Monolithic 3D Integration

using Standard Fab & Standard Transistors

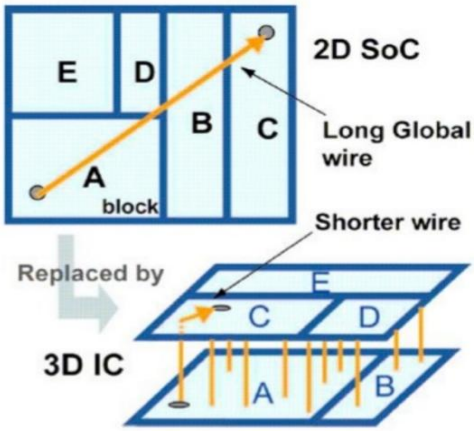
Zvi Or-Bach
CEO
MonolithIC 3D Inc.

MonolithIC3D™

3D Integration

Through Silicon Via ("TSV"), Monolithic

- Increase integration
  - Reduce interconnect total length
  - Not subject to improved litho. and process
- Heterogeneous integration
  - Silicon, Optoelectronic, Sensors, ..
  - I/O, Logic, Memory, …
- New functionality utilizing the third dimension
  - Redundancy (WSI)
  - Shared Litho. And processing (BiCS)



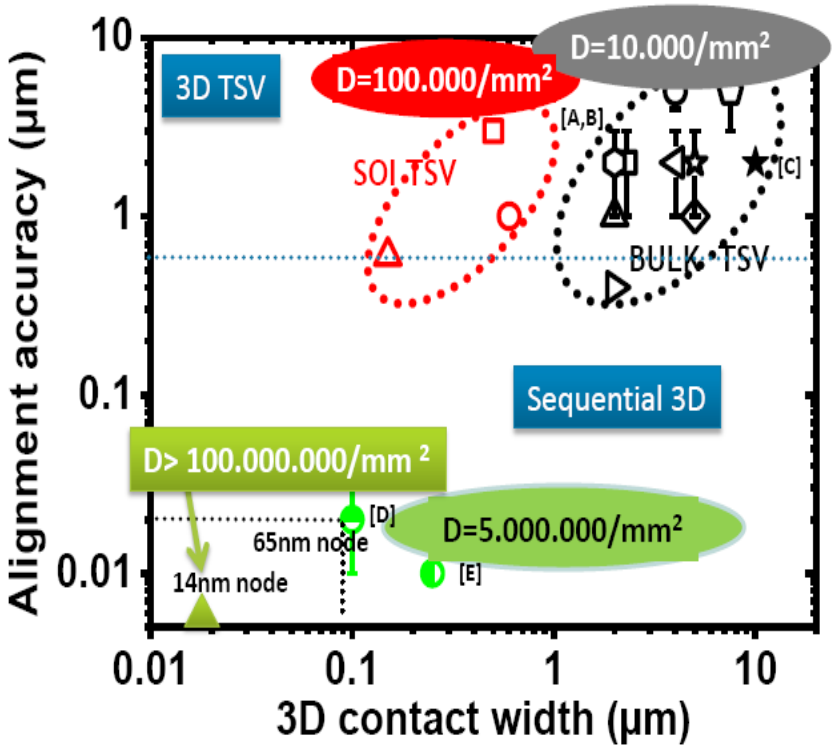
Monolithic 3D

10,000x the Vertical Connectivity of TSV

	TSV (Through Silicon Via)	Monolithic
Layer Thickness	~50µ	~50nm
Via Diameter	~5µ	~50nm
Via Pitch	~10µ	~100nm

CoolCube unique features

High density of vias



The Monolithic 3D Challenge

Why is it not already in wide use?

- ***Thermal Budget of Upper Layer Processing***
  - ***< 550°C over advanced node transistors***
  - ***< 400°C over Copper or Aluminum interconnect***
    - **How to bring mono-crystallized silicon on**
    - **How to fabricate state-of-the-art transistors on top**
- ***Misalignment of pre-processed wafer to wafer bonding step used to be ~1µm***
    - ***How to achieve 100nm or better connection pitch***
    - ***How to fabricate inter-layer vias of ~50nm***

MonolithIC3D™

Outline: Layer Formation

- **Deposit**
  - **Poly silicon + crystallize by laser (Taiwan NAR Lab)**
  - **Non silicon**
- **Layer Transfer**
  - **Bond SOI wafer and etch/grind back**
    - **IBM, MIT Lincoln Lab**
  - **Ion-Cut (Smart Cut®)**
  - **'Smart' Substrate**
    - **ELTRAN (Epitaxial Layer Transfer)**
    - ***New Breakthrough***

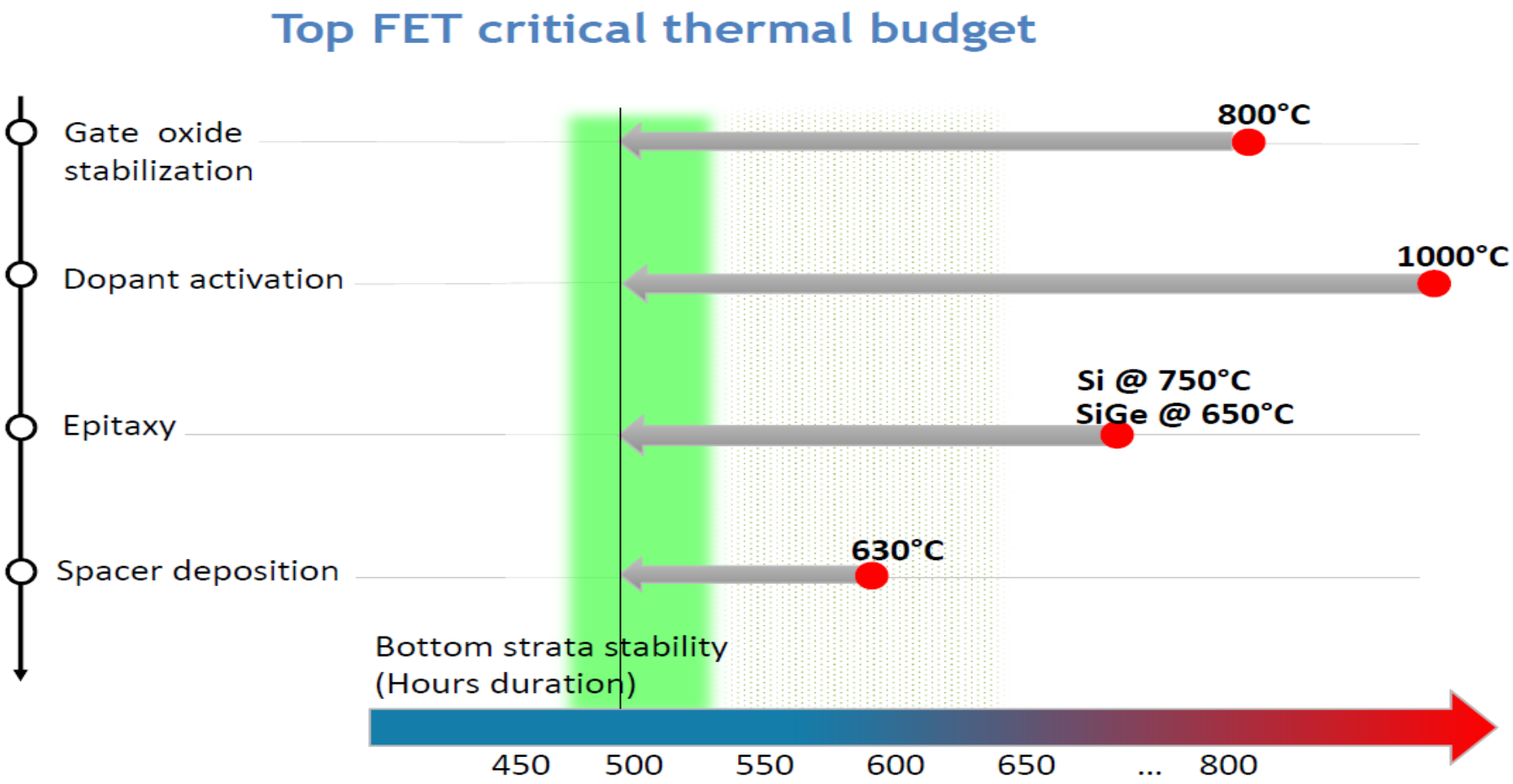
MonolithIC3D™

Outline: Overcoming the Process Heat

- Non silicon transistor
- Modified process to less than 550^{o}***C (CoolCube)***
- Break transistor formation into two phases,
  - Hot-Cold
- Isolate Top layer from base– “Crème Brulee”
- ‘nano-TSV’ – Precise ultra thin stratum transfer

