# Cryptanalysis of the ESSENCE Family of Hash Functions

Nicky Mouha[1,2], Gautham Sekar, Jean-Philippe Aumasson, Thomas Peyrin, Søren S. Thomsen, Meltem Sönmez Turan, Bart Preneel

[1]ESAT/SCD-COSIC
Katholieke Universiteit Leuven, Belgium
[2]IBBT, Belgium
nicky.mouha@esat.kuleuven.be

December 13, 2008

# Contents

# Outline

# Hash function basics

- Cryptographic hash function $h : \{0,1\}^* \rightarrow \{0,1\}^n$
  - Collision: $m, m'$ where $m \neq m'$: $H(m) = H(m')$, finding collision should be 'infeasible'
  - Also: finding (second) preimage 'infeasible'
- Birthday attack
  - Generic attack, collision after about $2^{n/2}$ $h$-evaluations
- Specialized attacks
  - Collision in less than about $2^{n/2}$ evaluations through weaknesses in $h$
  - X. Wang found attacks for MD5, SHA-1

# SHA-3 Competition

- SHA-1 broken
- SHA-2 unbroken, but similar design
- NIST announces SHA-3 competition

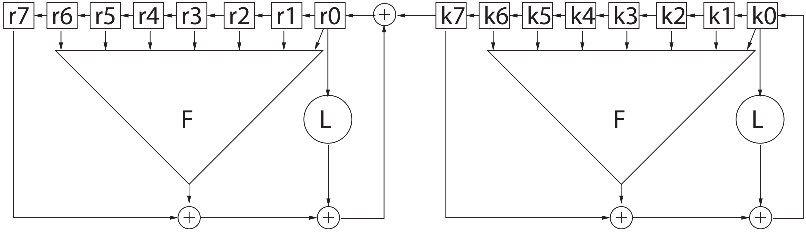
- ESSENCE = design by Jason Worth Martin
- Submitted to (ongoing) SHA-3 competition
- Advanced to first round

# Outline

# ESSENCE Hash Function

- Message is split into 256- or 512-bit blocks, depending on digest size
- Each message block is input to compression function
- Can use Merkle trees to increase parallelism (not used in SHA-3 submission)

# ESSENCE Compression Function




- 32- or 64-bit registers, for the 256- or 512-bit digest size respectively
- 8 $r_i$ registers loaded with the IV or chaining value
- 8 $k_i$ registers loaded with the 256- or 512-bit message block
- After 32 rounds + Davies-Meyer feedforward: $r_i$ contains new chaining value

# Description of $F$ and $L$

- The function $F$:
  - $F(a, b, c, d, e, f, g)$ is non-linear Boolean function from $GF(2^7)$ to $GF(2)$
  - Works in parallel ("bit-sliced") on all 32 or 64 bits of every register
- The function $L$:
  - $L$ is Linear Feedback Shift Register (LFSR)
  - Different $L$-function for 256- or 512-bit hash

# Outline

# Attack Description

- Semi-Free-Start collision
  - Same chaining value for $(m, m')$ in collision pair
  - Chaining value chosen by attacker
- ESSENCE design claim
  - Resistant to linear and differential cryptanalysis (24 rounds)
  - Analysis only for one-bit differences
- Our result: attack for 31 rounds using multiple-bit differences

# Attack Complexity

- Difference $A$: best possible difference for our characteristic (next slides)
- Characteristic to find collision in $2^{254.65}$ compression function calls
- Faster than generic attack ($2^{256}$)
- But: requires message pairs for first nine rounds with negligible complexity

