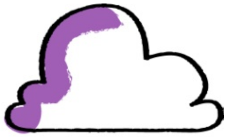
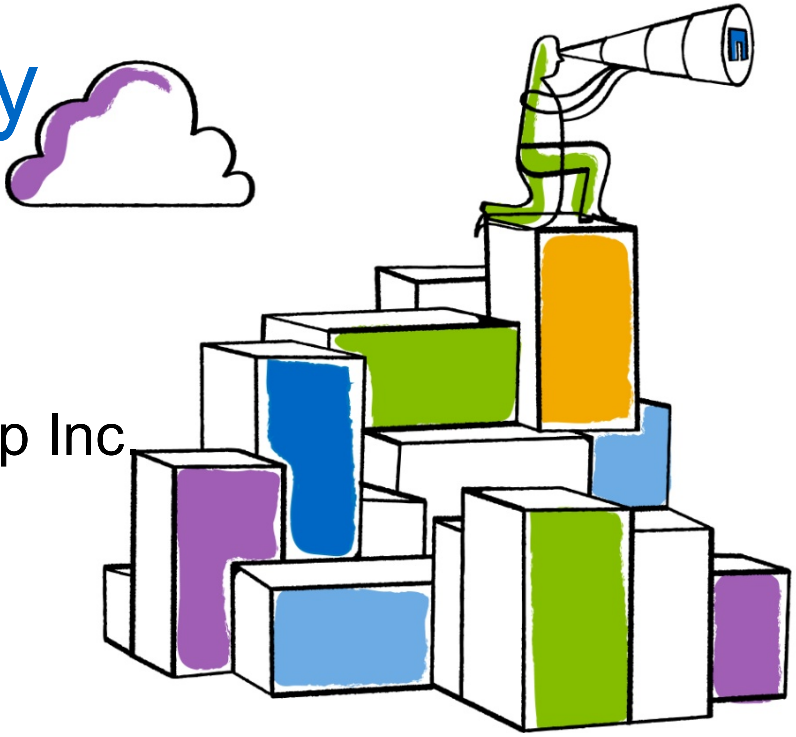
Go further, faster®



