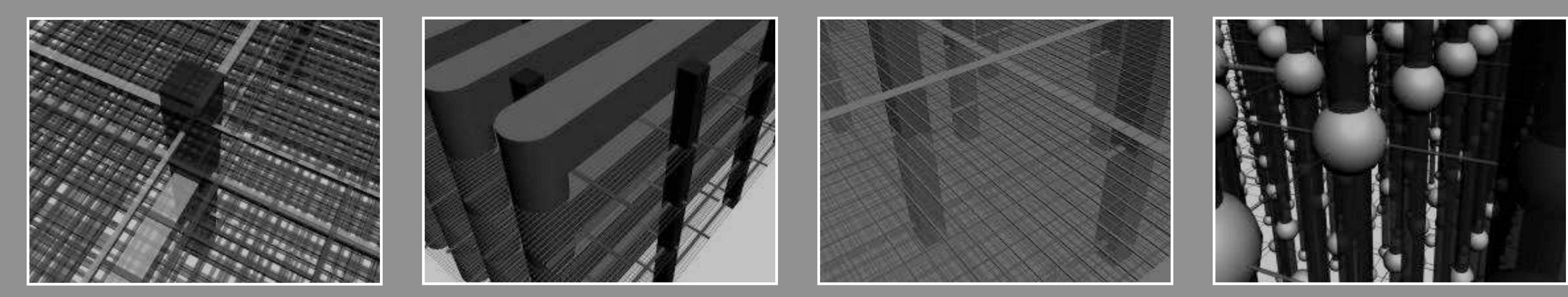
## Georgia College of Tech Engineering

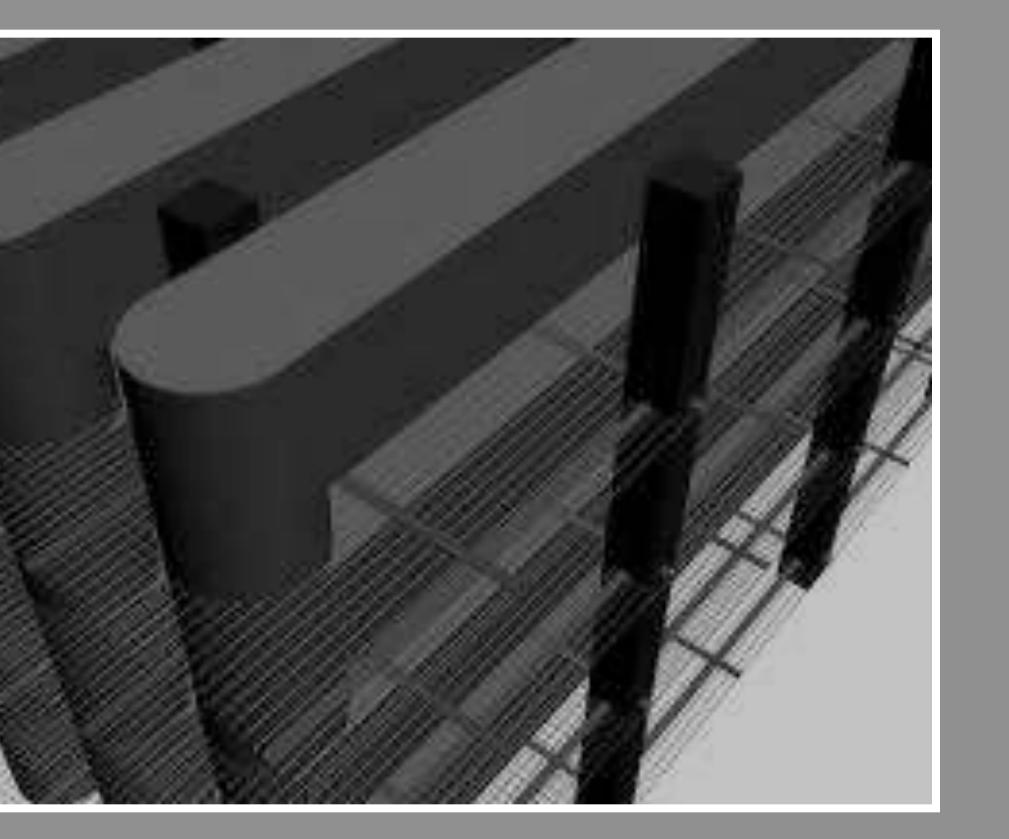
## Low Power Monolithic 3D IC Design of Asynchronous AES Core

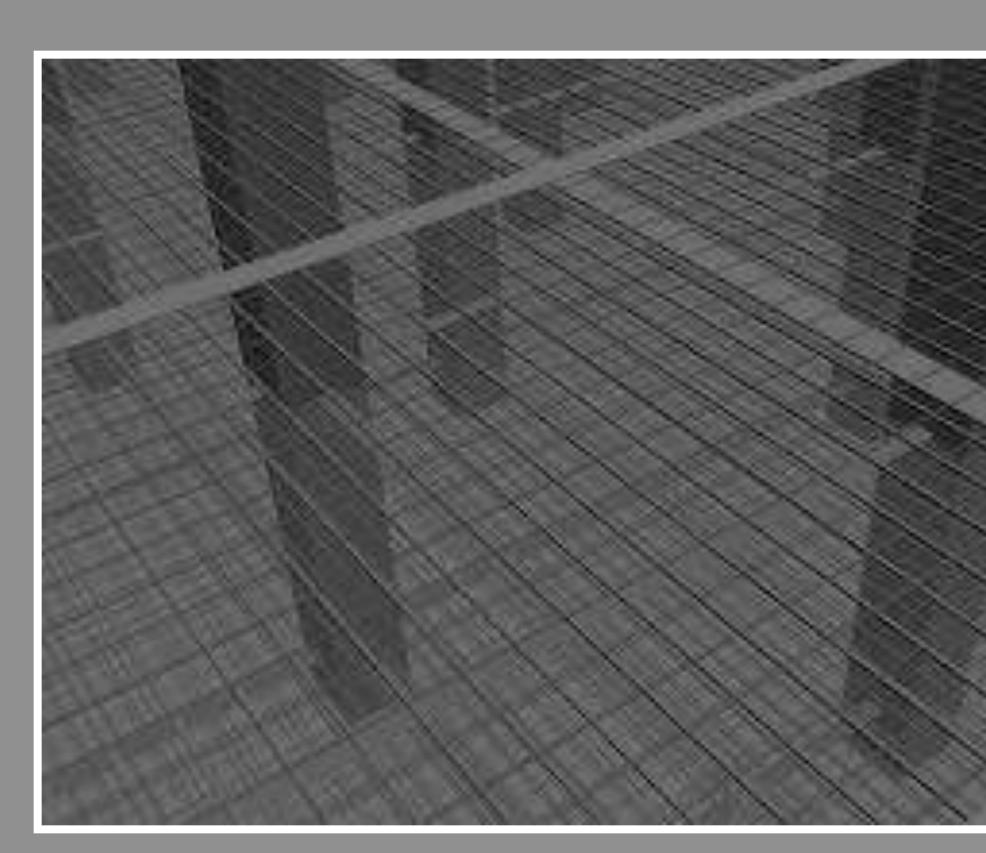


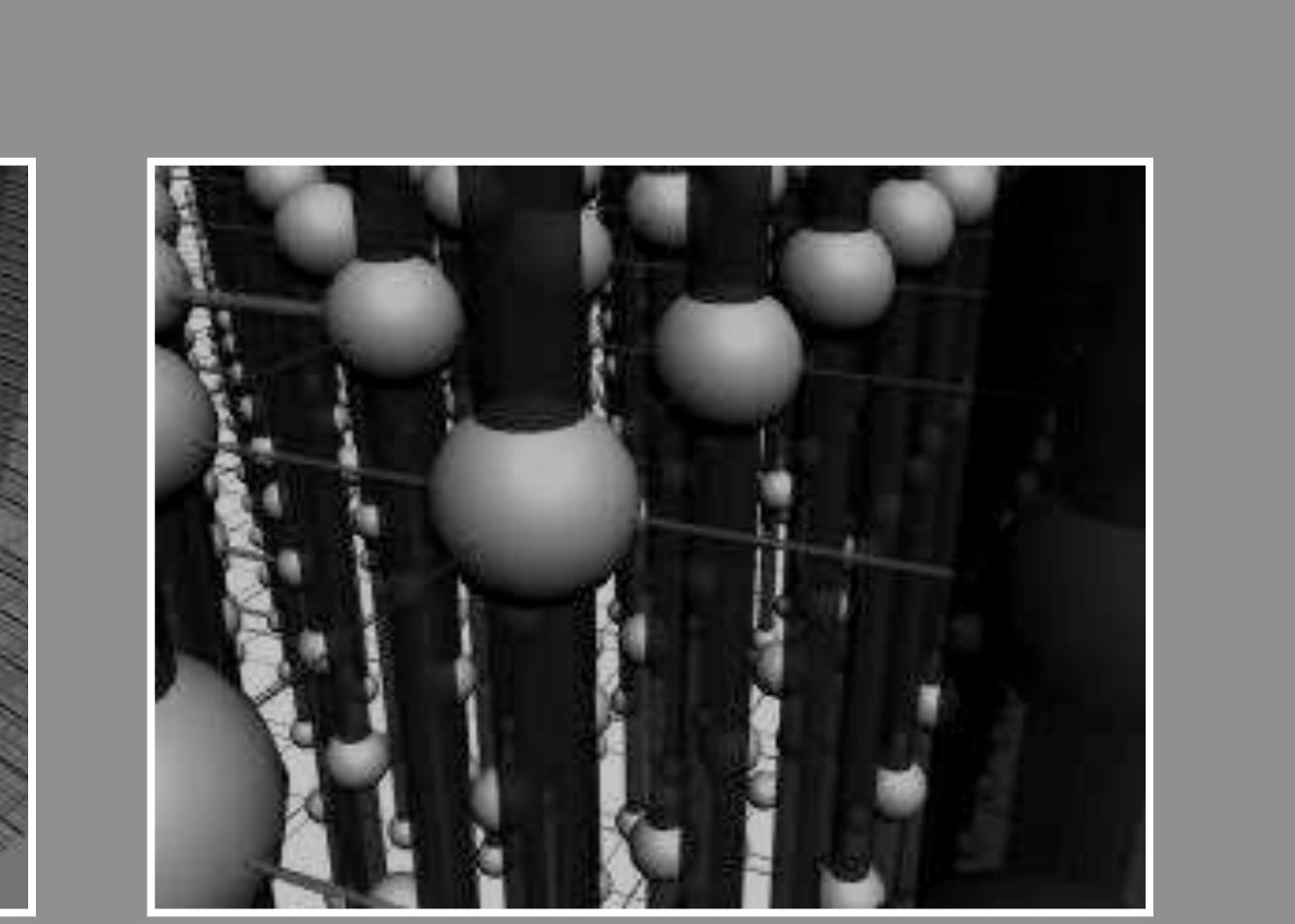
## Neela Lohith Penmetsa<sup>1</sup>, Christos Sotiriou<sup>2</sup>, and Sung Kyu Lim<sup>1</sup>

