Memory Efficient Sanitization of a Deduplicated Storage System

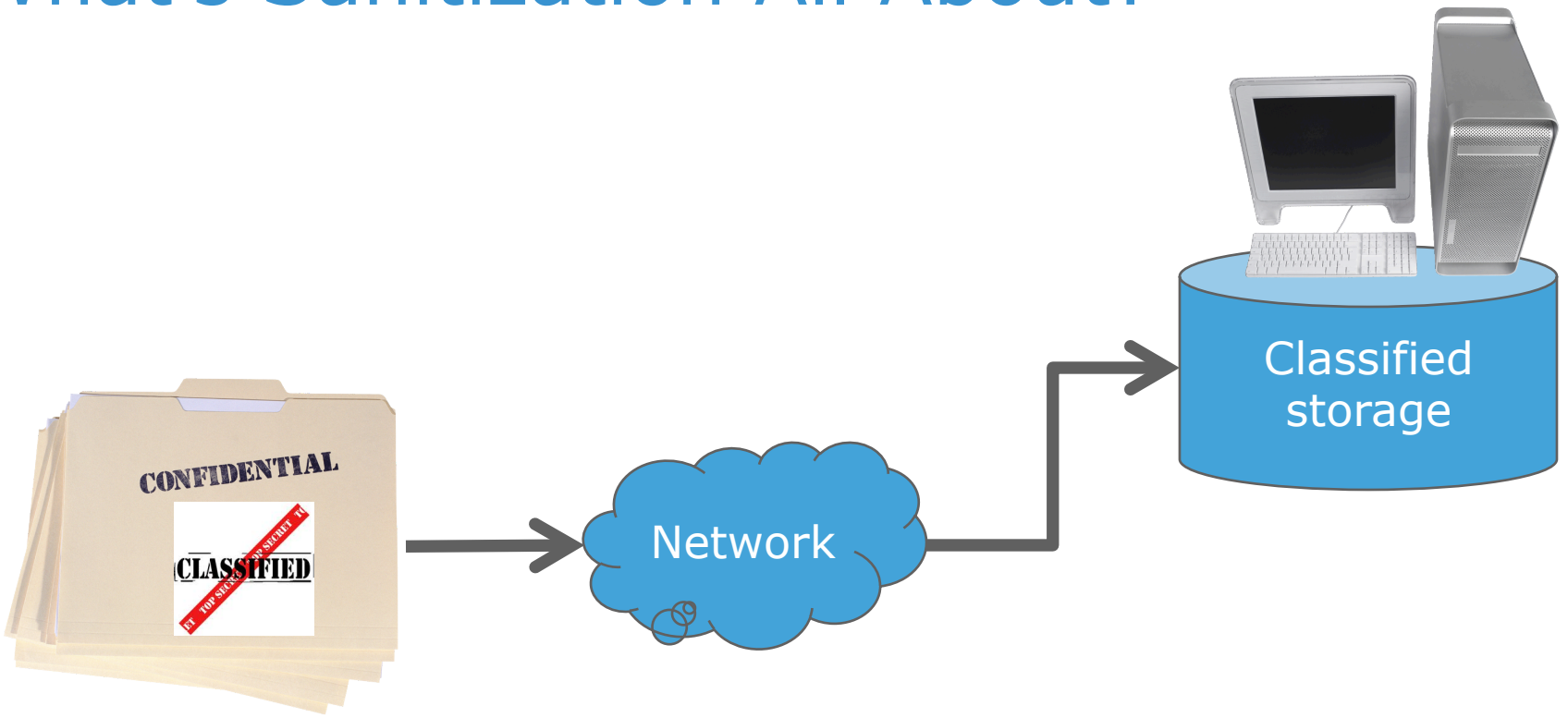
Fabiano C. Botelho
Philip Shilane
Nitin Garg
Windsor Hsu

FAST 2013, San Jose, February 12-15

EMC²

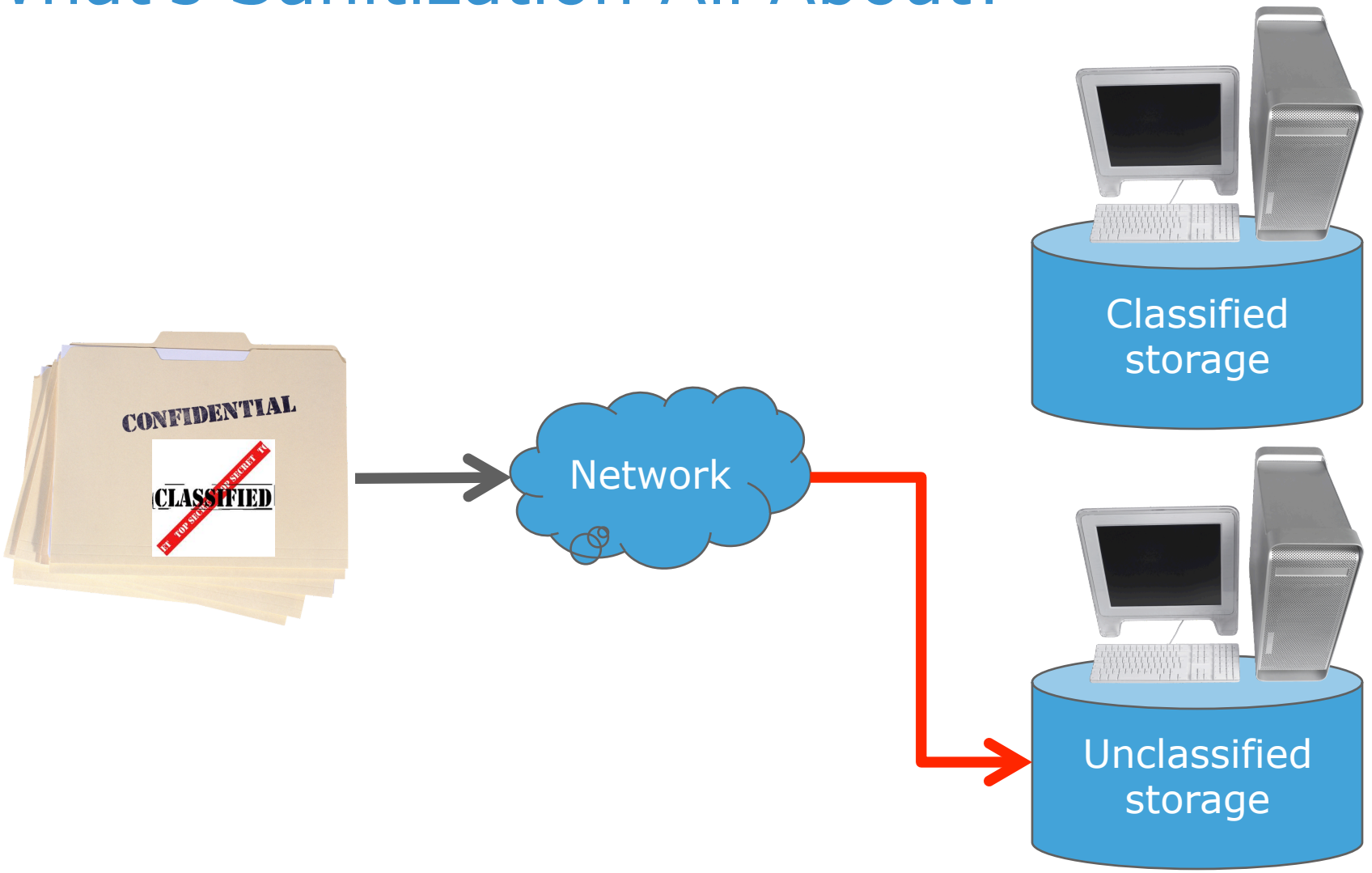
© Copyright 2013 EMC Corporation. All rights reserved.

What's Sanitization All About?



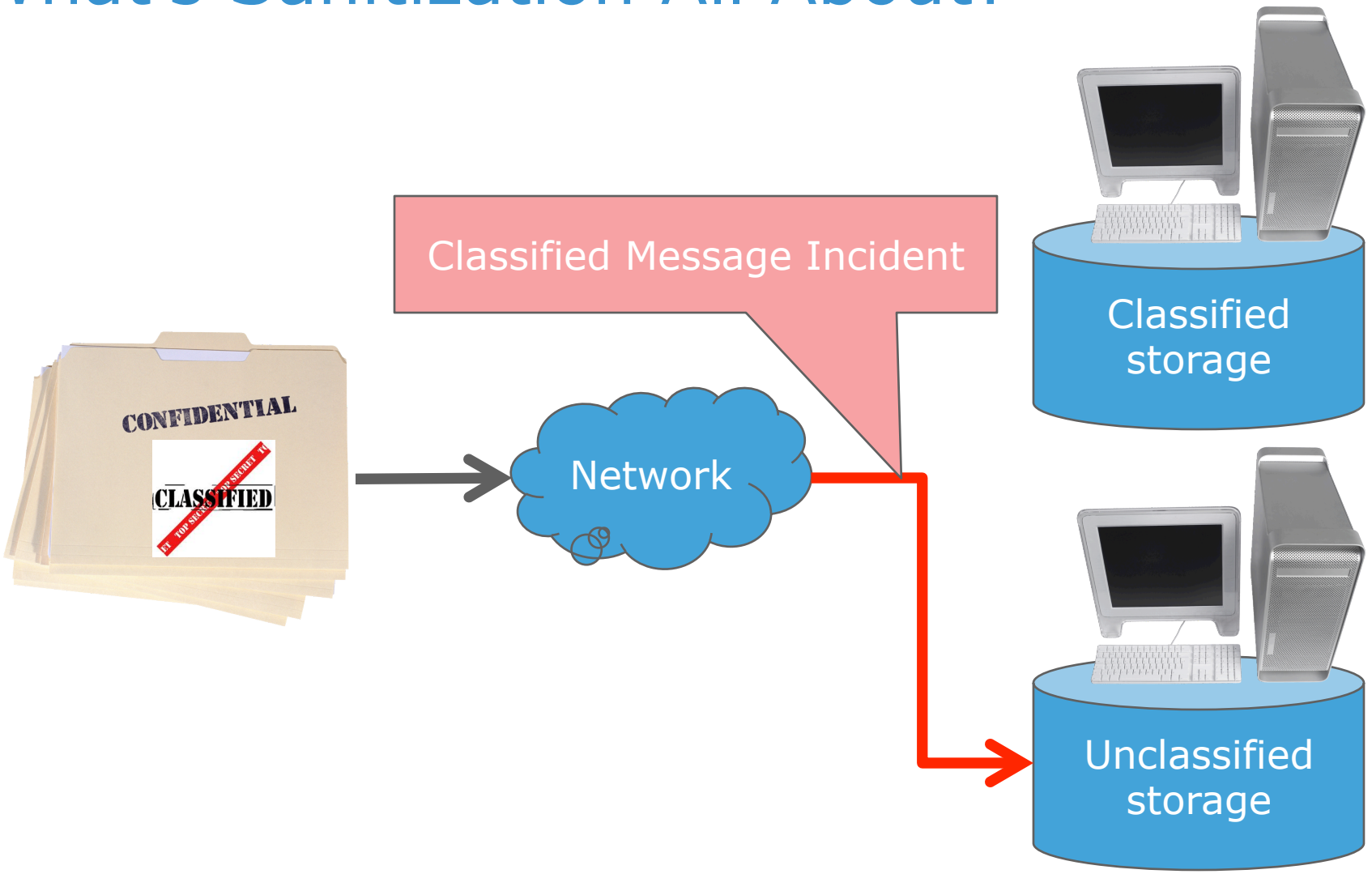
© Copyright 2013 EMC Corporation. All rights reserved.

