Intro to Databases (COMP_SCI 339)

05 Hash Tables

Northwestern University

WINTER 2024

Andrew Crotty

ADMINISTRIVIA

Project #2 is due Sunday 2/4 @ 11:59pm

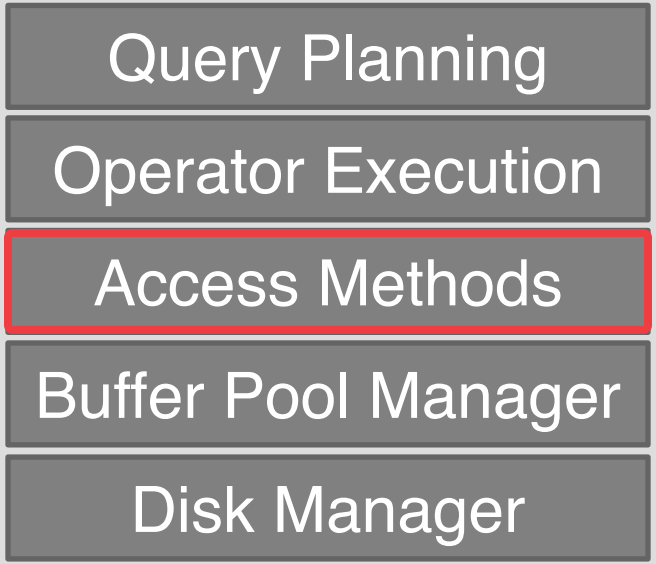
Exam #1 will be on 1/29 from 3:30-4:50pm

Homework #1 will be released tonight

COURSE STATUS

We will now talk about how to support the DBMS's execution engine to efficiently 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 lay out data structure in memory/pages and what information to store 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 unordered associative array that maps keys to values.

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

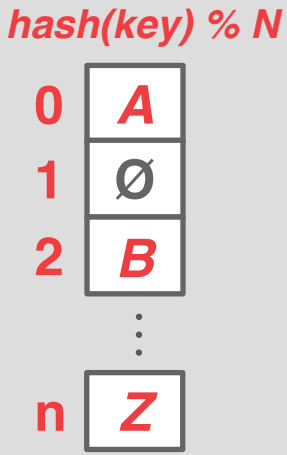
Space Complexity: **O(n)**
Time Complexity:
→ Average: **O(1)**
→ Worst: **O(n)**

DBMS needs to care about constants!

STATIC HASH TABLE

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

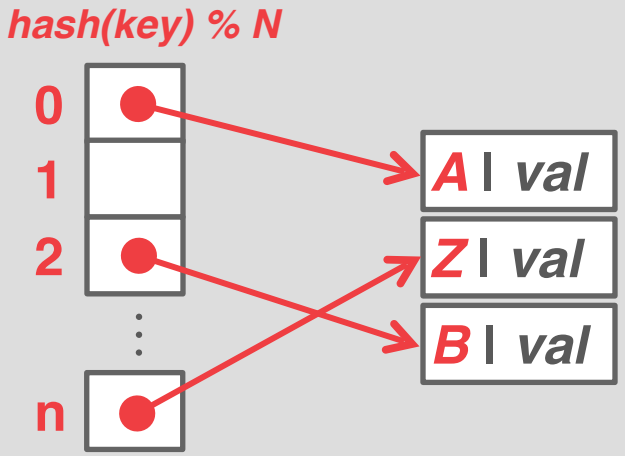
To find an entry, mod the key by the number of elements to find the offset in the array.



STATIC HASH TABLE

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

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



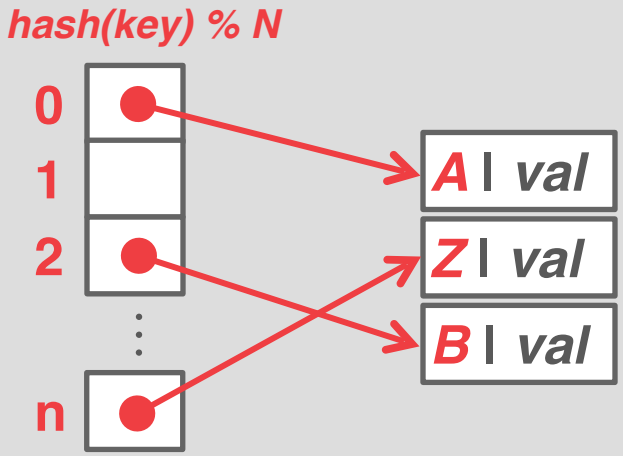
ASSUMPTIONS

Assumption#1: Number of elements is known ahead of time and fixed.

Assumption #2: Each key is unique.

Assumption #3: Perfect hash function.

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



HASH TABLE

Design Decision #1: Hash Function

→ How to map a large key space into a smaller domain.
→ Trade-off between being fast vs. collision rate.

Design Decision #2: Hashing Scheme

→ How to handle key collisions after hashing.
→ Trade-off between allocating a large hash table vs. additional instructions to get/put keys.

TODAY'S AGENDA

Hash Functions
Static Hashing Schemes
Dynamic Hashing Schemes

HASH FUNCTIONS

For any input key, return an integer representation of that key.

We do not want to use a cryptographic hash function for DBMS hash tables (e.g., SHA-2).

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

HASH FUNCTIONS

CRC-64 (1975)
→ Used in networking for error detection.

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

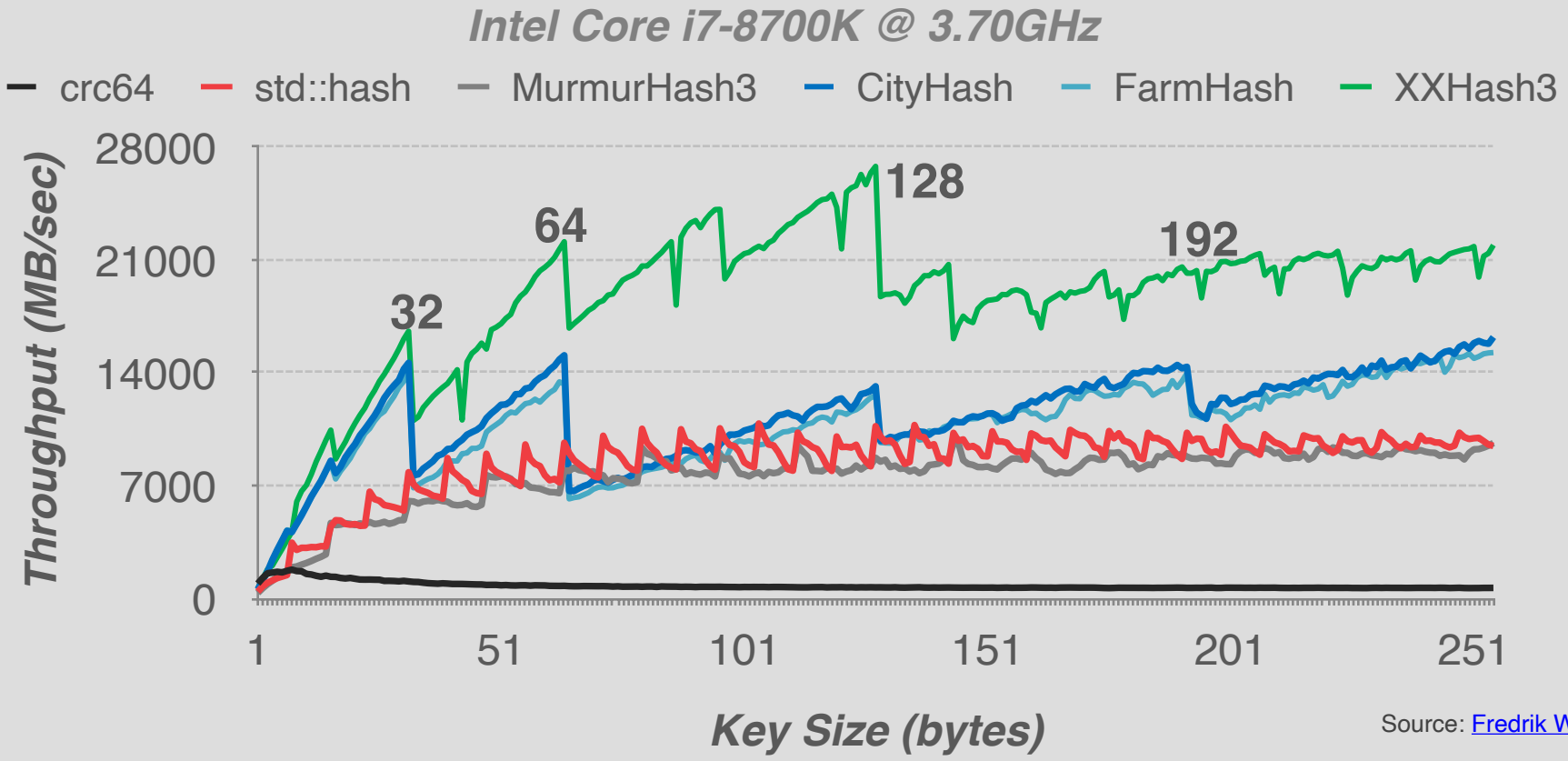
Google CityHash (2011)
→ Designed to be faster for short keys (<64 bytes).

Facebook XXHash (2012)
→ From the creator of zstd compression.

← State-of-the-art

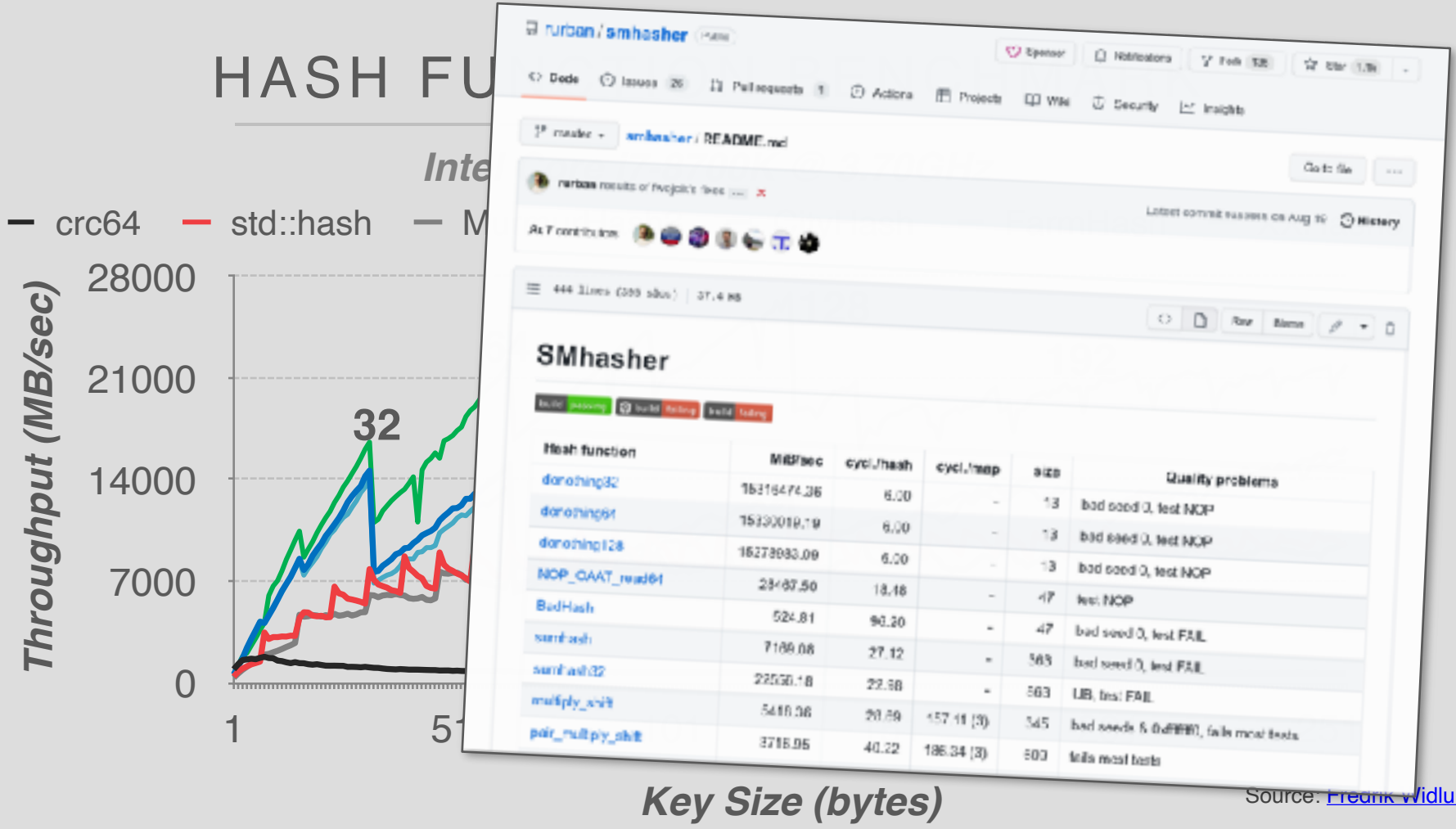
Google FarmHash (2014)
→ Newer version of CityHash with better collision rates.

HASH FUNCTION BENCHMARK



Source: Fredrik Widlund

HASH FU



Source: Fredrik Widlund

STATIC HASHING SCHEMES

Approach #1: Linear Probe Hashing

Approach #2: Robin Hood Hashing

Approach #3: Cuckoo Hashing

LINEAR PROBE HASHING

Single giant table of slots.

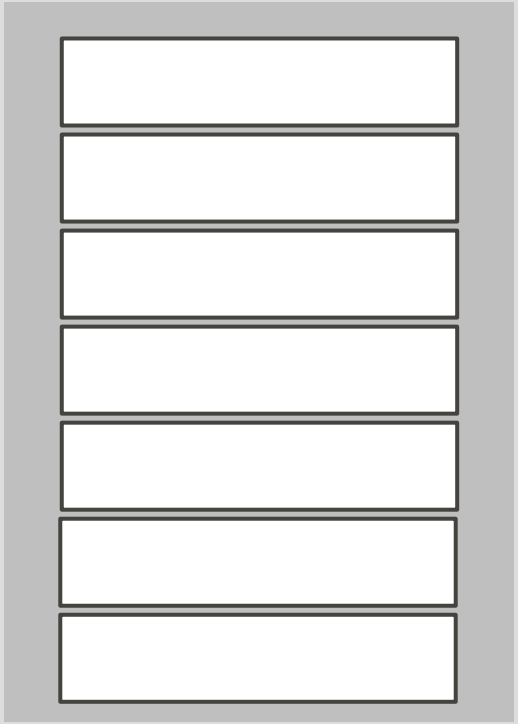
Resolve collisions by linearly searching for the next free slot in the table.

→ To determine whether an element is present, hash to a location in the index and scan for it.
→ Must store the key in the index to know when to stop scanning.
→ Insertions and deletions are generalizations of lookups.

