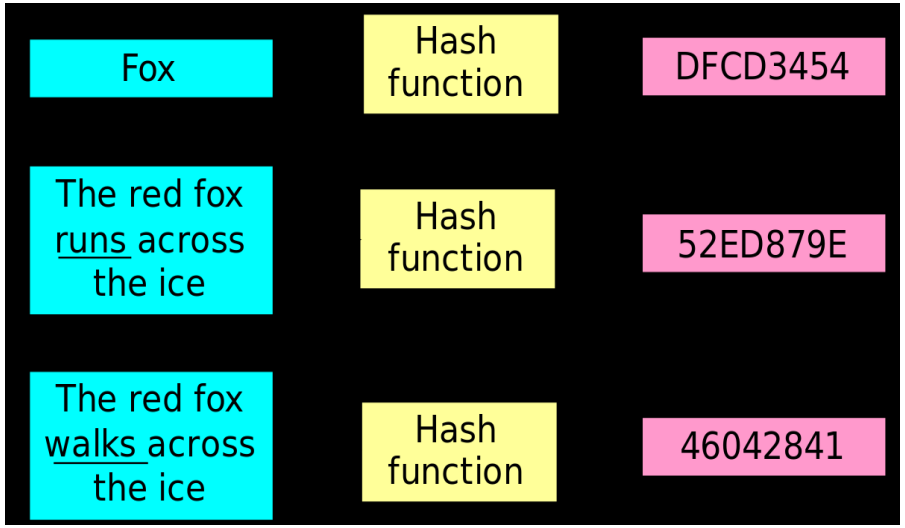
Hashing and Caching

JOSHUA OBERNESSER

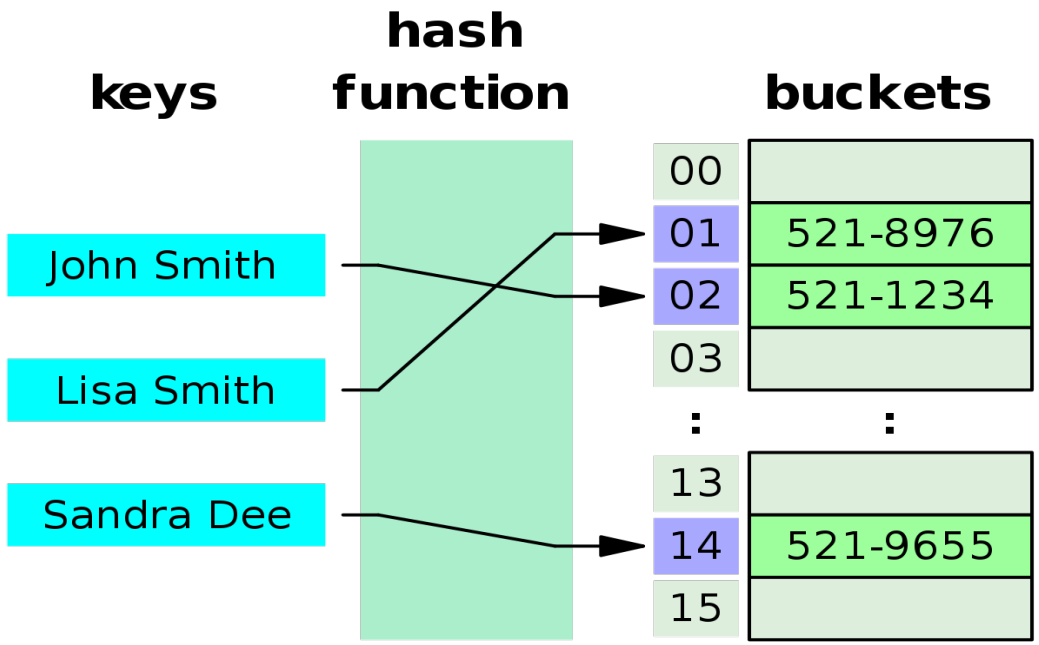
What is Hashing

- Hashing is the process of generating a value or values from a string of text using a mathematical function.
- It has two significant applications; data integrity, and memory management.



