DEUTSCHE BÖRSE GROUP

Precise Timing in Financial Markets

Time Distribution in Deutsche Börse's T7® Trading Network

CS 349F Technology for Financial Systems | October 2020 | Stanford

About the speakers

Andreas Lohr

- Degree in computer science from University of Applied Sciences in Darmstadt (1994)
- Working in technology for financial trading since
- Various roles in two investment banks (Frankfurt, London)
- Performance engineer in a technology-driven trading firm (Amsterdam)

Joined Deutsche Börse in 2016

- Trading IT, monitoring, infrastructure, co-location, and part-time chief cable measurer
- Deutsche Börse operates Xetra, the reference market for exchange trading in German shares and ETFs, and Eurex, a leading global derivatives exchange

Sebastian Neusüß

- Diploma in Theoretical Physics from Johannes Gutenberg University in Mainz (1999)

Joined Deutsche Börse in 1999

- Co-authored the messaging system that drives the current trading system of Deutsche Börse (T7)
- Designed the latency monitoring for T7
- Heads the trading system analytics team
- Does real time application monitoring for the T7 platform, market microstructure analysis, data analysis and chasing nanoseconds

We will go deep and move from business logic to ASIC design.

Agenda

3

Overview

The New York Times

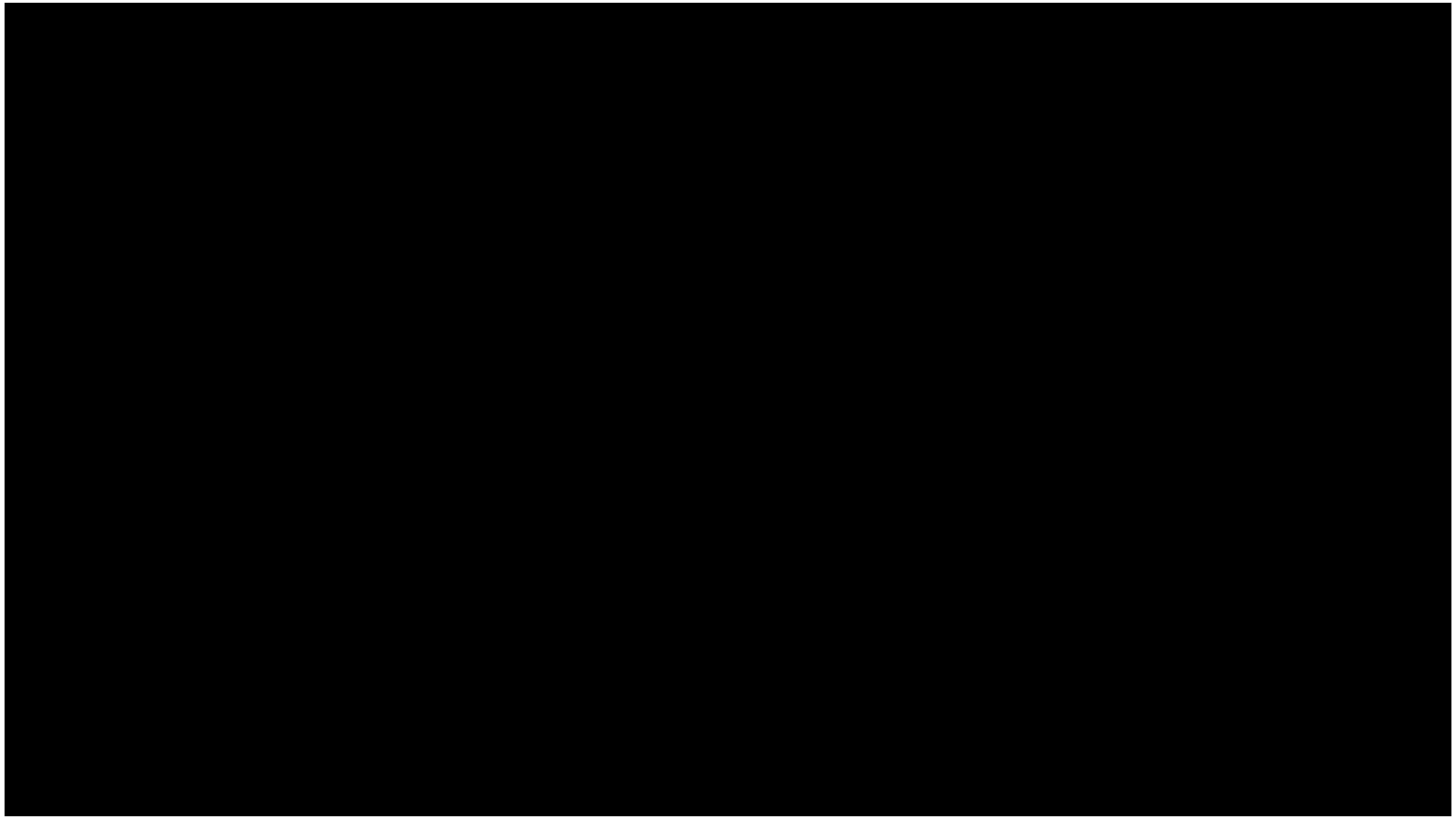
“Time Split to the Nanosecond Is Precisely What Wall Street Wants”

“The financial industry has easily become the most obsessed with time”

-Balaji Prabhakar, Stanford University

https://www.nytimes.com/2018/06/29/technology/computer-networks-speed-nasdaq.html

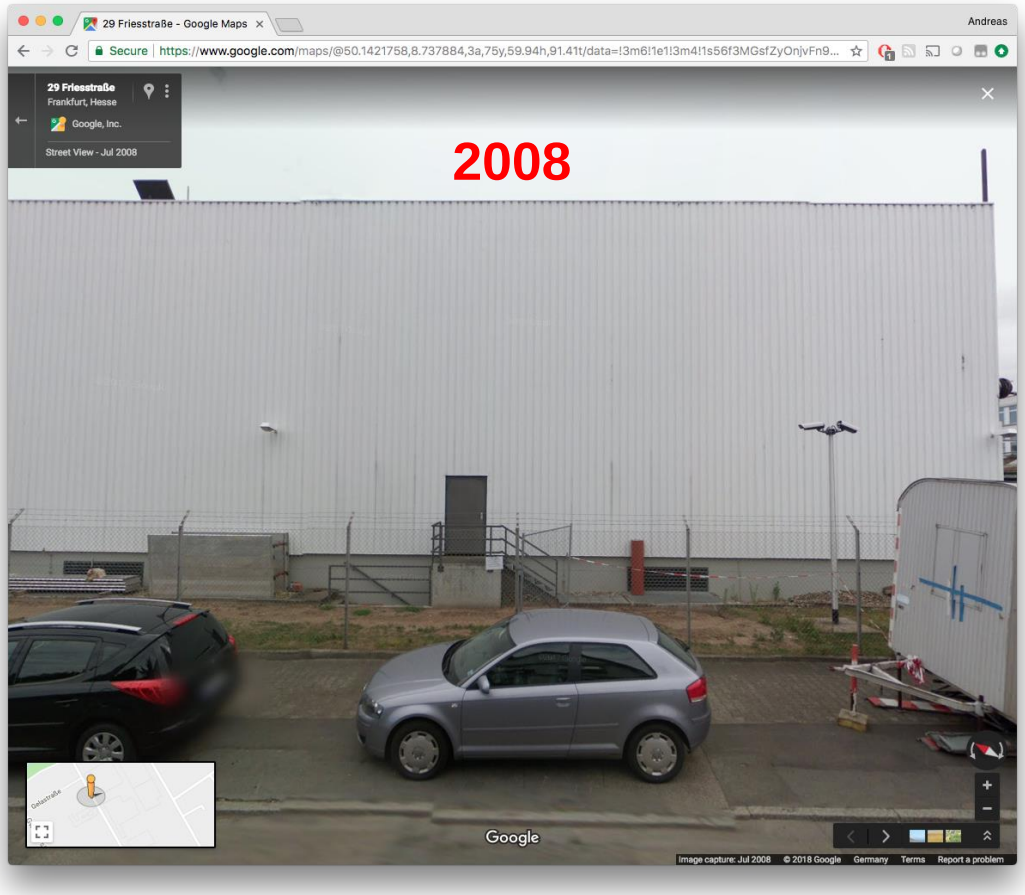
The Hummingbird Project (2018), fiction based on a true story



https://www.youtube.com/watch?v=YtAz6CT-1yA

Obsessed with Time?

Data center in Frankfurt, Germany



GPS (and other similar systems) disseminate precise time

History of financial trading in Frankfurt

1150 Medieval trade fair

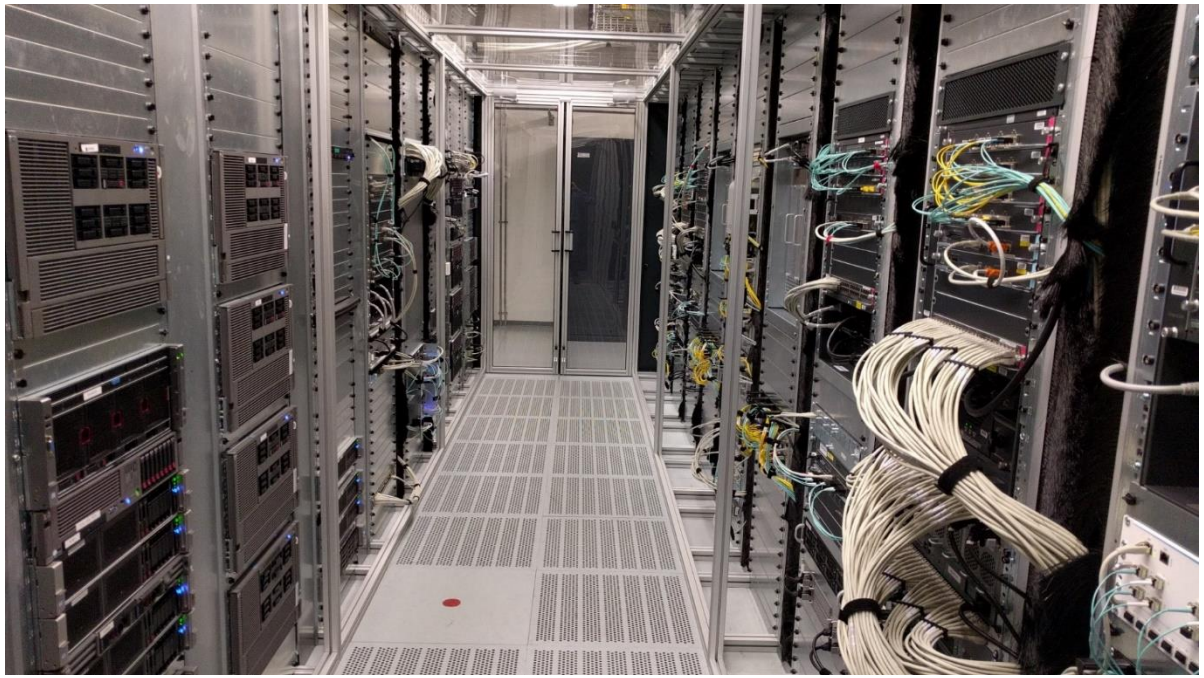
1585 Frankfurt merchants set uniform exchange rates for the first time

1820 The first share is listed in Frankfurt (shares of the Austrian National Bank)

1843 The first stock exchange building

1990 First electronic-only market for derivatives (DTB, Deutsche Terminbörse, now Eurex)

1997 Introduction of Xetra as the electronic stock market

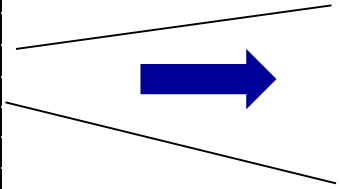


8

Central Limit Order Book

Central Limit Order Book (1)

LastTime	LastVolume	Last			
10:20:10.123	165	73.04			
BidCount	**BidVolume**	**Bid**	**Ask**	**AskVolume**	**AskCount**
3	624	73.04	73.05	3081	5
10	1474	73.03	73.06	2662	15
14	2505	73.02	73.07	5854	17
16	2843	73.01	73.08	4284	14
10	1925	73.00	73.09	6388	15
12	2328	72.99	73.10	1908	11
11	2814	72.98	73.11	1151	5
7	967	72.97	73.12	534	4
10	1803	72.96	73.13	4090	5
8	938	72.95	73.14	1104	4



Ask	AskVolume	QueuePos
73.05	100	1
73.05	791	2
73.05	1368	3
73.05	685	4
73.05	137	5

- List of **buy** orders on the “Bid” side
- List of **sell** orders on the “Ask” side
- Sorted by price

- Priority on each price level by time
- Price-time priority is the most used electronic execution model

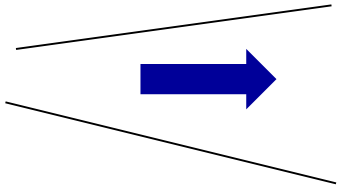
https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-1-141016
https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-II--141112

Central Limit Order Book (2)

LastTime	LastVolume	Last			
10:20:10.123	165	73.04			
BidCount	**BidVolume**	**Bid**	**Ask**	**AskVolume**	**AskCount**
3	624	73.04	73.05	3081	5
10	1474	73.03	73.06	2662	15
14	2505	73.02	73.07	5854	17
16	2843	73.01	73.08	4284	14

New order: Buy 100 @ 73.05

LastTime	LastVolume	Last			
10:20:10.352	100	73.05			
BidCount	**BidVolume**	**Bid**	**Ask**	**AskVolume**	**AskCount**
3	624	73.04	73.05	**2981**	**4**
10	1474	73.03	73.06	2662	15
14	2505	73.02	73.07	5854	17
16	2843	73.01	73.08	4284	14



Ask	AskVolume	QueuePos
73.05	100	1
73.05	791	2
73.05	1368	3
73.05	685	4
73.05	137	5

Example:

We want to buy 100 shares at a price of 73.05

Q: Which of the 5 sell orders is “matched” with our buy order?

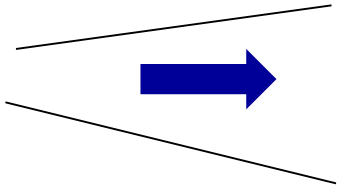
https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-1-141016
https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-II--141112

Central Limit Order Book (3)

LastTime	LastVolume	Last			
10:20:10.123	165	73.04			
BidCount	**BidVolume**	**Bid**	**Ask**	**AskVolume**	**AskCount**
3	624	73.04	73.05	3081	5
10	1474	73.03	73.06	2662	15
14	2505	73.02	73.07	5854	17
16	2843	73.01	73.08	4284	14

New order: Buy 100 @ 73.05

LastTime	LastVolume	Last			
10:20:10.352	100	73.05			
BidCount	**BidVolume**	**Bid**	**Ask**	**AskVolume**	**AskCount**
3	624	73.04	73.05	**2981**	**4**
10	1474	73.03	73.06	2662	15
14	2505	73.02	73.07	5854	17
16	2843	73.01	73.08	4284	14



Ask	AskVolume	QueuePos
73.05	100	1
73.05	791	2
73.05	1368	3
73.05	685	4
73.05	137	5

Example:

We want to buy 100 shares at a price of 73.05

Q: Which of the 5 sell orders is “matched” with our buy order?

