MD5 and SHA-1

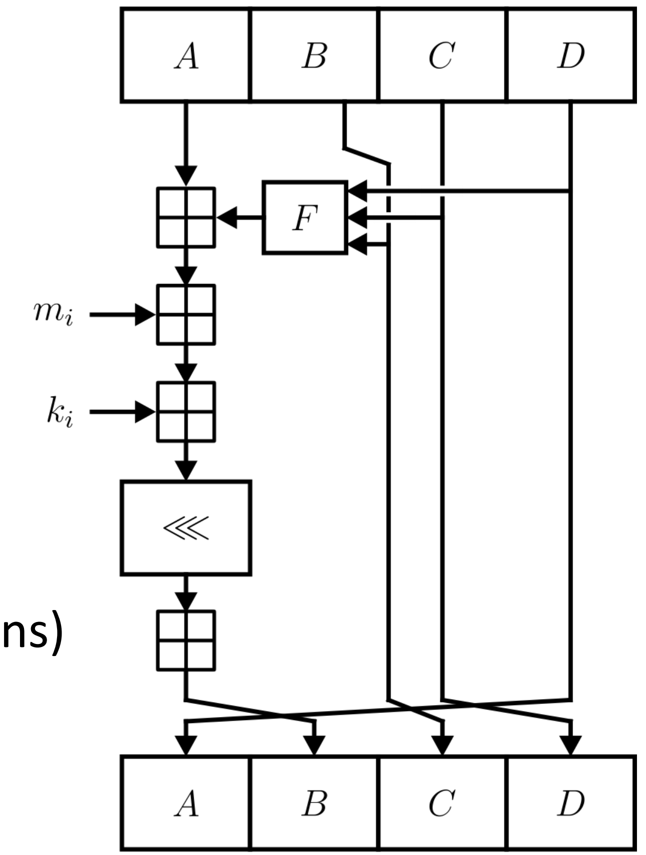
Hyoungshick Kim
Department of Software
College of Software
Sungkyunkwan University

MD5

- MD5 (Ron Rivest, 1991): still widely used, has 128-bit block.
  - So finding a collision would take about 2^{64} effort (via birthday attacks)
- Flaws found by Dobbertin and others; collision existence in 2004; fake SSL certificates in 2005 (two public keys with same MD5 hash); now collision attack takes only a minute

MD5 hash function

- Output = 128-bit
- Operates on 512-bit blocks (sixteen 32-bit words)
- Padding
  - Append ‘1’ to the end
  - Append zeros up to (512-64) bits
  - Append the message length (64-bit)
- 128-bit state divided into 4x32-bit words
- Four rounds (each of which comprises 16 operations)
  - Non-linear function F
  - modular addition
  - left shift

