4. Methods of Distributing Time and Frequency: A Review

Judah Levine
Time and Frequency Division
National Institute of Standards & Technology
325 Broadway
Boulder, CO 80303 USA
Tel: (303) 497-3903; Fax: (303) 497-6461
E-mail: jlevine@boulder.nist.gov

1. Abstract

Precise time and frequency information plays an increasingly important role in many areas of science and technology. Examples include high-speed digital communications, aircraft navigation, and wide-area computer networks. This paper outlines some of the methods that are currently used to distribute time and frequency information, and discusses how these techniques are likely to evolve in the next few years, in response to the demands for increases in both precision and accuracy. In particular, the use of digital-telecommunication systems is increasing very rapidly, and dissemination through such systems is therefore a major topic of consideration.

2. Introduction

The distribution of time and frequency information has become substantially more sophisticated since its inception, when it was confined largely to relatively simple radio broadcasts of the time of day, standard time intervals, and similar information. Some of these changes have been driven by the increasing sophistication of users, and by their requirements for higher and higher accuracy. The development of more advanced digital-communications methods, and more robust methods of evaluating the statistics of both the communication channels and the clocks at the nodes, are also important factors. Many new uses for time services have also been developed. An example is the Digital Signature and Time-Stamp system that we discuss below. Such uses make heavy use of both the improvement in digital-time networks, and in the increasingly powerful computers that are now available at modest cost.

Although time and frequency are closely related quantities, it is important to recognize the differences between them, because these differences affect how they are disseminated, and what uncertainties are associated with these processes. Time distribution is, of course, directly affected by the delay in the transmission channel between the clock and the user, and uncertainties in this delay enter directly into the error budget. A frequency, on the other hand, is a time interval (rather than an absolute time), so that the uncertainty in its distribution is limited more by temporal fluctuations in the transmission delay, rather than by its absolute magnitude. In other words, the time delay of a channel must be accurately known, if that channel is to carry time information, while the delay need only be stable, if frequency information is to be transmitted. This distinction is not of purely academic interest: many common channels have large delays, which change slowly enough with time that they may be thought of as constant for appreciable periods. Although the delays are constant to a very good approximation, they can be measured only by interrupting the channel, or by using relatively expensive ancillary hardware, which must itself be calibrated.

There are a number of techniques for quantitatively estimating the fluctuations in a clock itself, or in the delay of the channel between the clock and the user. The time-domain estimates, often called the two-sample [*Barnes et al., 1971*] or Allan variance [*Allan et al., 1988*], are defined as an IEEE standard [*IEEE, 1988*]. It is also useful to characterize the fluctuations in the frequency domain using techniques derived from Fourier analysis [*Percival, 1991*; *Rutman and Walls, 1991*]

3. Accuracy of time transfer

3.1 General considerations

The international standard of time is called Coordinated Universal Time (UTC). It is computed by the Bureau International des Poids et Mesures (BIPM) in Sèvres, France, using data supplied by a world-wide network of laboratories [*Quinn, 1991*]. The BIPM does not produce a physical realization of UTC, but rather publishes the differences between UTC and the time signals generated by the contributing timing centers. These differences are published after the fact, with a delay of somewhat less than two months.

Each timing center uses a different technique to generate its time signals. UTC(NIST), for example, is realized by a computer-controlled phase-stepper, using information from an ensemble of cesium standards and hydrogen masers. The time of UTC(NIST) is steered towards UTC, using discrete frequency changes on the order of $\pm 1 \times 10^{-14}$. If needed, these changes are applied at 0 hours on the first day of any month, and are published in advance in the NIST *Time and Frequency Bulletin*.

The difference between UTC, as computed by the BIPM, and the real-time realization of any laboratory, is typically tens of nanoseconds, and may be much larger in some cases. Timing information that is to be traceable to UTC must have this correction applied (after the fact), and any uncertainty in this correction represents a fundamental limit to the practical dissemination of UTC in real time. Since the uncertainties in these differences are on the order of nanoseconds, this represents a lower bound to how well UTC itself can be realized, at the present time, in any practical sense.

In addition to these problems of time-scale definition, users who are not located in the immediate vicinity of a timing laboratory must correct for the delay due to the finite transit time of the information along the transmission channel. This delay is on the order of nanoseconds per meter, and can be tens of milliseconds for a user a moderate distance away. The uncertainty with which it can be determined often represents the largest contribution to the overall uncertainty of the distribution process. A number of methods, which either measure the delay directly or minimize its impact on the error budget of the system, have been developed. We now turn to describing these methods in some detail.

3.2 Simple models of the delay

The simplest way of dealing with the transmission delay is to ignore it, and many users with modest timing requirements do just that. Many users of the time signals broadcast by NIST radio stations WWV and WWVH, for example, fall into this group. A somewhat more sophisticated strategy is to approximate the actual delay using some average value–perhaps derived from some simple parameter characterizing the distance between the user and the transmitter. Many of the users of the NIST Automated Computer Time Service [*Levine et al., 1989*] find the fixed-delay assumption that is used by the transmitters by default adequate for their needs. Similarly, users of the signal from NIST radio station WWVB often do not measure the transmission delay itself, but rather model it, based on average atmospheric parameters and the physical distance between themselves and the transmitter. The relatively simple models used to correct transmissions from high-frequency radio stations are not very accurate, and they are adequate only for users with very modest timing requirements. This is a result of the inadequacy of the models, however, and is not a fundamental limitation of the technique. Transmissions from GPS satellites, for example, can be corrected to much higher accuracy, using much more complex models of the satellite orbit and the ionospheric and atmospheric delays [*Lewandowski et al., 1991*].

