

Sorting Techniques-

AI → set of algorithm

- Bubble sort
- Insertion sort
- Selection sort
- Merge sort
- Quick sort

Time Sort:  $O(n)$  - length of the list.  
 Not Optimized

Online shopping → Cookies → data related to our browsing -

byproduct of internet

→ built this business empire.

→ analyse the cookies and show the relevant data, google, & b are doing w/ (ml).

Mall - (Share your tracking location info) (GPS)

Enters to mall, location, many shops, promotional messg like 30%, 20% off, 50% to that  
↓  
lockstitch showroom.

increasing mode - tracking you |

mobile (eye-wear)  
↓  
sport medical  
Lover  
style

2 - 1, 3 -

Right to Roject → ① Jagaj finance - emails. unsubscribe -

not a date

hi. Date Privacy

Date Protection Bill

70% of ad market -

①

storing

storing

Malware

insights

turn to volume

Terradate (RDBMS).

→ RDBMS - mysql, sql server, oracle

structured data

spark and all (processing tool).

unstructured data

→ Big data

semi structured data

xml, html, tag associated to

PowerCenter

Information

Developer - Motion

Eng. mail

Ennas Mobil → (Name of the company)

M Seal

ELT VS ETL.

Moving data  $\rightarrow$  Analysis

~~Extract~~

~~Load~~

Transform

static data / stored data - Analysis is

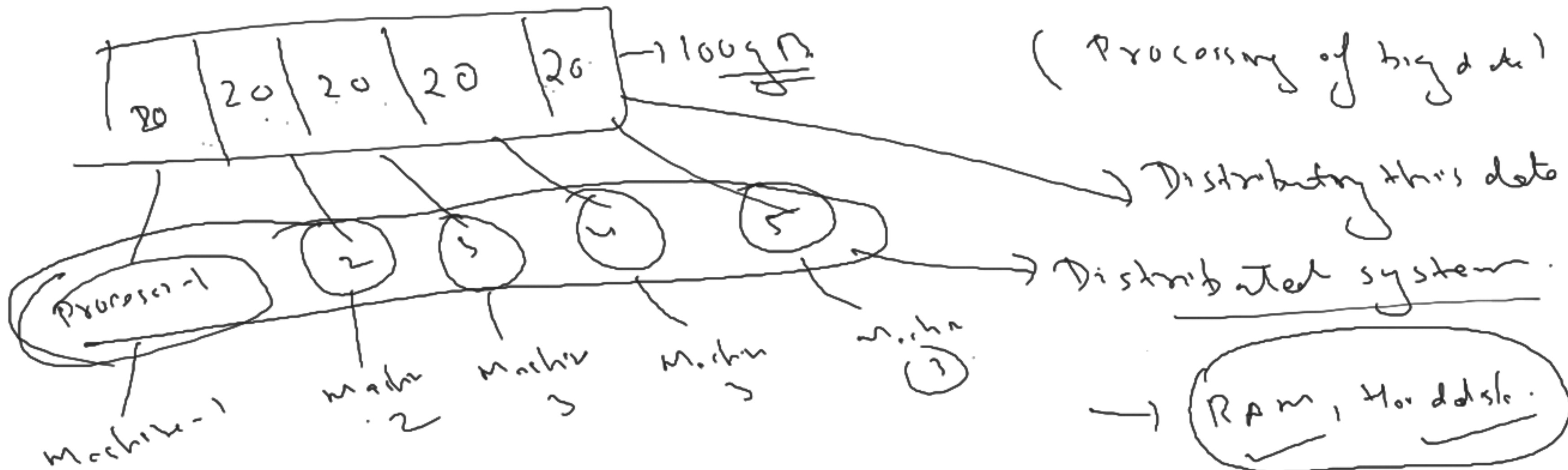
based upon stored data

① Streaming

② Batch processing  $\rightarrow$  Analysis

Storage cost < < < < processing data  
(min, ssn) < < < < (name)

100GB < < < < 100GB



Octa core - Single machine.  
Multiple machines - multiple processor.

① Serialization. ② De-Serialization } → PySpark.

converting your code into bytes.

