Challenges in Advanced Computing and Functionalities International Cooperation on Semiconductors

Technology Highlights in Advanced Computing

Nadine Collaert (imec) and Olivier Faynot (CEA-LETI)

Presented by Ray Duffy (Tyndall National Institute, UCC)



On Semiconductors



Outline

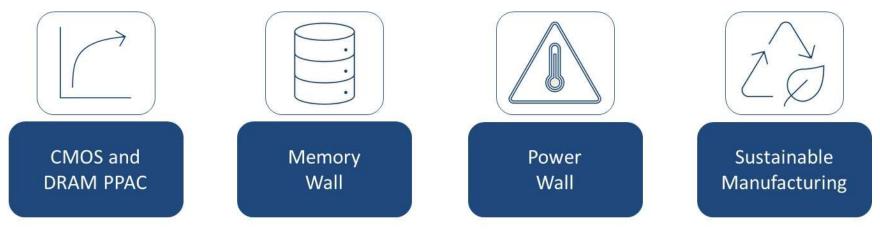
- Introduction: trends and challenges
- Computing roadmap:
 - CMOS device architecture
 - 2D materials for FEOL
 - New materials for BEOL
 - Lithography
 - CMOS 2.0
- Memory technologies
- Beyond Von Neumann disruptive approaches:
 - Near or in-memory computing
 - Quantum computing
- Heterogeneous integration: from chiplets to functional backside
- Analysis: EU and non-EU actors
- Conclusions





Summary

- Compute needs are growing at an unprecedented speed
- Innovations across different levels needed to enable chip & system performance enhancement
- Sustainability an increasingly important metric for evaluating technology choices
- Europe is very strong in R&D in all advanced compute areas



PPAC=Power-Performance-Area-Cost





Diversity of Applications and Workloads

GPUs for Training



AR/VR



Autonomous driving

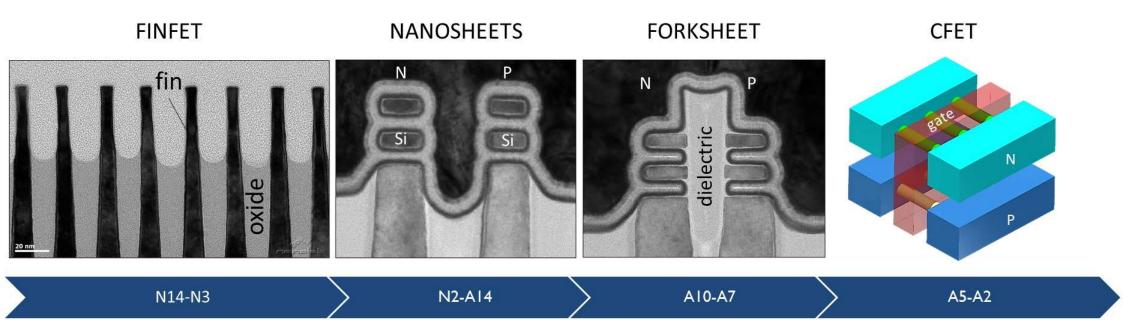


High throughput parallel compute Very high memory bandwidth Very high GPU-GPU bandwidth Low power Ultra low latency High memory bandwidth Small form factor Multi-sensor fusion Distributed real-time computation Reliable and explainable AI





