

# ARE WE THEN YET?

## Defining time sync requirements for electronic trading

Intercontinental Exchange



**WOJCIECH OWCZAREK**

WSTS 2015, San Jose, CA, 9-12 March 2015

## Background – ICE



- › Global stock exchange operator
- › ICE + NYSE group of companies
- › Multiple exchanges in US and Europe
- › Own clearing houses in US and Europe
- › Global financial market access network operator
- › State of the art DC: UK and multiple in US
- › Global dark fibre backbone, highly resilient
- › Nodes in all large finance hubs in EU and US
- › Primarily 10GE, 40GE rollout since 2014

## Timing in electronic trading - specifics

### › All about phase

### **No hard limits, poor sync harder to track:**

- › No dropped calls, no power grid collapse
- › Can go unnoticed unless specifically monitored
- › Most react, few detect

### **Exchanges vs. HFT players / hedge funds:**

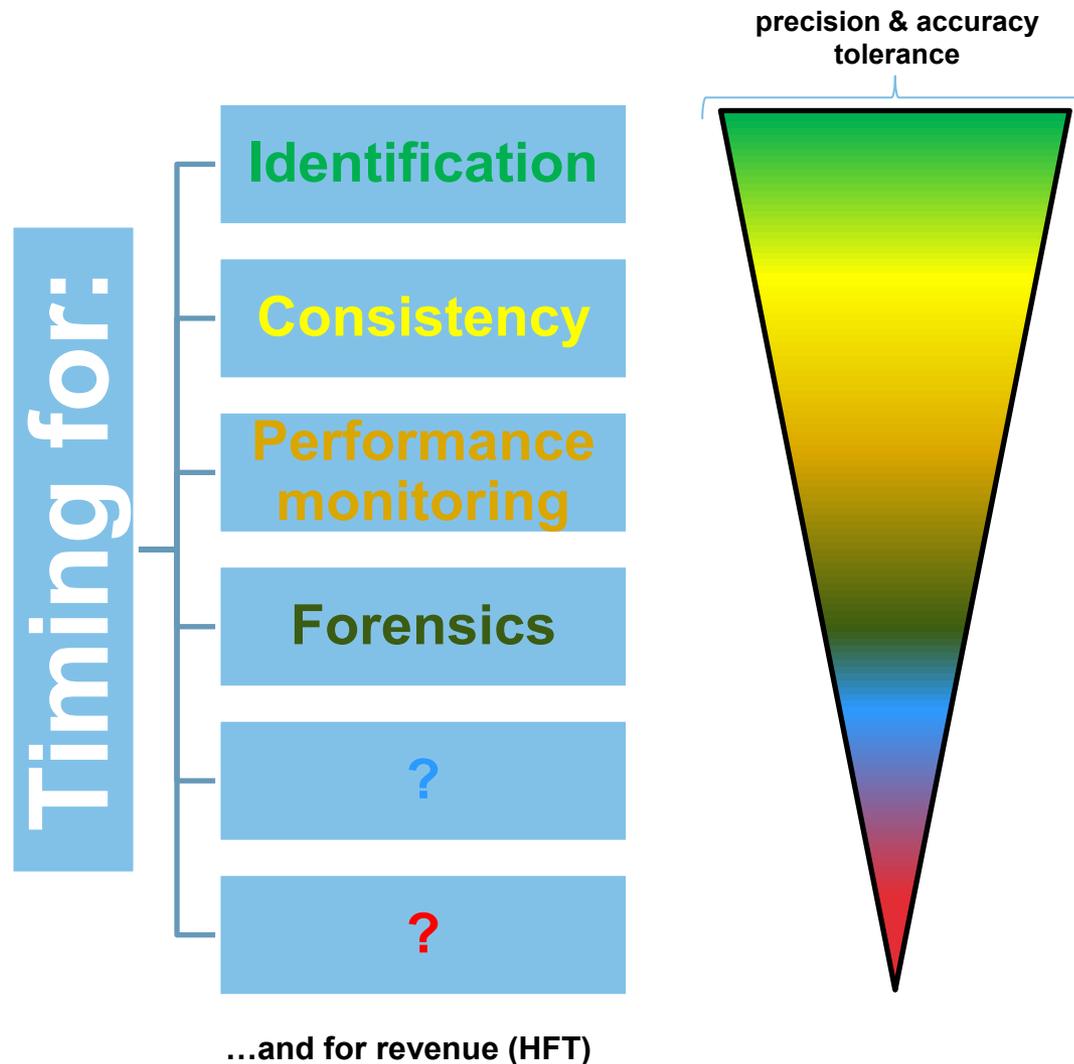
- › *Timing not critical for revenue*
- › Less frequent H/W refresh
- › Message rates can be lower