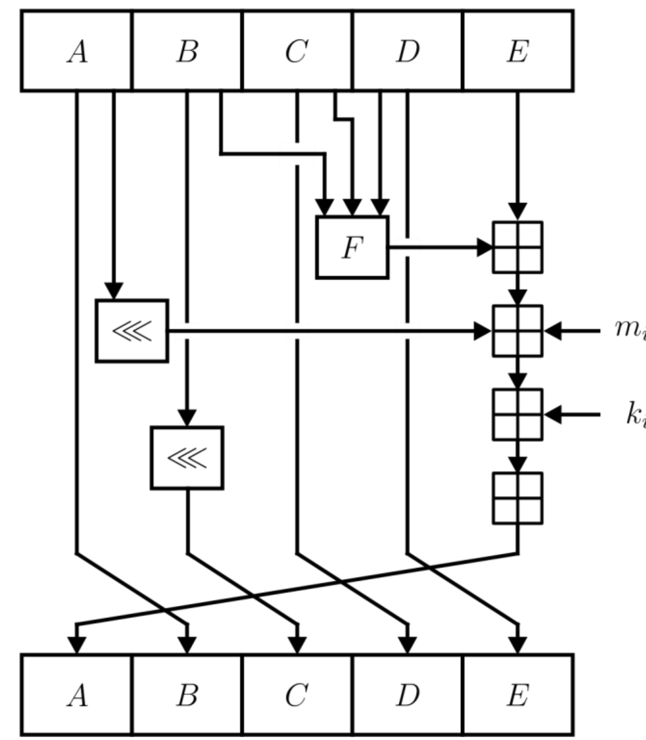


SHA-family hash functions

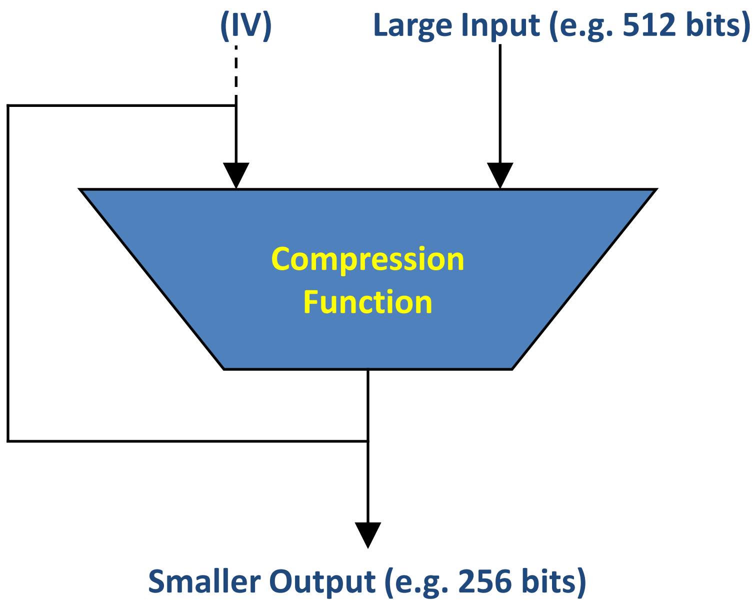
- Secure Hash Algorithm (SHA)
- SHA-0 (1993)
  - Weaknesses found but were not published
- SHA-1 (1995)
- SHA-2 (2001)
  - SHA-256, SHA-384, and SHA-512
  - SHA-224 (2004)
- SHA-3 (2012)
  - Keccak

SHA-1

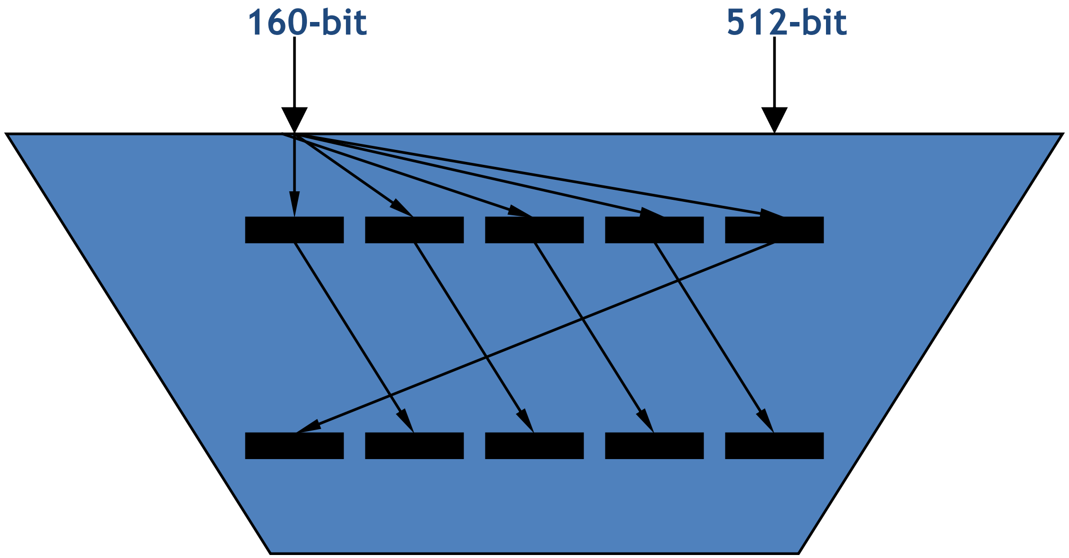
- 160-bit (based on MD4)
- Some similarity with MD5, but slower
- Relies on “linear recurrence” to stretch 16 words into 80 words
- Collisions can be found in 2^{80} steps (via birthday attacks)