A: The sell order which entered the order book the **earliest**.

Implication:

The faster you are, the higher the probability of an order being executed at the desired price. **True for both sides of the trade** (buyer and seller)

https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-1-141016
https://www.eurexchange.com/exchange-en/find/news/The-art-of-designing-markets.-Part-II--141112

Matching Algorithms

Price Time

- Priority is determined by price then **time**
- FIFO queue per price level
- Higher queue priority increases probability of being matched
- The most common matching algorithm

Other matching algorithms

- Pro-Rata (allocation based on volume)
- Auction models. Collect orders over a period of time, then match buys and sells at equilibrium price.

13

Asset Prices

Asset Prices

The efficient-market hypothesis (EMH) says that asset prices fully reflect all available information.

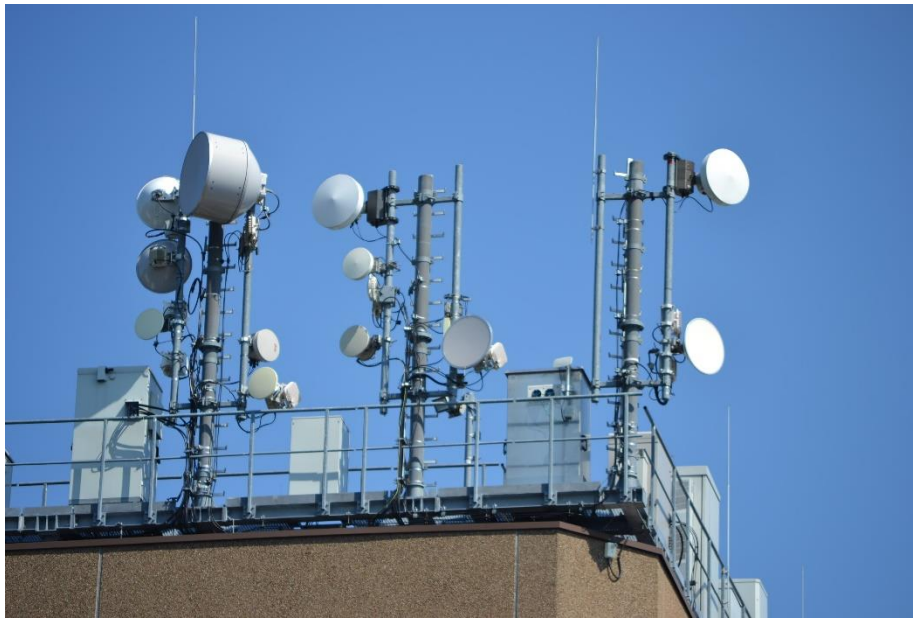
The EMH has critics, but we can agree on that new information (events) affect prices.

Examples:

- News
- Changes in interest rates
- Release of economic indicators (e.g. unemployment figures, non-farm payrolls)
- The central limit order book itself (e.g. more buyers than sellers)
- Prices of correlated instruments

Event source can be:

- Remote (need to send information over WAN and/or radio links)
- Local (LAN, co-location)

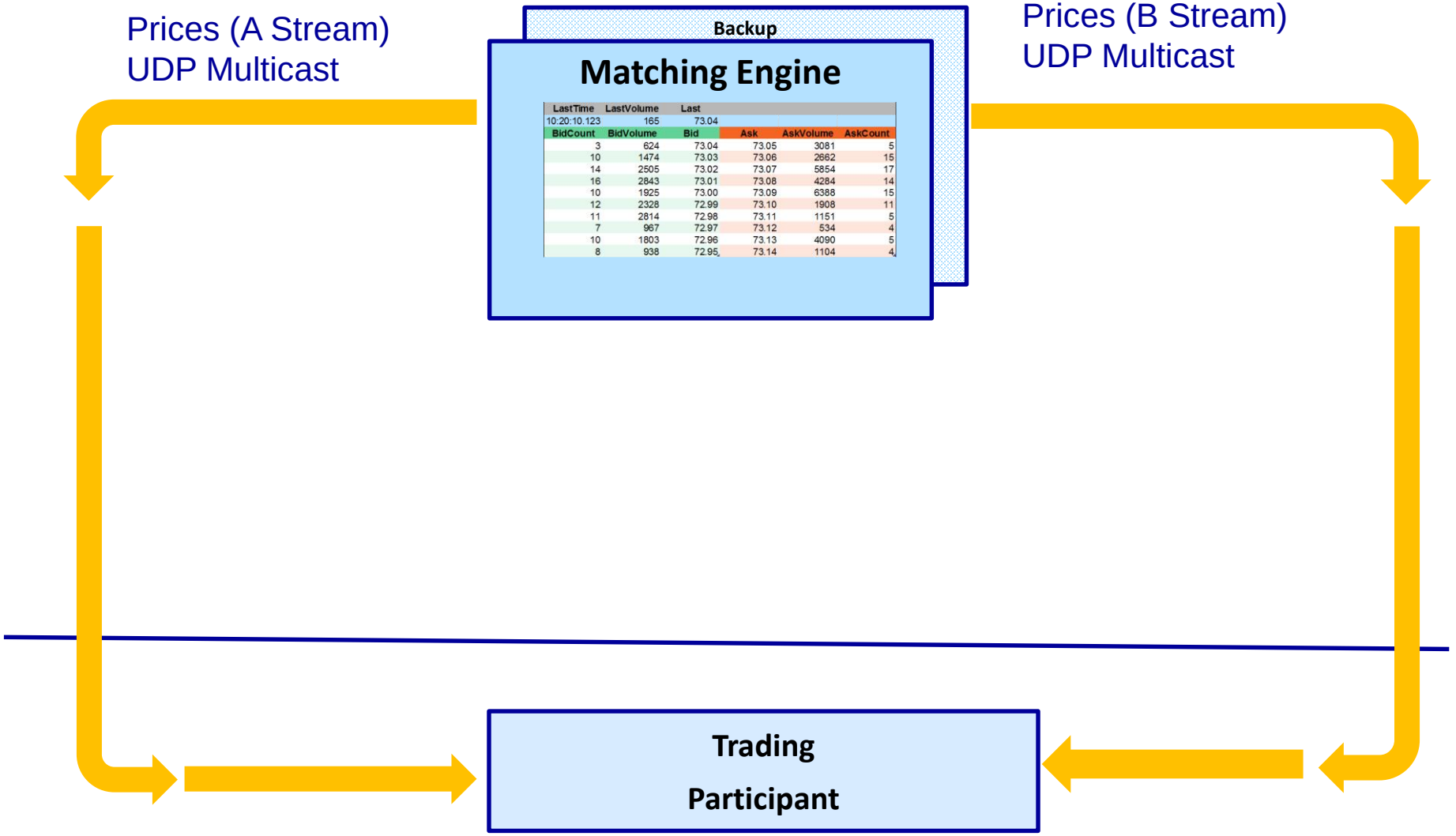


Correlated Asset Prices – Example

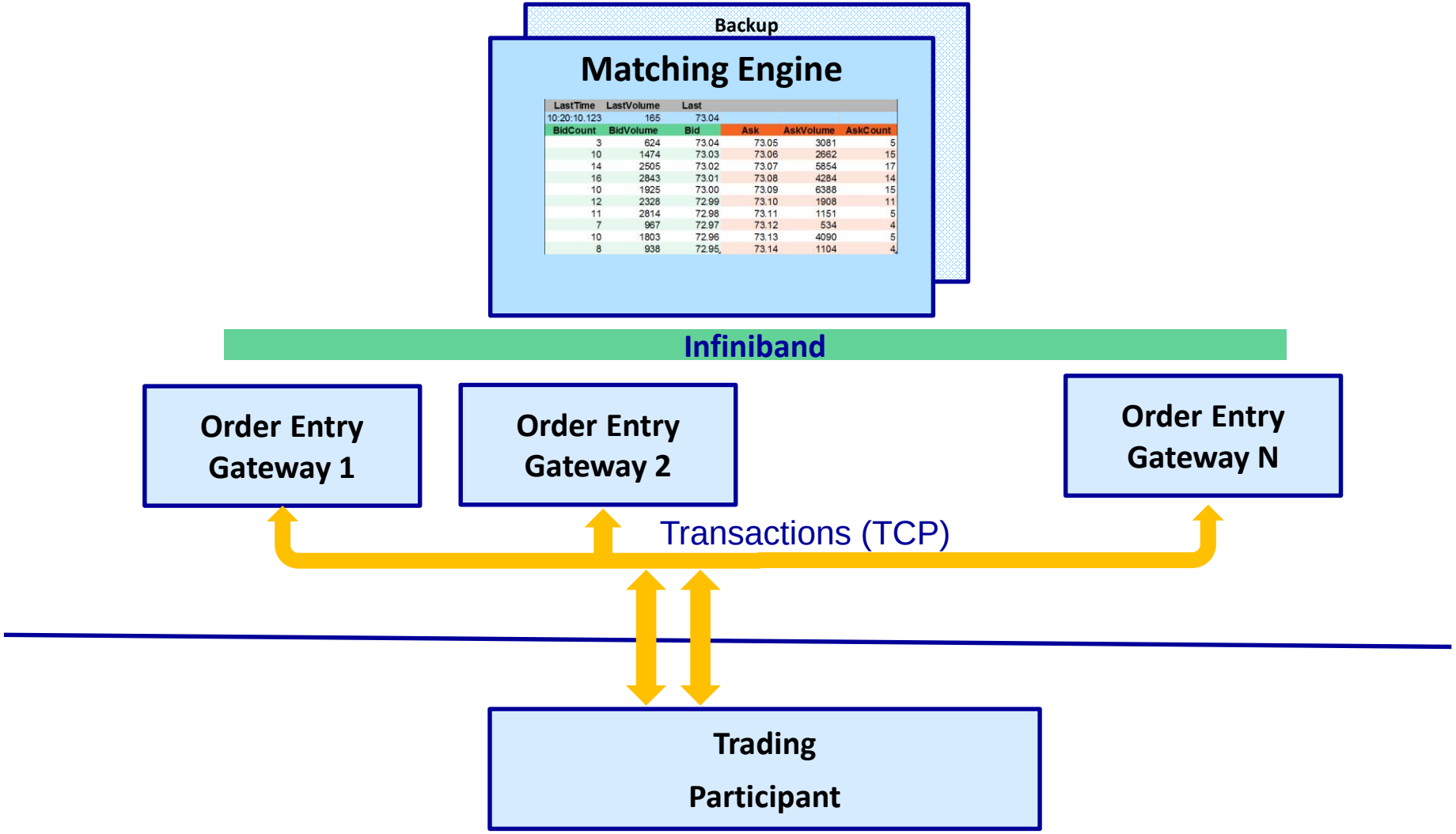
16

T7® Co-Location

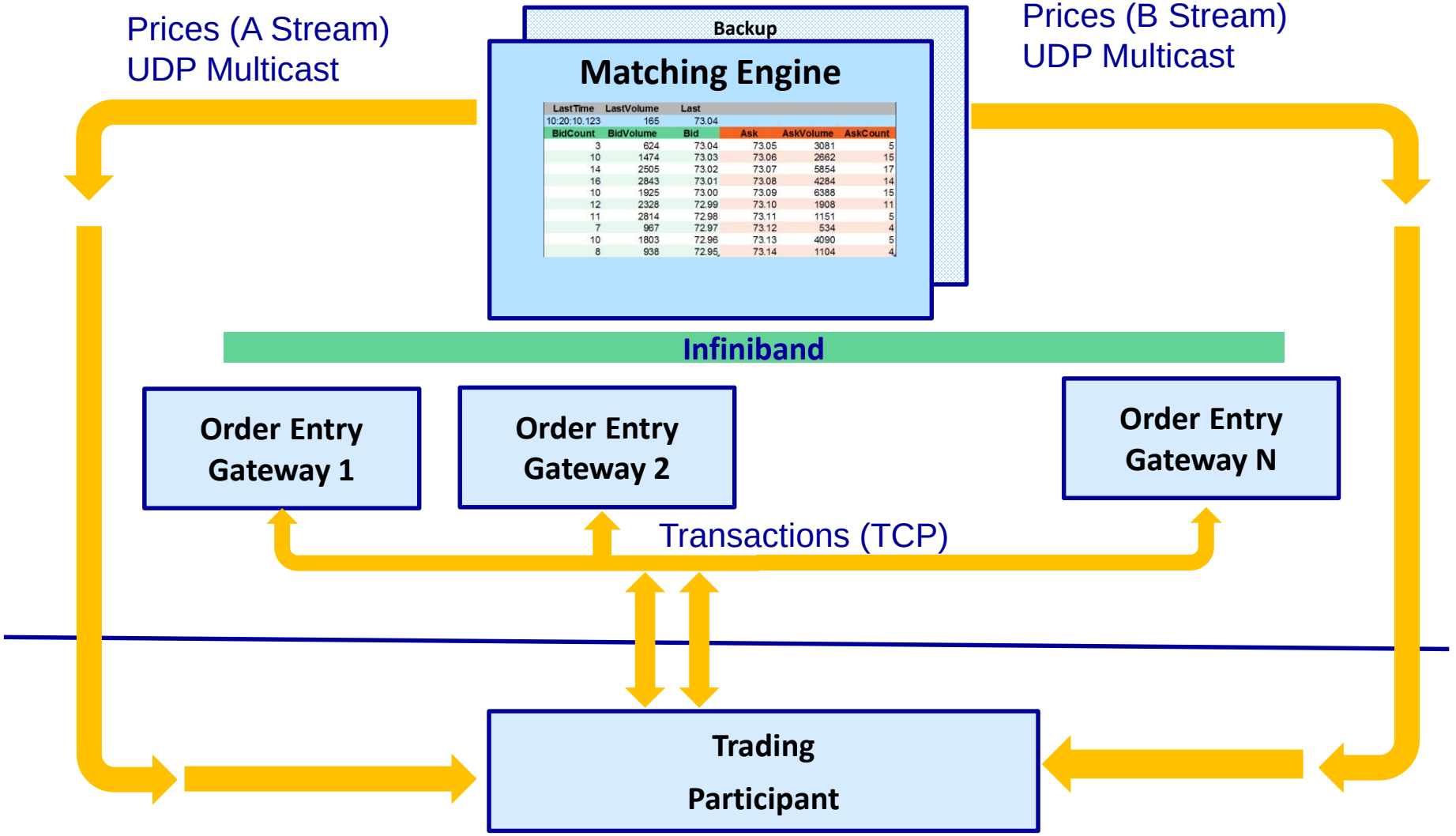
T7® Architecture – Price Information



T7® Architecture – Interaction (placing orders)