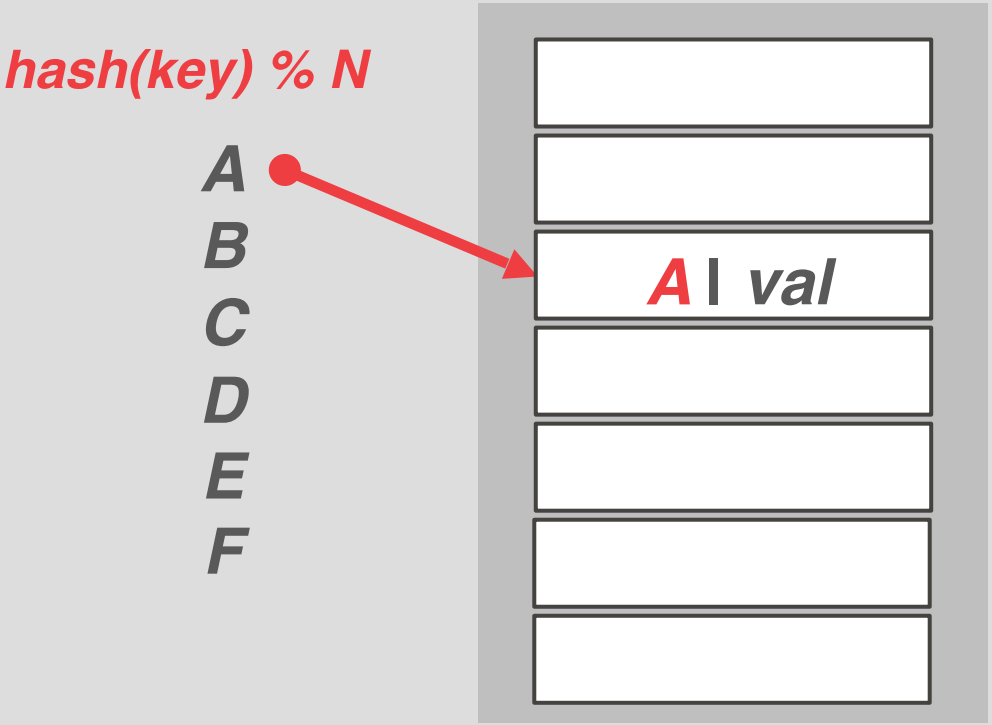
LINEAR PROBE HASHING

hash(key) % N

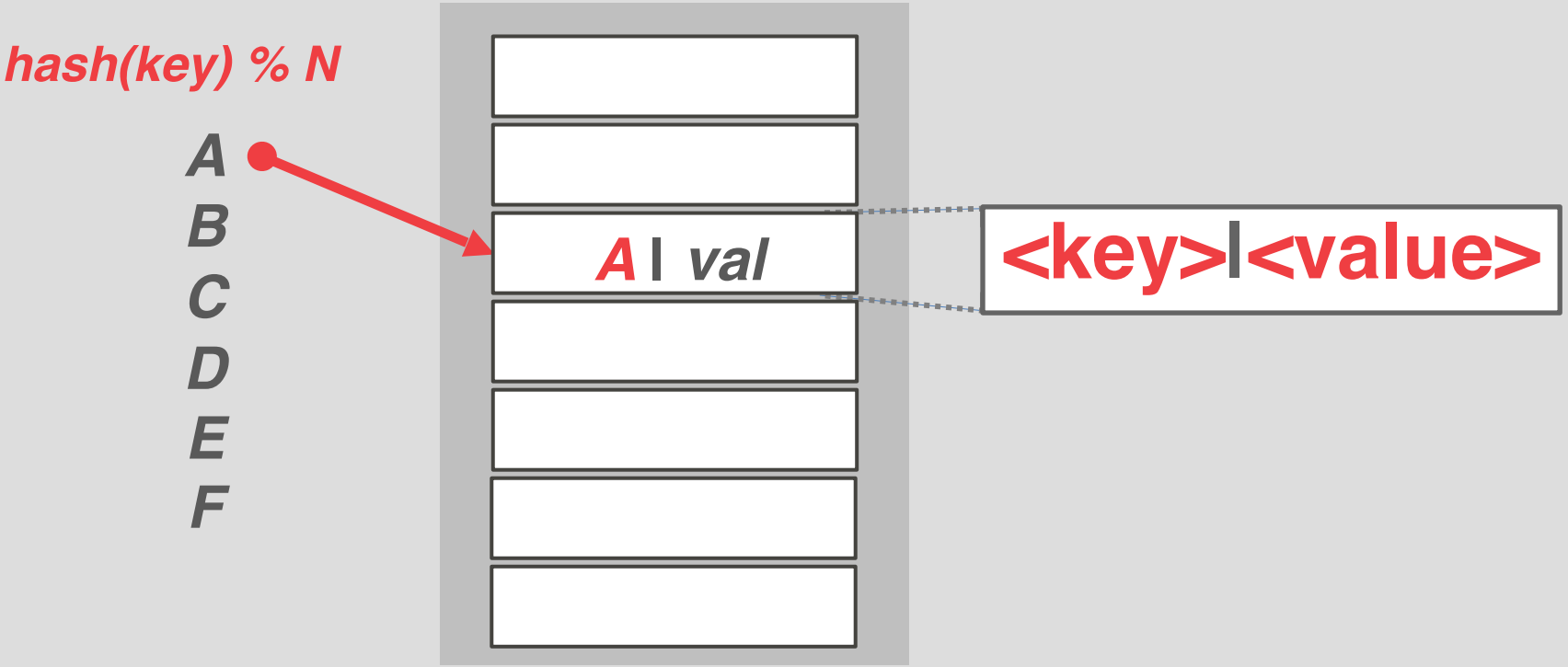
A
B
C
D
E
F



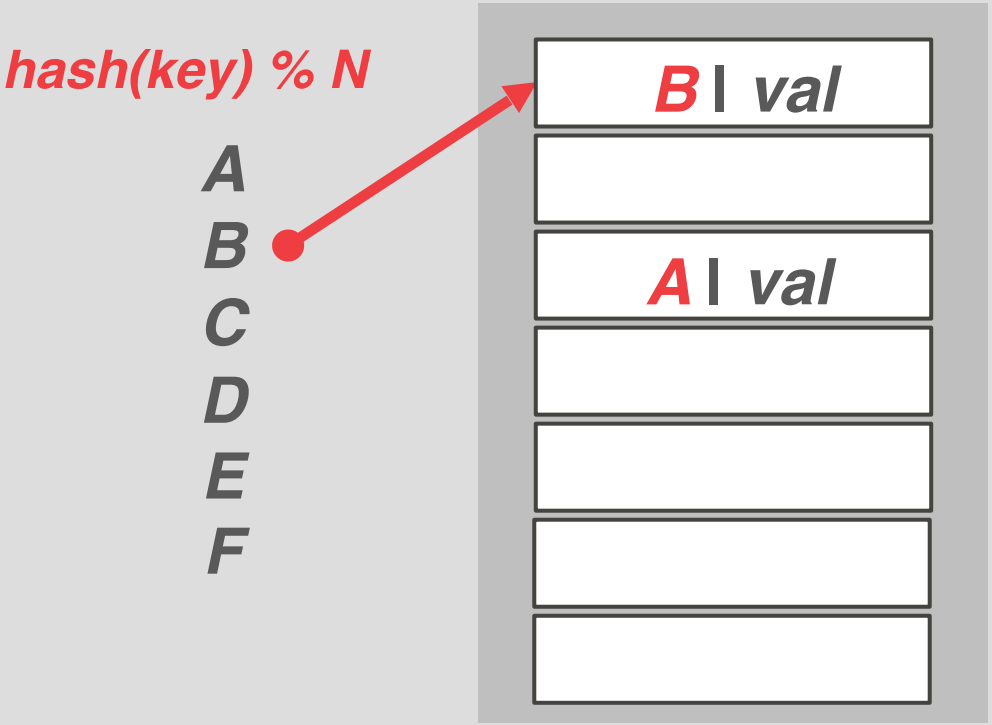
LINEAR PROBE HASHING



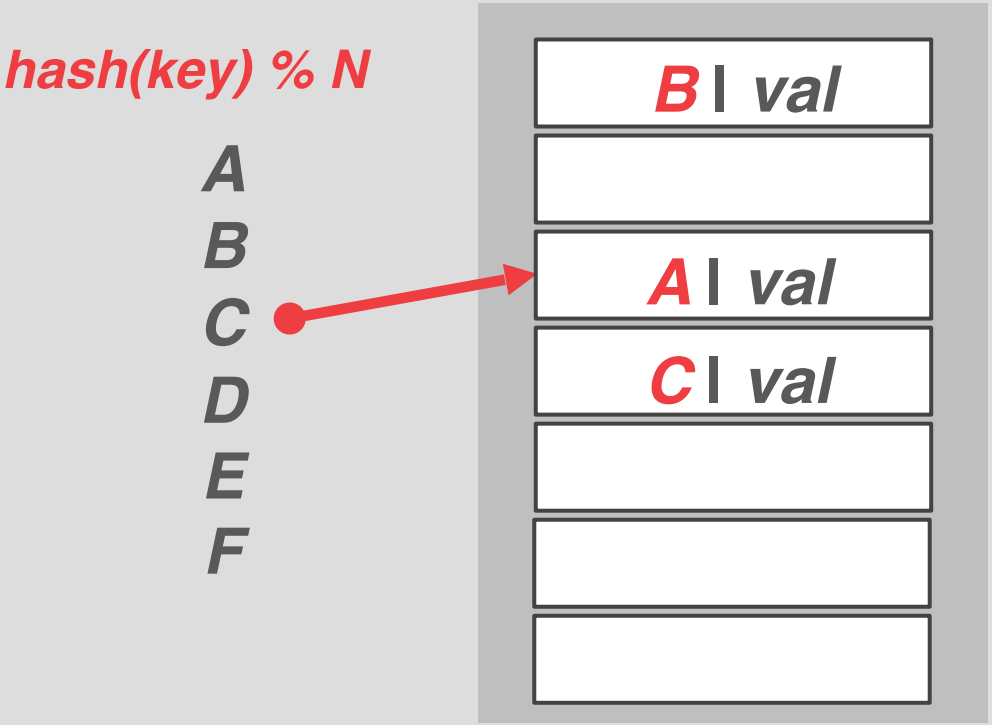
LINEAR PROBE HASHING



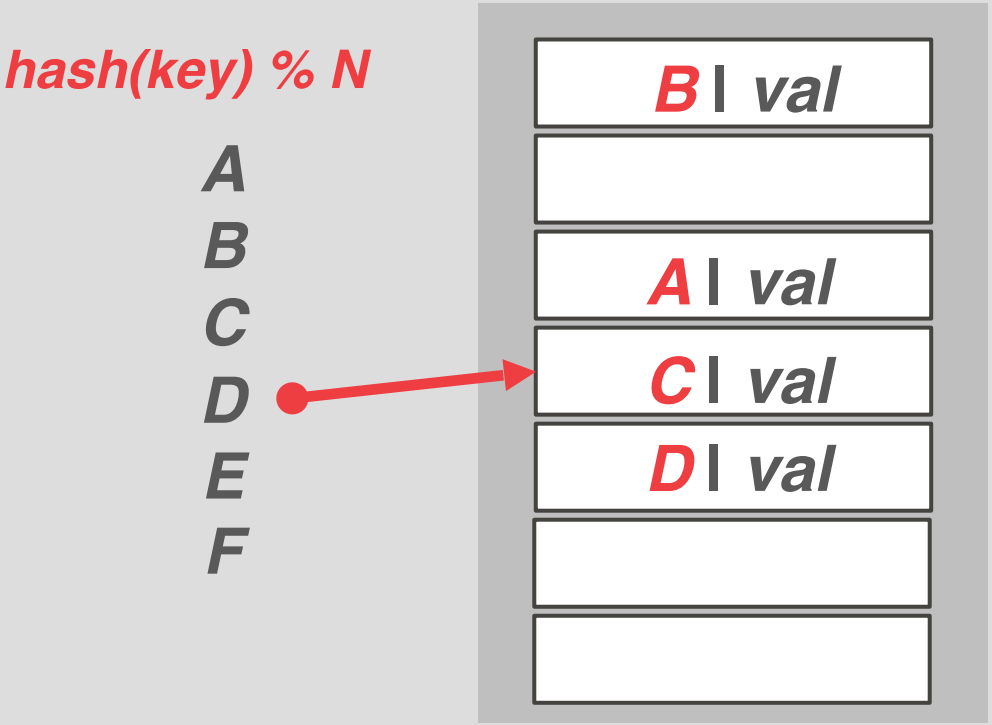
LINEAR PROBE HASHING



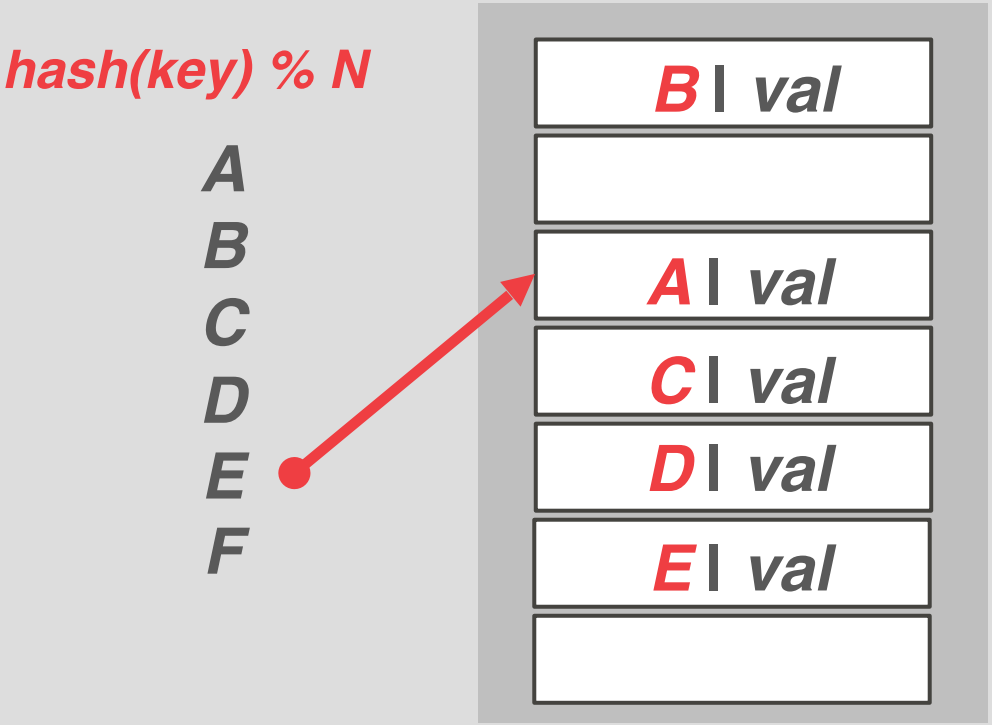
LINEAR PROBE HASHING



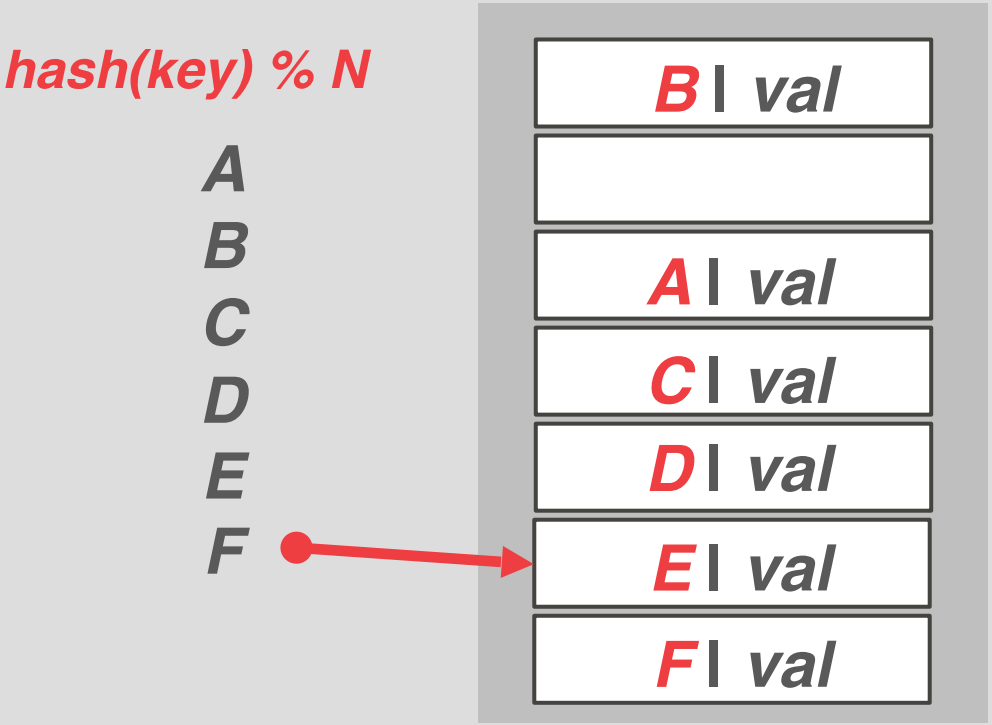
LINEAR PROBE HASHING



LINEAR PROBE HASHING



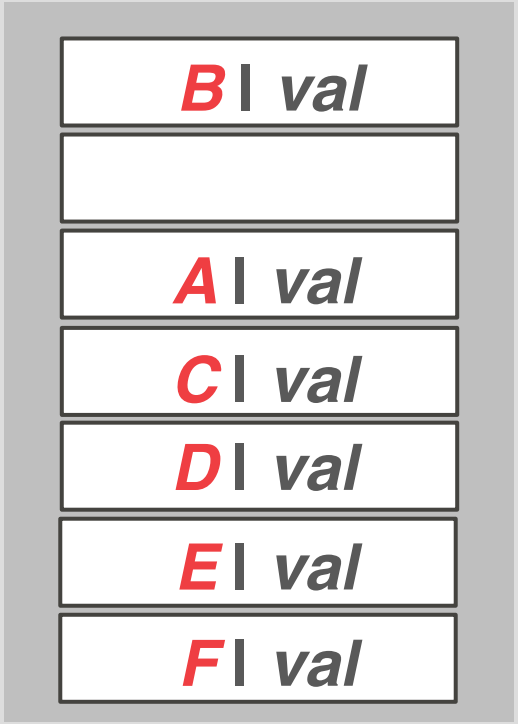
LINEAR PROBE HASHING



LINEAR PROBE HASHING – DELETES

hash(key) % N

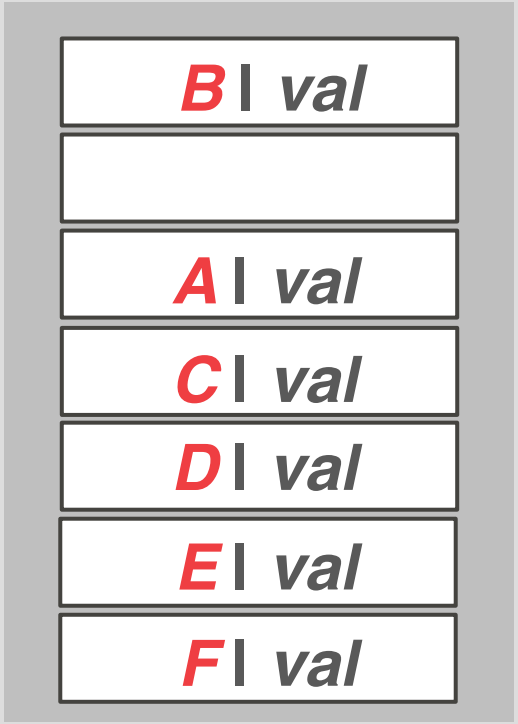
A
B
C
D
E
F



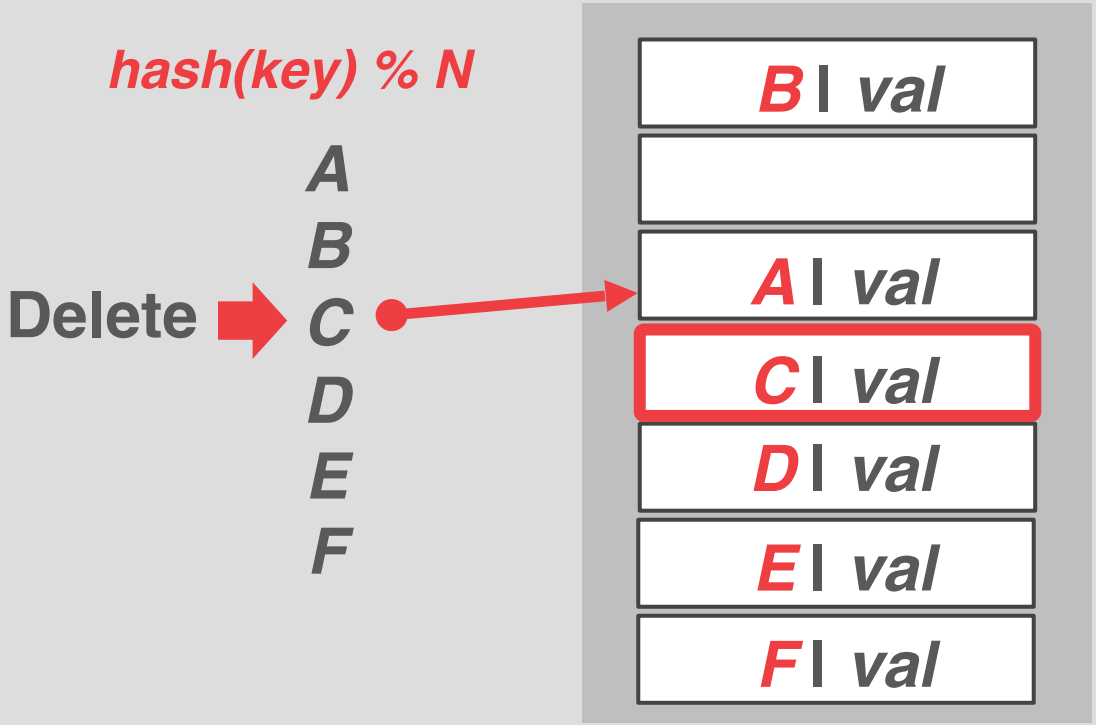
LINEAR PROBE HASHING – DELETES

hash(key) % N

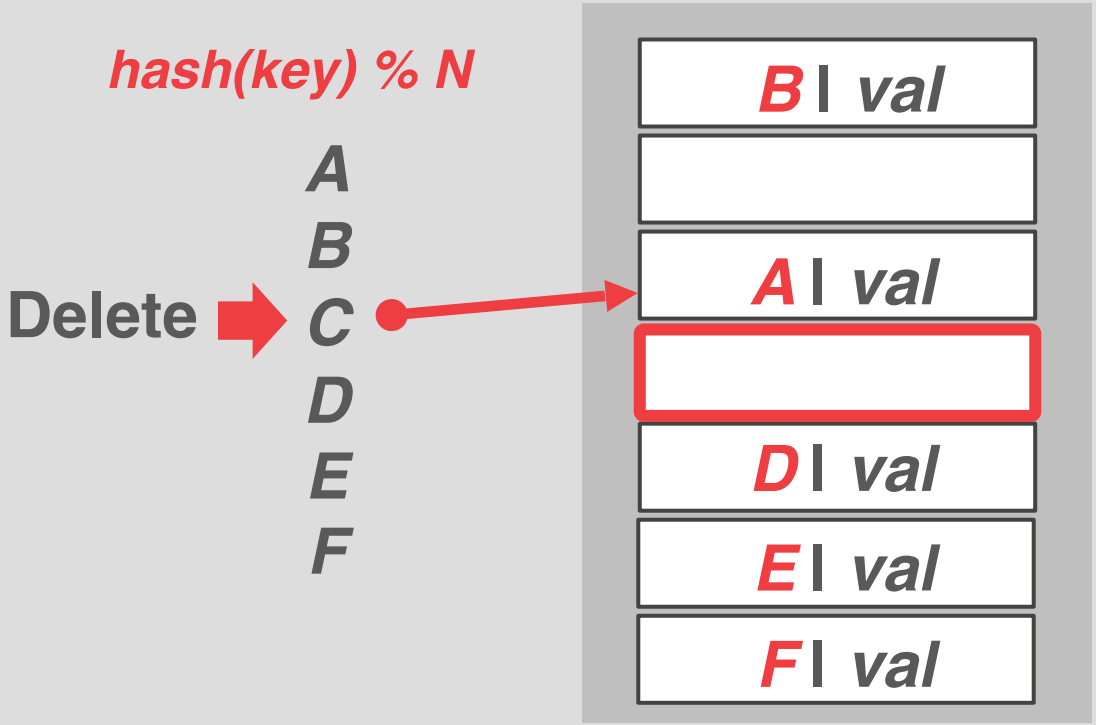
A
B
Delete → *C*
D
E
F



LINEAR PROBE HASHING – DELETES



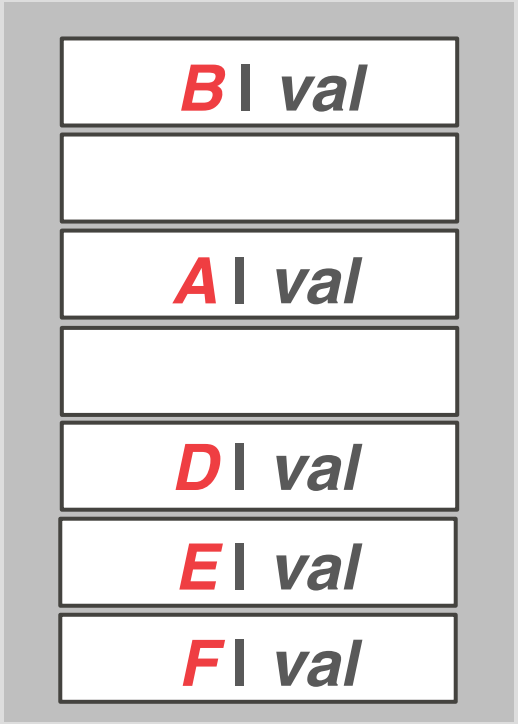
LINEAR PROBE HASHING – DELETES



LINEAR PROBE HASHING – DELETES

hash(key) % N

A
B
C
D
E
F

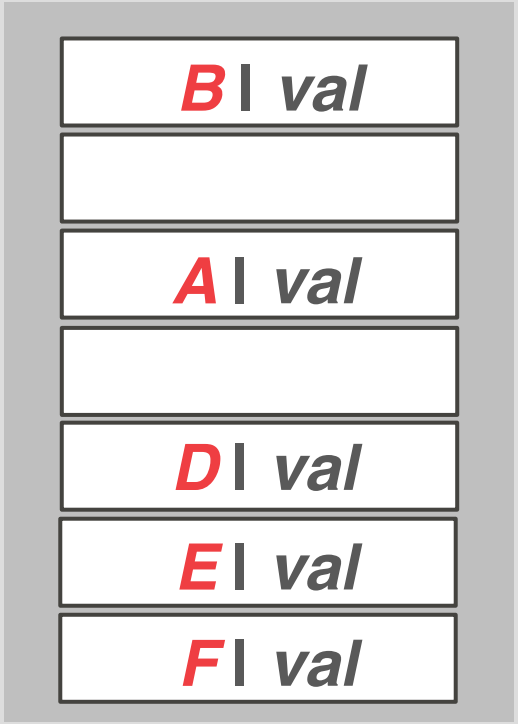


LINEAR PROBE HASHING – DELETES

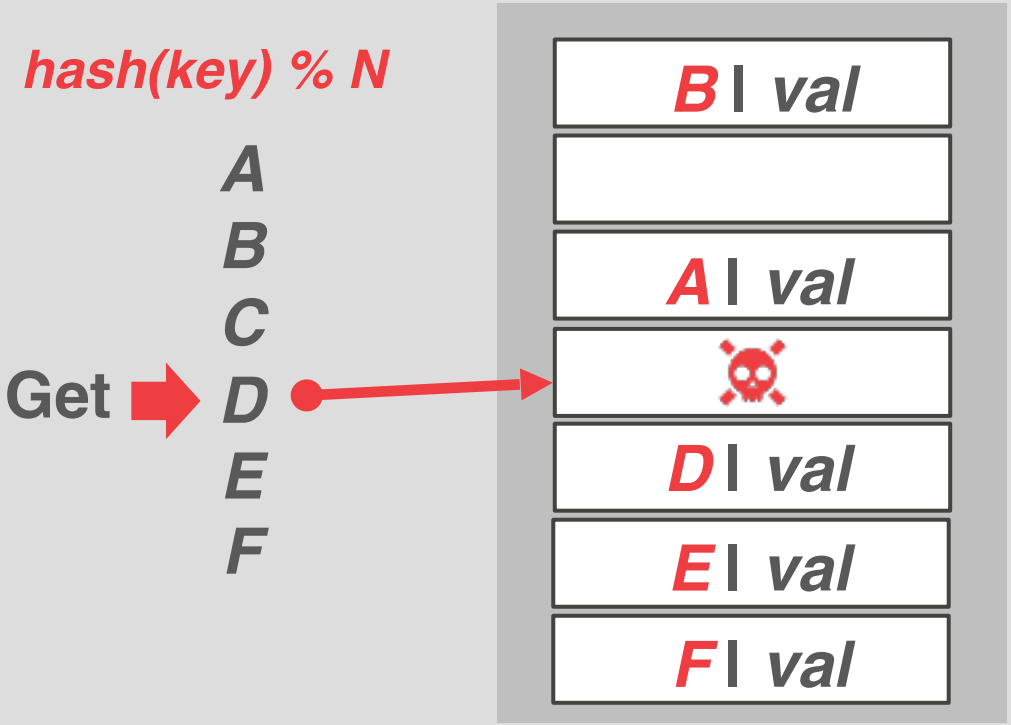
hash(key) % N

Get

A
B
C
D
E
F



LINEAR PROBE HASHING – DELETES

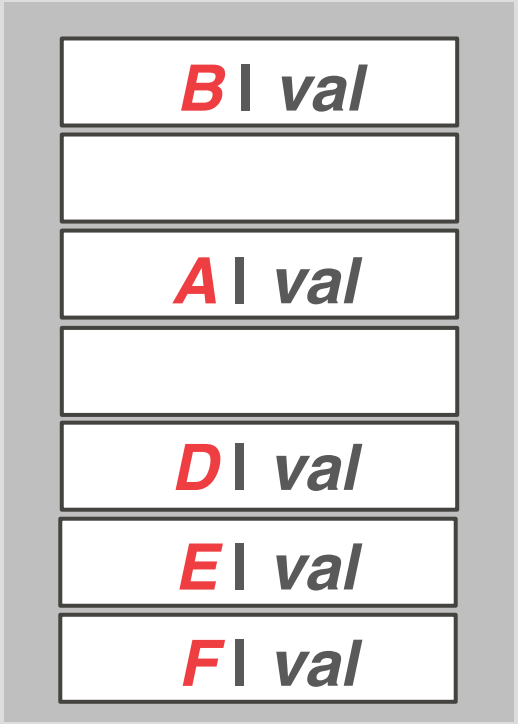


LINEAR PROBE HASHING – DELETES

hash(key) % N

Get

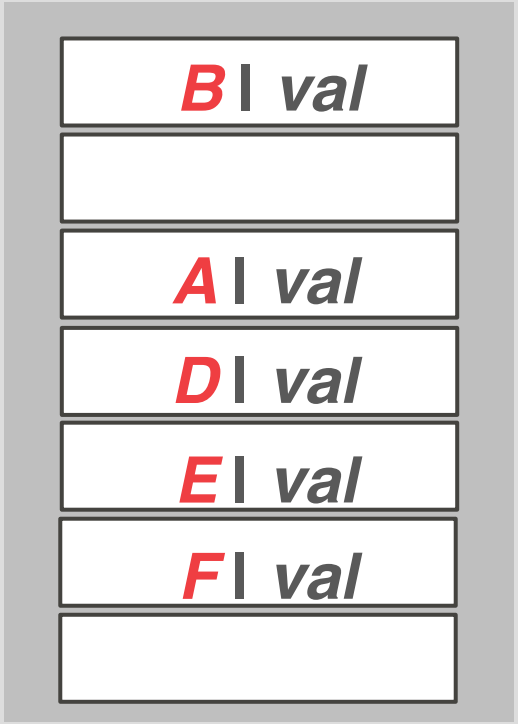
A
B
C
D
E
F



LINEAR PROBE HASHING – DELETES

hash(key) % N

A
B
C
Get → *D*
E
F

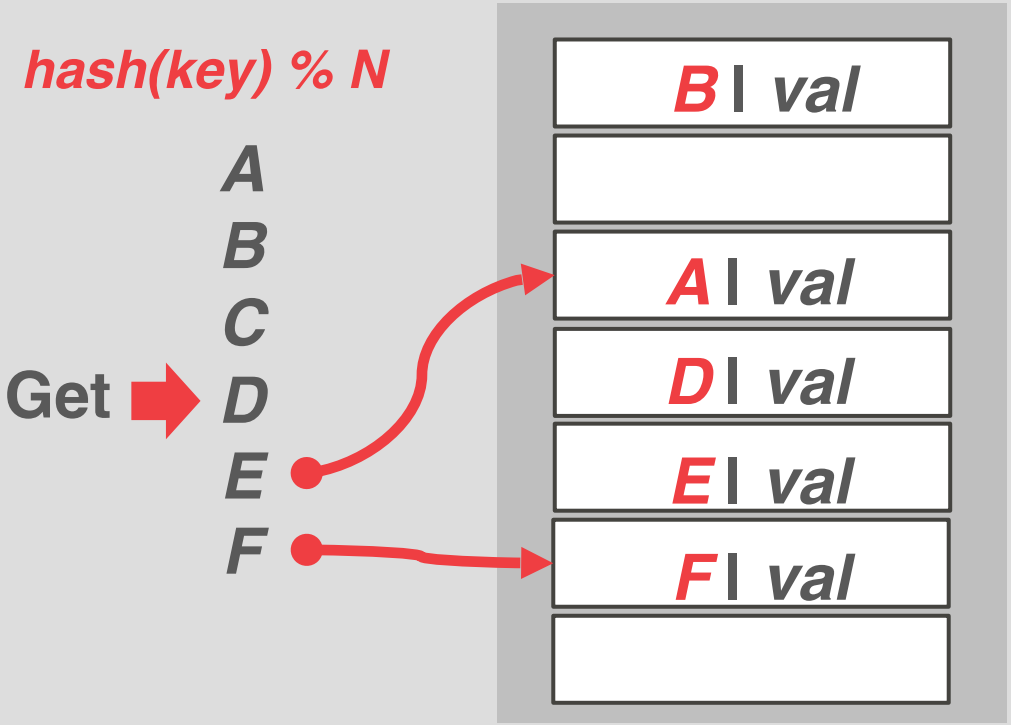


Approach #1: Re-hash

→ Re-hash and update the slots for all keys.

→ Nobody actually does this.

LINEAR PROBE HASHING – DELETES



Approach #1: Re-hash

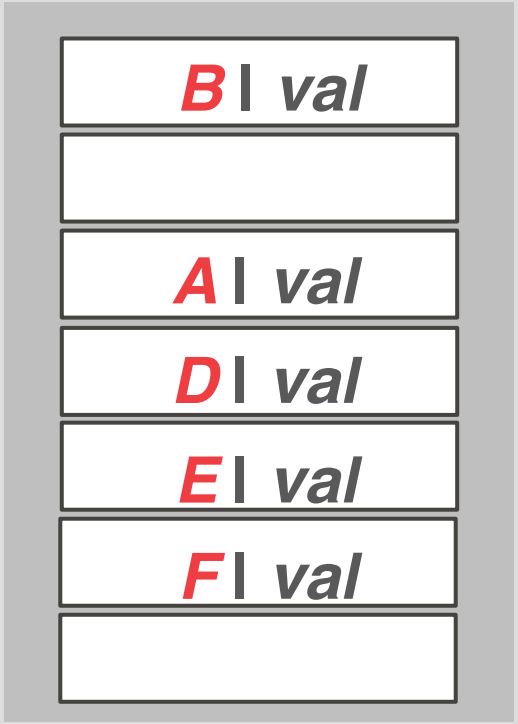
→ Re-hash and update the slots for all keys.

→ Nobody actually does this.

LINEAR PROBE HASHING – DELETES

hash(key) % N

A
B
C
Get → *D*
E
F



Approach #1: Re-hash

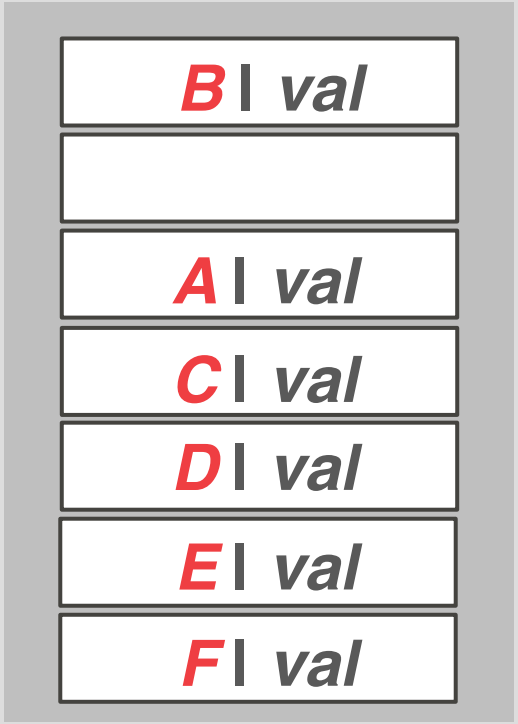
→ Re-hash and update the slots for all keys.

→ Nobody actually does this.

LINEAR PROBE HASHING – DELETES

hash(key) % N

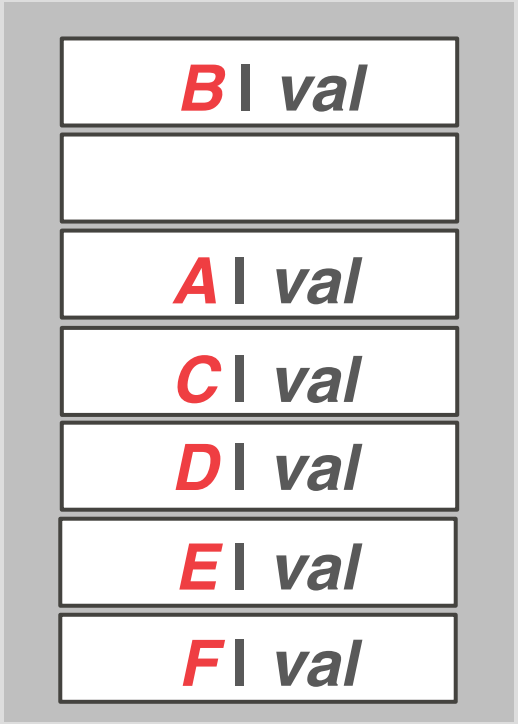
A
B
C
D
E
F



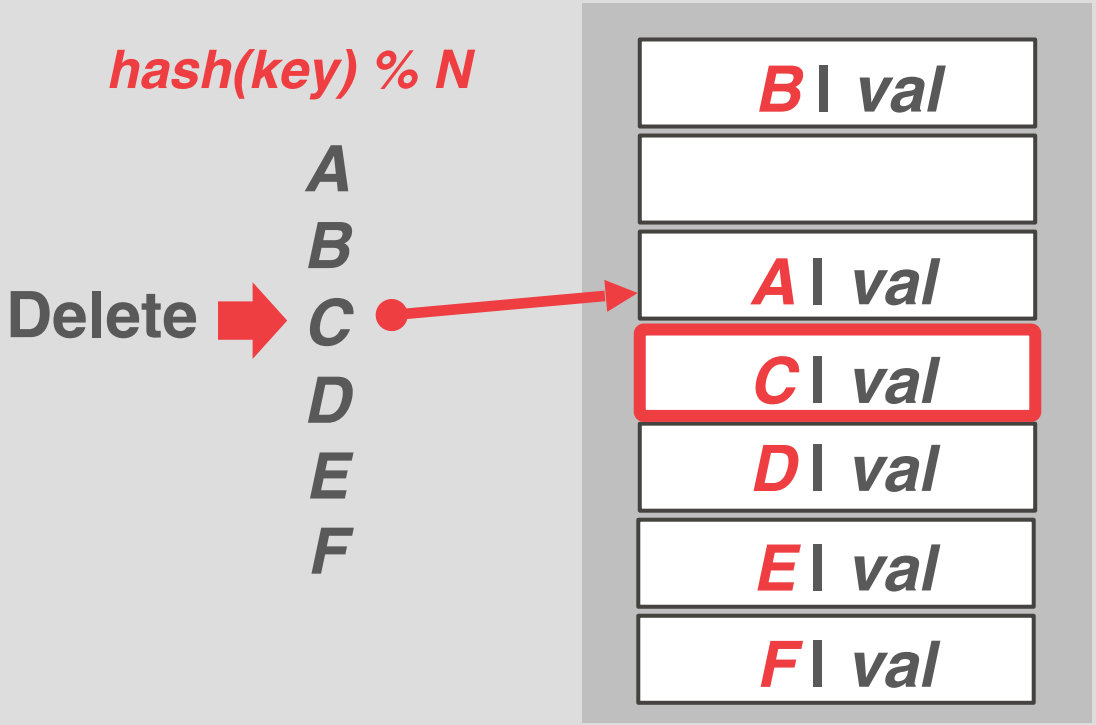
LINEAR PROBE HASHING – DELETES

hash(key) % N

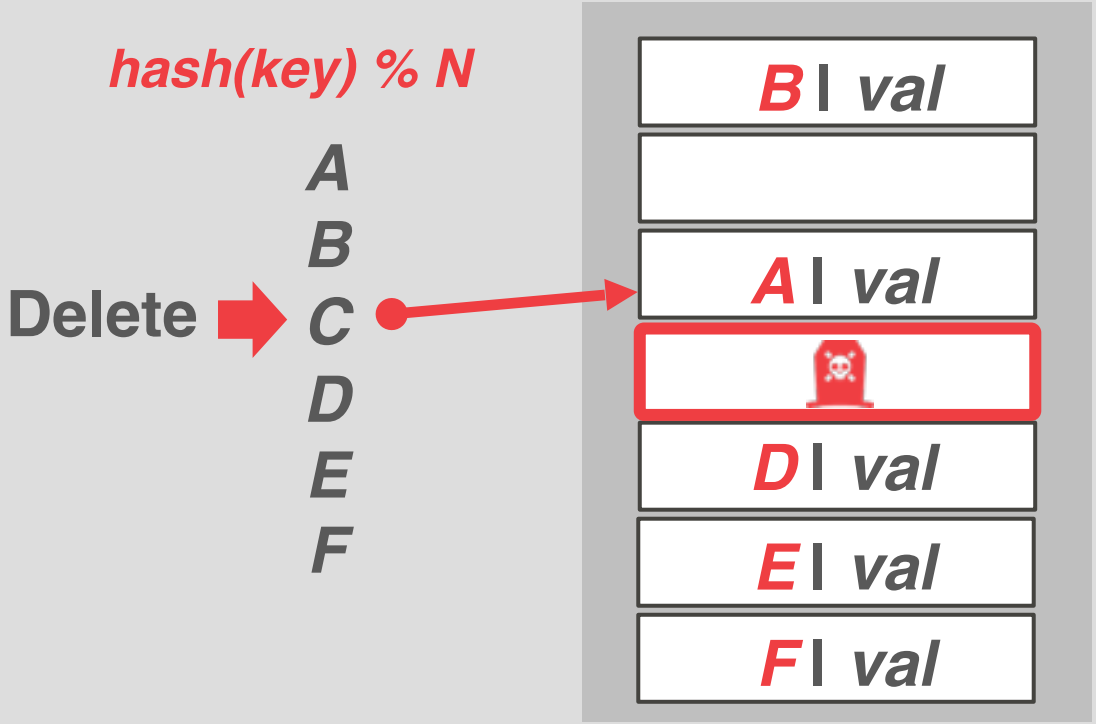
A
B
Delete ➡ *C*
D
E
F



LINEAR PROBE HASHING – DELETES



LINEAR PROBE HASHING – DELETES



Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

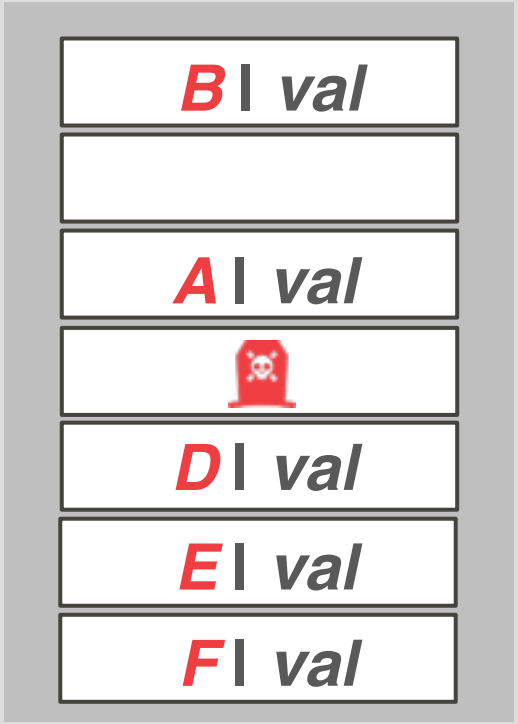
→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES

hash(key) % N

A
B
C
D
E
F



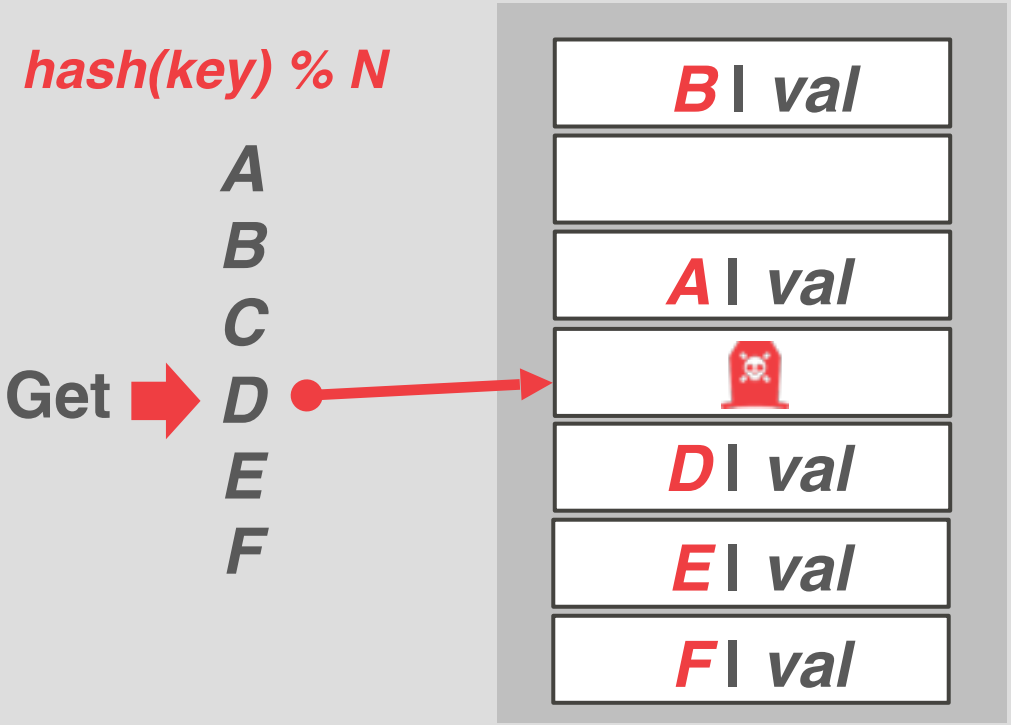
Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES



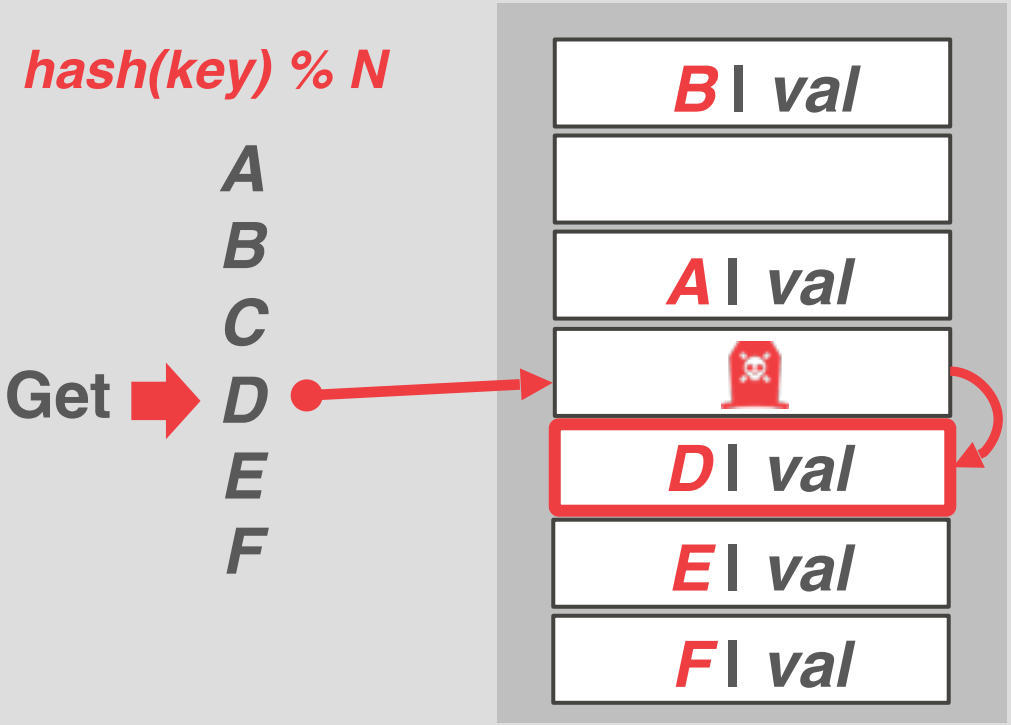
Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES



Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

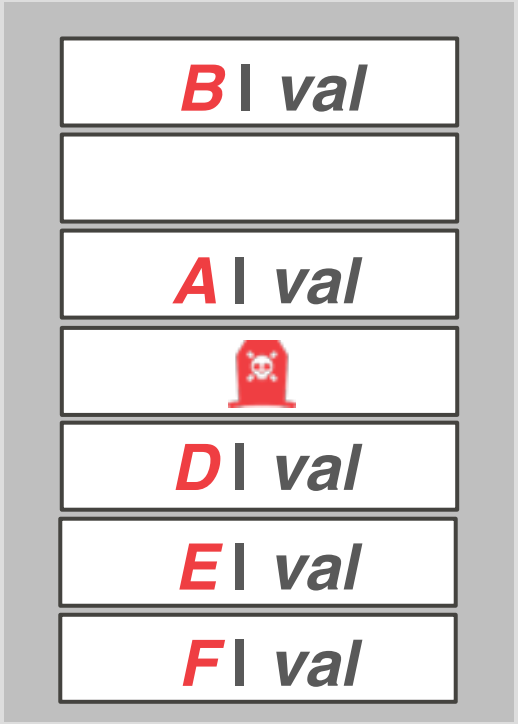
→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES

hash(key) % N

A
B
C
D
E
F



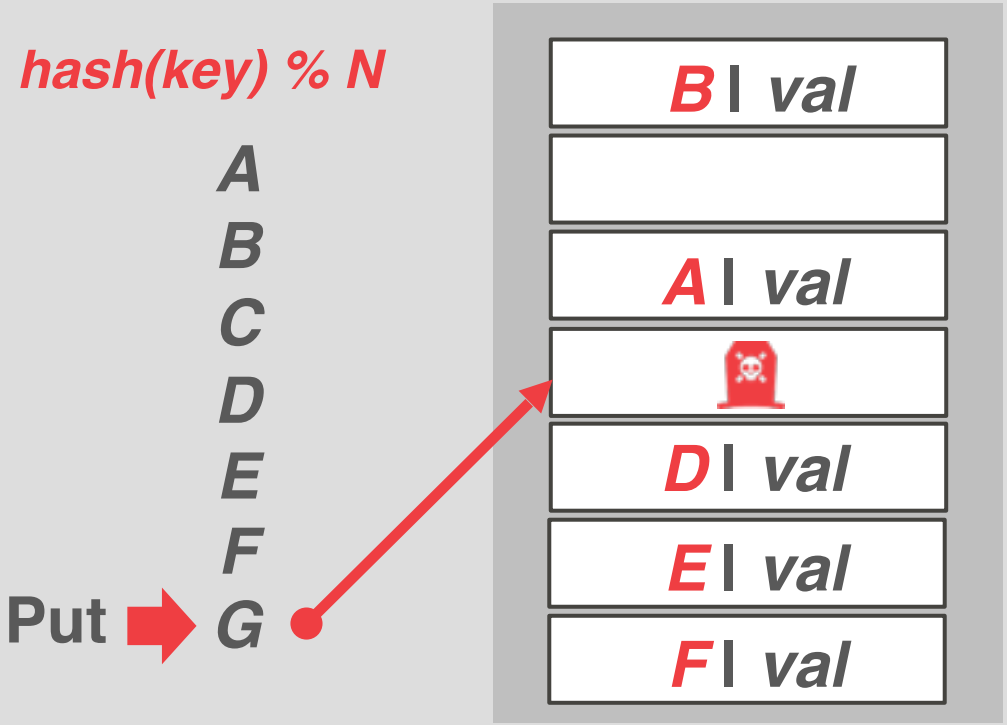
Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES



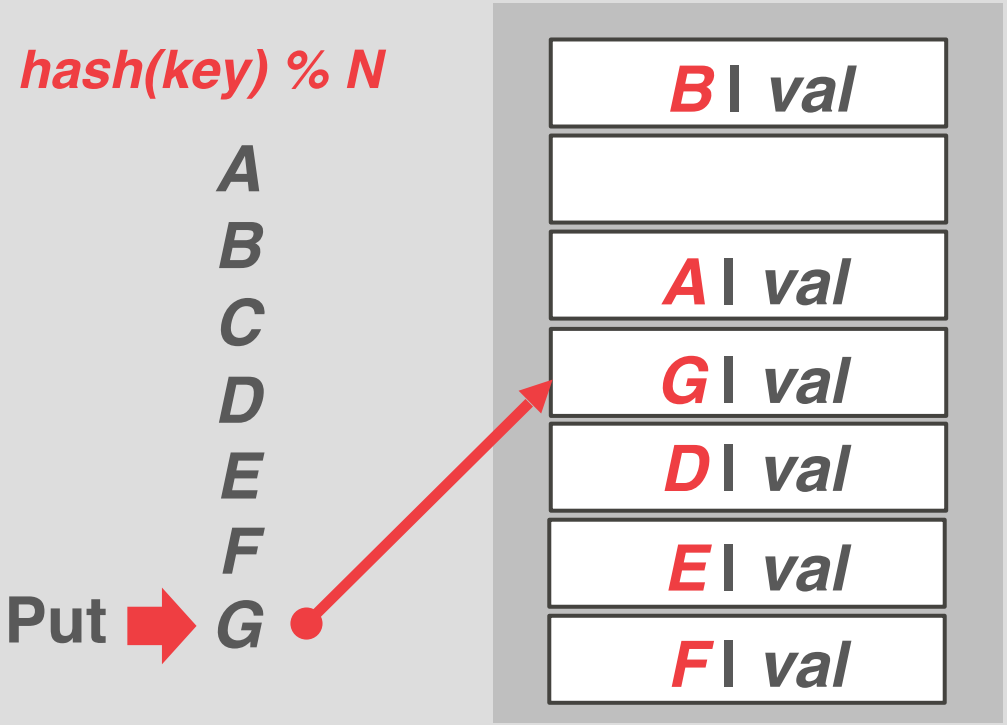
Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

LINEAR PROBE HASHING – DELETES



Approach #2: Tombstone

→ Set a marker indicating that the entry in the slot is logically deleted.

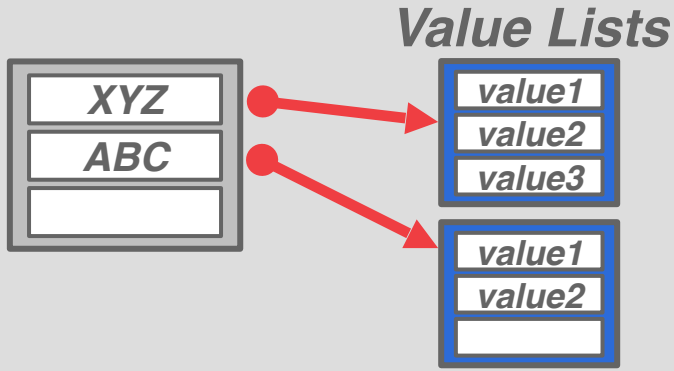
→ You can reuse the slot for inserting new keys.

→ May require periodic garbage collection.

NON-UNIQUE KEYS

Choice #1: Separate Linked List

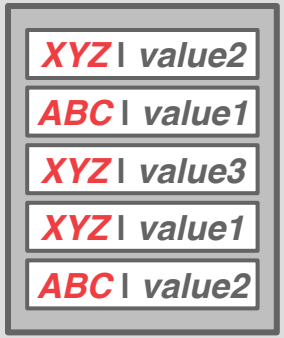
→ Store values in separate storage area for each key.



Choice #2: Redundant Keys

→ Store duplicate keys entries together in the hash table.

→ This is easier to implement so this is what most systems do.



ROBIN HOOD HASHING

Variant of linear probe hashing that steals slots from "rich" keys and give them to "poor" keys.

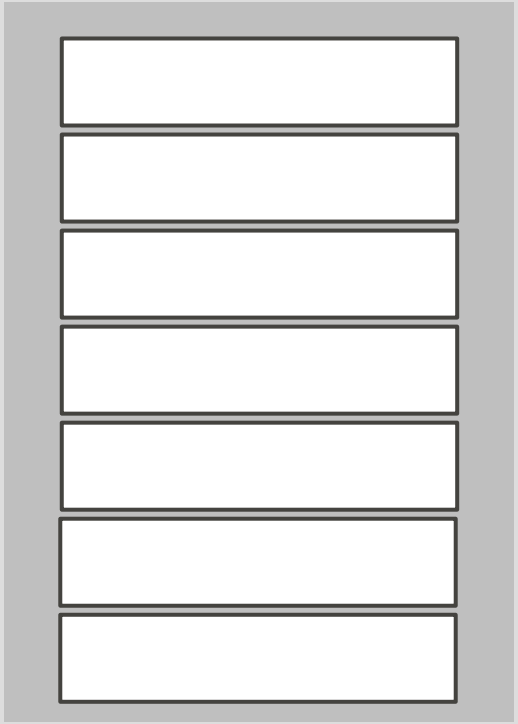
→ Each key tracks the number of positions they are from its optimal position in the table.

