MD5 to be considered harmful today.

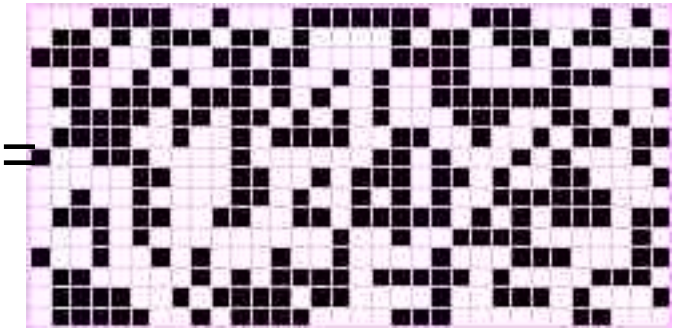
(should we say 10 years ago?)

Timeline

- 1992 – MD5 invented. Ronald Rivest.
- 1993 – Differential Cryptanalysis. Biham, Shamir.
- 1993 – First attacks. den Boer, Bosselaers.
- 1996 – Further attacks. Hans Dobbertin.
- 8/2004 – MD5 collisions presented. Wang et. al.
- 10/2004 – Method reversengineered. Hawkes et. al.
- 5/3/2005 – Method reengineered. Vlastimil Klima
- 6/3/2005 – Original method released. Wang et. al.

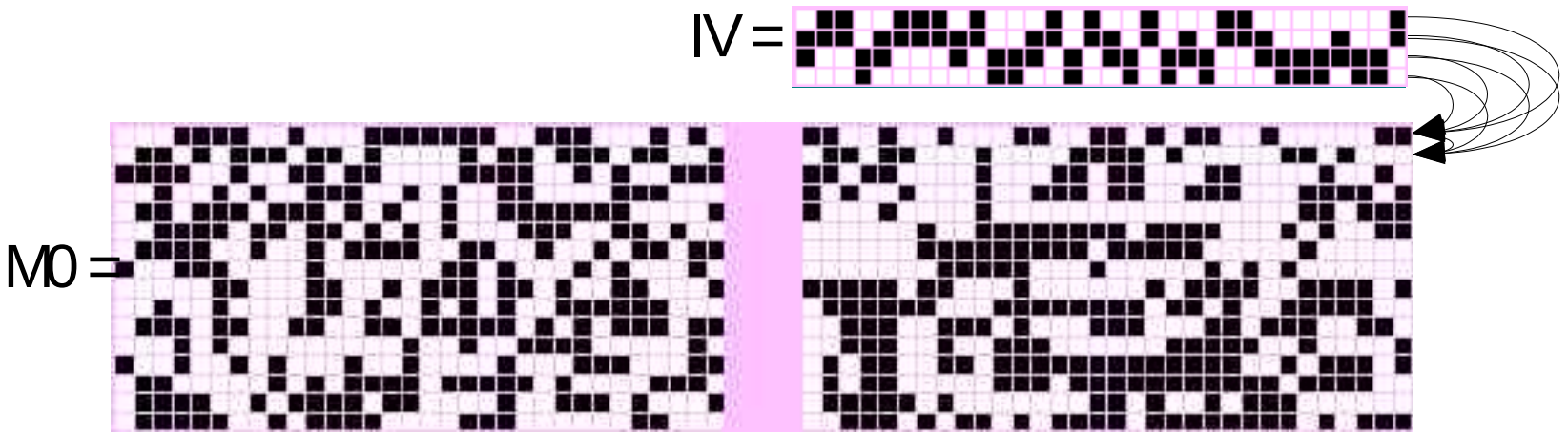
MD5 explained – 512 bits input per round