MonolithIC3D™

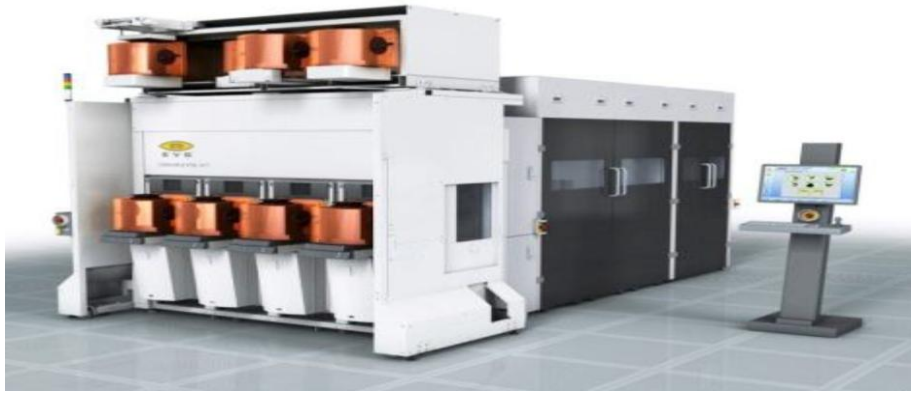
Modified the process for low temp



Rmk: Bottom strata stability window is wider for shorter anneal durations (up to 800°C, ms)

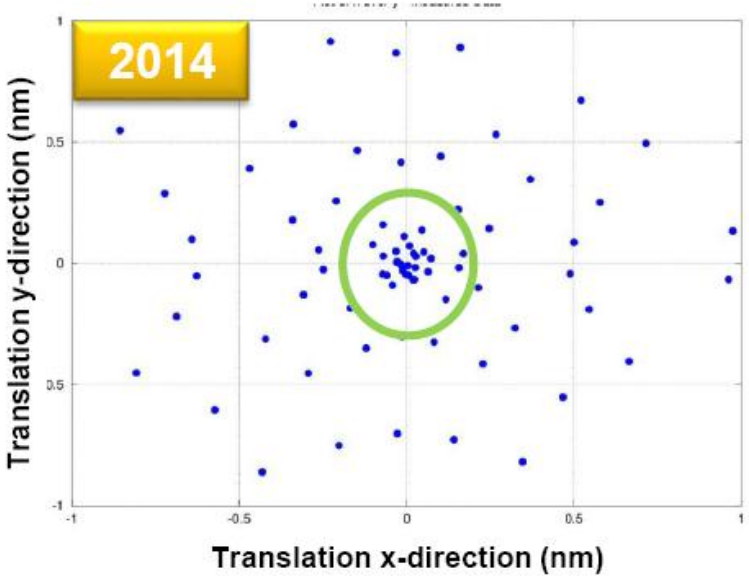
Precision Bonder – Multi-Stratum M3D

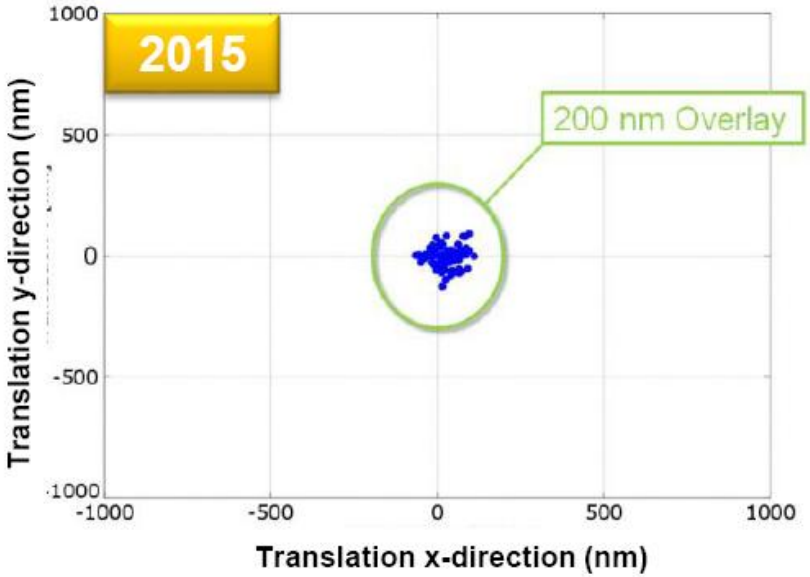
- **Utilizing the existing front-end process !!!**
  - **<200 nm (3σ)**
- Achieving 10,000x vertical connectivity as the transferred strata will be thinner than 100 nm
  - Mix – Sequential/Parallel M3D
  - Low manufacturing costs