The fundamental limitation to modeled corrections is often due to the fact that the adequacy of the model cannot be evaluated independently of the transmission itself. The GPS system is a notable exception, because of the redundancy provided by being able to observe several satellites simultaneously, and this technique is especially powerful when the satellites are located at different azimuths with respect to the antenna.

When modeling is not sufficiently accurate to estimate the path delay, users must use some technique to measure the delay with an uncertainty small enough for the task at hand. It is important to remember the obvious fact that the delay is a function of the channel between the transmitter and the user. There is generally no feasible way of correcting a broadcast-type service at the transmitter end for this reason: each user must compute an individual delay, and make an individual correction. In addition, measuring the delay usually requires the active collaboration of both the transmitter and receiver. Broadcast-type services are generally not amenable to making accurate delay measurements for these reasons, and may therefore have somewhat lower accuracy than what could be achieved with point-to-point transmissions between cooperating stations, using the same quality hardware.

3.3 Portable clocks

The simplest method for measuring the delay is to carry a clock between the two end points, and many timing laboratories used this technique in the past. It is limited, in practice, by the finite frequency stability of the portable clock while it is being transported along the path. It is also limited by uncertainties in a number of relativistic corrections that may have to be applied if the path is long, or if the end-points are at different heights above the geoid. While this method can be used to calibrate a static-transmission delay, it is too slow and cumbersome to evaluate the temporal fluctuations in the delay through real transmission media (due to changes in temperature or other environmental parameters, for example).

Even though various electrical- or electromagnetic-distribution methods are usually faster and more convenient, the portable-clock method remains useful. This is because it tends to be affected by a very different set of systematic uncertainties than the other methods to be described below, and it may provide estimates of these effects that are difficult to achieve in any other way [*Gifford and Barthomew, 1994* ; *Wheeler et al., 1994*].

3.4 Two-way time transfer

Another common method for measuring delay is two-way time transfer, in which signals are sent in both directions along the path. If the path is reciprocal, the one-way delay is estimated as one-half of the round-trip-transit time. Real paths are unlikely to be perfectly reciprocal: there is likely to be a small lack of reciprocity even in the atmosphere [*Jespersen, 1989*], and the receivers and transmitters at the two end points are unlikely to be perfectly balanced [*Kirchner, 1991*]. Even if this balance could be achieved, the delays through the different components are unlikely to have identical temperature coefficients, so that the balance cannot be maintained in normal operations.

Two-way time transfer can be further characterized by how transmissions in the two directions are related. A half-duplex system uses a single transmission path that is "turned around" to transmit in the reverse direction. A full-duplex system, on the other hand, transmits in both directions simultaneously, using either two nominally identical unidirectional channels, or some form of multiplexing on a single bi-directional path. All of these arrangements have advantages and drawbacks. The half-duplex system is more likely to be reciprocal, since the same path is used in both directions. However, that advantage is partially offset by the fact that fluctuations in the delay may degrade the time-delay measurement–especially if these fluctuations have periods on the order of the turn-around time. A full-duplex system is not degraded in this way, since the measurements in the two directions are more nearly simultaneous. However, incoherent fluctuations in the delays of the two unidirectional channels (or the cross talk between the two counter-propagating signals in the case of a single, bi-directional channel) may be an equally serious limitation.

The potential advantages of the two-way method must also be balanced against the fact that the hardware at the two end points must be nominally identical. Both ends must be active, and there is no real difference in the cost or complexity of the "transmitter" and "receiver." (Some savings in hardware can be achieved, if the signal received at one end is passively reflected–rather than actively transmitted–back along the path to the sending end. This advantage is likely to be offset by the poorer signal-to-noise ratio in the passively reflected beam and, possibly, by the need for an ancillary communications channel.) As we pointed out above, the symmetry of the two-way method generally is not compatible with broadcast-type services, which tend to be inherently asymmetrical and hierarchical.

Several proposals have been advanced for addressing this limitation. In one version of the idea, a time server transmits time signals, and also advertises its willingness to enter into two-way peer-type relationships for calibrating the delay to an individual client. The broadcast and two-way modes are supported simultaneously, by a single fast computer. Clients who do not need the increased accuracy that can be achieved by measuring the delay, or who are unable to invest in the overhead needed to determine it, can simply use the broadcast messages "as is." Those who want measured-delay mode use the broadcast messages to discover the name and address of the server, and to make a preliminary estimate of their time offset. These ideas are only being studied now, but something like them is likely to be developed in the next few years, for transmitting time over digital networks.

Two-way methods can be used in many different hardware environments. *Kihara and Imaoka* [*1993*] realized frequency synchronization of better than 10^{-13}, and time synchronization of better than 100 ns, using the two-way method in a 2000 km synchronous-digital hierarchy, based on optical fibers. Even higher accuracy has been realized using communications satellites to transfer time between timing laboratories [*DeYoung et al., 1994*; *Klepczynski et al., 1994*], and using dedicated optical-fiber circuits [*Primas et al., 1989*].

