SNIA NSF | NETWORKING STORAGE

Not Again! Data Deduplication for Storage Systems

Live Webcast

November 10, 2020

10:00 am PT

Today's Presenters

Alex McDonald
Moderator
Independent Consultant
Vice Chair SNIA NSF

Abhishek Rajimwale
Distinguished Member of Technical Staff
Dell Technologies

John Kim
Chair, SNIA NSF
NVIDIA

SNIA-At-A-Glance

SNIA-at-a-Glance

185
industry leading organizations

2,000
active contributing members

50,000
IT end users & storage pros worldwide

Learn more: **snia.org/technical** **@SNIA**

Technologies We Cover

- Ethernet
- iSCSI
- NVMe-oF
- InfiniBand
- Fibre Channel, FCoE
- Hyperconverged (HCI)
- Storage protocols (block, file, object)
- Virtualized storage
- Software-defined storage

SNIA NSF | NETWORKING STORAGE

SNIA Legal Notice

- The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced in their entirety without modification
  - The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

  NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

Agenda

- Deduplication Basics
- Concepts and Considerations
- Deduplication in Action

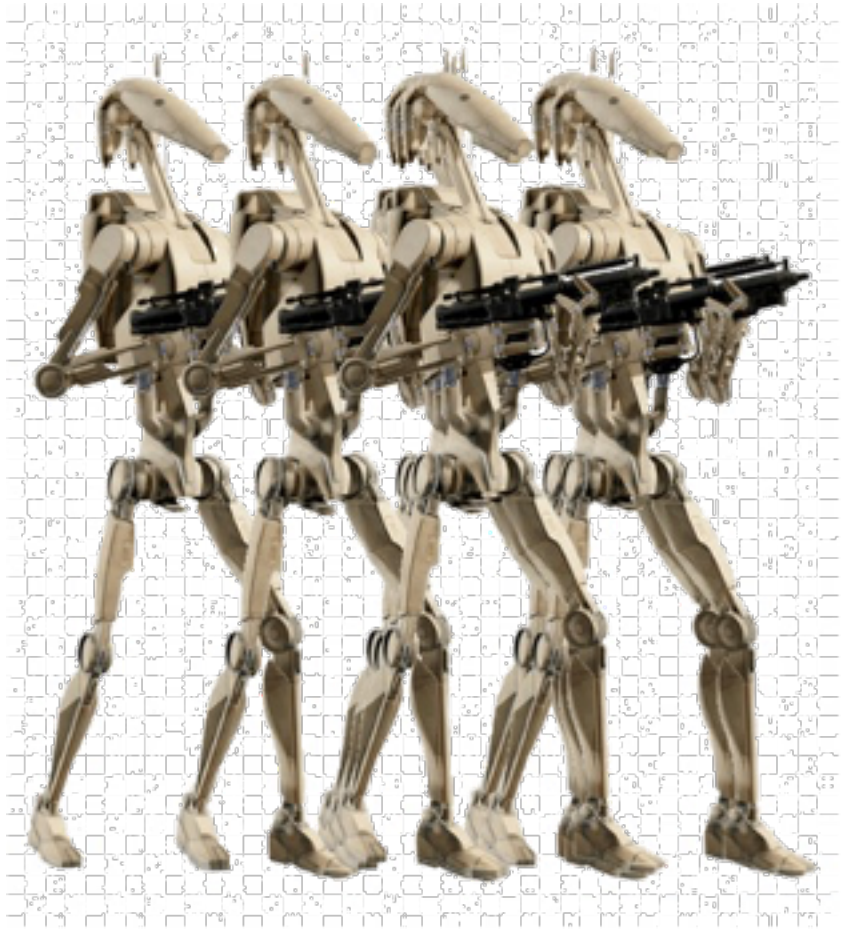
Deduplication Basics

Protection Storage Use-Case

Why Data Is So Often Duplicated

- Email attachments
- Users' local copies of data
- Test/development projects
- Backups and mirroring

- Result:
  - Hundreds of exact duplicates of files
  - Many files or objects that are nearly identical



Data Duplication Basic Idea

- Save (or transmit) first occurrence of each data chunk;
- Find reoccurring data chunks
- When duplicate chunk is encountered, save (or send) pointers or symbols referencing recurring data chunks, instead of sending storage or sending the data earlier.

Why Deduplicate Data?

Benefits

- Save space
- Save bandwidth
- Reduce wear on flash
- Save time on backups

Not Usually Benefits

- Improve performance on primary storage
- Improve security

Deduplicate or Compress?

Deduplication

- Usually works on files/objects (blocks for backup)
- Operate broadly/globally
- Scope is at on block level or above (coarse grained)

Compression

- Usually works on files, objects or storage devices (drives)
- Operates locally
- Scope is at block level or below (fine grained)

Who or What Does the Deduplication?