A Deduplication Study for Host-side Caches

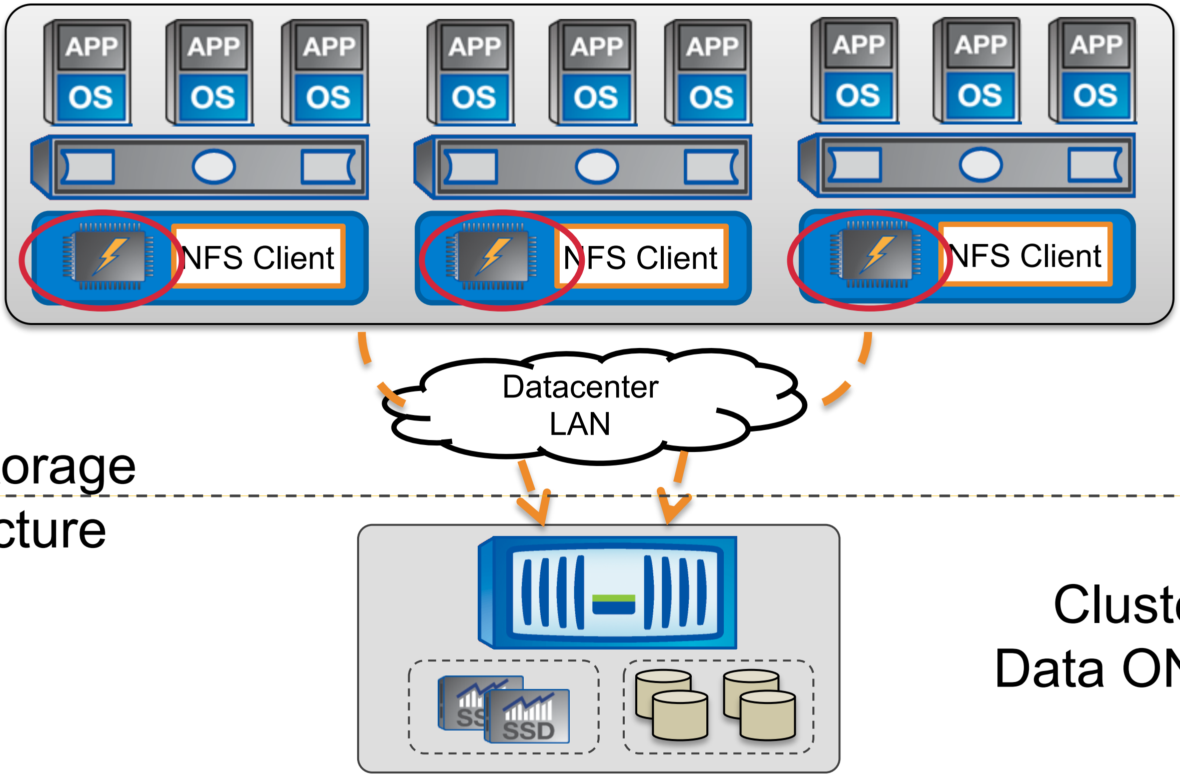


Jingxin Feng, Jiri Schindler

Advanced Technology Group, NetApp Inc.

May 9th, 2013

Virtualized Data Centers



Flash Accel Cache

Hosts

Shared Storage Infrastructure

Clustered Data ONTAP®

Single host-side cache contains data from many VMs running similar workload

Problems in Virtualized Environment

- Host-side caches may not be efficient in a dynamic environment
