

Talk given at Bank of England, April 12 2016

Financial services regulation of technology, theory and practice.

MiFID II Clock regulation approaches



April 2016

What happens as firms address a major regulatory change

Recent market events sparked calls for new regulations



MiFID II and MiFIR according to ESMA

Regulation on Markets in Financial Instruments, commonly referred to as MiFID II and MiFIR, were adopted by the European Parliament on 15 April 2014, by the Council of the European Union on 13 May 2014 and published in the EU Official Journal on 12 June 2014.

Building on the rules already in place, these new rules are designed to take into account developments in the trading environment since the implementation of MiFID in 2007 and, in light of the financial crisis, to improve the functioning of financial markets making them more efficient, resilient and transparent.

Clocks are a small, but key, part of MiFID II

Operators of trading venues and their members or participants shall establish a system of traceability of their business clocks to UTC. This includes ensuring that their systems operate within the granularity and a maximum tolerated divergence from UTC as per RTS 25. Operators of trading venues and their members or participants shall be able to evidence that their systems meet the requirements. They shall be able to do so by documenting the system design, it's functioning and specifications.

[Translation to follow]



RTS 25 requirement is simplicity itself

**High frequency operation:
100 microsecond accuracy on
timestamps
1 microsecond granularity.**

**General automated trading:
1 millisecond accuracy, 1
millisecond granularity.**

RTS 25 numbers are a big step back from original proposal

- ESMA acknowledges however, that at present it may not currently be feasible to expect trading venues to synchronise their clocks or time stamp events to a granularity which is less than nanoseconds. As a result, ESMA has proposed capping the granularity and accuracy requirements at the nanosecond level.

Light travels one foot in a nanosecond!

The units less than nanoseconds are called PICO SECONDS.

Pushback forced ESMA to go from “less than nanoseconds” to “100 microseconds”

100,000 times weaker

two questions

Why does ESMA
care about clock
synchronization
in the first place?

Why is it hard to
synchronize
clocks?

Since I do work in the United States

- **DISCLAIMER:** there is no legal or regulatory advice in this document.
- This is an explanation of one way to interpret regulations from technical and business perspective.



Why does ESMA care about business clocks so much?

- Electronic trading is inherently high speed
- Not specific to high frequency trading
- Clocks that are wrong damage “data integrity” of trading records.



Trading records

Timestamp

12:02.314007 Rx Order A Buy 100 AAL ...

12:02.314102 Tx Order A to LSE

12:02.314271 Rx Confirm A ...

LSE can complete
transactions in
microseconds

A single trading firm can have thousands of computer servers



A single multi-step transaction can produce records from many different computer servers

If the clock in server 1 is 100 microseconds fast & server 30 is 100 microseconds slow the records show confirmation before order

- 12:02.314271 Rx A 192.168.1.1
- 12:02.314272 Rx B 192.168.1.72
- 12:02.314474 Tx Order A 192.168.1.1
- 12:02.314374 Rx C 192.168.1.21
- 12:02.314420 Rx A Conf 192.168.1.30

And then each trading firm may have many data centers



And complex trading strategies involving multiple venues

Correct clocks are essential for regulation

to improve the functioning of financial markets making them more efficient, resilient and transparent.

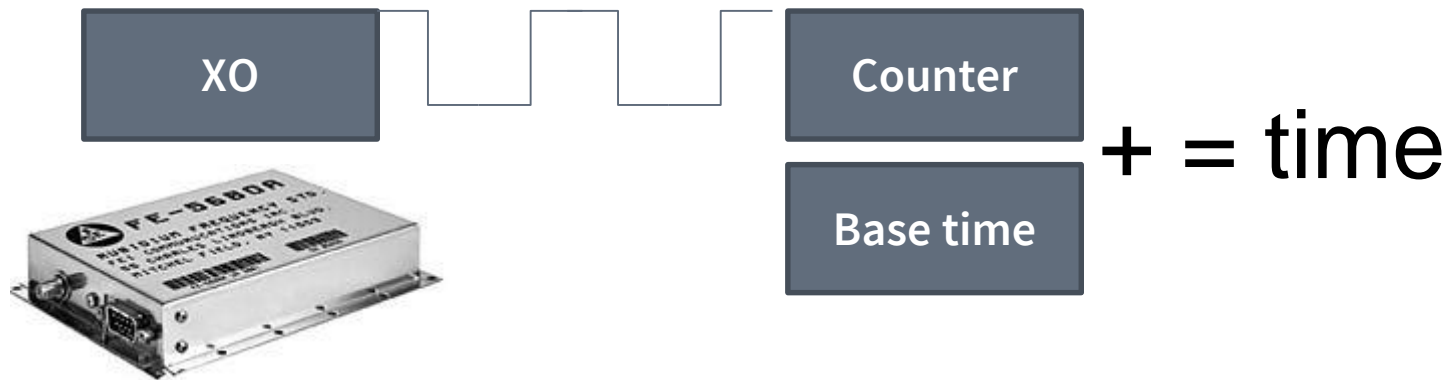
If the hundreds of clocks used to create records are not synchronized – the records don't make sense.

