




Monolithic 3D Technology: Nu-3D IC



A Semiconductor Industry Game-Changer



Zvi Or-Bach

President & CEO

NuPGA™

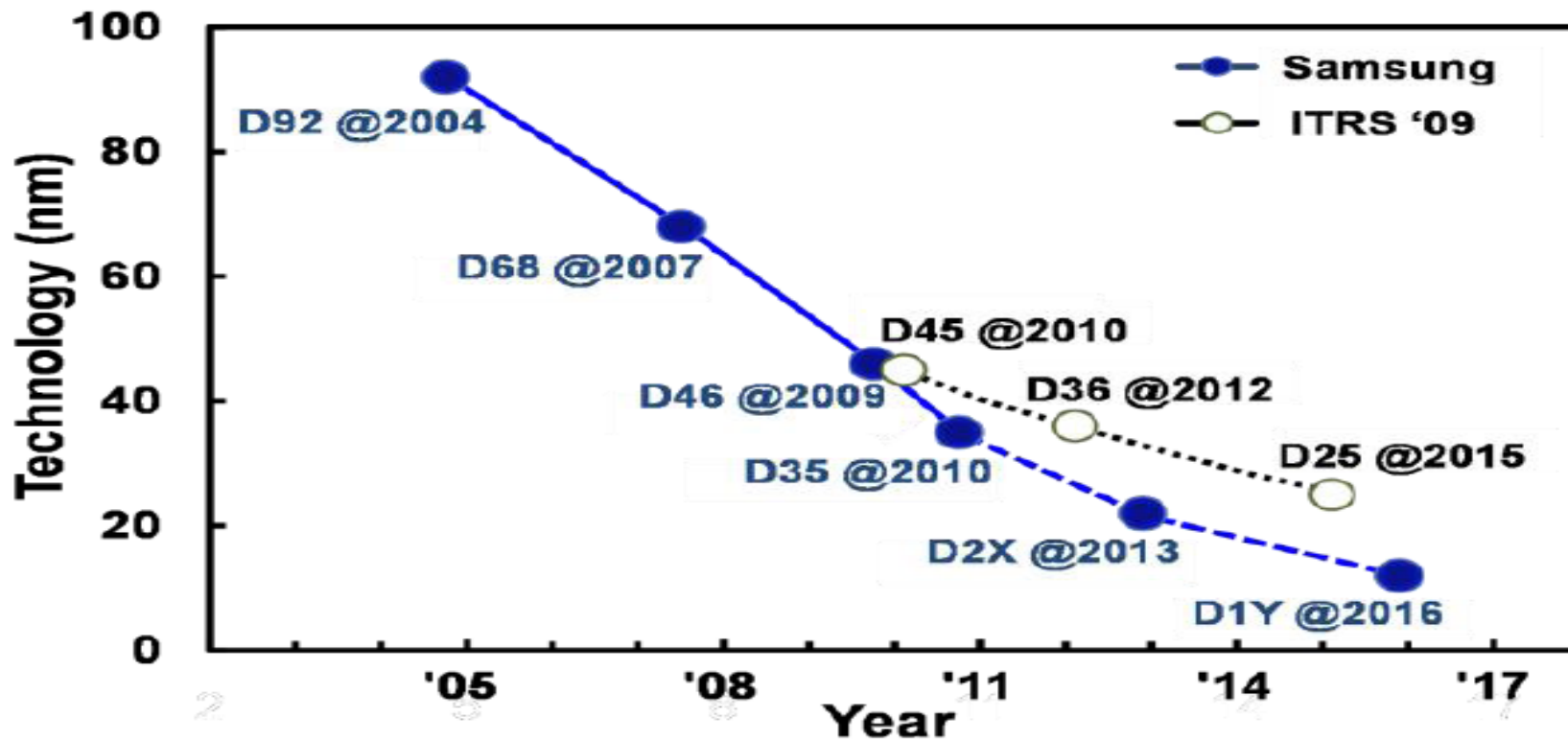
Agenda



- Scaling and Interconnect Challenges
- 3D – TSV vs. Monolithic
- NuPGA Solution
 - Path 1 – RCAT Transistor
 - Path 2 – State of the art MOSFET (Gate-Last)
- Summary

Moore's Law is Still in Effect – But at escalating Costs

“Even with its many challenges there is not a fundamental limitation in scaling down to several nm” * Samsung keynote at IEDM 2010





The Dilemma of the Semiconductor Industry

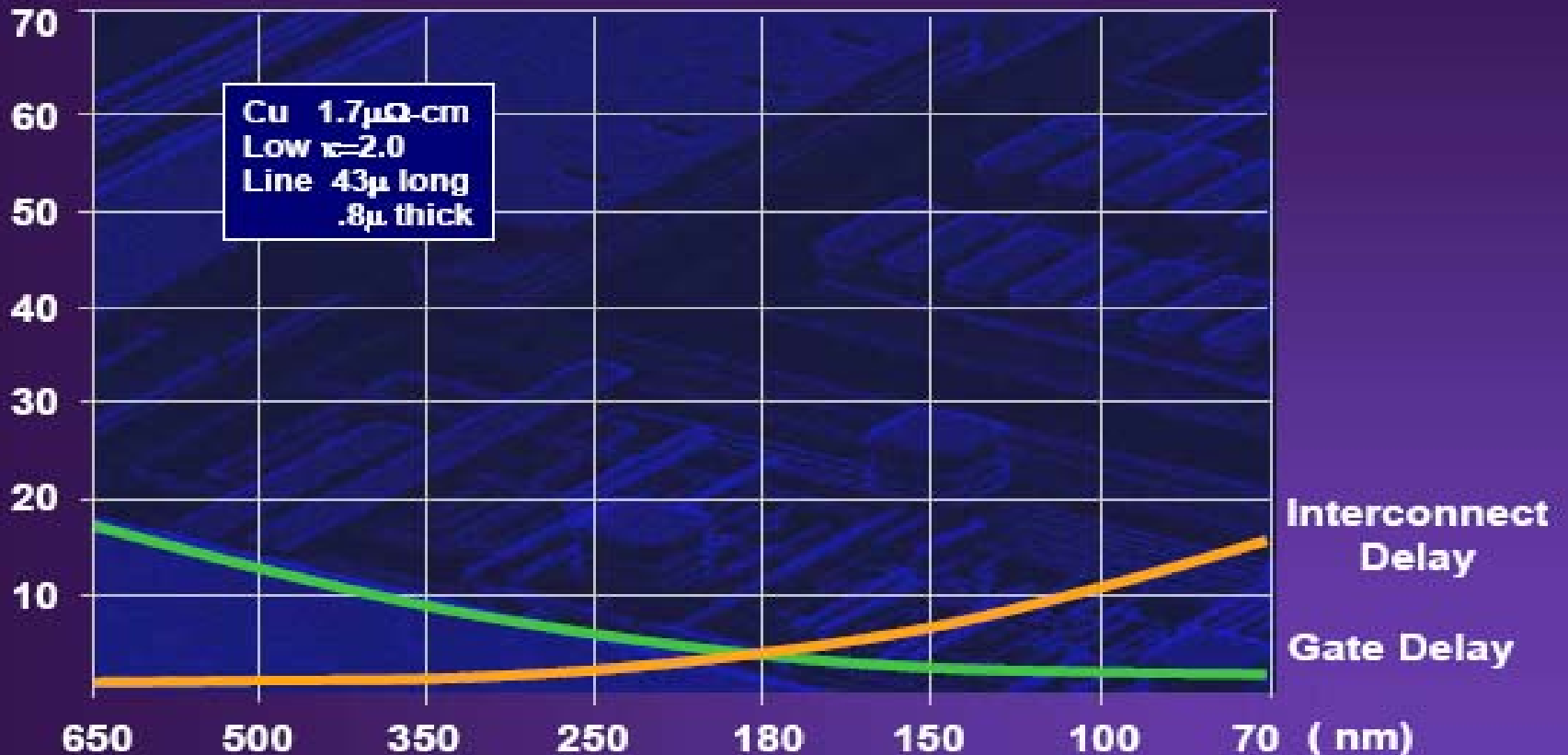
- Chip-makers need to keep pace with technology and focus on design
- ...while chip manufacturing and technology R&D continue to grow in cost and complexity



