

Need some help on a pivot_root() syscall implementation

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/hackers/2006-04/msg00359.html>

- *From:* Adrian Steinmann <ast@xxxxxxxxx>
 - *Date:* Wed, 26 Apr 2006 18:40:46 +0200 (CEST)
-

I have been working on an implementation of pivot_root() system call which should be a clone of the same linux system call.

Description from my code (and linux/fs/namespace.c):

- * pivot_root Semantics:
- * Moves the root file system of the current process to the directory put_old,
- * makes new_root as the new root file system of the current process, and sets
- * root/cwd of all processes which had them on the current root to new_root.
- *
- * Restrictions:
- * The new_root must be a mountpoint and not the current root; put_old must be
- * under new_root and no other file system may be mounted on put_old.

I have gotten as far as swapping, say, a hierarchy "/" which "/".

The problem appears in the second part – having moved "/" to "/a/mnt" (which becomes "/mnt") when part of the fs does not work correctly.

I have had versions where 'df' looks ok after pivot_root but 'ls /mnt' showed nothing, i.e. the mountpoint /mnt with the old root wasn't really there (or not visible). Now, I think I have a better version, but it panics when I do 'ls /mnt' (see below)

```
# df
Filesystem 1K-blocks Used Avail Capacity Mounted on
/dev/md0 7054 5284 1770 75% /
devfs 1 1 0 100% /dev
/dev/ad0s1a 30856 14666 13722 52% /mnt
/dev/md1 7054 5282 1772 75% /a
devfs 1 1 0 100% /a/dev
# usage: pivot_root new_root put_old
# pivot_root /a /a/mnt
pivot_root: enter
pivot_root: ok, put_old has no mount points
pivot_root(/a, /a/mnt): locked Giant
pivot_root: start lookup new_root '/a'
pivot_root: start lookup put_old '/a/mnt'
pivot_root: rootvnode 0xc221cc30; put_vp 0xc2253208 new_vp 0xc222d618
pivot_root: new_mp removed and reinserted in head of mountlist
```

Need some help on a pivot_root() syscall implementation

```
pivot_root: new root adjusted
pivot_root: f_mntonname /a => /
pivot_root: old root mountpoint adjusted
pivot_root: f_mntonname / => /mnt
pivot_root: mountpoint /dev under old root adjusted
pivot_root: f_mntonname /dev => /mnt/dev
pivot_root: mountpoint /mnt under old root adjusted
pivot_root: f_mntonname /mnt => /mnt/mnt
pivot_root: mountpoint /a/dev under new root adjusted
pivot_root: f_mntonname /a/dev => /dev
pivot_root: mountlist adjustments done
pivot_root: put_vp done
pivot_root: new_vp done
pivot_root: leaving, error = 0
# df
Filesystem 1K-blocks Used Avail Capacity Mounted on
/dev/md1 7054 5282 1772 75% /
/dev/md0 7054 5284 1770 75% /mnt
devfs 1 1 0 100% /mnt/dev
/dev/ad0s1a 30856 14666 13722 52% /mnt/mnt
devfs 1 1 0 100% /dev
# ls /mnt
panic: userret: Returning with 1 locks held.
cpuid = 0
KDB: enter: panic
[thread pid 70 tid 100040 ]
Stopped at kdb_enter+0x2b: nop db> show lockedvnodes
Locked vnodes

0xc222d618: tag ufs, type VDIR
usecount 10, writecount 0, refcount 12 mountedhere 0
flags (VV_ROOT)
lock type ufs: EXCL (count 1) by thread 0xc2207360 (pid 70)
ino 2, on dev md1
```

Does anyone have an idea what could be causing this and in which direction I should look?

I can supply a patch against -current to anyone who is interested in investigatig this further together with me, just drop me an Email!

Adrian

freebsd-hackers@xxxxxxxxxxx mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-hackers>

To unsubscribe, send any mail to "freebsd-hackers-unsubscribe@xxxxxxxxxxx"