Operating System

Dr. GuoJun LIU

Harbin Institute of Technology

http://guojunos.hit.edu.cn

Chapter 11

File Management

文件管理

Learning Objectives

- Describe the basic concepts of files and file systems
- Understand the principal techniques for file organization and access
- Explain file directories
- Understand the requirements for file sharing
- Understand the concept of record blocking
- Describe the principal design issues for secondary storage management

Outline

- **Overview**
- **File Organization and Access**
- **File Directory**
- **File Sharing**
- **Record Blocking**
- **Secondary Storage Management**
  - File Allocation
  - Free Space Management

Overview

- **Files and File Systems**
- **File Structure**
- **File Management Systems**
  - Definition
    - Goal
    - File Management System Objectives
    - Minimal User Requirements for each user
  - File System Software Architecture
  - Elements of File Management

Files and File Systems

- **Files**
  - Data collections created by users
  - The File System is one of the most important parts of the OS to a user
  - Desirable properties of files
    - Long-term existence
    - Sharable between processes
    - Structure

Files and File Systems

- **File Systems**
  - Provide a means to store data organized as files as well as a collection of functions that can be performed on files
  - Maintain a set of attributes associated with the file
    - owner, creation time, time last modified, access privileges
  - Typical operations include
    - Create
    - Delete
    - Open
    - Close
    - Read
    - Write

Structure Terms

- **Field**
  - basic element of data
  - contains a single value
  - fixed or variable length
- **Database**
  - collection of related data
  - relationships among elements of data are explicit
  - designed for use by a number of different applications
  - consists of one or more types of files
- **File**
  - collection of **similar records**
  - treated as a single **entity**
  - may be **referenced by name**
  - access control restrictions usually apply at the **file level**
- **Record**
  - collection of related fields that can be treated as a unit by some application program
  - fixed or variable length

File Management Systems

- **Definition**
  - A file management system is that set of **system software** that provides **services to users and applications** in the use of **files**
  - **the only way** that a user or application may access files is through the file management system
- **Goal**
  - This **relieves** the user or programmer of the necessity of developing **special-purpose software** for **each application**
  - and provides the **system** with a **consistent, well-defined means** of **controlling** its **most important asset**

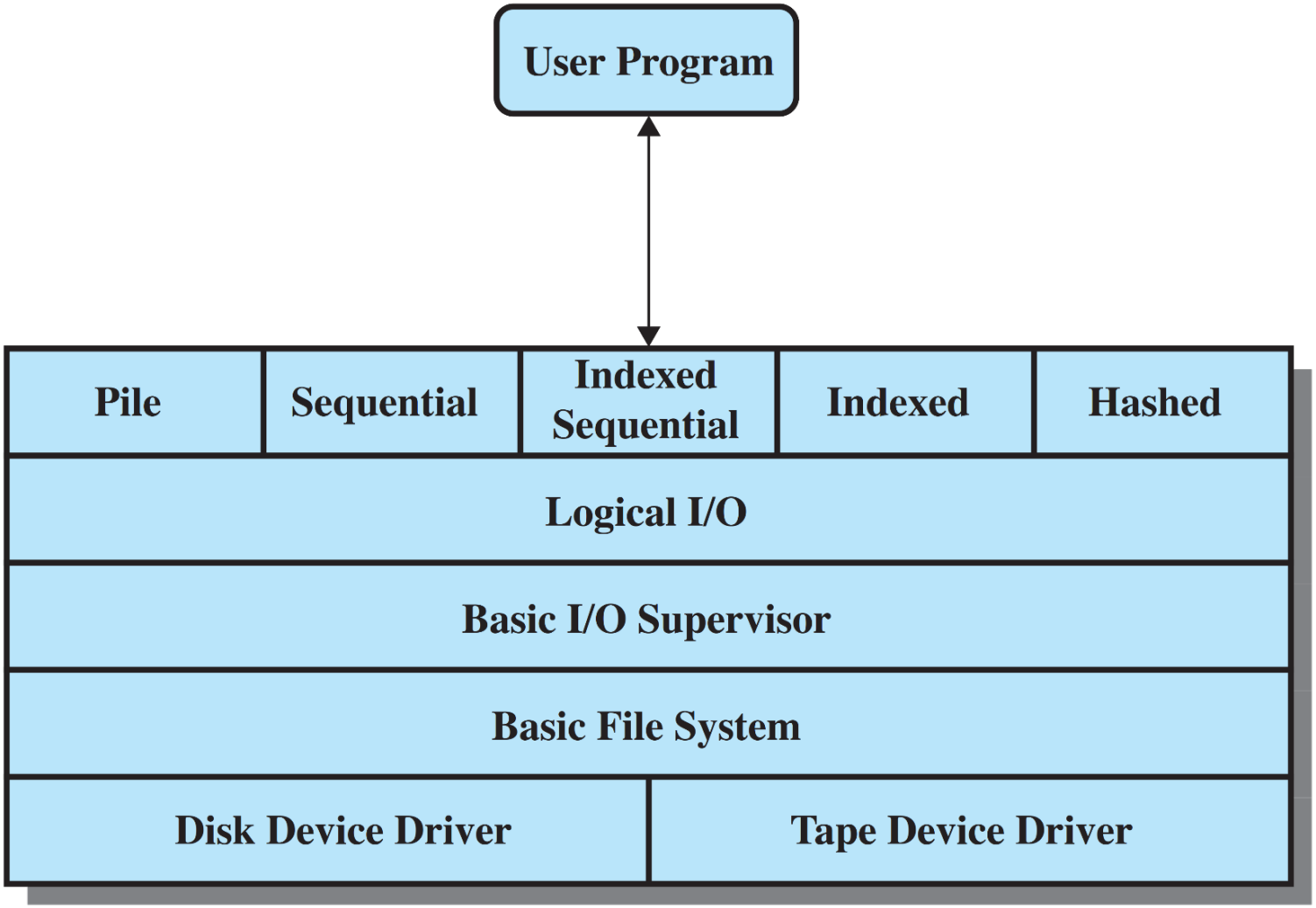
File Management Systems

- **File Management System Objectives**
  - **Meet the data management needs of the user**
  - Guarantee that the data in the file are valid
  - Optimize performance
  - Provide I/O support for a variety of storage device types
  - Minimize 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 in the case of multiple-user systems

