FILE SYSTEM RELIABILITY

Module Number 4. Section 5
COP4600 – Operating Systems
Richard Newman

FILE SYSTEMS TOPICS

- Introduction
- Directories
- File Allocation
- Block Management
- File System Reliability
- File System Optimization

FILE SYSTEM RELIABILITY

File systems are expected to be reliable

1. Persistent storage is expected to persist!
2. Data loss (denial of service) can be very damaging
3. Hardware can fail

   Head crashes once were commonplace
4. Attackers can destroy data

   Or ransomware can make it unavailable

HOW TO ACHIEVE RELIABILITY?

Redundancy!

1. Backups
   - Save to archival medium
   - When to back up?
   - Where to store?
2. Hardware Redundancy
   - Mirrored File Systems
   - RAID

FILE SYSTEM BACKUPS

Potential problems solved by backups:

1. Recover from disaster.
2. Recover from stupidity.
3. Recover from maliciousness.

Typical:	Which is the greatest source of loss?
10-15%	External attackers
15-20%	Internal attackers
5-10%	Natural disaster
60-70%	Dumb mistakes ←

BACKUP ISSUES

1. Backup all or part of the system?
   - Selective dumps (e.g., not swap or temp)
2. Don't back up file if not changed
   - Incremental backups
3. Compression of backup or not?
4. Difficulty of backup while file system active
5. Physical security of backup media
   - Maintaining access control (read/write)
   - Maintaining availability/integrity (Kobe)

FILE SYSTEM BACKUPS (2)

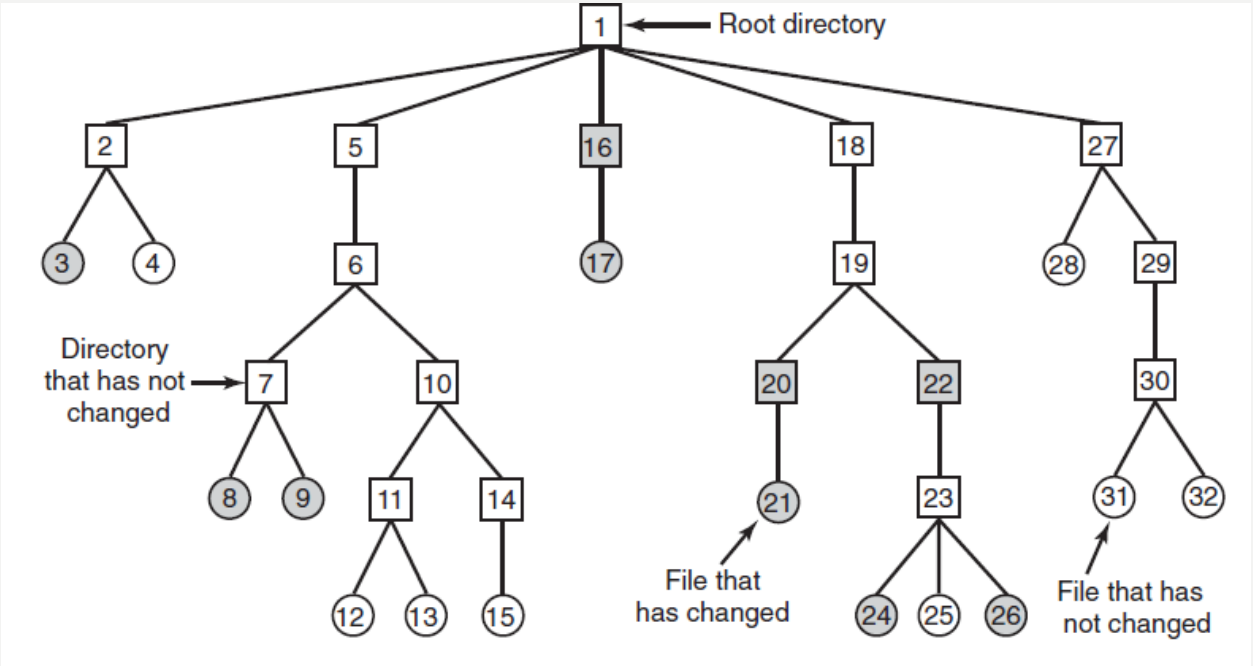


Figure 4-25. A file system to be dumped. The squares are directories and the circles are files. The shaded items have been modified since the last dump. Each directory and file is labeled by its i-node number.

