NC STATE UNIVERSITY

DRAPER

Trusted Fabrication through 3D Integration

Paul Franzon, Steve Lipa

Department of Electrical and Computer Engineering,

NC State University

slipa@ncsu.edu , paulf@ncsu.edu

Lisa McIlrath,

Draper

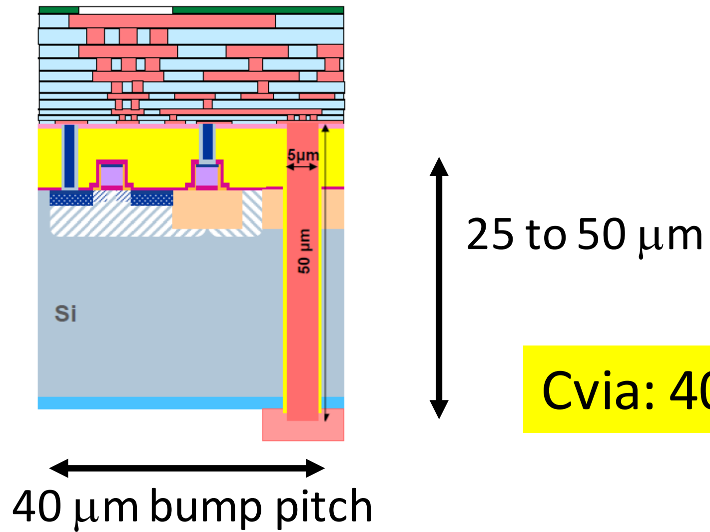
lmcilrath@draper.com

Outline

- 3D Technology Review
- 3D Projects at NCSU
- Trusted Fabrication Project
  - Opportunity and Threats
  - Obfuscation
  - Technology
  - Simulation Demonstration
  - Future Work

3D Technology Set

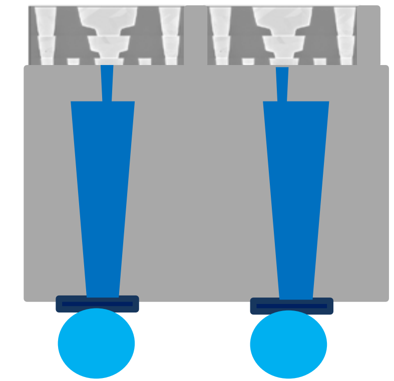
- 3DIC with TSVs



Tomorrow:
- DRAM 20 µm long TSV
- Logic: 5 µm long TSV
- 1- 2 µm pitch or below
- 25 µm bump pitch

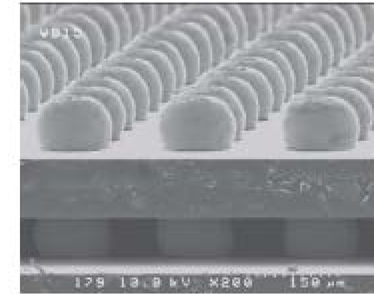
Cvia: 40 fF today → 2 fF tomorrow

- Interposers:
  - 50 µm thick 100 µm TSV pitch
  - Today: 10 µm wire features
  - Tomorrow: sub 1 µm wire features



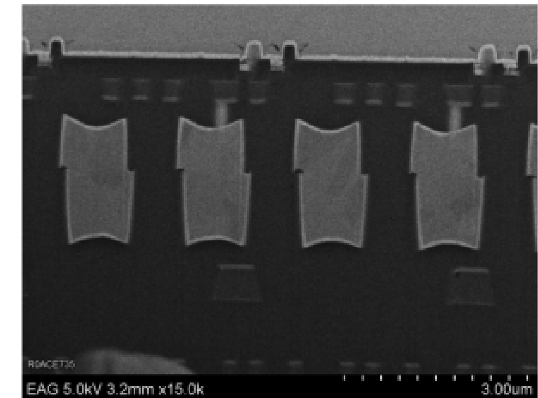
Attachment technologies

- Solder microbumps
  - Today typically 30 – 40 µm pitch
  - Potential for 5 µm pitch



IBM

- Copper-copper
  - Thermo-compression
    - @ high temperature (> 400 C)
  - Hybrid bonding
    - @ low temperature (Ziptronix DBI)
  - Typical 2 – 5 µm pitch
  - Potential for sub-1 µm pitch



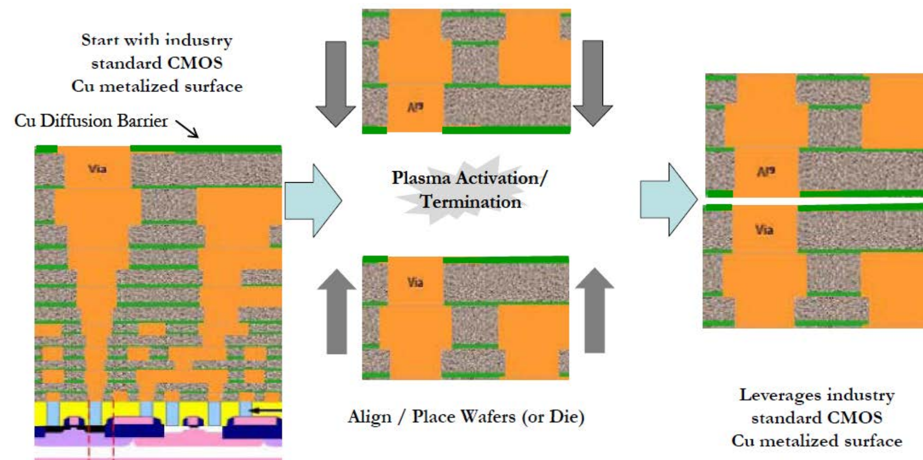
Ziptronix

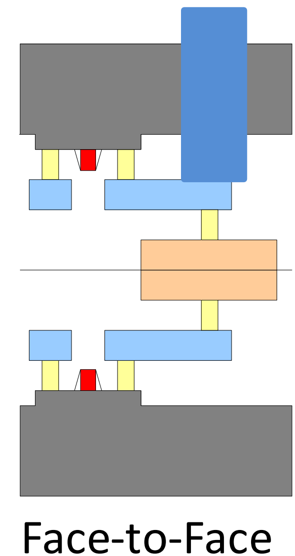
NC STATE UNIVERSITY

DRAPER

Hybrid Bonding

- Direct Bond Interface Steps:
  1. Terminate wafer processing on a via layer
  2. CMP surface
  3. Plasma active bonding
  4. Bond at slightly elevated temperature

