File Management

Chapter 9

1

1

Objectives

- To explain the function of file systems
- To describe the interfaces to file systems
- To discuss file-system design tradeoffs, including access methods, file sharing, file locking, and directory structures
- To explore file-system protection
- To describe the details of implementing local file systems and directory structures
- To describe the implementation of remote file systems
- To discuss block allocation and free-block algorithms and trade-offs

2

File Concept

- Contiguous logical address space
- Types:
  - Data
    - numeric
    - character
    - binary
  - Program

3

3

File Attributes

- **Name** – the only information kept in human-readable form
- **Identifier** – unique tag (number) identifies file within file system
- **Type** – needed for systems that support different types
- **Location** – pointer to file location on device
- **Size** – current file size
- **Protection** – controls who can do reading, writing, executing
- **Time, date, and user identification** – data for protection, security, and usage monitoring
- Information about files are kept in the directory structure, which is maintained on the disk

4

4

File Management

- File management system consists of system utility programs that run as privileged applications
- Input to applications is by means of a file
- Output is saved in a file for long-term storage

5

5

File System Properties

- Long-term existence
  - Stored on disk or secondary/tertiary storage
- Sharable between processes
  - Access can be controlled, with permissions
- Structure
  - Depending on the file structure, a file can have internal structure convenient for a particular application.
  - Files can be organized in hierarchy or more complex structure – to reflect relationships among them.

6

6

File Operations

- Create – define new file and position it within file structure.
- Delete – remove from the file structure and destroyed.
- Open – to allow a process to perform functions on it.
- Close – close with respect to a process.
- Read – read all or a portion of a file.
- Write (update) – add new data, or change values.

7

7

Terms Used with Files

- Field
  - Basic element of data
  - Contains a single value
  - Characterized by its length and data type
- Record
  - Collection of related fields
  - Treated as a unit
    - Example: employee record (Fields: name, emp_num, job_class)
  - May be fixed or variable length

8

8

Terms Used with Files

- File
  - Collection of similar records
  - Treated as a single entity
  - Have file names
  - May restrict access

9

9

Terms Used with Files

- Database
  - Collection of related data
  - Relationships exist among elements
  - Typical DB operations:
    - Retrieve_All
    - Retrieve_One
    - Retrieve_Next
    - Retrieve_Previous
    - Insert_One
    - Delete_One
    - Update_One
    - Retrieve_Few

10

10

File Management Systems

- A set of system software.
- The way a user of application may access files is through the FMS
- Programmer does not need to develop file management software

11

11

Objectives for a File Management System

- Meet the data management needs and requirements of the user
  - Storage, ability to perform operations
- Guarantee that the data in the file are valid
- Optimize performance
  - System throughput, response time (user's view)
- Provide I/O support for a variety of storage device types

12

12

Objectives for a File Management System

- Minimize or eliminate the potential for lost or destroyed data
- Provide a standardized set of I/O interface routines to user processes
- Provide I/O support for multiple users

13

13

Minimal Set of Requirements

- Each user should be able to create, delete, read, write and modify files
- Each user may have controlled access to other users' files
- Each user may control what type of accesses are allowed to the users' files
- Each user should be able to restructure the user's files in a form appropriate to the problem

14

14

Minimal Set of Requirements

- Each user should be able to move data between files
- Each user should be able to back up and recover the user's files in case of damage
- Each user should be able to access the user's files by using symbolic names

15

15

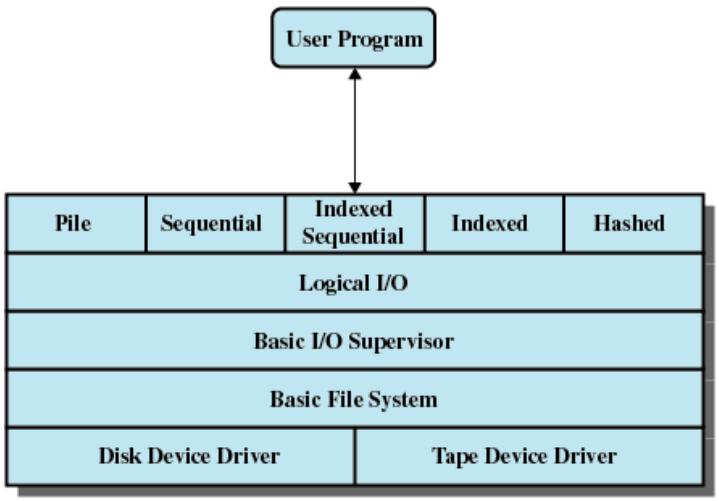


Figure 12.1 File System Software Architecture

16

16

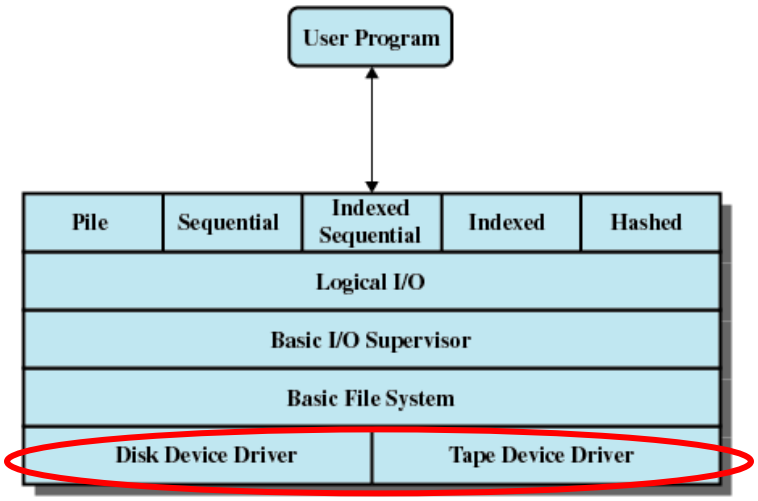
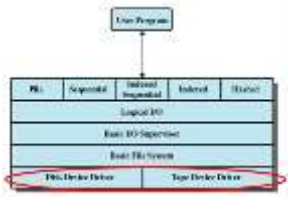


Figure 12.1 File System Software Architecture

17

17

Device Drivers



- Lowest level
- Communicates directly with peripheral devices or their controllers or channels
- Responsible for starting I/O operations on a device
- Processes the completion of an I/O request
- Typical device controlled (for file operation):
  - disk drives, tape drives
- Usually considered as part of OS

18

18

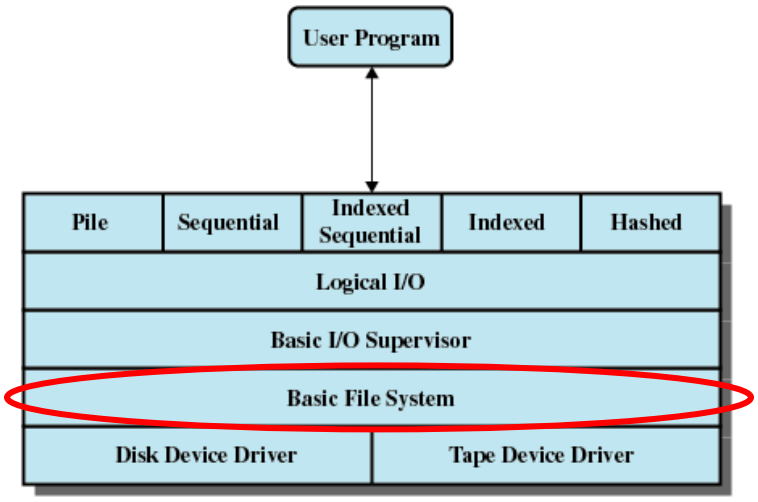


Figure 12.1 File System Software Architecture

19

19

Basic File System

- A.k.a Physical I/O
- Deals with exchanging blocks of data
- Concerned with the placement of blocks
- Concerned with buffering blocks in main memory
- Does not understand the content of data or the structure of the files involved.
- Also part of the OS.