FILE SYSTEM BACKUPS (3)

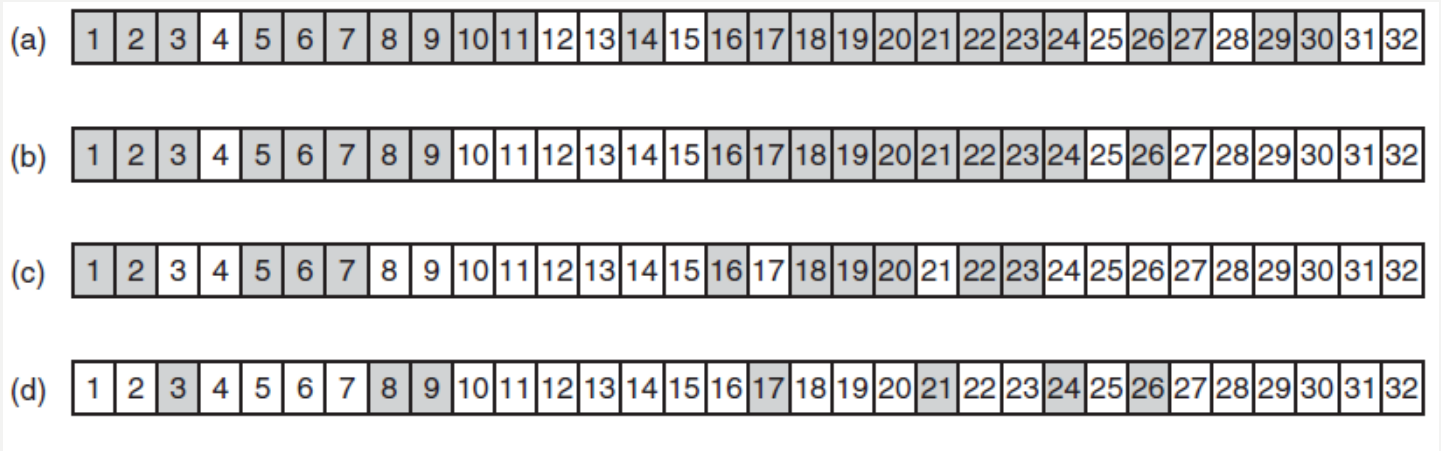
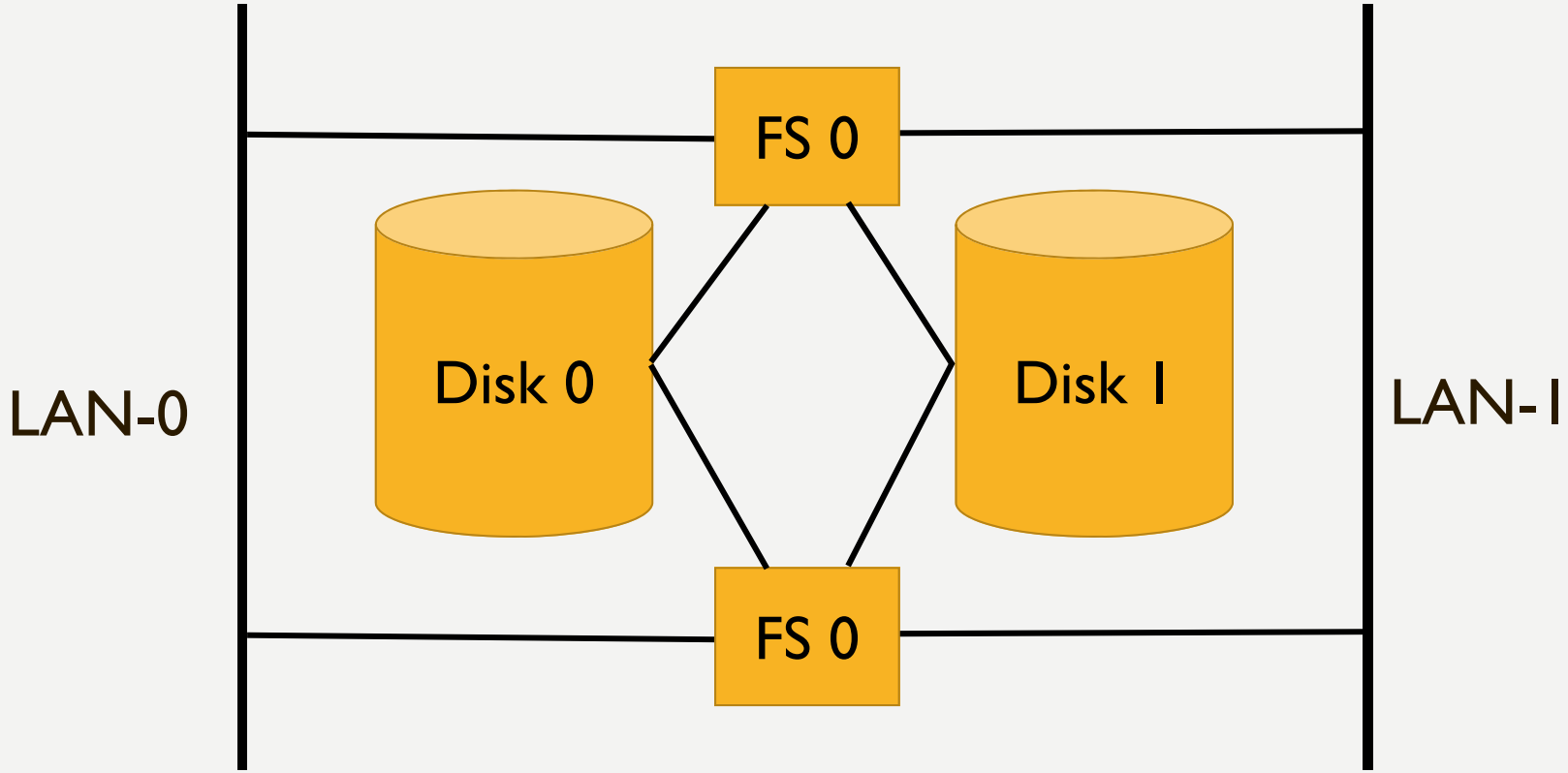