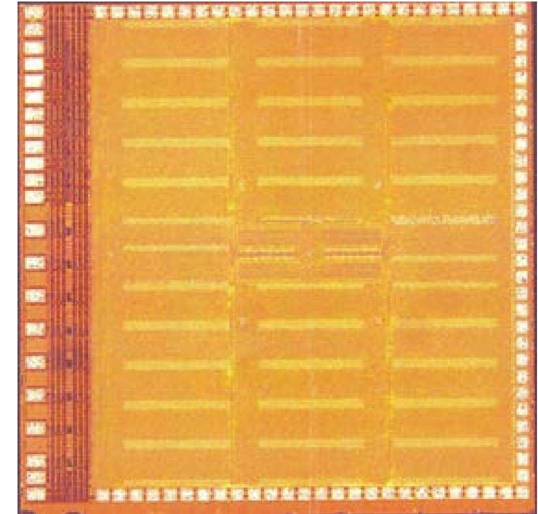




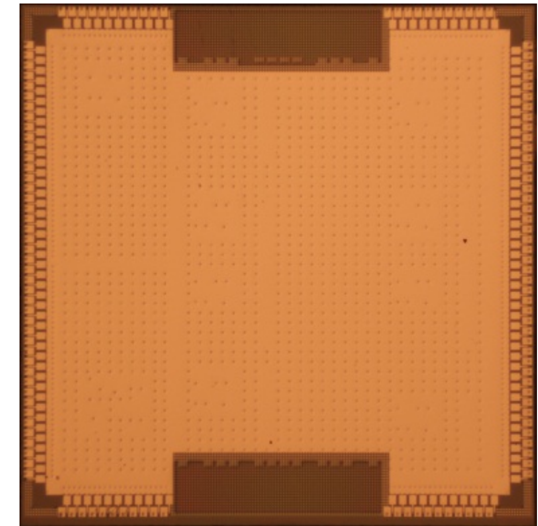
Face-to-Face

Some Past 3D Projects at NCSU

- 3D FFT with 60% power savings

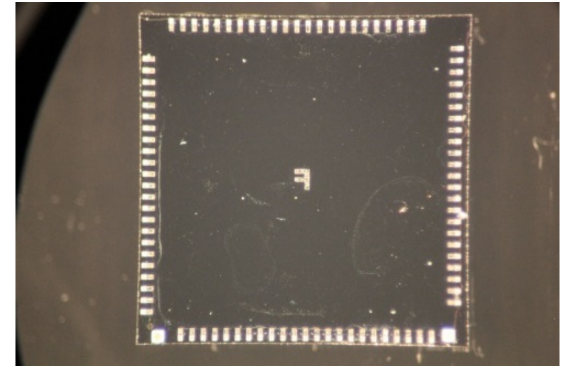


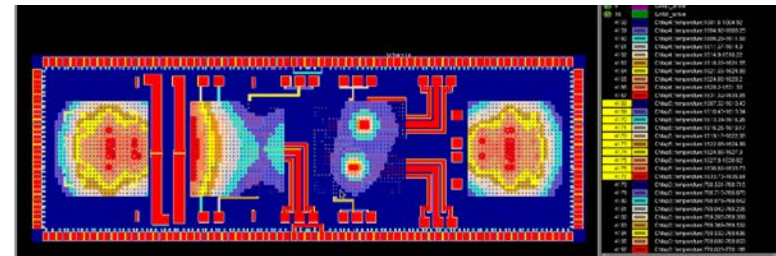
- SAR processor that achieves one Moore's law generation of scaling
  - 22% improvement in performance/power



Some Past 3D Projects at NCSU

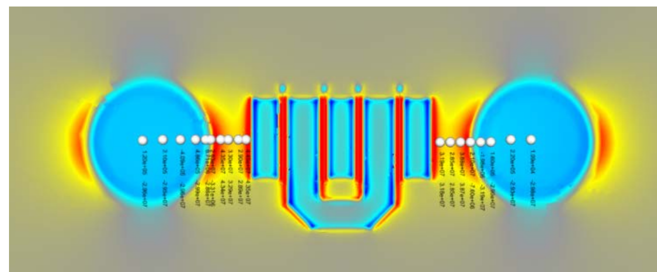
- Two core heterogeneous processor
  - 25% improvement in performance/power
- SIMD Machine Intelligence accelerator
- Thermal and stress analysis of GaN & InP on Silicon





Top view

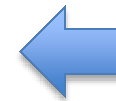
Side view



Just starting / Future Projects

- Demonstration of obfuscation through 3D integration
  - Continuation of this project
- 2.5D integration of chiplets for machine learning and machine intelligence (DARPA chips program)
- Reliability evaluation
  - With Vanderbilt (future)

Outline

- 3D Technology Review
- 3D Projects at NCSU
- Trusted Fabrication Project 
  - Opportunity and Threats
  - Obfuscation
  - Technology
  - Simulation Demonstration
  - Future Work

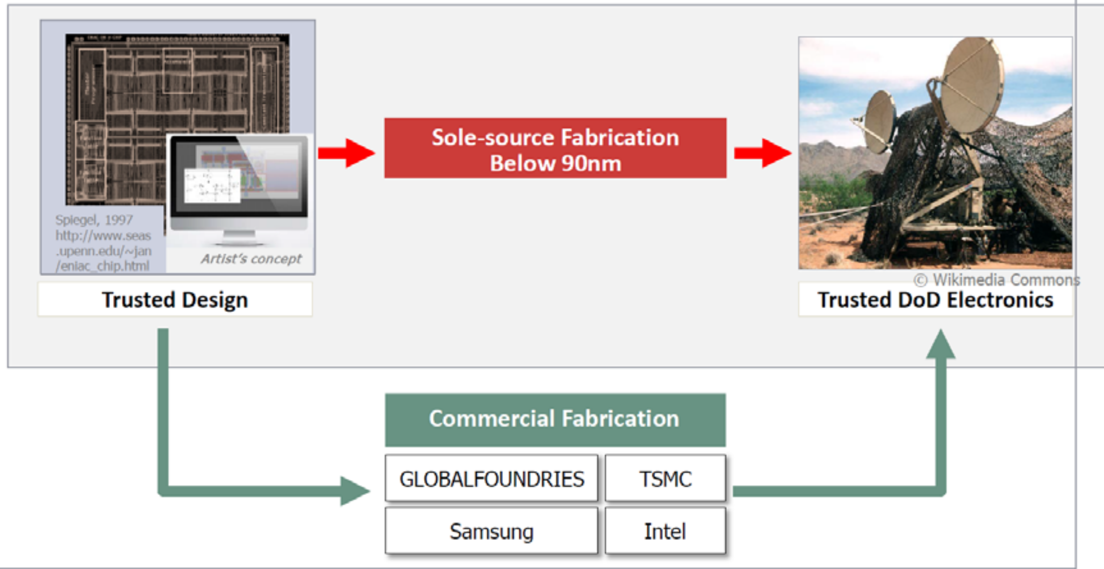
DOD Opportunity

- Use advanced commercial processes in DOD systems
- Requires complete design being sent to fab

It is the right time for DoD to reflect on its strategy

Today: DoD relies on a single, sole-source supplier for leading-edge microelectronics

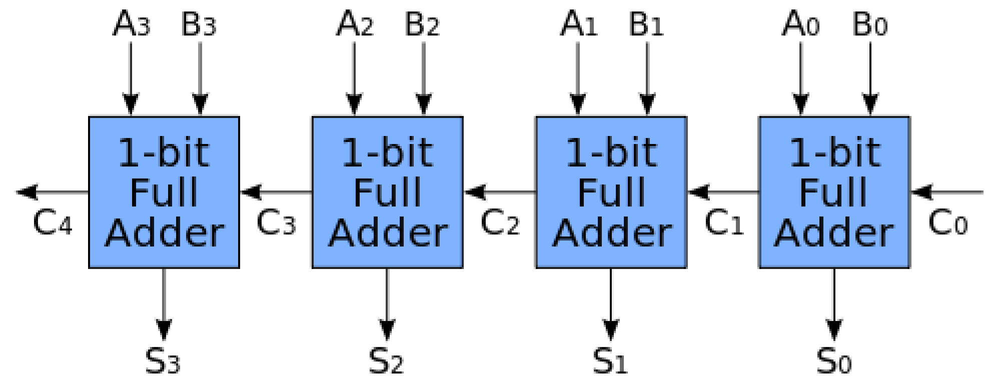
Tomorrow: Technology-driven security techniques can enable new DoD options for acquiring state-of-the-art, commercial microelectronics



Ken Plaks, DARPA

Threats

Capability Discovery



Trojan Insertion

Counterfeiting

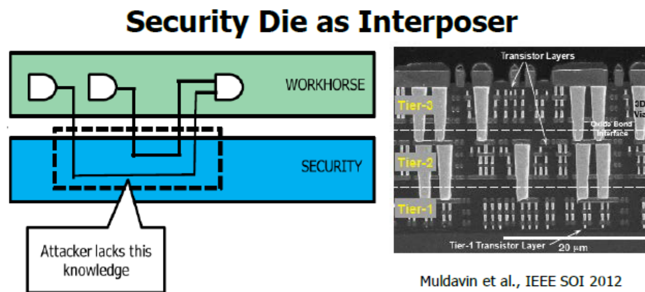
Obfuscation

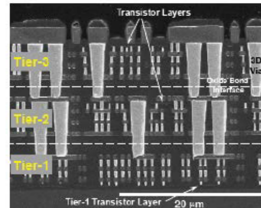
- Potential solution to exponentially complicate reverse engineering

Goal: Make it harder to insert hardware Trojan and other malicious logic into hardware

Mitigate risk through design obfuscation

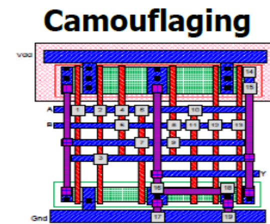
Security Die as Interposer





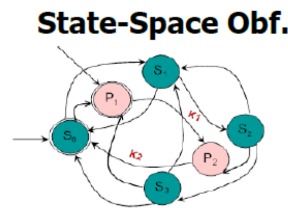
Muldavin et al., IEEE SOI 2012

Camouflaging



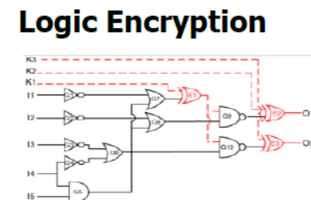
Rajendran et al., ACM SIGSAC 2013

State-Space Obf.



