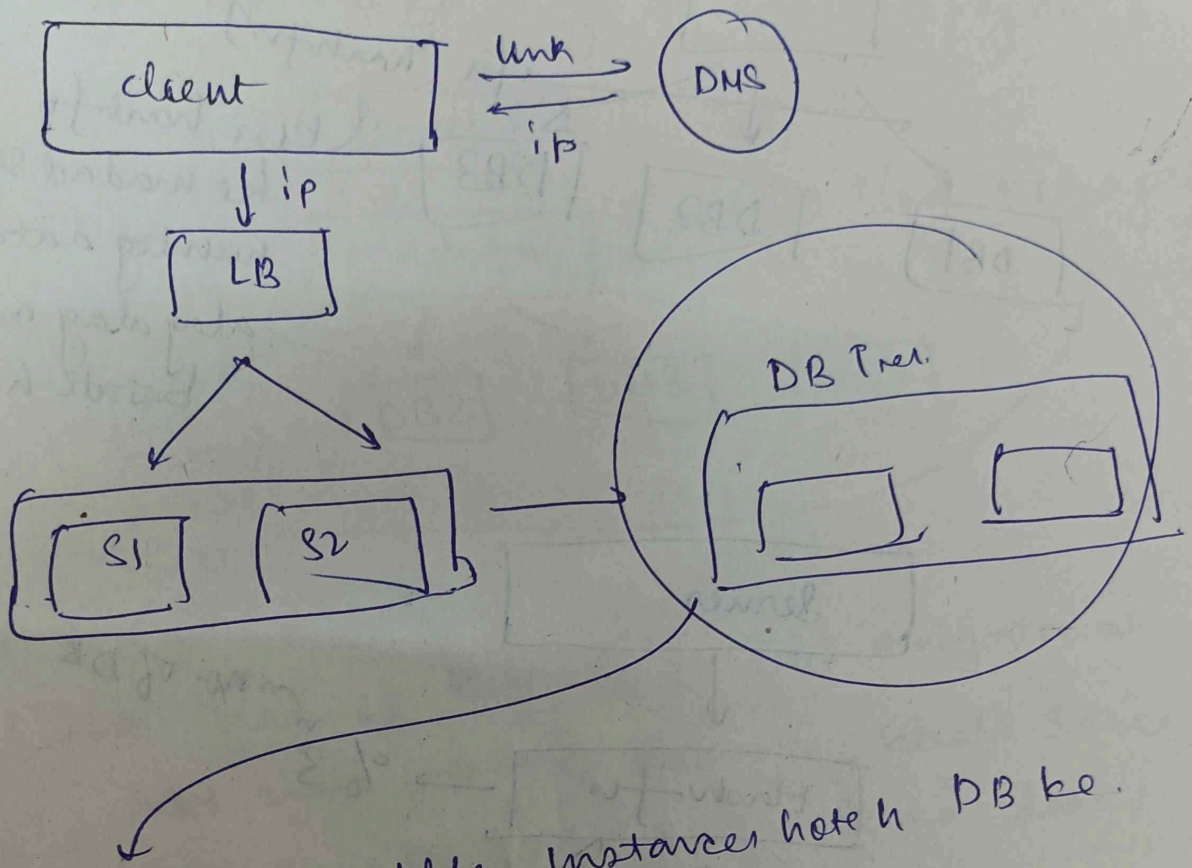
Lec04:- DB Sharding - Consistent Hashing - Load Balancer.