(bytes to your code)

Data moves inside a network -

(bytes (1024 bits) → (0,1))



② Install a wifi → taking a rust (4k, 500).

machine → spyder → fiber optic cable → glasser → your modem (by his hand) → (1.5 lakhs)

Conversion ice to water (for proper transportation) - serialization  
Conversion of water to ice (desired output) - deserialisation

Q.L

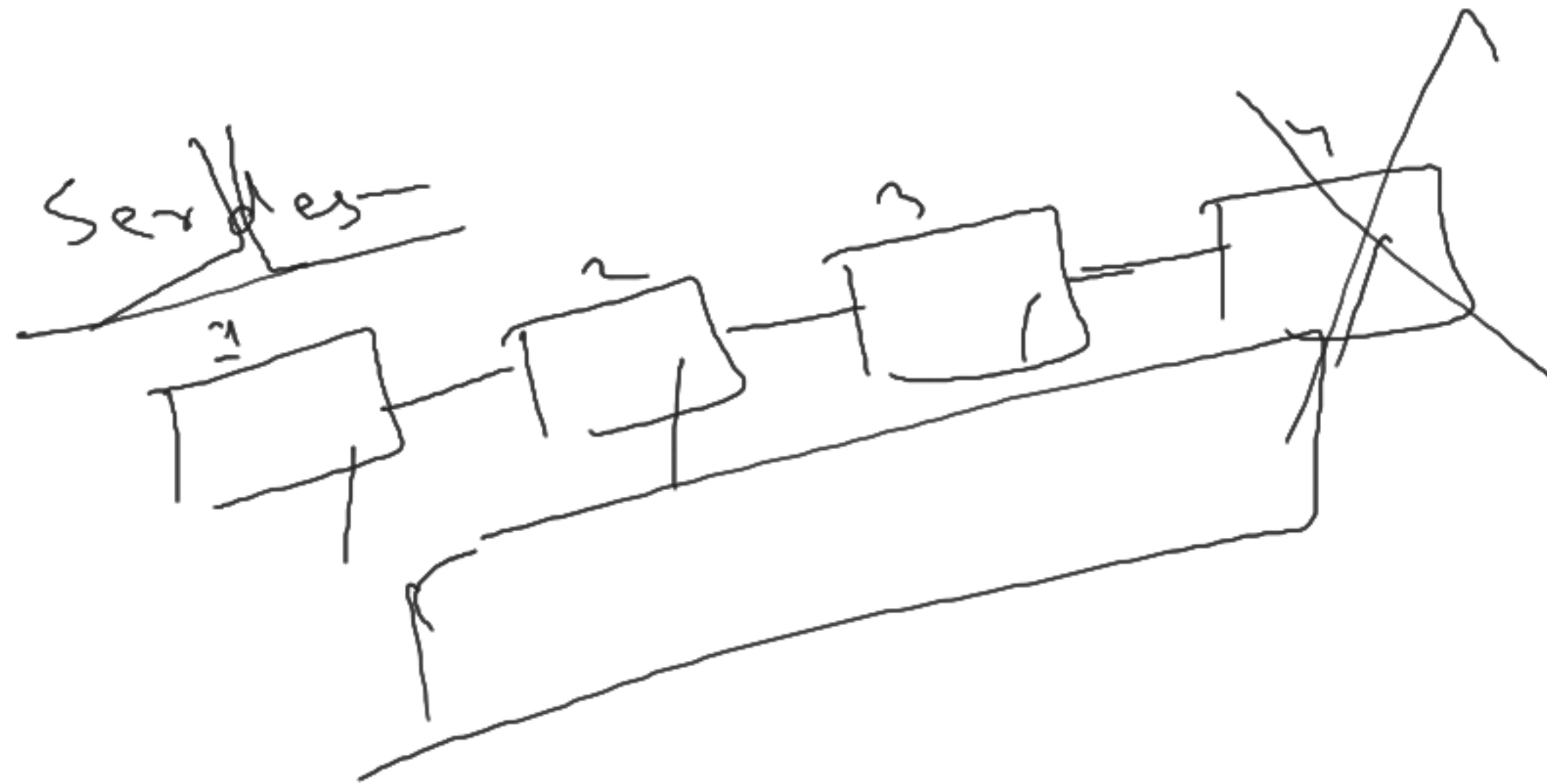
value

Pickling / Serialization / Deserialization / Unpickling / Marshalling / Un-marshalling

① It would be people generally difficult terms to explain a basic thing.

It makes it unique.

(Eventual consistency).



bits loss -  
electricity.

① Popular tool (Hadoop) - (HDFS) → Hadoop Distributed Filesystem.  
→ Google (GFS) → solved issue with big data.

② Kubernetes - (Google) → it was 10 years now (with a different name).

① Teacher/Trainee - Distribution (review)

② MySQL - (windows).

mysql/mysql workbench

③ Hadoop - open source (3.X)

↳ Any train or it

② Installation hard

↳ 4GB.

Virtual - linux.





