Short Chosen-Prefix Collisions for MD5 and the Creation of a Rogue CA Certificate

Marc Stevens *CWI, Netherlands*

Alexander Sotirov *New York, USA*

Jacob Appelbaum *Noisebridge/Tor, SF*

Arjen Lenstra *EPFL, Switzerland & Alcatel-Lucent Bell Labs*

David Molnar *UC Berkeley, USA*

Dag Arne Osvik *EPFL, Switzerland*

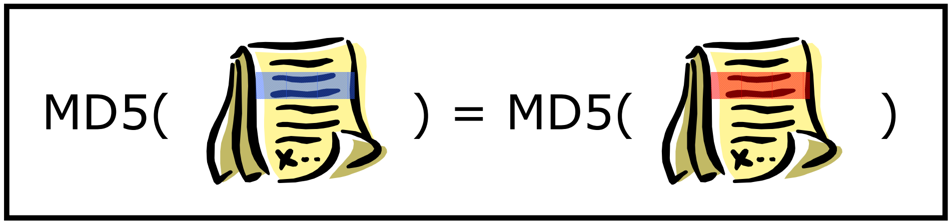
Benne de Weger *TU/e, Netherlands*

Collisions for MD5

2004: First collision for MD5 [Wang,Yu]:

  - Two 128 byte messages with same MD5 hash value

- *Identical prefix* collision attack
  - Messages differ only in 128 consecutive ‘random’ bytes
  - Bytes before or after may not differ
  - Currently: <1 sec on single pc core

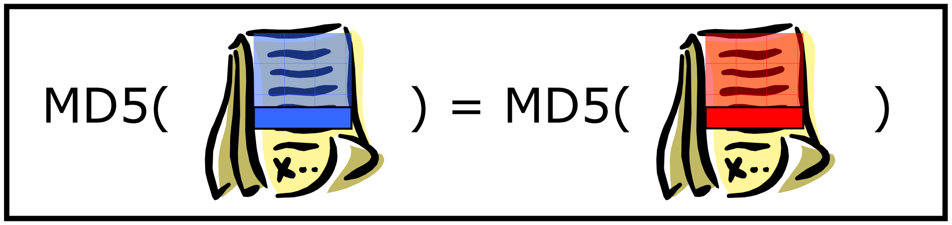

MD5( ) = MD5( )

- Same MD5 hash value ⇒ same signature

Chosen-Prefix Collisions

2006: *Chosen-prefix* collision (CPC) attack

- [Stevens, Lenstra, de Weger]
  - New stronger type of collisions
  - Choose two arbitrary files (same length)
  - Make them collide by appending 716 ‘random’ bytes
  - Currently: 1 day on quad-core pc w/ only 588 bytes

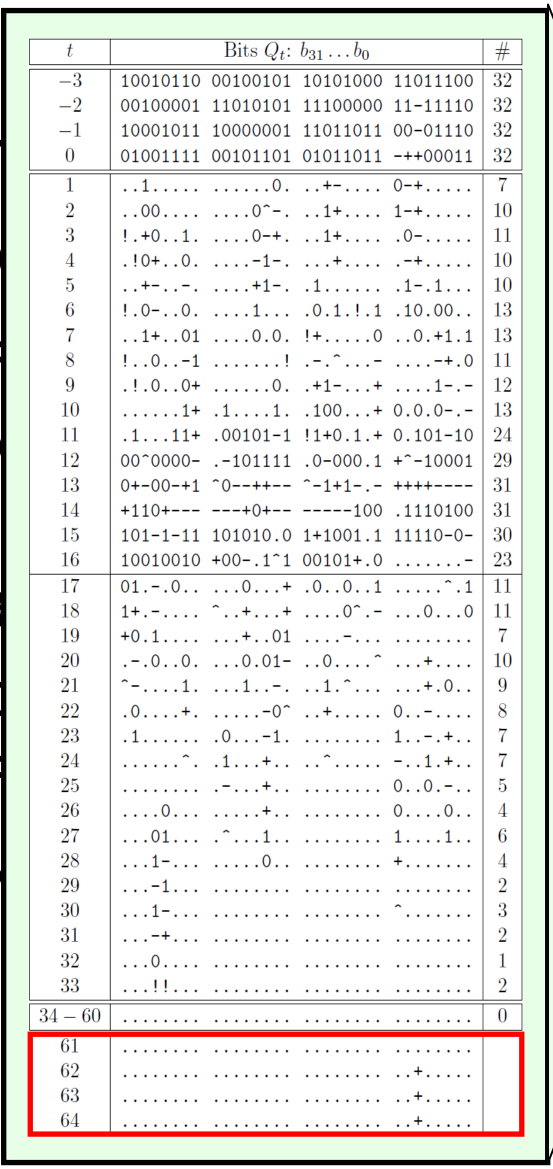


MD5( ) = MD5( )

