

① Physical / Colocation data centers

- (i) Instead of ~~buying~~ bldg costly capital intensive data IT companies (bunch of em) come together and rent spaces in shared facilities
- (ii)节省 capital for more flexib. uses as compared to real estate

② Virtualized data center

- (i) Comps. match the parts of a physical data center; servers, disks etc.
- (ii) But now there are virtual devices separately manageable from underlying hardware.
- (iii) Resources used $\xrightarrow{\text{more}}$ efficiently $\xrightarrow{\text{flexibly}}$.
- (iv) You can still buy, house, maintain infra. but you still gotta guess how much hardware you'll need & when, settin it up and keepin it runnin

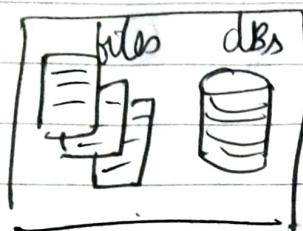
③ Cloud (~~serverless~~ GCP provides this in a serverless manner)

- (i) In internal cloud, services auto-provision and config the infra used to run familiar Google apps.
- (ii) Uses ~~cont~~ container (cnr) based architecture
- (iii) Provides wide variety of services (pre-configured) to manage & get data at large scale.

Cloud Computing

Delivery Model for services like

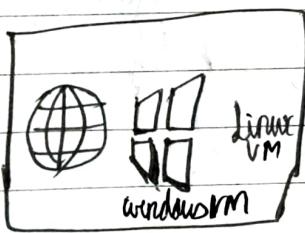
- ① Storage → unstructured or structured data in files or DBs



Cloud delivers user 100s of services to do so.

Also stores all the services and tools need to migrate your data to the cloud.

- ② Compute Power → web apps

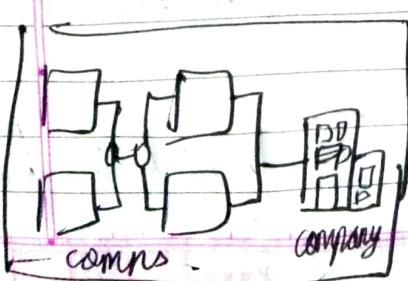


Windows, Linux VMs

Any of 100s of services available in the cloud cuz cloud cuz cloud is all about creating apps → web

→ AI
→ ML
→ reporting
etc.

- ③ Networks



Connect your corporate to each other securely

Cloud delivers ~~is~~ a set of facilities to create secure networks b/w these comps,

④ Analytics



performance data

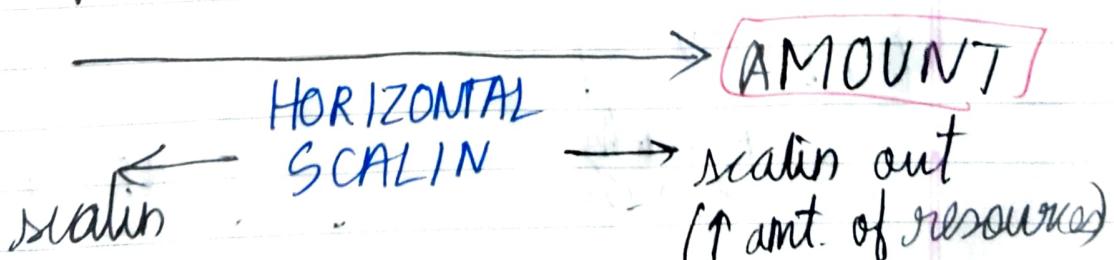
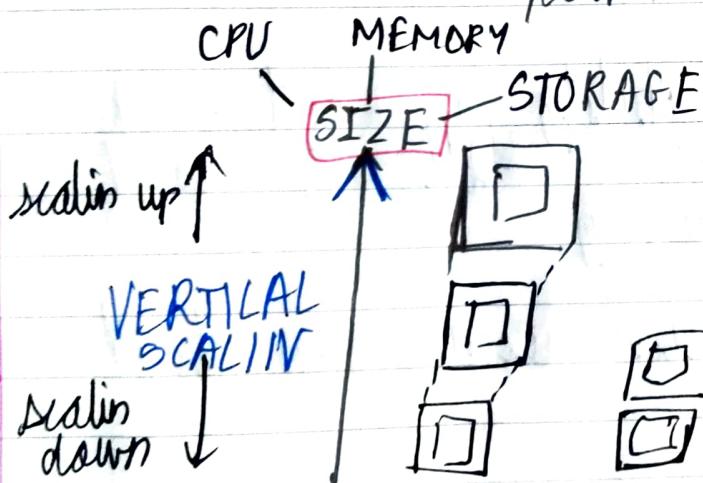
→ to review performance, telemetry data for your services so you can perform read. op's