M0 =



- (white) = bit is 0
- (black) = bit is 1

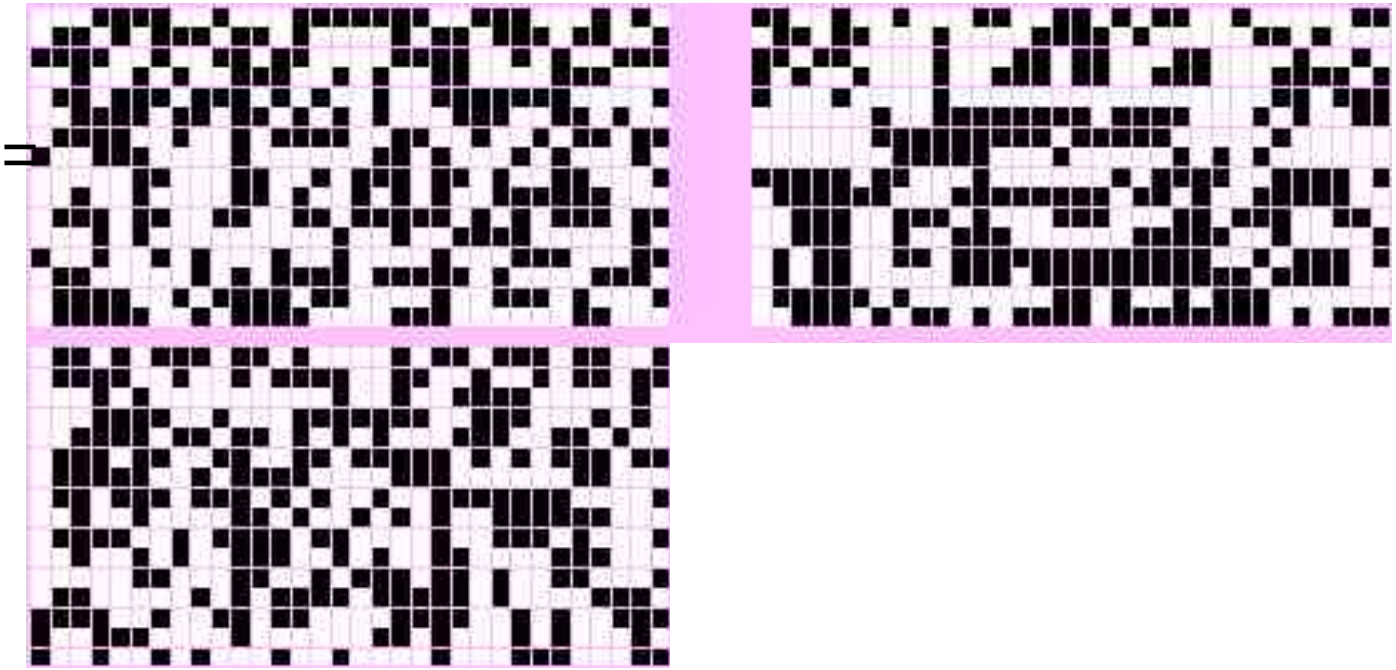
MD5 explained – computing internal state



