

## SHA-3 Competition: The Quest for Long-Term Security in Cryptographic Hashing

Bart Preneel  
 COSIC – Katholieke Universiteit Leuven, Belgium  
 bart.preneel(AT)esat.kuleuven.be

BeNeLux OWASP Day  
 2 December 2009

**ECRYPT II**  
 11th EU FP 6th IST


www.ecrypt.eu.org




## Hash functions

- MDC (manipulation detection code)
- Protect short hash value rather than long text

This is an input to a cryptographic hash function. The input is a very long string, that is reduced by the hash function to a string of fixed length. There are additional security conditions: it should be very hard to find an input hashing to a given value (a preimage) or to find two colliding inputs (a collision).



2

## Hash function flavours

cryptographic hash function

- MAC
- MDC *this talk*
  - OWHF
    - UOWHF (TCR)
    - CRHF

3

## Outline

- definitions and background
- generic attacks
- attacks on iterated constructions: the rise and fall of MD
- MD4, MD5, SHA-0, SHA-1
- SHA-2
- SHA-3: the NIST AHS competition
- SHA-4
- conclusions

4

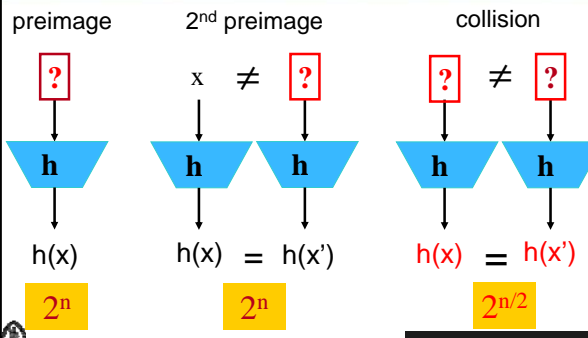
## Informal definitions (1)

- no secret parameters
- input string  $x$  of arbitrary length  $\Rightarrow$  output  $h(x)$  of fixed bitlength  $n$
- computation "easy"
- One Way Hash Function (OWHF)
  - preimage resistance
  - 2<sup>nd</sup> preimage resistance
- Collision Resistant Hash Function (CRHF): OWHF +
  - collision resistant

5

## Security requirements (n-bit result)

preimage      2<sup>nd</sup> preimage      collision



2<sup>n</sup>      2<sup>n</sup>      2<sup>n/2</sup> birthday paradox

## Applications

- digital signatures: OWHF/CRHF, 'destroy algebraic structure'
- information authentication: protect authenticity of hash result
- protection of passwords: preimage resistant
- confirmation of knowledge/commitment: OWHF/CRHF
- pseudo-random string generation/key derivation
- micropayments (e.g., micromint)
- construction of MAC algorithms, stream ciphers, block ciphers
- (redundancy: hash result appended to data before encryption)

Until 2005: 800 uses of MD5 in Windows

7

## Applications (2)

- Collision resistance is not always necessary
- Other properties are needed:
  - pseudo-randomness if keyed (with secret key)
  - near-collision resistance
  - partial preimage resistance
  - multiplication freeness
  - indifferentiable from random oracle
- Hard to formalize

8

## Brute force (2<sup>nd</sup>) preimage

- Finding a single second preimage:  $2^n$
- Multiple (2<sup>1</sup>) target second preimage (1 out of many):  $2^{n-1}$
- Multiple target second preimage (many out of many) [Hellman'80]
  - Precomputation:  $2^n$
  - Storage:  $2^{2n/3}$
  - inversion of one message in time:  $2^{n/3}$
- answer: randomize hash function
  - salt, spice, "key": parameter to index family of functions

9

## Brute force attacks in practice

- parallel (2<sup>nd</sup>) preimage search
  - $n = 128$ : 90 M\$ for 1 year if one out of  $2^{48}$  targets
  - $n = 128$ : 90 B\$ for 1 year if one out of  $2^{38}$  targets
- parallel collision search
  - $n = 128$ : 1 M\$ for 12 hours (or 1 year on 60K PCs)
  - $n = 160$ : 90 M\$ for 1 year
  - need 256-bit result for long term security (25 years or more)

10

## Quantum computers

- In principle exponential parallelism
- Inverting a one-way function:  $2^n$  reduced to  $2^{n/2}$  [Grover'96]
- Collision search:  $2^{n/2}$  no improvement in spite of claims



11

## Outline

- definitions and background
- generic attacks
- attacks on iterated constructions: the rise and fall of MD
- MD4, MD5, SHA-0, SHA-1
- SHA-2
- SHA-3: the NIST AHS competition
- SHA-4
- conclusions

13

### Hash function: iterated structure

Split messages into blocks of fixed length and hash them block by block with a compression function  $f$

Efficient and elegant  
But many problems...

