EE309 Advanced Programming Techniques for EE

Lecture 21: Message Authentication Codes (MACs) and Hashes

INSU YUN (윤인수)

School of Electrical Engineering, KAIST

[Slides from Introduction to Computer Security (18-487) at CMU]

Message Integrity

Goal: ***integrity*** (not secrecy)

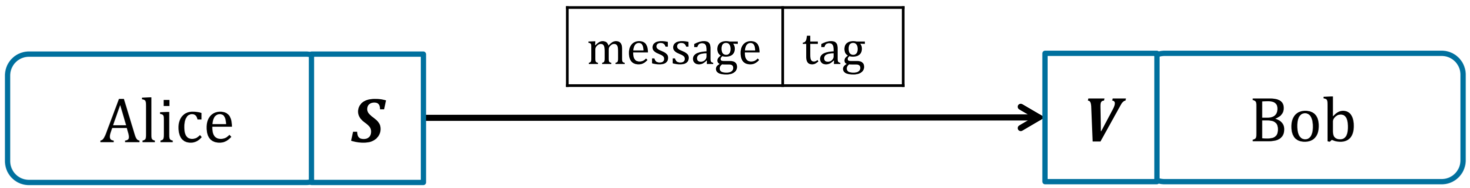
Examples:

- Protecting binaries on disk.
- Protecting banner ads on web pages

Security Principles:

- Integrity means no one can forge a signature

CRC (Cyclic Redundancy Check)



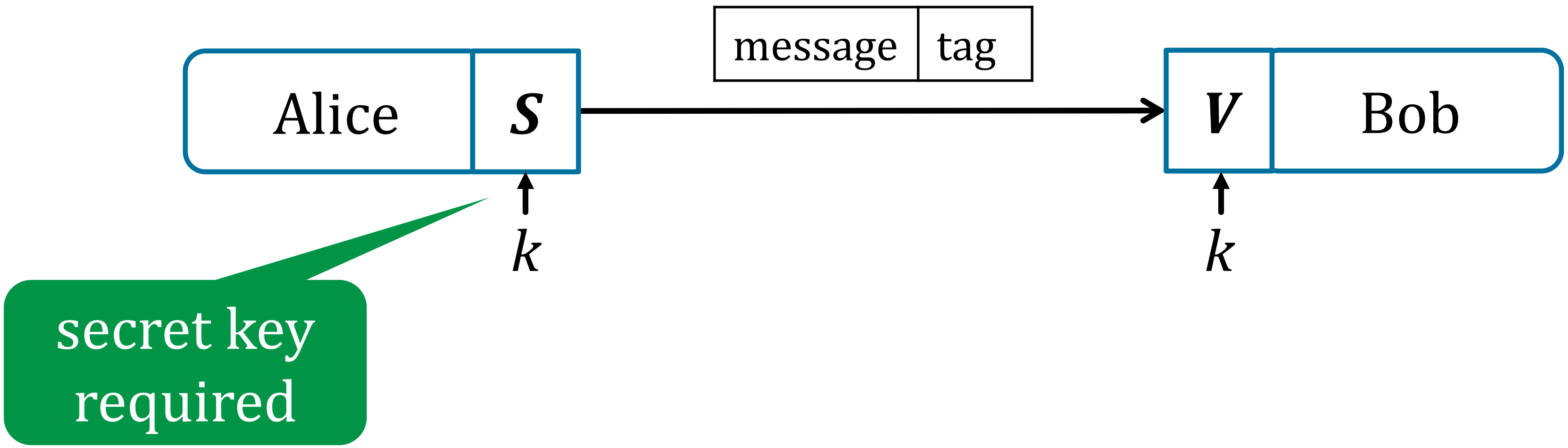
Generate tag:
tag $\leftarrow$ CRC(m)

Verify tag:
CRC(m, tag) ?= ‘yes’

Is this Secure?

- No! Attacker can easily modify message m and re-compute CRC.
- CRC designed to detect *random errors*, not malicious attacks.

