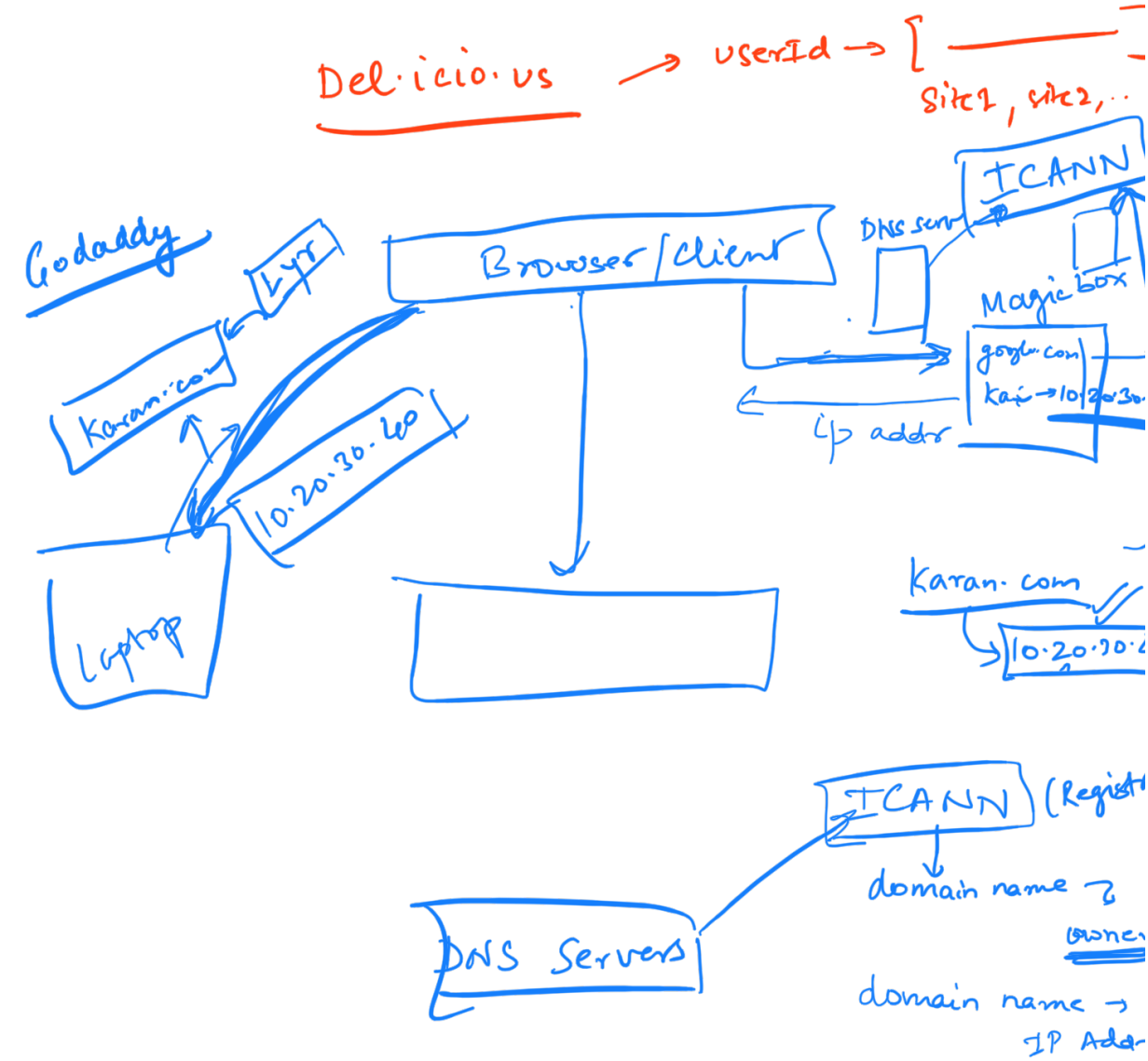
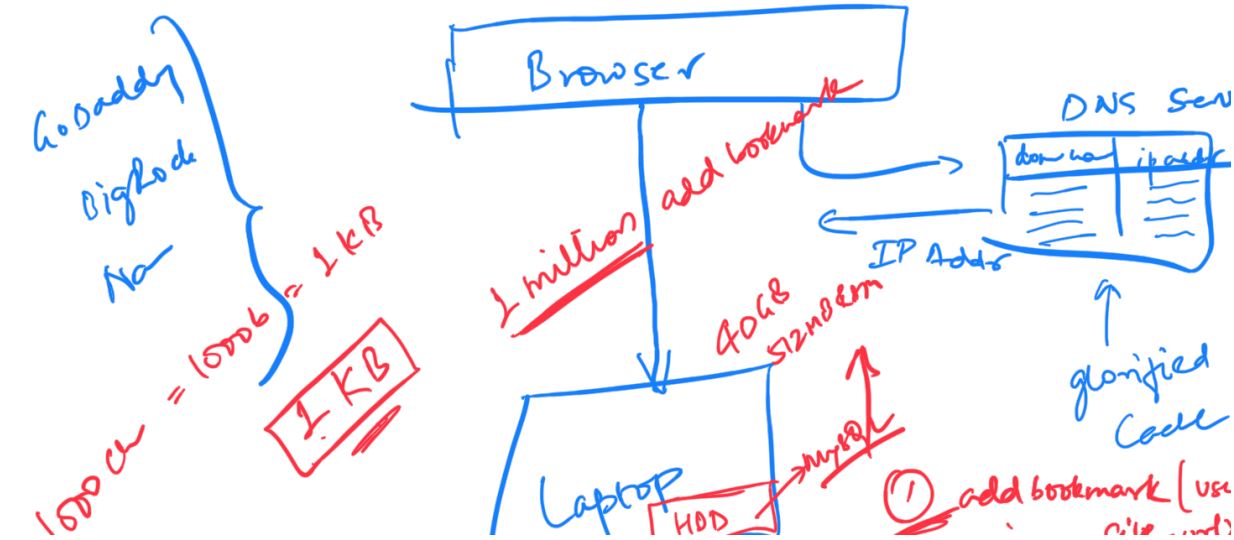
