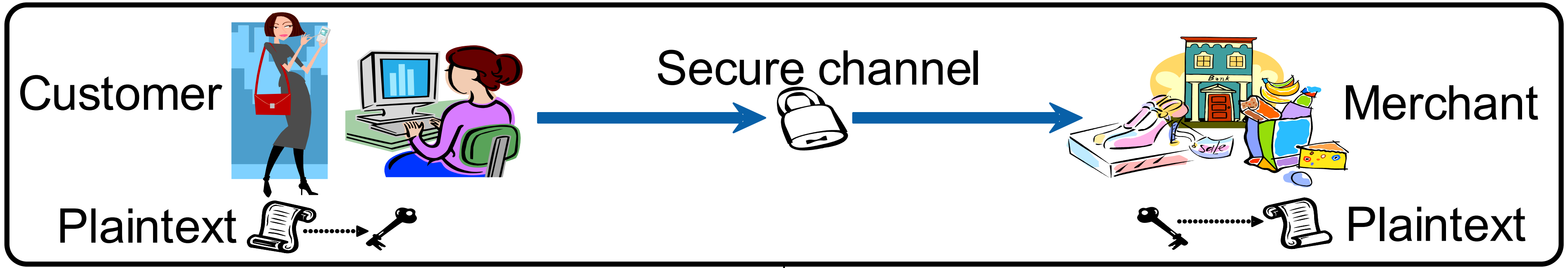
GLSVLSI 2022

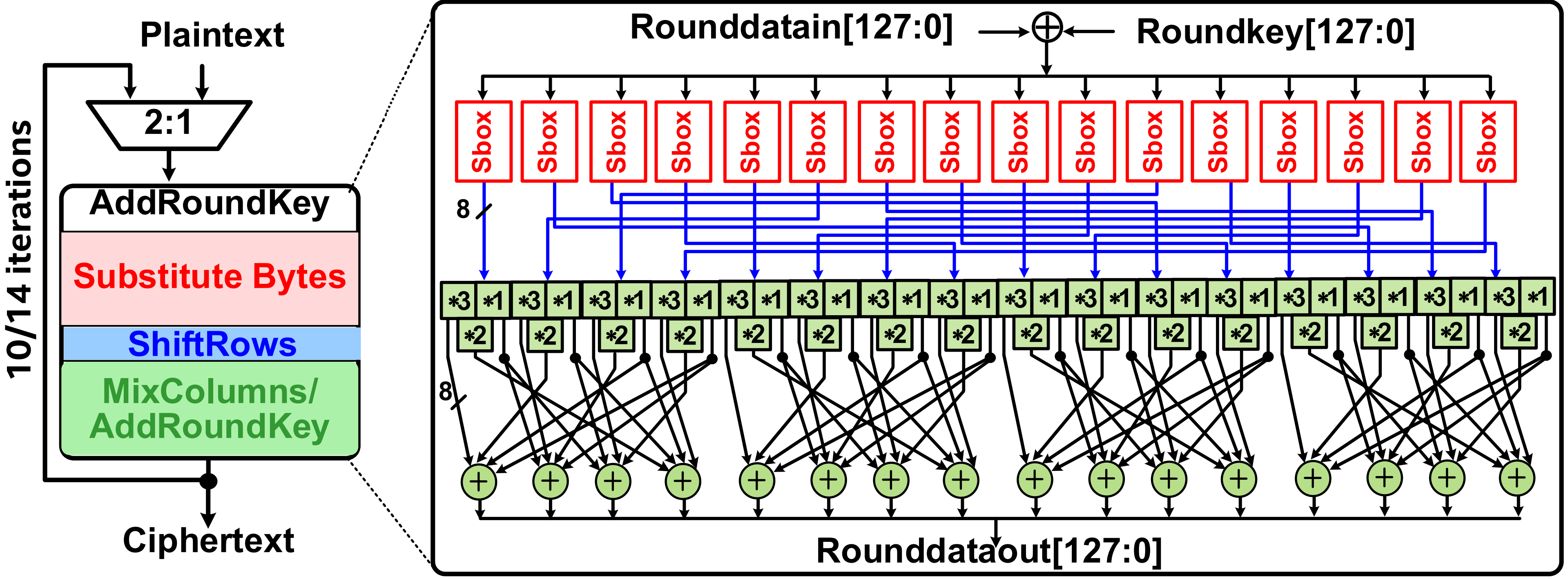
Attack Resistant Crypto Circuits

Sanu Mathew

Circuits Research Labs
Hillsboro, Oregon

Motivation: Securing Data

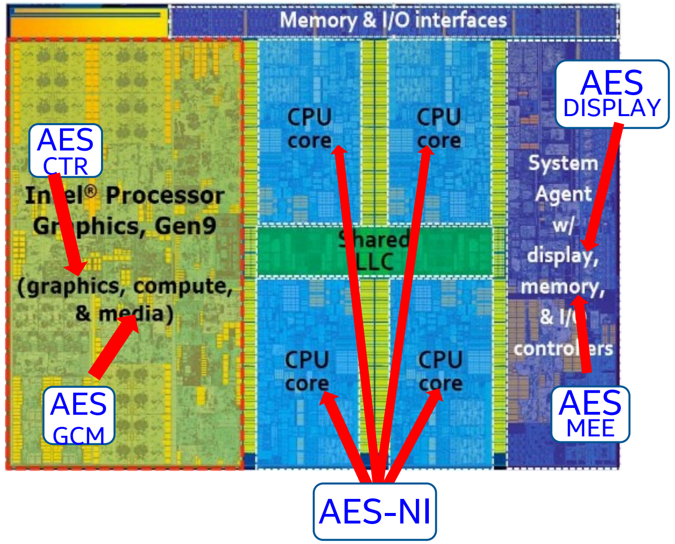




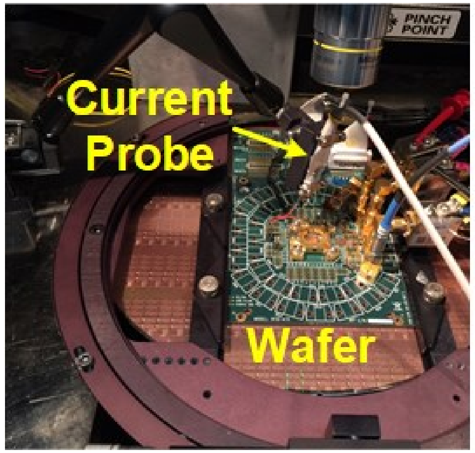
AES: de-facto symmetric-key encryption algorithm

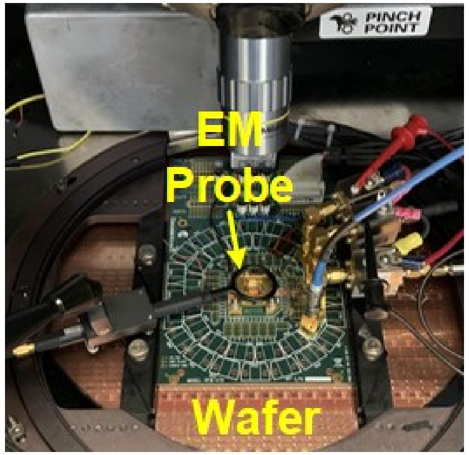
- 10-14 round iterative cipher encrypting 128b plaintext using secret key

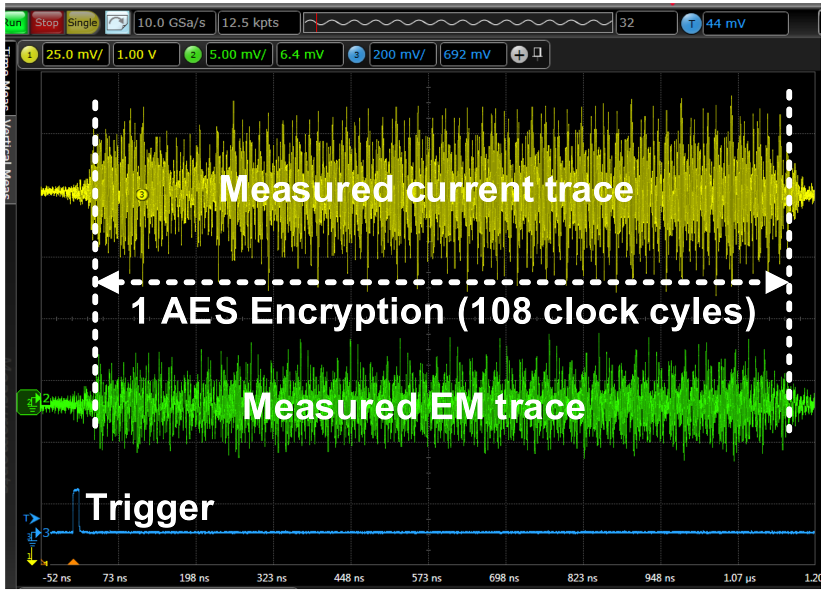
Power/EM Side-channel Attacks on AES



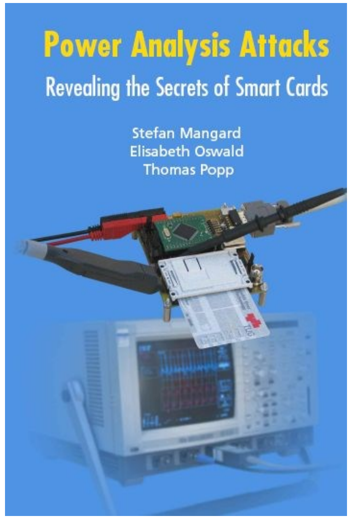
Memory & I/O interfaces
AES CTR
Intel® Processor Graphics, Gen9
(graphics, compute, & media)
AES GCM
CPU core
CPU core
Shared LLC
CPU core
CPU core
AES DISPLAY
System Agent w/ display, memory, & I/O controllers
AES MEE
AES-NI





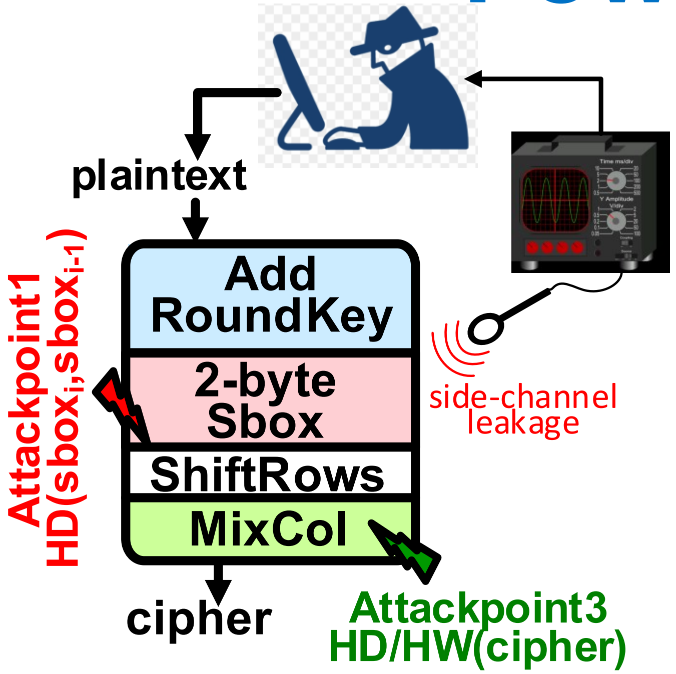


- Multiple AES instances for AES-NI, display, memory, IO, secure DFx, media.
- Key-related switching activity present in current and EM signatures.
- Correlation power analysis extracts correlation between keys and measured traces.



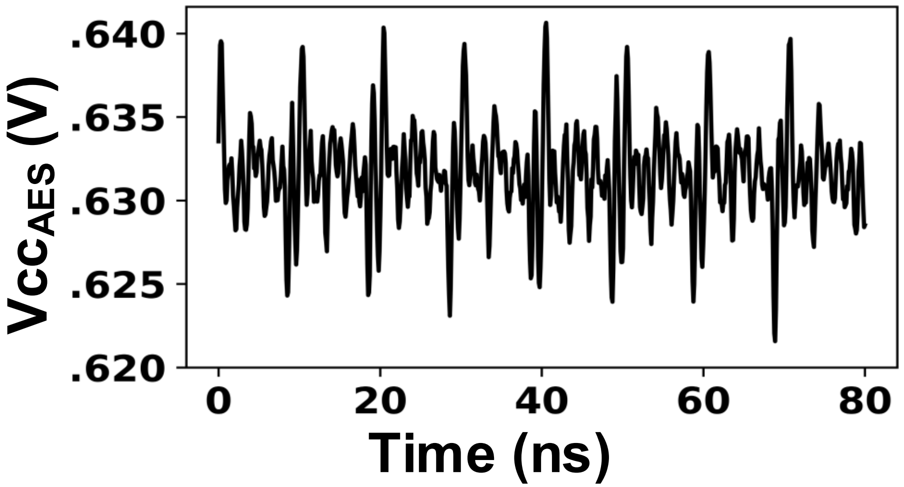
Physical access to device implies access to embedded secrets

Power/EM Side-channel Attacks

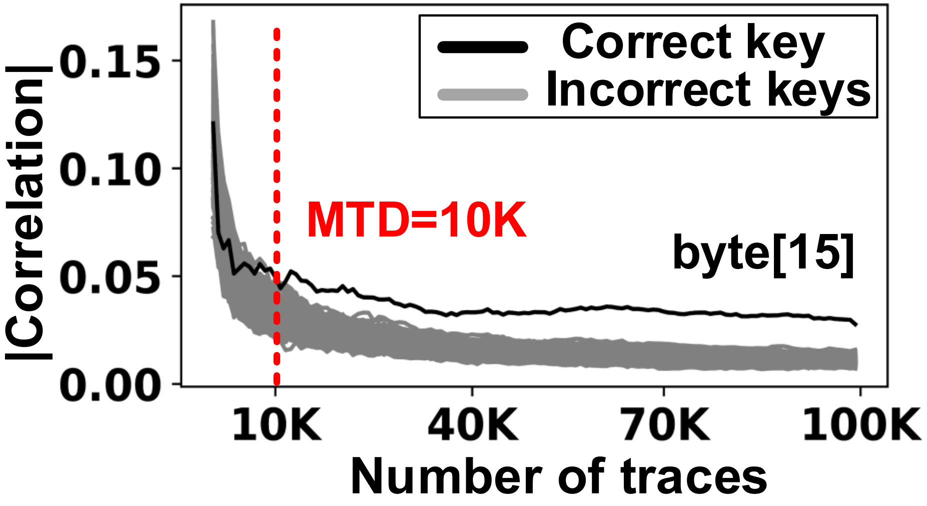


HD - Hamming distance
HW – Hamming weight

Measured current traces



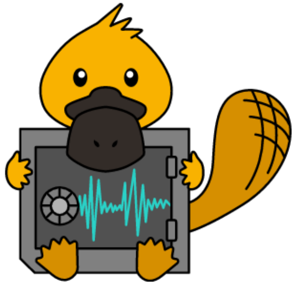
14nm CMOS, 0.75V, 100MHz, 25°C



- Attacker sends in random plaintext and builds power models with key guesses
- Collects current/EM traces as the chip encrypts data
- With large number of traces, the model with correct key guess shows high correlation

AES keys extracted within 1hr

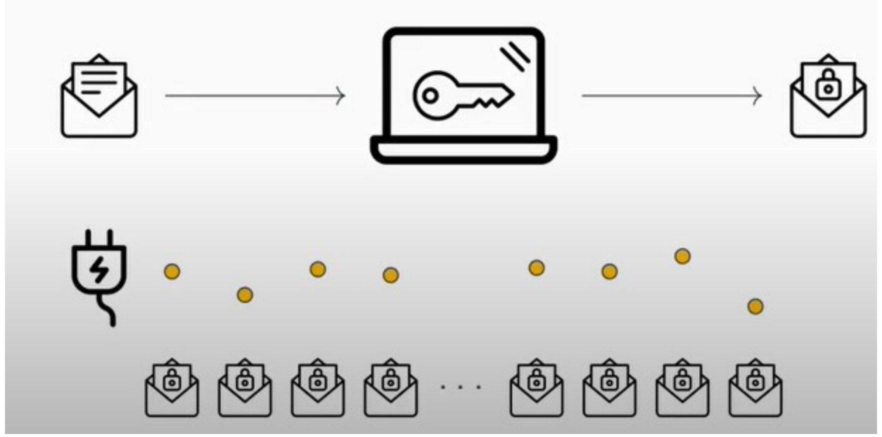
Remote Side-channel Attack on AES: Platypus

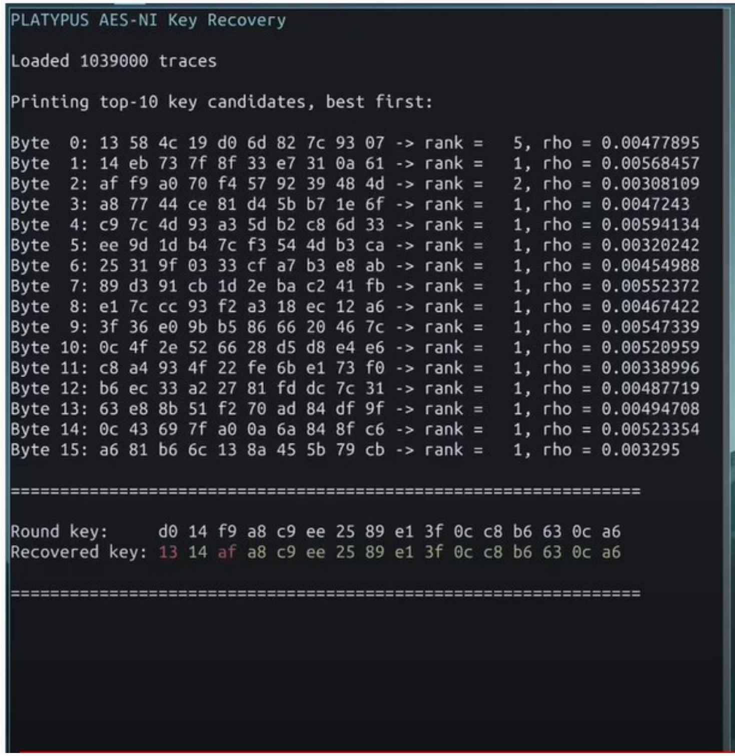


PLATYPUS

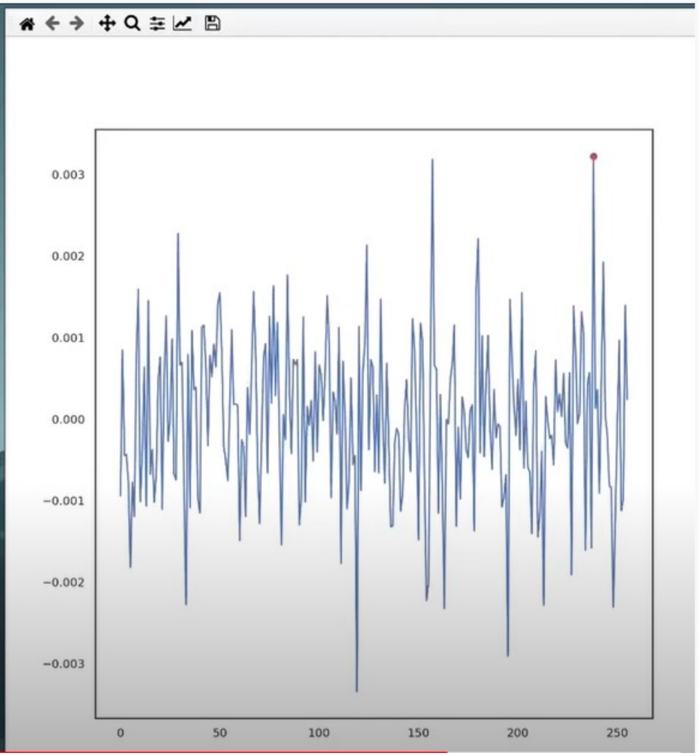
WITH GREAT POWER COMES GREAT LEAKAGE

With PLATYPUS, we present novel software-based **power side-channel attacks** on Intel server, desktop and laptop CPUs. We exploit the **unprivileged access** to the Intel RAPL interface exposing the processor's power consumption to **infer data** and **extract cryptographic keys**.





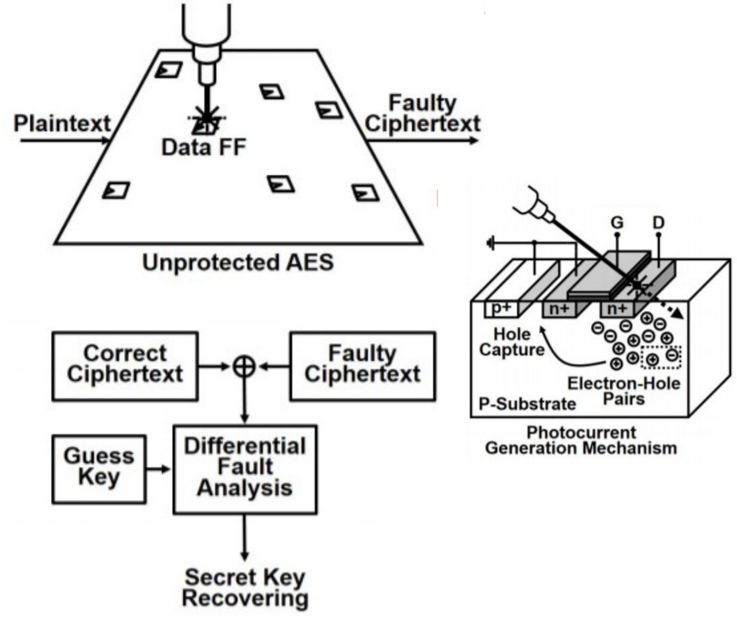
```
PLATYPUS AES-NI Key Recovery

Loaded 1039000 traces

Printing top-10 key candidates, best first:

Byte  0: 13 58 4c 19 d0 6d 82 7c 93 07 -> rank =   5, rho = 0.00477895
Byte  1: 14 eb 73 7f 8f 33 e7 31 0a 61 -> rank =   1, rho = 0.00568457
Byte  2: af f9 a0 70 f4 57 92 39 48 4d -> rank =   2, rho = 0.00308109
Byte  3: a8 77 44 ce 81 d4 5b b7 1e 6f -> rank =   1, rho = 0.0047243
Byte  4: c9 7c 4d 93 a3 5d b2 c8 6d 33 -> rank =   1, rho = 0.00594134
Byte  5: ee 9d 1d b4 7c f3 54 4d b3 ca -> rank =   1, rho = 0.00320242
Byte  6: 25 31 9f 03 33 cf a7 b3 e8 ab -> rank =   1, rho = 0.00454988
Byte  7: 89 d3 91 cb 1d 2e ba c2 41 fb -> rank =   1, rho = 0.00552372
Byte  8: e1 7c cc 93 f2 a3 18 ec 12 a6 -> rank =   1, rho = 0.00467422
Byte  9: 3f 36 e0 9b b5 86 66 20 46 7c -> rank =   1, rho = 0.00547339
Byte 10: 0c 4f 2e 52 66 28 d5 d8 e4 e6 -> rank =   1, rho = 0.00520959
Byte 11: c8 a4 93 4f 22 fe 6b e1 73 f0 -> rank =   1, rho = 0.00338996
Byte 12: b6 ec 33 a2 27 81 fd dc 7c 31 -> rank =   1, rho = 0.00487719
Byte 13: 63 e8 8b 51 f2 70 ad 84 df 9f -> rank =   1, rho = 0.00494708
Byte 14: 0c 43 69 7f a0 0a 6a 84 8f c6 -> rank =   1, rho = 0.00523354
Byte 15: a6 81 b6 6c 13 8a 45 5b 79 cb -> rank =   1, rho = 0.003295

================================================================

Round key:     d0 14 f9 a8 c9 ee 25 89 e1 3f 0c c8 b6 63 0c a6
Recovered key: 13 14 af a8 c9 ee 25 89 e1 3f 0c c8 b6 63 0c a6

================================================================
```



