Spanner: Google's Globally-Distributed Database

The paper covers the features and architecture of Spanner which is a globally-distributed semi-relational database developed by Google. Spanner was developed to overcome the shortcomings of its predecessors with features like, ease of sharding data across multiple continents, along with highly consistent database operations. It successfully implements all of these at a production level and now hosts the data for Google's Ad network. It achieves these through a novel take on the idea of a globally consistent clock and externally consistent read-write operations.

Spanner introduces the concept of the TrueTime API which shows that reifying clock uncertainty in the time API makes it possible to build distributed systems with much stronger time semantics and that we should no longer depend on loosely synchronized clocks and weak time APIs in designing distributed algorithms.

Spanner also successfully implements the concept of external consistency and ensures that the commit order adheres to the timestamp order which adheres to a consistent global clock time.

At the same time, Spanner's scope will eventually get limited by the development of better CPUs as currently most of the clock failures are overshadowed by CPU failures. As better CPUs are developed, it is likely that the clock failures themselves could become the next bottleneck.

Furthermore, although Spanner is scalable in the number of nodes, the node-local data structures have relatively poor performance on complex SQL queries, because they were designed for simple key-value accesses.

We would like to ask the authors about improvement potential of the TrueTime API as well as the storage costs for performing a snapshot read.

It would be interesting to see a more robust and accessible database management system emerge out of Spanner's distributed backend. It would be one with its own powerful query language that that can perform many powerful operations that are not conventionally possible in distributed databases with a lack of consistency.