COMPUTER SYSTEMS FUNDAMENTALS (4COSC009C)

Contact details

- Module Leader:

FILE SYSTEMS

This video:

- File types
- File operations
- Unix file systems
  - *File permissions*
  - *Access control*
  - *Indexing*
- Disk scheduling

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

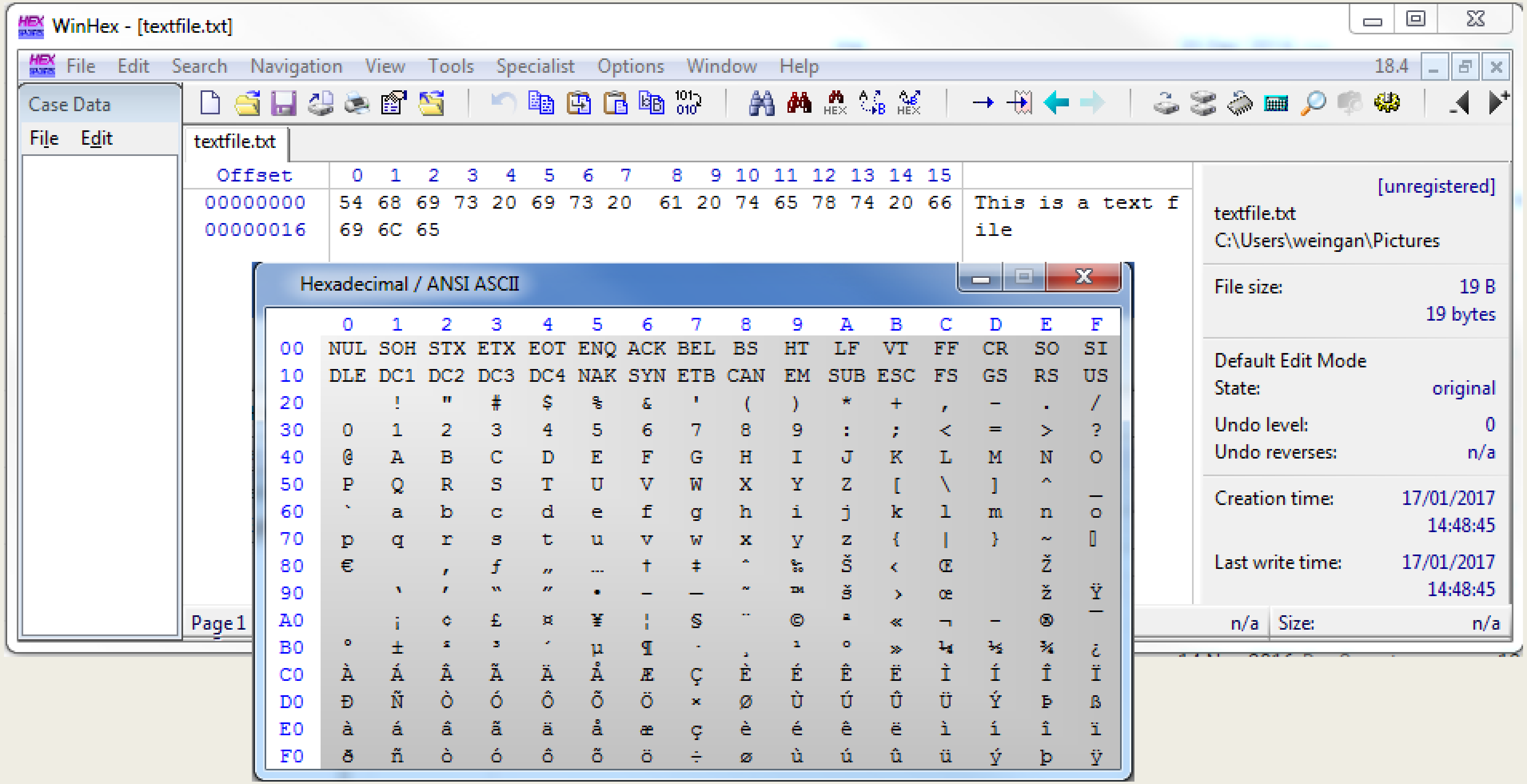
- Most files contain a specific types of information
  - *A Java program*
  - *A JPEG image*
  - *A BITMAP image*
  - *An MP3 clip*
- The kind of information is the **file type**
  - *So the File System knows which operations it can do*
  - *Most OS have associations between file types and applications*

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


WinHex - [textfile.txt]
File Edit Search Navigation View Tools Specialist Options Window Help
18.4
Case Data
File Edit
textfile.txt
Page 1

```
  Offset    0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000   54 68 69 73 20 69 73 20  61 20 74 65 78 74 20 66   This is a text f
00000016   69 6C 65                                           ile
```