Transistors no Longer Dominate – Metal Interconnections Took Over

Interconnect Delay Creates the Timing Closure Problem

Delay (ps)



Scaling vs. Interconnect - ITRS 2007 data

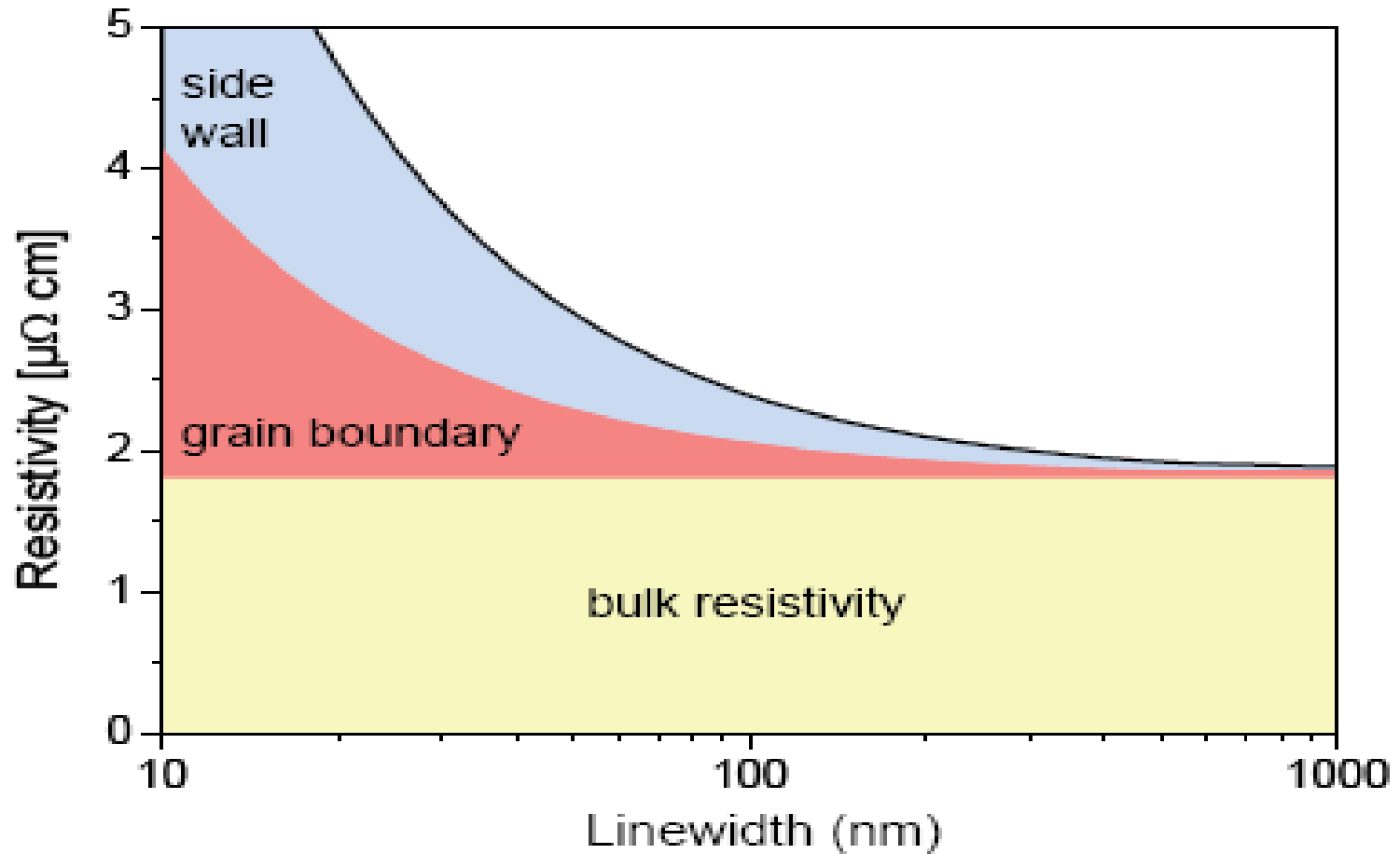


Figure INTCl Cu Resistivity

Scaling vs. Interconnect - ITRS 2007 data

(Fishbach et al., SLIP 2009)

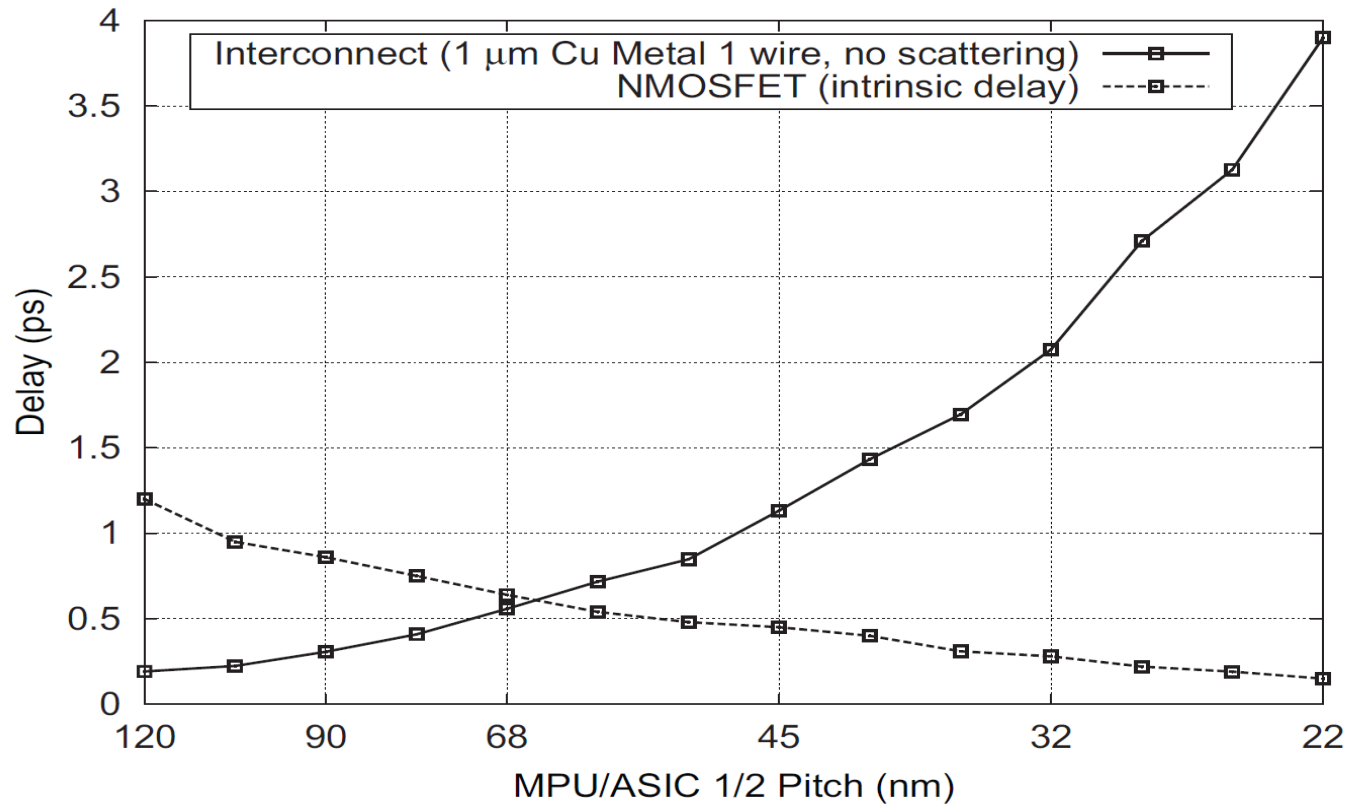


Figure 1: The increasing interconnect delay for various technology nodes according to the ITRS roadmap [1].

