

# The Dawn of the New Electronics Industry (Surfing on Chips Acts Waves)

Paolo Gargini

Chairman IRDS

Chairman ICOS IAB\*

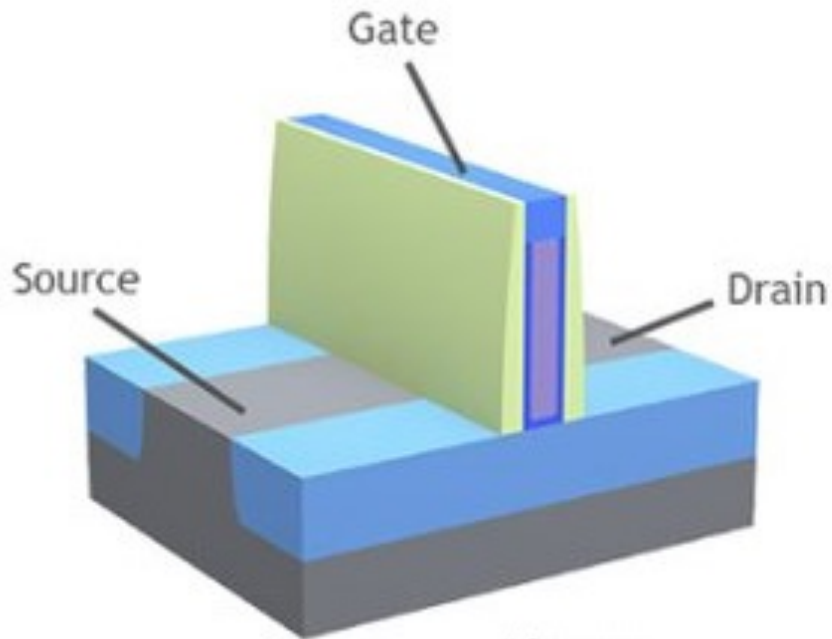
Life-Fellow IEEE, Fellow I-JSAP

\* International Advisory Board

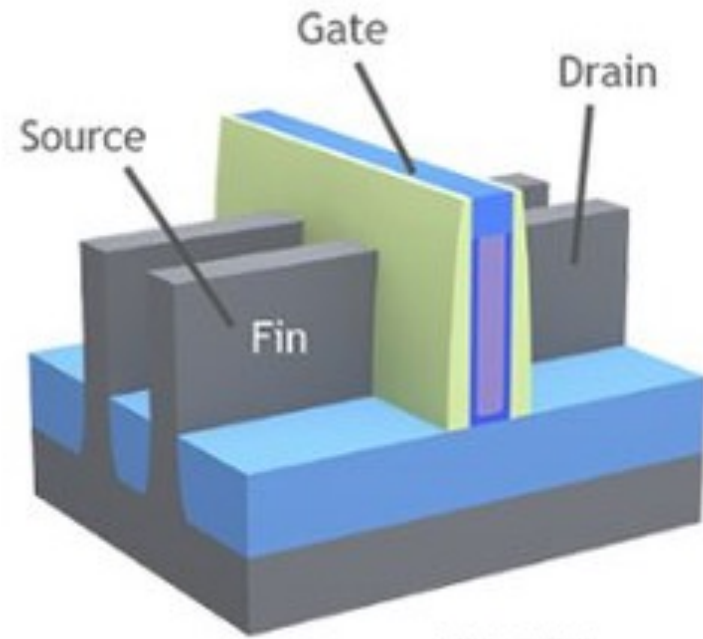
# Shockers involving Governments occur every 20 years

1. 1957: Sputnik
2. 1987: US DRAM Collapse
3. 1997: Gate oxide vanishing by 2005
4. 2020: Pandemic

# 1965-2011: From Planar Transistor to FinFET



Planar



FinFET

# *Fourth Shocker*

2024



P.Gargini

4



# Governments Investing Hundreds of \$Billions

May 2022

**ポスト5G情報通信システム基盤強化研究開発事業** October 2021 戦略情報政策局 情報産業課 03-3501-6944

**基金総額 2,000.0億円** (令和元年度補正予算額 1,100億円 + 令和2年度3次補正予算額 900億円)

**事業の内容**

**事業目的・概要**

- 第4世代移動通信システム(4G)と比べてより高度な第5世代移動通信システム(5G)は、現在各国で商用サービスが始まっていますが、さらに超伝送速や多数同時接続といった機能が強化された5G(以下、「ポスト5G」)は、今後、工場や自動車といった多様な産業用途への活用が見込まれており、我が国の競争力の核となり得る技術と期待されます。
- 本事業では、ポスト5Gに対応した情報通信システム(以下、「ポスト5G情報通信システム」)の中核となる技術を開発することで、我が国のポスト5G情報通信システムの開発・製造基盤強化を目指します。
- 具体的には、ポスト5G情報通信システムや当該システムで用いられる半導体を開発するとともに、ポスト5Gで必要となる先端的な半導体を将来的に国内で製造できる技術を開発するため、先端半導体の製造技術の開発に取り組みます。

**成果目標**

- 本事業で開発した技術が、将来的に我が国のポスト5G情報通信システムにおいて活用されることを目指します。

**条件(対象者、対象行為、補助率等)**

国(補助(基金造成)) → (中間評価) → 産学連携共同開発機構(NEDO) → 委託(補助(1/2)) → 民間企業・研究機関・大学等 → Industry Research University

**事業イメージ**

**(1) ポスト5G情報通信システムの開発(委託)**

- ポスト5Gで求められる性能を実現する上で、特に重要なシステム及び当該システムで用いられる半導体等の関連技術を開発。

**(2) 先端半導体製造技術の開発(補助)**

NEDO is a national research and development agency that creates innovation by promoting technological development necessary for realization of a sustainable society.

**Investments to support the Chips Act Europe**

The Chips Act itself should result in additional public and private investments of more than €15 billion.

These investments will complement:

- existing programmes and actions in research and innovation in semiconductors such as Horizon Europe and the Digital Europe programme
- announced support by Member States

In total, more than €43 billion of policy-driven investment will support the Chips Act until 2030, which will be broadly matched by long-term private investment.

The Chips Act proposes:

- Investments in next-generation technologies
- Providing access across Europe to design tools and pilot lines for the prototyping, testing and experimentation of cutting-edge chips
- Certification procedures for energy-efficient and trusted chips to guarantee quality and security for critical applications
- A more investor-friendly framework for establishing manufacturing facilities in Europe
- Support for innovative start-ups, scale-ups and SMEs in accessing equity finance
- Fostering skills, talent and innovation in microelectronics
- Tools for anticipating and responding to semiconductor shortages and crises to ensure security of supply
- Building semiconductor international partnerships with like-minded countries

## CHIPS Act At A Glance

### Creating Helpful Incentives to Produce Semiconductors for America (CHIPS Act)

\$52 billion total budget over 5 years

**Financial Incentives Programs**

\$39 billion

**Research and Development**

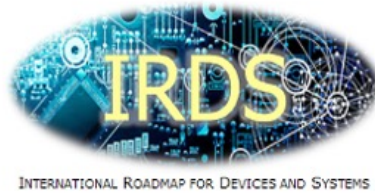
\$11 billion  
Technology Center  
Packaging Program  
MFG USA Institute(s)  
Metrology program

**Workforce Development**

US Code: 15 USC 4652: Semiconductor Incentives



2016



# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- New Products
- New Architecture
- SOC: Endless Die Size
- SOC and SIP



# ITRS, ITRS 2.0, and IRDS FORECAST

## 3D Power Scaling

2015->2025-2040



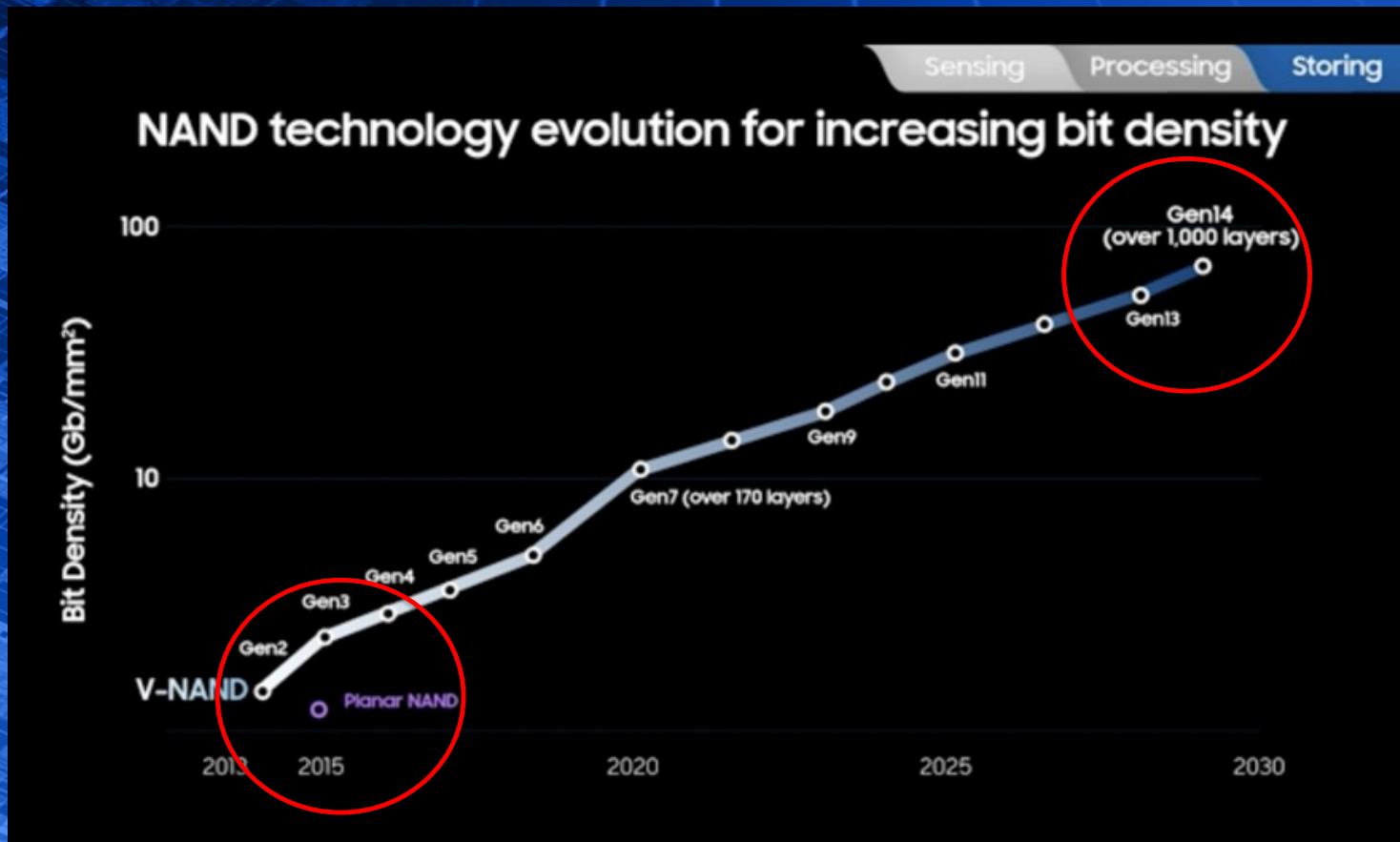
MEMORY



LOGIC



# Samsung, IEDM 2021



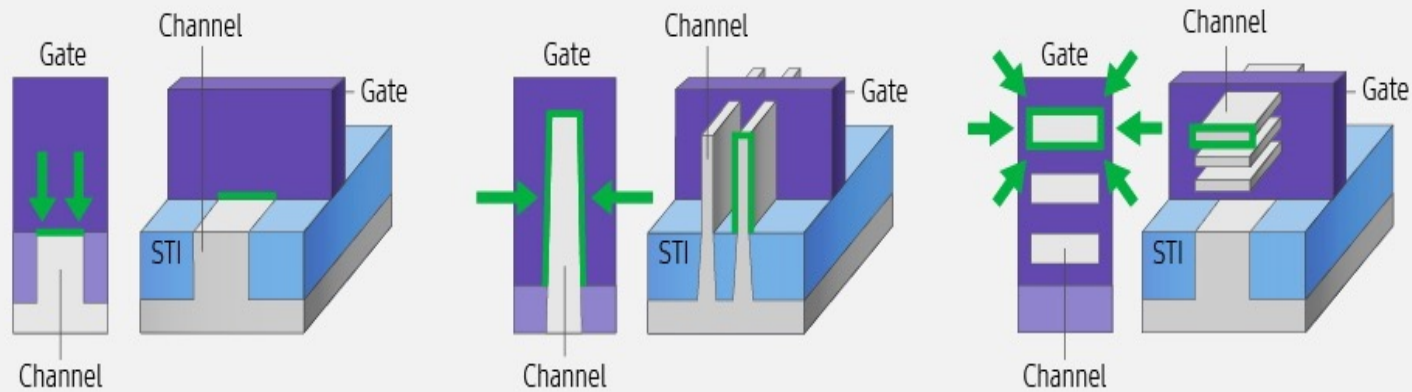
# ITRS, ITRS 2.0, and IRDS FORECAST

## 3D Power Scaling

2015->2025-2040



## 3 Steps to Surrounding the Channel



**Planar FET**

1 Gate on channel

**FinFET**

3 Gates on channel

**Gate-All-Around**

4 Gates on channel

# 2018 IRDS

YEAR OF PRODUCTION	2018	2020	2022	2025	2028	2031	2034
Logic industry "Node Range" Labeling (nm)	G54M36	G48M30	G45M24	G42M21	G40M16	G40M16T2	G40M16T4
Logic industry "Node Range" Labeling (nm)	"7"	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
IDM-Foundry node labeling	i19-i7	i7-55	i5-i3	i3-i2.1	i2.1-i1.5	i1.5e-i1.0e	i1.0e-i0.7e
Logic device structure options	FinFET	finFET	finFET	LGAA	VGAA	VGAA	VGAA
Mainstream device for logic	finFET	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
<b>LOGIC DEVICE GROUND RULES</b>							
Mx pitch (nm)	40	36	32	24	20	16	16
M1 pitch (nm)	36	32	30	21	20	20	20
M0 pitch (nm)	36	30	24	21	16	16	16
Gate pitch (nm)	54	48	45	42	40	40	40
L <sub>g</sub> - Gate Length - HP (nm)	20	18	16	14	12	12	12
L <sub>g</sub> - Gate Length - MD (nm)	22	20	18	14	12	12	12
Channel overlap ratio - two-sided	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Spacer width (nm)	8	7	6	6	6	6	6
Contact CD (nm) - finFET, LGAA	18	16	17	16	16	16	16
Contact CD (nm) - VGAA							
<b>Device architecture key ground rules</b>							
FinFET pitch (nm)	32.0	28.0	24.0				
FinFET Fin width (nm)	8.0	7.0	6.0				
FinFET Fin height (nm)	40	50	60				
Footprint drive efficiency - finFET	2.75	3.32	5.25				
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				25	20	15	10
LGAA width (nm) - MD				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				4.80	4.50	5.52	5.00
Device effective width (nm) - HP	88.0	107.0	126.0	192.0	156.0	160.0	120.0
Device effective width (nm) - MD	88.0	107.0	126.0	132.0	102.0	88.0	88.0
Device lateral pitch (nm)	32	28	24	22	20	20	20
Device height (nm)	40.0	50.0	60.0	53.0	48.0	57.0	57.0
Device width (nm) - HP	8	7	6	25	20	15	10
Device width (nm) - MD	8	7	6	15	11	6	6
Device width (nm) - SRAM	8	7	6	7	6	6	6

Acronyms used in the table (in order of appearance): FDSOI—fully-depleted silicon-on-insulator, LGAA—lateral gate-all-around-device (GAA), VGAA—vertical GAA, 3DVLSI—fine-pitch 3D logic sequential integration.



# ITRS, ITRS 2.0, and IRDS FORECAST

## 3D Power Scaling

2015->2025-2040



# What Comes Next?

2024

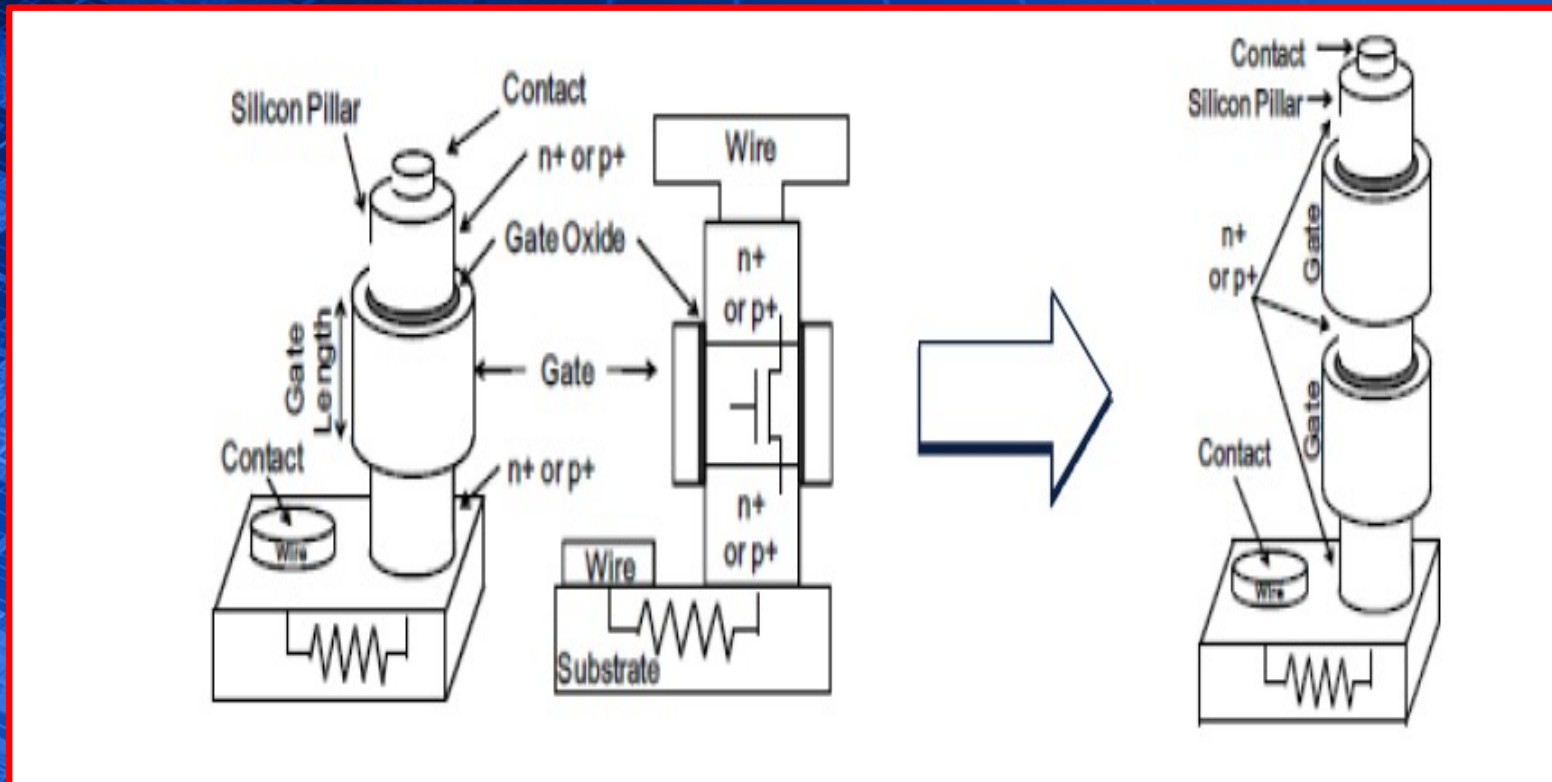


P.Gargini

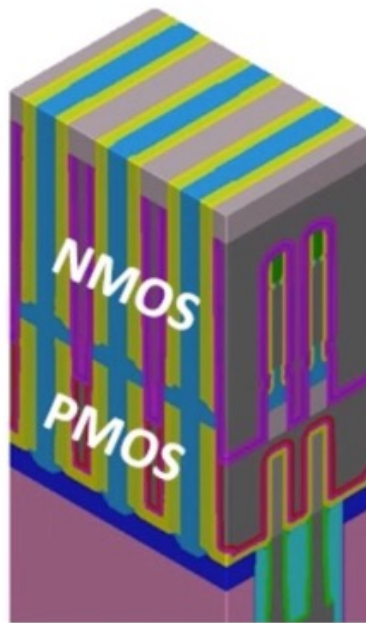
16



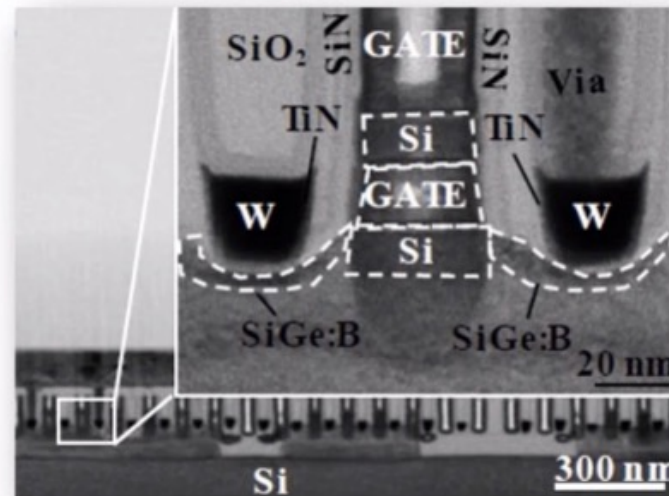
# Vertical Logic Architecture 2015



# Complementary FET (CFET)



CFET



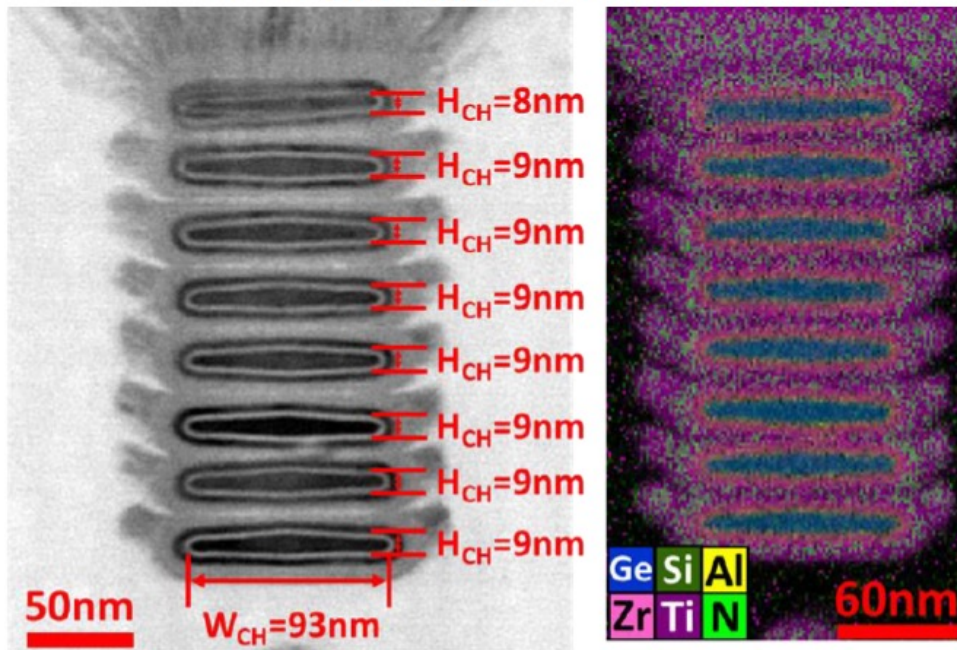
Monolithic Approach

S. Subramanian, VLSI 2020

# Highly Stacked Channels

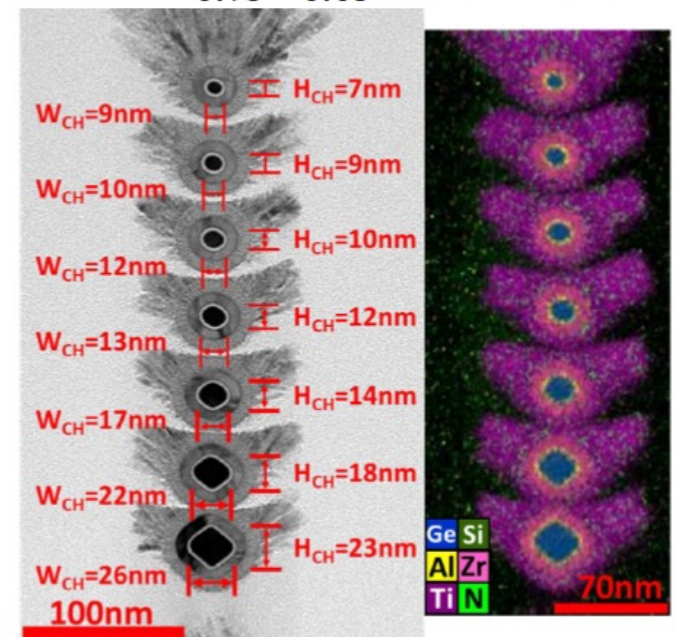
$$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} V_{OV}^2 : \text{Highly stacked channels}$$

8 Ge<sub>0.75</sub>Si<sub>0.25</sub> Nanosheets



Yi-Chun Liu *et al.*, *VLSI*, 2021, T15-2.

