Towards Adaptive Replication for Hot/Cold Blocks in HDFS using MemCached

Pinchao Liu, Adnan Maruf, Farzana Beente Yusuf, Labiba Jahan, Hailu Xu, Boyuan Guan, Liting Hu, Sitharama S. Iyengar

Florida International University

FIU | Computing & Information Sciences

ELVES research laboratory

ELVES Research Laboratory

FIU Computing & Information Sciences

ELVES research laboratory

HOME RESEARCH PEOPLE PUBLICATIONS NEWS EVENTS GALLERY CONTACTS US

THIS IS EXPERIMENTAL AND VIRTUALIZED SYSTEM (ELVES) RESEARCH LABORATORY

ABOUT US

Our lab conducts experimental systems research, specifically addressing high performance and enterprise applications on virtualized and distributed platforms; focusing on big data analytics systems, specific research interests include runtime adaptation, methods for online resource management, and system support solving enterprise motivated real-world problems that embody complicated data analytics; focusing on systems virtualization, research interests range from basic technologies to innovative technology application.

JOIN US

Our lab targets training students towards becoming first-class systems researchers and system builders. We are looking for passionate students who are interested in big data systems, systems virtualization and distributed systems to join us!

RECENT PUBLICATIONS

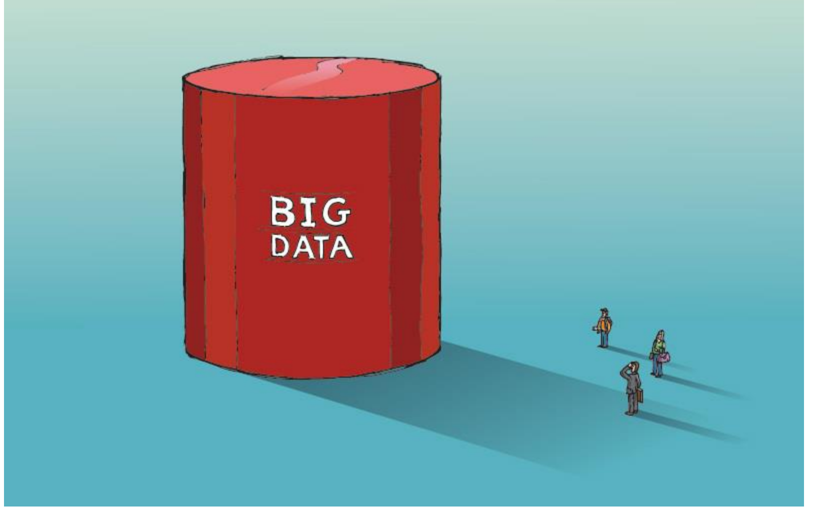
- CLOUD'18: "Oases: An Online Scalable Spam Detection System for Social Networks"
- CLOUD'18: "A Toolset for Detecting Containerized Application's Dependencies in CaaS Clouds"
- ICAC'17: "GOVERNOR: Smoother Stream Processing Through Smarter Backpressure"

Outline

- Problem Description & Challenges
  - Problem description
  - Challenges
- System Design
  - The basic workflow
  - Approach implementation
- Experimental Evaluation
  - Experiments showing functionality
  - Experiments showing improves application performance
- Conclusion & Future Works

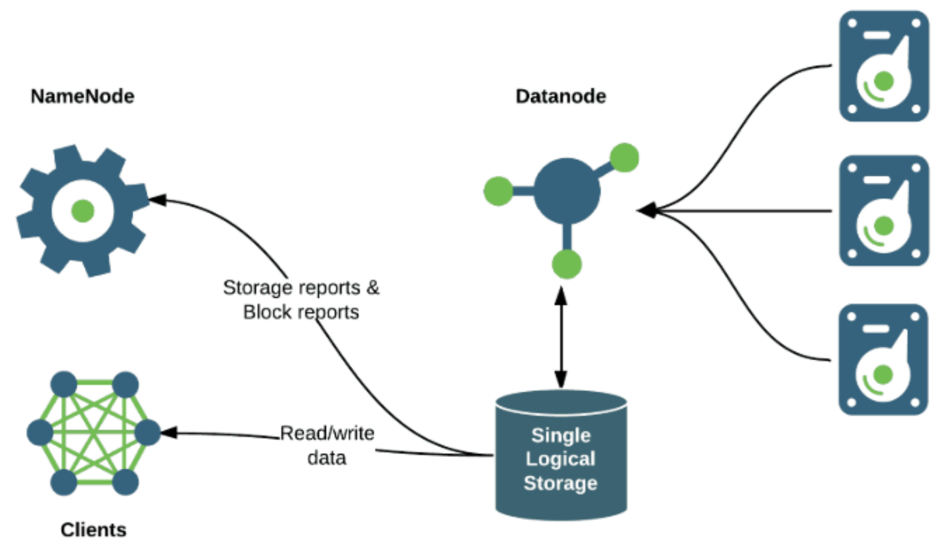
Problem Description

- Big Data Becomes A Big Topic
  - Multiple industries use big data ranging from
    - Banking
    - Healthcare
    - Energy
    - Technology
    - Consumer
    - Manufacturing
    - etc.



