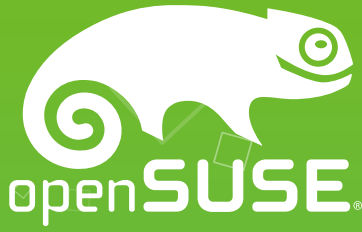


Converging image and package based OS updates

Ludwig Nussel

SUSE

lnussel@suse.de

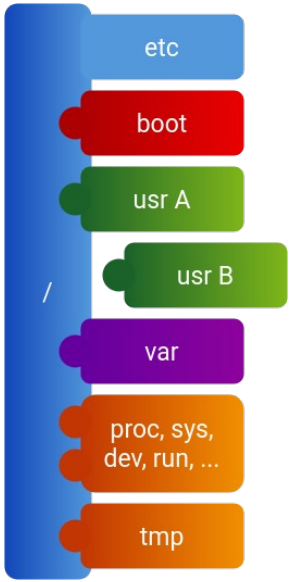
Image based vs Package based

- Packages
  - Individual components pre-built
  - Client decides what components are needed
  - Install may not require reboot
- Image
  - Full Linux system pre-built on server side
  - File system or tarball
  - Install requires reboot
  - Not extendable

Image based layout

- The OS is read-only in /usr
- /etc on / writable
- /boot is ESP
- Data in /var

Image based A/B updates



- Have several partitions for /usr
- + real read-only /usr
- + easy, partitions
- + rsync/casync for deltas
- + whole image signed
- - no deduplication
- - limited amount
- - preallocated space

BTRFS for image storage

- + data sharing
- + rsync/casync for deltas
- + flexible amount of versions
- - read-only state just a btrfs flag
- ? verification

How distros build images

- Images build with packages
- Packages can’t ship stuff outside /usr
- Scriptlets can’t read or modify /etc and /var
  - → Systemd presets, sysusers etc
- Need to track bootloader entries

Verification and Introspection

- Images consist of RPMs
  - Headers are signed
  - Headers have file checksums
  - List of %name-%version-%release.%arch
- The RPM DB is a binary blob though

Replace the RPM “DB”

- Just store RPM headers as files

```
/usr/lib/sysimage/packages
├── glibc-2.36-2.1.x86_64.rpm
└── bash-5.1.16-8.3.x86_64.rpm
```

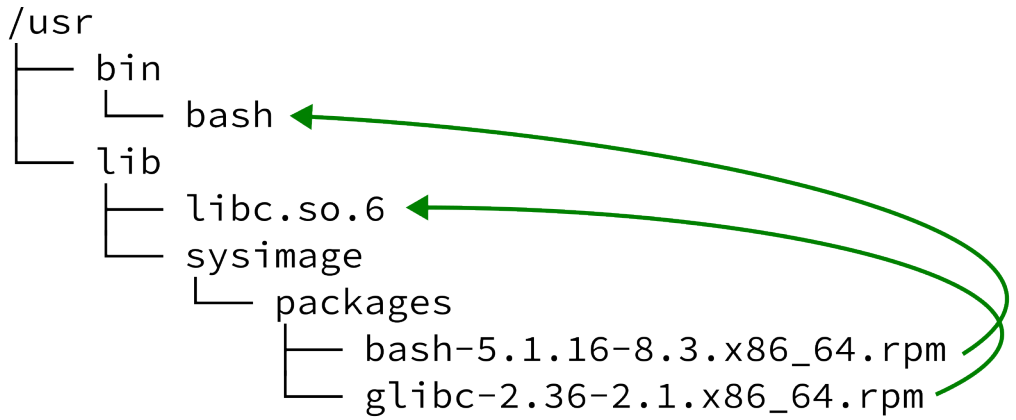
- Instrospect via `ls`
- `diff` shows updated packages
- `/usr` fully verifiable

What do we need the image for?

- `/usr` is RPM content anyway
- Just `rpm -i –noscripts`?
- Compressed payload bad for deltas

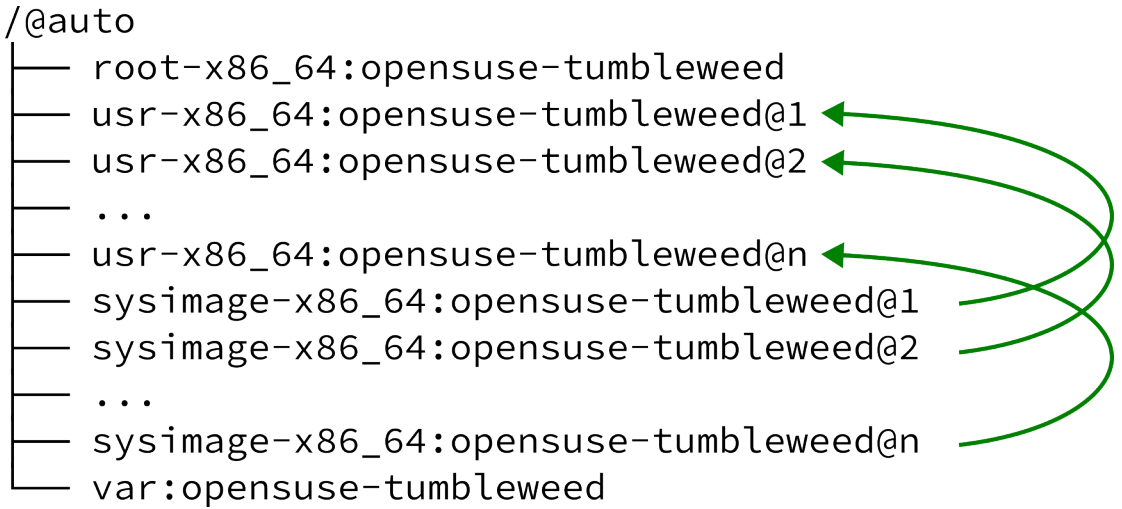
Reflinkable packages

- Uncompressed payload
- Align for ioctl_fideduperange(2)
- RPM “database” is the image



Subvolumes for reflinks

- /usr is just a view



Summary

- We can build image-like system leveraging btrfs
- Behavior is similar to image-based systems
- Software flexibility on client-side retained without hacks

Prototype/PoC

- https://github.com/lnussel/reflink/
- https://github.com/lnussel/rpm2extents/
- https://github.com/lnussel/busybox/tree/reflink
- Server side solving
  - An “image” is a directory of rpms
- Rsync for updates
- Busybox instead of rpm
- Uses systemd’s kernel-install to make bootable

TODO

- Fix packages!
  - No /etc and /var
  - No scriptlets
- Standardize btrfs volume names upstream
- Standardize reflink payload upstream
- Enhance systemd’s kernel-install
- Handle rollback of /
- Reconsider casync
- Adjust package management (rpm, zypp)

Questions?


openSUSE®