The Role of Atomic Clocks in Data Centers: How the Atom went from Data's Worst Enemy to its Best Friend

By David Chandler, Senior Product Marketing and Applications Manager for the Frequency and Timing Systems business unit at Microchip Technology

Timing from atomic clocks is now an integral part of data center operations. The atomic clock time transmitted via Global Position System (GPS) and other Global Navigation Satellite System (GNSS) networks is synchronizing servers across the globe and atomic clocks are deployed in individual data centers to preserve synchronization when the transmitted time is not available. This high level of synchronization is vital to ensure the zettabytes of data collected around the globe every year can be meaningfully stored and used in many applications, whether due to system requirements or regulatory compliance. The quantum nature of an atom enables the precision time and is a critical part of ensuring more data at faster speeds will be processed in the future, which is ironic as the quantum nature of the atom was seen as the ultimate death of this increase in data processing and speed just a few years ago.

In 1965 Gordon Moore predicted the transistor count on an integrated circuit would double every year [1]. This was eventually revised to doubling every two years. Along with this increase in transistor density came an important increase in speed as well as decreases in cost and power consumption. It may have been hard in 1965 to imagine there would be any real world need to have a semiconductor with 50 billion transistors on it in 2021 [2], but as semiconductor technologies kept up with the law, so did application demands. Cell phones, financial trading and DNA mapping are all applications that rely heavily on the number of operations per second a microprocessor can execute, which is closely tied to the transistor count on a chip.

Figure 1: Satirical image of engineer trying to keep up with Moore's Law

Unfortunately, Moore's Law is rapidly coming to an end due to a limit imposed by physics. With wafer fabrication now in the sub 10 nm technology nodes [3], the transistor sizes are only about ten to fifty times bigger than the diameters of a silicon atom. At this scale the size and quantum properties of atoms and free electrons significantly prohibit further size reduction. In essence, you could think of the atom as the ultimate court that struck down the law. But while Moore's Law will come to an end, the thirst for increased processing power continues to grow. With the advent of the Internet of Things (IoT), streaming services, social media posts and autonomous self-driving cars, the amount of data generated every day continues to exponentially increase. In 2021, an estimated 2.5 exabytes (2,882,303,761,517,120,000 bytes) was generated every day [4]. Exabyte databases managing more than 100,000 transactions per second (a transaction consists of multiple operations) are currently in use, and the size of the databases and the transactions per second will continue to grow for the foreseeable future.

This explosive growth in the volume of data, coupled with the speed the data must be written, read, copied, analyzed, manipulated, and backed up, required data center architects to find a way around the end of Moore's Law. The architects employed horizontal scaling in a data center with distributed databases, where instead of an entire database residing on one server, the database is distributed over multiple servers in a cluster. In this configuration the cluster essentially functions as one giant machine, hence the size and speed of the system now becomes limited by the physical size of a data center rather than size of an atom (take that atom!). Software engineers now make careers writing code that enables horizontal scaling. But for all the software to work, all the machines must be synchronized. If they aren't, violations in a concept called causality occurs.

What is causality? It is easiest to explain through an example. Suppose you have two cameras to record images for a 100-meter dash, each with its own internal clock. The first camera is at the starting blocks. The second camera is at the finish line. Both sensors are continually firing and stamping each image with the time from their respective clocks, a process referred to as time-stamping. To determine the official time of the winning sprinter in the race, the first camera's images are reviewed for the point in time the first runner left the block and this time-stamp is subtracted from the time-stamp on the last camera's image for the first runner crossing the finish line. For this to work, both cameras must be synchronized to an acceptable level of uncertainty. If the synchronization of the clocks is only ±0.05 seconds you would be unable to determine if someone who was recorded as running 9.6 seconds actually broke the world record of 9.58 seconds. What if they were only synchronized to ± 5 seconds from the stadium clock? Imagine this scenario. Observed from the main stadium clock, a race starts at exactly 12:00:00:00 pm. The first runner crosses the finish line at 12:00:09:60 pm. From the perspective of the main stadium clock the official race time was 9.6 seconds. But what if the first camera's clock was exactly 5 seconds fast and the second camera's clock was exactly 5 seconds slow? The race would officially start at 12:00:05:00 pm, and finish at 12:00:04:60 pm. The race would officially finish 0.4 seconds before it started, the world record would be shattered, the laws of physics would be broken, and the current record holder would most likely be wrongfully dropped by all his sponsors.