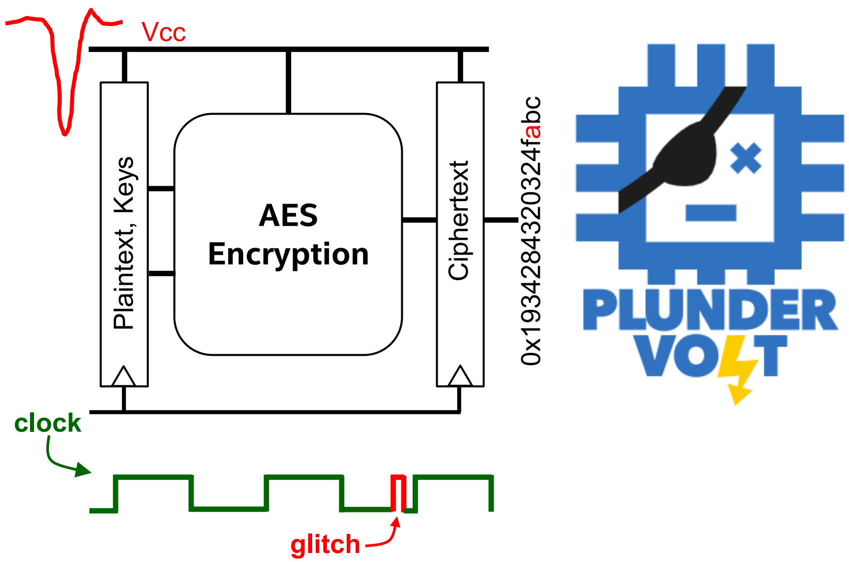
- Platypus attack published by in 2021
- Exploits RAPL (Running Average Power Limit) interface to monitor CPU power consumption
- AES keys recovered in <1M traces

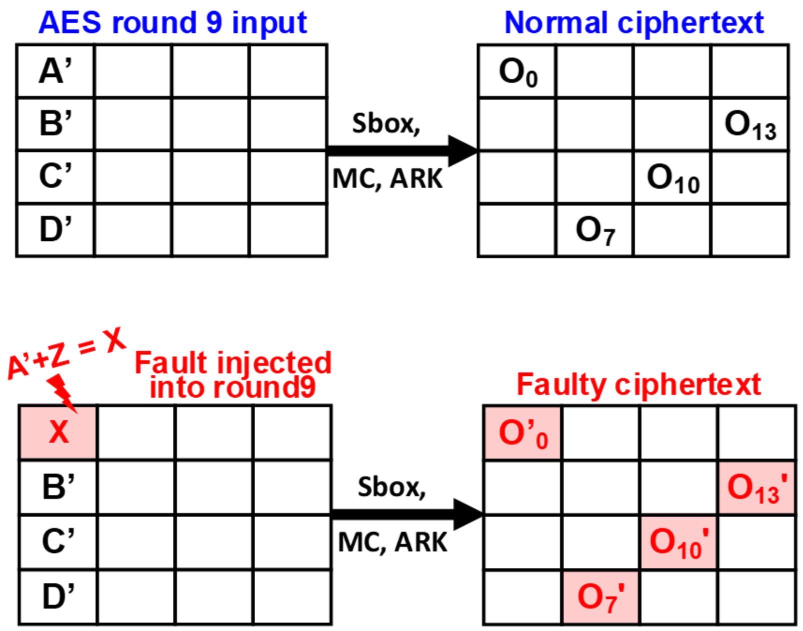
Fault-injection Attacks on AES

Laser fault injection



Voltage/Clock glitch attack

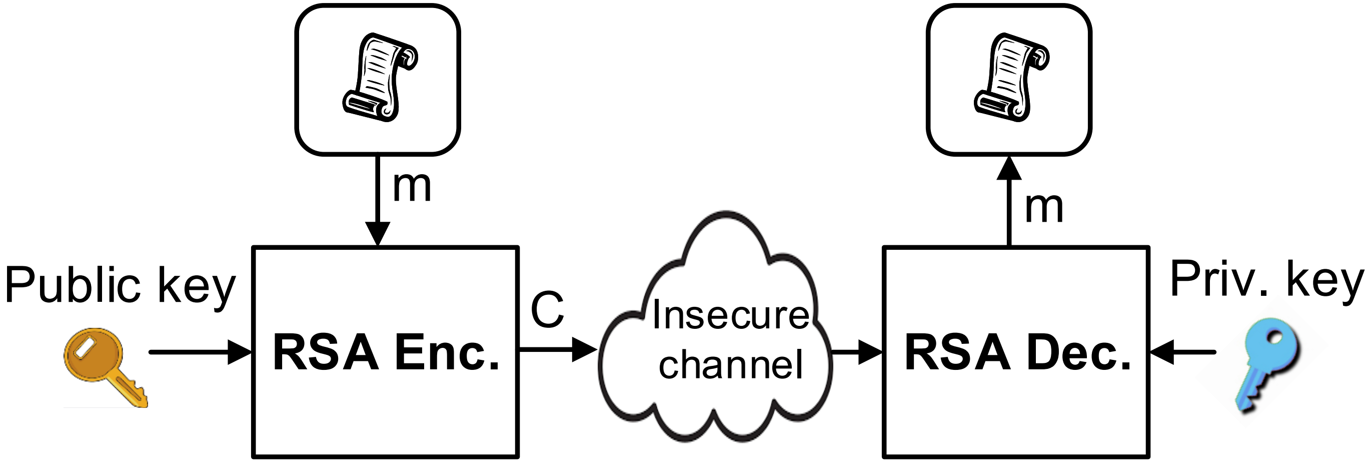




Differential cryptanalysis compares normal and faulty outputs

- Malicious fault-injection using voltage/clock-glitch, laser attack
- PlunderVolt: manipulates DVFS to inject faults in MEE
- One injected fault corrupts 4 ciphertext output bytes
- Reduces AES key security from 128b to 32b

Public Key Cryptography with RSA



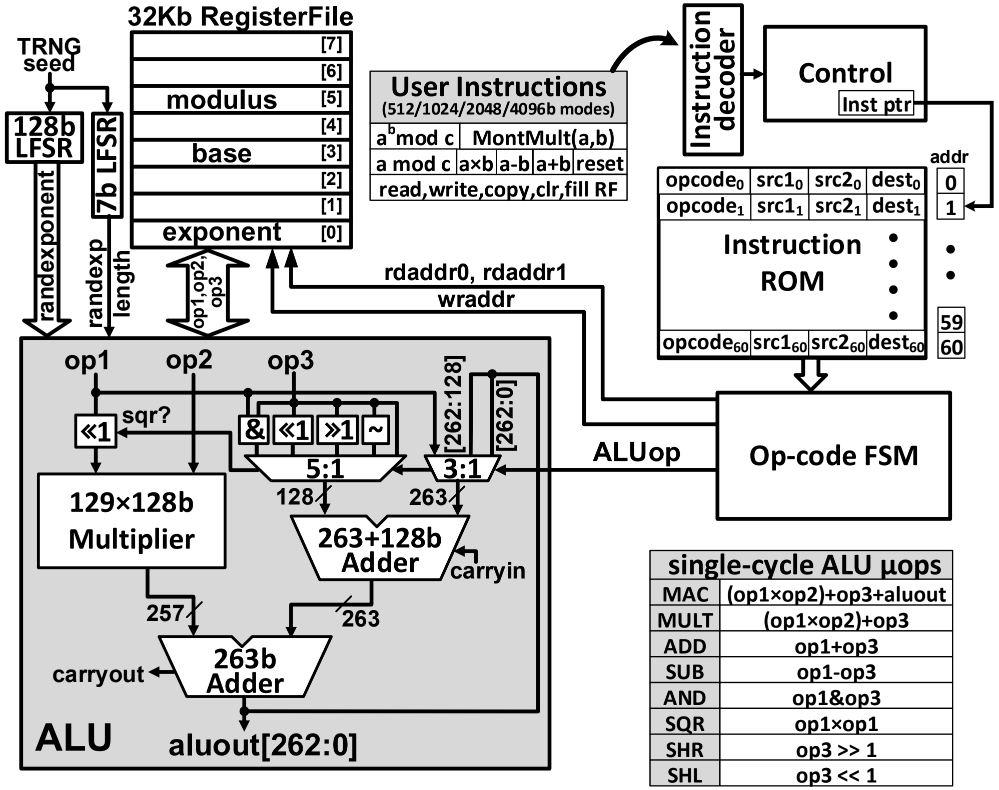
Public key crypto is a critical component of secure systems

- Applications: Key exchange, digital signatures, authentication, etc.

RSA is gaining renewed interest with quantum computer attacks

- ECC prone to quantum attacks due to shorter key lengths
- Key lengths > 4K currently employed in cryptosystems
- RSA performance determines sign/verification latencies

Side-channel Attacks on Public-key RSA

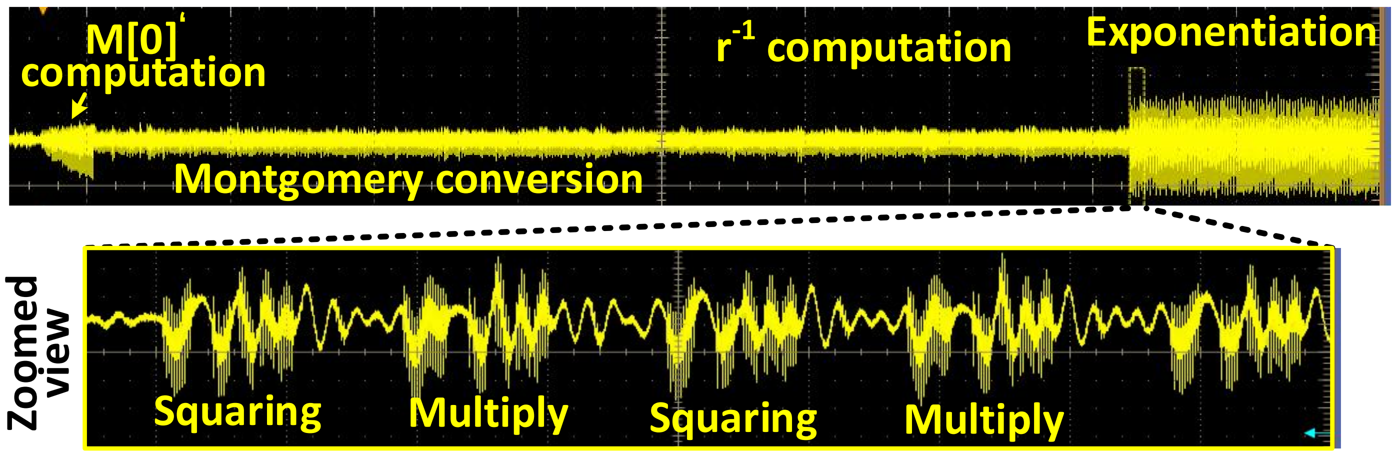


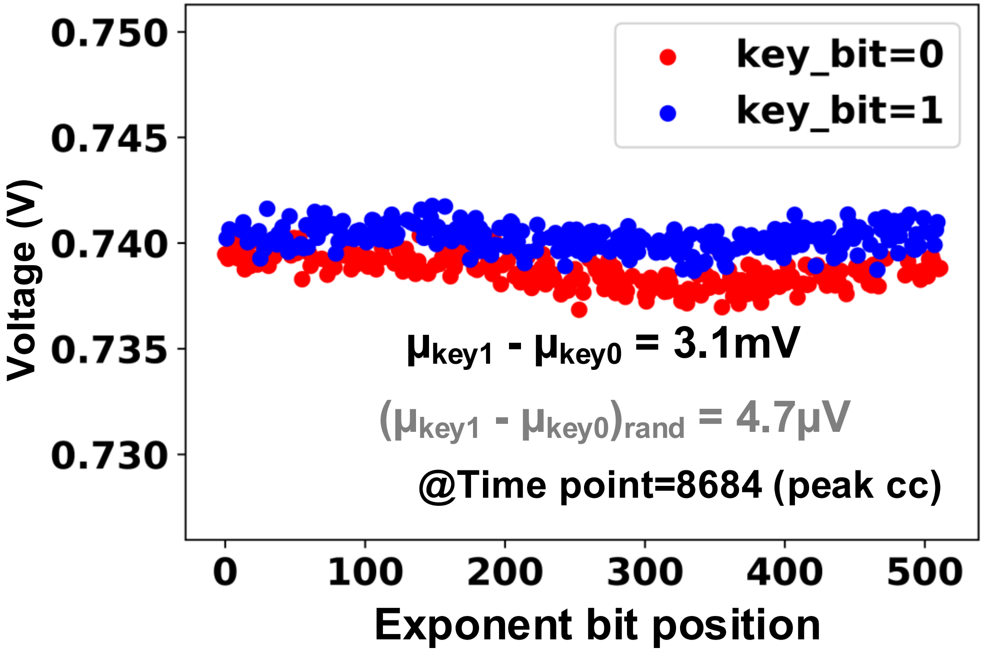
User Instructions (512/1024/2048/4096b modes)

a^b mod c	MontMult(a,b)			
a mod c	a×b	a-b	a+b	reset
read,write,copy,clr,fill RF				

opcode	src1	src2	dest
$opcode_0$	$src1_0$	$src2_0$	$dest_0$
$opcode_1$	$src1_1$	$src2_1$	$dest_1$
$opcode_{60}$	$src1_{60}$	$src2_{60}$	$dest_{60}$

single-cycle ALU µops	
MAC	(op1×op2)+op3+aluout
MULT	(op1×op2)+op3
ADD	op1+op3
SUB	op1-op3
AND	op1&op3
SQR	op1×op1
SHR	op3 >> 1
SHL	op3 << 1





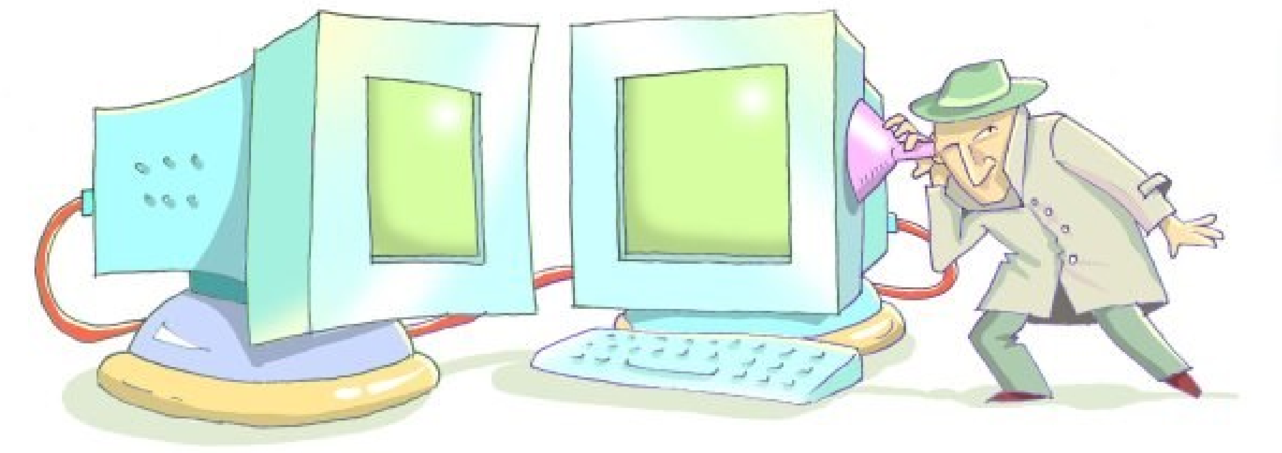
$\mu_{key1} - \mu_{key0} = 3.1mV$

$(\mu_{key1} - \mu_{key0})_{rand} = 4.7\mu V$

@Time point=8684 (peak cc)

- Square-multiply signature clearly visible in power traces
- RSA-512 keys extracted within 40 traces

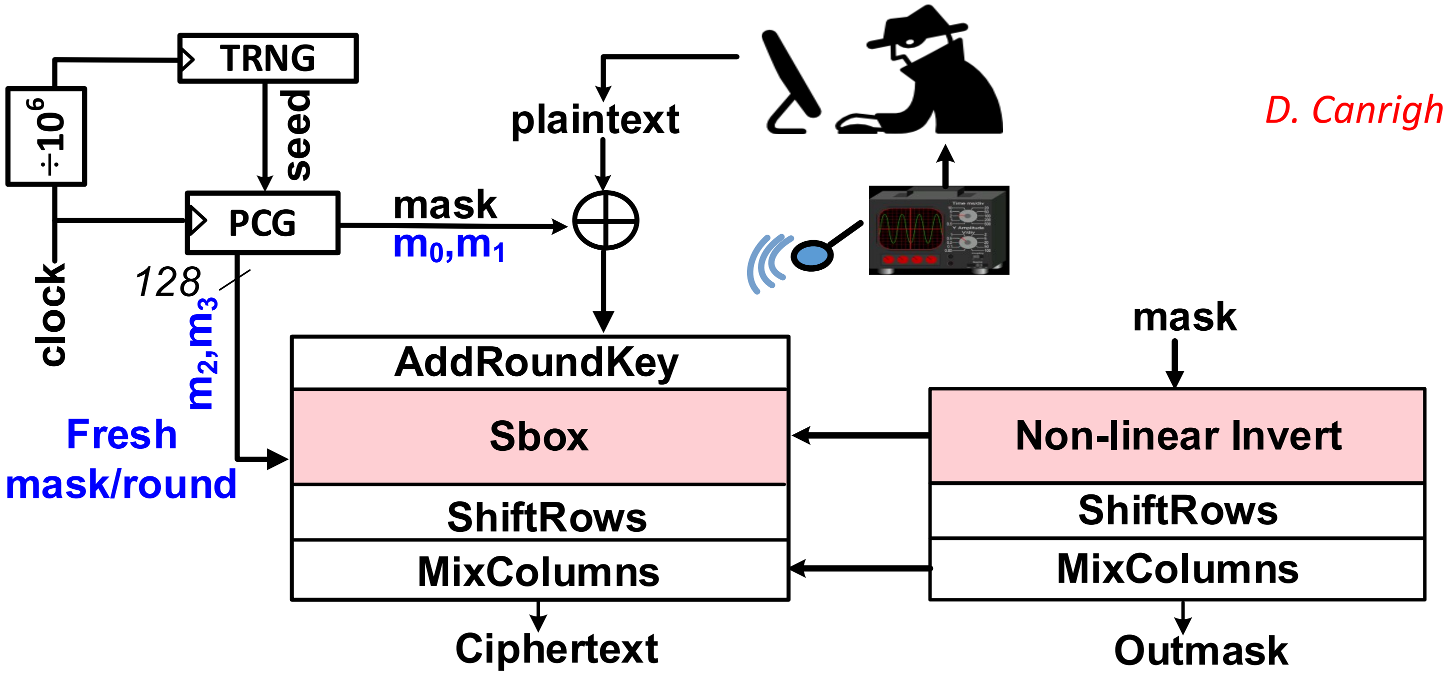
Attack-Resistant Crypto Circuits



- Physical side-channel attacks are a clear and present danger
- SCA undermines security value proposition of crypto HW

Credible countermeasures are a design imperative

Random Additive Masking

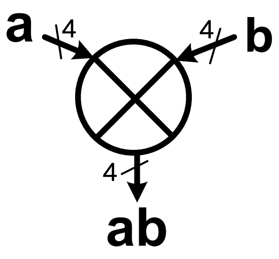


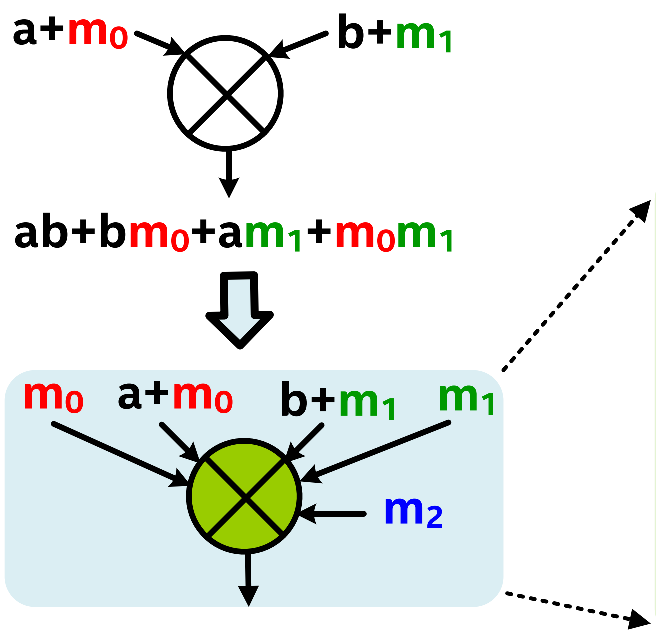
D. Canright and L. Batina, ACNS2008

- Random mask added to plaintext before key addition operation
- Mask inversion factors computed in parallel to compensate final ciphertext
- New masks/round are generated by permuted congruential generator (PCG)
  - PCG reseeded for every 1M cycles from a TRNG to prevent attacks on PRNG

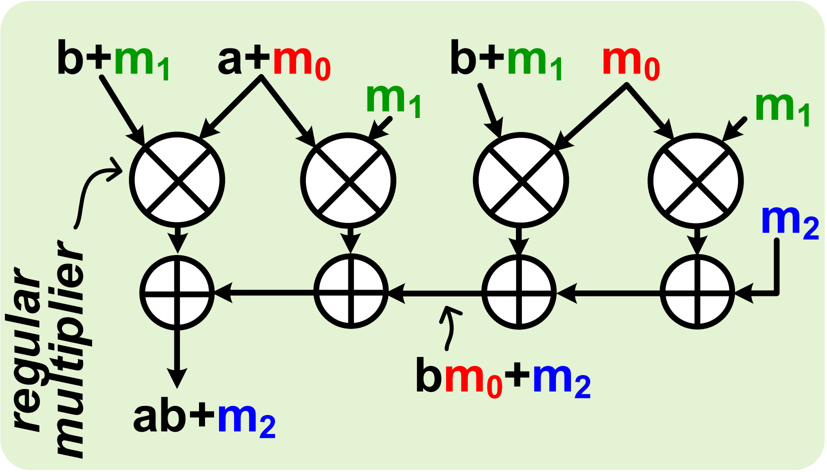
Masked GF(2^4) Multiplier

Regular GF(2^4) multiplier



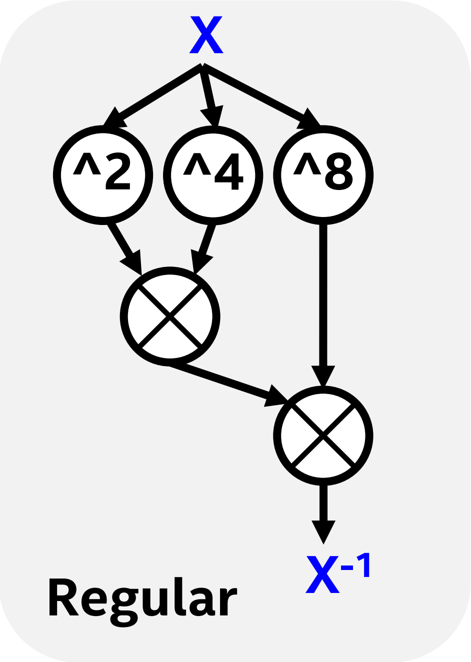


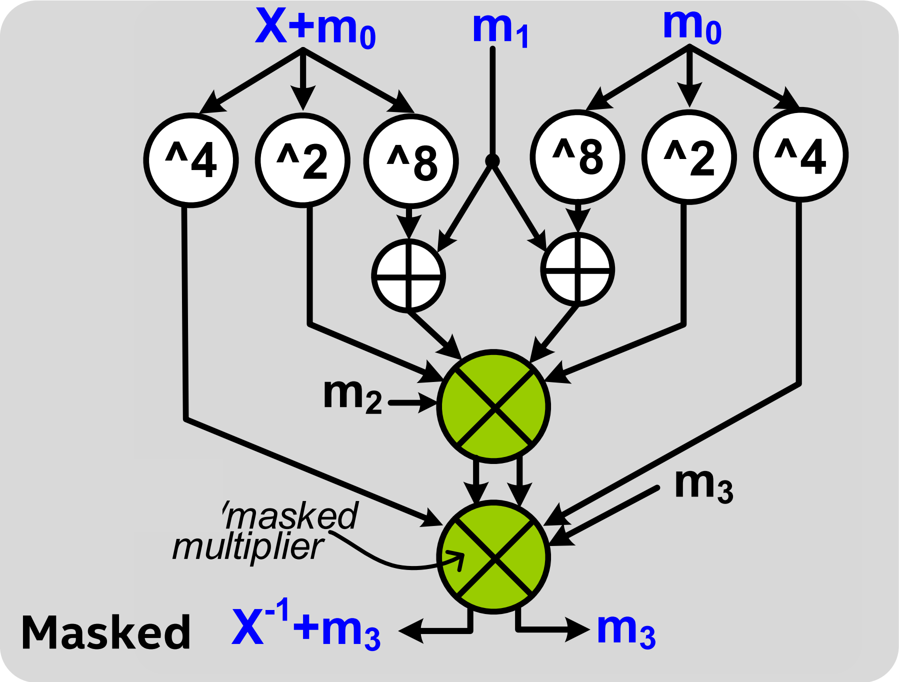
Masked GF(2^4) Multiplier



- Masked multiplier computes the product of masked GF(2^4) operands
  - 3 additional regular GF(2^4) multipliers compute masking byproducts in parallel
- Area overhead: 4x
- Delay overhead: 1.5x