What's Sanitization All About?



© Copyright 2013 EMC Corporation. All rights reserved.

What's Sanitization All About?



© Copyright 2013 EMC Corporation. All rights reserved.

How Do We Define Sanitization?

A process to restore the storage system to a state as if the classified message incident had never occurred

© Copyright 2013 EMC Corporation. All rights reserved.

Threat Models

1. Casual Attacks:
   - Access through regular file system interfaces
   - NFS, CIFIS, etc
2. Robust Keyboard Attack
   - Access through non-regular interfaces
   - Reading blocks directly from disk, swap areas, or unallocated blocks
3. Laboratory Attack
   - Access through exotic laboratory techniques
   - Require specific disk format knowledge and specialized hardware
   - Even after overwrites, the disk may retain magnetic values indicating a previous state

EMC²

© Copyright 2013 EMC Corporation. All rights reserved.

NIST, DoD Guidelines For Sanitization

1. Clearing Level
   - Single overwrite of affected areas is enough to protect against casual and keyboard attacks.
2. Purging Level:
   - Devices must be either degaussed or destroyed to protect against laboratory attacks

© Copyright 2013 EMC Corporation. All rights reserved.

Why Not Crypto Sanitization

1. Crypto sanitization:
   - Encrypt each file with a different key and throw away the key of the affected files.
2. Deduplication is challenging because blocks are shared
3. Key management is a huge hassle
4. NIST and DoD guidelines explicitly say that encryption is not acceptable
5. Sacrifices performance of normal FS operations
   - Read, write, replication

© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Of A System vs A Device

1. Device:
   - either overwrite each sector or degauss it.
2. In-place System:
   - following meta data references to the physical location within the storage system, overwriting the values one or more times, and erasing the meta data as well as other locations that have become unreferenced
3. Deduplicated System:
   - Usually they are log-structured with large units of writes
   - No in-place erasure of sub-units.
   - Copy forward live data and then erase an earlier region

EMC²

© Copyright 2013 EMC Corporation. All rights reserved.

Bulk Sanitization

1. Sanitizing individual files is challenging
    1. Need to track blocks/chunks that only belong to the file
    2. What if the file has already been deleted?
2. We sanitize the entire system

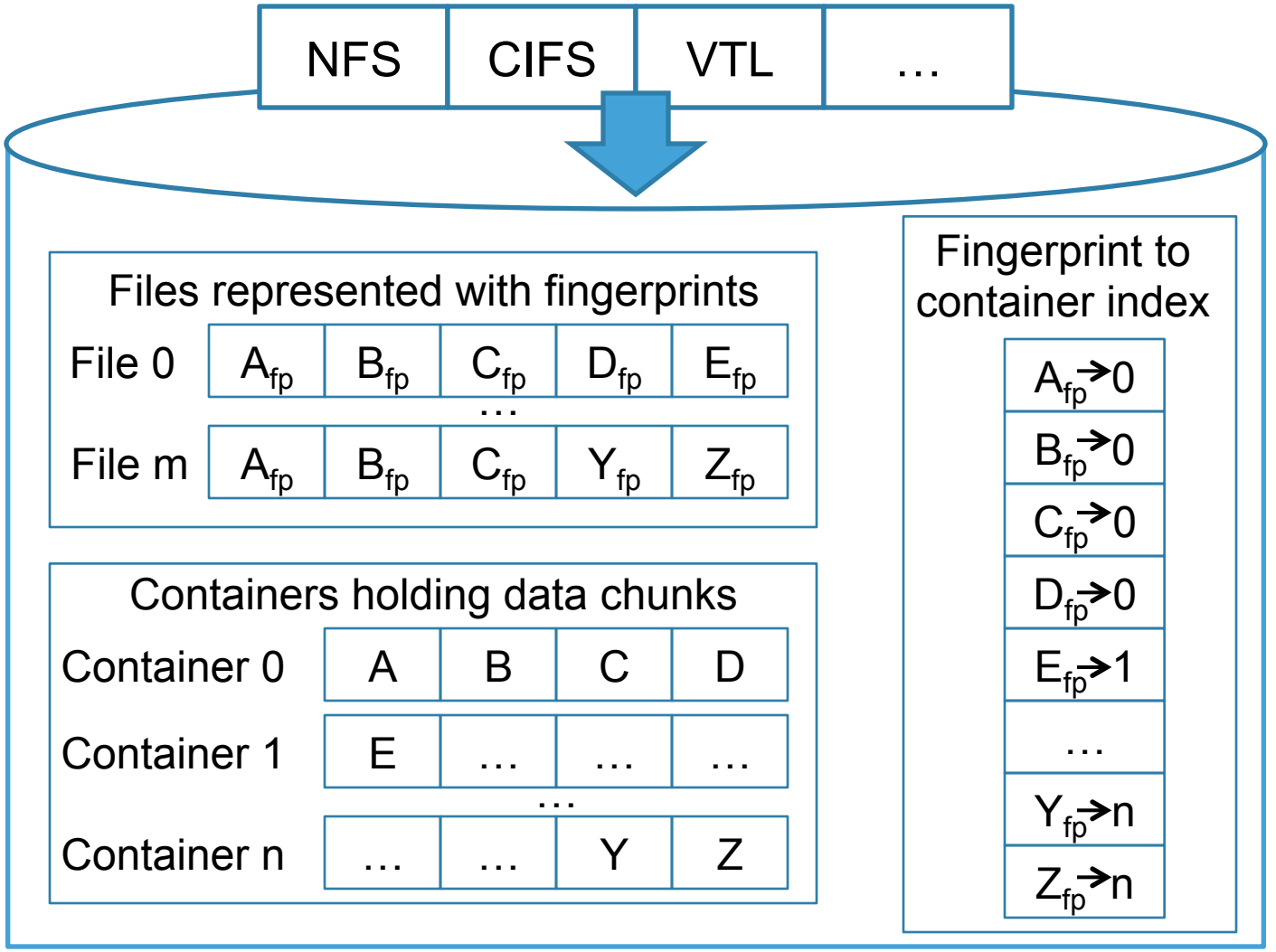
© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Requirements

1. All deleted data are erased
2. All live data are available
3. Sanitization is efficient
4. The storage system is usable while sanitization runs.

© Copyright 2013 EMC Corporation. All rights reserved.

