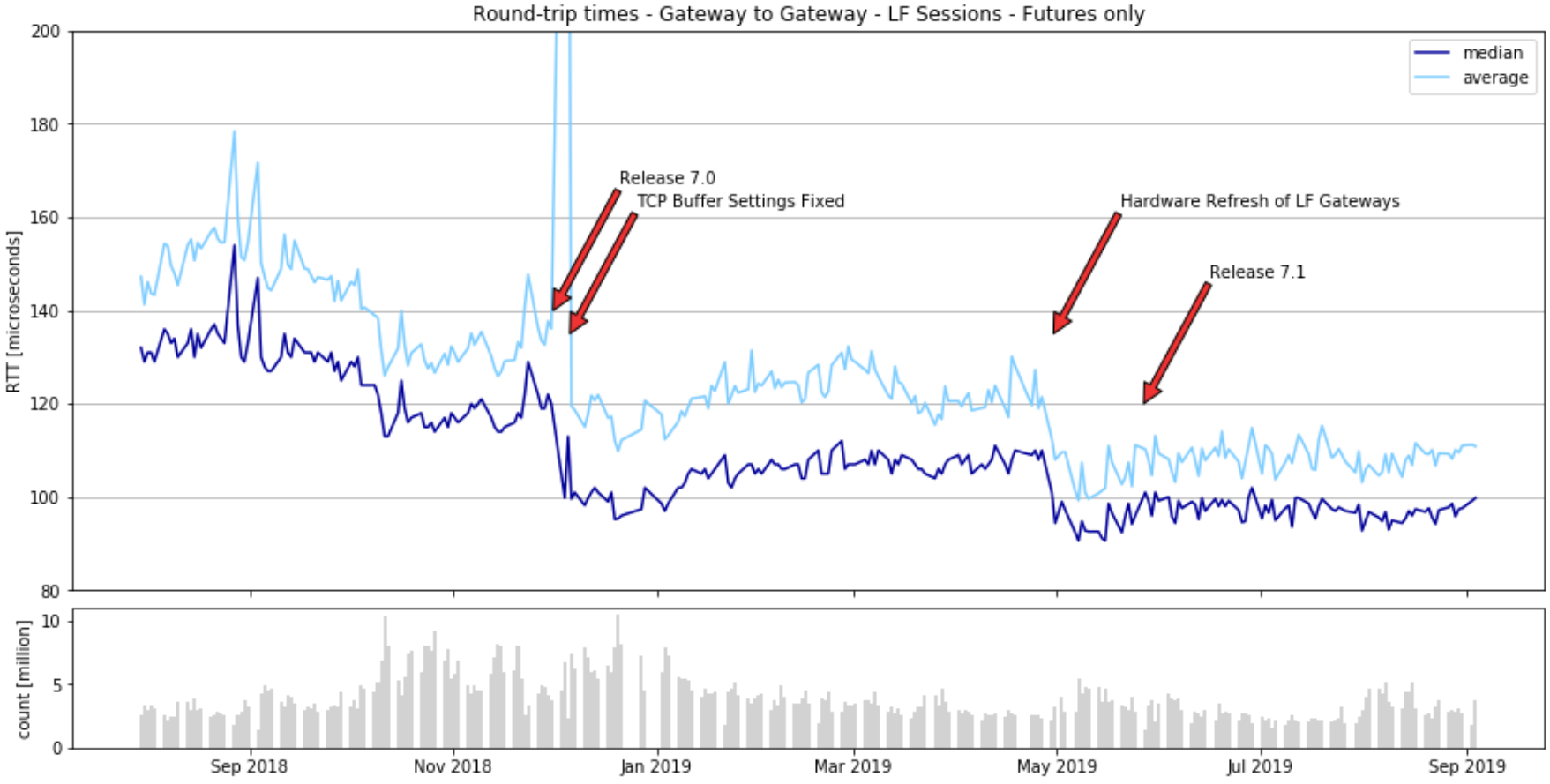
- EOBI volume in options much higher than in futures (avg of 8.5 GB vs 0.4 GB)
- EOBI faster in than 90% of trades and traded orders



Developments since last Open Day

Hardware refreshes of LF gateways in April 2019

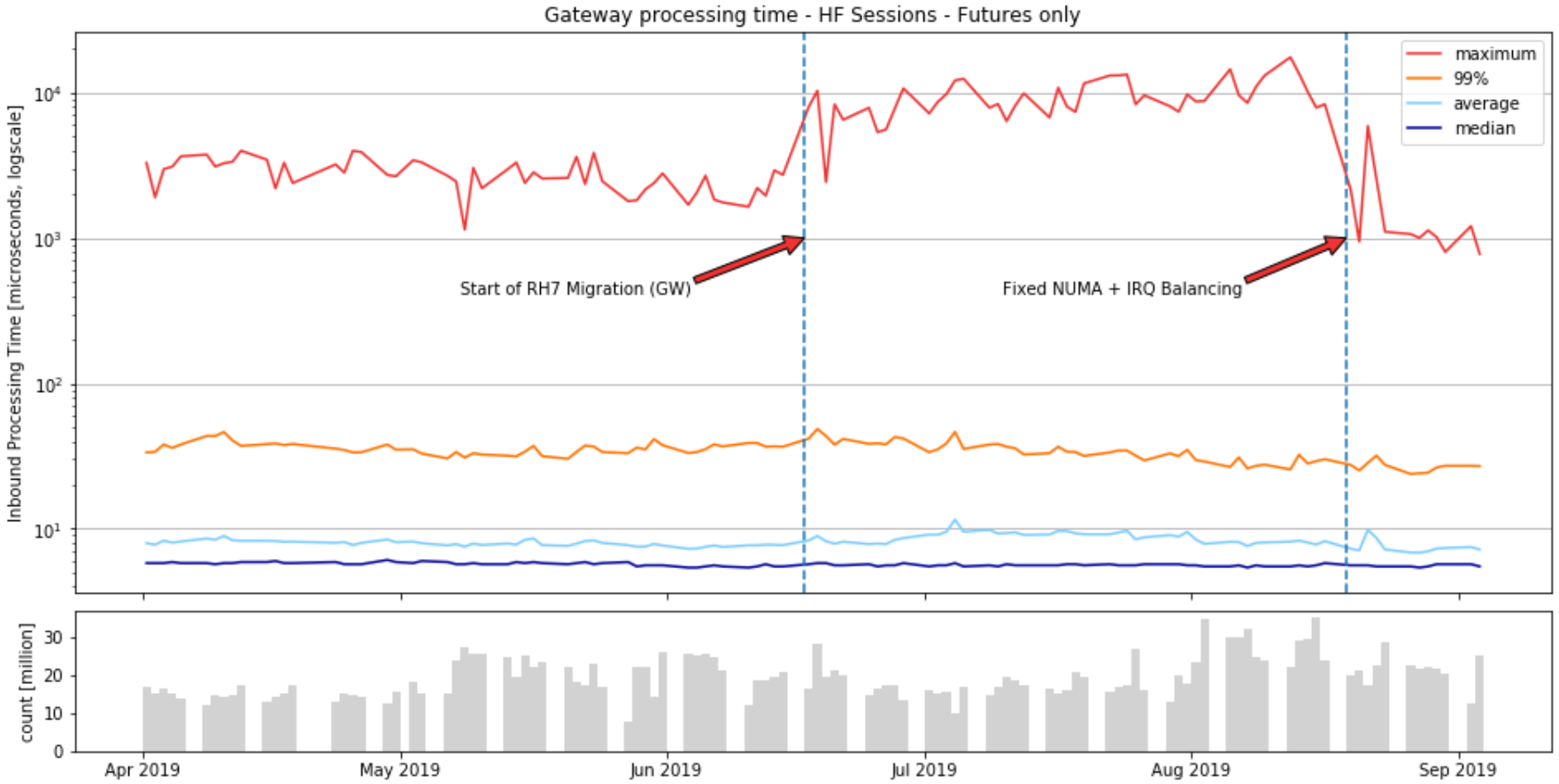
- Replaced LF gateways (and all other non-latency critical servers) by new HW
- Decrease of the roundtrip on LF gateways by ~20µs



Developments since last Open Day

Linux kernel migration

We are migrating from RedHat 6.9 (real-time kernel) to version 7.6 (standard kernel)



Developments since last Open Day

Inter-Product Spread Strategies

On 16 September 2019, Eurex launched Inter-Product Spreads (IPS) for fixed income futures as a standardised futures product. See Eurex circular 067/2019.