Masked GF(2^4) Inverse

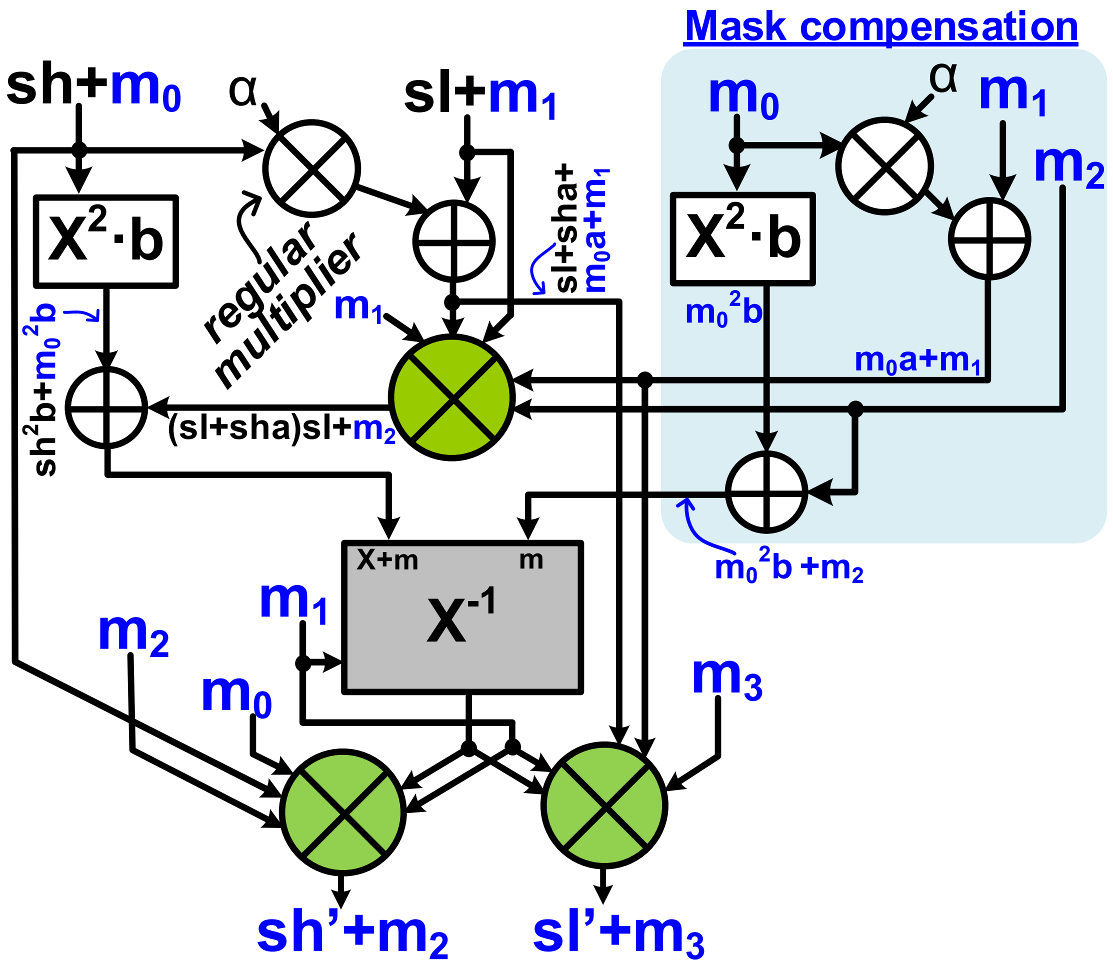




- Fermat's little theorem ($X^{-1}=X^{14}$) used to compute masked GF(2^4) inverse
- Parallel datapath computes the mask compensation factor (m^{14})
- Area overhead: 2x
- Delay overhead: 2x

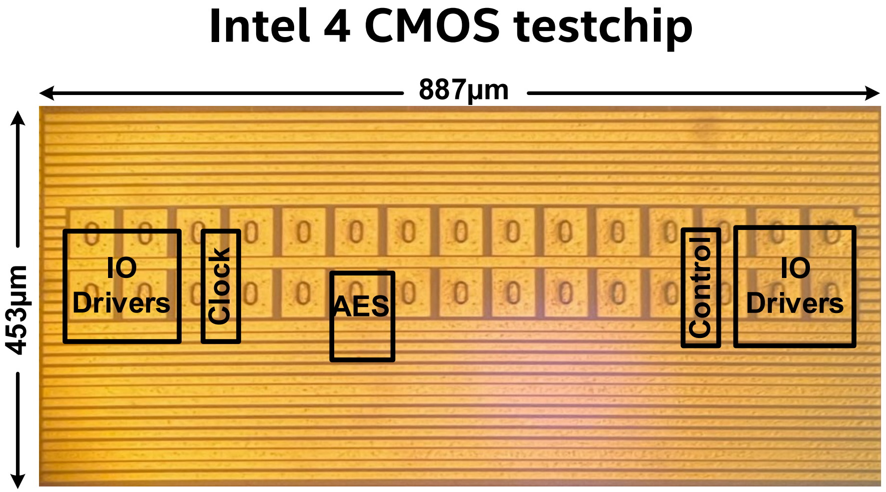
Masked AES Sbox

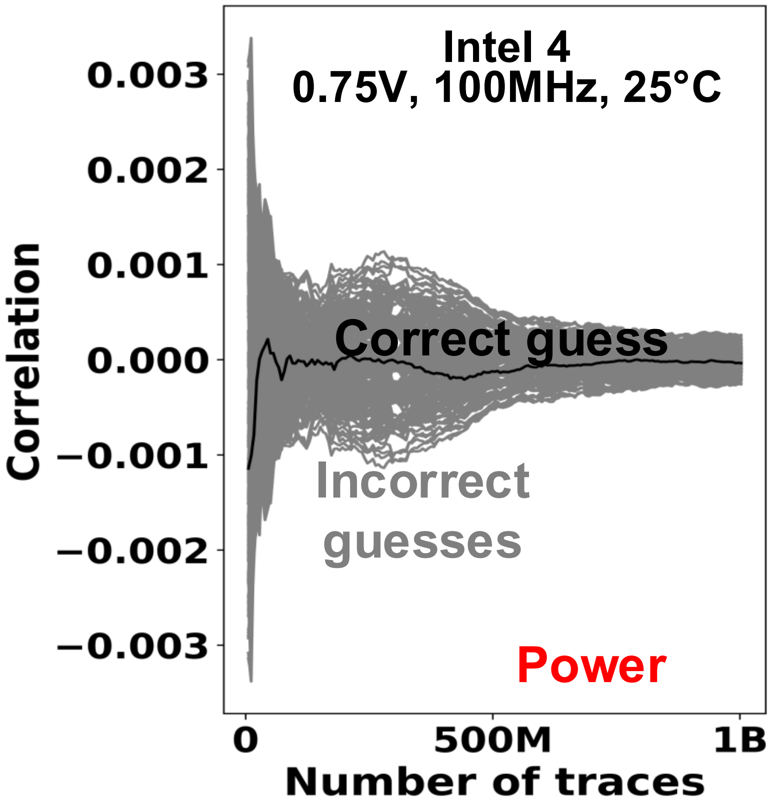
- Masked Sbox datapath implements Sbox(X+mask$_{in}$) = Sbox(X) + mask$_{out}$
- Regular multipliers and inverse blocks are replaced by masked counterparts
- All internal circuit nodes contain a mask component m_i

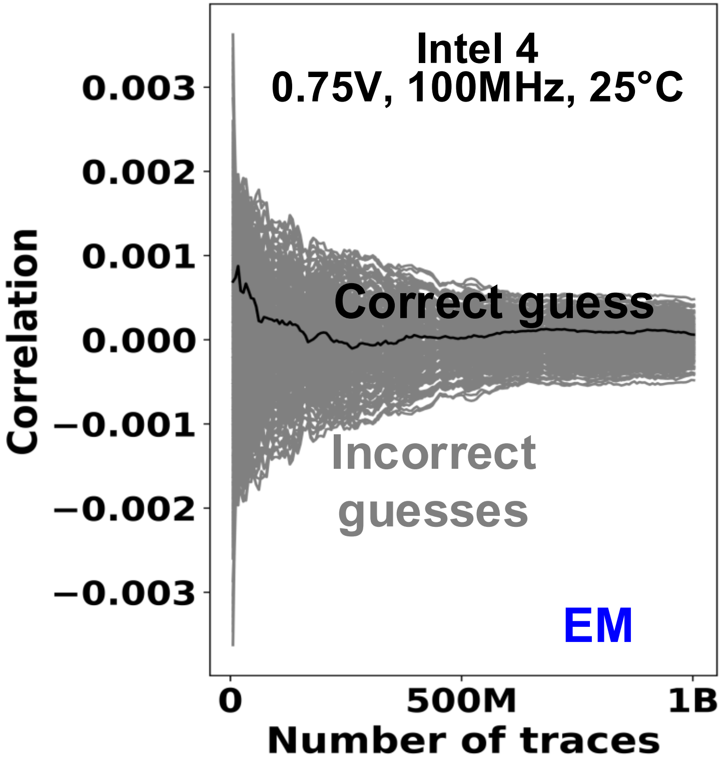


Internal switching activity opaque to external observer

SCA Attacks on Masked AES

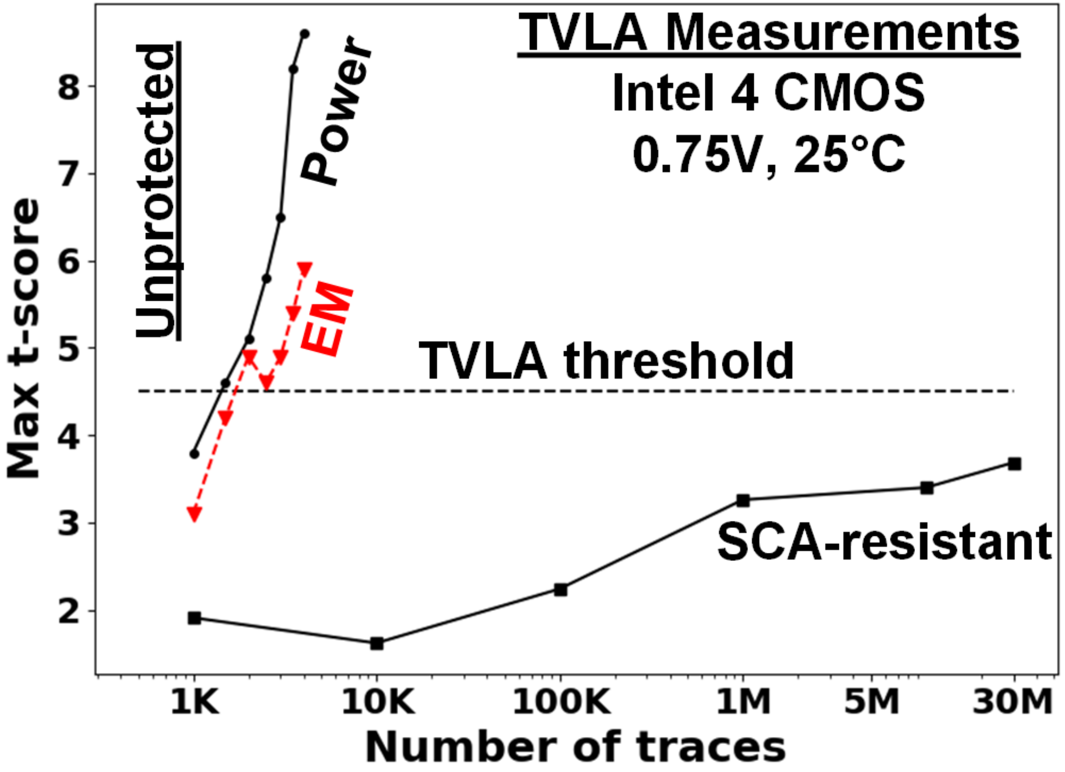






- Correlation power and EM attacks mounted with SCA-resistant mode enabled
- CPA/CEMA attacks unsuccessful after 1B encryptions (>40,000× MTD increase)

SCA Attacks on Masked AES



- Model independent leakage analyzed using TVLA fixed vs. random vectors
- |t|-score < 4.5 after 30M encryptions (>27,000× MTD increase)
- Area overhead: 2x-6x
- Delay overhead: 1.5x

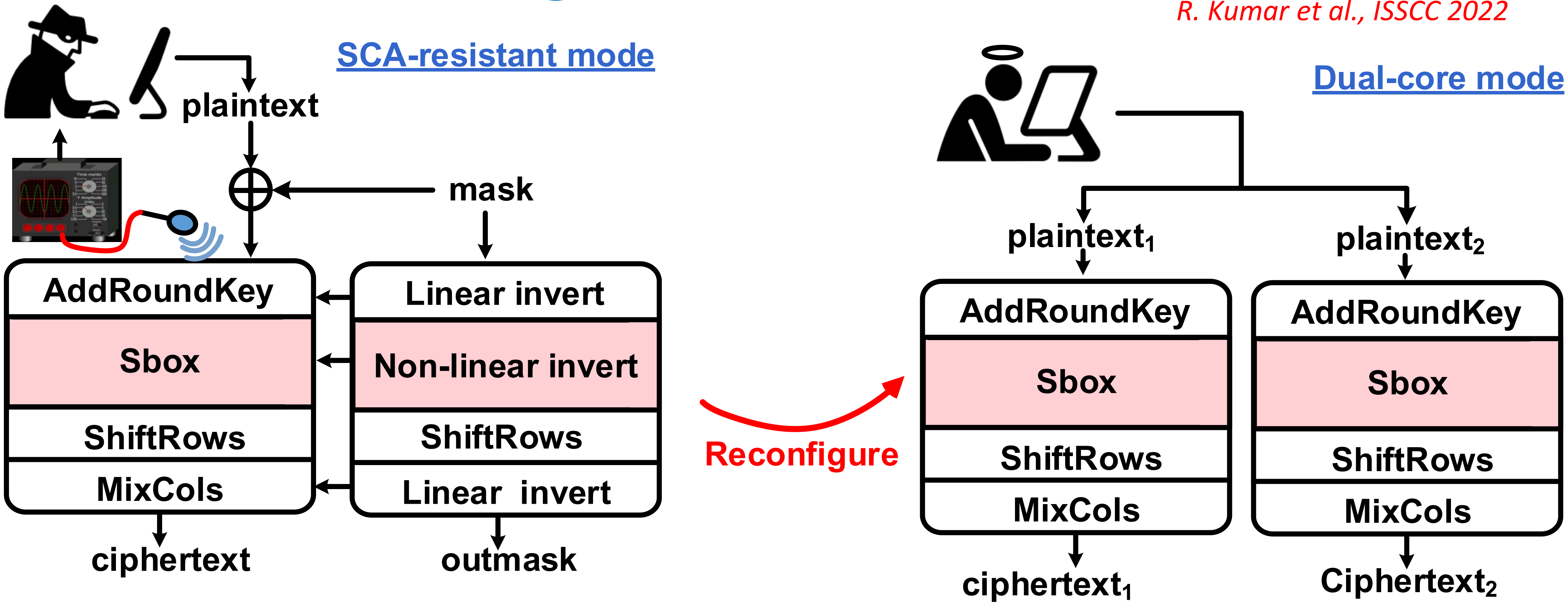
Provably secure, but comes at high cost!

SCA-resistance when/if you need it

Reconfigurable AES

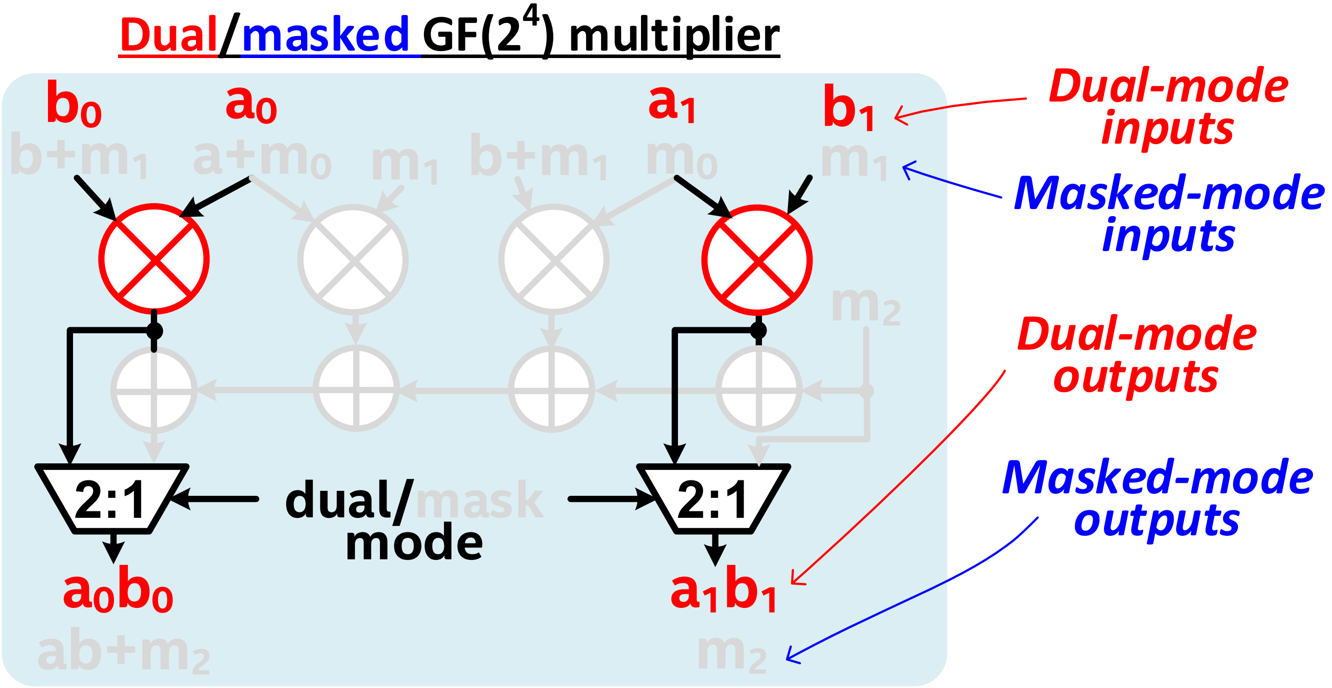
Reconfigurable Dual-core AES

R. Kumar et al., ISSCC 2022



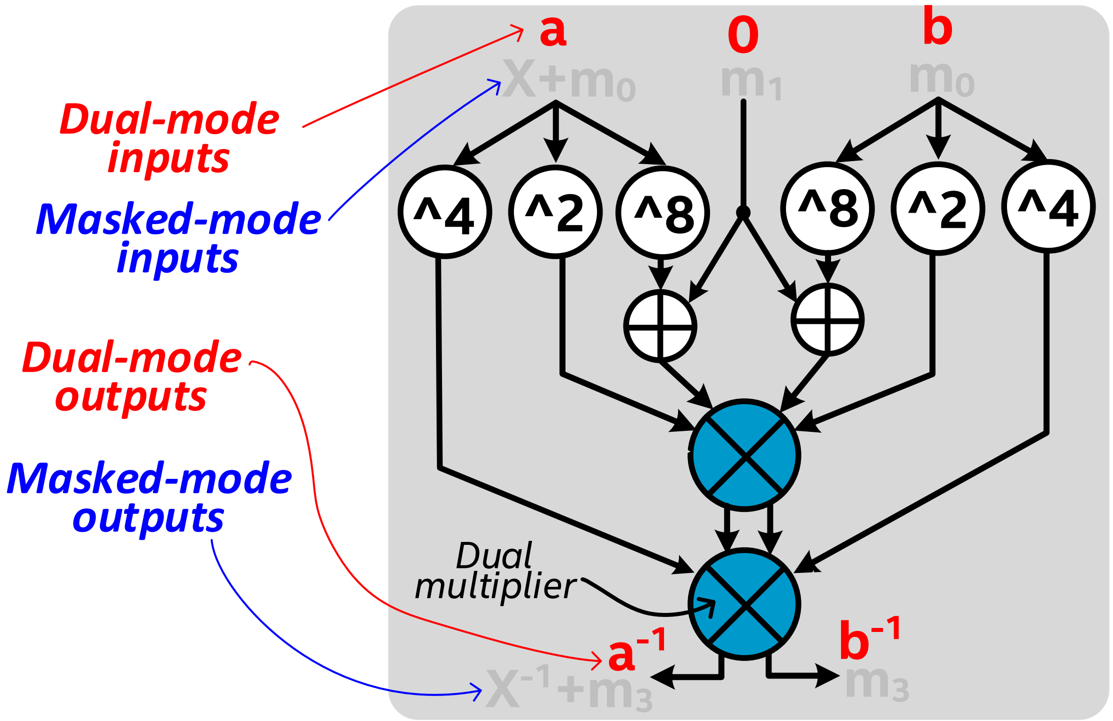
- Mask compensation datapath unused when SCA-resistance is not required
- Dual-core mode repurposes mask datapath to enable 2×higher throughput
  - Enables user to improve performance when operating in trusted environments

Dual-Core GF(2^4) Multiplier



- Masked GF(2^4) multiplier reconfigured to accept second pair of operands
- <1% area overhead incurred by reconfiguration multiplexers

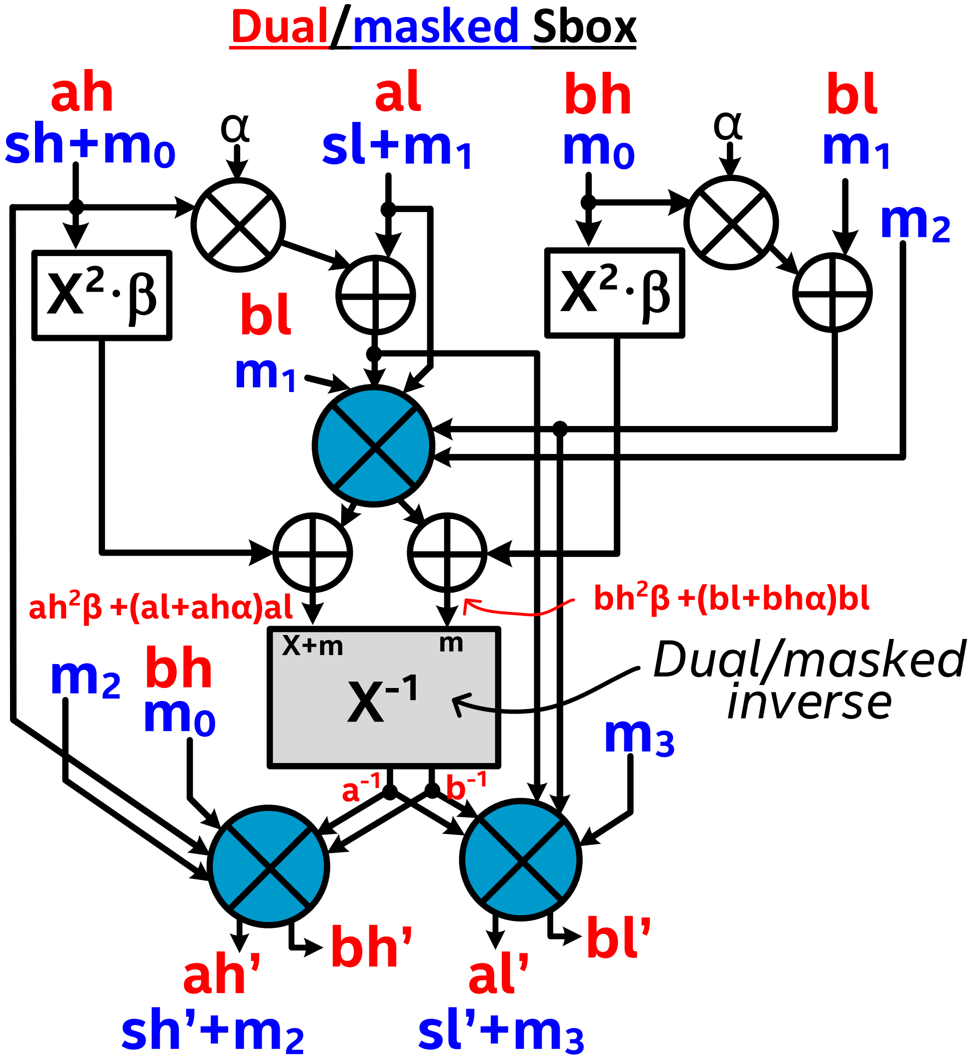
Dual-Core GF(2^4) Inverse



- Masking datapath repurposed to accept the second operand (*b*) through mask pin
- <2% area overhead incurred by reconfiguration logic