DB Sharding

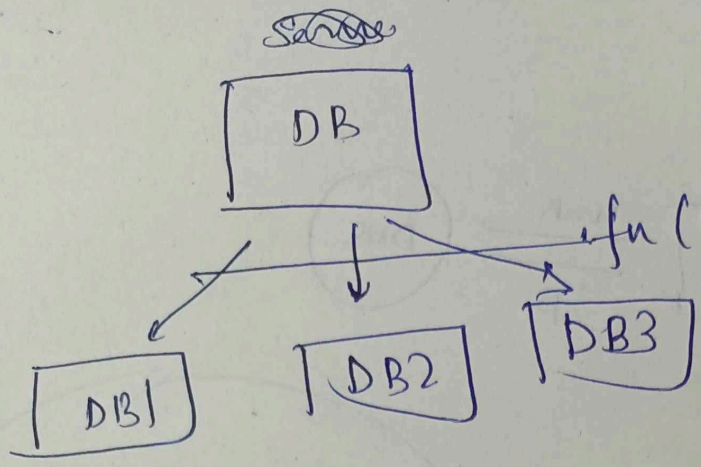


Hamare pas multiple instances hote h DB ke.

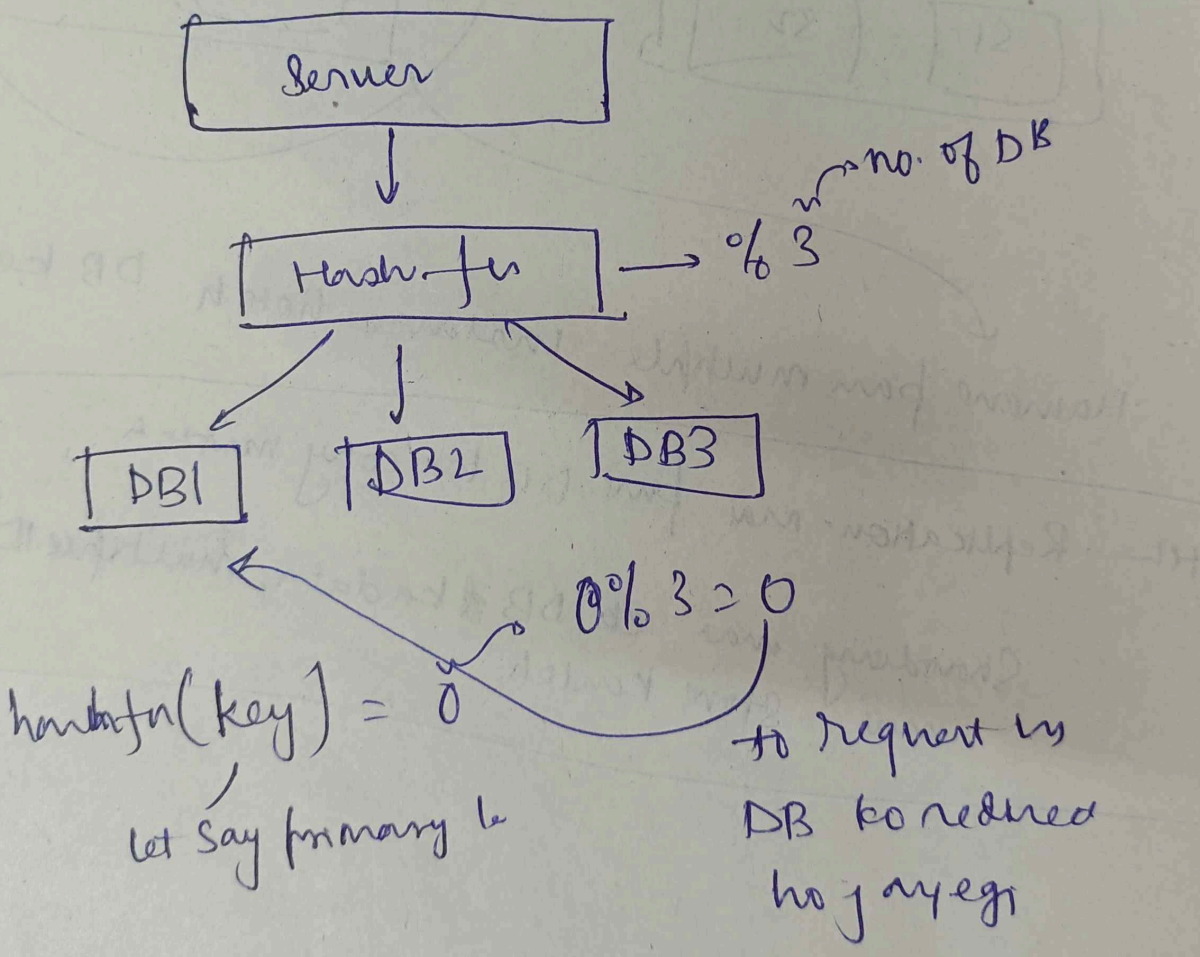
#

Note:- Replication mai pure DB ka copy marte h.