File Management Systems

- **Minimal User Requirements for each user**
  - should be able to create, delete, read, write and modify files
  - may have controlled access to other users' files
  - may control what type of accesses are allowed to the files
  - should be able to restructure the files in a form appropriate to the problem
  - should be able to move data between files
  - should be able to back up and recover files in case of damage
  - should be able to **access** his or her files **by name rather than by numeric identifier**

File System Software Architecture



Device Drivers

- Lowest level
- Communicates directly with peripheral devices
- Responsible for starting I/O operations on a device
- Processes the completion of an I/O request
- Considered to be part of the operating system

Basic File System

- Also referred to as the physical I/O level
- Primary interface with the environment outside the computer system
- Deals with blocks of data that are exchanged with disk or tape systems
- Concerned with the placement of blocks on the secondary storage device
- Concerned with buffering blocks in main memory
- Considered part of the operating system

Basic I/O Supervisor

- Responsible for all file I/O initiation and termination
- Control structures that deal with device I/O, scheduling, and file status are maintained
- Selects the device on which I/O is to be performed
- Concerned with scheduling disk and tape accesses to optimize performance
- I/O buffers are assigned and secondary memory is allocated at this level
- Part of the operating system

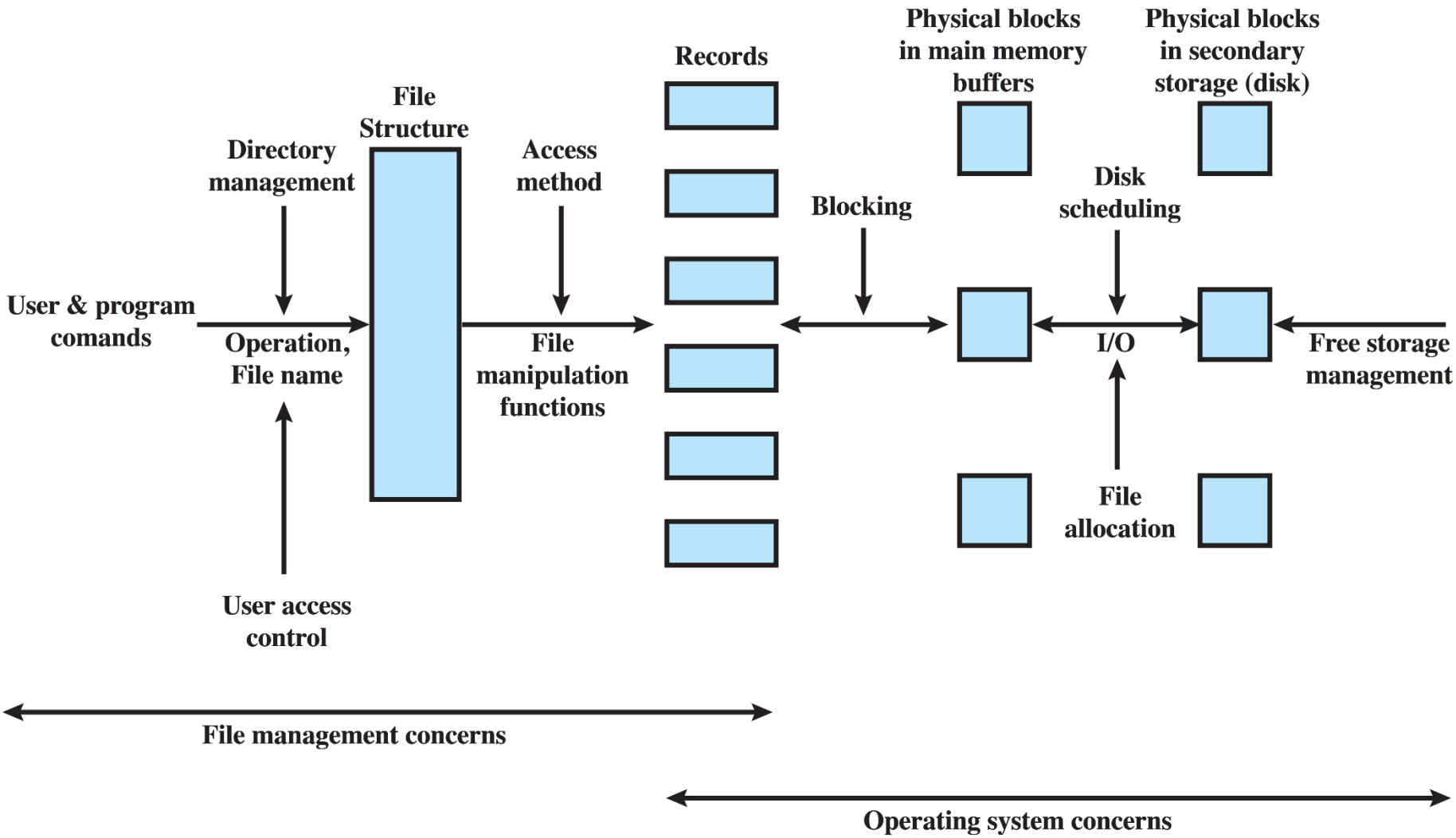
Logical I/O

- Enables users and applications to access records
- Provides general-purpose record I/O capability
- Maintains basic data about file

Access Method

- Level of the file system closest to the user
- Provides a standard interface between applications and the file systems and devices that hold the data
- Different access methods reflect different file structures and different ways of accessing and processing the data

Elements of File Management



File Organization and Access

- **File organization**
  - is the **logical structuring of the records** as determined by the way in which they are accessed
- **In choosing a file organization, important criteria**
  - **short access time**
  - ease of update
  - **economy of storage**
  - simple maintenance
  - reliability