## <sup>1</sup>School of ECE, Georgia Tech, Atlanta <sup>2</sup>University of Thessaly, Greece

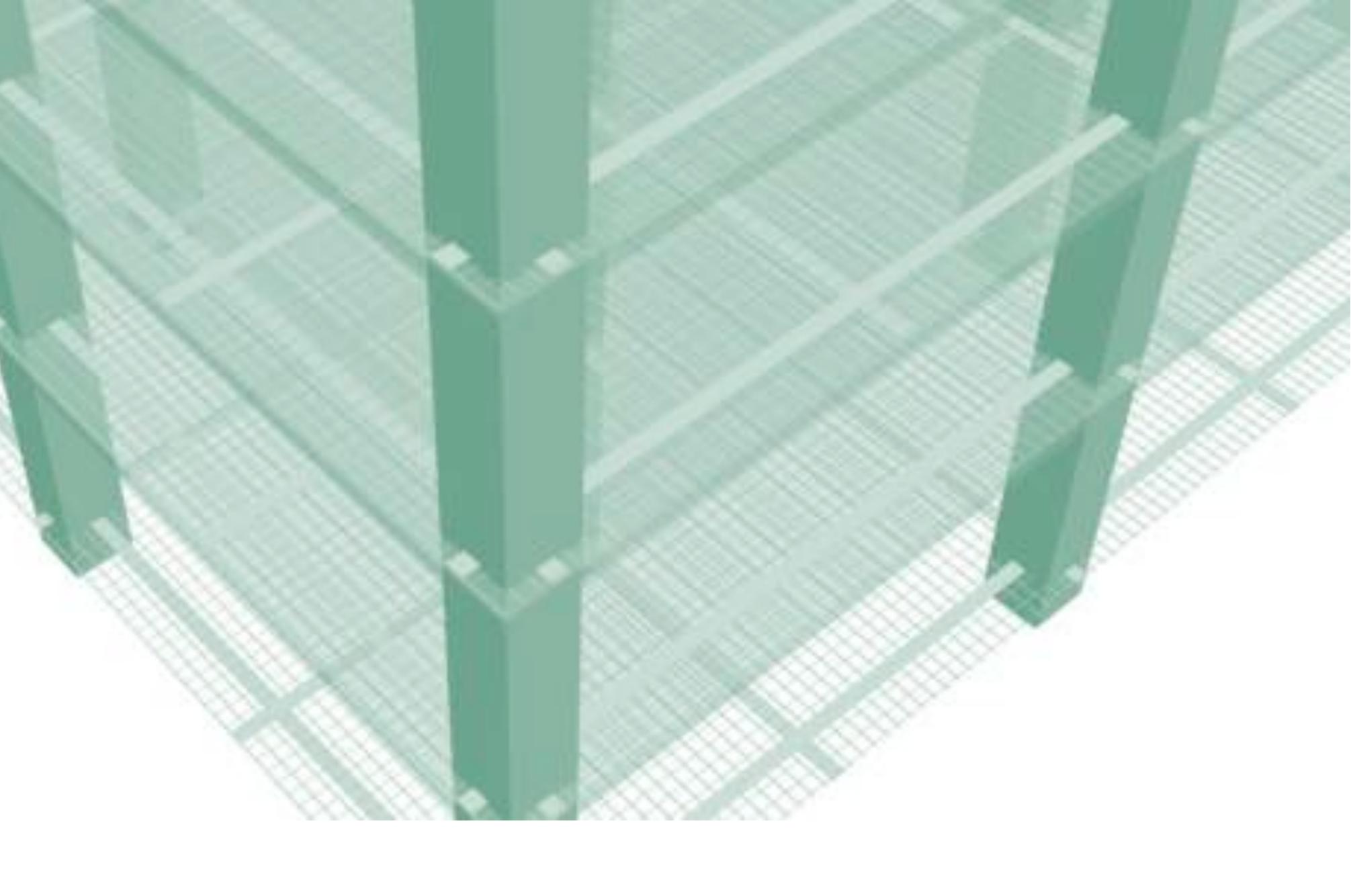










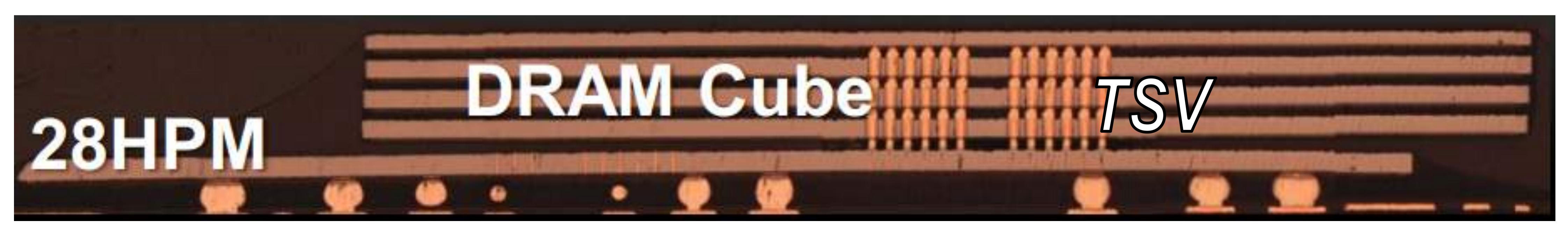




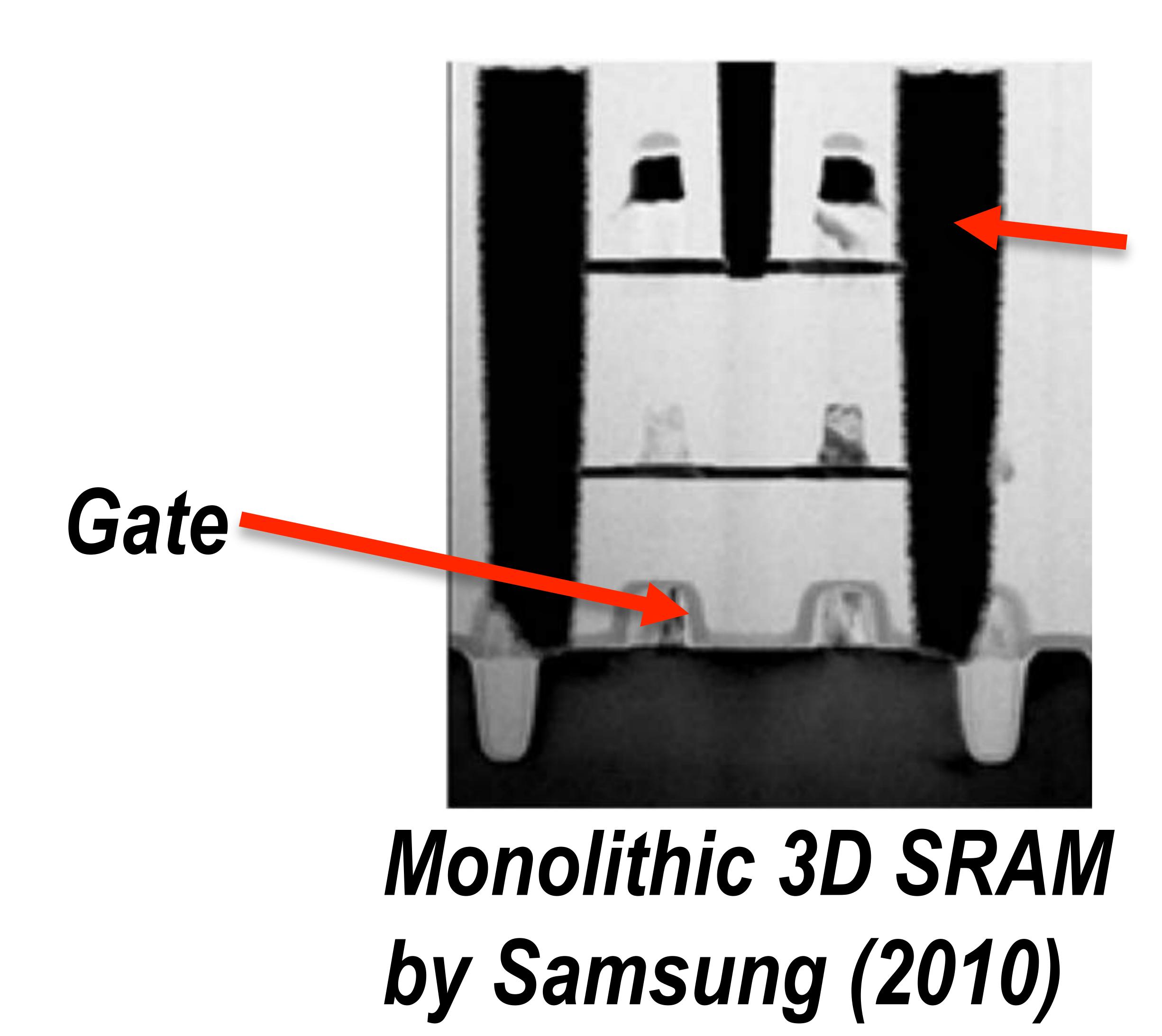


## Monolithic 3D ICs – An Emerging 3D Technology

### **Conventional TSV-based 3D**



### Monolithic 3D



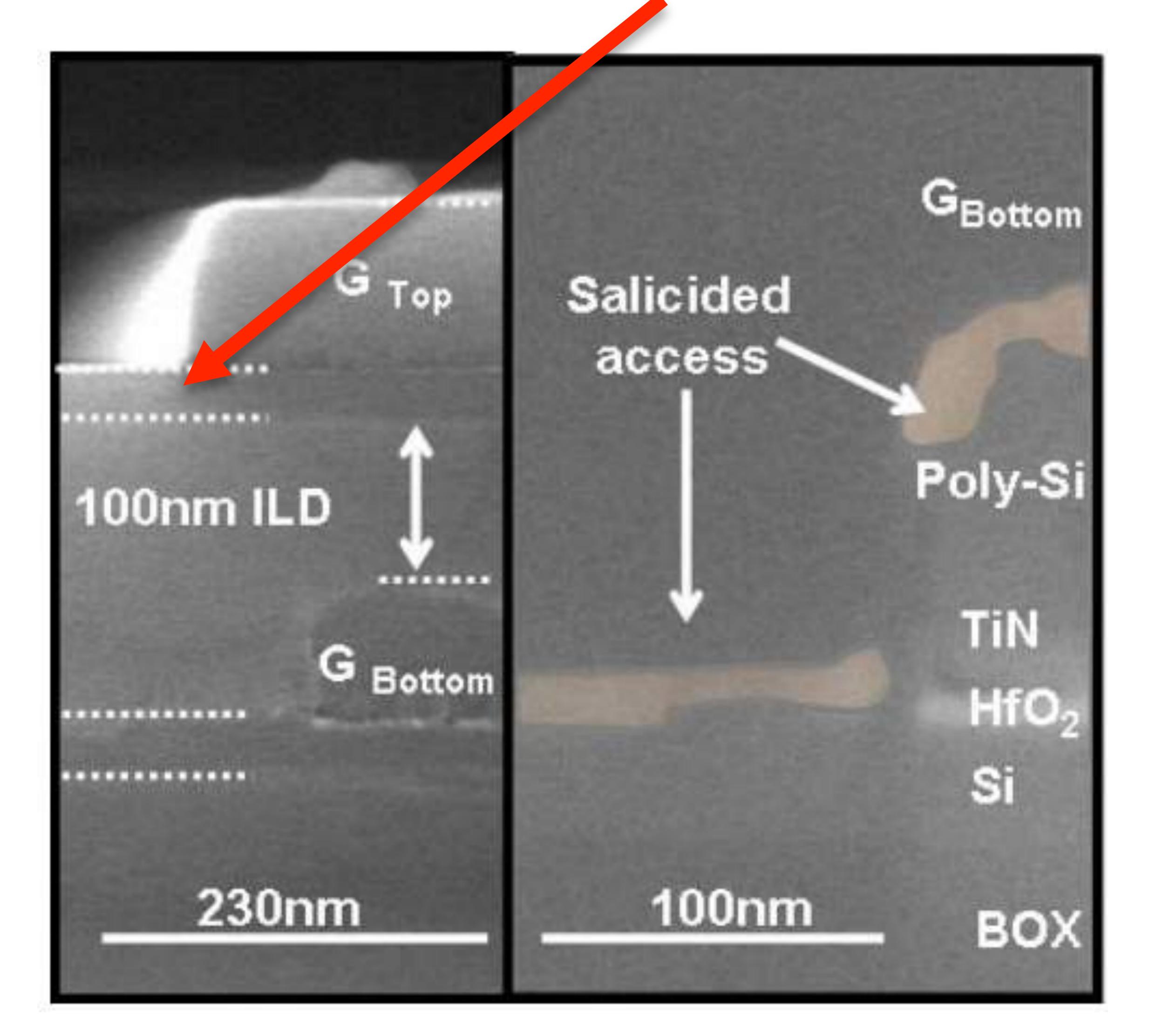
### **TSMC (ISPD 2014)**

### Monolithic inter-tier via







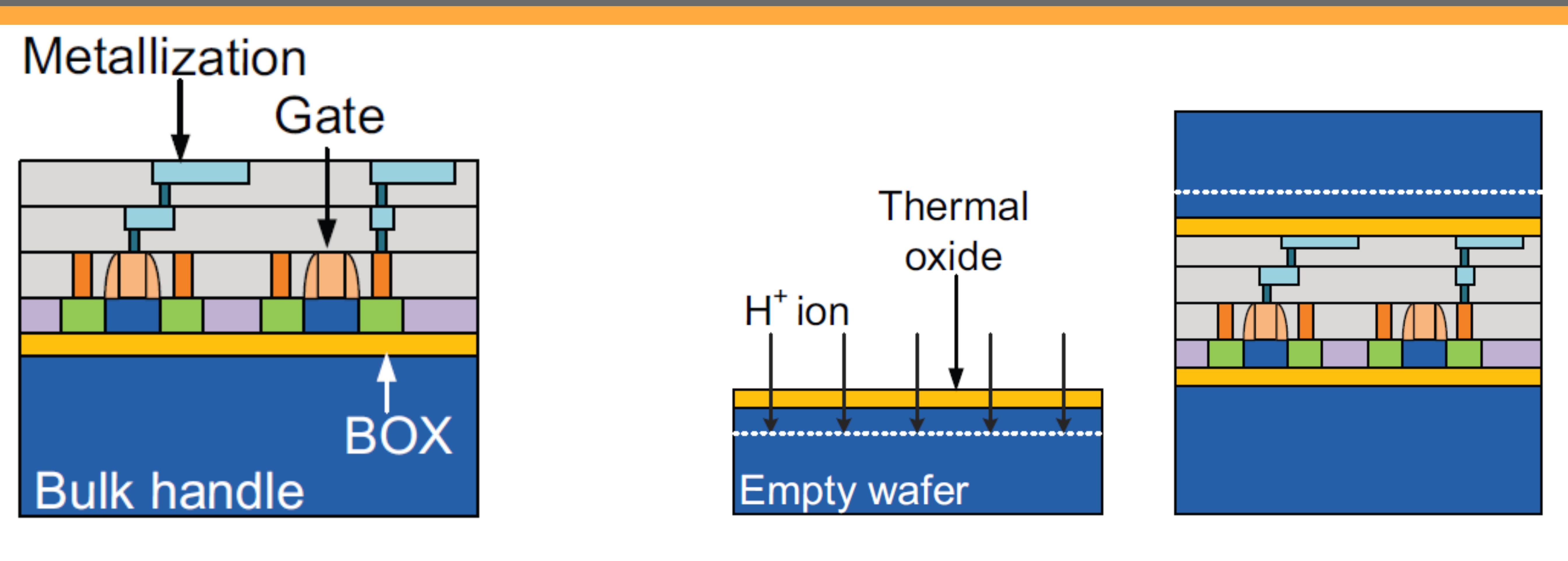


Monolithic 3D for general logic by LETI (2011)