Sharding mai ek DB ka data, multiple DB mai store karte h.



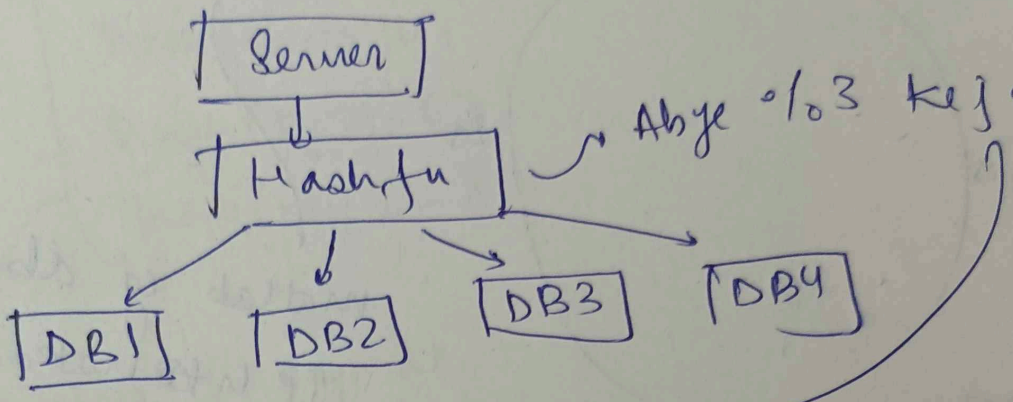
fn (hashfn)
(kisi hashfn ke madad se. numlog data to alag alag meri bataté h



Sharding ke limitations. 18:00

① Scaling

(Agar mai kal ko naya db laya (Lets say db4)

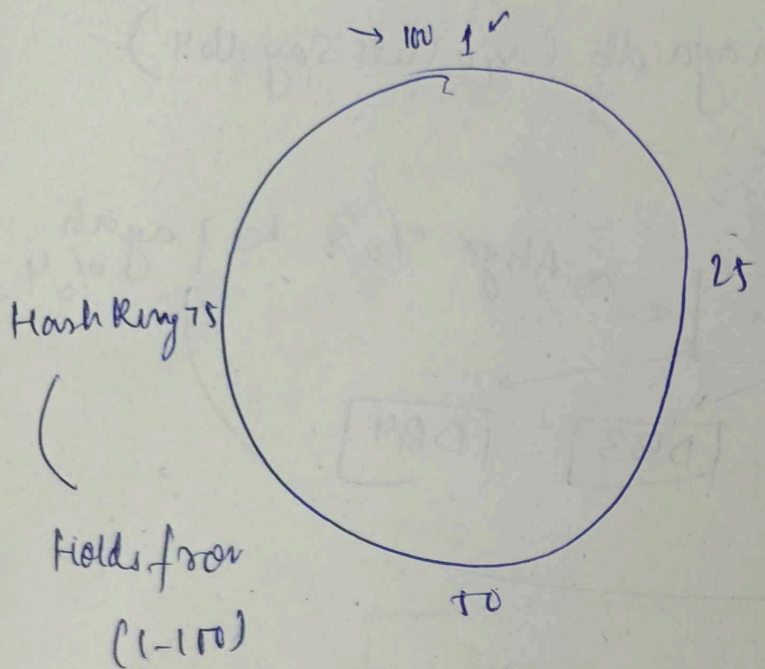


hoga

Isse kaaran sg key 3 jo DB1 mai jata abb DB3 mai jayega, aur aur kaafi saare migrations honge, kaafi saara data yaha ka waha hoga.

aur saare agar mai kor db hata du.

