- **DTCO/STCO**
- **Technology driven**
- **Design driven**
- **Application driven**

Status Quo and future of IC technologies

IC/semiconductors(Logic/Memory/Analog/Micro +OSD)

3D x3D x3D

Chee Wee Liu (劉致為)

Distinguished/Chair Professor / IEEE Fellow,
National Taiwan University

cliu@ntu.edu.tw

http://nanosioe.ee.ntu.edu.tw

High mobility/high K/FinFET

PPACR (performance/power/area/ cost/reliability)

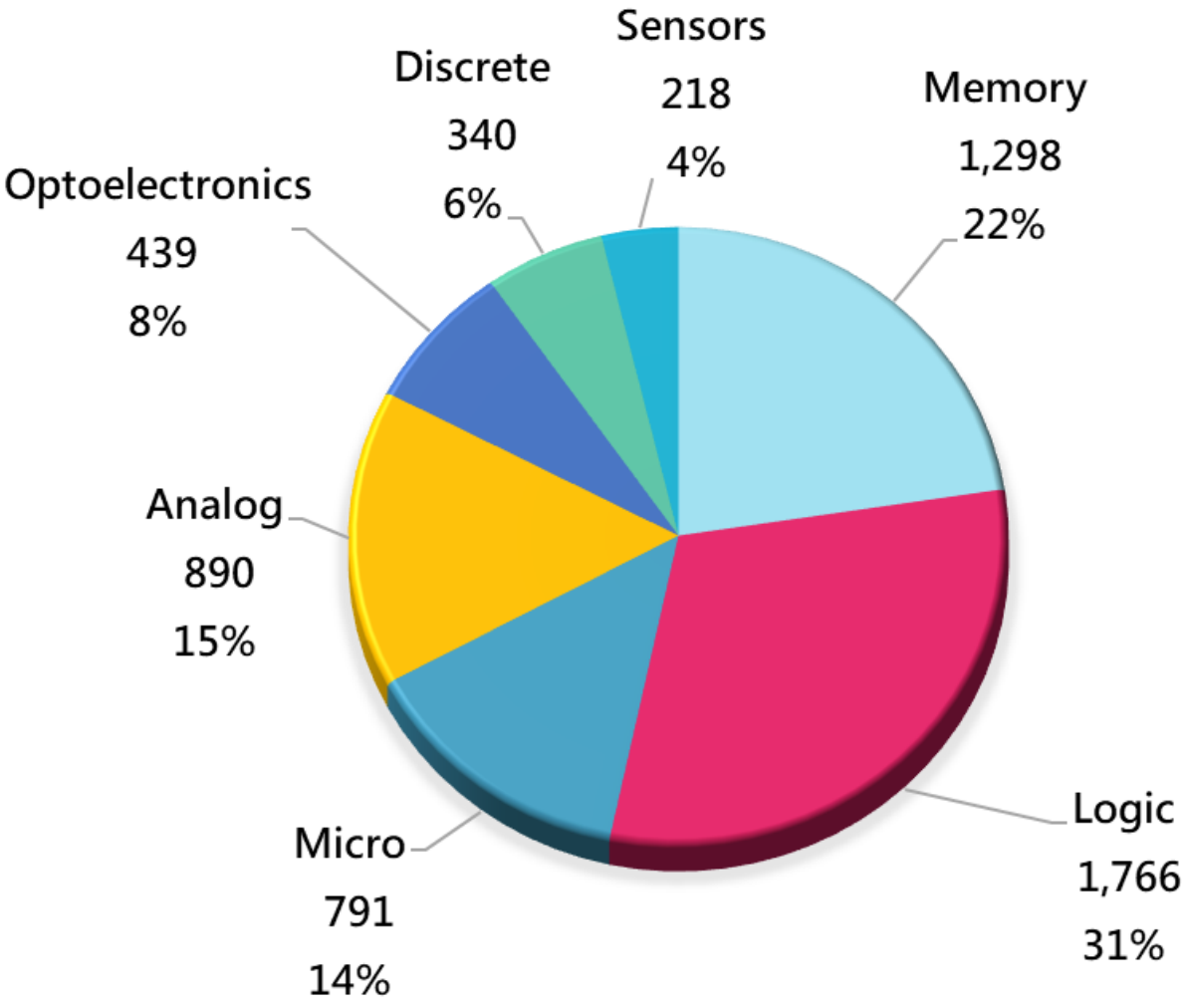
Google Scholar Page

Citation :8305/ H index:41

i10 index:201

2022 WW Semiconductor Market

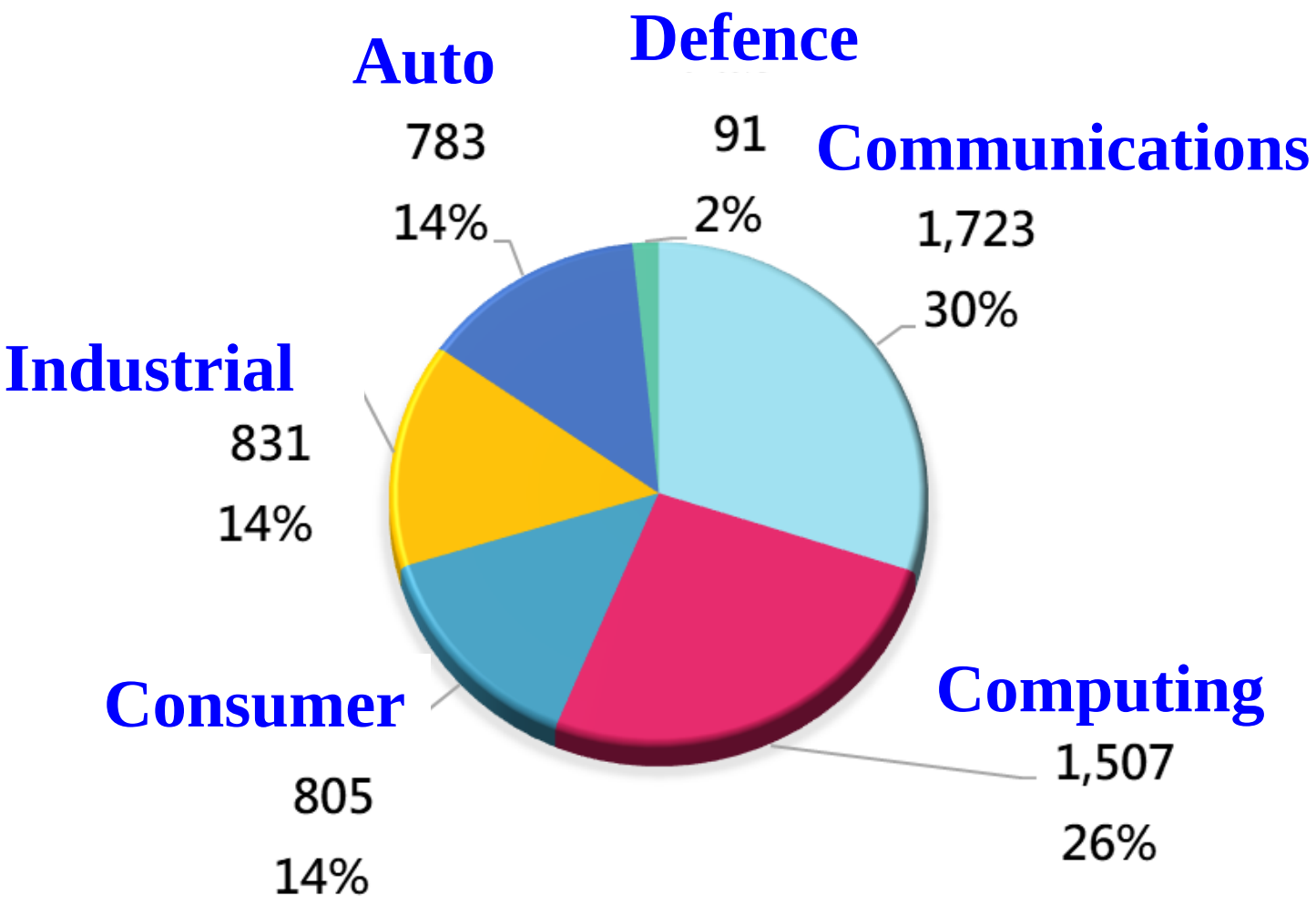
WW: 574.1B US$



Source: WSTS、ITRI/ISTI (2023/04)

2022 WW Semiconductor Market – Applications

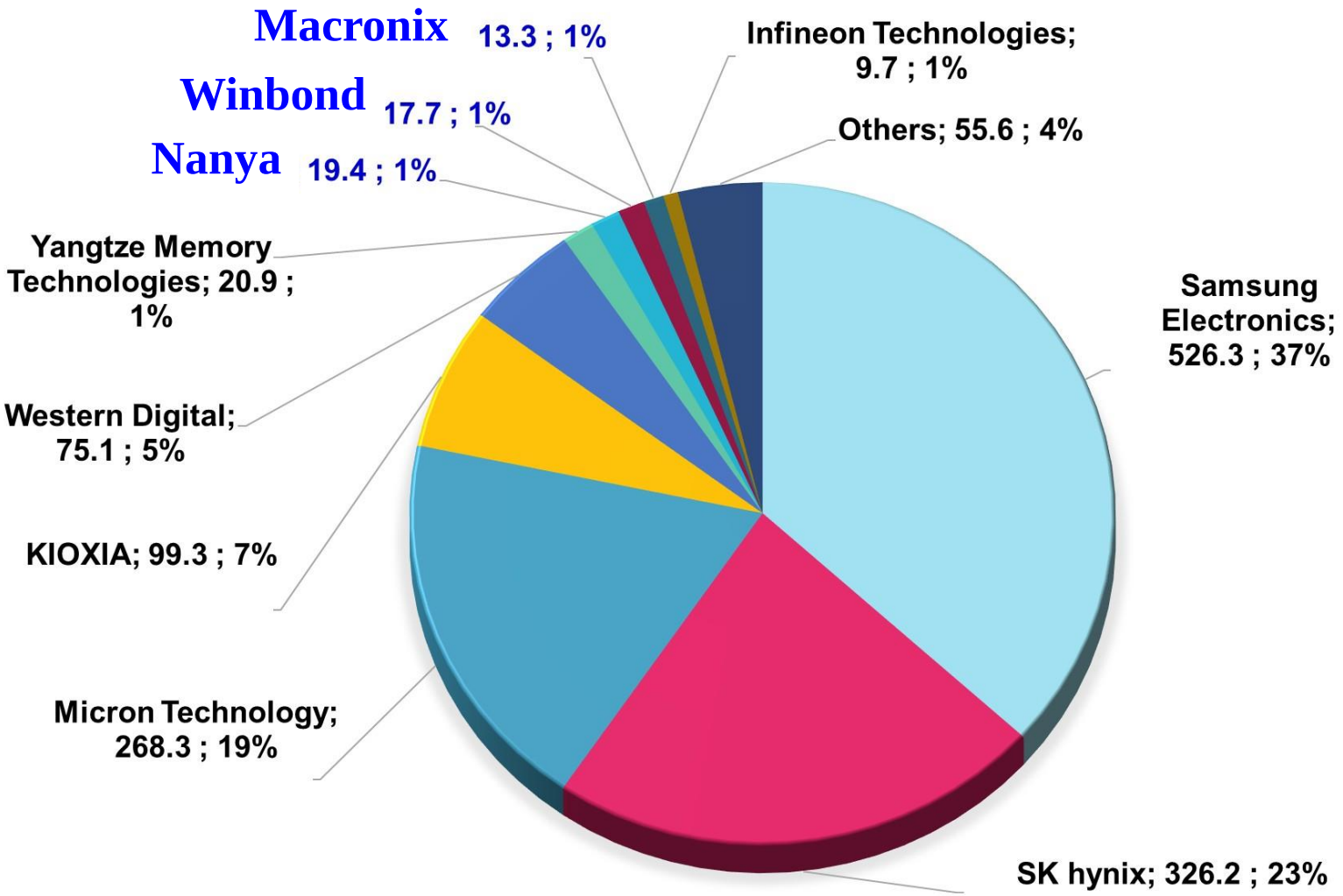
WW: 574.1B US$



Source: WSTS、ITRI/ISTI (2023/04)

2022 WW Market Share: Memory (e)

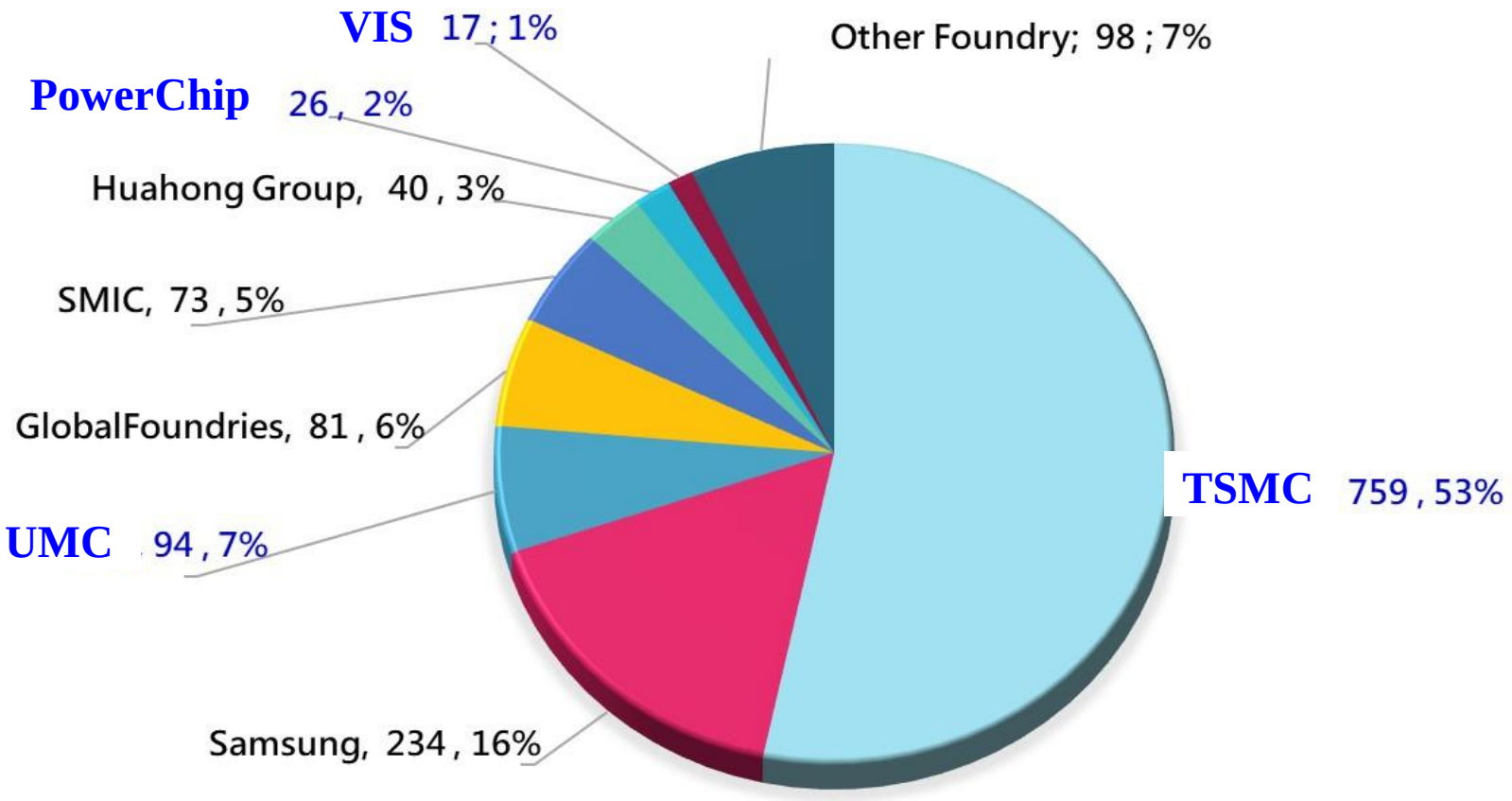
WW Market: 143.2B US$; TW: 3.9%



Source: Gartner、ITRI/ISTI (2023/04)

2022 WW Market Share: Foundry (e)

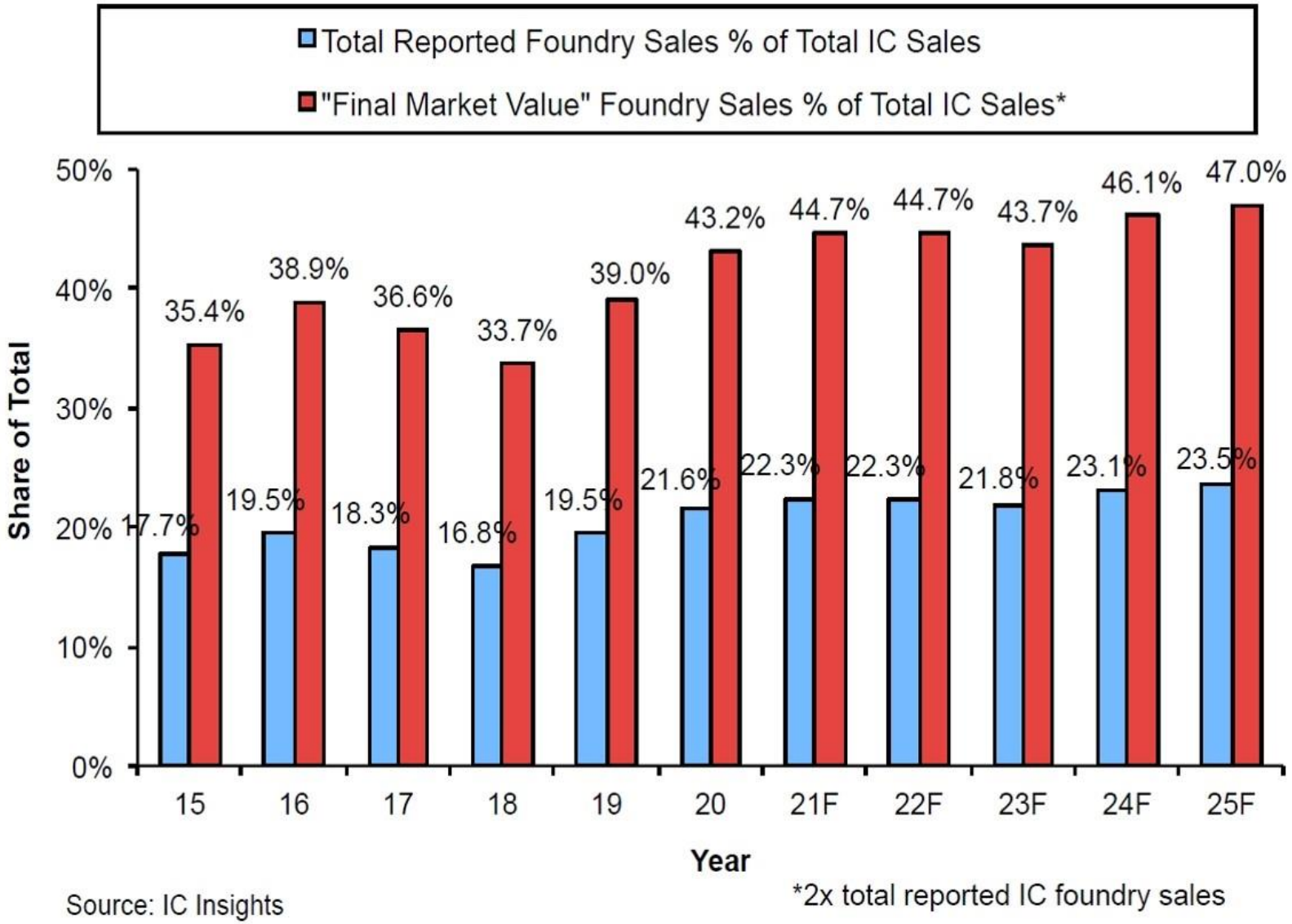
WW: 142.1B US$; TW: 63.4%



Taiwan manufactured 31% of semiconductor chips (2022)!

Source: TechInsights、ITRI/ISTI (2023/04)

Foundry Sales as a Percent of Total IC Sales



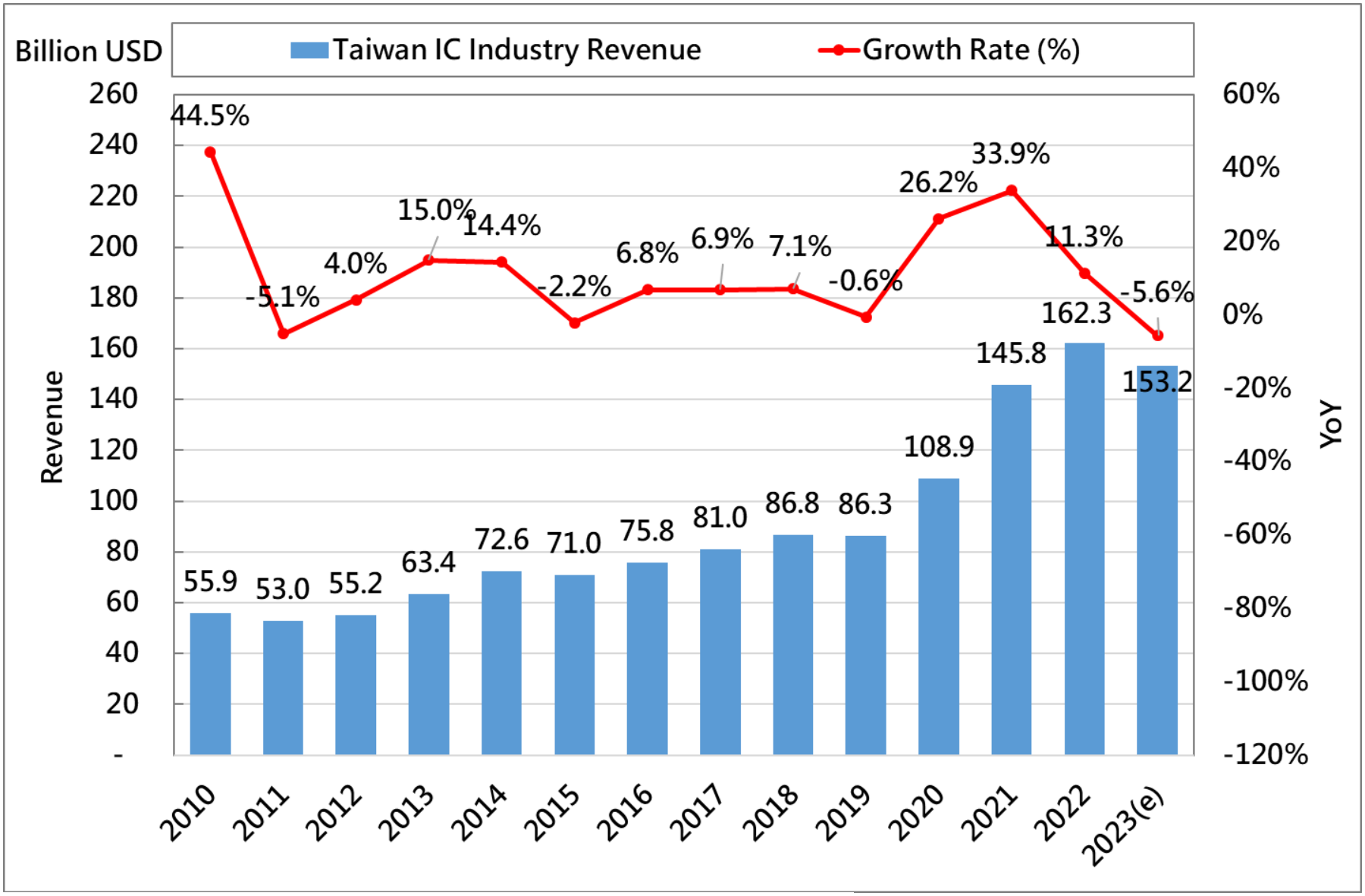
Source: IC Insights (2021/03)

2022 Taiwan Semiconductor Industry

- **Fabless: 41.3B US$ (25.5%)**
- **Manufacturing: 98.0B US$ (60.4%)**
  - **Foundry: 55.5%**
  - **Memory: 4.9%**
- **Pkg & Test: 22.9B US$ (14.1%)**
  - **Pkg: 15.6B US$ (9.6%)**
  - **Test:7.3B US$ (4.5%)**

Source: ITRI/ISTI (2023/03)

Taiwan IC Industry Sales Revenue



Source: ITRI/ISTI (2023/03)

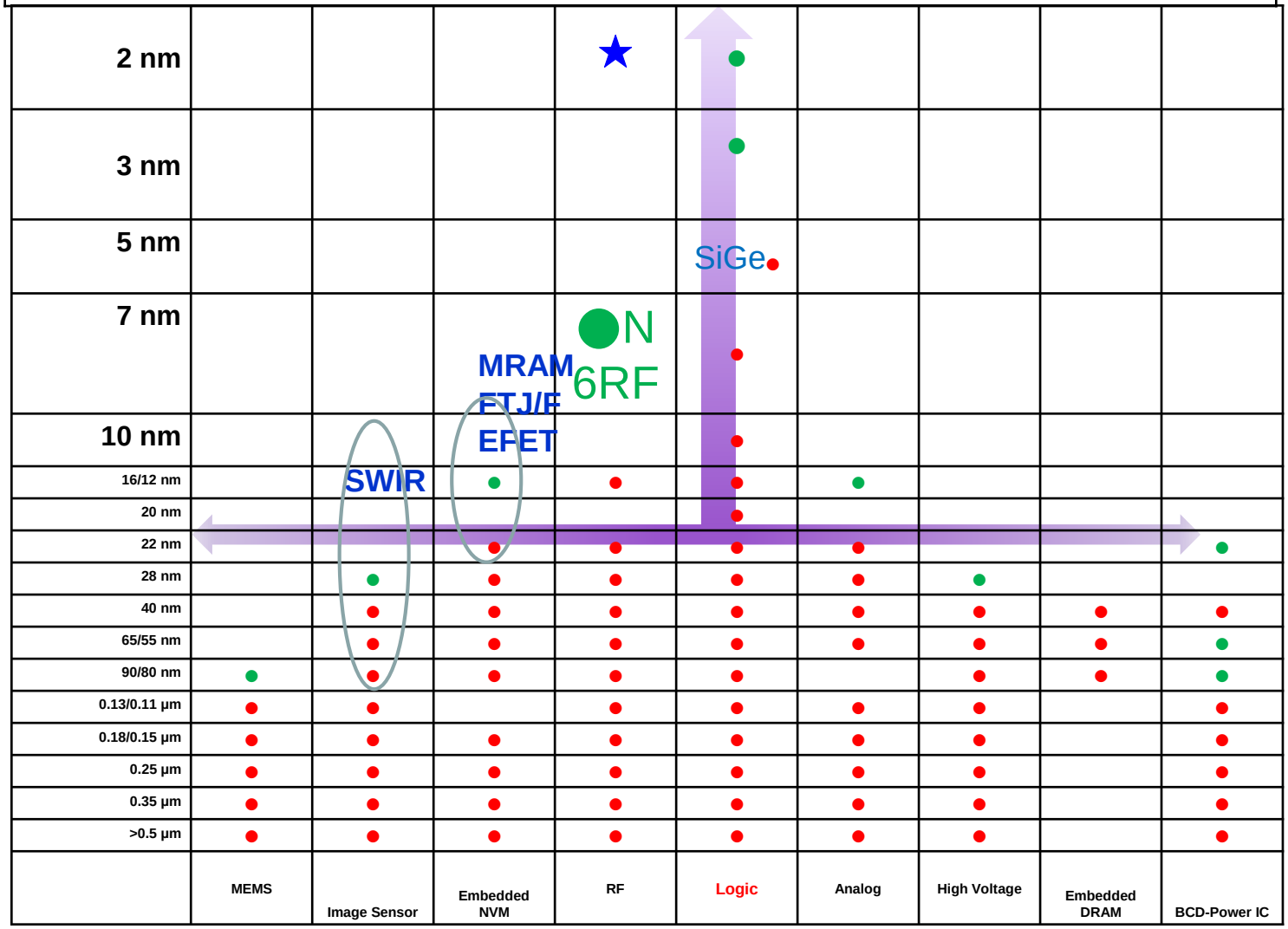
Moving Forward: Advanced Technologies

- **Manufacturing**
  - **Logic: 2nm and beyond**
  - **Specialty: high speed/freq、high/low power、III-V、…**
  - **Heterogeneous integration、3D、chiplet**
- **Design (incl. EDA、Algorithms、S/W)**
  - **High performance computing: cloud、AI、VR/AR、autonomous driving、...**
  - **Wireless computing (5G、6G)**
  - **High efficiency power management**
  - **High performance/precision data conversion**
  - **…**

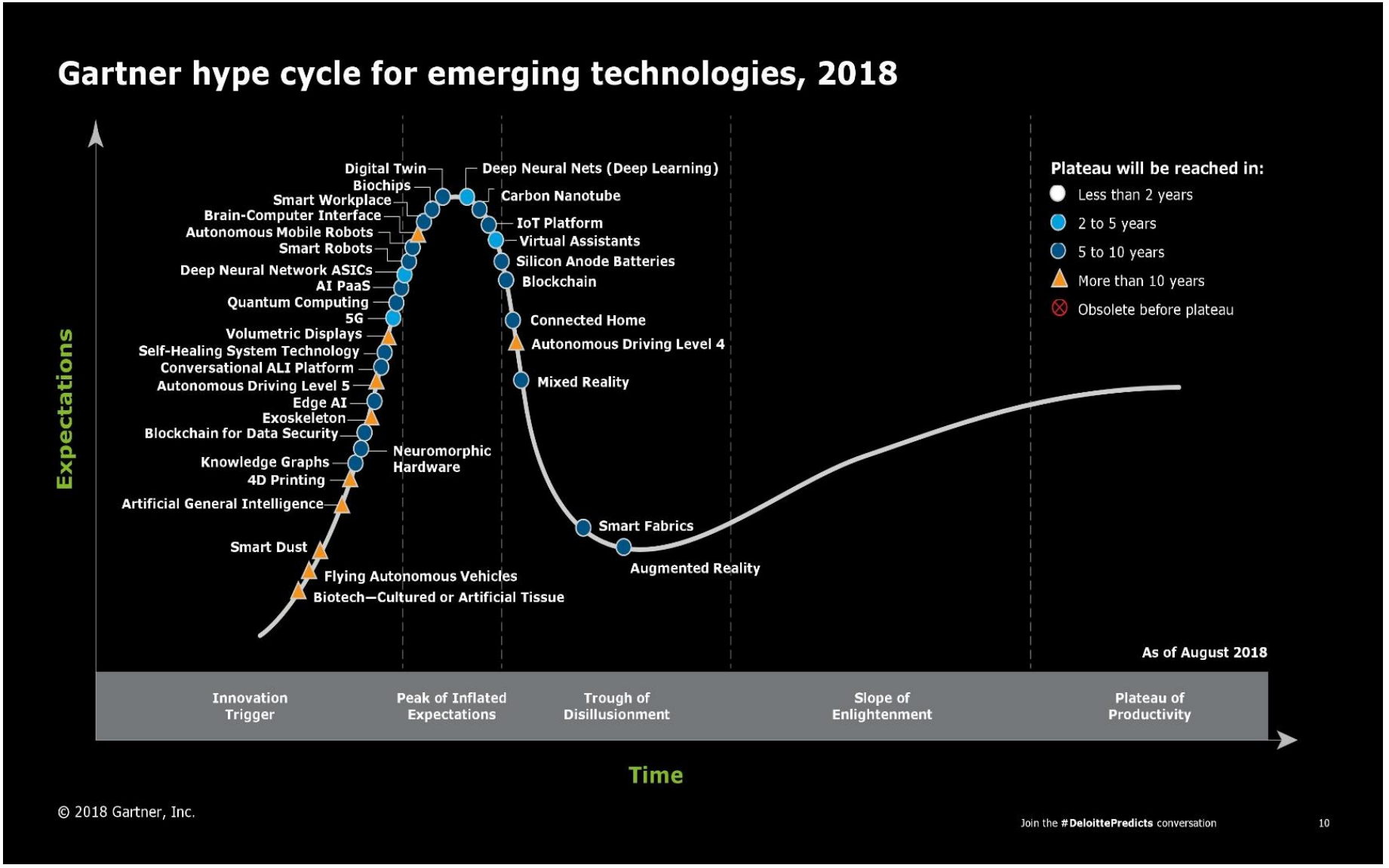
Tsmc Technology Portfolio

● → Available
● → In Development
★ → NTU Research

★1.4 nm & beyond **highly stacked nanosheet/ ultrathin body/treeFETs/CFET/GeSi/GeSn**



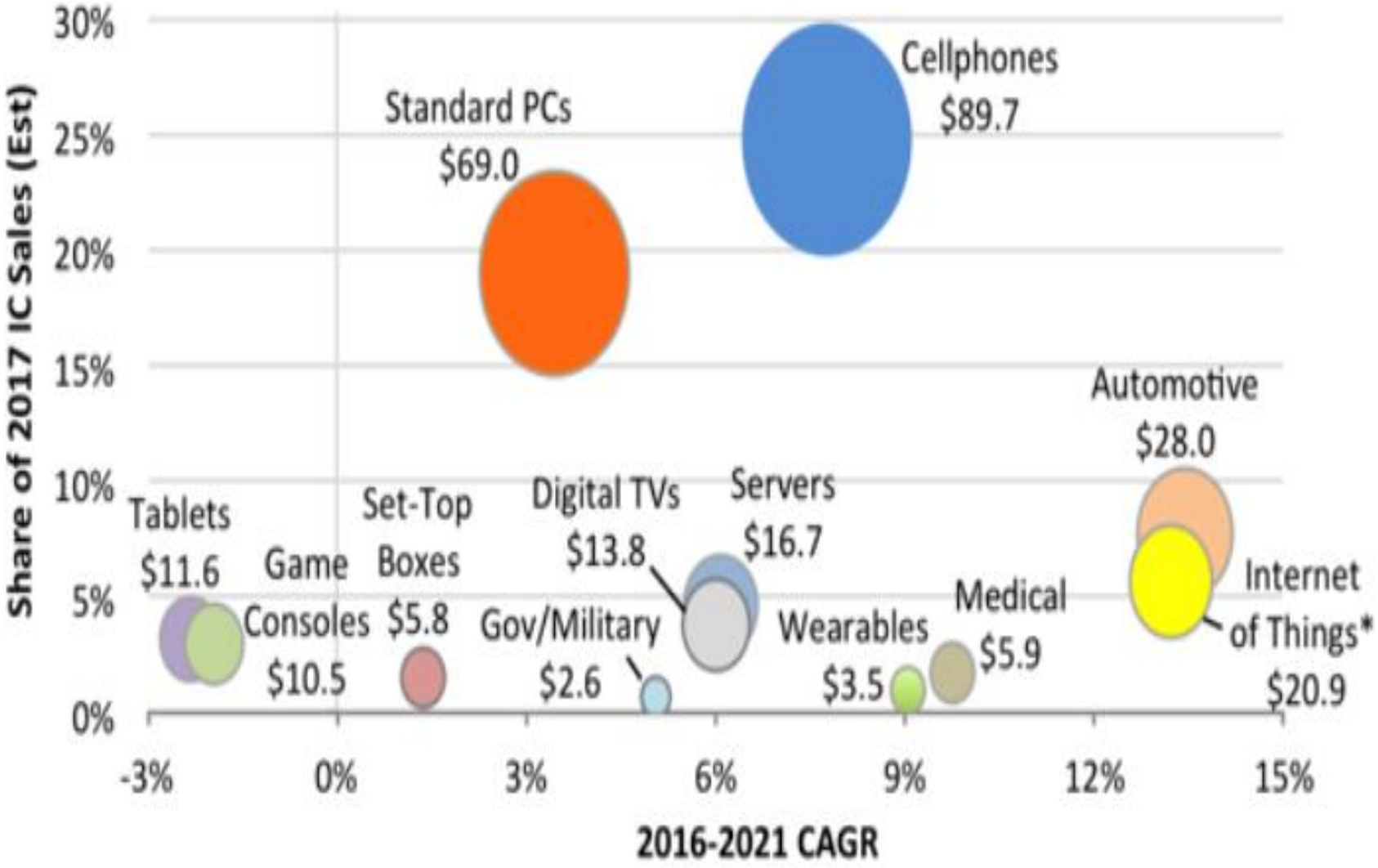
- More on Backend transistor (IGZO)/ Si photonics/ DRAM



5G + AI chips (AI PaaS, DNN), IOT , ADAS

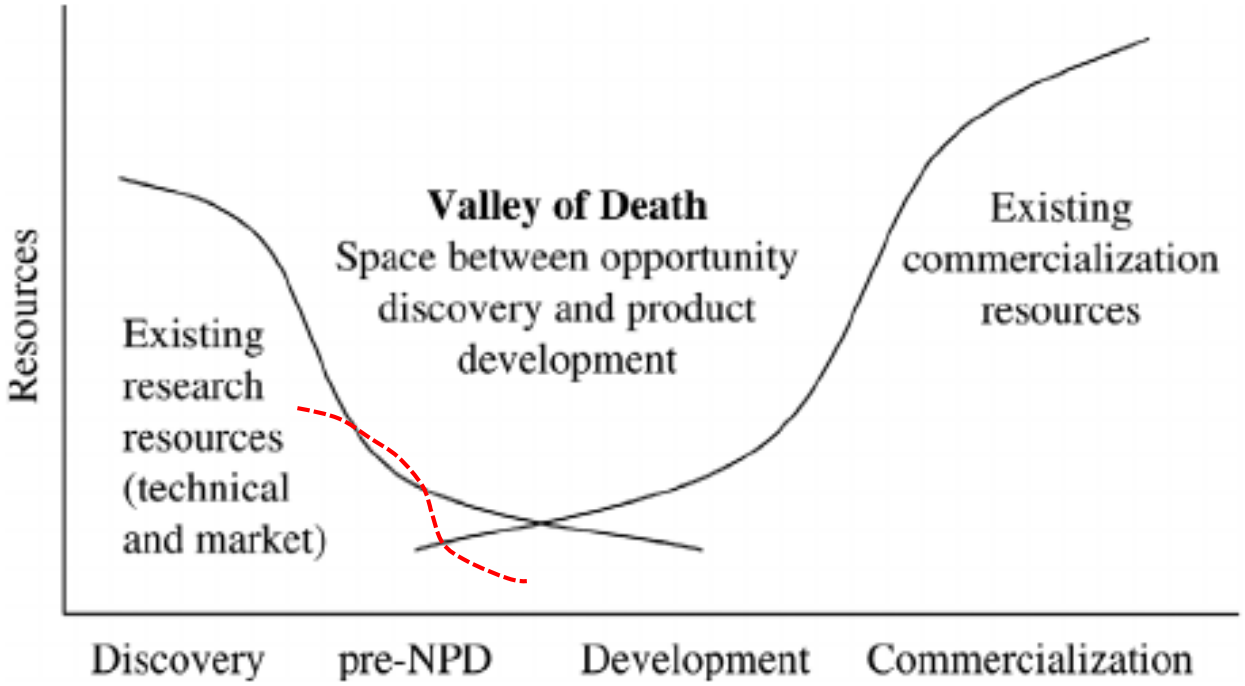
Future (ABC)

- A: AI (logic + memory)
- B: Big Data (memory, server)
- C: connectivity (RF, mm wave, TeraHz) 5G, 6G



*Covers only the Internet connection portion of systems.

