UFS/FFS Optimisations: Softupdates, Dirpref and Dirhash.

David Malone

November 2001
**Softupdates**

**Problem** Keeping on-disk filesystem metadata recoverably consistent. Historically uses sync writes.

**Solution** Allow on-disk and in-memory versions to differ. Reorder and sequence writes to allow async but maintain consistency.

**Authors** McKusick, Ganger, Patt.

**Pros & Cons** Big win where files are being created, removed or extended: updates almost as fast as async. Semantics of `fsync` maintained. Phantom full disk. Currently no strict NFS semantics.
Introduced OpenBSD 2.3+ (Nov ’97),
FreeBSD 4.0+ (Mar ’98),
NetBSD 1.5+ (Oct ’99).

Enabling Enabled at mount time.
Applies to all subsequent writes.
On recent {Net,Open}BSD use a
fstab/mount option ‘softdep’.
On FreeBSD and older
{Net,Open}BSD use ‘tunefs -n
enable’.

Tuning Several parameters exposed by
sysctl, but no user-serviceable parts.

Testimonial X410src-1.tgz Untar: 233s
to 70s, rm: 177s to 17s.
MH 33k files Create: 645s to 70s,
pack: 1030s to 240s, rm: 279s to 4.7s.
Dirpref

**Problem** Directories placed evenly throughout disk resulting in long seeks between parent and child directories.

**Solution** Bias directory allocation to place related directories close together.

**Author** Grigoriy Orlov.

**Pros & Cons** Big win for lots of directory traversal. No known down side? (though old `fsck` may complain).
Introduced OpenBSD 2.9+ (Apr ’01),
    FreeBSD 4.5+ (Apr ’01),
    NetBSD 1.6+ (Sep ’01).

Enabling Just use any recent kernel.
    Applies to subsequent directory
    layout. To apply to old directory tree
    you will need to rebuild it (cp, rm, mv).

Tuning To maintain even allocation on
    disk, estimates of average file size and
    average files-per-directory needed.
    Defaults to 16kB and 64. Can be set
    with tunefs.

Testimonial X410src-1.tgz Untar: 70s
to 49s, rm: 17s to 7.3s.
MH 33k files No change.
Dirhash

Problem Directory lookups are a linear search. Slow for large directories.

Solution Build in-memory hash table for directories when first accessed.

Author Ian Dowse.

Pros & Cons Big win when you repeatedly access directories with lots of entries. Can be a pessimisation if directory is not accessed again.
**Introduced** FreeBSD 4.4+ (Jun ’01).

**Enabling** Build a kernel with options `UFS_DIRHASH`. Applies to directory access with such a kernel.

**Tuning** Some `sysctl` settings. Make sure `vfs.ufs.dirhash_docheck` set to 0.
Amount of memory available for hashes `vfs.ufs.dirhash_maxmem` and smallest directory worth hashing `vfs.ufs.dirhash_minsize`.
Defaults 2MB and 2.5kB.

**Testimonial** `X410src-1.tgz` No change.

`MH 33k files` Create: 70s to 2.5s,
pack: 240s to 2.5s, rm: 4.7s to 2s.