Introduction to Hash Tables

By the end of this video you will be able to…

- Describe why hash tables are valuable
- Understand the role of a hash function

Thought Experiment

You have to keep track of integers which are part of a set. You know the possible integers will always be between 0 and 999,999. If you store this as an array of 1,000,000 booleans, how long does it take to add an integer to your set?

Thought Experiment

You have to keep track of integers which are part of a set. You know the possible integers will always be between 0 and 999,999. If you store this as an array of 1,000,000 booleans, how long does it take to add an integer to your set?

A. O(N)
B. O(logN)
C. O(1)

Thought Experiment

You have to keep track of integers which are part of a set. You know the possible integers will always be between 0 and 999,999. If you store this as an array of 1,000,000 booleans, how long does it take to add an integer to your set?

A. O(N)
B. O(logN)
C. O(1)

Thought Experiment

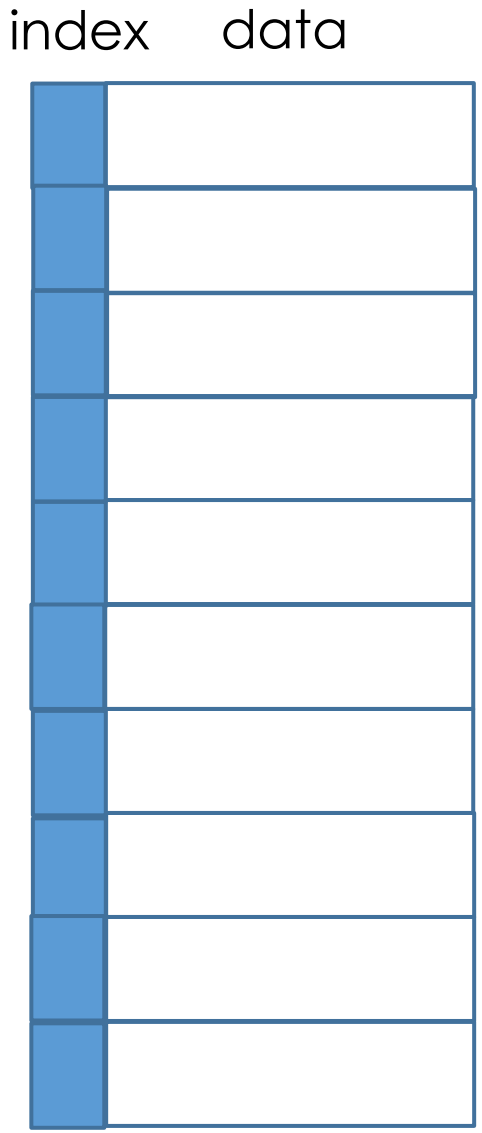
You have to keep track of integers which are part of a set. You know the possible integers will always be between 0 and 999,999. If you store this as an array of 1,000,000 booleans, how long does it take to add an integer to your set?