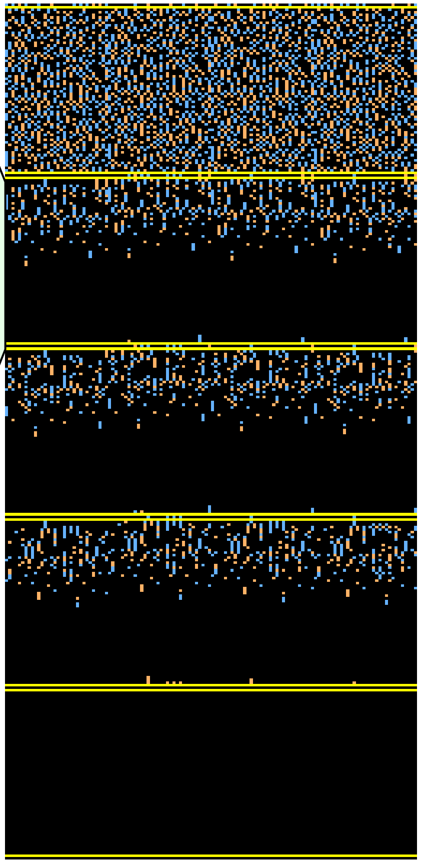
- Example:
  - Colliding certificates with different identities
- MD5 harmful for digital signatures

Chosen-Prefix Collisions

- MD5 Compr … vs IHV’, M’
- Analyze pro … rences
- Choose δM=
  - Which achi … elimination
- Construct se
  - Sufficient c
- Solve set of
  - Actual M, M
- Repeat unti