Chakraborty & Bhunia, ICCAD 2008

Logic Encryption



Rajendran et al., IEEE Tcomp 2015

Key Research Questions:

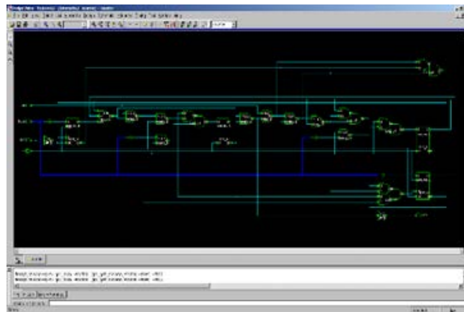
- What are the metrics & how do we measure effectiveness of obfuscation techniques?
- What design tools are needed to facilitate implementation of obfuscation?
- Is the solution scalable? What is the solution's impact on performance?

Ken Plaks, DARPA

NC STATE UNIVERSITY

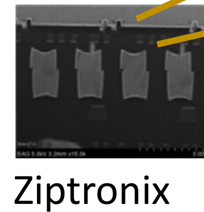
DRAPER

Obfuscation Through 3D

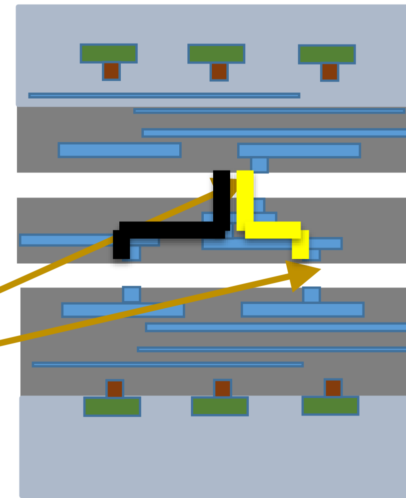


Netlist

3D partitioning and place & route



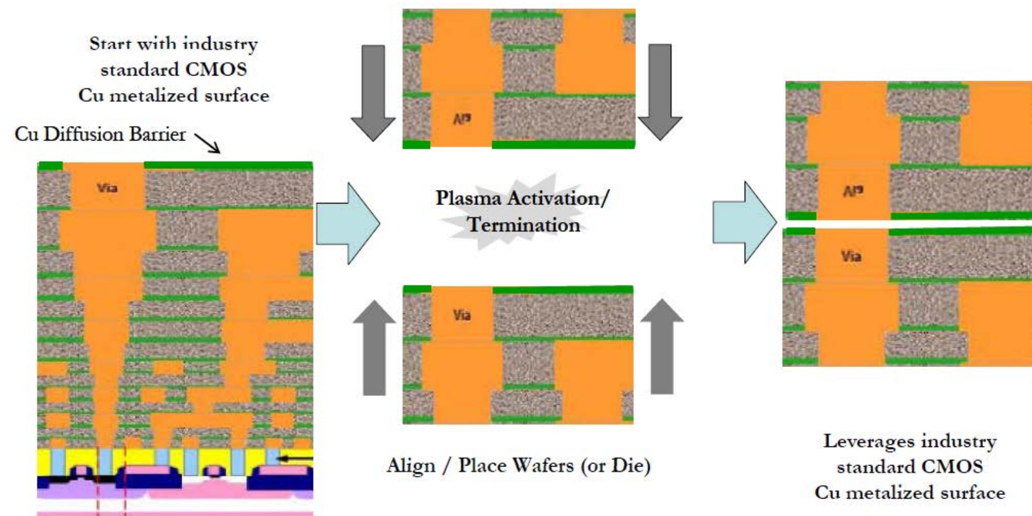
Ziptronix



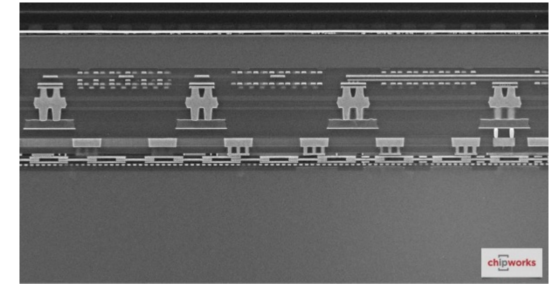
- Trust is created by not knowing the connections between the two untrusted tiers
- Secure fabs available for trusted wiring tier manufacture
- Difficult to delayer

Hybrid Bonding

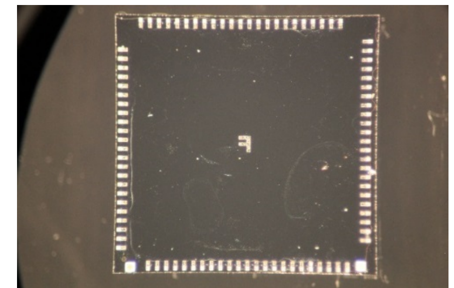
- Steps:
  1. Terminate wafer processing on a via plug layer
  2. CMP surface
  3. Plasma active bonding
  4. Bond at slightly elevated temperature



Hybrid Bonding features



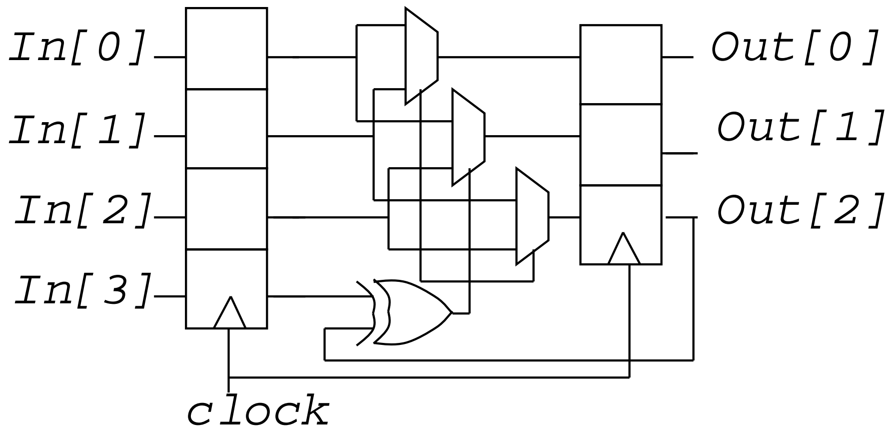
- 3 – 10 μm pitch in products today
  - Used by several companies
  - Used in many cell phone cameras (6 μm pitch)
- 1 μm pitch achievable today
- Active research into sub - 1 μm pitch at IMEC and elsewhere
- Potentially $500/wafer-pair in volume
  - Low volume cost higher
- Direct experience @ NCSU



