جامعة نيويورك أبوظبي
NYU | ABU DHABI

CENTER FOR CYBER SECURITY

DFX LAB
DESIGN FOR EXCELLENCE

Trojan Insertion versus Layout Defenses for Modern ICs: Red-versus-Blue Teaming in a Competitive Community Effort

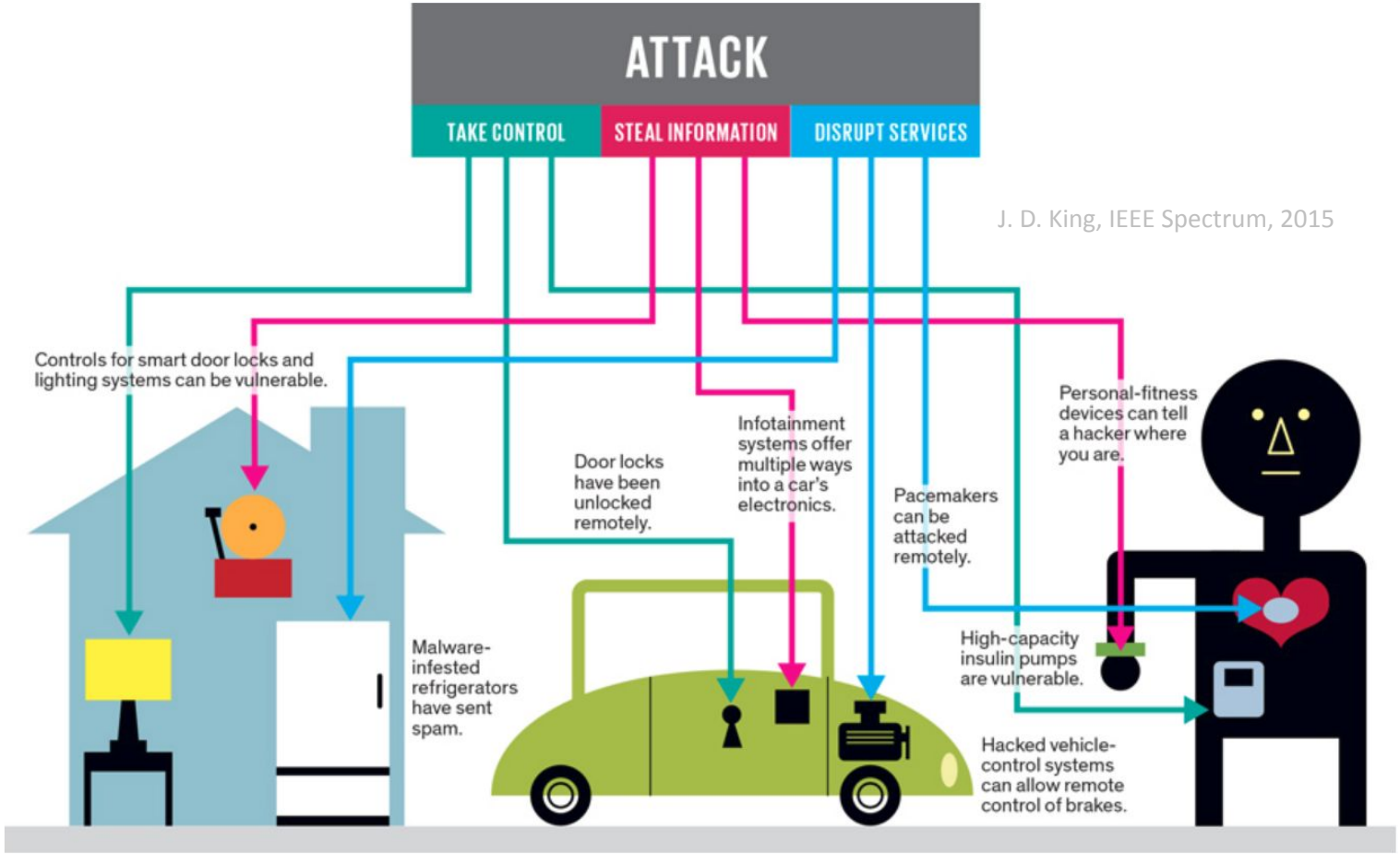
Johann Knechtel

johann@nyu.edu

wp.nyu.edu/johann

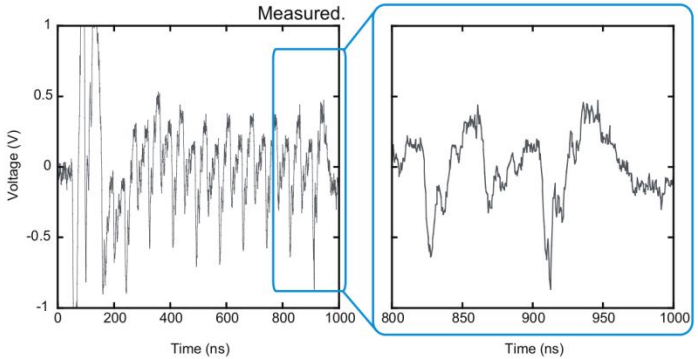
HARRIS 2025

Introduction

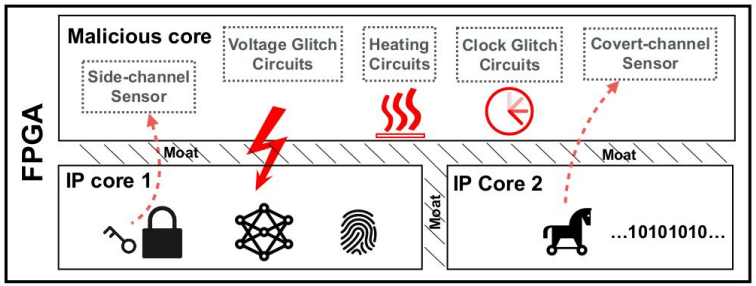


J. D. King, IEEE Spectrum, 2015

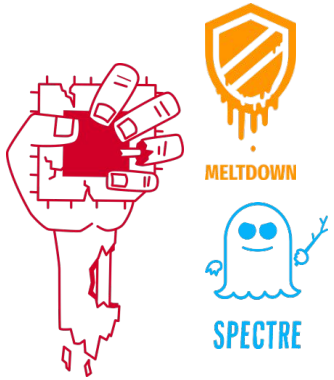
Data and Computation at Risk – Right at the Hardware

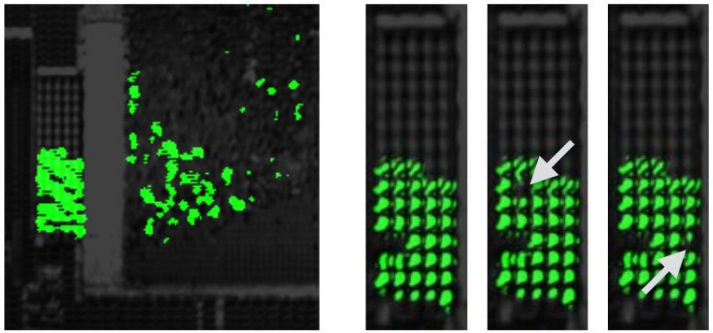


Fujimoto et al., EMC 2014



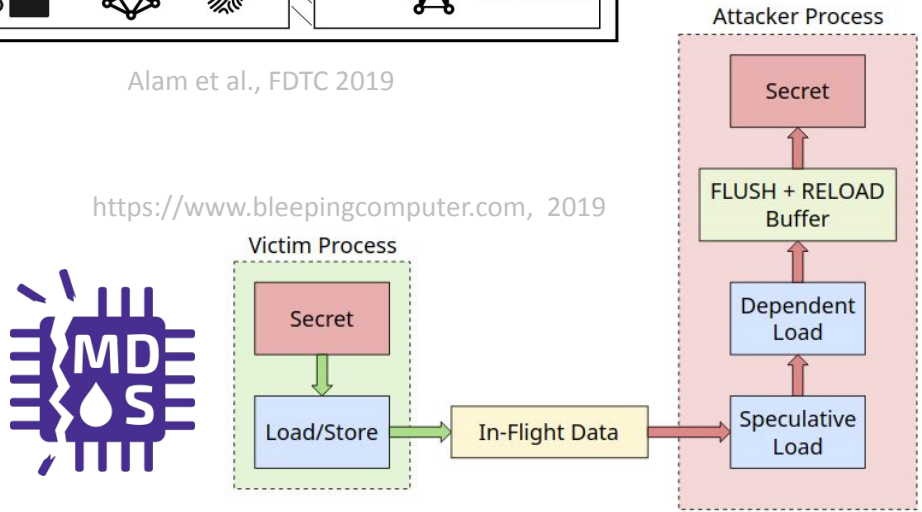
Alam et al., FDTC 2019





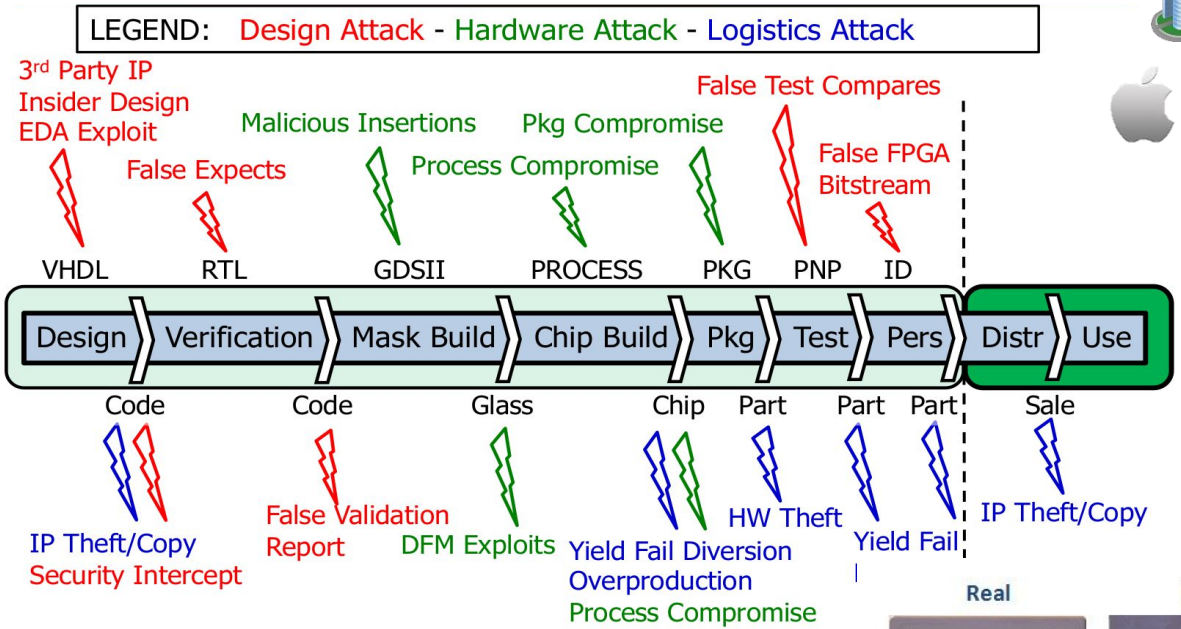
Tajik et al., CCS, 2017

https://www.bleepingcomputer.com, 2019

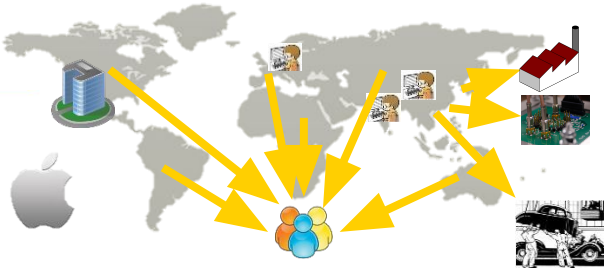


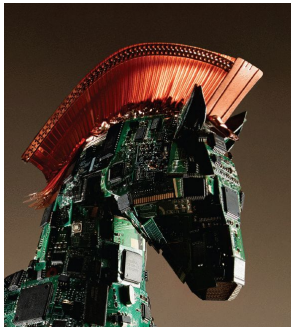
Knechtel, "Trojan Insertion versus Layout Defenses for Modern ICs: Red-versus-Blue Teaming in a Competitive Community Effort," HARRIS 2025

