Transient-Steady Effect Attack on Block Ciphers

Yanting Ren, An Wang and Liji Wu

Institute of Microelectronics, Tsinghua University, China

CHES 2015

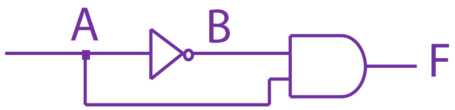
September 14th, 2015

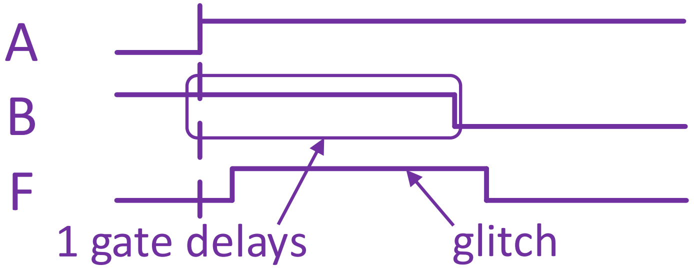
Outline

- Preliminaries
  - Glitches in combinational circuits
  - Clock-glitch-based fault attack
- Transient-steady effect attack
  - Basic idea
  - Attack on masked an unmasked S-Boxes
  - Experiments
- Further discussion
  - Attack scenario of parallel AES implementation
  - Attack scenario of WDDL-AES
  - Glitch injection
- Conclusion

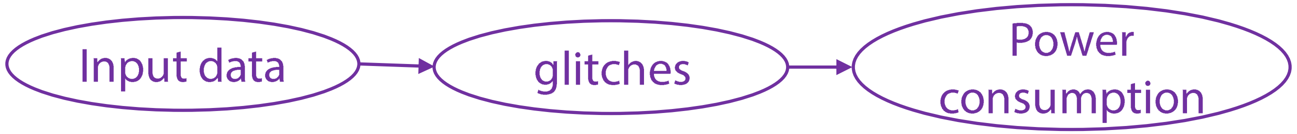
Glitches in combinational circuits

- Gates have inherent delays
- Glitches are unintended pulses at the output of a combinational circuit





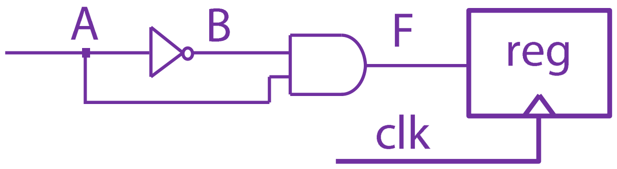
- Glitches can leak side-channel information
  - Glitches depend on the input patterns
  - The number of glitches affects the power consumption of the circuit

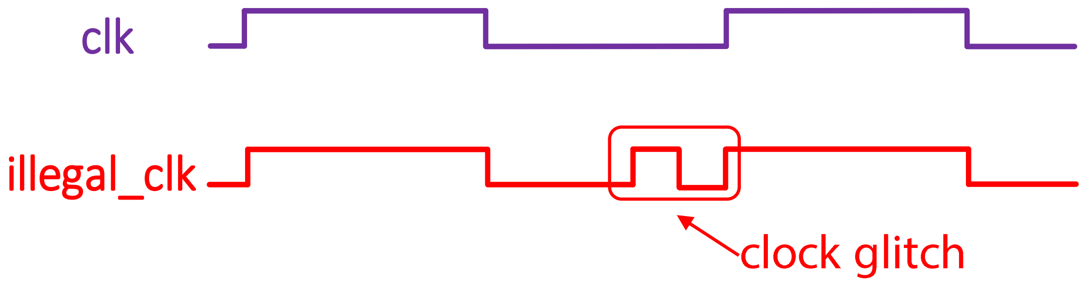


- What if we know the value of the glitch? How?

Clock-glitch-based fault attack

- Basic idea
  - By increasing the clock frequency, the attacker can get information from the abnormal behavior of the device
- One cycle fault
  - clock glitch





Outline

- Preliminaries
  - Glitches in combinational circuits
  - Clock-glitch-based fault attack
- Transient-steady effect attack
  - Basic idea
  - Attack on masked an unmasked S-Boxes
  - Experiments
- Further discussion
  - Attack scenario of parallel AES implementation
  - Attack scenario of WDDL-AES
  - Glitch injection
- Conclusion

Transient-steady effect

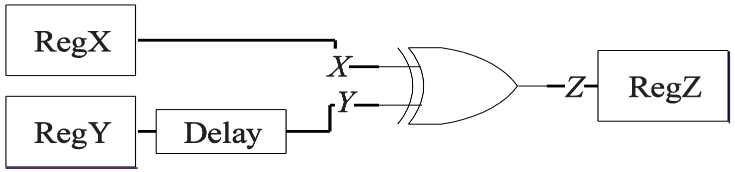
- Definition:
  - the output of a gate turns to a temporal value and keeps steady for a while before it switches to the final steady value
- The difference of propagation delays is large

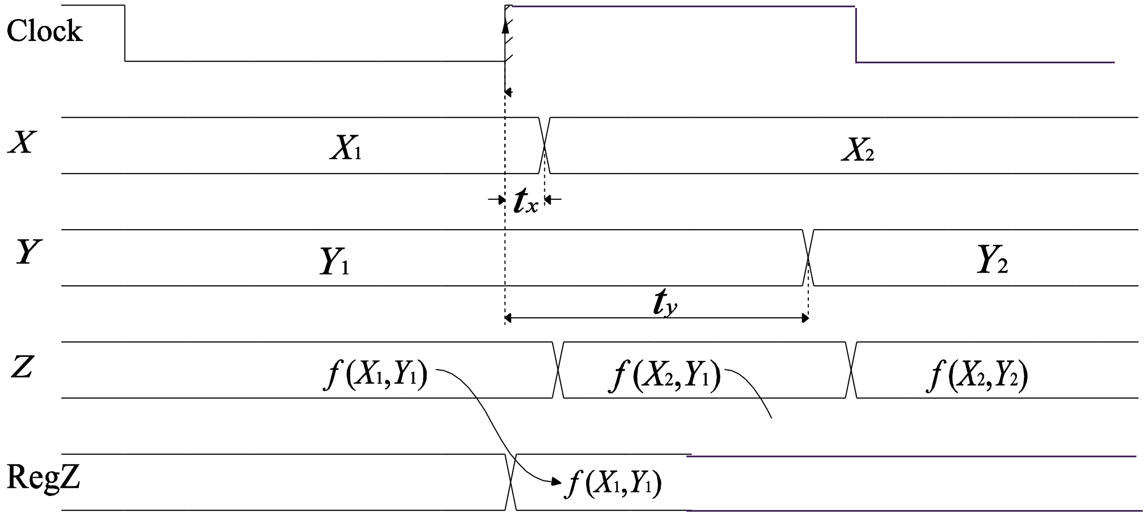
  → the glitch lasts long enough

  → transient-steady effect

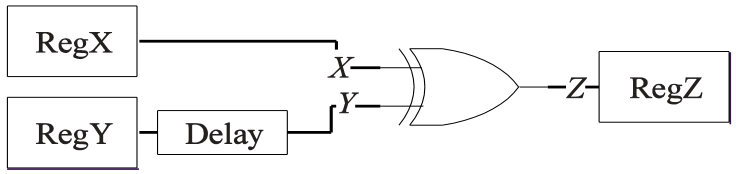
- Transient-steady effect + clock-glitch-based fault attack = Transient-steady effect attack (TSE attack)