Baffling



Why is this so hard?

- Trading software asks the operating system (Linux, Windows, UNIX ...) for the time.
- The operating system maintains a clock.
- The clock depends on an oscillator



The counter works something like a “stopwatch”.
Elapsed time plus base time = current time.

First question is what should we set the base time to?

- Ideally, all market participants have clocks that are “synchronized” to the same start time.
- This is why MiFID II uses “UTC” as the universal time standard.

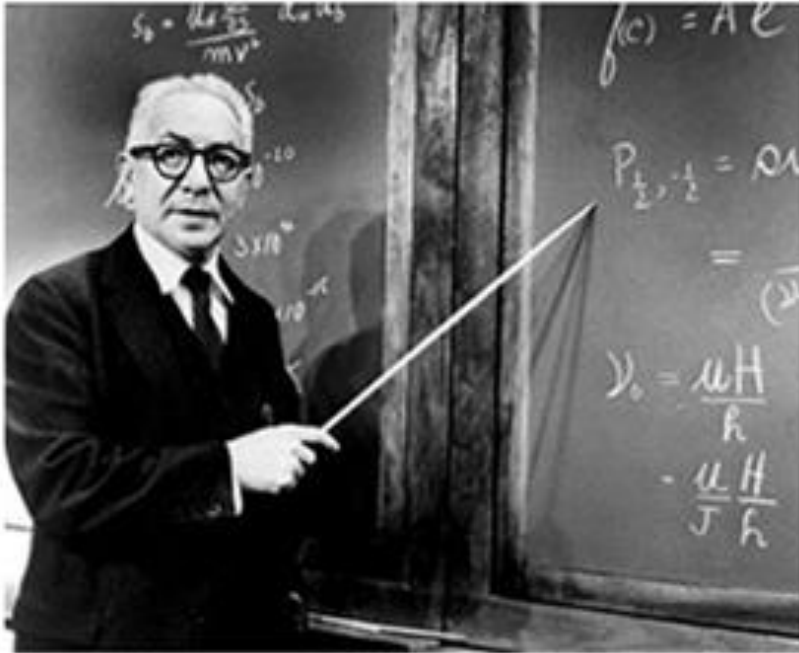
UTC time is the foundation

Operators of trading venues and their members or participants shall establish

a system of traceability of their business clocks to UTC.

*This includes ensuring that their systems operate within the granularity and a **maximum tolerated divergence from UTC** as per RTS 25. Operators of trading venues and their members or participants shall be able to evidence that their systems meet the requirements. They shall be able to do so by documenting the system design, it's functioning and specifications.*

