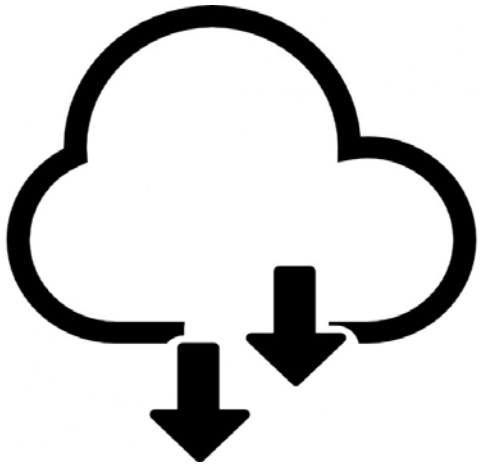
CStorage: An Efficient Classification-based Image Storage System in Cloud Datacenters

Haiying Shen, Heng Zhou^*

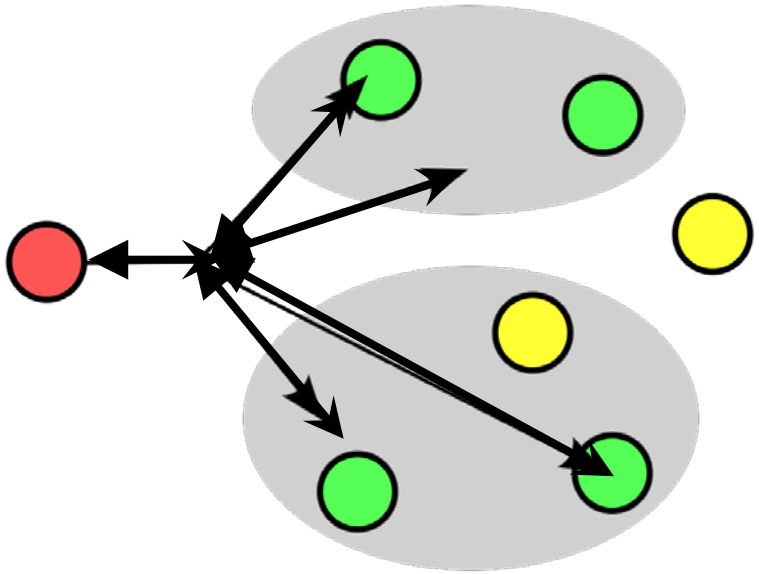
*University of Virginia, ^Clemson University

Outline

- Introduction
- Approach description
- Evaluation
- Conclusion



Incast Congestion



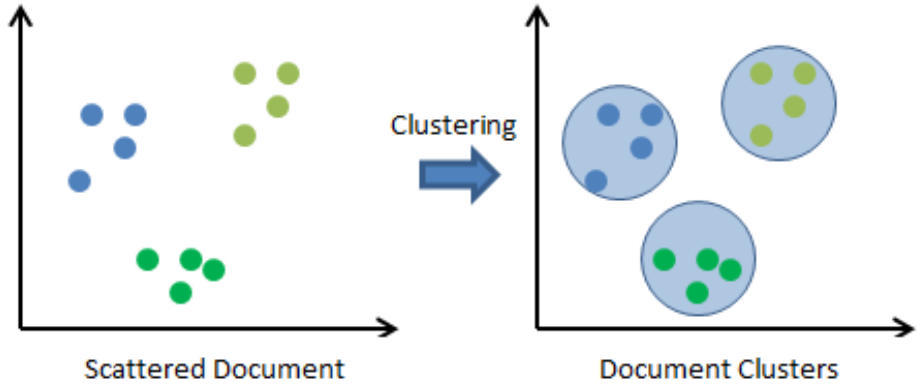
Incast is a many-to-one communication pattern commonly found in cloud data centers. It begins when a singular parent server places a request for data objects to a large number of servers simultaneously.

