Carnegie Mellon University

Hash Tables

Lecture #08

Database Systems
15-445/15-645
Fall 2017

Andy Pavlo
Computer Science Dept.
Carnegie Mellon Univ.

ADMINISTRIVIA

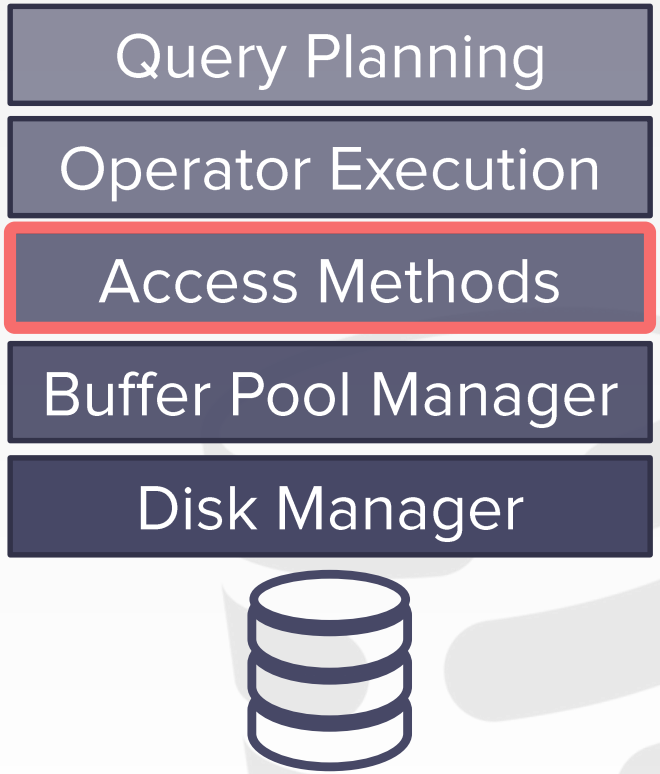
Project #1 is due Monday
October 2nd @ 11:59pm

Homework #3 is due Wednesday
October 4th @ 11:59pm

STATUS

We are now going to talk about how to support the DBMS's execution engine to read/write data from pages.

Two types of data structures:
→ Hash Tables
→ Trees



DATA STRUCTURES

Internal Meta-data
Core Data Storage
Temporary Data Structures
Table Indexes

DESIGN DECISIONS

Data Organization

→ How we layout memory and what information to store inside of the data structure to support efficient access.

Concurrency

→ How to enable multiple threads to access the data structure at the same time without causing problems.

HASH TABLES

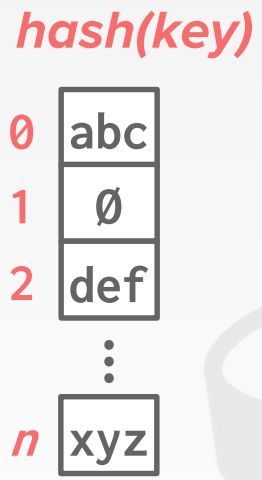
A hash table implements an associative array abstract data type that maps keys to values.

It uses a hash function to compute an offset into the array, from which the desired value can be found.

STATIC HASH TABLE

Allocate a giant array that has one slot for every element that you need to record.

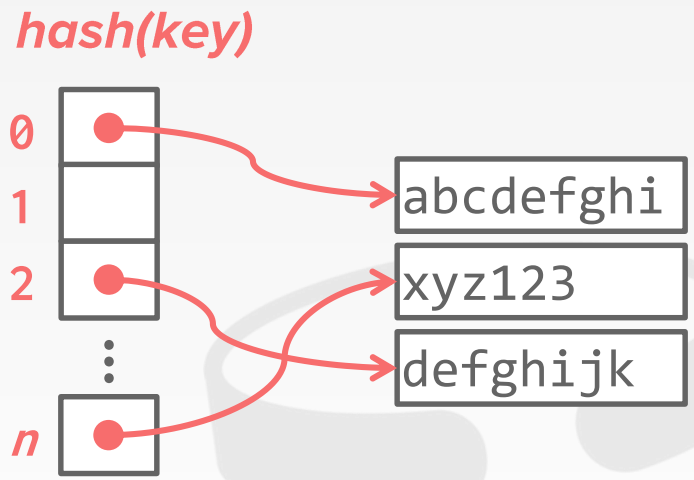
To find an entry, mod the key by the number of elements to find the offset in the array.



STATIC HASH TABLE

Allocate a giant array that has one slot for every element that you need to record.

To find an entry, mod the key by the number of elements to find the offset in the array.



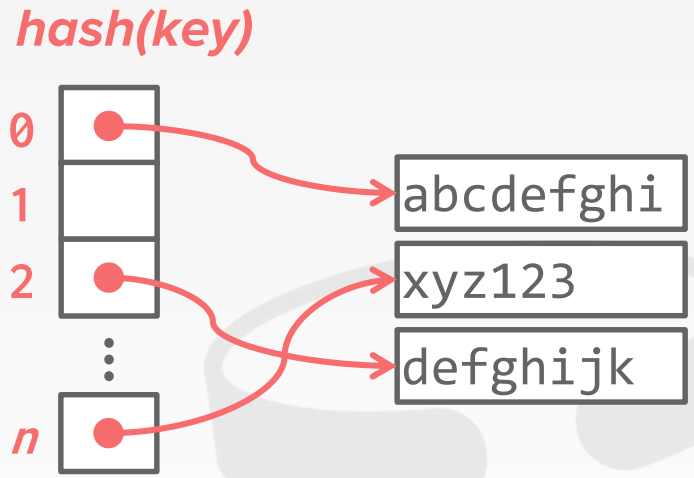
ASSUMPTION

You know the number of elements ahead of time.

Each key is unique.

Perfect hash function.

→ If **key1≠key2**, then **hash(key1)≠hash(key2)**



TODAY'S AGENDA

Chained Hashing
Open Hashing
Cuckoo Hashing
Extendible Hashing
Linear Hashing
Hash Functions

CHAINED HASHING

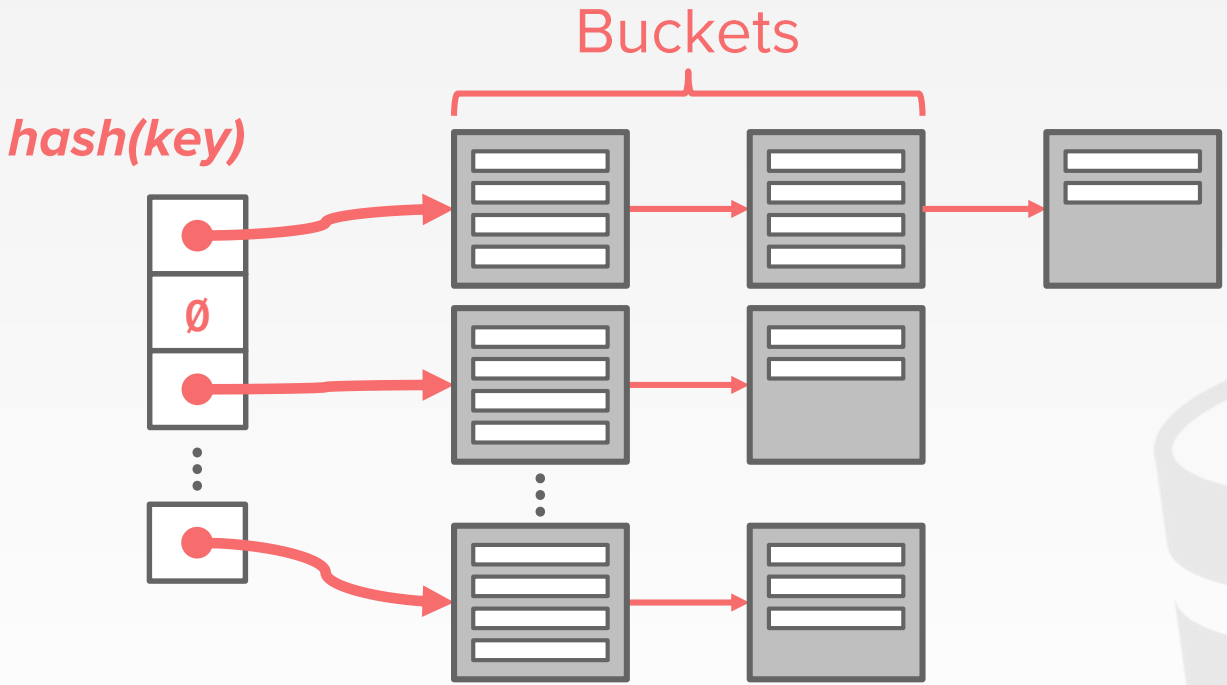
Maintain a linked list of buckets for each slot in the hash table.

Resolve collisions by placing all elements with the same hash key into the same bucket.

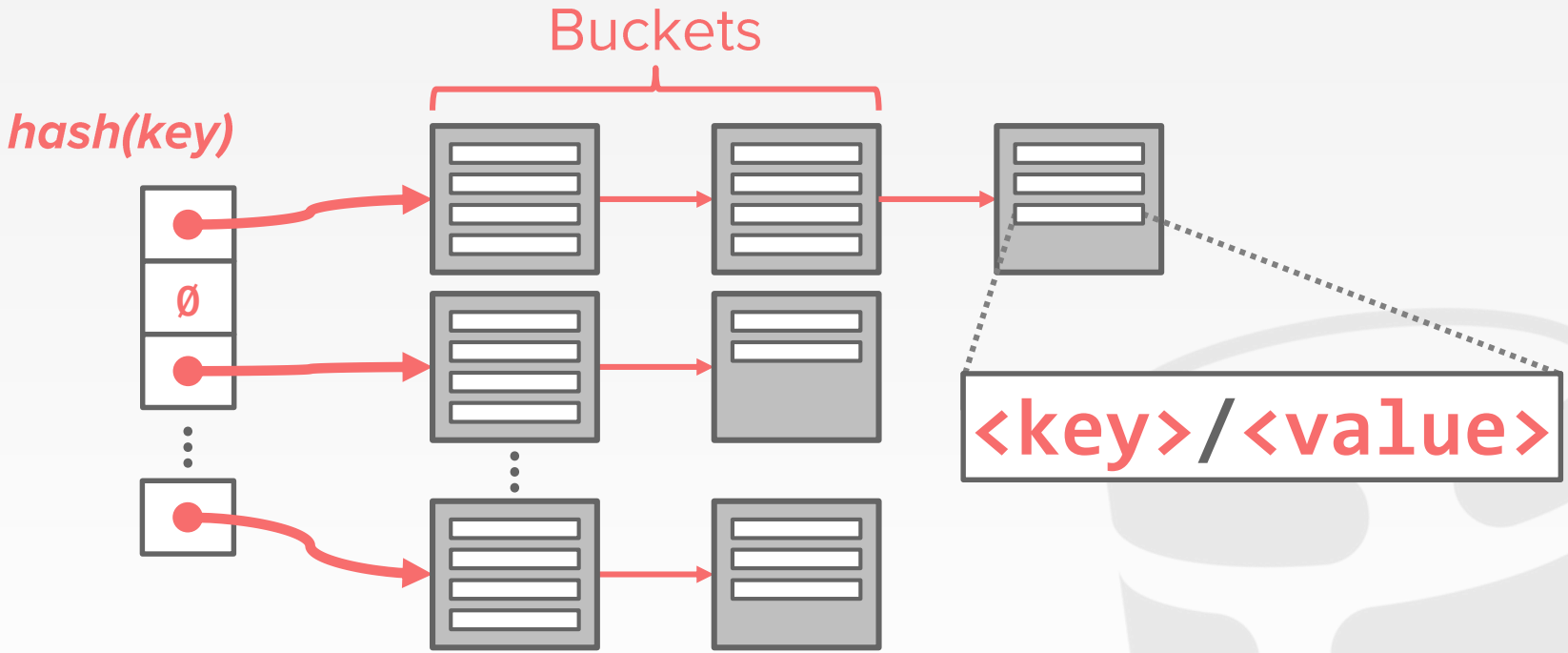
→ To determine whether an element is present, hash to its bucket and scan for it.

→ Insertions and deletions are generalizations of lookups.

CHAINED HASHING



CHAINED HASHING



CHAINED HASHING

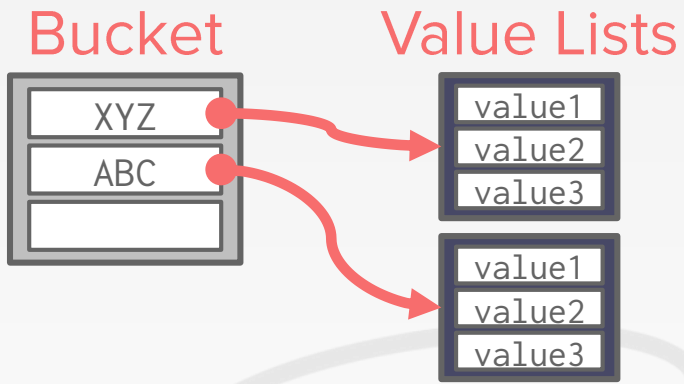
The hash table can grow infinitely because you just keep adding new buckets to the linked list.