UTC begins with atomic clocks that provide a precise reference second



The physics of cesium clocks is left as an exercise for the audience



UTC is the official time

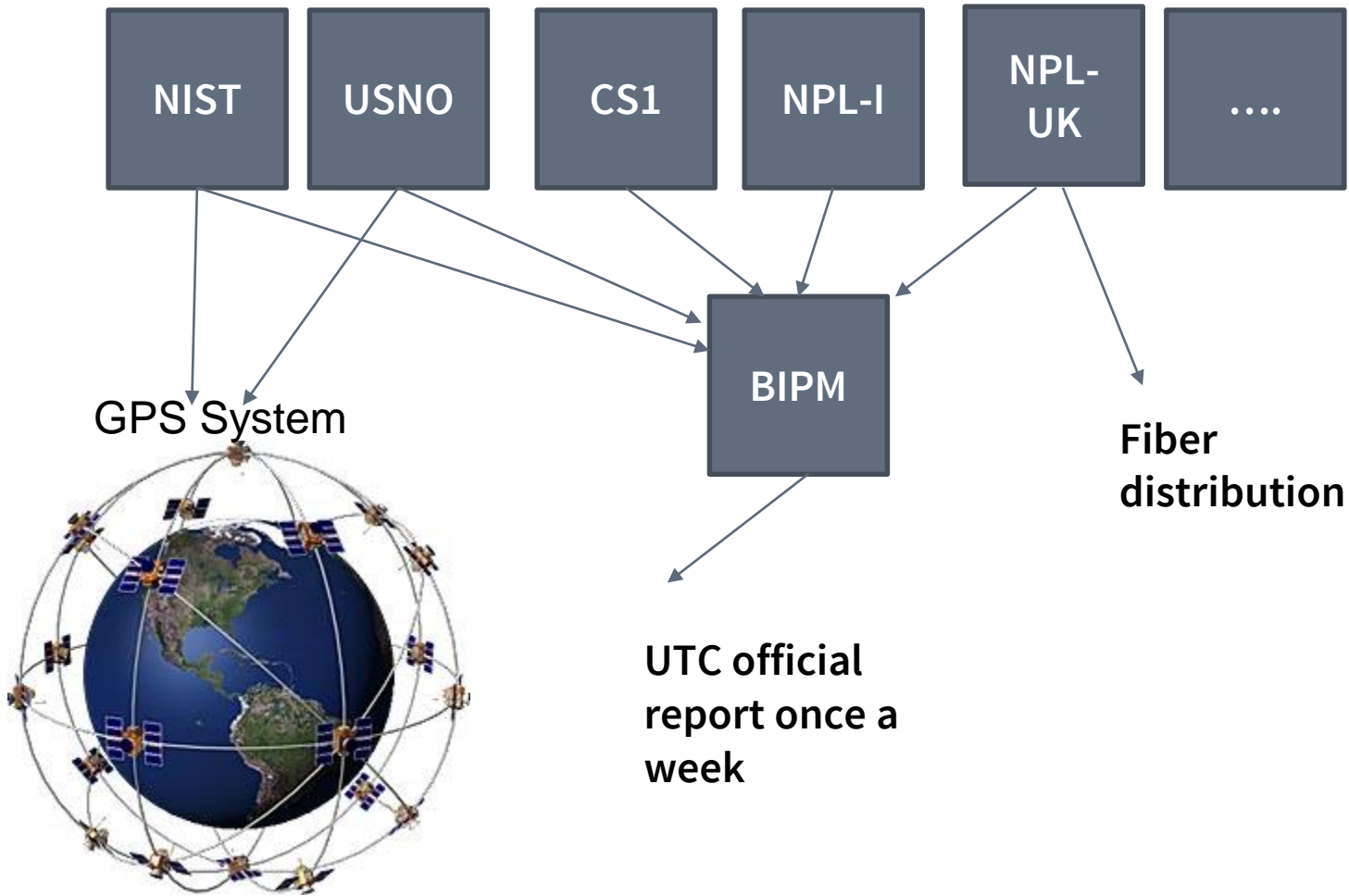
- It is only known in retrospect: as the atomic clocks of national physics labs are combined under a weighted average by the International Bureau of Weights and Measures.



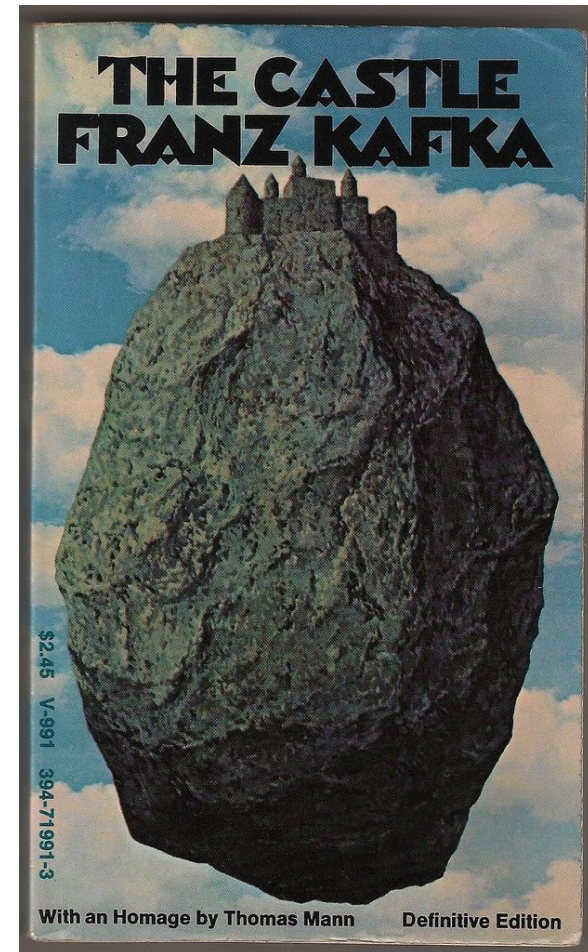
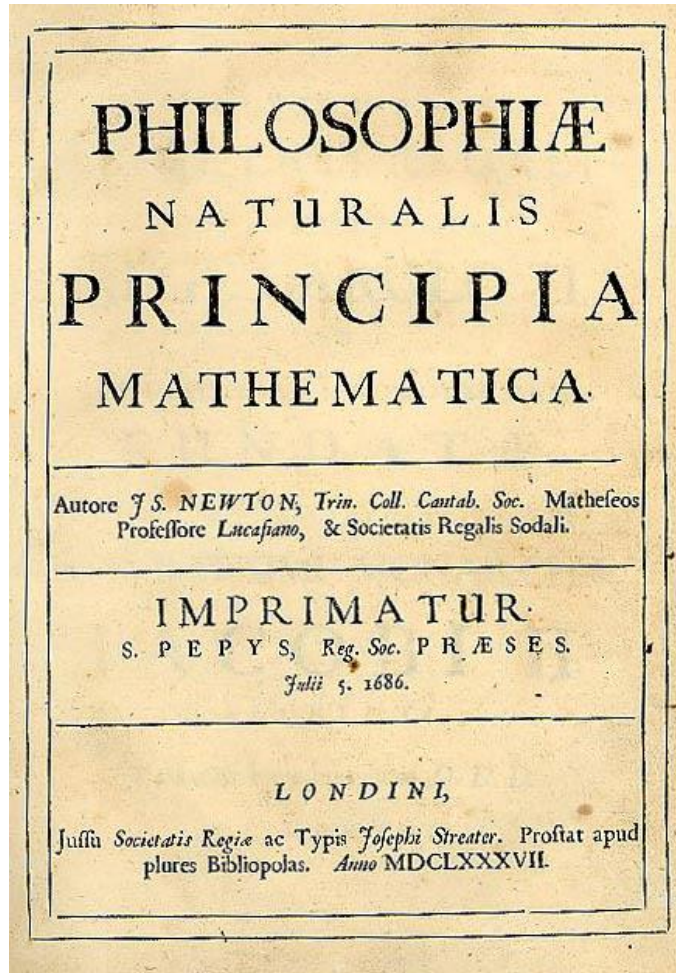
The Bureau of Weights and Measures itself.

