

# On the Security of COMET Authenticated Encryption Scheme

NIST Lightweight Workshop '19

---

Shay Gueron<sup>1,2</sup>, Ashwin Jha<sup>3</sup>, Mridul Nandi<sup>3</sup>

<sup>1</sup> University of Haifa, Israel

<sup>2</sup> Amazon Web Services, USA

<sup>3</sup> Indian Statistical Institute Kolkata, India

## Lightweight Authenticated Encryption Design

- Block cipher based.
- Rate-1.
- Small state size (close to  $(n + \kappa)$ -bit).
- Simple design (simple operations like XOR, shifts and rotations).

## Design Summary

- Rate-1 and Feedback-based authenticated encryption mode.
- Combined feedback function:

input is a function of current output and next plaintext block.

- Nonce and block counter-based rekeying.
- Parametrized by the block size,  $n \in \{64, 128\}$ . Tag size  $t = n$ .
- Two variants:
  - COMET-128: Here  $n = 128$ , key size  $\kappa = 128$ , nonce size  $r = 128$ .
  - COMET-64: Here  $n = 64$ , key size  $\kappa = 128$ , nonce size  $r = 120$ .

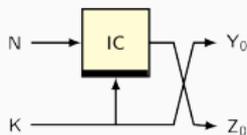
## Nonce-based Initial State Derivation

- For COMET-128:

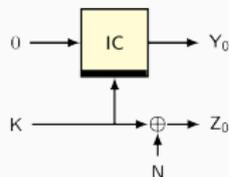
$$(Y_0, Z_0) := (K, IC_K(N))$$

- For COMET-64:

$$(Y_0, Z_0) := (IC_K(0), K \oplus N \parallel 0^{32})$$

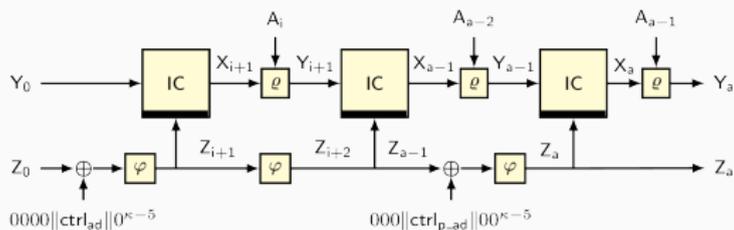


COMET-128



COMET-64

## Associated Data Processing

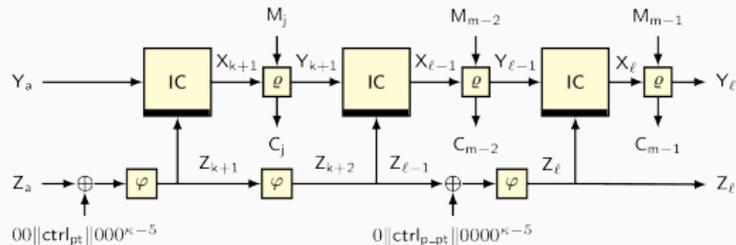


Here,  
 $0 \leq i \leq a-3$

$$ctrl_{ad} = \begin{cases} 1 & \text{if } |A| > 0, \\ 0 & \text{o.w.} \end{cases}$$

$$ctrl_{p\_ad} = \begin{cases} 1 & \text{if } |A_{a-1}| < n, \\ 0 & \text{o.w.} \end{cases}$$

## Plaintext Processing



Here,

$$0 \leq j \leq m-3,$$

$$k = a + j, \text{ and}$$

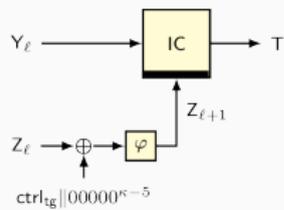
$$\ell = a + m$$

$$ctrl_{pt} = \begin{cases} 1 & \text{if } |M| > 0, \\ 0 & \text{o.w.} \end{cases}$$

$$ctrl_{p-pt} = \begin{cases} 1 & \text{if } |M_{m-1}| < n, \\ 0 & \text{o.w.} \end{cases}$$

Ciphertext processing is symmetrically defined.