## Timing in electronic trading - specifics

### **Adding timing to the network:**

- › Requires on-path hardware support
- › Need to use enterprise hardware
- › Separate timing network = wasted ports
- › Hardware just about ready, not every tier yet
- › Network architects must keep timing in mind
- › Can't get away from HW
- › Half measures short-lived

# Different facets of timing in electronic trading



## Electronic trading – timing for... Identification

### Timing for identification - orders:

- › Timestamp each step: enter, route, execute
- › Resulting public market data, timestamped
- › Market data feeds back algorithms

### Minimal requirement - resolution:

- › Busiest market: 200,000+ mps / line (FIF, 2014)
- › Peaks at over 10,000 mps / symbol
- › **Sub-100 $\mu$ s resolution needed *at minimum***
- › Many protocols still at 1ms timestamp fields
- › Many moved or moving to 1 ns resolution

# Electronic trading – timing for... Consistency

Story of a quick trade...

## Negative delays (very bad):

time @ node

- › Bid submitted (gateway or ME) 12:35:12.782000
- › Trade made (Matching Engine) 12:35:12.780000
- › Path delay (including software) 00:00:00.000100

GW vs. ME out of sync or time drift or spike on ME

App Timestamps usually generated in software

- › Order filled 1.9 ms in the future?

## Problems maintaining and measuring order lifetime

Max phase error allowed:  $\leq$  path delay  $_{GW \rightarrow ME}$

Easy? Fun begins here.

# Electronic trading – timing for... Consistency

- › Faster media, CPUs, NICs, ASICs... Path delay keeps shrinking

Host wire to application (OS)	single ToR switch	single Core switch Increased for m/cast and many line cards
2010: 20 $\mu$ s (OS only)	2003: 15 $\mu$ s+ GE ToRs S&F	2005: 12 $\mu$ s+ GE core switch chassis
2010: sub-10 $\mu$ s (offload engine)	2005-2008: 4-6 $\mu$ s+ Enterprise GE ToRs S&F	2008: 10 $\mu$ s+ 10GE core switch chassis
2012: 2-3 $\mu$ s (offload engine)	2008: 3-5 $\mu$ s first 10GE ToRs S&F	2010: 4 $\mu$ s 10GE core switch chassis
2011...: sub-1 $\mu$ s (FPGA off switching ASIC)	2010: sub-1 $\mu$ s first ULL 10GE ToR C/T	2012: 4 $\mu$ s Clos topology built with 10G TORs
2014: sub-2 $\mu$ s (offload engine)	2013: sub-500 ns next-gen 10GE ASICs C/T	2013...: sub-1 $\mu$ s 10/40 GE ToRs capable of core role
(RoCE or RDMAoIB usually internal)	2013: sub-1 $\mu$ s 40GE ASICs C/T	

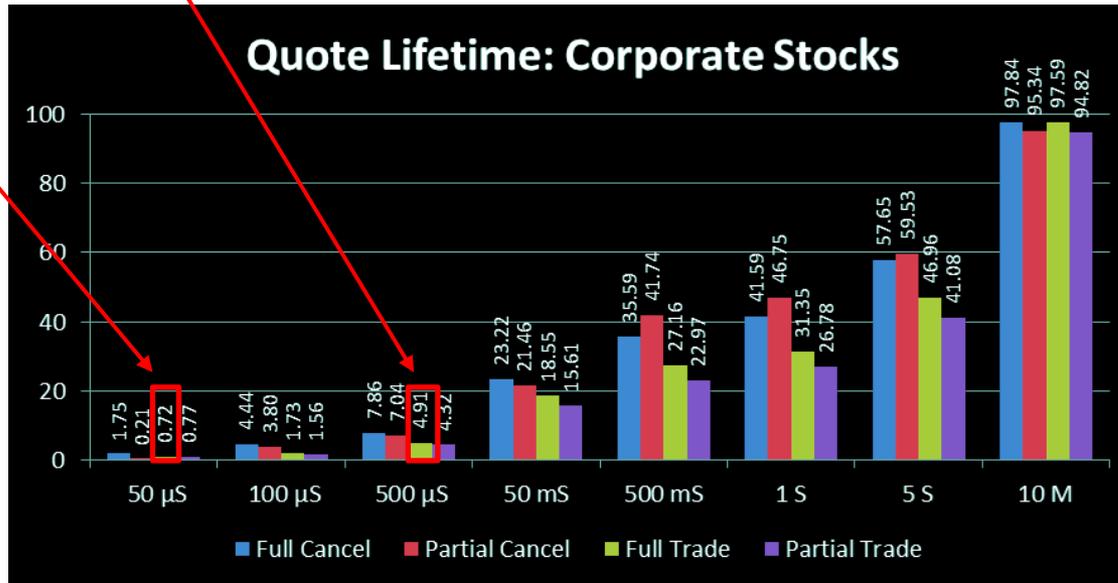
**Evolution of data delivery latencies (LAN/DC) – application has caught up with the network**

