Substrait: Rethinking DBMS Composability

(Keynote)
Jacques Nadeau
Co-founder & CEO Sundeck
jacquesnadeau@gmail.com

Abstract

We will give an overview of Substrait, a new (started September 2021) Apache-Licensed open source effort to develop a standardized way to represent compute operations and query plans. We'll start by talking about the dynamics that led to the creation of Substrait (including LLVM IR and JVM bytecode inspirations). We'll then give an overview of how Substrait primitives allow expression of common plan patterns. We'll then discuss Substrait's model of disciplined extensibility. Lastly we'll cover the large number active Substrait integration efforts including those underway in Ibis, Arrow, Velox, Presto, Spark, and DuckDB.

Biography

Jacques Nadeau is co-founder & CEO of Sundeck. Previously, he was CTO and co-founder of Dremio. Jacques is co-creator and founding PMC chair of Apache Arrow, co-creator of Substrait and an Apache Calcite PMC member.