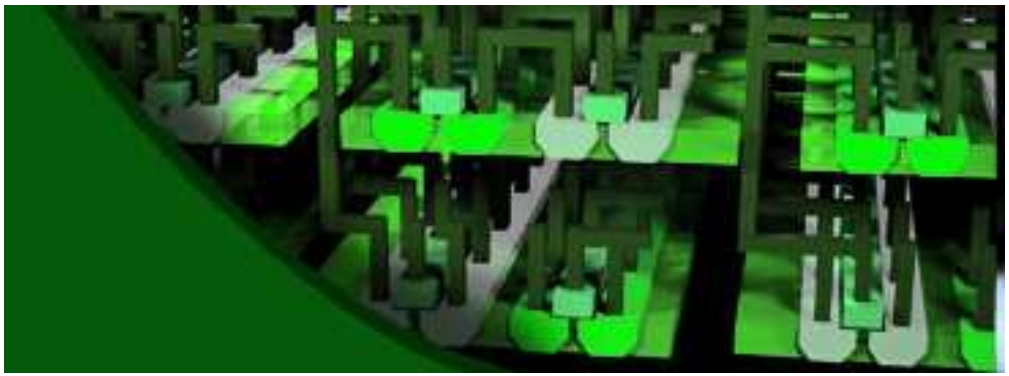
A DISRUPTOR TO THE SEMICONDUCTOR INDUSTRY

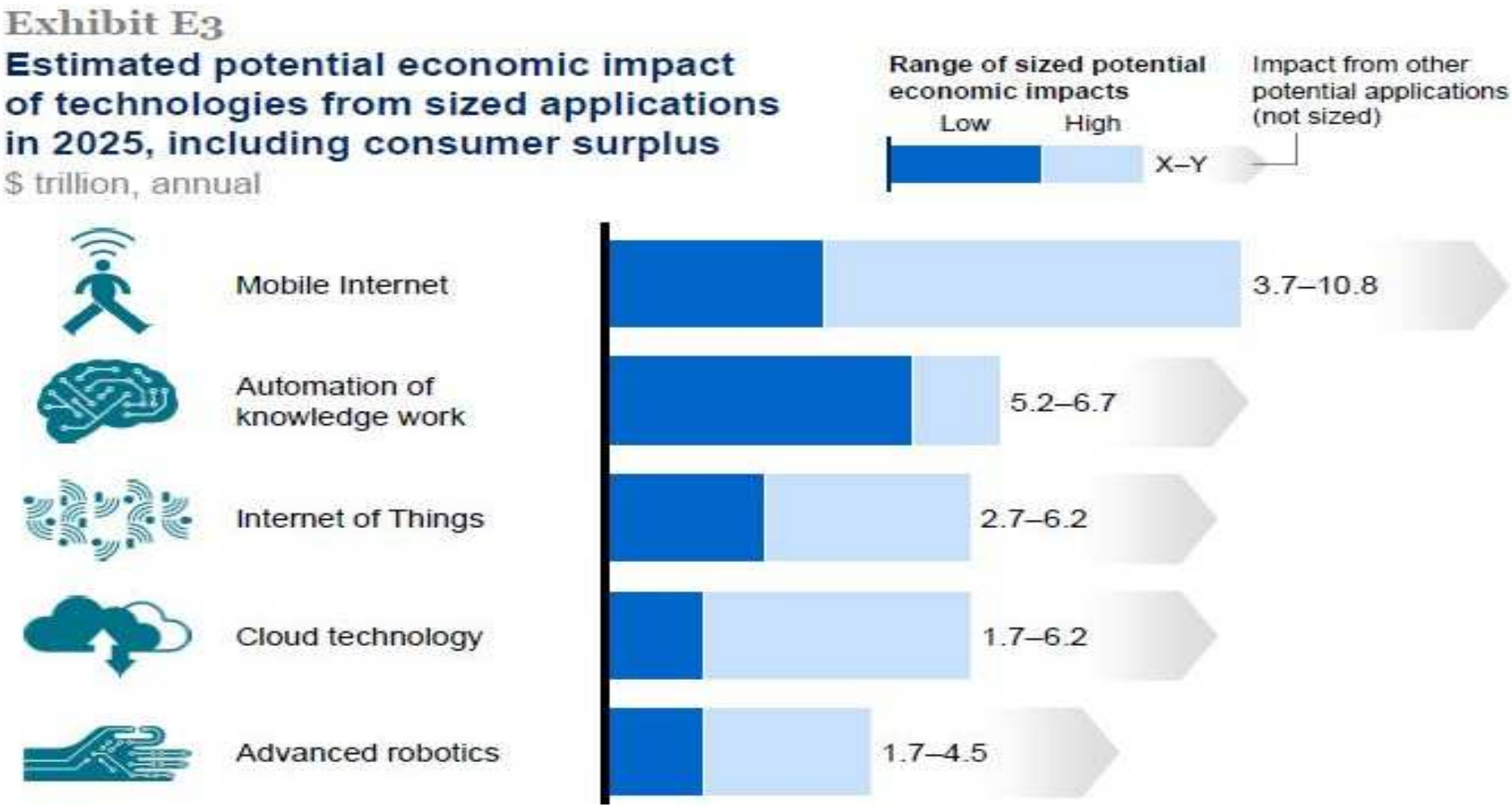
THE MONOLITHIC 3D-IC



Agenda:

- The Semiconductor future is exciting
- But we are reaching an inflection point
- Monolithic 3D IC – the emerging path for the next generation technology driver
- The challenge and solution for the fabrication of monolithic 3D IC

$15-34 trillion, annual =>~$5T Semi /year



Source: McKinsey Global Institute Analysis 2013

Cisco sees $19 Trillion opportunity in IoT

"CES LIVE: Cisco's Chambers Says Internet of Everything, $19 Trillion Opportunity, Is Next Big Thing" 1/7/14 <ttp://www.forbes.com/sites/connieguglielmo/2014/01/07/ces-live-cisco-ceo-chambers-to-deliver-keynote/>

$19 trillion: that's the opportunity he says for the Internet of Everything in the private and public sector combined. Breakout is $14.4 trillion in private sector and $4.6 trillion in public sector of new revenue generation or new savings. That's a conservative number he says for public sector.

"This will be bigger than anything done in high tech in a decade."

"As many as 50 billion devices will be connected to the Internet by 2020, creating a $14.4 trillion business opportunity" said Rob Lloyd, president of sales and development at Cisco, <http://www.eetimes.com/electronics-news/4409928/Cisco-sees--14-trillion-opportunity-in-Internet-of-Things>

Monolithic3D™

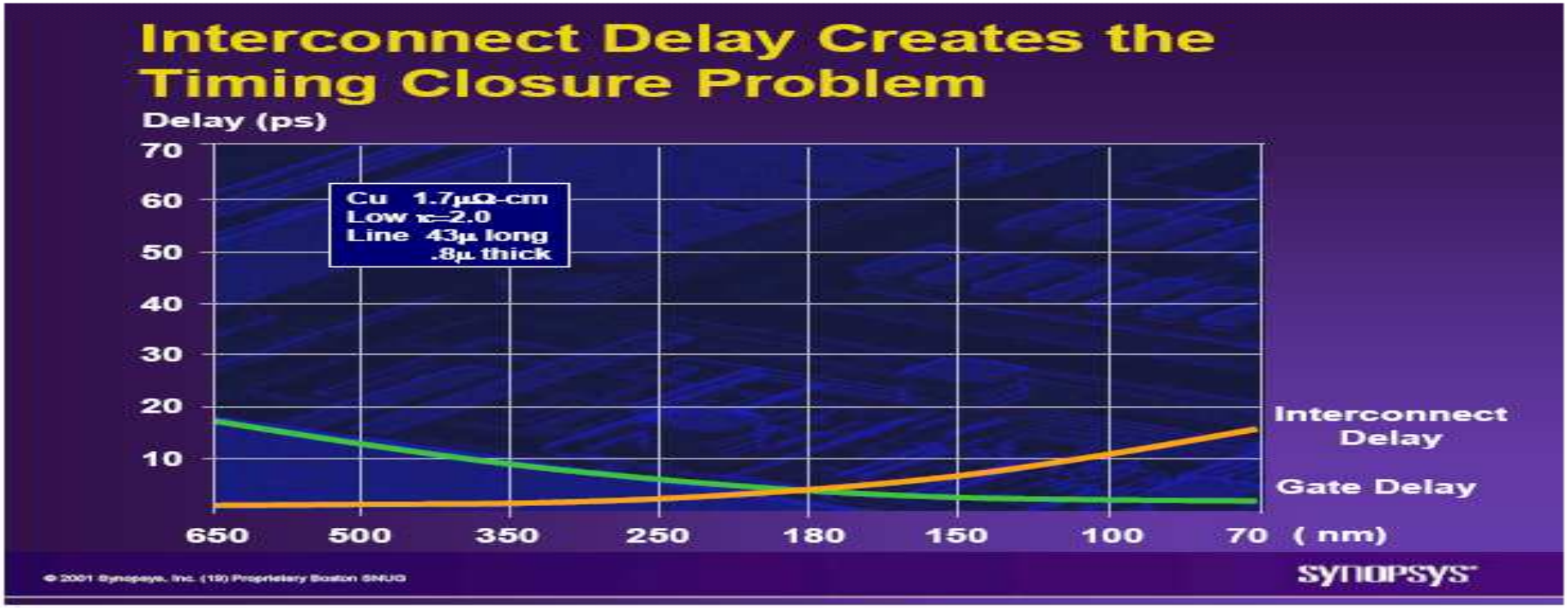
Semiconductor Industry is Facing an Inflection Point

Dimensional Scaling has reached Diminishing Returns

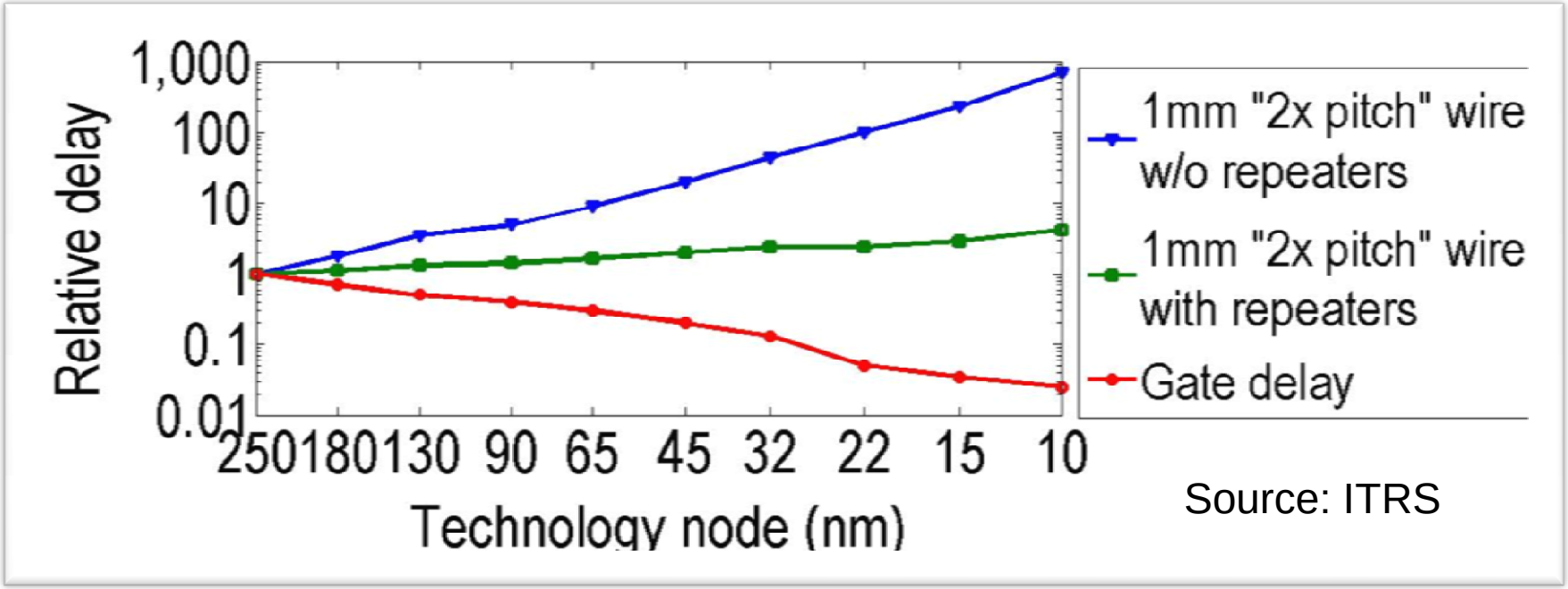
MonolithIC3D™

The Current 2D-IC is Facing Escalating Challenges - I

- On-chip interconnect is
  - Dominating device power consumption
  - Dominating device performance
  - Penalizing device size and cost



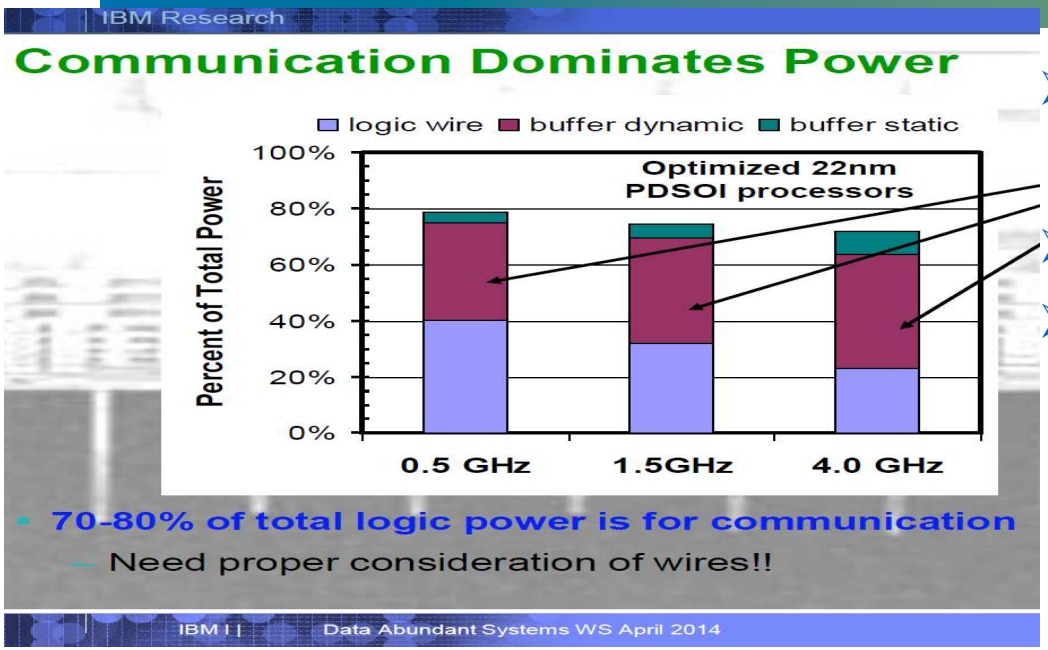
Interconnect Delay
A Big Issue with Scaling