Consistent Hashing



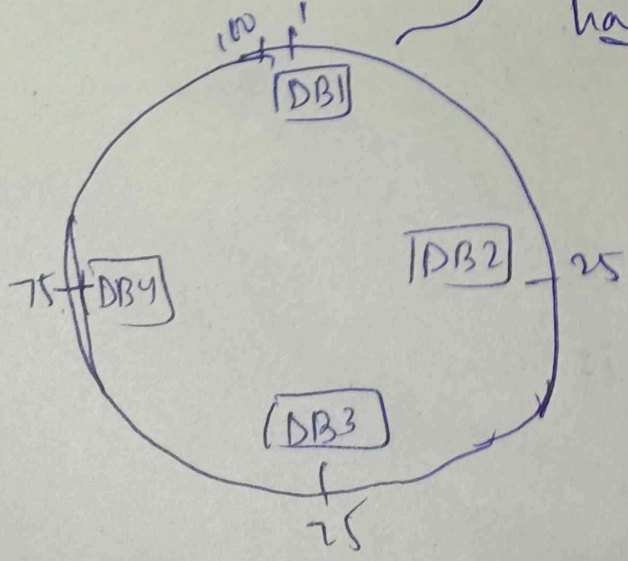
Sabse phle hum DB ka hash lenge

matlab eg db ka IP ka uska hash le lenge

hash(IP) ∈ (1, 100)

assume

Let's say hash(DB1 ka IP): 1
hash(DB2 ka IP): 25
hash(DB3 ka IP) = 50
hash(DB4 ka IP): 75
hash(DB4 ka IP): 1

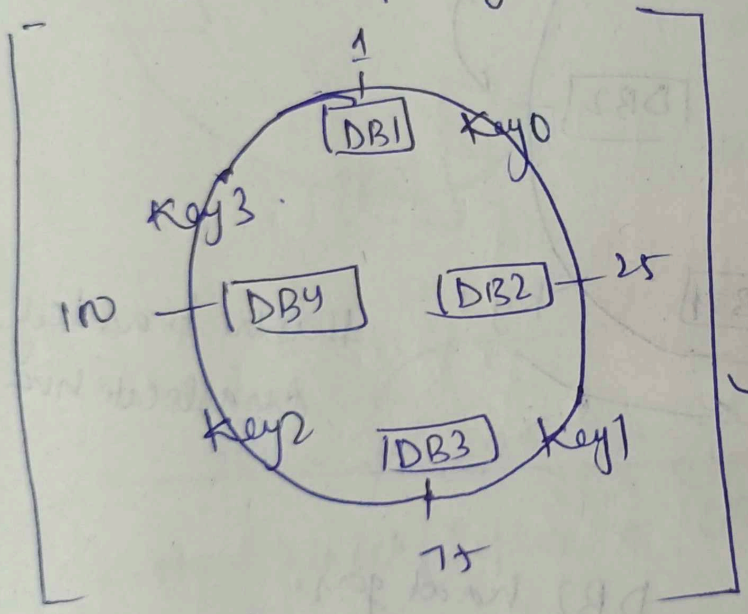


Iska perfect ho ye zaruri nhi.

