Legal Traceability of Time

Speaker/Author:

Judah Levine
Time and Frequency Division and JILA
NIST and the University of Colorado
325 Broadway
Boulder, Colorado 80305
Voice: (303) 497 3903
FAX: (303) 497 6461
jlevine@boulder.nist.gov

Abstract

A "traceable" measurement is one that can be related to national or international standards using an unbroken chain of measurements, each of which has a stated uncertainty. This paper describes the traceability of time signals, with a special focus on the legal aspects of time. It provides examples of when the legal aspects of time become important. It also shows that legal traceability is not a purely technical question, since the legal and technical definitions of time are not precisely the same in the US at present.

Introduction

A "traceable" measurement is one that can be related to national or international standards using an unbroken chain of measurements, each of which has a stated uncertainty. In this paper I will describe the traceability of time signals, with a special focus on the legal aspects of this question. Legal traceability is not a purely technical question – the legal and technical definitions of time are not precisely the same (at least in the US at present), and this difference could be significant in practice.

Although an unbroken chain of measurements is a necessary requirement for traceability, the uncertainties in the measurements that make up the links of the chain are not really sufficient to characterize a subsequent measurement – we must have additional information on the stability of these links, and especially on the stability of the clocks at both ends of the chain. Without this information, the fact that a previous time stamp from some source was found to be within an acceptable tolerance of a reference standard does not by itself imply (much less guarantee) that the current one will also satisfy the same requirement. Information on stability is particularly important when the individual links in the chain of traceability are not measured simultaneously. This lack of simultaneous evaluation is common in the time and frequency business.

We can address this issue statistically by combining the time that has elapsed since the last calibration with a statistical estimate of the frequency stability of a clock at the end of the chain (or the stability of the delay through a link in the middle) to arrive at some confidence interval for the current measurement. The result will be a function of the interval since the last calibration, the noise model of the clock and the channel, and the precision of the measurements themselves. After some period of time has elapsed, the resulting uncertainty of the overall process (which is at best a statistical extrapolation based on an assumed model and not an actual measurement) might or might not still satisfy our initial requirements.

In all of the commonly used clock models, the time dispersion resulting from the stochastic fluctuations in the frequency of a clock increases as some power of the interval since the last calibration. The best devices (those whose noise spectrum is dominated by nothing more divergent than white frequency noise) have a time dispersion that increases only as the square root of this interval. Since the average cost of repeated calibrations depends on the reciprocal of the time interval between them, a cost-benefit analysis suggests that the interval between calibrations should be increased as long as the resulting uncertainty still satisfies our initial requirement for the overall uncertainty in the measurement chain. The largest interval that satisfies this initial requirement results in optimum performance, in the sense that the required root-mean-square uncertainty is realized with the fewest calibrations. These methods are particularly useful when calibrations are expensive or when the calibration channel itself makes a significant contribution to the overall uncertainty.[1,2]

Unfortunately, all oscillators have flicker and random-walk frequency fluctuations, and this favorable dependence of the cost/benefit ratio on the interval between calibrations becomes less and less appropriate as this interval increases because the contributions of these processes dominate the variance. The performance degrades faster than the cost decreases in this regime, so that this domain is suitable only for undemanding applications.

The uncertainty associated with basing traceability on previous calibrations exists for a mechanical artifact (a voltmeter, for example) as well, but our confidence in the stability of a properly maintained and locally available mechanical artifact usually is much higher than it is for a complex system based on a remote reference standard. The fact that the channel between our device and the reference is usually not under our direct control does not help matters.

Applications that require traceable measurements usually have documentation requirements as well as technical ones. Depending on the details of the application, these requirements might range from maintaining a simple log of the calibrations to real-time oversight or auditing by a disinterested third party using encrypted and digitally-signed messages. Systems that can support these requirements can become quite complex, since they must be protected both against outsiders and insiders. NIST (and some other national laboratories) provide only some of these services; other services (especially those intended for satisfying commercial or financial requirements) are (or will be) provided by private third-parties using time signals that are traceable to national standards.