20

20

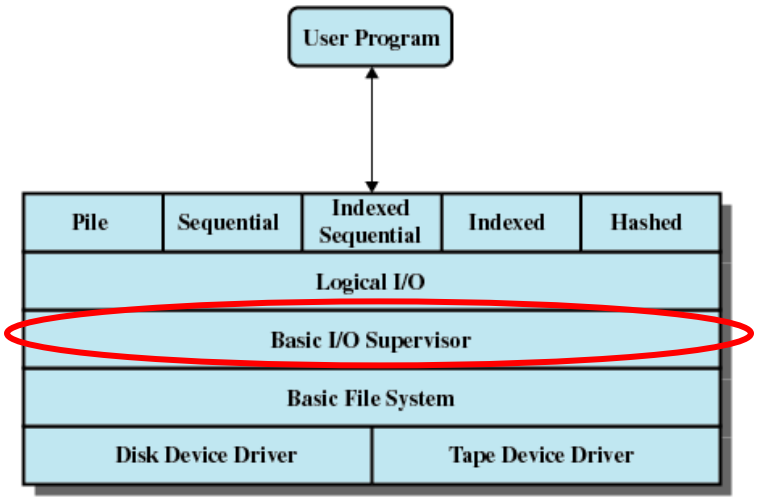
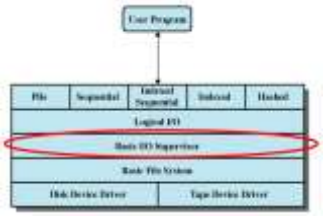


Figure 12.1 File System Software Architecture

21

21

Basic I/O Supervisor



- Responsible for file I/O initiation and termination
- Control structures are maintained
- Concerned with selection of the device on which file I/O is to be performed
- Concerned with scheduling access to optimize performance
- Part of the operating system

22

22

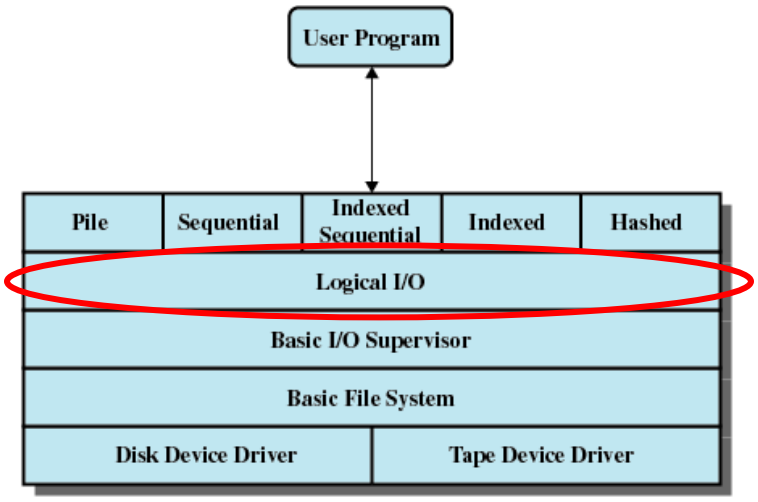


Figure 12.1 File System Software Architecture

23

23

Logical I/O

- Enables users and applications to access records
  - Thus, whereas the basic file system deals with blocks of data, the logical I/O module deals with file records.
- Provides general-purpose record I/O capability
- Maintains basic data about file

24

24

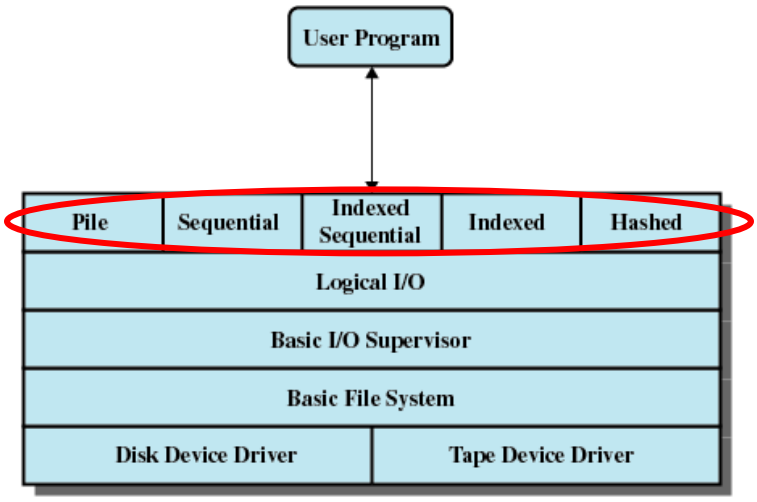
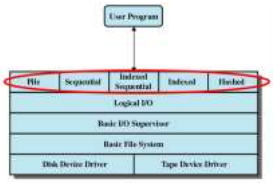


Figure 12.1 File System Software Architecture

25

25

Access Method



- The level of file system closest to the user is often termed as access method
- Reflect different file structures
- Different ways to access and process data
- Provides standard interface between applications and the file system and the devices that hold the data.

26

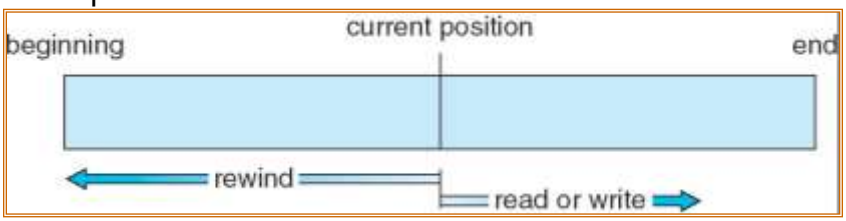
26

Access Methods

- **Sequential Access**
  - read next
  - write next
  - reset
  - no read after last write
    - (rewrite)
- **Direct Access**
  - read *n*
  - write *n*
  - position to *n*
    - read next
    - write next
  - rewrite *n*

n = relative block number

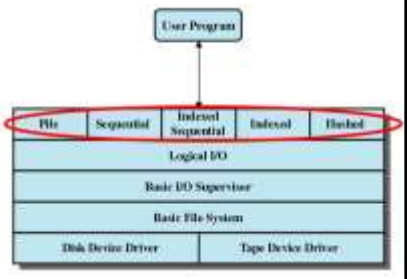
Sequential access



27

File Organization

- … is the logical structuring of records as how they are accessed.
- 5 structures:
  - Pile
  - Sequential file
  - Indexed sequential file
  - Indexed file
  - Direct or hashed file

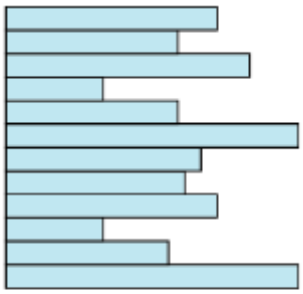


28

28

The Pile

- Least complicated form
- Data are collected in the order they arrive
- Purpose is to accumulate a mass of data and save it
- Records may have different fields
- No structure
- Record access is by exhaustive search
- Easy to update
- But unsuitable for most applications
- Used when data are collected before processing, or when data are not easy to organize
  - Uses space well



