CIMStore

Content-Aware Integrity Maintaining Storage

Charles B. Morrey III and Dirk Grunwald
University of Colorado, Boulder
Department of Computer Science

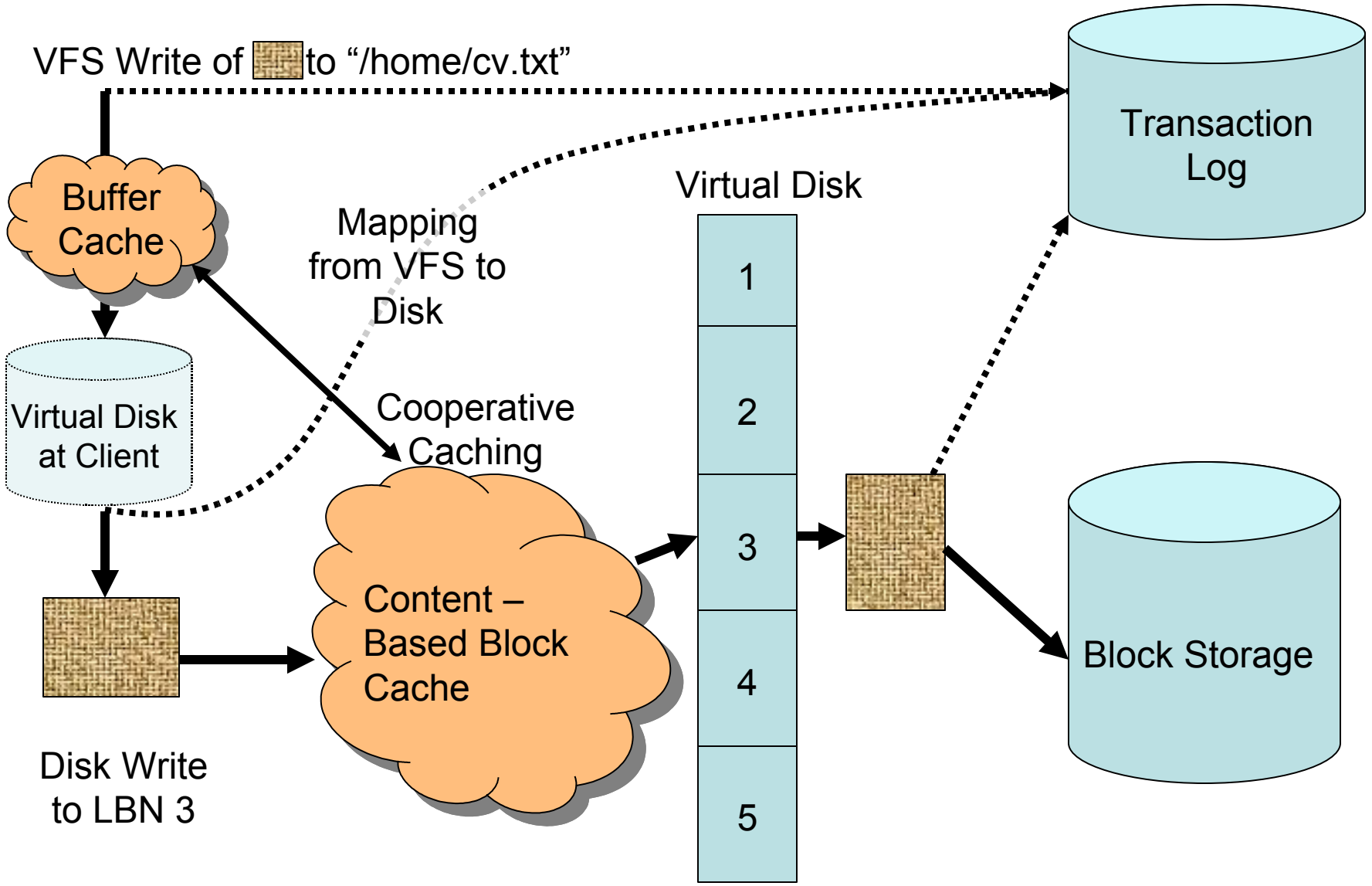
Thesis

CIMStore provides integrity maintenance for storage consumers in the form of “non-overwriting” block storage. To make “non-overwriting” storage feasible, *content-based coalescing* helps reduce space requirements and improves performance.