# Differential characteristic (1/4)

| Round | Register $R$ | Register $K$ | Pr for $CV$ | Pr for $m$ |
|---|---|---|---|---|
| 0 | 0 0 0 0 0 0 0 0 | A 0 0 0 0 0 0 0 | 1 | 1 |
| 1 | 0 0 0 0 0 0 0 A | 0 0 0 0 0 0 0 A | 1 | 1 |
| 2 | 0 0 0 0 0 0 A 0 | 0 0 0 0 0 0 A 0 | $2^{-17}$ | $2^{-17}$ |
| 3 | 0 0 0 0 0 A 0 0 | 0 0 0 0 0 A 0 0 | $2^{-17}$ | $2^{-17}$ |
| 4 | 0 0 0 0 A 0 0 0 | 0 0 0 0 A 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 5 | 0 0 0 A 0 0 0 0 | 0 0 0 A 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 6 | 0 0 A 0 0 0 0 0 | 0 0 A 0 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 7 | 0 A 0 0 0 0 0 0 | 0 A 0 0 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| ... | ... | ... | ... | ... |

- $0 = $ 0000000000000000, $A = $ 0A00102190303 6C3

# Differential characteristic (2/4)

| Round | Register $R$ | Register $K$ | Pr for $CV$ | Pr for $m$ |
|---|---|---|---|---|
| ... | ... | ... | ... | ... |
| 8 | A 0 0 0 0 0 0 0 | A 0 0 0 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 9 | 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 0 A | 1 | 1 |
| 10 | 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 A 0 | 1 | $2^{-17}$ |
| 11 | 0 0 0 0 0 0 0 0 | 0 0 0 0 0 A 0 0 | 1 | $2^{-17}$ |
| 12 | 0 0 0 0 0 0 0 0 | 0 0 0 0 A 0 0 0 | 1 | $2^{-17}$ |
| 13 | 0 0 0 0 0 0 0 0 | 0 0 0 A 0 0 0 0 | 1 | $2^{-17}$ |
| 14 | 0 0 0 0 0 0 0 0 | 0 0 A 0 0 0 0 0 | 1 | $2^{-17}$ |
| 15 | 0 0 0 0 0 0 0 0 | 0 A 0 0 0 0 0 0 | 1 | $2^{-17}$ |
| ... | ... | ... | ... | ... |

# Differential characteristic (3/4)

| Round | Register $R$ | Register $K$ | Pr for $CV$ | Pr for $m$ |
|---|---|---|---|---|
| ... | ... | ... | ... | ... |
| 16 | 0 0 0 0 0 0 0 0 | A 0 0 0 0 0 0 0 | 1 | $2^{-17}$ |
| 17 | 0 0 0 0 0 0 0 A | 0 0 0 0 0 0 0 A | 1 | 1 |
| 18 | 0 0 0 0 0 0 A 0 | 0 0 0 0 0 0 A 0 | $2^{-17}$ | $2^{-17}$ |
| 19 | 0 0 0 0 0 A 0 0 | 0 0 0 0 0 A 0 0 | $2^{-17}$ | $2^{-17}$ |
| 20 | 0 0 0 0 A 0 0 0 | 0 0 0 0 A 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 21 | 0 0 0 A 0 0 0 0 | 0 0 0 A 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 22 | 0 0 A 0 0 0 0 0 | 0 0 A 0 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| 23 | 0 A 0 0 0 0 0 0 | 0 A 0 0 0 0 0 0 | $2^{-17}$ | $2^{-17}$ |
| ... | ... | ... | ... | ... |

# Differential characteristic (4/4)

| Round | Register $R$ | Register $K$ | Pr for $CV$ | Pr for $m$ |
|---|---|---|---|---|
| ... | ... | ... | ... | ... |
| 24 | A 0 0 0 0 0 0 0 | A 0 0 0 0 0 0 R | $2^{-17}$ | 1 |
| 25 | 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 R S | 1 | 1 |
| 26 | 0 0 0 0 0 0 0 0 | 0 0 0 0 0 R S T | 1 | 1 |
| 27 | 0 0 0 0 0 0 0 0 | 0 0 0 0 R S T U | 1 | 1 |
| 28 | 0 0 0 0 0 0 0 0 | 0 0 0 R S T U V | 1 | 1 |
| 29 | 0 0 0 0 0 0 0 0 | 0 0 R S T U V W | 1 | 1 |
| 30 | 0 0 0 0 0 0 0 0 | 0 R S T U V W X | 1 | 1 |
| 31 | 0 0 0 0 0 0 0 0 | R S T U V W X Y | 1 | 1 |