## Tag Generation



Here,

$$\ell = a + m \quad \text{ctrl}_{\text{tg}} = \begin{cases} 1 & \text{for tag generation,} \\ 0 & \text{o.w.} \end{cases}$$

## Design Features

- **Design simplicity:** Only requires shift and XOR operations apart from block cipher calls.
- **Small state size:** Possibility of close to  $(n + \kappa)$ -bit state size in area optimized implementation.
- **Inverse free:** No need for block cipher decryption.
- **Dynamic key updation:** No two blocks share the same key non-trivially.
- **Efficiency:** Single-pass scheme.

## Submissions to NIST LwC Standardization Project

- COMET-128\_AES-128/128 instantiated with AES-128/128. [Primary]
- COMET-128\_CHAM-128/128 instantiated with CHAM-128/128.
- COMET-64\_Speck-64/128 instantiated with Speck-64/128.
- COMET-64\_CHAM-64/128 instantiated with CHAM-64/128.

# COMET : Security Claims

Submissions	Confidentiality		Integrity	
	Time	Data (in bytes)	Time	Data (in bytes)
COMET-128_AES-128/128	$2^{119}$	$2^{64}$	$2^{119}$	$2^{64}$
COMET-128_CHAM-128/128	$2^{119}$	$2^{64}$	$2^{119}$	$2^{64}$
COMET-64_Speck-64/128	$2^{119}$	$2^{64}$	$2^{112}$	$2^{45}$
COMET-64_CHAM-64/128	$2^{119}$	$2^{64}$	$2^{112}$	$2^{45}$

We focus on the security of COMET-128 .

## AEAD Security Game

- Indistinguishability game between the ideal system  $\mathcal{O}_0$  and real system  $\mathcal{O}_1$ , where

$$\mathcal{O}_0 := (\$, \perp, IC^\pm) \quad \mathcal{O}_1 := (\text{COMET-128.E}_K, \text{COMET-128.E}_K, IC^\pm).$$

- Advantage of any adversary  $\mathcal{A}$  against COMET-128 is defined as:

$$\text{Adv}_{\text{COMET-128}}^{\text{aead}}(\mathcal{A}) := \left| \Pr \left[ \mathcal{A}^{\mathcal{O}_1} = 1 \right] - \Pr \left[ \mathcal{A}^{\mathcal{O}_0} = 1 \right] \right|.$$

- $\mathcal{A}$  is **computationally unbounded**, but **bounded in number of queries** to its oracle.
- $\mathcal{A}$  operates under two restrictions:
  - **Nonce-respecting**: No two encryption query share the same nonce.
  - **Non-trivial forger**: An encryption query  $(N, A, M)$  yields  $(C, T)$ , a decryption query  $(N, A, C, T)$  is not allowed.

## Theorem

For  $\sigma_e, \sigma_d < 2^{127}$ ,  $q_p < 2^{127}$ , and  $(q_e, q_d, \sigma_e, \sigma_d, q_p)$ -adversary  $\mathcal{A}$  we have

$$\mathbf{Adv}_{\text{COMET-128}}^{\text{aead}}(\mathcal{A}) \leq \frac{4\sigma_c^2}{2^{256}} + \frac{14\sigma_c q_p}{2^{249}} + \frac{3\sigma_c^2}{2^{128}} + \frac{3.01q_p}{2^{121}} + \frac{4\sigma_c}{2^{128}} + \frac{q_c}{2^{64}} + \frac{6q_p\sigma_d}{2^{188.5}}$$

- $q_e$  and  $q_d$  denote the number of queries to COMET-128.E<sub>K</sub> and COMET-128.D<sub>K</sub>, respectively.
- $\sigma_e$  and  $\sigma_d$  denote the sum of input (associated data and message) lengths across all encryption and decryption queries, respectively;  
 $q_c = q_e + q_d$  and  $\sigma_c = \sigma_e + \sigma_d$ .
- $q_p$  denotes the number of direct queries to the block cipher.