### TSV Size = 5-10umMIV Size = 0.07 - 0.1 um

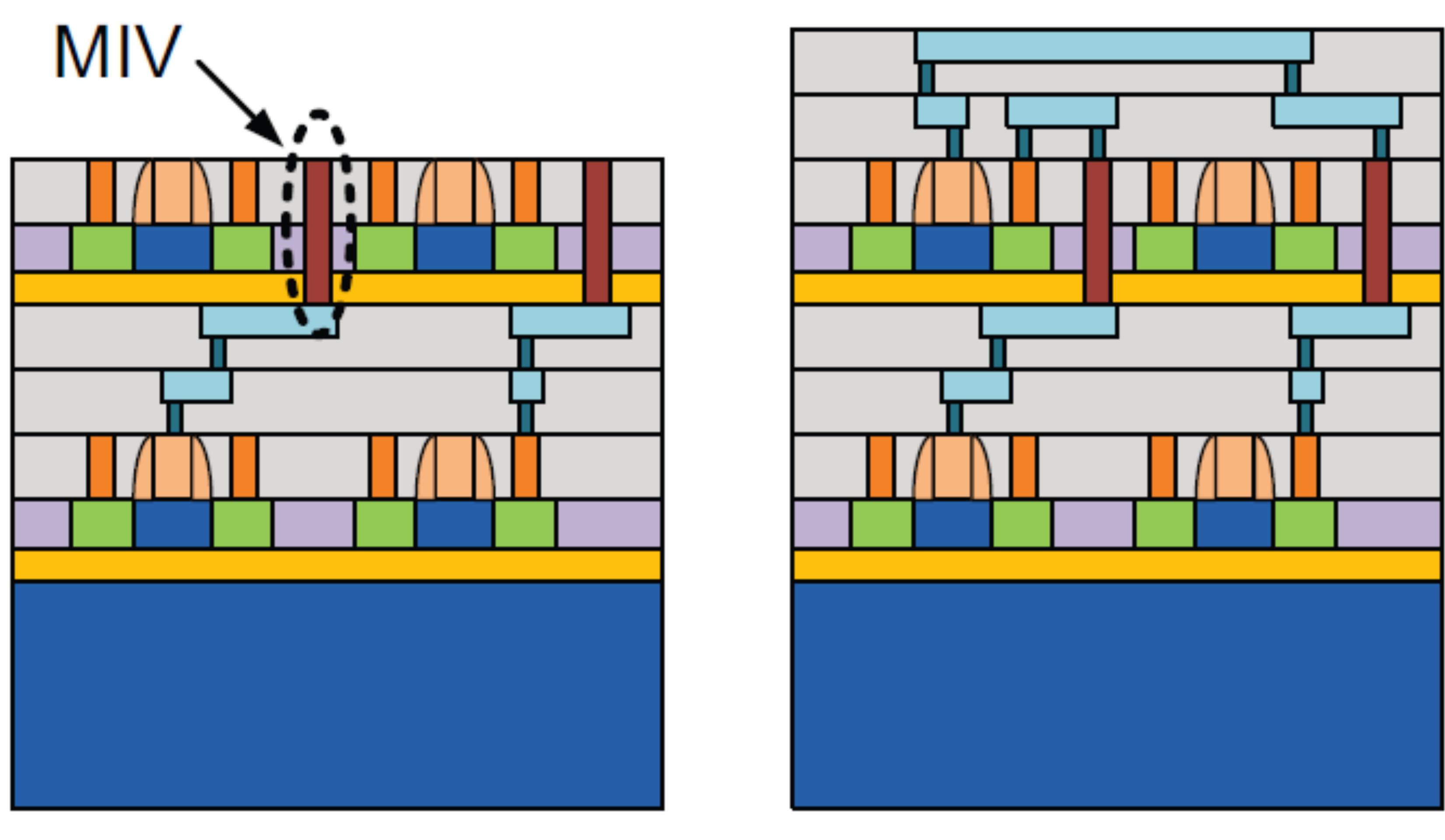
## High quality thin silicon



### **Bottom tier is created as usual**

### MIV = Monolithic-Inter Tier VIA

## Monolithic 3D ICs – Fabrication Process

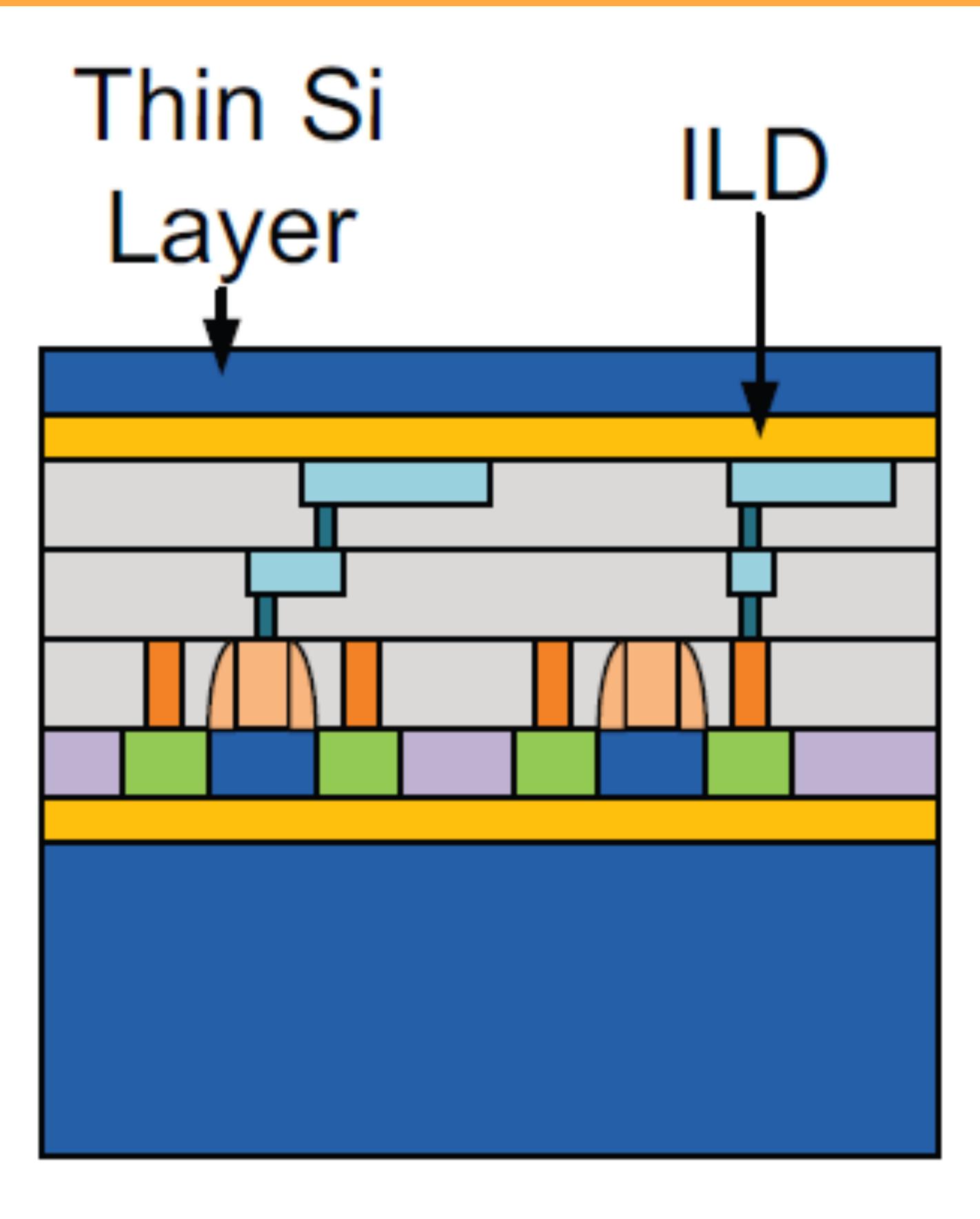


### Fabricate top tier devices + interconnects

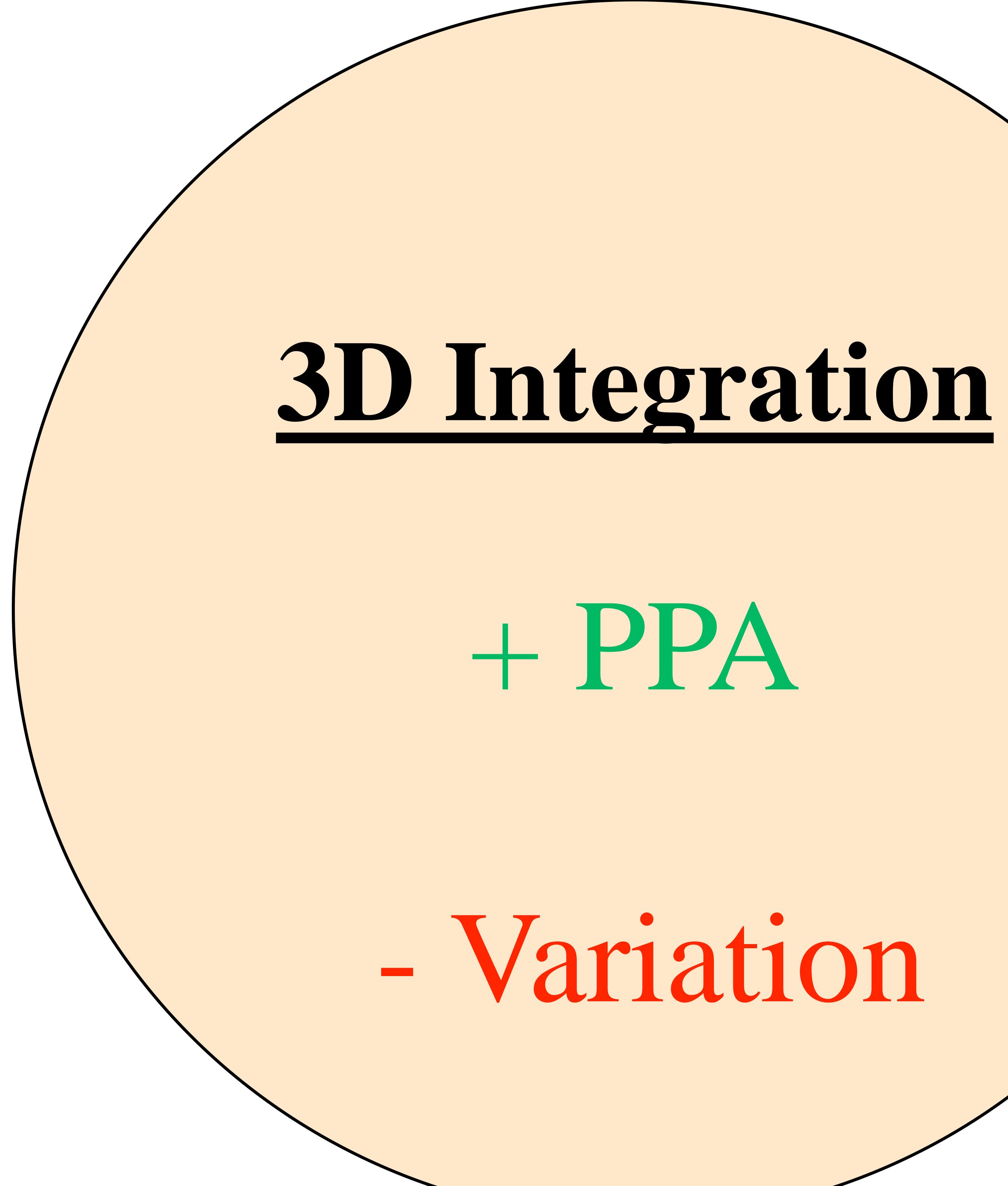
### Thin Si Layer is attached

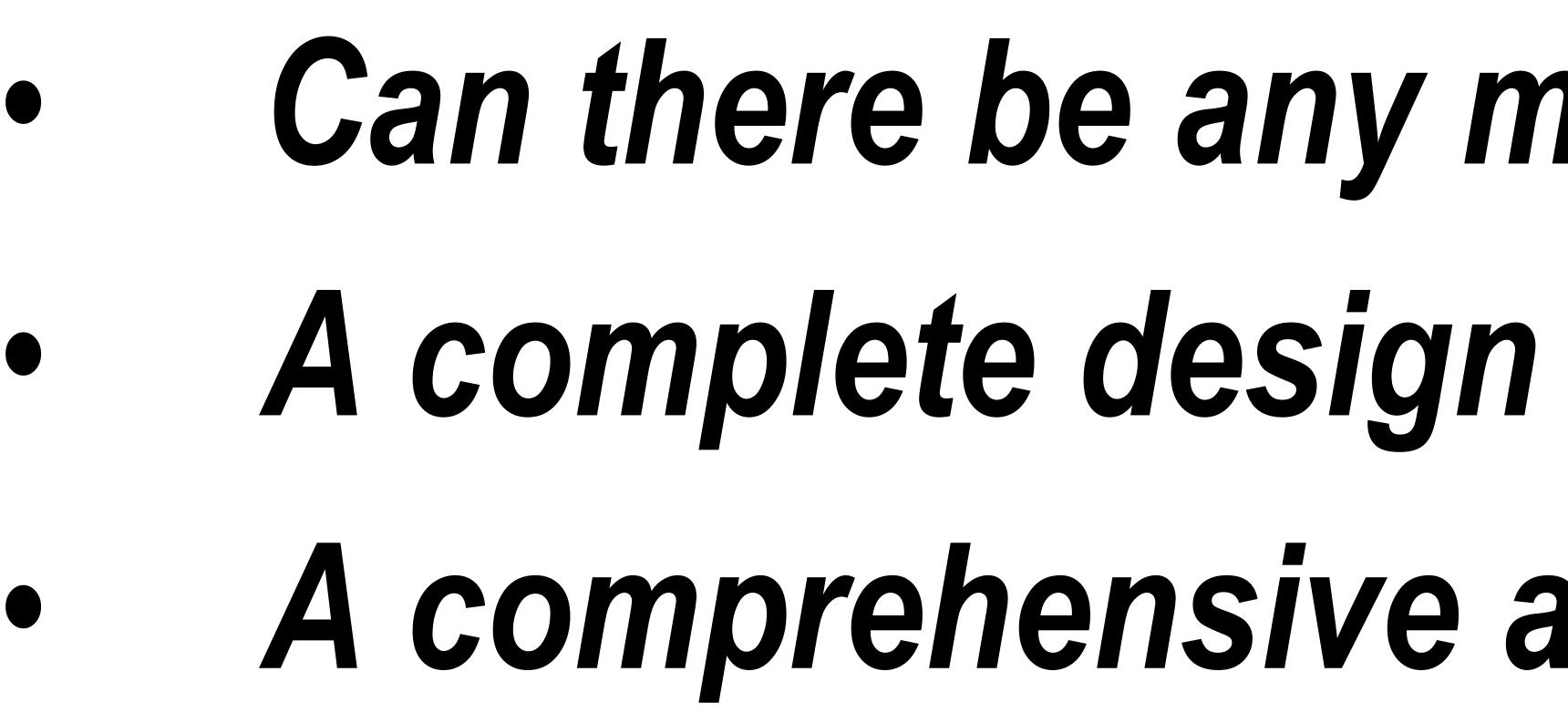


### 3/19



### [1] Batude et. al. IEDN'09

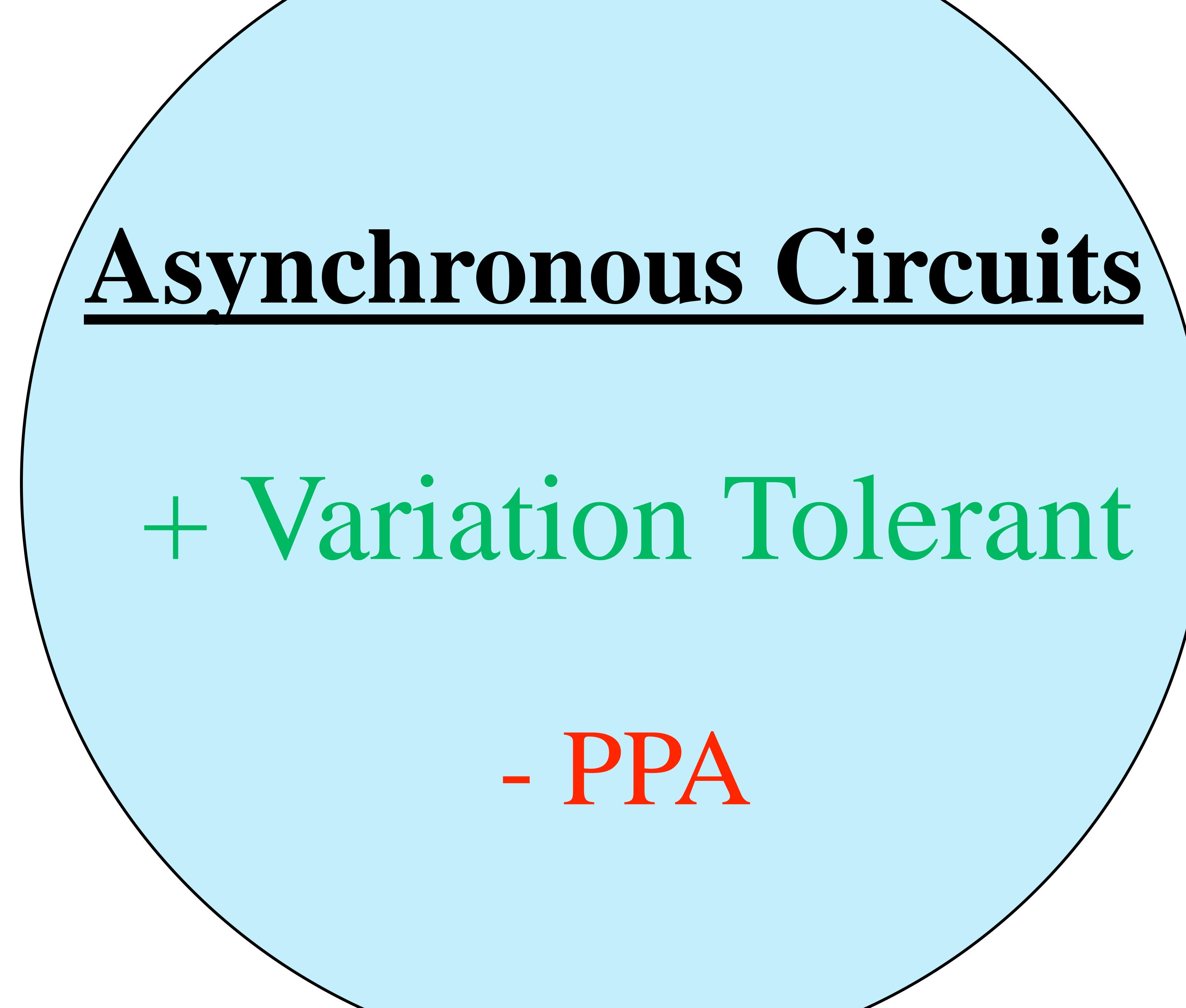




## **Motivation and Objectives**

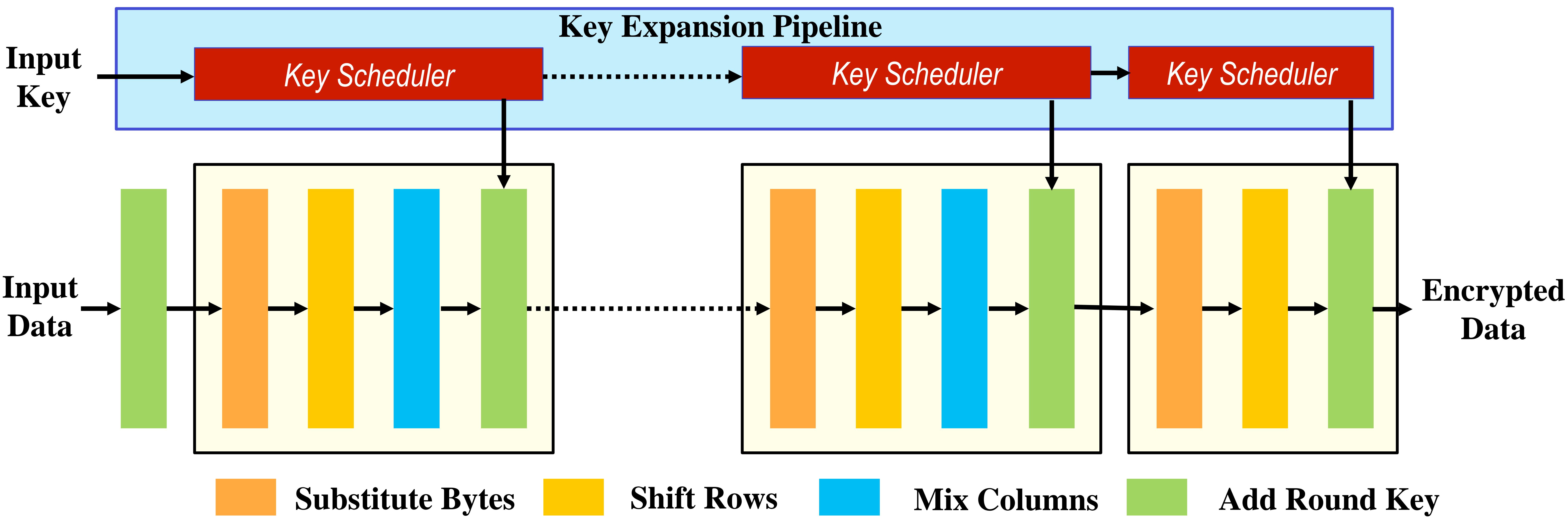
Can there be any mutual benefits by combining them ? A complete design methodology for 3D Integration of Asynchronous Circuits (28-nm PDK) A comprehensive analysis based on GDS II Layouts and standard sign-off flows