Figure 4-26. Bitmaps used by the logical dumping algorithm.

MIRRORED STORAGE



Every write is duplicated

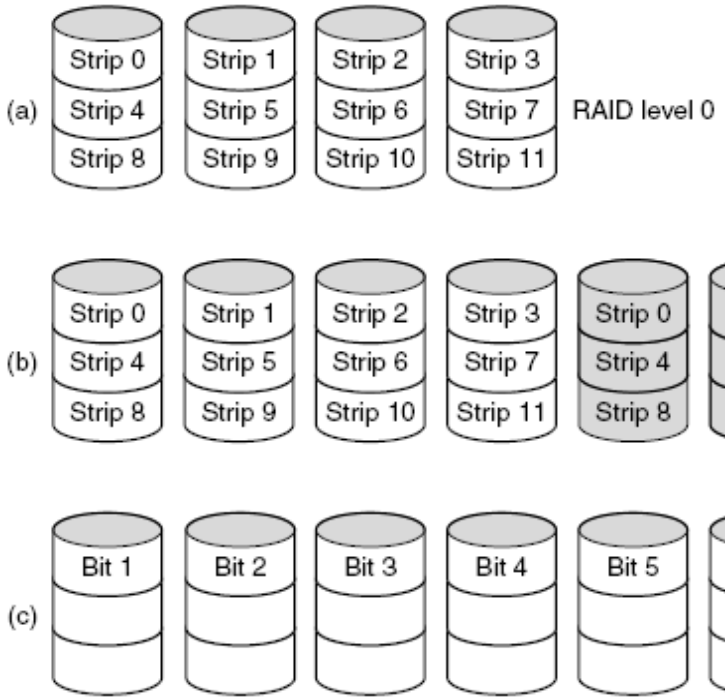
Reads are to whichever system is least loaded