https://it.wikipedia.org/wiki/Funzione_di_hash

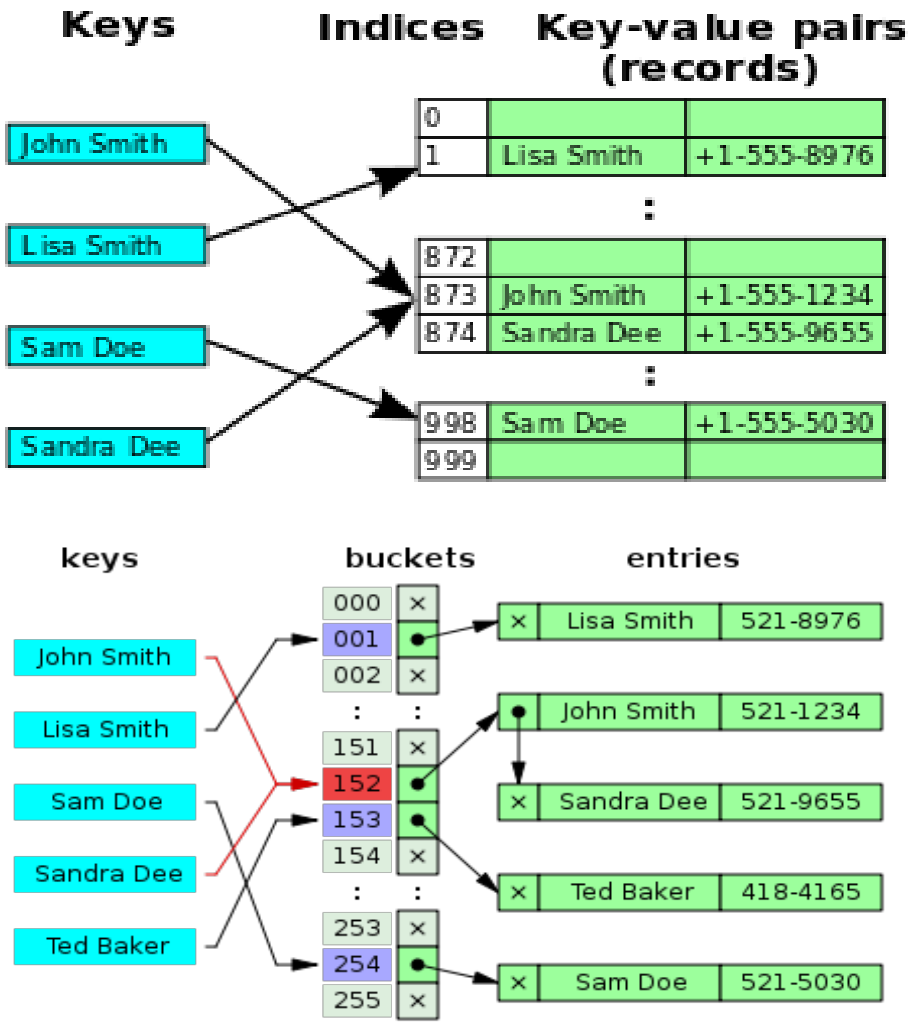
Hash Table Example



https://it.wikipedia.org/wiki/Hash_table

Probing and bucketing

- Dealing with Hash table collisions.
  - Probing is the method of looking for an empty spot elsewhere in the table.
  - Bucketing is where a set number of items can be stored at the same index.



https://en.wikipedia.org/wiki/Hash_table https://en.wikipedia.org/wiki/Linear_probing

Hash table runtimes

- Hash tables have a worst case run time of O(n) However that is not usually their runtime.
- Hash tables actually run on average much better than the worst case scenario. They are often optimal and require limited iterations.
- For insertion, deletion, and searching, the average run time is O(1).

