SSD-based Energy Efficient Cloud Storage

Salma Rodriguez

Florida International University
srodr063@fiu.edu

November 27, 2012

Motivation

Solid State Technology

High capacity EEPROM devices have been shown to reduce energy consumption when used as local cache for hard disk drives [2, 3].

Distributed SSD Caching

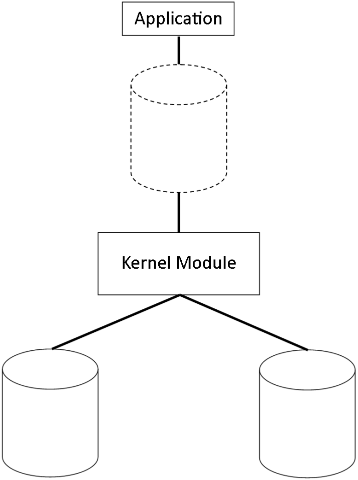
With a distributed caching solution, we hope to increase the working set in flash memory and keep server disks spun down.

Properties to Explore

We explore the properties of dynamic spin down of storage server disks and replication of cold pages.

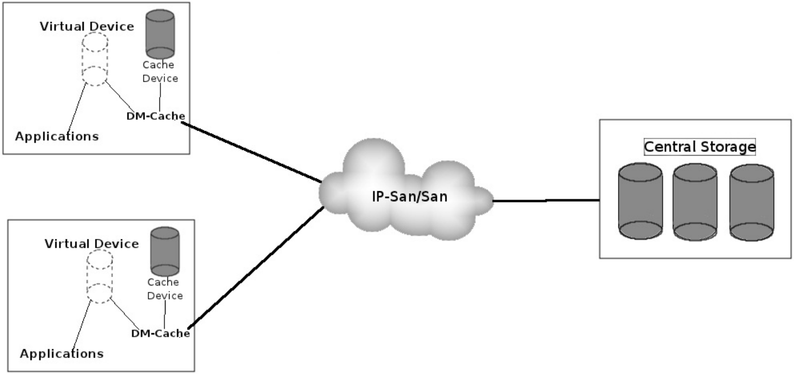
Linux Device Mapper

- device mapper facilitates mapping between two devices
- only pseudo device in volatile memory visible to applications
- kernel module uses device mapper to map bios sent to pseudo device onto real devices (source and target)
- devices may not necessarily be locally attached



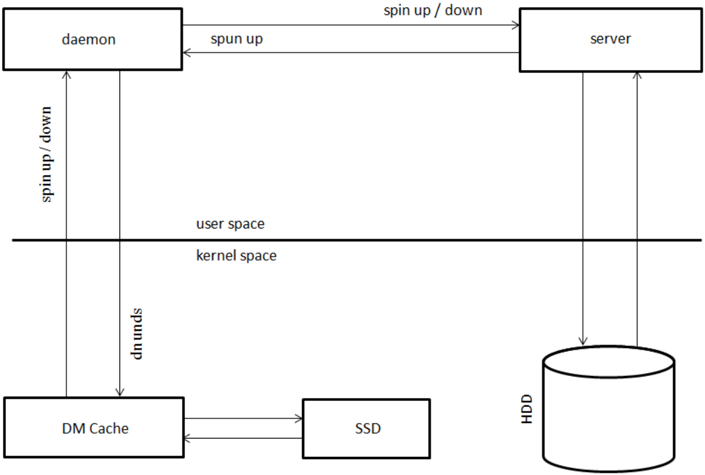
Device Mapper Cache

- Client machines communicate through iSCSI or AoE network.
- DM Cache takes advantage of spatial and temporal locality by storing popular data in a local cache.



Design

Figure 1 The state of each disk is controlled dynamically.



Implementation

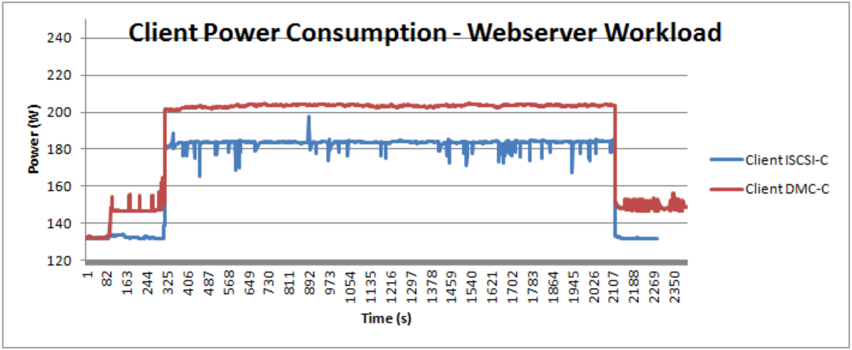
Algorithm 1 spinning the disk up or down dynamically

Precondition: storage server disk is spinning

```
procedure SPIN UP OR DOWN
    T ← constant specified by user
    while true do
        if disk is spinning then
            k ← current time in seconds
            c ← time since last cache miss
            if c + T ≤ k then
                spin down the disk and change state to not spinning
        else                                          ▷ disk is not spinning
            if DM Cache is blocking on a cache miss then
                spin up the disk and change state to spinning
                unblock DM Cache
        schedule a new process
```

Spin Down Results

Figure 2 dynamic disk spin down results for write-intensive workload The red plot is DM Cache with the spin down daemon. The blue plot is iSCSI setup without DM Cache.



Spin Down Results

Evaluation

- The baseline for our results is a simple iSCSI storage setup.
- We ran a 30 min test using a webserver-based write-intensive workload.

Interpretation

- DM Cache with the spin down daemon increases power by 14 Watt more than the baseline.
- Spin down worker thread performs too many context switches.
- Hypothesis: context switches are a performance bottleneck and also consume more energy over time.
- Next: use sleep instead of schedule system call.

New Spin Down Policy

Algorithm 2 new policy using sleep instead of schedule

Precondition: storage server disk is spinning

procedure SPIN UP OR DOWN
    $T \leftarrow$ constant specified by user
    while true **do**
        sleep for T seconds
        if disk is spinning **then**
            $k \leftarrow$ current time in seconds
            $c \leftarrow$ time since last cache miss
            if $c + T \leq k$ **then**
                spin down the disk and change state to not spinning
        else ▷ disk is not spinning
            if DM Cache is blocking on a cache miss **then**
                spin up the disk and change state to spinning
                unblock DM Cache

Cooperative Caching

Basic Idea

- Save power by sending unused clean data to neighboring caches instead of evicting from the cache to the storage server.
- Assumption: not all clients will have their cache full to capacity.
- Tradeoff between free space and caching fairness:

$$\text{benefit} = \frac{\text{free space}}{\text{space taken by other caches}}$$

Challenges

- Choose efficient cache replacement policy to send pages that are less popular to other caches.
- Cache cooperatively and handle nodes as they dynamically join and exit the network of peer caches.

References

- I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. *Special Interest Group on Data Communication (SIGCOMM '01)*, San Diego, California, Aug. 2001.
- J.D. Garcia, J. Carretero, and F. Garcia. Saving power in flash and disk hybrid storage system. *Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS 2009)*, London, U.K., Sep. 2009.
- T. Bisson and S. A. Brandt. Reducing Energy Consumption with a Non-Volatile Storage Cache. *International Workshop on Software Support for Portable Storage (IWSSPS) held in conjunction with the IEEE Real-Time and Embedded Systems and Applications Symposium (RTAS 2005)*, Mar. 2005.