14

### Security relation between $f$ and $h$

- Iterating  $f$  can degrade its security
  - trivial example: 2<sup>nd</sup> preimage

15

### Security relation between $f$ and $h$

- Solution: Merkle-Damgard (MD) strengthening (popular!)
  - fix IV, use unambiguous padding and insert length at the end
- [MD'89]  $f$  is collision resistant  $\Rightarrow$   $h$  is collision resistant
- [Lai-Massey'92]  $f$  is 2<sup>nd</sup> preimage resistant  $\Leftrightarrow$   $h$  is 2<sup>nd</sup> preimage resistant

16

### Attacks on MD: 1999-2006

- multi-collision attack [Joux'04]
  - the concatenation of 2 iterated hash functions is as most as strong as the strongest of the two (even if both are independent)
  - cost of collision attack against  $g$  at most  $n_1 \cdot 2^{n_2/2} + 2^{n_1/2} \ll 2^{(n_1 + n_2)/2}$

$g(x) = h_1(x) \parallel h_2(x)$

- other problems:
  - long message 2<sup>nd</sup> preimage attack [Dean-Felten-Hu'99], [Kelsey-Schneier'05]
  - herding attack [Kelsey-Kohn'06]

17

### Improving MD iteration

- degradation with use: salting (family of functions, randomization)
- strong output transformation  $g$  (which includes total length and salt)
- counter in every iteration
- larger internal memory: known as wide pipe

Many concrete proposals in the SHA-3 competition

18

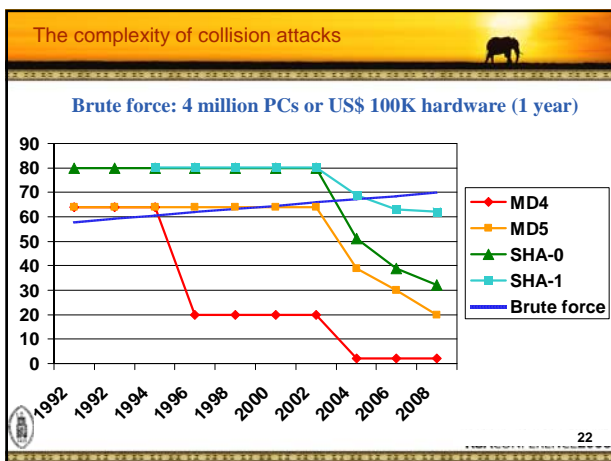
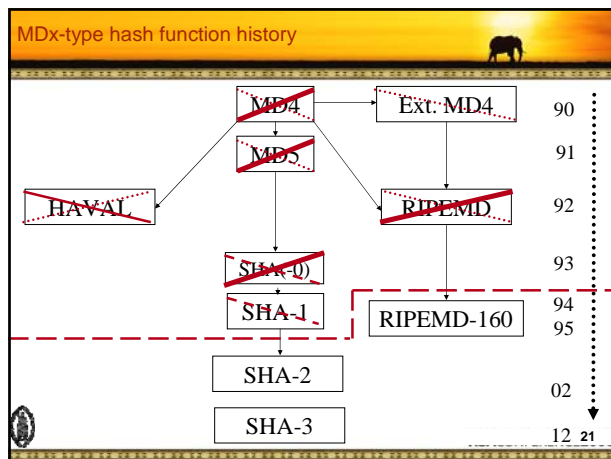
### Summary

19

### Outline

- definitions and background
- generic attacks
- attacks on iterated constructions: the rise and fall of MD
- MD4, MD5, SHA-0, SHA-1
- SHA-2
- SHA-3: the NIST AHS competition
- SHA-4
- conclusions