Hashing functions

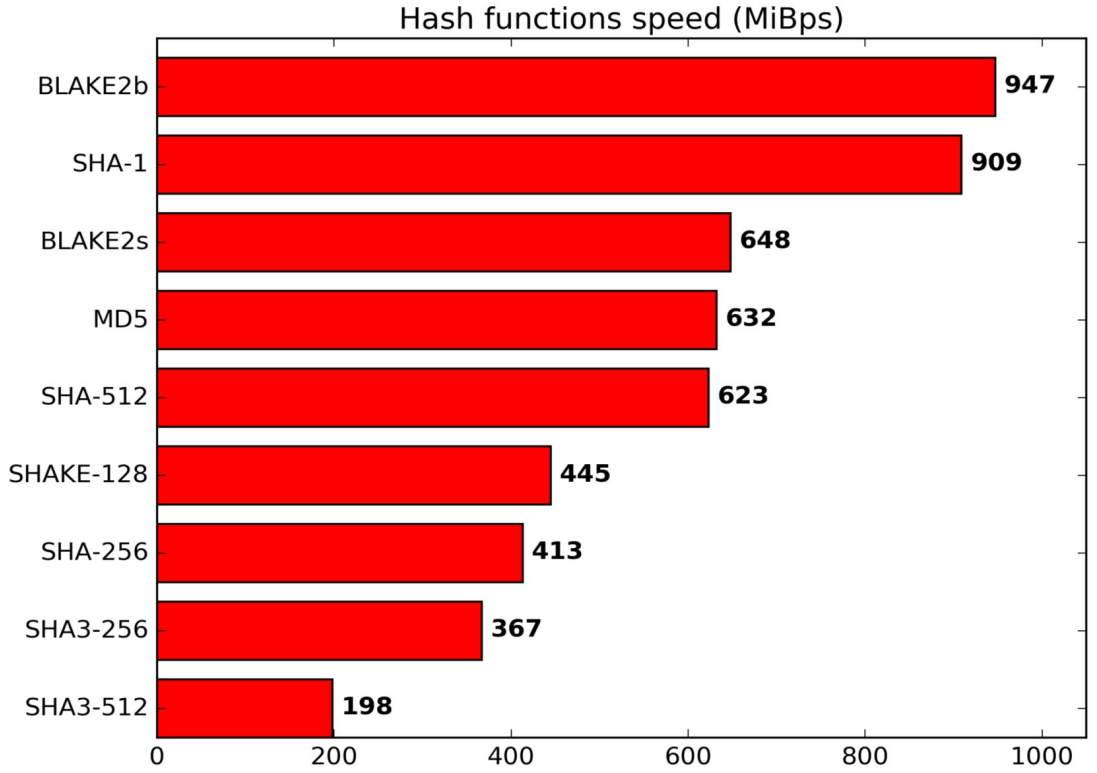
- There are many different hash functions, but there is some key criteria for what makes a good or bad one.(2)
  - 1. It should be very efficient to compute.
  - 2. The function needs to be determinist; you get the same output for the same input each time it is run.
  - 3. For encryption purposes, the output should not give any indication of the input.
  - 4. Finally it should avoid collisions, no two inputs should produce the same output. (Or it should be practically impossible.)

Hashing Functions cont

- The message direct family, most common is MD5(very vulnerable)
- SHA Family(Secure Hash Algorithm), currently up to SHA 3, but SHA 2 is the most common.
- The BLAKE Family, most recent BLAKE3 announced January 9, 2020

https://decrypthash.ru/en/review-of-the-md5-algorithm/, https://cyberhackersz.blogspot.com/2018/01/use-sha-256-hash-to-verify-your.html

https://blake2.net/



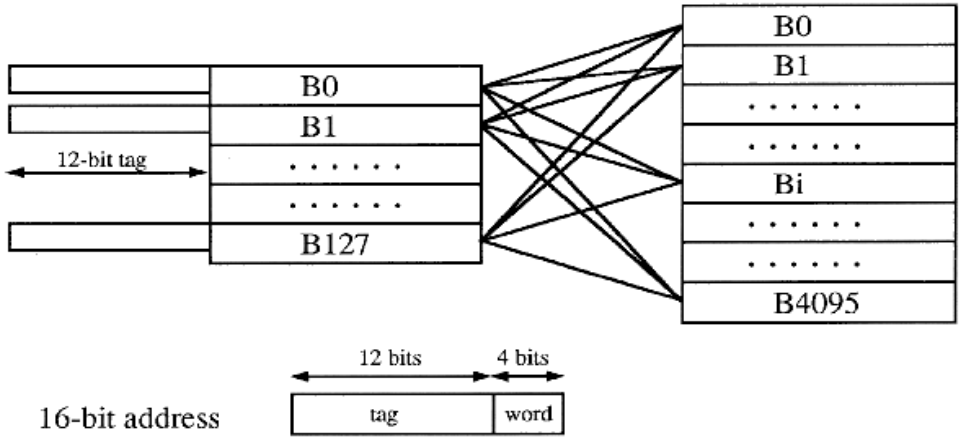
Caching

- Caching is the process of storing data so that it can accessed or used faster.
- Data is stored with more focus on locality, in reference to other data that has been called recently.
- It has many applications and uses, the most prominent one many people know of is browser caching.

https://askleo.com/whats_a_browser_cache_how_do_i_clear_it_and_why_would_i_want_to/

Cache Placement Policies

- Cached memory can be organized and utilized several different ways.
  - Fully Associative (A single cache set with multiple lines)
  - Direct Mapping (multiple cache sets each with one cache line)
  - Set Associative (A middle ground between Direct and Full)