# Electronic trading – timing for... Consistency

Over 0.5% trades executed in sub-50  $\mu$ s

Close to 5% trades executed in sub-500  $\mu$ s

Not all orders filled immediately



(US Securities and Exchange Commission, 2014)

**Max phase error allowed:  $\min(T_{\text{fill}})$**   
**(sync across whole exchange)**

# Electronic trading – timing for... Monitoring

## Technology:

- › Passive taps, timestamp at TapAgg layer
- › Port mirror + ingress timestamp on ToR
- › ToR Switches with built-in analytics
- › *OR* Applications using s/w timestamps for monitoring



## Latency monitoring probes:

- › Data from taps at different locations
- › Or timestamped data from TapAgg
- › Data hashed, correlated
- › End-to-end latencies measured as low as 10 $\mu$ s

**For 10% uncertainty margin:  
1 $\mu$ s precision or better**

## Time sync needed across monitoring kit only

## Electronic trading – timing for... Monitoring

### Internal latency monitoring (within trading system):

- › Self-computed figures not reliable unless...
- › At sub-1  $\mu\text{s}$  error no extra tools needed
- › At sub-10  $\mu\text{s}$  error extra tools mostly not needed

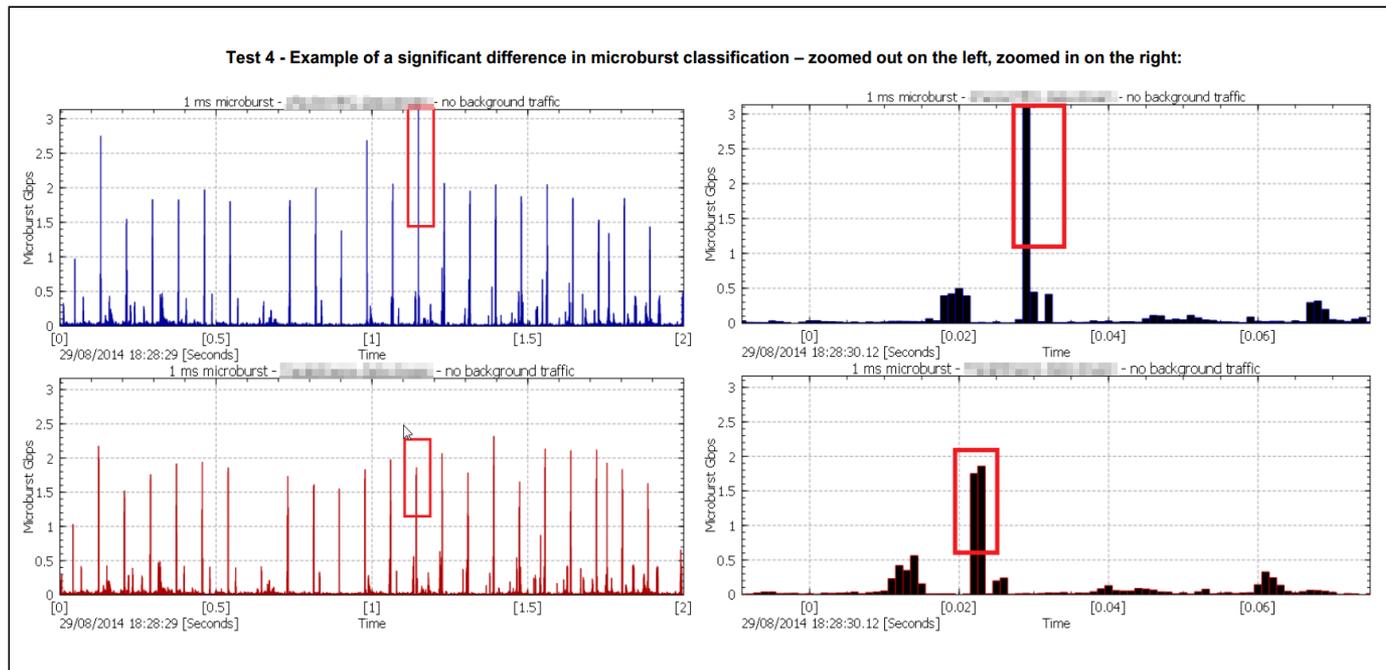
### Capacity monitoring:

- › Ultra low-latency switches (ULL), small buffers
- › Burst can cause packet loss, needs monitored
- › Capacity increased when hitting bursts often
- › Line rate microburst measurement
- › Derived from data rate seen in 1 ms

# Electronic trading – timing for... Monitoring

## Capacity monitoring:

- › Tools out of sync: false positives or negatives:



**Example - two tools reporting almost 50% difference:  
Caused by 160  $\mu$ s phase error (ICE/NYSE lab, 2014)**

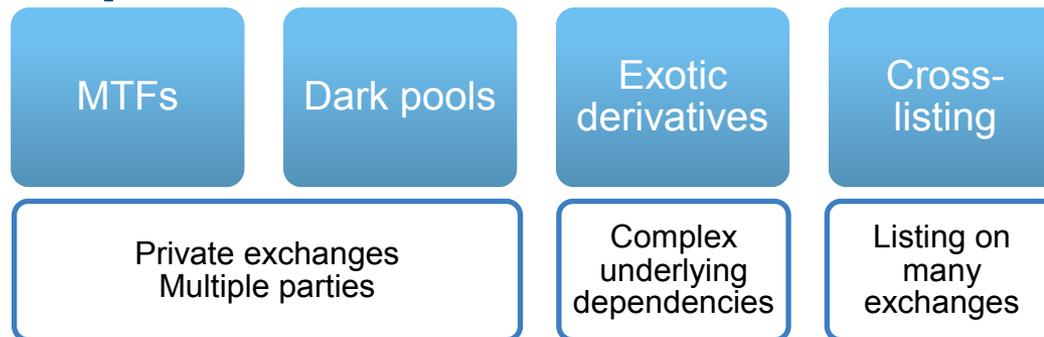
**For 10% margin,  
max error appr. 30  $\mu$ s**