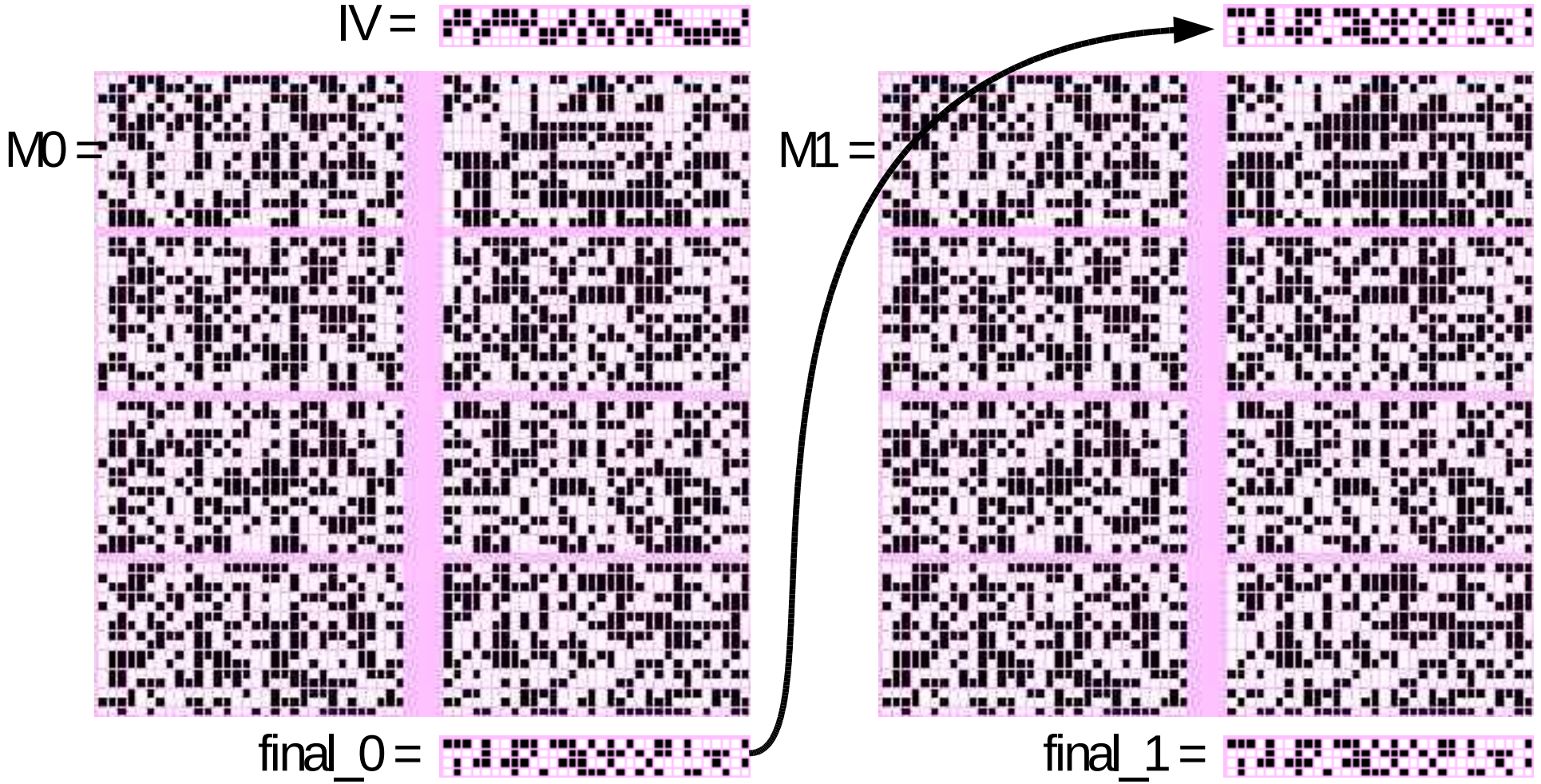
MD5 explained – input is used again

IV =

M0 =



MD5 explained – second 512 bits use final_0



M – Wang's c

wang_0 =

wang_1 =

MD5 explain[illegible]ons

wang_0 ≈ wang_1

MD5 explained – Appending data

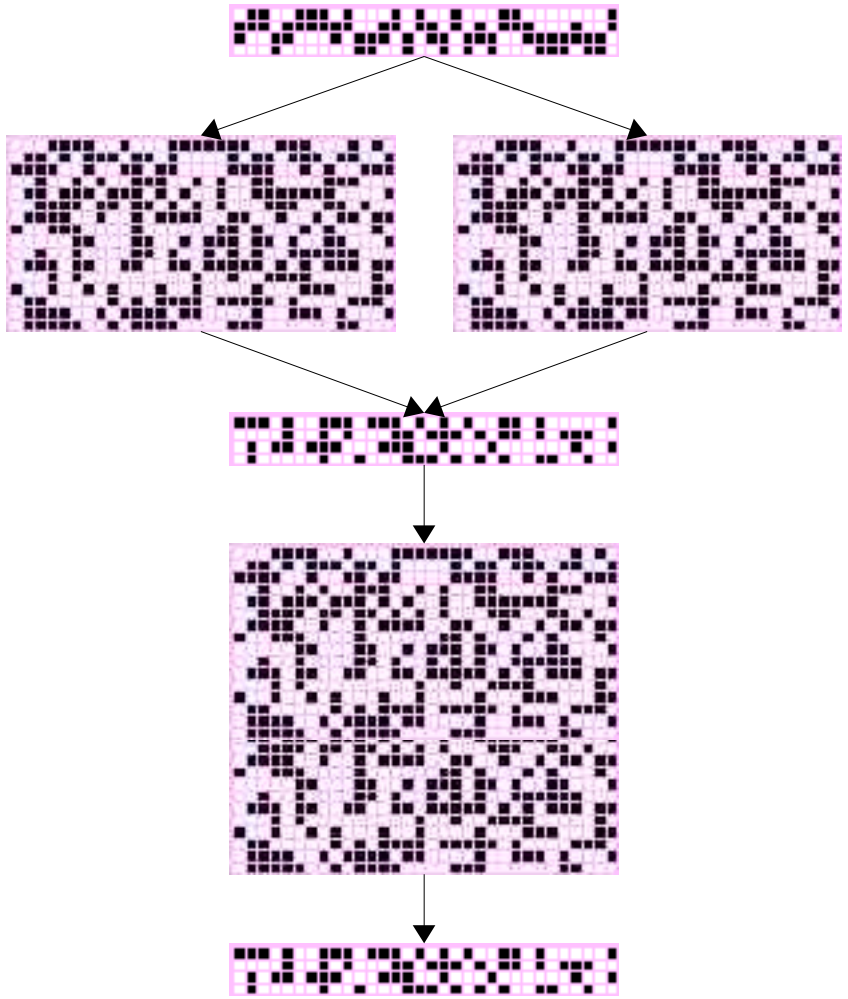
MD5(wang_0) = MD5(wang_1)

⇕

final(wang_0) = final(wang_1)

⇓