The Treaty of the Meter

The Treaty of the Meter is the basis for all international cooperation on questions of standards and precision metrology. The treaty was signed in 1875 in Paris and was ratified by the US Senate in 1878. The treaty was modified in 1921, and the modified version was ratified by the US Senate in 1923. The modifications did not make any substantial changes to the original document.

The treaty established the International Bureau of Weights and Measures (Bureau International des Poids et Mesures or BIPM), which is currently located in Sèvres, a suburb of Paris. The BIPM is managed by the International Committee of Weights and Measures (Comité International des Poids et Mesures, or CIPM). The President of the CIPM is currently Professor J. Kovalevsky, who is at the Observatoire de la Côte d'Azur, and the US representative is Dr. Karen Brown, the Acting Director of NIST.

The organizational structure defined in the Treaty of the Meter was initially intended to deal with maintaining and calibrating artifact standards, such as the standard meter and kilogram. The responsibilities of the BIPM were expanded over time to include other standards activities. In the following discussion, we will discuss only the current arrangement, which dates from 1987 when the responsibility for dealing with standards of time and frequency was transferred to the BIPM from the Bureau International de l'Heure. The older arrangements are described in the literature.[3] The CIPM appoints a number of consultative committees to provide technical advice on questions that are referred to them. The committee that is important for this discussion is the Consultative Committee on Time and Frequency (CCTF), which was formerly called the Consultative Committee for the Definition of the Second (CCDS). The CCTF in turn appoints a number of working groups and sub-committees as necessary. These sub-groups deal with specific questions. Two that are important for this discussion are the Working Group on International Atomic Time (TAI) and the sub-group on GPS and GLONASS Time Transfer Standards. The web page of the BIPM (www.bipm.fr) has additional details about these organizational details. In particular, that page describes other relevant working groups, such as the one that is concerned with the realization of primary frequency standards and another that deals with time and frequency transfer using non-GPS methods such as two-way satellite time transfer.

Time, Frequency and the BIPM

Although time and frequency were originally thought of as distinct quantities with independent definitions, this distinction has not been significant for about 30 years. When frequency standards based on atomic transitions were first developed in the 1950s, the initial plan was to use these standards to realize the standard of frequency but to maintain the standard of time astronomically. This method proved to be very cumbersome, and the realization of the standards for time and frequency were unified into their current configuration on 1 January 1972.

Since 1972, the length of the second has been defined using the frequency of a hyperfine transition in the ground state of the cesium atom.[4] International Atomic Time (TAI, using the French word order) is a time scale based on this definition of the second. The length of the day defined in this way is somewhat shorter than the length of the day defined by astronomical methods, which is called UT1. Leap seconds are introduced into atomic time to keep the absolute magnitude of the time difference less than 0.9 seconds. The resulting time scale (TAI + leap seconds) is called Coordinated Universal Time (UTC, or TUC using the French word order), and it is the basis for all civilian time keeping. (This process of adding leap seconds is what makes UTC a "coordinated" time scale.) The rates of TAI and UTC are identical between leap seconds. Currently, the length of the UTC day is shorter than the average astronomical day by about 2.6 ms (a fractional offset of about 3×10^{-8}), so that leap seconds are required about every 12 or 18 months. The most recent leap second, which made TAI-UTC = 32 s, was added at the end of December, 1998. The difference between the lengths of the UT1 and UTC days is currently somewhat smaller than the average value quoted above, and no additional leap seconds have been announced as of March, 2001.

Although UTC is sometimes called Greenwich Mean Time (GMT) in the older literature, the two are not the same in principle. The GMT time scale is based solely on astronomical observations whereas UTC is derived from TAI and uses the astronomical data only to decide on the need for a leap second. Neither GMT nor UTC reflect the meridian transit of the actual sun at Greenwich because of the annual variation in the orbital speed of the earth. The difference (called the "equation of time"), can be as large at ±15 minutes. The maximum difference occurs early in November, and the maximum rate of change is near the end of December (at the Winter Solstice).

