Competency 5

Uses operating systems to manage the functionality of computers

5.2 Explores how an operating system manages directories/folders and files in computers.

Time: 6 periods

Learning Outcomes

- Describes files and directories
- Briefly describes the need of disk formatting
- Identifies the need for file types
- Lists attributes of file and directories
- Describes the structure of the file systems
- Illustrates the organization of files and directories
- Briefly describes the methods used in file security
- Briefly describes how an operating system manages file security
- Briefly explains
    - Contiguous allocation
    - Linked allocation and
    - Indexed allocation
- Describes defragmentation and explains how it occurs

Files and directories

Files

- File is a named collection of data stored in secondary storage.

- There are different types of files such as text files, data files, directory files, binary and graphic files, and these different types of files store different types of information.

- Two views of a file.
    - Logical view (User's view) – How the users see the file? Collection of records (database file), Sequence of characters (text file), collection of light intensity values (image files).
    - Physical view (OS's view) – How the file is stored on the secondary storage? Blocks of secondary storage space either allocated or free?

File Attributes

Each file has an associated collection of information (attributes)

- File name
- Type (e.g., source, data, executable)
- Owner
- Location(s) on the secondary storage.
- Organization (e.g. sequential, indexed, random)
- Access permissions – who is permitted to read/write/delete data in the file.
- Time and date of creation, modification, last access
- File size

File Type

One of the possible implementation techniques of file type is to include the type as an extension to the file name. There are many different file formats as there are different programs to process the files.

File can be classified into various types based on the content.

- Executable(.exe)
- Text(.txt, .docx, …etc)
- Image(.bmp, .png, .jpeg, …etc)
- Video (.vob, .flv, .swf,…etc)
- Audio (.wav, .mp3,…etc)
- Compressed(.rar, .zip,…etc)

Naming Files

- The files, containing information, should have proper names to identity for later use.
- When you name a file consider following guidance:
    - The file name should contain only letters of alphabet (Character) and some special character.
    - The name should not be longer than 8 characters.
    - Extension of file is optional with a maximum of 3 characters.
- The files should not be named arbitrarily. It is advisable to have so meaningful name so that it represents the contents inside the file.
- The following extension are specifically used:
    - Com/exe: Command or executable file.
    - BAT: Batch files
    - BAK: Backup files of some other files.

Directories

A directory is special type of file that contains other files. It is also called as folder. A group of files can be stored in a directory.

