Materialize: A platform for building scalable event based systems

Frank McSherry
Materialize Inc.
mcsherry@materialize.com

ABSTRACT

Materialize is a system that presents to users as SQL against continually changing data. It transforms inbound streams of *change data capture* events into streams that exactly correspond to transformed data, and maintains indexed representations of the results for efficient access and operation. SQL over changing data is surprisingly (for me) expressive: Materialize can operate on unbounded data, implement data-driven windows, and perform event-based queries, all with ANSI standard SQL. We will discuss what an event-based SQL system looks like, SQL idioms that give rise to traditionally stream-exclusive behavior, and how one architects such a system to scale across multiple dimensions.

1 BIOGRAPHY

Frank McSherry is Chief Scientist at Materialize. He was initially a graduate student at the University of Washington where he worked with Anna Karlin on Spectral Graph Theory, then a Researcher at Microsoft Research SVC where he co-developed differential privacy and led the Naiad research project, and later a Visiting Researcher at ETH Zurich where he honed the dataflows timely and differential.