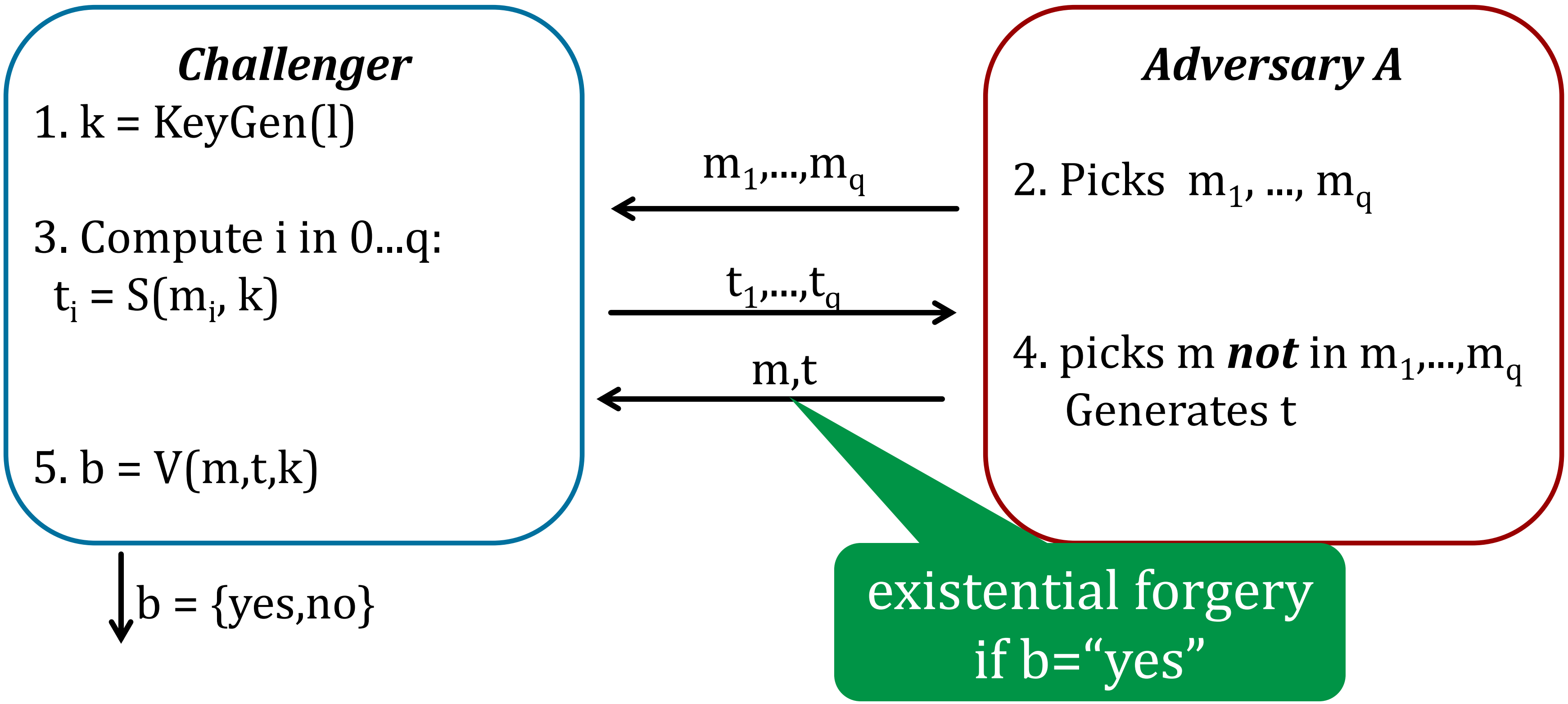
Message Authentication Codes (MAC)



Defn: A *Message Authentication Code (MAC)* MAC = (S,V) defined over (K,M,T) is a pair of algorithms:

- S(k,m) outputs t in T
- V(k,m,t) outputs \`yes' or \`no'
- V(k, m, S(k,m)) = 'yes' (consistency req.)

Secure MAC Game



Security goal: ***A*** cannot produce a valid tag on a message

– Even if the message is gibberish

Secure MAC Game

Challenger

1. k = KeyGen(l)

3. Compute i in 0...q:
   $t_i = S(m_i, k)$

5. b = V(m,t,k)

$\downarrow$ b = {yes,no}

$\xleftarrow{m_1,\ldots,m_q}$

$\xrightarrow{t_1,\ldots,t_q}$

$\xleftarrow{m,t}$

Adversary A

2. Picks $m_1, \ldots, m_q$

4. picks m ***not*** in $m_1,\ldots,m_q$
   Generates t

Def: I=(S,V) is a ***secure MAC*** if for all "efficient" A:

$\mathrm{Adv}_{\mathrm{MAC}}[A,I] = \Pr[\text{Chal. outputs } 1] < \varepsilon$

Let I = (S,V) be a MAC.

Suppose an attacker is able to find $m_0 \neq m_1$ such that

$$S(k, m_0) = S(k, m_1) \quad \text{for } ½ \text{ of the keys } k \text{ in } K$$

Can this MAC be secure?

