DS 593: Privacy in Practice

Computers and Data at rest

Instructor: Alishah Chator

News?

Apple removes advanced data protection tool in face of UK government request

Apple says removal of tool after government asked for right to see data will make iCloud users more vulnerable

Business live – latest updates

https://www.theguardian.com/technology/2025/feb/21/apple-removes-advanced-data-protection-tool-uk-government

Last time

- More about privacy harms
- Group privacy

Today

- Computers!

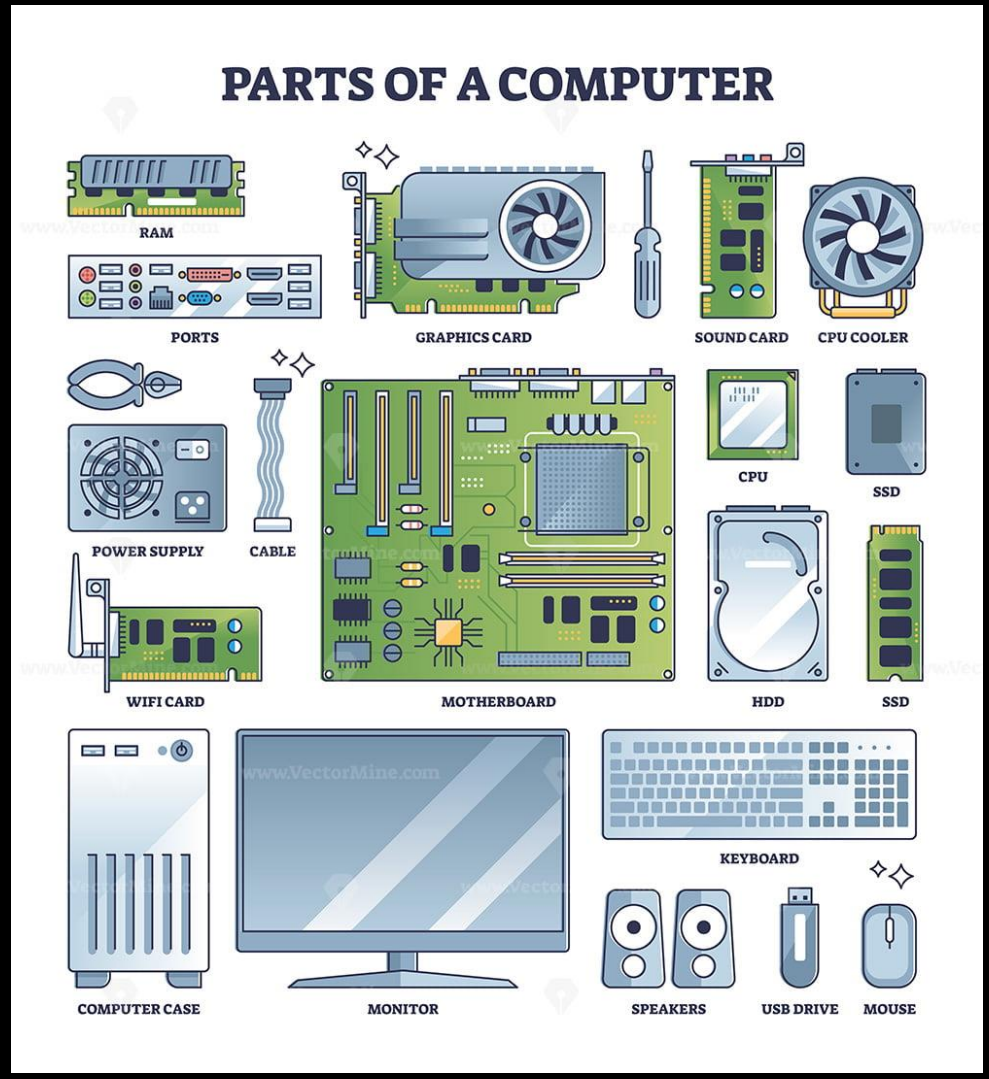
Roadmap

- Switching gears from how privacy is defined and enforced by social systems to how it is defined and understood in our technologies
- Building towards seeing how these two perspectives intersect and impact each other
- But, in order to do this we need firm grounding on the relevant background for our privacy technologies