Merkle-Damgård Construction

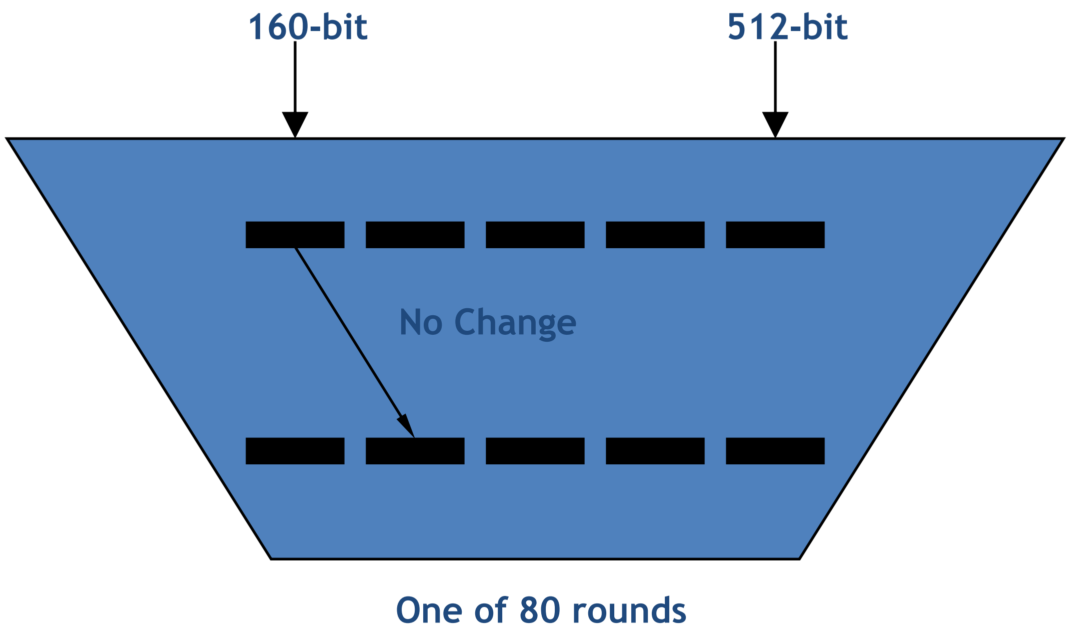


Merkle-Damgård Construction for SHA-1

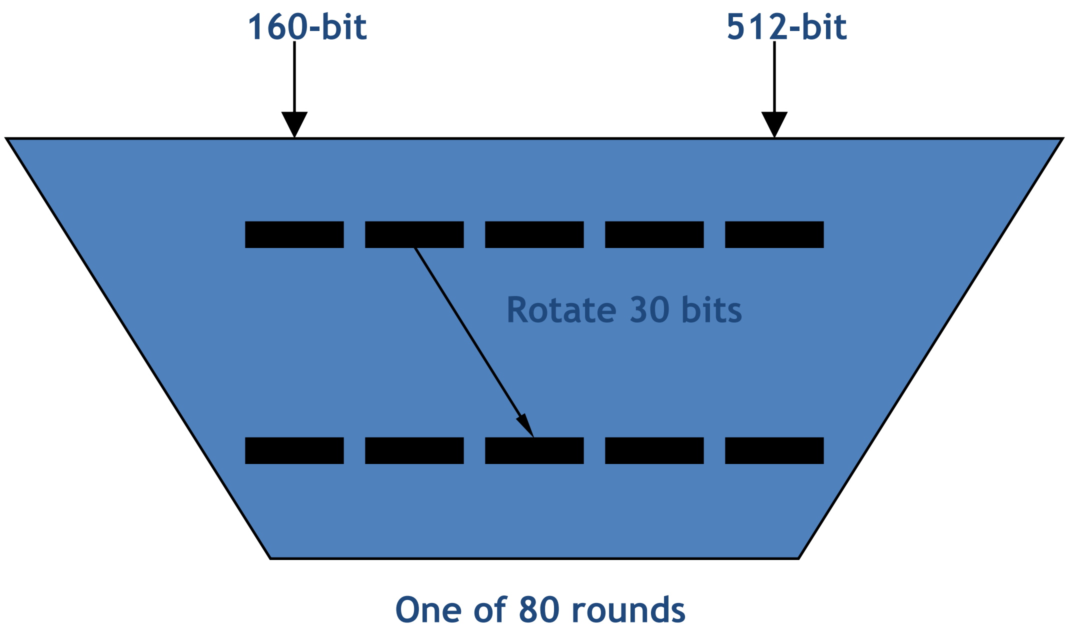


One of 80 rounds

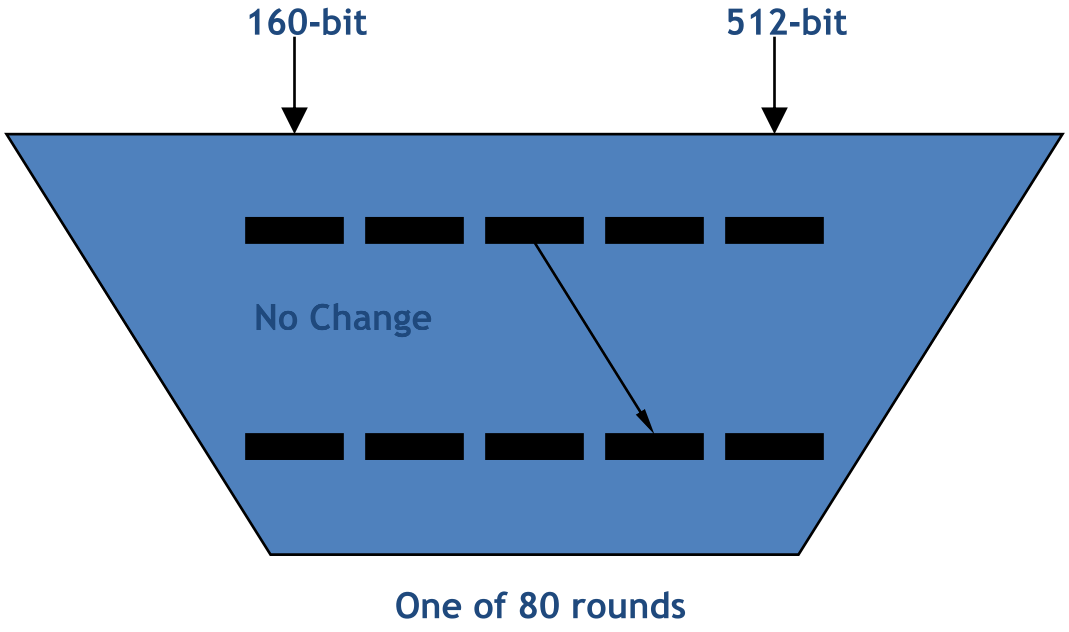