## Proof tool: Coefficient-H Technique

- Concentrates on the query-response tuple, called the **transcript**, generated by  $\mathcal{A}$ 's interaction with the oracle at hand.
- Let  $\Theta_1$ : transcript random variable corresponding to  $\mathcal{O}_1$ .
- Let  $\Theta_0$ : transcript random variable corresponding to  $\mathcal{O}_0$ .
- Identify a set of **bad** transcripts,  $\Omega_{\text{bad}}$ .
- Compute  $\Pr [\Theta_0 \in \Omega_{\text{bad}}] \leq \epsilon_{\text{bad}}$ .
- Show that  $\frac{\Pr [\Theta_1 = \omega]}{\Pr [\Theta_0 = \omega]} \geq (1 - \epsilon_{\text{ratio}})$  for all  $\omega \notin \Omega_{\text{bad}}$ .
- Then,  $\text{Adv}_{\text{COMET-128}}^{\text{aead}}(\mathcal{A}) \leq \epsilon_{\text{bad}} + \epsilon_{\text{ratio}}$ .

## Notational Conventions

- Variables in encryption queries are defined as per the figures.
- Variables in decryption queries are defined analogously, topped with a bar.
- Variables in primitive queries are defined analogously, topped with a hat.

## Oracle description

- Real oracle: Faithfully responds to encryption, decryption and primitive queries.
- Ideal oracle:

For the encryption query: samples  $X_1, \dots, X_\ell, T \leftarrow_s \{0, 1\}^n$ , and sets  $(Y_j, C_j) = \varrho(X_{a+j+1}, M_j)$  for all  $0 \leq j \leq m$ . Sets  $Y_j = X_j \oplus A_j$  for  $1 \leq j \leq a$ . Returns  $(C, T)$ .

For decryption query: Returns  $\perp$  symbol.

For primitive query: Responds faithfully using  $IC^\pm$ .

- After the query phase, both the oracles **release all encryption query internal variables and the secret key**.

## Identifying bad events

- Kcoll (key guessing/recovery):

B1:  $\exists i \in [q_e], j \in [m^i]$ , such that  $Z_j^i = K$ .

B2:  $\exists i \in [q_d], j \in [\bar{m}^i]$ , such that  $\bar{Z}_j^i = K$ .

B3:  $\exists i \in [q_p]$ , such that  $\widehat{Z}^i = K$ .

B4:  $\exists i \in [q_e]$ , such that  $Z_0^i = *||0^{n/2}$ .

B5:  $\exists i \in [q_d]$ , such that  $\bar{Z}_0^i = *||0^{n/2}$ .

B6:  $\exists (i, j) \in [q_e] \times [m^i], (i', j') \in [q_d] \times [\bar{m}^{i'}]$ , such that  $N^i \neq \bar{N}^{i'}$  and  $Z_j^i = \bar{Z}_{j'}^{i'}$ .

- EEmatch (encryption-encryption state matching):

B7:  $\exists (i, j) \in [q_e] \times [m^i], (i', j') \in [q_e] \times [m^{i'}]$ , such that  $(Z_j^i, Y_j^i) = (Z_{j'}^{i'}, Y_{j'}^{i'})$ .

B7:  $\exists (i, j) \in [q_e] \times [m^i], (i', j') \in [q_e] \times [m^{i'}]$ , such that  $(Z_j^i, X_j^i) = (Z_{j'}^{i'}, X_{j'}^{i'})$ .

## Identifying bad events

- EPmatch (encryption-primitive state matching):

B9:  $\exists(i, j) \in [q_e] \times [m^i]$  and  $i' \in [q_p]$ , such that  $(Z_j^i, Y_j^i) = (\widehat{Z}^{i'}, \widehat{Y}^{i'})$ .

B10:  $\exists(i, j) \in [q_e] \times [m^i]$  and  $i' \in [q_p]$ , such that  $(Z_j^i, X_j^i) = (\widehat{Z}^{i'}, \widehat{X}^{i'})$ .

- EPKcoll (technical requirement: key exhaustion via primitive query):

B11:  $\exists(i, j) \in [q_e] \times [m^i]$  such that  $|\{j \in [q_p] : \widehat{Z}^j = Z^i\}| \geq 2^{n-1}$ .

