해시함수

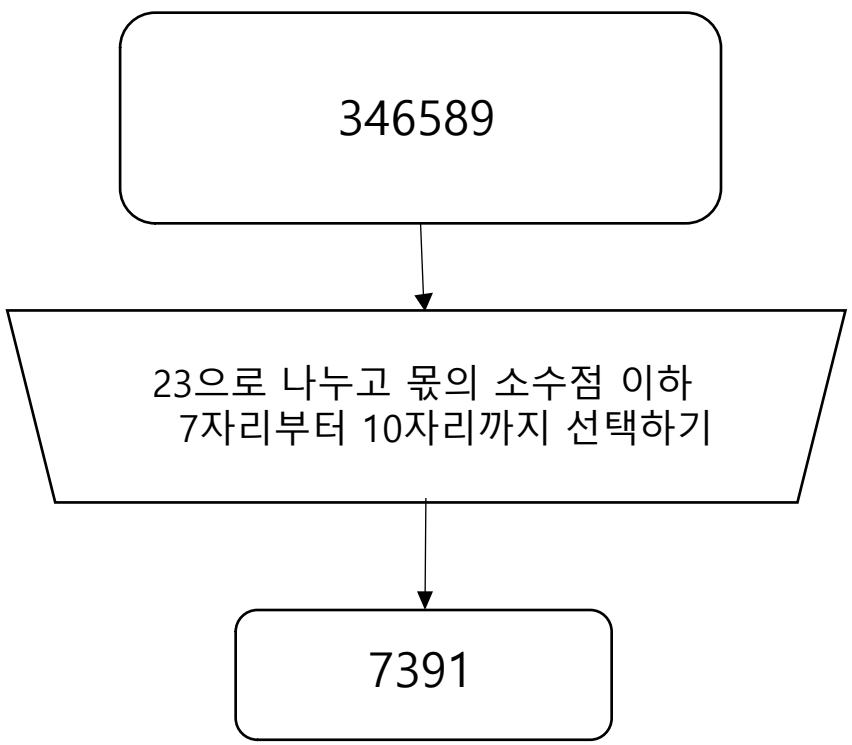
https://youtu.be/Wia5lTIc9g0

Contents

CryptoCraft LAB

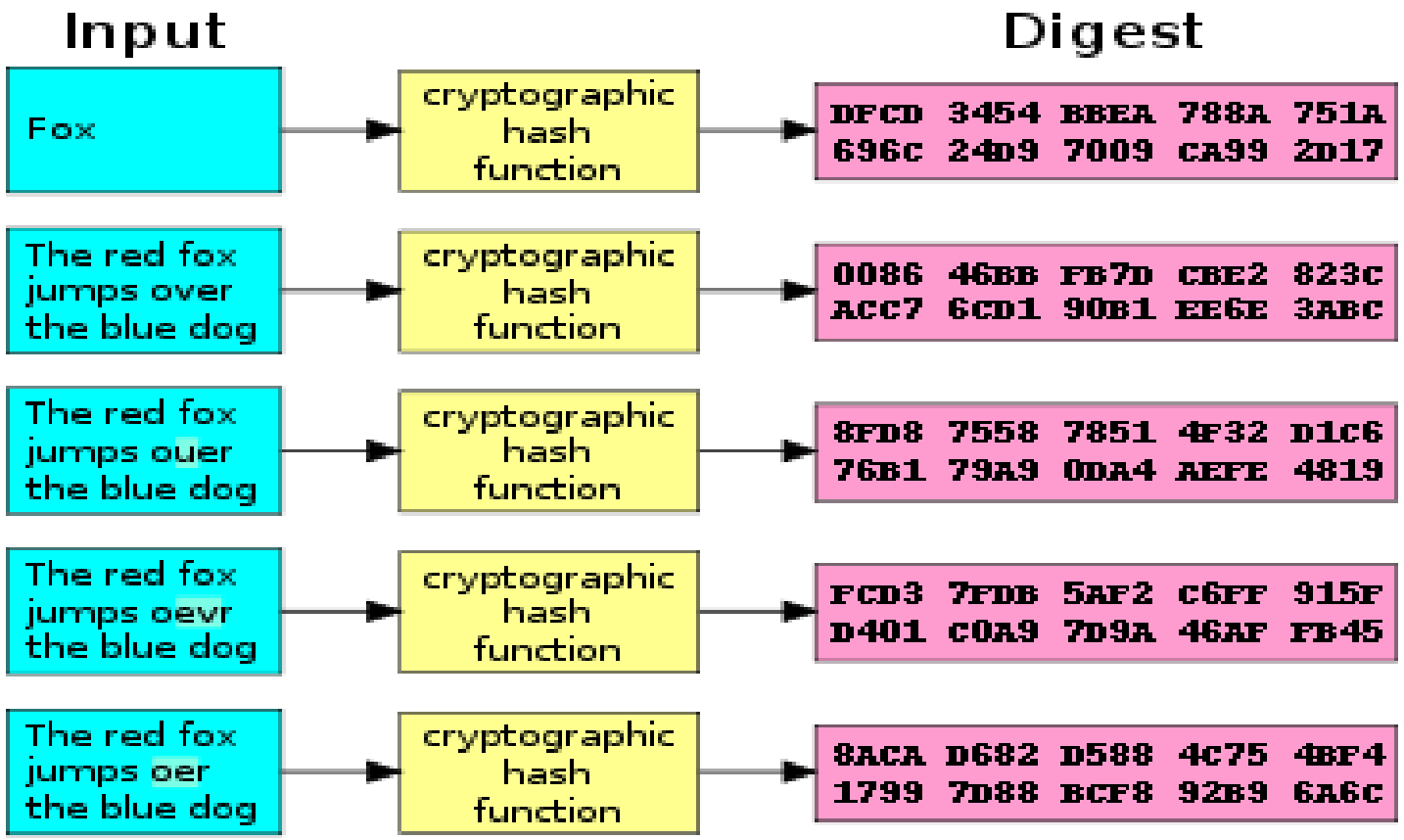
해시함수란?