Where Deduplication Can Happen	How commonly is it done?	Granularity	Typical Scope	Comment
Application	Rare (except in email systems)	Files or objects	One application's data	Applications compression is more common
Local file system	Very Rare	Blocks	Local file system	File system compression is more common
Network device	Rare except for WAN optimization	Blocks	Transmitted data	Common in WAN optimization appliances
Storage device (drive)	Very Rare (so far)	Blocks	Data on that device	Could happen soon with computational storage
Storage system	Common	Blocks	Data on that system	Can combine w/compression
Backup system	Very common	Blocks and/or files	All backed-up data	Often dedupes globally then compresses at media layer

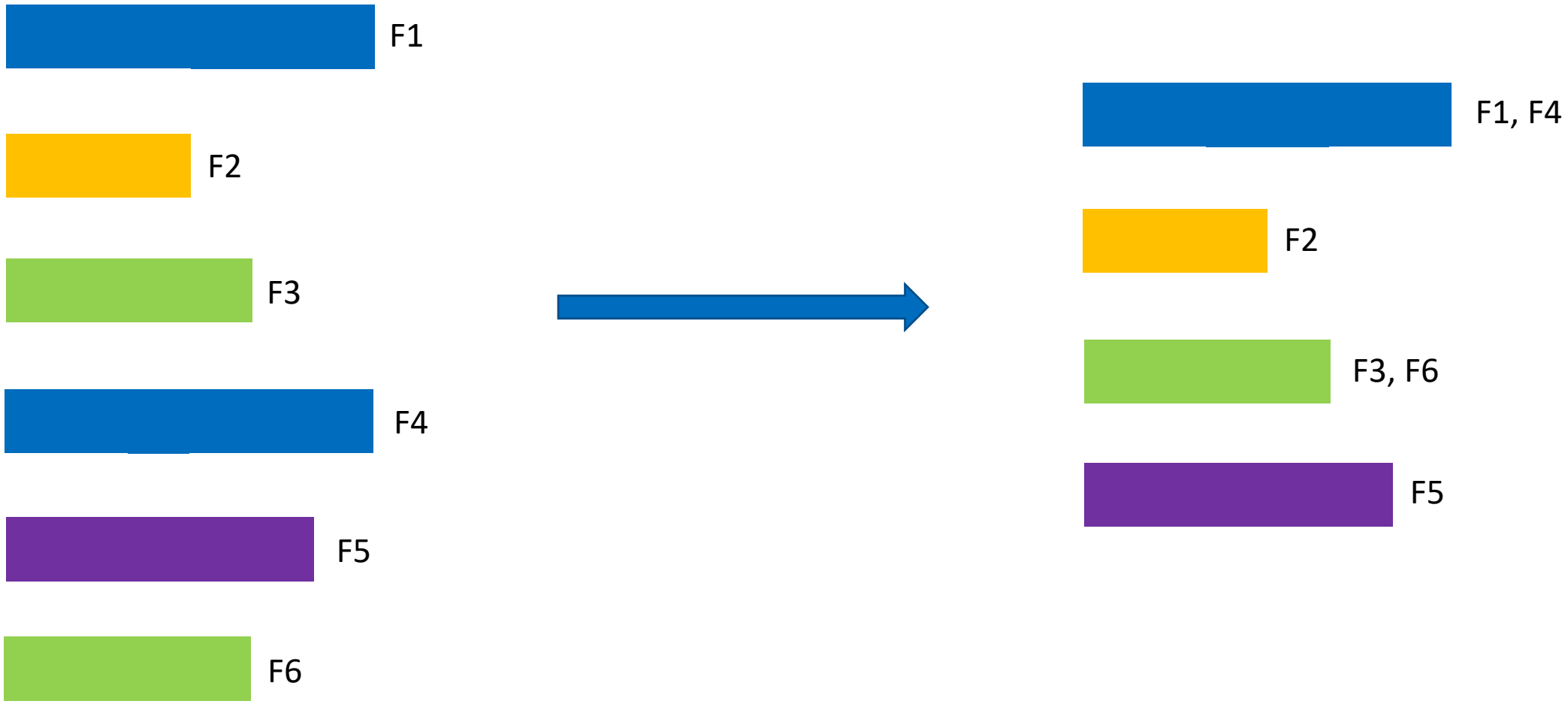
Using Deduplication with Compression and Encryption

- Many storage systems do both dedupe and compression
- Storage system can compress or dedupe first
  - Compressing first locally can hurt global dedupe
  - Or application compresses (e.g. audio/video/photo) then storage dedupes
- Strong encryption blocks deduplication
  - Encrypt after data reduction
  - If encrypt locally, must decrypt to perform global dedupe for backup