## Identifying bad events

- Chain (valid forgery via primitive (and encryption) queries):

Let  $\text{domain}(\omega_p) := \{(\hat{Z}_i, \hat{Y}_i)\}_{i \in [q_p]}$  and  $\text{range}(\omega_p) := \{(\hat{Z}_i, \hat{X}_i)\}_{i \in [q_p]}$ .

$$\delta_i := \begin{cases} \max_{\bar{C}_{0 \dots k-1}^i = C_{0 \dots k-1}^j} (\bar{a}^i + k) & \text{if } \bar{A}^i = A^j \wedge (\bar{A}^i, \bar{C}^i) \neq (A^j, C^j) \\ \max_{\bar{A}_{0 \dots k-1}^i = A_{0 \dots k-1}^j} (k) & \text{otherwise.} \end{cases}$$

$$\delta'_i := \begin{cases} \max_{\bar{X}_{\delta_{i+1}}^i, \dots, \bar{X}_j^i \in \text{range}(\omega_p)} (j) & \text{if } \bar{X}_{\delta_{i+1}}^i \in \text{range}(\omega_p) \\ \delta_i & \text{otherwise.} \end{cases}$$

B12: chain using primitive queries

$\exists i \in [q_d]$  such that  $\delta_i \geq 0$ ,  $\delta'_i = \bar{\ell}^i$  and  $\bar{X}_{\bar{\ell}^i+1}^i = \bar{T}^i$ .

B13: partial chain using primitive queries followed by encryption query

$\exists i \in [q_d], (i', j') \in [q_e] \times [m^{i'}]$  such that  $0 \leq \delta_i < \delta'_i < \bar{\ell}^i$  and  $(\bar{Z}_{\delta'_i}^i, \bar{Y}_{\delta'_i}^i) = (Z_{j'}^{i'}, Y_{j'}^{i'})$ .

## Bounding $\Pr [\Theta_0 \in \Omega_{\text{bad}}]$

- $\Pr [\text{Kcoll}]$ : using the fact that  $K \leftarrow_{\$} \{0, 1\}^{\kappa}$

$$\Pr [\text{B1}] \leq \frac{\sigma_e}{2^{\kappa}}; \quad \Pr [\text{B2}] \leq \frac{\sigma_d}{2^{\kappa}}; \quad \Pr [\text{B3}] \leq \frac{q_p}{2^{\kappa}}.$$

$$\Pr [\text{B4} | \neg \text{B3}] \leq \frac{q_e}{2^{n/2}}; \quad \Pr [\text{B5} | \neg \text{B3}] \leq \frac{q_d}{2^{n/2}}; \quad \Pr [\text{B6}] \leq \frac{\sigma_e \sigma_d}{2^{\kappa}}.$$

- $\Pr [\text{EEmatch} | \neg \text{Kcoll}]$ : using the fact that  $K \leftarrow_{\$} \{0, 1\}^{\kappa}$  and  $X_j^i, X_{j'}^{i'} \leftarrow_{\$} \{0, 1\}^n$ .

$$\Pr [\text{B7}] \leq \frac{\sigma_e^2}{2^{n+\kappa}}; \quad \Pr [\text{B8}] \leq \frac{\sigma_e^2}{2^{n+\kappa}}.$$

**Bounding**  $\Pr [\Theta_0 \in \Omega_{\text{bad}}]$

- $\Pr [\text{EPmatch} | \neg \text{Kcoll}]$ :

- Primitive query occurs before encryption query:

$$\Pr [\text{EPmatch} | \neg \text{Kcoll}] \leq 2q_p \sigma_e / 2^{n+\kappa}.$$

- Primitive query after encryption query:

Let,  $\text{mcoll}(x) := |\{X_j^i = x : (i, j) \in [q_e] \times [m^i]\}|$  and  $\text{Mcoll}$  denote the event  $\max_x \text{mcoll}(x) \geq n$ . Then,

$$\begin{aligned} \Pr [\text{EPmatch} | \neg \text{Kcoll}] &\leq \Pr [\text{Mcoll}] + \Pr [\text{EPmatch} | \neg (\text{Kcoll} \vee \text{Mcoll})] \\ &\leq \frac{\sigma_e}{2^{n-1}} + \frac{2nq_p}{2^\kappa}. \end{aligned}$$

- $\Pr [\text{EPKcoll}]$ : using the fact that the number of keys which are repeated in primitive queries at least  $2^{n-1}$  times is at most  $q_p / 2^{n-1}$ .

$$\Pr [\text{EPKcoll}] \leq \frac{2\sigma_e q_p}{2^{n+\kappa}}.$$

**Bounding**  $\Pr [\Theta_0 \in \Omega_{\text{bad}}]$

- $\Pr [\text{Chain} | \neg(\text{Kcoll} \vee \text{EEmatch} \vee \text{EPmatch})]$ :

Using graph-based analysis (similar to Beetle).

