Jedi: A Storage Manager for SIMD-aware, Worst-case Optimal Join Processing

Christopher Ré,
Stanford University

ABSTRACT

This talk describes a new graph-pattern engine called Jedi. Using a recent simplification of worst-case optimal join algorithms due to Ngo et al., Jedi translates join queries into a series of set intersection and union operations. Such set operations are ideally suited to modern CPUs that provides single-instruction, multiple data (SIMD) instructions. Using these ideas, we demonstrate that Jedi outperforms specialized graph engines by over an order of magnitude and relational systems by over two orders of magnitude on standard graph processing queries over real data.

Short Bio

Christopher (Chris) Re is an assistant professor in the Department of Computer Science at Stanford University and a Robert N. Noyce Family Faculty Scholar. His work's goal is to enable users and developers to build applications that more deeply understand and exploit data. Chris received his PhD from the University of Washington in Seattle under the supervision of Dan Suciu. For his PhD work in probabilistic data management, Chris received the SIGMOD 2010 Jim Gray Dissertation Award. He then spent four wonderful years on the faculty of the University of Wisconsin, Madison, before moving to Stanford in 2013. He helped discover the first join algorithm with worst-case optimal running time, which won the best paper at PODS 2012. He also helped develop a framework for feature engineering that won the best paper at SIGMOD 2014. In addition, work from his group has been incorporated into scientific efforts including the IceCube neutrino detector and PaleoDeepDive, and into Cloudera's Impala and products from Oracle, Pivotal, and Microsoft's Adam. He received an NSF CAREER Award in 2011, an Alfred P. Sloan Fellowship in 2013, and a Moore Data Driven Investigator Award in 2014.