Variable-length records
Variable set of fields
Chronological order

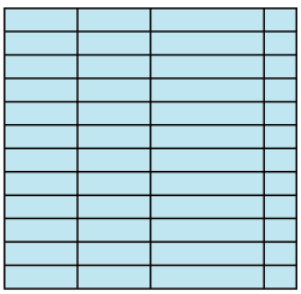
(a) Pile File

29

29

The Sequential File

- Most common
- Fixed format used for records
- Records are the same length
- All fields the same (order and length)
- Field names and lengths are attributes of the file
- One field is the key field (usually the first)
  - Uniquely identifies the record
  - Records are stored in key sequence
- New records are placed in a log file or transaction file
- Batch update is performed to merge the log file with the master file



Fixed-length records
Fixed set of fields in fixed order
Sequential order based on key field

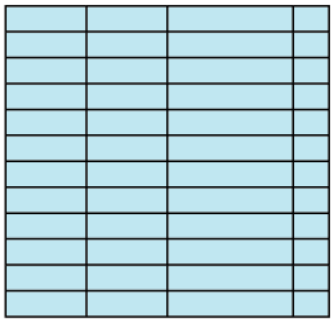
(b) Sequential File

30

30

The Sequential File

- Used in batch applications – optimum if involve processing all records
  - E.g. billing, or payroll applications
- The only file organization that can be stored on tape (as well as disk)
- Poor performance in terms of searching.



Fixed-length records
Fixed set of fields in fixed order
Sequential order based on key field

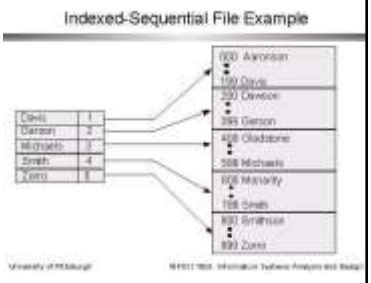
(b) Sequential File

31

31

Indexed-Sequential File

- Index provides a lookup capability to quickly reach the vicinity of the desired record
  - Contains key field and a pointer to the main file
  - Indexed is searched to find highest key value that is equal to or precedes the desired key value
  - Search continues in the main file at the location indicated by the pointer
- A popular approach to overcome the disadvantages of sequential file.
- Maintains the key characteristics of sequential file – records are organized in sequence based on the key field.



32

32

File Organization

Comparison of sequential and indexed sequential

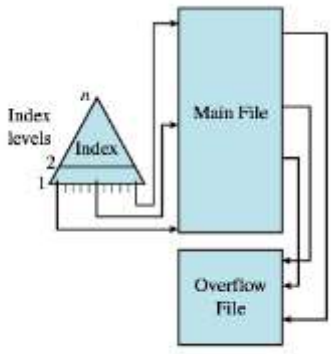
- Example: a file contains 1 million records
- **Sequential:**
  - On average 500,000 accesses are required to find a record in a sequential file
- **Indexed sequential**:
  - If an index contains 1000 entries, it will take on average 500 accesses to find the key, followed by 500 accesses in the main file. Now on average it is 1000 accesses.

33

33

Indexed-Sequential File

- New records are added to an overflow file
- Record in main file that precedes it is updated to contain a pointer to the new record
- The overflow is merged with the main file during a batch update
- Multiple indexes for the same key field can be set up to increase efficiency



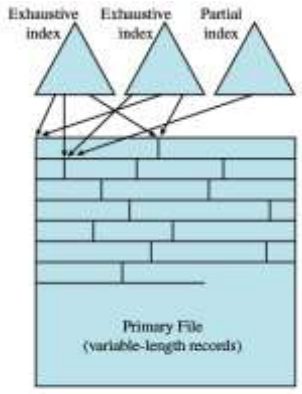
(c) Indexed Sequential File

34

34

Indexed File

- Uses multiple indexes for different key fields
- May contain an exhaustive index that contains one entry for every record in the main file
  - The index is organized as a sequential file for ease of searching
- May contain a partial index – contains entries to records where the field of interest exists.
- Used where timeliness of the info is critical and where data are rarely processed exhaustively
  - E.g. airline reservation syst, inventory control syst.



(d) Indexed File

35

35

Direct or Hashed File

- Directly access a block at a known address
- Key field required for each record
- Used where very rapid access is required,
- Or where fixed-length records are used,
- Or where records are always accessed one at a time.
- E.g. directories, pricing tables, schedules, name lists.

36

36

Table 12.1 Grades of Performance for Five Basic File Organizations [WIED87]

	Space		Update		Retrieval		
	Attributes		Record Size				
File Method	Variable	Fixed	Equal	Greater	Single record	Subset	Exhaustive
Pile	A	B	A	E	E	D	B
Sequential	F	A	D	F	F	D	A
Indexed sequential	F	B	B	D	B	D	B
Indexed	B	C	C	C	A	B	D
Hashed	F	B	B	F	B	F	E

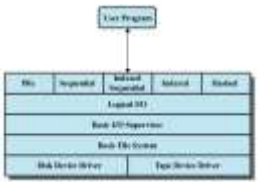
A	= Excellent, well suited to this purpose	$\approx O(r)$
B	= Good	$= O(o \times r)$
C	= Adequate	$\approx O(r \log n)$
D	= Requires some extra effort	$\approx O(n)$
E	= Possible with extreme effort	$= O(r \times n)$
F	= Not reasonable for this purpose	$\approx O(n^{>1})$

where
r = size of the result
o = number of records that overflow
n = number of records in file

37

37

File Management



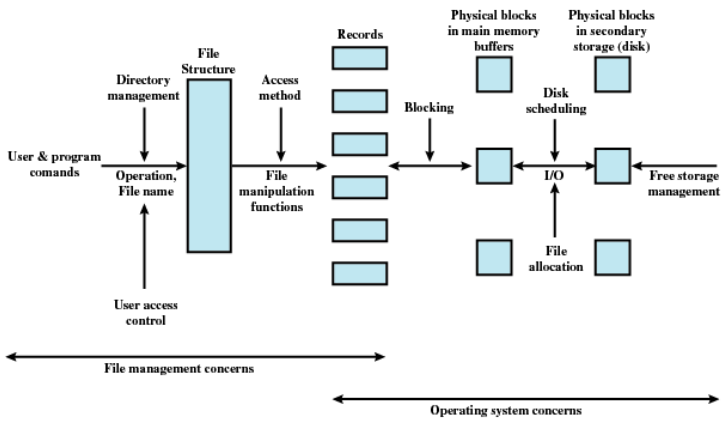


Figure 12.2 Elements of File Management

38

38

File Management Functions

- Identify and locate a selected file
- Use a directory to describe the location of all files plus their attributes
- On a shared system describe user access control
- Blocking for access to files
- Allocate files to free blocks
- Manage free storage for available blocks

39

39

Criteria for File Organization

- Short access time
  - Needed when accessing a single record
  - Not needed for batch mode
- Ease of update
  - File on CD-ROM will not be updated, so this is not a concern
- Economy of storage
  - Should be minimum redundancy in the data
  - Redundancy can be used to speed access such as an index
- Simple maintenance
- Reliability