- 특성 (1) : 일방향성



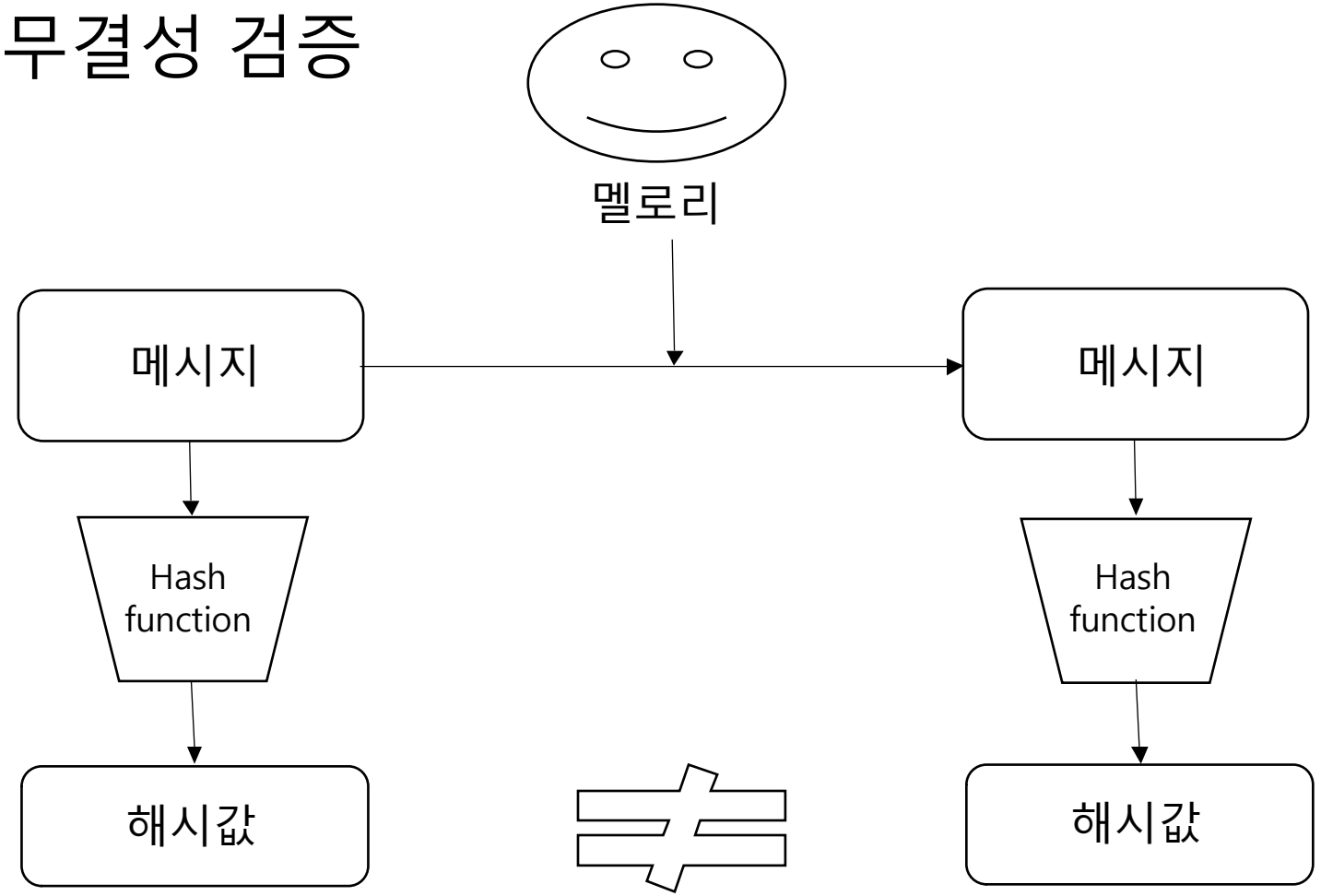
해시함수란?

