

NIST Technical Note 2189

**An Evaluation of Dependencies of
Critical Infrastructure
Timing Systems on the
Global Positioning System (GPS)**

Michael A. Lombardi

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2189>

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

NIST Technical Note 2189

**An Evaluation of Dependencies of
Critical Infrastructure
Timing Systems on the
Global Positioning System (GPS)**

Michael A. Lombardi
*Time and Frequency Division
Physical Measurement Laboratory*

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2189>

November 2021



U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
*James K. Olthoff, Performing the Non-Exclusive Functions and Duties of the Under Secretary of Commerce
for Standards and Technology & Director, National Institute of Standards and Technology*

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

National Institute of Standards and Technology Technical Note 2189
Natl. Inst. Stand. Technol. Tech. Note 2189, 65 pages (November 2021)
CODEN: NTNOEF

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2189>

Abstract

Accurate and reliable time signals are an essential part of critical infrastructure systems in the United States. Because these systems often cannot function properly without accurate time signals, timing system failures can have serious consequences, with the potential implications including economic loss, reduced safety and security, and loss of human life. The primary time synchronization sources for these systems are signals broadcast by Global Positioning System (GPS) satellites, which has understandably led to concerns about our dependency on GPS timing. This report provides an overview of how timing systems work and defines some basic terminology and specifications. It then discusses the regulatory timing requirements and GPS timing dependencies of United States critical infrastructure systems operating in the financial, telecommunications, and electric power sectors.

Keywords

Coordinated Universal Time; critical infrastructure systems; GPS; synchronization; time transfer.

Table of Contents

I. Introduction	1
II. Basic Terminology of Timing Systems and Timing Requirements.....	4
II.A. Time and Frequency Units and the Relationship between Frequency and Time	4
II.B. Coordinated Universal Time, the World’s Reference Clock	5
II.C. Traceability	6
II.D. Categories of Time and Frequency Measurements Specified in Requirements Documents.....	7
II.D.1 Time Synchronization of an OTM.....	7
II.D.2 Time Stamping the Occurrence of an Event.....	8
II.D.3 Frequency Syntonization	8
II.E. Timing Specifications: Accuracy, Stability, and Resolution	8
II.F. Free Running versus Disciplined Clocks	11
III. An Overview of Time Transfer Methods	15
III.A. One-Way Time Transfer.....	15
III.B. Loop-Back Time Transfer	16
III.C. Common-View Time Transfer	18
III.D. Two-Way Time Transfer.....	20
III.E. Dependencies of Other Time Distribution Systems on GPS	21
III.E.1 Restricted Access Time Distribution Systems.....	22
III.E.2 Public Access Time Distribution Systems.....	22
IV. Timing Requirements and Dependencies of Stock Exchanges.....	25
IV.A. The Importance of Time Synchronization to Stock Exchanges	25
IV.B. Time Synchronization Requirements of Stock Exchanges	27
IV.C. GPS Timing Dependencies of Stock Exchanges.....	30
V. Timing Requirements and Dependencies of the North American Power Grid.....	35
V.A. The Importance of Time Synchronization to the Power Grid	36
V.B. Time Synchronization Requirements of the Power Grid.....	39
V.C. GPS Timing Dependencies of the Power Grid	40
VI. Timing Requirements and Dependencies of Telecommunication Systems	42

VI.A. The Importance of Time Synchronization of Telecommunication Systems.....	43
VI.B. Time Synchronization Requirements of Telecommunication Systems	46
VI.C. GPS Timing Dependencies of Telecommunication Systems.....	47
VII. Summary.....	49
VIII. Acknowledgements.....	49
IX. References	50

List of Tables

Table 1. Time units and abbreviations	4
Table 2. Frequency units and abbreviations	4
Table 3. Relationship of frequency offset to time offset in a clock	5
Table 4. A summary of non-GPS public access time distribution systems	24
Table 5. Summary of stock market synchronization requirements	30
Table 6. Summary of power grid synchronization requirements	39
Table 7. Summary of telecommunications synchronization requirements	47
Table 8. GPS compared to other global navigation satellite systems	48

List of Figures

Fig. 1. A portion of the BIPM Circular T	6
Fig. 2. The rising or falling edge of a 1 pps electrical signal can serve as an OTM.....	8
Fig. 3. Estimating frequency accuracy from time difference data.....	9
Fig. 4. Time deviation of a GPS disciplined clock, indicating the clock’s stability.....	10
Fig. 5. Time stamp resolution	11
Fig. 6. Performance of a radio-controlled clock that corrects time but not frequency	12
Fig. 7. Block diagram of a GPS disciplined clock.....	12
Fig. 8. UTC(USNO) – UTC(NIST) via Circular T and via a GPSDC	13
Fig. 9. Performance of a GPSDC before and after the loss of GPS reception.....	14
Fig. 10. Performance of a previously disciplined cesium clock during holdover mode.....	14
Fig. 11. A one-way time transfer system	15
Fig. 12. A loop-back time transfer system.....	17
Fig. 13. A common-view time transfer system	19
Fig. 14. A two-way time transfer systems that utilizes a geostationary satellite	21
Fig. 15. U. S. stock volume from 1996 to 2018, and the contribution of HFT.....	26
Fig. 16. Block diagram of NIST Disciplined Clock (NISTDC)	31
Fig. 17. NISTDC common-view link to UTC(NIST)	32
Fig. 18. Accuracy of a NISTDC operated by a major U. S. stock exchange.....	32
Fig. 19. Accuracy of an NTP time server at a major U. S. stock exchange.....	33
Fig. 20. Regions and interconnections in the North American power grid	35
Fig. 21. The relationship between a phasor and a sine wave.....	37
Fig. 22. A synchrophasor measurement.....	37
Fig. 23. A total vector error (TVE) of 1%	38
Fig. 24. Map of PMUs in the North American power grid as of May 2017.....	39
Fig. 25. U. S. mobile phone versus land line subscriptions, 2000 to 2018.....	42
Fig. 26. A plesiochronous connection between networks that maintain their own PRS	43
Fig. 27. An accumulated time error of 125 μ s results in a cycle of frame slip.....	44
Fig. 28. Time division multiple access, users share the same frequency at different times	45
Fig. 29. Code division multiple access, users share the same frequency at the same time	46

I. Introduction

Although best known as a positioning and navigation system, the Global Positioning System (GPS) is also the world's primary system for the distribution of accurate (sub-microsecond) time. As such, GPS provides the time reference for numerous critical infrastructure systems in both the public and private sectors. The huge investments made in GPS research and development, beginning shortly after the launch of the first GPS satellite in 1978 [1], were greatly accelerated after the United States Air Force declared GPS to be a fully operational system available for civilian usage in 1993 [2]. This made it possible to embed GPS timing capability into nearly every type of instrument and system at a very low cost, in a way that is mostly transparent to users. Thus, sub-microsecond timing systems, which were difficult to construct and expensive to maintain prior to the invention of GPS, became almost trivial to build and almost free to use afterwards. Industry quickly exploited this tremendous new government resource by developing many new products and technologies.

The societal benefits of GPS are far reaching and hard to overstate, but the immense value it added to critical infrastructure systems eventually led to the question – what would happen to these systems if GPS went away? That question was formally raised in 1998 in *Presidential Decision Directive 63* which examined critical infrastructure protection. This directive called for “a thorough evaluation of the vulnerability of the national transportation infrastructure that relies on the Global Positioning System” and asked for an “independent, integrated assessment of risks to civilian users of GPS-based systems” [3]. This request resulted in the 2001 *Volpe Report* [4], which did not limit itself to studying the vulnerability of transportation systems to GPS positioning and navigation, but also brought the extensive use of GPS as a timing source to the nation's attention. The *Volpe Report* noted that “GPS-based timing synchronization is being used for transportation-related digital communication links and other applications such as telecommunications, banking, commerce, and the Internet,” [5], that GPS was already the “most frequently selected method for precise synchronization” in telecommunications systems, [6], and noting that backup systems were necessary:

“The GPS system cannot serve as a sole source for position location or precision timing for certain critical applications. Public policy must ensure that safety is maintained, even in the event of the loss of GPS.”

“Backups for positioning and precision timing are necessary for all GPS applications involving the potential for life-threatening situations or major economic or environmental impacts.” [7]

In 1997, prior to the publication of the *Volpe Report*, a significant effort had begun to develop a backup system for GPS by modernizing the existing Loran-C system. Loran-C was a radio navigation system that had preceded GPS by several decades, with the first stations beginning operation during World War II. Although its function was similar to GPS, its form was different. Its signals originated from ground-based transmitters rather than satellites and were broadcast in the low frequency (LF) part of the radio spectrum at 100 kHz, as opposed to GPS which operated in the ultra-high frequency (UHF) region (1575.42 MHz). In addition, Loran-C signals were broadcast at high power levels, with stations sometimes radiating as much one megawatt, as opposed to the sub-nanowatt spread-spectrum GPS signals. These differences were part of the appeal of Loran-C, allowing it to operate independently of GPS, and free from the sources of radio interference that affected GPS. The modernization effort included the addition of a digital time code to the legacy Loran-C signal, new timing hardware installed at the existing stations, and greatly improved modulation techniques [8], resulting in a new system called enhanced Loran, or eLoran.

From a timing systems viewpoint, eLoran compared favorably to all other existing systems as a potential alternative or backup system to GPS [9]. This conclusion was reached in several studies, most notably in the work of an Independent Assessment Team (IAT) that was organized via an inter-governmental effort to study the GPS vulnerability problem. The IAT report concluded that:

“After reviewing all prior studies and conducting detailed interviews, the IAT found that eLoran was the only system which could provide position, navigation, time, and frequency backup capability for all current and potential needs.” [10]

Despite these recommendations, support for eLoran waned. The system had not yet been fully implemented, nor was there any real assurance that industry planned to integrate eLoran into future critical infrastructure timing systems. It was true that due to the widespread acceptance and success of GPS, very few Loran-C users remained. Therefore, with the notable exception of one major telecommunications provider who was utilizing Loran-C as a backup frequency source for GPS [11], it was difficult to find examples of Loran-C being part of critical infrastructure timing systems. It was also true that even though eLoran performance would improve upon legacy Loran-C and could meet current timing requirements, a sizable performance gap would still exist between eLoran and GPS, and eLoran might be only a short-term solution. For these and many other factors, the decision was made to turn off all Loran-C stations in the United States in 2010 [12]. In the subsequent years, several global navigation satellite systems (GNSS) that are outside of U. S. control have become fully operational, and other terrestrial-based timing systems have been proposed as a GPS backup in the U. S., including a revitalization of eLoran [13], but as of this writing (February 2020) none have received sufficient support or endorsement to be included in critical infrastructure timing systems.

A 2019 study conducted by RTI International on behalf of the National Institute of Standards and Technology (NIST) estimated the economic consequences of 30-day GPS outages in various critical infrastructure sections. The estimated economic loss caused by a loss of GPS timing was severe in the telecommunications sector, with losses ranging between \$5.5 and \$14.2 billion. In the electric power sector, the estimated loss ranged from \$211.6 to \$338 million, with no substantial economic impact in the financial sector [14]. Overall, the report estimated that the combined loss of GPS position, navigation, and timing (PNT) services would have an adverse impact on the U. S. economy of at least \$1 billion per day [15].

On February 12, 2020, an executive order signed by the President brought the subject of GPS dependency to the forefront once again at the highest levels of government. The executive order notes that the PNT services provided by GPS have become a “largely invisible utility” and that the “disruption or manipulation of these services has the potential to adversely affect the national and economic security of the United States.” Most significantly, from a timing perspective, the executive order states that

“Within 180 days of the date of this order, the Secretary of Commerce shall make available a GNSS-independent source of Coordinated Universal Time, to support the needs of critical infrastructure owners and operators, for the public and private sectors to access.” [16]

The executive order reminds us that, nearly 20 years after its publication, the concerns expressed in the *Volpe Report* remain and the problem of GPS timing dependency remains unsolved. In fact, these concerns are now amplified because our dependence on GPS timing has continued to grow each year as new technologies are introduced. This report was written not to suggest alternatives or provide recommendations (a topic to be covered by NIST in a subsequent report), but rather to examine the current level of timing dependency in several critical infrastructure areas. Section II is an overview of the terminology of timing systems and timing requirements, providing a discussion of units and the

fundamentals necessary so that timing dependencies can be fully understood. Section III provides a basic overview of GPS and other time transfer and distribution systems. Sections IV through VI cover regulatory timing requirements and GPS timing dependencies in three critical infrastructure sectors; stock exchanges, the electric power grid, and telecommunications. When reading these sections, it is important to remember that the specifications for many industrial timing systems were written based on the level of timing accuracy that GPS could readily provide. Thus, in many cases, GPS dependency was built-in from the beginning. Section VII provides a brief summary.

II. Basic Terminology of Timing Systems and Timing Requirements

This section covers the fundamentals and basic terminology of timing systems, providing the information necessary to understand both timing requirements and dependencies. It begins by discussing measurement units and the categories of time and frequency information that are found in requirements documents. It then progresses to a discussion of Coordinated Universal Time (UTC), followed by discussions of traceability, the types of specifications included in timing requirements documents (specifically accuracy, stability, and resolution), and the characteristics of free running and disciplined clocks.

II.A Time and Frequency Units and the Relationship between Frequency and Time

The second, whose unit abbreviation is the small letter (s), is the standard unit for time interval, and the basis for all measurements of frequency and time. The second is one of the seven base units of the International System of Units, known as the SI. Since 1967, it has been defined as “The duration of 9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium-133 atom” [17]. Thus, the best physical realizations of the SI second are produced by cesium atomic clocks.

Most critical infrastructure timing systems are required to measure time intervals with durations much smaller than one second. The sub-second time interval units, and their unit abbreviations, are listed in Table 1. Currently, the microsecond, or 10^{-6} s, is the unit most often mentioned when discussing critical infrastructure timing requirements [18].

Table 1. Time units and abbreviations.

Unit Name	Unit Abbreviation	Duration (in seconds)
second	s	1
millisecond	ms	10^{-3}
microsecond	μ s	10^{-6}
nanosecond	ns	10^{-9}
picosecond	ps	10^{-12}

The hertz, abbreviated as Hz, is the standard unit for frequency. It represents the number of events that occur per second (the events are repeating pulses or cycles in an electrical signal). Pulses or square waves with a frequency of 1 Hz, or 1 pulse per second (pps) signals, are common in timing systems, because their period represents the standard unit of time interval, and because their arrival can be synchronized to agree with the time of a reference clock. However, the oscillators found inside of clocks produce faster signals, so frequency is usually expressed in multiples of the hertz (Table 2). The 1 pps timing signals are usually obtained by dividing a higher frequency signal from an oscillator, for example by dividing 10 MHz by 10^7 .

Table 2. Frequency units and abbreviations.

Unit Name	Unit Abbreviation	Events per second
pulse per second	pps	1
kilohertz	kHz	10^3
megahertz	MHz	10^6
gigahertz	GHz	10^9
terahertz	THz	10^{12}

Oscillators provide the heartbeat for all clocks. Oscillators generate a frequency, f , that is the reciprocal of the period of oscillation, T ; therefore $f = 1 / T$, where T is a time interval. Conversely, the period is the reciprocal of the frequency, $T = 1 / f$. A clock keeps time by measuring and counting the time intervals.

This relationship between time and frequency applies to all clocks but is perhaps easiest to visualize with a pendulum clock. Because the pendulum swings back and forth once per second, its period, T , is 1 s, and its frequency, f , is 1 Hz. The clock keeps time by assuming that the duration of one swing of the pendulum equals one second and counts pendulum swings (seconds) to measure longer intervals such as minutes or hours. A free running clock can only be as good as its oscillator. This means that if the pendulum frequency varies from 1 Hz, the clock will either gain or lose time, because it is counting seconds that are either shorter or longer than the SI second. A frequency offset will cause the clock to gradually accumulate a time offset during the period that it runs. This happens to all clocks, and even those with arbitrarily accurate frequencies will eventually require adjustment. Table 3 shows the relationship between frequency offset and time offset.

Table 3. Relationship of frequency offset to time offset in a clock.

Frequency Offset	Period that Clock Runs	Accumulated Time Offset
$\pm 1.00 \times 10^{-3}$	1 s	± 1 ms
$\pm 1.00 \times 10^{-6}$	1 s	± 1 μ s
$\pm 1.00 \times 10^{-9}$	1 s	± 1 ns
$\pm 2.78 \times 10^{-7}$	1 h	± 1 ms
$\pm 2.78 \times 10^{-10}$	1 h	± 1 μ s
$\pm 2.78 \times 10^{-13}$	1 h	± 1 ns
$\pm 1.16 \times 10^{-8}$	1 day	± 1 ms
$\pm 1.16 \times 10^{-11}$	1 day	± 1 μ s
$\pm 1.16 \times 10^{-14}$	1 day	± 1 ns

II.B Coordinated Universal Time, the World’s Reference Clock

The official, internationally agreed upon reference for world time is Coordinated Universal Time, abbreviated as UTC. The Bureau International des Poids et Mesures (BIPM) in Sèvres, France is the organization that maintains and distributes UTC [19]. The BIPM is an intergovernmental organization that was established by the Metre Convention of 1875. It was then given the mandate to provide a single, coherent system of measurements by establishing an International System of Units (SI).

UTC is an international weighted-average ensemble time scale, which simply means that it is obtained by computing a weighted average of time kept by other time scales located around the world (an ensemble is a group of items that is viewed collectively rather than individually) [20]. As of December 2019, a total of 82 timing laboratories located in 62 nations contribute data to UTC [21]. Four of these laboratories are located in the United States, the National Institute of Standards and Technology (NIST) in Boulder, Colorado, the United States Naval Observatory (USNO) and the Naval Research Laboratory (NRL), both located in Washington, DC, and the Applied Physics Laboratory (APL) located in Laurel, Maryland.

Each contributor to UTC maintains its own time scale, which is known as $UTC(k)$, where k represents the name of the laboratory, for example $UTC(NIST)$ or $UTC(USNO)$. Many of the $UTC(k)$ time scales also keep time by averaging an ensemble, but unlike UTC, where the ensemble consists of a group of times scales located around the world, the $UTC(k)$ ensemble is a group of atomic clocks, usually all located in the same laboratory or facility. Via their local $UTC(k)$ time scales, a total of 414 atomic clocks contributed data to UTC in December 2019. Most of these are cesium clocks, but other types of atomic clocks, most notably hydrogen masers, also contribute to UTC.

The BIPM distributes UTC via its *Circular T* document [22], which has been published monthly since 1988. The *Circular T* (Fig. 1) shows the time difference between UTC and each contributing laboratory, or $UTC - UTC(k)$, at 5-day intervals. Because UTC is a virtual or “paper” clock whose time is only known after the fact and that does not produce any physical signals, no exact physical realization of UTC exists. Fortunately, however, the $UTC(k)$ time scales do produce physical timing signals, in the form of electrical pulses or sine waves, and routinely serve as reference clocks for time distribution systems. Many of the $UTC(k)$ time scales are very close approximations of UTC, often differing from UTC by just a few nanoseconds, as the *Circular T* data indicates.

Bureau International des Poids et Mesures		CIRCULAR T 384		ISSN 1143-1393						
		2020 JANUARY 09, 15h UTC								
BUREAU INTERNATIONAL DES POIDS ET MESURES THE INTERGOVERNMENTAL ORGANIZATION ESTABLISHED BY THE METRE CONVENTION PAVILLON DE BRETEUIL F-92312 SEVRES CEDEX TEL. +33 1 45 07 70 70 tai@bipm.org										
The contents of the sections of BIPM <i>Circular T</i> are fully described in the document " Explanatory supplement to BIPM Circular T " available at ftp://ftp2.bipm.org/pub/tai/publication/notes/explanatory_supplement_v0.2.pdf										
1 - Difference between UTC and its local realizations $UTC(k)$ and corresponding uncertainties. From 2017 January 1, 0h UTC, $TAI-UTC = 37$ s.										
Date	2019/20 0h UTC	NOV 27	DEC 2	DEC 7	DEC 12	DEC 17	DEC 22	DEC 27	Uncertainty/ns	Notes
		MJD 58814	58819	58824	58829	58834	58839	58844	u_A u_B u	
Laboratory k		[UTC-UTC(k)]ns								
AOS (Borowiec)		6.4	5.1	3.7	3.0	2.7	2.3	2.3	0.4 4.2 4.3	
APL (Laurel)		0.2	0.6	1.4	2.5	1.3	4.1	-	0.3 11.2 11.2	
AUS (Sydney)		-371.4	-372.7	-369.4	-381.2	-378.4	-391.5	-382.5	0.3 6.6 6.6	
BEV (Wien)		-0.6	-1.7	3.2	8.1	9.9	7.3	5.8	0.5 3.4 3.5	
BIM (Sofiya)		11897.0	11893.8	11914.7	11948.3	11973.1	12025.6	12063.0	0.7 7.1 7.2	
BIRM (Beijing)		-3.6	0.0	-2.0	-5.0	-7.1	-6.0	-2.0	0.3 2.9 2.9	
BOM (Skopje)		-2859.6	-2875.5	-2891.3	-2933.5	-2962.4	-2981.6	-2995.5	1.5 8.5 8.7	(1)
BY (Minsk)		0.2	-0.7	-0.9	-0.7	-0.6	-0.7	0.1	1.5 12.2 12.3	
CAO (Cagliari)		-15910.5	-16014.6	-16126.0	-16227.3	-16341.5	-16450.0	-16563.5	1.5 20.0 20.1	
CH (Bern-Wabern)		3.5	2.7	3.2	4.2	3.1	2.9	3.1	0.3 2.3 2.3	
CNES (Toulouse)		-1.2	-0.2	1.3	2.4	2.8	3.3	3.8	0.3 2.9 2.9	
CNM (Queretaro)		-2.1	-9.8	-2.8	-4.4	-3.6	-5.8	-3.8	2.5 11.2 11.5	
CNMP (Panama)		-2.1	-96.3	-14.0	-9.3	-4.0	16.9	14.4	0.7 7.4 7.5	
DFNT (Tunis)		15025.6	15236.7	15469.1	15670.3	15888.4	16094.6	16312.7	0.7 20.0 20.0	
DLR (Oberpfaffenhofen)		-	1177.5	1194.0	1205.9	1217.6	1232.4	1244.7	0.7 3.5 3.5	
DMDM (Belgrade)		4.8	9.2	5.0	6.4	12.9	11.4	10.6	0.3 3.5 3.5	
DTAG (Frankfurt/M)		7.8	8.9	9.8	8.5	8.1	11.5	10.9	0.3 2.9 2.9	
EIM (Thessaloniki)		-3.4	6.5	-1.7	6.6	-1.5	3.6	4.1	3.0 11.2 11.6	
ESTC (Noordwijk)		1.1	0.5	-0.3	-1.0	-1.4	-1.1	-0.3	0.3 3.2 3.2	
HKO (Hong Kong)		144.9	150.4	158.8	169.7	166.6	171.9	170.6	1.0 2.9 3.0	
ICE (San Jose)		0.8	10.1	15.4	22.5	12.3	11.0	9.7	5.0 20.0 20.6	
IDN (Serpong-Tangerang)		1561.2	1562.8	1572.8	1593.6	1643.1	1646.7	1663.3	0.7 2.9 3.0	
IFAG (Wetzell)		-818.6	-818.9	-823.3	-823.8	-834.6	-833.9	-831.7	0.3 3.0 3.0	

Fig. 1. A portion of the BIPM *Circular T*.

II.C Traceability

Traceability is an important characteristic of all critical timing systems, because it ensures that all measurements of time, regardless of where they are made, use the same measurement units and ultimately link back to the same reference. The *International Vocabulary of Metrology (VIM)* defines metrological traceability in Section 2.41 (6.10) as “the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty” [23].