7 Ge<sub>0.95</sub>Si<sub>0.05</sub> Nanowires



Yi-Chun Liu *et al.*, *VLSI*, 2021, T15-2.

- The higher number of stacked channels can further improve the  $I_{ON}$  for the fixed footprint.

# The New Electronics Industry

- 3D Transistors
- **Beyond Silicon**
- 3D Everything
- System In Package
- New Products
- New Architecture
- SOC: Endless Die Size
- SOC and SIP

# GAA Beyond Silicon

2024

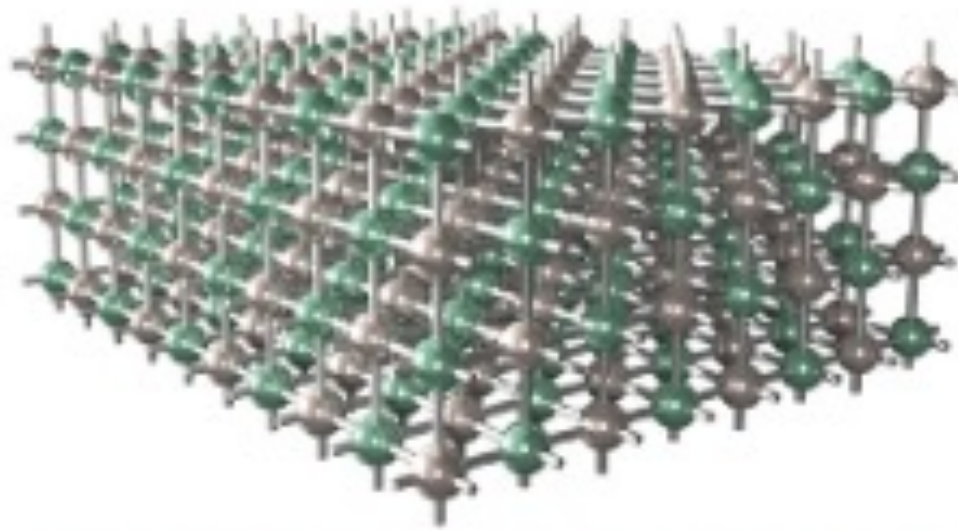


P.Gargini

21

# Intrinsic Challenges of Bulk Material

Bulk materials

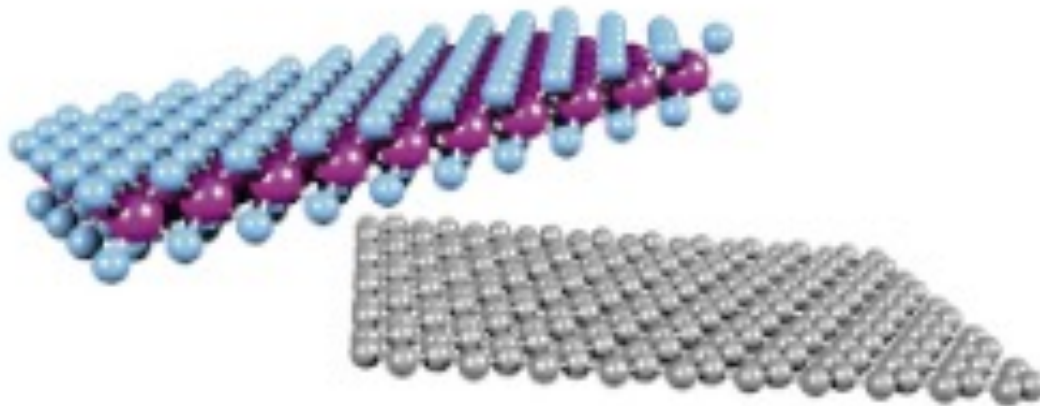


Dangling bonds

Rough surface

# 2D Materials

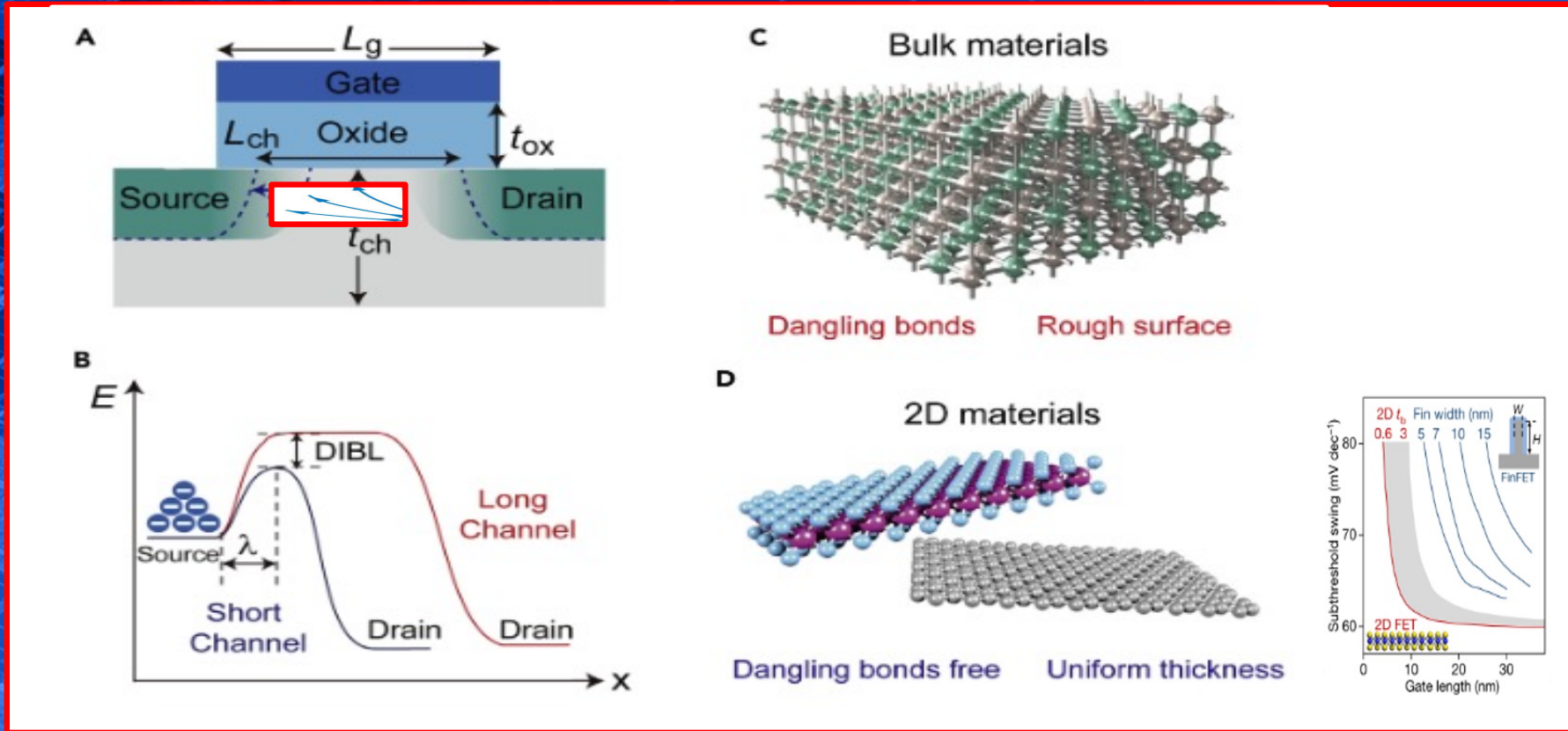
2D materials



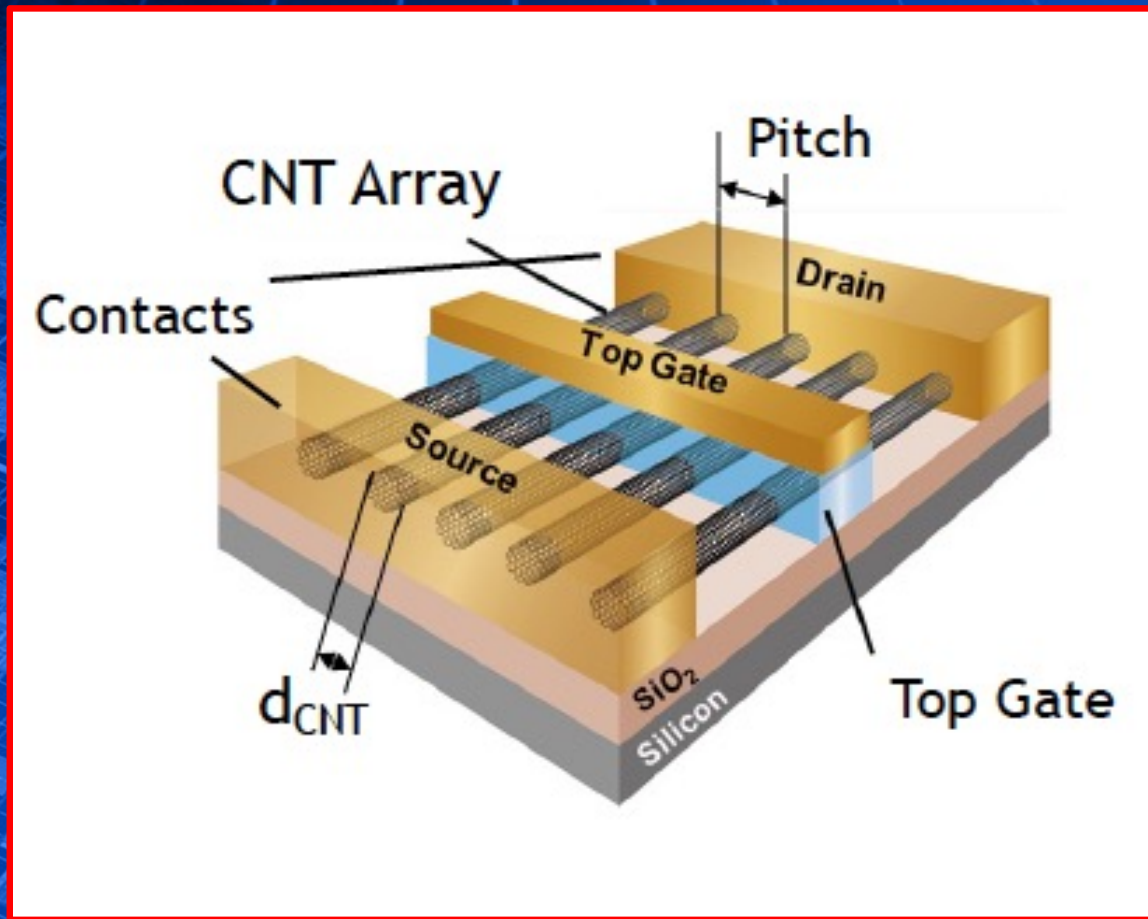
Dangling bonds free

Uniform thickness

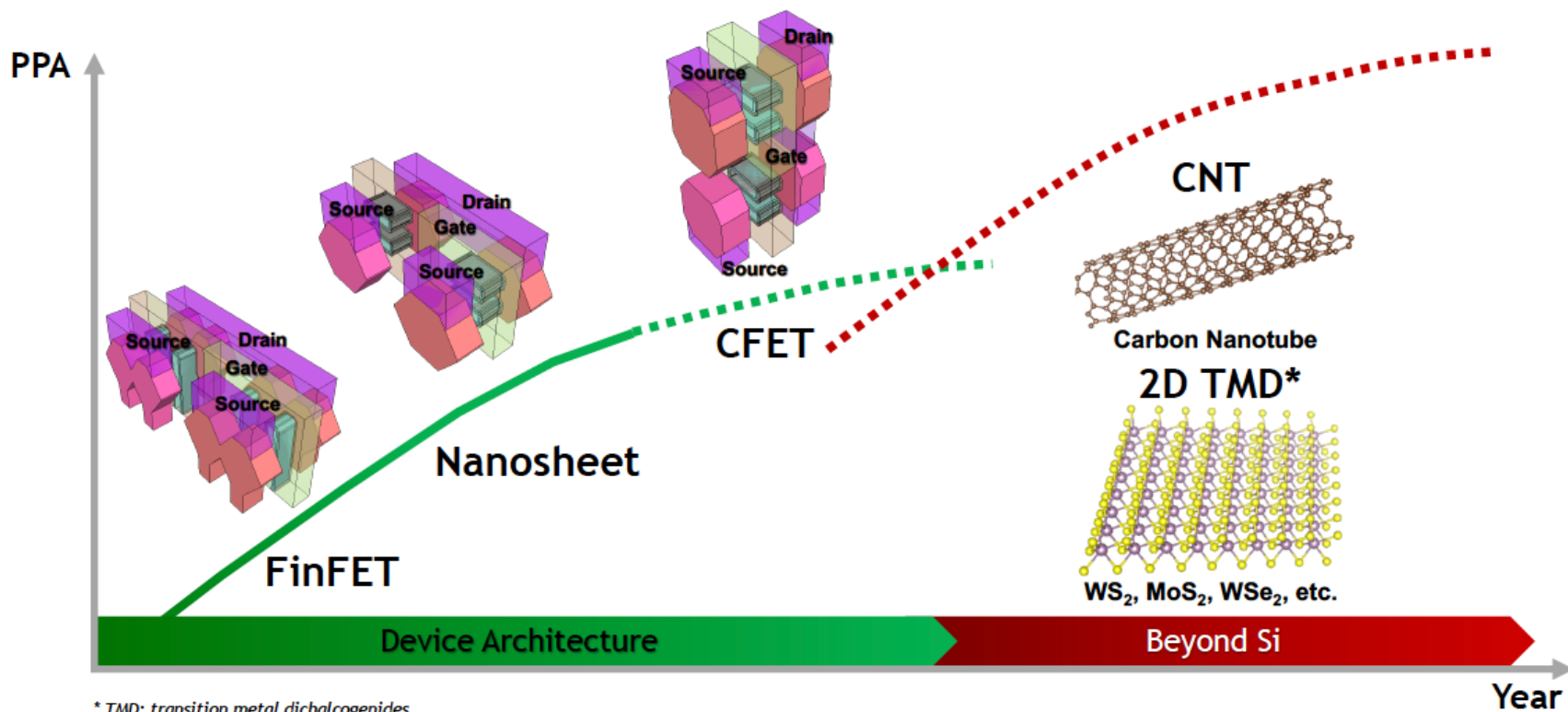
# Towards Elimination of Dangling Bonds







# Device Architecture Outlook



# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- New Products
- New Architecture
- SOC: Endless Die Size
- SOC and SIP