These are the 4 MAIN services delivered by cloud computing over the internet

other characteristics Cloud Computing fulfills

① Scalability → Addin more power via

→ More CPU, mem, faster storage
More no. of machines



Scalin → process of addin or removin resources

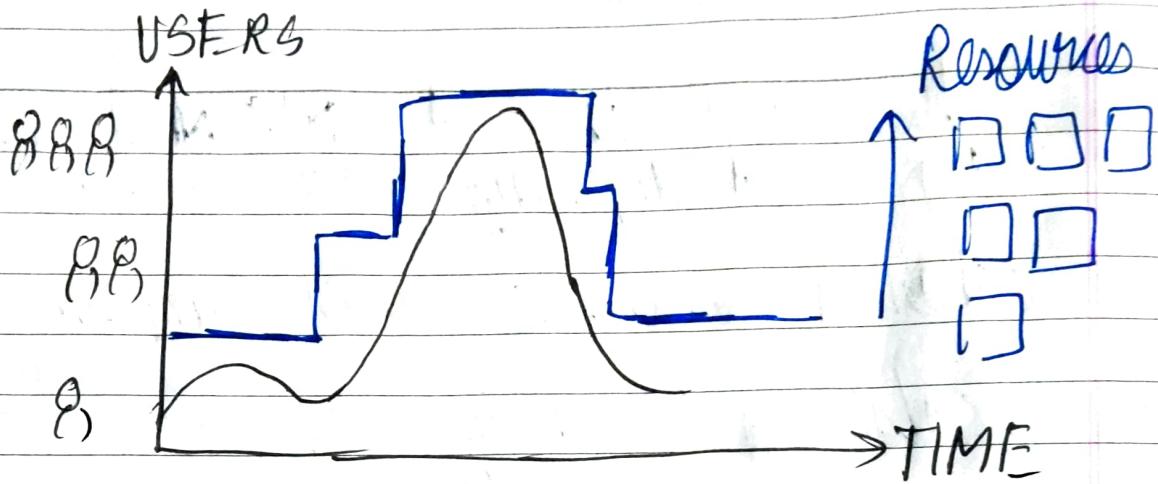
Eg:- RAM is 64 GB (size)

↑ No. of 8 GB RAM devices to 8 (horiz.)

size
vertical
amt. of instances of the resource
horiz.