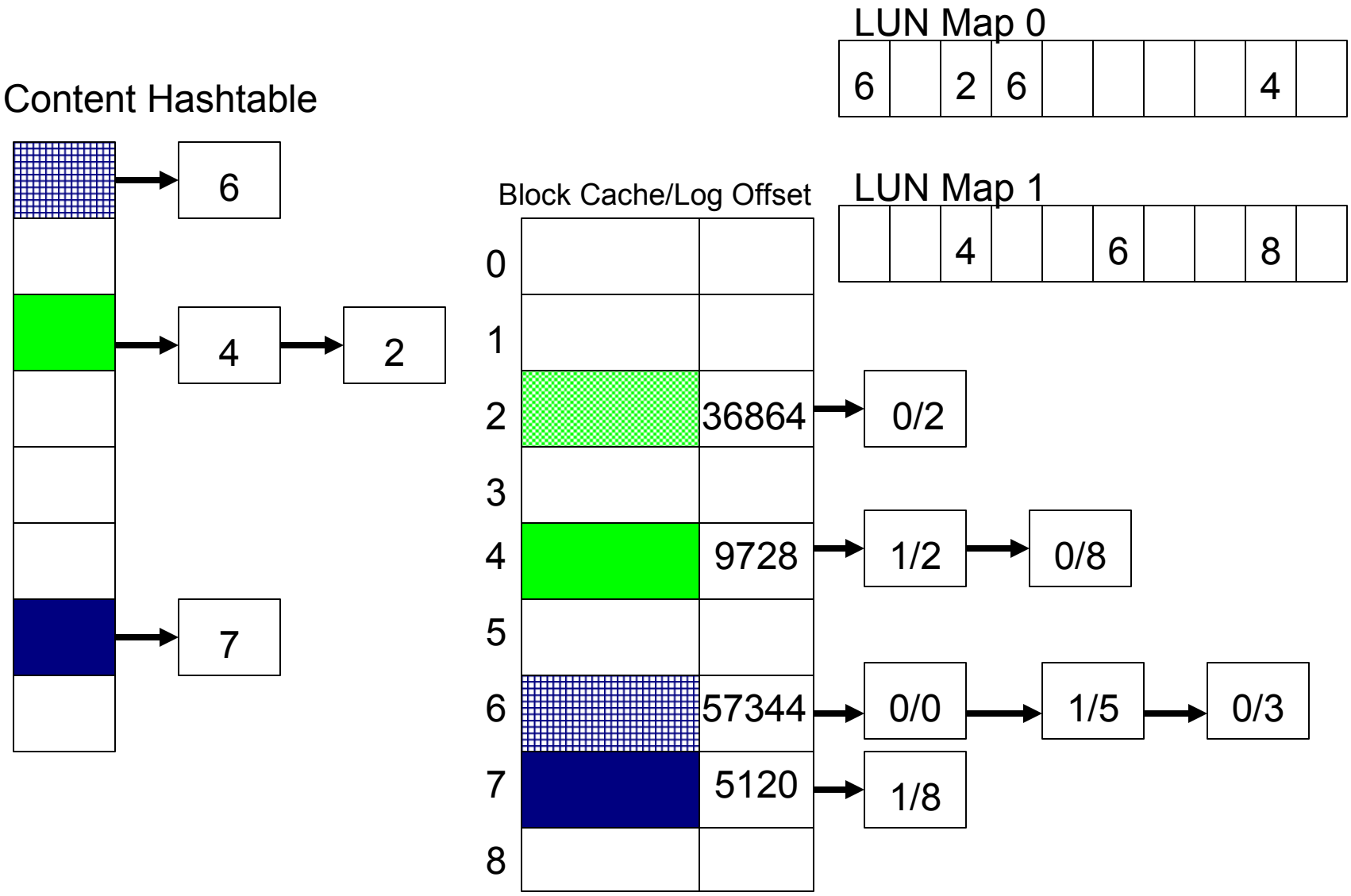
Motivation

- Why bother with non-overwriting block storage?
  - Exposes history of all data for search/browsing
    - Great Security Audit tool
    - Novel mechanism for versioning/temporal storage
  - Additional benefits
    - Provides an integrated backup solution
    - Potentially improved storage access performance

