CS & IT ENGINEERING

Data Structure

Hashing
Lec-01

Chapter 07

By- Pankaj Sharma sir

TOPICS TO BE COVERED

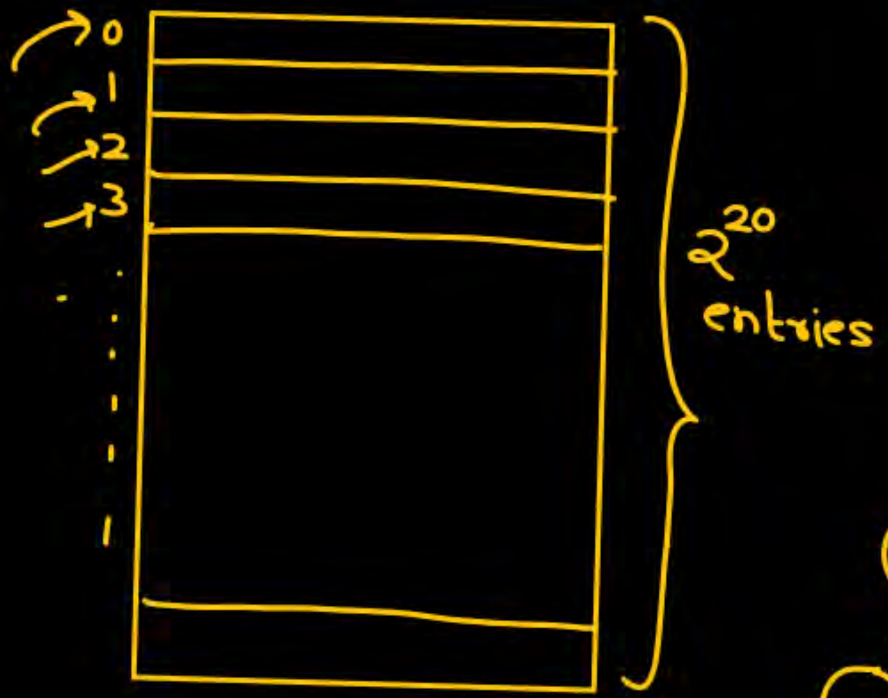
Hashing-I

DBMS

Hashing

1. $n = 2^{20} = 2^{10} \times 2^{10} \Rightarrow 1024 \times 1024 \cong 10^6$

comparision = $n = O(n)$



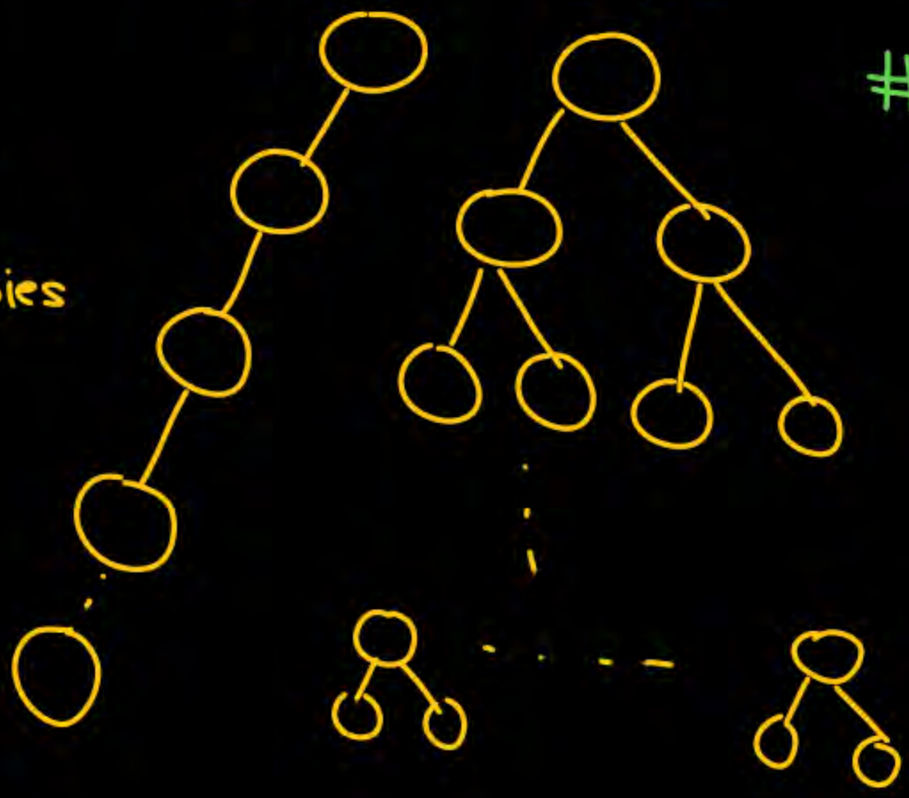
2. $\boxed{\text{BST}}$

#comp = $O(h)$

skewed → $O(n)$ ✓

balanced → $O(\log_2 n)$