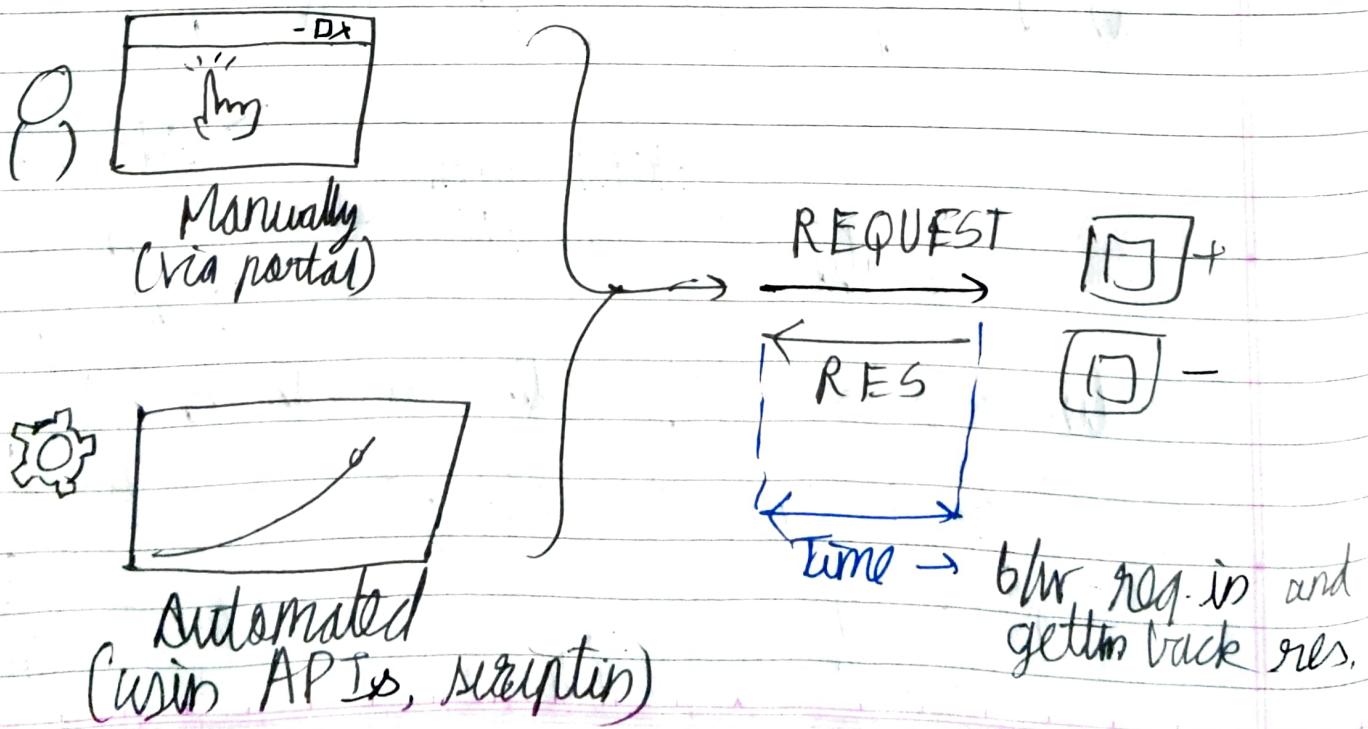
② Elasticity :- dynamic allocation/removal of resources as reqd.
aka auto-scaling

e.g. - adding/removing resources as user workload for that day changes.



③ Agility → The ability to react quickly i.e. ability to allocate & deallocate (scale) resources quickly!