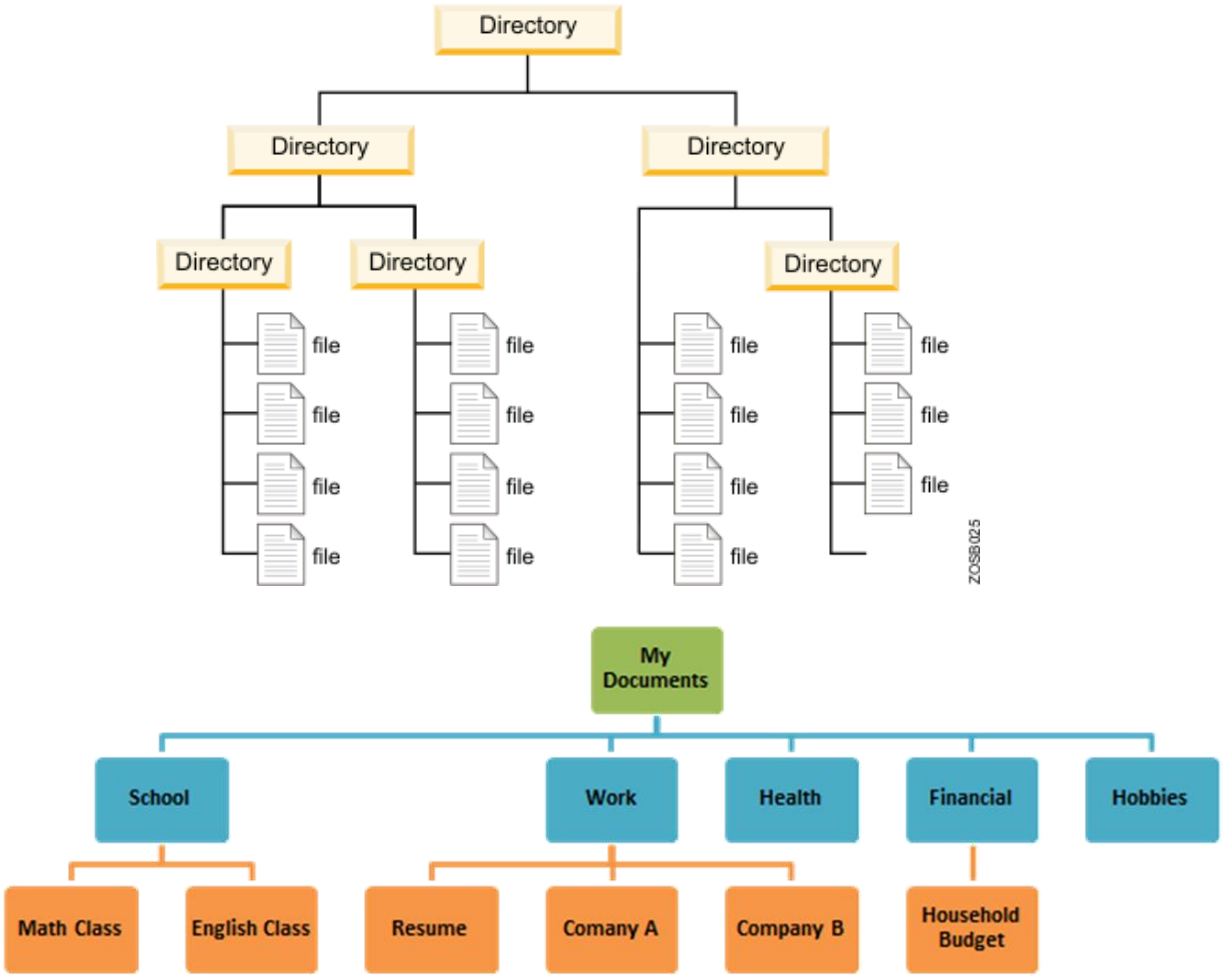
File Management System

- File management system is a part of OS.
- It provides services to users and applications to use files.
- It creates the illusion that there are files and directories being stored in the secondary storage.

File Organization

File organization is the way you organize your folders on your computer.

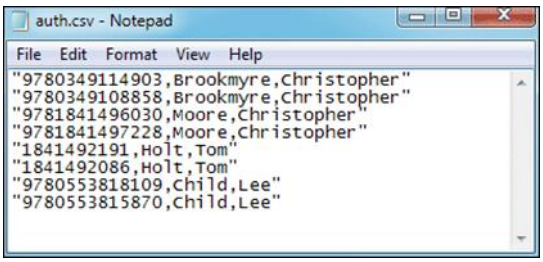
Keeping all your documents under one folder is a great way to keep organized and make things easy for backing up your computer.



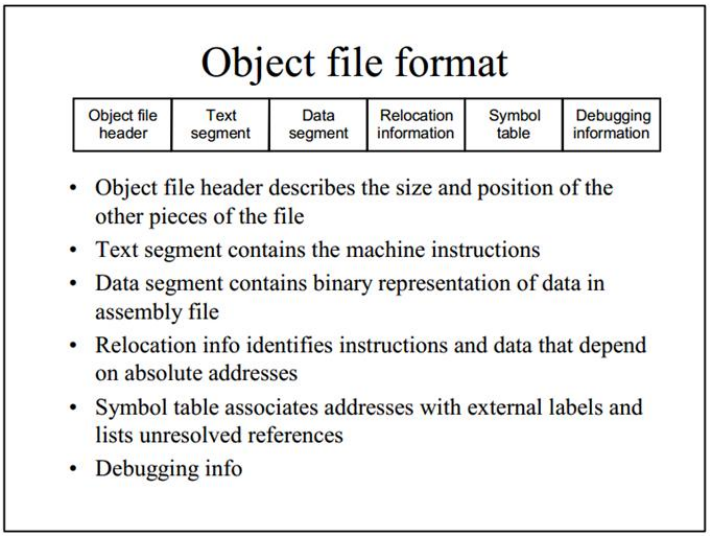
File Structure

A file has a certain defined structure according to its type.

- A text file is a sequence of characters organized into lines.