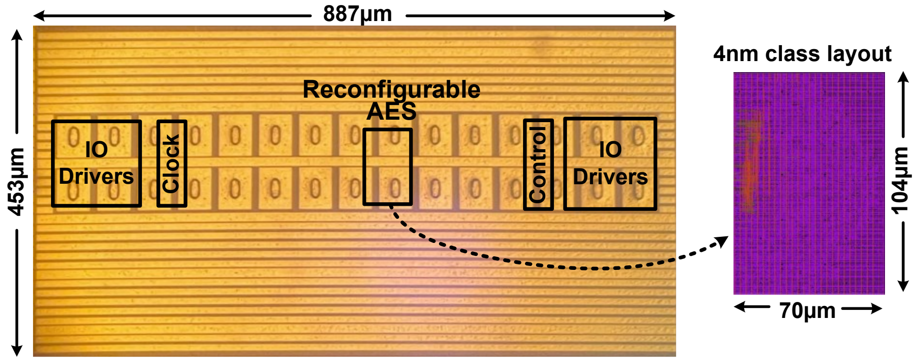
Dual-Core AES Sbox

- Mask compensate logic accepts second pair of operands (*bh*, *bl*) in dual-core mode
- Reconfiguration adds 4 multiplexers to the round critical path
  - 6% area overhead



AES Throughput Measurements

Reconfigurable AES fabricated in Intel 4

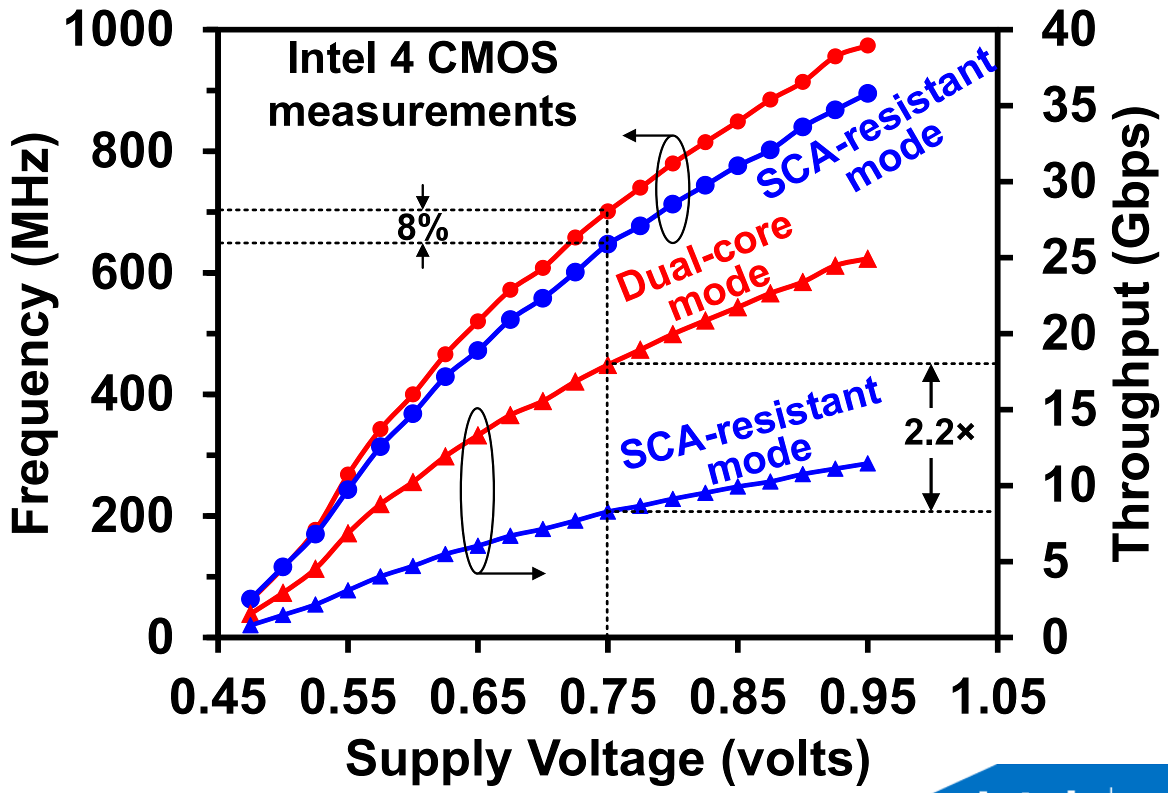


SCA-resistant mode:

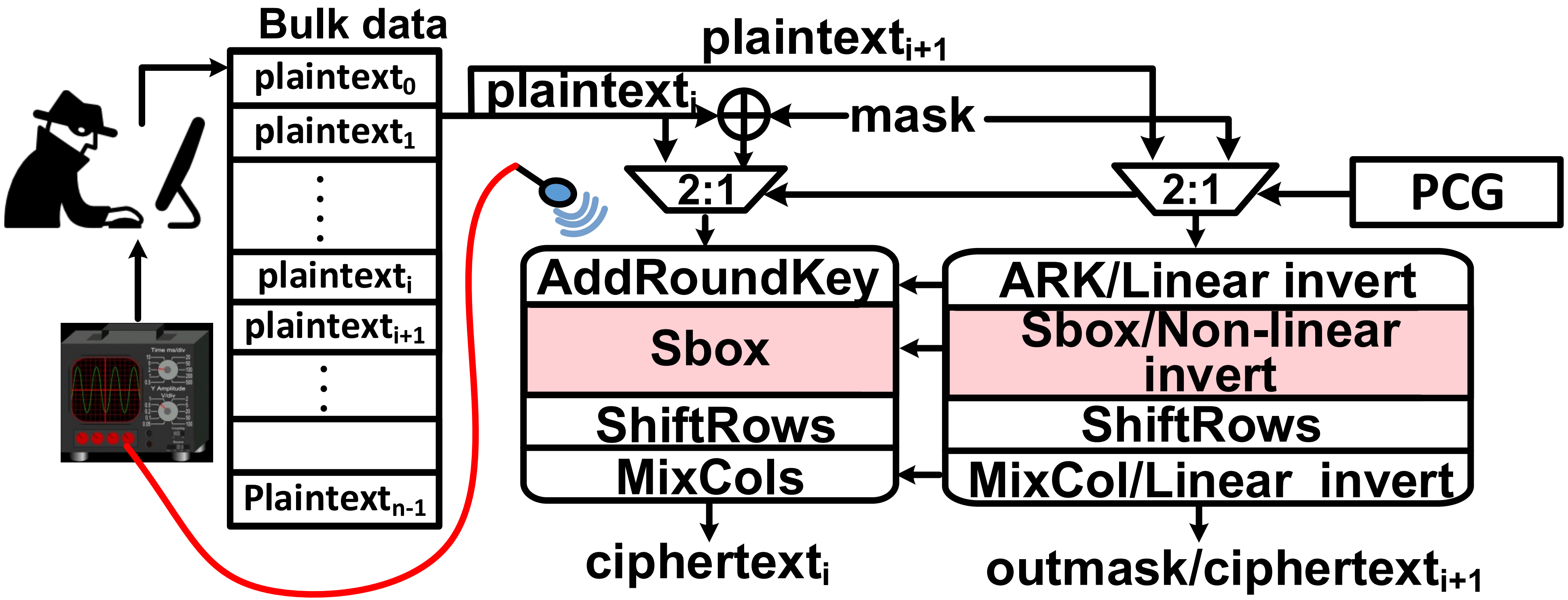
- Fmax of 647MHz
- AES throughput of 8.3Gbps

Dual-core mode:

- Fmax of 701MHz (8%↑)
- AES-128 throughput of 18Gbps (2.2x↑)

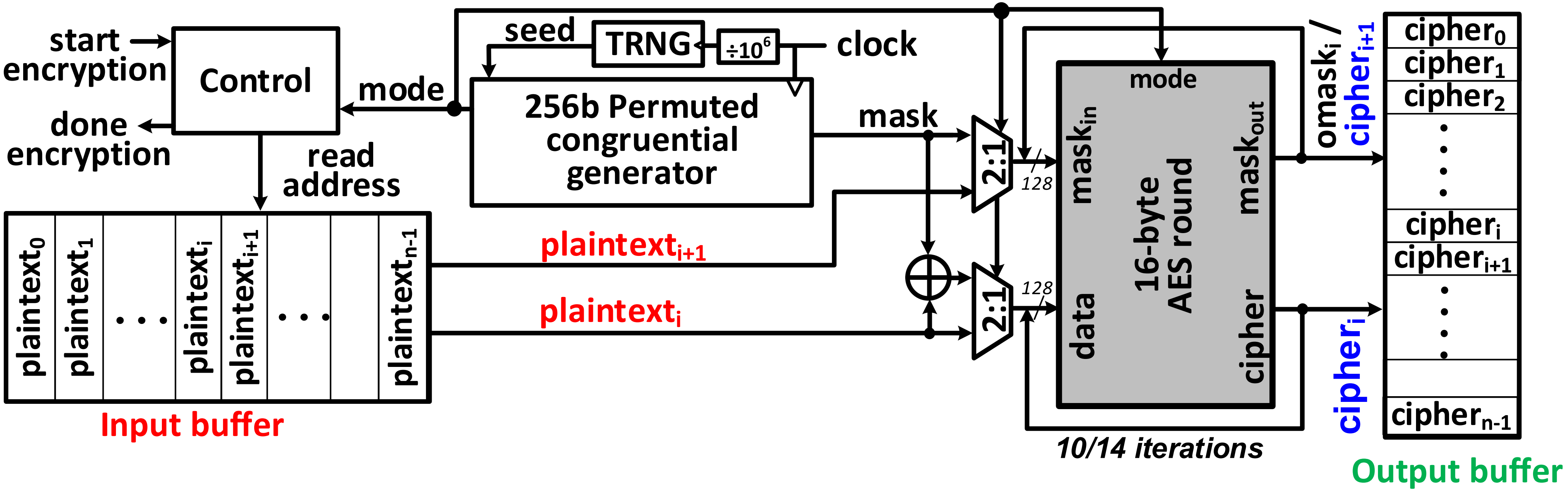


Blind-Bulk Mode



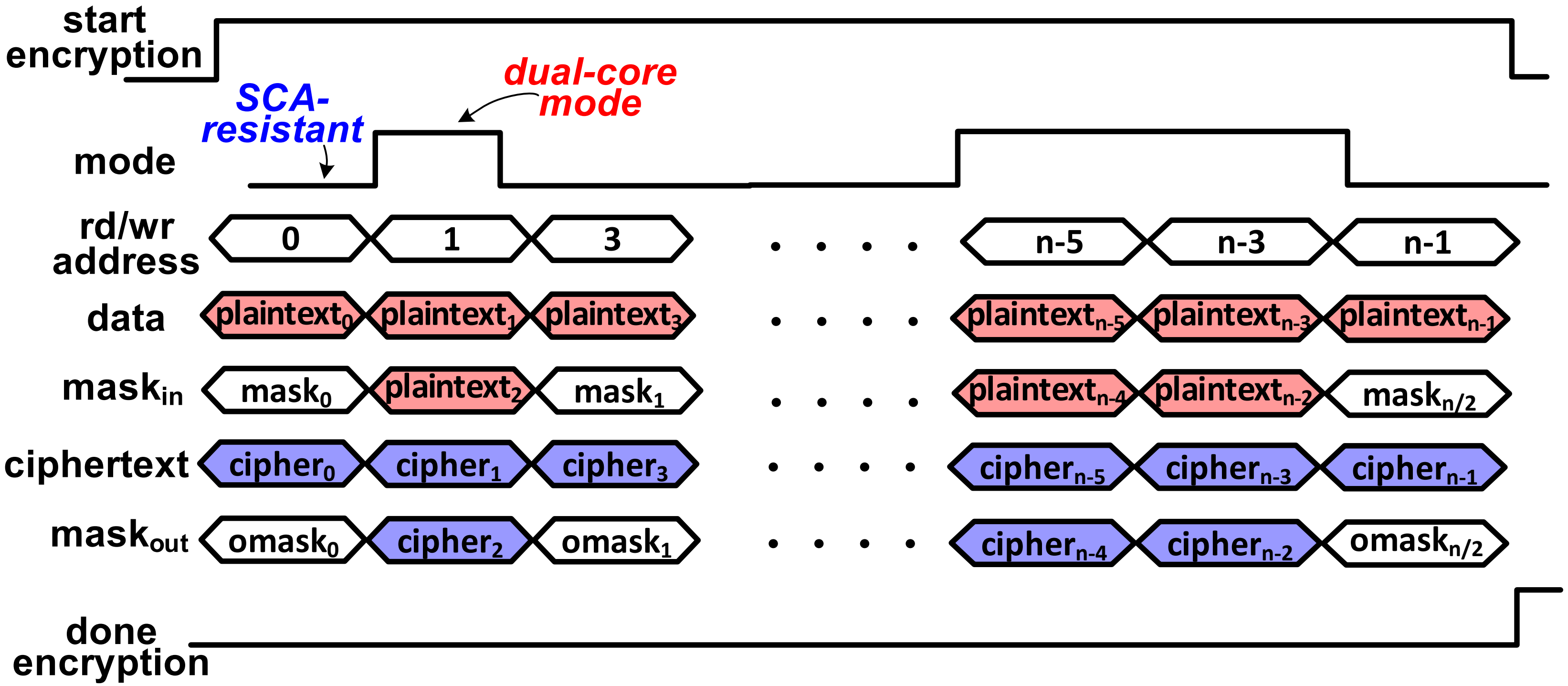
- Bulk data encryption is critical for encrypting memory, hard drives, SSDs, etc.
  - Proposed AES operates in blind-bulk mode to encrypt bulk data
- Random switching between SCA-resistant and dual-core modes
- Enables throughput vs security trade-offs by adjusting #SCA-resistant encryptions

Blind-Bulk Mode Datapath



- Control block randomly switches between SCA/dual-core modes
  - Plaintext loaded from input buffer is added with random mask in SCA mode
  - Pair of plaintexts are fetched from buffer in dual-core mode
- 256b PCG controls the ratio (*p*) of SCA-resistant mode

Blind-Bulk Mode Operation



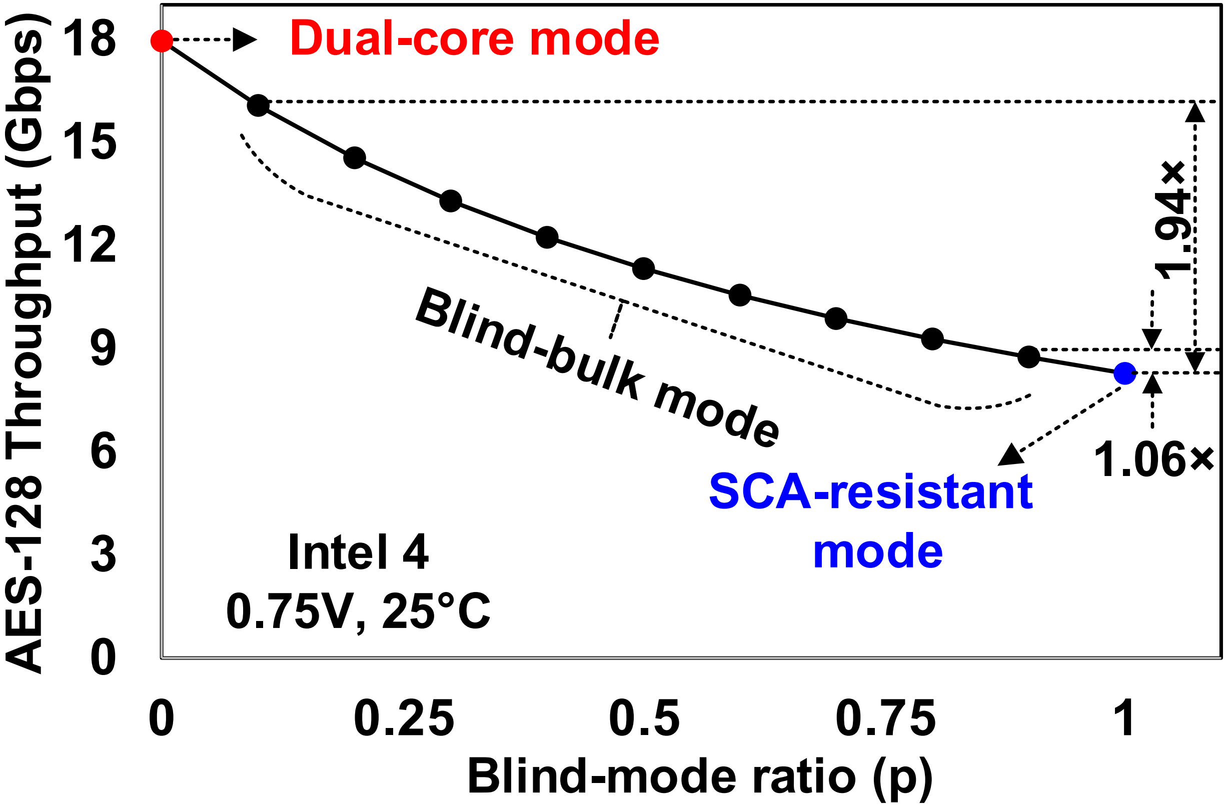
$$p = \frac{\textit{SCA-resistant encryptions}}{\textit{Total encryptions (n)}}$$

p	Throughput	SCA Resistance
1	1×	High
0.75	1.14×	Medium
0.5	1.33×	
0.25	1.6×	
0	2×	None

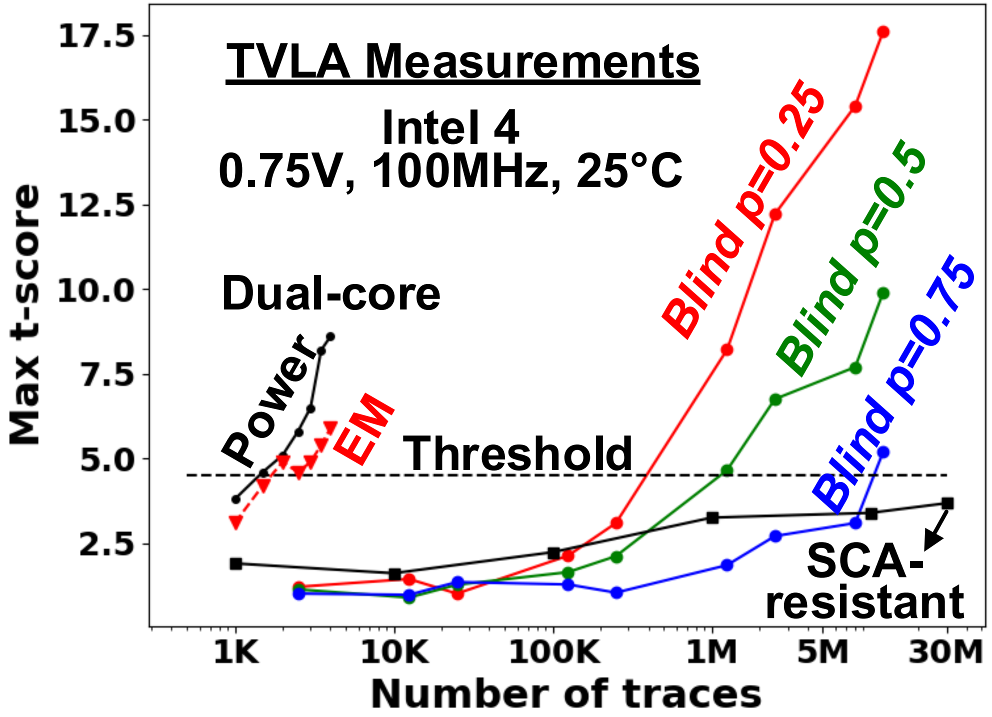
- Blind-bulk mode is time-invariant to prevent timing attacks
  - Total latency for *n* AES-128/256 encryptions: $5/7*n*(1+p)$
- Blind mode ratio (*p*) enables throughput vs side-channel resistance trade-offs

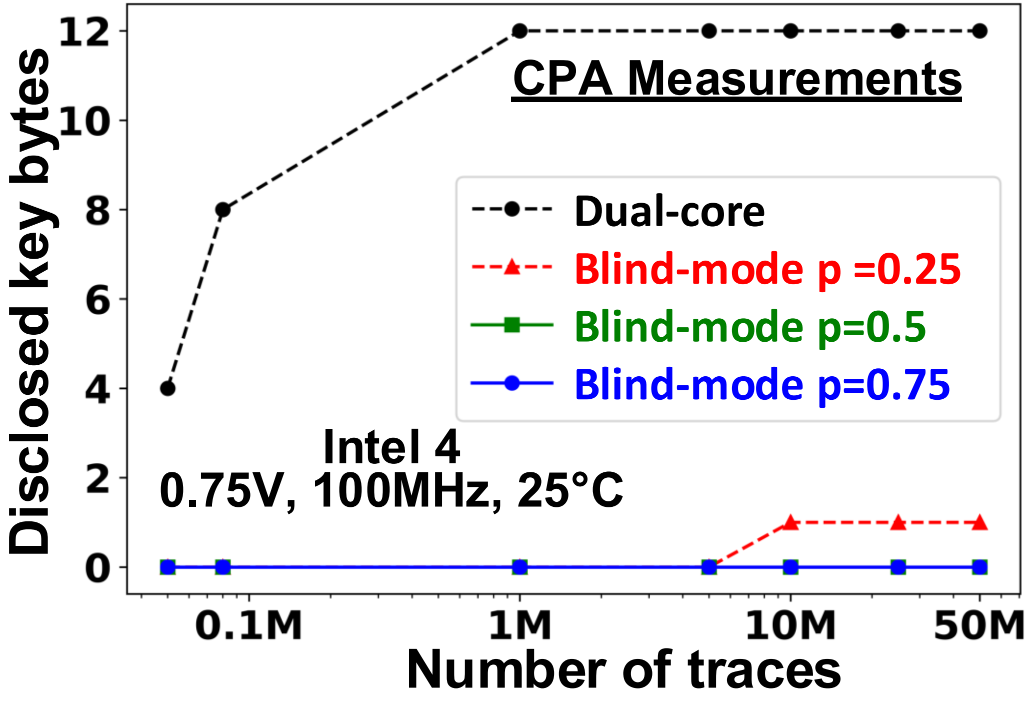
Blind-Bulk AES Throughput

- SCA-resistant mode ratio (p) swept between 0.1-0.9 for AES throughput
- 1.06-1.94× improvement in encryption throughput
- Enables throughput vs SCA-resistance trade-offs



Blind-Bulk Mode SCA Analysis





- SCA-resistant mode ratio (p) shows linear increase in TVLA MTD
  - **400/1200/12000×** improvement in MTD for p=0.25/0.5/0.75
- 400× improvement in CPA/CEMA MTD for p=0.25
- No key bytes revealed for p=0.5/0.75 after 50M traces (>2000× increase)