Let  $\mathcal{G}_{\omega_p} = (\mathcal{V}, \mathcal{E})$  be an edge-labeled graph where  $\mathcal{V} = \text{domain}(\omega_p)$  and  $((\hat{Z}_j, \hat{Y}_i), (\hat{Z}_j, \hat{Y}_j), C^*) \in \mathcal{E}$  if and only if

$$(\hat{Z}_j, \hat{Y}_j) = (\text{IC}_{\hat{Z}_i}(\hat{Y}_i), \text{IC}_{\hat{Z}_i}(\hat{Y}_i) \oplus C^*)$$

A walk  $\mathcal{W}$  from vertex  $W_0$  to  $W_k$  with label  $C = (C_1, \dots, C_k)$ , denoted  $W_0 \xrightarrow{C} W_k$ , is

$$W_0 \xrightarrow{C_1} W_1 \cdots W_{k-1} \xrightarrow{C_k} W_k.$$

# COMET-128 : Security Proof Sketch

**Bounding**  $\Pr [\Theta_0 \in \Omega_{\text{bad}}]$

- $\Pr [\text{Chain} | \neg(\text{Kcoll} \vee \text{EEmatch} \vee \text{EPmatch})]$ :

A multi-chain with label  $C = (C_1, \dots, C_k)$ , denoted  $\mathcal{C}_C$ , is a set of labeled walks  $\{\mathcal{W}_1, \dots, \mathcal{W}_s\}$  such that for all  $1 \leq i \leq s$ ,

$$\mathcal{W}_i : (\hat{Z}_0^i, \hat{Y}_0^i) \xrightarrow{C} (\hat{Z}_k^i, \hat{Y}_k^i) \wedge \hat{Y}_0^1 = \dots = \hat{Y}_0^s \wedge \hat{X}_{k+1}^1 = \dots = \hat{X}_{k+1}^s.$$

$$\mathcal{W}_1 : (\hat{Z}_0^1, \hat{Y}_0^1) \xrightarrow{C_1} (\hat{Z}_1^1, \hat{Y}_1^1) \xrightarrow{C_2} (\hat{Z}_2^1, \hat{Y}_2^1) \xrightarrow{C_3} (\hat{Z}_3^1, \hat{Y}_3^1) \xrightarrow{C_4} (\hat{Z}_3^1, \hat{Y}_4^1) \overset{\text{IC}}{\dashrightarrow} \hat{X}_5^1$$

$$\mathcal{W}_2 : (\hat{Z}_0^2, \hat{Y}_0^2) \xrightarrow{C_1} (\hat{Z}_1^2, \hat{Y}_1^2) \xrightarrow{C_2} (\hat{Z}_2^2, \hat{Y}_2^2) \xrightarrow{C_3} (\hat{Z}_3^2, \hat{Y}_3^2) \xrightarrow{C_4} (\hat{Z}_3^2, \hat{Y}_4^2) \overset{\text{IC}}{\dashrightarrow} \hat{X}_5^2$$

⋮

$$\mathcal{W}_s : (\hat{Z}_0^s, \hat{Y}_0^s) \xrightarrow{C_1} (\hat{Z}_1^s, \hat{Y}_1^s) \xrightarrow{C_2} (\hat{Z}_2^s, \hat{Y}_2^s) \xrightarrow{C_3} (\hat{Z}_3^s, \hat{Y}_3^s) \xrightarrow{C_4} (\hat{Z}_3^s, \hat{Y}_4^s) \overset{\text{IC}}{\dashrightarrow} \hat{X}_5^s$$

$$\Pr [\text{B11} | \neg(\text{Kcoll} \vee \text{EEmatch} \vee \text{EPmatch})] \leq \sum_{i \in [q]_d} \Pr \left[ |\mathcal{C}_{\bar{c}_{\delta_i, \dots, \bar{m}_i}}| \geq \lambda_i \right] + \frac{\lambda_i}{2^\kappa}.$$

**Bound on  $\Pr \left[ |\mathcal{C}_{\bar{c}_{\delta_i \dots \bar{m}^i}}| \geq \lambda_i \right]$  and  $\lambda_i$**

- Three ways to construct a multi-chain structure:

- Forward-only: all queries of the form  $(\hat{Z}_i, \hat{Y}_i)$ .

$$\Pr \left[ \mathcal{C}_{\text{fwd}} \geq n \left\lceil \frac{q_p}{2^n} \right\rceil \right] \leq \frac{1}{2^n},$$

(by bounding the multicollisions on  $\hat{X}_j$ )

- Backward-only: all queries of the form  $(\hat{Z}_i, \hat{X}_i)$ .

$$\Pr \left[ \mathcal{C}_{\text{bck}} \geq n \left\lceil \frac{q_p}{2^n} \right\rceil \right] \leq \frac{1}{2^n}.$$

(by bounding the multicollisions on  $\hat{Y}_j$ )

- Both forward and backward type queries:

reduced to multicollision event at some index  $1 \leq i \leq \bar{\ell}^i$  (using Pigeonhole-principle).

$$\Pr \left[ \mathcal{C}_{\text{fwd-bck}} \geq \bar{\ell}^i \frac{2\sqrt{n}q_p}{2^{n/2}} + \frac{2q_p}{2^n} \right] \leq \frac{1}{2^n}.$$

- $\Pr \left[ |\mathcal{C}_{\bar{c}_{\delta_i \dots \bar{m}^i}}| \geq \bar{\ell}^i \frac{2\sqrt{n}q_p}{2^{n/2}} + 2n \left\lceil \frac{q_p}{2^n} \right\rceil + \frac{2q_p}{2^n} \right] \leq \frac{3}{2^n}.$

## COMET-128 : Security Proof Sketch

- $\Pr [B11 | \neg(Kcoll \vee EEmatch \vee EPmatch)] \leq \frac{2\sqrt{n}\sigma_d q_p}{2^{\kappa+n/2}} + \frac{2nq_d}{2^\kappa} \left[ \frac{q_p}{2^n} \right] + \frac{2q_d q_p}{2^{n+\kappa}} + \frac{3q_d}{2^n}$ .
- $\Pr [B12 | \neg(Kcoll \vee EEmatch \vee EPmatch)]$  can be bounded in a similar fashion.

$$\Pr [B12 | \neg(Kcoll \vee EEmatch \vee EPmatch \vee B11)] \leq \frac{2\sqrt{n}\sigma_d q_p}{2^{\kappa+n/2}} + \frac{2nq_d}{2^\kappa} \left[ \frac{q_p}{2^n} \right] + \frac{2q_d q_p}{2^{n+\kappa}} + \frac{3q_d}{2^n}.$$

$$\text{Finally, } \Pr [Chain | \neg(Kcoll \vee EEmatch \vee EPmatch)] \leq \frac{6\sqrt{n}\sigma_d q_p}{2^{\kappa+n/2}} + \frac{6nq_d}{2^\kappa} \left[ \frac{q_p}{2^n} \right] + \frac{4q_d q_p}{2^{n+\kappa}} + \frac{6q_d}{2^n}.$$

## Good transcript analysis

Given any good transcript  $\omega$ :

$$\frac{\Pr[\Theta_1 = \omega]}{\Pr[\Theta_0 = \omega]} \geq \left( 1 - \frac{2\sigma_d(\sigma_e + q_p)}{2^{\kappa+n}} - \frac{2q_d}{2^n} \right).$$

- First term bounds the probability that for some decryption query  $i$  an intermediate input  $(\hat{Z}_j^i, \hat{Y}_j^i)$  collides with some encryption/primitive input, for  $j > \delta_i$ .
- The second term bounds the probability that som decryption forgery succeeds given that all intermediate inputs are fresh.

This completes the proof.

**Thank you. Questions...**

# Acknowledgments

This work is supported in part by The Ministry of Science and Technology, Israel, and the Department of Science and Technology, Government of India, DST/INT/ISR/P-20/2017.

Shay Gueron is supported by

The Israel Science Foundation (grant No. 1018/16);

NSF-BSF Grant 2018640;

The BIU Center for Research in Applied Cryptography and Cyber Security, in conjunction with the Israel National Cyber Bureau in the Prime Minister's Office;

The Center for Cyber Law & Policy at the University of Haifa in conjunction with the Israel National Cyber Directorate in the Prime Minister's Office.

The authors would like to thank Mustafa Khairallah for sharing his observations on bad conditions B4 and B5.