Fast File System

How to improve file system performance?

By Matthew Tancreti
For COM S 352
Iowa State University

Faster File System

Original UNIX file system had very poor performance (used only 2% of disk bandwidth)

The most significant physical limitation is the difference between random and sequential latencies

Optimizations need to take the disk into account

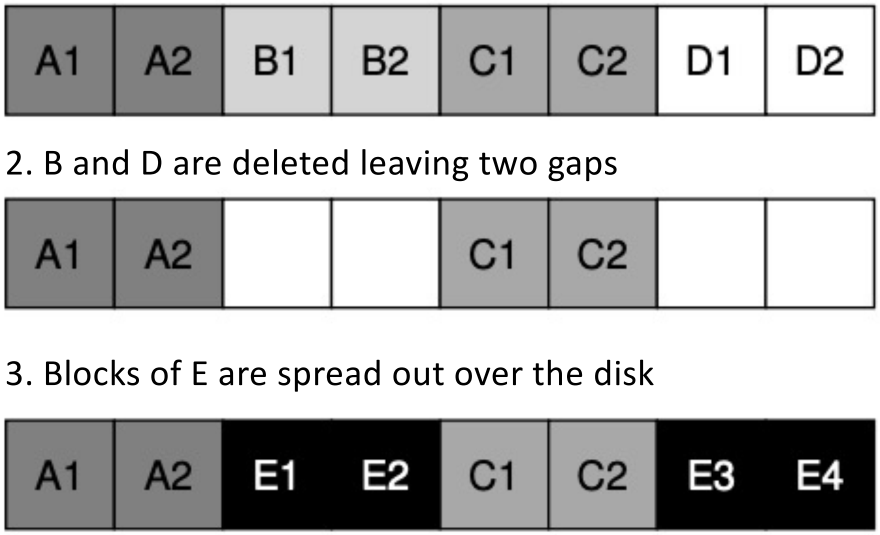
How to improve file system performance?

Hard drive internal [source]