The Tyranny of Interconnects



Delay Ratio	@100nm	@35nm
Routing (1mm)	30 ps	100
Transistors	5 ps	1
	$\frac{30}{5} = 6$	
Power Ratio		
Routing	5	30
Transistors	1	1



James Meindl, 2004

(<http://www.informationweek.com/news/showArticle.jhtml?articleID=18902137>)

The Solution - 3D IC

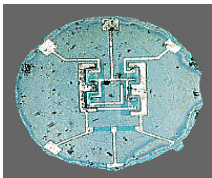
1950s

Too many interconnects to manually solder
→ interconnect problem

Solution: The **(2D)** integrated circuit



Kilby version:
Connections not integrated

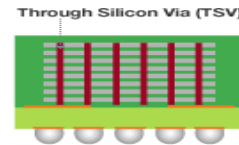


Noyce version
(the monolithic idea):
Connections integrated

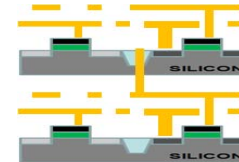
Today

Interconnects dominate performance and power and diminish scaling advantages
→ interconnect problem

Solution: The **3D** integrated circuit

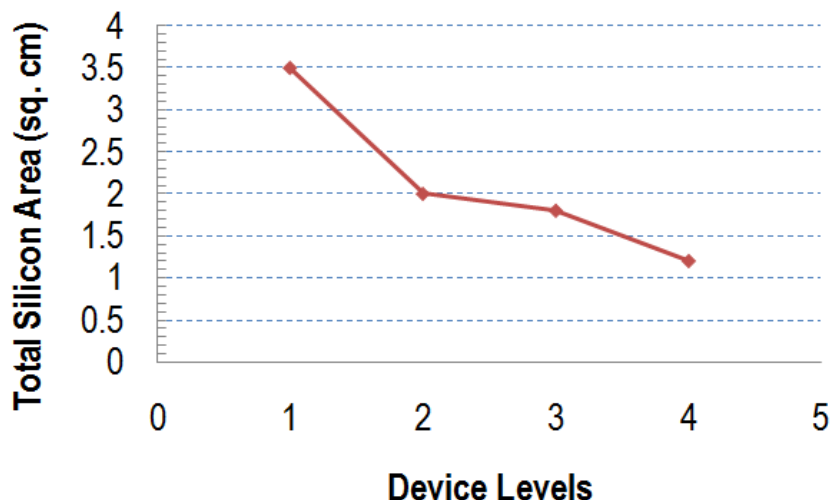


3D with TSV: TSV-3D IC
Connections not integrated



Monolithic 3D: Nu-3D IC
Connections integrated

Monolithic 3D Integration: A Much-Sought After Goal



From J. Davis, J. Meindl, et al., Proc. IEEE, 2001
Frequency = 450MHz, 180nm node, ASIC-like chip

Monolithic 3D Integration → Theoretically analyzed by [Joyner; Meindl], [Souri; Saraswat] and [Rahman; Reif] in late 90s.

Tremendous benefits when vertical connectivity ~ horizontal connectivity.

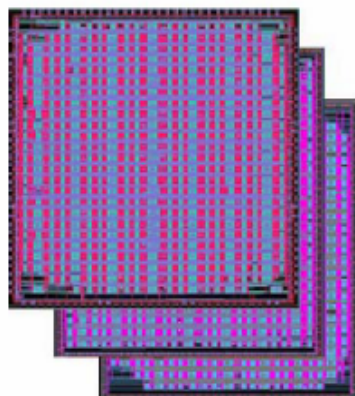
3x reduction in silicon area compared to a 2D implementation, @ 180nm !

Large Systems Benefits from 3D-IC Integration

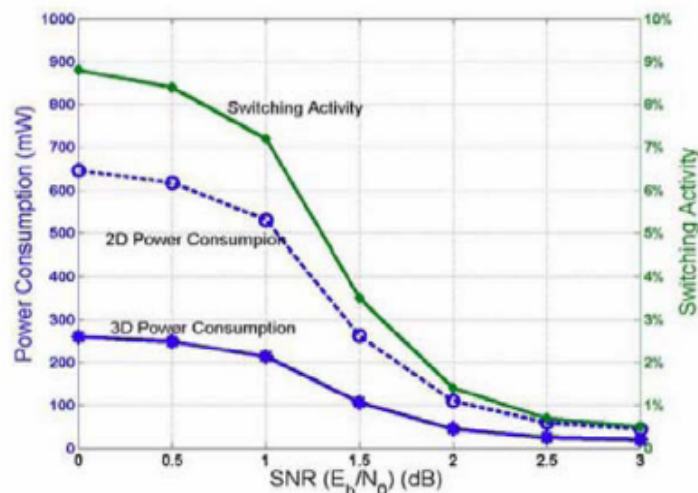
"Implementing a 2-Gbs 1024-bit $\frac{1}{2}$ -rate Low-Density Parity-Check Code Decoder in Three-Dimensional Integrated Circuits"

Lili Zhou, Cherry Wakayama, Robin Panda, Nuttorn Jangkrajarn, Bo Hu, and C.-J. Richard Shi
University of Washington

International Conference on Computer Design, ICCD, Oct. 2007



Final layout view of 3D LDPC structure.



Post-layout power of the LDPC decoder (2D vs 3D).

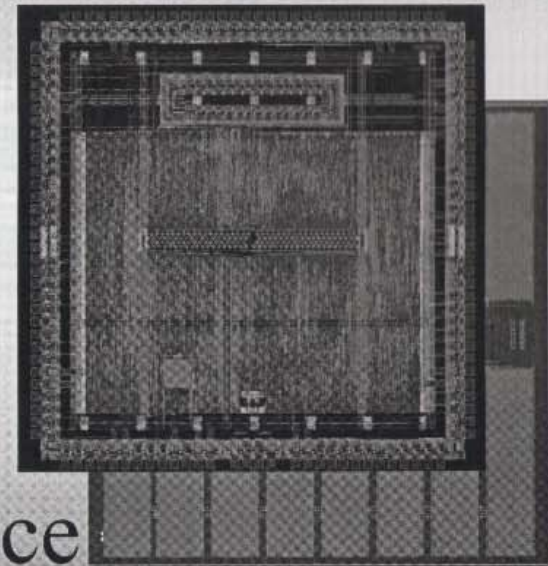
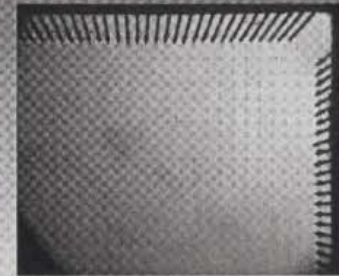
Comparison between 3D and 2D designs

	2D design	3D design
Area (mm*mm)	18.238*15.92 =290.35	(6.4*6.227)*3 = 119.56
Total wire length (m)	182.42	22.39+22.57+22.46 =67.42
Max WL before buffer insertion (mm)	13.82	8.68
Max WL after buffer insertion (mm)	4	4
Buffer used	32900	24636
Clock skew (ns)	2.33	1
Power dissipation (mw)	646.2	260.2

Performance Factor (Area * Timing * Power) = 14

An Illustration: CPU/Memory Stack

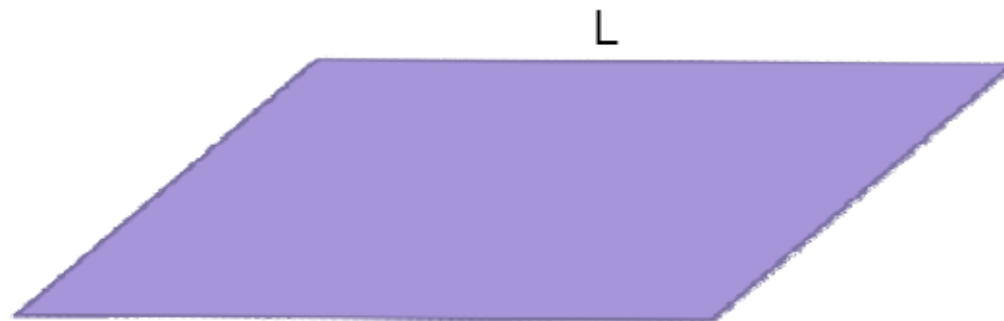
- R8051 CPU
 - 80MHz operation; 140MHz Lab test (VDD High)
 - 220MHz Memory interface
 - IEEE 754 Floating point coprocessor
 - 32 bit Integer coprocessor
 - 2 UARTs, Int. Cont., 3 Timers, ...
 - Crypto functions
 - 128KBytes/layer main memory
-
- 5X performance
 - 1/10th Power