# 3D Everything

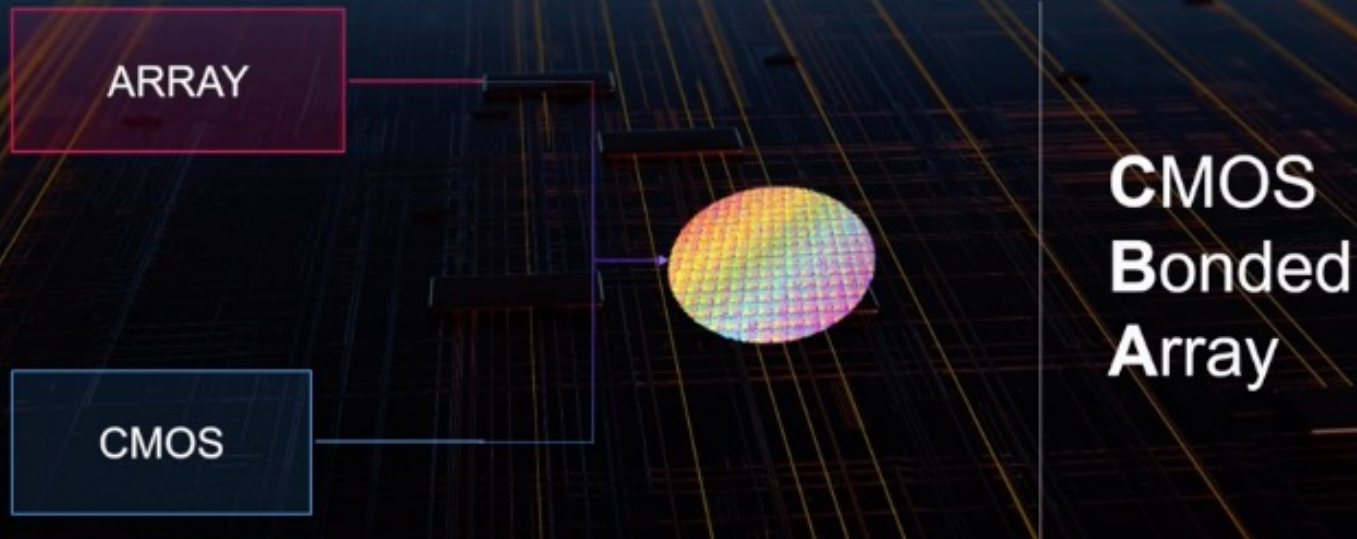
2024



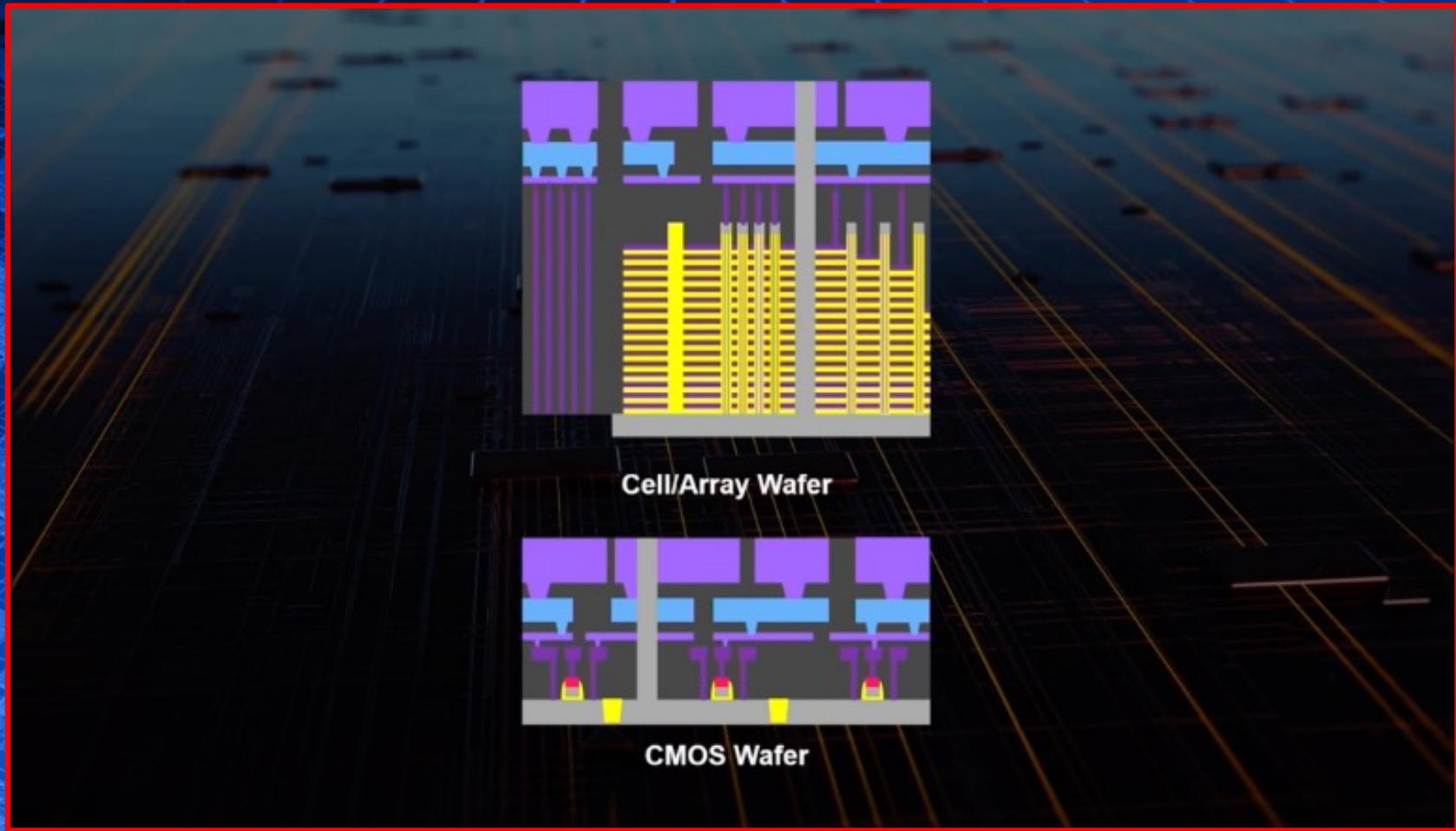
P.Gargini

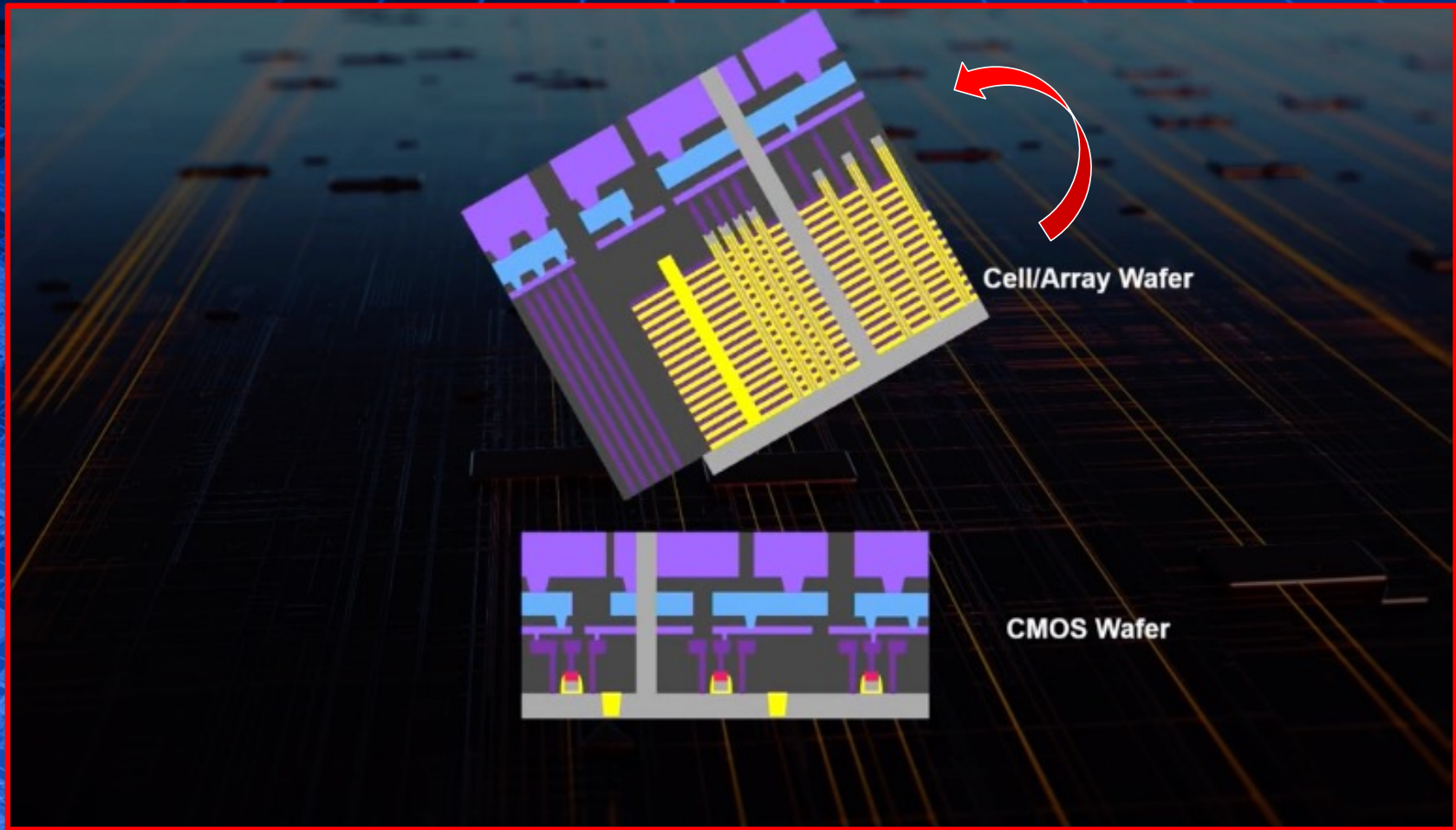
28

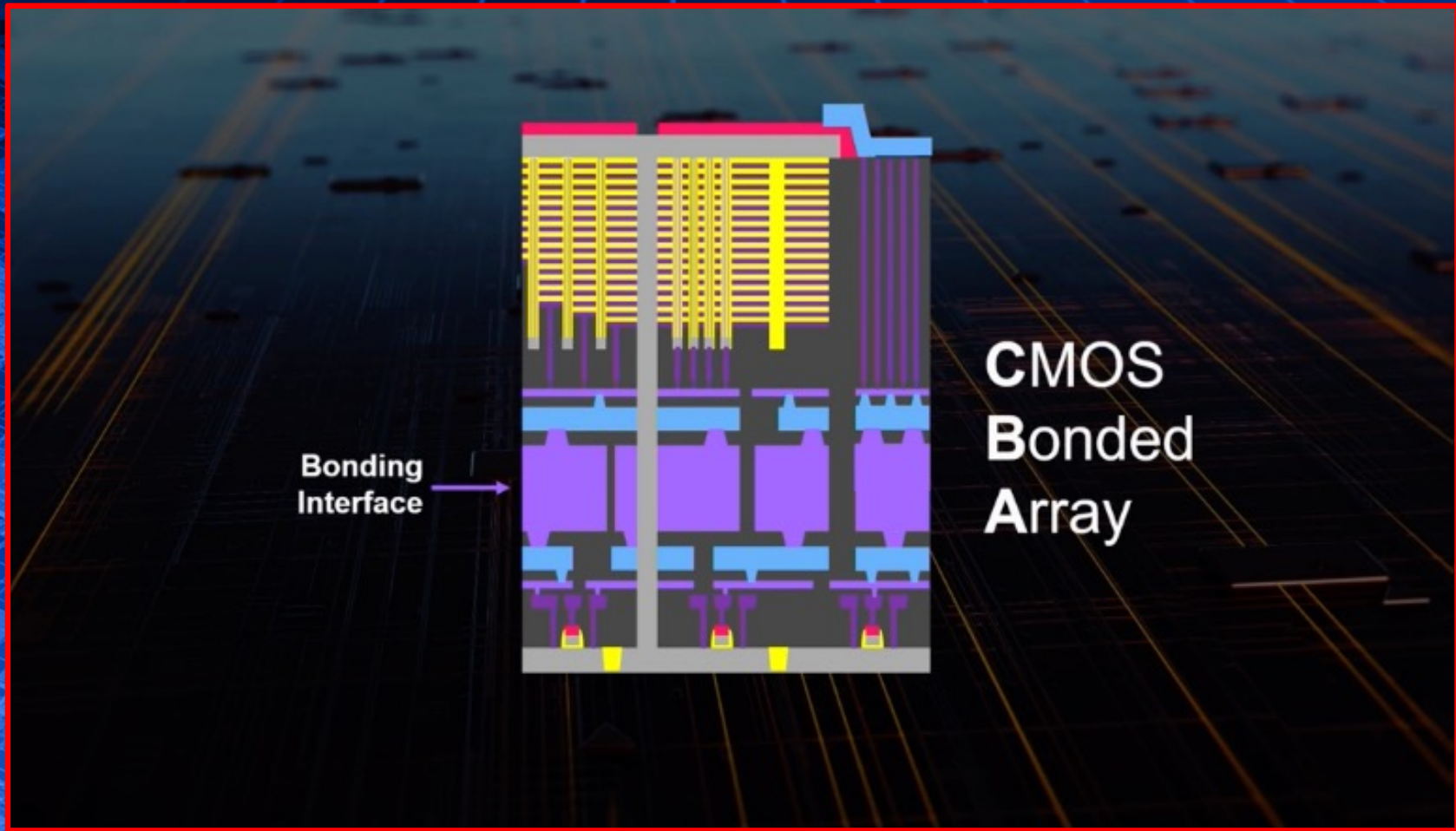
## Vertically Stacked NAND Challenge



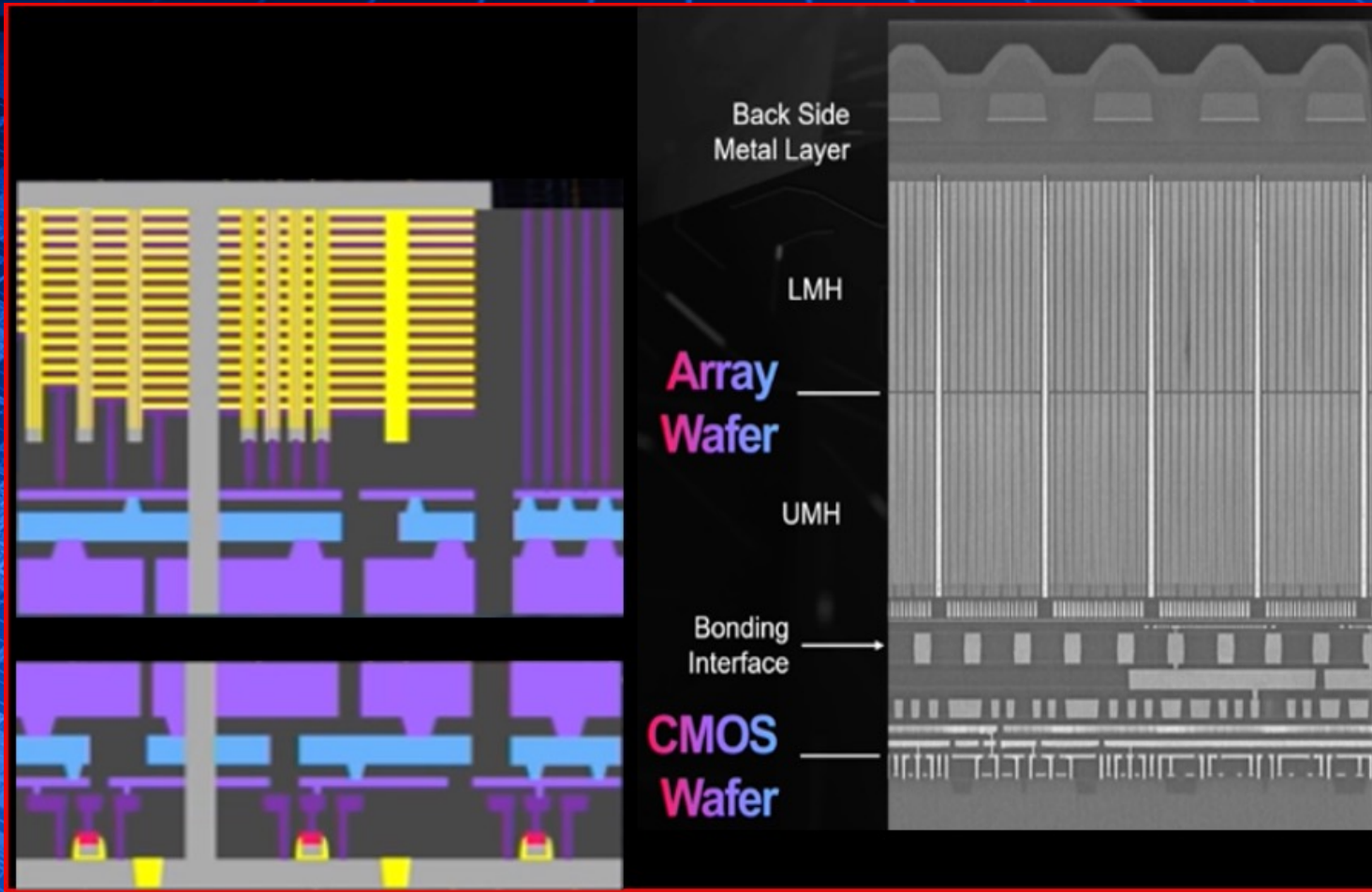
**CMOS circuits are not affected by high temperature Array processes**



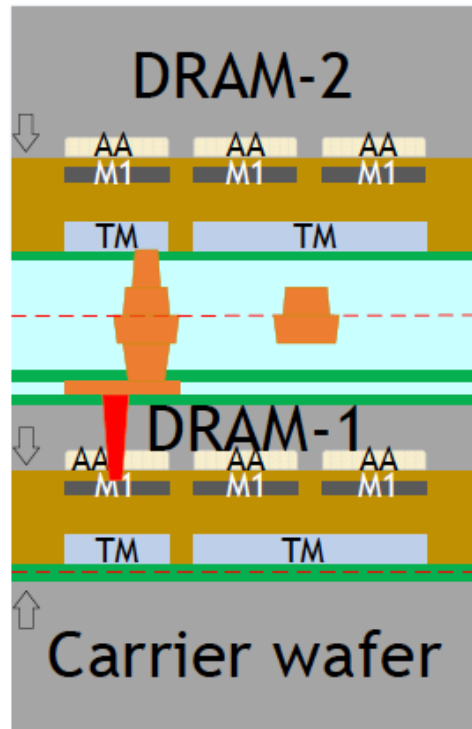




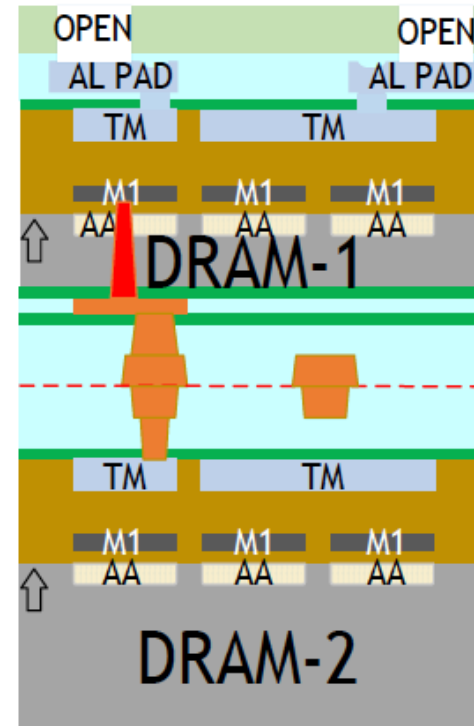




# Process of Proposed SeDRAM

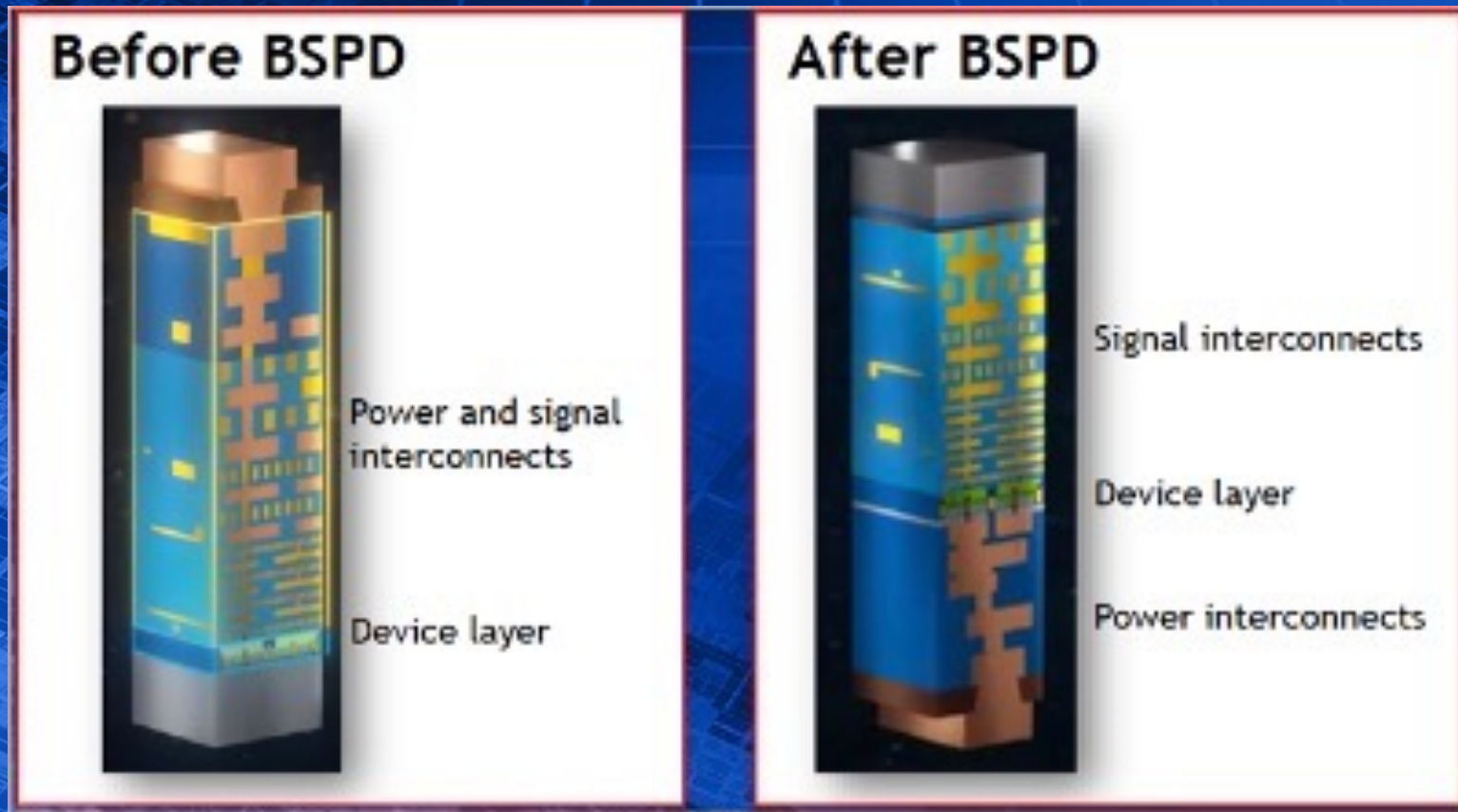


4. DRAM-1 and DRAM-2 bonding



5. Remove carrier, and then pad out on DRAM-1

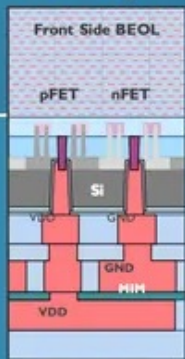
# Back Side Power Distribution



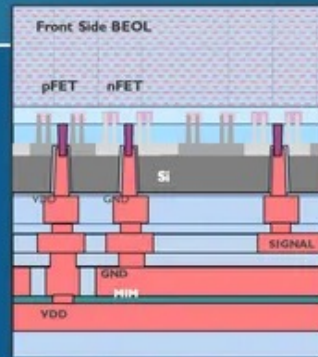
Source: 2023 VLSI Symposium, Intel

# Backside PDN paving the way to a truly functional backside

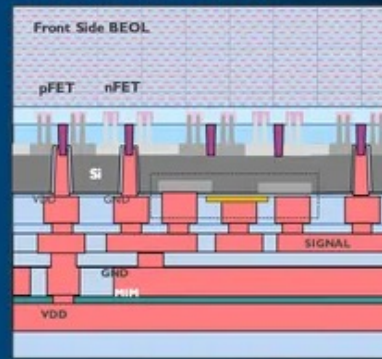
## Backside Power Delivery



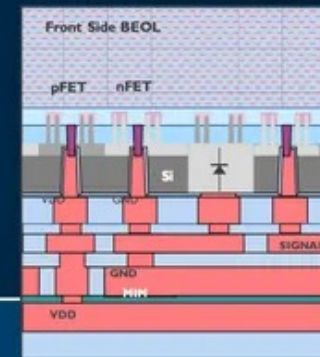
## Backside Global Interconnect



## Backside Devices



## Device Backside Extension



*Enhancing system performance by migrating system functions to the backside*

imec

ITF WORLD

# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- **System In Package**
- New Products
- New Architecture
- SOC: Endless Die Size
- SOC and SIP

