2nd Cryptographic Hash Workshop (2006/8/24-25, Santa Barbara, California)

How to Construct Double-Block-Length Hash Functions

Shoichi Hirose

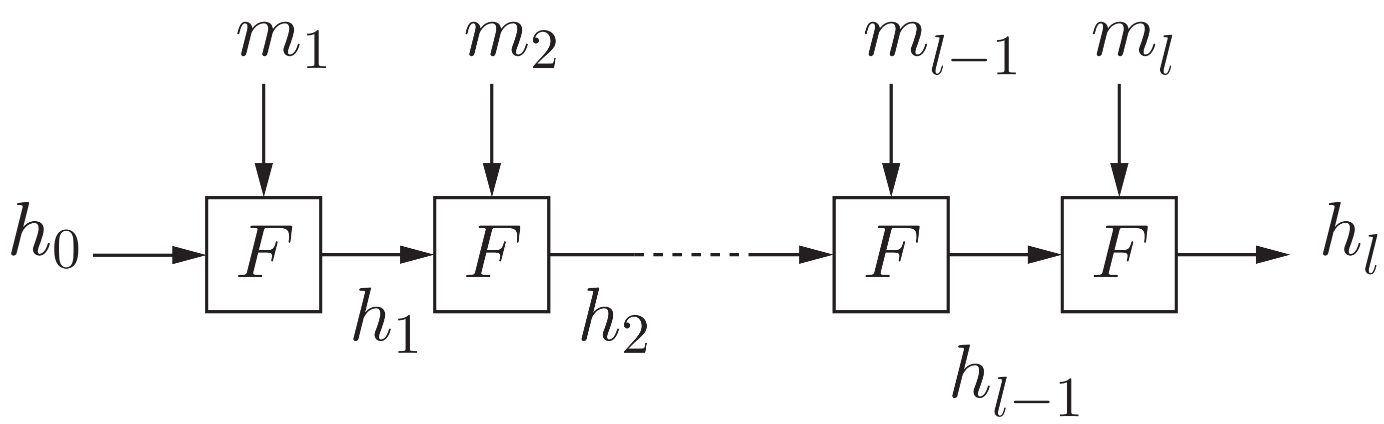
University of Fukui, Japan

24th Aug. 2006

Iterated Hash Function

- Compression function
  $F : \{0,1\}^{\ell} \times \{0,1\}^{\ell'} \longrightarrow \{0,1\}^{\ell}$
- Initial value $h_0 \in \{0,1\}^{\ell}$

Input $m = (m_1, m_2, \ldots, m_l)$, $m_i \in \{0,1\}^{\ell'}$ for $1 \leq i \leq l$



$H(m) = h_l$

Motivation

How to construct a compression function using a smaller component?

E.g.) Double-block-length (DBL) hash function

- The component is a block cipher.
- output-length $= 2 \times$ block-length
- abreast/tandem Davies-Meyer, MDC-2, MDC-4, ...

Cf.) Any single-block-length HF with AES is not secure.

- Output length is 128 bit.
- Complexity of birthday attack is $O(2^{64})$.

Result

- Some plausible DBL HFs
  - Composed of a smaller compression function
    - $F(x) = (f(x), f(p(x)))$
      p is a permutation satisfying some properties
    - Optimally collision-resistant (CR) in the random oracle model
  - Composed of a block cipher with key-length $>$ block-length
    - AES with 192/256-bit key-length
    - Optimally CR in the ideal cipher model
- A new security notion: Indistinguishability in the iteration

Def. (optimal collision resistance)
Any collision attack is at most as efficient as a birthday attack.

Related Work on Double-Block-Length Hash Function

- Lucks 05
  - $F(g,h,m) = (f(g,h,m), f(h,g,m))$
  - Optimally CR if f is a random oracle
- Nandi 05
  - $F(x) = (f(x), f(p(x)))$, where p is a permutation
  - Optimally CR schemes if f is a random oracle

Other Related Work

Single block-length

- Preneel, Govaerts and Vandewalle 93
  PGV schemes and their informal security analysis

- Black, Rogaway and Shrimpton 02
  Provable security of PGV schemes in the ideal cipher model

