Fast Software For Fast Hardware: How Analytics Platform MapD Exploits The Massive Parallelism Of GPUs

In this talk Todd Mostak, Founder and CEO of MapD Technologies, will speak about the technical underpinnings of MapD, an open-source GPU-accelerated SQL engine and visual analytics platform that enables querying and visualization of large datasets (tens of billions of records) in milliseconds, without indexing or or other forms of pre-computation. While some of this performance can be attributed to the system’s use of GPUs, a large part is also due to a focus around extracting the full performance and parallelism of all available hardware (including CPUs). Todd will focus his talk on a tour of the interesting architectural features of the system, including its ability to manage data placement across disk, CPU, and multiple GPUs, its LLVM-based operator fusing/query compilation engine, the strategies it takes around minimizing hash table size for group bys and joins, and its in-situ rendering engine which makes use the GPU’s native rendering pipeline to visualize large result sets without copying them off-chip. Todd will conclude with why he thinks GPUs, by virtue of their high compute throughput, memory bandwidth, and rendering capabilities, will become a mainstay for both database processing and general-purpose analytics.