1. Yes, the attacker cannot generate a valid tag for m_0 or m_1
2. No, this MAC can be broken using a chosen msg attack
3. It depends on the details of the MAC

1. A sends m_0, receives (m_0, t_0)
2. A wins with (m_1, t_0)
3. Adv[A,I] = ½ since prob. of key is ½.

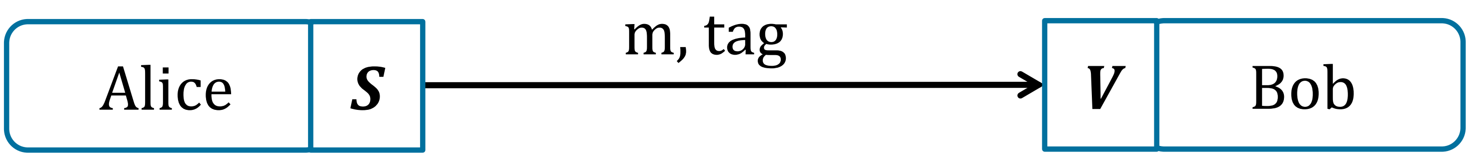
MACs from PRFs

Secure PRF implies secure MAC

For a PRF $F: K \times X \longrightarrow Y$, define a MAC $I_F = (S,V)$ as:

– $S(k,m) = F(k,m)$

– $V(k,m,t)$: if $t = F(k,m)$, output 'yes' else 'no'



$tag \leftarrow F(k,m)$

accept msg if

$tag = F(k,m)$

Security

Thm: If $F: K \times X \longrightarrow Y$ is a secure PRF and $1/|Y|$ is negligible (i.e., $|Y|$ is large), then I_F is a secure MAC.

In particular, for every eff. MAC adversary **A** attacking I_F, there exists an eff. PRF adversary **B** attacking F s.t.:

$$\mathrm{Adv}_{\mathrm{MAC}}[\mathbf{A}, I_F] \leq \mathrm{Adv}_{\mathrm{PRF}}[\mathbf{B}, F] + 1/|Y|$$

Proof Sketch

Let f be a truly random function

Challenger

2. f from FUNS[X,Y]
3. Calculates

$t_i = f(k, m_i)$

$\xleftarrow{m_1,\ldots,m_q}$

$\xrightarrow{t_1,\ldots,t_x}$

$\xleftarrow{m,t}$

b

Adversary A

1. Picks $m_1, \ldots, m_q$

4. picks m ***not*** in $m_1,\ldots,m_q$. Generates t

A wins iff t=f(k,m) and m not in $m_1,\ldots,m_q$

PR[***A*** wins] = Pr[***A*** guesses value of rand. function on new pt]

= 1/|Y| same must hold for F(k, x)

Question

Suppose $F: K \times X \longrightarrow Y$ is a secure PRF with $Y = \{0,1\}^{10}$

Is the derived MAC I_F a *practically* secure MAC system?

1. Yes, the MAC is secure because the PRF is secure
2. No tags are too short: guessing tags isn't hard
3. It depends on the function F

Adv[A,F] = 1/1024
(we need |Y| to be large)

Secure PRF *implies* secure MAC

$$S(k,m) = F(k,m)$$

Assuming output domain Y is large

So AES is already a secure MAC....

... but AES is only defined on 16-byte messages

Building Secure MACs

<ins>Given:</ins> a PRF for shorter messages (e.g., 16 bytes)

<ins>Goal:</ins> build a MAC for longer messages (e.g., gigabytes)

Construction examples:

- CBC-MAC: Turn small PRF into big PRF
- HMAC: Build from collision resistance

HMAC (Hash-MAC)

Most widely used MAC on the Internet.

… but, we first we need to discuss hash function.

Hash Functions

Collision Resistance

Let H: X $\rightarrow$ Y be a hash function (|X| >> |Y|)

A **collision** for H is a pair m_0 , $m_1 \in M$ such that:

$$H(m_0) = H(m_1) \quad \text{and} \quad m_0 \neq m_1$$

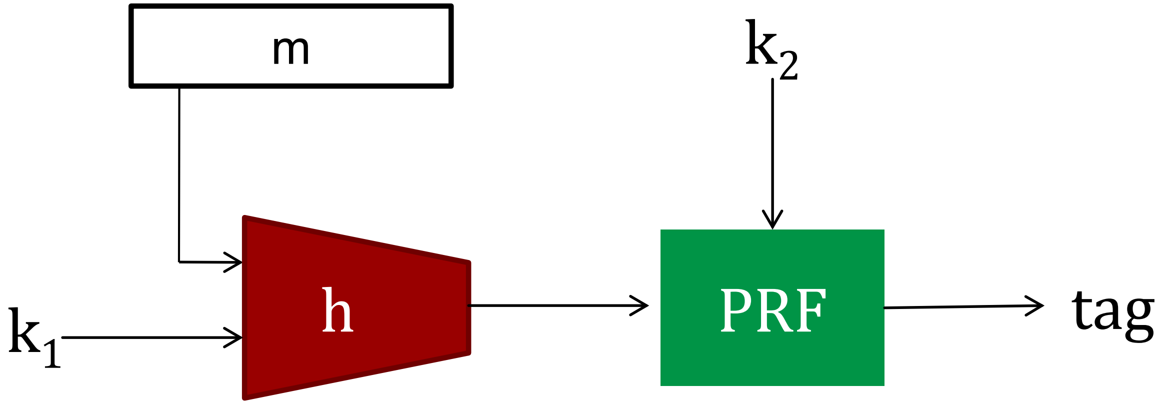
A function H is **collision resistant** if for all (explicit) "eff" algs. A:

$$\mathbf{Adv_{CR}[A,H] = Pr[\ A\ outputs\ collision\ for\ H]}$$

is "negligible".