R&D Strategy and Overcoming “the Valley of Death”



External Research (Pre-competitive Gov./Academia/ Consortium)

Internal Research

Pathfinding

TD

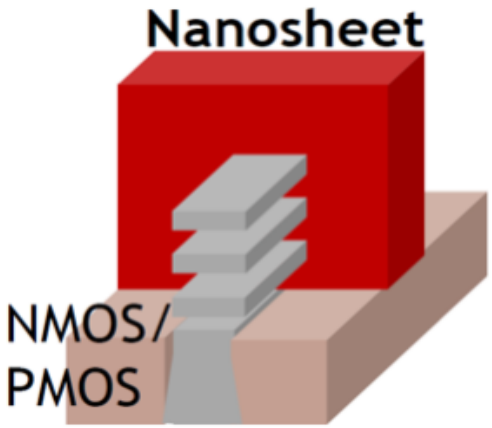
Manufacturing (Fab)

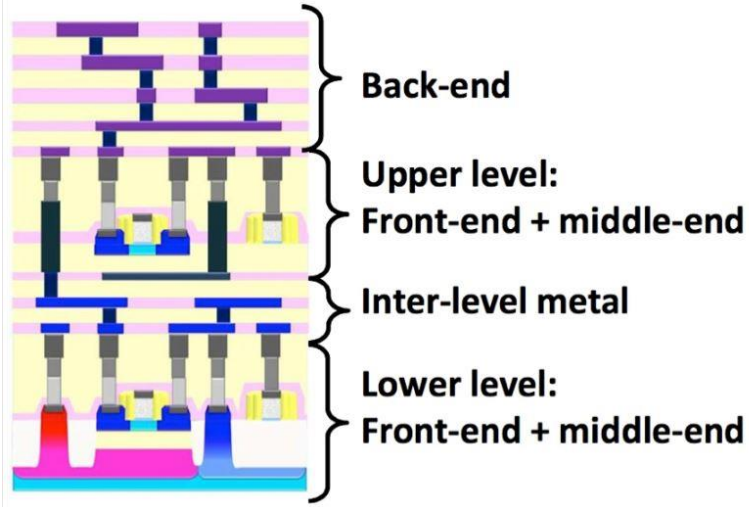
Source: Design Management Journal, Volume: 14, Issue: 1, Pages: 28-41, First published: 08 December 2019, DOI: (10.1111/dmj.12052)

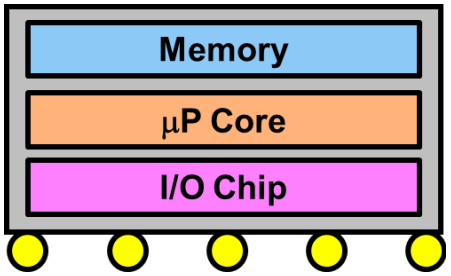
NPD: new product development

3Dx3Dx3D

Channel x Transistor(IC on IC) x Chiplet stacking







- **Channel stacking: stacked nanosheet (NS) can enhance the I_{on} at fixed footprint.**
- **Transistor stacking: stacking devices vertically can further increase device density. Monolithic vs Sequential (TFT)**
- **Chiplet stacking: dies with different technologies and applications can stack vertically→ area saving and low energy consumption**

「白米滿足人類食物需求；電晶體滿足人類精神所需」

蔡明介董事長

outline

- **Intro: node name/gate length high mobility/high K/FinFET(3D FETs)**
- **Roadmap: IRDS/IMEC/ tsmc/Intel**
- **Status quo of 2/3 nm**
- **Beyond 3 nm: nanosheet/CFET**
- **Si-based vs 2D**

Moore's Law

Density 2x/generation(18 month/2 yrs/depends)

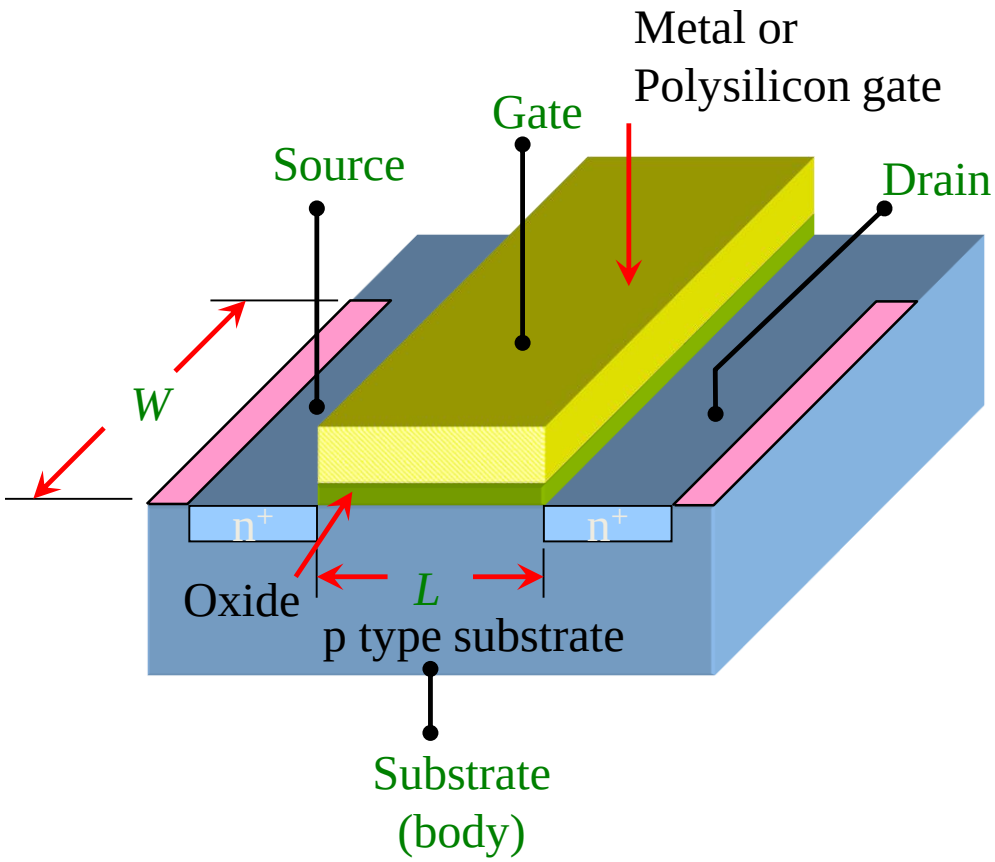
from n node to n+1 node

- Device/SRAM footprint (next generation) =Linear dimension (current generation) x 0.7

- 5nm has ~2x density of 7 nm
- Node name is density marker

Planar MOSFET structure upto 20/22nm

Smallest dimension?



Historically, L is the technology node up 250 nm/350 nm

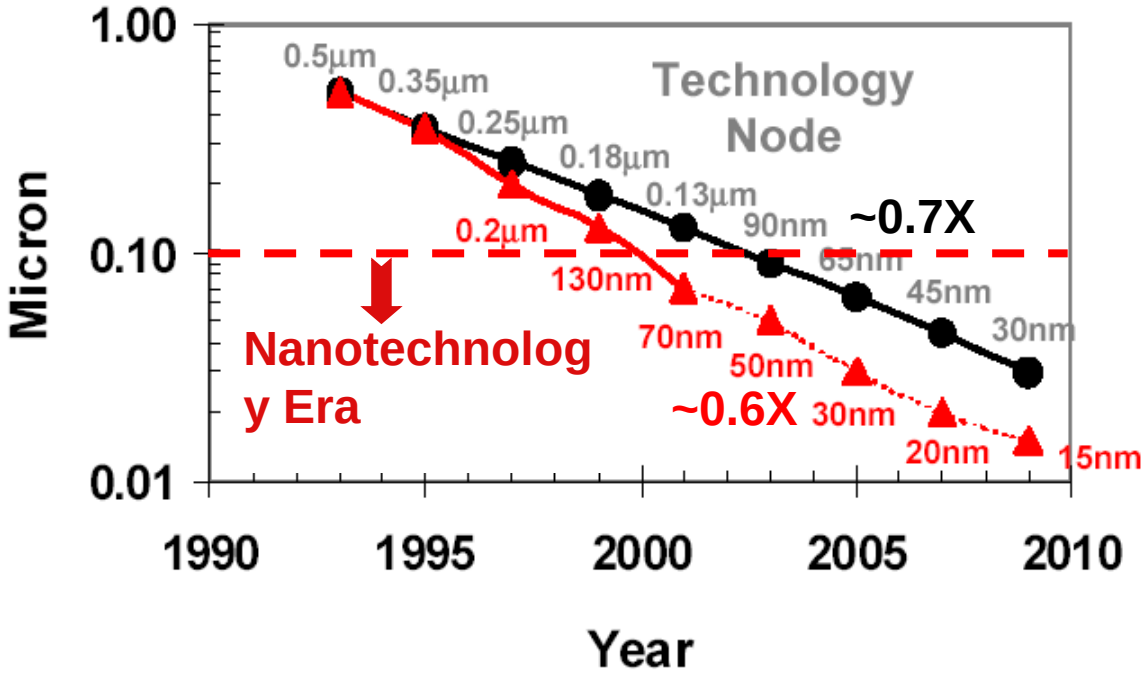
Why advanced nodes?

Table 9: Calculation of foundry sale price per chip in 2020 by node

Line	Node (nm)	90	65	40	28	20	16/12	10	7	5
1	Mass production year and quarter[220]	2004 Q4	2006 Q4	2009 Q1	2011 Q4	2014 Q3	2015 Q3	2017 Q2	2018 Q3	2020 Q1
2	Capital investment per wafer processed per year	$4,649	$5,456	$6,404	$8,144	$10,356	$11,220	$13,169	$14,267	$16,746
3	Net capital depreciation at start of 2020 (25.29% / year)	65%	65%	65%	65%	65%	65%	55.1%	35.4%	0.0%
4	Undepreciated capital per wafer processed per year (remaining value at start of 2020)	$1,627	$1,910	$2,241	$2,850	$3,625	$3,927	$5,907	$9,213	$16,746
5	Capital consumed per wafer processed in 2020	$411	$483	$567	$721	$917	$993	$1,494	$2,330	$4,235
6	Other costs and markup per wafer	$1,293	$1,454	$1,707	$2,171	$2,760	$2,990	$4,498	$7,016	$12,753
7	Foundry sale price per wafer	$1,650	$1,937	$2,274	$2,891	$3,677	$3,984	$5,992	$9,346	$16,988
8	Foundry sale price per chip	$2,433	$1,428	$713	$453	$399	$331	$274	$233	$238

Node name is NOT gate length

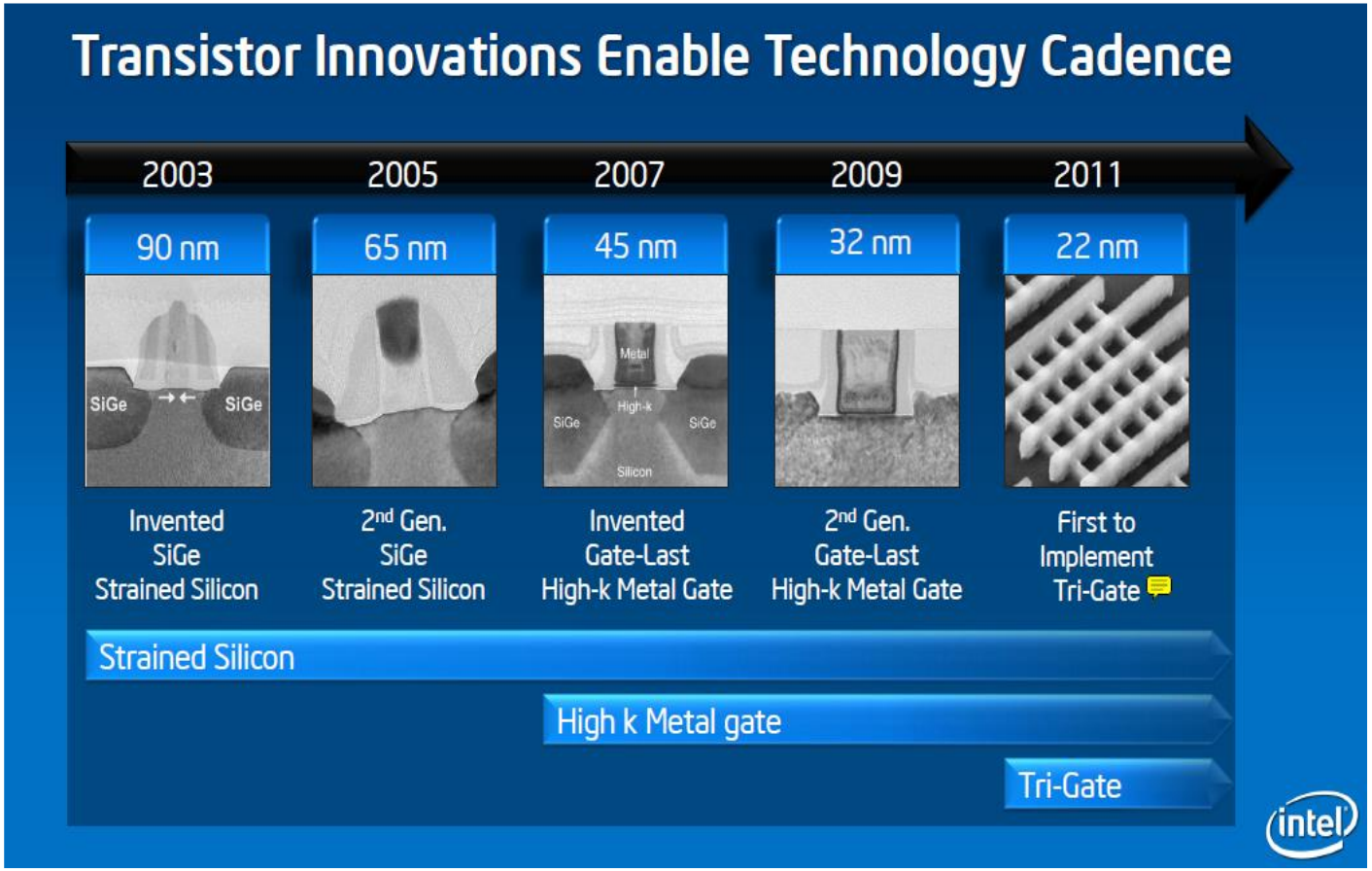
Evolution in Channel Length



Nanotechnology is in production today

(~20 yr ago)

Strained Si starting from 90 nm node

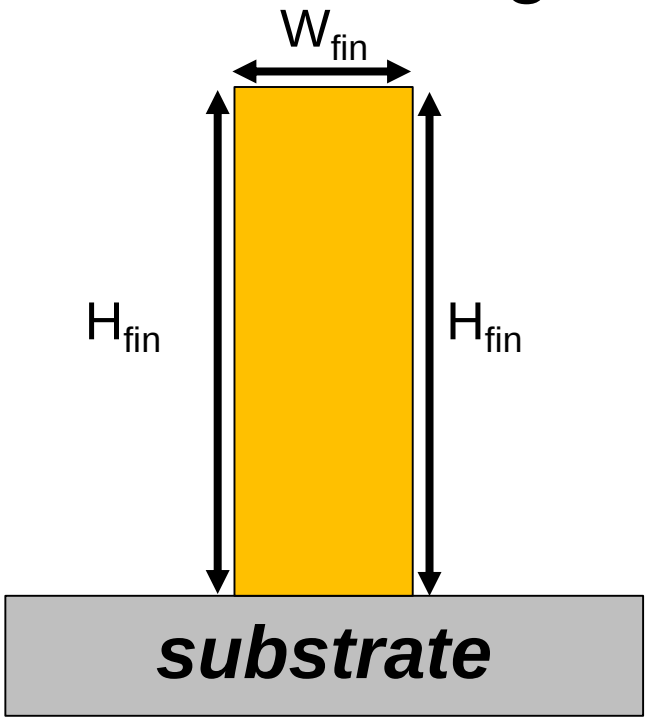


- Enabling technologies: high mobility, high-k/metal gate, 3D transistor

Make your FinFET/GAA/Nanosheet

I_{ON} Improvement -- W_{eff}

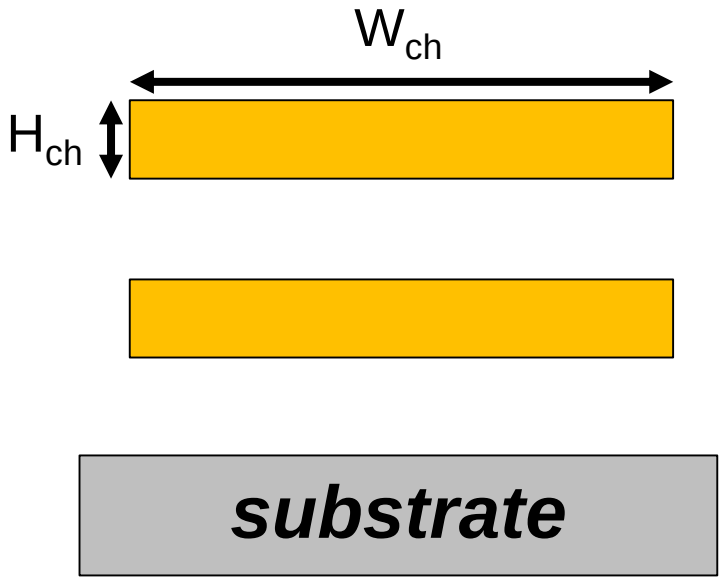
FinFET = Tri-gate



$W_{eff} = 2H_{fin} + W_{fin}$

Footprint = W_{fin} x L

Nanosheet/nanoribbon/
Multi bridge channel = GAA



$W_{eff} = (2H_{ch} + 2W_{ch})$ x floor number

I_{ON} Improvement by technologies

$$I_d = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{gs} - V_t)^a \times \text{stacking number (floor numbe}$$

→ Nanosheet (2nm)

FinFET (16nm) : $2H_{FIN}+W_{FIN}$

Intel : 22nm

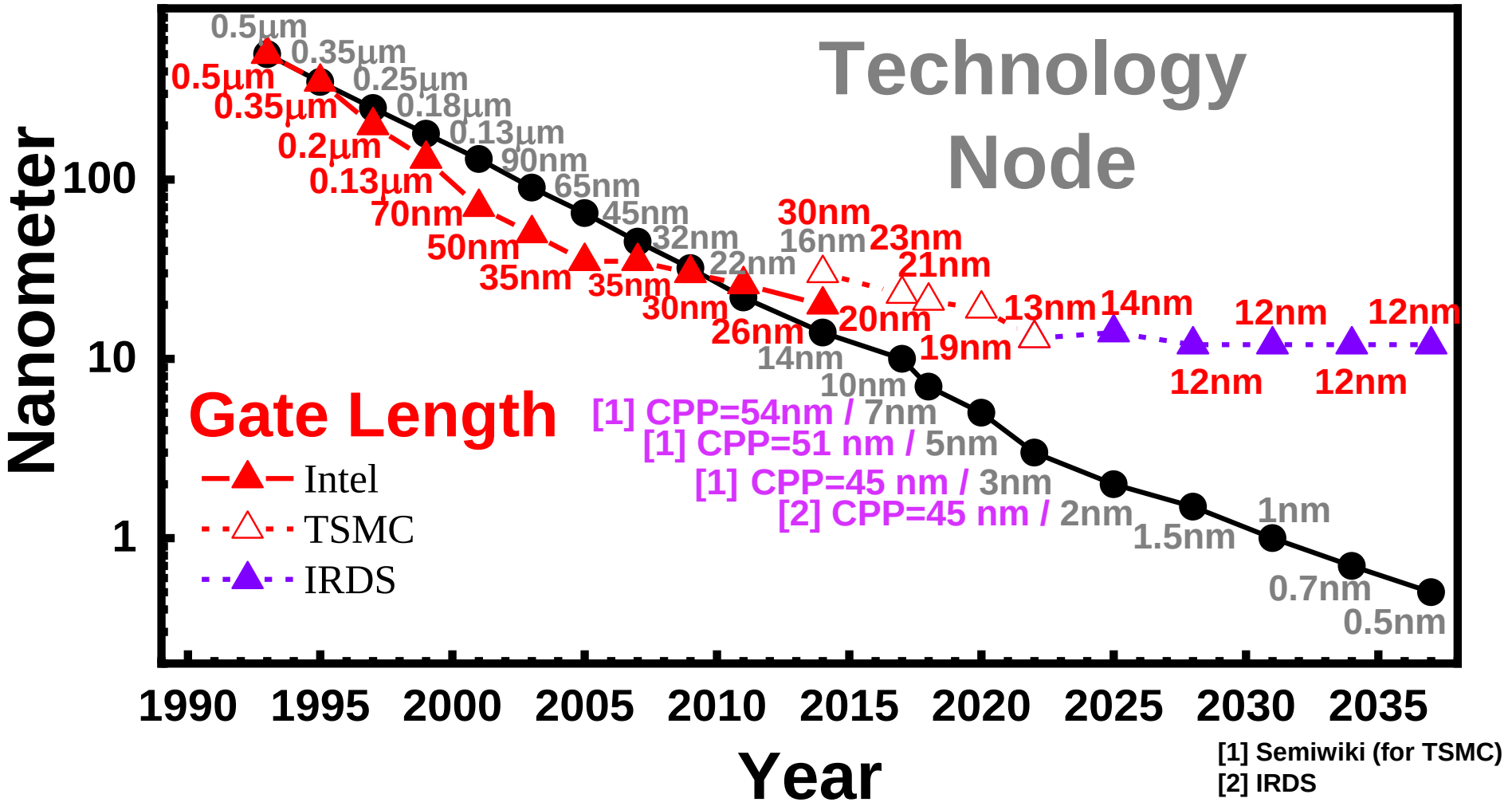
High mobility $\frac{\varepsilon_{ox}}{t_{ox}} \rightarrow HK(45nm\ node)$

$$\frac{\varepsilon_{ox}}{t_{ox}} = \frac{\varepsilon_{HK}}{t_{HK}} = \frac{\varepsilon_{ox}}{EOT} \qquad \varepsilon = k\varepsilon_0$$

t_{ox} >1nm. Otherwise, I_g is significant

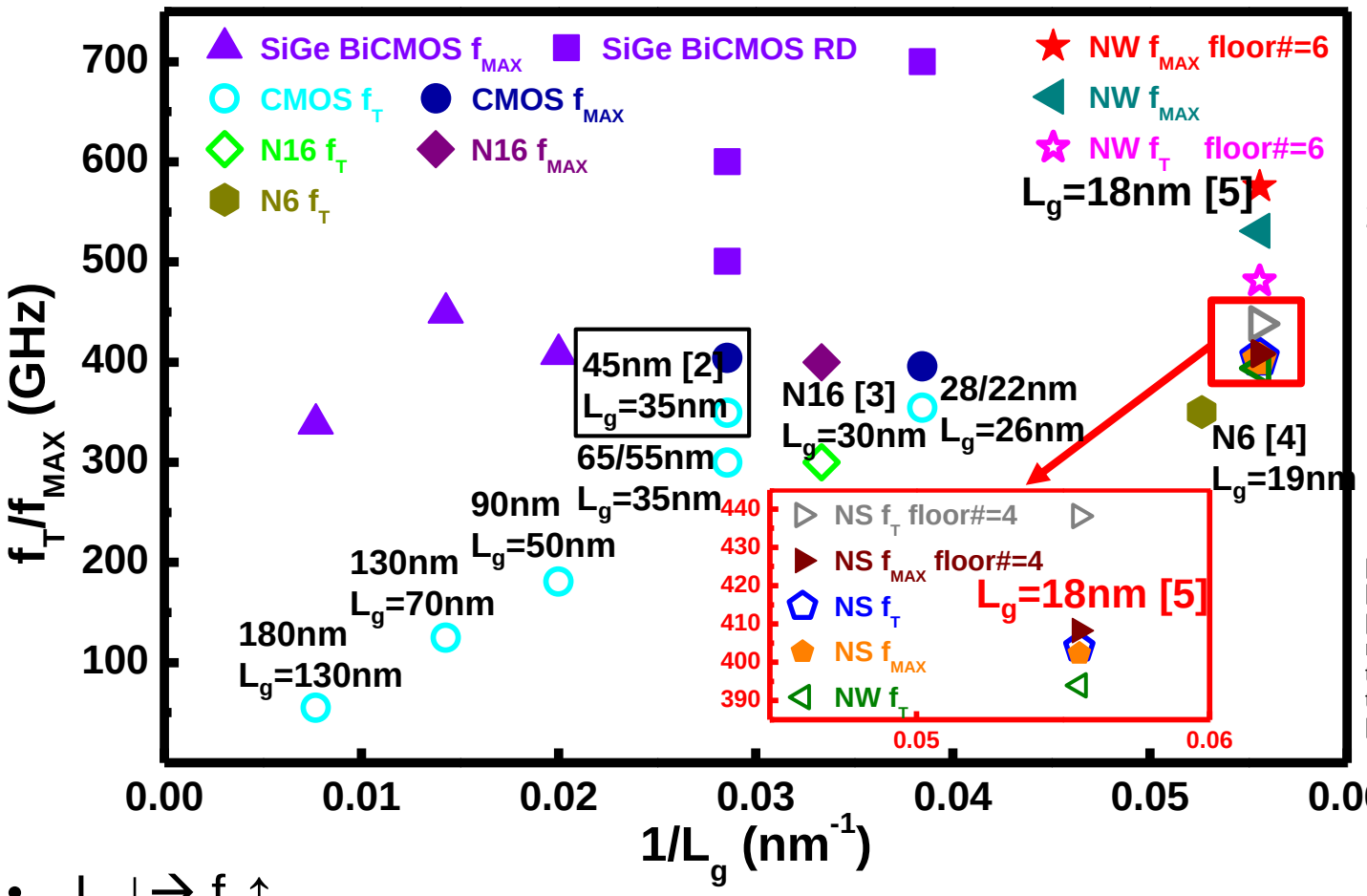
1. Strained Si (90nm)
2. High mobility channel (5nm SiGe p-channel)

L_g scaling with different nodes



- L_g can be larger than node name due to narrow width in footprint

Benchmark of f_T and f_{MAX}



N. Cahoon et al., IRPS, 2022.

[2] K. Mistry et al., IEDM, 2007.
[3] https://zhuanlan.zhihu.com/p/199004138
[4] https://semiwiki.com/semiconductor-manufacturers/tsmc/299944-highlights-of-the tsmc-technology-symposium-2021-silicon-technology/ (TSMC 2021 symposium)
[5] H.-C. Lin et al., IEEE EDL, 2022.

- $L_g\downarrow \rightarrow f_T\uparrow$
- FinFET [3, 4] with low electron mobility → f_T can not improve with $L_g\downarrow$
- NSs and NWs have better f_T/f_{MAX} than previous CMOS RF [5]

Back to the Basics : Power Consumption

" Keys are reduction of Voltage, Parasitic Capacitance, and Leakage current "

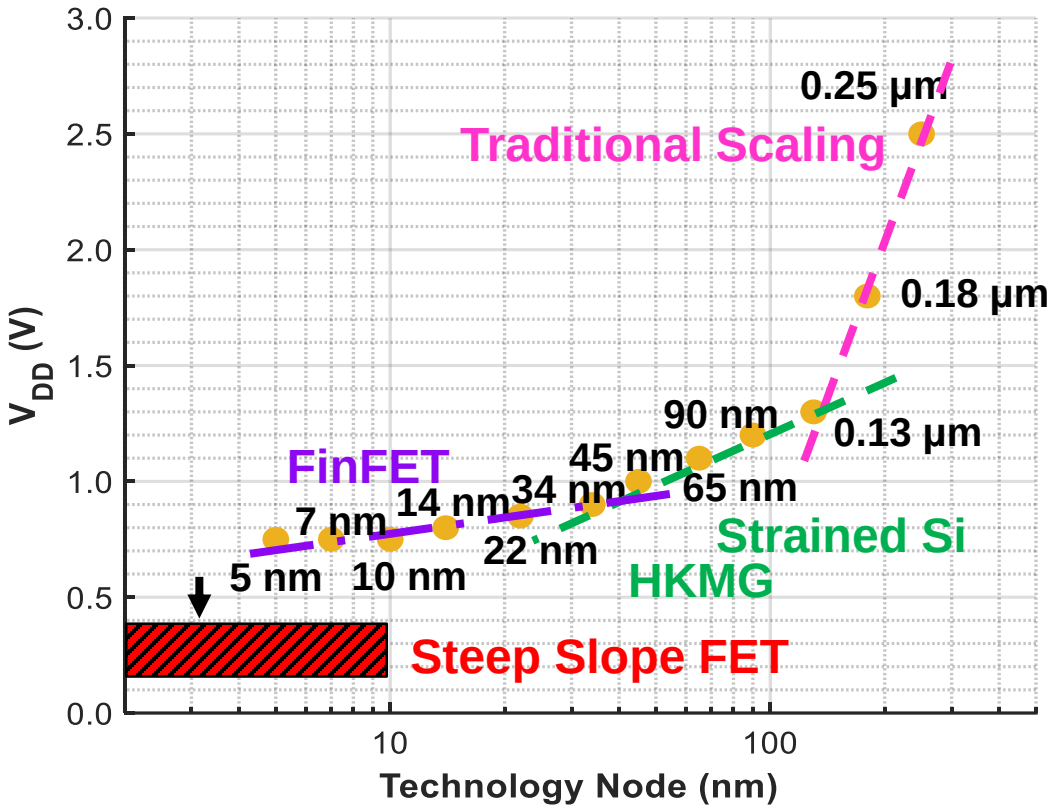
$$\text{Power} \equiv \underbrace{V_{dd}}_{①} \cdot \underbrace{I_{sb}}_{③} + \alpha \cdot \Sigma \underbrace{C_i}_{②} \cdot \underbrace{V_{dd}^2}_{①} \cdot F$$

α : Average Percentage of Switching Capacitance (~ activity factor)
ΣC_i : Sum of Gate and Interconnection Capacitances
Vdd : Supply Voltage, F : Clock Frequency

Example: For NAND only when input=(1,1) output will switched

$$\alpha = \frac{3}{4} * \frac{1}{4} = \frac{3}{16}$$

V_{DD} Scaling Trend vs Technology Node



Ref:

10 nm: IRDS 2017 EDITION MORE MOORE

7 nm:

[1] IRDS 2018 UPDATE MORE MOORE

[2] Hung-Li Chiang et al., Design Technology Co-Optimization for Cold CMOS Benefits in Advanced Technologies, 2021 IEDM

[3] J. Chang *et al*., ISSCC, 2017(TSMC)

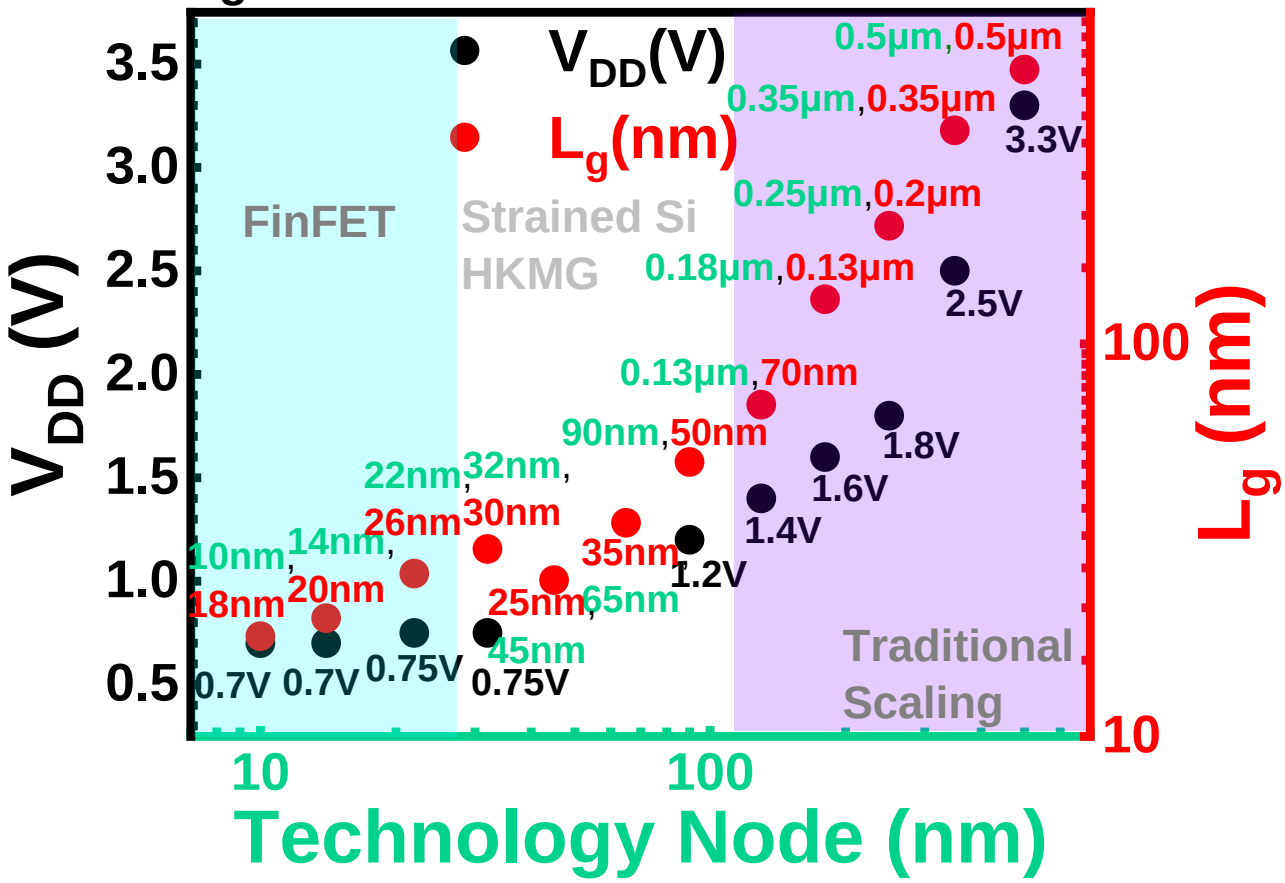
5 nm:

[1] IRDS 2021 UPDATE MORE MOORE

[2] J. Chang *et al*., ISSCC, 2020(TSMC)

- iMOS
- TFET
- NCFET

V_{DD} and L_g Scaling Trend vs Technology Node



- V_{DD} not available for 65nm and 45nm node

Intel technology node

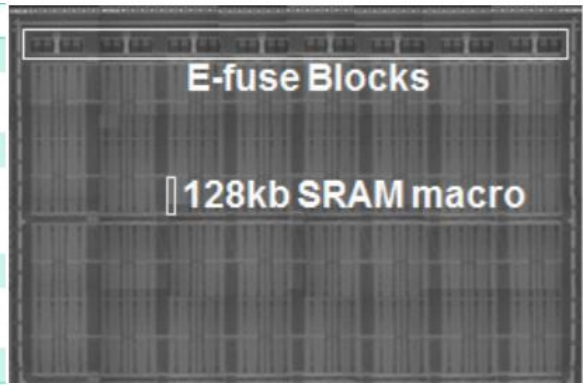
1993	P852	0.5 µm	4	P5			T_{ox}	8 nm	Gate Dielectric	SiO_2
							V_{dd}	3.3 V	SRAM	44 µm²
							L_g	500 nm		
							CPP		MMP	
1995	P854	0.35 µm	4	P6			T_{ox}	6 nm	Gate Dielectric	SiO_2
							V_{dd}	2.5 V	SRAM	20.5 µm²
							L_g	350 nm		
							CPP	920 nm	MMP	880 nm
1997	P856	0.25 µm	5	P6			T_{ox}	4.08 nm	Gate Dielectric	SiO_2
							V_{dd}	1.8 V	SRAM	10.26 µm²
							L_g	200 nm		
							CPP	500 nm	MMP	640 nm
1998	P856.5	0.25 µm	5	P6			T_{ox}	4.08 nm	Gate Dielectric	SiO_2
							V_{dd}	1.8 V	SRAM	9.26 µm²
							L_g	200 nm		
							CPP	475 nm	MMP	608 nm
1999	P858	0.18 µm	6	NetBurst			T_{ox}	2.0 nm	Gate Dielectric	SiO_2
							V_{dd}	1.6 V	SRAM	5.59 µm²
							L_g	130 nm		
							CPP	480 nm	MMP	500 nm
2001	P860	0.13 µm	6	Pentium M			T_{ox}	1.4 nm	Gate Dielectric	SiO_2
							V_{dd}	1.4 V	SRAM	2.45 µm²
							L_g	70 nm		
							CPP	336 nm	MMP	345 nm
2003	P1262 (CPU) P1263 (SoC, I/O)	90 nm	7	Pentium M			T_{ox}	1.2 nm	Gate Dielectric	SiO_2
							V_{dd}	1.2 V	SRAM	1.00 µm²
							L_g	50 nm		
							CPP	260 nm	MMP	220 nm
2005	P1264 (CPU) P1265 (SoC, I/O)	65 nm	8	Core, Modified Pentium M			T_{ox}	1.2 nm	Gate Dielectric	SiO_2
							V_{dd}		SRAM	0.570 µm²
							L_g	35 nm		
							CPP	220 nm	MMP	210 nm
2007	P1266 (CPU) P1267 (SoC, I/O)	45 nm	9	Penryn, Nehalem			T_{oxe}	1 nm	Gate Dielectric	High-κ
							V_{dd}		SRAM	0.346 µm²
							L_g	25 nm		
							CPP	160 nm	MMP	180 nm
2009	P1268 (CPU) P1269 (SoC, I/O)	32 nm	10	Westmere, Sandy Bridge			T_{oxe}	1 nm	Gate Dielectric	High-κ
							V_{dd}	0.75 V	SRAM	0.148 µm²
							L_g	30 nm		
							CPP	112.5 nm	MMP	112.5 nm
2011	P1270 (CPU) P1271 (SoC, I/O)	22 nm	11	Ivy Bridge, Haswell			T_{oxe}	0.9 nm	Gate Dielectric	High-κ
							V_{dd}	0.75 V	SRAM	0.092 µm²
							L_g	26 nm		
							CPP	90 nm	MMP	80 nm
							P_{fin}	60 nm		
							W_{fin}	8 nm	H_{fin}	34 nm
2014	P1272 (CPU) P1273 (SoC, I/O)	14 nm	12	Broadwell, Skylake, Kaby Lake, Coffee Lake			T_{oxe}		Gate Dielectric	High-κ
							V_{dd}	0.70 V	SRAM	0.0499 µm²
							L_g	20 nm		
							CPP	70 nm	MMP	52 nm
							P_{fin}	42 nm		
							W_{fin}	8 nm	H_{fin}	42-46 nm
2019	P1274 (CPU) P1275 (SoC, I/O)	10 nm	12-13	Cannon Lake, Icelake, Tigerlake, Sapphire Rapids			T_{oxe}		Gate Dielectric	High-κ
							V_{dd}	0.70 V	SRAM	0.0312 µm²
							L_g	18 nm		
							CPP	54 nm	MMP	36 nm
							P_{fin}	34 nm		
							W_{fin}	7 nm	H_{fin}	44-55 nm
2021	P1276 (CPU) P1277 (SoC, I/O)	7 nm		Granite Rapids						
2024	P1278 (CPU) P1279 (SoC, I/O)	5 nm								

- **Intel technology node**

SRAM Core VDD of Test Chip (TSMC)

7nm:

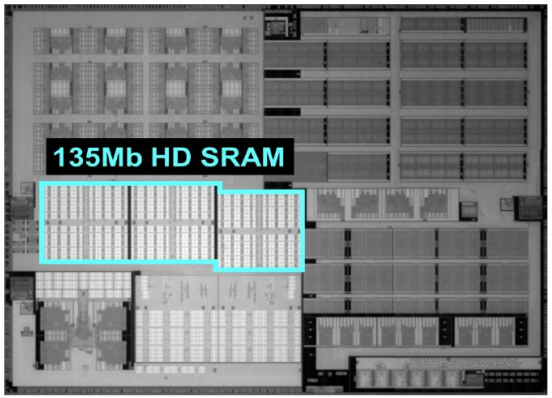
Technology	7nm HK-MG FinFET
Metal scheme	1P7M
Supply voltage	Core: 0.75V IO: 1.8V
Bit cell size	0.027μm²
SRAM macro configuration	4096x32 MUX=16 258 bits/BL, 272 bits/WL
SRAM capacity	256Mb
Test Features	Row/Column Redundancy Programmable E-fuse
Chip size	5903μm x 7223μm = 42mm²



J. Chang *et al*., ISSCC 2017(TSMC)

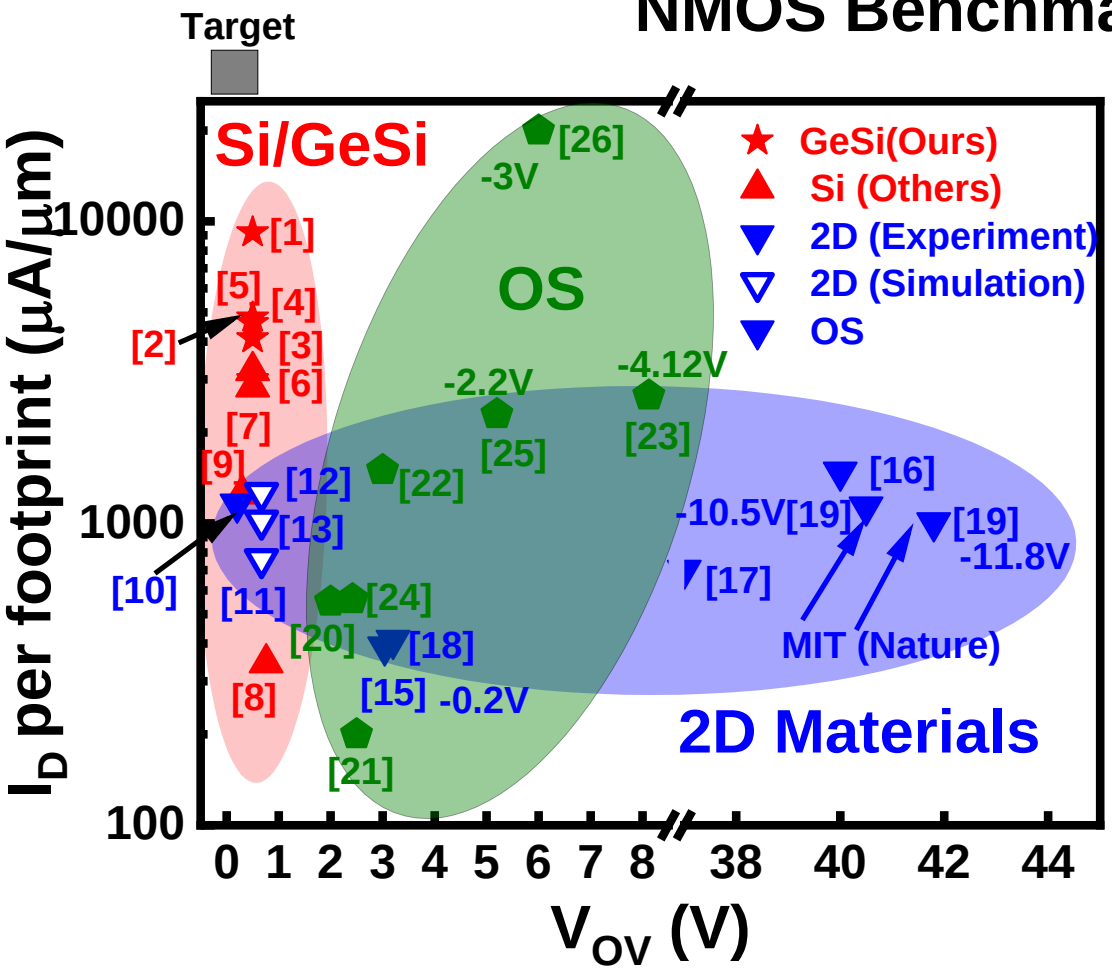
5nm:

Technology	5nm HK-MG FinFET
Supply voltage	Core: 0.75V IO: 1.2V
Bit cell size	0.021μm²
SRAM macro configuration	1024x144 MUX4 256 bits/BL, 288 bits/WL
SRAM capacity	135Mb
Test Features	Column Redundancy Programmable E-fuse
Chip size	10mm x 7.98mm = 79.8mm²



J. Chang *et al*., ISSCC, 2020(TSMC)

NMOS Benchmark



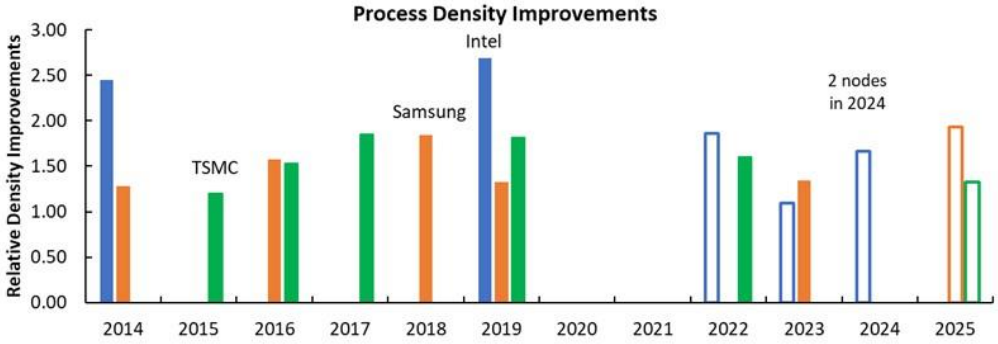
- **Negative Vt for NMOS is not acceptable for IC design**

[1] GAA(8 Stacked NW($Ge_{0.95}Si_{0.05}$) Y. -
[2] GAA(8 Stacked NS($Ge_{0.95}Si_{0.05}$) Y. -
[3] GAA(7 Stacked nanowire($Ge_{0.95}Si_{0.05}$) Y. -
[4] GAA(6 Stacked NW($Ge_{0.95}Si_{0.05}$ Y.-C
[5] GAA(2 Stacked NW($Ge_{0.98}Si_{0.02}$ Y.-
[6] GAA (2 Sacked NW) S. E STI
[7] SiGe Forksheet H. M
[8] GAA poly-Si F.-
[9] TSMC 7nm FinFET S. A. C
[10] Carbon nanotube (CNT) TG C. C
[11] $MoTe_2$ M.
[12] WS_2 M.
[13] WSe_2 M.
[14] Bi-1L MoS_2 (BG) She
[15] MoS_2 back gate (BG) A.-S
[16] MoS2 Back Gate Y. L
[17] MoS2 Back Gate C.
[18] GAA MoS2 Y.
[19] Bi-1L MoS_2 (BG) P.C.
[20] IWO double gate (DG) H.
[21] IWO BG W.
[22] DG IGZO W.
[23] BG In2O3 Lia
[24] BG ZnO Cha
[25] DG ITO Wa
[26] GAA In2O3 Z. Z

Density Jumps

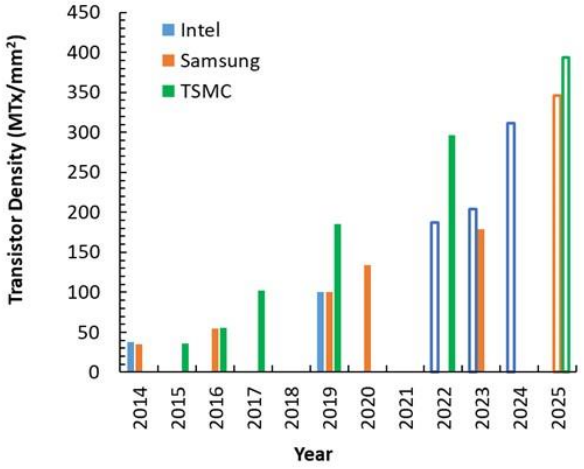
IC KNOWLEDGE LLC

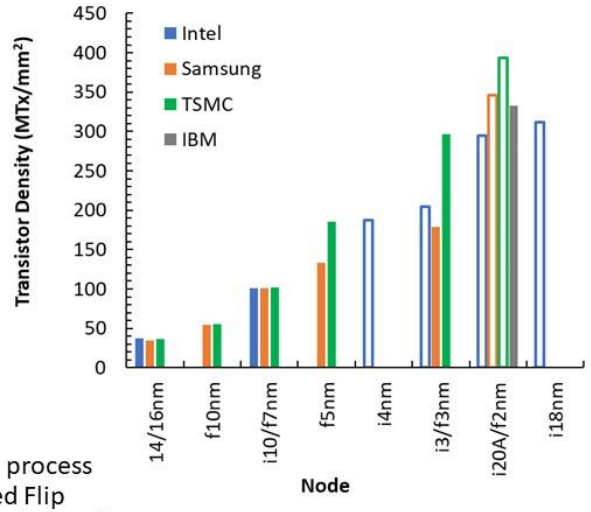
- 2014 to 2019, Samsung and TSMC each do four nodes while Intel does two.
- 2020 to 2025, Samsung and TSMC each do two nodes while Intel does four?
- Intel flips the script on the foundries.



