Hashing

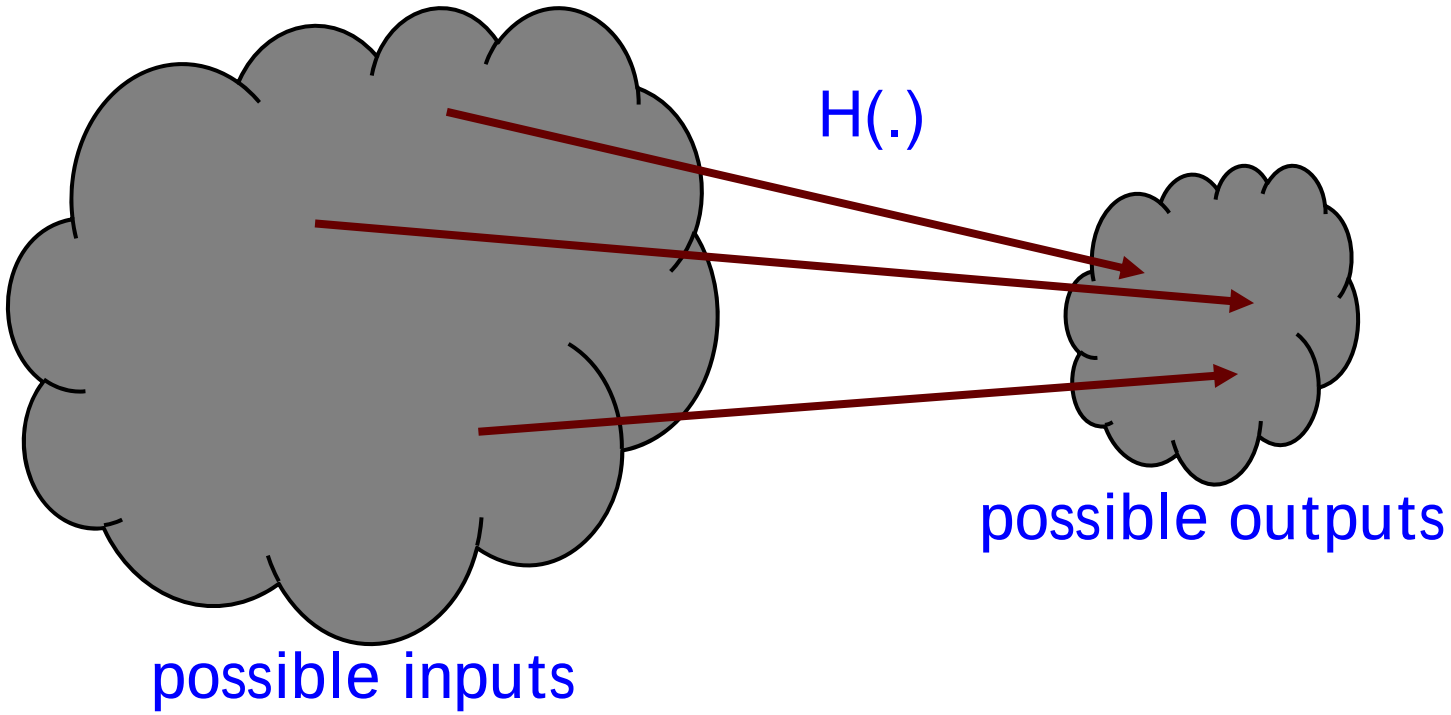
Hash function:

mathematical function

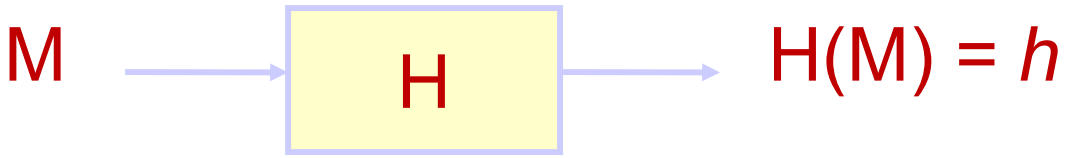
takes any string as input

fixed-size output (we'll use 256 bits)

efficiently computable (say, O(n))