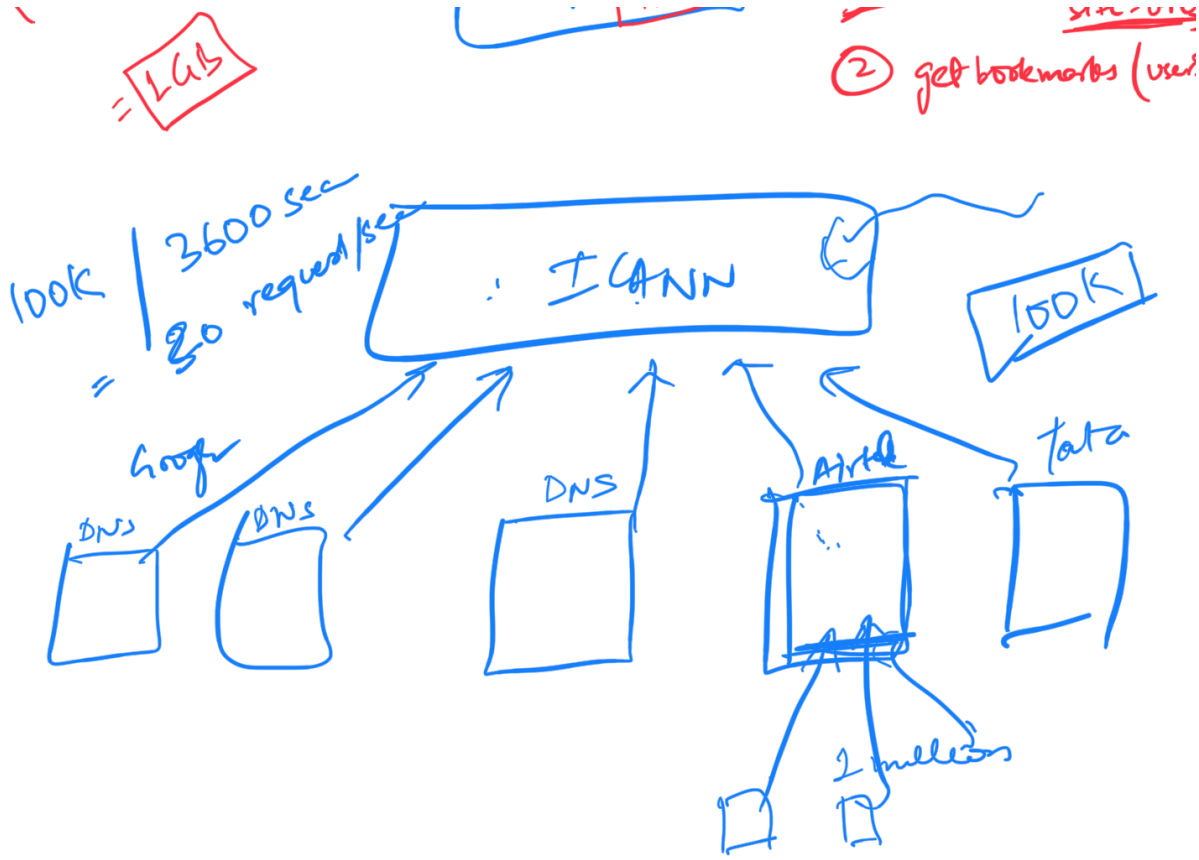
HLD-I