3.5 Common-view time transfer

The third commonly-used method is called common-view. Two receivers, which are approximately equidistant from a timing source, receive the same time information approximately simultaneously. The two receivers can be synchronized with an uncertainty that is proportional to the difference between the two paths back to the transmitter. Common-mode fluctuations in the two paths, and fluctuations in the timing source itself, cancel out (at least in first order).

Common-view techniques lend themselves naturally to broadcast services, since many receivers can observe a single transmitter. They have been an important technique in time-transfer for many years. Early systems used common-view observations of television signals, Loran stations, and similar transmissions. There have even been proposals to use local power-distribution systems for common-view time transfer within a single distribution region. In all of these cases, the common-view transmitter is actually designed for some other purpose, and it does not even need to "know" that it is being used for time distribution.

3.6 Multiple-frequency dispersion

If the transmission delay is dispersive (that is, if it depends on the frequency of the carrier used to transmit the information), and if the dispersion is separable into a part that depends on the carrier frequency, and a part that is a function of the physical parameters of the path, then it is often possible to estimate the transit time of the timing information along the path by measuring the dispersion: the apparent difference in the arrival times of the same information, transmitted using two different carrier frequencies. This difference is parameterized as a function of the same parameters that influence the delay itself, so that these parameters can be estimated and used to correct the delay.

There are a number of practical problems with this idea, but the most serious is the fact that the dispersion is usually quite a bit smaller than the delay itself. It must therefore be measured with higher accuracy than would be necessary if the delay could be measured directly. Since the method depends on special properties of the transmission medium, it is difficult to characterize its limitations in general terms.

In addition to the relatively poor signal-to-noise ratio that tends to characterize dispersion measurements, the signals are often subjected to a form of "chromatic aberration," which limits the usefulness of the method. If the medium is dispersive and spatially inhomogeneous (the atmosphere at optical wavelengths, for example), then the signals at the two wavelengths diverge in space, in addition to being differentially delayed because of refraction at each inhomogeneity. If this divergence becomes large enough that the two signals actually sample regions with different delays, then the method breaks down in a fundamental way, since the differential delay is due to a different value for the index of refraction, and not to dispersion. These problems are less serious at sub-optical frequencies, since the longer-wavelength signals tend to average these inhomogeneities.

3.7 Limitations of the methods

Each of these methods can be used to measure the delay, and each will be limited by unmodeled time fluctuations that violate its principal assumption: in the portable-clock method, it is fluctuations in the clock itself; in the two-way method, it is fluctuations in the reciprocity of the path; and in the common-view method, it is incoherent fluctuations in the two, physically separated one-way paths. The multiple-wavelength method is very powerful, but it can be used only in special situations, as discussed above, and its limitations depend on the details of the transmission medium.

4. Frequency distribution

A large class of systems need stable frequency references, while a much smaller class requires highly accurate frequency information that is traceable to the fundamental definition of the SI frequency [*BIPM, 1987*; *Guinot, 1989*]. An example of the former is a synchronous-data network, the operation of which depends on the fact that all of its nodes operate at the same clock rate (or perhaps at well-defined multiples of a single rate). An example of the latter might be an experiment to test the quantitative predictions of quantum mechanics [*Petley, 1991*] or relativity [*Vessot, 1991*; *Taylor, 1991*].

Any distribution method that can transmit periodic events can function as a frequency-distribution network, but "pure" frequency-distribution networks are rare. This is because the comparison between the distributed signal and the local clock that is to be evaluated is usually made in terms of phase (that is, time) differences, with some averaging or integration interval. There are usually sound statistical reasons for doing it this way. The phase comparisons are usually limited by noise that is approximately white: the effect of this white phase noise on the frequency determination can be attenuated by increasing the averaging time. As the averaging time is increased, the uncertainties in the phase measurements at each end of the interval are divided by a larger and larger time interval and, therefore, have less effect on the frequency determination.

As with time distribution, distributing frequencies that must be traceable to the SI definition has important consequences for the transmitting station. The accuracy of doing this may be limited by how accurately the frequency of the transmitter is known, relative to the SI standard. The need for traceability does not really impose any special requirements on the channel, beyond the requirement of adequate delay stability, which is basic to all frequency-distribution systems.

The primary standards that are currently used to realize the SI frequency have fractional uncertainties on the order of 10^{-14} [*Drullinger et al., 1993*; *Bauch et al., 1993*], and designs which may support even higher accuracies are under study (see [*Michaud et al., 1993*; *Clairon et al., 1995*]). Transmitting this frequency to a user without degrading its accuracy is a complex matter. Most of the high-accuracy methods that are currently available (GPS or two-way methods using communications satellites) are really time-transfer systems, so that the accuracy with which a frequency can be transmitted depends on the unmodeled temporal fluctuation in the path delay, as discussed above. If the frequency is to be transmitted at its full accuracy with an averaging time of 1 day, for example, then the transmission delay (and all other systematic errors) must be kept constant within 1 ns or less. This is a non-trivial task at the present time, and longer averaging times are generally used. Increasing the averaging time raises the real possibility that non-white fluctuations in the channel delay will begin to become important. The accuracy of the distribution process may no longer improve with increased averaging time, when this starts to happen.

These considerations suggest that there is often an optimum averaging time for transmitting frequency information. Shorter averaging times are not as good, because the effects of white phase noise on the time-difference measurements at each end point have not been averaged as well as they could be. Measurements using longer averaging times are degraded, because non-white processes begin to dominate the noise spectrum of the channel.

