

Comparison of filesystems in Linux

PV208 Advanced Topics of Linux Administration

Bc. Pavol Babinčák



March 2008

Types of filesystems

- Disk file systems:
 - ext2, ext3
 - ReiserFS
 - XFS
 - *many more*
- Shared disk file systems (known also as SAN or cluster filestems):
 - GFS
 - GPFS
 - Lustre
 - *others*

Extended file system (ext2, ext3)

Designed for Linux. Default file system for many popular Linux distributions. Second extended file system:

- very stable
- through fault can hurt file system
- unfortunately repair is easy ... but quite slow

ext3 = ext2 + journaling:

- compatible with ext2
- repair is fast ... but some metadata operations are slow
- automatic consistency check
- immutable files and append-only files

ReiserFS

Third version of ReiserFS in Linux kernel. Reiser4 not yet merged into mainline.

- good for small files
- not so stable
- less support

Features in Reiser4:

- plugin driven filesystem
- transactions

XFS

One of first journaling FS under UNIX. In Linux kernel since 2.4.

- best for large files, big directories, big filesystems
- slow repair
- creation/deletion of directory entries are slow

Features in XFS:

- delayed allocation for reducing fragmentation
- native backup/restore utilities able to make FS dump without unmounting

Global File System (GFS)

GFS2 is available in kernel since 2.6.19.

- cluster filesystem
- all nodes are equal, running are controlling access to shared resources
- failure cluster member affects only other members using shared resources

Features:

- direct I/O support allows databases to achieve high performance
- dynamic multi-path routing around failed components in SAN

General Parallel File System (GPFS)

- proprietary, generally bundled with IBM hardware
- used on very large clusters (up to 2000 nodes)
- high performance and grids

GPFS features:

- SQL based syntax policies for file placement and management
- shared disk or network block I/O configuration
- offer clustered NFS (High Availability)
- snapshot by copy-on-write

Lustre

Named from words Linux and Cluster. It is not in main kernel but patches to support it exists. Architecture:

- uses modified ext3 as storage FS
- single metadata target
- typically 2-8 object storage servers
- clients accessing to data

Features:

- support for high availability, recovery, transparent reboots
- data blocks striped across objects (bandwidth aggregation, not limited by size of target object)

Resources

- Choosing and Tuning Linux File Systems:
<http://www.valhenson.org/review/choosing.pdf>
- Design and Implementation of the Second Extended Filesystem:
<http://e2fsprogs.sourceforge.net/ext2intro.html>
- Symmetric Cluster Architecture and Component Technical Specifications (about GFS):
<http://people.redhat.com/~teigland/sca.pdf>

Wikipedia keywords: Reiserfs, XFS, IBM General Parallel File System, Lustre (file system), Shared disk file system

Thanks for your attention