The acronym is a compromise to get word order wrong in both English and French.

Traceable



UTC is a little strange



Computer networks deliver time corrections from reference clock devices to computers running applications.



Reference clock
Gets UTC from
GPS or some
terrestrial feed

Application
computers
synchronize their
clocks using
network time

Synchronizes local
oscillator and then
distributes time over
network via some
protocol: PTP or NTP



The back of a GPS network clock



10 Gigabit network ports

GPS antenna on the roof or fiber feed from terrestrial source



How GPS works

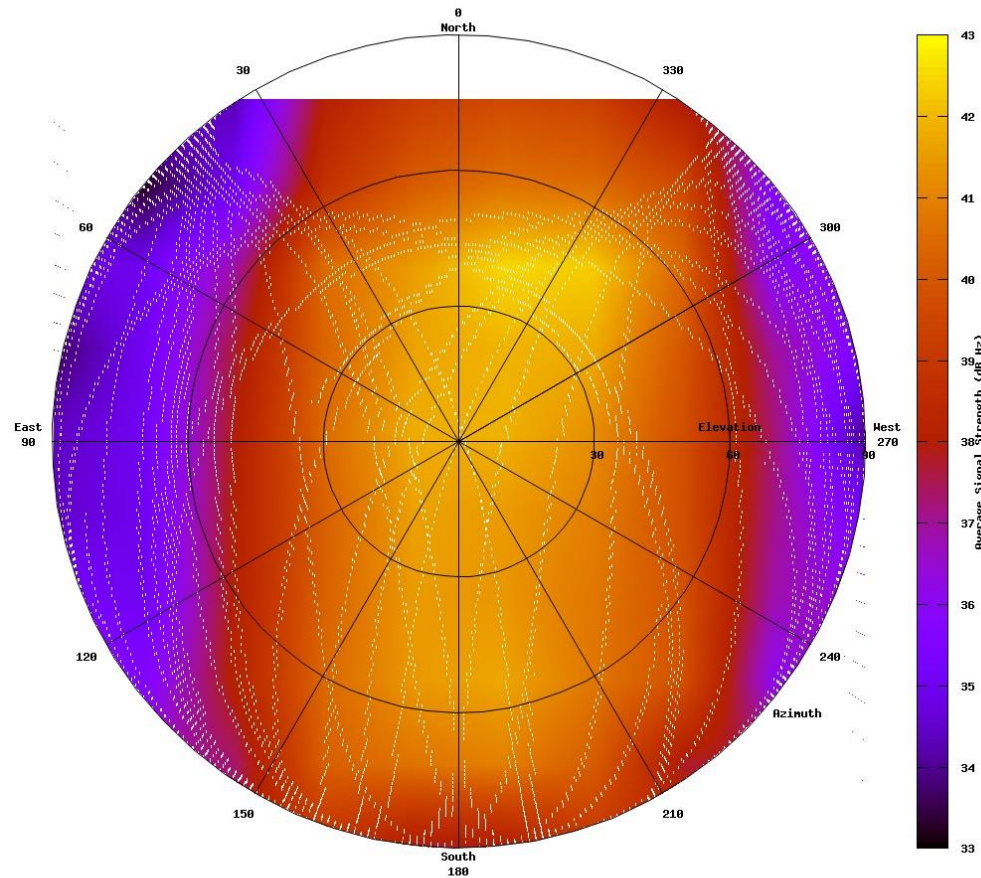
For GPS the antenna comes down to a device combining a radio, an oscillator, and a microcontroller