Example: SHA-256 (outputs 256 bits)

General Idea



Hash then PRF construction

MACs from Collision Resistance

Let $I = (S,V)$ be a MAC for short messages over (K,M,T) (e.g. AES)

Let $H: X \to Y$ and $S: K \times Y \to T$ $\qquad (|X| >> |Y|)$

Def: $I^{big} = (S^{big}, V^{big})$ over (K, X^{big}, Y) as:

$$S^{big}(k,m) = S(k,H(m)) \quad ; \quad V^{big}(k,m,t) = V(k,H(m),t)$$

<u>Thm</u>: If I is a secure MAC and H is collision resistant, then I^{big} is a secure MAC.

Example: $S(k,m) = AES_{2\text{-block-cbc}}(k,\ \text{SHA-256}(m))$ is secure.

MACs from Collision Resistance

$S^{big}(k, m) = S(k, H(m))$; $V^{big}(k, m, t) = V(k, H(m), t)$

Collision resistance is necessary for security:

Suppose: adversary can find $m_0 \neq m_1$ s.t. $H(m_0) = H(m_1)$.

Then: $\mathbf{S^{big}}$ is insecure under a 1-chosen msg attack

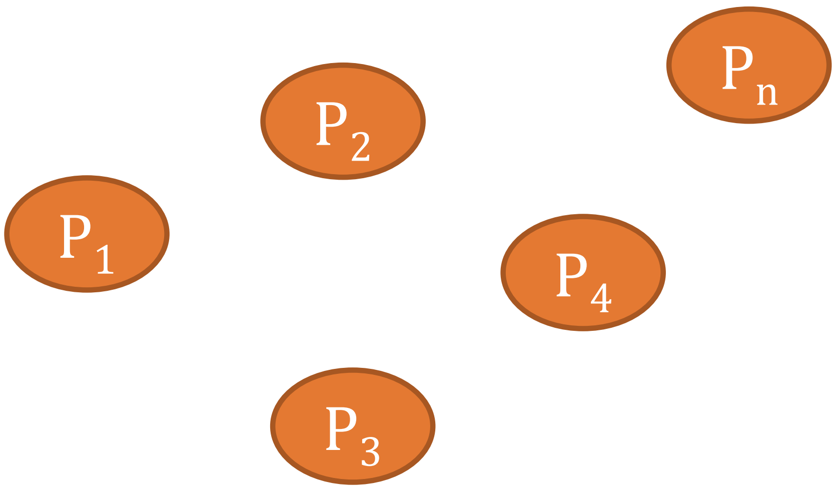
step 1: adversary asks for $t \longleftarrow S(k, m_0)$

step 2: output (m_1, t) as forgery

Collisions and the Birthday Paradox

Birthday Paradox

Put n people in a room. What is the probability that 2 of them have the same birthday?



$PR[P_i = P_j] > .5$ with 23 people.
(Think: n^2 different pairs)

Birthday Paradox Rule of Thumb

Given N possibilities, and random samples x_1, …, x_j, $PR[x_i = x_j] \approx 50\%$ when $j = N^{1/2}$

Generic attack on hash functions

Let $H: M \to \{0,1\}^n$ be a hash function ($|M| >> 2^n$)

Generic alg. to find a collision **in time** $\mathbf{O(2^{n/2})}$ hashes

Algorithm:

1. Choose $2^{n/2}$ random messages in M: $m_1, \ldots, m_{2^{n/2}}$ (distinct w.h.p)
2. For i = 1, …, $2^{n/2}$ compute $t_i = H(m_i) \in \{0,1\}^n$
3. Look for a collision ($t_i = t_j$). If not found, got back to step 1.

How well will this work?