→ On insert, a key takes the slot of another key if the first key is farther away from its optimal position than the second key.

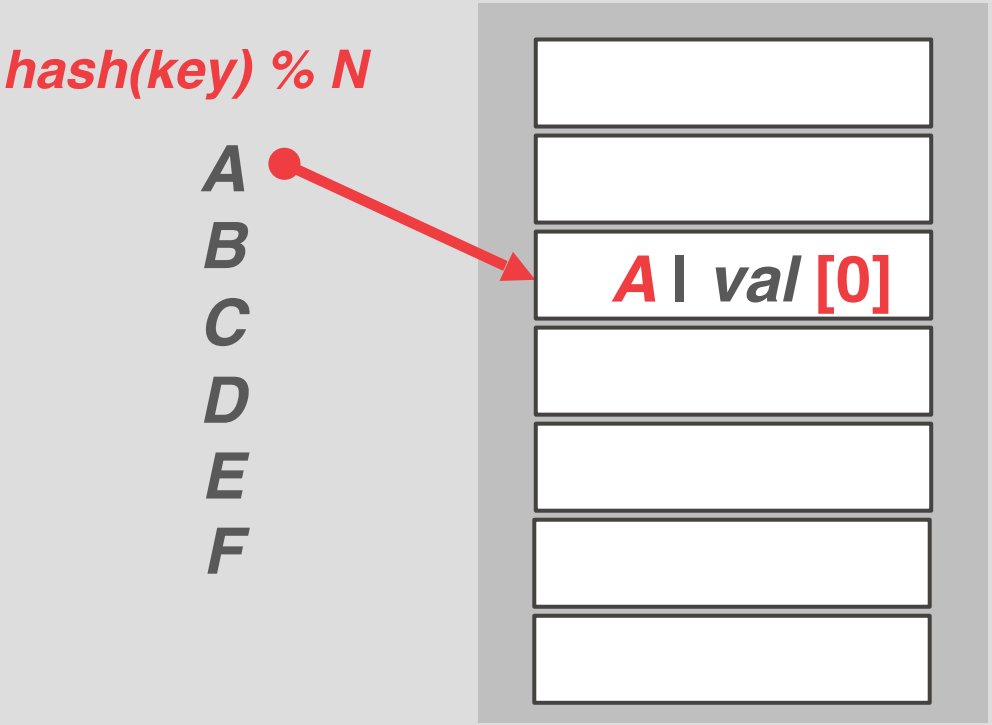
ROBIN HOOD HASHING

hash(key) % N

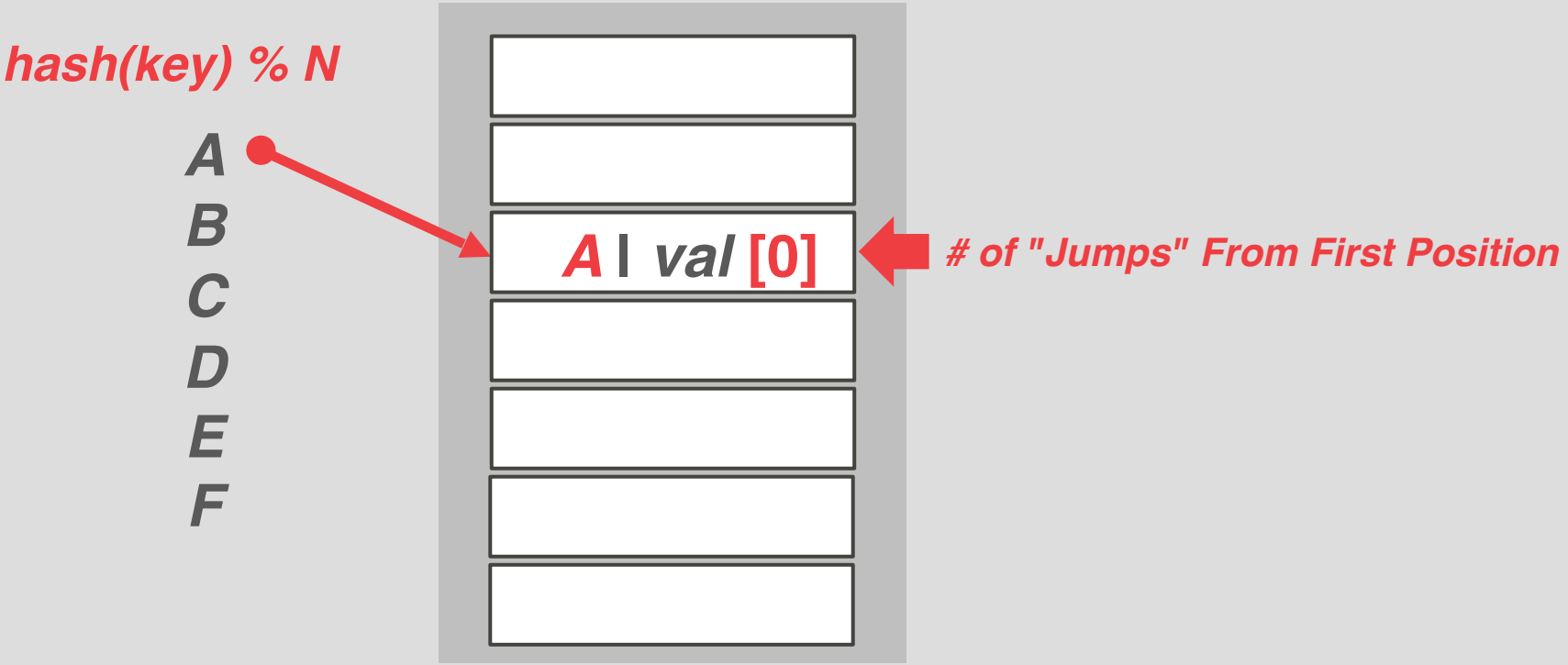
A
B
C
D
E
F



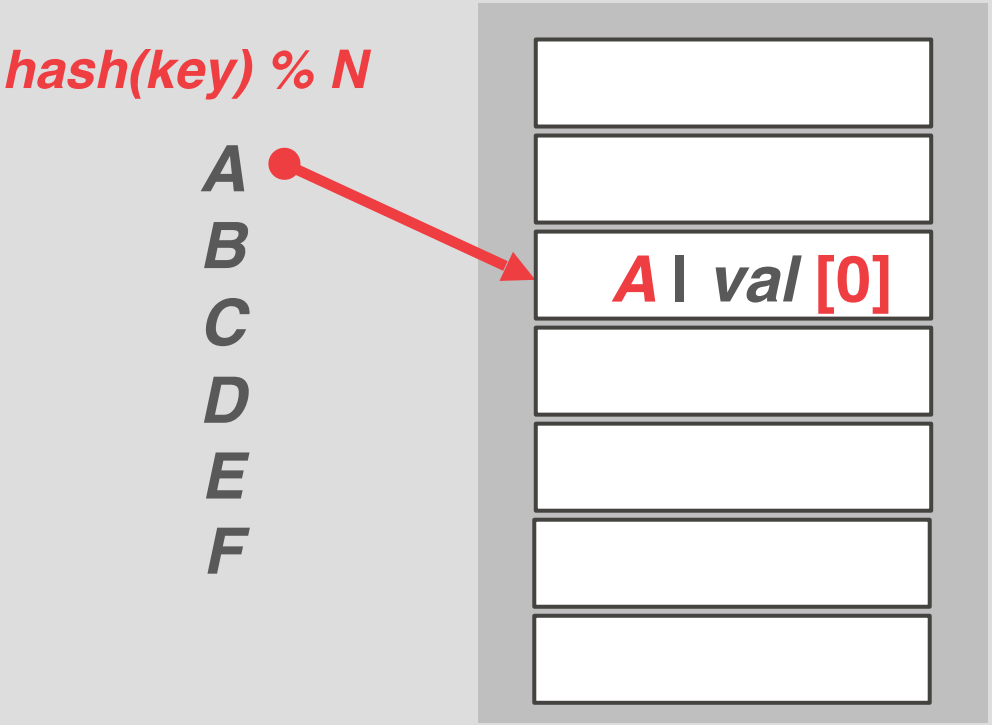
ROBIN HOOD HASHING



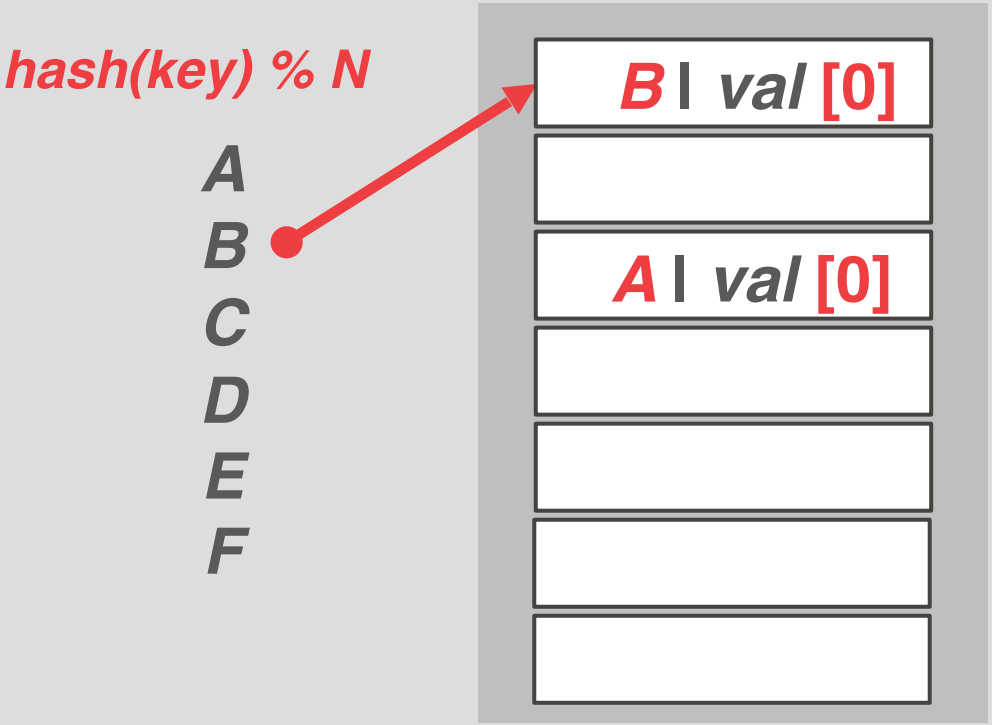
ROBIN HOOD HASHING



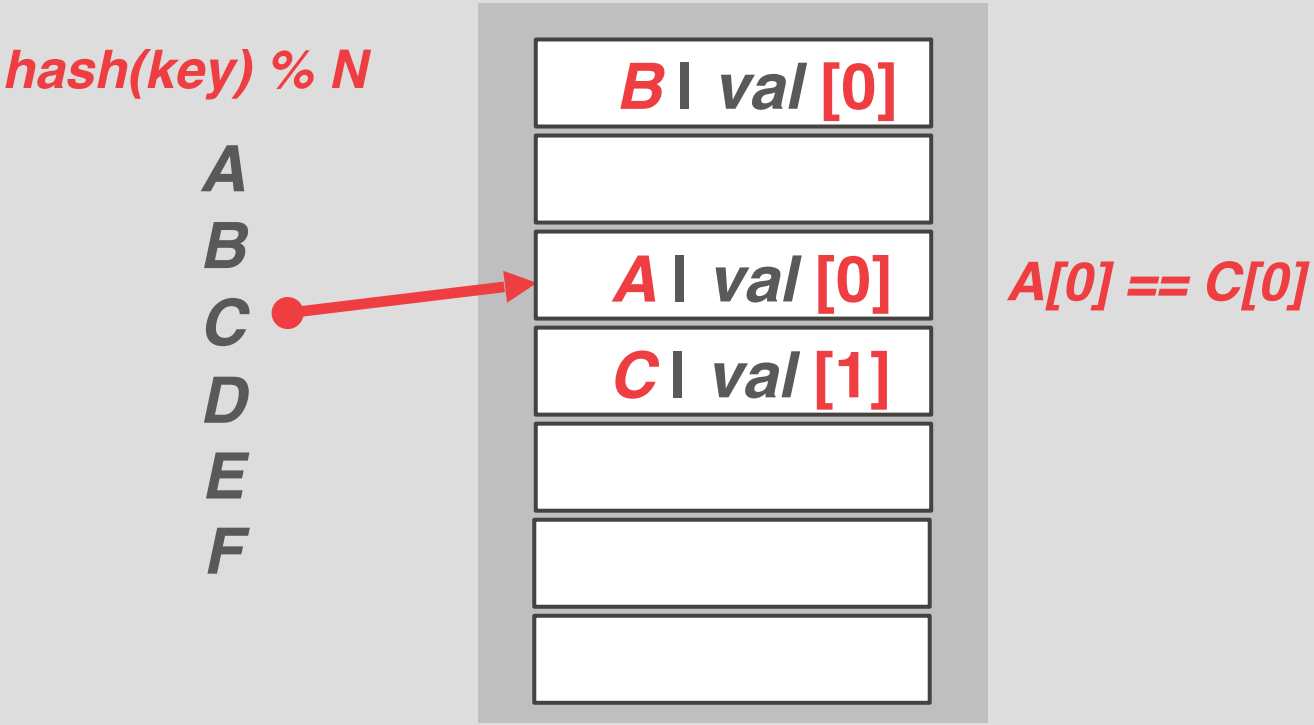
ROBIN HOOD HASHING



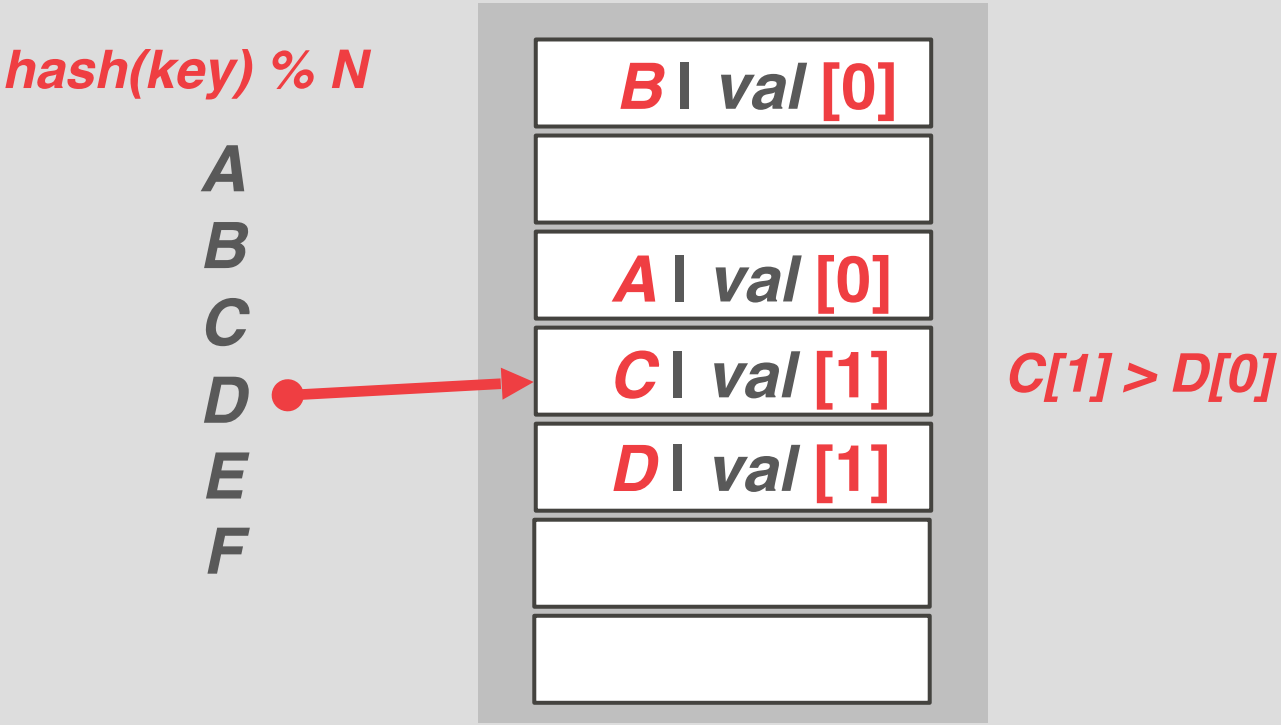
ROBIN HOOD HASHING



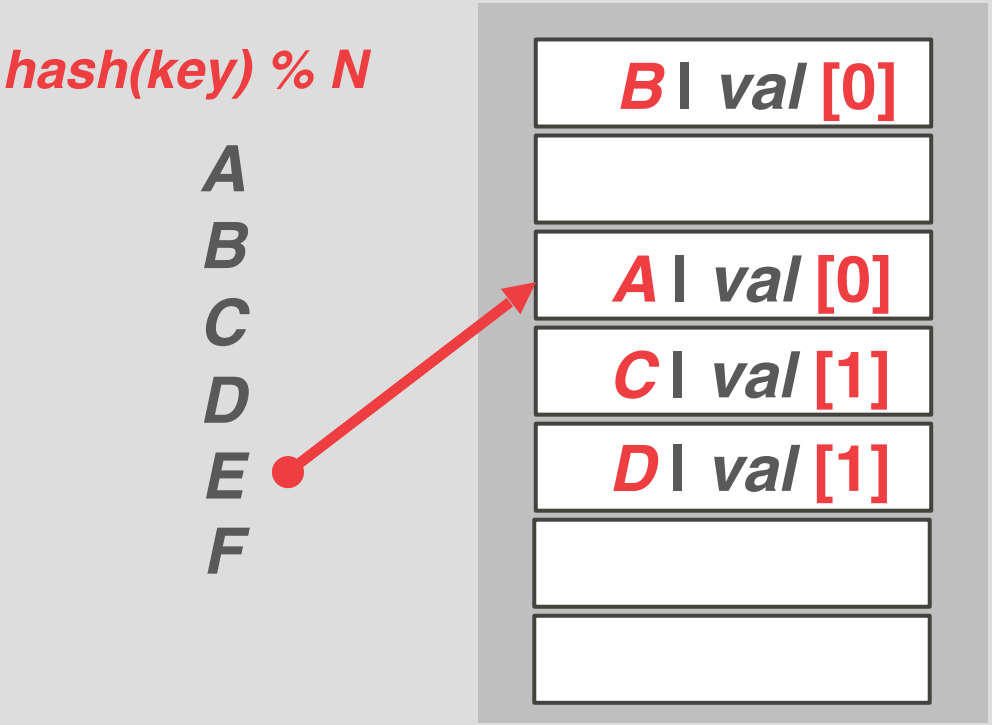
ROBIN HOOD HASHING



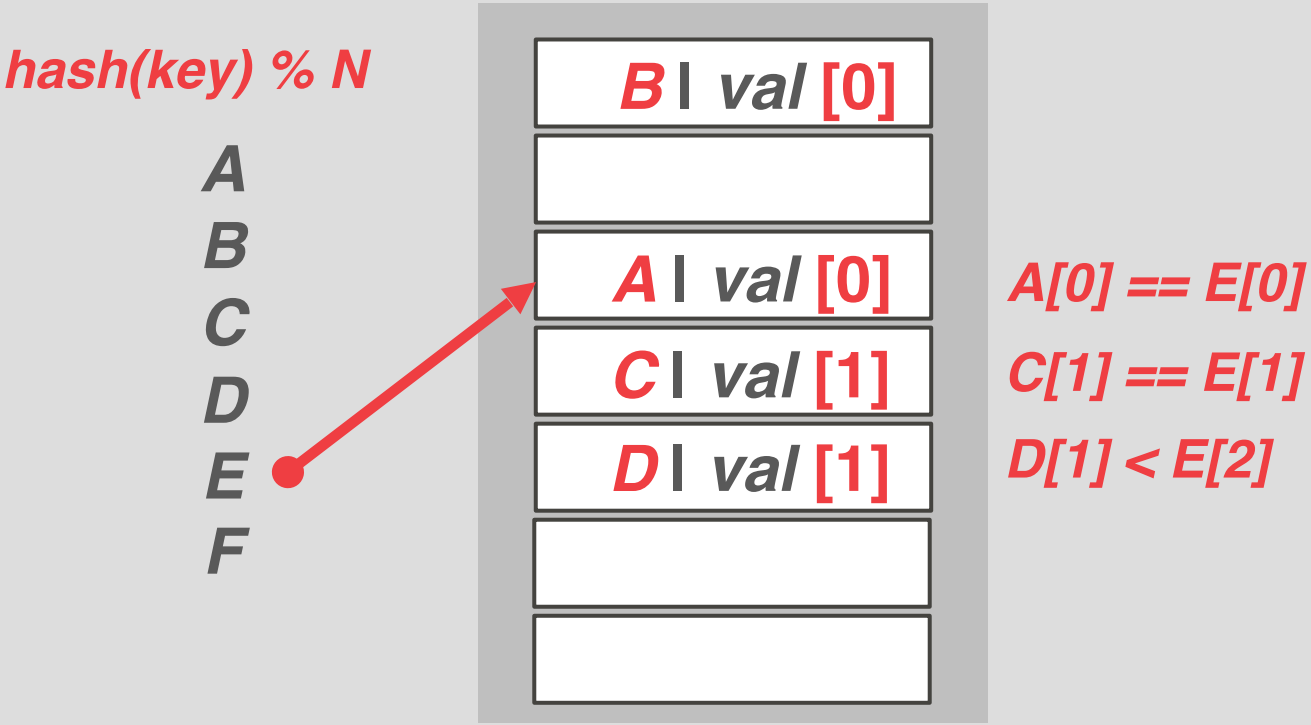
ROBIN HOOD HASHING



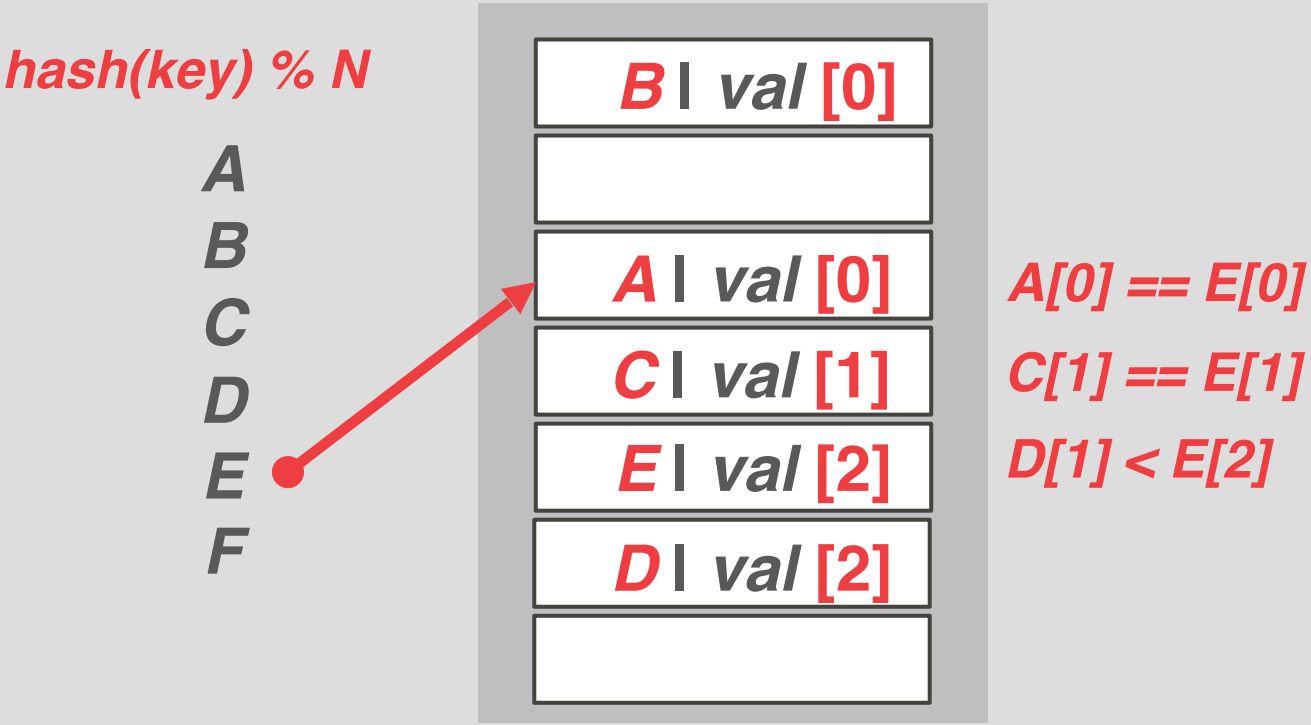
ROBIN HOOD HASHING



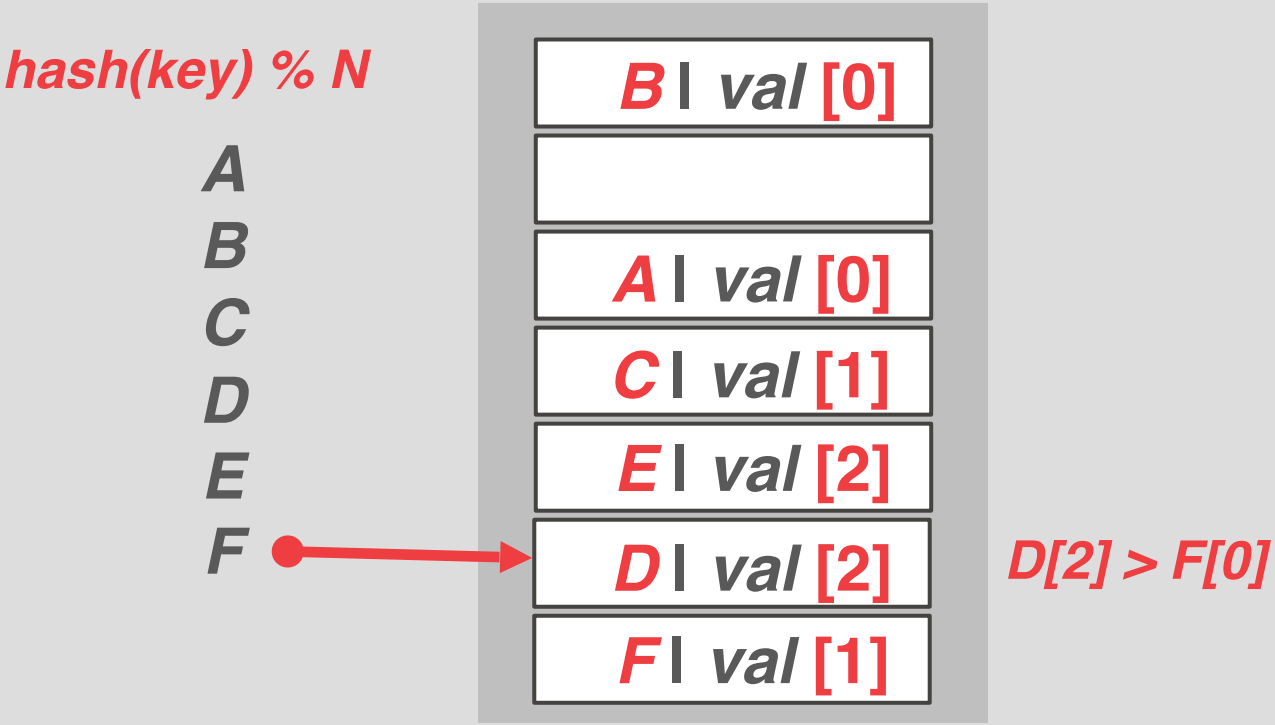
ROBIN HOOD HASHING



ROBIN HOOD HASHING



ROBIN HOOD HASHING



CUCKOO HASHING

Use multiple hash tables with different hash function seeds.

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

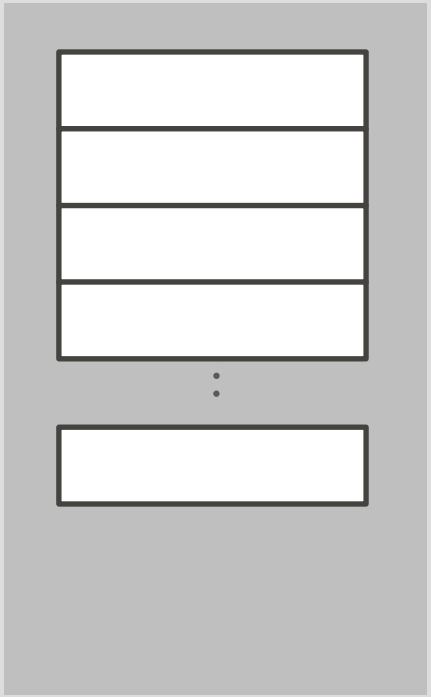
→ If no table has a free slot, evict the element from one of them and then re-hash it find a new location.

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

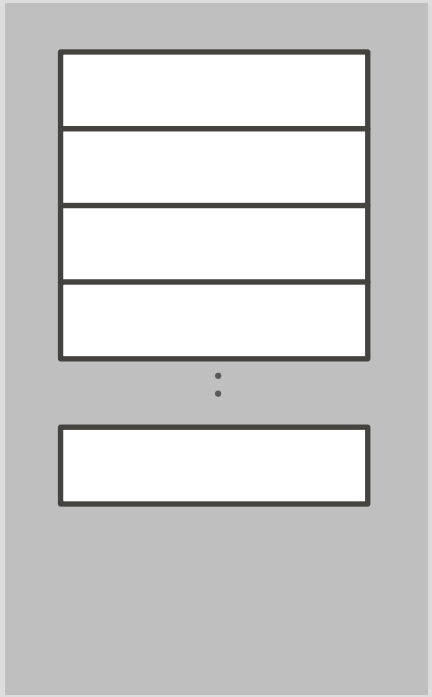
Best open-source implementation is from CMU.