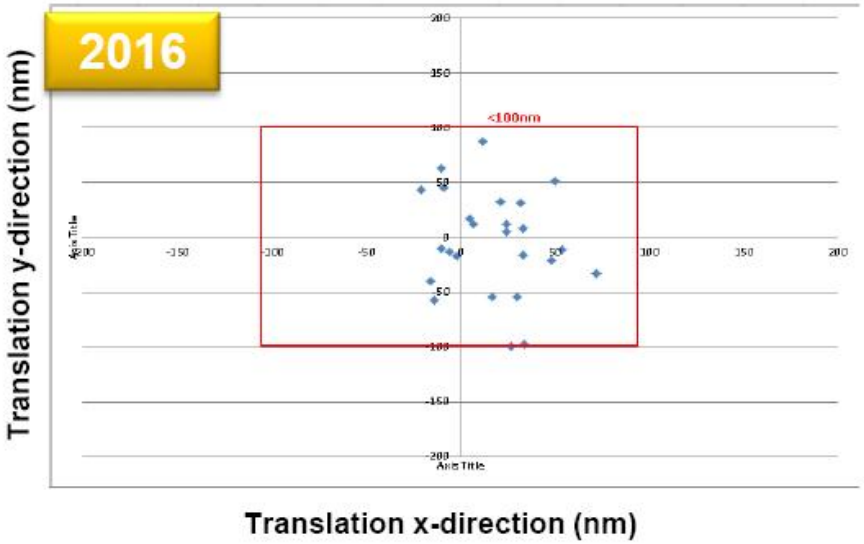


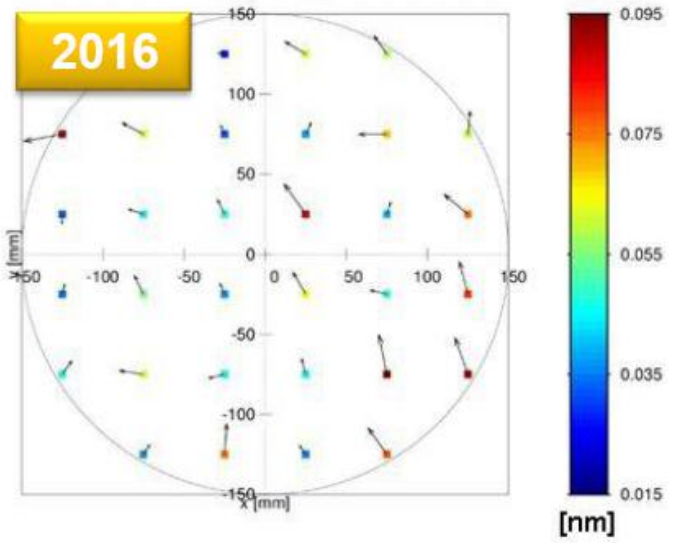
MonolithIC3D™

Fusion Bonding – Overlay Evolution

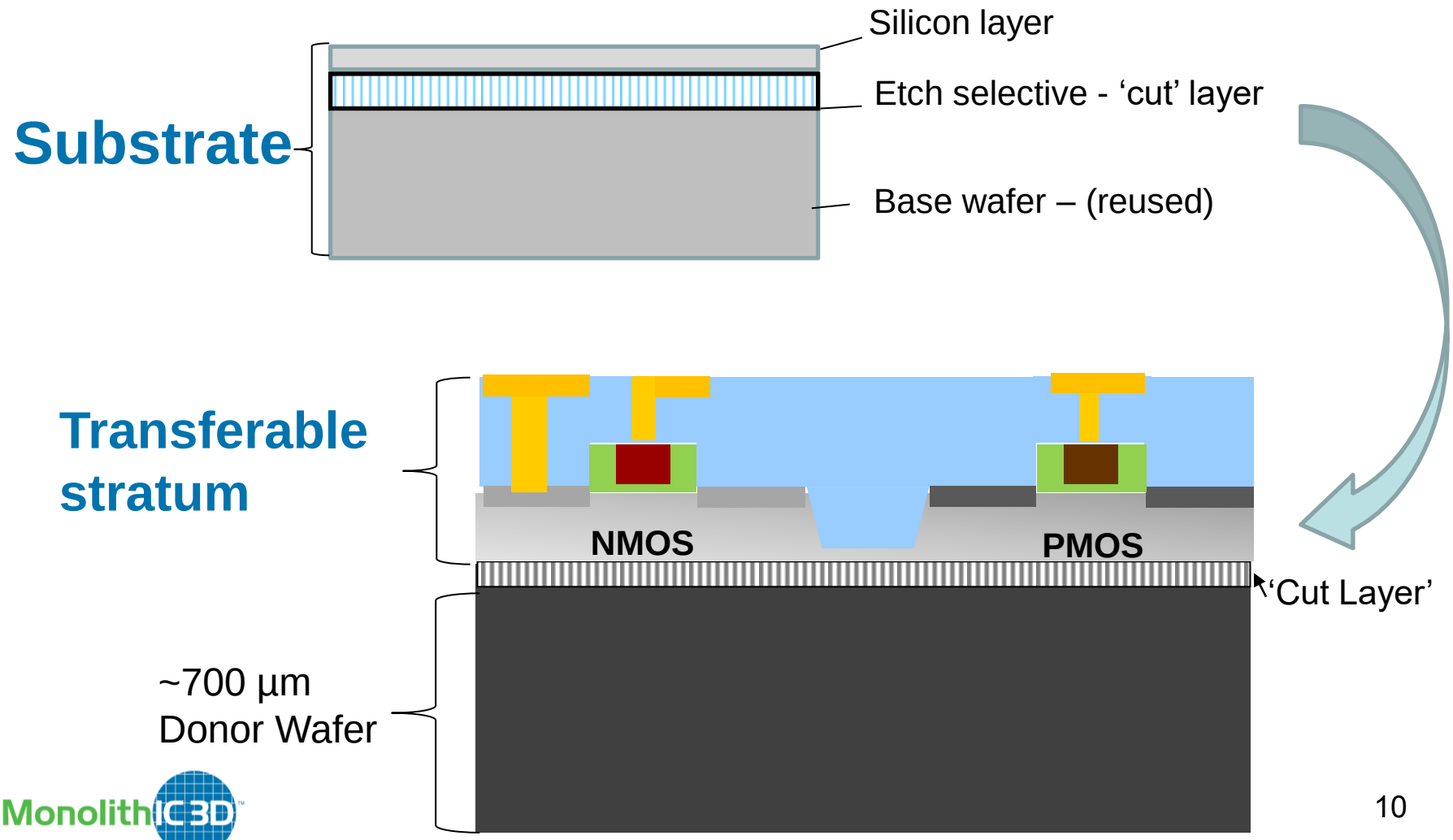




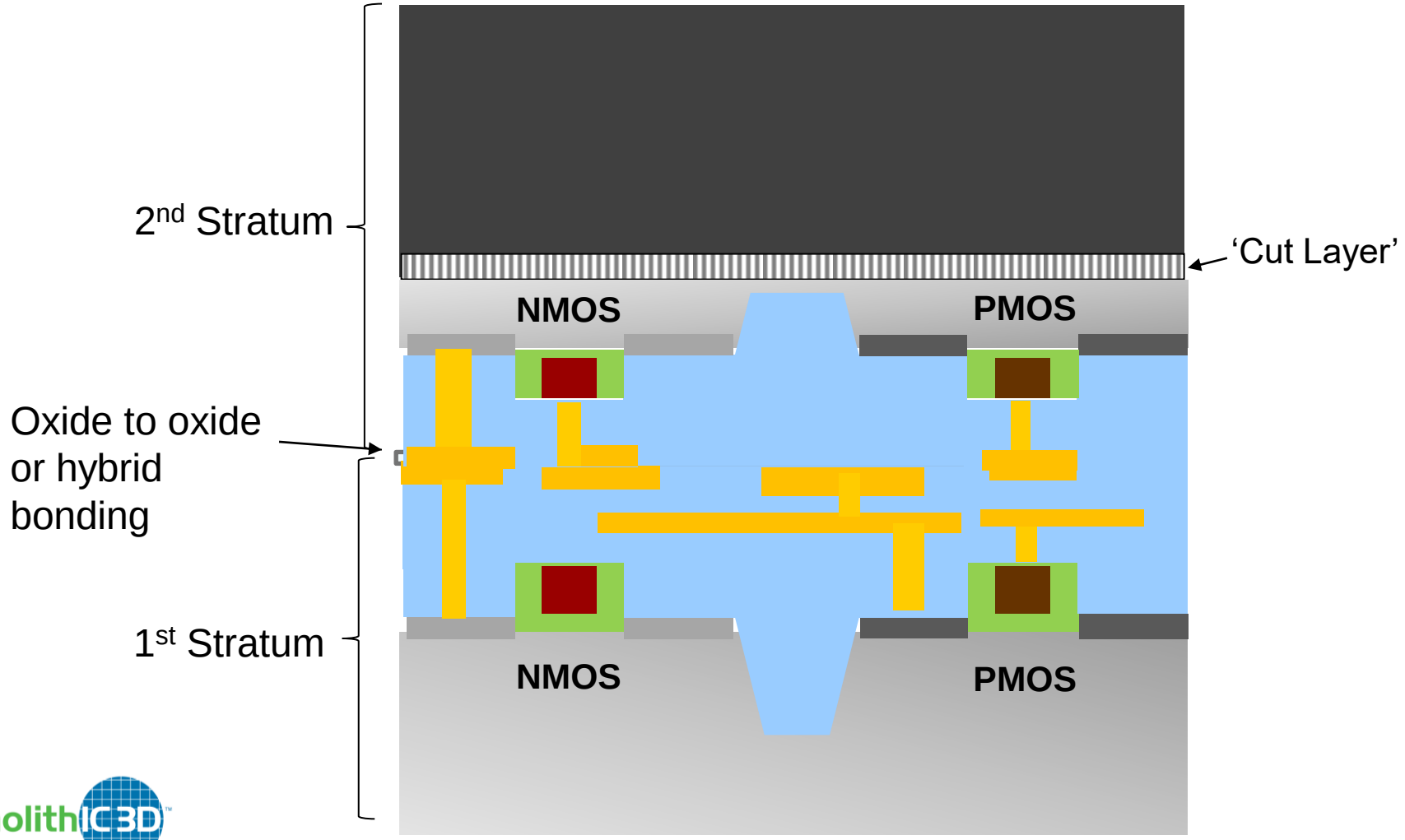




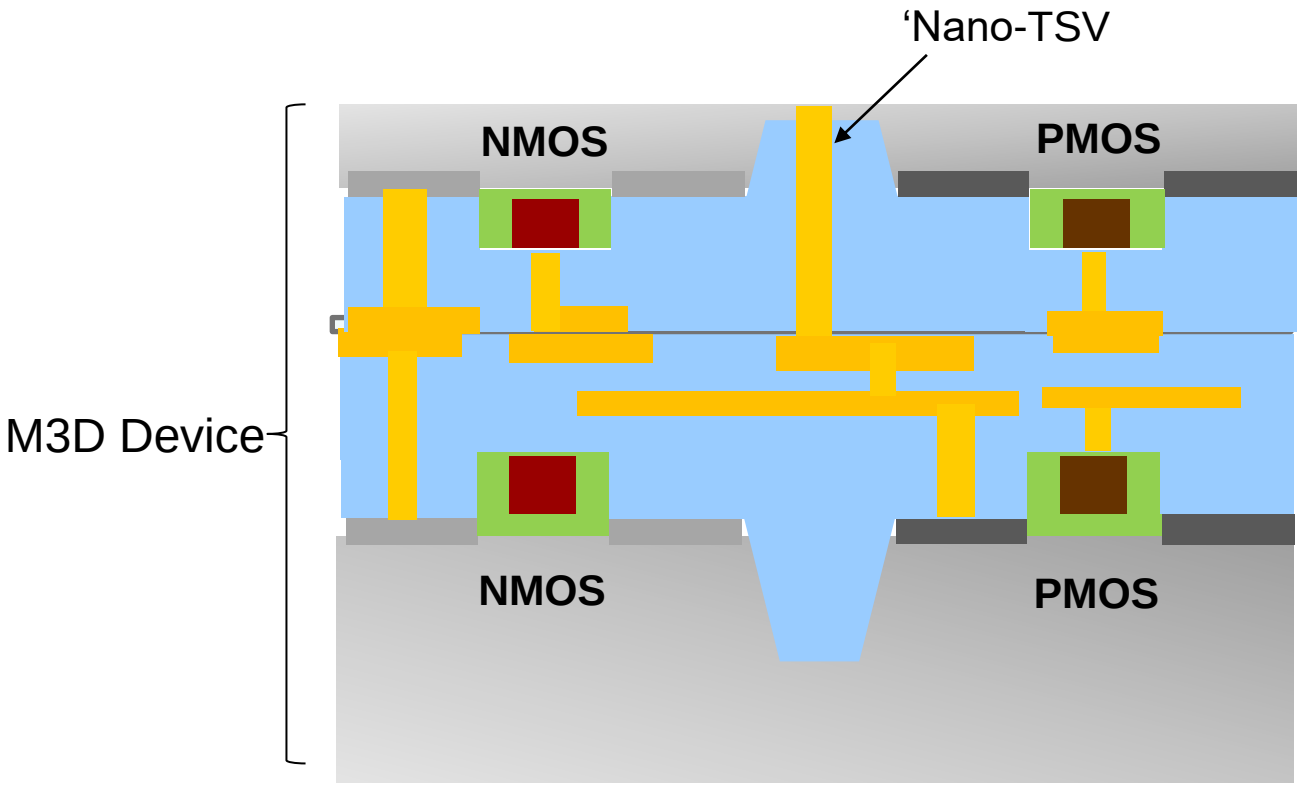
2nd Stratum Processed on 'Cuttable' wafer



Flip and Bond 2nd Stratum on Top of 1st Stratum



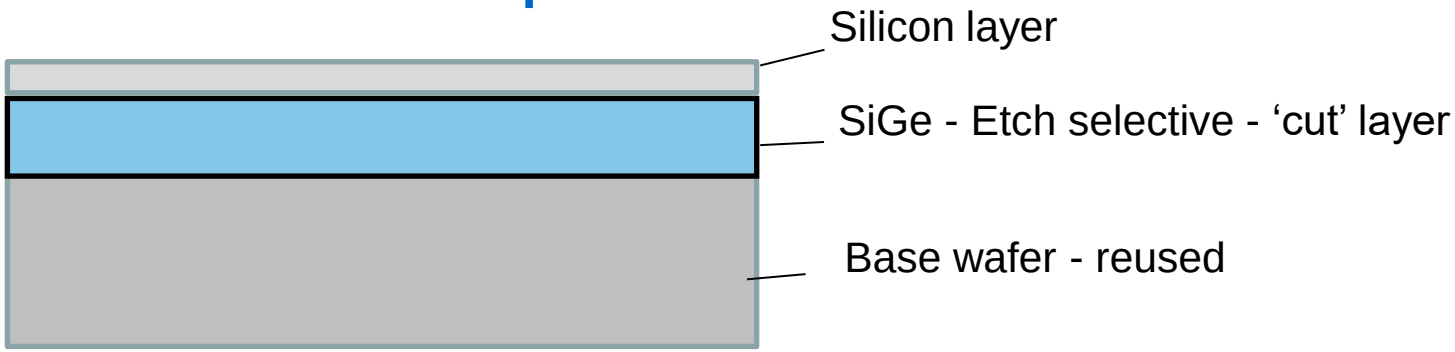
‘Cut’ Carrier Wafer and Connect with Nano-TSV



MonolithIC3D

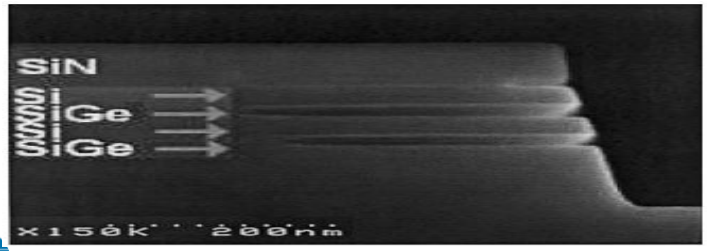
The Breakthrough –
SiGe – Fab Compatible 'Cut Layer'

- Use SiGe as etch stop



- Use selective etch of SiGe (Dry, ~1:100) for reuse

Samsung's



Applied Material's Selectra

MonolithIC3D™

Monolithic 3D using 'Cuttable' Substrate

- Utilizes existing transistor process (first ever)
- Leverages proven silicon processes
- Very competitive cost structure

MonolithIC3D™

The Monolithic 3D **Advantages**

– *The 1,000x*

- Enables increased integration without size reduction
- Reduce average wire length by ~0.7x
- **Heterogeneous Integration**
  - Type of wafer, Si, GaAs, GaN,..
  - Mix fabs and nodes, 7nm, 28nm, 130nm,…
  - **Mix process, logic, memory**, analog, RF,..
- Multiple layers Processed Simultaneously - ‘BiCS’
- **New functionality leveraging rich 3D connectivity**
  - Redundancy
- Allow connectivity from both sides

MonolithIC3D™

EDA for Monolithic 3D

- Commercial EDA is not available – yet.
- Preliminary tools and flow has been developed
  - Georgia Tec
  - Qualcomm
  - CEA Leti…
- **2D are sufficient for the 1,000x applications !!!**

Summary

- We have reached an inflection point
- Monolithic 3D would provide unparalleled advantage for most market segment
- Multiple practical paths to monolithic 3D exist
  - **Monolithic 3D could be now implemented by any fab without changes to front-line**
  - **In 2017 30% of all transistors made will be M3D** (3D NAND)