Case I Case II



3. AVL-Tree (Bal. BST)

#comp = $O(\log_2 n)$

$\log_2 2^{20} = 20$

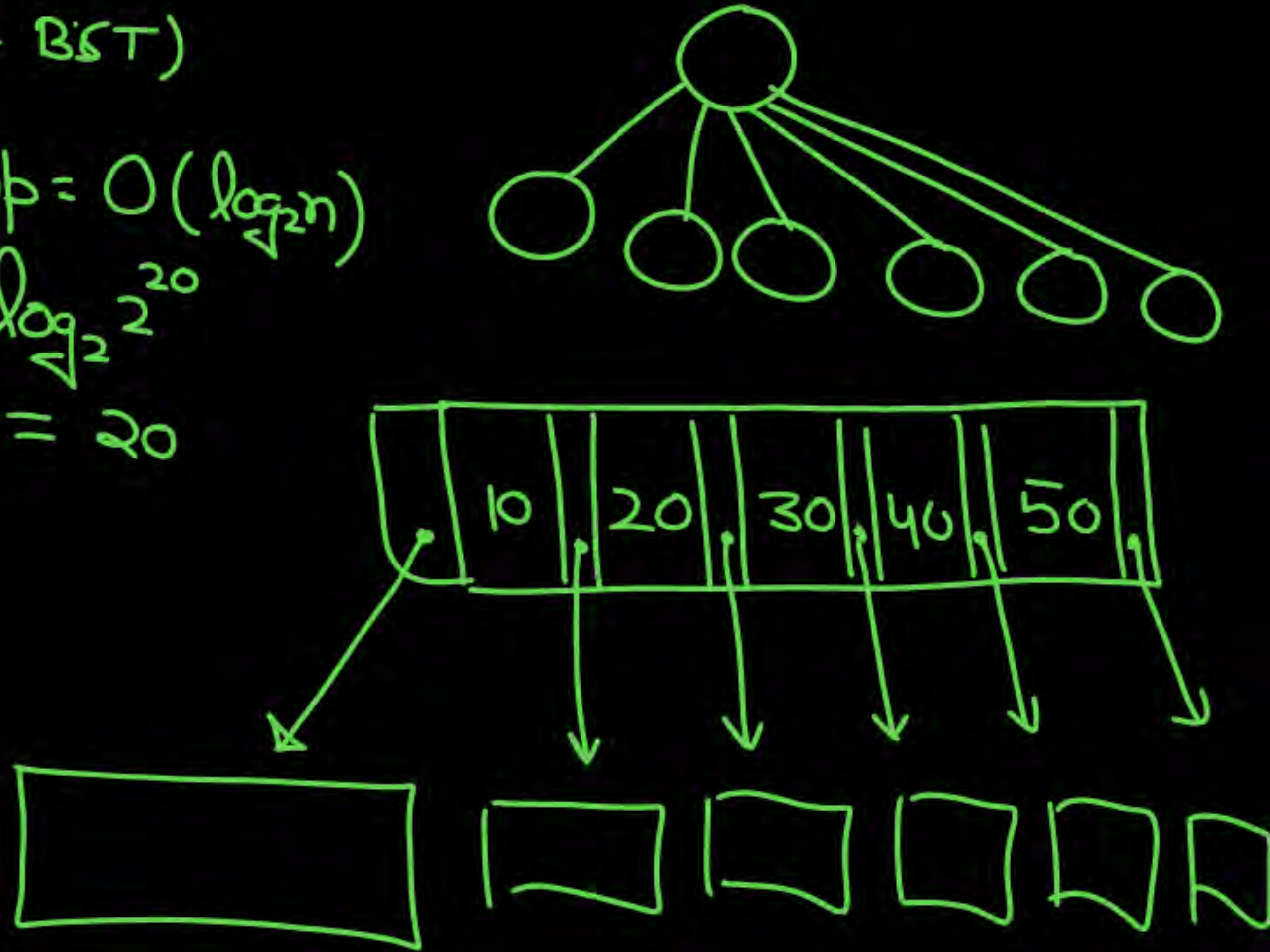
order 8

$O(\log_m n)$

$\log_8(2^{20}) \Rightarrow \frac{20}{3} \approx 7$

4. B-Tree & B+-Tree

m : Order



$y = f(x)$ input

$O(1)$

Keys : 12, 11, 18, 15, 13, 19, 16, 14

$m = 10$

$h(12) = 12 \bmod 10 = 2$

$h(11) = 11 \bmod 10 = 1$

12 →

Key → Hash function → (0-9)

