

Persistence: Review

CS 537: Introduction to Operating Systems

Louis Oliphant

University of Wisconsin - Madison

Fall 2023

IO Devices

- System Architecture
 - IO Busses (PCI, SCSI, SATA, USB)
 - Direct Memory Interface, IO Chip
- OS Communication
 - Polling, Interrupts
 - Programmed IO, Direct Memory Access
- Memory Mapped IO vs. Explicit IO Instructions
- Device Drivers

Disks & Scheduling

- Disk Structure
 - Sectors and blocks , tracks, platters, surfaces, cylinders, read/write head, skew
 - Address Space of disk (0 to n-1)
 - RPM and rotation time, Arm movement and seek time
- Disk Performance
 - transfer time: $T_{I/O} = T_{seek} + T_{rotation} + T_{transfer}$
 - transfer rate: $R_{I/O} = \frac{Size_{transfer}}{T_{I/O}}$
 - Random read/write, Sequential read/write
 - Schedulers: SJF, SSTF, SCAN, C-SCAN, SPTF
 - Work-conserving vs. non-work-conserving schedulers

RAID Systems

	RAID-0	RAID-1	RAID-4	RAID-5
	Striping	Mirroring	Parity	Parity w/ Rotation
Capacity	$N \cdot B$	$(N \cdot B)/2$	$(N - 1) \cdot B$	$(N - 1) \cdot B$
Reliability	0	1 (maybe more)	1	1
Sequential Read	$N \cdot S$	$(N/2) \cdot S$	$(N - 1) \cdot S$	$(N - 1) \cdot S$
Sequential Write	$N \cdot S$	$(N/2) \cdot S$	$(N - 1) \cdot S$	$(N - 1) \cdot S$
Random Read	$N \cdot R$	$N \cdot R$	$(N - 1) \cdot R$	$N \cdot R$
Random Write	$N \cdot R$	$(N/2) \cdot R$	$\frac{1}{2} \cdot R$	$\frac{N}{4} \cdot R$
Latency Read	T	T	T	T
Latency Write	T	T	2T	2T

File Systems

- Files, Directories, inodes, Directory Tree, relative/absolute pathnames
- Creating, Reading, Writing Files & Directories
- Permissions, access control lists (with afs)
- strace, lseek, fsync
- sharing file table entries with `fork()` and `dup()`
- hard/soft link & reference counts
- making & mounting file systems

Implementing File Systems

- superblocks, bitmaps (data and inodes), inode table, data blocks
- inode information (including pointers to data)
- directory data (dirent)
- User Operations (e.g. `creat()`) and translations to Disk reads/writes

Fast File System

- Cylinder and Block groups
- Change in layout of data structures
- Namespace locality
- Policies for placement of directories and files
- Large file exception
- sub-blocks, parameterization, long file names

FCK & Journaling

- Crash Consistency
- FCK checks
- Journal structure
- Protocol to maintain consistency
- Data and Metadata journaling
- Recovery process
- Other approaches (soft updates, COW, backpointer-based consistency, optimistic crash consistency)

Log Structured File Systems

- Layout on disk – checkpoint region, segments (data, inodes, imap, segment summary),
- Memory caching – imap and buffered writes
- Garbage Collection – block liveness, which blocks to clean
- Crash Recovery – multiple CRs, roll forward

Flash-based SSDs

- Physical Storage System
- Flash-based Operations
 - Read (a page), Erase (a block), Program (a page)
- Log-Structured FTL
- Garbage Collection / Mapping Tables
- Wear Leveling / Over Provisioning

Data Integrity

- Disk Failure Modes
- Handling Latent Sector Errors
- Detecting Corruption
 - Checksum Functions
 - xor, addition, Fletcher checksum, CRC
 - Checksum Layout
- Misdirected Writes
- Lost Writes
- Scrubbing