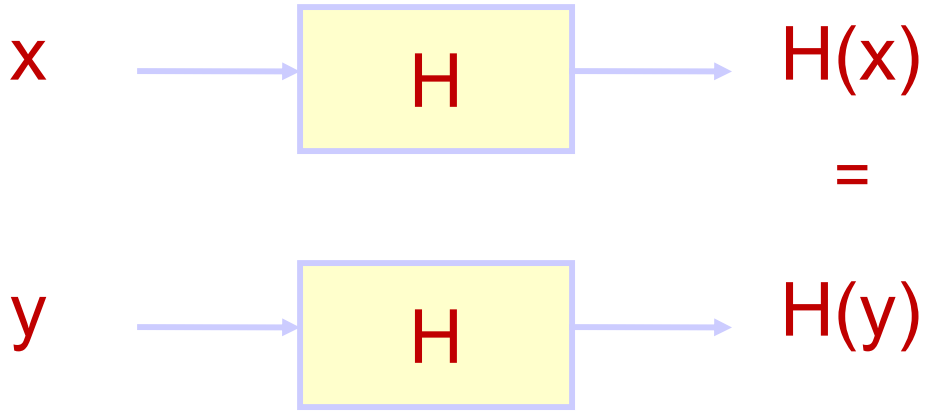
One-Way Hash Functions



Example

- **M = "Elvis"**
- **H(M) = ("E" + "L" + "V" + "I" + "S") mod 26**
- **H(M) = (5 + 12 + 22 + 9 + 19) mod 26**
- **H(M) = 67 mod 26**
- **H(M) = 15**