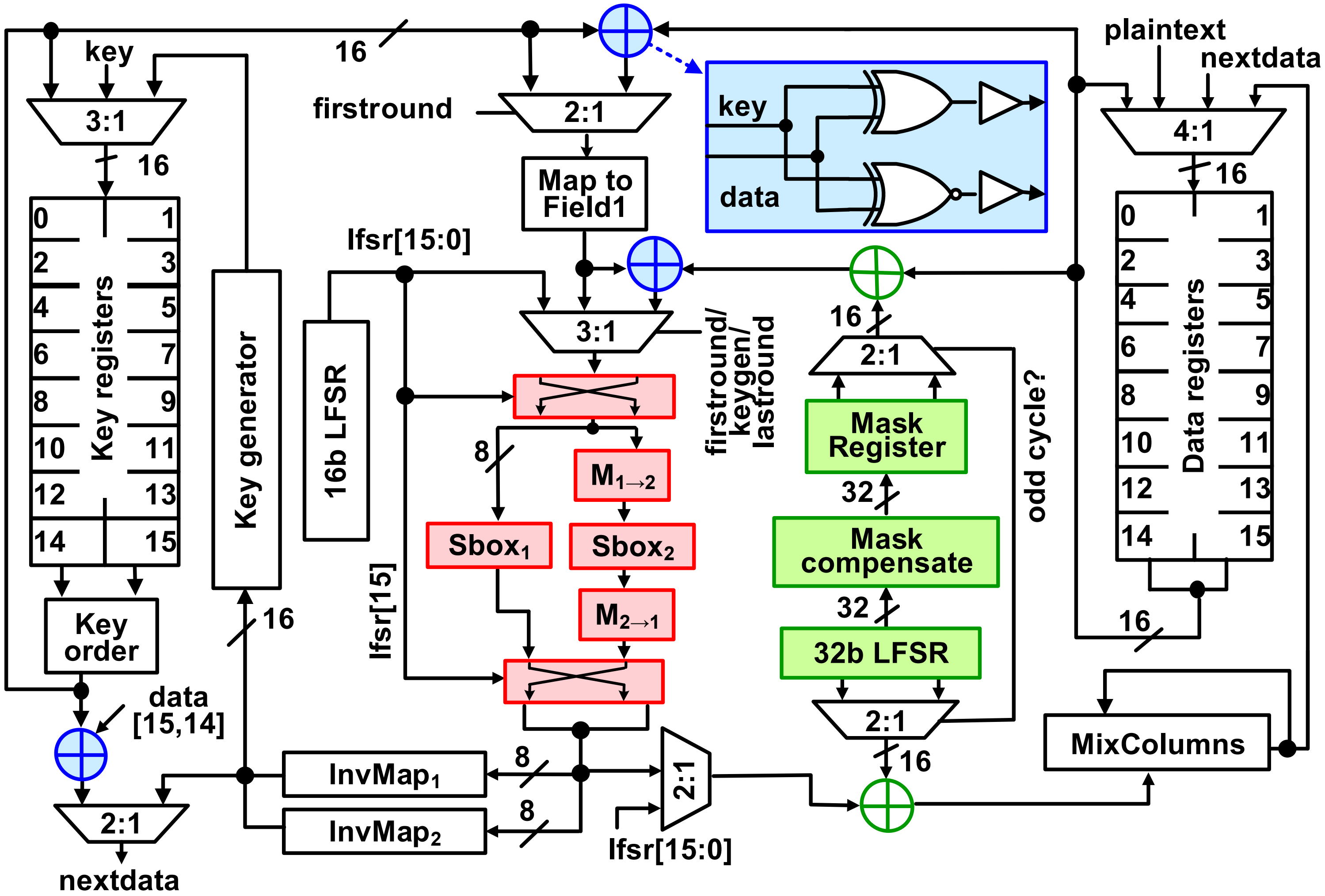
Lightweight SCA-resistant AES

Heterogenous Sbox AES

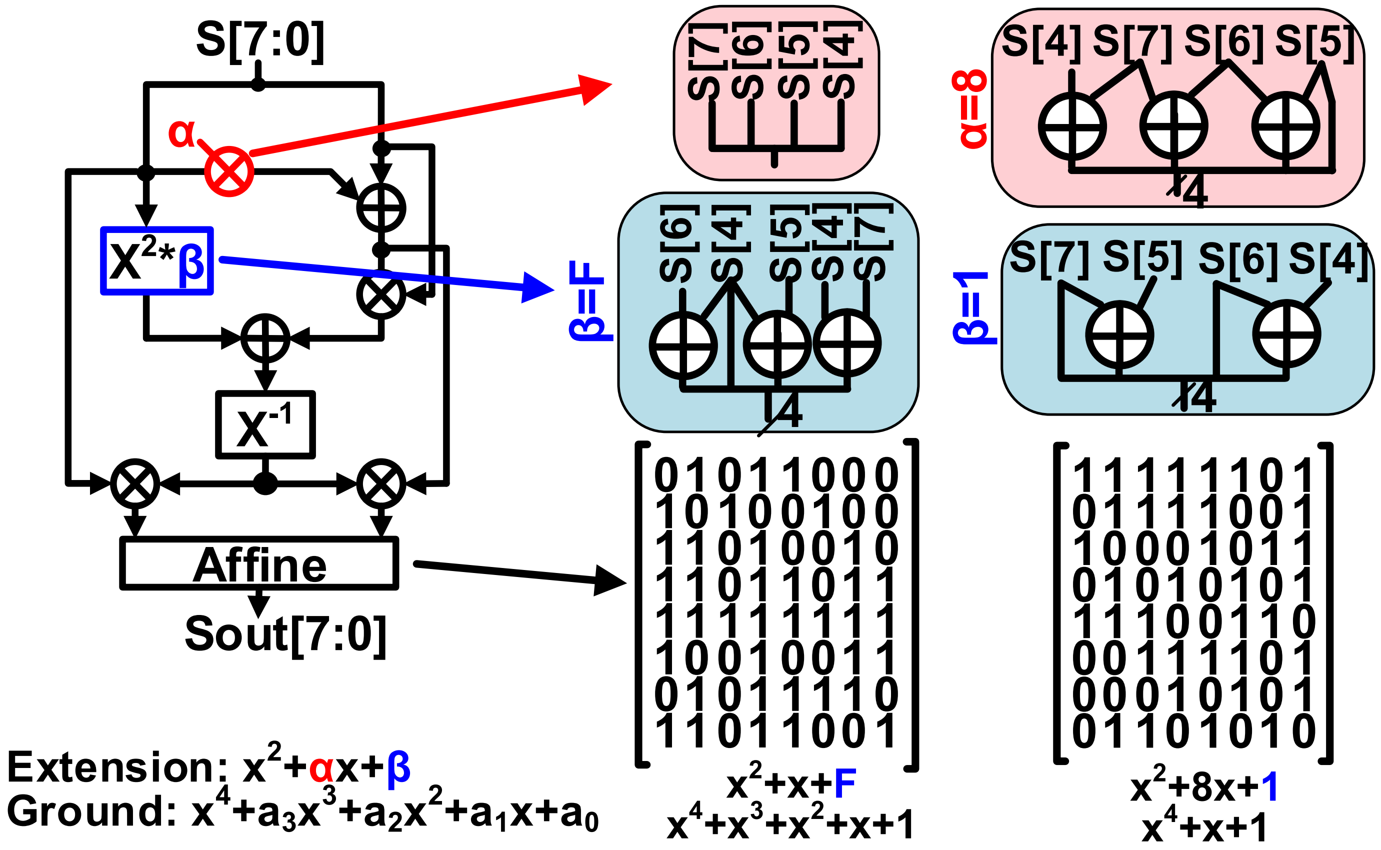
SCA-resistant Heterogenous-Sbox AES

R. Kumar et al., VLSI 2019

- 16b serial datapath
- Three SCA-resistance features:
  - Dual-rail key addition
  - Heterogeneous Sboxes
  - Masked MixColumns



Composite-field Sbox Datapath



$$\begin{bmatrix} 0&1&0&1&1&0&0&0\\ 1&0&1&0&0&1&0&0\\ 1&1&0&1&0&0&1&0\\ 1&1&0&1&1&0&1&1\\ 1&1&1&1&1&1&1&1\\ 1&0&0&1&0&0&1&1\\ 0&1&0&1&1&1&1&0\\ 1&1&0&1&1&0&0&1 \end{bmatrix}$$

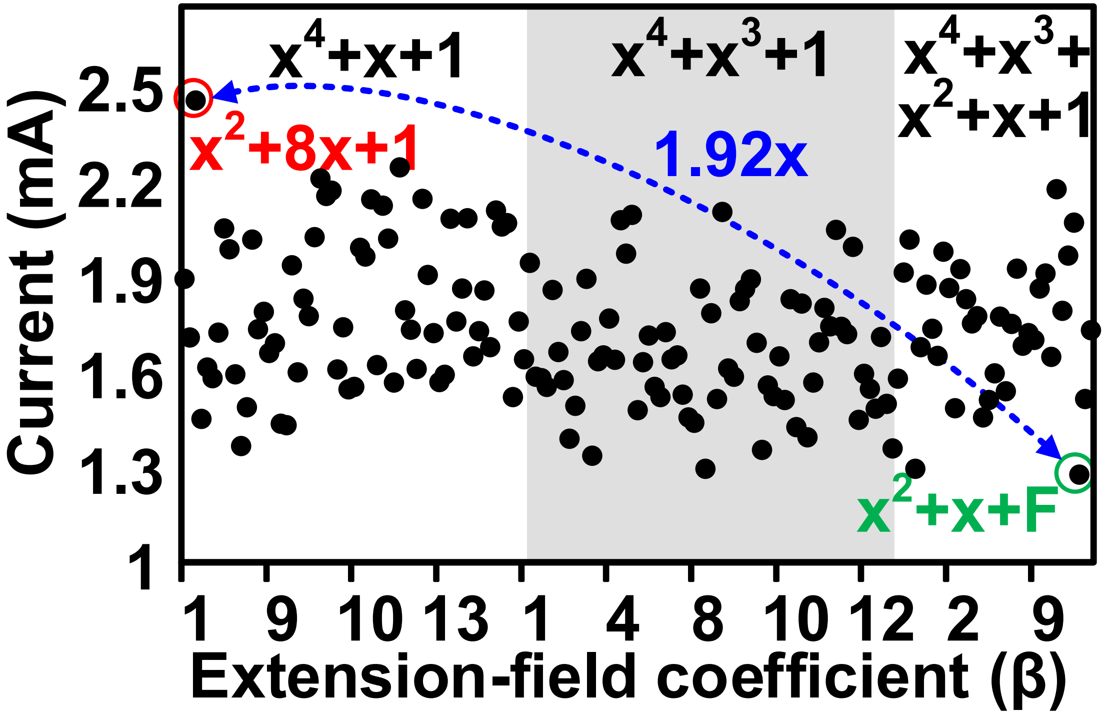
x^2+x+F
$x^4+x^3+x^2+x+1$

$$\begin{bmatrix} 1&1&1&1&1&1&0&1\\ 0&1&1&1&1&0&0&1\\ 1&0&0&0&1&0&1&1\\ 0&1&0&1&0&1&0&1\\ 1&1&1&0&0&1&1&0\\ 0&0&1&1&1&1&0&1\\ 0&0&0&1&0&1&0&1\\ 0&1&1&0&1&0&1&0 \end{bmatrix}$$

x^2+8x+1
x^4+x+1

Extension: $x^2+\alpha x+\beta$
Ground: $x^4+a_3x^3+a_2x^2+a_1x+a_0$

14nm CMOS Simulations
160 area-sorted polynomials

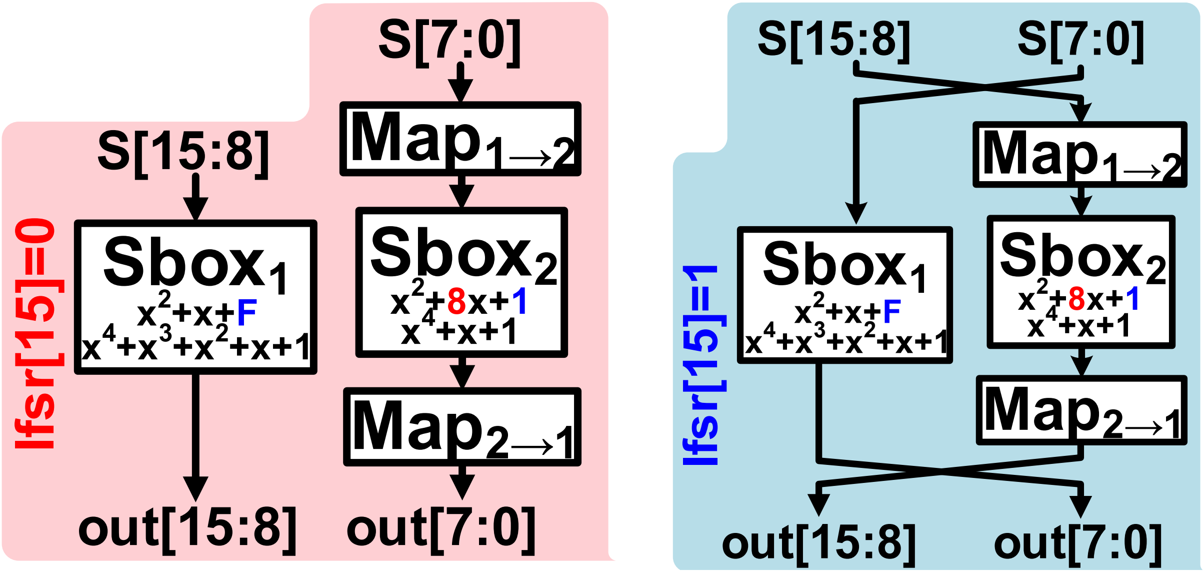


Composite-field GF$(2^4)^2$ polynomials have strong influence on Sbox circuits

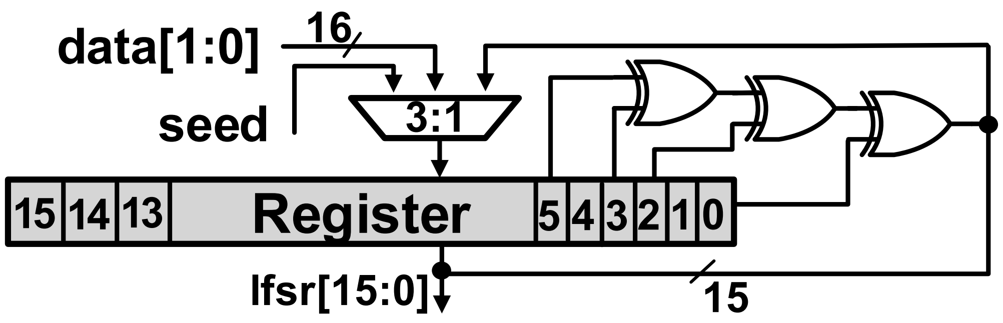
1.92x spread in Sbox power across 160 area-sorted polynomials

Heterogeneous Sbox Dataflow

Randomized dataflow switching



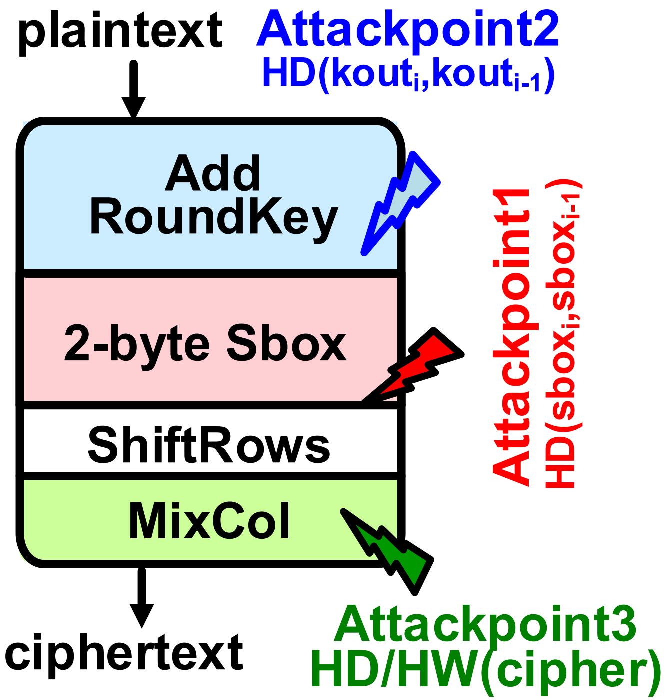
LFSR reseeding datapath



- Disrupts power correlations by random dataflow through heterogenous Sboxes
- 16b LFSR updates dataflow direction every cycle
- LFSR reseeded with intermediate ciphertext to prevent averaging attacks
  - Least significant bytes of data used as reseed
  - LFSR reseeded at end of key expansion

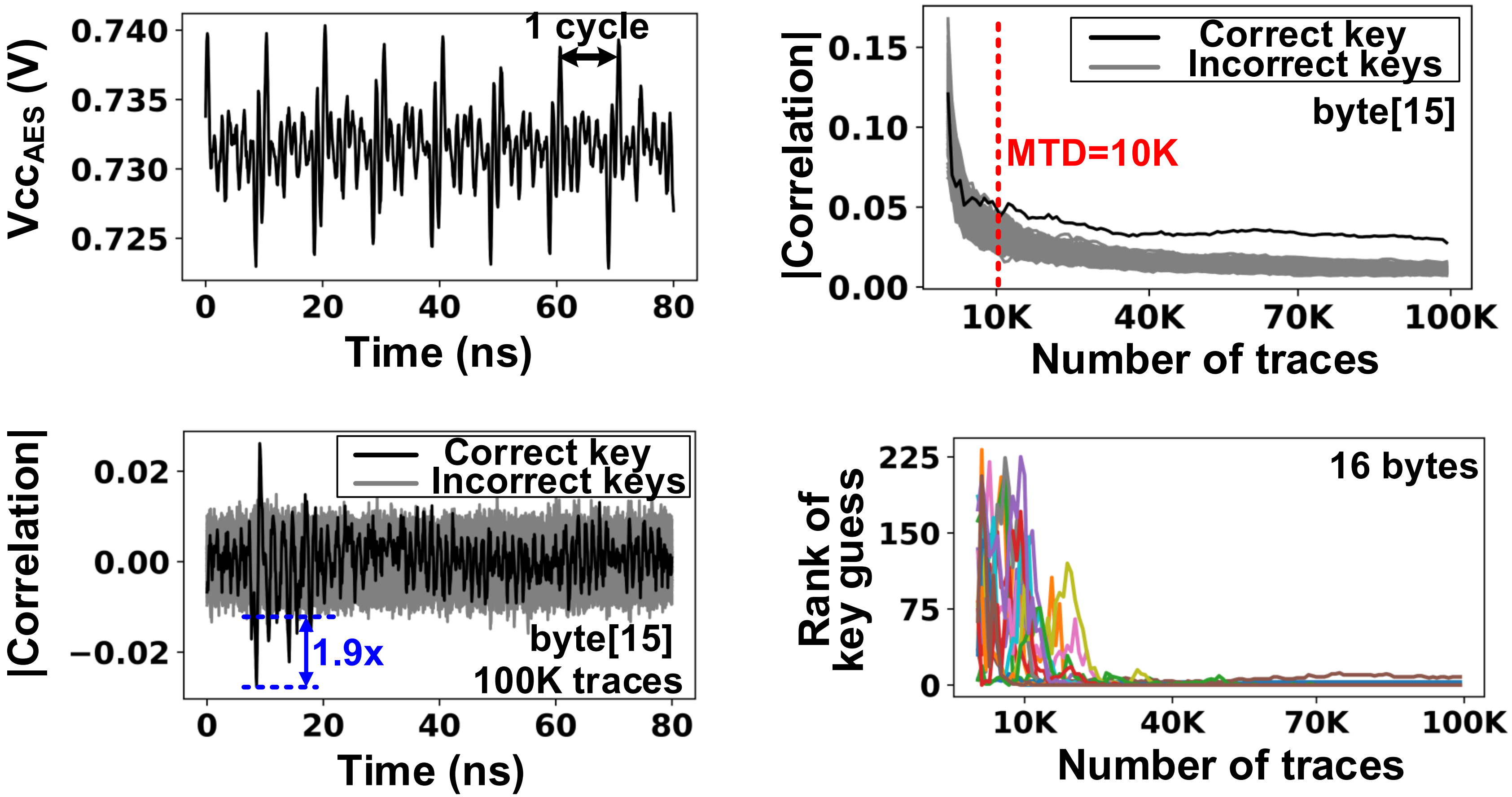
SCA Attacks on Unprotected AES

Attackpoints in unprotected AES



HD - Hamming distance
HW – Hamming weight

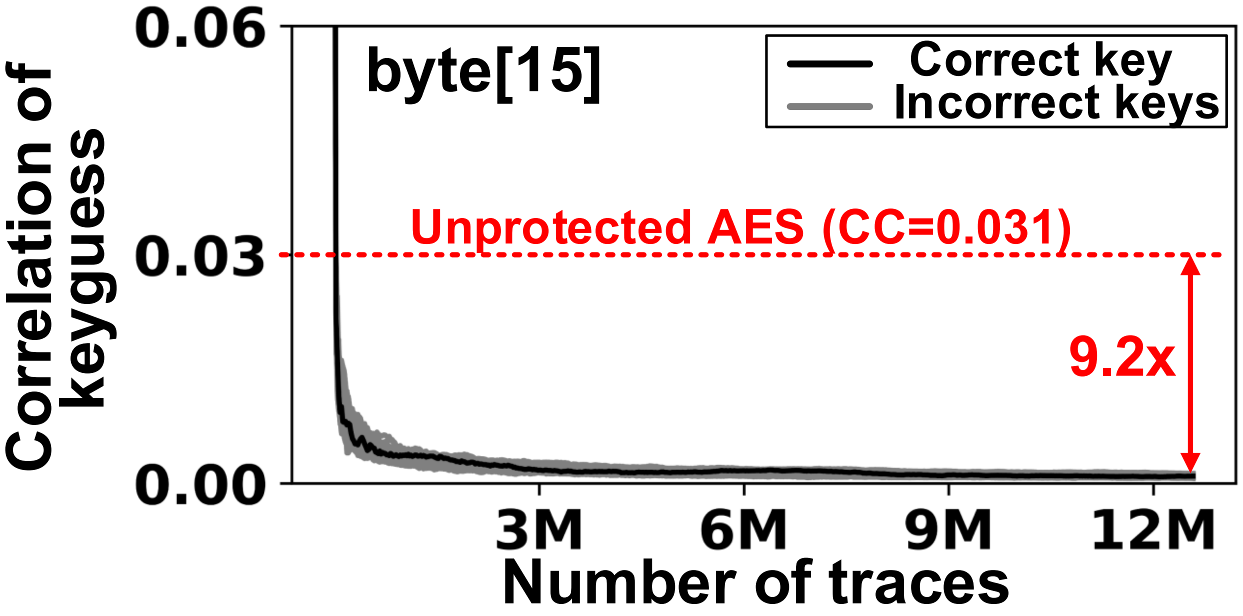
14nm CMOS measurements, 0.75V, 100MHz, 25°C

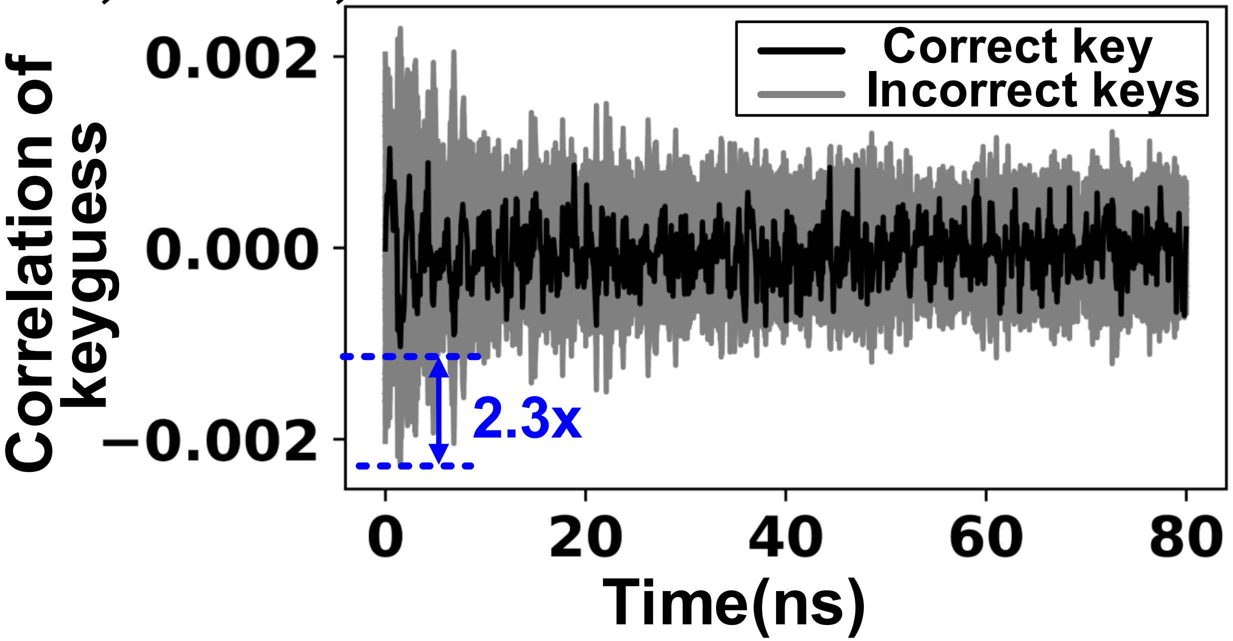


- Unprotected 16b serial, 108cycle latency AES fabricated in 14nm CMOS
- Three attack points cover all attack surfaces within the design
- Successful CPA attack with minimum traces to disclosure (MTD) of 10,000