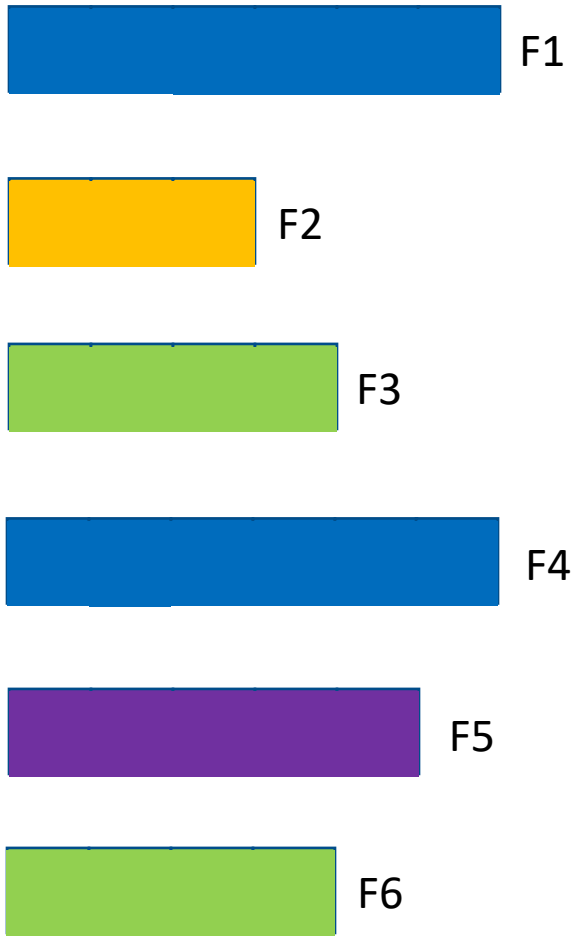
Data Deduplication

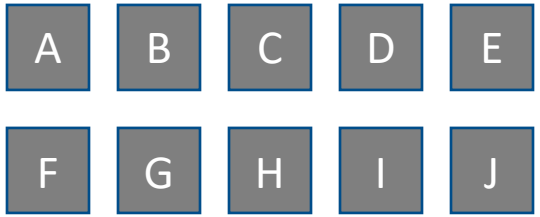
Concepts and Considerations

File/Object-based Dedupe

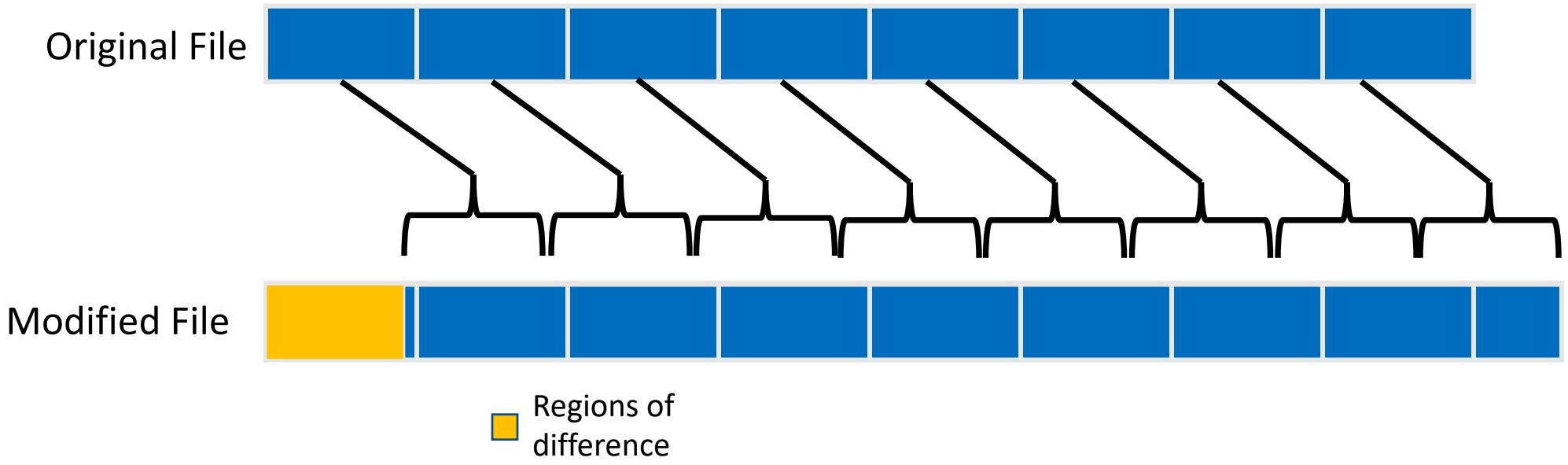


Block-based Dedupe

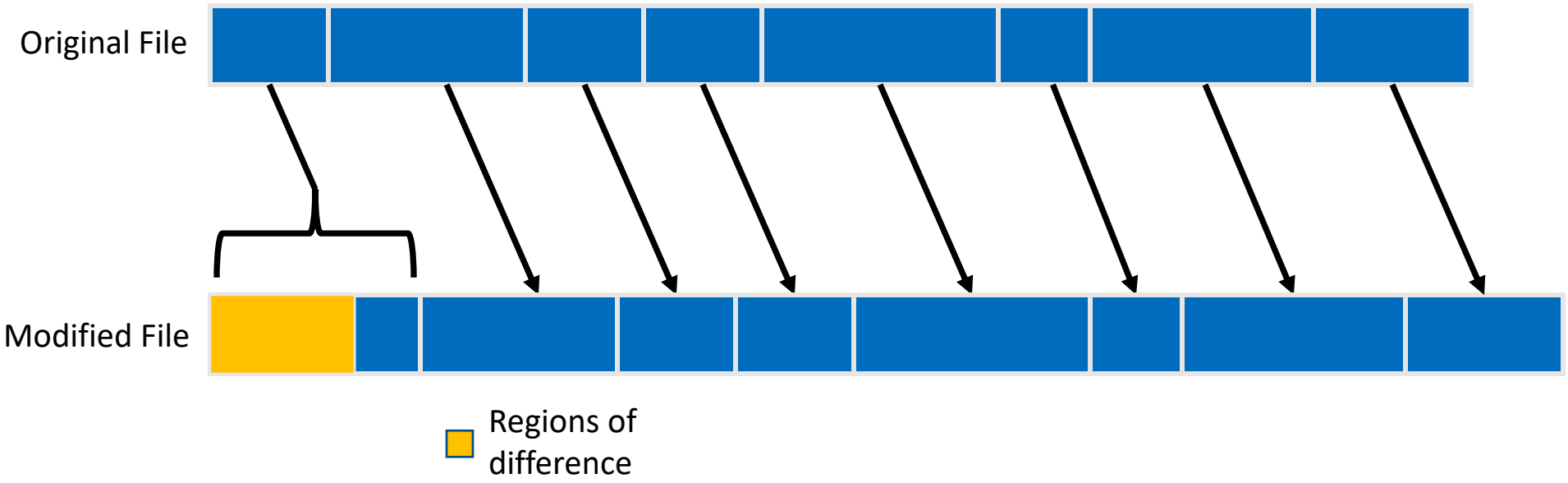