- $R$ to $Y$ are arbitrary differences

# Outline

# Finding Message Pairs

- Add linear equations to the inputs of $F$, such that outputs of $F$ are linear
- Then: finding message pairs = solving underdetermined system of linear equations
- One possible linearization: $2^{60}$ message pairs, very fast to enumerate
- Technique is similar to multi-message modification (MD5) or amplified boomerang attack (SHA-1), but obtained in fully automated way

# Conditions for $F$

| |
|---|
| $x_0 \oplus x_2 = 0$ |
| $x_1 = 0$ |
| $x_3 = 1$ |
| $x_4 = 1$ |
| $x_5 = 1$ |
| $x_7 = 0$ |
| $x_8 = 1$ |
| $x_9 = 0$ |
| $x_{10} = 0$ |
| $x_{12} = 1$ |

| |
|---|
| $F(x_0, \ldots, x_6) = 1$ |
| $F(x_1, \ldots, x_7) = x_2 \oplus 1$ |
| $F(x_2, \ldots, x_8) = 0$ |
| $F(x_3, \ldots, x_9) = 0$ |
| $F(x_4, \ldots, x_{10}) = 1$ |
| $F(x_5, \ldots, x_{11}) = 1$ |
| $F(x_6, \ldots, x_{12}) = 0$ |

# Conforming Message Pair

| $i$ | $m_i$ | $m'_i$ | $m_i \oplus m'_i$ |
|---|---|---|---|
| 0 | FFFFFFFFFFFFFFFF | FFFFFFFFFFFFFFFF | 0000000000000000 |
| 1 | 1A001021983836CB | 1A001021983836CB | 0000000000000000 |
| 2 | 5809832A1DEA2458 | 5809832A1DEA2458 | 0000000000000000 |
| 3 | 8AEF5FEBEB9FDAAB | 8AEF5FEBEB9FDAAB | 0000000000000000 |
| 4 | 32F9D8578015D297 | 32F9D8578015D297 | 0000000000000000 |
| 5 | 0D031372423B91AC | 0D031372423B91AC | 0000000000000000 |
| 6 | B804AC08CD97E348 | B804AC08CD97E348 | 0000000000000000 |
| 7 | E8BB8E649DC3B35F | E2BB9E450DF3859C | 0A0010219030 36C3 |