CUCKOO HASHING

Hash Table #1



Hash Table #2

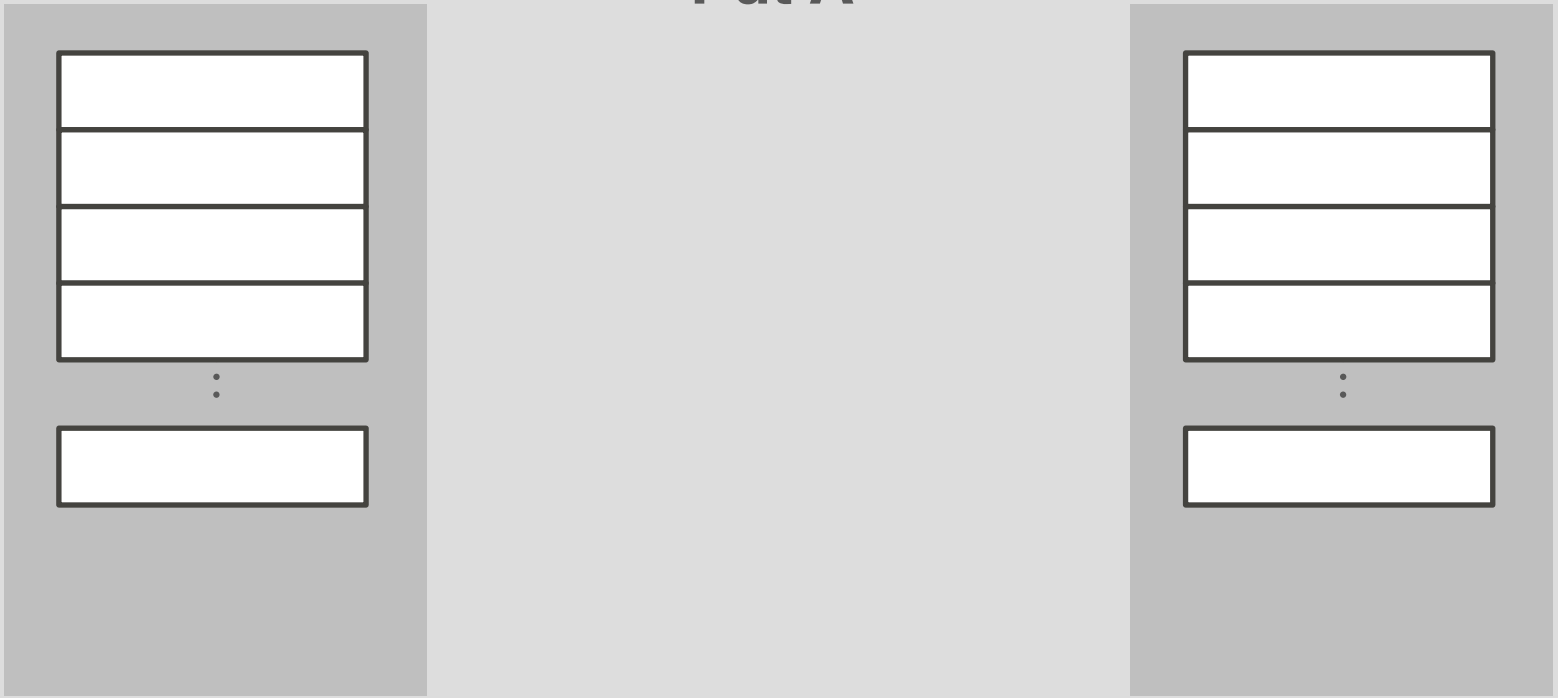


CUCKOO HASHING

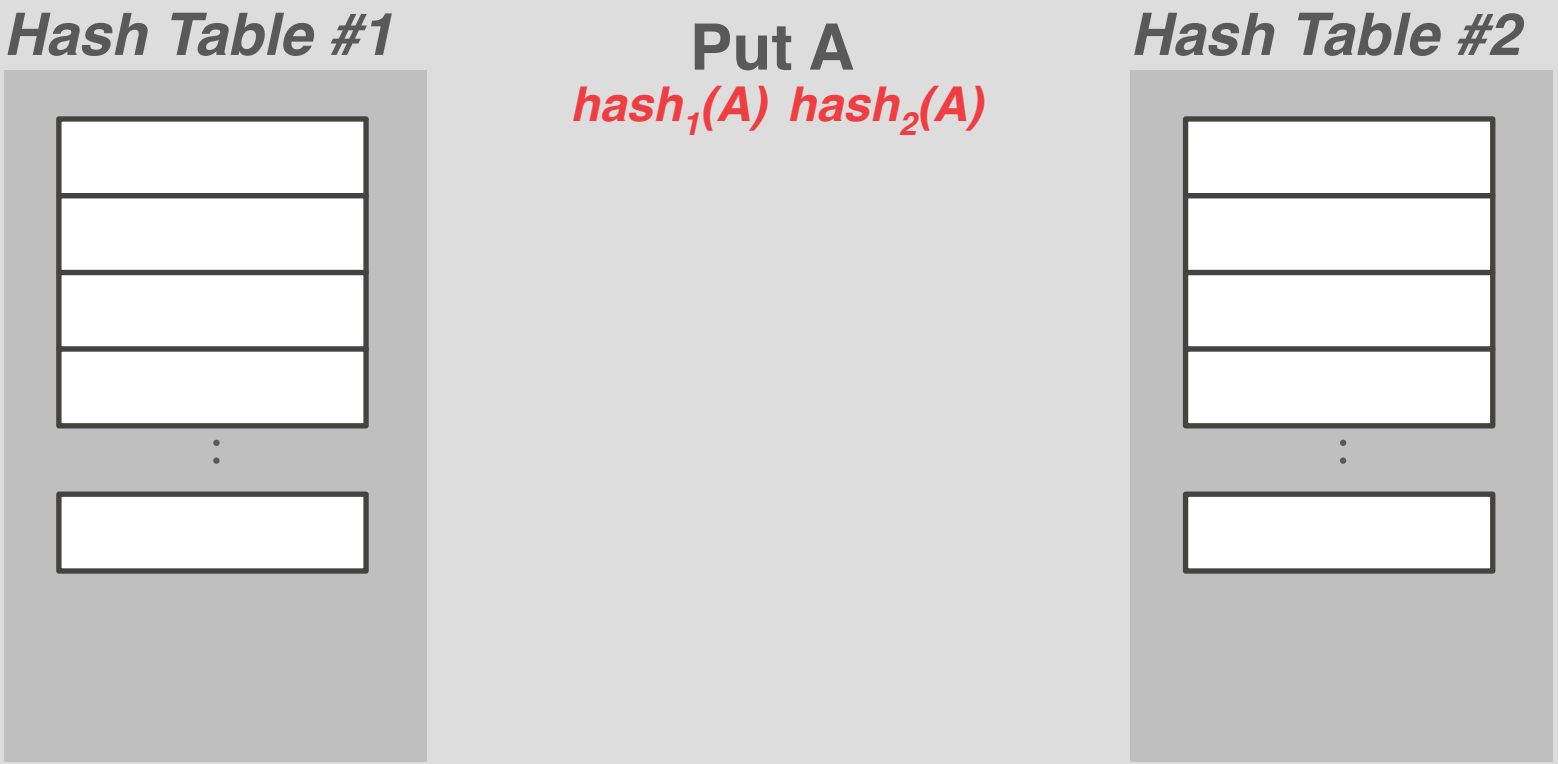
Hash Table #1

Put A

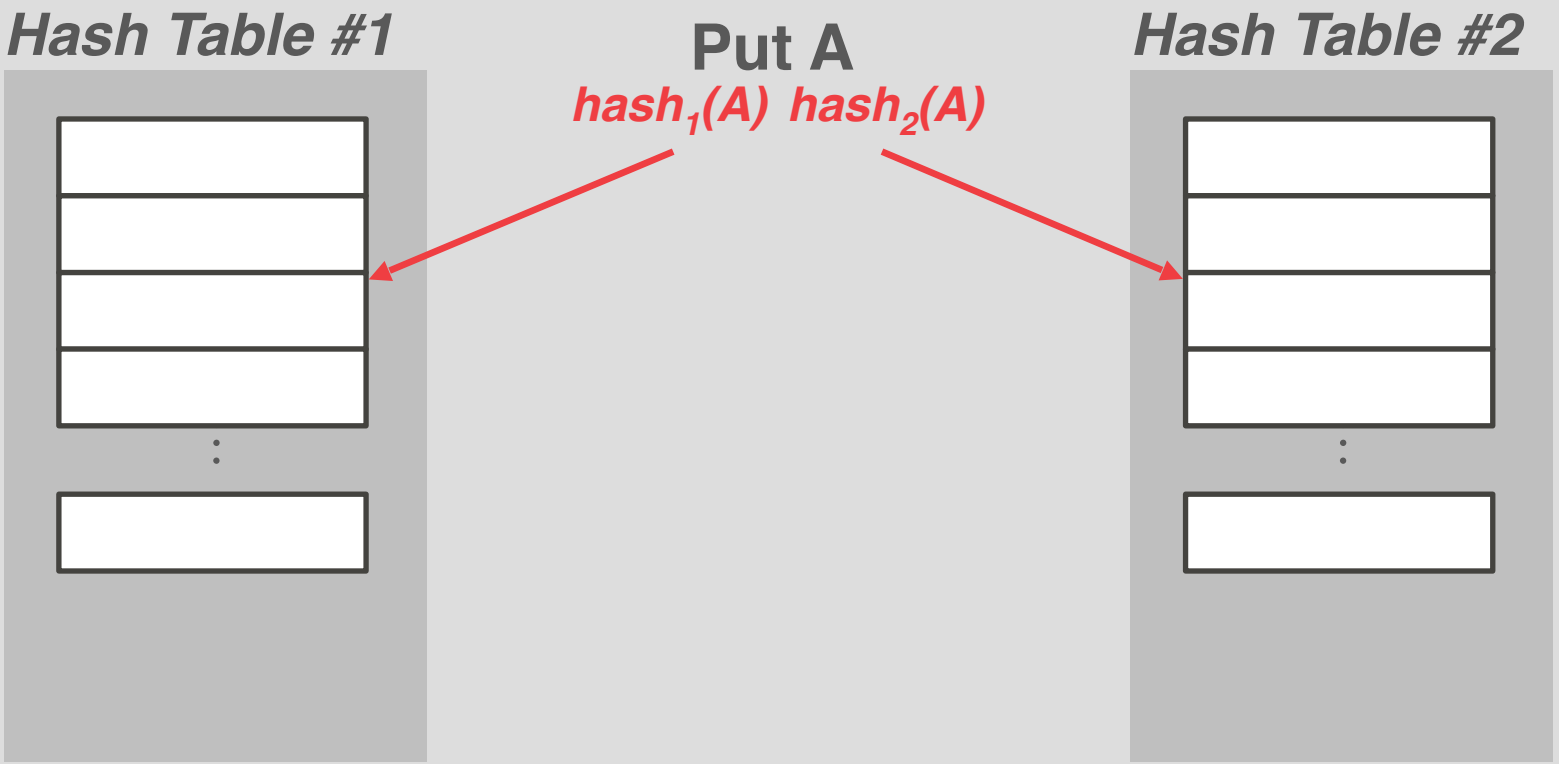
Hash Table #2



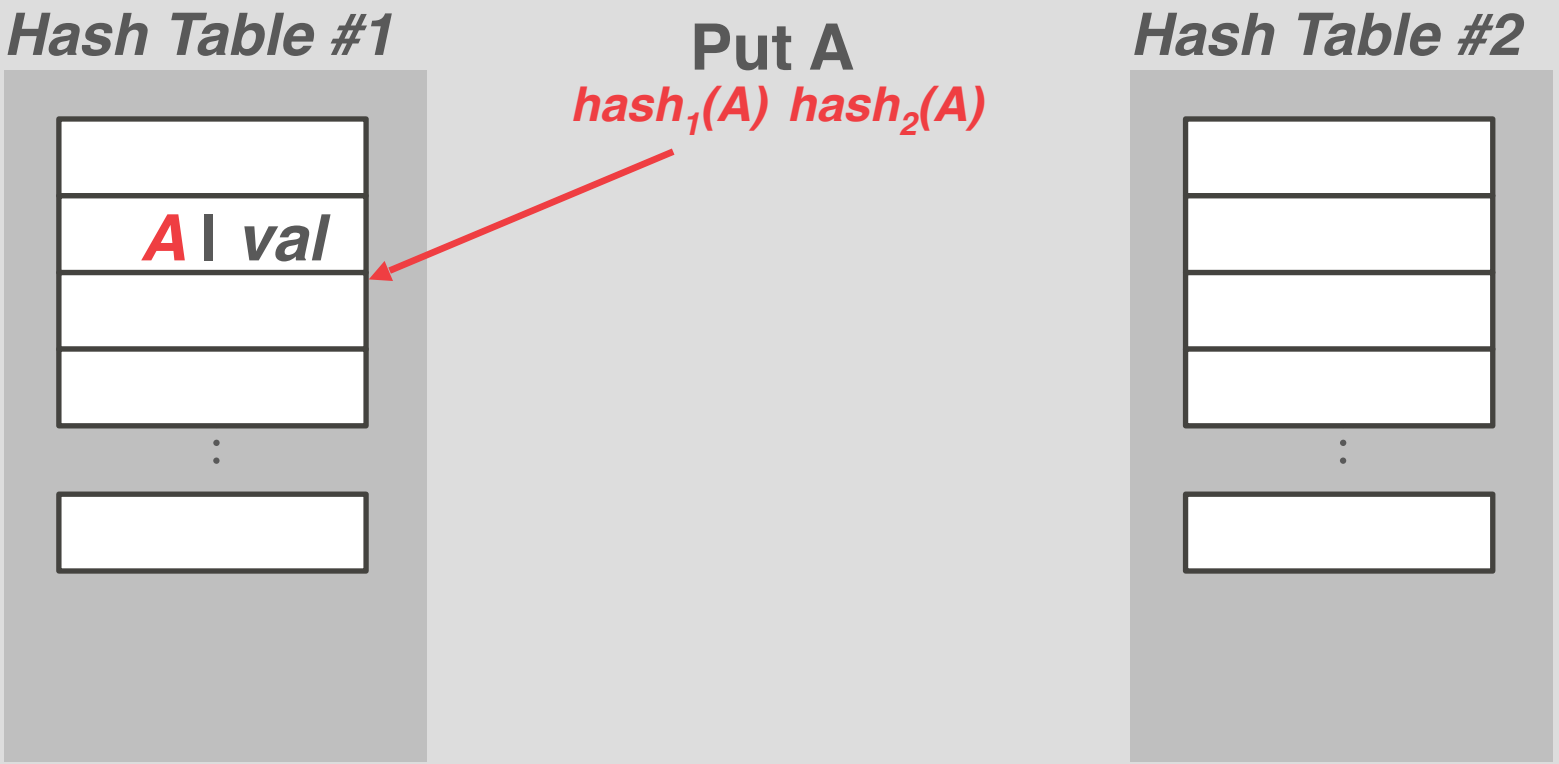
CUCKOO HASHING



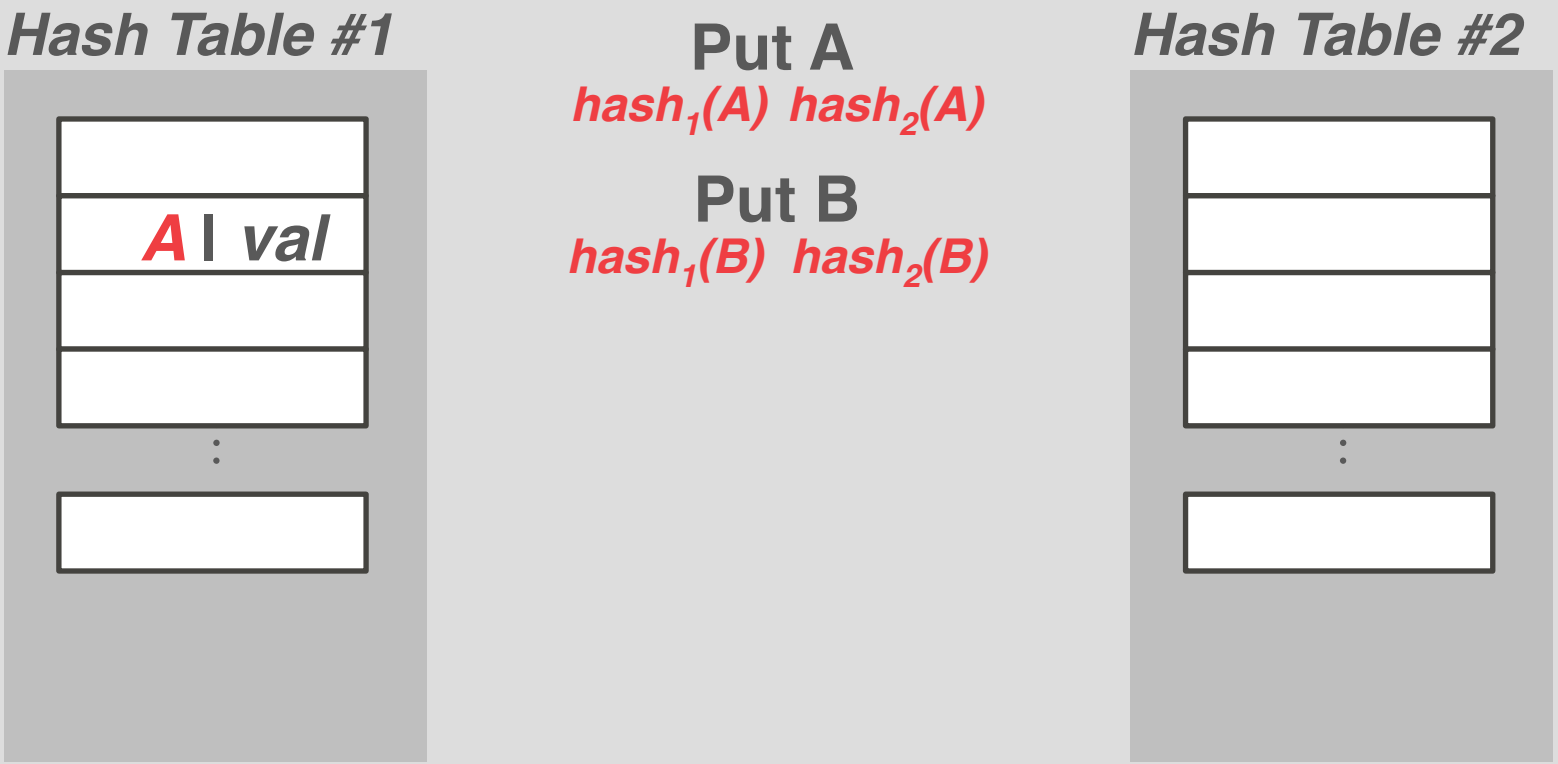
CUCKOO HASHING



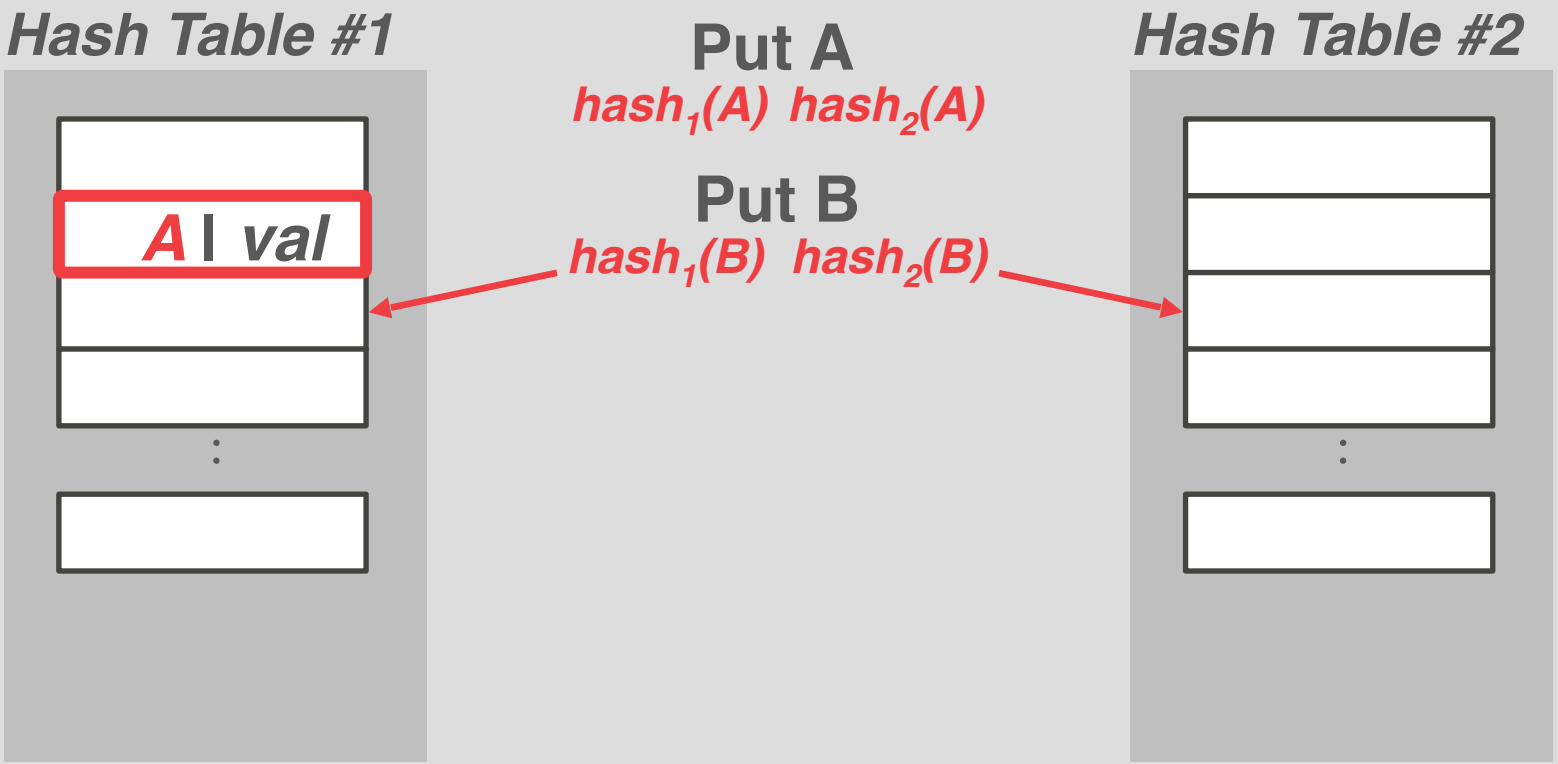
CUCKOO HASHING



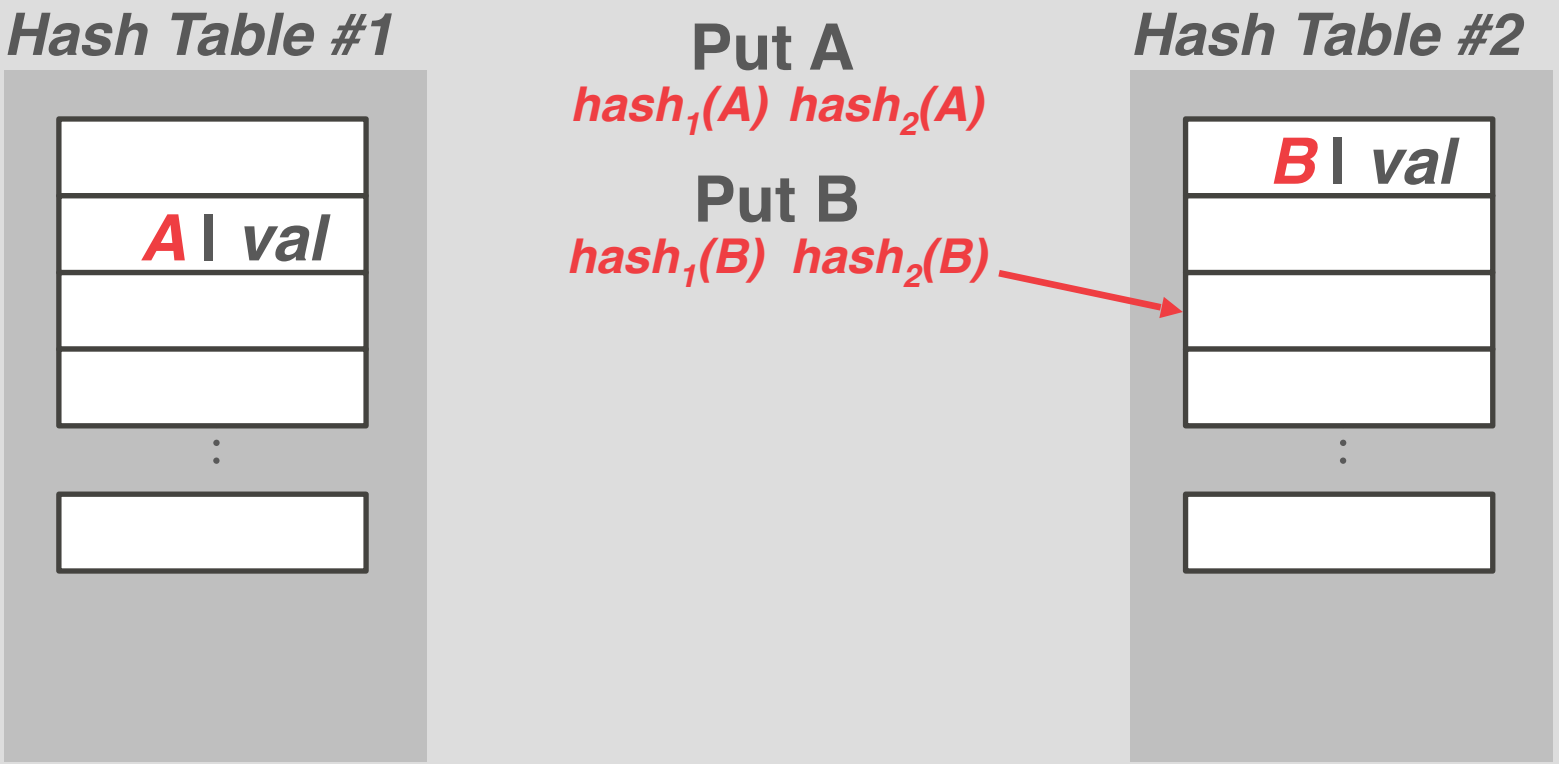
CUCKOO HASHING



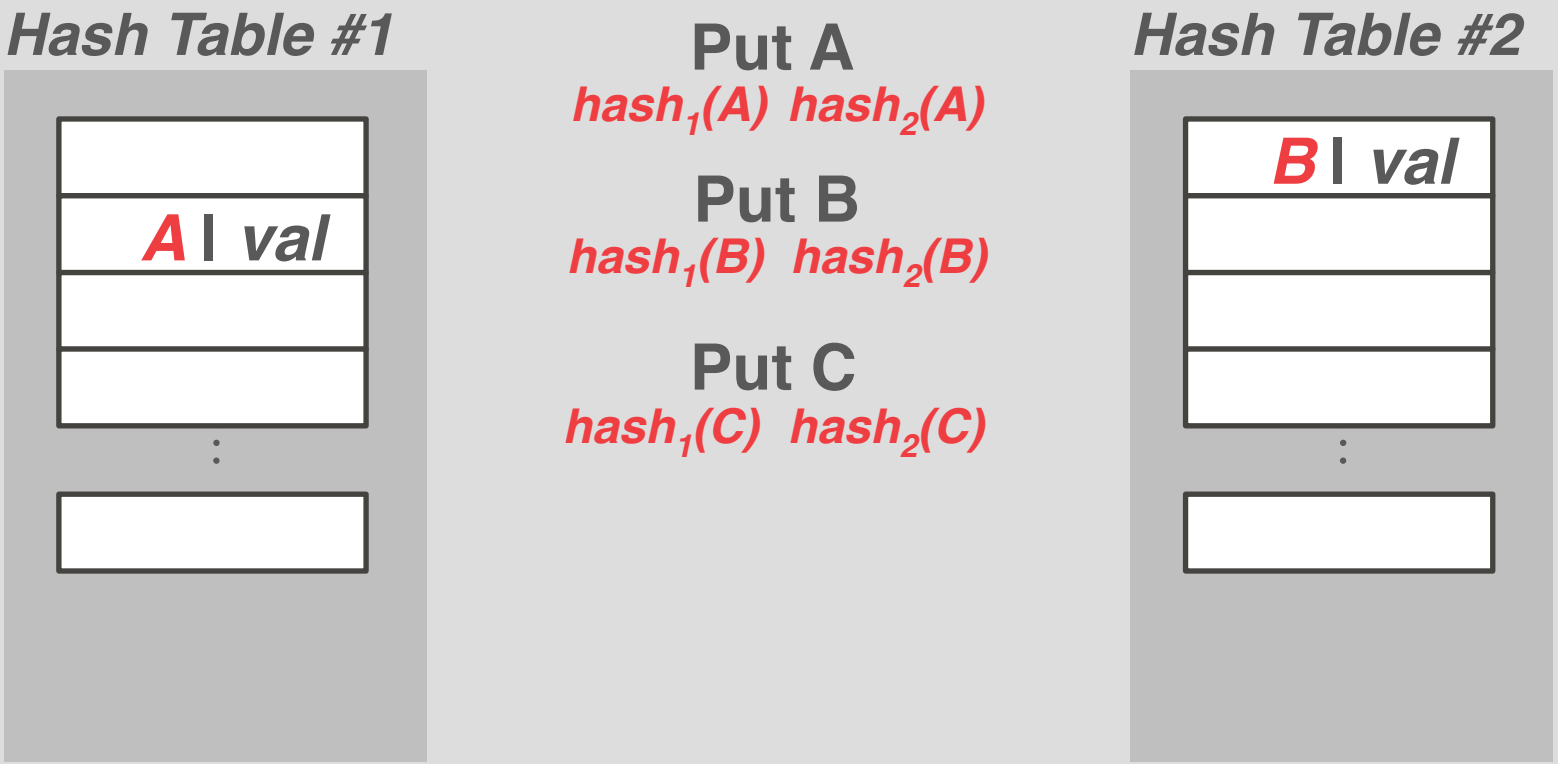
CUCKOO HASHING



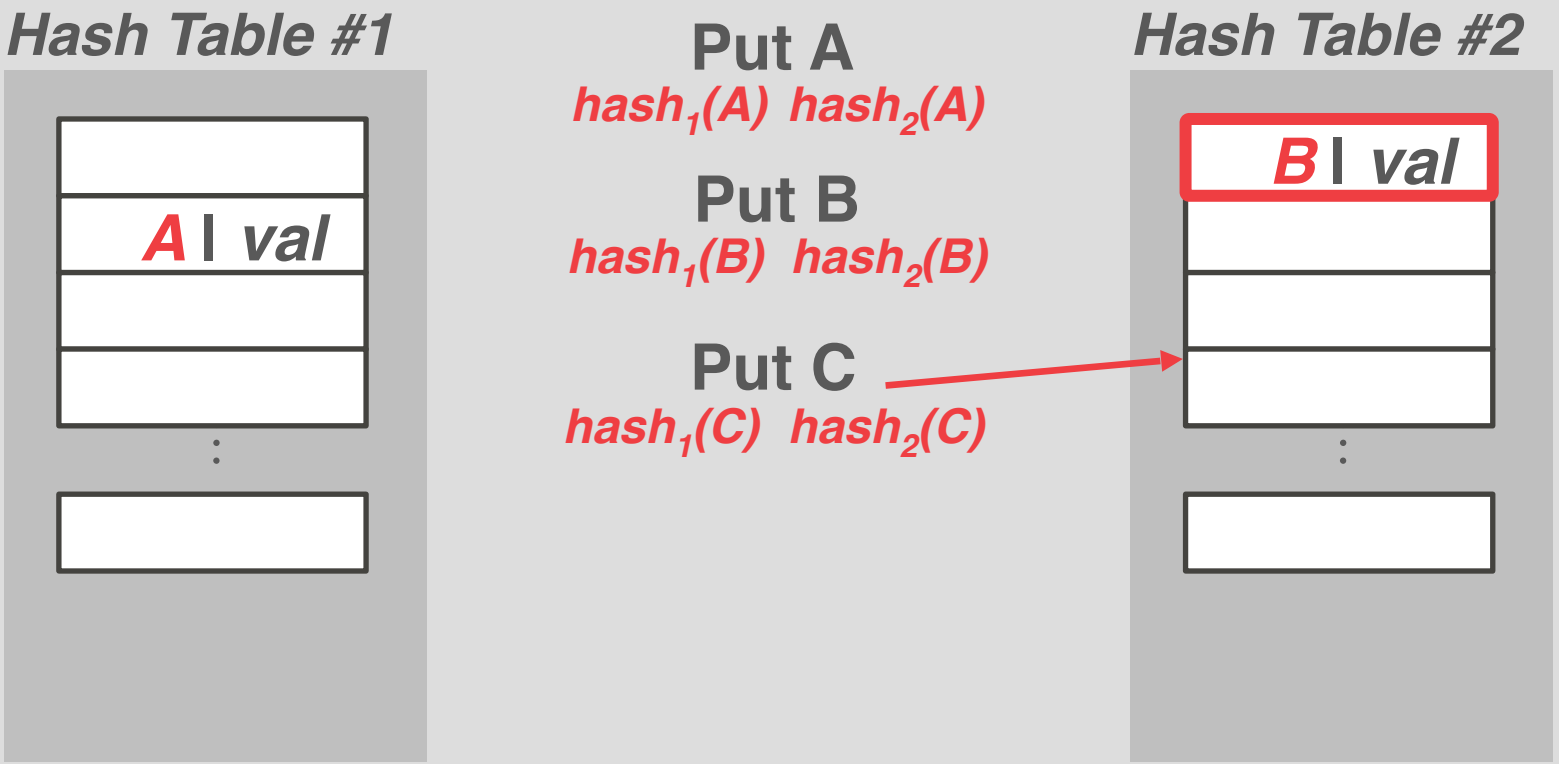
CUCKOO HASHING



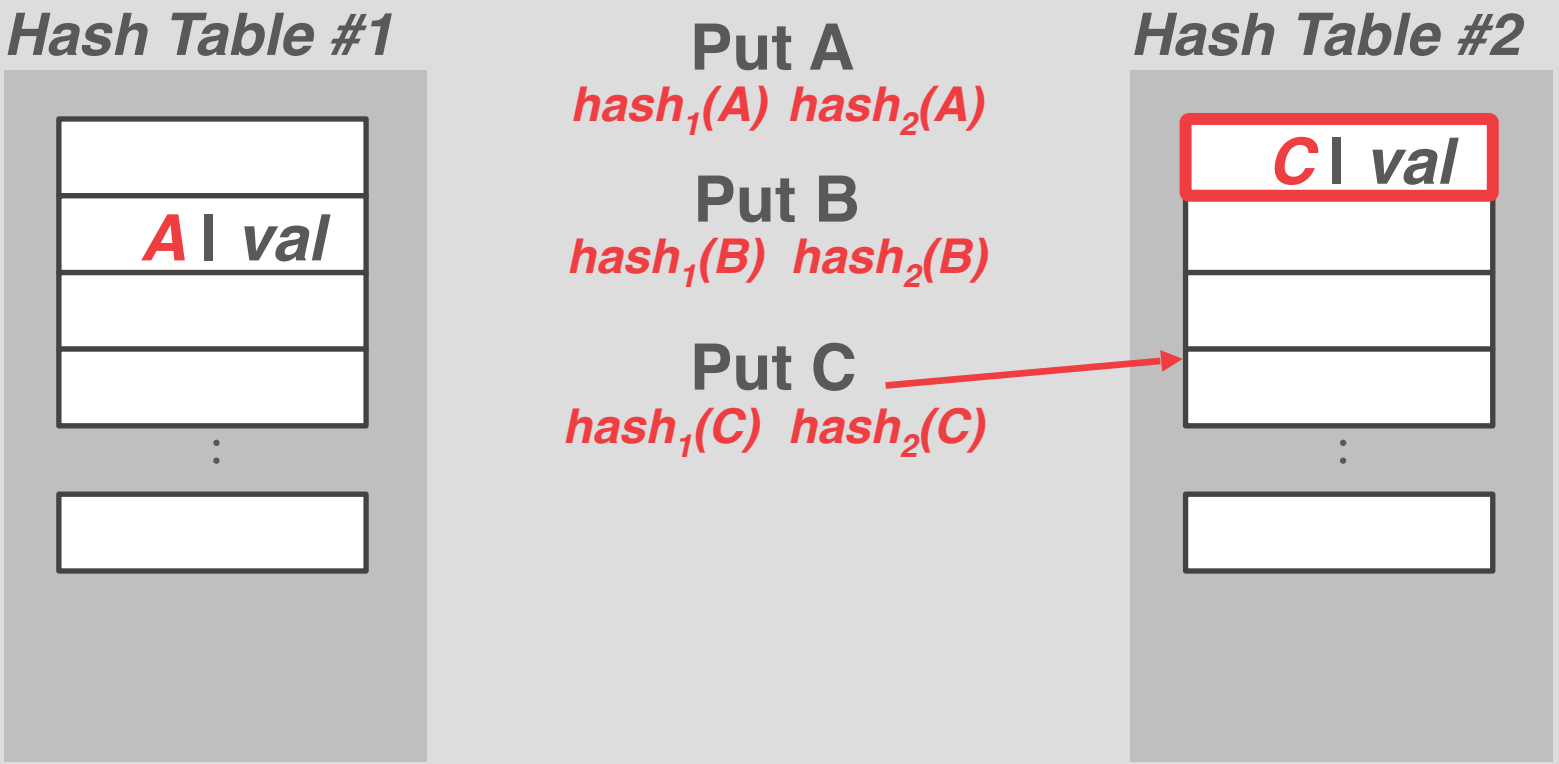
CUCKOO HASHING



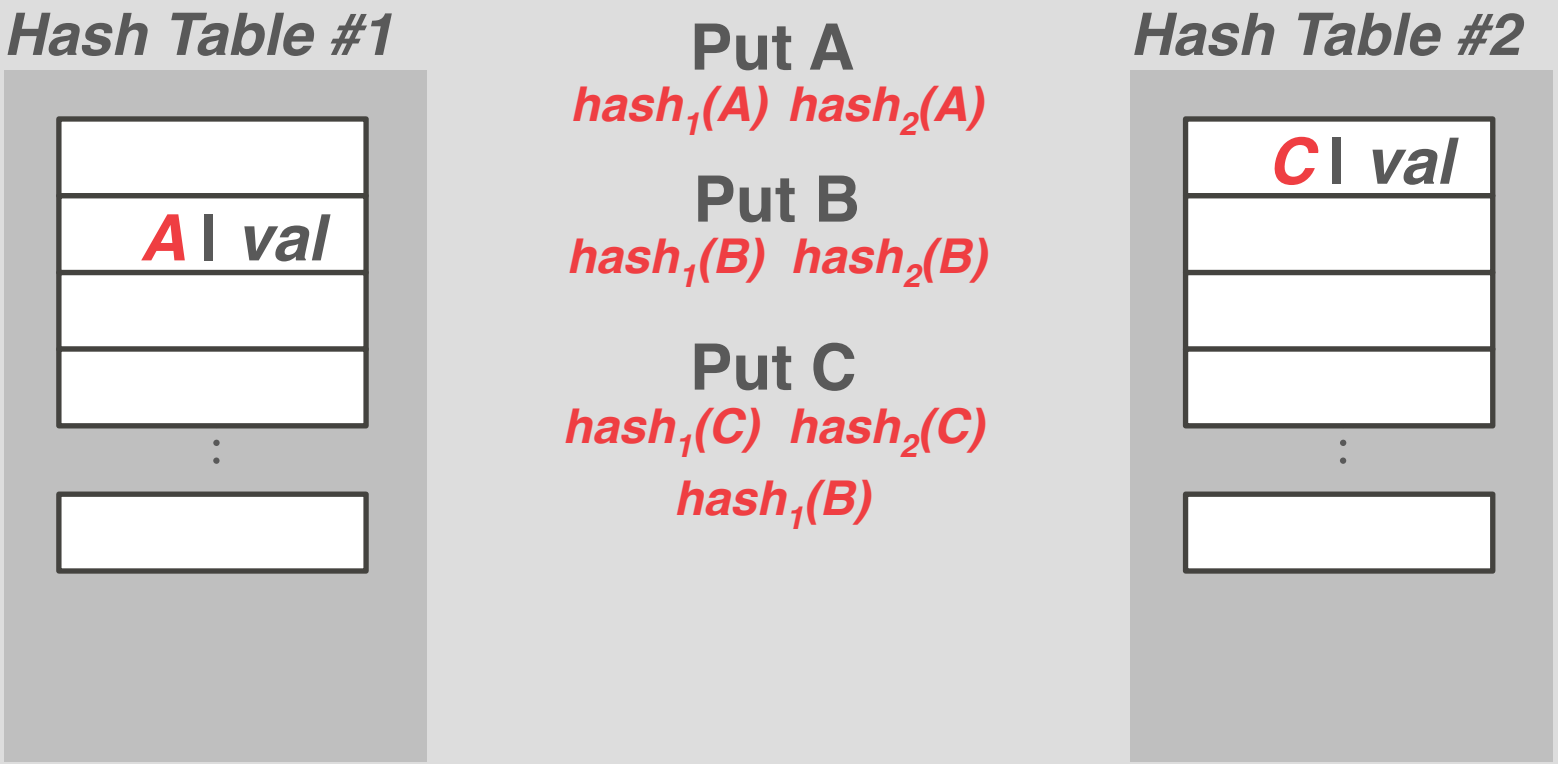
CUCKOO HASHING



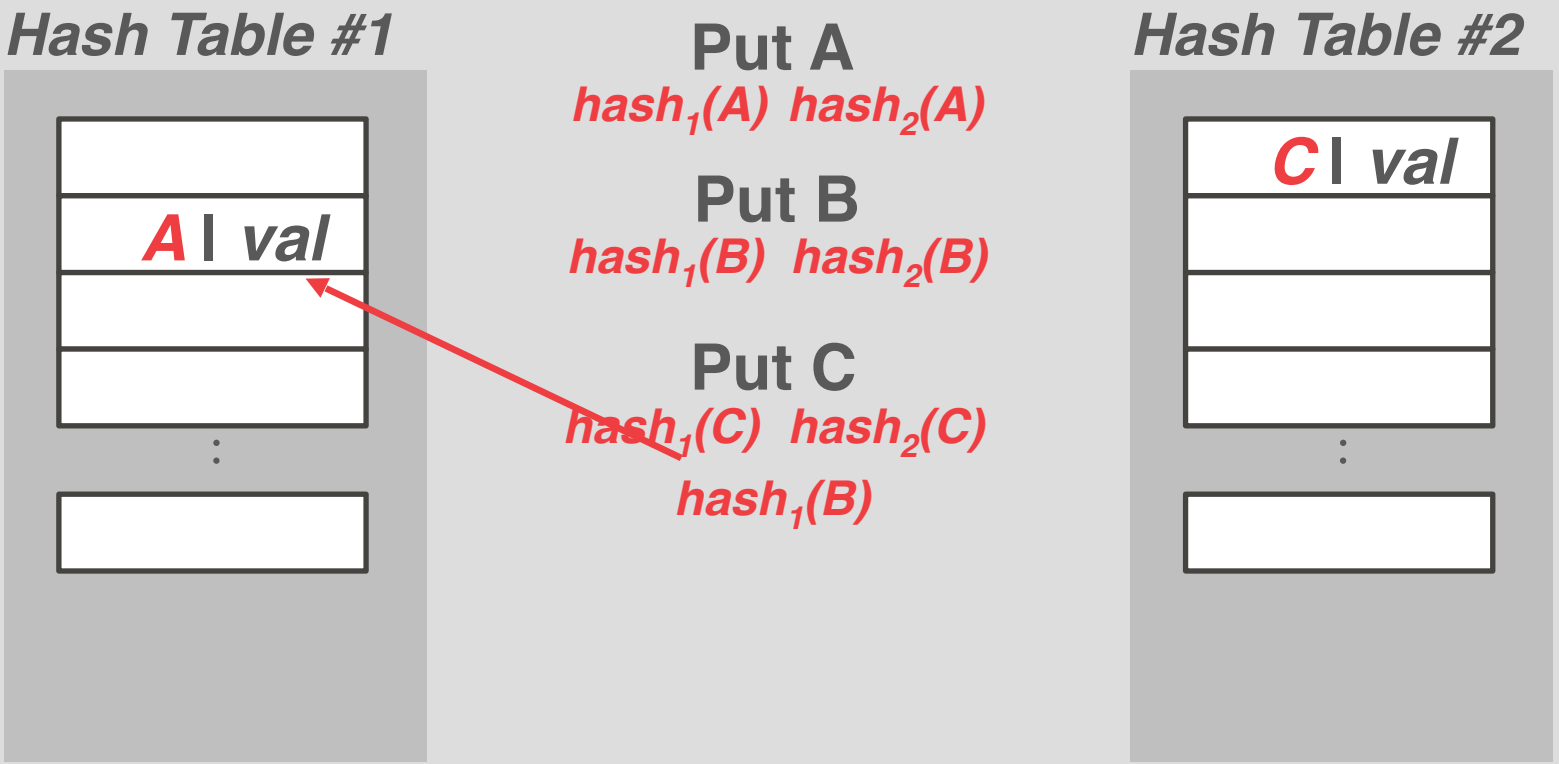
CUCKOO HASHING



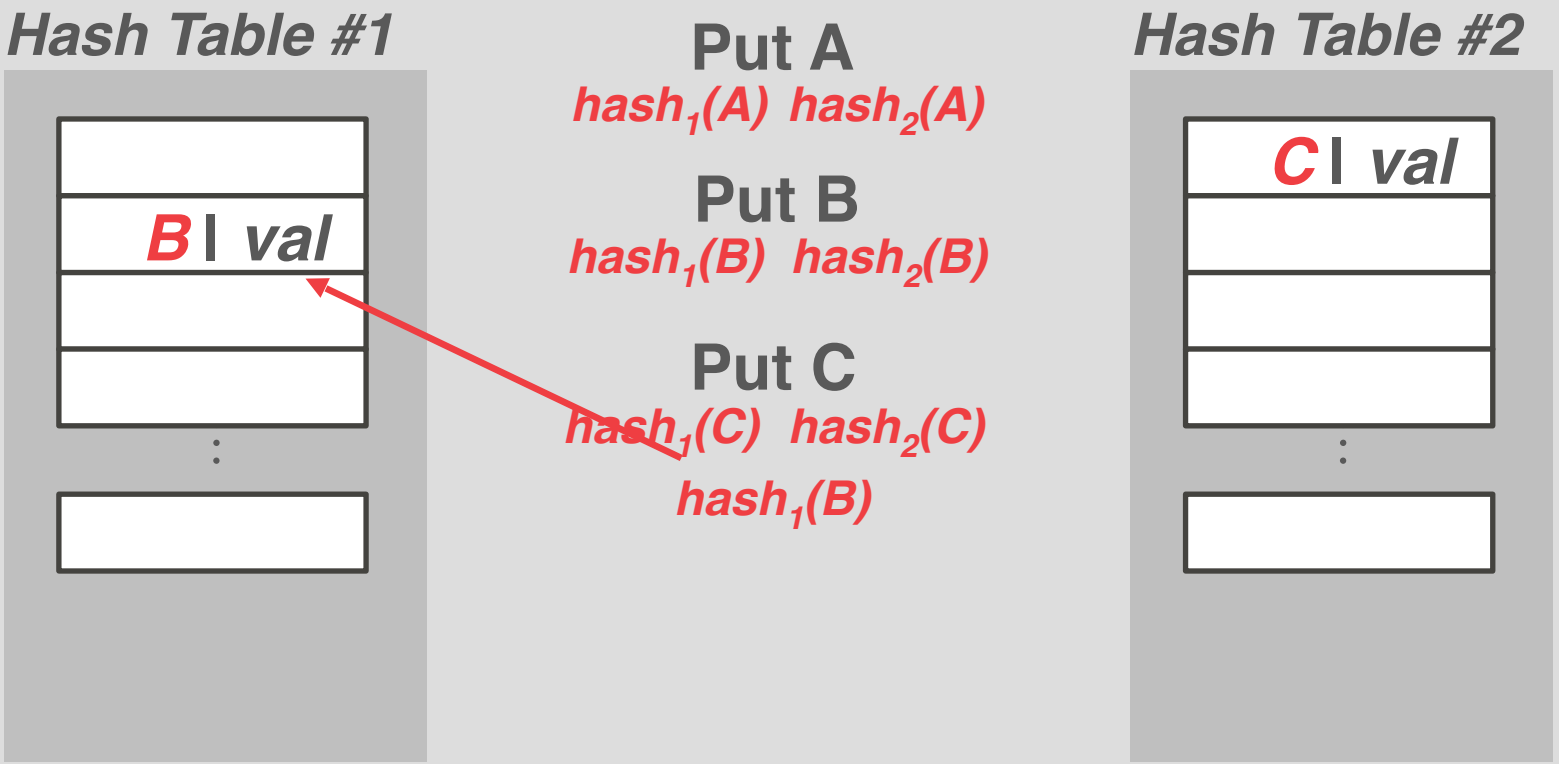
CUCKOO HASHING



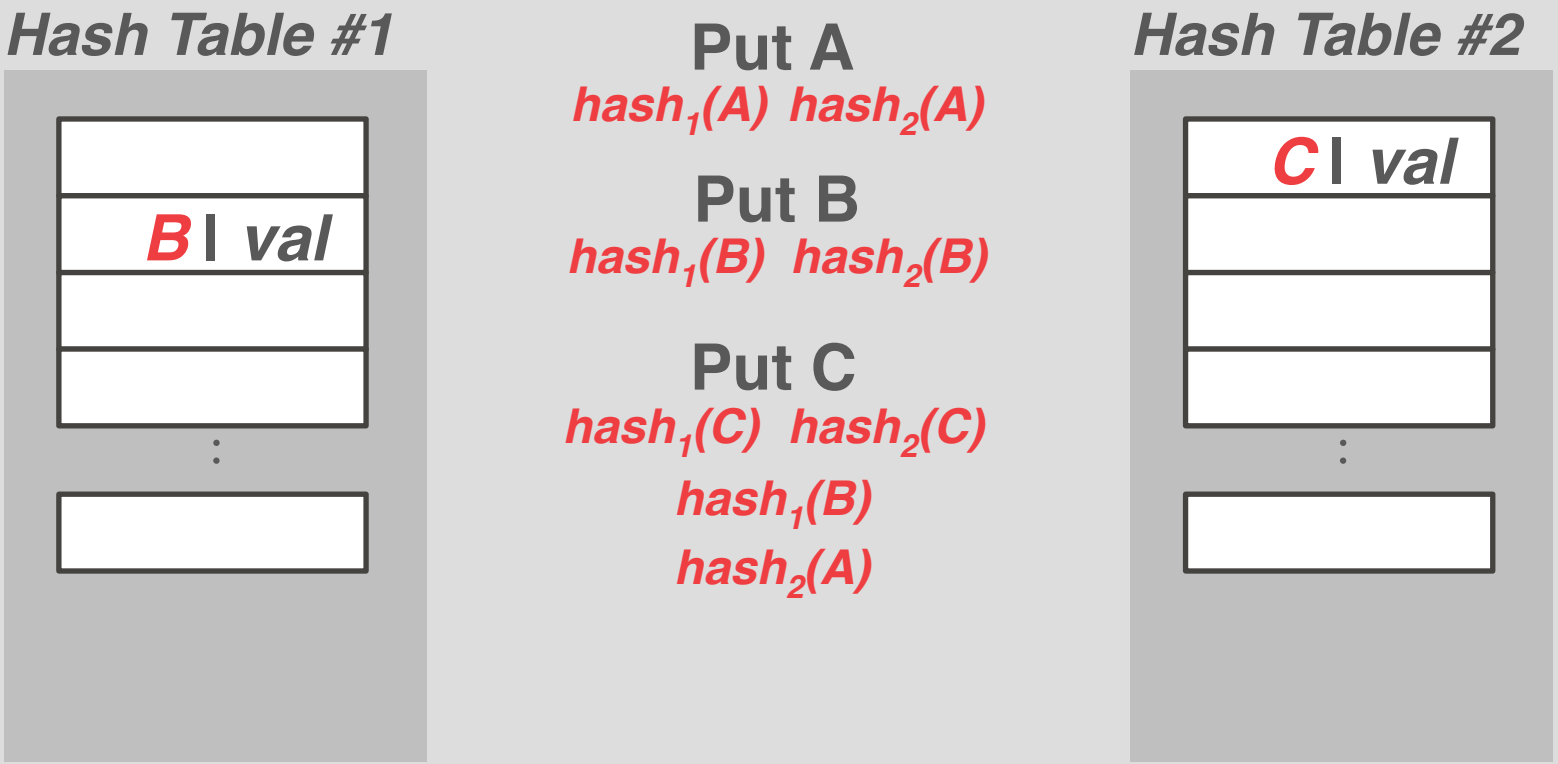
CUCKOO HASHING



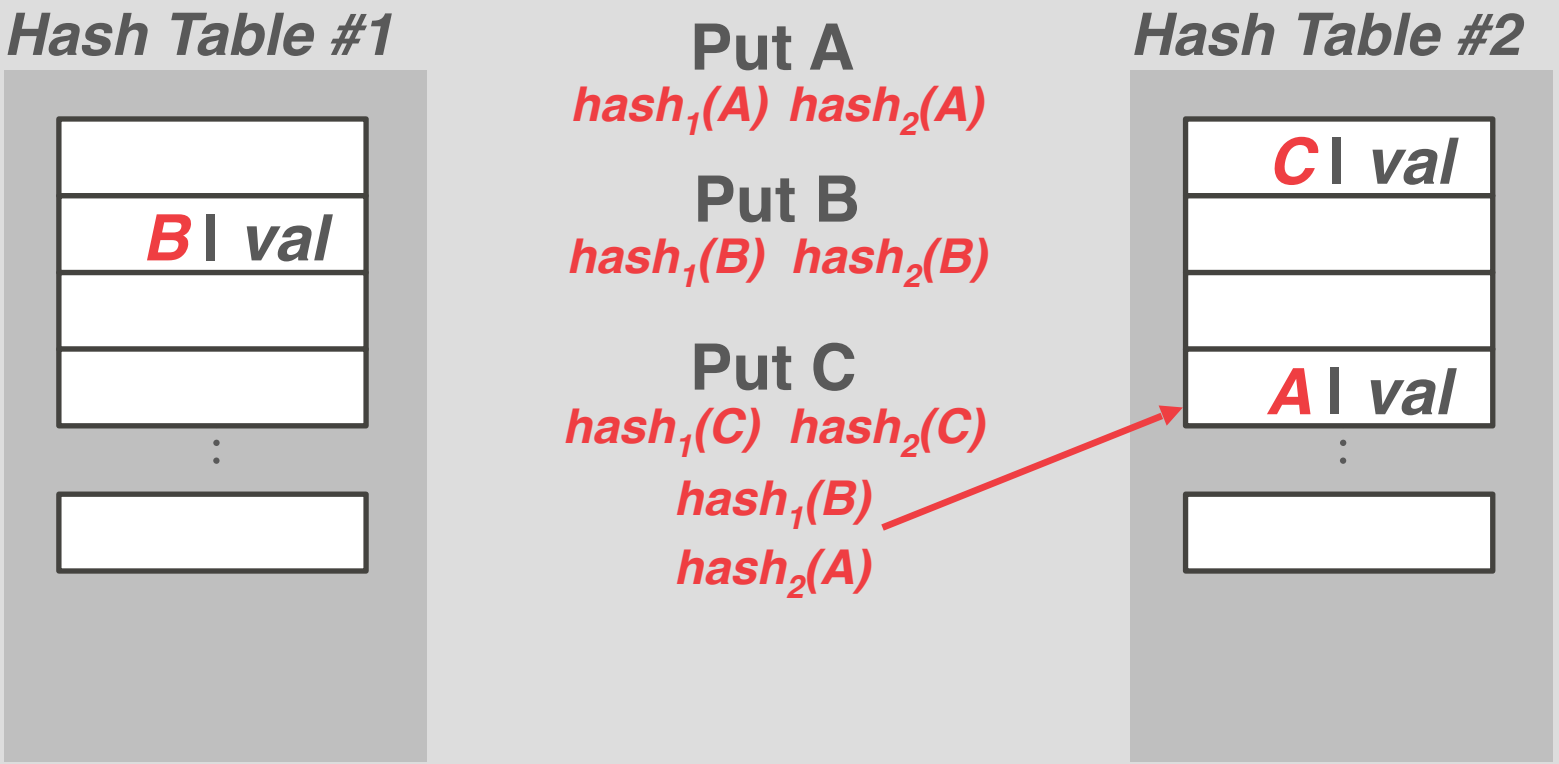
CUCKOO HASHING



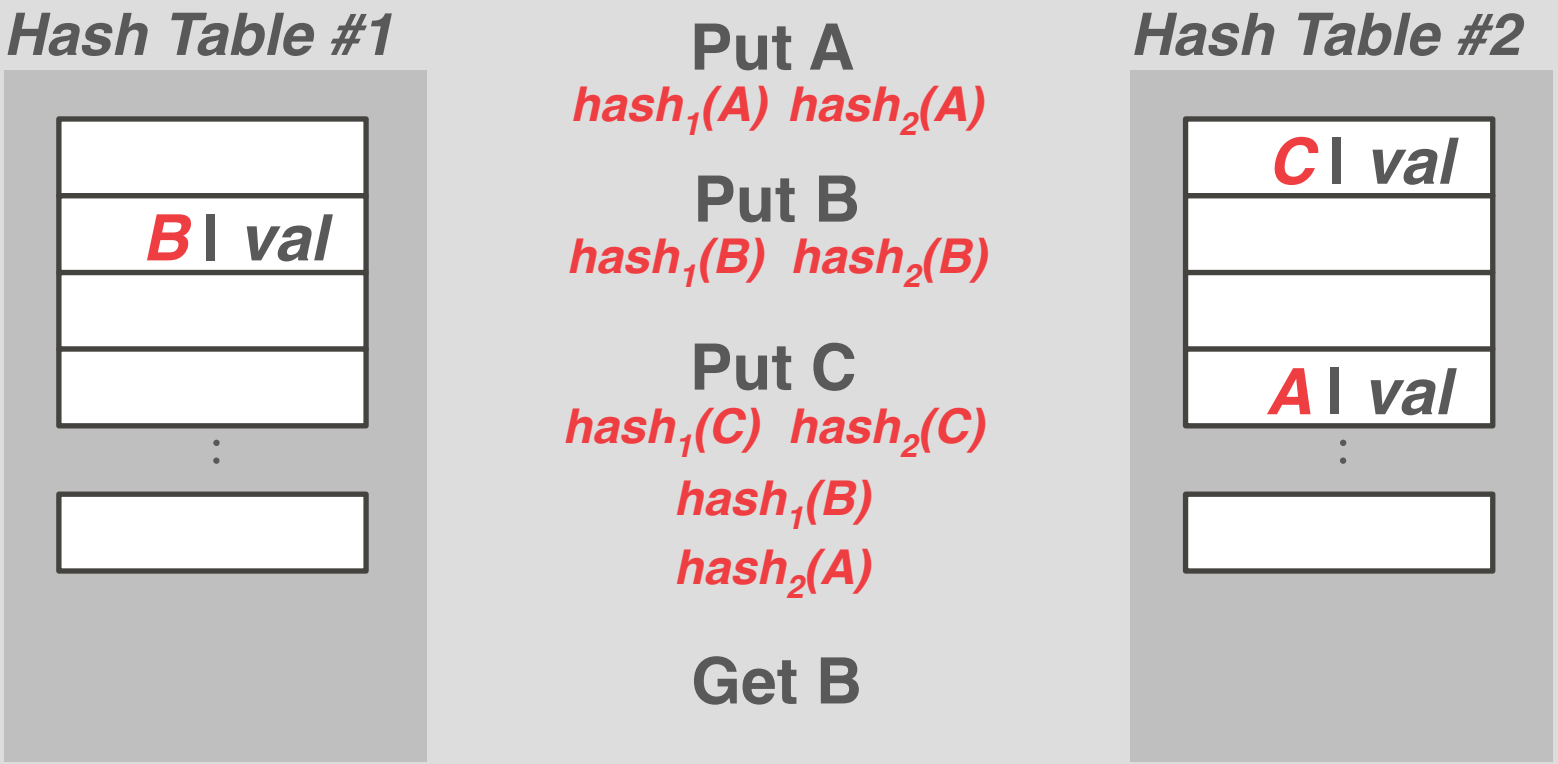