[unregistered]
textfile.txt
C:\Users\weingan\Pictures

File size: 19 B
19 bytes

Default Edit Mode
State: original

Undo level: 0
Undo reverses: n/a

Creation time: 17/01/2017 14:48:45

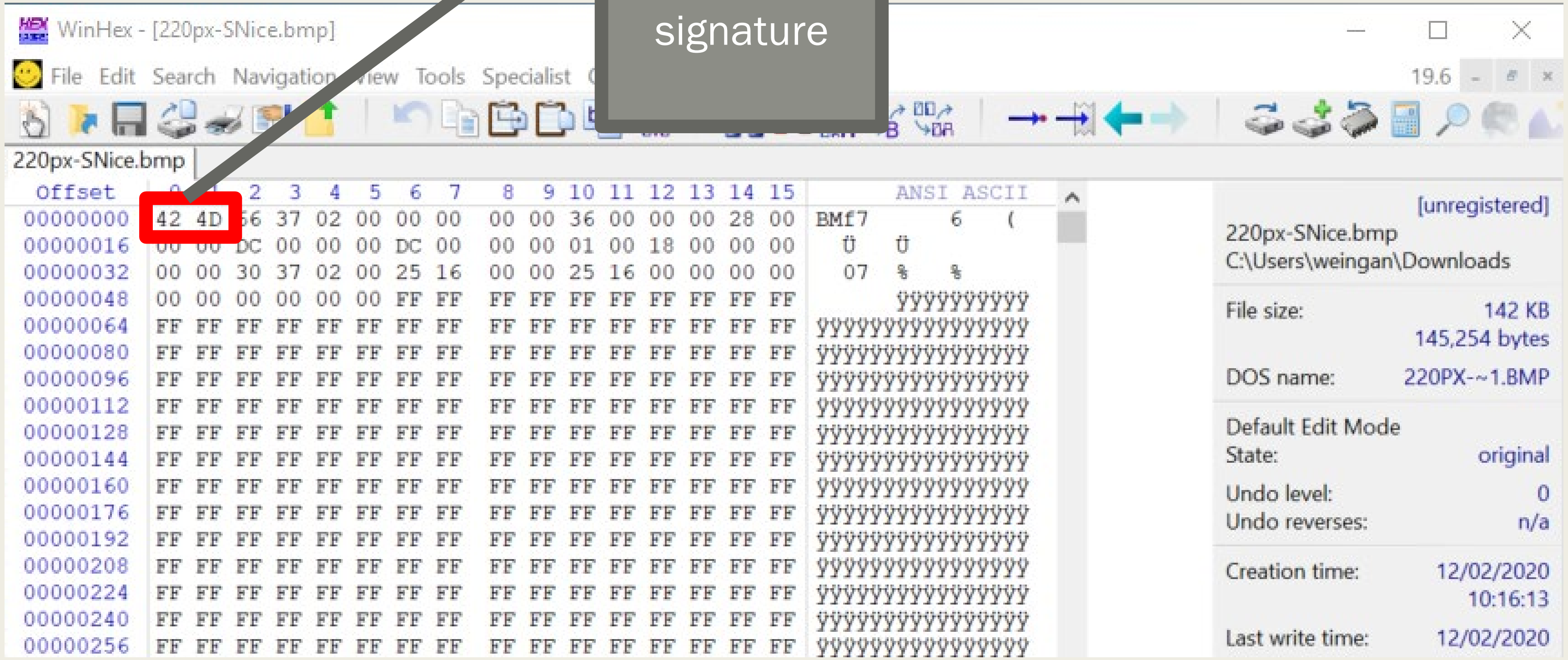
Last write time: 17/01/2017 14:48:45

n/a Size: n/a

Hexadecimal / ANSI ASCII

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
10	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
20		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
30	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
40	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
50	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
60	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
70	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	▯
80	€		‚	ƒ	„	…	†	‡	ˆ	‰	Š	‹	Œ		Ž	
90		‘	’	“	”	•	–	—	˜	™	š	›	œ		ž	Ÿ
A0		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
B0	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
C0	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
D0	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
E0	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
F0	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Revisiting BMP files



WinHex - [220px-SNice.bmp]
File Edit Search Navigation View Tools Specialist
19.6
220px-SNice.bmp

File signature

```
  Offset    0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15   ANSI ASCII
00000000   42 4D 66 37 02 00 00 00  00 00 36 00 00 00 28 00  BMf7      6   (
00000016   00 00 DC 00 00 00 DC 00  00 00 01 00 18 00 00 00    Ü   Ü
00000032   00 00 30 37 02 00 25 16  00 00 25 16 00 00 00 00    07  %   %
00000048   00 00 00 00 00 00 FF FF  FF FF FF FF FF FF FF FF        ÿÿÿÿÿÿÿÿÿÿ
00000064   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000080   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000096   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000112   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000128   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000144   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000160   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000176   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000192   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000208   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000224   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000240   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000256   FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
```