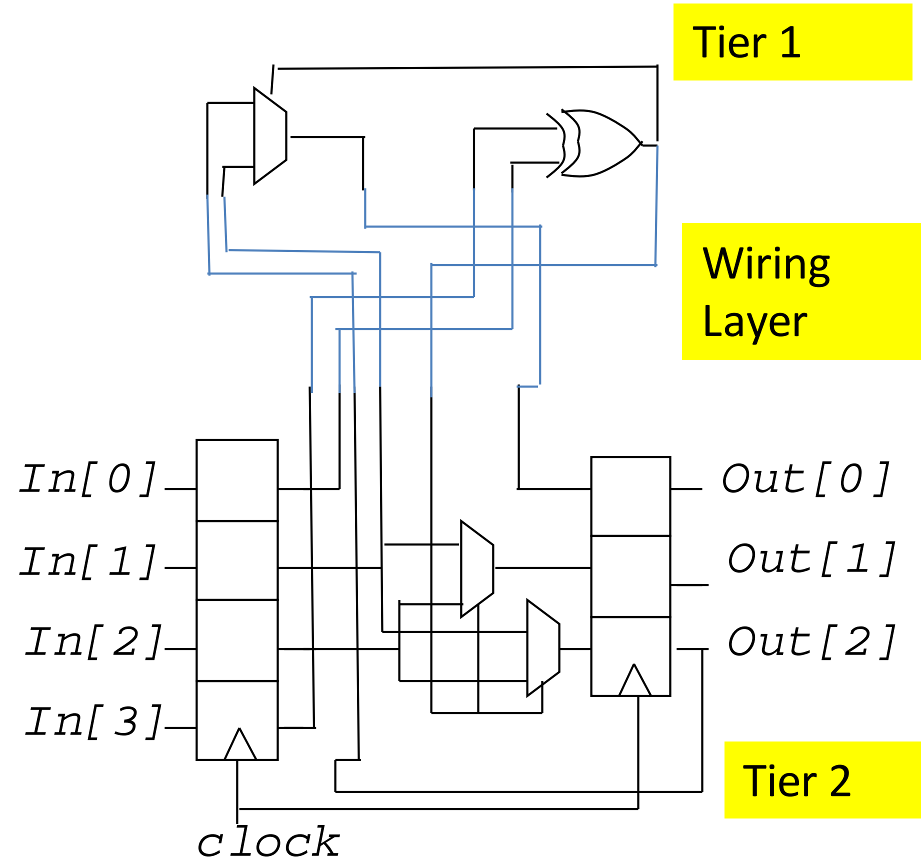
NC STATE UNIVERSITY

DRAPER

Example

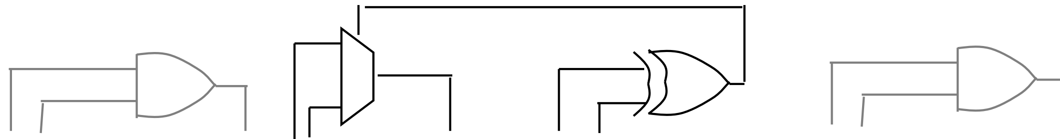


(a) Netlist

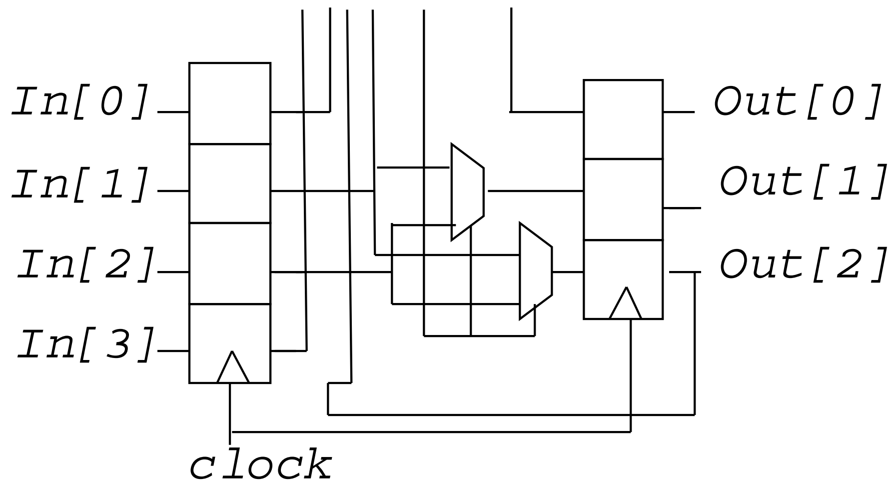


(b) Partitioned Netlist

...Example



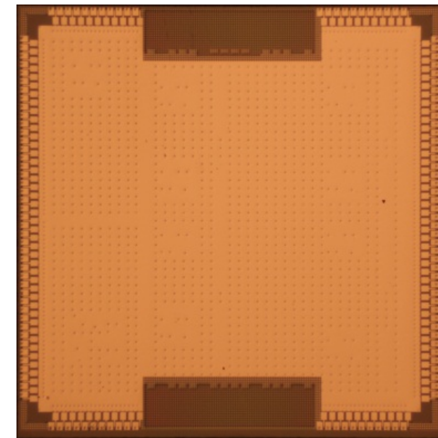
Tier 1



Tier 2

(c) What adversary sees (at best)
(with unrelated netlists mixed in)

Performance Improved!



- Two tier chip, 6.6 μm bond pitch
- Impact on Performance and Power:

	Total Wire Length (% Change)	Max Frequency (% Change)	Parasitic Po[illegible]	
PE 3D Seq.	-17.1%	+7.1[illegible]	[illegible]	[illegible]%
PE 3D Sim.	-17.7%	[illegible]	[illegible]	-7.7%
PE 3D True	-21[illegible]	[illegible]	-45.2%	-12.9%
AES 3D Seq.	[illegible]	[illegible]	-19.6%	-2.6%
MIMO 3D Seq.	[illegible]	+17.1%	-34.9%	-5.1%

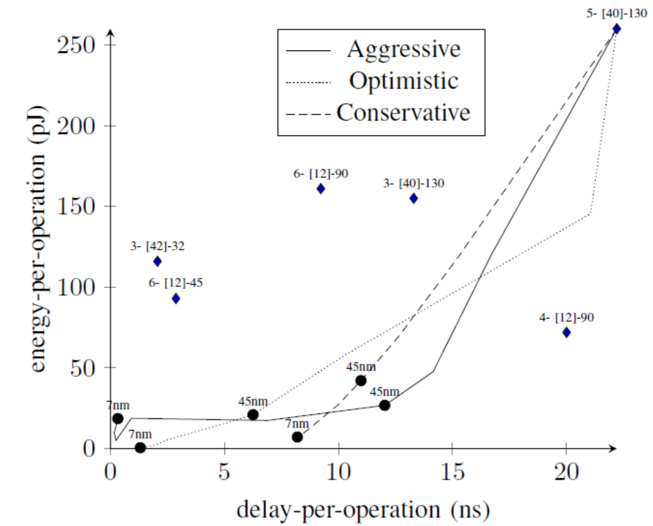
18% - 35% improvement in Performance/Power (21% for SAR (“PE Seq”))

T. Thorolfsson, S. Lipa and P. D. Franzon, "A 10.35 mW/GFlop stacked SAR DSP unit using fine-grain partitioned 3D integration," *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference*, San Jose, CA, 2012, pp. 1-4.

T. Thorolfsson, G. Luo, J. Cong and P. D. Franzon, "Logic-on-logic 3D integration and placement," *3D Systems Integration Conference (3DIC), 2010 IEEE International*, Munich, 2010, pp. 1-4.

Equivalent to one node of scaling

- One node:
  - Conservative scaling = Intel
  - FP Mult-Accumulate
  - 45 nm – 7 nm
    - 4x energy/op; 1.5x delay
    - 6x power/performance
    - Nodes: 45:30:22:15:7 (5 generations)
    - i.e. ~1.2 per generation
  - **3DIC folding = ONE GENERATION OF IMPROVEMENT**



A Review of Low-Power Digital Design

Joshua Schabel, *Student Member, IEEE*, Jong Beom Park, *Student Member, IEEE*, William Rhett Davis, *Fellow, IEEE*, and Paul D. Franzon, *Fellow, IEEE*

Experiment Conducted

- Took OpenCore 256 point FFT design
- Partitioned it between the two CMOS tiers
  – Kept all the flip-flops in one of the tiers so as to not have to write a 3D clock routing tool
  – Tried to balance the area between top and bottom tiers
  – “Dumb” (automatic) partitioner