NOTE: 2 ways to provision resources in cloud



CloudvsOn-PremiseSee
Mth
hr

day

week

month

Cloud responds with the resource you req'd very fast
∴ It is agile

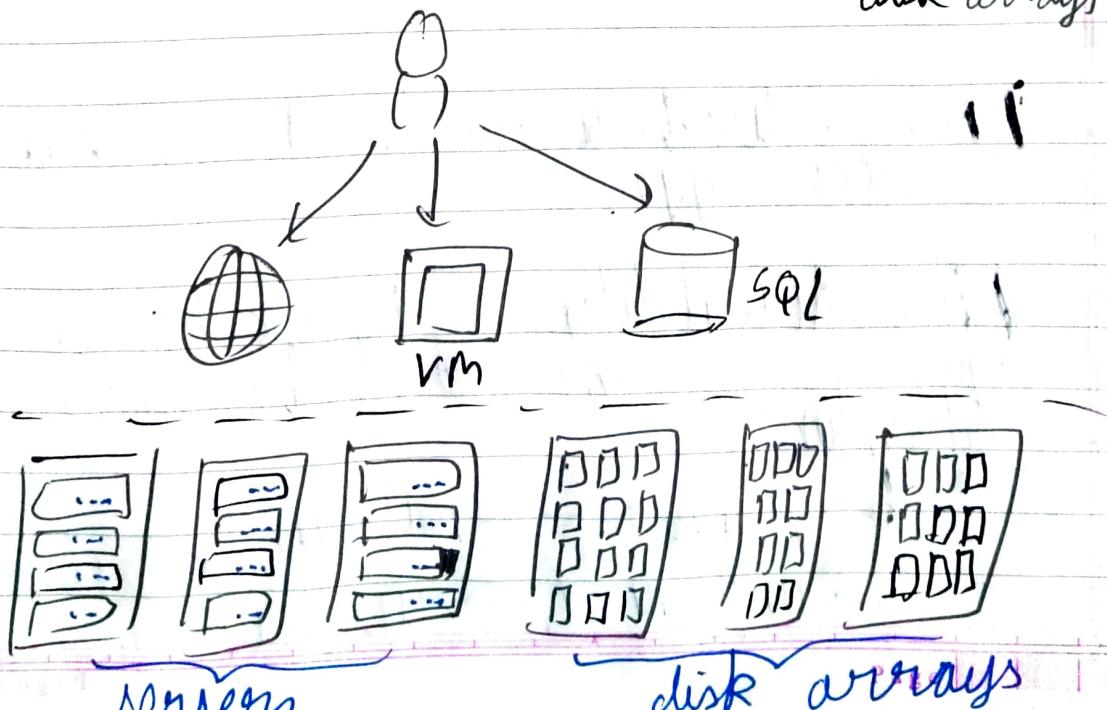
(4)

Fault Tolerance → Whenever there is any service or comp. failures on servers or when your data in disk array crashes → you get immediately assigned another one. s.t. ~~Pattern~~ no data is lost.

Fault tolerance → ability of sys to remain up & running during comp & service failures.

NOTE : Whenever you purchase a web app, VM or SQL db from a service, regardless of the service or you choose or interface you work with, ~~you~~ under the hood services run on  servers (do computers)

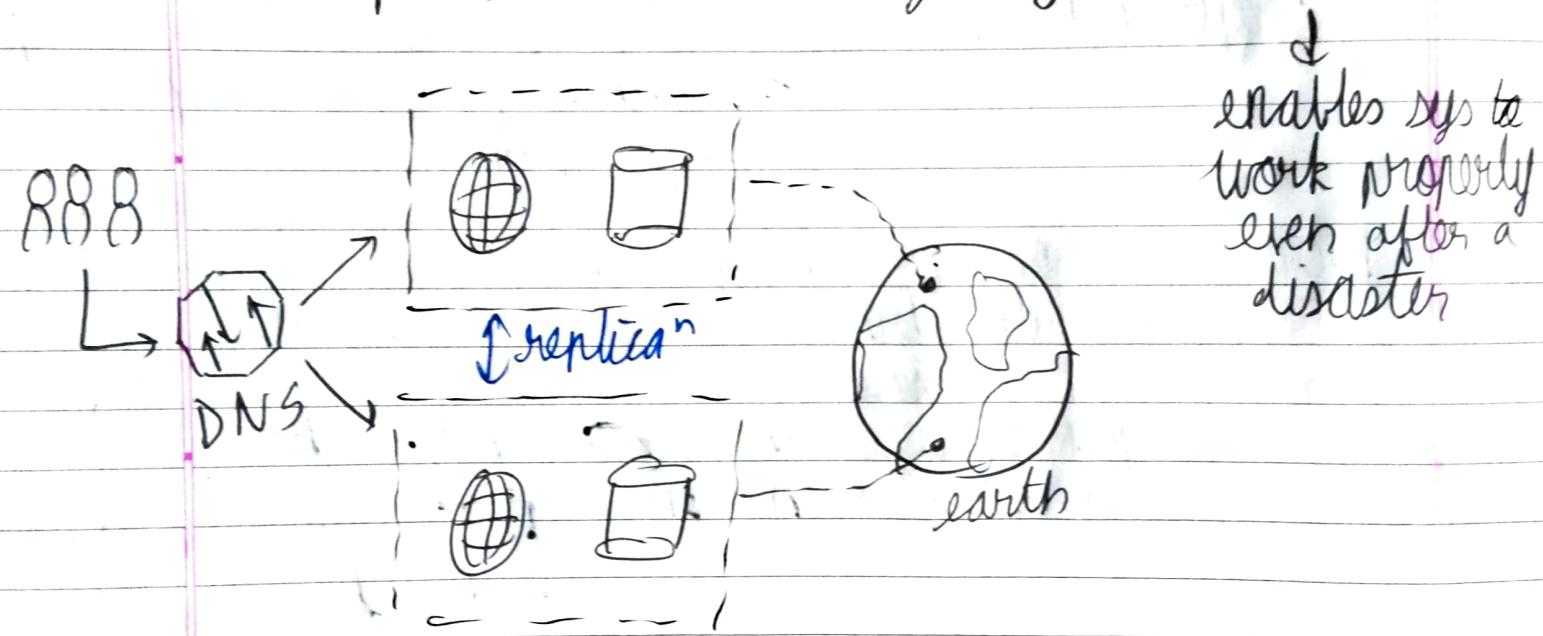
 disk arrays (store data)