Deduplicated Storage



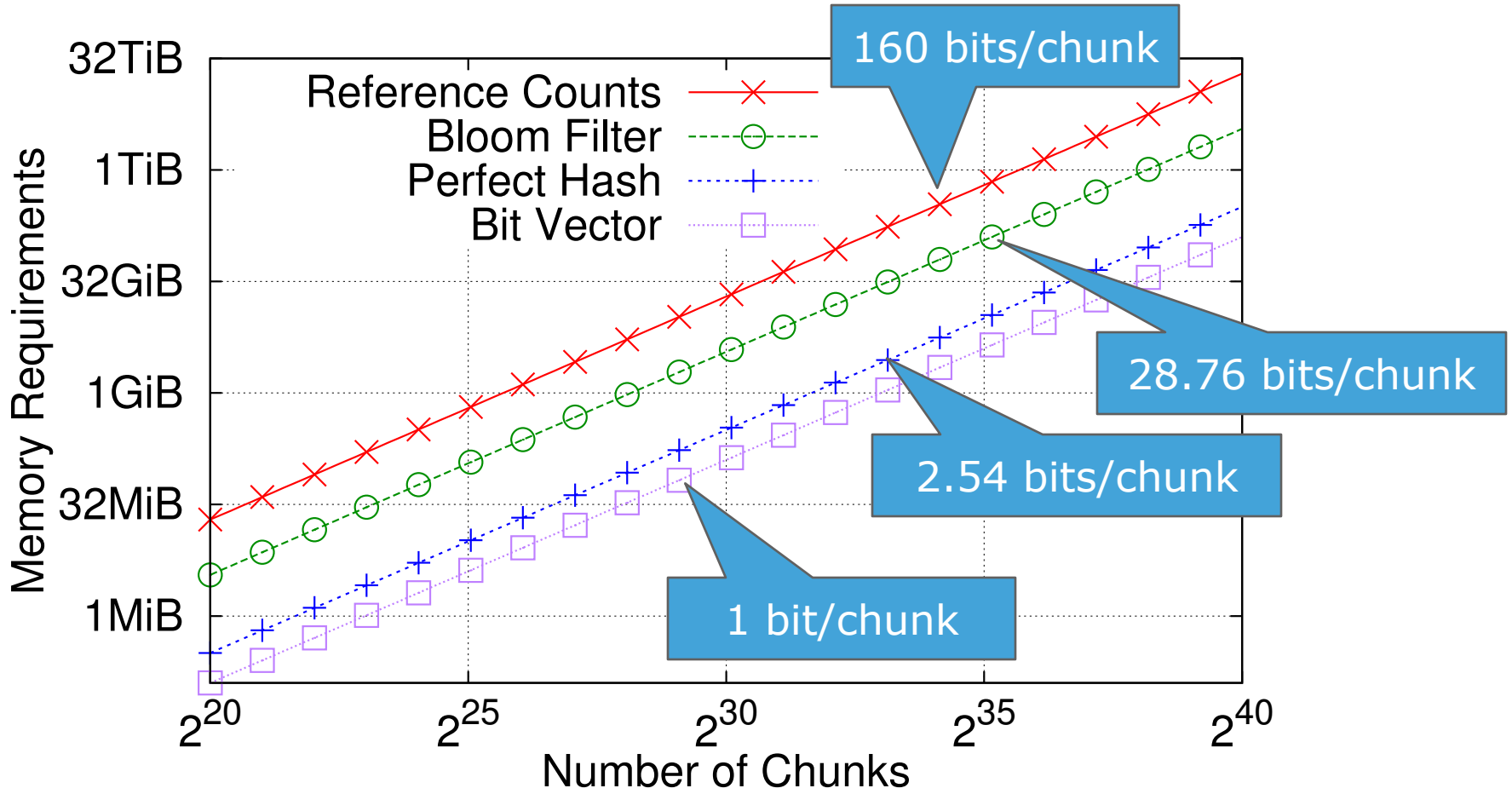
© Copyright 2013 EMC Corporation. All rights reserved.

Challenge With Chunk References

- Huge Fingerprint Set In High-End Systems
  - Physical capacity: 560TiB
  - Avg chunk size: 8 KiB, 4KiB after compression
  - Number of chunks: 140 billions

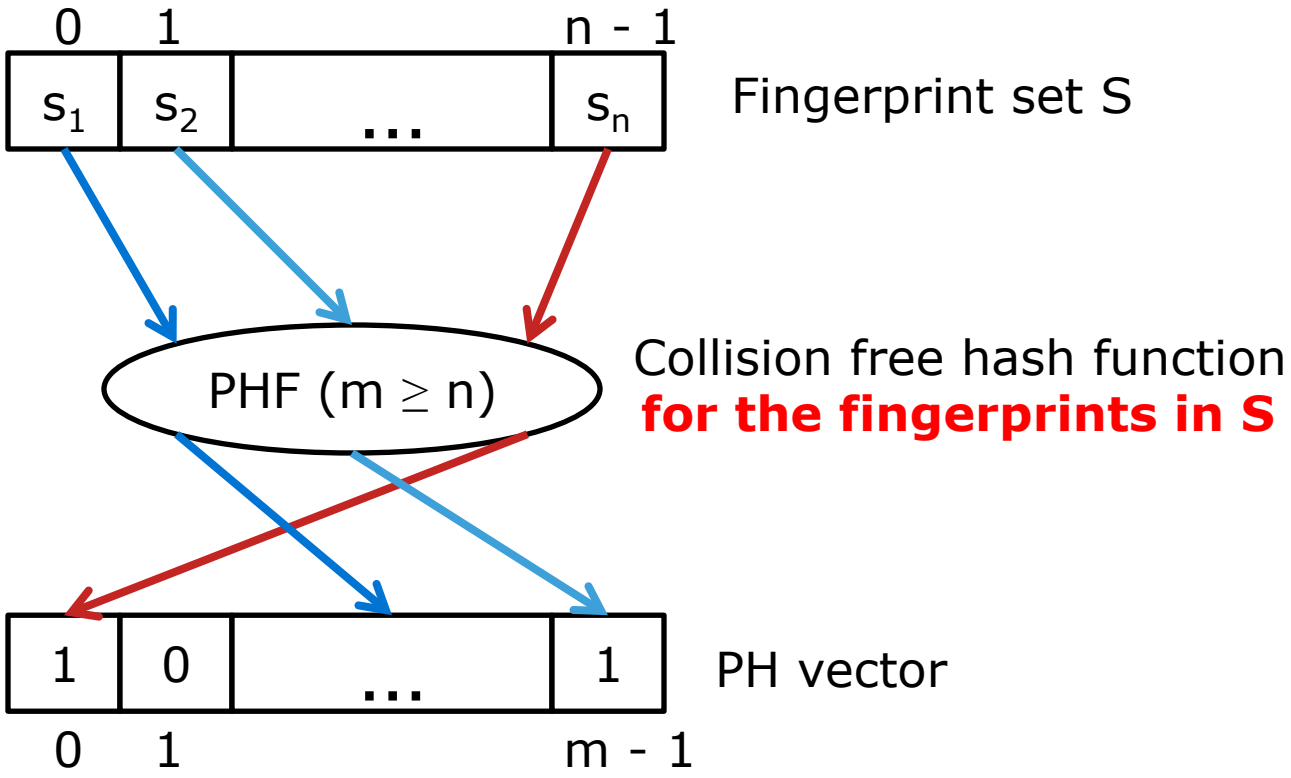
© Copyright 2013 EMC Corporation. All rights reserved.

Memory Usage Of Each Approach



© Copyright 2013 EMC Corporation. All rights reserved.

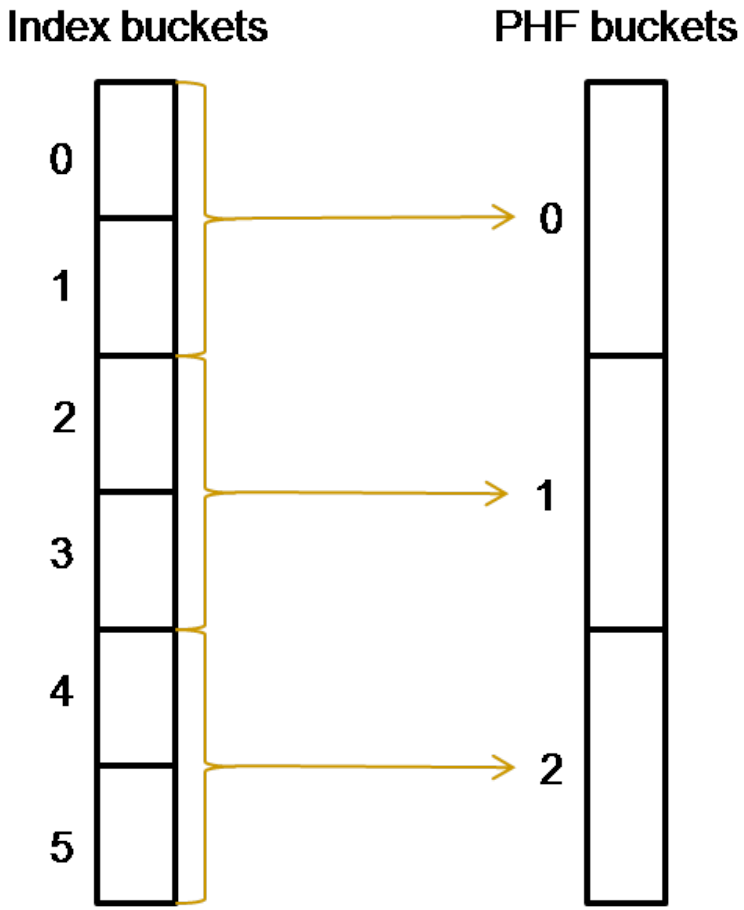
Perfect Hashing Vector (PH_{vec})



