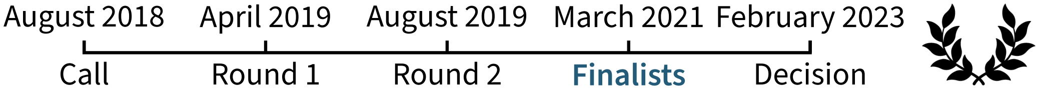
Finding Collisions for Round-Reduced Romulus-H

Marcel Nageler, Felix Pallua, Maria Eichlseder

FSE 2023 – Kobe 🇯🇵

The Romulus Family

- Authenticated Encryption + Hash function by Guo et al. [GIK+18]
- Hash function Romulus-H designed for NIST LWC
  - 10 Finalists including Romulus



August 2018 — Call

April 2019 — Round 1

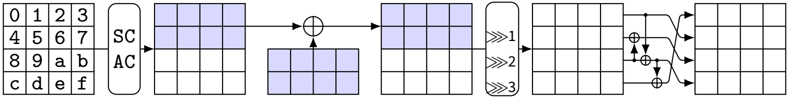
August 2019 — Round 2

March 2021 — **Finalists**

February 2023 — Decision

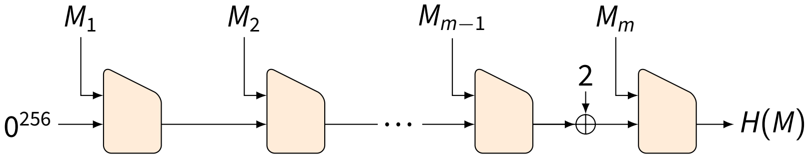
Skinny Specification [GIK+18]

- Romulus uses Skinny-128-384 with 40 rounds (instead of 56)
  - 128-bit blocks
  - 384-bit tweakey



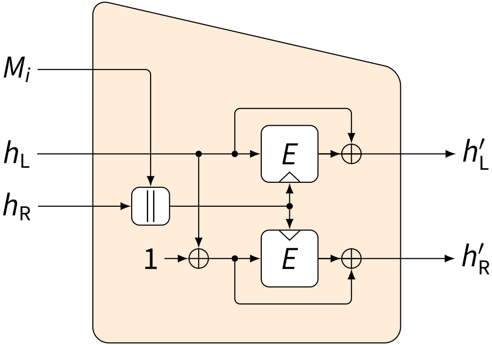
Romulus-H Mode [GIK+18]

- Merkle-Damgård with Permutation [HPY07]



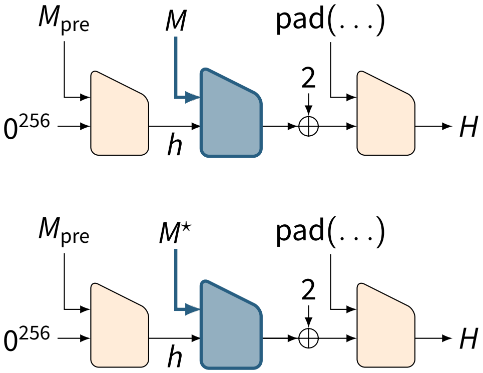
Romulus-H Compression Function [GIK+18]

- Hirose Double-Block-Length Construction [Hir06]
  - Two nearly equal block cipher calls
- Free-start collisions for 23 rounds by Dong et al. [DHS+21]
  - 2^{124} time, 2^{124} memory



Attack Goals

- Find good differential characteristics
- Find semi-free-start collisions
  - Collision on compression function with constant h
- Find hash collisions
  - Connect semi-free-start collision with prefix M_{pre}

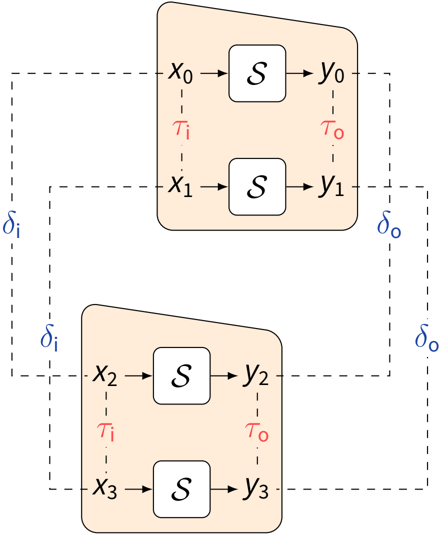


Our Results

Bounds on the number of active S-boxes based on different models.