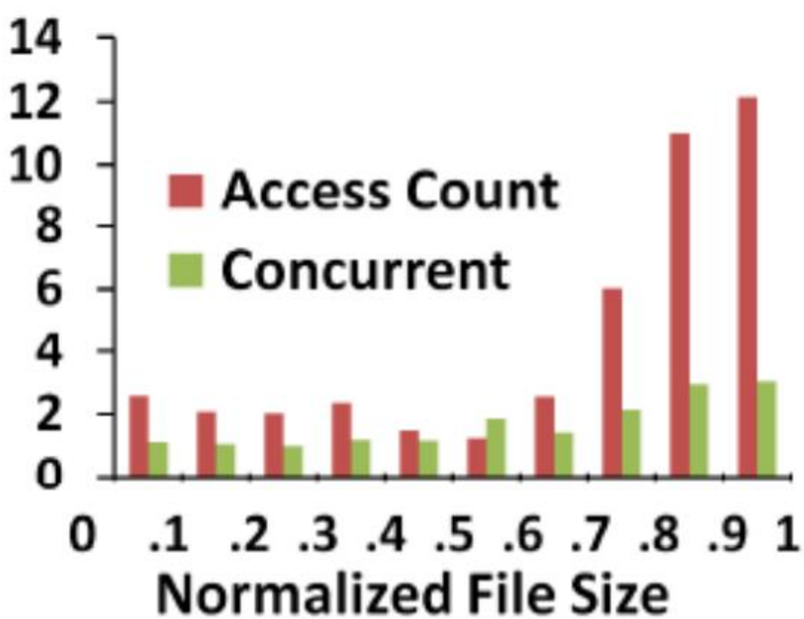
Problem Description

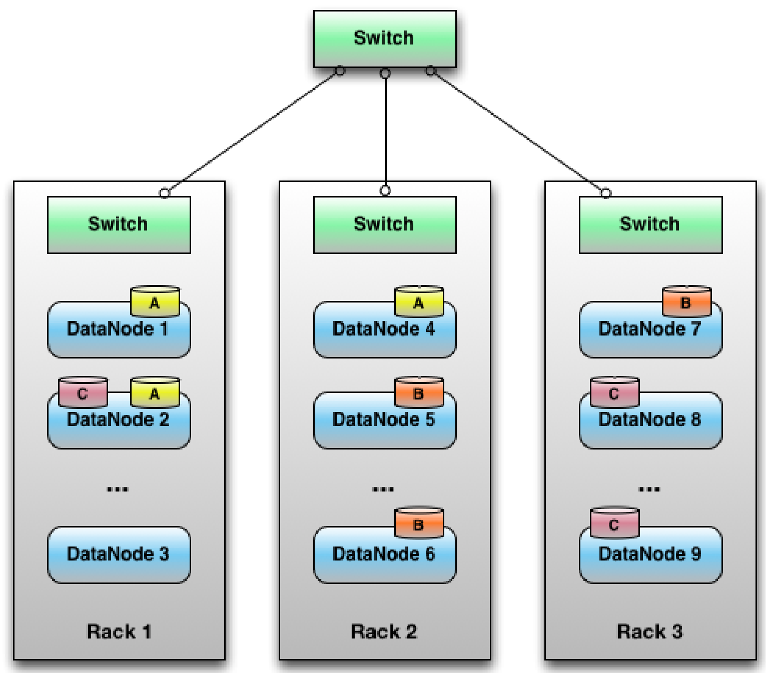
- HDFS is the Major Player in Big Data Storage
  - Low cost per byte
  - High bandwidth to support MapReduce workloads
  - Solid data reliability



Problem Description

- Skewed Popularity of Data Access
  - A Review on Hadoop from Yahoo!
  - Scarlett from Microsoft





Problem:

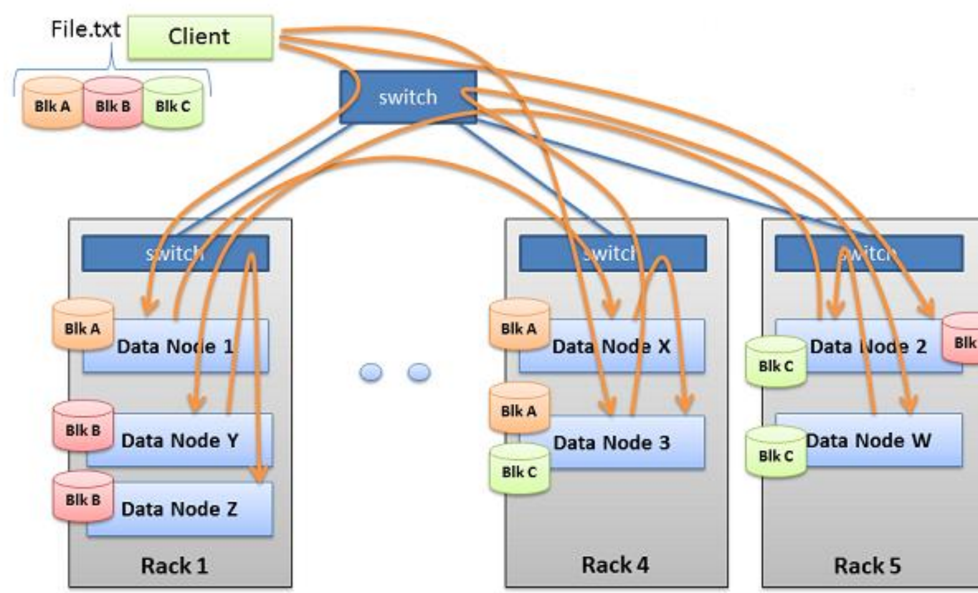
Create a strategy to cater to the skewed popularity in HDFS?

Key Challenges

- Reduce resource contention.
  - Network resource contention.
  - Disk resource contention.

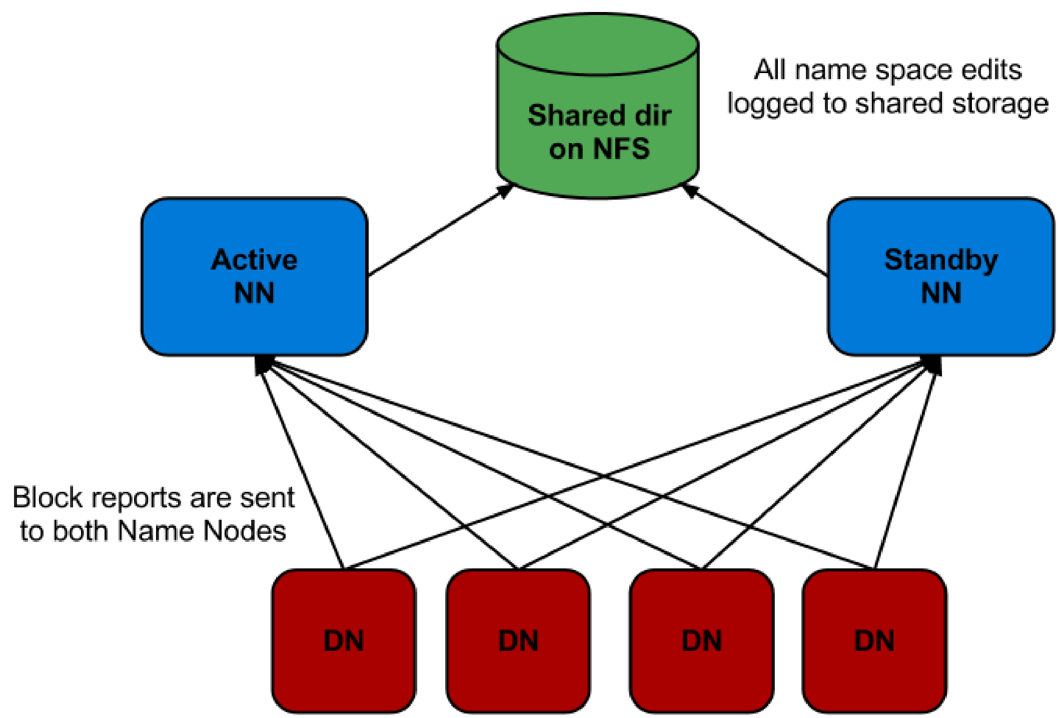
Key Challenges

- Reduce response time.
  - A typical DRAM has a transfer rate of 2-20GB/s, whereas typical SSDs have a transfer rate of 50MB-200MB/s. It's one to two orders of magnitude slower.