$|PH_{vec}| = |PHF| + |PH\ vector|$

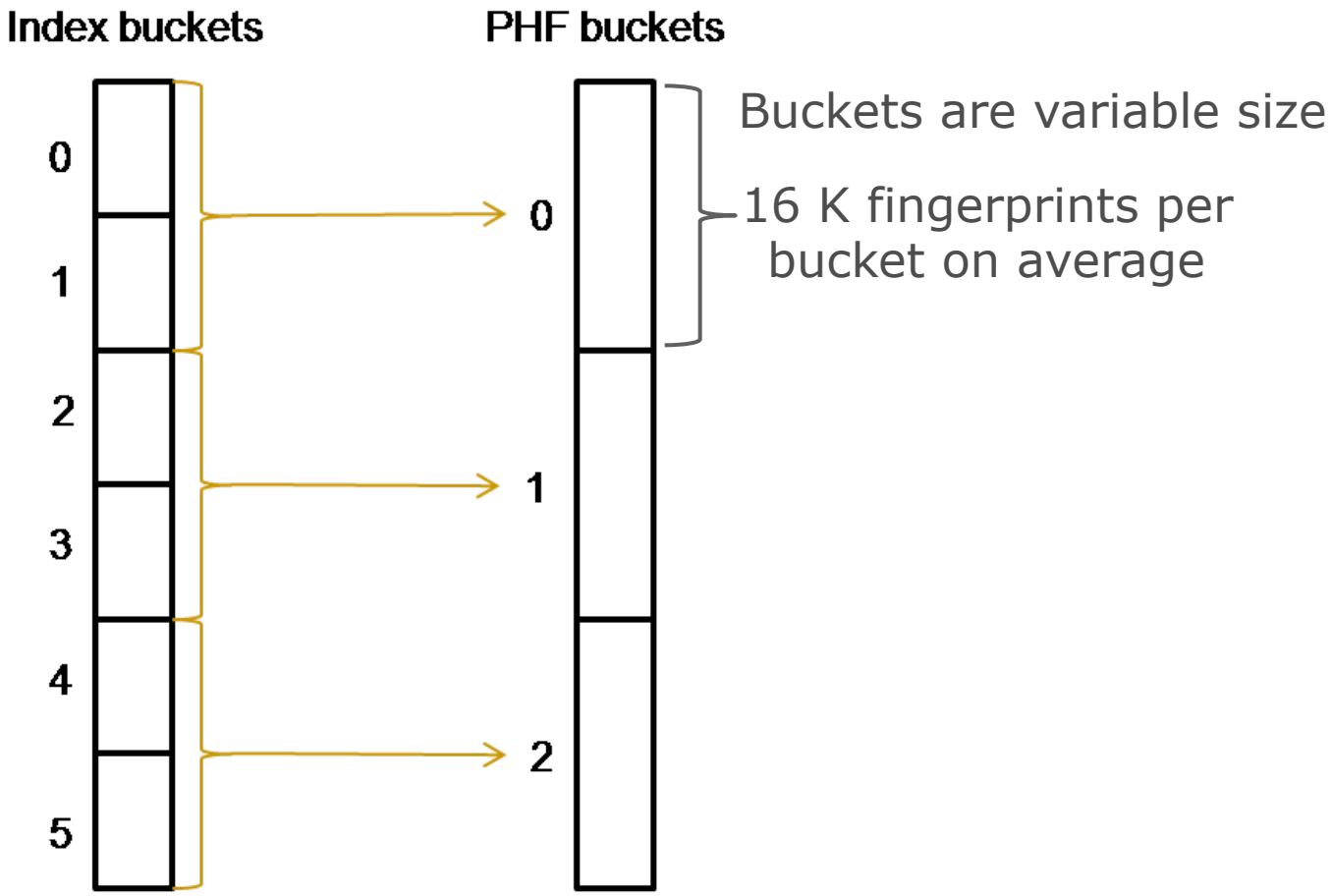
© Copyright 2013 EMC Corporation. All rights reserved.

Bucketizing The Huge Fingerprint Set



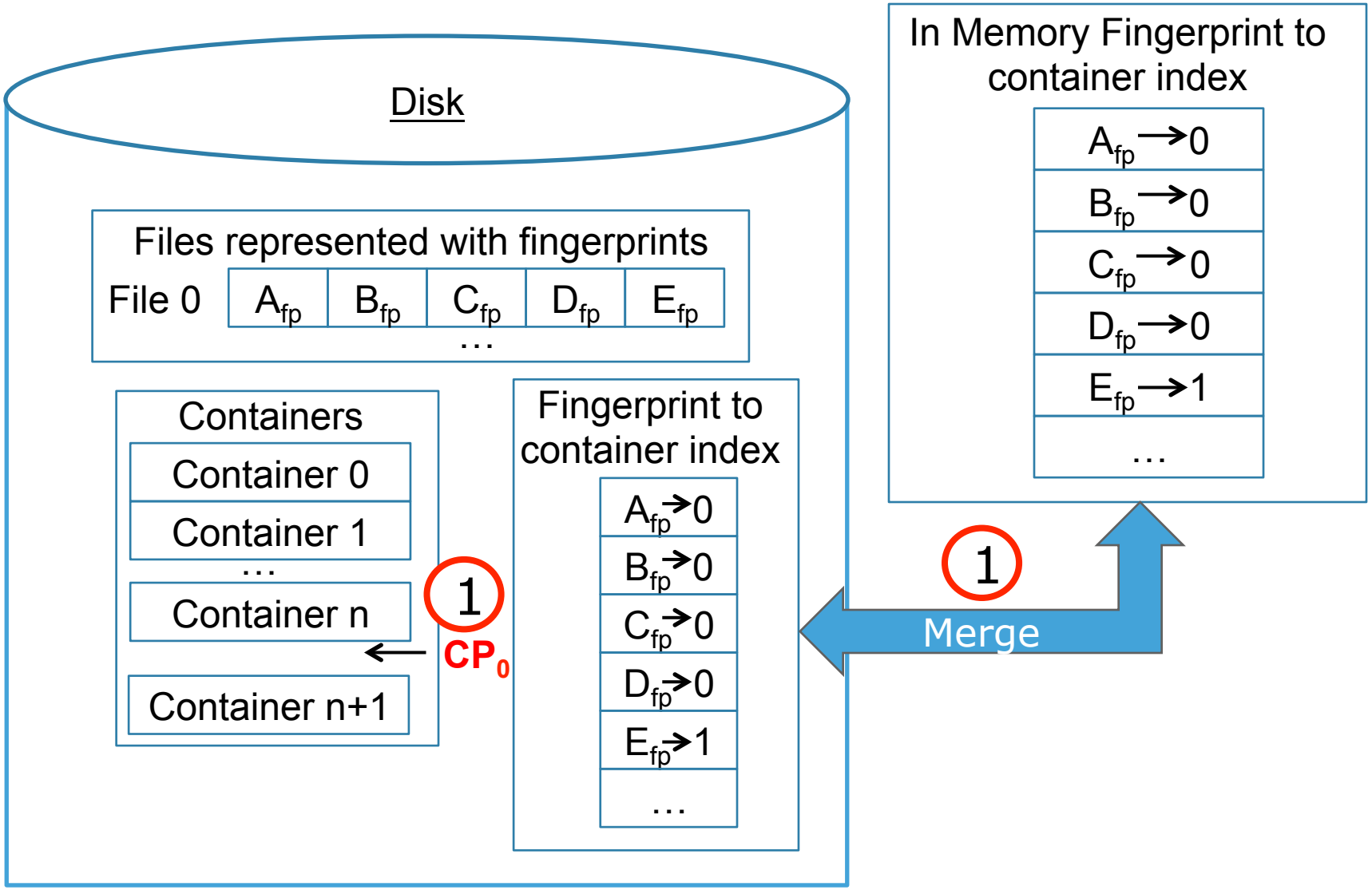
© Copyright 2013 EMC Corporation. All rights reserved.

Bucketizing The Huge Fingerprint Set



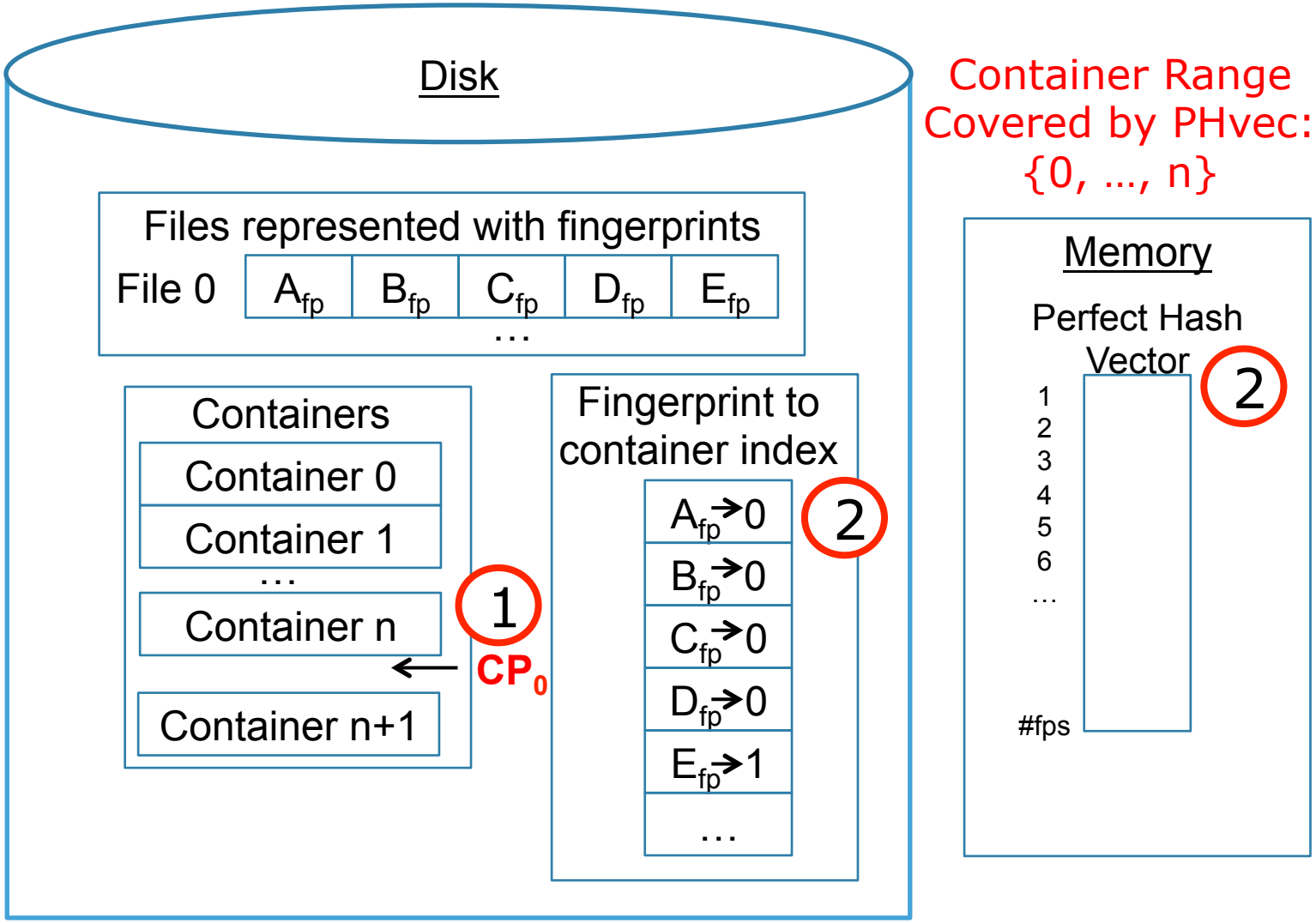
© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Process – Merge (1)



