BPAC: AN ADAPTIVE WRITE BUFFER MANAGEMENT SCHEME FOR FLASH-BASED SOLID STATE DRIVES

Guanying Wu
Ben Eckart
Xubin He

Tennessee Tech University

Introduction: Problem Definition and Goal

- Inspect the features that cause the poor random write performance of SSD's.
- Write Buffer: Locate the drawbacks of existing solutions on the cache management scheme.

- Goal: an on-disk cache management scheme for SSD's, which is:
  - Efficient: reduce the number of cache destages; reduce the overhead on the FTL.
  - Adaptive: automatically adapt to the workloads.

Motivation: Random Write Problem

- Low programming rate
- Write Amplification:
  - Garbage collection process
- Random vs. Sequential:
  - Log block consumption
    - Repeated writes
    - Non-sequential writes
  - Overhead: Switch/Partial merge vs. Full merge
- What cache can do:
  - Absorb repeated writes: eliminate Full merge
  - Make the workload more “sequential”: reduce the # of destages (batch); reduce the overhead of Partial merge.

Current Solutions and Their Drawbacks

- Cache Schemes for SSDs: BPLRU, FAB, CLC
  - Common Point: Page Cluster list (e.g., block size is 64; page 64~127 belong to block 1; workload 1, 2, 5, 34, 76, 10, 89, 128, will make three clusters)
  - Why: Batched outputs (sequential).
  - Differences:
    - BPLRU: Recency (LRU)
    - FAB: Size
    - CLC: Hot clusters in BPLRU-order; Cold ones in FAB-order
  - Drawbacks:
    - BPLRU: efficiency of LRU
    - FAB: early eviction
    - CLC: non-adaptive-ness on differentiating Hot and Cold clusters

Current Solutions and Their Drawbacks

- Drawbacks:
  - Data structure: pure clusters may cause a waste of cache space if hot pages go with cold blocks.
  - Temporal vs. Spatial
    - Hot pages vs. Cold cluster
- "Hot page in Cold cluster" is popular in real-world traces.

BPAC Design: the Techniques

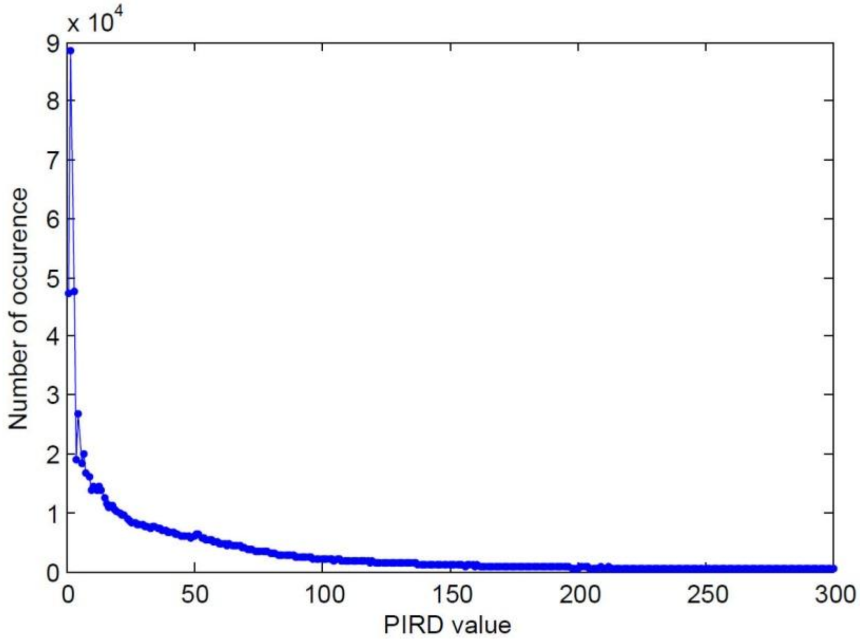
- Dual Lists:
  - Page list: pages with high temporal locality
  - Block list: contains page clusters with low temporal but high spatial locality, and clusters with low spatial locality will be evicted according to their size.

BPAC Design: Adaptive Partitioning

- P-list and B-list: how to determine whether a page is no longer hot?
- Hot clusters (high spatial locality) and Cold clusters (low spatial locality): how to determine a cluster is out of spatial locality?
- Solution: exploit the regularities of the history
  - The duration of temporal or spatial locality of a unit
- Locality Metrics:
  - Temporal: PIRD, Page-level Inter-Reference Distance
    - The distances between consecutive references to the same page.
    - Why: time interval between repeated accesses
  - Spatial: BIRD, Block-level Inter-Reference Distance
    - The distances between consecutive references to the same cluster.
    - Why: BIRD records the time interval between the time two pages join the cluster.

Analysis on PIRD Distribution

- PIRD distribution:
  - Observation: with a small threshold (250), the most of the distribution can be covered.
  - Most re-access distances of hot pages (PIRD sequences) fall in a short range (0~250).
  - If a hot page has not been accessed for 250 (PIRD_thd), the chance that there will be more accesses to the page in the future can be considered minimum. And thus, this page is no longer considered "hot", and is removed from P-list.
- BIRD_thd: the same idea



Evaluation and Results: Methodology

- Methodology
  - Trace-driven simulation
- Traces:
  - OLTP workloads: Financial 1 and 2
  - HP Cello99: disk3 and disk8
- Cache size: 8~128MB
- Comparison: BPLRU, FAB and CLC(optimal)
- Performance Metrics
  - Eviction Count (# of destages)
  - Average size of evicted/destage clusters
  - Overhead (time) on FTL

Conclusions

- Conclusions:
  - By fully exploiting the feature of SSD's, BPAC is presented as an adaptive cache scheme, which is consisted of a hybrid data structure, mechanisms of adaptive partitioning for cache space, and advanced replacement policy for heterogeneous access patterns.
  - BPAC reduces the number of evictions and increases the size of destages which in turn reduces the overhead on the FTL.
  - BPAC outperforms the existing solutions.

Acknowledgment

- This research is sponsored in part by National Science Foundation grants CCF-0937799 and CNS-0720617.

Questions?