CUCKOO HASHING



CUCKOO HASHING

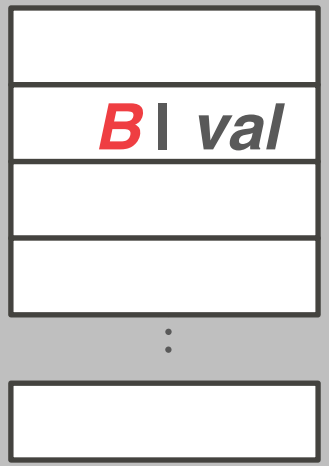


CUCKOO HASHING



CUCKOO HASHING

Hash Table #1



Put A

$hash_1(A)$ $hash_2(A)$

Put B

$hash_1(B)$ $hash_2(B)$

Put C

$hash_1(C)$ $hash_2(C)$

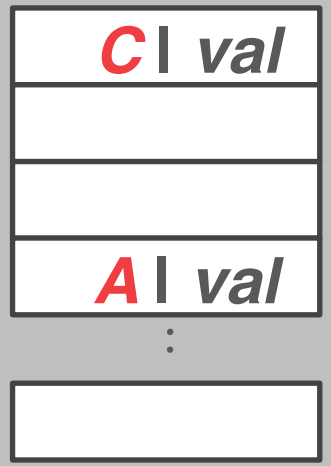
$hash_1(B)$

$hash_2(A)$

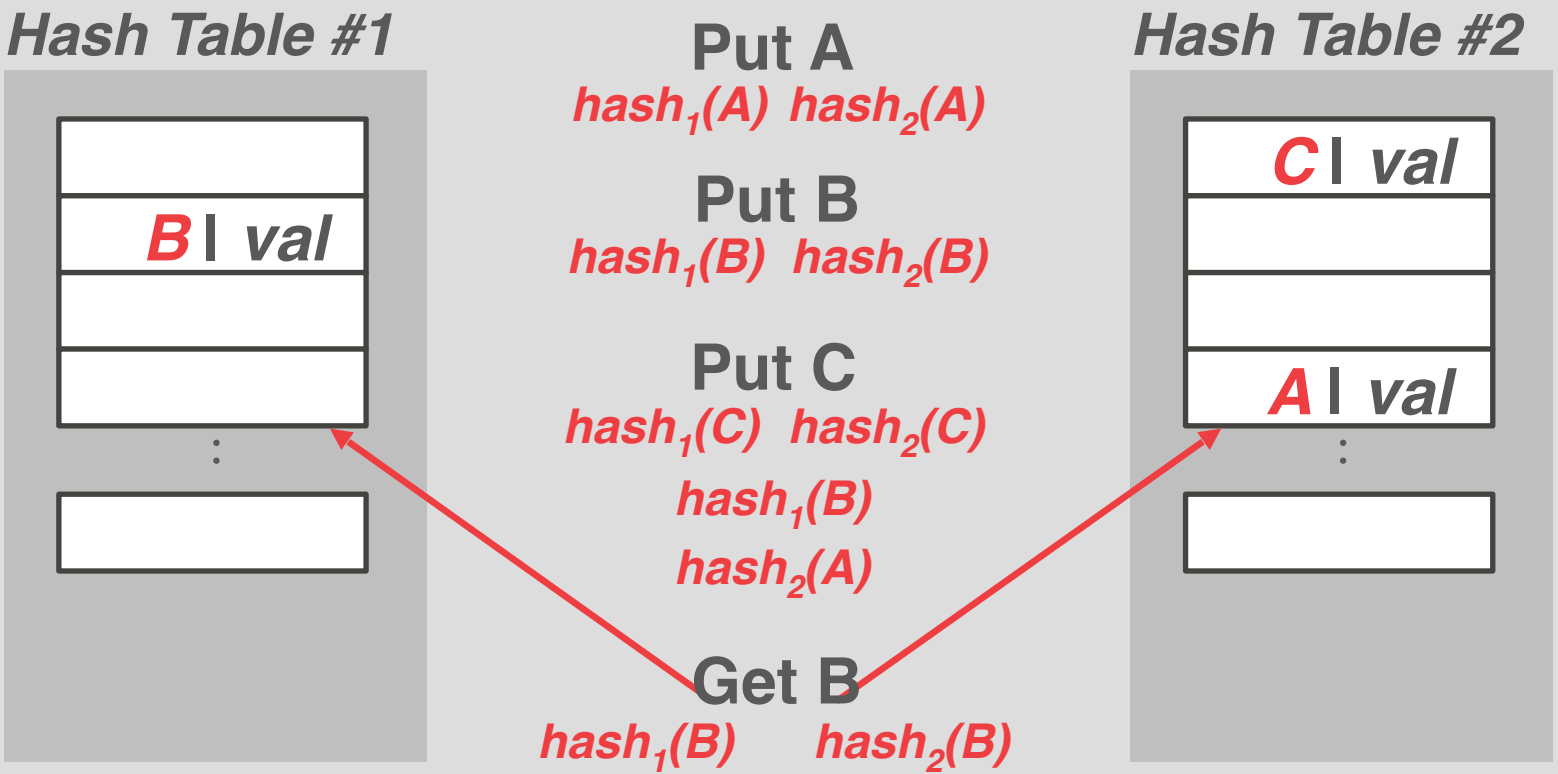
Get B

$hash_1(B)$ $hash_2(B)$

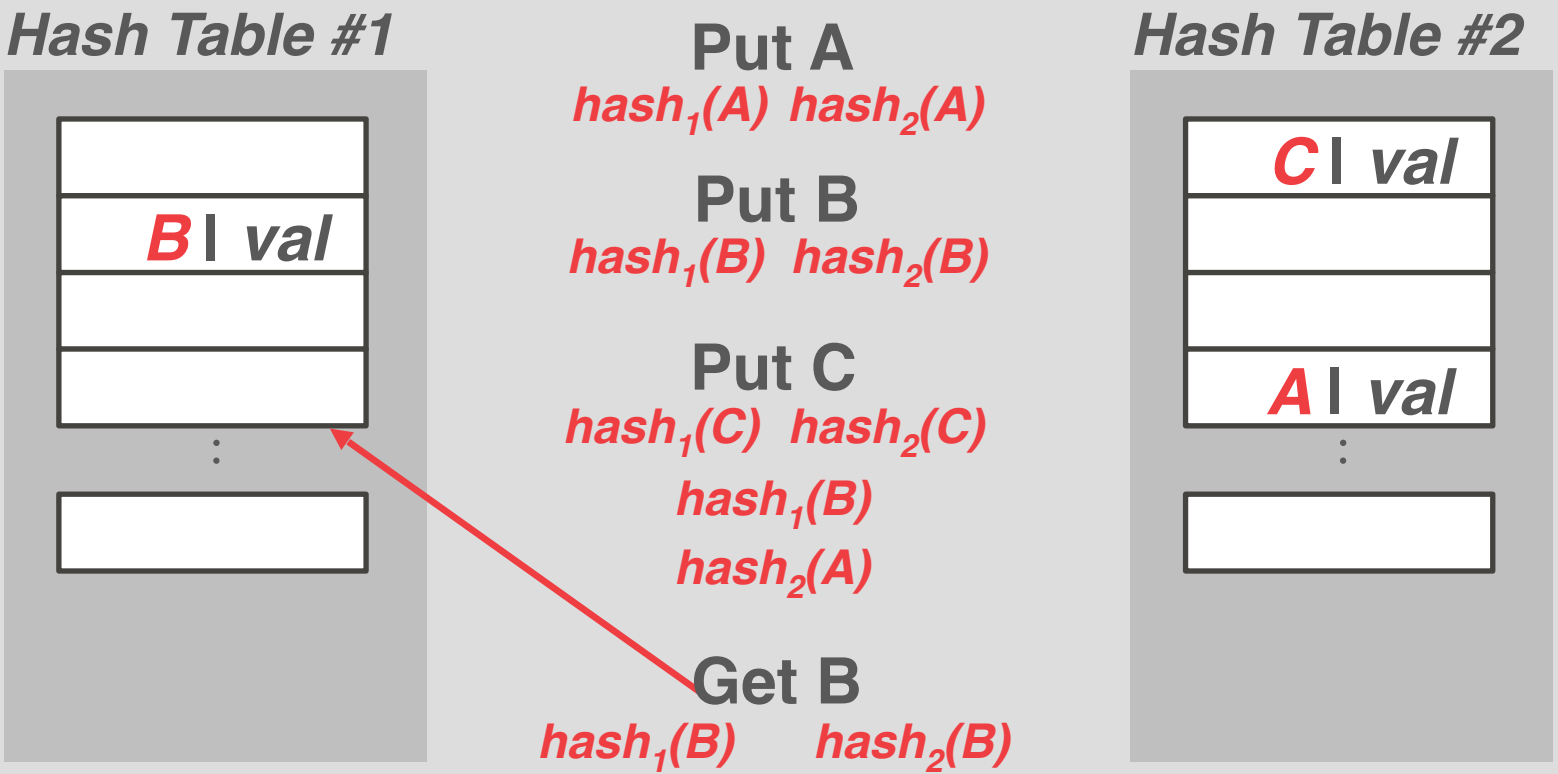
Hash Table #2



CUCKOO HASHING



CUCKOO HASHING



OBSERVATION

The previous hash tables require the DBMS to know the number of elements it wants to store.

→ Otherwise, it must rebuild the table if it needs to grow/shrink in size.

Dynamic hash tables resize themselves on demand.

→ Chained Hashing
→ Extendible Hashing
→ Linear Hashing

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.

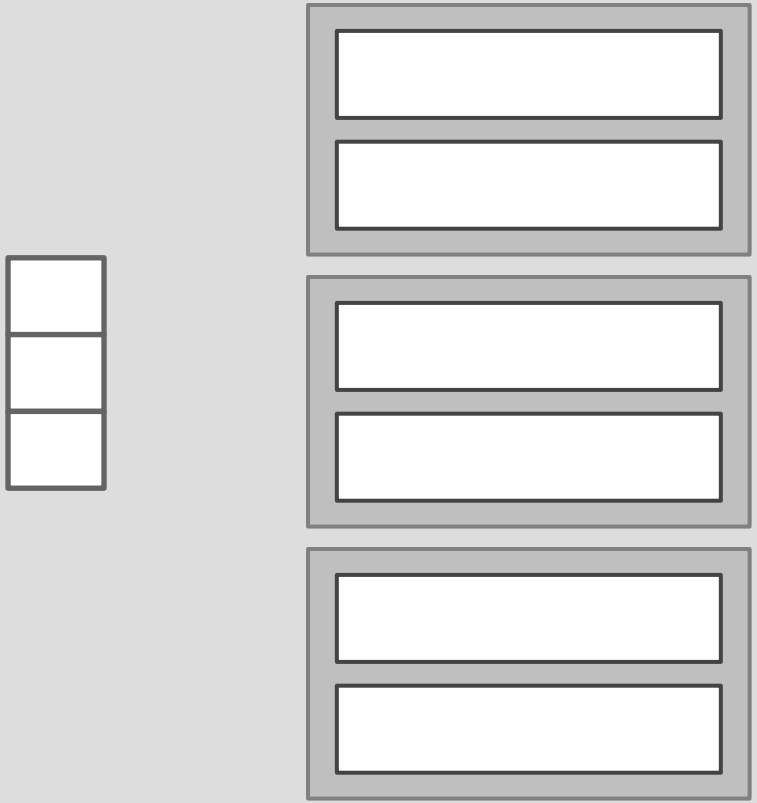
→ To determine whether an element is present, hash to its bucket and scan for it.

→ Insertions and deletions are generalizations of lookups.