Del.icio.us
userId
site1, site2, ..
Godaddy
Browser/Client
ICANN
DNS serv
Magic box
google.com
Kai → 10.20.30.
ip addr
Karan.com
10.20.30.40
Laptop
Karan.com
10.20.30.
ICANN
(Regist
domain name
owner
DNS Servers
domain name →
IP Add



Browser
DNS Serv
domain
ip addr
IP Addr
1 million add bookmark
40GB
512MB RAM
glorified Cache
Laptop
HDD
mysql
① add bookmark (use
Godaddy
BigRock
Na
1000 ch = 1000b = 1KB
1KB



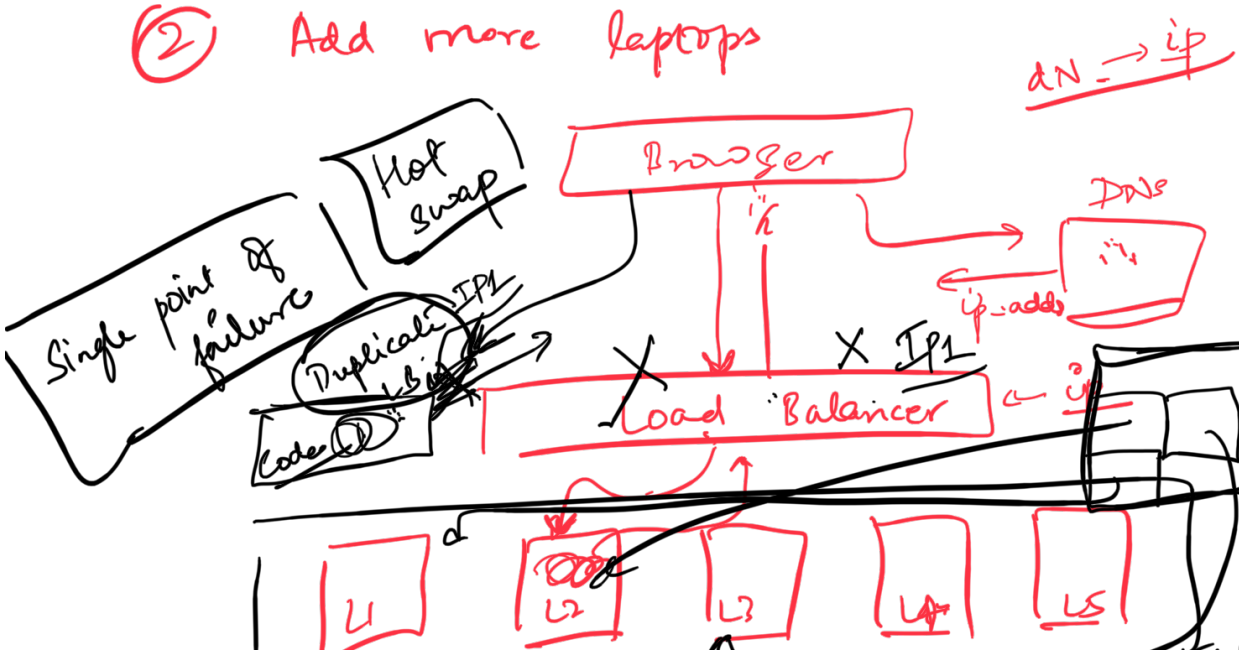
① Buy a better laptop (External HDD)