Merkle-Damgård Construction for SHA-1



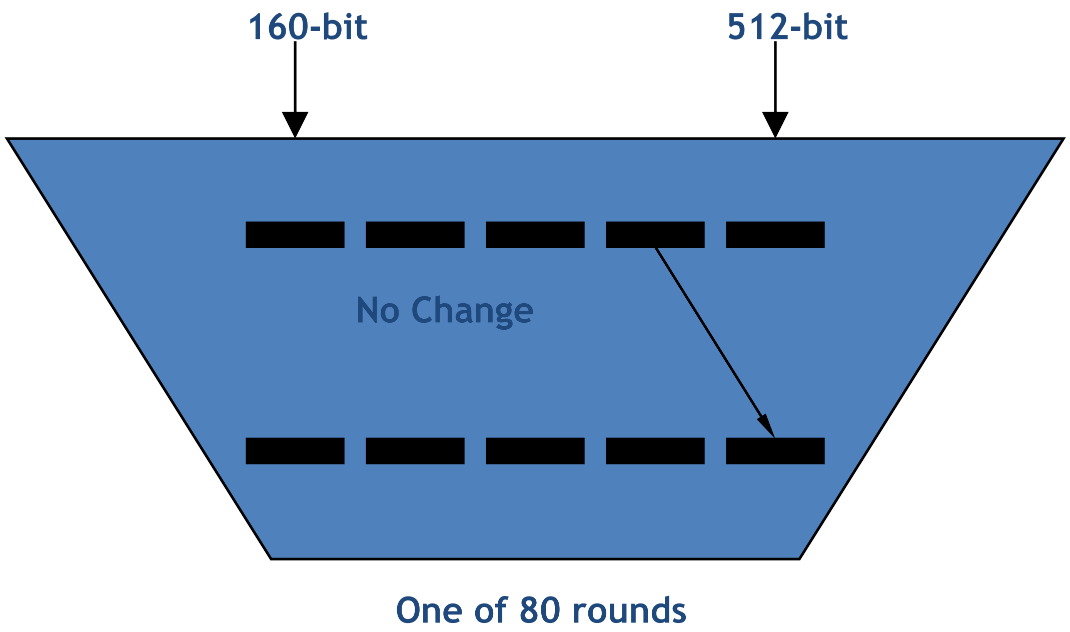
Merkle-Damgård Construction for SHA-1



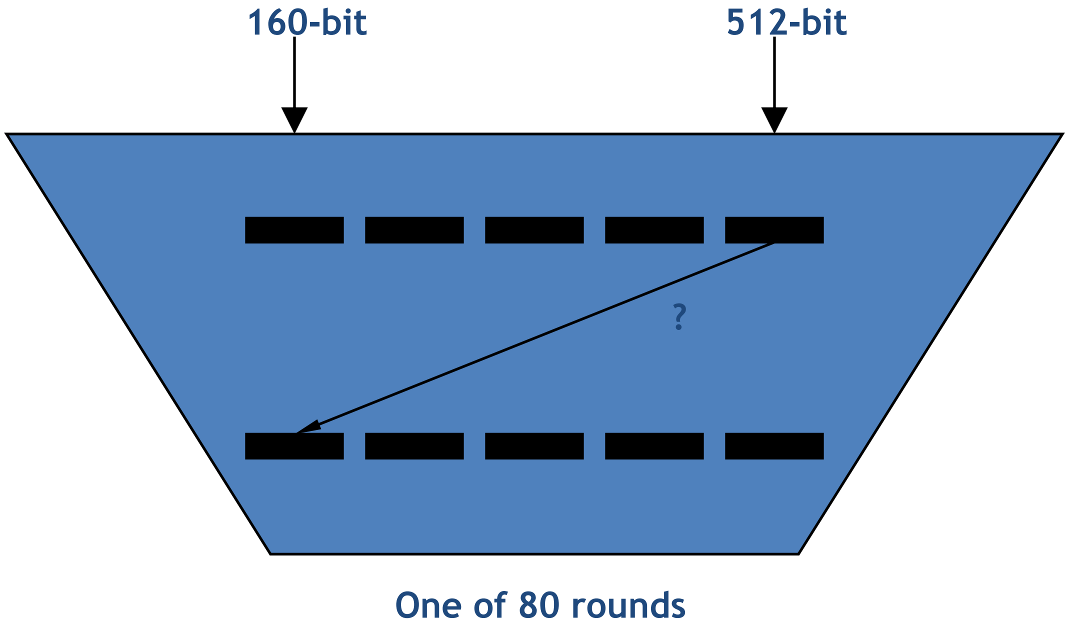
Merkle-Damgård Construction for SHA-1