- T7® release 7.1 requires the IPS legs to reside on the same partition
- T7® release 8.0 does not have this requirement
- FGBX and FBTP were re-allocated (Eurex circulars 029/2019 and 055/2019) on 10 June 2019

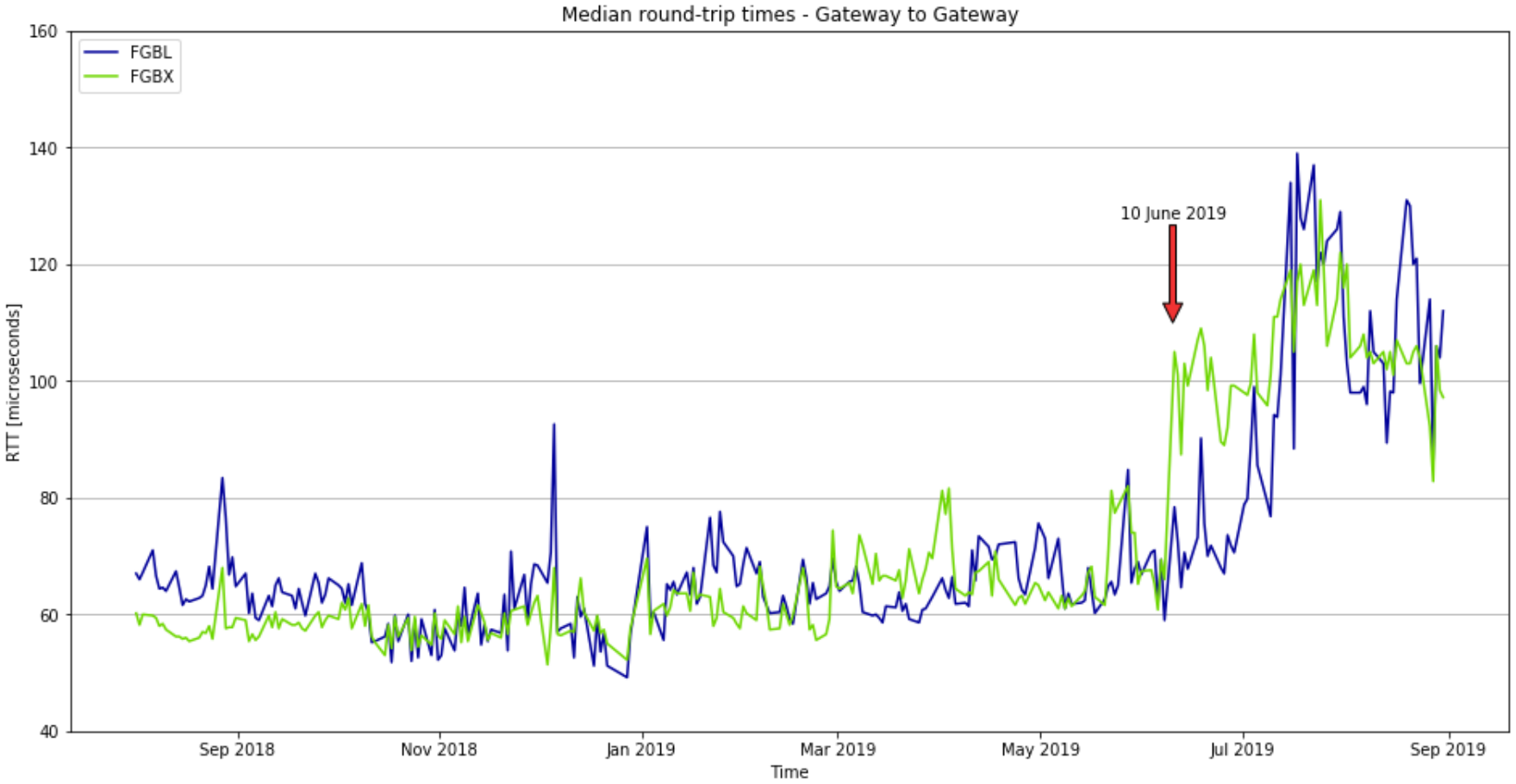
Product	Product Id	Leg 1	Partition 1	Leg 2	Partition 2
Schatz vs. Bund Spread	IPSL	FGBS	6	FGBL	2
Schatz vs. Bobl Spread	IPSM	FGBS	6	FGBM	4
Bobl vs. Bund Spread	IPMX	FGBM	4	FGBL	2
Bund vs. Buxl Spread	IPLX	FGBL	2	FGBX	**6 ➔ 2**
BTP vs. Bund Spread	IPPL	FBTP	**3 ➔ 2**	FGBL	2
OAT vs. Bund Spread	IPTL	FOAT	1	FGBL	2
BTP vs. OAT Spread	IPPT	FBTP	2	FOAT	1
BTP vs. BONO Spread	IPPO	FBTP	2	FBON	6
Schatz vs. BTS Spread	IPS2	FGBS	6	FBTS	6

Go-Live on 16 Sep 2019 on T7® current release 7.1

Candidates for go-live after introduction of T7® release 8.0

Developments since last Open Day

Inter-Product Spread Strategies - latency impact



The partition allocation of fixed-income futures will be reviewed after the introduction of T7® release 8.0.

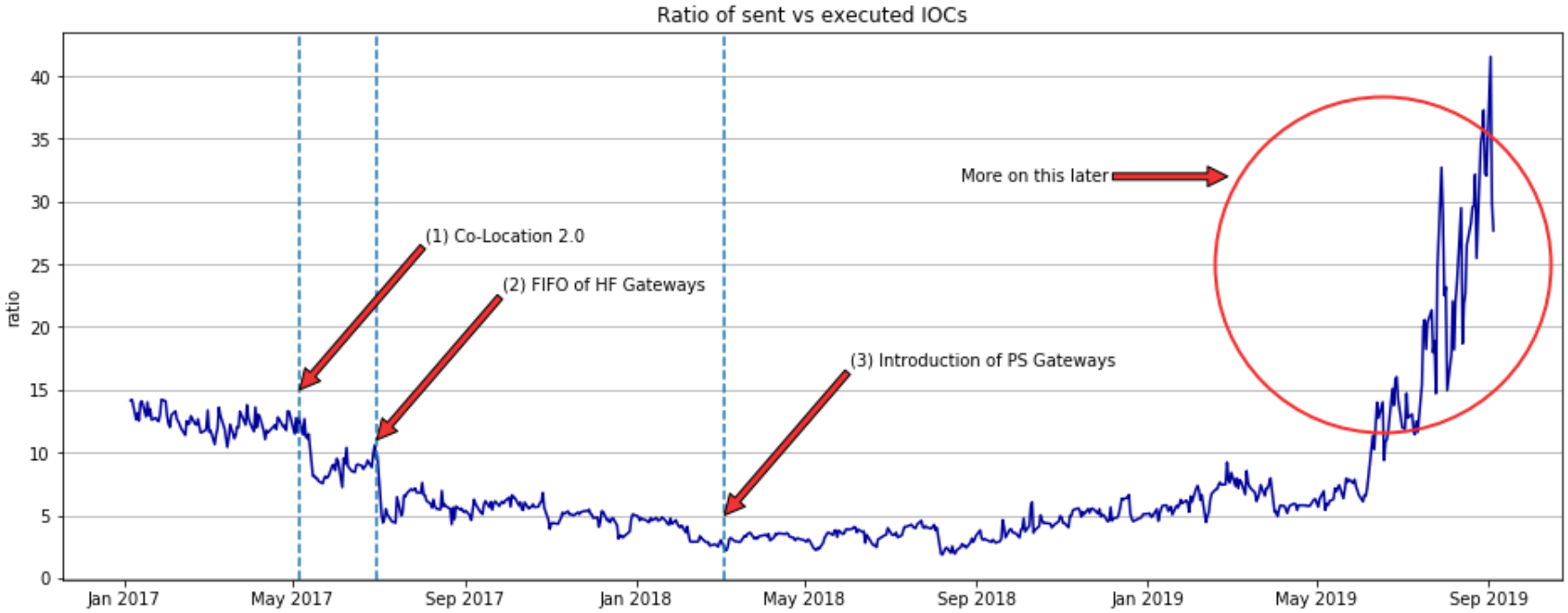
Developments since last Open Day

“Multiplicity”

Latency jitter on parallel inbound paths had incentivized multiplicity (i.e. usage of multiple parallel paths) 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.

Recent competition in the ultra low latency space has raised the ratio of sent vs. executed IOCs again (see below chart). Note however that the recent increase (4) is not driven by multiplicity – we rather see reactions on many more market data events than before. More on this topic later.



Developments since last Open Day

Launch of Passive Liquidity Protection (PLP) pilot phase

- Aggressive orders transactions, i.e. orders that are executable upon arrival in the matching engine, will be delayed before they are able to interact with the order book
- The deferral time is specific by product segment
- Pilot phase started on 3 June 2019
- No technical issues

Product Scope	Deferral Time
all German OSTK incl. weeklies	1 ms
all French OSTK incl. weeklies	3 ms

16

Topology changes for LF gateways

Topology changes for low frequency gateways

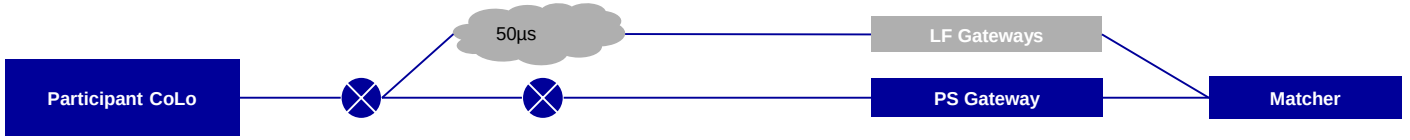
Tunnelling

The PS gateway is the single low-latency order entry point for T7® Xetra and Eurex.