40

40

File Directories

- Contains information about files
  - Attributes
  - Location
  - Ownership
- Directory itself is a file owned by the operating system
- Provides mapping between file names and the files themselves

41

41

Table 12.2 Information Elements of a File Directory

	Basic Information
File Name	Name as chosen by creator (user or program). Must be unique within a specific directory.
File Type	For example: text, binary, load module, etc.
File Organization	For systems that support different organizations
	Address Information
Volume	Indicates device on which file is stored
Starting Address	Starting physical address on secondary storage (e.g., cylinder, track, and block number on disk)
Size Used	Current size of the file in bytes, words, or blocks
Size Allocated	The maximum size of the file
	Access Control Information
Owner	User who is assigned control of this file. The owner may be able to grant/deny access to other users and to change these privileges
Access Information	A simple version of this element would include the user's name and password for each authorized user.
Permitted Actions	Controls reading, writing, executing, transmitting over a network

42

42

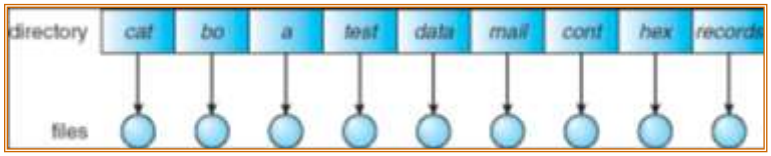
Usage Information	
Date Created	When file was first placed in directory
Identity of Creator	Usually but not necessarily the current owner
Date Last Read Access	Date of the last time a record was read
Identity of Last Reader	User who did the reading
Date Last Modified	Date of the last update, insertion, or deletion
Identity of Last Modifier	User who did the modifying
Date of Last Backup	Date of the last time the file was backed up on another storage medium
Current Usage	Information about current activity on the file, such as process or processes that have the file open, whether it is locked by a process, and whether the file has been updated in main memory but not yet on disk

43

43

Simple Structure for a Directory

- List of entries, one for each file
- Sequential file with the name of the file serving as the key
- Provides no help in organizing the files
- Forces user to be careful not to use the same name for two different files
- A single directory for all users
  - Naming problem
  - Grouping problem

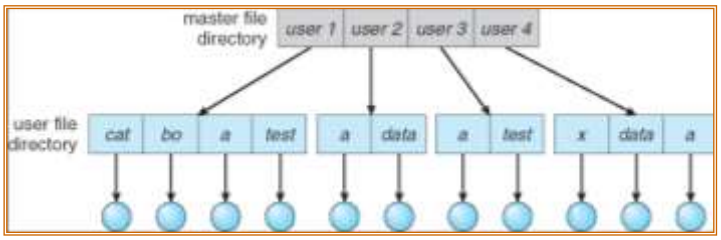


44

44

Two-level Scheme for a Directory

- One directory for each user and a master directory
- Master directory contains entry for each user
  - Provides address and access control information
- Each user directory is a simple list of files for that user
- Still provides no help in structuring collections of files

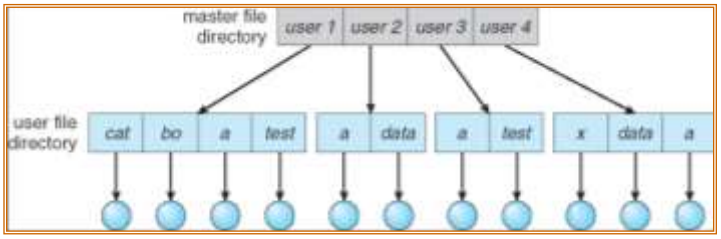


45

45

Two-Level Directory

- Separate directory for each user



- Path name
- Can have the same file name for different user
- Efficient searching
- No grouping capability

46

46

Hierarchical, or Tree-Structured Directory

- Master directory with user directories underneath it
- Each user directory may have subdirectories and files as entries

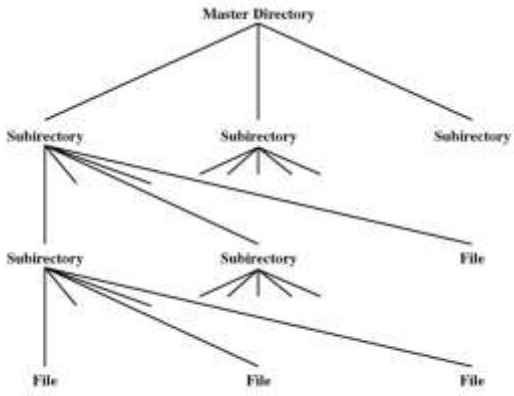
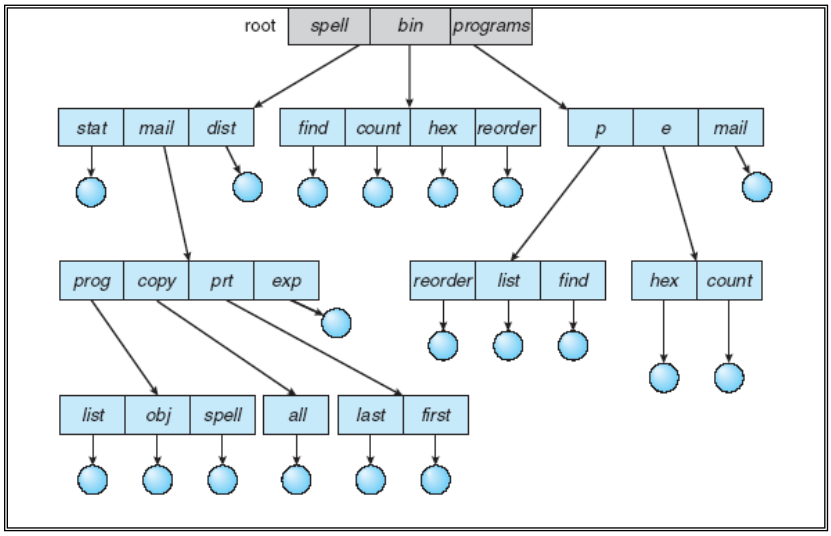


Figure 12.4 Tree-Structured Directory

47

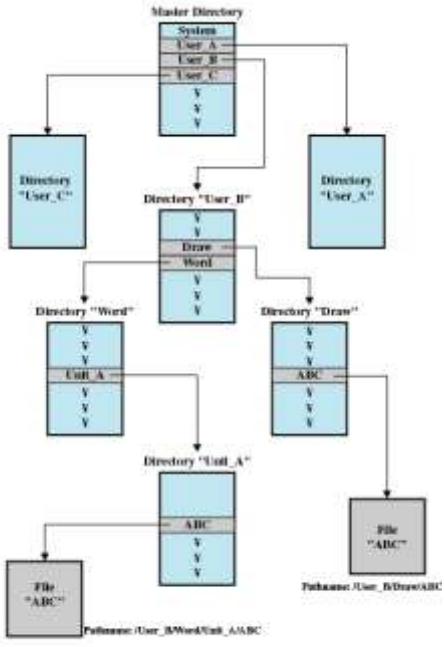
47

Tree-Structured Directories



48

48



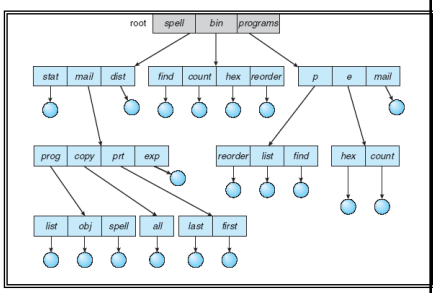
49

Figure 12.5 Example of Tree-Structured Directory

49

Hierarchical, or Tree-Structured Directory

- Files can be located by following a path from the root, or master, directory down various branches
  - This is the pathname for the file
- Can have several files with the same file name as long as they have unique path names
- Current directory is the working directory
- Files are referenced relative to the working directory
- **Absolute** or **relative** path name