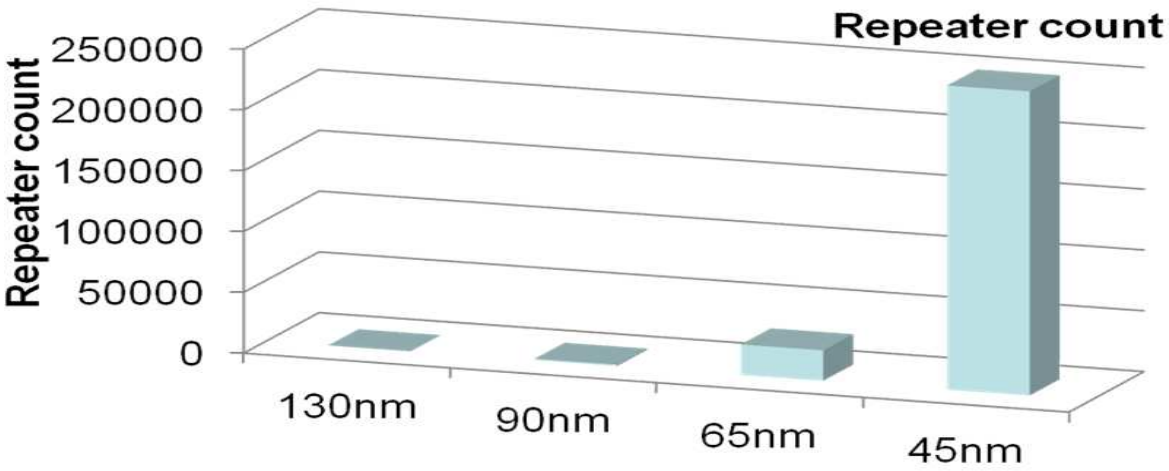
- ***Transistors improve with scaling, interconnects do not***
- ***Even with repeaters, 1mm wire delay ~50x gate delay at 22nm node***

Connectivity Consumes 70-80% of Total Power @ 22nm

Repeaters Consume Exponentially More Power and Area



- *At 22nm, on-chip connectivity consumes **70-80% of total power***
- *Repeater count increases exponentially*
- *At 45nm, repeaters are > 50% of total leakage*

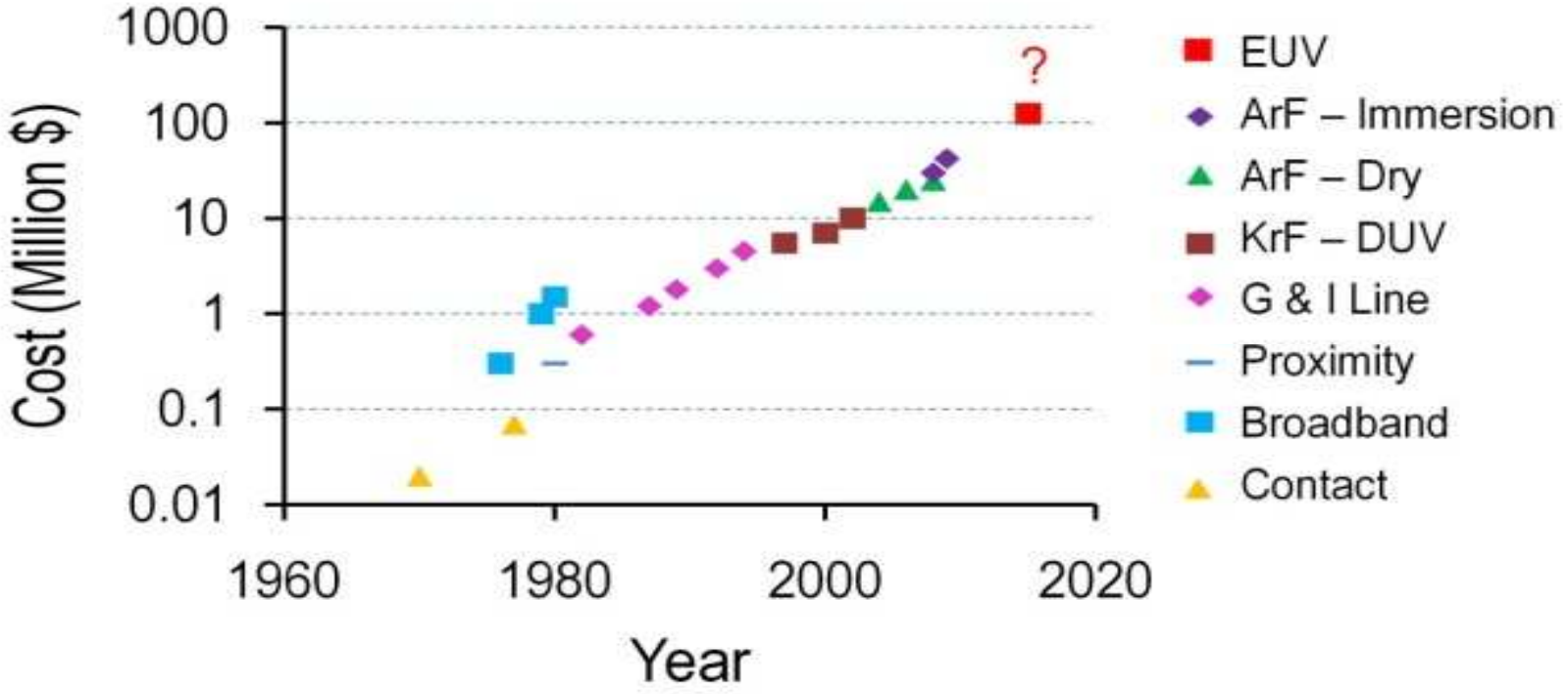


Source: IBM POWER processors
R. Puri, et al., SRC Interconnect Forum, 2006

The Current 2D-IC is Facing Escalating Challenges - II

- Lithography is
  - Dominating Fab cost
  - Dominating device cost and diminishing scaling's benefits
  - Dominating device yield
  - Dominating IC development costs

A Challenge: Lithography



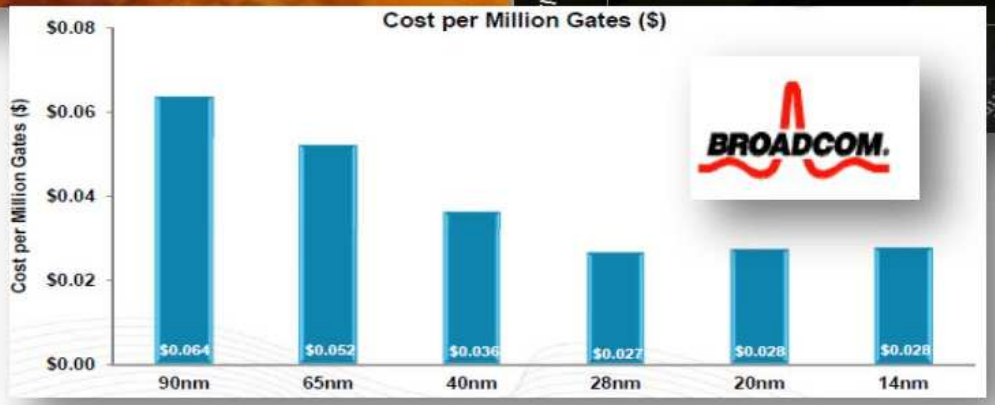
- **Quad-patterning next year → costly. EUV delayed, costly.**
- ***Can we get benefits of scaling <u>without relying on lithography</u>?***

Martin van den Brink -*EVP & CTO*, ASML
ISSCC 2013 & SemiconWest 2013

Cost becomes a concern post 28 nm

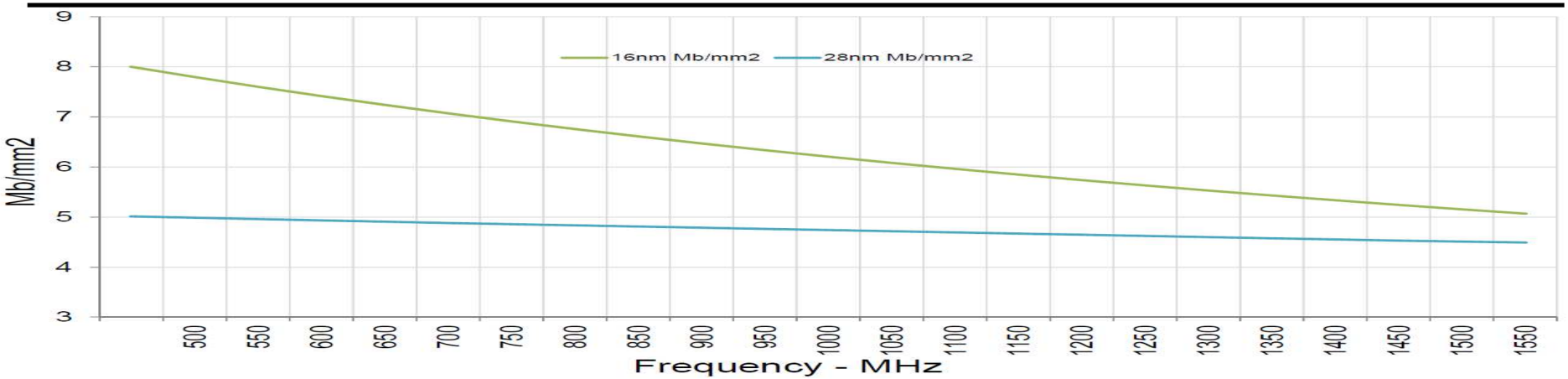
ASML

Public



Sources: nVidea, ITPC, nov, 2011
Broadcom, IMEC, may 2012
GF, ISS, jan 2013

Embedded SRAM isn't Scaling Beyond 28nm (1.1x instead off 4x)
eSRAM > 60% of Die Area => End of Dimension Scaling !



Memory density at 1500MHz and above scales by ~1.1x or less from 28nm to 16nm

For 400G ASSP/ASIC, need to double the SRAM density

- But density improvement from process node N to N+1 is not 2x anymore but by 1.1x
- 500mm^2 die in 28nm for 200G with 60% SRAM ported to 16nm for 400G will be ~ 745mm^2
- Die size close to reticle limit - exacerbates yield & cost of lower end segments

At > 400G, embedding all the SRAM would make the die size bigger than reticle limit

Dinesh Maheshwari, CTO, Memory Products Division at Cypress Semiconductors, ISSCC2014

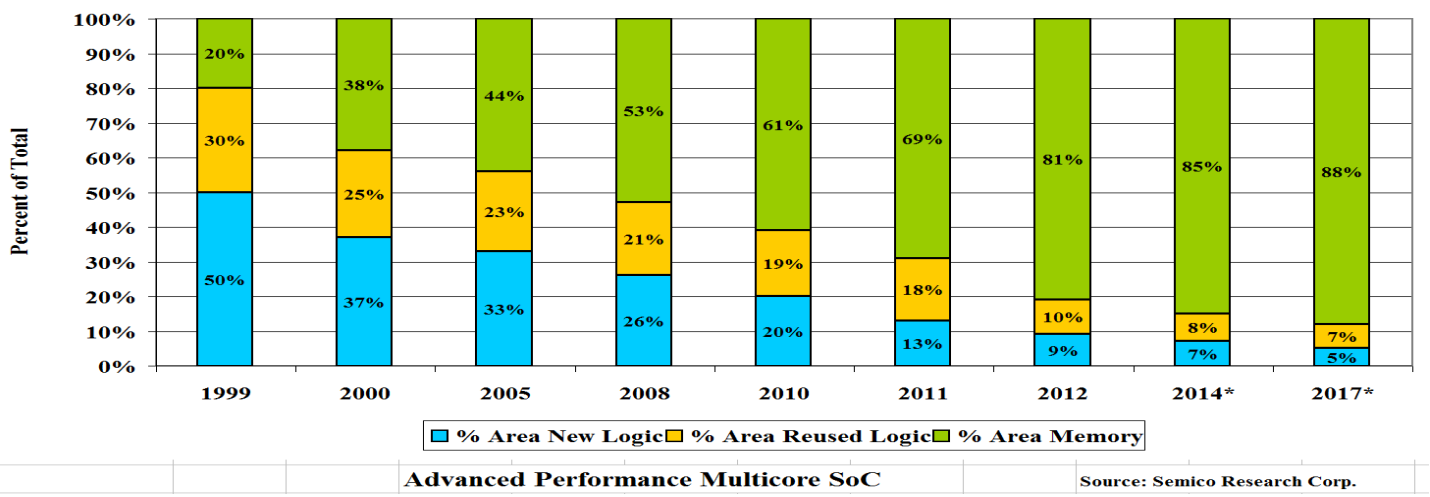
Embedded SRAM isn't Scaling Beyond 28nm

eSRAM > 60% of Die Area => End of Dimensional Scaling !

*

	Early production	2011 - 2012	2013 - 2014	2015 - 2016	2017 - 2018	2019 - ...
		22 – 20nm	16 -14nm	10nm	7nm	5nm
LOGIC	Memory (um^2)	SRAM 0.09-0.08	SRAM 0.08-0.07	SRAM 0.06-0.05	SRAM < 0.05;	SRAM < 0.05; (STT-MRAM)
	Device	Planar, FinFET	FinFET, FDSOI	FinFET	FinFET; (LOCSOI, GAA)	GAA FinFET; (NW)
	Gate EOT (nm)	HKMG 0.9	HKMG 0.8	HKMG 0.7	HKMG 0.7	HKMG 0.7
	Channel n/p	Si	Si / SiGe	Si / (SiGe)	Si / SiGe; (IIIV / Ge)	Si / SiGe; (IIIV / Ge)
	S/D Strain	yes	yes	yes	yes	yes
	Vdd (V)	0.8	0.8	0.8-0.7	0.7-0.5	0.7-0.5