You only need to take a latch on the bucket to store a new entry or extend the linked list.

NON-UNIQUE KEYS

Choice #1: Separate Linked List

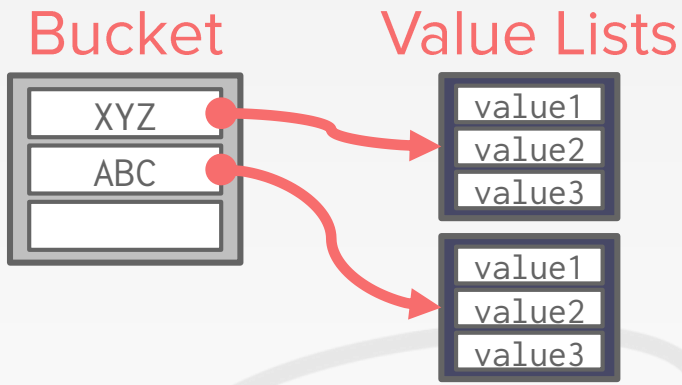
→ Store values in separate storage area for each key.



NON-UNIQUE KEYS

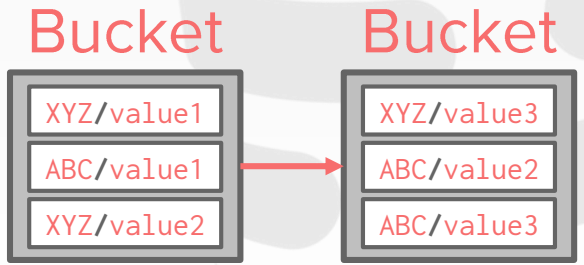
Choice #1: Separate Linked List

→ Store values in separate storage area for each key.



Choice #2: Store in Bucket

→ Store duplicate keys entries in the same buckets.



OPEN-ADDRESSING HASHING

Single giant table of slots.

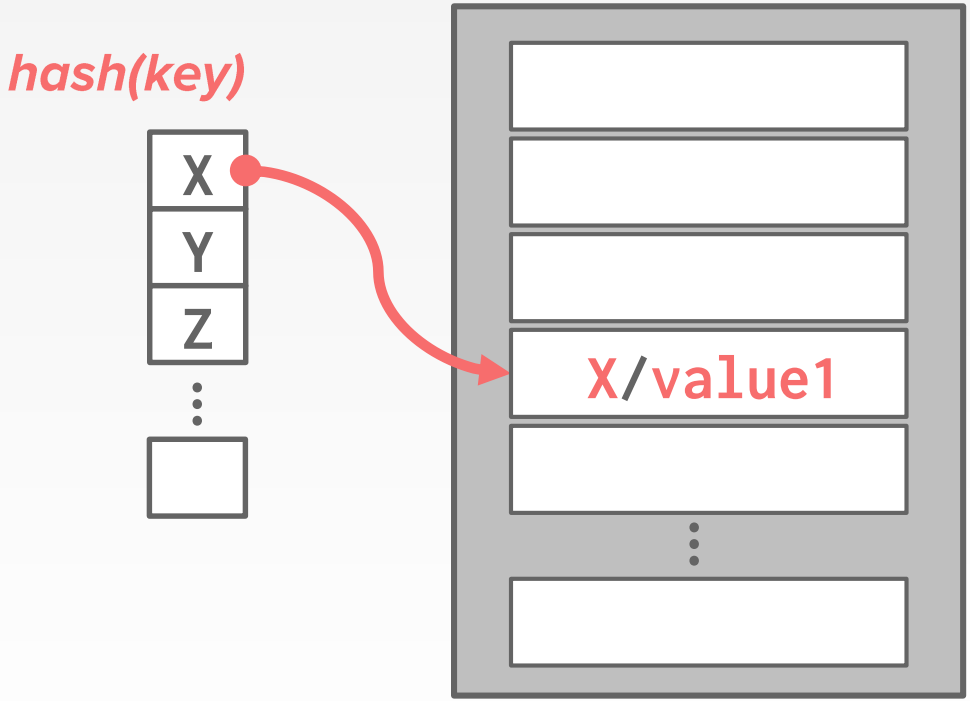
Resolve collisions by linearly searching for the next free slot in the table.

→ To determine whether an element is present, hash to a location in the index and scan for it.

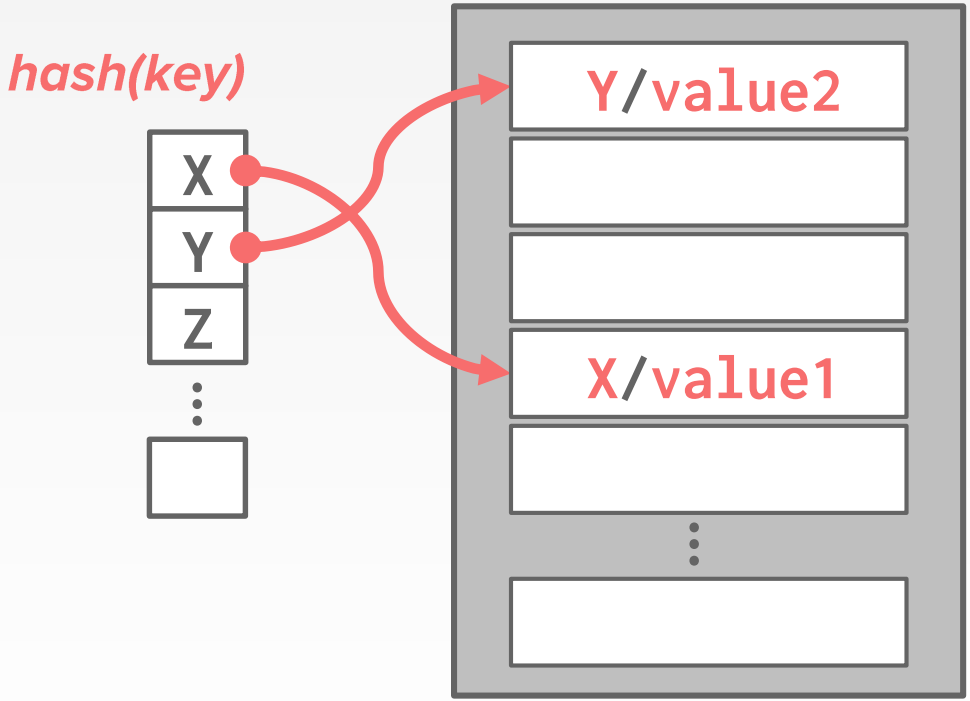
→ Have to store the key in the index to know when to stop scanning.

→ Insertions and deletions are generalizations of lookups.

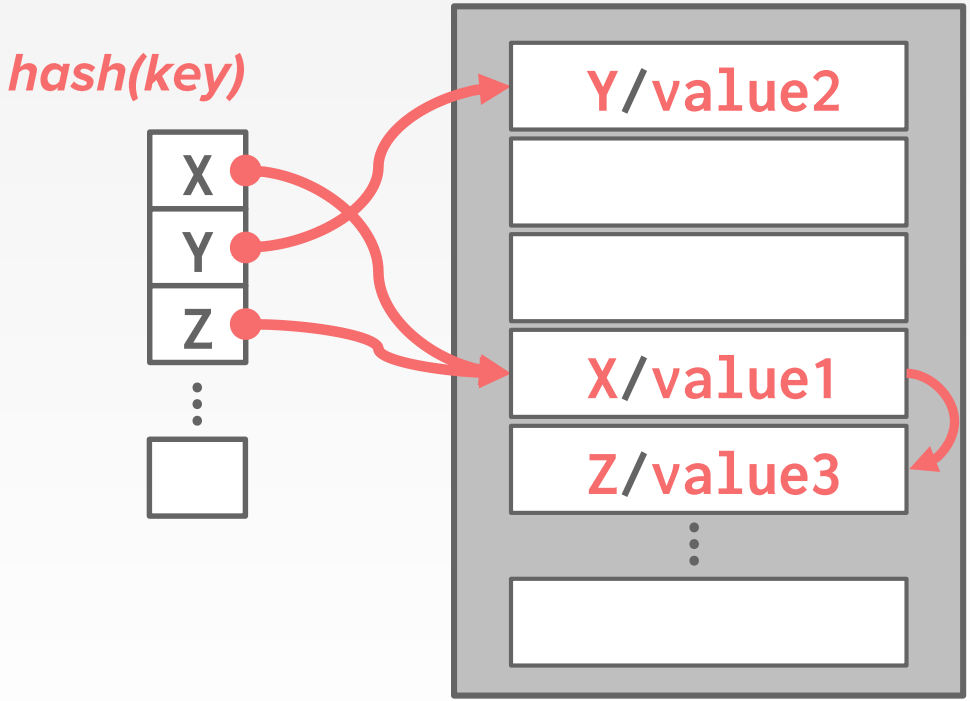
OPEN-ADDRESSING HASHING



OPEN-ADDRESSING HASHING



OPEN-ADDRESSING HASHING



OBSERVATION

To reduce the # of wasteful comparisons, it is important to avoid collisions of hashed keys.

This requires a hash table with ~2x the number of slots as the number of elements.

CUCKOO HASHING

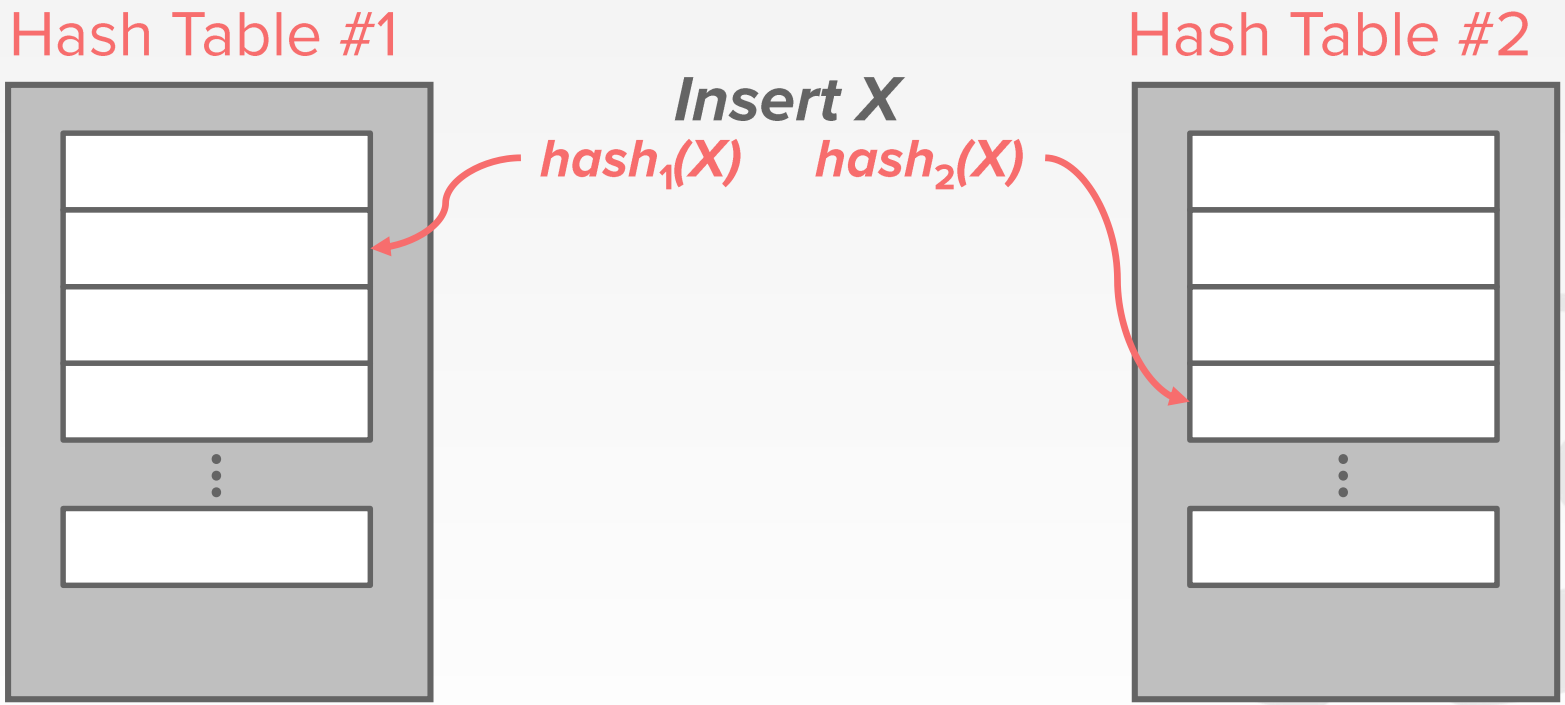
Use multiple hash tables with different hash functions.

→ On insert, check every table and pick anyone that has a free slot.