- An object file is a sequence of bytes organized into blocks that are understandable by the machine.



File Systems

A file system is used to control how data is stored and retrieved.

File systems for Windows

Microsoft Windows employs two major file systems:

1. **FAT (File Allocation Table)**, which was inherited from old DOS and has exFAT as its later extension.

2. **NTFS (New Technology File System)**, the primary format most modern versions of this OS use by default, and

In addition, the **ReFS file system -Resilient File System** was developed by Microsoft as a new generation file system and it requires Windows 8 or Windows Server 2012.

- This file system aims to improve data integrity as well as reduce data corruption and downtime.
- Fixes problems while online to reduce downtime.

FAT File System (Around the 70's)

- FAT is the file systems introduced with Microsoft Disk Operating System (MS DOS).
- It supports only 8 characters long file name.
- Partition and file size can be up to 4 GB.
- Disk can get fragmented thus slowing down the access.
- It is not very reliable since it does not support bad cluster mapping.

- Windows uses 2 types of FAT file systems.

- **FAT32**
  - Offers no security/ compression/Encryption
  - Does not have modern features like self-healing and transaction logging
  - Less overhead than NTFS/ReFS
  - Compatible with everything -A hard disk formatted with FAT32 can be read in just about any OS.
  - Often use with removable media

- **exFAT(Extended File Allocation Table)**
  - Microsoft file systems (proprietary and patented)
    - Support in windows XP and above
  - Optimized for flash drives
  - Better large drive performance than FAT32

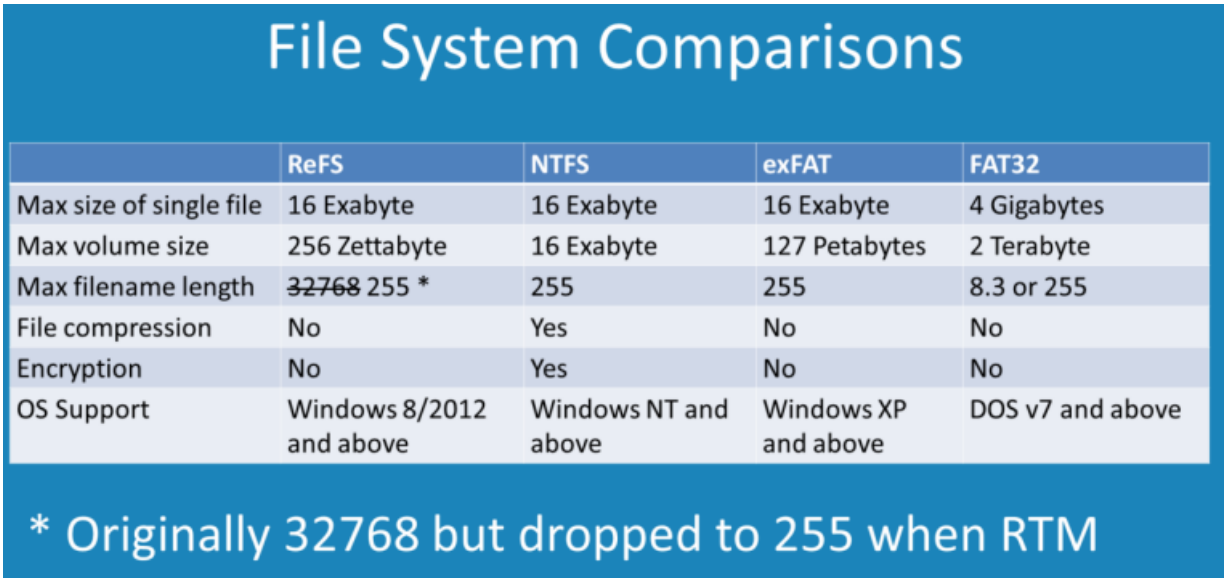
NTFS File System (Introduced in 1993)

- It **provides security for both local and remote users**. The security is provided to the level of files and folders.
- It supports 255 characters **long file name**.
- **Partition size** can be up to 16 Exabyte.
- It supports **file compression**.
- It provides **lesser possibility of fragmentation**.
- It is highly reliable since it supports bad cluster mapping and **transaction logging.**

NTFS is a proprietary file system developed by Microsoft. This is improvement of FAT. This improvement includes

- The capability to recover from some disk-related errors automatically, which FAT cannot.
- Support with Unicode encoding system
- Improved support for larger hard disks.
- Better security as permissions and encryptions are used to restrict access to specific files to approved users.

- ❖ FAT32 is still widely used for memory cards and USB sticks.
- ❖ The system is supported by smartphones, digital cameras and other portable devices.



File Security

What is File security?

- Protecting files from unauthorized user is referred to as file security.
- Commonly used file Security methods
  - Password
  - Access Privileges

Password

- On multi-access computer systems many different types of user are logged into a vast filing system.
- In order to gain access to the system a user is supplied with a user name or ID together with a personal password.
- Some files may be open for reading by anyone, others only by those users in the same department, and others only for reading by individuals, such as 'mail'.
- The main computer's operating system makes sure that file access is appropriate for the user's password.

Access Privileges

Group Name	Access Level
superuser	Full access to all system operations.
provisioning	Access to most network management configuration tasks. This level does not have access to some administrative tasks.
surveillance	Read-only access to monitor system operations.

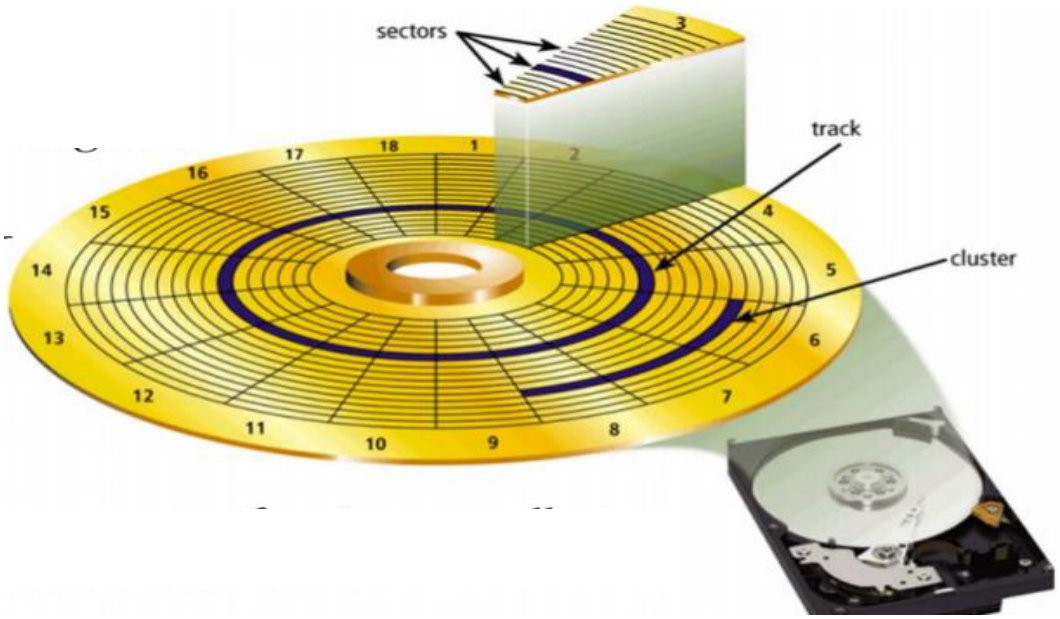
Authentication

- Authentication refers to identifying each user of the system and associating the executing programs with those users.

- It is the responsibility of the Operating System to create a protection system which ensures that a user who is running a particular program is authentic.
- Operating Systems generally identifies/ authenticates users using following ways:
  - ✓ **Username / Password** - User need to enter a registered username and password with Operating system to login into the system.
  - ✓ **User attribute - fingerprint/ eye retina pattern/ signature** - User need to pass his/her attribute via designated input device used by operating system to login into the system.

File Storage Management

- On secondary storage, stores data organized using :
  - Tracks – rings on platters
  - Sectors - tracks are divided into sectors
  - Clusters – a group of sectors, smallest unit allocated to store a file



- The Operating System's point of view a file is nothing but a collection of blocks of space on secondary storage.
- File management system of OS allocates blocks of space to files and also keeps the track of unallocated blocks on secondary storage.
- OS uses a data structure called File Allocation Table (FAT) to keep the track of allocated and free blocks.
- In this lesson describes alternatives for file allocation on a single disk.

File Allocation Methods

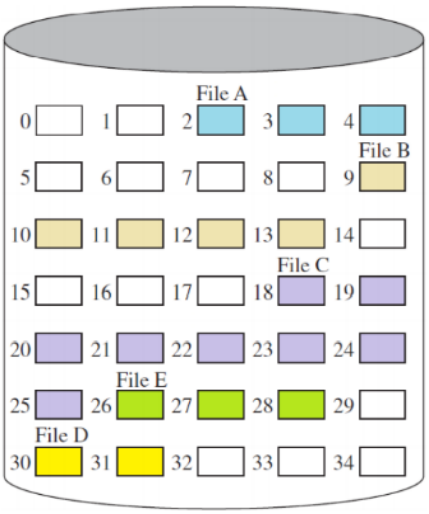
Three Methods are commonly use:

1. Contiguous Allocation
2. Linked Allocation
3. Indexed Allocation

Contiguous Allocation

- With contiguous allocation, a single contiguous set of blocks is allocated to a file at the time of file creation.
- FAT keeps a single entry for each file, showing the starting block and the length of file.

File	Starting block	Length
A	2	3
B	9	5
C	18	8
D	30	2
E	26	3



Features:

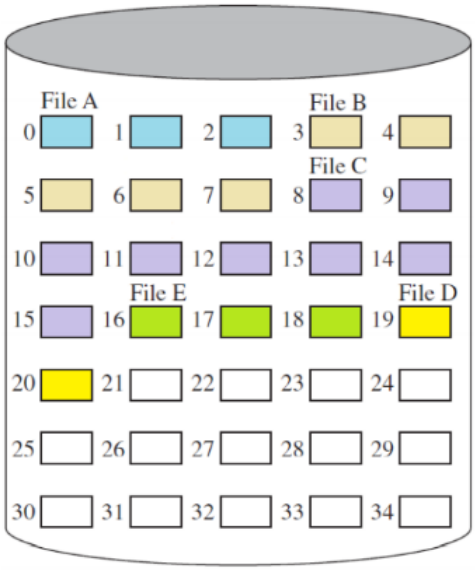
Strengths-

- Simple.
- Easy Access.

Weaknesses

- File size should know at the time of file creation.
- Extending file size is difficult (difficult to find contiguous blocks to accommodate growing size
- External fragmentation(unusable free blocks between allocation)
- Need periodic compaction/defragmentation (system overhead)
- After periodic compaction/defragmentation (after moving allocated contiguous blocks of files to a one end of secondary storage leaving free unallocated blocks at the other end.)

File	Starting block	Length
A	0	3
B	3	5
C	8	8
D	19	2
E	16	3

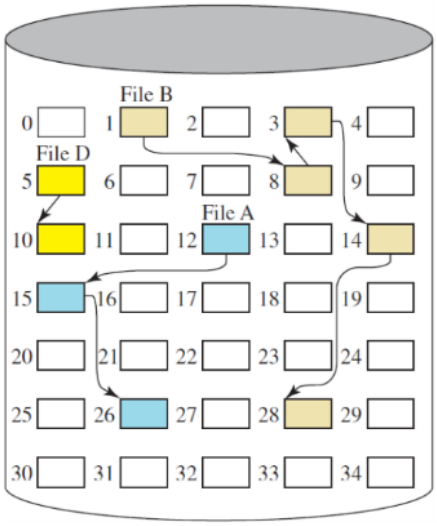


Linked /Chained Allocation

- At the opposite extreme from contiguous allocation is linked allocation. Inside each block, a link is maintained to point to where the next block of the file is.

- FAT keeps a single entry for each file, showing the starting block and the length of file.

File	Starting block	Length
A	12	3
B	1	5
-	-	-
D	5	2
-	-	-



Features

Strengths

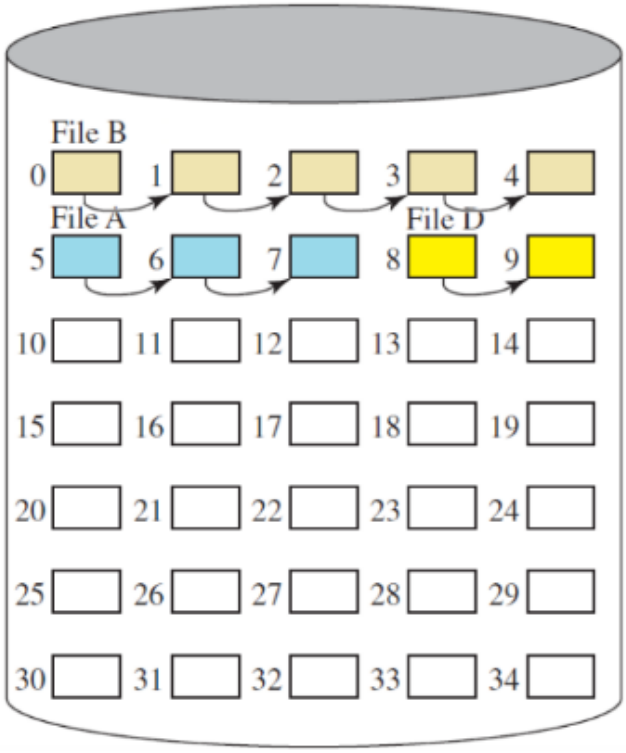
- No need to know the file size at the time of creation
- No external fragmentation to worry about because block by block is allocated at a time as needed.
- Files can grow dynamically and easily.

Weaknesses

- Many seeks are required to access file data (Seek time is the problem. Because reduce the performance)
- Periodic consolidation can solve this but adds system overhead.

After periodic consolidation (after rearrangement of blocks of each file contiguously to a one end of secondary storage leaving free unallocated blocks at the other end)

File	Starting block	Length
A	5	3
B	0	5
-	-	-
D	8	2
-	-	-



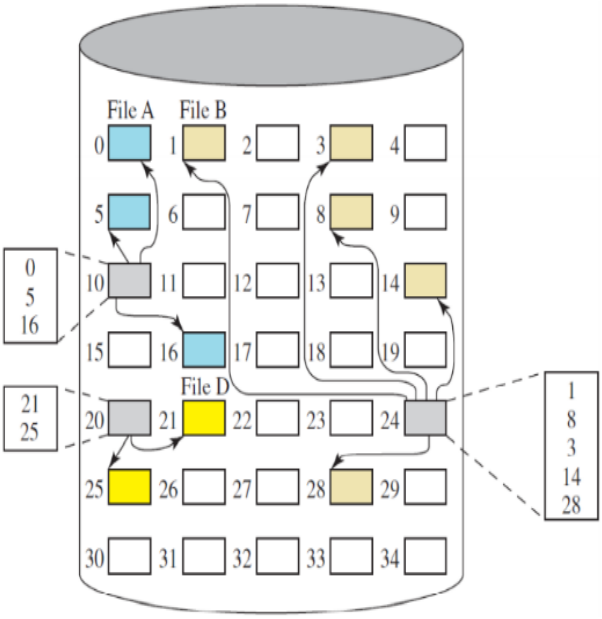
Example:

MSDOS FAT file system and early versions of Windows used Linked allocation method)

Indexed Allocation

- Indexed allocation is the most popular form of file allocation.
- The Indexed allocation addresses many of the problems of contiguous and linked allocation.
- FAT contains entry to index block of each file.
- The index block has entries for each block allocated to file.
- The index table is also saved in a block/s.
- Example: UNIX file system - used indexed allocation with contiguous blocks
- NTFS – used indexed allocation method

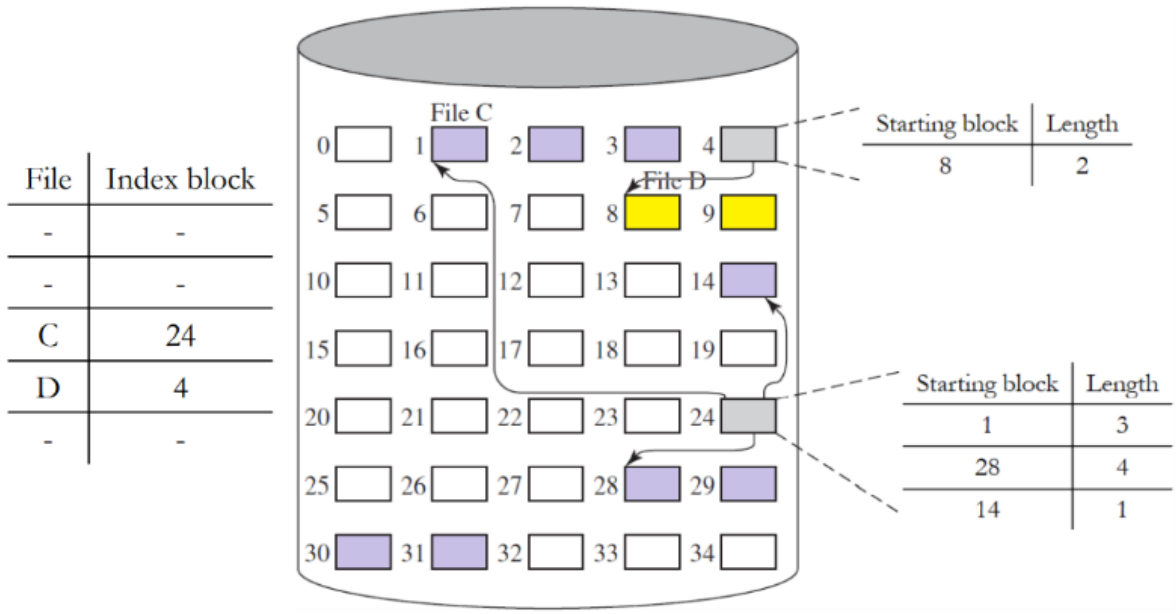
File	Index block
A	10
B	24
-	-
D	20
-	-



Features

- No need to know the file size at the time of creation.
- Files can grow dynamically and the index is updated accordingly.
- File ends at nil pointer
- Block by block allocation eliminates external fragmentation.
- Allocating contiguous blocks whenever possible reduces number of seek operations and improves performance.
- Periodic consolidation can reduce the index size and thereby index search time if contiguous blocks are allocated
- No external fragmentation
- Each block contains pointer to next block
- No compaction, external fragmentation

- Index allocation with contiguous blocks



The Secondary storage

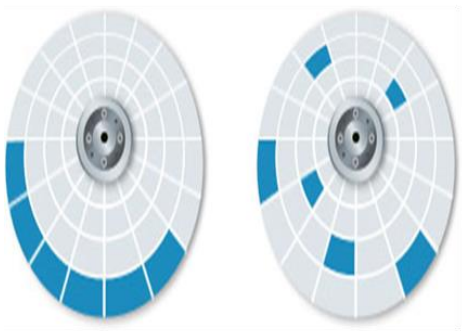
- Secondary storage is the non-volatile repository for both user and system data and programs.
- Secondary storage is typically used to store
  - Source program
  - Executable programs
  - Data for the program
  - Temporary data

Maintenance of Secondary storage

- Defragmentation
- Disk Cleanup

Disk Fragmentation

- A hard drive has a number of sectors on it, each of which can contain data.
- Files, particularly large ones, must be stored across a number of different sectors.
- Let's say you save a number of different files on your disk. Each of these files will be stored in contiguous sectors on a clean disk.
- Later, if you update one of the files you originally saved, increasing the file's size, the file system will attempt to store the new parts of the file right next to the original parts.
- Unfortunately, if there's not enough uninterrupted space, the file must be split into multiple parts.
- When your hard disk reads the data, its heads must skip around between different physical locations on the hard drive to read each extent, and this can have an impact on performance, especially if you have large files with many extents.