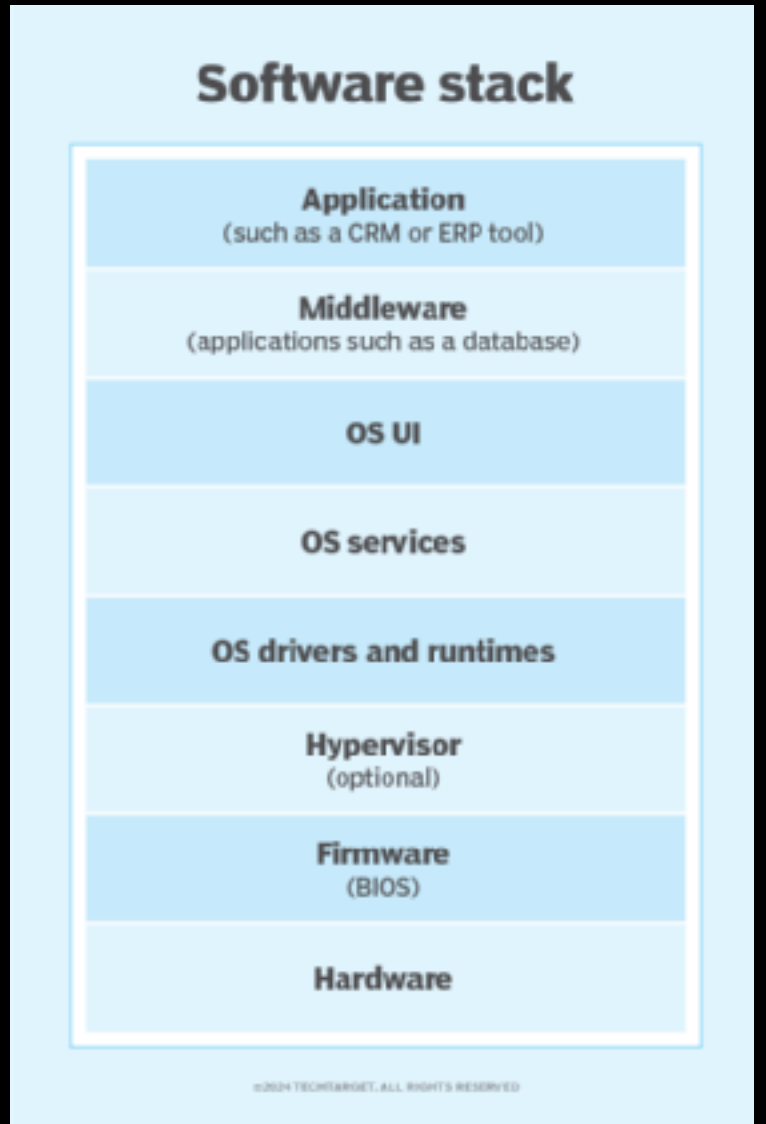
“There are no airplanes, only computers that fly. There are no cars, only computers we sit in. There are no hearing aids, only computers we put in our ears.”- Cory Doctorow

What exactly is a computer?

