Synchronising Uncertainty: Google's Spanner and Cartographic Time

Brian House

Introduction

In the following text, I discuss contemporary, large-scale, network-distributed databases, exemplified by the largest of all, Google's Spanner — so named because it circumscribes the entire planet. Though largely unknown to the public, Spanner is the infrastructure behind Google search, Google's advertising platform, and applica-tions like Gmail that billions of people use every day.

To operate at such scale, Spanner must synchronize time over the extent of the globe, and I situate this endeavour within a genealogy of Western timekeeping strategies extending from astronomical observations in the age of maritime navigation to the various electromagnetic media that have coordinated the clocks of railroads and satellites. This lineage demonstrates how evolving notions of temporality are inexorably bound to geography and to the material practice of cartography.

I argue that *random access*, a fundamental property of individual hard drives, is already cartographic by virtue of how it encapsulates the contingencies of time—this is what maps aspire to do. By physically extending this principle across the planet, Spanner explicitly links such data cartography with geographic mapmaking.

Further, random access also marks a shift in the evolution of time synchronisation. With Spanner, the ambition to establish an absolute measure of time itself is superseded by the need for synchronic slices—time is executed as "logical snapshots" of globally consistent data. By negotiating a contingent sense of time in order to posit a discrete one, Google extends strategic modes of knowledge that are inseparable from histories of industrialisation, colonialism, and militarism to our everyday interactions with its products.

"Cartographology"

I would like to begin with the hard drive which sits inside every internet server and on which, arguably, contemporary network culture is predicated. Jacques Derrida famously noted that writing is not secondary to spoken language, but that the means of inscription produces its own meaning (Derrida 1980). In an essay entitled "Extreme Inscription: Towards a Grammatology of the Hard Drive", Matthew Kirschenbaum extends this notion by articulating the material characteristics of the disk as a writing technology. Briefly, those are that the drive is

- a *signal processor* that converts between digital and analog signals
- *differential*, in that it both depends on the measurement of difference in the physical media, and, by extension, that it represents difference
- *chronographic* because the physical act of reading and writing data takes time
- *volumetric* since the disk platters take up space
- *rationalized* because every part of the disk has an address
- *motion-dependent* as the read/write head mechanically moves
- *planographic* because "the surface of the disk, in order to fly scant nanometers beneath the air bearings, must be absolutely smooth"
- and *non-volatile* because a disk does not forget anything when it is turned off

Some of these properties may be more or less relevant with newer technologies (solid state drives, for example, have no moving parts, so the idea of motion-dependency has to be loosened). But it is significant that most of these properties describe temporal processes inherent in the operation of the device—it is precisely these material contingencies in time that the hard drive encapsulates and attempts to conceal.

Such encapsulation is exemplified by *random access*[1] — another of Kirschenbaum's properties that more or less incorporates all the rest. The term refers to how the data of a storage medium can be accessed without regard to the order in which the data have been written. This differs fundamentally from sequential storage media such as magnetic tape in which information is arranged linearly and order is directly related to access time (imagine fast-forwarding and rewinding a cassette to get to your favourite song). To quibble, in any given situation certain data may in fact be quicker to access than others. But the goal of random access is to minimize the average time taken for a program to read or write an unpredictable sequence of data. In effect, this abstracts the details of the storage mechanism so that access time can be treated as a constant by the software that uses the disk. "Random" as "unpredictable" thus sits alongside its colloquial usage as "irrelevant"—constant time means everything in the data space is treated the same.

This planer, addressable, timeless surface functions in a way analogous to a geographic map. As Michel de Certeau beautifully puts it, maps transform

> *the* temporal articulation *of places into a spatial sequence of points. A graph takes the place of an operation. A reversible sign is substituted for a practice indissociable from particular moments and "opportunities" … it is thus a mark* in place of *acts.* (Certeau 1984, 35)

The map gains its power from this atemporality — that the flow of time has been deferred elsewhere means it can be "seized as a whole by the eye in a single moment" (Certeau 1984, 35), and it is this that enables strategic planning. This is not so different from how we think of data as a field of knowledge laid out before us. Us, or an algorithm — both the search routine that interprets the past and the artificially intelligent program that predicts the future depend on a static, map-like representation on which they can operate. Therefore, what I'm proposing is that what's at stake with storage technology is not only a matter of grammatology, as in the study of writing, but of what might be called *cartograph-ology* and the equally inscriptive cultural practice of mapmaking. If Kirschenbaum has elucidated the cartographic techniques of the hard drive, what are those of a distributed database such as Spanner?