For nearly all areas of metrology, including time and frequency, the SI units serve as ultimate measurement references. The SI units are definitions of ideal values and as such are “perfect”, meaning that they have a measurement uncertainty of 0. Systems that physically realize the units, by generating time or time signals, will of course introduce some measurement uncertainty. Because UTC is the world’s best physical approximation of the SI second, all time and frequency measurements should be referenced to UTC, and all traceability chains should originate with UTC [24]. However, as previously noted, UTC is not a physical standard, and actual time measurements will need to be made with respect to one of the local UTC(*k*) time scales, such as UTC(NIST) in the United States. This is usually not a problem, because the *Circular T* regularly publishes the UTC – UTC(*k*) time differences, which completes the traceability chain back to UTC and the SI unit of time.

One of the many advantages of GPS time is that it is inherently traceable to the SI and UTC, because it is referenced to UTC(USNO). The GPS navigation messages (subframe 4, page 18) [25] contain the parameters necessary to convert GPS time to UTC(USNO) and nearly all GPS receivers apply these corrections by default. The time obtained from the GPS signal in space as transmitted by the satellite can be considered directly traceable to UTC(USNO), with an uncertainty of a few nanoseconds [24], and most GPS receivers can produce time within 1 μ s of UTC without any need for calibration [18].

II.D Categories of Time and Frequency Measurements Specified in Requirements Documents

Three categories of time and frequency measurements are important to critical infrastructure timing systems: time synchronization of an on-time marker (OTM), time stamping the occurrence of an event, and frequency syntonization. Each category of measurements should be traceable to a UTC reference, as specified in the requirements document. The three categories are described below.

II.D.1 Time Synchronization of an OTM

For the purposes of critical infrastructure systems, time synchronization can be defined as the process of either measuring the time offset between the clock under test and a reference UTC clock, adjusting the clock under test to agree with a reference clock, or doing both things (measuring and then adjusting). In some cases, just knowing that the measured time difference is small enough to meet the requirement is enough, because it indicates that the clock under test is within a specified tolerance. However, in other cases the clock under test must be adjusted to bring it or keep it within tolerance. This is done by issuing a correction to either its time or its frequency in a way that reduces its time offset with respect to the reference clock to as close to zero as possible.

The reference clock outputs either an OTM, a time code, or both, at a time coincident with the UTC second. In many cases, the reference clock generates a 1 pps signal, and the OTM is sent on either the rising or falling edge of the square wave pulse (Fig. 2). Or, the OTM can be sent as part of the time code, typically at the beginning or end of the time code transmission.

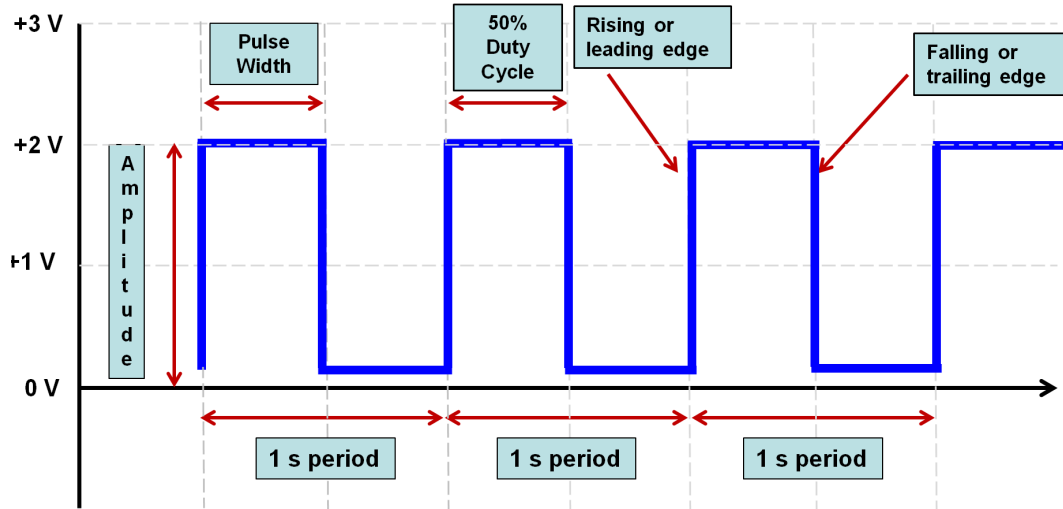


Fig. 2. The rising or falling edge of a 1 pps electrical signal can serve as an OTM.

II.D.2 Time Stamping the Occurrence of an Event

Many critical infrastructure timing systems need to time stamp when an event occurred, by recording and storing time-of-day information. The time stamp, sometimes referred to as a time tag, typically contains time-of-day information such as the UTC hour, minute, and second, and fractional parts of the second. The time stamp may also include date information, sometimes in the form of year, month, day, or in a form where the date can be calculated by looking at the number of seconds, days, or weeks, from a given epoch. The time stamp essentially labels the OTM, in other words it indicates the time-of-day when the OTM was generated. The time difference between the generation of the time stamp and the reference clock, must be within the tolerance of the timing requirement.

II.D.3 Frequency Syntonization

Some critical infrastructure timing systems require the oscillator in the clock under test to be within a specified tolerance of the UTC frequency. This is known as syntonization and is analogous to the time synchronization of the OTM. It can involve either measuring the frequency offset between the clock under test and a reference UTC clock, adjusting the clock under test frequency to agree with a reference clock, or doing both things (measuring then adjusting). The frequency offset is usually expressed in scientific notation as a unitless number as was shown in Table 3. The International Telecommunication Union (ITU) definition of a primary reference clock (PRC) is one of the best known examples of a frequency requirement:

The long-term accuracy of the PRC should be maintained at 1 part in 10^{11} or better with verification to coordinated universal time (UTC)..... The maximum allowable fractional frequency offset for observation times greater than one week is 1 part in 10^{11} , over all applicable operational conditions.
[26]

II.E Timing Specifications: Accuracy, Stability, and Resolution

In international metrology, accuracy is defined as the “measure of agreement between a measured quantity value and a true quantity value” [23]. For the purpose of time accuracy as it applies to specifications, the measured quantity value is obtained by comparing the clock under test to UTC, which represents the true quantity. The time difference between the clock under test and UTC indicates the accuracy. In critical

infrastructure systems, the time accuracy requirement often equals or approaches 1 μs . In most cases, this requirement is not an average, but rather a threshold that should not be exceeded. In other words, it does not mean that a clock can meet the requirement by keeping time to within 1 μs of UTC on average, but instead means that a clock should never deviate by more than $\pm 1 \mu\text{s}$ from UTC. For this reason, statistics such as MTIE, or maximum time interval error, are sometimes included in requirements documents for telecommunication systems, to indicate the peak time deviation of a clock, or worst case scenario [27].

For the purpose of frequency accuracy, the measured quantity value can be obtained by looking at how the time accuracy changes over time, or $\Delta t / T$, where Δt indicates the change in time of a clock during an interval, and T indicates the duration of the interval. If the reference is a UTC source of frequency, then the frequency offset with respect to UTC and the frequency accuracy are equivalent. Frequency accuracy over a given interval can be estimated by fitting a linear least squares line to a series of time difference measurements, and then using the slope of the least squares line to estimate Δt . To illustrate this, Fig. 3 shows a sample phase graph of an oscillator that was compared to a reference for a period of 7 days. During this period, the total accumulated time difference, Δt , was about 900 ns, as indicated by both the actual data and the least squares line that was fitted to the data. From the slope of the least squares line, the frequency accuracy can be estimated as 1.5×10^{-12} .

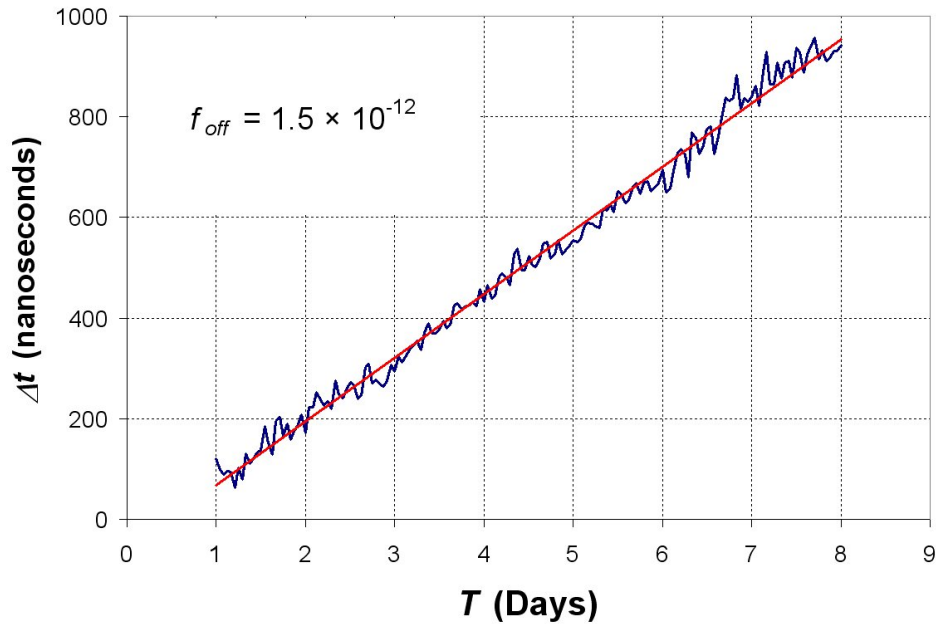


Fig. 3. Estimating frequency accuracy from time difference data.

Stability differs from accuracy because it does not indicate how closely the time or frequency of clock under test agrees with UTC. Instead, it indicates the change in the time offset or frequency offset during a given time interval. The stability of a clock indicates the potential accuracy of the clock if calibrated and by doing so establishes the limit of its accuracy. This is because a clock's accuracy during a given interval cannot be better than its stability during that same interval. The standard statistics for estimating frequency stability are the Allan deviation (ADEV), $\sigma_y(\tau)$, and the Modified Allan deviation (MDEV), $Mod \sigma_y(\tau)$. The standard statistic for estimating time stability is the Time deviation (TDEV), $\sigma_x(\tau)$. The symbol σ , or sigma, denotes standard deviation, y denotes frequency, x denotes time, and τ , or tau, denotes the duration of the averaging or observation period [27, 28].

Each of these statistics can help identify the type of noise that causes the frequency or time of a clock to change. The chief advantage of using MDEV instead of ADEV is that it can distinguish between the two types of phase noise (white and flicker). Because it relates to time and synchronization, TDEV is more likely to appear in requirements documents than ADEV or MDEV, particularly in telecommunication requirements [25]. However, TDEV is closely related to MDEV, and is obtained by simply multiplying $Mod \sigma_y(\tau)$ by $(\tau / \sqrt{3})$ [27].

Figure 4 shows the TDEV of a GPS disciplined clock (GPSDC) for intervals ranging from 1 minute to more than one day. The time stability is about 6 ns at $\tau = 1$ hour, but just 1.4 ns at $\tau = 1$ day. As noted earlier, these values provide an indication of both the potential and the limits of its accuracy. For example, if this clock was uncalibrated (meaning that no compensation has been made for receiver and antenna delays) and had a daily time offset with respect to UTC of 500 ns, TDEV is indicating that it could be calibrated to be much closer to UTC. In fact, by compensating for delays it could be calibrated until its time offset with respect to UTC is near 0, because the variation in its daily average time is only about 1 ns.

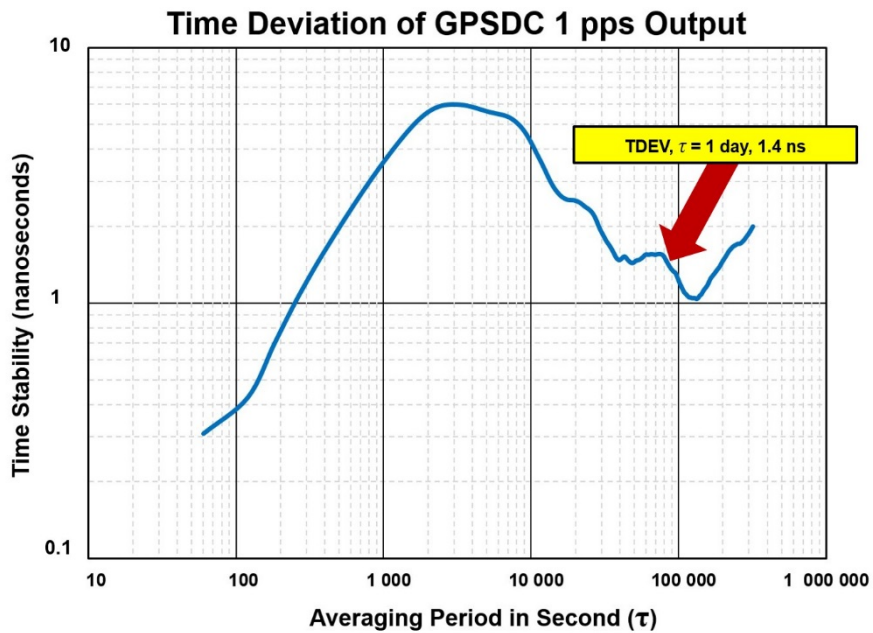


Fig. 4. Time deviation of a GPS disciplined clock, indicating the clock’s stability.

Resolution is defined as the “smallest change in a quantity being measured that causes a perceptible change in the corresponding indication” [23]. It is often referred to in requirements documents, typically when referring to time stamps, but is sometimes denoted with other names such as granularity or precision. When time stamping the occurrence of an event, the time stamp must have enough resolution to display the required number of digits. For example, if a time stamp is required to have 1 μ s resolution, six digits are required to the right of the decimal point, and nine digits to the right of the decimal point would be required for 1 ns resolution, as shown in Fig. 5. This, of course, requires the time stamp to be generated by a clock capable of incrementing in steps at least as small as the resolution requirement, so that all digits, including the least significant digit, contain meaningful information.

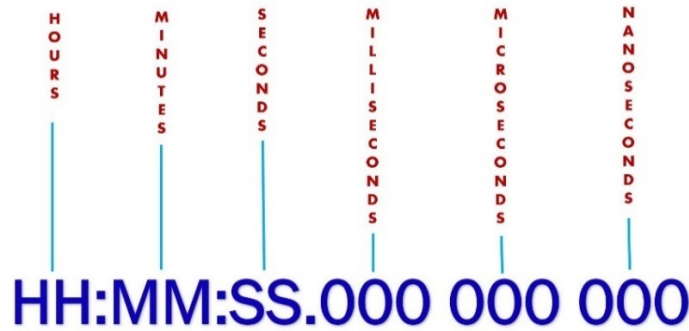


Fig. 5. Time stamp resolution.

II.F Free Running versus Disciplined Clocks

A free running clock is a clock whose frequency and time are not being adjusted, either manually or automatically, and thus keeps time commensurate with the frequency accuracy of its oscillator. Because no oscillator is perfect, free running clocks always accumulate a time offset (Table 3). For this reason, free running clocks are seldom utilized in critical infrastructure systems for time synchronization unless the requirements are modest, or unless cesium clocks are deployed. They are, however, often used to meet frequency requirements. For example, relatively low-priced rubidium clocks can provide frequency accurate to within parts in 10^{10} over long intervals while free running and are widely deployed in telecommunication networks. This still results in an accumulated time offset of tens of microseconds per day which makes them unsuitable for many synchronization requirements. Cesium clocks provide frequency accuracy of about 1×10^{-12} in the worst case and near 1×10^{-14} in the best case, an accumulated time offset ranging from about 1 to 100 ns per day, or small enough to meet many synchronization requirements over long intervals. Unfortunately, however, cesium clocks are too expensive, too large, and require too much maintenance to be considered for wide-scale deployment.

There are two basic methods that can keep a clock from accumulating a time error. The first method is to issue a periodic time step that makes the clock temporarily agree with a reference clock. After the time step, the clock will immediately begin to accumulate a time offset, but that offset will eventually be removed again by the next time step. This method is used, for example, by the low-cost radio-controlled clocks that receive 60 kHz signals from NIST radio station WWVB and utilize small and inexpensive quartz crystals as their oscillators. Figure 6 shows an example of a radio controlled wristwatch that synchronizes to a time code from WWVB every hour from midnight until 4 a.m. The clock then free runs for 20 hours until the next synchronization on the following midnight. During the period between synchronizations, it accumulates a time offset of 450 ms, indicating a poor frequency accuracy of about 6×10^{-6} [29]. The “sawtooth” pattern shown in Fig. 6 is found in many timing systems, even in some GPS systems, and with better oscillators and more frequent time steps this method can meet many synchronization requirements.

The second and preferred method is to adjust the oscillator frequency in a way that removes the frequency offset so that a time offset no longer accumulates. This is the method used by disciplined clocks, including the GPSDCs that are the workhorses of critical infrastructure timing systems. A GPSDC has at least three parts: a local oscillator (LO), a receiver and antenna that receive timing signals from the GPS satellites, and a frequency or phase comparator. The comparator measures the phase or time difference between the LO and GPS and converts this difference to a frequency correction that is periodically applied to the LO. By continuously repeating this process, the LO is locked to the GPS reference and can largely replicate its performance. No manual adjustment of the LO is ever necessary.

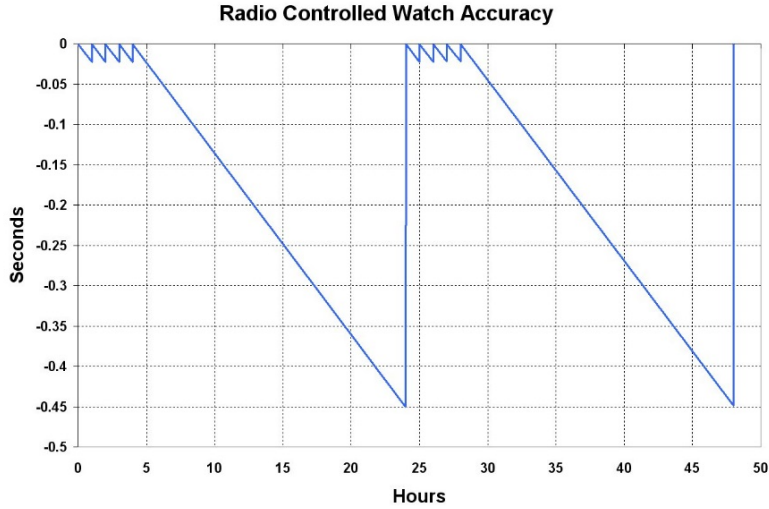


Fig. 6. Performance of a radio-controlled clock that issues time steps but does not adjust oscillator frequency.

A few basic elements are present in most GPSDC designs. The LO is usually a quartz oscillator, but more expensive models include an atomic rubidium oscillator. The GPS receiver is nearly always a single-frequency (L1 band, 1575.42 MHz) instrument that decodes the coarse acquisition (C/A) code broadcast by the satellites. The receiver is connected to a small antenna and typically outputs 1 pps or a similar low frequency signal. Various types of phase comparators are used to measure the difference between the GPS signal and the LO signal. The output of the phase comparator is read by a microcontroller (MCU) whose firmware executes a control loop, which is often some variation of a proportional-integral-derivative (PID) controller, and the control loop keeps the LO locked to GPS by continually issuing frequency corrections that keep the phase difference as small as possible. In a simple GPSDC design, the LO might be a voltage-controlled oscillator (VCO) and frequency corrections are sent by varying the control voltage, as shown in Fig. 7. The LO provides disciplined output signals, typically 1 pps for timing, and 10 MHz for frequency [30].

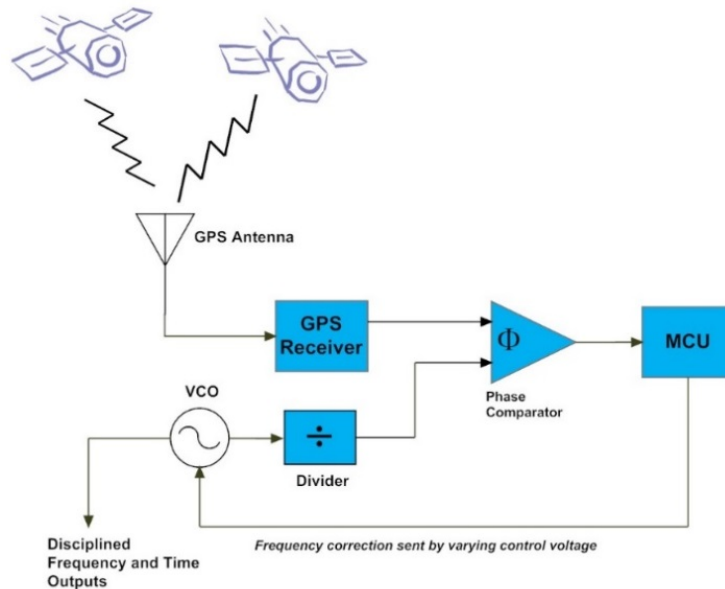


Fig. 7. Block diagram of a GPS disciplined clock.

A GPSDC can keep accurate time indefinitely, or for as long as GPS signals can be received [30, 31]. Because the GPS signals are referenced to UTC(USNO), GPSDCs are self-synchronizing, inherently accurate, inherently stable, and inherently traceable for both time synchronization and frequency. Unlike all free running clocks, a GPSDC does not accumulate any significant time offset with respect to UTC. To illustrate this, Fig. 8 shows the time differences between UTC(NIST) and UTC(USNO) over a 12.5 year period via two different methods. The first method, shown in red on the graph, obtains the time UTC(USNO) – UTC(NIST) difference from the *Circular T*. The second method, shown in blue, obtains UTC(USNO) from a calibrated GPSDC that was directly compared to UTC(NIST) during the entire period (January 2006 to June 2018). The GPSDC ran continuously during that period, but to match the reporting interval of the *Circular T*, only one value (a 24 hour average) is shown every five days. The GPSDC measurement has more outliers, but the structure of the data varies only slightly. Both methods show that UTC(USNO) and the GPSDC agreed to within ± 25 ns of UTC(NIST) for more than a decade.

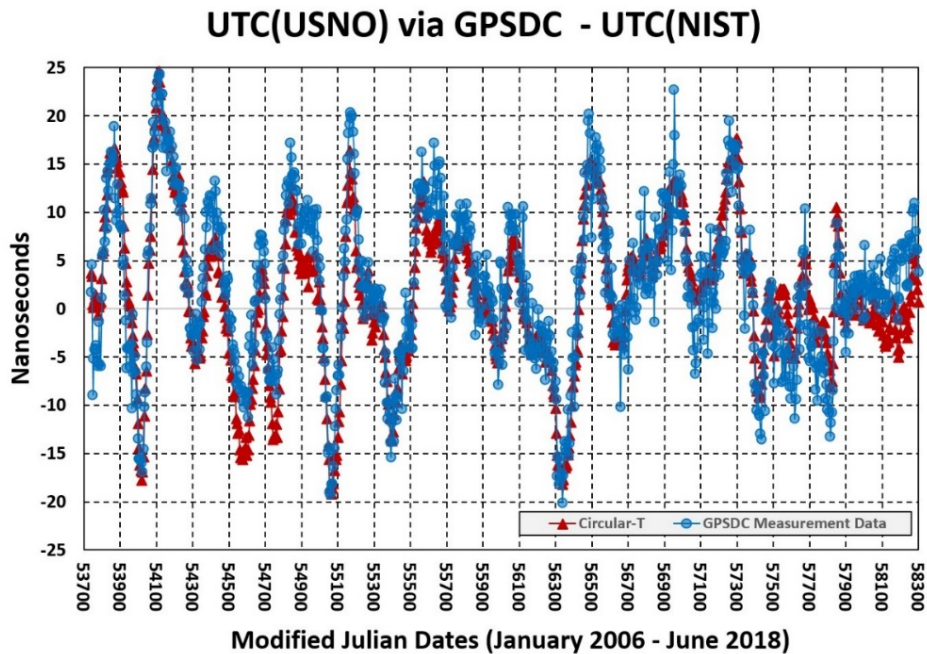


Fig. 8. UTC(USNO) – UTC(NIST) via Circular T and via a GPSDC.