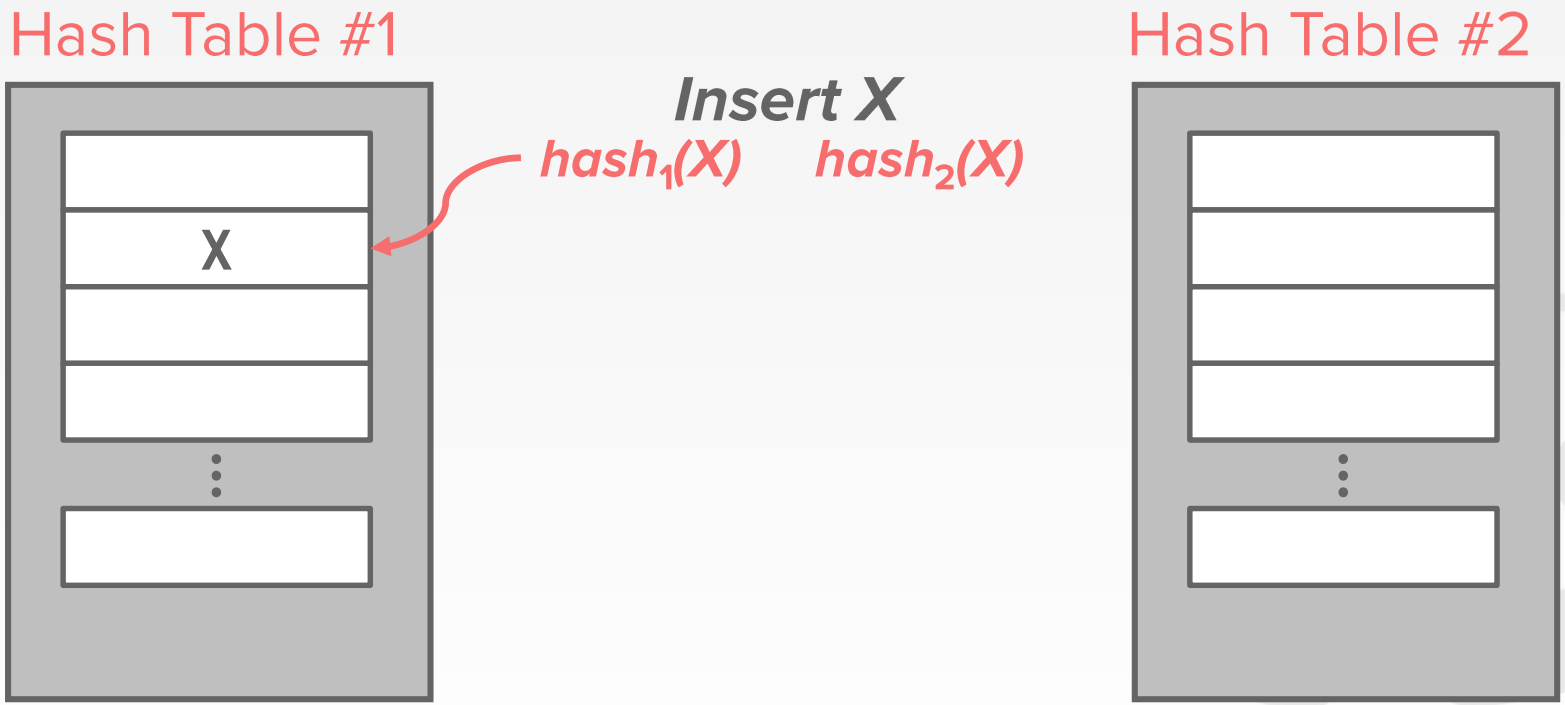
→ If no table has a free slot, evict the element from one of them and then re-hash it find a new location.

Look-ups and deletions are always O(1) because only one location per hash table is checked.