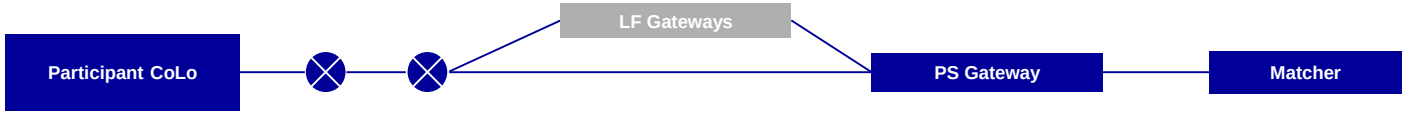
The LF gateways are usually slower than the PS gateways. However, the PS gateway queues requests during high loads. When this happens, requests sent to LF gateways may overtake PS gateway requests.

Solution: route all messages entered via LF gateway through the PS gateway.

Before



After



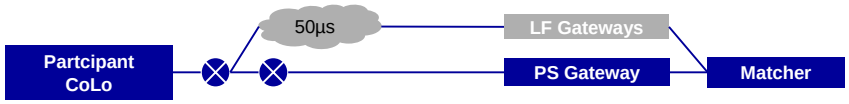
Topology changes for low frequency gateways

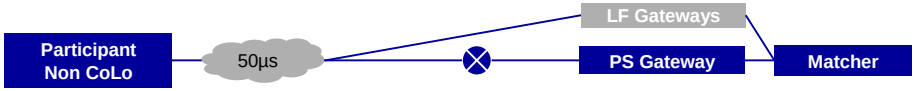
Tunnelling

Co-Location 2.0 (10 Gbit/s)

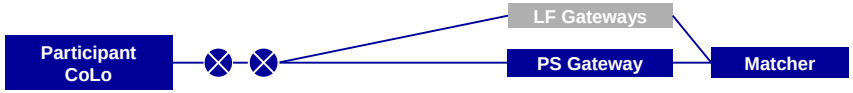
Non Co-Location

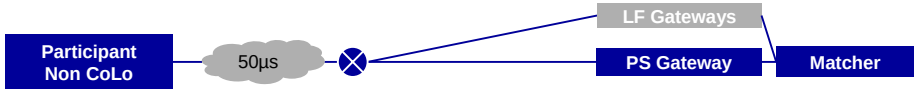
Current state



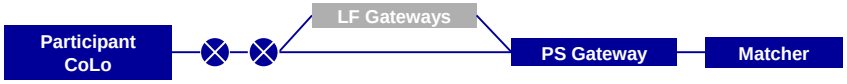


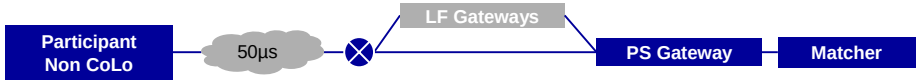
Step1 (Preparation, network changes)





Step 2 (Activation Q1 2020)





Topology changes for low frequency gateways

Tunnelling impact

Latency

- Co-location access via LF gateways ≈30µs faster than before
- Access via correct network side (A/B) more important than before
- Penalty for accessing “A” LF gateways via co-location “B” line (and vice-versa) is ≈50µs
- Penalty for accessing “A” partitions via “B” LF gateways in co-location (and vice-versa) is ≈50µs
- Matching priority is effectively assigned at PS gateway ingress
- Therefore no additional latency penalty for LF sessions compared to previous situation in co-location

ETI

- The RequestTime field in the ETI response will be filled with the PS gateway timestamp once tunnelling is activated
- LF gateway responses will contain the PS gateway timestamps

Market Data

- In EMDI and EOBI all fields referencing the matching engine in timestamp will be referencing the PS gateway in timestamp instead
- Exception: AggressorTime in the EOBI execution summary message

Topology changes for low frequency gateways

Timeline

Preparation (network changes to move LF gateways closer to PS gateways)

- Eurex Dec 2019
- T7® Xetra Oct 2019

Activation

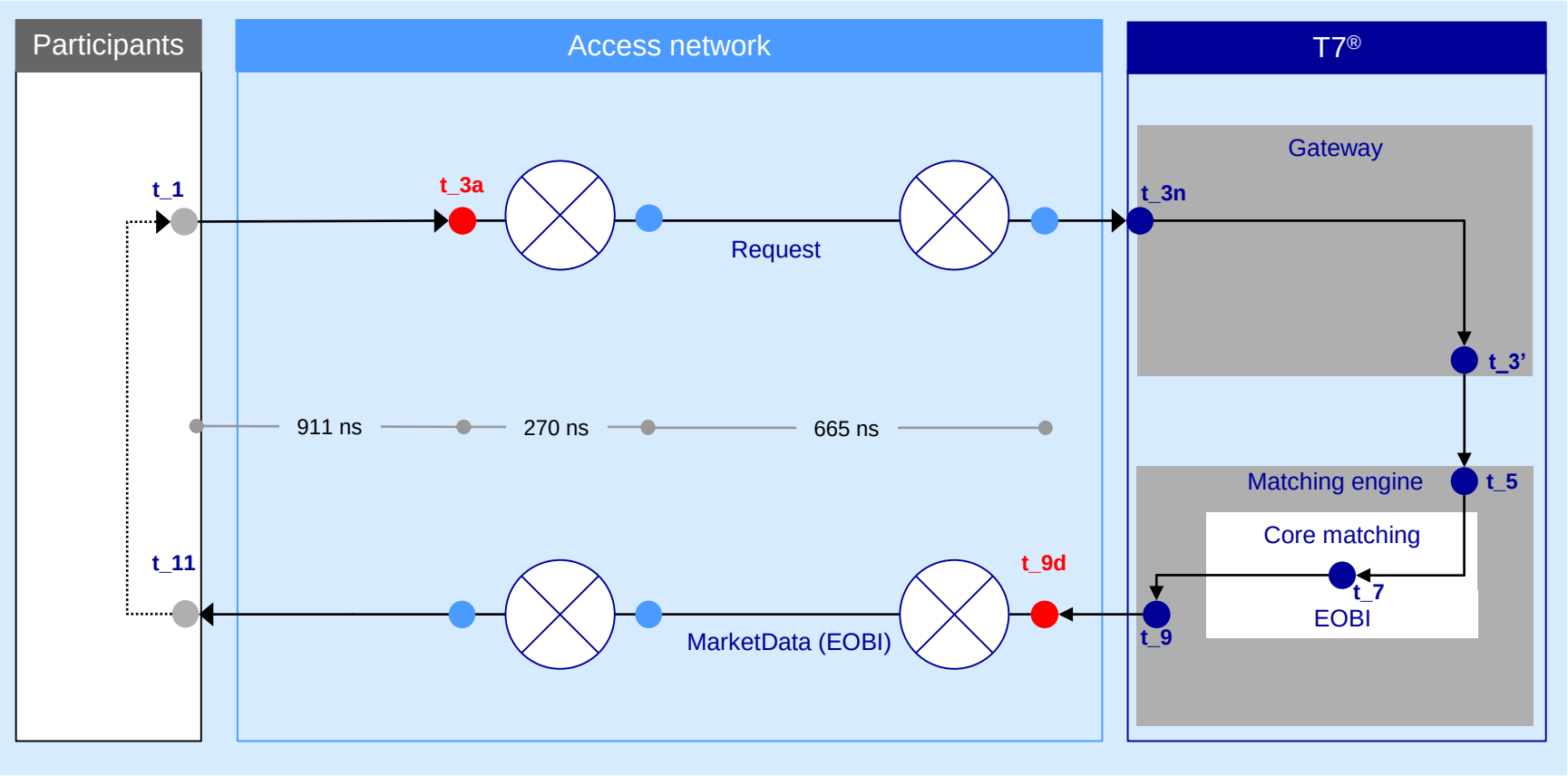
- Eurex Jan 2020
- T7® Xetra Feb 2020

21

High Precision Timestamp File

High Precision Timestamp File

Contains network times t_9d and t_3a for all trades



Use case: the signal generated by T7® leads to reactions by multiple trading participants. HPT allows to calculate reaction time differences with higher precision.