Key Challenges

- Increase Data Availability
  - Avoid single point of failure

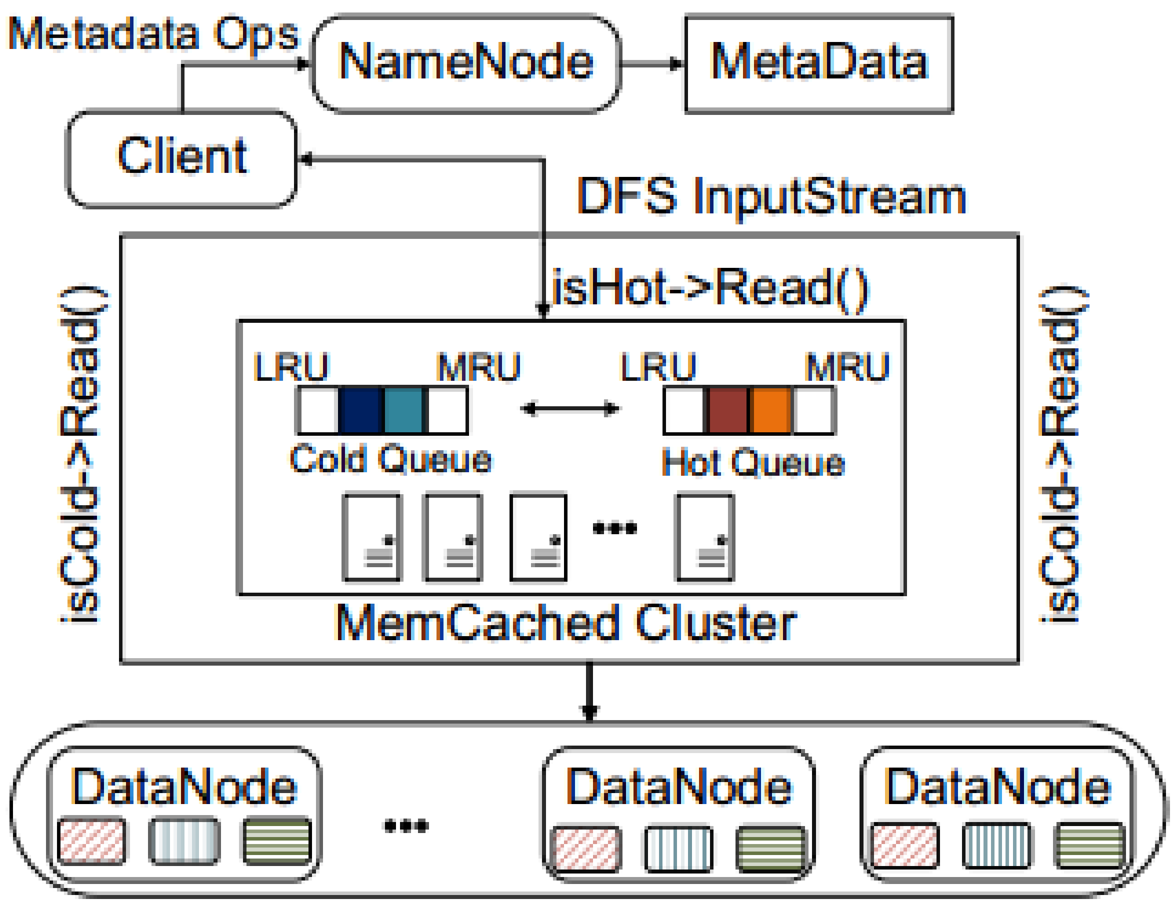


Challenge:

The system needs to have the properties of reducing resource contention, reducing response time and keep the availability.