5. Delay measurements in digital and virtual circuits

Everything that we have said so far, about measuring the transmission delay, assumed implicitly that the path between the transmitter and receiver had the well-defined characteristics that we intuitively associate with a physical circuit, or a well-defined path through the atmosphere. Even though our measurements of the delay show that it is variable, we think of the delay as having a well-defined value, with the variability assigned to measurement noise. These ideas are still applicable to most atmospheric paths, but transmission circuits are increasingly packet-switched rather than circuit-switched. This means that there is no end-to-end physical connection. Packets are sent from the transmitter to the receiver over a common high-bandwidth physical channel, which is shared among many simultaneous connections using time-domain multiplexing (or some other equivalent strategy). Each packet passes through a series of routers, which read the destination address, and direct the packet accordingly. (The circuit is termed virtual for this reason. Data appear to enter at one end and emerge at the other, but there is no actual connection between the two. See [*Maxemchuk and El Zarki, 1990*]). The queuing delays in these routers are load-dependent and unpredictable. In an extreme case, consecutive packets may travel over very different paths, because of congestion or a hardware failure. This variability frustrates any attempt to measure the effective delay of the circuit.

It is reasonable to suppose that the queuing and routing variability always increase the delay over some baseline "true" value. Thus, the shortest apparent delays would be closest to this "truth." They would also have smaller variability, since they waited on fewer queues on the average [*Mills, 1989*]. There is considerable evidence that this is a useful strategy on the average, and the Network Time Protocol uses a variation on this idea for choosing the source of time information [*Mills, 1993*].

The time-division multiplexing associated with packet protocols is implemented in a logical sense: the underlying physical circuit may or may not support more than one simultaneous virtual connection. If more than one connection is supported, then there is time-domain multiplexing in the hardware as well. In the simplest form of this arrangement, a single bit from each circuit is combined in a multiplexor to form a message. The bits are unpacked at the other end, using a de-multiplexor running at the same rate [*Pan, 1972*]. More complex systems are commonly used, but most of them use the basic round-robin-sampling scheme we have outlined. This technique can work only if all of the inputs and outputs are running at the same speed, and if the sampling rates of the multiplexor and de-multiplexor are an exact multiple of this base speed. This is difficult to achieve in practice, and arrangements of this type usually are plesiochronous (almost synchronous), rather than absolutely synchronized [*Kartaschoff, 1991*]. The short-term differences in the input and output rates at a circuit element are usually absorbed by storage elements called "slip-buffers." The delays through these buffers are variable, of course, and they are likely to vary widely between different circuits–even between two adjacent circuits operating in opposite directions. These slip-buffers degrade both common-view and two-way-delay estimates of the delay.

These problems are likely to become more important as multiplexed digital circuits become more widely used. It is in the interest of the common carriers to minimize the delay fluctuations introduced by these processes, and to make the spectrum of whatever is left as white as possible, so that it can be attenuated through averaging. It is tempting to assert that the absolute magnitudes of these problems will be smaller on higher-speed channels, but this is by no means a sure thing.

6. Using received timing information

In the simplest distribution system, the receiver is a simple, "dumb" device. Its internal state (whether time or frequency) is simply reset whenever a signal is received. The signal may be corrected for transmission delay, or some other systematic correction might be used, but the important point is that the previous state of the receiver makes little or no contribution to the process. (In this simplest case, "integrating" and "filtering" the data that are received does not really change this argument. The filtering procedures are usually aimed at attenuating channel and measurement noise, rather than at combining the received and local data in a statistically robust manner.) This may be the best strategy, if the uncertainty in the receiver state is so high as to make it effectively worthless, but it is important to recognize that the state of the receiver before the new information is received also contains information that may be useful.

The performance of the receiver clock can be characterized in terms of deterministic and stochastic components. The deterministic aspects of its performance can be used to predict the data that will be received from the transmitter at some future time. This prediction is never perfect, of course, because the deterministic components may not be known precisely; because even if they are known at one instant, they are likely to evolve in time; and because the stochastic components of the model are always present, and add noise to all measurements. Although this prediction might not be as accurate as we would like, it nevertheless can often provide an important constraint on the received data.

The most useful situation arises when the spectrum of the noise in the receiver clock has little or no overlap with the spectrum of the noise in the transmission channel. This happy situation supports the concept of "separation of variance:" the ability to distinguish between a "true" signal, which should be used to adjust the state of the receiver clock, and a "false" signal, which is due to an unmodeled fluctuation in the channel delay, and which should therefore be ignored. In an extreme situation, this strategy could be used to detect a failure in the transmitter clock, itself.

Decisions based on separation of variance are inherently probabilistic (rather than deterministic) in nature, and will only be correct in an average sense, at best. Whether or not they improve matters in a particular situation depends on the validity of the assumption of separation of variance, and on the accuracy of the various variance estimators. Since the estimators must be based on the data themselves, a certain amount of ambiguity is unavoidable.

This method obviously fails in those spectral regions where separation is impossible. There is no basis for deciding whether the observed signal is "real," or represents transmission noise and should be ignored. If the resulting ambiguity results in unacceptably large fluctuations, then either the receiver or the channel must be improved. The crucial point is that neither the receiver clock nor the transmission channel need be good everywhere in the frequency domain: it is perfectly adequate if they have complementary performance characteristics, and this may, in fact, be the "optimum," least-cost alternative for achieving a given accuracy in the distribution of time or frequency.