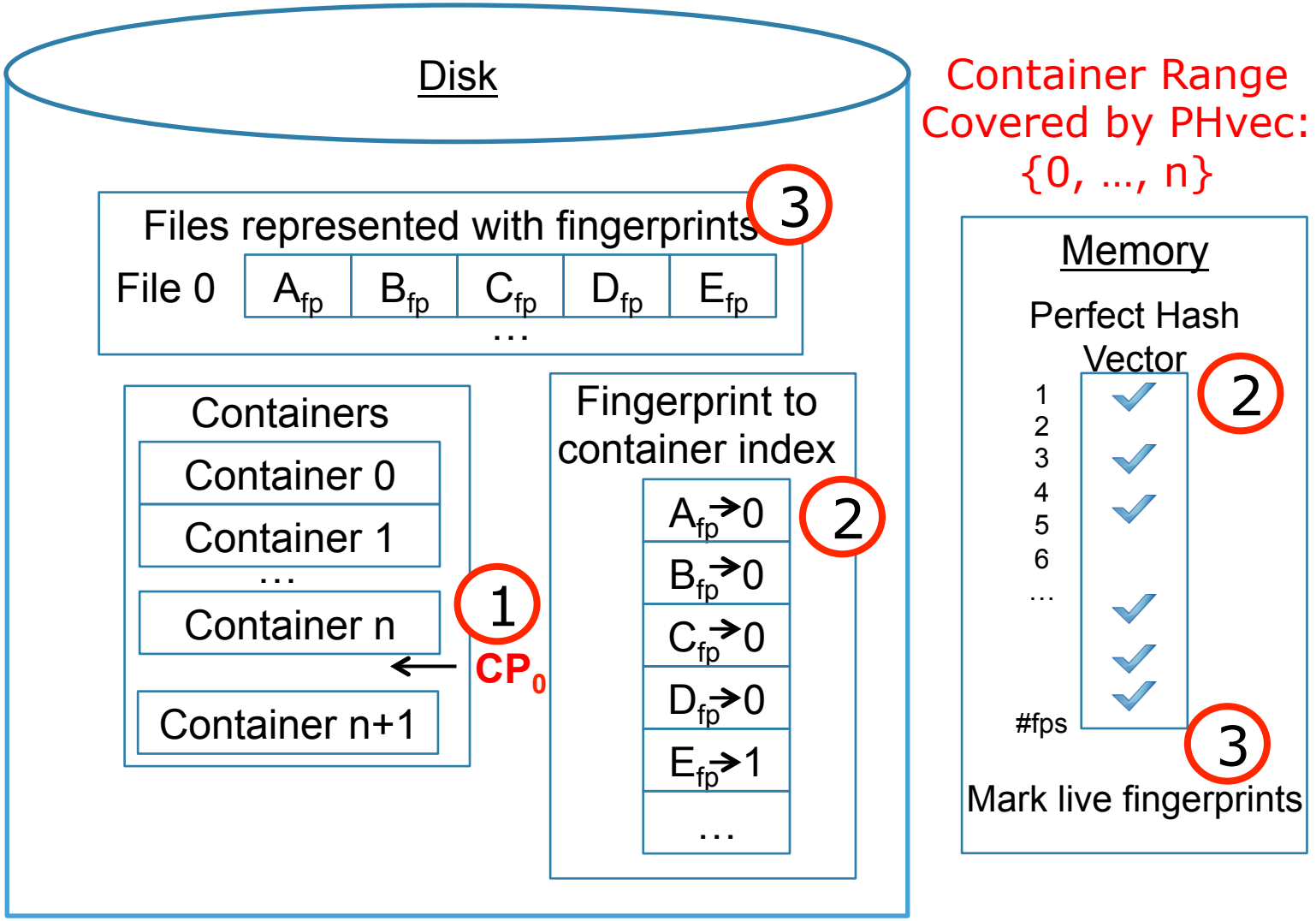
© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Process – Analysis (2)



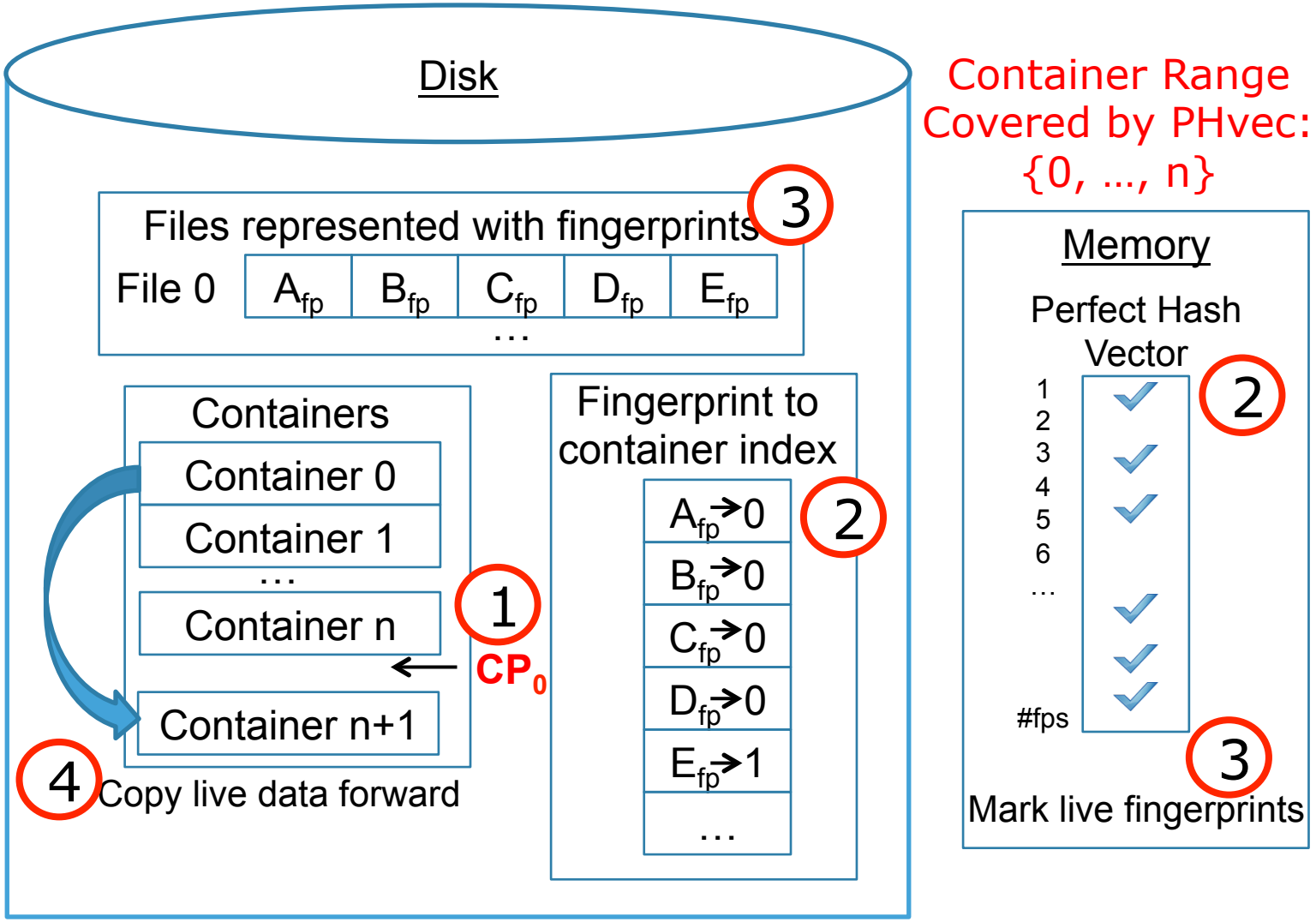
© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Process – Enumeration (3)