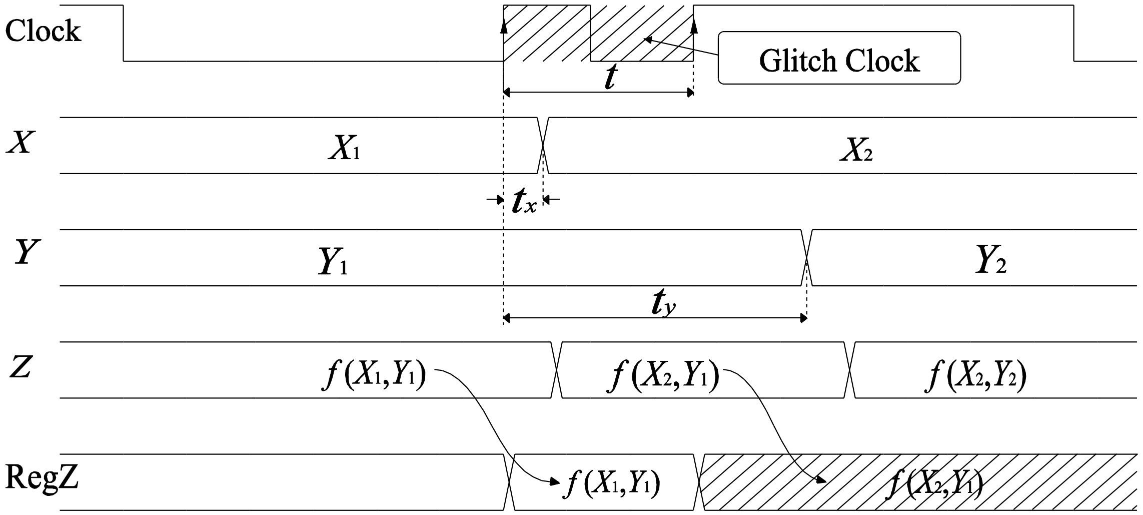
Basic idea


RegX
RegY
Delay
X
Y
Z
RegZ


Clock
X
X_1
X_2
t_x
Y
Y_1
Y_2
t_y
Z
$f(X_1,Y_1)$
$f(X_2,Y_1)$
$f(X_2,Y_2)$
RegZ
$f(X_1,Y_1)$

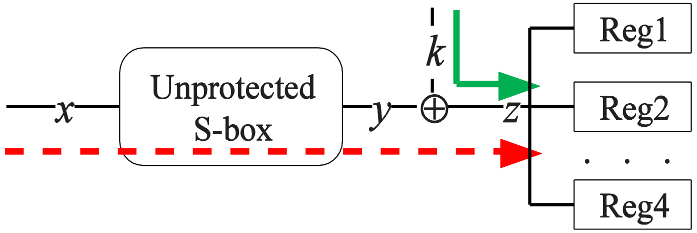
Basic idea


RegX
RegY
Delay
X
Y
Z
RegZ


Clock
Glitch Clock
t
X
X_1
X_2
t_x
Y
Y_1
Y_2
t_y
Z
$f(X_1,Y_1)$
$f(X_2,Y_1)$
$f(X_2,Y_2)$
RegZ
$f(X_1,Y_1)$
$f(X_2,Y_1)$

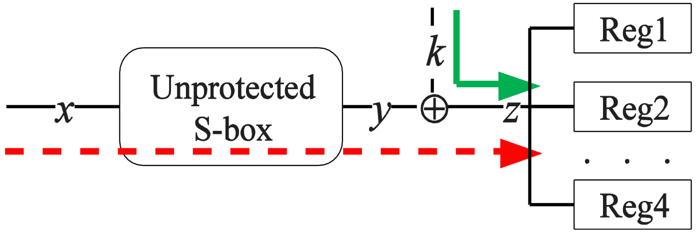
Attack on unmasked S-Box

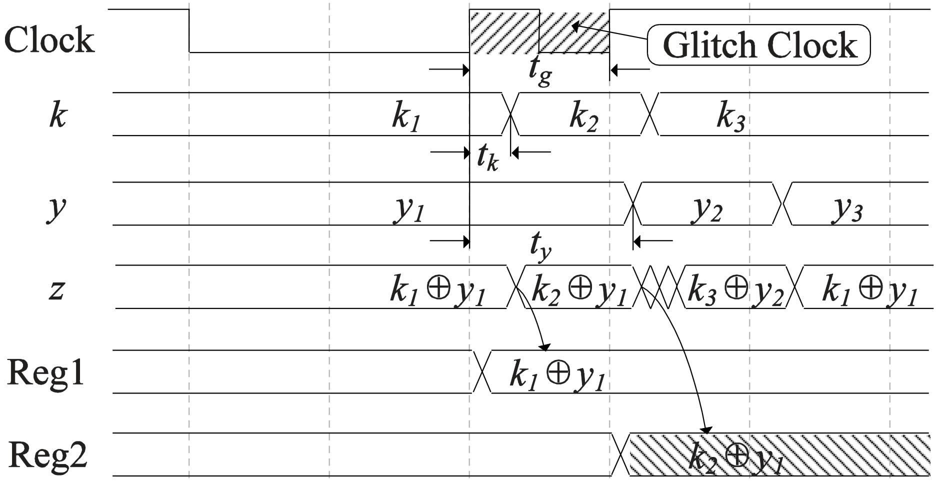
- The serial implementation
- The final AES round
- No specific requirement of the structure of the S-box



Attack on unmasked S-Box

- The serial implementation
- The final AES round
- No specific requirement of the structure of the S-box