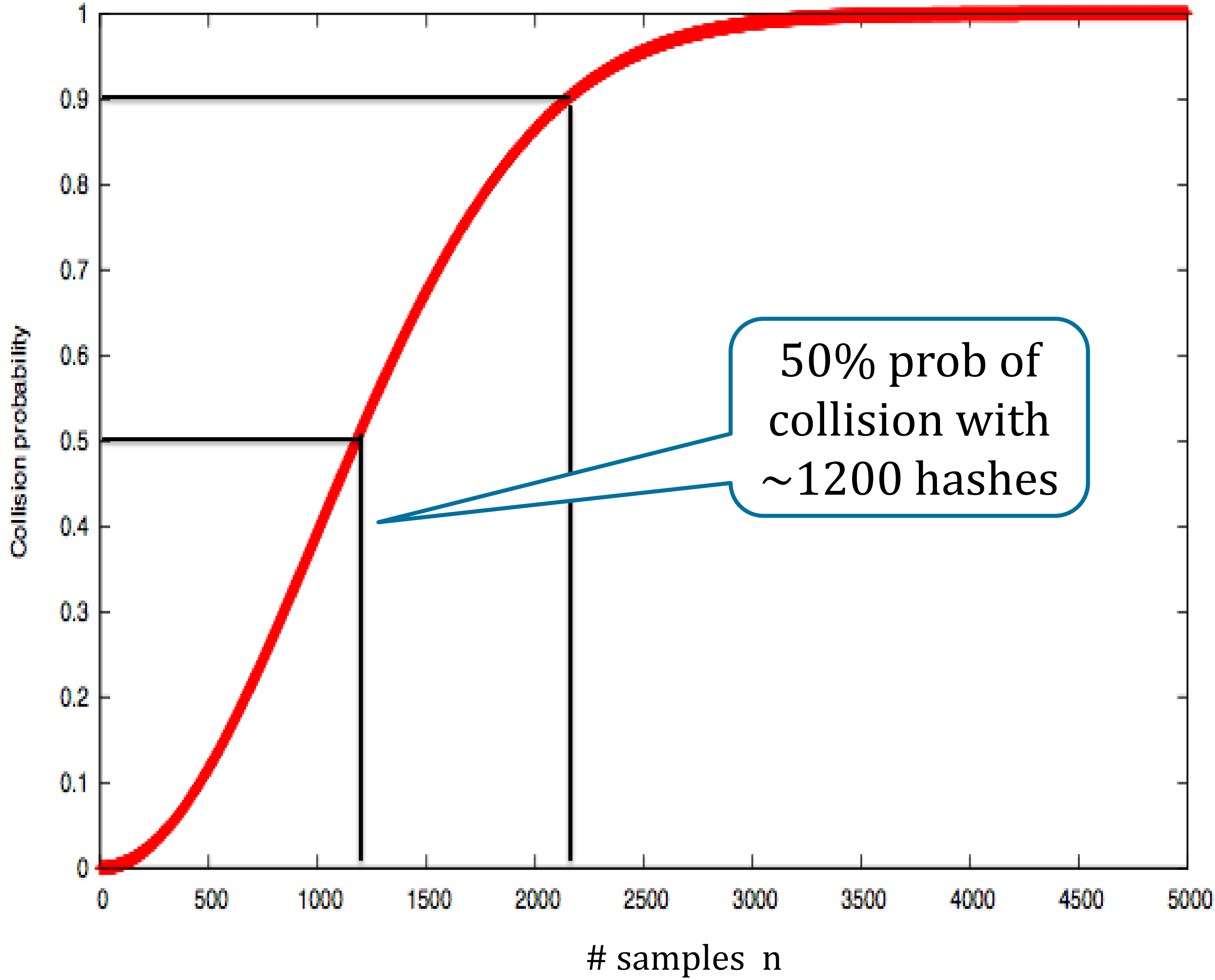
The birthday paradox

Let $r_1, \ldots, r_i \in \{1,\ldots,n\}$ be indep. identically distributed integers.

Thm:
when **i**$= 1.2 \times$ **$n^{1/2}$** then $\Pr[\ \exists\, i \neq j:\ r_i = r_j\ ] \geq ½$

If H: M-> $\{0,1\}^n$, then
Pr[collision] ~ ½
with $n^{1/2}$ hashes

B=10^6



Collision probability
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
50% prob of collision with ~1200 hashes
samples n

Sample Speeds Crypto++ 5.6.0 [Wei Dai]

AMD Opteron, 2.2 GHz (Linux)

	function	digest size (bits)	Speed (MB/sec)	generic attack time
NIST standards	SHA-1	160	153	2^{80}
	SHA-256	256	111	2^{128}
	SHA-512	512	99	2^{256}
	Whirlpool	512	57	2^{256}