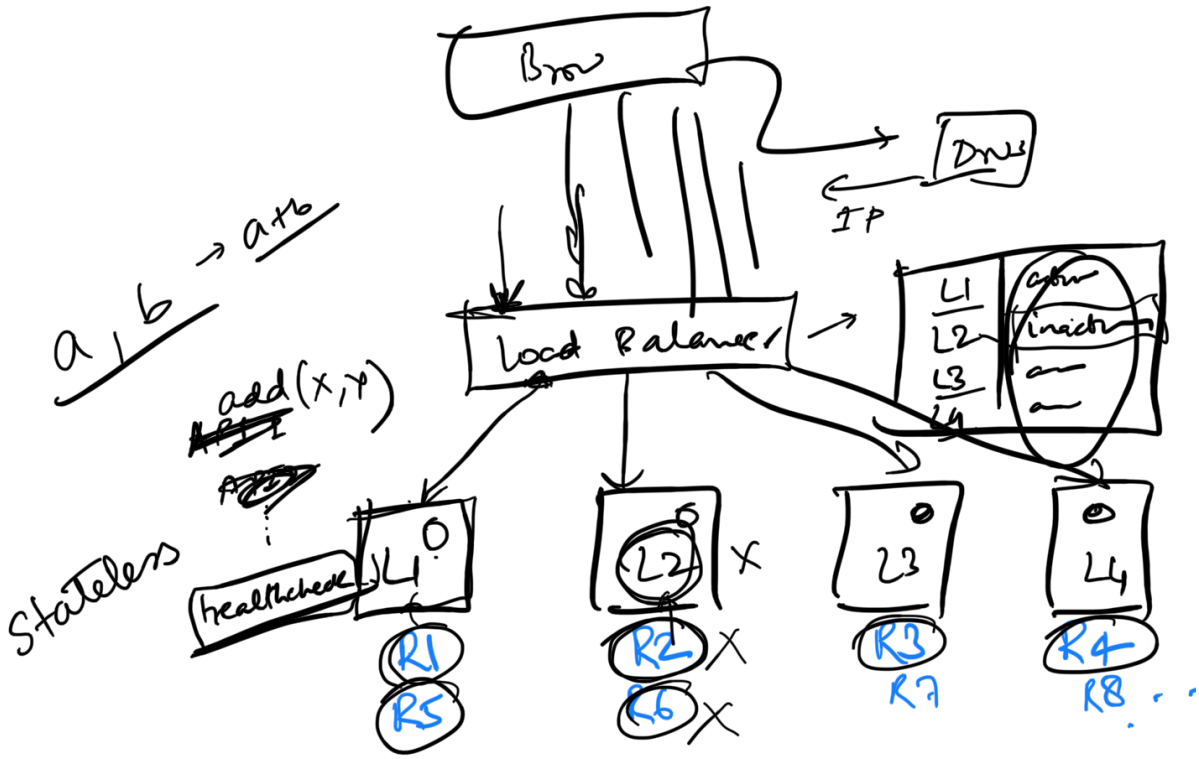
Vertical scaling

40 GB → 100 GB

40 days 100 days

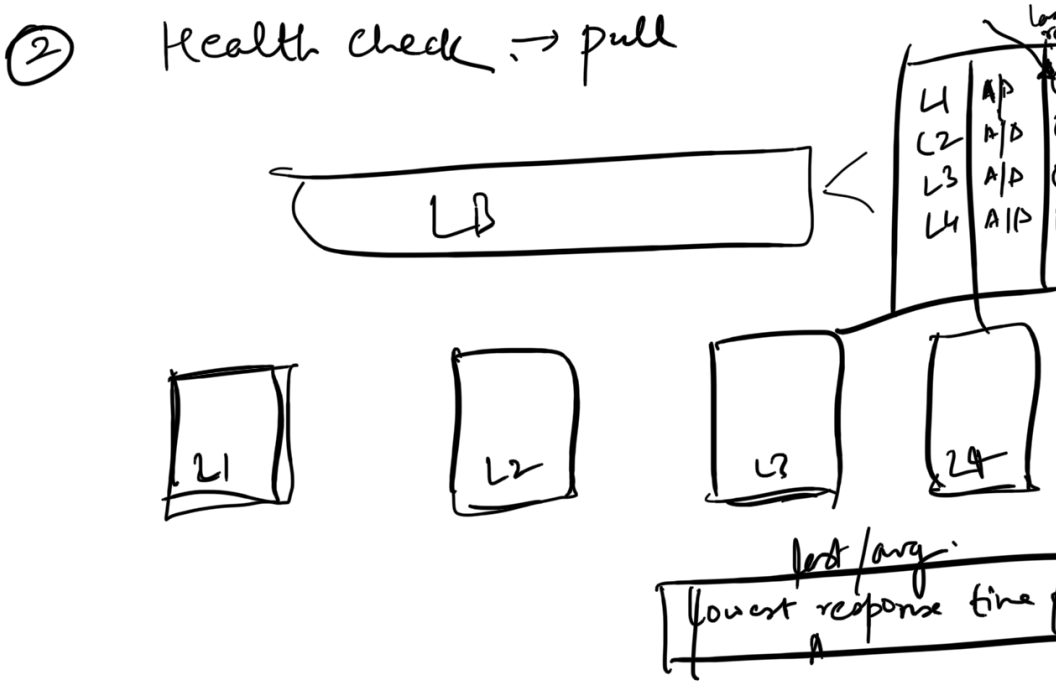
② Add more laptops

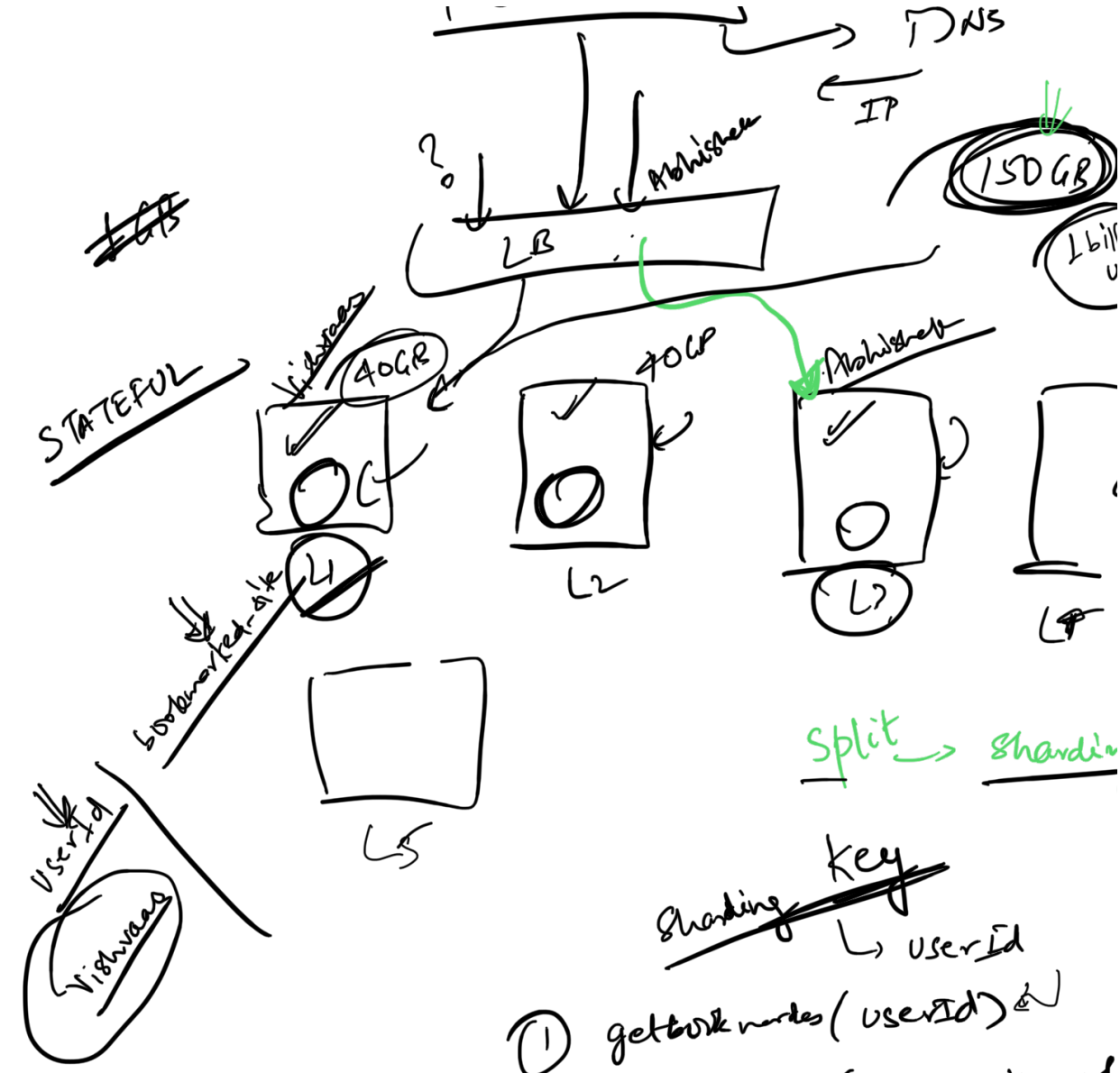




① Heart beat → push Pull / Push

② Health check → pull





Split → Shardin

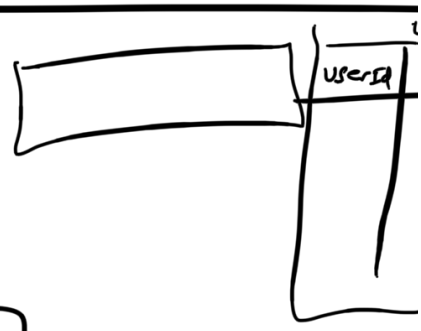
Sharding Key
↳ UserId

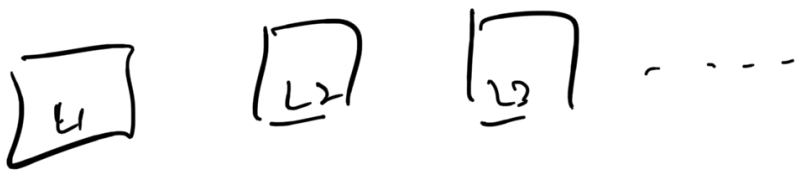
1. getbookmarks(UserId)
2. addbookm(UserId, site_url