- 특성 (2) : 쇄도 효과(Avalanche Effect) -> 무결성 검증, 전자서명



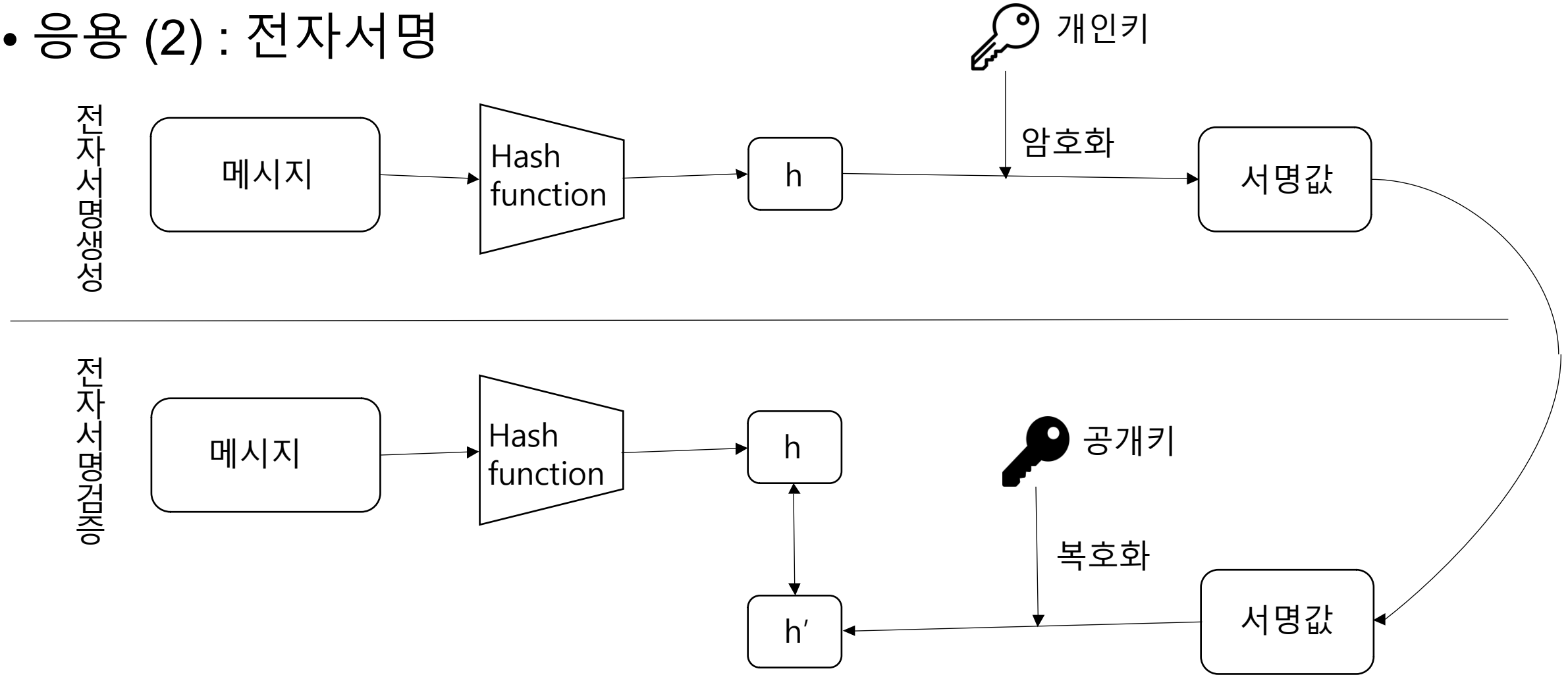
해시함수란?

- 응용 (1) : 무결성 검증



해시함수란?