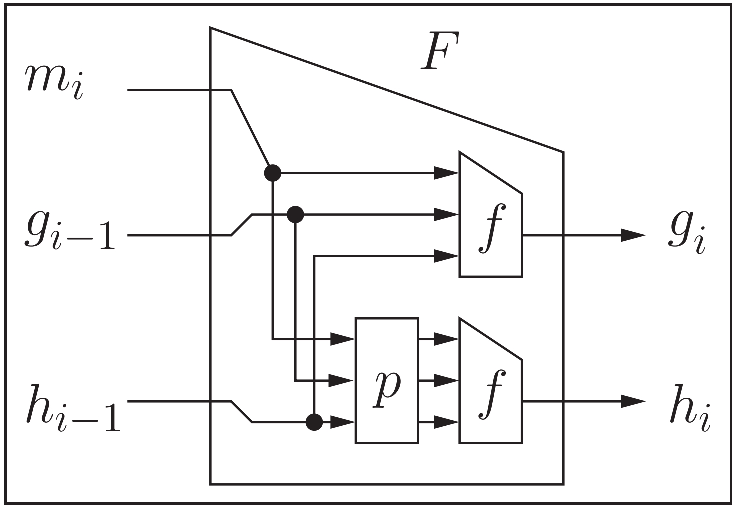
Double block-length

- Satoh, Haga and Kurosawa 99
  Attacks against rate-1 HFs with a $(n, 2n)$ block cipher

- Hattori, Hirose and Yoshida 03
  No optimally CR rate-1 parallel-type CFs with a $(n, 2n)$ block cipher

DBL Hash Function Composed of a Smaller Compression Function

- f is a random oracle
- p is a permutation
  - Both p and p^{-1} are easy
  - $p \circ p$ is an identity permutation



$$F(x) = (f(x), f(p(x)))$$
$$F(p(x)) = (f(p(x)), f(x))$$

$f(x)$ and $f(p(x))$ is only used for $F(x)$ and $F(p(x))$.

We can assume that an adversary asks x and $p(x)$ to f simultaneously.

Collision Resistance

Th. 1 Let $F : \{0,1\}^{2n+b} \to \{0,1\}^{2n}$ and $F(x) = (f(x), f(p(x)))$.

Let H be a hash function composed of F.

Suppose that

- $p(p(\cdot))$ is an identity permutation
- p has no fixed points: $p(x) \neq x$ for $\forall x$

$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \stackrel{\mathrm{def}}{=} \text{success prob. of the optimal collision finder for } H$$
$$\text{which asks } q \text{ pairs of queries to } f.$$

Then, in the random oracle model, $\mathbf{Adv}_H^{\mathrm{coll}}(q) \leq \frac{q}{2^n} + \left(\frac{q}{2^n}\right)^2$.

Note) MD-strengthening is assumed in the analysis.

Proof Sketch

F is CR $\Rightarrow$ H is CR

Two kinds of collisions:

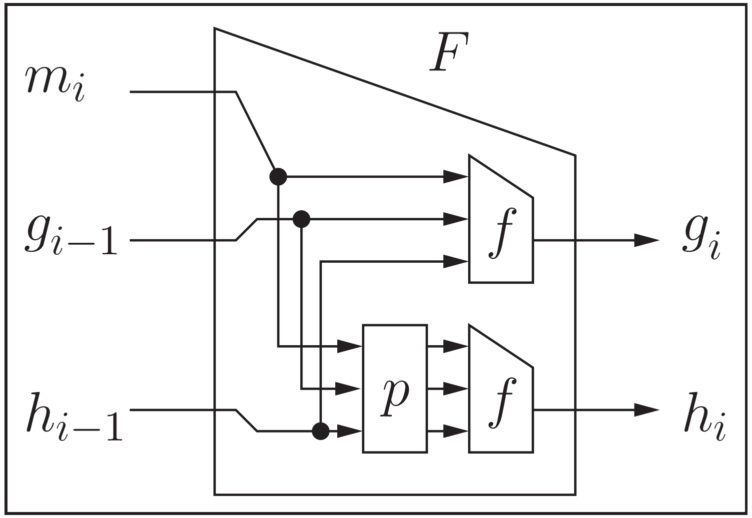
$$\Pr[F(x) = F(x') \mid x' = p(x)]$$
$$= \Pr[f(x) = f(x') \wedge f(p(x)) = f(p(x'))] = \left(\frac{1}{2^n}\right)^2$$
$$\Pr[F(x) = F(x') \mid x' = p(x)] = \Pr[f(x) = f(p(x))] = \frac{1}{2^n}$$

The collision finder asks q pairs of queries to f: x_j and $p(x_j)$ for $1 \le j \le q$.