Hardware Perspective

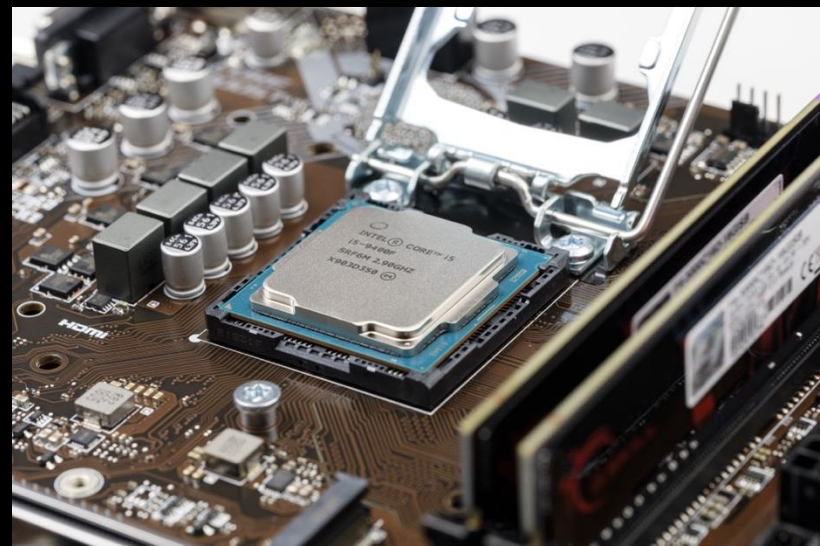


Software Perspective



Central Processing Unit

- The heart and brain of a computer
  - Lives on the motherboard
- This is where all of the “computing” happens
- Each architecture has its own language of instructions
  - Define what operations the CPU can do
  - Generally used to transform some memory
- There are also GPUs, these are less relevant for us
  - Used for computing graphics and visuals for display

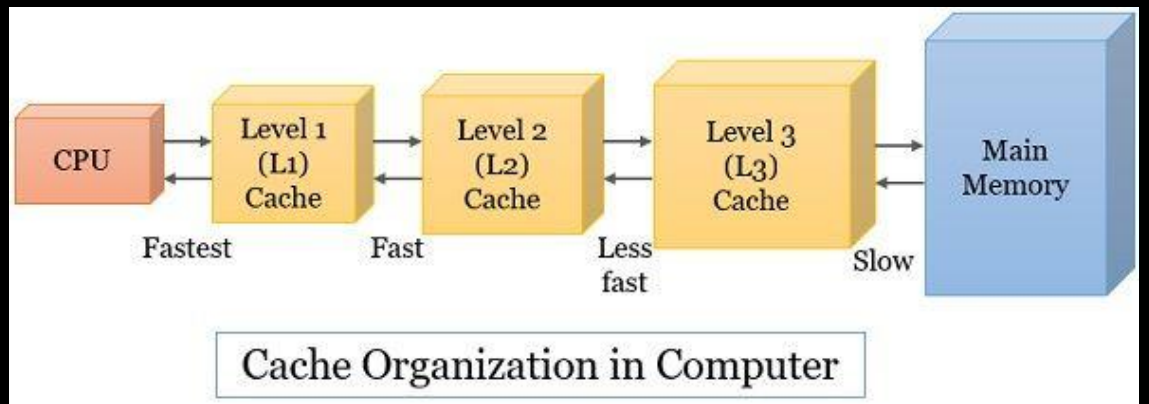


Memory

- Where data “lives”
  - Every piece of data has an “address”
- Different types of memory that vary in efficiency and persistence
  - Caches
  - RAM
  - Secondary
    - HDDs
    - SSDs
    - External Drives, USBs, SD cards, CDs, etc

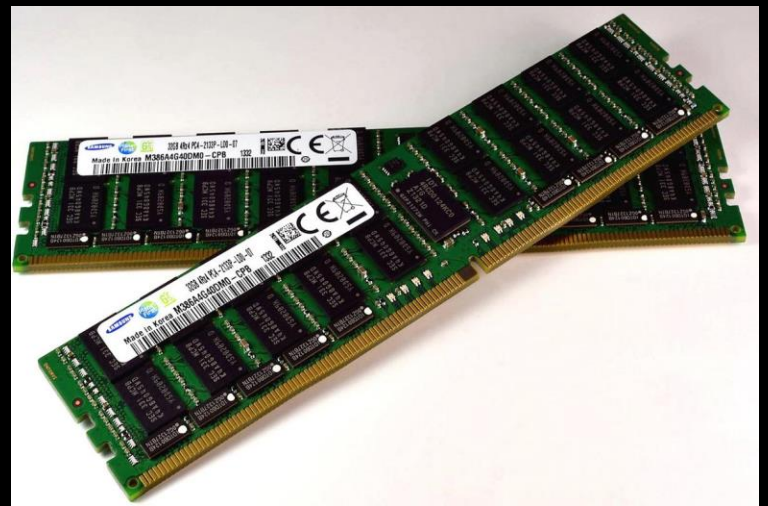
Cache

- Fastest to access but most temporary data store
- Several different types of caches
- Can think of them as a place the processor puts frequently useful data



Main Memory

- Colloquially used interchangeably with RAM
- Like Caches, it is *volatile* memory
- Used to quickly access and store data that is currently in use



Secondary Memory

- Persistent Memory- Data at Rest
- Where data lives when it is not in use
- Slower to access
- Where all your files are