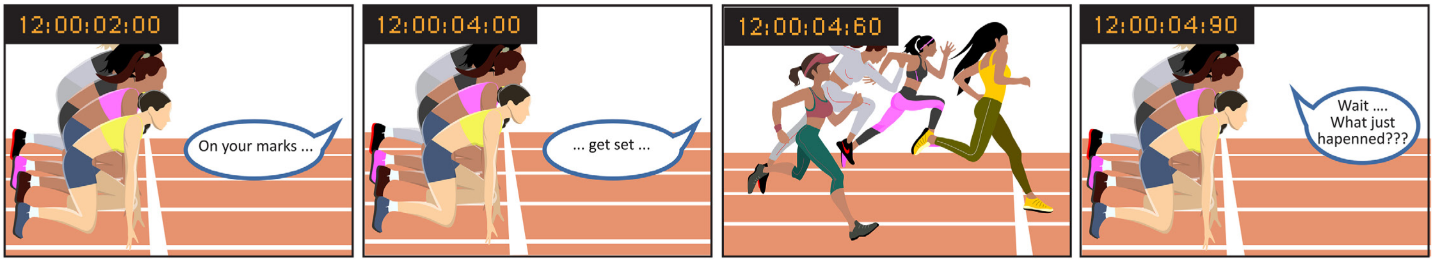

Figure 2: Clock uncertainty causes issues with causality. In this case a race officially finished before it started.

The same principle of causality is important in a database. Transactional record updates must appear in the database in the sequential order in which they occurred. If you counted on the direct deposit of your paycheck arriving prior to having a direct withdrawal to pay your monthly mortgage, and the bank's database didn't record these in the correct sequence, you are going to be charged an overdraft fee. On one machine causality errors are easy to prevent but on multiple servers, each with their own internal clock, the servers must be synchronized, and time-stamp every transaction. To achieve this, one server must act as a reference clock, much like the stadium clock, and it must be distributed to each server in a way that minimizes the time error of each server clock. The uncertainty of each time-stamp (± 5 seconds in the race) forms a time envelope that is twice the uncertainty of the clock (10 seconds for the race). For a distributed database, the number of nonoverlapping time-envelopes that can fit into a second should be at least on the order of the number of transactions per second expected for the system. Probability, criticality of causality, and cost of implementation will ultimately all play a role in the final solution, but this relationship is a good starting point. A system with time-stamp uncertainties of ± 1 millisecond, would have time-envelopes of 2 milliseconds, and a maximum of 500 non-overlapping time-envelopes would fit in one second. This system could support approximately 500 transactions per second. Time over ethernet technologies known as Network Time Protocol (NTP) and Precision Time Protocol (PTP) are used to synchronize all the servers in a distributed database in a data center. These protocols can ensure a local area network can distribute time with submillisecond (NTP) or submicrosecond (PTP) uncertainties, enabling thousands (NTP) or millions (PTP) of transactions per second.

Unfortunately, even with these solutions that enabled a detour around the atom-imposed demise of Moore's Law, physics has thrown another roadblock in the path of distributed databases in the form of the speed of light. Imagine a well-synchronized distributed database operating with PTP in San Jose California, happily executing 100,000 transactions per second with no causality issues. One of the database architects is sitting in his office in New York and his boss asks him to update a large series of records. The architect wants to be able to exploit his new database to its full extent and show off the system capabilities. He plans on executing 100,000 transactions per second. To update records per the request, he creates a simple transaction that adds the value of one record to a second record only if the value of the first record is greater than the second record. To accomplish this, he must issue a read to both records. His local machine in New York will then compare the values, then send a write command to the second record when needed. After completing this he then wants to execute the next transaction that compares a third value to the new sum. If the new sum is greater than the third record, then the third record is replaced with the sum. He wants to repeat this for 6 million records. Since the database is capable of 100,000 transactions per seconds, he thinks it will be done in roughly a minute. He tells his boss he will have the records updated in 5 minutes. He leaves to get a cup of coffee. While drinking his coffee, he reads a story about how the new 100-meter dash record is negative 0.4 seconds which defies the laws of physics, and that the previous record holder is suing the stadium officials because he has lost all his endorsement money. The architect laughs to himself and thinks the stadium should have hired him as the synchronization expert. He comes back to his desk 5 minutes later and is dismayed to see that his database update has completed less than 1500 transactions. He sadly realizes his mistake and prepares his resume to send it over to the stadium where he hopes his PTP deployment won't have the same problem.