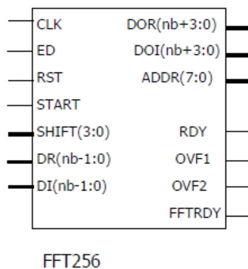
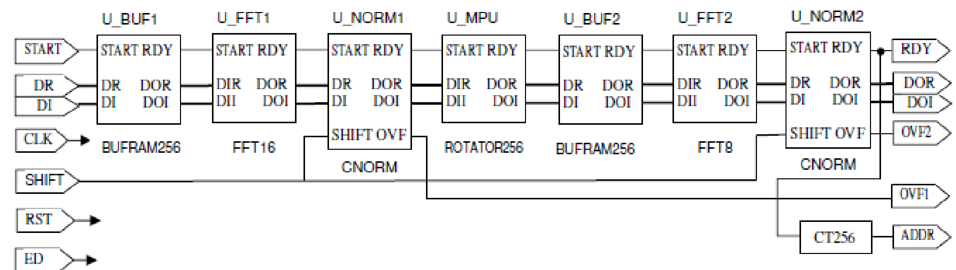
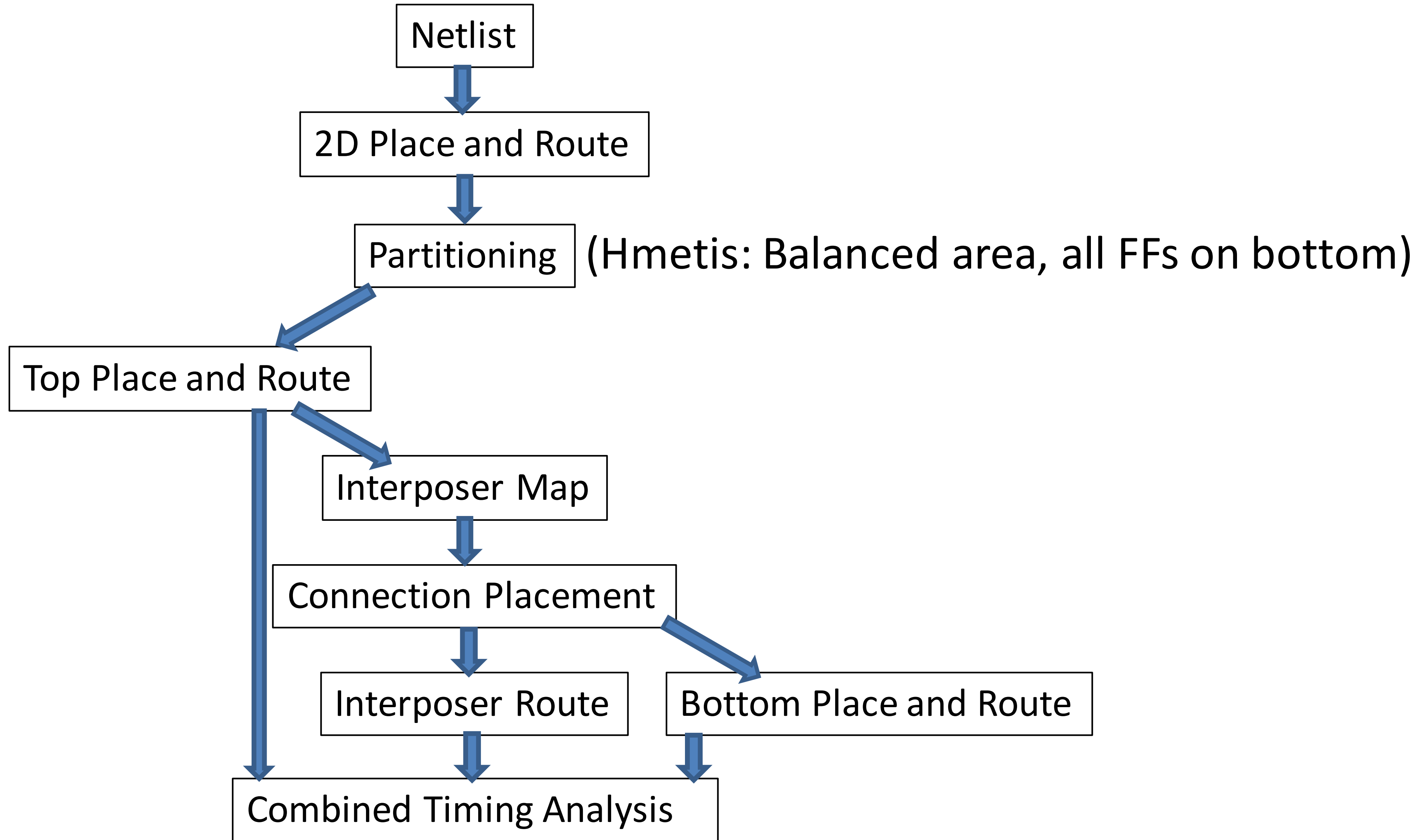


Figure 2. FFT256 symbol.

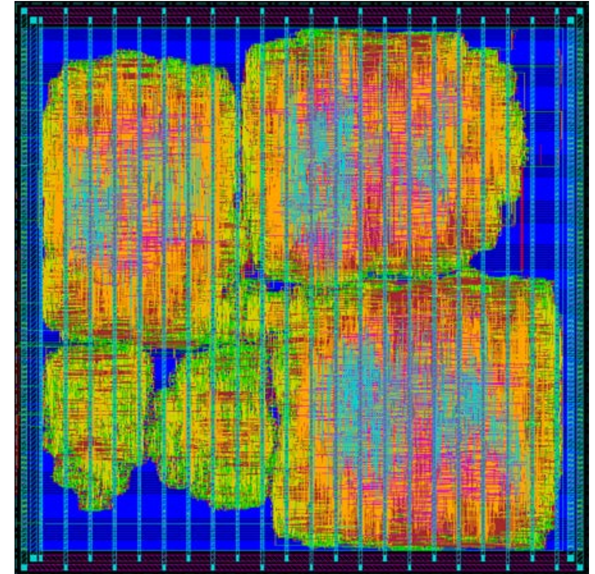


CAD Flow

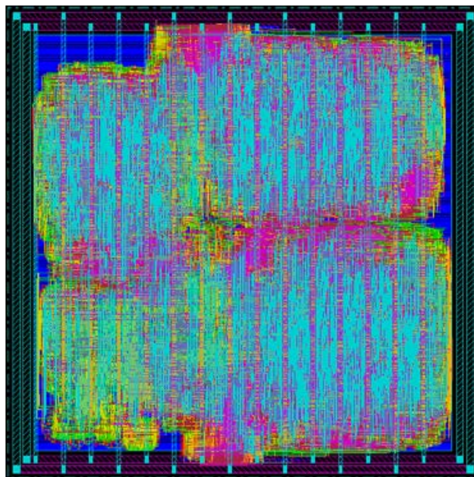


Design

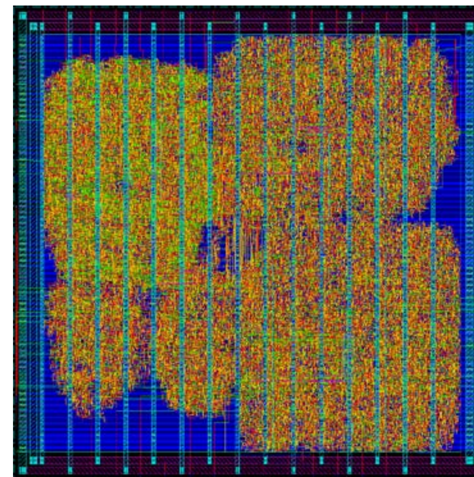
- 256 point radix-8 FFT
- 850x850 (0.72 sq. mm) 2D
- 600x600 (0.36 sq. mm) 3D
- 389K gates, including 95,760 flops and 147,128 total nets
- 99K face-to-face bumps for nets (1 in 4)
- 9 LM, 1 micron bump pitch
- **Netlist complexity: > 400!** ^{900}



2D



TOP



BOTTOM

Netlist Complexity

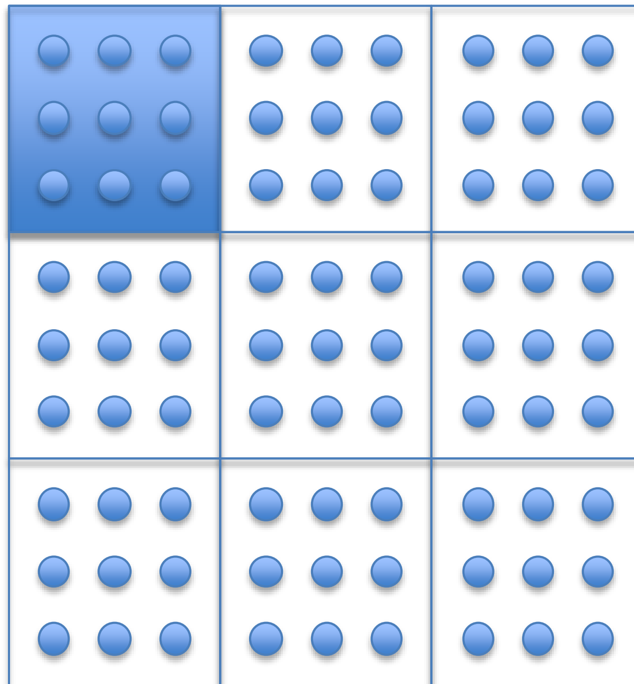
Candidate Netlists > R! ^ (N/R)

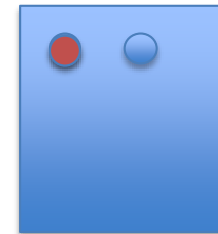
N = Number of crossing nets

R = Number of reachable connections

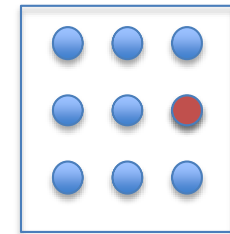
E.g. R = 9

N= 81





Top



Bottom

Net 1: 9 choices
Net 2: 8 choices
Etc.