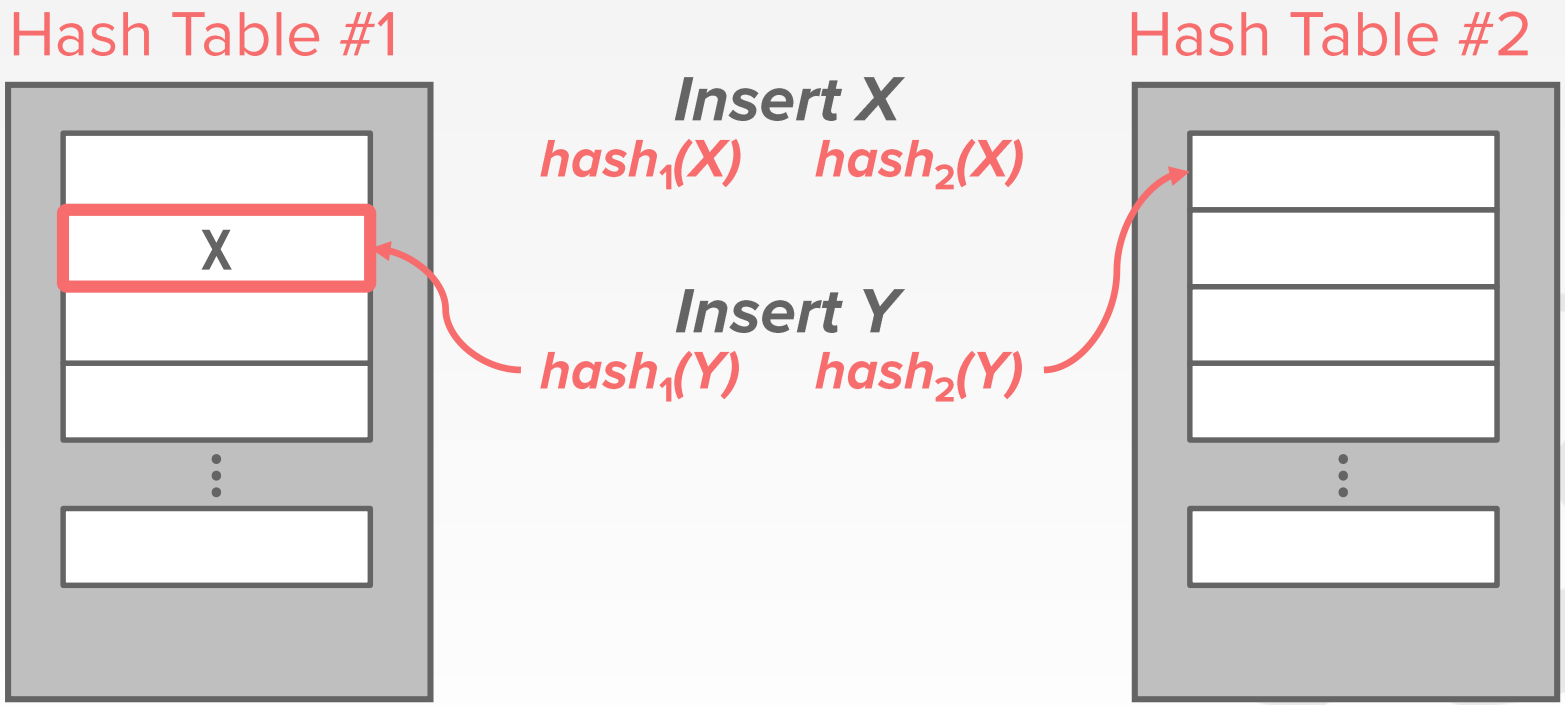
CUCKOO HASHING



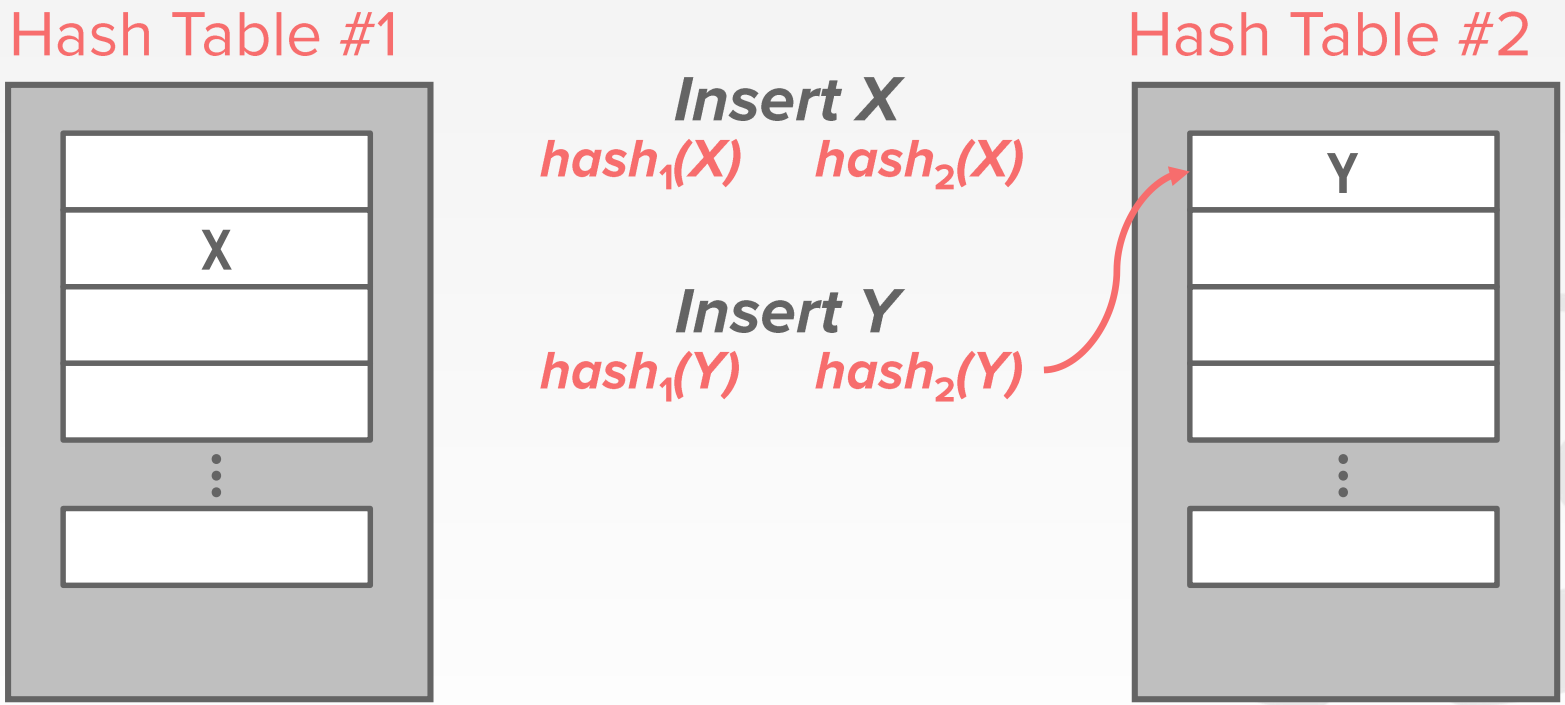
CUCKOO HASHING



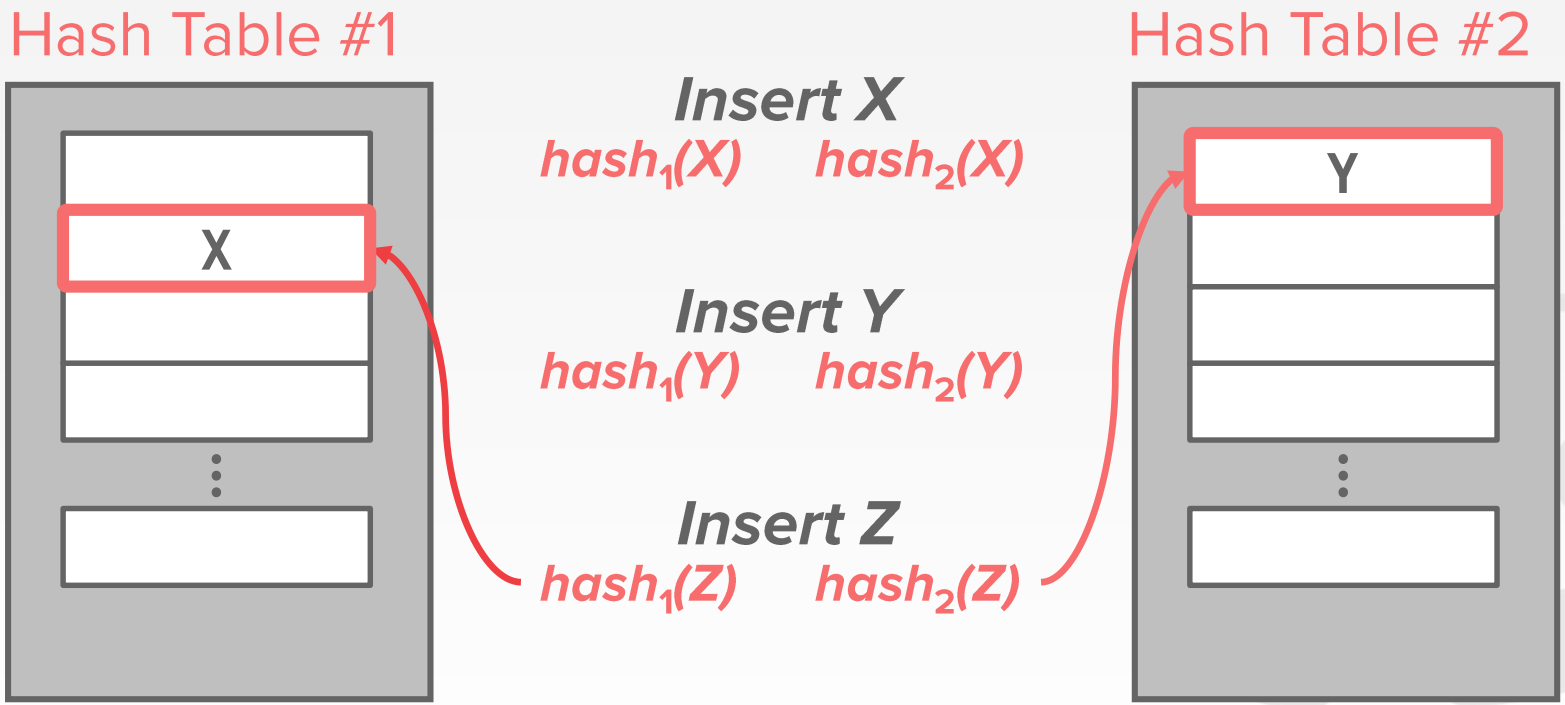
CUCKOO HASHING



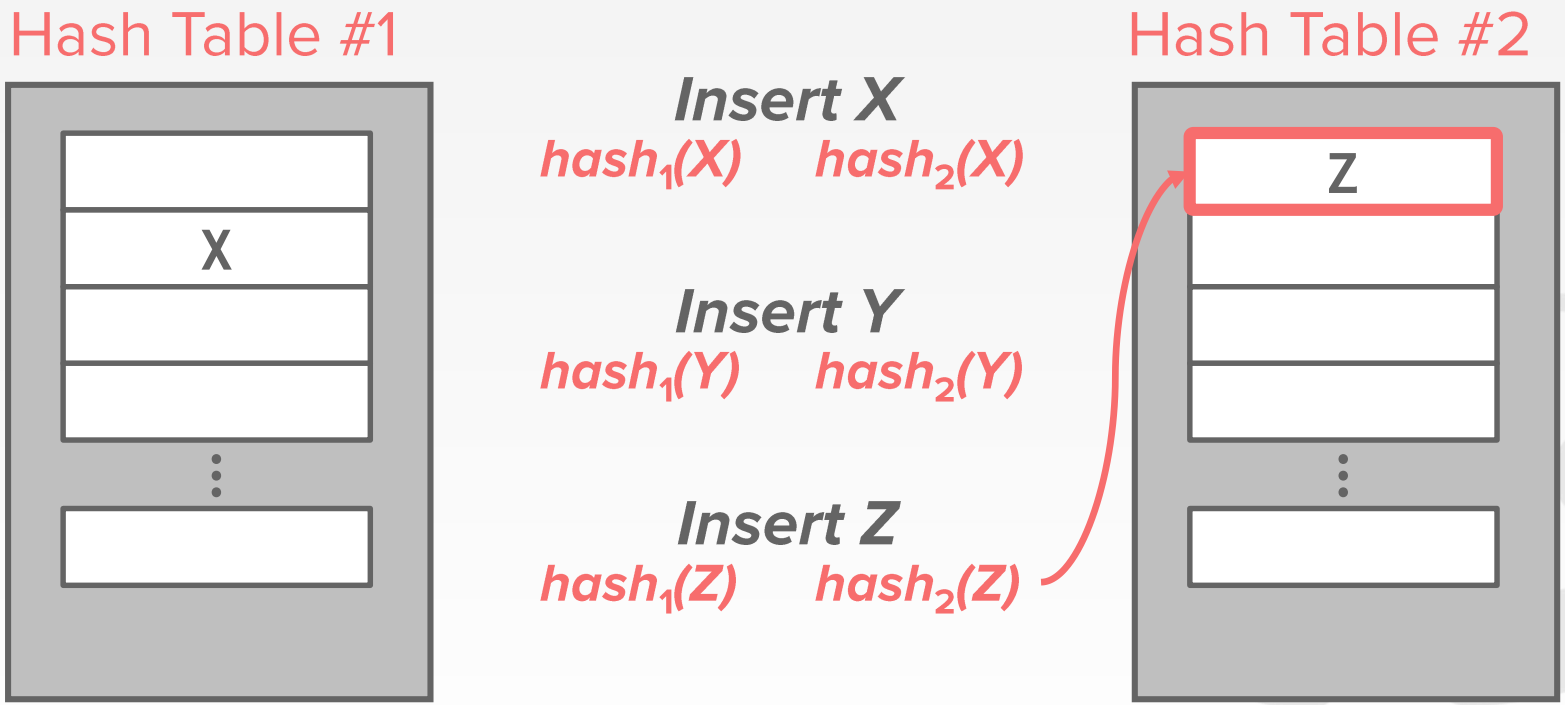
CUCKOO HASHING



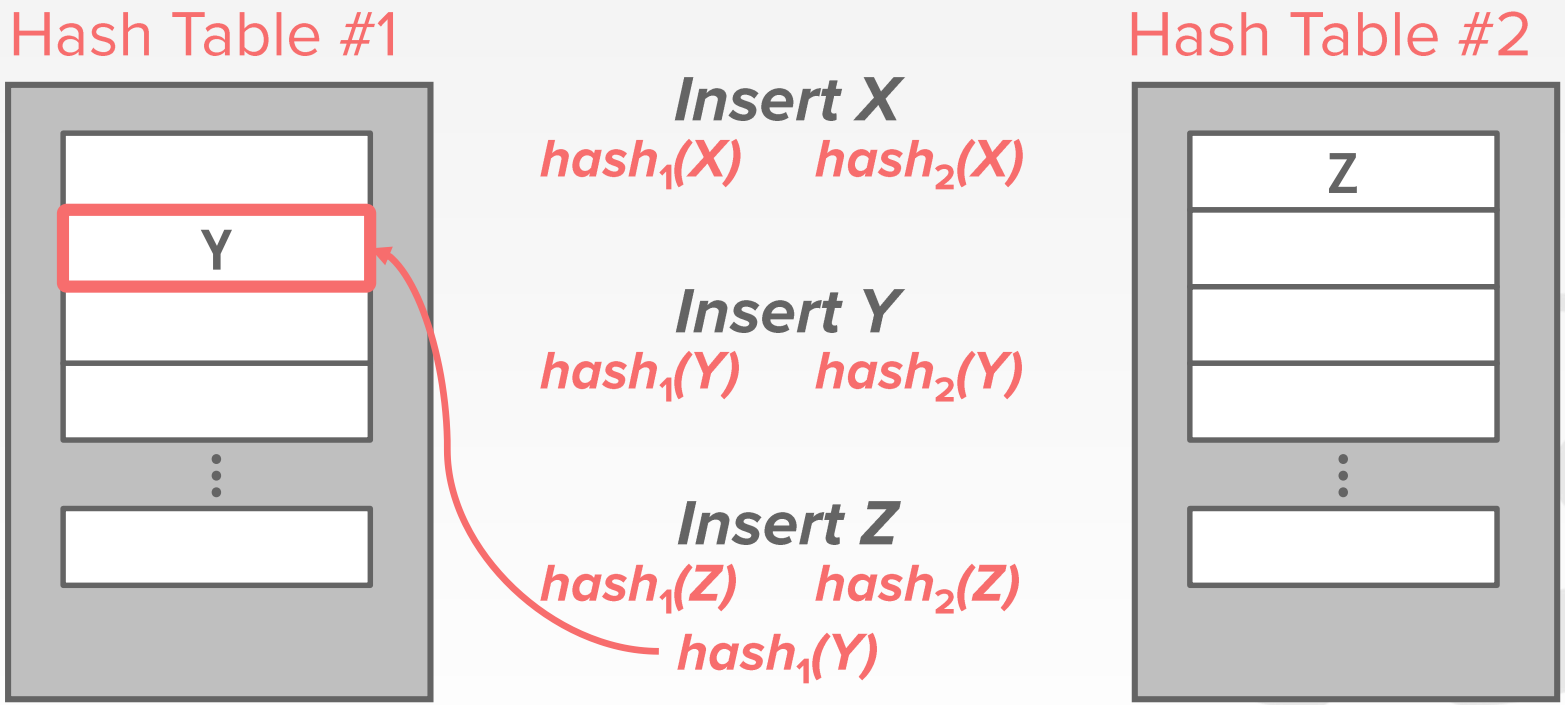
CUCKOO HASHING



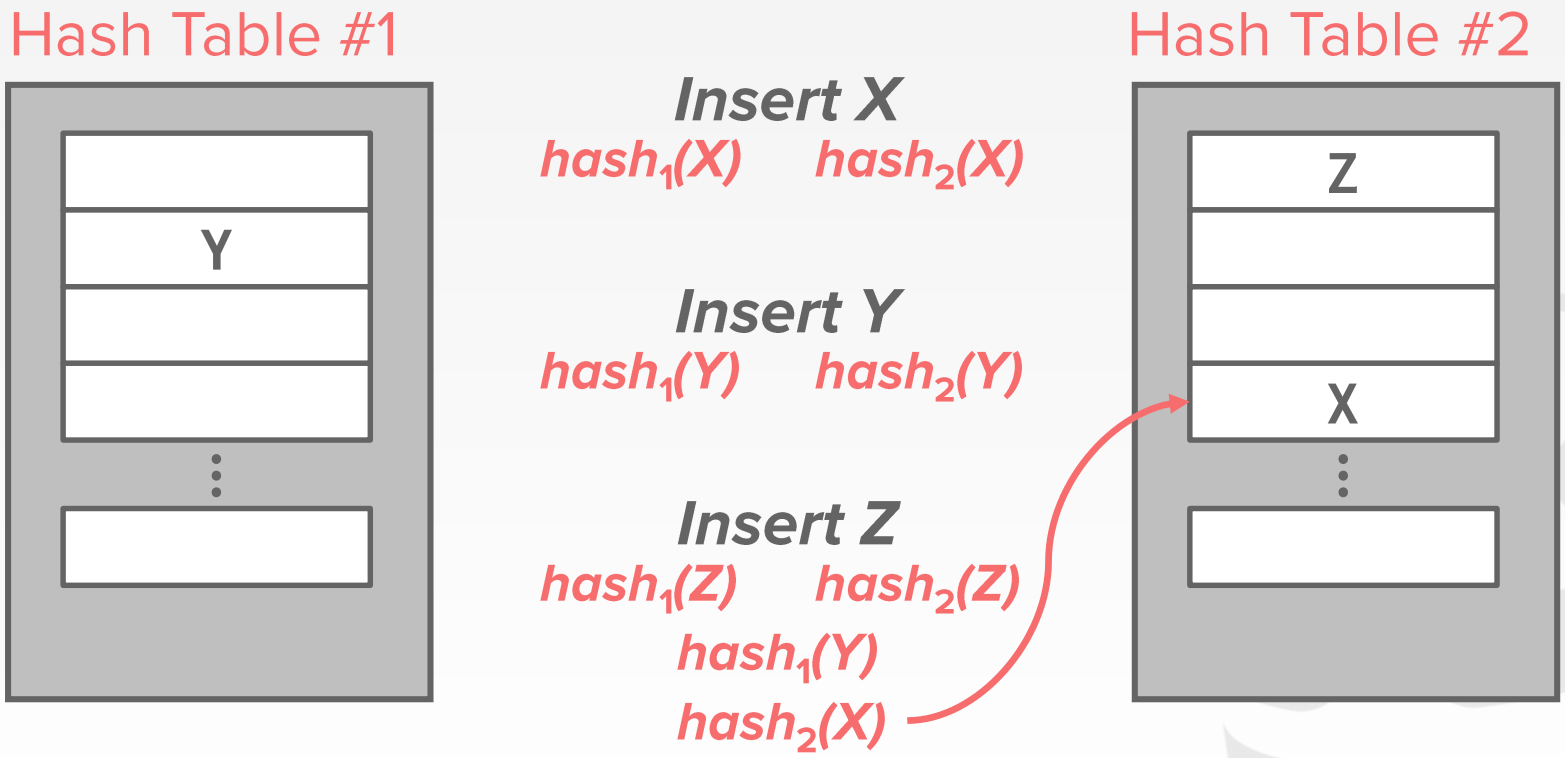
CUCKOO HASHING



CUCKOO HASHING



CUCKOO HASHING



CUCKOO HASHING

Make sure that we don't get stuck in an infinite loop when moving keys.

If we find a cycle, then we can rebuild the entire hash tables with new hash functions.

→ With **two** hash functions, we (probably) won't need to rebuild the table until it is at about 50% full.

→ With **three** hash functions, we (probably) won't need to rebuild the table until it is at about 90% full.

OBSERVATION

The previous hash tables require knowing the number of elements you want to store ahead of time.