4/19

## **Benchmark Design: AES - 128**



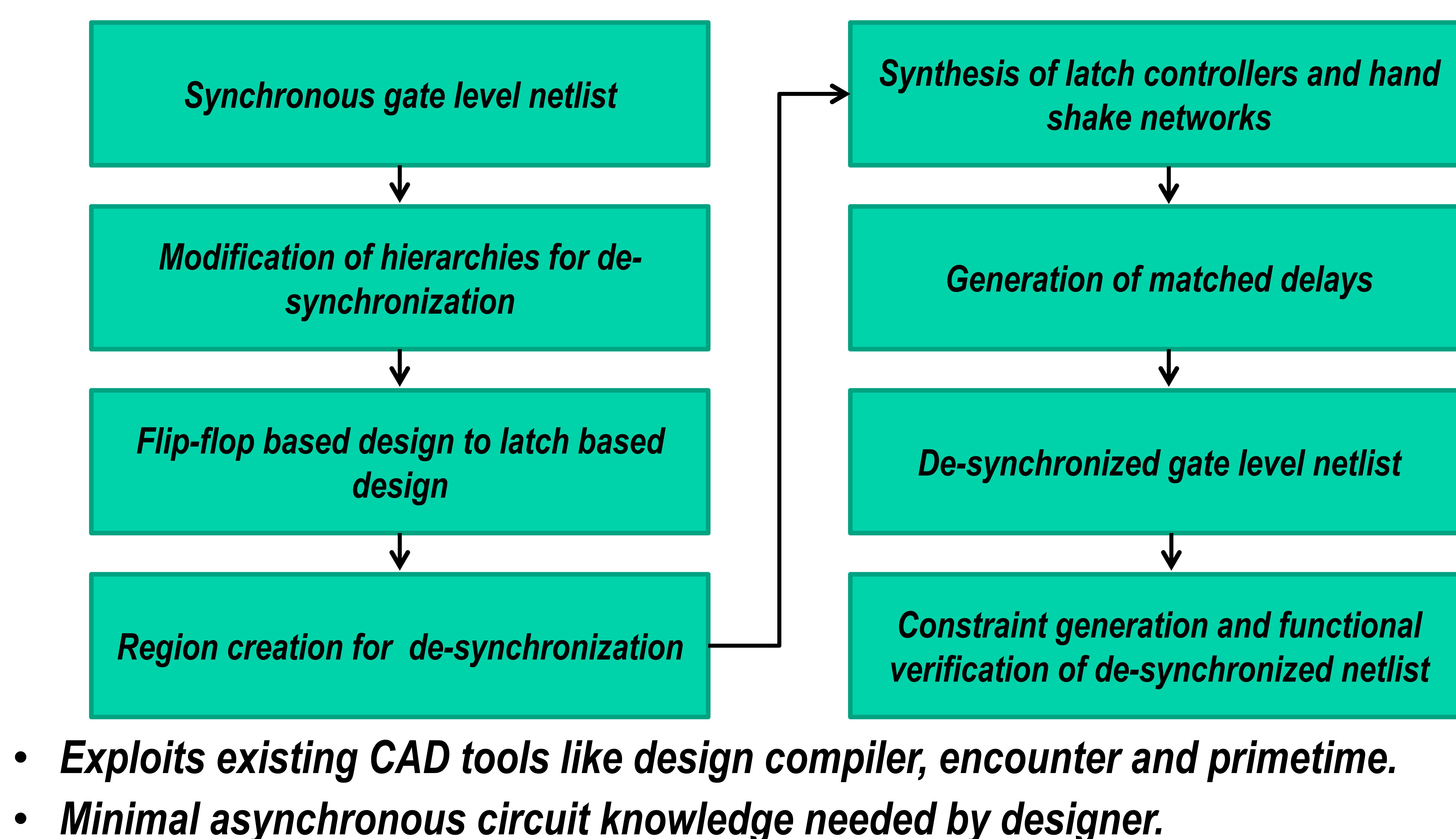
Ubiquitous design 

and use cases.

• Custom RTL for high speed (41 stage – 4Ghz). Verification methodology

Variety of implementation architectures

## **De-synchronization Flow Overview**

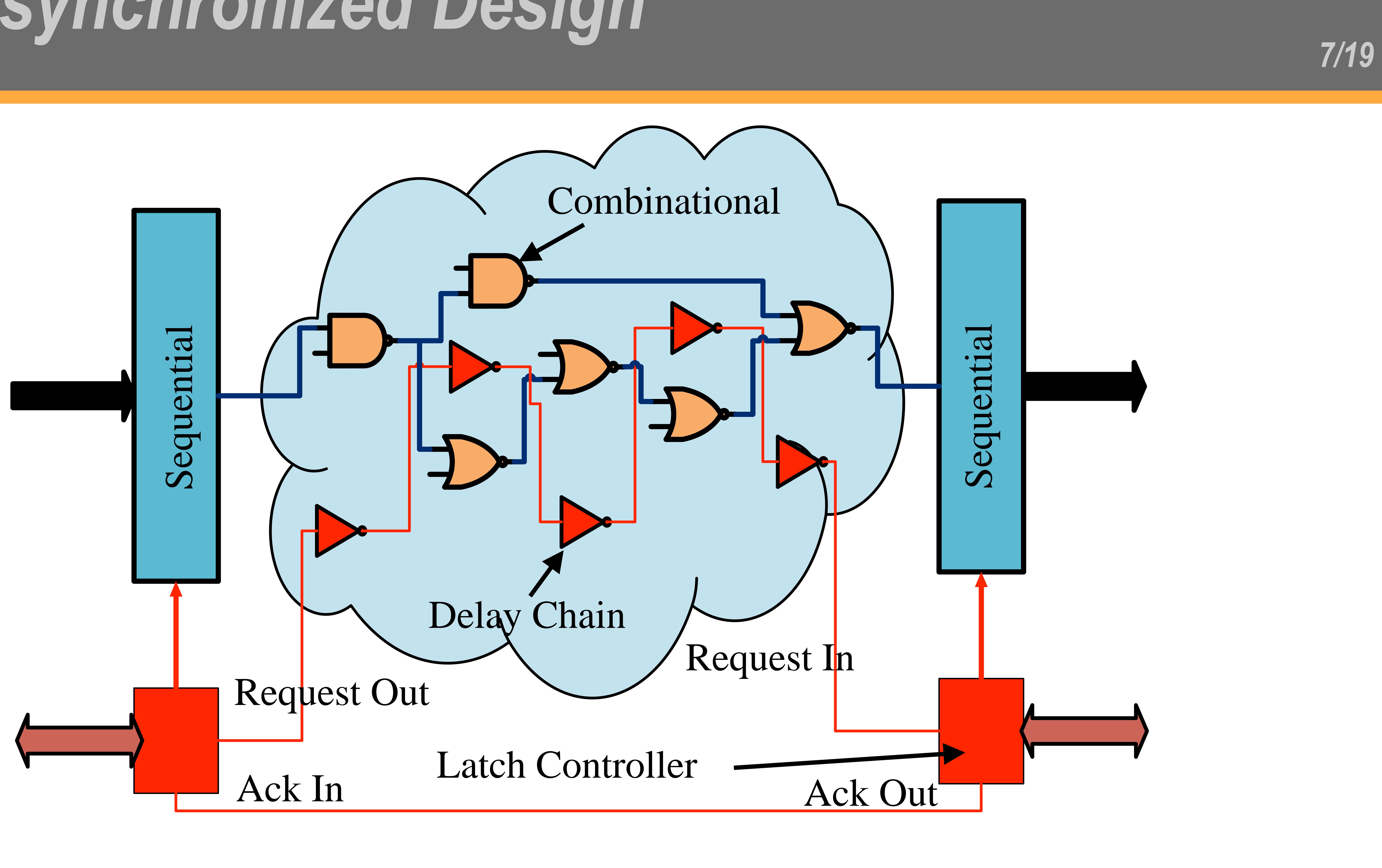


# Desynchronized circuits proved to be variation tolerant.

J.Cortadella et. Al, IEEE Trans on Computer Aided Design of Integrated Circuits & Systems, 2006



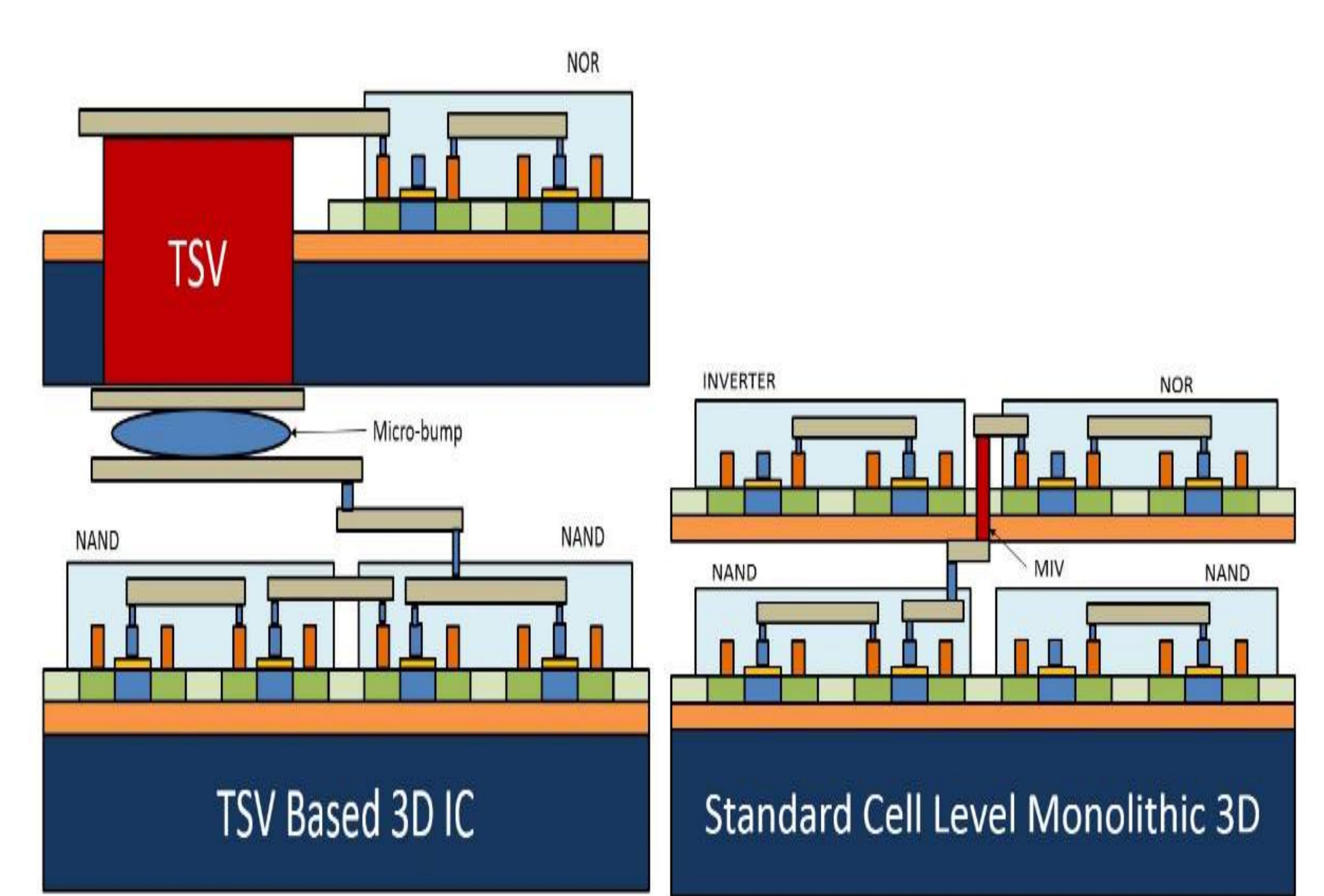
## De-synchronized Design



• 2-phase latch controllers are used in the de-synchronization flow.

## **3D Integration TSV vs Monolithic**

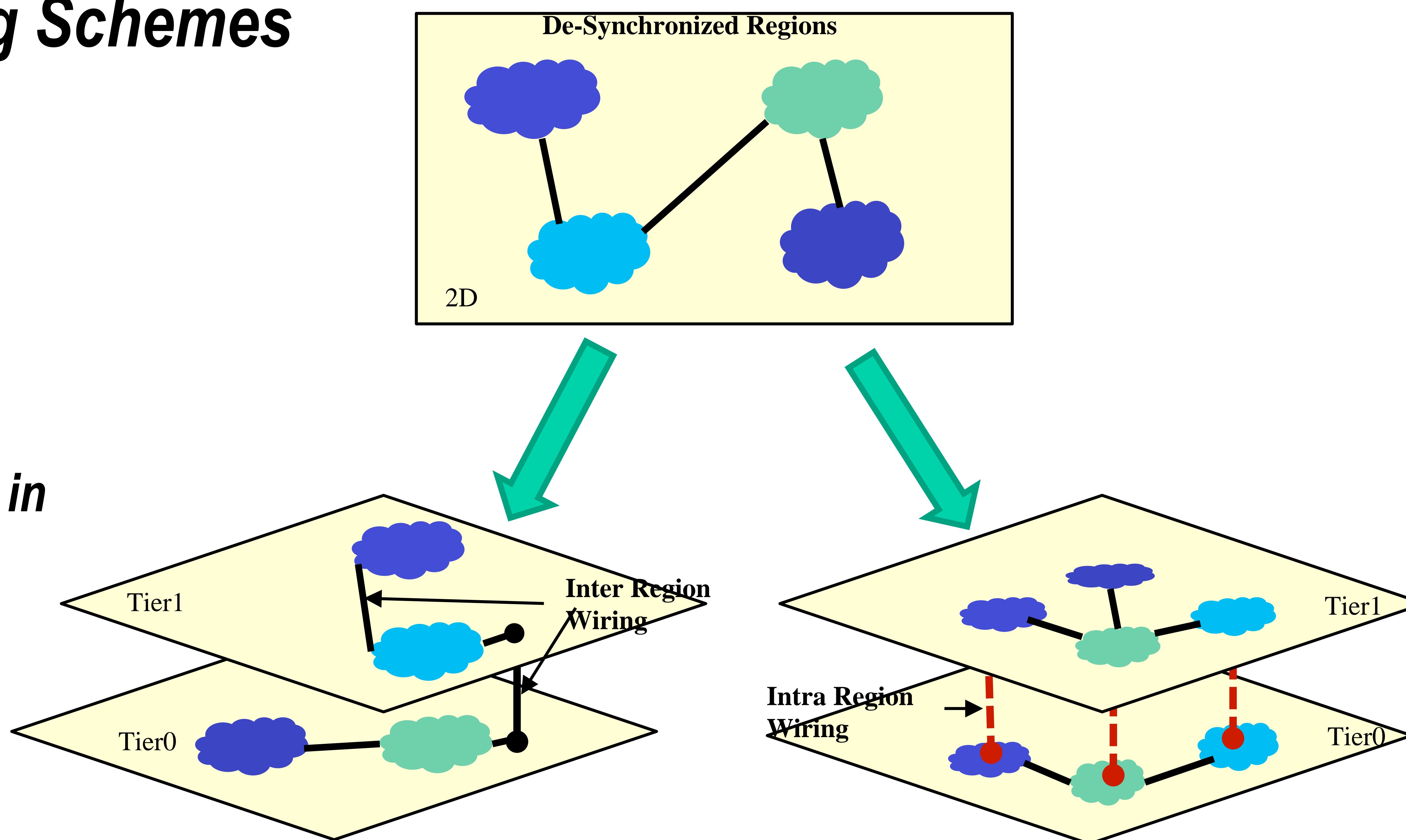
- Initial tests with both integration styles.
- 2D De-synchronized designs have about 15% area overhead due to additional circuitry.
- TSV integration adds to this overhead limiting the number of tier-tier connections.
- Higher integration density can be achieved through monolithic integration.