I/O

- The devices that allow computers to interact with humans/ the outside world

- Inputs
  - Mouse
  - Keyboard
  - Mic
  - Webcam
- Outputs:
  - Monitor
  - Speakers
  - Printer

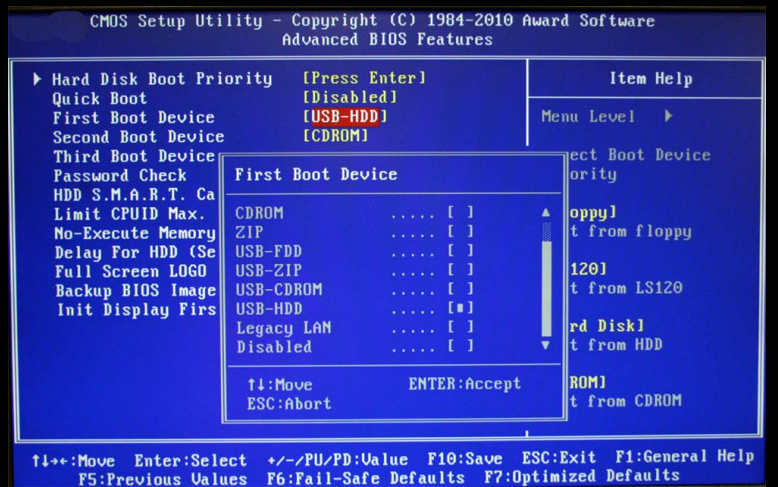
Network Card/Bluetooth/ETC

- By default, a computer cannot communicate with other computers
- Network cards provide the interface to facilitate this communication
- Allows for wired communication over Ethernet or for wireless communication over WiFi
- Other hardware dongles or antennae can support protocols like Bluetooth

BIOS and firmware

- Firmware refer to the software that controls hardware at a low level
  - This software usually exists in some read-only memory on the device
  - Some firmware can be changed in a process called flashing

- BIOS (basic I/O System) is the firmware on the motherboard that handles booting into the operating system

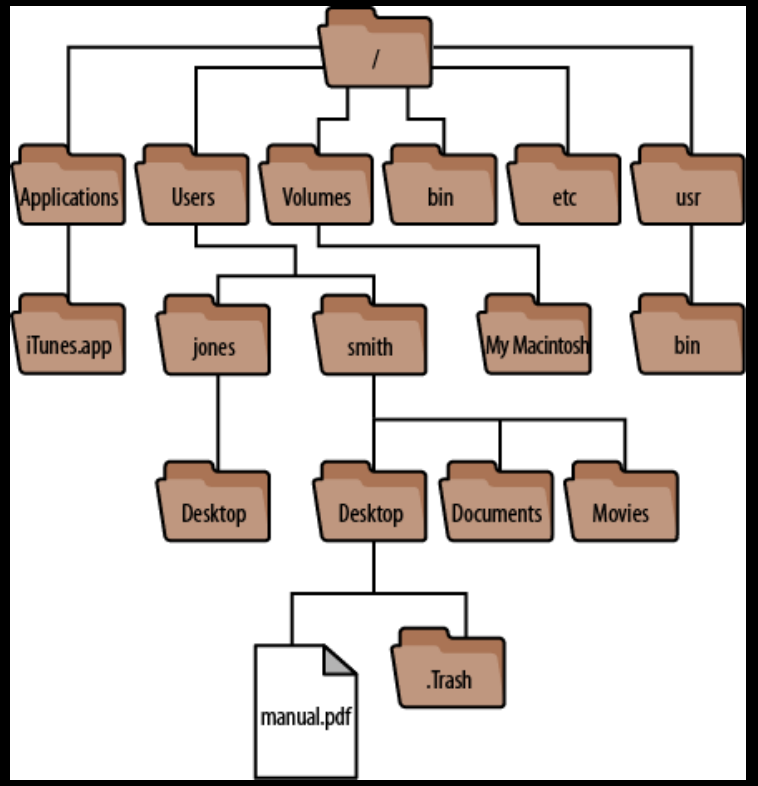


Operating System

- The ”Main” software for a computer
- Sometimes also called the Kernel
- Provides the interface between the hardware, network, and any applications
- Hands the administration of the system
  - Manages the file system
  - Allocates resources to tasks
  - Schedules tasks
  - Provides the user interface