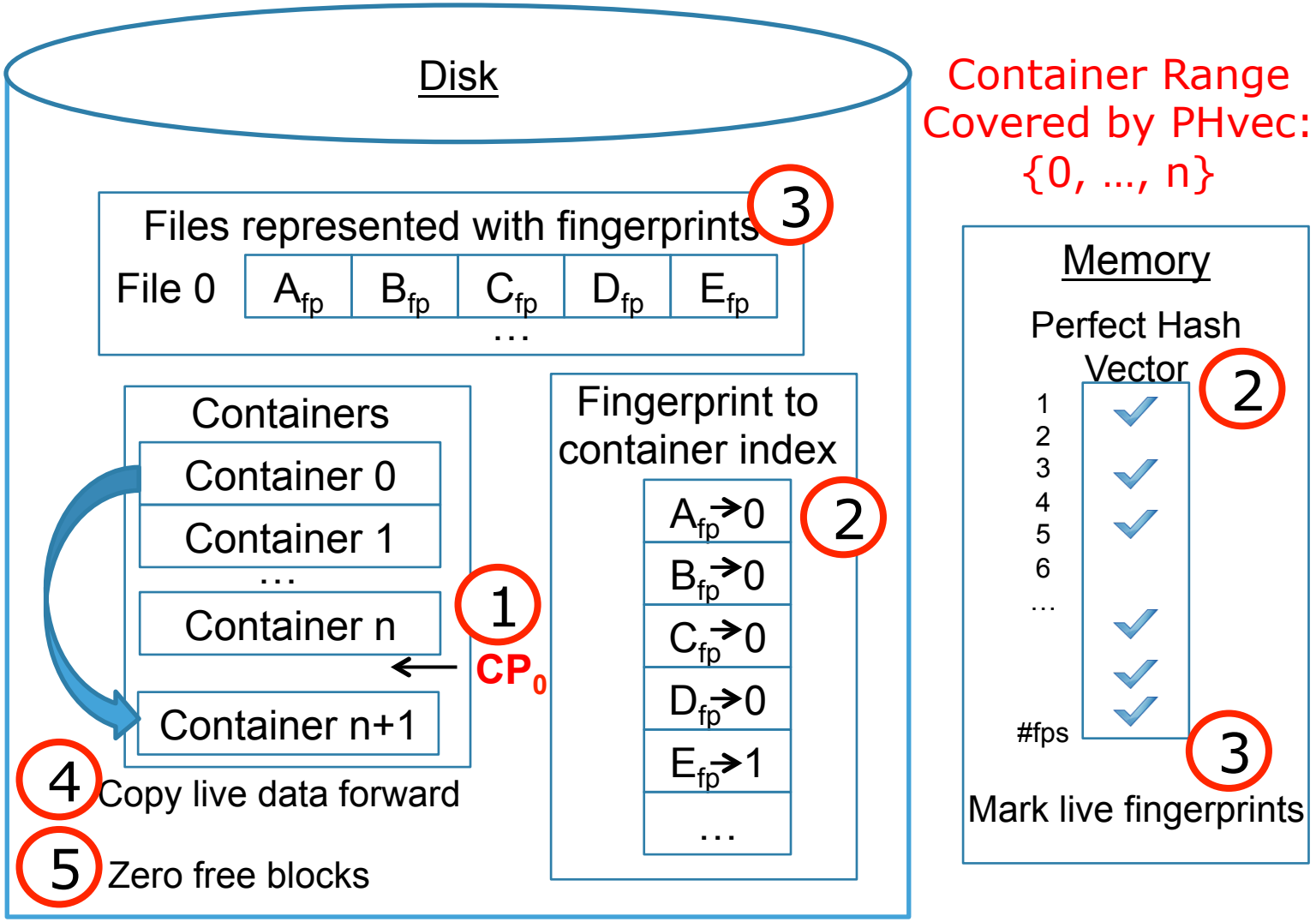
© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Process – Copy (4)



© Copyright 2013 EMC Corporation. All rights reserved.

Sanitization Process – Zero (5)



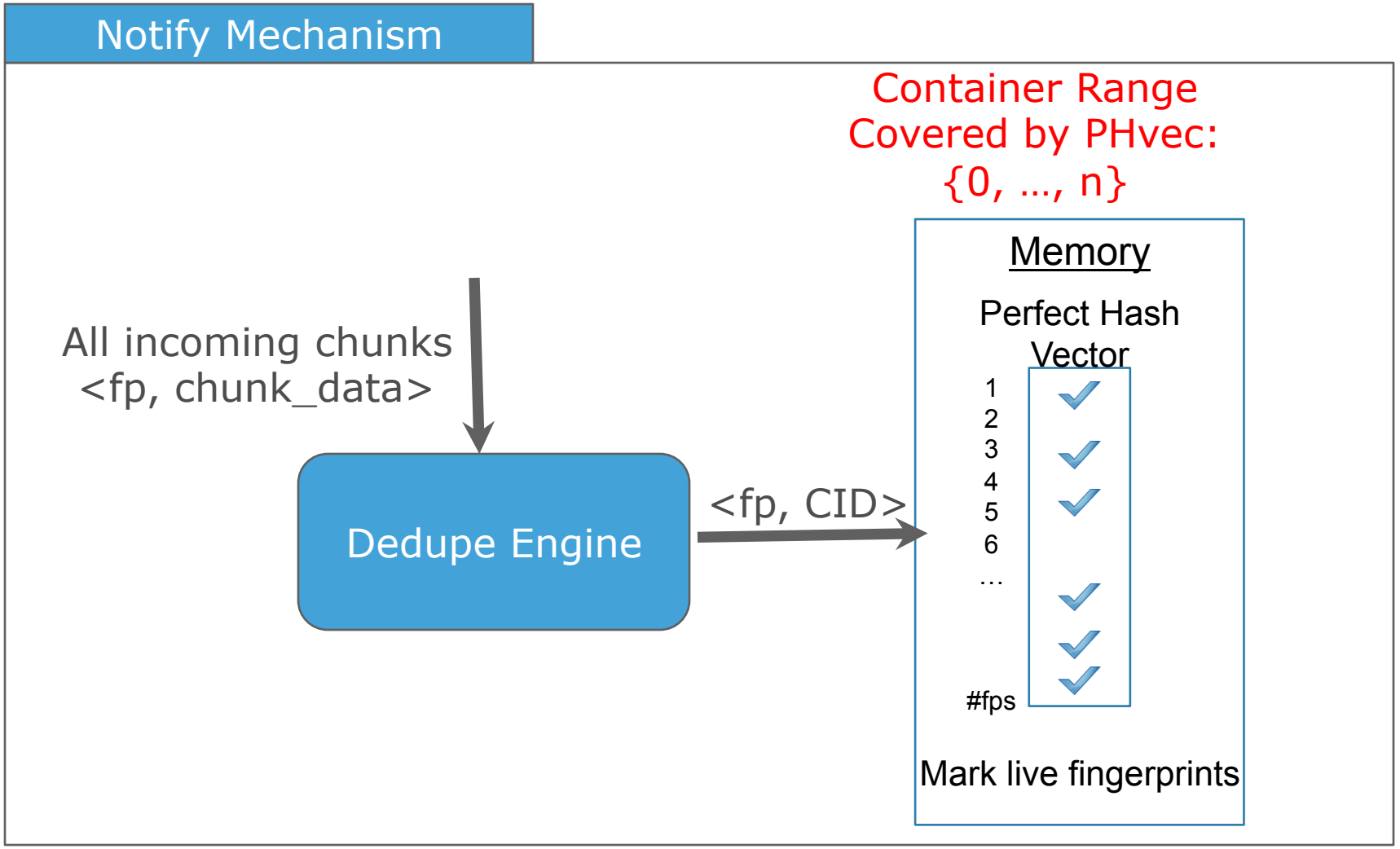
© Copyright 2013 EMC Corporation. All rights reserved.

Issues to Support Read-Write Mode

1. How do we handle resurrections?
2. How do we update the PHvec structure for fingerprints that came in after CP0 has been taken but before the PHvec structure was constructed in the analysis phase?

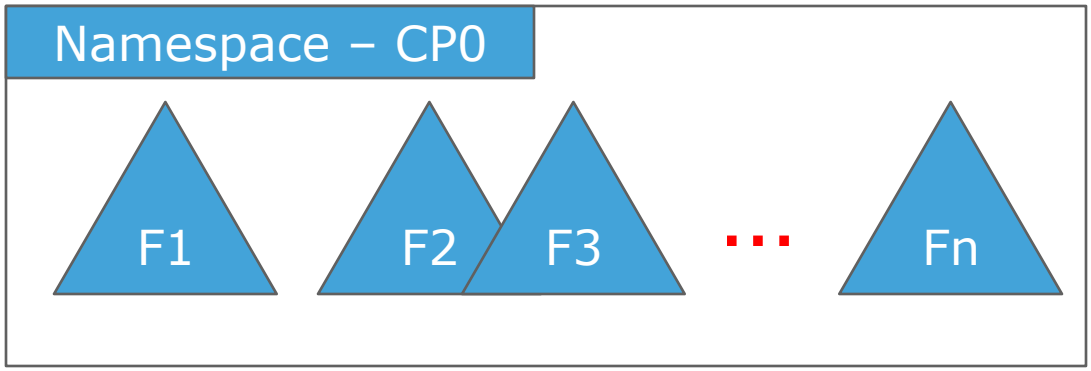
© Copyright 2013 EMC Corporation. All rights reserved.

Handling Resurrections



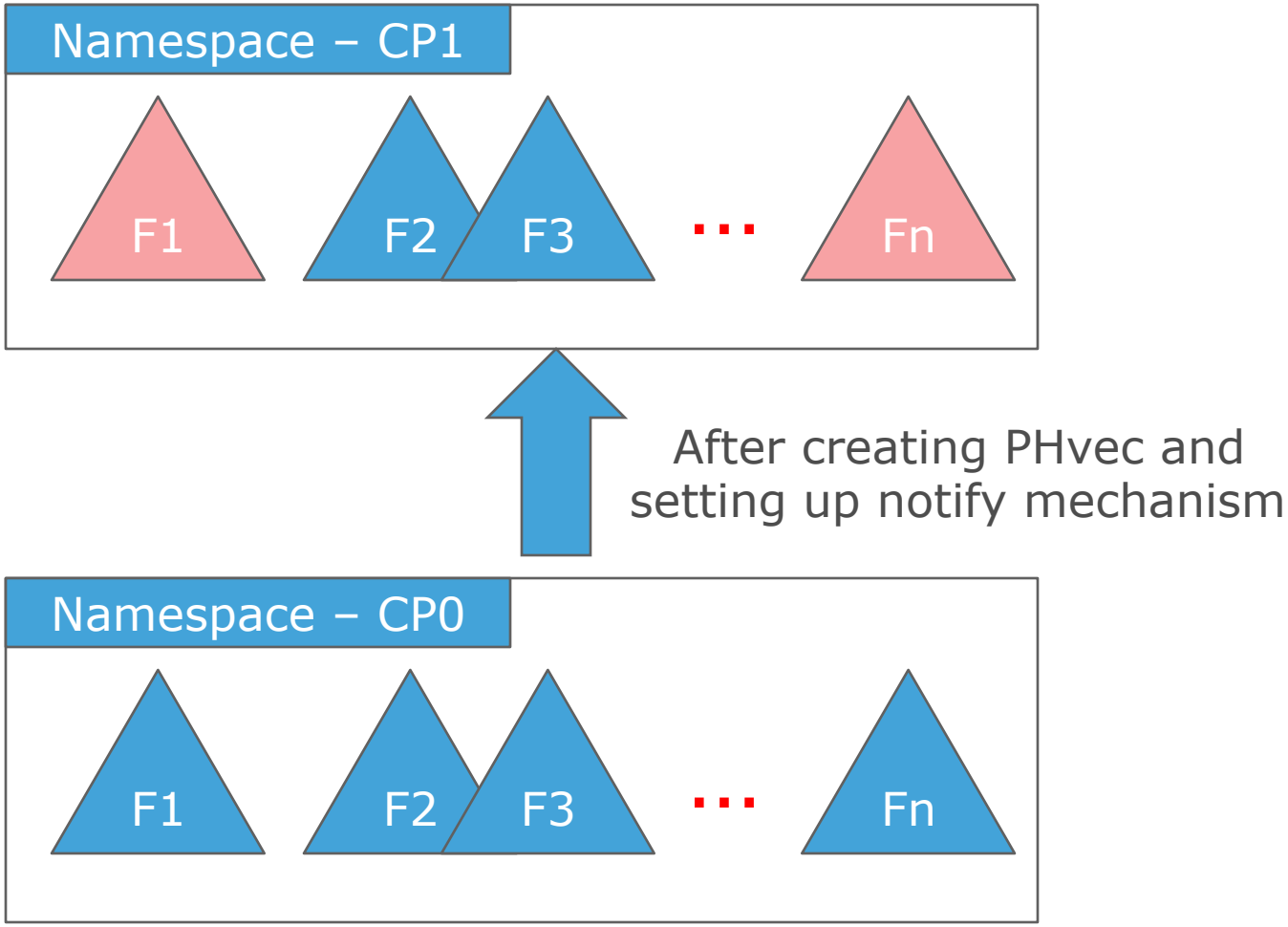
© Copyright 2013 EMC Corporation. All rights reserved.