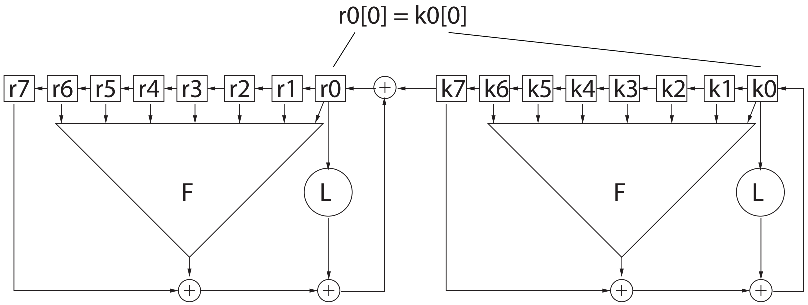
# Outline

# Weakness of $F$

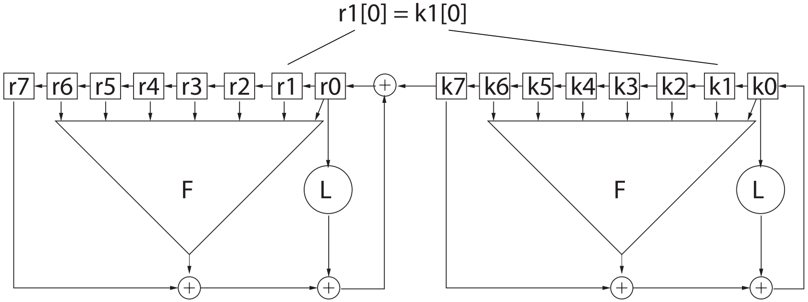
$$\begin{aligned} F(a,b,c,d,e,f,g) = {} & abcdefg + abcdef + abcefg + acdefg + abceg + \\ & abdef + abdeg + abefg + acdef + acdfg + \\ & acefg + adefg + bcdfg + bdefg + cdefg + abcf + \\ & abcg + abdg + acdf + adef + adeg + adfg + \\ & bcde + bceg + bdeg + cdef + abc + abe + \\ & abf + abg + acg + adf + adg + aef + aeg + bcf + \\ & bcg + bde + bdf + beg + bfg + cde + cdf + def + \\ & deg + dfg + ad + ae + bc + bd + cd + ce + df + \\ & dg + ef + fg + a + b + c + f + 1 \end{aligned}$$

- ANF of $F$ contains highest degree monomial $\Rightarrow F$ is unbalanced
- If $a, \ldots, g$ are uniformly distributed, then $\Pr[F(a,b,c,d,e,f,g)[j] = 0] = 0.5 + 2^{-7}$