20



### MD5 [Rivest'91]

- 4 rounds (64 steps)
- pseudo-collisions [denBoer-Bosselaers'93]
- collisions for compression function [Dobbertin'96]
- collisions for hash function
  - [Wang+'04] – 15 minutes
  - [Stevens+'09] – milliseconds
  - brute force ( $2^{64}$ ): 1M\$ 10 hours in '09
- 2<sup>nd</sup> preimage in  $2^{123}$  [Sasaki-Aoki'09]
- Popularity:
  - License-free
  - 10 times faster than DES
  - Export restrictions in the 90s less stringent for hash functions than for block ciphers

23

### MD5

- Advice (RIPE since '92, RSA since '96): **stop using MD5**
- Largely ignored by industry (click on a cert...)

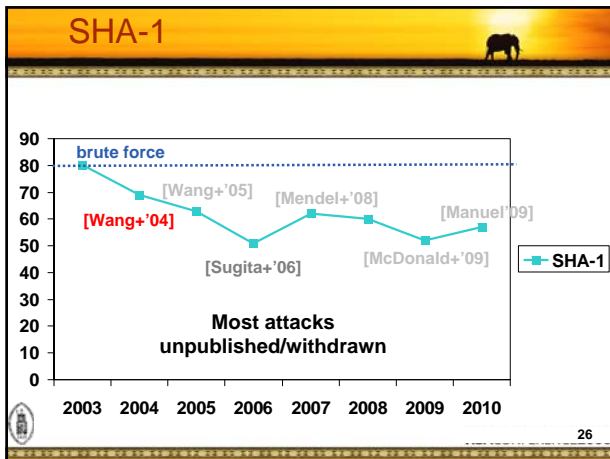
24

### SHA-1 [NIST'95]

- fix to SHA-0 (published '93, today collisions in 1 hour)
- collisions for SHA-1
  - 70 steps in  $2^{44}$  – highly structured [De Cannière-Rechberger'06]:
  - 70 steps  $2^{39}$  (4 days on a PC) [Joux-Peyrin'07]
  - $2^{69}$  [Wang+'05]
  - Can this be improved? [Wang+'05 - unpublished], [Sugita+'06 ], Mendel+'08 - unpublished], [McDonald+'09 - unpublished]
- preimages for 48/80 steps in  $2^{160-\epsilon}$  [Aoki-Sasaki'09]

Prediction: collision for SHA-1 in the next 12-18 months

25



### Impact of collisions

- collisions for MD5, SHA-0, SHA-1
  - 2 messages differ in a few bits in 1 to 3 512-bit input blocks
  - limited control over message bits in these blocks
  - but arbitrary choice of bits before and after them

- what is achievable for MD5?
  - 2 colliding executables/postscript/gif/... [Lucks-Daum'05]
  - 2 colliding RSA public keys – thus with colliding X.509 certificates [Lenstra+'04]
  - chosen prefix attack: different IDs, same certificate [Stevens+'07]
  - 2 arbitrary colliding files (no constraints) in 12 hours for 1 M\$**

28

### Rogue CA attack

[Sotirov-Stevens-Appelbaum-Lenstra-Molnar-Osvik-de Weger '08]

- request user cert; by special collision this results in a fake CA cert (need to predict serial number + validity period)

impact: **rogue CA** that can issue certs that are trusted by all browsers

- 6 CAs have issued certificates signed with MD5 in 2008:
  - Rapid SSL, Free SSL (free trial certificates offered by RapidSSL), TC TrustCenter AG, RSA Data Security, Verisign.co.jp

### Other ways to fool CAs

- [Moxie Marlinspike'09] Black Hat
  - browsers may accept bogus SSL certs
  - CAs may sign malicious certs
- certificate for [www.paypal.com](http://www.paypal.com) \0.kuleuven.be will be issued if the request comes from a kuleuven.be admin
- response by PayPal: suspend Moxie's account
  - [http://www.theregister.co.uk/2009/10/06/paypal\\_banishes\\_ssl\\_hacker/](http://www.theregister.co.uk/2009/10/06/paypal_banishes_ssl_hacker/)
- Related (independent) work at Financial Cryptography 2010