*imec's 2013 International Technology Forum,



MonolithIC3D™

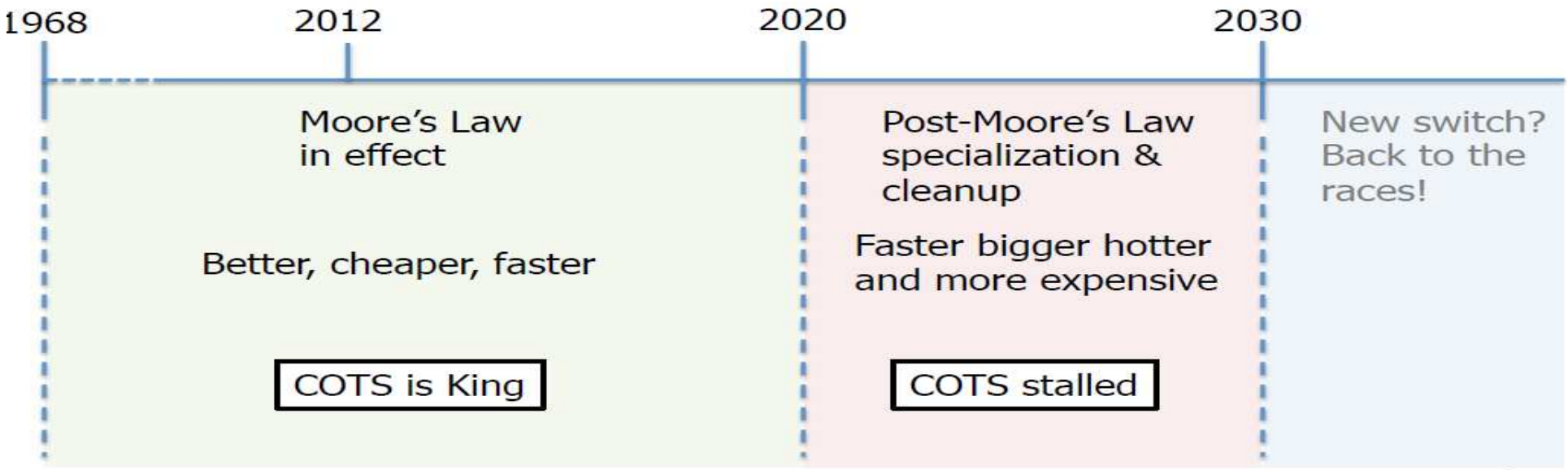
Moore's Law Dead by 2022*

Bob Colwell, Director MTO, DARPA

DARPA

My model: During and After Moore's Law

1. COTS is both problem & opportunity for DoD for next 10 years.
2. Then COTS stalls out. (But DoD doesn't have to!)



MonolithIC3D

*http://www.eetimes.com/document.asp?doc_id=1319330

*CRA/CCC & ACM SIGDA, Pittsburgh, March 2013

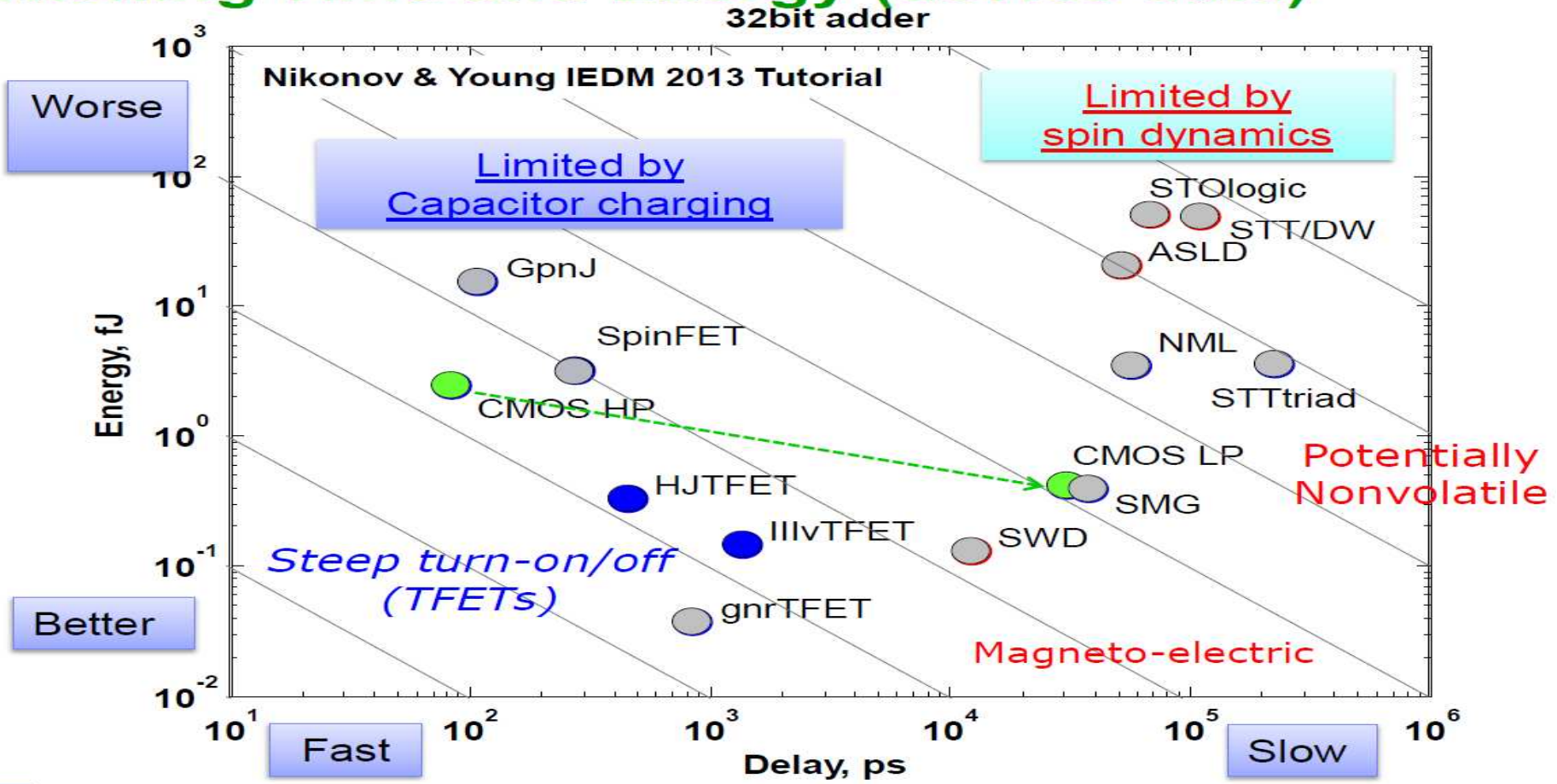
Conclusions:

- Dimensional Scaling (“Moore’s Law”) is already exhibiting diminishing returns
- The road map beyond 2017 (7nm) is unclear
- While the research community is working on many interesting new technologies (see below), none of them seem mature enough to replace silicon for 2019
  - Carbon nanotube
  - Graphene
  - Nanowire
  - Photonics
  - Indium gallium arsenide
  - Spintronics
  - Molecular computing
  - Quantum computing
  - ’2D’ devices: MoS2, etc
- 3D IC is considered, by all, as the near term solution, and **Monolithic 3D IC** is well positioned to be so, as it uses the existing infrastructure
  - **It is safe to state that Monolithic 3D is the only alternative that could be ready for high volume in 2019**

Monolithic 3D™

CMOS is the Best Device Option

Switching Time and Energy (device data)



3D and EDA need to make up for Moore's Law, says Qualcomm*

- "Qualcomm is looking to monolithic 3D and smart circuit architectures to make up for the loss of traditional 2D process scaling as wafer costs for advanced nodes continue to increase. .. Now, although we are still scaling down it's not cost-economic anymore"
- "Interconnect RC is inching up as we go to deeper technology. That is a major problem because designs are becoming interconnect-dominated. Something has to be done about interconnect. What needs to be done is monolithic three-dimensional ICs."
- "TSV...are not really solving the interconnect issue I'm talking about.

  So we are looking at true monolithic 3D. You have normal vias between different stacks."

* Karim Arabi Qualcomm VP of engineering Key Note DAC 2014
<http://www.techdesignforums.com/blog/2014/06/05/karim-arabi-monolithic-3dic-dac-2014/>

Monolithic 3D™

“CEA-Leti Signs Agreement with Qualcomm to Assess Sequential (monolithic)3D Technology”

Business Wire December 08, 2013

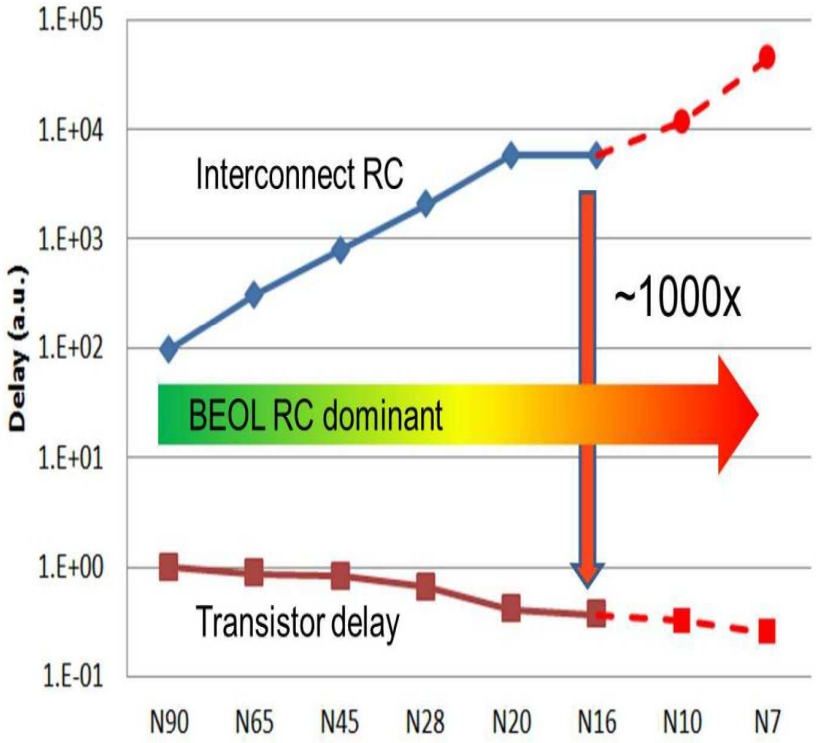


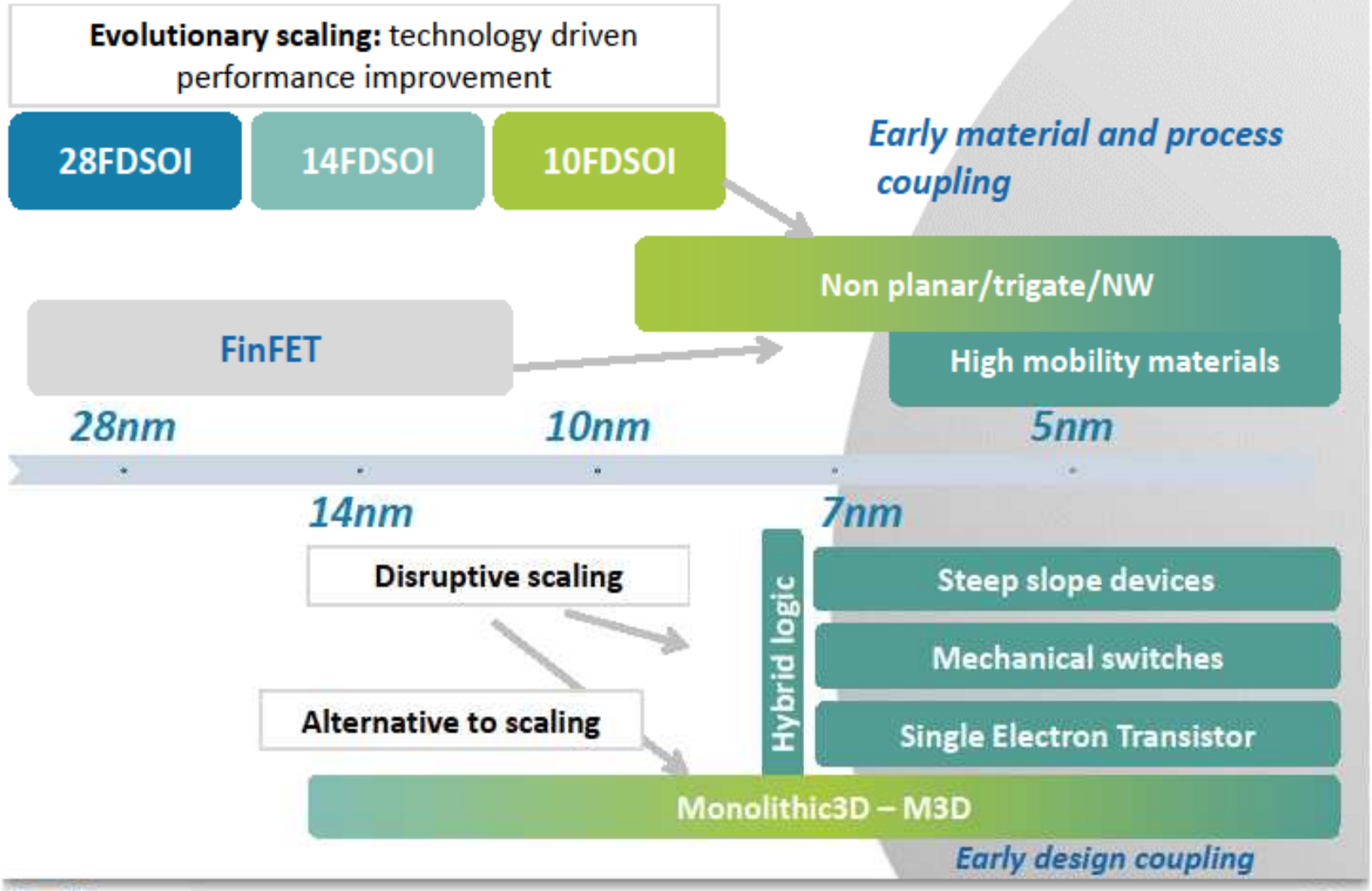
Fig. 17: BEOL performance/area/cost scaling is the foremost issue for 10nm/7nm nodes.

“Monolithic 3D (M3D) is an emerging integration technology poised to reduce the gap significantly between transistors and interconnect delays to extend the semiconductor roadmap way beyond the 2D scaling trajectory predicted by Moore’s Law.”

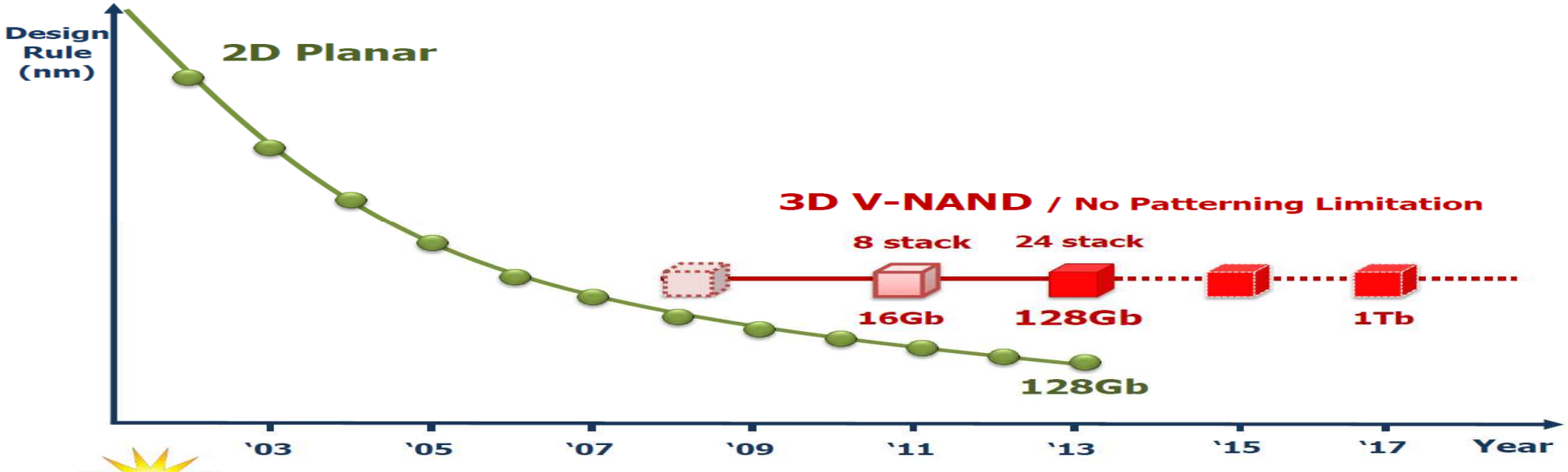
Geoffrey Yeap,
VP of Technology at Qualcomm,
Invited paper, IEDM 2013

MonolithIC3D™

Device technology roadmap



V-NAND Era for the Future



Flash Memory SUMMIT

SAMSUNG

Paradigm Shift from Drive to Fly

Innovative Technology

Material

Structure

Integration

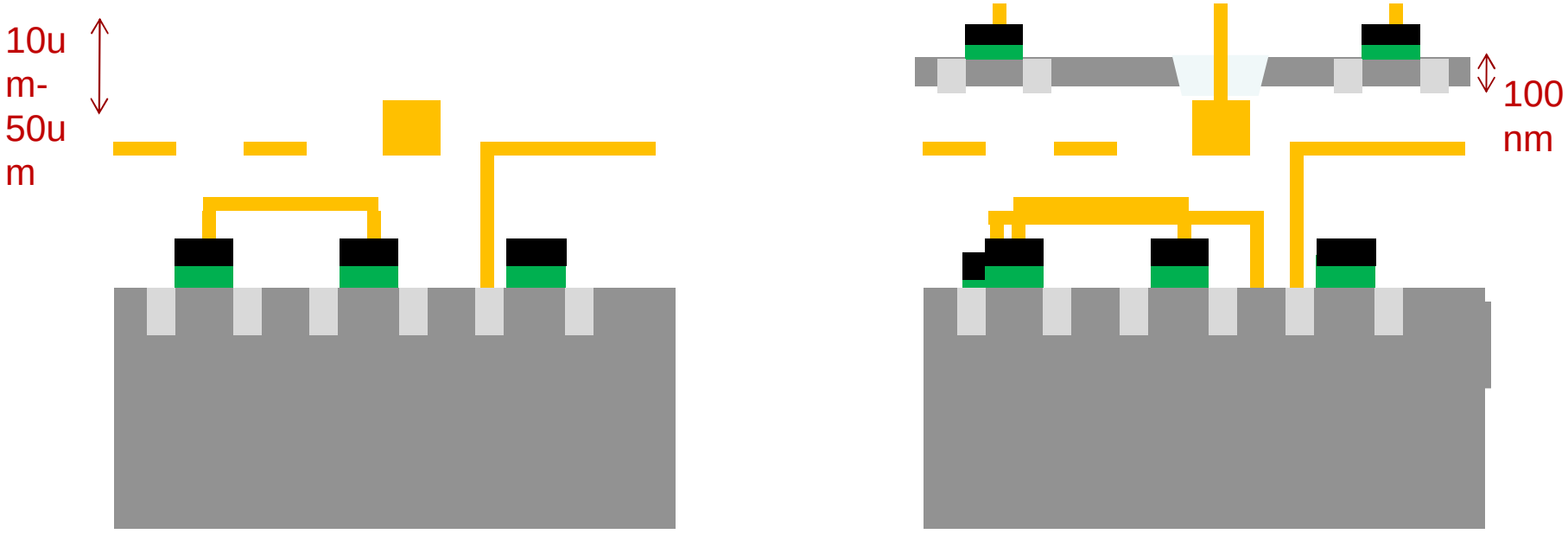
Flash Memory SUMMIT

SAMSUNG

Two Types of 3D Technology

3D-TSV
Transistors made on separate wafer @ high temperature, then thin + align + bond

