

Overlayfs and Containers

Is user #1 interested in the latest overlayfs features?

Overlayfs and containers developers should communicate more!





Containers wish list

- shiftfs/fsid mapping
- Userns root mount (coming up)
- "volatile" overlay (coming up)
- Containerized syncfs(2) (RFC)





Why use overlayfs?

- Because it's the default :-)
- Shared page cache
- Fast instance startup time
- Native diff => faster image build





Overlayfs historical issues

- Non-standard behavior:
 - inconsistent file data
 - inconsistent object identifier
 - broken hardlinks
 - failure to rename directory
- Userspace mitigations: yum-plugin-ovl, recursive mv





Overlayfs recent progress

- Rename directories (redirect_dir=on) (v4.10)
- Preserve hardlinks (index=on) (v4.13)
- NFS export (nfs_export=on) (v4.16)
- Unique object identifier (xino=auto) (v4.12..v4.17)
- Consistent file data (excluding mmap) (v4.19)
- Metadata only copy up (metacopy=on) (v4.19)





Why are the new features disabled?

- Backward compatibility with old kernels
- Enabling xino only impacts 32bit compat st_ino
- Depending on the use case:
 - Local instance
 - Migratable instance
 - Portable image





Local instance

- Stays on one container host and storage driver
- Can safely enable: redirect_dir, metacopy
- Enabling index, nfs_export may cause longer mount time and exposes mount leak bugs





Migratable instance

- Migrate instance between container hosts with same native storage driver (overlay2)
- Include trusted.overlay xattrs in image layers
- No need to convert opaque xattr and whiteout chardev to portable .wh. files
- Can safely enable: redirect_dir, metacopy
- The index and nfs_export features are NOT "natively migratable"





Portable image

- vnd.oci.image.layer.v1 conforms to decade old overlay/aufs whiteouts and opaque directories
- Enabling redirect_dir, metacopy requires more work
- vnd.oci.image.layer.v2 with redirect_dir, metacopy support?
- Downgrade to vnd.oci.image.layer.v1 format with overlayfs tools
 - IF is_metacopy_file_or_redirect_dir(upperdir/\$path):
 - rm -rf imagedir/\$path
 - cp -a mountdir/\$path imagedir/\$path

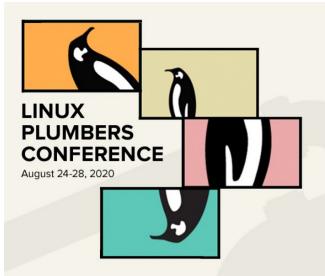




Recommendations

- Enable redirect_dir,metacopy for new instances
- Allow user to configure NFS exportable instances (index,nfs_export)
- Disable index during instance migration
- Downgrade redirect_dir,metacopy during image creation





Questions?

"I am glad that X was fixed!"

"I did not know that we had issue Y!"

"I want to follow up on Z..."

https://github.com/amir73il/overlayfs