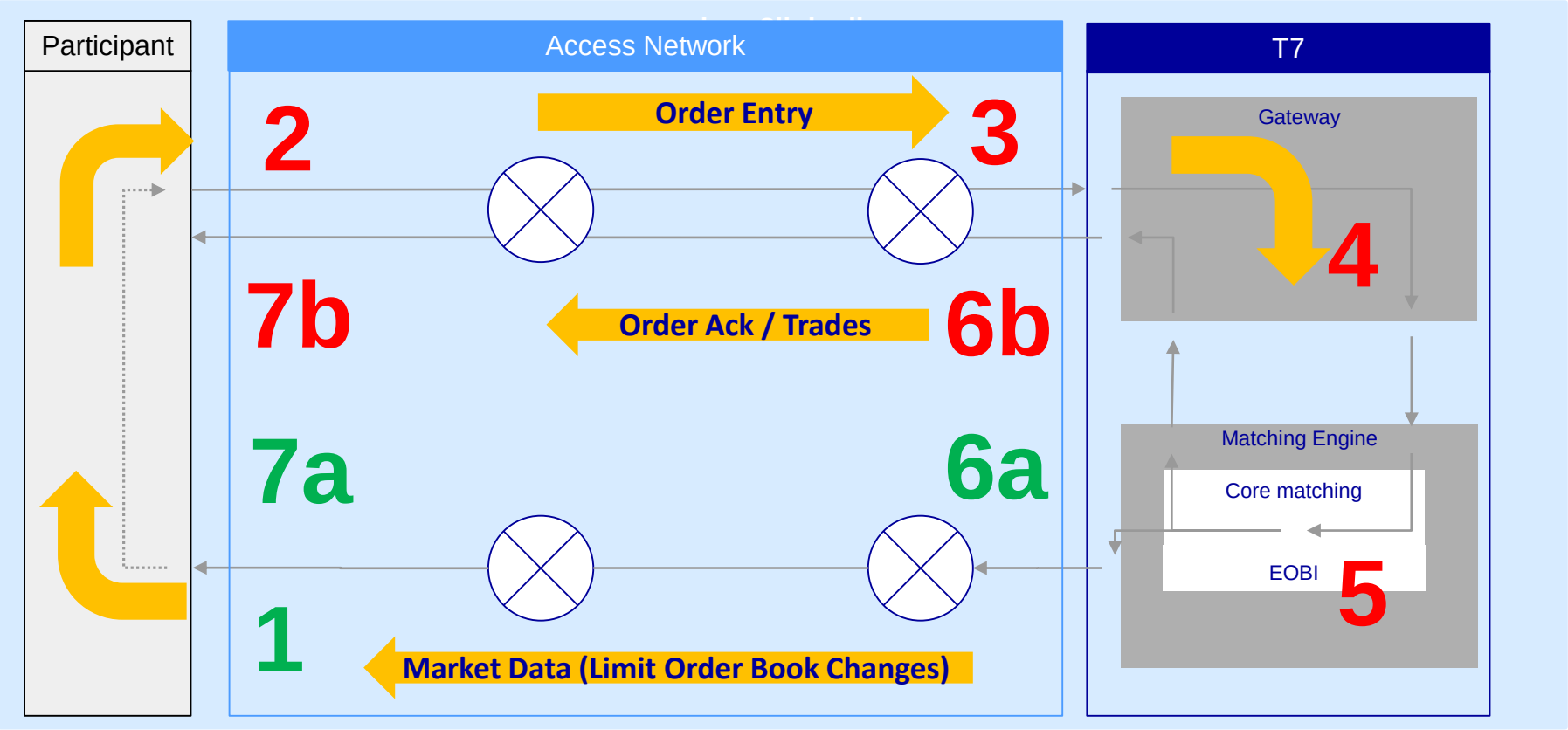


T7® Architecture



T7® Co-location

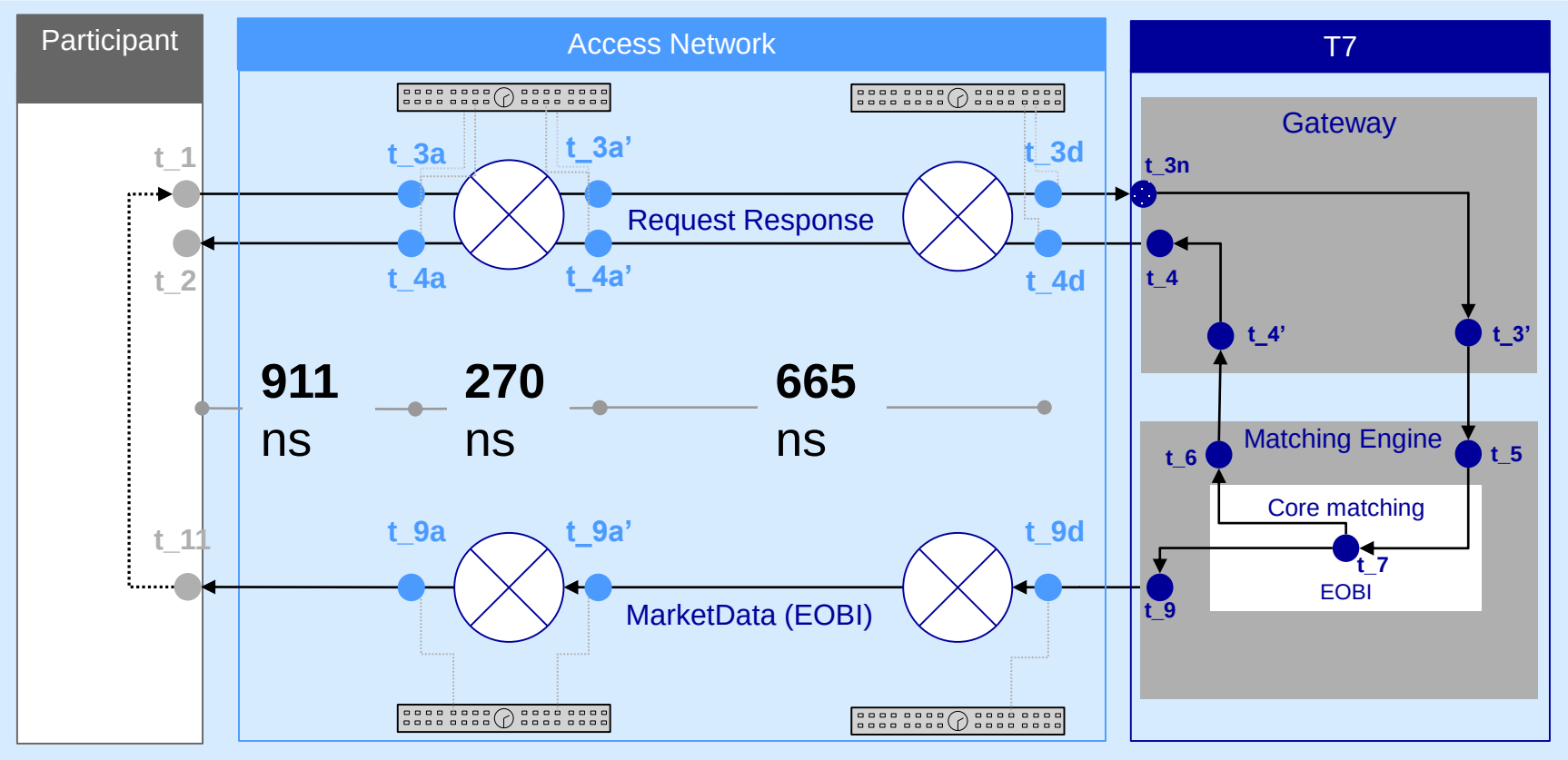
Zoom in to a single matching engine



Median round-trip from order entry to acknowledgement ≈ **60µs** (2) ➔ (7)
The fastest participants have **sub 100ns** response times (1) ➔ (2)

T7® Co-location

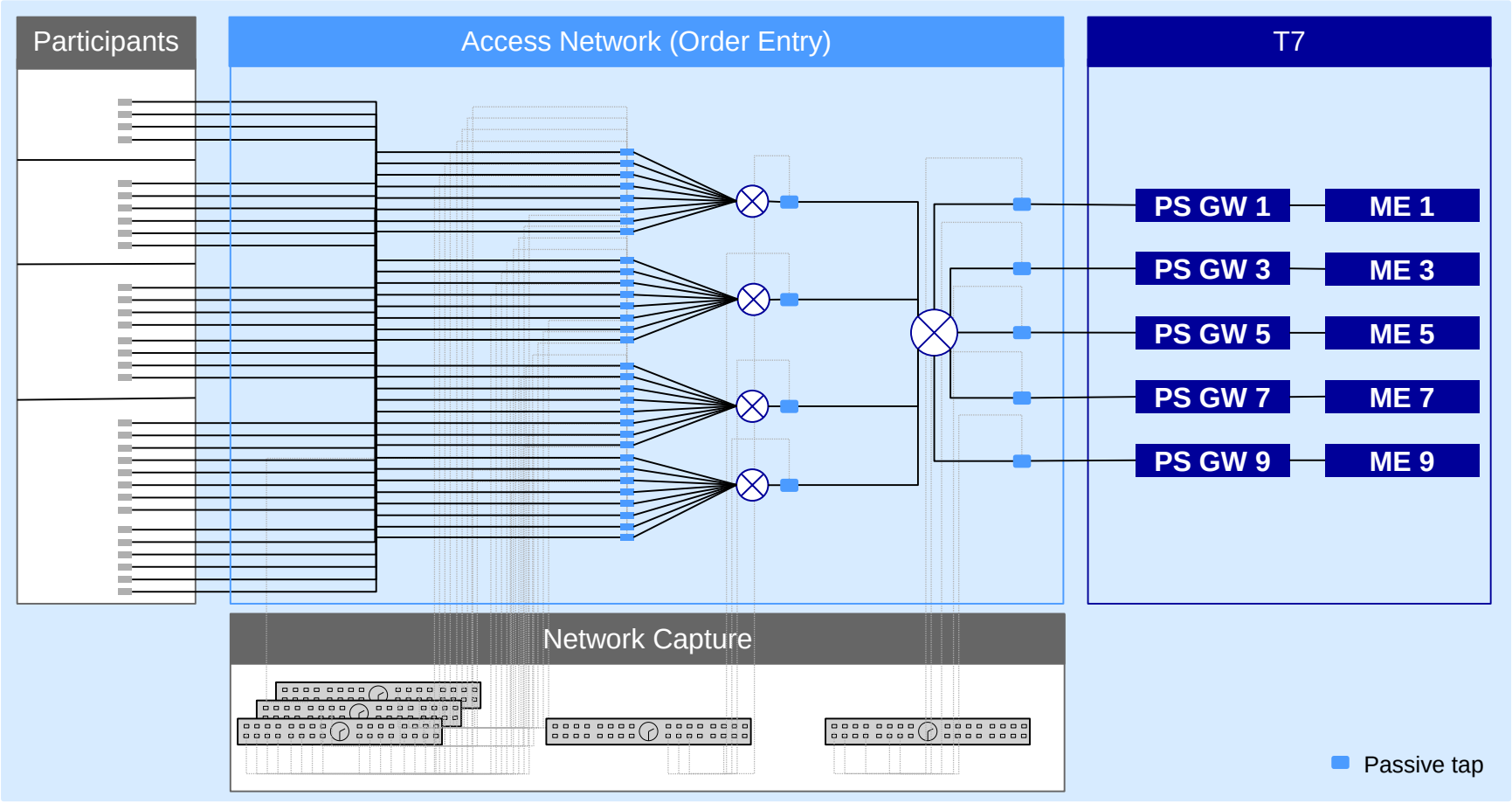
Timestamps



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

T7® Co-location

Scale (> 500 capture ports, > 60 timestamping devices)



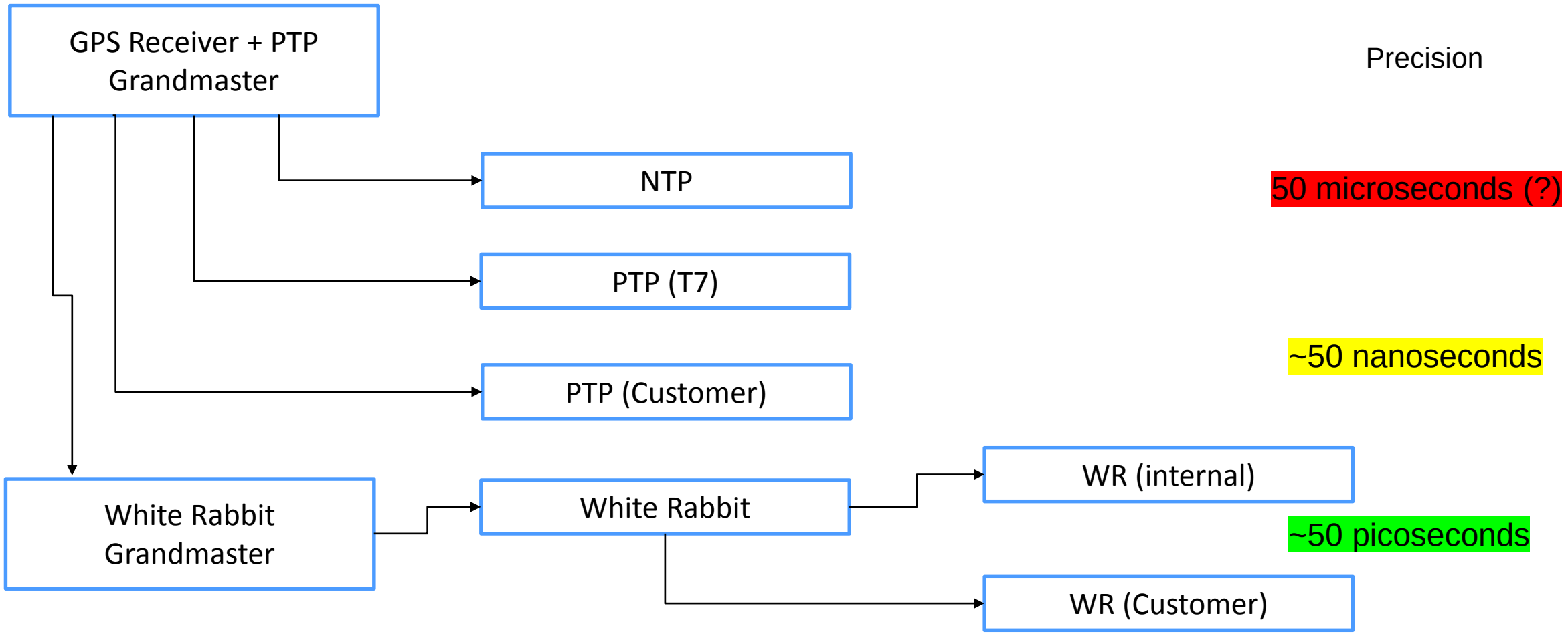
> 260 Order Entry lines captured (> 500 capture ports)

Identical setup regardless of participant room location and assigned access switch