$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \le \frac{q}{2^n} + \left(\frac{q}{2^n}\right)^2$$

Collision Resistance: A Better Bound

Th. 2 Let H be a hash function composed of $F : \{0,1\}^{2n+b} \rightarrow \{0,1\}^{2n}$. Suppose that



- $p(p(\cdot))$ is an identity permutation
- $p(g,h,m) = (p_{\mathrm{cv}}(g,h), p_{\mathrm{m}}(m))$
  - p_{cv} has no fixed points
  - $p_{\mathrm{cv}}(g,h) = (h,g)$ for $\forall(g,h)$

Then, in the random oracle model,

$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \leq 3\left(\frac{q}{2^n}\right)^2$$

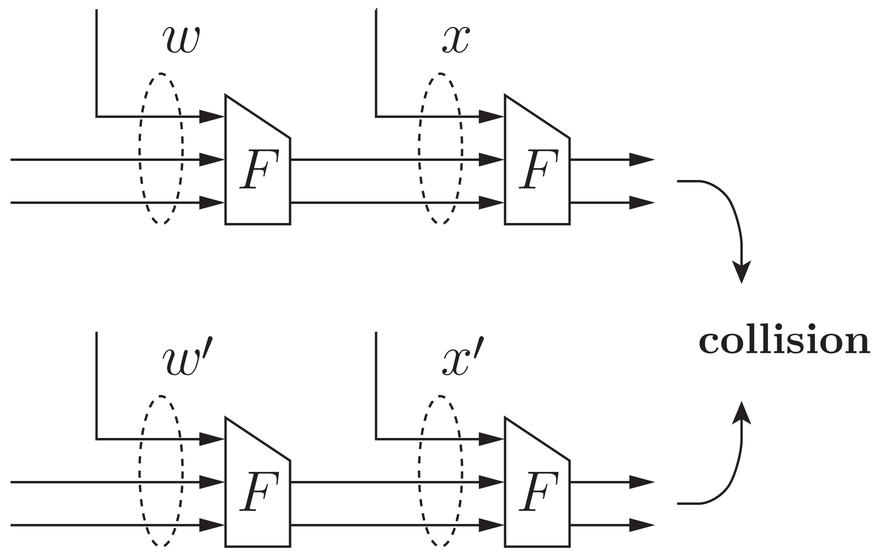
Proof Sketch

Two kinds of collisions:

$$\Pr[F(x) = F(x') \mid x' = p(x)] = \left(\frac{1}{2^n}\right)^2$$

$$\Pr[F(x) = F(x') \mid x' = p(x)] = \frac{1}{2^n}$$

However,



$$F(x) = F(x') \wedge x' = p(x) \Rightarrow F(w') = p_{\mathrm{cv}}(F(w)) \wedge w' = p(w)$$

$$\Pr[F(w') = p_{\mathrm{cv}}(F(w)) \mid w' = p(w)] = \left(\frac{1}{2^n}\right)^2$$

$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \leq 3\left(\frac{q}{2^n}\right)^2 = \left(\frac{q}{2^n}\right)^2 + 2\left(\frac{q}{2^n}\right)^2$$

Th. 1 vs. Th. 2

The difference between the upper bounds is significant.

E.g.) $n = 128$, $q = 2^{80}$

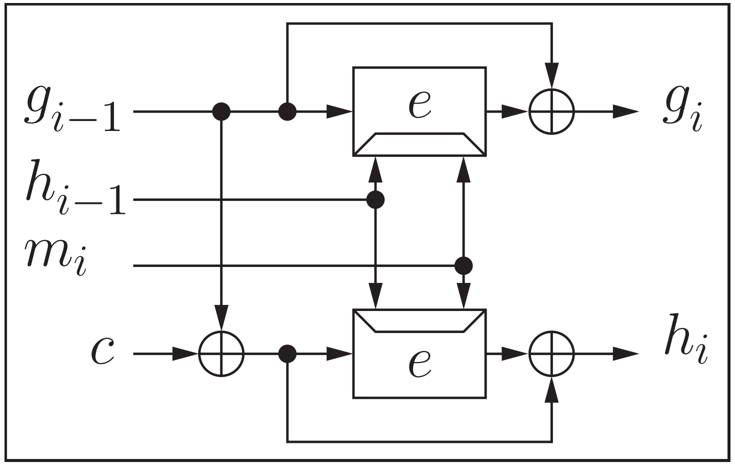
$$\textbf{Th. 1} \qquad \mathbf{Adv}_H^{\mathrm{coll}}(q) \le \frac{q}{2^n} + \left(\frac{q}{2^n}\right)^2 \approx 2^{-48}$$

