WAN Optimized Replication of Backup Datasets Using Stream-Informed Delta Compression

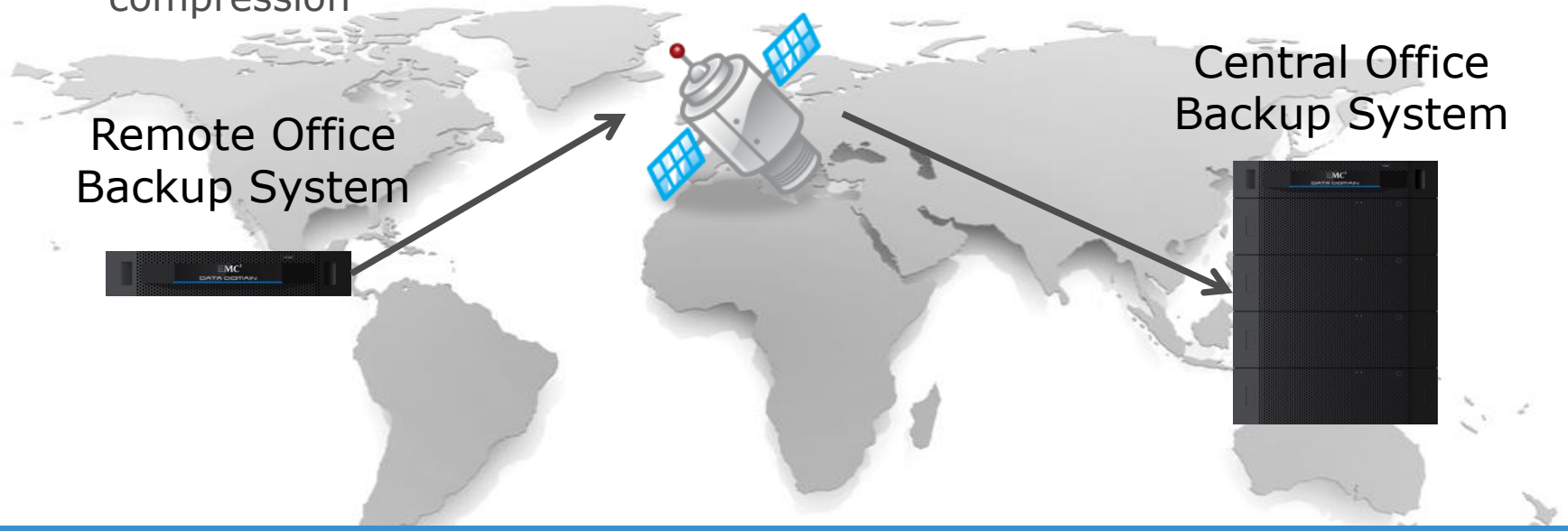
Philip Shilane, Mark Huang, Grant Wallace, & Windsor Hsu

Backup Recovery Systems Division
EMC Corporation

EMC²

Introduction

- 80% of our customers replicate most of their data off-site for disaster recovery
- WAN bandwidth often limits throughput
- Data reduction techniques increase effective throughput
  - Deduplication and local compression are effective
  - Delta compression with stream-informed caching adds 2X additional compression

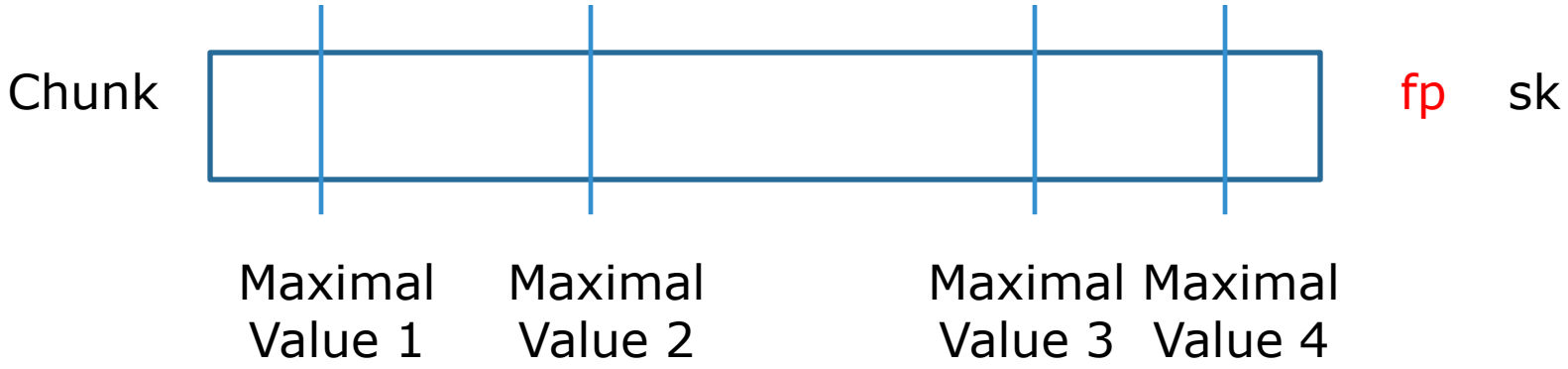


Example Of Deduplication And Delta Compression

Sketches based on Broder [97 & 00]

EMC²

Example Of Deduplication And Delta Compression



super_feature = Rabin_fp(feature$_1$…feature$_4$)

sketch is one or more super_features

Sketches based on Broder [97 & 00]

Example Of Deduplication And Delta Compression

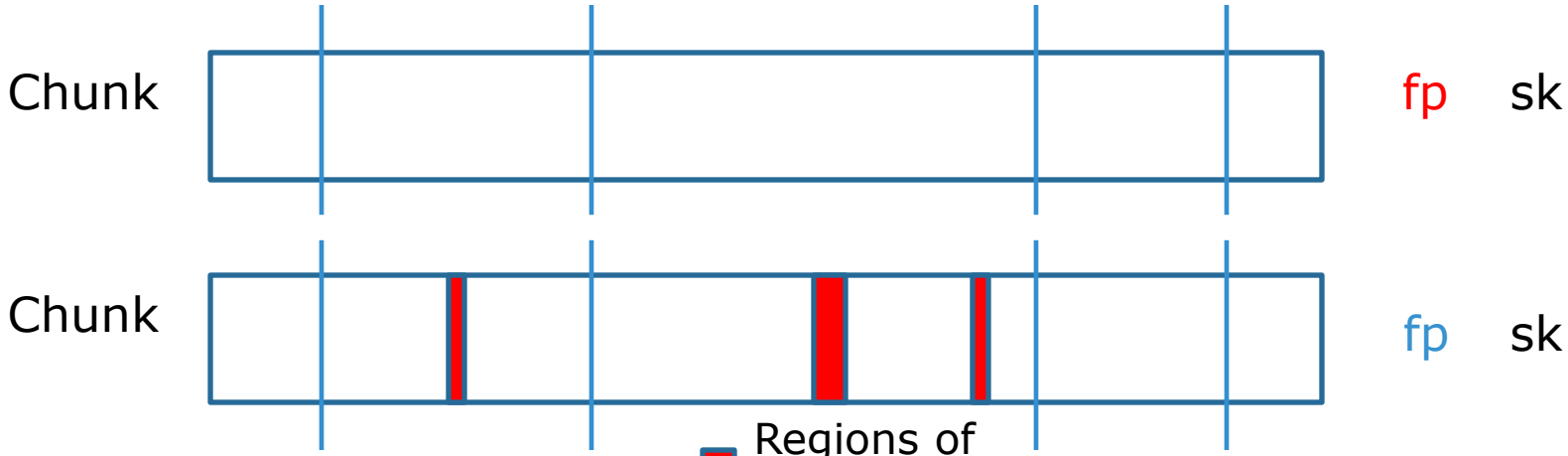
(duplicate of earlier chunk)

super_feature = Rabin_fp(feature$_1$…feature$_4$)

sketch is one or more super_features

Sketches based on Broder [97 & 00]

Example Of Deduplication And Delta Compression

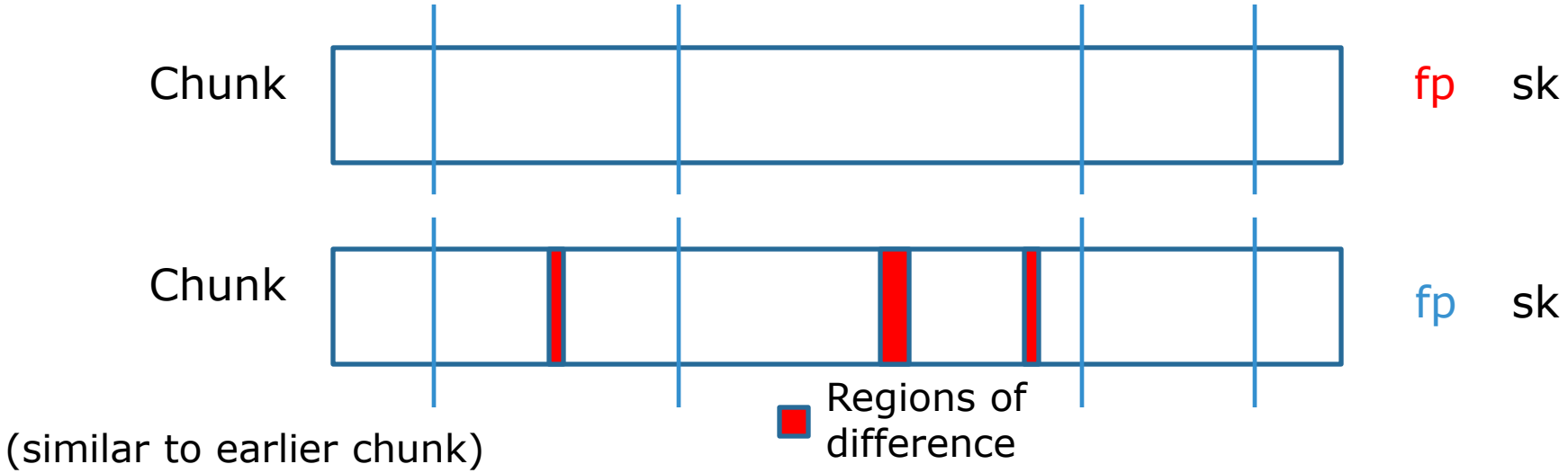


super_feature = Rabin_fp(feature$_1$…feature$_4$)

sketch is one or more super_features

Sketches based on Broder [97 & 00]

Example Of Deduplication And Delta Compression



Transmit fp and differences

Sketches based on Broder [97 & 00]

Sketch Index Options

- Full index (simple idea)
  - Requires IO
  - Difficult to update
  - Finds all similarity matches

256 TB capacity
8 KB chunks
16 byte record

0.5 TB index
per super-feature

Sketch Index Options

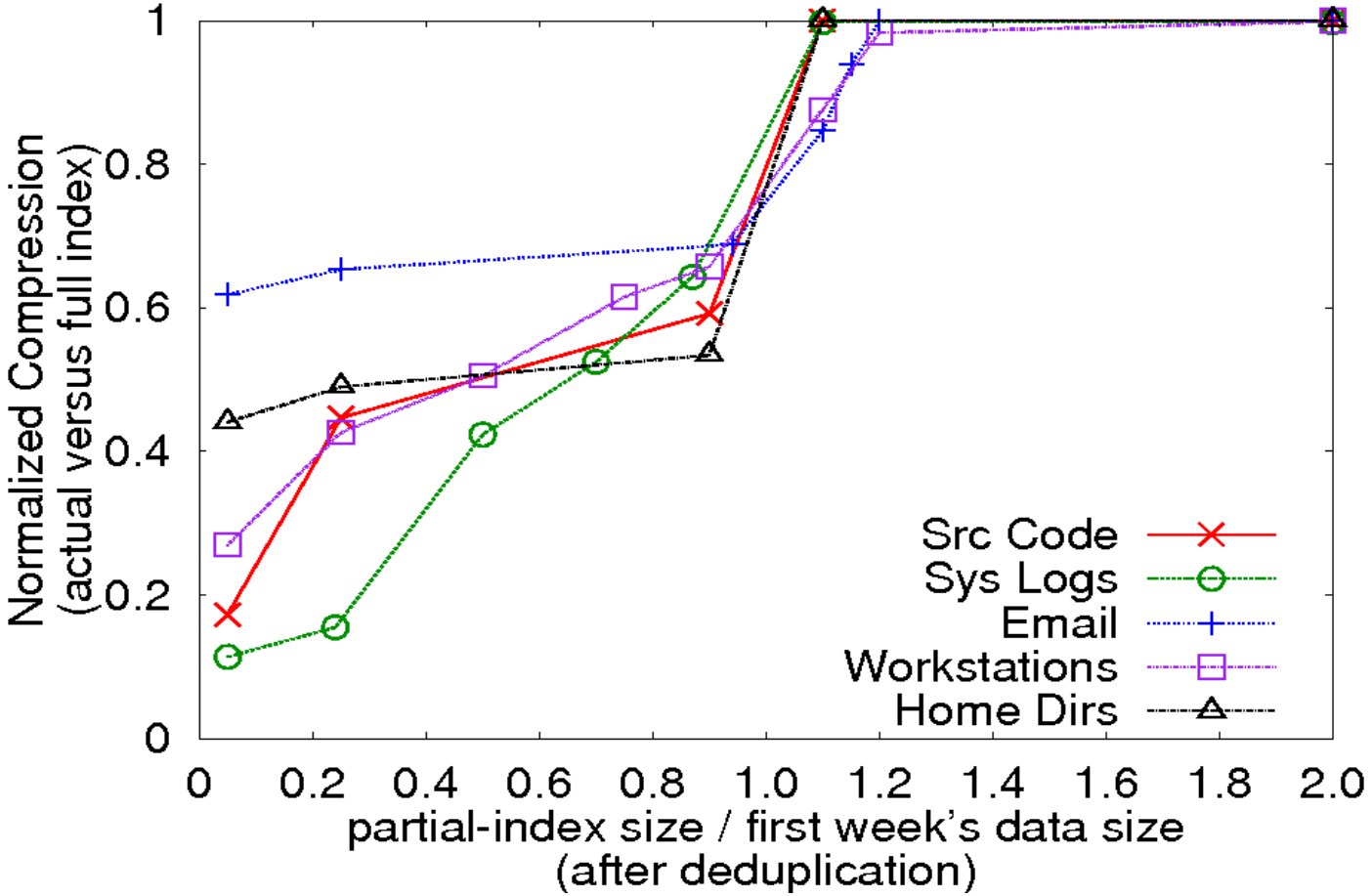
- Full index (simple idea)
  - Requires IO
  - Difficult to update
  - Finds all similarity matches
- Partial index (slightly better?)
  - Load and evict with LRU policy
  - Not persistent
  - Must be as large as full backup

256 TB capacity
8 KB chunks
16 byte record

0.5 TB index
per super-feature

Sketch Index Options

Partial index would have to hold a full backup to be effective



Sketch Index Options

- Full index (simple idea)
  - Requires IO
  - Difficult to update
  - Finds all similarity matches
- Partial index (slightly better?)
  - Load and evict with LRU policy
  - Not persistent
  - Must be as large as full backup

256 TB capacity
8 KB chunks
16 byte record

0.5 TB index
per super-feature

Partial index has to be large enough to index entire primary storage system

EMC²

Sketch Index Options

- Full index (simple idea)
  - Requires IO
  - Difficult to update
  - Finds all similarity matches

256 TB capacity
8 KB chunks
16 byte record

0.5 TB index
per super-feature

- Partial index (slightly better?)
  - Load and evict with LRU policy
  - Not persistent
  - Must be as large as full backup

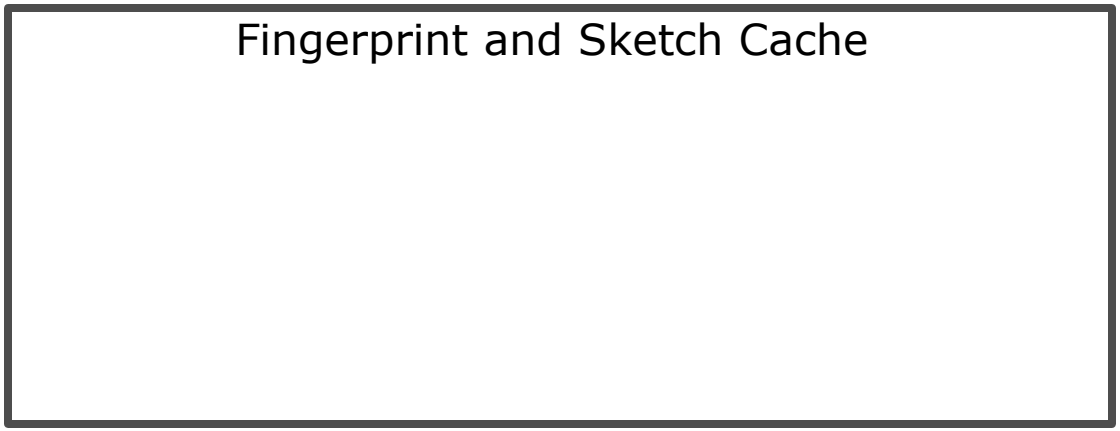
Partial index has to be large enough to index entire primary storage system