# Electronic trading – timing for... Market Forensics

**\*Exchange-driven requirements end here\***

- › Tracking down a cascade of events
- › Should be easy on one exchange

**Not so simple in real life:**



**“Algos gone wild”:**

**Algorithm places up to 30,000 quotes / second on the same stock (Feb 2014, NYSE – many other examples)**

# Electronic trading – timing for... Market Forensics

## Previous timing use cases:

- › Exchange time taken as is
- › Recording the order of trades good enough

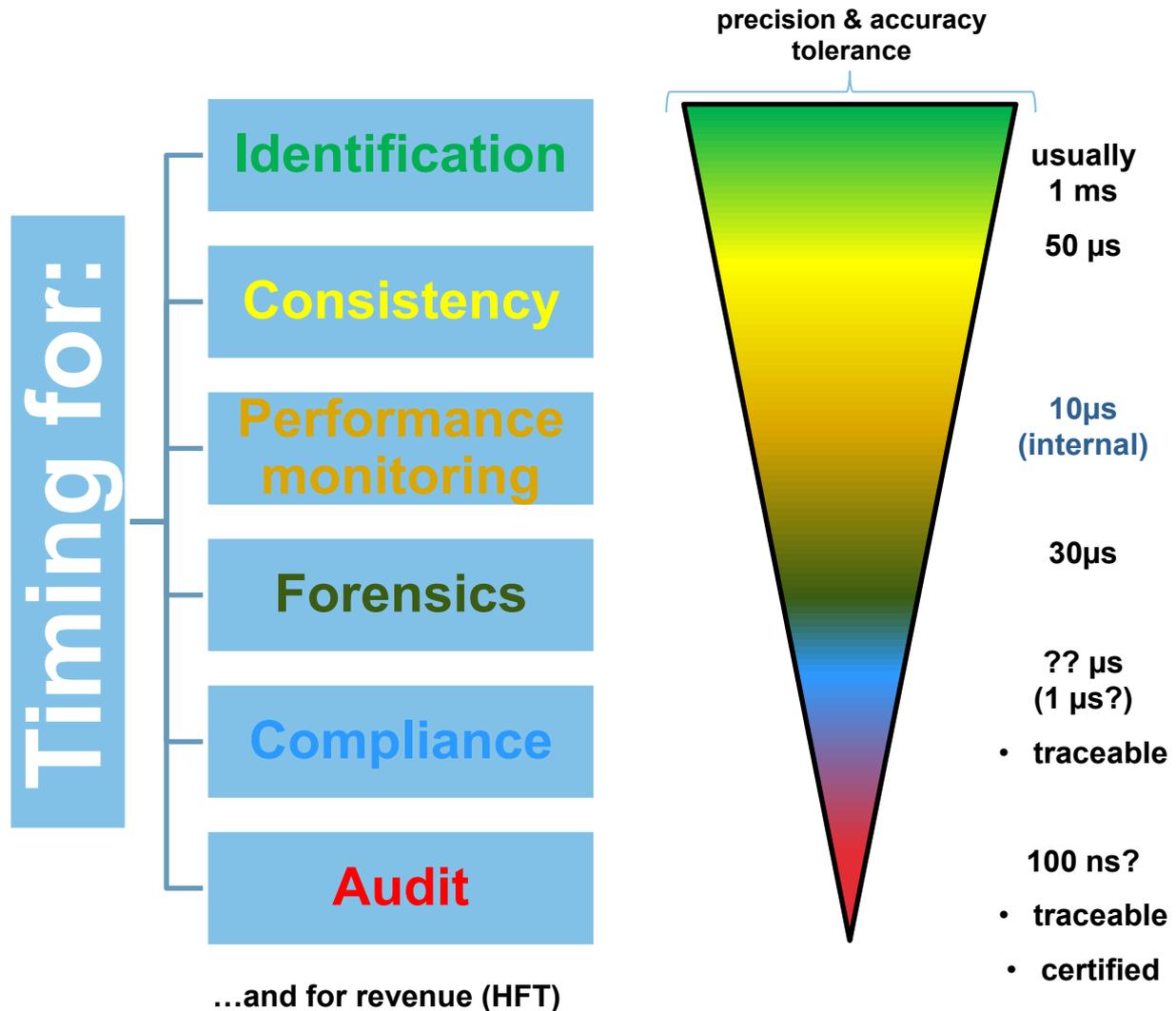
## Forensics in HFT land:

- › Multiple order entry points
- › Third parties behind members
- › Private mini-exchanges

## Required:

- › Global UTC sync of all parties
- › At least sub-30  $\mu$ s error and decreasing

# Different facets of timing in electronic trading



## Electronic trading – timing for... Compliance

### Current regulations on timestamps:

- › Can be met even with NTP over WAN
- › **FINRA: 1 s resolution. 1 s error vs. NIST:**
  - 1 ms resolution only since Sep 2014
    - off-exchange trades only
  - some pre-2000s rules such as: sync once per day
- › **SEC: 3 s from NIST or USNO until Feb 2014**
- › **SEC rule 613 (CAT) - Feb 2014:**
  - 1 ms resolution
  - max error 50 ms from NIST
- › **Europe - MiFID1: no mention of time sync**

# Electronic trading – timing for... Compliance

## ENTER MiFID 2:

### › May-Dec 2014

- Discussion Paper
- Call for comments

### › Dec 2014 - tech advice released:

- **1 ms resolution and accuracy**  
when venue's gateway-to-gateway time 1 ms+
- **1  $\mu$ s resolution and accuracy**  
when gateway-to-gateway time sub-1 ms (*many*)

### › **Effective Jan 2017 if approved mid-2015**



## Electronic trading – timing for... Compliance

### Timing questions in MiFID 2 DP:

- › National or pan-European source of sync?
- › Is 1  $\mu$ s resolution / precision acceptable?
- › Maximum error? How often to correct?