A. O(N)

B. O(logN)

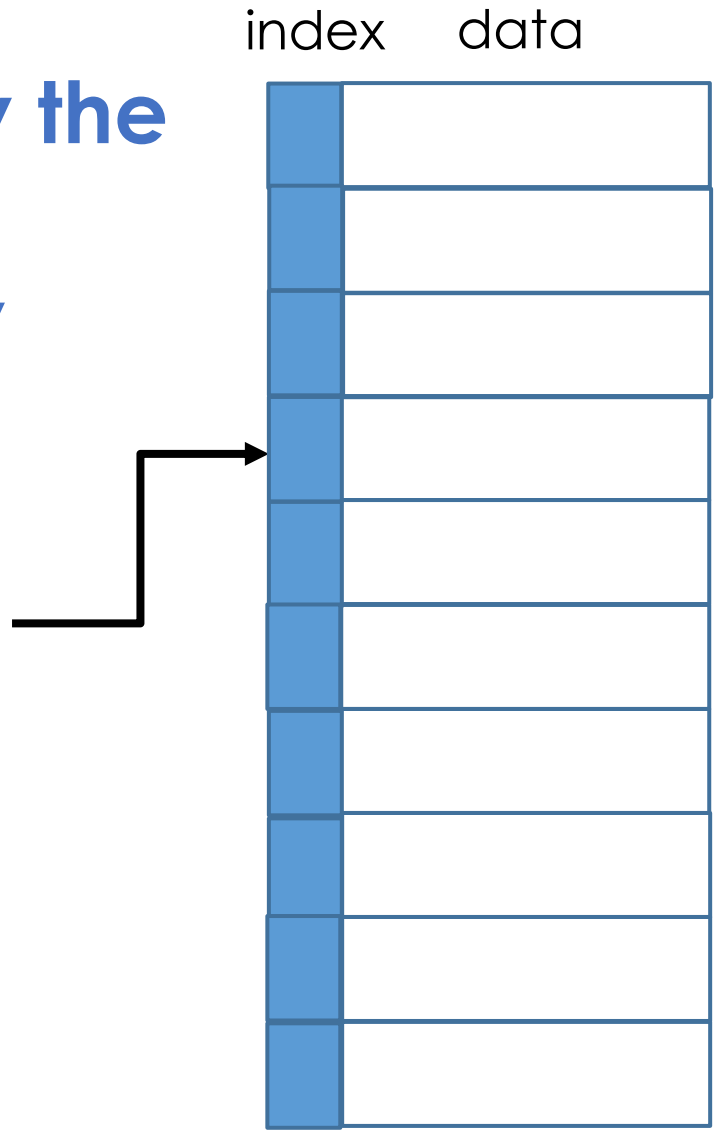
C. O(1)

Arrays are fast!



Arrays are fast!

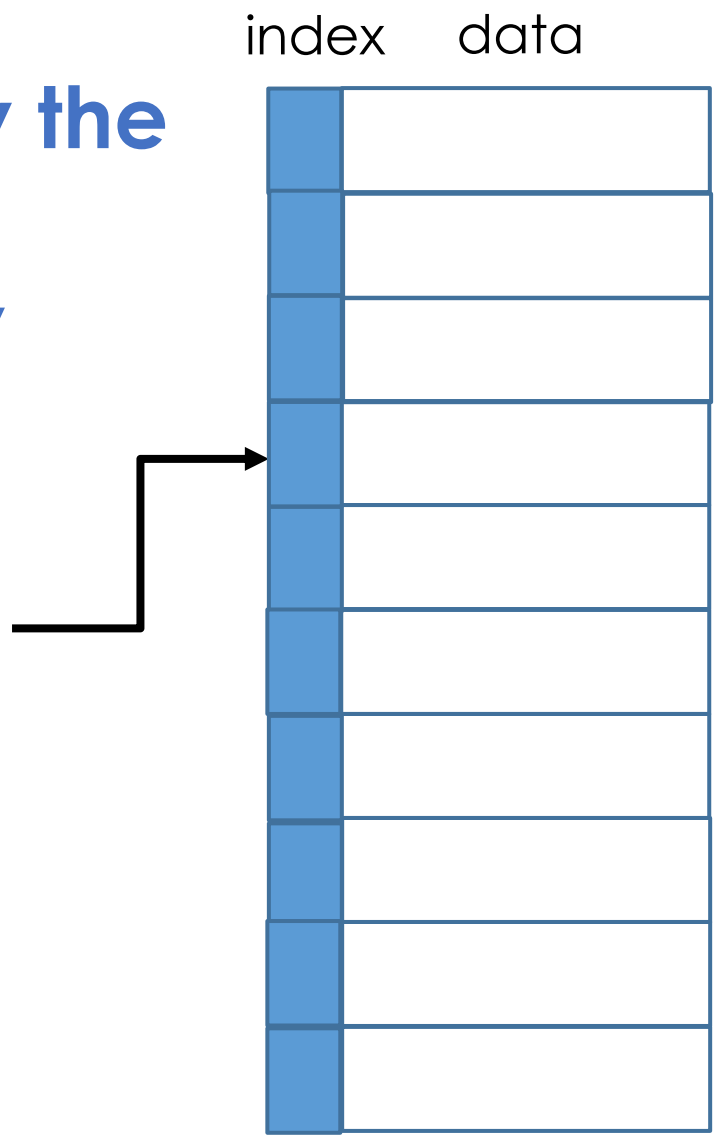
If you know where in memory the array starts, you can easily determine the address of any element using the index.