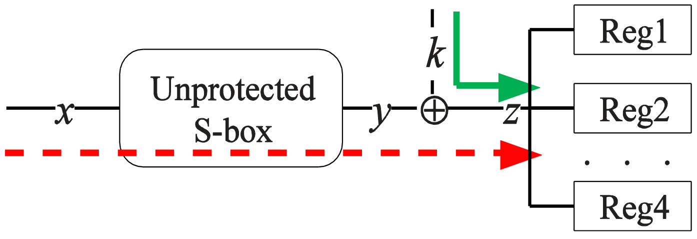


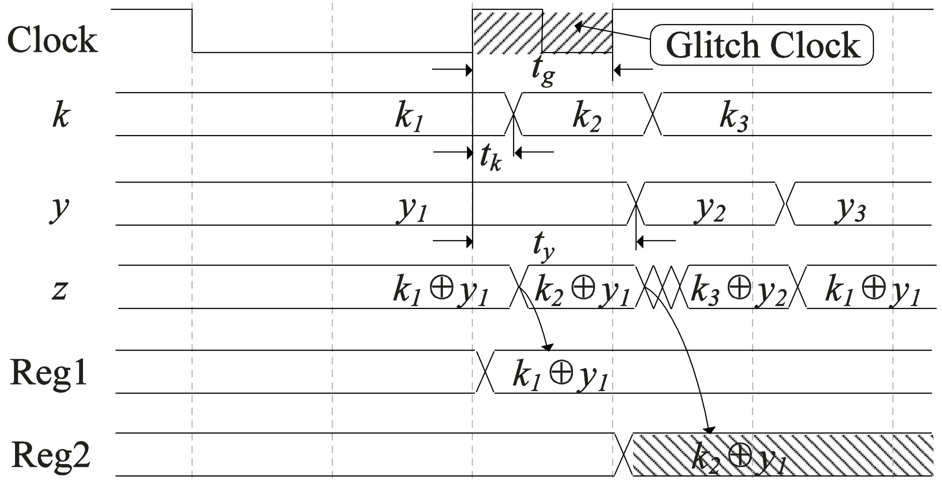


- $t_x < t < t_y$

Attack on unmasked S-Box

- The serial implementation
- The final AES round
- No specific requirement of the structure of the S-box





- $t_x < t < t_y$
- $t_x? \; t_y?$

$$z_1 \oplus \tilde{z}_2 = y_1 \oplus k_1 \oplus y_1 \oplus k_2 = k_1 \oplus k_2$$
$$= \Delta k_{1,2}$$

TSE Attack

- Step 1. Sweep the glitch frequency
  - At every frequency point, do encryptions with fixed x_1 and random x_2 for N_{pre} times, and record the outputs

TSE Attack

- Step 1. Sweep the glitch frequency
  - At every frequency point, do encryptions with fixed x_1 and random x_2 for N_{pre} times, and record the outputs
- Step 2. Find the feasible range of glitch frequency
  - With a fixed x_1, z_1 is a fixed value
  - If TSE attack succeeds, $\tilde{z}_2 = k_2 \oplus y_1$ is also a fixed value
  - Fixed output is the sign of feasible frequency

TSE Attack

- Step 1. Sweep the glitch frequency
  - At every frequency point, do encryptions with fixed x_1 and random x_2 for N_{pre} times, and record the outputs
- Step 2. Find the feasible range of glitch frequency
  - With a fixed x_1, z_1 is a fixed value
  - If TSE attack succeeds, $\tilde{z}_2 = k_2 \oplus y_1$ is also a fixed value
  - Fixed output is the sign of feasible frequency
- Step 3. Carry out TSE attack at a feasible glitch frequency
  - Do encryptions for N_{attack} times
  - Compute the attack result $z_1 \oplus \tilde{z}_2$ for every encryption
  - Choose the value with has the greatest occurrence rate in the attack results as the value of $\Delta k_{1,2}$
- Step 4. Repeat Step 3 for $\Delta k_{2,3}$, $\Delta k_{3,4}$ and so on

TSE Attack

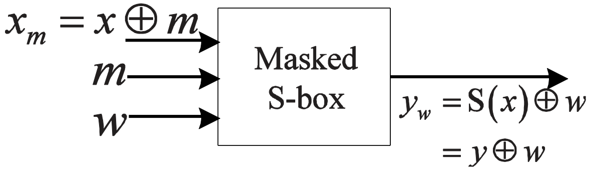
- Step 1. Sweep the glitch frequency
  - At every frequency point, do encryptions with fixed x_1 and random x_2 for N_{pre} times, and record the outputs
- Step 2. Find the feasible range of glitch frequency
  - With a fixed x_1, z_1 is a fixed value
  - If TSE attack succeeds, $\tilde{z}_2 = k_2 \oplus y_1$ is also a fixed value
  - Fixed output is the sign of feasible frequency

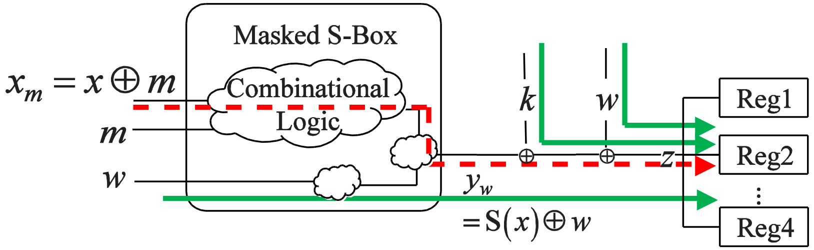
Pre-computation stage

- Step 3. Carry out TSE attack at a feasible glitch frequency
  - Do encryptions for N_{attack} times
  - Compute the attack result $z_1 \oplus \tilde{z}_2$ for every encryption
  - Choose the value with has the greatest occurrence rate in the attack results as the value of $\Delta k_{1,2}$
- Step 4. Repeat Step 3 for $\Delta k_{2,3}$, $\Delta k_{3,4}$ and so on