50

50

Tree-Structured Directories

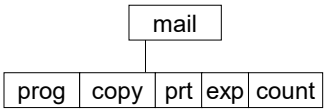
- Creating a new file is done in current directory
- Delete a file

  rm <file-name>
- Creating a new subdirectory is done in current directory

  mkdir <dir-name>

  Example: if in current directory /mail

  mkdir count



Deleting “mail” ⇒ deleting the entire subtree rooted by “mail”

51

51

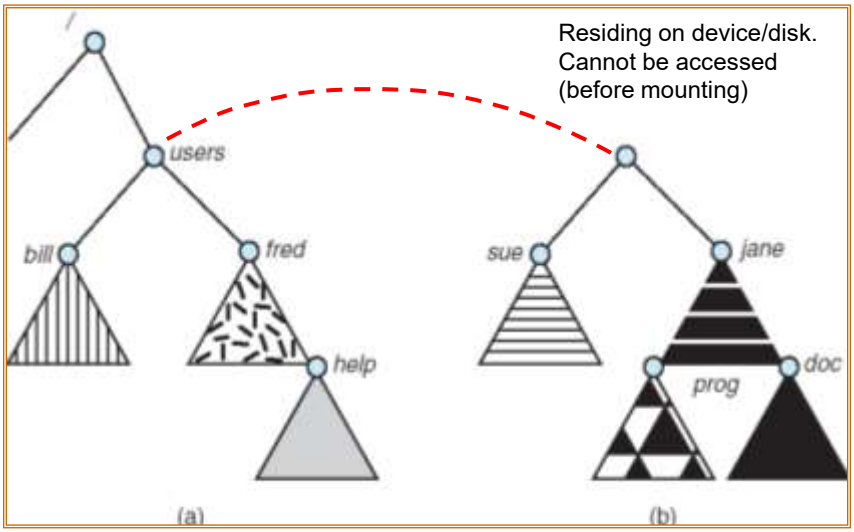
File System Mounting

- Just as a file must be opened before it can be used, a file system must be **mounted** before it can be accessed
- A unmounted file system (i.e. Fig. 11-11(b)) is mounted at a **mount point**.
- Mounting - the OS is given the name of the device and the mount point.
- The mount point is an empty directory.

52

52

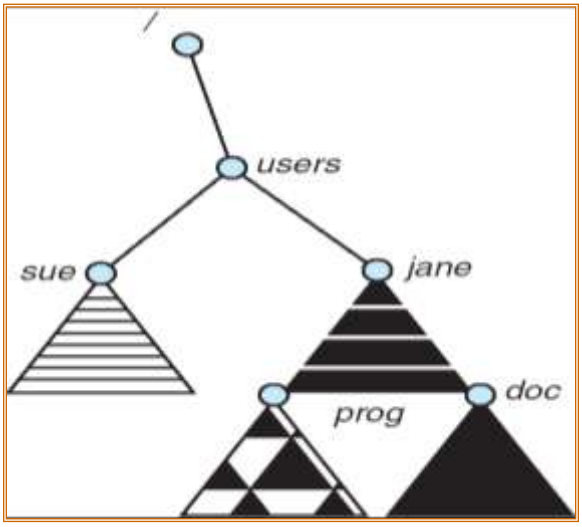
(a) Existing. (b) Unmounted Partition



53

53

Mount Point



54

54

File Sharing

- In multiuser system, allow files to be shared among users
- Sharing of files on multi-user systems is desirable
- Sharing may be done through a **protection** scheme
- On distributed systems, files may be shared across a network
- Network File System (NFS) is a common distributed file-sharing method

55

55

File Sharing – Multiple Users

- **User IDs** identify users, allowing permissions and protections to be per-user
- **Group IDs** allow users to be in groups, permitting group access rights

56

56

File Sharing – Remote File Systems

- Uses networking to allow file system access between systems
  - Manually via programs like FTP
  - Automatically, seamlessly using **distributed file systems**
  - Semi automatically via the **world wide web**
- **Client-server** model allows clients to mount remote file systems from servers
  - Server can serve multiple clients
  - Client and user-on-client identification is insecure or complicated
  - **NFS** is standard UNIX client-server file sharing protocol
  - **CIFS** is standard Windows protocol
  - Standard operating system file calls are translated into remote calls
- Distributed Information Systems **(distributed naming services)** such as LDAP, DNS, NIS, Active Directory implement unified access to information needed for remote computing

57

57

File Sharing – Failure Modes

- Remote file systems add new failure modes, due to network failure, server failure
- Recovery from failure can involve state information about status of each remote request
- Stateless protocols such as NFS include all information in each request, allowing easy recovery but less security

58

58

Protection

- File owner/creator should be able to control:
  - what can be done
  - by whom
- Types of access
  - **Read**
  - **Write**
  - **Execute**
  - **Append**
  - **Delete**
  - **List**

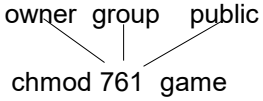
59

Access Lists and Groups

- Mode of access: read, write, execute
- Three classes of users

			RWX
a) **owner access**	7	⇒	1 1 1
b) **group access**	6	⇒	RWX 1 1 0
c) **public access**	1	⇒	RWX 0 0 1

- Ask manager to create a group (unique name), say G, and add some users to the group.
- For a particular file (say *game*) or subdirectory, define an appropriate access.

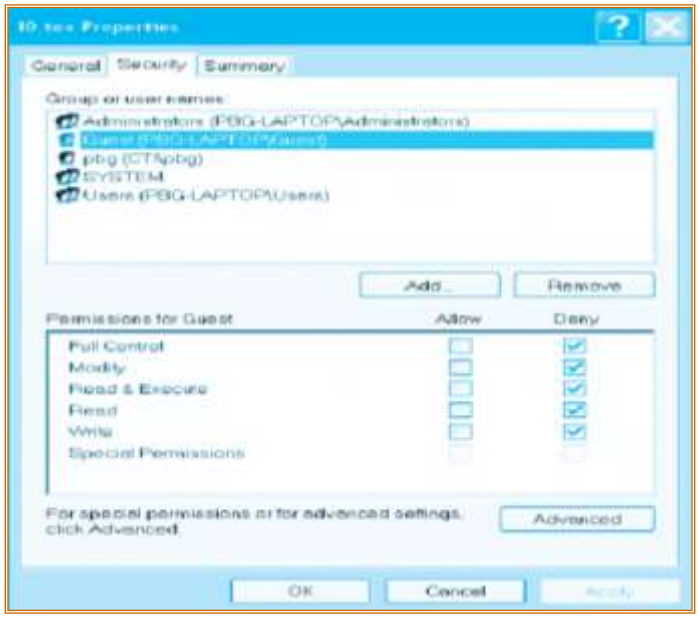


Attach a group to a file:

chgrp G game

60

Windows XP Access-control List Management



61

61

A Sample UNIX Directory Listing

-rw-rw-r--	1 pbg	staff	31200	Sep 3 08:30	intro.ps
drwx------	5 pbg	staff	512	Jul 8 09.33	private/
drwxrwxr-x	2 pbg	staff	512	Jul 8 09:35	doc/
drwxrwx---	2 pbg	student	512	Aug 3 14:13	student-proj/
-rw-r--r--	1 pbg	staff	9423	Feb 24 2003	program.c
-rwxr-xr-x	1 pbg	staff	20471	Feb 24 2003	program
drwx--x--x	4 pbg	faculty	512	Jul 31 10:31	lib/
drwx------	3 pbg	staff	1024	Aug 29 06:52	mail/
drwxrwxrwx	3 pbg	staff	512	Jul 8 09:35	test/