New device architectures to extend scaling



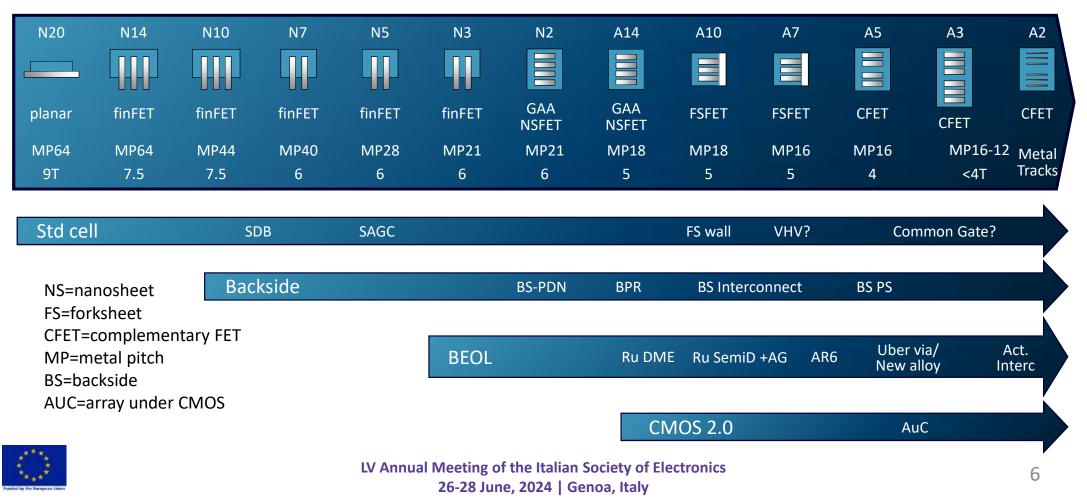
N. Collaert, "Advancements in IC Technologies: A look toward the future," in IEEE Solid-State Circuits Magazine, vol. 15, no. 3, pp. 80-86, 2023, doi: 10.1109/MSSC.2023.3280433.