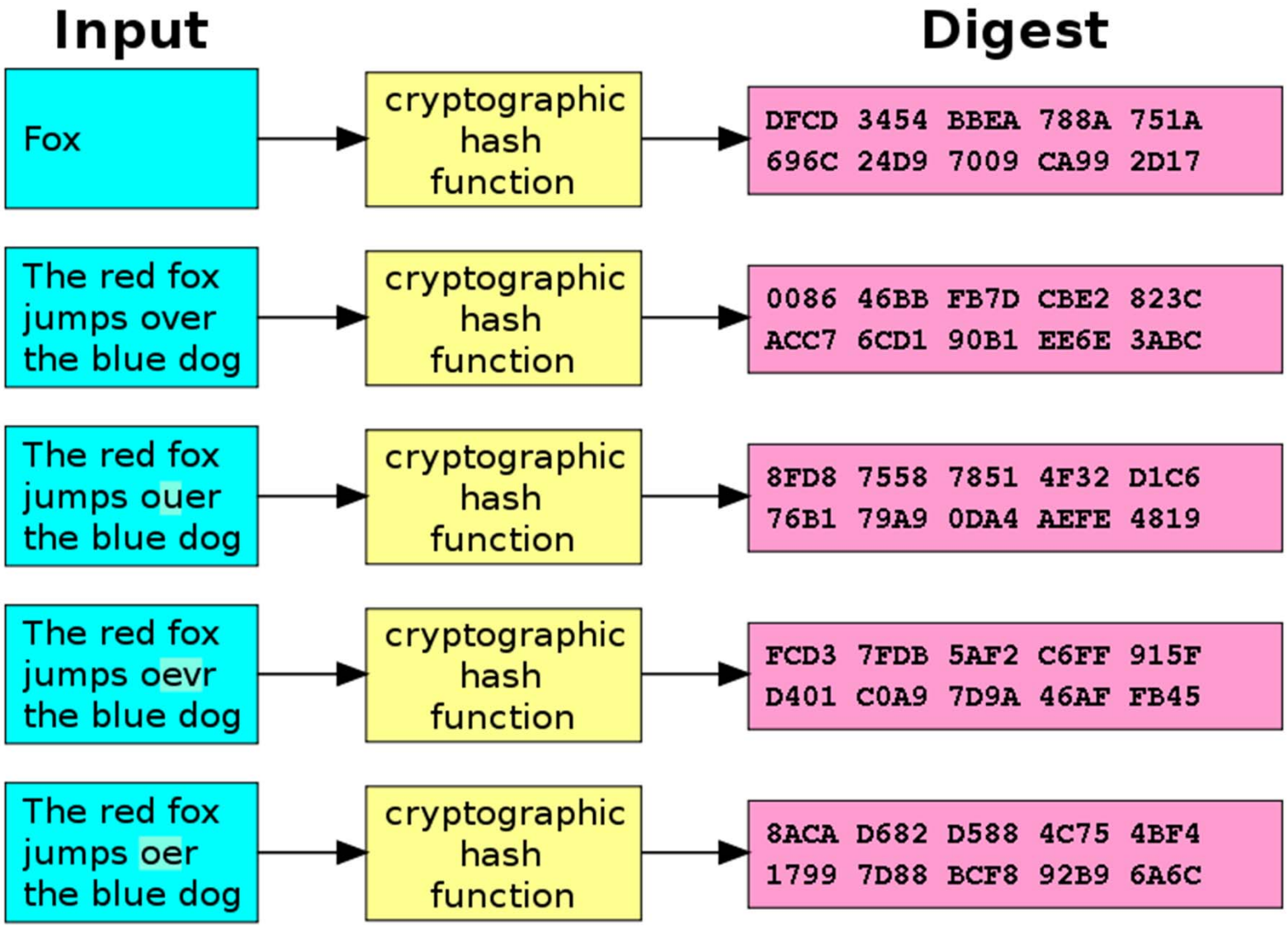
Collision



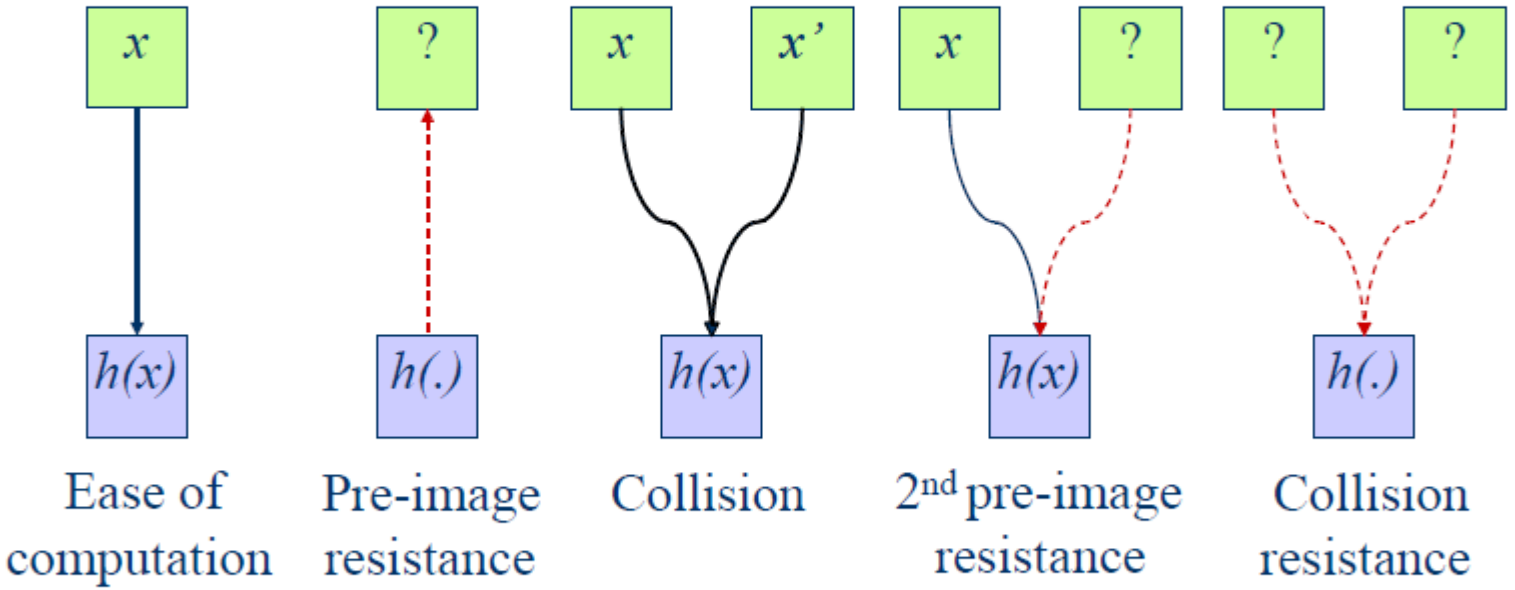
Example

- **x = "Viva"**
- **Y = "Vegas"**
- **H(x) = H(y) = 2**

avalanche effect



HASH PROPERTIES



- Collision resistance implies 2nd pre-image resistance
- Collision resistance does not imply pre-image resistance

Authentication

Basics

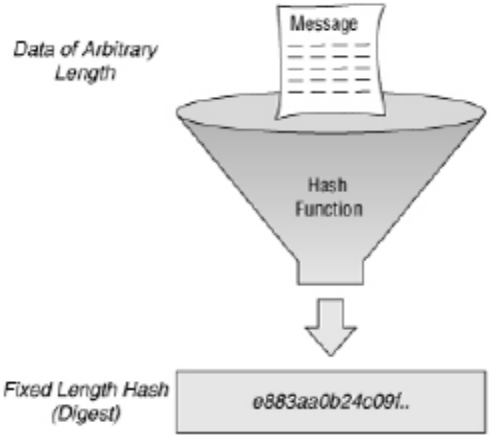
- **Authentication is the process of validating the identity of a user (Authenticity) or the integrity of a piece of data (Integrity).**