Attack stage

Attack on masked S-Box





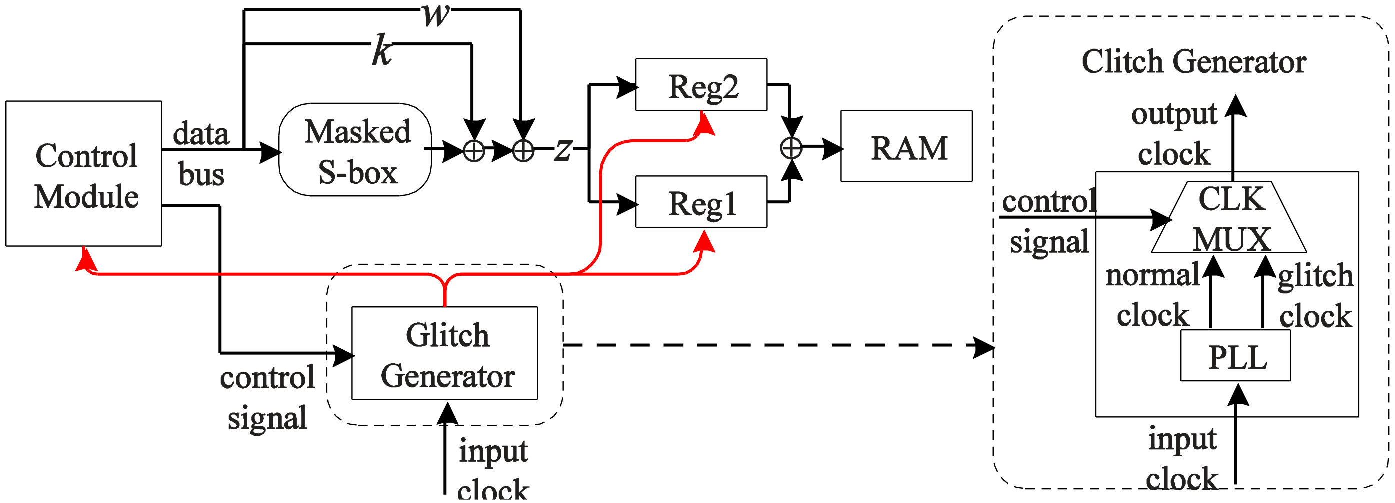
$$\begin{aligned} z_1 &= y_{w_1} \oplus k_1 \oplus w_1 \\ &= S(x_1) \oplus w_1 \oplus k_1 \oplus w_1 \\ &= S(x_1) \oplus k_1\ . \end{aligned}$$

$$\begin{aligned} \tilde{z}_2 &= \tilde{y}_{w_2} \oplus k_2 \oplus w_2 \\ &= S(x_1) \oplus w_2 \oplus k_2 \oplus w_2 \\ &= S(x_1) \oplus k_2 \end{aligned}$$

$$\begin{aligned} z_1 \oplus \tilde{z}_2 &= S(x_1) \oplus k_1 \oplus S(x_1) \oplus k_2 \\ &= k_1 \oplus k_2 \\ &= \Delta k_{1,2}\ . \end{aligned}$$

Experiments

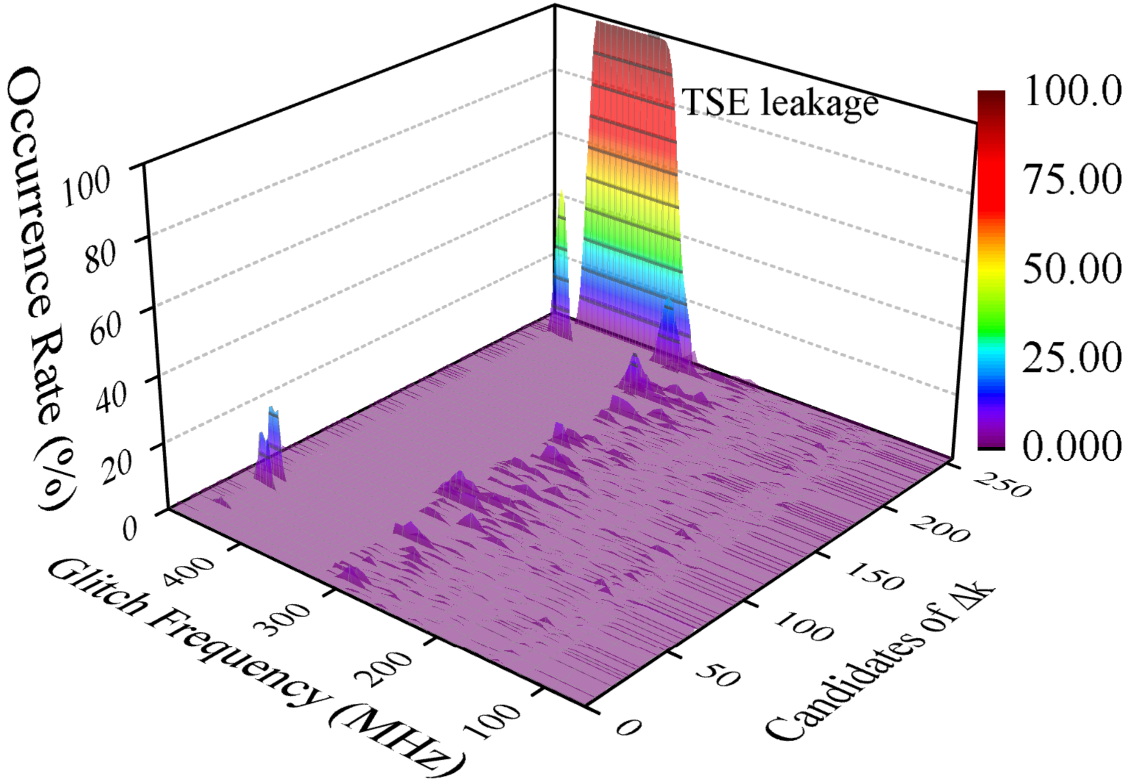
- Two unmasked S-boxes, one masked S-box
- DE2-115 FPGA board



- Without fault: $z_1 \oplus z_2 = S(x_1) \oplus S(x_2) \oplus k_1 \oplus k_2$
- TSE attack succeeds: $z_1 \oplus z_2 = k_1 \oplus k_2 = \Delta k_{1,2}$