http://datashop.deutsche-boerse.com/High-precision-timestamps

High Precision Timestamp File

X-ray for network dynamics and ultra low latency reactions

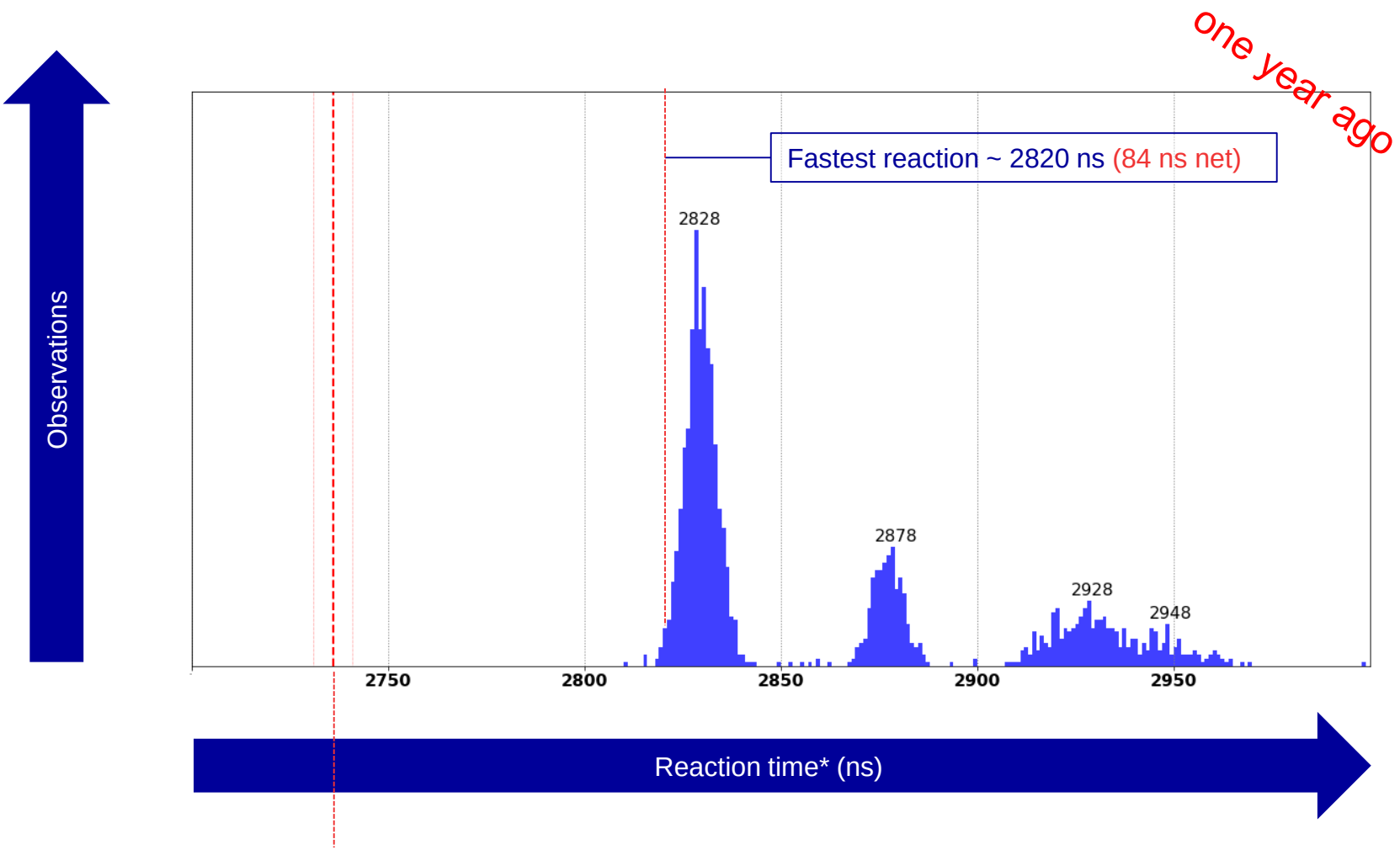
Extended the HPT file service

- Launched in 2018: executions only
- Since March 2019: all EOBI data for futures & cash
- Since August 2019: all EOBI data for Eurex (incl. options)

https://www.mds.deutsche-boerse.com/mds-en/data-services/analytics/high-precision-timestamps

High Precision Timestamp File

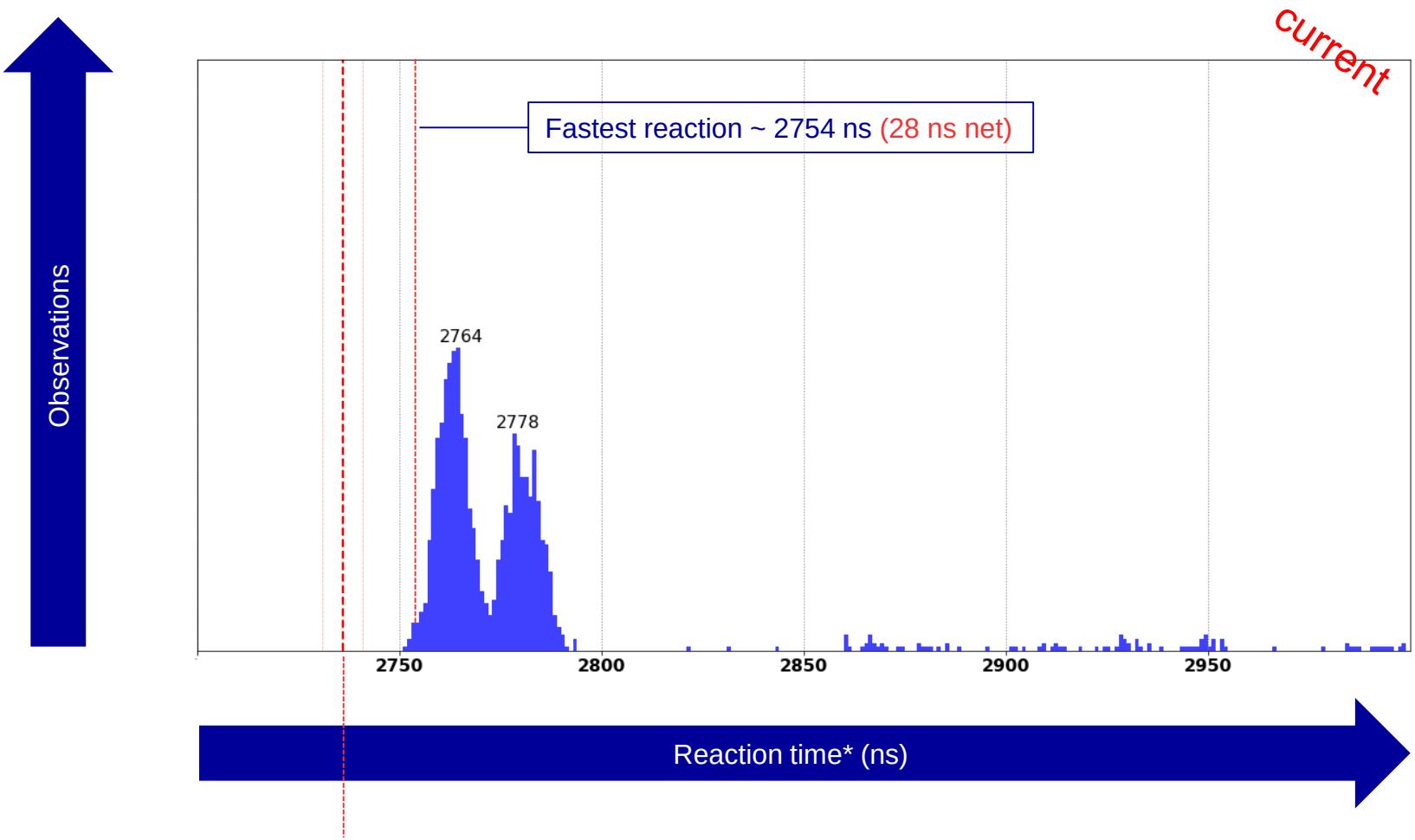
Reaction time based on t_9d and t_3a (HPT)



*t_3a – t_9d for FDAX trades triggered by FESX

High Precision Timestamp File

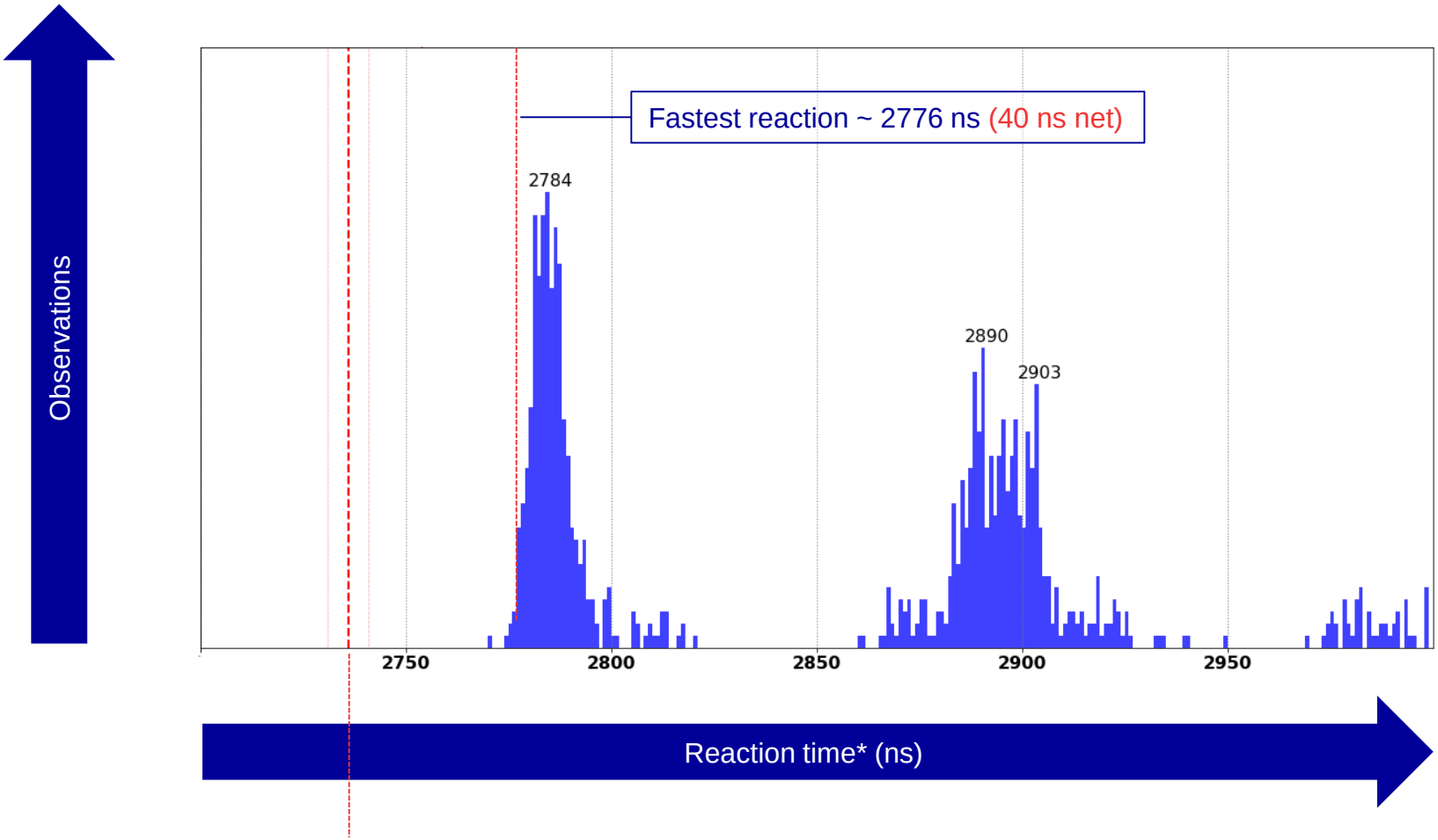
Reaction time based on t_9d and t_3a (HPT)