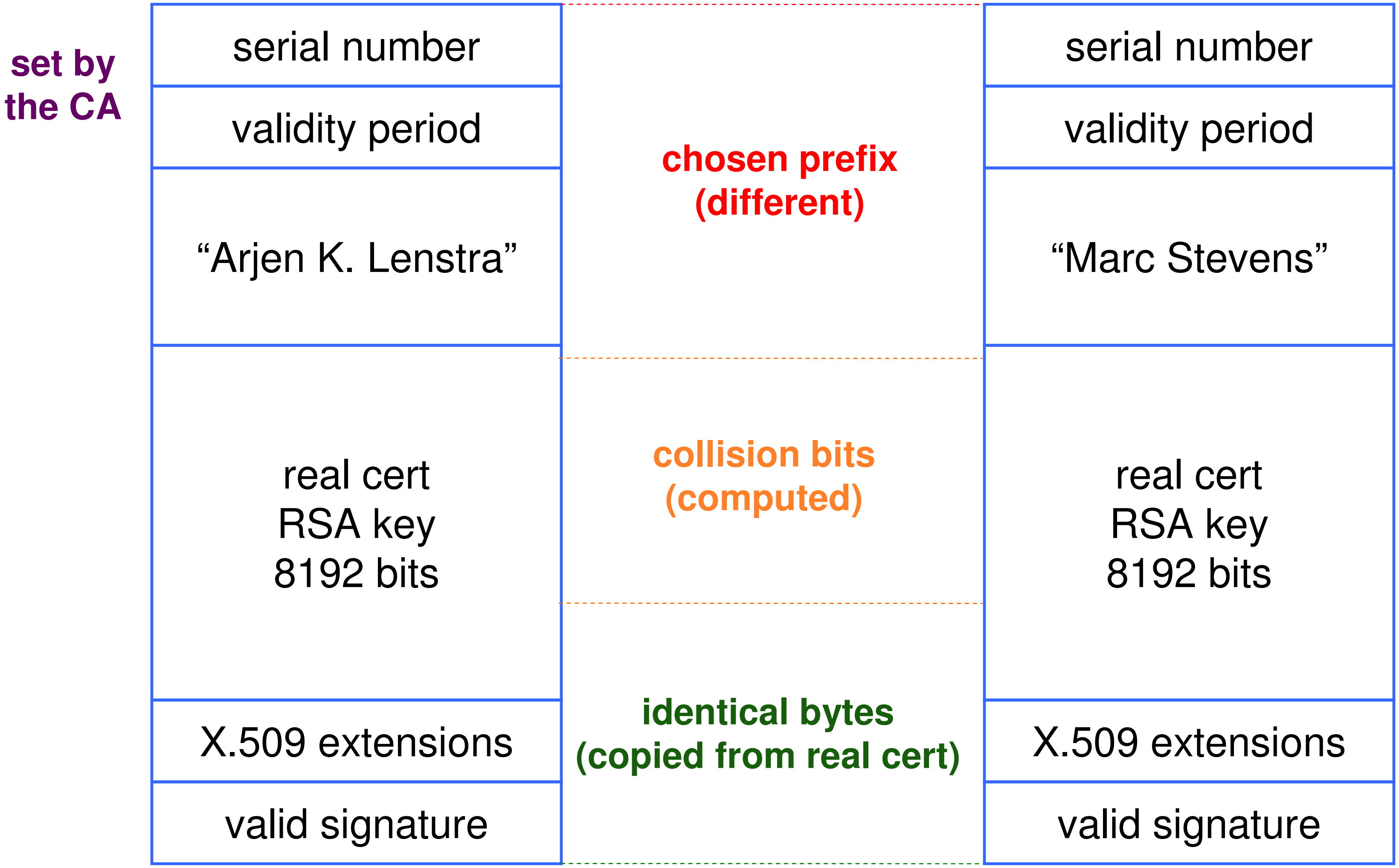
t	Bits Q_t: $b_{31} \dots b_0$	#
−3	10010110 00100101 10101000 11011100	32
−2	00100001 11010101 11100000 11-11110	32
−1	10001011 10000001 11011011 00-01110	32
0	01001111 00101101 01011011 -++00011	32
1	..1..... ......0. ..+-.... 0-+.....	7
2	..00.... ....0^-. ..1+.... 1-+.....	10
3	!.+0..1. ....0-+. ..1+.... .0-.....	11
4	.!0+..0. ....-1-. ...+.... .-+.....	10
5	..+-..-. ....+1-. .1...... .1-.1...	10
6	!.0-..0. ....1... .0.1.!.1 .10.00..	13
7	..1+..01 ....0.0. !+.....0 ..0.+1.1	13
8	!..0..-1 ....! .-.^...- ....-+.0	11
9	.!.0..0+ ......0. .+1-...+ 1-.-	12
10	......1+ .1....1. .100...+ 0.0.0-.-	13
11	.1...11+ .00101-1 !1+0.1.+ 0.101-10	24
12	00^0000- .-101111 .0-000.1 +^-10001	29
13	0+-00-+1 ^0--++-- ^-1+1-.- ++++----	31
14	+110+--- ---+0+-- -----100 .1110100	31
15	101-1-11 101010.0 1+1001.1 11110-0-	30
16	10010010 +00-.1^1 00101+.0 ....-	23
17	01.-.0.. ...0...+ .0..0..1 .....^.1	11
18	1+.-.... ^..+...+ ....0^.- ...0...0	11
19	+0.1.... ...+..01 ....-...	7
20	.-.0..0. ...0.01- ..0....^ ...+....	10
21	^-....1. ...1..-. ..1.^... ...+.0..	9
22	.0....+. .....-0^ ..+..... 0..-....	8
23	.1...... .0...-1. .... 1..-.+..	7
24	......^. .1...+.. ..^..... -..1.+..	7
25	....-...+.. .... 0..0.-..	5
26	....0... .....+.. .... 0....0..	4
27	...01... .^...1.. .... 1....1..	6
28	...1-... .....0.. .... +.......	4
29	...-1...	2
30	...1-... .... ^.......	3
31	...-+...	2
32	...0....	1
33	...!!...	2
34 – 60	....	0
61	....	
62	....+.....	
63	....+.....	
64	....+.....	



Chosen-Prefix Collisions

- Not all δIHVs can be eliminated
- First perform birthday search
  - Find δIHVs of specific form
    e.g. $\delta\text{IHV}=(0,x,x,y)$
  - Extend search to lower # near-collision blocks
- Appends 64 to 96 bits to prefixes
- Then iteratively eliminate differences in δIHV
- Till $\delta\text{IHV}=(0,0,0,0)$