Transistor Density Trends

IC KNOWLEDGE LLC



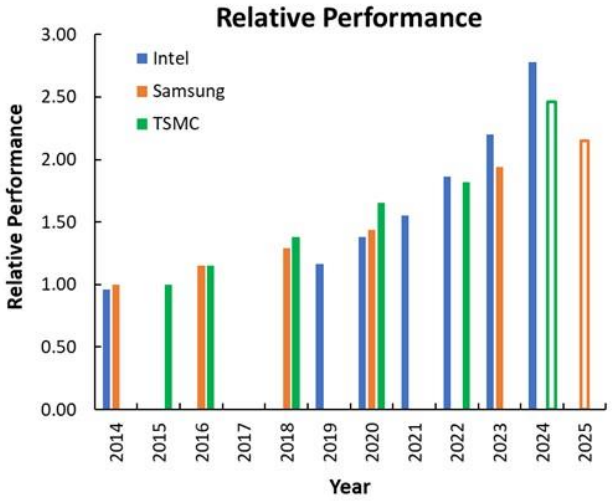


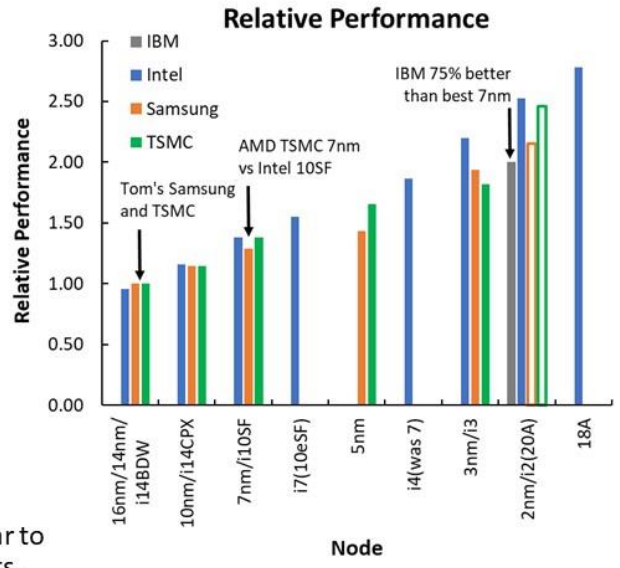
To calculate a standard metric for transistor density for each process we consider a weighting of 60% NAND cells and 40% Scanned Flip Flop cells. We use the M2P, Tracks, CPP and DDB/SDB to calculate cell sizes

- **TSMC to maintain the density lead through 2025.**

Relative Performance Trends

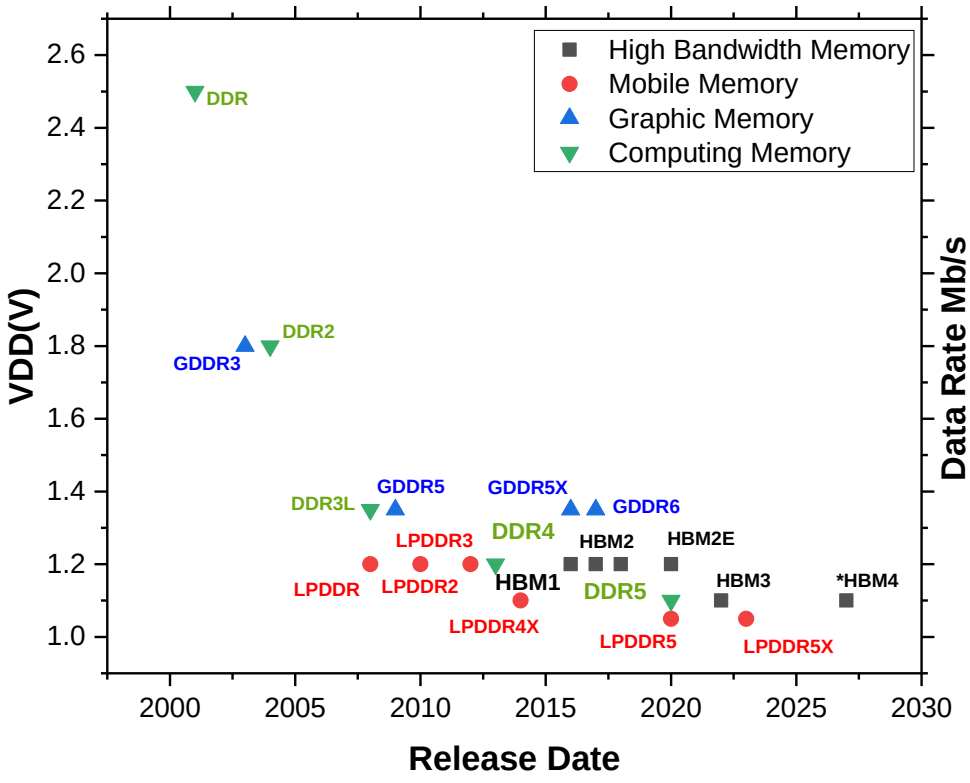
IC KNOWLEDGE LLC

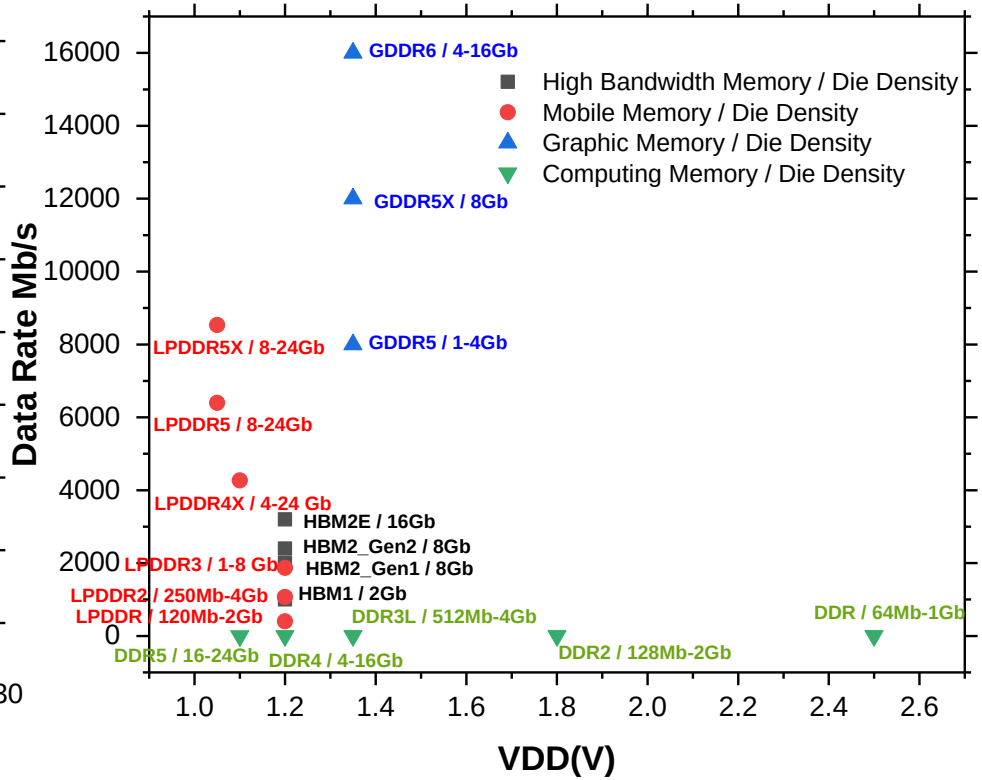




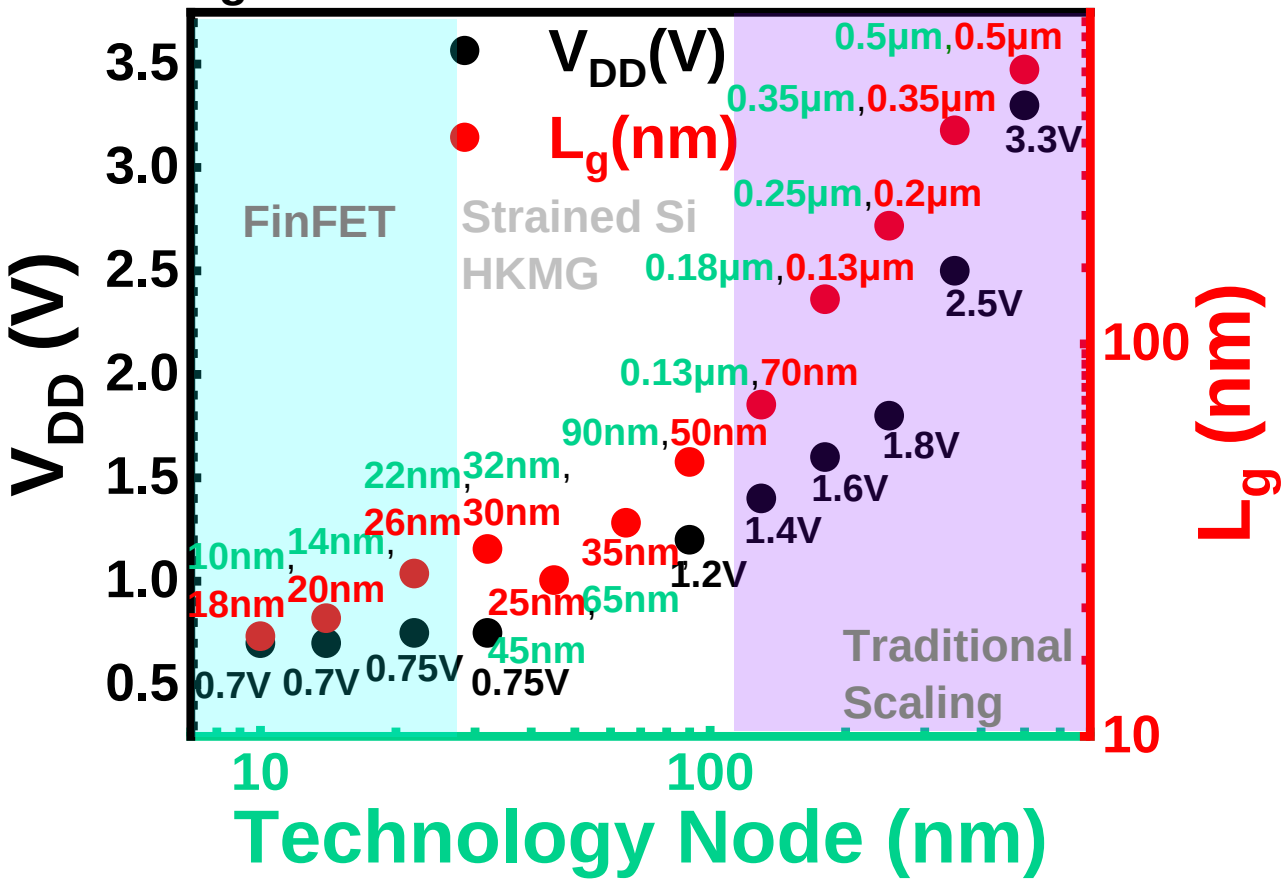
TSMC/Samsung - 16/14nm per Tom's hardware, then scaled based on company announcements. Intel based on 10SF similar to TSMC 7nm and then scaled based on company announcements. IBM based on their 2nm announcement.

- **Intel may take the performance lead both on a year basis and a node basis.**


High Bandwidth Memory
Mobile Memory
Graphic Memory
Computing Memory
DDR
DDR2
GDDR3
DDR3L
GDDR5
GDDR5X
GDDR6
LPDDR3
DDR4
HBM2
HBM2E
LPDDR
LPDDR2
HBM1
DDR5
HBM3
*HBM4
LPDDR4X
LPDDR5
LPDDR5X
VDD(V)
Release Date
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
2000
2005
2010
2015
2020
2025
2030


High Bandwidth Memory / Die Density
Mobile Memory / Die Density
Graphic Memory / Die Density
Computing Memory / Die Density
GDDR6 / 4-16Gb
GDDR5X / 8Gb
GDDR5 / 1-4Gb
LPDDR5X / 8-24Gb
LPDDR5 / 8-24Gb
LPDDR4X / 4-24 Gb
HBM2E / 16Gb
HBM2_Gen2 / 8Gb
HBM2_Gen1 / 8Gb
HBM1 / 2Gb
LPDDR3 / 1-8 Gb
LPDDR2 / 250Mb-4Gb
LPDDR / 120Mb-2Gb
DDR3L / 512Mb-4Gb
DDR / 64Mb-1Gb
DDR5 / 16-24Gb
DDR4 / 4-16Gb
DDR2 / 128Mb-2Gb
Data Rate Mb/s
VDD(V)
16000
14000
12000
10000
8000
6000
4000
2000
0
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6

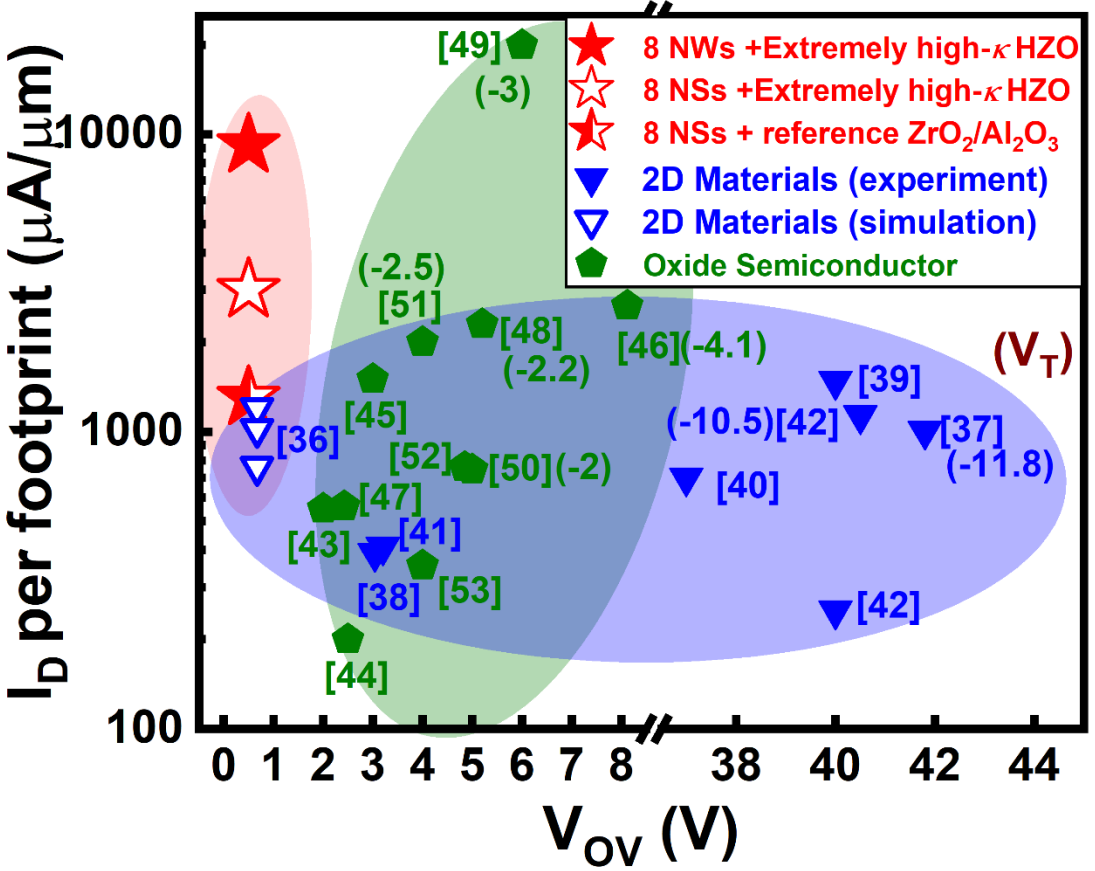
V_{DD} and L_g Scaling Trend vs Technology Node



- V_{DD} not available for 65nm and 45nm node

Intel technology node

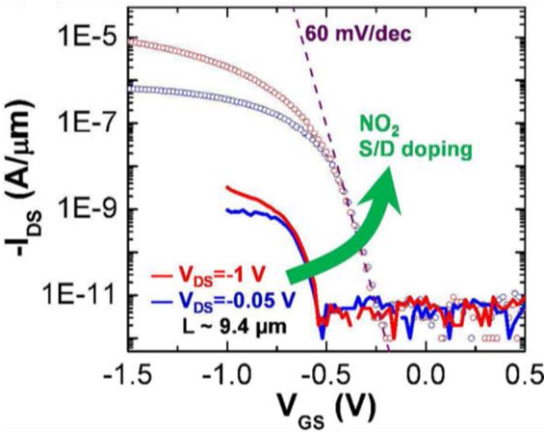
Honest Benchmark (I_D vs V_{OV})



2D	[36]	2D simulation	M. Luisier et al., IEDM, 2016
	[37]	Bi-1L MoS_2 (BG)	Shen, PC. et al., Nature 593, 211–217, 2021.
	[38]	MoS_2 back gate (BG)	A.-S. Chou et al., VLSI, 2020.
	[39]	MoS2 Back Gate	C. J. Mcclellan et al., ACS Nano, 2021, Stanford
	[40]	GAA MoS2	Y. Y. Chung et al. IEDM 2022, TSMC
	[41]	Bi-1L MoS_2 (BG)	P.C. Chen et al., Nature, 2021, MIT
	[42]	MoS_2	S. Ghosh et al., VLSI, 2023
OS	[43]	IWO double gate (DG)	H. Ye et al., IEDM, 2020, 28.3.
	[44]	IWO BG	W. Chakraborty et al., VLSI, 2020.
	[45]	DG IGZO	W. Lu et al IEDM 2022, CAS
	[46]	BG In2O3	Liao et al VLSI 2022, Purdue
	[47]	BG ZnO	Chand et al VLSI 2022, NUS
	[48]	DG ITO	Wahid et al IEDM 2022, Stanford
	[49]	GAA In2O3	Z. Zheng et al IEDM 2022, Purdue, 4-3
	[50]	In2O3	D. Zheng et al., VLSI, 2023
	[51]	ITO	Y. Kang et al., VLSI, 2023,
	[52]	ITO-IGZO	S. Hooda et al., VLSI , 2023
	[53]	IGZO	J. Zhang et al., VLSI, 2023

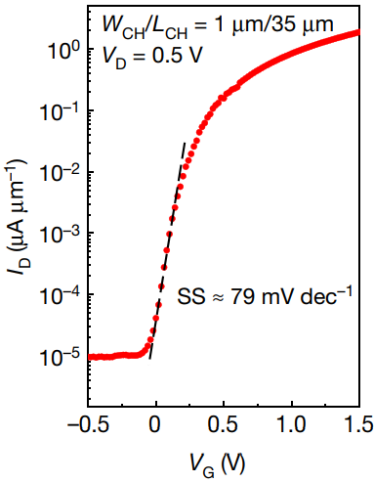
- For logic, low VDD and n/p FETs are needed → oxide semiconductor (OS) and 2D material is not applicable now
- For RF, higher VDD and only nFET for large output power, OS and 2D material have more chance

UC Berkeley
2D (WSe_2)
SS=60mV/dec



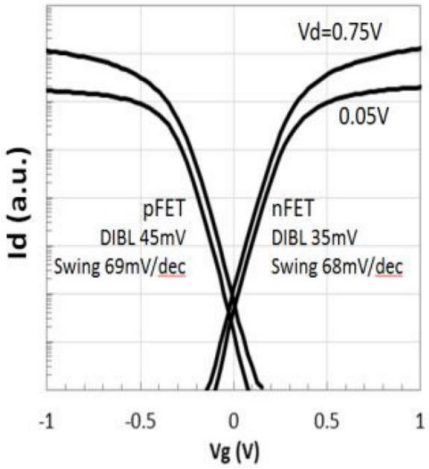
Fang, H. *et al.*, Nano Letters, 2012, 12(7), 3788–3792.

NTU 邱雅萍教授
2D (MoS_2)
SS=79mV/dec



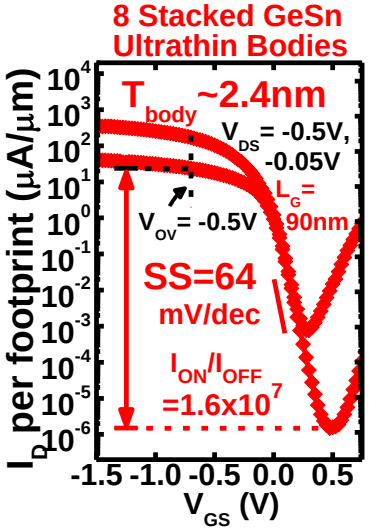
Huang, JK., Wan, Y., Shi, J. *et al.*
Nature 605, 262–267 (2022).

TSMC 5nm
nFET
SS=68mV/dec
pFET



Geoffrey Yeap *et al.*, IEDM, 2019, 36.7.

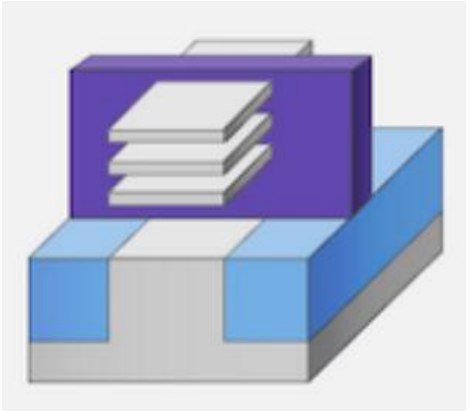
Our ultrathin body pFET
SS=64mV/dec



Chung-En Tsai *et al.*, VLSI, 2022.

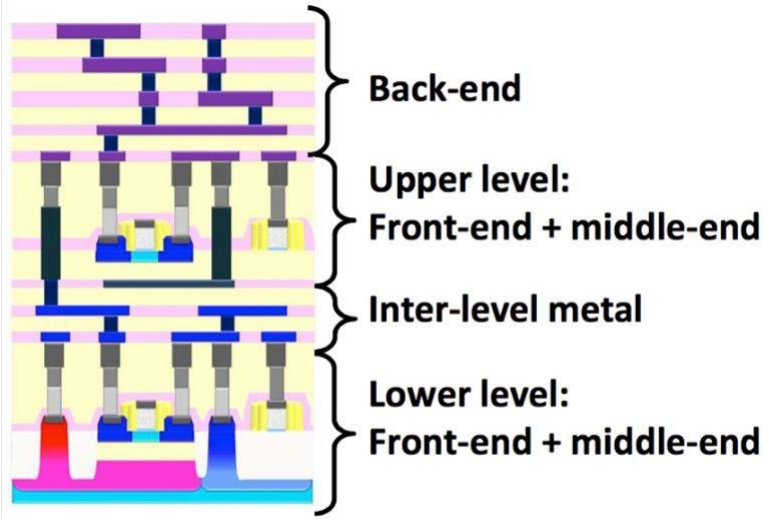
3Dx3Dx3D

3D



Channel stacking

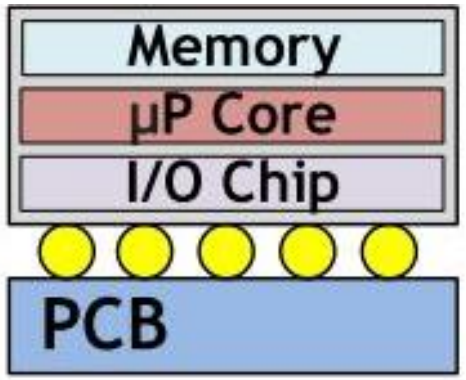
3D



Transistor stacking

Monolithic /Sequential (w/ BEOL)

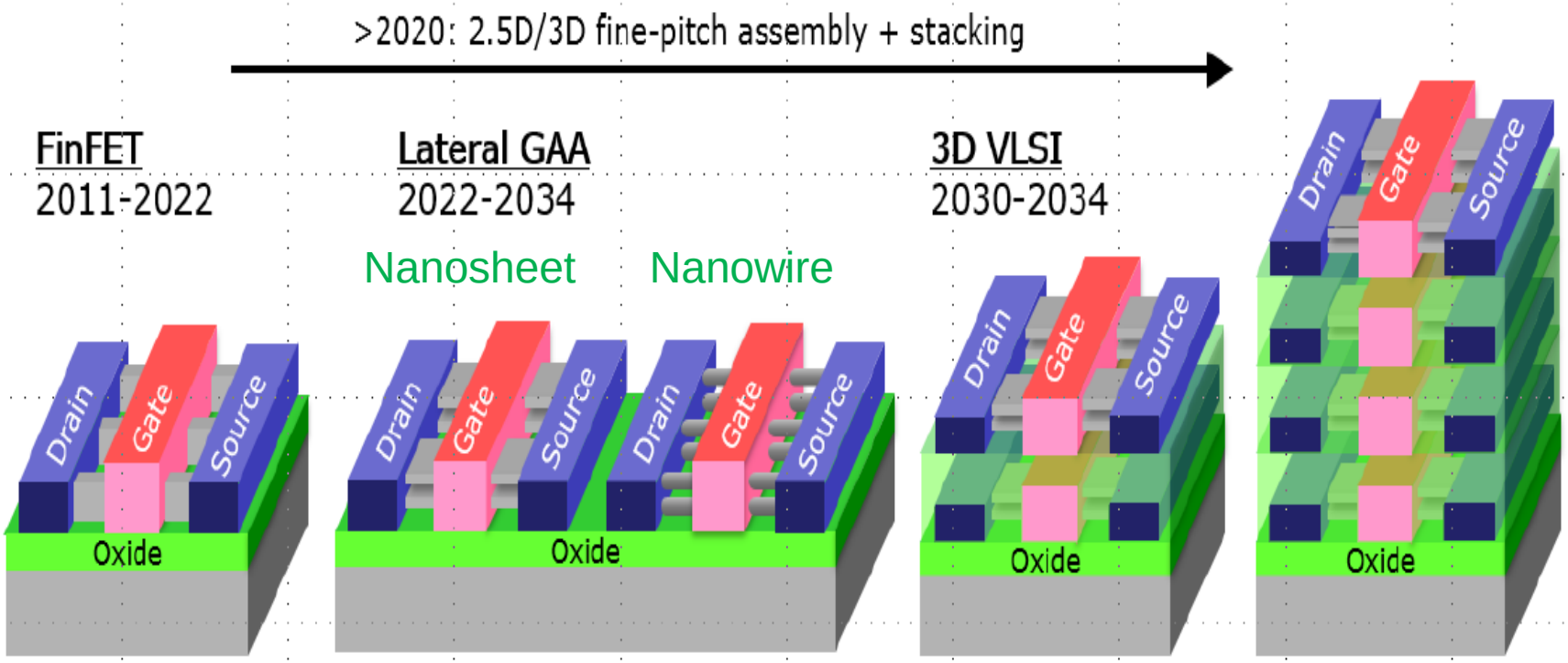
3D (Far backend/ Heterogeneous integration)



Dielets/Chiplets Stacking

- SOIC (tsmc)
- X-Cube (Samsung)
- FOVEROS (Unique and special in Greek) /EMIB (Intel)

IRDS,2020 Edition More Moore



- Increasing drive by taller fin
- Better channel control for better perf-power

- Increasing drive by stacked devices
- Better channel control for better perf-power
- Reduced footprint stdcell

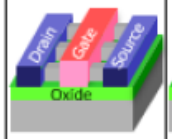
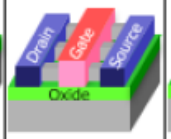
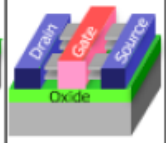
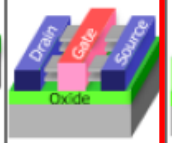
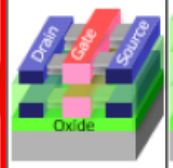
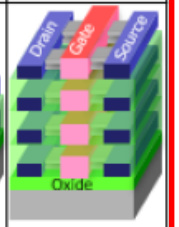
- Sequential integration/fine-pitch stacking (e.g. logic, memory, NVM, analog, IO, RF, sensors)


台積電-臺灣大學聯合研發中心
TSMC - NTU JOINT RESEARCH CENTER
Planar FET
GAAFET
FINFET
GAAFET

Roadmap : IRDS,2021 Edition

More Moore

- *Channel stacking starts at 2 nm ndoe*
- *Transistor stacking starts at 1 nm node.*

YEAR OF PRODUCTION	2021	2022	2025	2028	2031	2034
	G51M30	G48M24	G45M20	G42M16	G40M16/T2	G38M16/T4
Logic industry "Node Range" Labeling (nm)	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
IDM-Foundry node labeling	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
Logic device structure options	FinFET	finFET LGAA	LGAA	LGAA	LGAA-3D	LGAA-3D
Platform device for logic	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
						
LOGIC DEVICE GROUND RULES						
Mx pitch (nm)	36	32	24	20	16	16
M1 pitch (nm)	34	32	23	21	20	19
M0 pitch (nm)	30	24	20	16	16	16
Gate pitch (nm)	51	48	45	42	40	38
Lg: Gate Length - HP (nm)	18	16	14	12	12	12
Lg: Gate Length - HD (nm)	20	18	14	12	12	12
Channel overlap ratio - two-sided	0.20	0.20	0.20	0.20	0.20	0.20
Spacer width (nm)	7	6	5	4	4	4
Contact CD (nm) - finFET, LGAA	19	20	21	22	20	18
Contact CD (nm) - VGAA						
Device architecture key ground rules						
FinFET pitch (nm)	28.0	24.0				
FinFET Fin width (nm)	6.0	5.0				
FinFET Fin height (nm)	50	64				
Footprint drive efficiency - finFET	3.79	5.54				
Lateral GAA lateral pitch (nm)			22.0	20.0	20.0	20.0
Lateral GAA vertical pitch (nm)			18.0	16.0	14.0	14.0
Lateral GAA (nanosheet) thickness (nm)			7.0	6.0	5.0	5.0
Number of vertically stacked nanosheets			3	3	4	4
LGAA width (nm) - HP			30	25	20	15
LGAA width (nm) - HD			15	11	6	6
LGAA width (nm) - SRAM			7	6	6	6
LGAA total height (nm)			53	48	57	57
Footprint drive efficiency - lateral GAA - HP			4.93	4.77	5.88	5.52
Device effective width (nm) - HP	106.0	133.0	222.0	186.0	200.0	160.0
Device effective width (nm) - HD	106.0	133.0	132.0	102.0	88.0	88.0
Device lateral pitch (nm)	28	24	22	20	20	20
Device height (nm)	50.0	64.0	53.0	48.0	57.0	57.0
Device width (nm) - HP	6	5	30	25	20	15

PERFORMANCE BOOSTERS

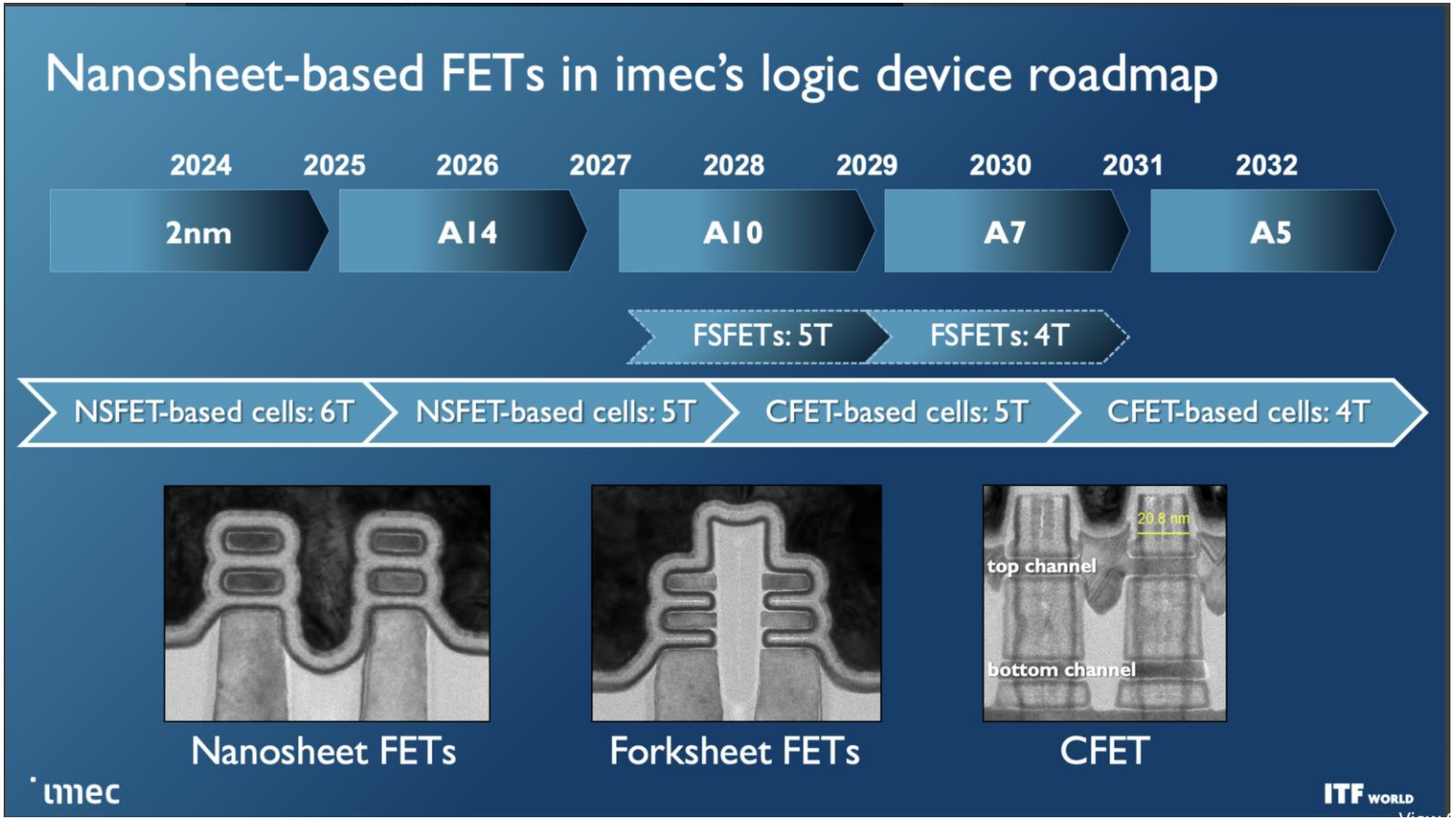
IRDS,2021 Edition More Moore

YEAR OF PRODUCTION	2021	2022	2025	2028	2031	2034
	G51M30	G48M24	G45M20	G42M16	G40M16/T2	G38M16/T4
Logic industry "Node Range" Labeling (nm)	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
IDM-Foundry node labeling	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
Logic device structure options	FinFET	finFET LGAA	LGAA	LGAA	LGAA-3D	LGAA-3D
Platform device for logic	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
Frequency scaling - node-to-node	-	0.02	0.16	0.09	-0.08	-0.01
CPU frequency at constant power density (GHz)	3.13	2.83	3.53	2.50	1.48	0.86
Power at iso frequency - node-to-node	-	-0.16	-0.27	-0.05	-0.06	-0.08
Power density - relative	1.00	1.12	1.04	1.59	2.51	4.27
LOGIC TECHNOLOGY ANCHORS						
Patterning technology inflection for Mx interconnect	193i, EUV DP	193i, EUV DP	193i, EUV DP	193i, High-NA EUV	193i, High-NA EUV	193i, High-NA EUV
Beyond-CMOS as complimentary to platform CMOS	-	-	-	2D Device, FeFET	2D Device, FeFET	2D Device, FeFET
Channel material technology inflection	SiGe25%	SiGe50%	SiGe50%	Ge, 2D Mat	Ge, 2D Mat	Ge, 2D Mat
Process technology inflection	Conformal Doping, Contact	Channel, RMG	Lateral/Atomic Etch	Non-Cu Mx	3DVLSI	3DVLSI
Stacking generation inflection	2D	3D-stacking: W2W, D2W Mem-on-Logic	3D-stacking: W2W, D2W Mem-on-Logic	3D-stacking, Fine-pitch stacking, P-over-N, Mem-on-Logic	3D-stacking, 3DVLSI: Mem-on-Logic with Interconnect	3D-stacking, 3DVLSI: Logic-on-Logic