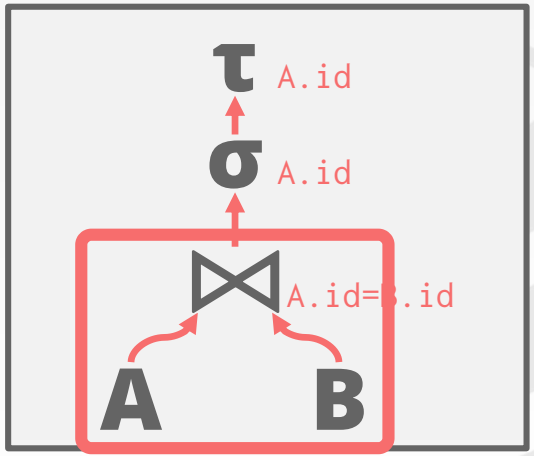
→ Otherwise you have rebuild the entire table if you need to grow/shrink.

OBSERVATION

The previous hash tables require knowing the number of elements you want to store ahead of time.

→ Otherwise you have rebuild the entire table if you need to grow/shrink.

```
SELECT A.id
  FROM A, B
 WHERE A.id = B.id
ORDER BY A.id
```



OBSERVATION

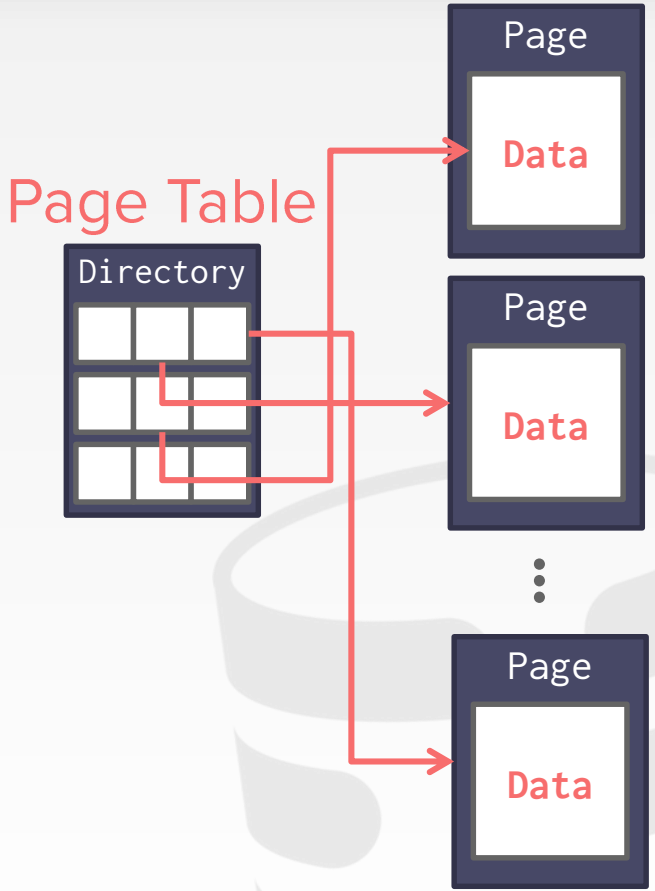
The previous hash tables require knowing the number of elements you want to store ahead of time.

→ Otherwise you have rebuild the entire table if you need to grow/shrink.

Dynamic hash tables are able to grow/shrink on demand.

→ Extendible Hashing

→ Linear Hashing

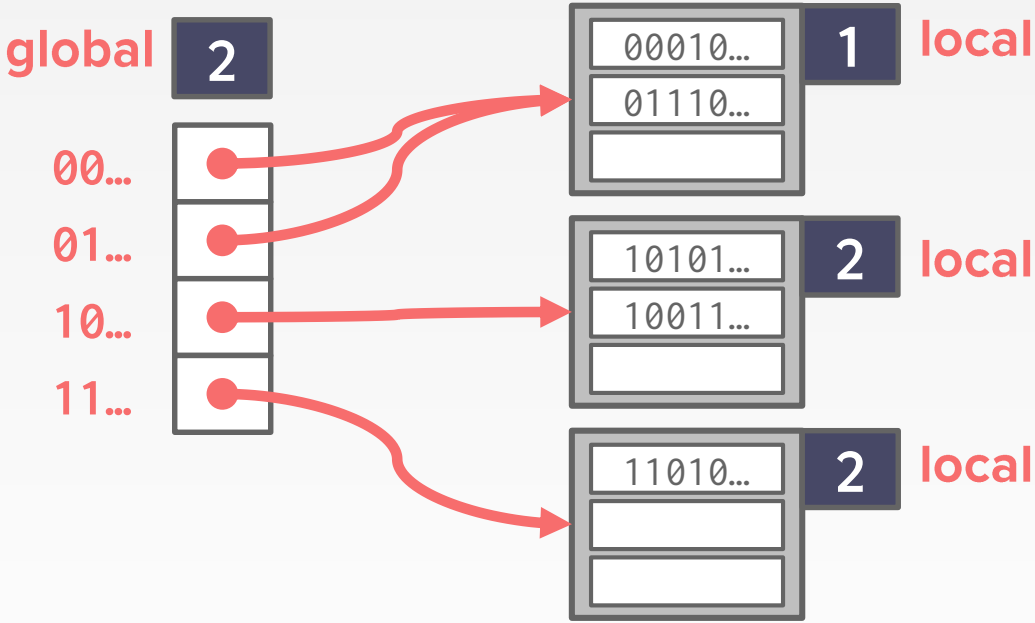


EXTENDIBLE HASHING

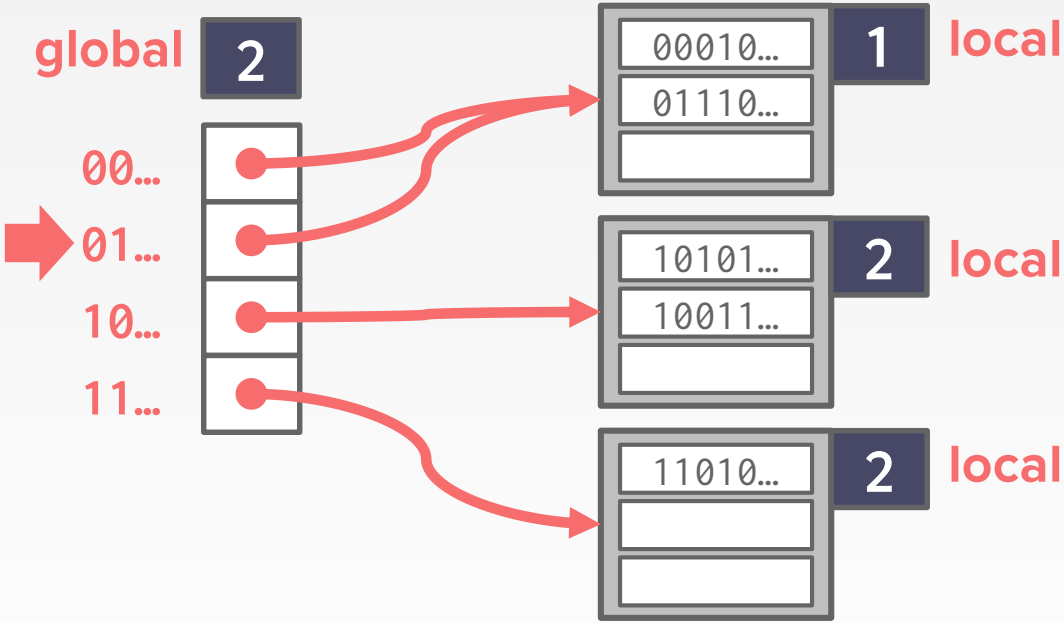
Chained-hashing approach where we split buckets instead of letting the linked list grow forever.

This requires reshuffling entries on split, but the change is localized.

EXTENDIBLE HASHING



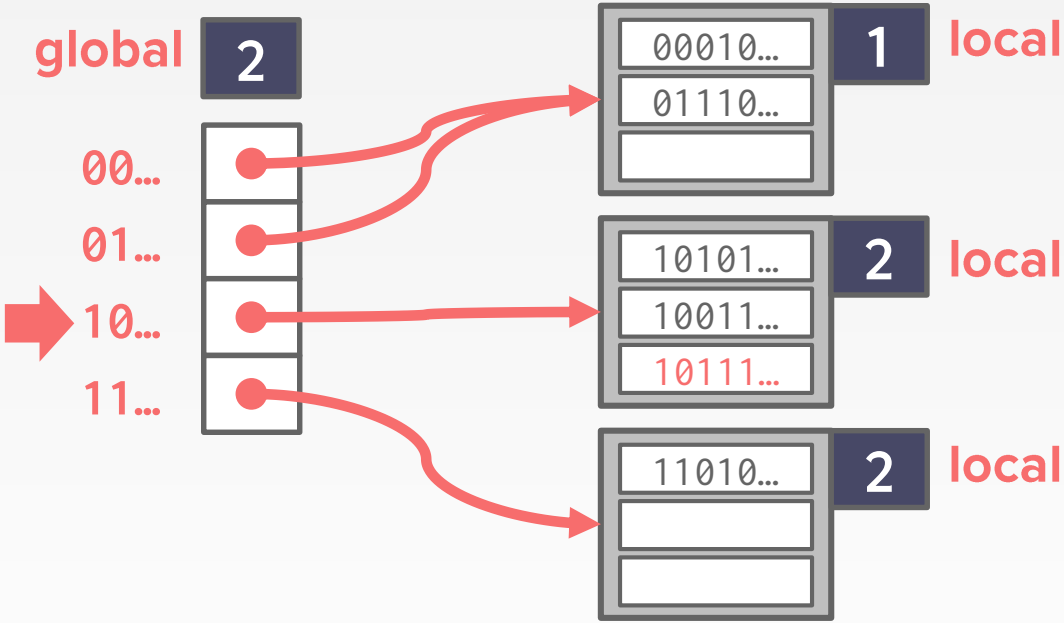
EXTENDIBLE HASHING



Find X

hash(X) = 01110...

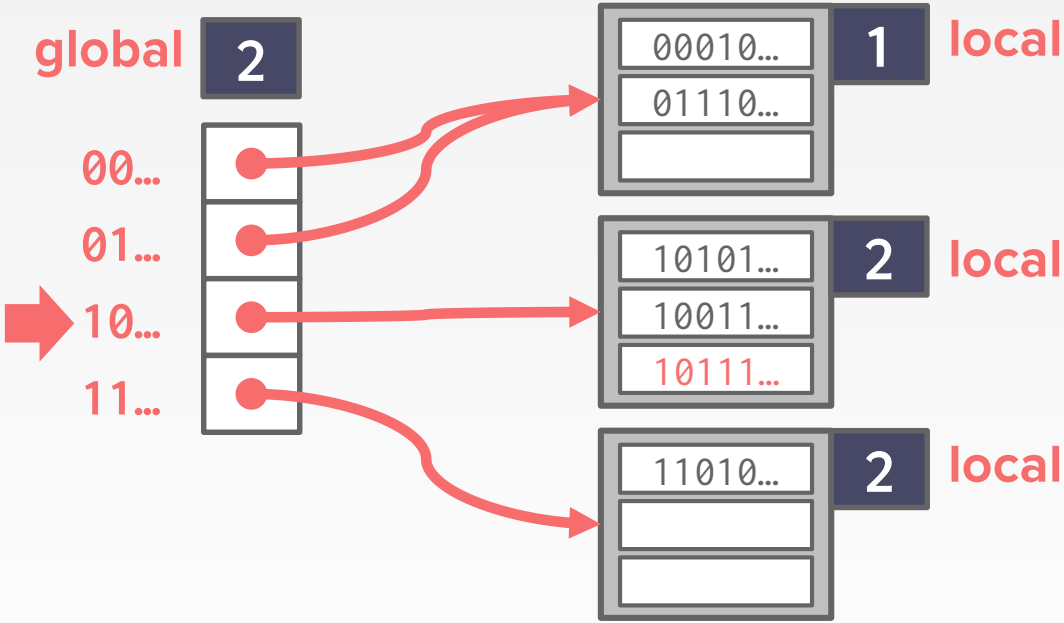
EXTENDIBLE HASHING



Find X
hash(X) = 01110...

Insert Y
hash(Y) = 10111...