If GPS cannot be received, a GPSDC will go into holdover mode, where its accuracy will now be limited by the quality of its LO, and if one is present, the quality of the holdover algorithm embedded in the GPSDC’s firmware. Holdover algorithms work by predicting current time errors, based on the history of the local oscillator and the adjustments it received when GPS was available, and continuing to issue corrections to compensate for those errors [32]. Because not all GPSDCs include a holdover algorithm, some immediately become free running clocks when GPS signals are lost. In that case, a GPSDC with a rubidium LO will depart from UTC at a slower rate than a GPSDC with a quartz LO. Figure 9 shows a disciplined rubidium clock that maintained 1 μ s synchronization for about 73 hours after its antenna was disconnected. The time error then began to accumulate more rapidly, reaching about 5 μ s after about 110 hours before it began to relock when the antenna was reconnected.

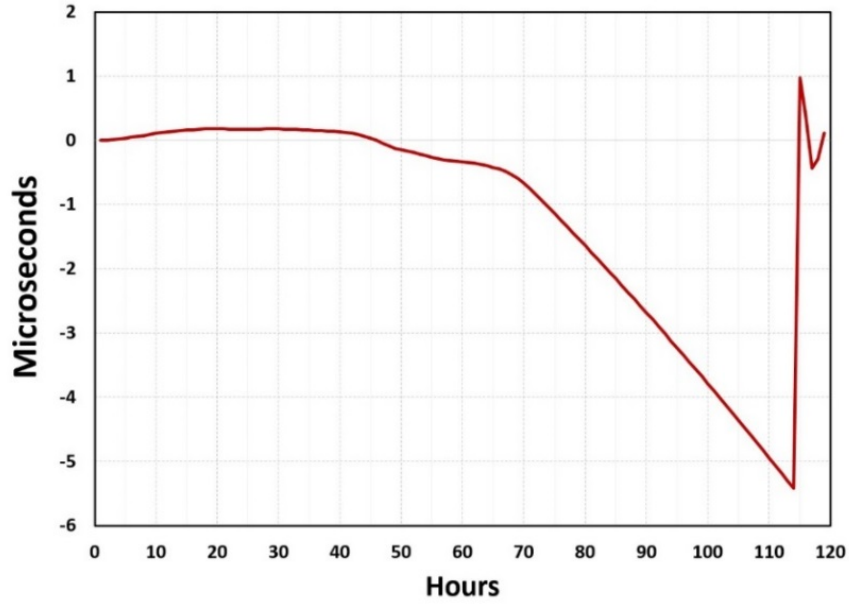


Fig. 9. Performance of a disciplined rubidium clock before and after the loss of GPS reception.

For the reasons discussed earlier in this section (cost, size, reliability) only a small number of cesium clocks are found in critical infrastructure systems, which is unfortunate because their holdover capability far exceeds that of a rubidium clock. Figure 10 shows the performance of a cesium clock in holdover mode that had previously been locked to UTC(NIST). It remained within 300 ns (0.3 μ s) of UTC(NIST) after free running for about eight months. Because the cesium clock frequency had been optimally adjusted while it was locked, the time offset increased at a rate of just 1.2 ns per day (frequency offset of $\sim 1 \times 10^{-14}$) while in holdover mode. Even if the cesium clock frequency has not been optimally adjusted, the time offset is likely to increase at a rate of less than 10 ns per day.

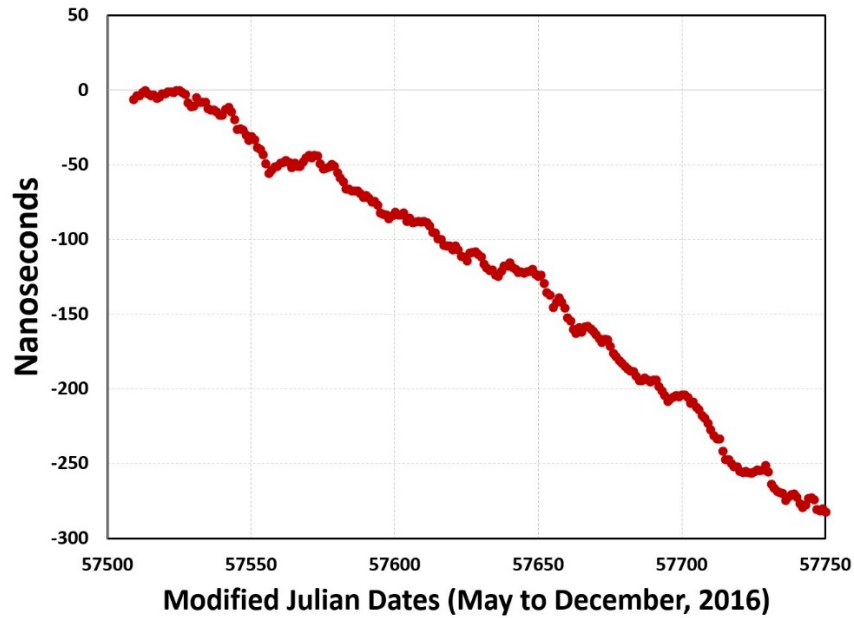


Fig. 10. Performance of a previously disciplined cesium clock after going into holdover mode.

III. An Overview of Time Transfer Methods

Time transfer is the practice of transferring time from a reference clock at one location and using it to measure or synchronize a clock at another location. Whenever a clock is compared to another clock or synchronized with another clock, time transfer is taking place. In addition to the satellite signals that critical infrastructure systems are so highly dependent upon, time is also transferred through a variety of other mediums; including terrestrial-based radio signals, coaxial cables, optical fibers, telephone lines, and computer networks.

The various methods of transferring time can be organized into four general categories known as one-way, loop-back, common-view, and two-way. It is helpful to note that the first two methods, one-way and loop-back, are the most common methods used by publicly accessible systems, such as GPS, that are routinely used to synchronize clocks. Therefore, they generally deliver both a time code, containing time-of-day information, and an OTM. The second two methods, common-view and two-way, are usually associated with high accuracy time transfer, and generally just deliver an OTM in the form of a 1 pps signal and not a time code. Because the time difference between two unlabeled 1 pps signals cannot exceed ± 0.5 s, a 1 pps signal from a clock under test cannot be fast or slow by more than 0.5 s with respect to UTC. Therefore, high accuracy time transfer systems often operate with the assumption that the remote clock already has correct time-of-day information with respect to UTC, and that this information can be used to correctly label the OTM. To make this assumption true, common-view and two-way systems often receive time-of-day information from a one-way or loop-back system. For example, a system that transfers time via common-observation of satellites might obtain time-of-day from a loop-back system via the Internet. The next four sections discuss the four categories of time transfer systems.

III.A One-Way Time Transfer

All time transfer systems have a reference clock at their source (point *A*). Information from the reference clock is encoded on a signal that is transmitted through a wired or wireless medium to its destination (point *B*), where a remote clock is located. In the simplest form of time transfer, known as the “one-way” method (Fig. 11), the remote clock is then synchronized with the time from the reference clock. The one-way method is typically employed by broadcast systems such as GPS that distribute time to multiple receivers (the number of receivers is unknown to the transmitter) that reside within the coverage area of the signal.

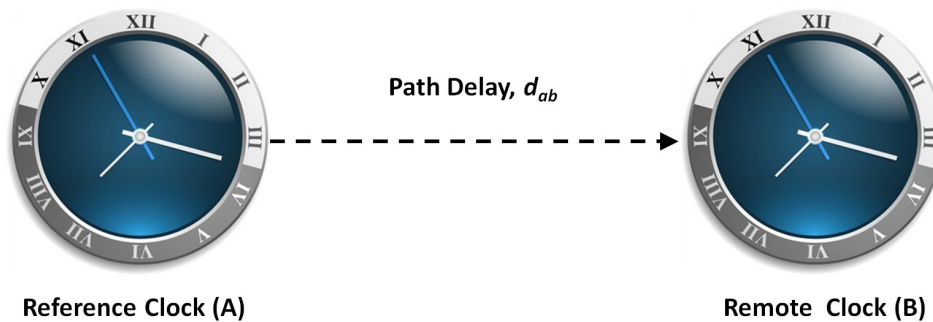


Fig. 11. A one-way time transfer system.

The path delay through the medium between points *A* and *B* is indicated by the variable, d_{ab} . Even if the reference clock is a perfect time source, the accuracy of the time transferred to the remote clock can be no better than the uncertainty of the path delay measurement, or no better than our knowledge of d_{ab} [33]. This simple fact can be thought of as the first principle of all time transfer systems.

To illustrate the concept of path delay, consider two clocks separated by a distance of 1000 km. A radio signal containing the time from a reference clock is transmitted across this 1000 km path, where it is received and used to synchronize another clock. Radio signals travel at the speed of light, which is 299 792 458 m / s, or roughly 3.3 μ s / km. Therefore, the time will be 3.3 ms late when received by the remote clock because $d_{ab} = 3.3$ ms.

For some one-way time transfer systems, the path delay is not considered important and is simply ignored. For example, the purpose of a consumer grade radio-controlled wall clock or wristwatch that receives WWVB is simply to display the time-of-day, and it is unlikely that anyone viewing the clock will be interested in or need time accurate to better than 1 s. Adjusting the clock's display to compensate for a path delay that is likely to be no more than 20 ms would provide no advantage as it would not be detectable to the human eye [29]. When better accuracy is required, other types of one-way time transfer systems have compensated for a portion of the path delay by sending the OTM out early, a method that works in systems where an average or minimum path delay can be estimated for all remote clocks. For example, fixed OTM advances have been implemented in time transfer systems operating over telephone lines [34], where delays through telephone circuits that exceed tens of milliseconds are usually unavoidable. Custom systems where a fixed path delay is introduced by a coaxial cable or fiber optic line can also benefit from this method. For example, if a time signal is sent between two buildings on the same campus via a coaxial cable, the delay of the cable can be measured or estimated, and the OTM can be advanced by that amount.

For critical infrastructure timing systems, either ignoring or coarsely estimating path delay is not an option; instead d_{ab} must be accurately measured and compensated for before correcting the time of the remote clock. GPS does exactly that, which is why its accuracy easily exceeded all the one-way time transfer systems that preceded it. The actual path delay is quite large – the GPS satellites are in semi-synchronous orbit at an altitude of about 20 200 km (about half the height of geostationary orbit) and it takes at least 65 ms, or slightly more than 1/16 of a second, for their signals to reach a clock on Earth. However, because the satellites are at known positions, and because the speed of light is a known constant, a GPSDC can measure and remove this path delay. It does so by making a series of range measurements between the its local clock and multiple satellites, a process known as trilateration, and using this information to compute its position on Earth. Once the receiver position is known, the distance between the GPSDC and the satellites can be calculated and converted to a time delay. Additional, and much smaller, corrections are applied to this time delay to obtain the final estimate of d_{ab} and to make the GPSDC even more accurate. For example, the satellite signals are delayed as they pass through the ionosphere and troposphere. Corrections that partially compensate for both delays are usually automatically made with algorithms contained in the receiver's firmware [35, 36].

As a result of these transparent path delay corrections, nearly all GPSDCs produce time within 1 μ s of UTC straight out of the box without any effort on the part of the user. This fact, coupled with their low cost, their ability to be easily embedded in other hardware, their small antennas, and their widespread availability, explains why GPSDCs are so widely deployed and so heavily depended upon in critical infrastructure timing systems. Their performance is often taken for granted and one microsecond accuracy is a modest estimate in most cases. If the GPSDC antenna position was properly surveyed, a process that many units perform automatically, and if the user enters carefully estimated delays for the receiver, antenna, and antenna cable, then an accuracy of < 0.1 μ s with respect to UTC is usually easy to achieve [30].

III.B Loop-Back Time Transfer

This method of time transfer employs a loop-back test to measure round-trip path delay (Fig. 12). For example, an OTM is sent from a reference clock (A) to a remote clock (B) over the path d_{ab} . The remote clock (B) then sends the OTM back to the transmitter (A) over the path d_{ba} . The one-way path delay is then assumed to be one half of the measured round-trip delay, or $(d_{ab} + d_{ba}) / 2$. This method is often easy to

implement in point-to-point applications, for example, to send time from a server clock to a client clock via a telephone or computer network; but is less practical to use through a wireless medium.

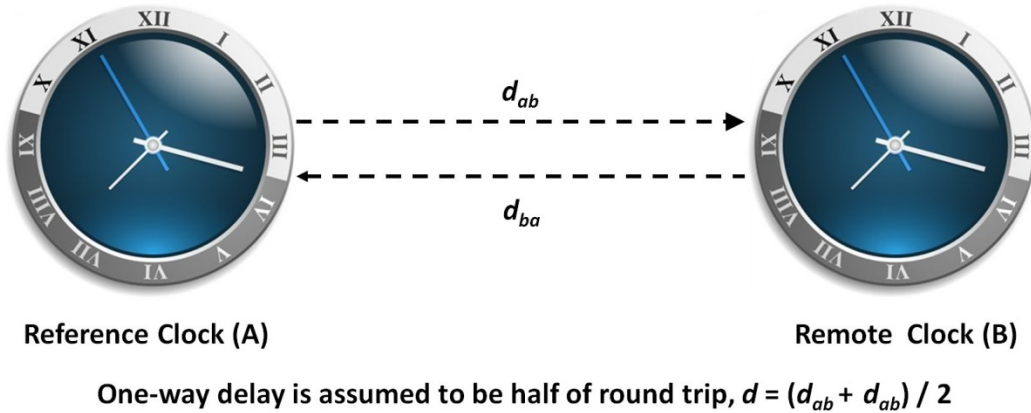


Fig. 12. A loop-back time transfer system.

The method used by a loop-back system to compensate for path delay varies slightly depending upon whether the round-trip delay is known by the server clock or the client clock. The round-trip delay is only known to the server clock (the reference clock) in a method used by the NIST Automated Computer Time Service (ACTS) [29], which requires the client to request a time code by making a telephone call to the server. The server answers the call and sends a time code and OTM to the client at a time T_1 , and then waits for the client to return the OTM, recording its arrival at time T_2 . The server now has an estimate of the round-trip delay, or $T_2 - T_1$. The server clock then advances the next OTM sent to the client by a time interval equal to $(T_2 - T_1) / 2$, which is an estimate of the one-way path delay.

The most common method used to synchronize computer clocks is the Network Time Protocol (NTP) [37, 38], which has widely been implemented on the public Internet. Like ACTS, NTP also works by having the client request the time from a server. However, the client records when the time was requested, T_1 , and the server records when the request was received, T_2 . The server then sends a time code to the client at T_3 . Thus, $T_3 - T_2$ is the server processing time. The client receives the time code at T_4 . Thus, only the client (the remote clock) has an estimate of the round-trip delay, which is $(T_4 - T_1) - (T_3 - T_2)$. The client divides the round trip delay by two to estimate the one-way path delay, and adds this quantity to the received time, T_3 , to compensate for the path delay. This method has the advantage over ACTS of not requiring the server to measure round trip delay or to advance the OTM.

The Precision Time Protocol (PTP), defined by the IEEE-1588 standard [39] is another loop-back time transfer method designed to synchronize network clocks. It is capable of better accuracy than NTP for numerous reasons, including the use of more hardware (NTP is usually implemented entirely in software), more frequent synchronization requests, and the fact that it is usually implemented in a local area network (LAN), as opposed to NTP, which is usually implemented on the public Internet. However, the basic method that PTP uses to transfer time is similar to NTP. A reference clock sends a sync message to a secondary clock which includes a time code known as T_1 . The secondary clock records the sync message arrival time, or T_2 , and sends a delay request message back to the reference clock at a time recorded as T_3 . The reference clock receives the delay request message and records its arrival time as T_4 , then sends a delay response message back to the secondary clock that includes T_4 . When this transaction is complete, the secondary clock has access to all four times; T_1 , T_2 , T_3 and T_4 , and computes the one-way path delay, or d , as $(T_2 - T_1) + (T_4 - T_3) / 2$. The time difference between the reference and secondary clock is $T_2 - T_1 - d$ [40].

The accuracy of the loop-back method is always limited by the asymmetry of the network. For example, if the network was symmetric, then the path delay would be the same in both directions, meaning that $d_{ab} = d_{ba}$. If this were true, then the “divide by two” method practiced by ACTS, NTP, and PTP would provide an ideal estimate of the one-way path delay. In practice, however, networks are asymmetric, and d_{ab} and d_{ba} are not equal. Therefore, estimating the one-way path delay as $(d_{ab} + d_{ba}) / 2$ always adds some uncertainty to the received time. In some cases, especially when time signals are sent over a wide area network (WAN) such as the public Internet, d_{ab} and d_{ba} may be very different, because the outgoing and incoming signals may be routed over completely different paths, or because network traffic on one path introduces additional delays. Because the potential for large uncertainties increases when the round-trip delay increases, the loop-back method typically works best when the round-trip delay is small, for example, when implemented over a LAN. The maximum amount of uncertainty that could occur in a loop-back system is 50% of the round-trip delay. This, of course, could only occur in a hypothetical situation where 100% of the path delay was in one direction. In practice, the uncertainty usually is not more than a few percent of the round trip delay.

The loop-back method is sometimes confused with the two-way time transfer method. While it’s true that it involves two-way communication, the reference and remote clocks respond to requests from each other, and send messages to each other at different times, over what can be very different paths. Therefore, the loop-back method is highly susceptible to network asymmetry. A true two-way time transfer system requires the clocks at points A and B to simultaneously send and receive time signals across the same path and thus is much less affected by asymmetry. Two-way time transfer is discussed in Section III.D.

III.C Common-View Time Transfer

The one-way and loop-back time transfer methods generally send both an OTM and a time code to the remote clock so that it can be synchronized to agree with the reference clock. While the common-view method can be used to synchronize the OTM of a remote clock, for example to synchronize its 1 pps output to a reference clock, it does not deliver a time code. Its primary purpose is to compare clocks at two or more locations. It does so by simultaneously measuring the time difference between each clock involved in the comparison and a common-view signal (CVS), which is typically provided by a GPS satellite.

Common-view GPS measurements were first demonstrated at NIST, then called the National Bureau of Standards, in 1980 [41] and soon became the most widely used time transfer method for long-distance comparisons of atomic clocks [42]. There are many variations of the common-view GPS technique, some of them that use the pseudo random noise (PRN) codes broadcast as the CVS, and others that obtain the CVS from the GPS carrier frequency. In addition, the all-in-view technique is often practiced with GPS [43]. This simply means that the CVS is obtained by averaging data from every satellite received at each clock site, rather than from just one satellite. The set of satellites received at each site can be different, as it is not necessary to have any satellites that are in “common-view.” This method allows clocks to be compared to each other anywhere on Earth and works well because the time signals from all the GPS satellites closely agree with each other.

Figure 13 shows a common-view time transfer system where a single satellite serves as the CVS source and a reference clock is compared to a remote clock. The CVS is simultaneously received at sites A and B . Both sites have a local clock and a receiver that each produce a 1 pps signal, and these signals are connected to a time interval counter (TIC) for comparison. The measurement at site A compares the CVS signal received over the path d_{sa} to the reference clock, producing the time difference $Clock A - CVS$. The measurement at site B compares the CVS signal received over the path d_{sb} to the local clock and produces the time difference $Clock B - CVS$. The two measurements are then either exchanged or sent to a common place, where they can be subtracted from each other. The difference between the two measurements is the time difference between the two clocks as the time from the CVS falls out of the equation. Delays common to both paths

d_{sa} and d_{sb} cancel even if they are unknown, but delays that aren't common to both paths contribute measurement uncertainty, resulting in an error term of $d_{sa} - d_{sb}$, which represents the relative, or differential delay, between the two common-view systems. Thus, the basic equation for common-view measurements is

$$Clock_A - Clock_B = (Clock_A - CVS) - (Clock_B - CVS) + (d_{sa} - d_{sb}). \quad (1)$$

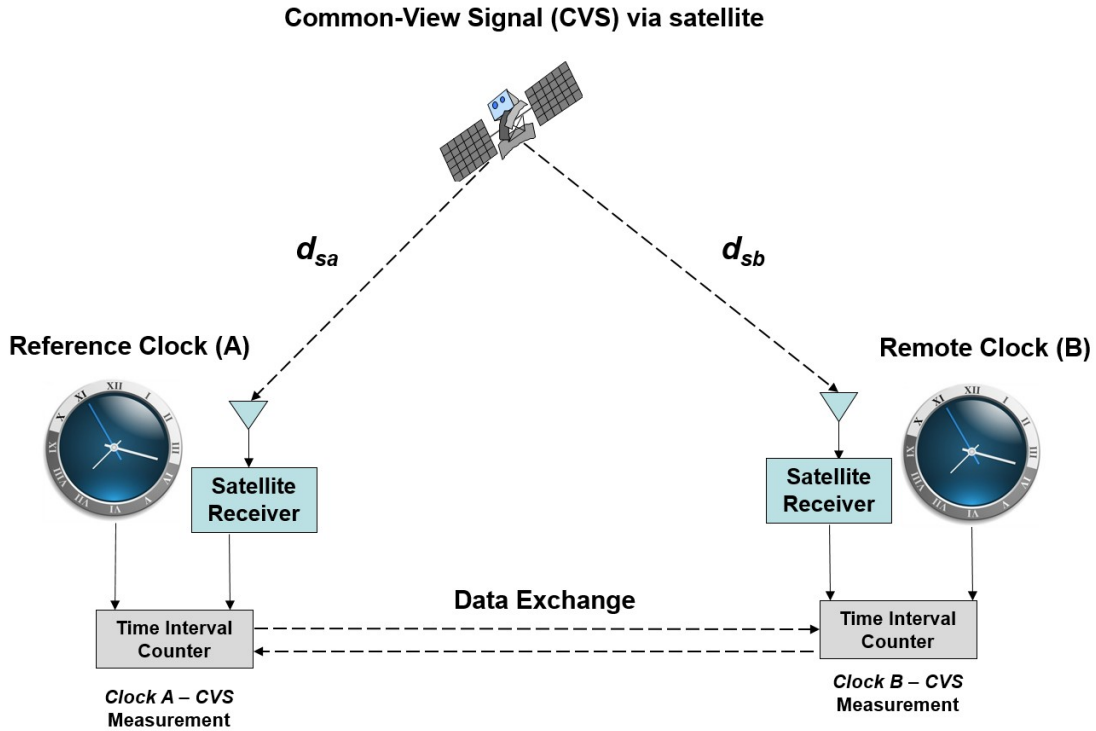


Fig. 13. A common-view time transfer system.

The delays that make up the $d_{sa} - d_{sb}$ error term can be measured or estimated and applied as a correction to the measurement. The delays include not only delays between the CVS and the receiving antennas, but also delays that take place after the signal is received. For example, a system like GPS compensates for nearly all the path delay between the satellite and the receiver on Earth, but some delays, typically measured in nanoseconds, still need to be measured or estimated. To get the best results, a common-view GPS system needs to account for delays added as the signal passes through the ionosphere and troposphere, for delays caused by multipath signal reflections, and for delays introduced by antenna coordinate errors. After the signal reaches the antenna, delays are introduced by the receiver, antenna, and antenna cable, and these delays must also be measured and compensated for to get the best results. The goal is to reduce $d_{sa} - d_{sb}$ to as close to zero as possible, and some common-view systems routinely transfer time with uncertainties of less than 10 ns.

Common-view is a passive, receive-only method. The time signals travel in just one direction, and the receiver in a common-view system, unlike in a loop-back or two-way system, does not exchange messages with the transmitter. They do, however, need to return their measurement data, as the data from all clocks participating in a common-view comparison must be collected and processed. For this reason, common-view systems sometimes cannot report results until long after the measurements are taken, and it is often

used as just a comparison method and not to control clocks. However, systems that exchange common-view measurements in real-time can discipline clocks in much the same way as a GPSDC (Section II.F) that utilizes one-way time transfer [44].

III.D Two-Way Time Transfer

The two-way time transfer method can potentially outperform all other time transfer methods because it can measure and compensate for path delay with very little uncertainty. It requires the two clocks being compared to each other to each transmit their own time signal, and to each receive the time signal sent by the other clock. The time signals simultaneously travel across the same path through the same medium, although different communication channels may need to be used to prevent the two signals from interfering with each other.

Like the loop-back method, the two-way method, due to expense and complexity, is more practical to use in a wired medium, such as a computer network, than it is with a wireless medium. However, wireless two-way time transfer via satellites is routinely used by NIST, the USNO, and timing laboratories in other countries to compare clocks located on different continents and to contribute data to the calculation of UTC. We briefly describe those satellite systems here to illustrate how the method works.