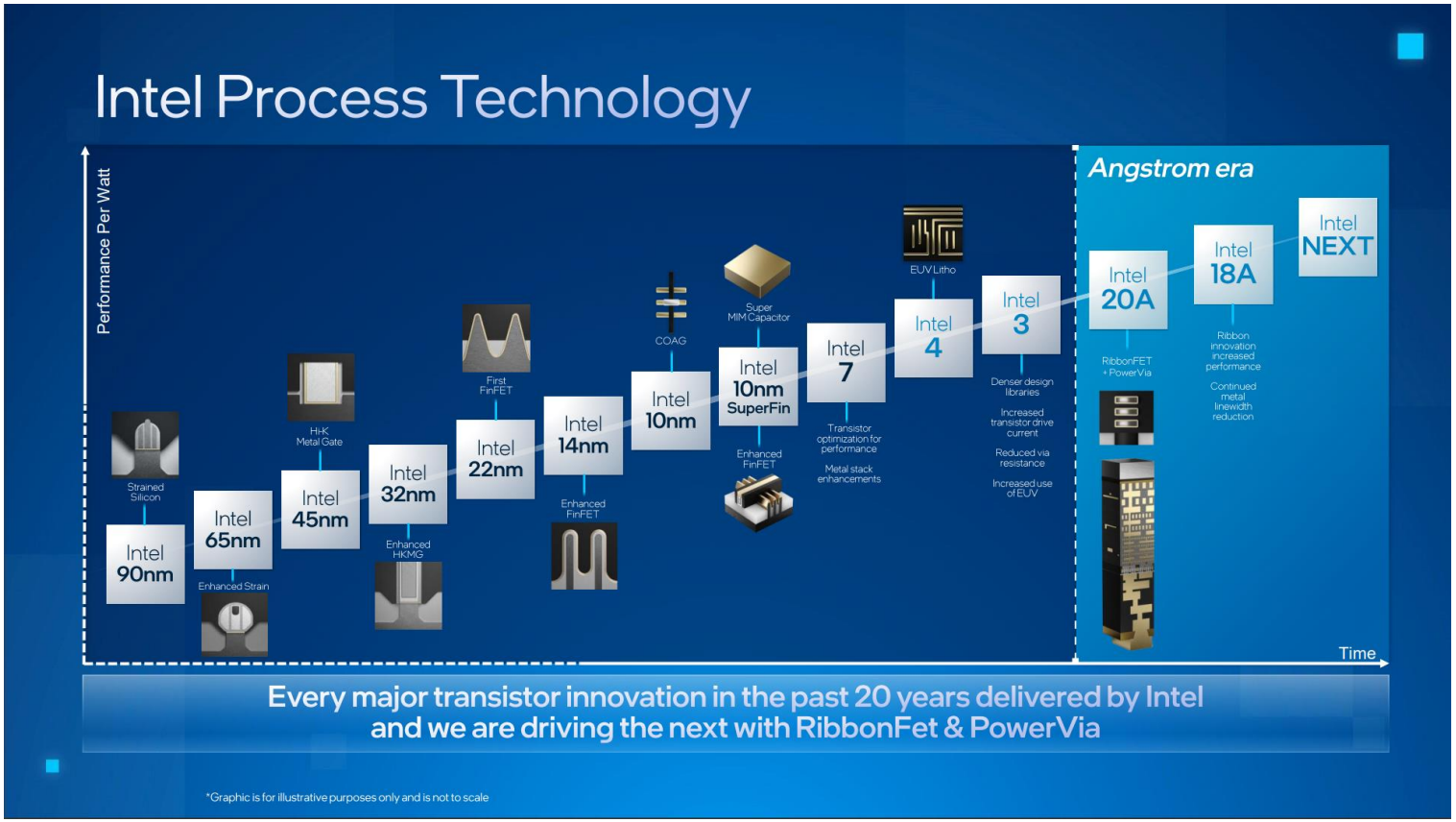
- *III V is gone*
- *FeFET means NCFET for SS < 60mV/dec.*
- *2D/FeFET starts at1.5 nm node???*

IMEC 2022

- CFET (transistor stacking, nFET on pFET or p on n) starts at A5 (0.5nm node)
- Atomic (2D) in A2 node (0.2 nm node)

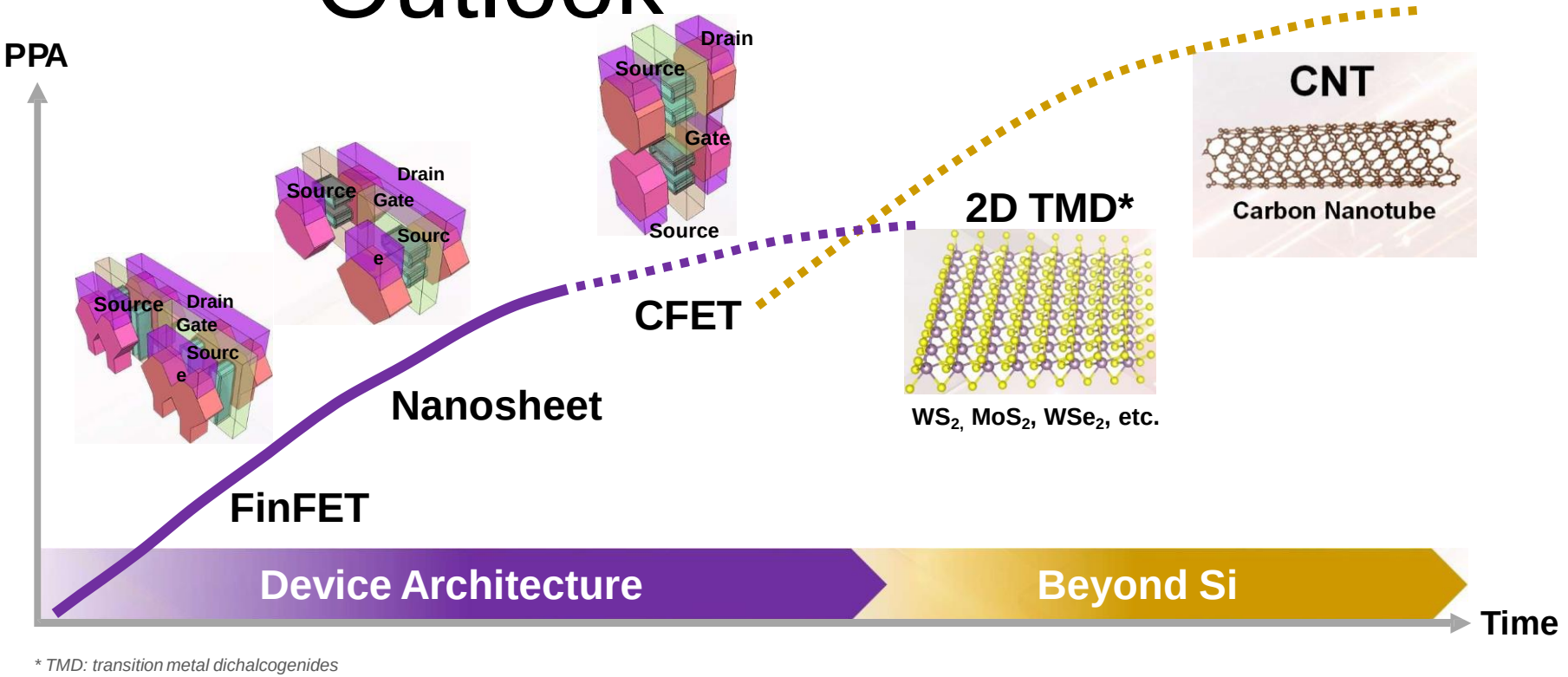


- **Gate All Around (GAA)/Nanosheet devices debut in 2024 with the 2nm node, replacing FinFETs**



- **Intel 20A & 18A(Angstrom Era) → 1st RibbonFET(4 stacked channels) + Power Via**
- **TSMC GAA 2nm/Samsung GAA 3nm**

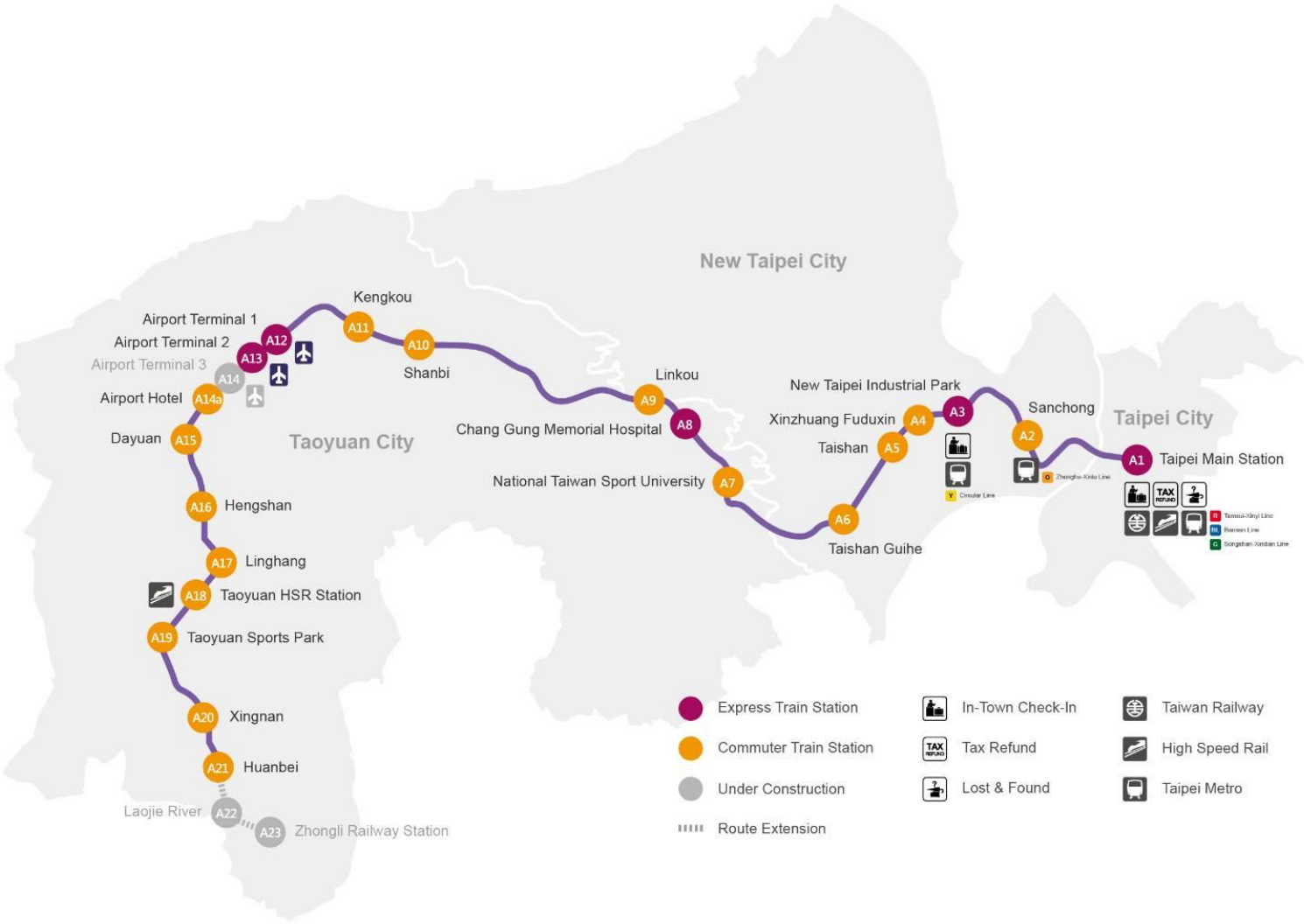
Device Architecture Outlook



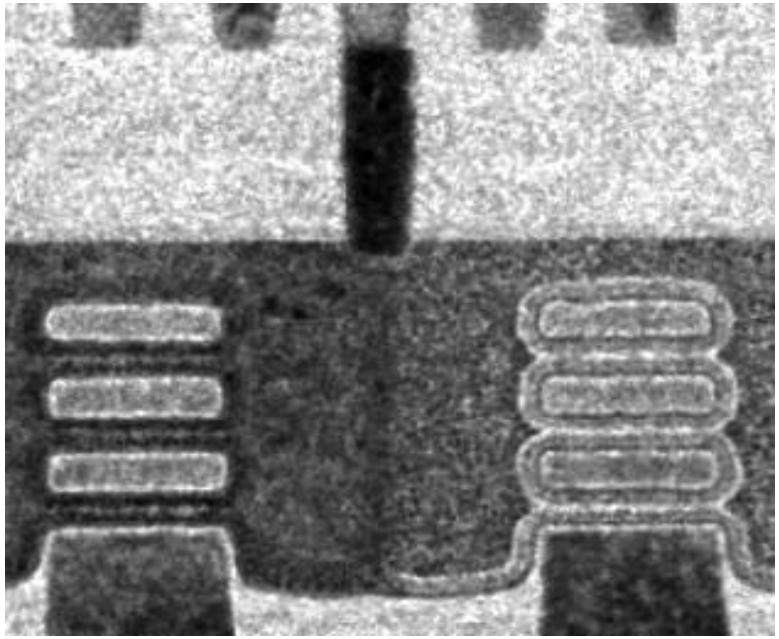
** TMD: transition metal dichalcogenides*

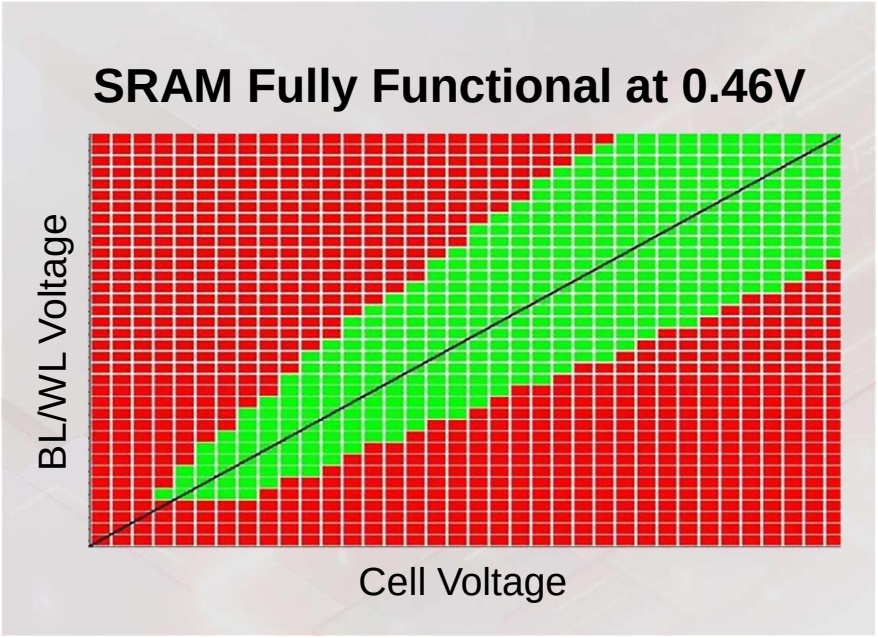
- **CFET after nanosheets**
- **2D after CFETs**

Roadmap of TW Government



Nanosheet Transistor

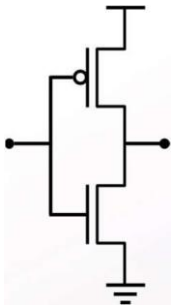


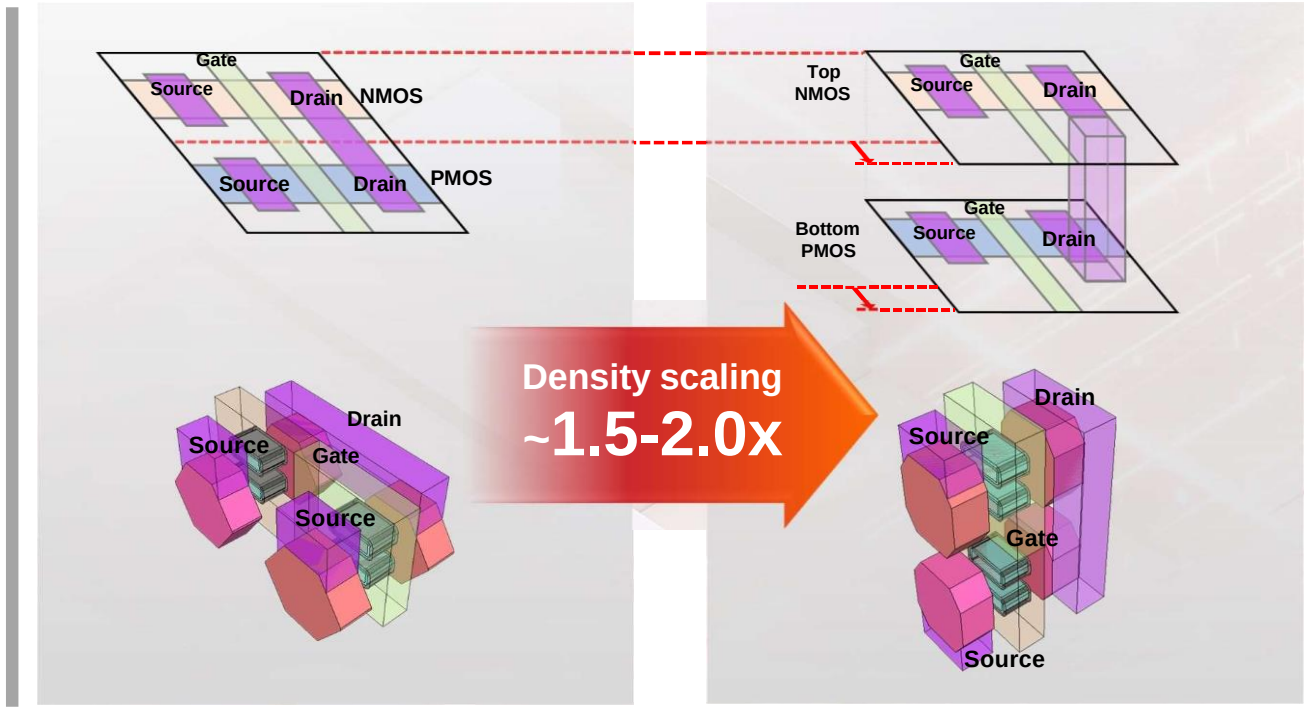


- Tight space sheet reduces para C.
- N/P are different
- Smaller SS and better short channel control to achieve low Vdd and low power.

CFET Density Scaling

Inverter





Nanosheet **CFET**

CFET Process Challenges

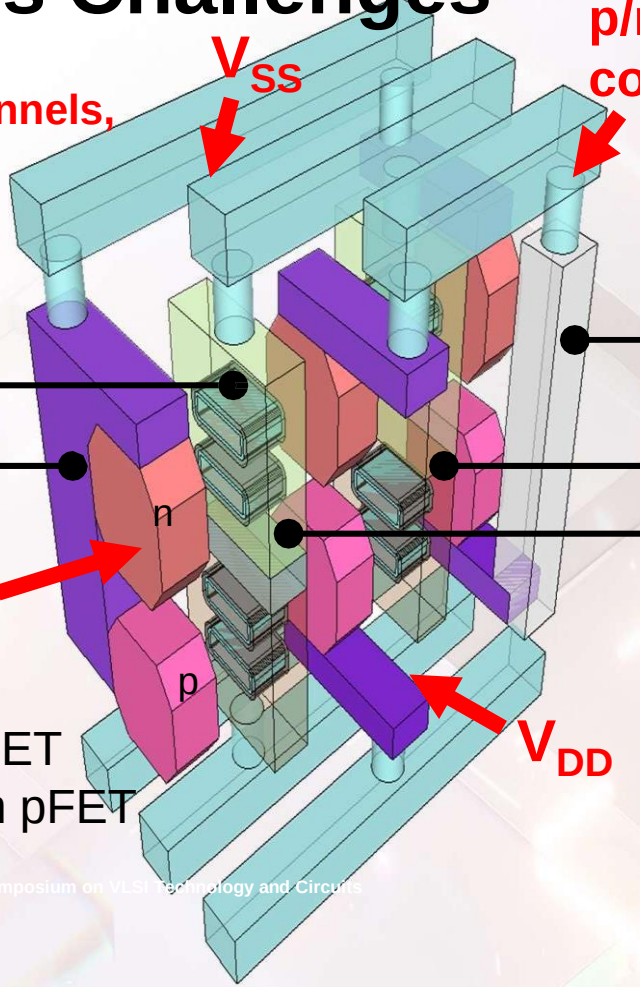
- **Fin height including n channels, p channels, and SLs**

High aspect ratio etch

(Same as nanosheet)

Low resistance metal gap-fill

(dielectric isolation)



Top nFET
Bottom pFET

p/nFET drain connection

High AR & Low-R metal gap-fill

- **Vertically stacked dual metal gate process is needed. (recess and deposition)**

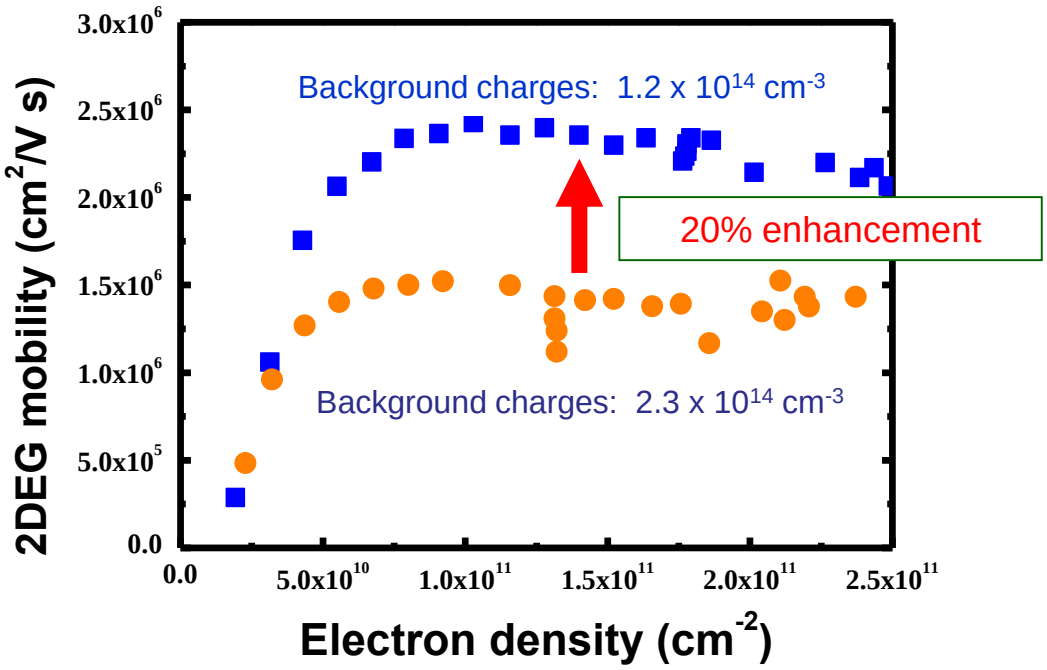
Multi-Vt for both n/pMOS

(independent gate control)

Gate isolation

- **Creating isolation between NMOS/PMOS gate**

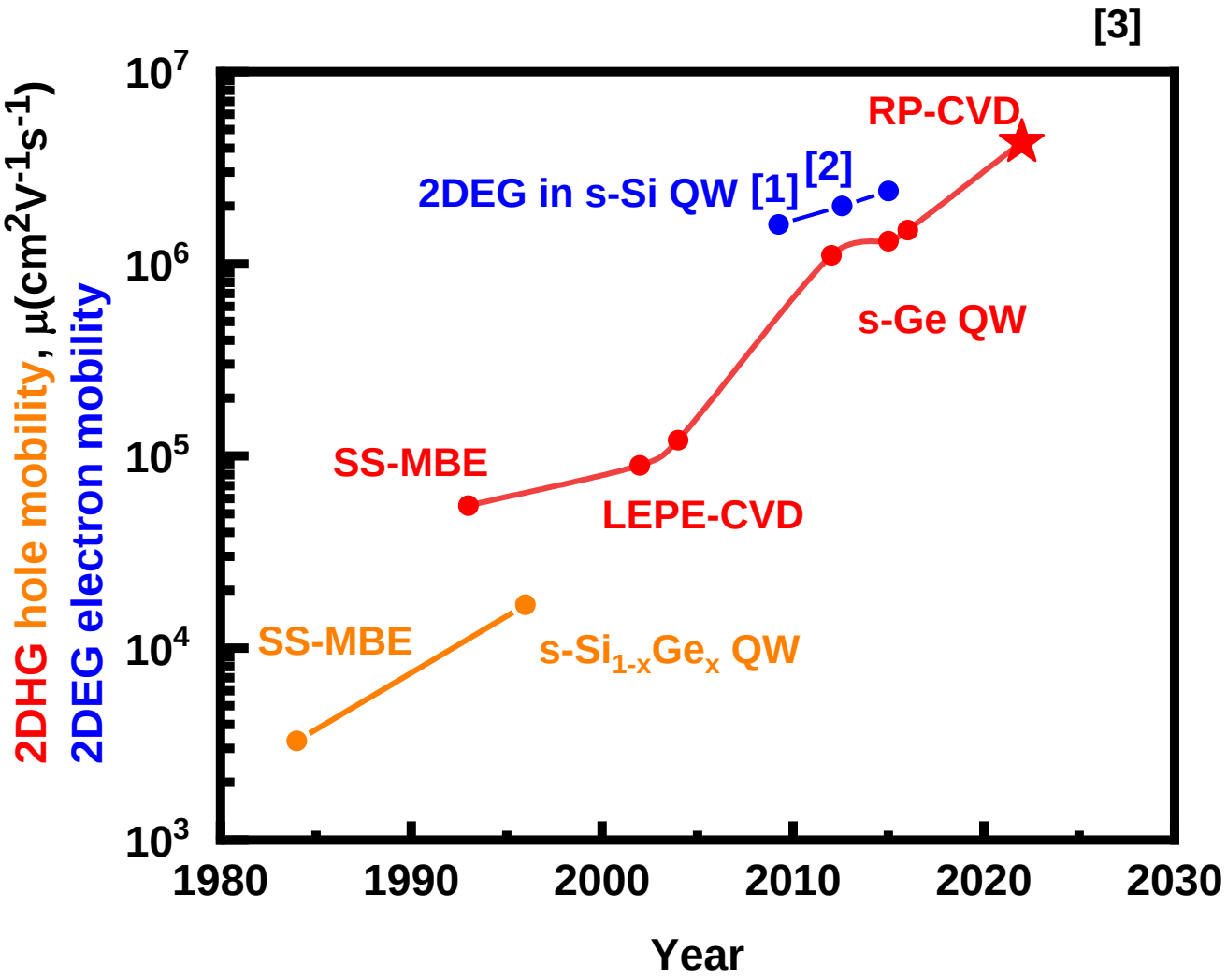
Record High 2DEG Mobility (2.4 M $cm^2/V\,s$)



M. Yu. Melnikov et al., APL, 106, 092102 (2015).

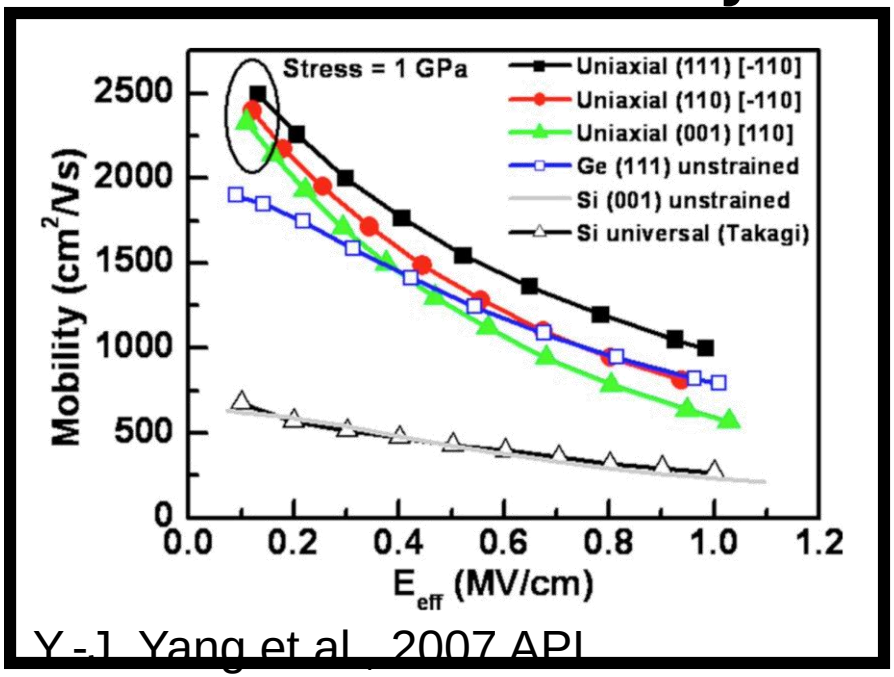
Collaborate with Prof. S. V. Kravchenko, Northeastern Univ.

- 2.4 x 10^6 $cm^2/V\,s$ by the reduction of background charges from 2.3 x 10^{14} cm^{-3} to 1.2 x 10^{14} cm^{-3}.



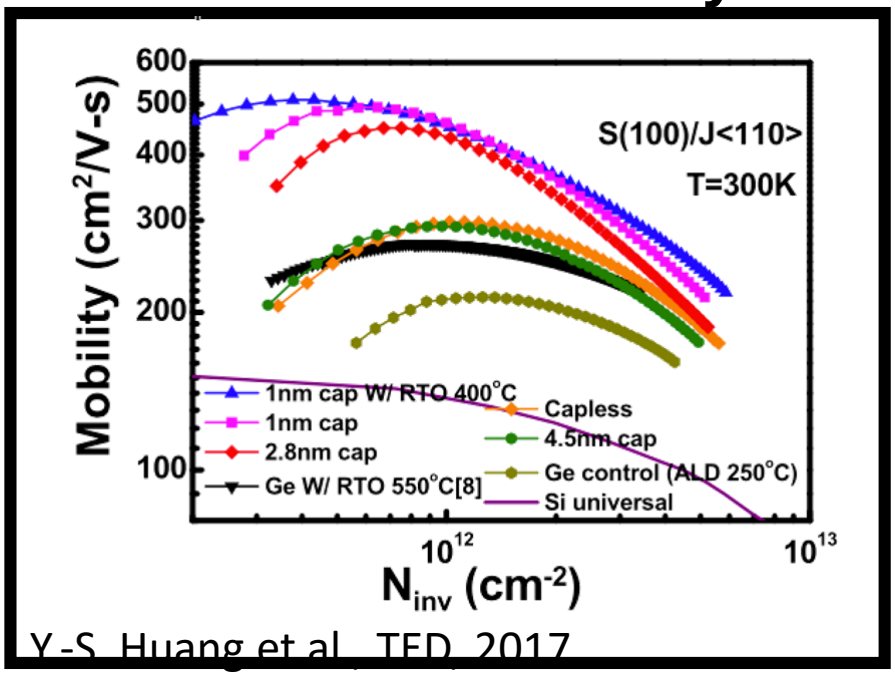
[1] T. M. Lu, D. C. Tsui, C.-H. Lee,2and C. W. Liu
[2] S.-H. Huang, T.-M. Lu, S.-C. Lu, C.-H. Lee, C. W. Liu, and D. C. Tsui, APPLIED PHYSICS LETTERS 101, 042111 (2012)
[3] Maksym Myronov, Jan Kycia, Philip Waldron, Weihong Jiang, Pedro Barrios, Alex Bogan, Peter Coleridge, and Sergei Studenikin, Small Sci. 2023, 2200094

Ge electron mobility



Y.-J. Yang et al., 2007 APL

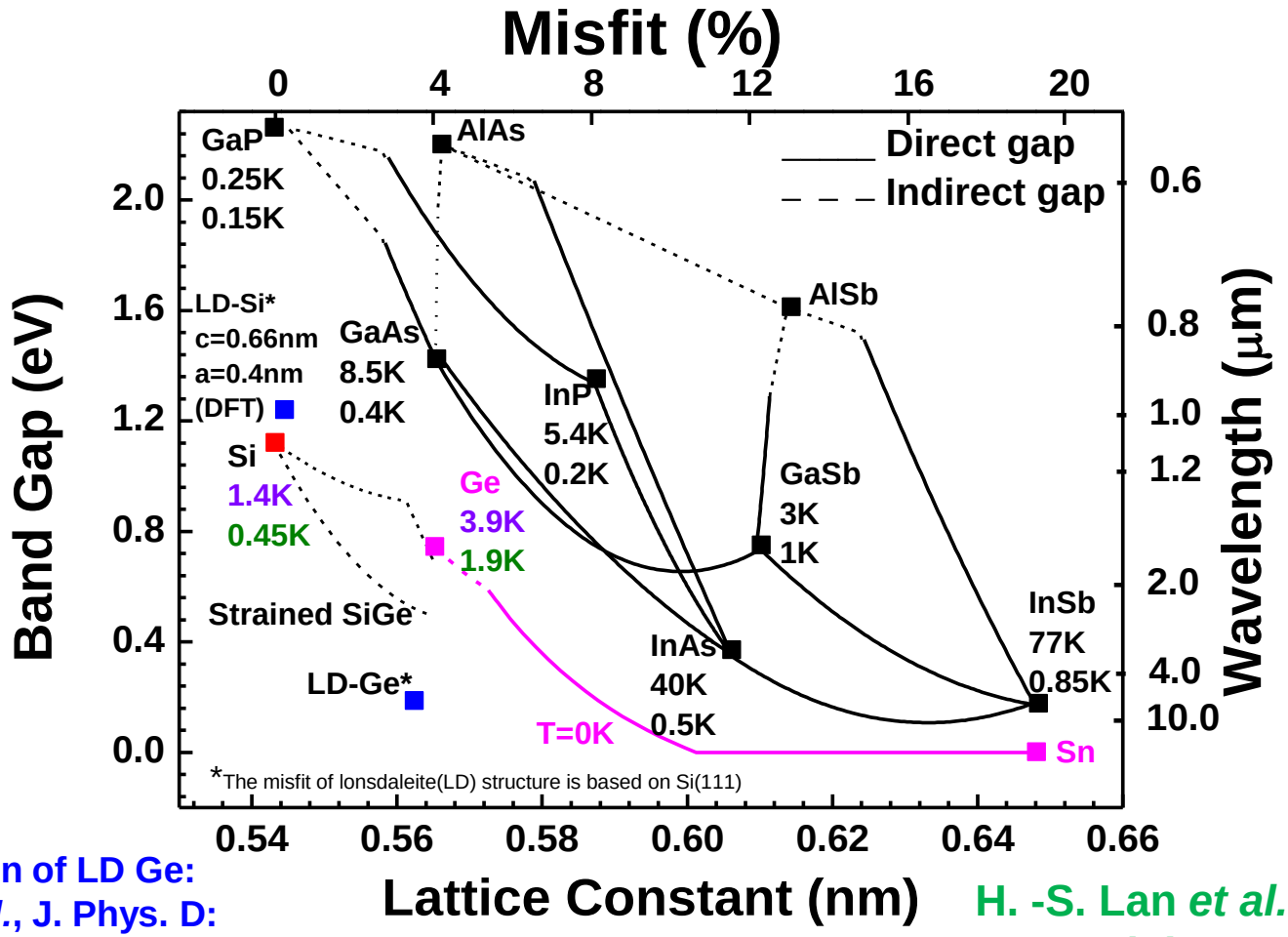
GeSn hole mobility



Y.-S. Huang et al., TED, 2017

- **Ge has higher electron mobility than Si.**
- **GeSn/Ge has higher hole mobility than Si.**

Something new about group IV



Band calculation of LD Ge:
P. -S. Chen *et al.*, J. Phys. D: Appl. Phys. 50, 015107, 2017.

H. -S. Lan *et al.*, PRB 95, 201201(R), 2017

- Ge has high electron and hole mobility but large misfit.
- y > 0.41 ➔ $Ge_{1-y}Sn_y$ becomes topological semimetal/ zero bandgap semiconductor (E_g=0)
- Hexagonal Si/Ge is possible with new properties

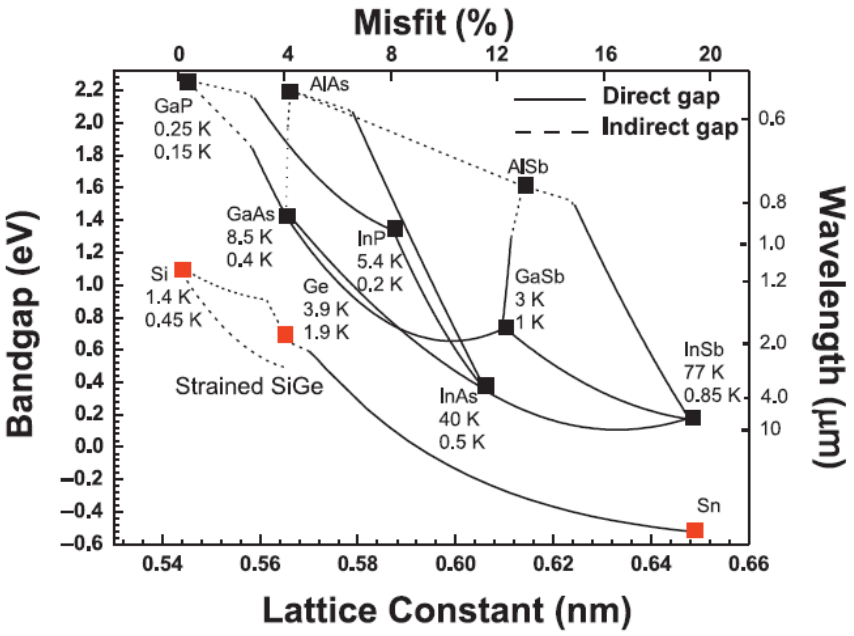
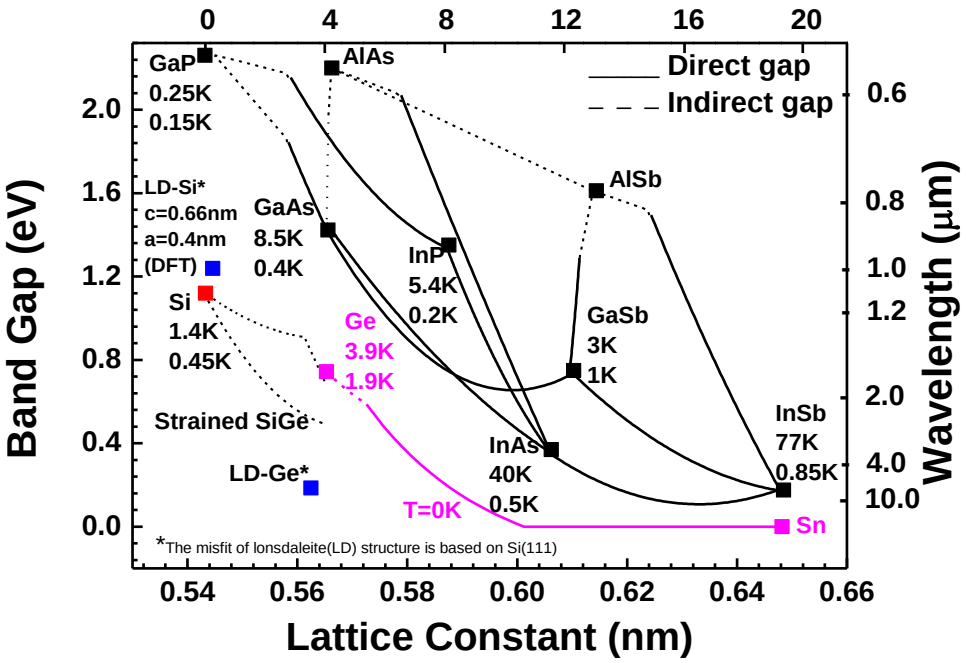


Figure 1. The plot of bandgap energy/wavelength versus lattice constant/misfit for III–V, Si/Ge/Sn. The numbers below the chemical symbols are electron and hole mobilities in units of cm^2/Vs. The solid lines indicate direct bandgaps.[25] (The red symbols are column IV elements.)



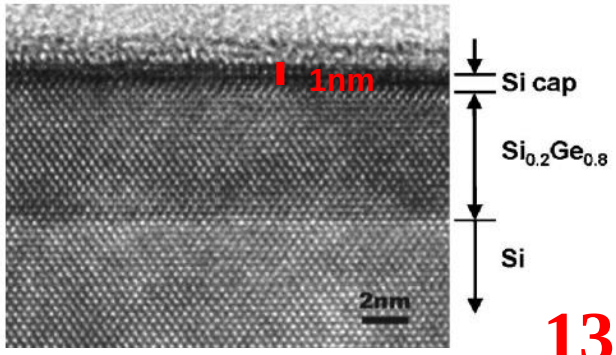
Lan *et al.*, PRB 95, 201201(R), 2017

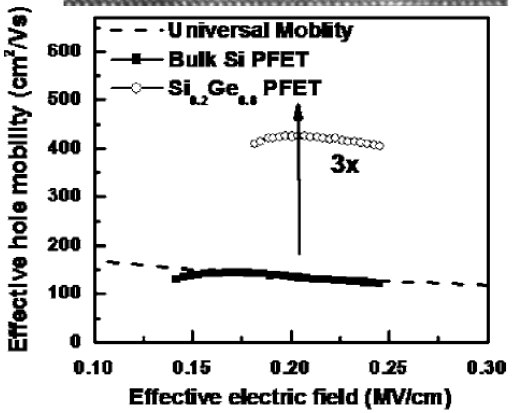
"New Materials for Post-Si Computing", MRS Bulletin, Vol. 39, No. 8, pp. 658-662, August, 2014

- Sn is zero bandgap semiconductor due to band inversion.
- y>0.41 → $Ge_{1-y}Sn_y$ becomes zero bandgap semiconductor (E_g=0)

MOST advanced FinFET on the market

2007 APL

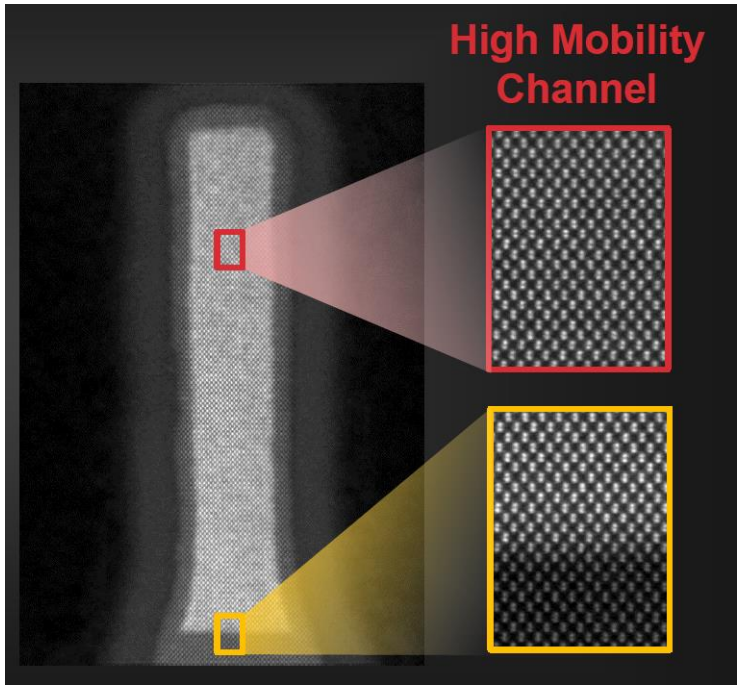




13 years

(2007-2020)

2021 TSMC ISSCC



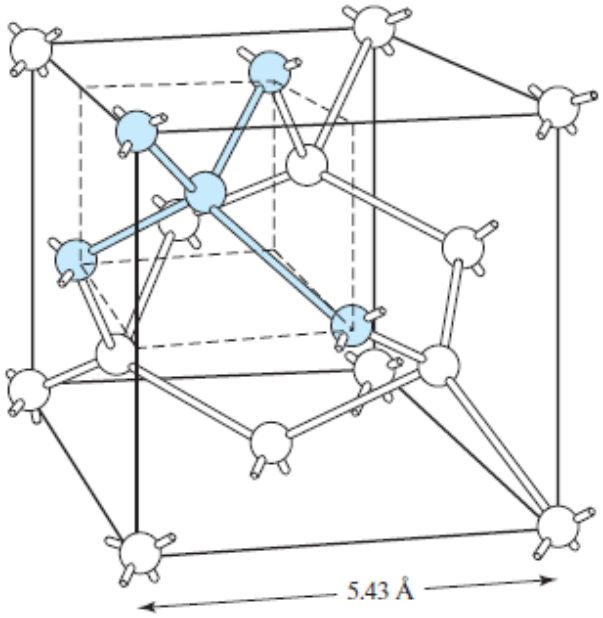
Mark Liu, Plenary Session 1.1, ISSCC 2021.

