

**Download it now!**[PDF](#) (543 KB)[Free Acrobat™ Reader](#)

## JFS layout

### How the Journaled File System handles the on-disk layout

Steve Best, Linux Technology Center, IBM

Dave Kleikamp, Linux Technology Center, IBM

May 2000

This article describes the on-disk Journaled File System (JFS) layout and the mechanisms used to achieve scalability, reliability, and performance using the on-disk layout structures. You'll learn about the policies and algorithms used to manipulate these structures and where JFS uses B+ trees throughout the file system to increase file system operations.

The JFS architecture can be explained in the context of its disk layout characteristics. The on-disk layout is the format used by JFS to control the file system. This paper covers extent-based file geometry, the directory formats, the formats of block allocation maps, inodes, and other characteristics of the layout structures. It provides detail and examples of the B+ tree data structures used for file layout. B+ trees were selected to increase the performance of reading and writing extents, the most common operations that JFS does.

#### Partitions, aggregates, allocation groups, filesets

Here is the "big picture" view of the on-disk layout.

#### Partitions

A JFS file system is built on top of a partition, which is the abstraction exported to JFS by FDISK.

A partition has:

- A fixed partition block size, with legal values of 512, 1024, 2048, or 4096 bytes. The partition block size defines the smallest unit of I/O supported on the partition. It corresponds to the underlying disk sector size of the physical device making up the partition, with 512 bytes being the most common size.
- A size, PART\_NBlocks, which is the number of partition disk blocks.
- An abstract address space, [ 0 .. PART\_NBlocks - 1 ], of partition disk blocks.

#### Aggregates

To support DCE DFS (Distributed Computing Environment Distributed File System), JFS separates the notion of a disk space allocation pool, called an aggregate, from the notion of a mountable file system sub-tree, called a fileset. The terms aggregate and fileset in this article correspond to their DFS usage. There is exactly one aggregate per partition; there may be multiple filesets per aggregate. In the first release, JFS only supports one fileset per aggregate; however, all of the meta-data has been designed for the fully general case.

[Figure 1](#) shows the layout of an aggregate with two filesets.

#### Contents:

[Partitions, aggregates, AGs, filesets](#)

[Extents, inodes, B+ trees](#)

[Block Allocation Map](#)

[Inode allocations](#)

[Fileset allocation inodes](#)

[File](#)

[Symbolic link](#)

[Directory](#)

[Access Control List \(ACL\)](#)

[Extended Attribute \(EA\)](#)

[Streams](#)

[Aggregate with a fileset](#)

[Summary](#)

[Resources](#)

[About the authors](#)