Handling Fingerprints Not Referenced by First Consistency Point (CP0)



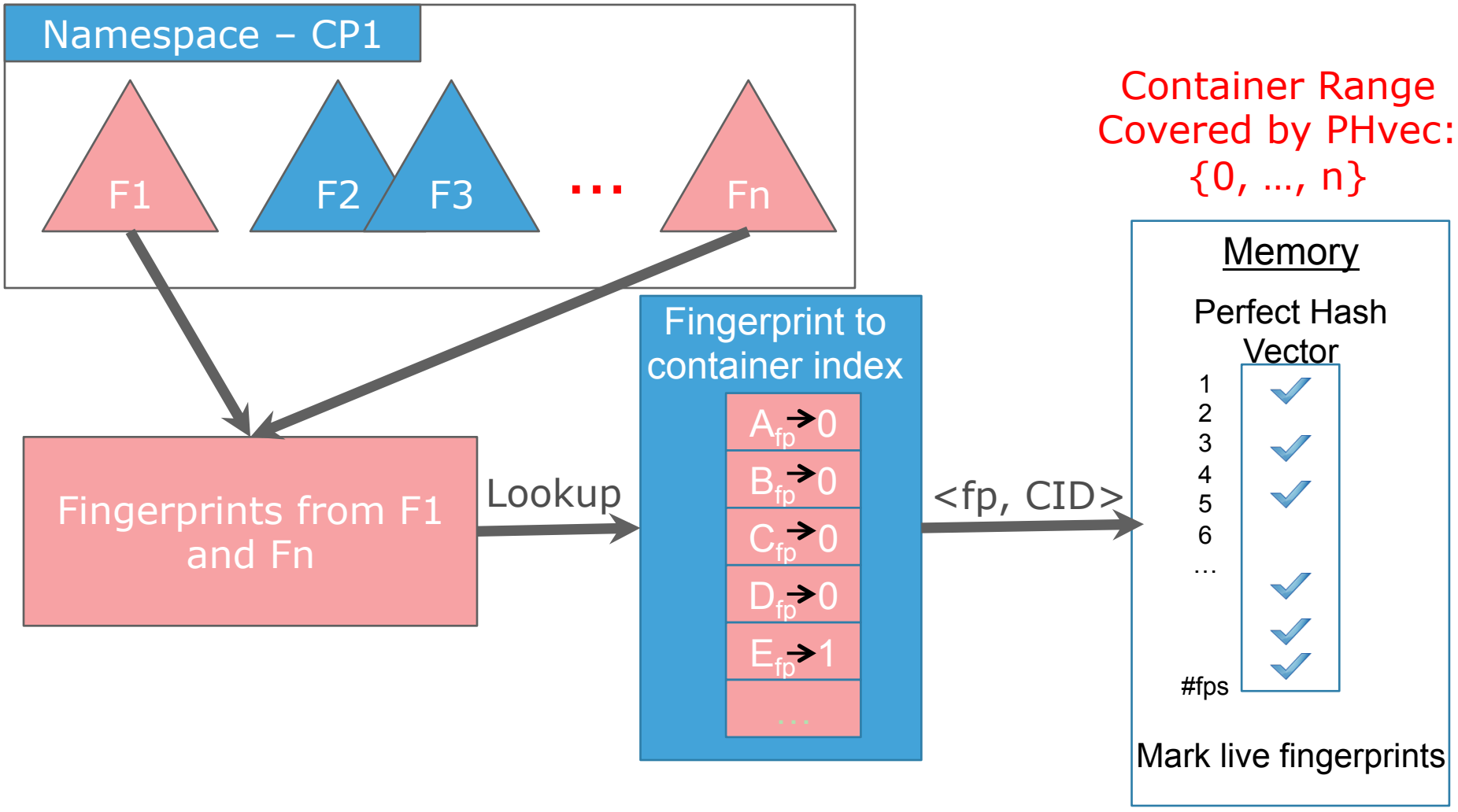
© Copyright 2013 EMC Corporation. All rights reserved.

Handling Fingerprints Not Referenced by First Consistency Point (CP0)



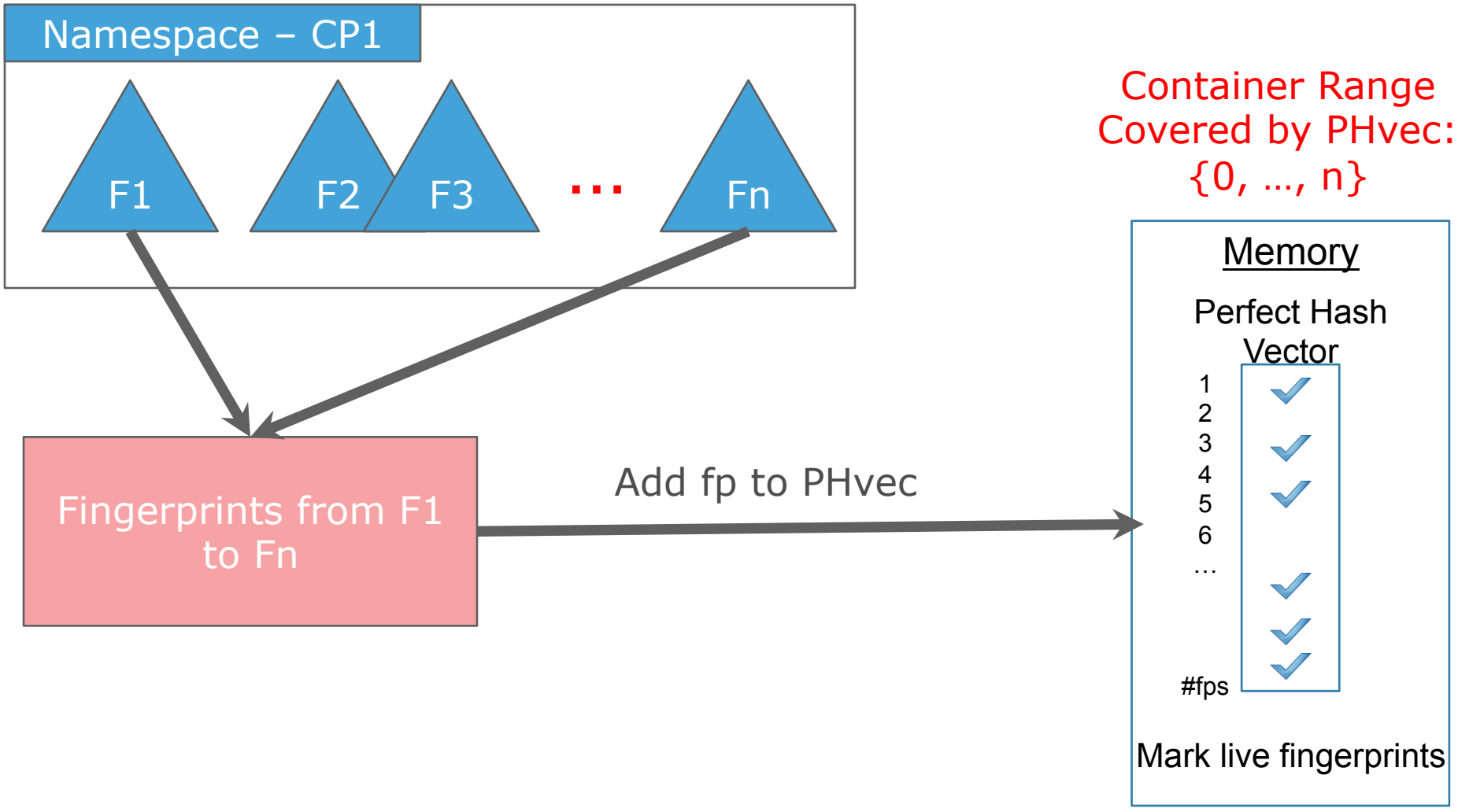
© Copyright 2013 EMC Corporation. All rights reserved.