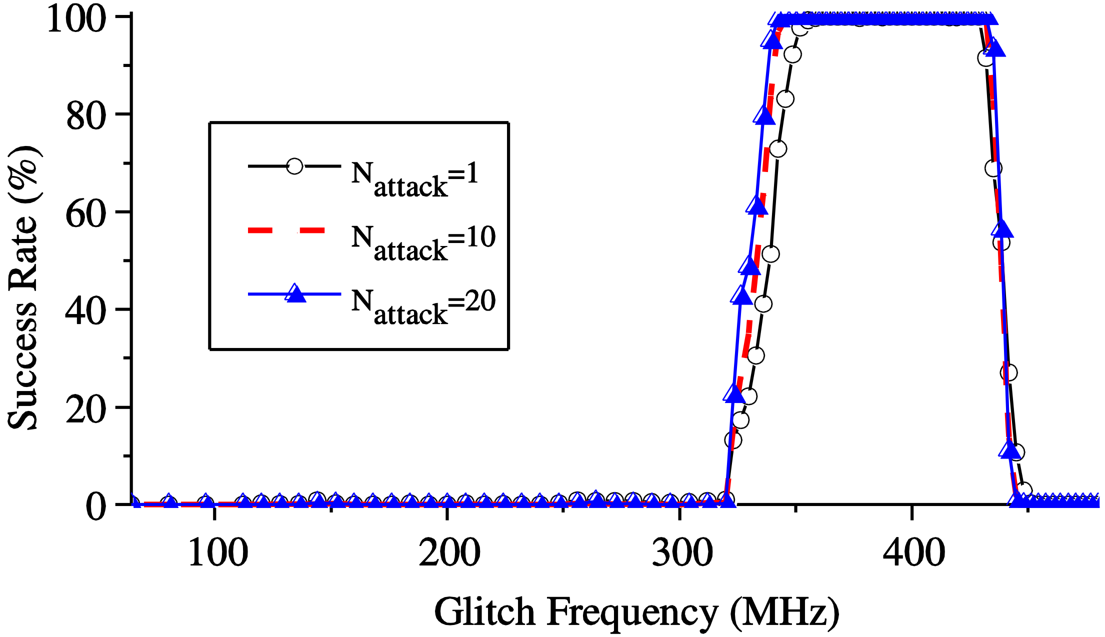
Experiment on unmasked S-box A

- $k_1 = 0XE2, k_2 = 0X19, \Delta k_{1,2} = 0XFB$
- Pre-computation stage
    - Sweep the frequency from 64MHz to 480MHz
    - $x_1 = 0X31, x_2 \in [0,255], N_{pre} = 65536$
- 360MHz ~ 430MHz



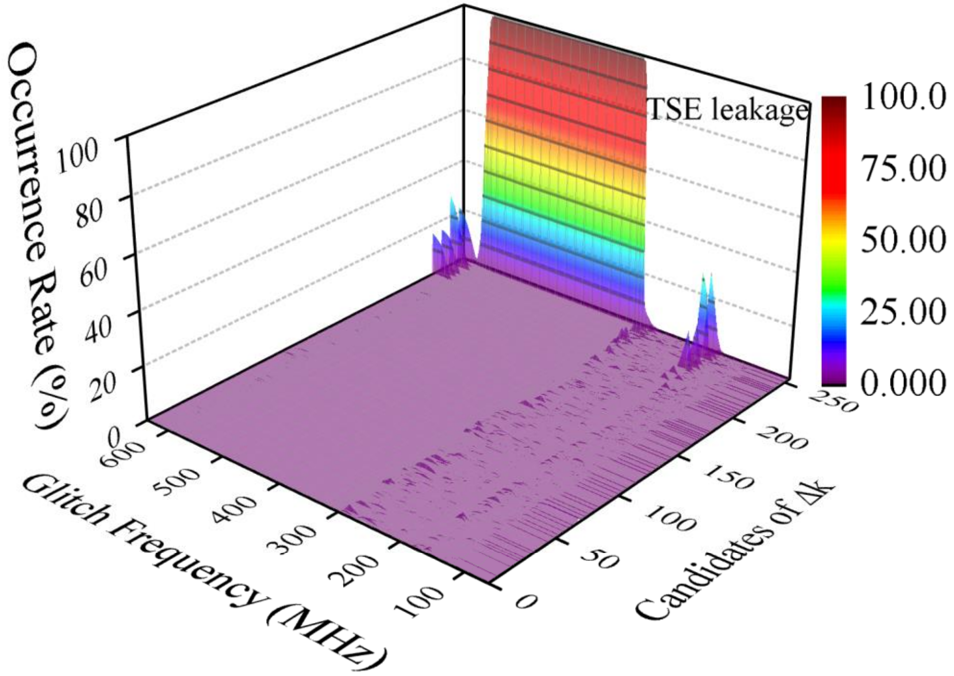
Experiment on unmasked S-box A

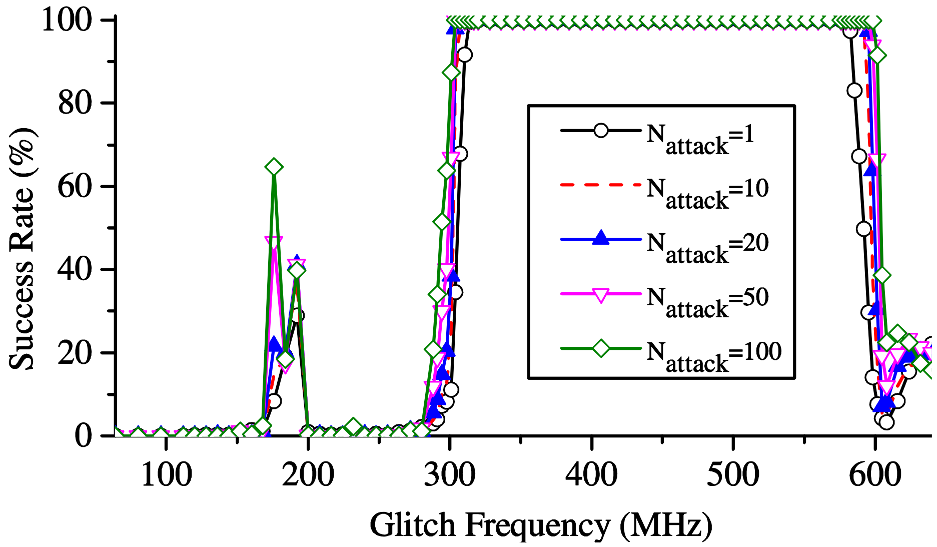
- $k_1 = 0\text{XE}2, k_2 = 0\text{X}19, k_1 \oplus k_2 = 0\text{XFB}$
- Attack stage
  - Feasible frequency range: 360MHz ~ 430MHz