[unregistered]
220px-SNice.bmp
C:\Users\weingan\Downloads

File size: 142 KB
145,254 bytes

DOS name: 220PX-~1.BMP

Default Edit Mode
State: original

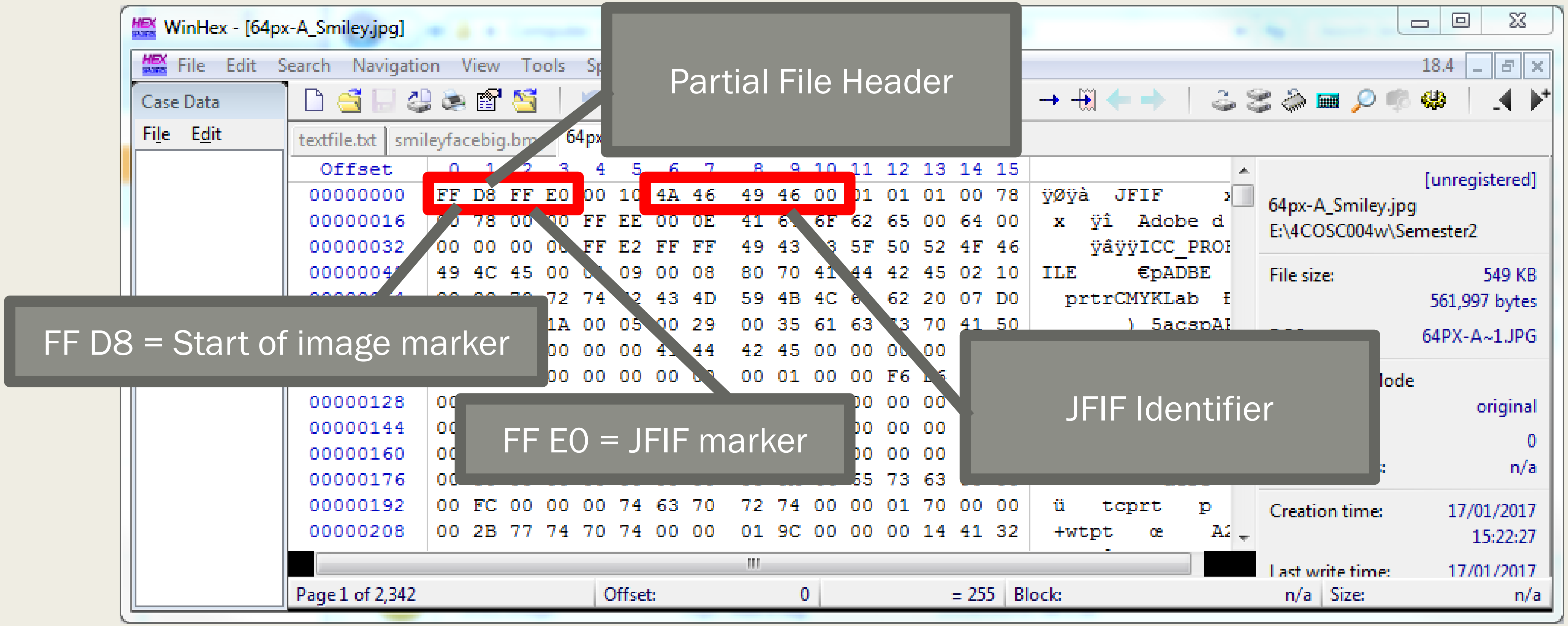
Undo level: 0
Undo reverses: n/a

Creation time: 12/02/2020
10:16:13

Last write time: 12/02/2020

Anatomy of a Binary file (*jpeg*) *File Headers*

(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   ÿØÿà  JFIF    x
00000016  00 78 00 00 FF EE 00 0E  41 64 6F 62 65 00 64 00    x  ÿî  Adobe d
00000032  00 00 00 00 FF E2 FF FF  49 43 43 5F 50 52 4F 46       ÿâÿÿICC_PROF
00000192  00 FC 00 00 00 74 63 70  72 74 00 00 01 70 00 00    ü   tcprt   p
00000208  00 2B 77 74 70 74 00 00  01 9C 00 00 00 14 41 32    +wtpt   œ    A2
```

[unregistered]
64px-A_Smiley.jpg
E:\4COSC004w\Semester2

File size: 549 KB
561,997 bytes