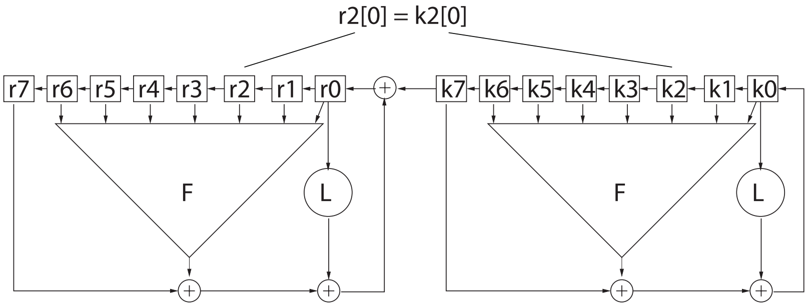
# Distinguisher: After 0 Rounds



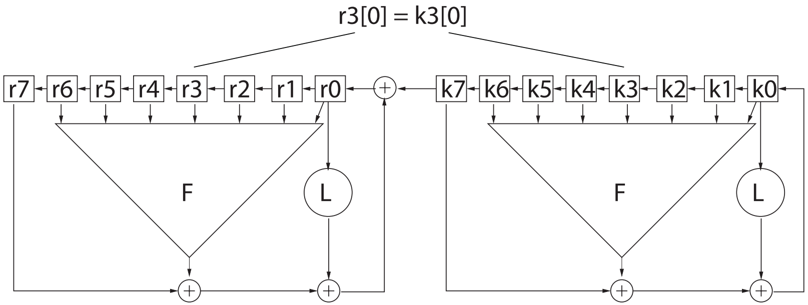
# Distinguisher: After 1 Round



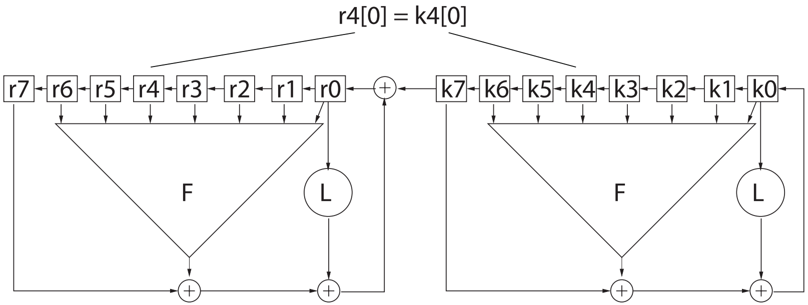
# Distinguisher: After 2 Rounds



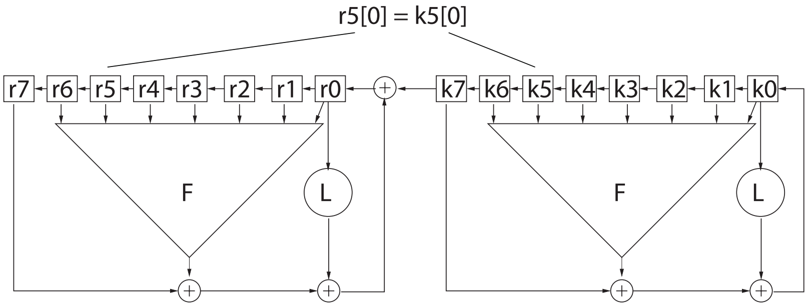
# Distinguisher: After 3 Rounds



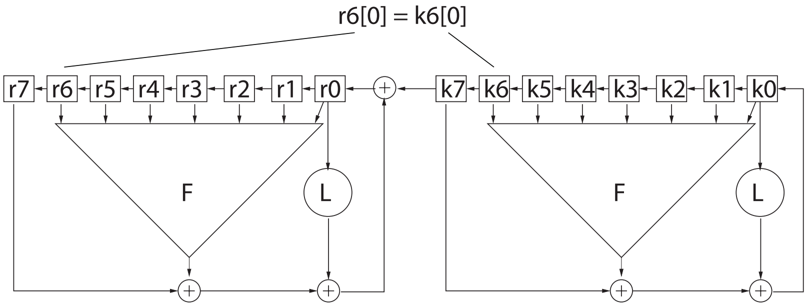
# Distinguisher: After 4 Rounds



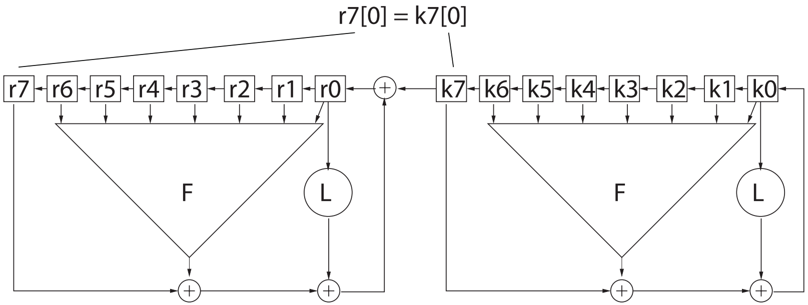
# Distinguisher: After 5 Rounds



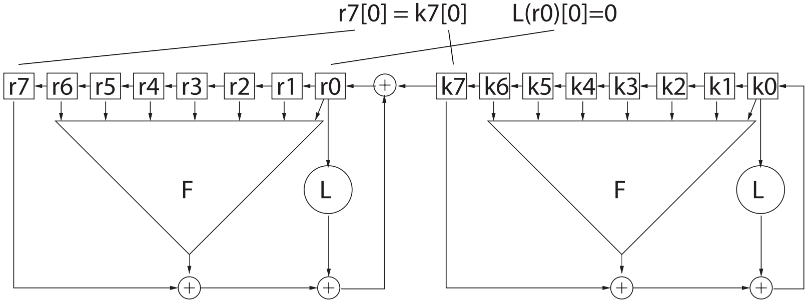
# Distinguisher: After 6 Rounds



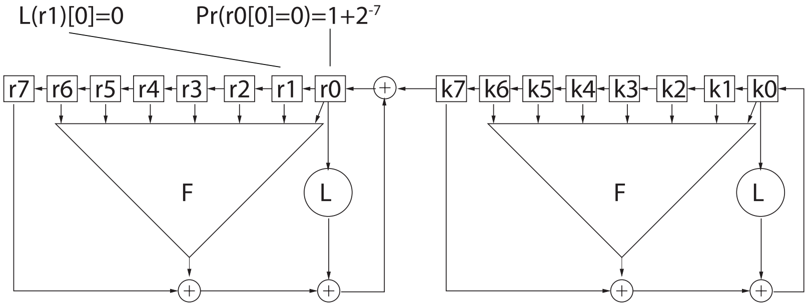
# Distinguisher: After 7 Rounds



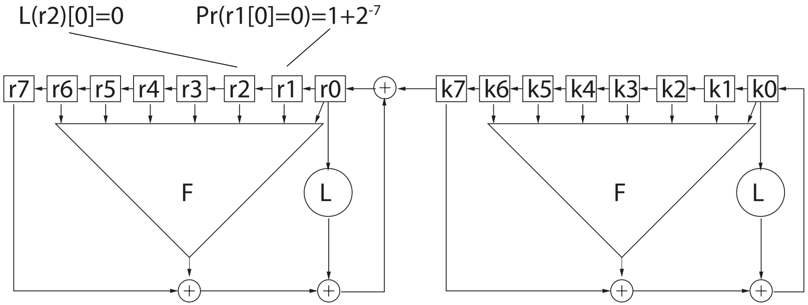
# Distinguisher: After 7 Rounds



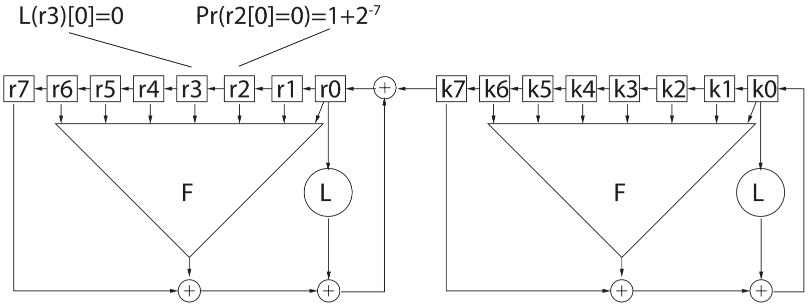
# Distinguisher: After 8 Rounds



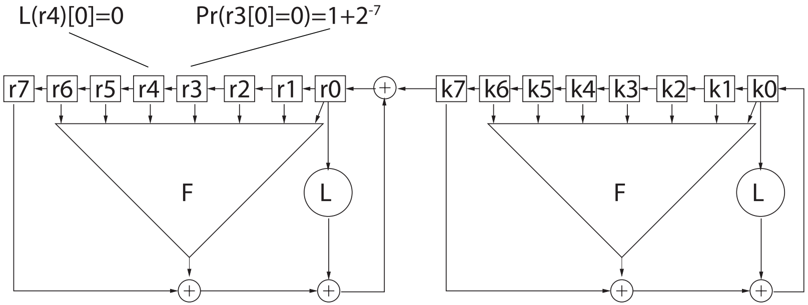
# Distinguisher: After 9 Rounds



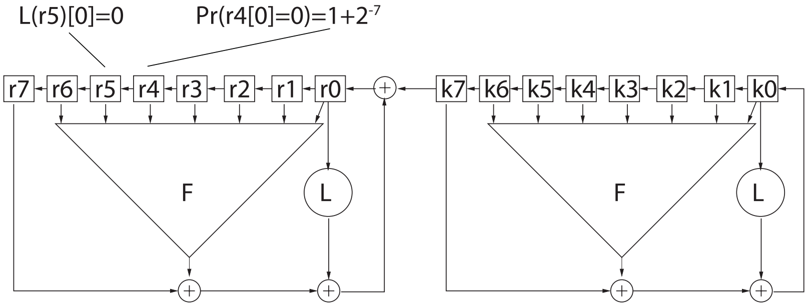