Merkle-Damgård Construction for SHA-1



Merkle-Damgård Construction for SHA-1

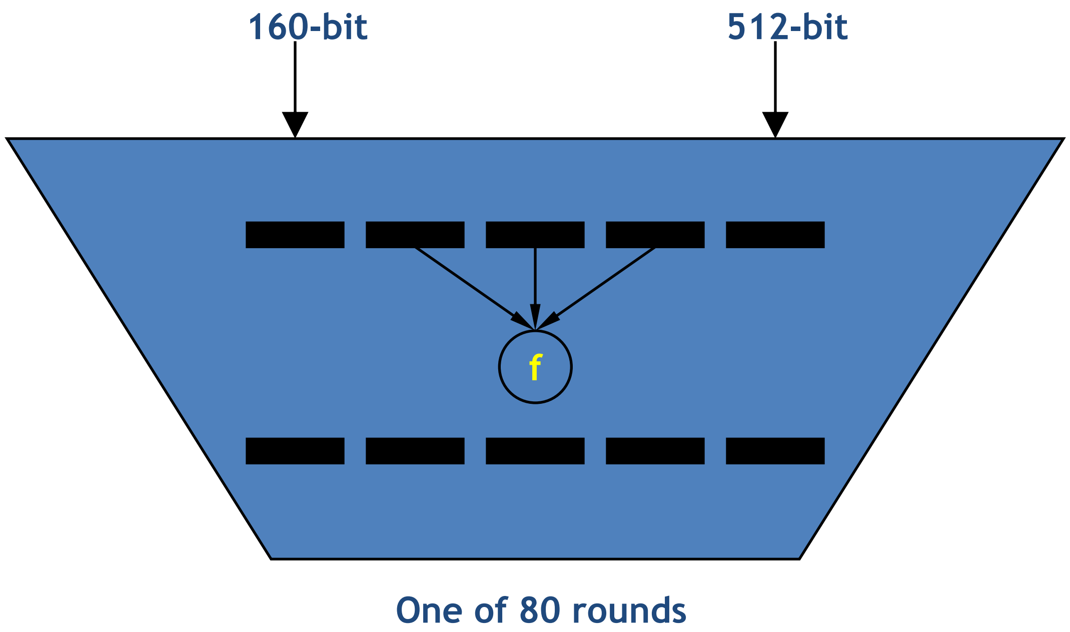


Merkle-Damgård Construction for SHA-1

What's in the final 32-bit transform?