final(wang_0 | X) = final(wang_1 | X)

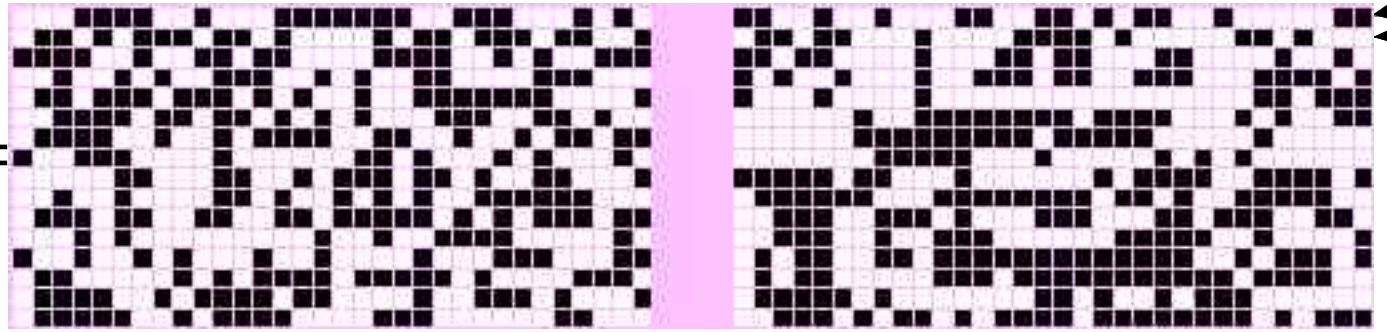


MD5 explained – Appending data

- 8/2004 – MD5 collisions presented. Wang et. al.
- 12/2004 – Kaminsky. MD5 harmful someday.
  - Custom made packer/unpacker
  - HTML/javascript page

MD5 explained – Prepending data

IV =



wang_0 =

MD5(X | wang_0) $\stackrel{?}{=}$ MD5(X | wang_1)