Consistency

Random access is technically straightforward to achieve when it comes to an individual disk within a single computer. But consider that Spanner is, as Google says, "designed to scale up to millions of machines across hundreds of datacenters and trillions of database rows" (Corbett 2012, 1). Further, these machines are not in the same place — there are data centers on six continents. Data in such a distributed system are *sharded*, which means that a single database must be coordinated across a network of storage devices. Sharding allows the system to scale — it abstracts the database from the disk in order to overcome the inherent size and speed limitations of individual pieces of hardware. This means that unlike Kirschenbaum's grammatology of a hard drive, a *cartographology* of a distributed database cannot be done purely on a mechanical level. Rather, it must account for the software architecture and processual techniques whereby that hardware is organised.

In that regard, we have to consider the big problem for any distributed database — maintaining *consistency*. A consistent database is one that is always in a valid state — that is, all information across the network is up-to-date, and at any given time all applications and users are accessing the same information. This is a necessary prerequisite if it is going to function as a map. Again, that is easy for a single disk, but transfer time across the distributed network, especially under global circumstances, makes this extremely difficult.

To address it, Google starts with the idea of the *logical snapshot*, whereby the data across all machines, in all data centres, across every continent, is known to be consistent at a given point in time in the past. To be able to do that, you need to know the order in which the data have been written, irrespective of which shards they have been written on. This is easier said than done — techniques developed prior to Spanner rely on "complicated coordination protocols" (Metz 2012) to let each other know about each write — but such complexity limits the scalability of the system and its capability to act as a truly unified whole.

Google's innovation at first seems almost banal — to determine the order of the data, simply record the time at which each was written. Assuming a "global wall-clock", a logical snapshot is just a temporal slice at some point in the near past — far enough *in the past* to account for the communication delay between all the shards. However, the existence of such a clock turns out to be a big assumption. Google's Andrew Fikes declares, "as a distributed systems developer, you're taught from — I want to say childhood — not to trust time" (Metz 2012). Fikes could also mean any given *representation* of time, but the conflation is revealing. It situates Google's drive to establish a global wall-clock, which is the central ambition of Spanner, within a genealogy of Western timekeeping strategies concerned with synchronization over expanding geographic areas.

A brief history of time(keeping)

Peter Galison has written a persuasive history tracing the relationship between geography, media, and synchronicity (Galison 2003). He explains how the emergence of the mechanical clock in Europe in the sixteenth century permitted the unbinding of time from location — that is, a clock, propelled by its own internal mechanism, may indicate what time it is somewhere else. As Galison discusses, this was a critical, if incrementally achieved, innovation for navigation and cartography. Consider that in order to understand the globe as a grid of latitude and longitude coordinates, one's position on the grid has to be observable. Navigation by star position provides a relatively straightforward way to determine latitude via the night sky — the star Polaris aligns with the north pole, and the Southern Cross can be used to triangulate the south. But because of the rotation of the earth, longitude can only be reliably fixed given the time of a known location. For example, if it is midnight in London and the stars where I am are shifted ninety degrees from what I would expect in the London sky, then I am a quarter way around the globe. Hence the rationalised sense of time as a constant, independent dimension that is the same everywhere also marks the birth of contemporary cartography. This continues to resonate in culture: time and space are separately thought, but practically bound.

Galison goes on to trace the progression whereby train routes maintained a unified "train time" which gradually reconciled the divergent timekeeping of regional metropolitan centres. This process was predicated by the emergence of electromagnetic media in the form of the telegraph and later the radio that allowed time synchronisation to happen over greater distances — the infrastructure that is the direct antecedent of the fiber optics and undersea cables that carry data today. Progressively, the observatory hubs anchoring clocks to local astronomical measurements surrendered to the international standard of Greenwich Mean Time and modern discrete time zones. And at each step, this was a political negotiation, from the municipal level all the way up to the empire-building of Britain, industrial expansion in the

US, and the extension of French Revolutionary values seeking rationalised standards. As Galison puts it, "beating overhead in church spires, observatories, and satellites, synchronized clocks have never stood far from the political order" (Galison 2003, 143).[2]

While Greenwich Mean Time was originally directly tied to measurements at the Royal Observatory in the UK, it turns out that the Earth's rotation is not constant — tidal friction and changes in the Earth's mass due to melting glaciers cause it to vary. Subsequently, a more accurate reference was needed. Decoupling the notion of the day from the transit of the sun, which happened on January 1, 1972, is a profoundly modernist gesture. 9,192,631,770 cycles of radiation from the caesium-133 atom is the current international standard for one second, and the atomic clock is the basis for Universal Coordinated Time, or UTC.