30

### Impact of collisions (3)

- digital signatures: only an issue if for **non-repudiation**
- none** for signatures computed before attacks were public (1 August 2004)
- none** for certificates if public keys are generated at random in a controlled environment
- substantial** for signatures after 1 August 2005 (cf. traffic tickets in Australia)

31

## And (2<sup>nd</sup>) preimages?

- security degrades with number of applications
- for large messages even with the number of blocks (cf. supra)
- specific results:
  - MD2:  $2^{73}$  [Knudsen+09]
  - MD4:  $2^{102}$  [Leurent'08]
  - MD5:  $2^{123}$  [Sasaki-Aoki'09]
  - SHA-1: 48 of 80 steps in  $2^{159.3}$  [Aoki-Sasaki'09]

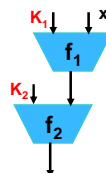


32

## HMAC

- HMAC keys through the IV (plaintext) [Kim+'06]
  - collisions for MD5 invalidate current security proof of HMAC-MD5
  - new attacks on reduced version of HMAC-MD5 and HMAC-SHA-1

	Rounds in f2	Rounds in f1	Data complexity
MD4	48	48	$2^{72}$ CP + $2^{77}$ time
MD5	64	33 of 64	$2^{126.1}$ CP
MD5	64	64	$2^{51}$ CP & $2^{100}$ time (RK)
SHA-1	80	53 of 80	$2^{98.5}$ CP



no problem yet for most widely used schemes



33

## Fixes/Alternatives

- One could fix SHA-1: more steps, message precoding, patch
- RIPEMD-160 seems more secure than SHA-1 😊
- Getting rid of SHA-1 and MD5 is much harder than expected: algorithm negotiation problem
  - Example: TLS1.2



34

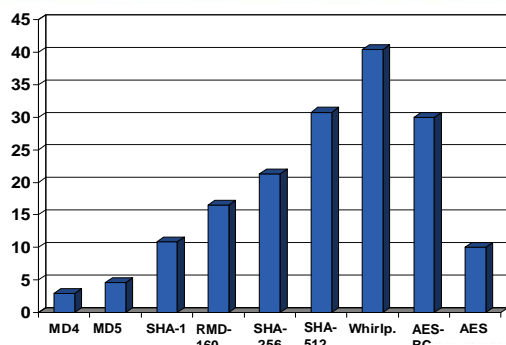
## SHA-2 [NIST'02]

- SHA-224, SHA-256, SHA-384, SHA-512
  - non-linear message expansion
  - more complex operations
  - 64/80 steps
  - SHA-384 and SHA-512: 64-bit architectures
- collisions: 24 steps [Sanadhyasarkar'08]
- implementations today faster than anticipated
- adoption
  - industry may migrate to SHA-2 by 2011 or may wait for SHA-3
  - very slow for TLS/IPsec (no pressing need)
- alternative: Whirlpool (AES-based), in ISO standard