62

62

File Sharing

Two issues in file sharing

- Access rights
- Management of simultaneous access

63

63

Access Rights

- None
  - User may not know of the existence of the file, much less access it
  - To enforce: User is not allowed to read the user directory that includes the file
- Knowledge
  - User can only determine that the file exists and who its owner is
  - User can then petition the owner for additional access rights
- Execution
  - The user can load and execute a program but cannot copy it
  - E.g. propriety program

64

64

Access Rights

- Reading
  - The user can read the file for any purpose, including copying and execution
  - Some system allow viewing, but not copying
- Appending
  - The user can add data to the file but cannot modify or delete any of the file's contents
- Updating
  - The user can modify, delete, and add to the file's data. This includes creating the file, rewriting it, and removing all or part of the data

65

65

Access Rights

- Changing protection
  - User can change access rights granted to other users
- Deletion
  - User can delete the file
- Owners
  - Has all rights previously listed
  - May grant rights to others using the following classes of users
    - Specific user
    - User groups
    - All for public files

66

66

Simultaneous Access

- User may lock entire file when it is to be updated
- User may lock the individual records during the update – finer grain
- Mutual exclusion and deadlock are issues for shared access

67

67

Record Blocking

- For I/O to be performed, records must be organized as blocks.
- Issues:
  - Should blocks be fixed or variable length?
    - Fixed on most systems
  - What should the relative size of blocks?
    - Large blocks – more records passed in one I/O operation
      - Good for sequential processing
      - Bad for random access – unnecessary transfer of unused records.
      - Also require larger buffer – difficult to manage.

68

68

Record Blocking

Three methods of blocking:

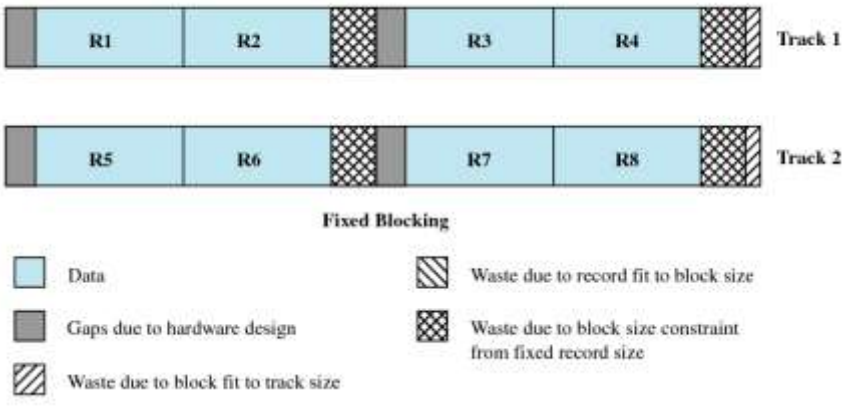
- Fixed blocking
- Variable-length spanned blocking
- Variable-length unspanned blocking

69

69

Fixed Blocking

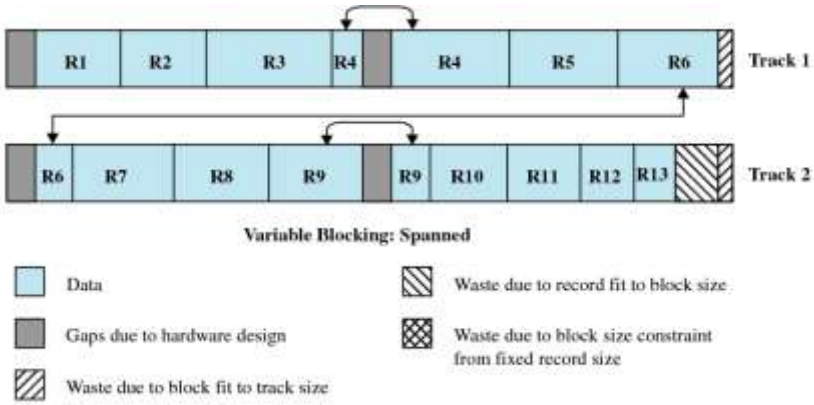
- Fixed length records.
- An integral number of records are stored in a block
- Possible internal fragmentation.



70

Variable Blocking: Spanned

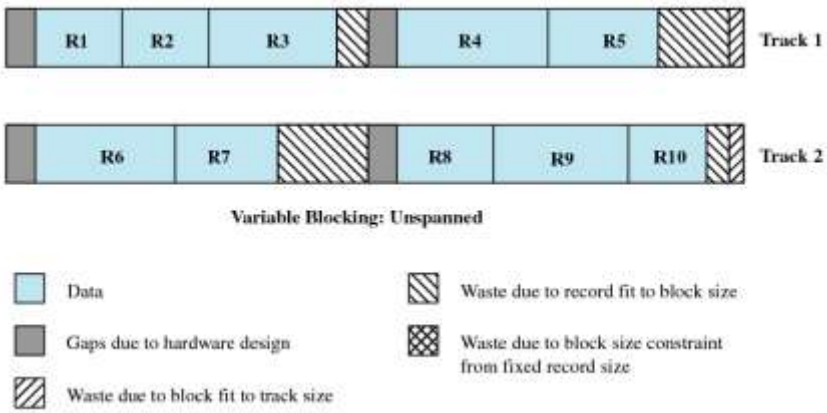
- Variable length records, no limit to record size.
- Packed into blocks with no unused space – some records must span two blocks, with the continuation indicated by a pointer.
- (-) Records spanning 2 blocks require 2 I/O operations.



71

Variable Blocking Unspanned

- Variable length records.
- No spanning employed – limits record size <= block size.
- Possible internal fragmentation.



72

72

Secondary Storage Management

- Space must be allocated to files
- Must keep track of the space available for allocation
- On secondary storage, file consists of a collection of blocks.

73

73

Preallocation

- Need the maximum size for the file at the time of creation
- Difficult to reliably estimate the maximum potential size of the file
- Tend to overestimated file size so as not to run out of space ➔ waste of unused space.
- Better to use dynamic allocation.