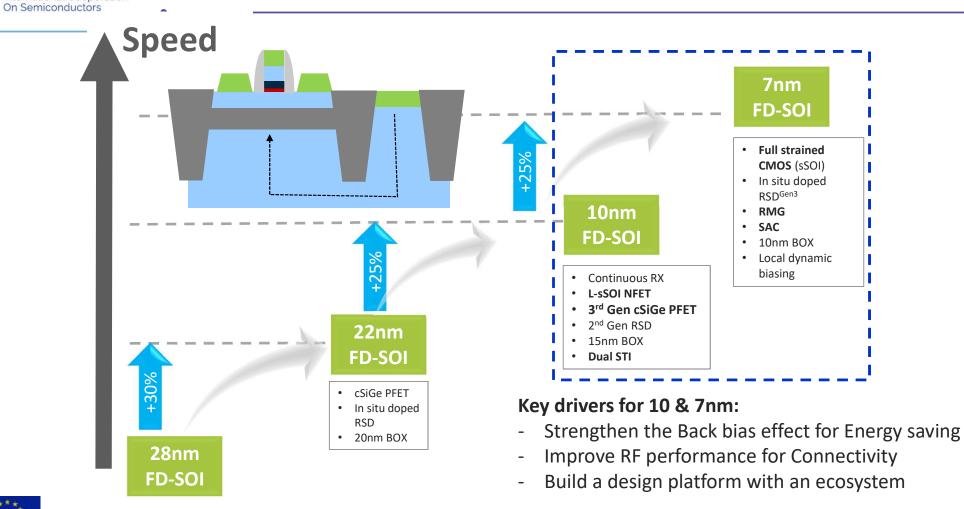


Logic scaling roadmap

A. Veloso, ICOS workshop, April 2023.





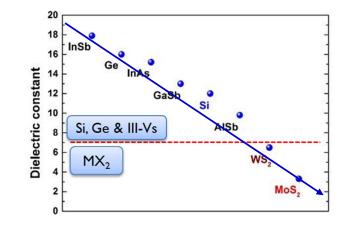


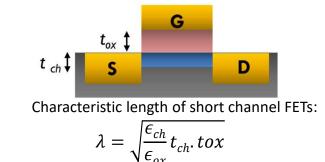


International Cooperation

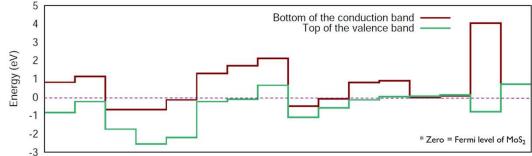


2D Channels: Next generation logic devices





Expect reduced short channel effects in planar devices Ultra-thin materials

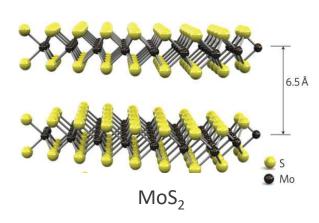


Choice of bandgaps and band alignment No/few dangling bonds at interfaces

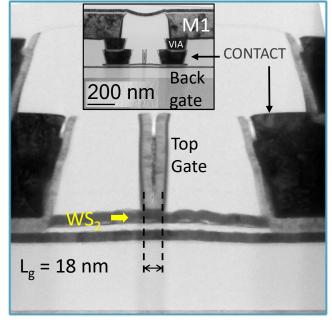
MoS₂ WS₂ ZrS₂ HfS₂ SnS₂MoSe₂WSe₂ZrSe₂HfSe₂SnSe₂MoTe₂WTe₂ZrTe₂HfTe₂ hBN graphene



2D Channels: Next generation logic devices



Monolayer channel thickness enables gate length scaling while keeping high performance