2006 Example Colliding Certificates



set by the CA

serial number

validity period

“Arjen K. Lenstra”

real cert
RSA key
8192 bits

X.509 extensions

valid signature

chosen prefix
(different)

collision bits
(computed)

identical bytes
(copied from real cert)

serial number

validity period

“Marc Stevens”

real cert
RSA key
8192 bits

X.509 extensions

valid signature

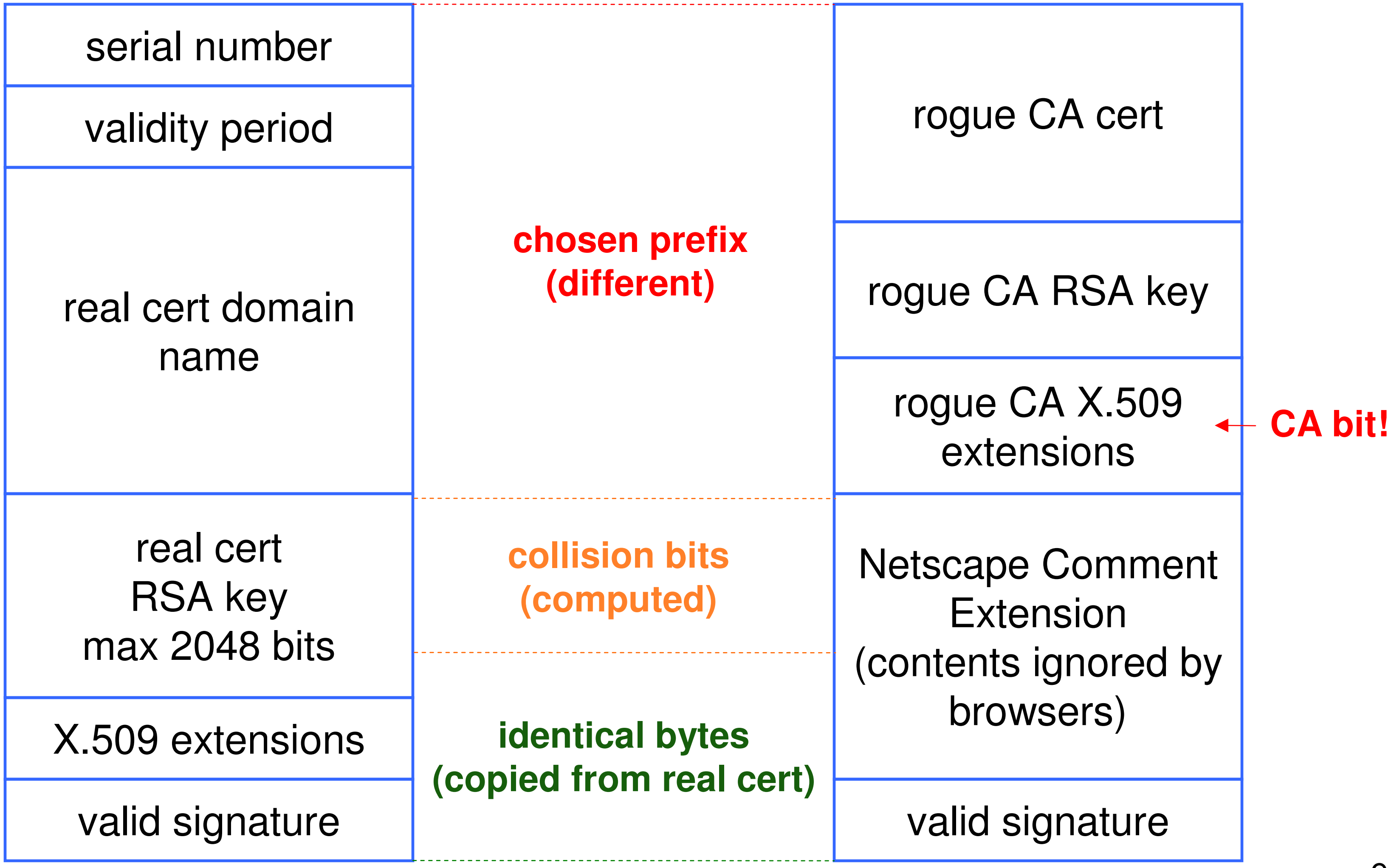
Certification Authorities

- Security and trust provided by CAs only as strong as the weakest CA
- Internet security may break down when even one CA is subverted
  - Man-in-the-Middle attacks
    - Impersonation of any secure website
    - Looks completely secure and as original website
    - Attacker has full control over all decrypted data
    - Phishing for private data
    - Or subtly alter data such as financial transactions
      - eBay, PayPal, online banking, etc.
  - Requires interception of connections
    - E.g. by subverting the insecure Domain Name System (DNS)
    - Local network access is already sufficient

Certification Authorities

- We were able to create a sub-CA signed by a known trusted CA (RapidSSL)
  - Not by default known by major web browsers
  - But is trusted as it is signed by a known CA
- Same effect as subverting a known trusted CA
- Possible because one particular commercial CA
  - used MD5 to create signatures
    - MD5 known to have significant weaknesses since 2004
  - had weaknesses in procedures

Creating a sub-CA



serial number

validity period

real cert domain name

real cert RSA key max 2048 bits

X.509 extensions

valid signature

chosen prefix (different)

collision bits (computed)

identical bytes (copied from real cert)

rogue CA cert

rogue CA RSA key

rogue CA X.509 extensions

CA bit!

