UNIVERSITY of WISCONSIN-MADISON
Computer Sciences Department

CS 537
Introduction to Operating Systems

Andrea C. Arpaci-Dusseau
Remzi H. Arpaci-Dusseau

File Allocation

Questions answered in this lecture:

- What are typical file access patterns?
- What are different approaches for tracking which blocks belong to a file?
- What are the advantages and disadvantages of each approach?

Workloads

Motivation: Workloads influence design of file system

File characteristics (measurements of UNIX and NT)

- Most files are small (about 8KB)
- Most of the disk is allocated to large files
  - (90% of data is in 10% of files)

Access patterns

- Sequential: Data in file is read/written in order
  - Most common access pattern
- Random (direct): Access block without referencing predecessors
  - Difficult to optimize
- Access files in same directory together
  - Spatial locality
- Access meta-data when access file
  - Need meta-data to find data

Goals

OS allocates LBNs (logical block numbers) to meta-data, file data, and directory data

- Workload items accessed together should be close in LBN space

Implications

- Large files should be allocated sequentially
- Files in same directory should be allocated near each other
- Data should be allocated near its meta-data

Meta-Data: Where is it stored on disk?

- Embedded within each directory entry
- In data structure separate from directory entry
  - Directory entry points to meta-data