### Industry's responses:

- › 34 out of 200+ responded – not all who should have
- › Some major exchanges and HFT did not.

## Electronic trading – timing for... Compliance

### MiFID 2 Discussion Paper - Industry's responses:

- › Nearly all agree on UTC
- › Many myths, many naïve answers (educate!)
- › Nearly all concerned with substantial costs (if 1 $\mu$ s mandated) – is it really that bad?
- › Huge disparity - anywhere from 1 $\mu$ s to seconds
- › Few agree on 1 $\mu$ s resolution / error
- › Most suggest 1 ms. Some: correct at start of day!
- › Very few suggest well-motivated 10  $\mu$ s
- › **Very few raise the need for timing audit**

# Electronic trading – timing for... Compliance + Audit

## MiFID 2 aftermath:

- › Even if not 1 $\mu$ s mandated, likely lower than 1 ms
- › Sooner or later SEC / FINRA will follow
- › **Like it or not – everybody gets precision timing**
- › Only systems in critical order path
- › **Traceability required**
- › Reporting and logging needed – time sync audit
- › **Time sync audit also requires *certified time***
- › Need to develop and deploy time monitor tools



Vendors :  
Sales opportunity

# Time sync for Electronic trading: getting there

## Yesterday – basic precision timing:

- › **GM, soft slaves or HW slaves**

## Today – partial on-path – separate timing network:

- › **GM → BC/TC (distribution) → ToR (BC) → HW Slv**
- › **GM → BC (distribution) → ToR (TC) → HW Slv**

## Tomorrow – full on-path support:

- › **DC core and WAN core gain TC and BC support**
- › **Not really enterprise-ready today**

## Day after tomorrow:

- › **Ubiquitous sync must happen, TAACCS effort**
- › **tod register and rdtod in CPU!**

# Time sync for Electronic trading: getting there

## Road to full sync:

- › **Most new NICs today, all tomorrow**
- › **Most network kit today, all tomorrow**
  - **IEEE 1588 as standard feature from many vendors**
- › **Realistically 10 $\mu$ s will do today (for trading systems)**
  - **Performance and monitoring need better**
- › **Organic transition if legislation is paced**
- › **Eventually in-hardware sync out of the box**

# Time sync for Electronic trading: getting there

## What we need from vendors:

- › **1PPS out and TIE measurement everywhere**
- › **Multi-input, arbitrating GMs:**
  - **GNSS, eLORAN, PTP in**
- › **10GE GMs! In fact – at least more GE GMs!**
- › **Core switches and core routers:**
  - **enterprise PTP, not Telecom or SyncE (unless BC can translate!)**
- › **Sync monitoring equipment (not lab kit)**
- › **Certified time services**

# ICE / NYSE LAB

Belfast : Sep 2010 – Jun 2014

Install



Build & Operate



Dismantle



Shut down



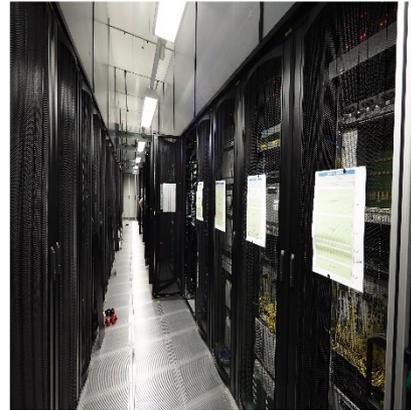
# ICE / NYSE LAB

Basildon DC: Jun 2014 onwards

## Secure & Move



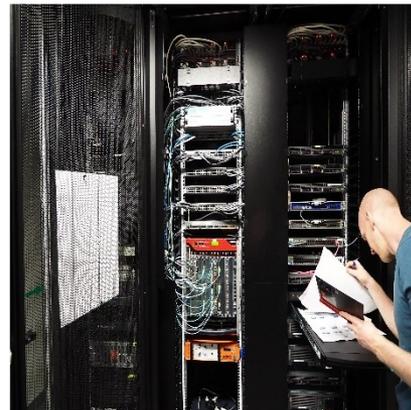
## Install



## Build



## Operate



**Time sync for Electronic trading: Thank you.**

**Questions?**

**wojciech.owczarek@theice.com**