Abstract

Non-volatile memory (NVRAM) is becoming available. With the availability of hybrid DRAM and NVRAM memory on the memory bus of CPUs, a number of experimental file systems on NVRAM have been designed and implemented. In this thesis we present the design and implementation of a file system on NVRAM called DurableFS, which provides atomicity and durability of file operations to applications. It provides ACID properties to transactions involving multiple files. Due to the byte level random accessibility of memory, it is possible to provide these guarantees without much overhead. We use standard techniques like copy on write for data, and a redo log for metadata changes to build an efficient file system which provides durability and atomicity guarantees to transactions. Benchmarks on the implementation shows that there is only a 7% degradation in performance due to providing these guarantees.

Since the storage devices are not only meant for files, we also have designed a Key-Value store for such hybrid systems. In this thesis we also present the design and implementation of a distributed key-value storage system on a cluster (DKVS) having NVRAM in each node. We have chosen a key-value system to enable different storage architectures to be implemented on top, such as file systems with hierarchical naming structures, relational database, NoSQL systems etc. Keeping RDBMS primarily in mind, DKVS provides support for transactions with ACID properties for sequences of operations on multiple key-value items. We provide resilience by allowing replicated copies of each key-value pair in another node. We have used techniques used in clustered file systems to provide caching of locations of key-values, and RDMA to fetch a value from remote node. The implementation is on a six-node system as a loadable Linux module. RDMA is implemented in software using Soft-RoCE. A portion of RAM was simulated as NVRAM using the “memmap” feature of Linux. We evaluated our implementation using our custom workload as well as with the YCSB benchmark. DKVS was compared with a Redis cluster having six nodes and found that our implementation is comparable with Redis with durability and consistency guarantees.