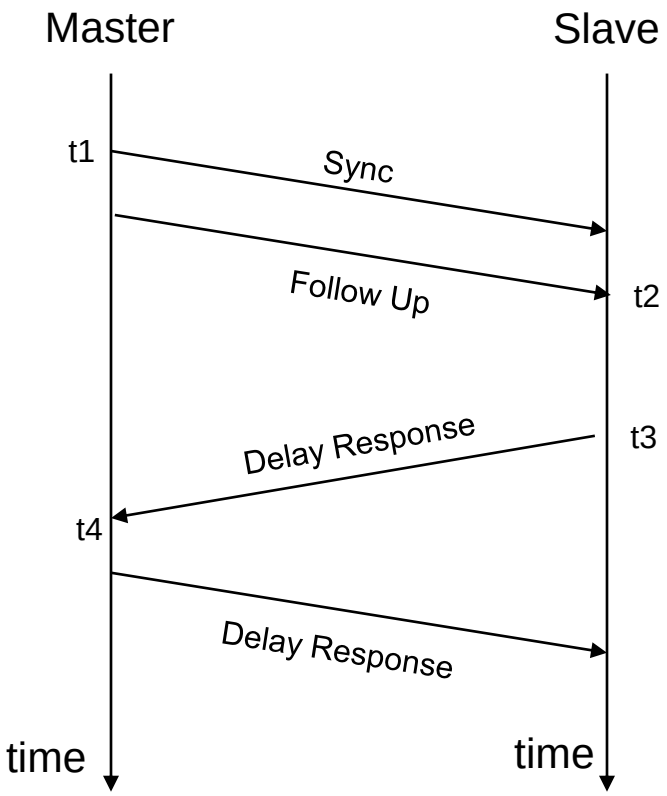
23

Time Distribution in T7

Time Distribution in T7



PTP v2 aka IEEE 1588-2008



- Slave has to know its offset from master (one-way delay)
- Calculates this with 4 timestamps

Path Delay = ((t4 – t1) – (t3 – t2)) / 2

Offset from Master = (t2 – t1) – Path Delay

- Works best on dedicated network
- Requires network path symmetry
- BMCA Best Master Clock Algorithm to select grandmaster

PTP v2 aka IEEE 1588-2008

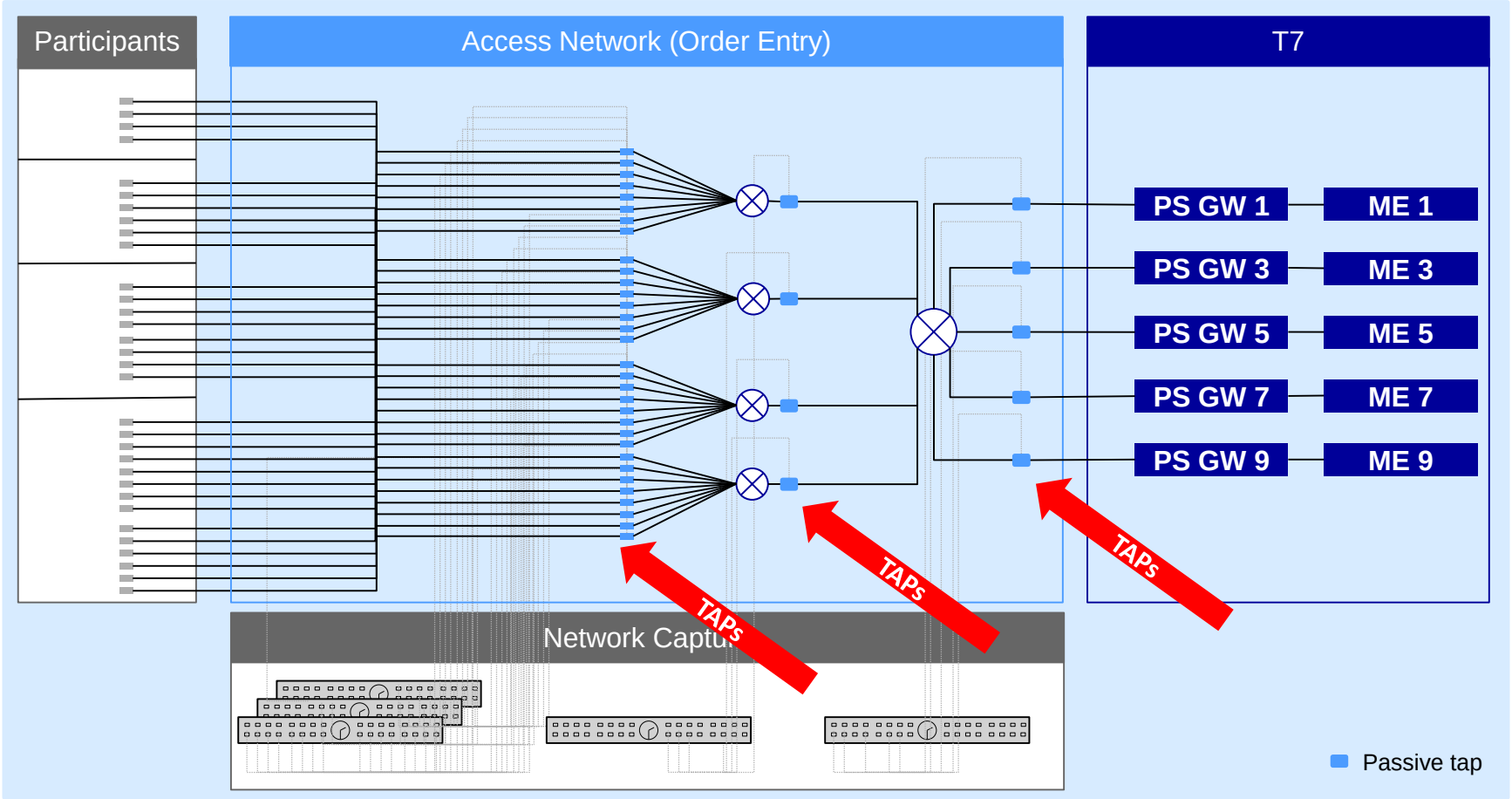
PTP Precision in Practice

Problem Statement – Why PTP is not good enough

How to capture and timestamp all customer cross-connects in co-location?

- 500+ capture ports
- 60+ capture devices
- 4 datacentre modules

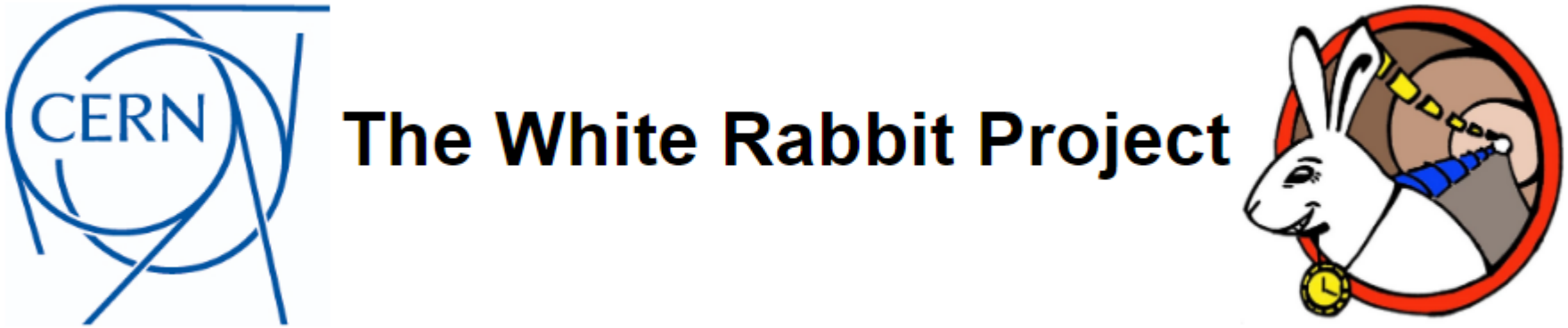
- PTP +/- 50ns jitter in our infrastructure at best
- Serialization time of order entry message = 120ns
- Goal of sub-10ns precision
- Distances too long for PPS over coax cables



28

White Rabbit at Deutsche Börse

White Rabbit



White Rabbit is a fully deterministic Ethernet-based network for general purpose data transfer and synchronization. It can synchronise over 1000 nodes with sub-ns accuracy over fibre lengths of up to 10 km.

White Rabbit provides **sub-nanosecond accuracy** and **picoseconds precision** of synchronisation for large distributed systems. It also allows for deterministic and reliable data delivery. White Rabbit allows you to precision time-tag measured data and lets you trigger data taking in large installations while at the same time using the same network to transmit data.

http://white-rabbit.web.cern.ch

http://www.ohwr.org/projects/white-rabbit/wiki

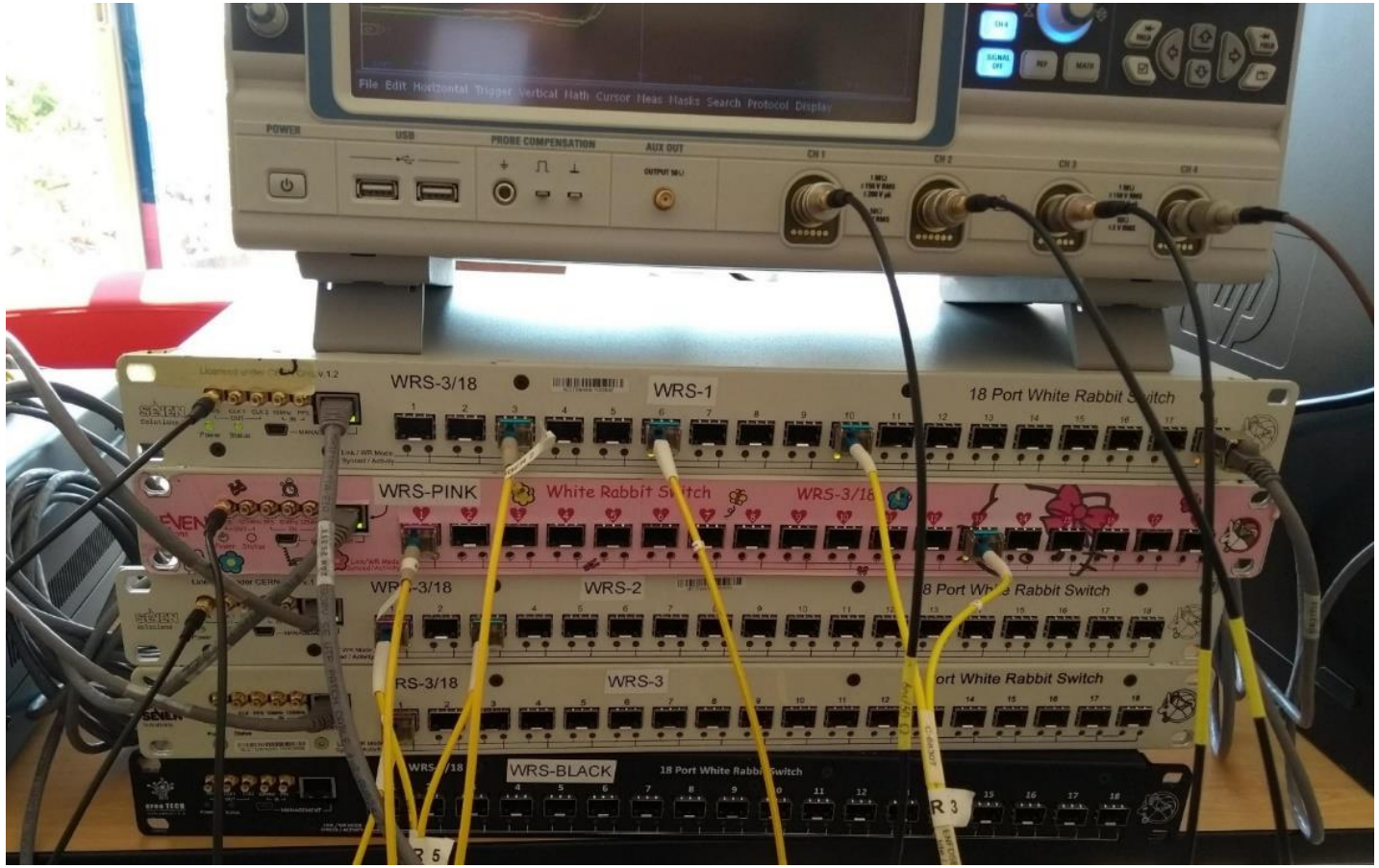
Used by CERN (Switzerland), GSI (Germany), Nikhef (Netherlands), and many more …

White Rabbit

- Came out the of the University of Warsaw
- Developed at CERN
- Provides sub-nanosecond accuracy and picoseconds precision of synchronization
- Tried and tested
- PTP over Synchronous Ethernet
- “High Accuracy” profile of the PTP standard IEEE1588-2019
- Commercially available (except the Hello Kitty model)

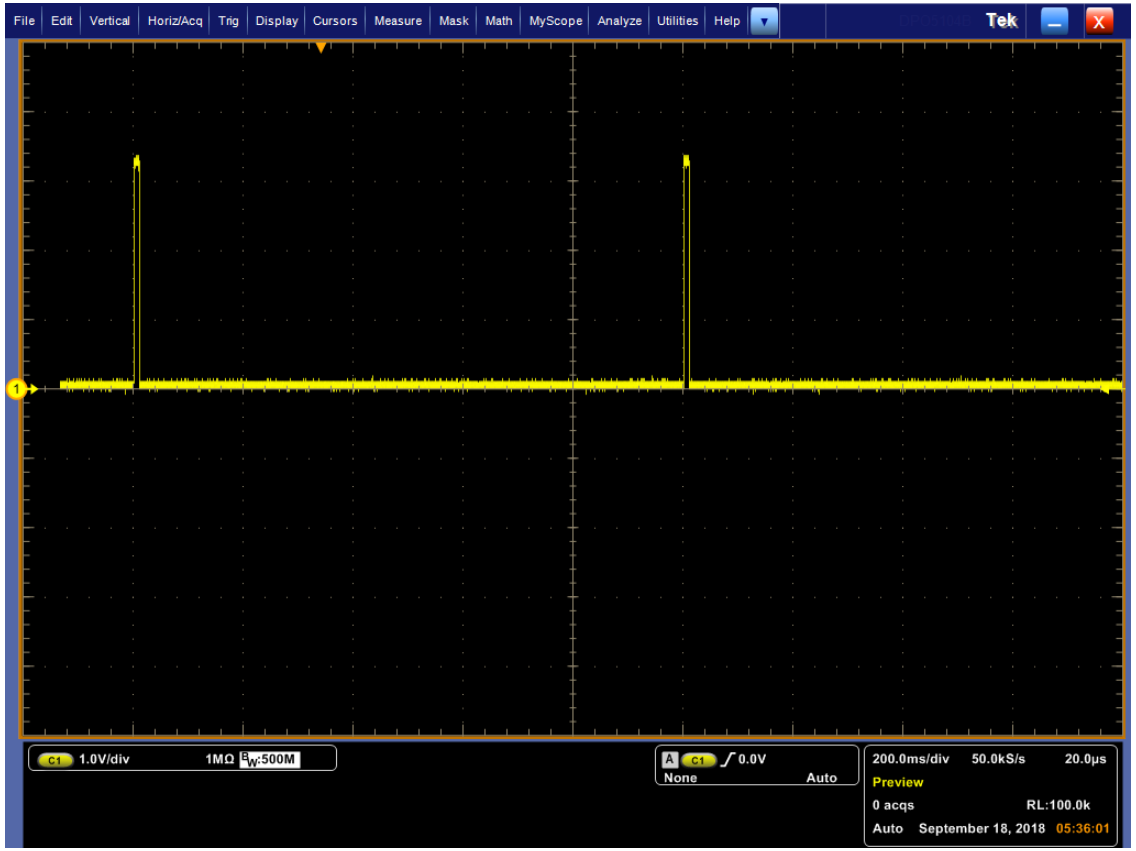
Issue for us:
No native White Rabbit support in NICs and switches

