



Overlayfs snapshots

Amir Goldstein

amir73il@gmail.com

github.com/amir73il/overlayfs/wiki

March 23, 2017



Overlays snapshots

- Introduction
 - Motivation: Next3 shortcomings
 - Comparison: btrfs snapshots, LVM thin snapshots
- Overlays overview
- Snapshots overview
- Demo
- Questions

Overlayfs snapshots - Motivation

- Next3 snapshots ship with CTERA products (since 2010)
 - Focus on reliability and performance of main fs
 - Multiple local snapshots for backup and previous versions
 - Snapshots use main fs free disk space
 - Ported to ext4 but upstream effort discontinued
- Overlayfs snapshots addresses Next3 shortcomings
 - Scale to larger storage
 - Directory sub-tree level snapshot
 - Maintainability (aiming for upstream)

Overlays snapshots - Comparison

	Btrfs	LVM thin	Overlays
Snapshot target	volume	volume	sub-tree
CoW unit	block	chunk	file
Writable and nested	V	V	X
Use snapshot as root (revert)	V	V	X
Snapshot any local fs	X	V	V
Share inode and page cache	X	X	V

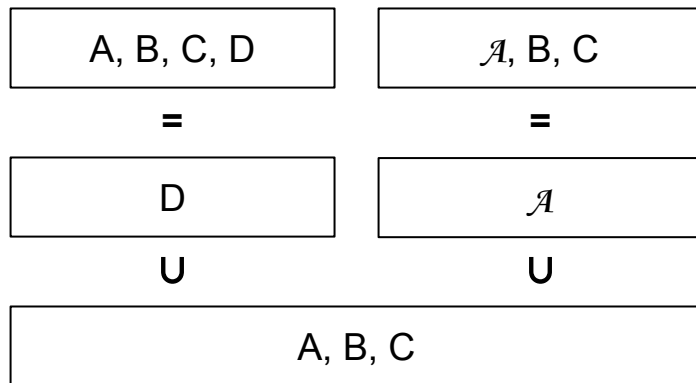
Overlays snapshots

- Introduction
- Overlays overview
 - Overlay use cases
 - Basics: copy up, whiteout, merge dir, opaque dir
 - New features: clone up, redirect dir, concurrent copy up, consistent fd
 - Redirect dir: lookup by name, redirect dir, redirect by file handle
- Snapshots overview
- Demo
- Questions

Overlays - Use cases

Writable clones

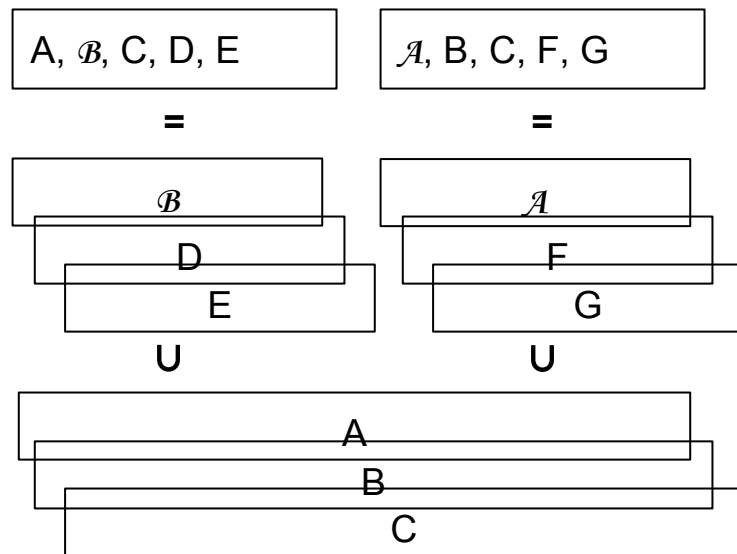
- Most files are in common lower layer
- Changes to upper layer are either persistent or disposable
- Efficient use of inode and page cache compared to block dev writable clones



Overlays - Use cases (cont.)