Monolithic 3D
Transistors made monolithically atop wiring (@ sub-400ºC for logic)



TSV pitch > 1*um**

TSV pitch ~ 50-100*nm*

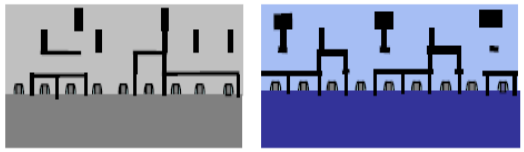
* [Reference: P. Franzon: Tutorial at IEEE 3D-IC Conference 2011]

MONOLITHIC
10,000x the Vertical Connectivity of TSV

Enables:	TSV	Monolithic
Layer Thickness	~50µ	~50nm
Via Diameter	~5µ	~50nm
Via Pitch	~10µ	~100nm
Wafer (Die) to Wafer Alignment	~1µ	~1nm
Overall Scale	microns	nano-meters

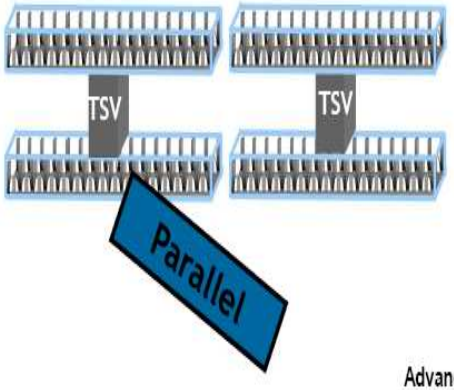
Monolithic/ Sequential 3D

- Wafers Processed separately



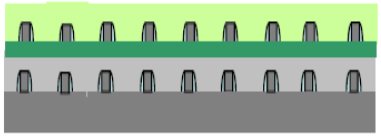
- Stacking and Contacting

- 1 TSV/ Block of 10,000 FETs

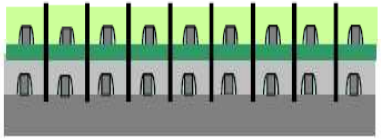


- Bottom transistor processing

- Top FET processing



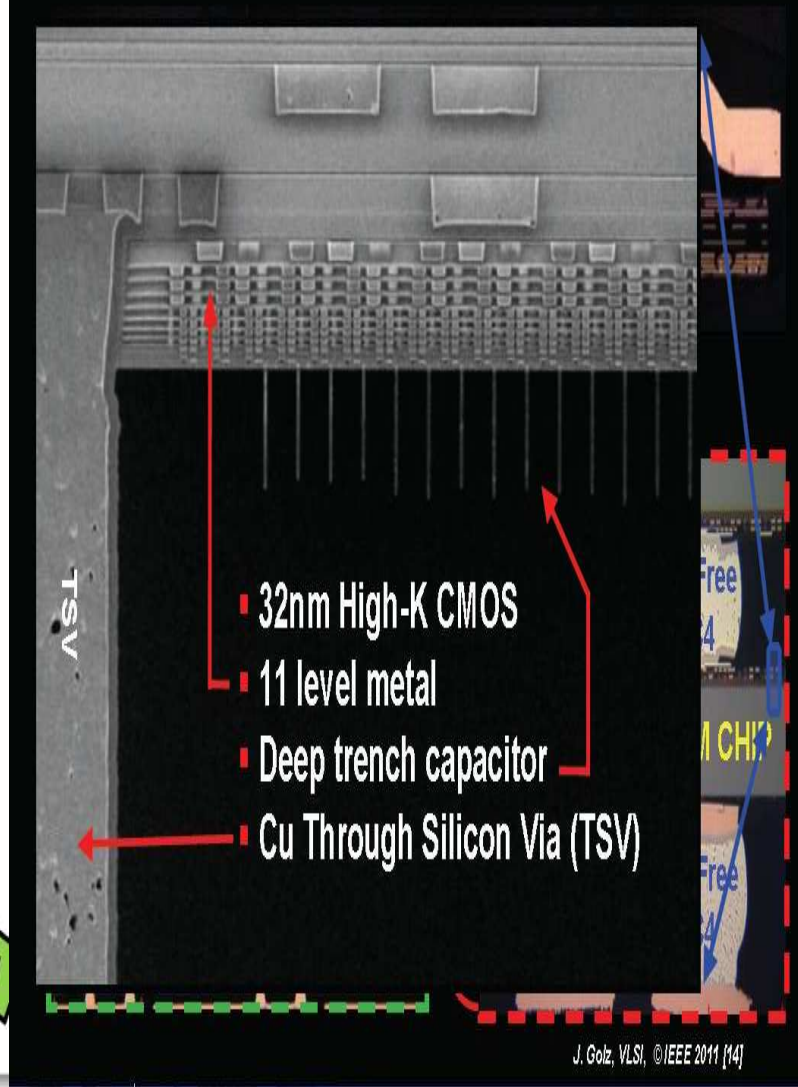
- Contacting



- 1 vertical contact/ FET

Advances, Challenges and Opportunities in 3D CMOS Sequential Integration.
P. Batude et al, IEDM 2011

IBM 32nm Qualification Test site: Module Panorama



3D ICs in older process (65*nm*) is better than 2D ICs built with a newer process (32*nm*)

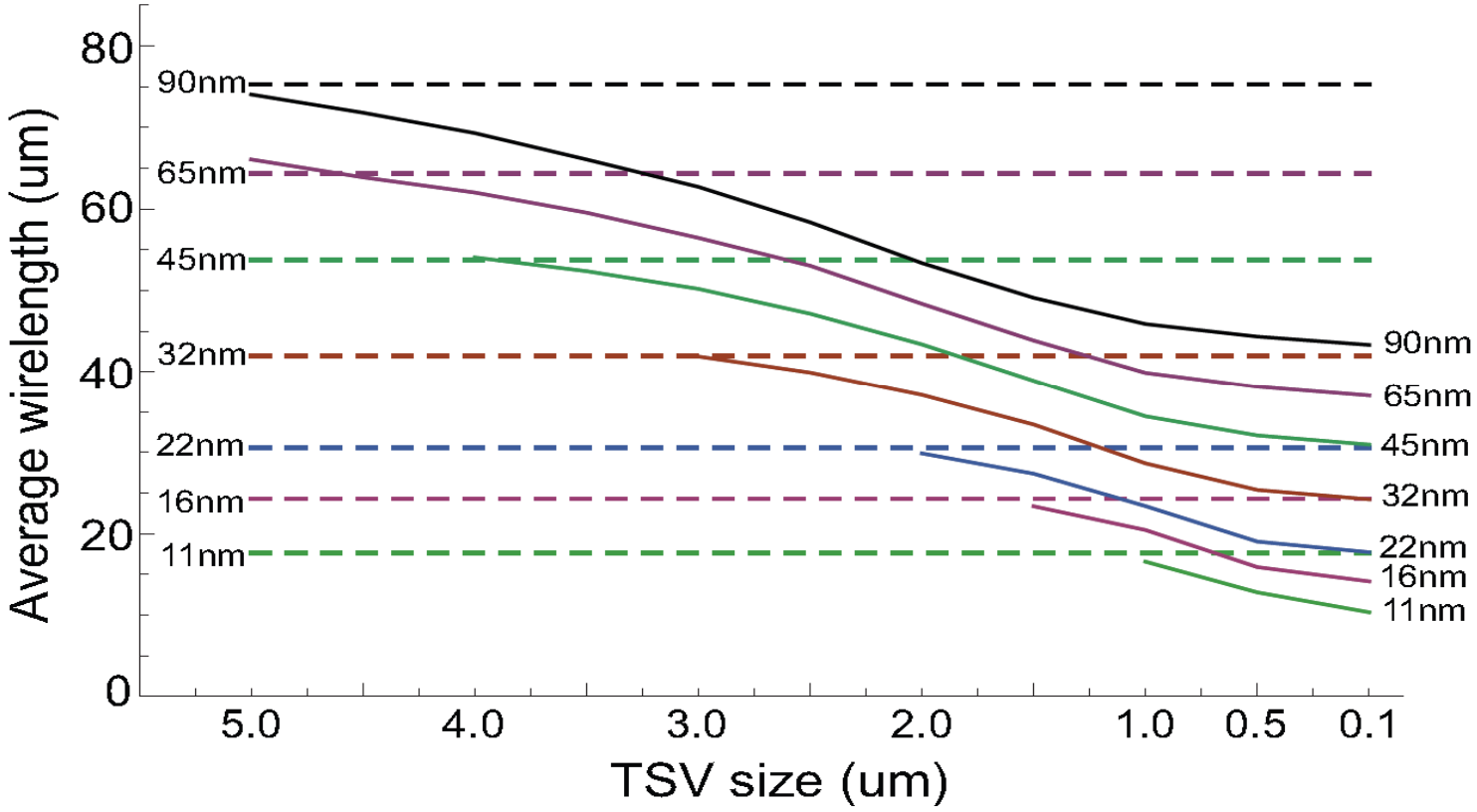


Fig. 6. Cross comparison among various 2D and 3D technologies. Dashed lines are wirelengths of 2D ICs. # dies: 4.

*IEEE IITC11 Kim

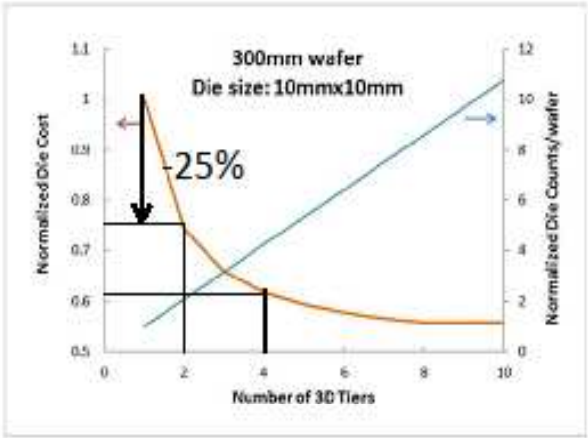
MonolithIC3D

Interest for M3D

Source: G. Bartlett, Global Foundries, SMC 2013

Without scaling avoid fab and process costs increase

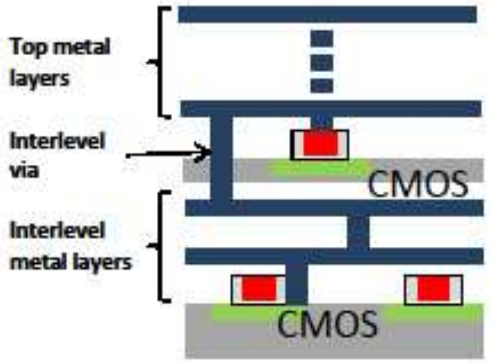
Source: R. Gilmore, Qualcomm VP, ESSIRC 2012



Stack 2 layers: 25% die cost reduction

- **Average gain benchmark for 6 circuits/planar**

1 node gain without scaling



3 metal levels

The Monolithic 3D Challenge

Why is it not already in wide use?

- ***Processing on top of copper interconnects should not make the copper interconnect exceed 400°C***
  - **How to bring mono-crystallized silicon on top at less than 400°C**
  - **How to fabricate state-of-the-art transistors on top of copper interconnect and keep the interconnect below at less than 400°C**
- ***Misalignment of pre-processed wafer to wafer bonding step is ~1um***
  - **How to achieve 100nm or better connection pitch**
  - **How to fabricate thin enough layer for inter-layer vias of ~50nm**

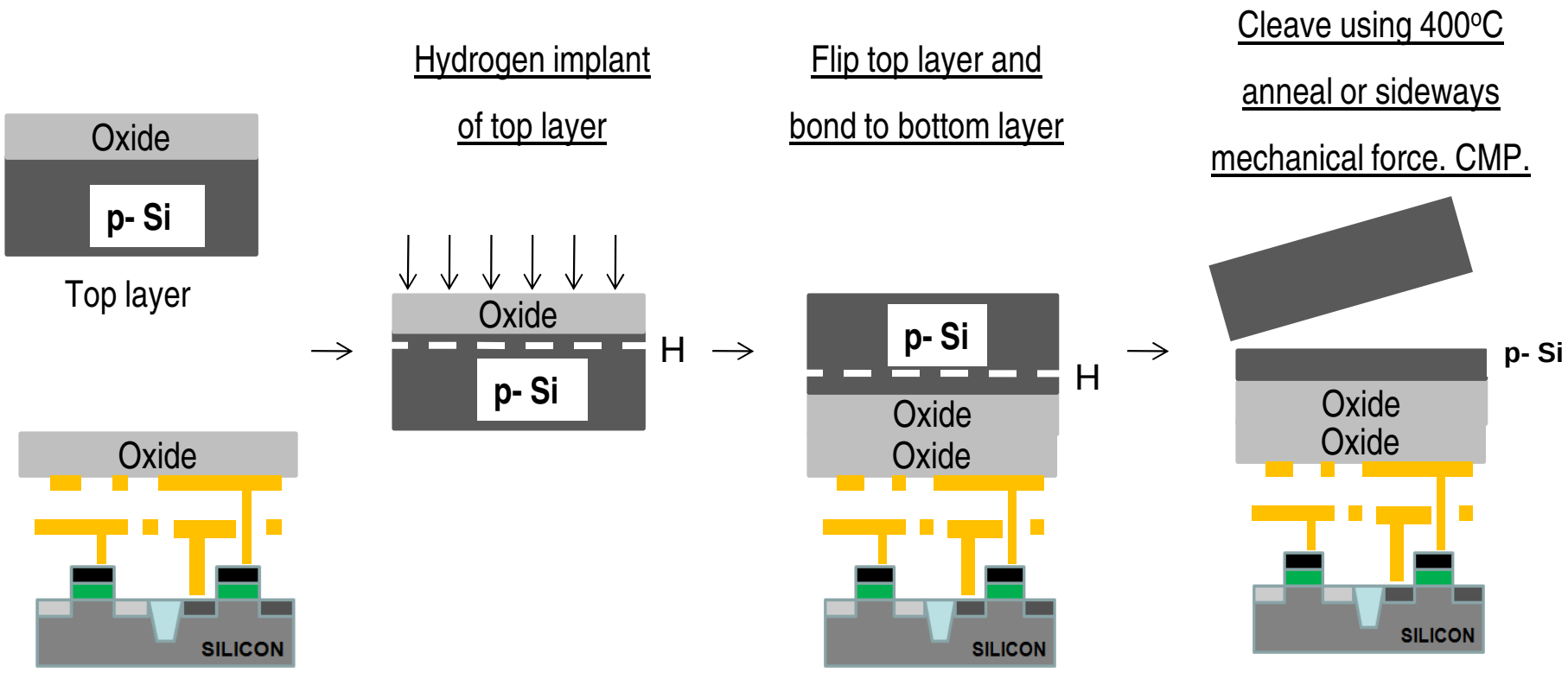
MonolithIC 3D – Breakthrough

3 Classes of Solutions (3 Generations of Innovation)

- **RCAT** *(2009)* – Process the high temperature on generic structures prior to ‘smart-cut’, and finish with cold processes – Etch & Depositions
- **Gate Replacement** *(2010)* (=Gate Last, HKMG) - Process the high temperature on repeating structures prior to ‘smart-cut’, and finish with ‘gate replacement’, cold processes – Etch & Depositions
- **Laser Annealing** *(2012)* – Use short laser pulse to locally heat and anneal the top layer while protecting the interconnection layers below from the top heat