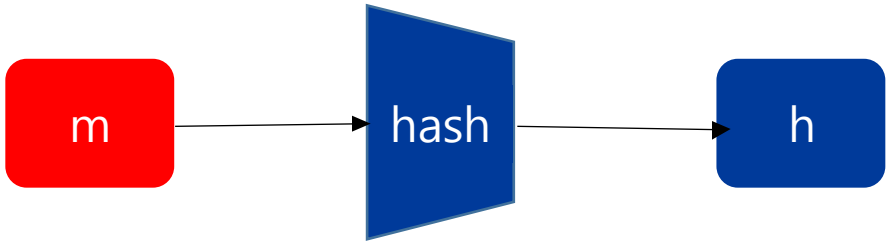
- 응용 (2) : 전자서명



해시함수 안전성

- **역상 저항성 (Preimage resistance)**
  - 어떤 해시 값이 주어졌을 때 h=H(m)를 만족하는 그 입력값 m을 찾는 것이 어렵다는 성질
  - 2^n보다 적은 복잡도

  Ex) SHA-1은 160bit의 해시값, 즉 2^{160}개의 메시지



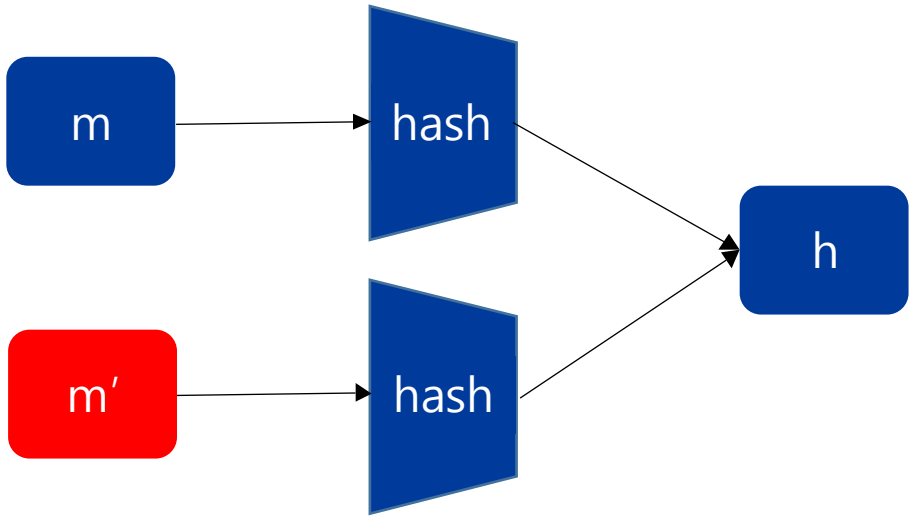
해시함수 안전성

- **제 2 역상 저항성 (Second preimage resistance)**
  - 어떤 입력값 m이 주어졌을 때, 동일한 해시 값 h가 나오는 다른 입력값 m′ 을 찾는 것이 어렵다는 성질 (즉, h=H(m) 일 때, h=H(m′) 인 m′을 찾는 것이 어렵다는 성질)

  - 2^n보다 적은 복잡도

  Ex) SHA-1은 160bit의 해시값, 즉 2^{160}개의 메시지

  - 무결성 검증, 전자서명에 중요

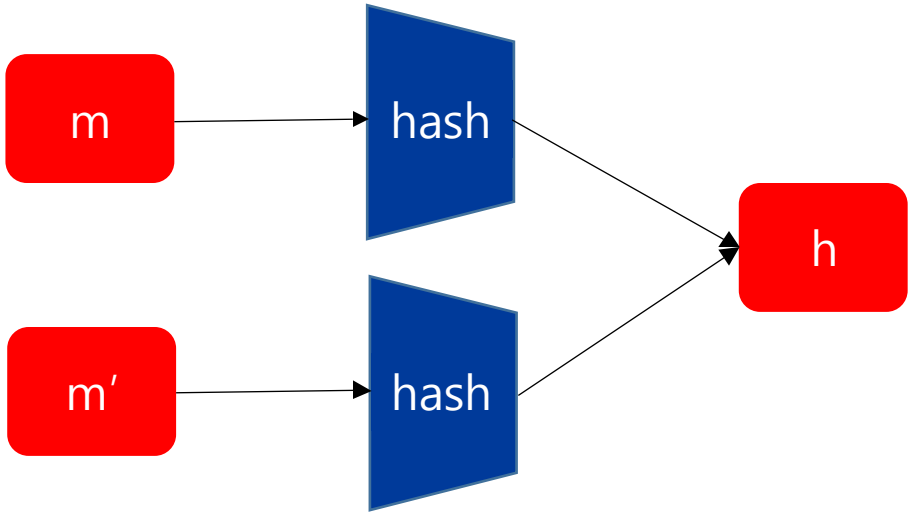


해시함수 안전성

- **충돌 저항성 (Collision resistance)**
  - 같은 해시값을 갖는 서로 다른 입력값 m, m′을 찾는 것이 어렵다는 성질 (즉, H(m)=H(m′) 을 만족하면서 m≠m 인 m, m′을 찾는 것이 어렵다는 성질)
  - $2^{n/2}$보다 적은 복잡도 (생일 역설)

  Ex) SHA-1은 160bit의 해시값, 즉 2^{80}개의 메시지

  => SHA-1 은 약 2^{60}~2^{63}의 복잡도로 충돌쌍 공격이 가능!



