

Velox: Unifying execution engines across Meta and beyond

(Invited Talk)

Pedro Eugenio Rocha Pedreira

Software Engineer

Meta Platforms Inc.

pedroerp@fb.com

Abstract

Velox is a novel open-source C++ database acceleration library created by Meta. Velox provides reusable, extensible, high-performance, engine and dialect-agnostic data processing components for building execution engines and enhancing data management systems. The library is currently integrated or being integrated with more than a dozen data systems at Meta, including analytical query engines such as Presto and Spark, stream processing platforms, message buses and data warehouse ingestion infrastructure, machine learning systems for feature engineering and data wrangling (PyTorch), and more. In this talk, the main motivations and value proposition behind Velox will be described, along with its main use cases and an outline of the library and its main optimizations.

Biography

Pedro Pedreira has worked as a Software Engineer at Meta for the last decade, focusing on Data Infrastructure. Currently, Pedro leads the Velox program, which is a major effort at unifying execution engines into an open-source library, spanning more than a dozen engines within Meta and beyond. In the past, he worked on log analytics engines (such as Scuba), and Cubrick - an in-memory analytical DBMS he and his team developed from the ground up based on multidimensional indexing ideas proposed in his Ph.D thesis. Pedro holds a PhD and an MS in Computer Science from the Federal University of Parana, in Brazil.