MonolithIC3D™

Layer Transfer ("Ion-Cut"/"Smart-Cut") → The Technology Behind SOI



Similar process (bulk-to-bulk) used for manufacturing all SOI wafers today

MonolithIC3D™

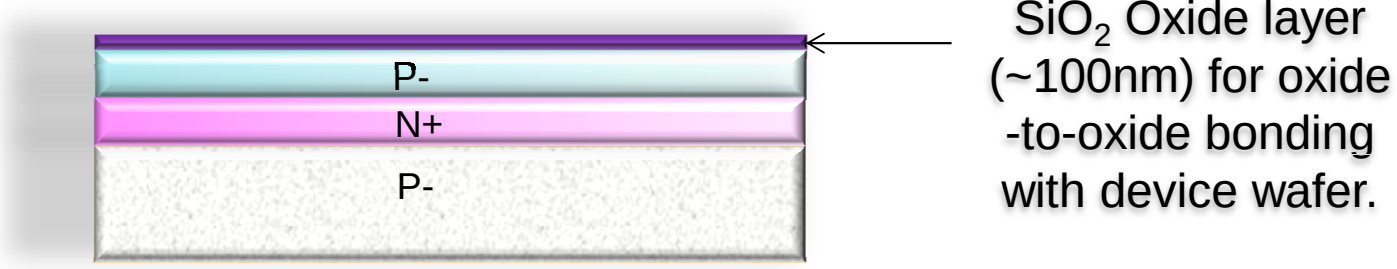
MonolithIC 3D - 3 Classes of Solutions

- **RCAT –** Process the high temperature on generic structure prior to ‘smart-cut’, and finish with cold processes – Etch & Depositions
- Gate Replacement (=Gate Last, HKMG) - Process the high temperature on repeating structure prior to ‘smart-cut’, and finish with ‘gate replacement’, cold processes – Etch & Depositions
- **Laser Annealing** – Use short laser pulse to locally heat and anneal the top layer while protecting the interconnection layers below from the top heat

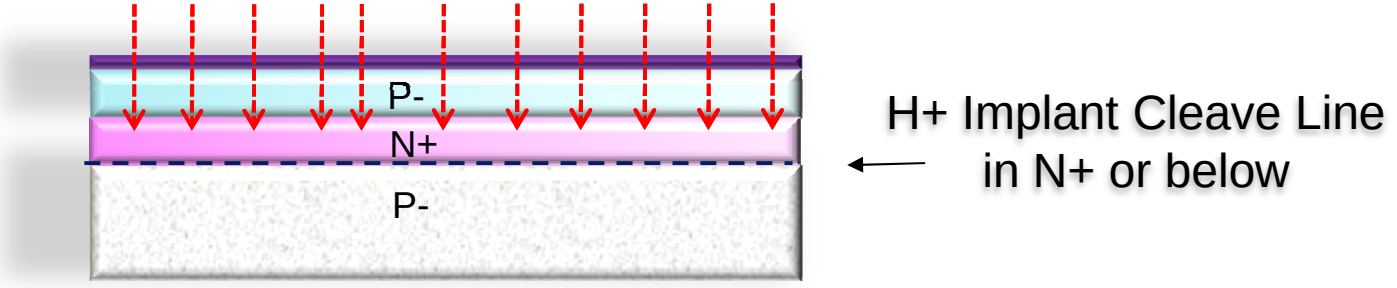
MonolithIC3D™

Step 1. Donor Layer Processing

Step 1 **- Implant and activate unpatterned N+ and P- layer regions in standard donor wafer at high temp. (~900°C) before layer transfer. Oxidize (or CVD oxide) top surface.**



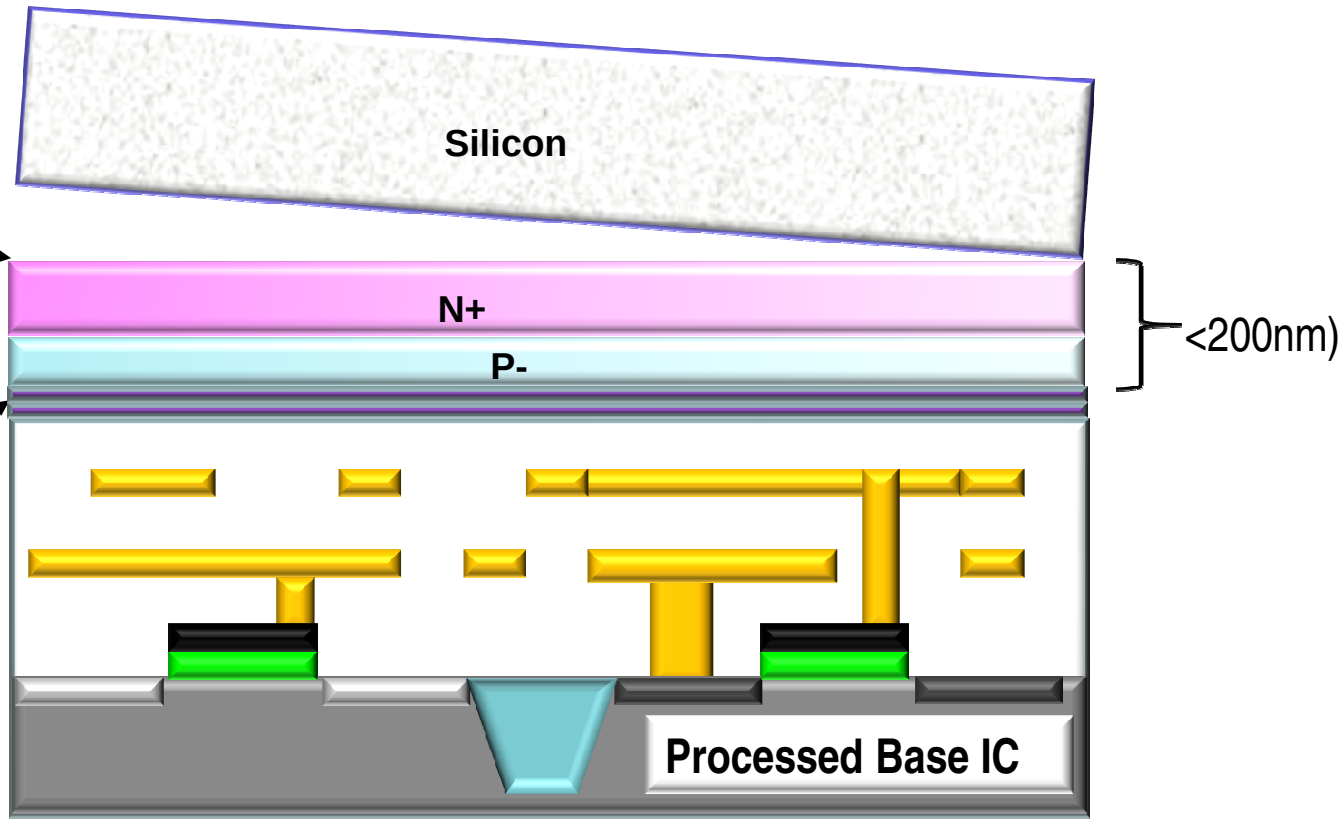
Step 2 **- Implant H+ to form cleave plane for the ion cut**



Step 3 - Bond and Cleave: Flip Donor Wafer and Bond to Processed Device Wafer

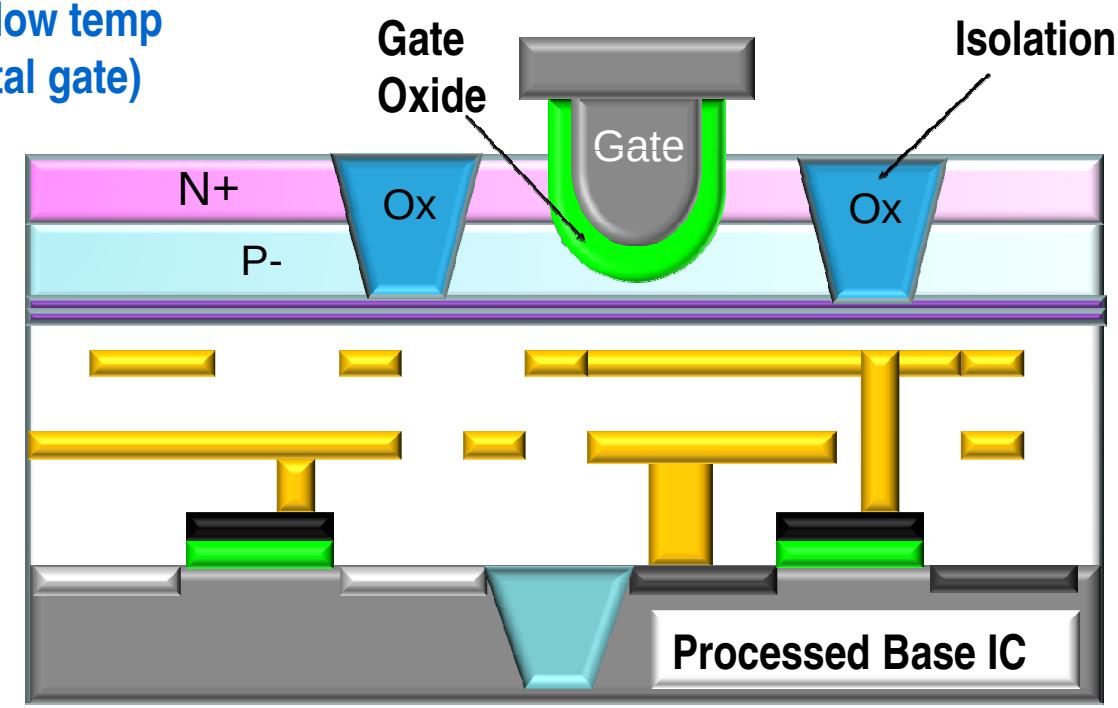
Cleave along H+ implant line using 400°C anneal or sideways mechanical force. Polish with CMP.

SiO_2 bond layers on base and donor wafers (alignment not an issue with blanket wafers)



Step 4 - Etch and Form Isolation and RCAT Gate

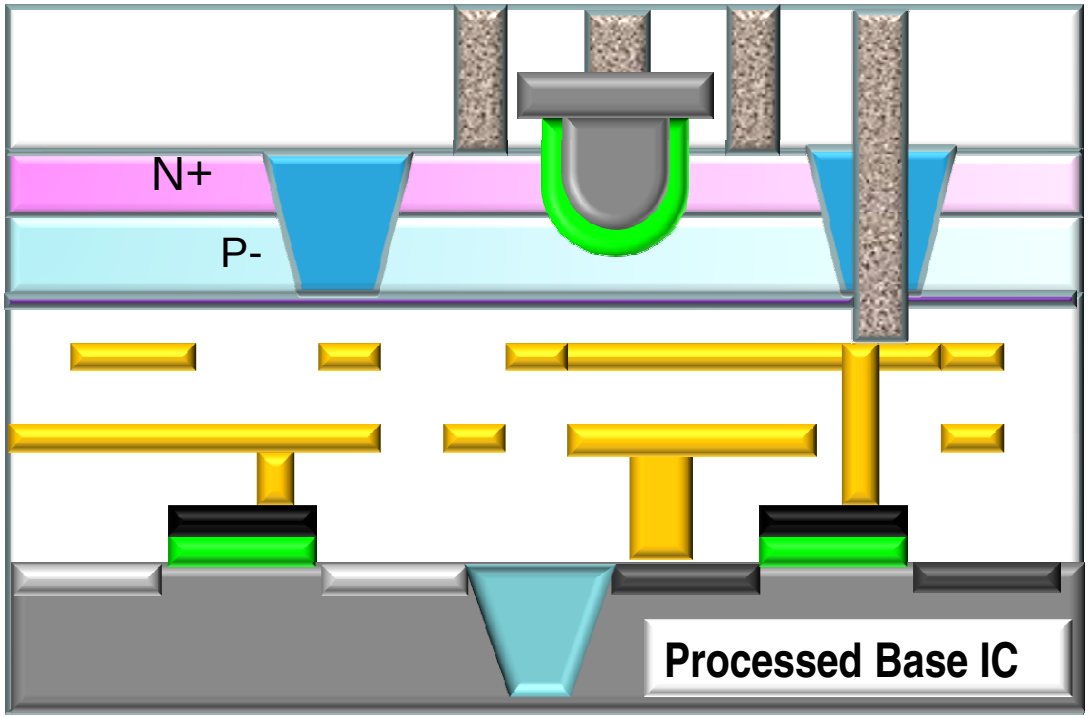
- Litho patterning with features aligned to bottom layer
- Etch shallow trench isolation (STI) and gate structures
- Deposit SiO_2 in STI
- Grow gate with ALD, etc. at low temp (<350º C oxide or high-K metal gate)



Advantage: Thinned donor wafer is transparent to litho, enabling direct alignment to device wafer alignment marks: no indirect alignment. (common for TSV 3DIC)

Step 5 – Etch Contacts/Vias to Contact the RCAT

- *Complete transistors, interconnect wires on 'donor' wafer layers*
- *Etch and fill connecting contacts and vias from top layer aligned to bottom layer*

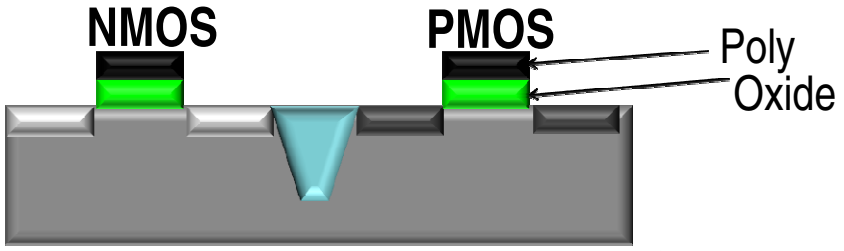


MonolithIC 3D - 3 Classes of Solutions

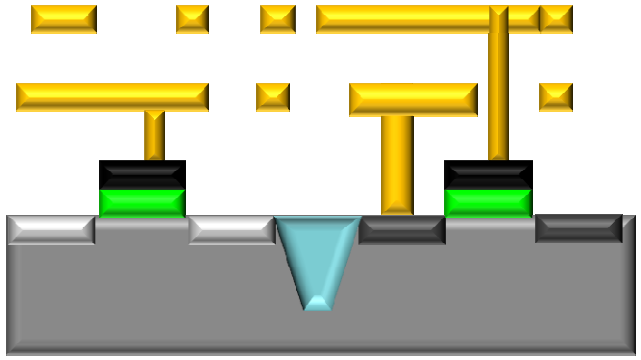
- RCAT – Process the high temperature on generic structure prior to ‘smart-cut’, and finish with cold processes – Etch & Depositions
- **Gate Replacement (=Gate Last, HKMG) -** Process the high temperature on repeating structure prior to ‘smart-cut’, and finish with ‘gate replacement’, cold processes – Etch & Depositions
- **Laser Annealing** – Use short laser pulse to locally heat and anneal the top layer while protecting the interconnection layers below from the top heat