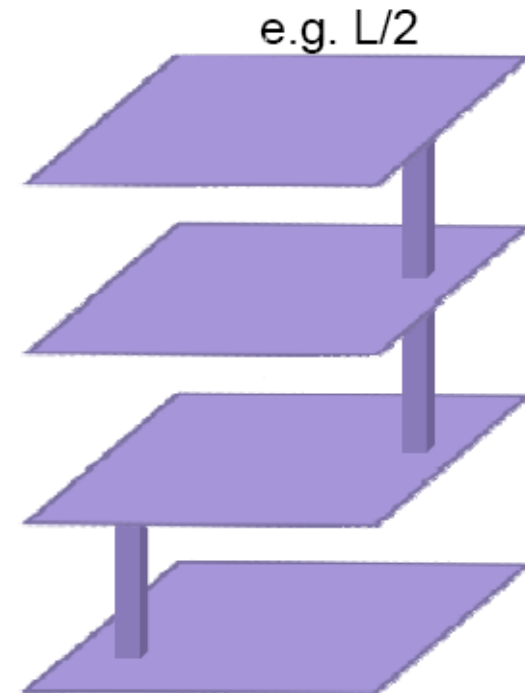
“3D” IC Integration Looks Great...

Technology Node n^{th} 2D \cong Technology Node $(n-2)^{\text{th}}$ 3D

- Much easier D and A&M/S integration
- Smaller footprint, higher bandwidth
- Shorter global interconnect
 - 3 tier \rightarrow -33%, 4 tier \rightarrow -50%
- Better timing and lower power



Silicon area = L^2 , Footprint = L^2
Corner to corner distance = $2L$



Silicon area = L^2 , Footprint = $L/4$
Corner to corner distance = $L + \epsilon$

Monolithic 3D vs. TSV

(1:10,000 vertical connectivity ratio)

	TSV TSV-3D IC	Monolithic: Nu-3D IC
Layer Thickness	~50 μ	50-100nm
Via Diameter	~5 μ	~50nm
Via Pitch	~10 μ	~100nm
Limiting Factors	Wafer handling (~5 μ) Aspect ratio (<10:1)	Lithography=> Will keep scaling
Wafer (Die) to Wafer Alignment	~1 μ	Layer to Layer Alignment => Will keep scaling

'Monolithic' 3D




*Is there some way for us to get monolithic 3D,
i.e., vertical connectivity ~ horizontal connectivity?*

YES

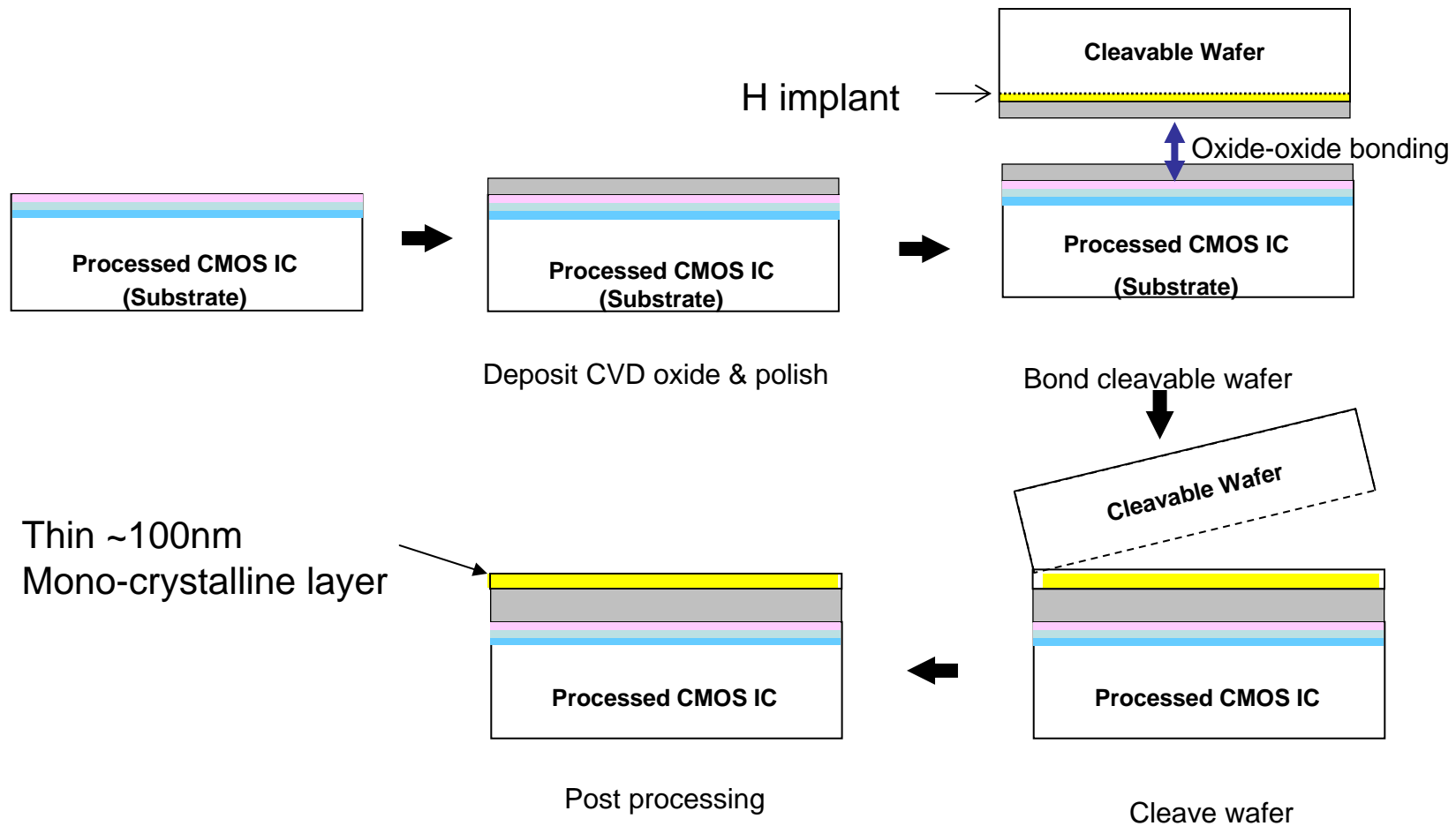
This can be done with NuPGA™ Technology: Nu-3D IC

The Monolithic 3D Challenge



- A process on top of copper interconnect should not exceed 400°C
 - How to bring mono-crystallized silicon on top below 400°C
 - How to fabricate advanced **transistors** below 400°C
- Misalignment of pre-processed wafer to wafer bonding step is $\sim 1\mu$
 - How to achieve 100nm or better connection pitch
 - How to fabricate thin enough layer for inter-layer vias of $\sim 50\text{nm}$

Adding a Thin Crystallized Silicon Layer On Top of the Base IC



Path 1: Define Transistors After doing Layer Transfer



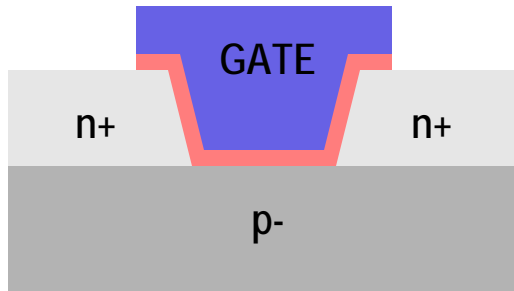
Solves: Alignment at lithography resolution
Challenge: Stay under 400°C