- Stream-informed cache (**our contribution**)
  - Experimentally demonstrate that delta locality closely matches deduplication locality for backup datasets
  - Updates handled by fingerprint system
  - Little extra memory
  - Finds most similarity matches

EMC²

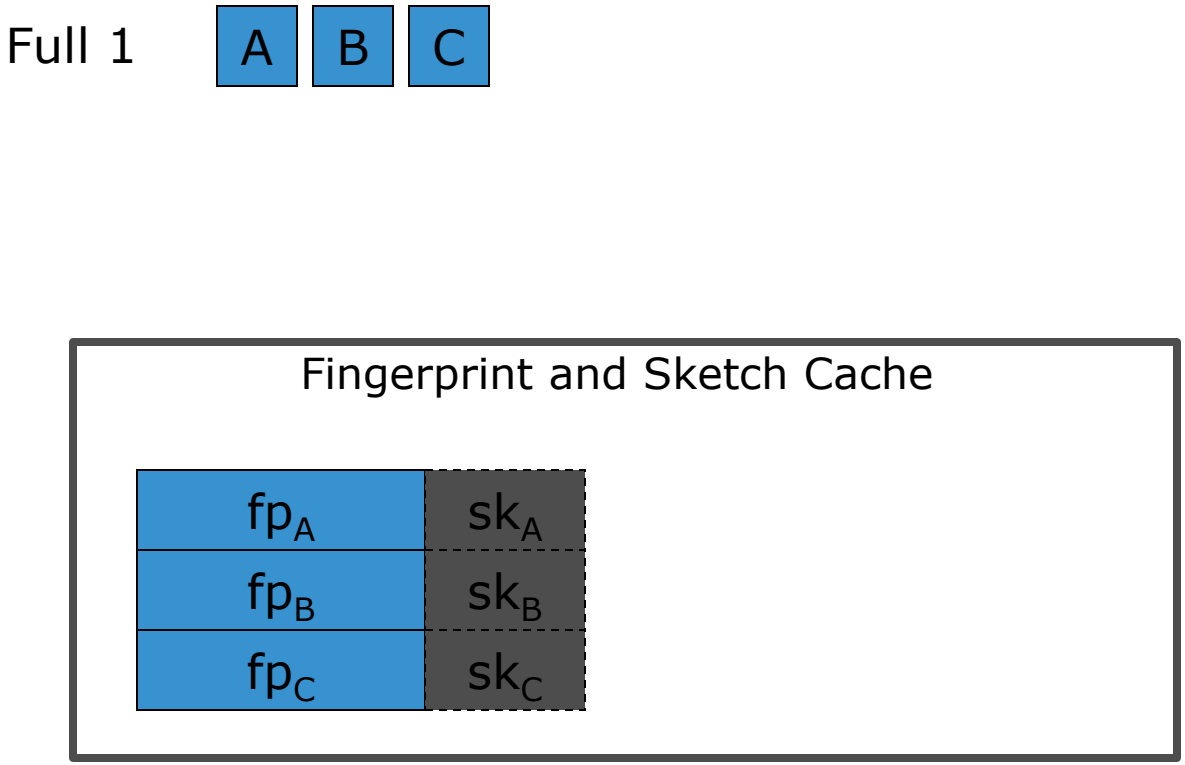
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



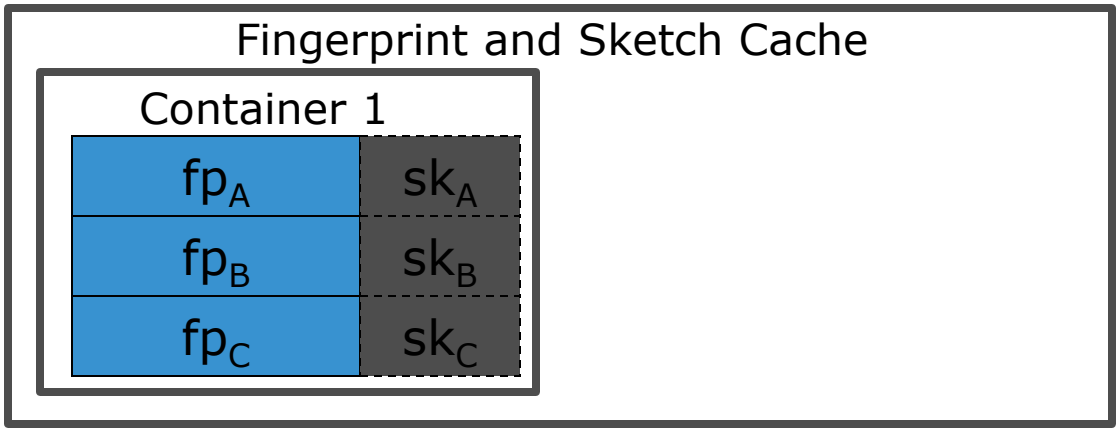
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



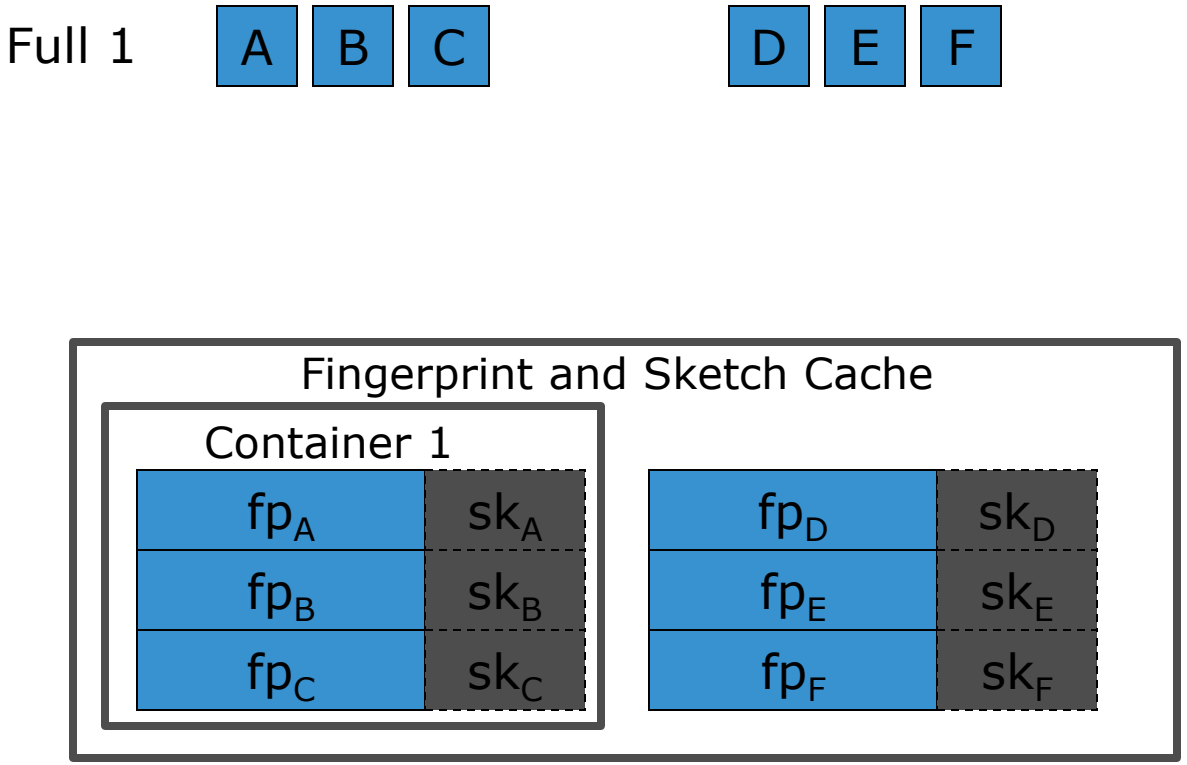
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



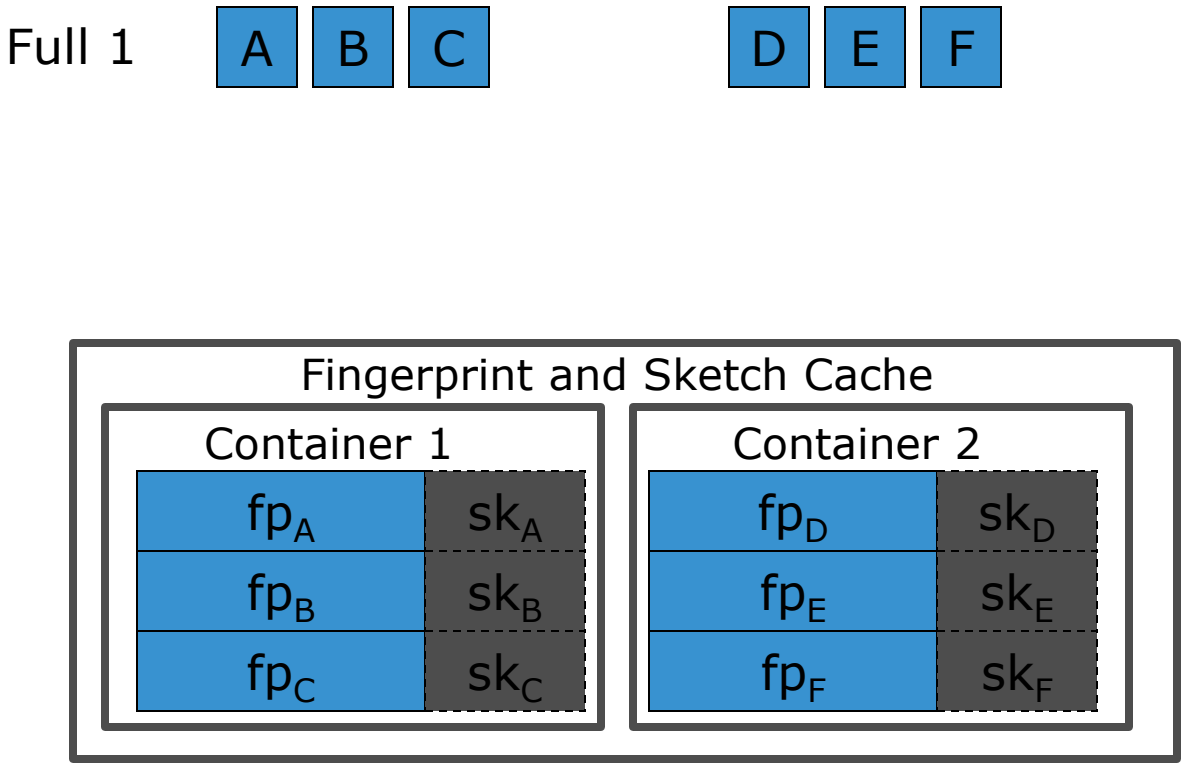
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



Stream-Informed Locality

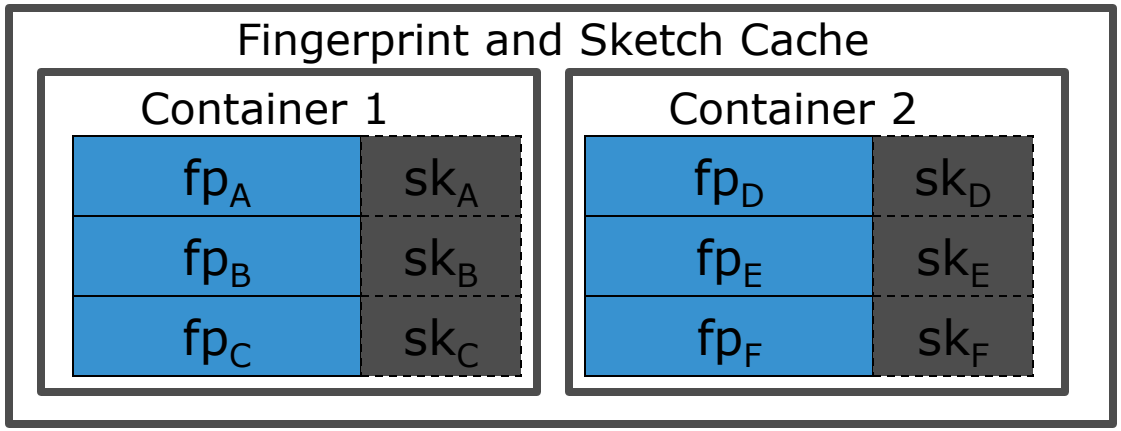
- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache

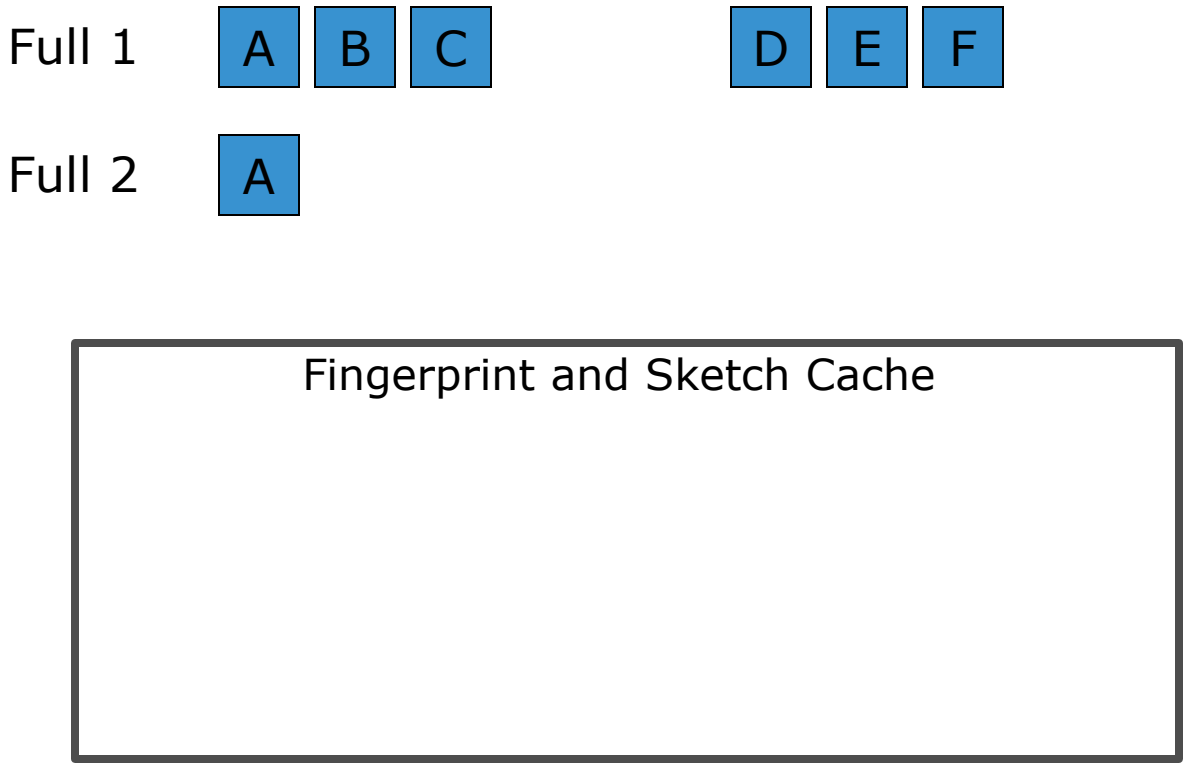
Store containers to disk



EMC²

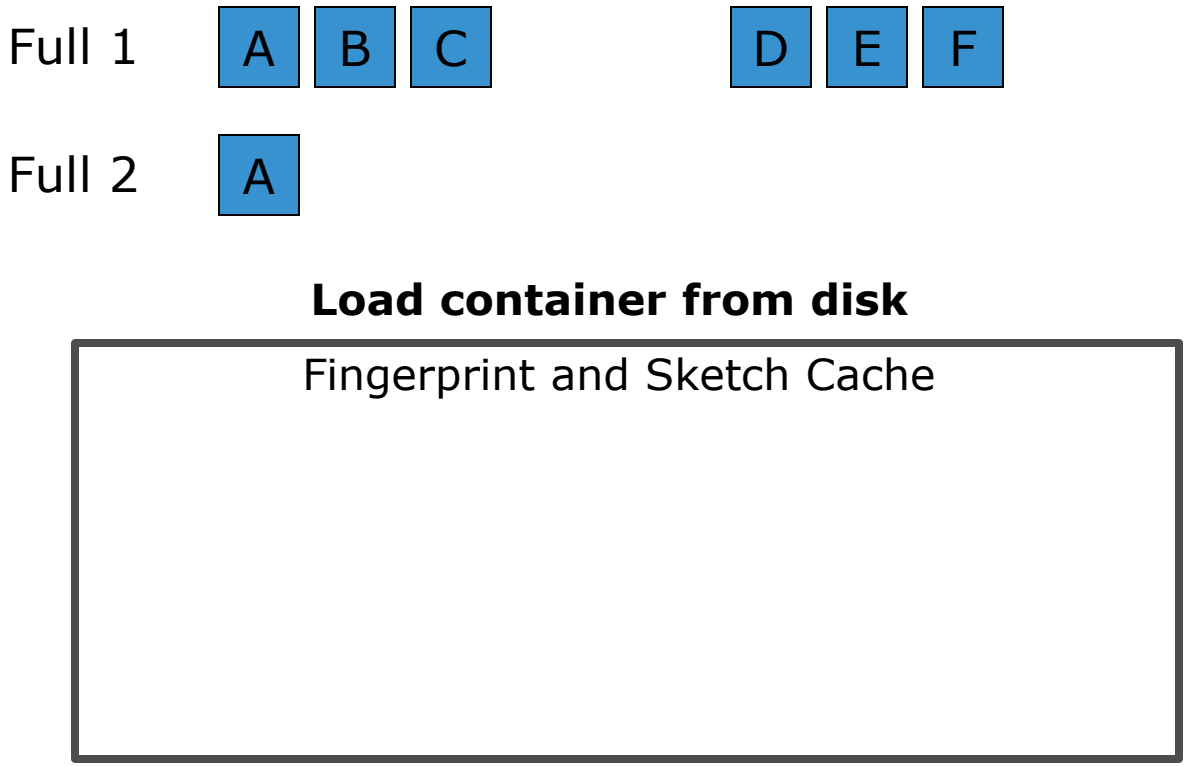
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