C.-Y. Peng, F. Yuan, C.-Y. Yu, P.-S. Kuo, M. H. Lee, S. Maikap, C.-H. Hsu, and C. W. Liu, Appl. Phys. Lett. 90, 012114 (2007) .(Cited No./Self Cited No.= 30/ 7)

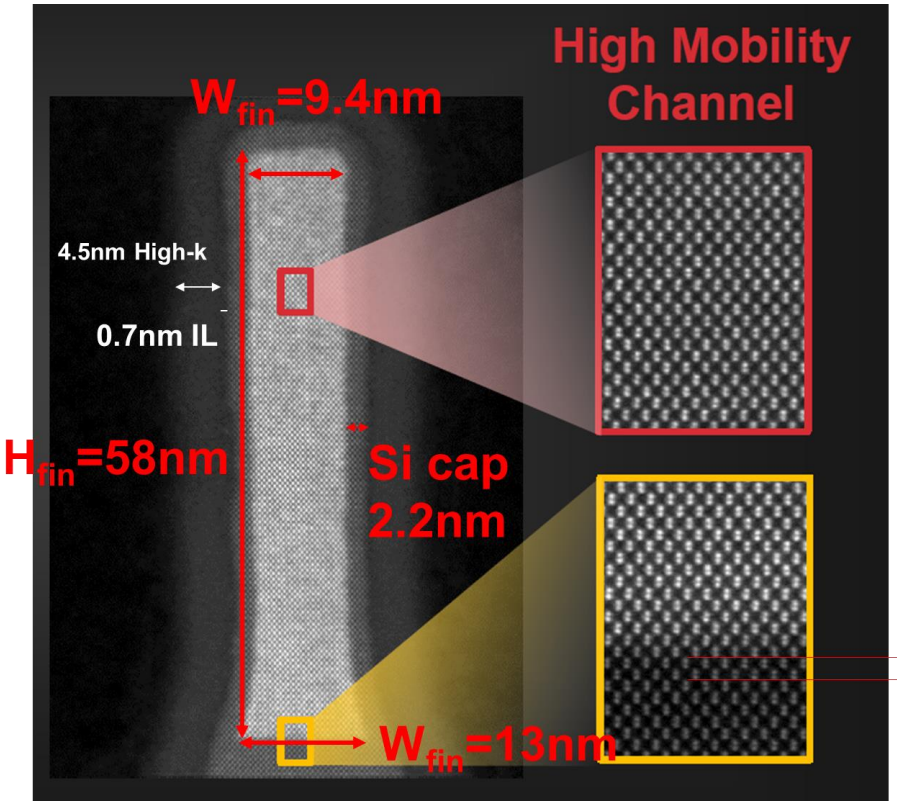
- **$Si_{0.2}Ge_{0.8}$ channel with 1nm Si cap, having higher mobility(3X) than Si.**
- **It takes 13 years to commercialize HMC.**

Silicon Crystal Structure (Ge, SiGe, GeSn)

- ***Unit cell*** of silicon crystal is cubic.

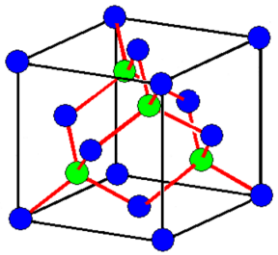


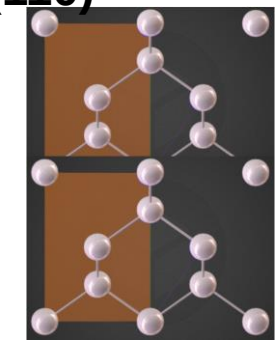
- ***Each Si atom has 4 nearest neighbors*.**
- **Si sp3**
- **Conduction band is ?**
- **Valence band is ?**

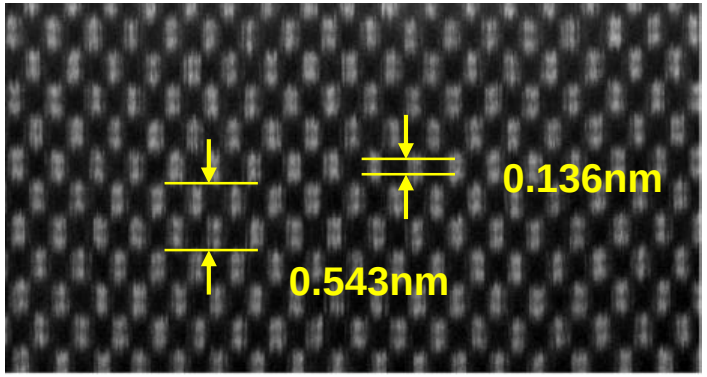


Ref: Mark Liu, Plenary Session 1.1, ISSCC 2021.

(110)



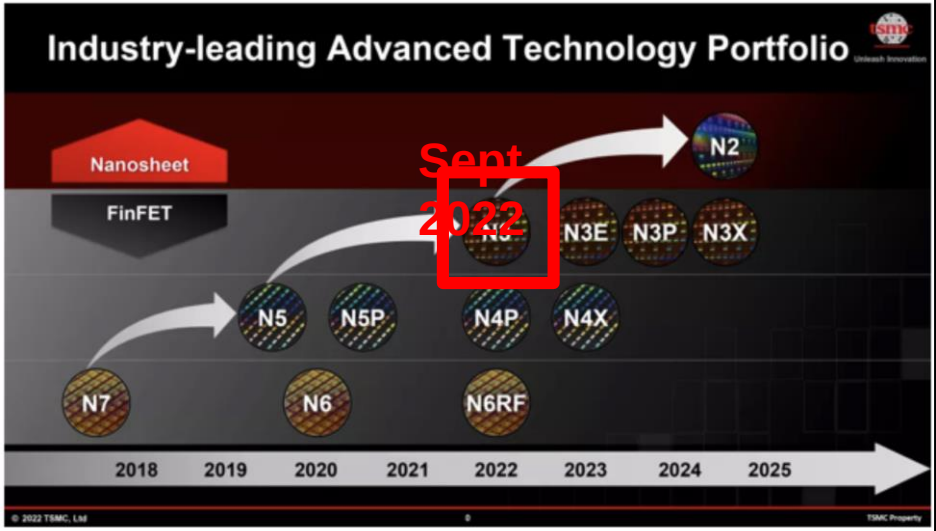




- The distance between the Si atoms is 0.136 nm.
- Dumbells

Ref: C.-N. Hsiao et al., Nanoscale Research Letters, 2014.

N5 vs N3

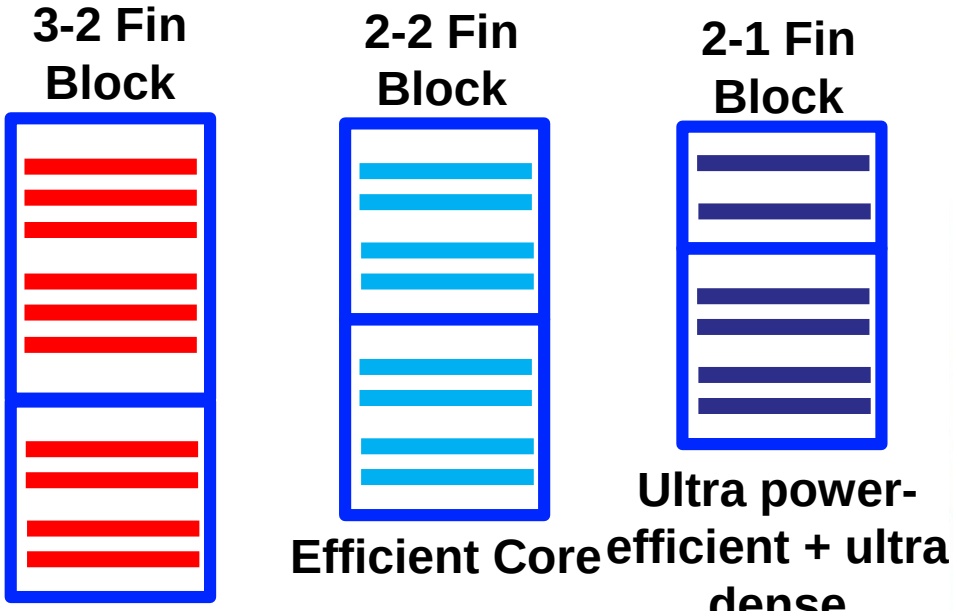


(Image credit: TSMC)

	N3E vs N5	N3 vs N5
Speed Improvement @ Same Power	+18%	+10% ~ 15%
Power Reduction @ Same Speed	-34%	-25% ~ -30%
Logic Density	1.7x	1.6x
HVM Start	Q2/Q3 2023	H2 2022

- **HVP N3 planned to be start at march 2022 ➔ delayed to sept 2022 ➔ iPhone 14 is not having N3.**
- **N3 node has a narrow process window ➔ expected to have lower yield.**
- **N3E ➔ have higher process window ➔ higher yield.**

FinFlex: N3 Node(2)



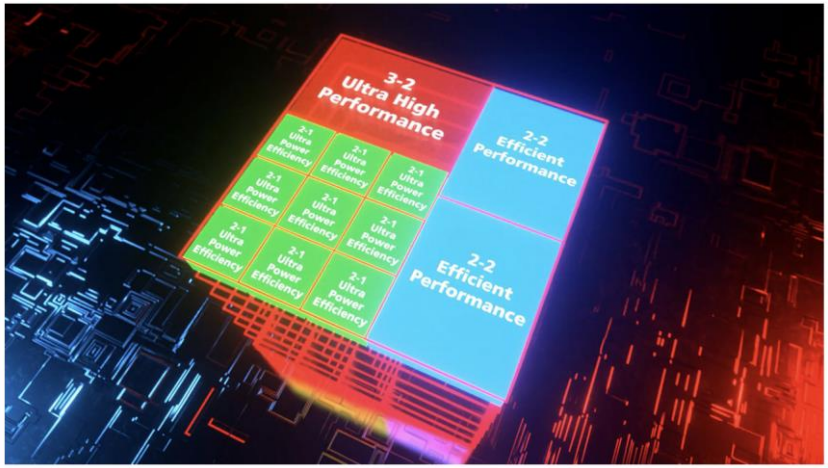
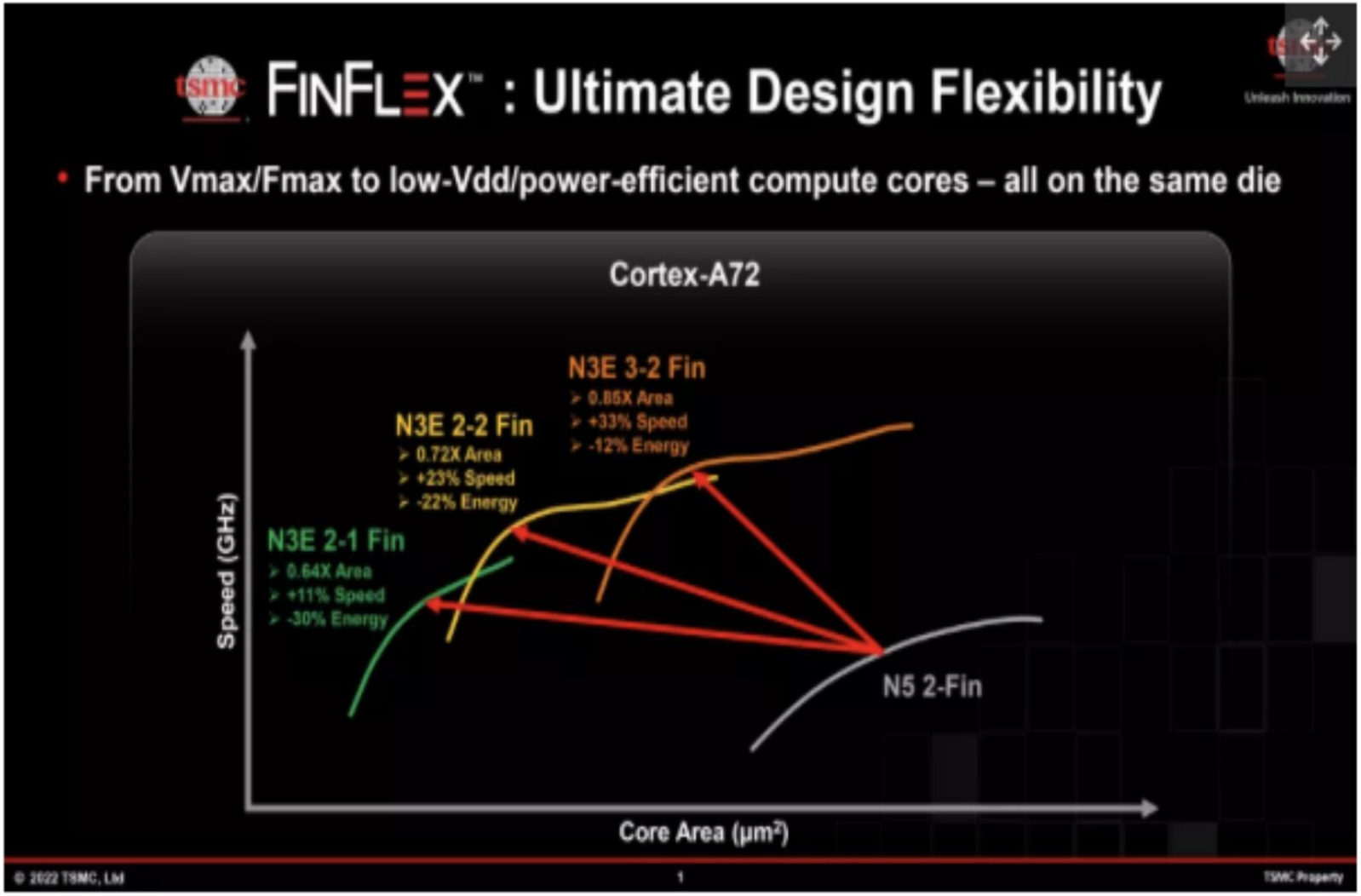


Diagram 2: N3 with FinFlex enables a designer to choose the ideal FIN configuration for each functional block on a chip.

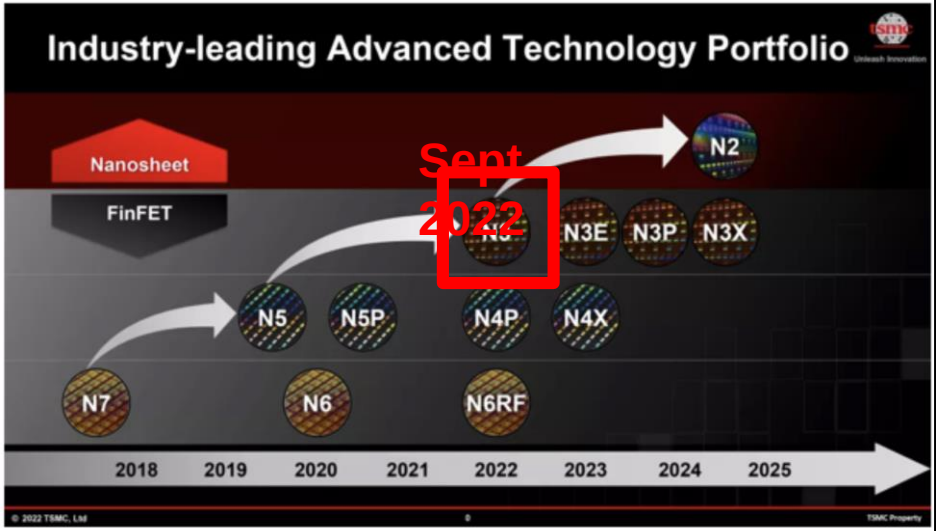
- **3-2 FIN** – Fastest clock frequencies and highest performance for the most demanding compute needs
- **2-2 FIN** – Efficient Performance, a good balance between performance, power efficiency and density
- **2-1 FIN** – Ultra Power Efficiency, lowest power consumption, lowest leakage and highest density
- **Finflex offers high-performance and power-efficient cores on the same die.**
- **All N3 nodes (N3, N3E, N3P and, N3X) will have the FinFlex capability.**

FinFlex: N3 Node(1)



(Image credit: TSMC)

N5 vs N3



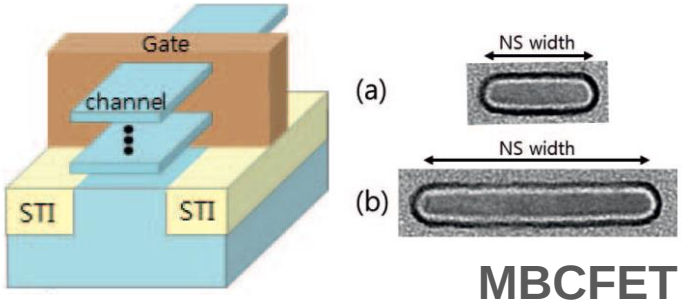
(Image credit: TSMC)

	N3E vs N5	N3 vs N5
Speed Improvement @ Same Power	+18%	+10% ~ 15%
Power Reduction @ Same Speed	-34%	-25% ~ -30%
Logic Density	1.7x	1.6x
HVM Start	Q2/Q3 2023	H2 2022

- **HVP N3 planned to be start at march 2022 ➔ delayed to sept 2022 ➔ iPhone 14 is not having N3.**
- **N3 node has a narrow process window ➔ expected to have lower yield.**
- **N3E ➔ lower transistor density to have higher process window ➔ higher yield.**

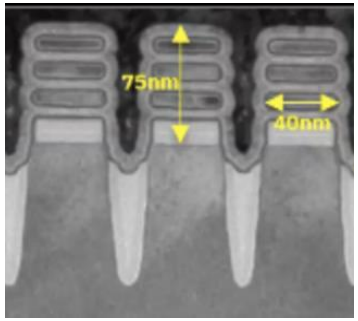
Status Quo of Industry for 3/2nm

2018 IEDM Samsung 3nm (?)



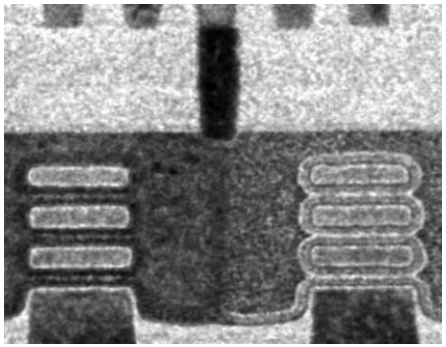
MBC: Multi-Bridge-Channel.

2021 IBM 2nm (3)



IBM 2nm technology node,
3 stacked nanosheets.

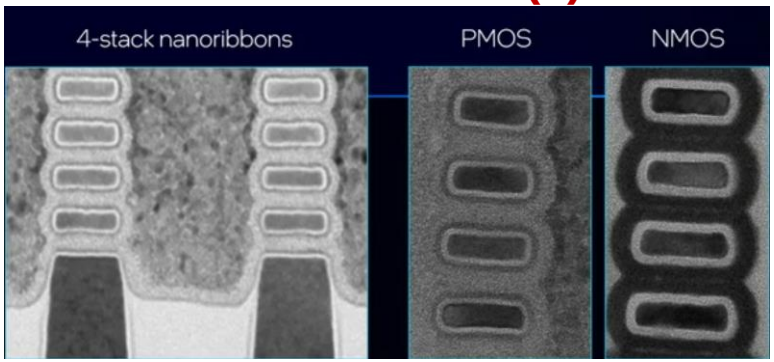
2021 ISSCC TSMC 2nm (3)



TSMC 2nm technology node,
3 stacked nanosheets.

Mark Liu, Plenary Session 1.1, ISSCC 2021.

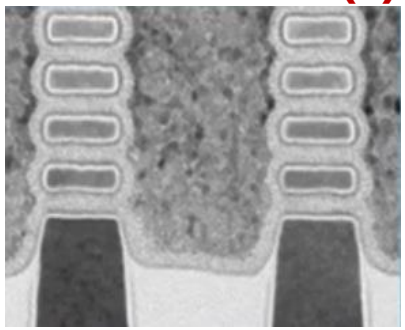
2024 Intel 20A (4)



RibbonFET, 4 stacked nanoribbons

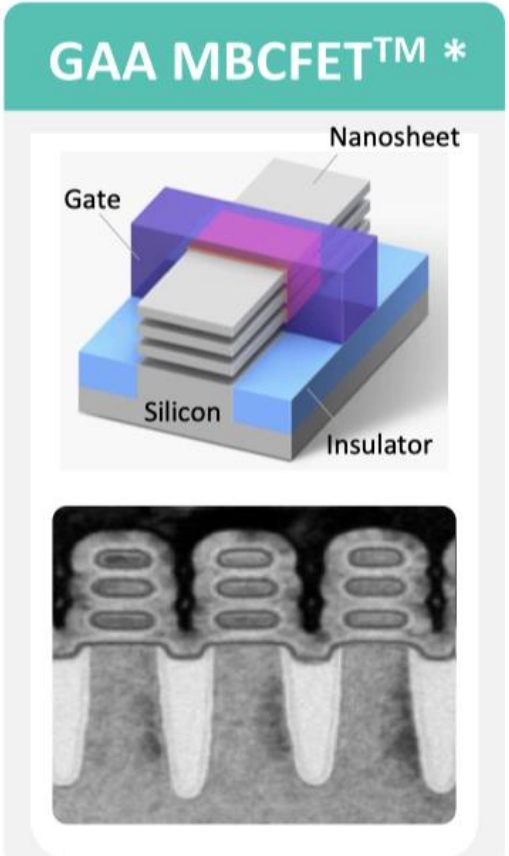
Ref: Intel Accelerated, July 26, 2021.

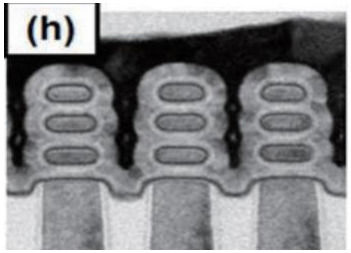
2025 Intel 18A (4)



4 stacked nanoribbons
(In early development)

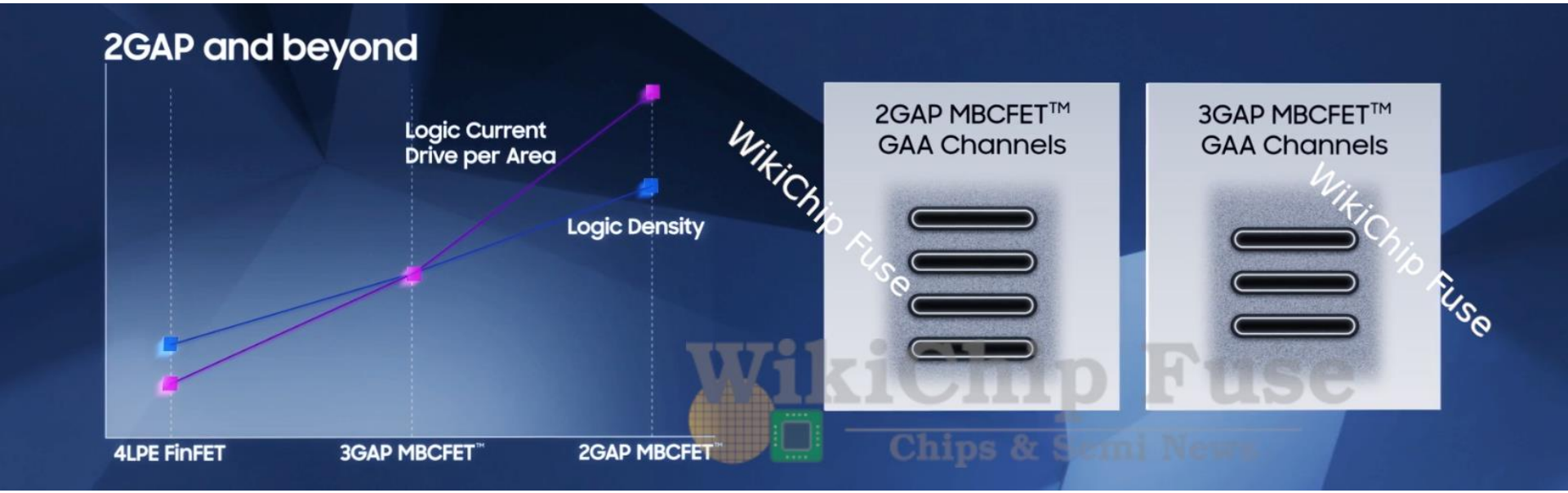
Samsung GAA MBCFET





Ref: N. Loubet et al., VLSI 2017, IBM, Samsung, GF

Ref: IEDM 2022 SC1 Samsung

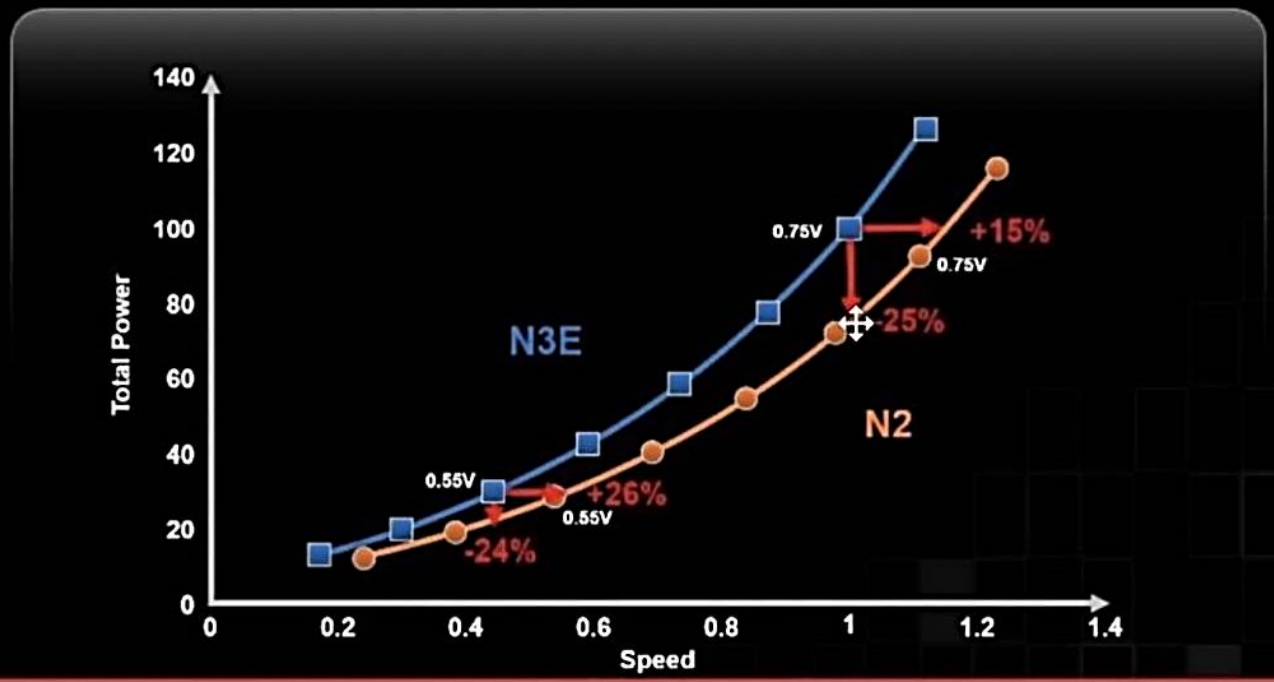


2-nanometer node is called 2GAP and is planned to enter mass production in 2025

Samsung says :

- **2GAP will include pitch scaling for improved transistor density.**
- **2GAP will add another nanosheet for a total of four stacked nanosheets in order to improve drive current.**

N2 Offers Full-node Performance and Power Benefit



2奈米之後的技術創新與方向

- 在新型電晶體結構、新材料、持續微縮和新導體材料方面出現的創新，標準半導體架構的演變從平面式電晶體轉至鰭式場效電晶體(FinFET)，並將再次進展到奈米片電晶體。
同時也期待在2D材料、1D奈米碳管方面的突破，在不斷微縮的同時，克服晶片行動性方面的挑戰。

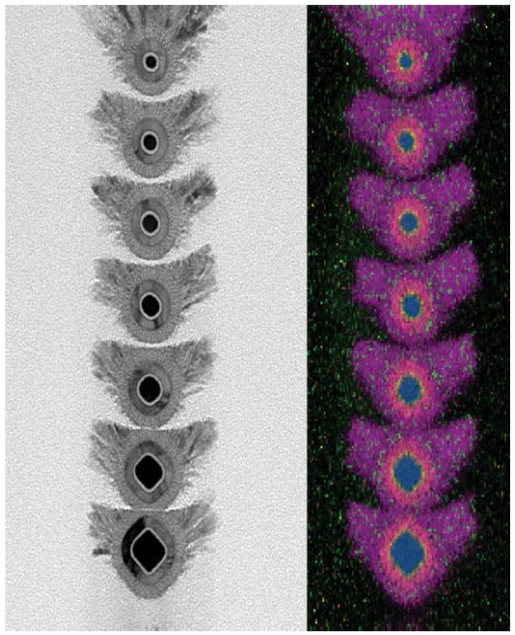
- **在2D材料中，我們的鉍(Bi)接觸二硫化鉬(MoS2)裝置實現了創紀錄的低接觸電阻，比過去2D材料提報的數據技術好上5倍，並產生了高通態電流。展望未來，台積電將繼續探索電晶體架構，並利用2D材料和碳奈米管等新材料。**

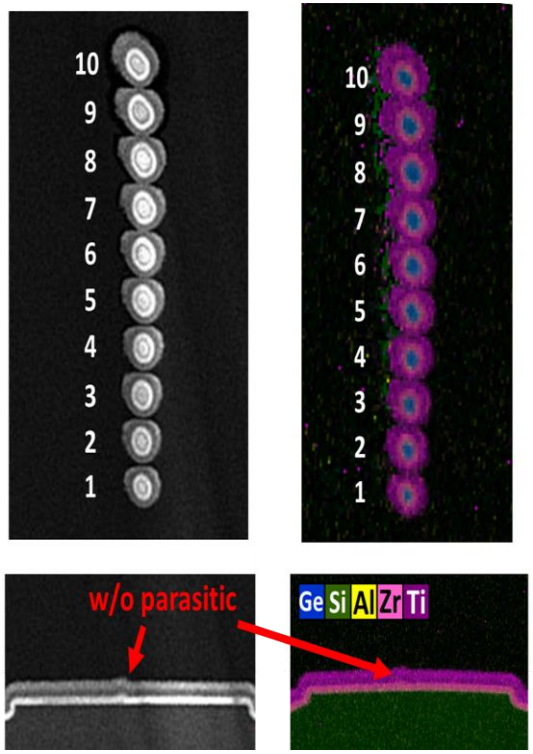
Source :tsmc

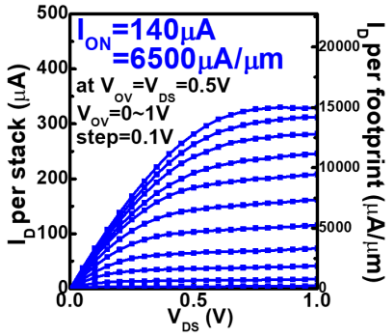
Beyond 3 nm

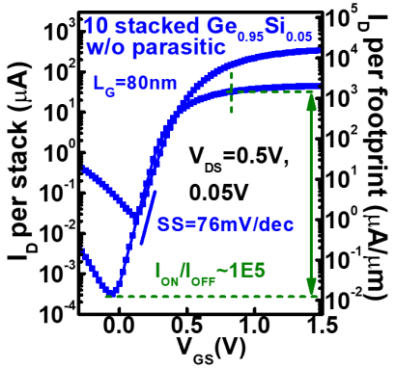
Post the 2021 Nature Electronics Research Highlight (July, 2021)/ VLSI 2021 highlight paper

7 stacked GeSi nanowires







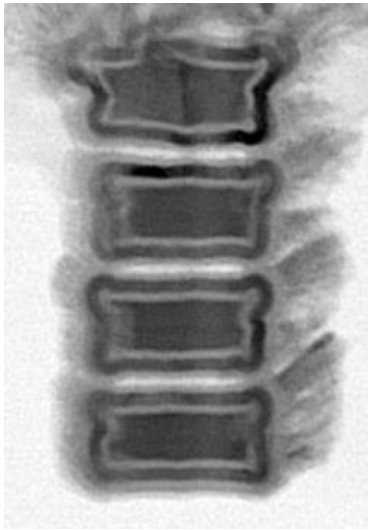


Nature Electronics, Vol. 4, July 2021, 452.

- **10 stacked $Ge_{0.95}Si_{0.05}$ nanowires without parasitic channels**
- **I_{ON} = 6500 μA/μm at V_{OV} = V_{DS} = 0.5V**
- **SS = 76 mV/dec and I_{ON}/I_{OFF}～1E5**

(Unpublished)

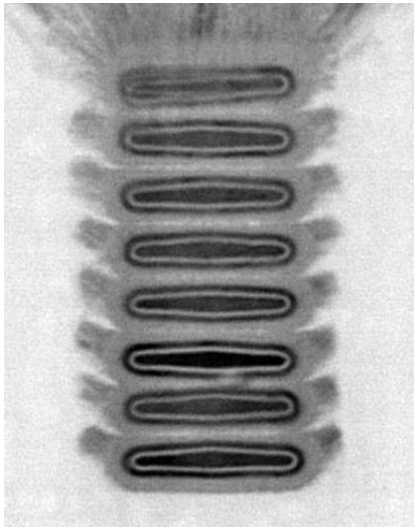
NTU 2020 IEDM
4 stacked GeSi



NTU 4 stacked nanosheets

Yu-Shiang Huang et al. and C. W. Liu, International Electron Devices Meeting (IEDM), 2020.

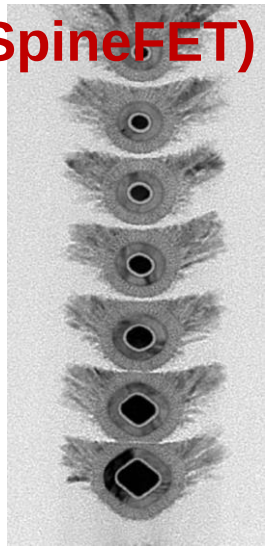
NTU 2021 VLSI
8 stacked GeSi



NTU 8 stacked nanosheets

Yi-Chun Liu et al. and C. W. Liu, Symposium on VLSI Technology (VLSI-Technology), 2021.

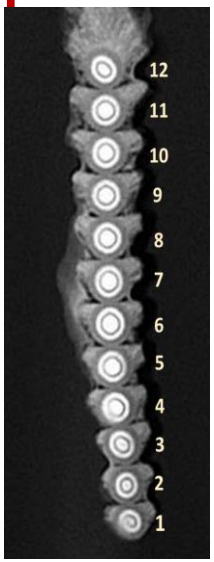
NTU 2021
7 stacked GeSi
(SpineFET)



NTU 7 stacked nanowires

Yi-Chun Liu et al. and C. W. Liu, Symposium on VLSI Technology (VLSI-Technology), 2021.

12 stacked GeSi
(SpineFET)

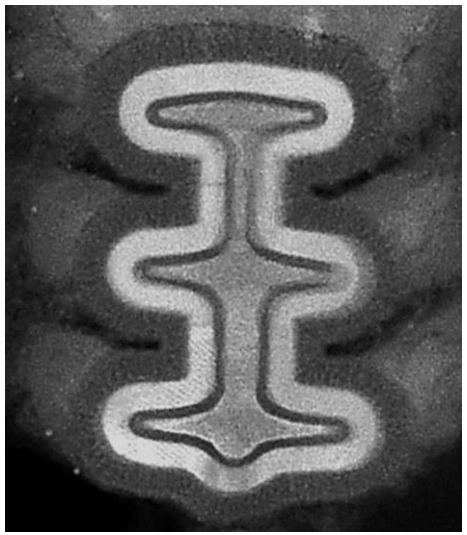


NTU 12 stacked nanowires

16 stacked GeSi
(SpineFET)

NTU 16 stacked nanowires

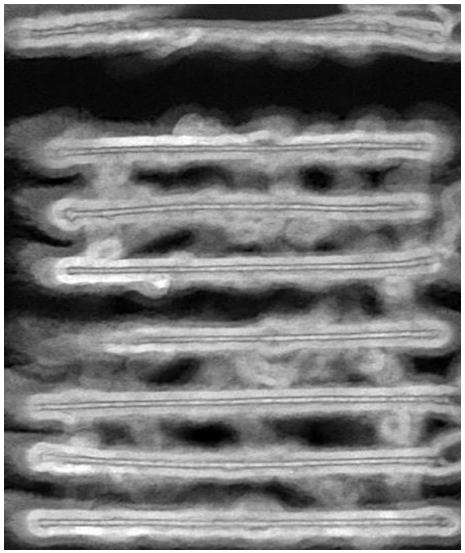
NTU 2021
3 stacked GeSi
{110} TreeFET (王)



NTU 3 stacked TreeFET

Chien-Te Tu et al. and C. W. Liu, EDL2022

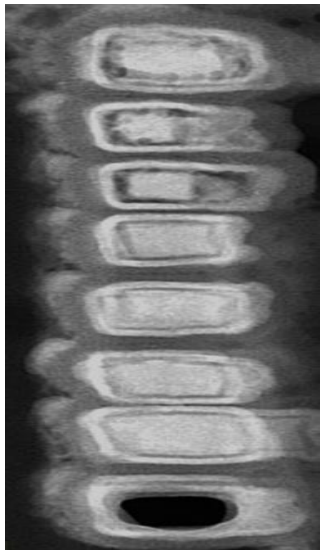
NTU 2021 (2D)
8 stacked GeSn
Ultrathin bodies



NTU 8 stacked ultrathin bodies

Chung-En Tsai et al. and C. W. Liu, International Electron Devices Meeting (IEDM), 2021.

NTU 2021
8 stacked GeSn
Nanosheets

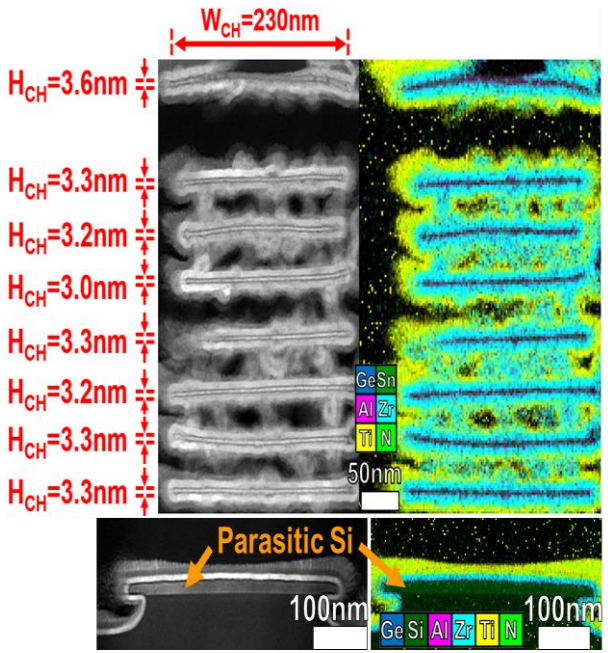


NTU 8 stacked nanosheets

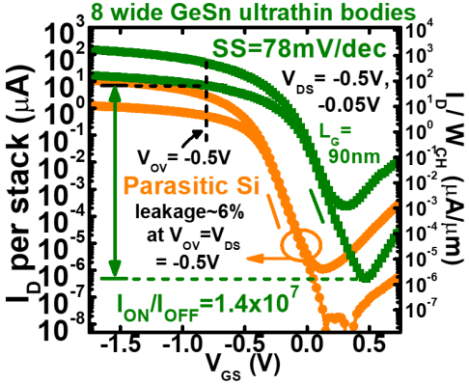
Chung-En Tsai et al. and C. W. Liu, International Electron Devices Meeting (IEDM), 2021.