Location

$$\boxed{h(k) = k \bmod 10}$$

0	
1	11
2	12
3	13
4	14
5	15
6	16
7	
8	18
9	19

$y = f(x)$ input

$O(1)$

$m=10$ Keys : 12, 14, 16, 32, 42, 24, 36

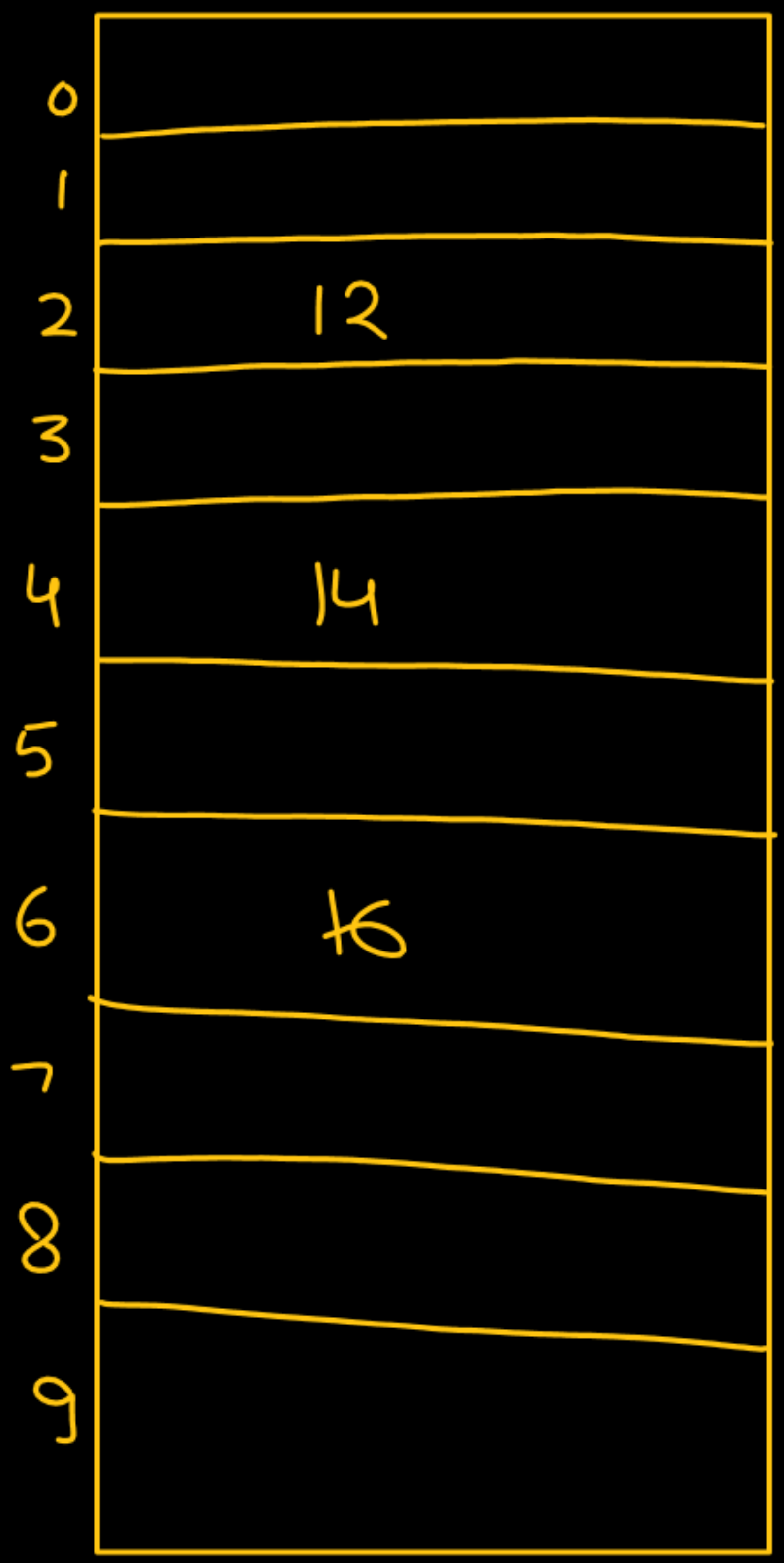
$h(12) = 2$

$h(14) = 4$

$h(16) = 6$

$h(32) = 2$

Collision



$$\boxed{h(k) = k \bmod 10}$$

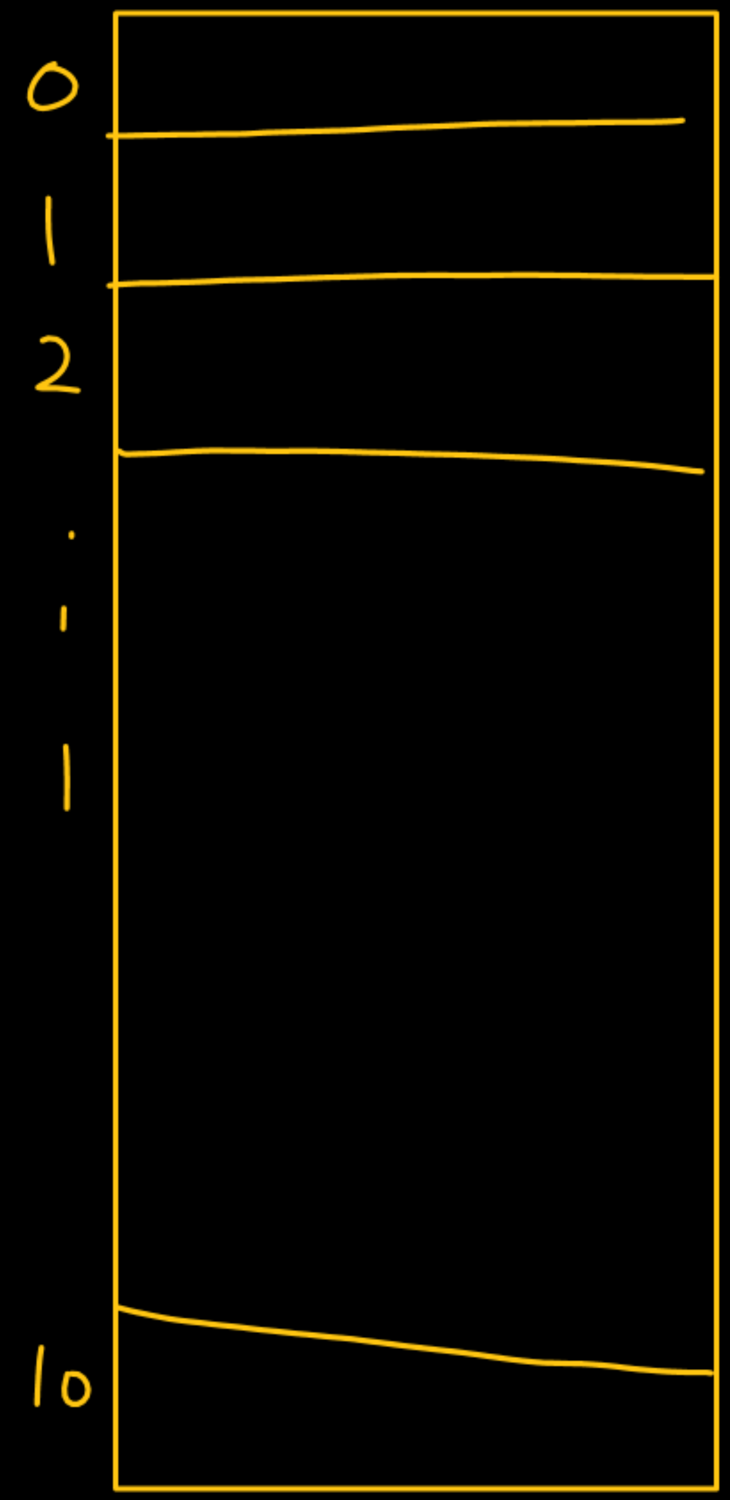
$h(k_1) \rightarrow L_1$

$h(k_2) \rightarrow L_1$

Collission

Good Hash Function

1. Easy to compute
2.



Collision

Collision Resolution Technique

- Linear Probing
- Quadratic Probing
- Double Hashing
- Separate chaining

Hash function