MonolithIC3D™

A Gate-Last Process for Cleave and Layer Transfer



Donor wafer



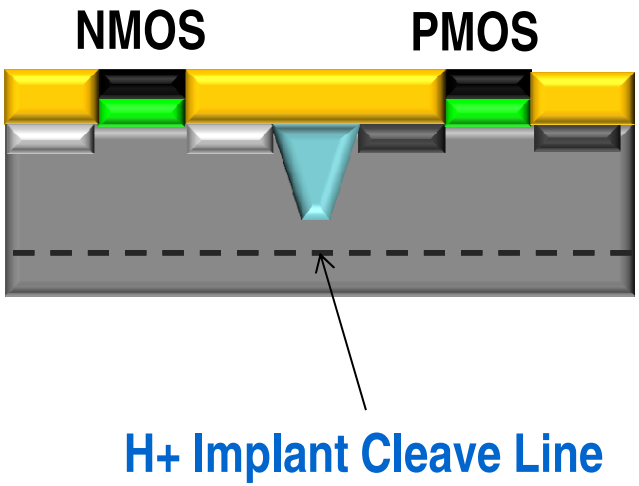
Device wafer

- Fully constructed transistors attached to each other; no blanket films
- proprietary methods align top layer atop bottom layer

A Gate-Last Process for Cleave and Layer Transfer

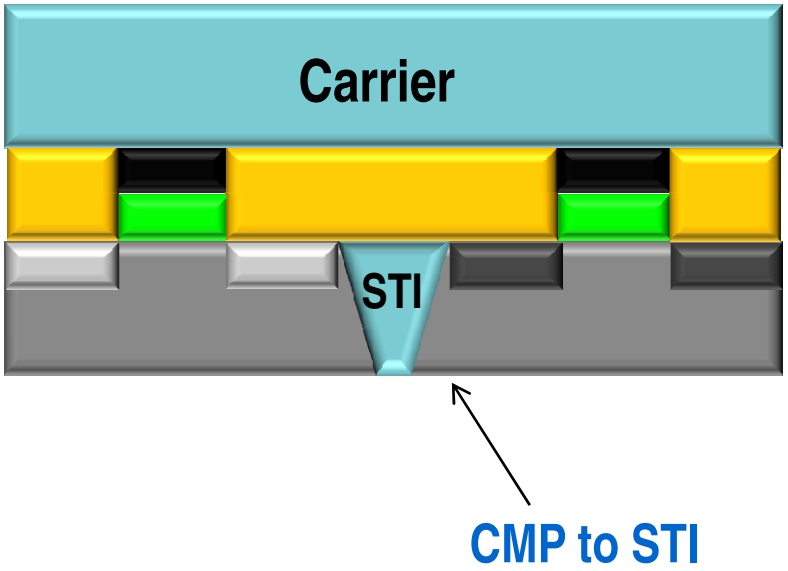
Step 3.

Implant H for cleaving

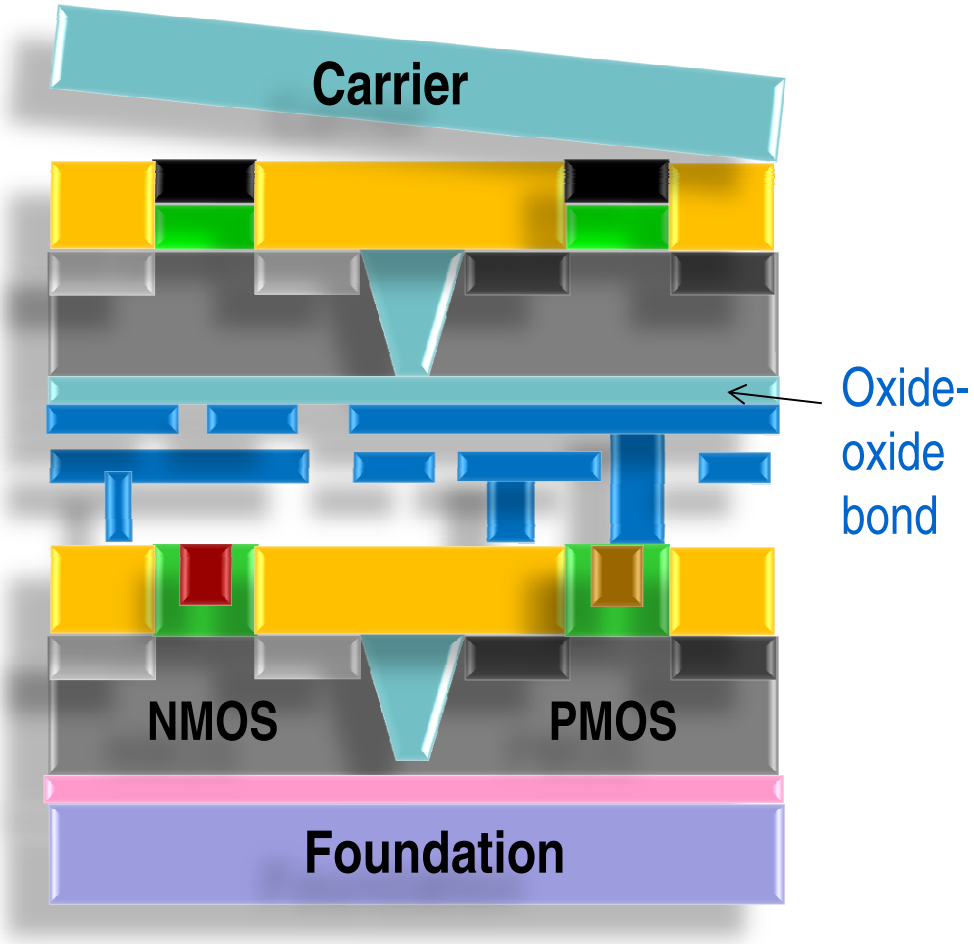


Step 4.

- Bond to temporary carrier wafer (adhesive or oxide-to-oxide)
- Cleave along cut line
- CMP to STI



A Gate-Last Process for Cleave and Layer Transfer



Step 5.

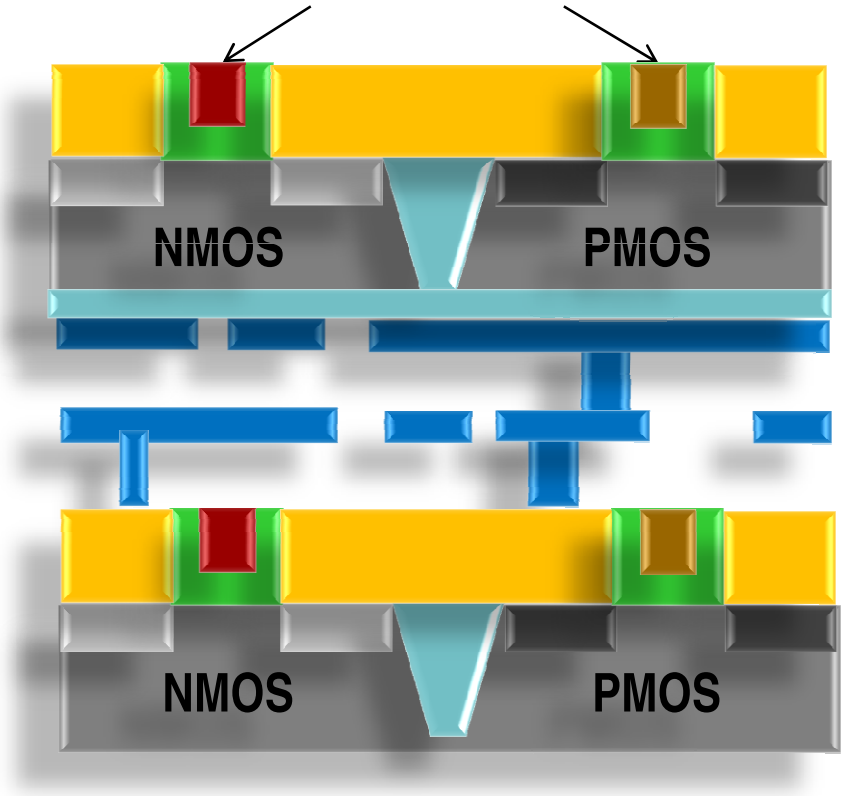
- Low-temp oxide deposition
- Bond to bottom layer
- Remove carrier

A Gate-Last Process for Cleave and Layer Transfer

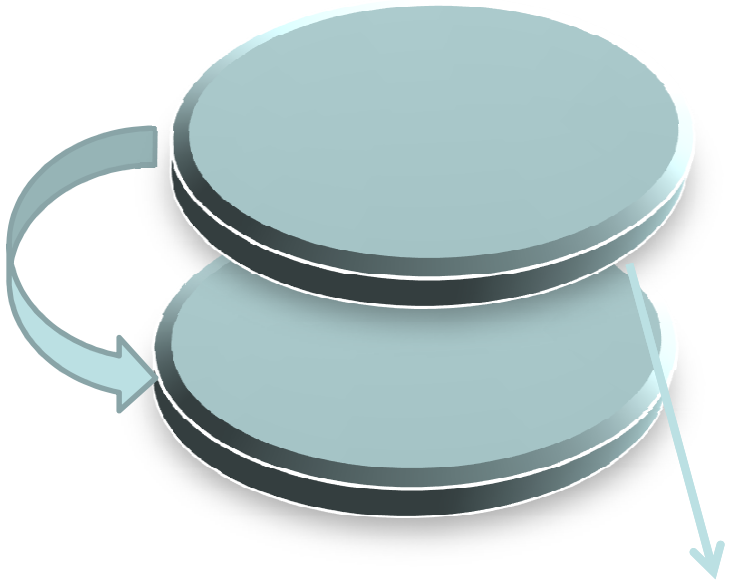
Step 6. On transferred layer:

- Etch dummy gates
- Deposit gate dielectric and electrode
- CMP
- Etch tier-to-tier vias thru STI
- Fabricate BEOL interconnect

Remove (etch) dummy gates, replace with HKMG



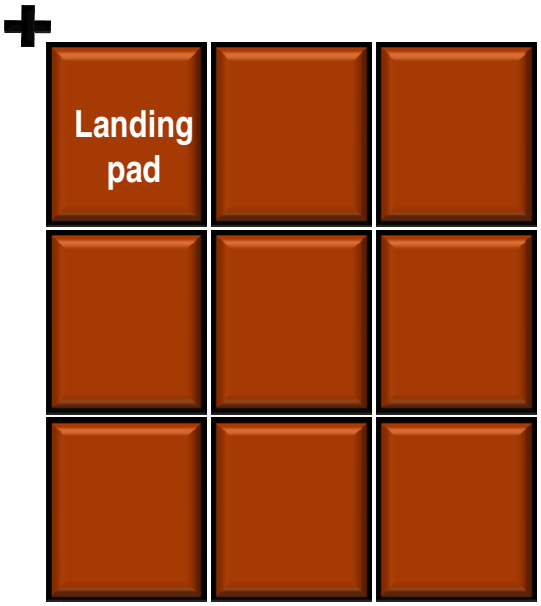
Path 2 – Leveraging Gate Last + Innovative Alignment

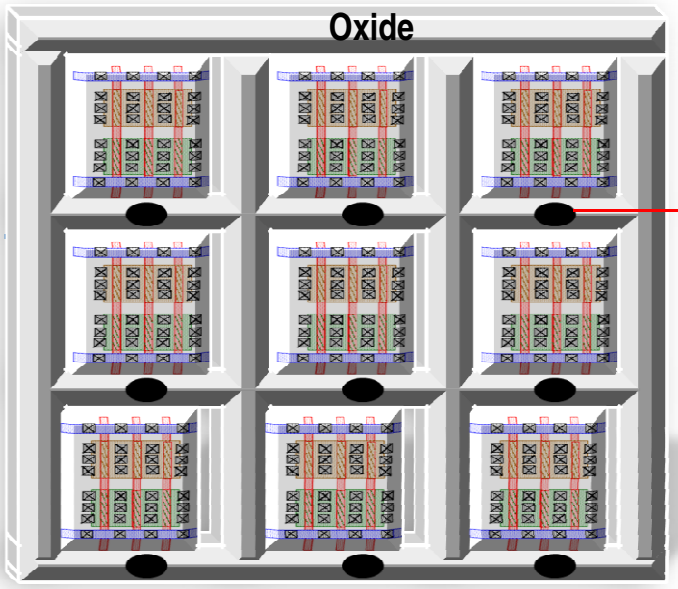


1µ Misalignment

- Misalignment of pre-processed wafer to wafer bonding step is ~1um
  - How to achieve 100nm or better connection pitch
  - How to fabricate thin enough layer for inter-layer vias of ~50nm

Novel Alignment Scheme using Repeating Layouts

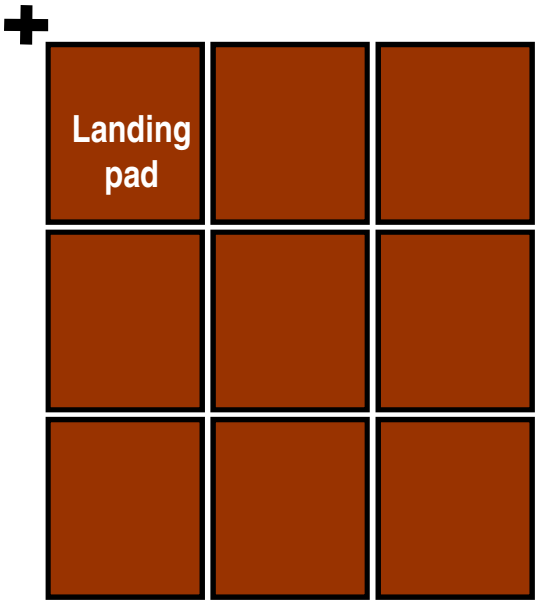




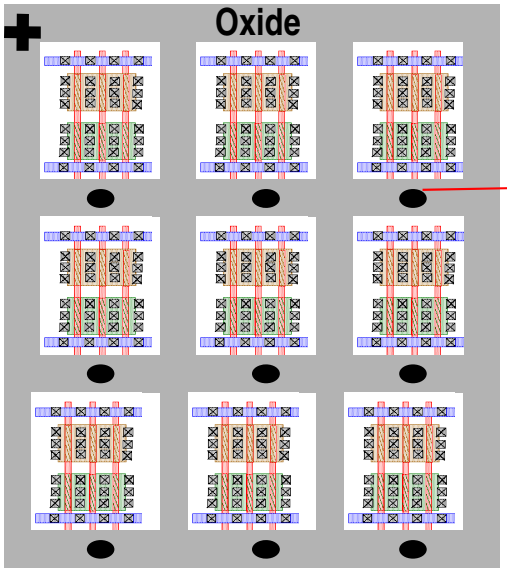
- Even if misalignment occurs during bonding → repeating layouts allow correct connections
- Above representation simplistic (high area penalty)