  - Working set changes
  - VM migrates
- Virtualized Desktop Infrastructure (VDI)
  - VMs with similar contents and workloads

Goal: Improve cache hit rates for host-side caches integrated with VM hypervisors

Summary of Our Work

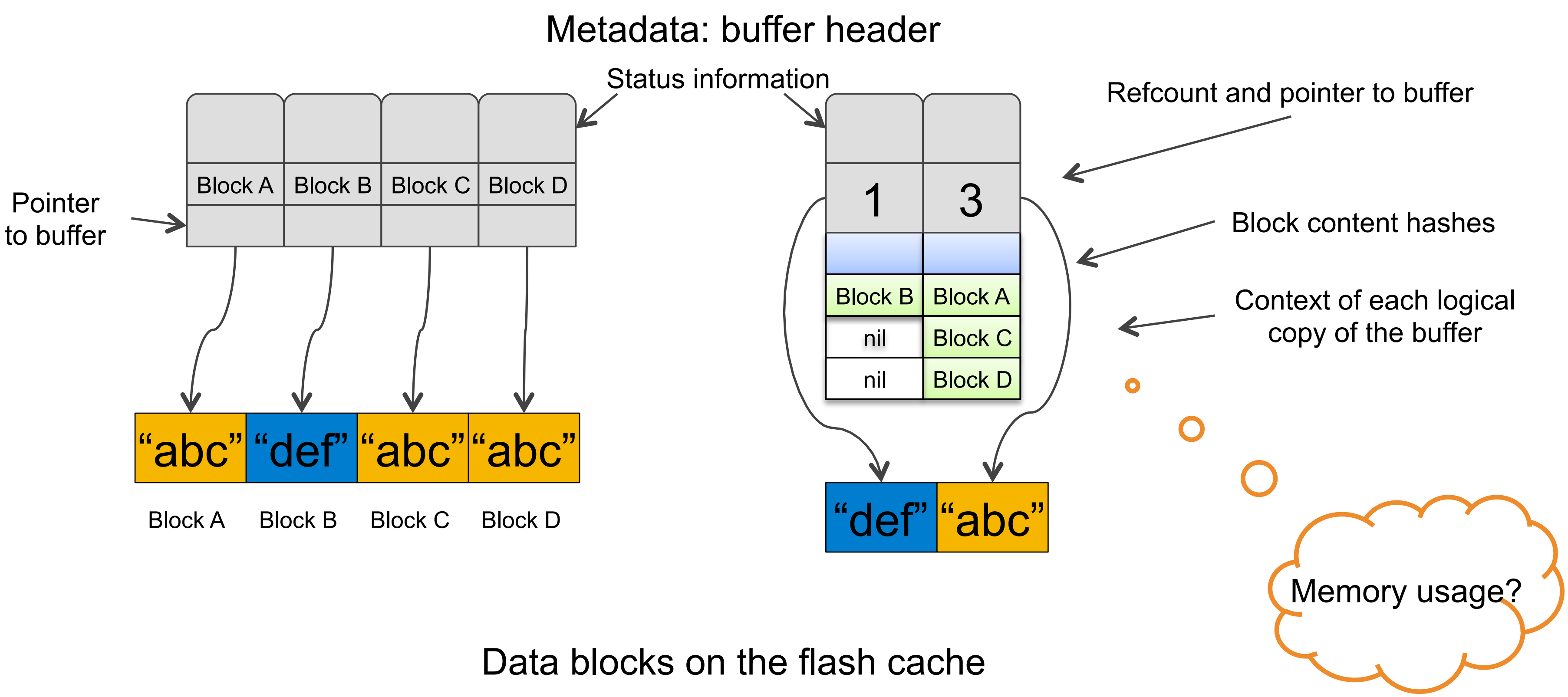
- Study traces of real-world workloads
- Evaluate the opportunity for deduplication
- Analyze improvements to cache efficiency
- Provide design hints for deduplicating caches