CPA Attacks on Heterogenous Sbox AES

14nm CMOS Measurements, 0.75V, 100MHz, 25°C





- Power traces collected with all 3 SCA-resistance techniques enabled
- No key bytes extracted after 12Million encryptions **(>1200x** improvement in MTD)
- 9.2x reduction in correlation over unprotected AES with an average rank of 139
  - **1100x** improvement in TVLA assessment

Lightweight SCA-resistant AES

Multiplicatively-masked AES

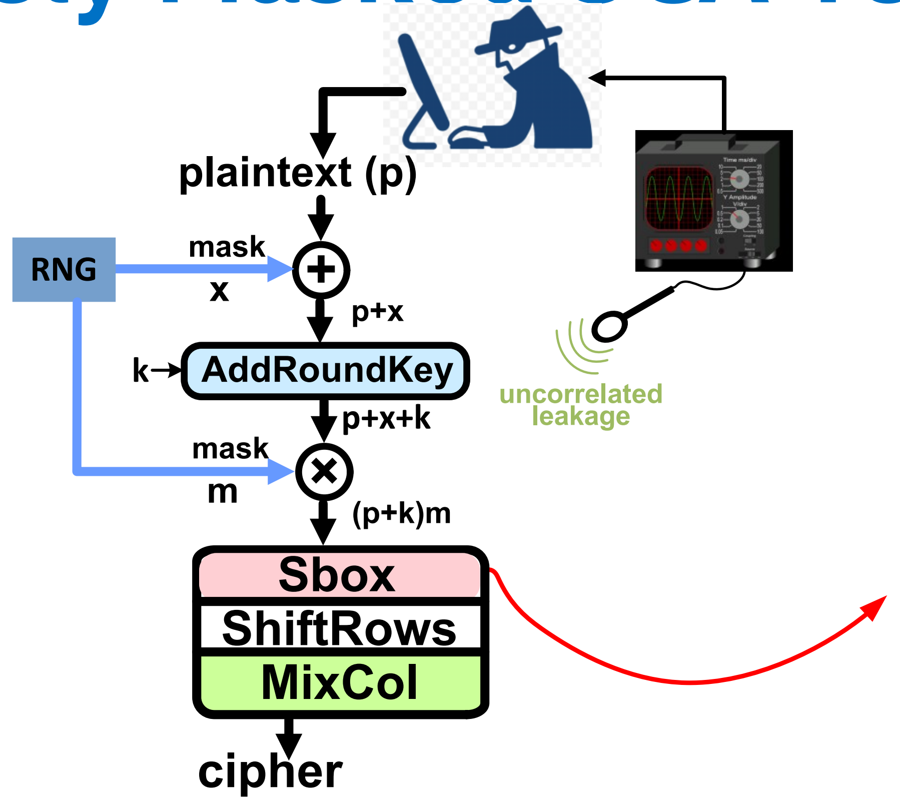
Multiplicatively Masked SCA-resistant AES

Additive masking

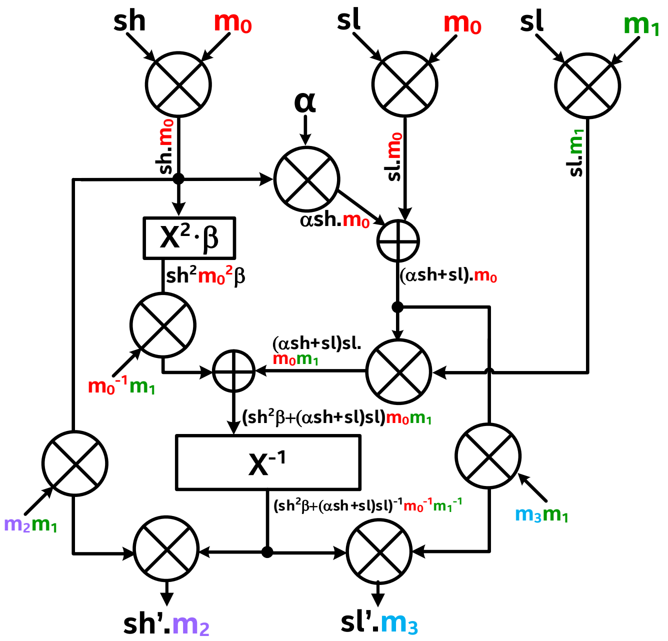
$$(x+m)^{-1} \neq x^{-1} + m^{-1}$$

Multiplicative masking

$$(x \bullet m)^{-1} = x^{-1} \bullet m^{-1}$$



Multiplicatively-masked Sbox

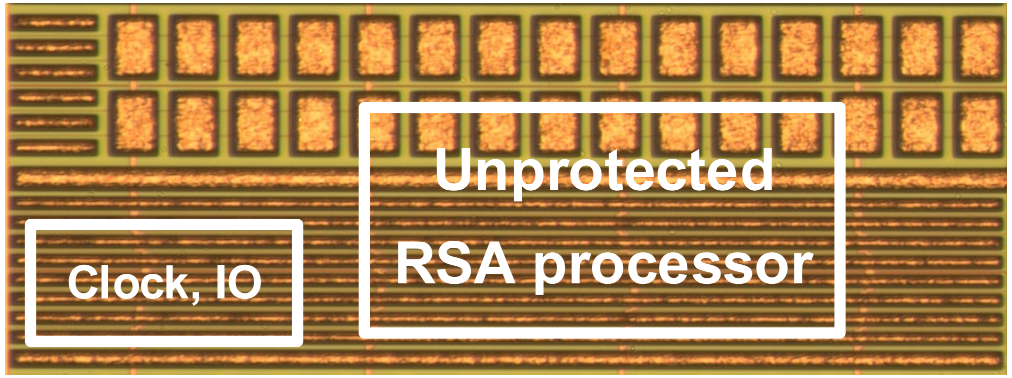


- Multiply a random 128b mask to the data prior to non-linear operation
- Mask compensation is relatively trivial (area overhead 25-50%)
- Vulnerable to zero-value attack (p+k=0)
- Preemptively detect zero value and obscure with random data

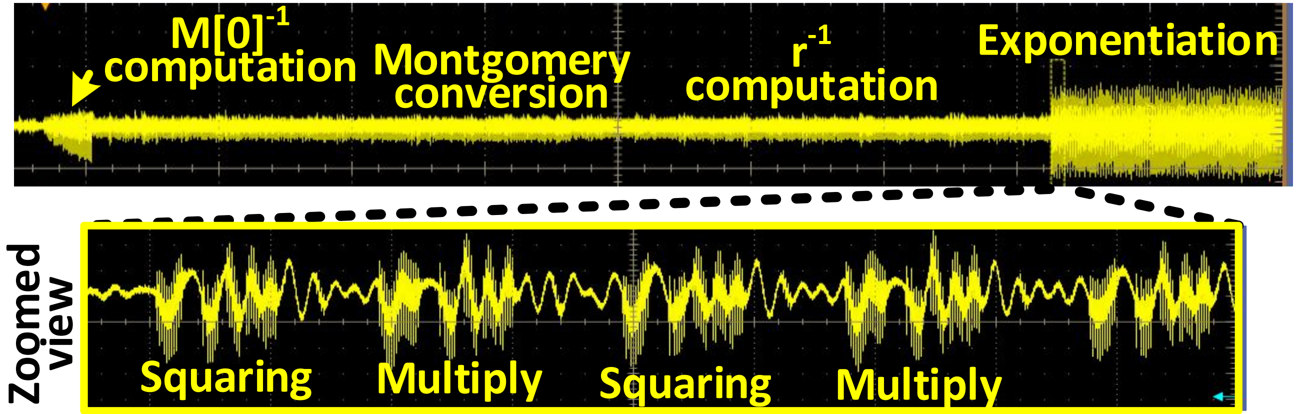
Protecting public-key Crypto

SCA-resistant RSA

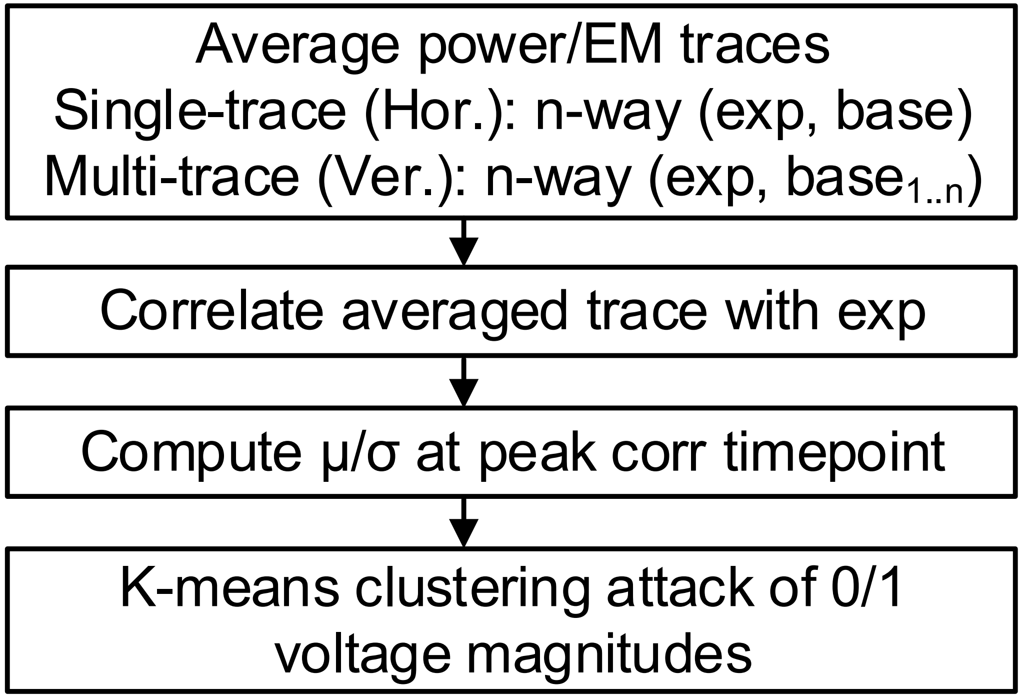
SCA Attacks on Unprotected RSA



Captured current traces



Horizontal and vertical profiling attacks



Average power/EM traces
Single-trace (Hor.): n-way (exp, base)
Multi-trace (Ver.): n-way (exp, $\text{base}_{1..n}$)

Correlate averaged trace with exp

Compute μ/σ at peak corr timepoint

K-means clustering attack of 0/1 voltage magnitudes

- Unprotected RSA-4K datapath fabricated in 14nm CMOS
  - Power/EM attacks mounted on unprotected RSA
  - Single-trace (Horizontal) and multi-trace (Vertical) attacks to quantify SCA leakage

SCA Attacks on Unprotected RSA

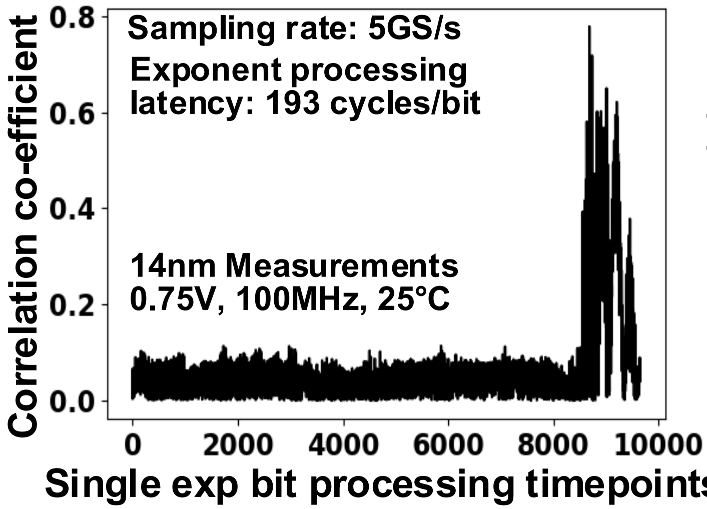
Peak correlation point identified on trace

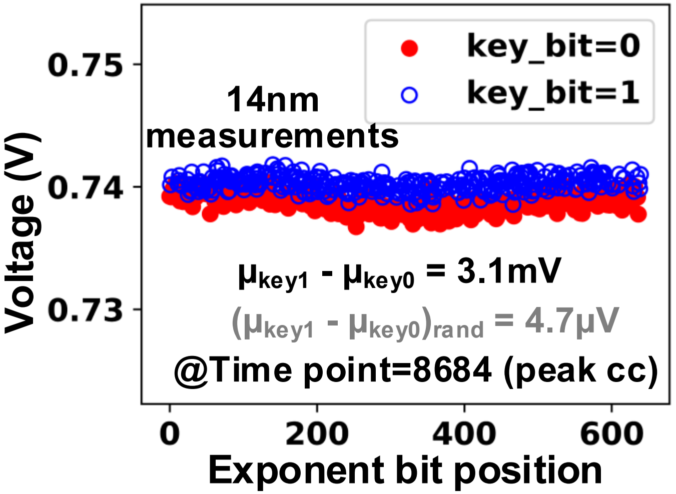
660× mean-separation over random binning of exponents

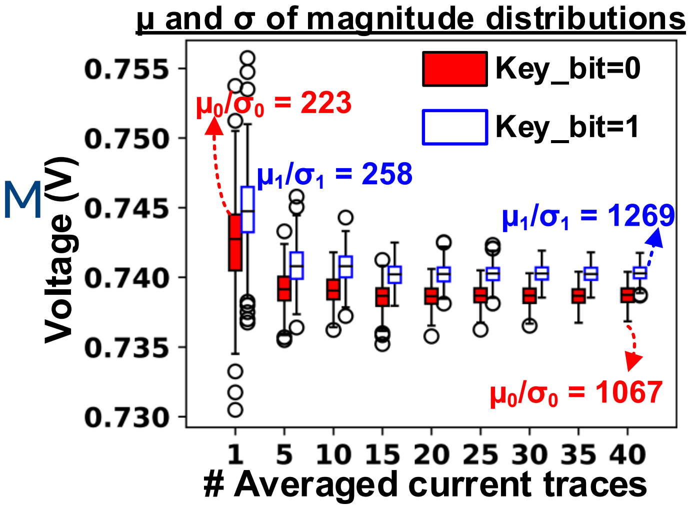
5.7× growth in μ/σ with multi-trace attacks

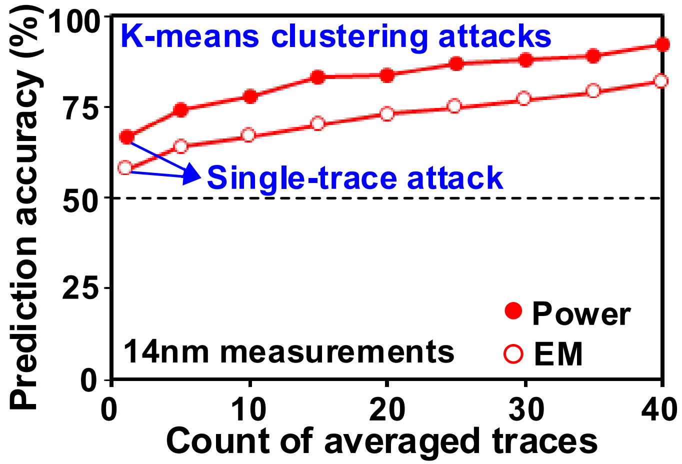
K-means cluster attacks show:

- 68/59% accuracy for single-trace power/EM attacks
- 91/80% accuracy for 40-way multi-trace power/EM attacks





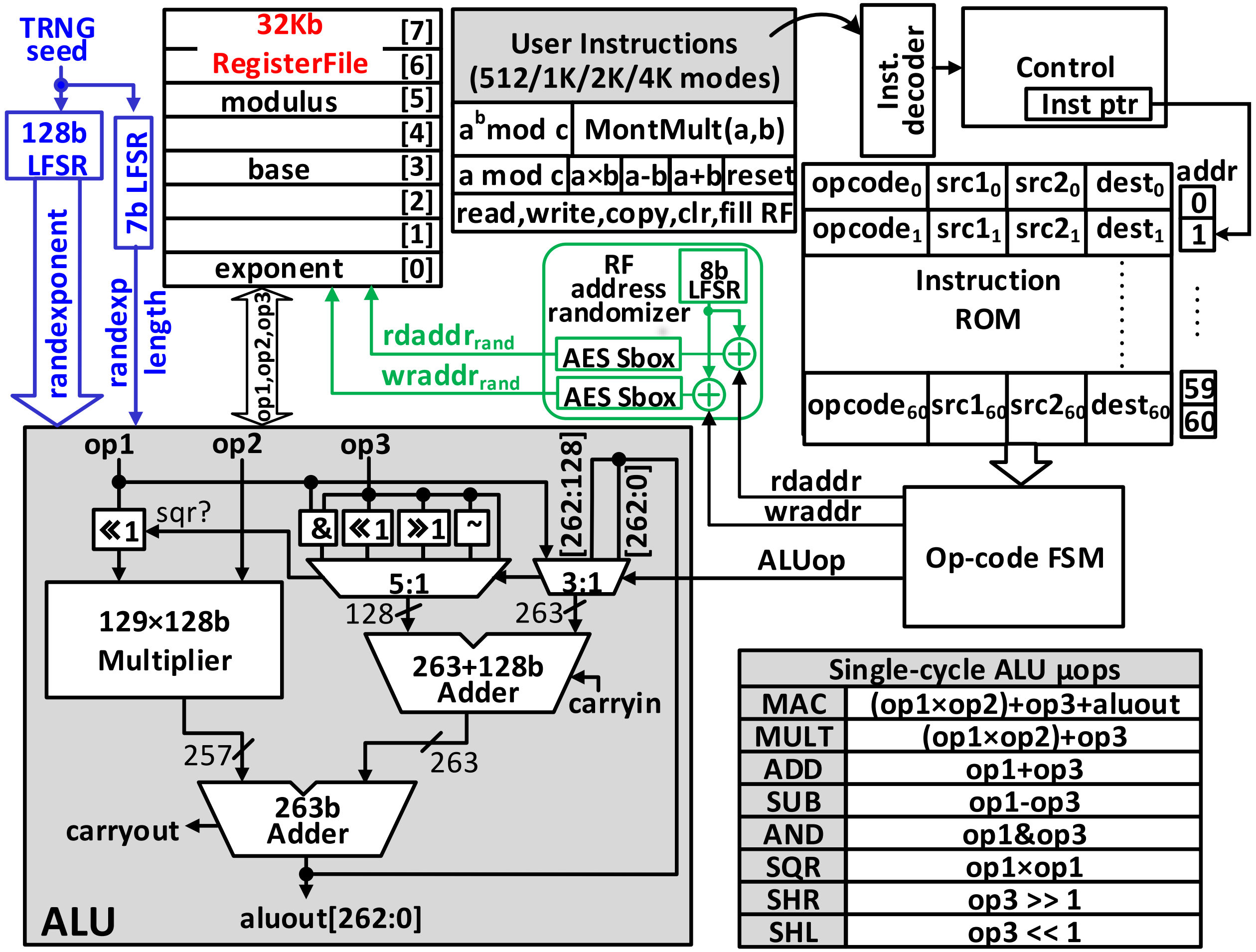




SCA-resistant RSA-4K Organization

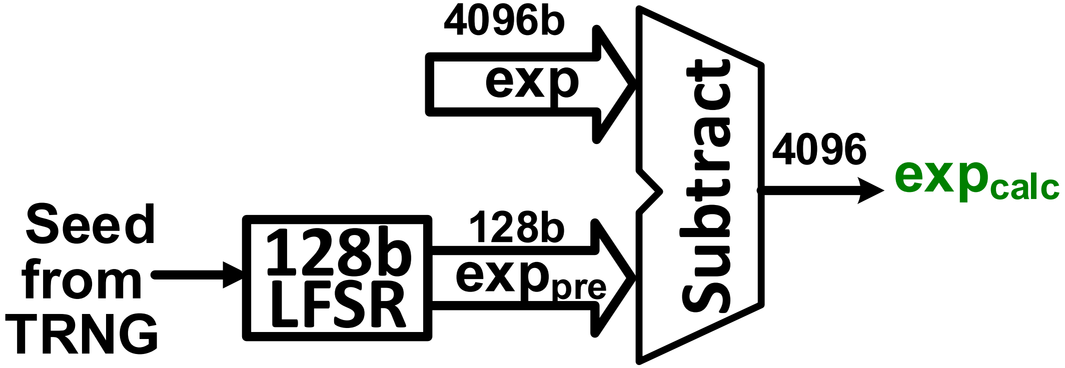
R. Kumar et al., VLSI 2020

- 128b ALU performs single-cycle multiply-add
- 32Kb RF stores operands and results
- User macro instructions are decoded to a program sequence
- Power/EM SCA-resistant features:
  - Exponent magnitude randomization
  - Exponent timing randomization



Single-cycle ALU µops	
MAC	(op1×op2)+op3+aluout
MULT	(op1×op2)+op3
ADD	op1+op3
SUB	op1-op3
AND	op1&op3
SQR	op1×op1
SHR	op3 >> 1
SHL	op3 << 1

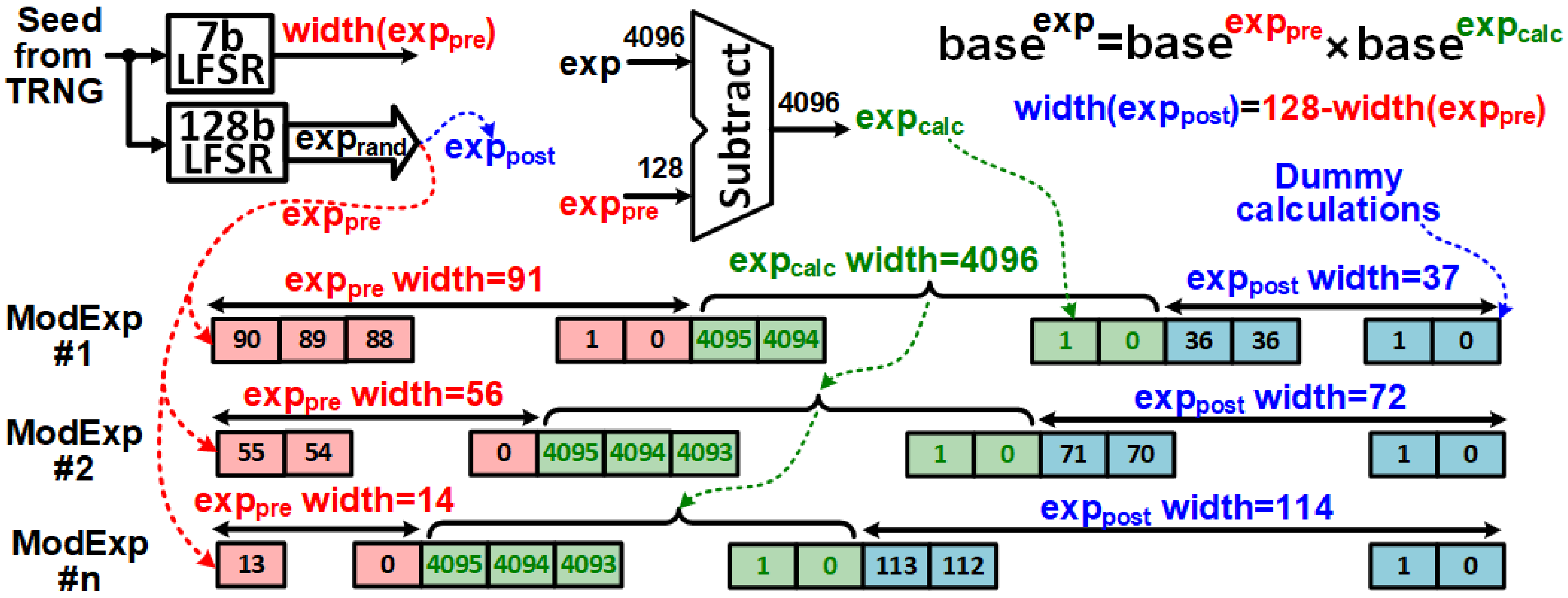
Exponent Magnitude Randomization



$$\text{base}^{\text{exp}} = \text{base}^{\text{exp}_{\text{pre}}} \times \text{base}^{\text{exp}_{\text{calc}}}$$