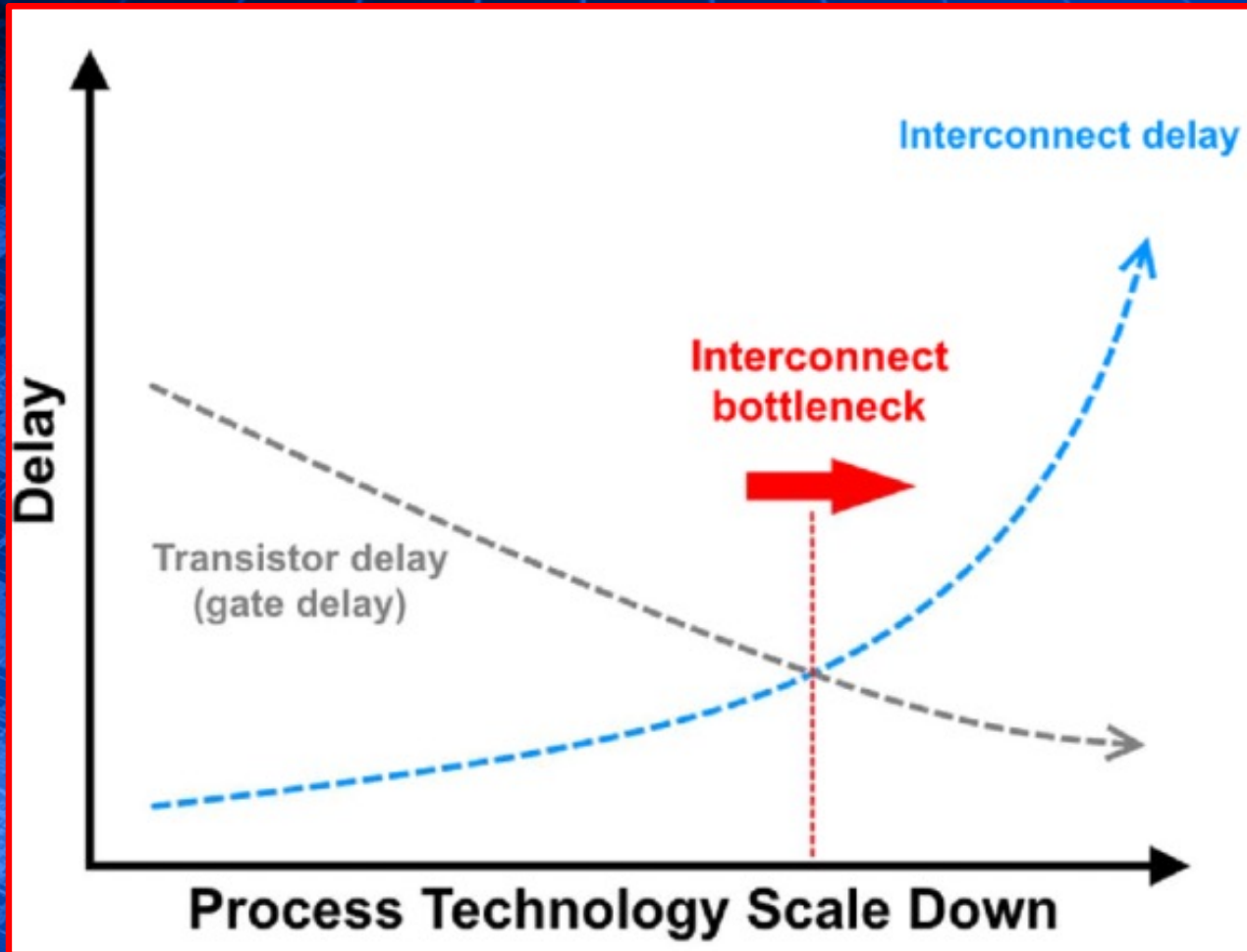
# System In Package SIP

2024



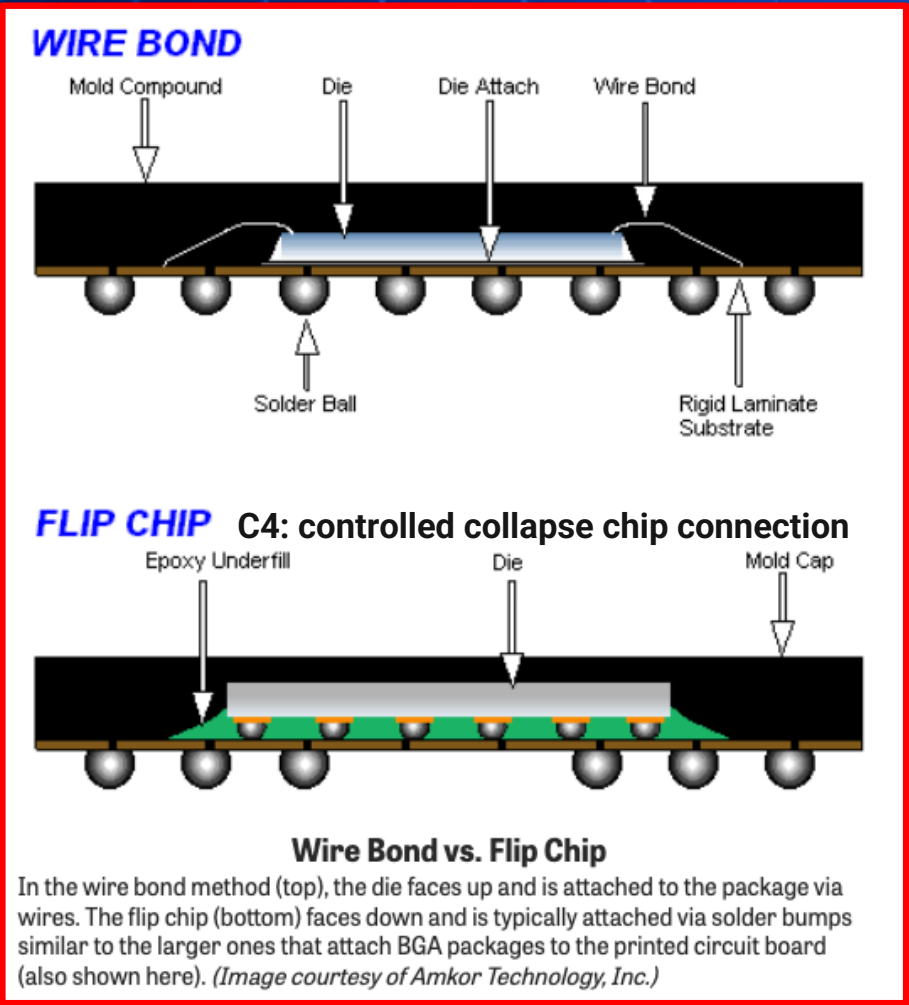
P.Gargini

38



# Develop a Technology to Produce an Infinitely “Large Die”

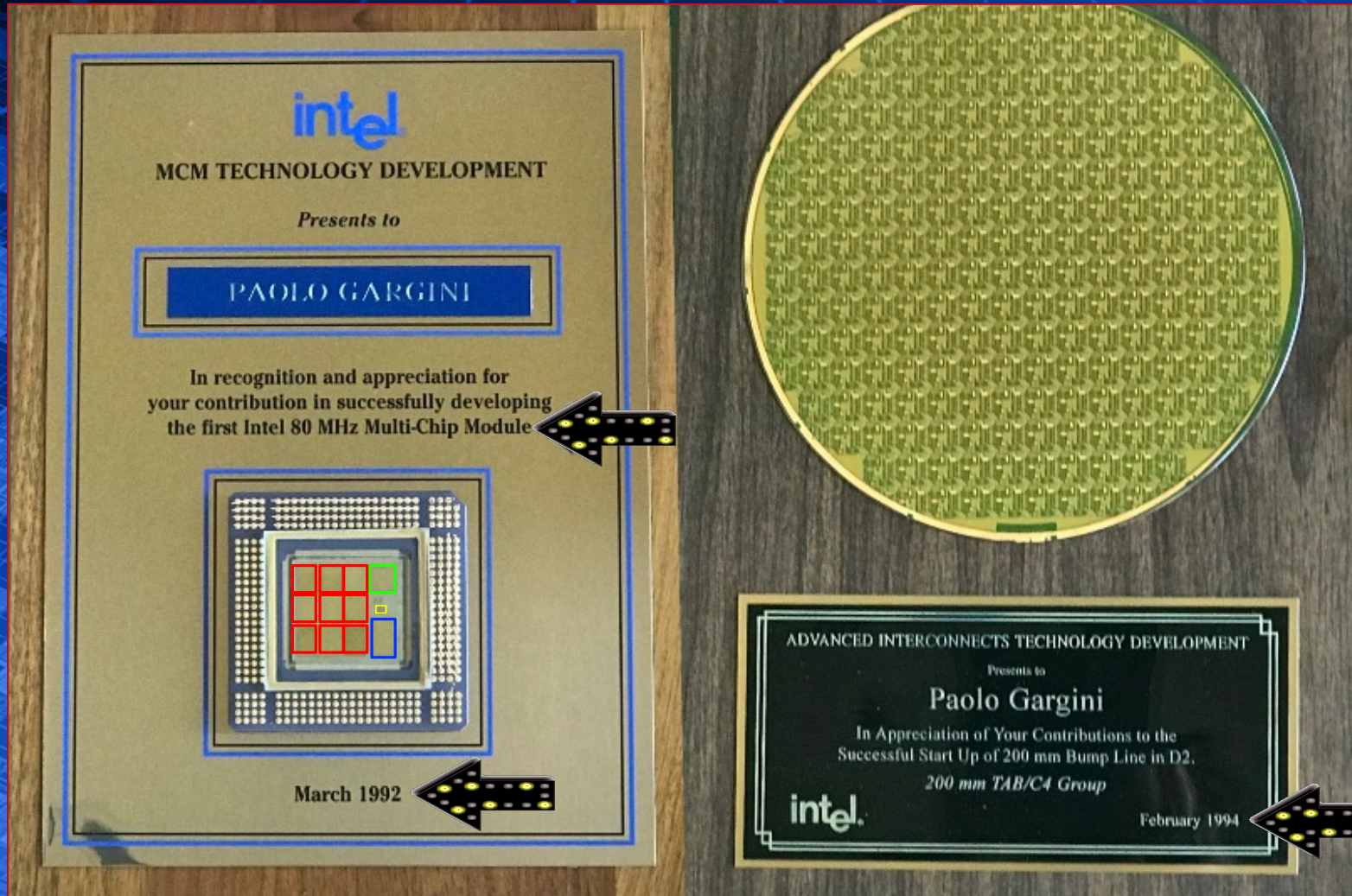
Melting point  
Pb  
327°C



Melting point  
Sn  
231°C

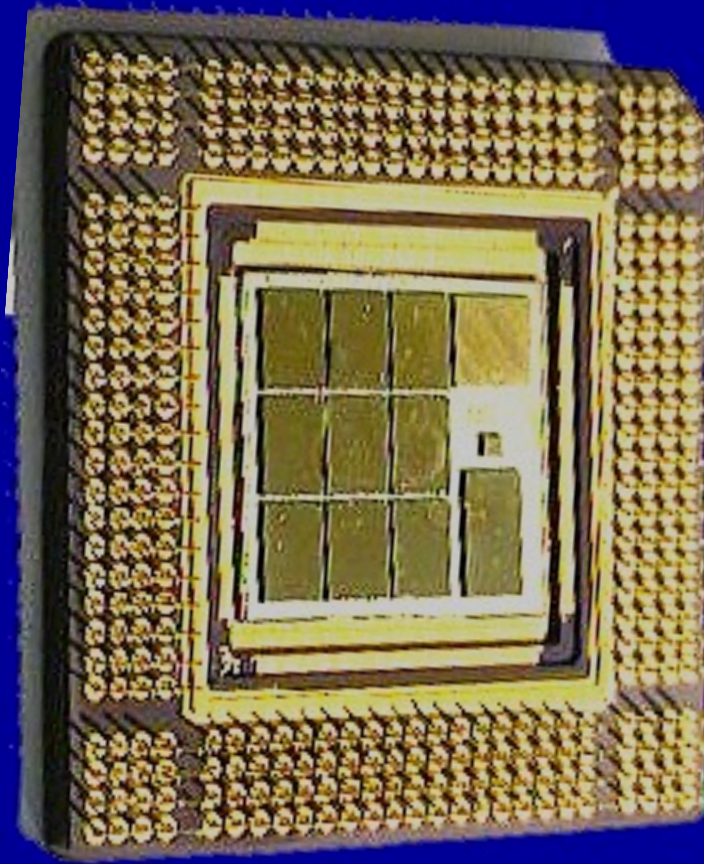


From co-evaporation to electrolytic deposition from solutions of metal salts



# MCM or Board?

DRAMs



SRAM

Voltage Regulator

80486

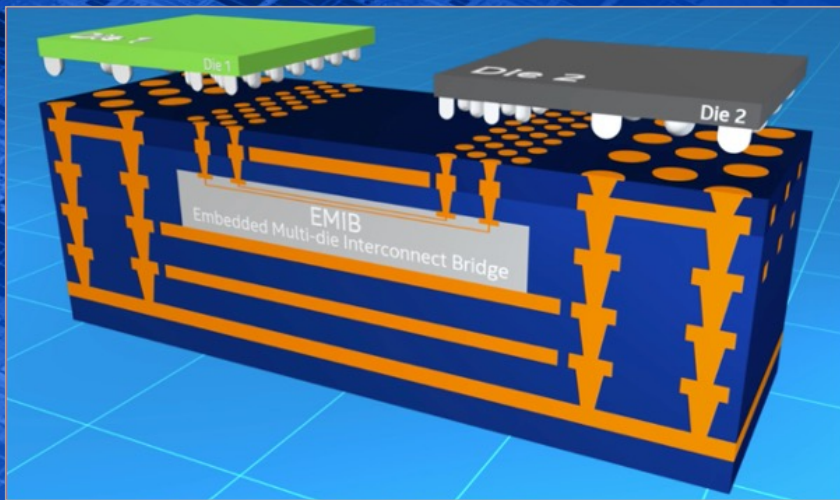
# Intel Publications and Patents Guidelines

1968-1996

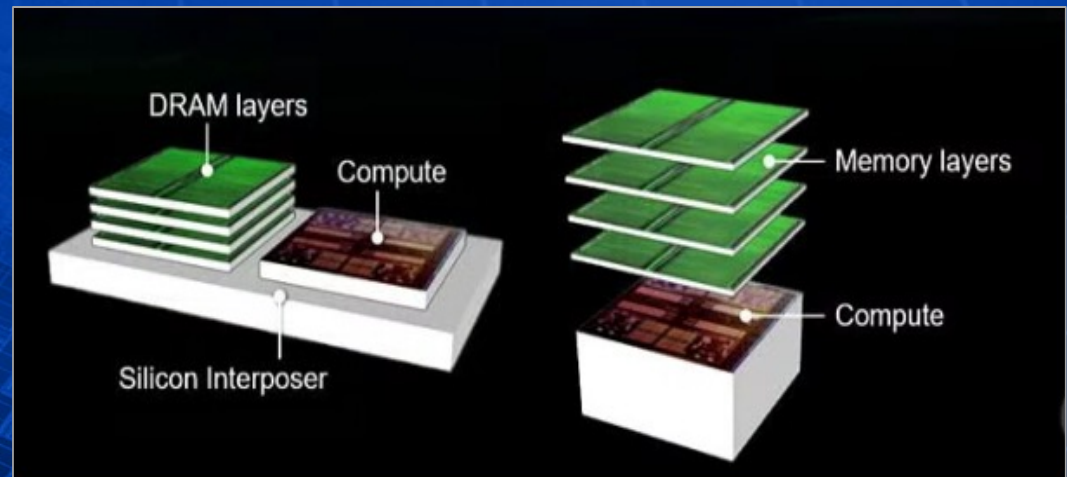
*Write it on Silicon  
NOT on Paper*

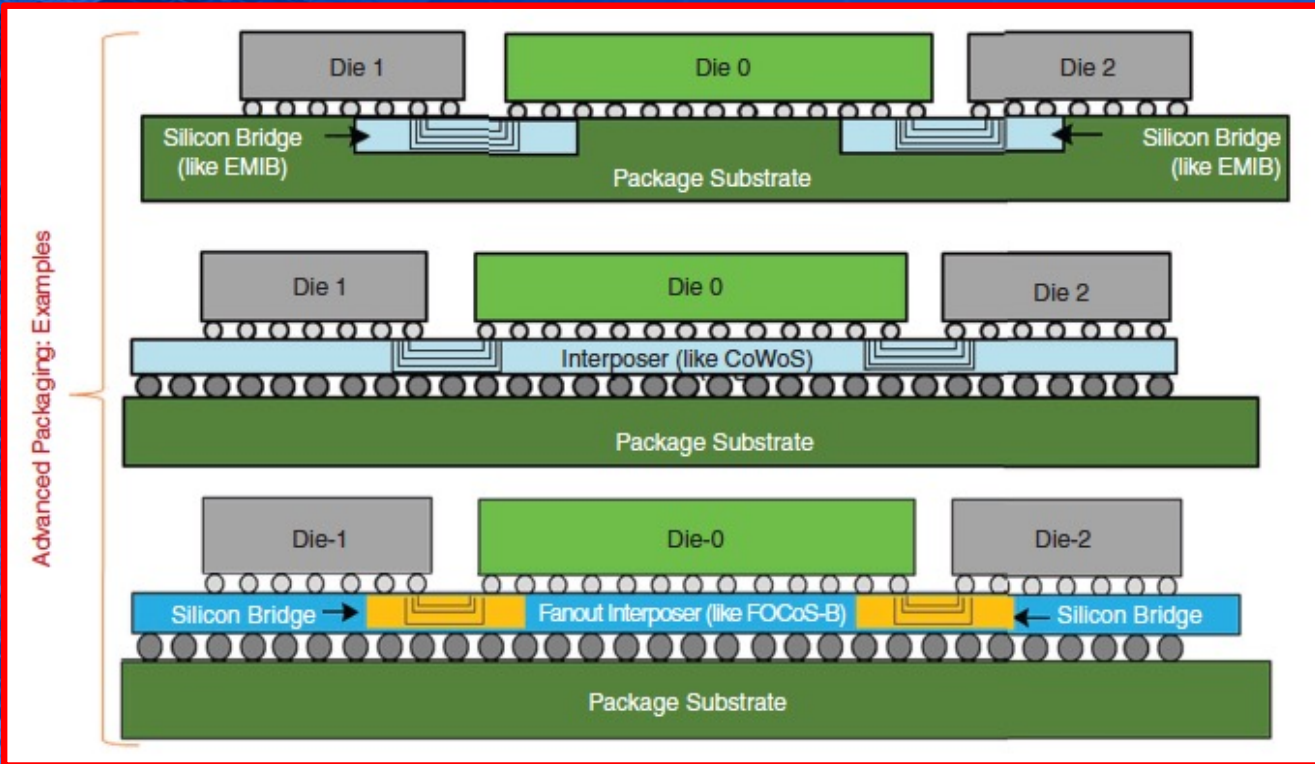
*Gordon Moore*

## The Horizontal Approach

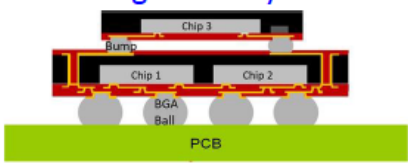
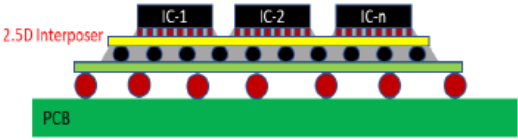
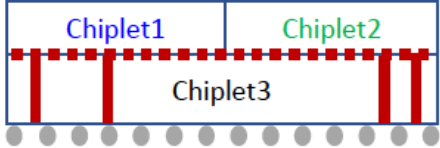


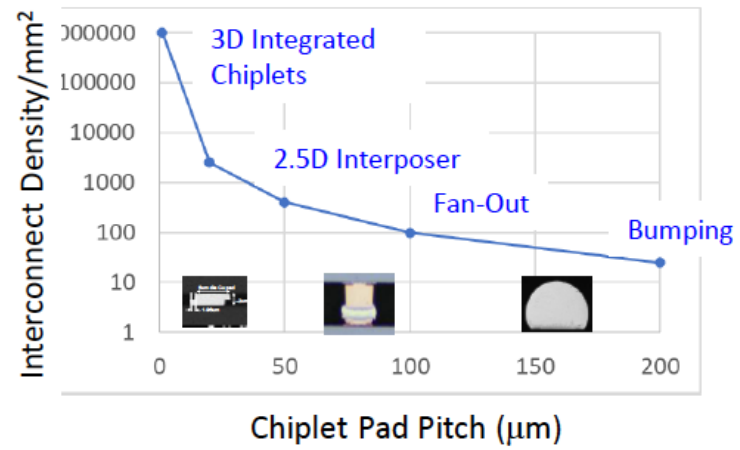
## The Vertical Approach





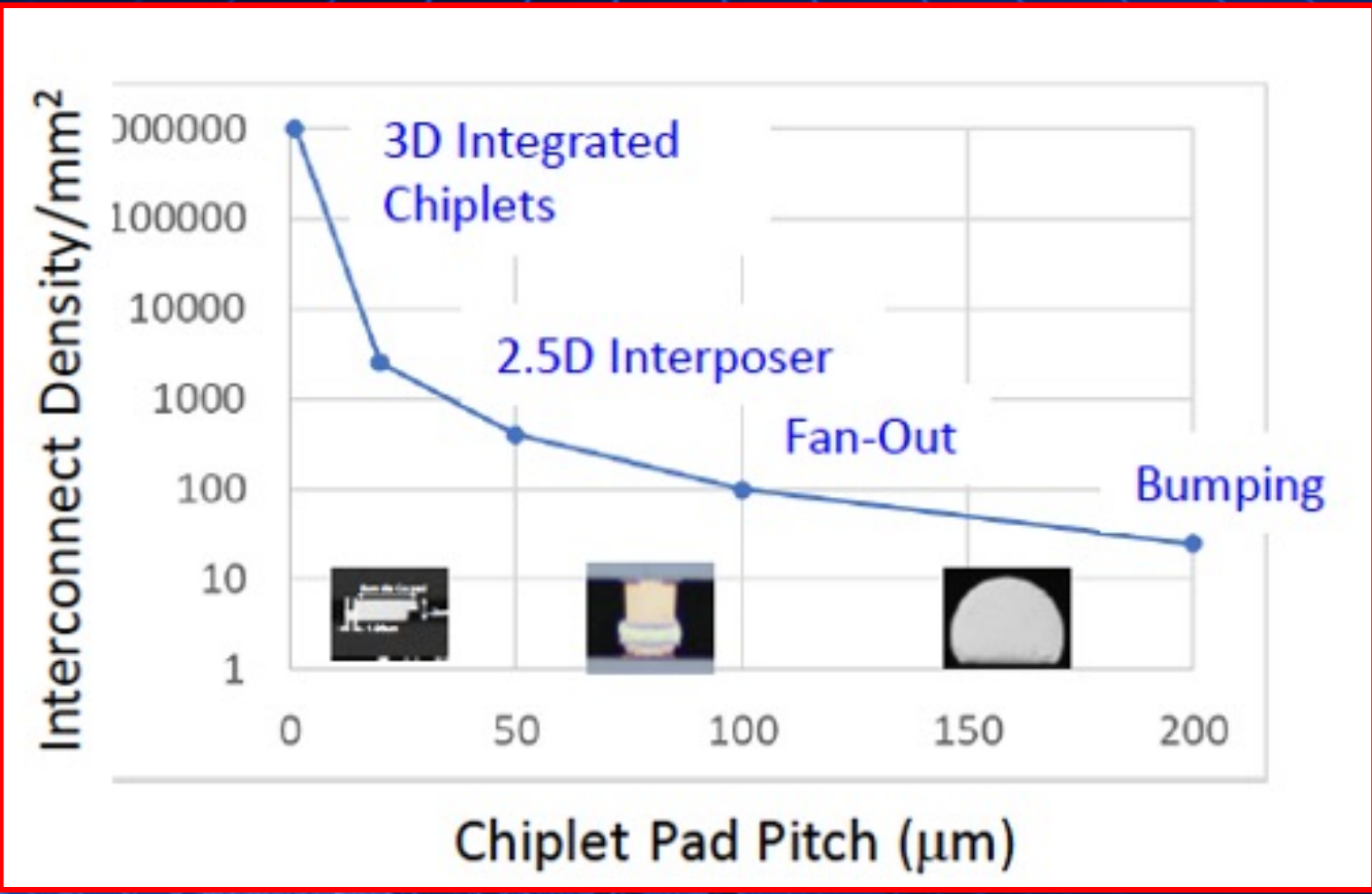
# Multi Chiplet Heterogeneous Integration (MCHI) Platforms

Target Applications.	Heterogeneous Integration Platforms
5G-Tx/Rx, mmWave Antenna, Automotive Radar AP+Memory, AR/VR.	<p><b>Embedded High Density Fan-Out WLP</b></p> 
Data Centre, HPC, AI, Chiplets	<p><b>2.5D Interposer</b></p> 
SOC Disintegration & reintegration, 3D Memory	<p><b>3D-Integrated Chiplets</b></p> 

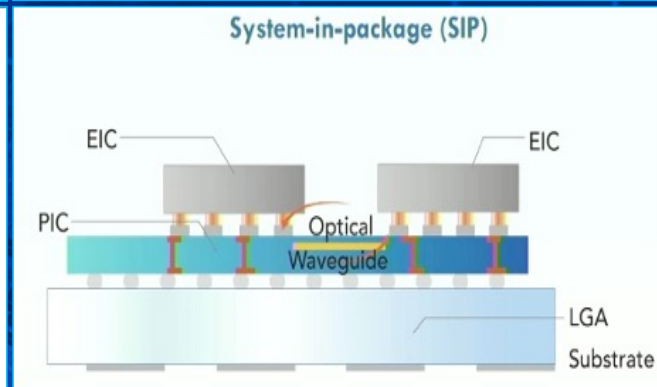
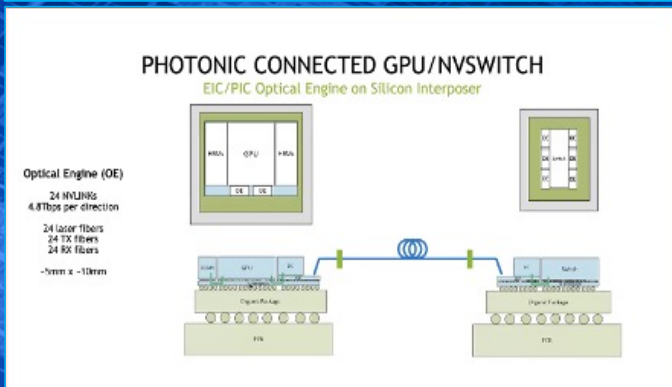
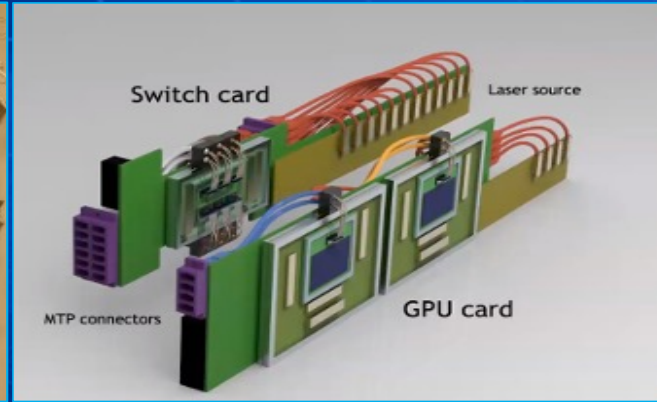
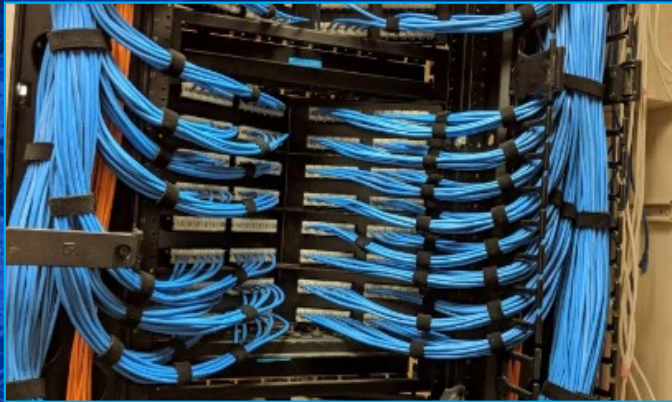


- Inter-chiplet wiring density increases
- 5x-10x from Bumping to FOWLP
  - 10x from FOWLP to 2.5D Interposer
  - >100x from 2.5D Interposer to 3DIC (hyb. Bonded)

**CHALLENGES:** Package Architecture, Design, Materials, Process Optimization



# Evolution of fiber optics from data centers (meters) to board (centimeters) to package-to-package (millimeters) to chip-to-chip (tens of microns)