35

## Performance of hash functions - Bernstein (cycles/byte) AMD Intel Pentium D 2992 MHz (f64)



36

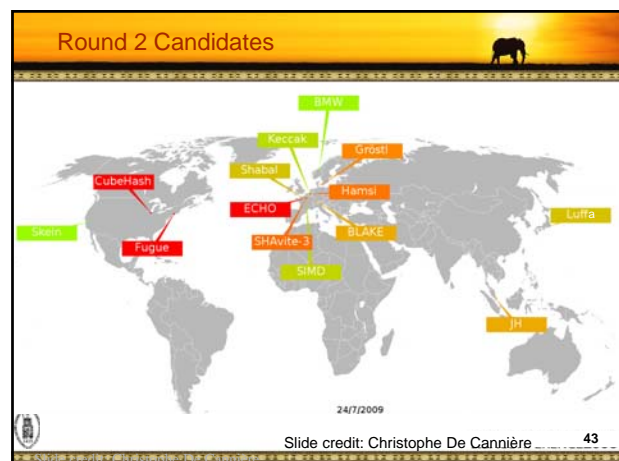
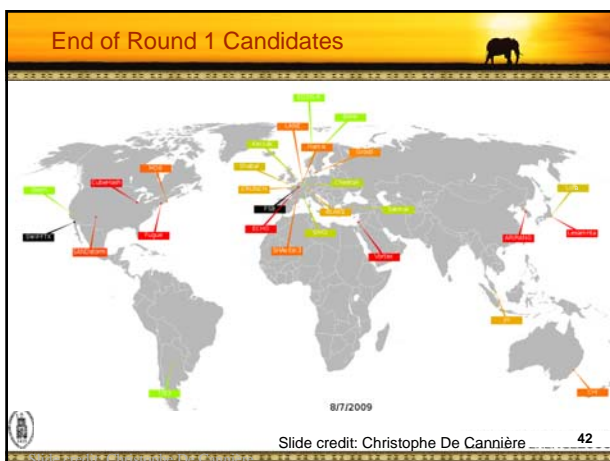
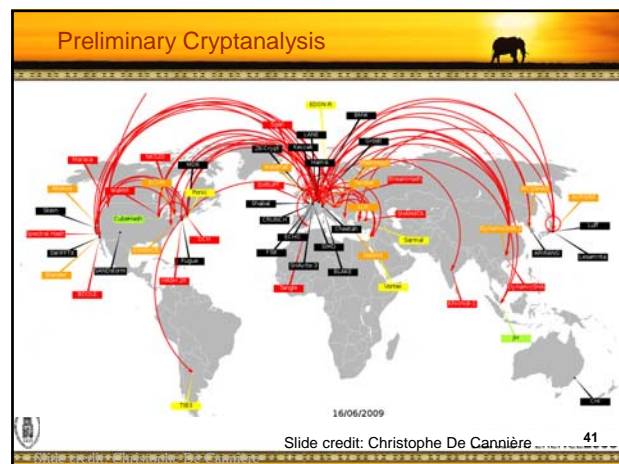
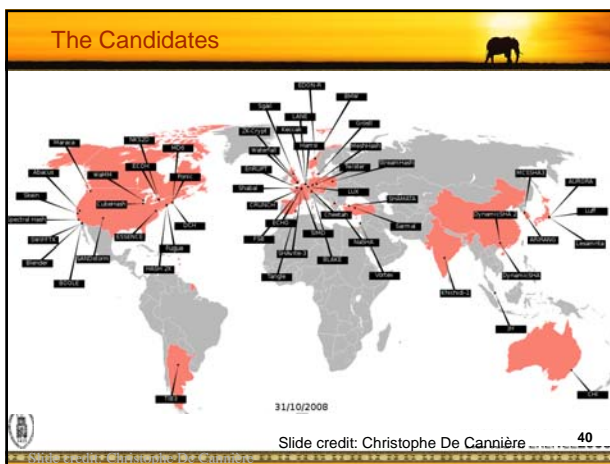
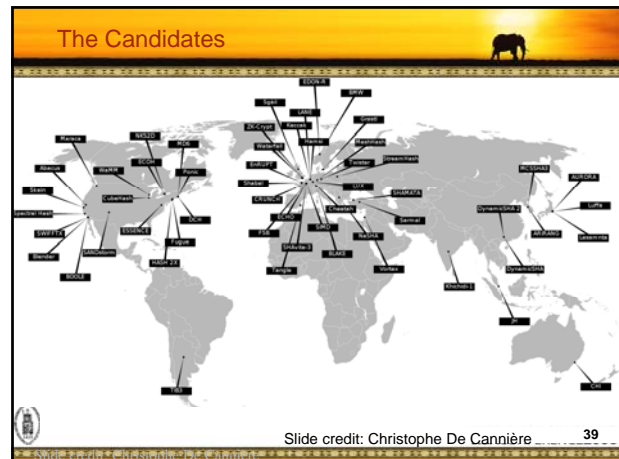
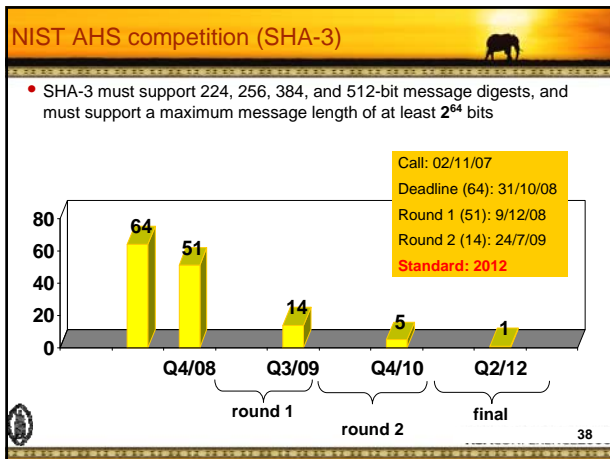
## Outline