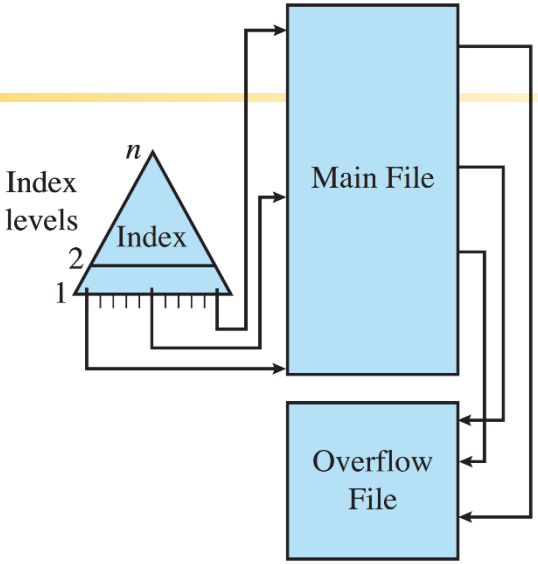
Indexes, redundancy, increasing the speed of access

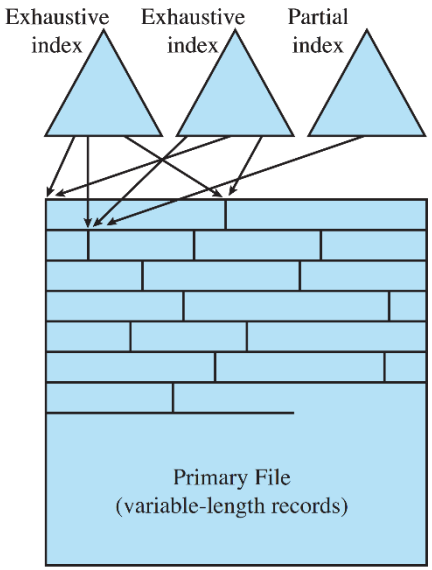
minimum redundancy

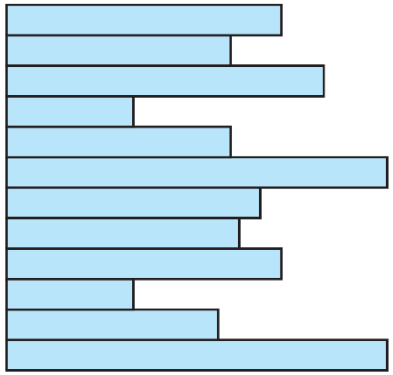
- **Priority of criteria depends on the application that will use the file**

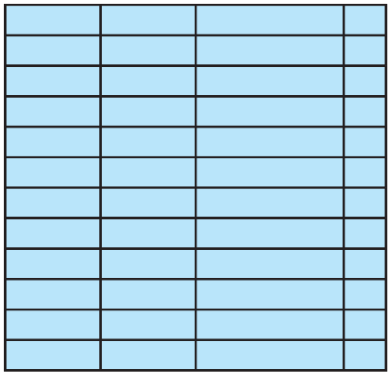
File Organization Types

- The pile
- The sequential file
- The indexed sequential file
- The indexed file
- The direct, or hashed, file