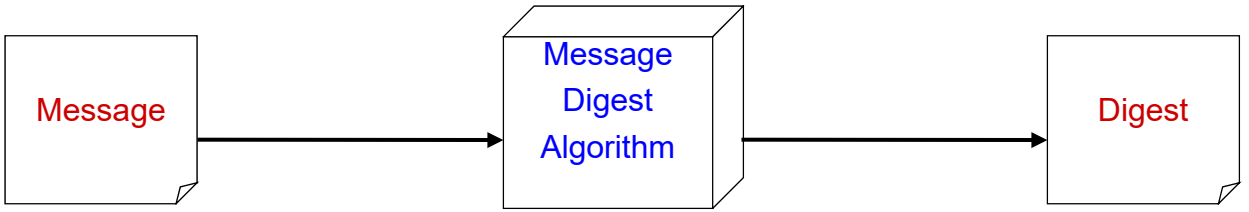
- **Technologies that provide authentication**
  - **Message Digests (MD)**
  - **Message Authentication Codes (MAC)**
  - **Digital Signatures**
  - **Others….**

Authentication

Message Digests (MD)

- **A message digest is a fingerprint for a document.**
- **Purpose of the message digest is to provide proof that data has not altered - Integrity.**
- **Process of generating a message digest from data is called** hashing

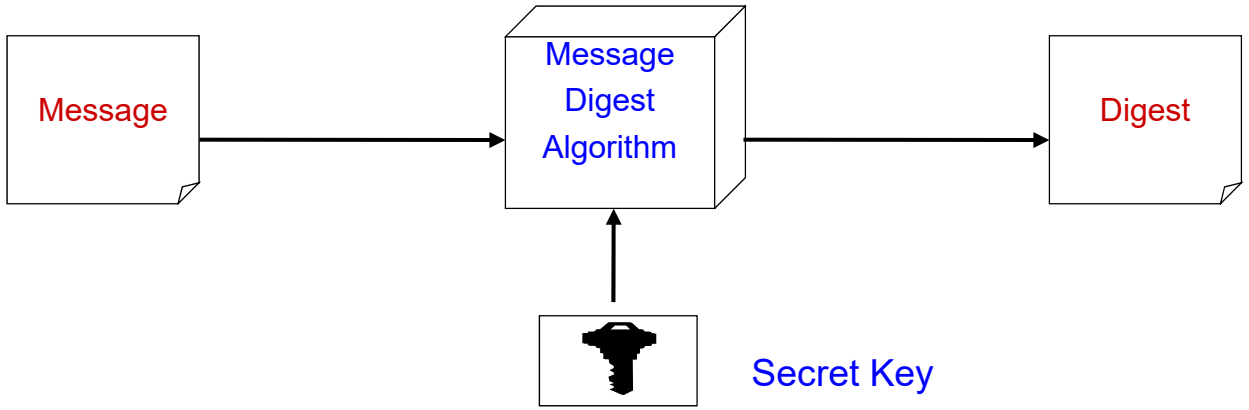




Authentication

Message Authentication Codes (MAC)

- **A message digest created with a key**
- **Creates security by requiring a secret key to be possesses by both parties in order to retrieve the message**
- **A MAC is a short string used to verify the message integrity and authentication**



COMMONLY USED HASH FUNCTIONS

- **MD** (Message Digest)
  - MD5
    - Max message < 2^{64}
    - Output: 128-bit
- **SHA** (Secure Hash Algorithm)
  - SHA-1
    - Max message < 2^{64}
    - Output: 160-bit
  - SHA-2
    - Max message < 2^{128}
    - Max output: 512-bit
  - SHA-3
    - Max message: Unlimited
    - Max output: 512-bit

Authentication

Message Digests (MD)

- **Standards:**

  MD5 : **128 bit hashing algorithm by Ron Rivest of RSA**

  - **Broken since 2004 by Xiaoyun Wang**
  - **Collisions can be constructed in seconds on a laptop**

  SHA & SHA-1 : **160 bit hashing algorithm developed by NIST**

  - **Considered insecure**
  - **Practically broken since 2005 by Xiaoyun Wang**

Do not use them in your security products!