- definitions and background
- generic attacks
- attacks on iterated constructions: the rise and fall of MD
- MD4, MD5, SHA-0, SHA-1
- SHA-2
- SHA-3: the NIST AHS competition
- SHA-4
- conclusions



37





## Issues arisen during Round 1

- Security:
  - controversy around pseudo-collision attacks and memory requirements
  - proofs have not helped much to survive
- Performance: roughly as fast or faster than SHA-2
  - tunable security/performance tradeoff: nominal parameters?
  - large memory (> 100 bytes) may be a problem for small devices
  - can we exploit 64 or 128 cores? Intel AES instruction?
- 14 Round 2 candidates
  - most are wide-pipe designs or sponge-like designs
  - two main types: AES-based and AXR (addition/xor/rotate)

44

## Security: SHA-3 Zoo

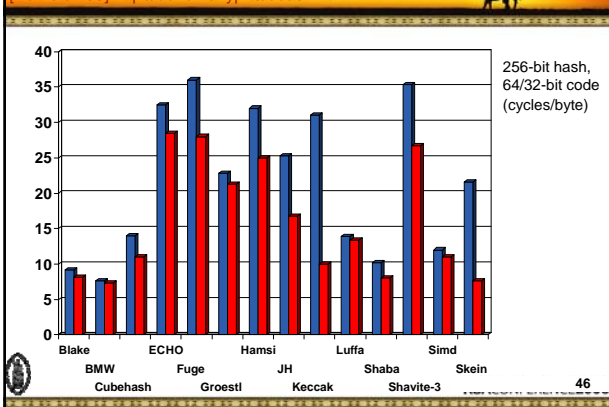
[http://ehash.iaink.tugraz.at/wiki/The\\_SHA-3\\_Zoo](http://ehash.iaink.tugraz.at/wiki/The_SHA-3_Zoo)



45

## Performance of hash functions

[Bernstein09] <http://bench.cr.yp.to/ebash.html>



46

## SHA-4

- an open competition such as SHA-3 is bound to result in new insights between 2009-2012
- only few of these can be incorporated using “tweaks”
- the winner selected in 2012 will reflect the state of the art in October 2008
- nevertheless, it is unlikely that we will have a SHA-4 competition before 2030
- for which applications will security proofs be relevant?

47

## Hash functions: conclusions

- SHA-1 would have needed 128-160 steps instead of 80
- recent attacks are only dramatic for a few applications: in practice we are saved by the fact that weaker security guarantees are mostly sufficient (but hard to assess for non-experts)
- upgrading a cryptographic algorithm is difficult
- because of faster block ciphers and stream ciphers, our addition to hash functions usage will decrease
- theory is developing for more robust iteration modes and extra features; still early for building blocks
- need for lightweight hash functions
- Nirwana: efficient hash functions with security reduction

48

The end


Thank you for  
your attention




49



### Hash functions: links



- NIST <http://csrc.nist.gov/groups/ST/hash/index.html>
  - first SHA-3 candidate conference: February 25-28, 2009, Leuven
  - workshop October 31-November 1, 2005 and August 24-25, 2006
- ECRYPT: <http://www.ecrypt.eu.org>
  - SHA-3 Zoo [http://ehash.iaink.tugraz.at/wiki/The\\_SHA-3\\_Zoo](http://ehash.iaink.tugraz.at/wiki/The_SHA-3_Zoo)
  - workshops in May 2007 and June 2005 + statement on hash functions
- The IACR eprint server <http://eprint.iacr.org>
- My 1993 PhD thesis <http://homes.esat.kuleuven.be/~preneel>
- Overview paper from 1998 (LNCS 1528)  
<http://www.cosic.esat.kuleuven.be/publications/article-246.pdf>

50