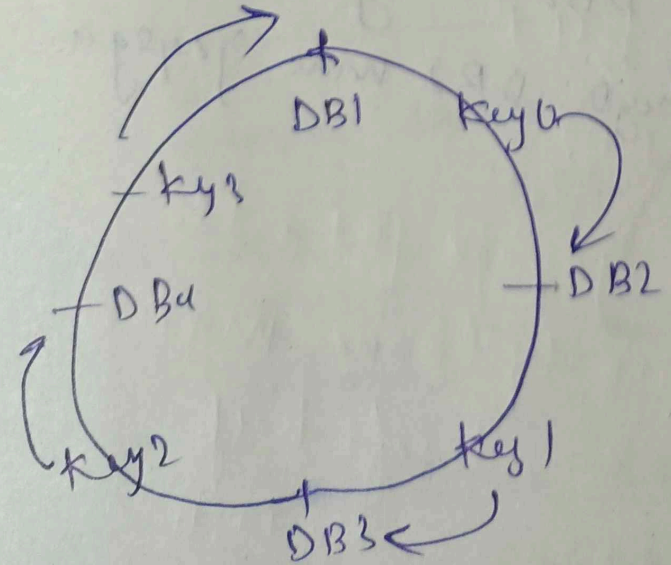
ab say Databases place ho gaye tab hum key's ka hash lenge jo 1-100 ani ke karenge aur usse circle pe place karenge.

eg. hash(key0) = 11
hash(key1) = 41
hash(key2) = 67
hash(key3) = 80.

ab circle pe laga de.



ab clockwise mai move karna h aur jo DB aaya pehle uss key ko uss DB mai daal dena h



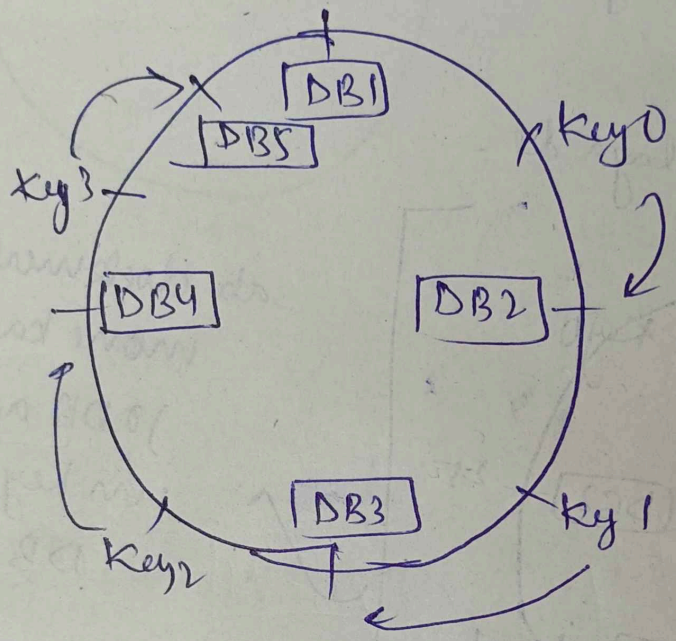
key0 gaya DB2 mai.
key1 gaya DB3 mai
key2 gaya DB4 mai.
key3 gaya DB1 mai

Advantages

→ Kalnaya DB aaya to kaafi kam keys deallocate honge.

↳ lets say db5 laa raha.

hash(db5 ka IP) = 85



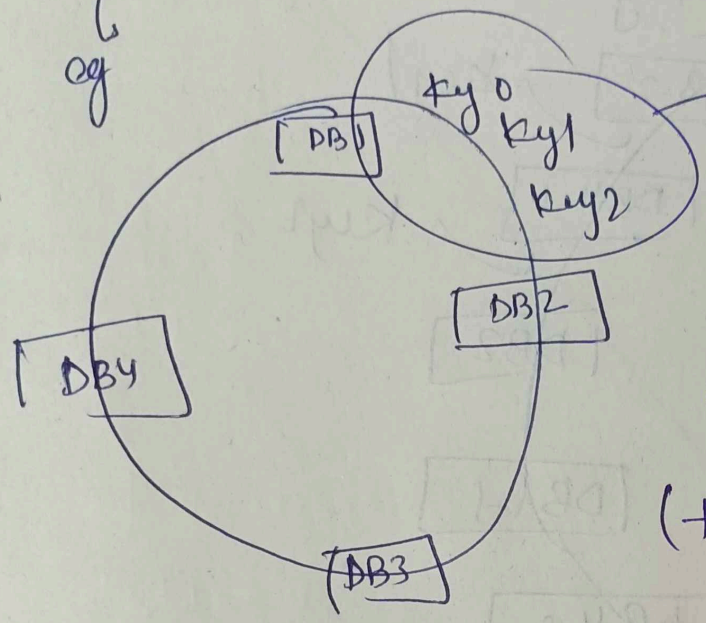
#Abhi bas key3 deallocate hua

aur lets say DB2 hat gaya.
to bas key0 DB2 mai jayega.

