Agenda

1) Challenges in Scaling
2) What & Why of HLD
3) Case Study → del.icio.us

Pre-requisites → DSA,
Databases,
Schema Design,
Networking,
Concurrency, etc.

50 — 60%

How do we build applications ?

Learn → Small Project / code on local system.

Real Life → Work with team
Large data
Test / Debug
change in requirements ...

Google Question (Architect / Principle Engineer)
>1 Cr CTC

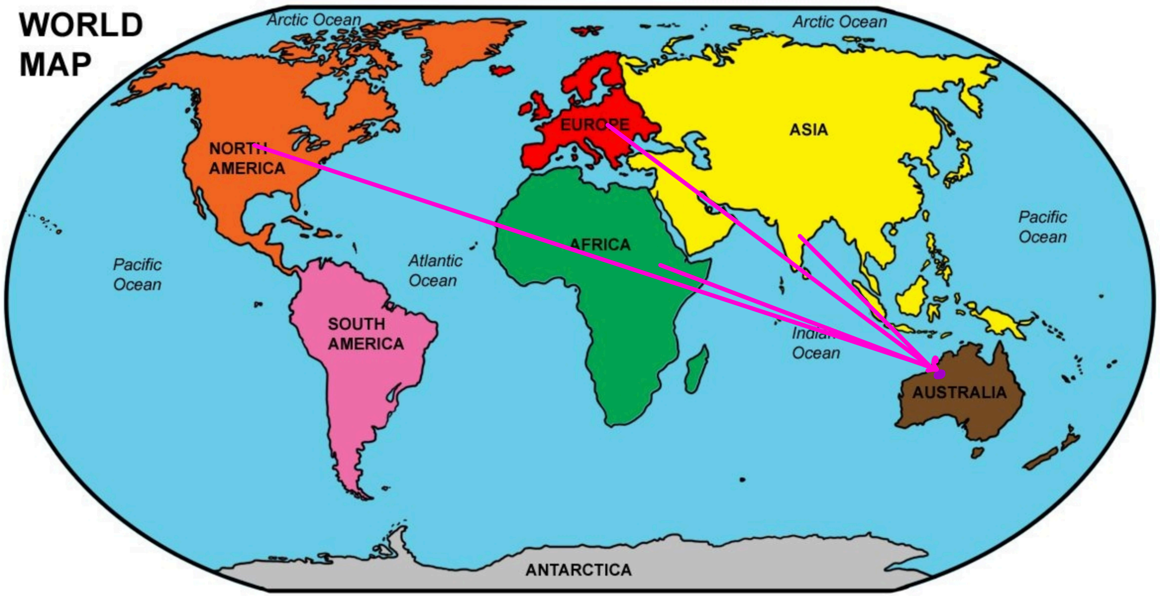
Q → Arrange a list of strings in dictionary order.

Java → Arrays.sort() / collections.sort()

size of data → 50 PB (v.v.v. large data)

Eg → Posts of instagram.

large data is distributed across the world.



Challenges

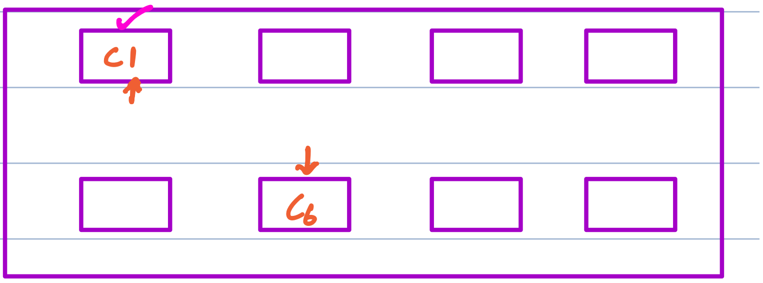
1) Internet can go down
2) Machine crash
3) Natural disaster
4) Human errors
5) Cyber attack
6) No electricity ...

solve all problems & provide best user experience. ✓

Case Study → del.icio.us (2003)

Bookmarking Platform

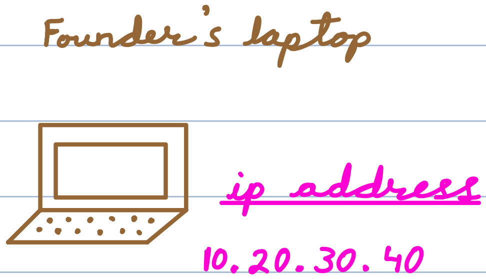
1) addBookmark(userId, url)
2) viewAllBookmarks(userId)



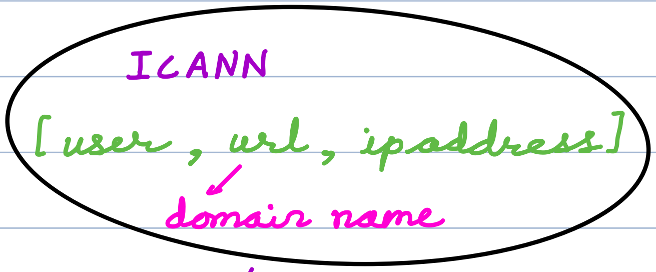
How can users interact ?