The clocks being compared are located near a satellite Earth station that typically contains a spread spectrum satellite modem, a dish antenna, a TIC, and radio transmitting and receiving equipment (Fig. 14). Both Earth stations (A and B) then simultaneously transmit time signals through a transponder on the same geostationary satellite. The transponder serves as a repeater, receiving the time signal from A and retransmitting it so it can be received by B, and vice versa. Each station then measures the time difference between two 1 pps signals, one generated by its local clock and the other received via satellite from the remote clock. Station A records $TIC_A = CLK_A - (CLK_B + d_{ba})$, where CLK_A is the time from the local clock, CLK_B is the time from the remote clock, and d_{ba} is the path delay from B to A, which includes the delays introduced by the transmitting modem, the satellite uplink, the satellite transponder, the satellite downlink, and the receiving modem. Station B records $TIC_B = CLK_B - (CLK_A + d_{ab})$, where d_{ab} is the path delay from A to B [45]. The two stations then exchange their measurements. The time difference between clocks A and B is calculated as

$$Clock_A - Clock_B = \frac{TIC_A - TIC_B}{2} - \frac{d_{ba} - d_{ab}}{2}. \quad (2)$$

If the two paths had identical delays, or if $d_{ab} = d_{ba}$, then only the part of the equation to the left of the minus sign would be necessary, and the time difference between the clocks would simply be $(TIC_A - TIC_B) / 2$. The part of the equation to the right of the minus sign, $(d_{ba} - d_{ab}) / 2$, reflects the differences in the two path delays that contribute uncertainty to the measurement, assuming that they have not already been applied as a correction to the measurement. Unlike the round-trip method, the path is reciprocal rather than asymmetric, but small differences in the two path delays can still be introduced by delays in the transmit and receive hardware that are different at the two sites, or if the signals in the two directions are transmitted on different frequencies [45, 46]. Even so, the uncertainty of the two-way method via geostationary satellites can be reduced to about 1 ns [47].

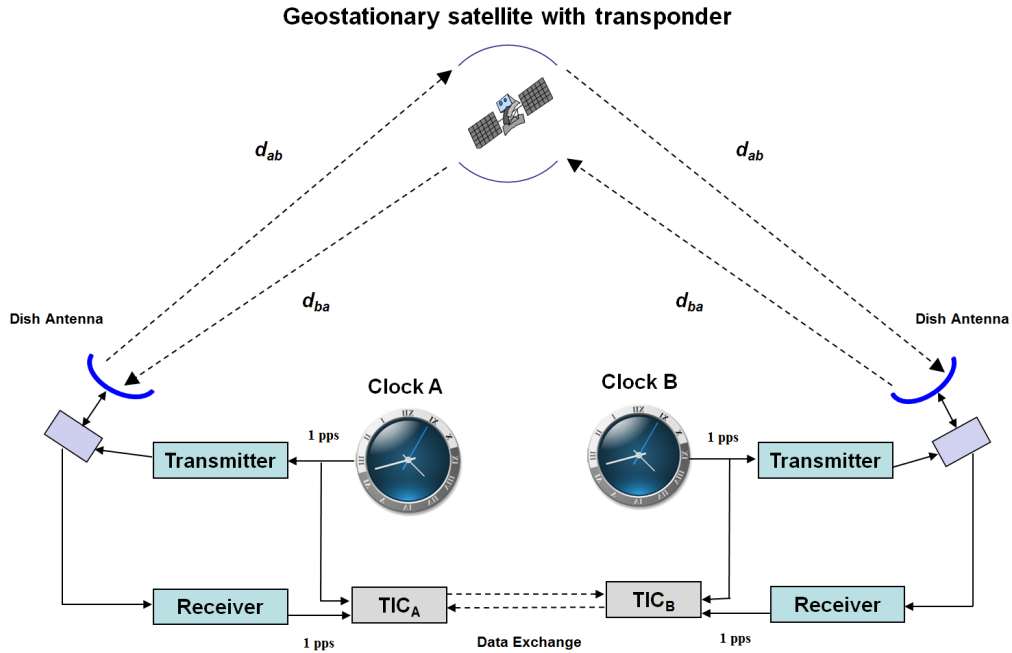


Fig. 14. A two-way time transfer system that utilizes a geostationary satellite.

In Section III.B, we noted that the loop-back method is sometimes confused with two-way method. However, the remote and reference clocks in a loop-back system exchange time signals at different times, a half-duplex method of communication. As illustrated by Fig. 14, two-way time transfer systems are full-duplex systems that require the simultaneous transmission of time signals from both clocks. As a result, they are less susceptible to fluctuations in the path delay.

If the clocks being synchronized are not too far apart, a two-way time transfer system via a wired network can be utilized to compare clocks located throughout a building or campus, or even across distances of hundreds of kilometers. A time transfer protocol now receiving widespread attention is White Rabbit [48], which combined two existing technologies, PTP [39] for time synchronization and Synchronous Ethernet (SyncE) [49] for frequency syntonization; and measures small phase changes between clocks with dual mixer time difference (DMTD) systems. The White Rabbit protocol can be implemented in a true two-way mode by having clocks send and receive time signals at different wavelengths through a single bidirectional optical fiber [50, 51]. Two-way time transfer systems implemented via wired networks, using White Rabbit, a variant, or a custom experimental protocol, have recently demonstrated smaller uncertainties than satellite-based time transfer systems. Time transfer uncertainties near or less than 100 ps (0.1 ns) have been demonstrated over relatively short distances of hundreds of meters via coaxial cables [52] and for distances of up to hundreds of kilometers via bidirectional optical fibers [50, 53, 54].

III.E Dependencies of Other Time Distribution Systems on GPS

Because GPS is such a tremendous resource, with signals that are free, accurate, and easy to receive anywhere on Earth; it is not surprising that numerous other time distribution systems are designed to use GPS as their reference clock. Both restricted access and public access time distribution systems often rely on GPS. The upside is that GPS is the enabling technology that made these systems possible. The downside is that the loss of GPS timing signals can cause these systems to fail. This section looks at GPS dependencies in both restricted and public access systems. The public access discussion is limited to systems that are available now and that are controlled by United States interests.

III.E.1 Restricted Access Time Distribution Systems

Restricted access timing systems can be classified into two broad categories. The first category includes timing systems designed for in-house usage by the occupants of an organization or facility. These systems are typically acquired by buying and installing commercially-available hardware and software. After the initial acquisition costs have been paid, these systems are free to use, except for periodic maintenance. The second category includes timing systems that are accessed by subscription only, where users pay to receive the time signal.

The first category is dominated by systems that provide time to private LANs via computer time protocols including NTP, PTP, and White Rabbit, but also includes, for example, systems that synchronize time displays (often with one of the Inter-Range Instrumentation Group (IRIG) formats [55]), or systems that require a Society of Motion Picture and Television Engineers (SMPTE) time code [56] to label frames of video or film. It is important to remember that none of these systems (NTP, PTP, White Rabbit, IRIG, and SMPTE) are reference clocks. They are simply standardized protocols for transferring time, and the time itself must originate from a reference clock. In practice, the reference clock is nearly always a GPSDC.

The level of GPS dependency and the potential impact of GPS failures on restricted access timing systems is high. The GPSDCs that control these time distribution systems are often embedded inside the equipment chassis. For example, they might be inside the chassis of a commercial PTP server, or they may be standalone clocks that interface to the distribution system. In either case, a GPSDC failure will eventually cause the distribution system to fail. There may be a domino effect that affects users who did not even know they were dependent upon GPS time. For example, a SMPTE system may synchronize to an NTP server located in another city or state that is in turn synchronized by a GPSDC. In this case a GPSDC failure would first cause the NTP time to be wrong which would in turn cause the SMPTE time code to be wrong. The domino effect is potentially very serious at sites where the rooftop installation of antennas is difficult, such as inside an office complex or data center located in a metropolitan area. Consider, for example, a hypothetical but not uncommon situation where just one antenna was installed to support a single GPSDC, which then synchronizes 10 NTP or PTP time servers inside the building, each of which in turn provides synchronization to 1000 client computers.

The second category of restricted access systems, those accessible by subscription only, usually do not rely on time from GPS, mainly because there would be no point in trying to charge for something that is freely available. These services reach a relatively small number of users. NIST, for example, distributes UTC(NIST) as generated in Boulder, Colorado to more than 50 customers through services that distribute frequency and time via the common-view method [57] described in Section III.C. These services are not affected if GPS time is wrong, because they do not distribute GPS time. They have a GPS dependency because the satellite signals are needed as a relay to deliver UTC(NIST) to customers. However, most of the customers that subscribe to these services have clocks with excellent holdover capability and are unaffected by short common-view outages. NIST began offering a time over fiber service in 2019 that originates from both the primary UTC(NIST) time scale in Boulder, Colorado [58] and from a secondary time scale in Gaithersburg, Maryland [59]. The Boulder service does not have a GPS dependency, but the secondary time scale in Gaithersburg, Maryland has a partial dependency because it is periodically synchronized with the primary time scale via GPS common-view, but again has excellent holdover capability.

III.E.2 Public Access Time Distribution Systems

Due in part to the success of GPS, which has at least indirectly led to the demise of eLoran and other systems, only a small number of free public access time distribution systems remain that are under U. S. control. All but one of these systems have at least one caveat when considered for critical infrastructure

usage, they are either not independent of GPS, not capable of microsecond-level accuracy, or both. A description of each system follows, with a summary provided in Table 4.

CDMA – CDMA clocks receive time codes from Code Division Multiple Access (CDMA) signals transmitted from mobile base stations. These time codes are synchronized to GPS and typically have sub-microsecond accuracy at their source, with some path delay added before they reach a CDMA clocks. However, because many thousands of base stations exist, it is highly likely that any CDMA clock located in a populated area will be within a few kilometers of a base station, and thus synchronized to within less than 10 μ s of UTC. CDMA clocks also have the advantage over GPS of being able to work indoors without an outdoor antenna, typically working anywhere a mobile phone will work, which makes them well suited for use in large metropolitan buildings or data centers [60]. However, CDMA clocks are completely dependent on GPS, and are even sometimes referred to as “indirect GPS”, because the base stations essentially function as GPS repeaters. More importantly, the two providers of CDMA services in the United States, Sprint and Verizon, are both expected to discontinue CDMA service, perhaps as soon as 2022 [61].

Public NTP Servers – A large number of public access NTP servers exist, both in the United States and other countries. The NTP Pool Project includes over 4000 servers as of January 2020, with more than 900 located in North America [62]. The percentage of these servers that are referenced to GPS is not known, but it is believed to be very high. For example, the Network Time Foundation database does allow searching for servers by synchronization method [63], and a query in January 2020 showed that of the 191 active servers listed, 185 (96.9%) were synchronized to GPS. The remaining six were synchronized to CDMA.

The largest block of public access NTP servers on the Internet are, however, synchronized by UTC(NIST) and independent of GPS. They comprise the Internet Time Service (ITS), operated by NIST, which handles billions of timing requests per day [64]. As of January 2020, the ITS currently consists of more than 20 servers that are located at NIST facilities in Colorado and Maryland [65]. Due to the problems of network asymmetry discussed in Section III.B, the accuracy obtained from these servers typically ranges from about 0.5 to 10 ms.

NIST Radio Stations (WWVB, WWV, and WWVH) – NIST radio stations WWV and WWVB, located near Fort Collins, Colorado, and WWVH, located on the island of Kauai in Hawaii, are the only radio stations located in the United States whose primary purpose is to distribute standard time and frequency signals. The stations are listed in the United States code of federal regulations as references sources for measurements of time and frequency [66]. WWV and WWVH both operate in the high frequency (HF) part of the radio spectrum, better known as shortwave. Both stations transmit on 2.5, 5, 10, and 15 MHz, with WWV offering two additional broadcasts at 20 and 25 MHz. WWVB operates in the low frequency (LF) part of the radio spectrum at 60 kHz and is well known as the synchronization source for low-cost consumer grade radio controlled clocks; tens of millions of these clocks have been sold in the United States and are commonly found in homes, schools, and offices.

The NIST stations are independent of GPS and multiple cesium clocks are kept at both sites that are kept in close agreement (< 20 ns) of UTC(NIST). However, the synchronization capability of the NIST stations is limited because they operate in one-way mode with no compensation for path delay. Thus, their signals are delayed as a function of distance from the transmitter, which can be delayed by as much as 15 ms at some locations in the continental United States, and which usually precludes them from critical infrastructure usage. The stable LF groundwave signals of WWVB, however, potentially make it useful as a frequency syntonization source with an accuracy of parts in 10^{12} [67], although currently no products exist that utilize WWVB for that purpose.

STL – The Satellite Time and Location (STL) service is operated by Satelles (www.satellesinc.com), a company headquartered in Reston, Virginia. Commercial STL clock products are now available to the

public, but because the technology is licensed and patented by Satelles, they can only be acquired through their authorized business partners. The STL signals are transmitted at frequencies ranging from 1616 MHz to 1626 MHz from 66 Low Earth Orbit (LEO) satellites that form the Iridium constellation. Iridium was originally designed for use with mobile phones, and like CDMA, STL does not require a rooftop antenna. The altitude of the satellites is just 781 km, as opposed to the 20 200 km altitude of the GPS satellites, and the signals as received on Earth are about 300× to 2400× stronger than GPS, making them usable indoors [68]. The timing accuracy specification for STL is ±500 ns (0.5 μs) which meets critical infrastructure requirements, and published measurements indicate an accuracy of less than 200 ns when compared to a GPSDC [69].

Table 4. A summary of non-GPS public access time distribution systems under U. S. control.

Public Access Time Distribution System	Capable of Microsecond Accuracy?	Independent of GPS?
CDMA	YES	NO
Public NTP Servers (non-NIST)	NO	NO
NIST Internet Time Service	NO	YES
WWV and WWVH	NO	YES
WWVB	NO	YES
STL	YES	YES

IV. Timing Requirements and Dependencies of Stock Exchanges

Stock exchanges are facilities where brokers and traders can buy and sell financial assets, such as shares in stocks and bonds. The two largest stock exchanges in the United States (and in the world) are the New York Stock Exchange (NYSE) and the NASDAQ, which was originally an acronym for the National Association of Securities Dealers Automated Quotations. Numerous smaller stock exchanges are also in operation - some provide their own listings while others handle a percentage of the volume of traded shares for companies listed on the NYSE and NASDAQ.

Both the NYSE and NASDAQ have headquarters in New York City. Shares of about 4000 companies are listed and available for public trading on the NYSE and NASDAQ, with the NYSE more likely to list the largest companies. Therefore, the combined market capitalization (the share price multiplied by the number of shares) of the NYSE (> \$20 trillion) is typically about double that of the NASDAQ (> \$10 trillion). Billions of shares are exchanged during each trading day, and thus stock exchange activities have a tremendous impact on both the U. S. and global economies. The following sections examine why time synchronization is important to stock exchanges, what the synchronization requirements are, and the dependency of stock exchanges on GPS.

IV.A The Importance of Time Synchronization to Stock Exchanges

The importance of high speed, low latency optical fiber networks to financial trading firms has been well chronicled, perhaps most notably in the book *Flash Boys* [70]. These networks can potentially give financial firms an “edge” by allowing them to transfer information to and from stock exchanges, and buy and sell stocks, more quickly than their competitors. This is a timing related activity, with the goal being to reduce the path delay as much as possible, but it is not illegal and thus not a concern for the agencies that regulate stock markets. Instead, regulators are concerned with the fraudulent activity and market manipulation that occurs when the stock trades are not processed by the exchanges in the order they are received. To prevent this from happening, each part of a stock market transaction must be recorded and time stamped. To keep the time stamps accurate, every clock involved in a stock market transaction should be synchronized to agree with a common reference clock that keeps accurate and internationally traceable time. In short, operating a fair and equitable stock exchange requires accurate time synchronization.

To better illustrate the importance of accurate time stamps to a fair and equitable stock exchange, it is worthwhile to describe how stocks are traded. In a stock transaction, the bid is the highest price that a buyer is currently offering to buy one share of stock. For example, if a buyer has submitted an order for 100 shares of ACME Corporation stock and is offering to pay \$10 per share, they are submitting a bid of \$10. The ask is the lowest price that a seller is currently willing to accept for one share of stock. In the above example, if the ask on ACME Corporation stock is \$10.05, the buyer’s bid won’t be accepted unless it is increased to match the ask. A stock trade only happens when the buyer and seller agree on a price. The general rule known to stock traders is simply — “buy at the ask and sell at the bid.” Of course, the bid and ask prices are always changing, based on the current level of demand between buyers and sellers.

The difference between the bid and ask is commonly known as the spread. The spread is the profit received by the organization that facilitates the trade, commonly known as the market maker. When a trader sells, a market maker buys, and when a trader buys, a market maker sells. Thus, market makers reverse what stock traders do; they buy at the bid and sell at the ask. If the market maker buys ACME Corporation stock for \$10 per share and sells it for \$10.05, their profit is \$0.05 per share [71]. Market makers are always motivated to adjust the spread in a way that maximizes profits, by either keeping the spread large and accepting less trading volume, or by reducing the spread to increase the trading volume.

Stocks were traded in the United States for about 200 years, beginning in the 1790s, before it finally became necessary to implement and regulate time synchronization requirements. The main reason that time synchronization finally became a concern was the widespread usage of electronic trading platforms and automated stock exchanges that began in the late 1990s. Electronic trading fundamentally changed stock exchanges; soon most market makers were no longer individuals working the telephones or physically waiting in line to place orders. Instead, computers began to automatically execute trades, and to buy and sell stocks to each other, based on software algorithms. These automated trading platforms greatly reduced the amount of time needed to execute a transaction making the markets move faster. In addition, automation led to smaller spreads, resulting in the decimalization of stock prices which began on the NYSE and NASDAQ in 2001. Prior to decimalization, stock prices were listed in fractional dollars, which kept the spread large; for example, if a stock traded in price increments of 1/8 of a dollar the spread would be at least 12.5 cents, but decimalization reduced the minimum price increment to one cent [72]. Smaller spreads made the market more liquid but created the incentive for market makers, now dealing with reduced profits from tighter spreads, to execute more trades [73].

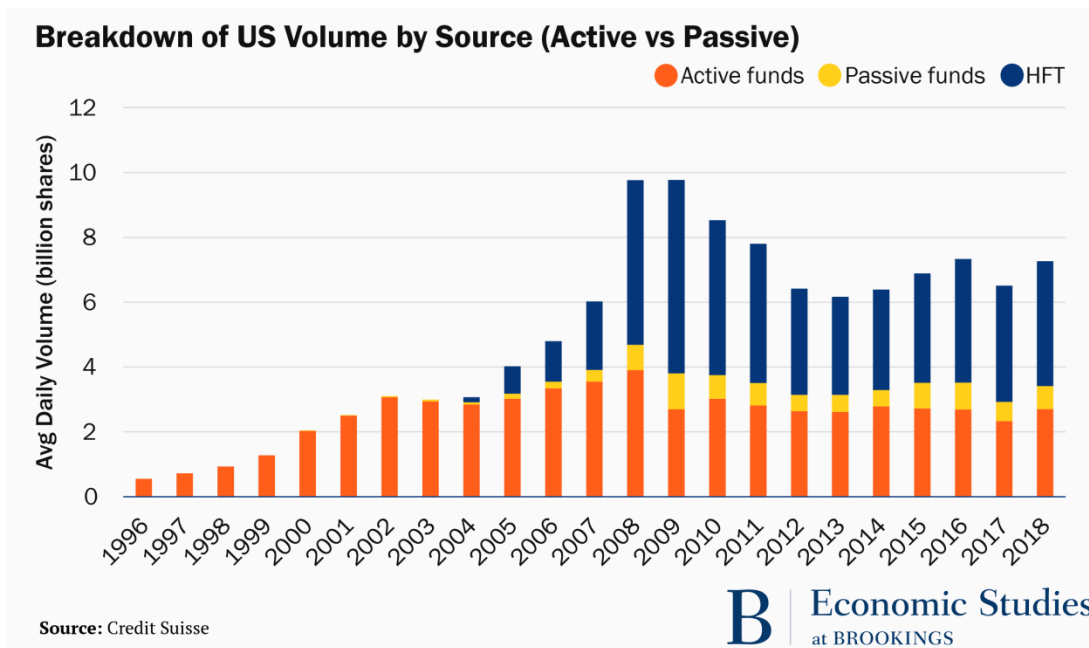


Fig. 15. U. S. stock volume from 1996 to 2018, and the contribution of high frequency trading (HFT).

Each of these factors led to the now widespread practice of high frequency trading (HFT) where a much larger number of transactions could occur in a given time period than before, because the trades executed by automated trading platforms can execute in intervals measured in microseconds [74]. At least half of the transactions in today's stock markets are the result of HFT. Figure 15 [75] shows the average trading volume, or the number of shares of stock traded per day, for U. S. stock exchanges from 1996 to 2018. Note that less than one billion shares per day were traded before electronic trading platforms began operation in the late 1990s. This number increased to nearly 10 billion shares per day in 2008 and 2009. Note also that HFT did not have a significant impact on trading volume until 2004, but since 2008 has accounted for at least half of the total volume. Finally, it is interesting to note that stock market volume has leveled off in recent years and remained about 6 billion and 7.5 billion shares per day each year from 2012 to 2018. Even so that is more than 10× the daily volume of the markets before the use of electronic trading platforms.

The practice of HFT made it essential for all stock exchanges and trading platforms to be able to document that their time stamps are accurate, to avoid stock market fraud and manipulation. To illustrate the importance of accurate time stamps, consider a hypothetical situation where a retail investor, trading from a home computer, is preparing to submit an order for 100 shares of ACME Corporation stock. Before submitting the order, the investor checks the display of their online trading software and notes that the current bid is \$9.23, and the current ask is \$9.25. The display also indicates that no one else is in line to buy the stock. Thus, they understandably assume that the ask price, \$9.25 per share, is the price they will pay and “click” to submit the order. A fraction of a second later a large investor (such as a hedge fund or investment bank), submits an order to buy 1,000,000 shares of ACME. This large order arrives at the stock exchange after the small order, but the exchange, for reasons unknown to the retail investor, fraudulently elects to execute the large trade first. The large trade immediately raises both the bid and the ask price by 15 cents, to \$9.38 and \$9.40, respectively. This action causes the retail investor to instantly lose money, regardless of whether they entered a market or limit order. If they entered a market order, meaning that they agreed to accept the current ask price regardless of what it is, they’ll pay \$9.40 for the stock they were expecting to buy at \$9.25. If they entered a limit order, indicating that \$9.25 was the most they will pay, they missed out on the trade completely because the large investor was allowed to buy their shares and their order was never filled. Meanwhile, the large investor who “cut ahead in line” has the option of selling at least some of the shares they just purchased at the new bid price, taking an instant profit [71].

Cutting in line is called “front running” and is illegal, but when computers are making the trades, it is essentially impossible to prove or regulate without accurate time stamps. It only becomes possible if distributed trading platforms rely entirely on clocks that are synchronized to a common reference clock. In addition, the resolution of the time stamps used to record each transaction must be finer than the execution times of the trade. For example, if trades are being stamped with a resolution of one second, it means that the time stamp can display 12:22:01 (12 h, 22 min, and 1 s) and 12:22:02, but nothing in between. Thus, a time stamp of 12:22:01.5 cannot be recorded. This is unacceptable with HFT, because a stock exchange can potentially execute many thousands of trades within a one second interval and each of those trades would have the same time stamp. This makes it possible for the market maker to reorder them in any way that suits their purposes, because there will be no way for an auditor reviewing the transaction logs to prove which orders arrived first [71]. Concern by regulators in the 1990s that trades were not always being executed in the best interest of customers led to the establishment of order audit trail systems (OATS) and the first rules for the synchronization of stock market clocks, as described in the next section.