Experiment on unmasked S-box B

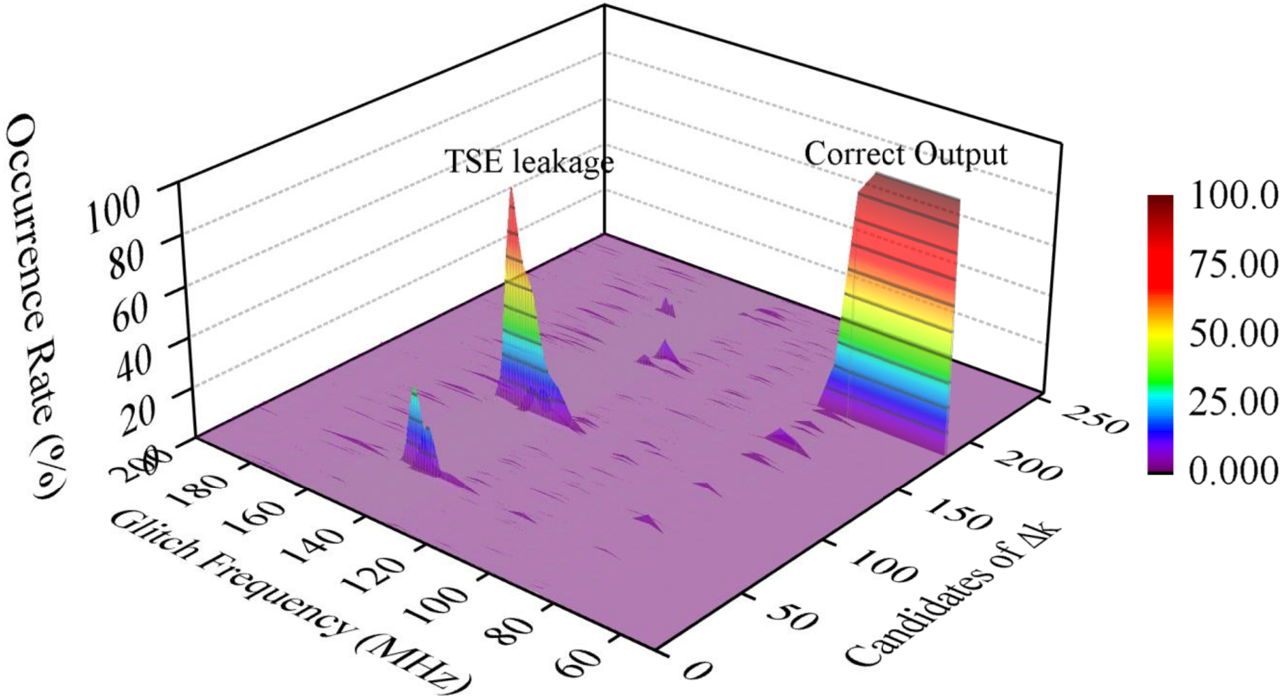
- Feasible frequency range: 320MHz ~ 580MHz



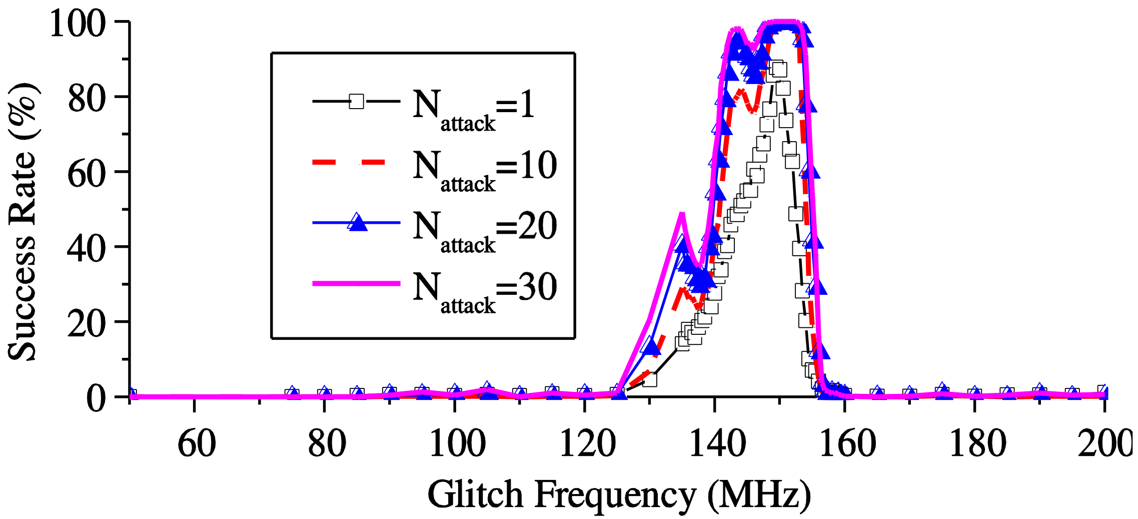


Experiment on unmasked S-box C

- $k_1 = 0X3F, k_2 = 0X58, \Delta k_{1,2} = 0X67$
- Pre-computation stage
  - $x_1 = 0X9D, x_2 = 0XE6$, masks are randomly chosen
  - Without fault: $0XB7$
  - Attack succeeds: $0X67$
- 145~150MHz



Experiment on unmasked S-box C



Efficiency Comparison

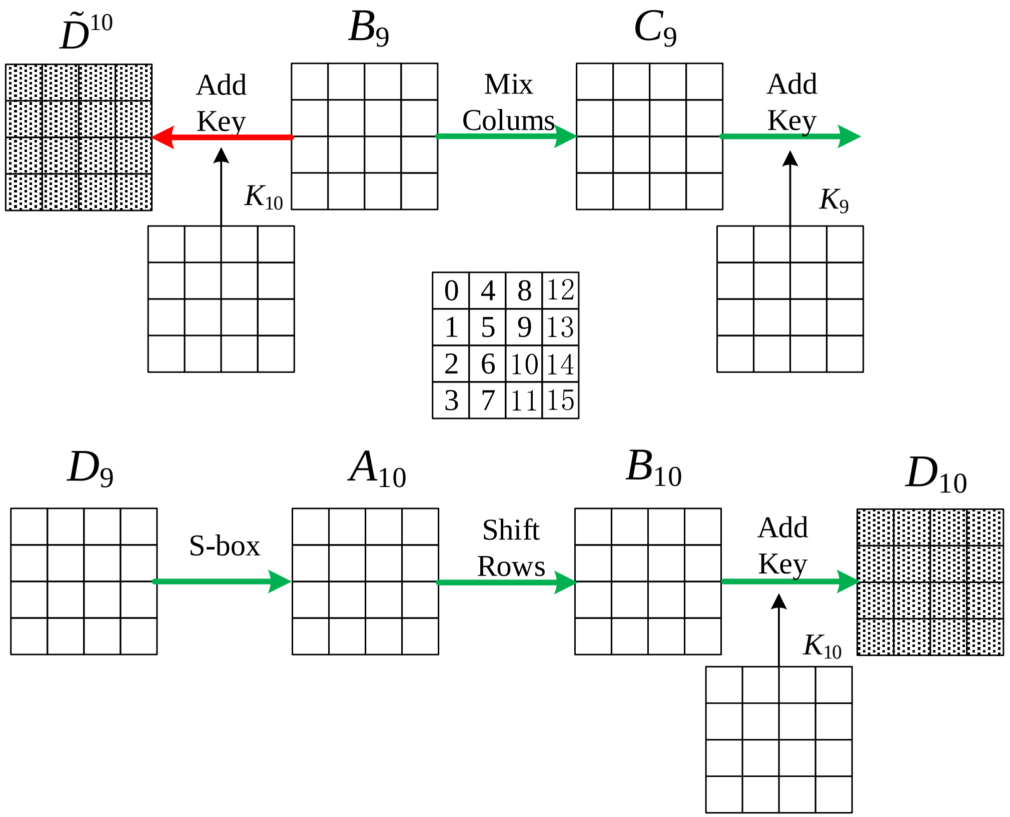
Method	FSA[9]	CTC[10]	FRA[12]	TSE Attack	TSE Attack
Target S-box	Unmasked	Masked	Masked	Unmasked	Masked
Num of Enc	840	1 000 000	80 000	1	20
Space (bytes)	120	2048	80	1	20
Offline Complexity	$256\,C_{\rho_7}$	$256\,C_{\rho_{256}}$	$1\,C_{div}$	≈ 0	≈ 0
Num of Pre-Enc	0	0	0	40 000	40 000