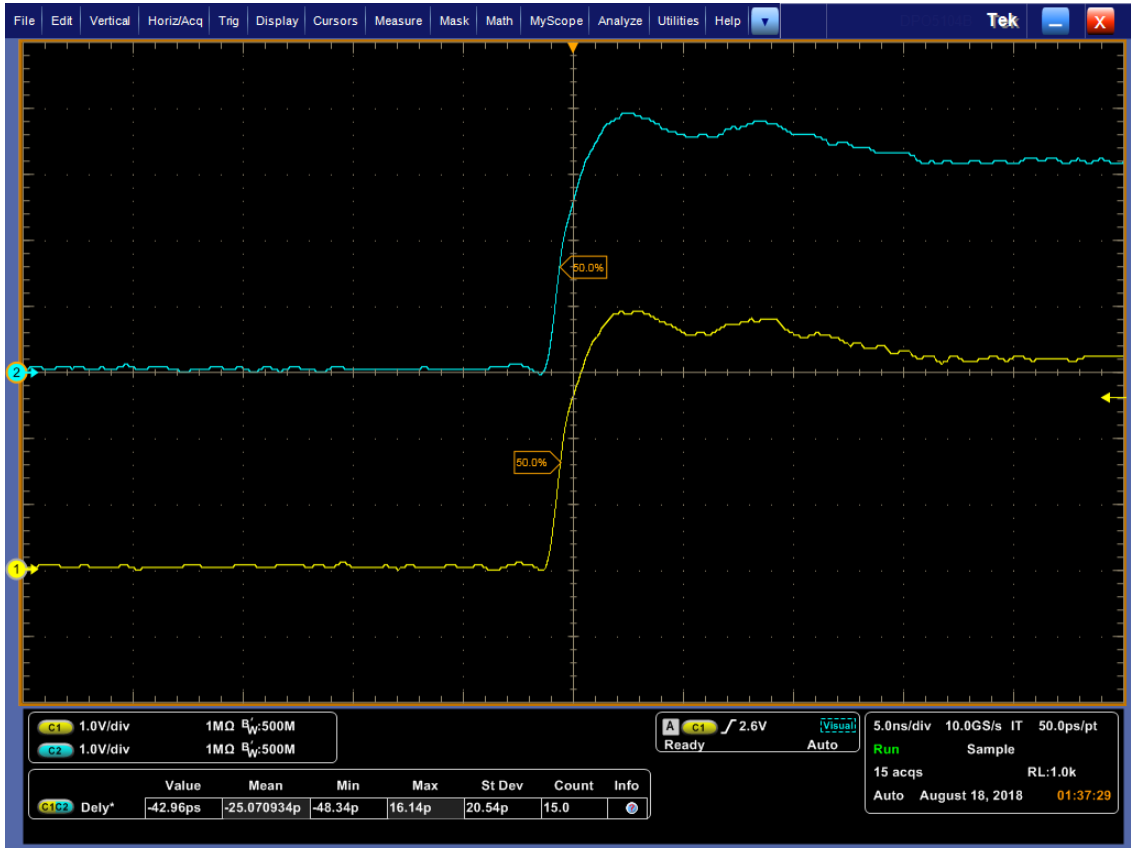
Solution:
We use White Rabbit to distribute 1PPS



White Rabbit

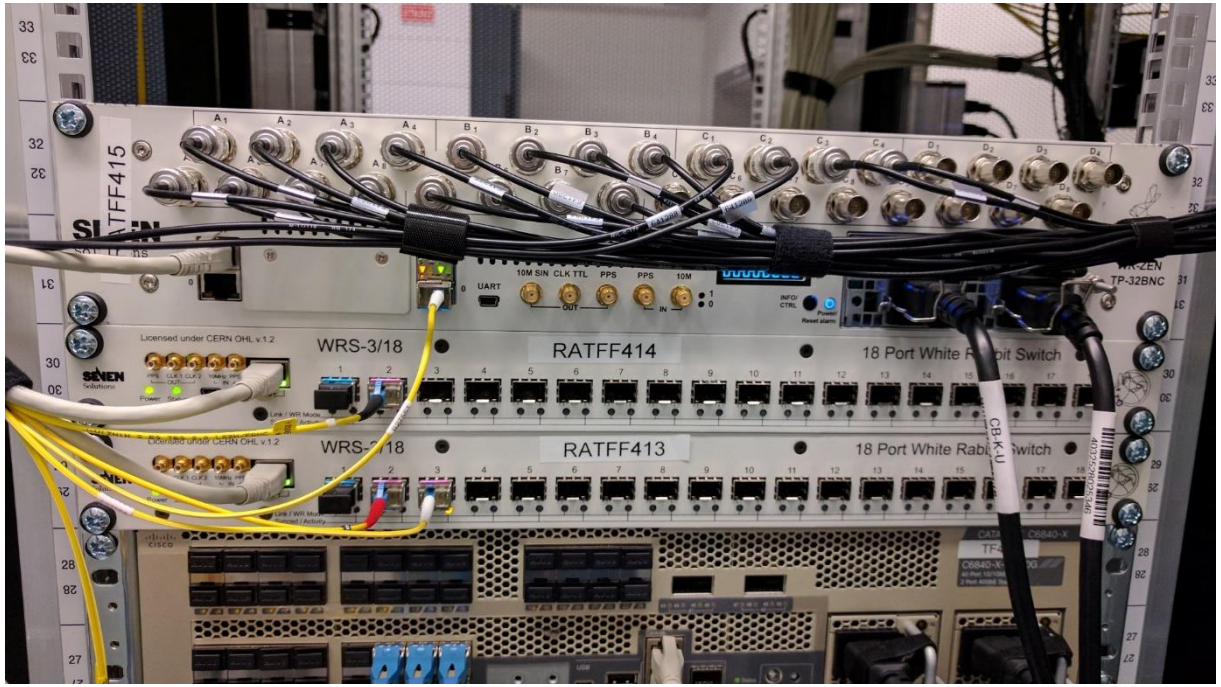
What is PPS?