File System

- How data and files are organized and stored on the system
- Can also be networked
- Generally comprises of folders and files
  - Separates data of different users
- Mainly lives on the secondary memory



Permissions

- Rules for who is allowed to access certain files, folders, or run certain programs
- Needed to control access to hardware resources, system-level data, user accounts
- Enforced by the operating system
  - Different users will be provided different permissions
  - Often there is a special “admin” or “superuser” account with full permissions

Threads

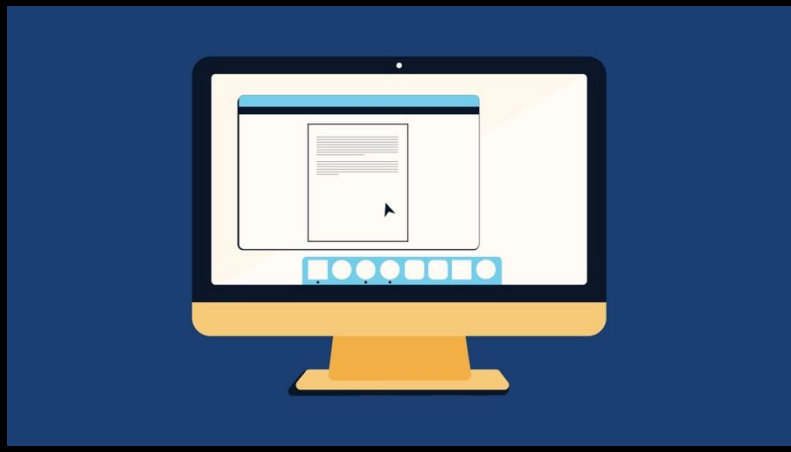
- The core unit of computing, a single task
- A sequence of programmed instructions that have been scheduled to be executed by the operating system
- Depending on CPU, threads can run sequentially or simultaneously

Processes

- All of the bookkeeping for a program that is running
  - The data it is using
  - What hardware resources it is using
  - The code that is being executed
  - The threads associated with this program
- Differs from a program, which is the passive list of instructions to execute that is typically stored as a file on disk
  - When this file is executed, a process is created
- Multitasking involves multiple processes sharing hardware resources and processors

Applications

- The fun stuff, higher-level software meant for end users
- Just like any other software, are computer programs and when running have their own computer processes
- Generally have restricted permissions



Basics of Data

- Computer data is allocated in blocks of bytes
- At its essence, all data is represented and stored in binary
- Files, directories, programs are all represented in this way
  - Everything that is not hardware is data
- Data can also be compressed or have redundancy added to it

Binary

Decimal	Binary
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

Encodings

- Binary by itself isn’t the most useful, so most programs will translate or encode this data into another representation
- Compact representations:
  - Higher bases (hexadecimal, base64) can allow for data to be described in fewer symbols of a higher base
- Apply meaning
  - A sequence of binary can be interpreted as a particular symbol for display
  - It can be interpreted as instructions for execution

Interpretations of Binary Patterns

Binary	Decimal	Hexadecimal	Excess	2's Complement
1111	15	F	7	-1
1110	14	E	6	-2
1101	13	D	5	-3
1100	12	C	4	-4
1011	11	B	3	-5
1010	10	A	2	-6
1001	9	9	1	-7
1000	8	8	0	-8
0111	7	7	-1	7
0110	6	6	-2	6
0101	5	5	-3	5
0100	4	4	-4	4
0011	3	3	-5	3
0010	2	2	-6	2
0001	1	1	-7	1
0000	0	0	-8	0