64PX-A~1.JPG

original
0
n/a

Creation time: 17/01/2017 15:22:27

Last write time: 17/01/2017

Page 1 of 2,342 | Offset: 0 | = 255 | Block: n/a | Size: n/a

WinHex - [64px-A_Smiley.jpg]
File Edit Search Navigation View Tools
Case Data
File Edit
textfile.txt
smileyfacebig.bm
64px
18.4

File signatures

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

File Operations

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

Unix File Systems: File permissions, Access control, Indexing

File protection

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

What happens when everyone can write to a file?

Cyber-security triad

- Three dimensions of cyber-security:
  - *Confidentiality*
  - *Integrity*
  - *Accessibility*

Confidentiality:

- Preventing access
- Keep the *bad-guy* out

Accessibility:

- Ensure access
- Make sure the *good-guy* can access the data

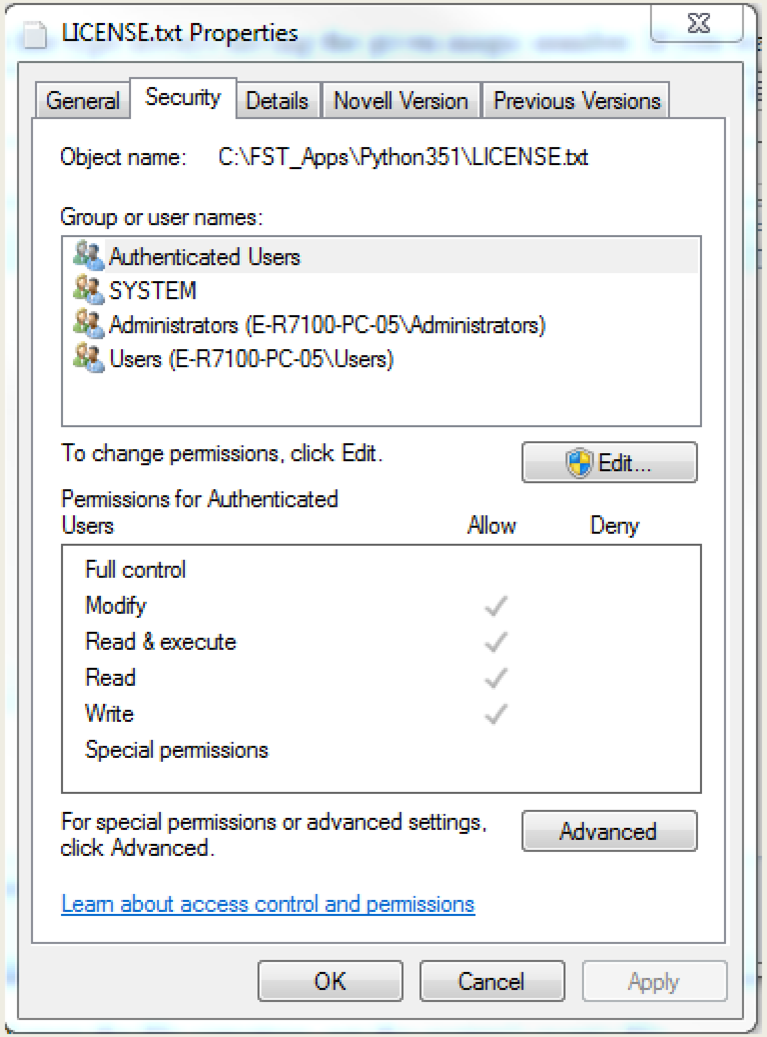
Integrity:

- 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



Object name: C:\FST_Apps\Python351\LICENSE.txt

Group or user names:

- Authenticated Users
- SYSTEM
- Administrators (E-R7100-PC-05\Administrators)
- Users (E-R7100-PC-05\Users)

To change permissions, click Edit.

Permissions for Authenticated Users	Allow	Deny
Full control		
Modify	✓	
Read & execute	✓	
Read	✓	
Write	✓	
Special permissions		

For special permissions or advanced settings, click Advanced.

Learn about access control and permissions

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*

Unix File Index System

- All files in a Unix system are defined by an **i-number**
  - *Index number*
  - *Unique in that File System*