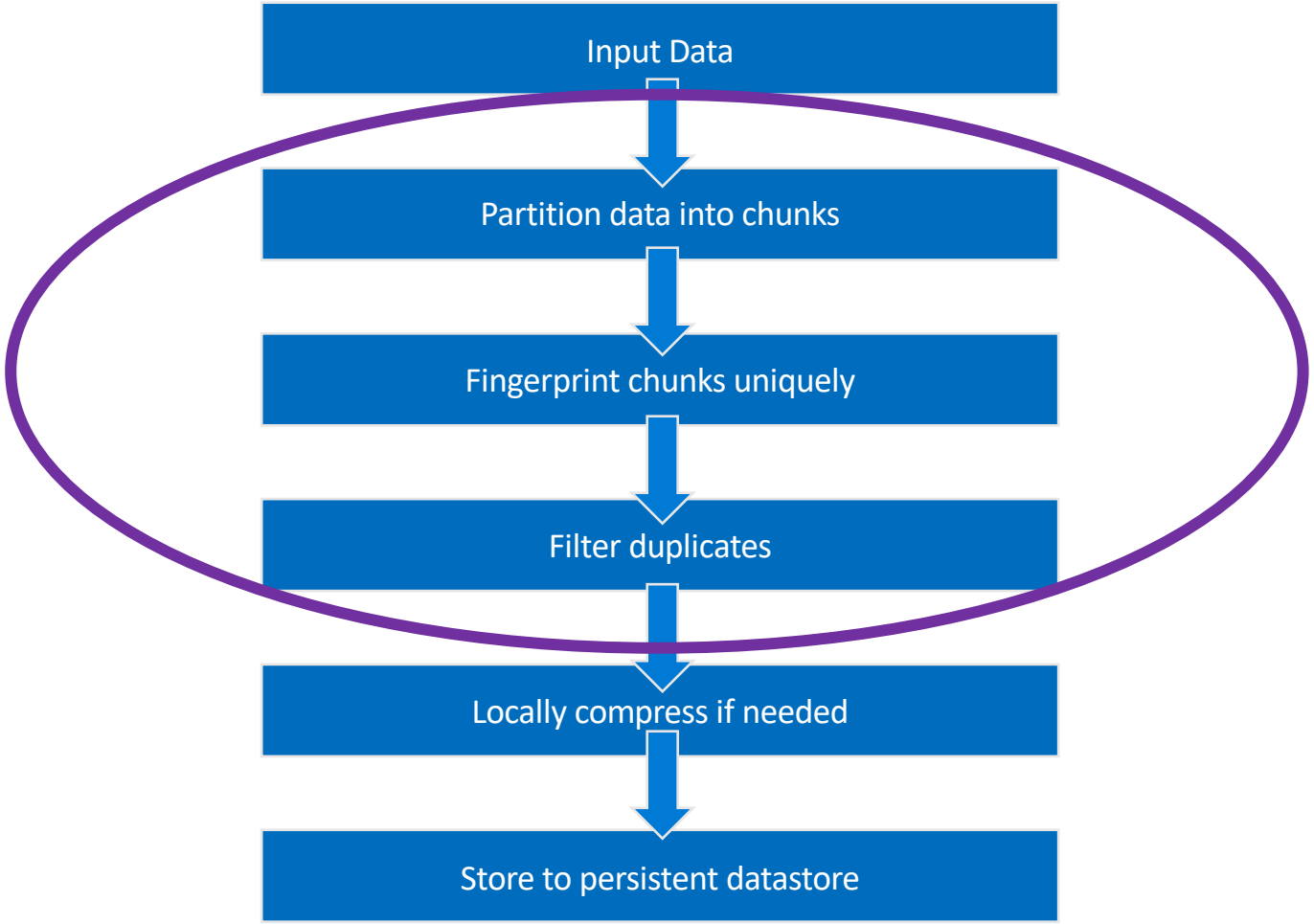
Fixed-size Blocks



Variable-size Blocks

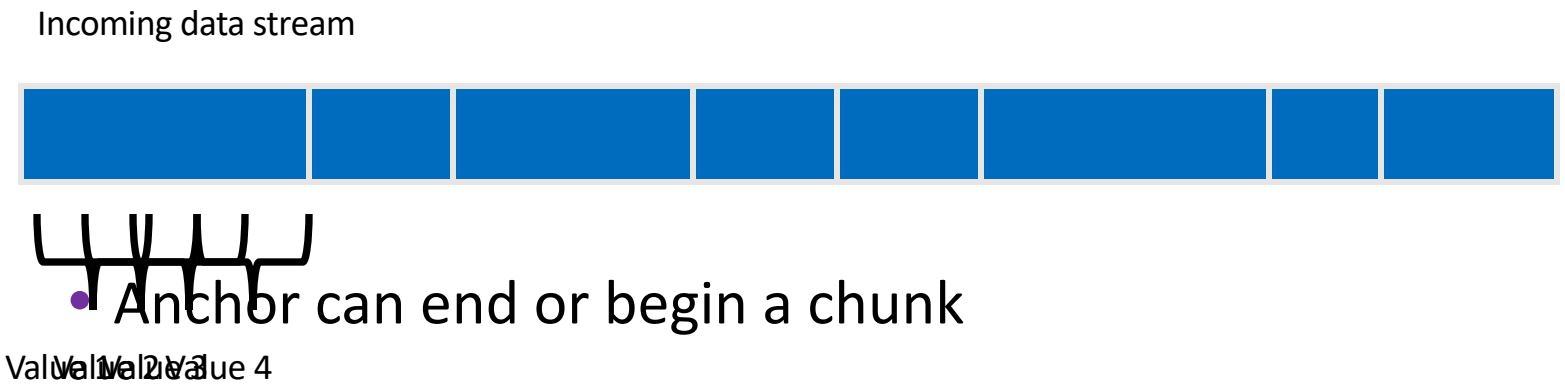


Typical Deduplication Storage Pipeline



Partition Data into Chunks

- Calculate a feature at rolling positions
- Select an anchor point with some property