CHAINED HASHING

hash(key) % N

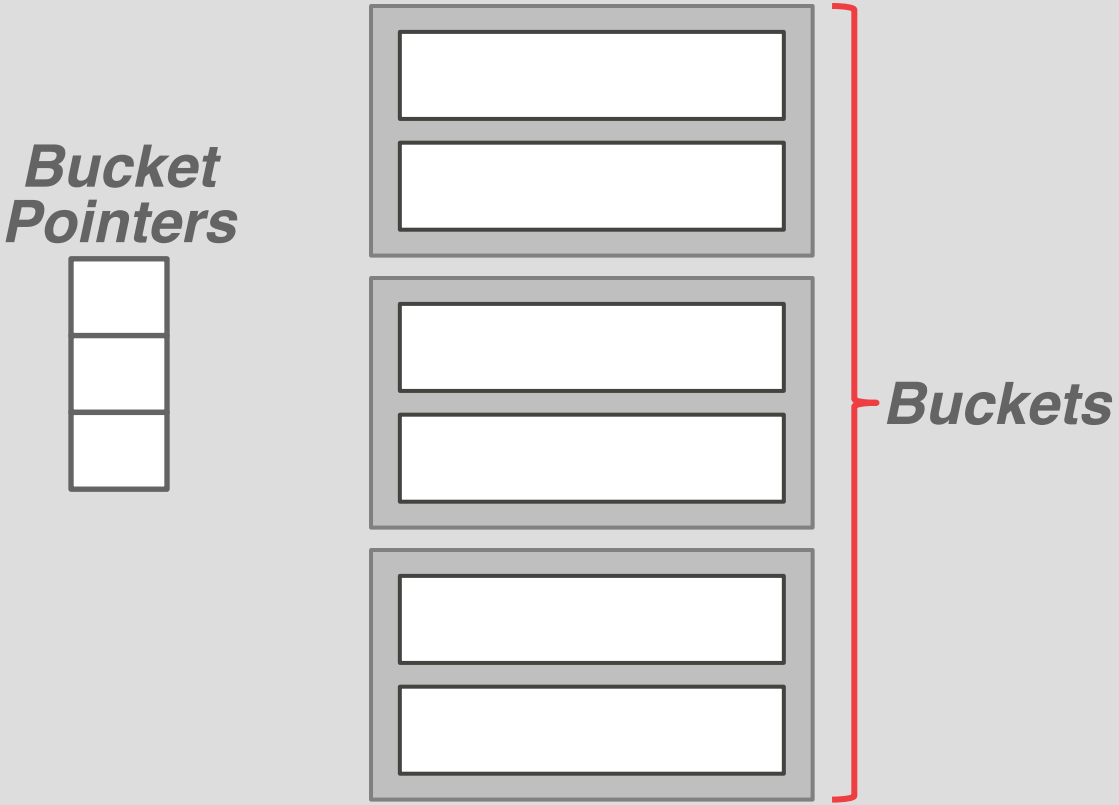
A
B
C
D
E
F



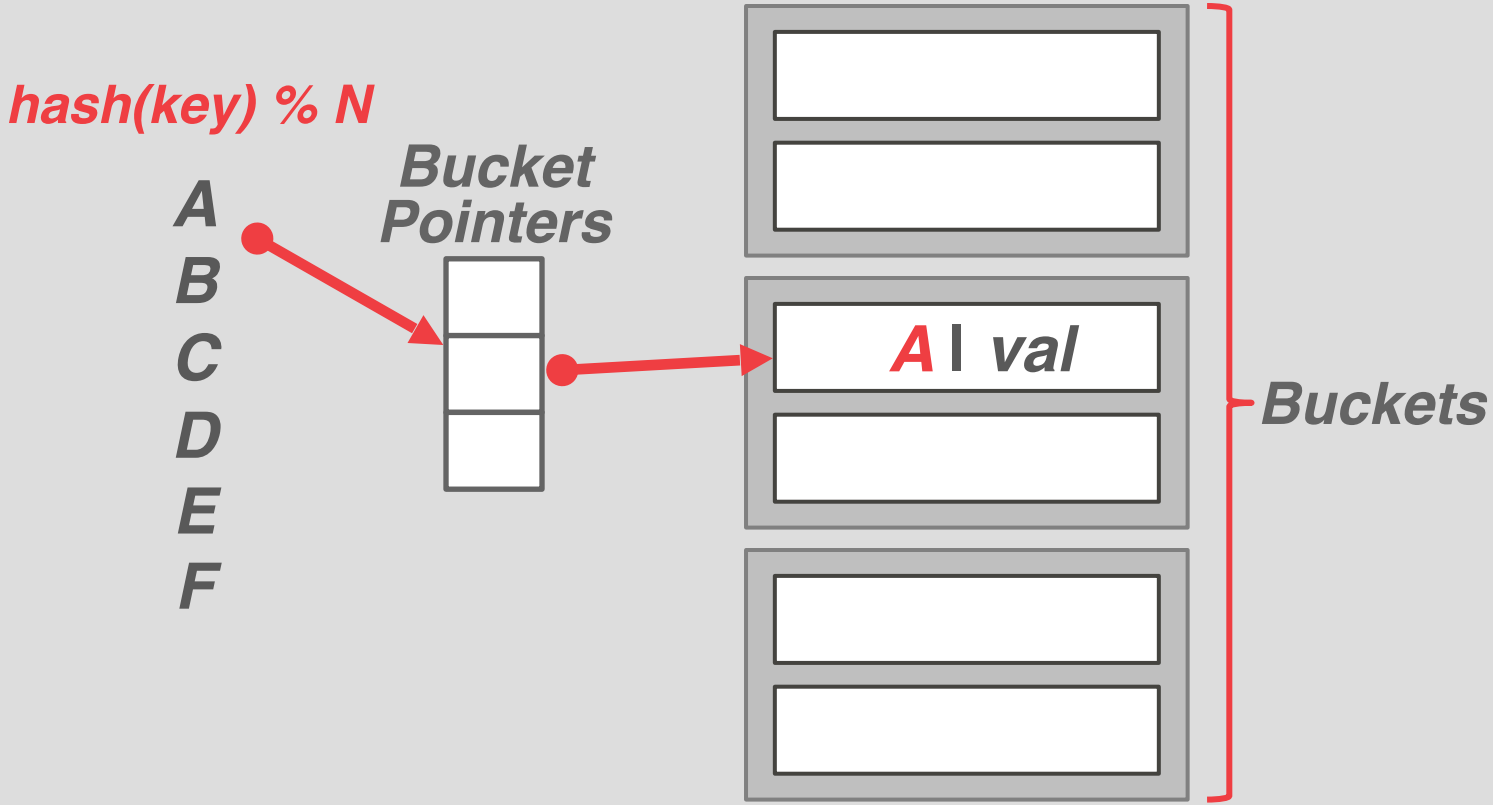
CHAINED HASHING

hash(key) % N

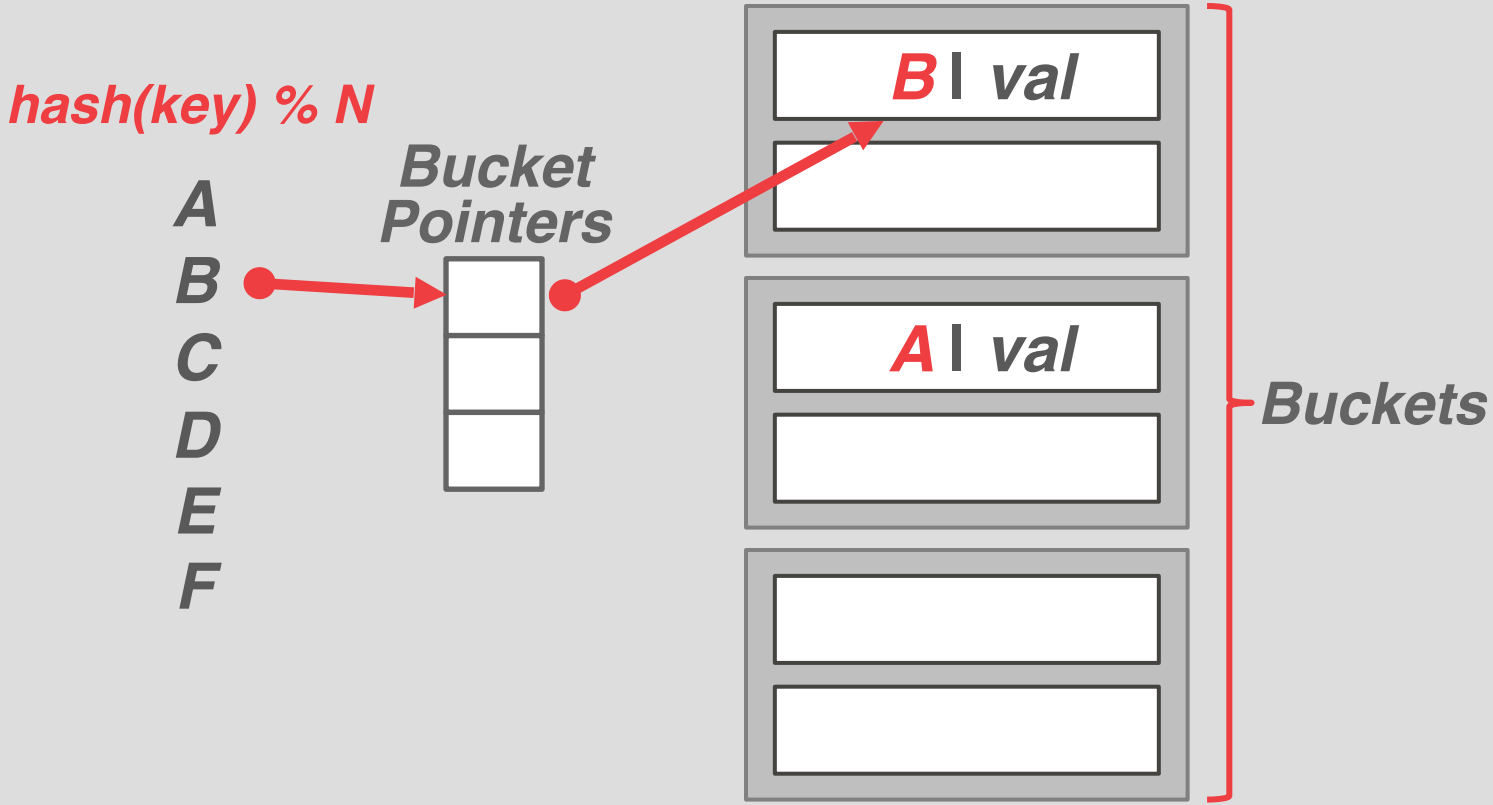
A
B
C
D
E
F



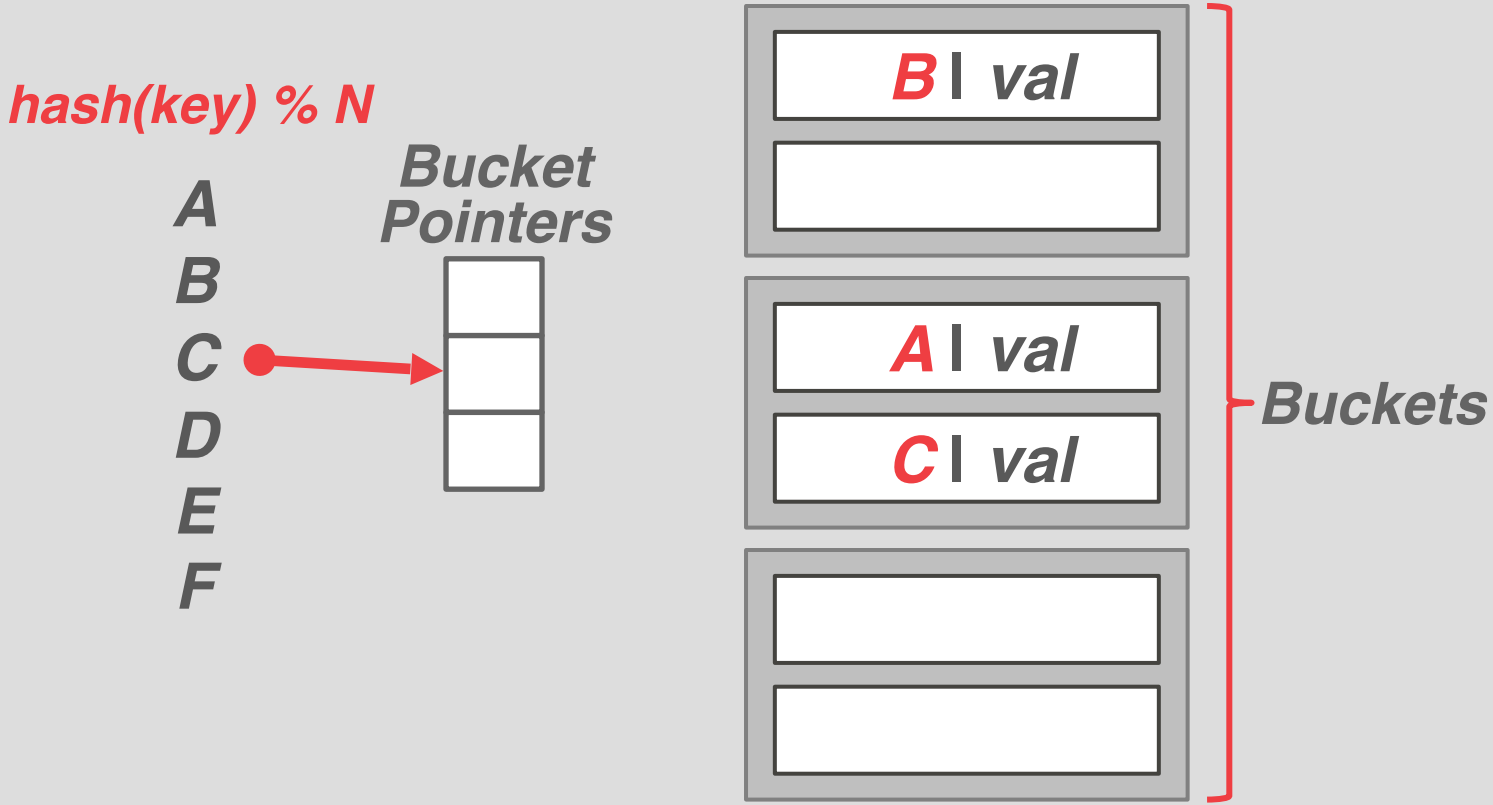
CHAINED HASHING



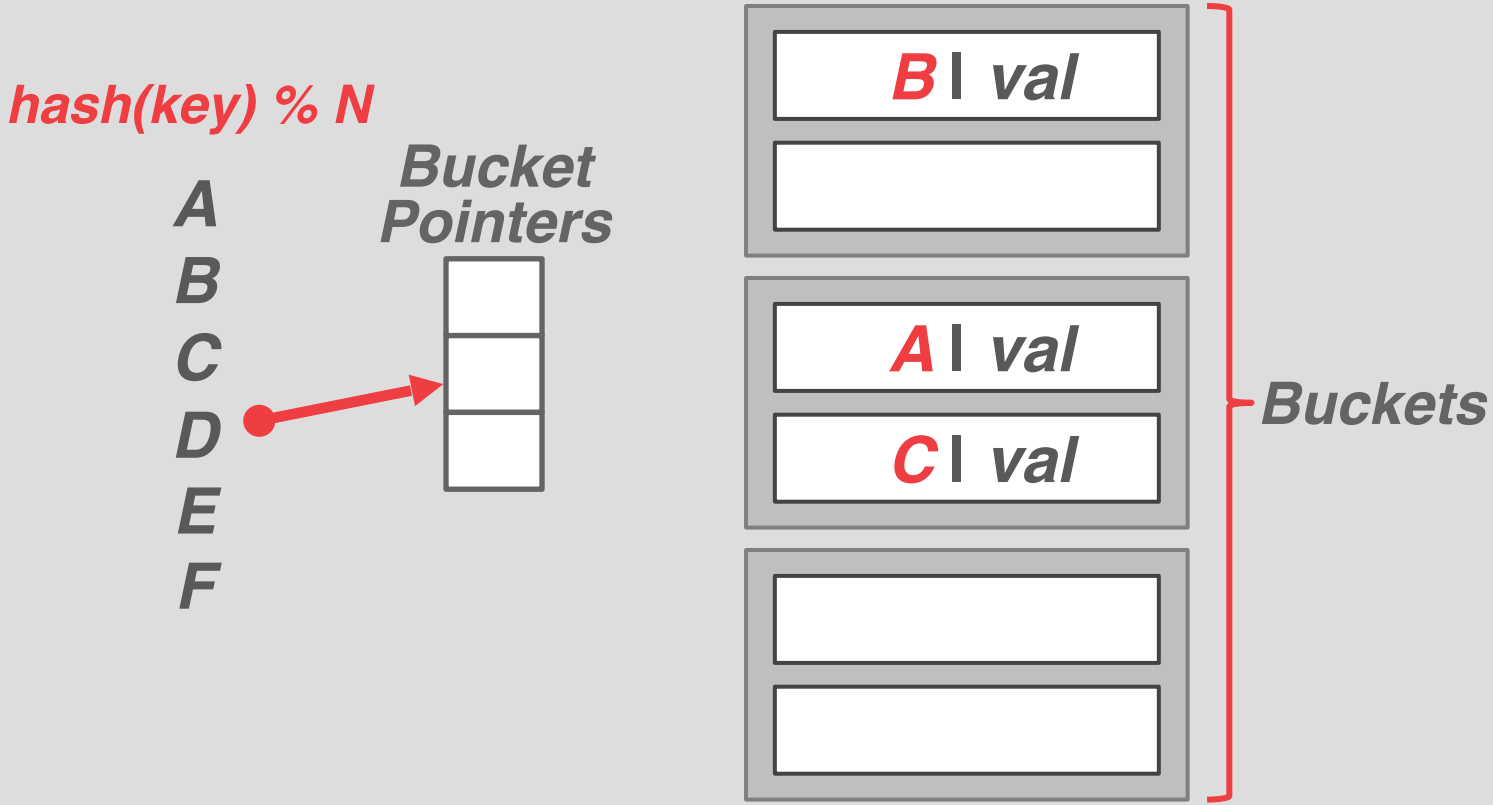
CHAINED HASHING



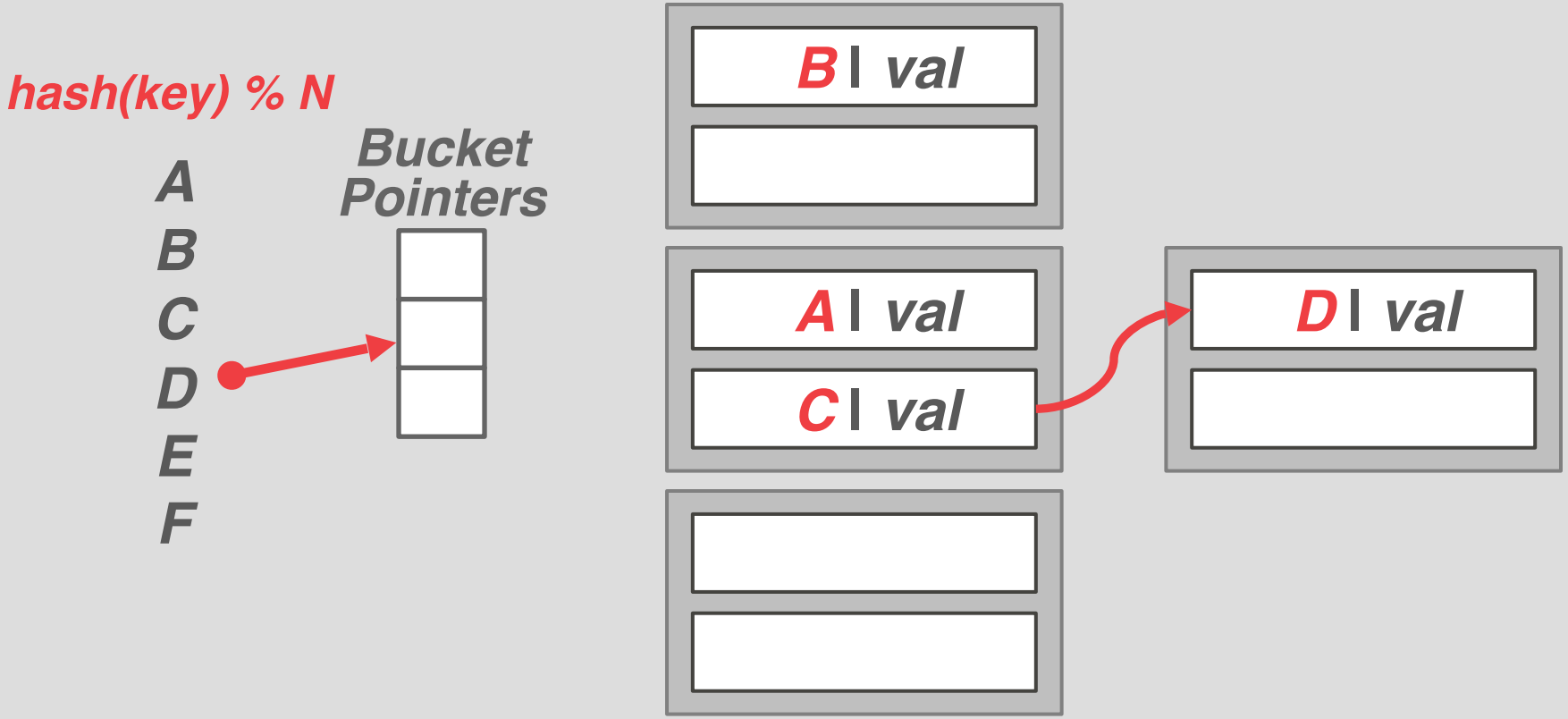
CHAINED HASHING



CHAINED HASHING



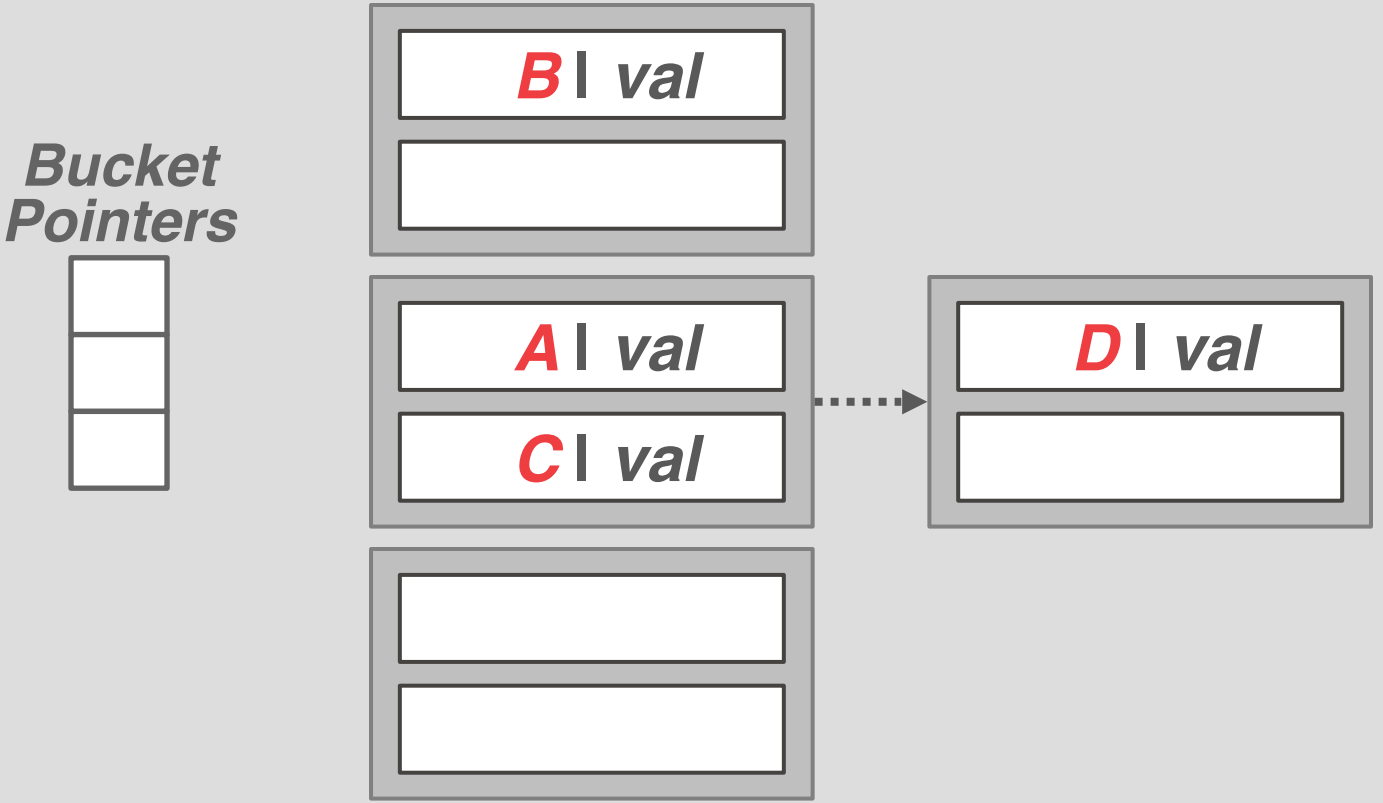
CHAINED HASHING



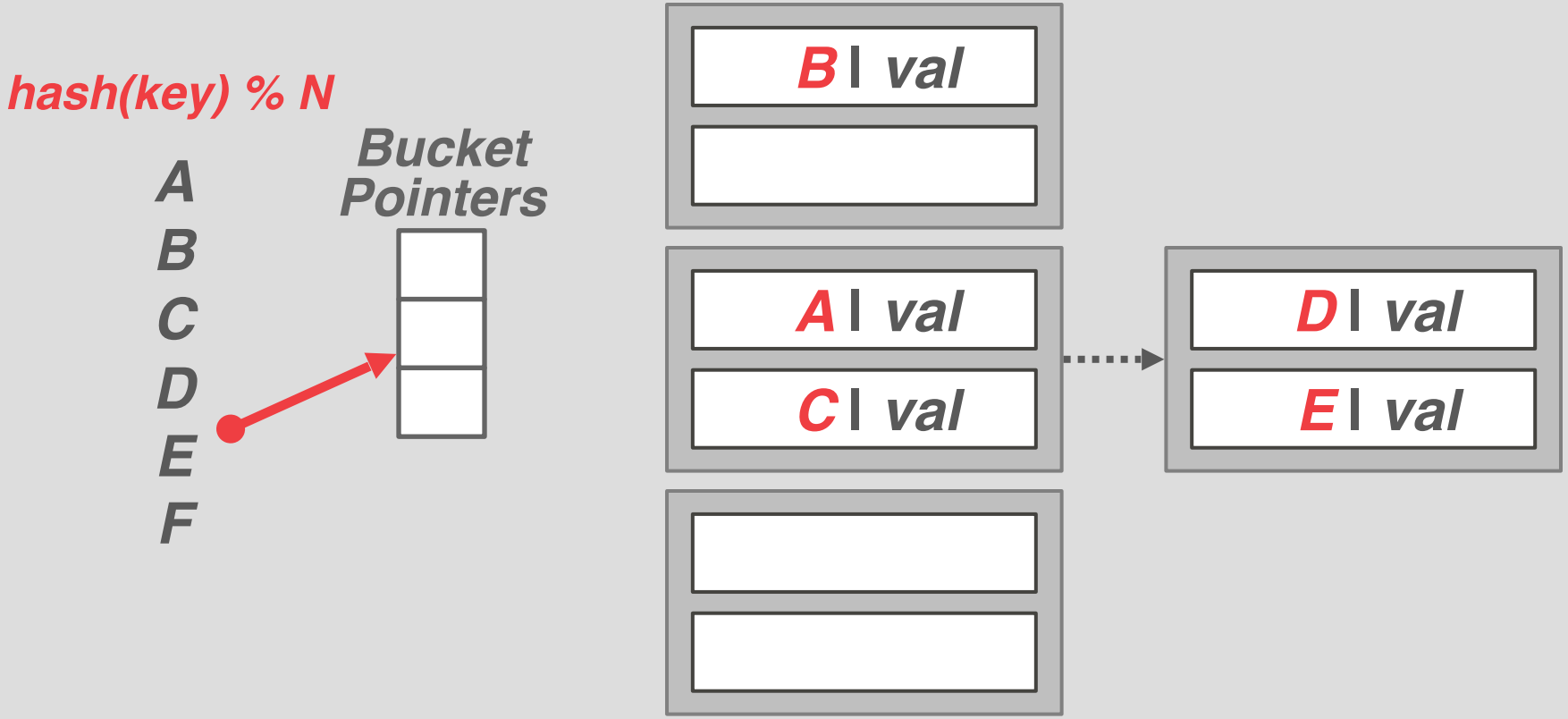
CHAINED HASHING

hash(key) % N

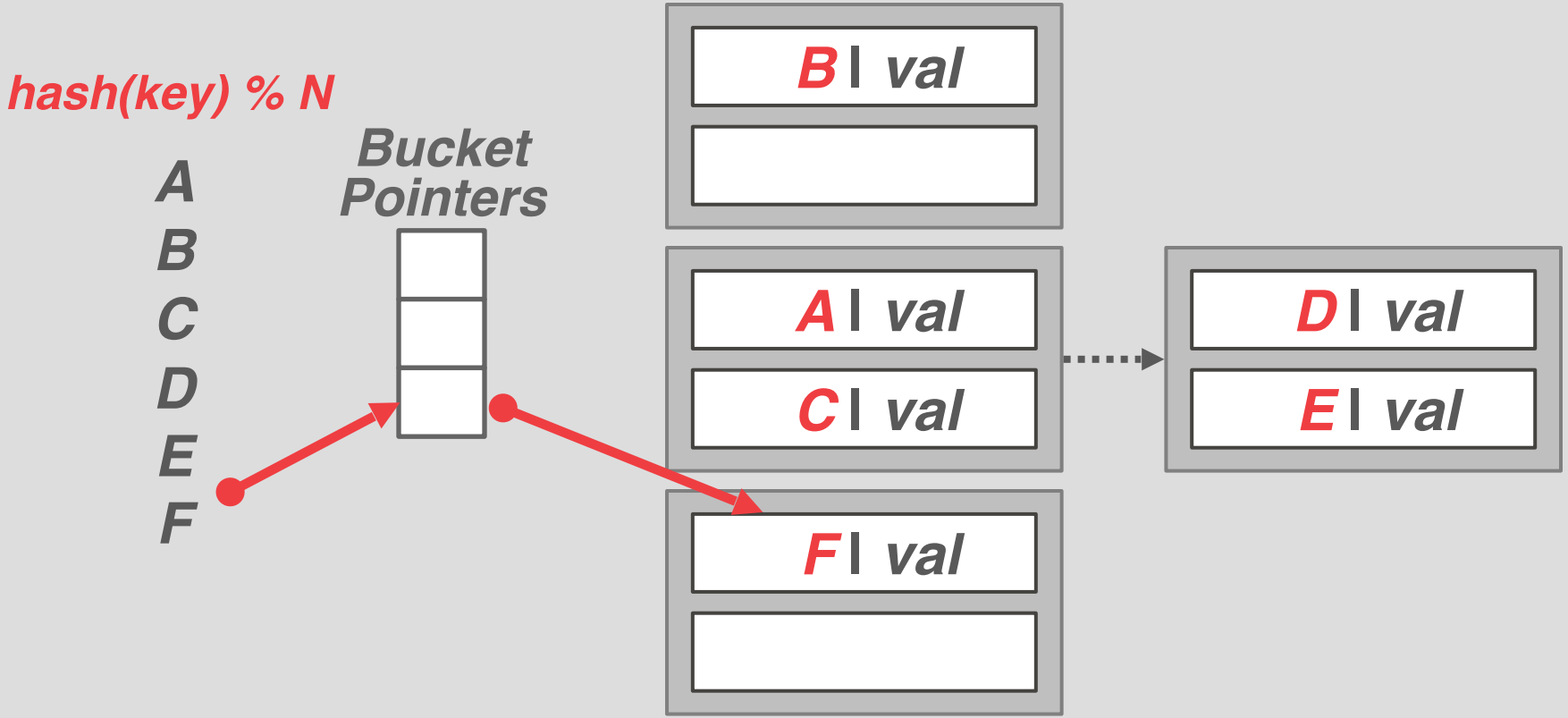
A
B
C
D
E
F



CHAINED HASHING



CHAINED HASHING



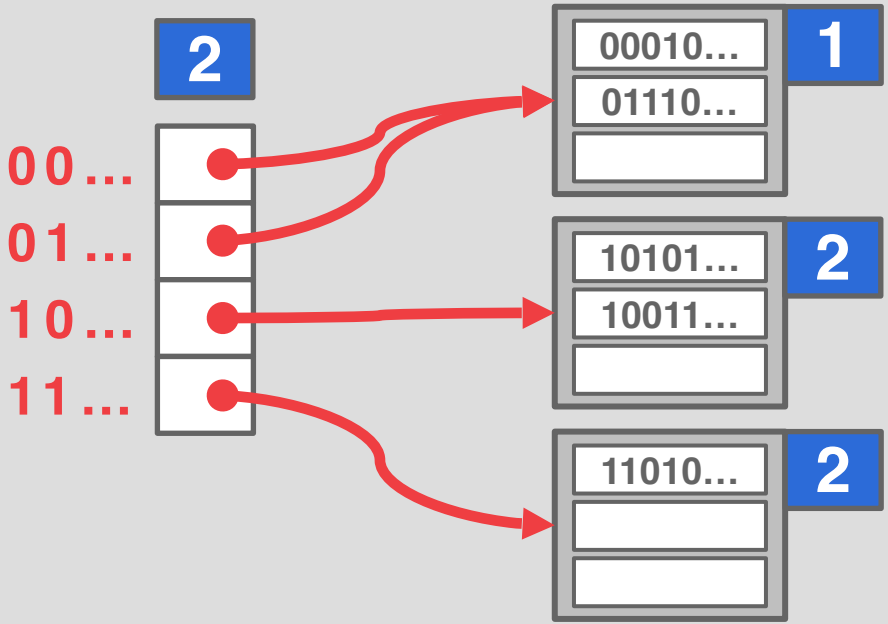
EXTENDIBLE HASHING

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

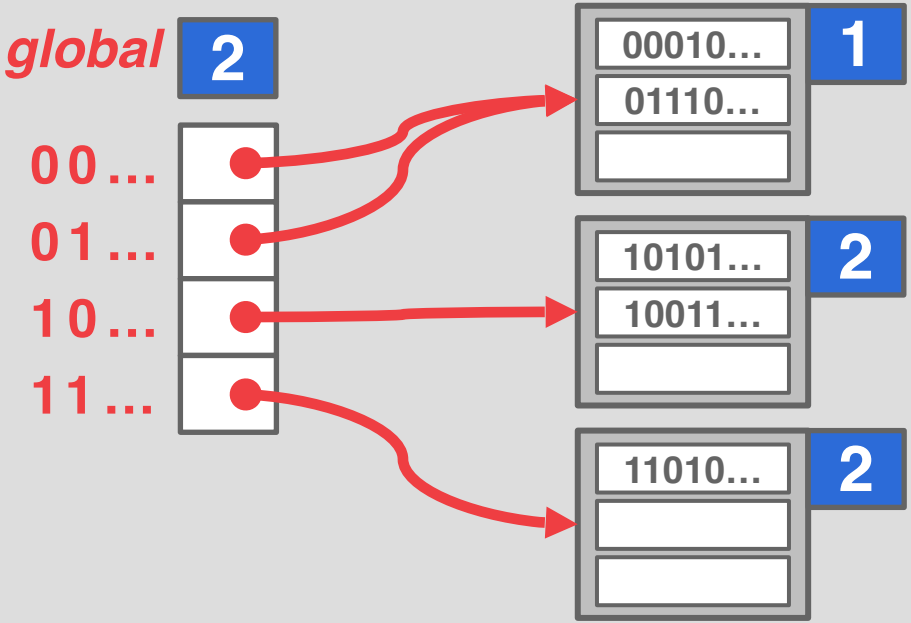
Multiple slot locations can point to the same bucket chain.

Reshuffle bucket entries on split and increase the number of bits to examine.
→ Data movement is localized to just the split chain.