These considerations explain why it may not be optimum to use the signal received over a noisy channel "as is." It may be possible to remove the channel noise, if the received data can be used to discipline an oscillator whose inherent noise performance can be characterized by a variance that separates cleanly from the noise added by the channel, in the sense we have discussed. These ideas can be used to filter the intentional degradations imposed on the GPS satellites because of Selective Availability [*Kusters et al., 1994*], and to improve the distribution of time on wide-area networks, such as the Internet [*Levine, 1995*].

7. Authentication of signals

It is an unfortunate fact of life that things that have value have some probability of being stolen or falsified, and there is no reason to think that time and frequency information are any exception. Two possible classes of attacks are those which deny service (jamming a radio signal is a trivial example), and "spoofing," where a bogus source masquerades as an authoritative source of time or frequency information. Denial-of-service attacks are less serious, in principle, because they are relatively easily detected, and can often be thwarted using parallel channels or multiple, separated sources. Spoofing may not be nearly so easily detectable or correctable.

Public-packet networks, such as the Internet, are particularly susceptible to spoofing, because they tend to have widely distributed administration, and little centralized authority. Unlike jamming (and its Internet analog), which is most effective when the source of the trouble is nearby, a bogus time server need not be local. The Network Time Protocol (NTP) supports access control via the NBS/NIST Data Encryption Standard (DES) [*NIST, 1977*]: when authentication is enabled, a machine will accept time information only if it is encrypted with the proper DES key. GPS signals include a sophisticated encryption system for the same reasons.

Symmetric encryption systems, like DES, are not necessarily optimum for this application. Since the encryption and decryption keys are the same, the security of the keys, and their management and distribution, are difficult problems, whose complexity grows very rapidly as the number of participating stations increases. A public-key algorithm is better suited to this task, in principle, since the decryption key can be made public, without revealing the encryption value. These algorithms are much slower than DES, however, and the delay needed to compute the encryption may be a significant bottleneck on a busy system. These delays are likely to become less serious as hardware speeds increase, or if the entire algorithm is realized in a special-purpose device. Several proposals are currently under study for improvements in all network services, and these will also be helpful in addressing these issues.

8. Time-stamp algorithms

There are many situations where it is important to be able to prove that a document existed on or before a certain date and time. (In principle, the time stamp applied to any document should be UTC, but this process does not really demand extremely high accuracy, and the UTC time scale realized by a recognized national timing center is likely to be more than adequate for the job.) Examples include patent disclosures, and commercial contracts where time is a factor. This need has been addressed by Notaries Public, who function as disinterested third parties to these transactions. This system is somewhat cumbersome, but was adequate when the document was a piece of paper. It is not well suited to documents, such as computer files, that exist solely (or perhaps primarily) in digital format. By analogy with a time stamp and signature on a physical document, we would like the digital version to have the following characteristics:

> 1. The signature, the time stamp, and the document should be bound together, in the sense that it should be computationally infeasible to alter the document or the time stamp without invalidating the signature, or to re-use either the time stamp or the signature string on another document [*Nechvatal, 1991*].

This requirement can be achieved by adding a time stamp to the document, and then "hashing" the combination. A number of hashing techniques have been discussed in the literature [*NIST, 1993*]. While they differ in detail, they all compute a large number (the hash value) by passing the input text through a complex, nonlinear transformation. The nonlinearity of the procedure insures that even the smallest change in the document will produce a very large change in the hash value, and that the original hash value can not be restored using an iterative approach. The size of the hash value is intended to make an exhaustive-search attack computationally infeasible.

> 2. The signature should be able to be independently verified as coming from a trusted source.

One way of realizing this requirement is to use public-key cryptographic techniques, in which the hashed value computed from the document is encrypted, with the output of this process acting as the signature [*NIST, 1994*]. As with hashing, there are a number of techniques in the literature, but they all are based on using two numerical keys: a "secret" key that is used to construct the signature, and a "public" key that is used to verify it [*Schneier, 1993*]. These keys are large numbers, to prevent exhaustive-search attacks, and it is computationally infeasible to compute the private key from the information that is publicly available.

In addition to adding a time stamp to each document, it is also possible to establish a time-ordered chain of documents, by incorporating some part of the signature from each document into the text of the next document that is processed [*Haber and Stonetta, 1991*; *Cipra, 1993*]. The hash value of the modified document is then computed, as above, and the result is signed. This procedure effectively binds all of the documents together in a "chain," the links of which form a time-ordered sequence that is independent of the accuracy of the individual time stamps. A number of variations of this idea have been suggested (see [*Haber and Stornetta, 1991*] and the references therein).

The authentication of digital documents is one of the first services that is made possible by a combination of improvements in the distribution of time in digital format, and the wide-spread availability of powerful, moderately-priced computers. It is likely that more applications of this type will be developed in the near future.

9. Summary, and a look to the future

The principal limitation to the accurate distribution of time and frequency is the measurement of the transit time of the information from the source to the user. As we have pointed out, this is not a trivial undertaking. Transit times–which may be tens of milliseconds or longer–must be measurable, and stable to better than a nanosecond, if the accuracy inherent in the best contemporary standards is to be transmitted to a user without appreciable degradation. The parameters that affect the transit time must also be understood well enough so that this level of stability can be achieved over long periods of time, which is perhaps even more difficult.