해시함수 안전성

- 생일 역설

$$\begin{aligned}\bar{p}(n) &= 1 \times \left(1 - \frac{1}{365}\right) \times \left(1 - \frac{2}{365}\right) \times \cdots \times \left(1 - \frac{n-1}{365}\right) \\ &= \frac{365 \times 364 \times 363 \times \cdots \times (365 - n + 1)}{365^n} \\ &= \frac{365!}{365^n (365 - n)!}\end{aligned}$$

$$p(n) = 1 - \frac{365!}{365^n (365 - n)!}$$

약 40%~50%를
넘어가면 안전하지
않다고 판단!

n	$p(n)$
1	0.0%
5	2.7%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
40	89.1%
50	97.0%
60	99.4%
70	99.9%
100	99.99997%
200	99.99999999999999999999999998%
300	$(100 - (6\times10^{-80}))\%$
350	$(100 - (3\times10^{-129}))\%$
365	$(100 - (1.45\times10^{-155}))\%$
366	100%
367	100%

해시함수 안전성

- 생일 역설

임의의 어느 두 사람의 생일이 다를 확률은

$$\Rightarrow \frac{365}{365} * \frac{364}{365} = \frac{365 * 364}{365^2} = \frac{364}{365}$$

n명 중에 임의의 두 사람을 짝지을 수 있는 경우의 수는

$$\Rightarrow {}_nC_2$$

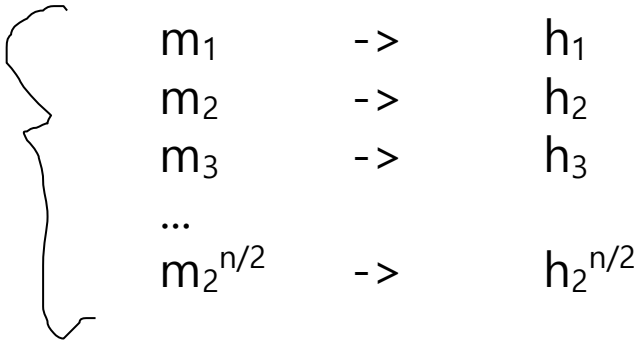
즉, n명 중에 임의의 두 사람을 짝 지었을 때 생일이 같을 확률은

$$\Rightarrow p(n) = 1 - \frac{364^{nC2}}{365}$$

해시함수 안전성

- 충돌쌍 공격의 복잡도가 $2^{n/2}$인 이유

$2^{n/2}$개



$m_1 \rightarrow h_1$
$m_2 \rightarrow h_2$
$m_3 \rightarrow h_3$
…
$m_{2^{n/2}} \rightarrow h_{2^{n/2}}$

For $1 < i < j < 2^{n/2}$
$h_i = h_j$

해시함수 안전성

- 충돌쌍 공격의 복잡도가 $2^{n/2}$인 이유

임의의 어느 두 해시값이 다를 확률은

$$=> \frac{2^n}{2^n} * \frac{2^n-1}{2^n} = \frac{2^n-1}{2^n} = 1+\frac{1}{-2^n}$$

임의의 두 해시값을 짝지을 수 있는 경우의 수는

$$=> {}_{2^{n/2}}C_2 = 2^{n-1}$$

즉, n개의 해시값에서 임의의 두 해시값을 짝 지었을 때 해시값이 다를 확률은

$$=> \bar{p}(n) = \left(1+\frac{1}{-2^n}\right)^{2^{n-1}} = \left(\left(1+\frac{1}{-2^n}\right)^{-2^n}\right)^{-2^{-1}} = e^{-1/2}$$

n개의 해시값에서 임의의 두 해시값을 짝 지었을 때 해시값이 같을 확률은

$1 - e^{-1/2}$ = 약 0.39

전용 해시함수

- 해시함수로써 사용되기 위해서 만들어진 함수
  Ex) MD4, MD5, SHA-series(SHA-0, 1, 256, 384, 512)

	SHA-1	SHA-256	SHA-384	SHA-512
메시지 다이제스트 길이 (해시값)	160	256	384	512
메시지 최대 길이	2^{64}	2^{64}	2^{128}	2^{128}
블록 길이	512	512	1024	1024
단어 길이	32	32	64	64
단계 수	80	64	80	80
충돌쌍 공격에 대한 복잡도	2^{80}	2^{128}	2^{192}	2^{256}

