

Linux Filesystems

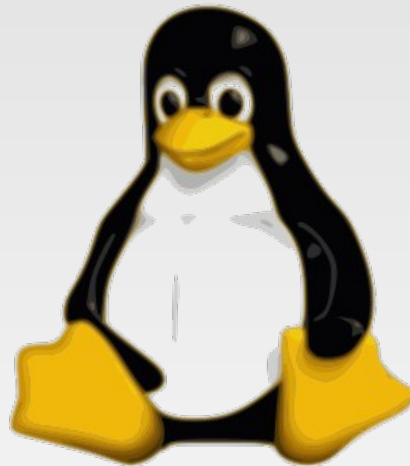
ext2

xfs

ext3

btrfs

ext4



reiserfs

What's Covered

- Major Linux Filesystems
 - ext2|3|4, xfs, reiserfs, btrfs
- Default Filesystems of Major Distributions
 - Debian, Fedora/Redhat, SuSE, Ubuntu
- Filesystem Performance
- Recommendations
 - What filesystem is best for where

What Isn't Covered

- FUSE-based Filesystems
 - zfs
- Network and Cluster Filesystems
 - SAMBA, NFS, OCFS2

ext2

- Introduced: 1993

- Limits:

Block Size	1KB	2KB	4KB	8KB
Max. File	16GB	256GB	2TB	2TB
Max. FS	4TB	8TB	16TB	32TB

- 2TB filesize on <2.6.17
- Distributions: All
- Preferred for SD media
- Extremely well-tested; stable.

ext3

- Introduced: 2001
- Limits: 16GB – 2TB (file), 2-16TB (FS)
- Implementation: B-Tree
- Slower than other Journaling FS'
- In-place upgrade from ext2

"It's Boring, but hey, that's what you want in a filesystem."

ext4

- Introduced: 2008 (stable)
- Limits: 16TB (file @ 4KB block)
1EB (FS limited to 16TB)
- Backwards compatible with ext2/ext3
- Extent-based
- Delayed allocation (allocate-on-flush)
- Much larger limits on sub directories (64k)



- Introduced: 1994 (SGI Irix), 2000 (Linux)
- Limits: 8EB (file), 16EB (FS)
- Online Defragmentation (xfs_fsr)
- Online Resizing (xfs_growfs)
- Disadvantages: growth, metadata operations

reiserfs

- Introduced: 2001
- Limits: 8TB (file – 32bit systems), 16TB (FS)
- Status: unsure?
- Distributions: Debian, Gentoo, SuSE, Ubuntu
- Small files = lots of performance
- Reiserfs v3 commonly available; v4 in -mm kernel.

btrfs

- Next-Gen Linux Filesystem
- Succession of ext2|3|4 family
- Stable: 2011 (hopefully!)
- Limits: 16EB (file), 16EB (FS)
- Features: volume resize, block device, defrag, balancing, transparent compression, subvolumes, snapshots, file cloning, RAID-like, data/metadata checksums, and more!

Filesystem Performance

- ext2/ext3 = slow
- Reiserfs = good, but unsure future
- ext4/xfs/btrfs = best; about same with ext4/btrfs coming ahead.

Filesystems for End Users

- What filesystem works best for end users?
- ext2/ext3/ext4
- xfs/reiserfs/btrfs

Filesystems for the Enterprise

- xfs
- ext4
- btrfs (soon!)
- reiserfs

Credits and Sources

The following is a list of sources I used during this presentation. If I missed one please let me know and I'll be sure to update it!

- ext2: <http://en.wikipedia.org/wiki/Ext2>
- ext3: <http://en.wikipedia.org/wiki/Ext3>
- ext4: <http://en.wikipedia.org/wiki/Ext4>
- btrfs: <http://en.wikipedia.org/wiki/Btrfs>
- xfs: <http://en.wikipedia.org/wiki/Xfs>
- reiserfs: <http://en.wikipedia.org/wiki/Reiserfs>,
<http://en.wikipedia.org/wiki/Reiser4>

Linux Filesystems

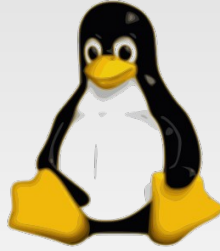
ext2

xf

ext3

btrfs

ext4



reiserfs

What's Covered

- Major Linux Filesystems
 - ext2|3|4, xfs, reiserfs, btrfs
- Default Filesystems of Major Distributions
 - Debian, Fedora/Redhat, SuSE, Ubuntu
- Filesystem Performance
- Recommendations
 - What filesystem is best for where

What Isn't Covered

- FUSE-based Filesystems
 - zfs
- Network and Cluster Filesystems
 - SAMBA, NFS, OCFS2

ext2

- Introduced: 1993

- Limits:

Block Size	1KB	2KB	4KB	8KB
Max. File	16GB	256GB	2TB	2TB
Max. FS	4TB	8TB	16TB	32TB

- 2TB filesize on <2.6.17
- Distributions: All
- Preferred for SD media
- Extremely well-tested; stable.

ext3

- Introduced: 2001
- Limits: 16GB – 2TB (file), 2-16TB (FS)
- Implementation: B-Tree
- Slower than other Journaling FS'
- In-place upgrade from ext2

"It's Boring, but hey, that's what you want in a filesystem."