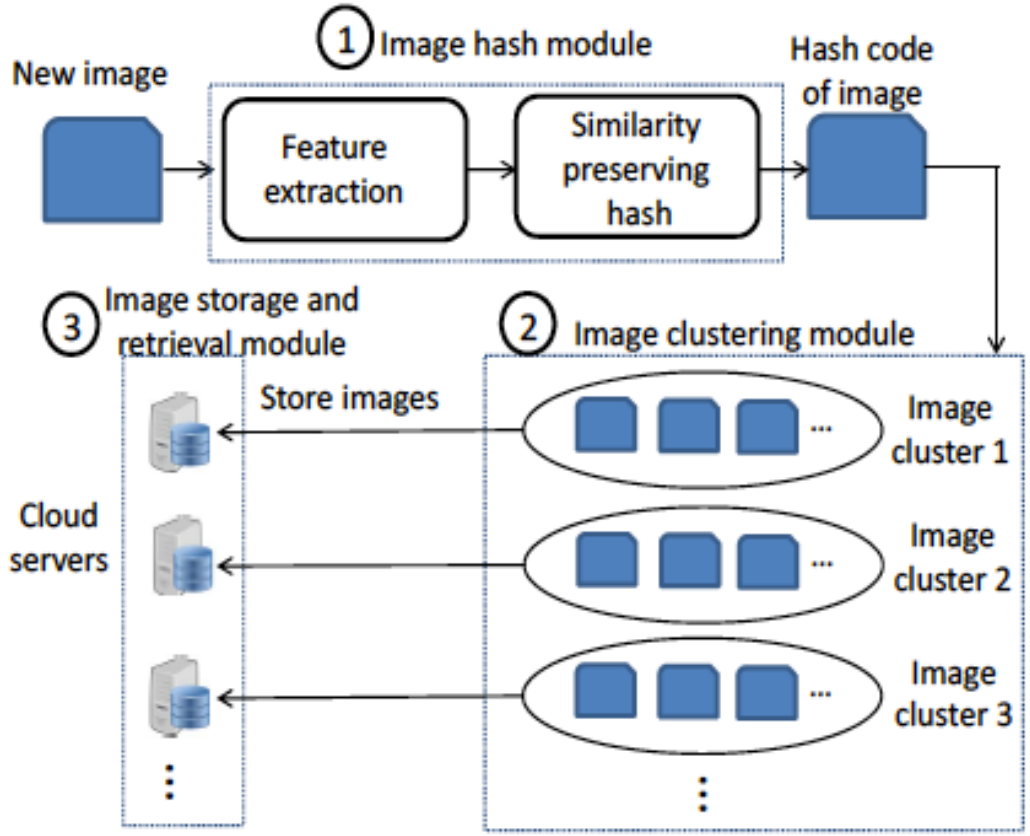
Outline

- Introduction
- Approach description
- Evaluation
- Conclusion

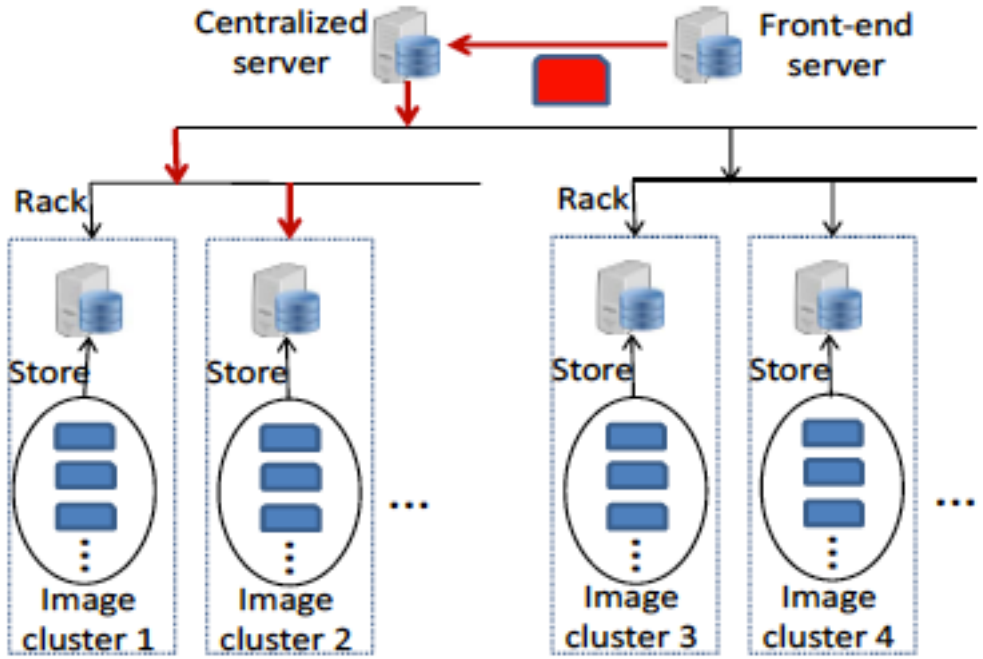
How to?