The device is embedded in a computer that connects to the network

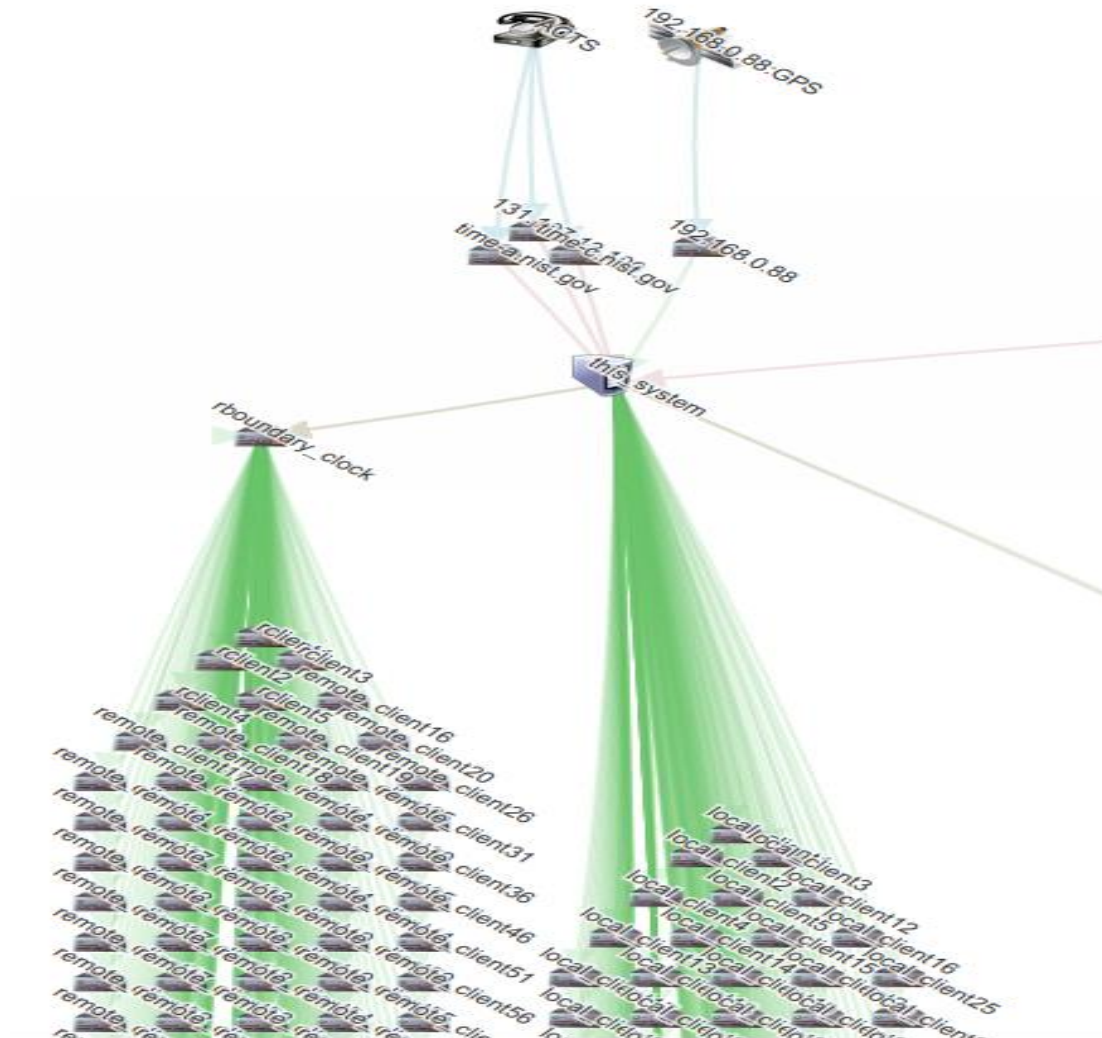


Diagnostic of GPS receiver quality

Barriers to east and west – New York City buildings in urban canyon environment.



Distributing time in a data center



Why is this so hard?

Microseconds are tiny increments of time

Those oscillators – not so great (even really expensive ones drift)

Clocks in computer servers can easily drift thousands of microseconds in an hour

So the clocks need to be adjusted every few seconds.

Corrections come over networks BUT computer networks are slow.