- Directories are just lists
  - *Of file names*
  - *And their i-number (reference to i-list)*
- Relate to an **i-list** of **i-node** entries

Unix i-node entry

- Each **i-node** entry in the **i-list** contains:
  - *The username & group id of the owner*
  - *Protection bits*
  - *Physical file address*
  - *File size*
  - *Code indicating if file is:*
    - Directory,
    - Special Link
    - Symbolic Link
    - Ordinary file
  - *Number of links to the file*
  - *……*

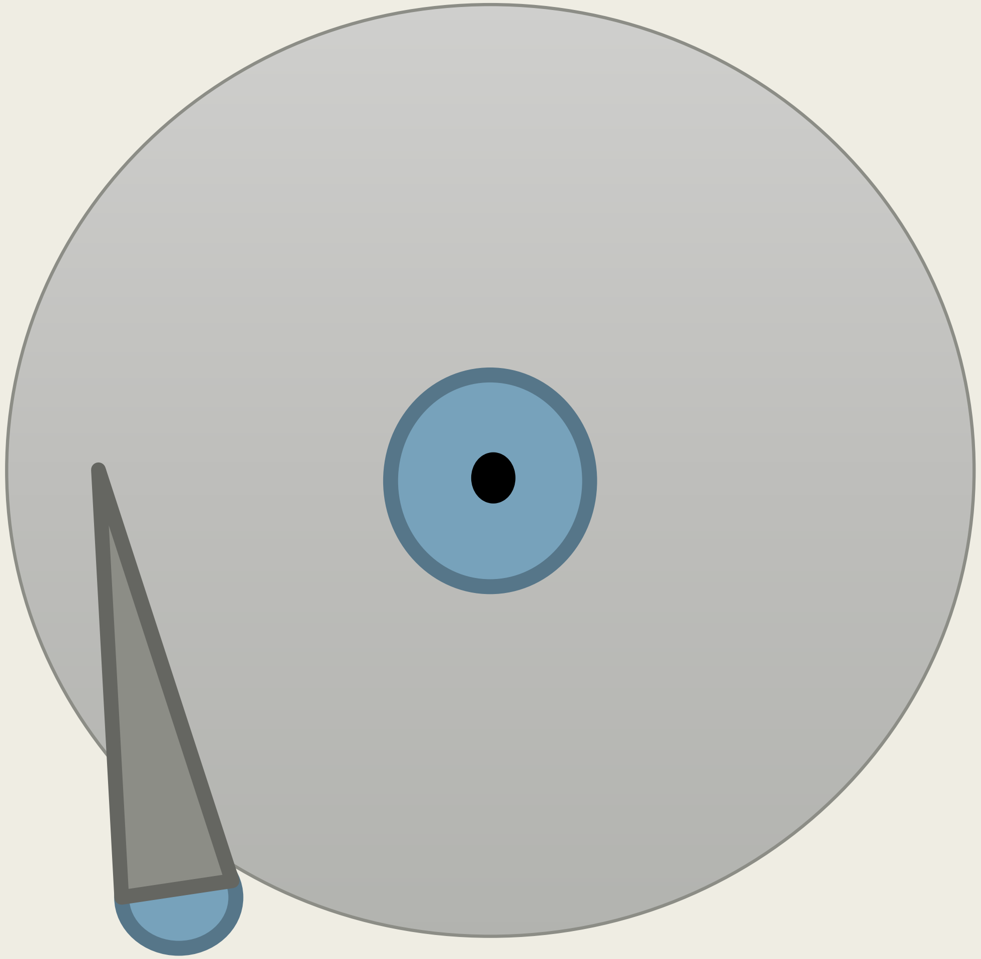
Disk scheduling

Disk scheduling

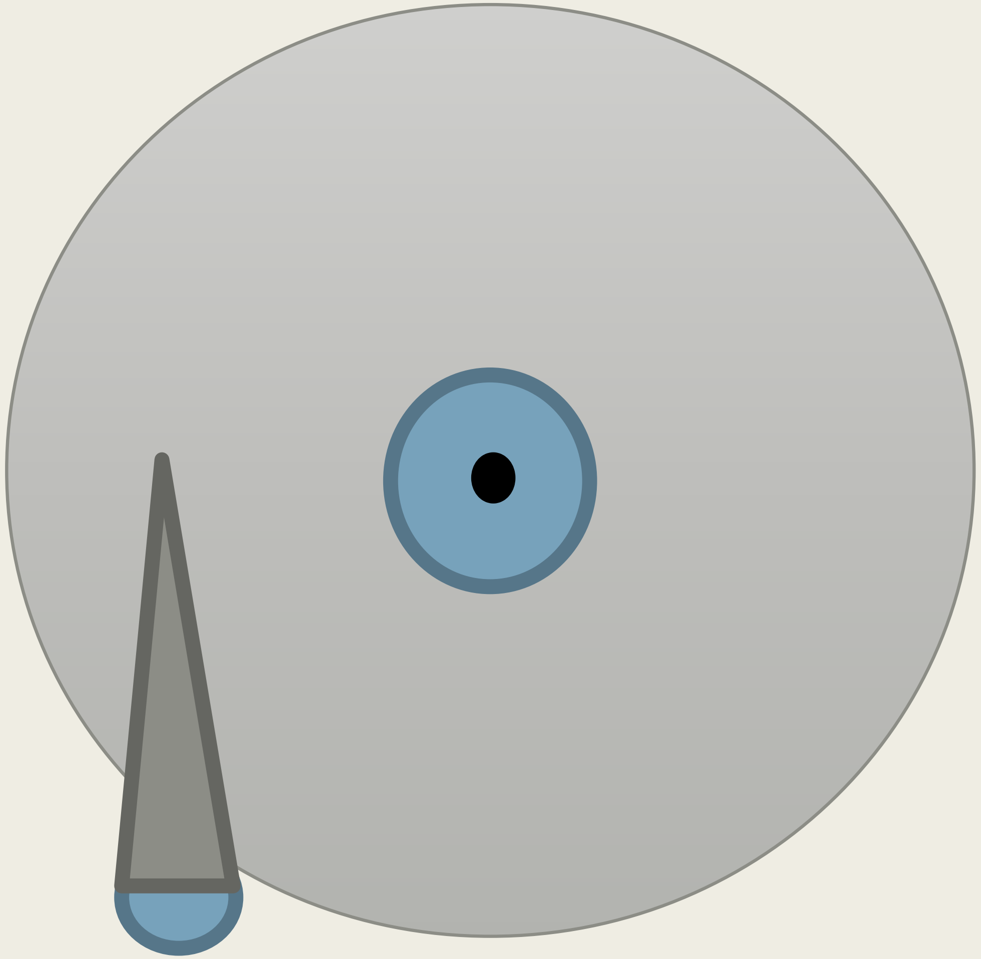
- Must be efficient
- Multiple processes, multiple requests to access disk
- Disk scheduling techniques to manage request:
  - *First-come, first-served (FCFS): Requests are serviced in the order they arrive, irrespective of positions of heads*