Our Goal: A deduplicating cache design

Tradeoff between deduplication and metadata overhead

Standard cache

A deduplicating cache



Metadata: buffer header

Status information

Refcount and pointer to buffer

Pointer to buffer

Block A | Block B | Block C | Block D

1 | 3

Block content hashes

Block B | Block A

nil | Block C

nil | Block D

Context of each logical copy of the buffer

“abc” “def” “abc” “abc”

Block A Block B Block C Block D

“def” “abc”

Memory usage?

Data blocks on the flash cache

Approach

- Analyze real world workloads
  - Two long-term CIFS traces
  - Six VDI traces of update, boot and login storms
- Explore intrinsic properties of traces regardless of cache size and specific replacement policies
- Provide an upper bound of cache space saved by deduplication

Metrics: Deduplication Degree

- Deduplication Degree (d)
  - Defined as the average number of links of data blocks
    - d = 1: no duplicates
    - d = 2: on average, each block has one duplicate
- The metric (d) directly expresses the average number of references to each data block
  - Estimates metadata overhead

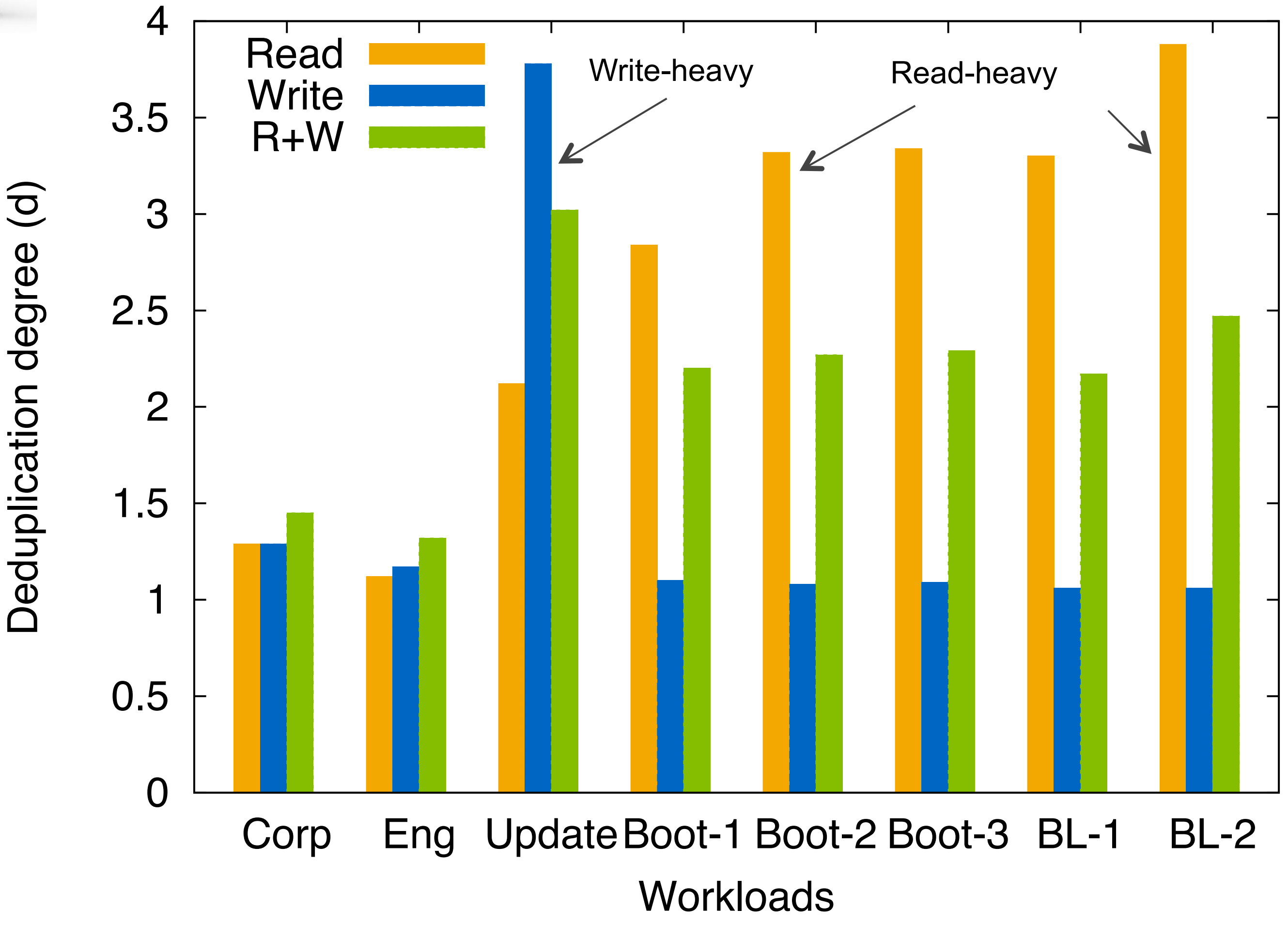
Corp & **Eng**: enterprise traces; **Update**: VDI update storm trace;