Outline

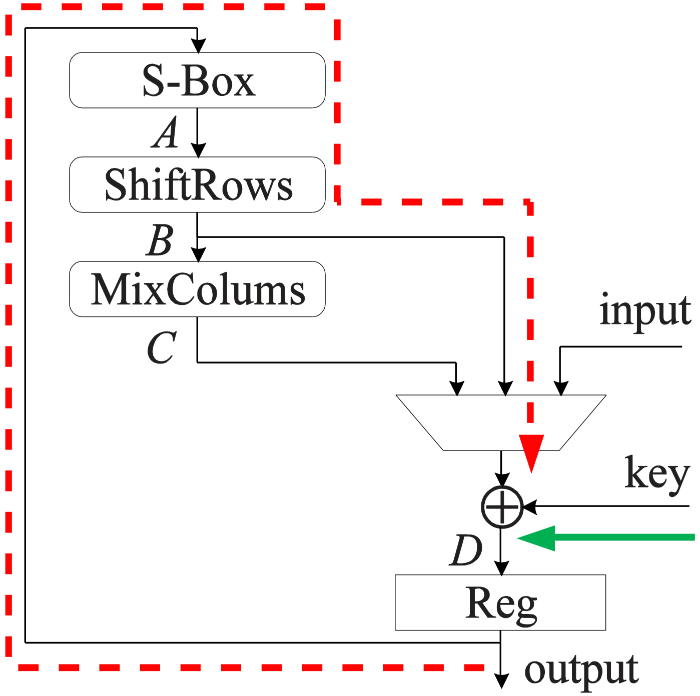
- Preliminaries
  - Glitches in combinational circuits
  - Clock-glitch-based fault attack
- Transient-steady effect attack
  - Basic idea
  - Attack on masked an unmasked S-Boxes
  - Experiments
- Further discussion
  - Attack scenario of parallel AES implementation
  - Attack scenario of WDDL-AES
  - Glitch injection
- Conclusion

Attack scenario of parallel AES implementation

- Without fault: $D_{10} = B_{10} + K_{10}$
- Attack succeeds: $\widetilde{D}_{10} = B_9 + K_{10}$



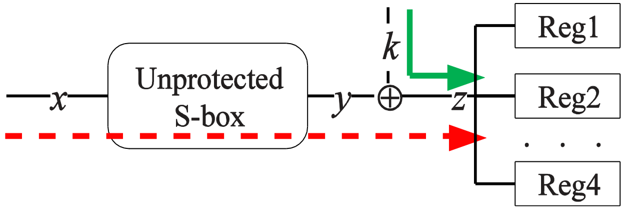
0	4	8	12
1	5	9	13
2	6	10	14
3	7	11	15

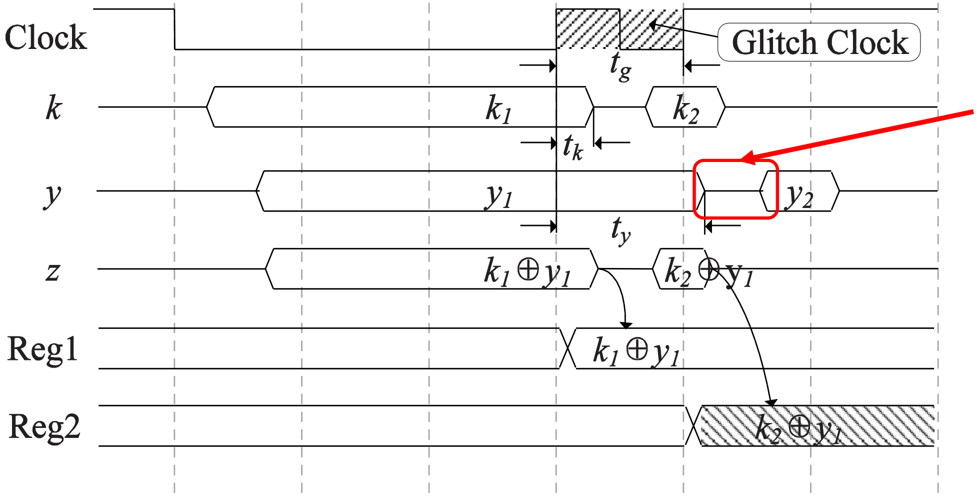


- One round AES
- Plaintext: D_{10}
- Ciphertext: $\widetilde{D}_{10}$

Attack scenario of parallel WDDL-AES

- Dual-rail precharge logic
  - Precharge phase: (0,0) (all-zero state)
  - Evaluation phase: (0,1) or (1,0)