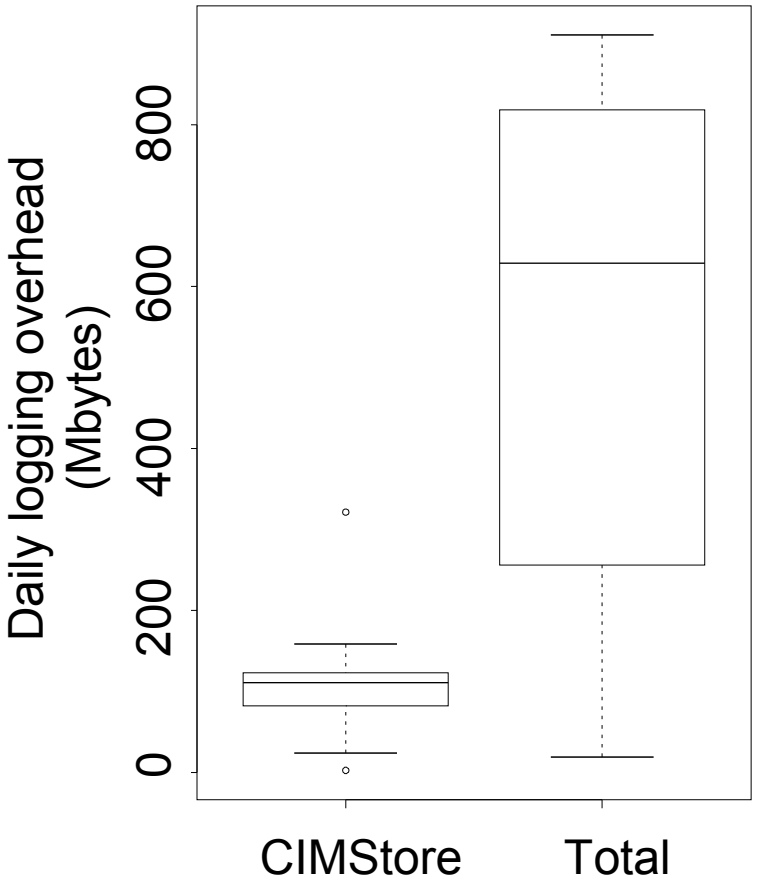
CIMStore System Components

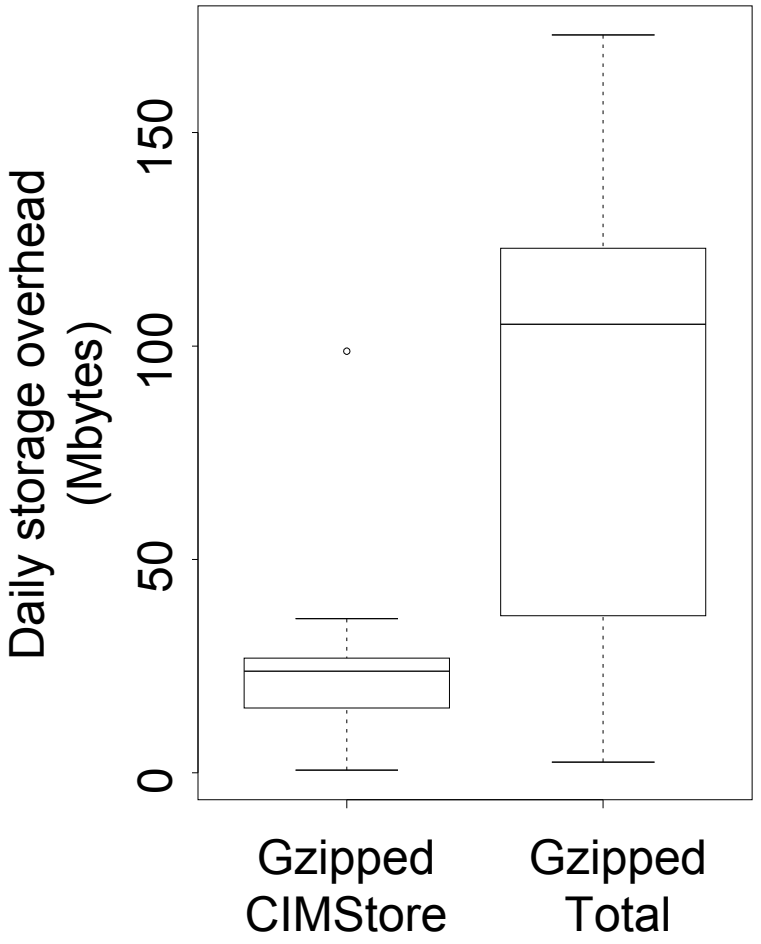


Modified Content-Based Block Cache

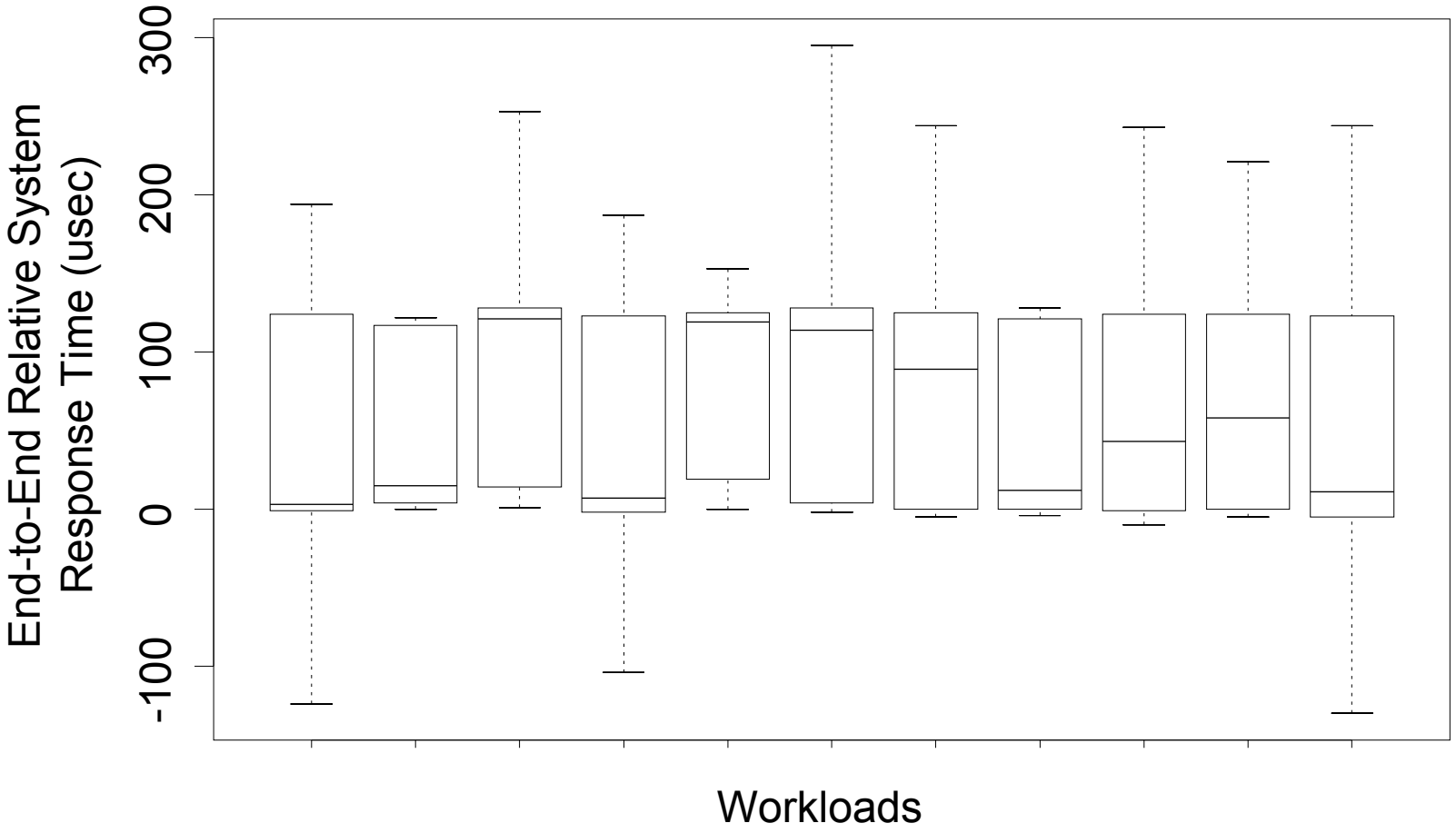


Storage System Overheads





Relative Response Time