*t_3a – t_9d for FDAX trades triggered by FESX

High Precision Timestamp File

Reaction time for OESX orders



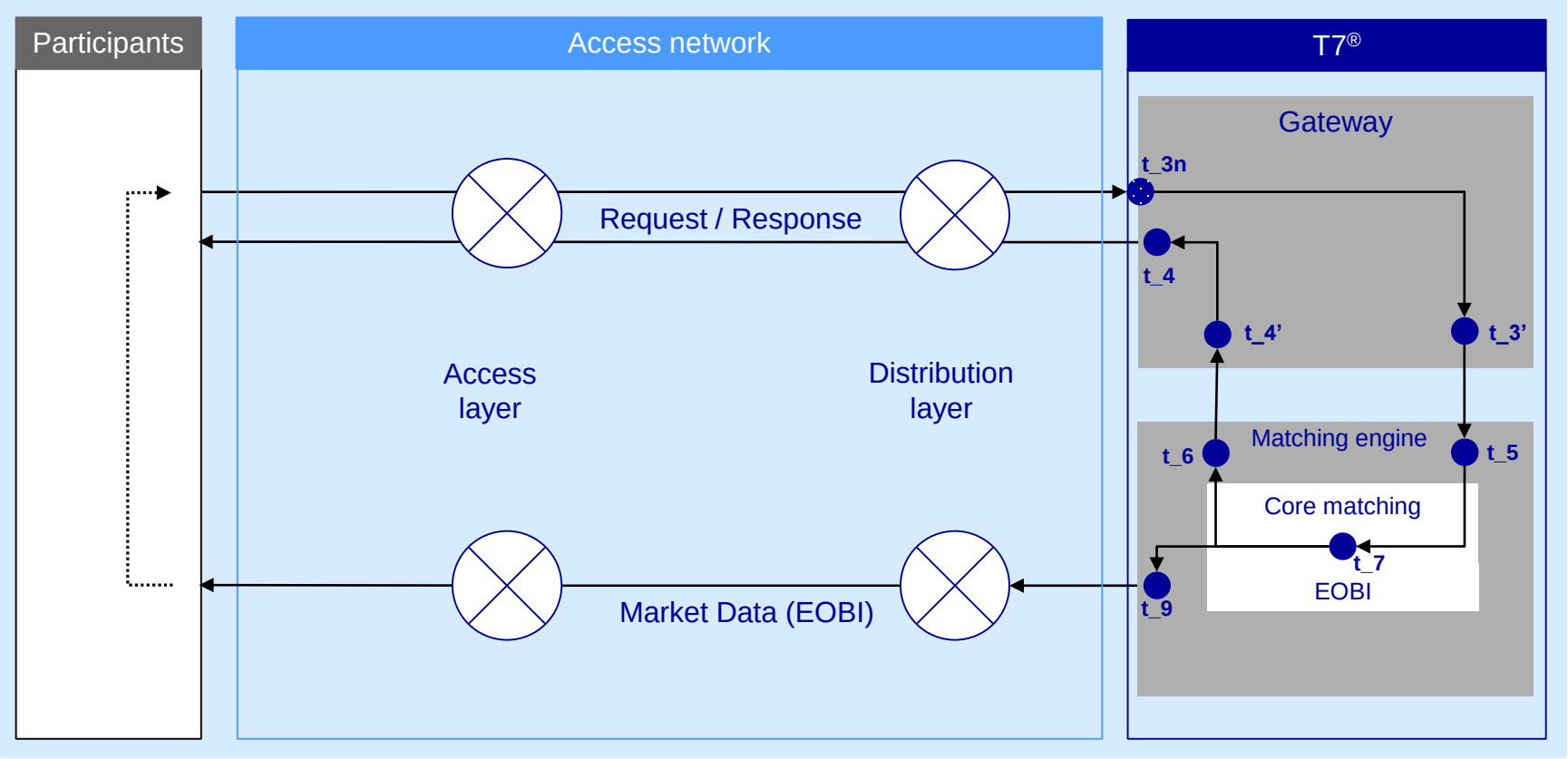
*t_3a – t_9d for OESX trades triggered by OESX orders

27

Network dynamics

Network dynamics

Simplified topology



- Timestamps provided in T7 API (in real time) in dark blue (t_3n: taken by network card, other: application level)
- Cisco 3548X switches operating in cut-through mode.

Network dynamics

Ethernet switching paradigms

Store and Forward

- Store the complete frame in memory
- Compute frame check sequence
- Decide where to forward to
- Forward

- ➢ Slow compared to cut-through
- ➢ Forwarding latency has more jitter than cut-through
- ➢ Drops any bad frames

Cut-Through (used in T7® co-location)

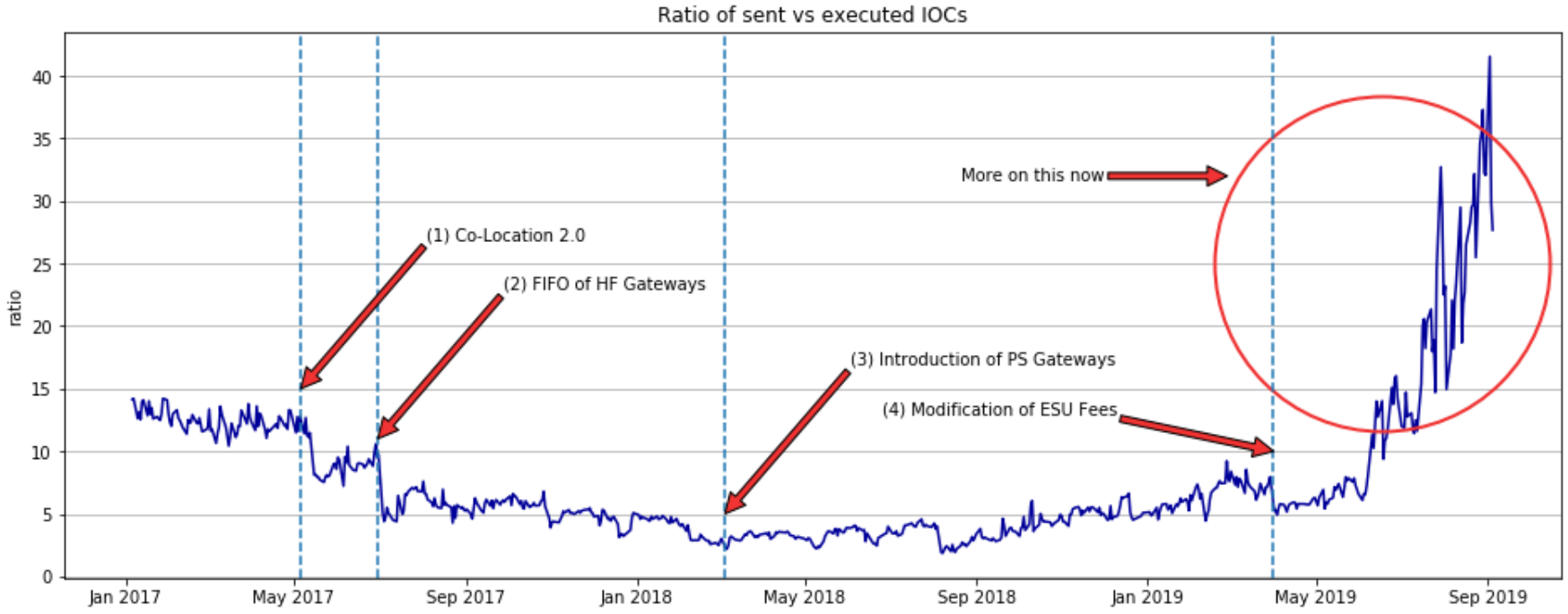
- Forward as soon as destination MAC is known (in theory)
- Actually reads a few more bytes (required to check ACLs for example)
- Only stores frame if egress port is currently blocked

- ➢ Faster than store-and-forward, less jitter
- ➢ Forwards bad frames (bad FCS, runts, etc.)
- ➢ In our case, extra burden of “rubbish disposal” put on gateways
- ➢ See T7® Network Access Guide for usage guidelines