IV.B Time Synchronization Requirements of Stock Exchanges

Prior to the advent of distributed electronic trading platforms and HFT, many of the clocks maintained by stock exchanges were mechanical devices that physically stamped the time, in ink, onto the paper documents used to record transactions. These clocks were not always synchronized and seldom had the ability to display seconds. This changed in August 1996, when the U.S. Securities and Exchange Commission (SEC) issued a report [76] that included findings from an investigation of the practices of the National Association of Securities Dealers (NASD) and the NASDAQ. The report alleged that the NASD and NASDAQ did not always act in the best interest of customers — some trades were improperly executed, and some collusion existed among market makers. A financial settlement was reached between the SEC and NASD that also included new regulations; the NASD was required to improve market surveillance and to develop an enhanced Order Audit Trail System (OATS). To new OATS rules went into effect in August 1998. One of the rules, rule number 6953, was entitled “Synchronization of Member Business Clocks” [77] and represented the first stock exchange synchronization requirement. It required computer systems and mechanical clocks to be synchronized every business day before the stock market opened to ensure that time stamps were accurate. The synchronization requirements were as follows:

All computer system clocks and mechanical time stamping devices must be synchronized to within three seconds of the National Institute of Standards and Technology (NIST) atomic clock. Any time provider may be used for synchronization, however, all clocks and time stamping devices must remain accurate within a three-second tolerance of the NIST clock. This tolerance includes all of the following:

- *The difference between the NIST standard and a time provider's clock;*
- *Transmission delay from the source; and*
- *The amount of drift of the member firm's clock.*

For example, if the time provider's clock is accurate to within one second of the NIST standard, the maximum allowable drift for any computer system or mechanical clock is two seconds [77].

The three second synchronization requirement was not stringent, but it was significant for two reasons: it finally required stock exchanges to synchronize their clocks, and it established NIST as the reference clock for U. S. stock exchanges. It was superseded by the Financial Industry Regulatory Authority (FINRA), which issued new requirements for stock market synchronization in 2008 for the NASDAQ and in 2011 for the NYSE. These requirements are contained in FINRA OATS Rule 7430 which had the same title as its predecessor, "Synchronization of Member Business Clocks." Like the previous requirements, the new requirements list NIST time as the official reference for stock market transactions. However, the new rule reduced the synchronization requirement by a factor of three, from within three seconds to within one second of NIST time. It also required one second synchronization to be maintained at all times when the markets were open, and for all clocks to have one second granularity, which was equivalent to one second resolution, and for firms to maintain copies of their clock synchronization procedures on-site [78].

Even though not all trading firms were in favor of more stringent synchronization requirements, the regulatory agencies were soon aware that reducing the synchronization requirement from three seconds to one second, while a step in the right direction, was too coarse a requirement to prevent HFT fraud and market manipulation. Thus, FINRA Regulatory Notice 14-47 [79] was sent out for comments in November 2014 and proposed tightening the synchronization requirement by a factor of 20, to within 0.05 s (50 ms) of NIST time. This led to the adoption of a new FINRA Rule 6820, now simply titled "Clock Synchronization", that superseded all previous rules [80]. Section a) of Rule 6820 reads as follows:

(a) Clock Synchronization

(1) Each Industry Member shall synchronize its Business Clocks, other than such Business Clocks used solely for Manual Order Events or used solely for the time of allocation on Allocation Reports, at a minimum to within a fifty (50) millisecond tolerance of the time maintained by the atomic clock of the National Institute of Standards and Technology ("NIST"), and maintain such synchronization.

(2) Each Industry Member shall synchronize (A) its Business Clocks used solely for Manual Order Events and (B) its Business Clocks used solely for the time of allocation on Allocation Reports at a minimum to within a one second tolerance of the time maintained by the NIST atomic clock, and maintain such synchronization.

(3) The tolerance for paragraph (a)(1) and (2) of this Rule includes all of the following:

- (A) The difference between the NIST atomic clock and the Industry Member's Business Clock;*
- (B) The transmission delay from the source; and*
- (C) The amount of drift of the Industry Member's Business Clock.*

(4) Business Clocks must be synchronized every business day before market open to ensure that timestamps for Reportable Events are accurate. To maintain clock synchronization, Business

Clocks must be checked against the NIST atomic clock and re-synchronized, as necessary, throughout the day [80].

In 2012, the SEC, through Rule 613, voted to require FINRA and U. S. stock exchanges to establish and conform to a consolidated audit trail (CAT) that would enable regulators to be able to monitor and analyze trading activity. The CAT would require FINRA and the stock exchanges to collect and accurately identify every order for all stocks and stock options across all U.S. markets, and to send complete documentation about the order to a central repository by 8 a.m. Eastern Time the day following the trade. Several subsequent revisions of the CAT plan followed, leading to the current version, published in November 2016. The synchronization requirements are equivalent to FINRA Rule 6820, 50 ms for automated orders and 1 s for manual orders, with respect to NIST. However, the CAT plan also includes a requirement for granularity (resolution), which must be at least one millisecond, and if the time stamps have resolution finer than 1 millisecond, time stamps that include all available digits should be recorded and sent to the central repository, so that all reportable events can be “adequately sequenced.” The millisecond resolution time stamps are required at five places in the audit trail; the time of order origination, the time when the order is routed, the time when the order is received, the time when the order was modified or cancelled, and the time when the order was executed [81].

Since January 3, 2018, the synchronization requirement for HFT in the European Union has been 100 μ s [82], or 500 \times more stringent than the U. S. requirement for accuracy, with a 1 μ s requirement for granularity (resolution), which is 1000 \times more stringent than the CAT requirement. Manual orders have a 1 s requirement, the same as in the U. S., but all other trades, meaning automated trades that do not qualify as HFT, have a 1 ms requirement. These requirements are commonly known as MiFID II, an acronym for the Market in Financial Instruments Directive (MiFID) where it originated. The reference clock for European stock exchanges can be any time scale, including NIST, that contributes to Coordinated Universal Time (UTC).

Some parties believe that equivalent, or even tighter, requirements will soon be adopted in the U. S., because 100 μ s synchronization is already being maintained by a high percentage of stock exchanges and trading firms and utilized in their internal operations [81]. However, it should be noted that the accuracy requirements are applicable to all clocks involved in a transaction, including the clocks in server and client computers. This is important, because whereas accuracy to within tens of nanoseconds is obtainable, for example, with a GPSDC installed at a stock market data center; the delay asymmetries found in computer networks, in networking hardware, and even in application software and operating systems, typically limit the accuracy of computer clocks to tens of microseconds, or 1000 \times worse. Thus, for the computer clocks that time stamp transactions, the 100 μ s accuracy requirement can be difficult for trading firms to achieve and perhaps even more difficult for auditors to verify.

Table 5 provides a summary of current stock market synchronization requirements. For the sake of completeness, and because of the international nature of financial markets, the European requirements, as well as the requirements of two other countries with known regulations, Australia [83] and Canada [84], have been included. Both Australia and Canada have requirements that are similar to those in the U. S.

Table 5. Summary of stock market synchronization requirements.

Region	Reference Clock	Time accuracy and granularity (resolution) requirements		
		Type of Trade	Accuracy	Granularity
US [80, 81]	UTC(NIST)	Automated orders	50 ms	1 ms
		Manual orders	1 s	1 s
EU [82]	Any time scale that contributes to UTC	Manual orders	1 s	1 s
		High frequency trading	100 μ s	1 μ s
		All other trading	1 ms	1 ms
Australia [83]	UTC(AUS)	All trading	20 ms	Not specified
Canada [84]	UTC(NRC), or any time scale that contributes to UTC	Automated orders	50 ms	Not specified
		Manual orders	1 ms	Not specified

IV.C GPS Timing Dependencies of Stock Exchanges

Complying with a 50 ms synchronization requirement with UTC(NIST) is not especially difficult and can be achieved with time signals that originate directly from NIST, including WWVB, ACTS, and ITS. Even so, these systems are seldom used to synchronize automated trading platforms. Instead stock exchange data centers in the U. S. and elsewhere routinely operate equipment, such as NTP and PTP time servers, that are referenced to GPSDCs. Because these products are readily available from numerous manufacturers, they often provide not only the most accurate solution available, but also the most economical, and the benefit of providing a turnkey solution to the problem. In addition, if the U. S. eventually adopts the 100 μ s HFT requirements of Europe, as some expect, or even a less demanding 1 ms requirement, then the use of WWVB, ACTS, and ITS will no longer be acceptable. Thus, installing GPS equipment now in their data centers prepares stock exchanges for the future.

The widespread use of GPS time to synchronize U. S. stock exchanges raises three important questions, two of which are regulatory questions and a third question that is related to critical infrastructure dependency. The first question is, does GPS meet the requirements of synchronization to NIST time? As we have seen, both FINRA Rule 6820 [80] and the SEC CAT Plan [81] specifically call for stock exchange clocks to be synchronized to NIST, and strictly speaking, GPS time is not NIST time, as it originates from UTC(USNO). As indicated in Fig. 8, however, the differences between the two time scales has seldom exceeded 10 ns for several years, which has no impact whatsoever on stock market synchronization requirements, and traceability to UTC(NIST) and UTC can be established through GPS [24]. If current verification is needed to show that NIST time agrees with GPS time, the current time difference can be accessed through the NIST GPS Archive, which is updated every 10 minutes [85]. Thus, a very strong claim can be made that GPS does satisfy the requirement of synchronization to NIST time, but an actual acceptance of the legality of that claim could only be determined for certain in a court of law.

If we accept that synchronization to GPS does satisfy the requirement of synchronization to NIST, a second regulatory question remains - how can we verify, at a given time and in a given place, that the time we are receiving from GPS and distributing through our time servers is accurate enough to meet our regulatory requirements? Even though it is highly likely that it is accurate enough, the only way to know with certainty is through continuous monitoring and verification of every clock. If the monitoring is not done, there may

be no warning signs when synchronization is not present. For example, unlike a telecom network which might indicate the loss of synchronization by failing or losing capacity, an unsynchronized stock exchange can continue to function, even though its order audit trail system has lost its integrity.

A method employed by NIST to monitor and verify that stock exchange clocks are always accurate and traceable involves a NIST disciplined clock (NISTDC) [57, 71] that is deployed to stock market data centers and that keeps a rubidium or cesium atomic clock continuously locked to the UTC(NIST) time scale via common-view GPS (Fig. 13). In addition to an atomic clock (rubidium clocks are integrated but cesium clocks are external to the chassis) and the hardware needed for common-view measurements, the NISTDC chassis includes a computer time server supplying both the network time protocol (NTP) and precision time protocol (PTP), an event timing board used to measure the accuracy of packets sent by stock exchange time servers, and an amplifier that distributes the atomic clock signals to time servers and other stock exchange clocks. The distribution amplifier provides three 10 MHz frequency outputs and eight 1 pulse per second (pps) time outputs that are synchronized to UTC(NIST), as shown in Fig. 16.

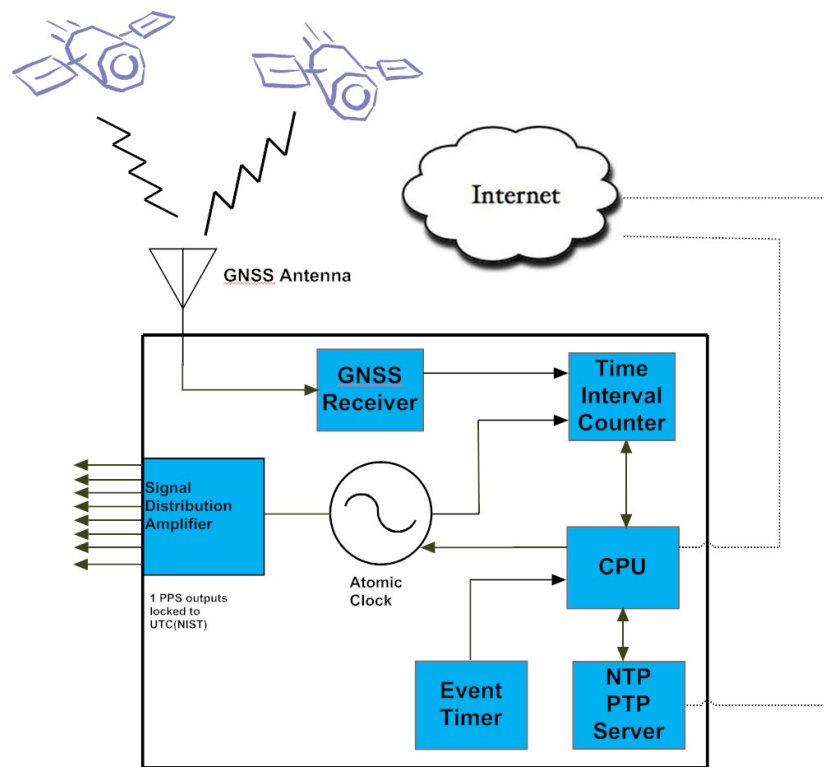


Fig. 16. Block diagram of NIST Disciplined Clock (NISTDC).

Time from the UTC(NIST) time scale in Boulder, Colorado is transferred to the NISTDC by use of the common-view disciplining method [44, 71, 86]. As illustrated in Fig. 17, a system at NIST measures the time difference $UTC(NIST) - GPS$, and measurements performed at the NISTDC site produce $NISTDC - GPS$. Every 10 minutes, each NISTDC and the NIST system simultaneously send their measurement data to an Internet cloud server where the NISTDC data are subtracted from the NIST data. This removes the contribution of GPS time and results in an estimate of the $UTC(NIST) - NISTDC$ time difference. The NISTDC converts this time difference to a frequency correction by use of an adaptive proportional-integral-derivative (PID) controller and then applies the correction to its local atomic clock. The process is continuously repeated to keep the NISTDC locked to UTC(NIST).

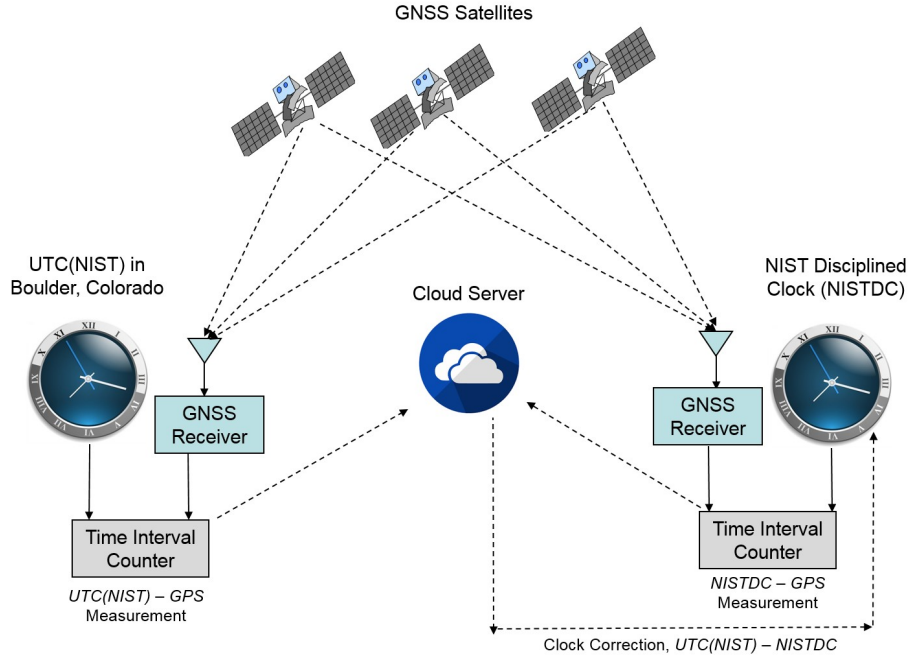


Fig. 17. NISTDC common-view link to UTC(NIST).

Common-view observations routinely show that a locked NISTDC seldom deviates by more than ± 10 ns ($\pm 0.01 \mu\text{s}$) from UTC(NIST) and that its average time offset is near 0. To illustrate this, Fig. 18 shows a 6-month (July to December 2019) comparison of a NISTDC, located at a major US stock exchange, to UTC(NIST). The peak-to-peak variation over the 6-month interval is ~ 25 ns, but most data points fall within ± 5 ns and the average time offset is less than 0.1 ns, or essentially 0.

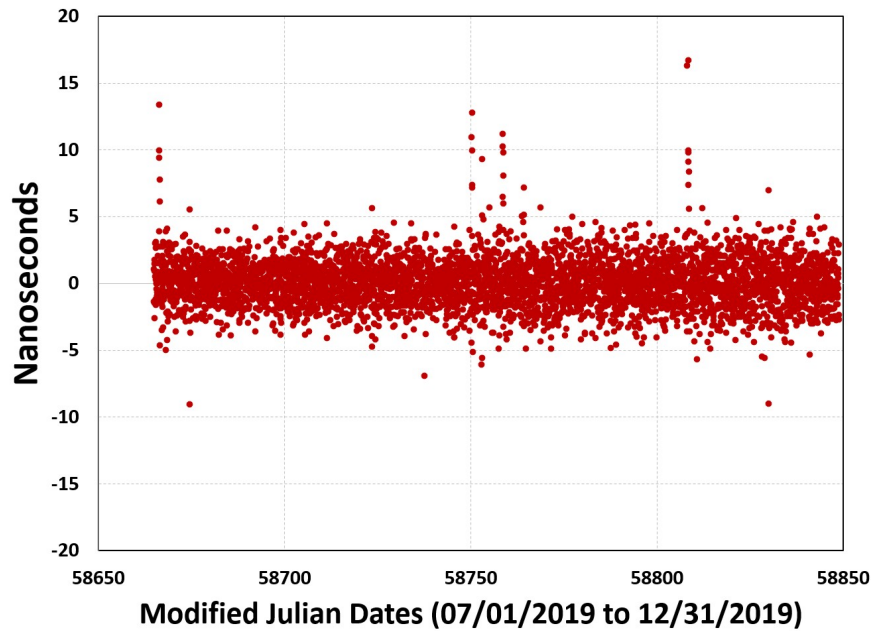


Fig. 18. Accuracy of NISTDC operated by a major U. S. stock exchange, with respect to UTC(NIST).

Time servers located in the same data center as a NISTDC can eliminate the need for connection to a GPS antenna, by using 1 pps signals distributed from the NISTDC via coaxial cable as their synchronization source and by obtaining time-of-day information from the NISTDC’s integrated NTP/PTP server. However, regardless of whether a time server is synchronized with UTC(NIST) or GPS, it can still be monitored by the NISTDC, which can simulate a client computer and send time requests to servers every 10 seconds. It then compares the time stamps in the received packets to the time kept by the NISTDC, a comparison made with 0.1 μ s resolution. When the NISTDC resides on the same local area network (LAN) as the time server, the measured time offset of a properly synchronized server typically ranges from a few microseconds to about 50 μ s when compared to the NISTDC. Figure 19 shows the time offset of an NTP server, operated by a major US stock exchange, when compared to UTC(NIST) during the last 100 days of 2019 (one data point per hour). The average time offset of the NTP server clock is 20 μ s and its time deviation (stability) is < 1 μ s at an averaging period of 1 day.

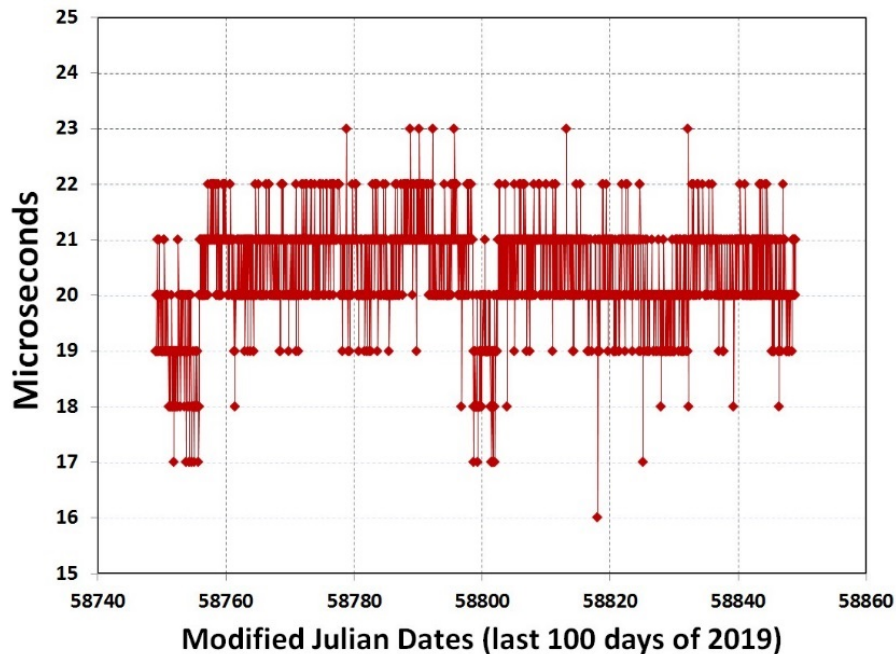


Fig. 19. Accuracy of an NTP time server at a major US stock exchange, with respect to UTC(NIST).

NISTDC units synchronize some of the world’s largest stock exchanges and are installed at data centers near New York City and Chicago in the U. S., as well as in London and Frankfurt in the European Union, and in Tokyo, Japan. Stock exchange clients who utilize the NISTDC have full 24/7 access to their clock data via a web portal. Other nations have designed systems to deliver time to stock exchanges and for monitoring and verifying clocks, including the National Physical Laboratory (NPL) in the United Kingdom, which utilizes dedicated fiber and transfers time via the loop-back method with PTP or the two-way method with White Rabbit [87], and the National Research Center in Canada, which utilizes a common-view disciplined clock method similar to the NISTDC [88].

The third and final question is, what will happen to stock exchanges if GPS is unavailable? A GPS failure would, of course, cause synchronization to be lost after a period commensurate with the holdover capability of the stock exchange clocks, if no other backup timing systems were available. This failure could be widespread, as GPSDCs are commonly found in stock exchange data centers. However, the loss of synchronization, unless it involved some extreme and unusual situations, such as clocks being targeted, spoofed, and falsely set by U. S. adversaries, would probably not cause the stock exchanges to stop trading

or cause markets to crash. Unlike some telecommunication systems (Section VI), which for technical reasons require synchronization requirements to be met to keep networks operational, the synchronization requirements of stock exchanges are regulatory requirements that were put in place to help protect investors. Thus, rather than resulting in a catastrophic incident that dominates the headlines; a synchronization outage would likely have the more subtle effect of simply providing less protection to investors against fraud and market manipulation, a situation which might be invisible and not known to those who incur monetary losses.

V. Timing Requirements and Dependencies of the North American Power Grid

The combined electric power transmission and distribution system in the U. S. is commonly known as the power grid or simply the “grid”. The grid is the most critical of our critical infrastructure systems, as it supplies the electricity that powers all other sectors. The grid is the backbone of our economic sector and essential to national safety and security, thus overestimating the importance of the grid as it applies to our quality of life in the United States is nearly impossible.

The North American Electric Reliability Corporation (NERC) is a not-for-profit international regulatory authority whose mission is to assure the effective and efficient reduction of risks to the reliability and security of the grid [89]. The grid that NERC oversees consists of four distinct interconnections, shown in Fig. 20, that supply power to the U. S. and Canada at a nominal frequency of 60 Hz. The two major interconnections are the Western, which supplies power to more than 80 million people in all or part of 14 western states and two Canadian provinces [90], and the Eastern, which supplies power than more than 200 million people (~68% of the U. S. population), in all or part of 36 states [91]. Most of the state of Texas is covered by the Electric Reliability Council of Texas (ERCOT) interconnection, and a fourth interconnection covers the Canadian province of Quebec. The grid systems in Alaska and Hawaii are not connected to the power system shown in Fig. 19, but the coordinating body in Alaska is an affiliate of NERC. All told, the U. S. power grid consists of more than 360,000 miles (~580,000 km) of transmission lines, including approximately 180,000 miles (~290,000 km) of high-voltage lines, connecting to about 7,000 power plants [92].