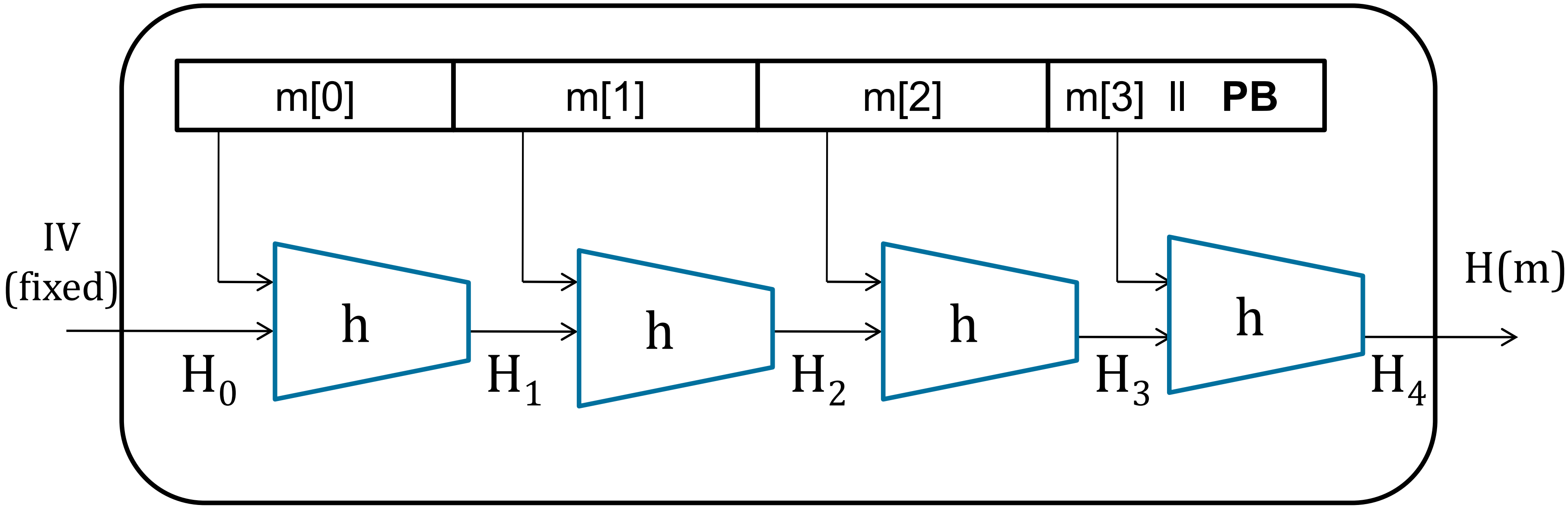
* best known collision finder for SHA-1 requires 2^{51} hash evaluations

Merkle-Damgard

How to construct collision resistant hash functions

http://www.merkle.com/

The Merkle-Damgard iterated construction



Given $\mathbf{h: T \times X \longrightarrow T}$ (compression function)

we obtain $\mathbf{H: X^{\leq L} \longrightarrow T}$. H_i - chaining variables

PB: padding block $\boxed{1000...0 \;\; \| \;\; \underbrace{\text{msg len}}_{\text{64 bits}}}$

If no space for PB
add another block

Security of Merkle-Damgard

Thm: if h is collision resistant then so is H.

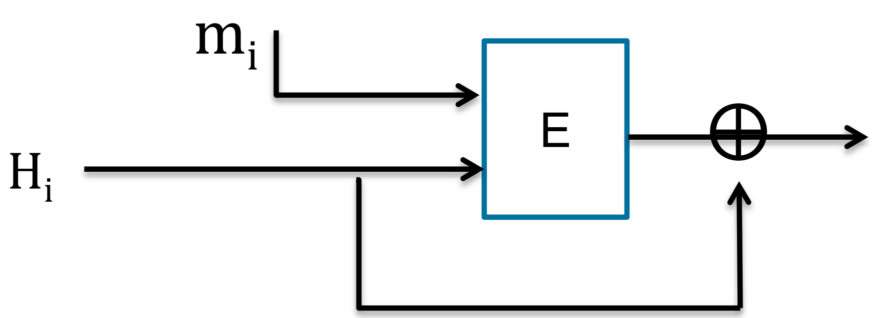
Proof Idea:
via contrapositive. Collisions on $H \Rightarrow$ collision on h

Suppose H(M) = H(M'). We build collision for h.

Compr. func. from a block cipher

E: K× {0,1}^{n} ⟶ {0,1}^{n} a block cipher.