Host a website

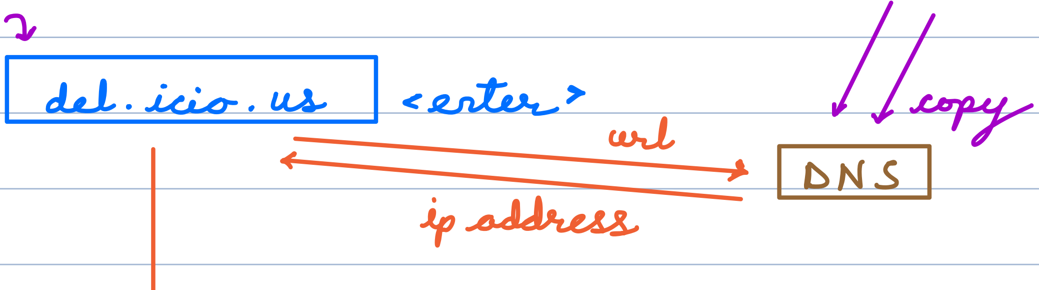
del.icio.us <enter>

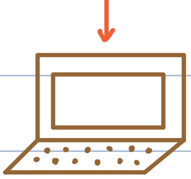


godaddy,
namecheap,
hostinger, ...
} Broker



user ↴





ip address

10.20.30.40

Scaling

1 M user requests per day to store bookmarks.

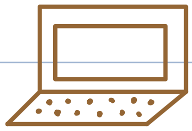
addBookmark (userId, url)

amazon.com/.../...

~1KB

data to be stored per day → 1 M * 1KB

$= 10^6 * 10^3 B = 10^9 B =$ 1GB



(~2004-05)

storage → 60 GB

storage full < 60 days

1) Purchase super computers with large storage.

storage → 128 GB

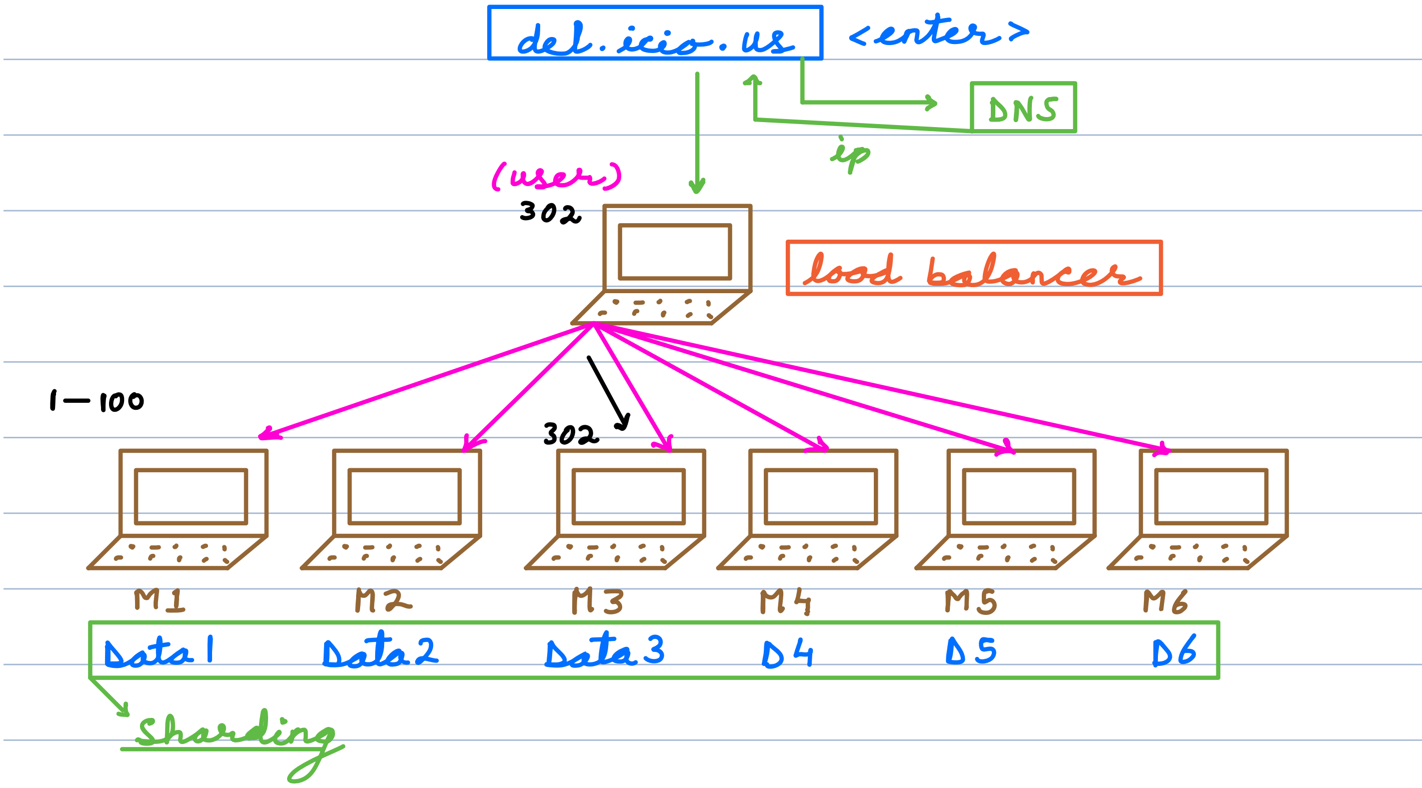
storage full < 128 days

Vertical Scaling

1) There is a limit wrt how much we can upgrade.

2) Single point of failure.

Horizontal Scaling



Challenges

1) load is distributed evenly
2) LB is able to re-direct the request to the right server.
3) Add / Remove servers.

Consistent Hashing