INCORPORATING SOLID STATE DRIVES INTO STORAGE SYSTEMS
Ph.d. Qualifying Examination Proposal
in
COMPUTER SCIENCE
by
Rosie Wacha
June 2009

Abstract

Big data stores are becoming increasingly important in a variety of domains including scientific computing, internet applications, and business applications. For price and performance reasons, such storage is comprised of magnetic hard drives. To achieve the necessary degree of performance and reliability, the drives are configured into storage subsystems based on RAID (Redundant Array of Independent Disks). Because of their mechanical nature, hard drives are relatively power-hungry and slow compared to most other computer components. Big data centers spend tremendous amounts on power, including cooling, adding significantly to their overall costs. Additionally, drives are orders of magnitude slower than electrical computer components, resulting in significant performance challenges any time disk I/O is required. Recently, SSDs (solid state drives) have emerged based on flash memory technology. Although too expensive to replace magnetic disks altogether, SSDs use less power and are significantly faster for certain operations.

This proposal examines a promising architecture that uses a limited number of SSDs to decrease the power consumption and either increase the performance or reliability of RAID storage subsystems. The research will examine the use of SSDs for parity storage in a disk-SSD hybrid RAID system. Because of its better performance characteristics, SSD parity storage reduces the disk overhead associated with parity storage and thereby significantly reduces the disk overheads caused by RAID. This decreases the power consumption and can be used to increase the performance of simple RAID schemes or increase reliability by enabling the use of higher order RAID schemes with less performance penalty. Storing parity on SSDs can reduce the average number of I/Os serviced by the remaining disks by up to 25-50%. Data storage systems are typically evaluated in terms of performance while reliability and power consumption are secondary concerns. We propose that by replacing some hard drives with SSDs, we reduce power and either improve performance or add reliability and maintain performance.