300mm Flow

I. Asselberghs et al, IEDM 2020

Layer Transfer1)2)153)55</td

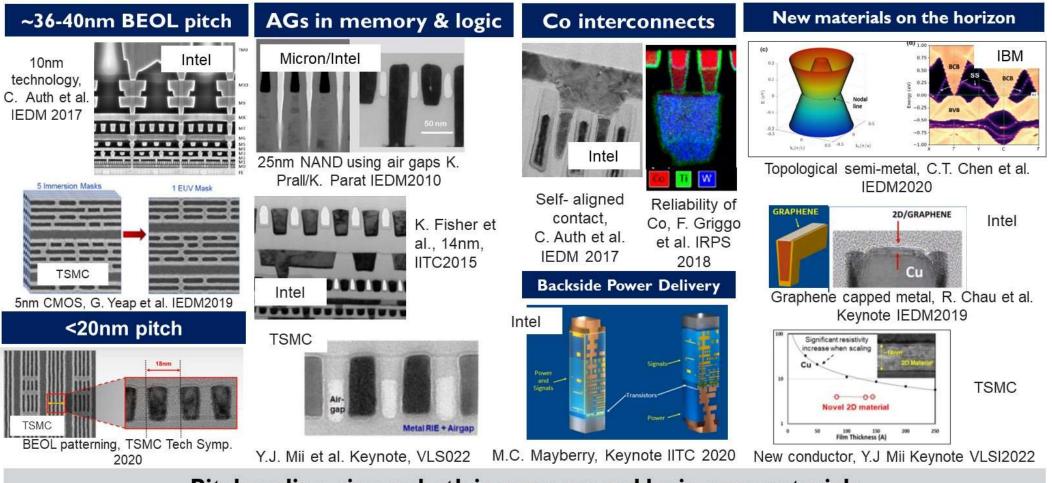
CEA-Leti, unpublished





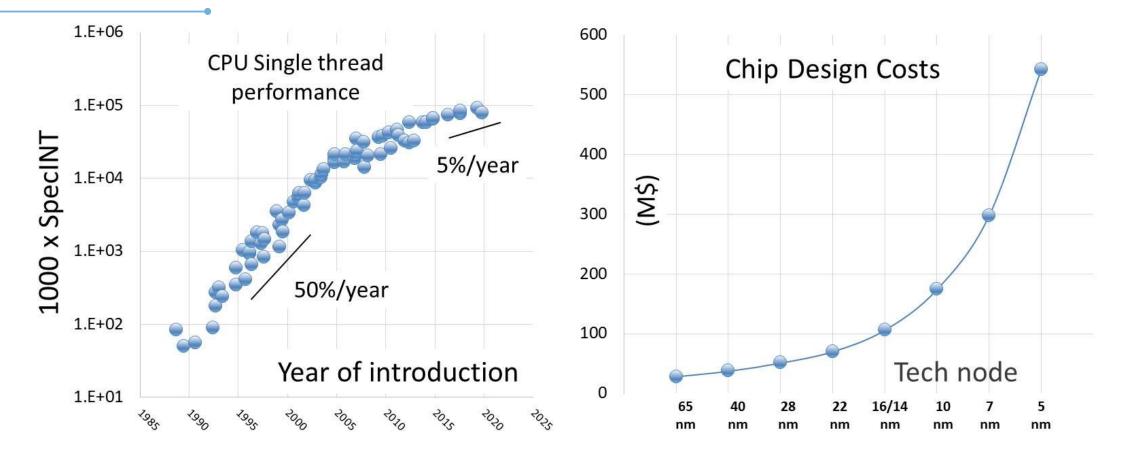
Industry BEOL trends

Courtesy: Zsolt Tokei (imec)



Pitch scaling, airgaps both in memory and logic, new materials

Slowdown in Performance and Increasing Costs





S S

International Cooperation On Semiconductors

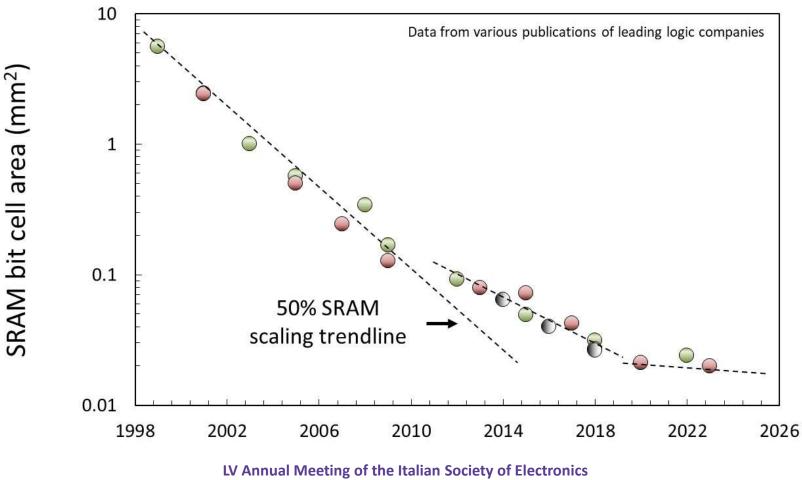
> Based on original data plotted by M. Horowitz, F. Labonte, O. Shachan, K. Olukotun, L. Hammond, C. Batten. Additional data compiled by K. Rupp