What went wrong? The speed of light limits the theoretical fastest possible transmission of data between New York and San Jose to 13.7 milliseconds.

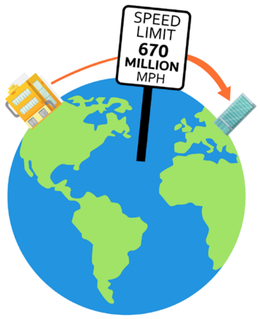

Figure 3: The speed of light imposes a theoretical limit to the speed data can be transferred between two points.

Unfortunately, real world numbers are even slower. Even with a dedicated fiber optic link between the two locations the refractive index of the fiber, the real-world path of the fiber, and other system issue make this transit time even slower. So just one transmission from New York will take 40 to 50 milliseconds to arrive in San Jose. However, in this transaction there are 4 unique operations. There are two read operations, which could happen in parallel, which then have to be sent back to New York. The round trip takes 80 to 100 milliseconds. Then, once both values are compared, a write operation is issued and a write acknowledgement must be sent back indicating the write operation completed before the next transaction can start. And suddenly it doesn't matter that the database can perform 100,000 transaction per second, because the distance is limiting the system to 5 transactions per second. To complete the 6 million transactions this system would take thirteen days, more than enough time for several more cups of coffee and to update a resume. This delay is referred to as communications latency.

But just like Moore's Law, database architects figured out how to circumvent latency. Database replications are created near the users, so that they can work with the data without having to send signals across the country. Periodically the replications are compared and reconciled to ensure consistency. During the reconciliation process, the transaction time-stamps are used to determine the actual sequence of transactions, and records are sometimes rolled back when there is an irreconcilable difference such as when the transactions time-envelopes overlap. Reducing clock uncertainty reduces the number of irreconcilable differences in replicated instances, as more time-envelopes reduce the probability of overlaps. This results in higher efficiencies and lower probabilities of data corruptions. But now the time-stamping has to be accurate not only within each data center, but also between the data centers, which can be separated by thousands of miles and connected via the cloud. This is a much more challenging task, as it requires an external reference with very low uncertainly that is readily available in both locations.

Enter the previous foe of the data base architect, the atom. While the atom was busy repealing Moore's Law, its subatomic particles were busy spinning. The neutrons and protons in the nucleus were rotating, while at the same time the electrons were busy orbiting about the nucleus, while also spinning on their own axes. This is analogous to earth orbiting around the sun while simultaneously spinning on its axis. The electrons can spin around their axes clockwise or counterclockwise. Considering there are roughly 7 octillion (7 with 27 zeros after it) atoms in a human [5], with all the subatomic particles spinning in our bodies, it is amazing we aren't permanently dizzy. (Note: The subatomic particles aren't really busy spinning and orbiting, they are really busy giving us probability wave functions and magnetic interactions that would give us results similar to if they were spinning and orbiting. But if the thought of all the spinning makes you dizzy, trying to comprehend the reality of quantum physics will make you positively nauseous.)

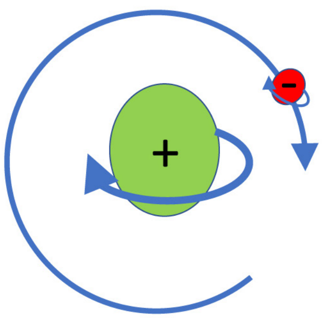

Figure 4: Conceptual atoms with nucleus and valence electron with nuclear spin, electron spin and orbital spin.