전용 해시함수

- SHA-1

  $\{0, 1\}^* \rightarrow \{0, 1\}^{160}$

  (단, $* < 2^{64}$)

	SHA-1
메시지 다이제스트 길이 (해시값)	160
메시지 최대 길이	2^{64}
블록 길이	512
단어 길이	32
단계 수	80
충돌쌍 공격에 대한 복잡도	2^{80}

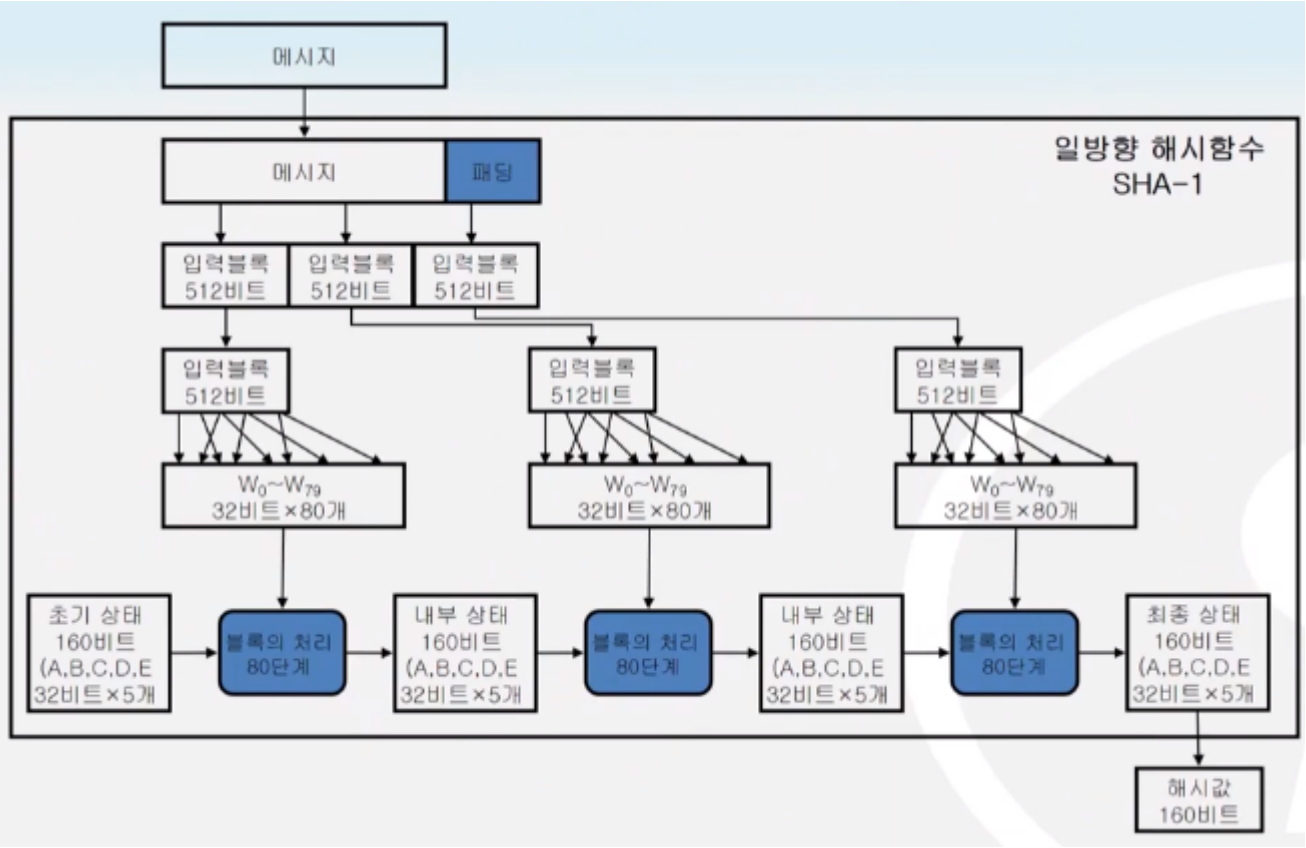
전용 해시함수

- SHA-1 과정
  (1) 패딩
  M = M||100..000|M의 길이정보|
  => 512의 최소의 배수
  (2) W_0 ~ W_{79} 계산
  (3) 블록 처리
  (4) 단계 별 처리

	SHA-1
메시지 다이제스트 길이 (해시값)	160
메시지 최대 길이	2^{64}
블록 길이	512
단어 길이	32
단계 수	80
충돌쌍 공격에 대한 복잡도	2^{80}

전용 해시함수

- SHA-1 구조 (Merkle-Damgard 구조)