Netscape Comment Extension (contents ignored by browsers)

valid signature

Obstacles

- Predicting serial number and validity period
- Total computation < a few days
- Max 204 collision bytes instead of 716
  - Limit by the CA RapidSSL
  - Greatly increases computational time
  - 17 months on 1000 pc cores

Predictions

- RapidSSL uses a fully automated system
- Certificate issued exactly 6 seconds after clicking

- RapidSSL uses sequential serial numbers:
  - Nov 3 07:44:08 2008 GMT 643006
  - Nov 3 07:45:02 2008 GMT 643007
  - Nov 3 07:46:02 2008 GMT 643008
  - Nov 3 07:47:03 2008 GMT 643009
  - Nov 3 07:48:02 2008 GMT 643010
  - Nov 3 07:49:02 2008 GMT 643011
  - Nov 3 07:50:02 2008 GMT 643012
  - Nov 3 07:51:12 2008 GMT 643013
  - Nov 3 07:51:29 2008 GMT 643014
  - Nov 3 07:52:02 2008 GMT ?

Predictions

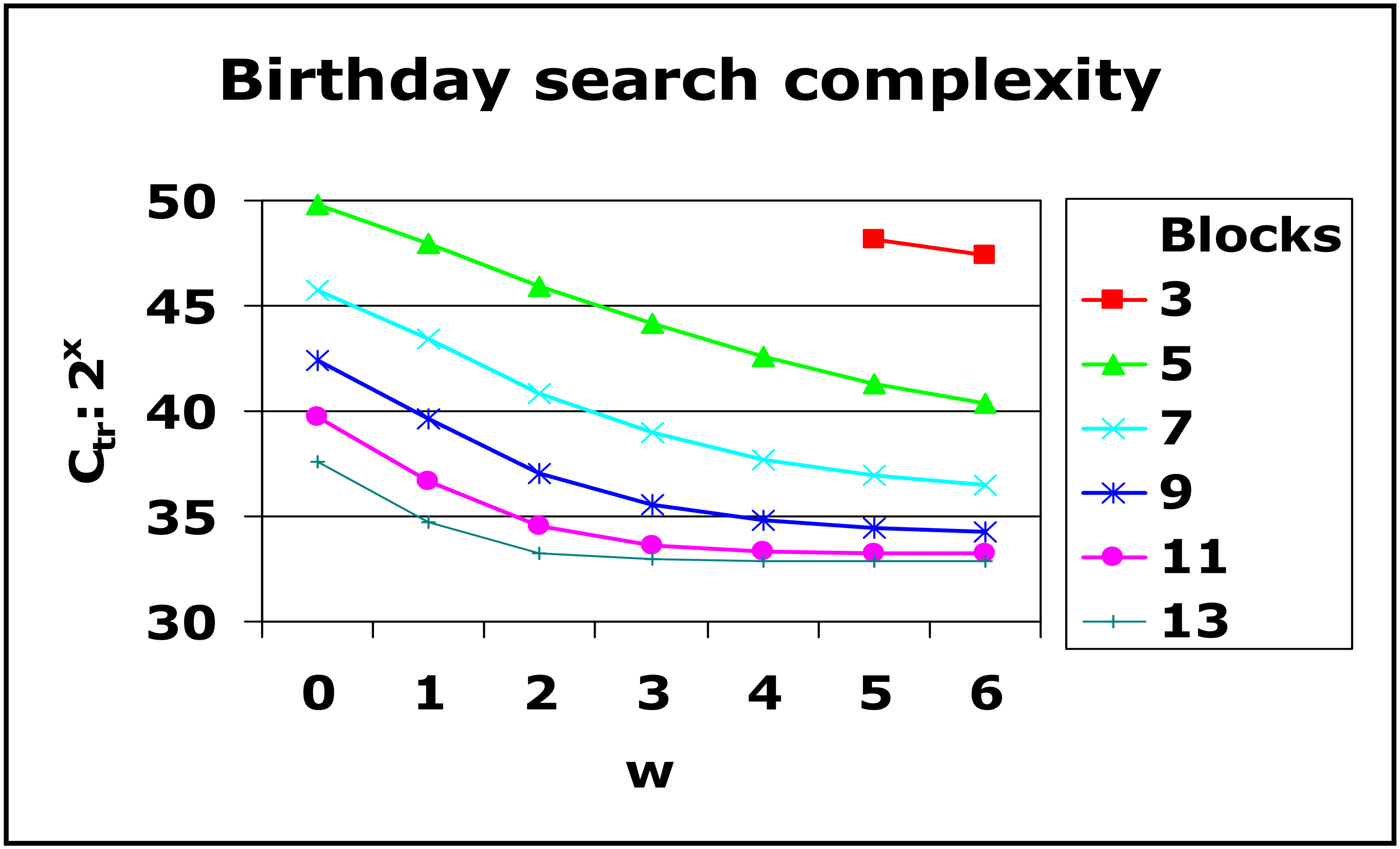
Estimate: 800-1000 certificates per weekend