Computer networks take microseconds or milliseconds to deliver packets



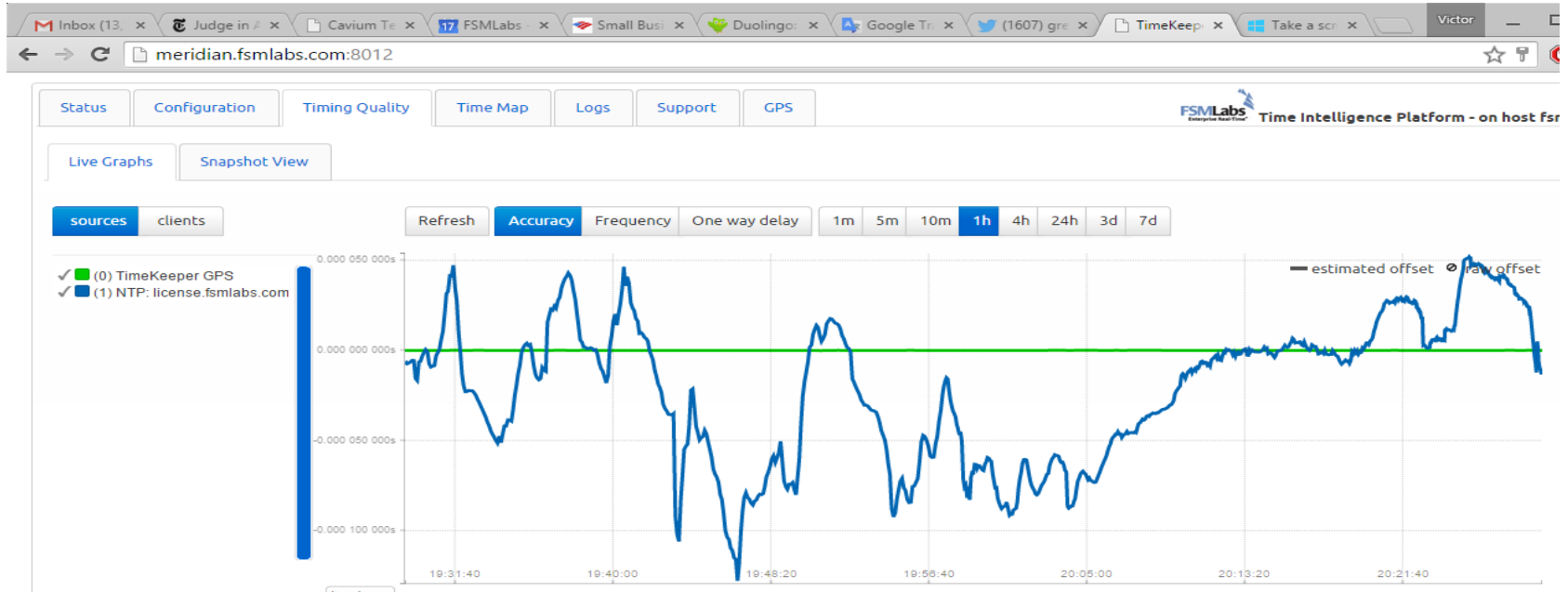
If the correct time is sent from a “reference clock” over the network to a server, it arrives late – by microseconds or even milliseconds.