Rounds	6	7	8	9	10	11	12	13	14	15	16
Semi-coll.	✓	✓	✓	✓	✓	✓	✓	—	✓	—	—
Collision	✓	✓	✓	✓	✓	—	—	—	—	—	—
#S-boxes (plain **2**×)	16	22	34	44	54	60	66	78	86	86	106
#S-boxes (equal 〓)	11	16	25	33	42	50	59	67	76	77	96
#S-boxes (joint 🔗)	11	16	25	33	41	46	54	59	69	73	74

Joint Differential Characteristics: Different Settings



2× plain: 2 SKINNY calls considered independent

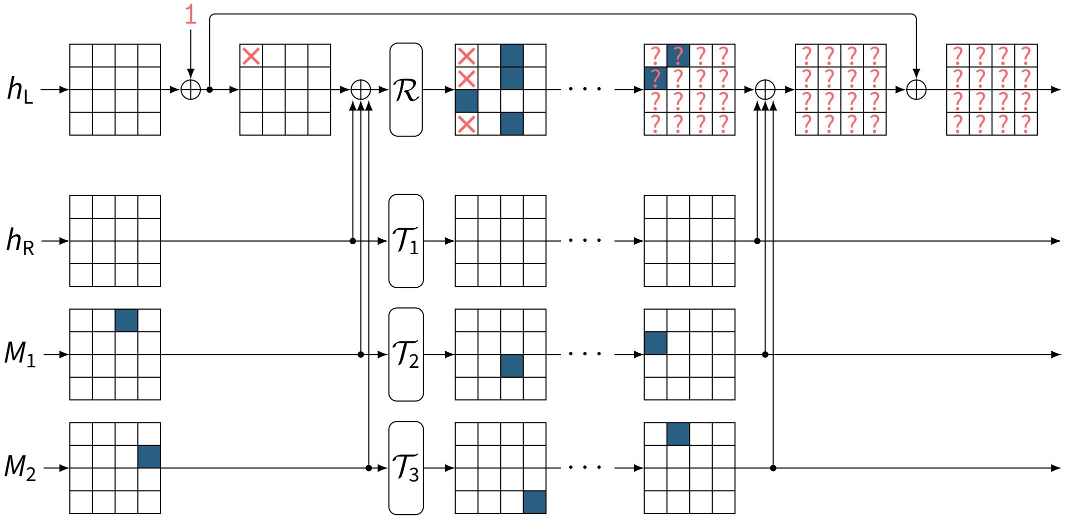
🔗 joint: add connecting difference τ

- $\tau \in \{0, \times, \textbf{?}\}$

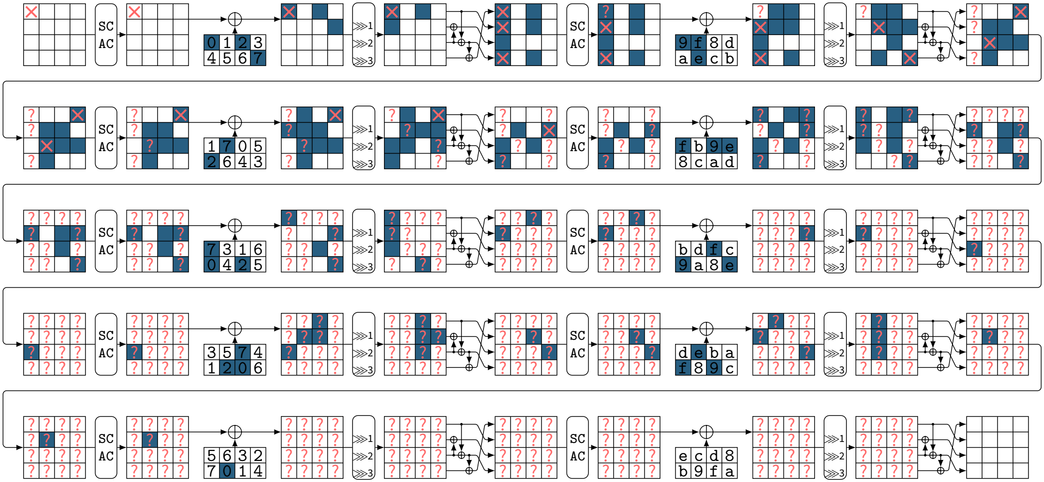
= equal: keep track of where 2 SKINNY calls are equal