Consequence of Fragmentation

Block based allocation means files can become spread out over the disk

Best performance is when files are written to contiguous memory



1. Assume files A, B, C and D are stored on disk

A1	A2	B1	B2	C1	C2	D1	D2

2. B and D are deleted leaving two gaps

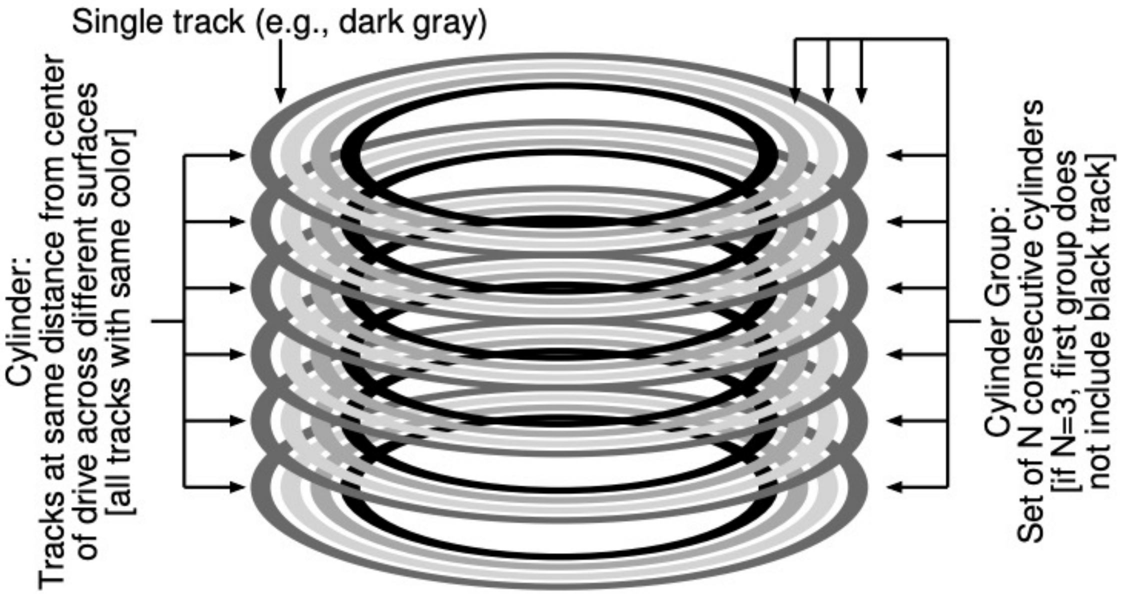
A1	A2			C1	C2		

3. Blocks of E are spread out over the disk

A1	A2	E1	E2	C1	C2	E3	E4

Cylinder Groups

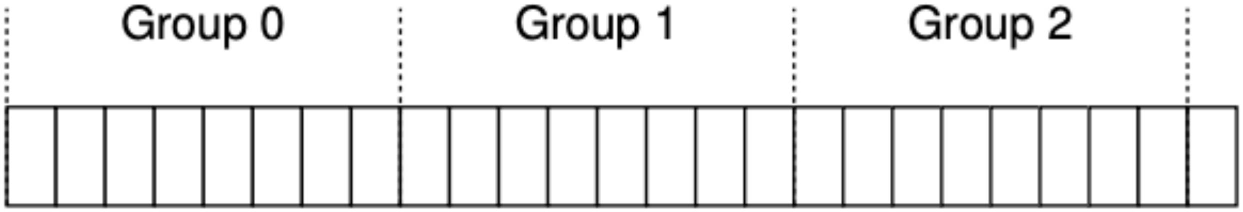
A **cylinder** is a set of tracks near to each other on the drive


Single track (e.g., dark gray)
Cylinder:
Tracks at same distance from center of drive across different surfaces
[all tracks with same color]
Cylinder Group:
Set of N consecutive cylinders
[if N=3, first group does not include black track]

Block Groups

Because most hard drives don’t provide enough information to chose cylinder groups, most file systems are organized by block groups

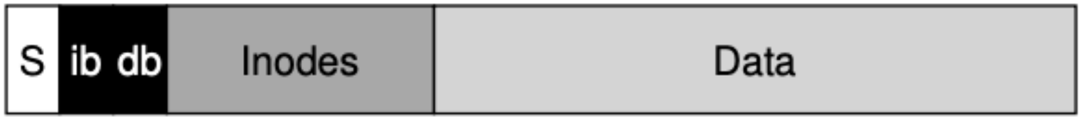
Block groups are consecutive portions of the disk’s address space

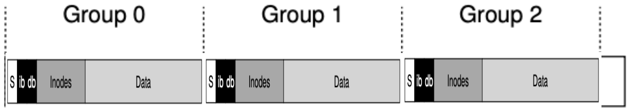


The Berkeley Fast File System (FFS)

Principle: *keep related stuff together*

A single block group (file system has many)





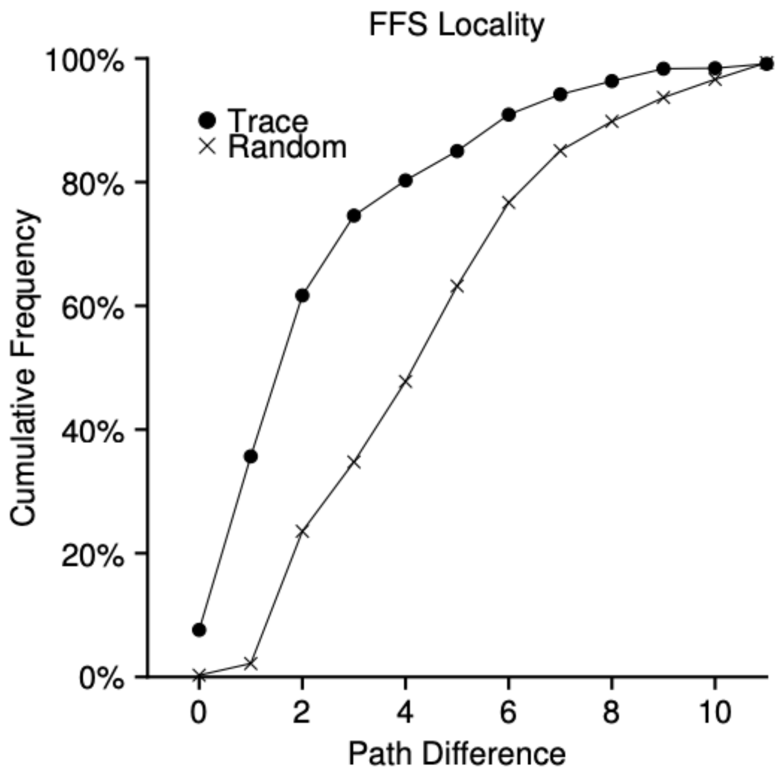
Two heuristics to improve performance:

- Try to allocate data blocks for a file in the same block group as the files inode
- Try to locate files that are together in a directory in the same block group

Path Locality

FFS heuristics are based on another form of locality

Path Locality – consecutive file accesses are likely to be to file paths that are near to each other



Large File Exception

Large file will completely fill block group, preventing files in same directory being in same group

Heuristic

- After blocks are allocated into the first block group (e.g., the 12 direct pointers), FFS places the next “large chunk” (e.g., those pointed to by the first indirect block) in another block group