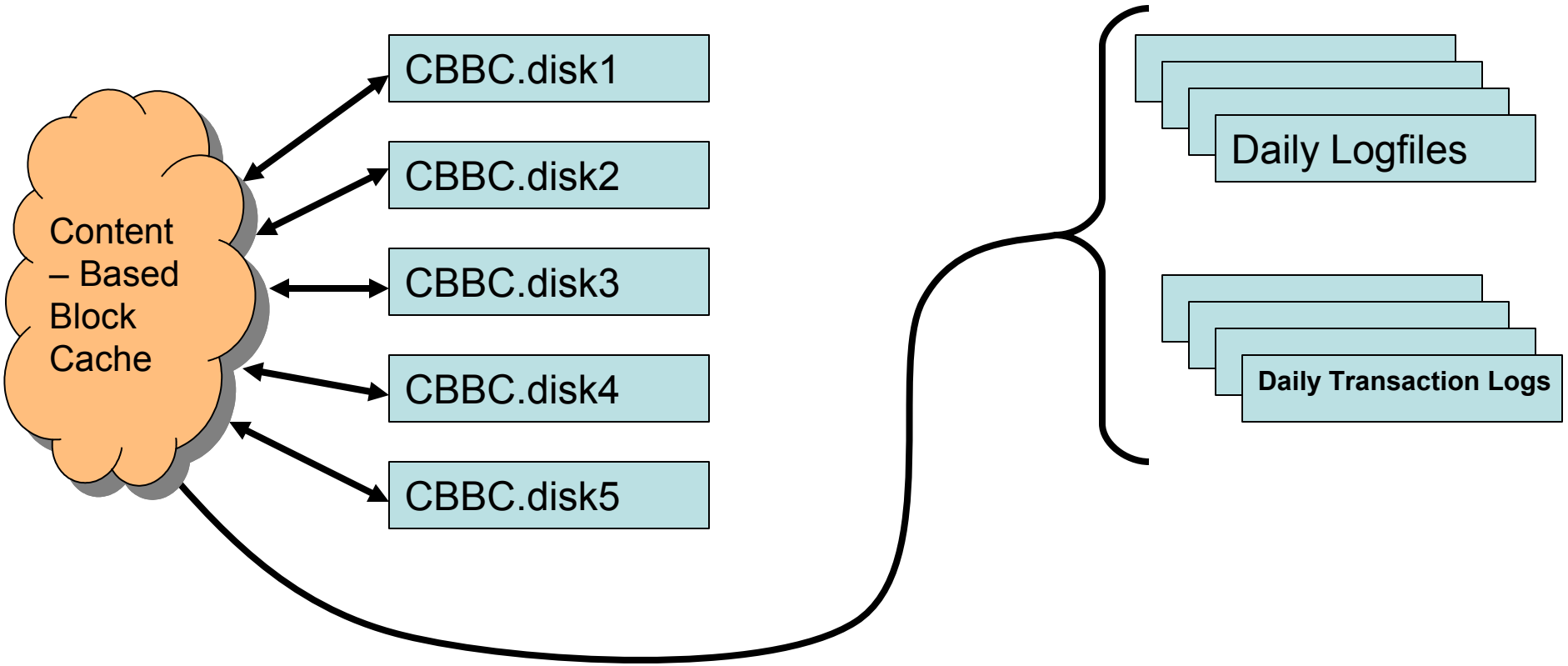
Conclusion and Future Work

- CIMStore saves all blocks with 5% space cost, and small performance overhead.
- Performance Tuning
- VFS-to-Disk Mapping implementation
- Content-Aware Cooperative Block Caching
- http://systems.cs.colorado.edu/~cbmorrey

Backup Slides

-

Content-Aware Block Storage

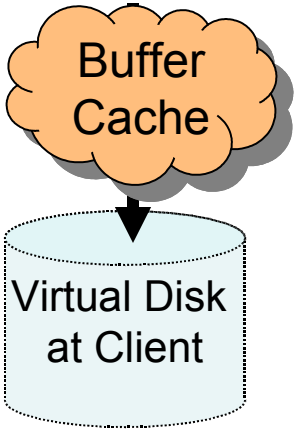


VFS-to-Disk Logging (Proposed)

VFS Write of  to offset 52 of “/home/cv.txt”

VFS Read 5 bytes at offset 2120 from “/home/cv.txt”

VFS Create “/home/jobapp.txt”