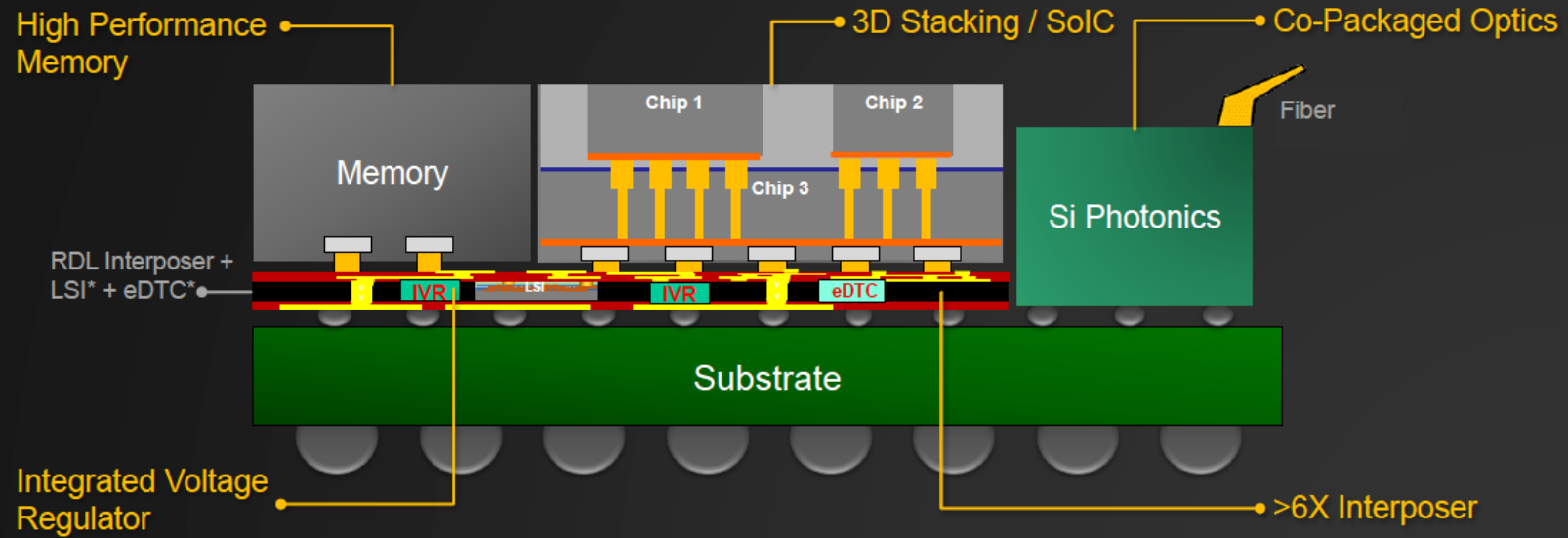


Sources: Fiber Broadband Association, NVIDIA, Lightelligence



# Technology Platform for HPC / AI

TOMORROW



\*: LSI: Local Silicon Interconnect; eDTC: embedded Deep Trench Capacitor

© 2024 IEEE International Solid-State Circuits Conference

22

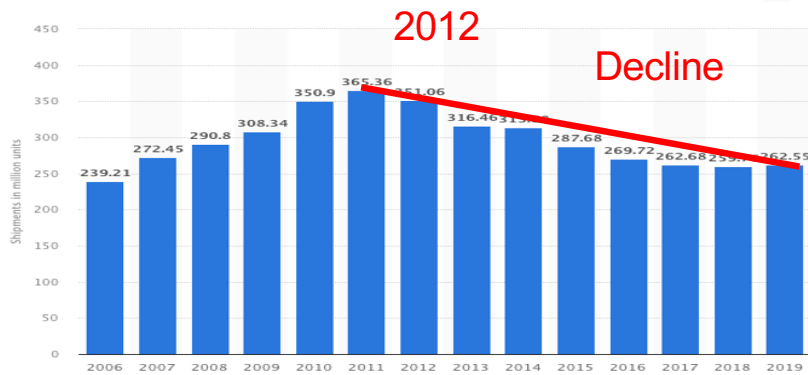
# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- **New Products**
- New Architecture
- SOC: Endless Die Size
- SOC and SIP

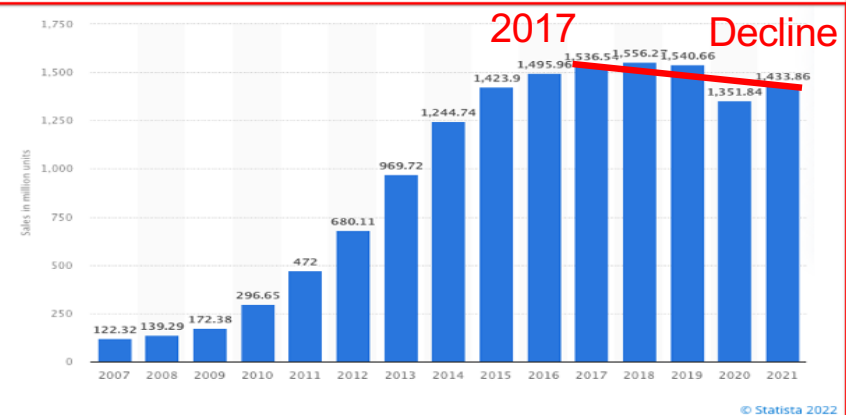
# What About products?

**Total unit shipments of personal computers (PCs) worldwide**

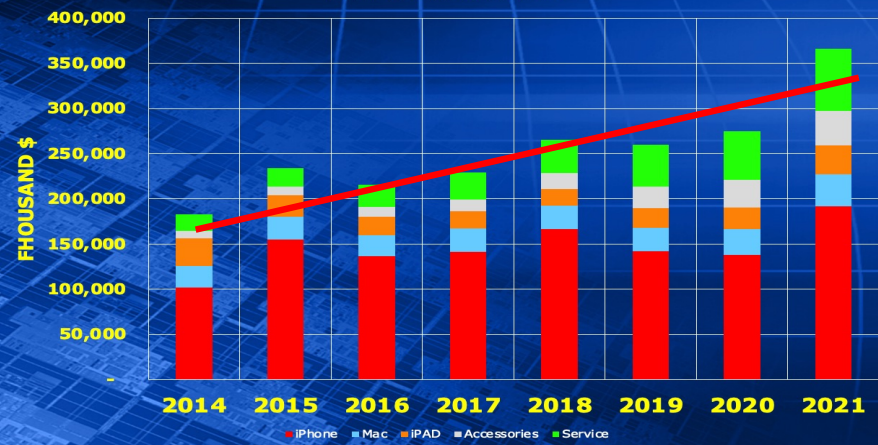
(in million units)



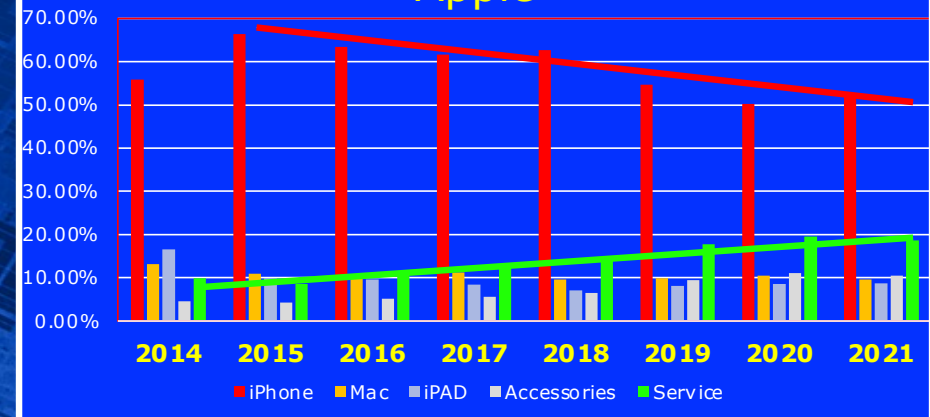
**Number of smartphones sold to end users worldwide from 2007 to 2021**



**APPLE REVENUE**



**Apple**



# The Next Wave of Products

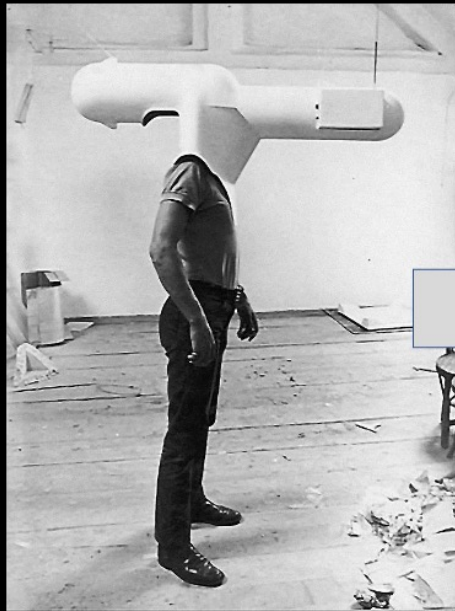
2024



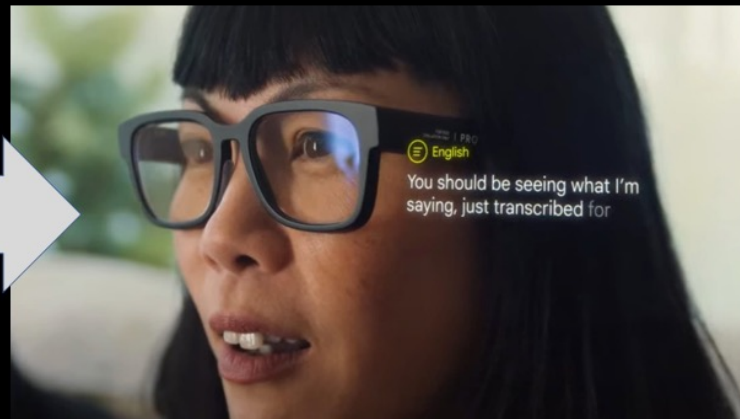
P.Gargini

53

Flat and thin optical systems are key to enabling compact, transparent and light weight AR display architectures



Walter Pichler's TV helmet (1967)



Google monocular AR display glasses (Google I/O, May 2022)

<https://www.istockphoto.com/search/2/image-film?page=13&phrase=ar%20glasses>

## Ray-Ban and Facebook introduce Ray-Ban Stories, first-generation smart glasses

After promising smart glasses for years, Facebook's first glasses **are** **disappointingly** familiar. **These aren't AR glasses at all** They don't have displays in them. Instead, they're a blend of technologies that have already been in other glasses: they have cameras in them, and microphones, and speakers. They're headphones and camera-glasses in one, and that's about it.



# Back to Basics

2024



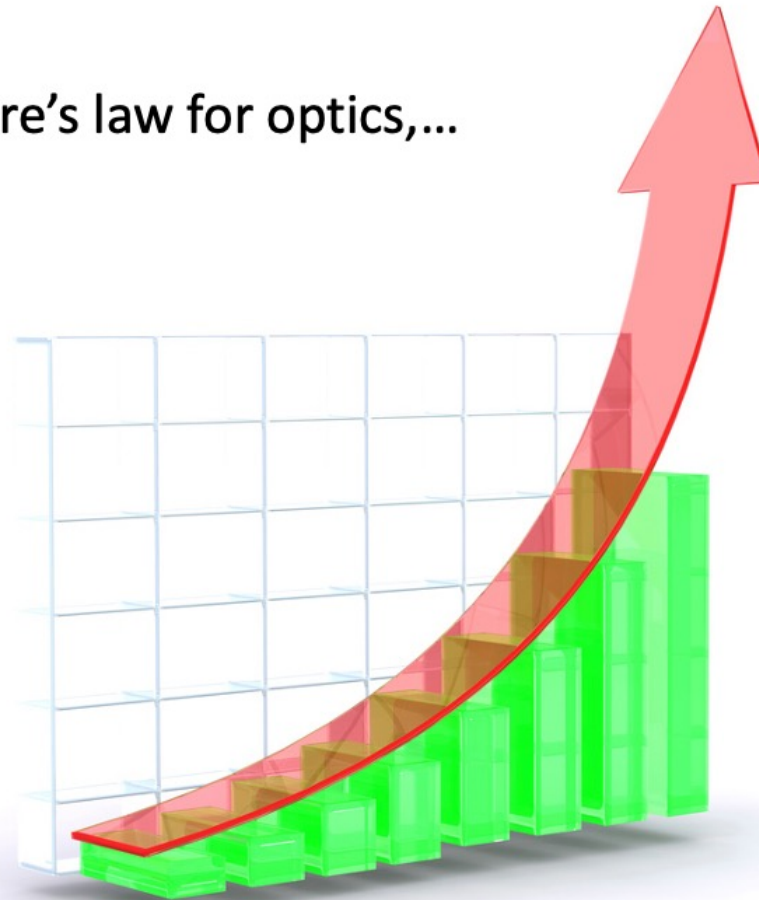
P.Gargini

56

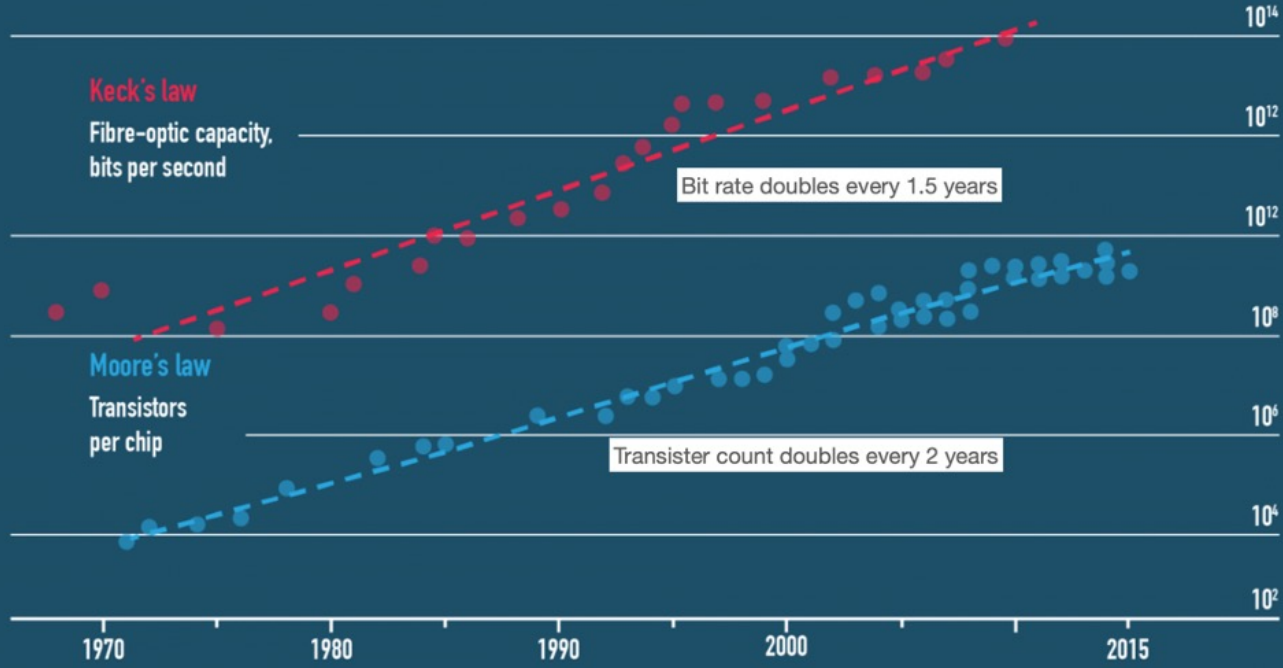


There might not be any Moore's law for optics,...

... but there are plenty of exponential growth laws in optics and photonics



# MOORE'S LAW VS KECK'S LAW, LOG SCALE



Source: Intel; Donald Keck - Via: The Economist

Gerd

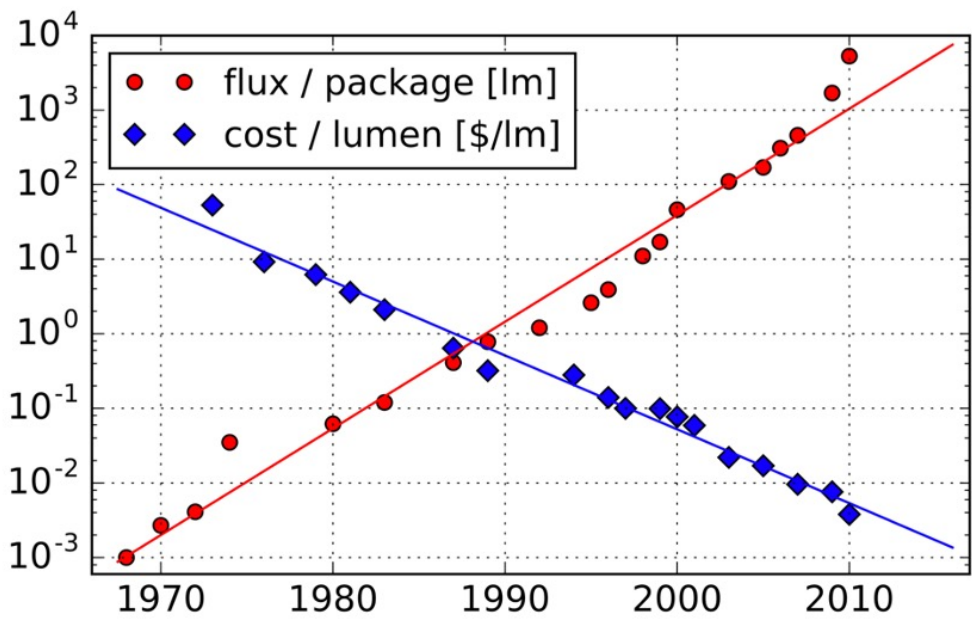
Donald Keck invented fiber optics in 1970 at Corning



Haitz's law is an observation and forecast about the steady improvement, over many years, of light (LEDs).

Roland Haitz  
Agilent Technologies

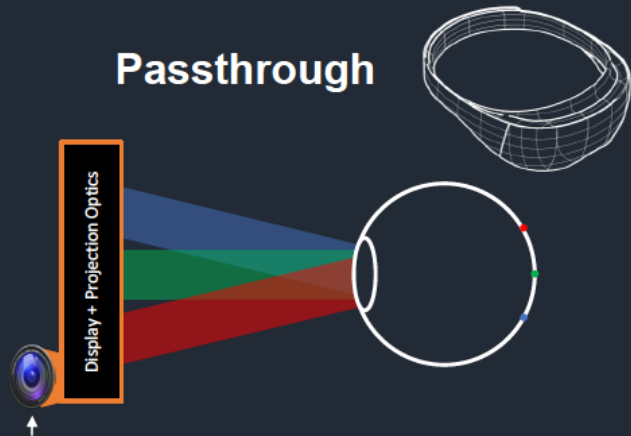
### Haitz's Law - Brightness and Cost Of Lighting





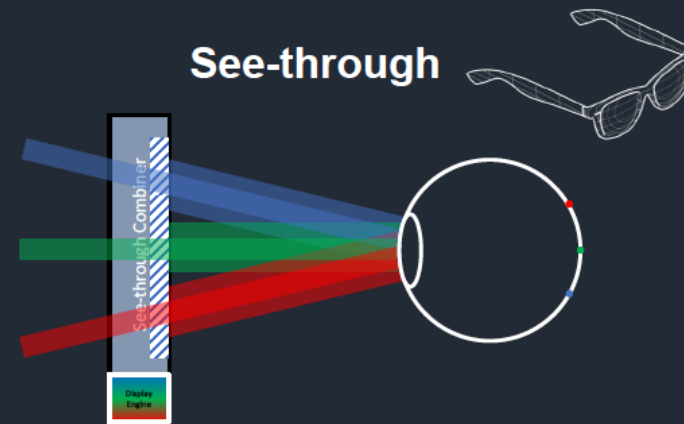
# AR: Passthrough vs. See-through

## Passthrough



Real world is captured through camera and re-displayed on screen together with other virtual content

## See-through



Real world is seen through a combiner that displays virtual content as well as being transparent



# The Future: A true see-through augmented reality experience that's all-day wearable



Meta

9



*Seamless UX*

*Contextual Awareness*

*High Resolution*

*Localized Audio*

*Bright*

**Immersive Experience**

**Super-power**

*Assistant*

*Large FOV*



*Stylish*

*Efficient*

**All-Day Wearable**

**Socially Acceptable**

*Transparent*

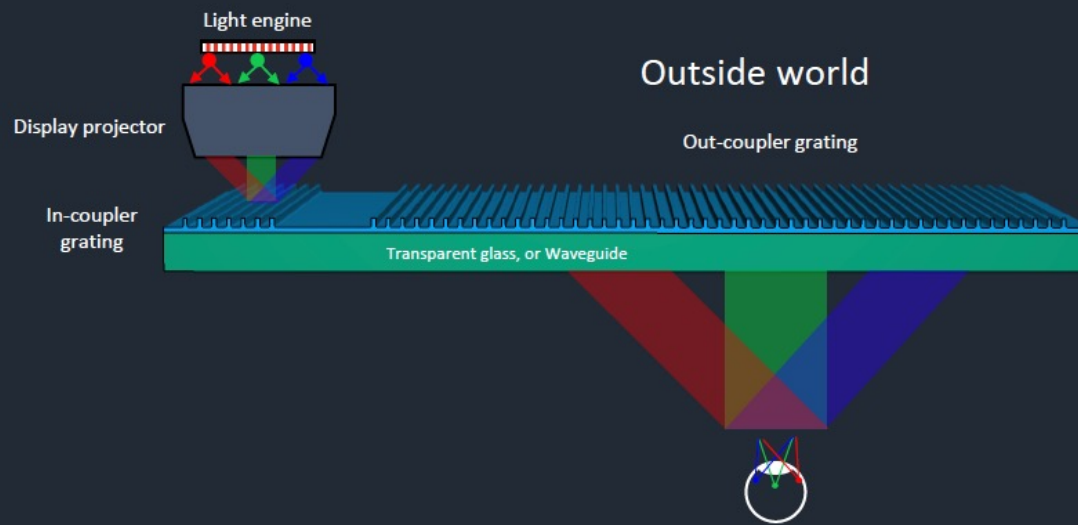
*Weight*

*Comfortable*

*Display visibility*

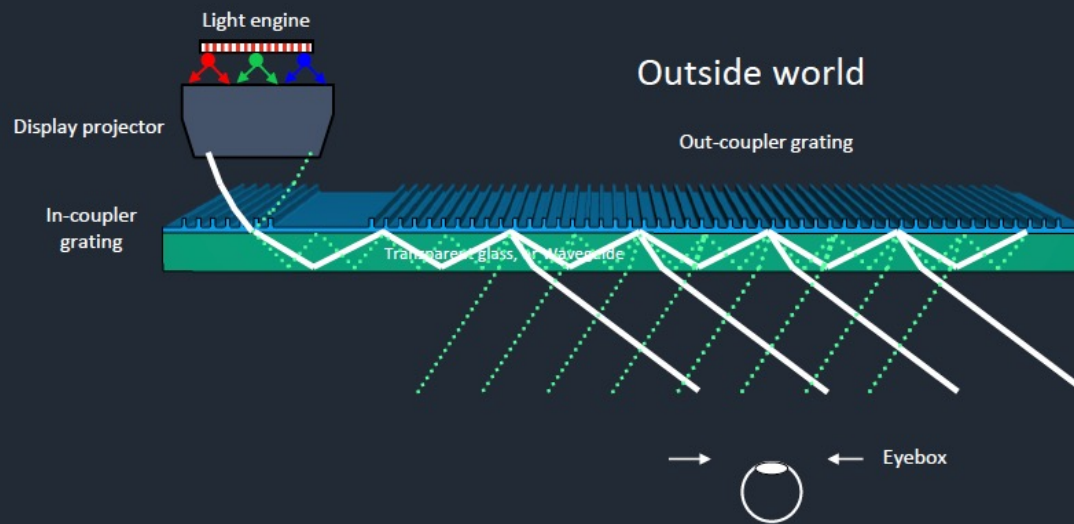
Meta

# How to enable see-through AR? – Waveguide in AR Glass





# How to enable see-through AR? – Waveguide in AR Glass





# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- New Products
- **New Architecture**
- SOC: Endless Die Size
- SOC and SIP