When microwave radiation at a very specific precise frequency is absorbed by an electron, the direction of spin about the electron axis can be changed. If this happened to the earth, the sun would suddenly set in the east and rise in the west! Atomic clocks are machines designed to detect the state of the electron spin, and then change that direction through microwave radiation. The frequency varies depending on the element, isotope, and the excitation state of the electrons. Once the machine determines the frequency, known as the hyperfine transition frequency, the period can be determined as the inverse of the frequency, and the number of periods can be counted to determine the elapsed time. The international definition of the second is defined by counting 9,192,631,770 periods of the radiation required to induce the hyperfine transition of an electron in the first orbital shell of a cesium atom [5].

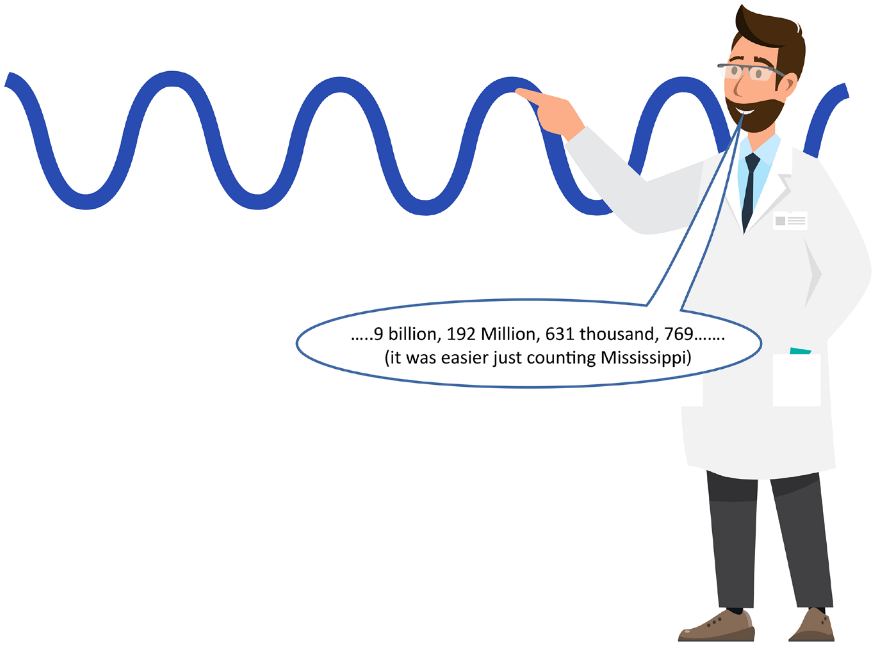

Figure 5: The unit second is defined by counting 9,192,631,770 cycles of the cesium hyperfine transmission radiation frequency.

Atomic clocks are the most stable commercially available clocks in the world. An atomic clock the size of a deck of cards called the chip scale atomic clock (CSAC) will drift a 1 millionth of a second in 24 hours, whereas an atomic clock the size of a refrigerator called a hydrogen maser will only drift 10 trillionths of a second in 24 hours. Coincidentally ten trillionths is also the approximate ratio of the radius of the hydrogen atom to the height of the sprinters in the 100-meter dash and the now unemployed data center architect in New York.