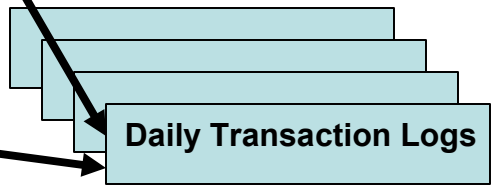


Journal Commit Bounds Associated VFS Operations and Block Modifications

Journal Operation #1

Write of  to offset 52 of LBN 0

Read 5 bytes at LBN 5



Commit Operation #1

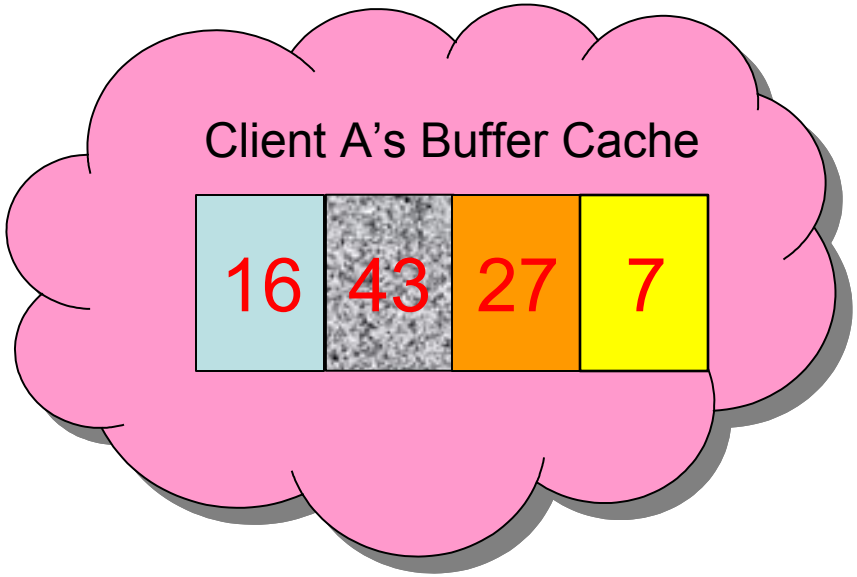
Journal Operation #2

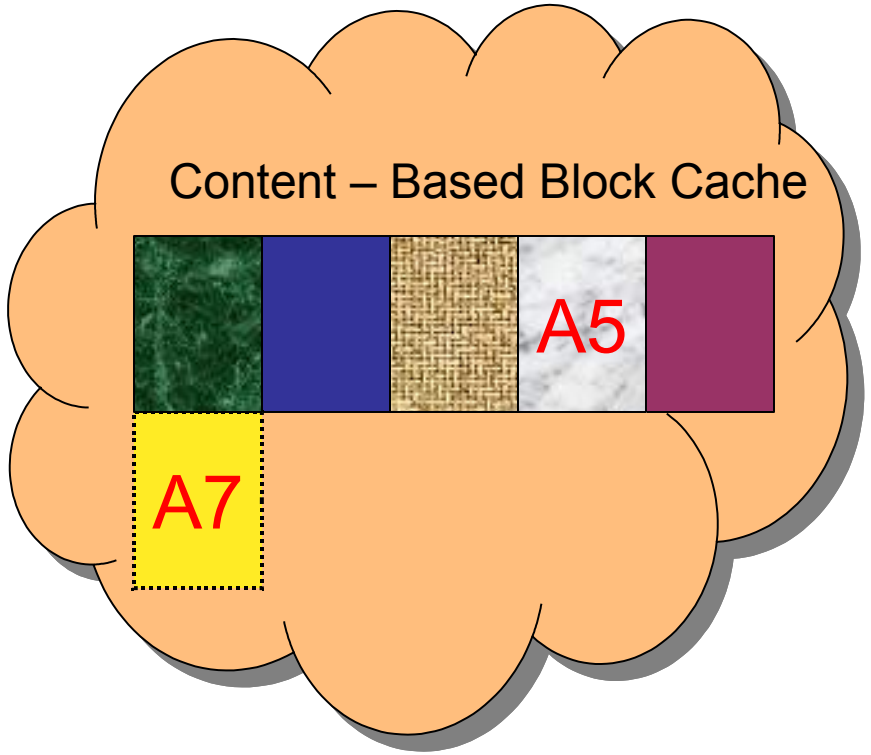
VFS Create

Content Aware Cooperative Block Caching (Proposed)

