

Chapter 15: Delta Sync

To identify which blocks have been modified and need to be delta-synced in a block storage environment, the system can use several methods:

1. Change Block Tracking (CBT)

- **What It Is:** CBT is a feature in many storage systems that tracks which blocks have changed since the last snapshot or sync.
- **How It Works:** Whenever a block is written or modified, the system flags it as "changed." When it's time to perform delta sync, the system only reads and sends the blocks marked as changed.
- **Example:** VMware's CBT (Change Block Tracking) or Amazon EBS snapshots with CBT.
- **Pros:** Efficient tracking without needing to scan the entire disk. Many enterprise storage solutions provide this feature.

2. Checksum Comparison

- **What It Is:** Checksum comparison involves calculating a checksum (hash) for each block and comparing it to the previous checksum for that block.
- **How It Works:** When a block is modified, its checksum changes. During a delta sync, the system calculates checksums for each block and identifies blocks with mismatched checksums as modified, syncing only those.
- **Pros:** Reliable for detecting changes without complex metadata tracking.

3. Bitmap Tracking

- **What It Is:** Bitmap tracking uses a bitmap (an array of bits) where each bit represents a block on the storage device.
- **How It Works:** When a block is modified, its corresponding bit is set to 1. During sync, the system checks the bitmap to identify modified blocks and resets bits after syncing.
- **Pros:** Lightweight and efficient. Common in backup software.

4. File System Journaling (if applicable)

- **What It Is:** Some file systems, such as NTFS or ext4, use journaling, which logs file changes.
- **How It Works:** By tracking the journal, the system can identify modified blocks. Journals record the addresses of blocks that were written or modified, so only those blocks are included in the delta sync.
- **Pros:** Useful for systems with file system-level access, though it adds some complexity.

5. Snapshot Comparison

- **What It Is:** Snapshot comparison involves taking periodic snapshots of the block storage.
 - **How It Works:** The system compares the current state with the last snapshot. Most systems supporting snapshots already store metadata about changed blocks, so the system can use this metadata to sync only the changed blocks.
 - **Pros:** Can leverage built-in snapshot capabilities in many storage platforms, though it may require more storage space.
-

Example Workflow for Delta Sync

1. **Modify Block Detection:** Using any of the above methods (e.g., CBT, checksums), the system flags modified blocks.
2. **Delta Sync Process:** The flagged blocks are then read, compressed (if needed), and sent to the destination.
3. **Post-Sync Update:** The system resets flags or updates checksums to prepare for the next delta sync cycle.

Using these approaches, a block storage system can efficiently detect modified blocks and apply delta sync to reduce network traffic.