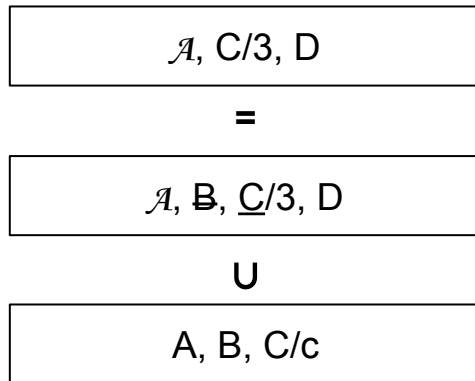
Stacking layers

- Repository of modular layers (Docker Hub)
- Image “source control”
- “Native” file level diff compared to block dev layers



Overlays - Basics

- Copy up
 - Large file copy duration on touch
 - Inconsistent fd before/after copy up
 - Global copy up lock
- Whiteout
 - Upper mark of a deleted file/dir in lower
- Merge dir
 - Union of upper files/whiteouts and lower files
 - EXDEV for rename of merge dir
- Opaque dir
 - New upper dir above deleted lower dir



Overlayfs - New features

- Clone up (v4.10)
 - Requires clone support in underlying fs (xfs, btrfs)
 - Requires lower and upper to be on same fs (like Docker)
 - Mitigates large copy up duration
- Redirect dir (v4.10)
 - Enables non-empty directory rename
- Concurrent copy up (v4.11-rc1)
 - Mitigates large file copy up DoS
- Consistent copy up fd (v4.12 ?)
 - Read-only lower fd ops intercepted after copy up

Overlays - Redirect dir

- Lookup by name
 - Lookup same name in all layers
 - Stop layer lookup for non-dir or opaque dir
- Redirect dir (v4.10)
 - Before merge dir rename, store src path in xattr
 - Lookup next layer by full path from layer root
- Redirect by file handle (RFC)
 - Lower and upper on same fs
 - On copy up, store lower dir file handle in xattr
 - Lookup next layer by decoding file handle

A/{a,1}, D/E/{b,2}, F/{c,3}

=

A/1, B, C, D/E_B/2, F_{&C}/3

U

A/a, B/b, C/c

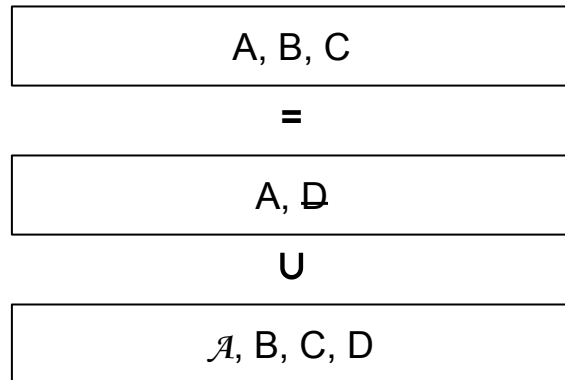
Overlays snapshots

- Introduction
- Overlays overview
- Snapshots overview
 - Snapshot use cases
 - Basics: snapshot overlay, snapshot mount, snapshot rules
 - Taking a snapshot
 - TODO
- Demo
- Questions

Overlays snapshots - Use cases

Short lived snapshot for backup

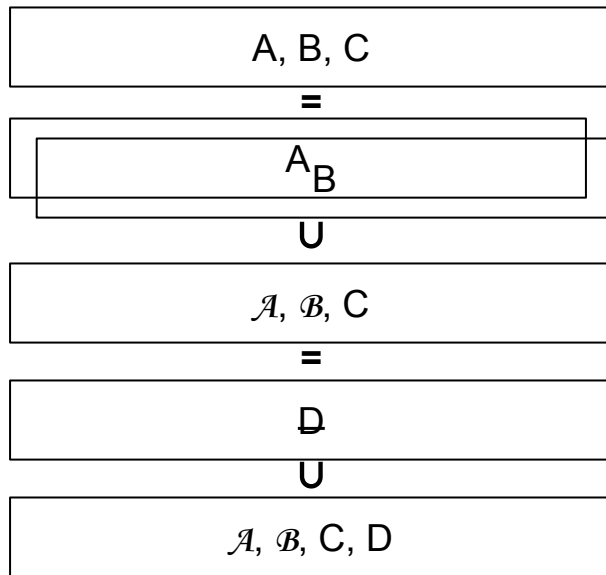
- File level CoW
- Changes to lower layer trigger copy up
- The content of the overlay snapshot remains stable
- Access to snapshot can tolerate performance hit



Overlays snapshots - Use cases (cont.)