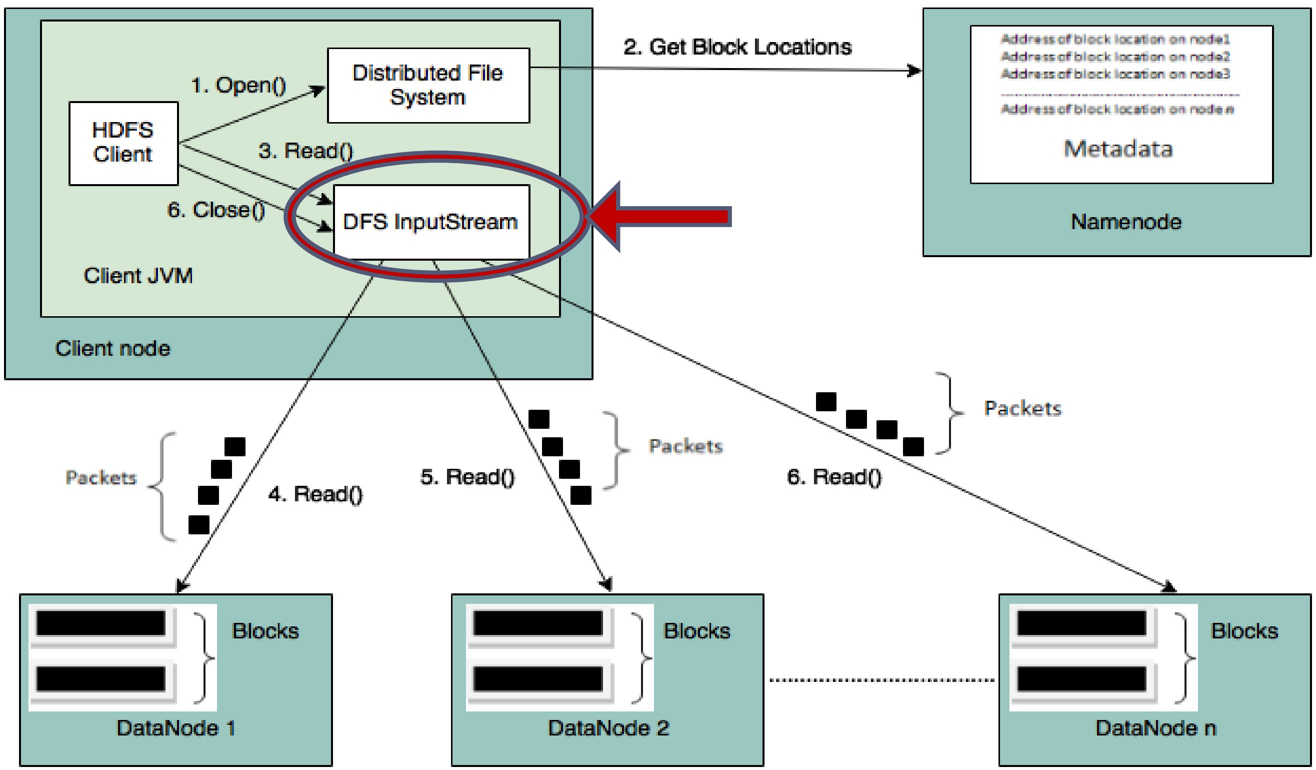
System Design

- Integrate Memcached with HDFS



System Design

- Integrate Memcached with HDFS



System Design

Vanilla Replication Algorithm

- Set threshold access value to get in/out Memcached.

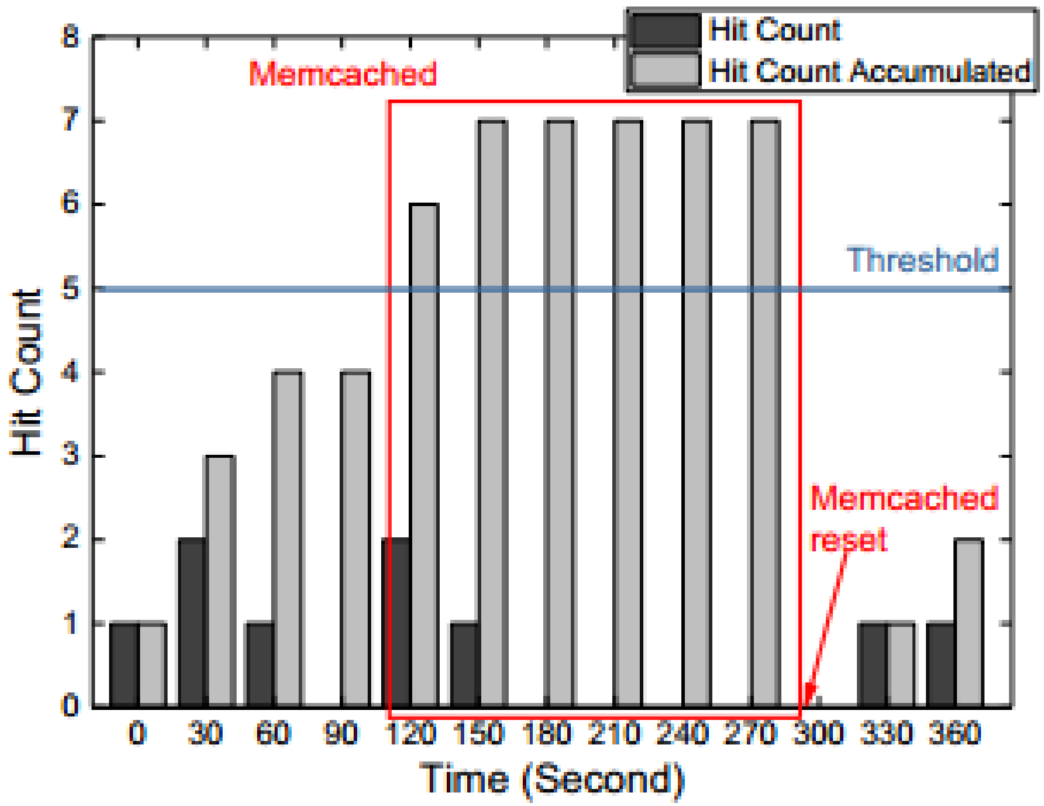
Set key := CurrentLocatedBlock.BlockID
IF (key exist in map)
Increment counter
ELSE Set counter