A More Sophisticated Alignment Scheme

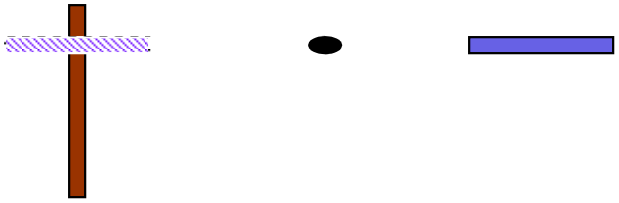
Bottom layer layout



Top layer layout



Through-layer connection



MonolithIC 3D - 3 Classes of Solutions

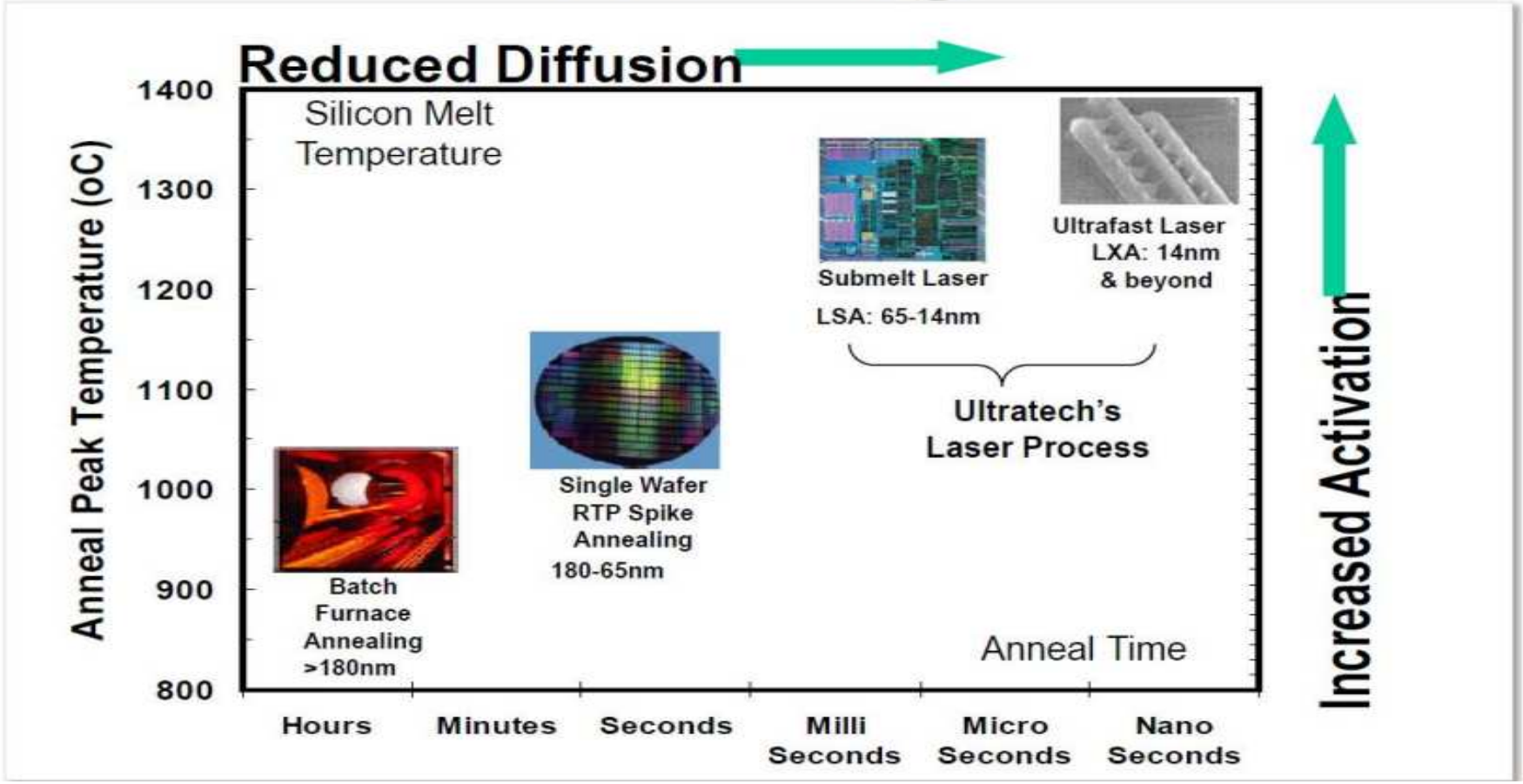
- **RCAT** – Process the high temperature on generic structures prior to ‘smart-cut’, and finish with cold processes – Etch & Depositions
- **Gate Replacement** (=Gate Last, HKMG) - Process the high temperature on repeating structures prior to ‘smart-cut’, and finish with ‘gate replacement’, cold processes – Etch & Depositions
- **Laser Annealing** – Use short laser pulse to locally heat and anneal the top layer while protecting the interconnection layers below from the top heat

MonolithIC3D™

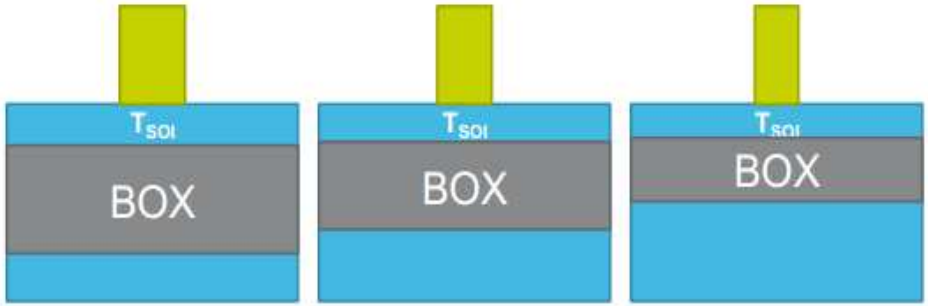
Annealing Trend with Scaling

Ultratech
TECHNOLOGY ACCEPTED

Thermal Annealing Evolution



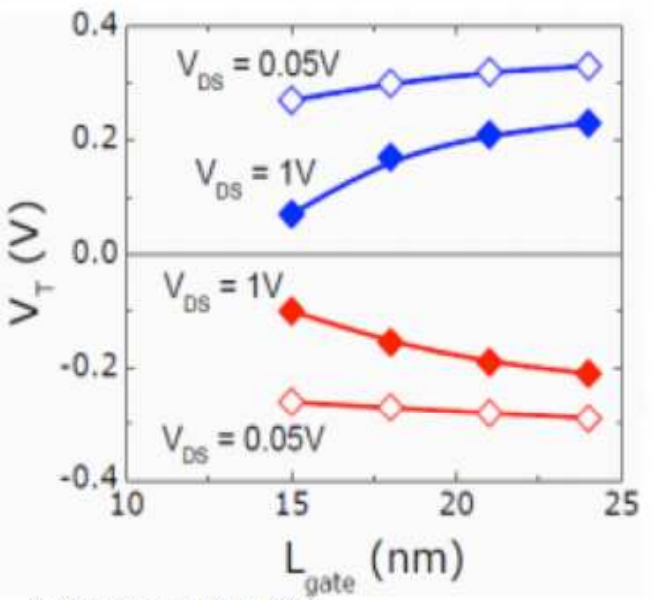
Planar UTBB FD-SOI Scalability: T_{SOI} & T_{BOX}



Node	28nm	14nm	10nm
T_{SOI} (nm)	7.5	6.5	6
T_{BOX} (nm)	25	20	15

O. Faynot et al, IEDM 2010

Si data for L_G=15nm:

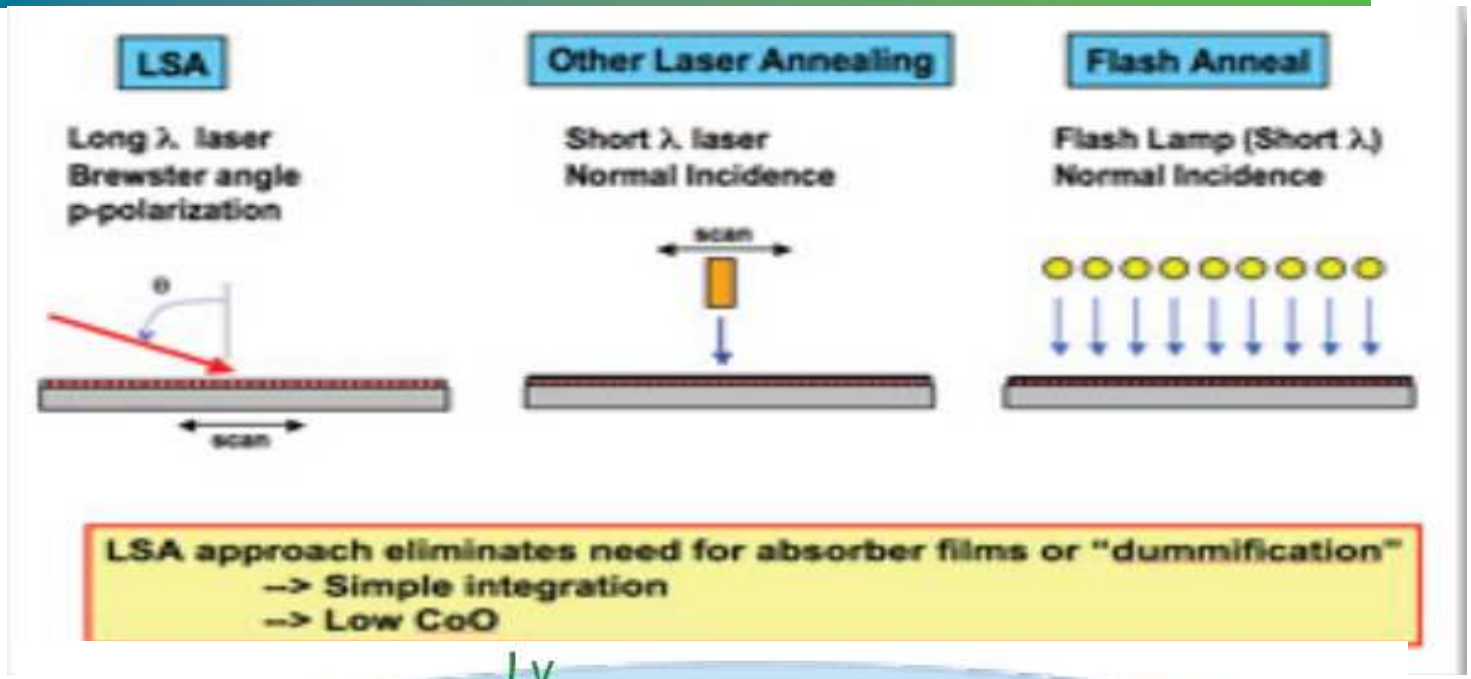


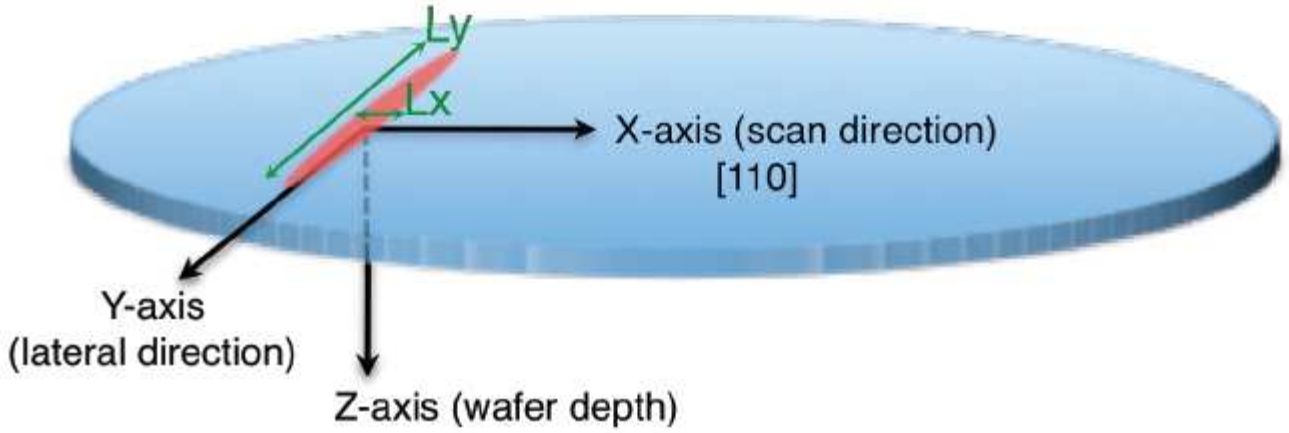
K. Cheng et al, VLSI 2011

- Electrostatic control improved by Thinning T_{BOX}
- Scalability down to 10nm node
- Devices already processed with 3.5nm SOI film!

LSA 100A – Short Pulse, Small Spot

Dwell time ~ 275µs





MonolithIC3D

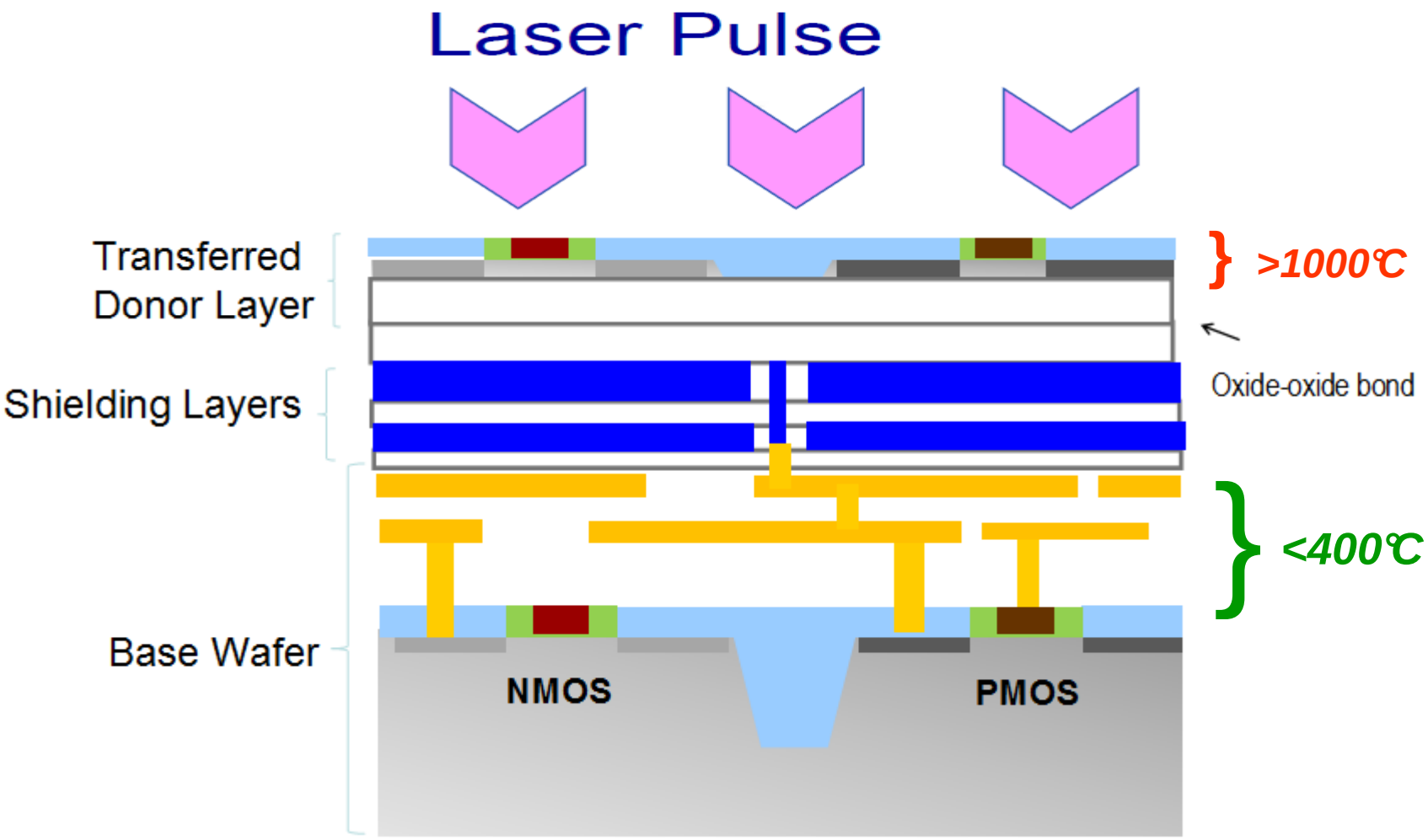
Two Major Semiconductor Trends help make Monolithic 3D Practical NOW

- As we have pushed dimensional scaling:
  - The volume of the transistor has scaled
    - Bulk um-sized transistors → FDSOI & FinFet nm transistors
  - Processing times have trended lower
    - Shallower & sharper junctions, tighter pitches, etc.