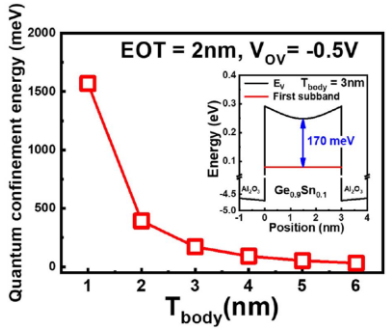
2021 IEDM Roger A. Haken Best Student Paper Award

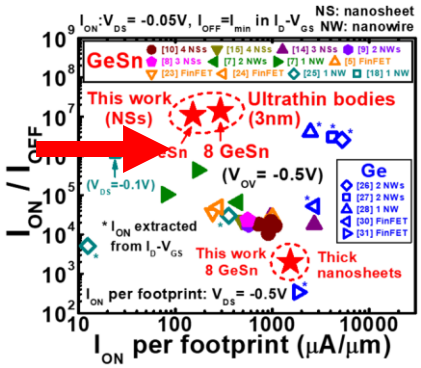
"Highly Stacked GeSn Nanosheets by CVD Epitaxy and Highly Selective Isotropic Dry Etching," IEEE Transactions on Electron Devices, 2022.

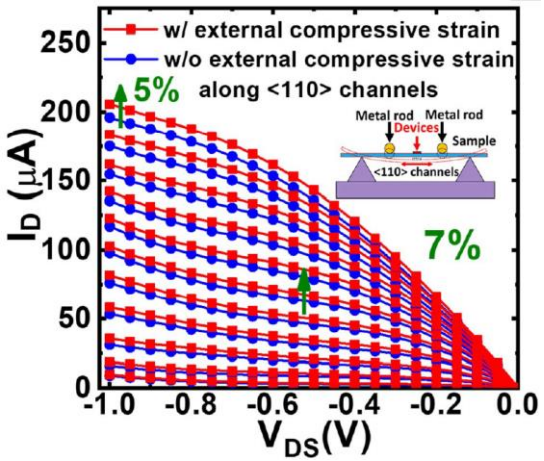


C.-E. Tsai et al. and C. W. Liu, IEDM, 2021.





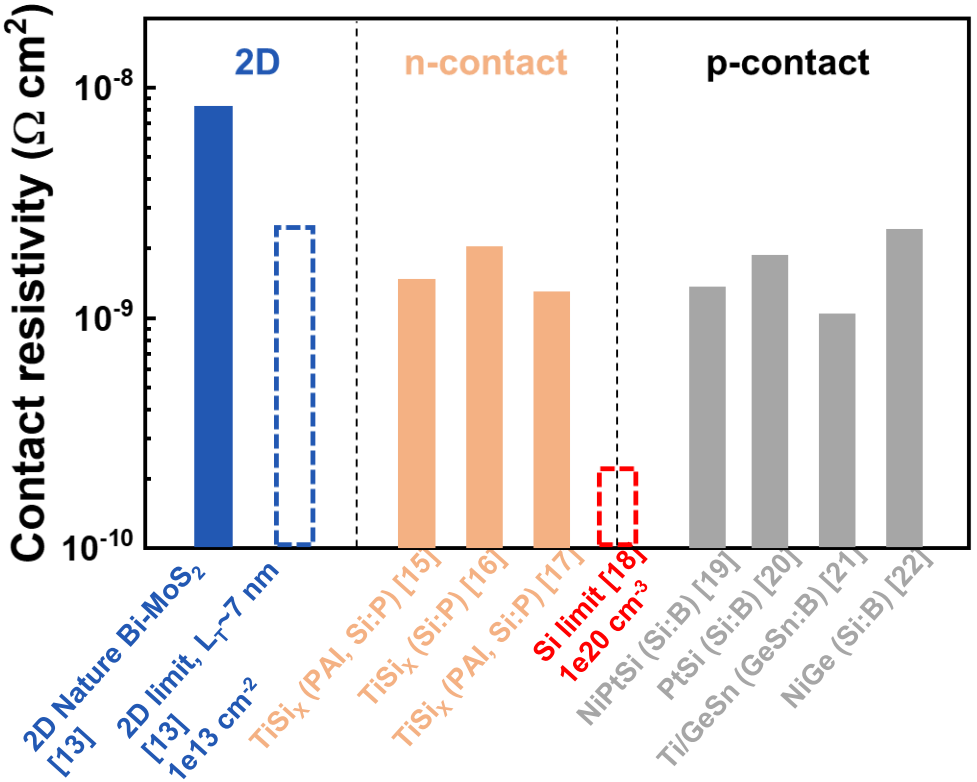




B.-W. Huang et al. and C. W. Liu, accepted by TED, 2022.

- The 8 stacked GeSn ultrathin bodies (~3nm) have record I_{ON}/I_{OFF} of $1.4x10^7$ at V_{DS}= -0.05V among GeSn/Ge 3D pFETs.
- T_{body} ↓ ➔ Quantum confinement energy ↑ ➔ I_{OFF} ↓
- Mechanical bending with external compressive strain ➔ I_{ON} ↑

Benchmark of ρ_C



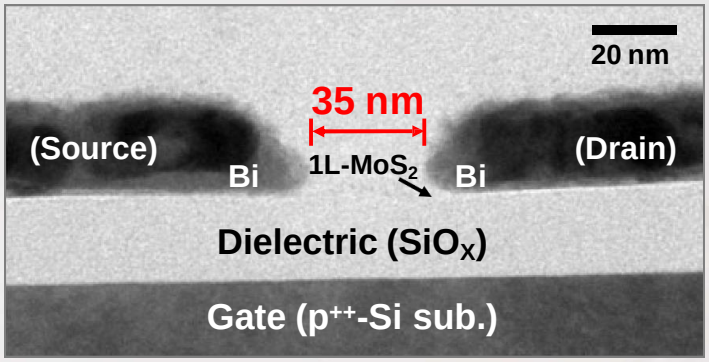
[13] P.-C. Shen et al., Nature, 2021.
[15] H. Yu et al., TED, 2016.
[16] E. Rosseel et al., ECS Trans, 2020.
[17] C.-N. Ni et al., VLSI, 2016.
[18] J. Maassen et al., APL, 2013.
[19] Z. Zhang et al., EDL, 2010.
[20] N. Stavitski et al., EDL, 2008.
[21] C.-E. Tsai et al., TED, 2020.
[22] H. Miyoshi et al., VLSI, 2014

- **2D materials have higher ρ_C than group-IV materials.**

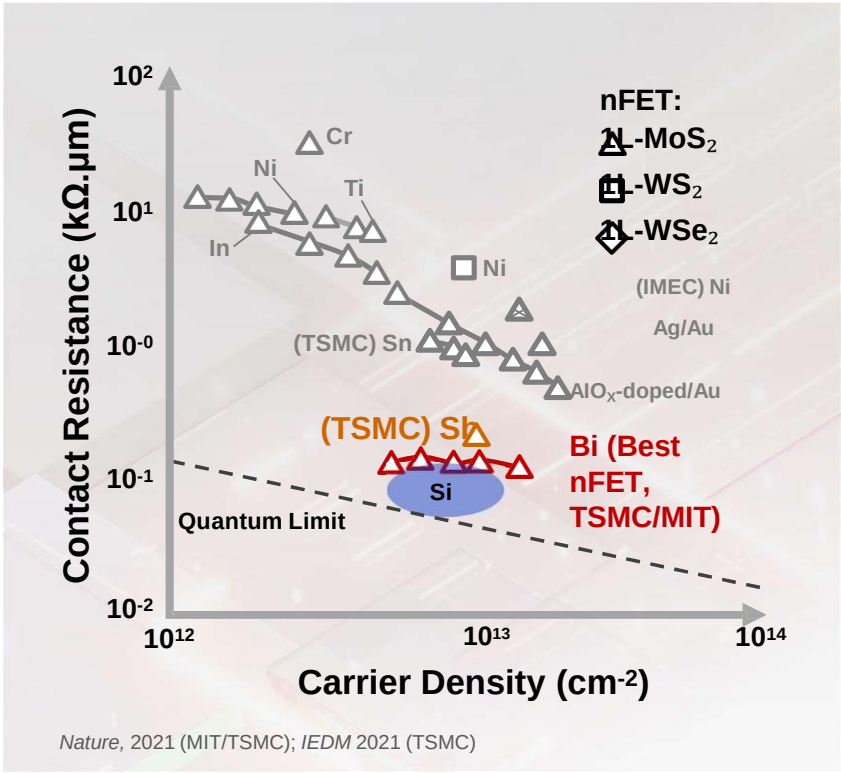
Innovative Semimetal Contact - Bismuth (Bi) and Antimony (Sb) for 2D Transistors

Semimetal (Bi) has zero Schottky barrier height due to near 0 DOS at E_F.

Bi-contacted MoS_2 FET



- **Bi: record low contact resistance and record high on-state current**
- **Sb: low contact resistance and enhanced thermal stability**



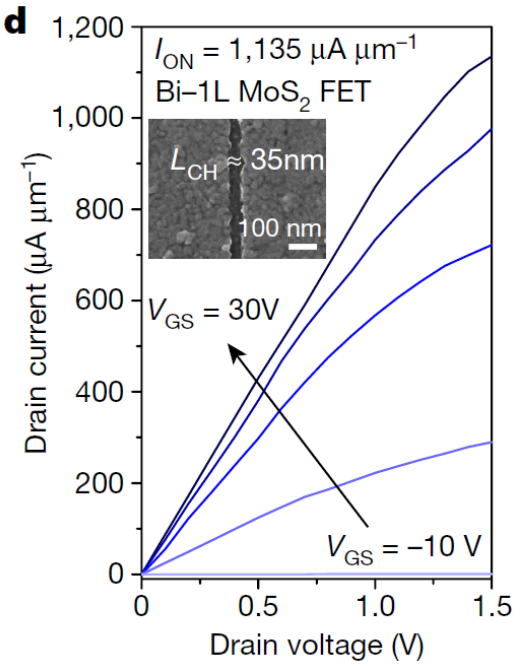
Nature, 2021 (MIT/TSMC); *IEDM* 2021 (TSMC)

- Top contact vs Edge contact

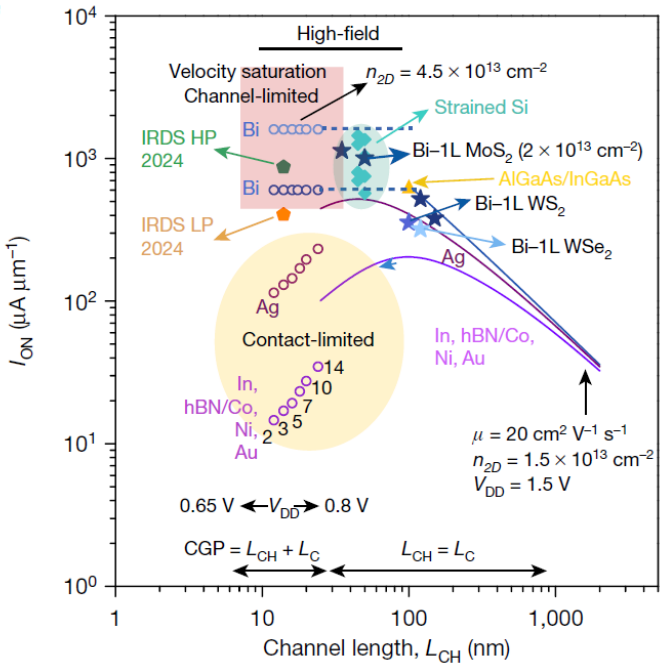
- 2D R_c is still worse than Si

I_{ON} of Bi-1L MoS_2 FET

(Fig. 4d)



(Fig. 4f)



Ref. P.-C. Shen et al., Nature, 593, 211-217, 2021.

- 2D FETs have lower I_{ON} than Si-based FETs (should be compared at same V_{GS}).
- VGS =30 V for beyond 1nm node???

Low I_{on} /high V_{dd} of 2D channels

Benchmark of I_{ON}

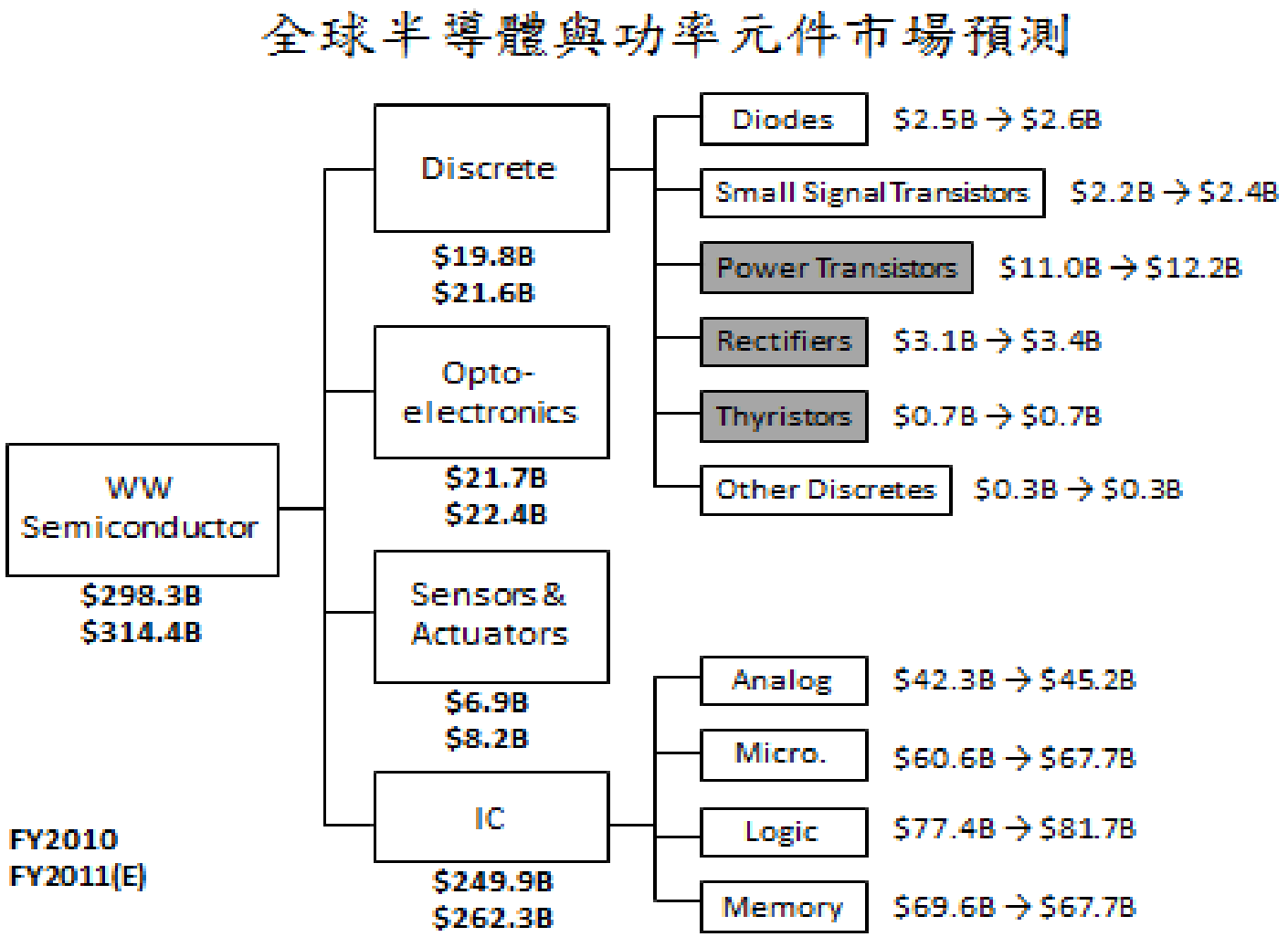
	nFET		pFET		
Reference	VLSI 2021 (NTU) [1]	Nature 2021 (MIT) [2]	IEDM 2021 (NTU) [3]	IEDM 2021 (NTU) [3]	IEDM 2021 (Intel) [4]
Channel	7 stacked $Ge_{0.95}Si_{0.05}$ nanowires	MoS_2	8 stacked $Ge_{0.9}Sn_{0.1}$ ultrathin bodies	8 stacked $Ge_{0.9}Sn_{0.1}$ thick nanosheets	WSe_2
I_{ON} at V_{DS}= ±1V	11000μA/μm at V_{OV}=V_{DS}= 1V	42μA/μm at V_{OV}=V_{DS}= 1V	910μA/μm at V_{OV}=V_{DS}= -1V	4400μA/μm at V_{OV}=V_{DS}= -1V	50μA/μm at V_G= -4V, V_{DS}= -1V

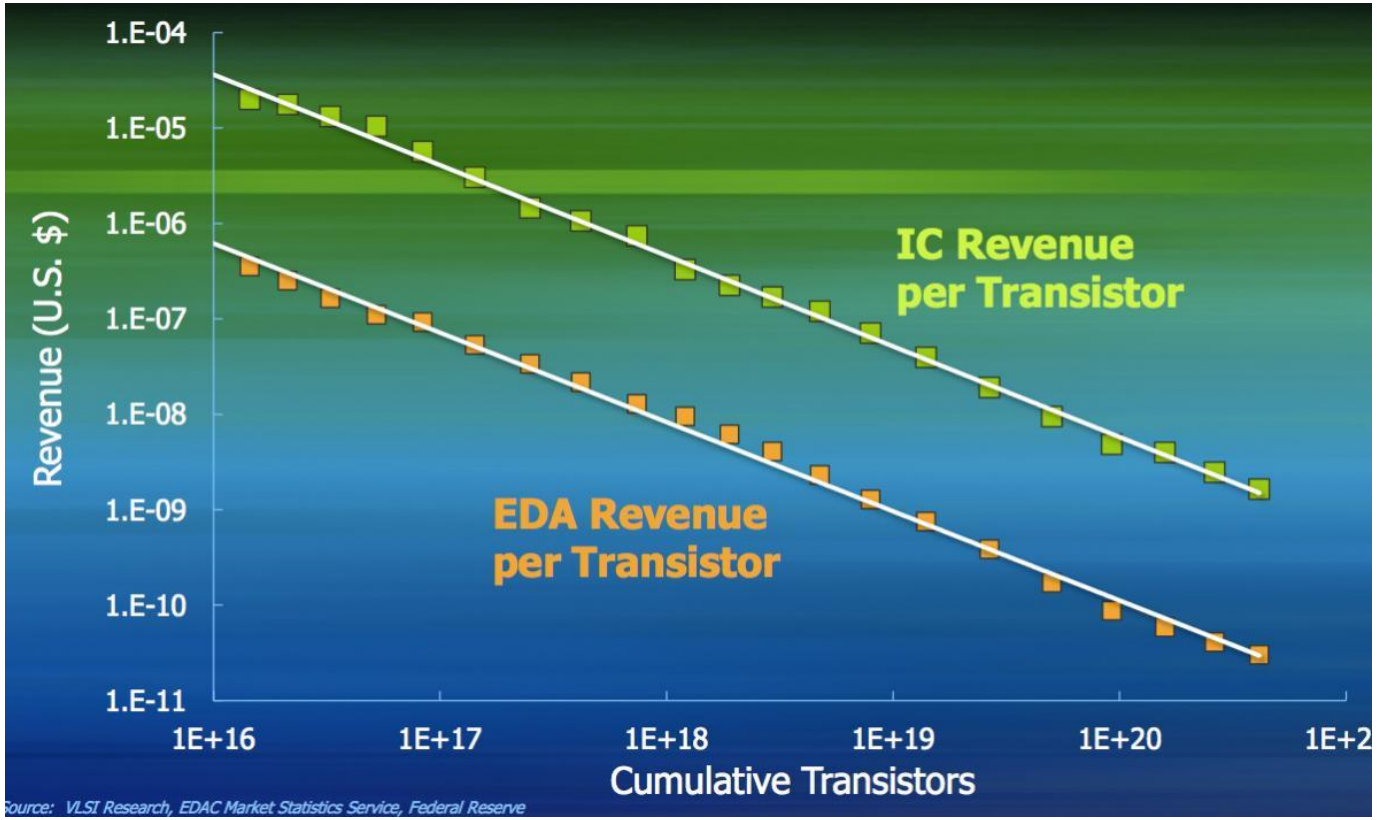
[1] Y.-C. Liu et al., VLSI, 2021, T15-2.
[2] P.-C. Shen et al., Nature, 593, 211-217, 2021.
[3] C.-E. Tsai et al., IEDM, 2021, pp. 569-572.
[4] K. P. O'Brien et al., IEDM, 2021, pp. 146-149.

What's Biggest Hurdle to Commercialization?

- **Getting industry to "buy in"**
  - — Industry is conservative & focused on next quarter's profits
  - — Academics are working with few students, few $$, and labs held together by duct tape
- **How is this going to happen?**
  - — Come up with novel, clever, **manufacturable** ideas
  - — Use **honest benchmarks** vs. incumbent technology
  - — Need deep & **sustained funding** (our funding system seems broken)
  - — Talk to industry friends, don't get shut out

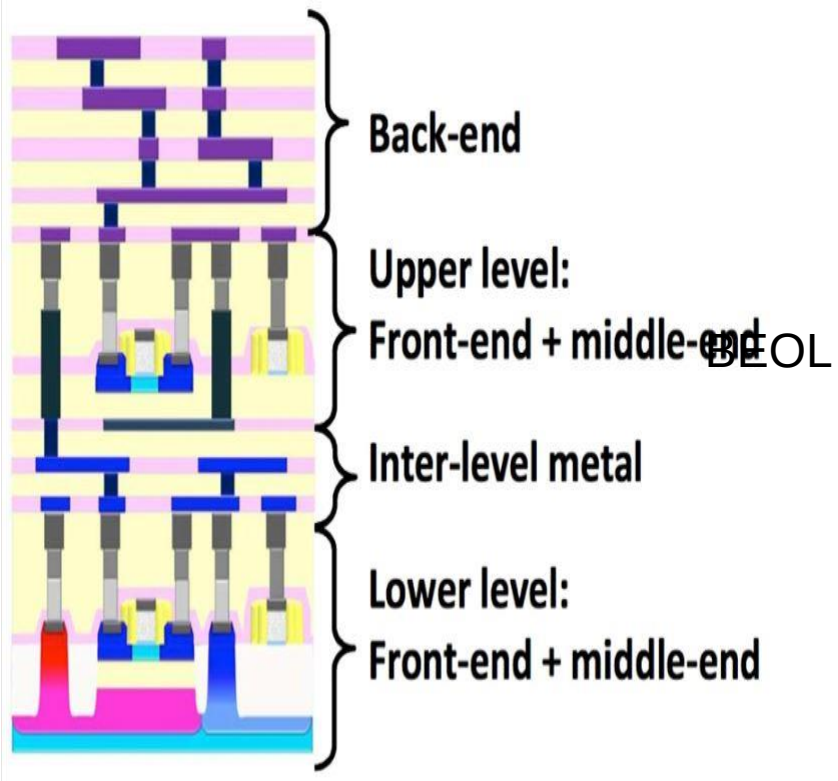
IC/Semiconductor





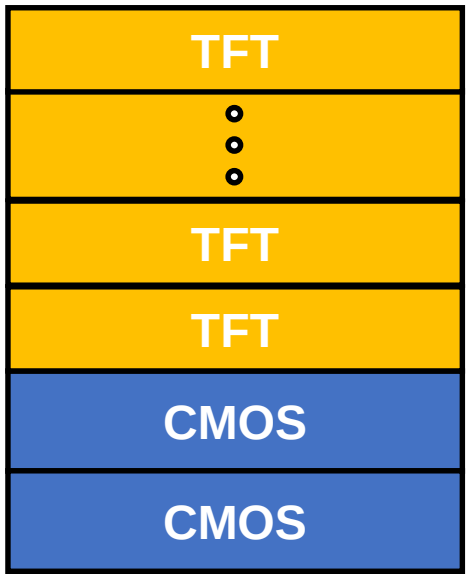
- **EDA 2% of Semi. 45,938 people 10B 2019**
- **Tsmc : 50,000 40 B (E) 2020**

Please Refer VLSI short course 2022 by Marko Radosavljevic, Intel (many sidles from his short course)



- **Stacked Transistors/ IC on IC**
- **Most difficult part: Battle field**

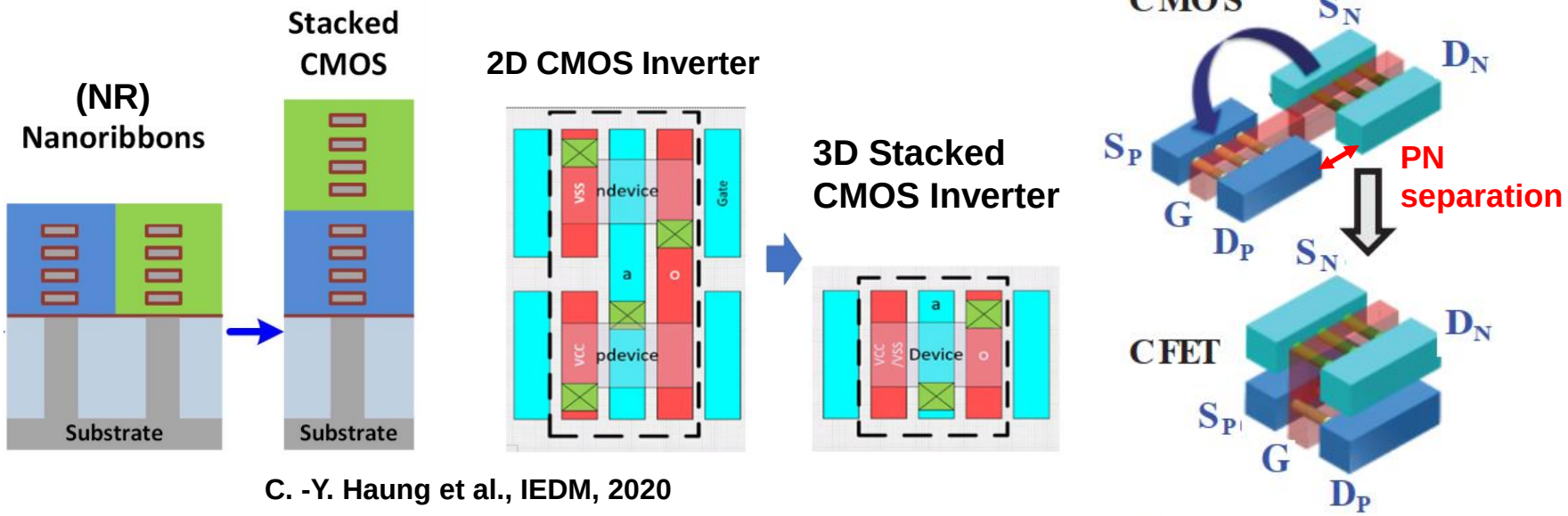
TFT on CMOS



- **a-IGZO**
  - **No grain boundary**
  - **Mobility ~ 20 cm^2/V-s**
    (Higher mobility is better)
- **Poly-Si**
  - **High mobility (~100 cm^2/V-s)**
  - **Grain boundary**
  - **Laser annealing**
- **2D material**
  - **Ultra high mobility**
  - **Size too small (cm x cm)**

- **CMOS is hard to realize transistor stacking due to epitaxy difficulties.**
- TFTs is easier to stack compared to CMOS due to low processing temperature.
- **a-IGZO TFT is promising due to high mobility compared to a-Si**

Transistors Stacking(monolithic-epi) – CFET (N+1 node)

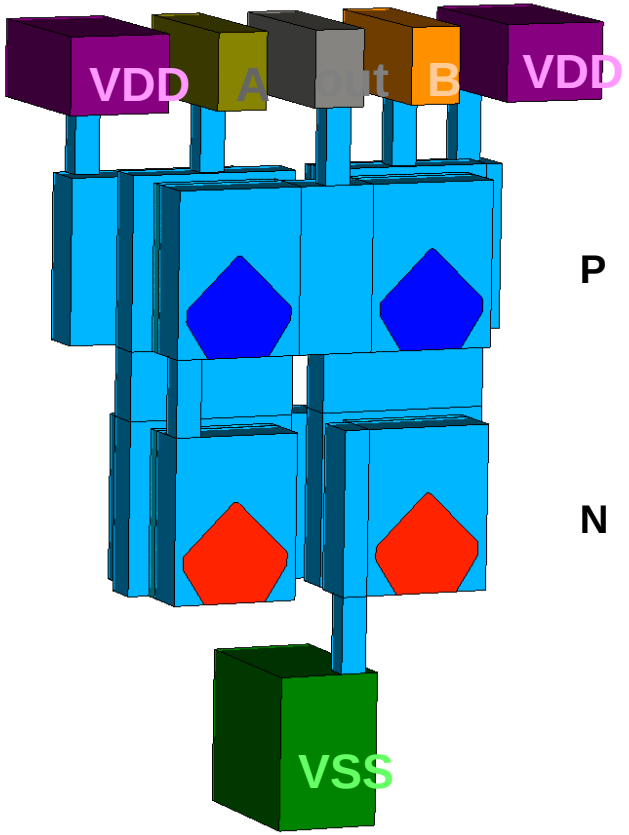


C. -Y. Haung et al., IEDM, 2020

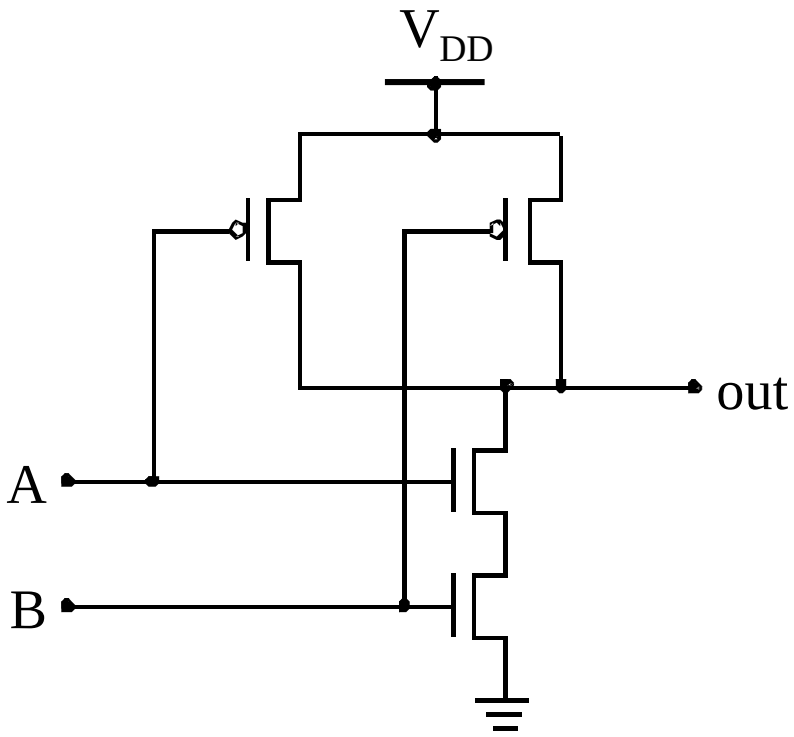
S.Subramanian et al., VLSI, 2020

- CFET architecture is the combination of NMOS and PMOS with devices vertically stacking and controlled with a common gate.
- CFET removes the lateral PN separation bottleneck and allows PMOS/NMOS folding.
- Transistor innovation of CFET architecture enables 50% area scaling by reducing cell height.

CFET NAND

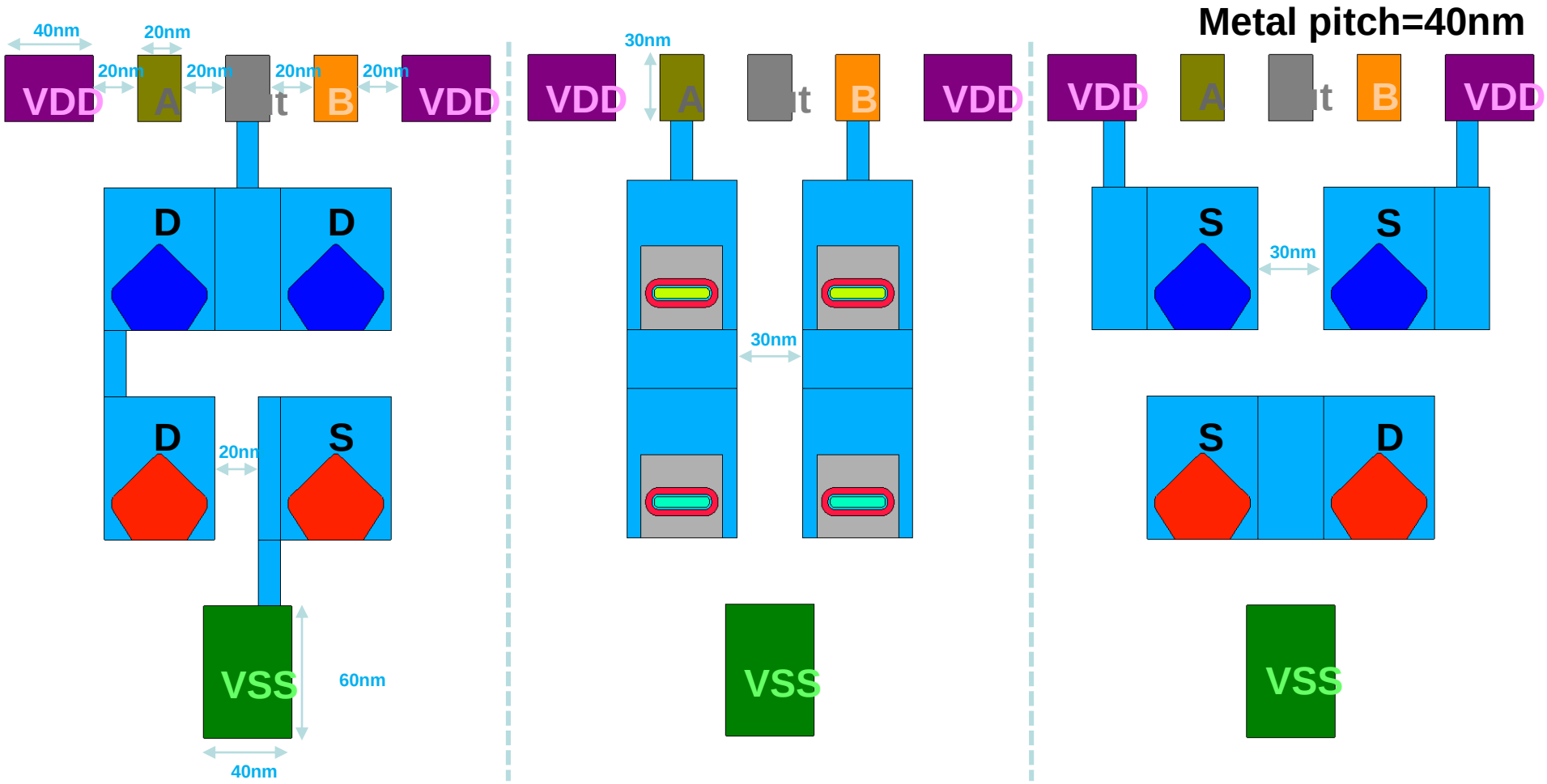


3D structure



circuit

Cross-section



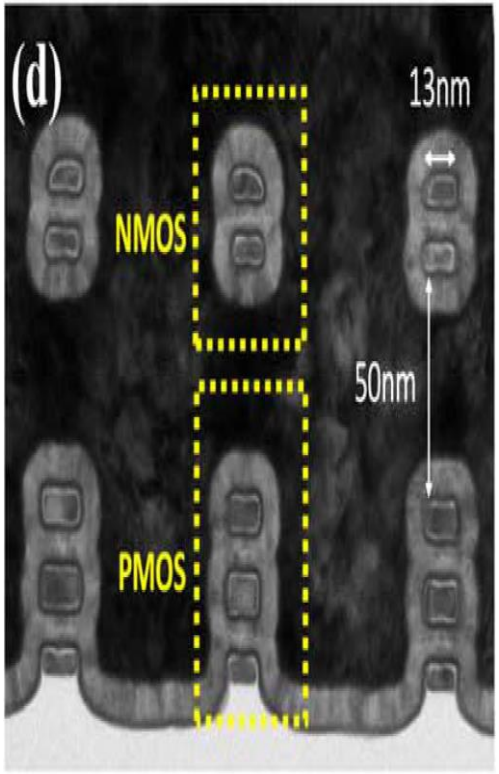
Transistor architectures beyond stacked NS ➔ footprint ↓

CFET (Complementary FET)

Intel

Sequential CFET

W. Rachmady et al., IEDM, 2019, pp. 697

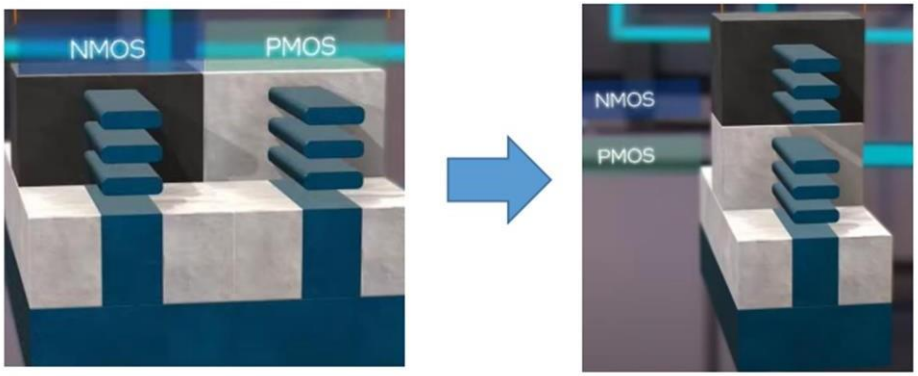


Intel

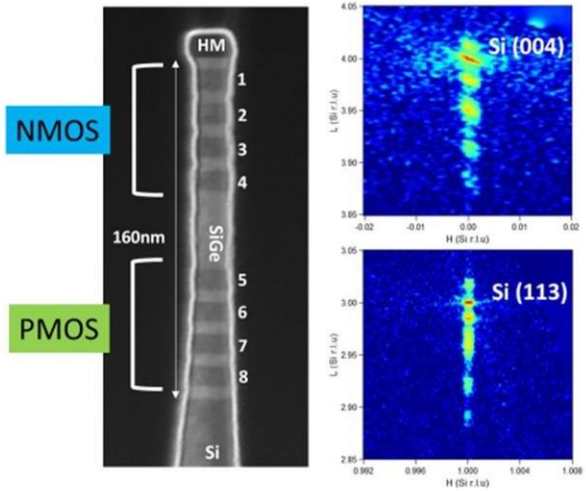
Monolithic CFET

C.-Y. Huang et al., IEDM, 2020, pp. 425

Self-Aligned 3D Stacked CMOS



Intel Innovation Day 2021

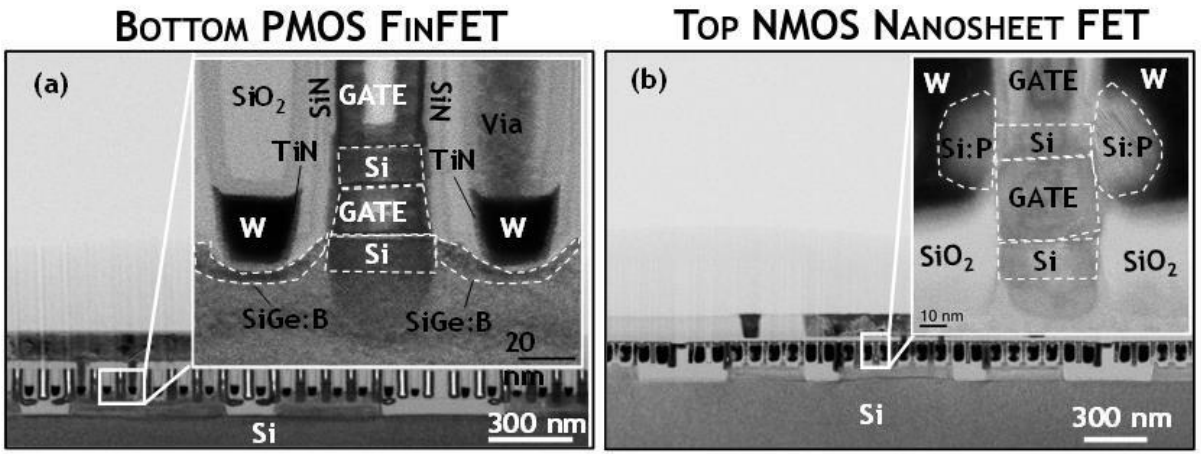


C.-Y. Huang et al, IEDM 2020

- Self-aligned 3D stacking eliminates the need for wafer bonding process and separate bottom/top device process → lower cost; more process flexibility;
- High aspect ratio leads to process complication; less flexibility for heterogeneous channel material

Self-Aligned 3D Stacked CMOS Example [1]

FINAL DEVICE

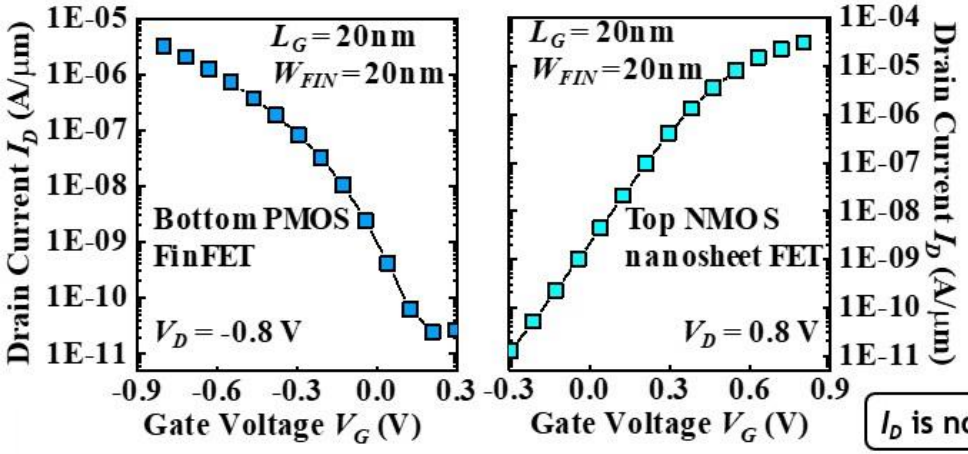


- TEM cross-section images of the **final bottom PMOS FinFET** and **top NMOS nanosheet FET.**
- Both PMOS and NMOS devices shown above have L_G**= 20nm**, fin pitch of 45nm and contact-poly pitch (CPP) = 90nm.