The **Davies-Meyer** compression function
$h(H, m) = E(m, H) \oplus H$



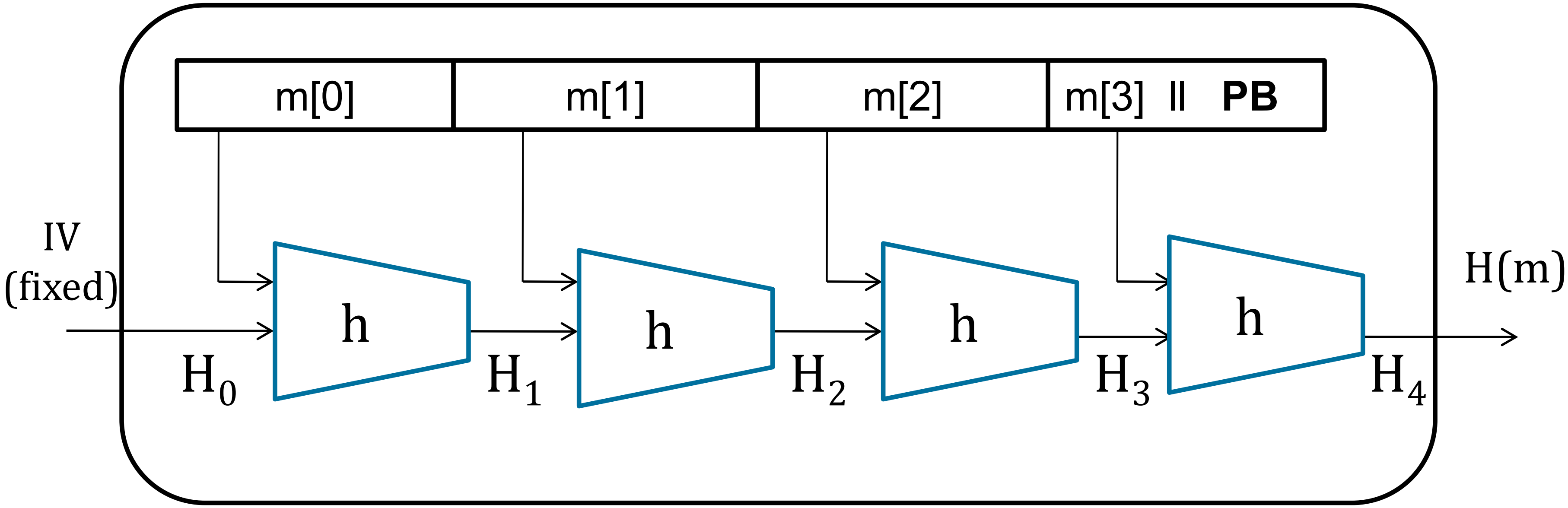
<u>Thm</u>: Suppose E is an ideal cipher (collection of |K| random perms.).

Finding a collision *h(H,m)=h(H',m')* takes $O(2^{n/2})$ evaluations of (E,D).

Hash MAC (HMAC)

Most widely used approach on the internet, e.g., SSL, SSH, TLS, etc.

Recall Merkel-Damgard



<u>Thm</u>:

h collision resistant implies H collision resistant

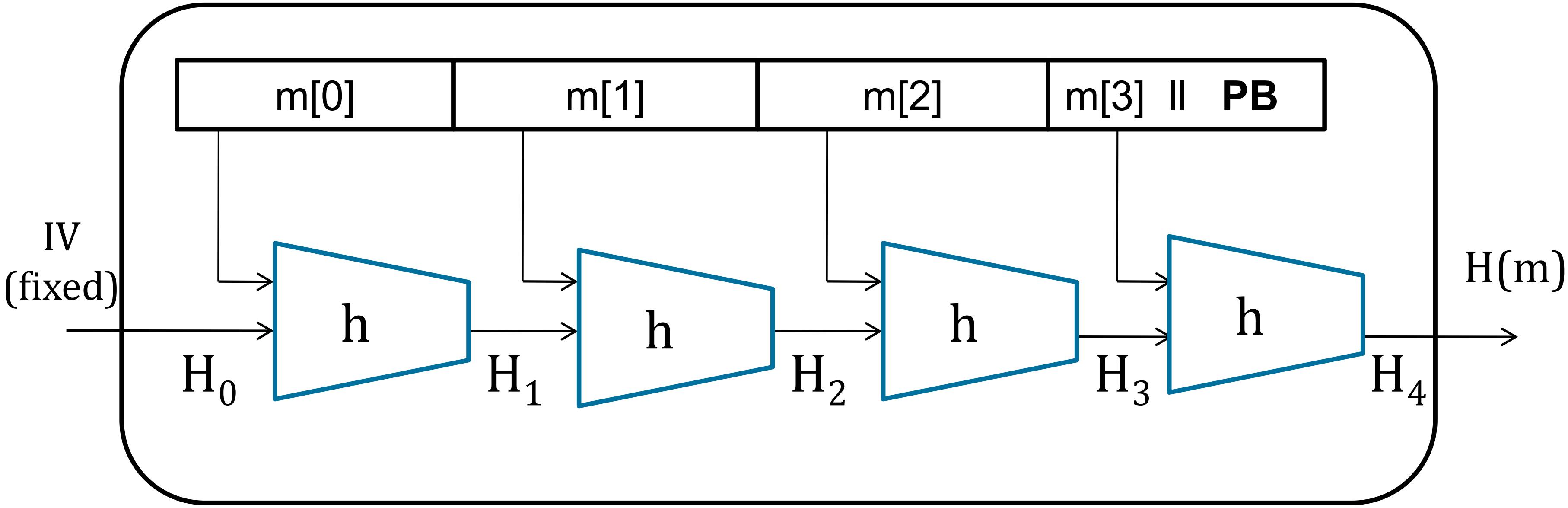
Can we build a MAC out of H?

Attempt 1

Let H: $X^{\leq L} \longrightarrow T$ be a Merkle-Damgard hash, and:

$S(k,m) = H(k||m)$

is this secure? no! why?