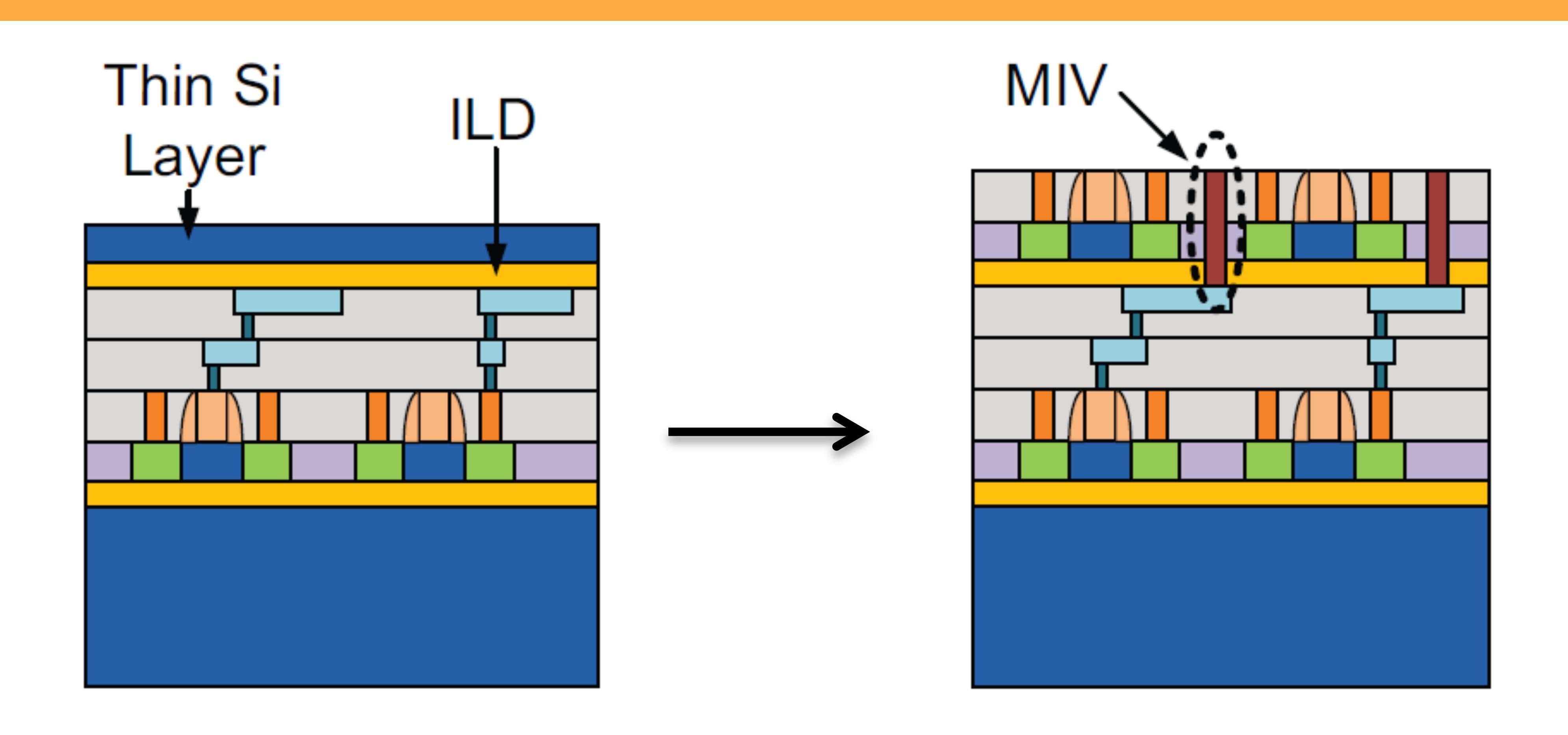
## **3D Integration TSV vs Monolithic**

## Folding/Partitioning Schemes

- 2 partitioning schemes examined.
  - Inter and Intra region partitioning.
- Simple mincut at top (Inter) resulted with 230 TSVs ; result in negative benefits.
  - Intra region folding is done through monolithic with MIV counts resulting up 41K.
- Intra region resulted in better PPA



## M3D: Source of Inter-tier Performance Variation



## Process improvement: < 625C without performance loss $\rightarrow$ still too high for Cu interconnect **Preventing damage to interconnects – Two options:** $\sim$ 400C processing on the top tier $\rightarrow$ Worse transistors on the top tier • Identical interconnects on both tiers (PMOS ~27% & NMOS ~16%) - Use Tungsten (W) on the bottom tier $\rightarrow$ Worse interconnects on bottom tier Identical devices on both tiers (3.1x bulk resistivity compared to Cu)

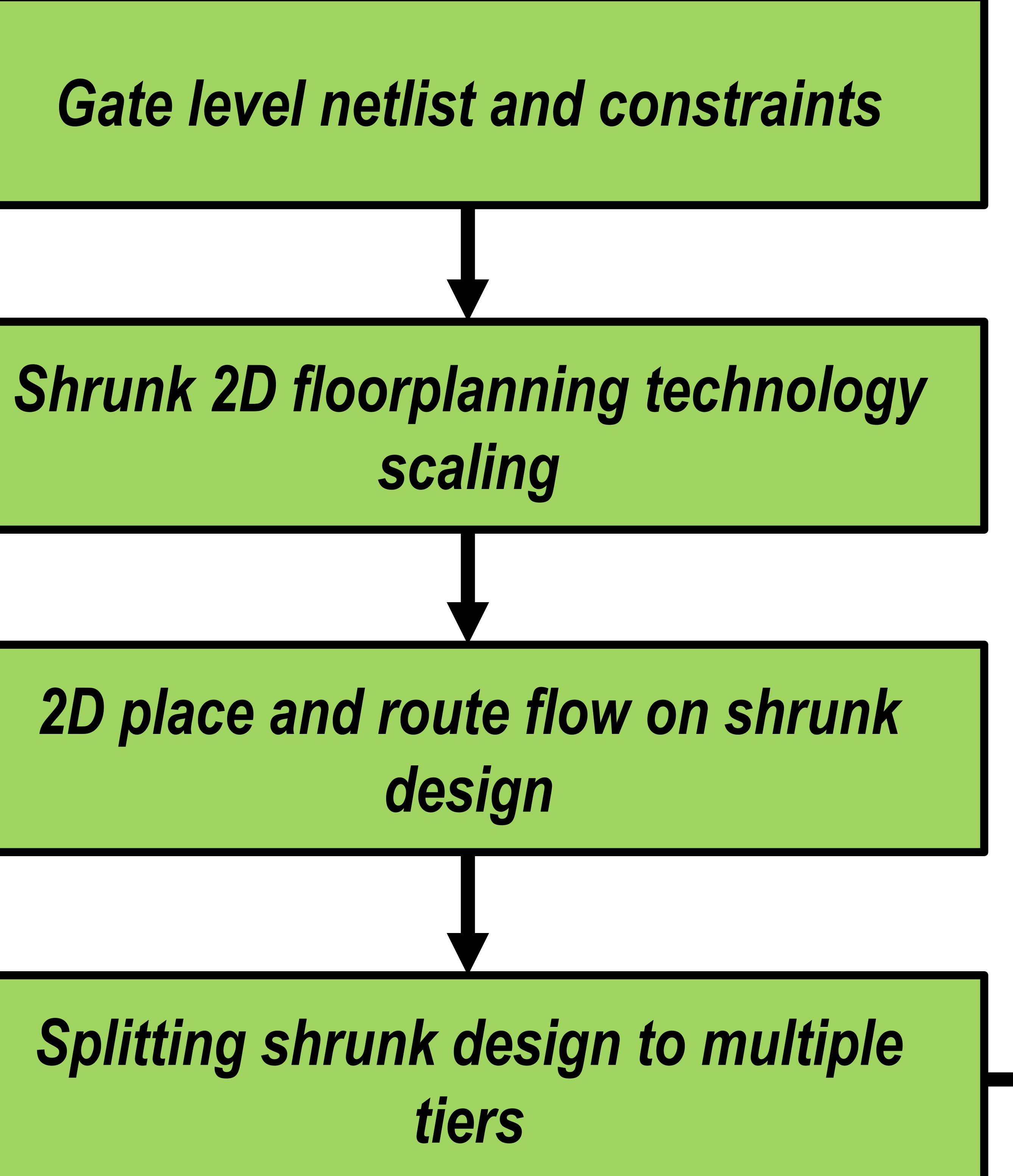
[1] Batude et. Al, Journal on Emerging and Selected Topics in Circuits and Systems, 2012 [2] Batude et AI, 3D Monolithic Integration, ISCAS 2011 [3] Improvements in low temperature ( < 625C ) FDSOI devices down to 30nm gate length, VLSI 2012 [4] Low Thermal Budget Processing for Sequential 3-D IC Fabrication, Rajendran et. AI, TED 2007

### FEOL processing of top tier • **RTA at 1200C will damage both** devices and interconnects



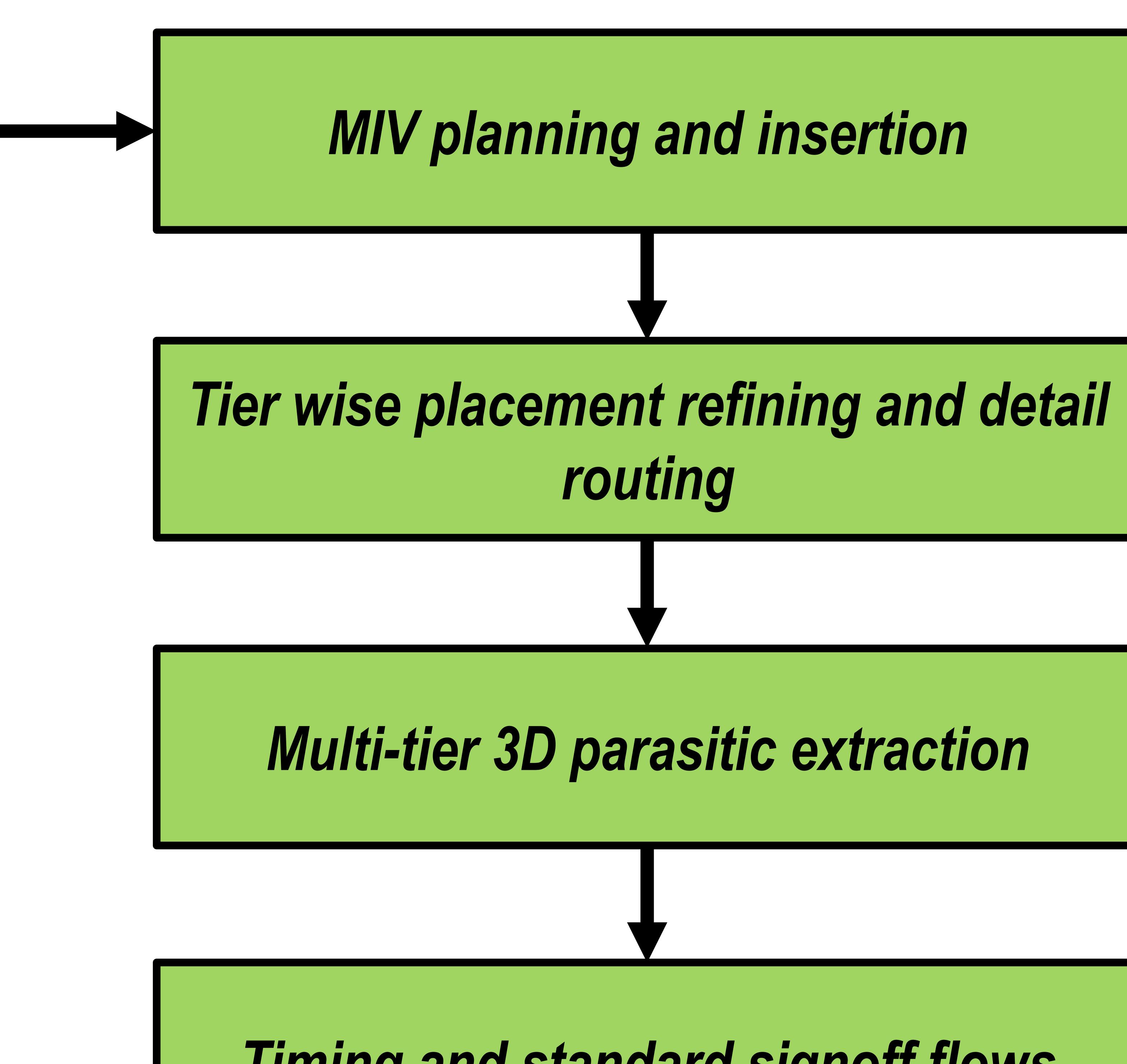
## Monolithic 3D IC Design Flow

 A 2 Tier Design • Uses Cadence Encounter and in house custom scripts



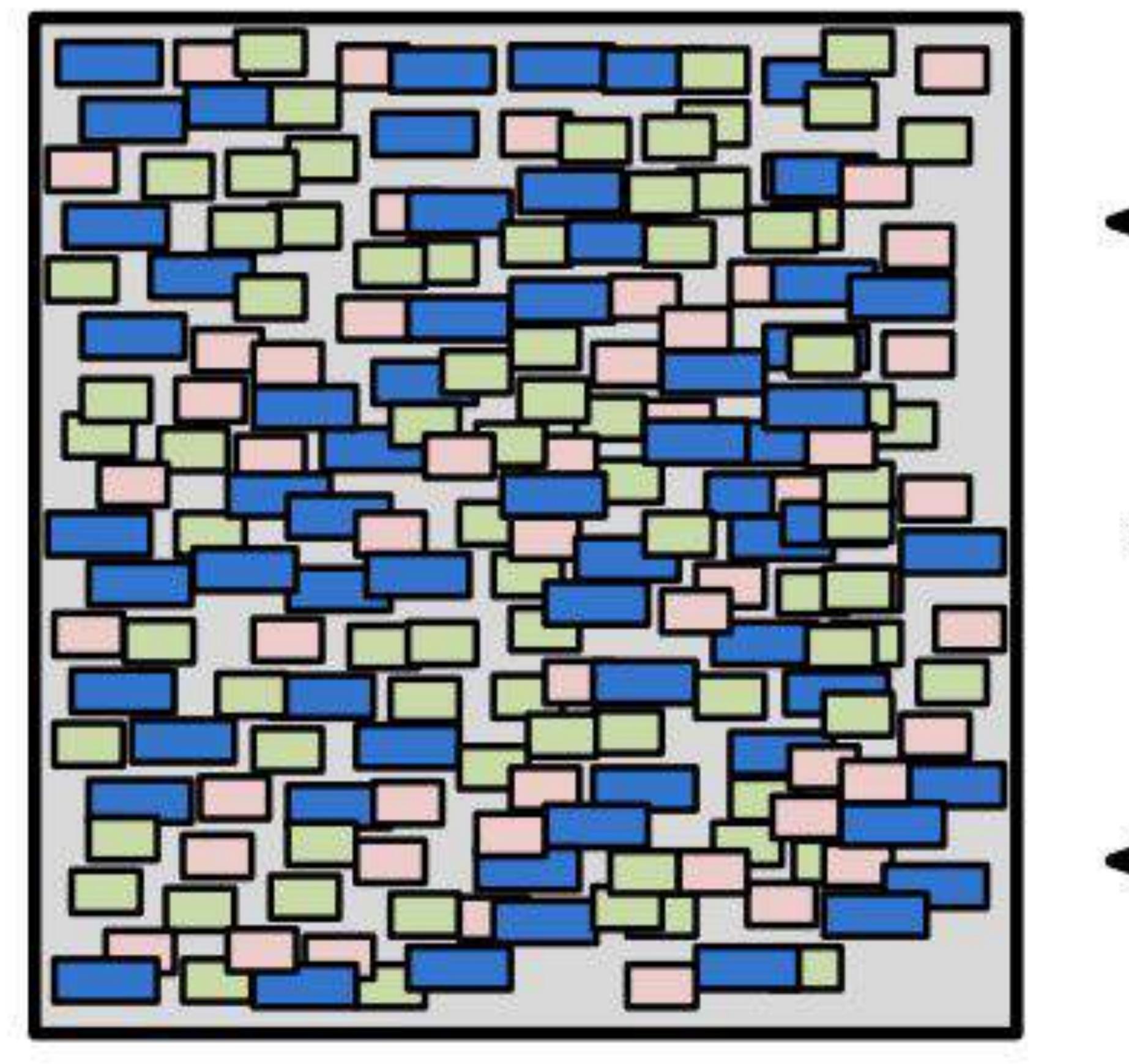
S.Panth et.AI, Placement-Driven Partitioning for Congestion Mitigation in Monolithic 3D IC Designs, ISPD 2014







### Timing and standard signoff flows



## Flattened 3D Placement

Shrink chip width to half the area.

objective on to multiple tiers.