Arrays are fast!

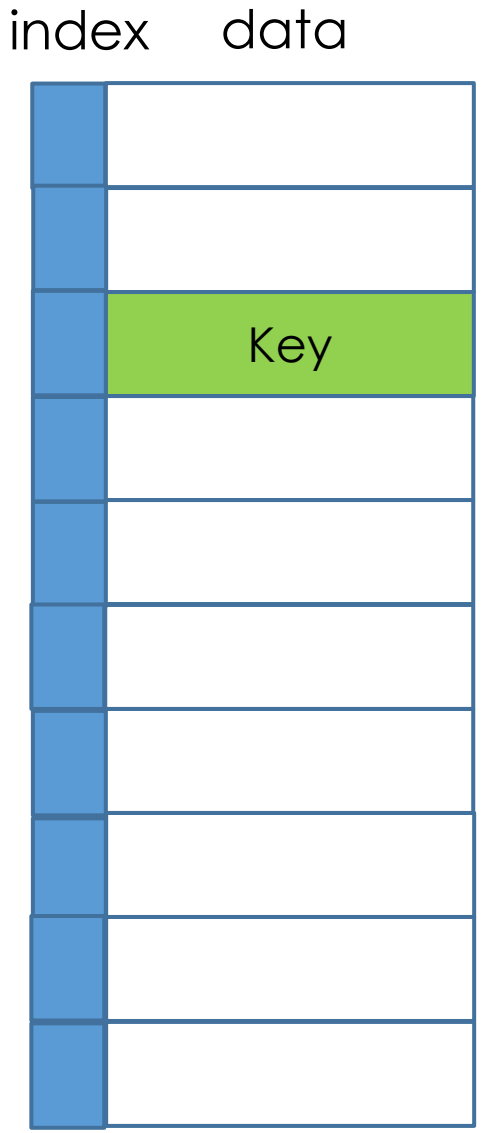
If you know where in memory the array starts, you can easily determine the address of any element using the index.