Uses a novel combination of four ideas

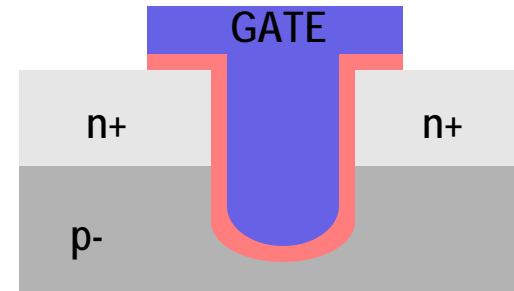
- Pre-implantation of the full wafer area
- High temperature activation of the full wafer area
- Low temperature face-down Layer Transfer with e.g., "Smart-Cut"
- Transistor definition only by 'cold' processes - Etch and Deposition

Recessed Channel Transistors

Recessed channel transistors → a family of transistors that can be defined @ <400°C



V-groove recessed channel transistor:
Used in the TFT industry today

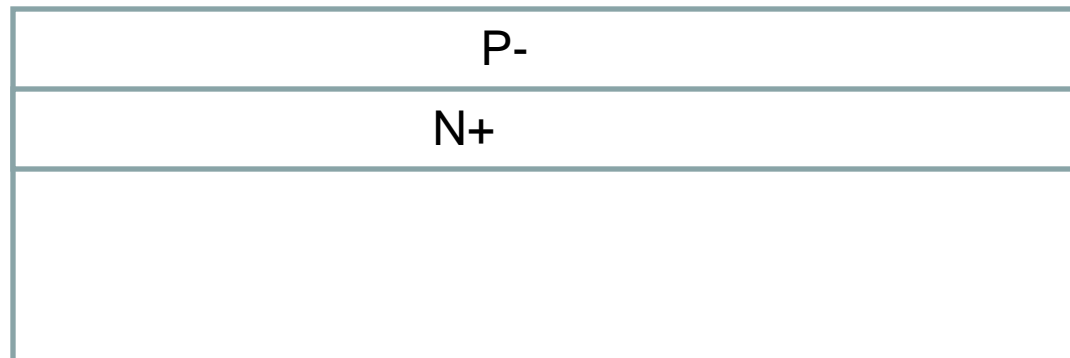


RCAT recessed channel transistor:

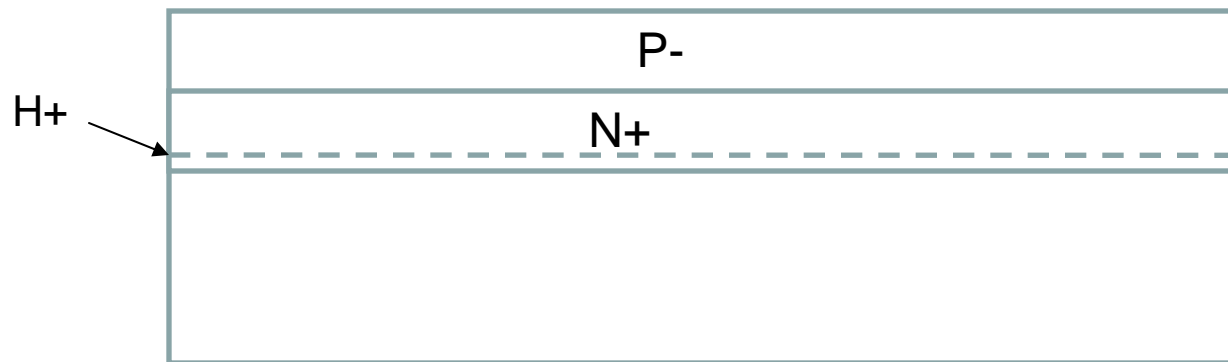
- Used in DRAM production @ 90nm, 60nm, 50nm nodes
- More capacitance, but less leakage, same drive current according to Samsung (for DRAM transistors)

J. Kim, et al. Samsung, VLSI 2003
ITRS

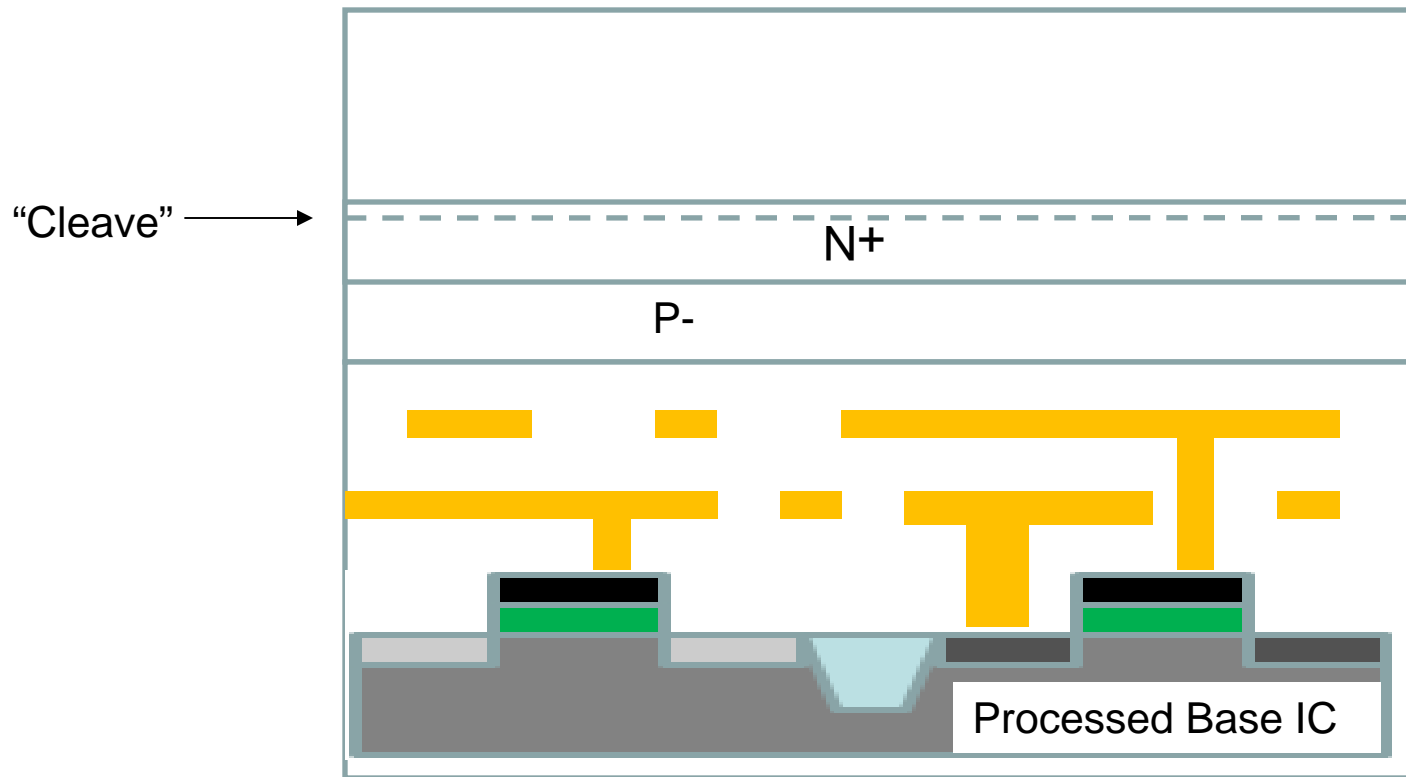
step 1 - Implant and Activate at High Temp. (~ 900°C) Before Layer Transfer