IF (key counter value > threshold)
Hot block detected!
Set in Memcached
ELSE Normal operation

< BlockID, counter>
<454768378, 1>
<563728272, 5>
<787365473, 3>

System Design

- Vanilla Replication Algorithm
  - Set threshold access value to get in/out Memcached.



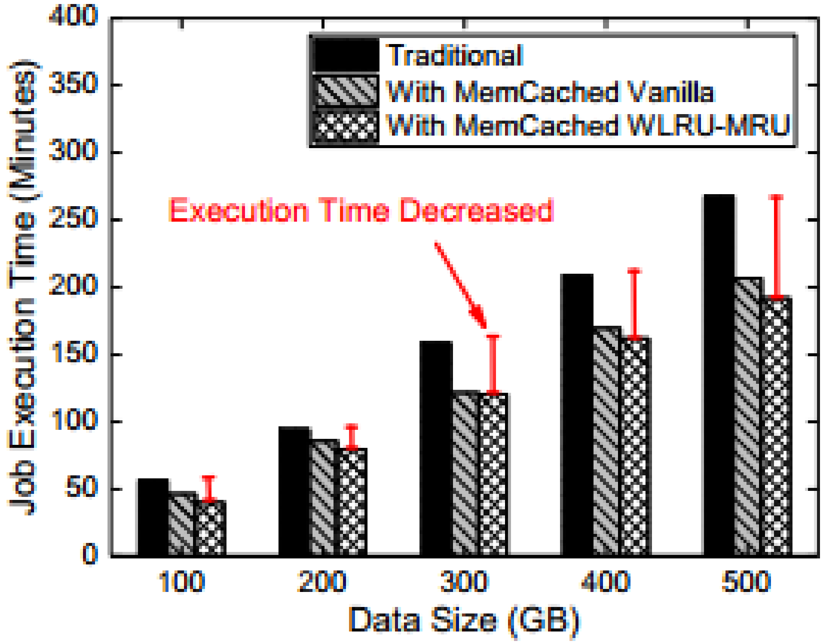
System Design

Dynamic Collaborative Replacement

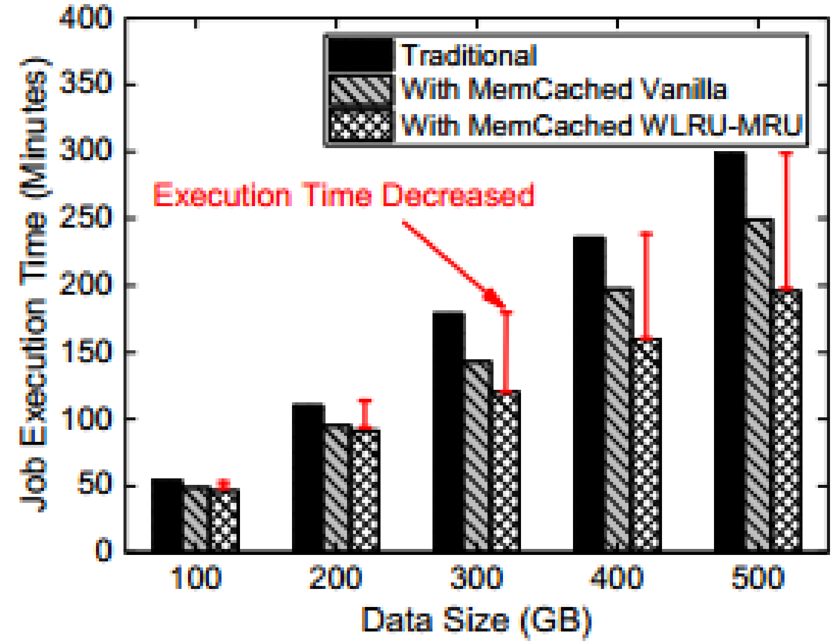
Algorithm 1: Dynamic collaborative replacement algorithm

```
Input: Label M, Label N
if (M==0) //datablock is not in MemCached. then
    if (N==0) //tag hit is 0 then
        LRU; //call LRU
        Replace(bottom); //replace the data at bottom
        MoveDown(other data); //other data move down
          in turn
    if (N==1) //tag hit is 1 then
        MRU;
        //call MRU
        Replace(top); //replace the data at top
        MoveUp(other data); //other data move up in
          turn
if (M==1) //datablock is in MemCached. then
    if (N==0) //tag hit is 0 then
        MRU; //call MRU
        minHeap.put(<hitData.id, hitData.weight++>);
        MoveUp(other data); //other data move up in
          turn
    if (N==1) //tag hit is 1 then
        LRU; //call LRU
        maxHeap.put(<hitData.id, hitData.weight++>);
        MoveDown(other data); //other data move down
          in turn
```