$$\textbf{Th. 2} \qquad \mathbf{Adv}_H^{\mathrm{coll}}(q) \le 3\left(\frac{q}{2^n}\right)^2 \approx 2^{-94}$$

E.g.) A permutation p satisfying the properties in $\textbf{Th. 2}$

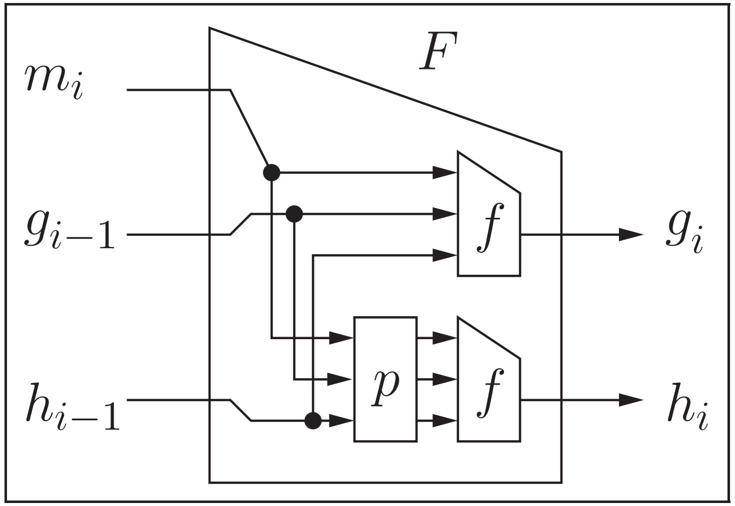
$$p(g, h, m) = (g \oplus c_1, h \oplus c_2, m), \text{ where } c_1 = c_2$$

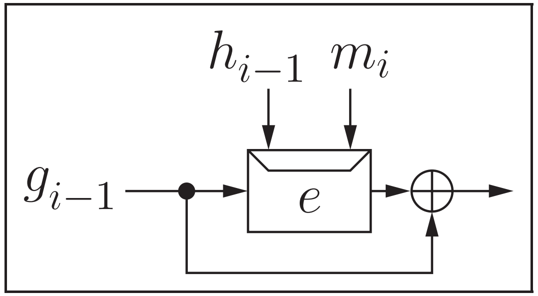
DBL Hash Function Composed of a Block Cipher

$F =$ 

c is a non-zero constant.

Cf.)