Hardware Itself Also at Risk



Kerry Bernstein, DARPA, 2016





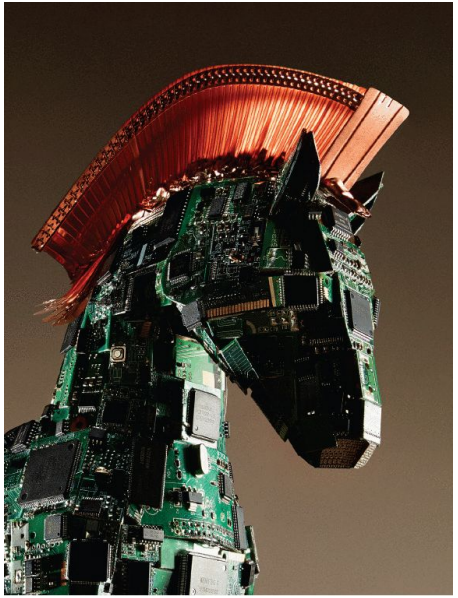
IEEE Spectrum, 2015

Contest at ISPD’23

- Motivation:
  - More and more threats are arising that affect hardware
  - Build up knowledge and experience in CAD community

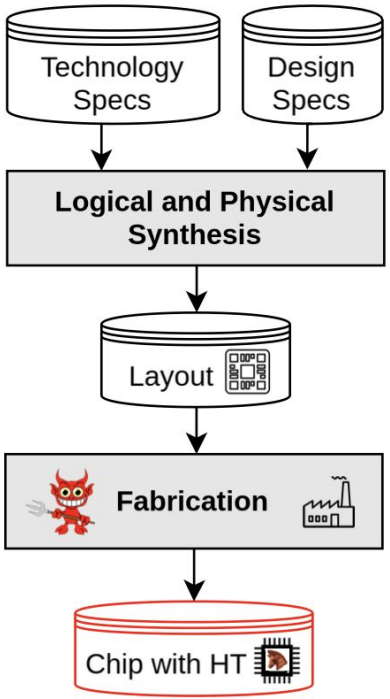
- Main theme:
  - Incorporating techniques for designing secure and trustworthy ICs in future CAD flows
  - This time, Hardware trojan horses

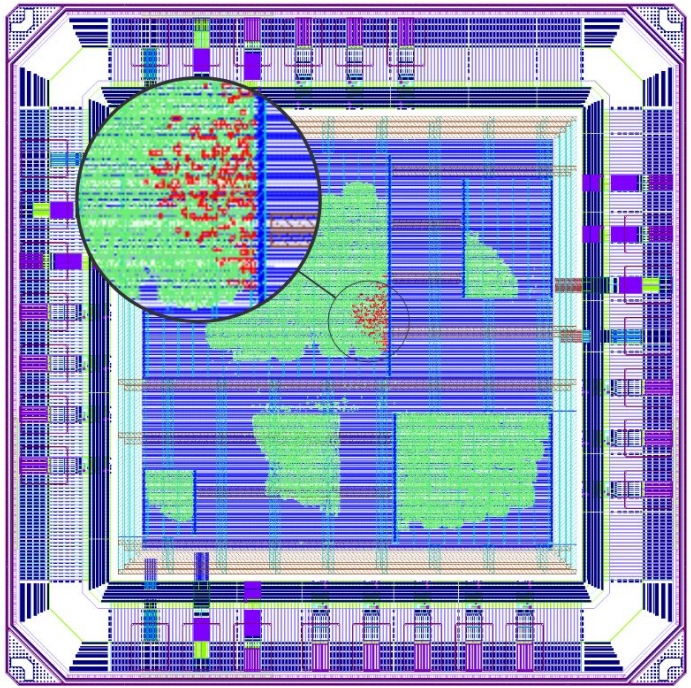
Full paper to be presented at CHES 2025



Main Part

Threat Model



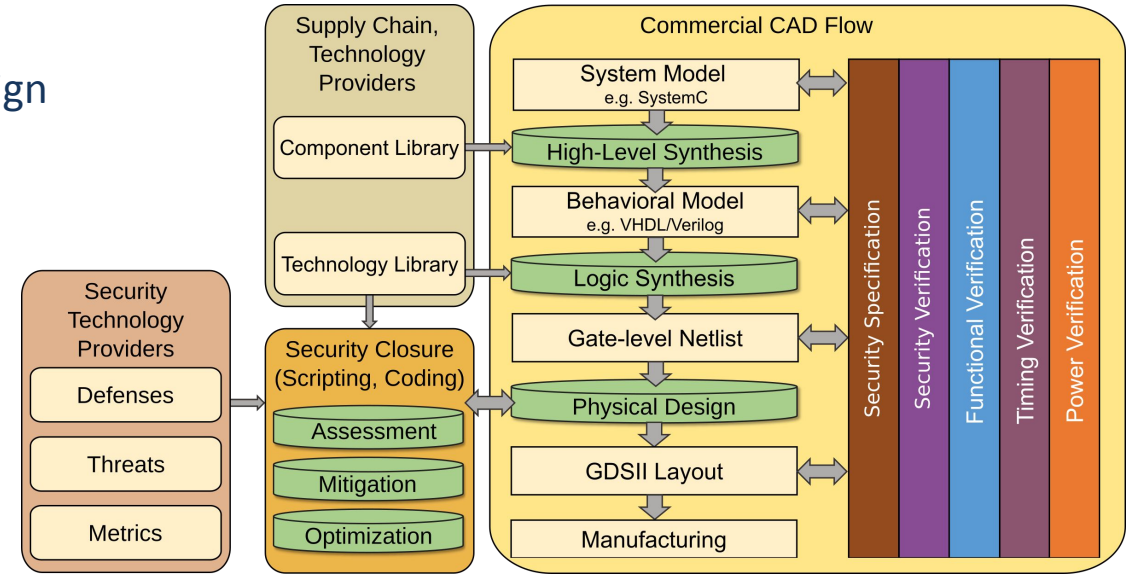


Contest Objectives (Blue Teams)

- Implement physical-design measures to proactively harden layouts against post-design Trojan insertion during mask generation or manufacturing

- Participants must enhance security while accounting for impact on design rules and PPA

- There is no single, right or wrong approach toward that end; complex multi-objective problem

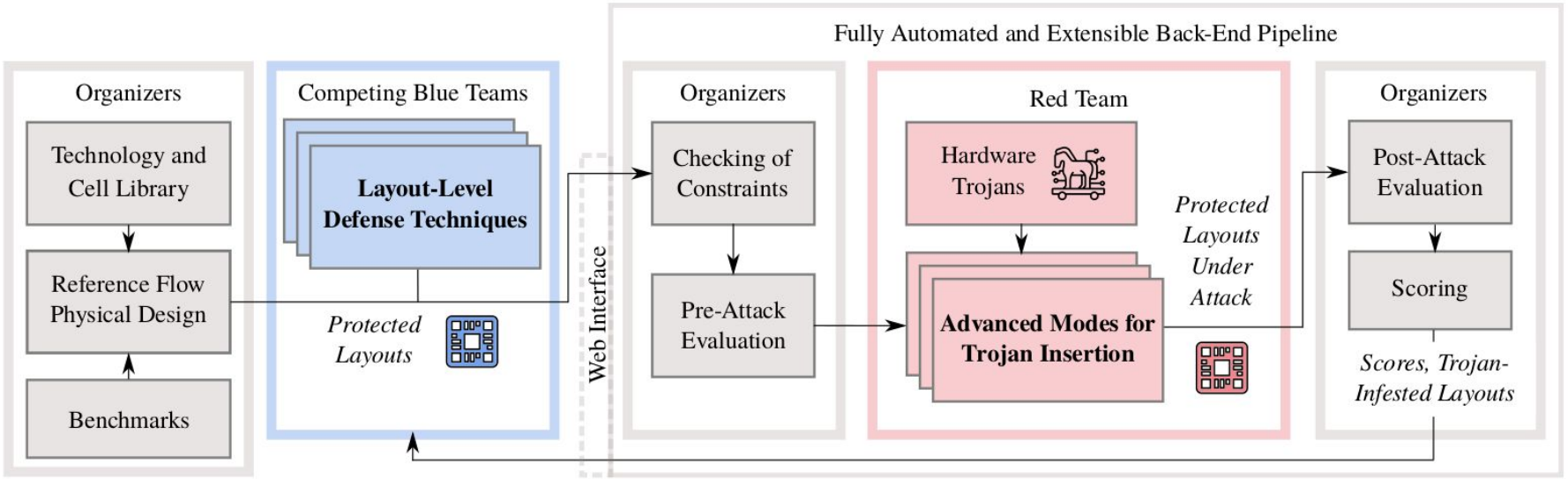


Prior Art

	[MGK+13]	[PMB+23]	[ST16]	[WZL24]	[PP22]	[GMMP20]	[KGB+21]	[TSBH20]	[WWF+23]	[TSBH23]	[GYT+23]	[EPP23]	[This]
Competition	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Red-versus-Blue	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Automated HT Insertion	✓/✗	✗	✗	✓/✗	✓	✓/✗	✗	✗	✗	✓/✗	✗	✗	✓
Automated HT Detection	✓	✓/✗	✓/✗	✓/✗	○	✓/✗	○	✓/✗	○	✓/✗	○	○	○
Actual HTs	✓	✗	✓	✓	✓	✓	✗	✗	✗	✓	✗	✗	✓
Placement Def.	✗	✗	✗	✓	✗	✗	✓	✗	✓	✗	✓	✓	✓
Routing Def.	✗	✗	✗	✗	✗	✗	✓	✗	✗	✓	✓	✓	✓
Rule Out Spares, Fillers	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Technology Nodes (nm)	180	90, 65, 40, 28	90	45	65	90, 65	45	45	45	45	45	65	**7**
IC Tape-Out	✓	✓	○	○	✓	✓	○	✗	○	✗	○	✓	○

Def. is short for defense. Symbols: ✓ means yes, ✗ means no, ✓/✗ means to some degree, and ○ means not applicable / out of scope.

Contest Components (Benchmarking Framework)



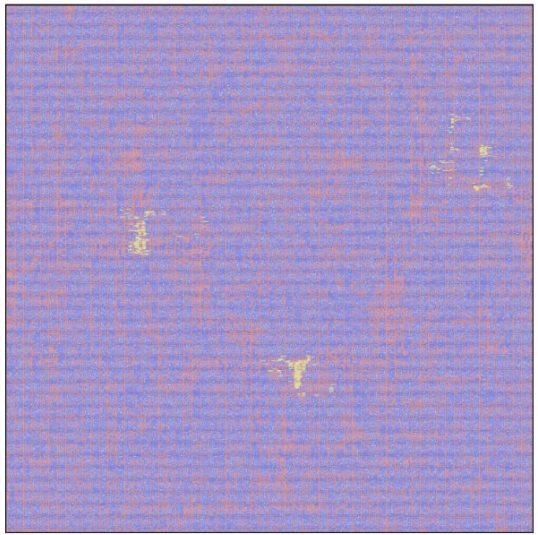
Benchmarks

- 6 different crypto cores as benchmarks
- Different sizes and complexities, ensuring different difficulty levels across the benchmarks
- Optimization is refrained from on purpose, to keep some margin for the teams to work with
- Cell assets: exemplary sensitive components, like key registers

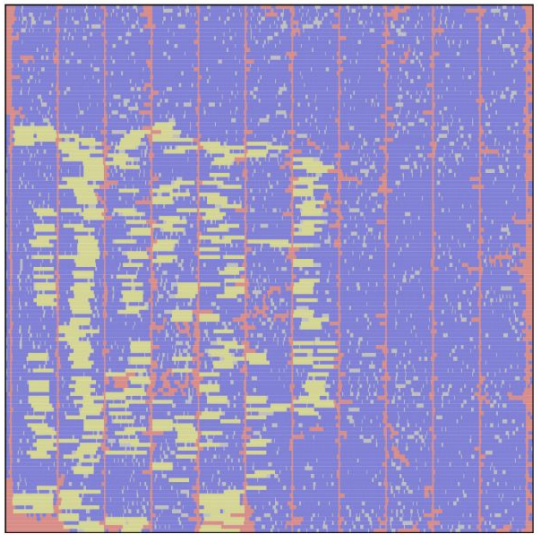
	Dimensions (µm)	# Cells	Util. (%)	# Cell Assets	CP: (ns)	WNS (ps): Setup / Hold	# Sites Across ERs: Sum / Max / Med	Routing Tracks: F. (%) / T.(#)
AES128	822.44 × 822.44	263,618	67.34	384	0.17	34.29 / 33.25	662,065 / 460,741 / 29	63.89 / 29,331
CAMELLIA	158.24 × 158.24	10,101	84.02	396	0.27	22.23 / 20.82	8,820 / 1,403 / 46	58.76 / 5,634
CAST	293.24 × 293.24	24,450	52.52	143	0.66	25.49 / 18.22	147,359 / 139,439 / 30	62.44 / 10,452
MISTY	174.44 × 174.44	10,558	68.77	332	1.85	05.00 / 20.97	26,039 / 5,932 / 33	65.54 / 6,215
SEED	206.84 × 206.84	17,334	76.65	127	1.02	34.22 / 29.35	25,478 / 3,854 / 33	65.77 / 7,369
SHA256	190.64 × 190.64	9,708	71.09	125	0.60	20.80 / 22.63	25,964 / 2,133 / 25	68.26 / 6,792

Util. is short for layout/placement utilization; *CP* is short for clock period, i.e., the global timing constraint; *ERs* is short for exploitable regions; *Med* is short for median; *F.* is short for free; *T.* is short for total.

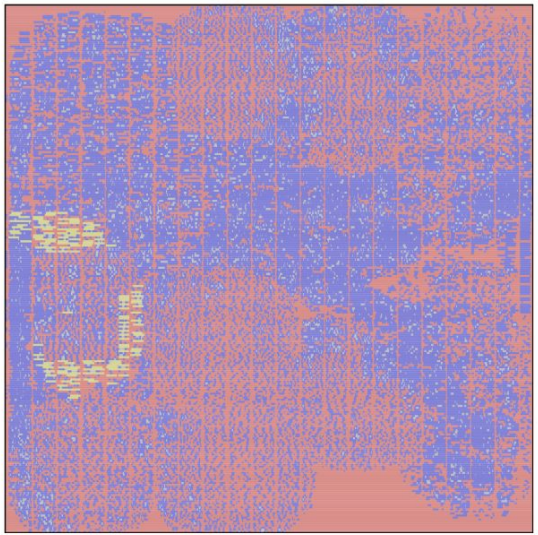
Benchmark Layouts



(a) AES128

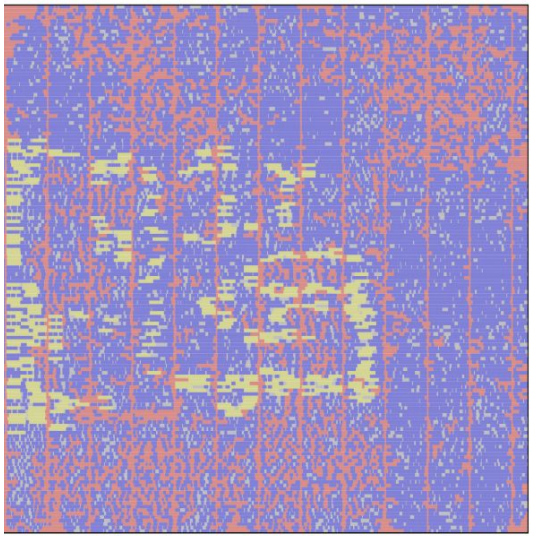


(b) CAMELLIA

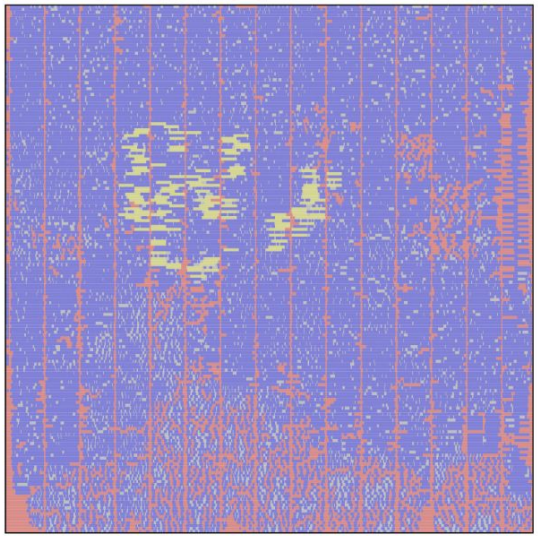


(c) CAST

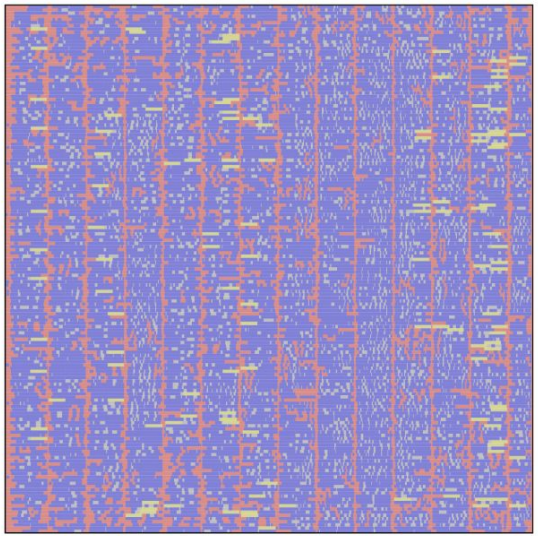
Benchmark Layouts



(d) MISTY



(e) SEED



(f) SHA256

Benchmark Implementation

ASAP7 PDK and Library:

- Originally developed by teams from Arizona State and ARM
- Likely the most complete PDK developed by and for academia; open-source
- Many files provided do resemble a commercial PDK, including multi-Vth cells, extraction decks, DRC decks
- A few modifications were made to the library:
  - Addition of colored metals
  - Introduced max density rules for all metal layers

- Replaces the Nangate 45nm Open Cell Library used in 2022

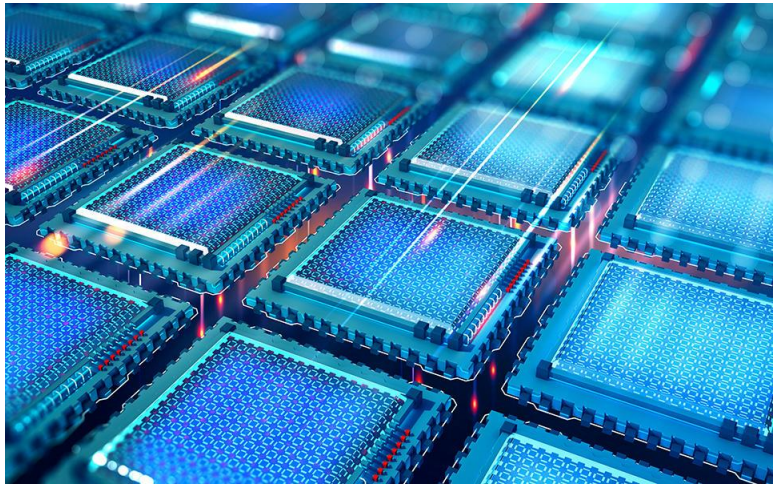
Benchmark Implementation and Reference Flow

1) Defining global variables: library files, design files, timing, etc.
2) Floorplanning and PDN planning: ring spacing, offset, and size; stripe-to-stripe distance
3) Pin Assignment: all input pins on the left side and all the output pins on the right side.
4) PDN: core rings in M6 and M7, follow pins in M1 and M2 (“stapled style”), vertical and horizontal stripes in M3 and M4
5) Place, CTS, and route
6) “Tape out”: Performing all necessary checks and verification, exporting layout, and post-route reports for PPA, etc.

Benchmark Release

The benchmark release includes:

- Post-route Verilog netlists and DEF files
- Design databases
- SDC timing files
- List of cell assets
- Scoring of the baseline layout
- Evaluation and scoring scripts

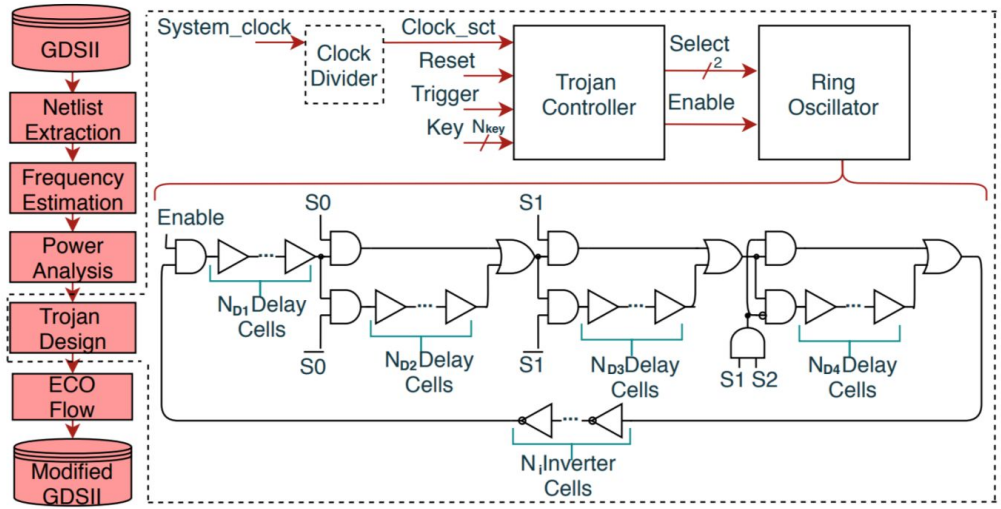


nist.gov

- https://wp.nyu.edu/ispd23_contest
- Benchmarks will stay online
- Best results, benchmarking framework, and reference flow also published

Security Metrics

1) Security for alpha/qualifying round: first-order metrics
   - Regions of 20+ continuous open placement sites
   - Free routing tracks
2) Security for final round: 1) + actual Trojan insertion, based on ECO design flow



T. Perez, M. Imran, P. Vaz and S. Pagliarini, "Side-Channel Trojan Insertion – a Practical Foundry-Side Attack via ECO," Proc. Int. Symp. Circ. Sys. (ISCAS), 2021, pp. 1-5

Scoring

$$score = [sec + des]/2 = [(1/2 \times sec\_ti\_gen + 1/2 \times sec\_ti\_ECO) + des]/2$$

(1) Trojan insertion, generic evaluation – *sec_ti_gen*

(a) 50%: open placement sites of exploitable regions (*sec_ti_sts*)

(b) 50%: free routing resources of whole layout (*sec_ti_f ts*)

(2) Trojan insertion, actual ECO insertion – *sec_ti_ECO*

- Score sheet, with lower scores for better defense / more difficulties for Trojans

  0—2 design failures; 5—7 DRC violations; 10—12 setup AND hold violations;
  15—17 setup XOR hold violations; 20—22 DRV OR clock check violations;
  25—27 no violations

- Normalized over worst-case (27); averaged across ECO modes;
  gap b/w categories intended

Scoring

(3) Design quality – des

(a) 33.3%: power (des_pwr)

(b) 33.3%: performance (des_prf)

- 50% weighted: ($des\_prf\_WNS\_set$)
- 50% weighted: ($des\_prf\_WNS\_hld$)

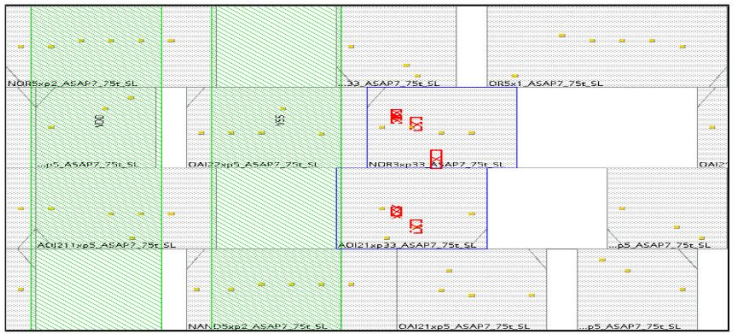
(c) 33.3%: area (des_ara)

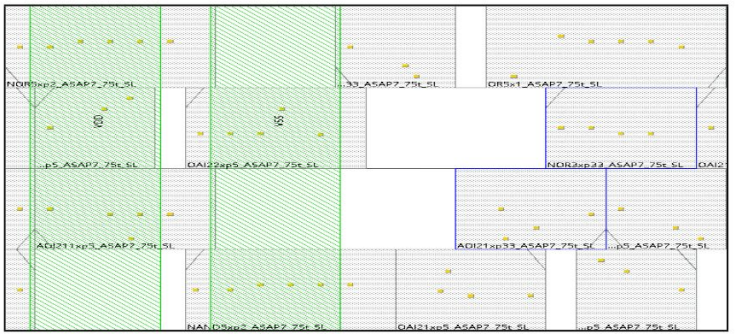
Constraints I

- Cannot incorporate trivial defenses --- filler, decap, and tap cells are purged
- Must meet setup, hold timing checks using the provided SDC files for timing analysis
- Must have 0 DRC violations
- Must maintain the assets
- Must maintain functional equivalence to the original design

Challenge for Defenses and Attacks: DRC Violations

- DRC violations are *expected,* for example for pin access around the power stripes
- They can become very challenging to manage for dense layouts, which is also part of the challenge put forward in this contest





Constraints II

- Cannot design custom cells
- Must maintain the general IO pin placement
- Cannot revise the metal layers/metal stack
- Must include a functional clock tree
- Must follow the PDN recipe provided in the reference flow

Trojan Insertion (Final Round)

There are 6 different Trojans per benchmark

Triggering conditions - activation mechanism

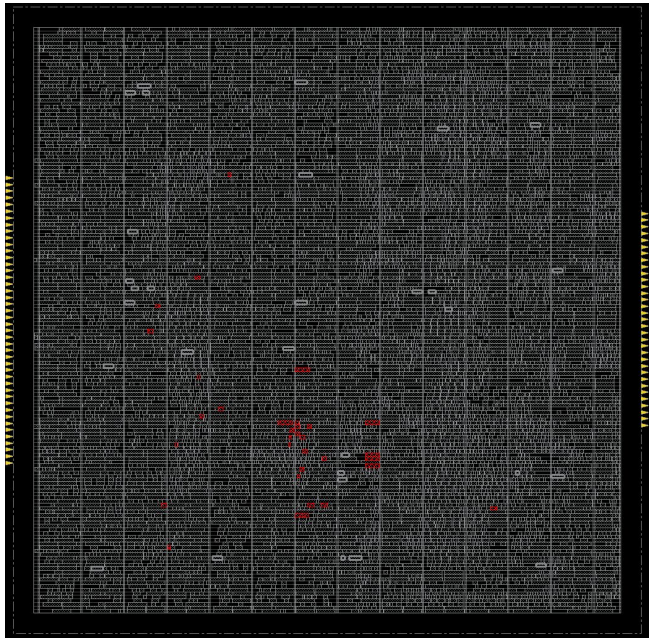
- Targeted: set of asset cells chosen from a common layout region
- Random: set of asset cells chosen randomly

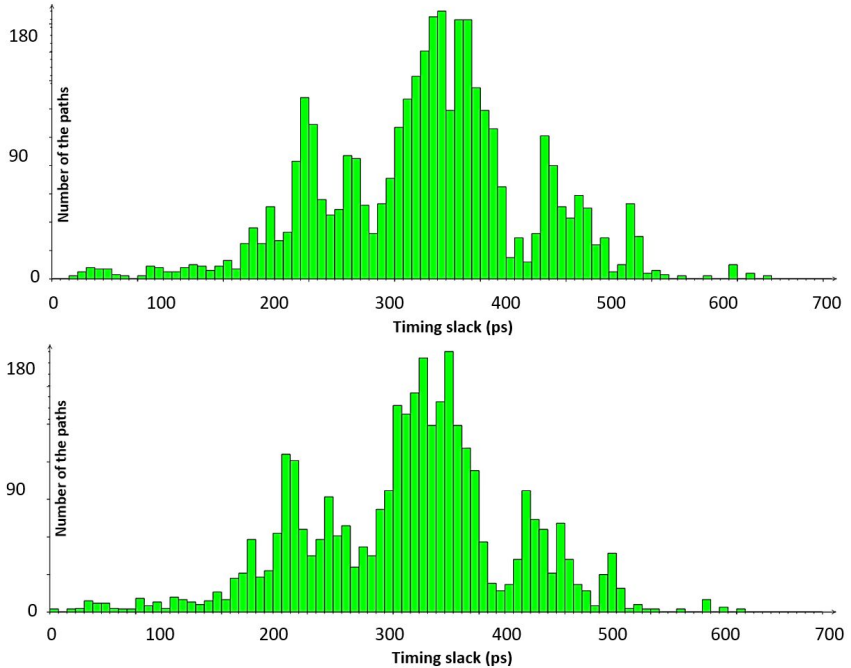
Trojan payload - effect of the Trojan

- Leak: Trojan is connected to the output of some sensitive asset FFs
- Modify: Trojan flips the output of some sensitive asset FFs
- Burn: Trojan adds a redundant FF chain / RO to consume more power

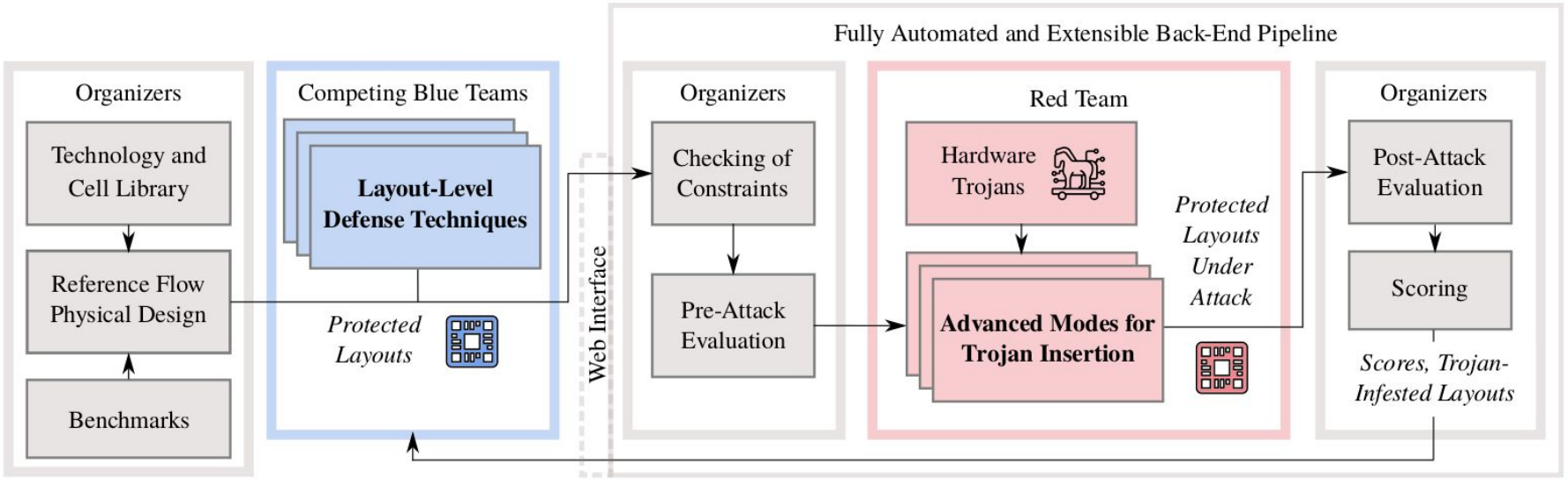
Trojan Insertion (Final Round)

An exemplary Trojan inserted into the SHA256 benchmark





Benchmarking Framework (Recap)



Defense Techniques

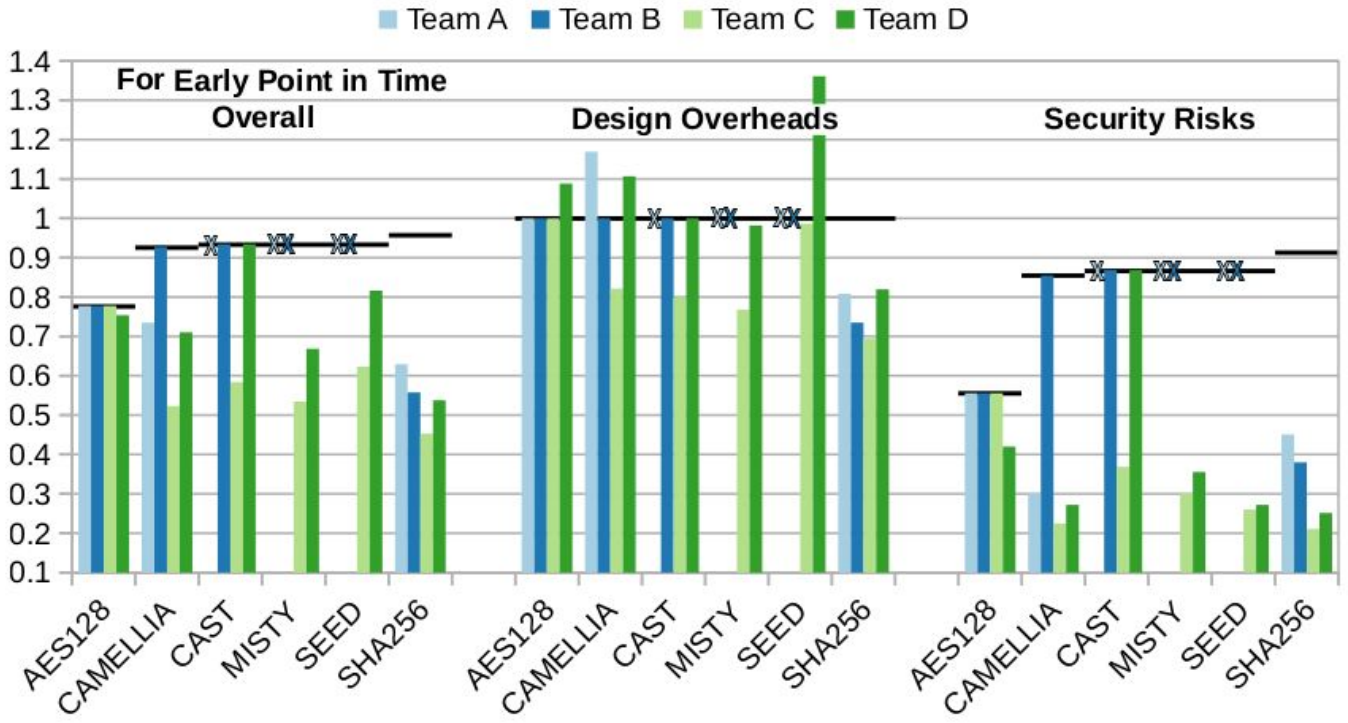
	Team			
	A	B	C	D
Shrink IC Outlines	✓	✓	✓	✓
Automated Parameter Tuning & Design-Space Exploration	✗	✓	✓	✗
Insertion of Functional Components	✓	✗	✗	✗
Insertion of Buffer Components	✓	✓	✗	✓
Insertion of Routing Detours	✓	✗	✗	✗
Re-Arrangement of Components	✓	✓	✓	✓

Note on final results: Team B 1st, Team C 2nd, Team A 3rd, Team D 4th

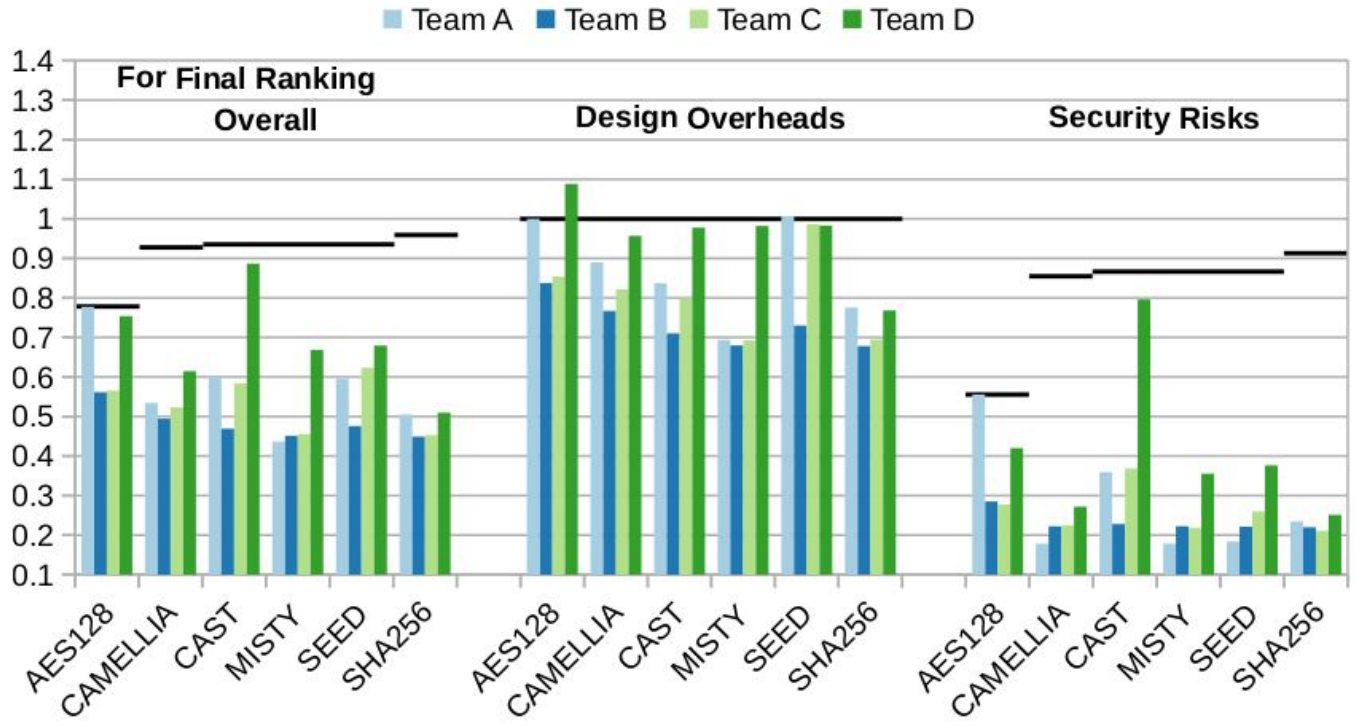
Submission Statistics

	Invalid Due to Violations For:						Valid
	Cell Assets	Functional Equivalence	Timing	DRCs	Additional Design & Technology Checks	Others	
Team A	0	1	6	10	97	7	54
Team B	0	0	4	15	48	1	37
Team C	3	1	0	5	25	12	94
Team D	0	0	3	6	3	3	62
Overall	3	2	13	36	173	23	247

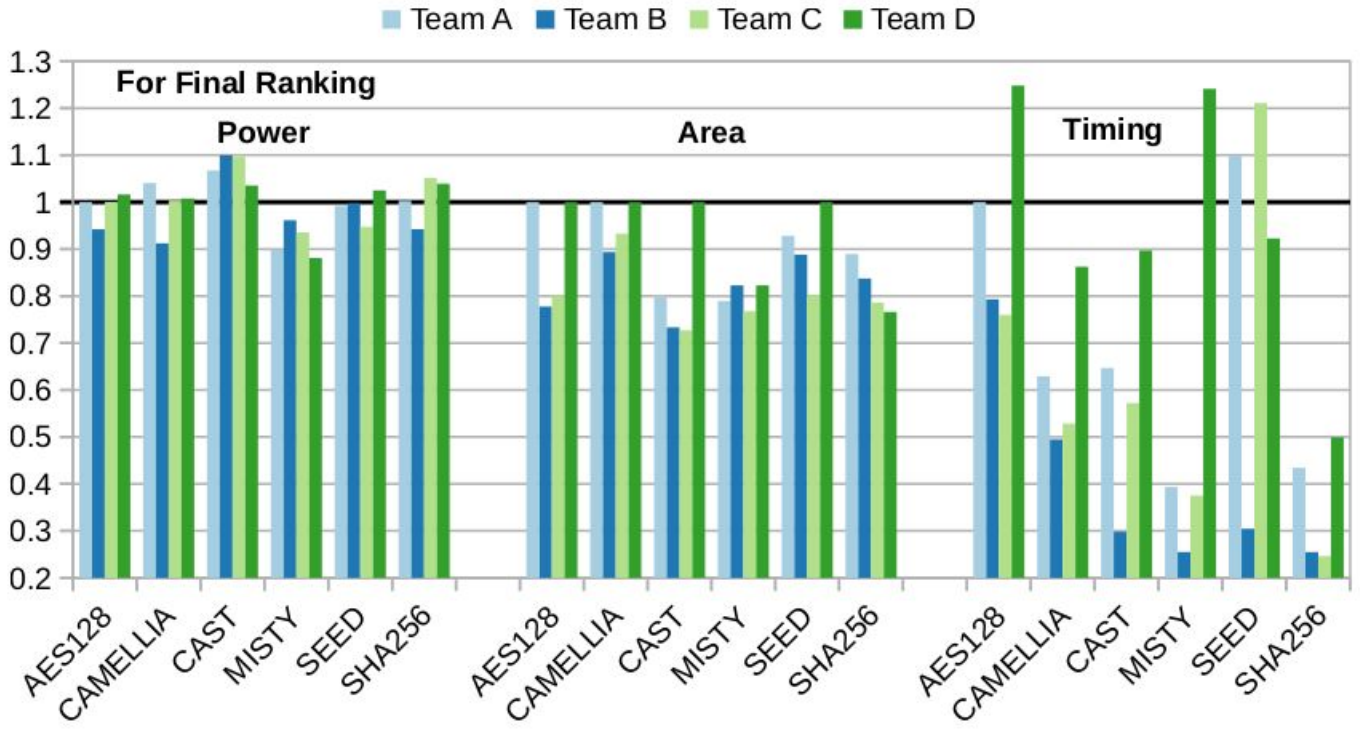
Results: Overview



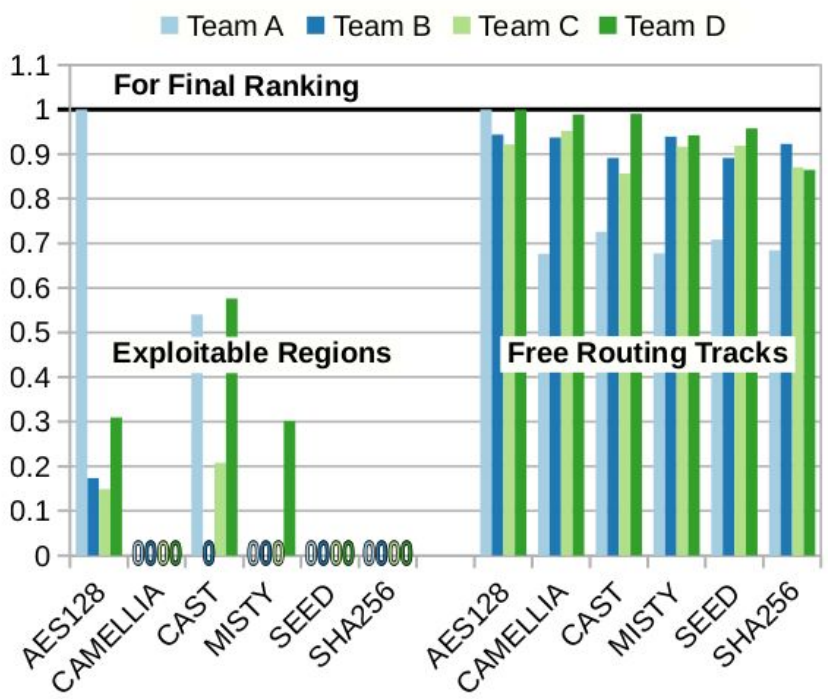
Results: Overview

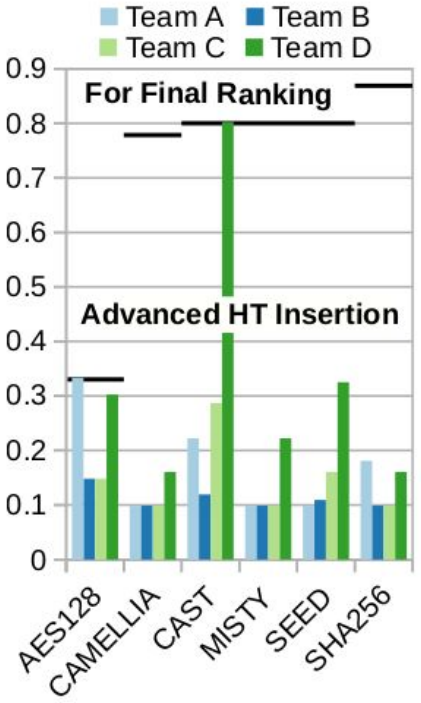


Results: Design

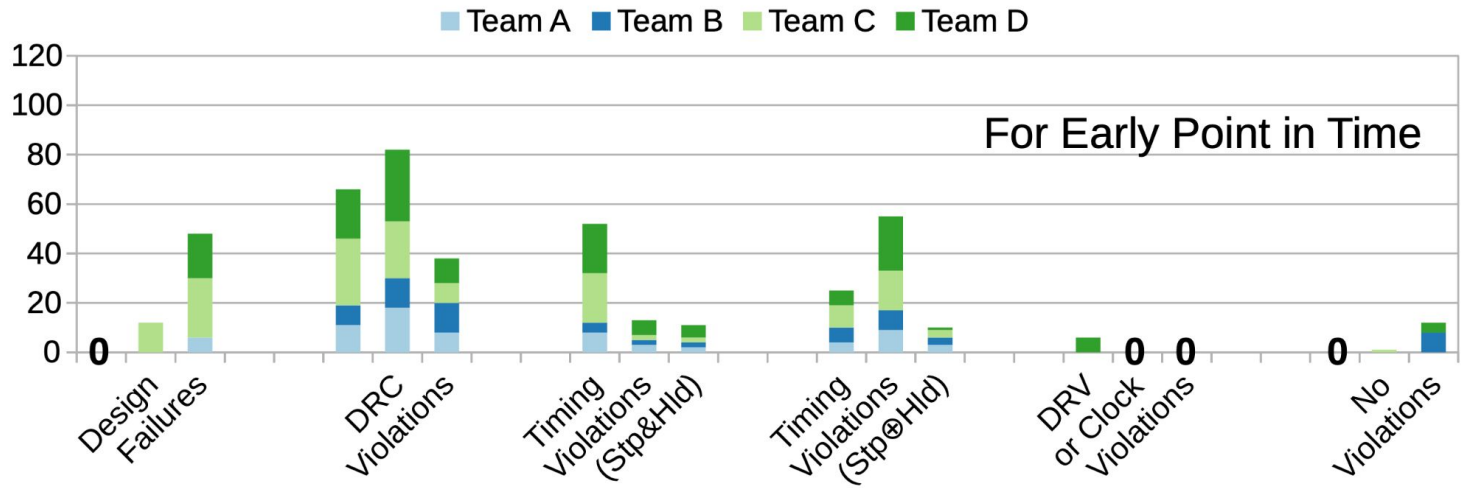


Results: Security



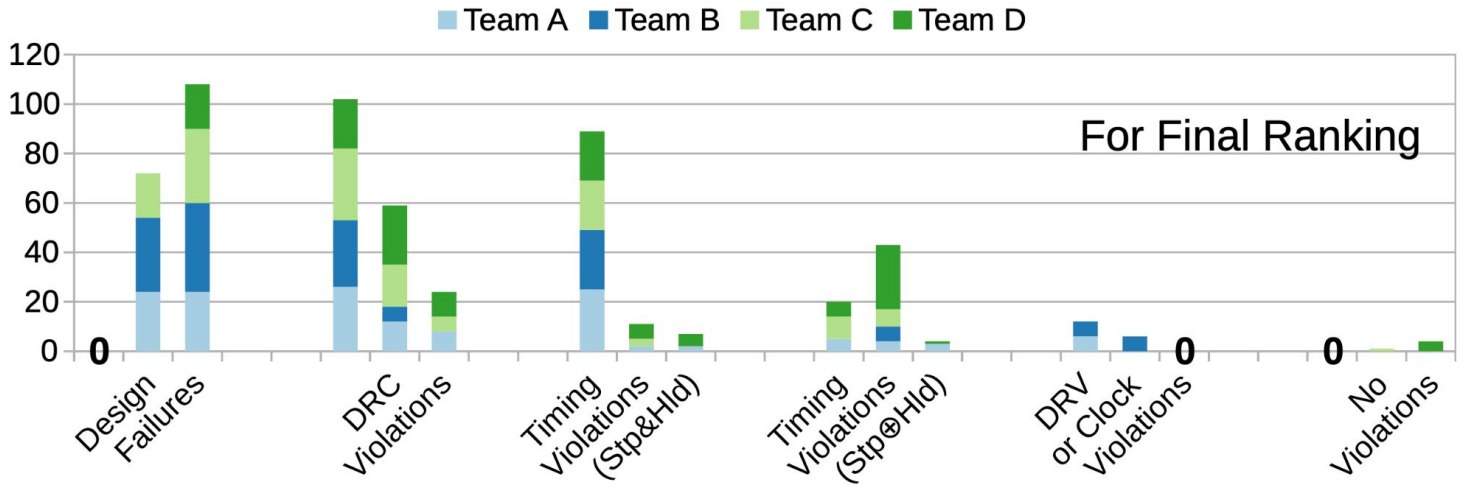


Progression of Defenses



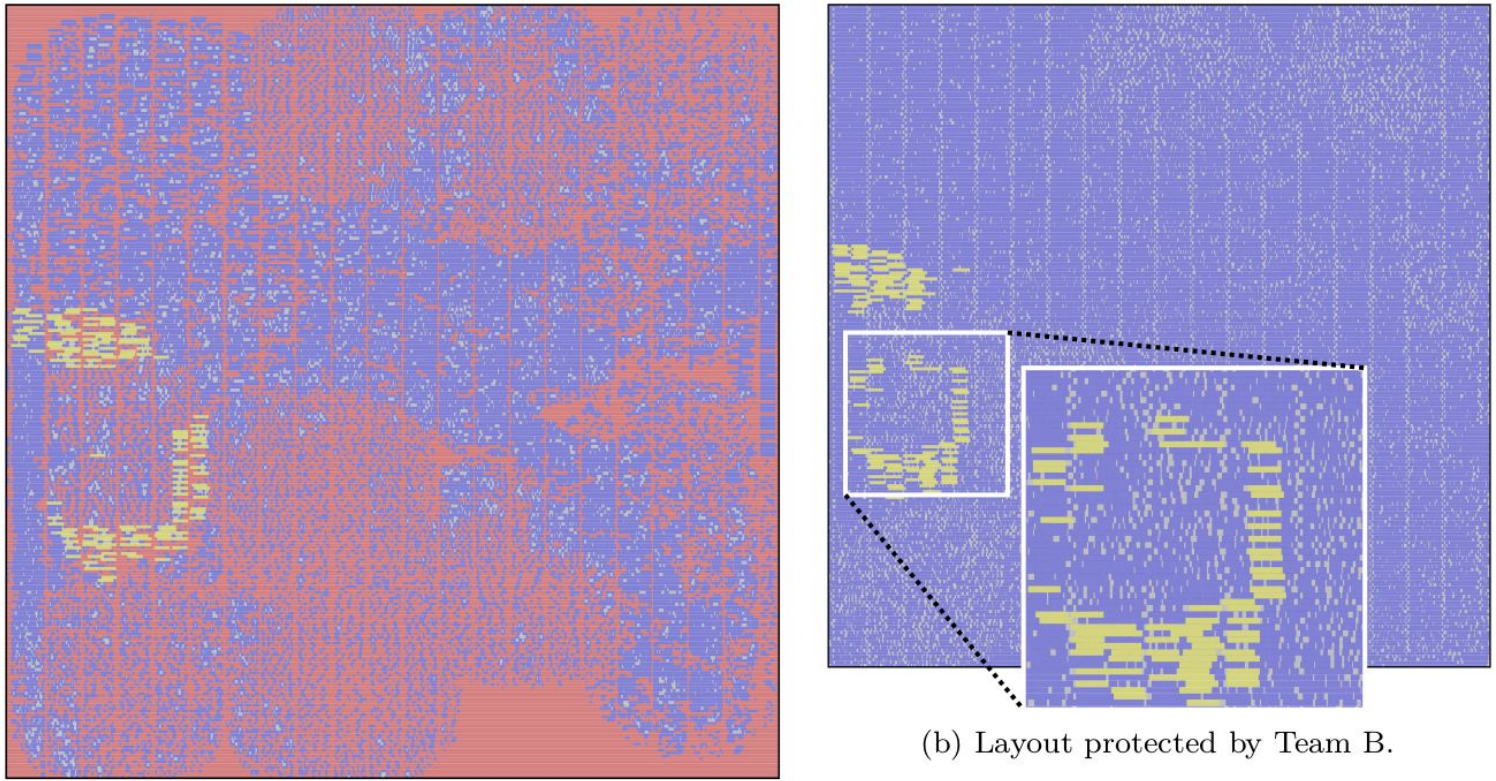
Note: For each category, from left to right: aggressive, moderate, conservative attack

Progression of Defenses



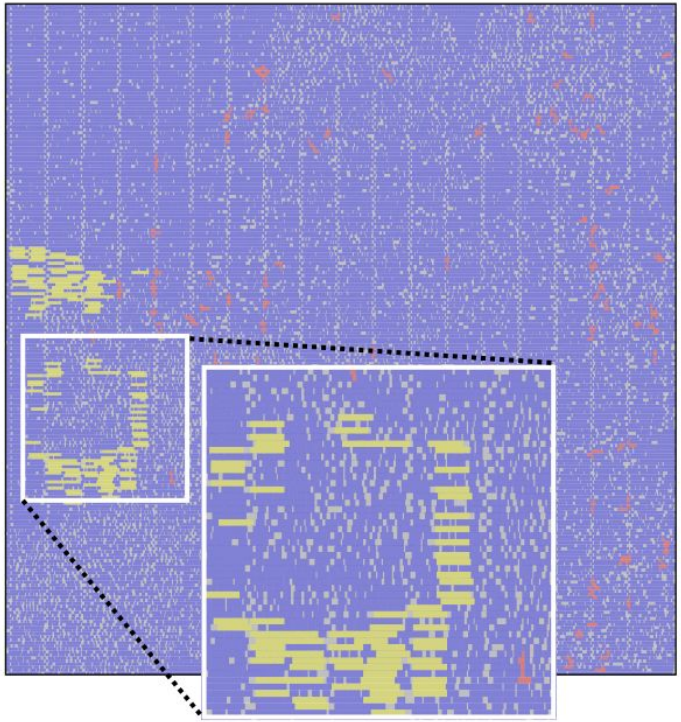
Note: For each category, from left to right: aggressive, moderate, conservative attack

Full Example

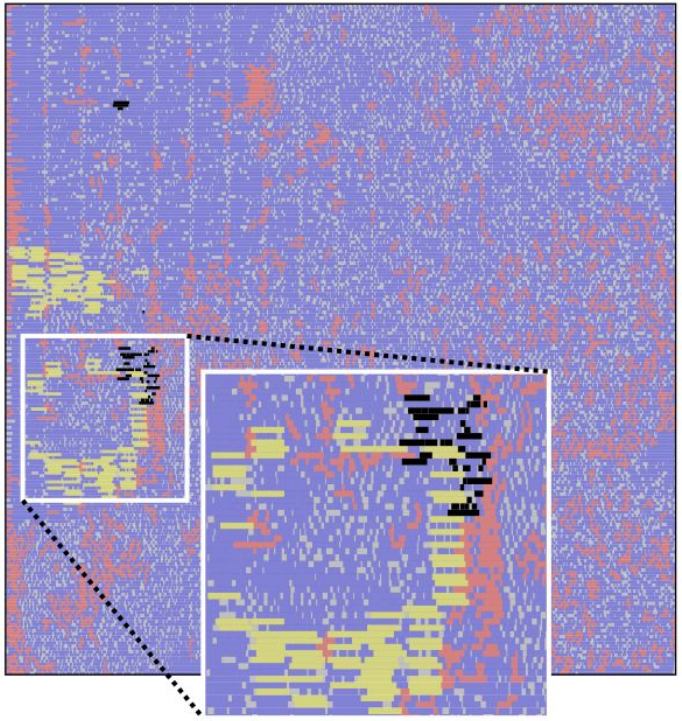


(b) Layout protected by Team B.

Full Example



(c) Protected layout under moderate attack.



(d) Protected layout under aggressive attack.

Winners

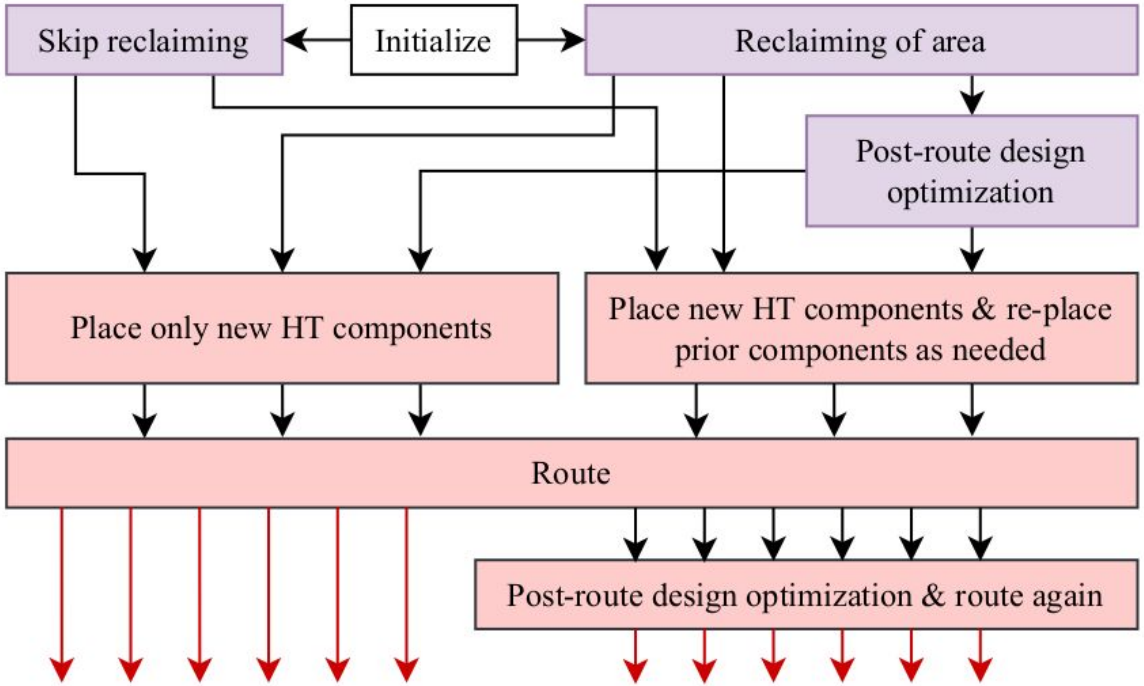
2nd Place: NTHU-TCLAB

1st Place: FDUEDA

3rd Place: CUEDA

4th Place: XDSecurity-II

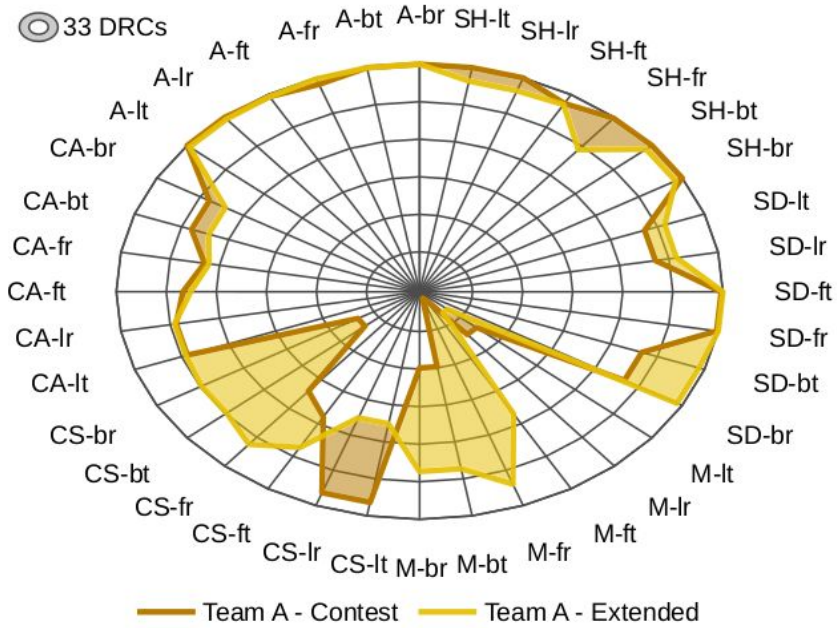
Post-Contest: Advanced Attacks

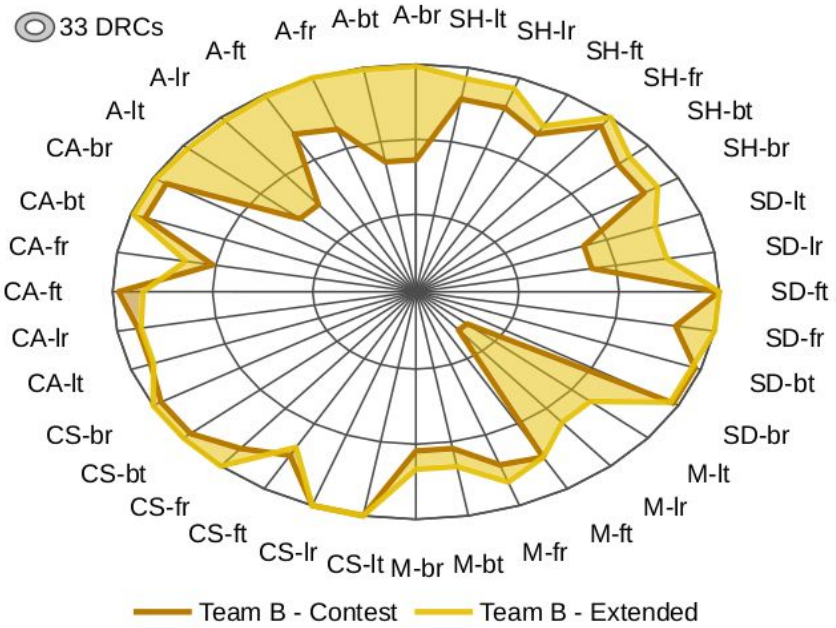


Post-Contest: Attack Results (w/o Manual Efforts)

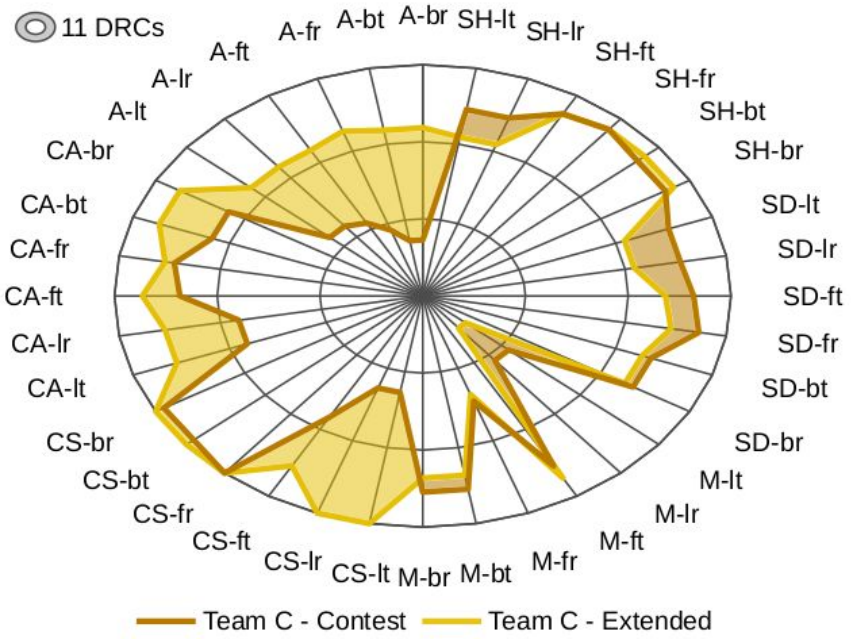
		Design Failures	DRC Vio.	Timing Vio. (Stp&Hld)	Timing Vio. (Stp⊕Hld)	DRV or Clock Vio.	No Vio.
Team A	AIC	0	28	21	11	6	0
	EXT	0	32	0	4	5	4
Team B	AIC	0	33	24	6	12	0
	EXT	0	27	0	2	0	9
Team C	AIC	0	35	20	12	0	1
	EXT	0	31	0	15	0	5
Overall	AIC	0	96	65	29	18	1
	EXT	0	90	0	21	5	18

Post-Contest: Improved DRC Handling for Attacks

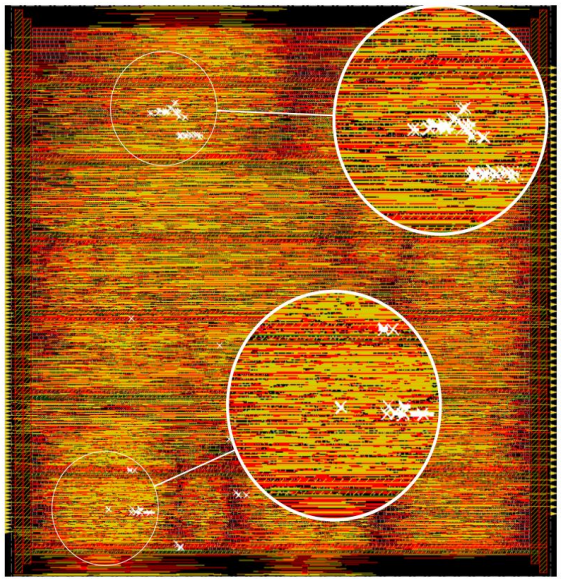




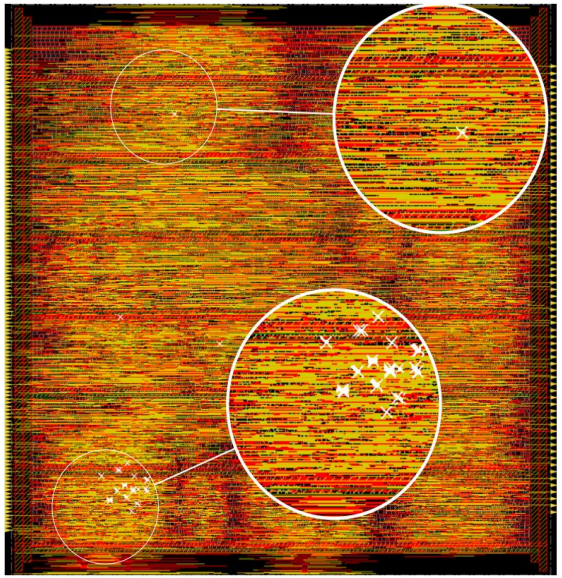
Post-Contest: Improved DRC Handling for Attacks



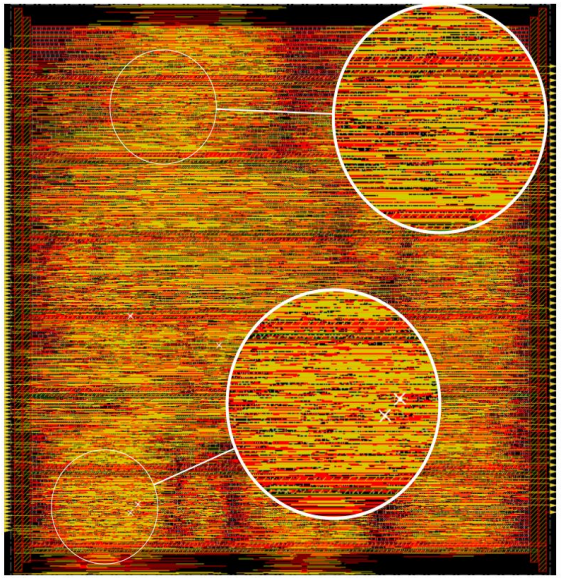
Post-Contest: Example for Manual DRC Closure



(a) Before fixing.



(b) After 1st round of fixing.



(c) After 2nd round of fixing.

Conclusions I

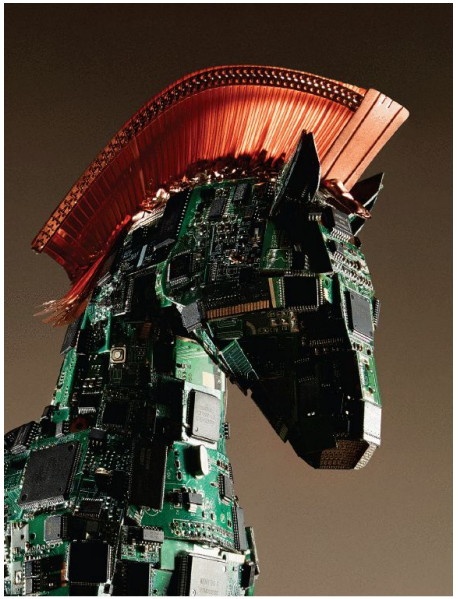
- Challenges for HT insertion are closely related to IC design, in more complex ways than prior art had recognized
- Regular, security-unaware IC design leave most layout resources exploitable
- Layout-level defenses are practical in general and, when done carefully, even without undermining design quality
- The (mis-)use of ECO techniques by the red team, an industry-wide standard for design modifications, is demonstrated as an effective and efficient attack approach

Conclusions II

- Motivation:
  - More and more threats are arising that affect hardware
  - **Build up knowledge and experience in CAD community**

- Congrats to all finalists! Really great efforts!

- Thanks to Samuel Pagliarini (CMU) and team!
- Thanks to everyone at ISPD committee for having us!



- https://wp.nyu.edu/ispd23_contest
- https://github.com/DfX-NYUAD/Trojan-Insertion-versus-Layout-Defenses

References

- Johann Knechtel, Jayanth Gopinath, Mohammed Ashraf, Jitendra Bhandari, Ozgur Sinanoglu, Ramesh Karri: Benchmarking Security Closure of Physical Layouts: ISPD 2022 Contest. ISPD 2022: 221-228
- Johann Knechtel, Mohammad Eslami, Peng Zou, Min Wei, Xingyu Tong, Binggang Qiu, Zhijie Cai, Guohao Chen, Benchao Zhu, Jiawei Li, Jun Yu, Jianli Chen, Chun-Wei Chiu, Min-Feng Hsieh, Chia-Hsiu Ou, Ting-Chi Wang, Bangqi Fu, Qijing Wang, Yang Sun, Qin Luo, Anthony W. H. Lau, Fangzhou Wang, Evangeline F. Y. Young, Shunyang Bi, Guangxin Guo, Haonan Wu, Zhengguang Tang, Hailong You, Cong Li, Ramesh Karri, Ozgur Sinanoglu, Samuel Pagliarini: Trojan Insertion versus Layout Defenses for Modern ICs: Red-versus-Blue Teaming in a Competitive Community Effort. (2024). IACR Transactions on Cryptographic Hardware and Embedded Systems, 2025(1), 37-77.
- Mohammad Eslami, Johann Knechtel, Ozgur Sinanoglu, Ramesh Karri, Samuel Pagliarini: Benchmarking Advanced Security Closure of Physical Layouts: ISPD 2023 Contest. ISPD 2023: 256-264
- Fangzhou Wang, Qijing Wang, Bangqi Fu, Shui Jiang, Xiaopeng Zhang, Lilas Alrahis, Ozgur Sinanoglu, Johann Knechtel, Tsung-Yi Ho, Evangeline F. Y. Young: Security Closure of IC Layouts Against Hardware Trojans. ISPD 2023: 229-237
- Guangxin Guo, Hailong You, Zhengguang Tang, Benzheng Li, Cong Li, Xiaojue Zhang: ASSURER: A PPA-friendly Security Closure Framework for Physical Design. ASPDAC 2023
- Jhih-Wei Hsu, Kuan-Cheng Chen, Yan-Syuan Chen, Yu-Hsiang Lo, Yao-Wen Chang: Security-aware Physical Design against Trojan Insertion, Frontside Probing, and Fault Injection Attacks. ISPD 2023
- Mohammad Eslami, Tiago Perez, Samuel Pagliarini: SALSy: Security-Aware Layout Synthesis. arXiv:2308.06201