Functionality Evaluation



(a) Word count execution time.

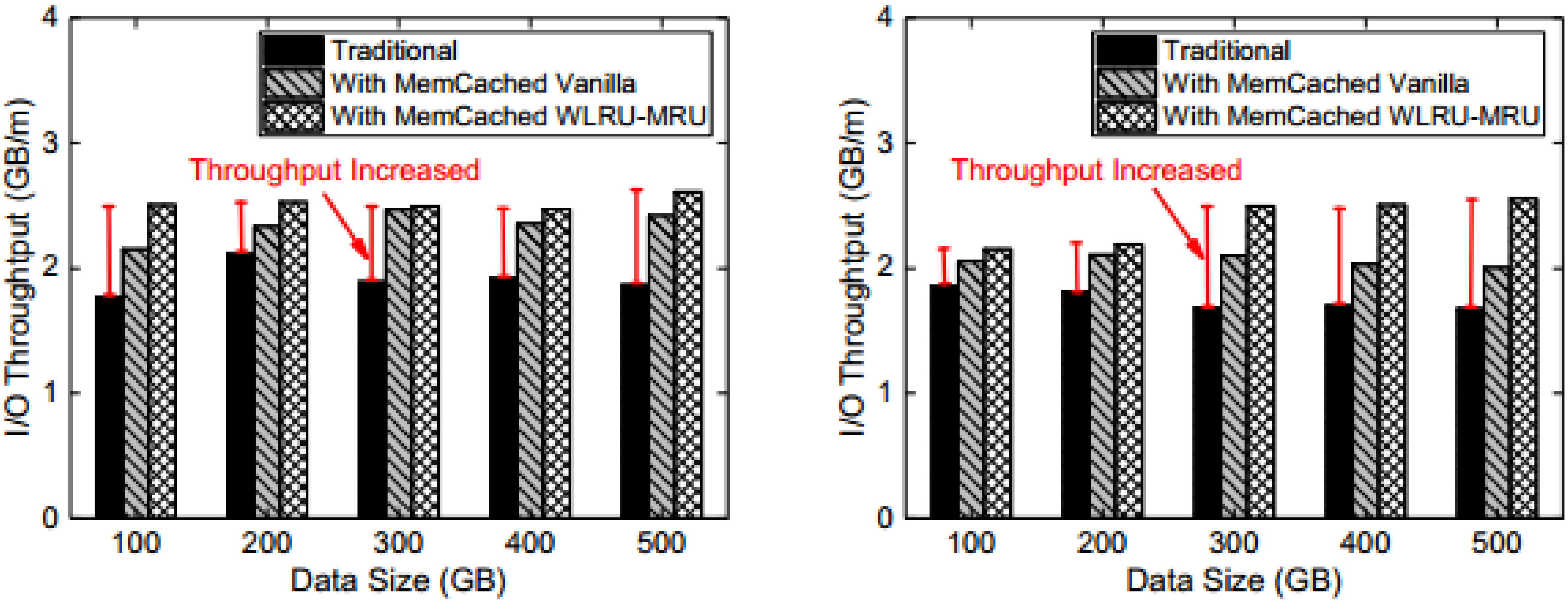


(b) Grep execution time.

Figure 4. Job execution time for word count and grep.