Existential forgery:
$H(k||m) = H(k||m||PB||w)$
(just one more h)

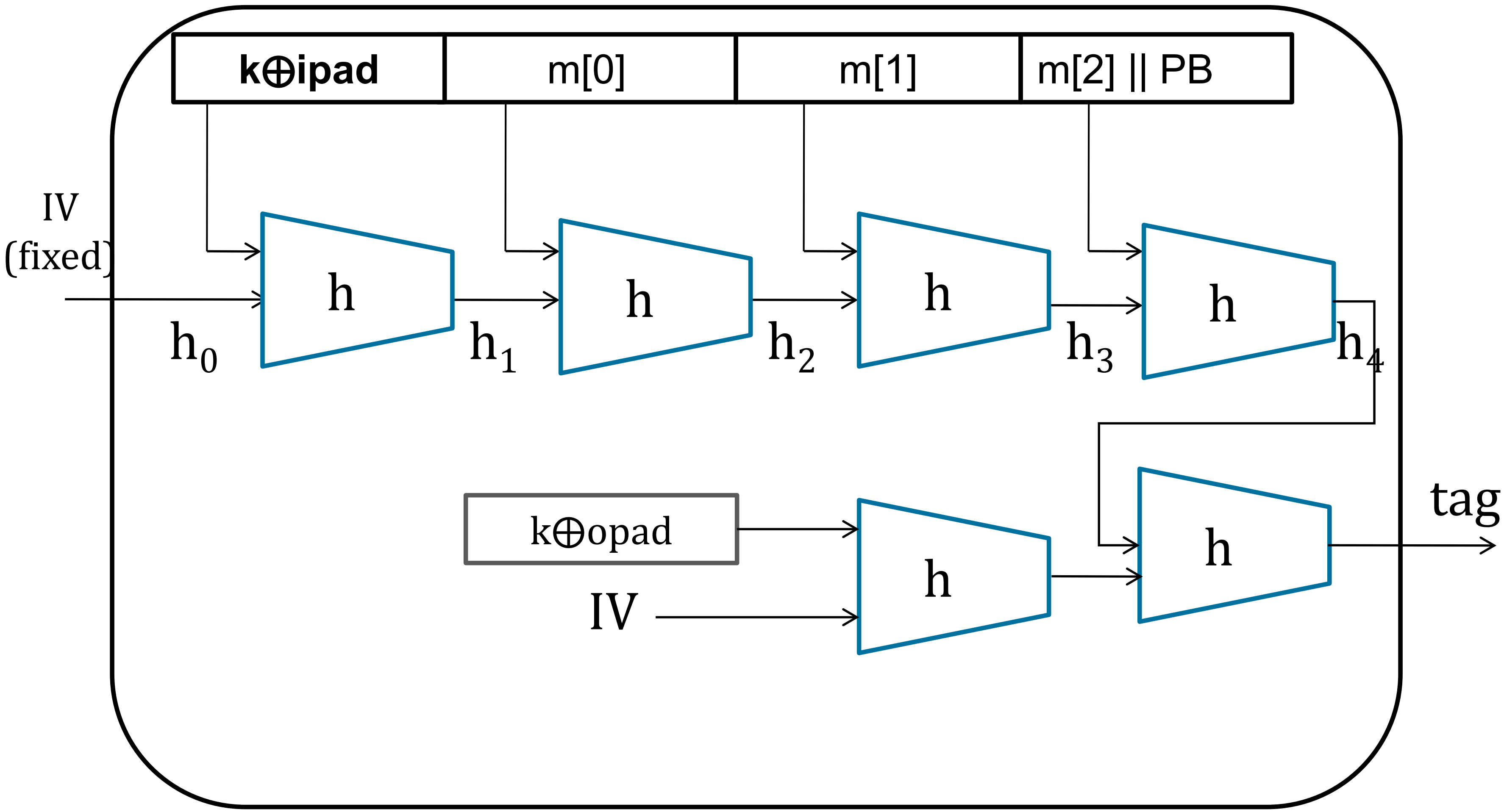
Hash Mac (HMAC)

Build MAC out of a hash

$$\text{HMAC:}\quad S(k, m) = H(\ k\oplus\text{opad}\ ,\ H(k\oplus\text{ipad} \,||\, m)\ )$$

- Example: H = SHA-256

HMAC



PB: Padding Block

Further reading

- J. Black, P. Rogaway: CBC MACs for Arbitrary-Length Messages: The Three-Key Constructions. J. Cryptology 18(2): 111-131 (2005)
- K. Pietrzak: A Tight Bound for EMAC. ICALP (2) 2006: 168-179
- J. Black, P. Rogaway: A Block-Cipher Mode of Operation for Parallelizable Message Authentication. EUROCRYPT 2002: 384-397
- M. Bellare: New Proofs for NMAC and HMAC: Security Without Collision-Resistance. CRYPTO 2006: 602-619
- Y. Dodis, K. Pietrzak, P. Puniya: A New Mode of Operation for Block Ciphers and Length-Preserving MACs. EUROCRYPT 2008: 198-219


Questions?

END