$$h(k) = k \bmod m$$

$$0, 1, 2, \ldots m-1$$

$$h(k) = (k \bmod m) + 1$$

$$1, 2, 3, \ldots m$$

Linear Probing

Let $h(K) = K \bmod m$

m = Table size

$h(k) = k \bmod m = L_1$] collission

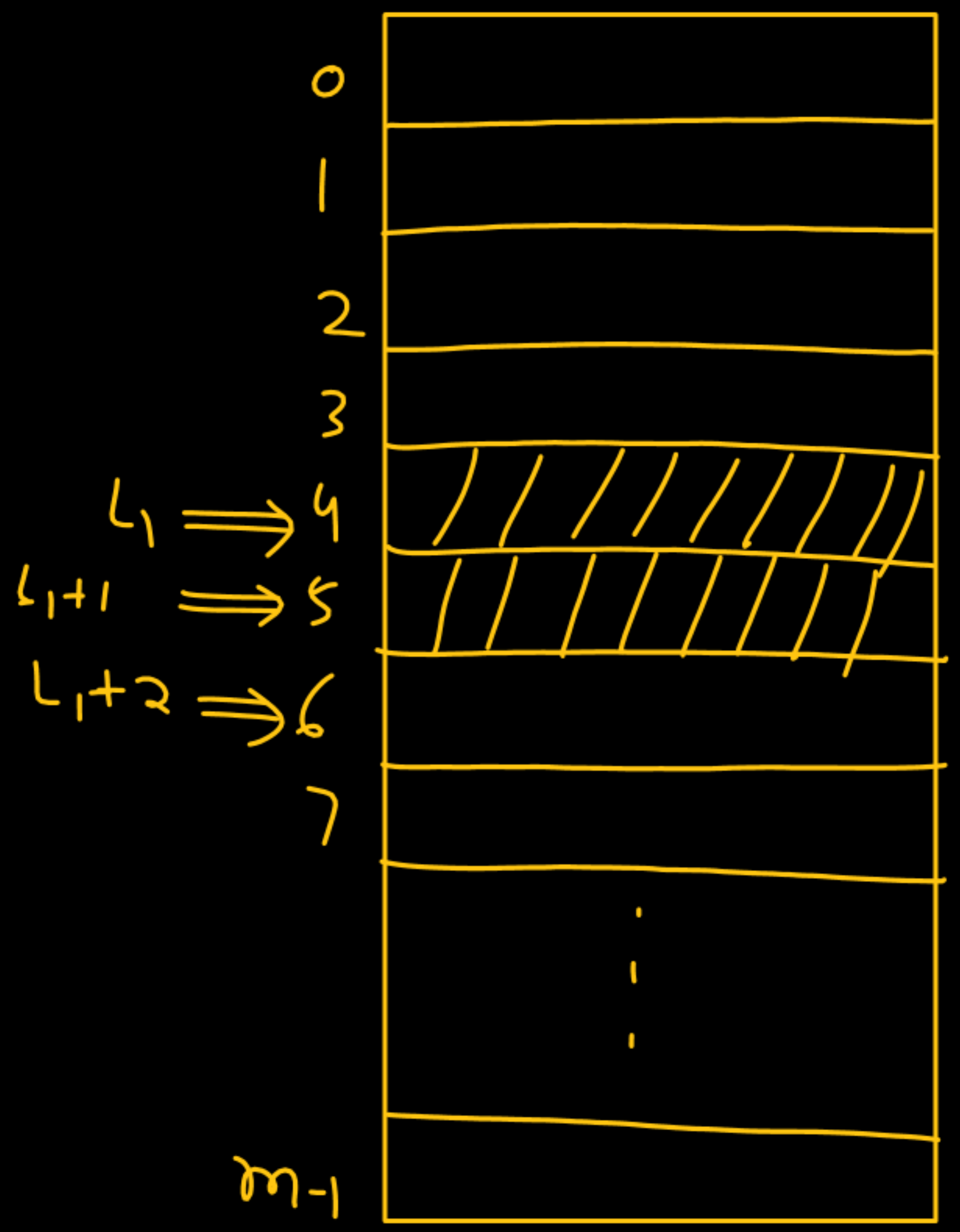
$H(k, i) = (h(k) + i)$

k → key

i → collission no. for particular key

$H(k,1) = (h(k) + 1) = L_1 + 1$

$H(k,2) = (h(k) + 2) = L_1 + 2$



$m=10$

Linear Probing

$h(k) = 4$ colission

$H(k,1) = (4+1) = 5$ ×

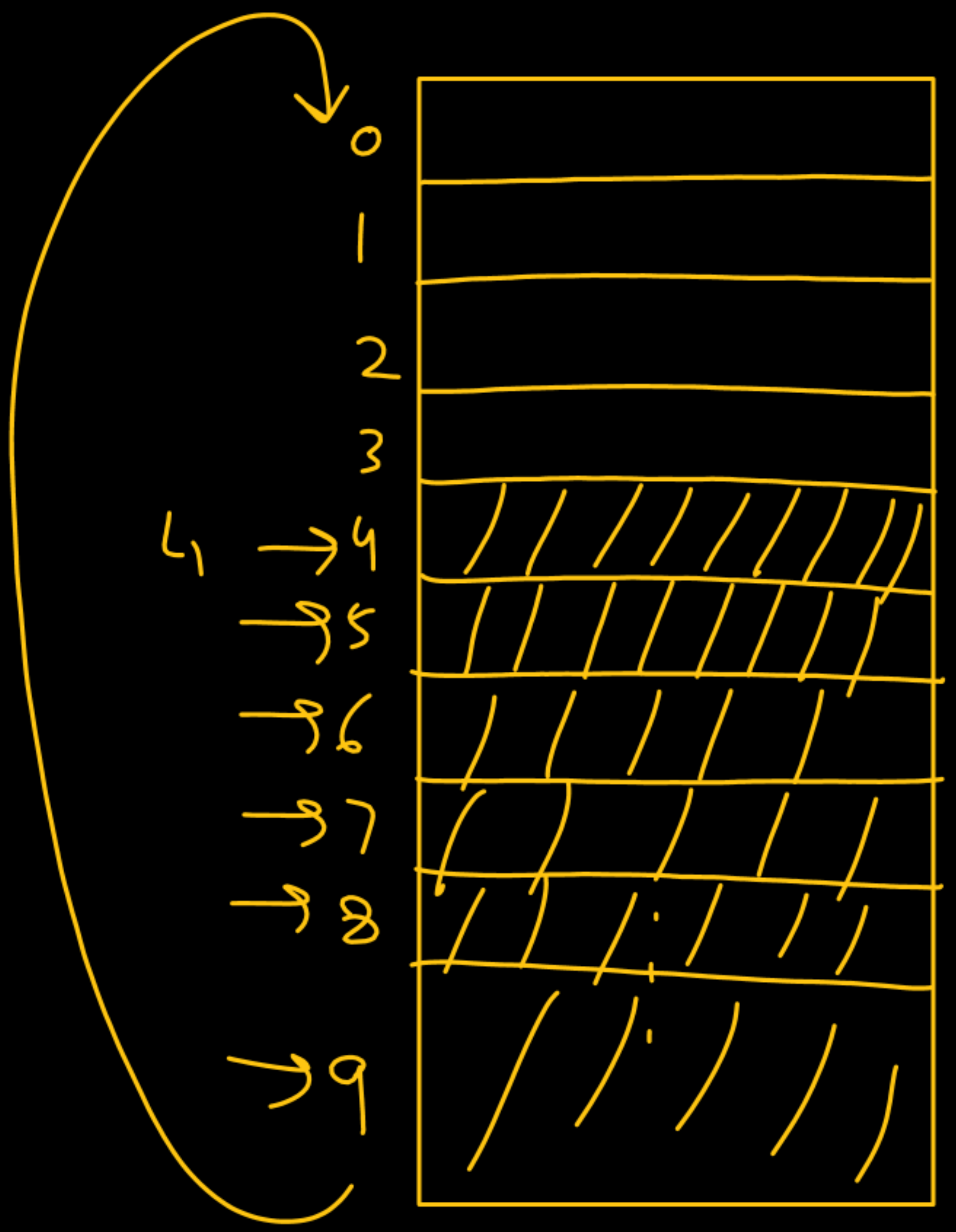
$H(k,2) = (4+2) = 6$ ×

$H(k,3) = (4+3) = 7$ ×

$H(k,4) = (4+4) = 8$ ×

$H(k,5) = (4+5) = 9$ ×