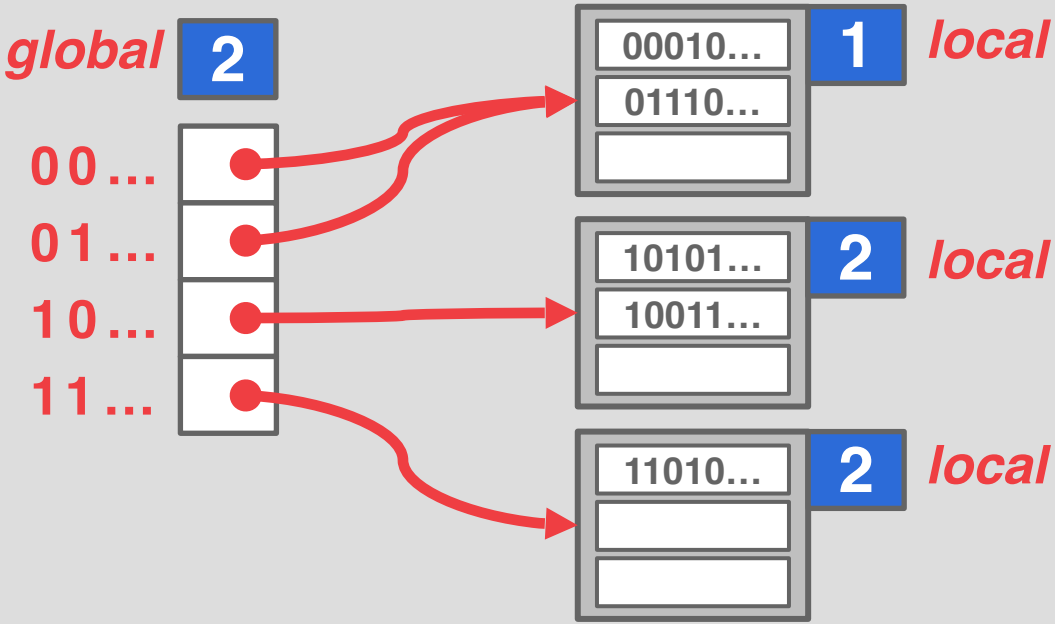
EXTENDIBLE HASHING



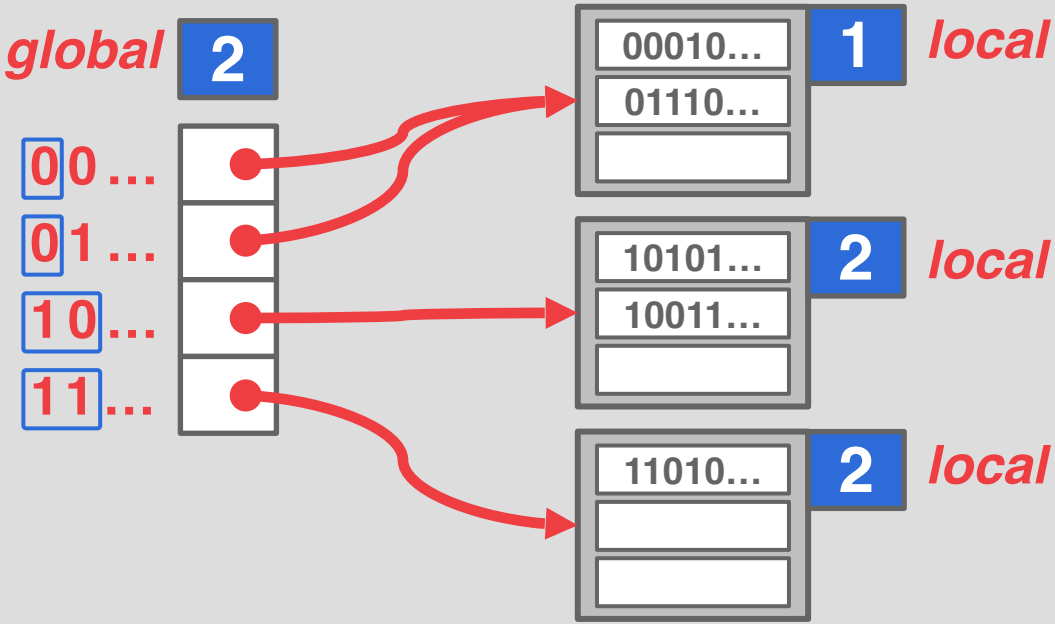
EXTENDIBLE HASHING



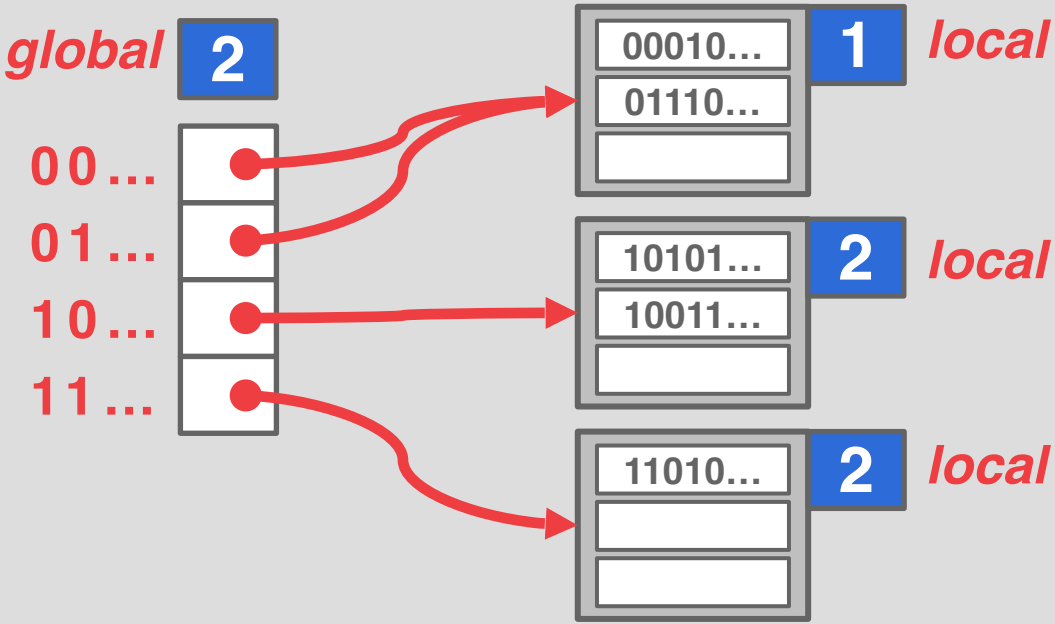
EXTENDIBLE HASHING



EXTENDIBLE HASHING



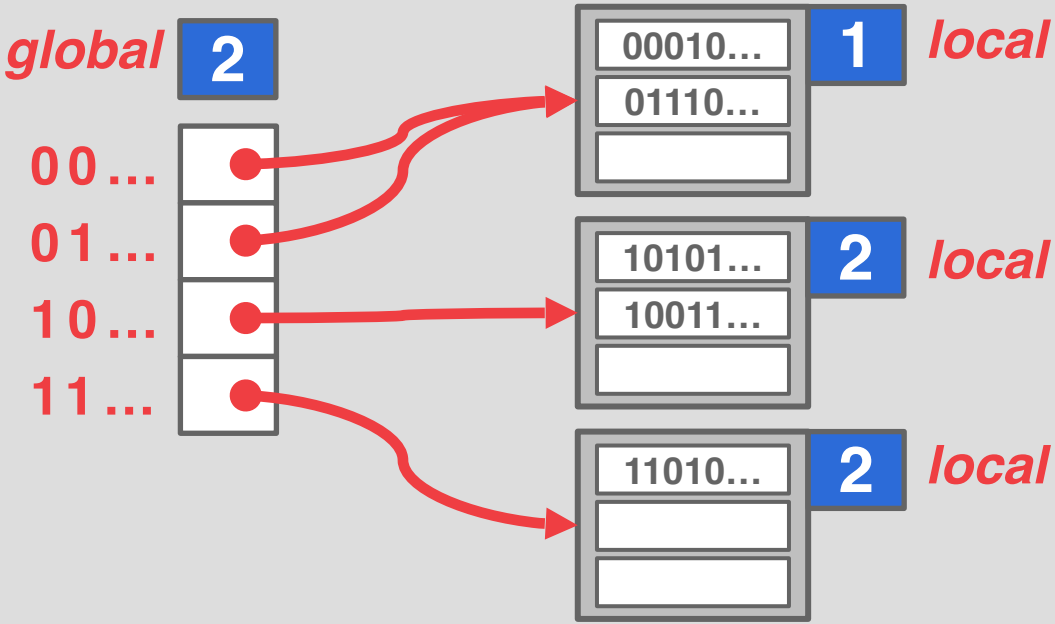
EXTENDIBLE HASHING



Get A

hash(A) = 01110…

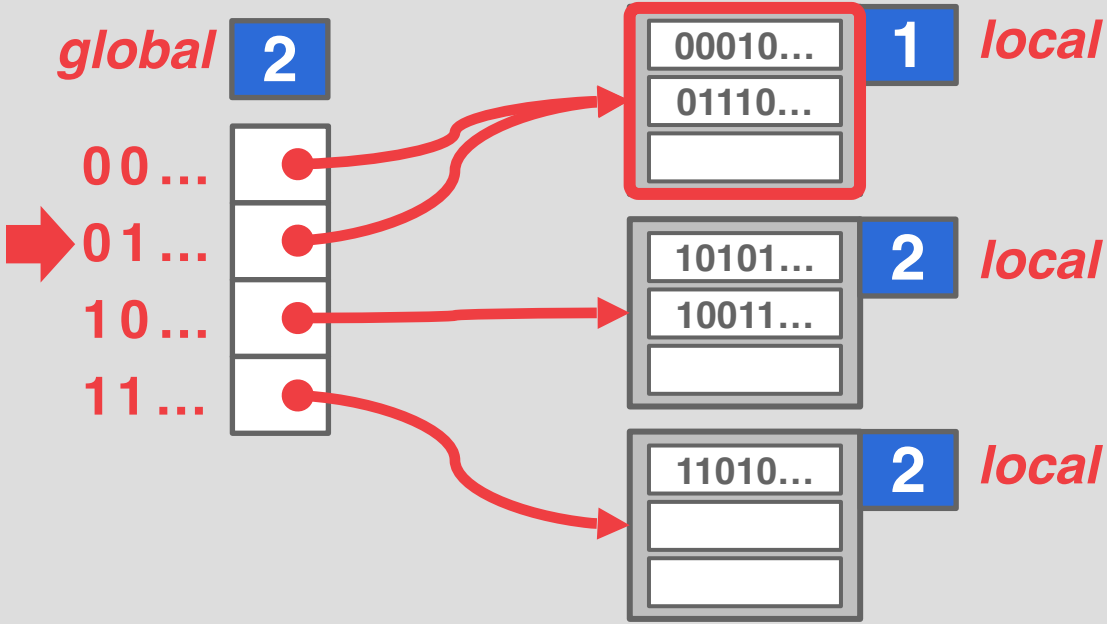
EXTENDIBLE HASHING



Get A

hash(A) = 01110…

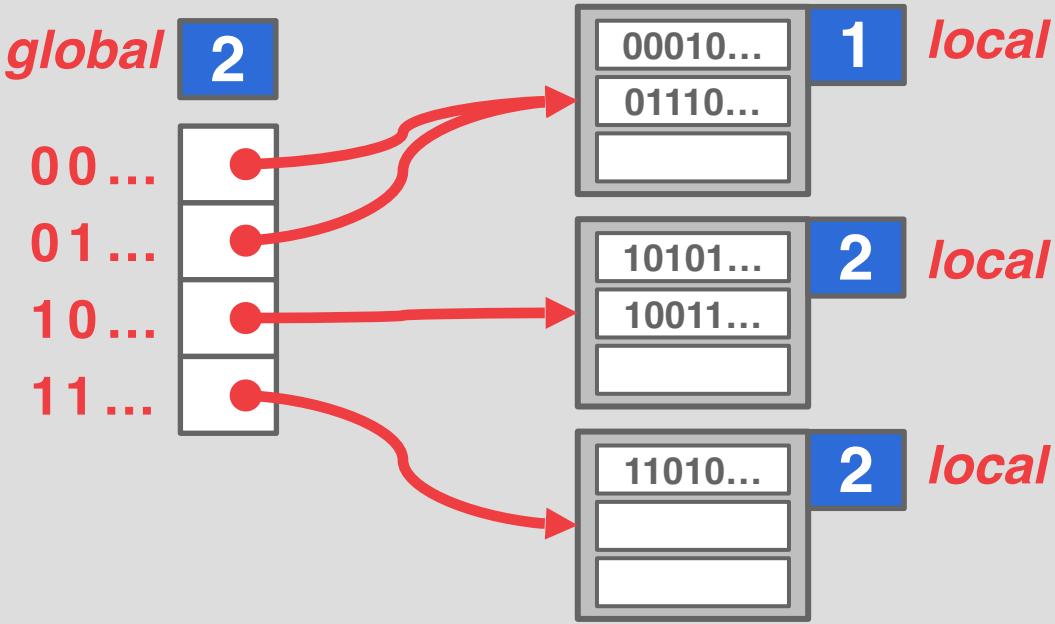
EXTENDIBLE HASHING



Get A

hash(A) = 01110…

EXTENDIBLE HASHING



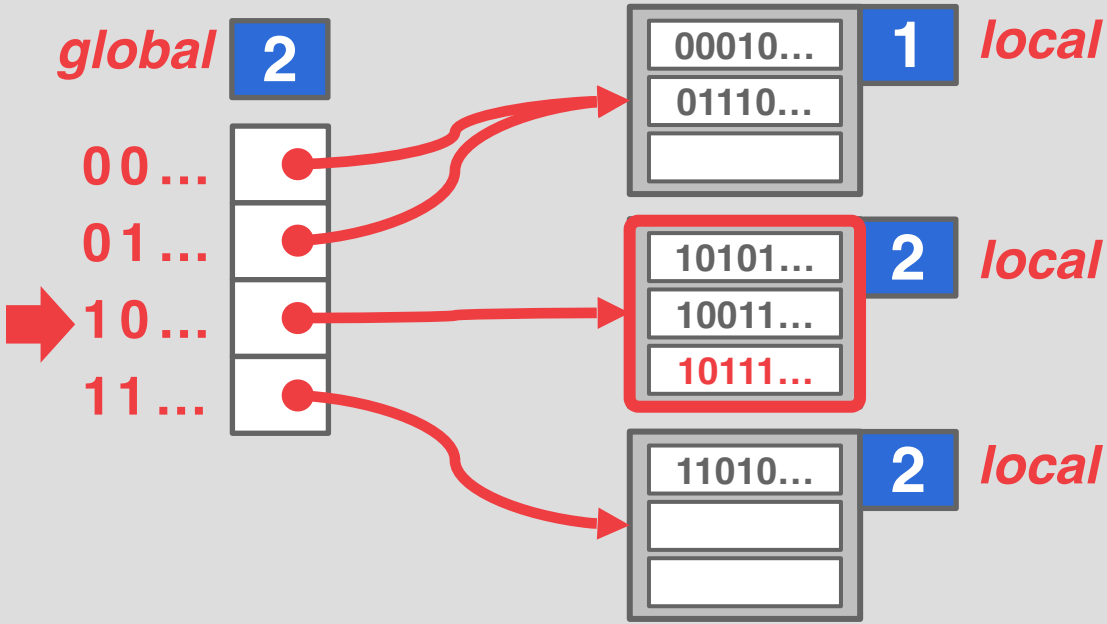
Get A

hash(A) = 01110…

Put B

hash(B) = 10111…

EXTENDIBLE HASHING



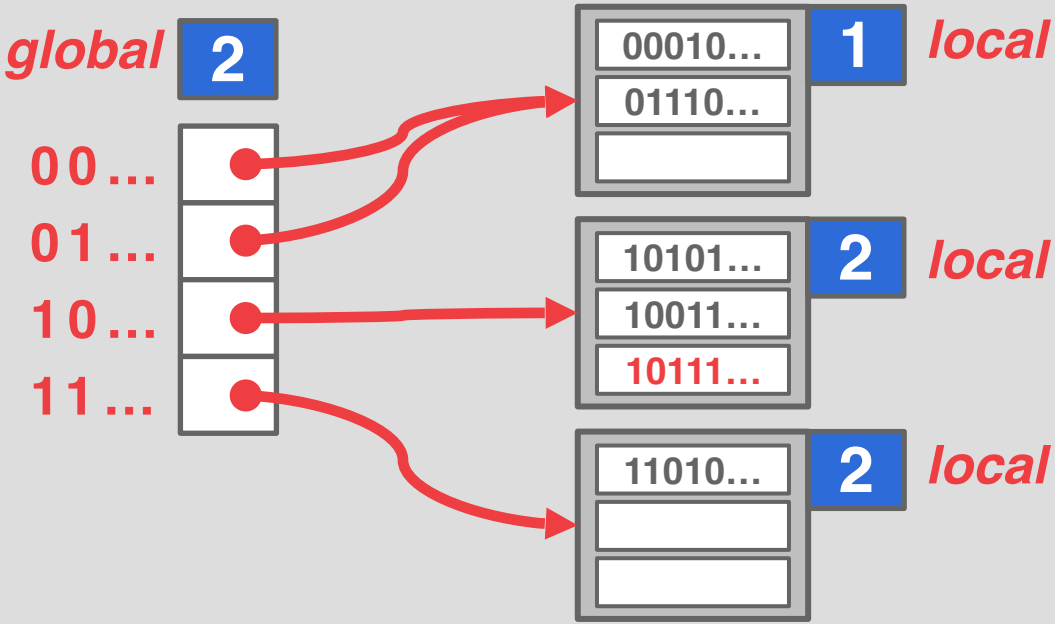
Get A

hash(A) = 01110…

Put B

hash(B) = 10111…

EXTENDIBLE HASHING

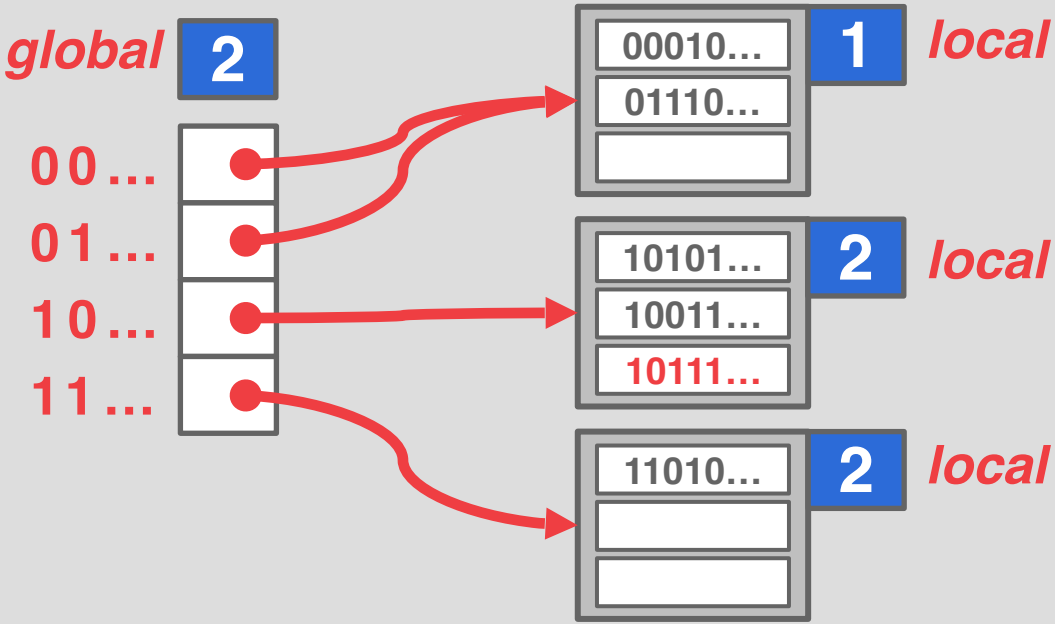


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING



Get A

hash(A) = 01110…

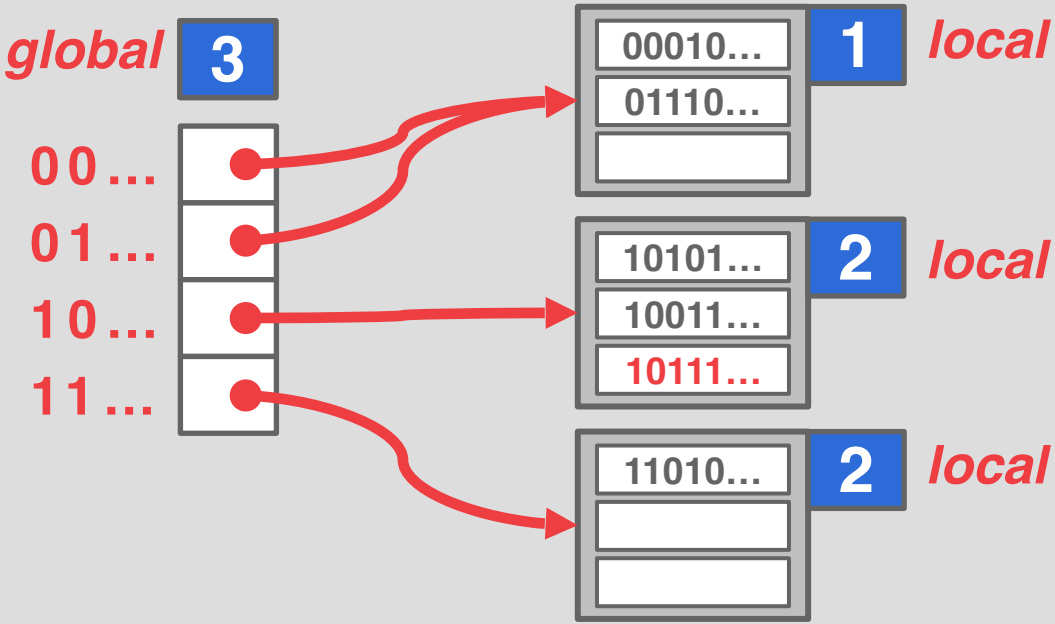
Put B

hash(B) = 10111…

Put C

hash(C) = 10100…

EXTENDIBLE HASHING



Get A

hash(A) = 01110…

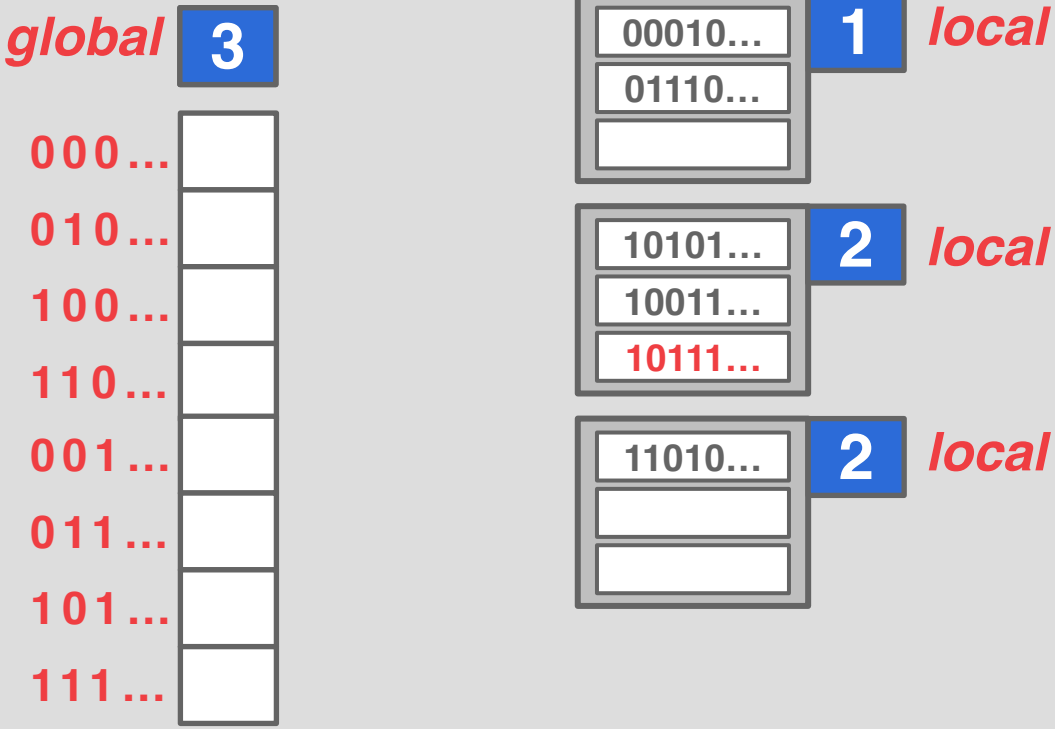
Put B

hash(B) = 10111…

Put C

hash(C) = 10100…

EXTENDIBLE HASHING



Get A

hash(A) = 01110...

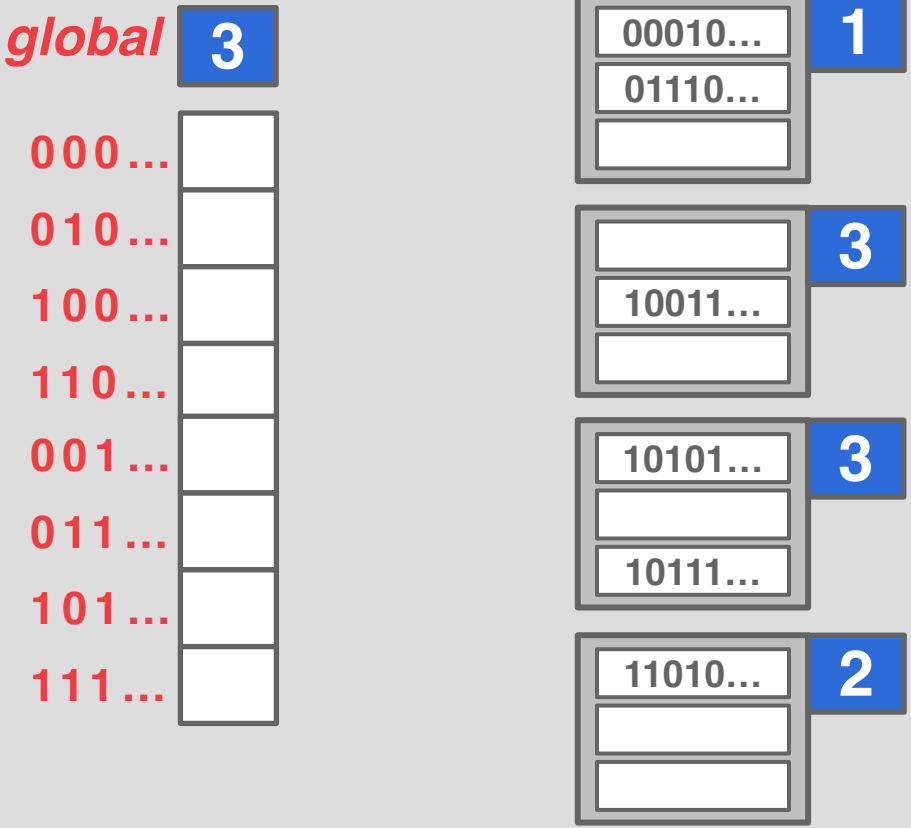
Put B

hash(B) = 10111...

Put C

hash(C) = 10100...

EXTENDIBLE HASHING

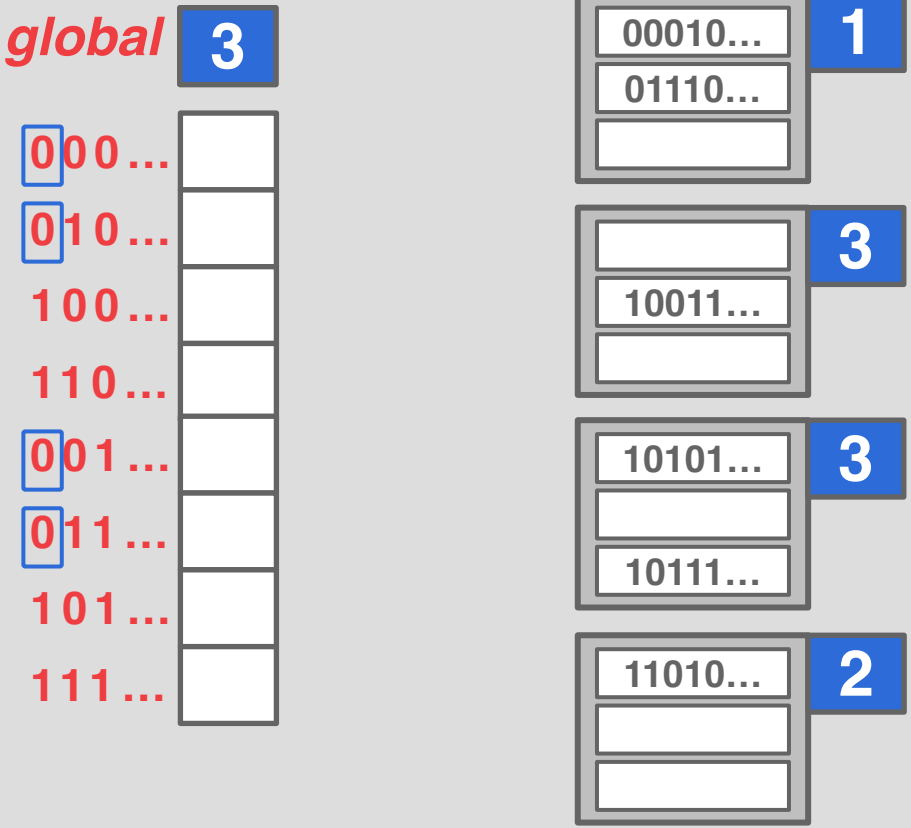


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING

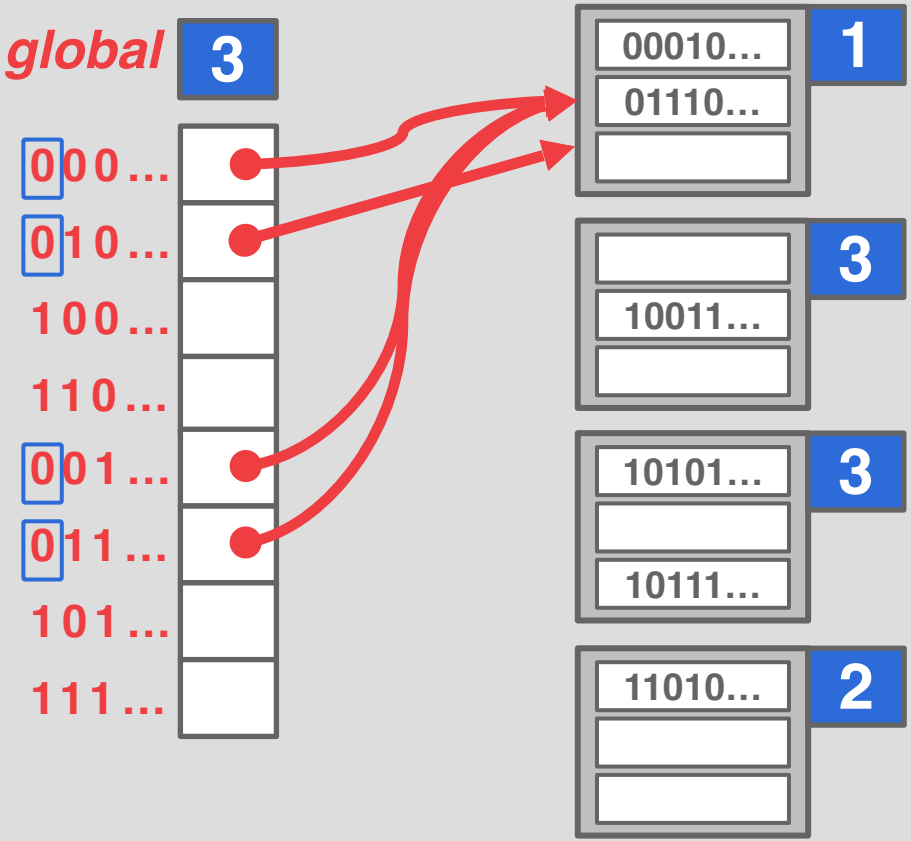


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING

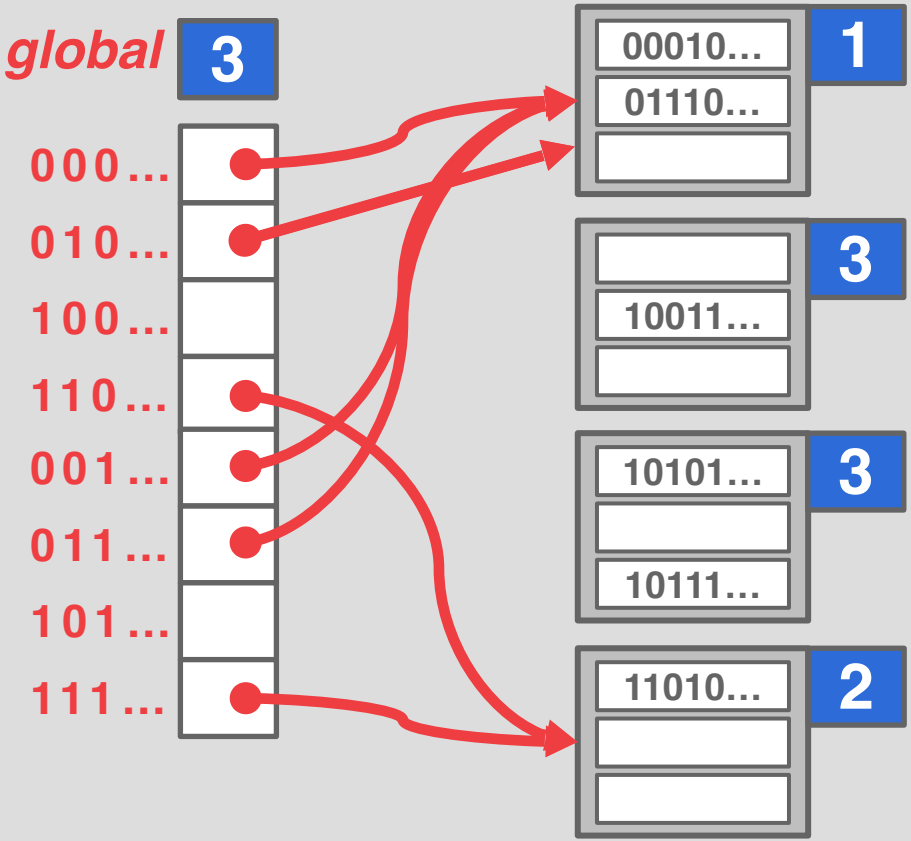


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING

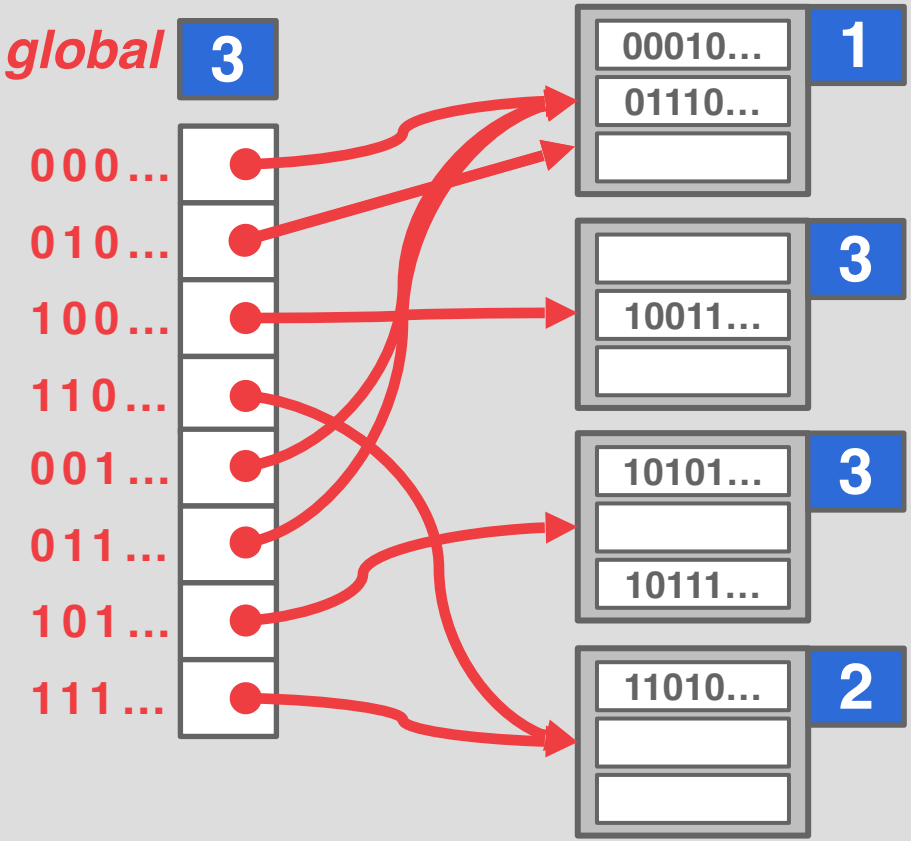


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING

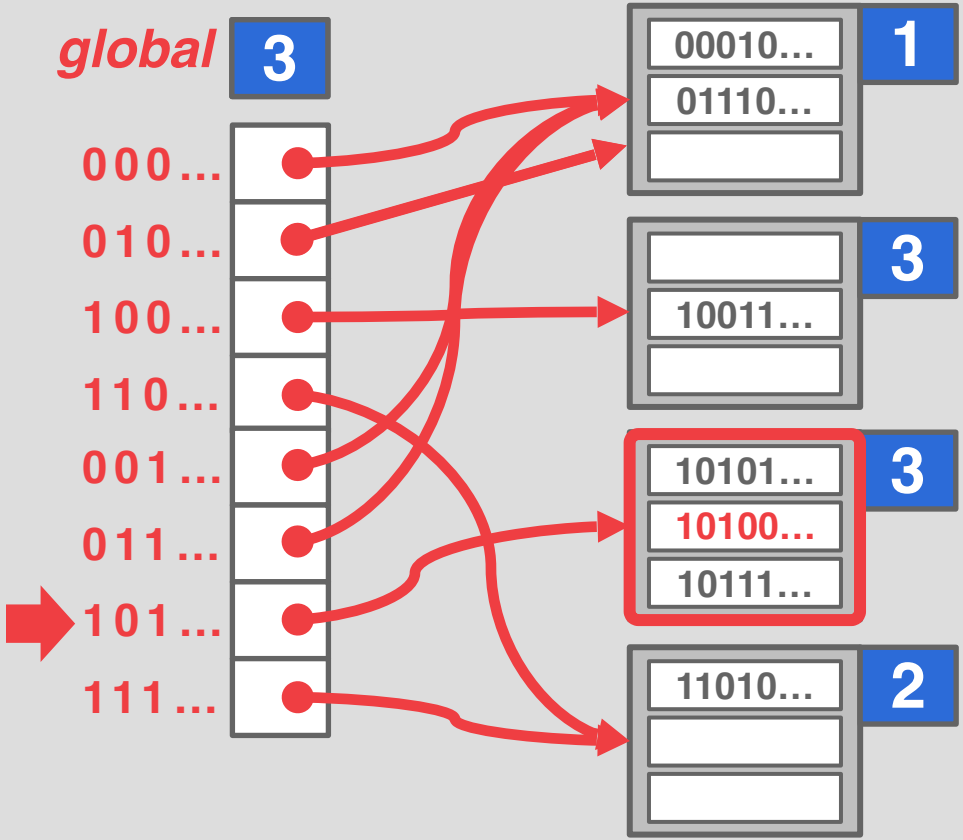


Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

EXTENDIBLE HASHING



Get A
hash(A) = 01110…

Put B
hash(B) = 10111…

Put C
hash(C) = 10100…

LINEAR HASHING

The hash table maintains a pointer that tracks the next bucket to split.

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

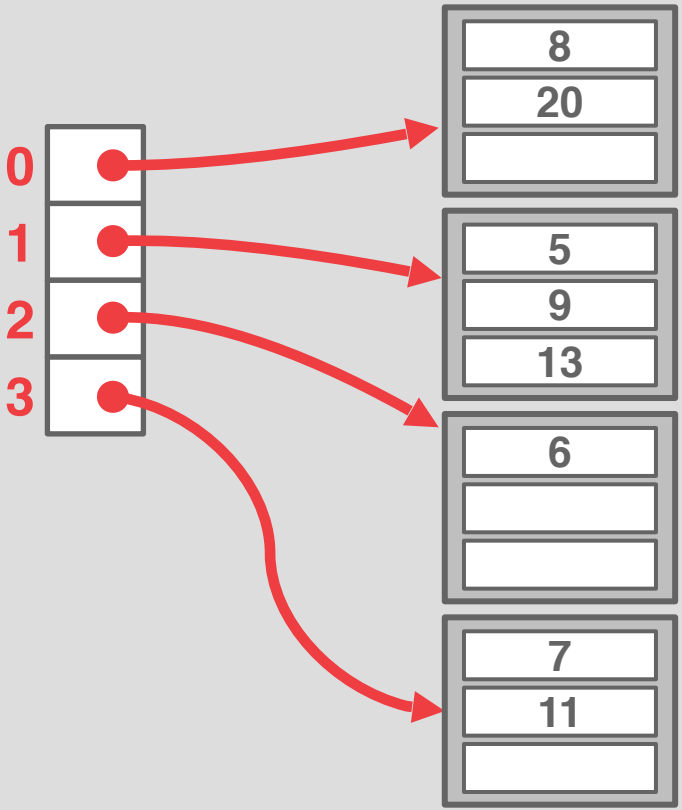
Use multiple hashes to find the right bucket for a given key.