$\Uparrow$

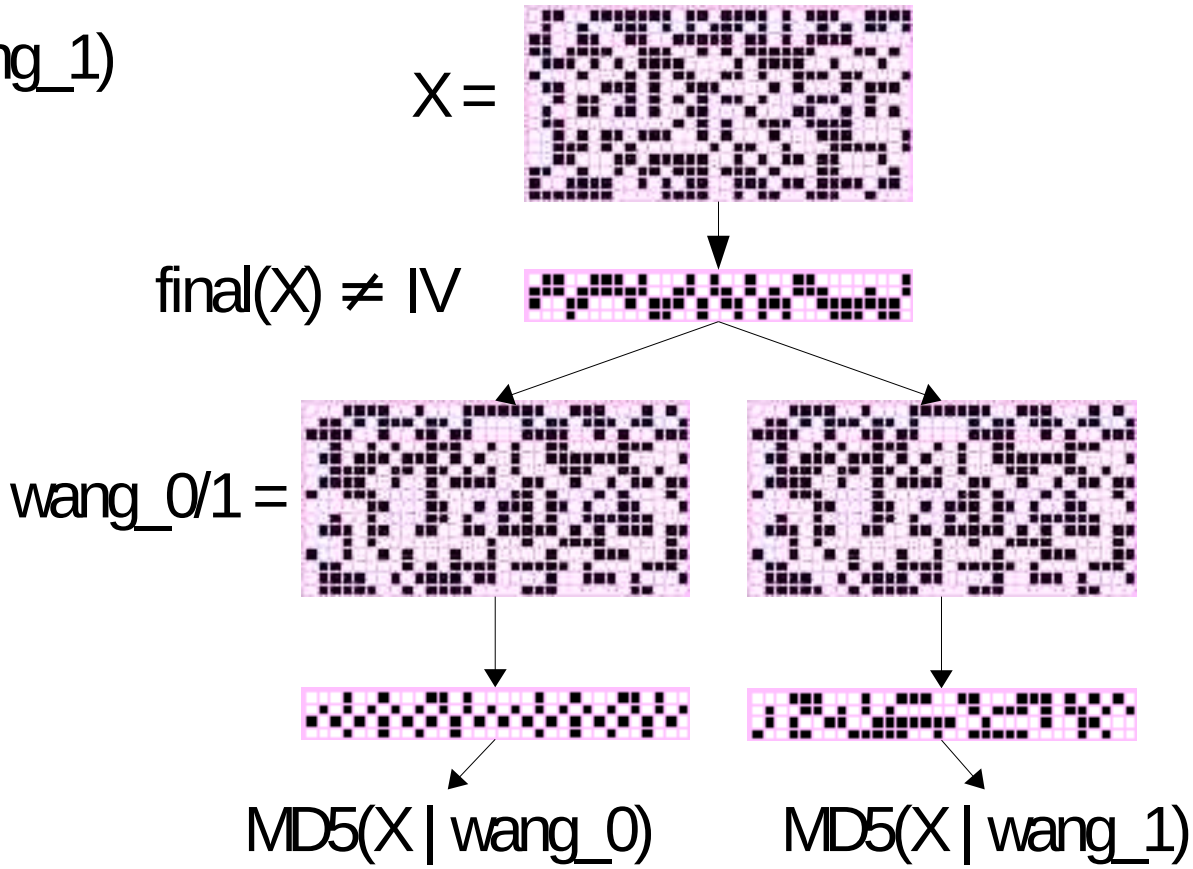
final(X) = IV

MD5 explained – Prepending data

$$MD5(X \mid wang\_0) \stackrel{?}{=} MD5(X \mid wang\_1)$$

$$\Updownarrow$$

$$final(X) = IV$$



MD5 explained – Prepending data

- 8/2004 – MD5 collisions presented. Wang et. al.
  - Method to find MD5 collisions
  - for arbitrary IV
- 10/2004 – Method reversengineered. Hawkes et. al.
- 2 weeks ago – Attack reimplemented.
- Two/Three days ago – Tool released. Patrick Stach