$H(k,6) = (4+6) = 10 \bmod 10 = 0$



Keys : 31, 26, 43, 27, 34, 46, 14, 58, 13

$m = 12$

$h(31) = 31 \bmod 12 = 7$

$h(26) = 26 \bmod 12 = 2$

$h(43) = 43 \bmod 12 = 7$ collission

$H(x, i) = (h(x) + i) \bmod m$

$H(43, 1) = (h(43) + 1) \bmod 12 = 8$

$h(27) = 27 \bmod 12 = 3$

$h(34) = 34 \bmod 12 = 10$

$h(46) = 46 \bmod 12 = 10$ coll.

$H(46, 1) = (h(46) + 1) \bmod 12 = (10 + 1) \bmod 12$
$= 11$

$h(14) = 14 \bmod 12 = 2$ collission

$H(14, 1) = (h(14) + 1) \bmod 12 = (2 + 1) \bmod 12$
$= 3$ collission

$H(14, 2) = (h(14) + 2) \bmod 12 = (2 + 2) \bmod 12$
$= 4$ ✓

2nd collission for key 14

$h(58) = 58 \bmod 12 = 10$ collision

$H(58, 1) = (h(58) + 1) \bmod 12 = 11$ coll.

$H(58, 2) = (h(58) + 2) \bmod 12 = 12 \bmod 12$
$= 0$

$h(13) = 13 \bmod 12 = 1$

0	58
1	13
2	26
3	27
4	14
5	
6	
7	31
8	43
9	
10	34
11	46

m = 12 7 2 ~~7~~ 8 3 10 ~~10~~ 11 ~~2~~ ~~3~~ 4 ~~10~~ ~~11~~ 0 1

Keys : 31, 26, 43, 27, 34, 46, 14, 58, 13

Problem : Primary clustering Problem

70

70 mod 12

= 10

34, 46, 58, 13, 26, 27, 14

Cluster

0	58
1	13
2	26
3	27
4	14
5	
6	
7	31
8	43
9	
10	34
11	46

m = 12   7   2   ~~4~~ 8   3   10   ~~10~~ 11   ~~2~~ ~~3~~ 4   ~~10~~ ~~11~~ 0   1

Keys : 31, 26, 43, 27, 34, 46, 14, 58, 13

Problem : Primary clustering Problem

$\frac{8}{12} \Rightarrow \frac{2}{3} \Rightarrow 66.67\%$

$\frac{8}{12}$ Prob. that a new elem. will get this slot

$\frac{8}{12} > \frac{1}{12}$

$\frac{8}{12} > \frac{3}{12}$

$\frac{1}{12}$

$\frac{3}{12}$

0	58
1	13
2	26
3	27
4	14
5	
6	
7	31
8	43
9	
10	34
11	46

CRT

Open addressing

1. Linear Probing
2. Quad. Probing
3. Double Hashing

Sep. Chaining

$\boxed{\text{2 PM}}$ ✓

Keys : 24, 17, 32, 2, 13, 50, 30, 61

$m = 11$


THANK YOU GW
SOLDIERS !