Red Teaming

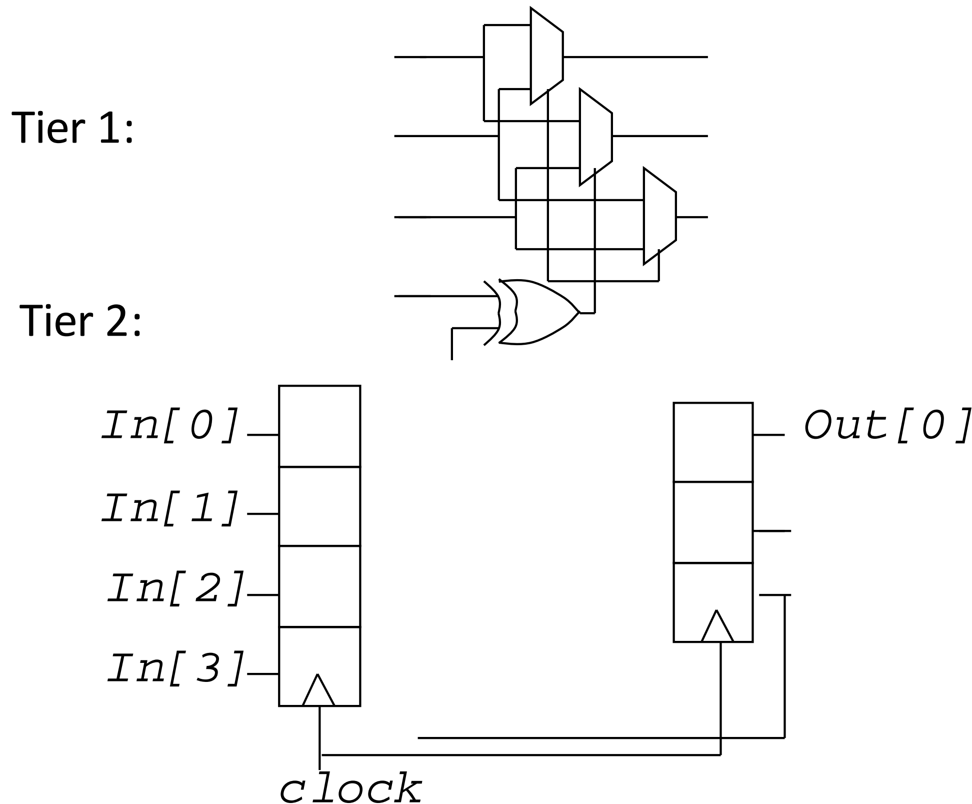
- 6.5 man weeks of effort
- Not able to reverse engineer netlist
- Was able to obtain functionality of certain combinational logic blocks
  - Could indicate partial capability and could be basis for Trojan insertion
  - Made easier by “dumb partitioning”

Reverse Engineering Approach

- Use Boolean satisfiability solver
  - Allows Boolean functions to be identified in sea of logic
- Certain functions easy to identify, E.g.
  - Adders (XOR gates);
  - Memory address decode logic;
  - counters;
- Other functions were difficult
  - Butterfly logic
  - Well partitioned state machines

Why some comb Logic was discoverable

- Flip-flops are large, so a number of equal area partitions came out like this



Gives:
- Adder bit lengths
- Memory decoder logic

Smart Partitioning

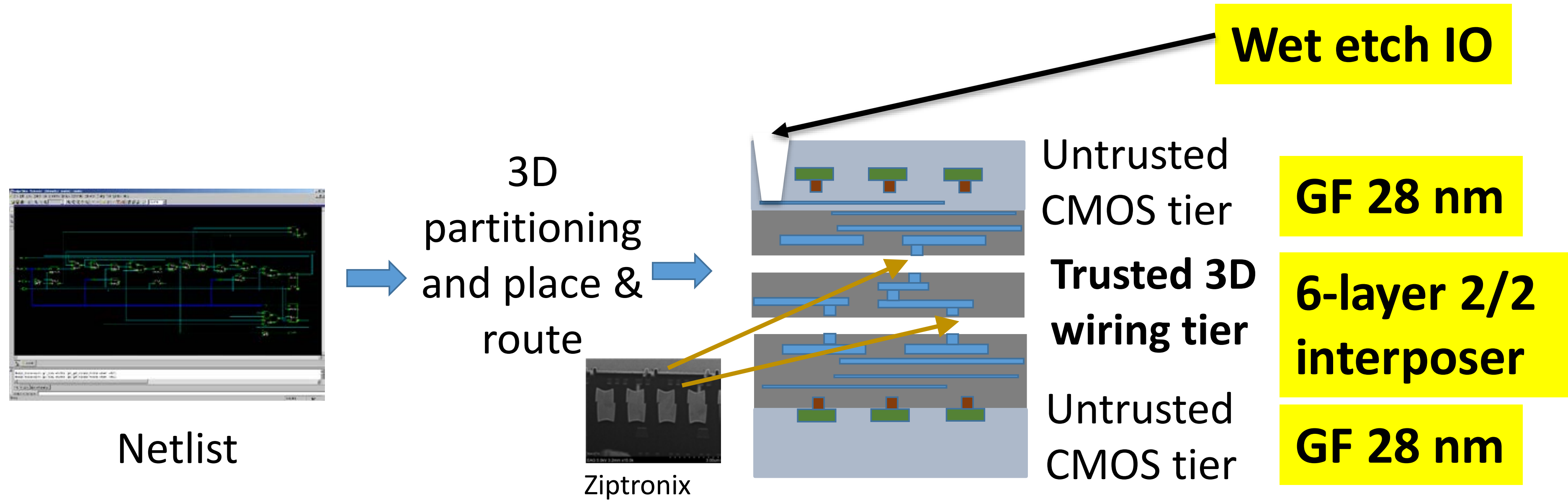
- Not everything is worth partitioning
  - E.g. Why partition a standard IP block
- Can focus on unique functions
  - But this itself needs to be obfuscated
- Can carefully choose what and where to partition

Smart Partitioning

Practice	Reason
Have flip-flops on both tiers and partition away from flip-flop	Results in exponential complexity
Partition hierarchically	Removes “crib” of different partitions for same function
Break up logic for arithmetic at carry chain	Obfuscate bit length (capability)
Partition decode logic and lookup tables	Uniform decode logic easy to identify
Partition complex interconnect	Very hard to reverse engineer
Partition state decode logic	Very hard to reverse engineer
Merge neighboring FSMs	Increased complexity

Next Steps

- Physical Demonstration using GF 28 nm and nHanced Semi Interposer



Next Steps

- Designs to be obfuscated
  - Arithmetic
  - Finite State Machines
  - AES core
  - SIMD compute core
- December tapeout for September demonstration