MD5 explained – Prepending data

MZ Header

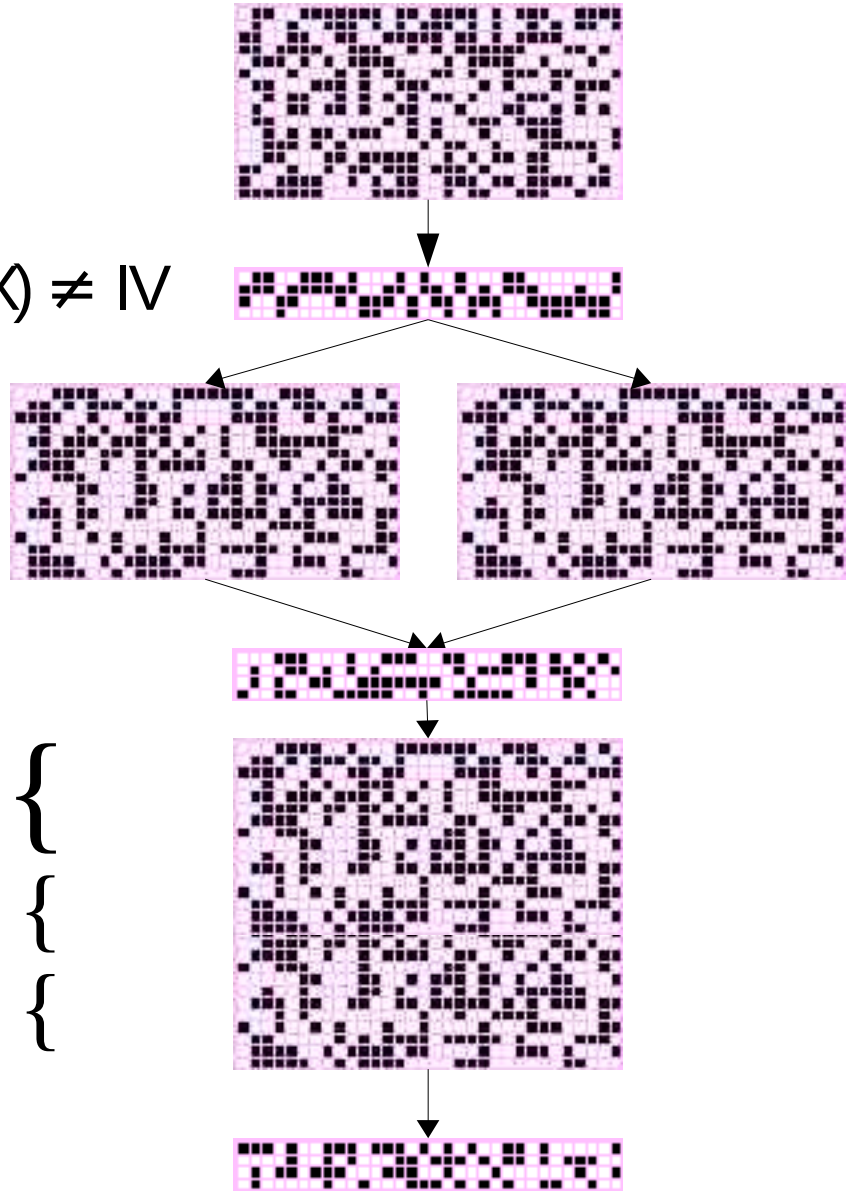
PE Program

Multiplexor code

.exe #1

.exe #2

final(X) ≠ IV



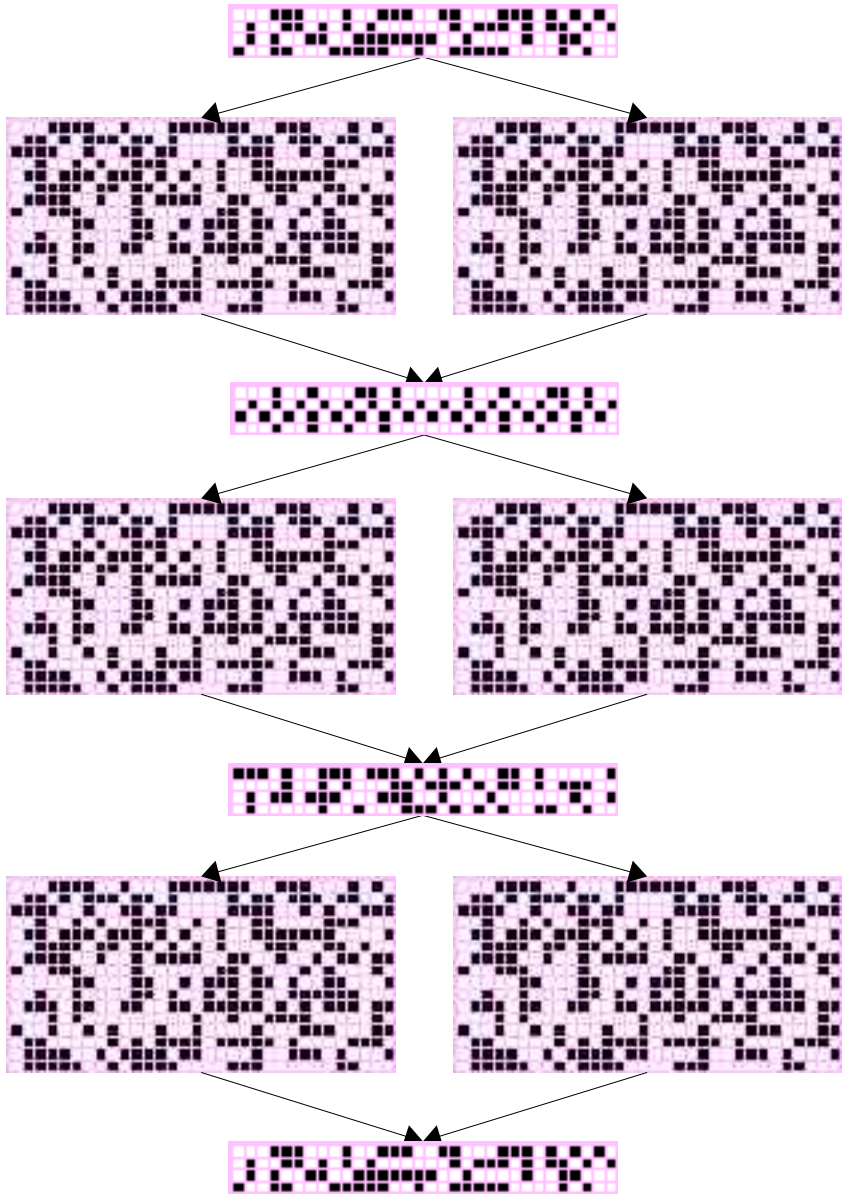
MD5 explained – 2^n Collisions

8 colliding blocks

1st collision

2nd collision

3rd collision



Other hash algorithms

- MD4 – broken in 2004 by Wang et. al.
- RIPEMD – broken in 2004 by Wang et. al.
- SHA-0 – broken in 2004 by Joux et. al.
  - Improved by Wang et. al.
- Haval-128 – broken in 2004 by Wang et. al.
- SHA-1 – Theoretically broken by Wang et. al.
  - 2/2005 – needs less than 2^{69} SHA-1 computations.
  - 8/2005 – needs less than 2^{63} SHA-1 computations.

Other hash algorithms

- SHA-256 – currently suggested as secure.
- Use 2 different algorithms at the same time.
  - MD5 + CRC32 – May be broken with 32 collisions?
  - No theoretical proof of better security.

MD5 to be considered harmful 30 years ago

Questions!?

Gerardo Richarte
gerardo.richarte@coresecurity.com

¡GRACIAS!

Gerardo Richarte
gerardo.richarte@coresecurity.com