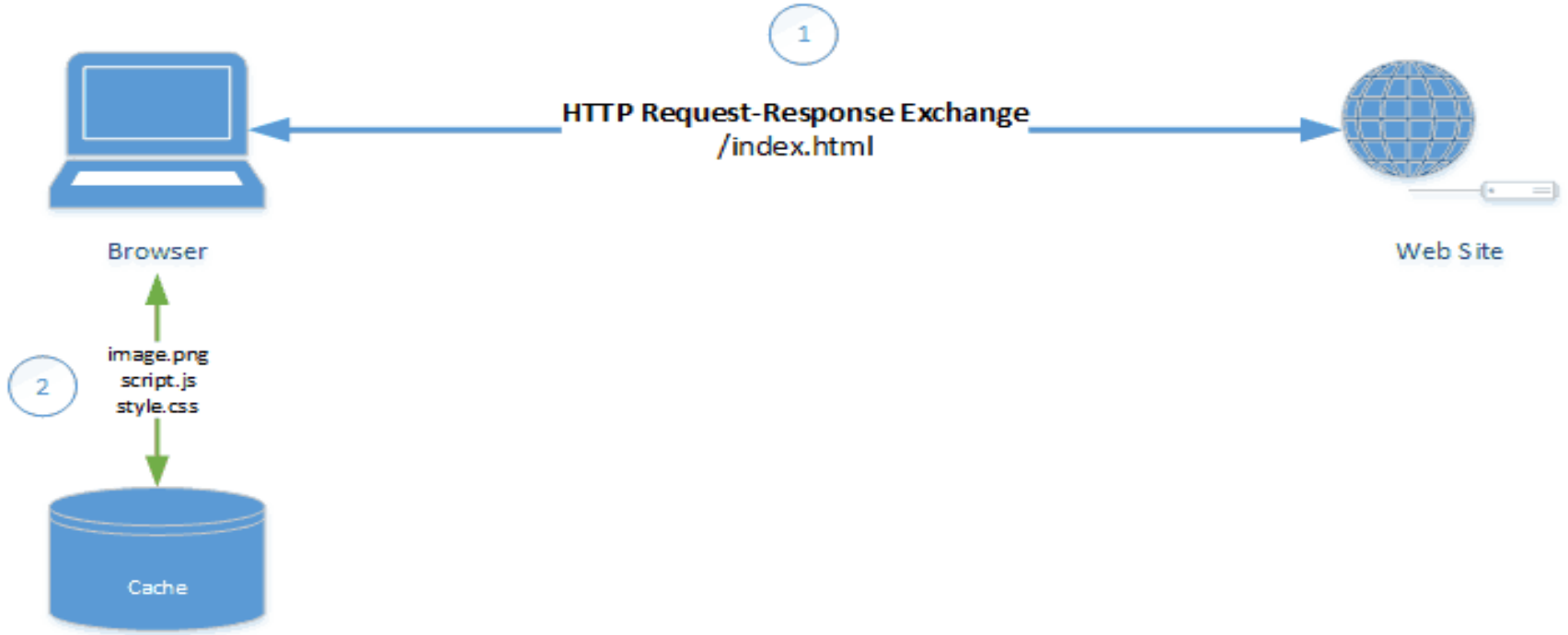


http://fourier.eng.hmc.edu/e85_old/lectures/memory/node4.html

Cache Memory Replacement Policy

- When blocks in the Cached memory needs to be replaced, there are several possible options that can be used.
  - Optimal Replacement(Replace the block which is no longer needed in the future.)
  - First in First out(FIFO)
  - Least recently used (LRU)
  - Random Selection

Browser caching



https://underconstructionpage.com/website-performance-optimization/

Thank you

Longwood Classes

- CMSC 162
- CMSC 201
- CMSC 242
- CMSC 360

- Wikipedia for all the excellent images.

Sources, any questions?

- https://it.wikipedia.org/wiki/Funzione_di_hash (image)
- https://komodoplatform.com/cryptographic-hash-function/
- https://en.wikipedia.org/wiki/Linear_probing https://en.wikipedia.org/wiki/Hash_table
- https://decrypthash.ru/en/review-of-the-md5-algorithm/
  https://cyberhackersz.blogspot.com/2018/01/use-sha-256-hash-to-verify-your.html
- https://tools.ietf.org/html/rfc7693 .https://blake2.net/
- https://en.wikipedia.org/wiki/Cache_placement_policies#Set-associative_cache
  .https://medium.com/breaktheloop/direct-mapping-map-cache-and-main-memory-d5e4c1cbf73e
- https://tiptopsecurity.com/what-is-cryptographic-hashing-md5-sha-and-more/
  https://www.toolsqa.com/data-structures/hash-tables-in-data-structures/
- https://pediaa.com/difference-between-cache-memory-and-virtual-memory/
  https://askleo.com/whats_a_browser_cache_how_do_i_clear_it_and_why_would_i_want_to/
- https://underconstructionpage.com/website-performance-optimization/
  http://fourier.eng.hmc.edu/e85_old/lectures/memory/node4.html