

FILE TYPES

Types of files - Confusion: Text files vs. Binary Files

- Text files (ASCII / UNICODE)
 - Bytes of data are organised as characters from respective character sets
- Binary files
 - Data in a specific format that requires interpretation.
- Text files vs. Binary Files
 - All files are in Binary
 - Text Files are formatted in chunks of 8 bits or 16 bits
 - Files in any other format are Binary Files

File Types

- Most files contain a specific types of information
 - A Java program
 - A JPEG image
 - A BITMAP image
 - An MP3 clips
- The kind of information is the file type
 - So, the File System knows which operations it can do
 - Most OS have associations between file types and applications

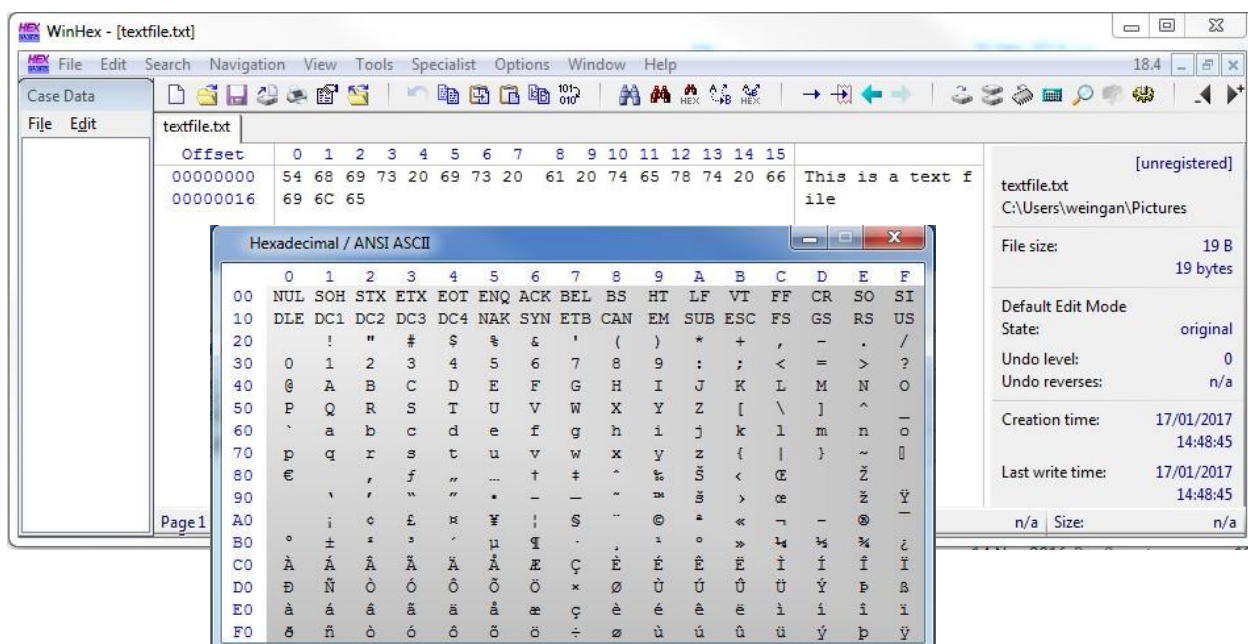
File Types Extensions

- File names are often separated by a full-stop into 2 parts
 - Main name
 - File extension
- The file extension was used by the OS to identify the type of file – But is not necessarily the actual file type

- Windows 10 will inspect the file to ascertain the actual file type – Looking at the file header

Extension	File Type
.txt	Text data file
.mp3, .au, .wav	Audio file
.gif , .tiff , .jpg	Image file
.doc , .odt	Word processing files
.java , .sql	Programming source file

Anatomy of an ASCII File



Revisiting BMP files



File Signature

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
00000000	42	4D	66	37	02	00	00	00	00	00	36	00	00	00	28	00
00000016	00	00	DC	00	00	00	DC	00	00	00	01	00	18	00	00	00
00000032	00	00	30	37	02	00	25	16	00	00	25	16	00	00	00	00
00000048	00	00	00	00	00	00	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000064	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000080	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000096	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000112	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000128	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000144	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000160	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000176	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000192	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000208	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000224	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000240	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
00000256	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF

220px-SNice.bmp
File size: 142 KB
145,254 bytes
DOS name: 220PX~1.BMP
Default Edit Mode: original
Undo level: 0
Undo reverses: n/a
Creation time: 12/02/2020 10:16:13
Last write time: 12/02/2020

Anatomy of a Binary file (jpeg) File Headers

Anatomy of a Binary file (jpeg) File Headers

(JFIF)
JPEG File Interchange Format

Partial File Header

FF D8 = Start of image marker

FF E0 = JFIF marker

JFIF Identifier

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
00000000	FF	D8	FF	E0	00	10	4A	46	49	46	00	01	01	01	00	78
00000016	00	78	00	00	FF	EE	00	0E	41	6A	6F	62	65	00	64	00
00000032	00	00	00	00	FF	E2	FF	FF	49	43	43	5F	50	52	4F	46
00000048	49	4C	45	00	09	00	08	80	70	41	44	42	45	02	10	
00000064	00	00	70	72	74	72	43	4D	59	48	4C	61	62	20	07	D0
00000080	1A	00	05	00	29	00	35	61	63	73	70	41	50			
00000096	00	00	00	41	44	42	45	00	00	00	00	00	00	00	00	00
00000112	00	00	00	00	00	00	01	00	00	F6	D5	00	00	00	00	00
00000128	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000144	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000160	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000176	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000192	00	FC	00	00	74	63	70	72	74	00	00	01	70	00	00	00
00000208	00	2B	77	74	70	74	00	00	01	9C	00	00	00	14	41	32

64px-A_Smiley.jpg
File size: 549 KB
561,997 bytes
64PX-A~1.JPG
Creation time: 17/01/2017 15:22:27
Last write time: 17/01/2017

File Signatures

- There file signature databases – Filesignatures.net
- Wikipedia often has high quality listings of the entire file header

File Operations

- Create a file ■ Delete a file ■ Open a file ■ Close a file ■ Read data from a file
- Write data to a file
- Reposition the current file pointer in a file
- Append data to the end of a file
- Truncate a file – ie. delete all or part of it
- Rename a file
- Copy a file

File Protection

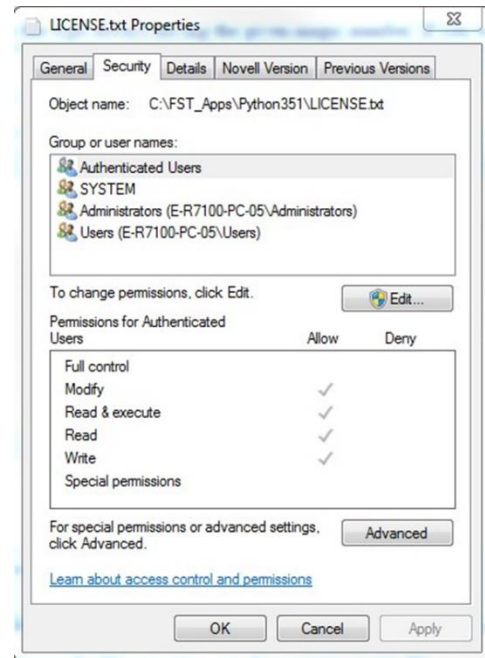
- Multi-user Systems
- Access control – Controls who can access files
- Who can read
- Who can write
- Who can execute

Cyber-Security Triad

- Three dimensions of cyber-security:
 1. Confidentiality:
 - Preventing access
 - Keep the bad-guy out
 2. Accessibility:
 - Ensure access
 - Make sure the good-guy can access the data
 3. Integrity:
 - Keep control of any changes made to the data
 - Who can change it
 - Keep track of any changes

File Permissions (Windows)

- NTFS:
 - Access Control Lists (ACL's)
- Each file has list of user identities with permissions
 - Explorer
- File Permissions
- Security
 - Different user, different permissions
- No multi-user security for FAT32



Windows Permissions Classifications

- Full control
 - File can be written to/read from
 - Permissions can be read and modified
 - Ownership can be changed
 - Folder can be listed and entries deleted
- Modify
 - Same as Full control
 - But cannot change permissions or ownership

Windows Permissions Classifications

- Read/Execute
 - File can be read or executed as a program
 - Folder can be listed and traversed
- Read
 - File can be read
- But not executed
 - Folder can be listed
- But not traversed

- Write
 - File can be modified
 - Files/subfolders can be created in a folder
- But NOT deleted
- List folder contents (for folders only)
 - Same as Read/Execute, but not available for files, and only inherited by folders

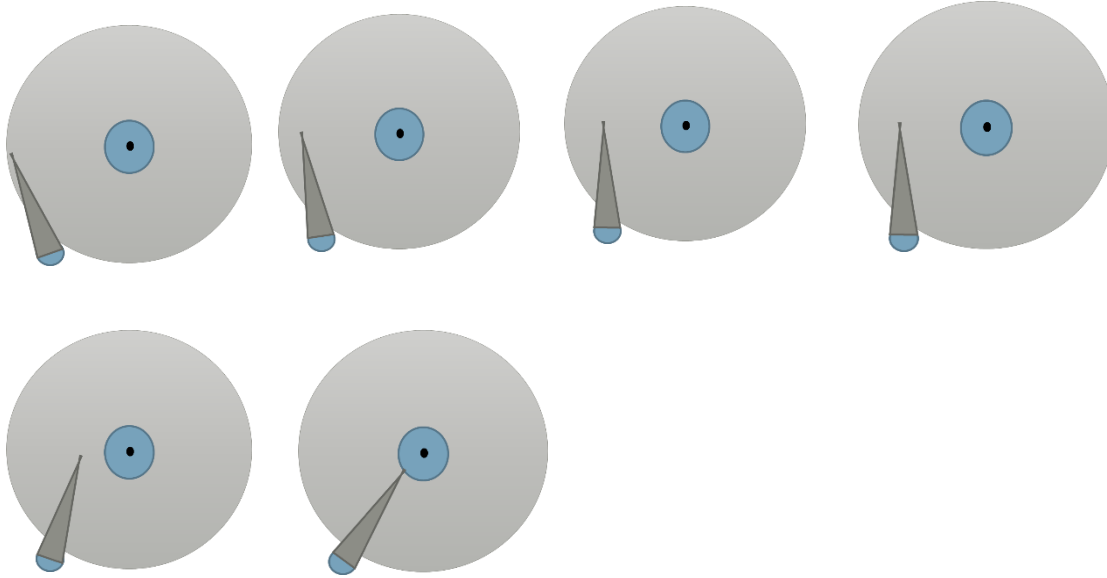
Security Inheritance

- Windows:
 - New file or subfolder created, will inherit it's parent's permissions by default
 - You can override
- Unix:
 - Permissions are not inherited for newly created files
 - Based on user's umask
 - Mask of permissions specific to that user – octal absolute format

DISK SCHEDULING

- Must be efficient
- Multiple processes, multiple requests to access disk
- Disk scheduling techniques to manage request:
 - First-come, first-served (FCFS): Requests are serviced in the order they arrive, irrespective of positions of heads
 - Shortest-Seek-Time-first (SSTF): Minimise movement of disk heads
 - SCAN: Disk heads continuously move in and out, servicing requests as the locations are found.
 - C-SCAN: Circular scan
 - Look: Like SCAN, but does not scan all the way to edge
 - C-Look: Like C-SCAN

Remember Anatomy of a Disk



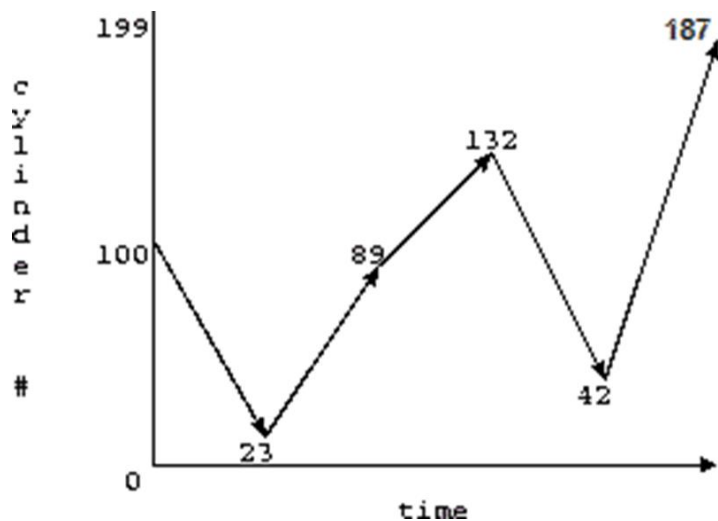
Example:

- Work queue: 23, 89, 132, 42, 187
- There are 200 cylinders 0-199
- The disk head starts at number 100

First-Come-First-Served:

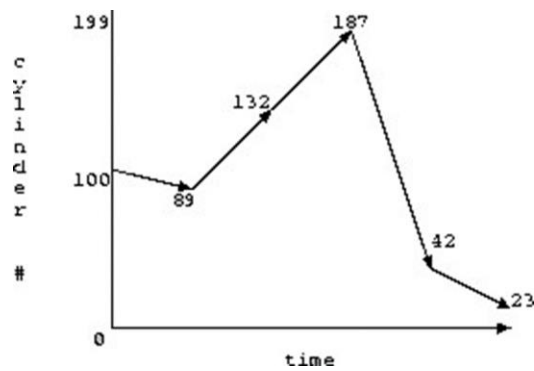
- Work queue: 23, 89, 132, 42, 187
- Total seek length:

$$23 - 100 + 89 - 23 + 132 - 89 + 132 - 89 + 42 - 132 + 187 - 42 = 421$$



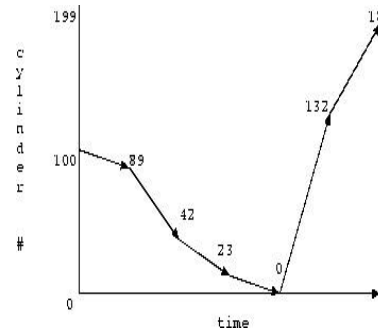
Shortest-Seek-Time-First:

- Work queue: 23, 89, 132, 42, 187
- Can be inefficient
 - Multiple changing directions
 - Starvation



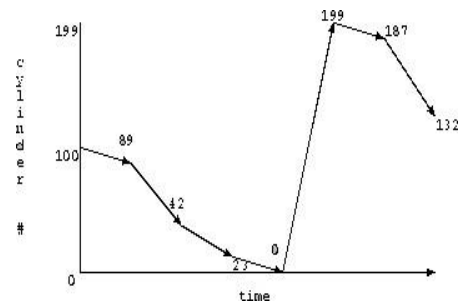
SCAN:

- Work queue: 23, 89, 132, 42, 187
- Elevator
- Sweeps the disk, to-and-fro
- LOOK is similar



C-SCAN:

- Work queue: 23, 89, 132, 42, 187
- Elevator
- Sweeps the disk, but one-direction



Performance:

- Depends on number of requests
- SCAN & C-SCAN are good for systems that place a heavy load on the disk, less likely to cause starvation
- Default: SSTF or LOOK
 - PRIORITY