We have discussed a number of methods that are used to estimate this transit time. All of these methods are currently used in disseminating time and frequency, and some systems use more than one method simultaneously. The GPS system, for example, is often used in common-view mode, and its dual-frequency-signal structure is designed to facilitate an estimate of the ionospheric refractivity, using the observed dispersion between the L1 and L2 signals. The tropospheric refractivity, on the other hand, usually cannot be estimated from the dispersion. The effect is very small, to begin with, and is not completely separable, because of the presence of a term due to tropospheric water vapor. Various models have been used to estimate the tropospheric refractivity [*Wells, 1987*], using data from balloon measurements and water-vapor radiometers, but none has proven completely successful. Corrections for the tropospheric index of refraction are already widely used in geodetic measurements using GPS, and they are likely to be necessary in time transfer, in the future.

For users who are more interested in frequency than in time, the ideal dissemination system would be one whose fluctuations could be characterized as white-phase noise at all averaging times. The uncertainty in the frequency transmitted by such a system can be made as small as desired, by increasing the averaging time until some other process (such as the performance of the clock itself) becomes important. No currently operating system satisfies this condition, and developing such a system will be one of the goals of the near-term future. Although it may be possible to improve the characteristics of the transmission media, further decreases in the uncertainty of the transit time may only be realized with more-sophisticated averaging schemes, perhaps including extensive post-processing.

At least in the short term, dissemination will probably depend on the two-way and common-view methods. Both of them are inherently symmetrical, and do not single out either participant as client or server. Any system based on these methods is able to support the peer concept that is likely to be the optimum way of interacting with customers with very high-quality-clock hardware. Common-view GPS [*Allan et al., 1985*] would seem to have an initial advantage, in terms of the number of sites that can participate in a simultaneous-measurement campaign, but this advantage may not be fully realized, in practice, because of the intentional degradations of the GPS signal. A two-way system using communications satellites is less likely than a common-view GPS system to be limited by fluctuations in the atmospheric component of the transit delay, but this advantage may be lost due to time-varying asymmetries in the hardware [*Hanson, 1989*; *Saburi, et al., 1976*].

Another trend is the increasing demand for moderately accurate time in various digital formats. These uses do not tax the state of the art from the point of view of accuracy, but satisfying the sheer volume of the requests, and their great diversity, requires some careful planning. Important issues for these users are often cost, reliability, ease of use and, possibly, legal traceability. The accuracy of the signal transmitted by WWV, for example, is adequate for many users. However, receiving the signal reliably often requires an outside antenna, and extracting the time in a digital format [*Beehler and Lombardi, 1991*] requires something more than a bottom-of-the-line receiver. A telephone time service, like the NIST Automated Computer Time Service, may be easier than WWV to install, and more reliable to operate, but each use requires a telephone call to the server. A user with many systems to synchronize may be faced with appreciable telephone costs, and a significant investment in server hardware will be required to provide an adequate level of service. A single ACTS server can handle perhaps 1500 calls/day; more than 500 servers would be needed if only 1% of the estimated 10^8 PC users used a service like ACTS only once per day. Similar scaling arguments apply to almost any publicly-available service that attempts to estimate the transit time of the signal along the path to the users. (The cost of running a radio service like WWV obviously does not depend on the number of users, but the path delay can only be estimated using tables of the average propagation characteristics at any site.)

The load on the servers could be reduced with better client oscillators that required less-frequent synchronization and more sophisticated use of the calibration data when they are acquired. The NIST Internet Time Service is a first attempt in this direction: it has demonstrated performance accuracies substantially better than 100 ms, using less than one calibration message per day. A single server can handle at least 20 000 requests per day, so that the required server hardware increases much more slowly with heavy use than would be true for an ACTS-type system. Although ACTS can deliver time signals more accurately than the Internet, the difference is not important, in many applications.

Both ACTS and the Internet time service are based on the two-way principles outlined above. The ACTS system is more accurate, because its signals travel over a dedicated telephone circuit that is more-easily characterized. Several Internet methods are based on a combination of two-way and common-view, in an attempt to compensate for the relatively poor characteristics of the Internet. Improvements here are quite likely, both in the speed of the network itself, and in the sophistication of the software at each client node.

Authentication and anti-spoofing are issues that are likely to become more important in the near future, as time stamps become more important, and their use in various digital formats increase. Several proposals have been advanced for adding authenticating digital signatures to messages transmitted over wide-area computer networks. Various public-key encryption systems, such as the method currently used for GPS transmissions, have been proposed for authenticating messages sent by other means. None of these methods is generally available or particularly easy to use. It is quite possible that simplicity and robustness will continue to be conflicting goals, and there may be no alternative to systems that are complex and somewhat awkward to use, as a result.

10. References

D. W. Allan, D. D. Davis, M. Weiss, A. Clements, B. Guinot, M. Granveaud, K. Dorenwendt, B. Fischer, P. Hetzel, S. Aoki, M. K. Fujimoto, L. Charron, and N. Ashby [1985], "Accuracy of International Time and Frequency Comparisons Via Global Positioning System Satellites in Common-View," *IEEE Trans. Instrum. Meas.*, **IM-34**, pp. 118-125.