Atomic clocks are also the foundation of contemporary map-making. Each of the satellites that make up the Global Positioning Service, or GPS, contains an atomic clock within it. In many ways, GPS — originally deployed by the US military — culminates the narrative of terrestrial time synchronization by literally rising above the earth. The system broadcasts clock signals to the ground, where receivers, ubiquitously embedded in things like mobile devices, triangulate their position — minute differences between the received times indicate varying distances to the known location of each satellite. This temporal negotiation smooths geographic space into the Cartesian grid postulated by post-Enlightenment thought — it is exemplified by the gesture of looking down at GPS-powered Google Maps on your iPhone in order to see the earth from above.[3]

True time

How does that iPhone keep time? Computing devices generally make use of a *real-time clock*, or RTC, which is based on a cheap crystal oscillator. An RTC will inevitably drift out of synchrony with other clocks due to temperature fluctuations and other physical factors. However, with systems connected to the internet, the RTC synchronizes with a *time server* using the Network Time Protocol, or NTP. Such servers are maintained by governments (time.nist.gov), independent foundations (pool.ntp.org), and large corporations (time.apple.com). In this case, synchronization happens via internet packets, and as such it is subject to network latency. For most systems, though, NTP is good enough.

However, when a Google engineer doesn't "trust time", it reflects practical experience that much can go wrong with NTP synchronization procedures. Communication may fail due to network variability, and, critically, machines distributed around the world will experience uneven latency in relation to a central time server. Clocks may or may not line up, and worse, there is no way to verify after the fact if this has happened.

Hence Spanner. First, Spanner eschews NTP and is linked explicitly to GPS — every data centre has a "time master" unit that is always receiving GPS time. There are also "Armageddon masters" within the system that have their own atomic clocks, in the extreme case that GPS should ever fail. Each machine continually updates its RTC by continuously polling a variety of these master clocks, both in the local data center and from across the network. The slightly differing times received from all the masters are combined to produce an optimal time estimation, an emergent consensus that is uniform across the globally-distributed database. This uniformity, however, comes with a level of calculated uncertainty, an artifact of all the aggregated network latency together with clock drift on individual machines.

This negotiated uncertainty is represented by what Google calls the TrueTime API. An API, or Application Programming Interface, is an essential programming concept based on obfuscation. Software components need not — and in fact, should not — know the implementation details of other components. Rather, an API provides stable terms through which software can reliably communicate while hiding the underlying, and potentially variable, mess. Application code that uses Spanner does so through the TrueTime API, which "explicitly represents time as … an interval" that indicates the earliest and latest points that an event could possibly have happened. In other words, the brilliance of the TrueTime API is that it "reif[ies] clock uncertainty" (Hsieh 2012).

Google describes this strategy as being Rumsfeldian — that is, "known unknowns are better than unknown unknowns." They abandon the naïve hope that fast is fast enough — instead, Spanner leverages statistical knowledge about its own vast hardware to gauge how confident it can be about time. In an industry obsessed with making things faster, a counter-intuitive feature of the system is that "if the uncertainty is large, Spanner *slows down* to wait out that uncertainty" (Hsieh 2012). All of this is done in service to having a global wall-clock that Google can depend on — it is what makes those logical snapshots possible.

Random access geography

Finally we can return to Kirschenbaum. Does the scale achieved with Spanner exceed the qualities of the individual hard drive? This is undeniably the case. Yet, in many ways, such a geographically totalising database infrastructure aspires to function as a single disk. Revisiting and reformulating Kirschenbaum's grammatology, or our cartographology, elucidates the comparison.

Spanner is certainly a *signal processor*, but that analog-to-digital conversion now happens multiple times across the network switches and undersea cables of distributed infrastructure. It is still as *differential* as its individual disks. TrueTime itself clearly marks Spanner as *chronographic*. If the hard drive is volumetric, Spanner's data centers are extremely so. It is a *rationalised* system, because any data across the space may be addressed, and, significantly, that location *is also a geographical place*. Is Spanner *motion dependent*? If the hard drive has the spinning disk, Spanner adds the orbit of GPS satellites, the oscillation of the caesium atom, and the packets traversing the network. *Non-volatility* maps to Spanner's robustness and those Armageddon masters. And *planographic* speaks to the data centers spread out over the surface of the earth. We can therefore construct an analog between how Kirschenbaum enumerates the technology of inscription that is the hard drive and this far larger system, supplementing purely mechanical elements with software and geographic processes.

What about *random access*? Spanner's logical snapshots accomplish the same thing — they render the notion of time itself secondary to a consistent plane of stable data. It is the felt quality experienced by the individual or application that is able to call up any piece of information from the database at will, regardless of the material conditions of its storage. This is Spanner's goal, for all data to be available from any point and time, at a geographic scale.