Write of  to offset 52 of LBN 0

Read 5 bytes at offset 84 LBN 5

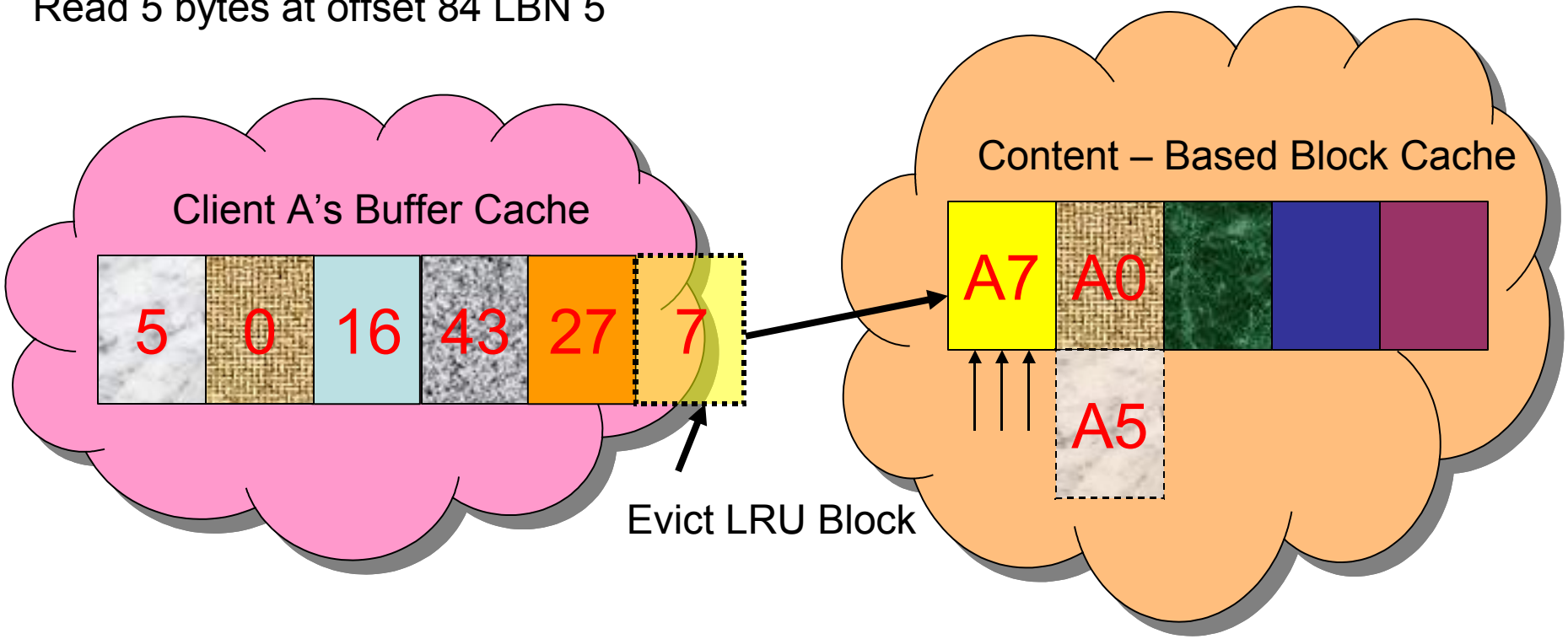




Content Aware Cooperative Block Caching (Proposed)

Write of  to offset 52 of LBN 0

Read 5 bytes at offset 84 LBN 5



Research Artifacts

- Disksim 3.0 Simulator modified to add Content-Based Block Cache (CBBC)
- iSCSI Target modified to add CBBC and log disk operations
- Several months of live system disk traces of Linux (ext3) and Windows (NTFS)

Research Artifacts

- Xen 3.0 server with Cisco iSCSI initiator serving multiple CIMStore-backed partitions
  - Webserver
  - SysAdmin Class Student Server Sandboxes

Definitions

- Non-Overwriting Block Storage
  - Changes to all blocks that reach stable storage (whether via file system flush or direct writes to raw disk) are logged
- Content-Based Coalescing
  - Only storing a single copy of any block content
    - Used both in memory and on disk
  - Uses hashing and other metadata for fast access

CIMStore Components

- Modified Content-Based Block Cache
- Content-Aware Block Storage (Peabody)
- Client-Side VFS-to-Disk operation logging (work-in-progress)
- Content-Aware Cooperative Block Caching (work-in-progess)