Authentication

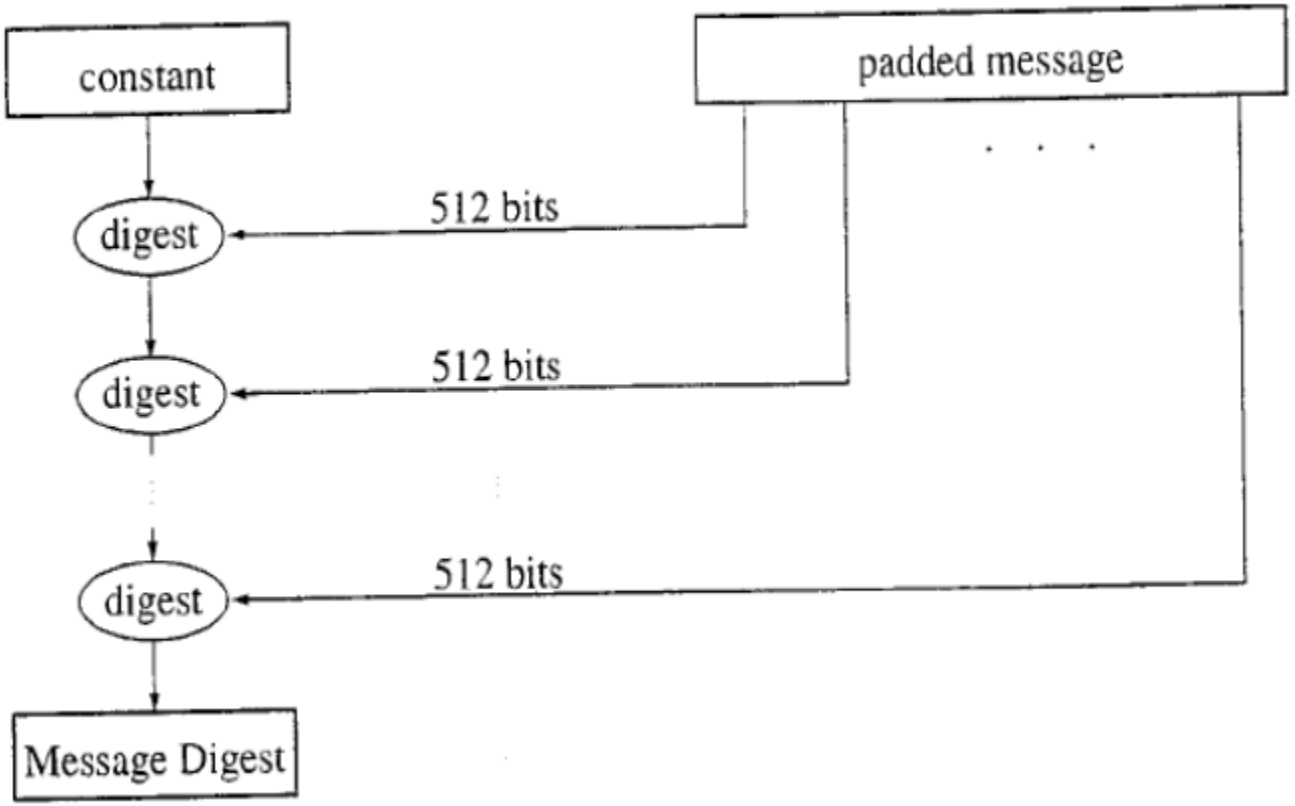
Message Digests (MD)

- **SHA-2 family: 224, 256, 384 or 512 bits**
- **SHA-3 family output can be arbitrary size**
  - **NIST started a competition for SHA-3 in 2007.**
  - **NIST's original concern was that SHA-2 would soon be broken, although in fact SHA-2 is still fine.**
- **SHA-3 (Secure Hash Algorithm 3) is the latest member of the Secure Hash Algorithm family of standards, released by NIST on August 5, 2015.**

MD5: Message Digest Version 5

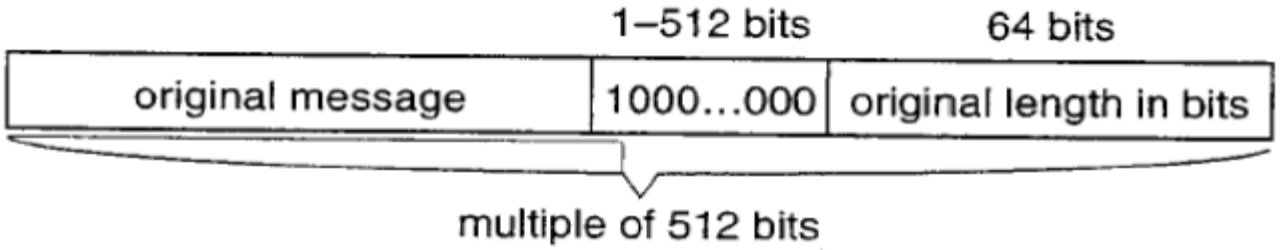
- Author: R. Rivest, 1992
- 128-bit hash
  - based on earlier, weaker MD4 (1990)
- **Collision resistance (B-day attack resistance)**
  - only 64-bit
- Output size not long enough today (due to various attacks)