Network dynamics

Speculative triggering

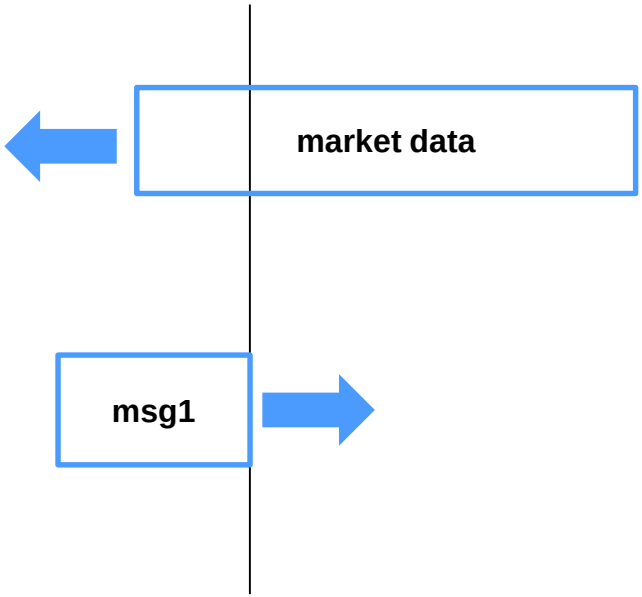
1. React immediately on market data packets by submitting first of N messages. Sometimes this message has no order book impact. Financially disincentived by changes in the Excessive System Usage (ESU) fee.
2. Followed-up by at least one more message in the very same Ethernet frame. The first message has the side effect of reserving priority in the network.



Network dynamics

Speculative triggering

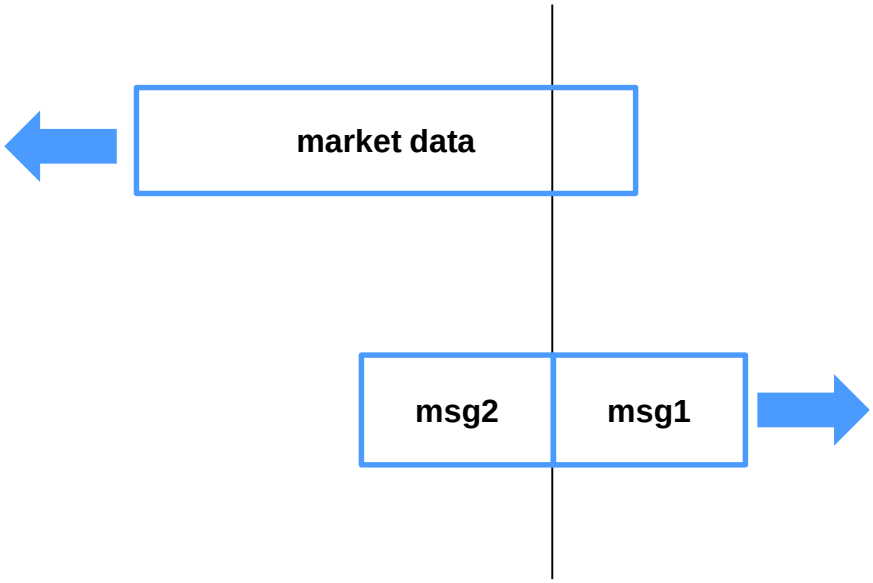
Step 1



t_0+ δ: react by sending msg1

t_0: start to receive market data

Step 2



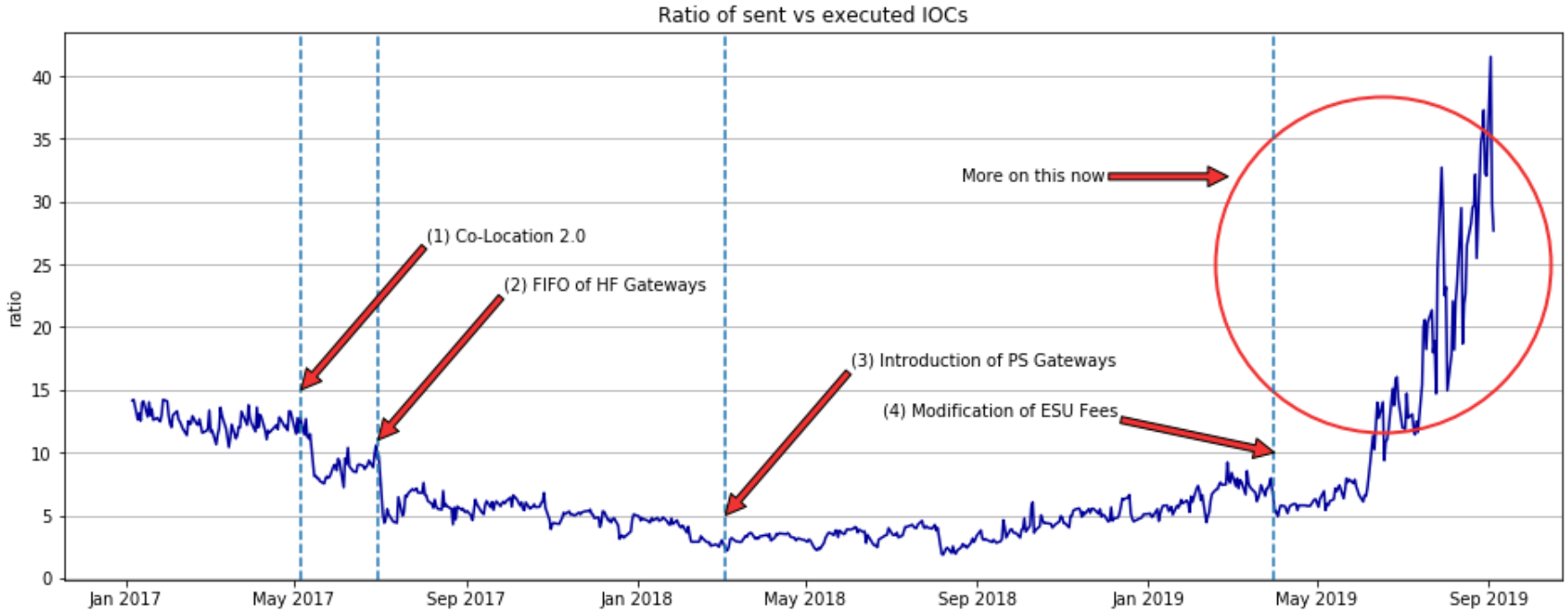
t_1: send msg2

msg1 + msg2 inseparable

Network dynamics

Speculative triggering

1. The most recent increase is because of similar method now applied in options.
2. Changes to ESU fees for option products under review



Network dynamics

Ethernet switching paradigms

- The T7® co-location network switches operate in cut-through mode
- This will not change at least until end of 2020
- None of the two paradigms store-and-forward or cut-through are ideal
- Each paradigm incentivizes unorthodox behaviour to gain or reserve priority on the network
- We plugged the obvious technical loopholes and monitor for unexpected behaviour.

