



Joy Arulraj

The Design and Implementation of a Non-Volatile Memory Database Management System

Friday, October 20, 2017 – 1:30 p.m. – NSH 3305

For the first time in 25 years, a new non-volatile memory (NVM) category is being created that is two orders of magnitude faster than current durable storage media. The advent of NVM will fundamentally change the dichotomy between volatile memory and durable storage in database systems. These new NVM devices are almost as fast as DRAM, but all writes to it are potentially persistent even after power loss. Existing database systems are unable to take full advantage of this technology because their internal architectures are predicated on the assumption that memory is volatile. With NVM, many of the components of legacy database systems are unnecessary and will degrade the performance of the data intensive applications.

This dissertation explores the implications of NVM for database systems. It presents the design and implementation of Peloton, a new database system tailored specifically for NVM. The dissertation focuses on three aspects of a database system: (1) logging and recovery, (2) storage management, and (3) indexing. Our primary contribution in this dissertation is the design of a new logging and recovery protocol, called write-behind logging, that improves the availability of the system by more than two orders of magnitude compared to the ubiquitous write-ahead logging protocol. Besides improving availability, we found that write-behind logging improves the space utilization of the NVM device and extends its lifetime. Second, we propose a new storage engine architecture that leverages the durability and byte-addressability properties of NVM to avoid unnecessary data duplication. Third, the dissertation presents the design of a latch-free range index tailored for NVM that supports near-instantaneous recovery without requiring special-purpose recovery code.

Thesis Committee:

Andy Pavlo, Chair

Garth Gibson

Todd Mowry

Samuel Madden, MIT

Donald Kossmann, Microsoft Research