=> Much less to heat and for much shorter time

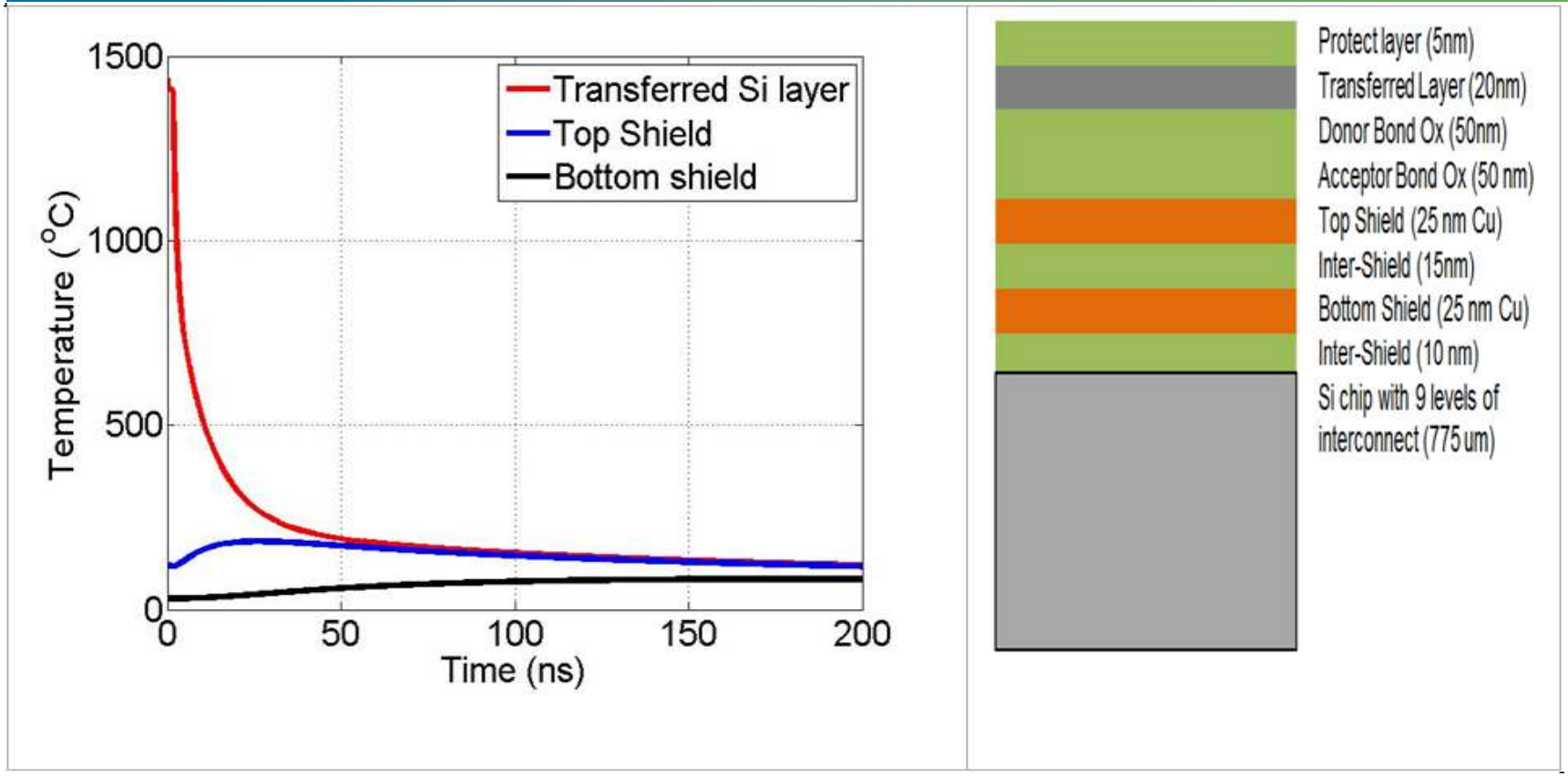
MonolithIC3D™

The Top Layer has a High Temperature >1000°C) without Heating the Bottom Layers (<400℃) !!!



MonolithIC3D™

Process Window Set to Avoid Damage



Temperature variation at the 20 nm thick Si source/drain region in the upper active layer during laser annealing. Note that the shield layers are very effective in preventing any large thermal excursions in the lower layers

MonolithIC3D™

The Monolithic 3D **Advantage**

II. Reduction die size and power – doubling transistor count
- **Extending Moore's law**

Monolithic 3D is far more than just an alternative to 0.7x scaling !!!

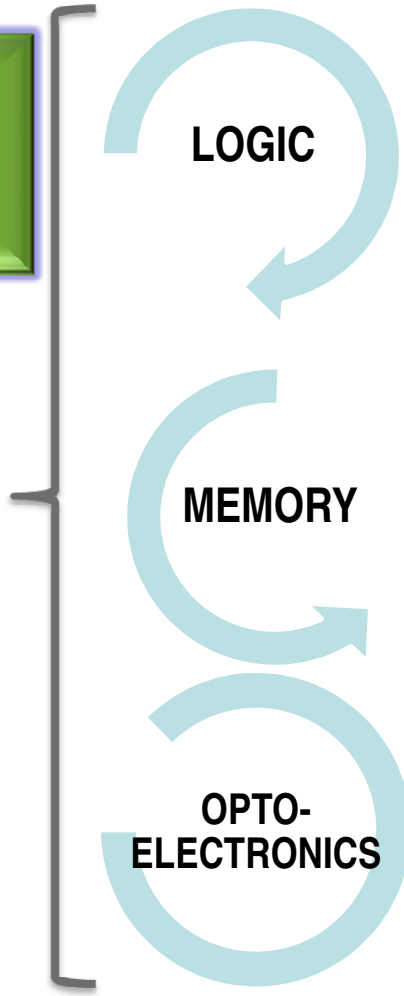
III. Significant advantages from using the same fab, design tools
IV. Heterogeneous Integration
V. Multiple layers Processed Simultaneously - Huge cost reduction (Nx)
VI. Logic redundancy => 100x integration made possible
VII. Enables Modular Design
VIII. Naturally upper layers are SOI
IX. Local Interconnect above and below transistor layer
X. Re-Buffering global interconnect by upper strata
XI. Others
    A. Image sensor with pixel electronics
    B. Micro-display

MonolithIC3D™

Monolithic 3D Provides an Attractive Path to…

Monolithic 3D Integration with Ion-Cut Technology

Can be applied to many market segments



LOGIC

- 3D-CMOS: Monolithic 3D Logic Technology
- 3D-FPGA: Monolithic 3D Programmable Logic
- 3D-GateArray: Monolithic 3D Gate Array
- 3D-Repair: Yield recovery for high-density chips

MEMORY

- 3D-DRAM: Monolithic 3D DRAM
- 3D-RRAM: Monolithic 3D RRAM
- 3D-Flash: Monolithic 3D Flash Memory

OPTO-ELECTRONICS

- 3D-Imagers: Monolithic 3D Image Sensor
- 3D-MicroDisplay: Monolithic 3D Display
- 3D-LED: Monolithic 3D LED

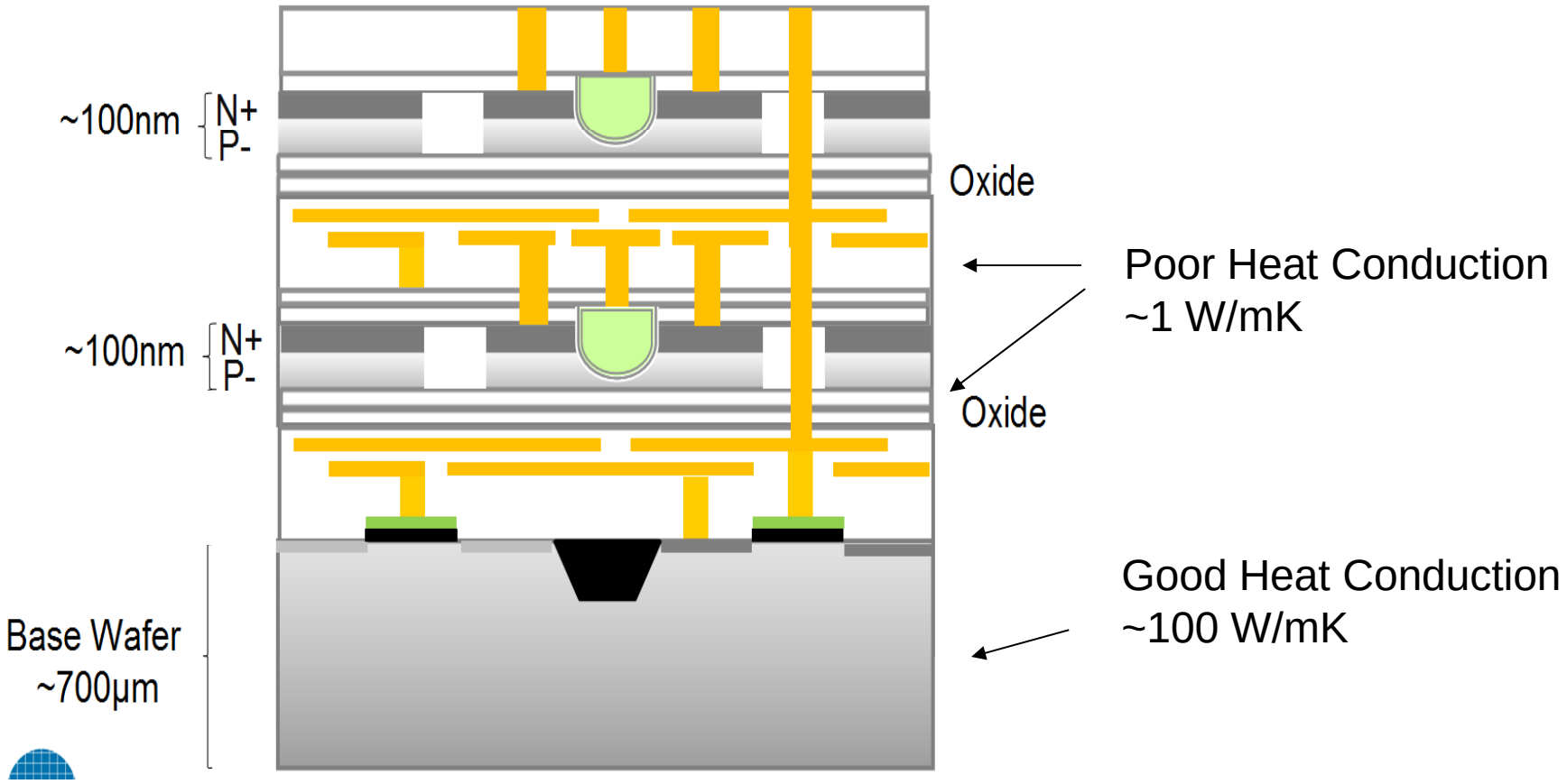
Summary

- Monolithic 3D is now practical and well positioned to keep Moore's Law alive for many years
- Multiple paths to process mono-crystal transistors over copper interconnect
- Monolithic 3D IC provides many opportunities for existing products and for new products & architectures

Back Ups

The Operational Thermal Challenge

- Upper tier transistors are fully surrounded by oxide and have no thermal path to remove operational heat away



The Solution

- Use Power Delivery (Vdd, Vss) Network (“PDN”) also for heat removal
- Add heat spreader to smooth out hot spots
- Add thermally conducting yet electrically non-conducting contacts to problem areas such as transmission gates

IEDM 2012 Paper

Cooling Three-Dimensional Integrated Circuits using Power Delivery Networks (PDNs)

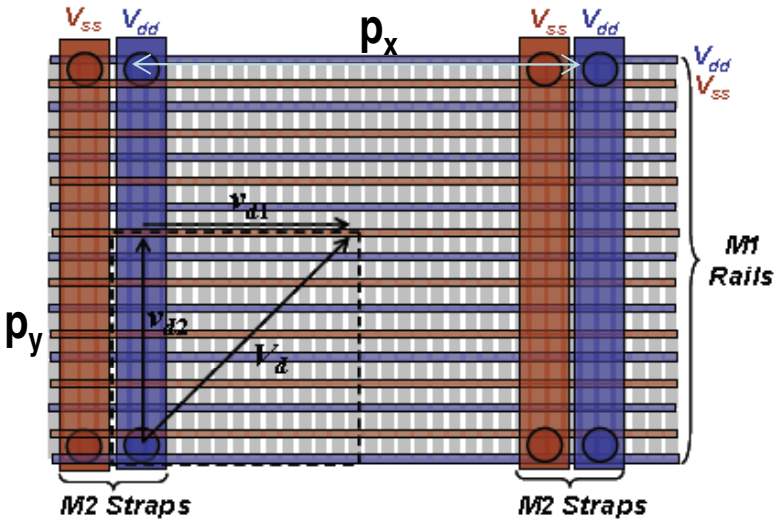
Hai Wei, Tony Wu, Deepak Sekar[+], Brian Cronquist*, Roger Fabian Pease, Subhasish Mitra

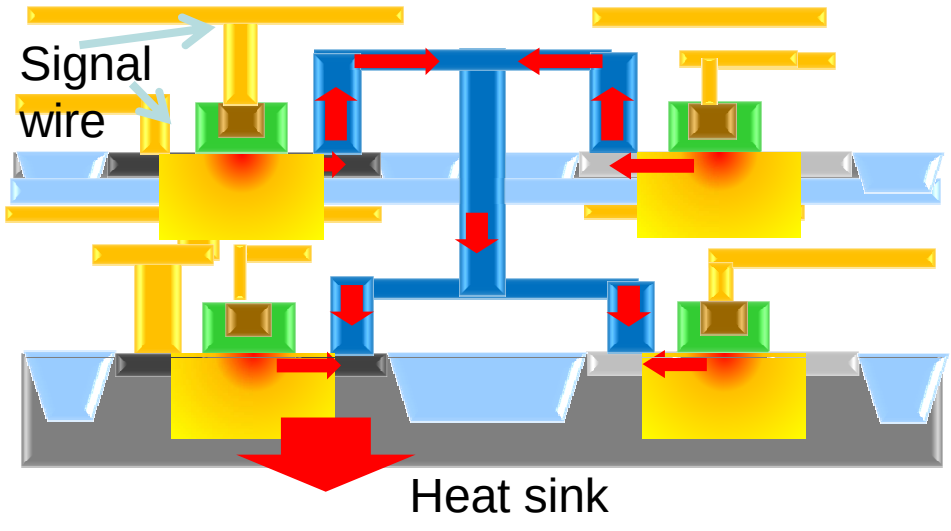
Stanford University, Rambus[+], Monolithic 3D Inc.*

Monolithic 3D™

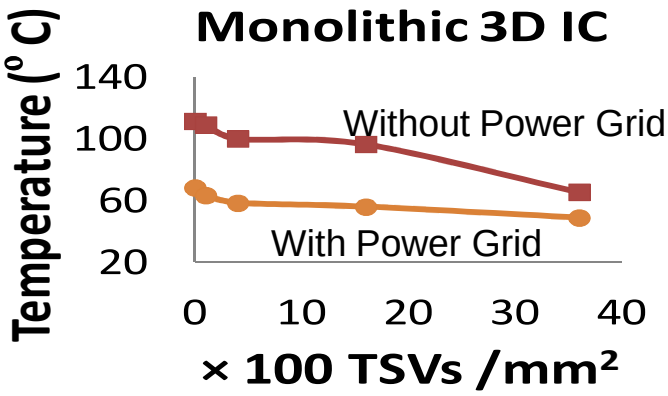
Monolithic 3D Heat Removal Architecture

(Achievable with Monolithic 3D vertical interconnect density)





- Global power grid shared among multiple device layers, local power grid for each device layer
- Local V_{DD} grid architecture shown above
- Optimize all cells in library to have low thermal resistance to V_{DD}/V_{SS} lines (local heat sink)



Monolithic 3D

Patented and Patent Pending Technology

Power Delivery (Vdd, Vss) Network Provide effective Heat Removal Path