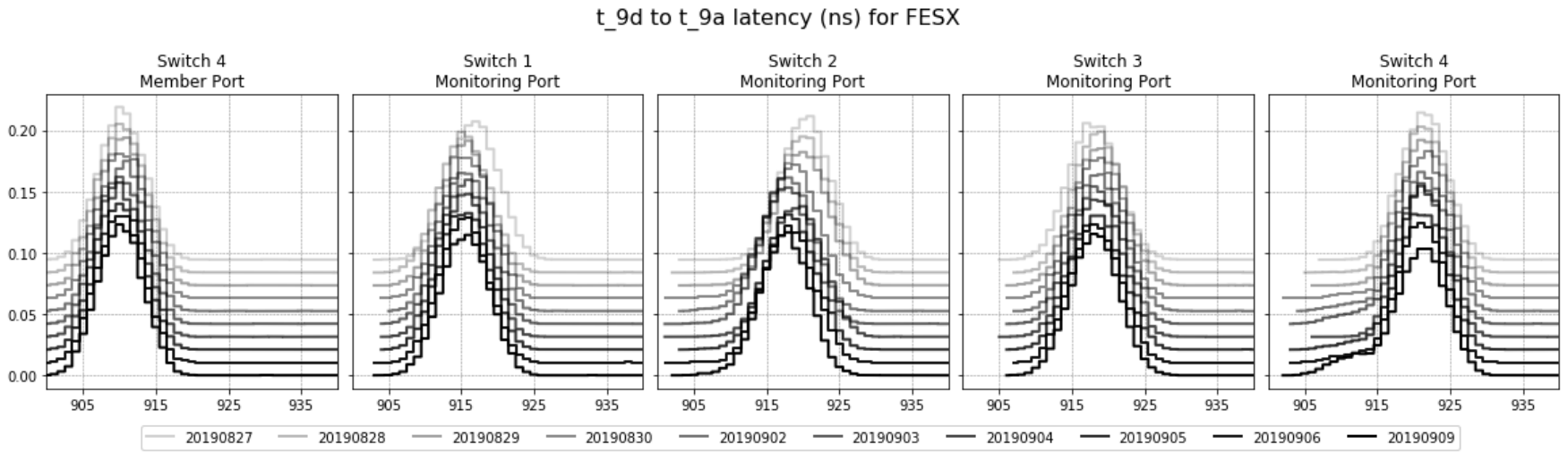
34

Market data

Market data

Port fairness

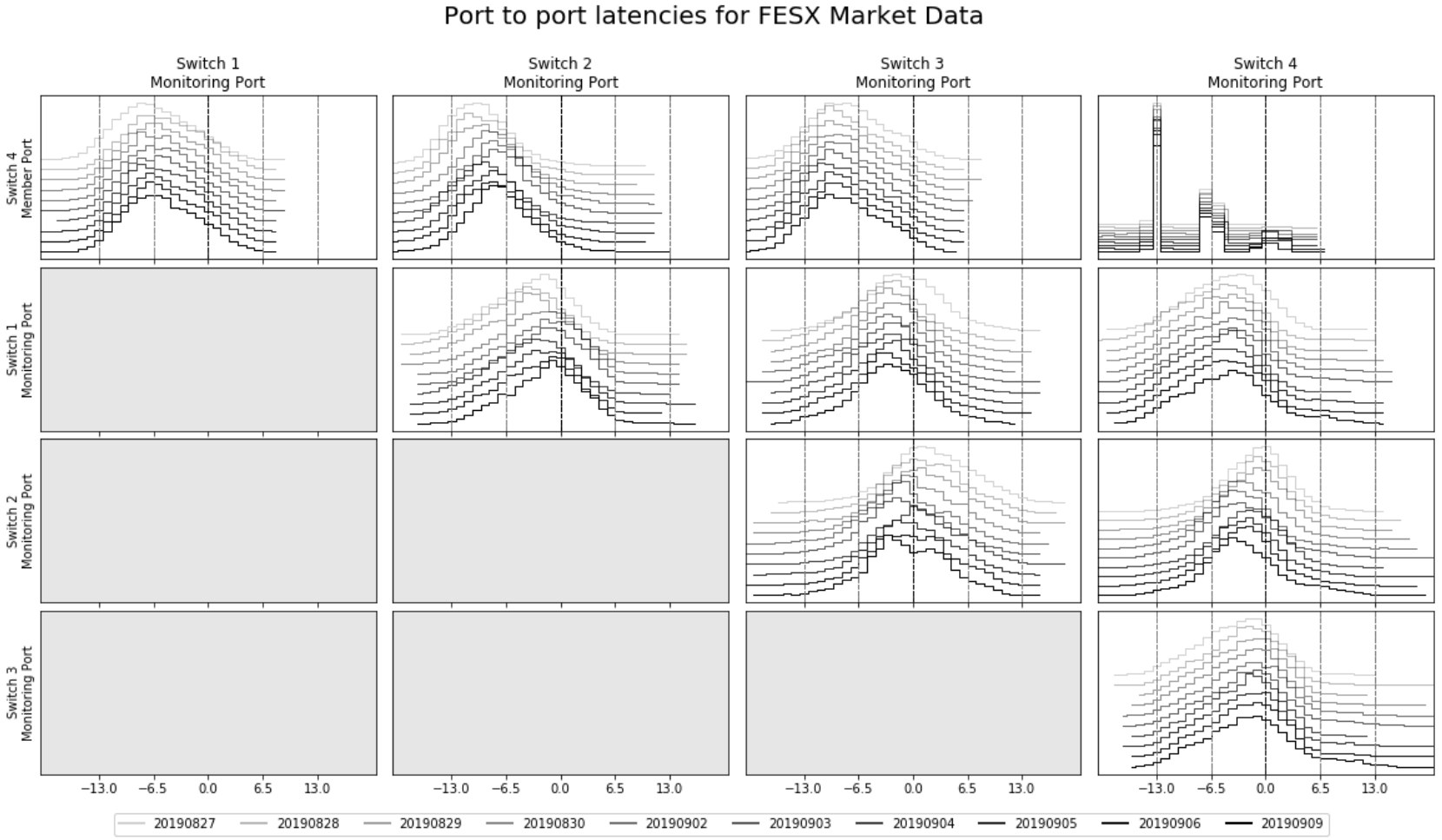
- Jitter from one packet to the next ≈15-20ns
- Some ports show a static offset ≈7-8ns (leftmost graph)
- The static offset is not caused by cable length differences
- The offset appears to depend on the multicast group subscription lists of a) the port under consideration and b) other ports on the same switch
- Analysis of this effect still ongoing
- Try different market-data cross-connects if those differences are relevant



Market data

Port fairness

Comparison of t_9a vs t_9a for different market data lines relative to each other

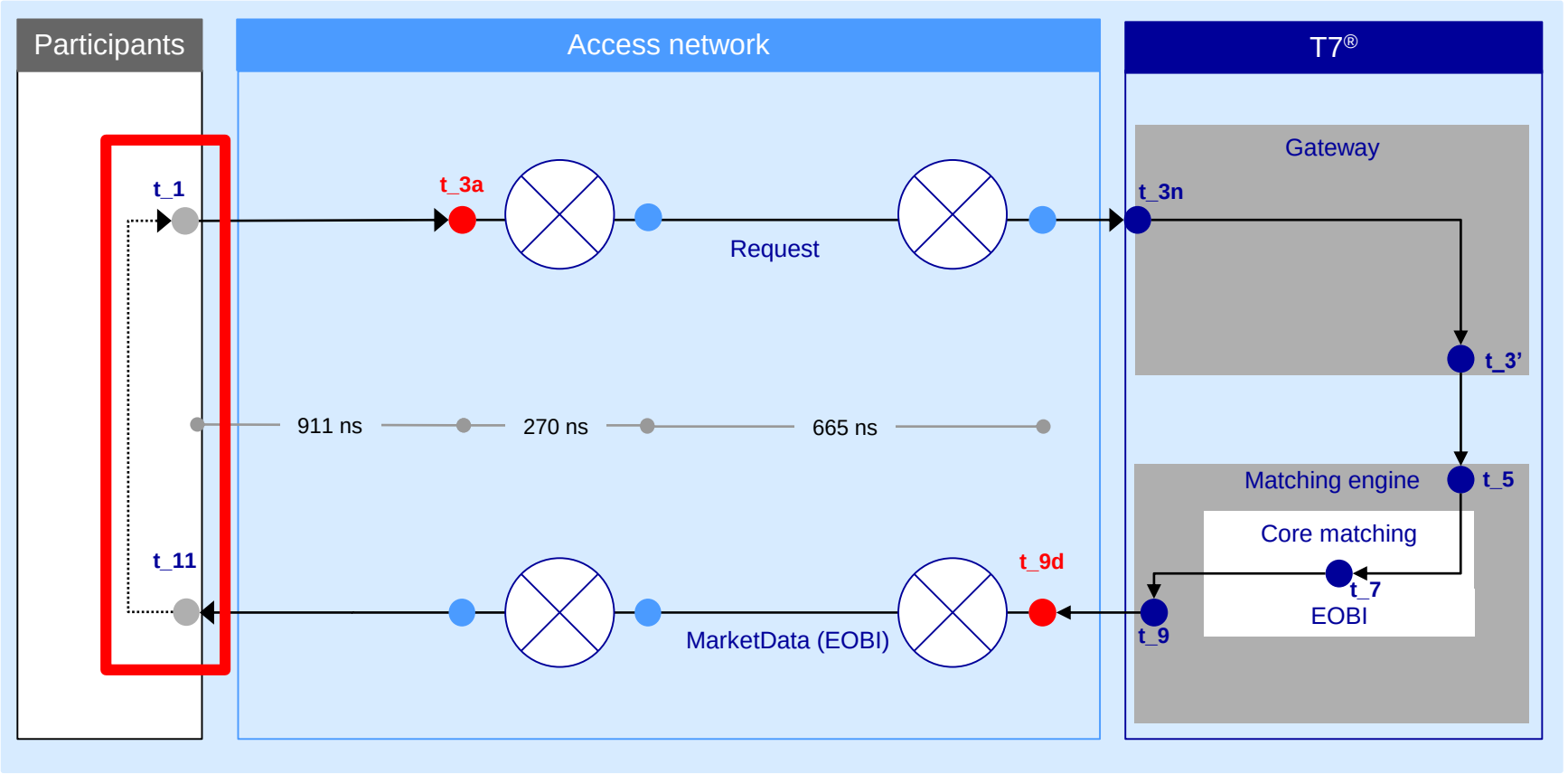


37

T7® Time Synchronisation White Rabbit

T7® Time Synchronisation

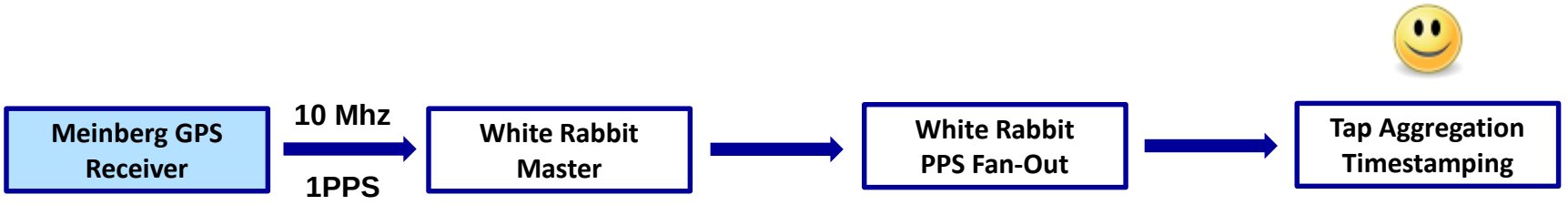
White Rabbit



- White Rabbit time service went live in Q3 2018 as pilot project
- Fully supported alternative to PTP since April 2019