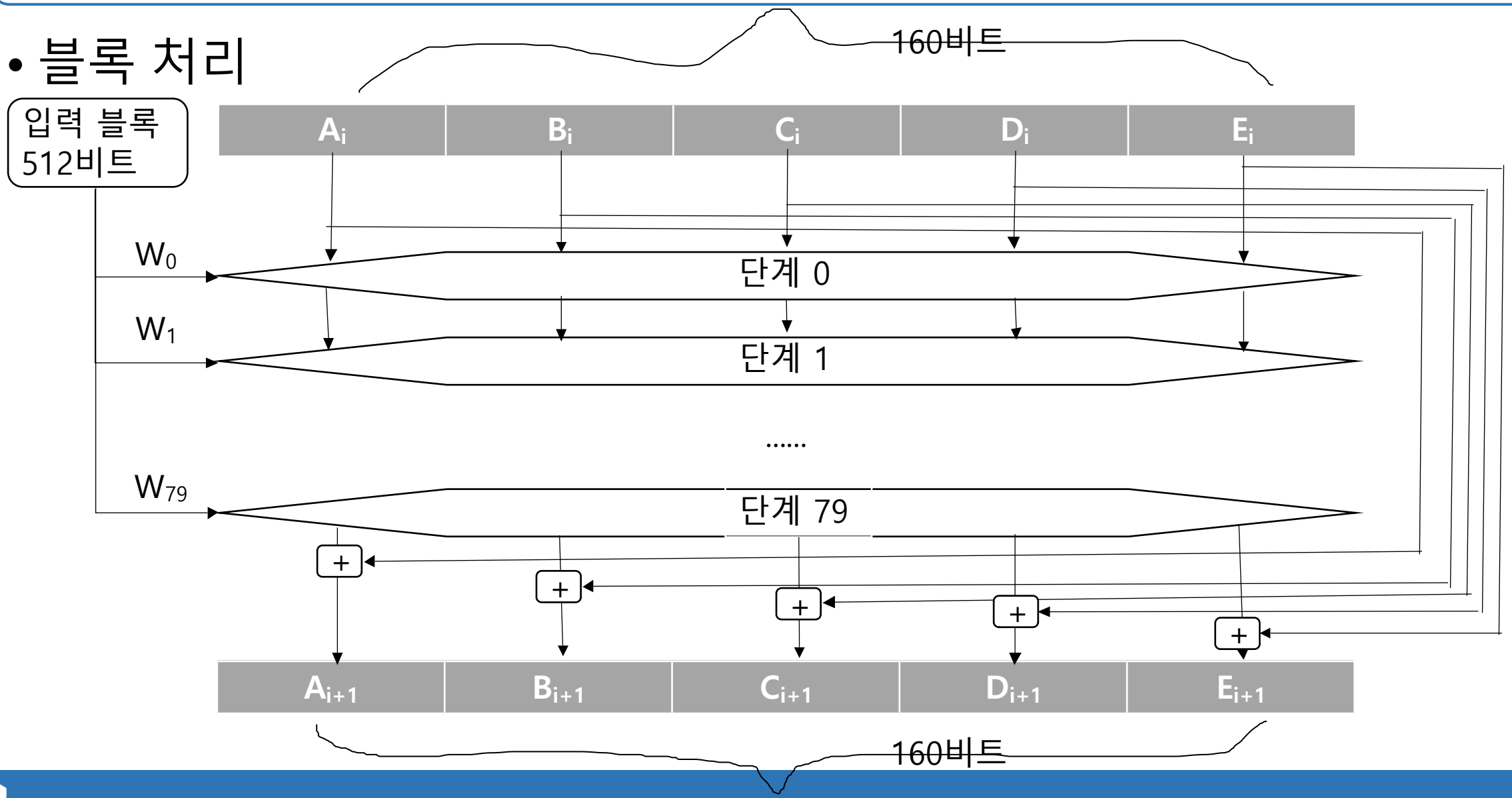


	SHA-1
메시지 다이제스트 길이 (해시값)	160
메시지 최대 길이	2^{64}
블록 길이	512
단어 길이	32
단계 수	80
충돌쌍 공격에 대한 복잡도	2^{80}