D. W. Allan, H. Hellwig, P. Kartaschoff, J. Vanier, J. Vig, G. M. R. Winkler, and N. Yannoni [1988], "Standard Terminology for Fundamental Frequency and Time Metrology," *42nd Annual Symposium on Frequency Control*, June 1-3, 1988, Baltimore, MD, USA, IEEE Catalogue No. 88CH2588-2 (available from IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA), pp. 419-425.

J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, Jr., W. L. Smith, R. L. Sydnor, R. F. C. Vessot, and G. M. R. Winkler [1971], "Characterization of Frequency Stability," *IEEE Trans. Instrum. Meas.*, **IM-20**, pp. 105-120.

A. Bauch, B. Fischer, T. Heindorff, R. Schroder, [1993], "The New PTB Primary Cesium Clock," *IEEE Trans. Instrum. Meas.*, **IM-42**, pp. 444-447.

R. E. Beehler and M. A. Lombardi [1991], *NIST Time and Frequency Services*, NIST Special Publication 432 (Revised 1990) (available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC 20402-9325, USA).

BIPM, Bureau International des Poids et Mesures [1991], *Le Système International d'Unités*, 6th edition (available from Director, BIPM, Pavillon de Breteuil, F-92310 Sèvres, France).

B. Cipra [1993], "Electronic Time-Stamping: The Notary Public Goes Digital," *Science*, **261**, pp. 162-163.

A. Clairon, Ph. Laurent, G. Santarelli, S. Ghezali, S. N. Lea, and M. Bahoura [1995], "A Cesium Fountain Frequency Standard: Preliminary Results," *IEEE Trans. Instrum. Meas.*, **IM-44** (in press).

J. A. DeYoung, W. J. Klepczynski, J. A. Davis, W. Powell, P. R. Pearce, C. Hackman, D. Kirchner, G. de Jong, P. Hetzel, A. Bauch, A. Soering, P. Grudler, F. Baumont, H. Ressler, L. Veenstra [1994], "Preliminary Results from the 1994 International Transatlantic Two-Way Satellite Time and Frequency Transfer Experiment," *Proceedings of the 26th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, December 5-8, 1994, Reston, VA, USA, pp. 39-50 (available from US Naval Observatory, Time Service, 3450 Massachusetts Ave., Washington, DC 20392-5420, USA; also available as NASA Conference Publication 3302, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA).

R. E. Drullinger, J. H. Shirley, J. P. Lowe and D. J. Glaze [1993], "Error Analysis of the NIST Optically Pumped Primary Standard," *IEEE Trans. Instrum. Meas.*, **IM-42**, pp. 453-456.

A. Gifford and T. Bartholomew [1994], "An Argument for Independent Calibration of GPS Time Transfer," presented at the 26th Annual Precise Time and Time Interval Applications and Planning Meeting, December 5-8, 1994, Reston, VA, USA; private communication.

B. Guinot [1989], "Atomic Time," *Reference Frames in Astronomy and Physics*, in J. Kovalevsky, I. I. Mueller and B. Kolaczek (eds.), New York, Kluwer, pp. 379-415.

S. Haber and W. S. Stornetta [1991], "How to Time-Stamp a Digital Document," *J. Cryptology*, **3**, pp. 99-111.

D. W. Hanson [1989], "Fundamentals of Two-Way Time Transfers by Satellite," *43rd Annual Symposium on Frequency Control*, May 31-June 2, 1989, Denver, CO, USA, IEEE Catalogue No. 89CH2690-6 (available from IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA), pp. 174-178.

IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology [1988], *IEEE Std 1139-1988* (available from IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA).

J. Jespersen [1989], "Impact of Atmospheric Non-Reciprocity on Satellite Two-Way Time Transfers," *43rd Annual Symposium on Frequency Control*, May 31-June 2, 1989, Denver, CO, USA, IEEE Catalog No. 89CH2690-6 (available from IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA), pp. 186-192.

P. Kartaschoff [1991], "Synchronization in Digital Communications Networks," *Proc. IEEE*, **79**, pp. 1019-1028.

M. Kihara and A. Imaoka [1993], "Accurate Network Synchronization for Telecommunications Networks Using Paired Paths", *Proceedings of the 7th*

European Frequency and Time Forum, March 16-18, 1993, Neuchâtel, Switzerland (available from EFTF Secretariat, FSRM, Rue de l'Orangerie 8, CH-2000, Neuchâtel, Switzerland), pp. 83-88.

D. Kirchner [1991], "Two-Way Time Transfer via Communication Satellites," *Proc. IEEE*, 79, pp. 983-990.

W. J. Klepczynski, J. A. Davis, K. Kirchner, G. de Jong, F. Baumont, P. Hetzel, A. Soering, C. Hackman, M. Granveaud, W. Lewandowski [1994], "Comparison of Two-Way Satellite Time Transfer and GPS Time Transfer During the Intelsat Field Trial," *Proceedings of the 26th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, December 5-8, 1994, Reston, VA, USA, pp. 89-94 (available from US Naval Observatory, Time Service, 3450 Massachusetts Ave., Washington, DC 20392-5420, USA; also available as NASA Conference Publication 3302, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA).

J. A. Kusters, R. P. Giffard, L. S. Cutler and D. W. Allan [1994], "A Globally Efficient Means of Distributing UTC Time and Frequency Through GPS," *Proc. of the 26th Annual Precise Time and Time Interval Applications and Planning Meeting*, December 5-8, 1994, Reston, VA, USA, pp. 235-245 (available from US Naval Observatory, Time Service, 3450 Massachusetts Ave., Washington, DC 20392-5420, USA; also available as NASA Conference Publication 3302, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA).