  - *Shortest-Seek-Time-first (SSTF): Minimise movement 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 the autopsy of the hard disk:

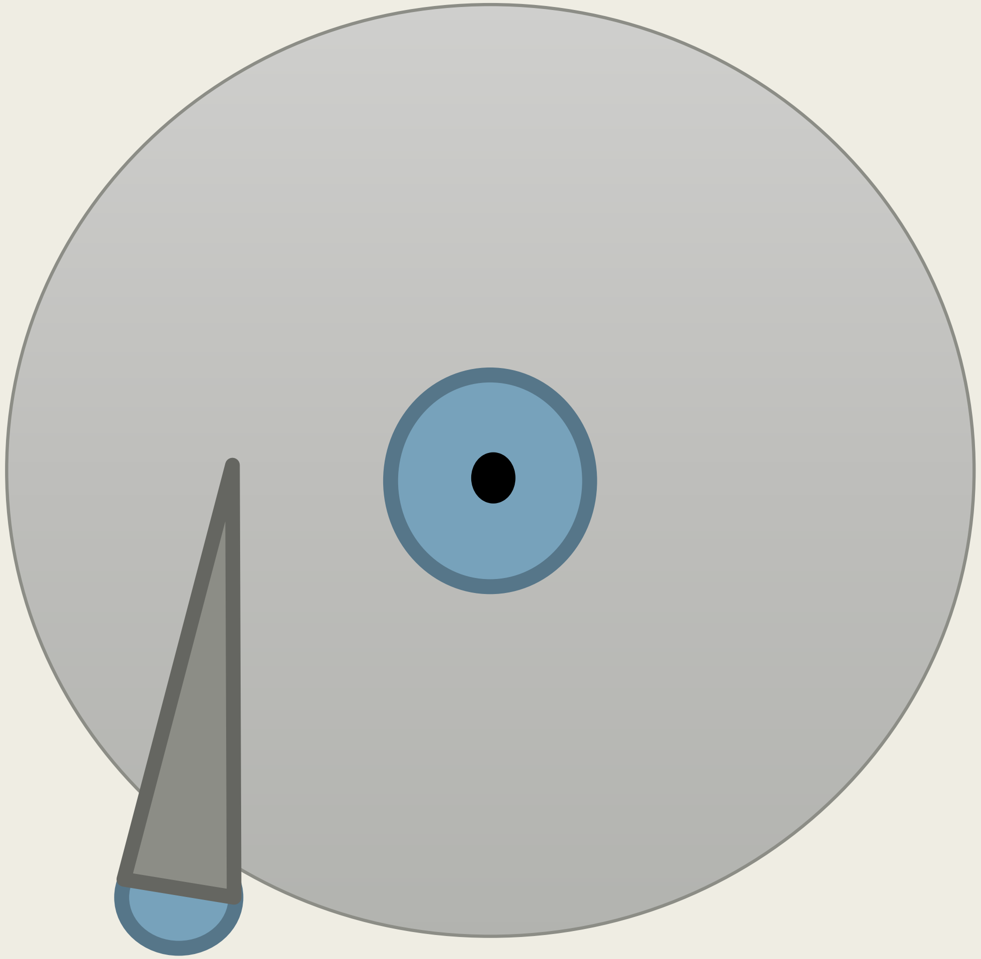
Remember the autopsy of the hard disk:



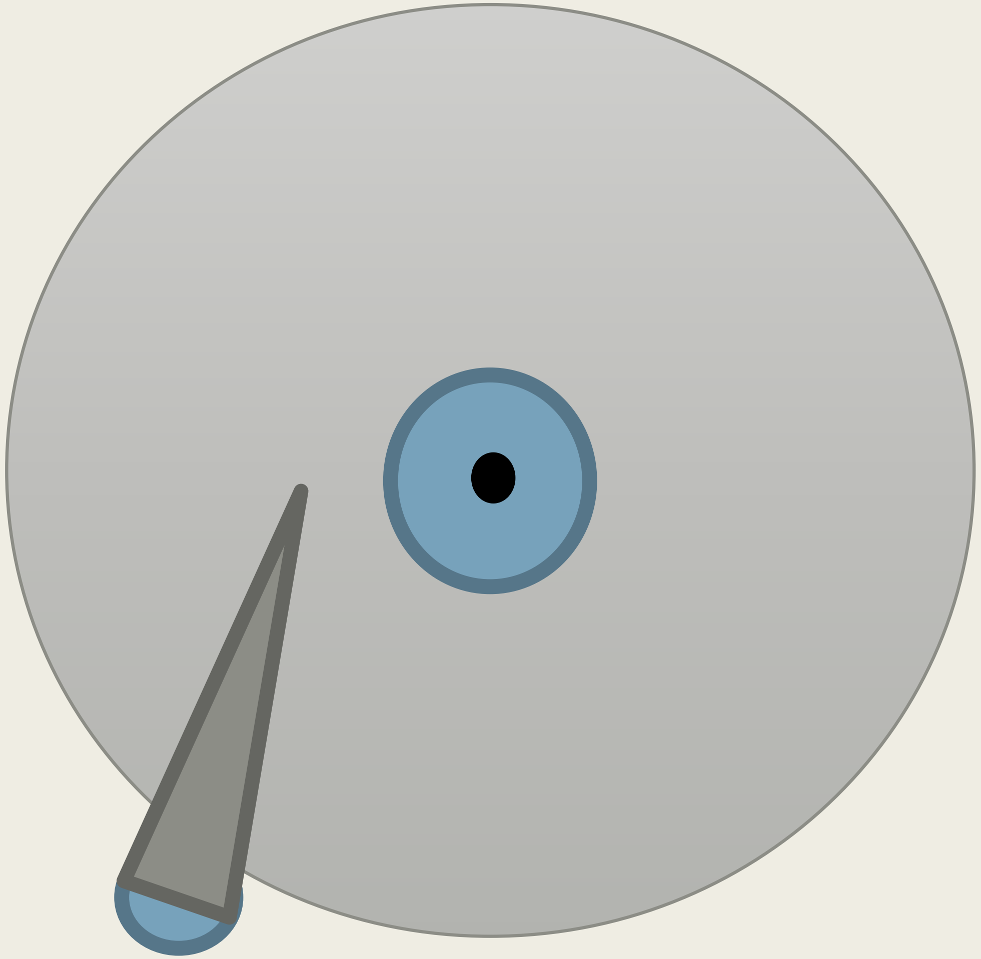
Remember the autopsy of the hard disk:



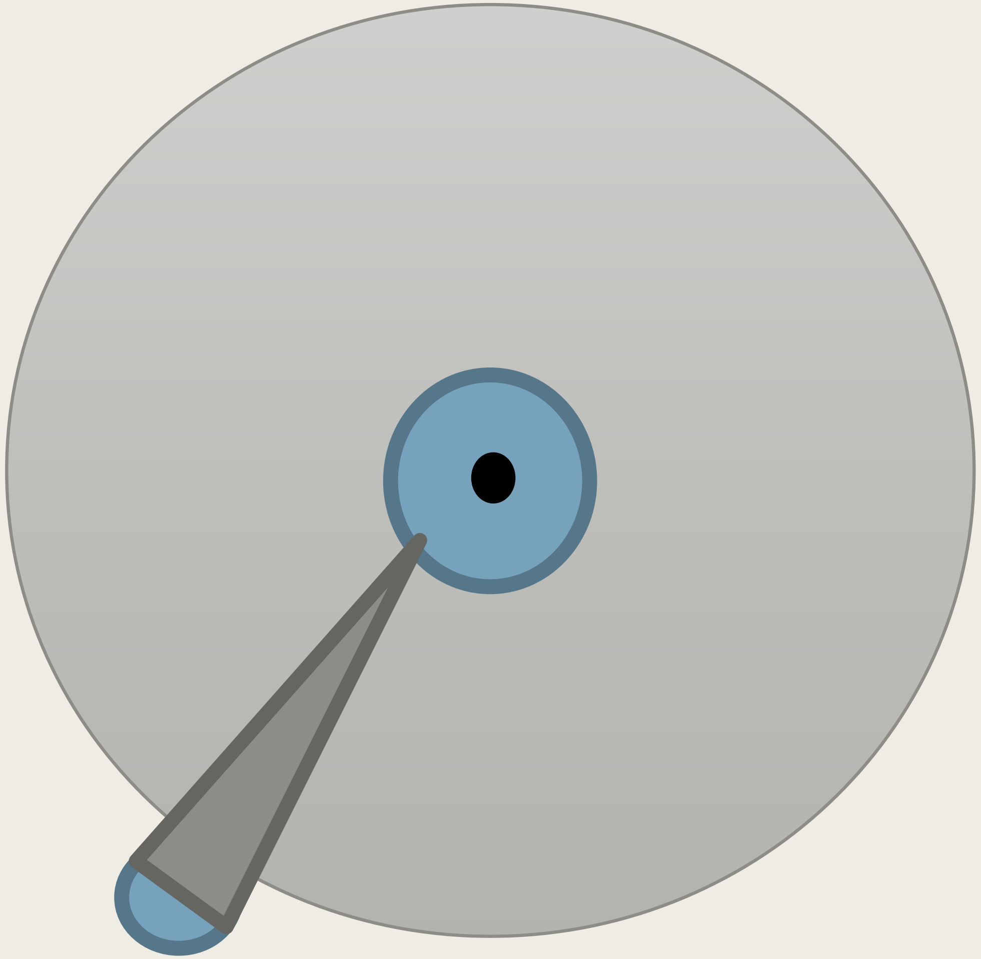
Remember the autopsy of the hard disk:



Remember the autopsy of the hard disk:



Remember the autopsy of the hard disk:



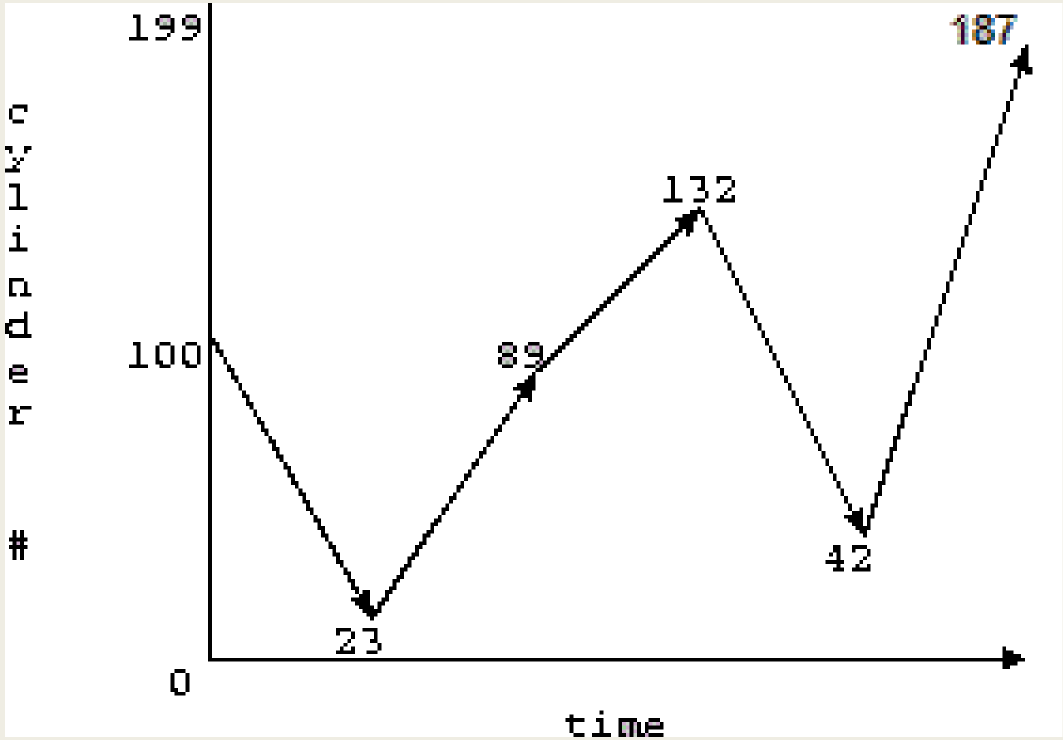
Example:

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

First-Come-First-Served:

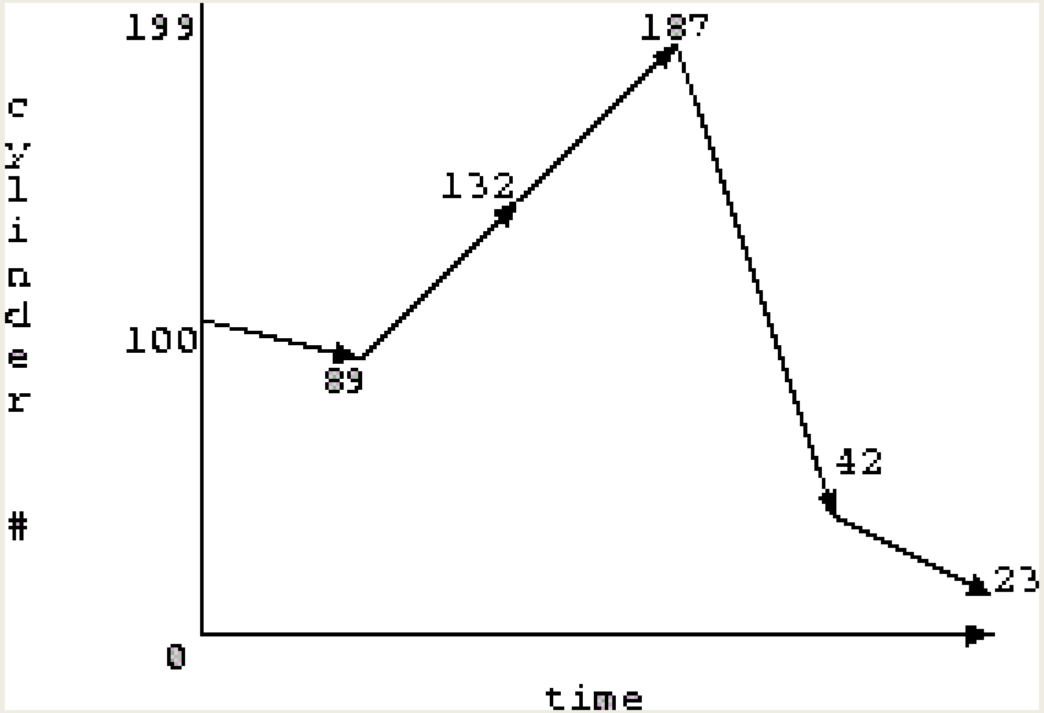
- Work queue: 23, 89, 132, 42, 187
- Total seek length:

$$|23-100|+|89-23|+|132-89|+|\quad + \qquad | \quad |42-132|+|187-42|=421$$