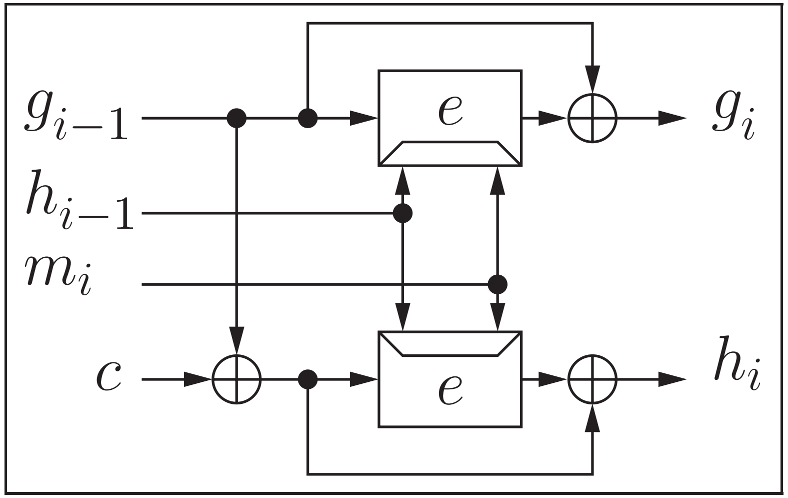


such that $f =$ 

$$p(g, h, m) = (g \oplus c, h, m)$$

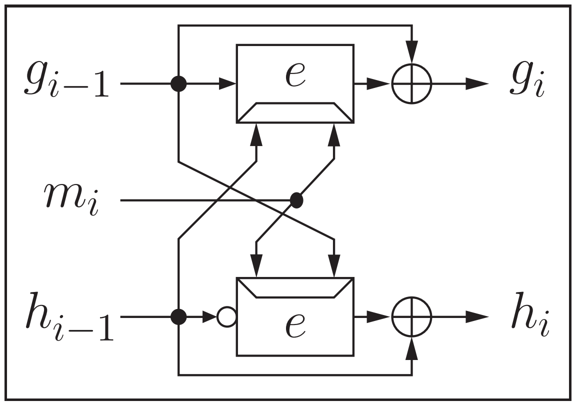
DBL Hash Function Composed of a Block Cipher

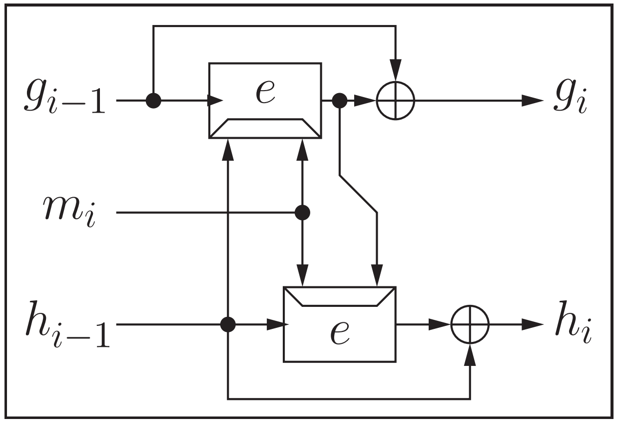
$F =$



- can be constructed using AES with 192/256-bit key
- requires only one key scheduling

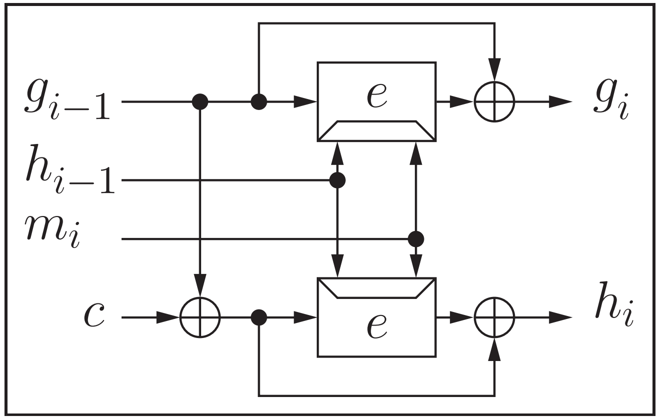
F is simpler than abreast Davies-Meyer and tandem Davies-Meyer





Collision Resistance

Th. 3 Let H be a HF composed of $F : \{0,1\}^{2n+b} \to \{0,1\}^{2n}$ such that

$F =$ 

.

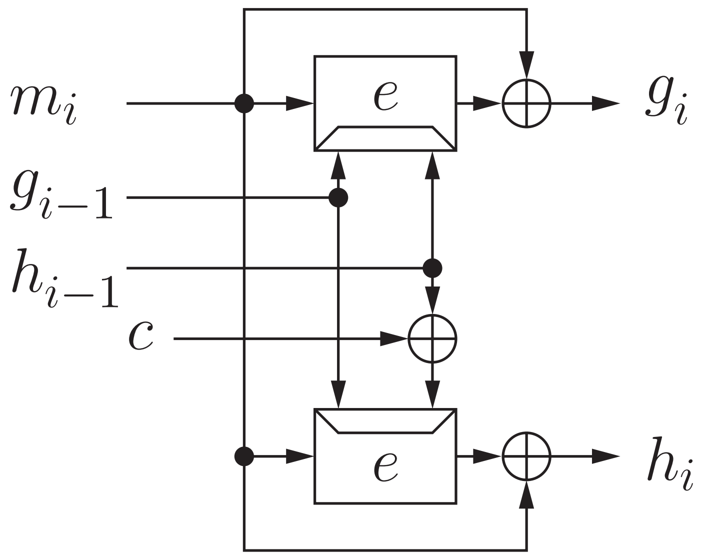
$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \stackrel{\mathrm{def}}{=} \text{success prob. of the optimal collision finder for } H$$

which asks q pairs of queries to (e, e^{-1}).

