Reliable Energy-Aware SSD based RAID-6 System  
(FAST 2012, February 14-17, San Jose, U.S.A.)

Mehdi Pirahandelh, Deok-Hwan Kim*  
Dept. of Electronic Engineering, Inha University, Incheon, South Korea  
mehdii@iesl.inha.ac.kr, deokhw法兰@inha.ac.kr, *Corresponding Author

Motivation and goals

- Challenges of SSD markets deal with  
  - Breakthroughs in terms of energy consumption, reliability, performance of SSDs  
  - Achieve optimal energy consumption for HDD, SSD and large storage systems  
  - Dynamic voltage measurement  
  - Auto power management  
  - Compiler directed energy optimization  
  - Performance and energy use of RAID systems with various type SSDs  
  - Reliability system and some criteria  
  - Cost for replication based schemes are expensive  
  - Erasure codes considering power management  
  - Need to measure repair transition rate  
  - Average data loss due to N iterations of IO operations on critically exposed sectors  
  - Lower utilization level produces lower failure rates

- Goals  
  - Proposes an energy aware algorithm  
  - Model which enables SSDs to increase the performance and decrease the level of power consumption

Core layer in the RAID-6 controller
- Encoding and decoding scheduler
- Random and sequential read and write modules with an internal failure detector using erasure codes
- Traces are updated using the log generator

Reliability-aware layer
- Procedures for prediction of SSD reliability
  - Use erasure codes generated from the core layer

Energy-aware layer
- Pre-processing procedure
  - Initialize the reliability measurement and utilization level
  - SSD selector and pre-processing procedure
- Set the status of selected SSD power mode to idle-sleep-active via imported traces
- Dynamic SSD scheduler
- Visualize the statistics of SSD energy consumption
- Update the power mode into idle, sleep or active.
- Host interface layer, flash translation layer and NAND flash chips.

Energy flow of sequential/random read/write operation using proposed model
- Data pages are segmented into large chunks
- Use power switching of SSDs after writing of each chunk is done
- After reading-writing of current chunk in SSD Sj, IO operation of next chunk is performed in SSD Sj+1
- Measure SSD reliability and choose the parity SSD Sj+4 and Sj+5
- with less utilization level
- Read operation needs to access four disks from Sj to Sj+3 sequentially and skips two parity SSDs
- When one disk fails, 1st parity SSD is “active” and 2nd parity SSD is in “sleep”
- When two disk failure occurs, both SSDs are in “active” mode
- Write operation requires accessing six SSDs using power switching modes

Table 1: pseudo code of reliable energy-fault aware algorithm

Table 2: SSD energy consumption for busy server-like workload

Contribution

- Improved approach for periodic estimating the energy consumption of SSDs
- The reliability estimation considered to enhance the energy efficiency on the SSD based RAID-6 system
- A layered architecture for reliable energy-aware RAID-6 system
- Reduce energy consumption of parallel access of SSDs
  - Dynamically switching of SSD power modes among active, idle, waken and sleep
  - Allowing one of parity SSDs in safe zone to be sleep mode
  - Segregating data pages into large chunks
  - Using the power switching of SSDs after writing or reading of each chunk is done
- Reduce energy consumption during active mode
  - Activating SSDs sequentially and minimizing delay time of power switching of SSDs
  - Avoiding repetitive accessing of same SSD by using large chunks
  - Increasing number of disks in idle mode

Acknowledgements

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0000114) and in part by the Ministry of Knowledge Economy(MKE) and Korea Institute for Advancement in Technology (KIAT) through the Workforce Development Program in Strategic Technology.