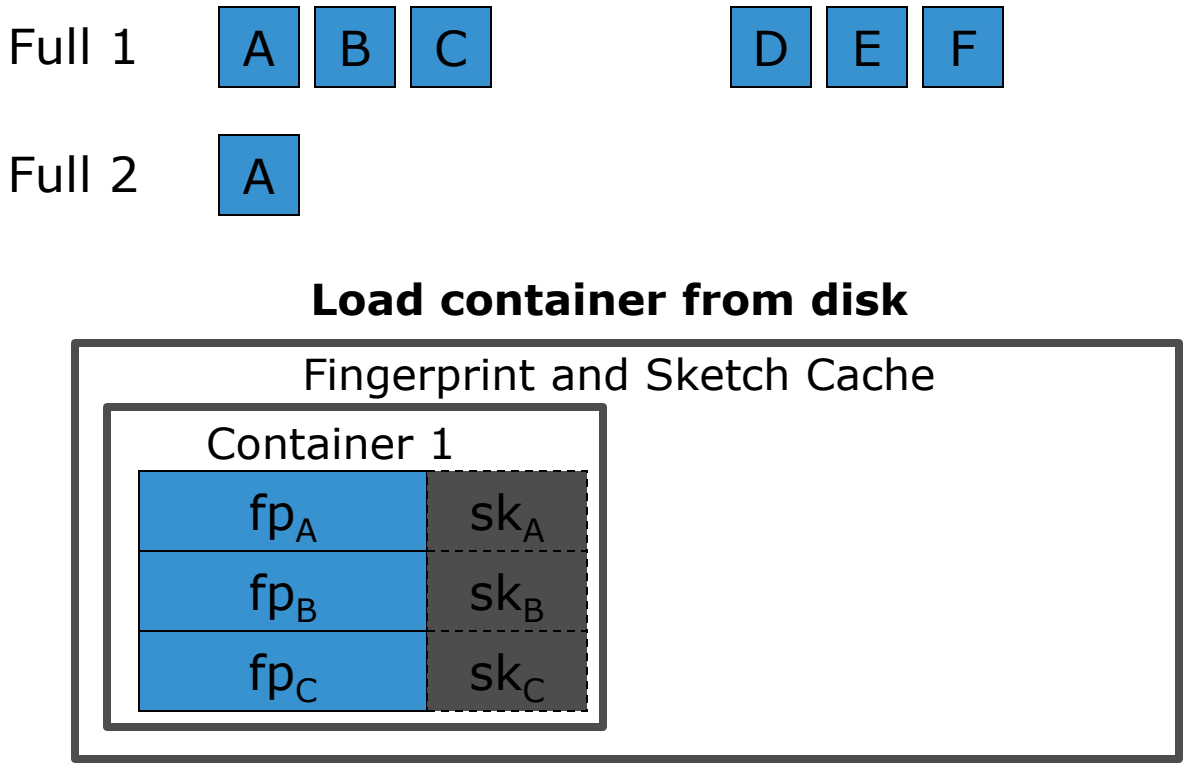
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



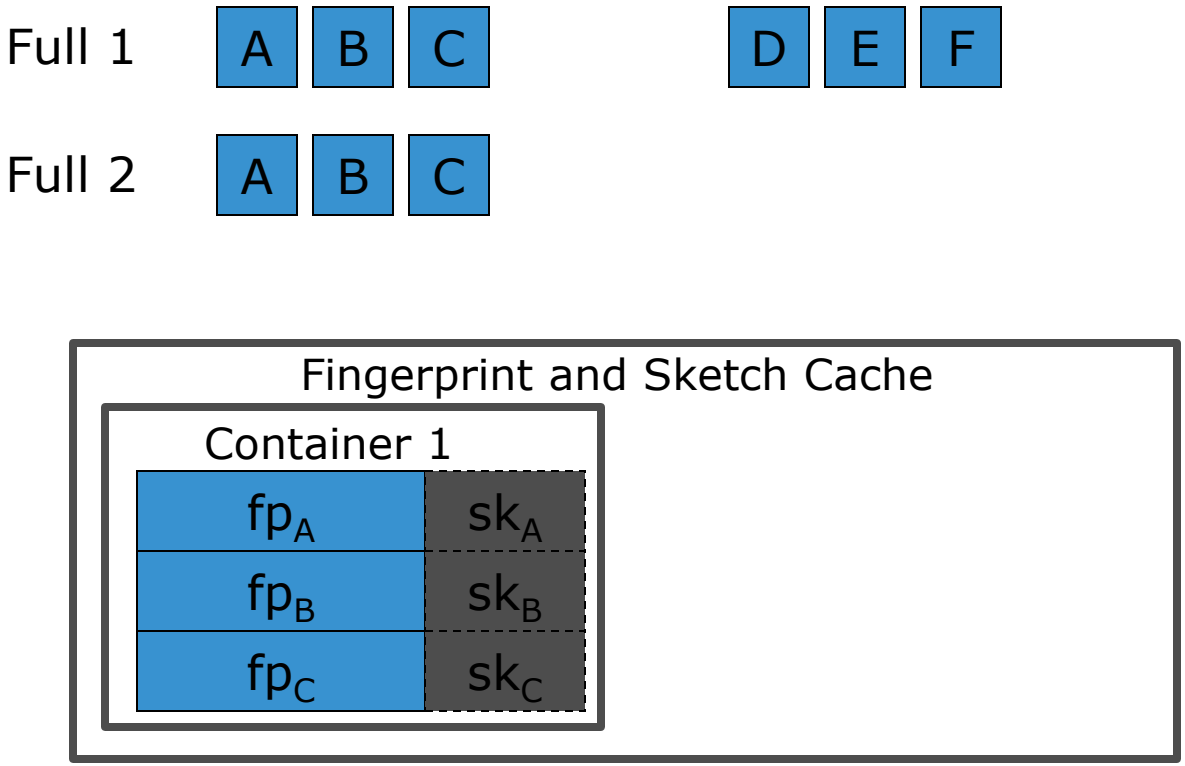
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



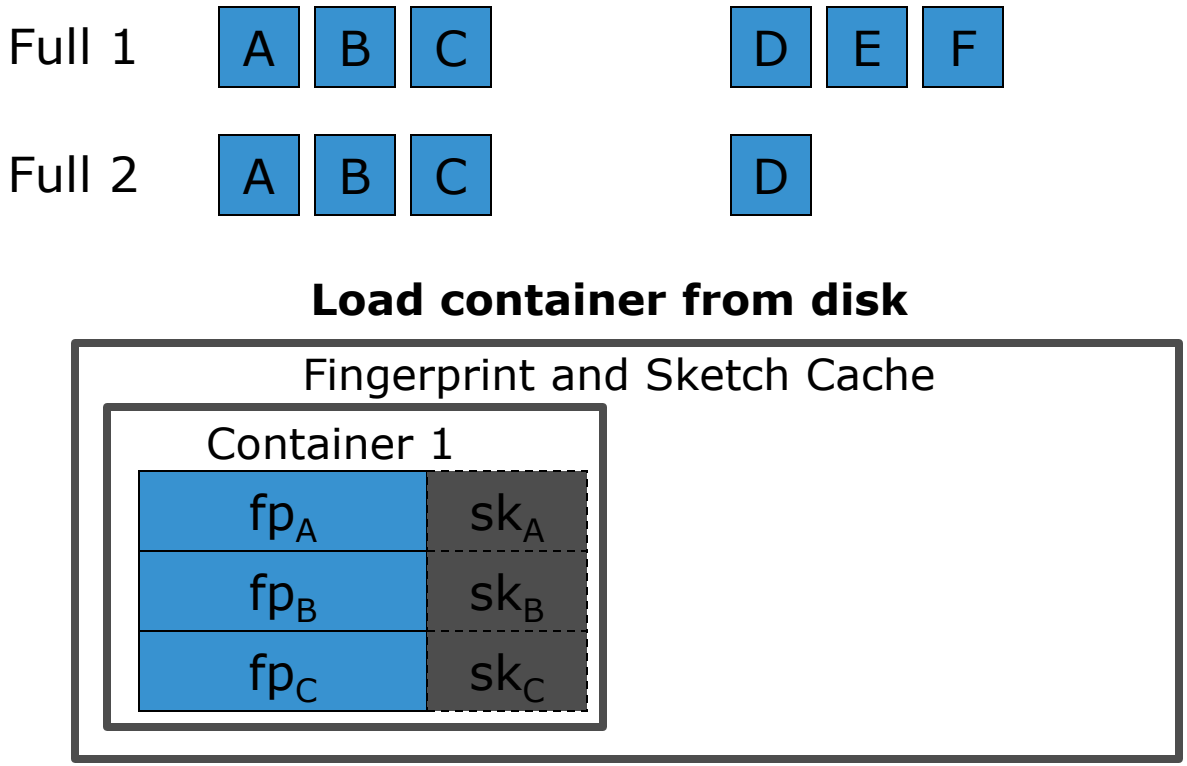
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



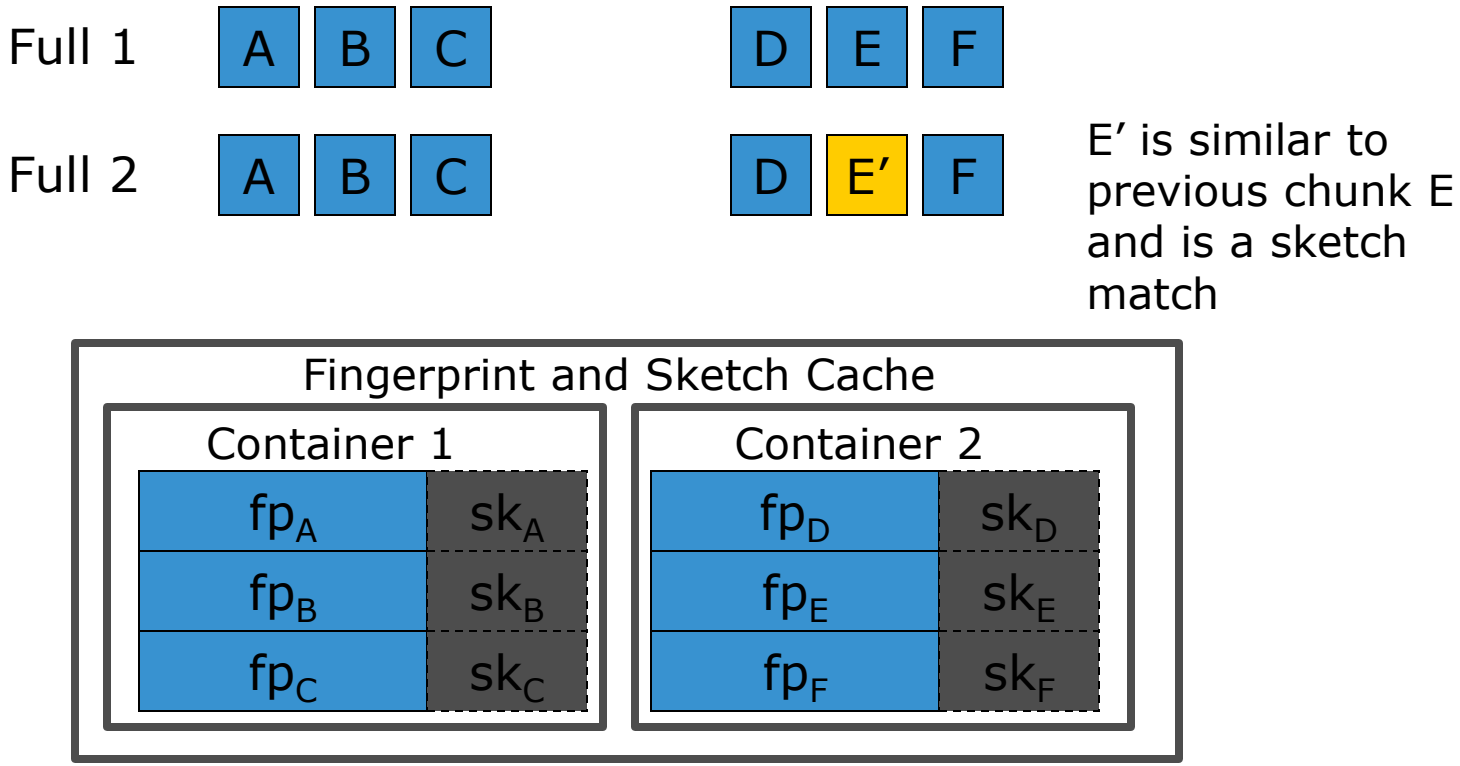
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache



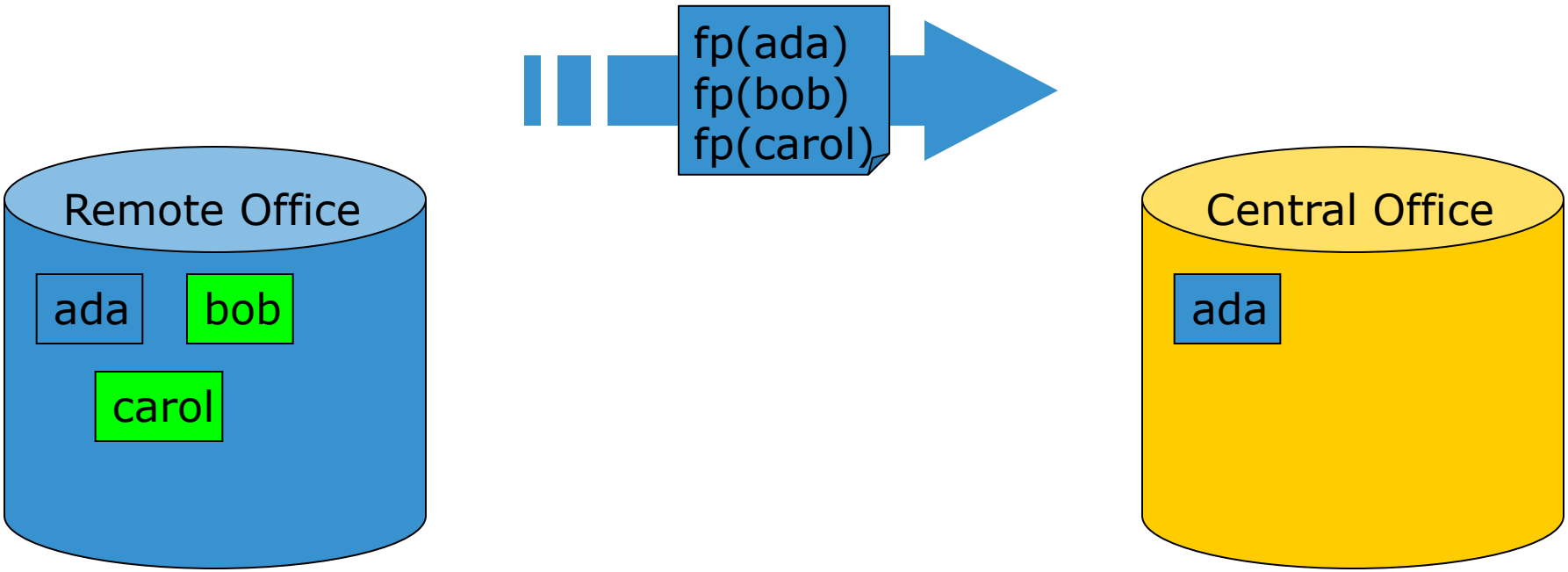
Stream-Informed Locality

- Similarity search can leverage deduplication locality
- Sketch cache loaded based on fingerprint cache