Image clustering





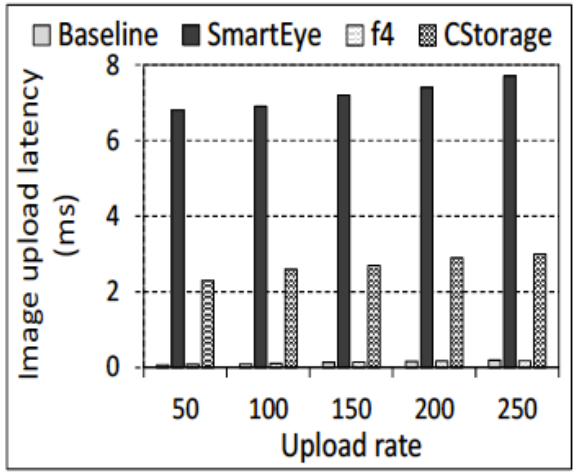
Overview of the CStorage image storage system

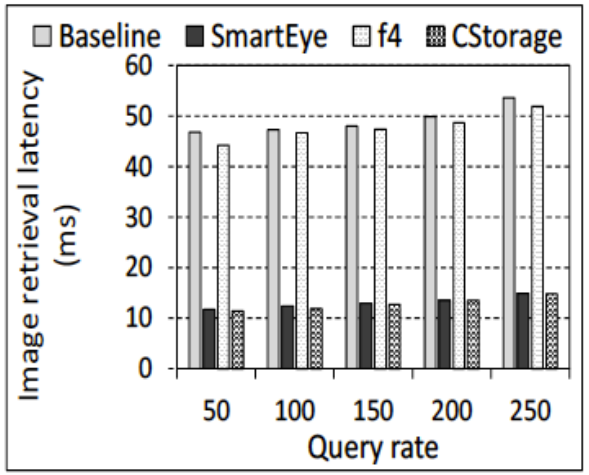


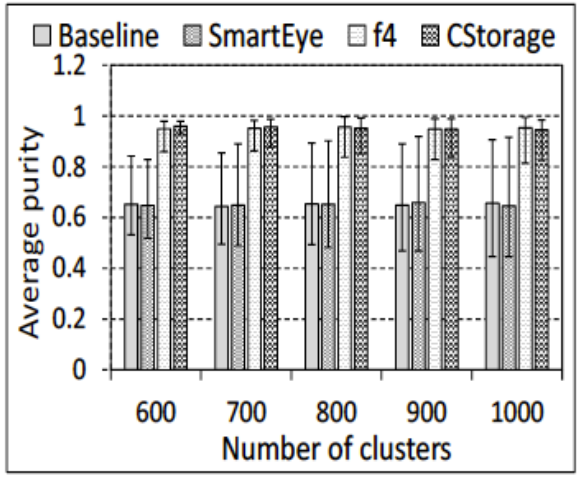
Overview of the image storage and retrieval module

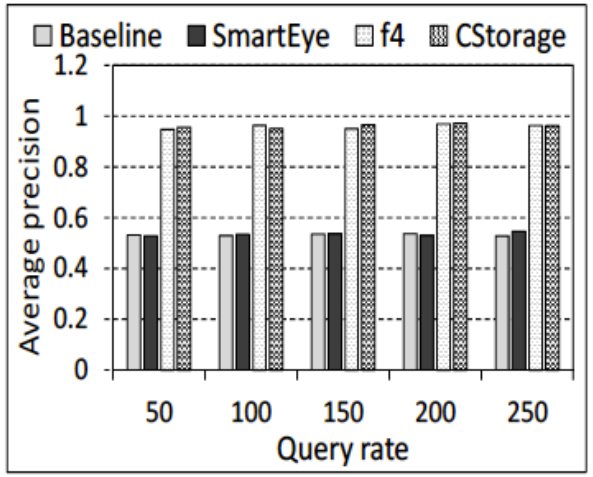
Outline

- Introduction
- Approach description
- Evaluation
- Conclusion


Baseline
SmartEye
f4
CStorage
Image upload latency (ms)
0
2
4
6
8
50
100
150
200
250
Upload rate


Baseline
SmartEye
f4
CStorage
Image retrieval latency (ms)
0
10
20
30
40
50
60
50
100
150
200
250
Query rate


Baseline
SmartEye
f4
CStorage
Average purity
0
0.2
0.4
0.6
0.8
1
1.2
600
700
800
900
1000
Number of clusters


Baseline
SmartEye
f4
CStorage
Average precision
0
0.2
0.4
0.6
0.8
1
1.2
50
100
150
200
250
Query rate

Outline

- Introduction
- Approach description
- Evaluation
- Conclusion

1) CStorage leverages the deep learning technique to extract image features, which are more accurate than handcrafted image features such as PCA-SIFT.

2) CStorage significantly reduces the image retrieval latency primarily caused by incast network congestion.

3) We conduct extensive experiment with millions of images. Experimental results show the effectiveness of CStorage in reducing image retrieval latency and improving retrieval accuracy.

Thank you!

Questions?