Procedure:

1. Get the serial number **S** on Friday
2. Predict the value for time **T** on Sunday to be **S+1000**
3. Generate the collision bytes
4. Shortly before time **T** buy enough certs to increment the counter to **S+999**
5. Send colliding request at time **T** and get serial number **S+1000**

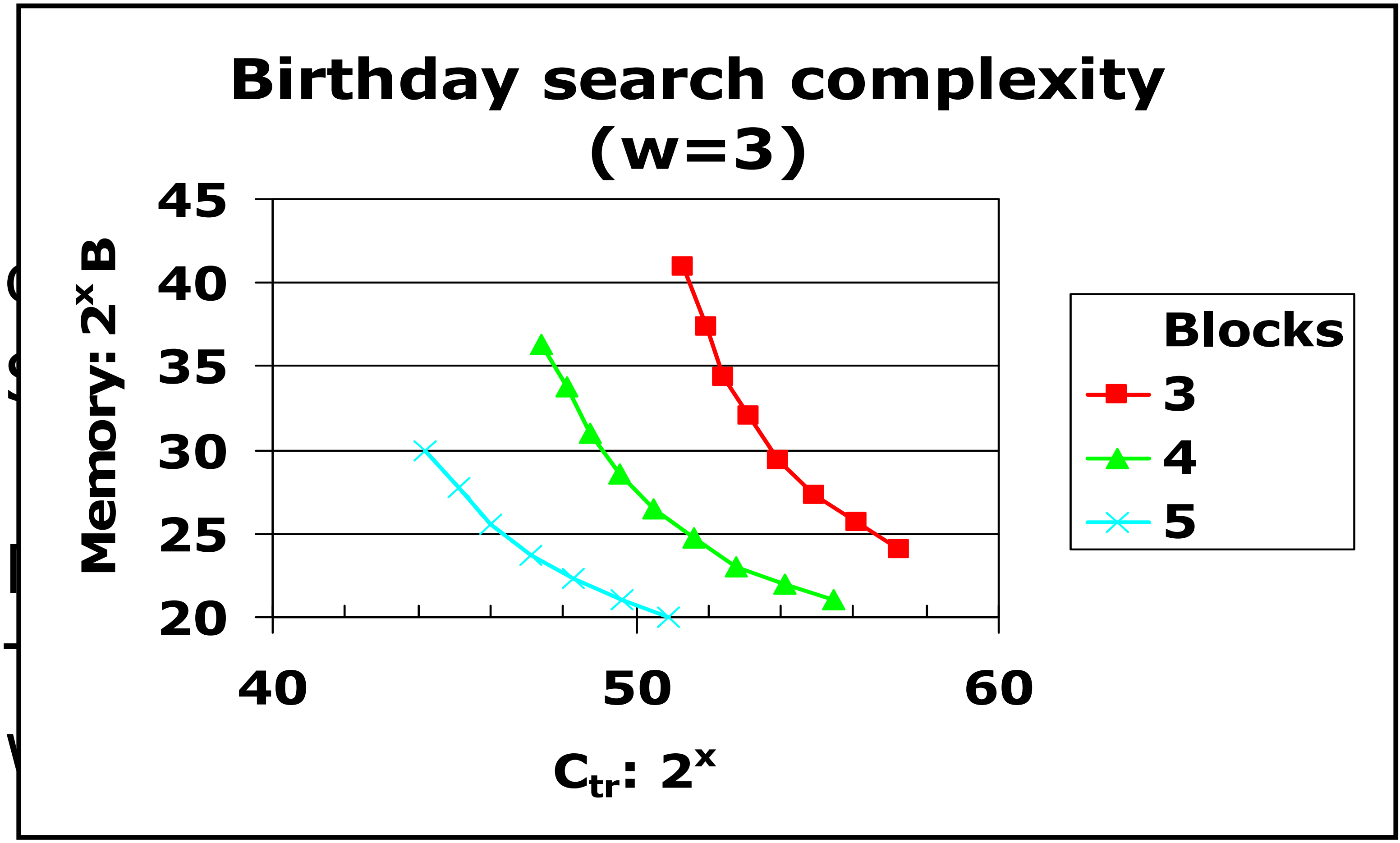
Collision Improvements

- Allow extra bit differences in last step



Collision Improvements

- Birthday search for δIHV=(δa,δb,δc,δd)

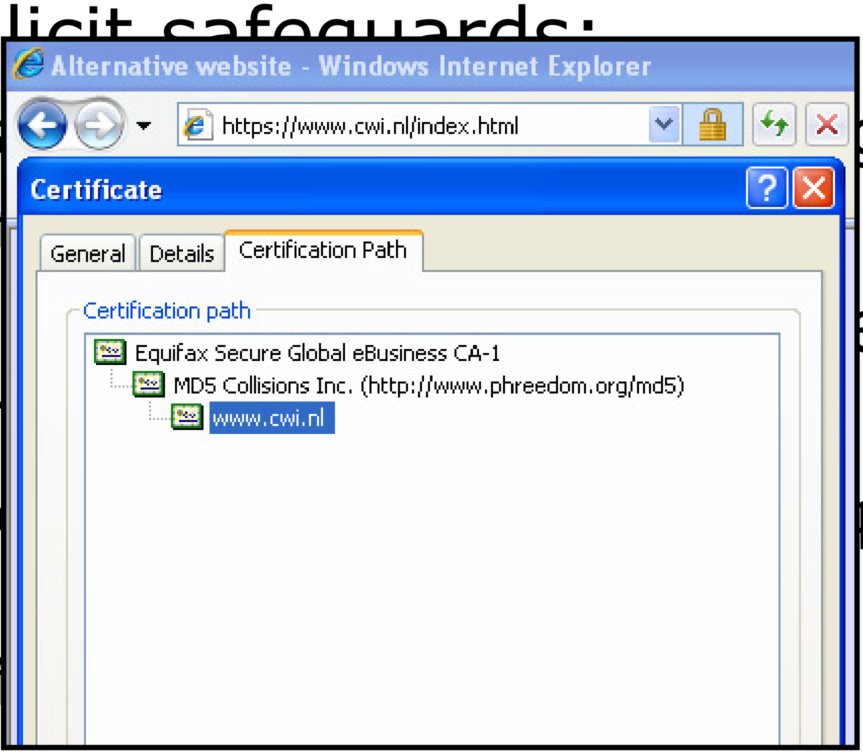


Collision Improvements

- Rogue CA construction (<2048 bits)
  - Cluster of 215 PlayStation3s
    - Performing like 8600 pc cores
  - Complexity 2^{50} using 30GB:
    - 1 day on cluster
  - Complexity $2^{48.2}$ using a few TBs:
    - 1 day on 20 PS3s and 1 pc
    - 1 day on 8 NVIDIA GeForce GTX280s
    - 1 day on Amazon EC2 at the cost of $2,000
- Normal CPC
  - Complexity approx. 2^{39} (<1 day on quadcore pc)

Result

- Success at 4th attempt
  - Generated CA signature for real cert also valid for rogue CA cert
- Explicit safeguards:
  - Va gust 2004
  - Pr
- Maj ed CAs wer
  - R quately
  - M rs af



