“We want to make it easy at Google for programmers to build their applications” is the reasoning given by Wilson Hsieh to build Spanner: Google’s Globally Distributed Database. Spanner was introduced in 2012 at OSDI, and an academic paper that is cited at the end of this paper. Hsieh says, “Spanner is used currently at Google on our most valuable data being our ads.” They transitioned from a sharded MySQL database to Spanner. The vision of Spanner was to merge the ideas of easy-to-use, semi-relational interface, transactions, and an SQL based query language of the database community with the ideas of scalability, automatic sharding, fault tolerance, and external consistency of the systems community.

Spanner provides everything most databases already do, namely schematized tables, semi-relational model, SQL query language. Spanner allows the data being stored in the database to be stored across the globe. A level deeper Google is sharding that data and distributing it to different data centers in those countries. Then they replicate that data more than one way, so it is backed up. “Store data across thousands of machines, tens of thousands of data centers, millions of rows that’s the scale we are talking about” proclaims Hsieh. The big research implications of Spanner are how to execute read only transactions across data centers in a lock-free manner. An example of a read only transaction is viewing the page of friend on a social network. That feature is possible because of the theoretical property that distributed transactions are externally consistent. Spanner is the first system to provide this property at world wide scale. A multi-version system is not used for the process of write to preserve concurrency. Synchronizing snapshots uses global wall-clock time to implement external consistency. This is essential for the consistency of the entire system. The database must satisfy the properties that timestamp order (of the social media post) is equal to the commit order, and the timestamp order also respects global wall-time order. This is very difficult to implement across multiple systems in multiple data centers. To help with the implementation Google created TrueTime API. The API exposes clock uncertainty, and the guarantees on Spanner’s timestamps depend on the bounds that the implementation provides. An implementation of the TrueTime API is produced by Google’s cluster management software. Making the clock uncertainty manifest in the API is one of, if it not the most, significant part of this project.

Spanner is a globally-distributed, high-availability storage system being used at Google today to host their most important feature, the advertising backend. A Spanner schema language is in the works along with automatic load based resharding. In the future Google is also looking to host full applications across distributed data centers. Although developed for internal use originally, Google now sells the service to enterprises by promising 99.999% availability, no planned downtime, and enterprise grade security. They sell the service as Cloud Spanner noting “benefits of relational database with non-relational horizontal scale… delivers high performance transactions and strong consistency.” This new state-of-the-art technique solves the challenge of efficiently managing app databases backends while simultaneously empowering developers with the tools they need to build apps. Everything is real-time and up to date synchronously thanks to the TrueTime API. Other features include multi-language support, relational semantics, and high availability and scalability. The most important feature however is the ability for developers to use in order sequencing in their app logic. The possibility to build distributes systems with much stronger time semantics is made possible by TrueTime API. Because of these developments the database community should stop depending on poorly synchronized clocks and weak time APIs.

Works Cited

<https://static.googleusercontent.com/media/research.google.com/en//archive/spanner-osdi2012.pdf>

<https://cloud.google.com/spanner/>