EXTENDIBLE HASHING

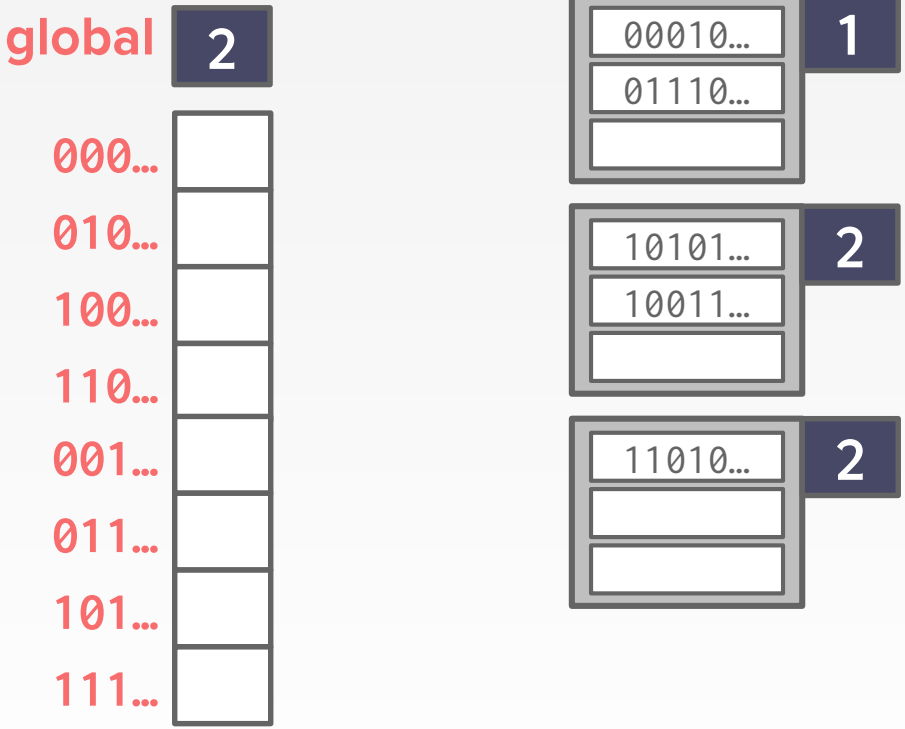


Find X
hash(X) = 01110...

Insert Y
hash(Y) = 10111...

Insert Z
hash(Z) = 10100...

EXTENDIBLE HASHING

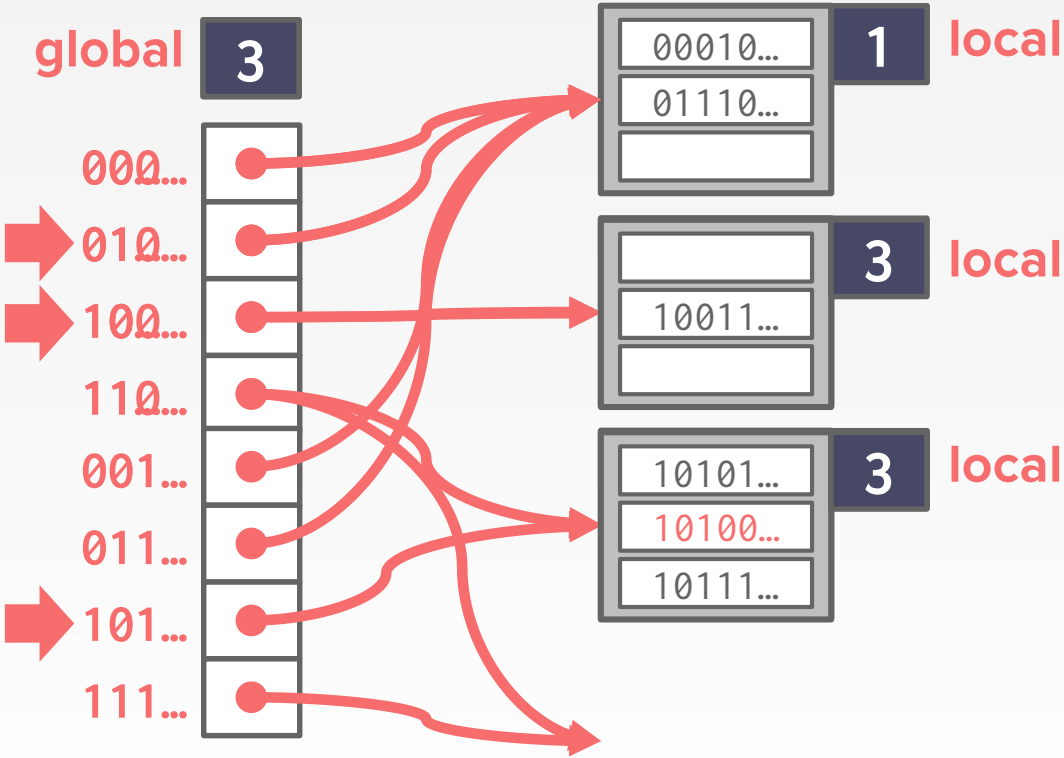


Find X
hash(X) = 01110...

Insert Y
hash(Y) = 10111...

Insert Z
hash(Z) = 10100...

EXTENDIBLE HASHING



Find X
hash(X) = 01110...

Insert Y
hash(Y) = 10111...

Insert Z
hash(Z) = 10100...

LINEAR HASHING

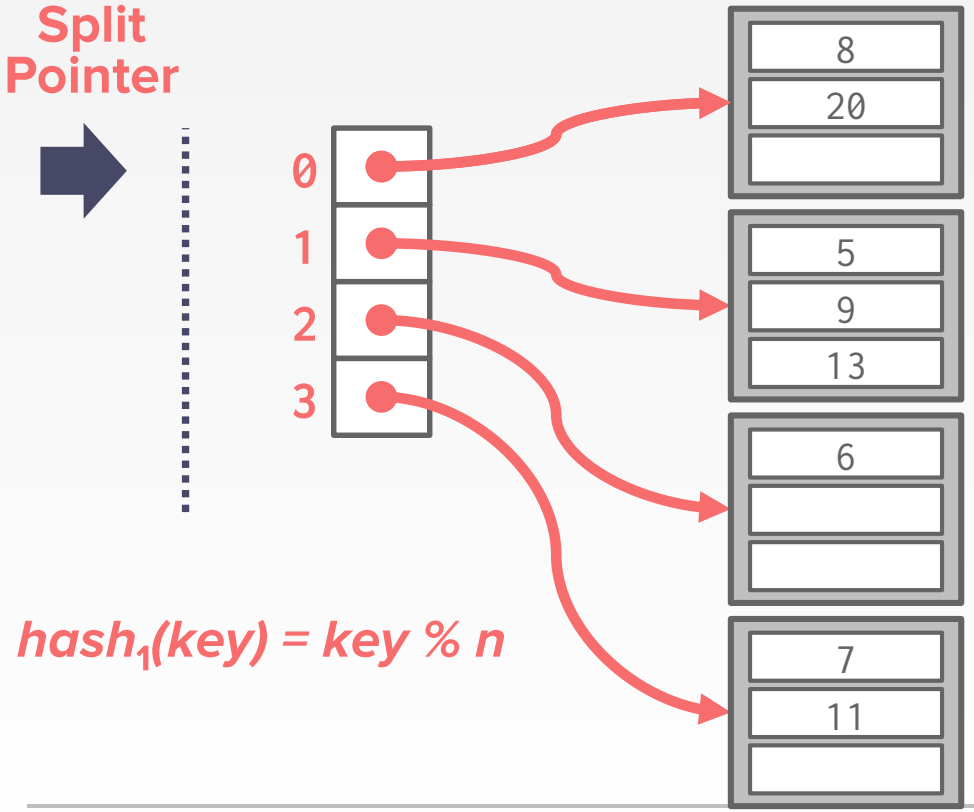
Maintain a pointer that tracks the next bucket to split.

When any bucket overflows, split the bucket at the pointer location.

Overflow criterion is left up to the implementation.

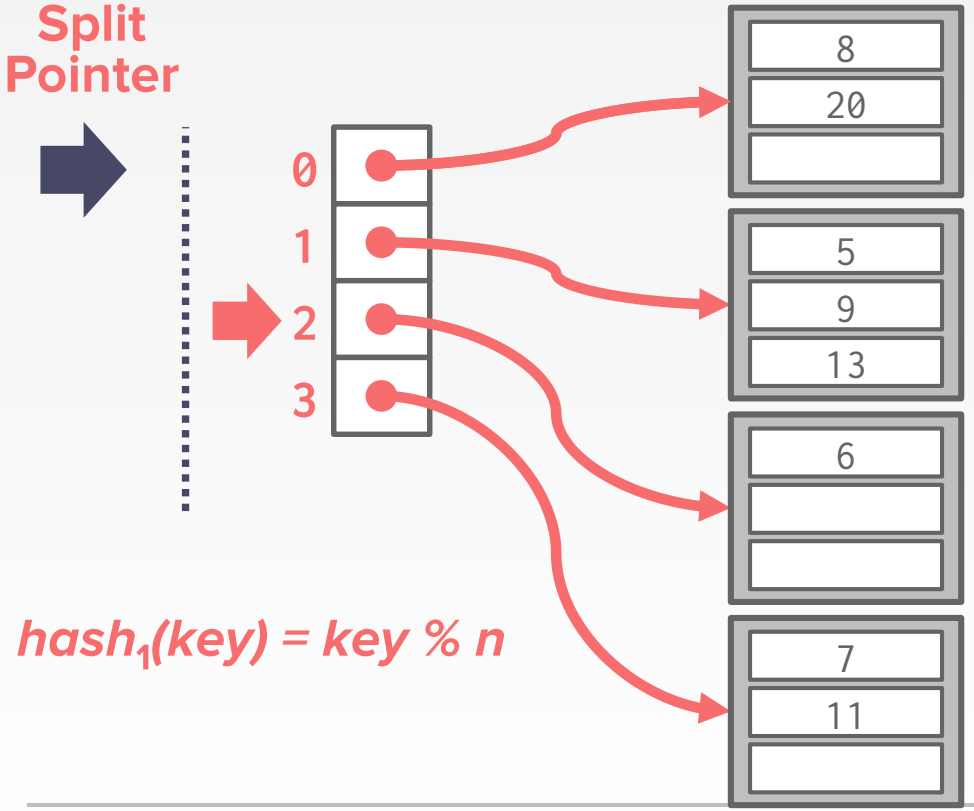
→ Space Utilization
→ Average Length of Overflow Chains

