Innovating with storage formats to push the limits in a hybrid database
(Invited Talk)

Eugene Kogan

Software Architect, SingleStore

Abstract

Data storage formats and also ways a query processor interacts with data structures in different storage tiers is one of the main design areas databases use to add value for customers and become more competitive. In this talk I'll explore SingleStore's proprietary storage formats and indexing techniques. I will also discuss how the data structures are used by the query processor, and illustrate that in many cases indexes and mixed column- and row-centric data representation do not lead to query planning complexity. SingleStore's case is interesting because of the variety of requirements to data storage and access methods, relative maturity and continued investments. SingleStore is a distributed hybrid transactional and analytical database that efficiently supports queries with low, medium and high selectivity, as well as DML with small, sparse and bulk data updates. I'll discuss reasoning for prior design decisions, and report on new unpublished work.

Speaker Biography: Eugene's relatively short 22+ years in databases took him from Microsoft SQL Server to Google Spanner, Google BigQuery, Snowflake, and, finally, to a Software Architect role at SingleStore. His biggest interests in the RDBMS field have been semi-structured data processing, data replication, distributed query coordination and execution, and, lately, hybrid system architectures that include transactional, analytical and complex event processing in a single database.