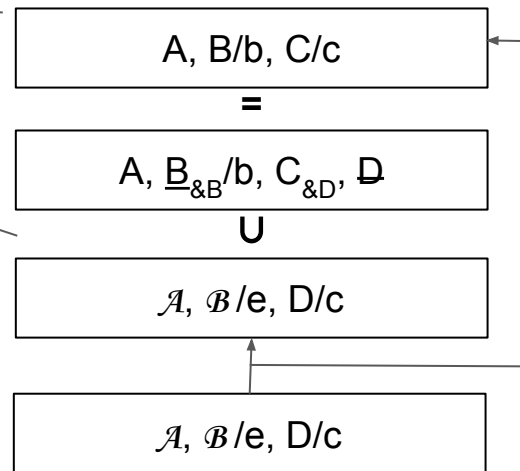
Previous versions

- Upper CoW layers are rotated **without** unmounting lower layer
- Composing previous upper layers produces previous snapshots
- “Native” file level diff between snapshots for incremental backup



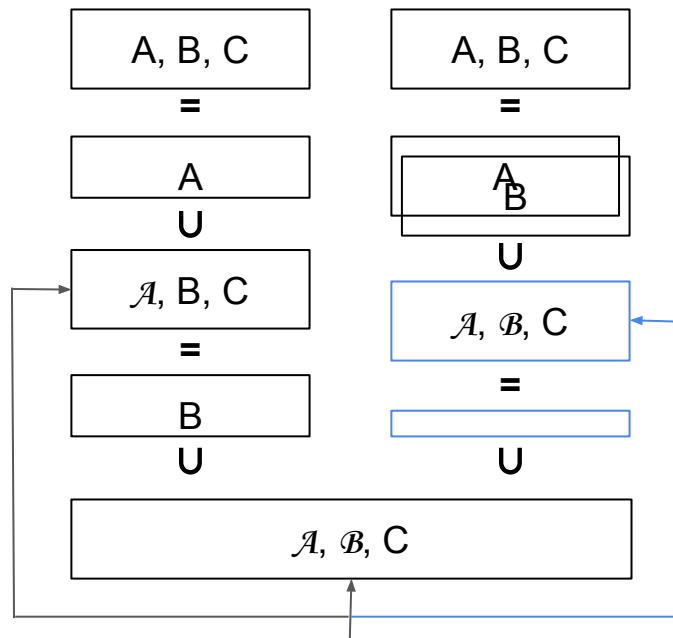
Overlays snapshots - Basics

- Snapshot overlay (overlay fs)
 - Enhanced overlays mount
 - Lower and upper on same fs
- Snapshot mount (snapshot fs)
 - Stackable fs over lower layer
 - Uses a snapshot overlay mount
- Rules of snapshot mount
 - Copy up before modify/delete lower
 - Whiteout upper before create lower
 - Redirect upper dir on rename lower dir
 - Set upper dir opaque on delete lower dir



Overlays snapshots - Taking a snapshot

1. Unmount previous snapshot overlays
2. Create new empty layer N
3. Mount new snapshot N: lower + layer N
4. Tell snapshot mount about snapshot N
5. Freeze snapshot mount
6. Tell snapshot mount to use snapshot N
7. Thaw snapshot mount
8. Mount previous snapshots incl. layer N-1



Overlayfs snapshots - TODO

- Handle open files during snapshot take
- Stabilize redirect dir before snapshot take
- Take a new snapshot without unmounting previous snapshot overlays
- Consolidate snapshots (merge layers)
- Upstream!

Overlays snapshots

- Introduction
- Overlays overview
- Snapshots overview
- Demo
- Questions

Overlays snapshots

- Introduction
- Overlays overview
- Snapshots overview
- Demo
- Questions
 - When & where get I can it?
 - How can I contribute to the project?



Overlayfs snapshots

Amir Goldstein

amir73il@gmail.com

github.com/amir73il/overlayfs/wiki

March 23, 2017