# Distinguisher: After 10 Rounds



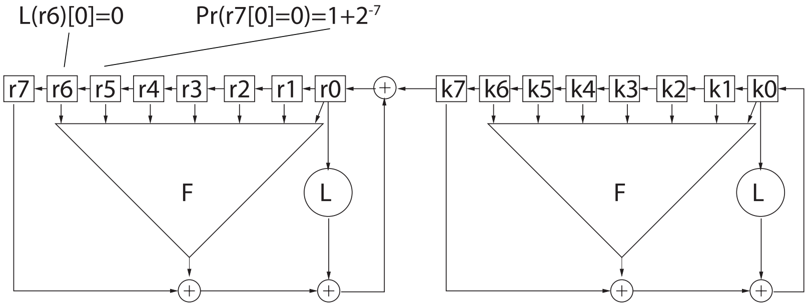
# Distinguisher: After 11 Rounds



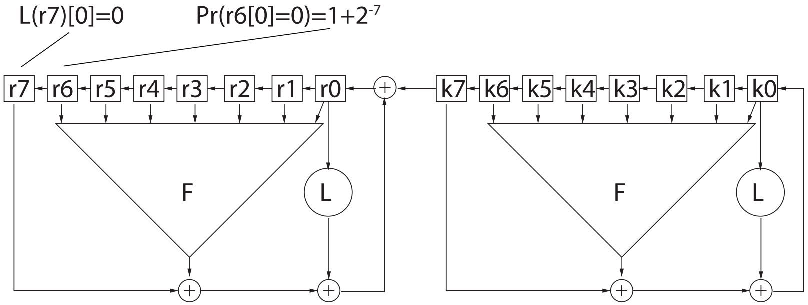
# Distinguisher: After 12 Rounds



# Distinguisher: After 13 Rounds



# Distinguisher: After 14 Rounds



# Distinguisher Results

- Complexity for the distinguisher: $2^{17}$ plaintexts for success probability .9772
- Distinguisher can be turned into key-recovery attack
  - Complexity: testing $2^{225.1}$ and $2^{450.1}$ keys
  - Exhaustive search: $2^{256}$ and $2^{512}$ keys
- By undoing Davies-Meyer feedforward:
  - Block cipher distinguisher extends to compression function

# Outline

# Slide Attacks

- Attack on ESSENCE compression function
- Works for any number of rounds

| | |
|---|---|
| $c, c'$ | 243F6A88 243F6A88 243F6A88 243F6A88 243F6A88 243F6A88 243F6A88 243F6A88 |
| $m$ | 00000000 00000000 00000000 00000000 00000000 00000000 00000000 F6B1EB63 |
| $m'$ | 094E149C 00000000 00000000 00000000 00000000 00000000 00000000 00000000 |
| $R$ | BE31AA01 EB6E9F07 EAD99889 6FE79B44 391CCD35 67FDB8B6 FC3AA0F6 6E80148E |
| $R'$ | F86D77C6 BE31AA01 EB6E9F07 EAD99889 6FE79B44 391CCD35 67FDB8B6 FC3AA0F6 |

# Fixed Points

- Fixed point: same values for internal registers after round function update

| | ESSENCE-256 | ESSENCE-512 |
|---|---|---|
| $c_0$ | 993AE9B9 | D5B330380561ECF7 |
| $m_0$ | 307A380C | 10AD290AFFB19779 |

- Conclusion slide attacks + fixed points: don't use ESSENCE in block cipher mode

# Outline

# Conclusion

- Several types of attacks
  - 31-Round Semi-Free-Start Collision
  - 14-Round Distinguisher and Key Recovery
  - Slide Attacks
  - Fixed Points
- ESSENCE not in second round of SHA-3 competition
- But:
  - ESSENCE is a simple design, easy to analyze and hardware friendly
  - We give countermeasures against our attacks
- Questions?