

OpenAFS Native Mountpoints on Linux

Andrew Deason

June 2019

OpenAFS Workshop 2019

Single mountpoints are pretty simple:

```
$ fs mkm mtpt1 root.cell
$ ls -ld mtpt1
drwxr-xr-x 7 root root 2048 May 28 17:25 mtpt1
$ cd mtpt1
$ cd ..
█
```

Background

What about multiple mountpoints?

```
$ fs mkm mtpt1 root,cell
$ mkdir dir
$ fs mkm dir/mtpt2 root,cell
$ cd mtpt1
$ cd ..
$ cd dir/mtpt2
$ cd ..
$
█
```

What is '..'?

Dentries and Inodes

- Inodes: files, dentries: links
- No directory hard-links
 - Why?
- Historically, AFS mountpoints *were* dir-links
 - or '..' moved on access

Dir-link workarounds

- What about `/proc/$pid/cwd`, `/proc/$pid/fd/$n`?

```
$ mkdir dir
$ cd dir
$ sleep 999 & pid=$!
[1] 13991
$ cd ..
$ rmdir dir
$ mkdir dir
$ touch dir/newfile
$ ls -ld /proc/$pid/cwd
lrwxrwxrwx 1 [+++] /proc/13991/cwd -> /home/adeason/dir (deleted)
$ ls /proc/$pid/cwd/
$ ls dir
newfile
$
[]
```

Dir-link workarounds

- `d_automount`, `follow_link`
- Commit: “Make dir dentry aliases act like symlinks”
 - <https://gerrit.openafs.org/7741>
- `find`
 - `./mtpt1 → ./mtpt1/subdir → ./mtpt1/subdir/mtpt2 → ./mtpt1/subdir/mtpt2/.. → ./`
- There must be a better way!

Real Mountpoints

- Linux VFS says: must use real mounts
- volumes/mtpts fit the model
- Why don't we already have it?
 - userspace: mount
 - kernel: GPLONLY `vfs_kern_mount()`

- Alternatives:
 - GPL wrapper
 - FUSE
 - kAFS
 - Userspace upcall

- Existing upcalls:
 - DNS (AFSDB)
 - OS X `mv` helper

Linux Native Mounts

```
$ cd /afs/vm.kupo.dson.org/deb-1
$ mount
sysfs on /sys type sysfs (rw,nosuid,nodev,noexec,relatime)
proc on /proc type proc (rw,nosuid,nodev,noexec,relatime)
[...]
```

AFS on /afs type afs (rw,relatime)

```
#vm.kupo.dson.org:root.cell on /afs/[...] type afs (rw,relatime)
#d1.root on /afs/vm.kupo.dson.org/deb-1 type afs (rw,relatime)
$
[]
```

Linux Native Mounts

```
$ df
Filesystem      1K-blocks    Used  Available Use% Mounted on
udev            505672         0    505672   0% /dev
tmpfs           102036    10688     91348  11% /run
[...]
```

Filesystem	1K-blocks	Used	Available	Use%	Mounted on
udev	505672	0	505672	0%	/dev
tmpfs	102036	10688	91348	11%	/run
[...]					
AFS	2147483647	0	2147483647	0%	/afs

```
$
█
```

Linux Native Mounts

```
$ mount -t afs '#root,cell' /mnt -o root_cell=1,root_vol=536870916
$ stat /mnt/ /afs/vm.kupo.dson.org/
  File: /mnt/
  Size: 2048          Blocks: 4          IO Block: 4096   directory
Device: 26h/38d Inode: 16842776   Links: 3
[...]
```

File:	/mnt/				
Size:	2048	Blocks:	4	IO Block:	4096
Device:	26h/38d	Inode:	16842776	Links:	3

```
[...]
```

File:	/afs/vm.kupo.dson.org/				
Size:	2048	Blocks:	4	IO Block:	4096
Device:	26h/38d	Inode:	16842776	Links:	3

```
[...]
```

```
$
[]
```

- MNT_EXPIRE in a loop
- 10k mounts, 100k, ...?
- Needs testing

- `/afs/:.mount/cell:vol:vnode:uniq`
 - Non-/afs mountpoints

- Separate devices
 - `df -h`

Top commit

<https://gerrit.openafs.org/13285>

All commits

<https://gerrit.openafs.org/#/q/topic:linux-native-mounts>

Slides

<http://dson.org/talks>

?