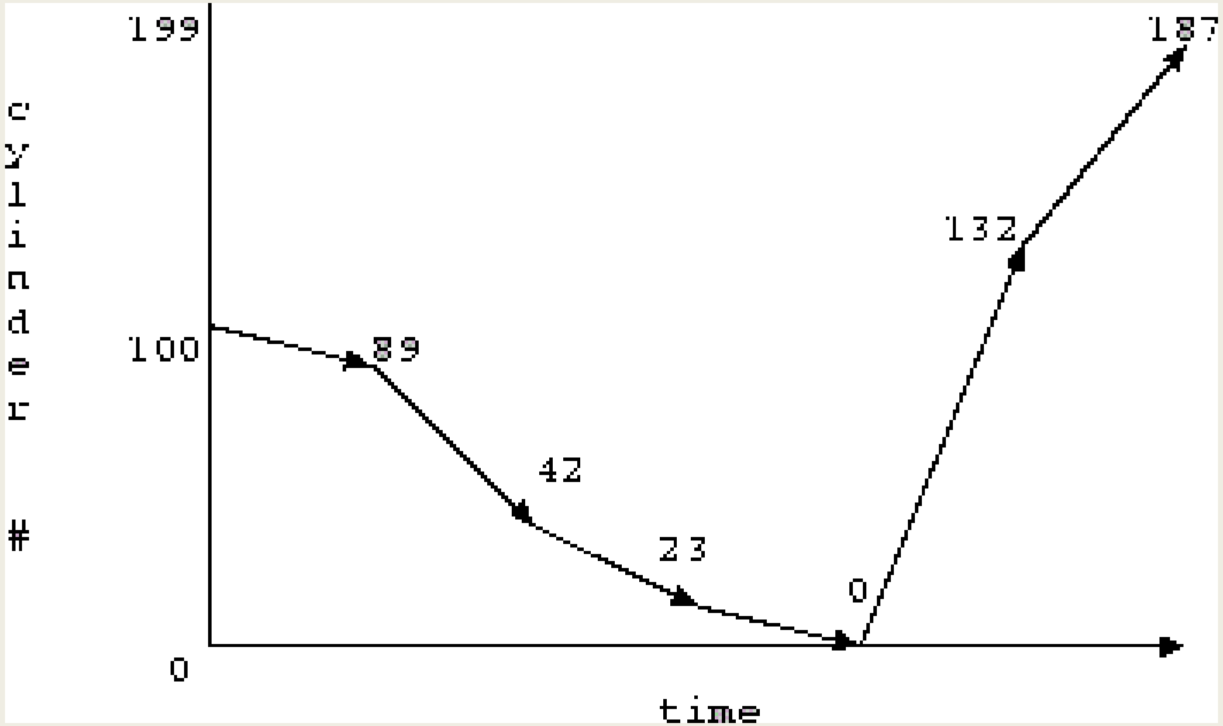
Shortest-Seek-Time-First:

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



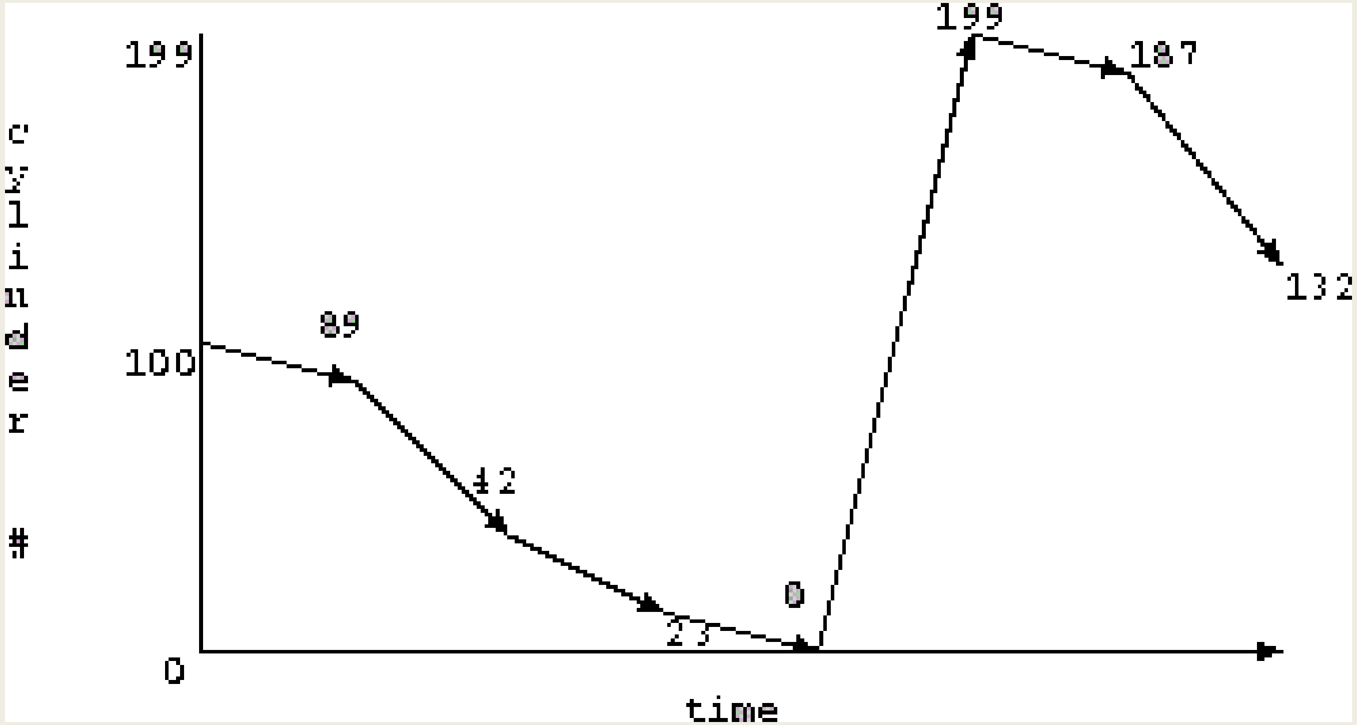
SCAN:

- 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*

What we covered:

- File types
- File operations
- Unix file systems
  - *File permissions*
  - *Access control*
  - *Indexing*
- Disk scheduling

Further reading:

- File signatures reference:
  - *https://filesignatures.net/*
  - *http://www.garykessler.net/library/file_sigs.html*
- Indexing & Disk scheduling:
  - *Operating Systems: Internals & Design Principles, Williams Stalling (7th ed.)*
    - PP. 550-552 & 510-512
    - Online, see Reading List

Tutorial exercise:

- File Types:
  - *How to establish exactly what the file type of a file is*
  - *Even if the extension is wrong*