- Provides online FS growth (can be very slow!)
- Provides a journal so you don't have to fsck the disk after an unclean shutdown.
- Can always be mounted as an ext2 if needbe (failsafe)
- No online defragmentation for ext3; offline only
 - requires mounting as ext2 and not all newer Features are supported; may corrupt data!
- No support for deleted files.
- No support for snapshots (LVM)
- No journal checksumming.

ext4

- Introduced: 2008 (stable)
- Limits: 16TB (file @ 4KB block)
1EB (FS limited to 16TB)
- Backwards compatible with ext2/ext3
- Extent-based
- Delayed allocation (allocate-on-flush)
- Much larger limits on sub directories (64k)

- Born as a series of backwards compatible ext3 extensions.
- Extents offer much better large file performance and reduces fragmentation because it is a range of contiguous physical blocks.
- Backwards compatibility allow you to mount ext2/ext3 volumes as ext4.
- Delayed allocation means ext4 waits to allocate the block(s) until it needs to flush the data to disk.
 - This can cause data loss!



- Introduced: 1994 (SGI Irix), 2000 (Linux)
- Limits: 8EB (file), 16EB (FS)
- Online Defragmentation (xfs_fsr)
- Online Resizing (xfs_growfs)
- Disadvantages: growth, metadata operations

reiserfs

- Introduced: 2001
- Limits: 8TB (file – 32bit systems), 16TB (FS)
- Status: unsure?
- Distributions: Debian, Gentoo, SuSE, Ubuntu
- Small files = lots of performance
- Reiserfs v3 commonly available; v4 in -mm kernel.

btrfs

- Next-Gen Linux Filesystem
- Succession of ext2|3|4 family
- Stable: 2011 (hopefully!)
- Limits: 16EB (file), 16EB (FS)
- Features: volume resize, block device, defrag, balancing, transparent compression, subvolumes, snapshots, file cloning, RAID-like, data/metadata checksums, and more!

- Linux's answer to ZFS
- Provides built-in functionality for RAID/LVM/FS
- many features available online (don't have to take the filesystem down to do them!)
-

Filesystem Performance

- ext2/ext3 = slow
- Reiserfs = good, but unsure future
- ext4/xfs/btrfs = best; about same with ext4/btrfs coming ahead.

- Phoronix.com has filesystem benchmarks (ext4 vs. Btrfs vs. Xfs)

Filesystems for End Users

- What filesystem works best for end users?
- ext2/ext3/ext4
- xfs/reiserfs/btrfs

- generally speaking, we're talking about < 4TB storage.
- not looking at need for "quick" online growth (can take fs offline to do so)
- looking for "boring" filesystems (ext's are best for this!)
- Distributions tend to lean towards ext4 for "end users".
- btrfs will be good when stable because it'll be one filesystem for everything (end users and enterprise alike)
- xfs/reiser are good but more for "power users".

Filesystems for the Enterprise

- xfs
- ext4
- btrfs (soon!)
- reiserfs

- XFS: provides best mix of features and allows for online growth and scalability. Personally my current favorite FS.
- ext4: provides best upgrade path from ext2/ext3 and migrates to btrfs. Performance is on-par with xfs. Scalability still a big limit!
- btrfs: not stable (yet). Hopefully will be **THE** FS of tomorrow for Linux. Will have best mix of features and performance will be with xfs or better. Allows upgrade from ext line of filesystems.
- reiserfs: will be around for some time. Not recommended.

Credits and Sources

The following is a list of sources I used during this presentation. If I missed one please let me know and I'll be sure to update it!

- ext2: <http://en.wikipedia.org/wiki/Ext2>
- ext3: <http://en.wikipedia.org/wiki/Ext3>
- ext4: <http://en.wikipedia.org/wiki/Ext4>
- btrfs: <http://en.wikipedia.org/wiki/Btrfs>
- xfs: <http://en.wikipedia.org/wiki/Xfs>
- reiserfs: <http://en.wikipedia.org/wiki/Reiserfs>,
<http://en.wikipedia.org/wiki/Reiser4>