The BIPM computes UTC and TAI using data from a world-wide ensemble of about 250 commercial cesium standards and hydrogen masers. Most of these clocks are located at National Metrology Institutes (NMIs). The clocks at the different locations are compared using a number of different techniques, including common-view observations of the satellites of the Global Positioning System (GPS) and two-way satellite time transfer. (Common-view GPS is described below.) The time scale identified as UTC (with no modifier) refers to the scale computed by the BIPM. This scale is sometimes referred to as UTC(BIPM) when there is a possibility for ambiguity.

The computation at the BIPM assigns a weight to each commercial clock based on its previous stability; the scale also includes data from a number of primary frequency standards. These data are used to make small adjustments to the rate of the scale; a typical adjustment would change the rate of TAI by about $\Delta f/f = 1 \times 10^{-15}$. These steering corrections are too small to be seen by most users, but they are comparable to the stability of the time scales maintained by many NMIs and must be included in local time scales.

The algorithm that is used by the BIPM to compute UTC is called ALGOS; it was introduced in the 1970s. The basic algorithm has remained unchanged, although its detailed operation has been modified a number of times since then. It is designed to optimize the long-term stability of the scale at the expense of real-time output.

The desire to optimize the long-term frequency stability of UTC and the mechanics of collecting the data from the contributing laboratories mean that both UTC and TAI are computed after the fact and are not available in real-time. Generally, the computing for any month is not completed until the 16th of the following month; although some of this delay could be reduced by more rapid data collection by the BIPM, some part of it is an inevitable consequence of the retrospective nature of the computation, which is driven by the goal of maximizing the long-term stability of the scale.

UTC(NMI) and Circular T

Since UTC and TAI are not available in real time, most National Metrology Institutes define a local realization of UTC using data from an ensemble of the atomic clocks at the laboratory. The data from these clocks are combined to compute a time scale that is used to realize an estimate of UTC in real-time. These real-time versions are identified as UTC(NMI) to distinguish them from UTC as computed by the BIPM. At NIST, for example, the average time of the local clock ensemble is computed using an algorithm called AT1 (which is similar in concept to ALGOS), and the UTC derived from this computation is identified as UTC(NIST).

The difference between UTC and each UTC(NMI) is published monthly by the BIPM in its Circular T. The magnitudes of these differences vary from month to month, but are on the order of nanoseconds for laboratories such as NIST or the US Naval Observatory. (These fluctuations are caused by the flicker and random-walk frequency changes that characterize both the clock ensembles at the laboratories and TAI itself.)

Most laboratories make small adjustments to UTC(NMI) to steer it to UTC. The steering algorithm used by each laboratory must be a compromise between the conflicting goals of timing accuracy and frequency smoothness. At NIST, for example, these steering corrections are made only at the start of a month and are announced in advance. The magnitude of the monthly frequency correction is generally of order ±1 ns/day ($\Delta f/f = \pm 1.2 \times 10^{-14}$) or less. A frequency change of this magnitude is somewhat smaller than the frequency stability of the best commercial cesium frequency standards, although it can be detected easily by a commercial hydrogen maser. It is therefore invisible to almost all users. Time adjustments are never used.

Mutual Recognition Arrangements

In 1999, the directors of many of the metrology laboratories that subscribe to the Treaty of the Meter agreed to establish agreements under which measurements and calibrations performed at one laboratory would be deemed equivalent (at some specified accuracy level), to an equivalent measurement at another laboratory. These agreements were a response to the increasingly international character of calibration and measurement activities. These agreements are still being developed, but there are already some prototype examples in the area of time and frequency metrology. Examples are NORAMET, the North and Central America Metrology Cooperative, which links NIST with laboratories in Canada and Mexico, and a memorandum of understanding between NIST and USNO regarding equivalence of time and frequency signals generated at the two laboratories.