# The Next Computing Architecture

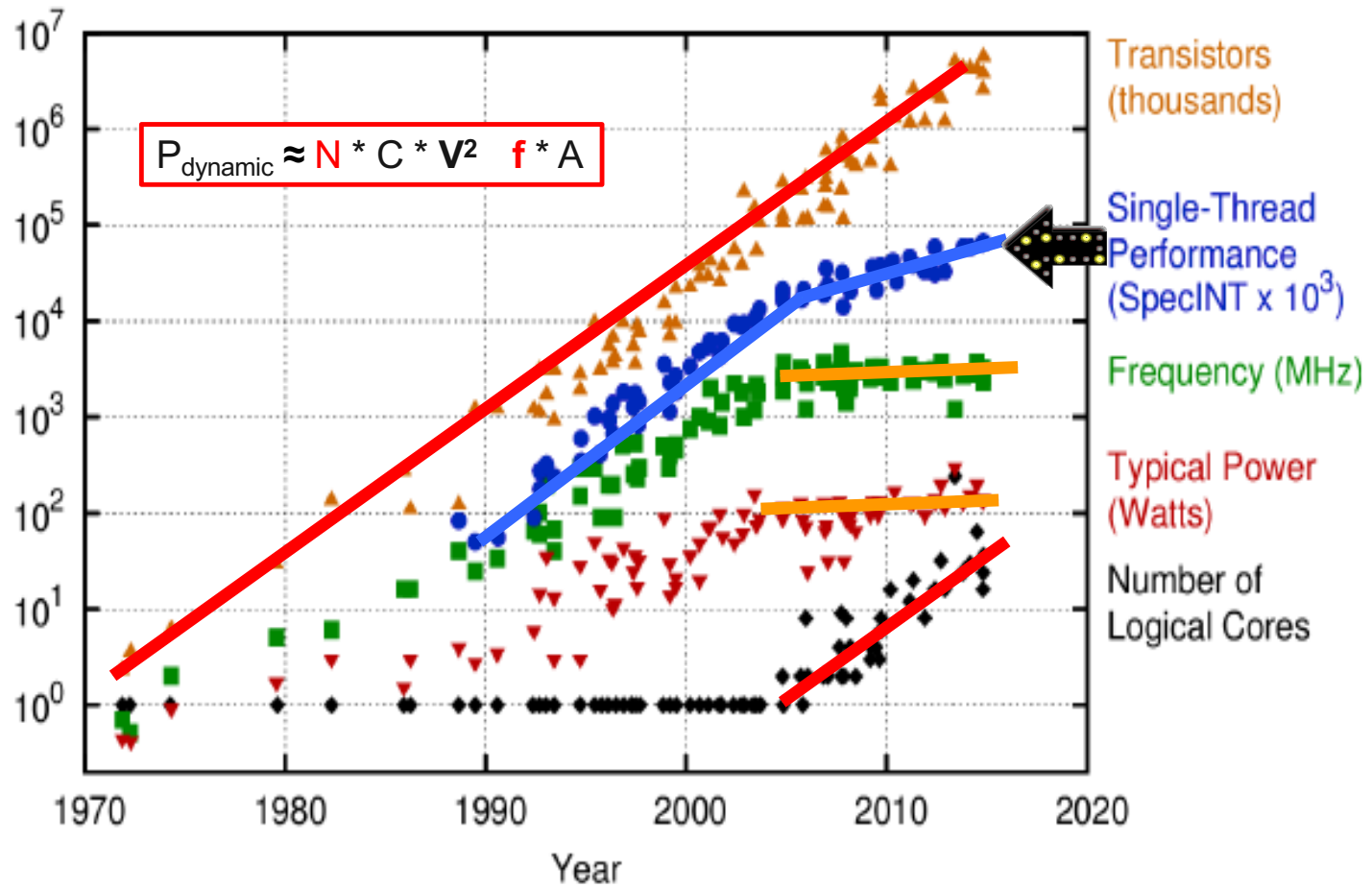
2024



P.Gargini

68

## 40 Years of Microprocessor Trend Data



2009

---

## Large-scale Deep Unsupervised Learning using Graphics Processors

---

Rajat Raina

Anand Madhavan

Andrew Y. Ng

Computer Science Department, Stanford University, Stanford CA 94305 USA

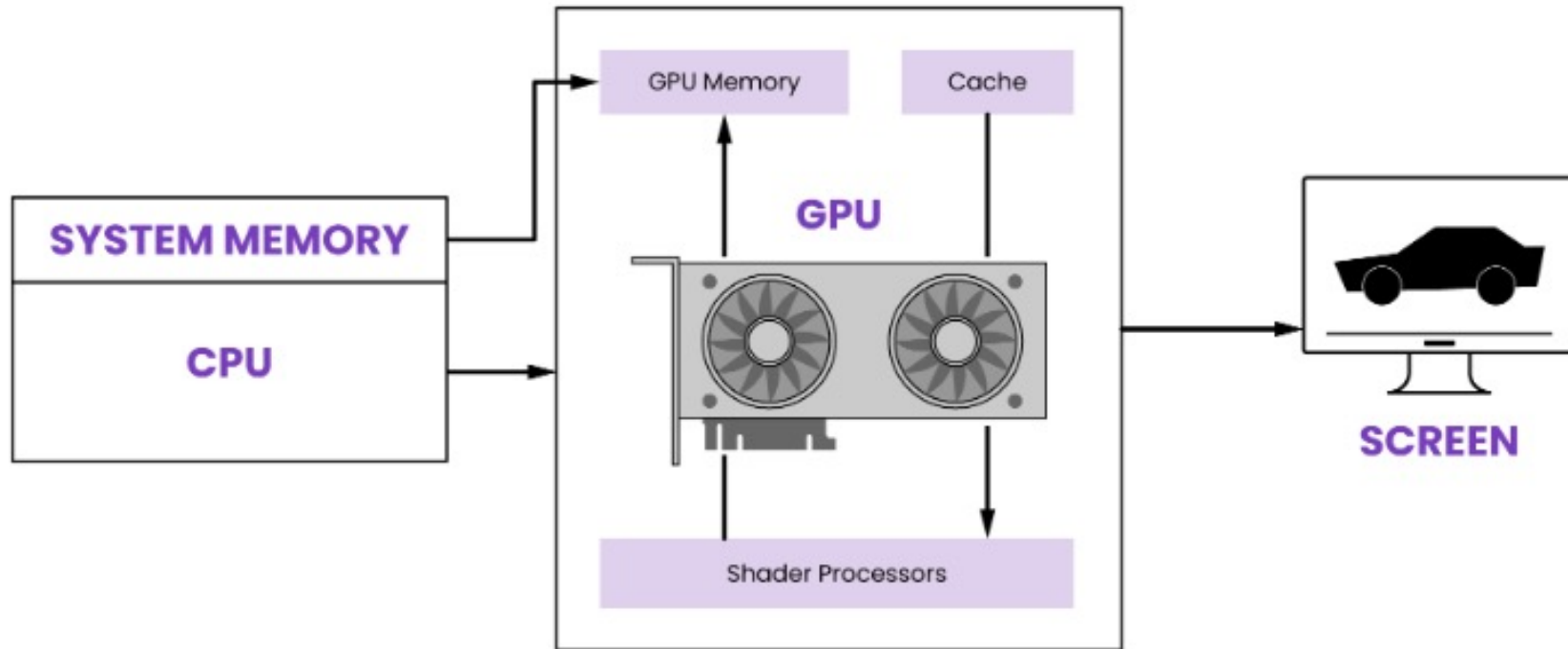
RAJATR@CS.STANFORD.EDU

MANAND@STANFORD.EDU

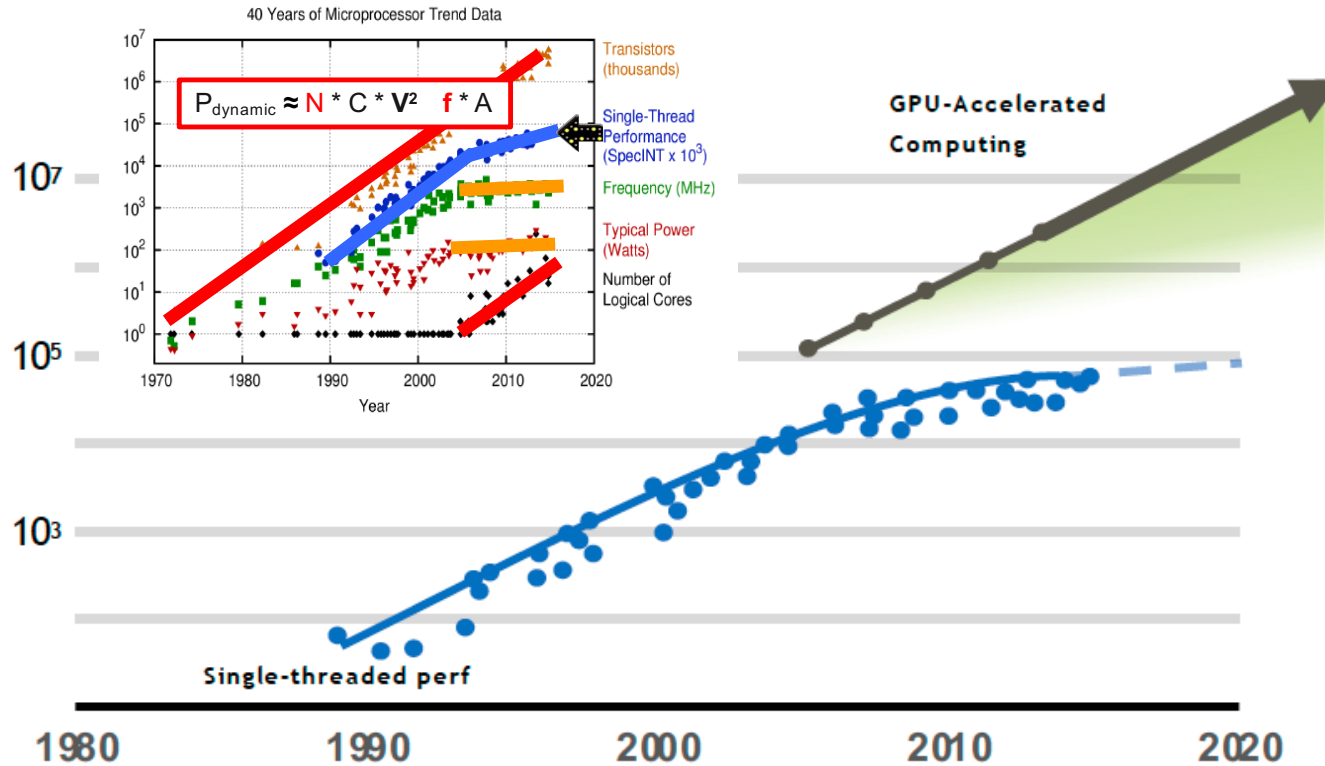
ANG@CS.STANFORD.EDU

### 7. Discussion

Graphics processors are able to exploit finer-grained parallelism than current multicore architectures or distributed clusters. They are designed to maintain thousands of active threads at any time, and to schedule the threads on hundreds of cores with very low scheduling overhead. The map-reduce framework (Dean & Ghe-



# A Solution to the Architectural Problem

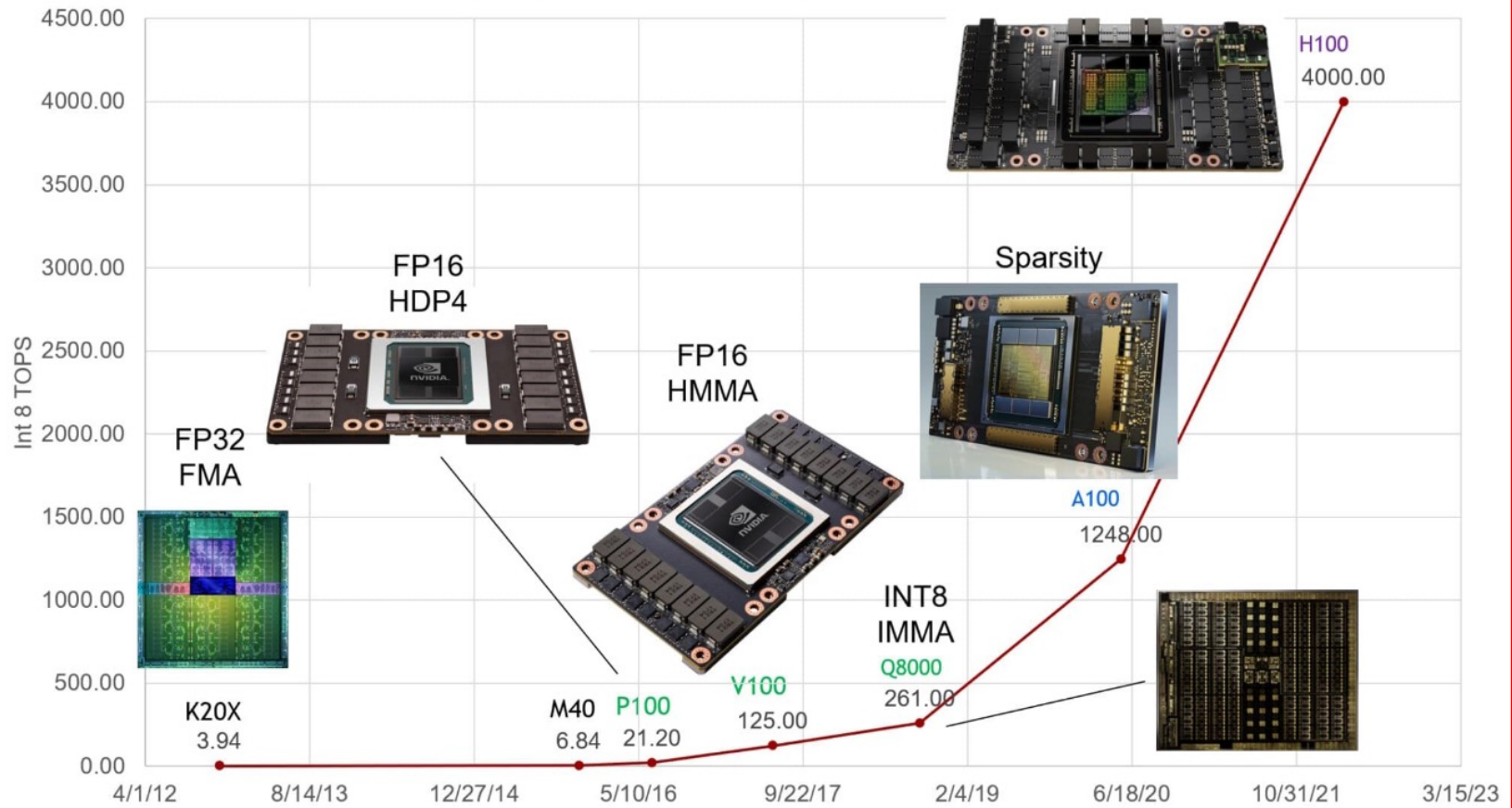


Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten  
 New plot and data collected for 2010-2015 by K. Rupp

Courtesy W. Dally, NVIDIA



### Single-Chip Inference Performance - 1000X in 10 years



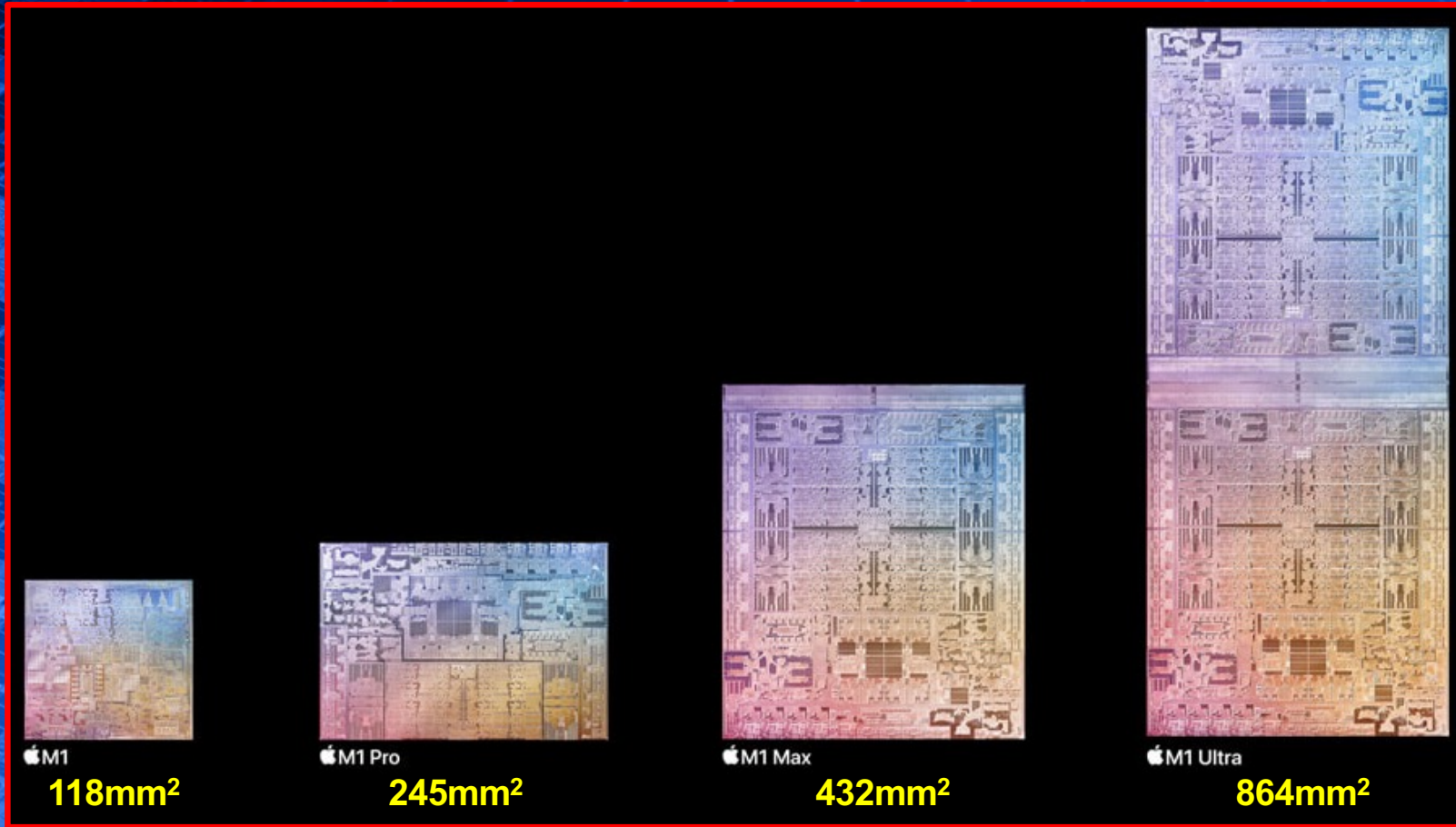
# The New Electronics Industry

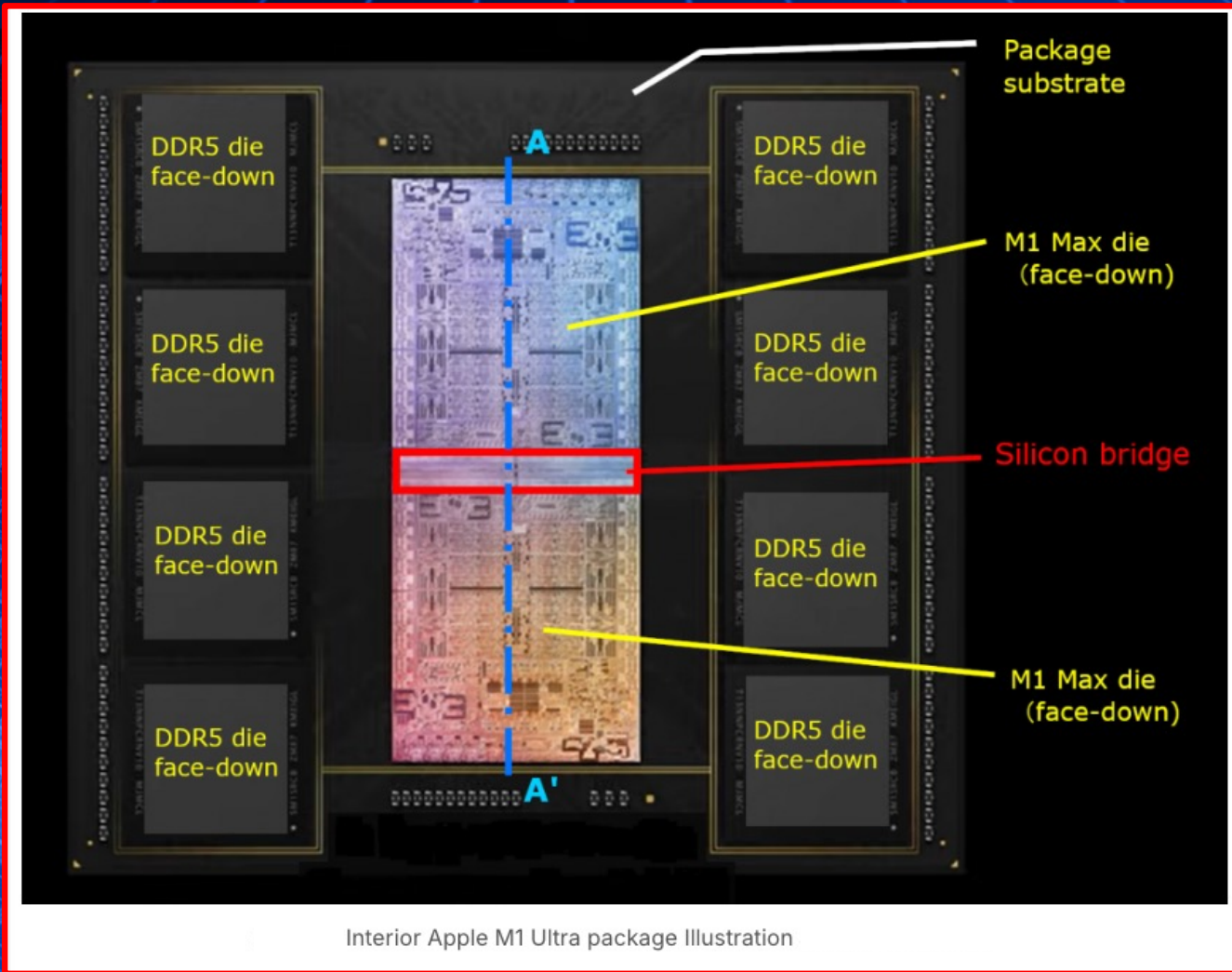
- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- New Products
- New Architecture
- **SOC: Endless Die Size**
- SOC and SIP