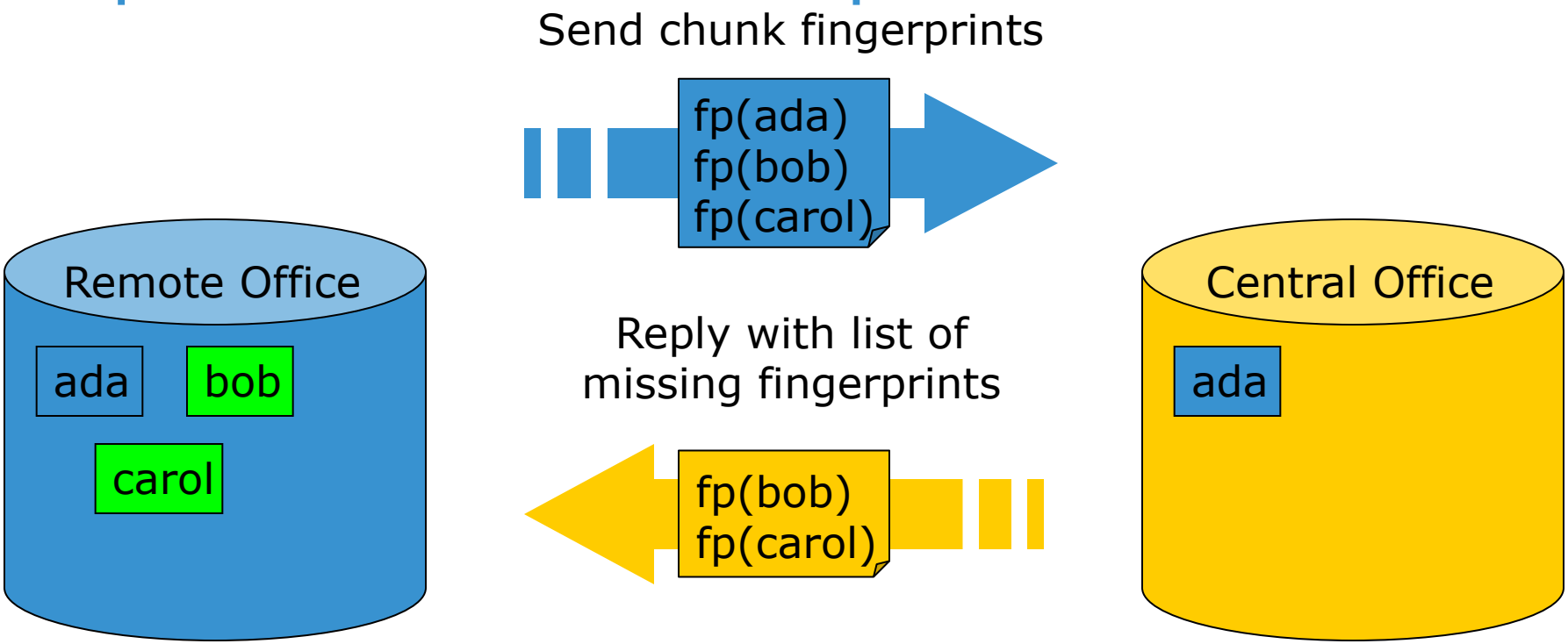


Replication With Deduplication

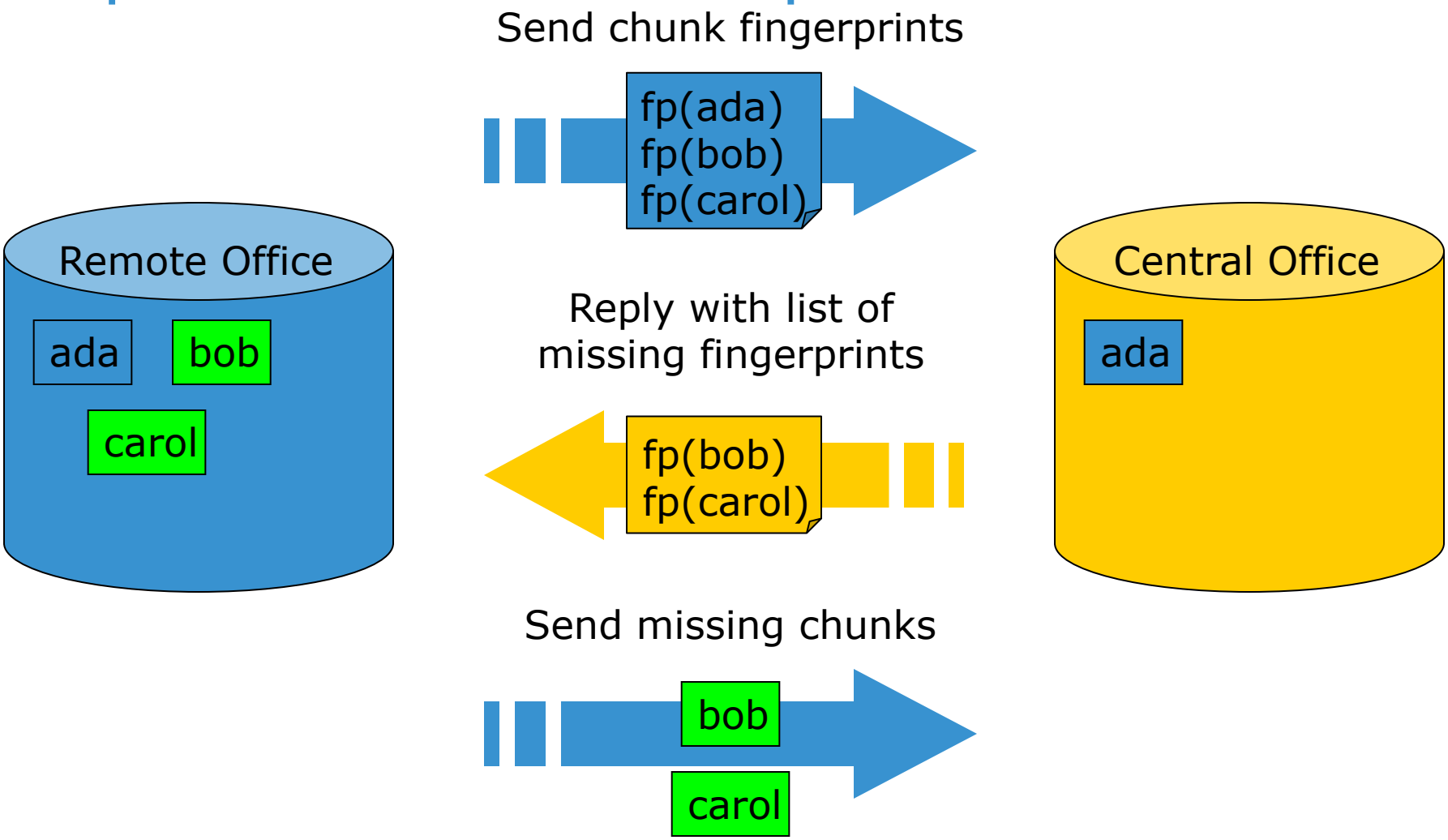
Replication With Deduplication



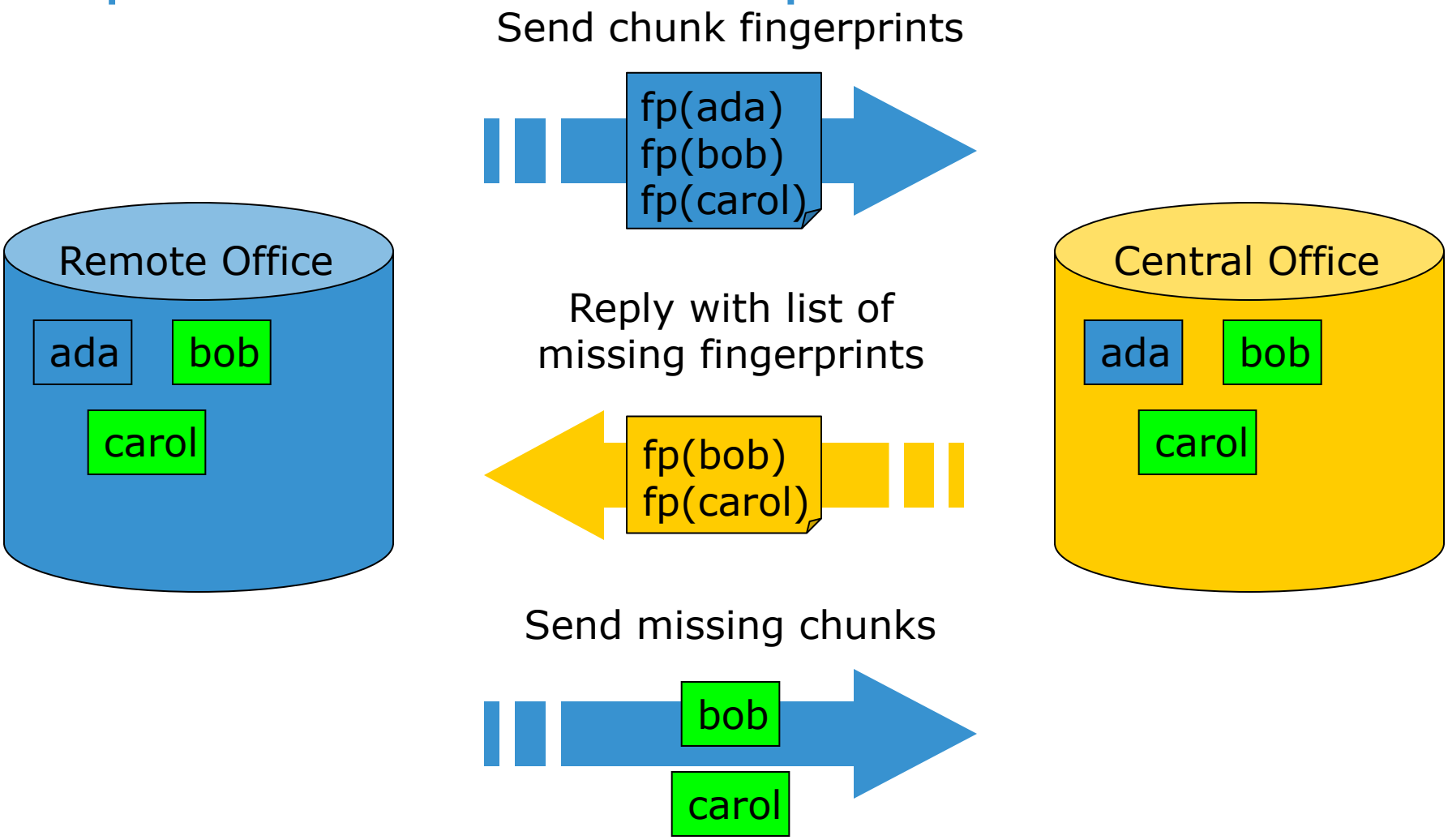
Replication With Deduplication



Replication With Deduplication

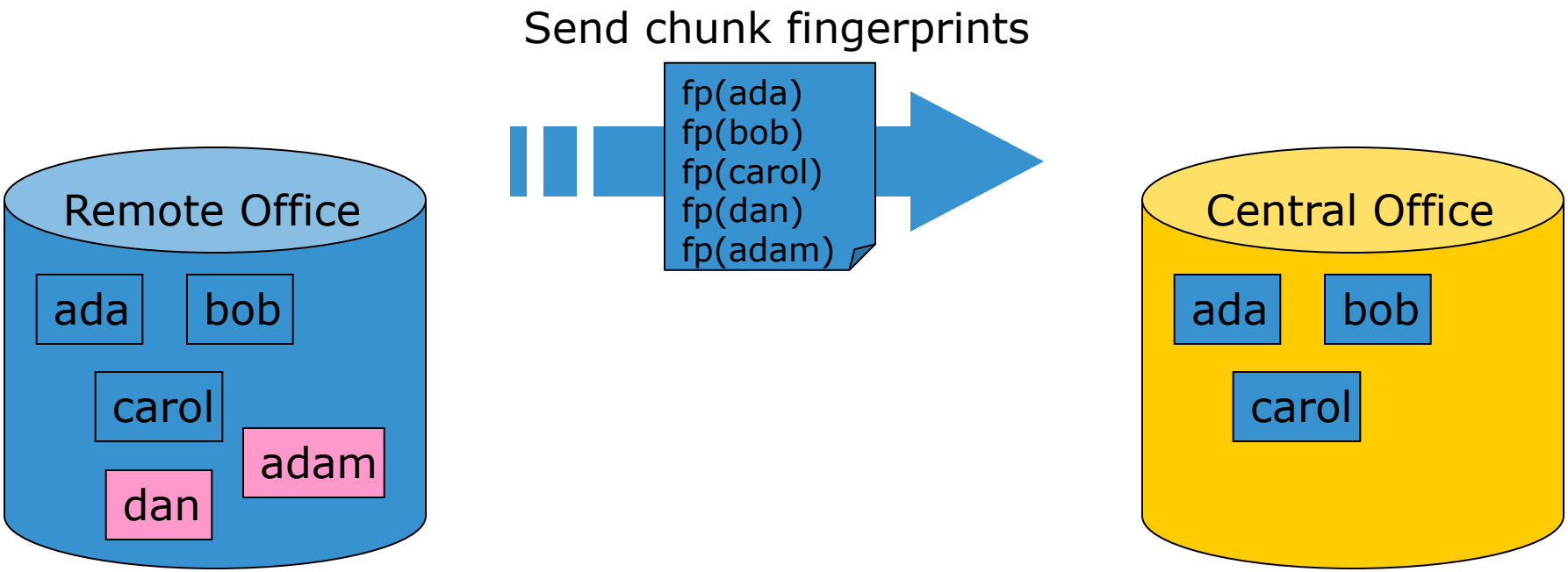


Replication With Deduplication

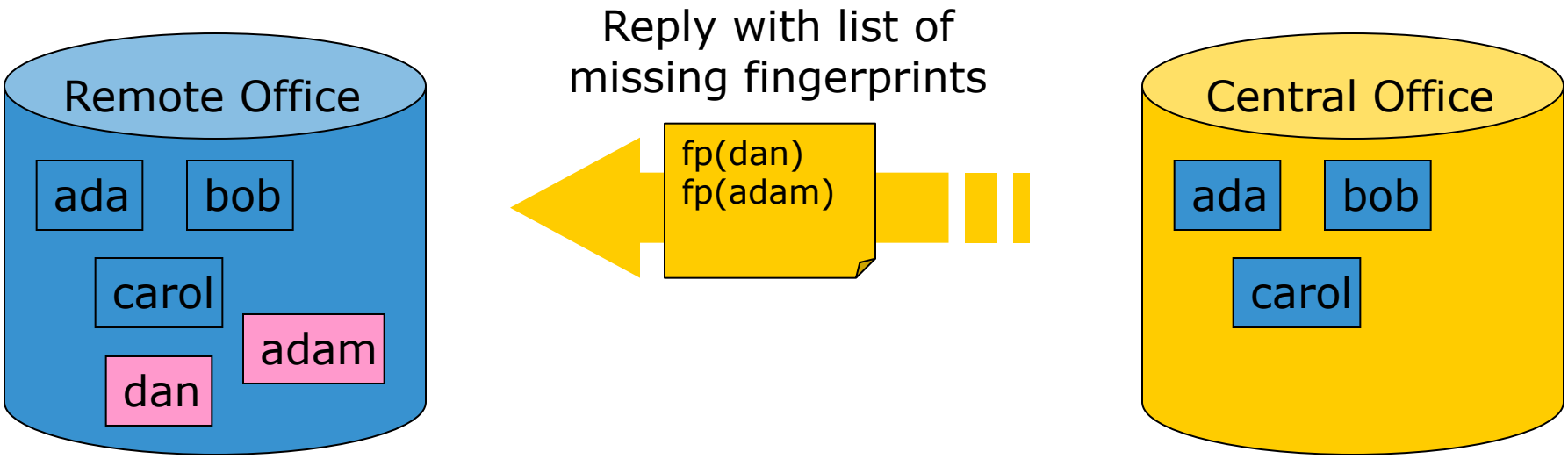


Replication With Deduplication And Delta

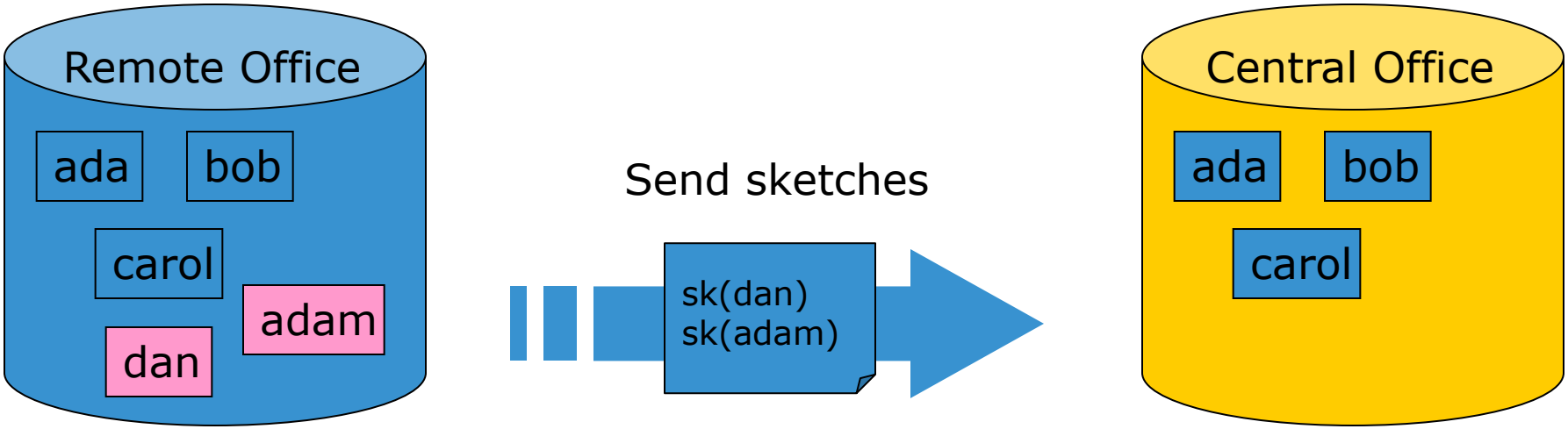