Can use different overflow criterion:

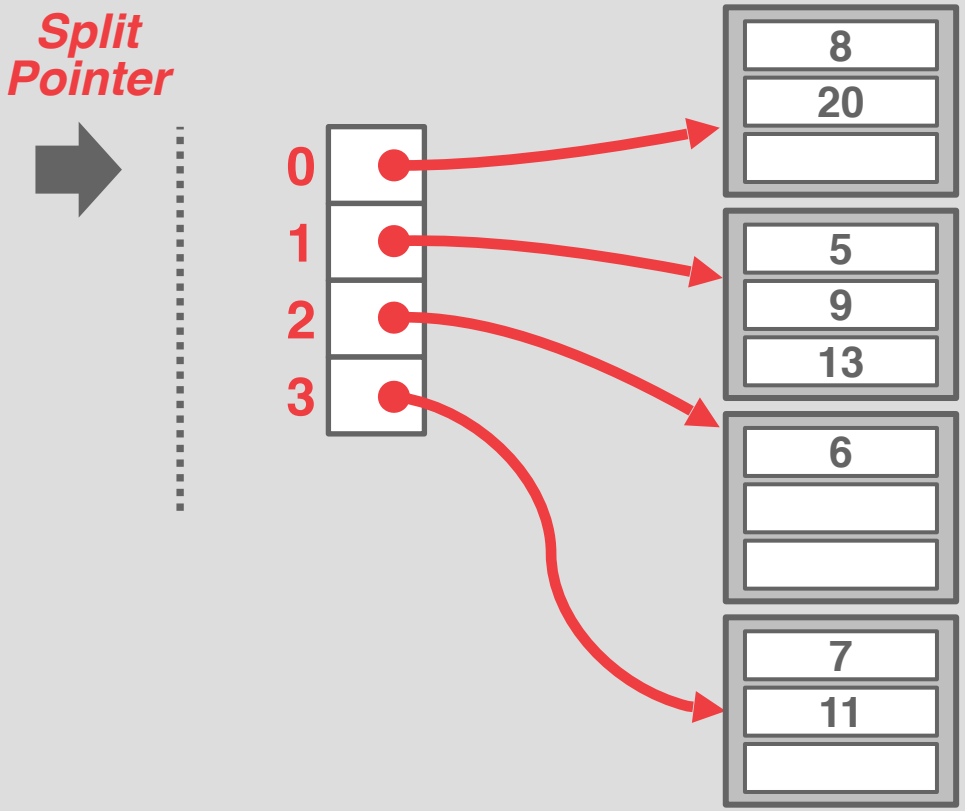
→ Space Utilization

→ Average Length of Overflow Chains

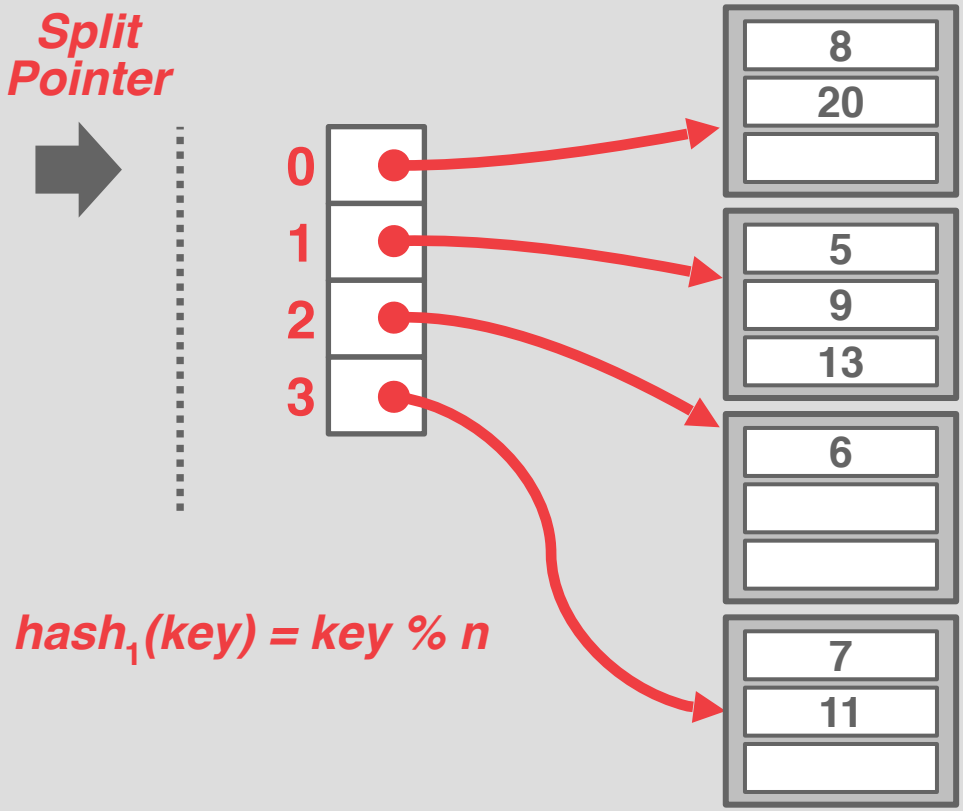
LINEAR HASHING



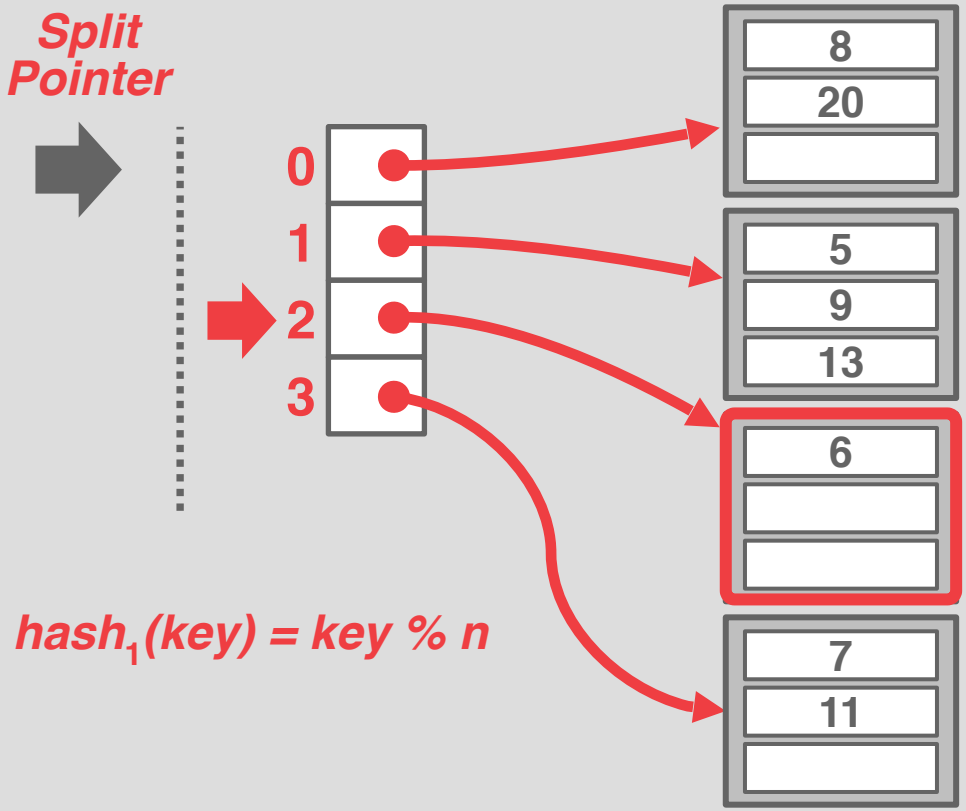
LINEAR HASHING



LINEAR HASHING



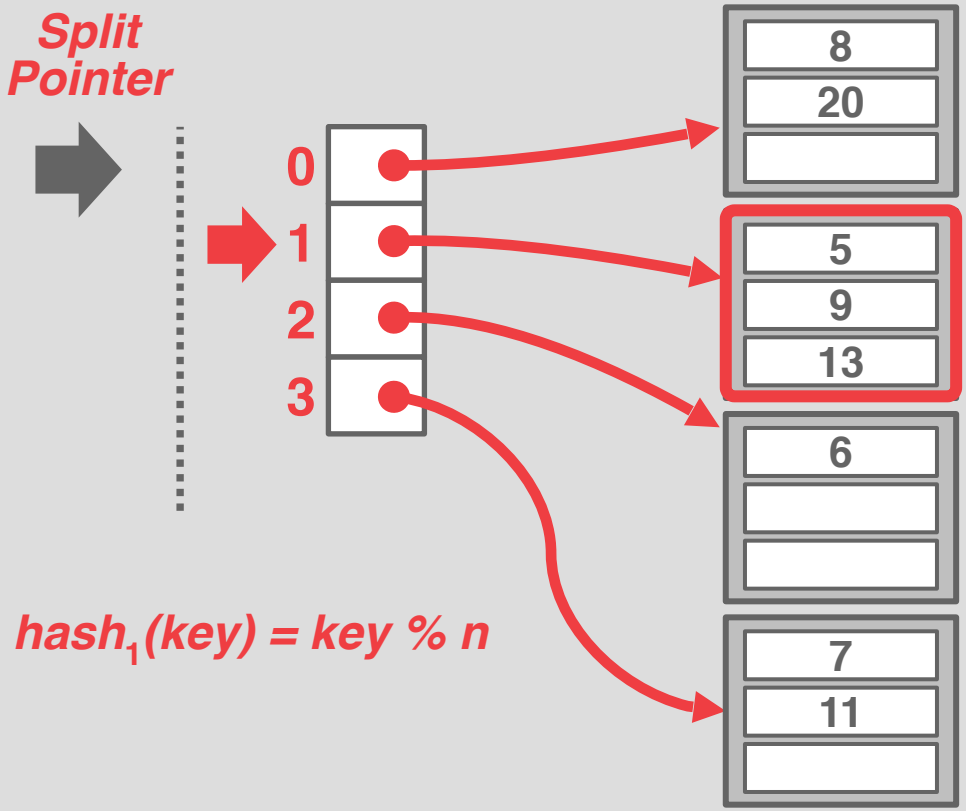
LINEAR HASHING



Get 6

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

LINEAR HASHING



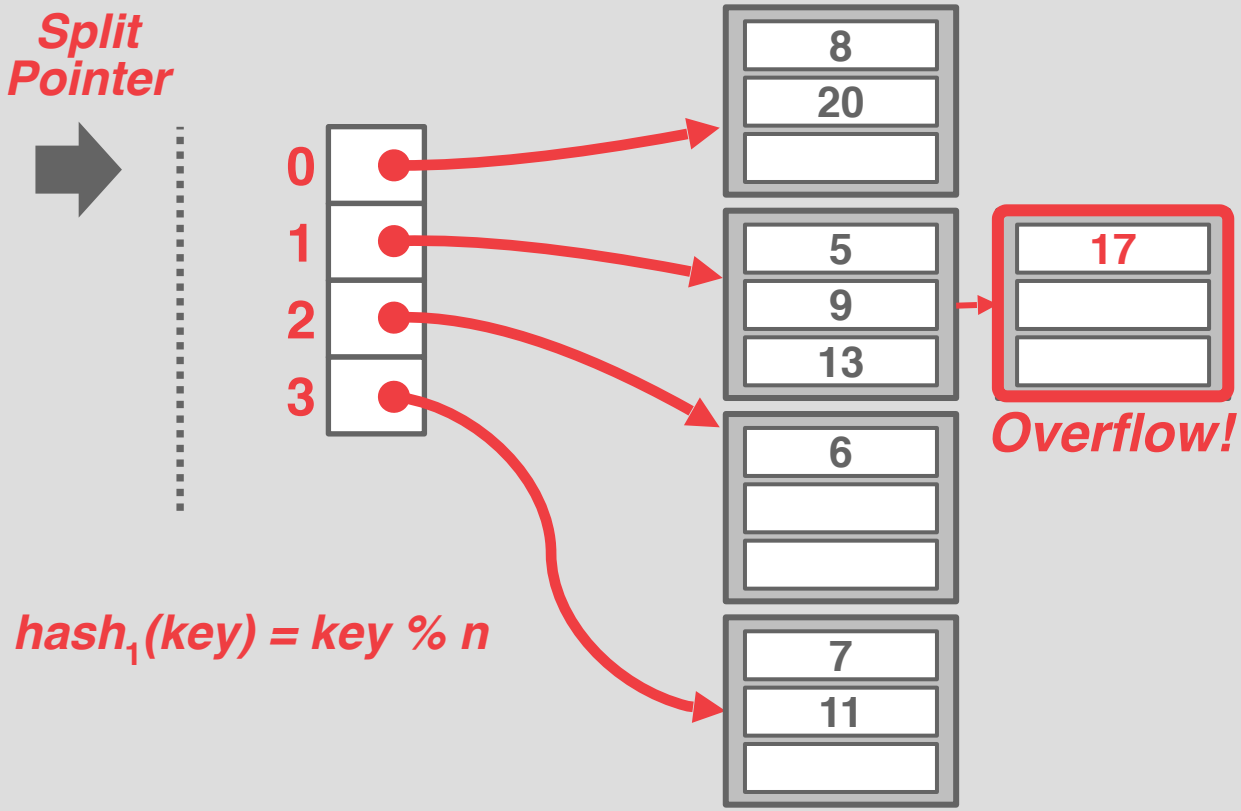
Get 6

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

Put 17

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

LINEAR HASHING



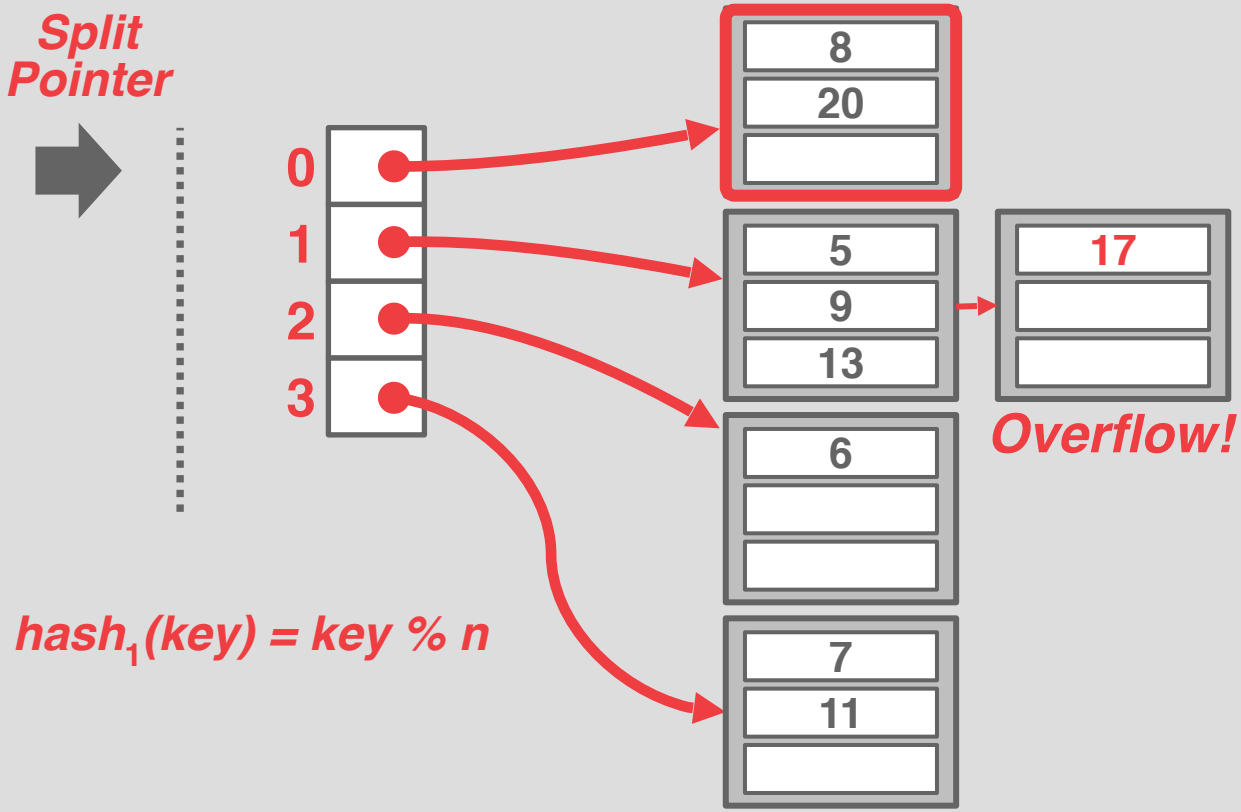
Get 6

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

Put 17

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

LINEAR HASHING



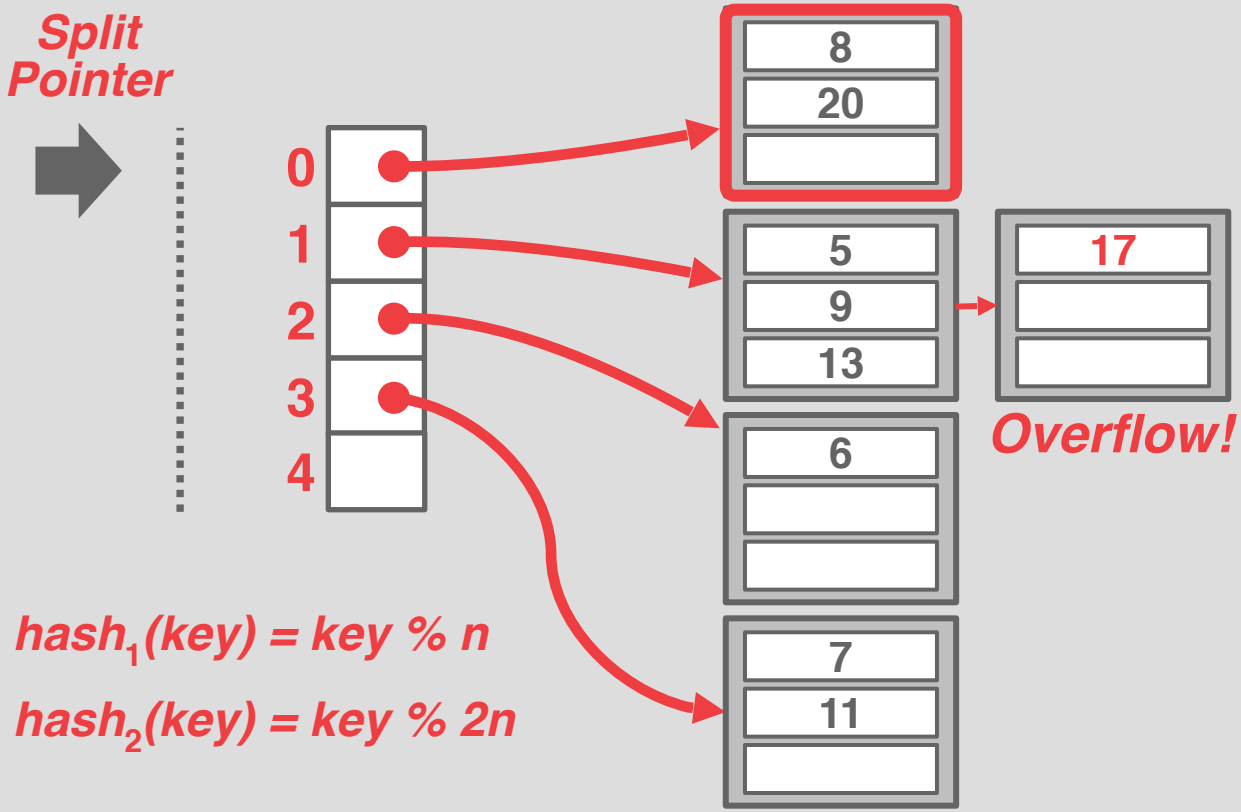
Get 6

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

Put 17

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

LINEAR HASHING



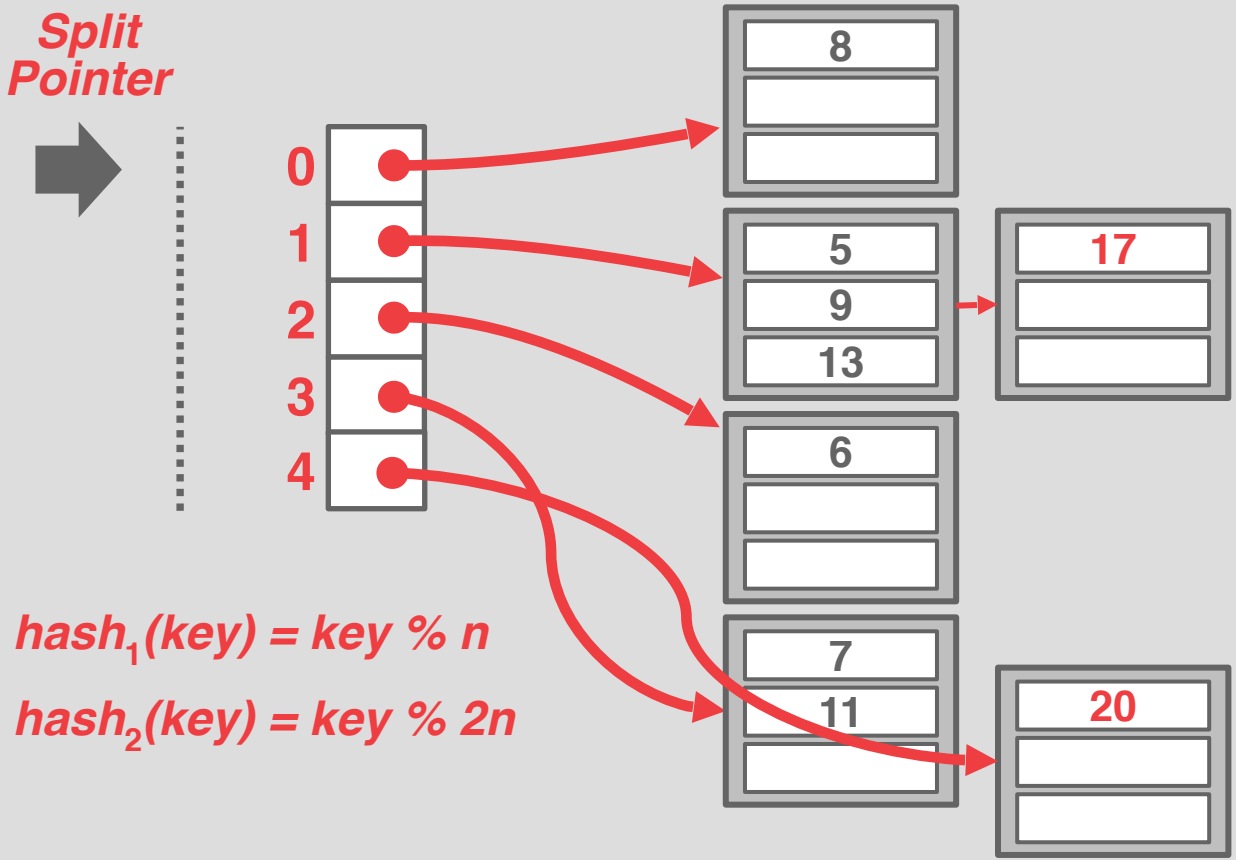
Get 6

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

Put 17

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

LINEAR HASHING



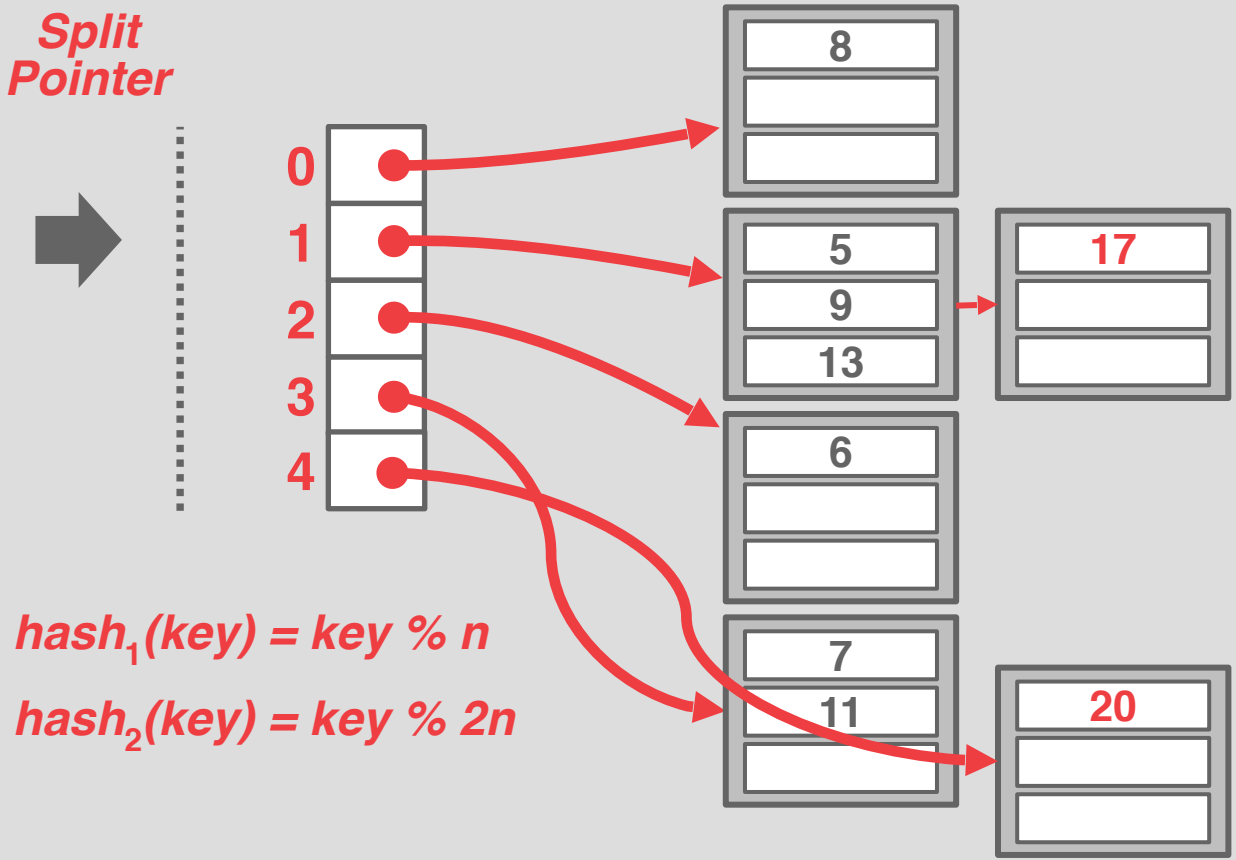
Get 6

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

Put 17

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

LINEAR HASHING



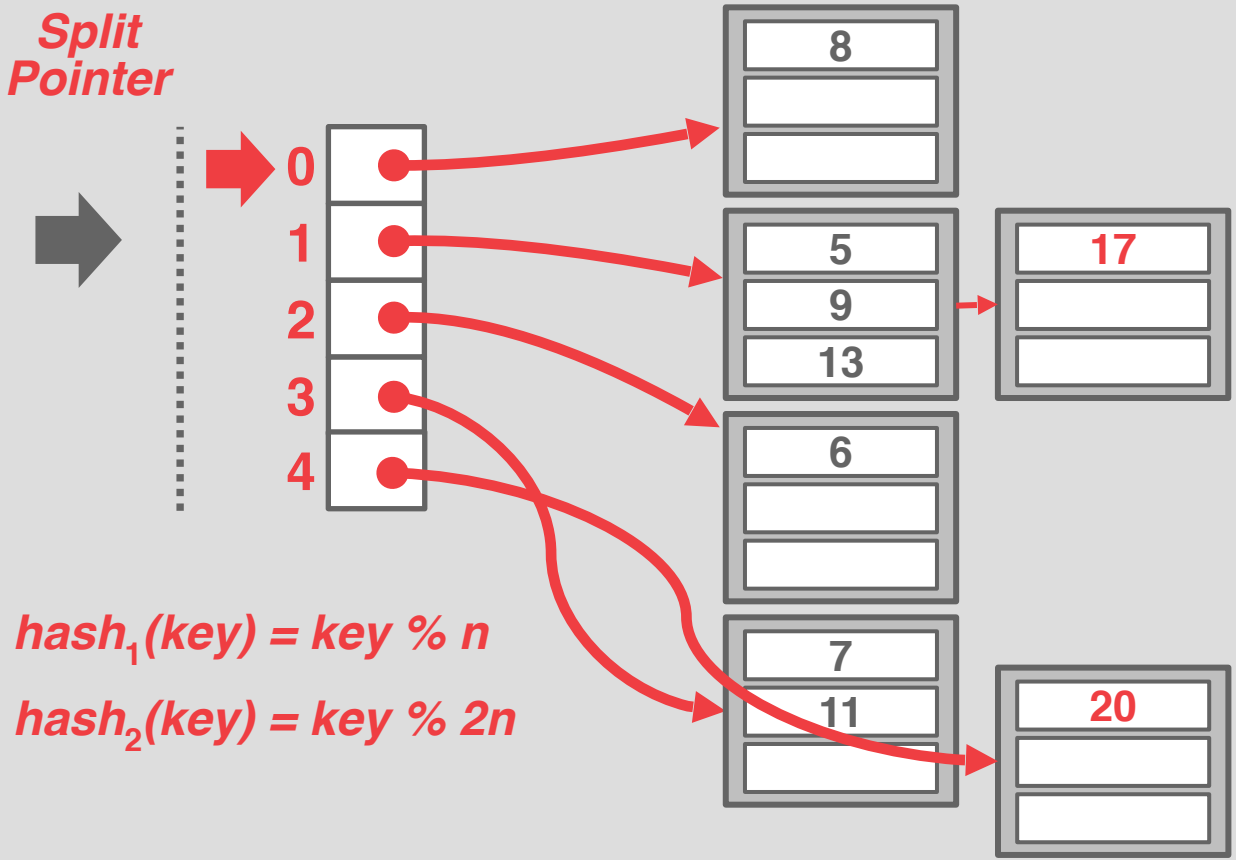
Get 6

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

Put 17

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

LINEAR HASHING



Get 6

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

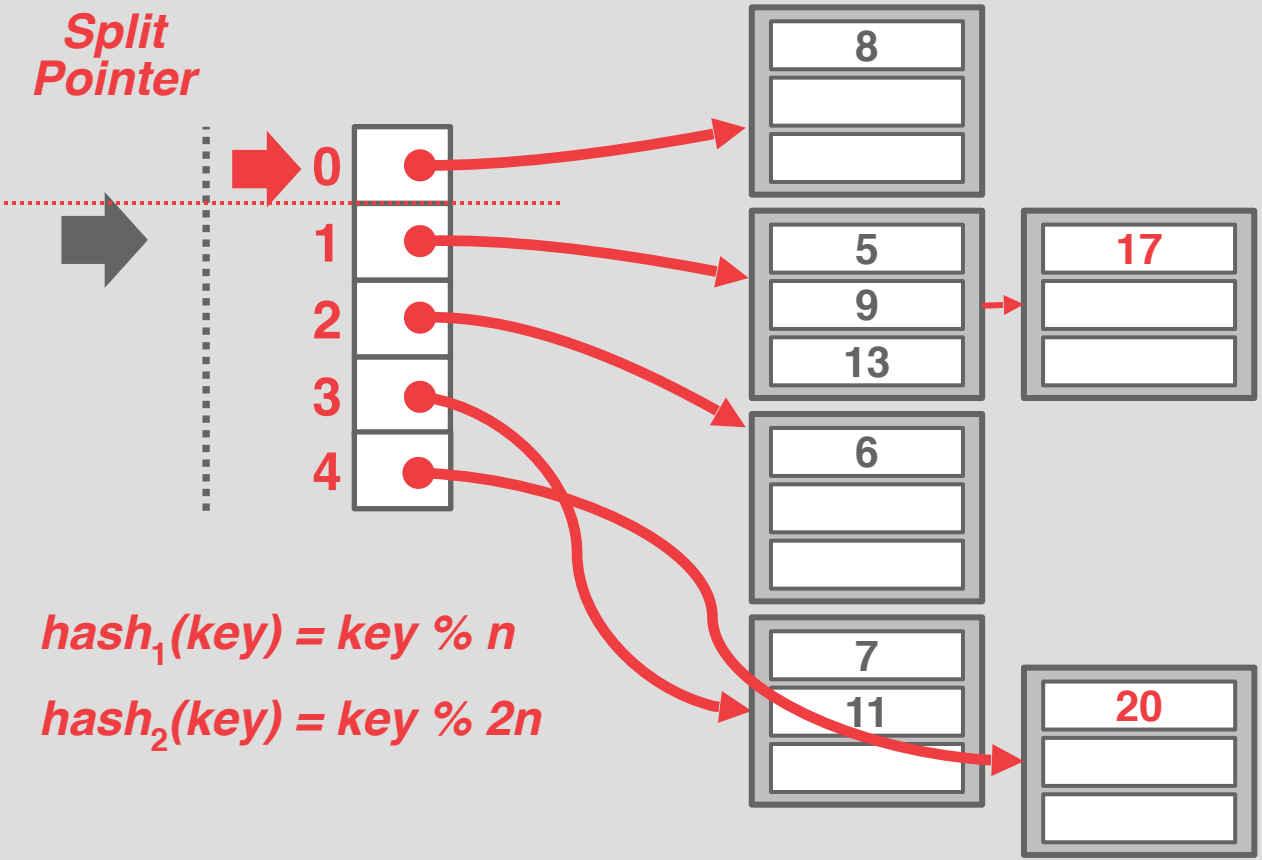
Put 17

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

Get 20

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

LINEAR HASHING



Get 6

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

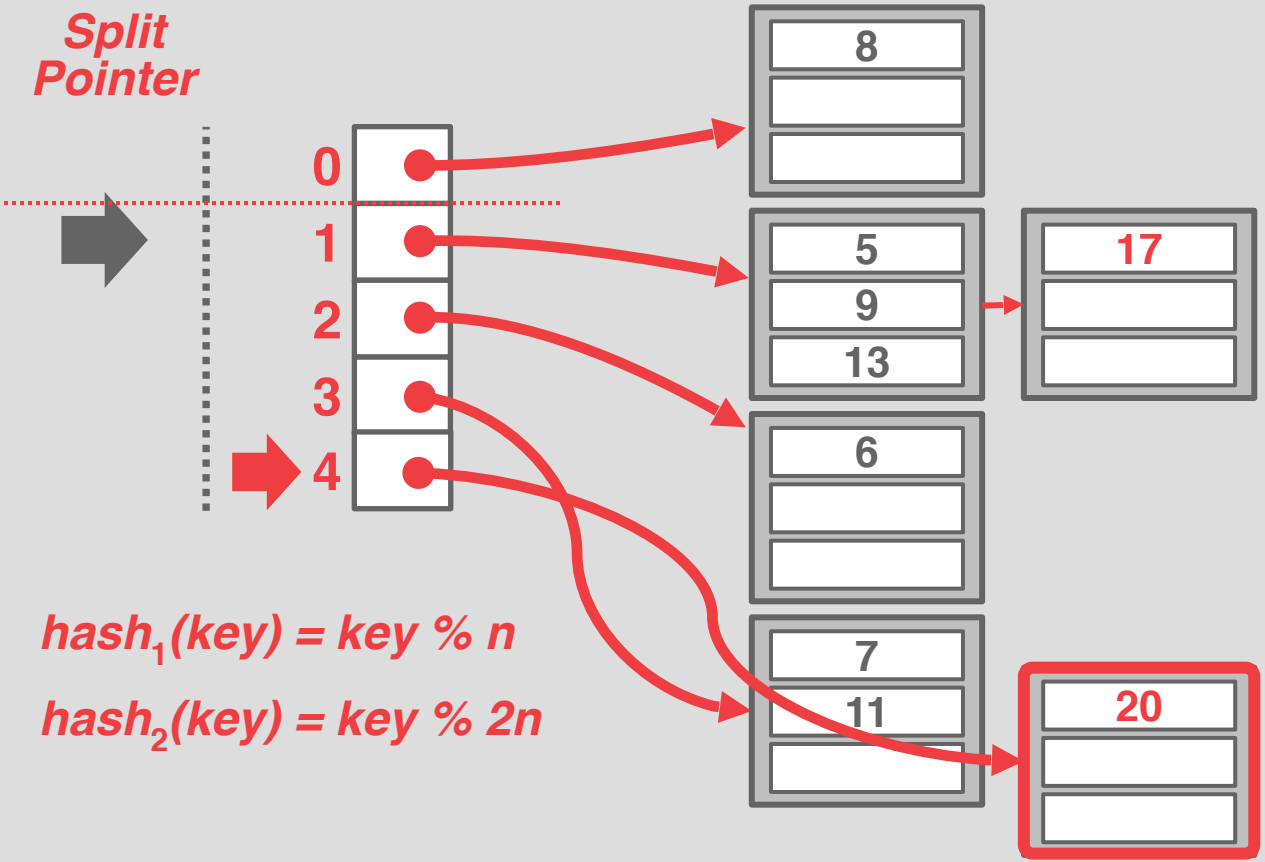
Put 17

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

Get 20

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

LINEAR HASHING



Get 6

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

Put 17

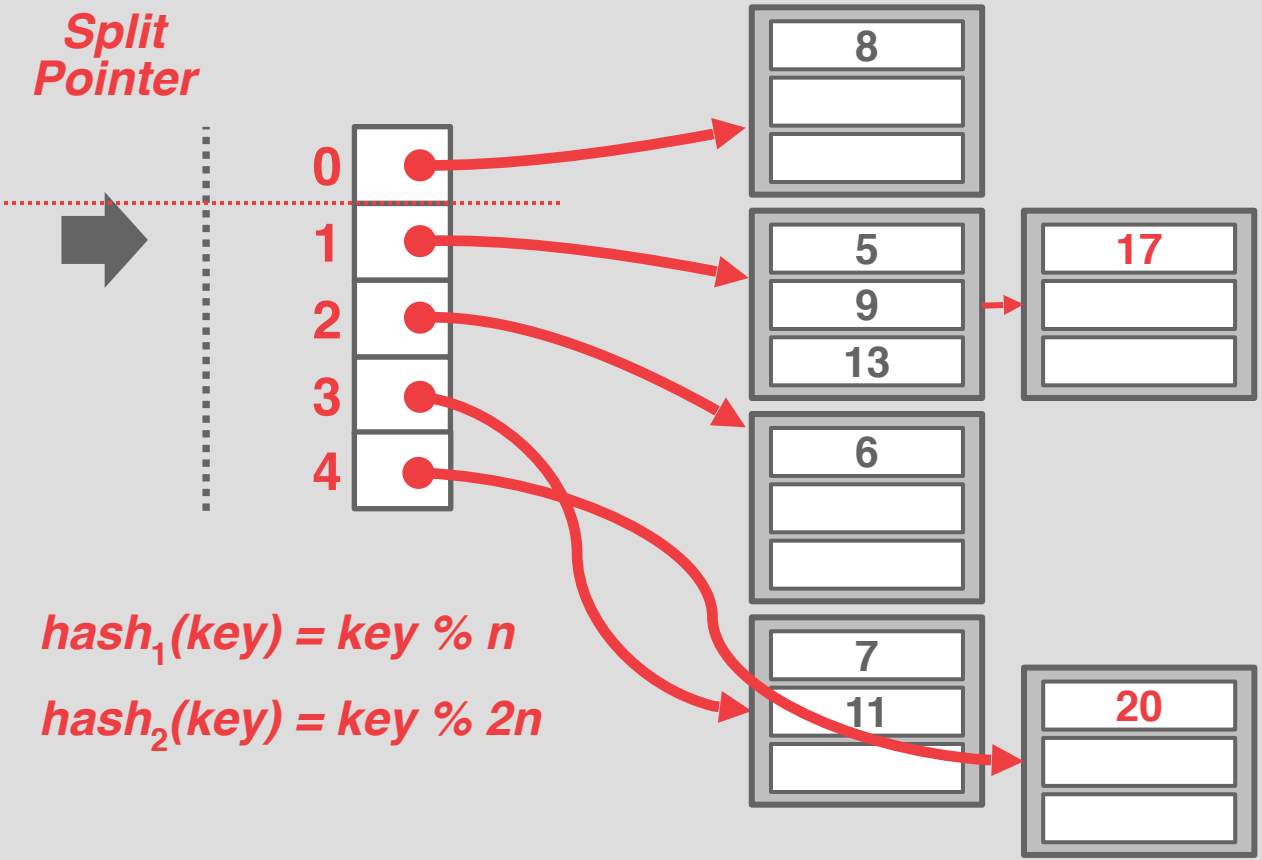
$hash_1(17) = 17 \% 4 = 1$

Get 20

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

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

LINEAR HASHING



Get 6

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

Put 17

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

Get 20

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

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

Get 9

$hash_1(9) = 9 \% 4 = 1$

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.

CONCLUSION

Hash tables are fast data structures that support **O(1)** lookups and used throughout DBMS internals.
→ Trade-off between speed and flexibility.

Hash tables are usually **not** what you want to use for a table index…

NEXT CLASS

B+Trees