- $\tau \in \{0, \textbf{?}\}$

Attack Setup



Cellwise Characteristic for 10 Rounds (Joint Setting 🔗, 41 active S-Boxes)



■ difference $\bar{\delta}$ ☒ connecting difference $\bar{\tau} = \mathrm{x}$? $\bar{\tau} = ?$ 0 TK lane index

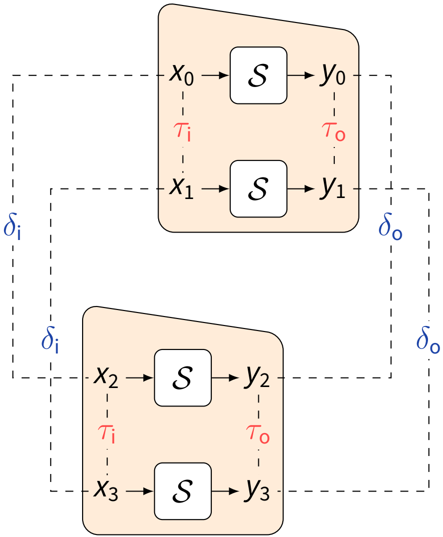
Finding Bitwise Characteristics

- Model CNF of DDT $\geq w$
- What to do when $\delta = \tau = \times$
  a) define DDT4($\delta_i, \delta_o, \tau_i, \tau_o$)
    - # of solutions to simultaneous transition $\delta_i \to \delta_o, \tau_i \to \tau_o$
    - model CNF of DDT4 $\geq w$

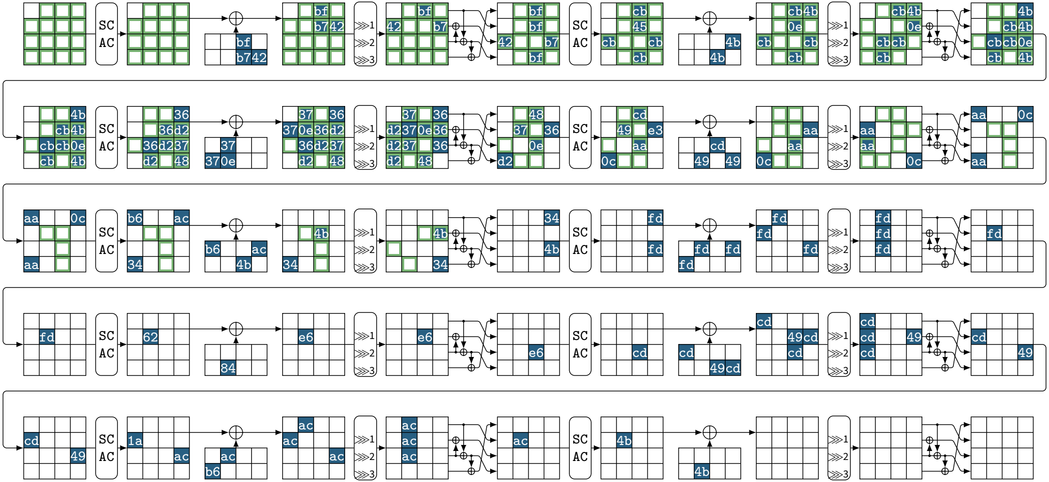
    $\to$ very expensive

  b) Switch to equality setting (**=**)

    $\to$ cheaper model as $\tau \neq \times$.