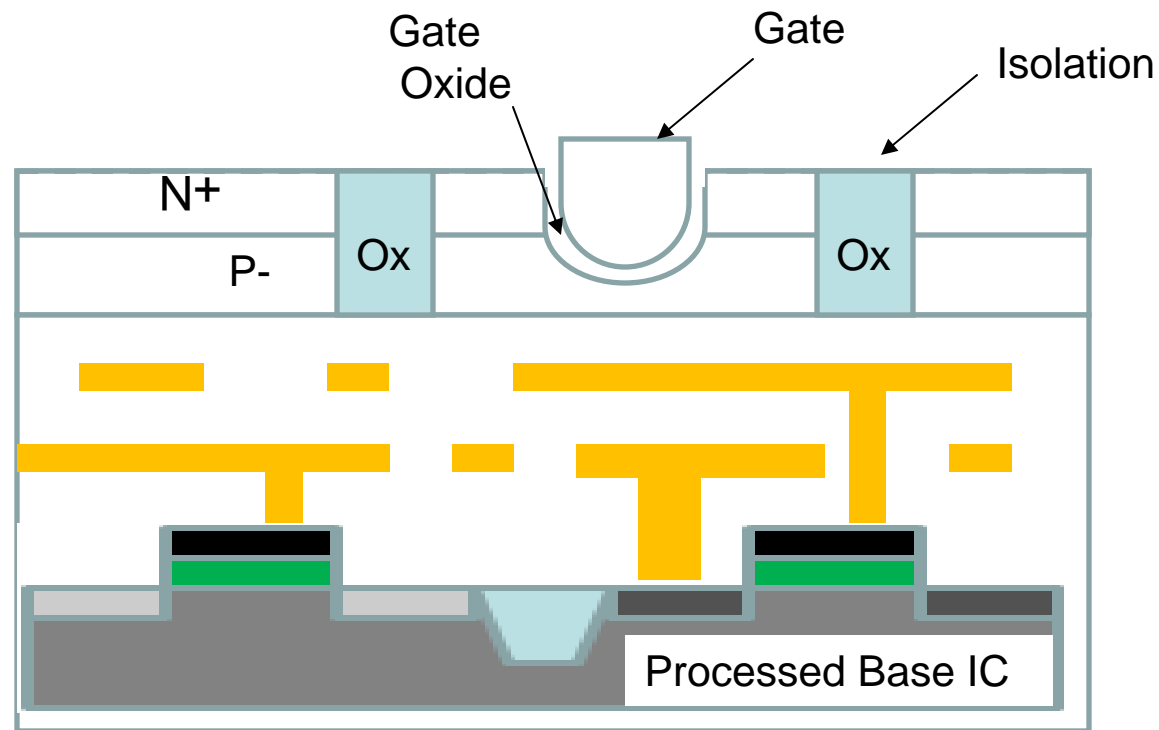
step 2 - Implant H+ for the Ion-Cut



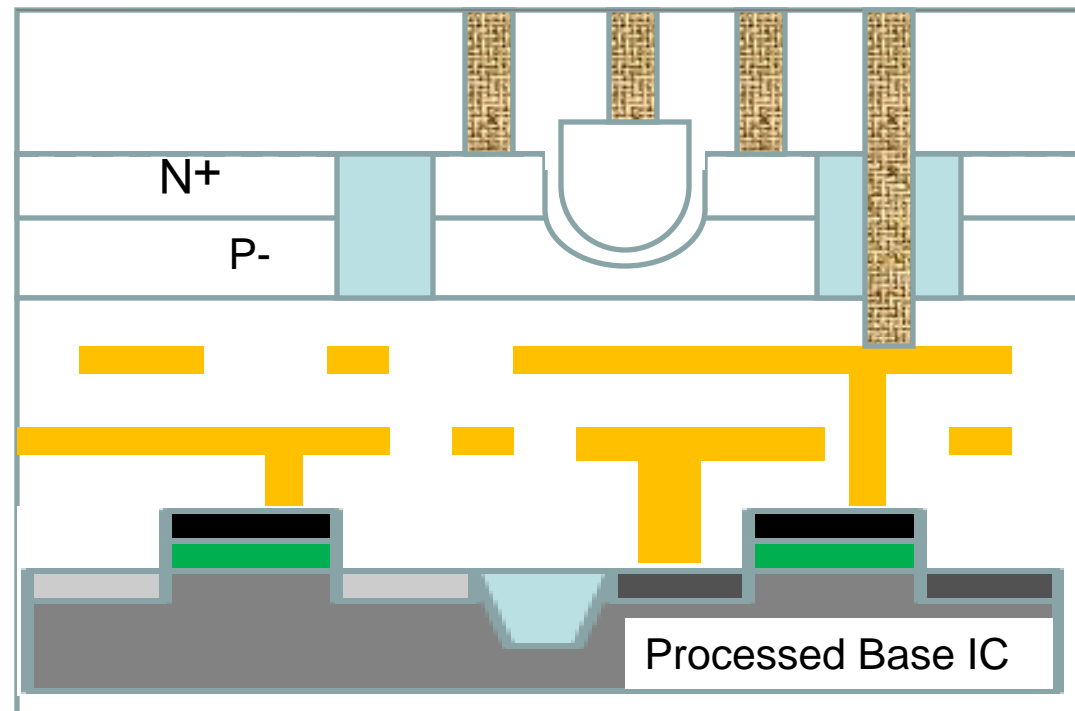
step 3 - Bond and Cleave



step 4 - Etch and Form Isolation and RCAT Gate



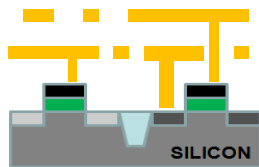
step 5 - Make Contact to the RCAT



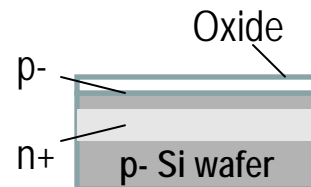
Summary: Sub-400°C Processed Transistors

Recessed Channel Transistors

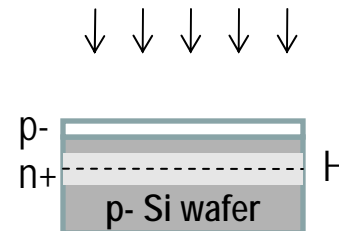
(1) Construct bottom layer with transistors and wires.



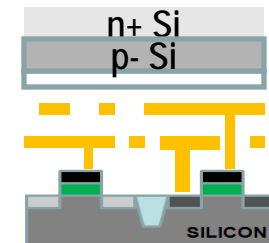
(2) Start processing top layer with a p- Si wafer. Form n+, p-. RTA. Oxidize top surface.



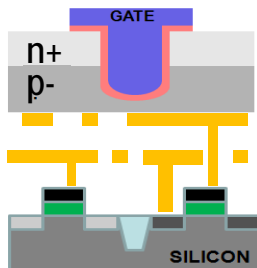
(3) Implant hydrogen into wafer at required depth



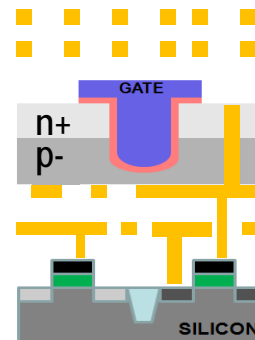
(4) Flip and bond atop bottom layer, cleave



(5) Etch, form gate with ALD, etc. Pattern with features aligned to bottom layer.



(6) Complete transistors, wires, through-vias on top layer aligned to bottom layer

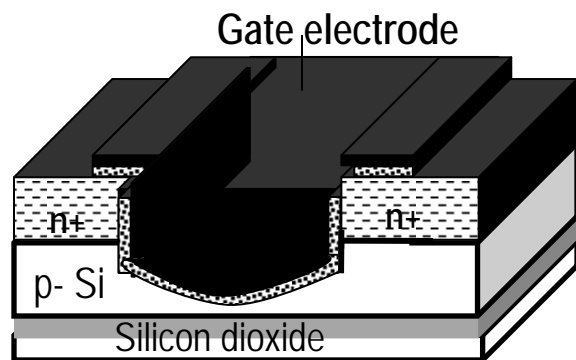


Note:
All steps after Next
Layer attached to
Previous Layer are
@ < 400°C!