Replication With Deduplication And Delta



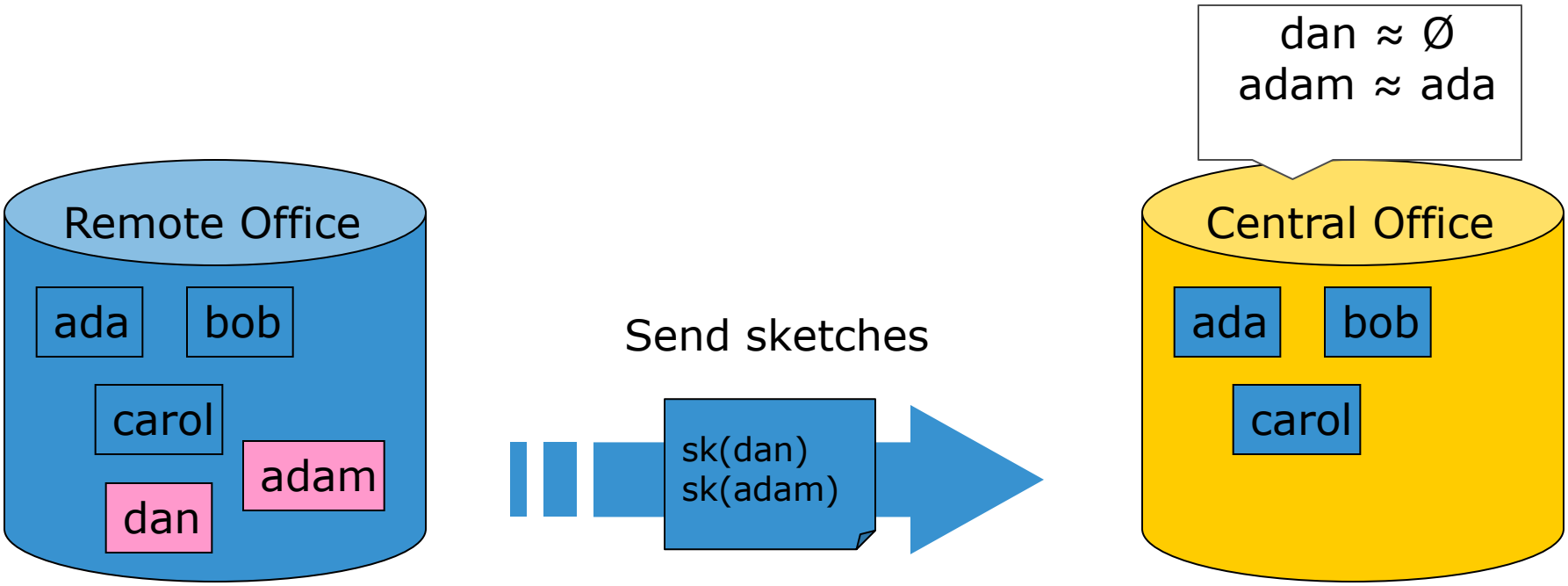
Replication With Deduplication And Delta



Replication With Deduplication And Delta

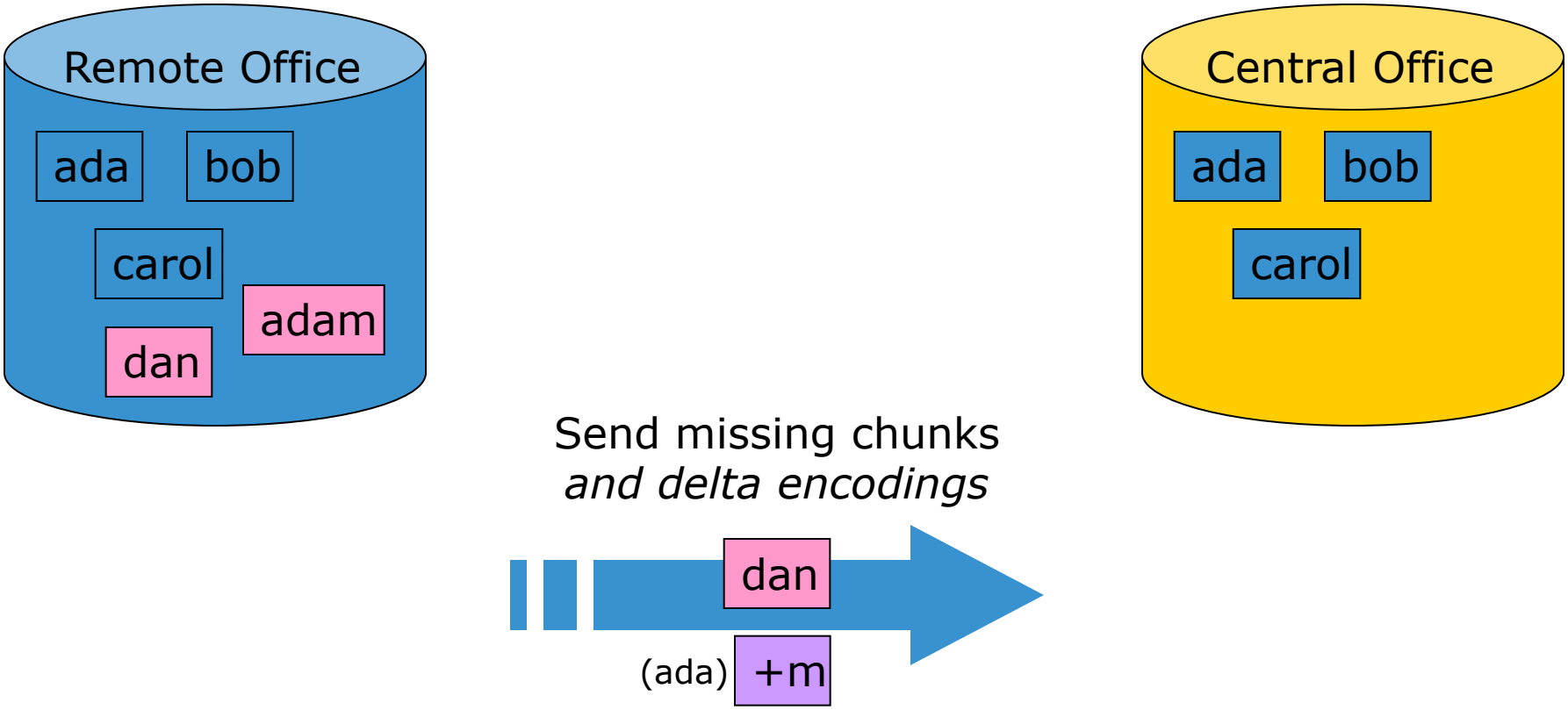


Replication With Deduplication And Delta

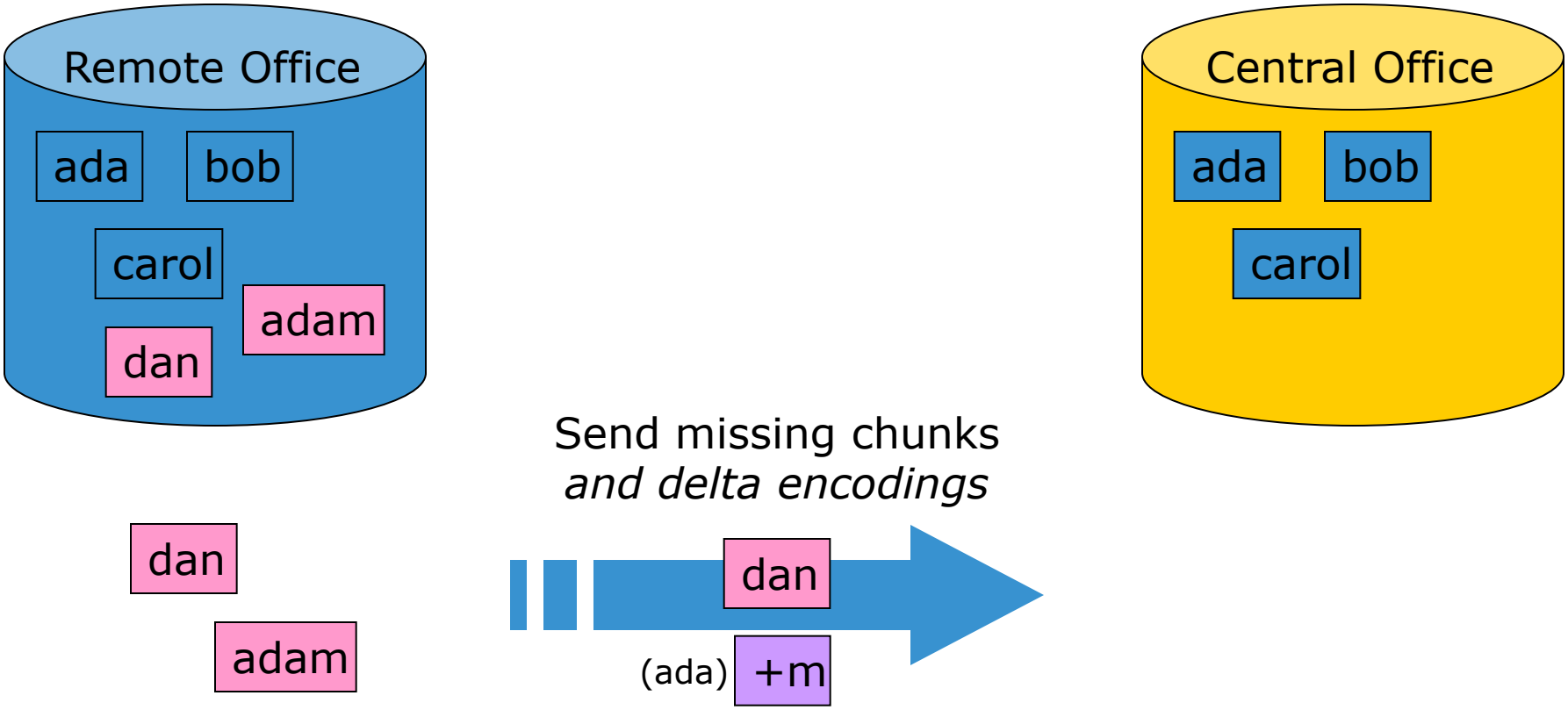


Replication With Deduplication And Delta

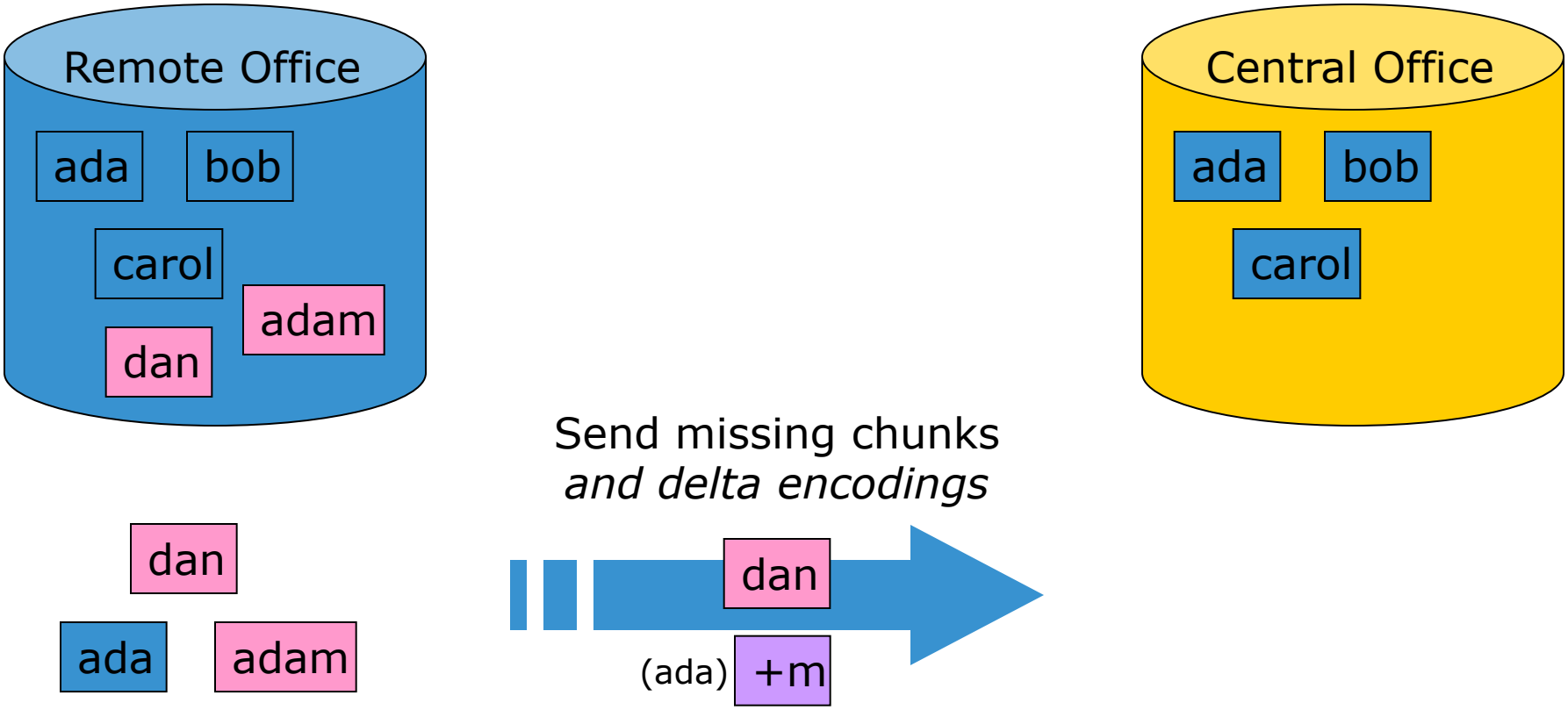
Replication With Deduplication And Delta