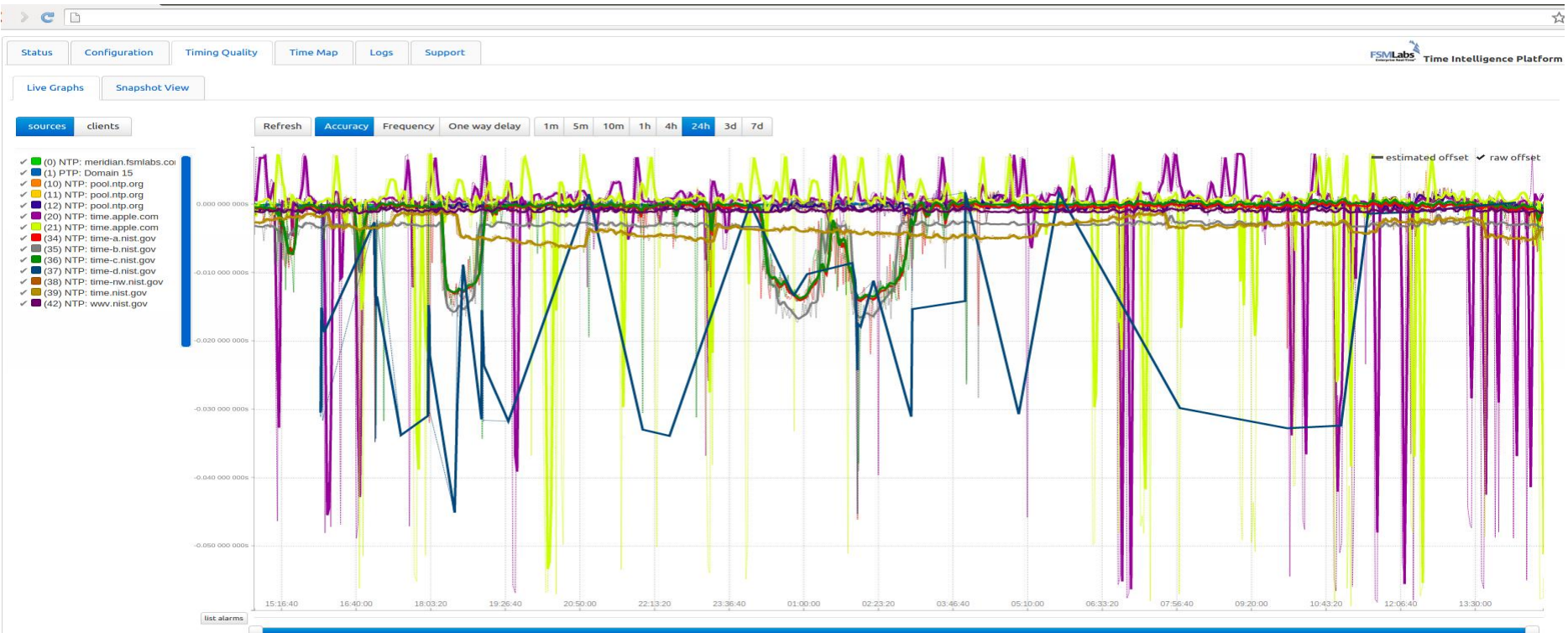
<http://www.computerhope.com>

Network performance changes due to changing load and increasingly complex routers and switches .

Even in a simple network: packet travel time wobbles.



Even worse in the cloud



Fortunately computers can run smart algorithms

- To detect and compensate for problems with networks, reference clocks or local oscillators
- Using sophisticated stochastic models (stochastic models never fail ... or well ...)

Unlike housing prices, time always increases which makes it easier to have robust models.

(although there are clock sync systems that go backwards sometimes!)

Key word

*Operators of trading venues and their members or participants shall establish a system of traceability of their business clocks to UTC. **This includes ensuring that their systems operate within the granularity and a maximum tolerated divergence from UTC as per RTS 25.** Operators of trading venues and their members or participants shall be able to evidence that their systems meet the requirements. They shall be able to do so by documenting the system design, its functioning and specifications.*

Does “ensure” imply fault-tolerance?

- **Fault tolerance is a hard problem**
 - Monitor multiple reference time sources
 - know when a reference is wrong!
 - Have to be able to switch rapidly
- **Fortunately two different bureaucracies, at least, work hard to make the problem more difficult.**



UTC is adjusted to compensate between astronomical and atomic time

- Astronomical time scale is “organic” and simply makes seconds longer or shorter to compensate for earth rotation: always 86400 seconds per day.
- UTC has an unchanging second and adds or subtracts “leap seconds” to avoid diverging
 - TAI (atomic seconds since epoch started Jan 1 1977)
 - $UTC = TAI + LeapSeconds$ (currently -36)
 - Last minute of some months is 61 seconds
 - Seemed like a good idea in 1960s
- World Radiocommunication Assembly can’t decide about abolishing leap seconds

Sometimes what seems like a good idea is not.



The March Hare took the watch and looked at it gloomily: then he dipped it into his cup of tea, and looked at it again: but he could think of nothing better to say than his first remark, 'It was the *best* butter, you know.'

-

There are two similar standards for carrying reference time information in a computer network