MD5 Overview



Similar for MD4/MD5/SHA-1

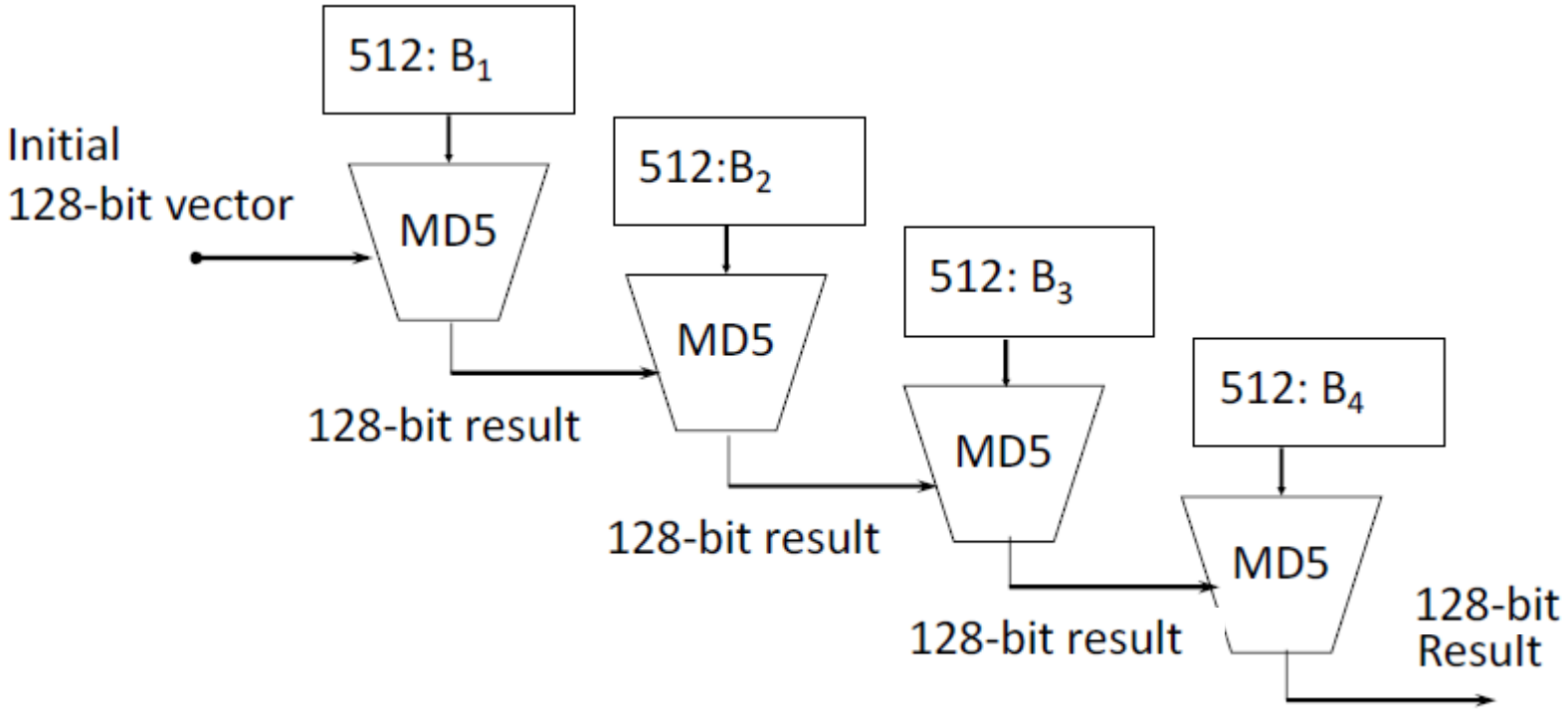
MD5: Padding



- Given original message M, add padding bits "100…" such that resulting length is 64 bits less than a multiple of 512 bits.
- Append *original length in bits* to the padded message
- Final message chopped into 512-bit blocks

MD5: Blocks

- As many stages as the number of 512-bit blocks in the final padded message



Secure Hash Algorithm (SHA)