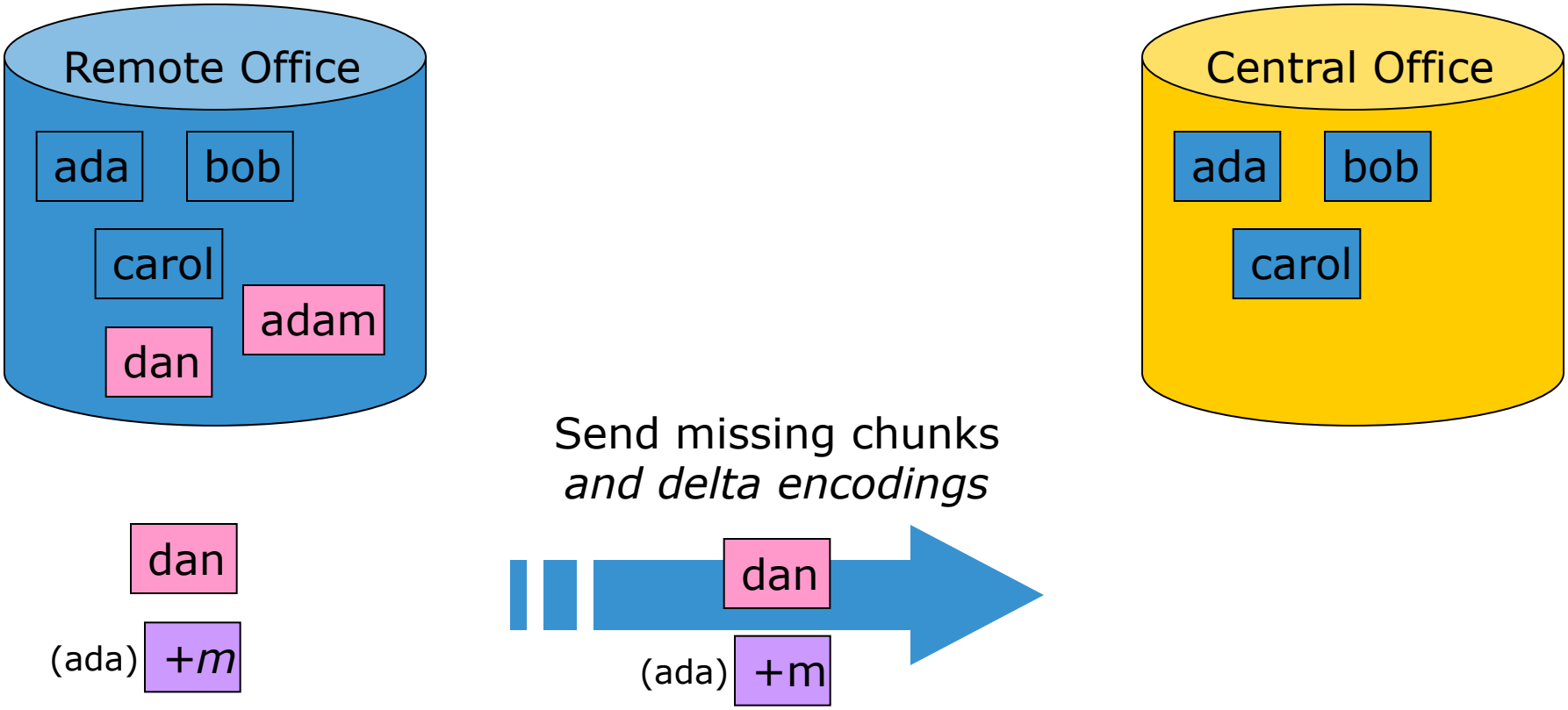
Replication With Deduplication And Delta



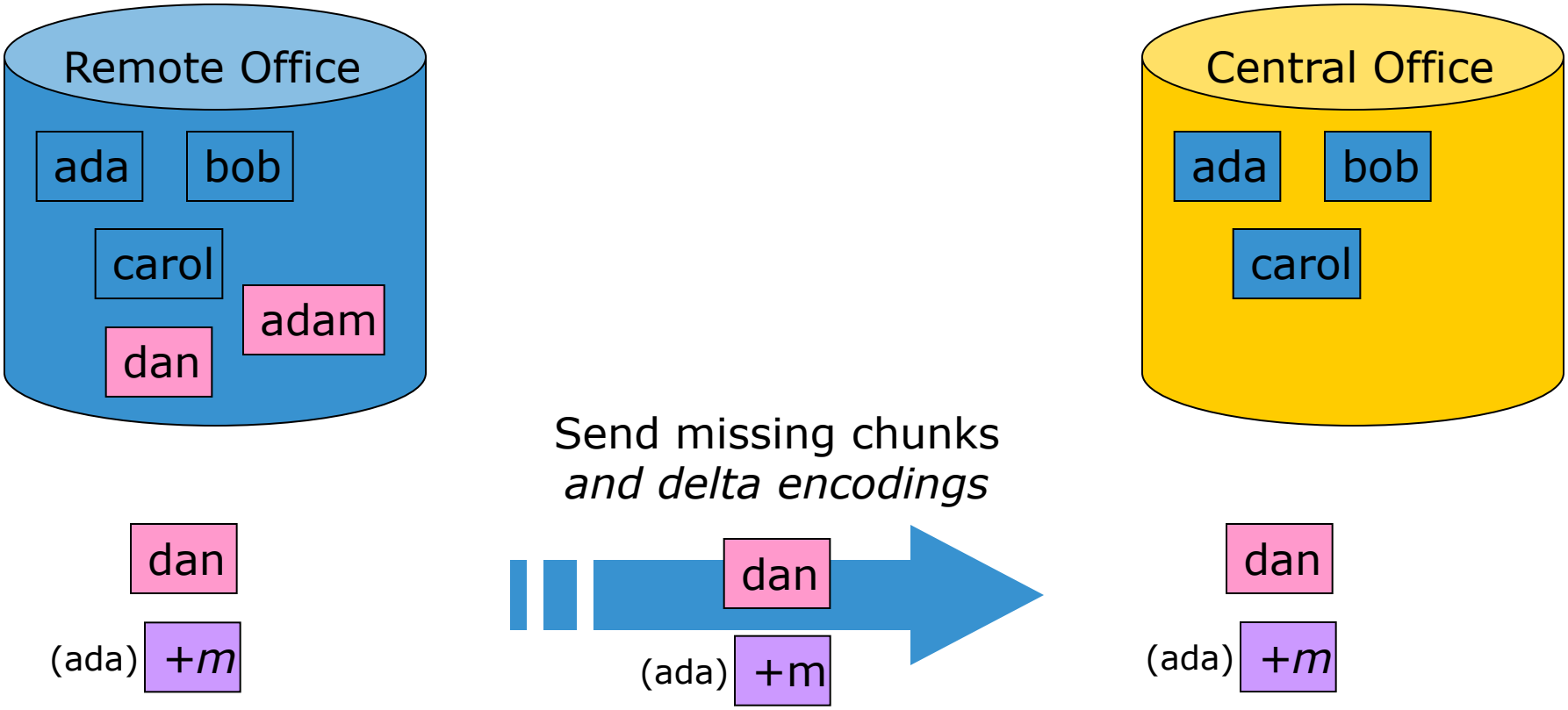
Replication With Deduplication And Delta



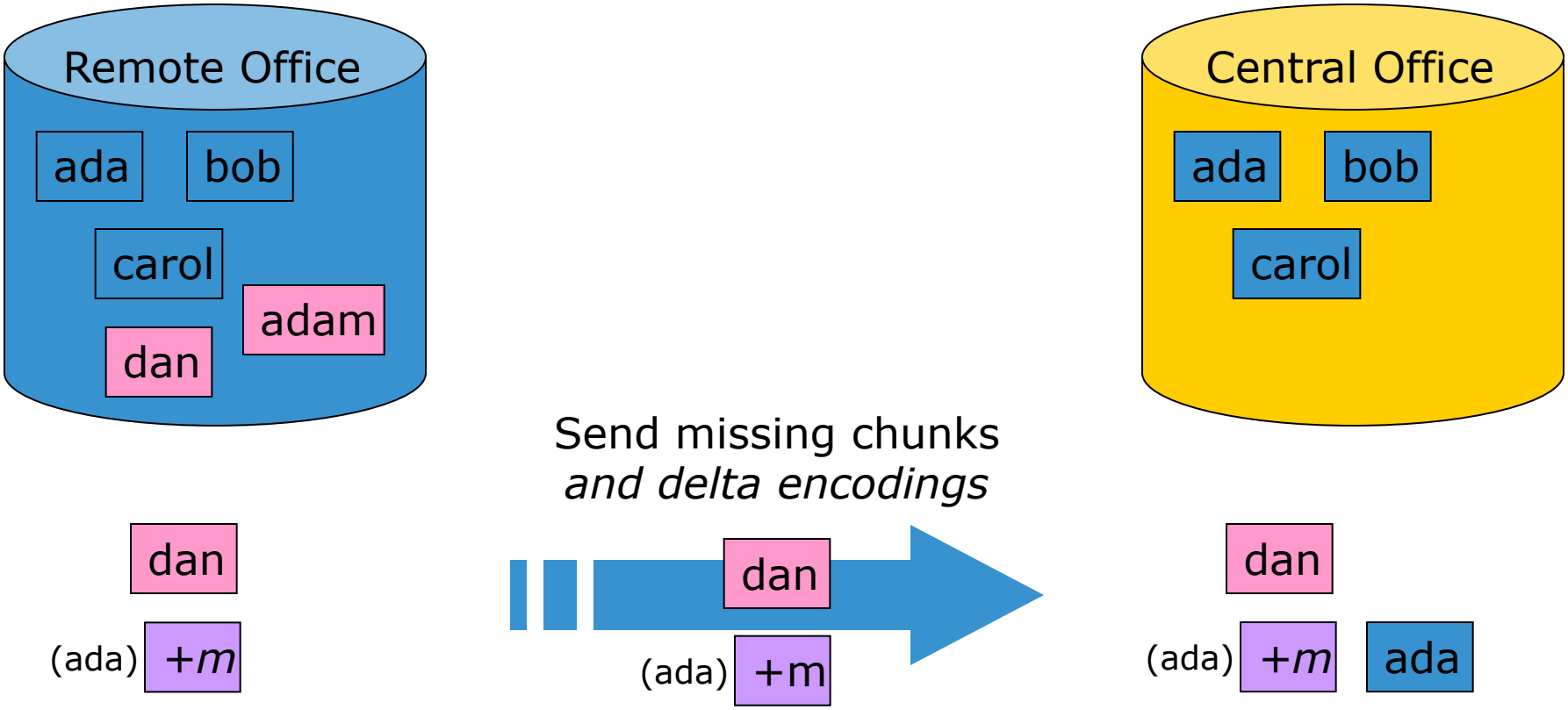
Replication With Deduplication And Delta



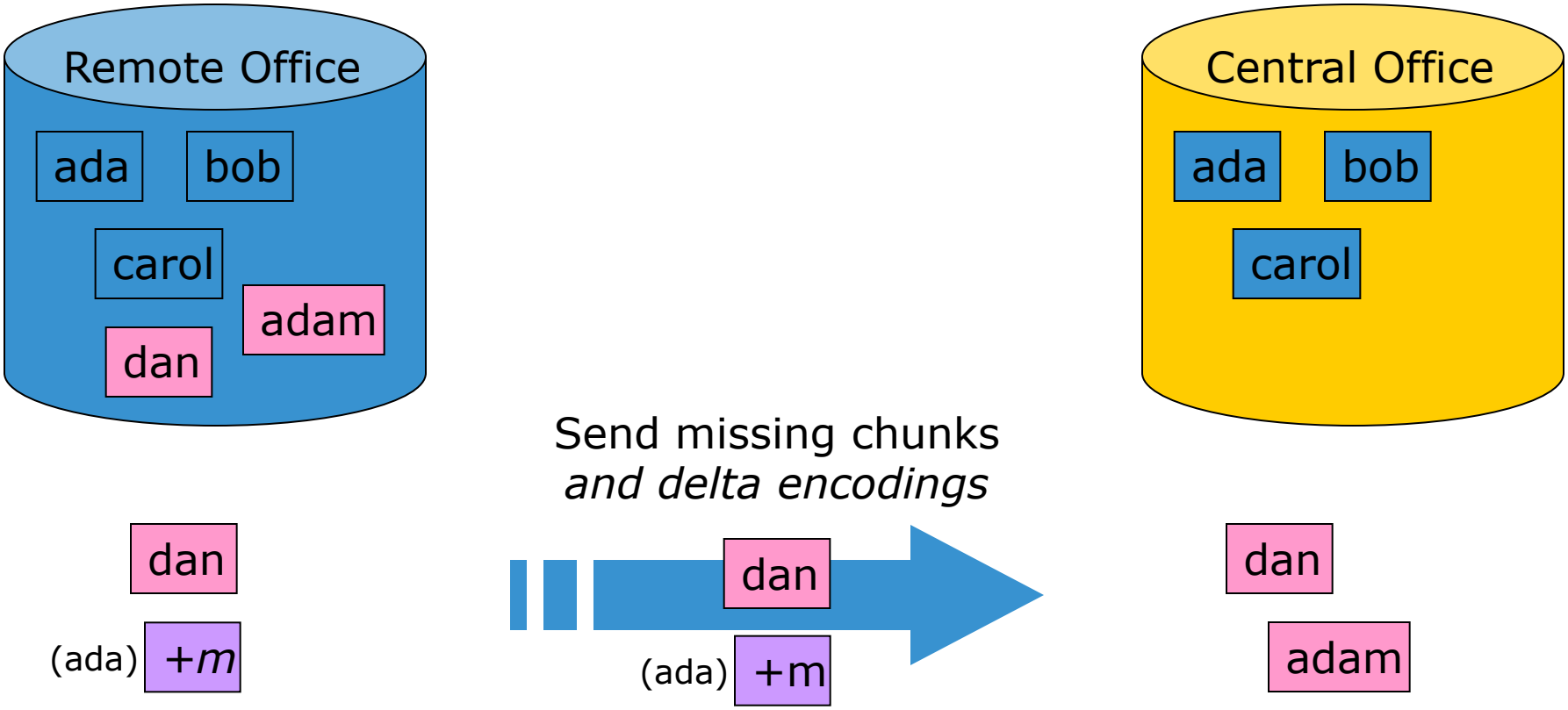
Replication With Deduplication And Delta



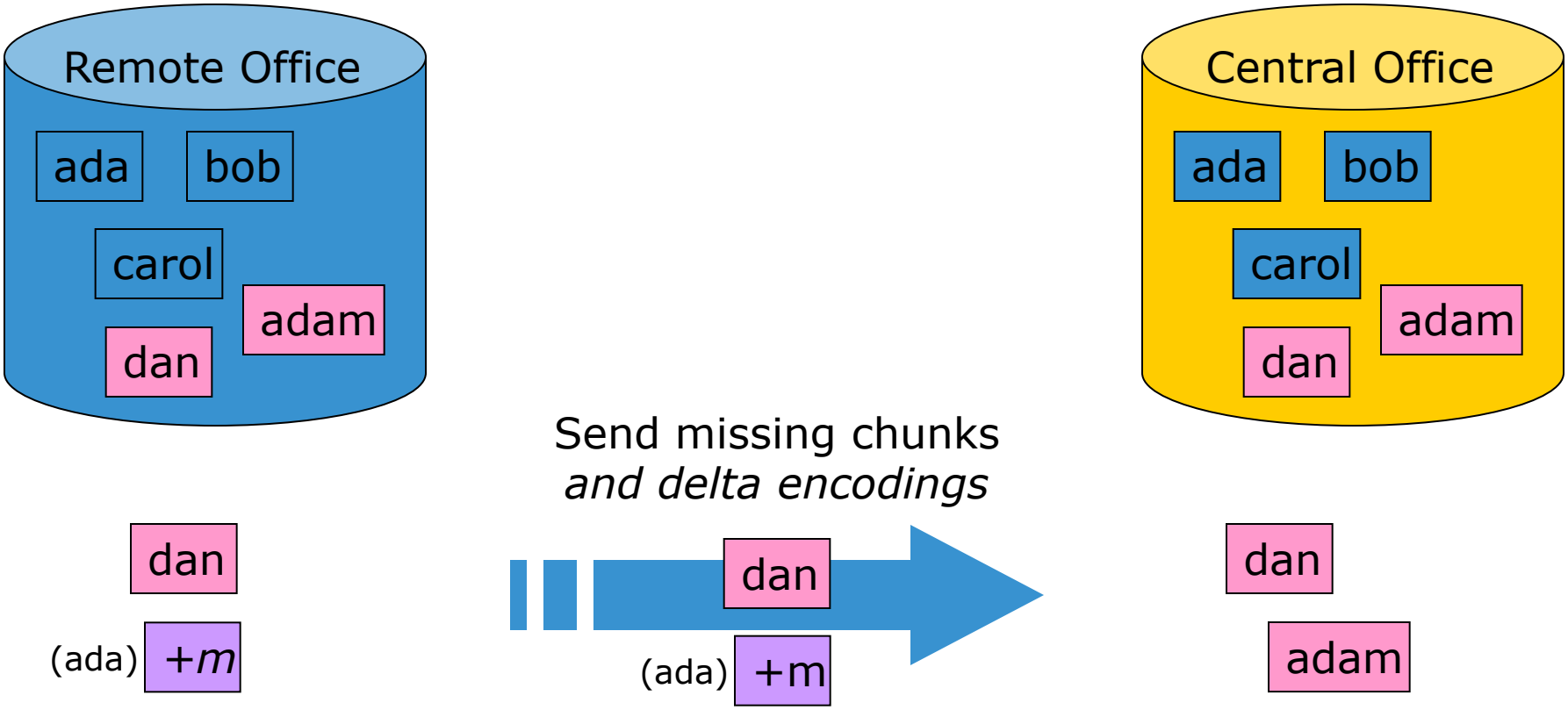
Replication With Deduplication And Delta