---

1 billion user

User ⇒



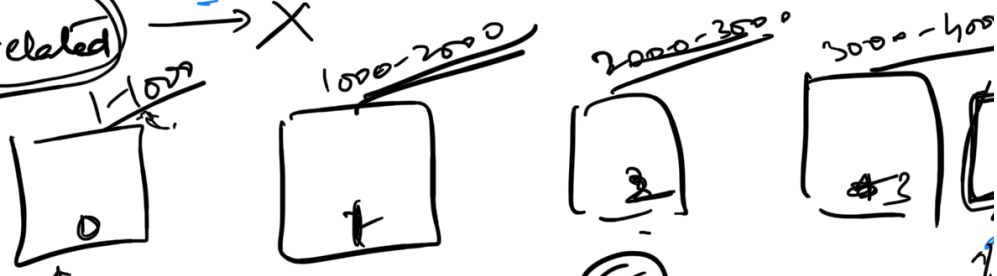


1. UserId % no. of machine

- get book (UserId
- addbook (UserId

② Range 1 - 1000 → 1
1001 - 200 → 2

③ Geo related → X



⑤

12	0	2
14	2	4
19	3	4
26	2	1
18	2	3
38	2	3
39	3	4
40	0	0
15	1	0

1-250
1000-125
2000-220
⋮

Range	machine
⋮	

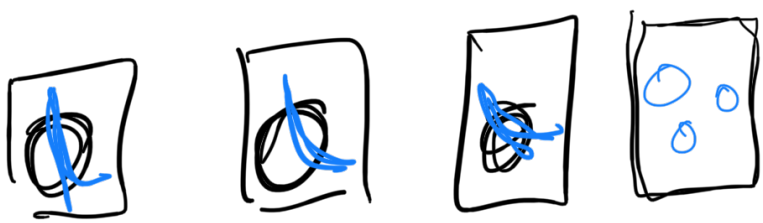