$$W_t = (W_{t-16} \oplus W_{t-14} \oplus W_{t-8} \oplus W_{t-3}) <<1$$

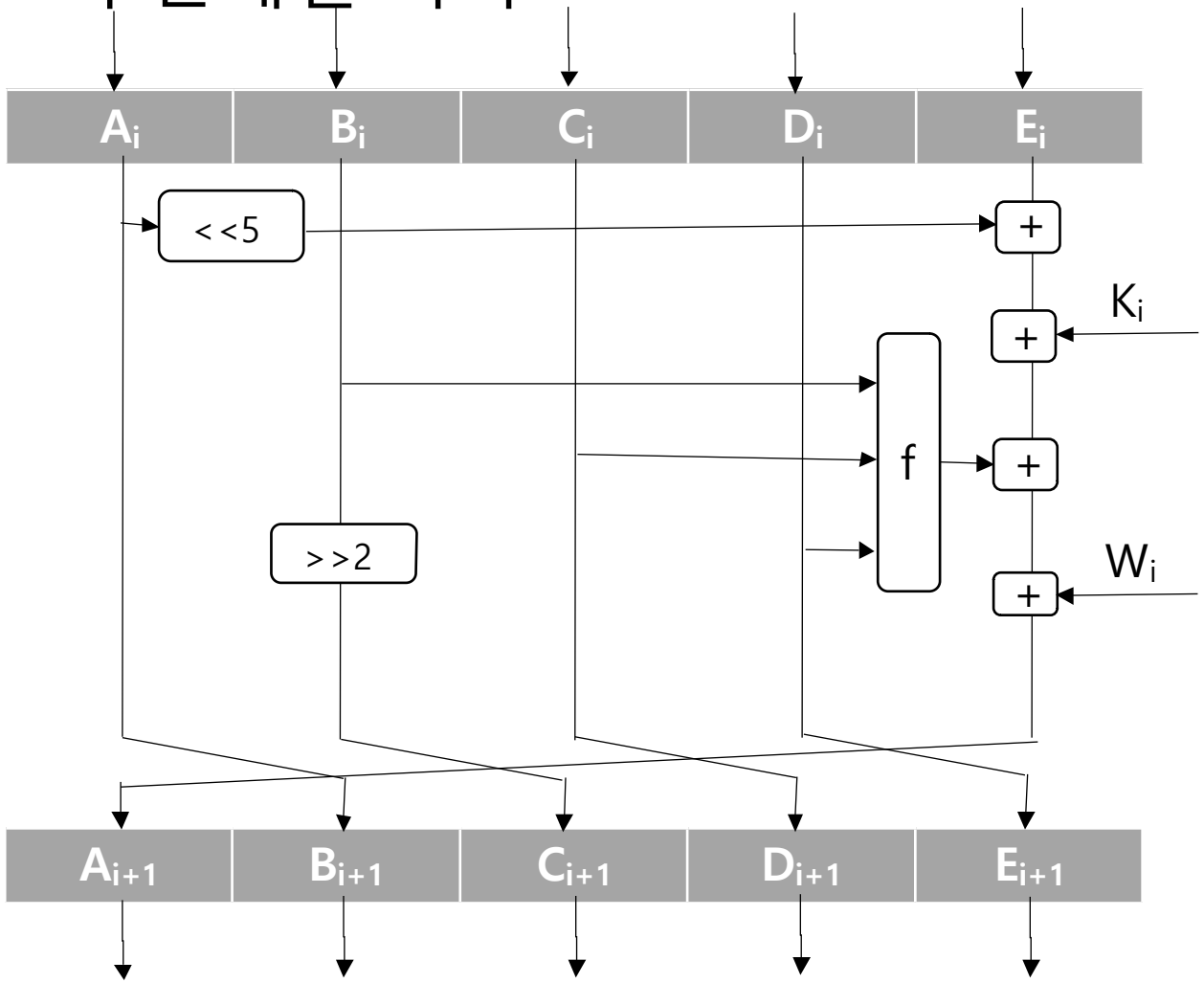
전용 해시함수

- 블록 처리


160비트
입력 블록
512비트
A_i
B_i
C_i
D_i
E_i
W_0
단계 0
W_1
단계 1
.....
W_{79}
단계 79
+
+
+
+
+
A_{i+1}
B_{i+1}
C_{i+1}
D_{i+1}
E_{i+1}
160비트

전용 해시함수

- 각 단계별 처리



A_0 = 67 45 23 01
B_0 = EF CD AB 89
C_0 = 98 BA DC FE
D_0 = 10 32 54 76
E_0 = C3 D2 E1 F0

$K_0 \sim K_{19}$ = 5A 82 79 99
$K_{20} \sim K_{39}$ = 6E D9 EB A1
$K_{40} \sim K_{59}$ = 8F 1B BC DC
$K_{60} \sim K_{79}$ = CA 62 C1 D6

$f_0 \sim f_{19}$ = (B and C) or (not B and D)
$f_{20} \sim f_{39}$ = B xor C xor D
$f_{40} \sim f_{59}$ = (B and C) or (C and D) or (D and B)
$f_{60} \sim f_{79}$ = B xor C xor D

Q & A