- Anchor can end or begin a chunk
- Has to be fast to update Rabin fingerprint and others
- Minimum and maximum chunk size thresholds

Fingerprint Chunks Uniquely

- Cryptographically secure fingerprint
- Collisions are extremely unlikely
- MD5Sum, SHA-1, SHA-256
- Example – 160 bit fingerprint
  - Chance of 1 or more collisions
  - C = number of chunks
  - b = number of bits
  - 530 EB, 8KB chunks, 160 bit fp = . 0000000000019%

$$1 - e^{\frac{(-C^2+C)}{2*2^b}}$$

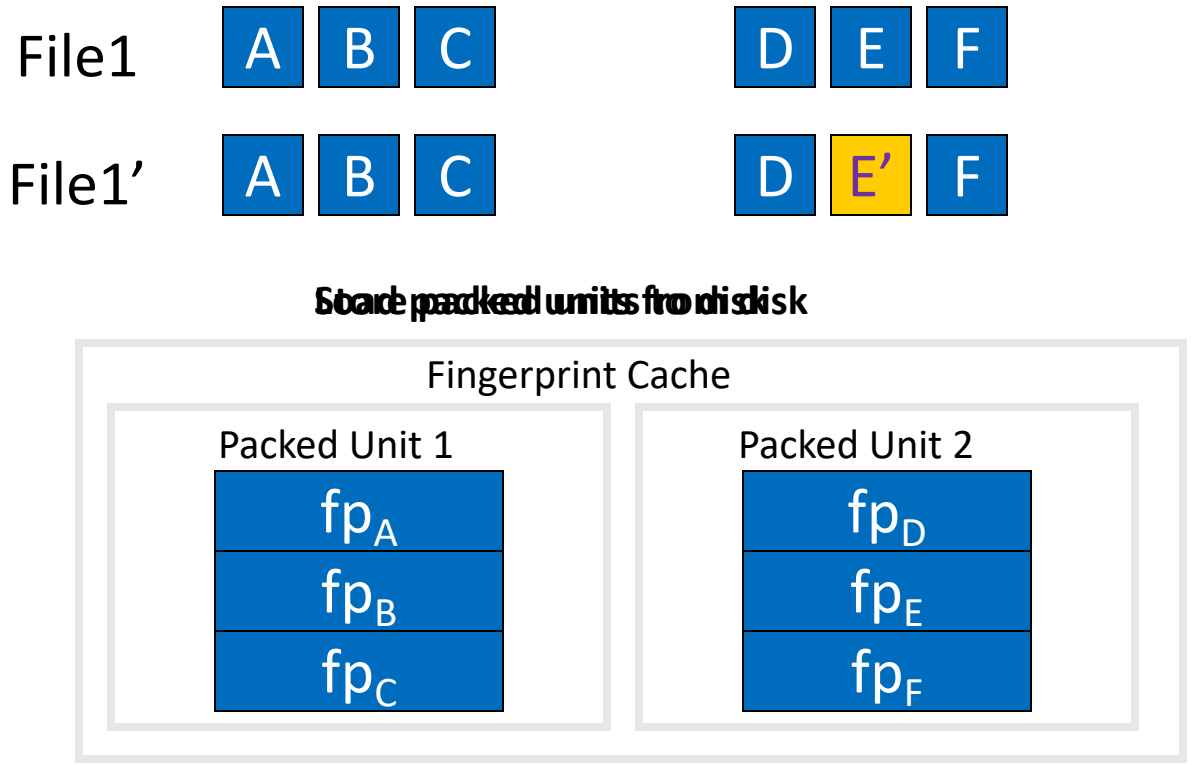
- Many more-likely errors to worry about instead:
  - RAID6 failures within 5 years = .001%
  - Struct by lightening = .0001%

Filtering Duplicates at High Speeds

- Disk Bottleneck Problem
  - Example Requirement: Ingest data at 1.2 GBps
  - Constraint: Cannot hold Index (fingerprint to location mapping) in memory
  - 8 KB chunk size requires 150 K index lookups per second
  - Index lookups are random
    - 7200 RPM SATA drive: <120 seeks/second
    - 120/second @ 8 KB segment: 0.96 MBps disk
  - Would need 1200 disks !!!
- Design Considerations:
  - Use locality of data on disk
    - Prefetching chunks fingerprints in memory cache for dedupe
  - Use smarts to avoid disk
    - Bloom filter in memory determines if a segment is unique

Data Locality

- Store consecutive chunks in “packed units” that will be loaded as one



Inline vs. Post-process

- Inline: Dedupe data as it is ingested, before it is stored persistently
  - More efficient in resource consumption – no landing zone
  - Data doesn't need to be read back and processed
  - Provides faster replicas for recovery and DR
  - Easier to size
- Post-process: Dedupe later by reading stored data back
  - May provide better dedupe as data can be analyzed lazily
  - Ingest speeds are not affected by dedupe processing delays
  - May be better if data needs to be read-back immediately

Global vs. Local Deduplication

- Global dedupe means unique data stored only once across multiple nodes or partitions
  - Create global deduplication domain by filtering across multiple storage domains
- Pros
  - More storage efficiency as chunks are stored only in once place
- Cons
  - Increased complexity to maintain a distributed Index across nodes
  - Possibility of performance bottlenecks and higher latencies
- Storage efficiency tradeoff based on size of a single dedupe domain
  - A local dedupe domain with 20x space savings Vs global dedupe with 10x !!