Source: AI Chips and why they matter", S. Khan and A. Mann, 2020

26-28 June, 2024 | Genoa, Italy



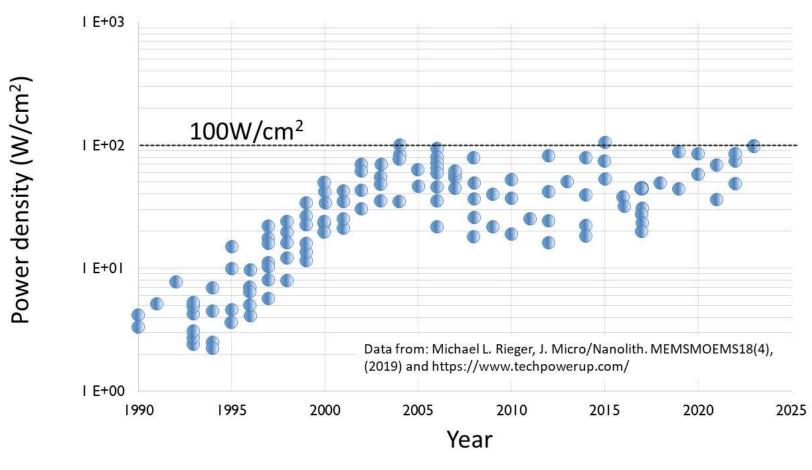
SRAM scaling has slowed down



26-28 June, 2024 | Genoa, Italy



Chip Cooling Limit Power Density of Chips



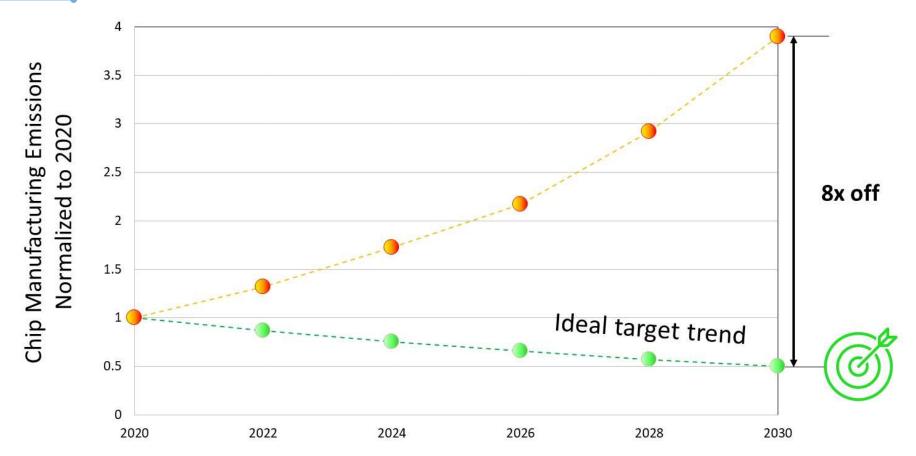
LV Annual Meeting of the Italian Society of Electronics 26-28 June, 2024 | Genoa, Italy



13



Carbon Emissions: "Do Nothing" Scenario

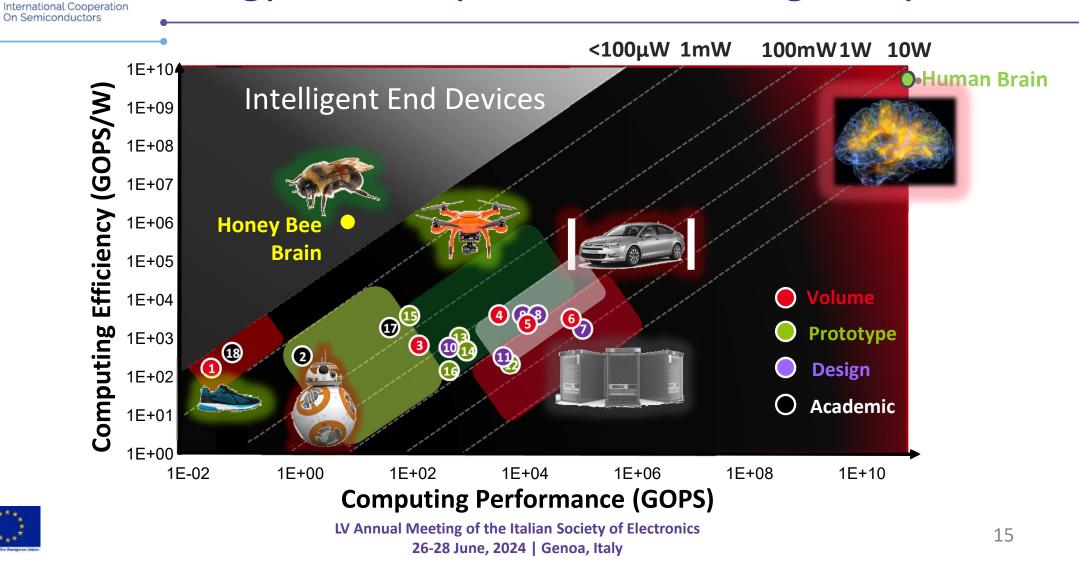


Constant electricity mix (0.49 kCO2eq/kWh), Abatement and GHG global warming potential according to IPCC assumed for the years 2020-2030. Volume technology mix from IBS "Foundry Market Trends and Strategic Implications" Vol 30, N 12, Dec 2021. Logic nodes only. *imec.netzero emissions estimate of imec process nodes representative of foundry nodes.