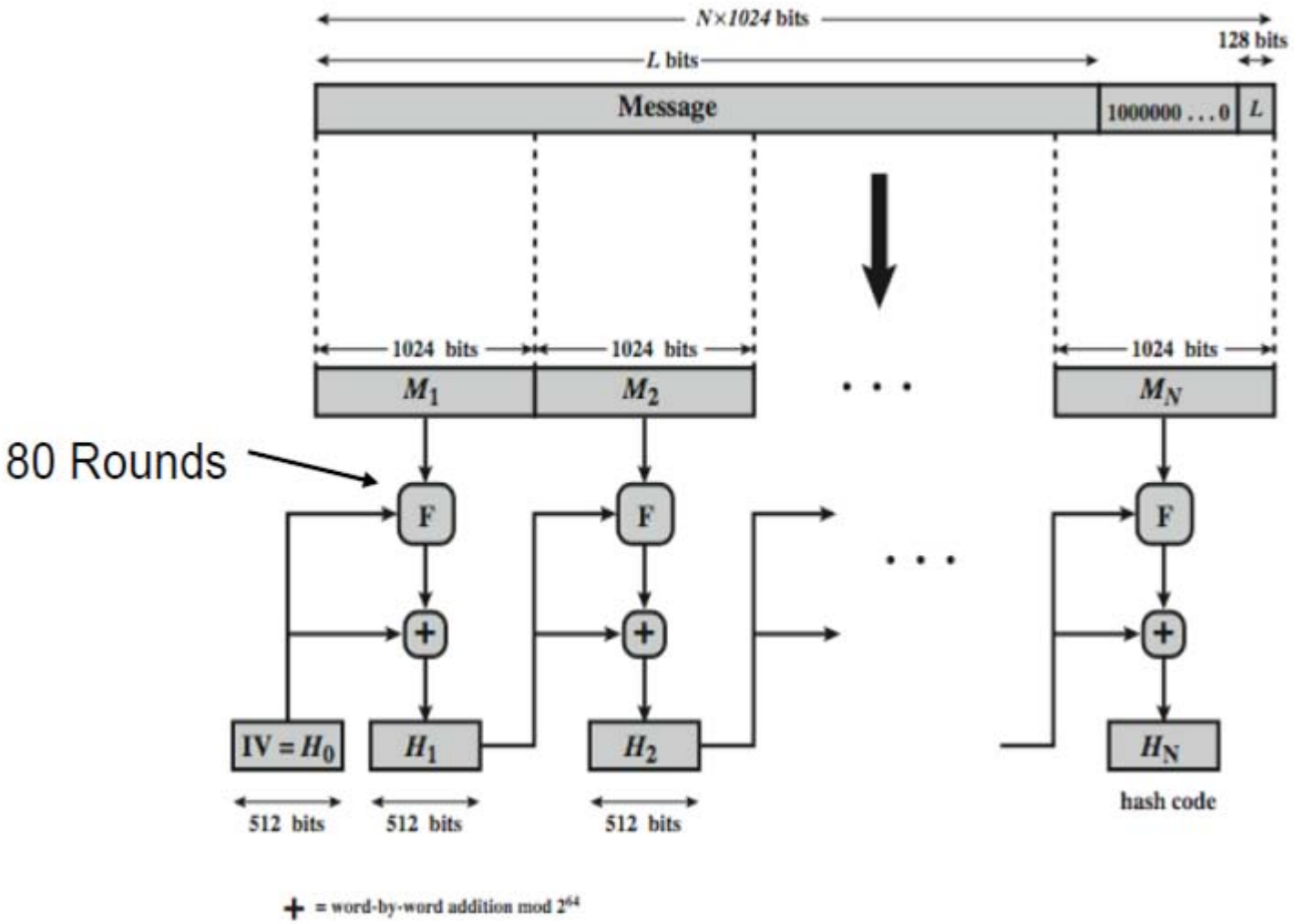
- **SHA originally designed by NIST & NSA in 1993**
- **was revised in 1995 as SHA-1**
- **US standard for use with DSA signature scheme**
  - **standard is FIPS 180-1 1995, also Internet RFC3174**

- **based on design of MD4 with key differences**
- **produces 160-bit hash values**
- **2005 results on security of SHA-1 raised concerns on its use in future applications**

Revised Secure Hash Standard

- **NIST issued revision FIPS 180-2 in 2002**
- **adds 4 additional versions of SHA**
  - **SHA-224, SHA-256, SHA-384, SHA-512**
- **designed for compatibility with increased security provided by the AES cipher**
- **structure & detail is similar to SHA-1**
- **hence analysis should be similar**
- **but security levels are rather higher**