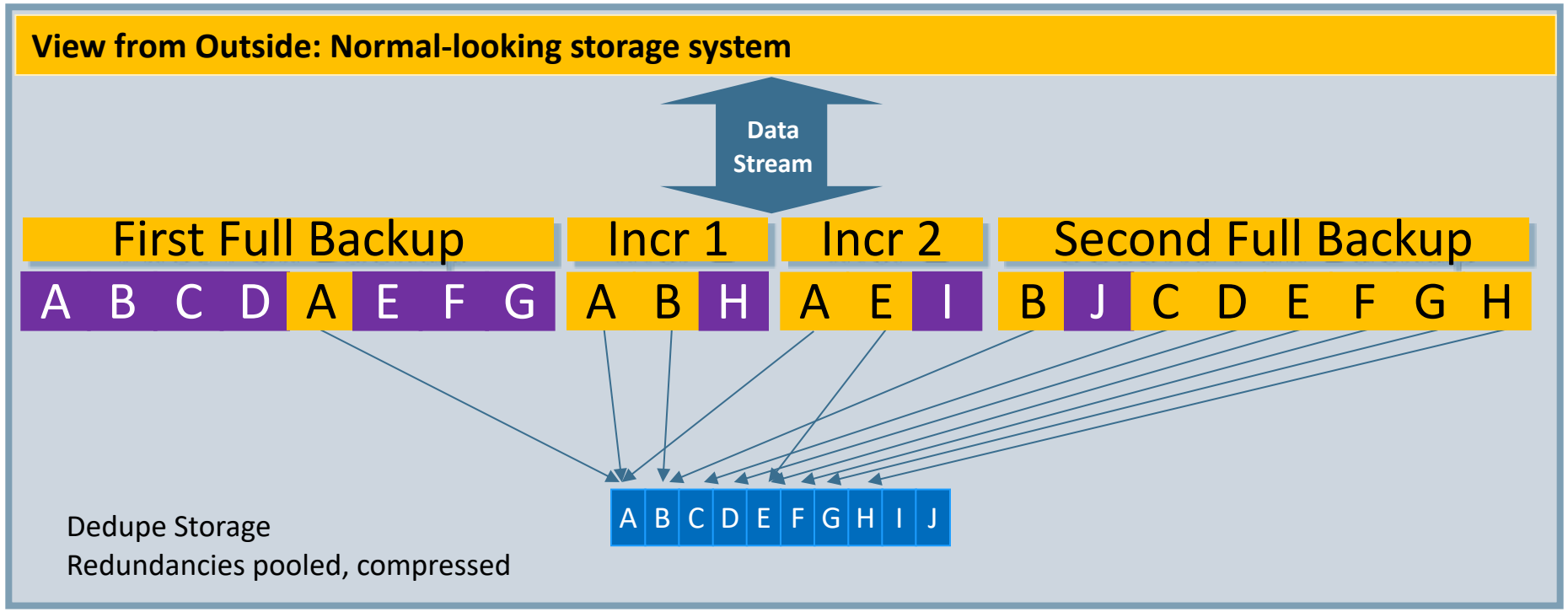
Role of Flash in Deduplication

- Flash as a cache
  - Can be used as a log when post-processing
  - Moving the index into flash accelerates lookups
  - Other metadata can move into flash
  - A large cache can absorb overwrites and accelerate reads
  - Useful in capacity optimized (protection/backup) storage
- Deduplicated flash storage:
  - Flash is expensive, so deduplication (and compression) lower the price
  - Flash wears out, so decreasing the writes extends the lifespan
  - Makes more sense for performance optimized (primary) storage