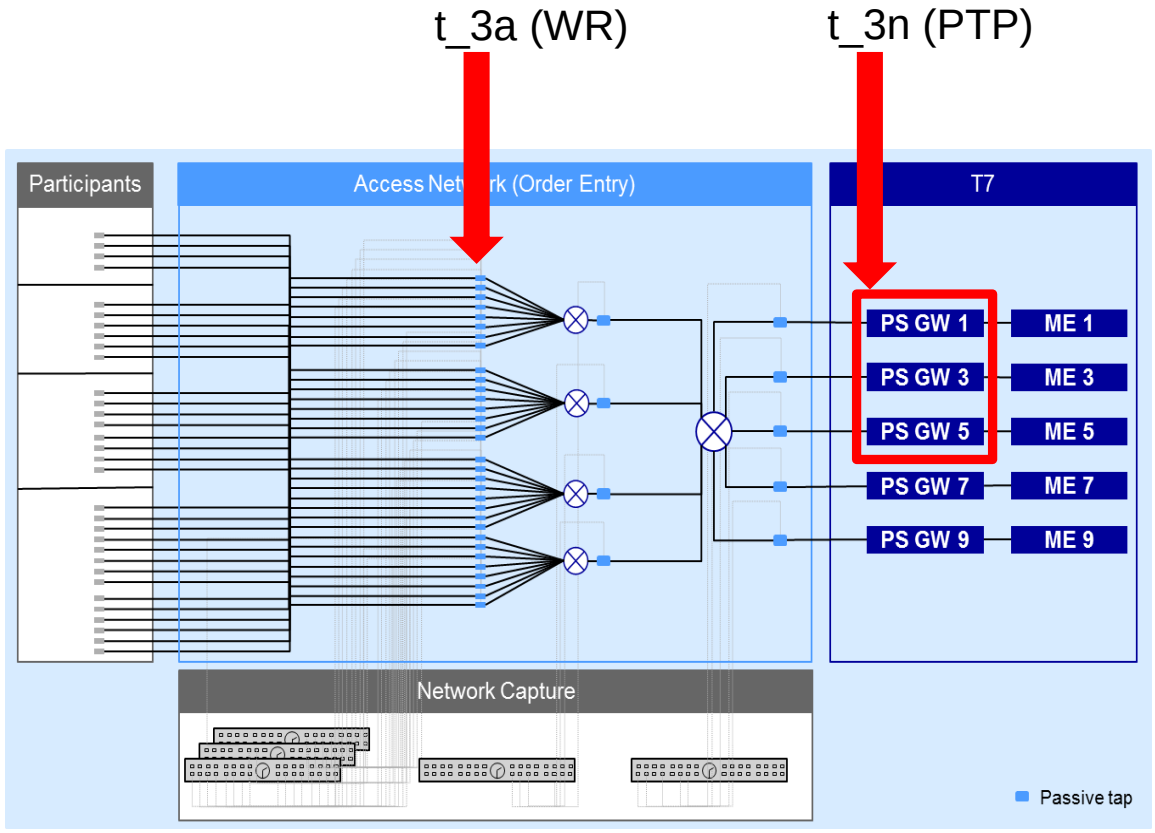
White Rabbit deployed hardware

Timestamping devices synchronized by 1PPS over White Rabbit



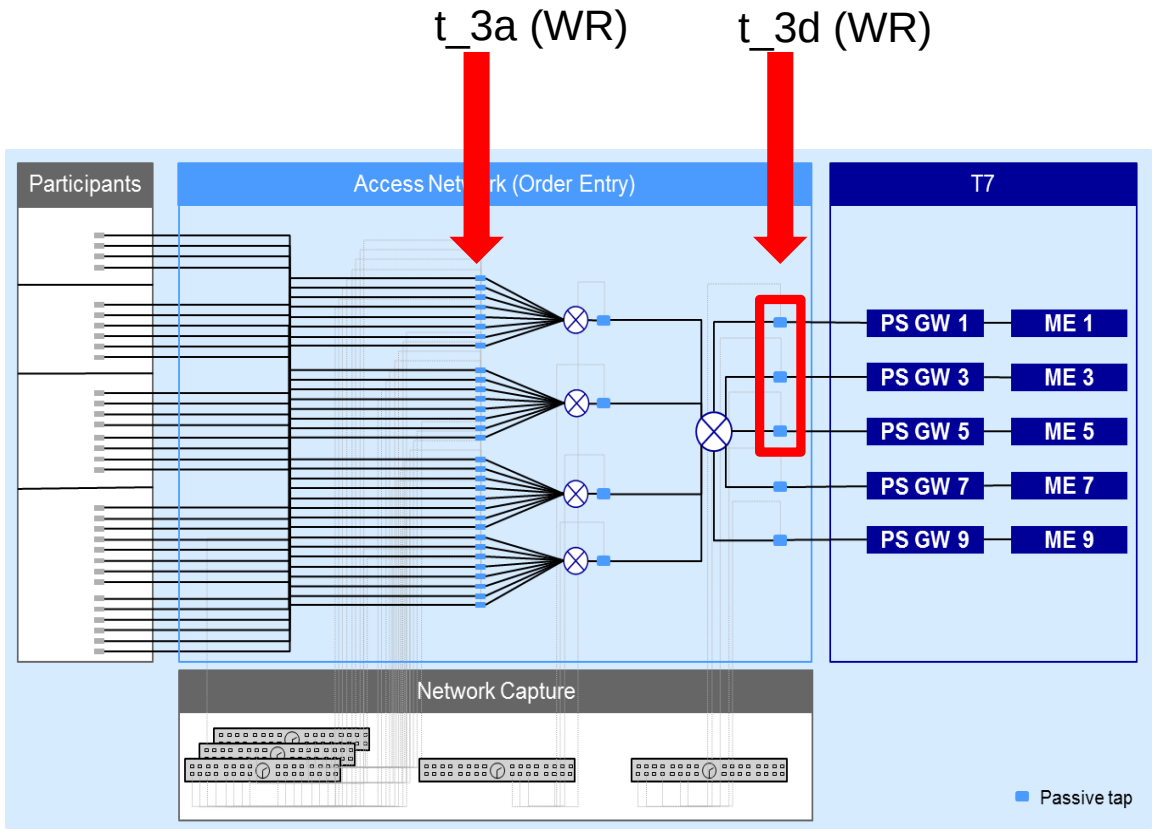
Time Synchronisation

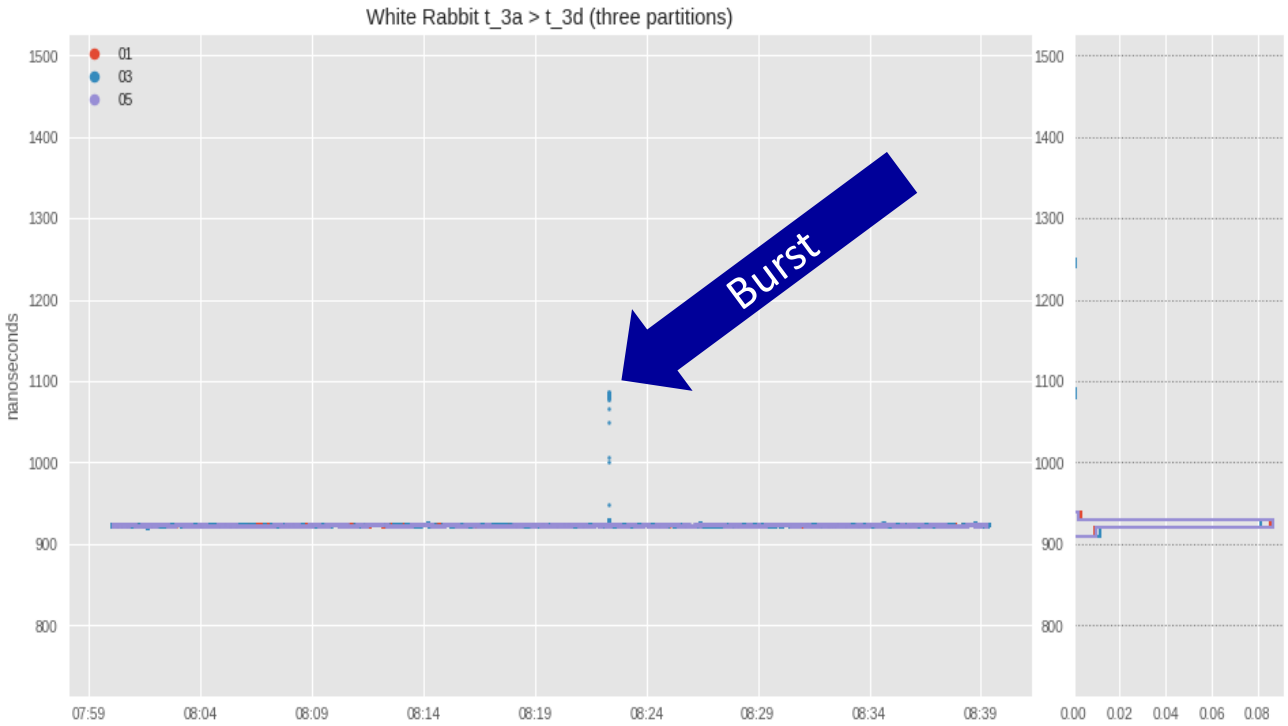
PTP



Time Synchronisation

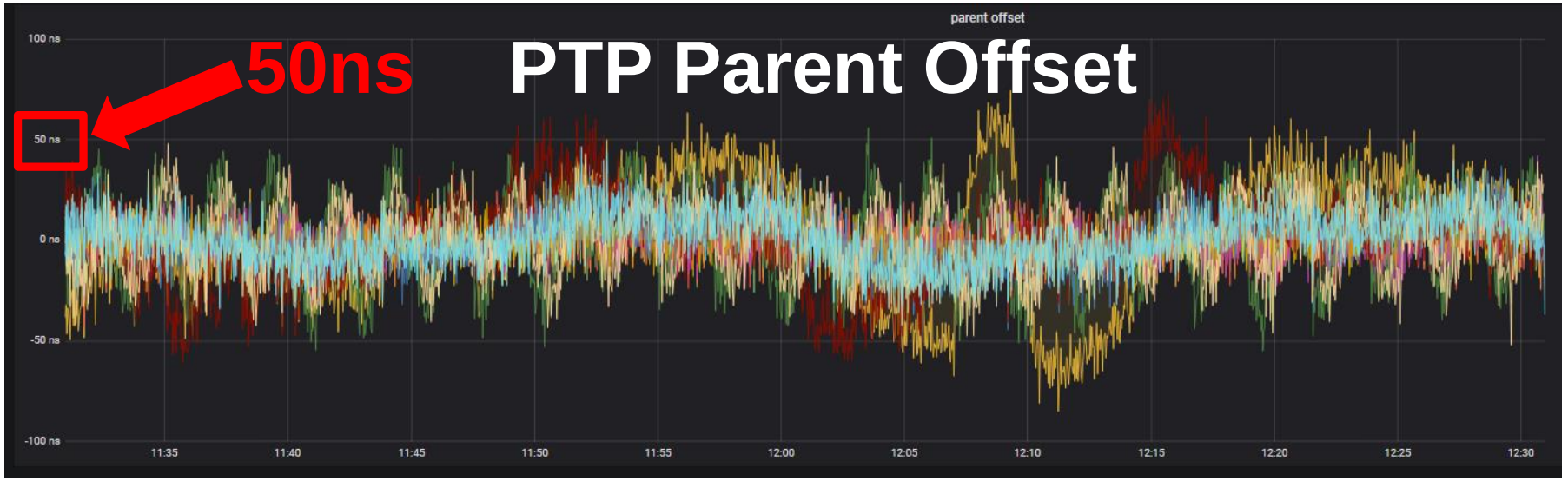
White Rabbit



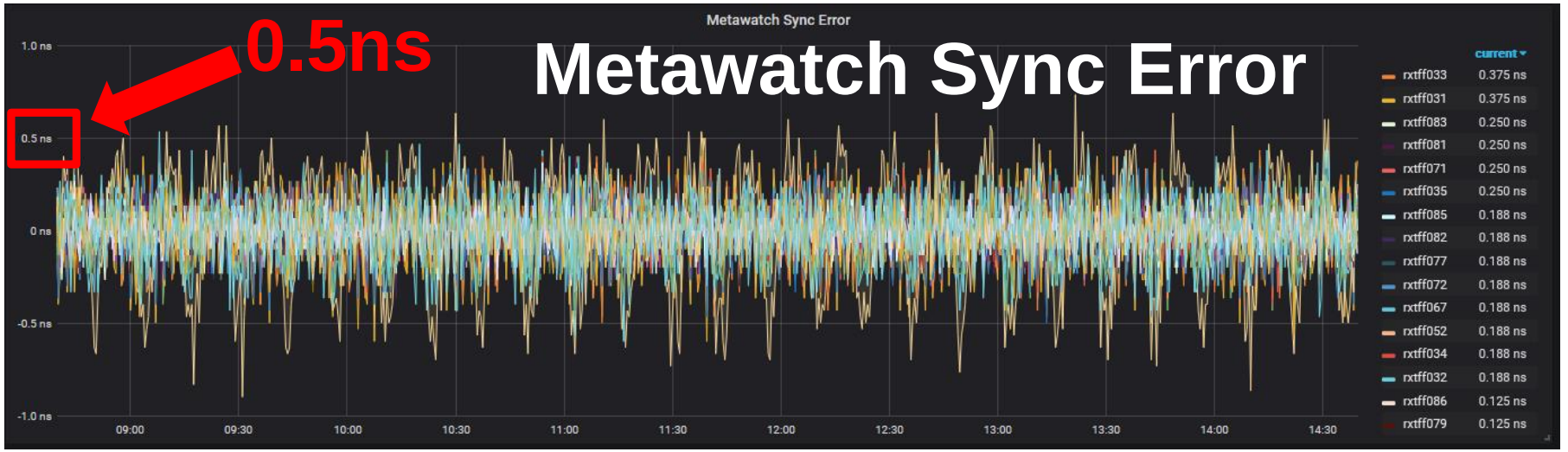


PTP vs White Rabbit

PTP



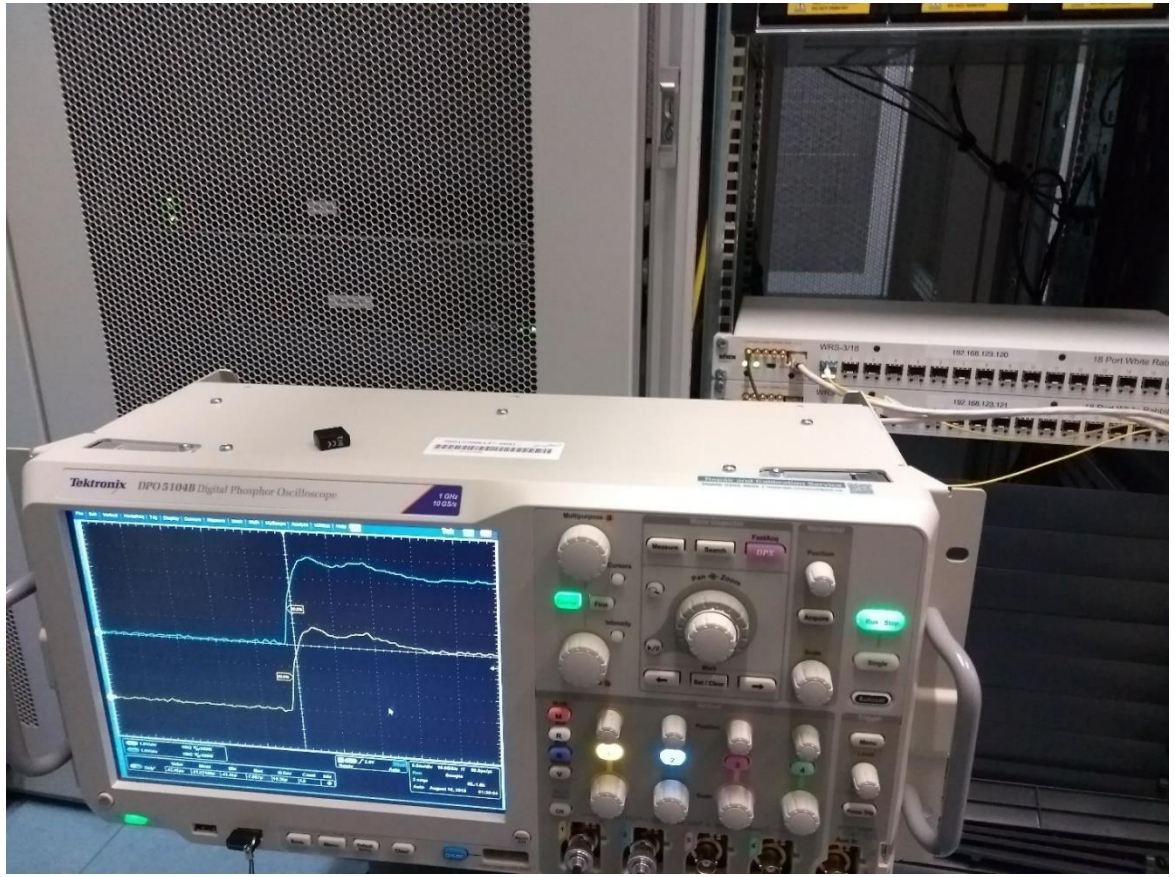
White Rabbit

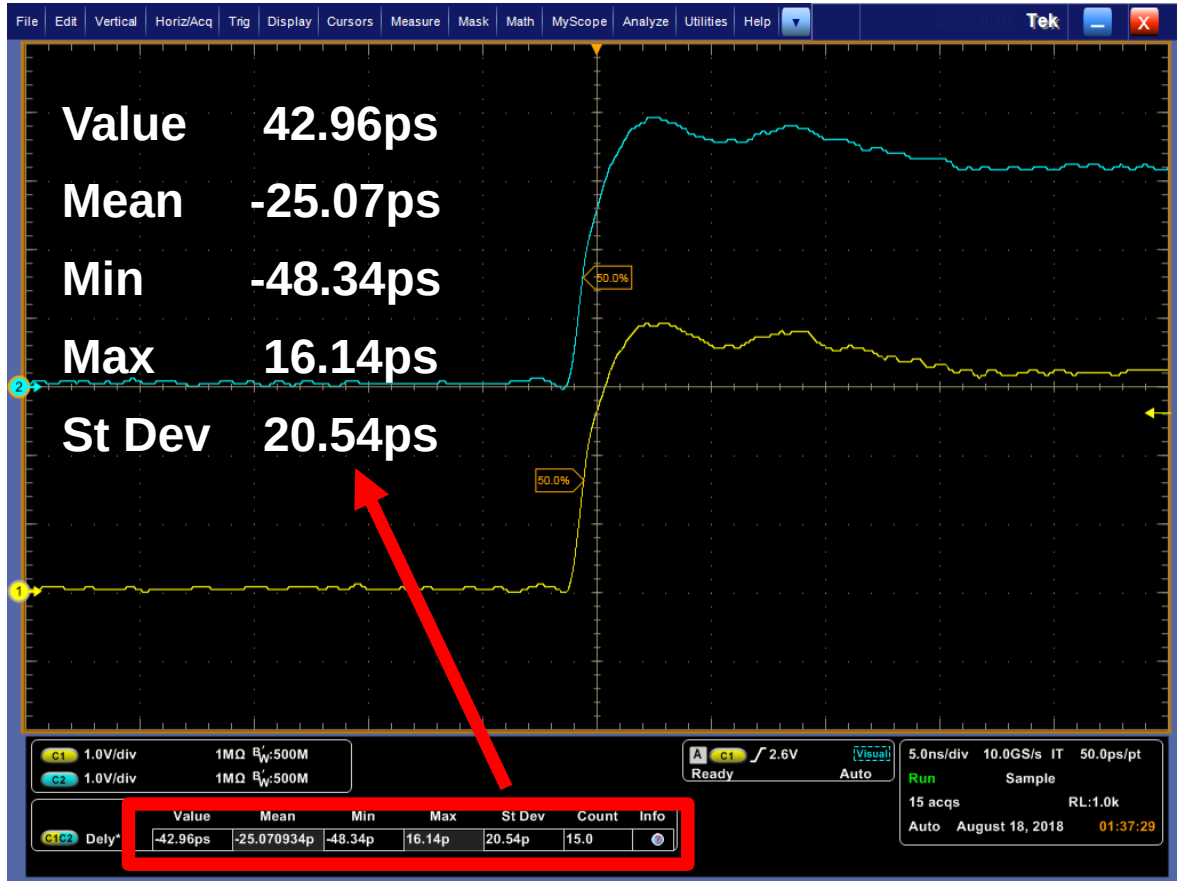


White Rabbit Services for our Trading Customers

Launched two WR based services

1. High-Precision Timestamp (HPT) File (more in a moment)
2. High Precision Time Service (connect to our WR network)





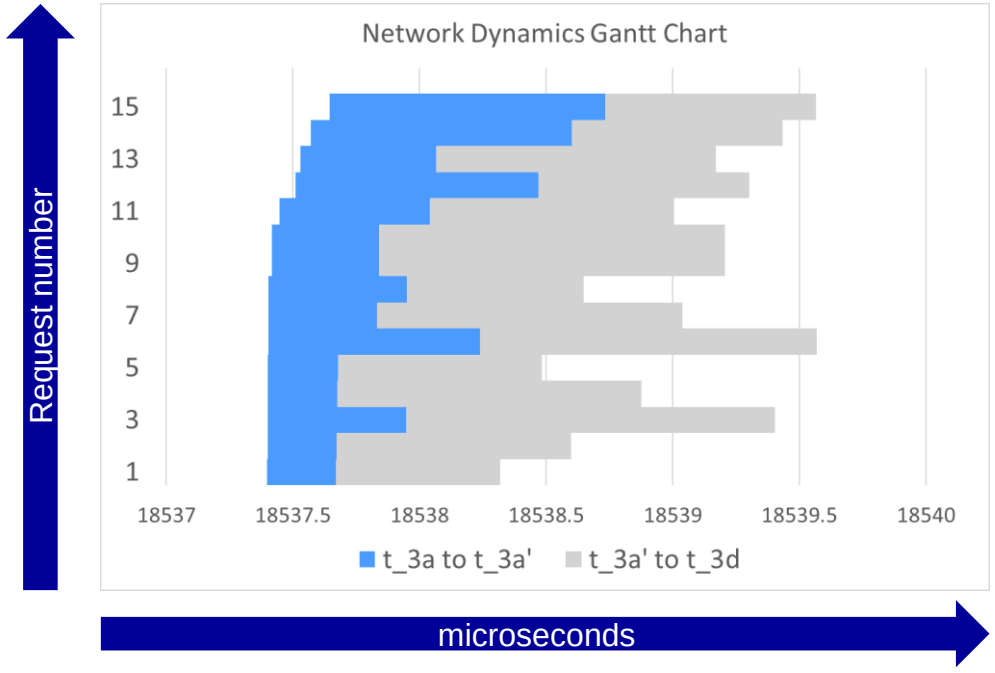
Precise Timestamping

Burst Analysis

Networking at our exchange is characterised by bursts

The most frequently asked question:

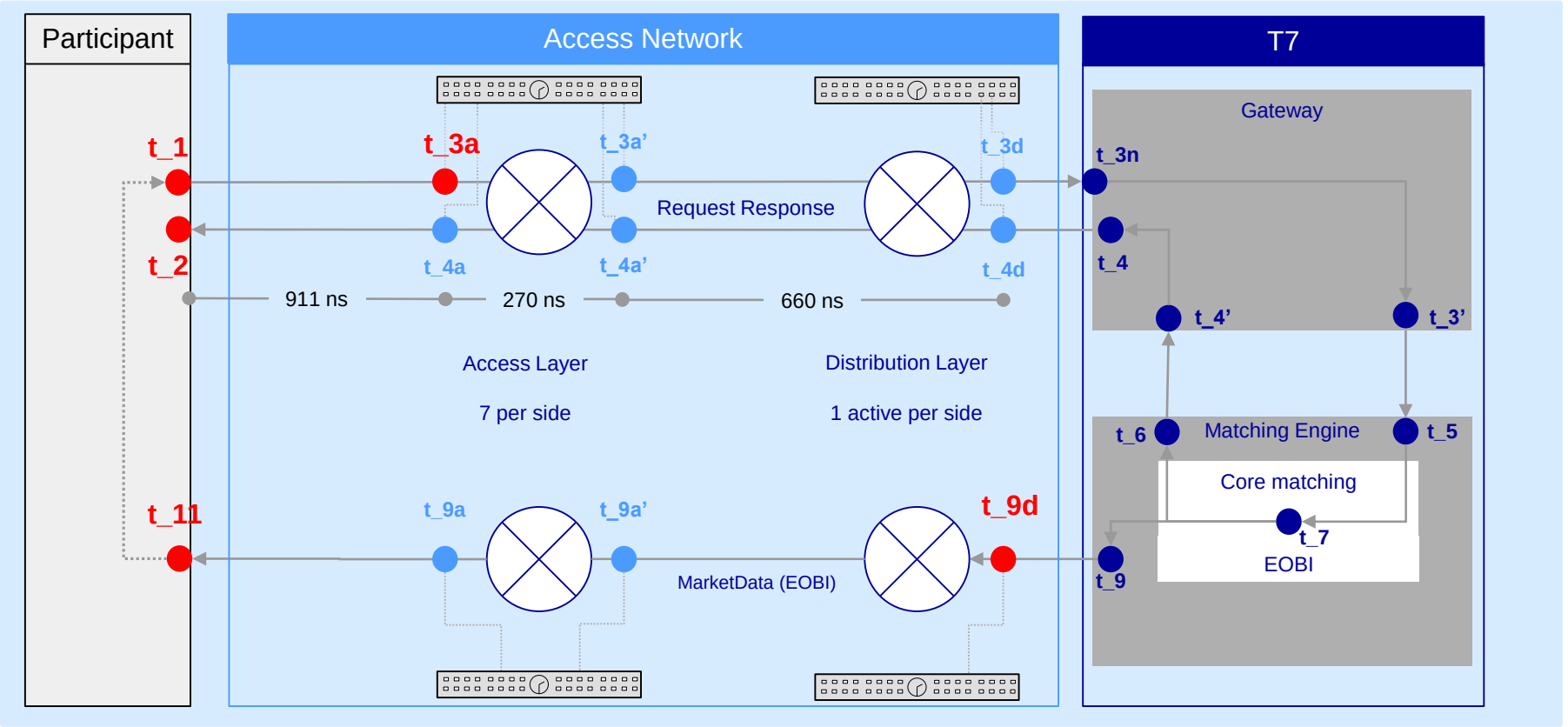
How far behind am I, really?



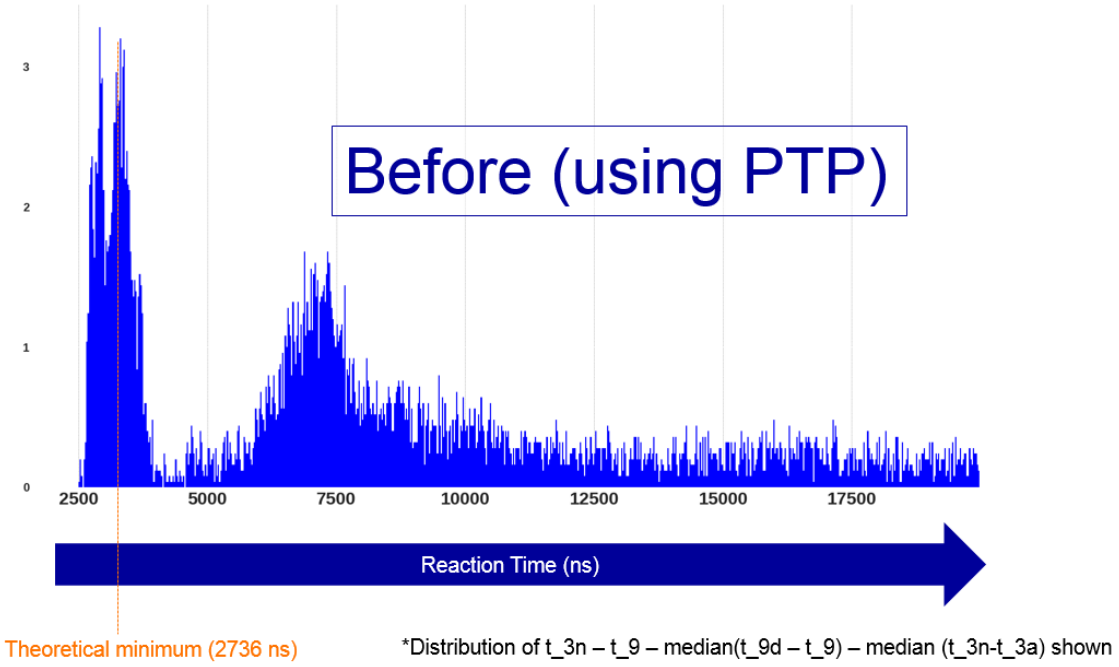
	t_3a (Access In)	t_3a' (Access Out)	t_3a to t_3a'	t_3a' to t_3d	delta.t3a	delta.t3a'	delta.t3d
1	11:05:00.018537384	11:05:00.018537659	275	650			
2	11:05:00.018537398	11:05:00.018537669	271	650	170	14	-4
3	11:05:00.018537399	11:05:00.018537672	273	925	206	1	2
4	11:05:00.018537399	11:05:00.018537947	548	1456	212	0	275
5	11:05:00.018537400	11:05:00.018537676	276	1199	204	1	-272
6	11:05:00.018537401	11:05:00.018537678	277	804	157	1	1
7	11:05:00.018537402	11:05:00.018538237	835	1332	224	1	558
8	11:05:00.018537403	11:05:00.018537832	429	1206	219	1	-406
9	11:05:00.018537404	11:05:00.018537950	546	698	169	1	117
10	11:05:00.018537416	11:05:00.018537839	423	1366	202	12	-123
11	11:05:00.018537416	11:05:00.018537839	423	1366	202	0	0
12	11:05:00.018537446	11:05:00.018538041	595	963	179	30	172
13	11:05:00.018537510	11:05:00.018538468	958	833	150	64	363
14	11:05:00.018537529	11:05:00.018538066	537	1104	161	19	-421
15	11:05:00.018537571	11:05:00.018538602	1031	833	166	42	494
16	11:05:00.018537644	11:05:00.018538734	1090	832	153	73	59
17	11:05:00.018581841	11:05:00.018582110	269	651	156	44197	-821

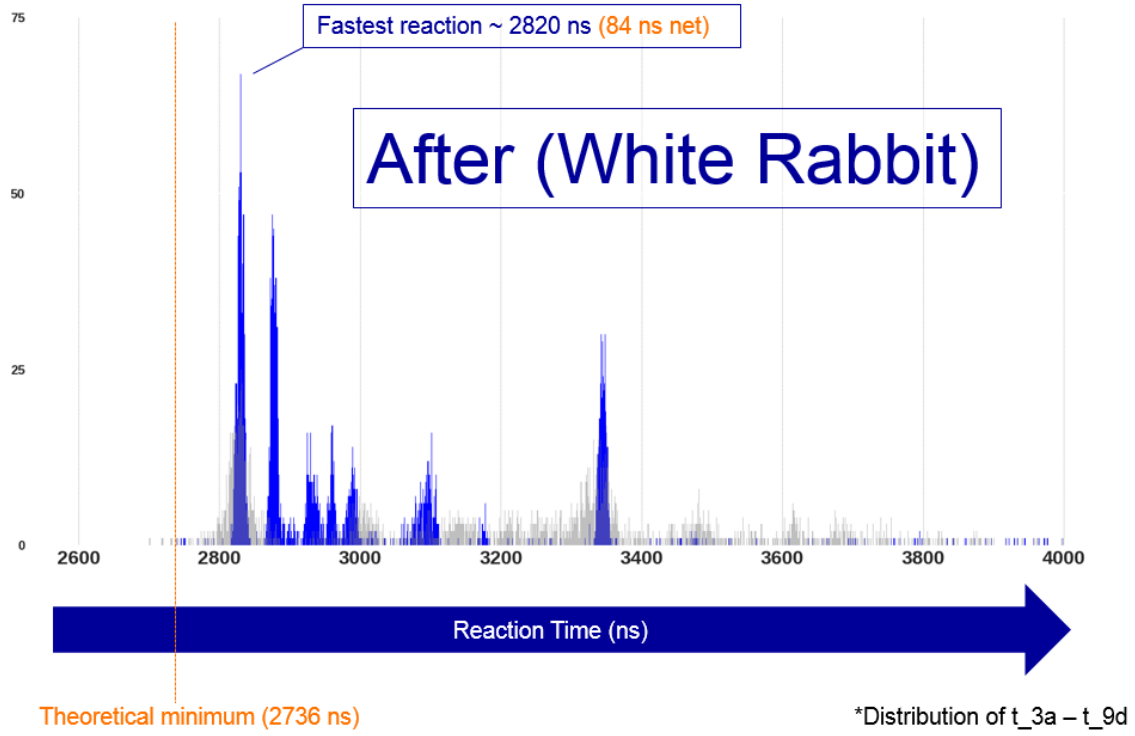
High Precision Timestamp File

For Modeling and Precision Latency Analysis



High Precision Timestamp File





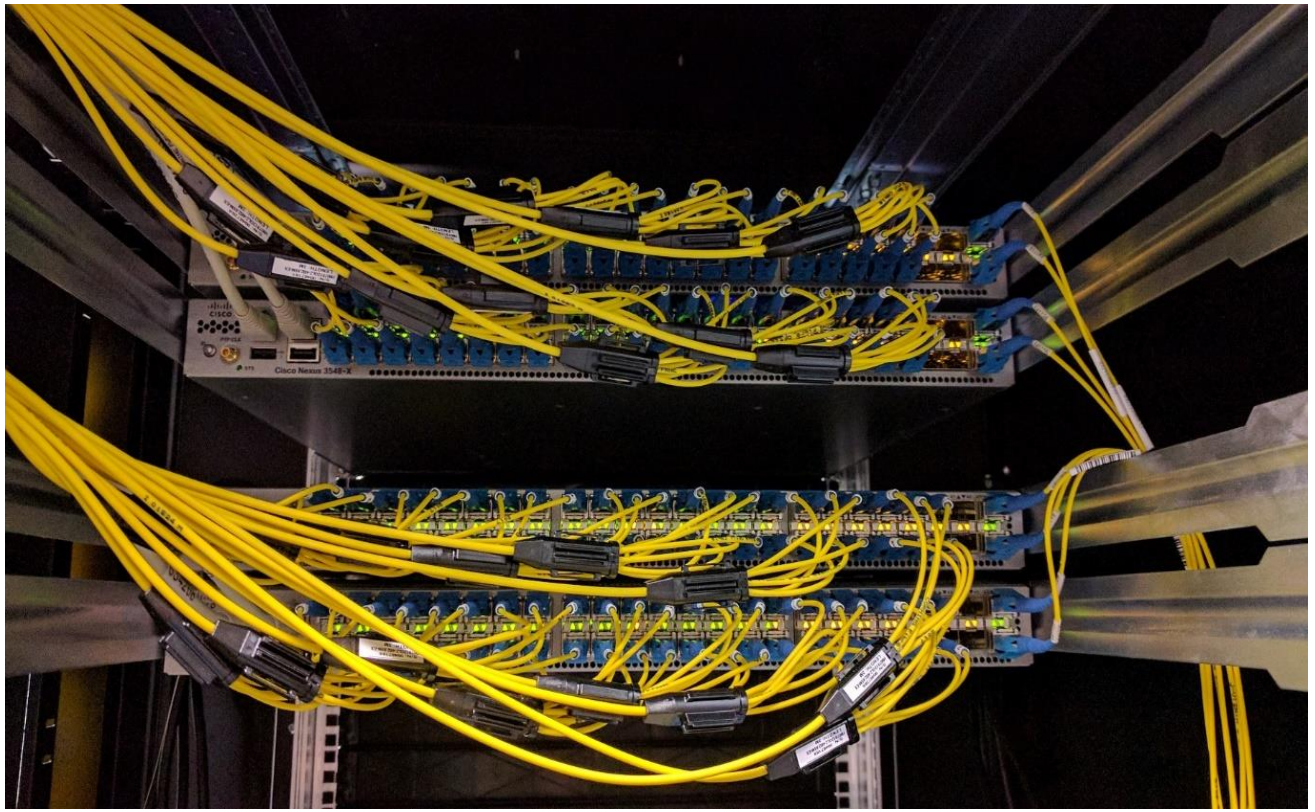
40

Market Data Distribution

Market Data Distribution

UDP Multicast using a Cisco 3548-X

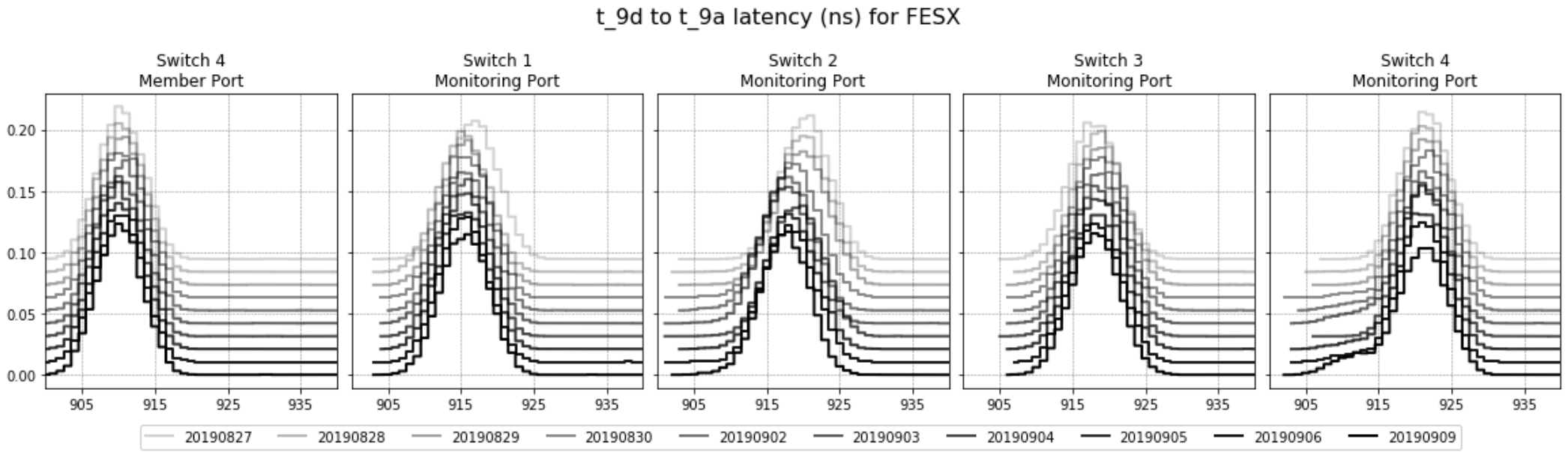
- ASIC designed in 2010
- Introduced in 2011
- Fully featured, low-latency switch which is popular in financial markets
- Cut-through
- ~200ns latency