74

74

Methods of File Allocation

- Contiguous allocation
- Linked allocation (Chained)
- Indexed allocation

75

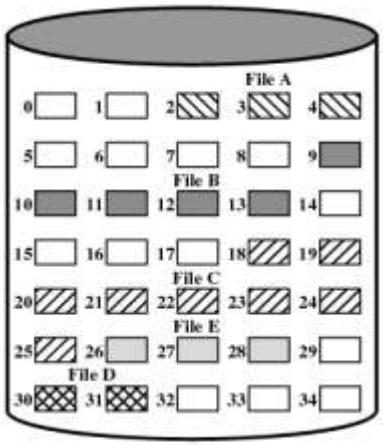
75

Contiguous Allocation

- Single set of blocks is allocated to a file at the time of creation
- Only a single entry in the file allocation table
  - Starting block and length of the file
- External fragmentation will occur
  - Need to perform compaction

76

76



File Allocation Table

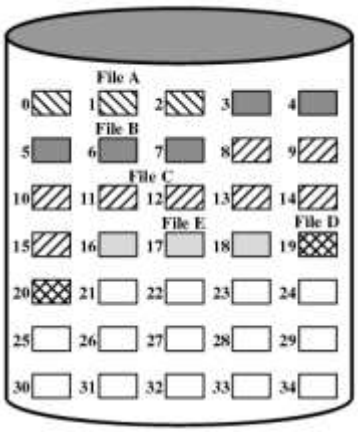
File Name	Start Block	Length
File A	2	3
File B	9	5
File C	18	8
File D	30	2
File E	26	3

Leads to external fragmentation

Figure 12.7 Contiguous File Allocation

77

77



File Allocation Table

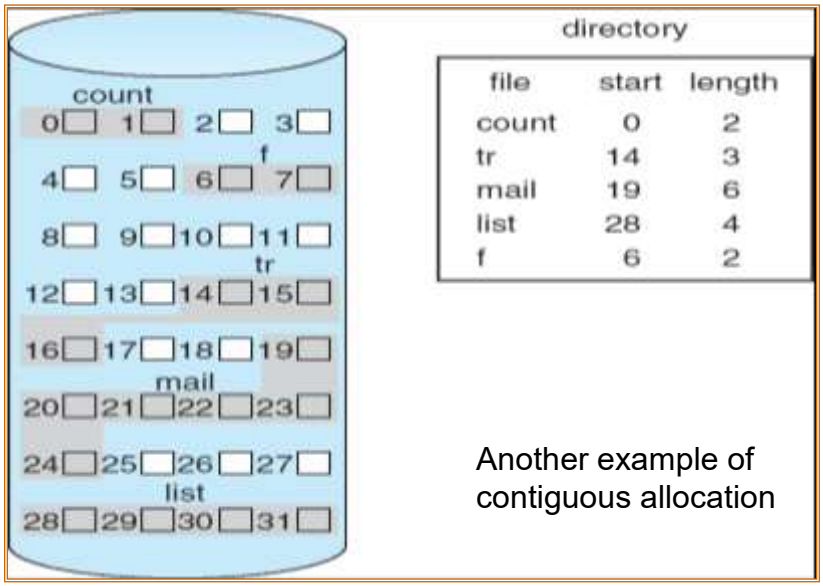
File Name	Start Block	Length
File A	0	3
File B	3	5
File C	8	8
File D	19	2
File E	16	3

- Best method for sequential file
- Easy to retrieve a single block

Figure 12.8 Contiguous File Allocation (After Compaction)

78

78



79

79

Linked/Chained Allocation

- Allocation on basis of individual block
- Each block contains a pointer to the next block in the chain
- Only single entry in the file allocation table
  - Starting block and length of file
- No external fragmentation
  - Any free block can be added to a chain
- Best for sequential files
- No accommodation of the principle of locality

80

80

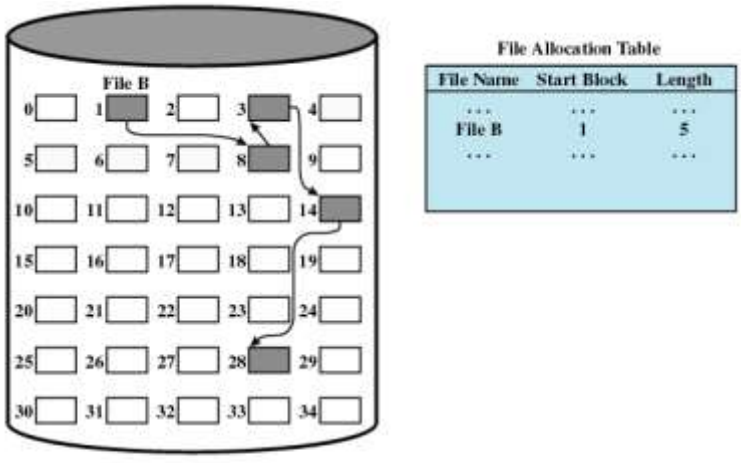


Figure 12.9 Chained Allocation

81

81

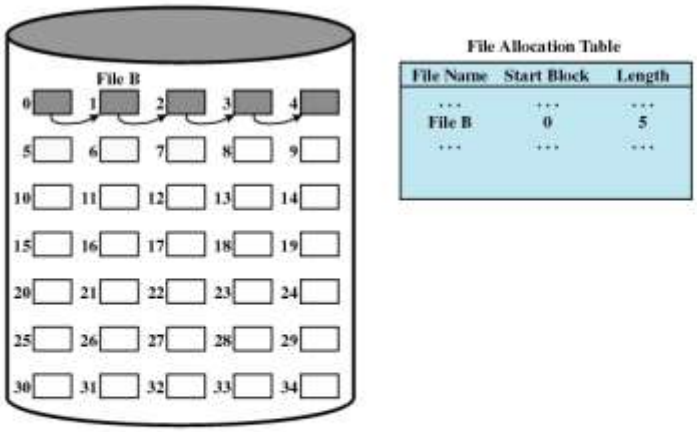
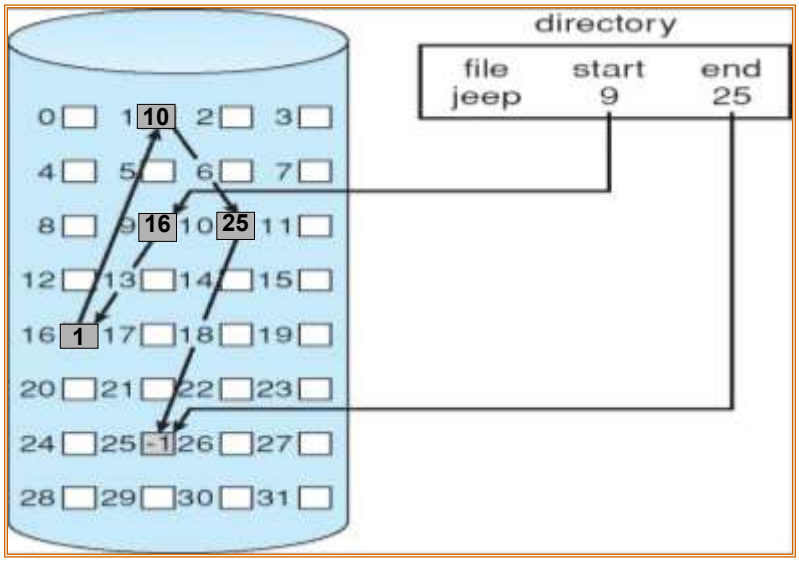


Figure 12.10 Chained Allocation (After Consolidation)