Localized failures do not interrupt your service
 as most of the time in the cloud all
 the services have circuit ~~seg~~ fault-tolerance
 sys (immune to localized failures)

- ⑤ Disaster → Failure of a v. big slab
 Recovery → disrupt of services caused by natural
 or human-induced causes
 e.g. - power grid failures, storms.

Setup disaster recovery by "replica"



DNS redirects users to the working version of your app.

- ⑥ High Availability → Metric that measures uptime (sys being accessible to users) vs
downtime
- planned → patching
 → unplanned → sys. failures

(alt)

availability - calc of uptime wrt a specific time period e.g. 1 yr, lifetime of service

uptime
uptime + downtime

Dependin upon client you calc alt per yr / month / day.

Alt.	Year
99%	3.65 days
99.9%	8.77 hrs
99.99%	52.6 mins

Alt → measure of uptime for users / services

High alt → ability of sys to keep services running for extended period of time with r. lit downtime.

⑦

⑥ Global Reach

Ability to reach audiences around the globe.



more customer base for your app.

Cloud services can have a presence in various regions across the globe, which your customer can access, giving you a presence in those regions even tho you may not have any infra in that region

⑧

Customer latency caps. → Cloud services hve the ability to deploy resources in datacenters around the globe which addresses any customer latency issues

Cloud services

~~Cloud~~ gives amazing latencies to services to react to cust ac's when the service are local to the cust.

Modern optic fibres + local service → faster than cloud.

(9)

Predictive cost consider'n → The ability for users to predict costs that they'll incur for a particular cloud service.

i) Cost for indiv services are made avail. and tools are provided to predict cost

ii) can perform analysis based on planned growth.

(10)

Technical skill reqmts & consider'n → A user can be an expert in the app they wanna run w/o requiring skills to build & maintain underlying hardware & software for workloads (as cloud provides it all) infra

(11)

↑ Productivity →

i) On prem. data centers require →

hardware setup,
aka "racking & stacking"
software patching

ii) Cloud Computing (CC) elims.

need for many of these tasks, allowing IT teams to spend time focussin

other time-consumin IT management chores

on achievin more imp business goals

(12)

Security → Cloud providers (CPs) offer a broad set of policies, techs, controls, expert tech skills → provide better security than most orgs.