S. Subramanian et al, VLSI2020

Self-Aligned 3D Stacked CMOS Example [2]

DEVICE PERFORMANCE

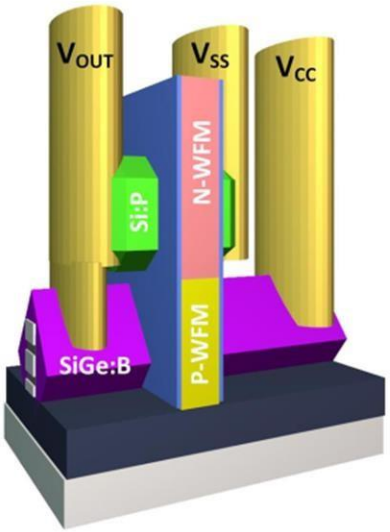


I_D is normalized to the device footprint.

- I_D-V_G of bottom PMOS FinFET and top NMOS nanosheet FET. **NMOS *I-V*** curve is from the wafer with **525 °C epi**.
- Device dimensions/architecture and process flow needs to be refined to achieve performance targets.

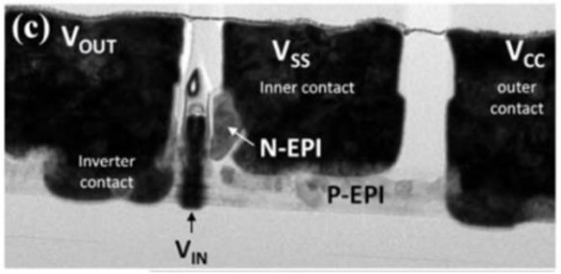
S. Subramanian et al, VLSI2020

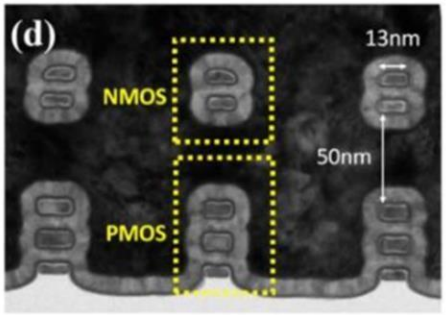
Self-Aligned 3D Stacked CMOS Process Flow



- Si/SiGe epitaxy stack
- Si/SiGe nanoribbon fin patterning
- Poly dummy gate patterning
- Gate spacer formation
- **Vertically stacked dual S/D EPI**
- Interlayer dielectric deposition and polish
- Poly dummy gate removal
- Si nanoribbon release (SiGe etch)
- High-k deposition
- **Vertically stacked dual metal gate**
- Inner contact/outer contact/inverter contact etch
- Contact metal/W fill and CMP

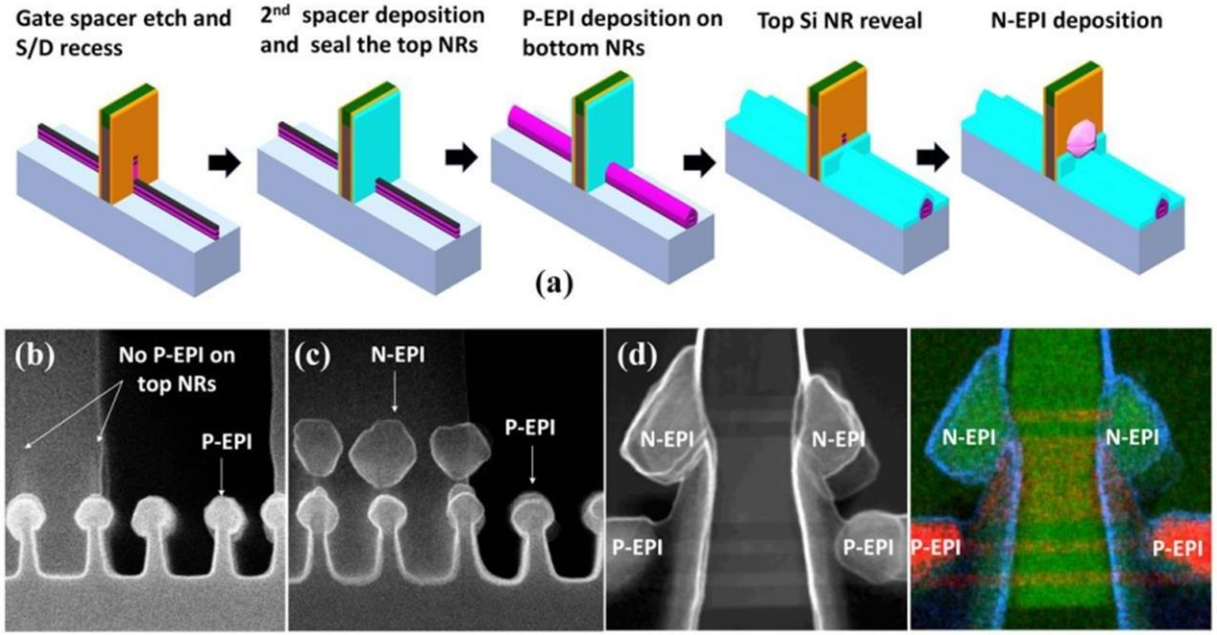
C.-Y. Huang et al, IEDM 2020





- Key module enablers for self-aligned 3D stacking are vertical stacked/isolated dual S/D epi and dual WF process within high aspect ratio trench depth

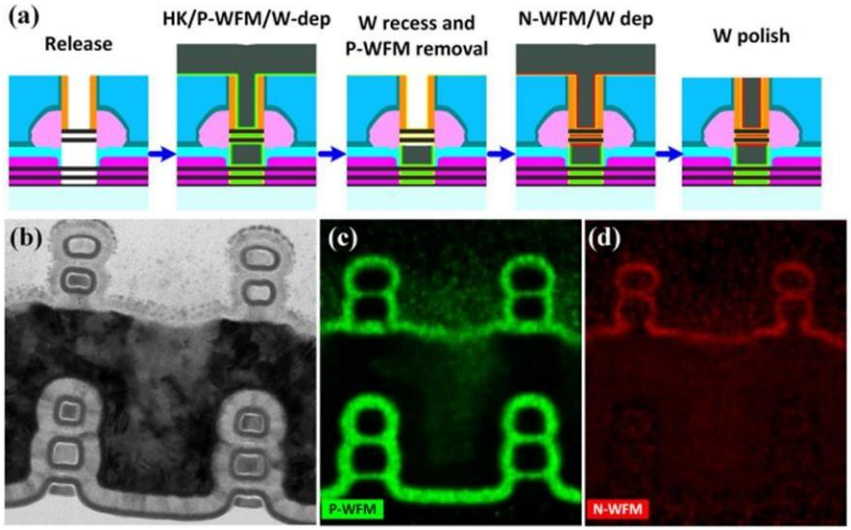
Self-Aligned 3D Stacked CMOS – Dual S/D Epi

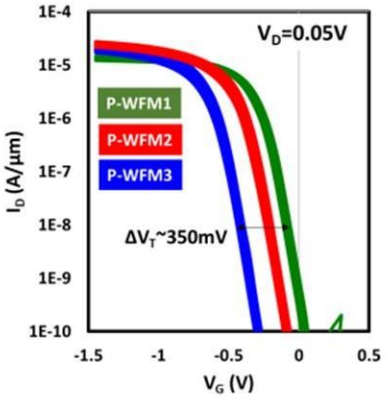


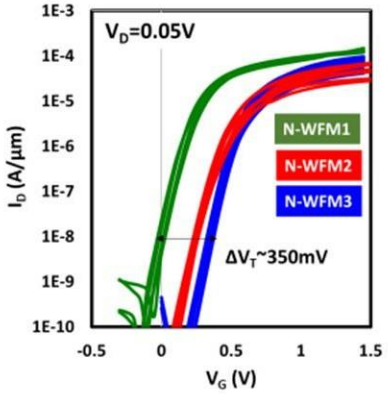
C.-Y. Huang et al, IEDM 2020

- Vertically stacked/isolated dual S/D epi process critical for performance, yield and reliability co-optimization

Self-Aligned 3D Stacked CMOS – Dual Metal Gate





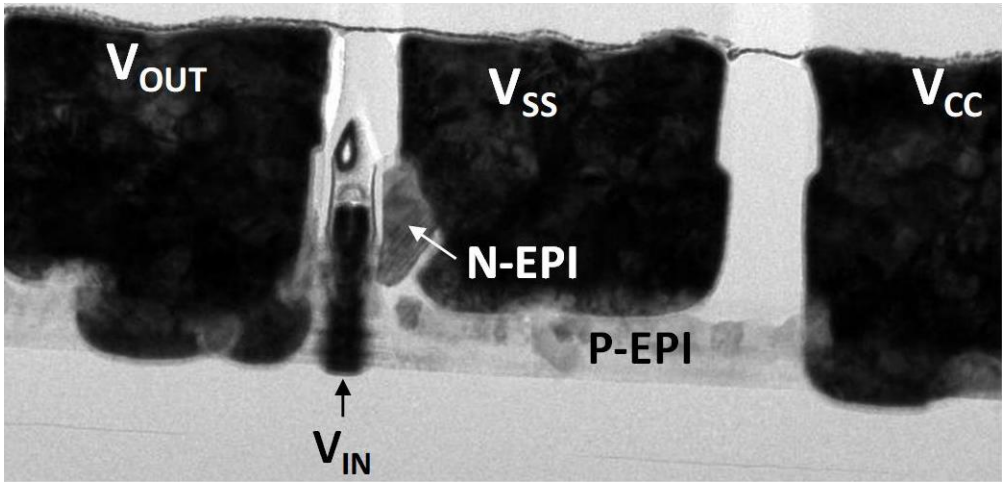


PMOS and NMOS V_T can be adjusted for a range of ~350 mV with different WFM and thickness.

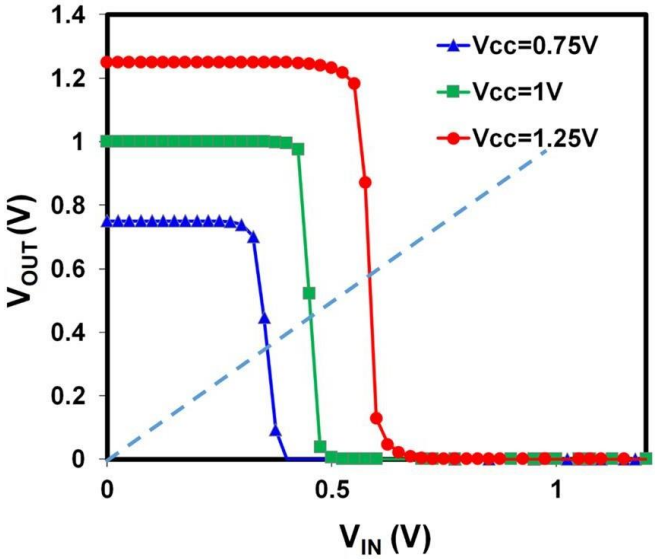
C.-Y. Huang et al, IEDM 2020

- Vertically stacked dual metal gate process requires integration/patterning scheme optimized for high aspect ratio stack.

Self-Aligned 3D Stacked CMOS Inverter

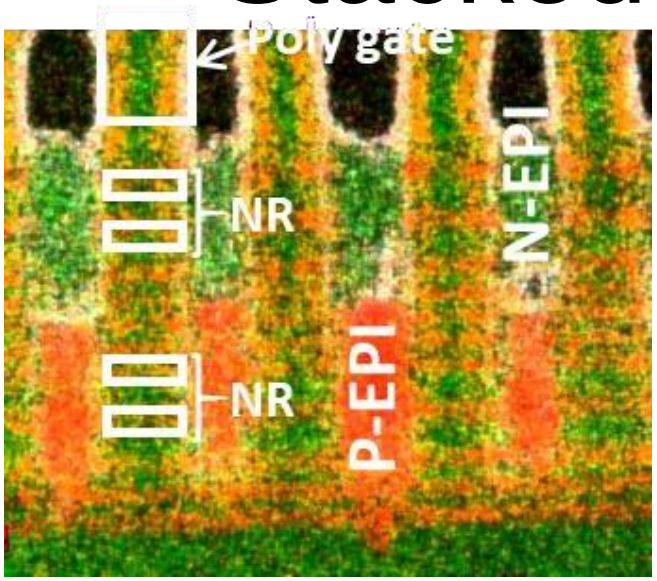


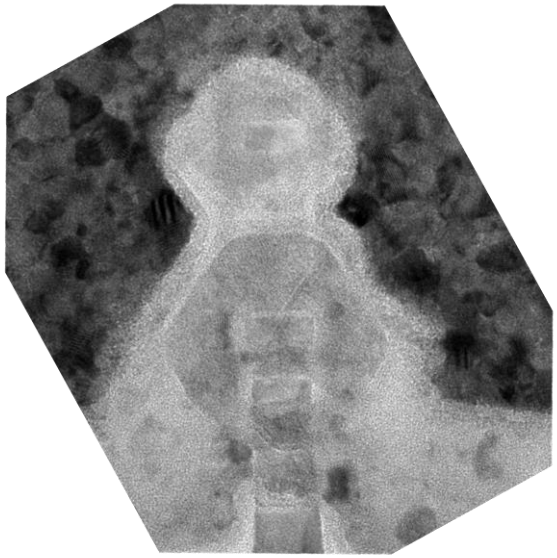
C.Y. Huang et al, IEDM 2020

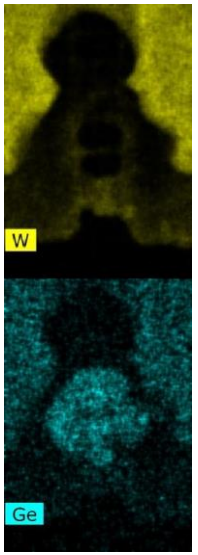


- Vertical interconnects enable demonstration of simple logic gates

Scaled Self-Aligned 3D Stacked CMOS



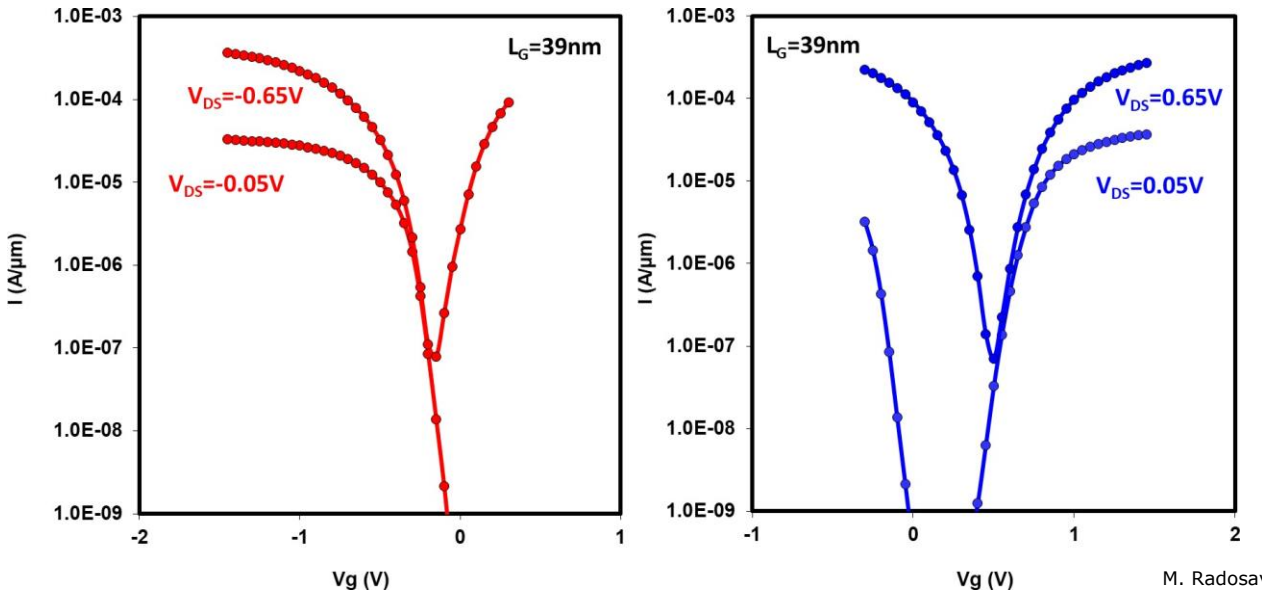




M. Radosavljevic et al, IEDM 2021

- Stacked dual epi demonstrated in scaled self-aligned devices

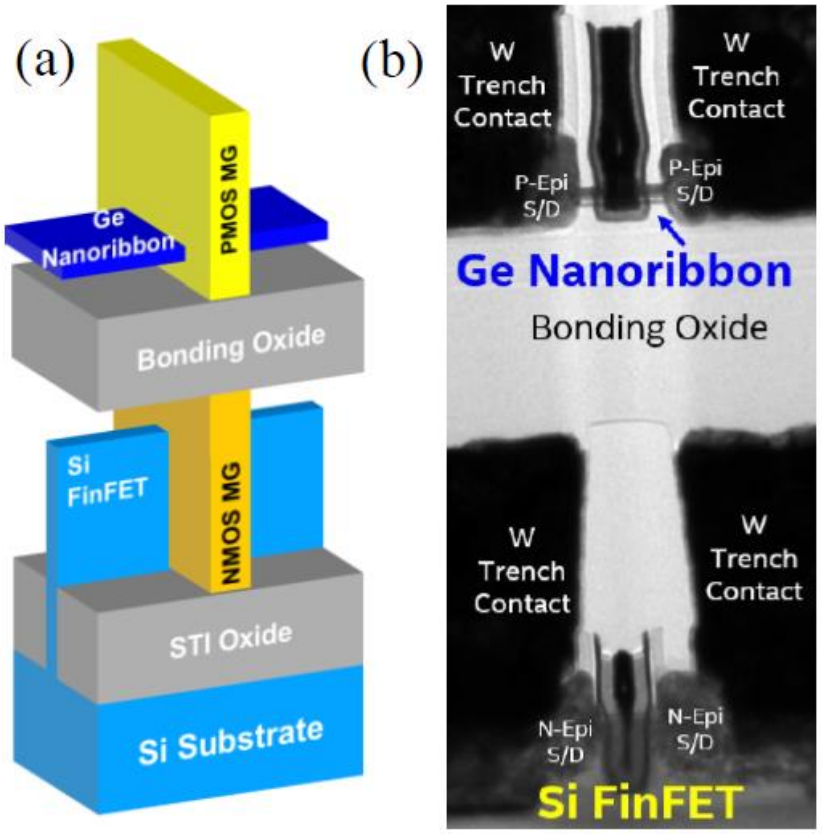
Scaled Self-Aligned 3D Stacked CMOS



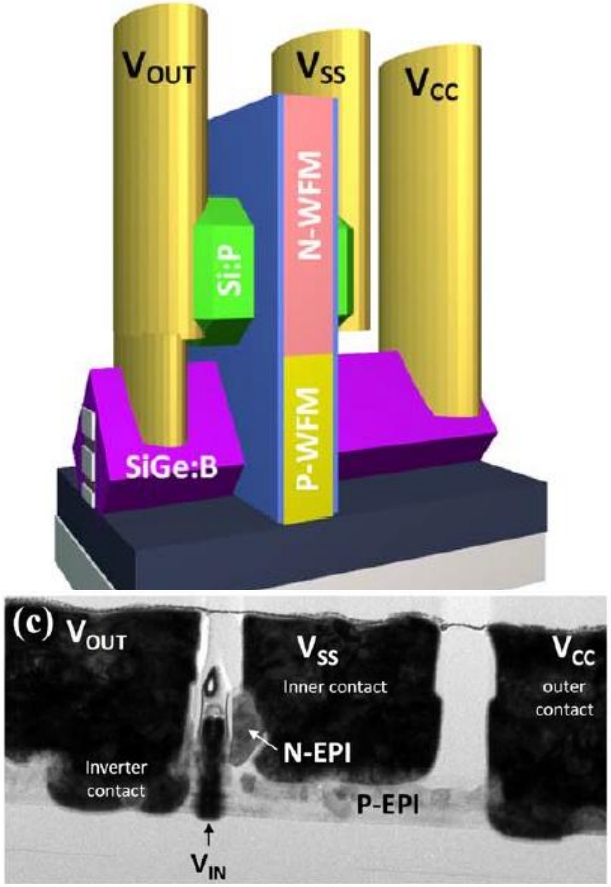
M. Radosavljevic et al, IEDM 2021

- Without split gate and split contact devices behave as both NMOS and PMOS based on biasing

IEDM 2019 Sequential (bonding)

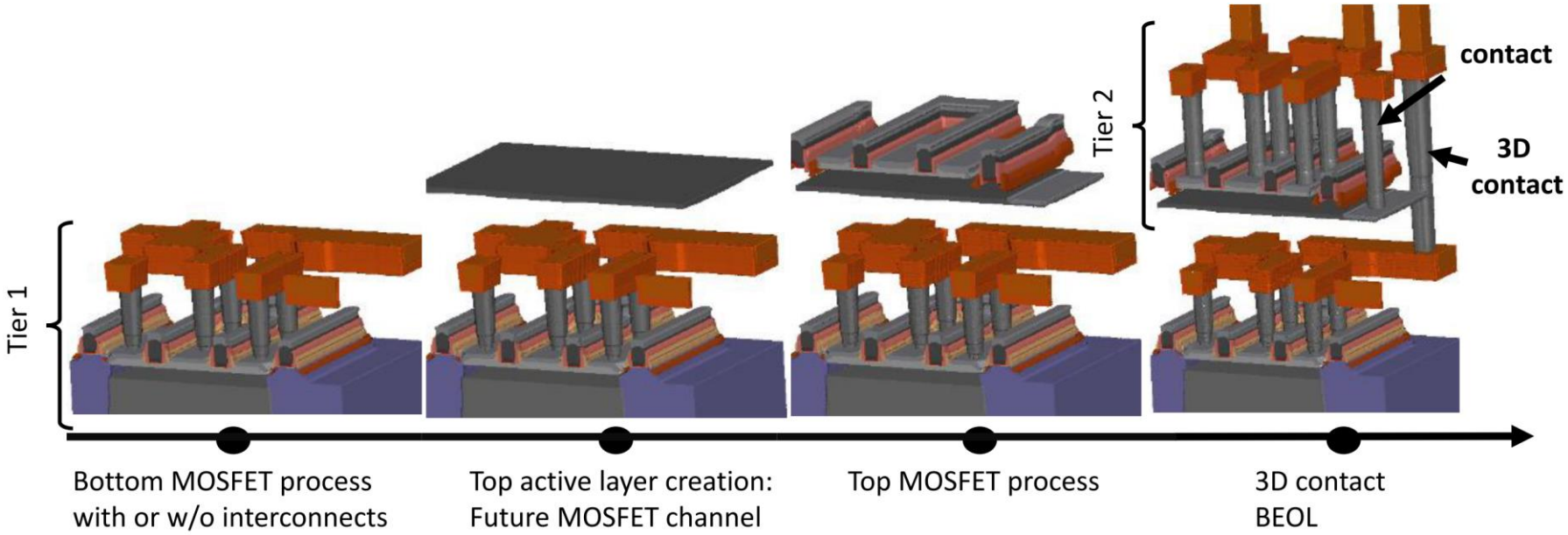


IEDM 2020 Monolithic (epi)



3D Sequential Integration (bonding)

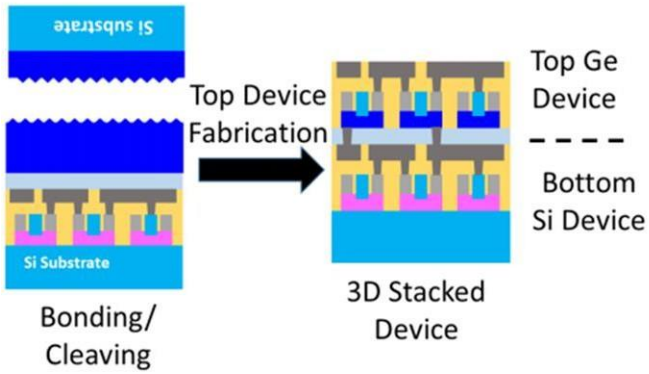
C. Fenouillet-Beranger et al., IEDM sc, 2020



- **Also named 3D monolithic, 3D VLSI … etc.**
- **Sequential process ➔ Thermal budget constraints for top devices processing.**

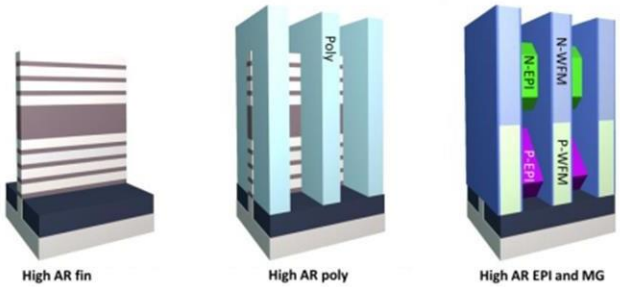
Sequential vs. Self-Aligned 3D stacked CMOS

Sequential 3D stacked CMOS



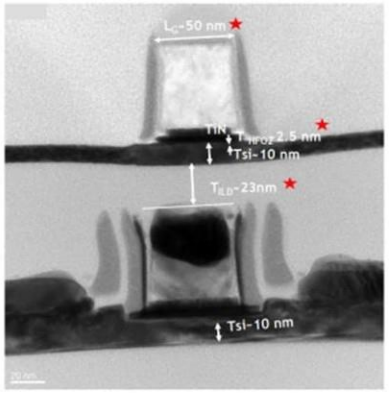
- Separate top and bottom layer processing
- Flexible to co-integrate heterogeneous channel materials
- Thermal budget limit
- Top devices not self-aligned to bottom devices
- Higher cost

Self-Aligned 3D stacked CMOS

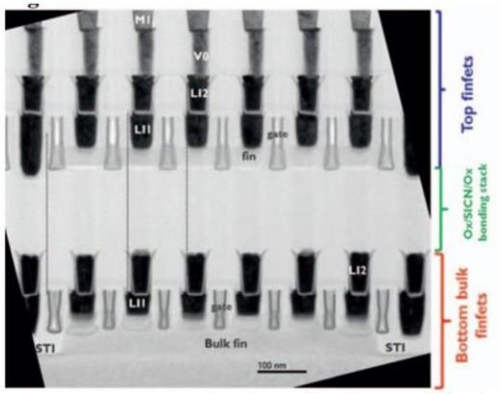


- Self-aligned fin and poly patterning for top and bottom devices
- Similar flow as conventional 2-D CMOS → lower cost
- Higher aspect ratio process complication

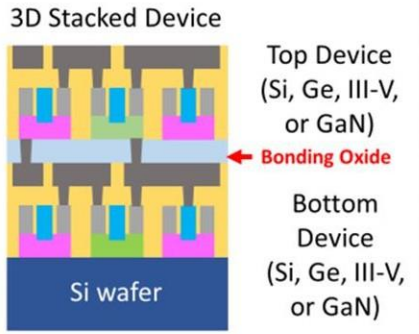
Sequential 3D Stacked CMOS Examples



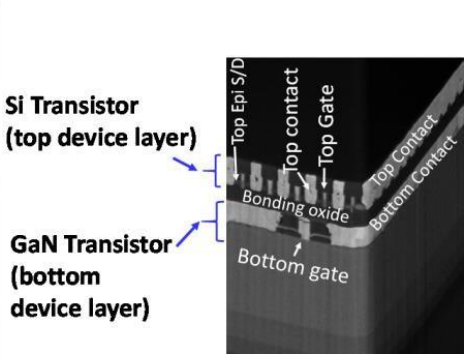
P. Batude et al, IEDM 2011



A. Vandooren et al, IEDM 2018



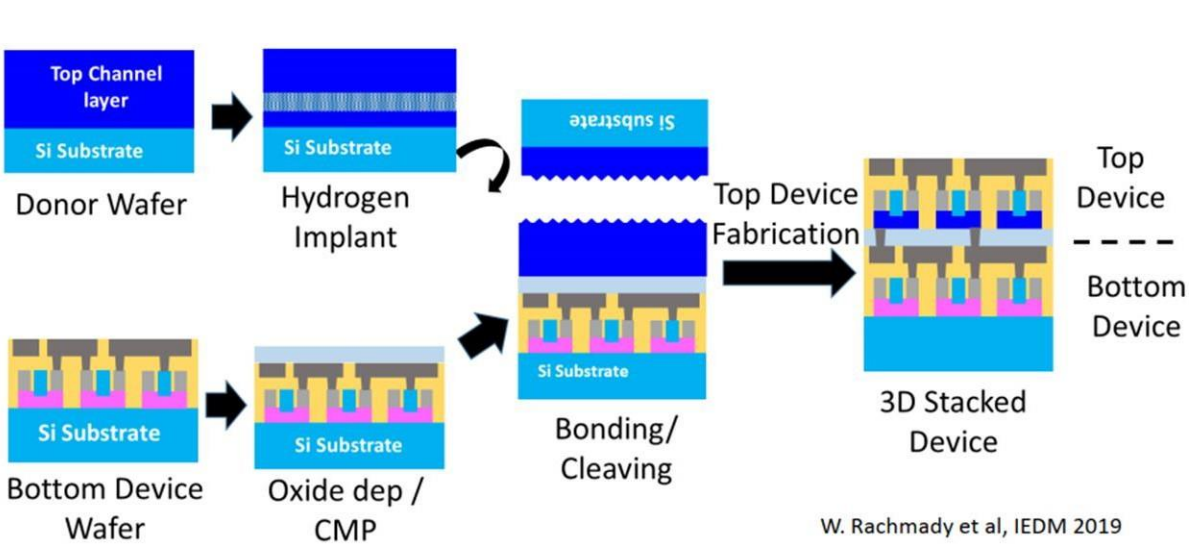
W. Rachmady et al, IEDM 2019

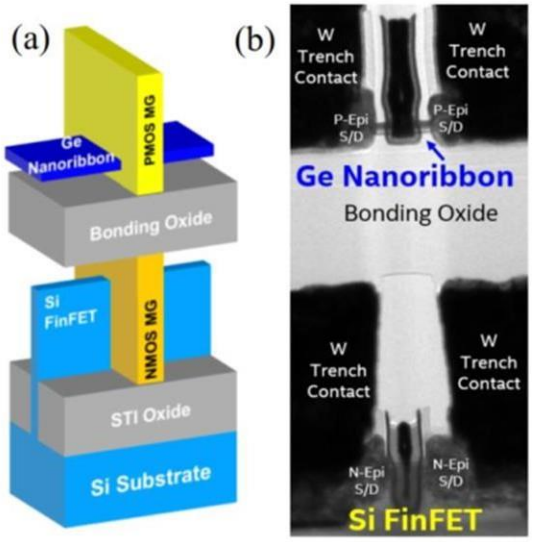


H. W. Then et al, IEDM 2019

- Sequential 3D Stacking opens up flexibility for independent optimized N and PMOS devices
- Require layer transfer technique for heterogenous integration
- Thermal budget of top transistor is limited – Gate stack quality, S/D dopant activation and interconnect

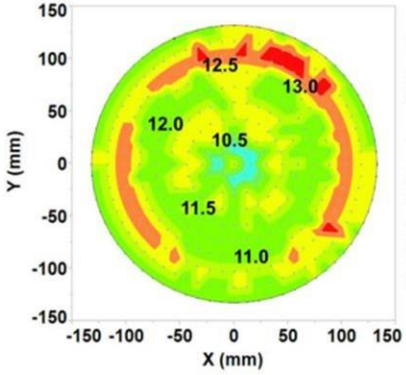
Sequential 3D Stacked CMOS Process Flow

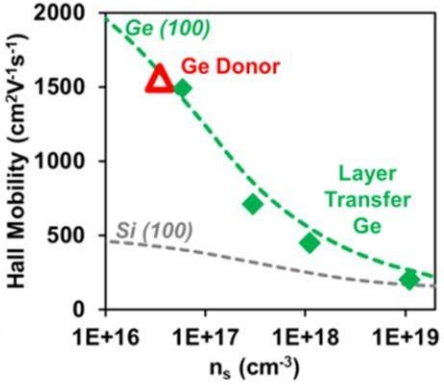


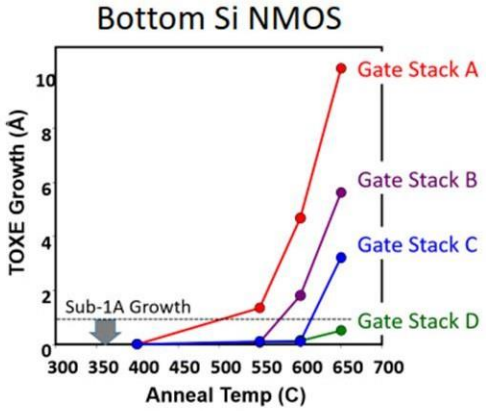


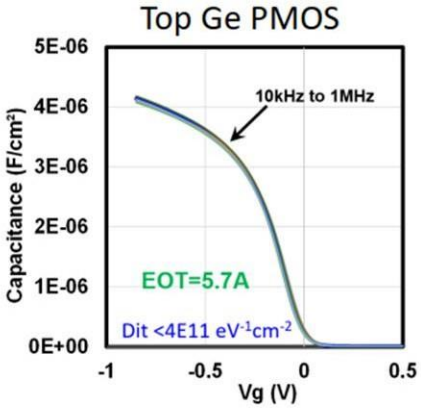
- Top channel layer uniformity/quality post layer transfer
- Gate stack quality to account for thermal budget, S/D dopant activation and interconnect

Sequential 3D Stacked CMOS – Key Modules [1]







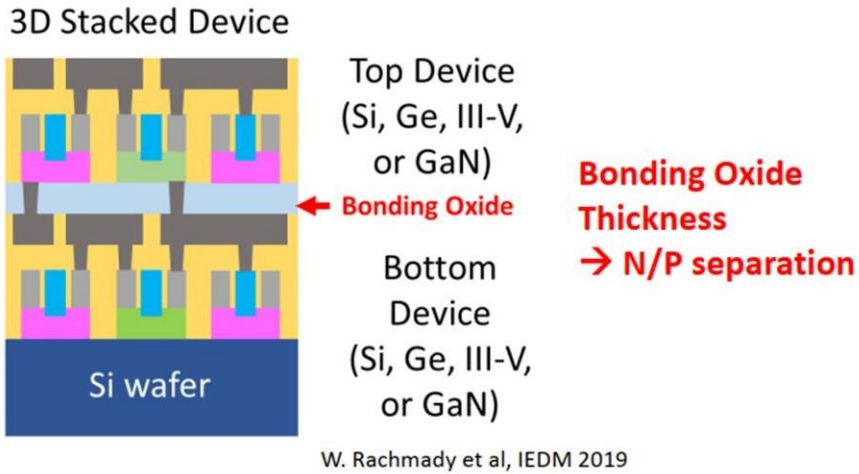


W. Rachmady et al, IEDM 2019

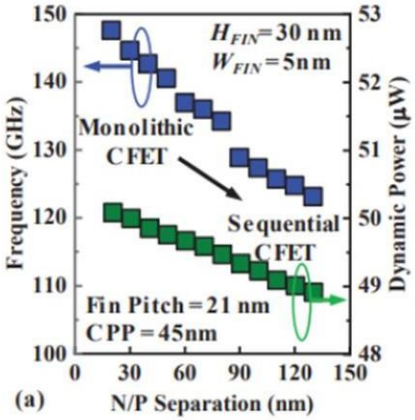
- Top layer Ge thickness uniformity post layer transfer critical for device variability control
- Maintained Ge layer mobility benefit over Si vs donor wafer post layer transfer

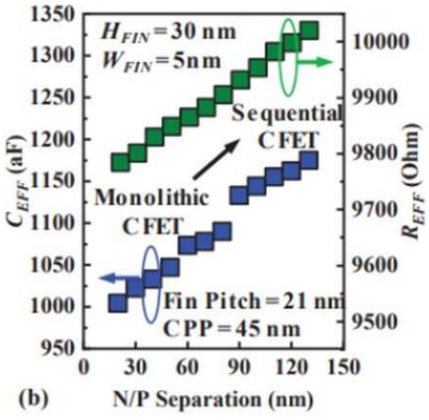
- Both bottom and top device requires careful gate stack process optimization to account for additional thermal budget for bottom device and low thermal budget for top device

Sequential 3D Stacked CMOS – Key Modules [2]



W. Rachmady et al, IEDM 2019

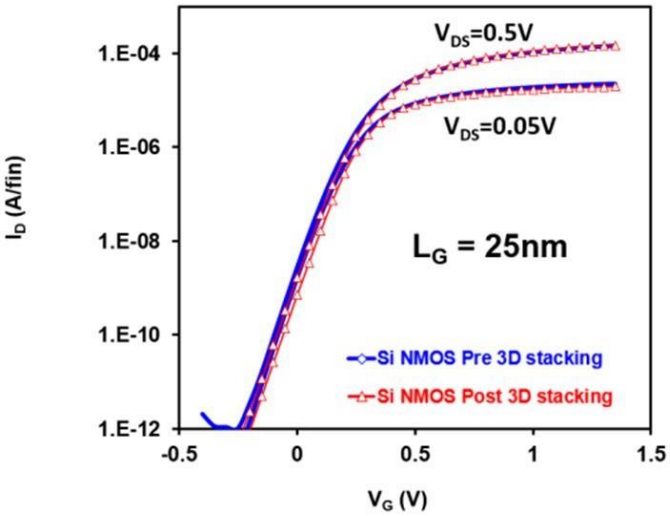


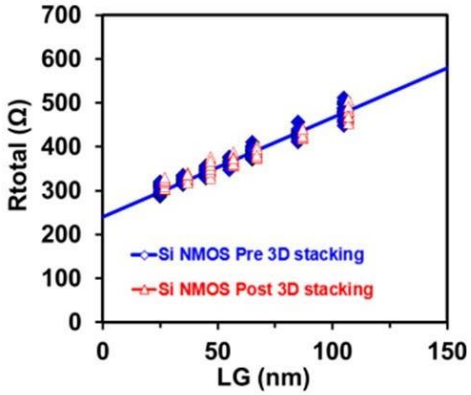


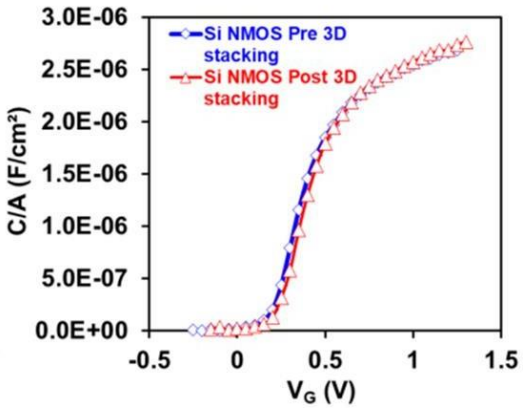
S. Subramanian et al, VLSI 2020

- Bonding oxide thickness leads to strong trade-off between bonding void defects and performance (parasitic capacitance and resistance modulation)
- Thinner bonding oxide without void defect is critical enabler for Sequential 3D performance gain

Sequential 3D Stacked CMOS – Device Response



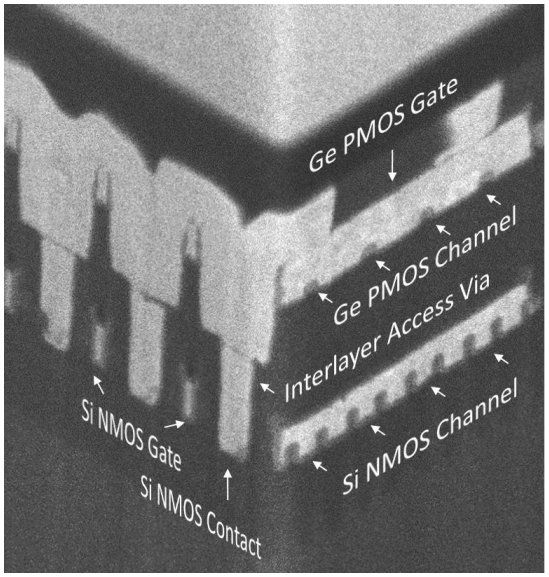


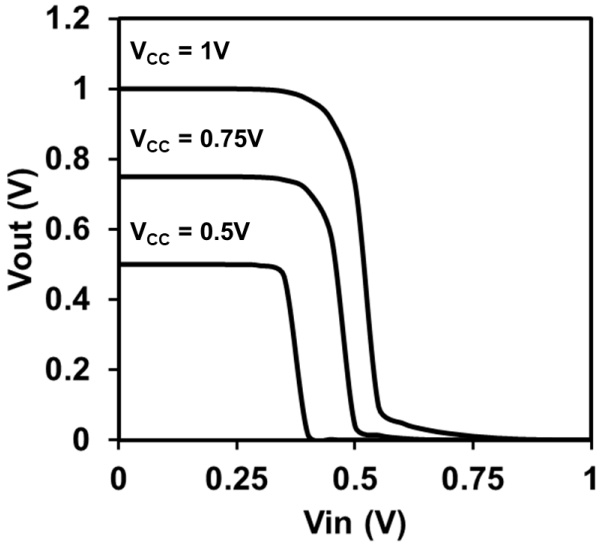


W. Rachmady et al, IEDM 2019

- Maintain bottom device integrity before and post 3D stacking is one of successful criteria
- Bottom device design requires consideration of additional layer transfer and thermal budget