Disadvantage

→ DB mai jo no. of keys jaa rahe uno generally uniform ni hote.
(Uneven distribution of keys)

eg



→ Let's say maine key ka hash nikala aur teeno isi range mai hi aa gaya.

(to yhi dikkat hki uniformity compromise hojati h)

Iska soln banaya gaya h.

Virtual Nodes (Replicas of Original Nodes)

humlog databases ka replica banayenge

DB1 → DB1-1
DB1-2
DB1-3

DB2 → DB2-1
DB2-2
DB2-3

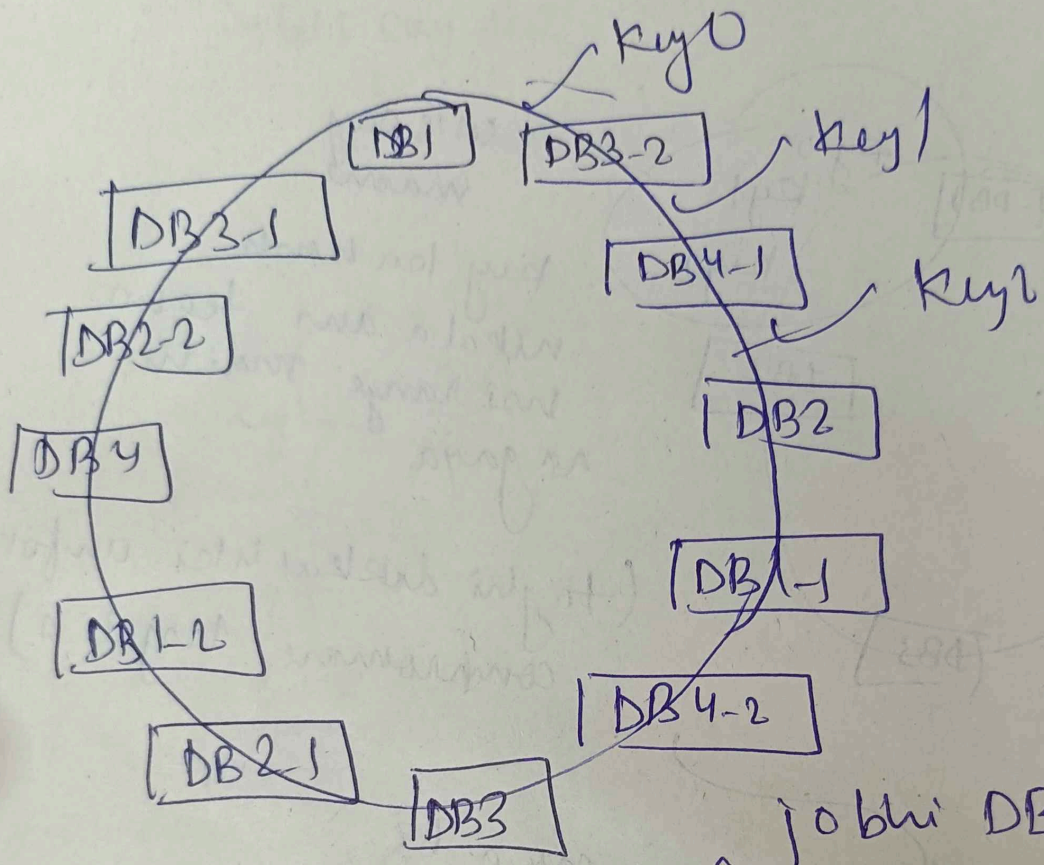
DB3 → DB3-1
DB3-2
DB3-3

OR DB ke do replica h,

Note: Data DB_1 ya DB_2 ya DB1 mai jaye, at the end teeno mai rahega

Ab mai wapas se fir replica DB ka hash nikalenge

(let's say mo ye aaye.



jo bhi DB3 hi hat ta end.