Boot-1,2,3: VDI boot storm traces; **BL-1,2:** VDI boot and login storm traces

NetApp

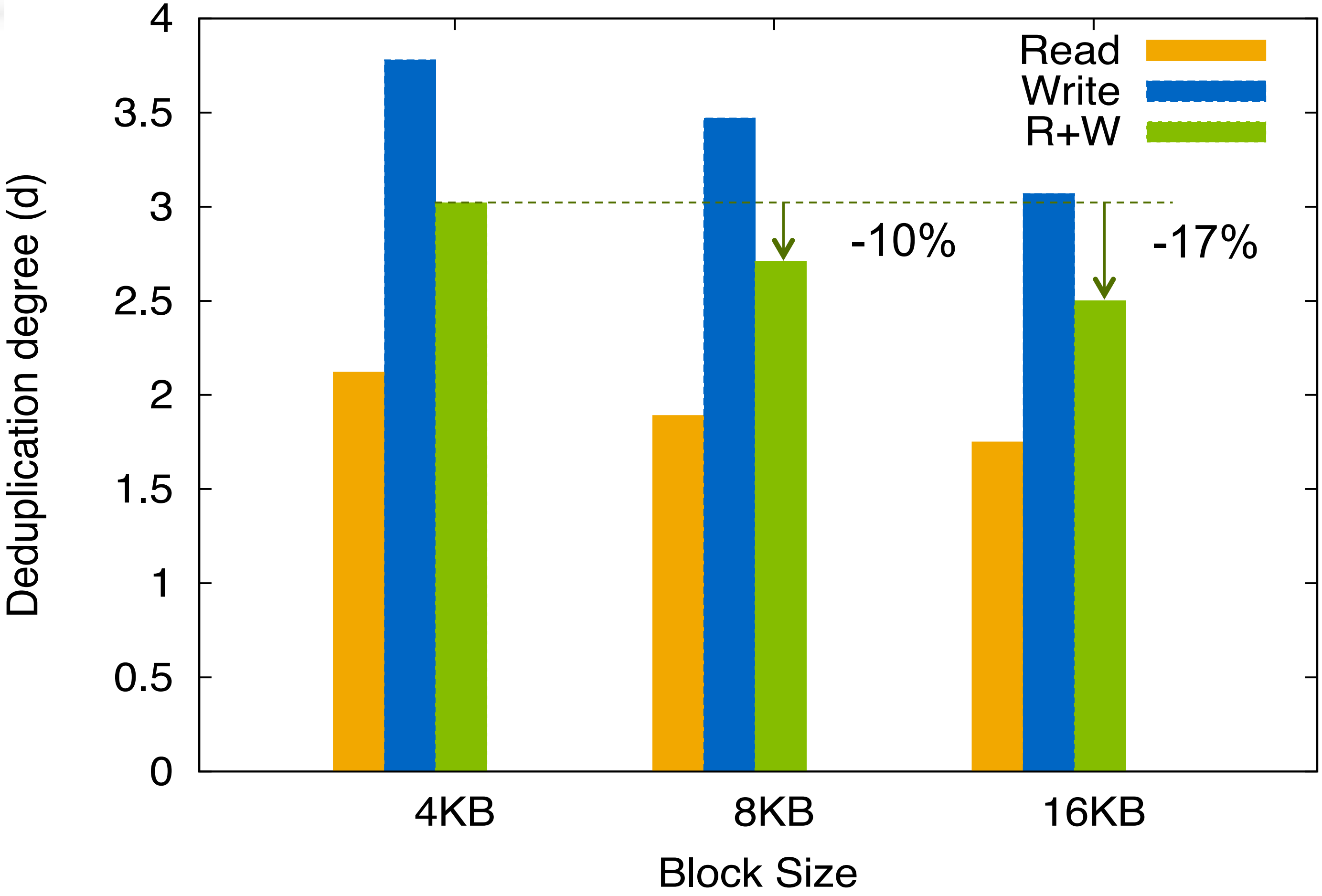
Deduplication Degree

Do opportunities for content duplication exist? How much?



Sensitivity to Cache Block Size -- Update

If we keep less metadata, how much deduplication will we lose?



Space savings: Fixed vs. Variable length

Do we design the cache as fixed or variable block size?

- Variable-length chunking
  - Believed to save more space
  - Increase complexity of cache design
- Fingerprints
  - Fixed: 4 KB block size
  - Variable chunking: Rabin fingerprints
    - min 2 KB, mean 4 KB, and max 16 KB; sliding window 48 bytes.
- Additional savings (0-15%) are achieved using variable-length chunking.

Conclusion: Design Hints

- Deduplication could improve the cache hit ratios in a dynamic environment
- Fixed-length chunking is good enough in terms of deduplication degree and implementation complexity
- Users need to make a tradeoff between deduplication ratios and metadata overhead.

NetApp

Thank you