=>Use correctly MonolithIC 3D would provide 1,000x better devices

MonolithIC3D

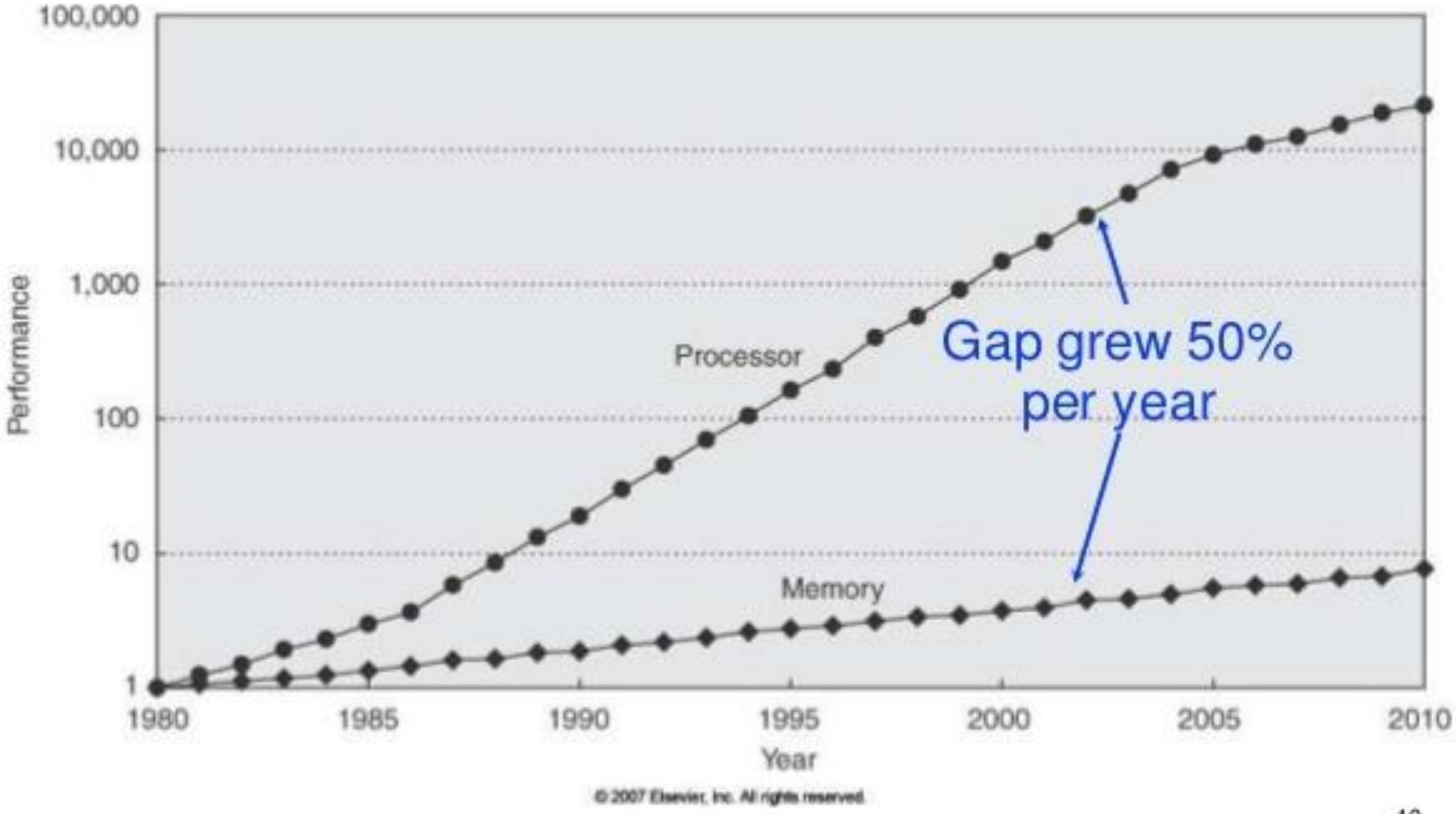
Monolithic 3D Integration

for

1,000x Better Computer

MonolithIC3D™

Processor Memory Gap



Source: Computer Architecture, A Quantitative Approach by John L. Hennessy and David A. Patterson

Why we need Exascale and why we won't get there by 2020

Horst Simon
Lawrence Berkeley National Laboratory

The Problem with Wires:

Energy to move data proportional to distance

- **Cost to move a bit on copper wire:**
  - power= bitrate * Length / cross-section area

- **Wire data capacity constant as feature size shrinks**
- ***Cost to move bit proportional to distance***

System Level Interconnect Gap Grow

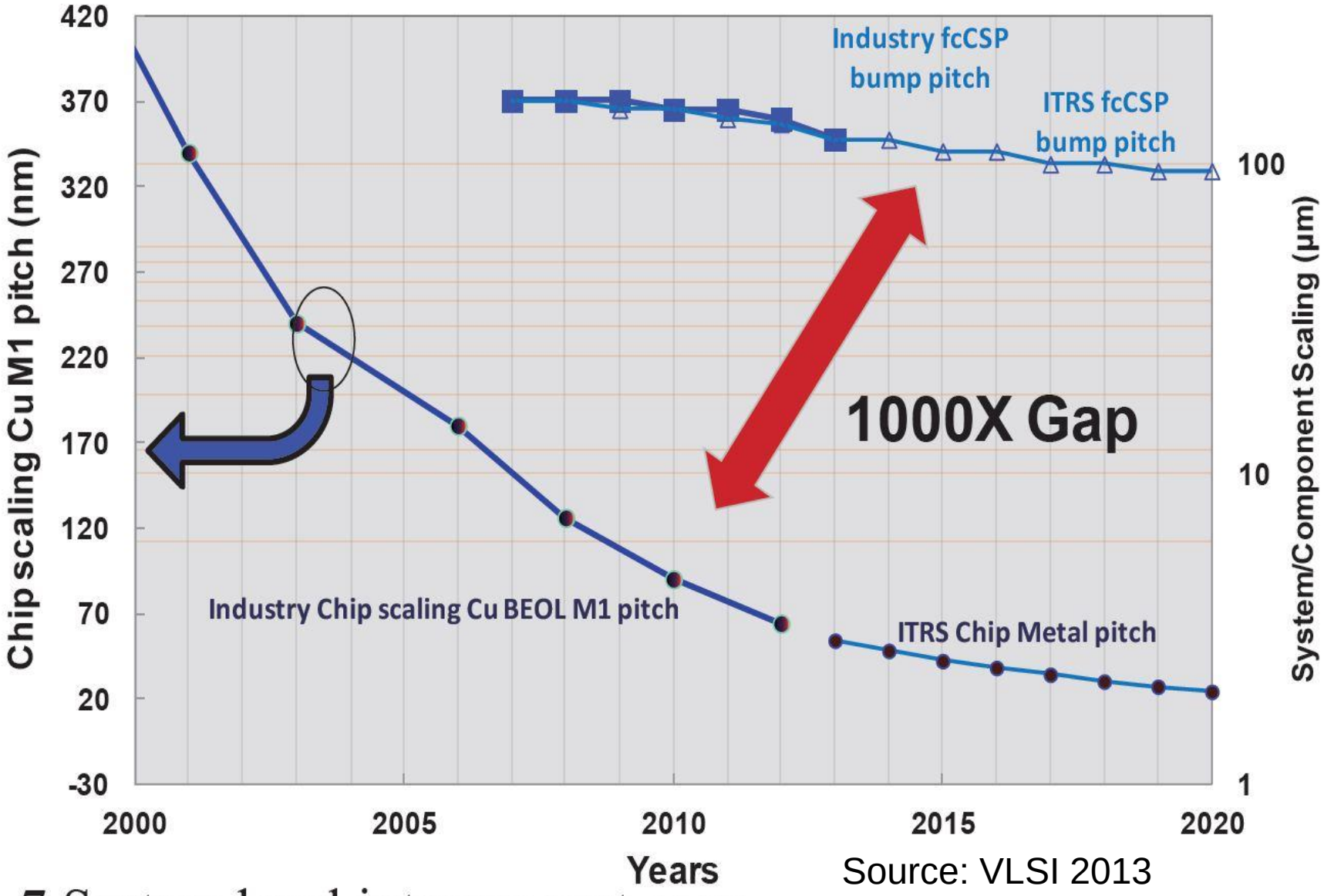
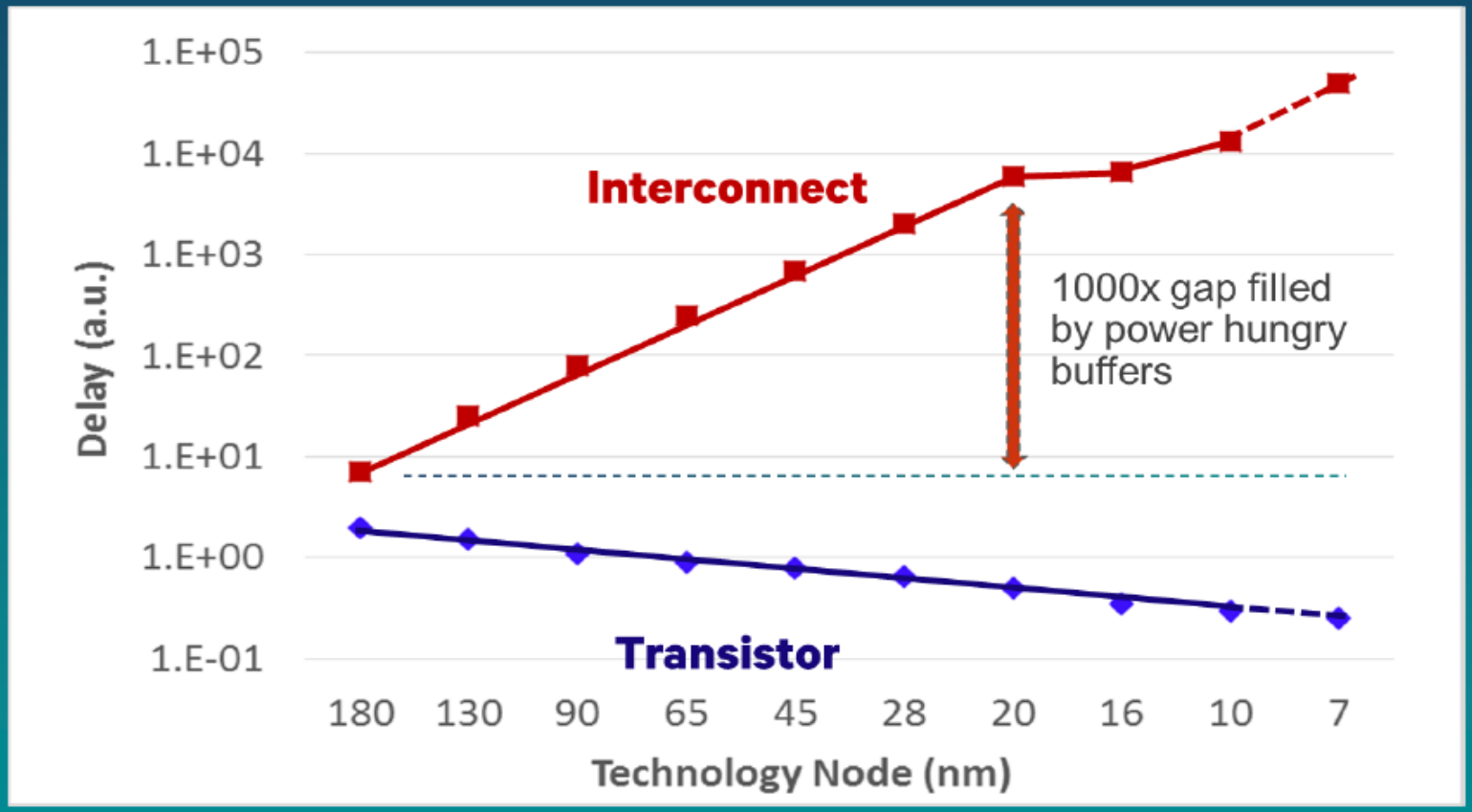