25

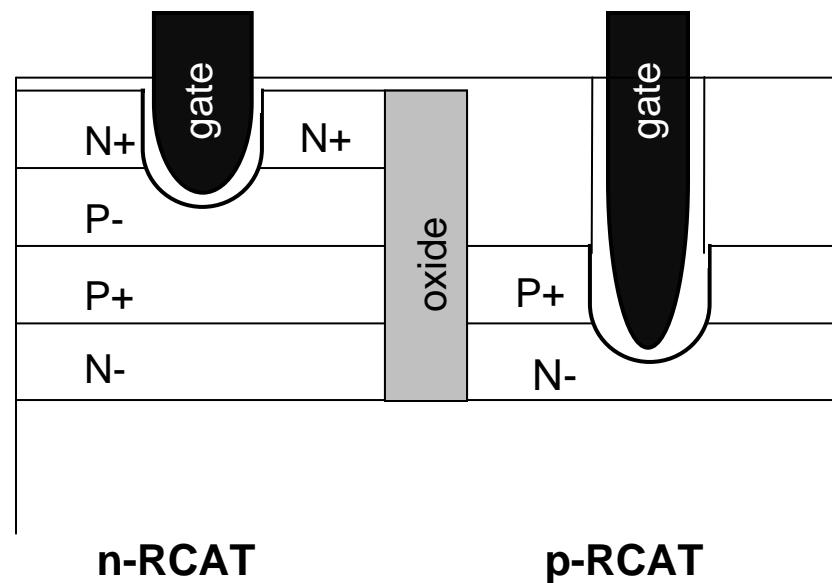
Full CMOS On One Layer is Also Possible

Build transistor layers above wiring layers monolithically
@ <400°C



Recessed Channel Transistors.

- Sub-400°C stacking possible.
- Used in DRAM and TFT applications today.



Path 2: Innovative Alignment with Standard Transistors



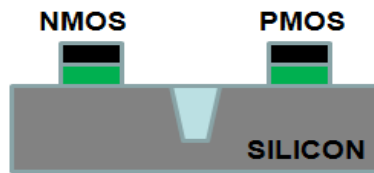
Uses a novel combination of four ideas

- Gate-Last HKMG process and proper sequence of Layer Transfer
- Low temperature face-up Layer Transfer
- Repeating layouts
- Innovative alignment

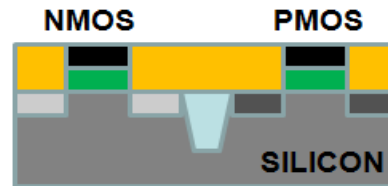
A Gate-Last Process with Repetitive Pattern



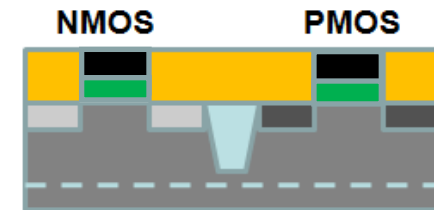
(1) Make dummy gates with oxide, poly-Si



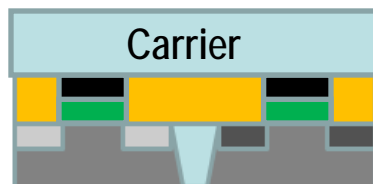
(2) S/D implants, SiGe S/D, high-T anneal, salicide/contact etch stop, dep./CMP ILD



(3) Implant H for cleaving



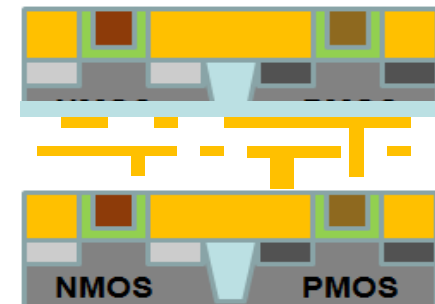
(4) Temporary bond to carrier wafer, cleave, grind to STI



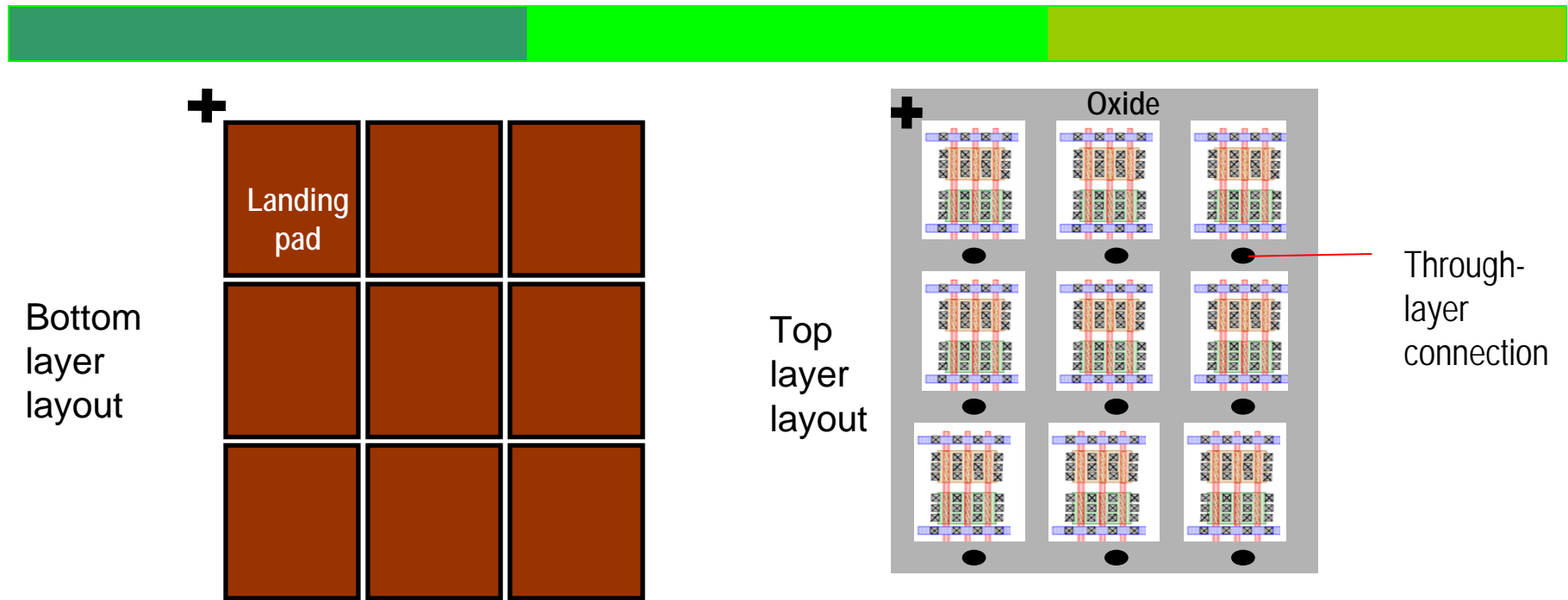
(5) Low-T oxide dep., bond to bottom layer, remove carrier