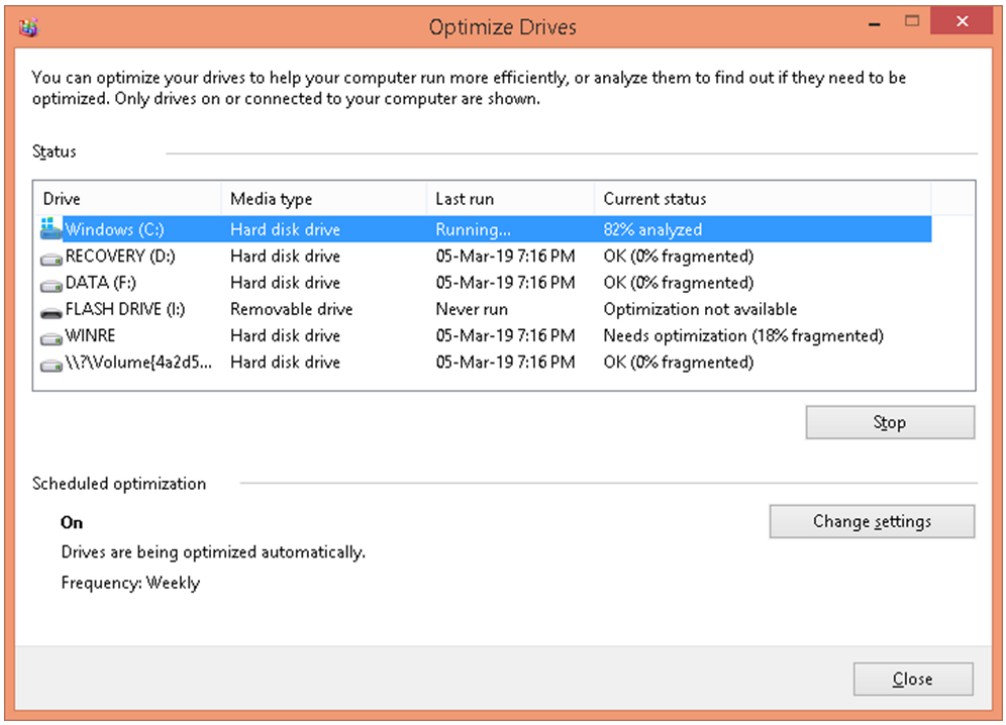


Defragmentation

- Defragmentation is a process that reduces the amount of fragmentation.
- Performing disk defragmentation (or "defragging") will reorganize the data on a hard disk, allowing it to operate more efficiently.
- File systems like NTFS have some built in buffers of free space that can reduce fragmentation issues, but files can and do still become fragmented.

- Starting with Windows 2008, Microsoft has included an automatic weekly Defrag in its scheduler, but it needs to be enabled manually:

Windows 8.1



Disk Cleanup

- Disk Cleanup is a Microsoft software utility first introduced with Windows 98 and included in all subsequent releases of Windows.
- It allows users to remove files that are no longer needed or that can be safely deleted.
- Removing unnecessary files, including temporary internet files (associated with Internet Explorer), downloaded program files, and offline webpages.
- Disk Cleanup also allows you to empty the Recycle Bin, delete temporary files, and delete thumbnails
- Can help to speed up and improve the performance of the hard drive and computer.
- Running Disk Cleanup at least once a month is an excellent maintenance task and frequency.

Disk formatting

- Formatting is the process of preparing a data storage device for initial use which may also create one or more new file systems.
- The first part of the formatting process that performs basic medium preparation is often referred to as "low-level formatting".
- Partitioning is the common term for the second part of the process, making the data storage device visible to an operating system.
- The third part of the process, usually termed "high-level formatting" most often refers to the process of generating a new file system.

Recovery of data from a formatted disk

- As file deletion is done by the operating system, data on a disk are not fully erased during every high-level format. Instead, links to the files are deleted and the area on the disk containing the data is retains until it is overwritten.

Answer the following Questions

1. What is a file?
2. What are the file attributes?
3. What are the file types?
4. What are the file operations?
5. What are the things you have to keep in mind when naming a file?
6. What is a directory?
7. What is a file structure? Describe using a diagram.
8. What is a file system?
9. What are the major file systems used in Windows operating system? Compare above listed file systems.
10. What is file security?
11. What are the commonly used file security methods?
12. Briefly explain above listed file security methods?
13. What are the commonly used file allocation methods? Briefly explain them.
14. Hard disk of your computer is do a major task. The operating system used methods to make it best performance. What are they? Briefly explain.

References

- Teachers Guide – 2017
- Chapter 3 and 9, W.Stallings, Operating System: Internals and Design Principles, 9th Edition, Pearson, 2017
- https://www.lifewire.com/what-is-a-file-2625878
- https://www.techopedia.com/definition/7199/file
- https://www.computerhope.com/jargon/f/file.htm
- http://osr507doc.xinuos.com/en/OSUserG/_Working_with_directories.html
- https://www.ufsexplorer.com/articles/file-systems-basics.php
- https://blog.wsol.com/utilities-for-handling-disk-fragmentation-in-linux