Replicating and Virtualizing Non-Volatile Main Memory

Fall 2014 CNS Review

Progress Report

Steven Swanson

10/8/14

Next-generation non-volatile memories (NVMs) promise DRAM-like performance, persistence, and high density. They can attach directly to processors to form non-volatile main memory (NVMM) and offer the opportunity to build very low-latency storage systems. These high-performance storage systems would be especially useful in large-scale, virtualized data center environments, where reliability and availability are critical. However, providing reliability and availability to NVMM is challenging, since the latency of data replication can squander the low latency that NVMM can provide.

We have built Mojim, a system that provides the reliability and availability that large-scale storage systems require, while preserving the performance of NVMM. Mojim achieves these goals by using a two-tier architecture in which the primary tier contains a mirrored pair of nodes and the secondary tier contains one or more secondary backup nodes with weakly consistent copies of data. Mojim uses highly-optimized replication protocols, software, and networking stacks. Our evaluation results show that Mojim provides replicated NVMM with 0.5 to 3.5x the bandwidth of unreplicated NVMM and reduces the average latency of the unreplicated NVMM by 29% to 73%. We have run three popular applications, including MongoDB, on Mojim. The resulting system is 3.4 to 4x faster than the MongoDB's native replication mechanism and 35 to 42x faster than unreplicated MongoDB.

We are currently exploring how to extend Mojim to support replication, migration, and fast fail-over in virtualized environments.

We have submitted the Mojim paper for publication and would be happy to provide a copy of the draft on request. Contact swanson@cs.ucsd.edu.