LINEAR HASHING



$hash_1(key) = key \% n$

LINEAR HASHING

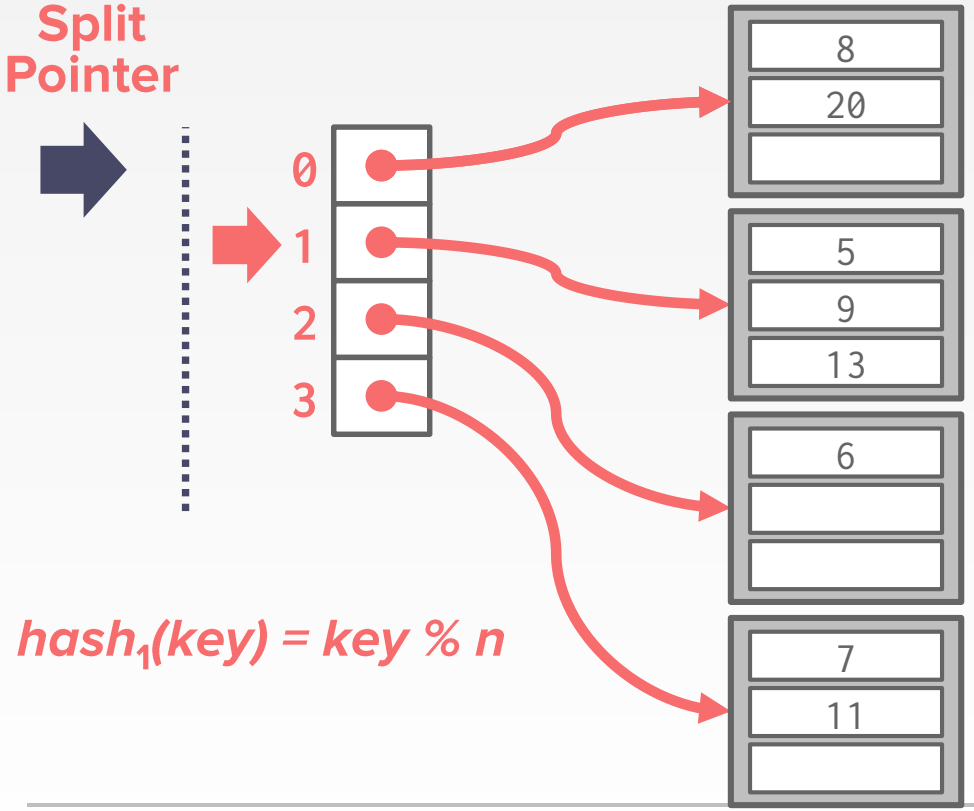


$hash_1(key) = key \% n$

Find 5

$hash_1(6) = 6\%4 = 2$

LINEAR HASHING



$hash_1(key) = key \% n$

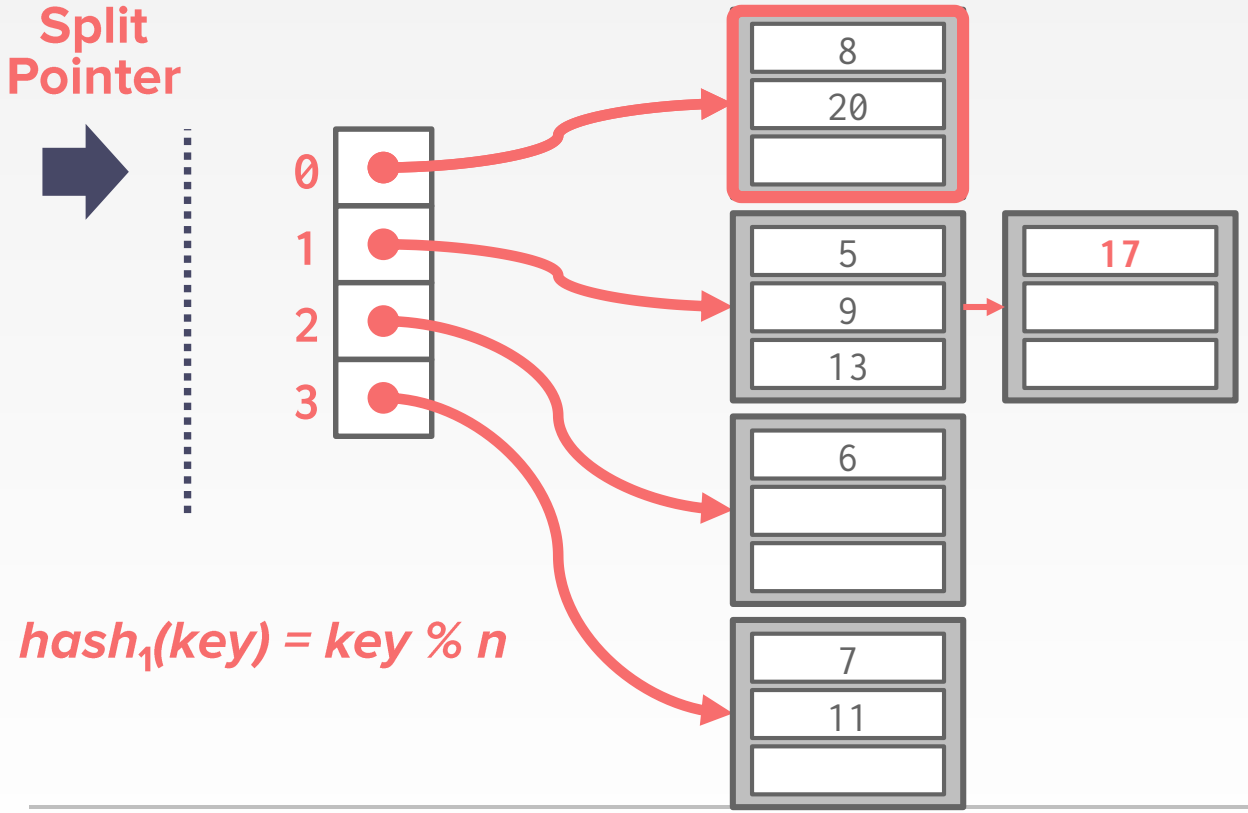
Find 5

$hash_1(6) = 6\%4 = 2$

Insert 17

$hash_1(17) = 17\%4 = 1$

LINEAR HASHING



Find 5

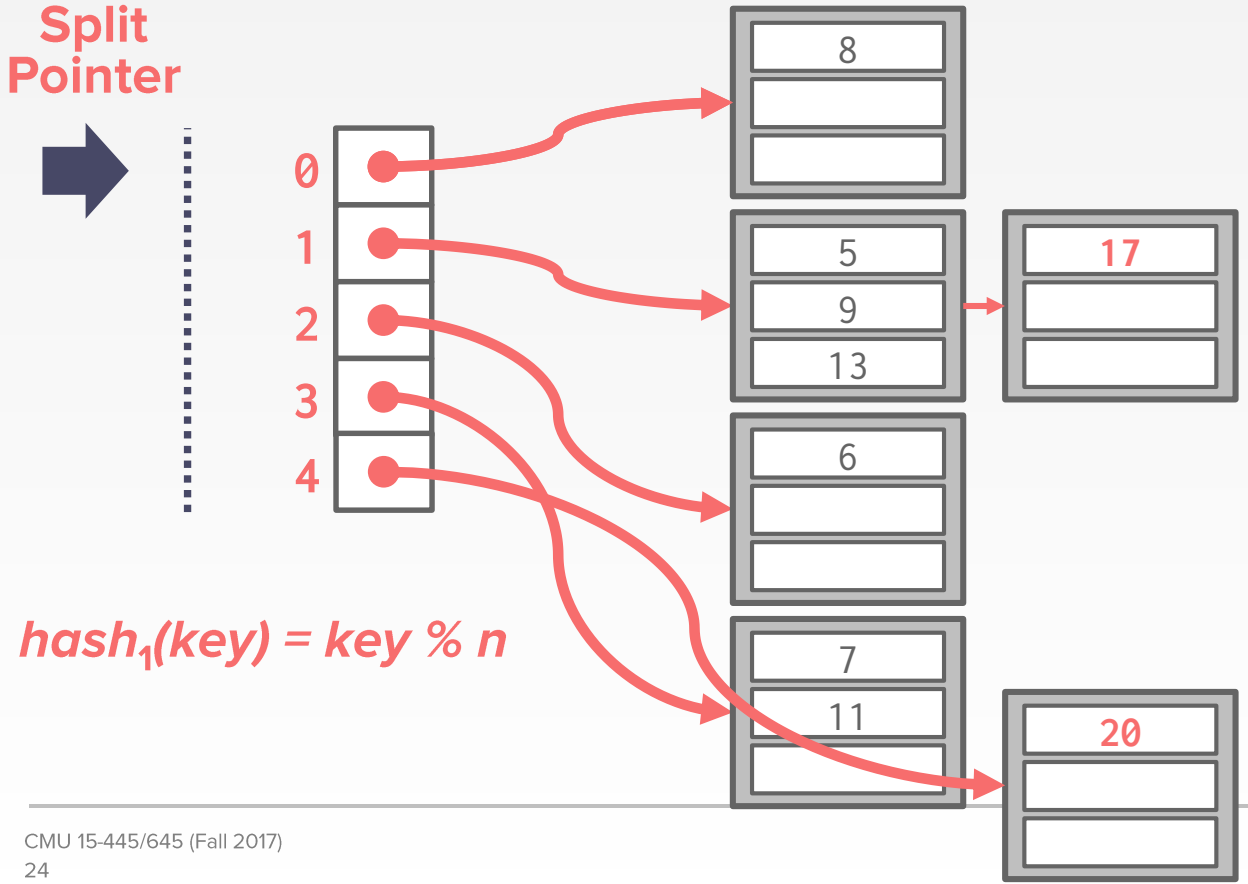
$hash_1(6) = 6\%4 = 2$

Insert 17

$hash_1(17) = 17\%4 = 1$

$hash_1(key) = key \% n$

LINEAR HASHING



Find 5

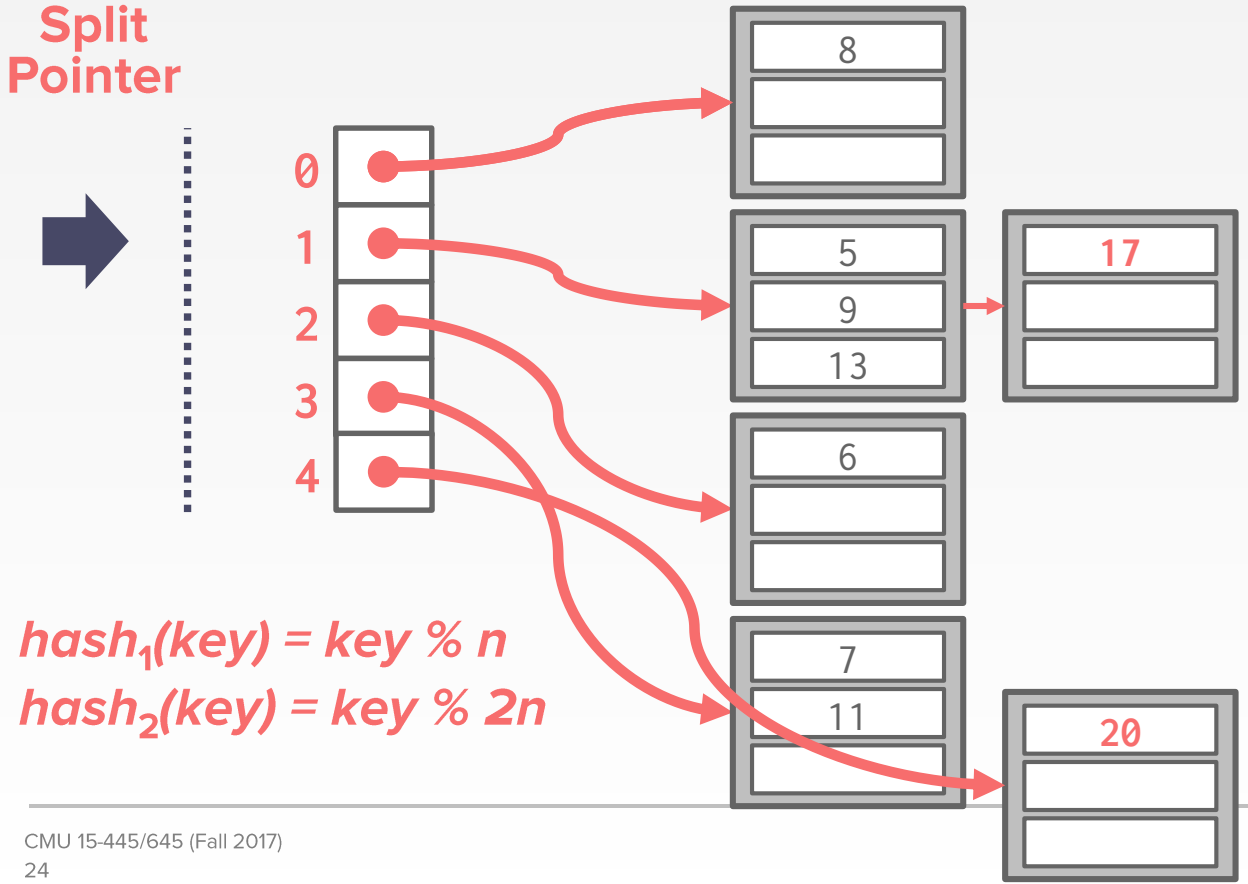
$hash_1(6) = 6\%4 = 2$

Insert 17

$hash_1(17) = 17\%4 = 1$

$hash_1(key) = key \% n$

LINEAR HASHING



Find 5

$hash_1(6) = 6\%4 = 2$

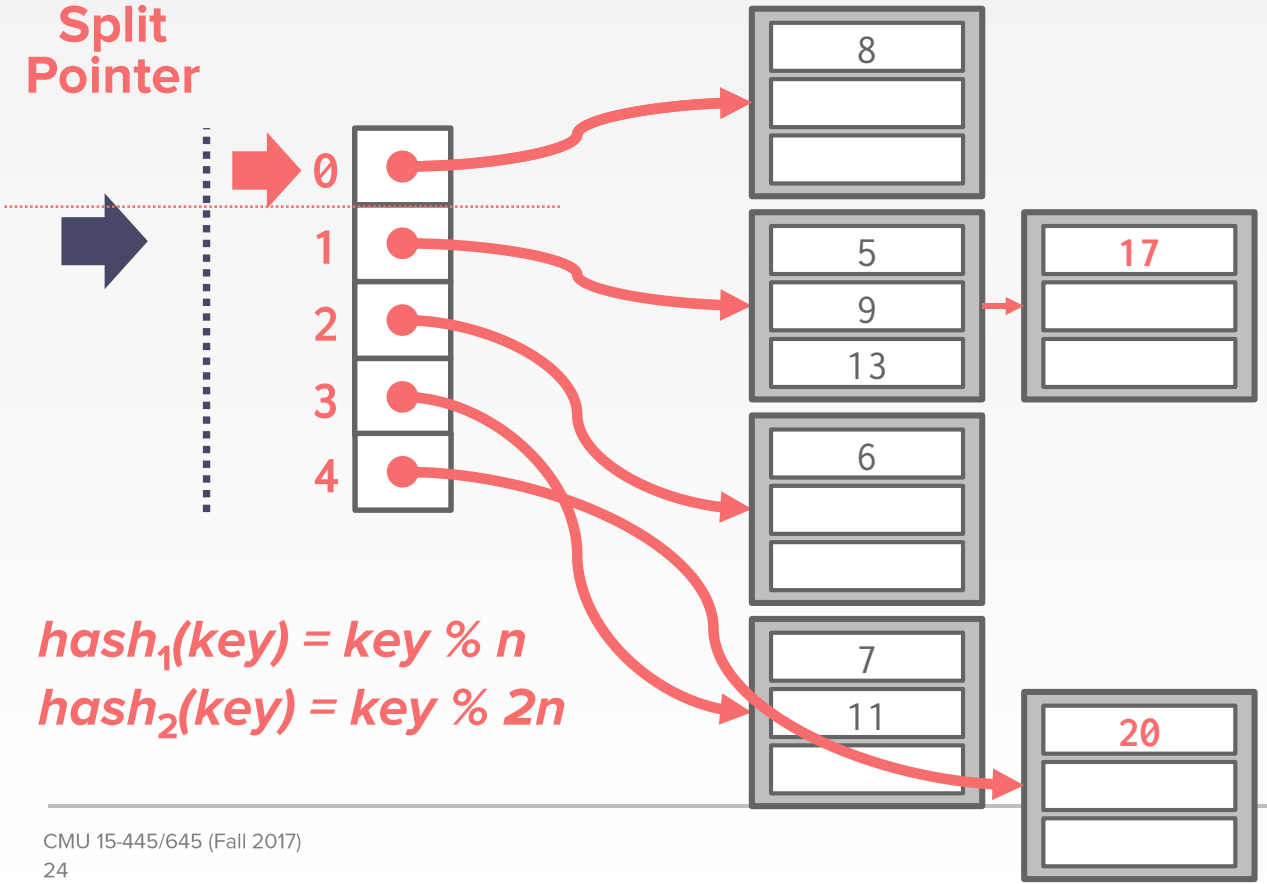
Insert 17

$hash_1(17) = 17\%4 = 1$

$hash_1(key) = key \% n$

$hash_2(key) = key \% 2n$

LINEAR HASHING



$hash_1(key) = key \% n$

$hash_2(key) = key \% 2n$

Find 5

$hash_1(6) = 6\%4 = 2$

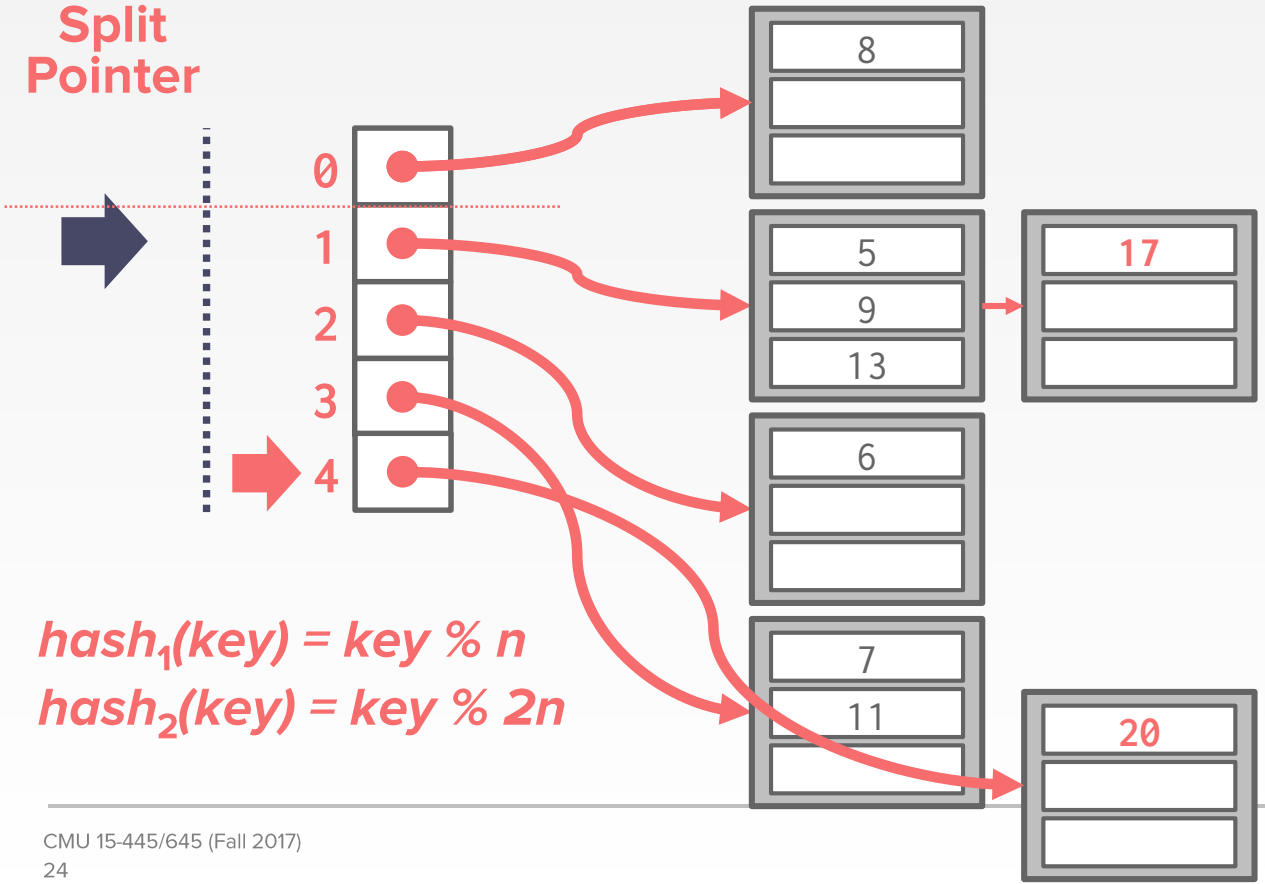
Insert 17

$hash_1(17) = 17\%4 = 1$

Find 20

$hash_1(20) = 20\%4 = 0$

LINEAR HASHING



$hash_1(key) = key \% n$

$hash_2(key) = key \% 2n$

Find 5

$hash_1(6) = 6\%4 = 2$

Insert 17

$hash_1(17) = 17\%4 = 1$

Find 20

$hash_1(20) = 20\%4 = 0$

$hash_2(20) = 20\%8 = 4$

LINEAR HASHING