82

82

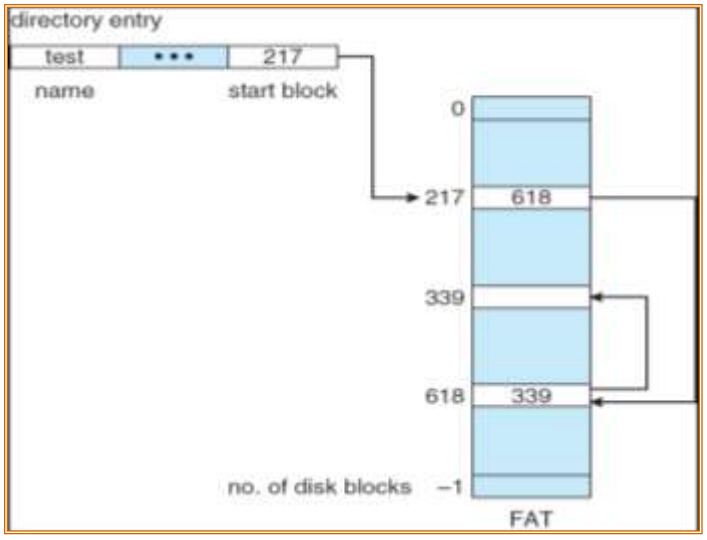
Linked Allocation



83

83

File-Allocation Table



84

84

Indexed Allocation

- File allocation table contains a separate one-level index for each file
- The index has one entry for each portion allocated to the file
- The file allocation table contains block number for the index

85

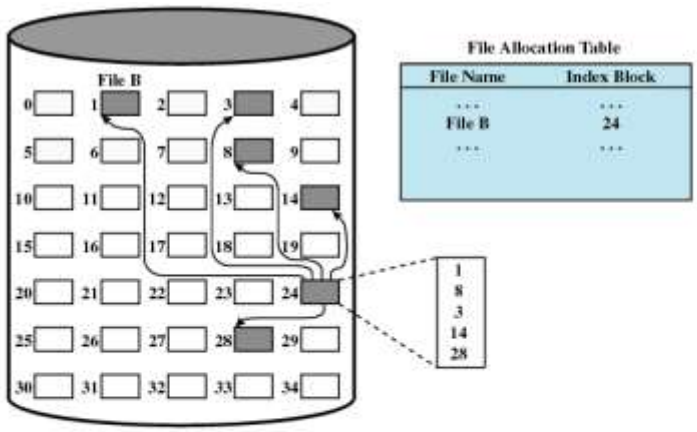


Figure 12.11 Indexed Allocation with Block Portions

86

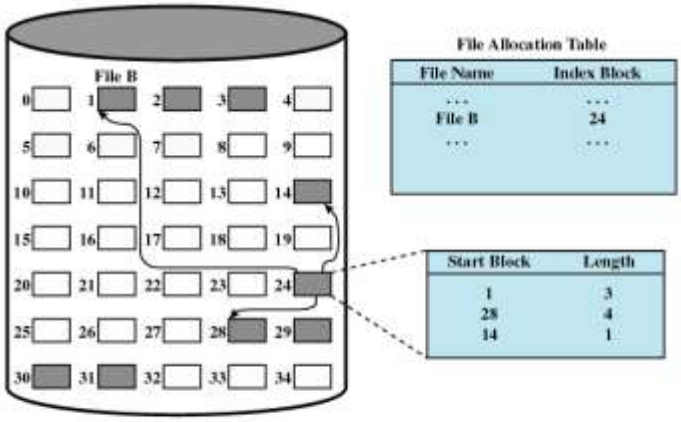
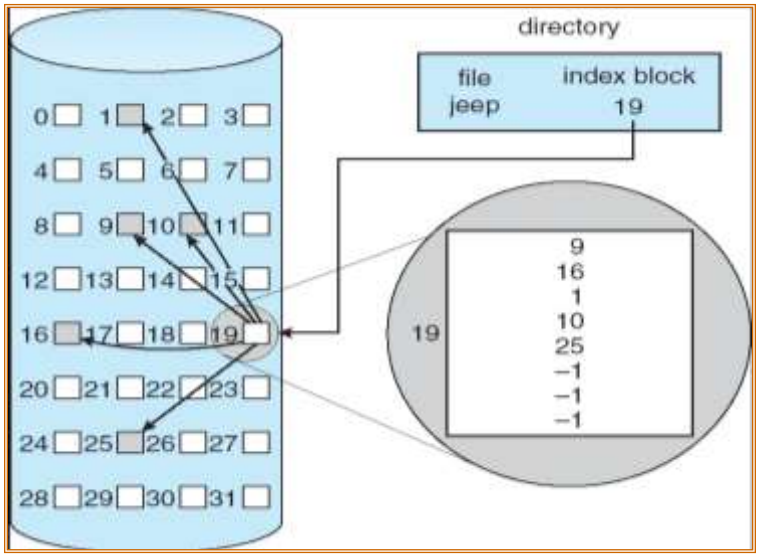


Figure 12.12 Indexed Allocation with Variable-Length Portions

87

87

Example of Indexed Allocation



88

88

Recovery

- Consistency checking – compares data in directory structure with data blocks on disk, and tries to fix inconsistencies
- Use system programs to **back up** data from disk to another storage device (floppy disk, magnetic tape, other magnetic disk, optical)
- Recover lost file or disk by **restoring** data from backup

89

89

Log Structured File Systems

- **Log structured** (or journaling) file systems record each update to the file system as a **transaction**
- All transactions are written to a **log**
  - A transaction is considered **committed** once it is written to the log
  - However, the file system may not yet be updated
- The transactions in the log are asynchronously written to the file system
  - When the file system is modified, the transaction is removed from the log
- If the file system crashes, all remaining transactions in the log must still be performed

90

90

The Sun Network File System (NFS)

- An implementation and a specification of a software system for accessing remote files across LANs (or WANs)
- The implementation is part of the Solaris and SunOS operating systems running on Sun workstations using an unreliable datagram protocol (UDP/IP protocol and Ethernet

91

91

NFS (Cont.)

- Interconnected workstations viewed as a set of independent machines with independent file systems, which allows sharing among these file systems in a transparent manner
  - A remote directory is mounted over a local file system directory
    - The mounted directory looks like an integral subtree of the local file system, replacing the subtree descending from the local directory
  - Specification of the remote directory for the mount operation is nontransparent; the host name of the remote directory has to be provided
    - Files in the remote directory can then be accessed in a transparent manner
  - Subject to access-rights accreditation, potentially any file system (or directory within a file system), can be mounted remotely on top of any local directory

92

92

NFS (Cont.)

- NFS is designed to operate in a heterogeneous environment of different machines, operating systems, and network architectures; the NFS specifications independent of these media
- This independence is achieved through the use of RPC primitives built on top of an External Data Representation (XDR) protocol used between two implementation-independent interfaces
- The NFS specification distinguishes between the services provided by a mount mechanism and the actual remote-file-access services

93

93

NFS Mount Protocol

- Establishes initial logical connection between server and client
- Mount operation includes name of remote directory to be mounted and name of server machine storing it
  - Mount request is mapped to corresponding RPC and forwarded to mount server running on server machine
  - Export list – specifies local file systems that server exports for mounting, along with names of machines that are permitted to mount them
- Following a mount request that conforms to its export list, the server returns a file handle—a key for further accesses
- File handle – a file-system identifier, and an inode number to identify the mounted directory within the exported file system
- The mount operation changes only the user's view and does not affect the server side

94

94

NFS Protocol

- Provides a set of remote procedure calls for remote file operations. The procedures support the following operations:
  - searching for a file within a directory
  - reading a set of directory entries
  - manipulating links and directories
  - accessing file attributes
  - reading and writing files
- NFS servers are **stateless**; each request has to provide a full set of arguments
  (NFS V4 is just coming available – very different, stateful)
- Modified data must be committed to the server's disk before results are returned to the client (lose advantages of caching)
- The NFS protocol does not provide concurrency-control mechanisms

95

95