Enumerate CP1



© Copyright 2013 EMC Corporation. All rights reserved.

Enumerate CP0



© Copyright 2013 EMC Corporation. All rights reserved.

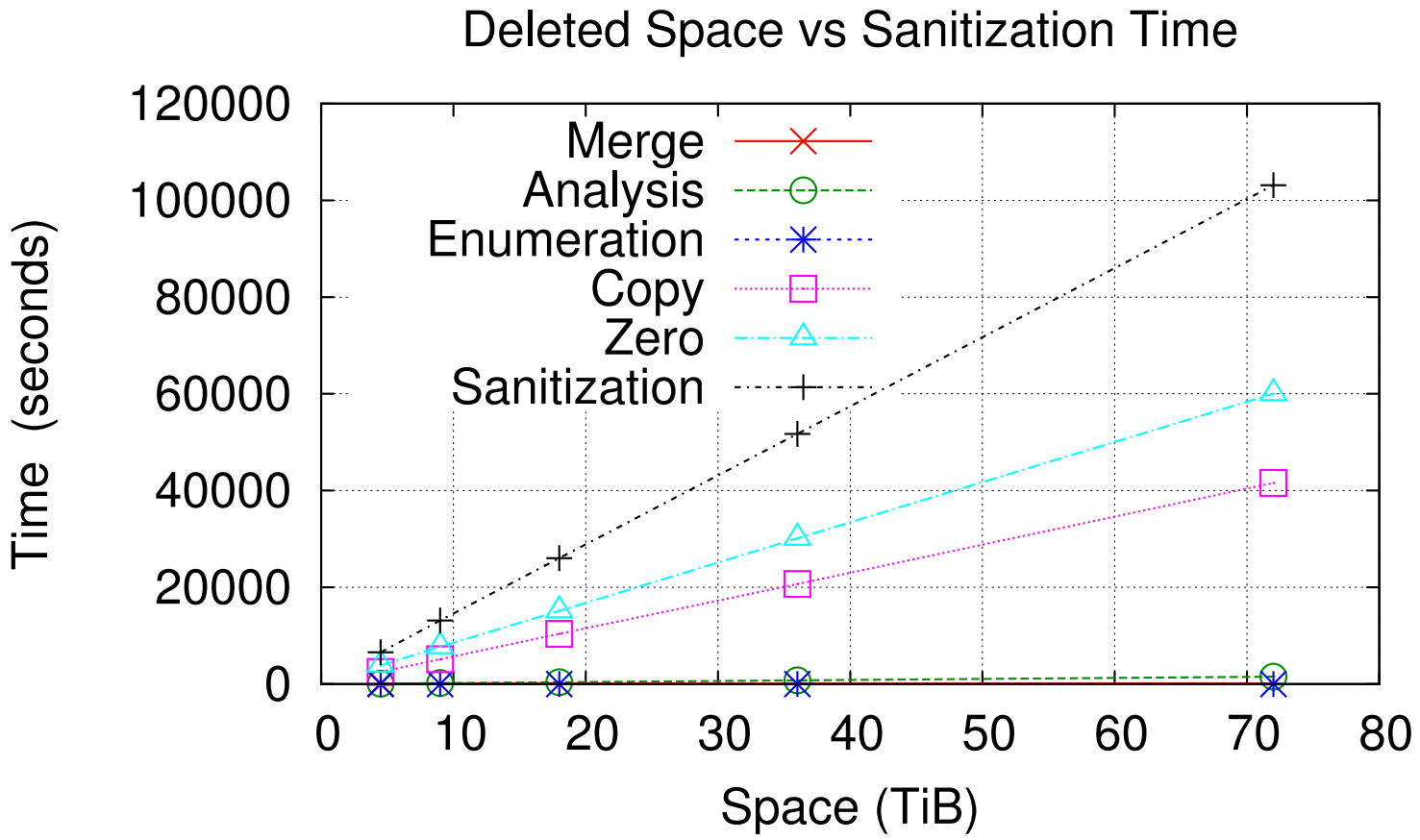
Experimental Setup and Synthetic Tool

- 16-Core Intel Xeon, 2.53 GHz, 8 MiB cache
- 6 RAID6 groups with 2TiB drives
  - 129.4 TiB of usable capacity
- 72 GiB of RAM
- Synthetic tool that mimics backup workload by leveraging our prior knowledge of such workloads

EMC²

© Copyright 2013 EMC Corporation. All rights reserved.

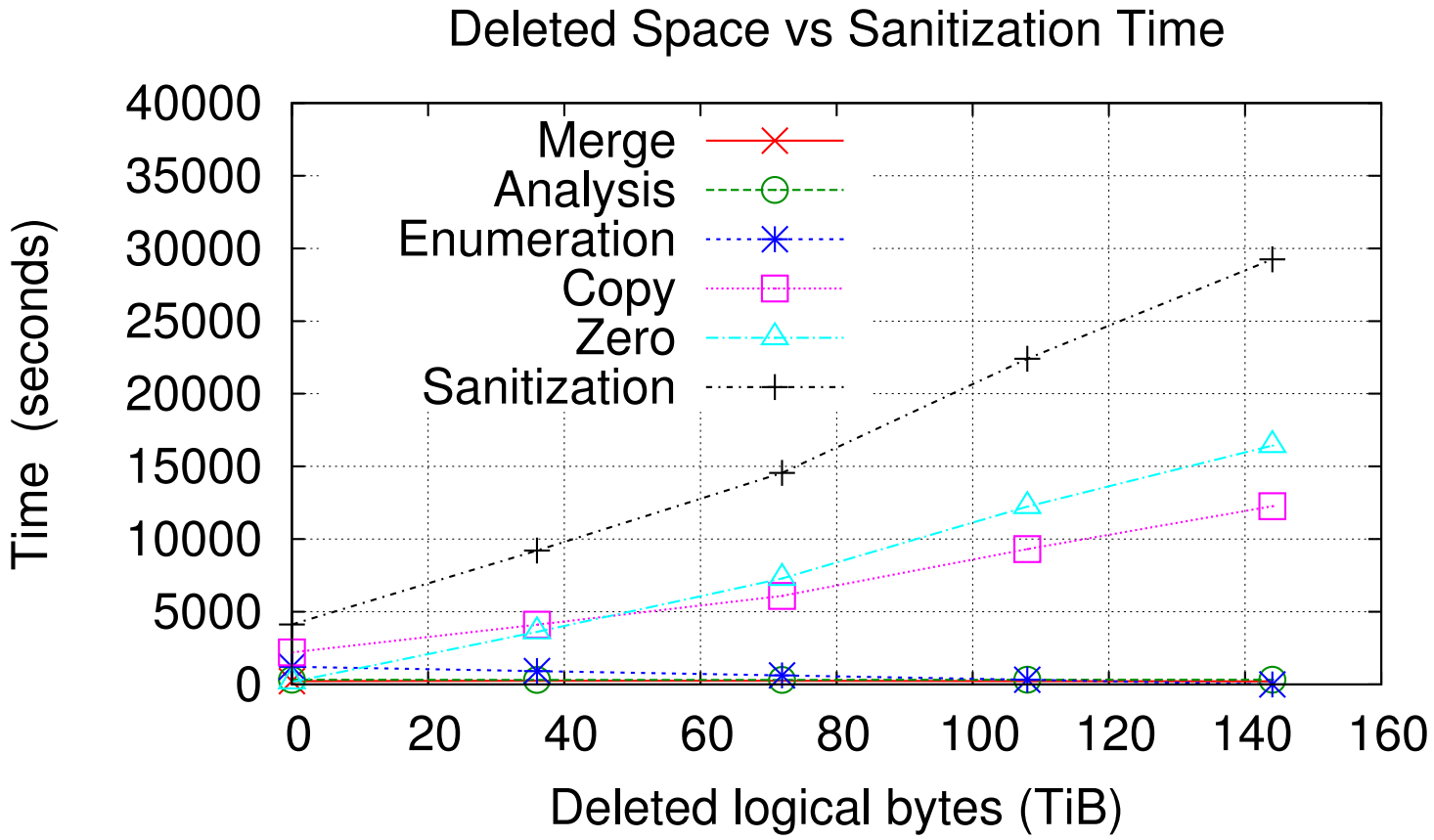
No Deduplication



731 MiB/second on average

© Copyright 2013 EMC Corporation. All rights reserved.

With Deduplication



5.06 GiB/second on average

© Copyright 2013 EMC Corporation. All rights reserved.

No Deduplication vs With Deduplication

- Deduplication factor 7.38X
- Throughput (No dedup): 731 MiB/second
- Throughput (with dedup): 5.06 GiB/second
- Speedup Factor: 7.1X

© Copyright 2013 EMC Corporation. All rights reserved.

Concurrent Data Ingest and Sanitization - with Deduplication

- Data ingest runs at 70% of its peak throughput
- Sanitization runs at least at 59% of its peak throughput
- Data Ingest: CPU intensive
- Sanitization: IO intensive

© Copyright 2013 EMC Corporation. All rights reserved.

Concurrent Data Ingest and Sanitization – No Deduplication

- Data ingest runs at 70% of its peak throughput
- Sanitization runs at least at 45% of its peak throughput
- Data Ingest: CPU and IO intensive
- Sanitization: IO intensive

© Copyright 2013 EMC Corporation. All rights reserved.

Conclusions

- Sanitization is a critical feature for security matters
- We have proposed a process to carry out bulk-sanitization of a storage system rather than devices
- We use perfect hashing to minimize memory and I/O requirements
- Nearly linear performance as storage grows
- Effective throughput multiplying with the deduplication factor.
- Sanitization without zero phase can be used as a process to reclaim dead space

EMC²

© Copyright 2013 EMC Corporation. All rights reserved.

Q&A

© Copyright 2013 EMC Corporation. All rights reserved.

EMC²