Splitting buckets based on the split pointer will eventually get to all overflowed buckets.

→ When the pointer reaches the last slot, delete the first hash function and move back to beginning.

The pointer can also move backwards when buckets are empty.

HASH FUNCTIONS

We don’t want to use a cryptographic hash function for our join algorithm.

We want something that is fast and will have a low collision rate.

HASH FUNCTIONS

MurmurHash (2008)
→ Designed to a fast, general purpose hash function.

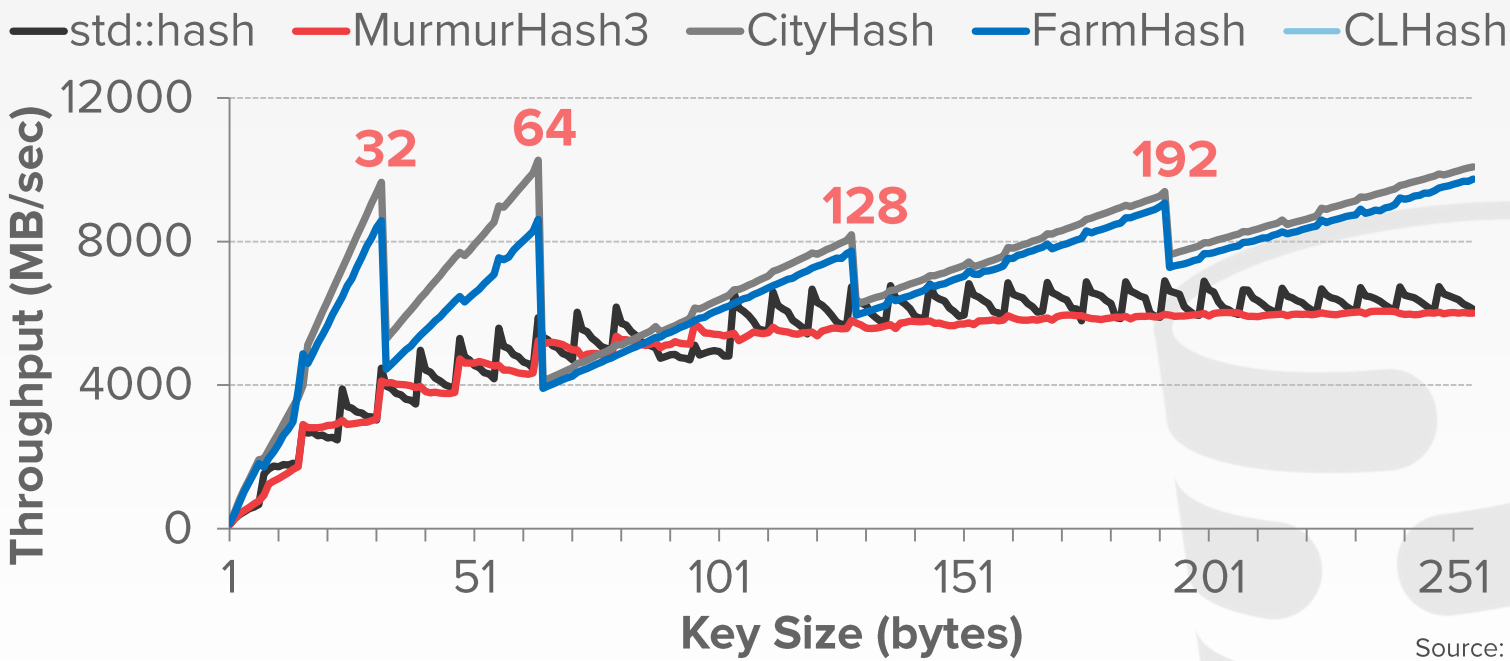
Google CityHash (2011)
→ Based on ideas from MurmurHash2
→ Designed to be faster for short keys (<64 bytes).

Google FarmHash (2014)
→ Newer version of CityHash with better collision rates.

CLHash (2016)
→ Fast hashing function based on carry-less multiplication.

HASH FUNCTION BENCHMARKS

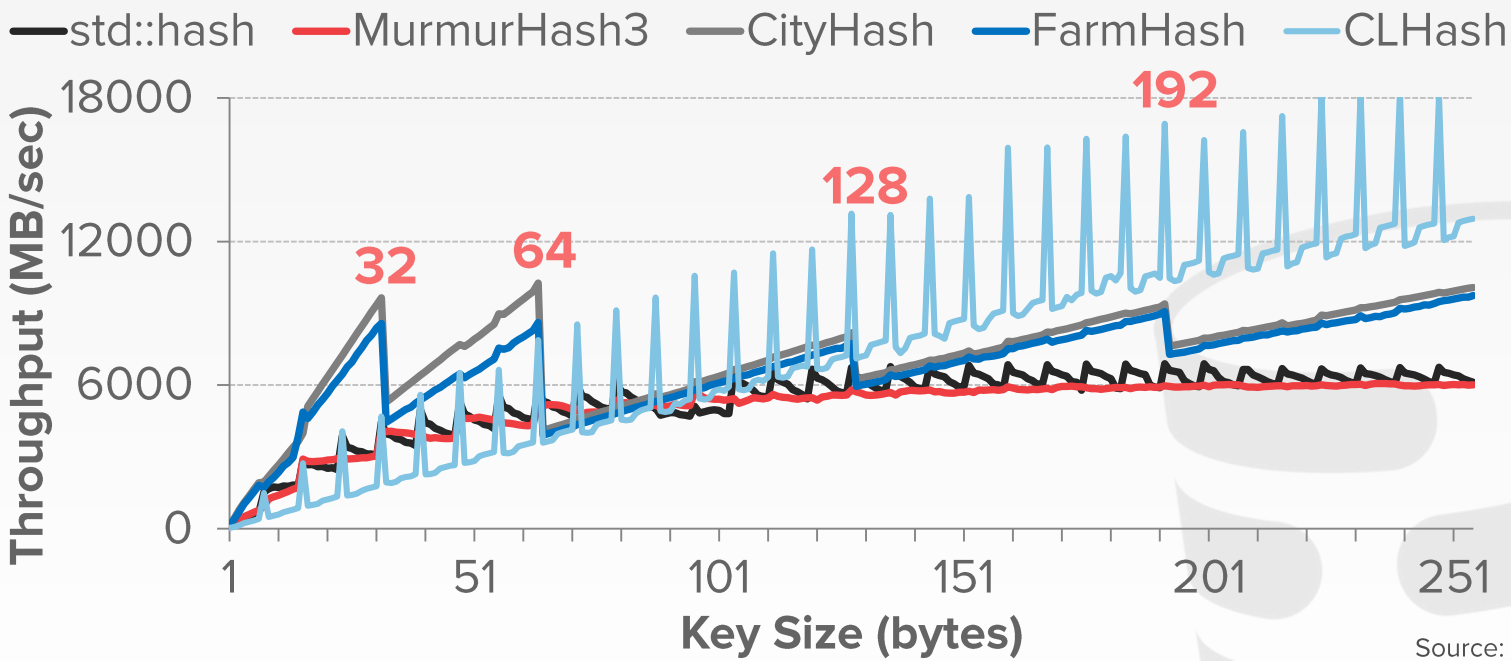
Intel Xeon CPU E5-2630v4 @ 2.20GHz



Source: Fredrik Widlund

HASH FUNCTION BENCHMARKS

Intel Xeon CPU E5-2630v4 @ 2.20GHz



Source: Fredrik Widlund

HASH TABLES

Fast data structures that support O(1) look-ups.

This they are usually **not** what you want to use for a table index…

Postgres Demo

CONCLUSION

Hash tables are important in databases.

Trade-off between speed and flexibility.

NEXT CLASS

B+Trees
Skip Lists
Radix Trees