Spanner makes the isomorphism of a hard drive to a map quite literal. This is even reflected by certain representations that Google puts forth, namely, Google Earth and Google Maps. The effortlessly spinning globe that one floats above in Earth might well serve as a metonym for the random access space coordinated by Spanner, as "Google Earth … can be understood as the aesthetic rendering … of the logic of Google search" (Munster 2013, 63). Search is, of course, the paradigmatic operation of random access and is inseparable from the rationalised qualities of the distributed database beneath it. A plain link is thus established between representation and infrastructure.

Conclusion

Clearly, though, there is a "sense" here, that is missing. Anna Munster's work on how we experience networks and data is particularly compelling in this respect. She explains how there is a difference between recognising something that is already within the parameters of what is knowable, as one does when pointing something out on a map, and the active, contingent process of experiencing some unknown potential unfold in time (Munster 2013, 43). The latter is, in short, the uncertainty that is exactly what Spanner urgently seeks to obfuscate. Where has the time gone? The TrueTime API extends the techniques of timekeeping in Galison's history — it is a synchronization procedure. But with Spanner, the quest to chase uncertainty down to ever finer intervals — even to the oscillations of the atom — is superseded by a concern with a sequence of logical snapshots that bypasses that uncertainty. Potential is abstracted away by an engineered lag behind the "now".

That the human experience of time is irreducible to modes of timekeeping should be self-evident — otherwise we would never have to check the clock. Consequently, as a totalising project, Spanner is aspirational. We are well acquainted with the "spinning beach ball of death" and other aesthetic ruptures we experience when technology can't quite keep up (see Winnie Soon's contribution in this volume) — the unresponsive hard drive, the stutter in the video stream, even the tone-deaf targeted ad — these moments reveal material contingencies that resist representation. In Spanner's case, "network lag" is a kind of shorthand for the physical resources and social structures required to build, connect, and maintain millions of computers across vast distances. They are left out of the map even as they are essential to the cartographic act. But when Spanner slows down the world to make it conform to its strategic view, that elision manifests in the micro-experiences of billions of users.

To "keep time" is to mark temporal experience, but to "keep" is also to withhold or suspend. To the extent that maps — whether of data, geography, or both — accomplish this, they reserve extraordinary power. But by understanding the practices of timekeeping that make such abstraction possible, we can rethink them as a particular construction of lived time and modulate our participation accordingly. After all, "keeping time" is also what drummers do in musical performance, and a distributed database, too, is a matter of temporal aesthetics rather than absolute measure.

Notes

1. Not to be confused with Random Access Memory.

2. Galison reprints a map of a French plan for synchronizing South America, with telegraph lines reaching Rio from Europe and encircling the continent, passing through Lima, and continuing north to the United States. It bears a remarkable similarity to an image in Wired accompanying its article on Spanner, an isomorphism which evinces similar ambitions.

3. See the work of Johnathan Hanahan, http://www.hanahan.works/pixel_posters.html.

References

Brewer, Eric. 2000. "Towards Robust Distributed Systems." PODC Keynote, July 19. http://www.cs.berkeley.edu/~brewer/cs262b-2004/PODC-keynote.pdf.

De Certeau, Michel. 1984. *The Practice of Everyday Life*. Trans. by Steven F. Rendall. Berkeley: University of California Press.

Derrida, Jacques. 1980. *Of Grammatology.* Trans. by Gayatri Spivak. Baltimore: Johns Hopkins University Press.

Galison, Peter. 2003. *Einstein's Clocks, Poincaré's Maps*. New York: W. W. Norton.

Hsieh, Wilson. 2012. "Google's Globally-Distributed Database." Paper presentation. See http://new.livestream. com/accounts/1545775/osdi12/videos/4646642 and also James C. Corbett et al (2012), *OSDI: Tenth Symposium on Operating System Design and Implementation Proceedings.*

Kirschenbaum, Matthew. 2008. "Extreme Inscription: A Grammatology of the Hard Drive" reedited from *Mechanisms: New Media and the Forensic Imagination*. Cambridge MA: MIT Press.

Marzullo, K. A. 1984. *Maintaining the Time in a Distributed System: An Example of a Loosely-Coupled Distributed Service.* Palo Alto: Stanford University.

Metz, Cade. 2012. "Google Spans Entire Planet With GPS-Powered Database." *Wired*, September 19. http://www.wired.com/2012/09/google-spanner/all/.

Munster, Anna. 2013. *An Aesthesia of Networks*. Cambridge MA: MIT Press.