(6) Etch dummy gates, gate dielectric and electrode deposition, CMP, BEOL

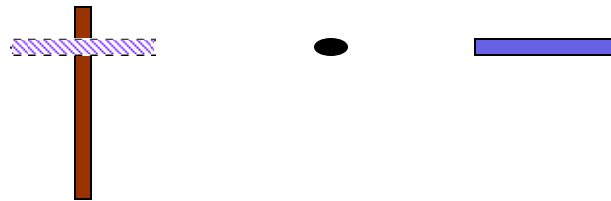
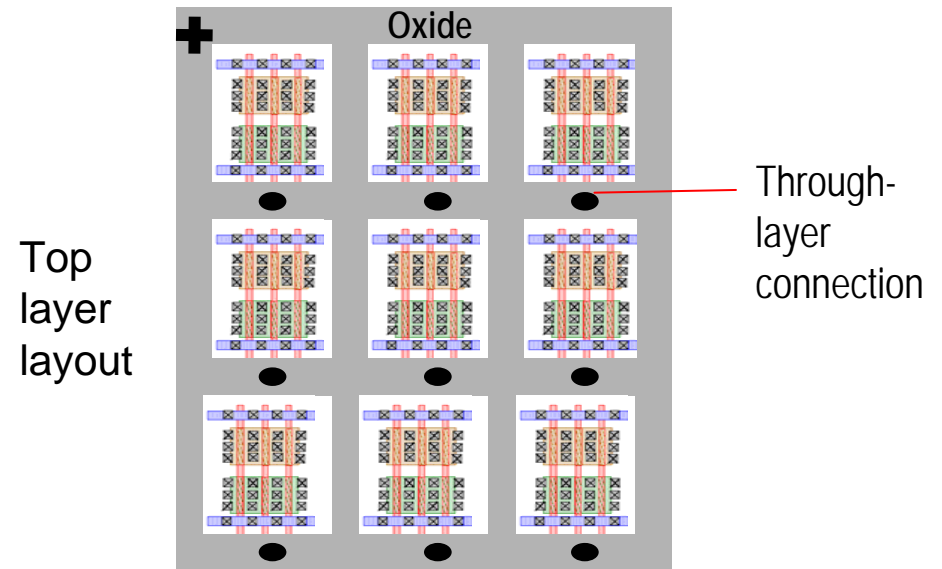
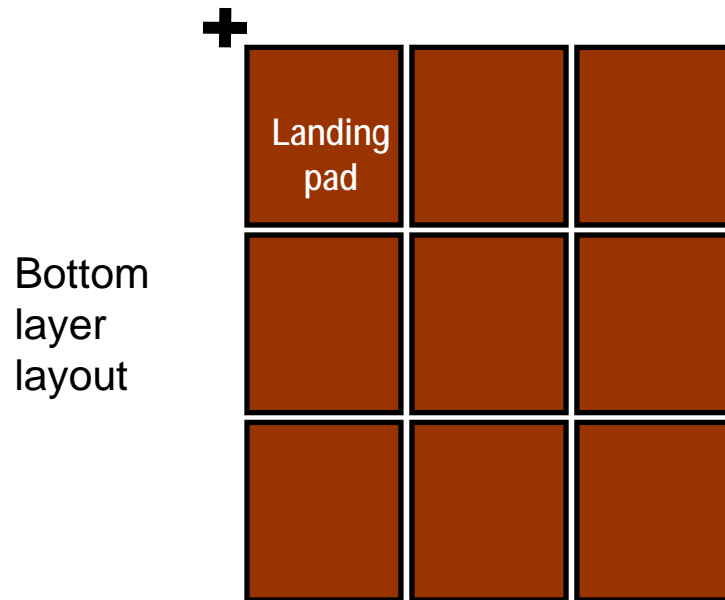


Repeating Layouts for Transistors

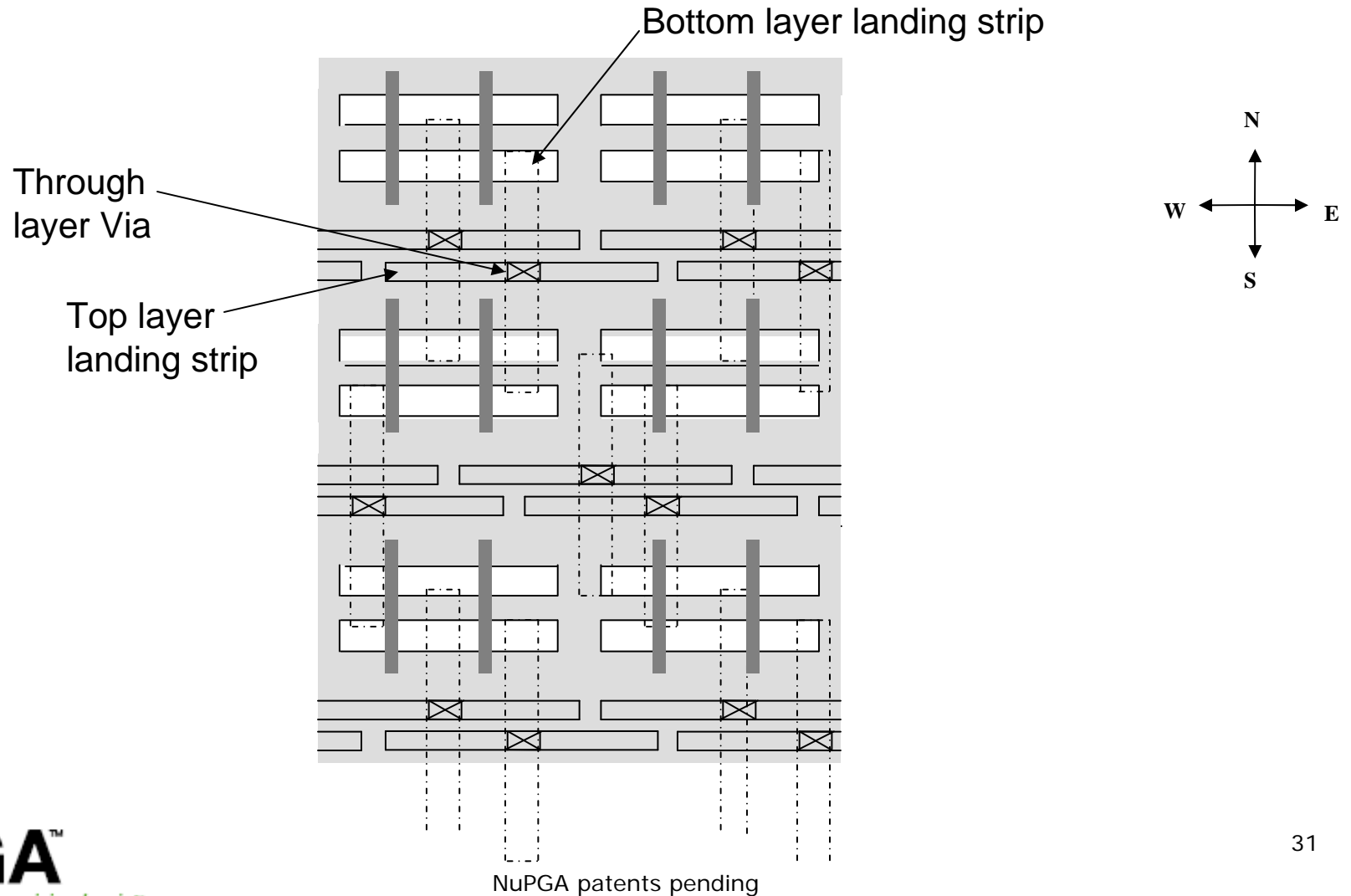


- Even if misalignment occurs during bonding → regular layouts allow correct connections.
- Above representation simplistic, high area penalty.
- *High density of connections between 3D stacked layers Using standard transistors*

Landing 'Pad' for Repeating Layouts



Repeating Cell with Two Orthogonal Strips for low penalty Vertical Connectivity



To Summarize the Two Approaches



Monolithic 3D with high density vertical connections is possible:

- Novel, manufacturing proven transistors that allow high-temperature processing before definition (e.g. Recessed Channel Transistors)
- Standard cutting-edge transistors and repeating layouts
- Other variations (not shown here)

Summary

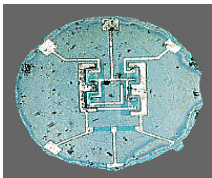
1950s

Too many interconnects to manually solder
→ interconnect problem

Solution: The **(2D)** integrated circuit



Kilby version:
Connections not integrated

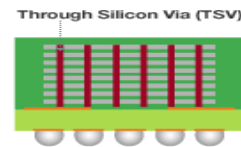


Noyce version
(the monolithic idea):
Connections integrated

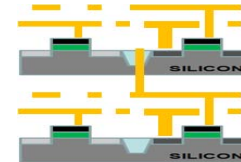
Today

Interconnects dominate performance and power and diminish scaling advantages
→ interconnect problem

Solution: The **3D** integrated circuit



3D with TSV: TSV-3D IC
Connections not integrated



Monolithic 3D: Nu-3D IC
Connections integrated

Back Up



MOSFET vs RCAT - IdVg

IdVg at Vd=1V