Fig. 7 System level interconnect gaps

Source: VLSI 2013
Dr. Jack Sun, CTO of TSMC

Growing interconnect challenge



Source: ITRS

SEMICONDUCTOR RESEARCH OPPORTUNITIES

An Industry Vision and Guide

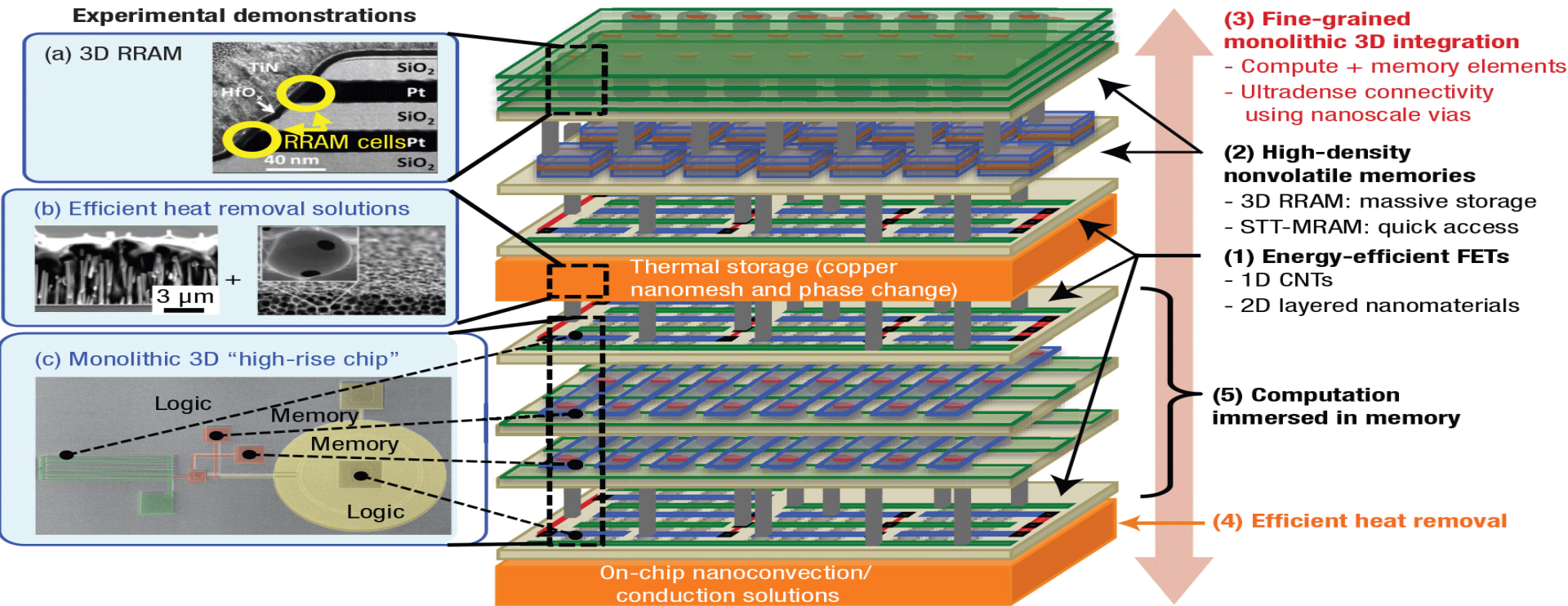
March 2017

SIA SEMICONDUCTOR INDUSTRY ASSOCIATION

SRC Semiconductor Research Corporation

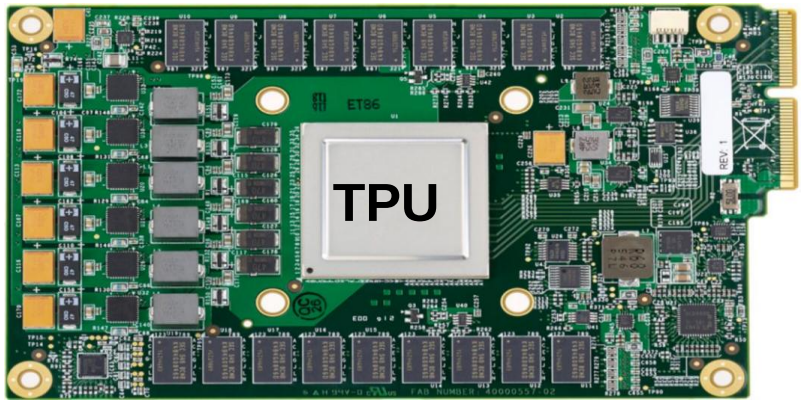
Devices enabling high-density, fine-grained, monolithic 3D systems for reducing data movement and communication cost:

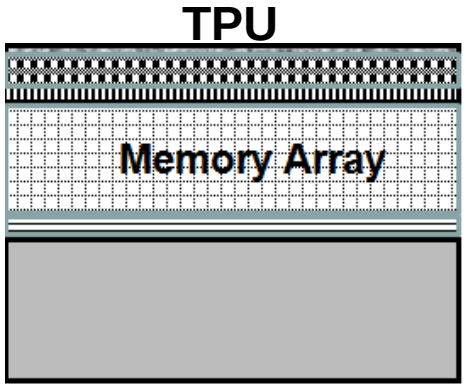
- High-performance devices for logic and memory suitable for low-temperature processing used for stacked layer integration; the resulting energy-delay product should be 1000X better than state-of-the-art conventional technology. Candidate materials for next-generation logic



Monolithic 3D for 1,000x better compute

- Processor to memory connectivity breakthrough:

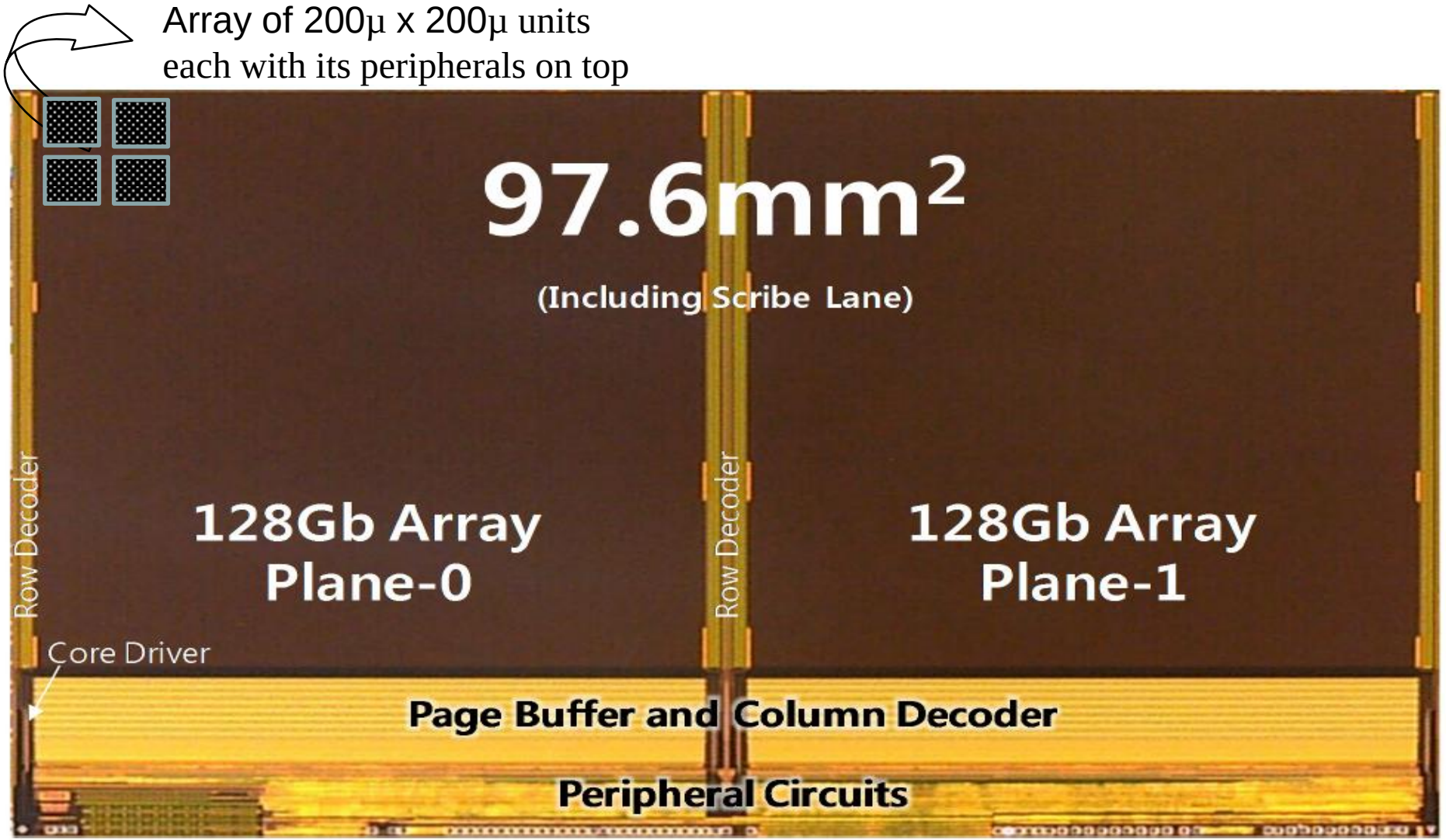




  - Number of wires: ~100 ⇨ ~100,000
  - Length of wires: ~20mm ⇨ ~20μm
- Applications:
  - Smart-Phones
  - IoT
  - Computers

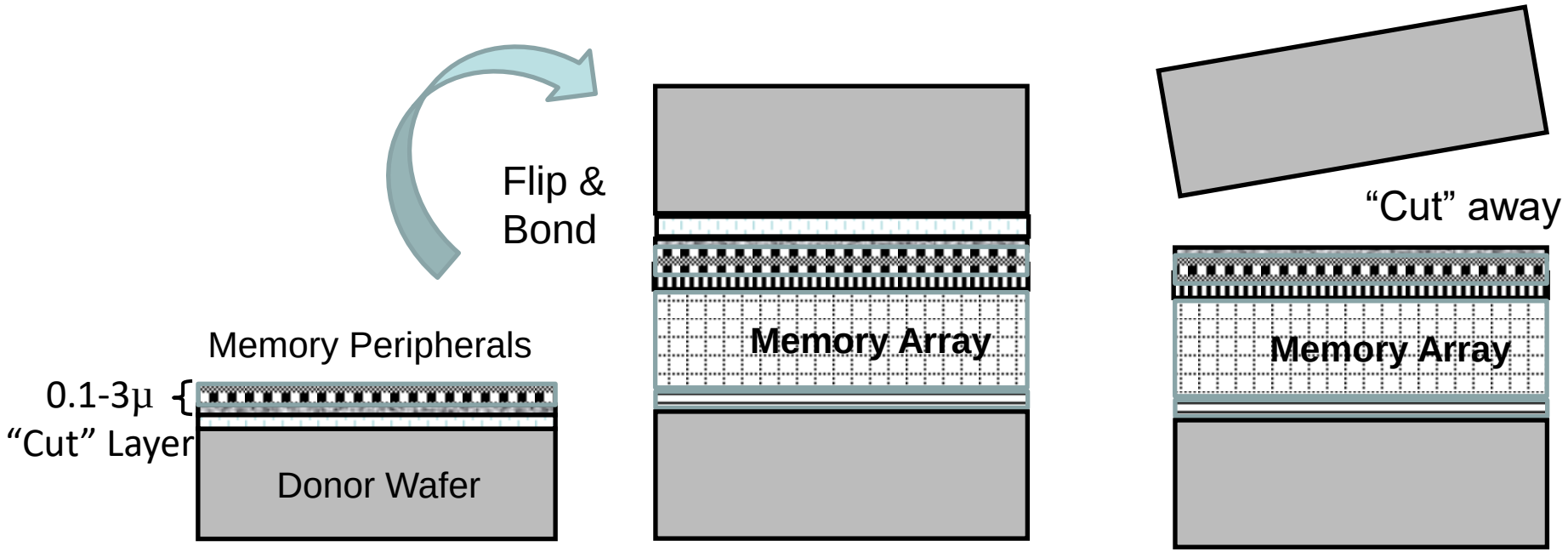
MonolithIC3D™

Replacing few mm long bus of <100 wires with few microns long bus of 2,500 x 100 wires



1,000x Better Compute System Peripherals over Memory Array

- Solving the ‘Memory Wall’
  - Close proximity of memory to the logic control
  - Allows a different process for the memory matrix than the peripherals