Requirement

① LB

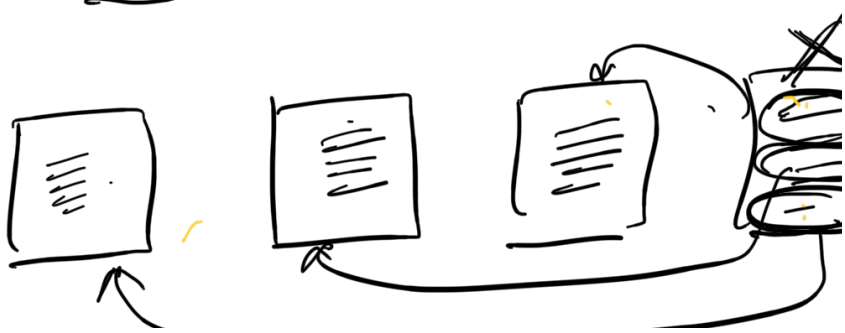
UserId → machine

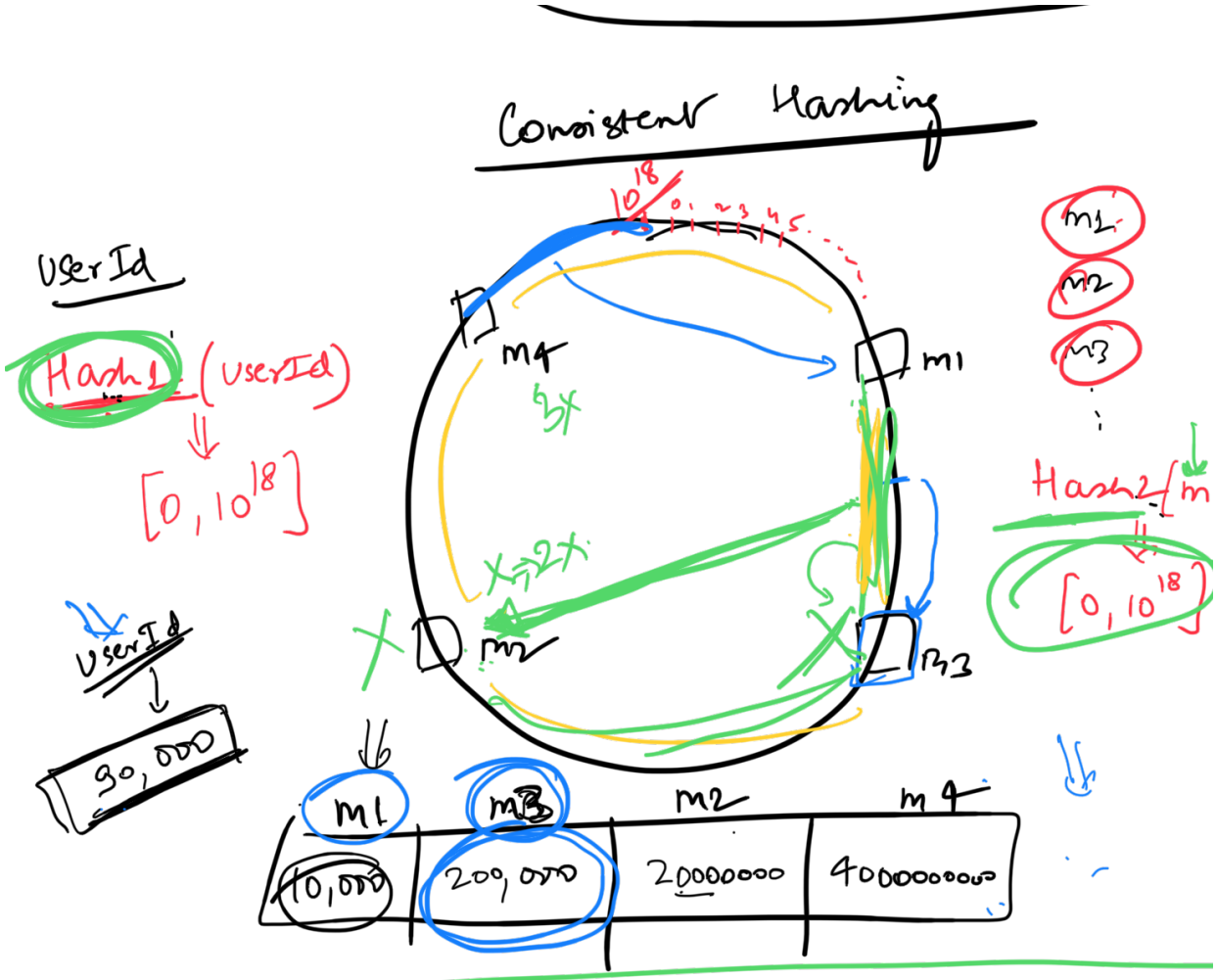
less memory
less CPU

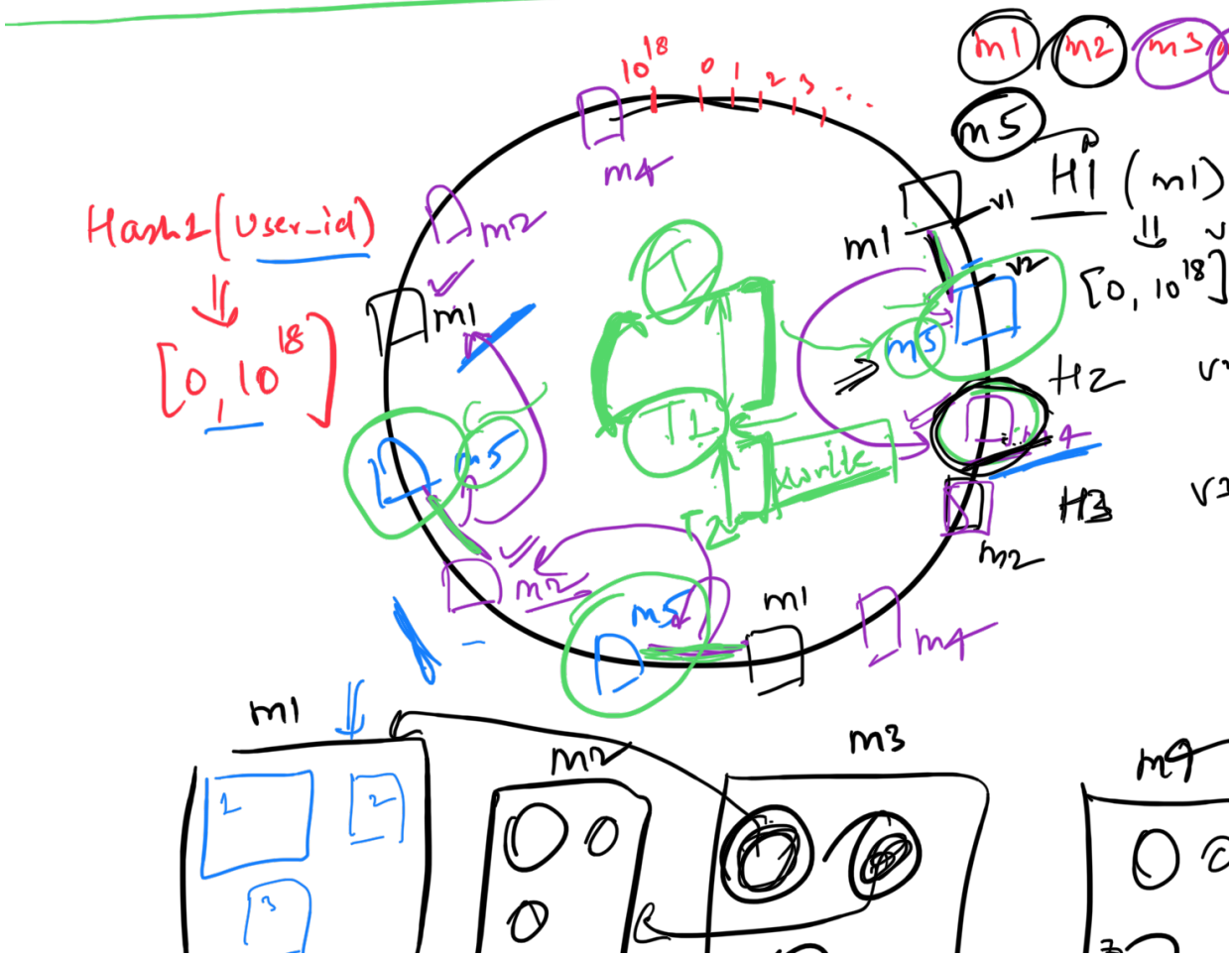
②

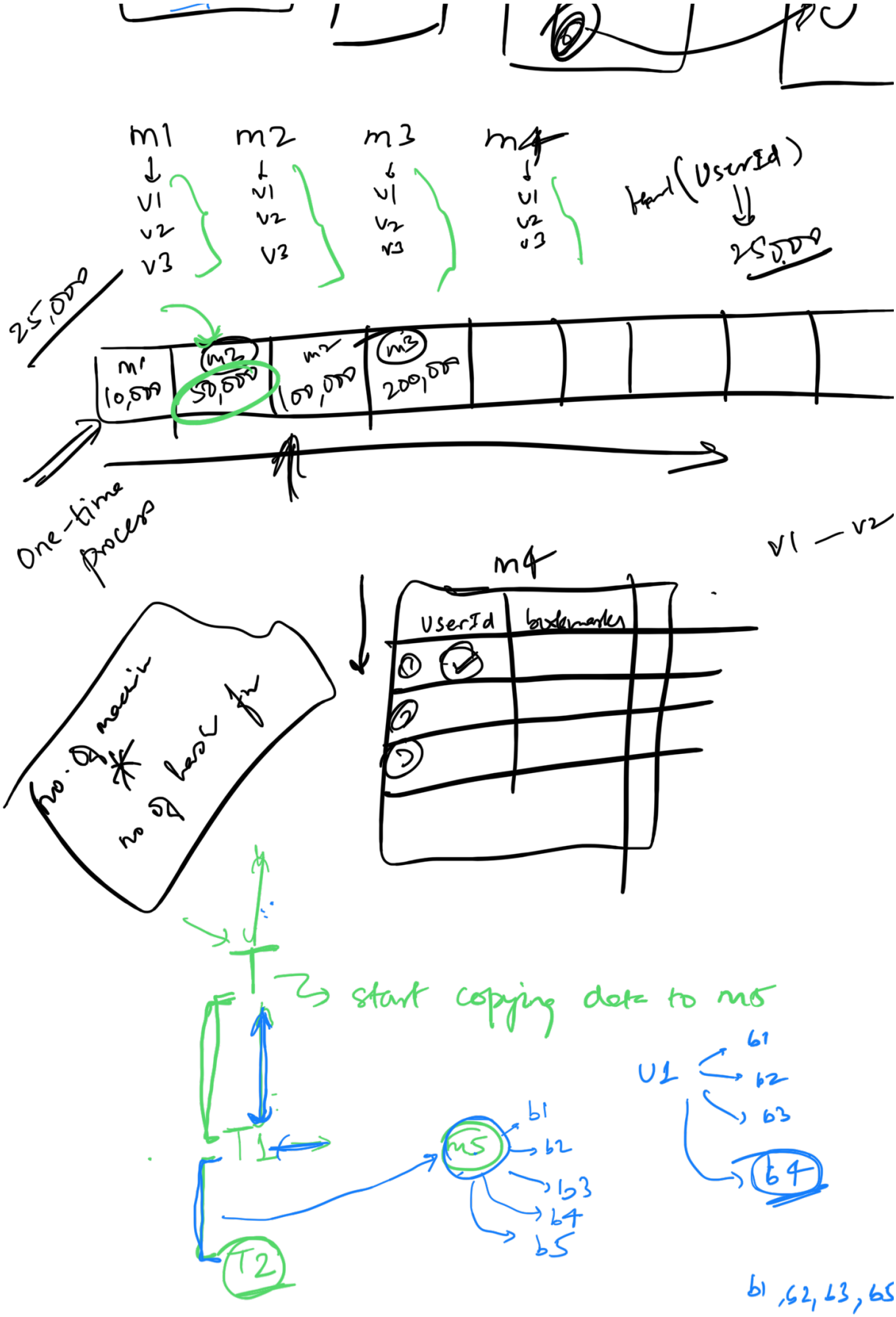


③




Consistent Hashing
UserId
Hash1 (userId)
[0, 10^18]
m1
m2
m3
m4
Hash2(m
[0, 10^18]
UserId
30,000
m1
m3
m2
m4
10,000
200,000
20000000
400000000


Hash2(user_id)
[0, 10^18]
m1
m2
m3
m5
H1 (m1)
[0, 10^18]
H2
H3
m4
m1
m2
m3
m4


m1
m2
m3
m4
V1
V2
V3
V1
V2
V3
V1
V2
V3
V1
V2
V3
hash(UserId)
25,000
25,000
m1
10,000
m2
50,000
m2
100,000
m3
200,000
One-time process
V1 — V2
m4
UserId
bookmarks
start copying data to m5
T1
T2
m5
b1
b2
b3
b4
b5
U1
b1
b2
b3
b4
b1, b2, b3, b5