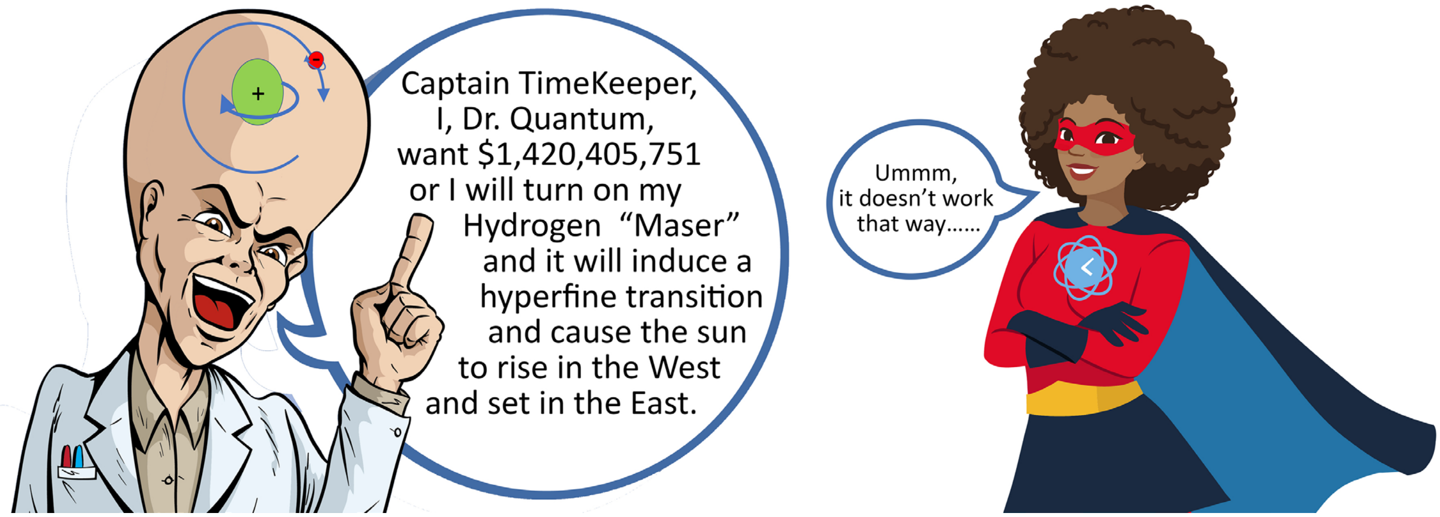

Figure 6 – The hyperfine transition frequency produced in a hydrogen maser, 1.420405751 GHz, will cause spin reversal in an electron.

With the accuracy provided by these atomic clocks, approximately 500,000 to ~ 50 billion nonoverlapping time-envelopes can be provided for a distributed database running in data centers in Tokyo, London, New York, Timbuktu or anywhere else in the world. But how does time get to all the data centers from these atomic clocks? Universal Coordinated Time (UTC) is a global time distributed by satellites, fiber optic networks, and even the Internet. UTC itself is derived from a collection of high precision atomic clocks located in national laboratories and timing stations around the world. Contributors to UTC receive a report that provides the UTC time from these clocks and their individual offset from calculated UTC. The labs and other facilities then transmit the time to the world.

The UTC report is published monthly and tells the national labs their miniscule timing offset from UTC during the previous month. Technically, we don't know precisely what time it was up until a month after the fact. And to make things worse, extra seconds are periodically added to UTC, called leaped seconds, which are inserted due to variations in the earth's rotation and our relative position to observable stars. While this aligns the earth to the universe, it causes havoc in data centers and 100-meter dashes.

Two common methods used by data centers to acquire UTC are via the internet using publicly available NTP time servers and via satellite using GPS or the GNSS networks. While timing through public NTP timeservers over the internet was common during early deployment of distributed databases, inherent performance, traceability and security issues have created the push to move away from this solution. Even though GPS and other GNSS satellite networks are typically thought of as positioning and navigation systems, they really are precision timing systems. Position and time at a receiver are determined by the transit time of signals traveling at the speed of light from multiple satellites to the receiver. Ironically this is another case of a physics principle causing a problem, in this case the speed of light instead of the atom, but also contributing to the solution. The satellites have their own onboard atomic clocks, which are synchronized to UTC that was transmitted to the satellites from ground stations. Acquiring UTC with this method can provide time uncertainties in the 5 nanosecond range, providing 100 million time-envelopes per second. This method is far more reliable and accurate than public NTP servers and while these signals can be interrupted by events like solar storms or intentional signal jamming, backup clocks that have been synchronized to the satellite signals when present can be placed in each individual data center to provide the desired uncertainty levels during these interruptions.