Conclusions

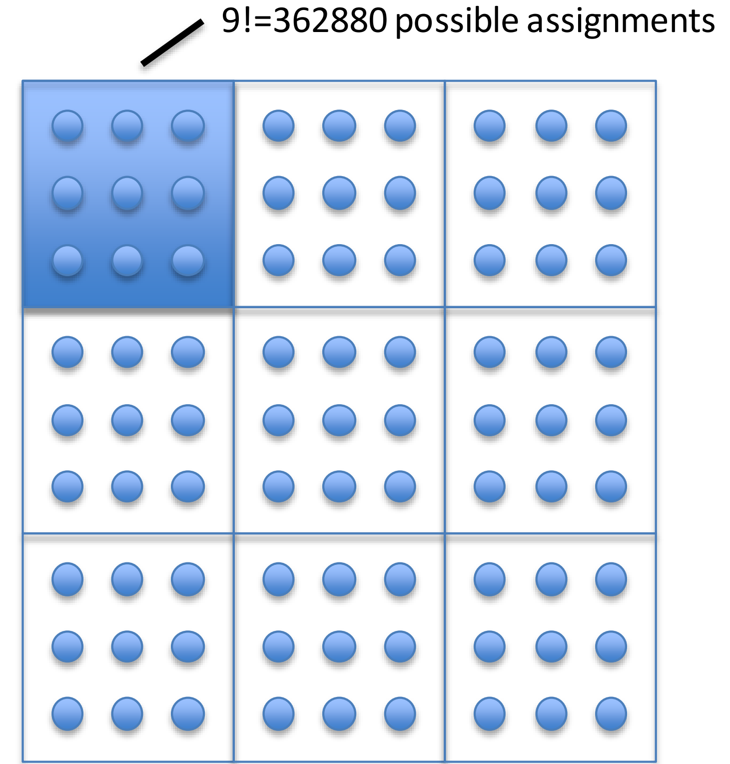
- 3D partitioning used a trusted routing layer has potential to greatly increase the cost of executing the following threats
  - Capability discovery with no trusted IC fab
  - Sophisticated Trojan insertion
- Preliminary experiment confirms netlist complexity
  - Exponential complexity with non uniform FF or area assignment
- Building Smart partitioner

Acknowledgements

- Funded by Dan Green (DARPA) through ONR

Netlist Complexity

- Imagine 81 nets in grid.
- Assume each bump on top can reach the nearest 9 bumps on the bottom.
- Assume oversimplified approach where bumps are assigned in 3x3 groups.
- Top left group: after first bump assigned, there are 8 choices left…
- Thus 9! different assignments possible for top left group.
- For each of these, there are 9! assignments possible for the group to the right.
- And so on…



$9!^9 = 1.09 * 10^{50}$ possible assignments (there are potentially many more…)

Next step in Seedling

- Produce new design that
  1. Assumes a 3D clock router
  2. Follows "smart partitioning" guidance as much as practical with current partitioning flow

Hand to red-team for analysis

CAD FLOW

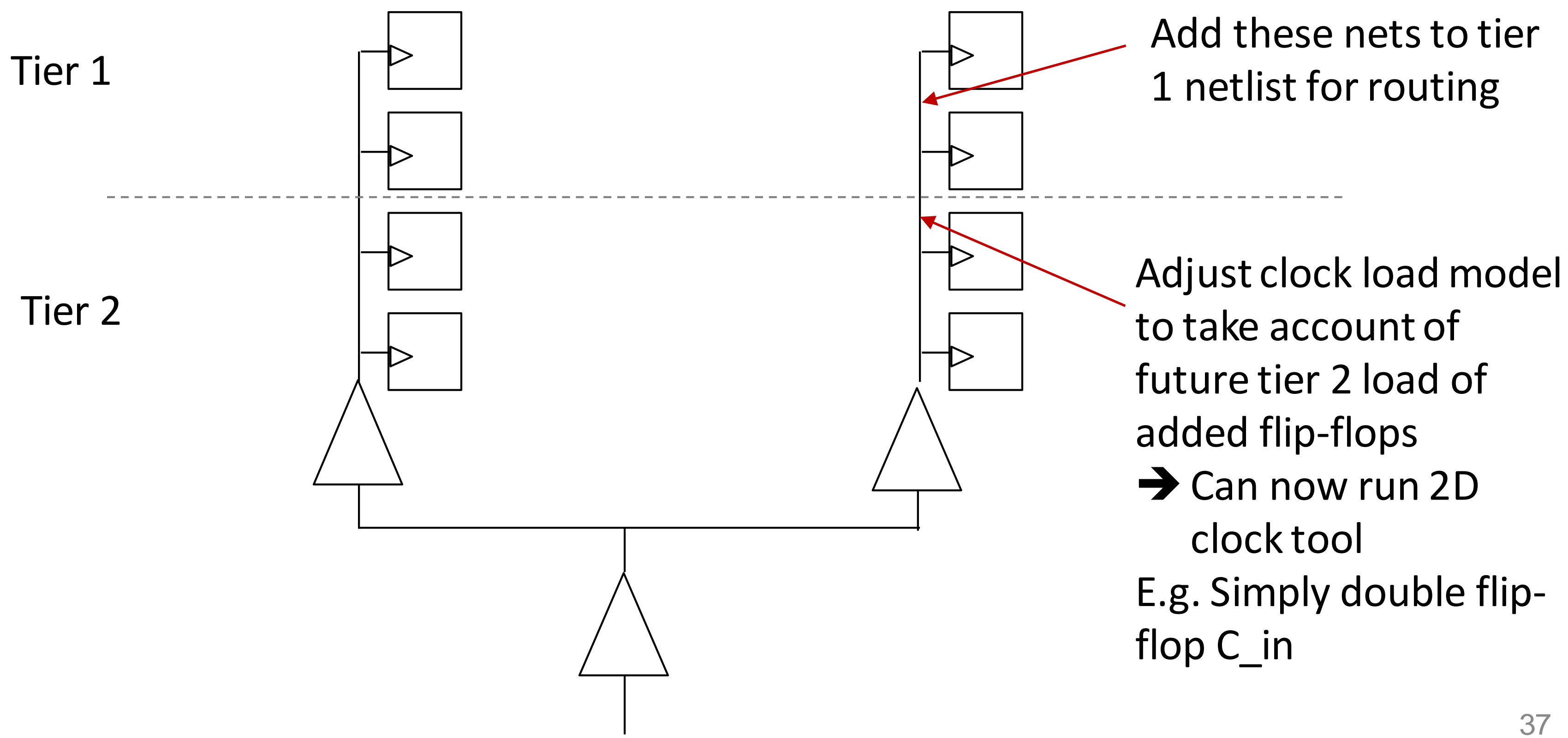
- Simplified flow:
  - synthesis
  - full 2D PR in area A
  - partition to TOP and BOTTOM
  - PR TOP in area A/2
    - interposer map created and used to limit f2f bump choices
    - f2f bumps assigned based on net centroids from 2D design scaled down
    - placement based on f2f bumps
  - PR BOTTOM in area A/2
    - f2f bumps assigned based on TOP bumps and interposer map
    - placement based on f2f bumps
  - TOP and BOTTOM routed netlists combined for timing analysis
  - TOP and BOTTOM tweaked as needed

Comment #1

- Keeping clock (i.e. flops) within one tier enabled identifying functionality of many logic blocks easy
- Flops take a lot of area so this forces logic into other tier, i.e.

Solution: 3D Clock Router

- An achievable commercial-process compatible 3D clock router could be built like this:



3D Clock Router

- Other benefits:
  - Would reduce the number of vertical signal connections by a factor of two
    - Only need one per partitioned net, not two
- Metrics improve with fewer partitions needed

Comment #2

- Partitioning more effective if more attention given to WHAT and where to partition
- Only partially possible with current Hmetis partitioner
- Could be possible with future “Smart partitioning” tool

NC STATE UNIVERSITY

DRAPER

256-point FFT Engine

- Top Level:

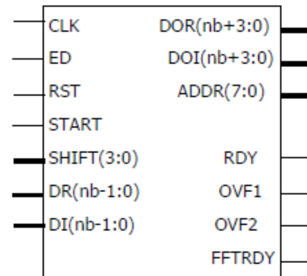
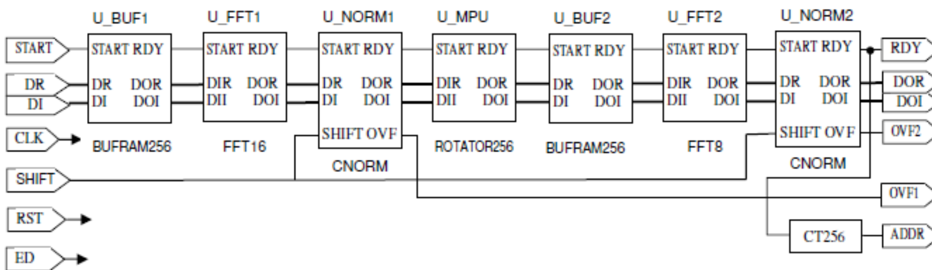
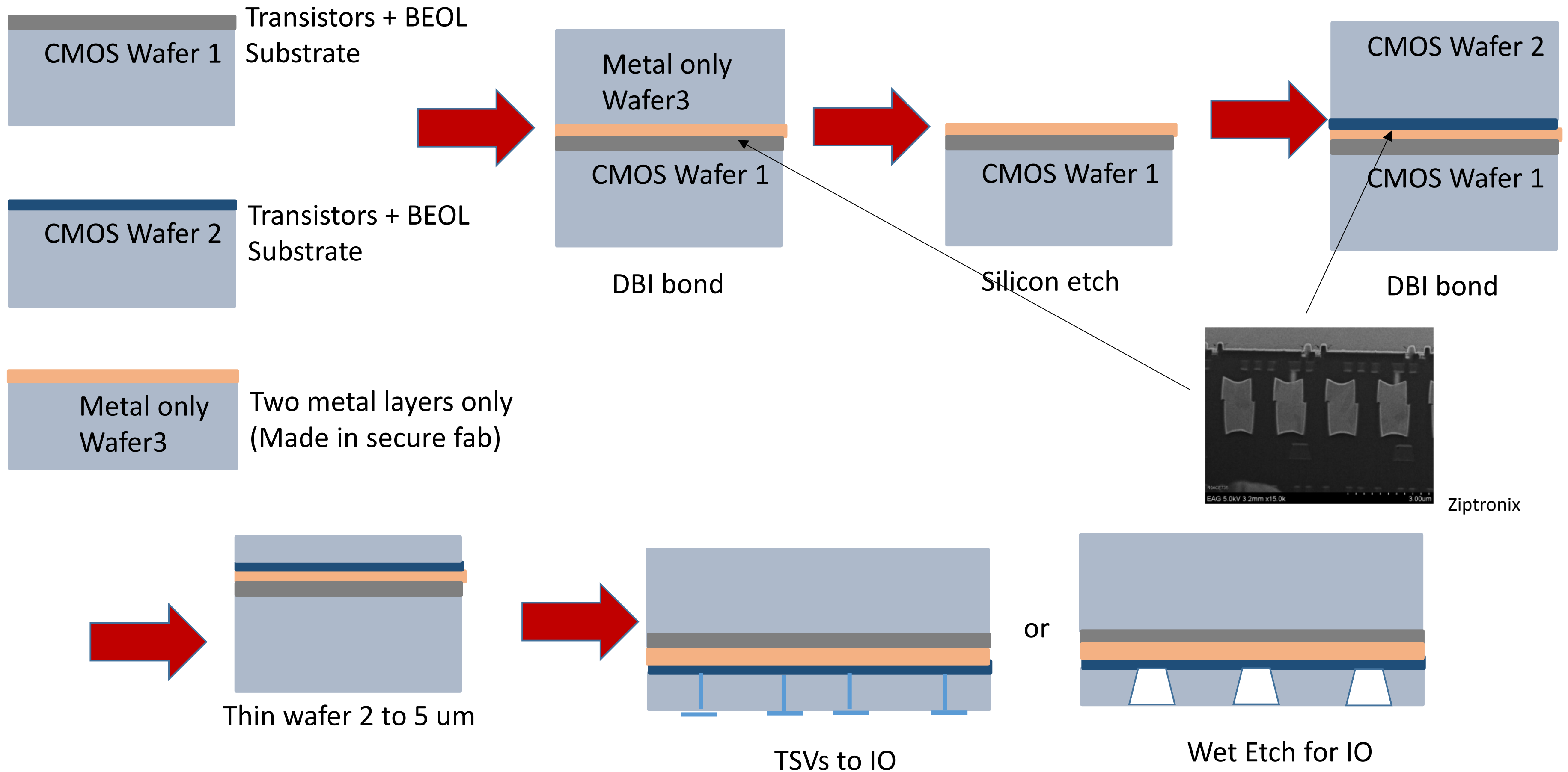


Figure 2. FFT256 symbol.

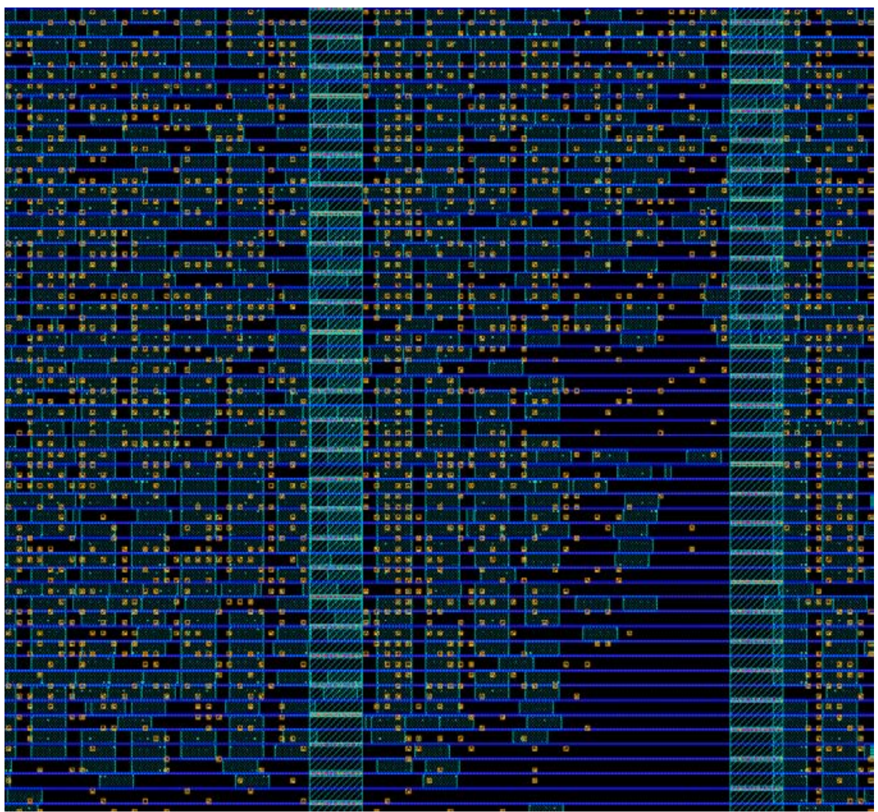
- Core:

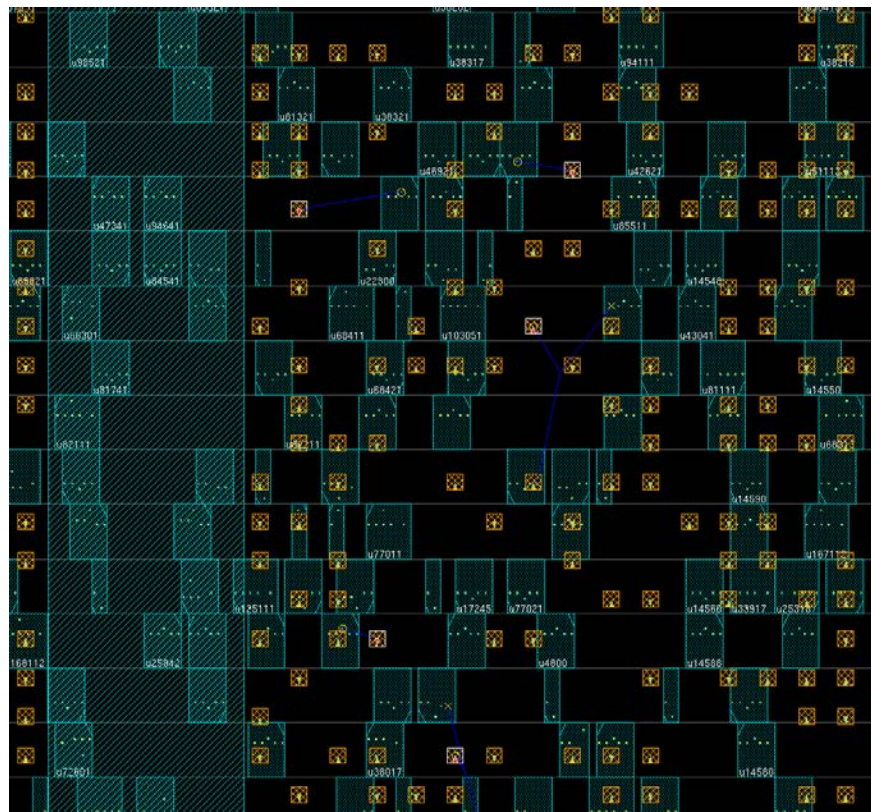


Manufacturing Flow



Layout Details





These pictures show how the bump assignments affect the placement…