SQream DB - Bigger Data On GPUs: Approaches, Challenges, Successes

This talk will present SQream’s journey to building an analytics data warehouse powered by GPUs. SQream DB is an SQL data warehouse designed for larger than main-memory datasets (up to petabytes). It’s an on-disk database that combines novel ideas and algorithms to rapidly analyze trillions of rows with the help of high-throughput GPUs. We will explore some of SQream’s ideas and approaches to developing its analytics database – from simple prototype and tech demos, to a fully functional data warehouse product containing the most important features for enterprise deployment. We will also describe the challenges of working with exotic hardware like GPUs, and what choices had to be made in order to combine the CPU and GPU capabilities to achieve industry-leading performance – complete with real world use case comparisons.

As part of this discussion, we will also share some of the real issues that were discovered, and the engineering decisions that led to the creation of SQream DB’s high-speed columnar storage engine, designed specifically to take advantage of streaming architectures like GPUs.