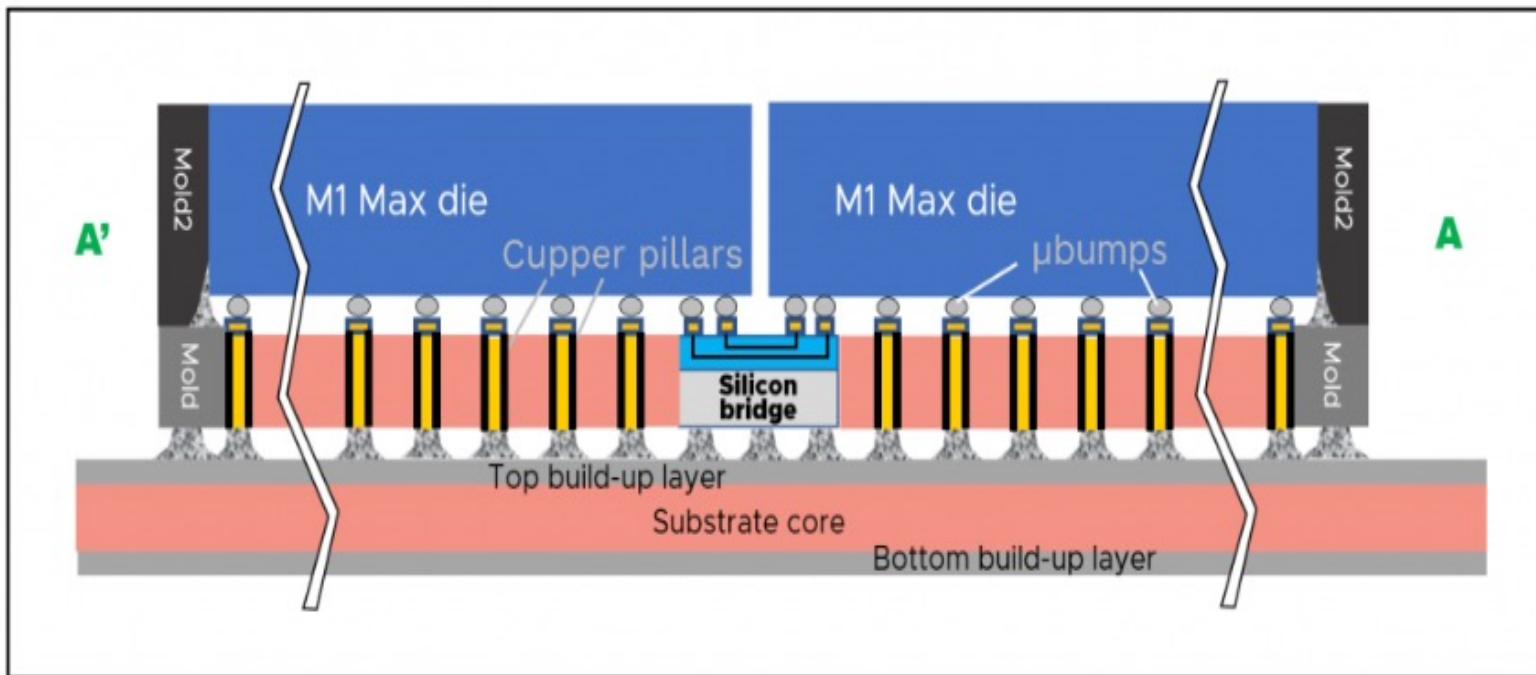
# SOC

## No Limits to Die Size

# APPLE Chips

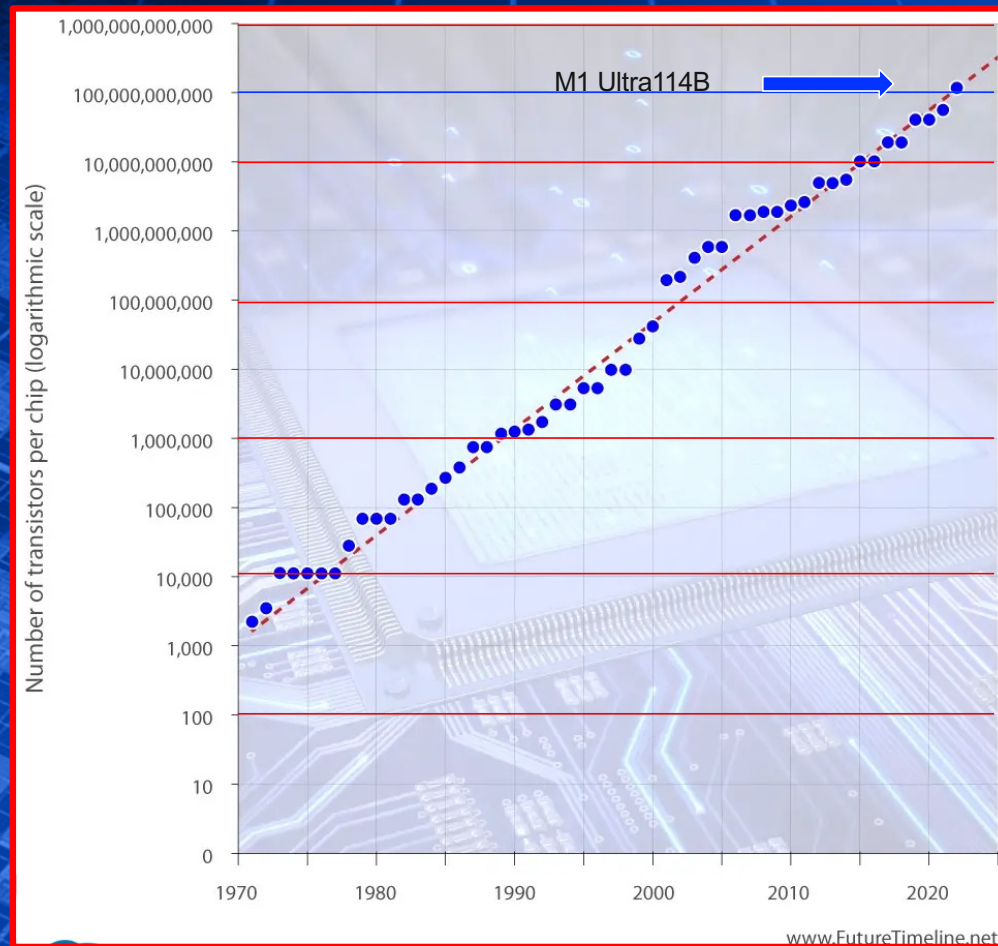




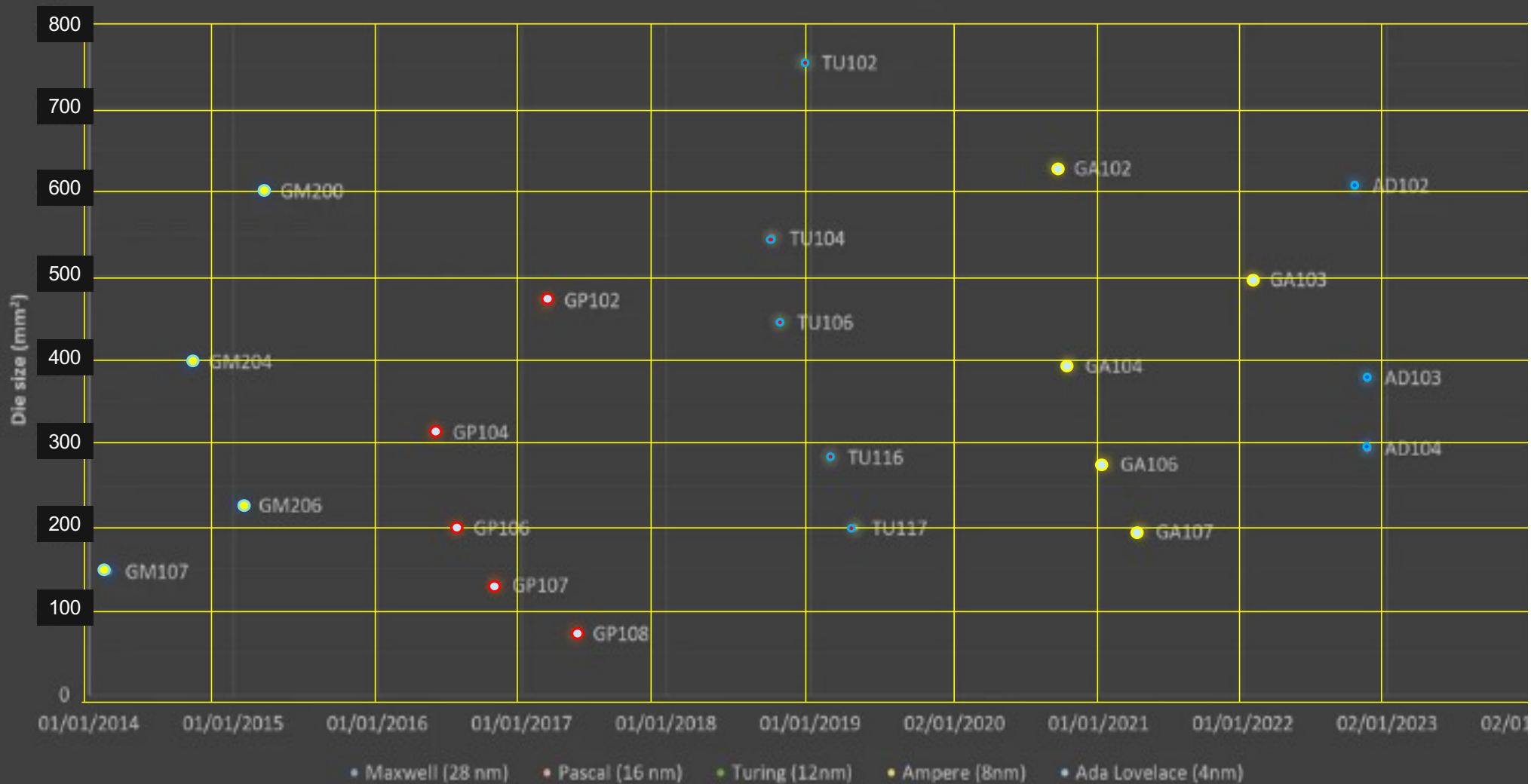


Vertical cross-section of the M1 Ultra package

# Over 100 Billion Transistors!



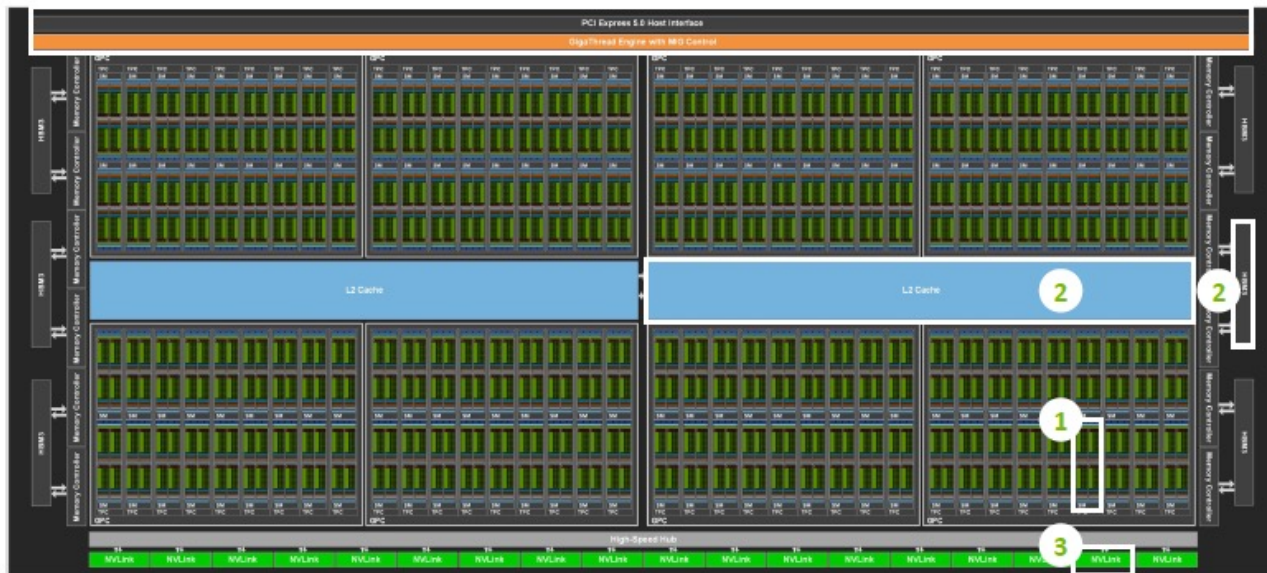
# Die sizes of Nvidia Gaming GPUs





# NVIDIA Hopper GPU — Full Implementation

80B Transistors, 814mm<sup>2</sup> in TSMC 4N



1. 144 SMs 2x Performance Per Clock 4<sup>th</sup> Gen Tensor Core Thread Block Clusters
2. New Memory System World's First HBM3e DRAM Larger 60MB L2
3. 4<sup>th</sup> Gen NVLink 900GB/s total BW New SHARP support NVLink Network

# Technologies are Reaching 100M Users Faster than Ever

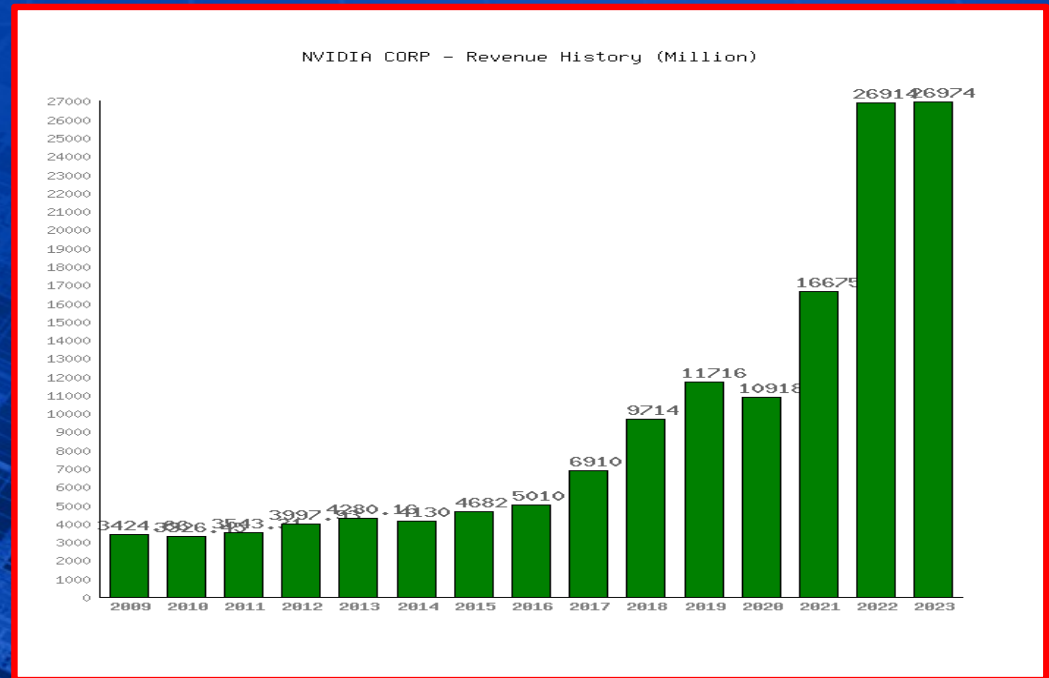
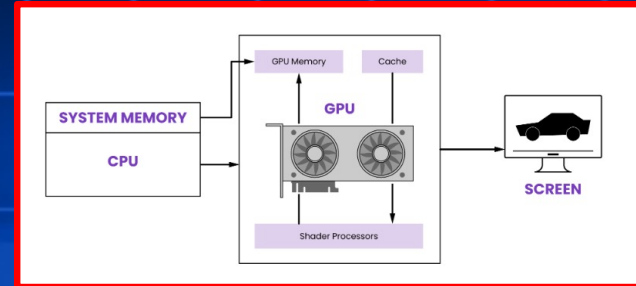


Source: <https://www.visualcapitalist.com/threads-100-million-users/>

© 2024 IEEE  
International Solid-State Circuits Conference

Nvidia DGX H100	
Component	AI Server
CPU	\$ 5,200
8 GPU + 4 NVSwitch Baseboard	\$ 195,000
Memory	\$ 7,860
Storage	\$ 3,456
SmartNIC	\$ 10,908
Chassis (Case, backplanes, cabling)	\$ 563
Motherboard	\$ 360
Cooling (Heatsinks+fans)	\$ 463
Power Supply	\$ 1,200
Assembly and Test	\$ 1,485
Markup	\$ 42,000
<b>Total Cost</b>	<b>\$ 268,495</b>
DRAM BOM %*	2.9%
NAND BOM %	1.3%
Memory BOM %*	4.2%

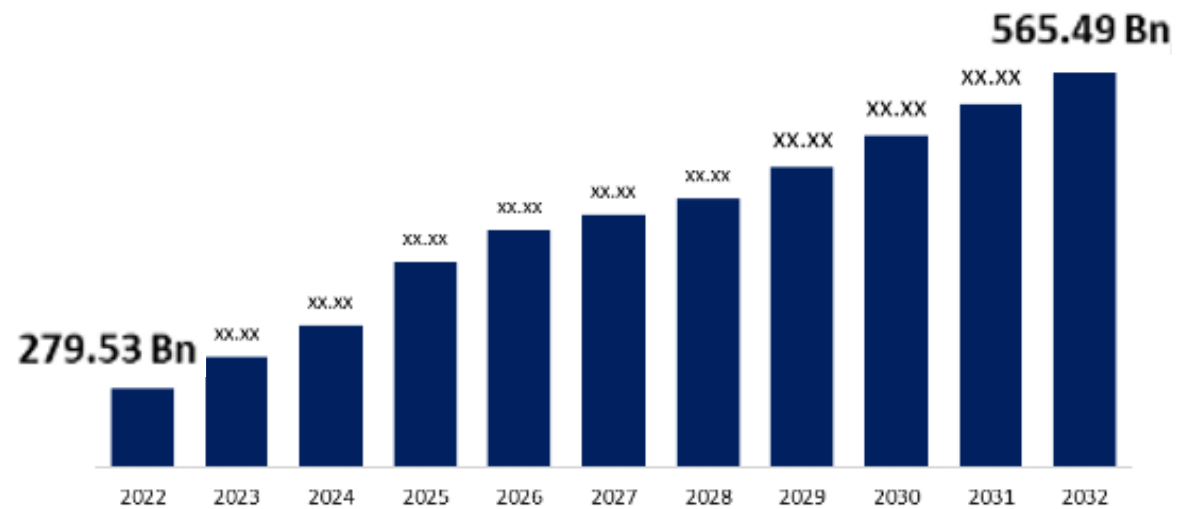
\*HBM Costs are not included as they are part of the GPU. We have broken the Nvidia BOM below for subscribers.



## Global Data Center Market Insights Forecasts to 2032

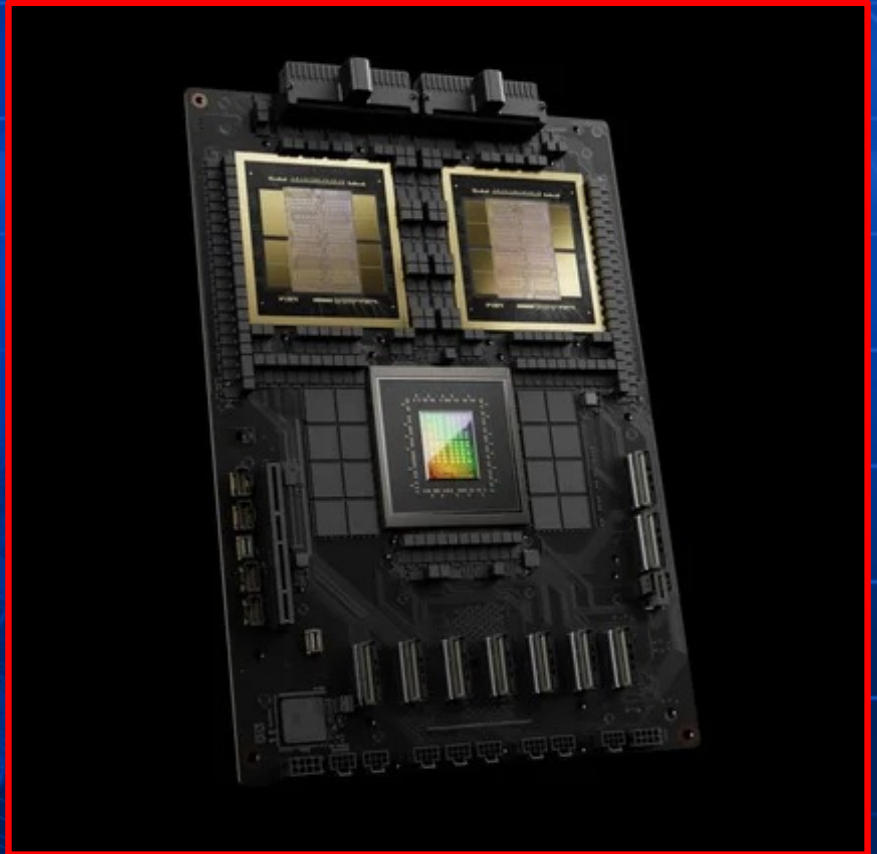
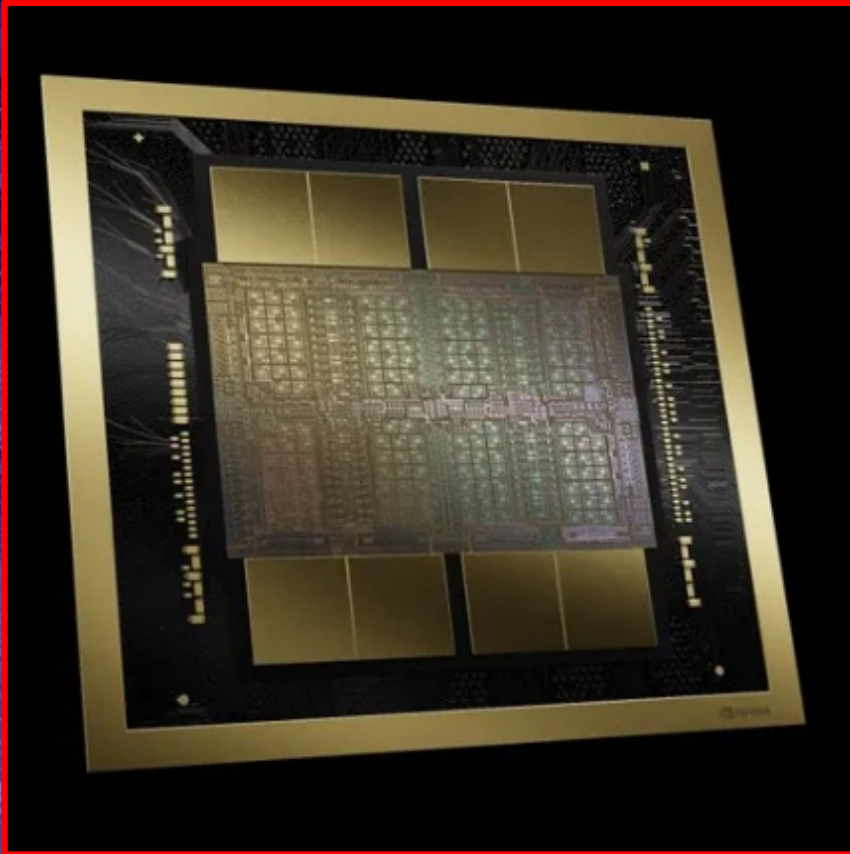
- The Global Data Center Market Size was valued at USD 279.53 Billion in 2022.
- The Market is Growing at a CAGR of 7.3% from 2023 to 2032
- The Worldwide Data Center Market Size is expected to reach USD 565.49 Billion by 2032
- Asia-Pacific is expected to Grow the fastest during the forecast period

### Global Data Center Market



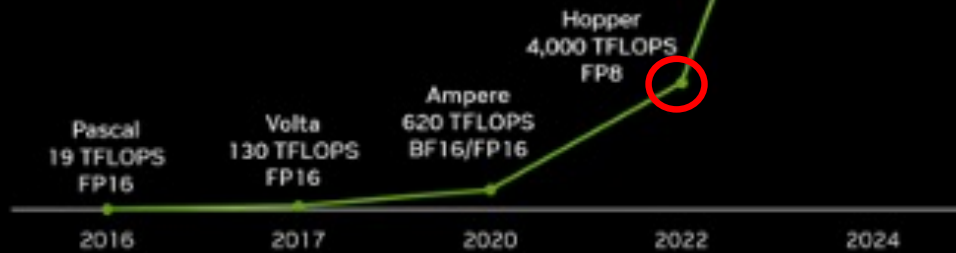
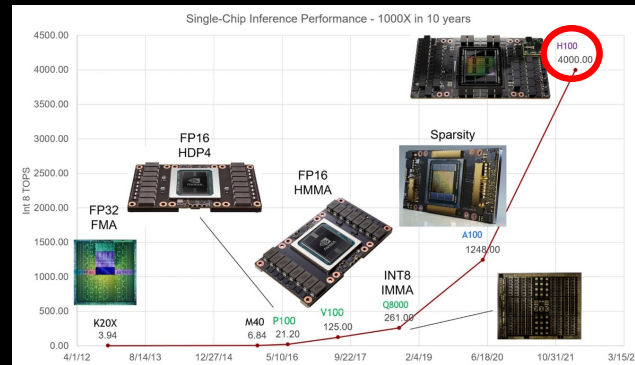
# BLACKWELL