Bitwise Characteristic for 10 Rounds (Equality Setting ▬, $p = 2^{-234}$)

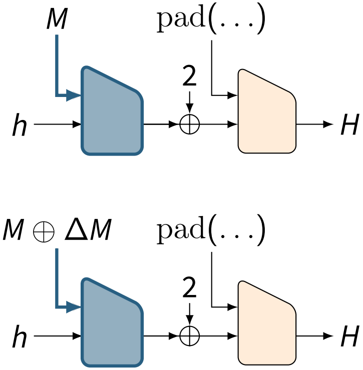


■ difference δ 🟩 connecting difference $\tau = 0$ □ connecting difference $\bar{\tau} =$ **?**

Finding Assignment for Characteristic

- Encode linear layer using Xor constraints of Z3 SMT solver
- Encode S-box as minified CNF of solution set
- Solve for M and h

  $\rightarrow$ Get semi-free-start collision
- Optimized model to reduce number of variables



Semi-Free-Start Collision Results

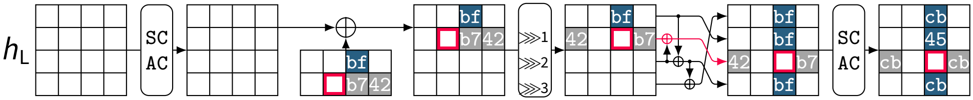
- Most characteristics are actually impossible

  → Generate many and verify

- For **14 rounds**
  - Generating characteristic takes ∼1 second
  - Verifying characteristic takes ∼30 seconds
  - After 32 hours on 1 CPU, we find a valid characteristic with $p = 2^{-420}$
    - Satisfiable using 512 degrees of freedom (256-bit message, 256-bit chaining value)

Finding Hash Collision

- Randomly choose an initial block
- Verify the characteristic in the first 2 rounds is satisfiable (in C++)
  - only then run SMT solver
  - $p \approx 2^{-11}$ that a given h_L is compatible



🟦 active, dep. on tk_2 □ inactive, dep. on tk_5 ▨ active, dep. on other tk_i

10-Round Collision

- Collision can be found in about 1 hour on 88 cores.
  - Based on characteristic with $p = 2^{-234}$ (256 degrees of freedom)

$$\begin{aligned}
M_{\text{pre}} &= \texttt{55554654434b5555 59495a41504a4c41 4c41545247414452 4a4447515247594c}, \\
M_1 &= \texttt{b63a14a596b5216e 97e6d7cc7b0b014d 1d533b4f882a2075 04dd06463e1f98ed}, \\
M_2 &= \texttt{b63aa4a596b52116 97e620cc50202a4d 1d534a4f882a20fc 04dd2d46dffe79ed}, \\
M_1 \oplus M_2 &= \texttt{0000b00000000078 0000f7002b2b2b00 0007100000000089 00002b00e1e1e100}.
\end{aligned}$$

$$H_{10}(M_{\text{pre}} \| M_1) = H_{10}(M_{\text{pre}} \| M_2)$$

Conclusion

- Differential model for `Romulus-H`
- Joint differential characteristics (δ, τ)

- ✓ Collisions for 10 rounds of `Romulus-H`
- ✓ Semi-free-start-collisions for 14 rounds

`github.com/IAIK/romulush_collisions`

Bibliography I

[DHS+21] Xiaoyang Dong, Jialiang Hua, Siwei Sun, Zheng Li, Xiaoyun Wang, and Lei Hu. **Meet-in-the-Middle Attacks Revisited: Key-Recovery, Collision, and Preimage Attacks.** CRYPTO 2021. Vol. 12827. LNCS. Springer, 2021, pp. 278–308. DOI: 10.1007/978-3-030-84252-9_10.

[GIK+18] Chun Guo, Tetsu Iwata, Mustafa Khairallah, Kazuhiko Minematsu, and Thmoas Peyrin. **Romulus.** Submission to NIST Lightweight Cryptography. https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/finalist-round/updated-spec-doc/romulus-spec-final.pdf. Aug. 2018.

[Hir06] Shoichi Hirose. **Some Plausible Constructions of Double-Block-Length Hash Functions.** FSE 2006. Vol. 4047. LNCS. Springer, 2006, pp. 210–225. DOI: 10.1007/11799313_14.

Bibliography II

[HPY07] Shoichi Hirose, Je Hong Park, and Aaram Yun. **A Simple Variant of the Merkle-Damgård Scheme with a Permutation.** ASIACRYPT 2007. Vol. 4833. Lecture Notes in Computer Science. Springer, 2007, pp. 113–129. DOI: 10.1007/978-3-540-76900-2_7.