- Exponent split to a random 128b wide pre-exponent and calculated exponent
  - $\text{exp} = \text{exp}_{\text{pre}} + \text{exp}_{\text{calc}}$
  - Partial results from pre- and calc- exponents multiplied to get final product
- **25/3%** latency overhead for RSA-512/4K mode
- Constant pre-exponent width results in information leakage about expcalc

Exponent Timing Randomization

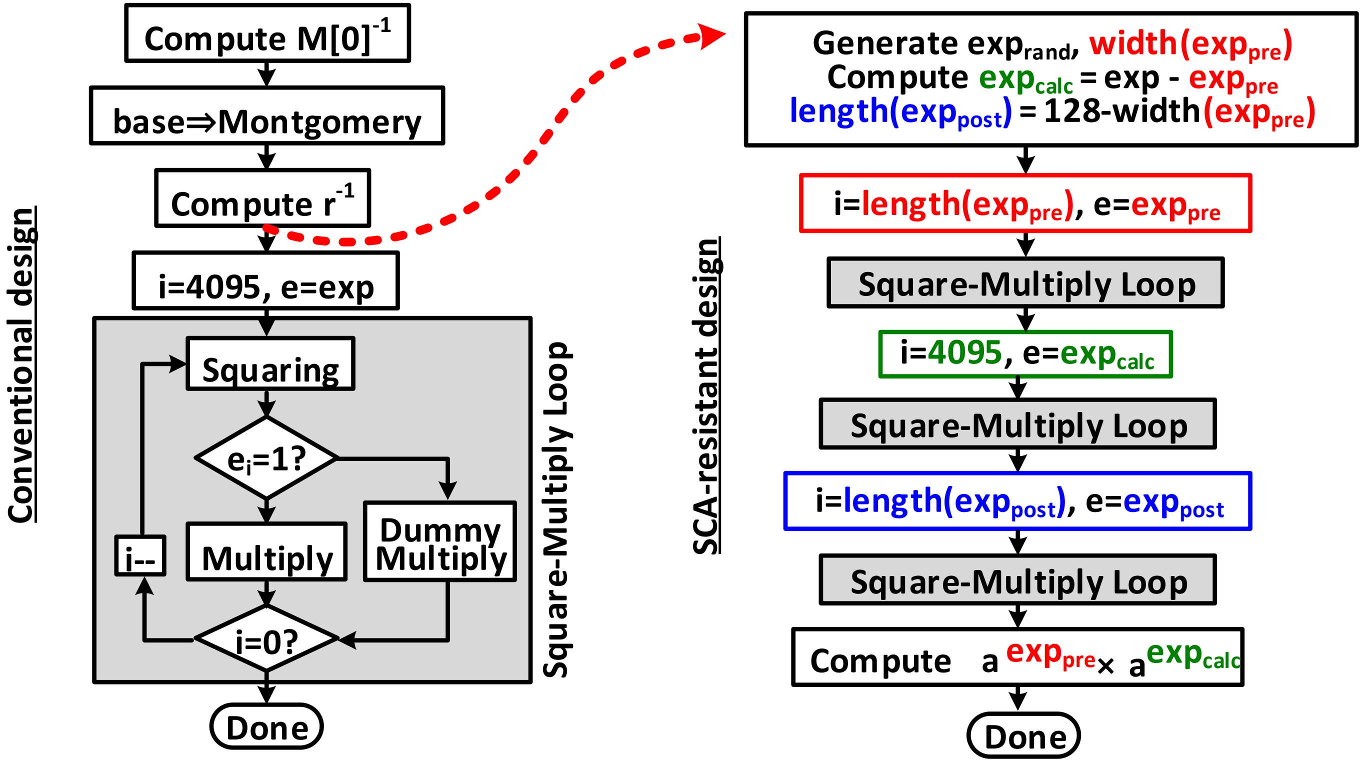


$$base^{exp} = base^{exp_{pre}} \times base^{exp_{calc}}$$

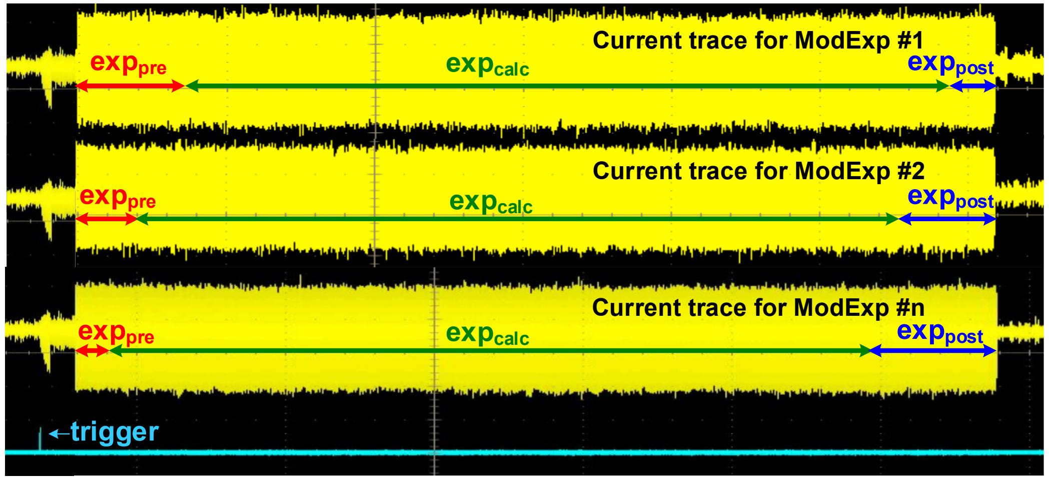
$$width(exp_{post}) = 128 - width(exp_{pre})$$

- Pre-exponent width randomized for every encryption run by a 7b on-chip LFSR
- Post-exponent width computed as 128-width(exp_{pre}) to produce time-invariant operation
  - $base^{exppost}$ results are written to memory locations not used in exponentiation
- Switching activities of exponents convoluted in single/multi-trace attacks

SCA-resistant Modular Exponentiation



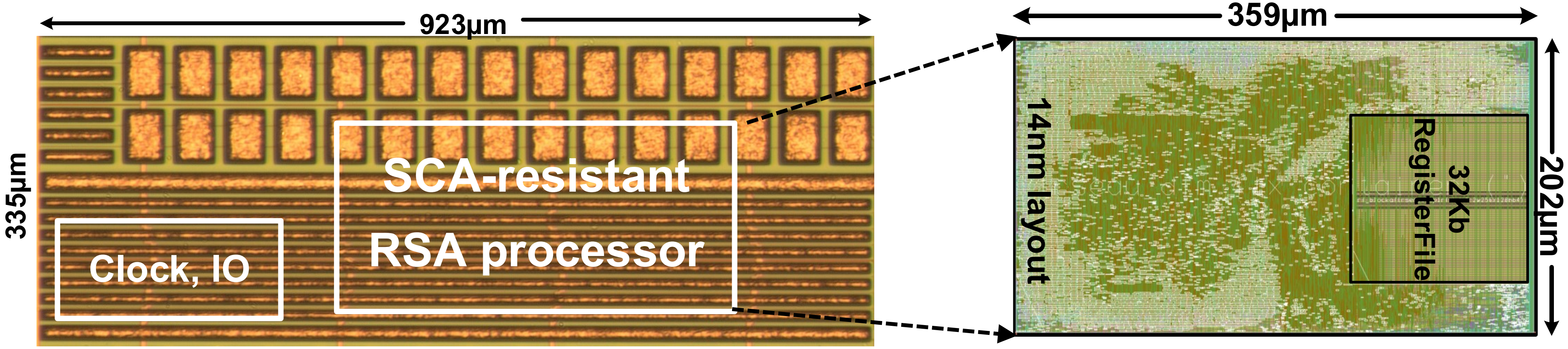
Measured current traces for different exponentiations



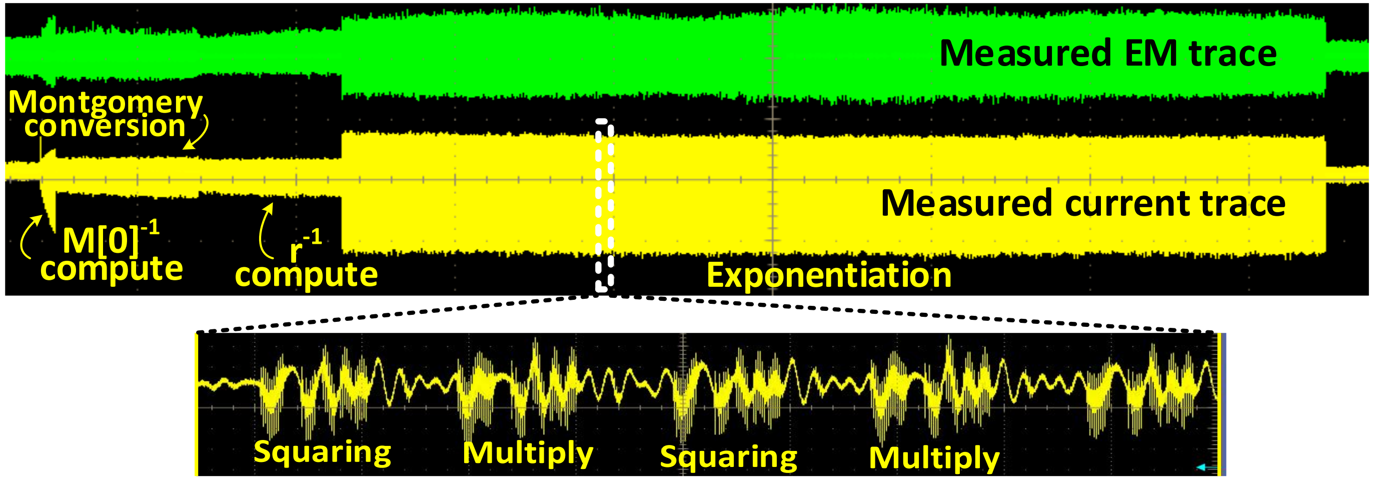
Main square-multiply loop interpolated between additional exppre and exppost loops

- Guaranteed total iteration count of 4224 in RSA-4K mode

14nm SCA Attack Measurements on RSA



- <0.05% area overhead from SCA-resistant features
- RSA-4K latency of 22M cycles (50ms sign/verification time)
- Langer RF-2/MFA probe set used for EM measurements

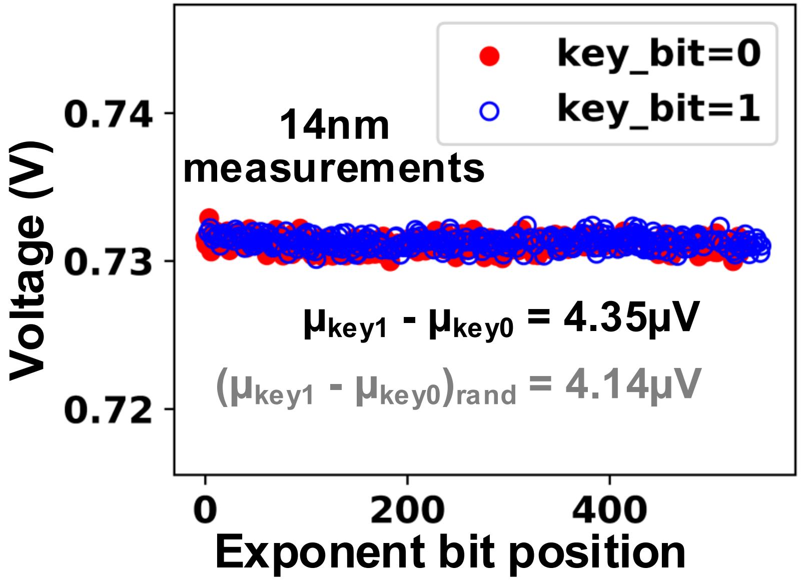


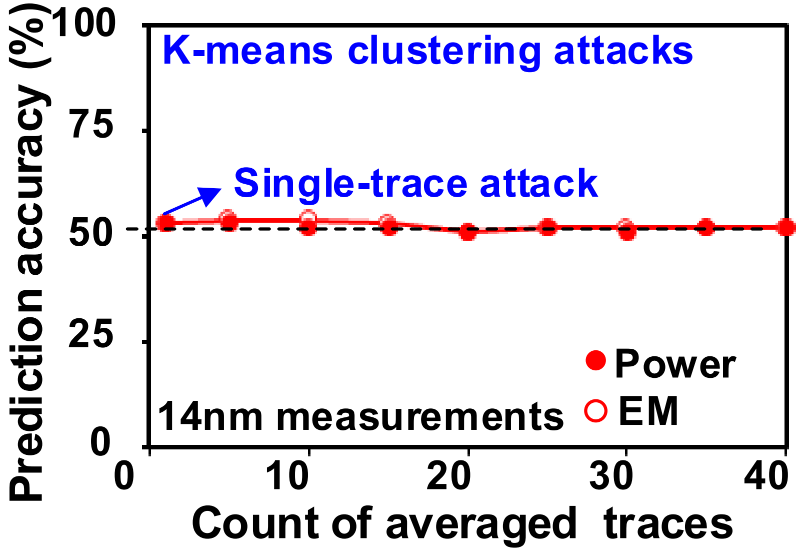
SCA Attack Measurements

711× suppression in means separation over unprotected RSA

K-means attack measurements:

- **52%** accuracy for single-trace
- **~50%** accuracy for multi-trace

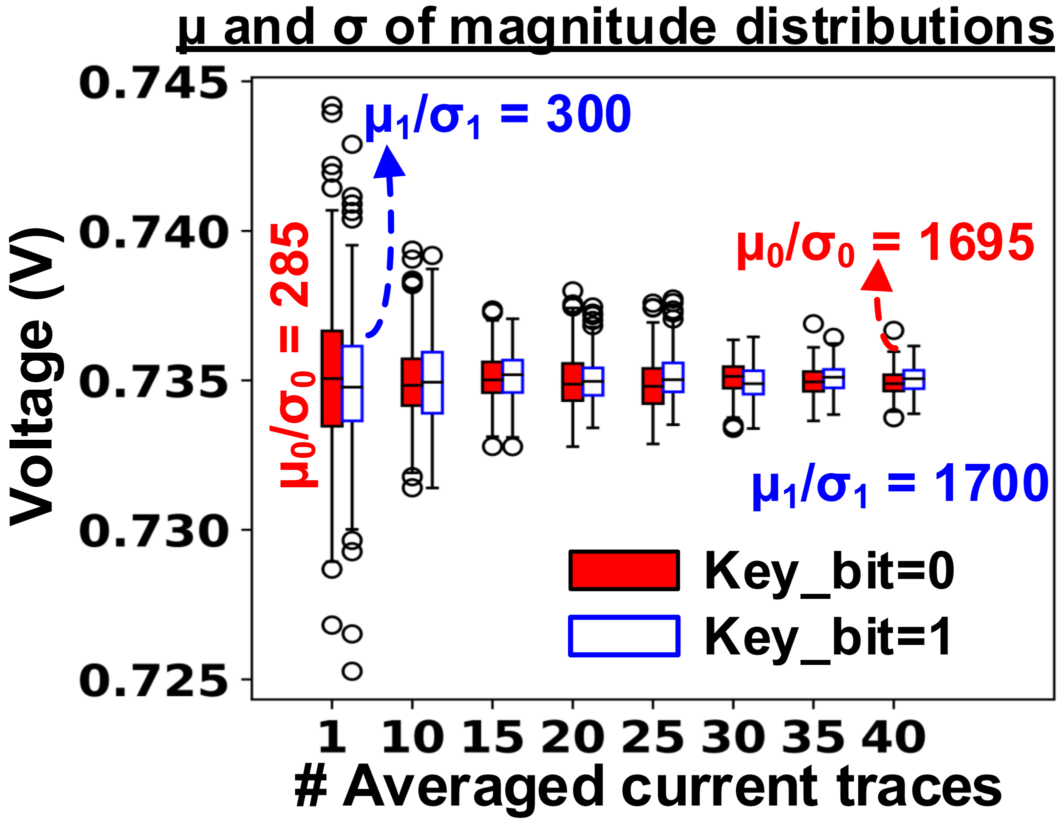




SCA Attack Measurements

Difference in μ/σ decreases by **3×** approaching a difference of zero

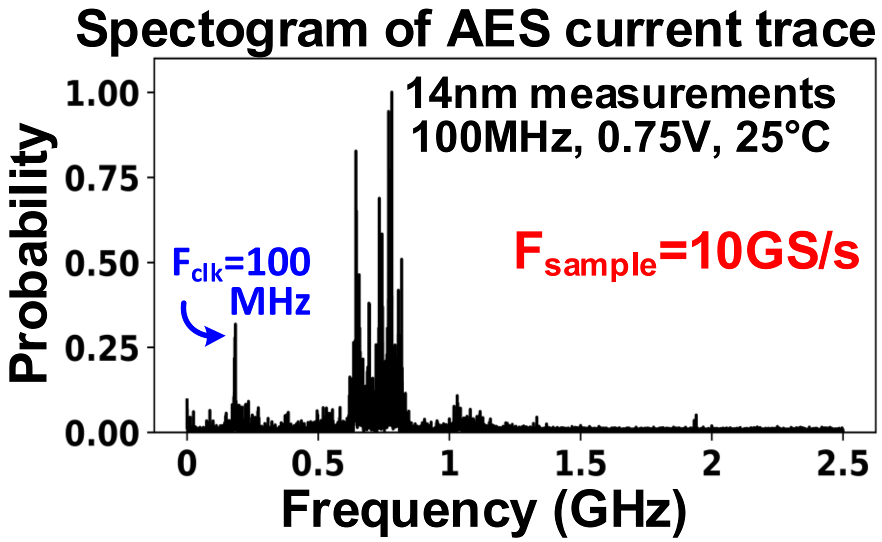
- No discernible separation in means for efficient binning



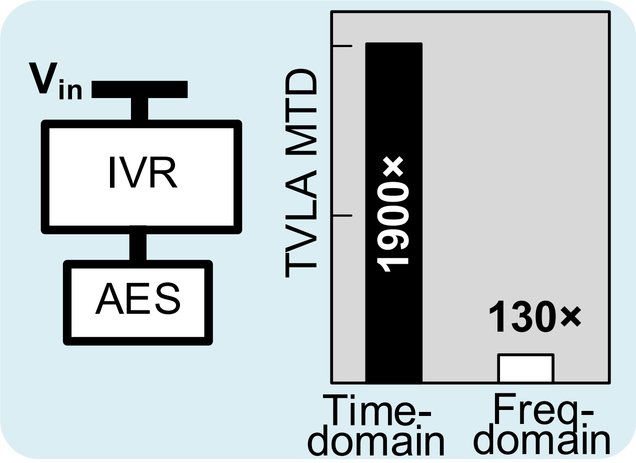
Algorithm-agnostic SCA-resistance

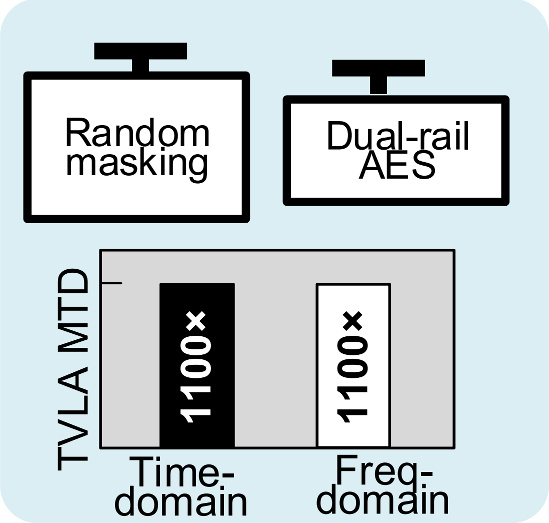
IVR-based SCA-AES

Non-linear DLDO Organization



Conventional SCA-resistant AES designs





AES current spectrogram shows information content in 600-800MHz band

Linear LDO leaks information in frequency-domain with modest SCA improvement

Non-linear LDO with fast response needed to mask high-frequency transients

Arithmetic techniques provide uniform improvements in time/frequency domain

Goal: >10^5× MTD improvement in both frequency/time domains

Non-linear DLDO Organization

R. Kumar et al., VLSI 2020

3-level non-linear controller

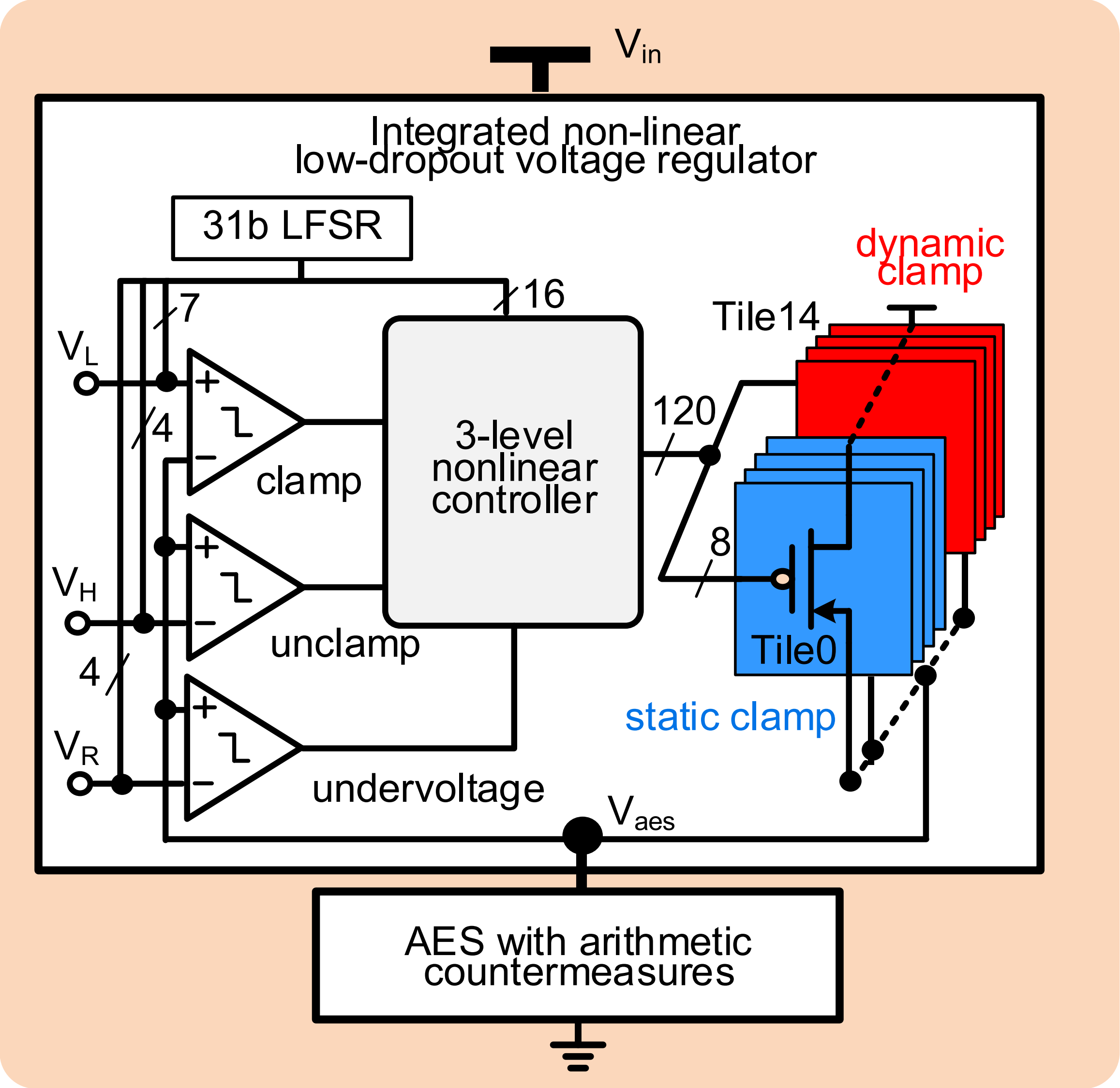
- **500ps** response time to fast transients

