

Journaling Filesystem Definition

A *journaling filesystem* is a [filesystem](#) that maintains a special [file](#) called a *journal* that is used to repair any inconsistencies that occur as the result of an improper shutdown of a [computer](#). Such shutdowns are usually due to an interruption of the power supply or to a software problem that cannot be resolved without a rebooting.

A filesystem is a way of storing [information](#) on a computer that usually consists of a hierarchy of [directories](#) (also referred to as the *directory tree*) that is used to organize files. Each [hard disk drive](#) (HDD) or other [storage](#) device as well as each [partition](#) (i.e., logically independent section of a HDD) can have a different type of filesystem if desired.

Journaling filesystems write [metadata](#) (i.e., data about files and directories) into the journal that is flushed to the HDD before each [command](#) returns. In the event of a system [crash](#), a given set of updates may have either been fully *committed* to the filesystem (i.e., written to the HDD), in which case there is no problem, or the updates will have been marked as not yet fully committed, in which case the system will read the journal, which can be *rolled up* to the most recent point of data consistency.

This is far faster than a scan of the entire HDD when rebooting, and it guarantees that the structure of the filesystem is always internally consistent. Thus, although some data may be lost, a journaling filesystem typically allows a computer to be rebooted much more quickly after a system crash.

In the case of non-journaling filesystems, HDD checks during rebooting after a system crash can take many minutes, or even hours in the case of large HDDs with capacities of hundreds of [gigabytes](#). Moreover, if an inconsistency in the data is found, it is sometimes necessary for intervention by a skilled technician to answer complicated questions about how to repair certain filesystem problems. Such downtime can be very costly in the case of big systems used by large organizations.

The most commonly used journaling filesystem for [Linux](#) is the *third extended filesystem* (ext3fs), which was added to the [kernel](#) from version 2.4.16 (released in January 1993). It is basically an extension of [ext2fs](#) to which a journaling capability has been added, and it provides the same high degree of reliability because of the exhaustively field-proven nature of its underlying ext2. Also featured is the ability for ext2 partitions to be converted to ext3 and vice-versa without any need for backing up the data and repartitioning. If necessary, an ext3 partition can even be [mounted](#) by an older kernel that has no ext3 support; this is because it would be seen as just another normal ext2 partition and the journal would be ignored.

The first journaling filesystem to be added to the kernel was actually *ReiserFS*, which was developed by Hans Reiser and others, and it has been the default filesystem on some Linux distributions. The lack of a journaling filesystem was formerly often cited as one of the major factors holding back the widespread adoption of Linux at the enterprise level.

As was the case with ext2, ReiserFS was designed from the ground up for use in Linux. However, unlike ext3, it was also designed from the ground up as a journaling filesystem rather than as an add-on to an existing filesystem, and thus it was widely considered to be the most advanced of the native Linux journaling filesystems. Features include high speed, excellent stability and the ability to pack small files into less disk space than is possible with many other filesystems. It is not clear what, if any, effect the October 2006 arrest of Hans Reiser as a suspect in the murder of his wife might have on the future of this filesystem, although shortly thereafter Novell decided to replace the default ReiserFS with ext3 in its SUSE Linux Enterprise edition.

JFS was originally developed by IBM in the mid-1990s for its AIX [UNIX](#) operating system. It was later ported to the company's OS/2 operating system, and support was added to Linux beginning with the 2.4.20 and 2.5.6 kernels after it was converted to [open source](#). JFS is currently used primarily on IBM enterprise [servers](#), and it is also a good choice for systems that multiboot Linux and OS/2.

XFS was developed in the mid-1990s by Silicon Graphics (SGI) for its 64-bit [IRIX UNIX](#) servers. These servers were designed for advanced graphics processing, and they thus feature the ability to accommodate huge file sizes. The company likewise converted XFS to open source, after which it was also adopted by Linux. Because it is a 64-bit filesystem, XFS features size limitations in the millions of [terabytes](#) (in contrast to the still generous 4TB limit of ext2). Support was added with the 2.5.36 kernel.

The newest journaling filesystem for Linux is [ext4](#), which is now in the final stages of development. Ext4 is designed to accommodate the rapidly approaching era of terabyte (1024 [gigabytes](#)) HDD capacities, and it features support for storage up to 1024 petabytes (1024 terabytes) per volume.

Created April 13, 2007.

Copyright © 2007 The Linux Information Project. All Rights Reserved.