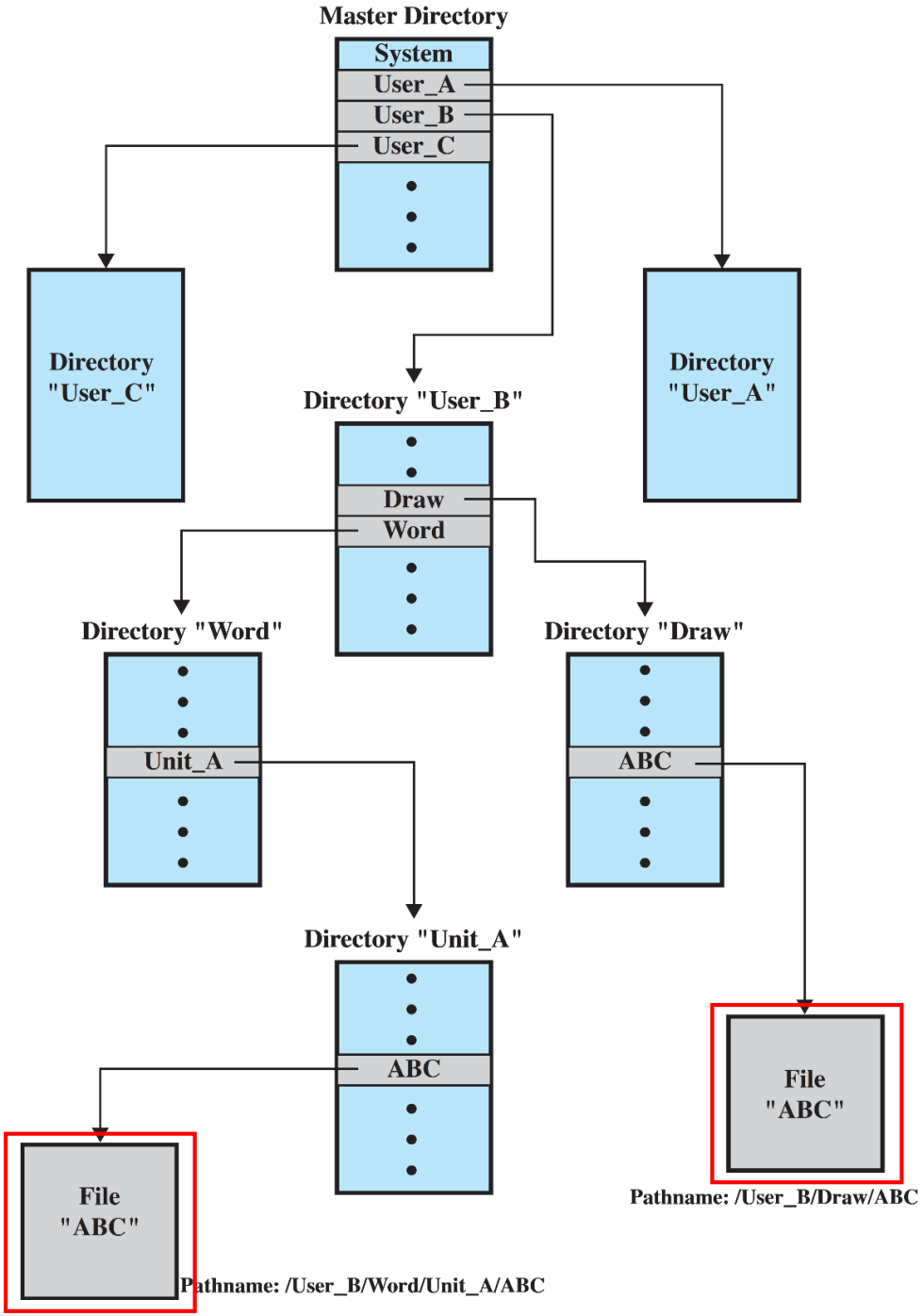
File Directory Information

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, and transmitting over a network
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

Operations Performed on a Directory

- **To understand the requirements for a file structure, it is helpful to consider the types of operations that may be performed on the directory**
  - Search
  - Create files
  - Delete files
  - List directory
  - Update directory

Tree-Structured Directory



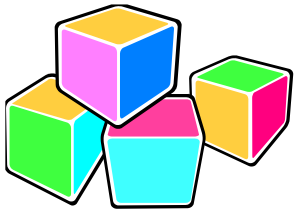
File Sharing

- **In a multiuser system, there is almost always a requirement for allowing files to be shared among a number of users**
- **Two issues arise**
  - access rights
  - the management of simultaneous access

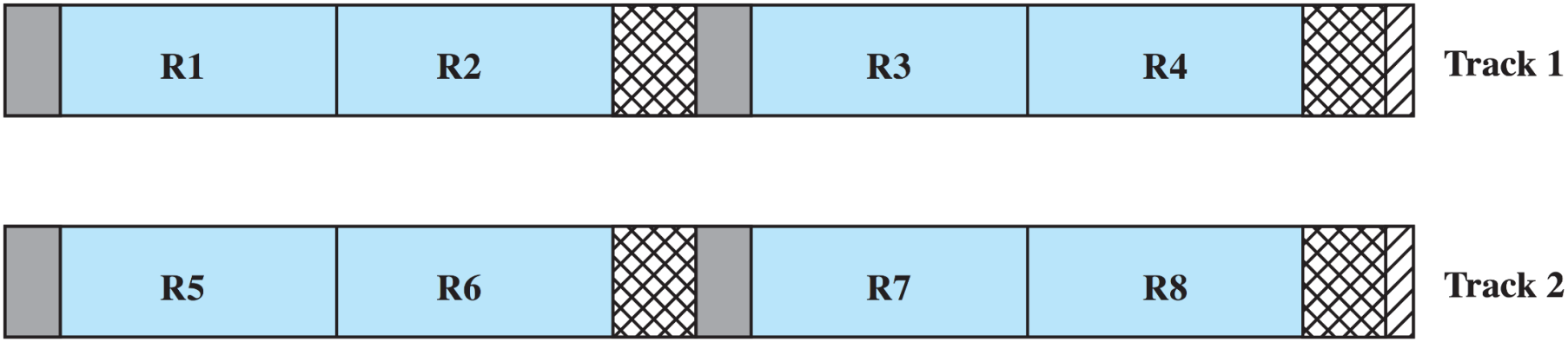
Record Blocking

- **Blocks are the unit of I/O with secondary storage**
  - for I/O to be performed **records must be organized as blocks**

- **Given the size of a block, three methods of blocking can be used**
  - Fixed-Length Blocking
  - Variable-Length Spanned Blocking
  - Variable-Length Unspanned Blocking