Automatic failover when disk, FS, or network fails

RAID – REDUNDANT ARRAYS OF ~~INEXPENSIVE~~ INDEPENDENT DISKS

- Provide performance gains through parallel hardware access
  - Original target = streaming video
  - Striping
- Provide reliability through redundant storage
- Fast recovery
- Hot swap is possible

RAID STORAGE (1)



Striping only (no redundancy)

RAID 1 = Mirroring only (no striping)

This is RAID 10 (Mirror & Strips)

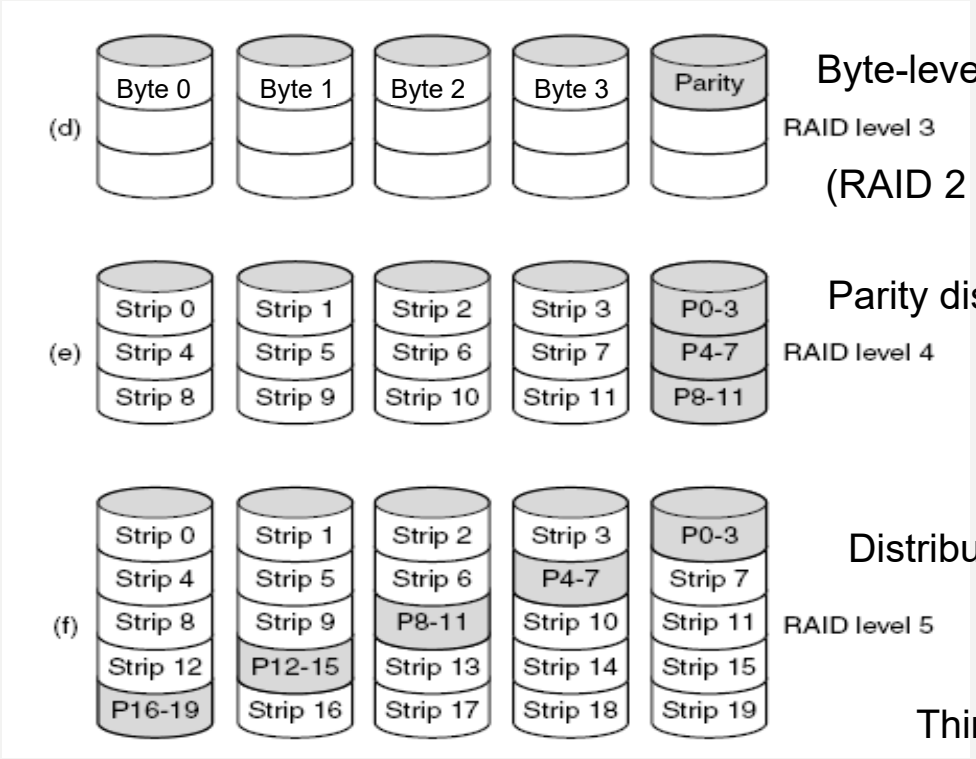
Bit level striping - Synchronous spin

Not used in practice

RAID levels 0 through 2.

Backup and parity drives are shown shaded

RAID STORAGE (2)



Byte-level striping with parity

(RAID 2 also has parity bit)

Parity disk

Distribute parity strips

Why?

Think about parity disk load

RAID levels 3 through 5.

RAID-6 is like RAID-5 but with two redundant drives

JOURNALING FILE SYSTEMS

Robust in face of failure

Steps to remove a file in UNIX:

1. Remove file from its directory.
2. Release i-node to the pool of free i-nodes.
3. Return all disk blocks to pool of free disk blocks.

No matter order of operations, problem if crash before done

1. Write log entry with intentions (idempotent)
2. Do operations
3. Remove entry

SUMMARY

- File systems are expected to be reliable
- Loss of files can be devastating
- Failures happen!
- Need Redundancy
- Backups
- Hardware Redundancy
  - Mirrored File Systems
  - RAID
- Journaling FS