Distributing time and frequency signals

At all timing laboratories, UTC(NMI) is defined at a point called the *reference plane*, and the time delay between this point and the user's equipment must be measured. Since this delay is at least 3 ns/m, it is often much larger than the difference between UTC and UTC(NMI) or between UTC(NMI_1) and UTC(NMI_2). In many configurations, the uncertainty in this delay measurement is the largest single contribution to the overall error budget of the time comparison.

There are a number of methods of measuring this delay, including direct calibration of the time-transfer equipment and estimating it as one-half of the measured round-trip delay. These estimates are complicated by the variability both in the delay and in its asymmetry, which limits the usefulness of round-trip methods. The spectrum of the variations often has both relatively rapid stochastic components (due to atmospheric turbulence, for example) and longer-period components (due to changes in ambient temperature, for example). These variations usually limit (and may even dominate) the overall error budget for the entire time-transfer process.

Although a quantitative estimate of this problem depends on the details of the time-transfer equipment, loosely speaking it is relatively easy to keep the overall uncertainty in the estimate of the channel delay to less than 1 μs but it is almost impossible to achieve an overall uncertainty of less than 1 ns. It is possible, but quite difficult, to achieve an overall uncertainty of less than 10 ns. Nearly-diurnal changes in the delay (usually caused by a sensitivity of the system to ambient temperature) can be attenuated by averaging, provided that the clock in the receiver is sufficiently stable to support this.

The uncertainties and fluctuations in the delays through the different channels between a user and the various national timing laboratories may limit the practical usefulness of the equivalence defined by a Mutual Recognition Arrangement.

Legal time in the United States

To complicate matters further, legal time in the United States is not UTC but mean solar time as referenced to the Greenwich meridian (with appropriate integral-hour offsets derived from the longitude).[5] As we have described above, the difference between mean solar time and UTC (often called dut1) has a sawtooth-like character, decreasing slowly between leap seconds and increasing precipitously when a leap second is inserted. The peak-to-peak amplitude of this variation can be as large as 1.8 s in the long term, but is typically less than this value.

The value of dut1 is transmitted by a number of time services, including the NIST digital telephone service (ACTS) and the NIST radio stations WWV, WWVB and WWVH. Since the correction currently changes by approximately 75 ms per month and is transmitted with a resolution of 0.1 s, it is possible to monitor these services occasionally and cache the value received. (The rate of change of the correction is currently less than this value, but this is probably just a temporary effect.) Depending on the details of this process, different clients might have values that differ by 0.1 s (or even more in the immediate vicinity of a leap second).

Although it would not be difficult to use mean solar time for legal time stamps in principle, this is rarely done in practice. Even if the correction were more widely available than at present, the relatively poor resolution at which it is transmitted would totally dominate the accuracy of the message. A simpler approach would be to change the legal definition of time to be UTC; perhaps this has not been done yet because of the lack of clients that need legally traceable time with a resolution of substantially better than 1 s.

Practical difficulties at leap seconds

Although it is not strictly an issue of legal traceability, many digital systems have difficulties assigning an unambiguous time stamp in the vicinity of a UTC leap second. The leap second is always added as the last second of the day, and UTC time stamps in the vicinity of the leap second are identified as follows in the left-hand column:

	UTC Time stamps	Equivalent Computer Time
Day N	23:59:58	23:59:58
	23:59:59	23:59:59
	23:59:60 (the leap second)	23:59:59
Day N+1	00:00:00	00:00:00

Most computer systems keep time internally as the number of seconds since some epoch (0000 UTC on 1 January 1970 or 1 January 1900 are common choices), and there is no way of representing the leap second in this format. In the case of computer clocks, the most common practice is to stop the clock for 1 s during the leap second, effectively transmitting 23:59:59 twice. This is shown in the right-hand column above. An event which happens during the leap second therefore receives a time stamp that is indistinguishable from an event that happened in the previous second.