Fixed Blocking



Data

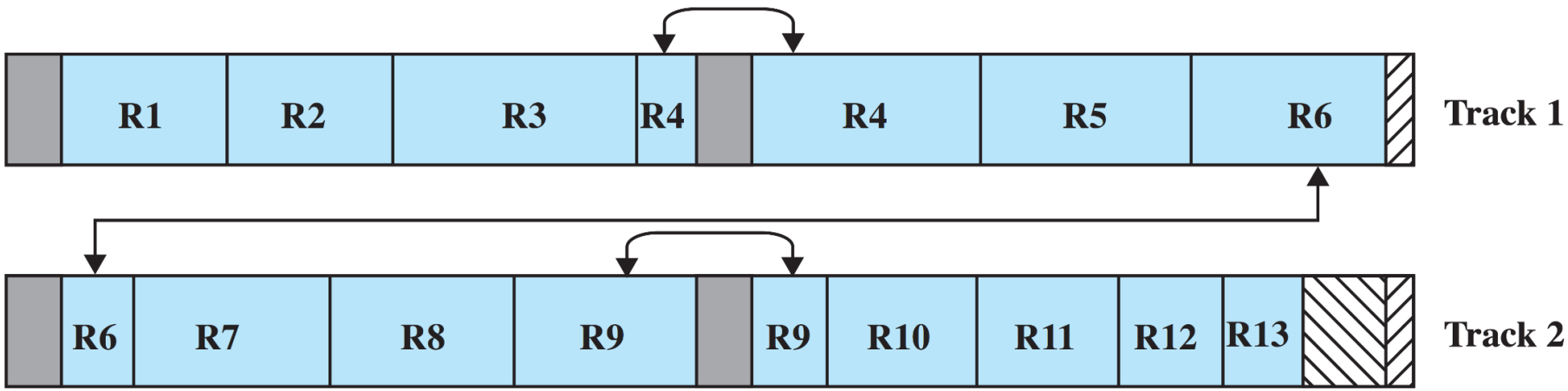
Gaps due to hardware design

Waste due to block fit to track size

Waste due to record fit to block size

Waste due to block size constraint from fixed record size

Variable Blocking: Spanned



Data

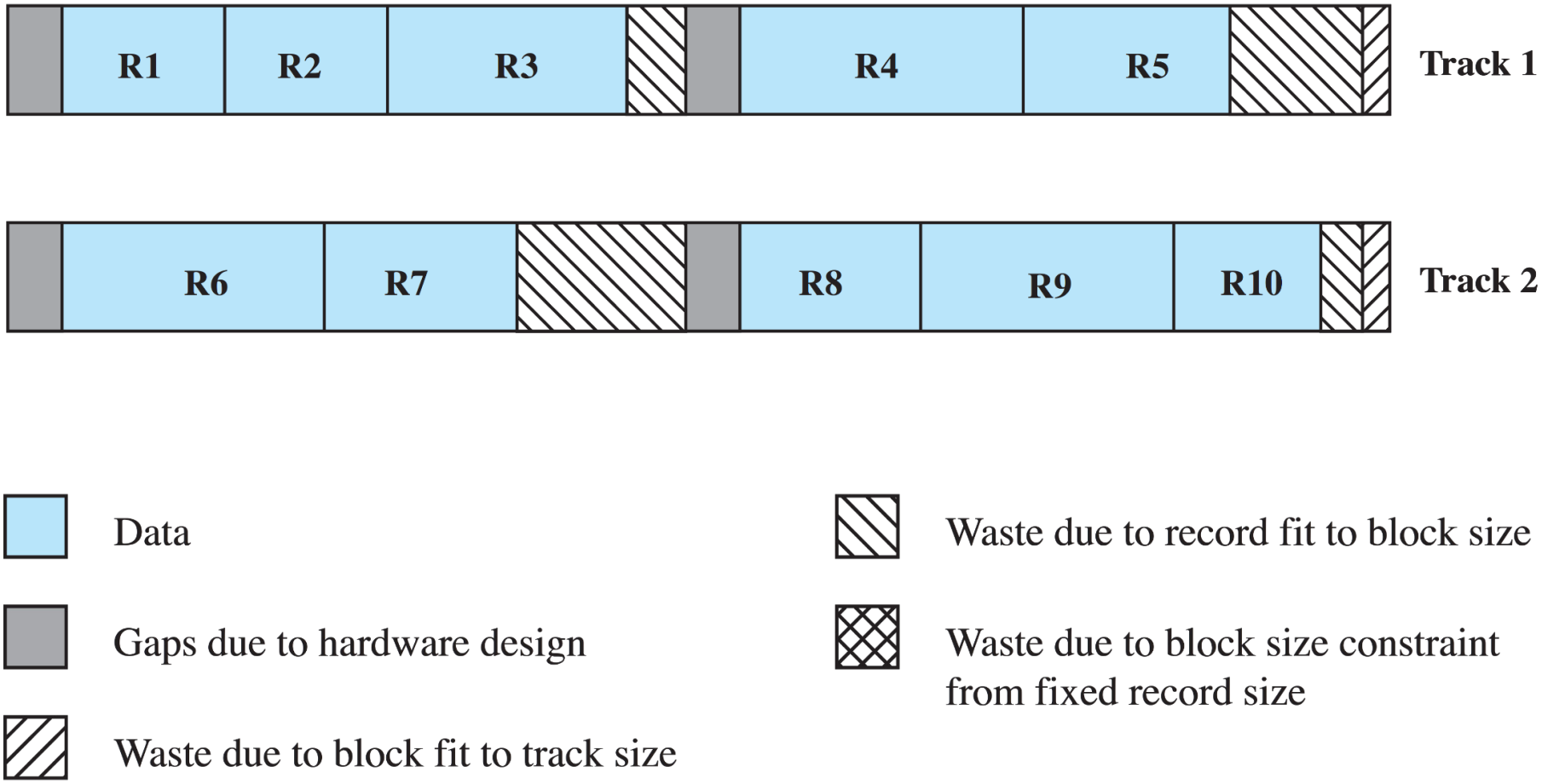
Gaps due to hardware design

Waste due to block fit to track size

Waste due to record fit to block size

Waste due to block size constraint from fixed record size

Variable Blocking: Unspanned



Secondary Storage Management

- **File Allocation**
  - On secondary storage, a file consists of a collection of **blocks**
  - The operating system or file management system is responsible for **allocating blocks to files**
  - The approach taken for file allocation may influence the approach taken for **free space management**
  - Space is allocated to a **file** as one or more **portions (contiguous set of allocated blocks)**
  - **File allocation table (FAT)**
    - data structure used to keep track of the portions assigned to a file