- Take the rightmost word.
- Add in the leftmost word rotated 5 bits.
- Add in a round-dependent function f of the middle three words.

Merkle-Damgård Construction for SHA-1

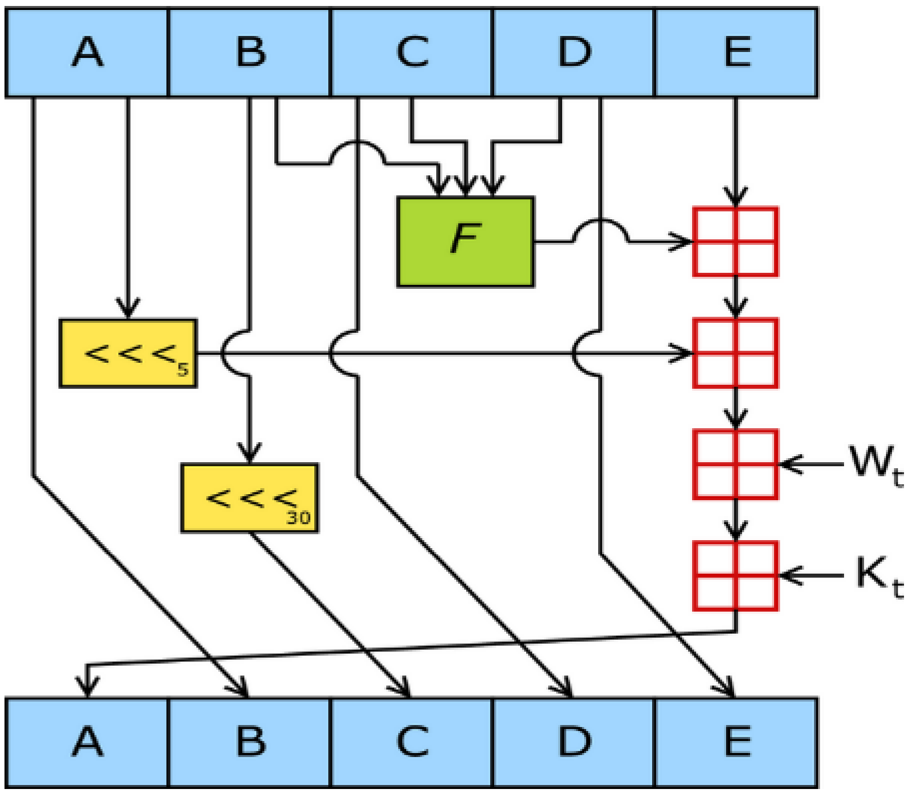


Merkle-Damgård Construction for SHA-1

What's in the final 32-bit transform?

- Take the rightmost word.
- Add in the leftmost word rotated 5 bits.
- Add in a round-dependent function f of the middle three words.
- Add in a round-dependent constant.
- Add in a portion of the 512-bit message.

Round Function in SHA-1



Picture from
Wikipedia

Non-linear function F in SHA-1

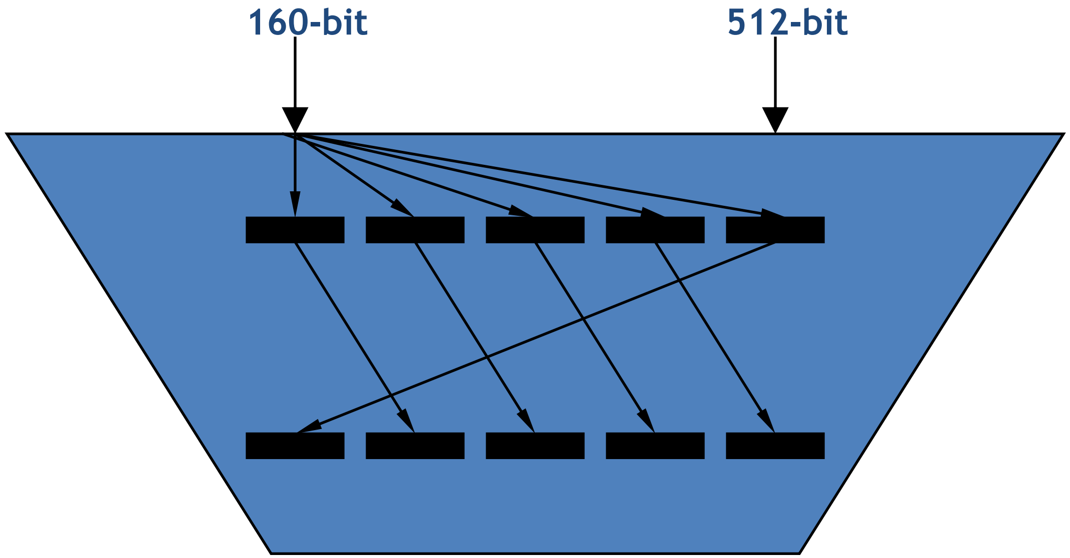
Depending on the round, the “non-linear” function f is one of the following.