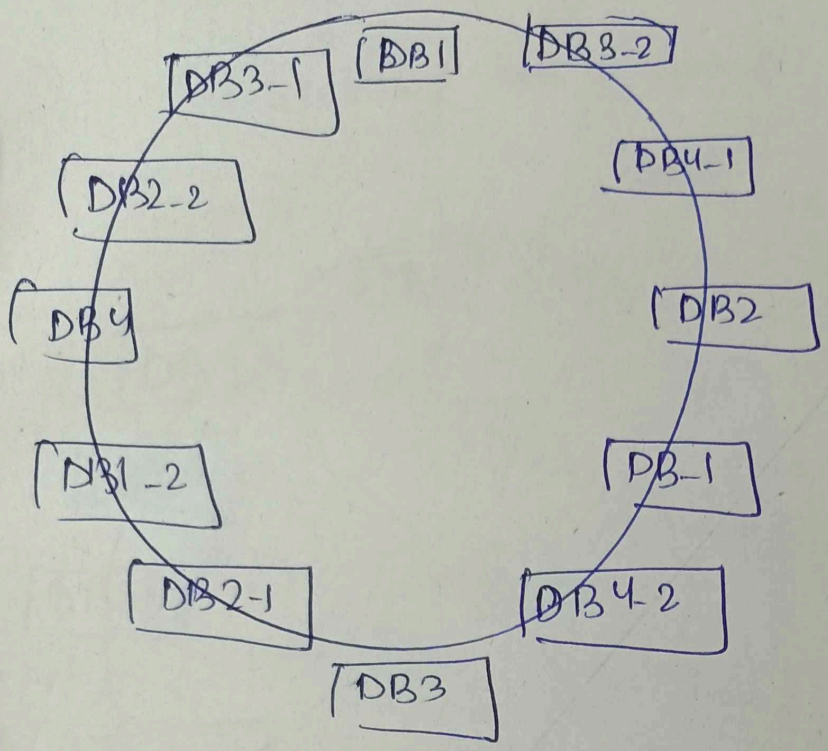
ab key0 DB3-2

key1 DB4-1

key2 DB2

mai jayega jo ki ~~th~~ uniformity laane mai kafi help karta h.

Celebrity Problem



→ famous kuch keys ek particular db mai aa jate h, jisse no. of read operation us DB pe kaafi badh jata h.

foreg:- Imta pe. SRK, Virat Kohli, Ronaldo, Modi kadata agar ek hi DB mai gaya to us DB pe no. of read operation kaafi badh jayeg

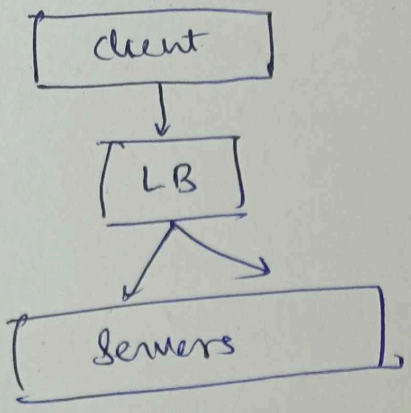
for egg

Key4: SRK
Key5: Virat Kohli
Key6: Ronaldo
Key7: modi

→ lets say ye saare keys DB 3 mai chale gayi, to kyunki inki profile to kaafi log fetch karte h to DB3 pe to kaafi jayda load padega.

load Balancing

passport serv



LB itna Smart kaise hota h?

Uska kaise pata chal jata h ki kaun se req ko kaha bhejna h

Applications of LB

- Scalability
- Availability
- Security
- Responsiveness

1:11:34