Deduplication in Action

Protection Storage Use-Case

Deduplication with Backups: Under the Hood



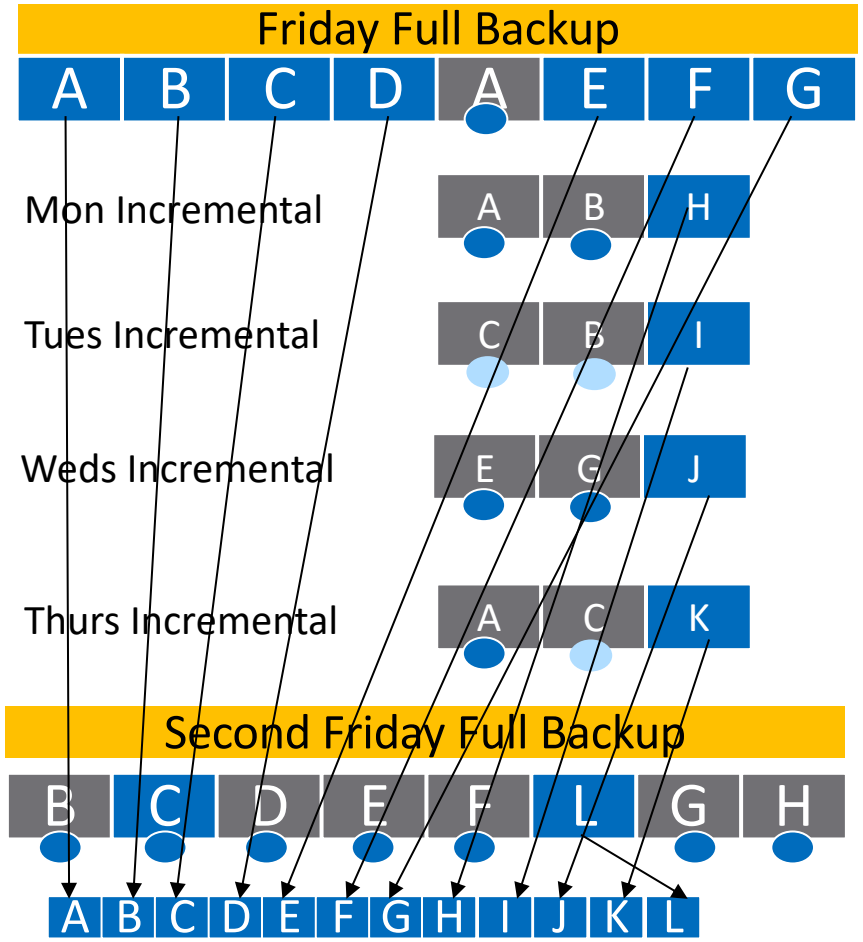
= Unique data chunks

= Redundant data chunks

= Compressed unique chunks

Dedupe Savings in Backup

- Store More Backups in a Smaller Footprint



Backup Data	Logical	Estimated Reduction	Physical
FRIDAY FULL	1 TB	2–4x	250 GB
Monday Incremental	100 GB	7–10x	10 GB
Tuesday Incremental	100 GB	7–10x	10 GB
Wednesday Incremental	100 GB	7–10x	10 GB
Thursday Incremental	100 GB	7–10x	10 GB
Second FRIDAY FULL	1 TB	50–60x	18 GB
TOTAL	2.4 TB	7.8x	308 GB

More Backups for Longer, with Less

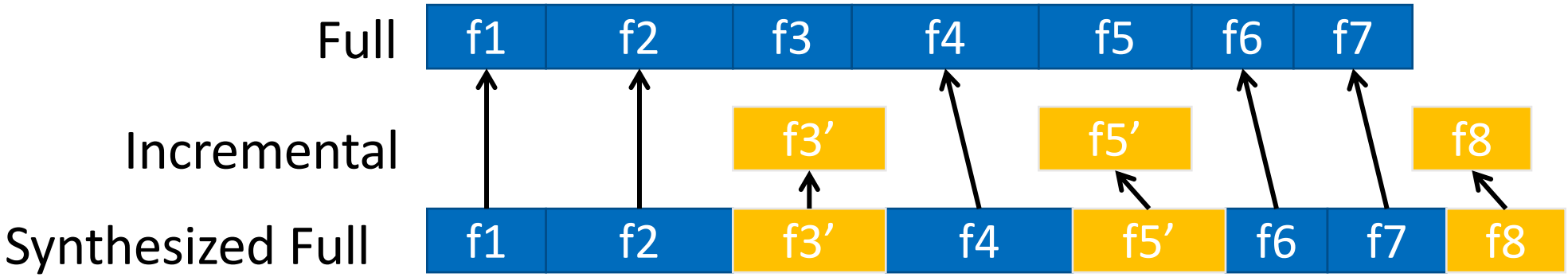
- Week 1
- Week 2
- Week 3
- Month 1
- Month 2
- Month 3
- Month 4

Backup Data	Cumulative Logical	Estimated Reduction	Physical
First Full	1 TB	4x	250 GB
April 7	2.4 TB	8x	308 GB
April 14	3.8 TB	10x	366 GB
April 21	5.2 TB	12x	424 GB
April 28	6.6 TB	14x	482 GB
May 31	12.2 TB	17x	714 GB
June 30	17.8 TB	19x	946 GB
July 31	23.4 TB	20x	1,178 GB
TOTAL	23.4 TB	20x	1,178 GB

Target-side and Source-side Dedupe

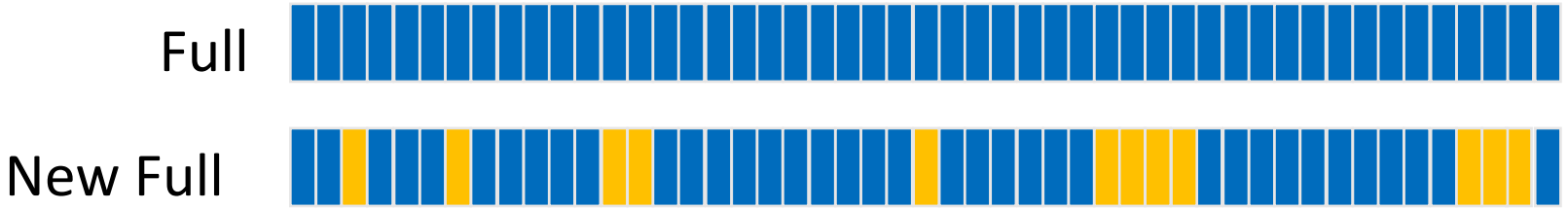
- Target-side dedupe
  - Storage system that dedupes data as it is ingested
  - Often useful as a drop-in target backup appliance
- Source-side dedupe
  - Dedupe at the source of the data
  - Eliminate redundant transfer of data to target (Backups)
- Example: Backup infrastructure with source-side dedupe support
  - Client or source-side support required
    - App-aware: Databases, VMs etc. can track changes
    - Agent-based: Agent talks to backup target to avoid sending data already sent
  - Advantage: Can result in smaller backup windows
  - Disadvantage: May not be ideal for high change rates
- Often a combination of both works best