$$f(X,Y,Z) = (X \wedge Y) \vee ((\neg X) \wedge Z)$$

$$f(X,Y,Z) = (X \wedge Y) \vee (X \wedge Z) \vee (Y \wedge Z)$$

$$f(X,Y,Z) = X \oplus Y \oplus Z$$

Merkle-Damgård Construction for SHA-1



One of 80 rounds

Attacks of SHA

- At Crypto 2005, a 2^{69} collision attack on SHA was published by Xiaoyun Wang et al., “Finding Collisions in the Full SHA-1” (birthday attack in 2^{80})
- As an interim measure, people are moving to SHA-2 (256-bit or 512-bit output lengths, no known significant weaknesses)

Google, CWI announce SHAttered attack against SHA-1

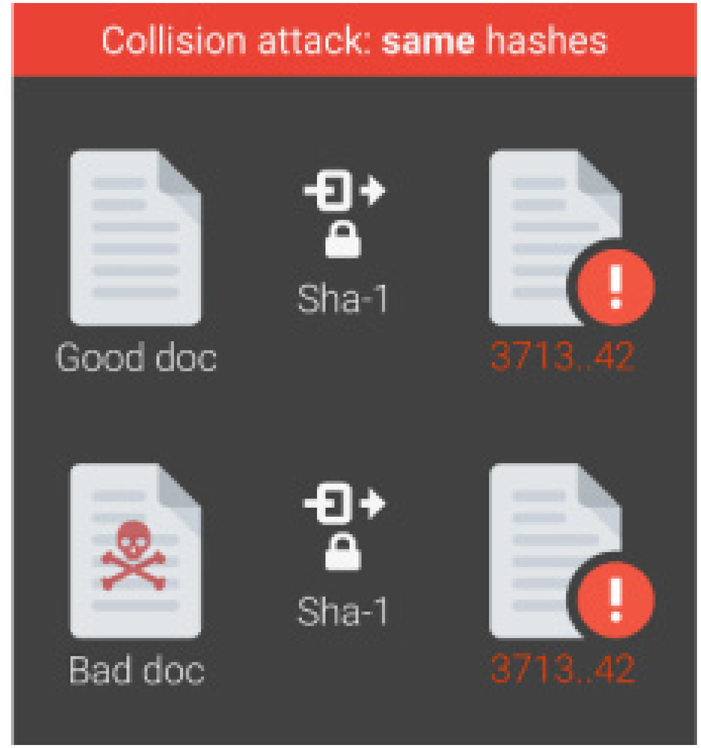
Published on 24th February 2017 by Gareth Halfacree

News | 0 Comments

The still-popular cryptographic hash function SHA-1 has been officially broken, two decades after it was introduced, by researchers working at Google and Centrum Wiskunde & Informatica (CWI) Amsterdam.