GPS system time does not include leap seconds at all, but the current and future leap second counts are transmitted by each satellite and can be subtracted from GPS time to construct UTC. Not all receivers parse this field correctly, and it cannot always be used to compute UTC in the past or very far into the future.

Since UT1, TAI and GPS system time do not have a discontinuity at the leap second, a number of proposals have been advanced to address the leap second problem using some combination of UTC and one of these scales to produce time stamps that are smooth and monotonically advancing across the leap second boundary. For example, Judah Levine and David Mills have proposed using the Network Time Protocol (NTP) to transmit a table giving the epochs of all past and currently-scheduled leap seconds.[6] They describe a procedure for combining this table with the UTC time-stamps transmitted by NTP to produce an equivalent of TAI – a time-scale that is both directly traceable to national standards and has no discontinuity at any leap second.

The difficulty could also be solved by simply abandoning the notion of adding leap seconds to UTC. The difference between UTC and UT1 would then increase by about 1 s/yr. The rate of this divergence is itself increasing, and the difference would probably exceed 2 minutes by the end of this century. Whether this divergence would or would not be more troublesome than handling the leap seconds themselves is a matter of current debate. However, if the definition of legal time in the US were not changed from mean solar time, then this difference would produce a noticeable divergence between legal time and the time defined by the Treaty of the Meter and maintained by all NMIs. The discrepancy could create significant problems for any application that required legally traceable time stamps.

Realization of UTC using GPS

Since signals from the GPS satellites are widely used for navigation, it may be helpful to discuss how the signals from these satellites can be used to produce time stamps that are legally traceable to national standards. Since GPS system time as transmitted by the satellites does not include leap seconds, I assume that any receiver first corrects for leap seconds using either the value transmitted as part of the GPS navigation message or using some other means. The following discussion is concerned with the procedure after that correction has been applied.

If the location of the receiver is known, receiving the signal from a single GPS satellite is enough to allow the receiver to solve for the difference between the local clock and GPS system time. At a minimum, this solution requires the ephemeris broadcast by the satellite, and it may also use other parameters in the navigation message such as the ionospheric model. If the receiver can process both the L1 and L2 frequencies, then the delay through the ionosphere can be estimated from the L1 – L2 dispersion. (If the location of the receiver is not known, it must be found using similar data from additional satellites.) The step from GPS system time to UTC can be accomplished in a number of different ways:

1. Using the offset between GPS time and UTC(USNO) broadcast by the satellites. This has the advantage that it requires no additional hardware or other data. However, the transmitted value is an extrapolation. The offset values transmitted by different satellites may differ because the parameters transmitted by the various satellites were uploaded at different times.

2. Estimate the difference between GPS time and UTC(NMI) using measurements from a timing laboratory rather than from the broadcast message itself. For example, NIST publishes the differences between GPS time and UTC(NIST) for each satellite that can be viewed from Boulder, Colorado with a delay of about 1 day; USNO and many other timing laboratories do something similar. Using these measurements to relate GPS system time to national standards could cancel or attenuate problems with the satellite clock, errors in the broadcast ephemeris, and ionospheric effects. This method does not depend on the extrapolations that form the basis of the previous method, but it requires ancillary data from another site and it cannot be completed in real time.

3. Use real-time common-view data with a National Metrology Institute. This method substantially improves the cancellation of the effects mentioned in the previous paragraph, since the two sites observe the satellite at the same time and use the same method to average the data. The common-view method is potentially the most accurate for this reason. However, it requires an active collaboration between the two sites. NIST provides this service to some customers using the common-view schedules published by the BIPM.