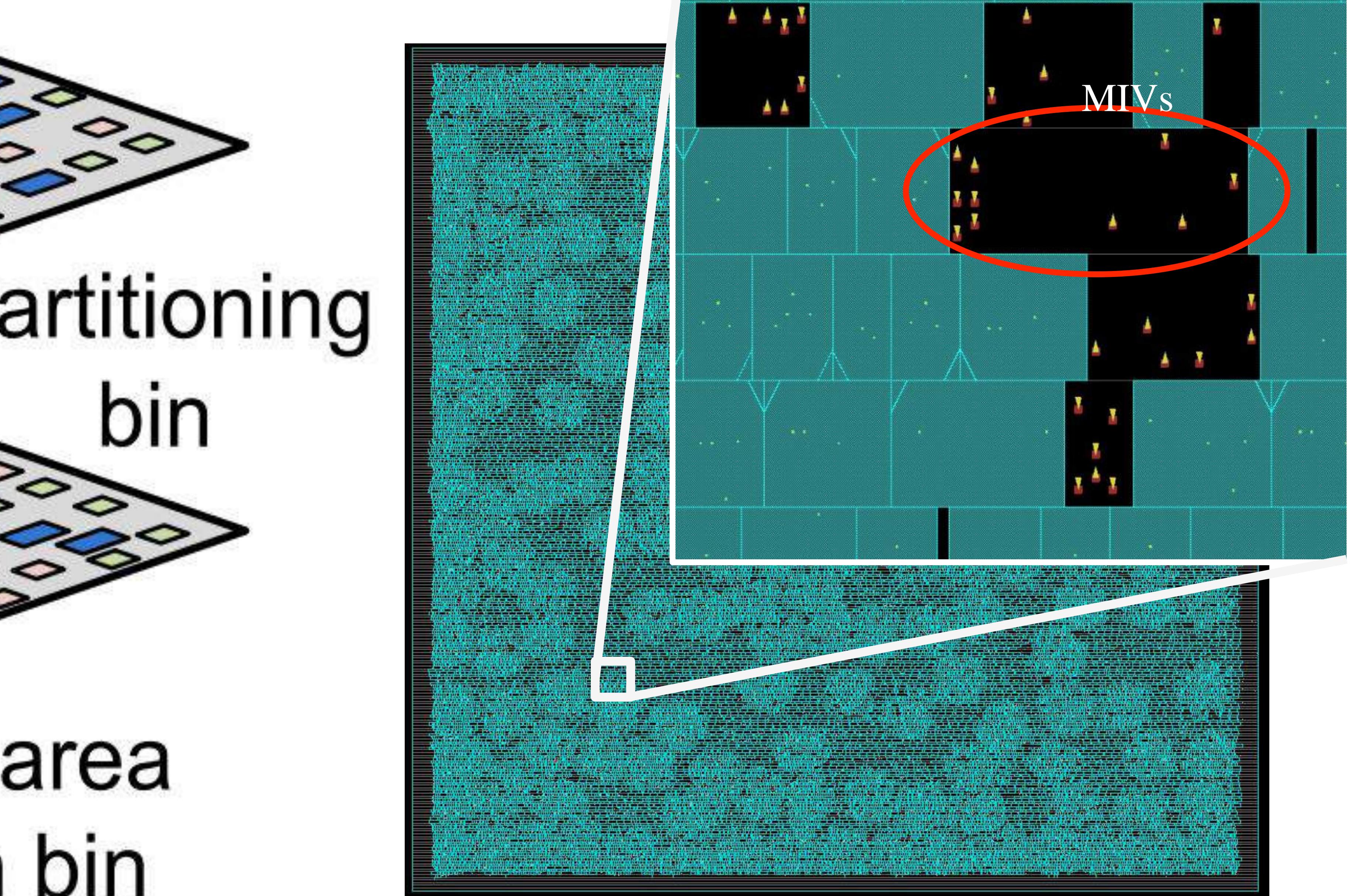
## **Nonolithic 3D Placement**

## Repartition with area balance in each bin

## Shrink cells with a scaling factor of 0.707 on a shrunk die. **Compressed cells are exploded to original sizes with area balance**

S.Panth et.Al, Placement-Driven Partitioning for Congestion Mitigation in Monolithic 3D IC Designs, ISPD 2014





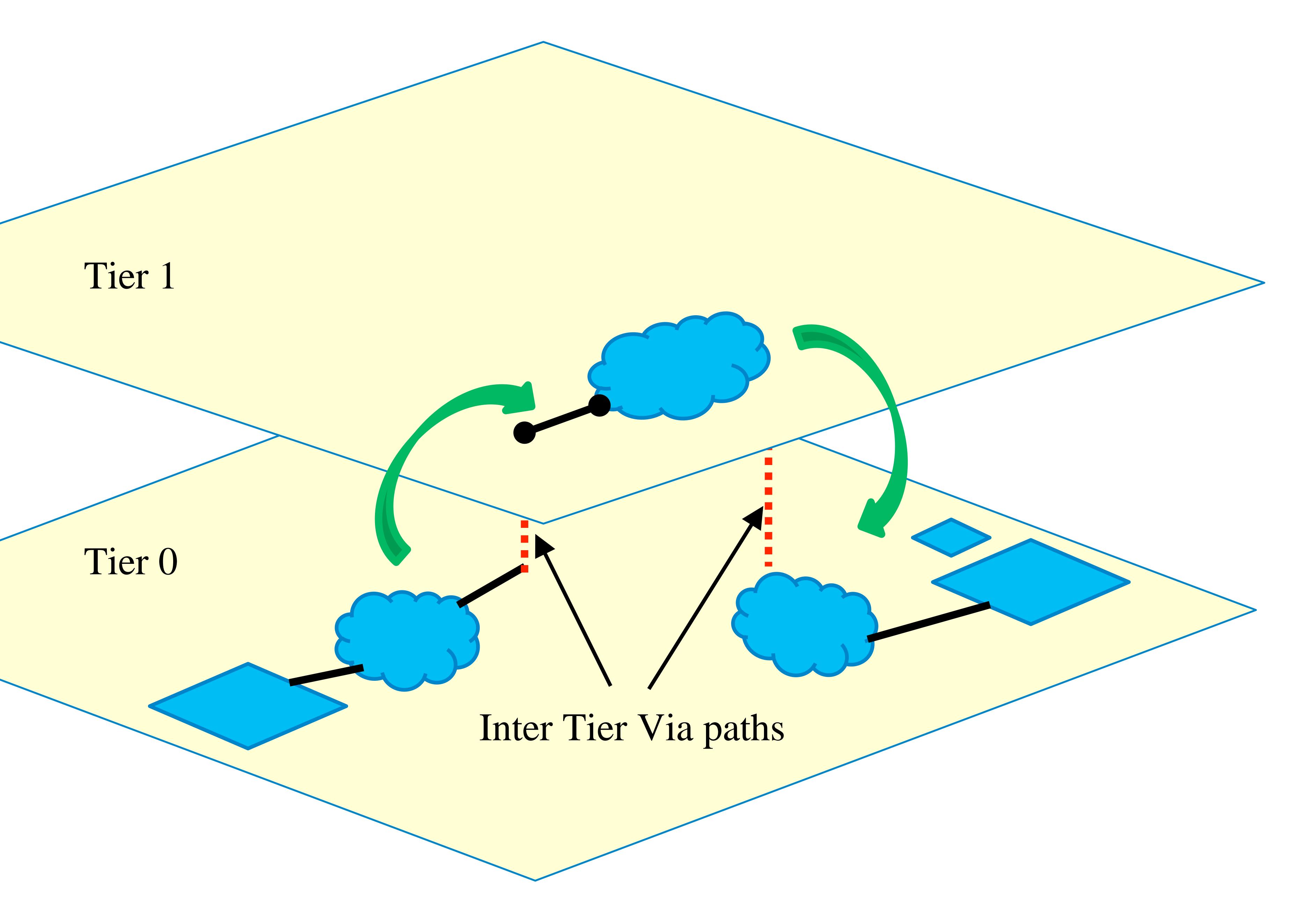
## **3D Integration of Delay Chains**

### Snaking paths timing optimization problem.

- Delay chains track inter tier combo logic paths.
- Currently supports only 2 tier designs.
- Delay elements keep track of variation on each die closely.





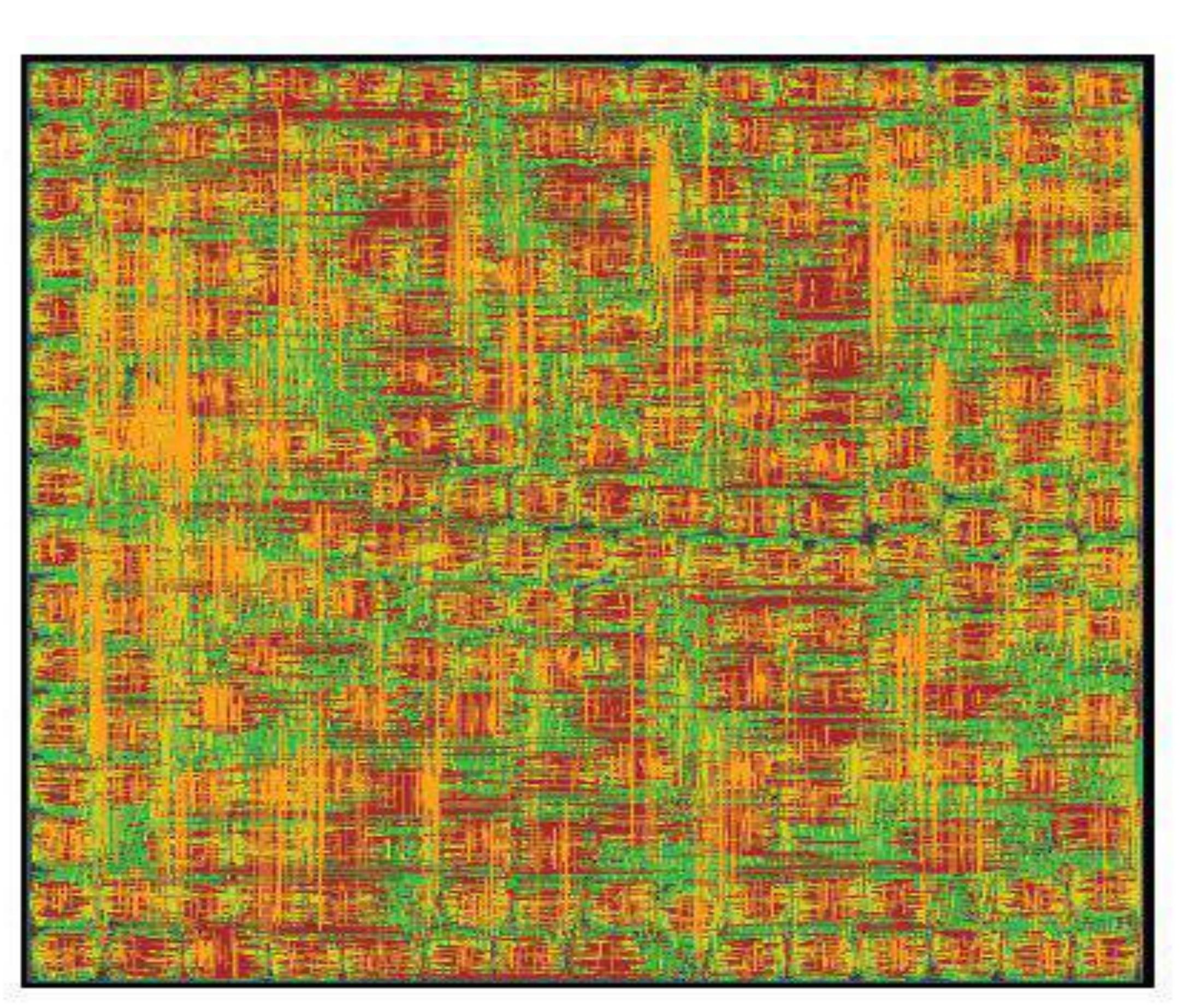


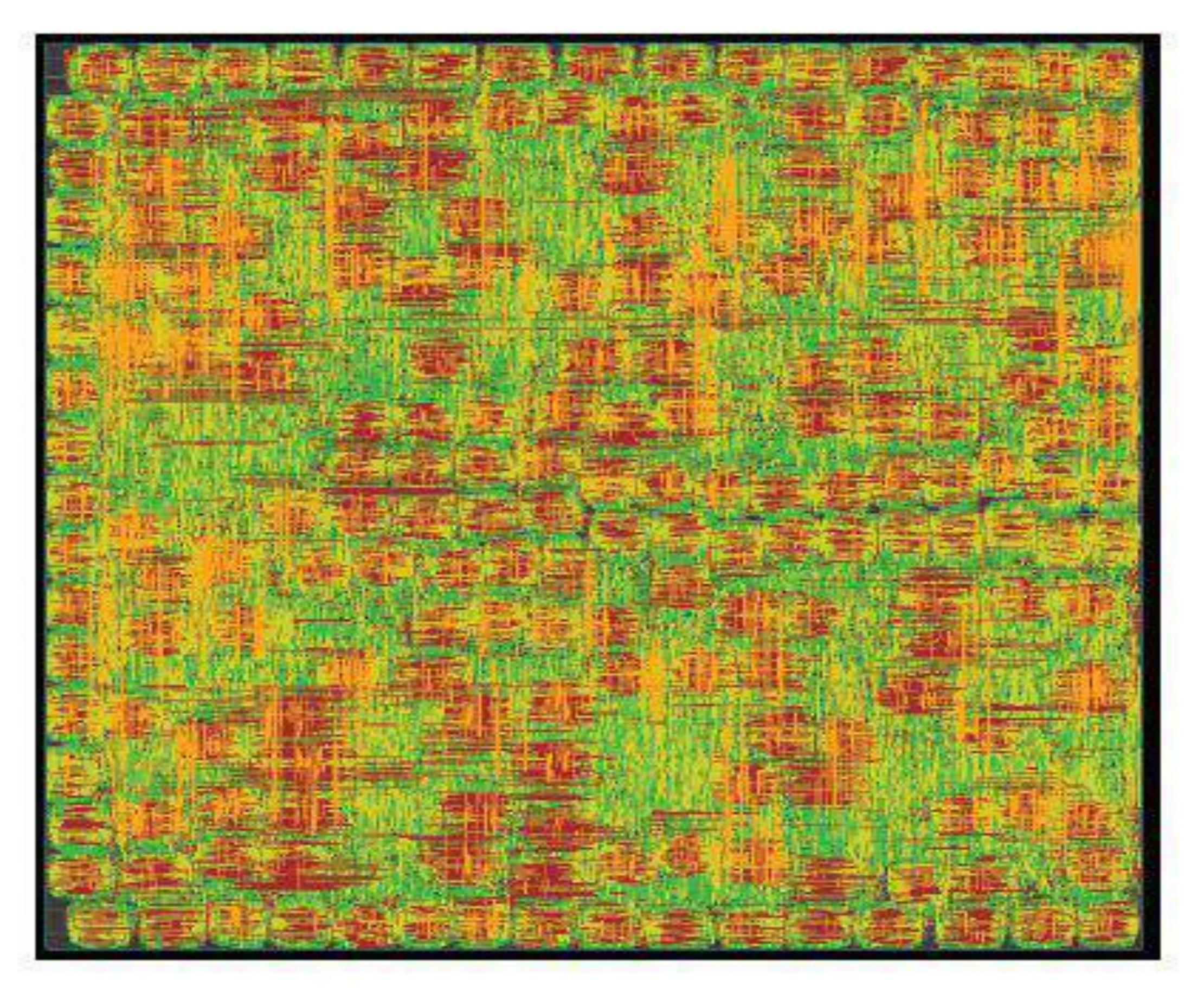
13/19

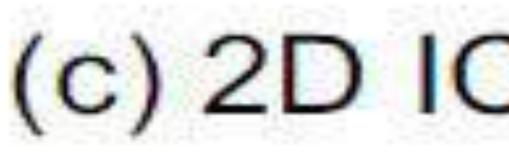
## **Die-Shots of Final Designs**

### **GDSII Layouts of 2D and 2-tier** 3D synchronous and desynchronized AES designs.

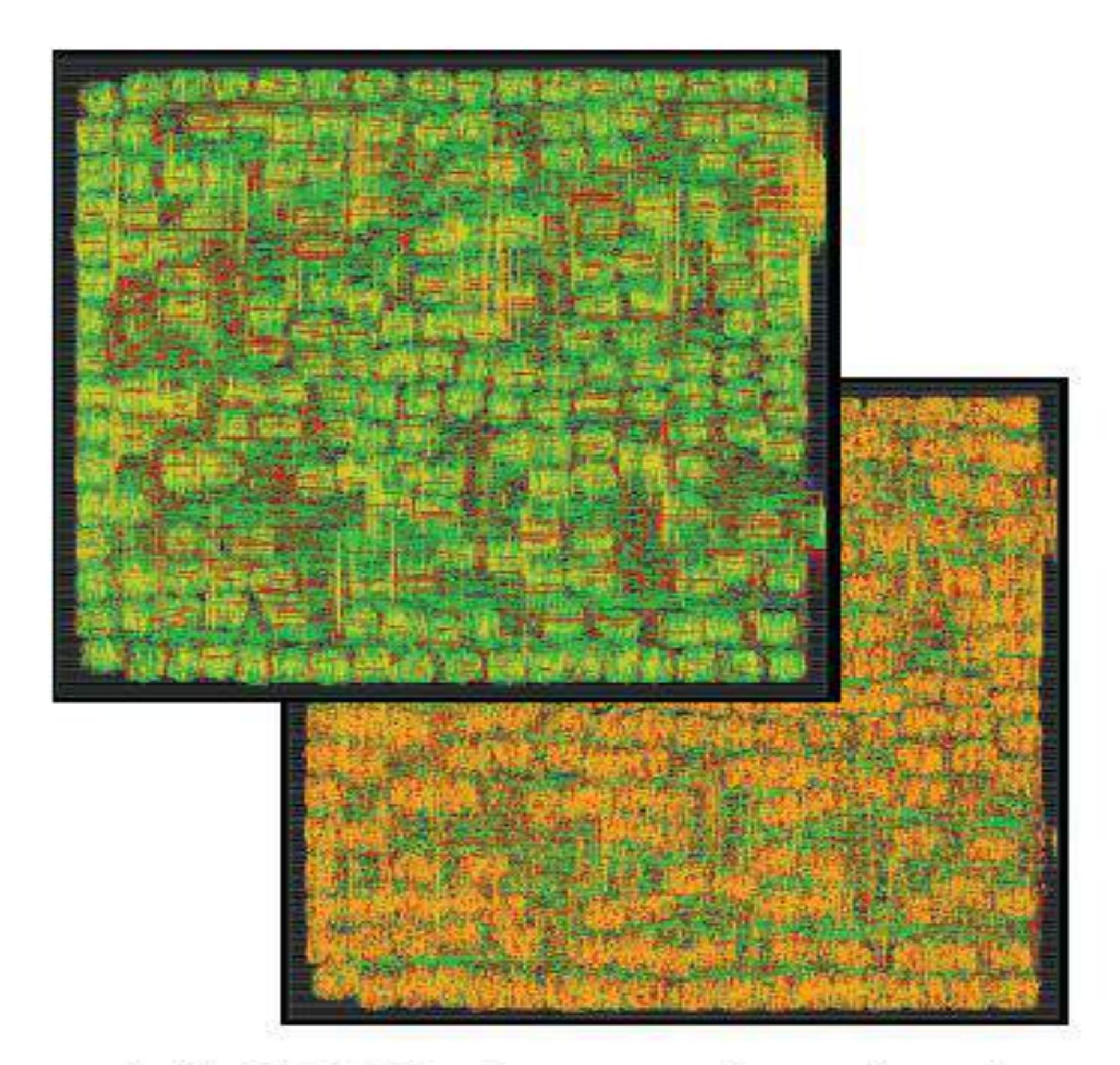
- 2D footprint is 710x710um, and 3D is 500x500um. We observe that de-synchronous has fewer global interconnects.
- Full chip DRC clean layouts (GDS)
- 3D parasitic extraction