Accessing an address is an O(1) operation



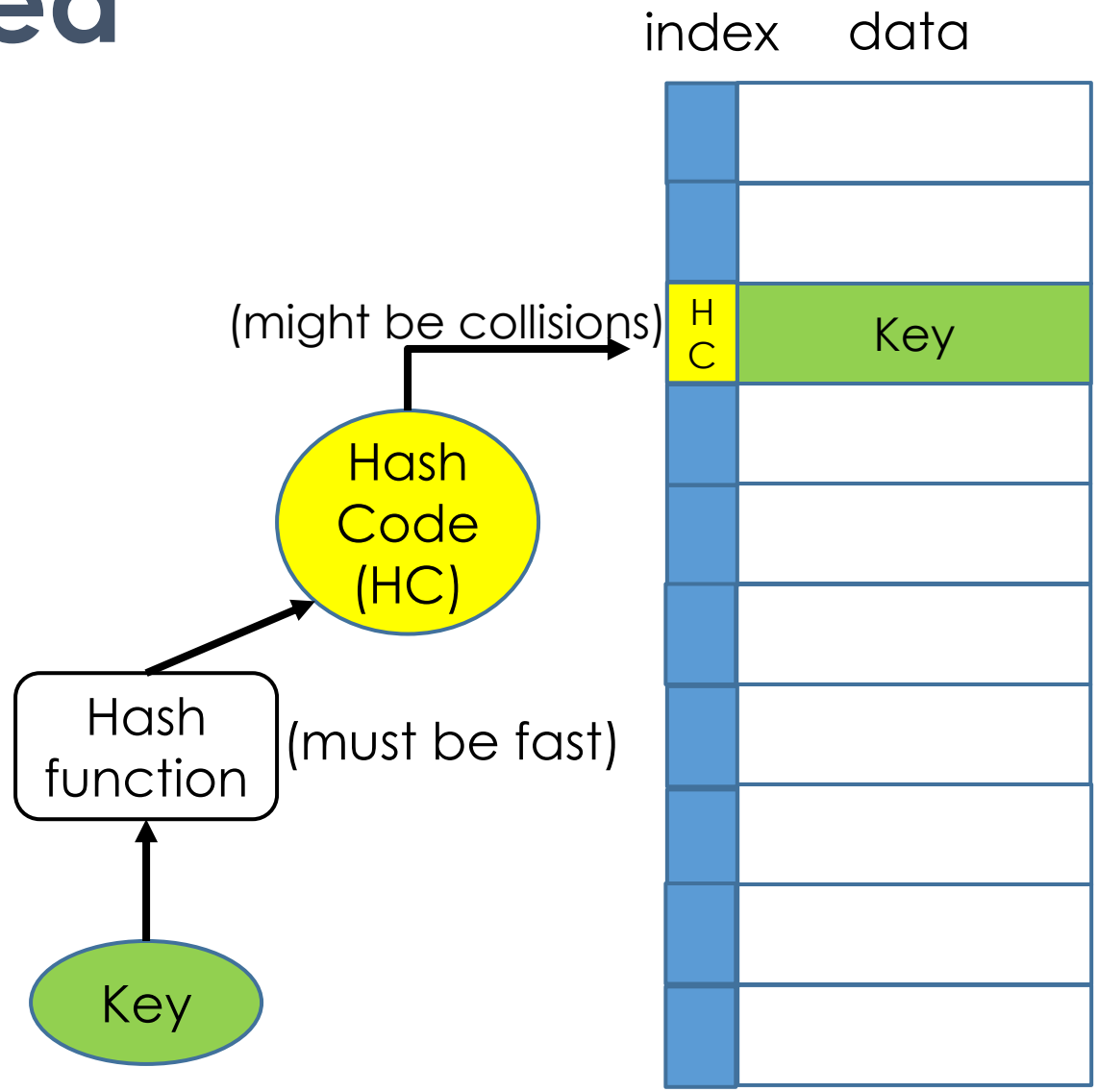
Hash Table Idea

If I want to add something into an array, could I find a way to translate it into an index?



