

CSE 5330/7330 FALL 2009
I/O and FILE SYSTEMS

Professor Margaret H. Dunham
Department of Computer Science and Engineering
Southern Methodist University
Dallas, Texas 75275
(214) 768-3087
fax: (214) 768-3085
email: mhd@seas.smu.edu
www: <http://www.seas.smu.edu/~mhd>

September 3, 2009

I/O And File Concepts

- What is a File?
- File Systems
- Other File Topics
- Secondary Storage Management
- Disk Scheduling
- Related Topics

TERMS

- Record - Logical unit
- Block (Physical record) - Physical unit of I/O
- Sector - Physical division of a track

-

$$\begin{array}{l} \text{Fixed length} \\ \text{Variable length} \end{array} \left. \vphantom{\begin{array}{l} \text{Fixed length} \\ \text{Variable length} \end{array}} \right\} \left\{ \begin{array}{l} \text{Records} \\ \text{Blocks} \end{array} \right.$$

- Blocking factor
- Spanned/Unspanned

ADDRESSES

- Logical:
 - Key - Field(s) in record whose value(s) is(are) unique across all records
 - Relative Record(Block) - Relative to start of file
 - Inefficient
- Physical:
 - Device/Cylinder/Track/Sector/Block/Offset
 - Efficient
 - Difficult to program
 - Pinned records
- Pointer on disk could be either physical or logical address
- Dangling pointer

LOGICAL vs PHYSICAL

- Logical:

- Attribute/Field
- Record
- File
- Database

- Physical:

- Bit
- Byte
- Word
- Block
- Sector
- Track
- Cylinder
- Device

WHAT IS A FILE?

- Named collection of persistent data.
- File Metadata: Name, Type (organization), Location, Size, Protection, Time of last use (update,create), Owner ID

FILE MANAGEMENT SYSTEM (FMS)

- Software that creates and provides access to data files on secondary storage
- Removes detail and specifics concerning I/O from program
- Statistics
- Tools
- Coordinate access - Simple locking
- Buffer Pool management
- Directory
- Placement on disk
- Organizations - Storage structure
- Access routines - Software to read/write data in files
- Protection

LAYERED FILE SYSTEM

- User(logical) to Device(physical)
- Generic I/O controllers - facilitate uniform view and access
- Device driver
- Interrupt handlers

BUFFER POOL

- Area in memory set aside to read(write) from(to) disk
- May contain many buffers
- Overlap filling of buffer and reading of another
- Double (Triple) Buffering
- Anticipatory I/O

BUFFER POOL MANAGEMENT

- More important for databases than files
- DBMS may provide own buffer pool and management
- Issues:
 - Size - Number of blocks
 - Fetching - When to bring records into buffer
 - * Minimize I/Os
 - * Anticipatory I/O - Anticipate future access based on past. Prefetch blocks. When need will already be there.
 - Replacement - When to replace records
 - * May depend on access.
 - * Keep hot set in buffer
 - * Fix some buffers in memory

DIRECTORY

- Index to files
- Indicates organization
- Indicates (by file) where stored
- Describes each extent
- Contains statistics about file
- Types:
 - Single-level
 - Two-level
 - Tree
 - Acyclic Graph
 - General Graph

PLACEMENT ON DISK

- Allocation:
 - Static vs Dynamic
 - Extent - Unit of allocation; group of blocks
 - Contiguous - All blocks in contiguous disk locations
 - Linked - Each block in file points to next
 - Indexed - Special index blocks point to extents
- Clustering - Placing data accessed together close together on disk
- Partitioning - Divide file and store sections separately. Horizontal and Vertical.

MULTIPLE SMALLER DISKS

- Disk access can become a bottleneck
- Increasing size and even speed does not always solve problem
- Channel limits the amount of data to be transferred to memory
- Solution: Many small disks
- RAID - Redundant Array of Inexpensive Disks
- Striping - Placing data across all disks in interleaved fashion
- Striping Granularity - Size of striped unit

ORGANIZATION vs ACCESS

- Organization:
 - Manner in which blocks are arranged on secondary storage
 - May or may not dictate access techniques allowed
- Access:
 - Software that provides logical access based on specific organization
- Choose based on access needs of applications
- Some organizations are difficult to maintain

ACCESS

- Sequential - Retrieve records in logical order
- Random - Retrieve record based on key
- Direct - Retrieve record based on physical address
- Relative - Retrieve record based on relative position in file
- Binary Search - Randomly retrieve record doing binary search
- Skip Sequential - Position at specific record in file then read sequentially

ORGANIZATION

- Sequential (Ordered, sorted) - Records stored in logical order of key. Access: sequential, relative, binary search.
- Heap (Unordered) - Records added to end. Access: direct.
- B Tree - Multiway balanced search tree. Access: sequential, random, skip sequential.
- Hashing - Store and access record based on address determined when key is hashed. Access - random.
- Accesses provided depend on FMS

PROTECTION

- Access Lists
- Access Groups (owner,group,all others)

OPEN and CLOSE

- OPEN:
 - Prepare file for access
 - Build tables
 - Allocate buffers
 - Request lock
 - Prefetch

- CLOSE:
 - Indicate program is finished with file
 - Free tables
 - Flush buffers
 - Free buffers
 - Unlock

OPEN FILE TABLE

- Directory information
- Pointer to buffers
- Disk address
- Read/write location
- File Descriptor

FILE MOUNTING

- Must open a file before access
- Must mount a file system before use
- Allows many different file systems
- Could have different file system per device
- May mount a directory. This integrates the directory into the directory hierarchy.

DISK SCHEDULING

- How to scheduling I/O requests to same disk
- No starvation
- Reduce waiting time (mean, per process)
- Increase throughput (reduce seek time)
- Reduce response time (mean, per process)
- Types:
 - FCFS
 - SSTF
 - SCAN
 - C-SCAN

FREE SPACE MANAGEMENT

- Free File
- Free List (linked list)
- Bit Maps
- Index

“BAD” BLOCK MANAGEMENT

- “Bad” File
- Bit map

RECOVERY

- Backup
- Roll Forward
- Roll Back
- Log
- Checkpoint