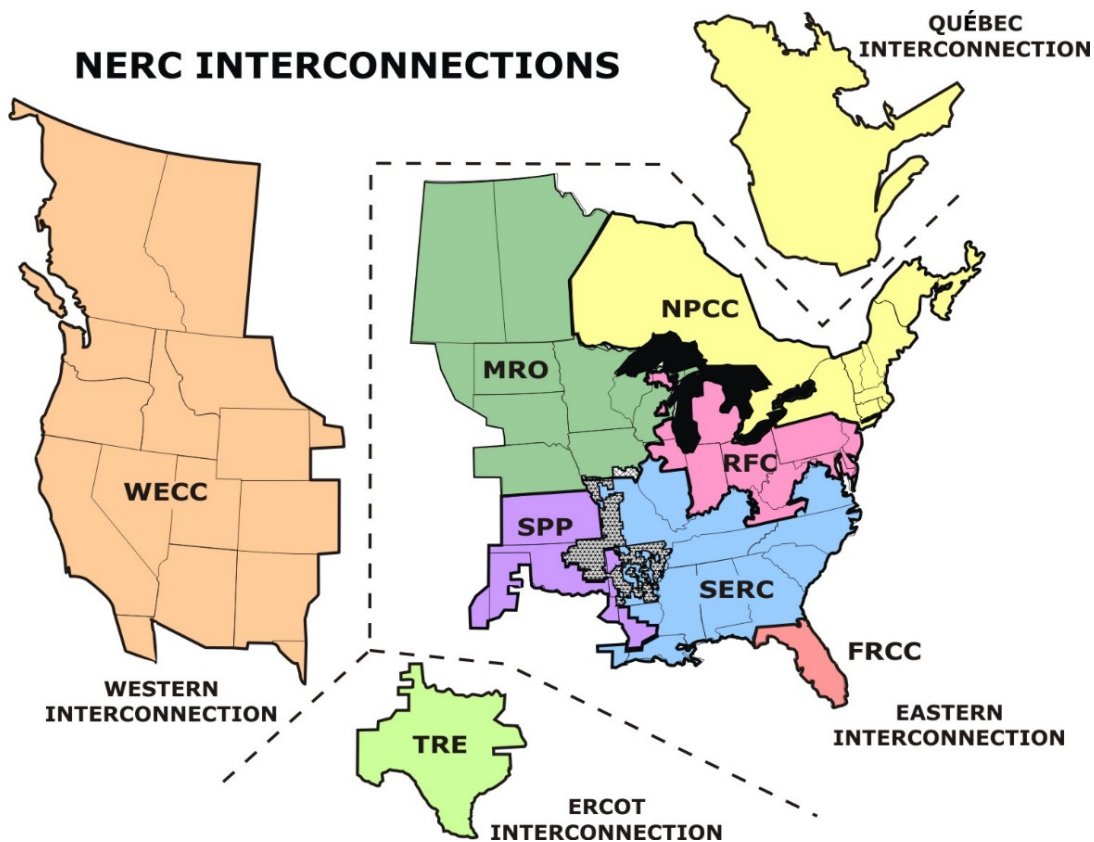


Fig. 20. Regions and interconnections in the North American power grid.

V.A The Importance of Time Synchronization to the Power Grid

Because the power grid encompasses many large and geographically separated systems that need to be synchronized, it requires access to a UTC reference clock that is easily accessible across a wide area. Time synchronization is needed for a variety of applications, including: continuous monitoring of the state of the power grid and its operating limits, protecting grid assets by quickly identifying and locating grid failures so they can be repaired before wide spread outages occur, and recently, as synchronization capability has continued to improve; advancing the grid's real-time measurement and control capabilities in ways that make better use of the grid's resources, so the needs of a growing population can continue to be met.

Locating faults on a transmission line is the application identified by power companies as having the most stringent synchronization requirements [93]. A fault on a transmission line can occur due to switching operations, a lightning strike, or other causes, and numerous methods of fault detection exist [94]. One method, known as travelling wave fault detection, is based on the principle that a traveling wave is generated at the point of a line termination or fault. Normally, the voltage and current waves travelling on a transmission line are related by the surge impedance of the line, but when a fault occurs, voltage collapses, tending towards 0, while the current wave doubles. This creates a high-frequency pulse that travels down the transmission line at nearly the speed of light. The arrival time of this pulse is measured at both ends of the line and time stamped by clocks that are synchronized to each other and that usually have a resolution of $0.1 \mu\text{s}$ [95]. This method became effective with the use of GPSDCs, which made it possible to identify the tower nearest to where a fault occurred. For example, if the transmission line is supported by high voltage towers that are spaced 300 m apart, locating the fault requires at least $1 \mu\text{s}$ synchronization, or the period of a 300 m wavelength [96].

Another synchronization requirement involves disturbance monitoring equipment (DME), which includes protective relays, sequence of event recorders, and digital fault recorders (DFRs). Relays are the workhorse devices used to protect grid assets; they measure the time interval between a triggering event such as a fault and respond quickly, for example by tripping a breaker, opening a line, or isolating a power plant, before significant damage occurs [96]. Time synchronization is normally provided by either GPS, or through a PTP device which is usually synchronized internally with a GPSDC and that may also distribute time via an IRIG time code [97]. Sequence of event recorders are used to record exactly when an incident, such as a breaker trip, occurs. They help answer questions about chronological relationships, allowing investigation of issues such as lightning strikes and downed poles. The current and voltage on the grid is typically monitored and sampled at high rates with DFRs, which time stamp data that is aligned with data from other recorders to help correlate and analyze power grid events. For example, the time stamped voltage and current data can be used to find the point of origin of a lightning strike. The NERC synchronization requirement for DME is $\pm 2 \text{ ms}$ with respect to UTC [98]. This also applies to the data recorded for post-mortem failure analysis [99]. However, in many cases utility companies rely upon and demand much tighter synchronization and categorize DME equipment based on guidelines presented in the International Technical Commission (IEC) 61850 series of standards, where the accuracy requirements for the time stamping of events are summarized in five time performance classes, T1 to T5, that range from 1 ms to $1 \mu\text{s}$ [100, 101].

Most of the attention in power grid synchronization, however, relates to phasor measurement units (PMUs), the instruments that perform synchrophasor measurements. A phasor is a phase vector that represents a sine wave, in this case the 60 Hz alternating current (AC) sine wave generated by the power grid. A complete sine wave can be constructed by a single vector that rotates counter clockwise in a circle. This relationship works because a sine wave is a projection of a rotating circle's radius. A phasor rotates at an angular velocity (ω), expressed in radians per second, of $\omega = 2\pi f$, where f is the frequency of the waveform. In this case, f is the grid frequency, or 60 Hz.

In the power industry, however, the term phasor normally refers not to a rotating phasor, but rather to a stationary phasor, recorded at a given point in time, that indicates the amplitude and direction of the sine wave. Figure 21 is a simplified diagram where the red vectors inside the circle are stationary phasors, or “snapshots” of the sine wave at a given time. The blue lines connect the stationary phasor to the corresponding point on the sine wave.

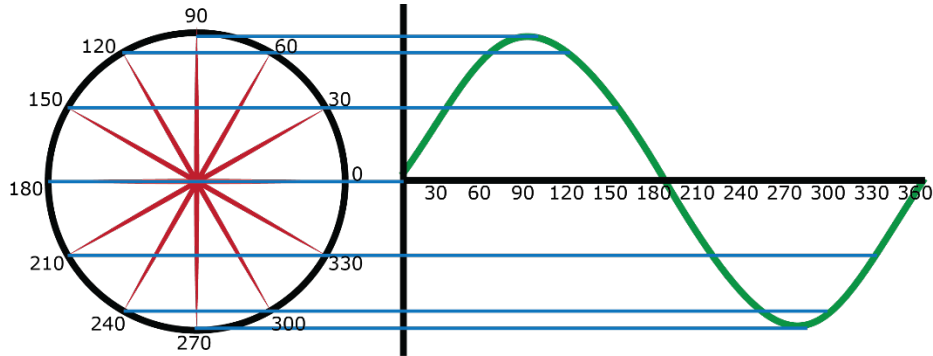


Fig. 21. The relationship between a phasor and a sine wave.

Synchronized phasors, or synchrophasors, are referenced to an absolute point in time by using UTC as a common clock, distributed by a GPSDC. Figure 22 shows the phase of a sine wave sampled at two points, δ_0 and δ_1 , with $\Delta\delta$ the phase angle difference between the two samples, as represented by the two stationary phasors inside the circle, and Δt is $1 / \text{sampling rate}$. The frequency, f , can be derived as $f = f_{\text{nom}} + \Delta f$, where $\Delta f = \Delta\delta / (360^\circ \times \Delta t)$.

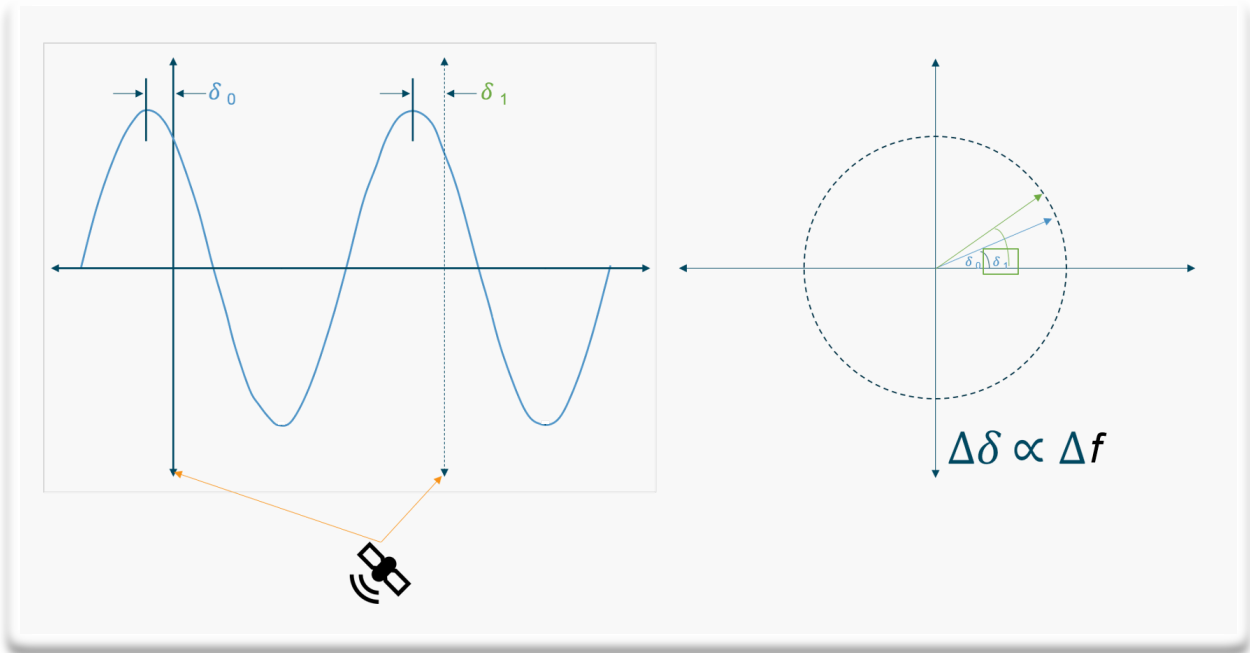


Fig. 22. A synchrophasor measurement.

A PMU measures the total vector error (TVE), which is the difference between a measured phasor and a true or ideal phasor, with no measurement uncertainty. The word “total” is used because the TVE includes all error sources, including phasor angle and phasor magnitude, in addition to the time synchronization error. The TVE is expressed as a percentage, and the acceptable PMU tolerance for a TVE measurement is 1%, as stated in the *IEEE C37.118.1* standard. A 1% TVE can be visualized as a circle drawn on the end of the phasor, as shown in Fig. 23, with the size of the circle exaggerated for clarity. The error reaches 1% even when the error in amplitude is zero, if the error in angle reaches 0.573° . PMUs are calibrated by generating a test signal of known accuracy and meet specifications if the samples they collect do not lie outside the circle [102].

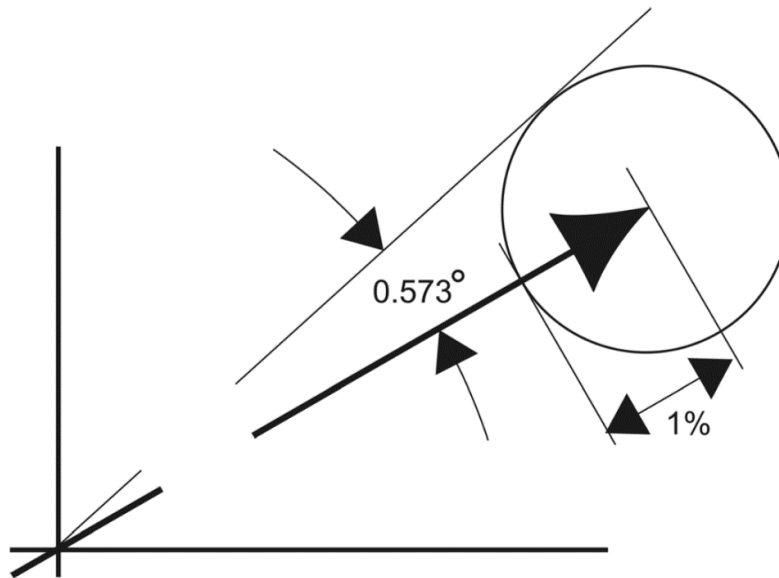


Fig. 23. A total vector error (TVE) of 1%.

Because the period of a 60 Hz sine wave is $16\,666\ \mu\text{s}$, an error in phase angle of 0.573° will occur if the time synchronization error exceeds $26.5\ \mu\text{s}$ ($\text{period} / (360 / 0.573)$). However, because Section 4.3 of the *IEEE C37.118.1* standard states that a time source “at least 10 times better than these values corresponding to 1% TVE is highly recommended”, the requirement is usually considered to be $2.6\ \mu\text{s}$ or less, with the desired accuracy stated as $1\ \mu\text{s}$, which corresponds to a 0.022° error in phase angle. Section 4.3 of the standard states that the resolution should be commensurate with this level of accuracy, by asserting that the “time tag shall accurately resolve time of measurement to at least $1\ \mu\text{s}$ within a specified 100 year period.” Section 4.3 also mentions traceability to UTC in three places, beginning with the clause stating that “the PMU shall be capable of receiving time from a reliable and accurate source, such as the Global Positioning System (GPS), that can provide time traceable to UTC” [102].

A PMU performs synchrophasor measurements and records time stamped voltage and current vector values that are sent very quickly to an operating center, typically at a rate of 30 to 120 samples per second. Data from multiple PMUs are sent to a centralized phasor data concentrator (PDC) where the readings are aligned and compared to provide a near real-time indication of the health of the power grid. This data can be used to make real-time decisions about power allocations within the grid. GPS was an enabling technology for synchrophasor measurements. Although the concept of a PMU had been explored long before the GPS satellites were launched, GPS was necessary to provide the necessary timing accuracy. The first prototype PMU was assembled by Phadke and his colleagues at Virginia Tech in 1988 with a GPS clock [103], and commercial units soon followed. As of May 2017, about 1,800 PMUs were in use in North America, as shown in the map (Fig. 24) provided by the North American SynchroPhasor Initiative (NASPI).

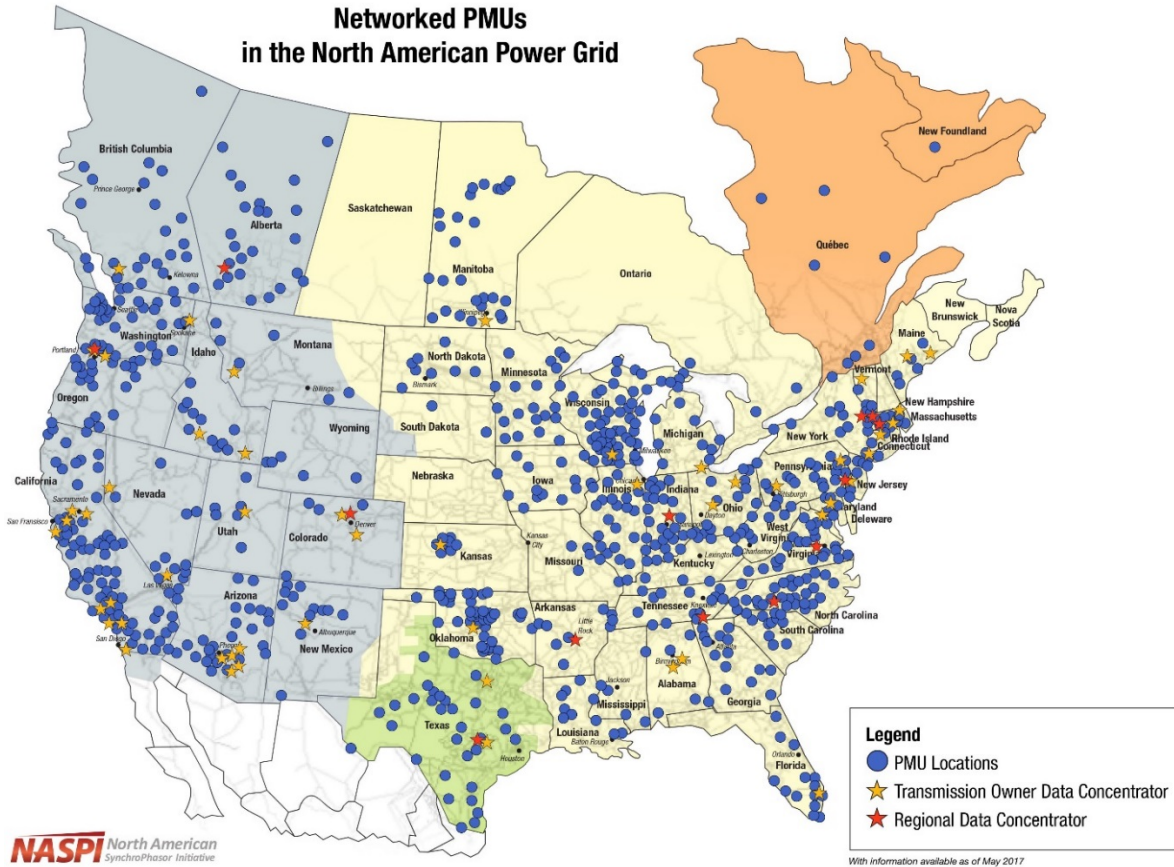


Fig. 24. Map of PMUs in the North American power grid as of May 2017.

V.B Time Synchronization Requirements of the Power Grid

Power grid synchronization requirements are essentially based on speed and response time, the equipment needs to report upon and respond to events quickly enough to prevent wide area failures. Table 6 summarizes the basic synchronization requirements.

Table 6. Summary of power grid synchronization requirements.

Grid application	Synchronization requirements (relative to UTC)
Sequence of events recorder	50 μ s to 2 ms [95, 98, 99]
Digital fault recorder	50 μ s to 2 ms [95, 98, 99]
Protective relays	< 1 ms to 2 ms [95, 98, 99]
Synchrophasor/PMU	2.6 μ s, 1 μ s resolution [102]
Traveling wave fault location	100 ns [95]

Timing requirements for communication systems used by power utilities are often based on PTP and are provided in [100], and general guidelines for designing power substation timing systems are provided in [104].

V.C GPS Timing Dependencies of the Power Grid

The synchronization requirements of the power industry vary (Table 6), because the entire system has evolved over many years. The older parts of the system have less stringent timing requirements because they were designed using technologies that predated the Global Positioning System (GPS). The newer parts of the system, in particular PMUs, were designed using GPS, and thus are dependent upon GPS – either through direct reception or through a network connection via PTP, to provide them with the necessary accuracy.

The power industry is well aware that bad things can happen when synchronization is lost. An often cited example is the massive blackout that occurred on August 14, 2003 in eight states and two Canadian provinces. The outage lasted for two days in some U. S. areas with rotating blackouts occurring for about two weeks in parts of Canada. Time synchronization was not the root cause of the blackout, but recommendation 12 from the detailed post-mortem report prepared by NERC was to “Install Additional Time-Synchronized Recording Devices as Needed,” noting that:

“A valuable lesson from the August 14 blackout is the importance of having time-synchronized system data recorders. NERC investigators labored over thousands of data items to synchronize the sequence of events, much like putting together small pieces of a very large puzzle. That process would have been significantly improved and sped up if there had been a sufficient number of synchronized data recording devices.

NERC Planning Standard I.F — Disturbance Monitoring does require location of recording devices for disturbance analysis. Often time, recorders are available, but they are not synchronized to a time standard. All digital fault recorders, digital event recorders, and power system disturbance recorders should be time stamped at the point of observation with a precise Global Positioning Satellite (GPS) synchronizing signal. Recording and time-synchronization equipment should be monitored and calibrated to assure accuracy and reliability.

Time-synchronized devices, such as phasor measurement units, can also be beneficial for monitoring a wide-area view of power system conditions in real-time, such as demonstrated in WECC with their Wide-Area Monitoring System (WAMS).” [105]

In today’s grid, where GPS clocks and PMUs are widely deployed (Fig. 24) and that, when working normally, have no trouble meeting synchronization requirements, the concern has shifted to the reliability of GPS and the dependence upon one timing source. For example, control applications, both automated or based on analytical decisions made by power grid operators, will be adversely affected if the time stamps from GPS are inaccurate or lost. If the time signal received by a PMU is jammed or spoofed, that error will cause false calculations of phase angle and a misalignment of measured grid conditions relative to other PMUs.

The *IEEE 2030.101* standard addresses timing system vulnerabilities in Section 4.14, noting that the “single best tool” for mitigating vulnerabilities is a high-quality time standard with sufficient holdover to allow “synchronization processes to continue unabated for extended periods of time” when the reference signals are lost. It also recommends installing multiple diverse time references, because it is “unlikely that diverse reference signals will be lost at the same time [104].” However, the signals it recommends, such as PTP, which is likely to be derived from GPS, and WWVB, which is unlikely to have sufficient accuracy, simply reminds us that diverse backup choices are limited.

In conclusion, power grid operators have benefitted tremendously from GPS time synchronization but concerns about GPS reliability have currently dissuaded them from pursuing fully automated operations, which appears to be a wise decision, especially when we consider the consequences of wide area outages. As the NASPI Time Synchronization Task Force stated in 2017:

“Power system owners and operators use primarily GPS as the source for timing and determining asset position. Today, GPS disruptions complicate (with higher cost, longer duration, and lower efficiency) but do not kill grid operations. For mission-critical time-synchronized applications in the future, however, GPS and alternate time sources (and the ways they are delivered and used) will need to become more reliable [95]”.

VI. Timing Requirements and Dependencies of Telecommunication Systems

The word telecommunication is derived from the Greek prefix, *tele*, which means “far off” or “from afar” or “from a distance” and the Latin word, *communicare*, which means to “make something common” or to “share”. The field of telecommunications is exceptionally broad, because it includes systems that share information of almost any nature across distances of nearly any length through a transmission medium. The transmission medium can be wired, for example, information sent over wires or cables, or wireless, involving electromagnetic radiation that travels through free space as radio waves or light.

Long established methods of sharing information, including one-way transmission or broadcast systems such as television and AM/FM radio, qualify as telecommunication systems. These systems do have frequency syntonization requirements, albeit modest by modern standards, typically parts in 10^5 [106]. However, the focus of this section will be on two-way telecommunication systems used by individuals for personal communications and to obtain information. These systems can be classified as critical infrastructure because human life and safety often depend upon them, especially in times of emergency. Mobile phone networks, first demonstrated by Motorola in 1973 [107] but used by relatively few Americans until the 1990s, have become the dominant system of this type. Figure 25, obtained from ITU data [108], graphs the number of mobile phone subscriptions in the U. S. from 2000 to 2018, a number that now exceeds 400 million; or more than one phone for every U. S. resident. The graph also shows how the number of fixed-telephone subscriptions, commonly known as land lines, has declined during this same period, and that mobile phone subscriptions surpassed land line subscriptions in 2004. Today, the ratio of mobile phone to land line subscriptions is about 4:1, indicating that many individuals and businesses now rely exclusively on mobile phones as a communication device. Surveys conducted in June 2019 indicate that about 96% of U. S. residents over the age of 18 have a mobile phone and for about 81% of residents, their mobile phone is a smartphone [109], which means that it provides access to numerous other telecommunication systems in addition to voice telephone calls, including functioning as a mobile Internet platform.