J. Levine, M. Weiss, D. D. Davis, D. W. Allan, and D. B. Sullivan [1989], "The NIST Automated Computer Time Service," *J. Res. Natl. Inst. Stand. Tech.*, 94, pp. 311-321.

J. Levine [1995], "An Algorithm to Synchronize the Time of a Computer to Universal Time," *IEEE/ACM Trans. Networking*, 3, pp. 42-50.

W. Lewandowski and C. Thomas [1991], "GPS Time Transfer," *Proc. IEEE*, 79, pp. 991-1000.

N. F. Maxemchuk and M. El Zarki [1990], "Routing and Flow Control in High-Speed Wide-Area Networks," *Proc. IEEE*, 78, pp. 204-221.

A. Michaud, M. Chowdhury, K. P. Zetie, C. J. Cooper, G. Hillenbrand, V. Lorent, A. Steane and C. J. Foot [1993], "Realization of a Frequency Standard Using an Atomic Fountain," *Proceedings of the 7th European Frequency and Time Forum*, March 16-18, 1993, Neuchâtel, Switzerland (available from EFTF Secretariat, FSRM, Rue de l'Orangerie 8, CH-2000, Neuchâtel, Switzerland), pp. 525-530.

D. L. Mills [1989], "Measured Performance of the Network Time Protocol in the Internet System," *Network Working Group*, Request for Comments No. 1128 (available on the Internet via ftp from nis.nsf.net or nic.ddn.mil; for more information contact Government Systems, Inc., Suite 200, 14200 Park Meadow Drive, Chantilly, VA 22021, USA), pp. 9-11.

D. L. Mills [1993], "Network Time Protocol (Version 3); Specification and Implementation," *Network Working Group*, Request for Comments No. 1305 (available on the Internet via ftp from nis.nsf.net or nic.ddn.mil; for more information contact Government Systems, Inc., Suite 200, 14200 Park Meadow Drive, Chantilly, VA 22021, USA), pp. 22-37.

NIST, National Institute of Standards and Technology [1977], *Announcing the Data Encryption Standard*, Federal Information Processing Standard No. 46 (available from National Technical Information Service, Springfield, VA 22161, USA).

NIST, National Institute of Standards and Technology [1993], *The Secure Hash Standard*, Federal Information Processing Standard No. 180 (available from National Technical Information Service, Springfield, VA 22161, USA).

NIST, National Institute of Standards and Technology [1994], *The Digital Signature Standard*, Federal Information Processing Standard No. 186 (available from National Technical Information Service, Springfield, VA 22161, USA).

J. Nechvatal [1991], *Public-Key Cryptography*, NIST Special Publication 800-2, (available from Superintendent of Documents, US Government Printing Office, Washington, DC 20402-9325, USA), pp. 25-35.

J. W. Pan [1972], "Synchronizing and Multiplexing in a Digital Communications Network," *Proc. IEEE*, **60**, pp. 594-607.

D. B. Percival [1991], "Characterization of Frequency Stability: Frequency Domain Estimation of Stability Measures," *Proc. IEEE*, **79**, pp. 961-973.

B. W. Petley [1991], "Time and Frequency in Fundamental Metrology," *Proc. IEEE*, **79**, pp. 1070-1076.

L. Primas, R. Logan, Jr., G. Lutes [1989], "Applications of Ultra-Stable Fiber Optic Distribution Systems," *Proceedings of the 43rd Annual Symposium on Frequency Control*, May 31-June 2, 1989, Denver, CO, USA, IEEE Catalog No. 89CH2690-6 (available from IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA), pp. 202-212.

T. J. Quinn [1991], "The BIPM and the Accurate Measurement of Time," *Proc. IEEE*, **79**, pp. 894-905.

J. Rutman and F. L. Walls [1991], "Characterization and Measurement of Frequency Stability," *Proc. IEEE*, **79**, pp. 952-960.

Y. Saburi, M. Yamamoto, K. Harada [1976], "High Precision Time Comparisons via Satellite and Observed Discrepancy of Synchronization," *IEEE Trans. Instrum. Meas.*, **IM-25**, pp. 473-477.

B. Schneier [1993], *Applied Cryptography*, New York, John Wiley & Sons, Inc., Chapter 12.

J. H. Taylor, Jr. [1991], "Millisecond Pulsars: Nature's Most Stable Clocks," *Proc. IEEE*, 79, pp. 1054-1062.

R. F. C. Vessot [1991], "Applications of Highly Stable Oscillators, to Scientific Measurements," *Proc. IEEE*, 79, pp. 1040-1053.

D. Wells (ed.) [1987], *Guide to GPS Positioning*, Fredericton, New Brunswick, Canadian GPS Associates.

P. Wheeler, D. Clamers, A. Davis, P. Koppang, A. Kubig and W. Powell [1994], "High Accuracy Time Transfer Synchronization," *Proceedings of the 26th Annual Precise Time and Time Interval Planning and Applications Meeting*, December 5-8, 1994, Reston, VA, USA, pp. 51-62 (available from US Naval Observatory, Time Service, 3450 Massachusetts Ave., Washington, DC 20392-5420, USA; also available as NASA Conference Publication 3302, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA).