Cost-efficient 3D Integration to Hinder Reverse Engineering During and After Manufacturing

Peng Gu*, Dylan Stow, Prashansa Mukim, Shuangchen Li and Yuan Xie*

Electrical and Computer Engineering Department

University of California Santa Barbara, CA, USA

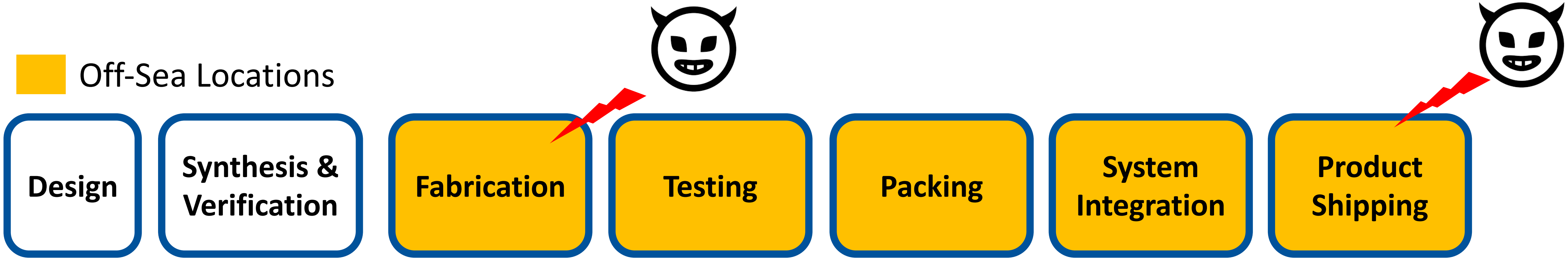
Outline

- Motivation & Background
  - Global Semiconductor Supply Chain Challenge
  - Existing Protection Mechanism
  - 3D Integration
- Key Idea
- Secure Min-Cutsize Partition Algorithm
- Secure 3D split-fab design flow
- Evaluation

Outline

- Motivation & Background
  - Global Semiconductor Supply Chain Challenge
  - Existing Protection Mechanism
  - 3D Integration
- Key Idea
- Secure Min-Cutsize Partition Algorithm
- Secure 3D split-fab design flow
- Evaluation

Global Semiconductor Supply Chain Challenge



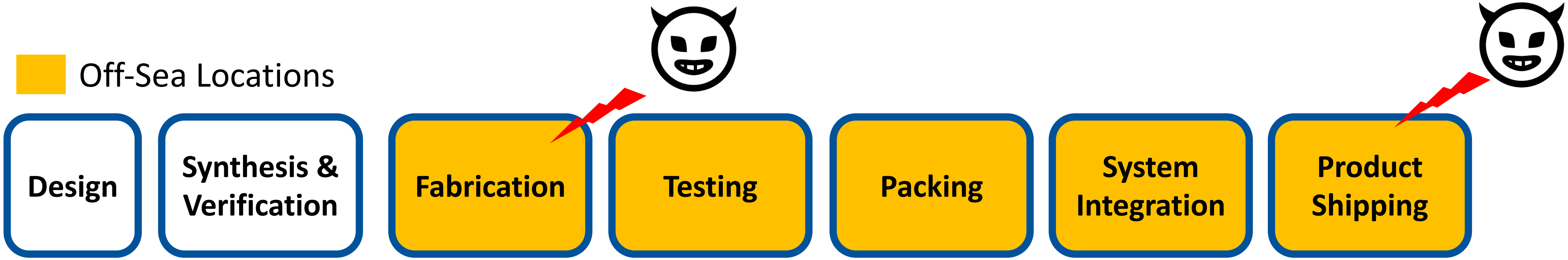
🟨 Off-Sea Locations

Design → Synthesis & Verification → Fabrication → Testing → Packing → System Integration → Product Shipping

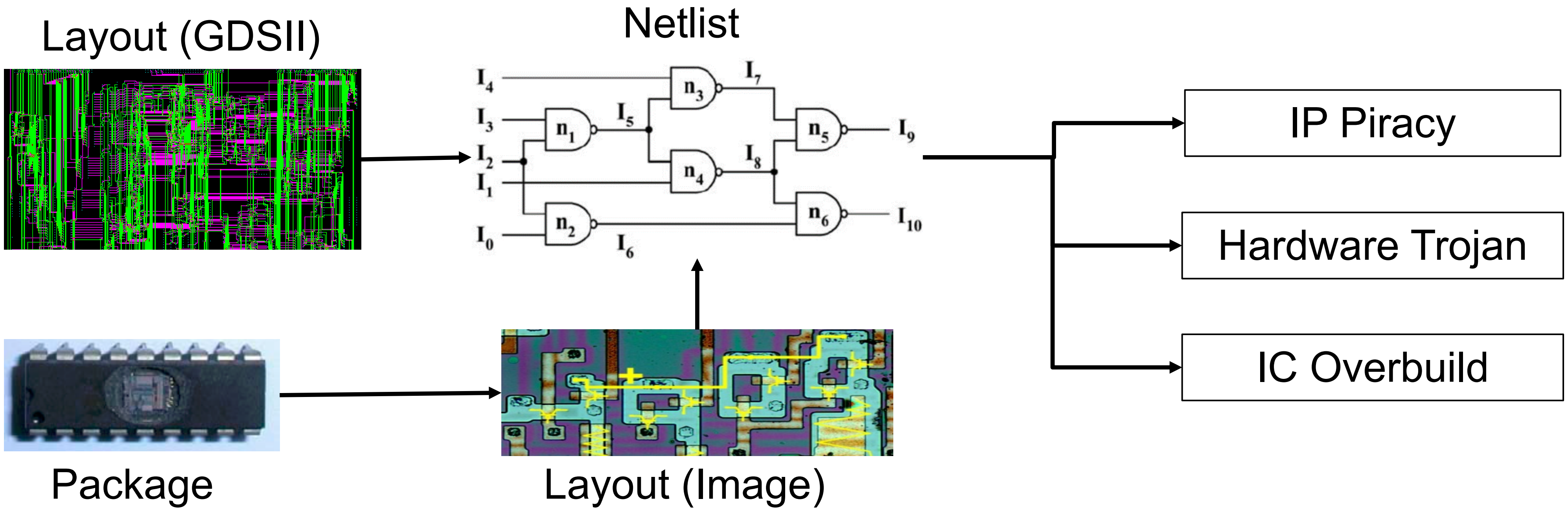
- Reverse engineering has become a serious threat
  - **<u>During fabrication:</u>**
    - These potentially malicious foundries can learn the functionality of the outsourced designs by reverse engineering layout files.
  - **<u>After manufacturing:</u>**
    - Adversaries may also acquire the whole chip and learn the layout and circuit netlist through chip delayering, imaging, probing, and netlist extraction.

Global Semiconductor Supply Chain Challenge

🟨 Off-Sea Locations



Design → Synthesis & Verification → Fabrication → Testing → Packing → System Integration → Product Shipping



Layout (GDSII)

Netlist

Package

Layout (Image)

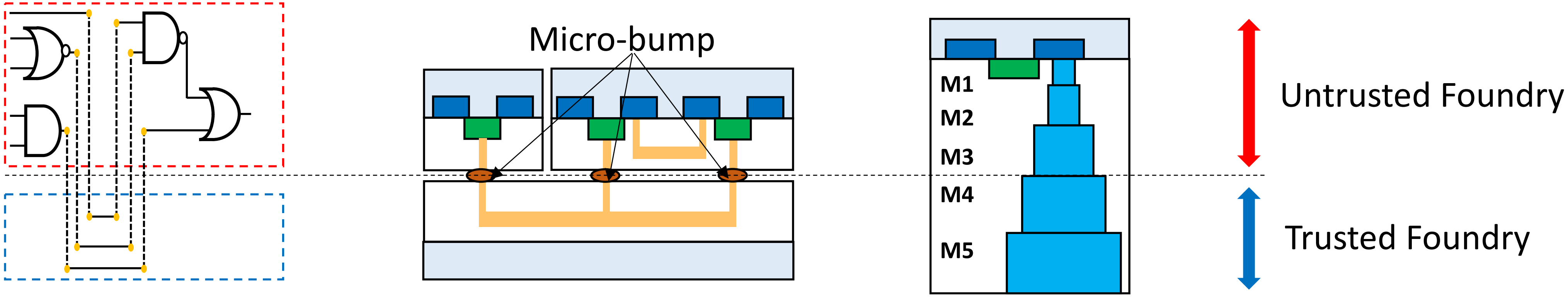
IP Piracy

Hardware Trojan

IC Overbuild

Existing Protection Schemes

- Split-Manufacturing
  - Provide “During Manufacturing” Protection



Netlist Partition

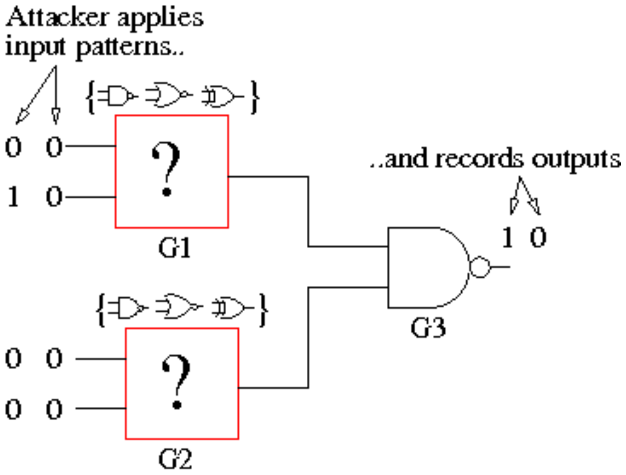
2.5D Passive Interposer Split-Fab [M. Jagasivamani 2014]

FEOL/BEOL Split-Fab [Y. Xie 2015]

- Limitations
  - Metal wires in the trusted tier are easy to be reverse-engineered after adversaries acquire the final product.
  - Large cutsize overhead and camouflaged routing overhead for 2.5D interposer split-fab.
  - Technology gap between available trusted and untrusted processes for FEOL/BEOL split-fab.

Existing Protection Schemes

- Circuit Camouflaging
  - Provide “After Manufacturing” Protection

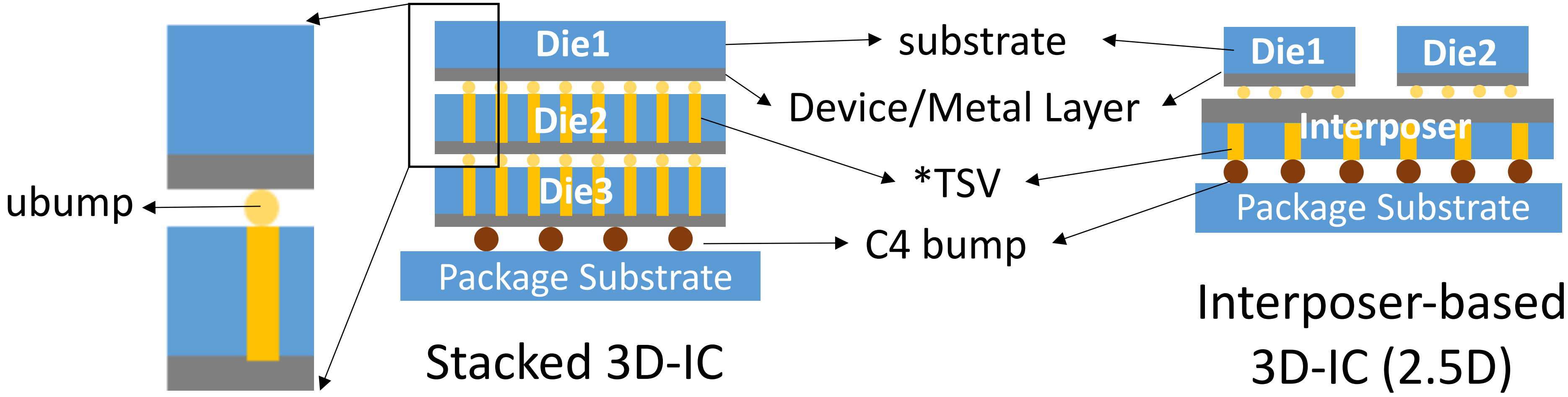


Gate Camouflaging
[J. Rajendran 2013]

- Limitations
  - Ineffective during manufacturing, since the untrusted foundries require the very detailed layout information to fabricate the circuit.

3D/2.5D Integration Fundamentals

- 3D integration is a technology that enables heterogeneous stacking of multiple dies in vertical dimension, connected by Through-Silicon-Vias (TSVs) and micro-bump (ubump).
  - 3D integration is already happening for HBM and HMC, and will be used for Intel’s next generation Feveros product

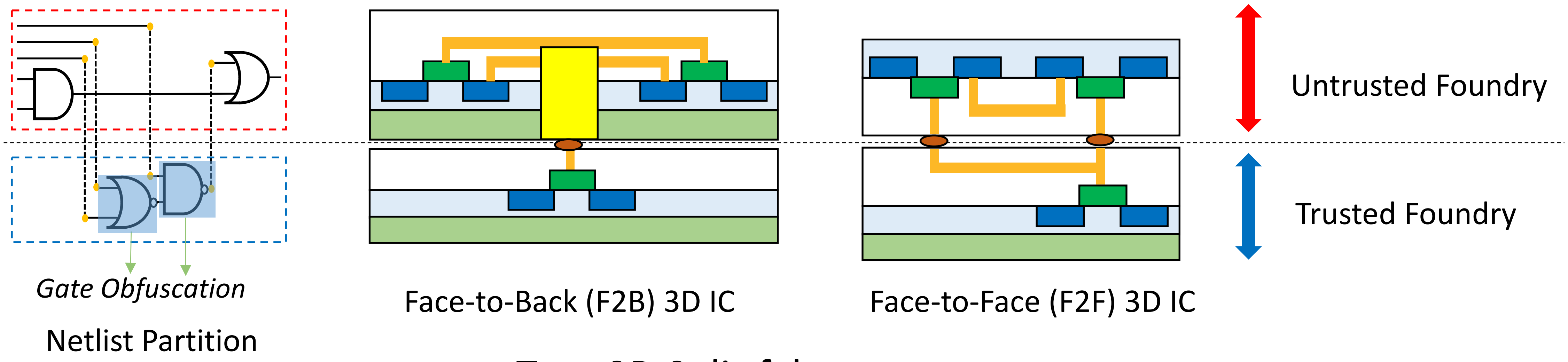


*TSV: Through-Silicon-Via

[P. Gu 2016]

Key Idea

- Use cost-efficient 3D integration to combine the concepts of **split fabrication** and **circuit camouflaging** so that IP is secured against reverse engineering attacks **during and after manufacturing**.



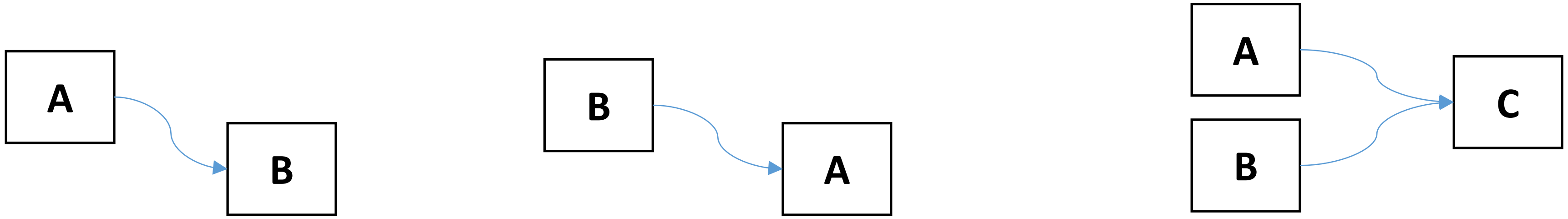
- True 3D Split-fab
- Circuit camouflaging on secure die
- Cost-effective & utilization of old technode

Outline

- Motivation & Background
  - Global Semiconductor Supply Chain Challenge
  - Existing Protection Mechanism
  - 3D Integration
- Key Idea
- Secure Min-Cutsize Partition Algorithm
- Secure 3D split-fab design flow
- Evaluation

Concept – Gate Interference

- If gate A is said to logically interfere with gate B, then either:
  - the inputs of A is on the output path of B, or if inputs of B is on the output path of A, **OR**
  - the primary output of A and B converges.

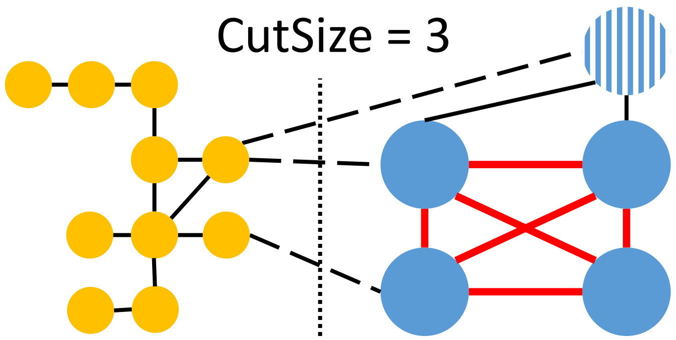


- To maximally enhance the effectiveness of circuit camouflaging, **the largest interference graph (theoretical maximum complexity)** is extracted from the original netlist, where **every gate in that graph is interfered with each other**. [J.Rajendran 2014]

Secure Min-Cutsize Partition Algorithm

GOAL: Maximize largest interference graph size

- Reduce partition cutsize
- Maintain partition ratio



CutSize = 3

Pratio = 1/3

N = 15

$C_{camouflaged} = C_{max\_clique} = 4$

$C_{untrusted} = 10$

$C_{trusted} = 5$

$$\eta_{se} = \frac{C_{camouflaged}}{C_{trusted}} = 80\%$$

Selection Efficiency

Given:

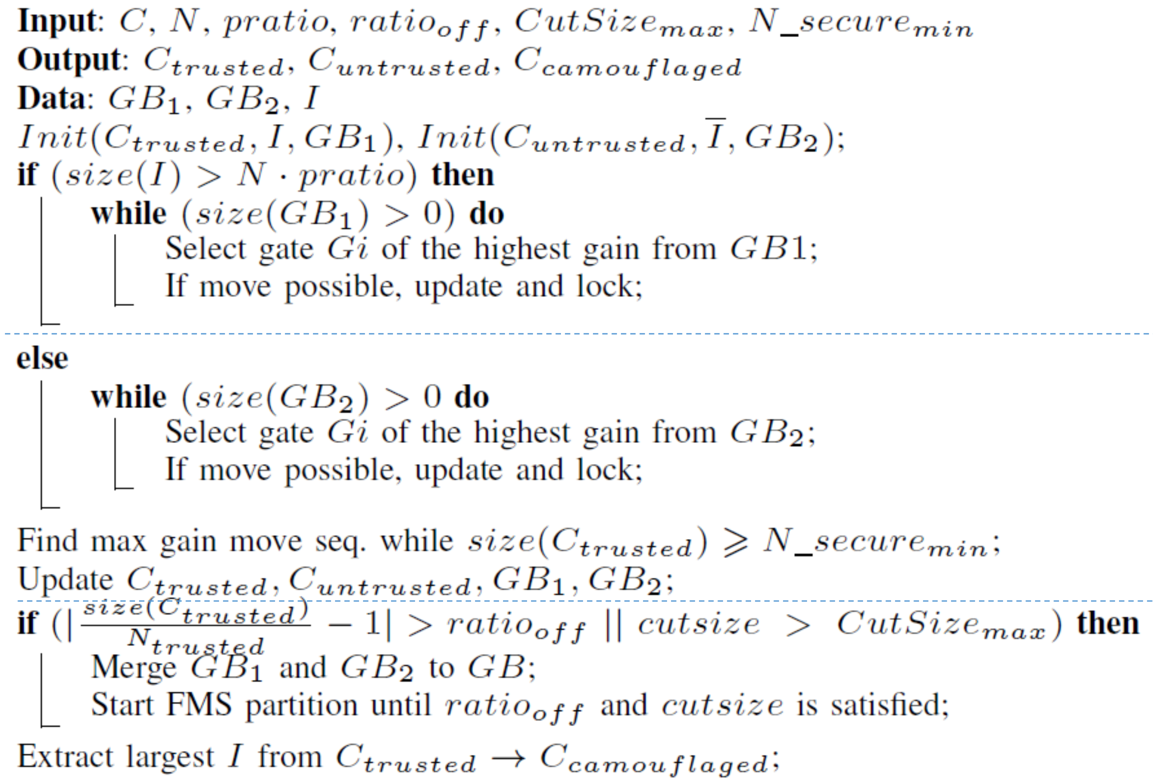
- the netlist of a circuit C with gate count N,
- partition ratio: $Pratio$,
- maximum cutsize: $CutSize_{max}$
- minimum number of fully interfered gates N_secure_{min}

Find partitions $C_{trusted}$, $C_{untrusted}$, and camouflaged gate list $C_{camouflaged}$

Secure Min-Cutsize Partition Algorithm

GOAL: Maximize largest interference graph size

- Reduce partition cutsize
- Maintain partition ratio



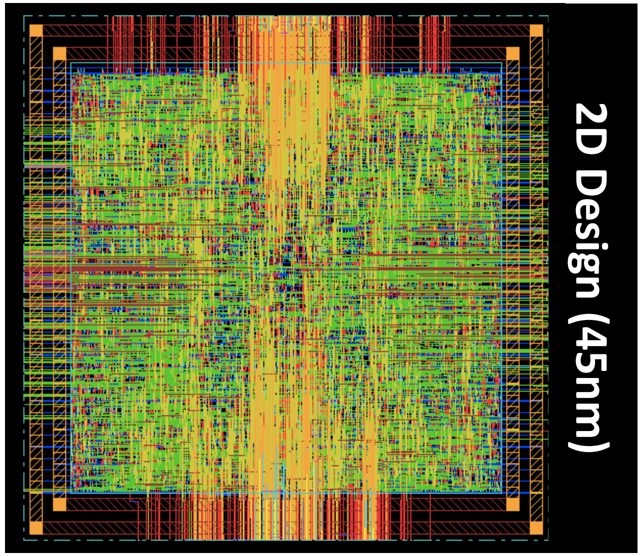
```
Input: $C$, $N$, $pratio$, $ratio_{off}$, $CutSize_{max}$, $N\_secure_{min}$
Output: $C_{trusted}$, $C_{untrusted}$, $C_{camouflaged}$
Data: $GB_1$, $GB_2$, $I$
$Init(C_{trusted}, I, GB_1)$, $Init(C_{untrusted}, \overline{I}, GB_2)$;
if $(size(I) > N \cdot pratio)$ then
    while $(size(GB_1) > 0)$ do
        Select gate $Gi$ of the highest gain from $GB1$;
        If move possible, update and lock;
else
    while $(size(GB_2) > 0$ do
        Select gate $Gi$ of the highest gain from $GB_2$;
        If move possible, update and lock;
Find max gain move seq. while $size(C_{trusted}) \geqslant N\_secure_{min}$;
Update $C_{trusted}, C_{untrusted}, GB_1, GB_2$;
if $(|\frac{size(C_{trusted})}{N_{trusted}} - 1| > ratio_{off} \,||\, cutsize > CutSize_{max})$ then
    Merge $GB_1$ and $GB_2$ to $GB$;
    Start FMS partition until $ratio_{off}$ and $cutsize$ is satisfied;
Extract largest $I$ from $C_{trusted} \rightarrow C_{camouflaged}$;
```

Partition Initialization

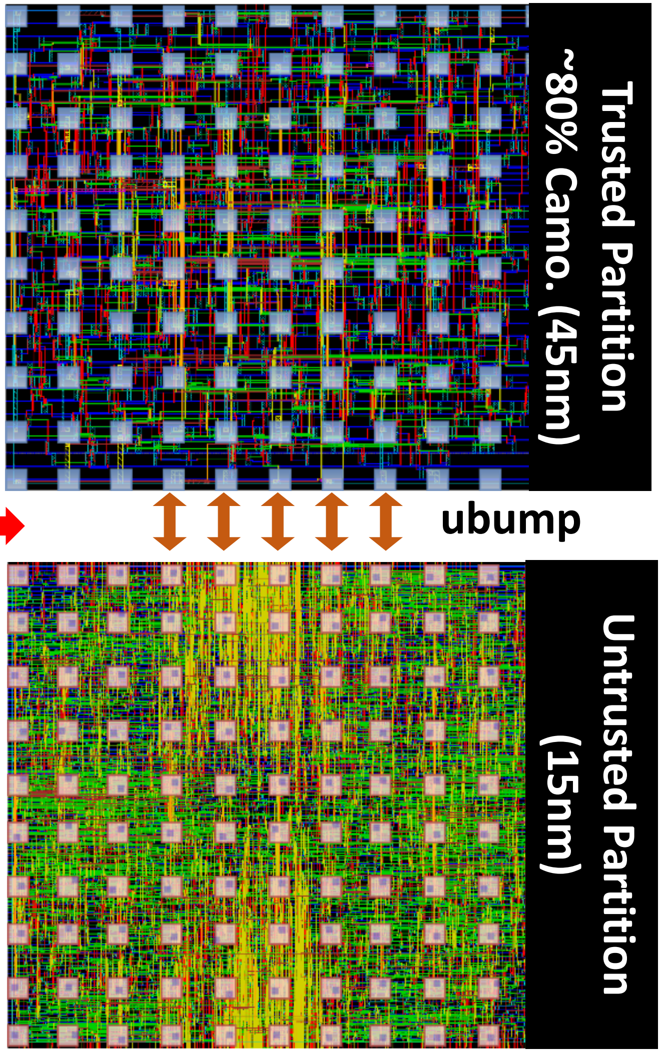
Unidirectional Gate Movement

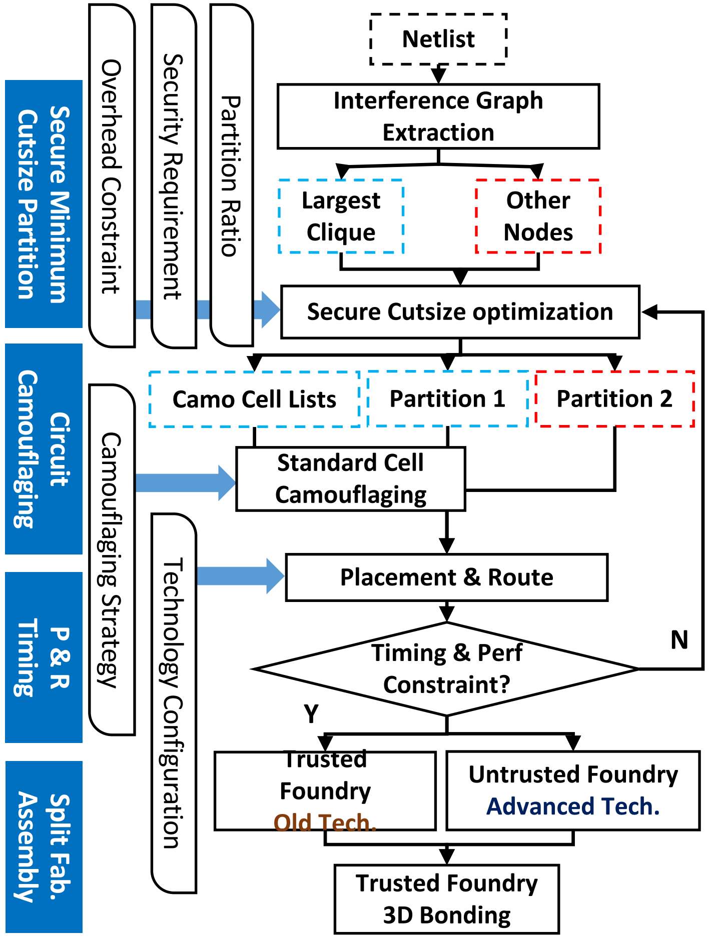
Bidirectional Gate Movement

3D split-fab design flow



Our work: 3D split-fab + gate camouflaging



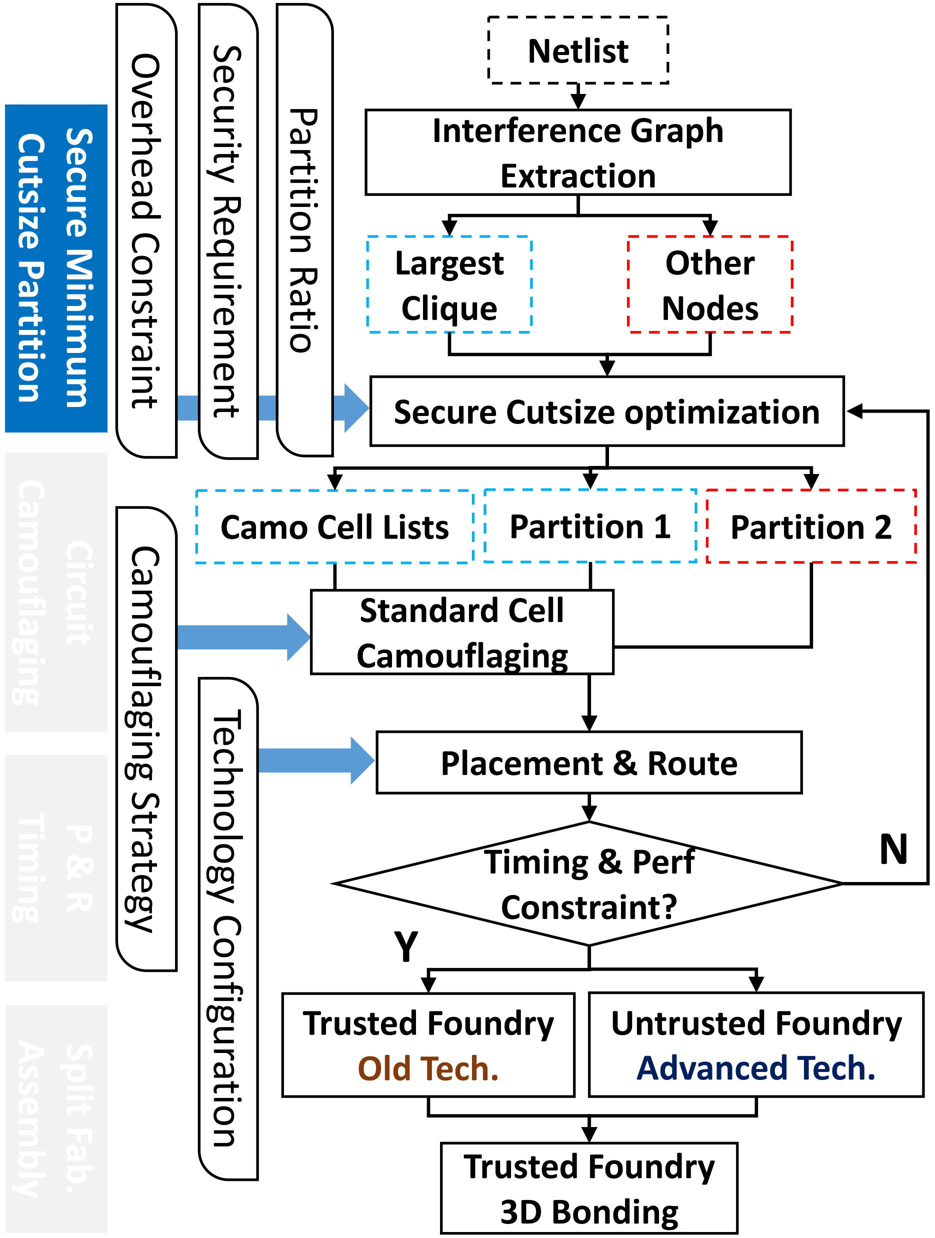


3D split-fab design flow

- Based on **gate interference**, the largest interference graph will be selected to form a clique.

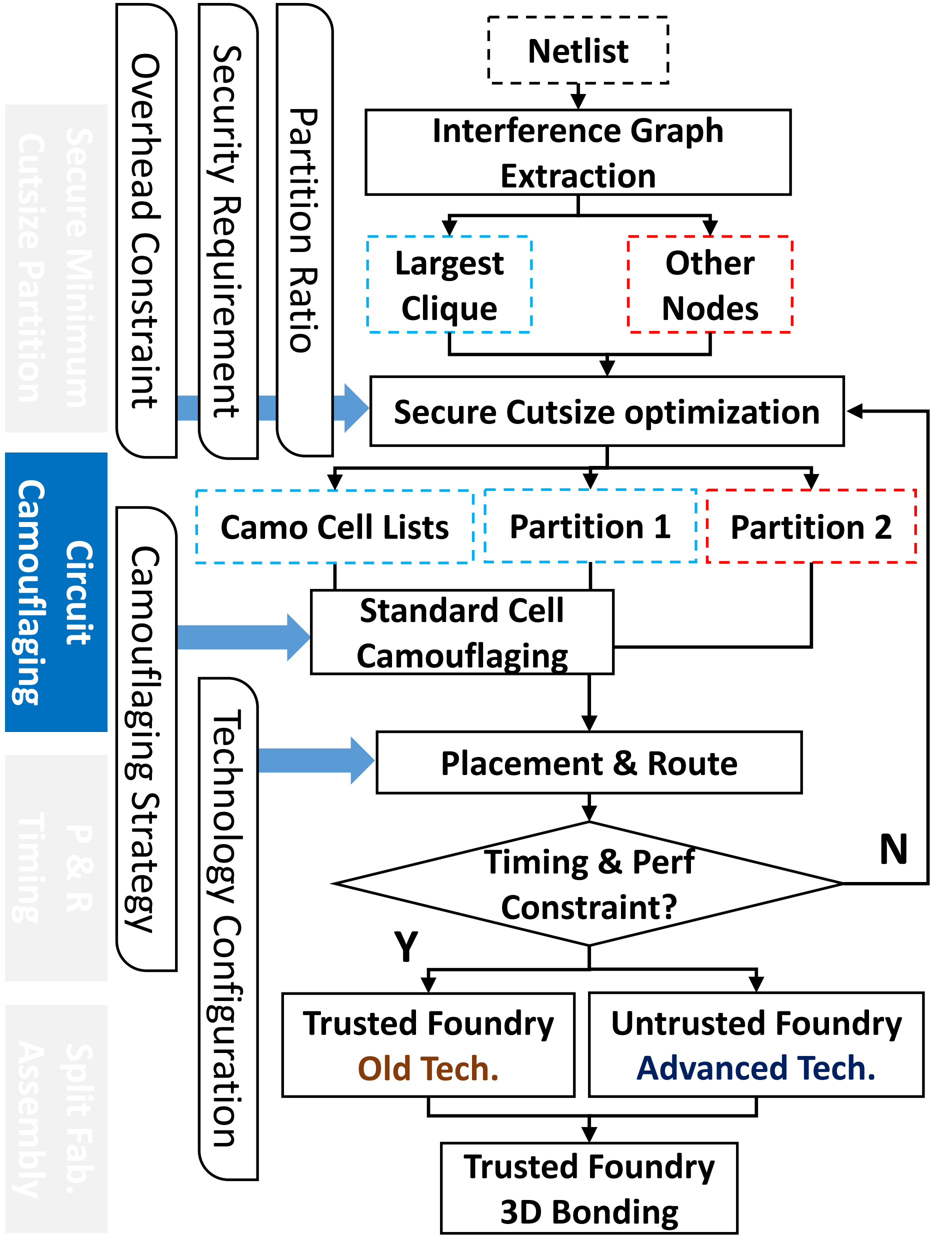
- Designer provides three parameters:
  - **Partition Ratio (**$Pratio$**)**, which is determined by the technology ratio used at trusted and untrusted die,
  - **Security Requirement (**N_secure_{min}**)**, which is the minimum number of fully interfered camouflaged gates that are placed on the trusted die, and
  - **Overhead Constraint (**$CutSize_{max}$**)**, which is the maximum partition cutsize allowed.

- **The security optimized min-cutsize algorithm** will use the largest clique to initialize the partition and optimize security and cutsize under the above constraints.



3D split-fab design flow

Partition 1 netlist ($C_{trusted}$) will be synthesized according to Camo Cell List ($C_{camouflaged}$) and the gate camouflaging strategy adopted by the trusted foundry.



Secure Minimum Cutsize Partition

Circuit Camouflaging

P & R Timing

Split Fab. Assembly

Overhead Constraint

Security Requirement

Partition Ratio

Camouflaging Strategy

Technology Configuration

Netlist

Interference Graph Extraction

Largest Clique

Other Nodes

Secure Cutsize optimization

Camo Cell Lists

Partition 1

Partition 2

Standard Cell Camouflaging

Placement & Route

Timing & Perf Constraint?

N

Y

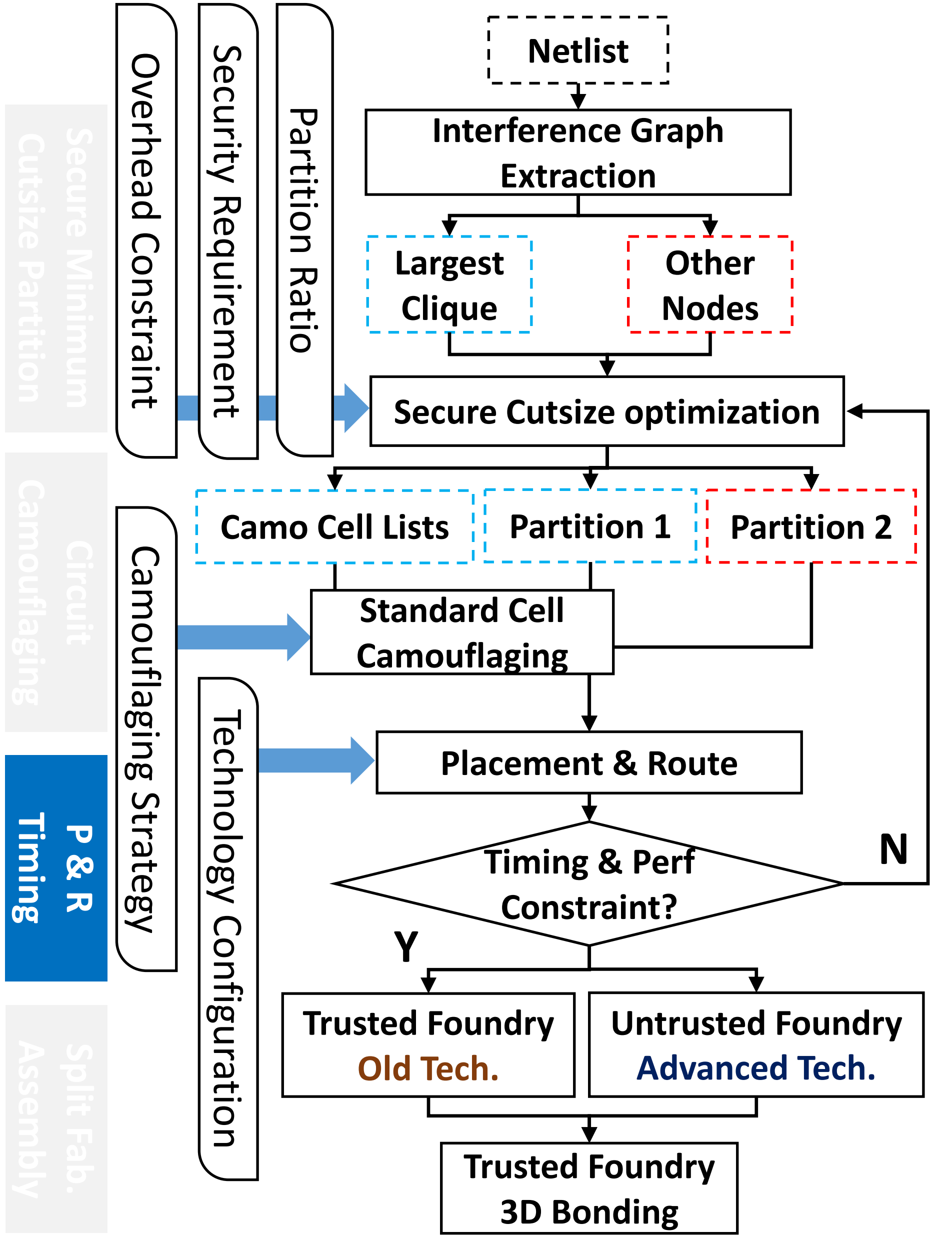
Trusted Foundry Old Tech.

Untrusted Foundry Advanced Tech.

Trusted Foundry 3D Bonding

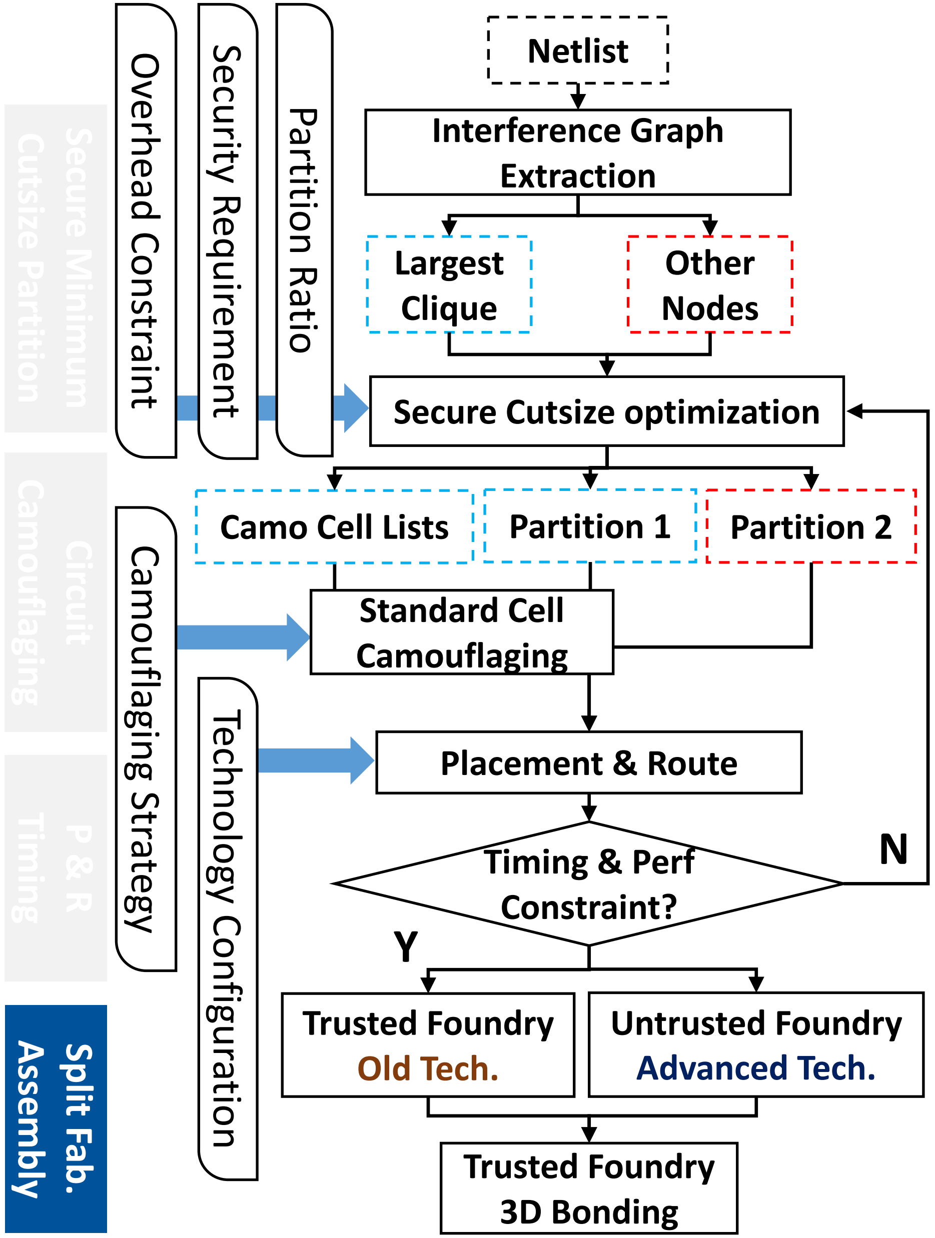
3D split-fab design flow

If the timing and performance of the wire length optimized placement and routing cannot be satisfied, then *Pratio* and N_secure_{min} will be relaxed in the first stage to re-generate the partition. This process will loop until a satisfying partition is achieved.



3D split-fab design flow

The final split fabrication is carried out and assembly as well as testing will be done in the trusted foundry



Outline

- Motivation & Background
  - Global Semiconductor Supply Chain Challenge
  - Existing Protection Mechanism
  - 3D Integration
- Key Idea
- Secure Min-Cutsize Partition Algorithm
- Secure 3D split-fab design flow
- Evaluation

Evaluation

- Evaluate the effectiveness against
  - proximity attacks **during manufacturing**
  - brute-force circuit decamouflaging attacks **after product shipping**
- 6 benchmarks from ISCAS’85 and ITC’99 (under different pratio)
  - Use FMS partitioning tool
  - Modify automatic pattern generation tool to find largest interference graph
- Area evaluation
- Cost evaluation

32nm/16nm, pratio=0.2
45nm/16nm, pratio=0.1
65nm/16nm, pratio=0.057
90nm/16nm, pratio=0.03
180nm, pratio=0.5

Metrics

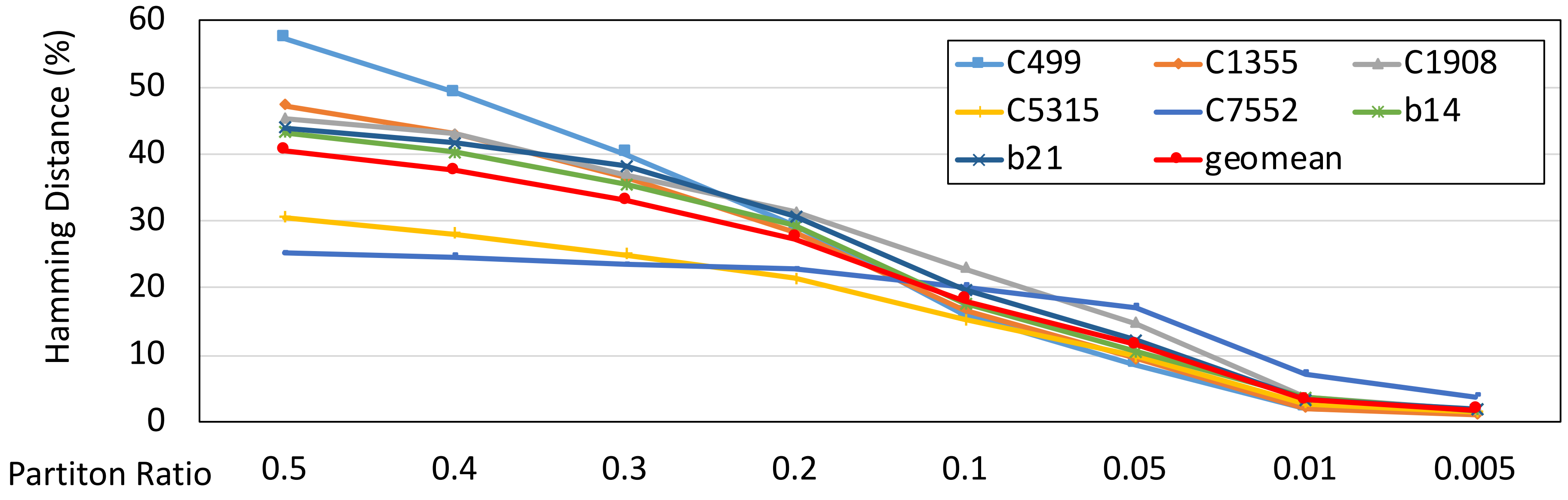
- Hamming distance:
  - A widely adopted metric to evaluate the protection against proximity attacks.
  - Given the same input vector, HD equals the normalized number of different output bits between the original netlist and the reconstructed netlist from the partial circuit.

$$HD(F, F') = \frac{1}{n} \sum_{x_i \in X} \frac{|F(x_i) - F'(x_i)|_{norm1}}{\#output\_bits}$$

- Complexity-to-Decamouflage (CtD):
  - the computational effort and the number of test patterns needed to learn the netlist using either brute force methods or SAT based attacks

$$\begin{aligned} CtD(F') = {} & log_{10}(min\{Brute\ Force\ Patterns, \\ & SAT\ Computation\ Steps + Query\ Patterns\}) \\ & \approx log_{10}(m^n) \end{aligned}$$

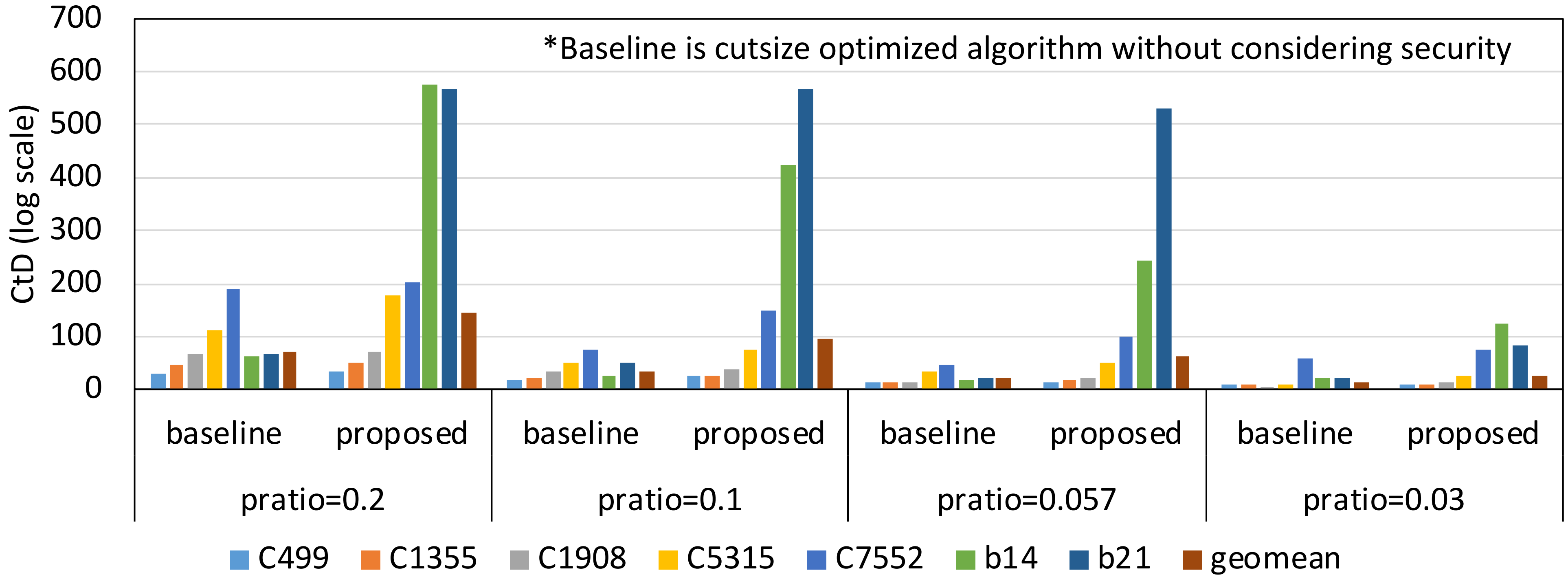
Proximity Attacks



Hamming Distance for different partition ratios

Reasonable split fabrication scheme between 32nm/15nm processes can achieve an average HD = 28% and an even split-fab ratio can have a very high average HD = 41%.

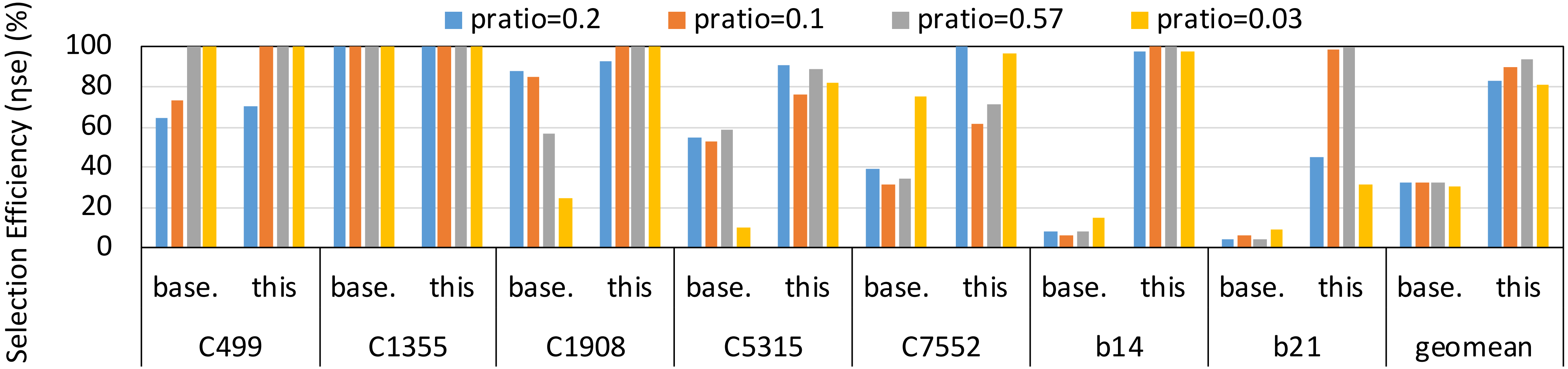
Brute-force-attack Complexity Comparison



- For small circuits (< 1000 gates), the improvement of CtD is not significant (~3 avg.) and for large circuits (> 10000 gates), the improvement of CtD is significant (~310 avg.).
- As partition ratio (pratio) becomes smaller (more advanced tech node and older tech node), our method shows more CtD improvements.

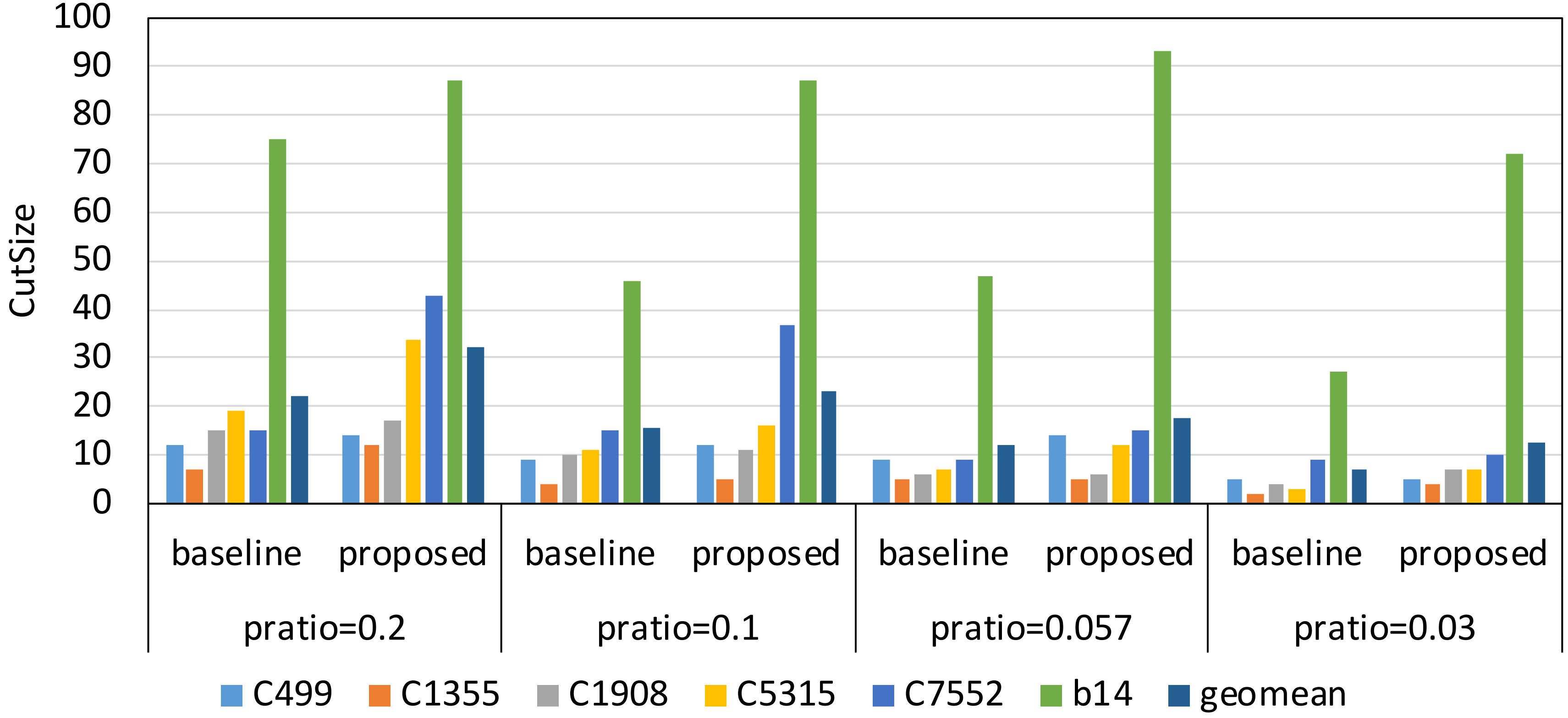
Partition Selection Efficiency

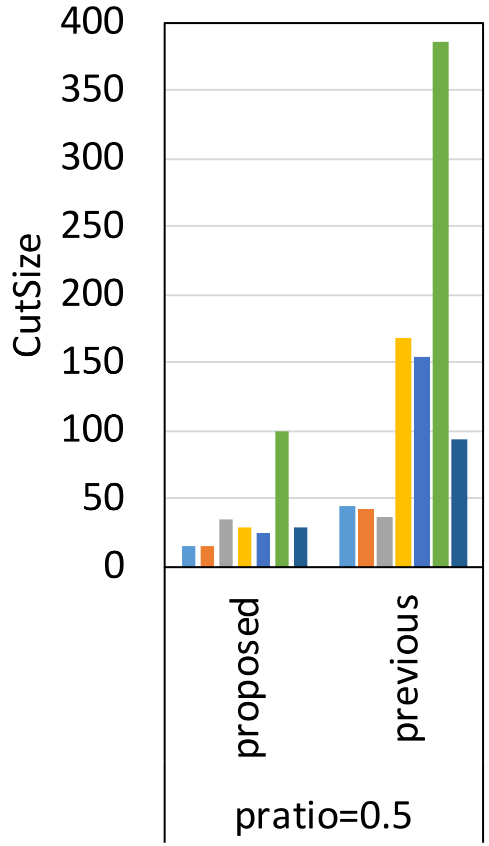
$$\eta_{se} = \frac{C_{camouflaged}}{C_{trusted}}$$



Compared with baseline, our proposed method can achieve higher selection efficiency (e.g. more gates on the trusted die are effectively camouflaged) for larger circuit benchmarks

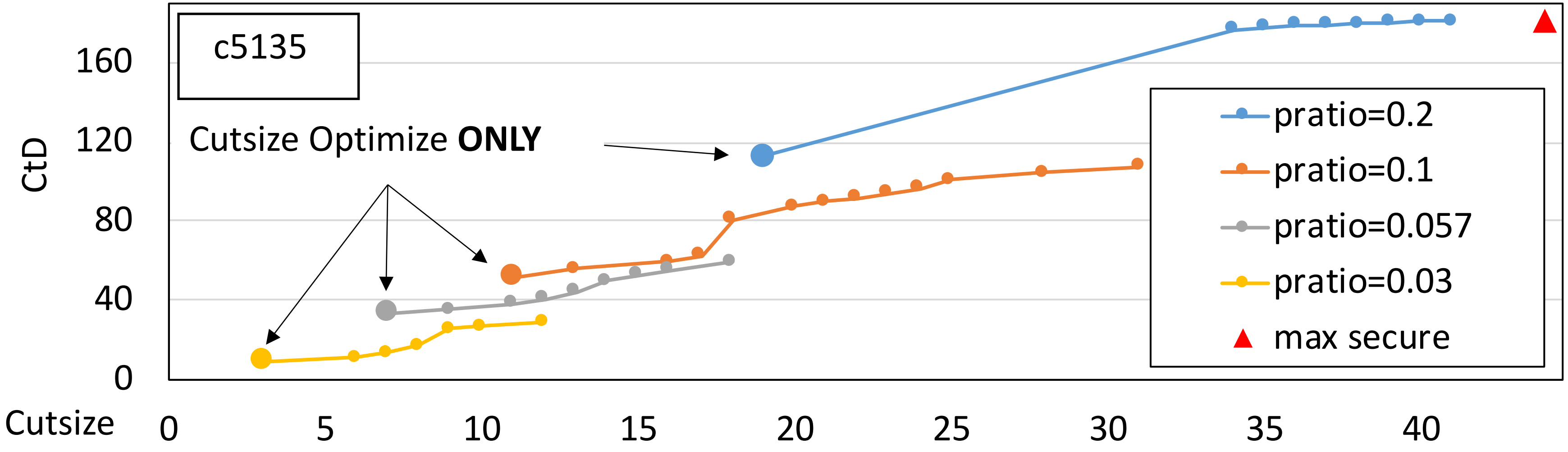
Cutsize Comparison





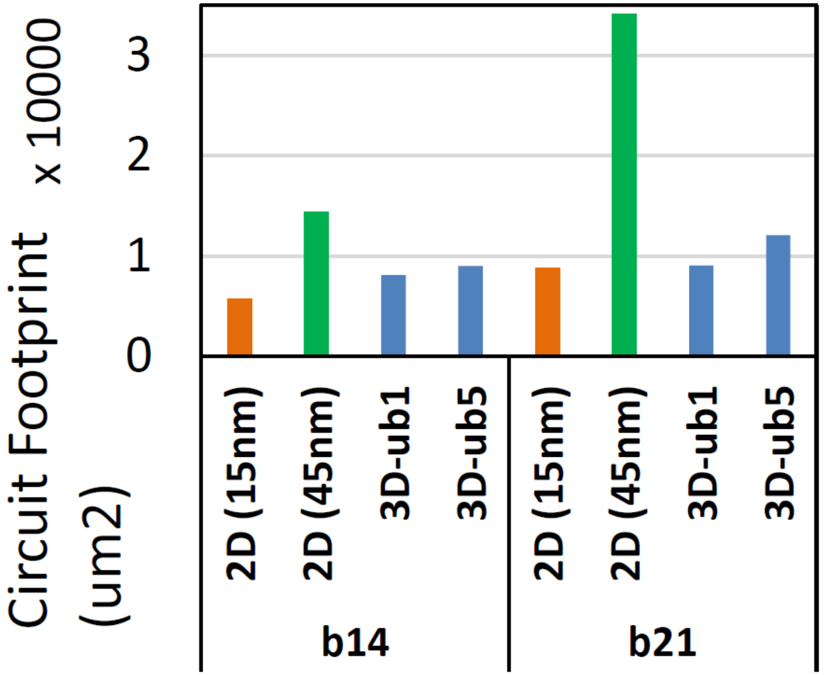
- Compared with baseline, the cutsize increase in our method is not significant (1.54X on average)
- Compared with previous 2.5D interposer based split-fab, we can achieve significantly lower cutsize (save 3.20X cutsize overhead)

Design Space Exploration



- The maximum security level (red triangle) can be achieved by putting the largest interference graph on the trusted tier with large cutsize overhead however.
- The proposed secure min-cutsize algorithm allows changing both $\boldsymbol{CutSize_{max}}$ and $\boldsymbol{N\_secure_{min}}$ to flexibly explore the design space.

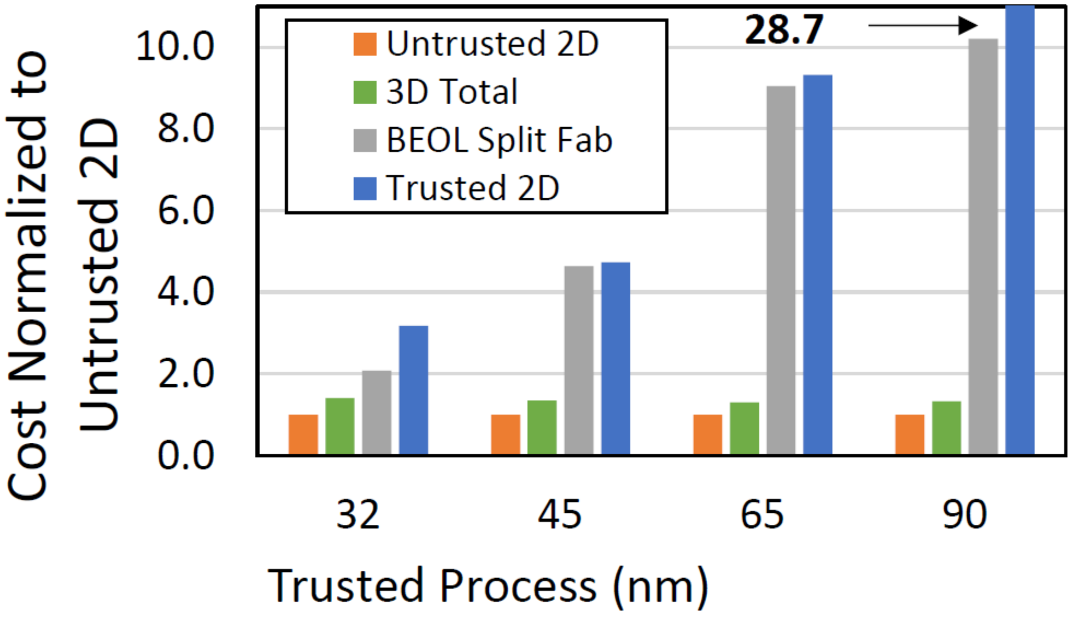
Area Evaluation

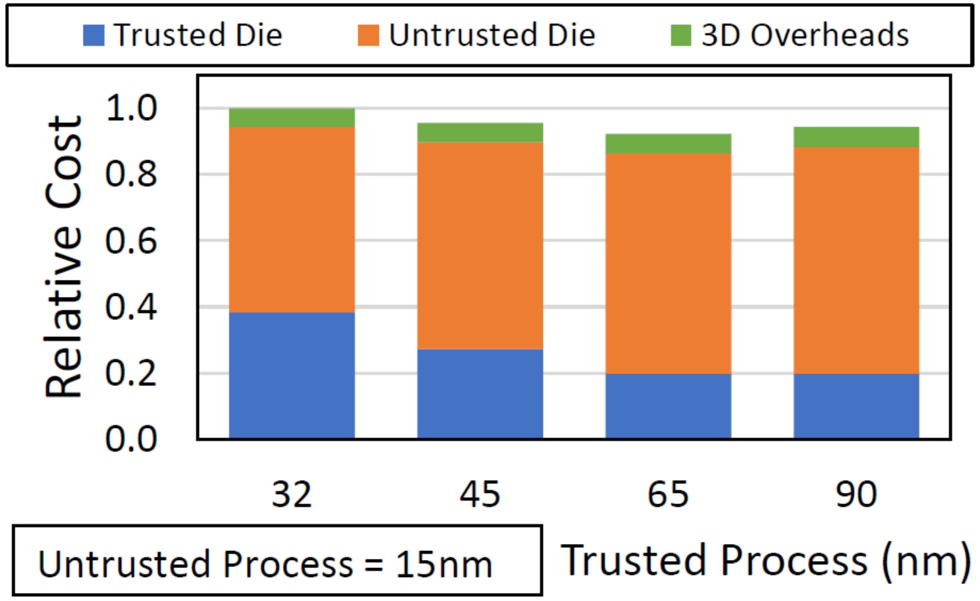


ubump size/pitch	bench mark	Trusted Area(um2) /Density	Unstrusted Area(um2) /Density	ubump Area (um2)
ub1 (1um/2um)	b14	8100/0.66	1293/0.68	289
	b21	9025/0.68	3021/0.7	399
ub5 (5um/10um)	b14	9018/0.18	8998/0.16	7225
	b21	12090/0.2	9981/0.17	9975

- Our proposed 3D split-fab introduces very low footprint overhead (22.6% avg.) compared with 2D 15nm baseline, and saves a lot area overhead (52.7% avg.) compared with 2D 45nm baseline.

Cost Evaluation





[D. Stow 2017]

Trusted Technology (nm)	90	65	45	32
Mask Cost Overhead	29%	30%	44%	68%

Trusted mask cost overheads (NRE)

- Our proposed **3D split-fab** introduces very low cost overhead (34% avg.) compared with untrusted 2D baseline, and is significantly cost-efficient compared with BEOL split-fab (400% avg.) and trusted 2D baseline (657% avg.).
- Cost breakdown shows that most of the cost belongs to advanced node (untrusted die, 65% avg.) and 3D overhead is relatively small (<10% avg.).
- Future work on IP reuse of trusted tier can further bring down NRE mask cost.

Summary

- We propose to securely select a partition to be fabricated in the advanced but untrusted foundry, while camouflaging part of the circuit at the trusted foundry to provide protection after manufacturing.
- Evaluation results show that our method can effectively improve security and optimize the cutsize with small overheads.
- Further, 3D cost analysis verifies that our method is cost-efficient compared to prior solutions.

Thank you!

Q&A

Reference

- [M. Jagasivamani 2014] Jagasivamani, Meenatchi, et al. "Split-fabrication obfuscation: Metrics and techniques." Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium on. IEEE, 2014.
- [Y. Xie 2015] Xie, Yang, Chongxi Bao, and Ankur Srivastava. "Security-aware design flow for 2.5 D IC technology." Proceedings of the 5th International Workshop on Trustworthy Embedded Devices. ACM, 2015.
- [J. Rajendran 2013] Rajendran, Jeyavijayan, et al. "Security analysis of integrated circuit camouflaging." Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. ACM, 2013.
- [P. Gu 2016] Gu, Peng, et al. "Leveraging 3D technologies for hardware security: Opportunities and challenges." Great Lakes Symposium on VLSI, 2016 International. IEEE, 2016.
- [D. Stow 2017] Stow, Dylan, et al. "Cost-effective design of scalable high-performance systems using active and passive interposers." Proceedings of the 36th International Conference on Computer-Aided Design. IEEE Press, 2017.