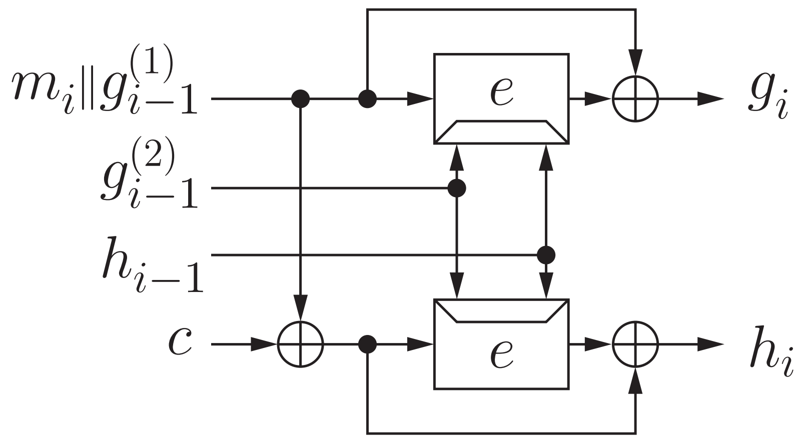
Then, in the ideal cipher model, for $1 \le q \le 2^{n-2}$,

$$\mathbf{Adv}_H^{\mathrm{coll}}(q) \le 3\left(\frac{q}{2^{n-1}}\right)^2$$

A Few More Examples of Compression Functions



for AES with 256-bit key

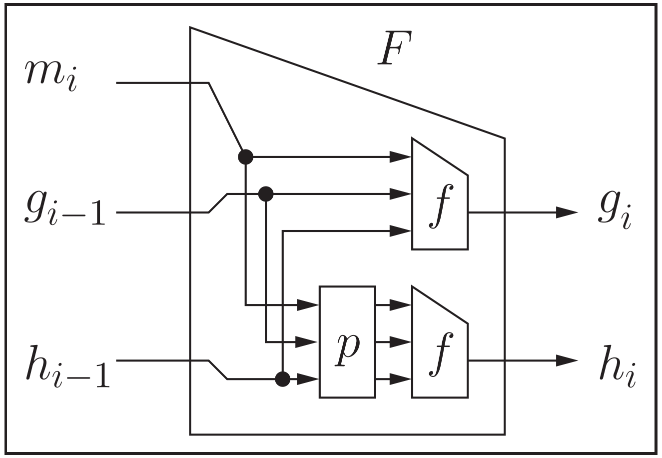


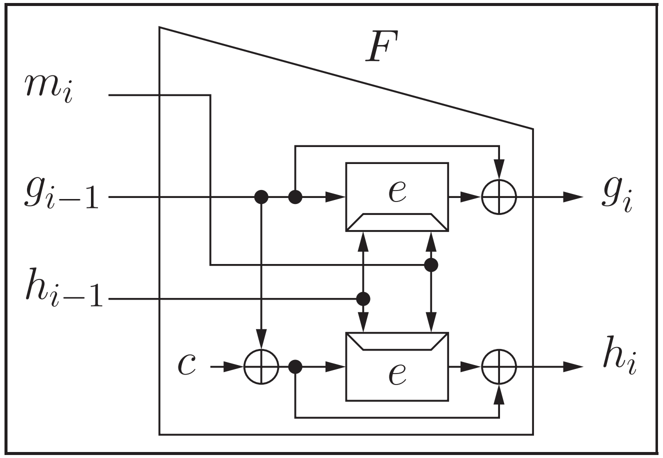
for AES with 192-bit key

Conclusion

- Some plausible DBL HFs
  - composed of

    a smaller compression function   or   a block cipher





$p \circ p$ is an identity permutation   key-length > block-length

  - optimally collision-resistant

- A new security notion: Indistinguishability in the iteration