ASCII Table

Dec	Hex	Char	Name	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
0	0	NUL	null	32	20	Space	64	40	@	96	60	`
1	1	SOH	start of heading	33	21	!	65	41	A	97	61	a
2	2	STX	start of text	34	22	"	66	42	B	98	62	b
3	3	ETX	end of text	35	23	#	67	43	C	99	63	c
4	4	EOT	end of transmission	36	24	$	68	44	D	100	64	d
5	5	ENQ	enquiry	37	25	%	69	45	E	101	65	e
6	6	ACK	acknowledge	38	26	&	70	46	F	102	66	f
7	7	BEL	bell	39	27	'	71	47	G	103	67	g
8	8	BS	backspace	40	28	(	72	48	H	104	68	h
9	9	HT	horizontal tab	41	29	)	73	49	I	105	69	i
10	A	LF	line feed	42	2A	*	74	4A	J	106	6A	j
11	B	VT	vertical tab	43	2B	+	75	4B	K	107	6B	k
12	C	FF	form feed	44	2C	,	76	4C	L	108	6C	l
13	D	CR	carriage return	45	2D	-	77	4D	M	109	6D	m
14	E	SO	shift out	46	2E	.	78	4E	N	110	6E	n
15	F	SI	shift in	47	2F	/	79	4F	O	111	6F	o
16	10	DLE	data link escape	48	30	0	80	50	P	112	70	p
17	11	DC1	device control 1	49	31	1	81	51	Q	113	71	q
18	12	DC2	device control 2	50	32	2	82	52	R	114	72	r
19	13	DC3	device control 3	51	33	3	83	53	S	115	73	s
20	14	DC4	device control 4	52	34	4	84	54	T	116	74	t
21	15	NAK	negative acknowledge	53	35	5	85	55	U	117	75	u
22	16	SYN	synchronous idle	54	36	6	86	56	V	118	76	v
23	17	ETB	end of transmission block	55	37	7	87	57	W	119	77	w
24	18	CAN	cancel	56	38	8	88	58	X	120	78	x
25	19	EM	end of medium	57	39	9	89	59	Y	121	79	y
26	1A	SUB	substitute	58	3A	:	90	5A	Z	122	7A	z
27	1B	ESC	escape	59	3B	;	91	5B	[	123	7B	{
28	1C	FS	file separator	60	3C	<	92	5C	\	124	7C	|
29	1D	GS	group separator	61	3D	=	93	5D	]	125	7D	}
30	1E	RS	record separator	62	3E	>	94	5E	^	126	7E	~
31	1F	US	unit separator	63	3F	?	95	5F	_	127	7F	DEL

From https://ajsmith.org/tools/ascii-table/

Opcode (binary)	Instruction meaning
00000	No operation!
00001	Store operand in A
00010	Add memory location contents to A
00011	Output A
00100	Decrement contents of memory location
01000	Jump if not zero
11111	Halt
00110	Store contents of memory location in A
00101	Increment contents of memory location
00111	Add operand to A

File types

- A file system will often have a convention of describing different file types with different .extensions (i.e. .txt, .py, .html)
- The operating system uses this to associate certain files with specific programs
- A file type is just a hint on how to interpret the binary of a file
  - A file can be named like one file type, but actually be structured for a different file type

Metadata

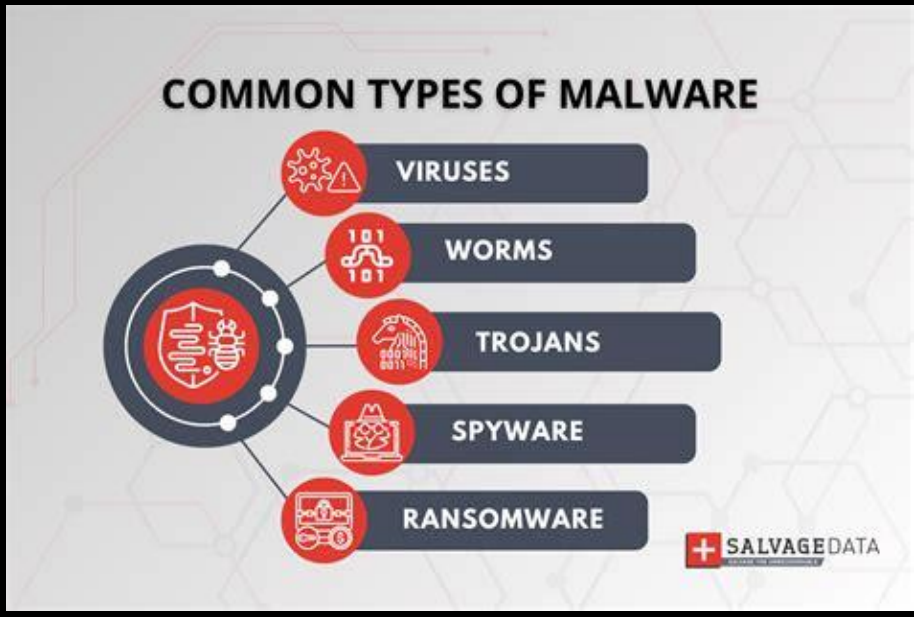
- The data of the data
- Generally used for reference by users or the operating system
- Examples:
  - Name
  - Size
  - Permissions
  - Changelog
  - Owner
  - Application specific info