MonolithIC3D

Summary

1,000x better compute Systems using M3D

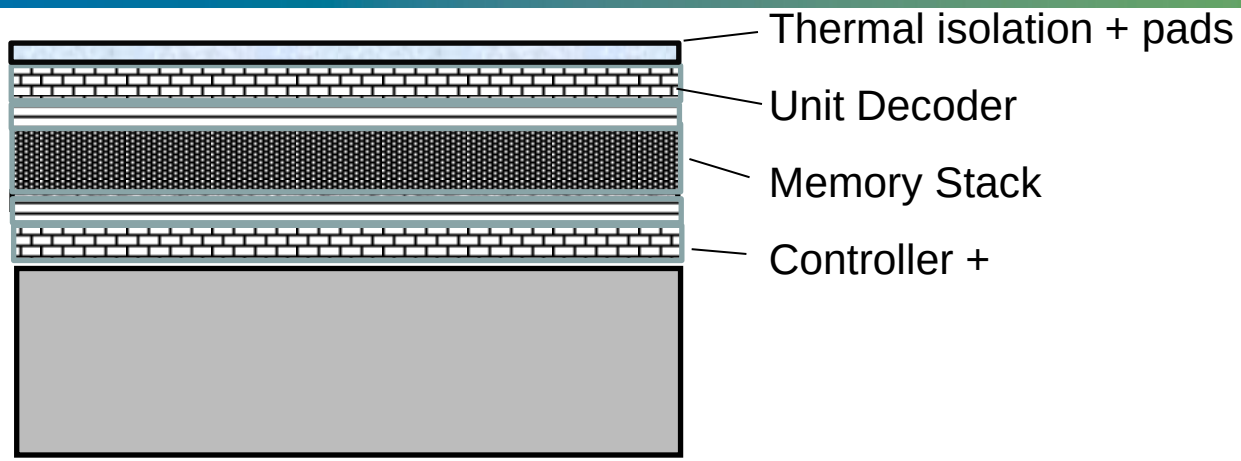
- Multi thin layers each with thousands memory units

  =>~100,000 wires @ ~20μ length

  => **1,000x better computer system**
- Utilizes existing memory (DRAM) and logic process
- Far cheaper than any other form of DRAM module
- Yield repaired in fabrication and in operation (field)
- Memory stack could serve as an generic end product
- Low-power (1/10), low size, low weight

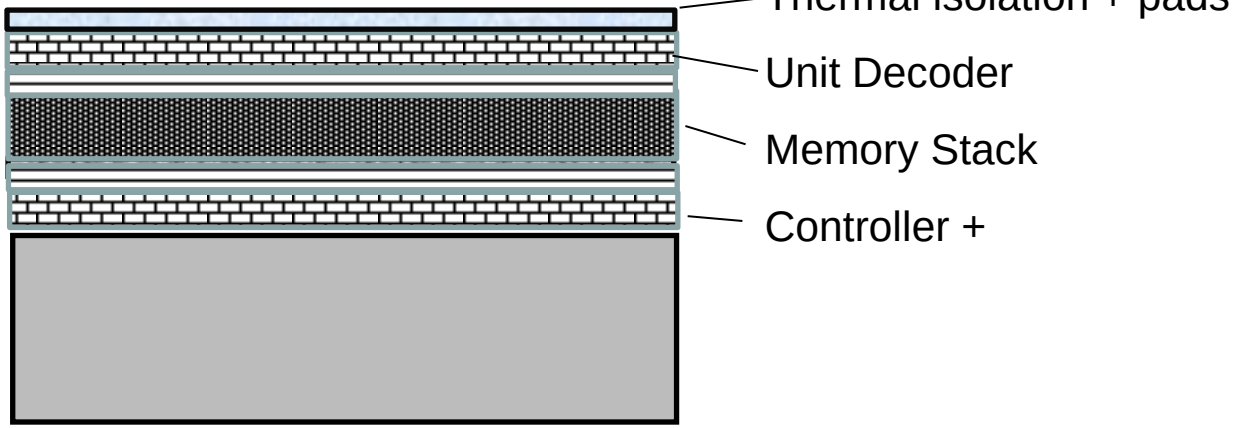
Monolith IC3D™

New Memory Product – Generic 3D Wafer

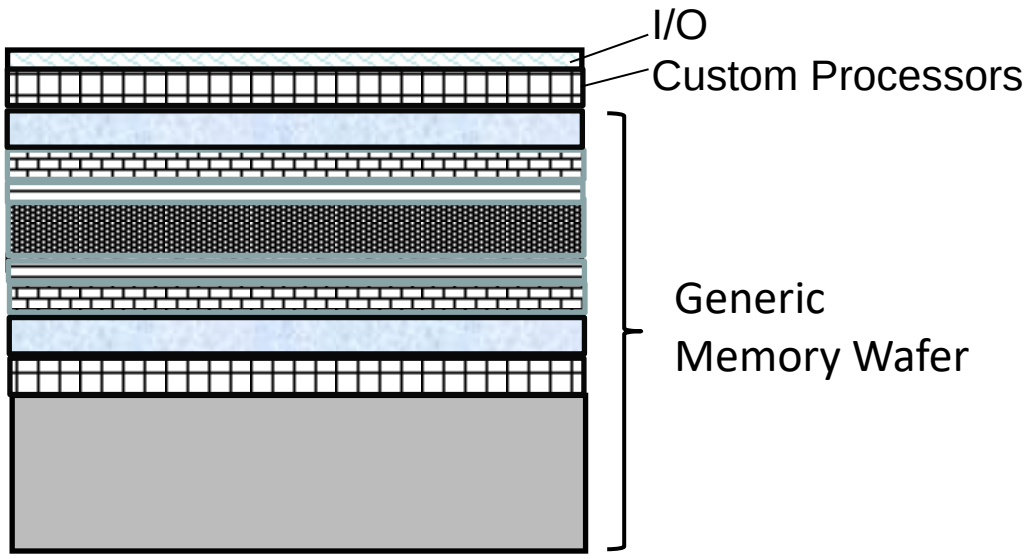


Generic Memory Wafer

New Memory Product – Generic 3D Wafer



Generic Memory Wafer



3D Integrated Computer

MonolithIC3D

Other Approaches

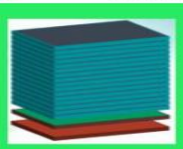
- HBM, HMC, Tezzaron – DiRAM4
  - 2-8x more expensive
  - Only solve the memory bank integration
    - Only for DRAM

leti ceatech

3D DRAM : COMPARISON

MULTICORE/MANY-CORE SYSTEMS-ON-CHIP

MCSoC'16 Conference, Lyon, France

	LPDDR4	WideIO/2	HBM	HMC	DiRAM4
		JEDEC  SK hynix	JEDEC  SK hynix SAMSUNG	Hybrid Memory Cube  Micron	Tezzaron  Tezzaron
Interface type	parallel	wide data	wide data	serial	wide data or serial
Data bus	16b DDR	64b DDR	128b DDR	16 lanes	64b
Channel	2	4-8	8	4-8	
I/O bandwidth	3.2Gbps @1600MHz	0.8Gbps @400MHz	1-2Gbps @500-1000MHz	10-15Gbps	
Total bandwidth	12.8GBps	25.6-51.2GBps	128-256GBps	160-320GBps	2TBps
Capacity	16GB	16GB	32GB Currently 1GB (Gen1) Next 4-8GB (Gen2)	32GB Currently 2-4GB	8GB
Total I/O	66	776	1616	256-512	
Integration / Packaging	POP, MCP	3D	2.5D	MCP	
Computing-In-Memory		NO	NO	**YES**	NO

Alternatives for Processor Memory Connectivity

	Processor	PCB	Module	DRAM
2D	5mm	10-30mm	20mm	5mm
2.5D-TSV	5mm	5-10mm	0	1mm
M3D	0.5mm	0	0	0.2mm

	#of Wires	M Cost
2D	100	1
2.5D	1,000	2-8
M3D	100,000	0.5