Sequential 3D Stacked CMOS Inverter

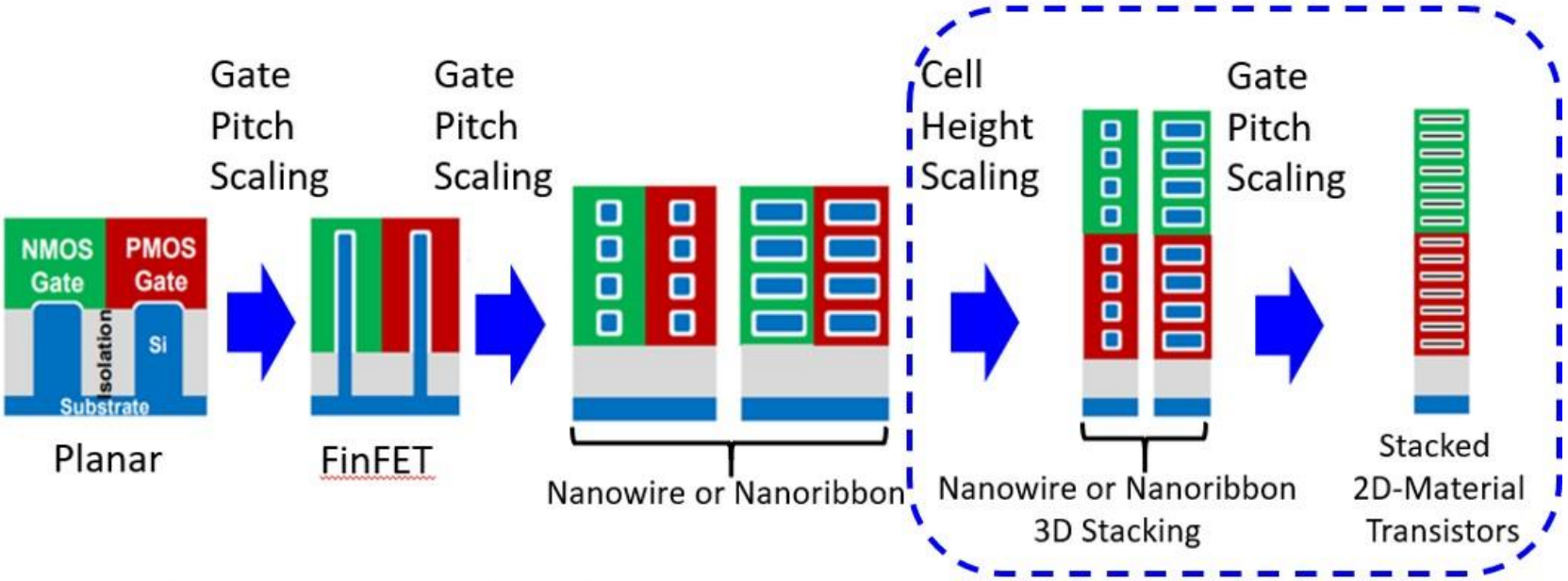




- Vertical interconnects enable demonstration of simple logic gates

W. Rachmady et. al., IEDM 2019

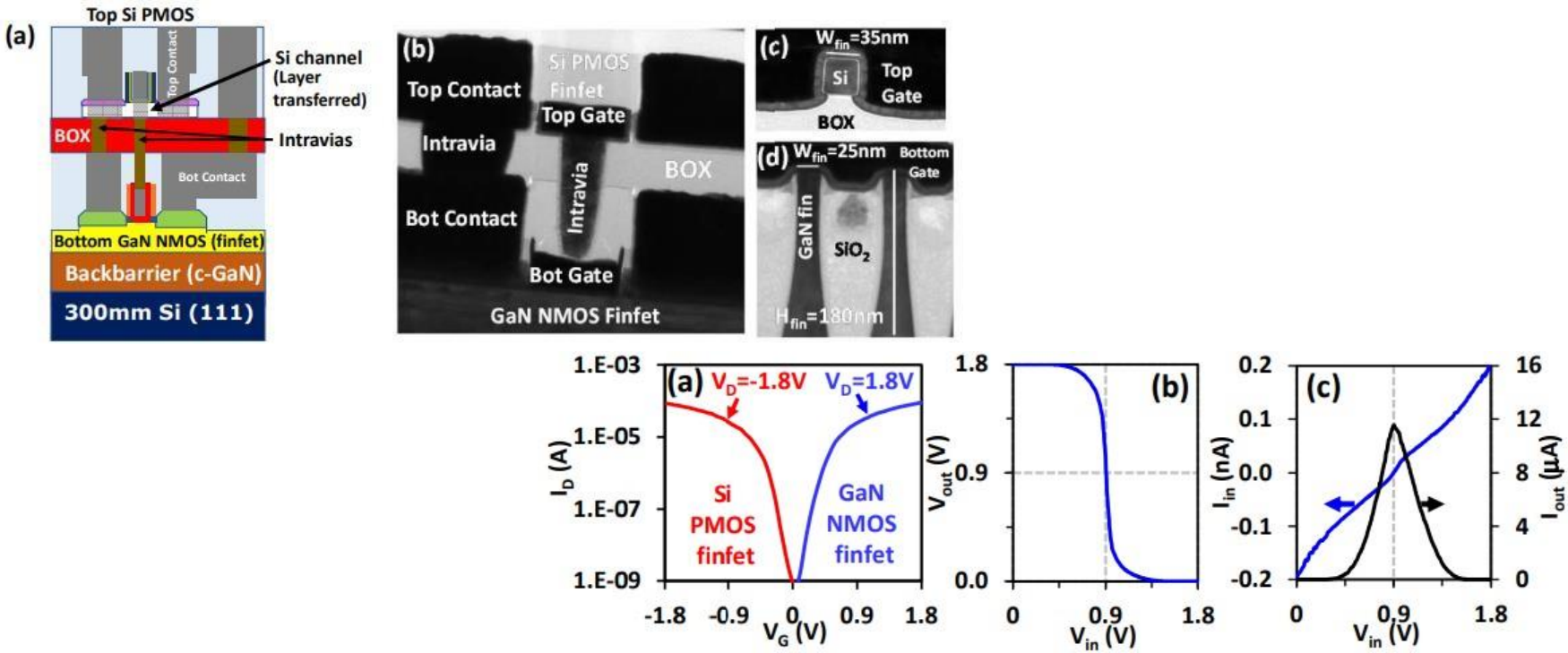
Transistor architecture innovation continues to drive Moore's law scaling



Process and Packaging Innovations for Moore's Law Continuation and Beyond.

Robert Chau, *IEDM*, 2019, pp. 1.1.1-1.1.6.

Sequential Heterogeneous 3D Stacked CMOS Inverter

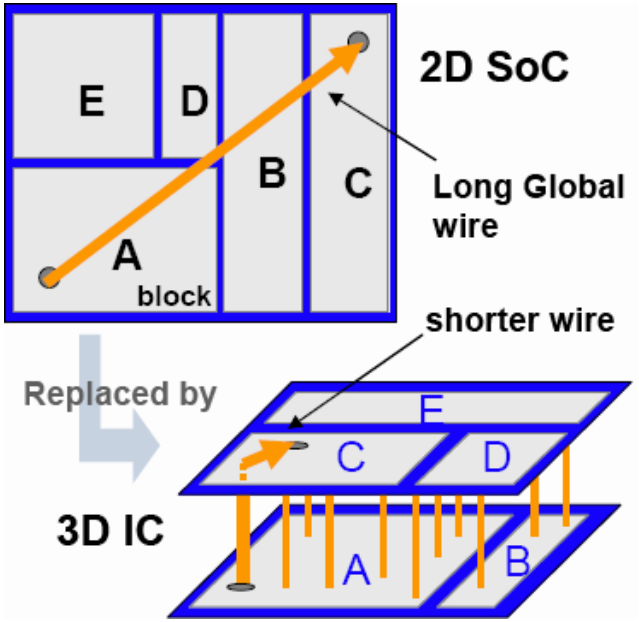


- Similar techniques can be applied to different materials and applications – e.g. in power delivery

H. W. Then et al, IEDM 2020

3D IC (package)

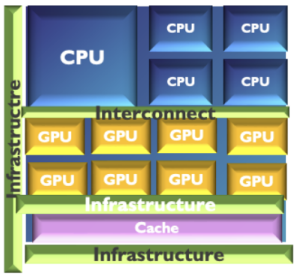
- Scaling faces physical and cost limits.
- Size reduction
- High interconnect density
  - Faster
  - Less power consumption
- Hetero-integration
  - Memory, logic, optical, MEMS, RF chip



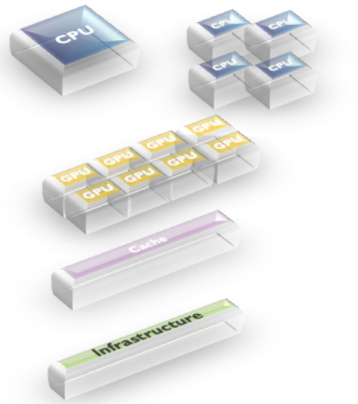
TSMC course

From high integration to "disintegration"

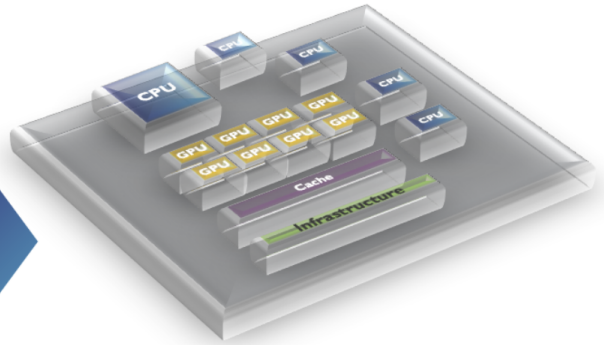
SoC



Breaking down the complex System-on-Chip (SoC) into smaller, more manageable components



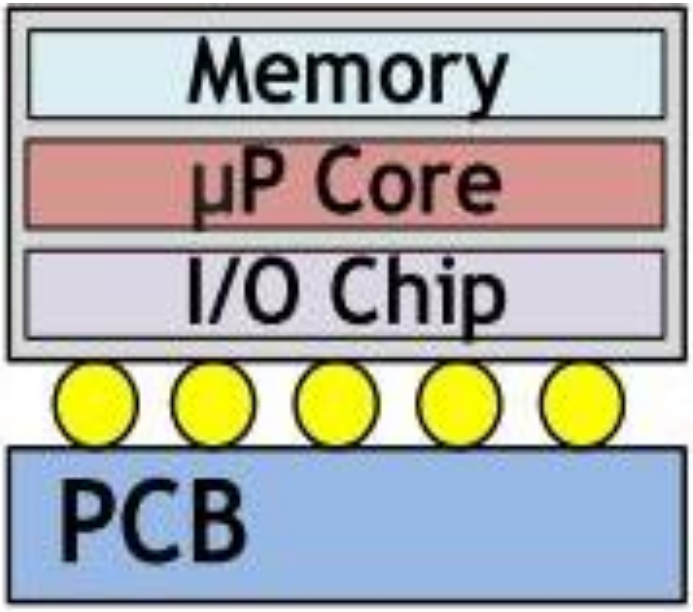
The blocks can be optimized separately at design stage but also at technology level



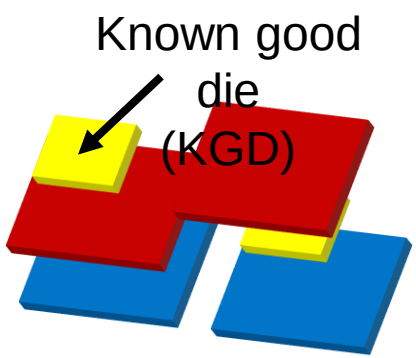
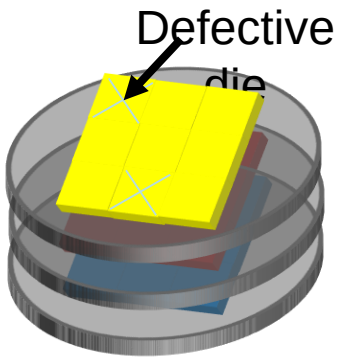
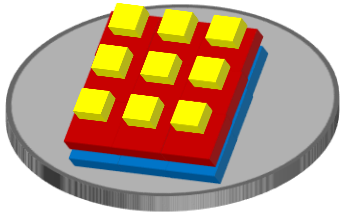
Reassembled using 2.5D (interposers) or 3D stacking

Advanced packaging for high-bandwidth and low-latency connectivity

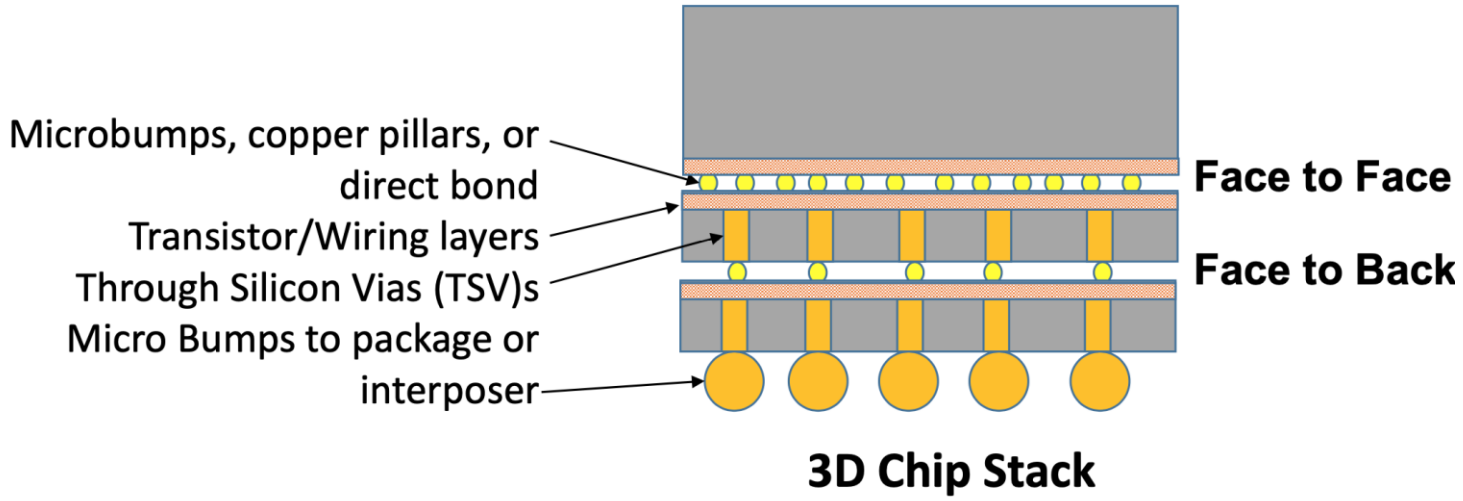
Chiplet Stacking/ Stacked Dielets



- **SOIC (tsmc)**
- **X-Cube (Samsung)**
- **FOVEROS (Unique and special in Greek) /EMIB (Intel)**

Bonding methods	Die-to-Die	Wafer-to-Wafer	Die-to-Wafer
			
Stacking procedure	Pick & Place	Wafer bonding	Pick & Place
High production throughput	No	Yes	Middle
High production yield	Yes	No	Middle
High flexibility in chip size	Yes	No	Yes
Application	DRAM(HBM)	CIS	Chiplets-on-Si interposer CIS(InGaAs-on-Si) SRAM-on-Logic

3DIC Technology

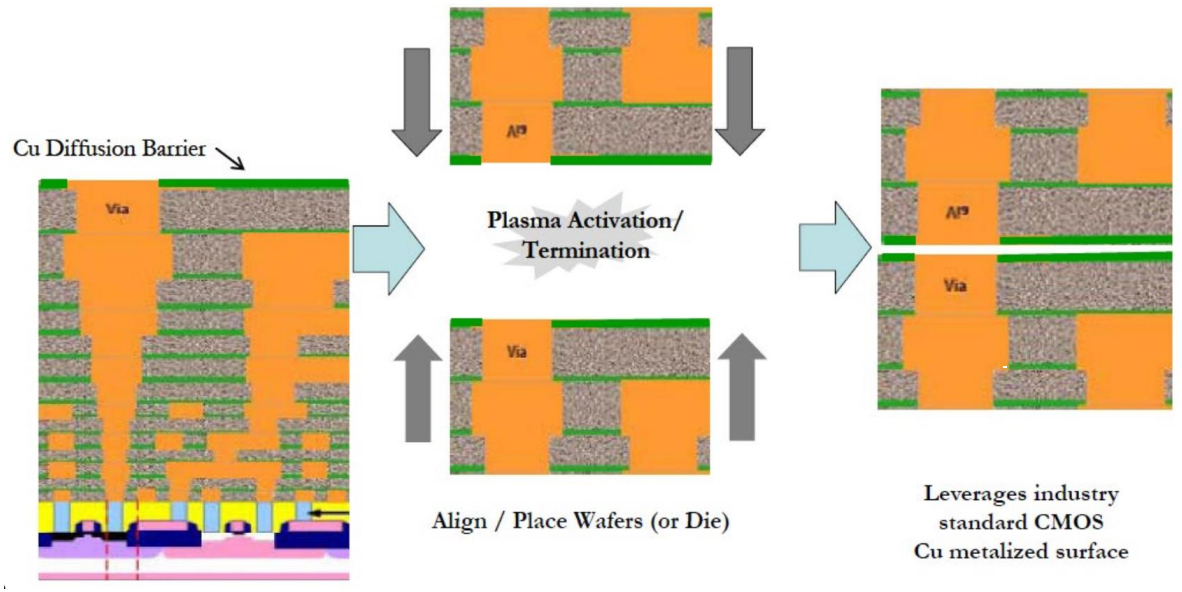


Typical Dimensions: Thinned Chip : 25 - 50 μm
Unthinned chip : 300 μm
Microbump pitch : 25 μm +
Direct Bond pitch : 1 μm +
Through Silicon Via (TSV) Diameter : 5 – 10 μm
TSV pitch : 25 μm +

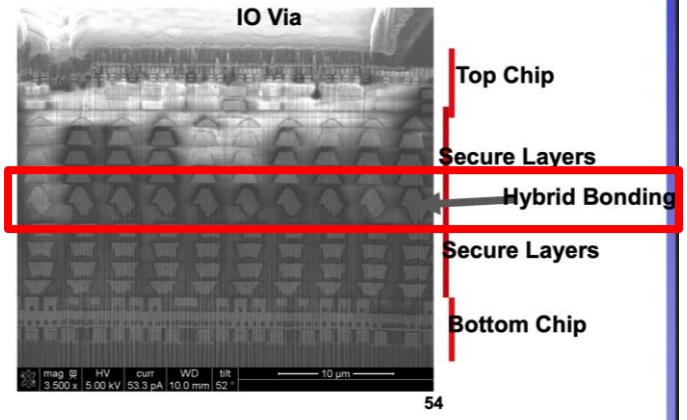
Reference: Paul Franzon, 3D-IC,2021

- **Smaller possible sizes for thinner wafer.**

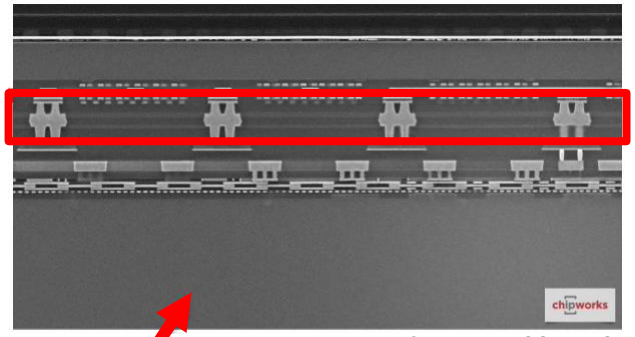
Hybrid Bonding



Reference: Paul Franzon, 3D-IC,2021



Reference:Nigussie, Franzon, 2021



Reference: Chipwork

- **Plasma active bonding**
- **1.6 – 10 µm pitch in products today, used in many cell phone cameras (6 µm pitch)**

Solder bonding Vs. Hybrid Bonding

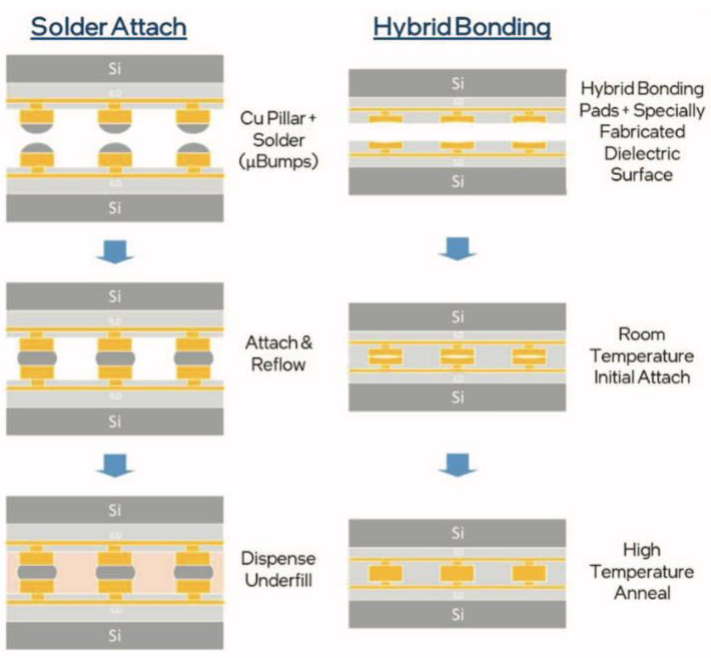
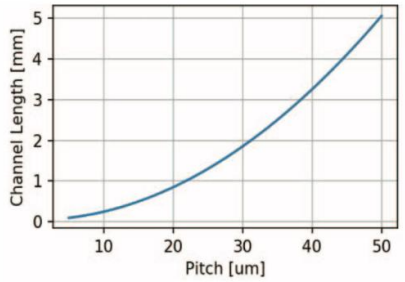


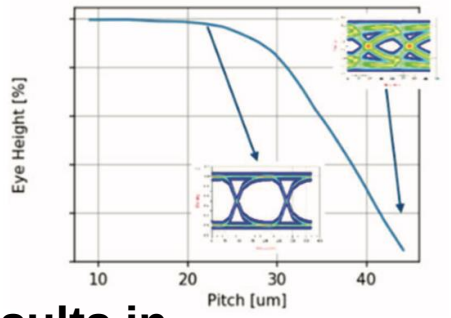
Figure 2: Hybrid Bonding vs. Solder Attach.

Table 1: Area consumed by die to die IO connections at different interconnect pitch values (40 µm – 10 µm).

	Current Generation Processing Core 1X Bandwidth	Next Generation Processing Core 1X Bandwidth	Next Generation Processing Core 2X Bandwidth
IO Area % @ 40 µm	33%	66%	Not Possible
IO Area % @ 20 µm	8%	16%	33%
IO Area % @ 10 µm	2%	4%	8%



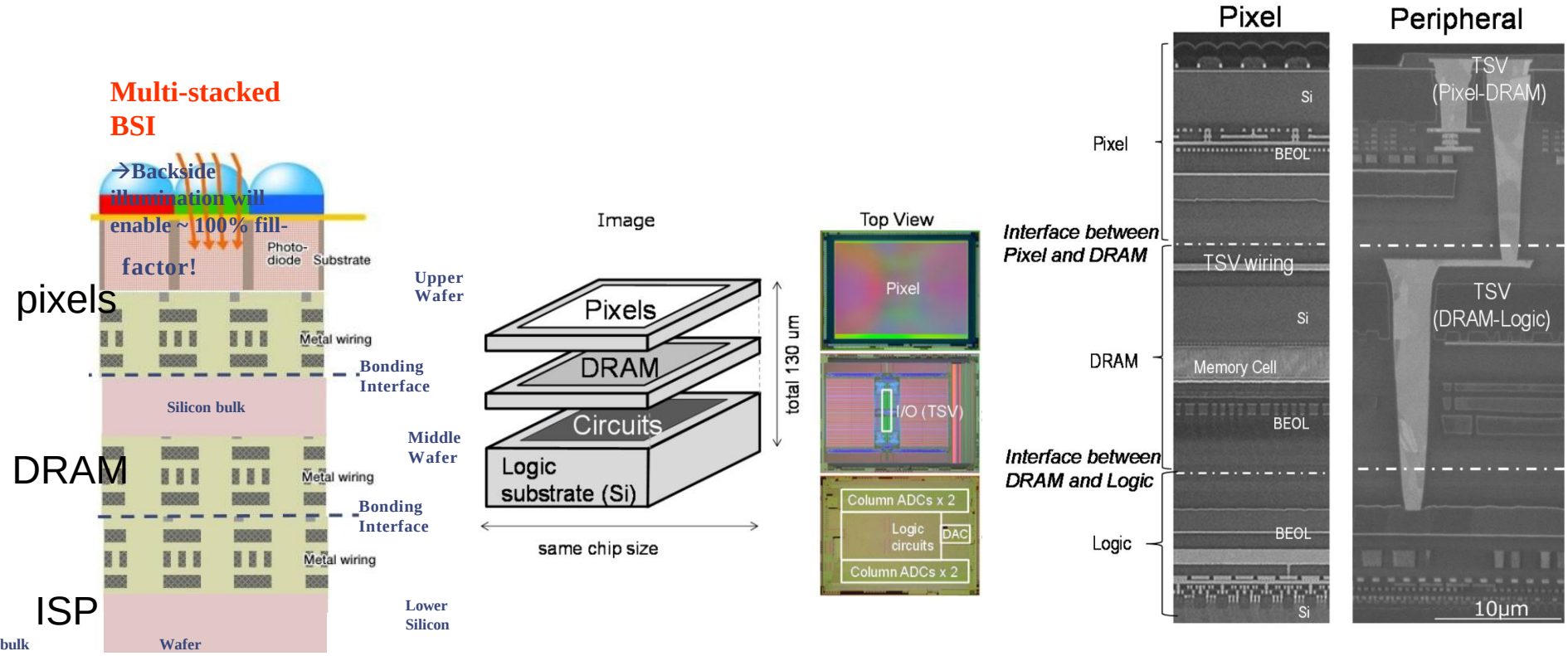
(a)



- **The larger area consumed by the IO connections results in core power integrity**
- **Interconnect pitch value ↑,channel length ↑.**
- **Interconnect pitch value ↑,the channel energy efficiency ↓.**

Referance: Adel Elsherbini, et al., ECTC2021 (intel)

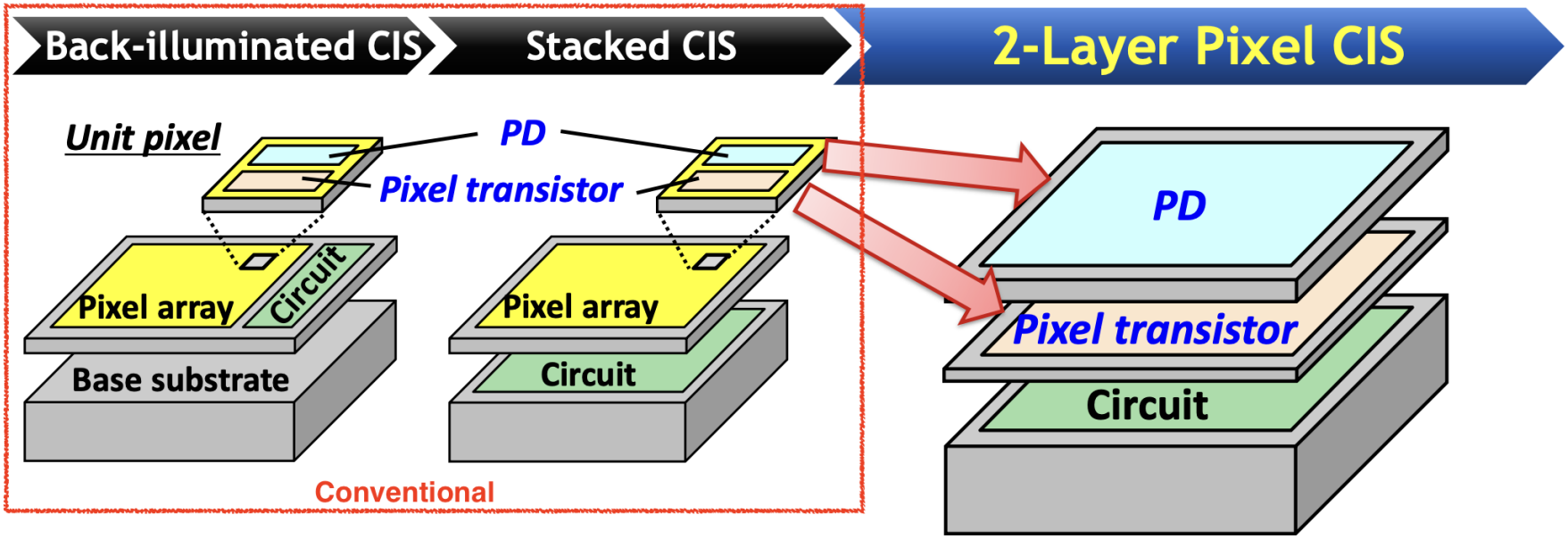
Sony 3-layer multi-stacked CIS Structure



- **3-layer stacked CIS of Sony include BSI, DRAM and logic using TSV.**

2-Layer Transistor Pixel Stacked CIS

2-Layer Pixel

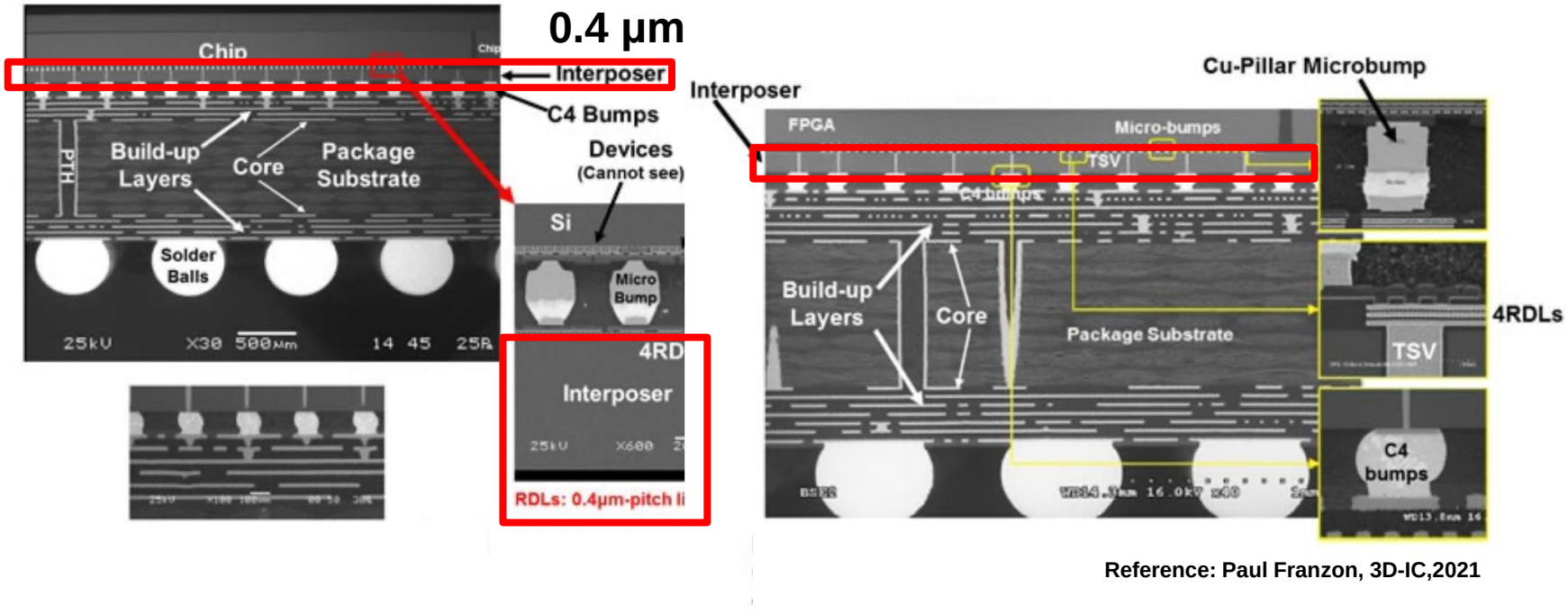


- ✓ 2-Layer Pixel has PD and pixel transistor on different silicon layers.
- ✓ We can increase PD volumes and FWC.

K.Zaitsu et. al., VLSI 2022, Sony

Interposers

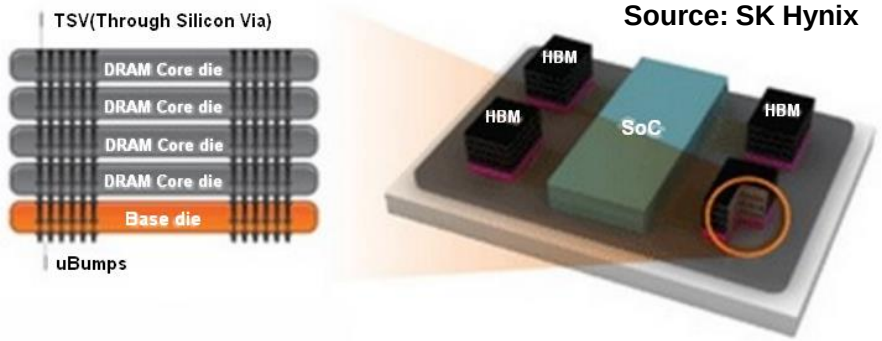
Xilinx



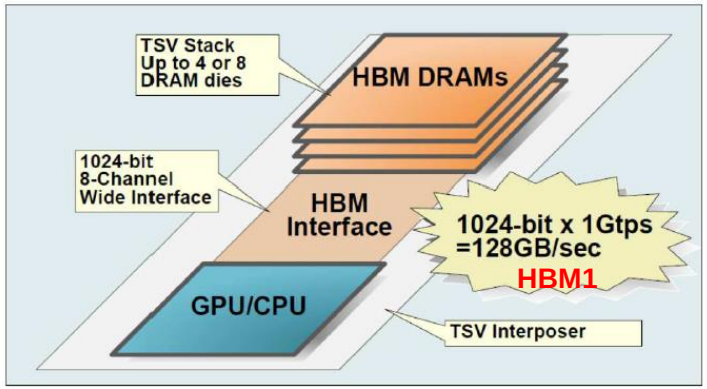
Reference: Paul Franzon, 3D-IC,2021

- Xilinx (TSMC) uesd 0.4 µm W/S interposer.
- High density silicon interposers are expensive.

High Bandwidth Memory



Source: SK Hynix

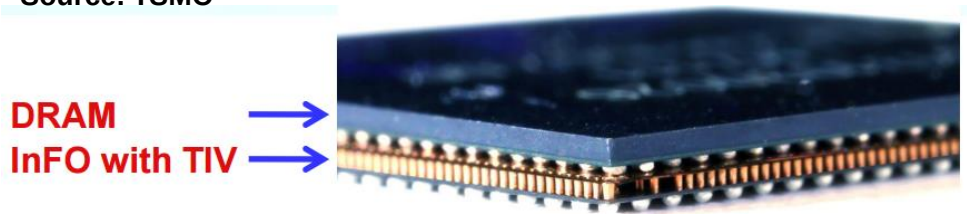


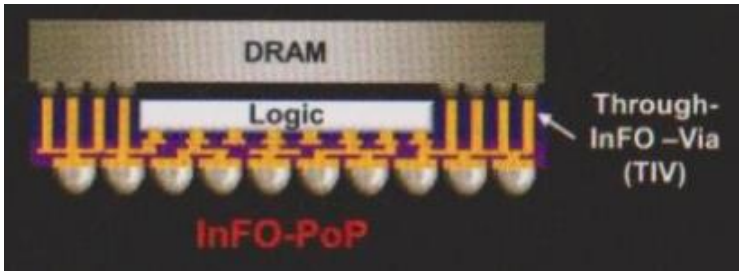
(256GB/sec for HBM2)

- Stacked DRAM dies, connected by TSVs
- Size reduction, but high density

DRAM on Application Processor

Source: TSMC

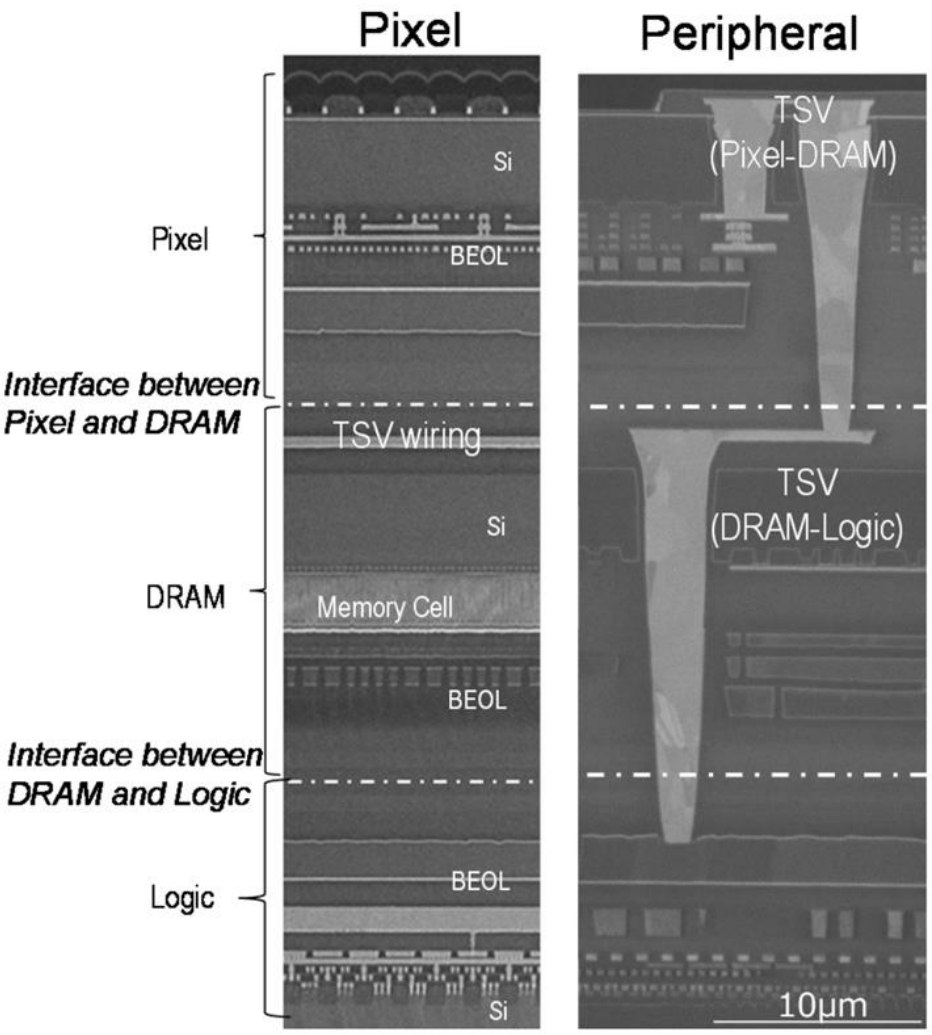




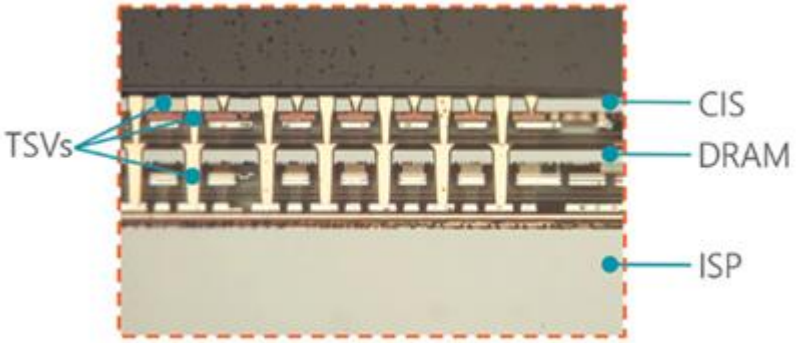
- **Vertical interconnects in molding compound (TIV)**
- **Stacked DRAM dies on application processor (AP)**
- **No additional substrate**

→ Low cost

3-layer stacked CIS



Fabrication Process	90 nm 1AL 5Cu CIS 30 nm 3AL 1W DRAM 40 nm 1AL 6Cu Logic
Number of effective pixels	5520 (H) x 3840 (V) 21.2 M pixels
Image Size	Diagonal 7.73 mm (Type 1/2.3)
Pixel Size	1.22 μm (H) x 1.22 μm (V)
Frame Rate (Still)	30fps at 4:3 19.3 Mpix
Pixel Readout Speed	8.48 msec at 19.3 Mpix (4:3)
Supply Voltage	2.5 V / 1.8 V / 1.1 V
Outputs	MIPI (CSI2) DPHY 2.2 Gbps/lane or CPHY 2.0 Gsps/lane
DRAM Size	1 Gbit



Ref: SONY, ISSCC 4.6, 2017.

- **Sony use TSV to connect CIS (TSV), DRAM (TSV), and ISP**

https://www.systemplus.fr/wpcontent/uploads/2017/07/SP17343_Sony_IMX400_Tri-layer_Stacked_CIS_Flyer_System_Plus_Consulti.pdf