20-20 Julie, 2024 | Gellua, Italy



Energy Efficiency is far from biological Systems

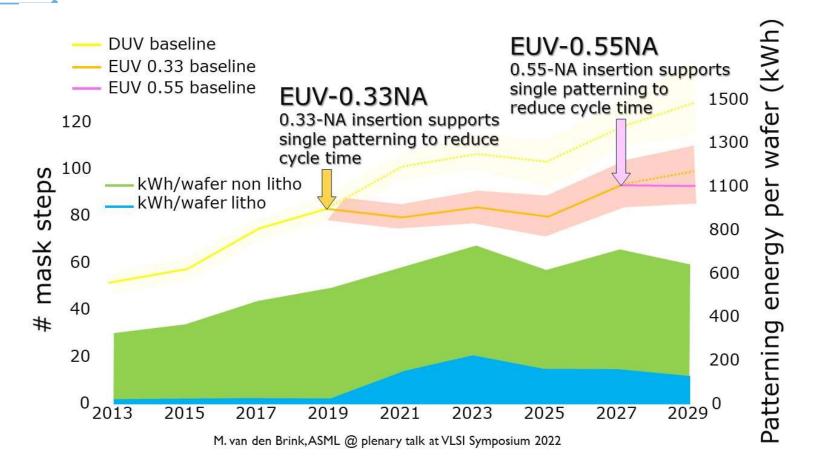




S



EUV lithography key enabler for scaling







Emerging Non-Volatile Memories

	NOR FLASH	MRAM	PCRAM	OxRAM	FeRAM (PZT)	FeRAM (HfO ₂)
Programming power	~200pJ/bit	~20pJ/bit	~300pJ/bit	~100pJ/bit	~10fJ/bit	~10fJ/bit
			14ns @ 2.5V			
Write speed	20 µs	20 ns	10-100 ns	10-100 ns	<100ns	(SONY) 4ns @ 4.8V (LETI)
Endurance	10 ⁵ - 10 ⁶	10 ⁶⁻ 10 ¹⁵	10 ⁸	10 ⁵ – 10 ⁶ on 16 kbit	> 10 ¹⁵	> 10 ¹¹ single device 10 ⁶ - 10 ⁷ on 16 kbit
Retention	> 125°C	85°C - 165 °C	165°C	> 150°C	125°C	125°C
Extra masks	Very high (>10)	Limited (3-5)	Limited (3-5)	Low (2)	Low (2)	Low (2)
Process flow	Complex	Medium	Medium	Simple	Simple	Simple
Multi-Level Cell	Yes	No	Yes	Yes	No	No
Scalability	Bad	Medium	High	High	Medium	Poor (2D) High (3D)

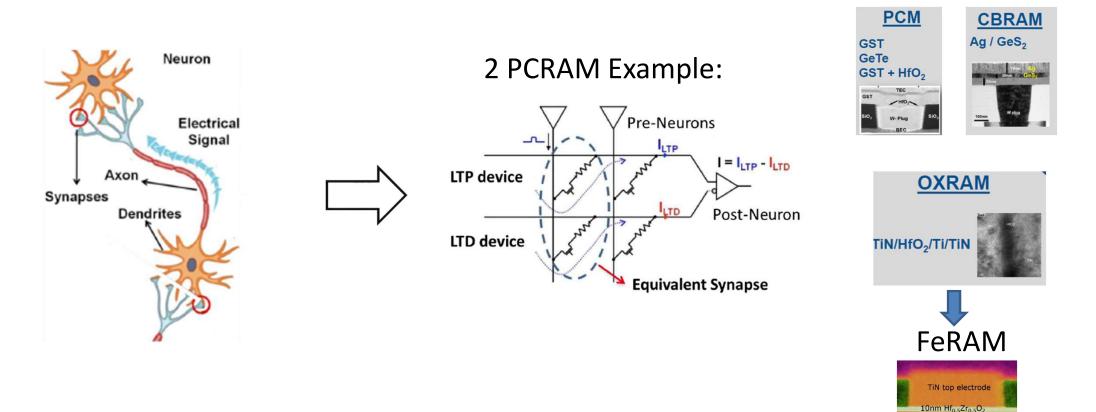
Memory activity focus on embedded NVM for NOR flash replacement