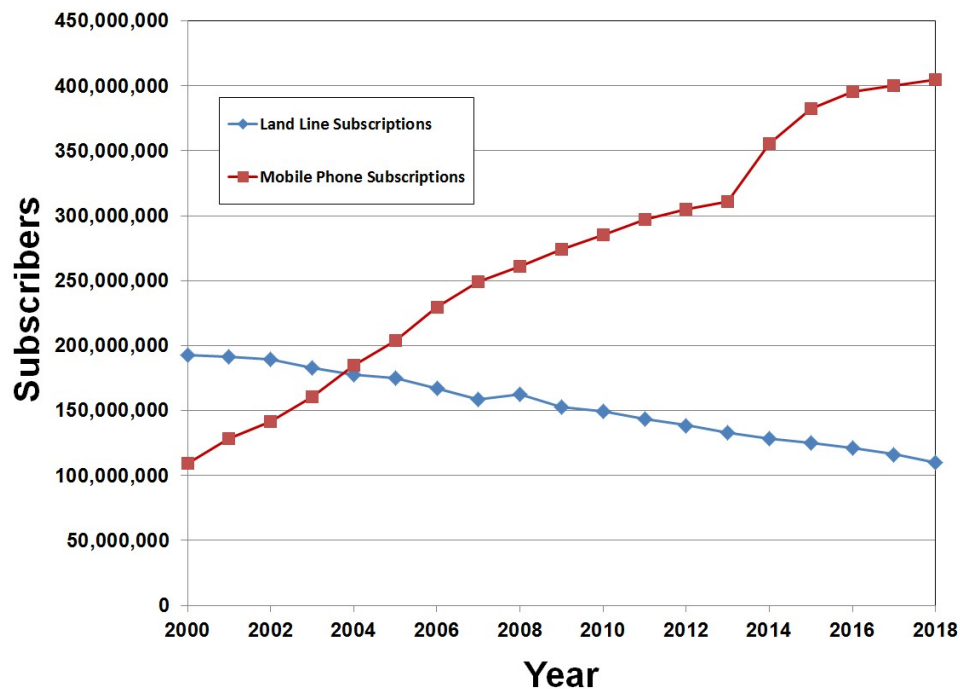


Fig. 25. U. S. mobile phone versus land line subscriptions, 2000 to 2018.

VI.A The Importance of Time Synchronization to Telecommunication Systems

Synchronization in telecommunications networks is the process of aligning the clocks of transmission and switching equipment so that operations occur at the correct time and in the correct order. Synchronization requires the receiver clock to acquire and track timing information in a transmitted signal that originated from another clock. The required level of synchronization/syntonation varies, of course, depending upon the type of network, but all networks will lose data or in some cases fail completely if the basic requirements are not met. More stringent synchronization can help to maximize the use of available bandwidth so that networks can operate at full capacity and make the best use of available resources.

The beginnings of the current high level of importance that synchronization has on telecommunication networks can be roughly traced to the divestiture of AT&T on January 1, 1984 [110]. Prior to this date, all telephone calls in the U. S. were routed through a single telephone company that was handling approximately 600 million calls per day. A very tiny percentage contained data, from data links between computer modems, but for all practical purposes everything carried by the network was voice communications. Synchronization for the entire network was provided by the Bell System Reference Frequency (BSRF), which was located underground in Hillsboro, Missouri. It consisted of three cesium clocks, had a frequency offset of less than 1×10^{-11} [110, 111], and, because the system originated prior to GPS, traceability was established to the USNO via a Loran-C link [110].

After the divestiture, AT&T was divided into seven Regional Bell Operating Companies (RBOCs). Each of these companies now had to maintain its own reference clock that had equivalent accuracy to the BSRF. These clocks were called the primary reference source (PRS) by the American National Standards Institute (ANSI) [112], or the primary reference clock (PRC) by the ITU [26]. In order for the various carriers to interconnect and exchange data with each other, each of their clocks had to appear to be synchronized with all of the others, even though no synchronization paths existed between carriers (Fig. 26). This is called plesiochronous operation, which means that the system is almost, but not quite, synchronized. It works with a minimal amount of data loss if each PRS stays within frequency tolerances defined with respect to UTC.

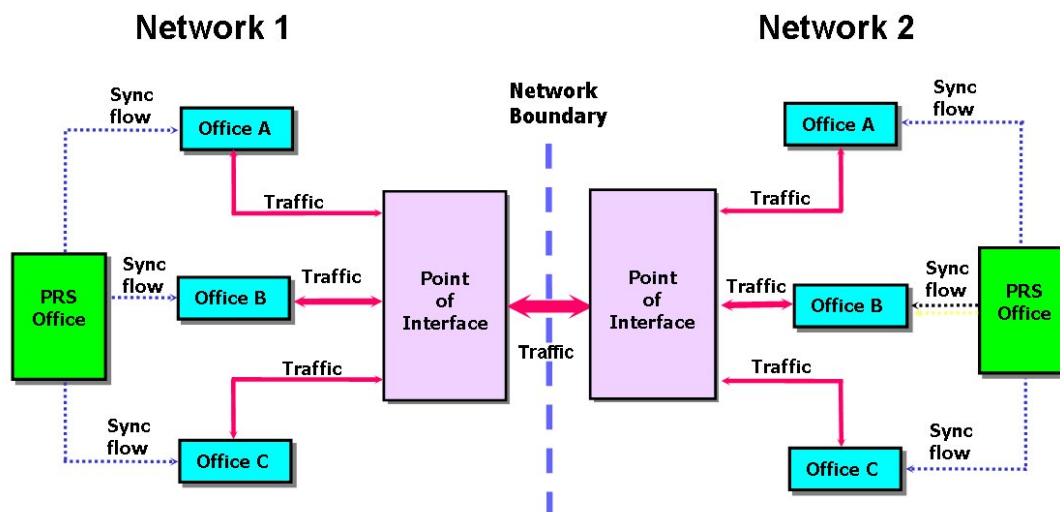


Fig. 26. A plesiochronous connection between networks that each maintain their own PRS.

To illustrate this, consider that the traffic exchanged in Fig. 25 is via a DS1/T1 connection [112], which remains the primary digital telephone circuit in the U. S. and Canada. It consists of a digital data stream clocked at a frequency of 1.544 MHz. This data stream is divided into 24 voice channels, each with 64 kHz of bandwidth. Each voice channel is sampled 8000 times per second. When the time difference between the two PRS units exceeds the period of the sampling rate, a cycle or frame slip occurs. This either results in a noisy call or a dropped call. The slip rate, SR , is calculated as

$$SR = \frac{T_{s\text{amp}}}{F_{\text{diff}}}, \quad (3)$$

where $T_{s\text{amp}}$ is the sampling period (125 μs), and F_{diff} is the frequency difference between PRS A and PRS B. If PRS A is high in frequency with respect to UTC by $+1 \times 10^{-11}$ and PRS B is low in frequency with respect to UTC by -1×10^{-11} , then the interval between slips is

$$SR = \frac{125 \times 10^{-6} \text{ s}}{2 \times 10^{-11}} = 6\,250\,000 \text{ s} = 72.5 \text{ days}. \quad (4)$$

Figure 27 depicts a slip as an accumulated time or phase error. Here the unit interval (UI) is equal to the period of the DS1/T1 bit frequency, or 647.7 ns. A slip occurs when a complete frame (193 bits) has been lost ($647.7 \text{ ns} \times 193 = 125 \mu\text{s}$, the period of the sampling rate). Even if one PRS was far more accurate than the requirement, the frequency error in the other PRS would eventually cause a slip, and thus all clocks must maintain good synchronization.

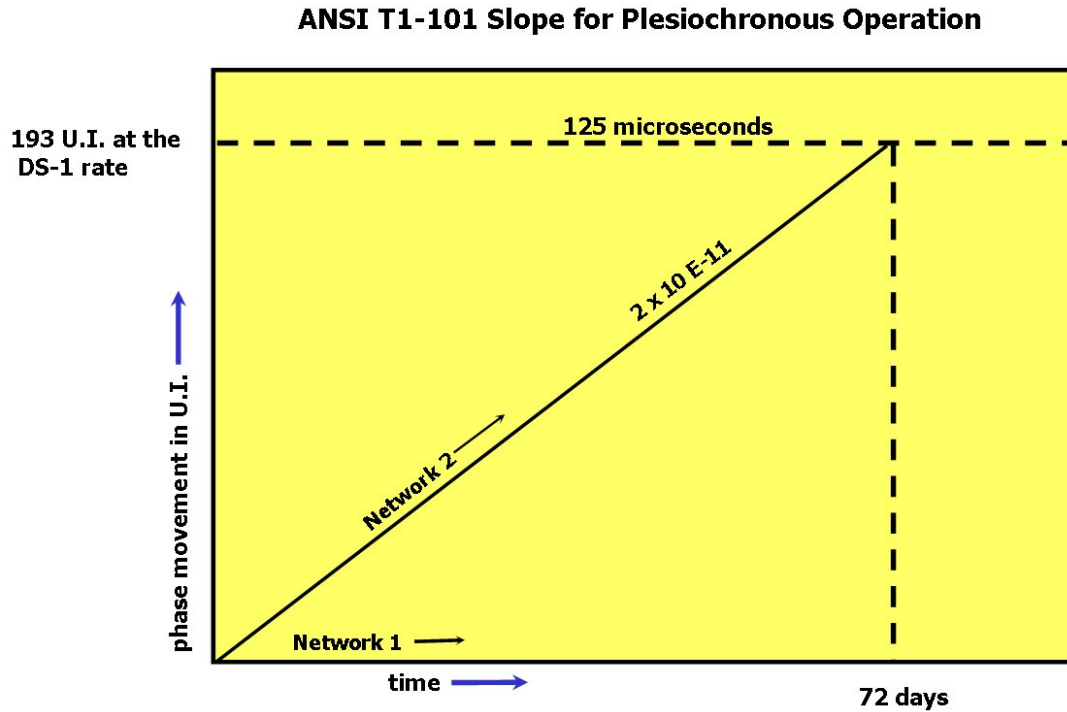


Fig. 27. An accumulated time error of 125 μs results in a cycle or frame slip.

Earlier, in Section II.D.3, it was noted that a PRC has a long-term frequency accuracy requirement of 1×10^{-11} or better with verification to Coordinated Universal Time (UTC) in the ITU standard [26], and the same specification applied in the ANSI/ATIS standards documents in the U. S. [112]. The ITU also now

has a standard for an enhanced primary reference clock (EPRC) which states that “the maximum allowable fractional frequency offset for observation times greater than one week is 1 part in 10^{12} , over all applicable operational conditions” [113]. The PRC/ERPC standards can be realized with an autonomous clock, that operates independently of all other sources, or a non-autonomous clock, which is disciplined by “UTC-derived precision signals received from a radio or satellite system” [113]. In most cases, the only autonomous clock that can meet these requirements is a cesium clock, which is typically prohibitively expensive. Thus, a non-autonomous clock, in nearly all cases a GPSDC, typically serves as the PRC.

Time synchronization, and not just frequency syntonization, has become a prerequisite in mobile phone networks. Mobile phone technology originally utilized two types of radio networks to communicate with cellular towers; GSM, an acronym for Global System for Mobile Communications, and CDMA. GSM is far more common with about 90% of mobile phones having GSM capability. GSM is a time division multiple access (TDMA) system that divides a radio frequency into time slots and then allocates slots to multiple calls. Each user has access to the entire frequency channel for the duration of the slot (Fig. 28). It requires the base transceiver station (BTS) to use a single frequency source of absolute accuracy of better than 0.05 parts per million (5×10^{-8}) for both the carrier frequency signal and for clocking the time base. However, the time synchronization of the BTS units with respect to each other is optional [114], even though they are usually synchronized due to the presence of a GPSDC at each BTS.

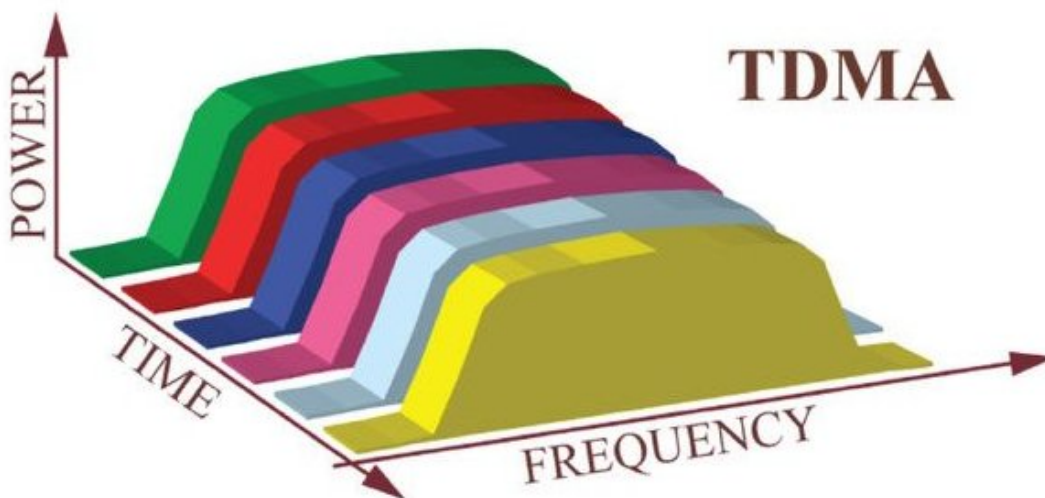


Fig. 28. Time division multiple access, users share the same frequency at different times.

In a CDMA system (Fig. 29), the frequency syntonization requirement is the same, 5×10^{-8} , but the BTS units must be synchronized in time. All users always share the same frequency spectrum and there are no time slots to allocate. However, all transmitted information is encoded by use of spread spectrum codes, and individual users are allocated unique codes that allow them to be identified. This allows the desired signal to be extracted while rejecting everything else as random noise. Every BTS in a CDMA network is aligned to CDMA system time, which originates from GPSDCs. Synchronization should be within $\pm 3 \mu\text{s}$, and if the BTS supports multiple simultaneous CDMA channels, within $\pm 1 \mu\text{s}$. The holdover requirement is $\pm 10 \mu\text{s}$ for a period of not less than eight hours, even if the external source of CDMA system time is disconnected [115].

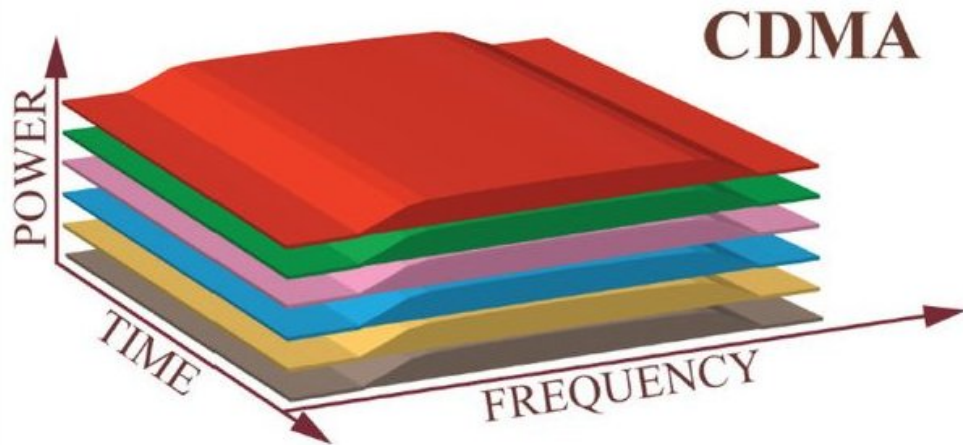


Fig. 29. Code division multiple access, users share the same frequency at the same time.

The original GSM and CDMA systems described above represented the 2nd generation of mobile phone technology known as 2G. The 2G networks were the first digital networks, as the original mobile phone networks, now known as 1G, were analog. Today's protocols transfer data at higher speeds and are thus better suited for use as mobile Internet platform, and are somewhat loosely defined as 3G, 4G, and 5G. The protocols are not backward compatible with each other, but some mobile phone providers still support older protocols, including 2G, to keep older equipment operational.

The 3G standards include the Universal Mobile Telecommunications Systems (UMTS), which blended the two technologies. It was based on GSM but utilizes Wideband Code Division Multiple Access (WCDMA) to obtain more bandwidth [116] and can transfer data faster than the original CDMA standard, which was known as IS-95. The 4G standards are defined by the *ITU IMT-Advanced* standard [117] include and include Long Term Evolution Advanced (LTE Time Division Duplex (TDD) or LTE-Advanced) [118, 119] and WirelessMAN Advanced, which has been commercialized using the name WiMAX [120]. The 5G networks, known as 5G New Radio (NR), began deployment in 2019, and follow a new standard that is still under development as of February 2020 [121].

The 4G and 5G networks are based on variations of the orthogonal frequency-division multiple access (OFDMA) technique, which transmits data at low rates across many closely spaced carriers. Because the signals are orthogonal to each other, they do not interfere. This technique makes better use of the spectrum than CDMA, because it uses less bandwidth and thus has better scalability. However, because it is not a spread spectrum technology, but rather a specialized form of multiplexing, OFDMA may be less secure due to the lack of spread spectrum codes. It is also more complicated to implement. The synchronization requirement is similar to CDMA, about 1 μ s.

VI.B Time Synchronization Requirements of Telecommunication Systems

The various mobile phone standards each have time synchronization requirements that are too stringent for free running clocks to meet without frequent calibration, typically near 1 μ s, as summarized in Table 7. The table also includes the previously discussed frequency syntonization requirements which can be met by a cesium clock, or even free running rubidium clock in the case of the GSM or CDMA carrier frequency requirement.

Table 7. Summary of telecommunications synchronization requirements.

Network application	Synchronization or syntonization requirements (relative to UTC)
Primary Reference Clock (PRC)	1×10^{-11} [26, 112]
Primary Reference Time Clock (PRTC)	100 ns [122]
Enhanced Primary Reference Clock (EPRC)	1×10^{-12} [113]
Enhanced Primary Reference Time Clock (EPRC)	30 ns [123]
GSM or CDMA carrier frequency	5×10^{-8} [114, 115]
LTE-TDD (wide area), 4G	10 μ s [119]
CDMA2000, 2G	3 μ s [115]
LTE-TDD (home area), 4G	3 μ s [119]
WCDMA-TDD, 3G	2.5 μ s [116]
WiMAX, 4G	~ 1 μ s [120]

VI.C GPS Timing Dependencies of Telecommunication Systems

It seems clear when looking at mobile phone network time synchronization requirements that these systems were enabled by GPSDCs and designed with their usage in mind. GPS is specifically mentioned in nearly all of the standards documents, and time distribution systems under U. S. control that can meet the accuracy requirements without being dependent upon GPS are rare (Table 4). It is also interesting to note that systems have not been designed that require synchronization better than the 1 μ s that GPS can provide without calibration, a not so subtle indication that no substitute timing source exists that can provide that type of accuracy at all locations. Therefore, it is easy to see that the mobile phone network has a very high level of dependency on GPS, much higher than the level of dependency in stock exchanges or the power grid.

To obtain much needed timing system diversity, some telecomm providers in the U. S. are turning to global navigation satellite system (GNSS) clocks [124] that receive multiple satellite constellations in addition to GPS, including GLONASS (Russia), Galileo (Europe), and Beidou (China). Table 8 summarizes these systems. There are obvious policy and security arguments against relying upon time signals that are controlled by non-U.S. interests, and there is also the technical argument that multi-constellation reception could be of limited value, because jamming or interference that impacts GPS may impact all GNSS systems with similar frequencies and signal levels. However, network engineers realize that a clock that receives multiple GNSS signals, or even a clock that receives signals on more than one

GPS frequency, will be less susceptible to accidental jamming than the single frequency GPS clocks that are typically in use.

Table 8. GPS compared to other global navigation satellite systems.

System	BeiDou	Galileo	GLONASS	GPS
Host Nation	China	European Union	Russia	United States
Altitude (km)	21,150	23,222	19,130	20,180
Orbit Period (hh:mm)	12:38	14:05	11:26	11:58
Center Frequencies (GHz)	1.561098 (B1) 1.589742 (B1-2) 1.20714 (B2) 1.26852 (B3)	1.57542 (E1) 1.191795 (E5) 1.27875 (E6)	~1.602 (L1) ~1.246 (L2) (FDMA, satellites use multiple frequencies)	1.57542 (L1) 1.2276 (L2) 1.17645 (L5)
Number of satellites	35	30	24	31

The impact of a long lasting, widespread GPS outage on mobile phone networks would likely be staggering, due to the large number of GPSDCs in use. The number of GPSDCs employed in U. S. mobile phone networks already numbers in the hundreds of thousands and may increase dramatically as 5G coverage improves. Mobile phone providers install their BTS equipment at shared cellular sites that are typically located on towers or rooftop platforms, but they do not share equipment. Therefore, a quick glance at any cellular site usually reveals multiple GPS antennas, which either indicates multiple providers or multiple systems operated by the same provider that each require synchronization. CTIA, a trade association that represents the wireless communications industry in the U. S., estimates that 349,344 cell sites were operational in 2018 [125]. Following the plausible assumption that on average at least one, and not more than three GPSDCs are located at each site, results in an estimate of 350,000 to 1 million devices now in service. A widespread rollout of 5G sites is now underway to meet the strong demand for faster data transfer. The 5G rollout involves “small cell” or “micro sites” that cover a limited geographic area and more sites will be needed to provide the necessary coverage. Some analysts predict that the number of cellular sites will triple or quadruple within a few years [126, 127], with perhaps a corresponding increase in the number of GPSDCs.

VII. Summary

In an attempt to provide a better understanding of the dependency of critical infrastructure systems in the United States on GPS time signals, the first half of this report, contained in Sections I through III, presented the fundamentals and terminology of time distribution systems, exploring concepts including Coordinated Universal Time (UTC) and traceability to the International System (SI) of units, as well as time synchronization, time stamping, frequency syntonization, accuracy, stability, resolution, and the differences between free running and disciplined clocks. From there, the various methods of distributing time from a reference clock so that it can synchronize clocks at other locations were explained, including the one-way, loop-back, common-view, and two-way methods of transferring time. This was followed by a summary of the available public and restricted access time distribution systems that are controlled by U. S. interests.

The second half of the report, Sections IV through VI, was dedicated to examining the time requirements and dependencies of three critical infrastructure sectors; stock exchanges, the electric power grid, and telecommunication systems. Each of these sections was divided into three parts. The first part studied the importance of time synchronization to the sector, the second part studied the sector's time synchronization requirements, and the third part evaluated the sector's dependency on GPS as its timing reference. All three sectors were found to heavily employ and rely on GPS disciplined clocks as part of their everyday operations. However, the level of dependency was found to be highest in the telecommunication sector, particularly in mobile phone networks, where the large number of GPS clocks in place, coupled with the stringent synchronization requirements, would make continued operation without GPS impossible. In contrast, stock exchanges would be likely to continue to operate without GPS, albeit with reduced protection for investors. The daily operation of the power grid would certainly become more labor intensive and difficult, as daily operations would be severely hampered and diagnostic tools would become less useful, but grid operation would likely continue as it did prior to the introduction of GPS.

The key takeaway from this study is that significant GPS dependencies exist in all three sectors, and that all sectors would benefit from improvements in both resiliency and diversity. Resiliency can be improved by establishing standard practices that manufacturers could implement to make GPS clocks less likely to fail, and diversity can be improved by recommending and developing new timing methods and sources to both backup and complement GPS clocks.

VIII. Acknowledgements

The author thanks and acknowledges the Department of Homeland Security for their sponsorship of this effort and for being given the opportunity to write this report. Special thanks are due to Brannan Villee, Rachel Wallner, and Ernest Wong of the United States Department of Homeland Security, Science and Technology Directorate (DHS S&T); and James Platt of the Cybersecurity and Infrastructure Security Agency's National Risk Management Center (CISA/NRMC).

This report was partially funded by DHS in their continued effort to support more secure and resilient infrastructure.