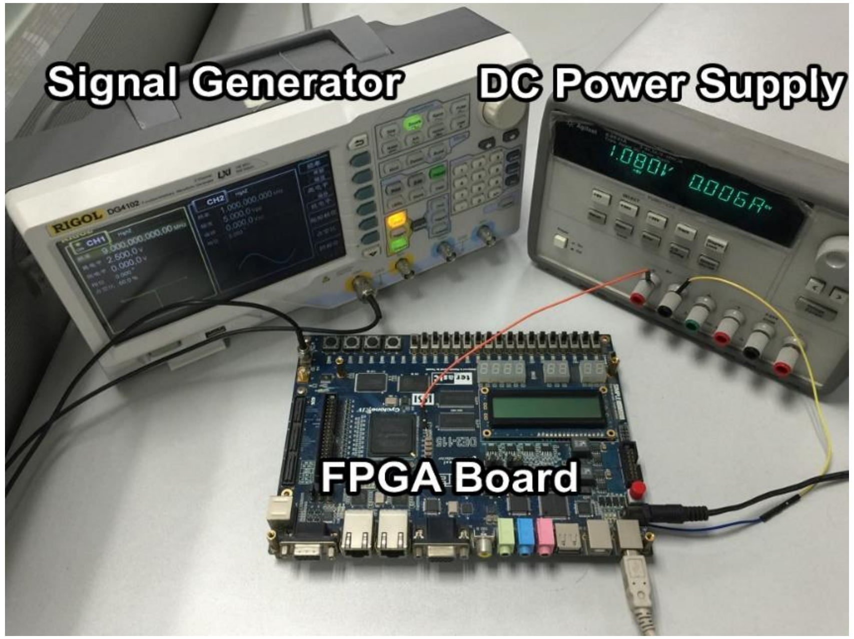




all-zero state

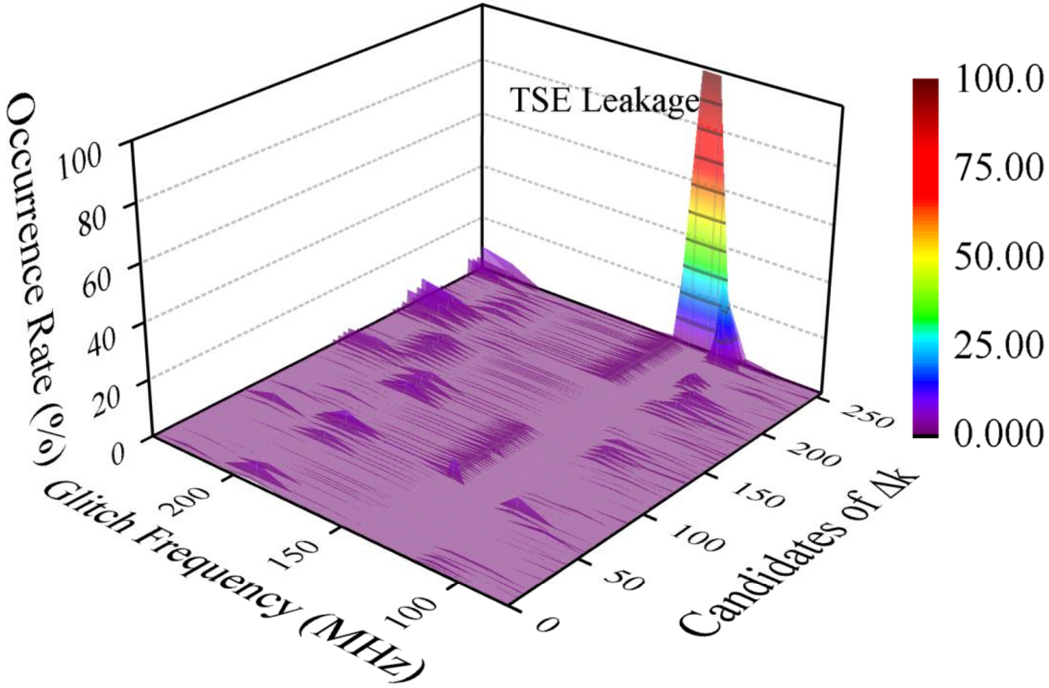
Glitch injection

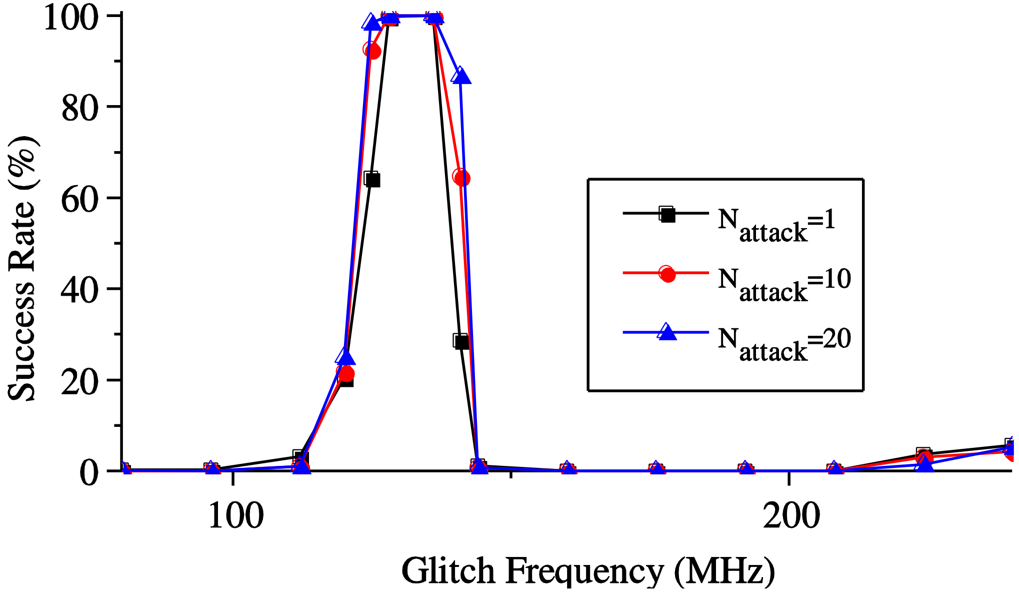
- The feasibility of injecting clock glitch externally
  - <2.8ns (S-box A)
  - May be filtered out when injected externally
- Solutions
  - Semi-invasive attack
  - Slow down the target circuit



Glitch injection

- Experiment on S-box A with reduced voltage
- 1.50V: 360 ~ 430MHz
- 1.08V: 125 ~ 136MHz





Conclusion

- We propose a new TSE attack based on the **transient-steady effect**
- We conduct experiments on **two kinds** of unmasked S-boxes and **one kind** of masked S-box
- Experimental results show that TSE attack can recover a key byte of an unmasked S-box with **1** encryption, and a masked S-box with less than **20** encryptions
- The attack scenarios on parallel AES implementation and WDDL-AES are also discussed
- The foundation of TSE attack is that the key's data path is obviously shorter than other signals'
- Countermeasure: **increase the delay of the key**

Thank you!