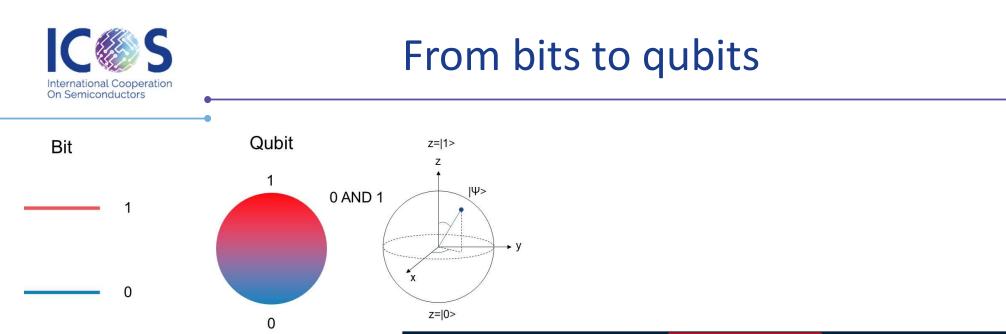
Neuromorphic based RRAM circuit

M. Suri et al, IEDM 2011.





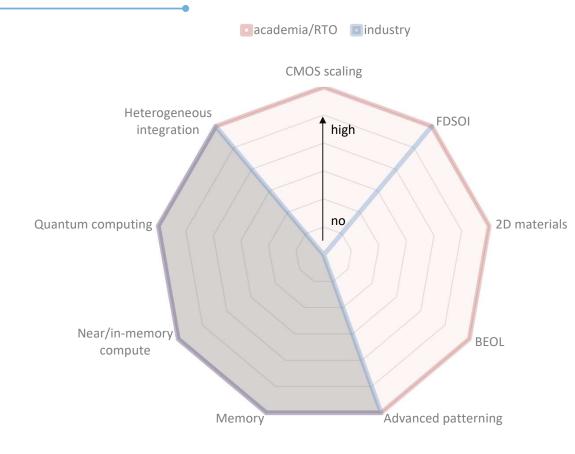
LV Annual Meeting of the Italian Society of Electronics 26-28 June, 2024 | Genoa, Italy TiN bottom electrode



	Superconductor	Si spin qubit	Trapped ion	Photon
Size*	(100µm)²	(100nm)²	(1mm)²	~(100µm)²
1qubit fidelity	99.96%	99.93%	99.98%	
2qubit fidelity	~99.3%	>99%	99.9%	50% (measurement) 98% (gates)
Speed**	12-400 ns	~1 µs	100 µs	1 ms
Variability	3%	0.1%-0.5%	0.01%	0.5%
T° of operation	15mK	1K	10K	4K/10K
Entangled qubits	433 (IBM)	3 (TU) (6 - QuTech)	32 (IonQ)	70 (Pan-China)



EU and non-EU actors - EU



- R&D very strong in all areas of compute
- Unique strong position in EUV lithography
- In general, industrial EU players lacking to take up R&D



Challenges in Advanced Computing and Functionalities International Cooperation on Semiconductors

Thank you for your attention

This project has received funding from the European Union's Horizon Europe research and innovation programme under GA No 101092562

icos-semiconductors.eu



On Semiconductors