- We get strengthened security to protect

- 1) data
- 2) apps
- 3) infra

} protected against potential threats

Principles of economics of scale

e.g.: - your own delivery company

Scale	 X 3	 X 300
Car	3rd party purchase : 10k	Bulk purchase : 9k
Maintenance	Andhr : 100	Contract : 90
Insurance	Andhr : 500	Bulk purchase : 400
Other services	Andhr : 100	Shared services : 80
Customer	8 x 5	8 x 5000
Price per unit (PPV)	₹ 10	9

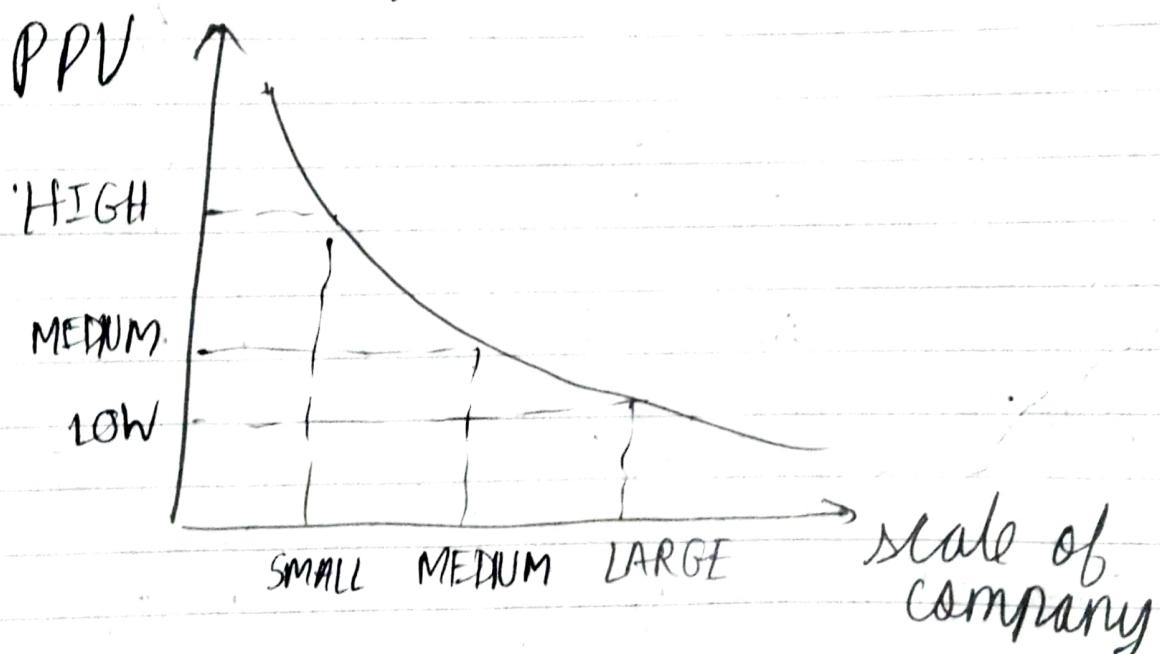
Basically as your company grows (scales up), your PPV falls due to bulk purchases, contracts, shared services

employ some 3rd party staff for maintenance etc.

large companies → scale is large → can handle more cust.ers w/o sacrificing quality of service

Large companies get a lot of internal benefits which it shares with cust.ers

Makes company more competitive to grow in market



This principle perfectly describes what's happening with the cloud ~~is~~ right now.

As the cloud grows & has more cust.ers, the prices of services go down

Reason → Big companies like Microsoft will get more efficient at what it does → building data centers → getting better ~~hard~~, advanced hardware

Thus,

Economies of scale → the ability to reduce costs and gain efficiency when operating at a larger scale.

Storage costs, in the past ~~decade~~ decade have ↓ed significantly due to the CP's ability to purchase larger amounts of storage at significant discounts

then, they use that storage more efficiently and pass on these benefits to end users at low prices