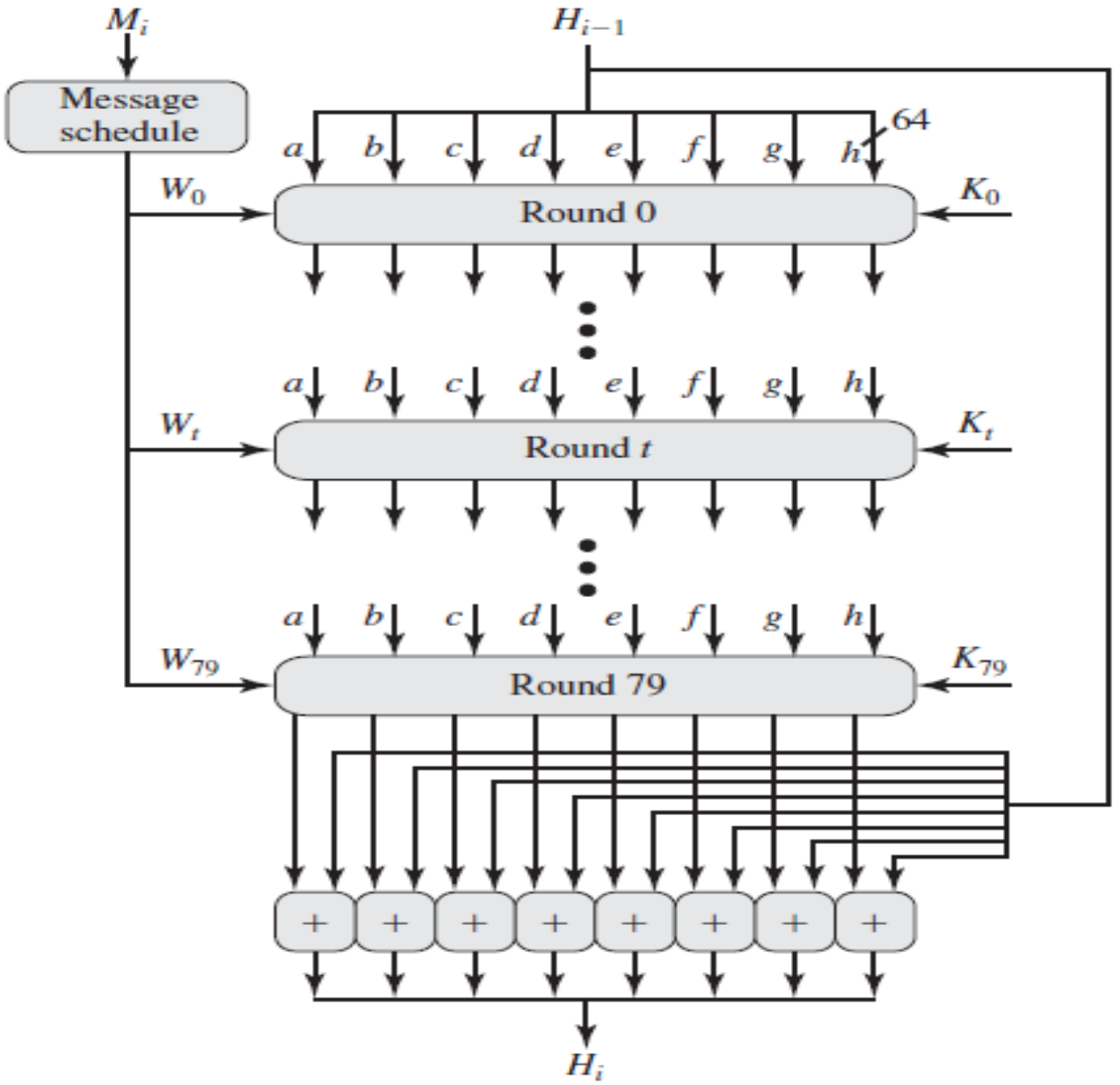
SHA-512 Overview



SHA-512 Compression Function

- heart of the algorithm
- processing message in 1024-bit blocks
- consists of 80 rounds
  - updating a 512-bit buffer
  - using a 64-bit value w_t derived from the current message block
  - and a round constant based on cube root of first 80 prime numbers

SHA-512 Compression Function



SHA

	Output size (bits)	Internal state size (bits)	Block size (bits)	Max message size (bits)	Word size (bits)	Rounds	Operations	Collisions found
SHA-0	160	160	512	$2^{64} - 1$	32	80	+, and, or, xor, rot	Yes
SHA-1	160	160	512	$2^{64} - 1$	32	80	+, and, or, xor, rot	None (2^{52} attack)
SHA-2	256/224	256	512	$2^{64} - 1$	32	64	+, and, or, xor, shr, rot	None
	512/384	512	1024	$2^{128} - 1$	64	80	+, and, or, xor, shr, rot	None