Replication With Deduplication And Delta



Replication With Deduplication And Delta



Properties Of Stream-Informed Delta Compression

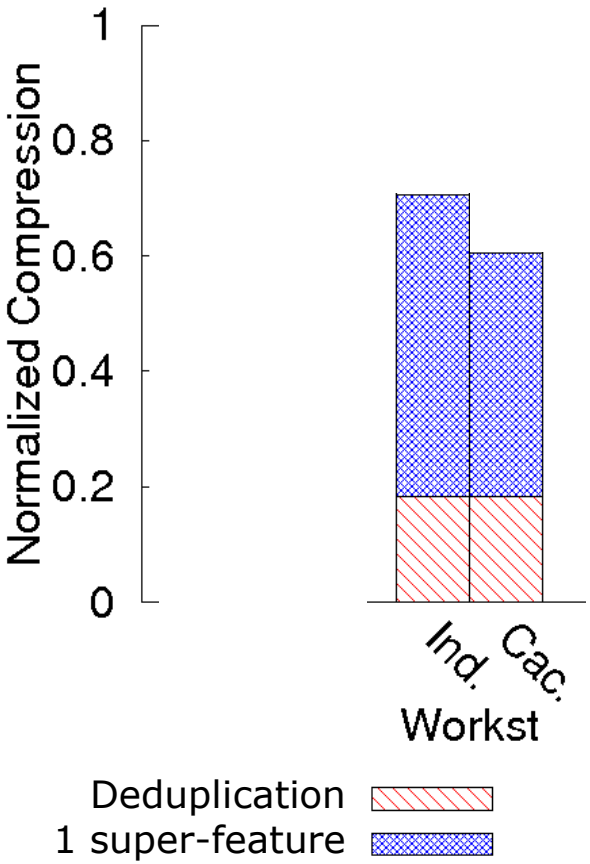
- Pros:
  - Eliminates on-disk sketch index
  - Improves performance – fewer disk reads than using a sketch index
  - Small memory footprint
  - Improves compression
- Cons:
  - Dependent on stream locality and caching to find similar chunks
  - Requires read IO and CPU to process delta chunks

Datasets

Dataset	Type	Backup Policy	TB	Months
Source Code	Version control repository	Weekly full Daily incremental	4.6	6
Workstations	16 desktops	Weekly full Daily incremental	4.9	6
Email	MS Exchange server	Daily full	5.2	7
System Logs	Server's /var directory	Weekly full Daily incremental	5.4	4
Home Directories	Engineers' home directories	Weekly full Daily incremental	12.9	3