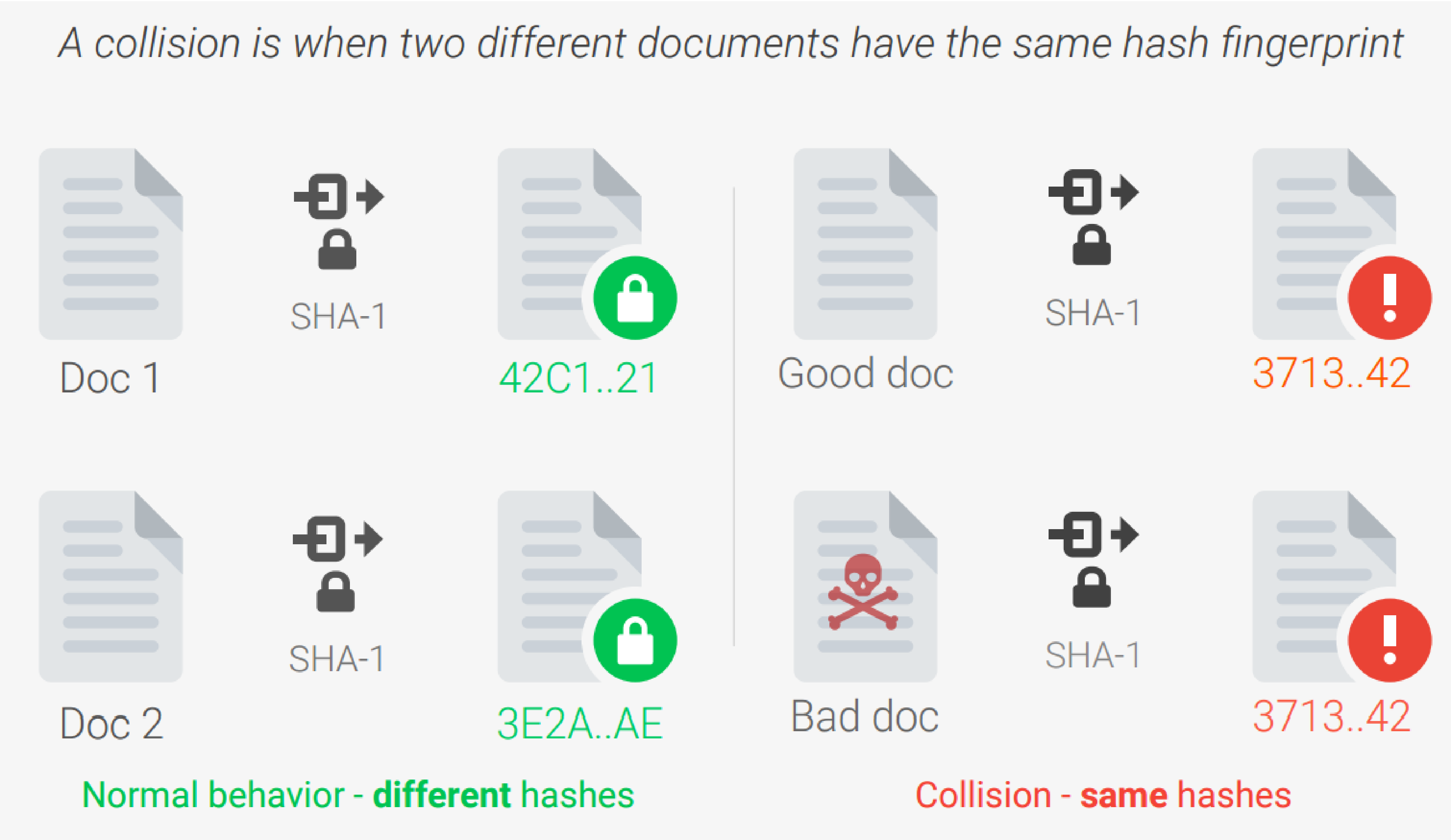
Designed by the US National Security Agency and adopted as a formal US Federal Information Processing Standard by the National Institute of Standards and Technology (NIST), SHA-1 (Secure Hash Algorithm 1) is a one-way hash function: given data, it spits out a fixed-length message digest which becomes invalid if the data is modified in any way. It's used for everything from checking that downloaded data hasn't been corrupted and ensuring that files uploaded to shared hosting have unique filenames to validating passwords against a database without having to store the password in plain text, digital signing, and the TLS cryptography standard.

Sadly, it has also been proven flawed: A team of researchers from Google and CWI Amsterdam have



The SHA-1 hash function has been officially broken, following an effort involving 6,500 years of CPU time and 110 years of GPU time.

http://shattered.it/ - information about practical SHA-1 collision


A collision is when two different documents have the same hash fingerprint
Doc 1
SHA-1
42C1..21
Doc 2
SHA-1
3E2A..AE
Good doc
SHA-1
3713..42
Bad doc
SHA-1
3713..42
Normal behavior - different hashes
Collision - same hashes

Attack complexity

9,223,372,036,854,775,808

SHA-1 compressions performed

Shattered compared to other collision attacks

MD5
1 smartphone
30 sec

SHA-1 Shattered
110 GPU
1 year

SHA-1 Bruteforce
12,000,000 GPU
1 year

Questions?

Sungkyunkwan
Campus
Social
Institute
Building
Korean
Sciences
University
Seoul
Suwon
School
established
university
campus
Graduate
English
constructed
Natural
edit