The author also thanks former NIST staff member Stefania Romisch for her initiation of this effort, and thanks current NIST staff members Elizabeth Donley, Victor Zhang, Jeff Sherman, Andrew Novick, Bijunath Patla, and Jonathan Hardis for their review of this manuscript and for many helpful comments and corrections.

IX. References

- [1] P. Jorgensen, "The Short Term Quality of NAVSTAR Tracking Data," *SAMSO-TR-79-28*, September 1978.
- [2] S. Pace, G. Frost, I. Lachow, D. Frelinger, D. Fossum, D. Wassem, and M. Pinto, *The Global Positioning System: Assessing National Policies*, RAND MR-614-OSTP, p. 246, 1995.
- [3] White Paper, "The Clinton Administration's Policy on Critical Infrastructure Protection: Presidential Decision Directive 63," May 22, 1998.
- [4] U. S. Department of Transportation, Volpe Center, "Vulnerability Assessment of the Transportation Infrastructure Relying on the Global Positioning System," *Volpe National Transportation Systems Center Report*, August 2001.
- [5] *Volpe Report*, p. 58.
- [6] *Volpe Report*, p. 20.
- [7] *Volpe Report*, p. 60.
- [8] S. Lo, B. Peterson, P. Enge, and P. Swaszek, "Loran Data Modulation: Extensions and Examples," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 43, no. 2, pp. 628-644, April 2007.
- [9] M. Lombardi, T. Celano, and E. Powers, "The Potential Role of Enhanced LORAN-C in the National Time and Frequency Infrastructure," *Proceedings of the 2005 International Loran Association (ILA) and Technical Symposium*, 15 p., October 2005.
- [10] B. Parkinson, J. Doherty, J. Darrah, A. Donahue, L. Hirsch, D. Jewell, W. Klepczynski, J. Levine, K. Lewis, E. Stear, P. Ward, P. Rambow, "Independent Assessment Team (IAT) Summary of Initial Findings on eLoran," *Institute for Defense Analyses Report*, 32 p., January 2009.
- [11] C. Norman, "Synchronization in the Global Telecommunications Network," *2006 NIST - ATIS Workshop on Synchronization in Telecommunication Systems (WSTS)*, March 2006.
- [12] Department of Homeland Security – Coast Guard, "Terminate Long Range Aids to Navigation (Loran-C) Signal," *United States Federal Register*, vol. 75, no. 4, p. 998, January 7, 2010.
- [13] G. Offermans, S. Bartlett, and C. Schue, "Providing a Resilient Timing and UTC Service Using eLoran in the United States," *Journal of the Institute of Navigation*, vol. 64, no. 3, pp. 339-349, Fall 2017.
- [14] A. O'Connor, M. Gallaher, K. Clark-Sutton, D. Lapidus, Z. Oliver, T. Scott, D. Wood, M. Gonzalez, E. Brown, and J. Fletcher, "Economic Benefits of the Global Positioning System (GPS)," *RTI Report Number 0215471*, sponsored by the National Institute of Standards and Technology, RTI International, 306 p., June 2019.
- [15] *RTI Report*, p. ES-4.
- [16] Title 3, The President, "Executive Order on Strengthening National Resilience through Responsible Use of Positioning, Navigation, and Timing Services," *Executive Order 13905*, Federal Register, vol. 85, no. 3, pp. 9359-9361, February 18, 2020.
- [17] Resolution 1 of the 13th *Conference Generale des Poids et Mesures (CGPM)*, 1967.
- [18] M. Lombardi, "Microsecond Accuracy at Multiple Sites: Is it Possible Without GPS?," *IEEE Instrumentation and Measurement Magazine*, vol. 15, pp. 14-21, October 2012.

- [19] T. Quinn, “Time, the SI, and the metre convention,” *Metrologia*, vol. 48, no. 4, pp. S121–S124, August 2011.
- [20] G. Panfilo and F. Arias, “The Coordinated Universal Time (UTC),” *Metrologia*, vol. 56, no. 4, 26 p., June 2019.
- [21] BIPM, *Circular T*, 384, January 9, 2020.
- [22] <https://www.bipm.org/en/bipm-services/timescales/time-ftp/Circular-T.html>
- [23] Joint Committee for Guides in Metrology (JCGM), “International vocabulary of metrology – Basic and general concepts and associated terms (VIM),” *JCGM 200*, 3rd edition, p. 29, 2008.
- [24] D. Matsakis, J. Levine, and M. Lombardi, “Metrological and legal traceability of time signals,” *InsideGNSS*, vol. 14, pp. 48-58, April 2019.
- [25] Global Positioning Systems Directorate, “Navstar GPS Space Segment/Navigation User Interfaces,” *Interface Specification IS-GPS-200H*, 2013.
- [26] International Telecommunication Union (ITU), “Timing characteristics of primary reference clocks: Amendment 1,” *ITU-T G.811 Amendment 1*, April 2016.
- [27] W. Riley, “Handbook of Frequency Stability Analysis,” *NIST Special Publication 1065*, 136 p., July 2008.
- [28] IEEE, “Standard definitions of physical quantities for fundamental frequency and time metrology—Random instabilities,” *IEEE Standard 1139*, 2008.
- [29] M. Lombardi, “How Accurate is a Radio Controlled Clock?,” *The Horological Journal*, vol. 152, pp. 108-111, March 2010.
- [30] M. Lombardi, “Evaluating the Frequency and Time Uncertainty of GPS Disciplined Oscillators and Clocks,” *NCSLI Measure: The Journal of Measurement Science*, vol. 11, pp. 30-44, 2016.
- [31] D. Piester, A. Bauch, T. Polewka, E. Staliunine, and K. Teichel, “Disciplined oscillators for traceable frequency and time in metrology and financial sectors,” *Journal of the Institute of Navigation*, vol. 66, no. 3, pp. 661-672, 2019.
- [32] Y. Shmaliy and L. Arceo-Miquel, “Efficient Predictive Estimator for Holdover in GPS-based Clock Synchronization,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, no. 10, pp. 2131-2139, October 2008.
- [33] J. Jespersen, “A Survey of Time and Frequency Dissemination Techniques,” *Proceedings of the 1970 Frequency Control Symposium*, pp. 322-324, 1970.
- [34] J. Levine, M. Weiss, D. Davis, D. Allan, and D. Sullivan, “The NIST Automated Computer Time Service,” *Journal of Research of the National Institute of Standards and Technology*, vol. 94, no. 5, pp. 311-321, September-October 1989.
- [35] E. Kaplan and C. Hegerty, *Understanding GPS: Principles and Applications*, 2nd ed., Artech House Publishers, 2005.
- [36] P. Misra and P. Enge, *Global Positioning System: Signals, Measurement, and Performance*, 2nd ed., Ganga-Jamuna Press, 2011.

- [37] D. Mills, J. Martin, J. Burbank, and W. Kasch, “Network Time Protocol Version 4: Protocol and Algorithms Specification,” *Internet Engineering Task Force RFC-5905*, June 2010.
- [38] D. Mills, *Computer Network Time Synchronization: The Network Time Protocol*, CRC Press, 2006.
- [39] IEEE, “Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems,” *IEEE Standard 1588-2019*, November 2019.
- [40] EndRun Technologies, “Precision Time Protocol (PTP/IEEE-1588),” *EndRun Technologies White Paper*, 2011 (<https://endruntechnologies.com/pdf/PTP-1588.pdf>).
- [41] D. W. Allan and M. A. Weiss, M. A., “Accurate Time and Frequency Transfer During Common-View of a GPS Satellite,” *Proceedings of the 34th Annual Meeting on Frequency Control*, pp. 334-346, May 1980.
- [42] W. Lewandowski, J. Azoubib, and W. Klepczynski, “GPS: Primary Tool for Time Transfer,” *Proceedings of the IEEE*, vol. 87, no. 1, pp. 163-172, January 1999.
- [43] G. Petit and Z. Jiang, “GPS All in View time transfer for TAI computation,” *Metrologia*, vol. 45, no. 1, pp. 35-45, December 2007.
- [44] M. Lombardi and A. Dahlen, “A common-view disciplined oscillator,” *Review of Scientific Instruments*, vol. 81, no. 5, pp. 055110/1-6, 2010.
- [45] D. Hanson, “Fundamentals of Two-Way Time Transfers by Satellite,” *Proceedings of the 1989 Frequency Control Symposium*, pp. 174-178, May 1989.
- [46] M. Desaintfuscien, *Data Processing in Precise Time and Frequency Applications*, Springer, 217 pages, 2007.
- [47] A. Bauch, “Time and frequency comparisons using radiofrequency signals from satellites,” *Comptes Rendus Physique*, vol. 16, no. 5, pp. 471-479, June 2015.
- [48] J. Serrano, M. Cattin, E. Gousiou, E. van der Bij, T. Wlostowski, G. Daniluk, and M. Lipinski, “The White Rabbit Project,” *Proceedings of the 2nd International Beam Instrumentation Conference*, Oxford, UK, pp. 936-942, Kobe, Japan, pp. 93–95, September 2013.
- [49] International Telecommunication Union (ITU), “Timing characteristics of synchronous equipment slave clock,” *ITU-T G8262/Y.1362*, November 2018.
- [50] E. Dierikx, A. Wallin, T. Fordell, J. Myyry, P. Koponen, M. Merimaa, T. Pinkert, J. Koelemeij, H. Peek, and R. Smets, “White Rabbit Precision Time Protocol on Long-Distance Fiber Links,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 7, pp. 945-952, July 2016.
- [51] J. Savory, J. Sherman, and S. Romisch, “White Rabbit-Based Time Distribution at NIST,” *Proceedings of the 2018 IEEE International Frequency Control Symposium*, 5 p., May 2018.
- [52] P. Panek, J. Kodet, and I. Prochazka, “Accuracy of two-way time transfer via a single coaxial cable,” *Metrologia*, vol. 50, no. 1, pp. 60-65, 2013.
- [53] J. Kodet, P. Panek, and I. Prochazka, “Two-way time transfer via optical fiber providing subpicosecond precision and high temperature stability,” *Metrologia*, vol. 53, pp. 18-26, 2016.
- [54] Ł. Śliwczyński, P. Krehlik, A. Czubla, Ł. Buczek, “Dissemination of time and RF frequency via a stabilized fiber optic link over a distance of 420 km,” *Metrologia*, vol. 50, no. 2, pp. 133-145, 2013.

- [55] Range Commanders Council: Telecommunications and Timing Group, “IRIG Serial Time Code Formats,” *IRIG Standard 200-16*, August 2016.
- [56] Society of Motion Picture and Television Engineers, “Television – Time and Control Code,” *SMPTE 12M-1-2008*, February 2008.
- [57] M. Lombardi, “The Reach and Impact of the Remote Frequency and Time Calibration Services at NIST,” *NCSLI Measure: The Journal of Measurement Science*, vol. 12, pp. 30-37, June 2019.
- [58] NIST Service ID 78110S (<https://shop.nist.gov>)
- [59] NIST Service ID 78100S (<https://shop.nist.gov>)
- [60] M. Korreng, “UTC Time Transfer for High Frequency Trading Using Is-95 CDMA Base Station Transmissions and IEEE-1588 Precision Time Protocol,” *Proceedings of the 42nd Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting*, pp. 359-368, November 2010.
- [61] Endrun Technologies, “CDMA Timing,” (<https://endruntechnologies.com/products/cdma-timing>).
- [62] NTP Pool Project (<https://www.ntppool.org/en/>), accessed January 2020.
- [63] Network Time Foundation Web Search, <http://support.ntp.org/bin/view/Servers/WebSearch>, accessed January 2020.
- [64] J. Sherman and J. Levine, “Usage Analysis of the NIST Internet Time Service,” *Journal of Research of the National Institute of Standards and Technology*, vol. 121, pp. 33-46, March 2016.
- [65] The current status of the NIST Internet Time Service is displayed at: <https://tf.nist.gov/tf-cgi/servers.cgi>
- [66] United States Code of Federal Regulations, “WWV-WWVH-WWVB Broadcasts,” *15 CFR § 200.107*, 2019.
- [67] G. Nelson, M. Lombardi, and D. Okayama, “NIST Time and Frequency Radio Stations: WWV, WWVH, and WWVB,” *NIST Special Publication 250-67*, 161 p., January 2005.
- [68] Satelles White Paper, <https://www.satellesinc.com/wp-content/uploads/2016/06/Satelles-White-Paper-Final.pdf>, 2019.
- [69] G. Gutt, D. Lawrence, S. Cobb, M. O’Connor, “Recent PNT Improvements and Test Results Based on Low Earth Orbit Satellites,” *Proceedings of the 2018 Precise Time and Time Interval (PTTI) Meeting*, pp. 72-79, January 2018.
- [70] M. Lewis, *Flash Boys: A Wall Street Revolt*, W. W. Norton Company, New York, 2014.
- [71] M. Lombardi, A. Novick, B. Cooke, and G. Neville-Neil, “Accurate, Traceable, and Verifiable Time Synchronization for World Financial Markets,” *Journal of Research of the National Institute of Standards and Technology*, vol. 121, pp. 436–463, 2016. <http://dx.doi.org/10.6028/jres.121.023>
- [72] H. Bessembinder, “Trade Execution Costs and Market Quality after Decimalization,” *The Journal of Financial and Quantitative Analysis*, vol. 38, no. 4, pp. 747-777, 2003. <http://dx.doi.org/10.2307/4126742>
- [73] A. Menkveld, “High frequency trading and the new market makers,” *Journal of Financial Markets*, vol. 16, no. 4, pp. 712-740, 2013. <http://dx.doi.org/10.1016/j.finmar.2013.06.006>

- [74] M. Goldstein, P. Kumar, and F. Graves, “Computerized and High-Frequency Trading,” *Financial Review*, vol. 49, no. 2, pp. 177-202. <http://dx.doi.org/10.1111/fire.12031>
- [75] Figure 14 appears in this article, attributed to Credit Suisse – A. Klein, “Congress wants to tax stock trades. Investors shouldn’t fret,” *Brookings Institution Web Site* (<https://www.brookings.edu/opinions/congress-wants-to-tax-stock-trades-investors-shouldnt-fret/>), June 10, 2019. An earlier version of this graph, with data through 2016 appears in: A. Avramovic, V. Lin, and M. Krishnan, “We’re All High Frequency Traders Now,” *Credit Suisse White Paper*, March 15, 2017.
- [76] United States Securities and Exchange Commission (SEC), *Report Pursuant to Section 21(a) of the Securities Exchange Act of 1934 Regarding the NASD and the NASDAQ Market*, (SEC, Washington, D.C.), 1996.
- [77] National Association of Securities Dealers (NASD), *OATS Reporting Technical Specifications*, (NASD, Inc. Rockville, MD), 1998.
- [78] Financial Industry Regulatory Authority (FINRA), *OATS Reporting Technical Specifications*, (FINRA, Rockville, MD), 2016.
- [79] Financial Industry Regulatory Authority (FINRA), “Equity Trading Initiatives: Synchronization of Business Clocks,” *FINRA Regulatory Notice 14-47*, (FINRA, Rockville, MD), 2014.
- [80] Financial Industry Regulatory Authority (FINRA), “Rule 6820. Clock Synchronization,” Adopted March 15, 2017.
- [81] United States Securities and Exchange Commission (SEC), “Joint Industry Plan; Order Approving the National Market System Plan Governing the Consolidated Audit,” *Release No. 34-79318, File No. 4-698*, November 15, 2016.
- [82] European Union, “Commission Delegated Regulation (EU) 2017/574 of 7 June 2016 supplementing Directive 2014/65/EU of the European Parliament and of the Council with regard to regulatory technical standards for the level of accuracy of business clocks,” *Official Journal of the European Union*, pp. L 87/148 – L 87/151, March 31, 2017.
- [83] Australian Securities & Investment Commission (ASIC), “Financial Markets: Domestic and Overseas Operators,” *ASIC Regulatory Guide 172*, pp. 112-113, May 2018.
- [84] Investment Industry Regulatory Organization of Canada (IIROC), “Guidance on Time Synchronization,” *IIROC Notice 16-0022*, February 2016.
- [85] The NIST GPS Data Archive (<https://www.nist.gov/pml/time-and-frequency-division/services/gps-data-archive>).
- [86] M. Lombardi, “Disciplined clock for providing a disciplined time and a disciplined frequency synchronous with a reference clock,” *United States Patent 10,466,363 B2*, granted November 5, 2019.
- [87] E. Laier-English, P. Whibberley, C. Langham, D. Hicks, and L. Lobo, “Update on the NPLTime® and Future Developments with White Rabbit,” *Proceedings of the 2017 Precise Time and Time Interval (PTTI) Meeting*, pp. 80-87, January 2017.
- [88] J. Bernard, A. Charbonneau, B. Hoger, H. Pham, and M. Gertsvolf, “NRC Remote Clock – A Secure and Traceable Time Source,” *Proceedings of the 2017 Precise Time and Time Interval (PTTI) Meeting*, pp. 73-79, January 2017.
- [89] North American Electric Reliability Corporation (NERC) web site (<https://www.nerc.com/>).

- [90] Western Electricity Coordinating Council (WECC) web site (<https://www.wecc.org/>).
- [91] M. Brown and Y. Wang, “Estimating the Energy-Efficiency Potential in the Eastern Interconnection,” U. S. Department of Energy, Oak Ridge National Laboratory, *Report ORNL/TM-2012/568*, 112 p., April 2013.
- [92] U. S. Department of Energy, “United States Electric Industry Primer,” Office of Electricity Delivery and Energy Reliability, *DOE/OE-0017*, 94 p., July 2015.
- [93] J. Allnutt, D. Anand, D. Arnold, A. Goldstein, Y. Li-Baboud, A. Martin, C. Nguyen, R. Noseworthy, R. Subramaniam, and M. Weiss, “Timing Challenges in the Smart Grid,” *NIST Special Publication 1500-08*, 28 p., January 2017.
- [94] IEEE, “IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines,” *IEEE Standard C37.114-2004*, June 2005.
- [95] NASPI Time Synchronization Task Force, “Time Synchronization in the Electric Power System,” *NASPI Technical Report NASPI-2017-TR-001*, PNNL-26331, 59 p., March 2017.
- [96] H. Lee and A. Mousa, “GPS Travelling Wave Fault Locator Systems: Investigation into the Anomalous Measurements Related to Lightning Strikes,” *IEEE Transactions on Power Delivery*, vol. 11, no. 3, pp. 1214-1223, July 1996.
- [97] W. Schenk, “Time Synchronization of protection relays to IEEE 1588/PTP,” *TICRO 100 Application Note*, Omicron Lab, 19 p. 2014.
- [98] North American Electric Reliability Corporation (NERC), “Disturbance Monitoring Equipment Installation and Data Recording,” *NERC Standard PRC-018-1*, 4 p., August 2006.
- [99] North American Electric Reliability Corporation (NERC), “Disturbance Monitoring and Reporting Requirements,” *NERC Standard PRC-002-2*, 38 p., September 2015.
- [100] International Electrotechnical Commission (IEC), “Communication networks and systems for power utility automation – Part 3: General Requirements,” *IEC 61850-3*, 2013.
- [101] C. Ozansoy, A. Zayegh, and A. Kalam, “Time Synchronisation in an IEC 61850 Based Substation Automation System,” *Proceedings of 2008 Australian Universities Power Engineering Conference (AUPEC)*, 7 p., January 2009.
- [102] IEEE, “IEEE Standard for Synchrophasor Measurements for Power Systems,” *IEEE Standard C37.118.1-2011*, 61 p., December 2011.
- [103] A. Phadke, “Synchronized Phasor Measurements ~ A Historical Overview,” *Proceedings of the 2002 IEEE/PES Transmission and Distribution Conference and Exhibition*, pp. 476-479, October 2002.
- [104] IEEE, “IEEE Guide for Designing a Time Synchronization System for Power Substations,” *IEEE Standard 2030.101-2018*, 118 p., March 2018.
- [105] North American Electric Reliability Corporation (NERC), “Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn?,” *Report by the NERC Board of Trustees to the NERC Steering Group*, 124 p., July 13, 2004.
- [106] United States Code of Federal Regulations, “Carrier Frequency Departure Tolerances,” §73.1545, 2001.
- [107] D. Steinbock, *The Mobile Revolution: The Making of Worldwide Mobile Markets*, Kogan Page Business Books, 352 p., June 2005.

- [108] International Telecommunication Union (ITU) web site, statistics page, accessed February 2020 (<https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>).
- [109] Pew Research Center, “Mobile Fact Sheet,” (<https://www.pewresearch.org/internet/fact-sheet/mobile/>), June 12, 2019
- [110] D. Glen, “Networks, Signaling, and Switching for Post-Divestiture and the ISDN,” *NTIA Technical Report 86-191*, National Telecommunications and Telecommunication Administration (NTIA), 120 p., February 1986.
- [111] R. Powers, “Reference Transmission Over Bell System Radio and Coaxial Facilities,” *Proceedings of the 28th Annual Symposium on Frequency Control*,” pp. 373-378, May 1974.
- [112] Alliance for Telecommunications Industry Solutions (ATIS), “Synchronization Interface Standard,” *ATIS-0900101.2013(R2018)*, 2018. This document supersedes the original *ANSI T1.101-1999* standard.
- [113] International Telecommunication Union (ITU), “Timing characteristics of enhanced primary reference clocks,” *ITU-T G.811.1*, August 2017.
- [114] European Telecommunications Standards Institute (ETSI), “Digital cellular telecommunications system (Phase 2+) (GSM); GSM/EDGE Radio subsystem synchronization (3GPP TS 45.010 version 14.2.0 Release 14),” *ETSI TS 145 010 V14.2.0*, 39 p., October 2017.
- [115] 3rd Generation Partnership Project 2 (3GPP2), “Recommended Minimum Performance Standards for cdma2000 Spread Spectrum Base Stations: Release B,” *C.S0010-B, Version 2.0*, February 2004.
- [116] 3rd Generation Partnership Project (3GPP), “Universal Mobile Telecommunications Systems (UMTS); Synchronisation in UTRAN Stage 2,” *3GPP TS 125.402 V6.2.0*, March 2005.
- [117] International Telecommunication Union Radiocommunication Sector (ITU-R), “Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-Advanced (IMT-Advanced),” *ITU-R M.2012.4*, November 2019.
- [118] 3rd Generation Partnership Project (3GPP), “Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9),” *3GPP TR 36.814 V9.2.0*, March 2017.
- [119] 3rd Generation Partnership Project (3GPP), “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management,” *3GPP TS 136.133 V13.3.0*, May 2016.
- [120] IEEE, “IEEE Standard for Air Interface for Broadband Wireless Access Systems,” *IEEE Standard 802-16-2017*, 2726 p., March 2018.
- [121] International Telecommunication Union Radiocommunication Sector (ITU-R), “Minimum requirements related to technical performance for IMT-2020 radio interface(s),” *ITU-R R15-IMT.2020*, in progress as of February 2020.
- [122] International Telecommunication Union (ITU), “Timing characteristics of primary reference time clocks,” *ITU-T G.8272/Y.1367*, November 2018.
- [123] International Telecommunication Union (ITU), “Timing characteristics of enhanced primary reference time clocks,” *ITU-T G.8272.1/Y.1367.1*, November 2016.
- [124] GPS World Staff, “ADVA introduces multi-band GNSS receiver for 5G timing accuracy,” *GPS World*, November 6, 2019 (<https://www.gpsworld.com/adva-introduces-multi-band-gnss-receiver-for-5g-timing-accuracy/>).

[125] CTIA, “2019 Annual Survey Highlights,” 7 p., 2019 (www.ctia.org).

[126] K. Rogers, “The drive to make 5G a reality has led to a surge in demand for tower climbers,” *CNBC web site*, May 3, 2019 (<https://www.cnbc.com/2019/05/03/the-drive-to-make-5g-a-reality-has-led-to-a-surge-in-demand-for-tower-climbers.html>).

[127] B. Snyder, “5G could require cell towers on every street corner,” IDG Communications, *CIO web site*, September 8, 2016 (<https://www.cio.com/article/3117705/5g-could-require-cell-towers-on-every-street-corner.html>).