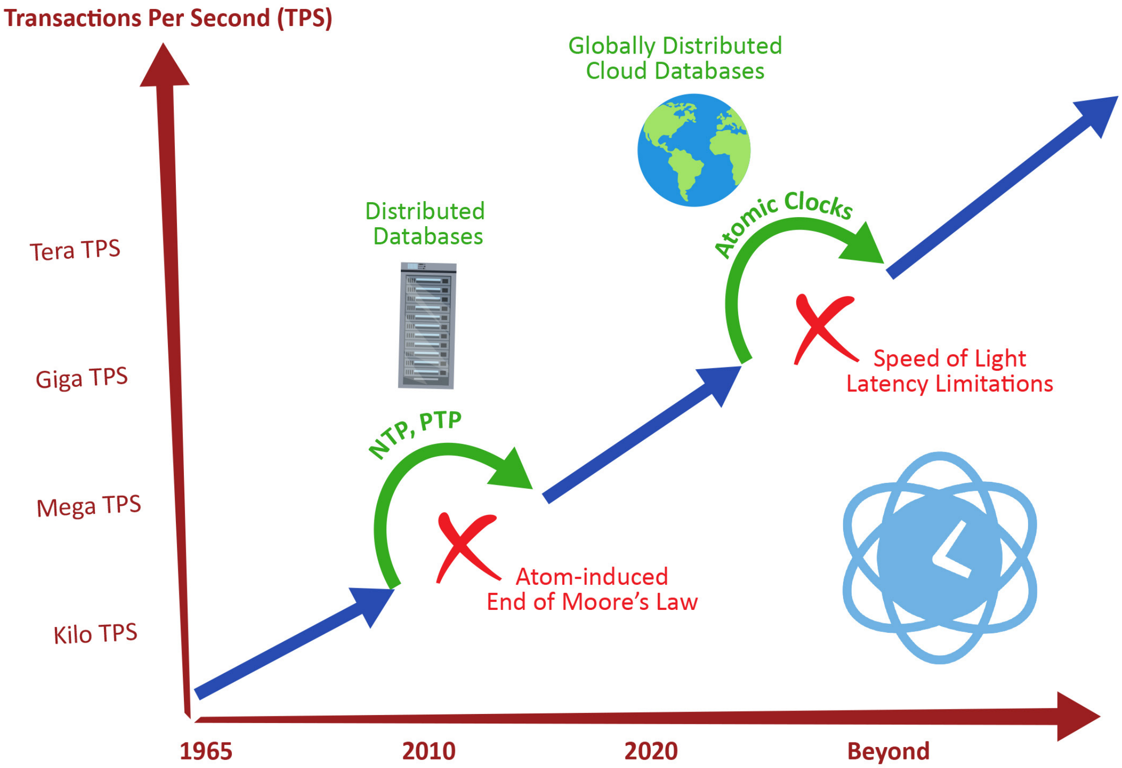

Figure 7: The evolution of database transaction rates and the enabling and disabling technologies.

As our quest to acquire, store, and transact data in the future continues to grow, novel atomic clock technologies and time transmission systems with lower uncertainties will be needed. Currently national timing labs are developing atomic clocks that work on the optical transitions that occur when an electron jumps orbital shells. These offer frequency stabilities measured in quintillionths (1 with 18 zeroes following it) and will eventually be used to redefine the unit second [6]. Signal transmission through dedicated fiber optic links or airborne lasers are already yielding improved transmission accuracy [6]. With the continued innovations data, the atom, and light will continue their complex love-hate relationship to enable ever larger quantities of data processed at ever increasing rates without consistency issues or causality casualties.

[1] https://newsroom.intel.com/wp-content/uploads/sites/11/2018/05/moores-law-electronics.pdf

[2] https://ourworldindata.org/uploads/2020/11/Transistor-Count-over-time.png

[3] https://semiengineering.com/knowledge_centers/manufacturing/process/nodes/

[4] https://time.com/6108001/data-protection-richard-stengel/

[5] https://www.thoughtco.com/how-many-atoms-are-in-human-body-603872

[6] https://spectrum.ieee.org/optical-atomic-clock-advantage-expands-electronics

DS00006209A