1. Network Time Protocol (NTP)
2. IEEE 1588 Precision Time Protocol (PTP)

There are the usual arguments about which is better

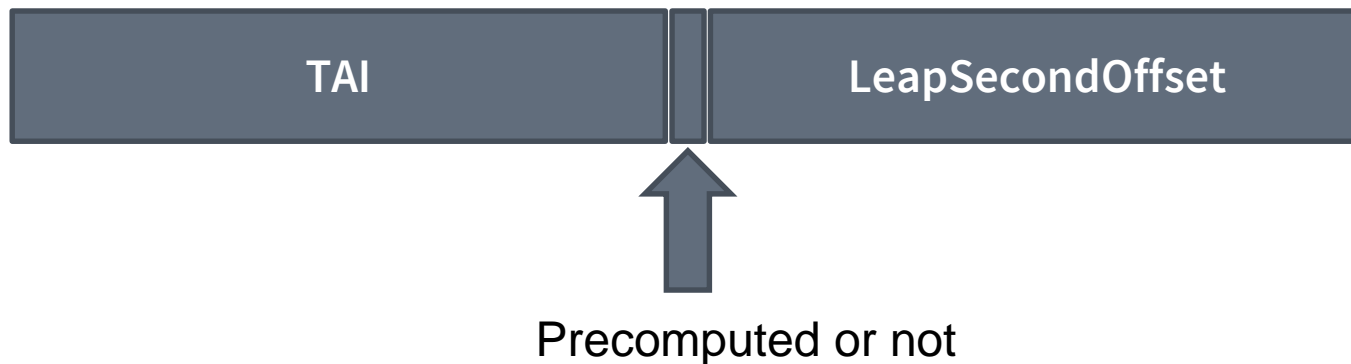


It is computed that eleven thousand persons have at several times suffered death, rather than submit to break their eggs at the smaller end. Many hundred large volumes have been published upon this controversy: but the books of the Big-endians have been long forbidden, and the whole party rendered incapable by law of holding employments.

- From Gullivers Travels (via Danny Cohen 1980 www.ietf.org/rfc/ien/ien137.txt)

IEEE 1588 Precision Time Protocol

- Permits time to be represented either
 - Pre-computed UTC
 - As two numbers: a TAI count and LeapSecond
 - There is a single binary digit in the packet that tells you which method is being used.



Sadly PTP is a politically retrograde standard

- NTP has “time servers” and “time clients”
- PTP has “masters” and “slaves”
- “Slaves” are not permitted to disagree – even if that single bit gets set wrong!



It has happened

That the time was sent in (TAI,LeapSecond) form

But the bit was wrong

And the clocks on the slaves jumped 35 seconds

35 seconds is 35,000,000 microseconds – only 350,000 times more than regulation.

That was just one example of the complexity of the task for clock synchronization tech.

- Network delays
- Network errors
- Asymmetric networks
- Deliberate attempts to compromise
- Heating and cooling of computers that speed up and slow down oscillators
- Failover from reference sources
- Network protocol failures
- GPS antenna problems
- Wrestle with imperfect standards



And make it all intelligible to very busy IT staff who are not experts in or even fascinated by clock sync

And then: Is “evidence” really a verb?

*Operators of trading venues and their members or participants shall establish a system of traceability of their business clocks to UTC. This includes ensuring that their systems operate within the granularity and a maximum tolerated divergence from UTC as per RTS 25. **Operators of trading venues and their members or participants shall be able to evidence that their systems meet the requirements. They shall be able to do so by documenting the system design, it's functioning and specifications.***

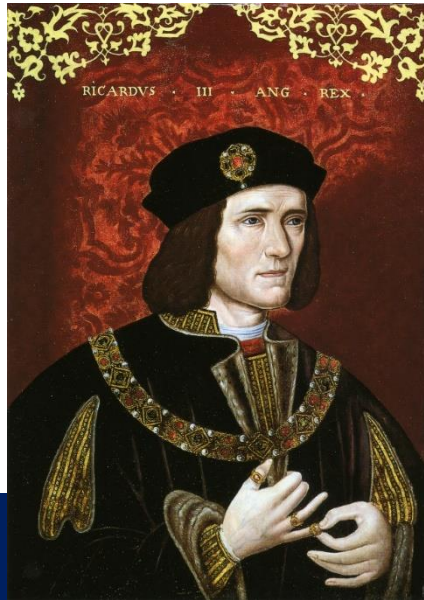
Firms are scrambling to define what “evidence” involves

- We think: “*logs show that during that moment our clock was within accuracy limits for 2 (3? 5?) sources*” is part of evidencing.
- How many years of records are needed?
- How much of network configuration needs to be documented ?
- Why didn't they just say “documented” instead of “evidenced”?
- ... ?

There is a giant record-keeping requirement so we can “evidence”



So it looks simple, but it's a big jump



Richard III:

*Why, then, I do but dream on sovereignty;
Like one that stands upon a promontory,
And spies a far-off shore where he would tread,
Wishing his foot were equal with his eye,
And chides the sea that sunders him from thence,
Saying, he'll lade it dry to have his way*

Summary

- These regulations are necessary and overdue: you can't have fair markets without reliable records.
- The regulations look simple
- The technology is hard – we've just looked at one part.
- All that horrible bureaucracy, tedious RFCs and drafts etc. , eventually approaches a workable balance (we hope).

Contact info

Victor Yodaiken
FSMLabs, Inc.
11701 Bee Caves Road, Suite 200
Austin, TX 78738
USA
yodaiken@fsmlabs.com

Telephone: 1-512-263-5530

Appendix

- Quote on slide 2 from <https://www.esma.europa.eu/policy-rules/mifid-ii-and-mifir>
- Slide 3 https://www.esma.europa.eu/sites/default/files/library/2015-1909_guidelines_on_transaction_reporting_reference_data_order_recording_keeping_and_clock_synchronisation.pdf

Bonus slide

Multiple Internet NTP sources pre-Leap Second 2015 as Google time servers “slew” off correct time