Index VS Stream-Informed Cache

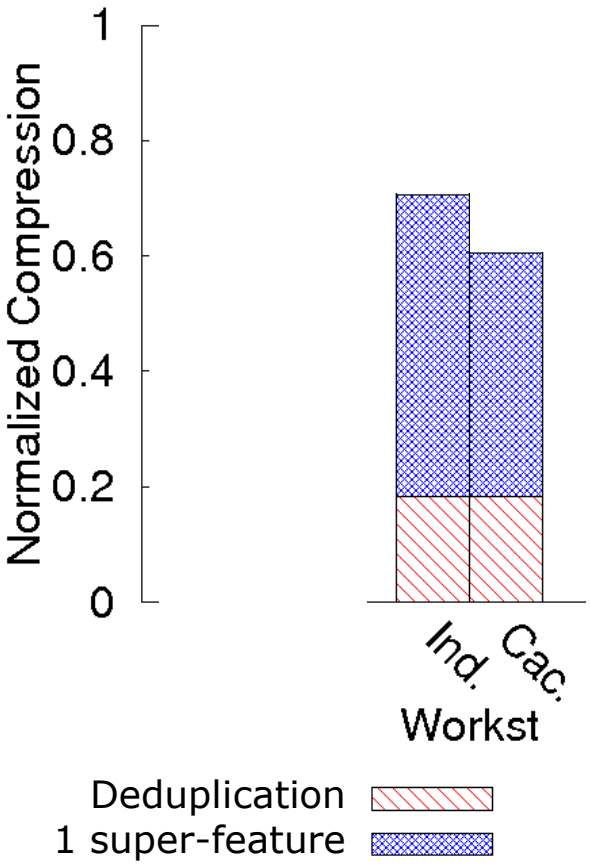
Cache sized at 12 MB per stream



Index VS Stream-Informed Cache

Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

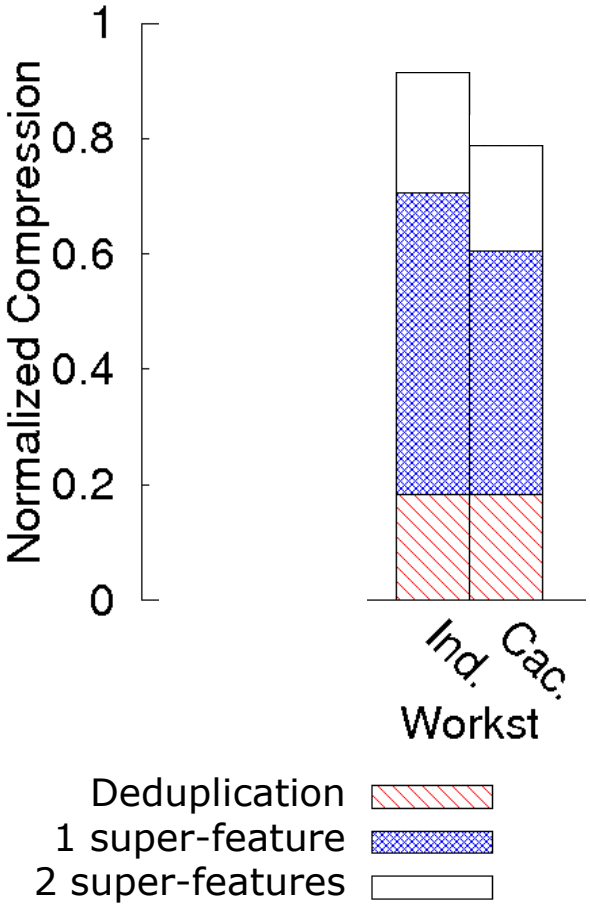


Index VS Stream-Informed Cache

Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

Two super-features in a cache is better than an index with one



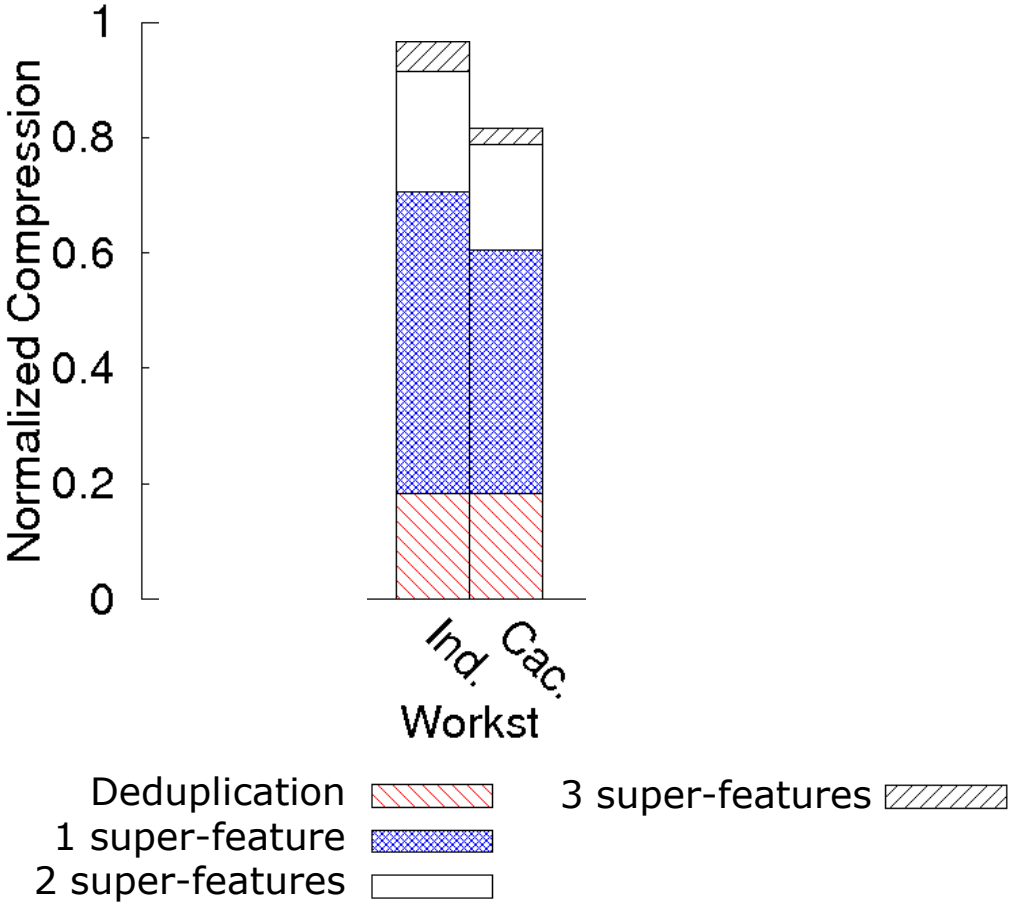
EMC²

Index VS Stream-Informed Cache

Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

Two super-features in a cache is better than an index with one

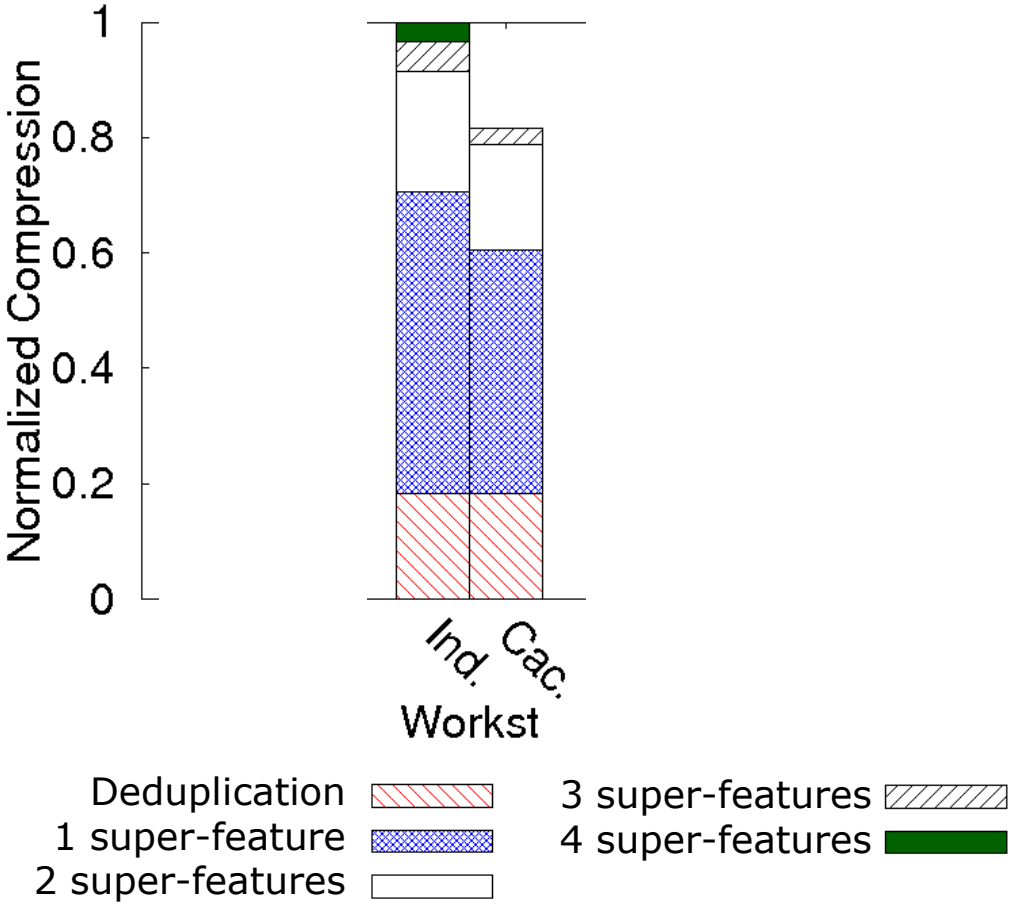


Index VS Stream-Informed Cache

Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

Two super-features in a cache is better than an index with one

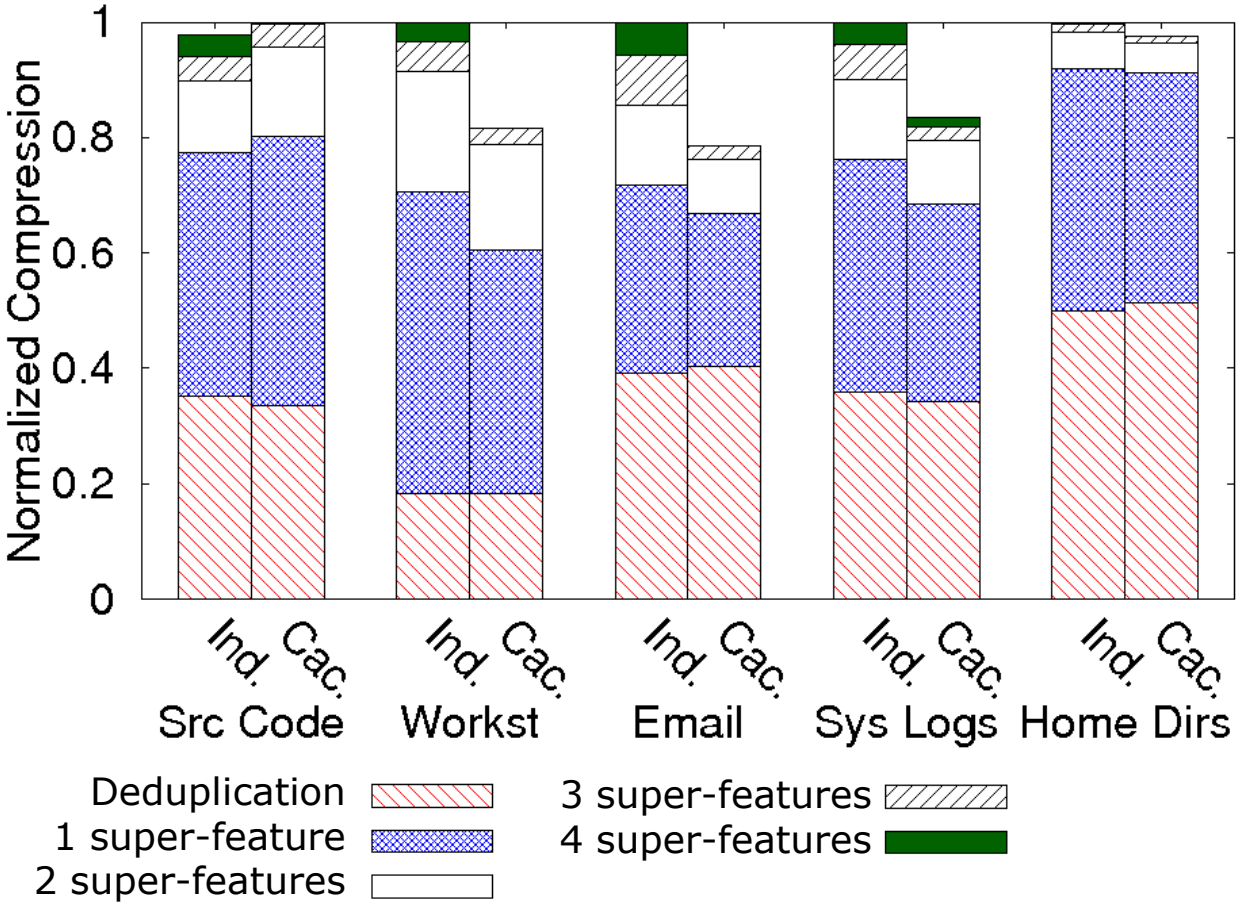


Index VS Stream-Informed Cache

Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

Two super-features in a cache is better than an index with one



Index VS Stream-Informed Cache

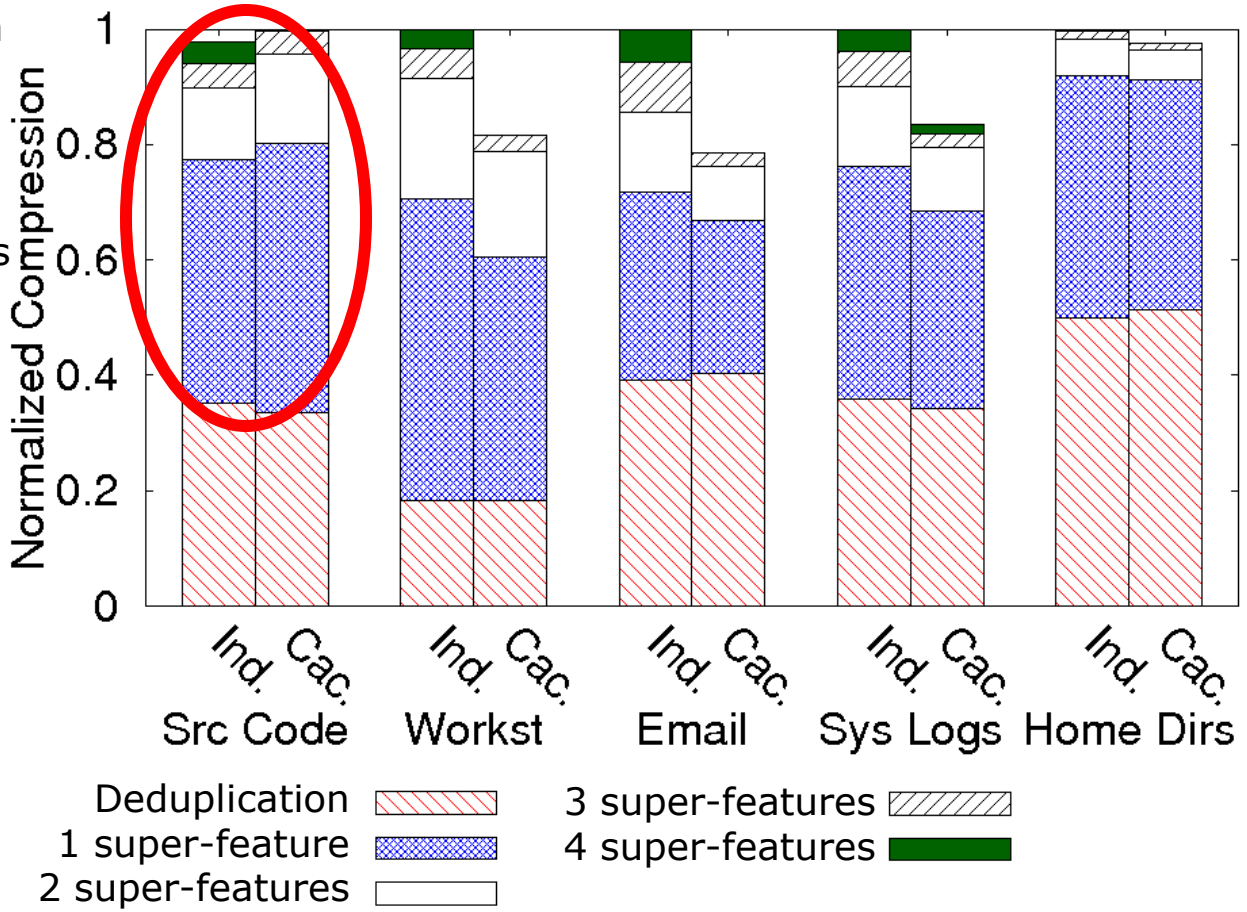
Cache sized at 12 MB per stream

For one super-feature, compression is within 14% of using an index

Two super-features in a cache is better than an index with one

1. Deduplication is slightly different because of different implementations

2. Delta compression in a cache was higher than a full index because of better delta encoding with a candidate in the cache



Delta Compression

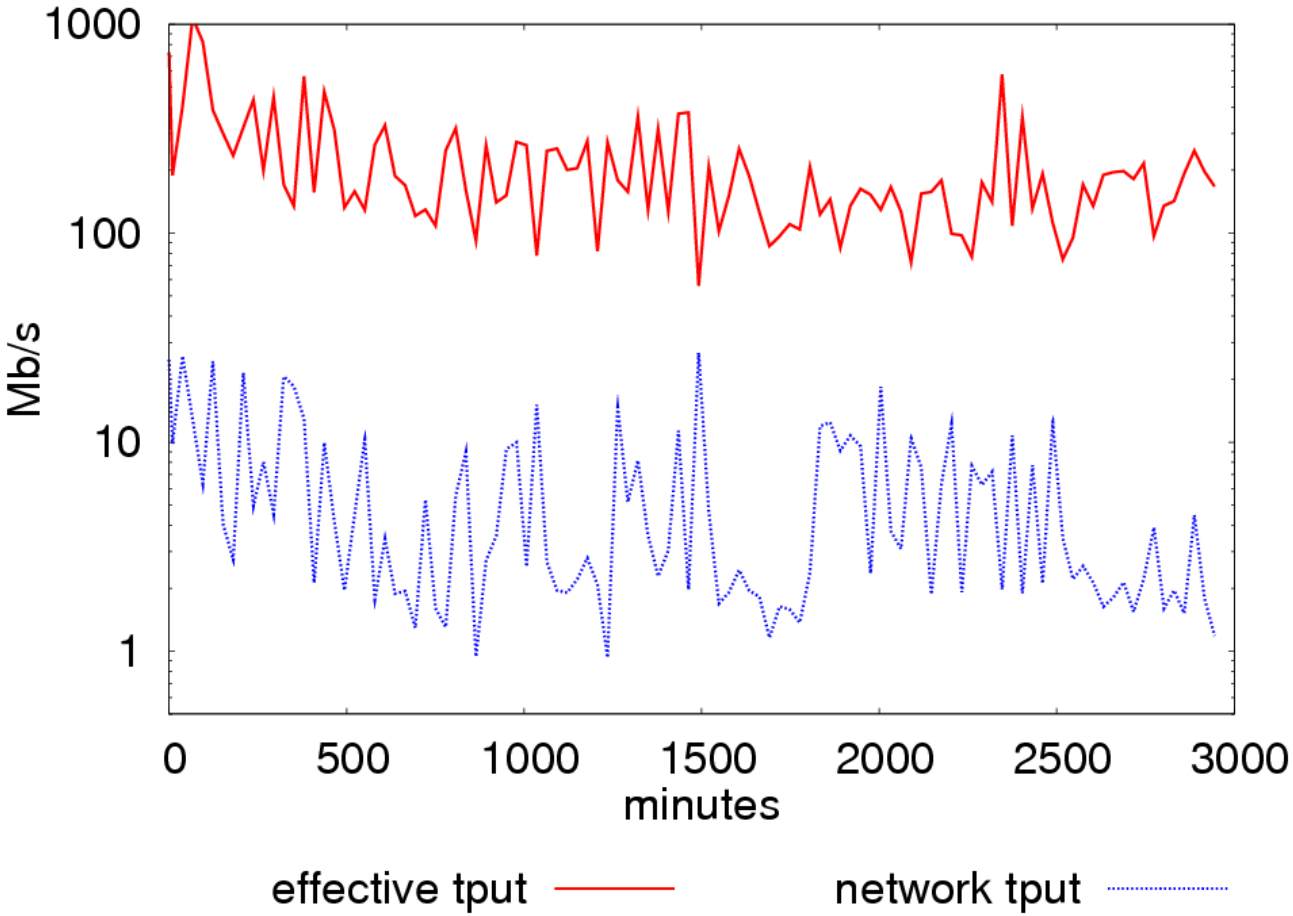
Typical delta improvement is 2X beyond deduplication and GZ compression

Dataset	Deduplication	Choose GZ or Delta with GZ: GZ	Choose GZ or Delta with GZ: Delta w/ GZ	Delta Improv.
Source Code	24.9X	7.2X	14.9X	2.1X
Workstations	5.7X	2.8X	8.8X	3.1X
Email	6.9X	3.1X	5.8X	1.9X
System Logs	57.9X	4.6X	10.2X	2.2X
Home Directories	31.7X	3.1X	5.5X	1.8X

Compression factors are presented after first week of seeding.

Network Throughput

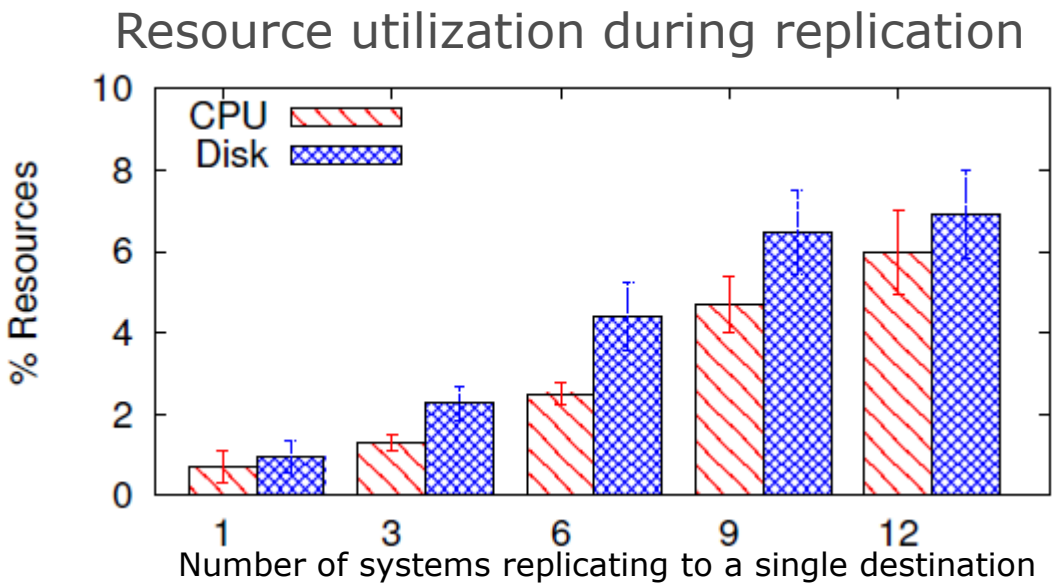
Effective throughput is 1-2 orders of magnitude faster



Overheads And Limitations

- Sketches take up about 20 bytes per non-duplicate chunk
- Uses read IO and CPU on source and destination
  - Sketching is a 20% slowdown on writes, but only for non-duplicates
  - Scales linearly at destination with number of streams

Overheads And Limitations

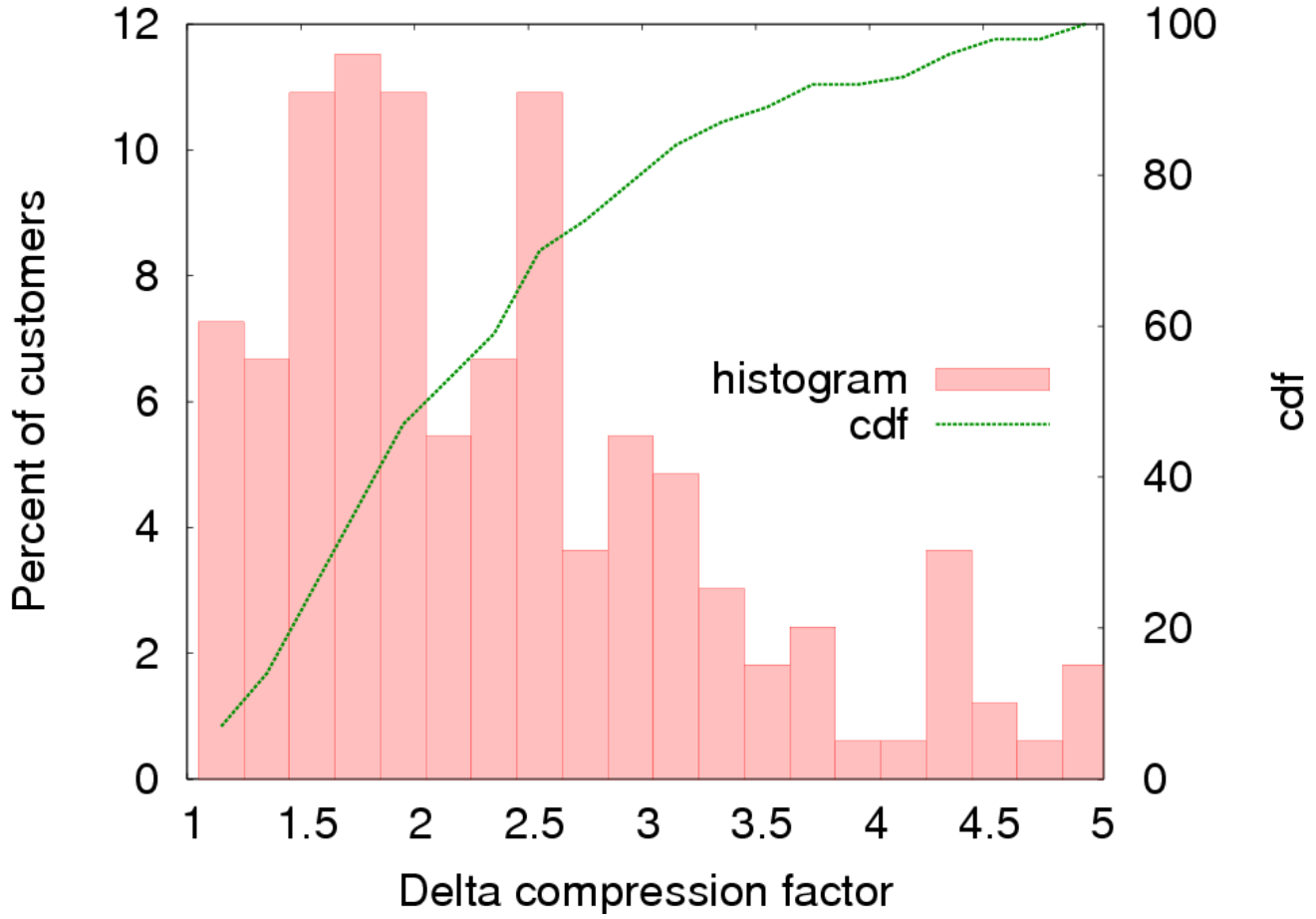


Overheads And Limitations

- Sketches take up about 20 bytes per non-duplicate chunk
- Uses read IO and CPU on source and destination
  - Sketching is a 20% slowdown on writes, but only for non-duplicates
  - Scales linearly at destination with number of streams
- Shared sketch cache affects compression
  - System sized to handle 20 streams
    - With 25 streams compression loss of 0-12%
    - With 50 streams compression loss of 0-27%

Customer Results

Median customer has 2X delta compression beyond deduplication



Related Work

- Optimized network transfer
  - Spring00, Muthitacharoen01, Eshghi07, and Park07
- File synchronization
  - Tridgell00, Suel04
- Delta compression
  - Burns97, Mogul97, Hunt98, Chan99, MacDonald00, Suel02, Trendafilov02, and Chen04
- Similarity detection
  - Brin94, Manber94, Broder[97 & 00], Douglis03, Kulkarni04, You04, Jain05, and Aronovich09
- Deduplicated storage
  - Policroniades04, and Bobbarjung06
- Stream-informed deduplication
  - Zhu08, Bhagwat09, Lillibridge09, Min10, Guo11, and Xia11

EMC²

Conclusion

- Delta locality closely matches deduplication locality for backup datasets
- Good scalability
  - Stream-informed delta compression is effective with a small cache
  - CPU and IO utilization is low
- Product allows customers to replicate and protect twice as much data across a WAN

Questions?

EMC²