Market Data Distribution

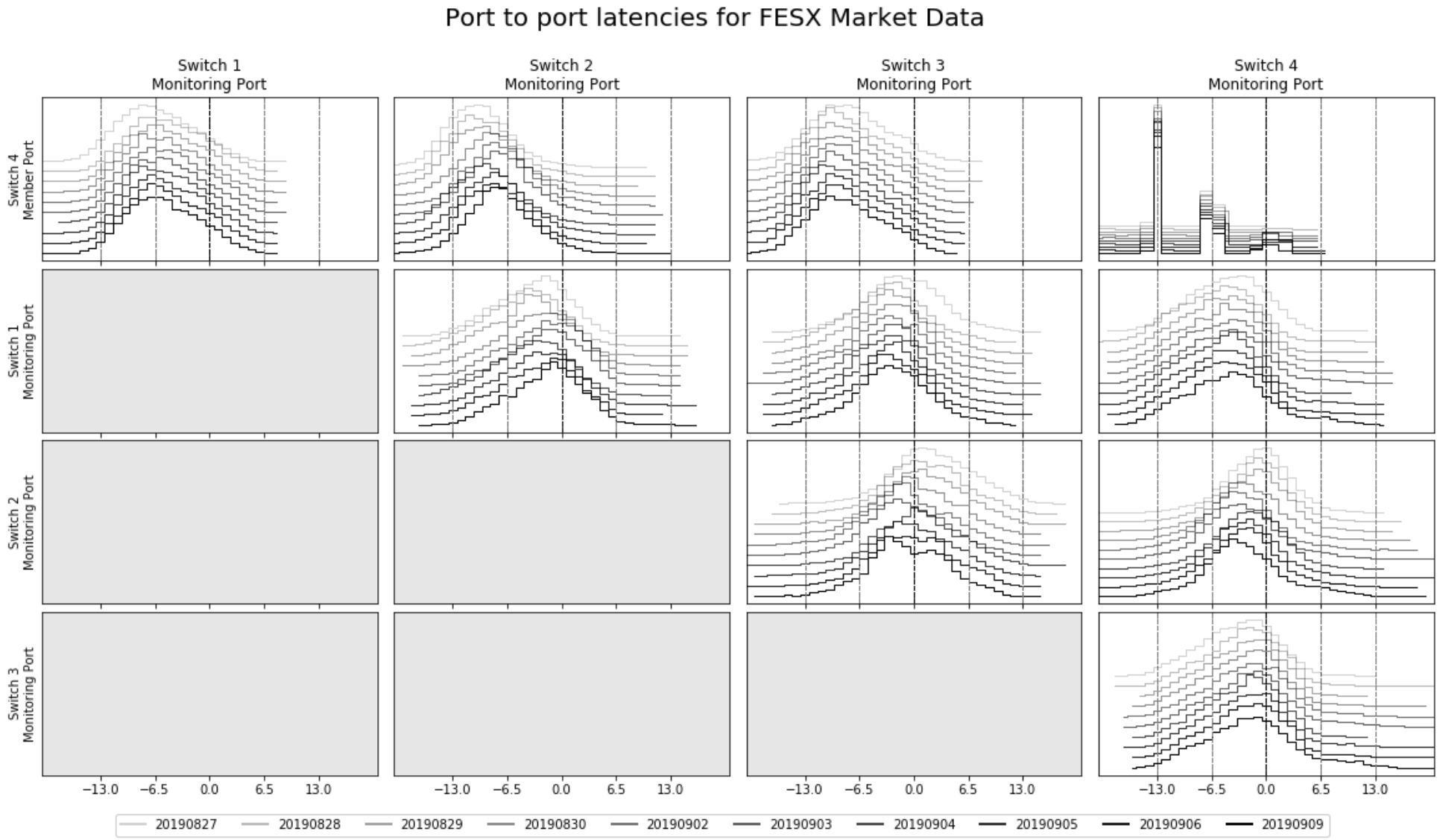
Effects

- 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



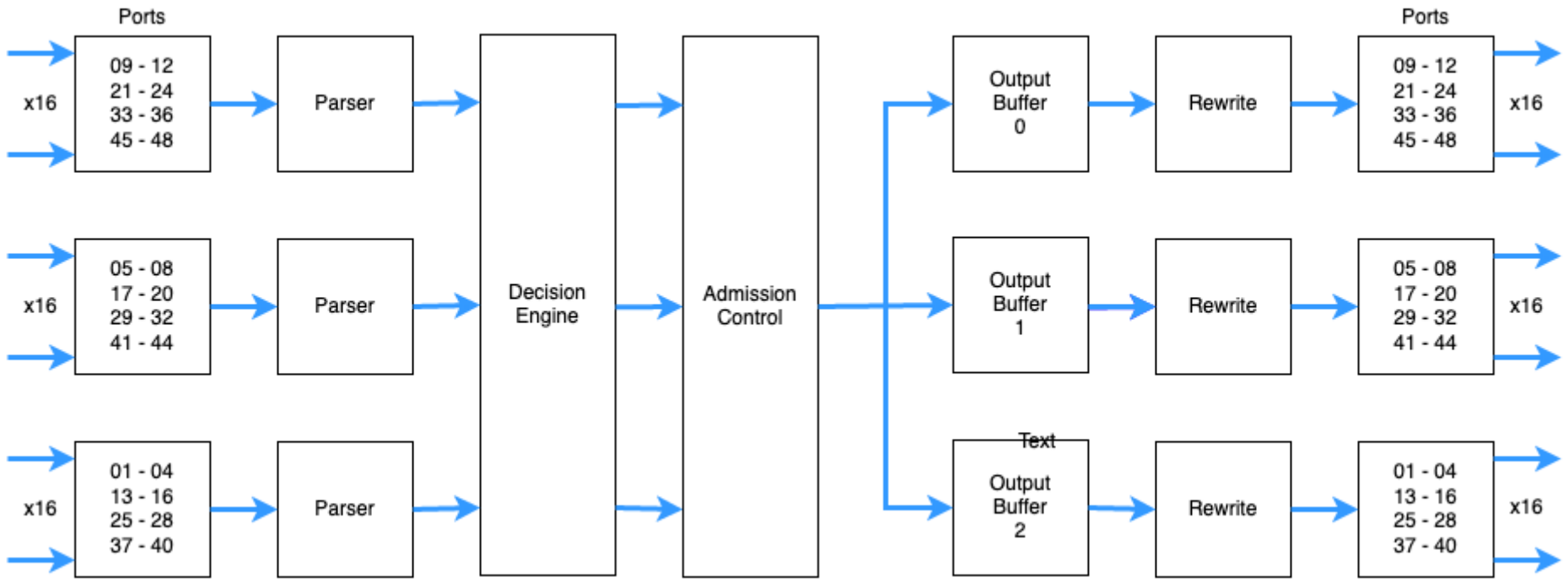
Market Data Distribution

Effects



Market Data Distribution

Monticello ASIC Block Diagram



Thank you for your attention

Contact

Andreas Lohr
Sebastian Neusüß

E-mail monitoring@deutsche-boerse.com
Phone +49 69 2111 8686

46

Appendix

Matching Algorithms
Measures to Reduce Need for Speed

Actual

- Various “Speed Bumps”
  - Magic shoe box = long box of coiled up cable
  - “Latency floor”

- Passive Liquidity Protection on Eurex
  - New orders that could match with resting orders in the limit order book will be deferred
  - Pilot phase launched German (1ms delay) and French (3ms delay) equity options

- Last Look
  - Liquidity provider has additional time to device whether to accept a trade or not
  - Used in foreign exchange markets (currencies)

Matching Algorithms
Measures to Reduce Need for Speed

Proposals

- Frequent Batch Auctions
  - https://faculty.chicagobooth.edu/eric.budish/research/HFT-FrequentBatchAuctions.pdf

© Deutsche Börse 2019
Deutsche Börse AG (DBAG), Clearstream Banking AG (Clearstream), Eurex Frankfurt AG, Eurex Clearing AG (Eurex Clearing) as well as Eurex Bonds GmbH (Eurex Bonds) and Eurex Repo GmbH (Eurex Repo) are corporate entities and are registered under German law. Eurex Zürich AG is a corporate entity and is registered under Swiss law. Clearstream Banking S.A. is a corporate entity and is registered under Luxembourg law. U.S. Exchange Holdings, Inc. is a corporate entity and is registered under U.S. American law. Deutsche Boerse Asia Holding Pte. Ltd., Eurex Clearing Asia Pte. Ltd. and Eurex Exchange Asia Pte. Ltd are corporate entities and are registered under Singapore law. Eurex Frankfurt AG (Eurex) is the administrating and operating institution of Eurex Deutschland. Eurex Deutschland and Eurex Zürich AG are in the following referred to as the "Eurex Exchanges".
All intellectual property, proprietary and other rights and interests in this publication and the subject matter hereof (other than certain trademarks and service marks listed below) are owned by DBAG and its affiliates and subsidiaries including, without limitation, all patent, registered design, copyright, trademark and service mark rights. While reasonable care has been taken in the preparation of this publication to provide details that are accurate and not misleading at the time of publication DBAG, Clearstream, Eurex, Eurex Clearing, Eurex Bonds, Eurex Repo as well as the Eurex Exchanges and their respective servants and agents (a) do not make any representations or warranties regarding the information contained herein, whether express or implied, including without limitation any implied warranty of merchantability or fitness for a particular purpose or any warranty with respect to the accuracy, correctness, quality, completeness or timeliness of such information, and (b) shall not be responsible or liable for any third party's use of any information contained herein under any circumstances, including, without limitation, in connection with actual trading or otherwise or for any errors or omissions contained in this publication.
This publication is published for information purposes only and shall not constitute investment advice respectively does not constitute an offer, solicitation or recommendation to acquire or dispose of any investment or to engage in any other transaction. This publication is not intended for solicitation purposes but only for use as general information. All descriptions, examples and calculations contained in this publication are for illustrative purposes only.
Eurex and Eurex Clearing offer services directly to members of the Eurex exchanges respectively to clearing members of Eurex Clearing. Those who desire to trade any products available on the Eurex market or who desire to offer and sell any such products to others or who desire to possess a clearing license of Eurex Clearing in order to participate in the clearing process provided by Eurex Clearing, should consider legal and regulatory requirements of those jurisdictions relevant to them, as well as the risks associated with such products, before doing so.
Eurex derivatives are currently not available for offer, sale or trading in the United States or by United States persons (other than EURO STOXX 50® Index Futures, EURO STOXX 50® ex Financials Index Futures, EURO STOXX® Select Dividend 30 Index Futures, EURO STOXX® Index Futures, EURO STOXX® Large/Mid/Small Index Futures, STOXX® Europe 50 Index Futures, STOXX® Europe 600 Index Futures, STOXX® Europe 600 Banks/Industrial Goods & Services/Insurance/Media/Travel & Leisure/Utilities Futures, STOXX® Europe Large/Mid/Small 200 Index Futures, Dow Jones Global Titans 50 IndexSM Futures (EUR & USD), DAX®/Mini-DAX®/MDAX®/TecDAX® Futures, SMIM® Futures, SLI Swiss Leader Index® Futures, MSCI World (FMWO, FMWP, FMWN)/Europe (FMEU, FMEP)/ Europe Value/Europe Growth/Emerging Markets (FMEM, FMEF, FMEN)/Emerging Markets Latin America/Emerging Markets EMEA/Emerging Markets Asia/China Free/India/Japan/Malaysia/South Africa/Thailand/AC Asia Pacific ex Japan Index Futures, TA-25 Index Futures, Daily Futures on TAIEX Futures, VSTOXX® Futures, Gold and Silver Futures as well as Eurex FX, property and interest rate derivatives).

Trademarks and Service Marks
Buxl®, DAX®, DivDAX®, eb.rexx®, Eurex®, Eurex Bonds®, Eurex Repo®, Eurex Strategy WizardSM, Euro GC Pooling®, FDAX®, FWB®, GC Pooling®,,GCPI®, MDAX®, ODAX®, SDAX®, TecDAX®, USD GC Pooling®, VDAX®, VDAX-NEW® and Xetra® are registered trademarks of DBAG. All MSCI indexes are service marks and the exclusive property of MSCI Barra. ATX®, ATX® five, CECE® and RDX® are registered trademarks of Vienna Stock Exchange AG. IPD® UK Annual All Property Index is a registered trademark of Investment Property Databank Ltd. IPD and has been licensed for the use by Eurex for derivatives. SLI®, SMI® and SMIM® are registered trademarks of SIX Swiss Exchange AG. The STOXX® indexes, the data included therein and the trademarks used in the index names are the intellectual property of STOXX Limited and/or its licensors Eurex derivatives based on the STOXX® indexes are in no way sponsored, endorsed, sold or promoted by STOXX and its licensors and neither STOXX nor its licensors shall have any liability with respect thereto. Dow Jones is a service mark of Dow Jones & Company, Inc. All derivatives based on these indexes are not sponsored, endorsed, sold or promoted by Dow Jones & Company, Inc. Dow Jones & Company, Inc. does not make any representation regarding the advisability of trading or of investing in such products. Bloomberg Commodity IndexSM and any related sub-indexes are service marks of Bloomberg L.P. All references to London Gold and Silver Fixing prices are used with the permission of The London Gold Market Fixing Limited as well as The London Silver Market Fixing Limited, which for the avoidance of doubt has no involvement with and accepts no responsibility whatsoever for the underlying product to which the Fixing prices may be referenced. PCS® and Property Claim Services® are registered trademarks of ISO Services, Inc. Korea Exchange, KRX, KOSPI and KOSPI 200 are registered trademarks of Korea Exchange Inc. Taiwan Futures Exchange and TAIFEX are registered trademarks of Taiwan Futures Exchange Corporation. Taiwan Stock Exchange, TWSE and TAIEX are the registered trademarks of Taiwan Stock Exchange Corporation. BSE and SENSEX are trademarks/service marks of Bombay Stock Exchange (BSE) and all rights accruing from the same, statutory or otherwise, wholly vest with BSE. Any violation of the above would constitute an offence under the laws of India and international treaties governing the same. The names of other companies and third party products may be trademarks or service marks of their respective owners.