Alternative website - Windows Internet Explorer

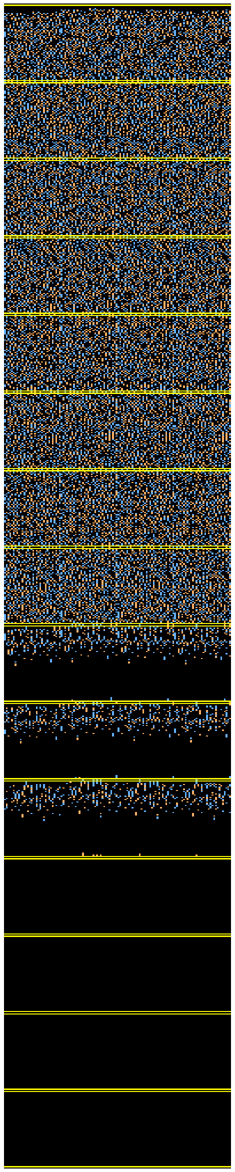
https://www.cwi.nl/index.html

Certificate

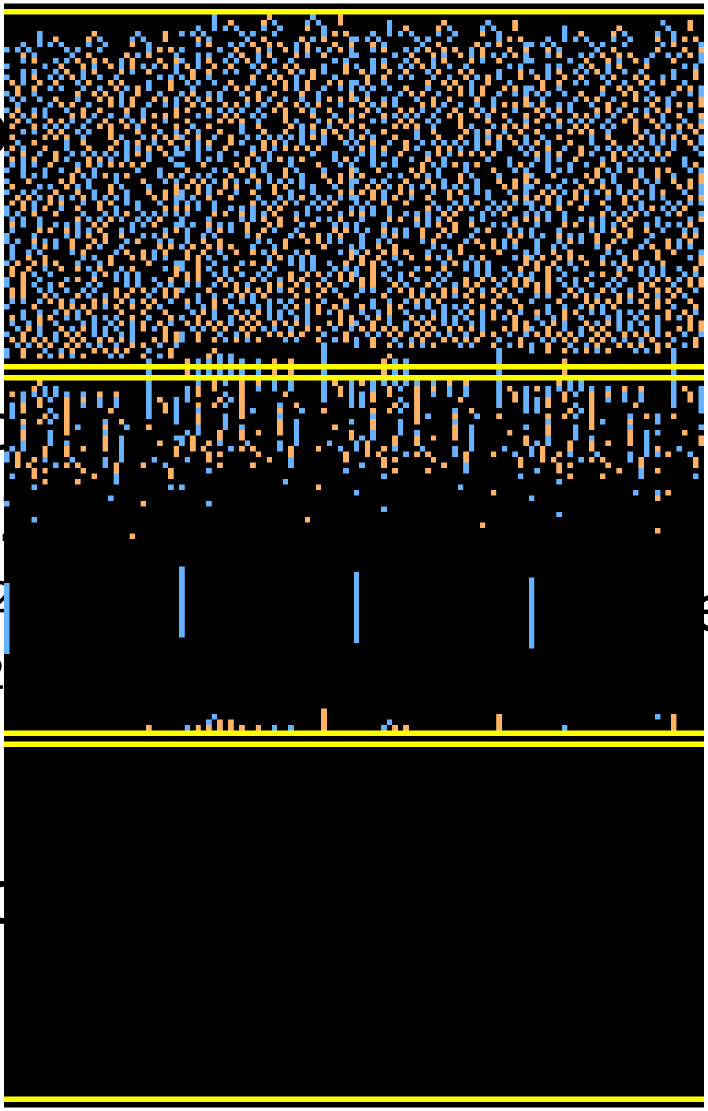
General | Details | Certification Path

Certification path

- Equifax Secure Global eBusiness CA-1
  - MD5 Collisions Inc. (http://www.phreedom.org/md5)
    - www.cwi.nl



Single block CPC



- Birthday search fo
  reduced to 0 with n block
- New approach:
  - New fastest near-c . 2^{15})
  - Allow extra factor compl.
  - Results in set of 2^{2} e form
    δa=-2^5, δd=-2^5+2^2
- Total complexity:
- Example single blo

Conclusion

- Collision attacks on MD5 form real threat
- Hard to replace broken crypto primitives
  - MD5 used by major CAs
    4 years after first collision attacks
  - Crypto primitives can be broken overnight
  - What to do when e.g. SHA-1 really falls, say yesterday?
  - How to make replacement of primitives easier?
- Source code implementation released:
  http://code.google.com/p/hashclash
  (Support for CELL/PS3 & CUDA)

Progress of Collision Attacks

Attack complexities for MD5, SHA-1 and SHA-2

	MD5		SHA-1		SHA-2(256)	
jaar	identical prefix	chosen prefix	identical prefix	chosen prefix	identical prefix	chosen prefix
– 2003	64	64	80	80	128	128
2004	40		69			
2005	37		63			
2006	32	49		80 - ε		
2007	25	42	61			
2008	21					
2009	16 New!	39	52 New?			

(logarithmic: 38 means $2^{38} \approx$ 1day on 1pc)