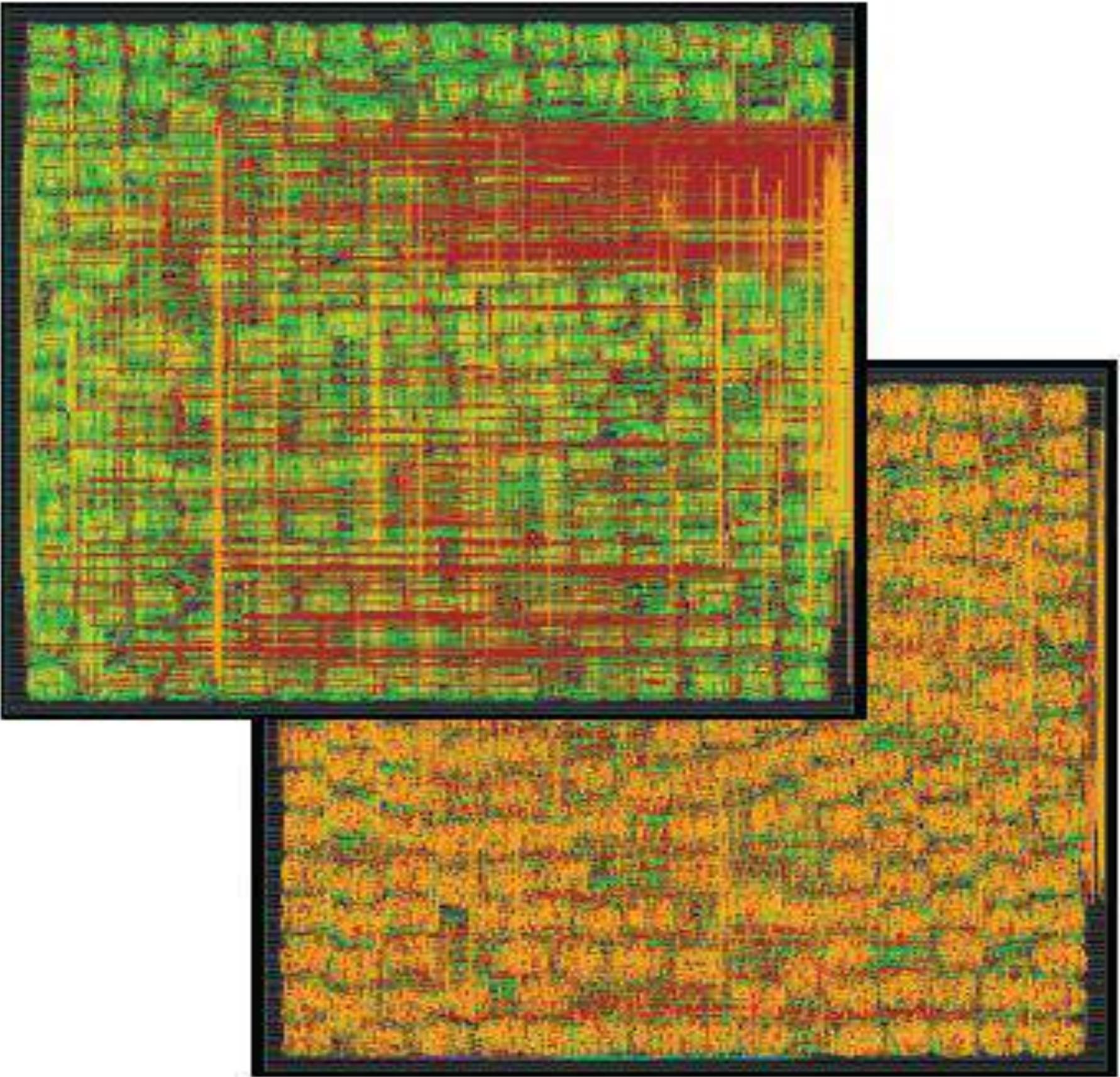


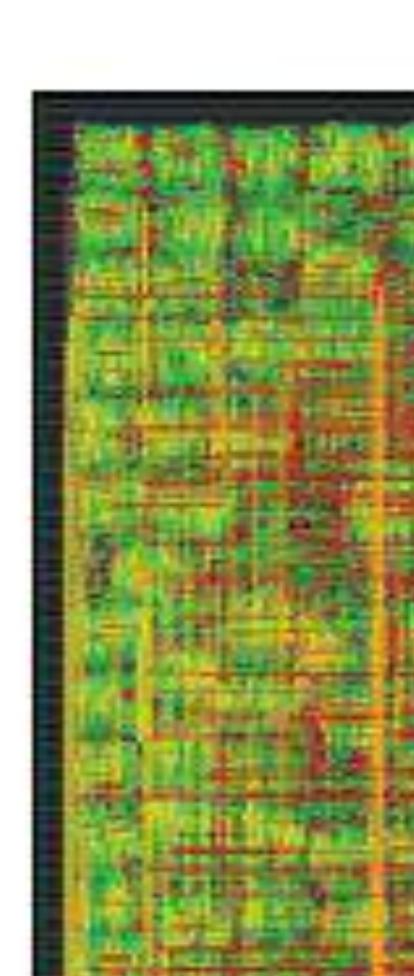


### (c) 2D IC de-synchronized



### (a) 2D IC synchronous





### 14/19

### (b) 3D IC synchronous

### (d) 3D IC de-synchronized

## Verification Methodology & Results

- complicated.
- to validate correctness.

## • Verification methodology for asynchronous circuits is very

## • Primetime based timing analysis is done with post layout parasitic information to generate delay models for GLS.

## • Full functional GLS simulations must be done and assertions written

## • Timing information is extracted based on the simulation waveforms. • A few thousand packets of data is encrypted and corresponding

## activity vectors are generated.

comparison with synchronous system.

## • Vector based power measurement with real work loads used for fair-



## Design Metrics: ISO-Performance (0.25ns)

<i>Metrics</i>	
	2D
footprint (mm2)	0.504
cell area (mm2)	0.400
buffer count	31757
total WL (m)	3.03
avg WL (um)	20.27

Low Latency: Single packet encryption in 3D Synchronous: 10.25ns De-Sync: 6.33ns

<b>Synchronous</b>			
	<b>3D</b>		
	0.25 (-50.3%)		
	0.373 (-6.80%)		
	26440 (-16.7%)		
	2.09 (-31.0%)		
	14.582 (-28.1%)		
	26440 (-16.7%) 2.09 (-31.0%)		

# • Area/WL/buffer count penalty due to de-synchronization.

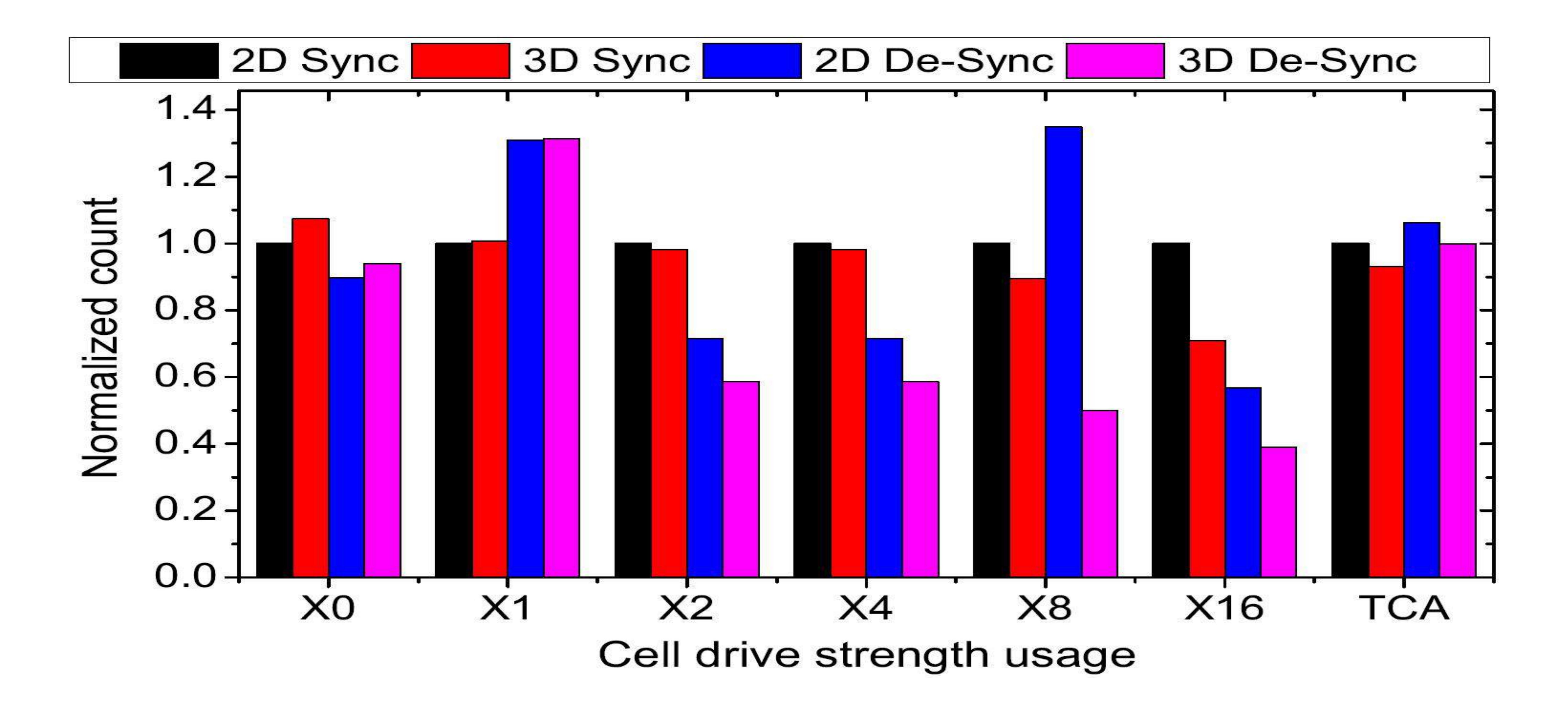
De-synchronous		DeSync 3D vs Sync 2D
<b>2D</b>	<b>3D</b>	%
0.504	0.25 (-50.3%)	-50.3%
0.425	0.399 (-6.06%)	<i>-0.25%</i>
34292	29834 (-13.0%)	-6.05%
3.06	2.01 (-34.3%)	-33.66%
18.20	13.18 (-27.5%)	-34.97%

• 2D Foot print based on synchronous designs. 2D de-sync can have higher utilization.





## Design Metrics: Cell size distribution



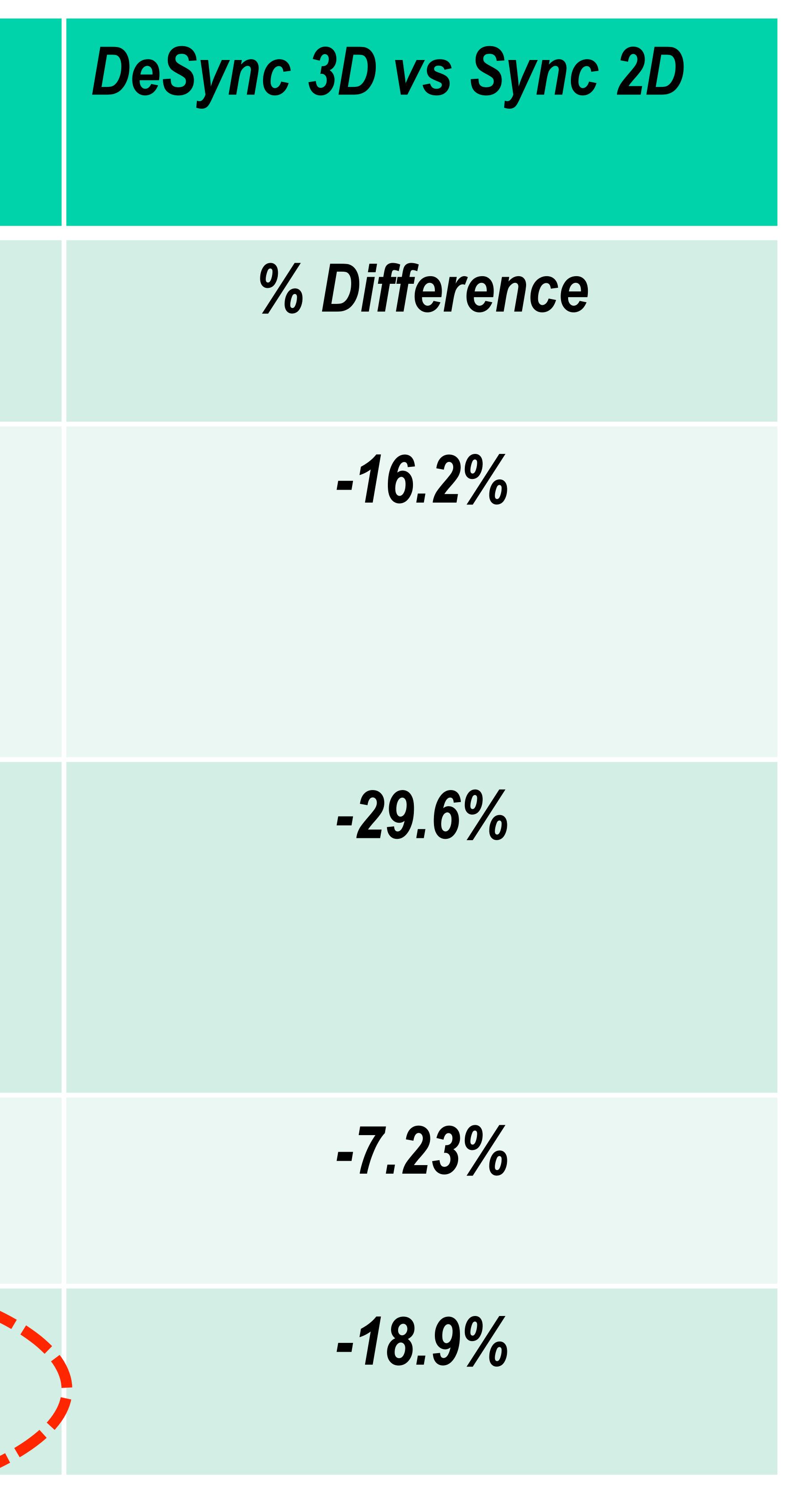
Smaller cells used in de-sync designs due to absence of global nets.
3D uses fewer gates over all.