Synthesized Backups



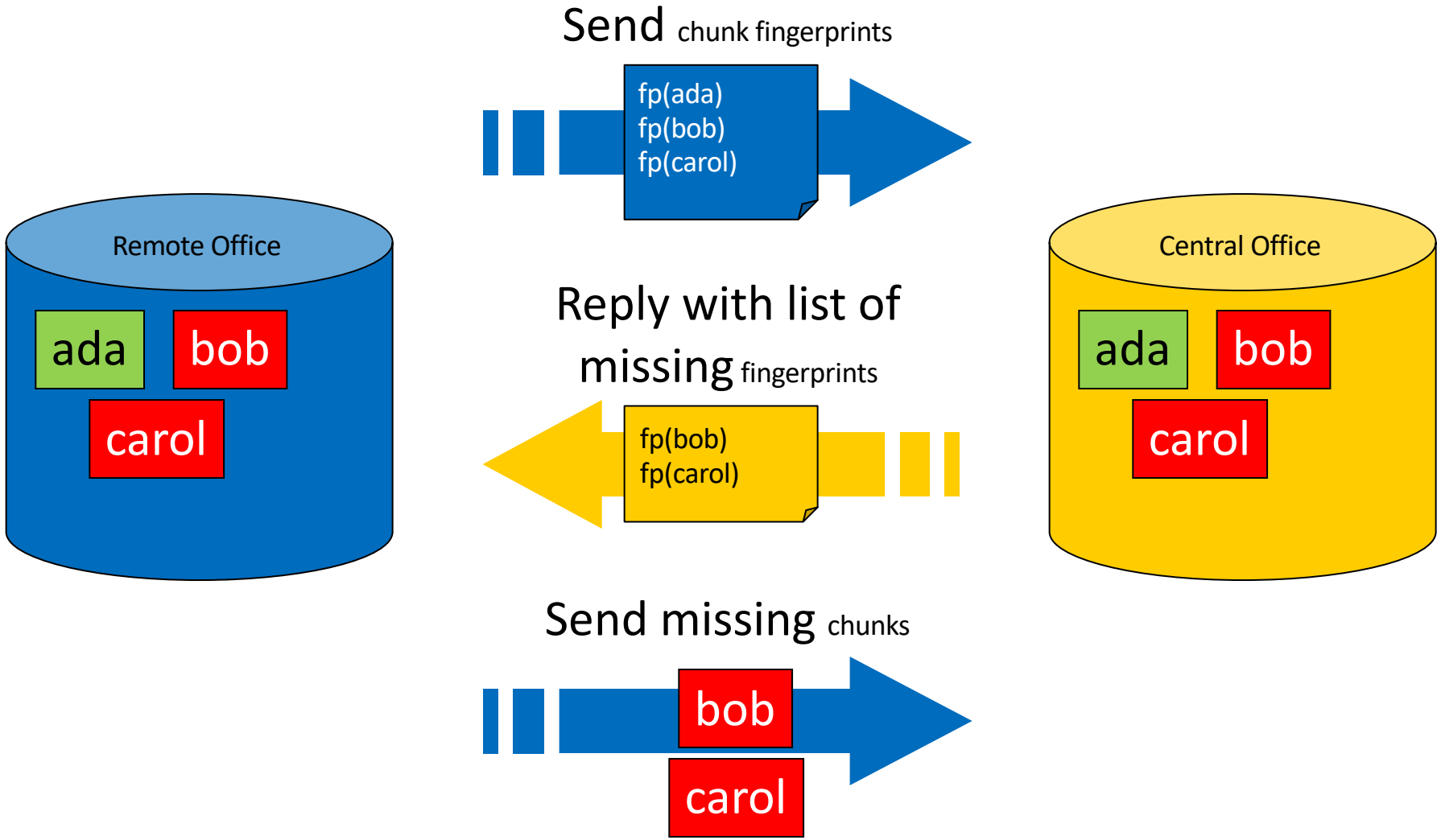
- Synthesize a region of data from a previous backup image to the new backup image
- Dedupe storage can optimize by copying chunk references instead of data
- Useful for file system backups

Change Block Tracking (CBT)



- The previous backup image is cloned
- New full image created by updating the clone by overwriting changed blocks
- Oracle Incremental Backups and VM backups use CBT

Replication with Deduplication



Summary

- “New York, New York”
  - So good, they named it twice
  - “New York” x2
- Deduplication is so common, it’s hard to find storage systems without it as a feature
- Works with files, block and object storage
- The big win: significantly reduces costs by many factors
  - “Space amplification”

Related Webcasts in this Series

- Everything You Wanted to Know About Storage But Were Too Proud to Ask: Data Reduction
  - https://www.snia.org/educational-library/everything-you-wanted-knowbut-were-too-proud-ask-data-reduction-2020
- Compression: Putting the Squeeze on Storage
  - https://www.snia.org/educational-library/compression-putting-squeeze-storage-2020

After this Webcast

- Please rate this webcast and provide us with your feedback
- This webcast and a copy of the slides will be available at the SNIA Educational Library https://www.snia.org/educational-library
- A Q&A from this webcast, including answers to questions we couldn't get to today, will be posted on our blog at https://sniansfblog.org/
- Follow us on Twitter @SNIANSF

Thank You