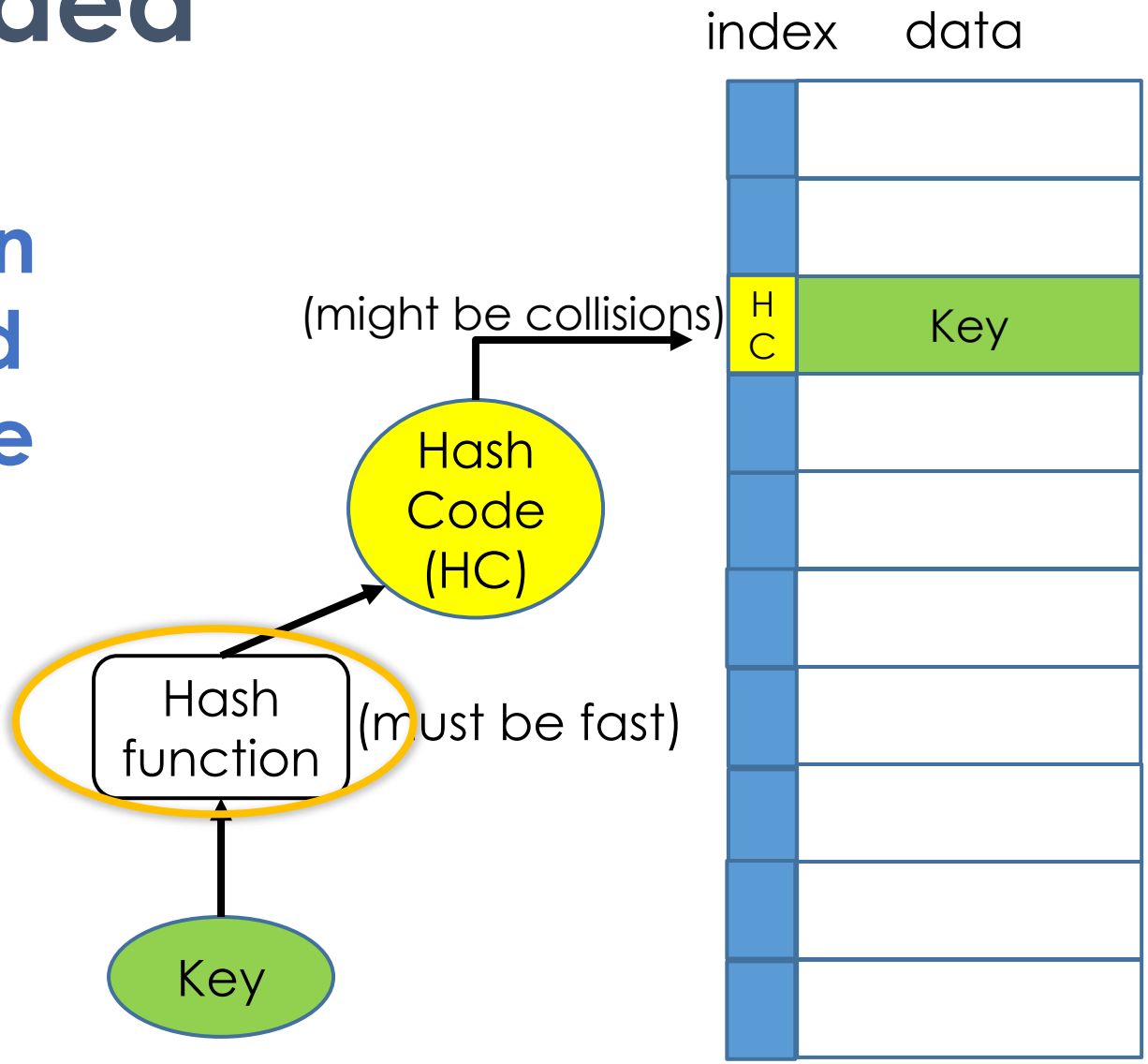
Hash Table Idea

If I want to add something into an array, could I find a way to translate it into an index?



Hash Table Idea

If I want to add something into an array, could I find a way to translate it into an index?



Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3		

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	=	3

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	=	3
11		

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	=	3
11	11 mod 5	1

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	=	3
11	11 mod 5	1

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	3 mod 5	3
11	11 mod 5	1

K mod N is a common hash function

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	3 mod 5	3
11	11 mod 5	1

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
'a'		

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
'a'	97 mod 5	2

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
"Hi"		

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
"Hi"	(72+105) mod 5	2

Which of these statements is true about a hash function?

A. If two elements are considered equal, then their hash values should also be the same
B. If two elements have the same hash value, then they should also be considered equal
C. Both A and B
D. Neither A nor B

Which of these statements is true about a hash function?

A. If two elements are considered equal, then their hash values should also be the same
B. If two elements have the same hash value, then they should also be considered equal
C. Both A and B
D. Neither A nor B

Which of these statements is true about a hash function?

A. If two elements are considered equal, then their hash values should also be the same
B. If two elements have the same hash value, then they should also be considered equal
C. Both A and B
D. Neither A nor B

Which of these statements is true about a hash function?

A. If two elements are considered equal, then their hash values should also be the same

B. If two elements have the same hash value, then they should also be considered equal

C. Both A and B

D. Neither A nor B

B is false because of "collisions"

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	3 mod 5	3
13	13 mod 5	3

So a key part of creating hash functions is trying to minimize collisions

Your array has 5 elements. How would you "hash"?

Key	Function	Hash Code
3	3 mod 5	3
13	13 mod 5	3

We'll address how to handle collisions next…