POWER COMPONENT	Synchronous		<b>De-synchrone</b>	DUS
(All units in W)	<b>2D</b>	<b>3D</b>	<b>2D</b>	<b>3D</b>
Switching power	0.1171	0.0824 (-29.6%)	0.1361	0.0981 (-27.9%)
Cell power	0.0529	0.0423 (-20.0%)	0.0513	0.0372 (-27.4%)
Leakage power	0.0221	0.0198 (-10.4%)	0.0225	0.0205 (-8.88%)
Total Power	0.1921	0.1444 (-24.8%)	0.2098	0.1557 (-25.7%)

• 3D sync has better power numbers.



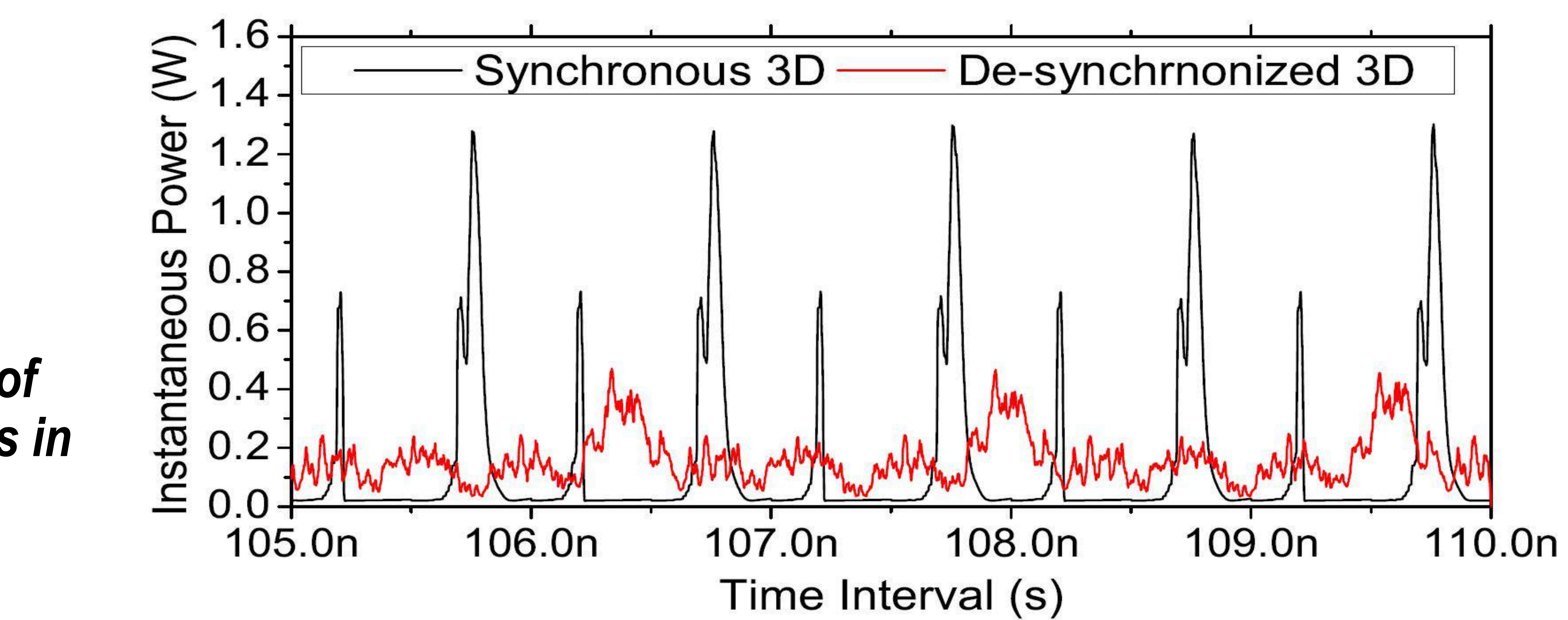




## Design Metrics: Instantaneous Power

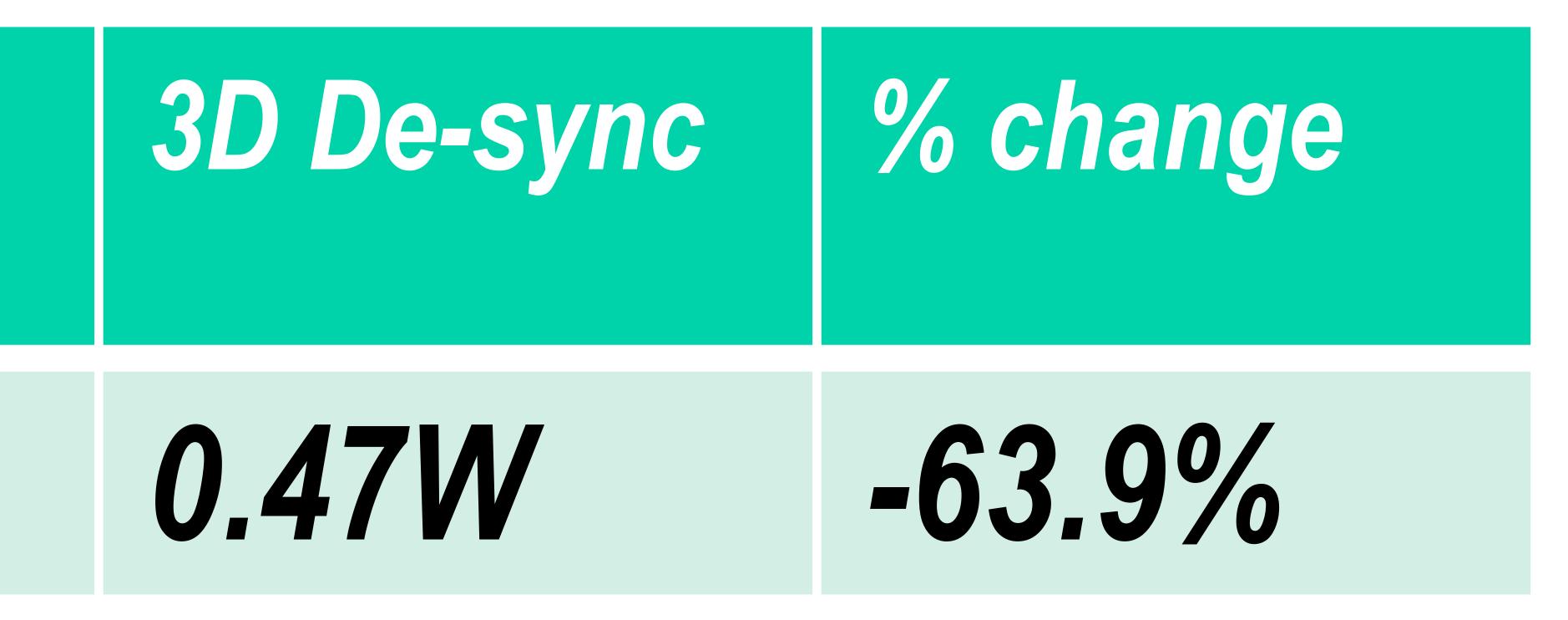
### Much lower EMI and noise.

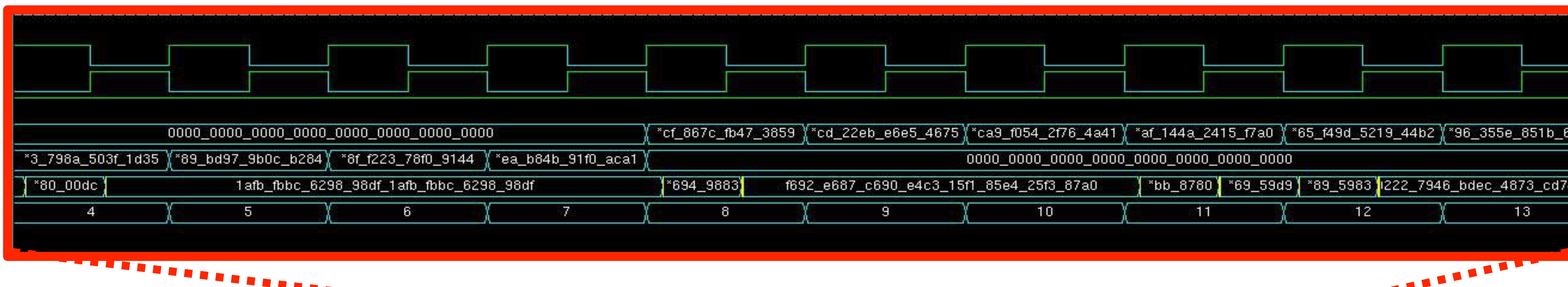
- Do not decrease the signal to noise ratio of adjacent analog parts in a Soc (Less EM emission/pollution)
- Wide operating voltages & can accept poor supply voltage quality
- Resistance to hardware attacks like DPA.



2D Sync	2D De-sync	% change	3D Sync
1.39W	0.602W	-56.6%	1.302W

### PEAK POWER CONSUMPTION IN WATTS

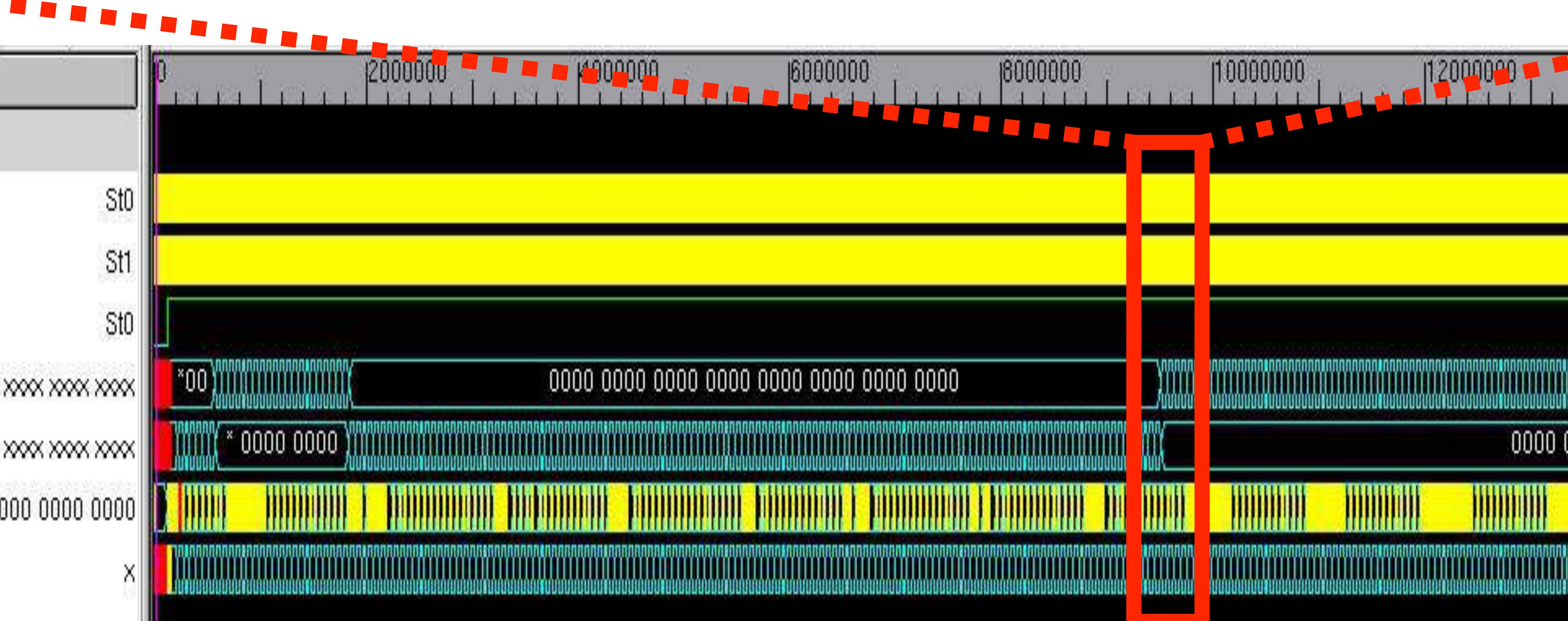


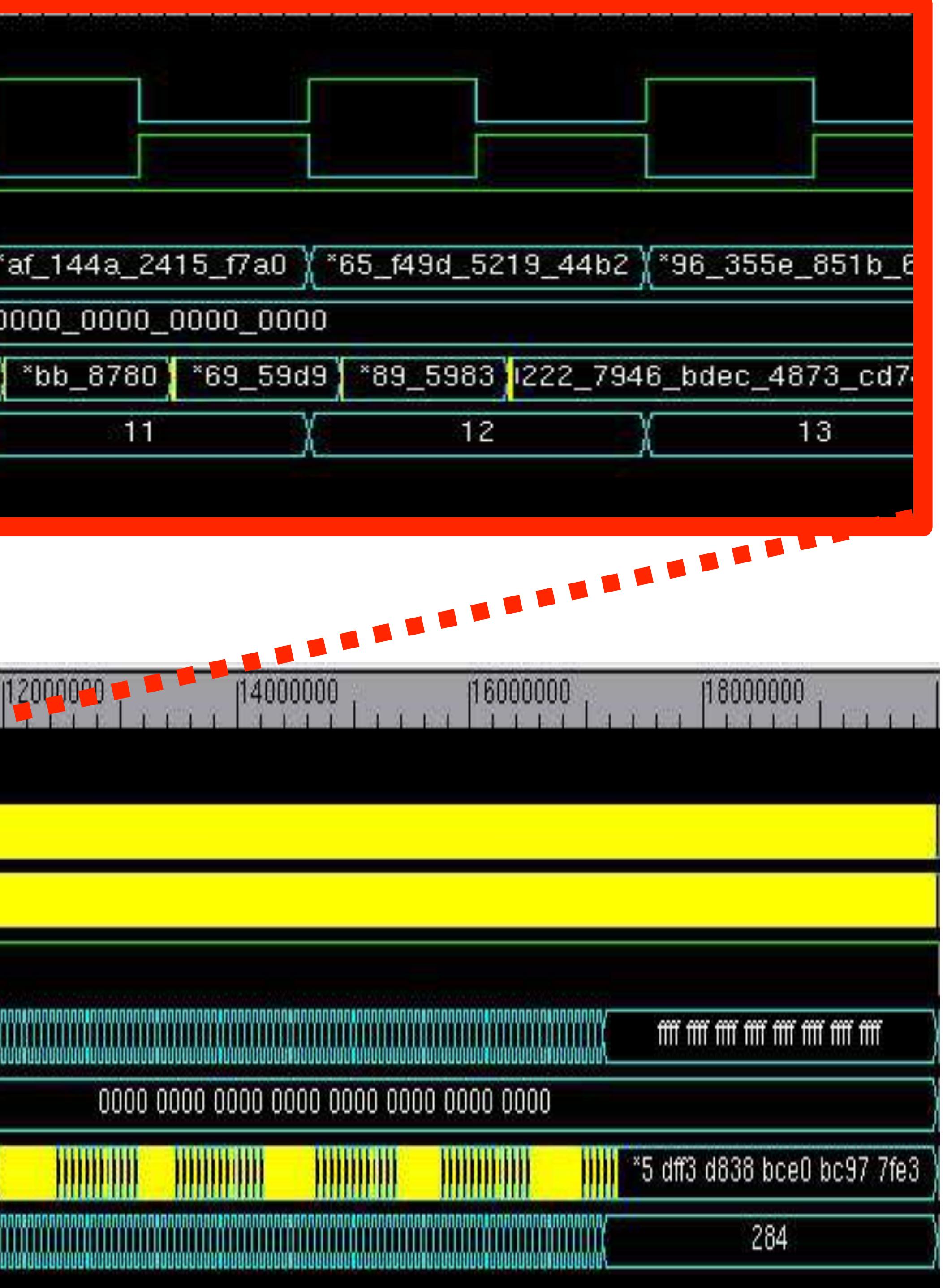


Name	Value	
🛛 req		
ack 🛛		
🛛 reset		
🕂 🛿 cipher_key[127:0]	128'hxxxx xxxx xxxx xxxx xxxx xxxx xxxx x	
	128'hxxxx xxxx xxxx xxxx xxxx xxxx x	
	128'h0000 0000 0000 0000 0000 0000 00	

### 3D Desynchronized AES system has correct functionality even with up-to 15% performance degradation. • 3D Synchronous fails at 2%

## Variation Aware Functional Analysis





## Summary

• We study the synergistic benefits of monolithic 3D and asynchronous circuits for the first time. • We propose a design methodology for 3D integration of desynchronized circuits. • We demonstrate PPA overhead in de-synchronized AES is significantly reduced through monolithic 3D integration. • We demonstrate significant power reduction in 3D circuits at ISOperformance comparison. • We observed that de-synchronized 3D design is more variation tolerant than 3D synchronous version.



## Questions ??