Executables

- Special files that can be interpreted as instructions for the OS or CPU
- Sometimes also referred to as a “binary”
  - Output of code that is compiled
- Permissions are important to control who can run an executable and what the executable is allowed to do
  - Can have an executable that can change system files but requires admin permissions
  - Can have an executable that anyone can run but is only allowed to edit the current user’s files
- There are also scripts which are parsed by other programs and then executed
  - i.e. python, bash, etc

What is Malware?

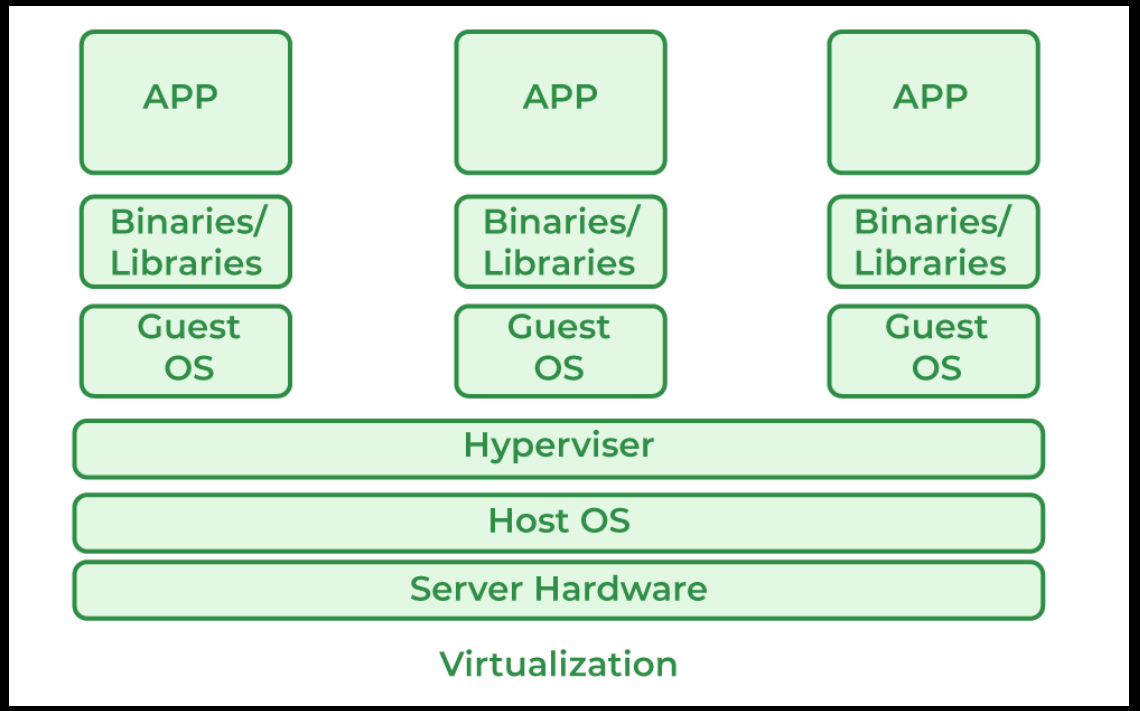
- Malicious Software
  - Tries to override permissions
  - Tries to hide its execution
  - Tries to escalate its privileges
  - Tries to directly access hardware
- High Level
  - Malware will try to exert control over software running at the same level or above it
  - It does this by trying to get a foothold on a lower level than it
    - A malware application will try to control the OS to allow it to expand its privileges and hide



Virtualization/Containerization

- Useful in many contexts, but virtualization is a way of providing isolation to the applications running on a machine

- This can be useful for “sandboxing”
  - What happens in one virtual machine or container does not affect other parts of the system
- Containerization is a newer approach that does application level isolation vs operating system level



Next Time

- The Internet!