Functionality Evaluation

FIU Computing & Information Sciences

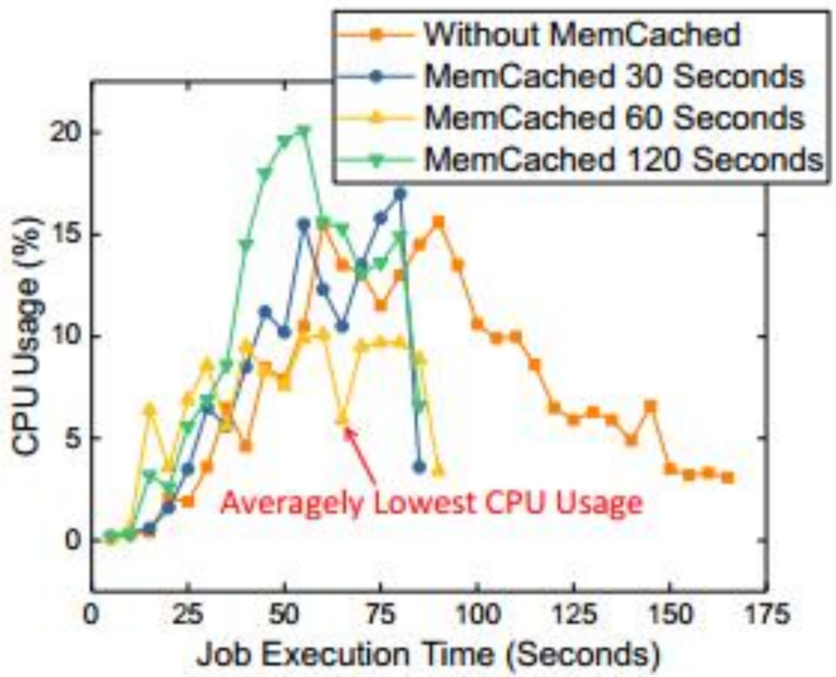


(a) Word count I/O throughput.

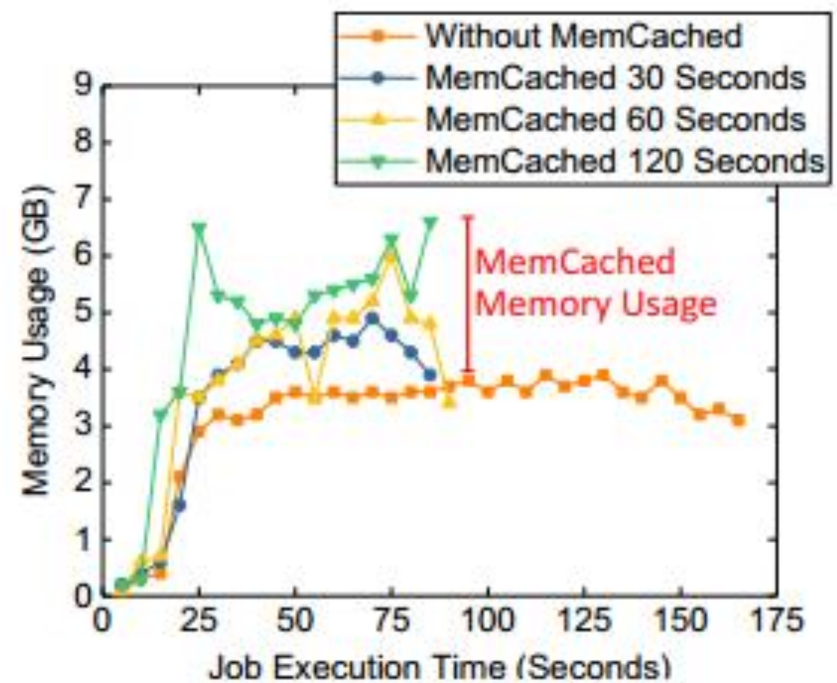
(b) Grep I/O throughput.

Figure 5. I/O throughput for word count and grep.

Overhead Evaluation



(a) CPU usage comparison.



(b) Memory usage comparison.

Conclusion

The designed system integrates Memcached successfully into HDFS, with the improved performance evaluation results.

Contributions compared to related work

- The designed system can benefit for the skewed data access pattern.
- The designed system can dynamically adjust the hot/cold blocks.
- The designed system can keep the high availability property.

Future Work

- Evaluate and manage the distributed file system at cloud-scale.
- Explore more efficient ways to adjust the blocks such as on-line learning method.
- Explore the ways to expend the strategy to other distributed file system.

Thank You.

Pinchao Liu
pliu002@fiu.edu

FIU | Computing & Information Sciences

ELVES research laboratory

Towards Adaptive Replication for Hot/Cold Blocks in HDFS using MemCached

Pinchao Liu, Adnan Maruf, Farzana Beente Yusuf, Labiba Jahan, Hailu Xu, Boyuan Guan, Liting Hu, Sitharama S. Iyengar

Florida International University

FIU | Computing & Information Sciences

ELVES research laboratory