The first method is clearly the easiest one to use, and it is almost always the best choice for a user who wishes to use the GPS signals and who requires traceability to UTC with an uncertainty of 1 μs or more. At the opposite end of the spectrum, users who need the best-possible link to UTC must use method 3. The overall uncertainty using this method can be 10 ns or less, and is usually dominated by the uncertainties in the delay through the receivers and by multipath effects at the receiving antennas.

Using any of these methods, the final step involves estimating the difference between UTC and UTC(NMI). These data are available in Circular T, which is published by the BIPM about one month after the fact. In many cases this difference is small enough to be ignored.

No matter how the offset between the GPS system time and UTC is calculated, the final data must be corrected for the delay through the receiving equipment including any offset between the GPS time computed internally by the receiver and the emission of the physical pulse that is used to calibrate an external device. The apparent delay may vary with time as the satellites move across the sky due to changes in the multipath delay of the antenna.

Receivers can be calibrated by operating them in common-view with a second receiver located nearby whose delay is known. A common reference clock is used for the two receivers, and the two antennas are placed near each other. Another method measures the response of the receiver to a signal generated by a satellite simulator. Both of these methods have advantages, but neither is simple. These calibrations generally have uncertainties of at least a few ns, and this uncertainty is often a significant fraction of the overall error budget.

Summary

National laboratories maintain real-time estimates of UTC and disseminate these time scales using a number of different methods. The GPS system is currently the method of choice when the highest possible accuracy is required. Depending on how the signals are used and how carefully the receiving equipment is calibrated, GPS signals can provide traceability to national and international standards with an accuracy between about 10 ns and 1 μs. At the highest levels, the accuracy is usually limited by uncertainties in the delay through the channel between the satellite and the equipment at the receiving station. Estimating these delays is complicated by the presence of time-varying effects including multipath reflections at the antenna and the sensitivity of the receiving equipment to changes in the ambient temperature. Uncertainties in this delay may limit the technical traceability of the time stamps at the receiver.

Many National Metrology Institutes also transmit UTC(NMI) using other methods, including radio broadcasts and methods that use digital transmissions over dial-up telephone and other public networks (such as the Internet). The details of these services have been published[7,8] and are also available on the World Wide Web.[9]

In addition to issues of technical traceability, legal traceability may impose additional auditing and documentation requirements on the client system. These requirements will vary with the application; in their most comprehensive form they may require external monitoring of the client system. This sort of arrangement cannot be realized using only a one-way broadcast system. It is also not clear how to combine auditing requirements expressed in worst-case terms with measurements whose uncertainties are expressed as root-mean-square values.

Finally, legal traceability is further complicated during a UTC leap second and by the current US legal definition, which references legal time to mean solar time rather than to UTC.

References

1. Levine, Judah, "Time Synchronization Using the Internet," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 45, 1998, pages 450-460.

2. Levine, Judah, "Time Synchronization over the Internet Using an Adaptive Frequency-Locked Loop," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 46, 1999, pages 888-896.

3. Quinn, Terry, "The BIPM and the Accurate Measurement of Time," Proc. of the IEEE, Vol. 79, 1991, pages 894-905.

4. Le Systeme International d'Unites, published by the Bureau International des Poids et Mesures, Sevres, France, 7th edition, 1998.

5. United States Code, Title 15, Chapter 6, subchapter IX, sections 260-267.

6. Levine, Judah and Mills, David, "Using the Network Time Protocol (NTP) to Transmit International Atomic Time," Proceedings of the 32nd Annual Precise Time and Time Interval Planning and Applications Meeting (PTTI), November, 2000, in press.

7. Beehler, Roger E. and Lombardi, Michael A., "NIST Time and Frequency Services," NIST Special Publication 432, June, 1991.

8. Annual Report of the BIPM Time Section, Volume 12, 1999, pages 77-92. Published by the Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sèvres, France.

9. See www.time.gov, www.boulder.nist.gov/timefreq and www.usno.navy.mil for information about the NIST and USNO services. See www.bipm.fr for information about the BIPM.