

Union mounts

<http://valerieaurora.org/union/>

Valerie Aurora

Red Hat, Inc.

<vaurora@redhat.com>

Overview

- The problem
- Existing solutions
- Union mounts
- Why not unionfs?
- Limitations of unionfs
- What you can do
- How to take over the world with union mounts and Puppet

What's the problem?

- Lots of nearly identical root file systems
- Virtual machines
- Any cluster
- LiveCD
- Could we share some of these?

Ways to share file systems

- Copy-on-write (COW) block devices
- NFS-exported root file system
- Other network/cluster file systems
- Snapshots

Why do we need union mounts?

- COW block device inefficient, divergent
- NFS server gets overloaded
- Cluster file system is needlessly heavyweight
- Snapshots are limited to a single system

Solution: file-level sharing

- Take one read-only file system (shared)
- Add one local writable file system
- Lookups “fall through” to lower read-only fs
- Writes go to topmost writable fs
- Writes to lower fs trigger copy-up to topmost fs

Sounds like... unionfs

- Unionfs and aufs panic the kernel
- Unionfs architecture is fundamentally broken
- Al Viro's canonical explanation here:
<http://lkml.indiana.edu/hypermail/linux/kernel/0802.0/0839.html>
- University research project
- Unmerged for 7 years and counting

Why union mounts is different

- Implemented in the VFS
- Designed with help of VFS maintainers
- Many fewer features

Limitations of union mounts

- `open(O_RDONLY)` gets you a different file from `open(O_RDWR)`
- `fchmod()/fchown()/futimensat()` on fd don't work
- Only one writable layer
- Requires on-disk format changes to topmost fs

What you can do

- Use `open(O_RDONLY)`
- Don't use `fchmod()/fchown()/futimensat()` on fd
- Don't access same file from multiple fds
- Let your engineers work on union mounts
- Review
- Test
- <http://valerieaurora.org/union/>

Puppet + Union mounts =
World Domination

Thank you

Jan Blunck

Miklos Szeredi

Felix Fietkau

Al Viro

Erez Zadok

Christoph Hellwig

J. R. Okajima

Ric Wheeler

and more

Union Mounts

<http://valerieaurora.org/union/>