March 18, 2024



206 Billion Transistors, two 800mm<sup>2</sup> fused dice

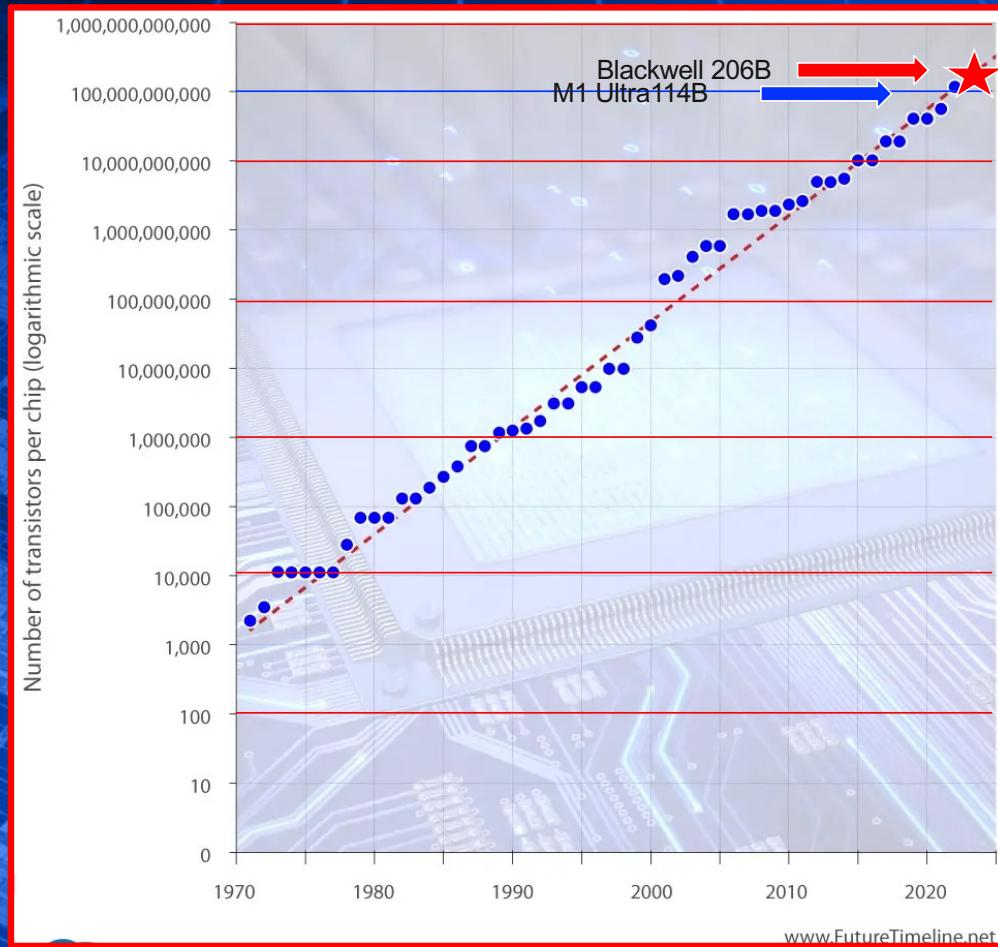
# 1000X AI Compute in 8 Years



**Blackwell  
20,000 TFLOPS  
FP4**



# Over 200 Billion Transistors!





# 4 TRILLION TRANSISTORS



Andrew Feldman

**900,000**

AI-Optimized Cores  
123x more cores

**44GB**

On-Chip SRAM  
1,000x more on-chip memory

**214Pb/s**

Interconnect Bandwidth  
45,000x more bandwidth

**21PB/s**

Memory Bandwidth  
12,800x more bandwidth

**1.2Tb/s**

System I/O

**15RU**

System Dimensions

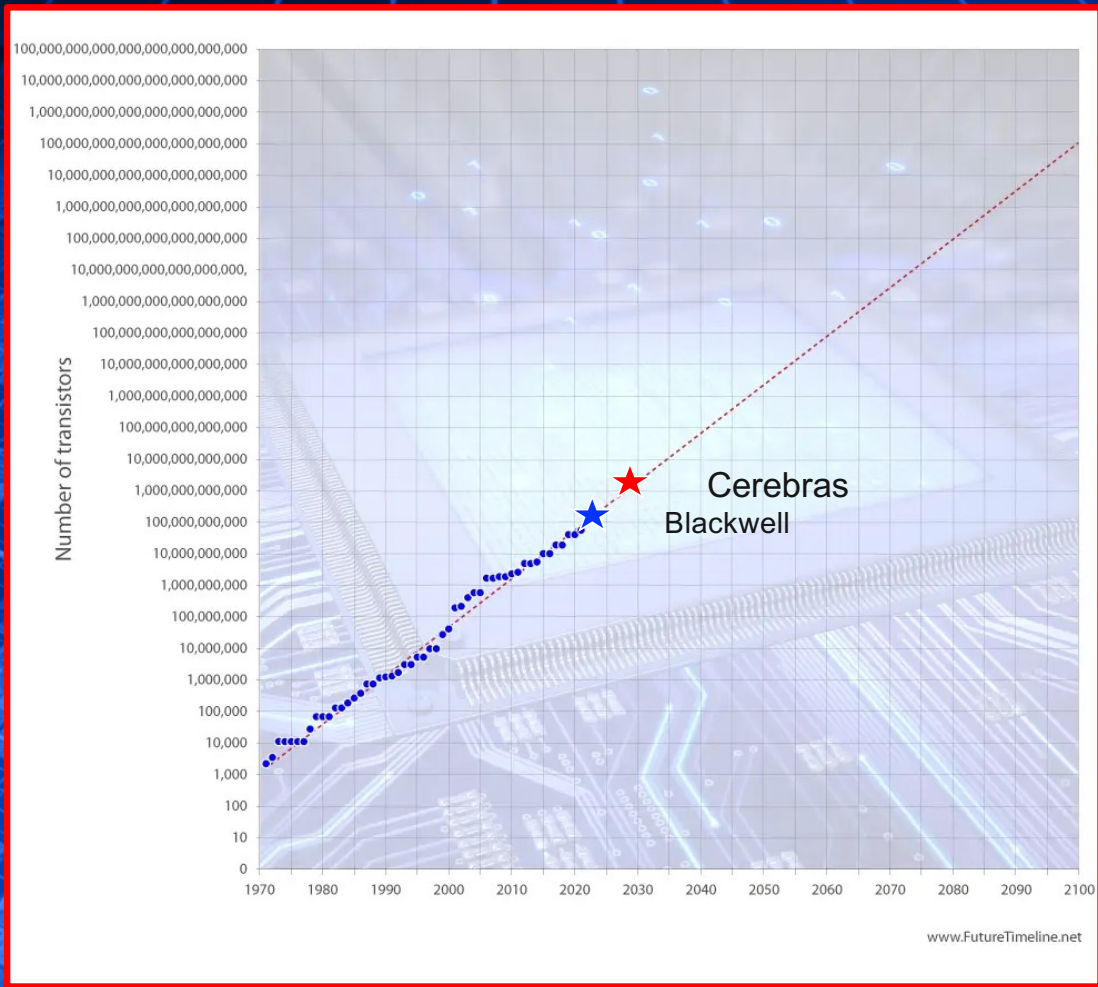




There is **no fundamental obstacle** to achieving device **yields of 100%**. At present, packaging costs so far exceed the cost of the semiconductor structure itself that there is no incentive to improve yields, but they can be raised as high as is economically justified. **No barrier exists** comparable to the thermodynamic equilibrium considerations that often limit yields in chemical reactions; it is **not even necessary** to do any fundamental **research** or to replace present processes.

**Only the engineering effort is needed**

Gordon Moore, 1965



# The New Electronics Industry

- 3D Transistors
- Beyond Silicon
- 3D Everything
- System In Package
- New Products
- New Architecture
- SOC: Endless Die Size
- SOC and SIP

# SOC and SIP

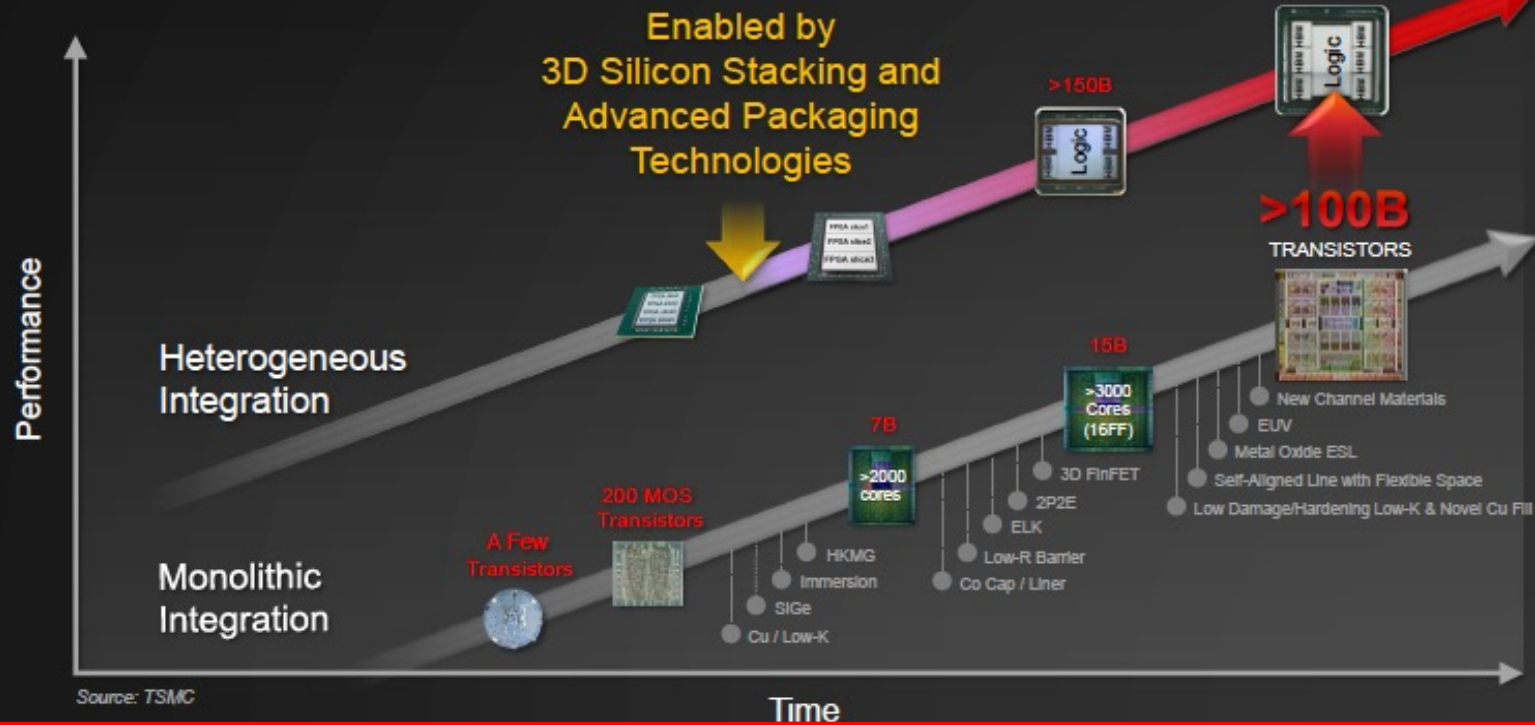
2024



P.Gargini

92

# Innovation Beyond Chip Level is a Must



Source: Kevin Zhang, Sr. VP of TSMC, February 19, 2024, ISSCC Plenary

**1975**

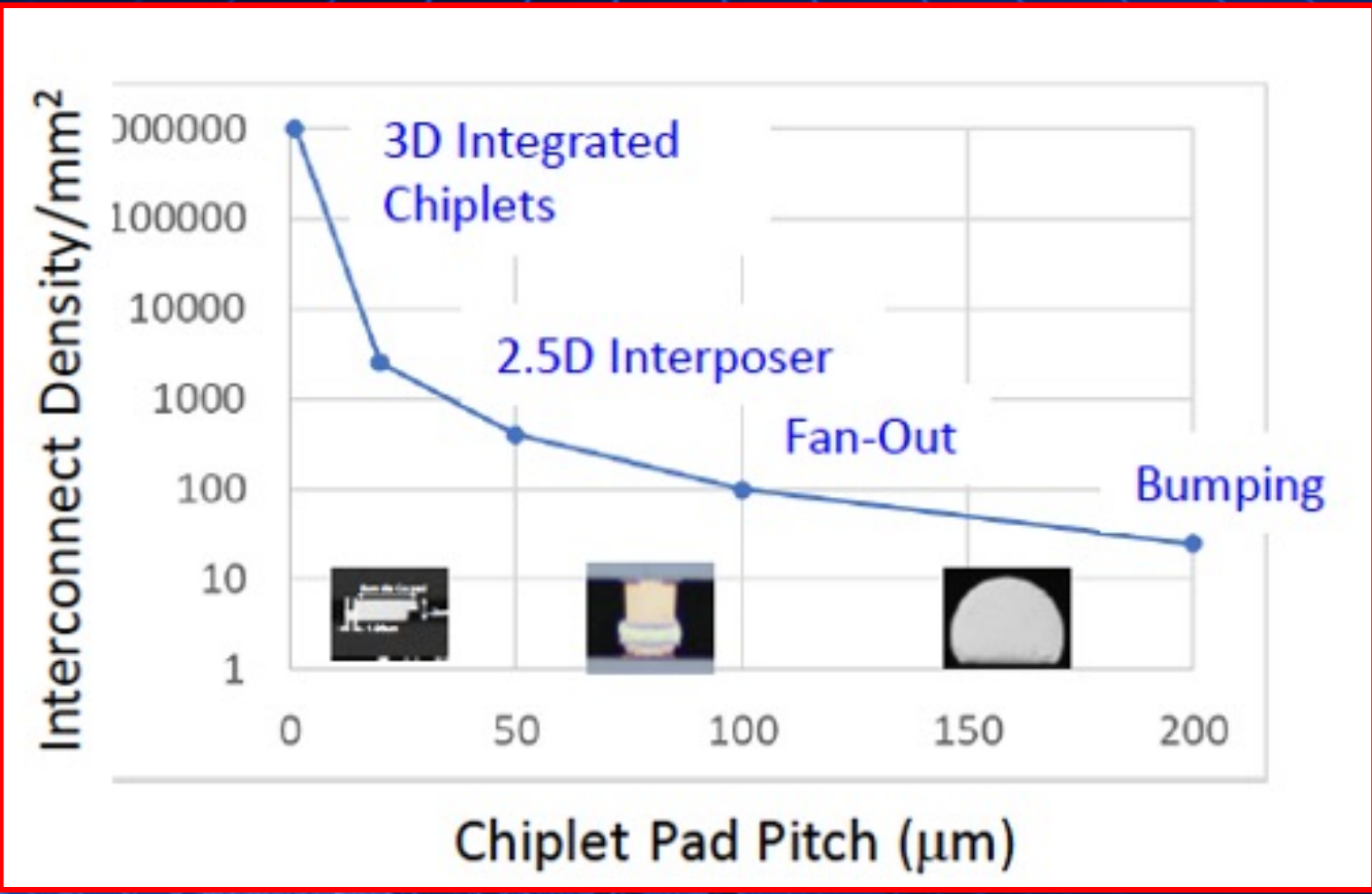
**THE NUMBER OF TRANSISTORS  
PRODUCED IN AN INTEGRATED  
CIRCUIT CAN COST EFFECTIVELY  
DOUBLE EVERY 2 YEARS**

**Gordon Moore**

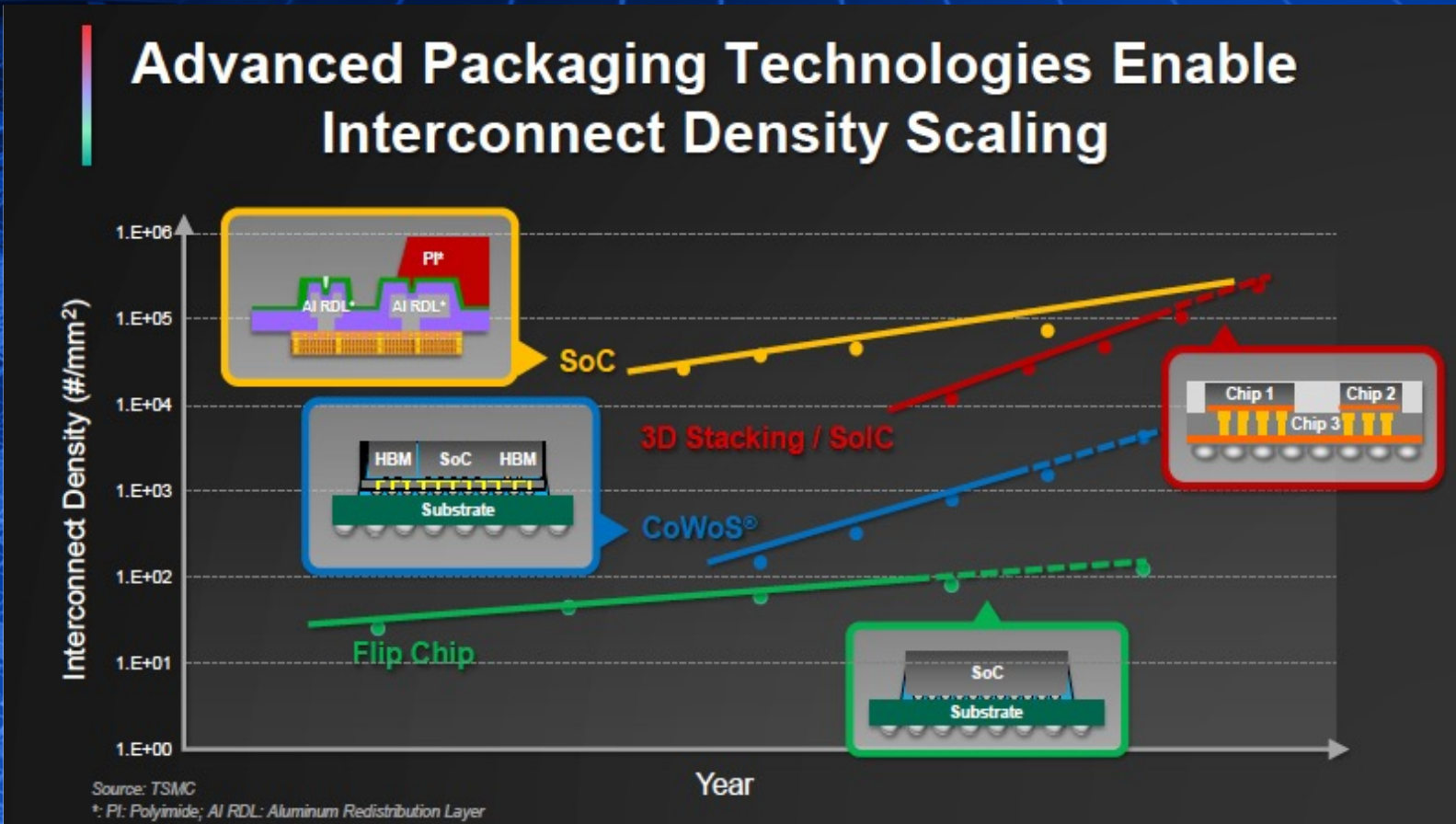
**2020**

**THE NUMBER OF BITS  
PRODUCED IN AN INTEGRATED  
PACKAGE CAN COST EFFECTIVELY  
DOUBLE EVERY 2 YEARS**

**Paolo Gargini**



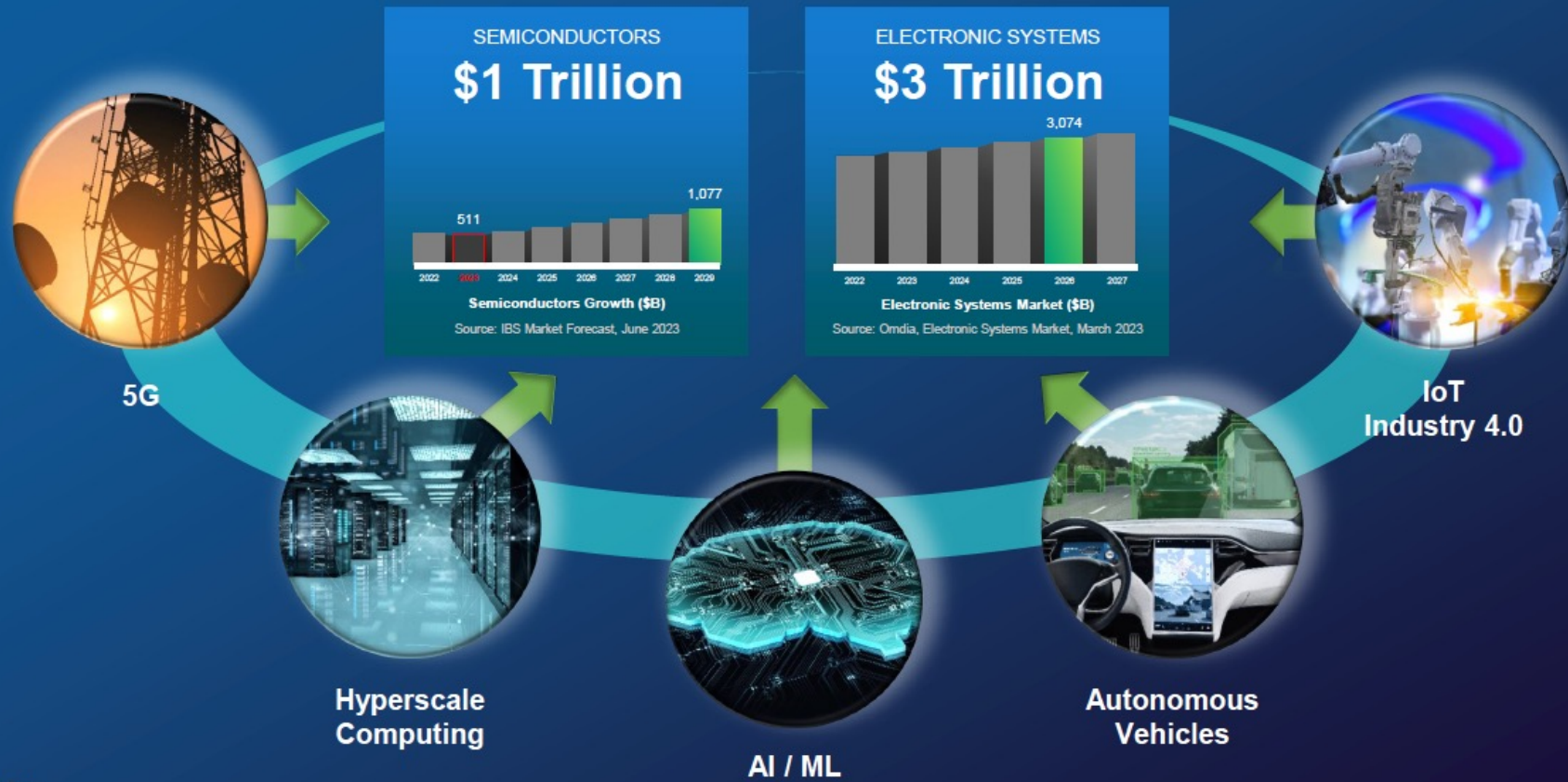
# Advanced Packaging Technologies Enable Interconnect Density Scaling



Source: Kevin Zhang, Sr. VP of TSMC, February 19, 2024, ISSCC Plenary



# Generational Semiconductor Drivers



© 2024 IEEE  
International Solid-State Circuits Conference

# Conclusions

# Transistor, Die, Package, Architecture, Products

- During the past 60 years the Electronics Industry has undergone and overcome 3 transformational inflection points
- In each case the Electronics Industry has created new semiconductors' devices and introduced new products that have transformed society
- The Electronics Industry is undergoing a top to bottom and bottom to top unprecedented revolution.
- SOC, SIP, GAA beyond silicon are all coming together to support unlimited possibilities
- AR/VR will give everybody a new "vision" of reality
- AI is going to make everybody omniscient
- Everything we have known will change in the new society