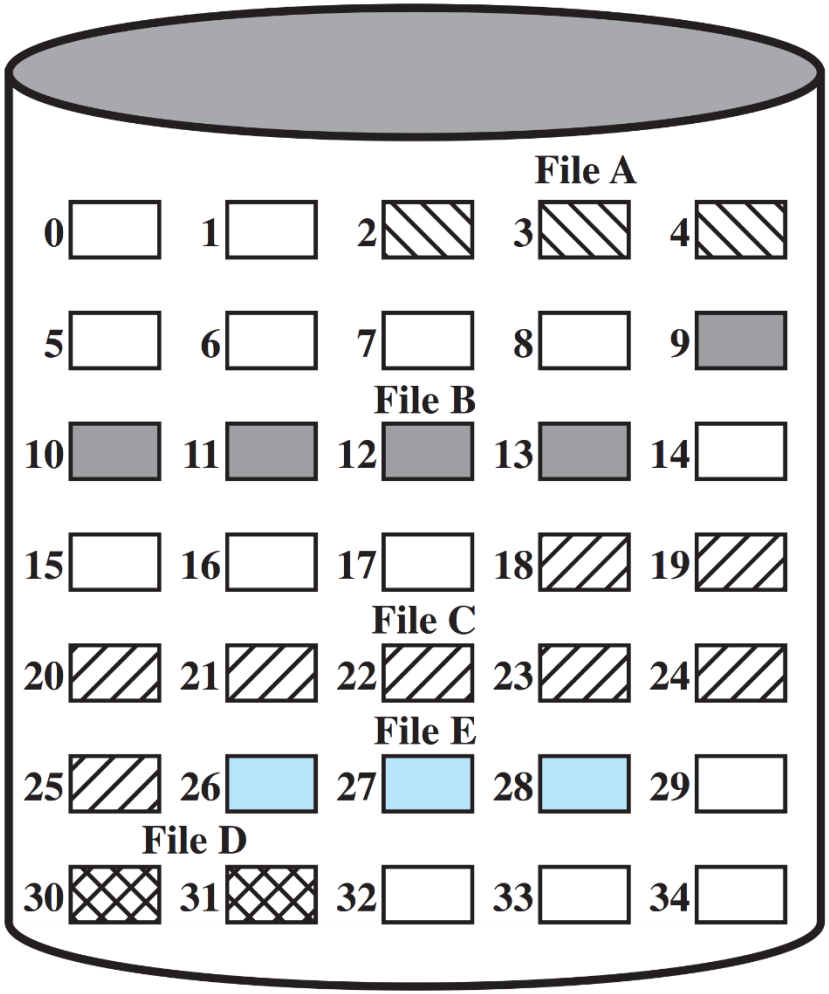
Issues of File Allocation

- **When a new file is created, is the maximum space required for the file allocated at once?**
  - Preallocation vs. Dynamic Allocation
- **What size of portion should be used for file allocation?**
  - Variable, large contiguous portions
  - Blocks
- **What sort of data structure or table is used to keep track of the portions assigned to a file?**
  - DOS： file allocation table (FAT)

File Allocation

- **Preallocation**
  - A preallocation policy requires that **the maximum size of a file** be declared at the time of the file creation request
  - For many applications it is **difficult to estimate** reliably the maximum potential size of the file
    - **tends to be wasteful** because users and application programmers tend to overestimate size
- **Dynamic Allocation**
  - Dynamic allocation allocates space to a file in portions as needed