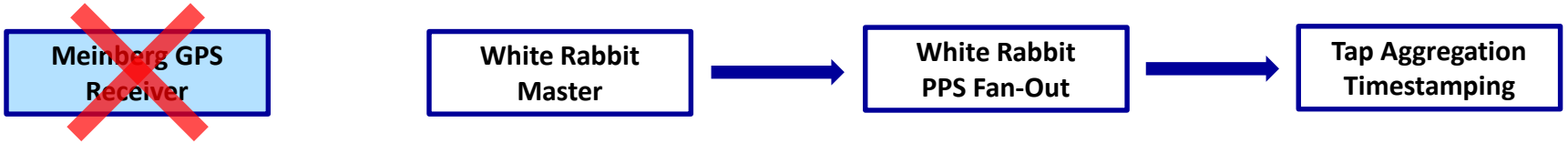
White Rabbit 2019-08-05

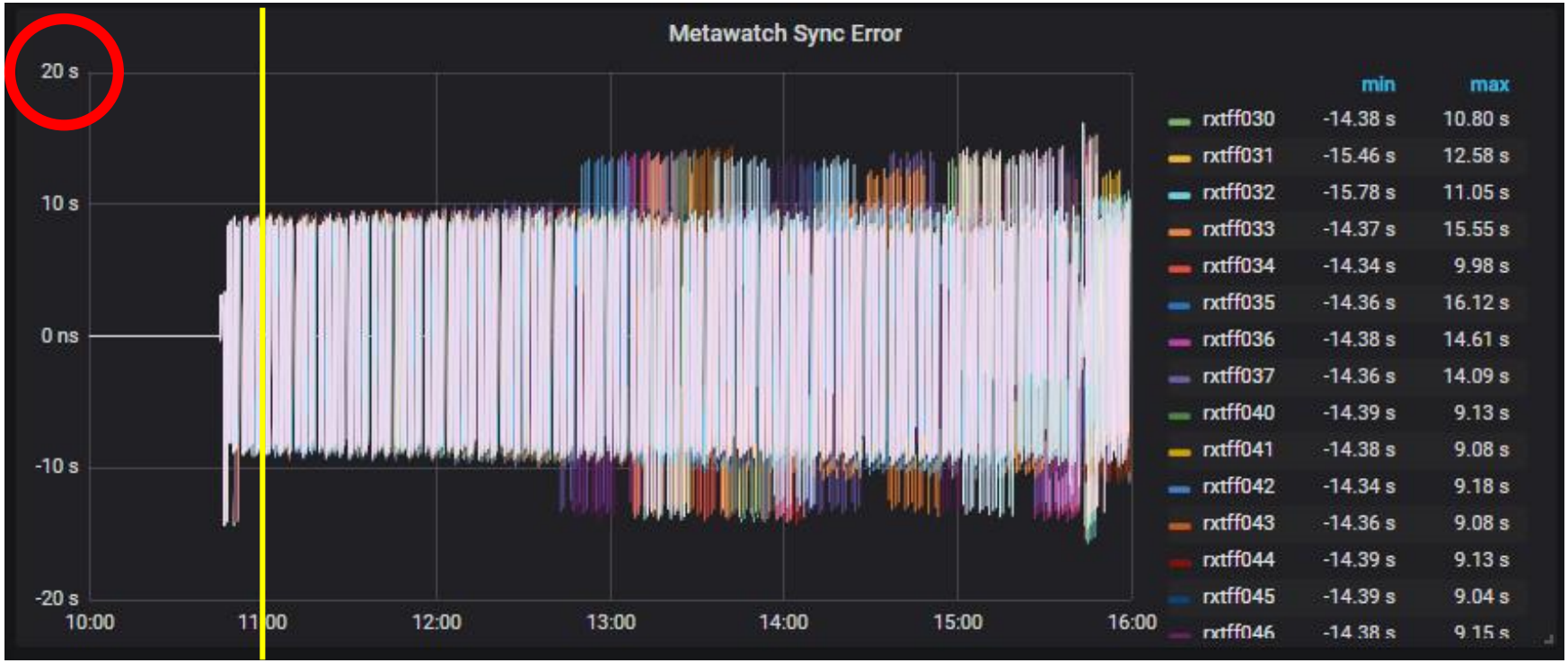
Normal Operation – Sync error in timestamping devices is less than +/- 1ns



White Rabbit 2019-08-05

- Planned work on GPS receiver on a Saturday
- GPS service was restored by 11:00 (yellow bar)
- Root cause known and solution is underway





Thank you for your attention.

Contact
Sebastian Neusüß
Andreas Lohr
E-mail monitoring@deutsche-boerse.com
Phone +49-(0) 69-2 11-1 86 86

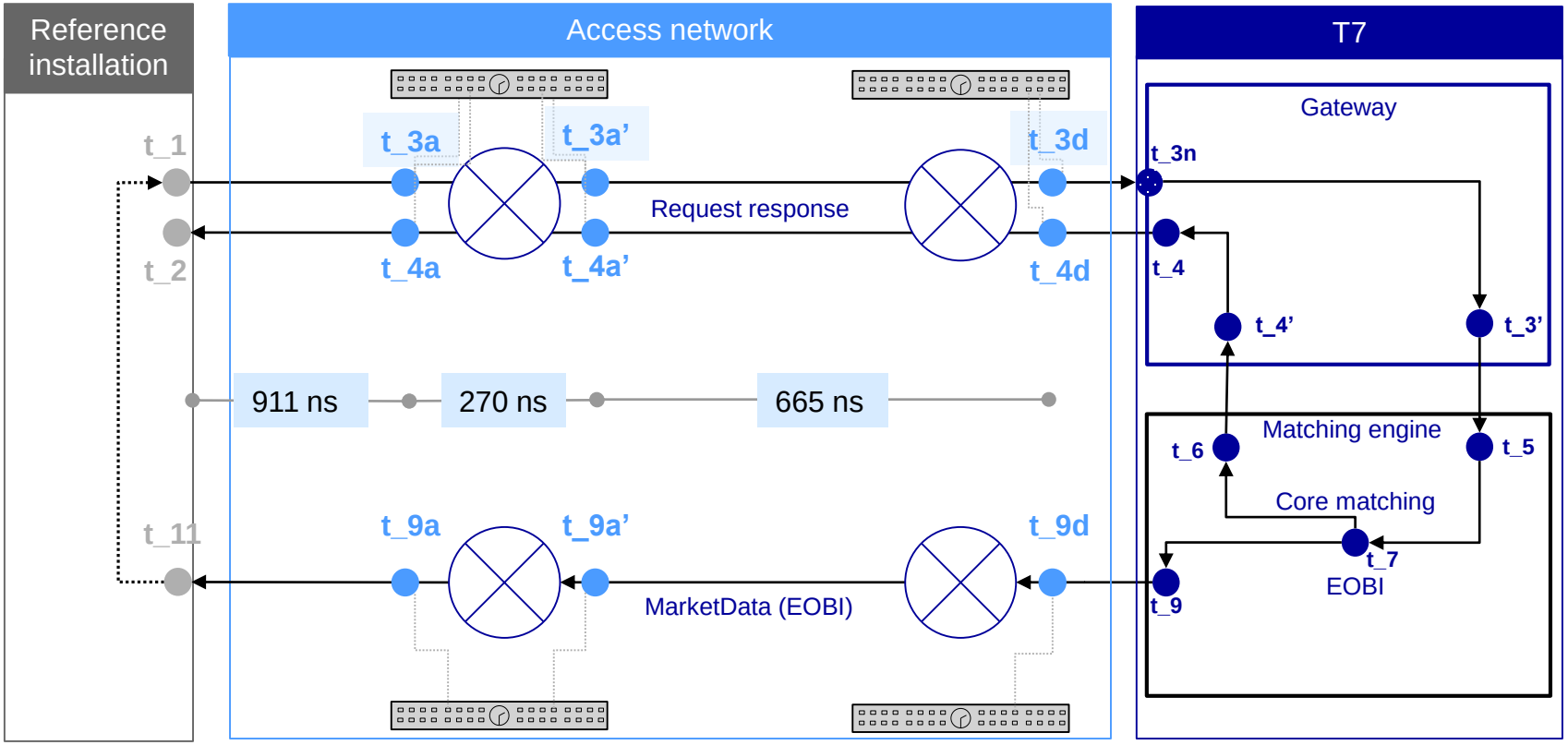
DEUTSCHE BÖRSE GROUP

26 September 2019

Disclaimer

Deutsche Börse AG opens up international capital markets for its customers. Its product and service portfolio covers the entire process chain – from pre-IPO services and the admission of securities, through securities and derivatives trading through the settlement of transactions and the provision of market information to the development and operation of electronic trading, clearing and settlement systems. With its process-oriented business model, Deutsche Börse increases the efficiency of capital markets. Committed employees are the key factor for innovation and further growth: without them, Deutsche Börse Group would not have developed into one of the most modern exchange organisations in the world. More than 5,000 employees work for the Group – a dynamic, motivated and international team.

T7® 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