15 PMOS tiles, each configured as:

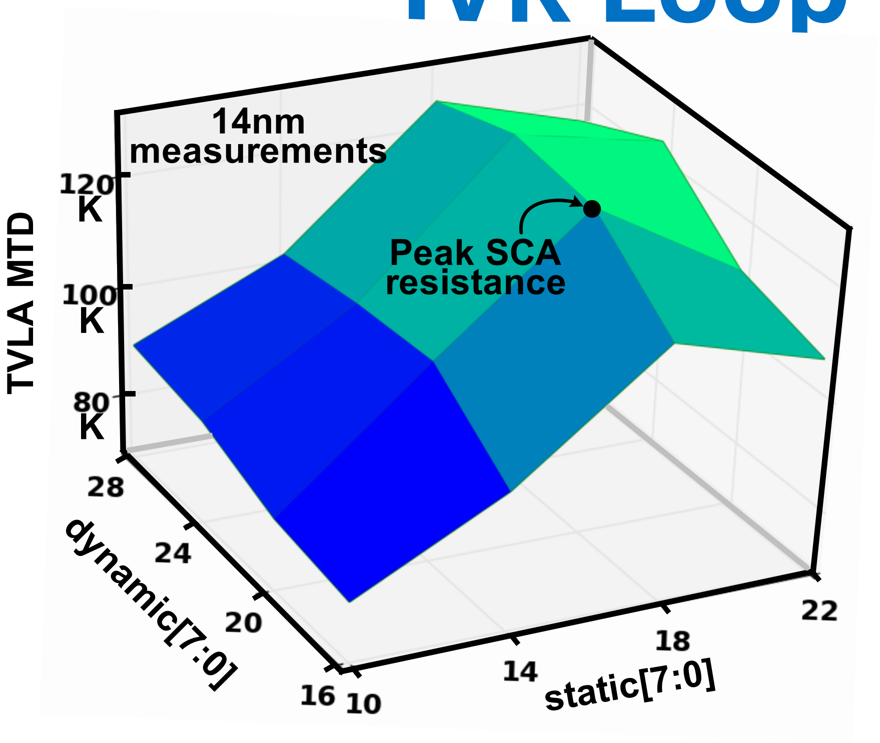
- Static current source
- Tunable dynamic clamp
- Variable combination of static/dynamic clamps

NLC loop triggered by 3 comparators:

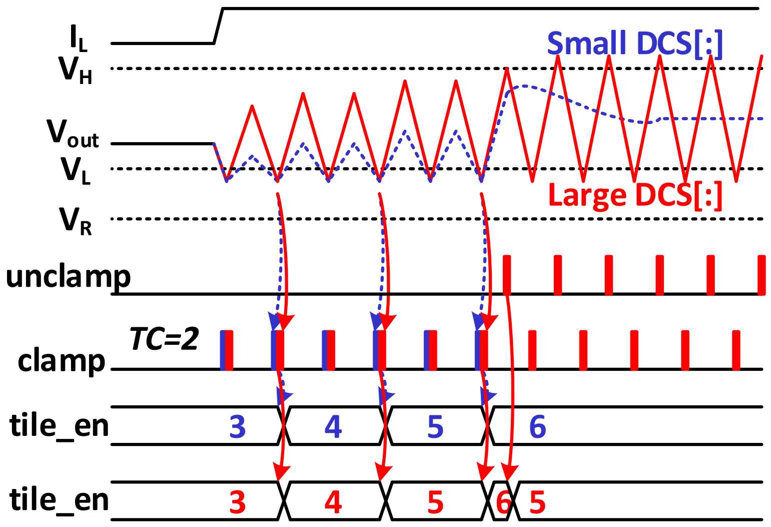
- V_H, V_L and reset (V_R)
- $V_{out}>V_H$ → disables all tiles
- $V_{out}<V_L$ → enables tiles at dynamic[7:0] strength
- $V_{out}<V_R$ → enables all tiles at full strength



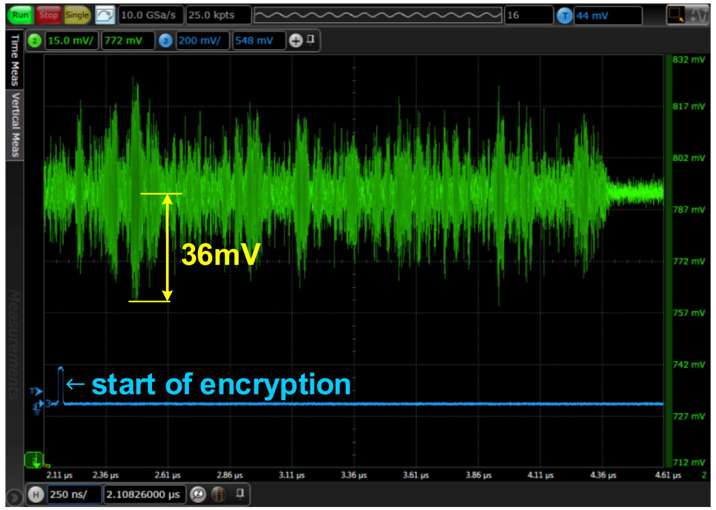
IVR Loop Parameters Exploration



Impact of clamp strengths



Measured V_{aes} waveform



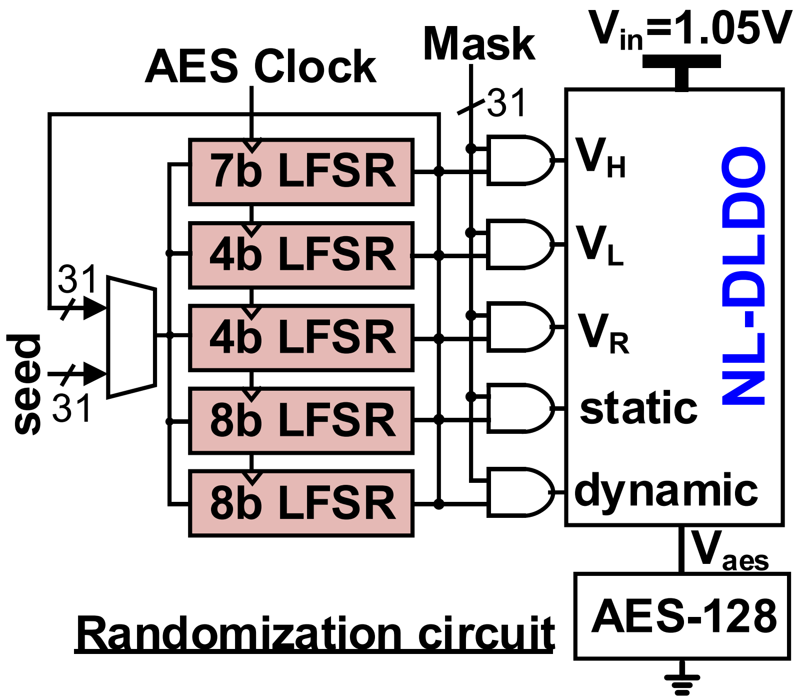
Static clamp strength modulates regulator output resistance

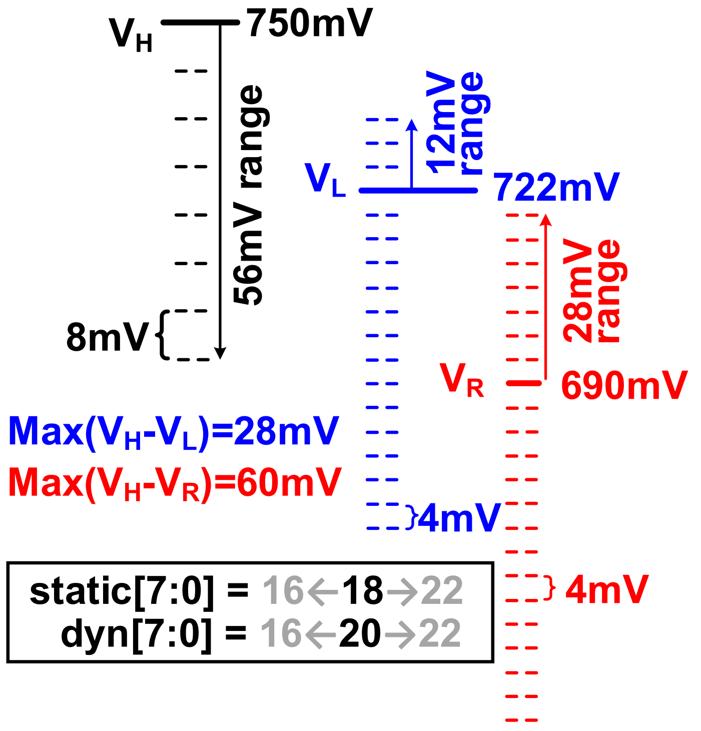
- Lower settings create frequent clamping, while higher settings propagate switching transients

Sweep of clamp strength shows peak SCA resistance at *TC*=7, *static*=18 and *dynamic*=20

Maximum droop of **36mV** observed across entire sweep range

Loop Parameters Randomization





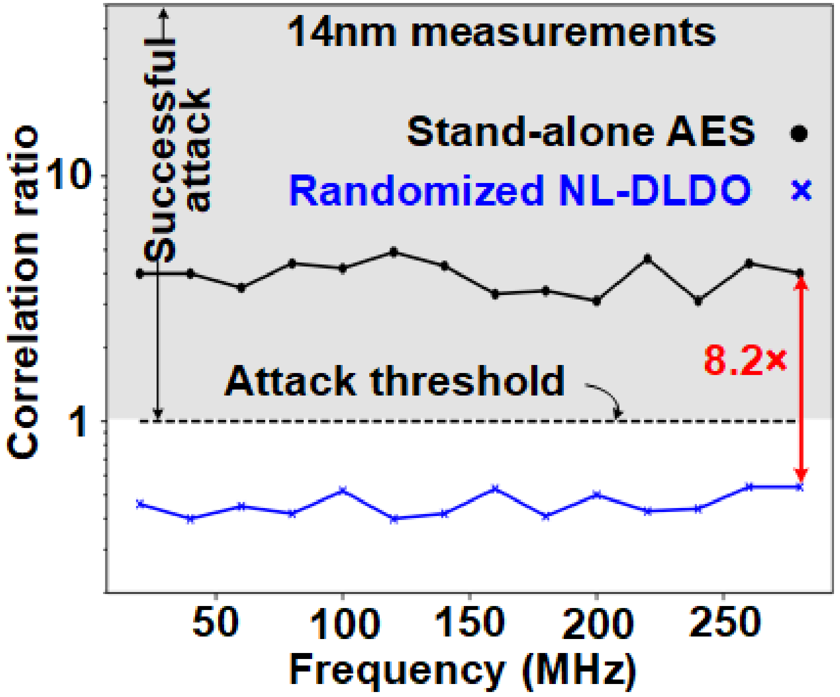
NL-DLDO parameters randomized with LDO biased at peak SCA resistance setting

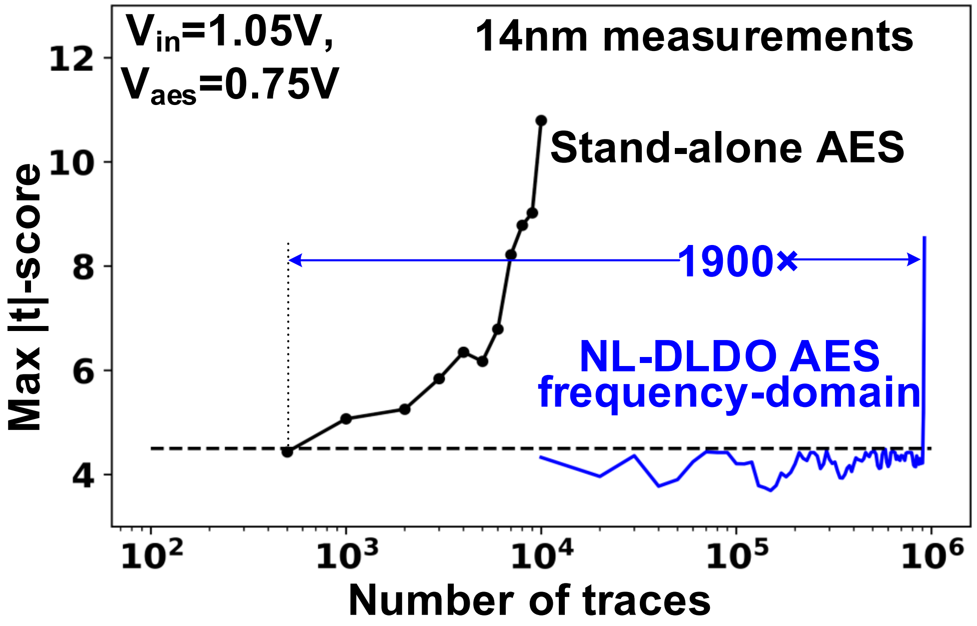
31b on-chip LFSR seeded by TRNG randomizes IVR control loop parameters

Randomized parameters: Clamp strengths, voltage thresholds (V_H, V_L and V_R)

31b mask value controls the randomization range for stable IVR operation

Loop Parameters Randomization





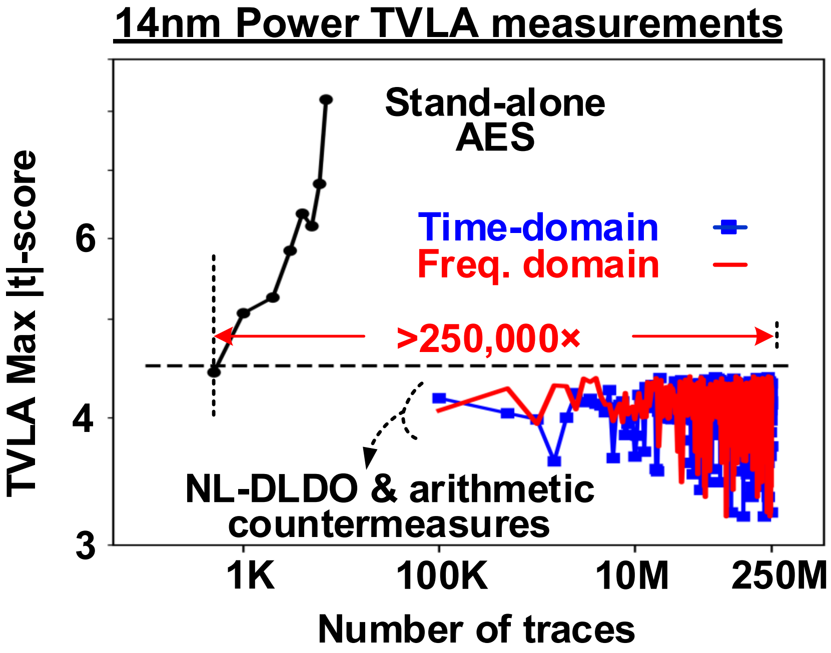
CPA attacks mounted with loop parameters randomized in stable operating range

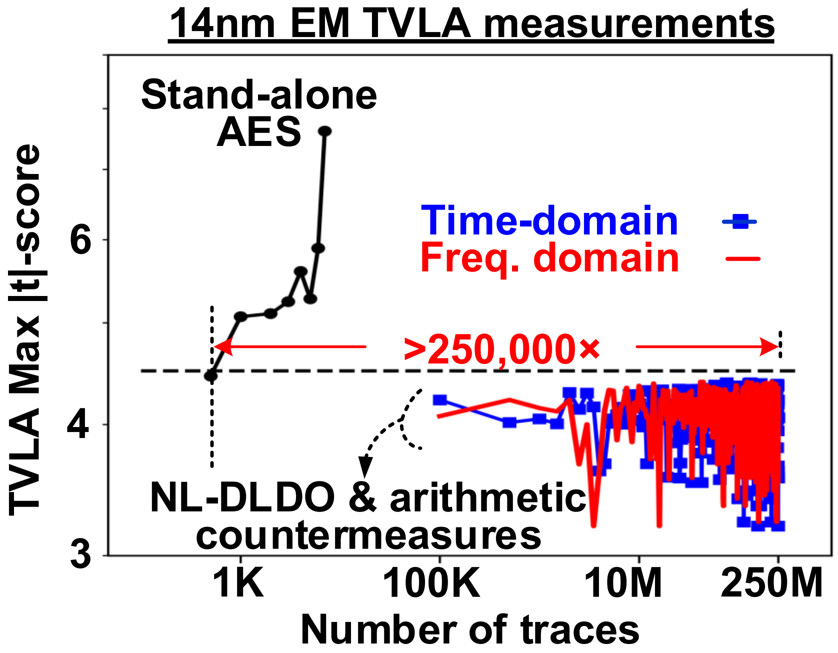
- Correlation ratio (CR) < 1 observed across entire frequency spectrum of load transients
- **8.2×** lower CR over standalone AES after 1M encryptions

1900× improvement in frequency-domain MTD over the unprotected AES implementation

Information leakage observed after 1M traces in frequency-domain

NL-DLDO with Arithmetic Countermeasures



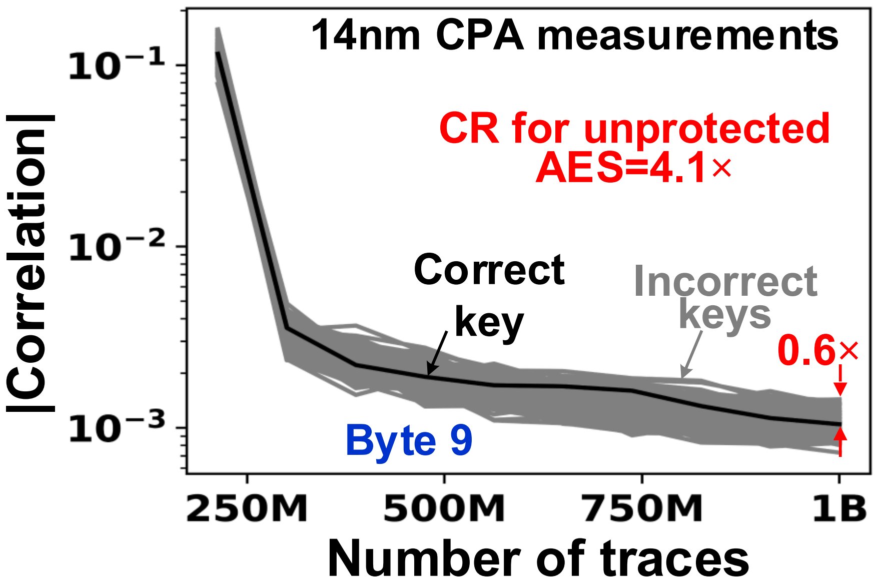


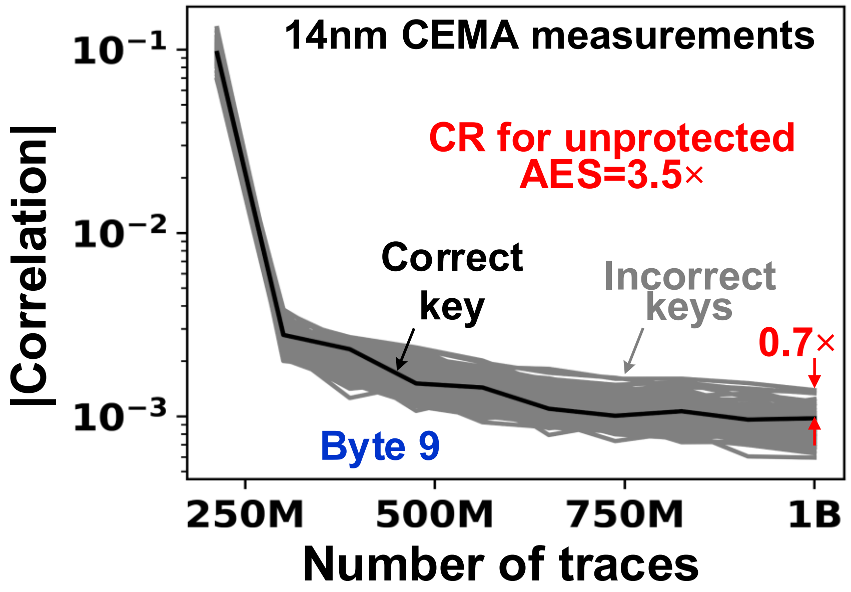
250M power and EM traces collected with loop parameter randomization and arithmetic countermeasures enabled

TVLA score < 4.5 for both power and EM measurements after 250M encryptions

- **>250,000×** improvement in MTD over unprotected AES implementation

CPA Measurements



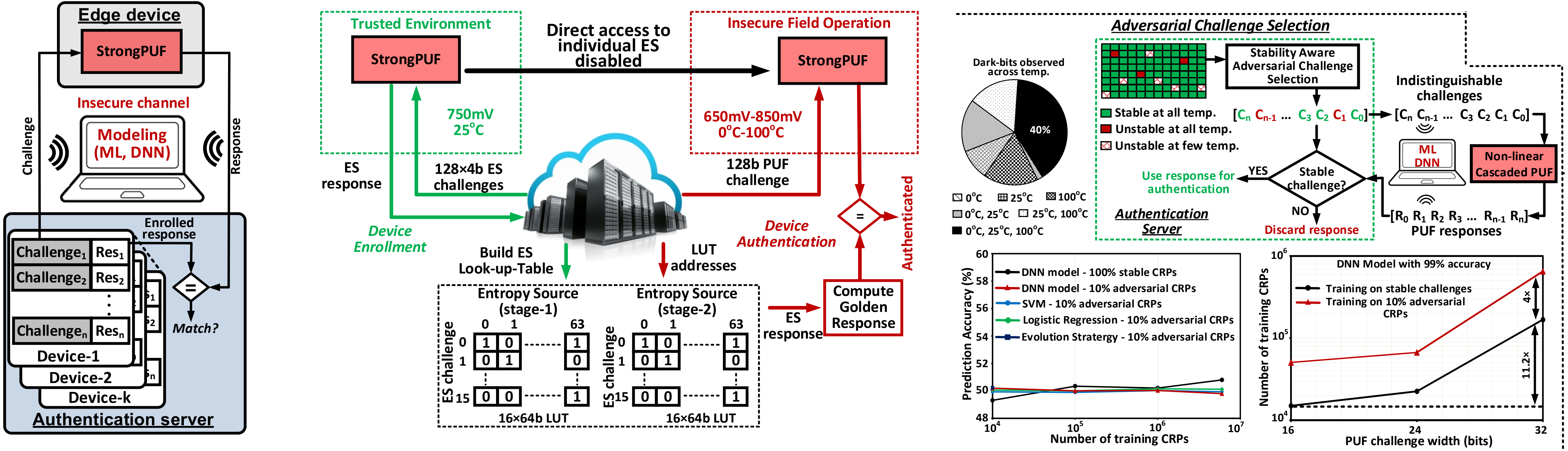


- **1Billion** power and EM traces collected
- CPA and CEMA attacks show no key bytes were extracted with 1B encryptions
  - **5-6.8×** suppression in trace correlation ratios

Machine-Learning Attacks

ML-resistant StrongPUFs

ML-modeling attack resistant StrongPUF



V. Suresh et al., VLSI 2020

- StrongPUF offer a lightweight solution to secure authentication
- They are vulnerable to machine-learning modelling attacks
- Stability-aware challenge pruning reduces BER to 0.26%
- 2-stage non-lineary cascaded PUF with adversarial challenge selection

Summary

- Attack-resistant crypto HW are the foundation of secure systems
- SCA-resistant AES
  - Random additive-masking
  - Reconfigurable AES with blind-bulk mode
  - Heterogenous Sbox AES
  - Multiplicatively masked AES
  - Non-linear IVR-coupled AES
- SCA-resistant RSA
  - Exponent magnitude/timing randomization
- ML-resistant Strong PUF
  - 2-stage non-lineary cascaded PUF with adversarial challenge selection

intel®