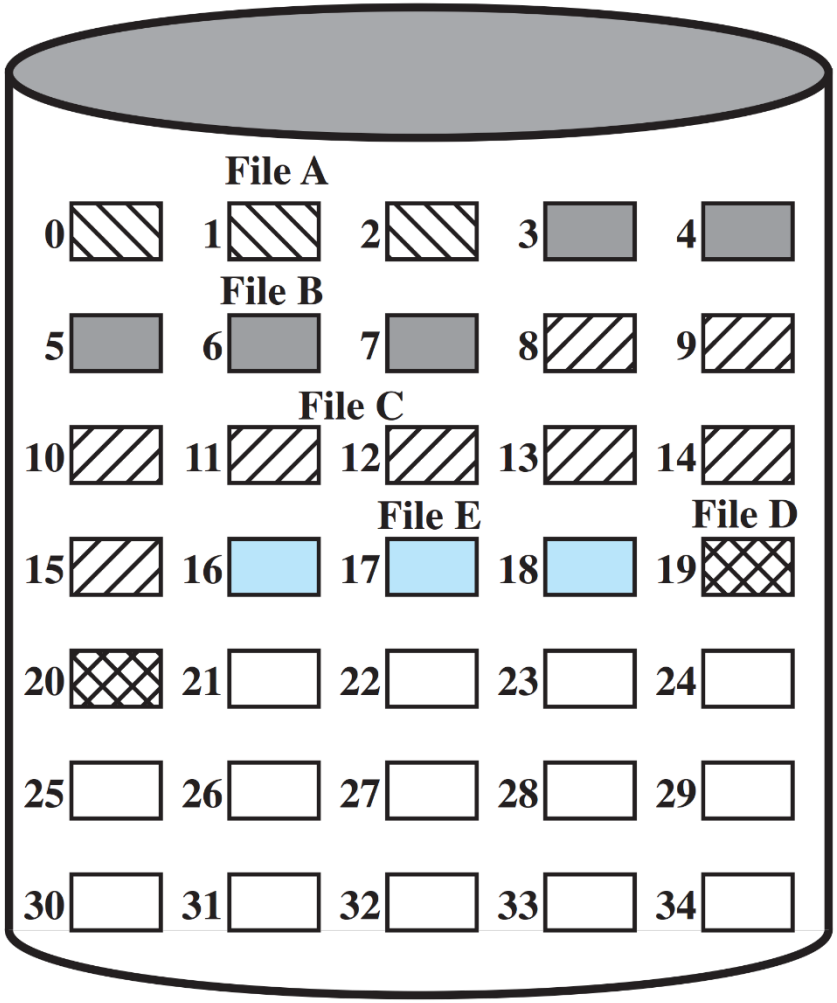
Contiguous File Allocation



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

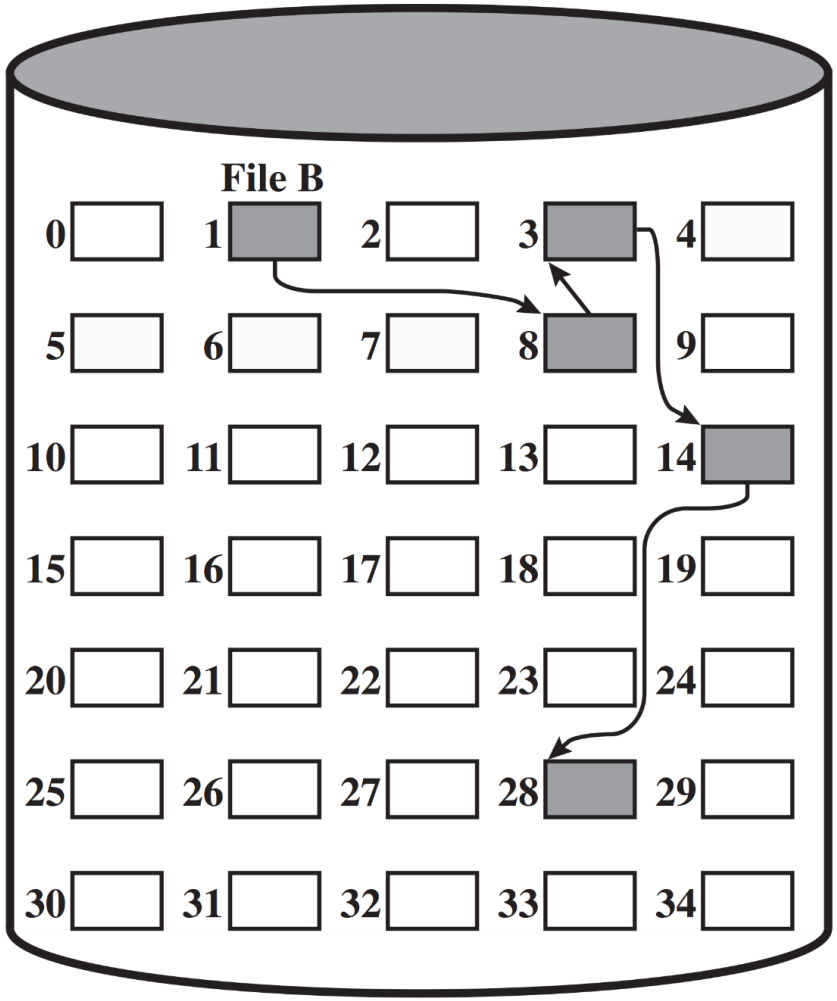
After Compaction



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

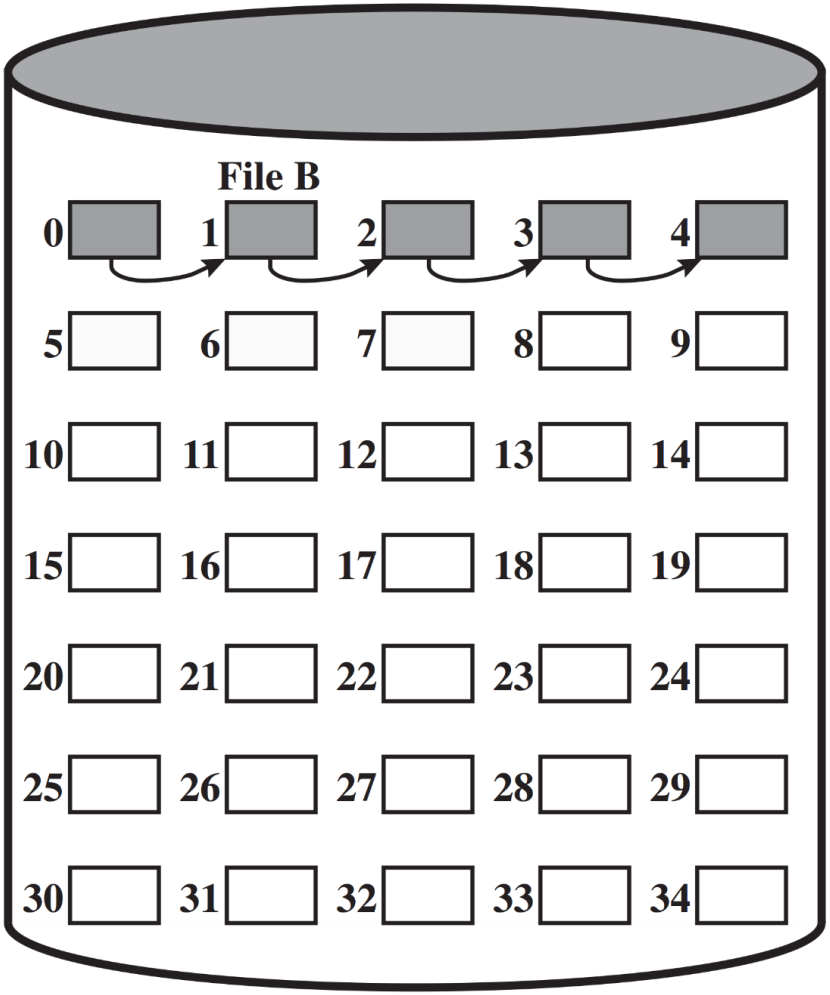
Chained Allocation



File Allocation Table

File Name	Start Block	Length
• • •	• • •	• • •
File B	1	5
• • •	• • •	• • •

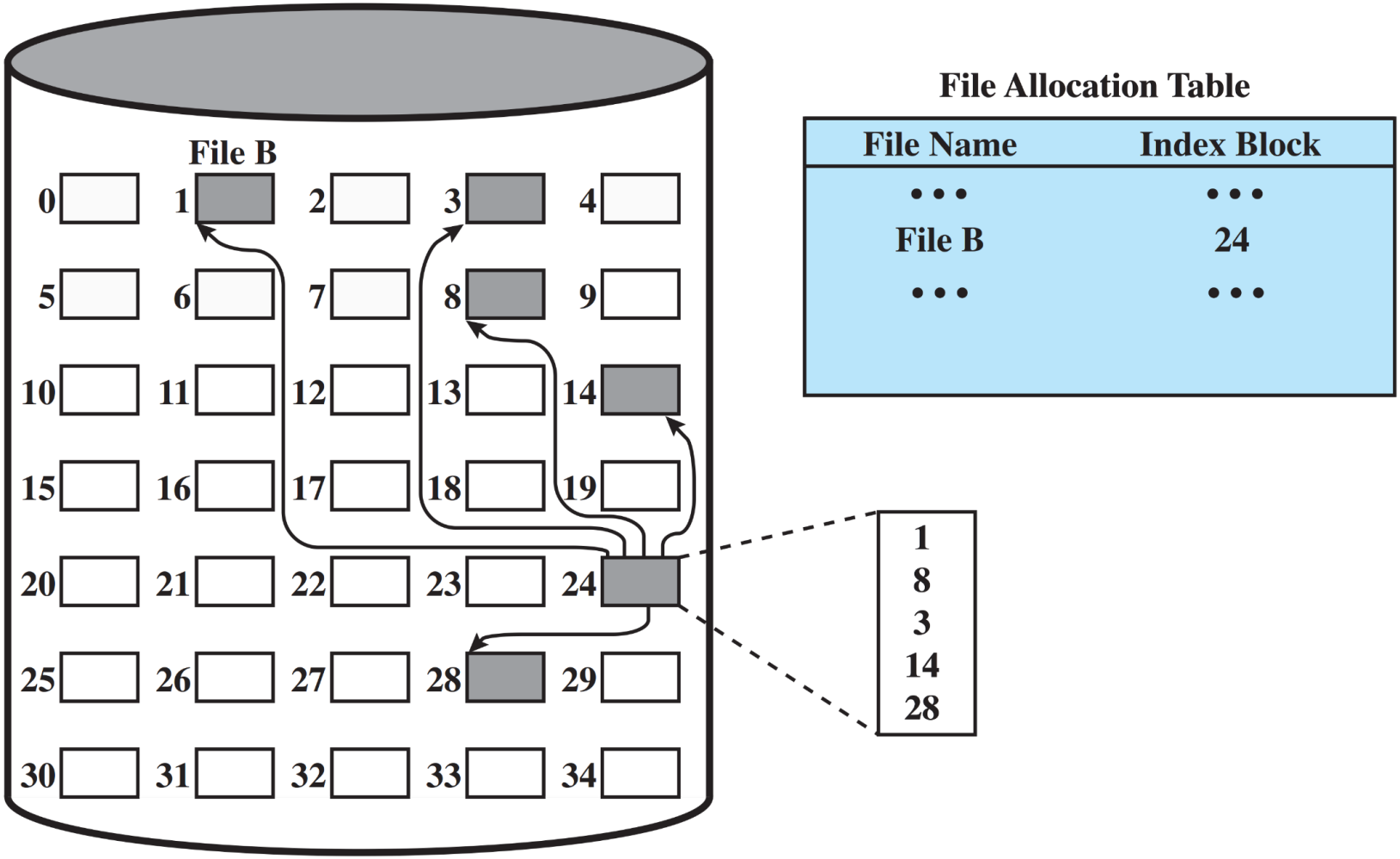
After Consolidation



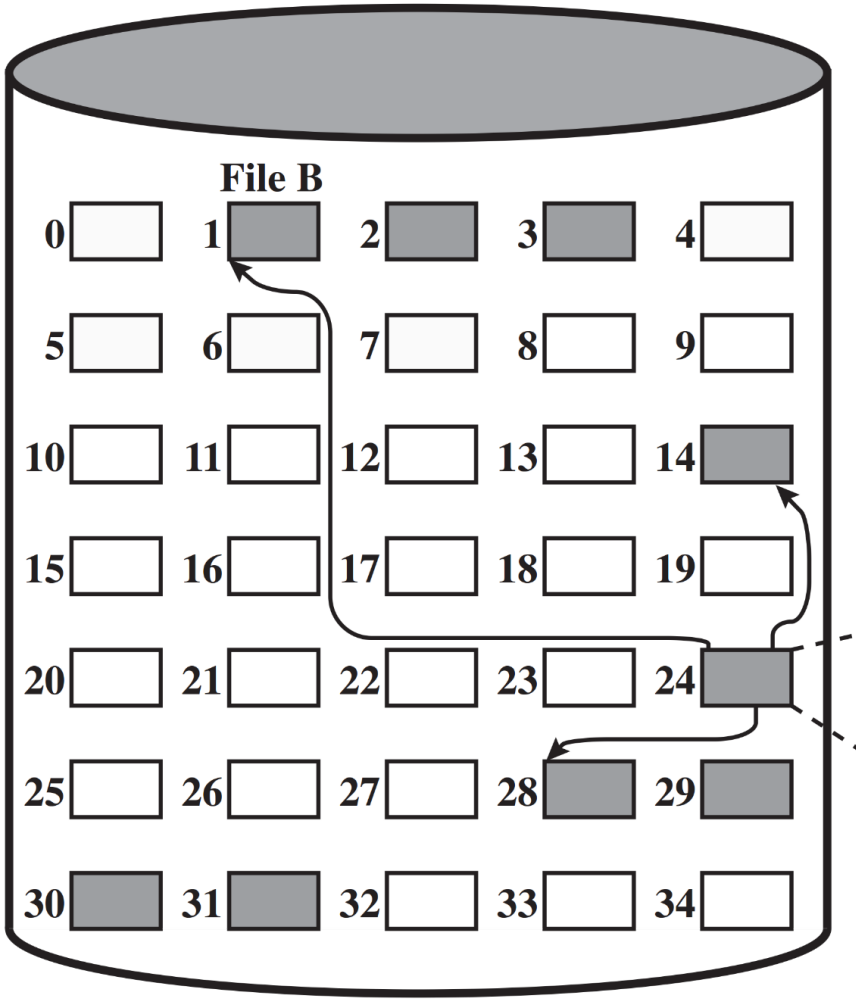
File Allocation Table

File Name	Start Block	Length
•••	•••	•••
File B	0	5
•••	•••	•••

Indexed Allocation with Block Portions



Indexed Allocation with Variable Length Portions



File Allocation Table

File Name	Index Block
• • •	• • •
File B	24
• • •	• • •

Start Block	Length
1	3
28	4
14	1

Free Space Management

- **Just as allocated space must be managed, so must the unallocated space**
- **To perform file allocation, it is necessary to know which blocks are available**
- **A disk allocation table (DAT) is needed in addition to a file allocation table**
  - Bit Tables
  - Chained Free Portions
  - Indexing
  - Free Block List

Summary

- **A file management system:**
  - is a set of system software that provides services to users and applications in the use of files
  - is typically viewed as a system service that is served by the operating system
- **Files:**
  - consist of a collection of records
  - if a file is primarily to be processed as a whole, a sequential file organization is the simplest and most appropriate
  - if sequential access is needed but random access to individual file is also desired, an indexed sequential file may give the best performance

Summary

- if access to the file is principally at random, then an indexed file or hashed file may be the most appropriate
- directory service allows files to be organized in a hierarchical fashion

- **Some sort of blocking strategy is needed**
- **Key function of file management scheme is the management of disk space**
  - strategy for allocating disk blocks to a file
  - maintaining a disk allocation table indicating which blocks are free